Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.01

Hippocrates Meets Democritus: A History of Psychiatry and Clinical Psychology

DONALD K. ROUTH
University of Miami, Coral Gables, FL, USA

1.01.1 INTRODUCTION 2
1.01.2 THE ANCIENT GREEKS AND ROMANS 3
1.01.3 CHINA AND INDIA 4
1.01.4 THE NEAR EAST 4
1.01.5 MEDIEVAL EUROPE 5
1.01.6 FROM THE RENAISSANCE TO THE SEVENTEENTH CENTURY 5
1.01.7 THE EIGHTEENTH CENTURY, INCLUDING THE EMERGENCE OF PSYCHIATRY 6
1.01.8 THE NINETEENTH CENTURY, INCLUDING THE EMERGENCE OF CLINICAL PSYCHOLOGY 9
1.01.9 THE TWENTIETH CENTURY: GUILD-RELATED ISSUES 13
  1.01.9.1 Psychiatry 13
  1.01.9.2 Psychoanalysis 14
  1.01.9.3 Clinical Psychology 15
  1.01.9.4 Social Work 18
  1.01.9.5 Religious Counseling 18
  1.01.9.6 Interdisciplinary Activities and Advocacy 18
1.01.10 RESEARCH ON PSYCHOPATHOLOGY IN THE TWENTIETH CENTURY 20
  1.01.10.1 General Paralysis of the Insane 20
  1.01.10.2 Dementia 21
  1.01.10.3 Dementia Praecox/Schizophrenia 21
  1.01.10.4 Schizoaffective Disorder 22
  1.01.10.5 Depression 22
  1.01.10.6 Neurosis, Including Anxiety 22
  1.01.10.7 Shell Shock/Post-traumatic Stress Disorder 23
  1.01.10.8 Stress 23
  1.01.10.9 Psychosomatic Disease 23
  1.01.10.10 Anorexia Nervosa and Bulimia 23
  1.01.10.11 Borderline Personality Disorder 24
  1.01.10.12 Antisocial Behavior 24
  1.01.10.13 Substance Abuse 24
  1.01.10.14 Sexual Dysfunction 24
  1.01.10.15 Mental Retardation 25
  1.01.10.16 Autism 26
  1.01.10.17 "Minimal Brain Dysfunction" (MBD)/ADHD 26
  1.01.10.18 Learning Disabilities 26

1.01.10.19 Child Abuse and Neglect 26
1.01.11 RESEARCH ON ASSESSMENT IN THE TWENTIETH CENTURY 27
1.01.11.1 Intelligence 27
1.01.11.2 Categorical Diagnosis of Mental Disorders 28
1.01.11.3 Interviews 29
1.01.11.4 Self-report Measures of Personality and Psychopathology 29
1.01.11.5 Rating Scales 30
1.01.11.6 Infant Assessment 30
1.01.11.7 The Rorschach Test 30
1.01.11.8 Human Figure Drawings 30
1.01.11.9 Thematic Assessment Methods 31
1.01.11.10 Sociometric Measures 31
1.01.11.11 Assessment of Social Competence 31
1.01.11.12 Assessment of Marital and Family Relationships 31
1.01.11.13 Neuropsychological Measures 31
1.01.11.14 Biological Assessment Procedures 31
1.01.12 RESEARCH ON TREATMENT IN THE TWENTIETH CENTURY 32
1.01.12.1 The Boston School of Psychotherapy 32
1.01.12.2 Psychoanalysis and its Derivatives 32
1.01.12.3 Play Therapy for Children 33
1.01.12.4 Japanese Therapeutic Approaches 34
1.01.12.5 Behavior Therapies 34
1.01.12.6 Behavior Analysis 35
1.01.12.7 Cognitive Therapy 35
1.01.12.8 The Interpersonal Approach 36
1.01.12.9 Research on Psychotherapy 36
1.01.12.10 The Psychotherapy Integration Movement 37
1.01.12.11 Group Therapy 37
1.01.12.12 Self-help Groups 37
1.01.12.13 Marriage and Family Therapy 37
1.01.12.14 Community Interventions and Prevention 38
1.01.12.15 Special Education 38
1.01.12.16 Rehabilitation 38
1.01.12.17 Drug Therapies 38
1.01.12.18 Psychosurgery 40
1.01.12.19 Electroconvulsive Shock Treatment 40
1.01.12.20 Other Biological Treatments and Preventive Measures 40
1.01.13 EPILOGUE 41
1.01.14 REFERENCES 42

1.01.1 INTRODUCTION

The story goes that Hippocrates (460–377 BC), the famous physician, was once summoned to the Greek city of Abdera by its citizens in order to investigate the seemingly irrational behavior of the philosopher Democritus. Democritus (460–370 BC) is best known today for his view that everything consists of tiny particles called atoms. On this occasion, Democritus was seated under a plane tree surrounded by the carcasses of dogs and cats. He was experiencing an episode of melancholy and was dissecting these animals in order to discover the source of black bile, considered to be responsible for such mental disturbances (Burton, 1621/1971). The term "melancholy" literally means black bile in Greek, and Hippocrates' theory of melancholy, like that of Democritus, was that it was caused by an imbalance of the "humors" of the body (e.g., Jackson, 1978). Hippocrates thus considered Democritus to be rational, not insane.

Hippocrates was not a psychiatrist, nor was Democritus a clinical psychologist. It would be many centuries before either of these specialized fields was to emerge. Nevertheless, the principal theme of this chapter is that even from early historic times, mental disorders were the concern not only of physicians but also of philosophers and others who did not practice medicine. This fact can serve as a sort of historic precedent for the existence of clinical psychology and related nonmedical fields concerned with mental disorders. As Democritus once said in a letter to Hippocrates:

> Wisdom is the sister of medicine: the one rescues the soul from passions, the other alleviates the disease of the body. The mind benefits from health, whereas ill health dampens the desire for virtue, and illness binds the soul. (Temkin, 1991, p. 70)

This history of "clinical psychology" is thus at the same time a history of aspects of medicine, psychiatry, and other related fields. It is concerned with a broad spectrum of pathophysiology, psychopathology, assessment, and various types of interventions for mental disorders.

The author's earlier inquiries into the history of clinical psychology concerned the organizations of clinical psychologists that developed in the USA beginning in 1917 and some of the practitioner pioneers in the field including Lightner Witmer, J. E. W. Wallin, and Leta Hollingworth (Routh, 1994, 1996, 1997). These writings gave relatively little recognition to events before the 1890s, to the contributions of adjacent disciplines such as psychiatry, or to the work of colleagues in other countries. The request from C. Eugene Walker to write the present chapter was interpreted as an invitation to explore not just the organizational history of clinical psychology but its substance as well, not just in recent times but from its origins. In its first 100 years, clinical psychology has become so intertwined with psychiatry (especially in the domain of research) that it made more sense to consider them jointly rather than separately. It soon became evident that writing such a history was nothing short of a lifetime project. However, the present chapter was prepared over only a two-year period and thus required the use of many secondary sources and even some tertiary sources in order for it to be completed at all. The main justification for trying to carry out such a task is that most clinical psychologists are remarkably ignorant about the history of their field.

The chapter begins with a brief history of ideas concerning mental illness in ancient, medieval, and early modern times. It then discusses the development of the specialty of psychiatry in the eighteenth century and after and of clinical psychology about 100 years later. Finally, it describes the further development of the two fields, considered as aspects of a larger whole, during the twentieth century, with a focus on developments in the areas of psychopathology, assessment, and treatment. Many histories of psychiatry and psychology have been written, beginning in the early nineteenth century (e.g., Marx, 1994). So far as the author is aware, however, this admittedly modest effort is the first attempt to write a combined history of psychiatry and clinical psychology.

1.01.2 THE ANCIENT GREEKS AND ROMANS

The Hippocratic writings of the fifth and fourth centuries BC and later (which have diverse authorship) consider mental disturbances to be illnesses like any other. Besides melancholy, the Hippocratic list of mental disorders included phrenitis, mania, hysteria, and paranoia. Plato (427–347 BC) added the category of dementia as well. Phrenitis (which literally means inflammation of the mind) referred to the situation in which a person with fever begins to talk nonsense; this was considered a temporary condition which would clear up as soon as the fever was over. Mania refered to a furious, acute mental disturbance in the absence of fever. Hysteria consisted of bodily complaints made by women, interpreted as being due to the wandering of the uterus. Paranoia described a condition in which the person suffered from abnormal thinking (in the absence of fever). Dementia refered to a deteriorating mental condition often seen in the elderly. Plato, in the *Republic*, describes in detail individual differences in learning ability. He was also aware of the existence of persons with congenital mental incompetence (Plato, 1933). The terminology of mental disorders often changes over time, as do concepts of psychopathology, although some of the phenomena these terms and concepts refer to may be timeless. It would thus be a mistake to identify the above concepts of the Greeks with those we use today, even when they are called by the same names.

Hippocratic medicine emphasized the importance of proper diet, exercise, and moderation in all things. It considered most illnesses to result from an imbalance of the "four humors": blood, black bile, yellow bile, and phlegm. Given this theory, one reasonable type of treatment tried to reduce the excessive amount of a particular humor, by bleeding the patient, using emetics or purgatives, or inducing sweating.

Different schools of thought resembling some present controversies in the philosophy of science were already present among Hippocratic physicians, as the following quotation makes clear:

> Two main sects had formed among his followers. For the so-called Dogmatists, or Logicians, Hippocrates was the founder of a scientific medicine based on natural philosophy, anatomy, and physiology—in short, on hidden causes, as their opponents, the Empiricists contended. In the opinion of the latter, nature could not be fathomed; hence speculation about these hidden causes, which included elements, humors, and pneuma, was useless. They held that medicine had to make do with a knowledge of obvious causes, such as hunger, cold, and sleep. The Empiricists thought of Hippocrates as himself an Empiricist, their great clinical teacher of the past. (Temkin, 1991, p. 6)

Galen (AD 130–200), the most famous successor of Hippocrates, was born in Greece but taught and practiced in Alexandria and Rome. He was an early "scientist–practitioner" in that he was an expert in anatomy as well as medicine. He carried out some of the first physiological experiments. For example, he demonstrated that one could eliminate a pig's squeal by sectioning a particular nerve (Sarton, 1954; Temkin, 1973). Some of the mistakes in Galen's writings on anatomy resulted from the dissection of apes rather than human cadavers (human dissection not being allowed).

1.01.3 CHINA AND INDIA

The study of psychopathology did not originate only in Greece but has other roots throughout the world. A well known example from two centuries BC in China is the *Yellow Emperor's book of internal medicine* (Veith, 1955). This book interprets disease (according to Veith, this included the kind we would consider to be mental illness) as resulting from an improper balance of the powers of "yin" and "yang." These are considered to be caused by moral transgressions of Tao, "the way." Traditional Chinese medicine as described in this book includes an elaborate description of procedures such as acupuncture, still in use in China and elsewhere today. According to these precepts, the person who lives properly has every reason to expect to reach age 100, in vigorous good health. It is proverbial that mental illness is rarely diagnosed in China, perhaps in part because it would involve public loss of "face." Also, children feel obliged to honor their parents and to care for them even if senile dementia has made them burdensome and foolish. The identified prevalence of mental illness is also low among the overseas Chinese. Veith (1963) noted that while yin and yang are implicated in the scholarly account of mental illness in China, the popular view is different. Uneducated people interpret disturbed behavior as being the result of supernatural causes such as spirits entering into the body of the sufferer. Exorcism is thus a common treatment.

An important nonmedical source of the assessment component of present-day clinical psychology and psychiatry is the tradition of formal mental tests established in China. Public examinations served for almost 1000 years as the principal way of recruiting and selecting "Mandarin" civil servants. In China, this system was abolished by the empress dowager only in 1905 (Hanson, 1993), which oddly enough, was just when psychology's involvement in psychometrics was beginning.

In India, at least according to Sigerist's (1945) admittedly dated account, many persons receive medical care, whether for physical or mental illness, only from indigenous practitioners. Some of these doctors follow the principles of the *Ayurveda*, the "science of longevity," as described by Brahmin texts dating back 3500 years. Others follow the Unani, or Greek school, as filtered through Arabian and Persian cultures.

1.01.4 THE NEAR EAST

Judaism, Christianity, and Islam have been competing religious and cultural forces for millenia in the Near East and elsewhere, as they continue to be. All of them have their own methods for treating illness, be it physical or mental. Consider, for example, Jesus' restoration of the sight of a man with a missing eye and raising Lazarus from the dead. As Temkin (1991) points out, secular medicine did not pretend to perform such feats. In Greece, Christianity eventually superseded the Asclepian religious approach to healing. This involved visiting the temples of Asclepias, consulting the priests, and perhaps sleeping there overnight. However, these religions each managed to coexist relatively peacefully with Hippocratic medicine and its descendents right up to the conquest of Alexandria (the center of Hellenistic medicine) in AD 642. After that event, Islamic physicians assimilated the medical teachings of the Greeks and made their own independent contributions to the field.

The most famous of the Islamic physicians was Avicenna (AD 980–1037), whose Arabic name was Ali al-Husayn ibn Sina. Magner (1992) cites some of the work of Avicenna as an early example of psychotherapy:

> [One] challenging case involved a young man who suffered from melancholy and the delusion that he was a cow. The man mooed loudly, refused to eat, and begged to be killed and made into stew. The patient cheered up immediately when Avicenna sent word that the butcher would come to slaughter him. Finally, Avicenna came to the sickroom with a butcher's knife and asked for the cow. Mooing happily, the young man was bound hand and foot, but after a thorough examination, Avicenna declared that the cow was too thin to be butchered. The patient ate so eagerly that he soon recovered his strength and was cured of his delusion. (p. 143)

According to Dols (1987) and Magner (1992), the Muslims also built hospitals including provisions for the care of the mentally ill. The largest of these was established in Cairo in 1286. Formerly a palace, this hospital had a capacity of 8000 patients. Other such hospitals were located in Baghdad and Damascus and existed as early as the ninth century AD. The treatments for mental patients included baths, massage, bloodletting, leeches, cautery, medication with purgatives, emetics, sedatives, and opium. There was also concern for the patient's diet, opportunities to participate in pleasant conversation, and diversions such as music, dance, and drama. If necessary, Dervishes performed exorcisms. Some violent patients were in chains, and whips were used on occasion.

The condition we know as mental retardation was also familiar. As Dols (1987) says:

> The village idiot, the court jester, and the wise fool were tolerated forms of possible derangement. The archetypal wise fool in Islamic literature is Buhlul, the entertaining and harmless critic of social conditions and mores. He typically lurks undisturbed in the cemetary, is molested in the streets by children, or is fettered and put in chains. Having abandoned everyday cares, he trusts in God's aid and men's charity. His intense sensitivity plays on the ambiguity between the unholy and the holy fool. (p. 13)

1.01.5 MEDIEVAL EUROPE

Medieval European times were not noted for scientific advances of any kind, certainly not with respect to the understanding and treatment of psychopathology. The medieval era does have its positive side in terms of the Christian emphasis on the virtue of caring for the unfortunate, including those with mental disorders. Thus, in the sixth century AD, Benedict of Nursia proclaimed the ideal of service toward those living in monasteries who were ill (Jackson, 1972). However, such care did not extend very far to those in the community, as Rosen (1964) notes:

> During the medieval period, public authorities took only limited responsibility for the mentally deranged. Mentally or emotionally disturbed members of a community were left at liberty as long as they caused no public disturbance. Custody of the mentally ill generally rested with their relatives and friends; only those considered too dangerous to keep at home, or who had no one to care for them, or who were socially disturbing, were dealt with by communal authorities. (p. 377)

The institutions for the mentally ill that developed in that era provided at most food, lodging, and basic physical care. They included general hospitals such as the Hotel-Dieu in Paris, the hospital of St. Mary of Bethlem in London (built in 1247, later notorious as "Bedlam"), and the madhouses that were part of the hospitals at Elbing (built in 1326) and Hamburg (built in 1375). Priests offered exorcism to persons considered to be possessed by devils, and both the secular and the ecclesiastical authorities sometimes subsidized pilgrimages by the mentally disordered to religious shrines such as Lourdes.

In the twelfth century a church was built at Gheel, a village north of Brussels, Belgium, to honor St. Nymphna. The village developed the pattern of boarding mentally disturbed patients in the homes of the residents, and this Shrine of Gheel became an enduring legend in community care. The famous psychiatrist Esquirol visited it in 1822.

One high point in medieval and renaissance thought concerning mental disturbance is the philosophical writings of the Portuguese King Duarte (1391–1438) known as *Leal Conselheiro* ("loyal counselor"). Duarte experienced mental depression and undertook to write about it as a form of self-therapy. What he had to say is well worth quoting:

> And it seems to me . . . that sorrow generally has six causes: first and chiefly, fear of death, dishonor, pain, or spiritual and bodily suffering. Second, anger . . . Third, strong desire unfulfilled or delayed. Fourth, unhappiness which we acquire because of dishonor, death, losses, imprisonment, illnesses, retention, and longing. Fifth, disordered compression, which truly is called an illness of the melancholy humor. Sixth, talk, conversation of sad persons . . . By each of these means, more or less, we become sad, according to the feelings and passions which are dominant in each of us. (Roberts & Sacks, 1954, p. 22)

As it turned out, Duarte's melancholy abated after his mother became ill, and he turned his thoughts away from himself to care for her.

An acknowledged horror story of the middle ages and renaissance (from the standpoint of the understanding and treatment of psychopathology) is the publication of *Malleus Malificarum* ("hammer of the witches") by two German Dominican monks, Heinrich Kraemer and Johann Sprenger (1486/1971). This was a manual for the identification, torture, and trial of witches. It was endorsed by the Pope and went into 19 editions over the next 300 years (Leigh, 1957). According to this book, witches, though they might appear to be merely lunatics, were actually in league with the devil and caused all manner of mischief, including hailstorms, crop failure, and sexual dysfunctions such as impotence in other persons. Most of the "witches" who experienced the regimen prescribed by these two doctors of the church were aging women.

1.01.6 FROM THE RENAISSANCE TO THE SEVENTEENTH CENTURY

The answer to the "witches' hammer" was Johann Weyer's book, *De Praestigiis Daemonum* ("the slight of hand of demons"), published in 1583 (Mora, 1991). Weyer, an experienced physician, was employed by a wealthy duke and seemed to be unafraid of the enemies he might make by his writings. He had considerable personal experience in examining various individuals who were accused of witchcraft and usually found that they were poor, frail, mentally unstable women who could not possibly have done the deeds of which they were accused. He thus disputed Kraemer and Sprenger's *Malleus Malificarum* very thoroughly. Weyer's book was put on the Papal index of forbidden writings.

Another important medical figure of the renaissance was Paracelsus (1493–1541), the name adopted by a flamboyant Swiss man named Philippe Aureole Theophrastus Bombastus von Hohenheim. Paracelsus was of modest background rather than being a member of the nobility (as his name might imply) and was not as well educated as many of his medical peers, whom he often scandalized by communicating in the vernacular rather than in Latin, the common language of scholars of that time. His father had worked in the mines, and Paracelsus knew a lot about minerals. An alchemist as well as a physician, he introduced chemical treatments. Temkin (1946) estimates that Paracelsus was as important to the history of medicine as Luther was to church history. The followers of Galen had always recommended herbal remedies, and now Paracelsus came along with chemical ones—the founder of a new school of thought known as "iatrochemistry." Paracelsus was also responsible for a very modern-sounding insight into the possible role of minerals in the drinking water in the etiology of goiter and Cretinism (a form of what we could now call mental retardation) that was so prevalent in Switzerland.

Robert Burton (1577–1640), the author of *The anatomy of melancholy* (Burton, 1621/1971), lived in this era. Burton was not a physician but a clergyman and an Oxford don. He was in a state of melancholy himself when he wrote this huge universal compendium of ancient knowledge partly as a way of distracting himself. Burton referred to himself as "Democritus Junior," thereby alluding to the pre-Socratic philosopher mentioned in the title of the present chapter. Burton evidently did not know of King Duarte's writings, but there are interesting parallels in the lives and thinking of Democritus, Duarte, and Burton. Clinical psychologists might do well to consider all of them as mentors.

The physician Sydenham (1624–1689) wrote a treatise on psychological disorders that has been described as the most important one of the seventeenth century. Sydenham reportedly considered psychological disorders to be as common as physical ones. He was also something of a legend as a therapist, as the following incident reveals:

> ... in dealing with a patient who stubbornly refused to get well, Sydenham suggested a consultation with the wonder-working Dr. Robinson at Inverness. The patient made a long trip on horseback only to find that there was no such doctor, but, as Sydenham expected, anticipation, exercise, and anger effected a cure. (Magner, 1992, p. 222)

Foucault's (1961) account of the classical age asserts that the seventeeth century was an age of increased "confinement" involving not only the mad but also criminals, the poor, and the indigent ill all over Europe. On a more positive note, Foucault credits Thomas Willis with being the first to make the observation that melancholia and mania tended to alternate in the same individuals. Willis wrote about this in 1681.

Finally, some mention should be made of a late-seventeenth century philosopher, John Locke (1632–1704), and of his influential book, *An essay concerning human understanding* (Locke, 1689/1824). Locke is well known to psychologists as one of the British Empiricists and as an associationist (Hoeldtke, 1967). In this book, Locke argued that the mind is a blank slate at birth and that its contents are acquired through sensory experiences and by reflection upon these. One assertion of his that was influential regarding mental disorders was that whereas idiots were incapable of associating ideas, madmen associate ideas incorrectly.

1.01.7 THE EIGHTEENTH CENTURY, INCLUDING THE EMERGENCE OF PSYCHIATRY

At the beginning of the eighteenth century, Edinburgh was a major intellectual center. It was associated with the Scottish Enlightenment and with common sense philosophy and psychology. The University of Edinburgh was also a center of medical education, with its influence reaching across the Atlantic (Carlson & Simpson, 1969) and into France. Vienna was also respected as a center of medical thought as was Montpellier in France, but the German states were in a rather backward condition, perhaps due to the after effects of the 30 Years War. At this time the dominant figure in Scottish medicine was perhaps William Cullen (1710–1790), who founded a school of "neuropathology." The modern sense of this word is somewhat misleading. It does not adequately convey the fact that Cullen was trying to represent the brain and the nervous system as the "master" influence on the health of the body (Riese, 1949) as an alternative to the "humoralism" still lingering from the ancient days of Hippocrates. Cullen's major work was a book, *First principles of the practice of physic*, published in 1776–1784. Naturally, the "neuropathology" approach had a lot to say about mental disorders.

This was also the time when the tradition of "mad doctors" emerged more fully in the British isles (Hunter & McAlpine, 1963). The first of these had been John Archer, who already in the seventeenth century became the proprietor of a private madhouse. The most eminent "mad doctor" of the eighteenth century was a man with the interesting name of William Battie (1704–1776), who published a *Treatise on madness* in 1758. According to Porter (1987), Battie's book contains the key ideas of moral therapy as subsequently practiced by William Tuke (1732–1822). The number of private madhouses increased greatly during the late eighteenth and early nineteenth centuries. Before the advent of moral therapy, many of their proprietors used fear and restraint, including chains, to deal with the inmates (Scull, 1993). According to Bynum (1974), the first act aimed at licensing private madhouses was passed in Britain in 1774. The law did not require the proprietors to be medically qualified but did require a medical certificate for a patient to be admitted. Notably, 30 years earlier than this, the prevailing law had implied madness to be a condition that anyone could recognize and that any magistrate could formalize; the legal recognition of special knowledge about madness is part of what confirms the emergence of what we now call psychiatry during the eighteenth century in Britain. As Ackerknecht (1986) noted, the title "doctor" when used for the proprietor of a madhouse could refer to a university-trained physician, a surgeon, an apothecary, or simply a quack.

Although the history of psychiatry has somewhat indistinct beginnings in various European countries, the specialty was well established on both sides of the Atlantic before 1800. As Porter (1987) says concerning England:

> Around 1700, the top practitioners for the disturbed were all-round physicians: within a century they were specialists. Mad-doctoring came of age on the day in 1788—5 December in fact—when the failure of the King's physicians-in-ordinary to master George III's delirium was acknowledged by the summons to the specialist, if very much despised, mad-doctor Francis Willis. (p. 175)

Some would probably not choose to honor Willis as the founder of this new professional specialty, since he represented the old-fashioned coercive approach rather than the newer one of moral treatment. In any case, the specialists known as "mad doctors" in Britain were called "alienists" in France, and ultimately termed "psychiatrists" in the German states and then all over the world. By the time of the American and French revolutions, the individuals who stood out as pioneer figures in this field included Benjamin Rush (1745–1873) in the United States, Phillipe Pinel (1745–1826) in France, Vincenzo Chiarugi (1759–1820) in Florence, and Francis Willis (1718–1807) in England.

It appears that the leadership in medicine shifted from Edinburgh to Paris at about the time of the French revolution in 1789 (Ackerknecht, 1967). The French physician Philippe Pinel, who is often considered to be the founder of psychiatry, was much influenced by the doctrines of William Cullen of Edinburgh. Pinel actually translated one of Cullen's books into French. Cullen's writings also influenced Benjamin Rush, who incidentally was a signer of the American Declaration of Independence.

How were mental patients treated in the eighteenth century? In some respects, the ancient ways of Hippocrates prevailed. Benjamin Rush was especially well known for his propensity to bleed his patients, and may have inadvertently hastened the death of many of them. Many eighteenth century "mad-doctors" in Britain emphasized coercive methods, including chains and fetters. The important therapeutic innovation of this time is what was known as "moral treatment." This phrase emphasized the attempt to use psychological and social influences and to treat the mentally disturbed with kindness rather than physical force, chains, or medications. The asylum began to be considered as a therapeutic intervention rather than simply a way to house and contain mental patients. In retrospect, moral treatment is usually portrayed as a tremendous advance, perhaps the very cornerstone of the new specialty of psychiatry. Indeed, it was obvious from the first that many patients did not require restraint after all, and often responded well to kindness and understanding. Who deserves the most credit for the development of moral treatment is somewhat unclear. Philippe Pinel became the focus of a legend, not only as a founder of psychiatry but also as the one who was the first to remove the chains from his patients at the Bicetre hospital in Paris. Actually, the evidence is that he was not the first to do this but served an apprenticeship to Citizen Jean B. Pussin in this respect (Weiner, 1979). Another contender for the credit of founding moral treatment is William Tuke, who was not a physician but a Quaker tea merchant. Tuke founded the York Retreat in 1796 as a way of providing a gentler approach than that offered by other available institutions. Tuke's work was influential and was copied, for example, by the founding of the Hartford Retreat in the United States. Finally, one must mention Vincenzo Chiarugi as one who opposed restraints and harsh measures. Chiarugi was born in Florence. His patron was Pietro Leopoldo, the Duke of Tuscany. Chiarugi directed the Hospital of Bonifacio in Florence and published three volumes on his work in 1793–1794, under the title, *Della Pazzia in Genere e in Specia* ("on madness in general and in specific").

In any case, not all scholars view the advent of "moral treatment" as an unmitigated improvement. Foucault (1961) became famous for noting that "moral treatment" could be considered simply a different kind of oppression. In making the mad responsible for their own self-control, Tuke and Pinel may have merely substituted guilt for external coercion, and unduly imposed demands for a certain type of bourgeois respectability.

Even with all the emphasis on the absence of chains in the era of moral treatment, Pinel and his colleagues in other countries did use restraint when necessary, including strait jackets and isolation rooms (Woods & Carlson, 1961). Pinel avoided punishment and was reluctant to use the traditional Hippocratic treatments of bleeding, emetics, and purgatives. He observed the progress of the patient and tried to provide a humane environment, hoping that the healing forces of nature would prevail over the illness. In Britain, the physician who became most famous for a policy of nonrestraint was John Conolly, who had charge of the Hanwell Asylum in England in the nineteenth century (Scull, 1993).

A countervailing force at this time was the emerging movement of romanticism, associated particularly with the name of the philosopher and social contract theorist, Jean Jacques Rousseau (1712–1778). In his influential novel, *Emile, or education*, and in his other writings, Rousseau (1911) stressed the importance of what emerged naturally in human development as opposed to what is contrived by the educational system. He emphasized the heart over head, in contrast to what he considered the overintellectual emphasis of the Enlightenment.

Mesmerism emerged during the eighteenth century, more as a treatment for medical problems than for mental disorders. This development came from the work of Franz Anton Mesmer (1734–1815), a physician who was originally trained in Vienna but who later flourished in Paris. Mesmer's theory about his therapy explained it in terms of magnetic forces being transmitted from the body of the doctor to that of the patient, curing all kinds of ailments. The theory was based on the study of electricity including the role of electrical stimulation on the movement of frogs' legs (the original "animal magnetism"). This was debunked as an explanation by more than one distinguished commission in Paris. The commissioners included such experts on electricity as Benjamin Franklin, then serving as ambassador from the United States. These scientific authorities tended to believe that the therapy was based merely on "imagination." At the time, the effect on Mesmerism was quite negative in terms of the opinions of educated persons, although its practice continued not only in France but also in the German states and elsewhere (Gauld, 1992). Ellenberger (1970), in his book *The Discovery of the unconscious*, portrays Mesmer's work as being an important link between the medieval use of priestly exorcism and the subsequent development of psychoanalysis by Freud.

It happened that the first cases of what was subsequently identified as neurosyphilis began to appear in the 1780s (its prevalence rose later, during the nineteenth century). This disease, which is known to have a latent phase lasting 10–15 years from the time of initial infection, was manifested in the form of the general paralysis of the insane (GPI) or paresis. Neurosyphilis included dementia paralytica, progressive paralyses, terminal convulsions, and death. Its early symptoms could include mania and locomotor ataxia (tabes dorsalis). Paresis most commonly affected middle-aged men.

Toward the end of the eighteenth century there were important new developments in the treatment of congenital mental handicaps. One innovator was Giacobbo (Jacob) Rodriguez Pereire (1715–1780). Pereire (or Pereira as his name is sometimes spelled), living in France at the time, developed methods for teaching the deaf to speak. He never revealed the particulars of his methods but was obviously successful in doing what he claimed, as revealed by the fact that his deaf pupils did learn to speak. A famous disciple of Pereire was the physician Jean Marc Gaspard Itard (1775–1838), who attempted to use some of the same approaches to train the Wild Boy of Aveyron. Itard considered his attempts to socialize this feral child and to teach him language to be unsuccessful, but in retrospect they have proven to be a model for educators working with children with mental retardation or autism. Itard's work was thus influential in the development of fields as diverse as otolaryngology, child psychiatry, special education, and clinical psychology.

As noted by Shorter (1997), in the late eighteenth century patients with minor nervous complaints tended to seek out treatment at spas. Thus, there grew up a cadre of "spa doctors" to care for them. In addition to bathing in or drinking natural mineral waters, such patients might also be subjected to actual "hydrotherapy," special diets, or to mild shocks from some special electric apparatus. One can see in the practice of such physicians much of what was more likely to be identified as clinical "neurology" for many of the years of the following century.

1.01.8 THE NINETEENTH CENTURY, INCLUDING THE EMERGENCE OF CLINICAL PSYCHOLOGY

Paris continued to dominate the medical scene throughout the first half of the nineteenth century, with its emphasis upon the hospital as the site of clinical investigation (Ackerknecht, 1967). It was there that treatment was carried out and there that the correlation of clinical symptoms with the findings of pathological anatomy were observed. Pinel was if anything as famous in France for his work in internal medicine as in psychiatry. Pinel's most famous pupil, Jean-Etienne-Dominique Esquirol (1772–1840), unlike his teacher, was able to devote his entire professional career to psychiatry.

Berrios (1996) identifies the second decade of the nineteenth century as a time when the previous global descriptions of types of mental disorders began to be elaborated into detailed symptom lists and definitions, or what is now called descriptive psychopathology. This process was essentially complete by the time of the World War I. Emil Kraepelin (1856–1926), the leader in this domain, published the final volume of the eigth edition of his textbook of psychiatry in 1915.

The romantic movement within German psychiatry continued to flourish in the 1820s and in places perhaps until 1848. An example of such a psychiatrist was Johann Heinroth (1773–1843), who was appointed to the first known chair of psychische Heilkunde ("psychotherapy") at the University of Leipzig in 1819. The type of psychotherapy favored by Heinroth made considerable use of Christian precepts such as sin, repentance, and religious conversion. The romantics, influenced by German nature philosophy, were "Psychikers" who battled against the "Somatikers."

Formal organizations of psychiatrists began at about the same time in France, Britain, and the United States during the 1830s and 1840s, when public mental asylums began to be constructed in these countries. In France, a law was passed in 1838 creating a national system of asylums (Goldstein, 1987). It was supported by liberals (who tended to be anticlerical) and to some extent replaced the previous system run by the Catholic orders. Even after the secularization of mental hospitals, many of the psychiatric nurses in the nineteenth century were supplied by Catholic religious orders such as the Sisters of St. Vincent de Paul (Santos & Stainbrook, 1949). The French psychiatric journal, *Annales Medico-psychologiques*, began publication in 1843, and in 1848 the Société medico-psychologique was founded (Dowbiggin, 1989). In England, which passed a similar law in 1845, there had previously been founded in 1841 an Association of Medical Officers of Asylums and Hospitals for the Insane (Scull, MacKenzie, & Hervey, 1996). In 1853 the Association's journal began, with Sir John C. Bucknill (1817–1897) as editor. Called the *Asylum Journal*, it eventually became the *British Journal of Psychiatry*. In the United States, Dorothea Dix (1802–1887) presented her famous Memorial to the Massachusetts legislature, urging the state to take responsibility for the insane by building and staffing proper asylums for them (Gallaher, 1995). In 1844 the forerunner of the American Psychiatric Association was founded and began to publish the *American Journal of Insanity*, eventually called the *American Journal of Psychiatry* (Grob, 1962).

In Paris, Edouard Seguin (1812–1880), a physician who was a pupil of Itard, in 1837 opened an innovative training school for imbeciles. Seguin later emigrated to the United States and became the most influential person in the founding, in 1876, of the Association of Medical Officers of American Institutions for Idiotic and Feebleminded Persons.

In Switzerland, in 1841, the physician Johann J. Guggenbuhl (1816–1863) founded an institution for Cretins at the Abendberg (Kanner, 1959). Guggenbuhl received praise from all over the world for this innovation, but the Abendberg was closed down after only a few years because of mismanagement. Theodor Kocher (1841–1917), a Swiss surgeon, discovered the cause of Cretinism (Schlick, 1994), despite his lack of any intitial interest in the problem of mental retardation. Kocher, in ablating the thyroid gland of certain of his patients with goiters (especially the children among them), noticed that he had inadvertently created a sort of artificial cretinism. He later made a visit to the Valais and the mountains near Bern in Switzerland to examine some natural cases of cretinism there to compare them with his surgical cases. Of 32 alleged cretins referred to him for examination, he concluded that only six fit the refined criteria his surgical experience led him to develop to diagnose this disease. The other patients, though mentally handicapped, failed to meet his criteria for such a thyroid-based disorder. Later, Kocher developed techniques for transplanting thyroid glands, the first such successful organ transplant. Thus, he discovered a curative therapy for cretinism as well as its cause. He won the Nobel prize for these discoveries after the turn of the century, in 1909.

In England, physician J. Langdon Down (1828–1896) wrote a notable paper on the mental retardation syndrome now associated with his name. He misinterpreted the condition as representing some kind of an evolutionary throw-back to "Mongol" ancestors.

In the late nineteenth century, let us say after 1848, the dominant role in medicine and psychiatry shifted to the German states (Ackerknect, 1967). The unified country of Germany emerged in 1871 after the Franco-Prussian war. German medicine was dominated by laboratory approaches rather than by work in the clinic or the hospital. As the German physician Bernard Naunyer (1839–1925) was reported to have said: "Medicine will be a science, or it will not exist at all" (Magnus-Levy, 1944). The leaders in German psychiatry would surely include Wilhelm Griesinger (1817–1868) and later Emil Kraepelin (mentioned above in relation to descriptive psychopathology). Griesinger published an influential book, *Mental pathology and therapeutics*, in 1845. He was the director of psychiatry at Berlin's Charite Hospital and held the chair in psychiatry at the University of Berlin, beginning in 1864. Unlike other psychiatrists of the time, Griesinger did not work in a mental hospital out in some isolated rural area. He tried to make psychiatry an integral part of a major urban university. Griesinger has become infamous among the psychodynamic psychiatrists and psychologists of our day for his rather dogmatic statement that mental diseases are brain diseases. What is not often realized by such critics is that Griesinger was quite sophisticated in the psychology of his time. He made extensive use of the concepts of the psychologist Johann F. Herbart (1776–1841) concerning the unconscious, the threshold of consciousness, and the "apperceptive mass" used by people in interpreting their experiences. Griesinger founded a new journal with the title [English translation] *Archives for Psychiatry and Nervous Diseases* in 1867 to focus on the type of neurologically oriented field of psychiatry he envisioned. However, he died the next year, with his vision still unrealized.

One German psychiatrist who took the road laid out by Griesinger was Theodore Meynert (1833–1892), a professor in Vienna, who was incidentally one of the teachers of Sigmund Freud (1856–1939). This was in the days when Freud still intended to become an academic neuroanatomist. Meynert was the investigator who distinguished the various layers of the cerebral cortex in terms of cell types. Although he had a poor reputation as a clinican, Meynert trained several students who themselves made contributions to neuroanatomy. In 1865 the French physician Pierre P. Broca (1824–1880) had developed his influential hypothesis that expressive aphasia results from damage to the third frontal convolution of the left hemisphere of the human brain. At that time, Broca had only a single case with an autopsy to report. One of Meynert's pupils, Carl Wernicke (1848–1905), later described a type of receptive aphasia that was quite distinct from Broca's aphasia, anatomically as well as functionally. The patients Wernicke identified had lesions in the posterior perisylvian region of the left hemisphere, were unable to understand spoken words, and spoke in a fluent but incomprehensible jargon. Julius Eduard Hitzig (1838–1907), another Meynert pupil, carried out research on dogs indicating that the brain responds to local electrical stimulation.

Franz Joseph Gall (1759–1828), a German who received his M.D. from Vienna, also began his career with well-respected research on the topic of the localization of function in neuroanatomy (Temkin, 1947). His viewpoint was opposed by Pierre Flourens (1794–1867), who did not find evidence for such localization in his research on the cerebral cortex of pigeons. Gall and a colleague named Spurzheim went on to try to establish a correlation between human "neuroanatomy" (as indexed by bumps on the skull) and psychological characteristics. Spurzheim called this field phrenology, which translates simply as the study of the mind. Phrenology went on to considerable popular and commercial success in England and the United States and at first was supported by a significant element within the medical and psychiatric community. For example, John Conally (1794–1866), the prominent British alienist and champion of the doctrine of nonrestraint, was also a supporter of phrenology. Some supporters managed to combine phrenology and hypnotism (Gauld, 1992). Thus, they developed a professional approach that included a sort of prototype of neuropsychological assessment and psychological treatment. Ultimately, however, phrenology was consigned to the status of a pseudoscience and its practice condemned as quackery.

Late nineteenth century Germany was also one place where the new discipline of psychology emerged. It was at first modeled upon experimental physiology, for example, the work of Ernst Weber (1795–1878) and Gustav Fechner

(1801–1887) leading up to Fechner's 1860 book, *Elements of psychophysics*. The founders of psychology as an independent academic discipline in the 1870s are generally considered to have included William James (1842–1910) at Harvard University in the United States and Wilhelm Wundt (1832–1920) at the University of Leipzig, Germany. Of these founders of psychology, the one who made the greatest contribution to the study of psychopathology was William James, as shown, for example, by his 1896 Lowell lectures on exceptional mental states (Taylor, 1982).

The American Psychological Association was founded in 1892 by Granville Stanley Hall (1844–1924). Hall tried his best to facilitate a relationship between psychology and psychiatry. A Ph.D., he served at one point as the superintendent of the Bayview mental hospital in Maryland. Later he taught psychology to psychiatry trainees at Worcester State Hospital in Massachusetts. After the turn of the century, in 1909, he was to invite Freud and Jung to America, the only time Freud ever visited the US.

Emil Kraepelin, ultimately a professor at the University of Munich, may have been the leading German psychiatrist of his day. He received his MD at Wurtzburg and served as a resident under Bernhard von Gudden (1824–1886) in Munich (Danek, Gudden, & Distel, 1989). He also studied psychology under Wundt at Leipzig. In his 1893 textbook, Kraepelin first introduced the concept of "dementia praecox." This term combined a heterogeneous group of symptoms including social withdrawal, hebephrenia (silly, childish language), paranoia (delusions), and catatonia (paralysis with waxy flexibility upon passive movement). The reason for combining these symptoms was Kraepelin's hypothesis that many persons with such symptoms tend to become chronic mental patients. Kraepelin considered this syndrome to be a kind of dementia in a young person that predicted the individual's ultimate mental deterioration. He wrote extensively about the entire spectrum of psychopathology and divided the syndromes into 13 major groups.

In 1826 the French physician Antoine L. Bayle (1799–1858) had found an association of GPI with chronic inflammation of the meninges of the brain. In 1851, two French psychiatrists simultaneously came up with the concept we now term bipolar disorder, confirming the statements of Thomas Willis in the seventeenth century. Jean-Pierre Falret (1794–1870) called this condition "folie circulaire," whereas Jules Baillarger (1809–1890) called it "folie a double forme" (Bourgeois & Geraud, 1997). Incidentally, Falret was also the first to coin the term "folie a deux" for a sort of psychosis of association in which an otherwise normal person acquires symptoms from being around someone with a mental disorder. Thus, a wife might adopt her paranoid husband's view that the house is surrounded by agents bent on harming them.

In England, the most influential scientific figure of the nineteenth century was Charles Darwin (1809–1882), who published his book *Origin of the species* in 1859. His cousin, Francis Galton (1822–1911), was the author of a book on *Hereditary genius*, published in 1869 (Galton, 1978), in which he developed a new field he called "eugenics" attempting to use the principles of genetics to improve the human species. It was Galton's American student, James McKeen Cattell (1860–1944), who first coined the term "mental test" (Cattell, 1890).

French psychiatry had a lengthy involvement with a concept later considered discredited, namely the theory of degeneration of Benedict A. Morel (1809–1884). This theory was a variation of Lamarck's theory of the inheritance of acquired characteristics (Walter, 1956). It led to the expectation that the offspring of disturbed or profligate persons would become worse in each generation. One follower of Morel was Cesare Lombroso (1835–1909) at the University of Turin in Italy. In 1876 Lombroso published a book, *L'uomo delinquente*, concerning what he described as the born criminal, with numerous "stigmata" of degeneration. Richard von Krafft-Ebing (1840–1902), a professor of psychiatry at the University of Vienna, also endorsed a version of degeneration theory as applied to sexual perversions. Krafft-Ebing published a book, *Psychopathia sexualis*, in 1886.

In the United States, George M. Beard (1839–1883), a New York neurologist (i.e., a specialist in nervous and mental diseases), published in 1880 a book about his concept of "neurasthenia," meaning simply weak nerves, with a vast array of symptoms. Beard was interested in electrotherapy and had been involved in the scientific study of hypnosis in the 1870s. Neurasthenia became an extremely popular medical diagnosis for people with chronic fatigue and vague nervous complaints (Rosenberg, 1962). The American Neurological Association had just established in 1871 its *Journal of Nervous and Mental Diseases*. The "neurologists" of those days were in competition for outpatients with psychiatrists, who were still for the most part in institutional practice. According to Shorter (1997, p. 130), neurasthenia was the "standard diagnosis for all functional nervous diseases" up to the time of World War I. These were the kinds of patients for whom the Philadelphia physician Silas Weir

Mitchell (1829–1914), of the Jefferson Medical College, prescribed his famous "rest cure," or as he tended to say, "Dr. Diet and Dr. Quiet." The American Neurological Association was founded in 1874. In 1894, Mitchell gave a famous address before the 50th Annual Meeting of the American Medico-Psychological Association (forerunner of the present American Psychiatric Association) in Philadelphia. As a neurologist who worked mainly with outpatients, Mitchell severely criticized his "alienist" colleagues for their lack of connection to the rest of the field of medicine and their lack of research:

> Where, we ask, are your annual reports of scientific study, of the psychology and pathology of your patients. ... Seriously we ask you experts, what have you taught us of the 91 000 insane whom you see and treat? (Mitchell, 1894, p. 101)

This talk, though harsh, facilitated the later integration of these two parts of present-day psychiatry.

In 1880 in New York, a rather short-lived organization was founded called the National Association for the Protection of the Insane and the Prevention of Insanity (NAPIPI). The membership consisted largely of neurologists and social workers who were critical of the asylum psychiatry of the day. In a sense this organization could be regarded as a forerunner of the antipsychiatry movement of the twentieth century (Dain, 1994).

Charles Locock, a London internist, was responsible for initiating the common use of bromides as a sedative in psychiatric practice. According to Shorter (1997, p. 200), Locock happened to mention at a medical meeting in 1857 that he had given 10 grains of potassium bromide per day for two weeks to a patient with "hysterical epilepsy," with the apparent effect of suppressing the epilepsy. Otto Liebreich, a professor of pharmacology in Berlin, discovered in 1869 that chloral hydrate could relieve insomnia in patients with anxiety and depression (Shorter, pp. 198–199). This is the substance known as "knock-out drops" or "Mickey Finns" and indeed does what it is supposed to do. It is considered more reliable than morphine and can be taken orally (though it has a bad taste). It was widely used in mental hospitals thereafter, despite having some danger of addiction.

The Russian psychiatrist Sergei Korsakoff was the first to describe the syndrome named after him: chronic psychosis and memory loss in patients with alcoholism. Carl Wernicke in Germany had described the acute form of this condition in 1881—mental confusion and staggering gait. These conditions were subsequently identified as being due to a dietary deficiency in thiamine, often found in poorly nourished alcoholics.

The French physician Jean-Martin Charcot (1825–1893) was a key figure in the history of several modern specialties, including neurology, psychiatry, psychoanalysis, and clinical psychology. Charcot, the chief physician at the Salpetriere hospital in Paris, was known for his work on clinical-neuropathological correlations as well as on hysteria and hypnosis. Charcot was responsible for elucidating the neuropathology of amyotrophic lateral sclerosis, now known in the United States as Lou Gehrig's disease (in France, it is called "Charcot's disease"). Hypnosis, meanwhile, had been rehabilitated somewhat by the Englishman James Braid (1795–1860) under new terminology. Braid's 1843 book was entitled, *Neurypnology, or the rationale of nervous sleep*. Charcot successfully defended hypnosis before the French Academy of Sciences on February 13, 1882 (Gauld, 1992, p. 311) and used it in his "experimental therapeutics" of hysteria. Charcot considered hypnotism to be an abnormal process in itself, characteristic only of persons with hysteria. Charcot appointed Pierre Janet as head of a newly established psychology laboratory at the Salpetriere Hospital in Paris in 1890.

Charcot's contemporaries in Nancy, France, A. A. Liebault (1823–1904) and Hippolyte Bernheim (1840–1919), considered hypnosis to apply to normal people as well as to those with hysteria. Sigmund Freud translated Bernheim's book on hypnosis into German. Freud also visited Charcot in Paris. He returned to Vienna and in 1895 published a book that was the beginning of his development of psychoanalysis as a treatment for hysteria that went beyond hypnosis (Breuer & Freud, 1895/1986).

Spiritualists and mental healers throughout the nineteenth century had been interested in psychology and often identified themselves with psychology in their writings. One development in the United States in the nineteenth century was the emergence of Christian Science as a religion, based on the writings of Mary Baker Eddy (1821–1910). She had experienced improvement in her own state of illness in the hands of the healer Phineas P. Quimby (1802–1866) (Gauld, 1992).

Toward the end of the nineteenth century a coalescence began of the traditional "alienists" or psychiatrists, who worked mostly in isolated mental hospital settings, and physicians who were neurologists (or "nerve specialists"). The latter had an outpatient, office clientele.

Late in the nineteenth and early in the twentieth century, a field of "abnormal psychology" emerged in France and was continued in the United States. This was the result of the work of physicians, physiologists, philosophers, and psychologists including Theodule Ribot (1839–1916), Pierre Janet (1859–1947), William James (1842–1910), Granville Stanley Hall (1844–1924), and Morton Prince (1854–1929). Much of the research done by these workers can be considered to represent a nonpsychoanalytic offshoot of some of Charcot's work with hypnotism and hysteria. Ribot, who was trained as a physiologist, became chair of the department of experimental psychology at the College de France, where he specialized in abnormal psychology. Ribot was among those who influenced Charcot. Charcot's pupil Janet, later a professor at the College de France, interpreted neurotic complaints such as hysteria in terms of his concepts of dissociation and fixed ideas. Janet wrote an influential book on the psychological treatment of hysteria. William James, the author of the classic textbook on psychology in 1890, also lectured on abnormal psychology (Taylor, 1982).

Historian Eugene Taylor (personal communications, January 13–14, 1998) mentions several MD–Ph.D.s as being among those who "defined the interface between psychology and psychiatry" in this era and subsequently, Boris Sidis, Hugo Munsterberg, and Pierre Janet. According to Taylor, there were already clinics in France, Germany, the Netherlands, and the United States by the 1890s devoted to "medico-psychological" treatment. Boris Sidis did psychotherapy with adult outpatients at the Massachusetts General Hospital beginning in 1885. Caplan and Caplan (1969, pp. 304–305) considered Boris Sidis to have been the first psychologist to speak to the American Medico-Psychological Association. In this 1900 address, his opening words were:

> Medical men and especially alienists seem to observe less strictly the law "not to admit an alien into the congregation of the Lord," and are willing to listen even to a psychologist. (Caplan, 1969, p. 305)

At the time, Sidis held the position of associate in psychology and psychopathology at the Pathological Institute of the New York State Hospital. Thus, there is justice to Taylor's claim that these developments constitute some of the roots of clinical psychology.

The above developments were independent of the work of Lightner Witmer, a Wundt Ph.D., who founded the first psychology clinic at the University of Pennsylvania in 1896 and began to train a whole generation of Ph.D. clinical psychologists there in Philadelphia. He emphasized the role of clinical psychologists in helping others rather than only studying them (McReynolds, 1997; Routh, 1996). Unlike the members of the Boston school, Witmer took his cues from Itard, Seguin, and Montessori rather from Charcot and Janet. He emphasized work with children rather than adults and focused on academic and cognitive problems more than emotional ones.

1.01.9 THE TWENTIETH CENTURY: GUILD-RELATED ISSUES

The twentieth century has been a time of diversification in the number of professional groups providing services to the public in the area of mental health. These include psychiatry, psychoanalysis, psychology, social work, and newer groups such as marriage and family therapy. It has also been a time of impressive scientific advances in the study of psychopathology, assessment, and treatment. The scientific advances seem to cut across these disciplines and therefore can be discussed in a unified way. The professional developments of the century seem to require separate treatment for each competing guild. Let us therefore discuss each profession separately before proceeding to an account of scientific developments.

1.01.9.1 Psychiatry

Psychiatry began the century with most of its practitioners working in institutions. After it began to take over work from neurologists, its outpatient practice steadily increased. Neurology was a primary care specialty in the United States up until World War II but after that withdrew from such activities to become a field of consulting specialists, researchers, and teachers (Bluestein, 1992).

Philanthropist Henry Phipps was encouraged by psychiatrist Adolf Meyer to endow a psychiatric clinic at Johns Hopkins University. It was the first university psychiatric clinic of its kind in the United States and opened in 1913. In England psychiatrist Henry Maudsley (1835–1918) in 1907 offered funds to the London Council to establish a teaching hospital. Completed in 1915, the Maudsley Hospital ultimately became the locale of the Institute of Psychiatry at the University of London. In 1917 Emil Kraepelin founded the German Research Institute for Psychiatry in Munich. Shorter (1997, p. 262) calls this "the first comprehensive center for investigating brain, mind, and their disorders." Kraepelin's organization was later made a Kaiser Wilhelm Institute by the German Empire. It subsequently received funds from the Rockefeller Foundation and today is one of the Max Planck Institutes supported by the

German government. The Rockefeller Foundation began funding liaison psychiatry at the Massachusetts General Hospital and about 10 other hospitals beginning in 1925.

In 1934, the American Board of Psychiatry and Neurology was established and began to grant "board certification" in psychiatry to those who passed its examination. This provided to psychiatrists greater parity with their colleagues in other medical specialties who had long since achieved the possibility of "board status."

Psychiatry was especially influenced by psychoanalysis during and after World War II. The first psychiatrist to attain the rank of General in the US Army was William C. Menninger, who had a psychoanalyic orientation. According to Shorter (1977, p. 307), by 1945 most chairs of departments of psychiatry in US medical schools were psychoanalysts. In 1946, an organization called the Group for the Advancement of Psychiatry (GAP) was founded within the American Psychiatric Association, under the leadership of Menninger. This group favored psychoanalytic approaches. By the mid-1970s, very few American departments of psychiatry selected psychoanalysts as chairs. The instruction of psychiatrists in intensive psychotherapy was also on the wane. According to some observers, an era of biological psychiatry and psychopharmacology had begun.

Of course psychiatry is anything but a monolithic field and accommodates many viewpoints and approaches to treatment. For example, psychiatrist Victor Frankl was among those incarcerated in a Nazi concentration camp during World War II and believed that he had been able to survive the experience because he was able to find some meaning in life (Frankl, 1962). Accordingly, he developed a treatment approach he called Logotherapy aimed at helping his patients to find meaning in their own lives.

The "antipsychiatry" movement is also worth mentioning in discussing the history of twentieth century psychiatry as a profession. Psychiatrist and psychoanalyst Ronald D. Laing came to voice strong criticism of both psychiatry and of the families of persons with schizophrenia (Laing, 1960). His view was that parents tended to tie their children in mental knots and threaten their identity. Laing thought of schizophrenia as being in many cases not so much a breakdown as a breakthrough, a necessary stage on the path of development. In 1967 Laing reviewed Michel Foucault's book, *Madness and civilization* (Foucault, 1961) in the *New Statesman* and thereafter considered Foucault to be a sort of ally. Foucault spoke of the construction of so many public mental hospitals after the French revolution as "the great confinement" and tended to view this as a sort of conspiracy by psychiatry against human freedom. Of course, at this time, Foucault was hardly alone in being critical of residential treatment for psychopathology. This type of criticism, together with the existing disillusionment with institutions, and the availability of effective neuroleptic medications, hastened the general trend toward deinstitutionalization. By this time most psychiatrists had established an outpatient practice in any case.

1.01.9.2 Psychoanalysis

Psychoanalysis virtually swamped the field of psychotherapy early in the century, relying on the treatment principles of free association and interpretation by the analyst. Freud's writings attracted many disciples. Also, outpatients were willing to pay for this kind of treatment for periods of months and years, if so advised. Psychoanalysis soon began organizing itself in a formal way. The first international psychoanalytic congress was held in Salzburg in 1908. Freud and Jung were invited to the United States in 1909 to speak at Clark University. Psychoanalysis also soon began a process of fission, with Alfred Adler (1870–1937) breaking away in 1911. Then in 1913 Carl Jung (1875–1961) left the fold. In that same year, Freud and certain colleagues set up a small group called "the Committee" to maintain more oversight of psychoanalytic orthodoxy. Interestingly, each of its offshoots seemed to follow Freud's pattern of elaborating a theory of psychopathology and treatment and establishing formal organizations that its adherents could join.

Freud thought of psychoanalysis as an independent field, separate from psychiatry and medicine, and was willing to train analysts (including psychologists) who lacked a medical background. The American Psychoanalytic Association, founded in 1911, had ideas different from Freud's about the relation of psychoanalysis to medical practice and after 1938 for 50 years restricted its accredited institutes to offering training to medical candidates only. In 1926, on the other side of the Atlantic, psychologist Theodore Reik had been tried for practicing medicine without a license, that is, for doing psychoanalysis. Reik was acquitted. In fact, Freud himself came to Reik's defense (S. Freud, 1927). In 1985 Bryant Welch (a psychologist–attorney) and three psychologist colleagues filed a lawsuit against the American Psychoanalytic Association for this practice as a violation of federal antitrust law. In 1988 the organization settled the suit out of court and began allowing social workers and psychologists to become trainees. According to Shorter (1997, p. 310), the number of nonmedical candidates in training soon rose to above 20%.

In France, psychoanalysis emerged as a highly influential approach, but a number of years after it had done so in the United States (Roudinesco, 1990). Psychoanalysis in France began with the work of Marie Bonaparte and her colleagues in the 1920s (Bertin, 1982). This included the founding of an institute and a professional journal as well as the translation of some of Freud's writings into French. Psychiatrist and psychoanalyst Jacques Lacan (1901–1981) was active on the French psychoanalytic scene beginning in the 1930s but began his influential seminars in Paris only in 1953. The legal issue of lay vs. medical analysis also arose there. After the student disturbances of 1968, psychoanalysis assumed a greater representation in French university curricula. In 1973 Lacan published a book on what he considered the four fundamental concepts of psychoanalysis. Lacan's "deconstructionist" approach was popularized in literary circles by Jacques Derrida. Even today, if one examines the shelves of a bookstore in Paris, the section on psychoanalysis is highly visible, but there may be no sections labeled psychology or psychiatry.

The influence of psychoanalysis in Britain and the United States had been evident well before that in France. The Rockefeller Foundation provided some funding for the Chicago Psychoanalytic Institute in 1925. Franz Alexander (1891–1964) was appointed professor of psychoanalysis at the University of Chicago in 1931, the first such academic recognition for the field (Mora, 1994). The role of psychoanalysis in the United States and Britain increased during the 1930s, with the immigration of many analysts from continental Europe fleeing Hitler. As the American interest in psychoanalysis cooled somewhat in the 1970s, Roy Schafer (1976) and others began to reconsider whether psychoanalysis ought to model itself after the natural sciences. Perhaps, Schafer suggested, it would be more appropriate to consider psychoanalysis to be part of the humanities. Its use of hermeneutic approaches in interpreting patients' utterances were seen as similar to those of literary scholars in interpreting poems and fictional narratives. The philosopher of science Adolph Grunbaum (1984, 1993), an outsider to psychoanalysis, criticized its use of a patient's subsequent behavior as a way of confirming the correctness of psychoanalytic interpretations. Other criticisms came from insiders, such as psychoanalyst–historian J. M. Masson (1984). Nevertheless, psychoanalysis continues as a viable profession and as an influence on scholars over much of the world. Its influence continues to spread in Latin America and elsewhere during the late twentieth century.

1.01.9.3 Clinical Psychology

The Boston physician Morton Prince viewed abnormal psychology as a subject belonging more to philosophy or the liberal arts than to medical school. He was influenced by Pierre Janet and William James and published a book on "Christine Beauchamp," a famous case of multiple personality disorder he had encountered. Prince founded the *Journal of Abnormal Psychology* in 1906 and later gave it to the American Psychological Association. He also founded the Harvard Psychological Clinic in 1926. It was a research facility rather than one for providing services to the public and later received funding from the Rockefeller Foundation. Another of Prince's projects was the American Psychopathological Association, open to membership for both psychiatrists and psychologists. If Prince had had his way, the fields of psychology and psychiatry would probably be more unified than they have turned out to be. His aims have been fulfilled much more in the domain of research than that of practice.

Eleven years after his clinic had been established, Lightner Witmer founded a journal called the *Psychological Clinic* in 1907 and in its inaugural issue elaborated on his concept of what the field should involve. This was the first known use of the term clinical psychology. Witmer's journal continued to publish about 23 volumes, well into the 1930s, mostly case histories of children treated in his own clinic at the University of Pennsylvania. Witmer's idea of clinical psychology focused on intervention but resembled what we now think of as special education more than psychotherapy.

According to Eugene Taylor (personal communication, January 13–14, 1998), the first official Ph.D. clinical psychologist at the Massachusetts General Hospital was L. Eugene Emerson, who as early as 1911 was using a type of psychotherapy derived from the work of Freud and Jung. Another Ph.D. psychologist, F. Lyman Wells (1884–1964), was doing psychotherapy with inpatients at the McLean Hospital at the same time. The theories of Freud, Jung, Adler, and other psychodynamic views also influenced academic psychology, especially in the clinical area but also in the areas of personality, developmental, and social psychology.

James McKeen Cattell (Ph.D. Leipzig, 1886) was one of the strongest supporters of the idea of psychology as a profession but did not have much interest in clinical issues. He founded the Psychological Corporation in 1921 as a non-profit organization to promote applications in psychology, mostly psychometric testing. His interests soon turned to science at large. He bought *Science* magazine from Alexander Graham Bell, and served as its editor for many years. Cattell also published *American Men of Science* (later known as *American Men and Women of Science*) and inaugurated the practice of highlighting the relative merit of scientists' contributions by putting a star beside the names of those considered the most notable. In this way he carried forward Galton's idea of the importance of individual differences.

In the era before World War II, clinical psychologists were known more as mental testers than as psychotherapists. The mainstay of their early activities in this area was the Binet test, first published in France in 1905, translated into English in 1908, and eventually standardized in the United States. Psychologist Henry Goddard, who first had the Binet translated, was the research director at the Vineland Training School in New Jersey (an institution for persons with mental retardation). He set up the first internship program in clinical psychology in that same year, 1908. Internships involving a year of full-time supervised experience soon became a part of the expected training of clinical psychologists. Goddard's research was supported financially by Samuel S. Fels, the soap maker, and some friends from the Ethical Culture Society of Philadelphia. Fels and Goddard were part of an informal group that met at Vineland for some years that called itself humorously, "the Feeble Minded Club" (Doll, 1988).

In 1917, the American Association of Clinical Psychologists was founded. It was the first organization in the field of clinical psychology and came about because its founder J. E. W. Wallin (1876–1969) wished to begin to establish some professional standards. In particular, he was concerned that Binet testing was being done by school teachers with no background in psychology on the one hand, and by academically trained psychologists with no clinical experience on the other. Only two years later, this organization disbanded itself in favor of a new Clinical Section within the American Psychological Association. For a brief time, the American Psychological Association tried to develop a procedure for certifying what were at the time called "consulting psychologists," meaning those who were adequately prepared to offer their services to the public. The person who was in charge of this effort was F. Lyman Wells, already mentioned as an influential figure in early clinical psychology. The APA's efforts in this domain did not suceed (only 25 persons were ever so certified) and were discontinued in 1927.

Psychologist Leta Hollingworth (1886–1939) published a little-cited paper in 1918 that was the first to mention the possibility of a different type of degree for professional psychologists, the "doctor of psychology" instead of the more usual Ph.D. (doctor of philosophy). The Psy.D. was discussed as a possibility for many years but did not come into much actual use until the 1960s. One of the first programs that granted the Psy.D. began in 1968 at the University of Illinois. In the United States, Britain, and some other European countries, the Psy.D. and other similar degrees such as the "doctor of clinical psychology" (D.Clin.Psy.) are now more often offered by training programs for practioners. Research-oriented graduate programs still give the Ph.D. in clinical psychology.

By the time of World War I, self-report questionnaires began to be developed for the assessment of personality and psychopathology. By the 1930s, projective tests such as the Rorschach inkblot test and the Thematic Apperception Test joined what was becoming the clinical psychologist's standard test battery.

In 1937 the American Association of Applied Psychologists (AAAP) was founded. The Clinical Section within the American Psychological Association disbanded itself, and its members formed the Clinical Section of the AAAP. In 1945 Division 12 of the American Psychological Association was founded, the present Division of Clinical Psychology. It was based on a merger of two groups: one of these was the psychologists from the Clinical Section of the AAAP, who in a sense represented the tradition established by Lightner Witmer of Philadelphia. The other group briefly in 1945 carried the the label of the Division of Abnormal Psychology and Psychotherapy and was to have been APA Division 11. This group could perhaps be seen as the descendents of the Boston school of abnormal psychology. Division 11 was never officially established; instead its potential members became part of Division 12, which was initially called the Division of Abnormal and Clinical Psychology.

In 1945, the American Psychological Association appointed David Shakow (1901–1981) to head a committee to discuss how clinical psychologists should be trained. The impetus for this committee's work was provided by the promise of massive funding of such training in all mental health fields by the Veteran's Administration (VA) and the new National Institute of Mental Health (NIMH). James Grier Miller who had been trained in both psychology and psychiatry was chief of the VA

Division of Clinical Psychology and Neuropsychiatry at the time. The Shakow Committee's recommendations were later endorsed by the so-called Boulder Conference in 1949. Shakow's model called for training clinical psychologists at the Ph.D. level as "scientist-practitioners," including a required one-year predoctoral internship, usually obtained during the trainee's third year of graduate school. As Levy (1962) later pointed out, many psychologists trained as "scientist–practitioners" never went on to publish any research; many did not even publish their Ph.D. dissertations. Thus, this training model did not seem to function as well as had been hoped. As time went on, more of the graduates of such programs also went into private practice, an outcome that no doubt would have horrified David Shakow and his colleagues.

The first state law certifying psychologists for independent practice was passed in Connecticut in 1945 (Heiser, 1945), and by 1977 all states had such certification or licensing laws. The distinction between these two types of legislation is that certification usually protects only the title (in this case "psychology"), whereas licensing is more likely to specify the actual content of the practice. Similar legislation was passed in each of the Canadian provinces. It is the goal of most national psychological associations in European countries to achieve full legal recognition for professional psychology, if indeed they had not already attained it (McPherson, 1986).

In 1947, the American Board of Examiners in Professional Psychology (ABEPP, later simply ABPP after the word "Examiners" was dropped from the Board's name) was founded. This board soon began to grant diplomas in clinical psychology and was modeled after boards that were well established in various medical fields including psychiatry. In 1948, Nicholas Hobbs (1915–1983) carried out a project initiated by the American Psychological Association to develop an explicit set of ethical standards for psychologists. Revisions of these ethical principles continue up to the present.

In London after World War II, psychologist Hans J. Eysenck set up a Department of Psychology within the Institute of Psychiatry at the University of London. This department became a major center for behavior therapy in Britain in subsequent decades.

In his 1959 book, psychologist Alan Ross (1921–1993) noted the fact that after the end of World War II, clinical psychology had moved far from some of its origins focusing on work with children. This was partly because of all the government funds that encouraged work with adults such as military veterans. Ross wished to reclaim what he viewed as the original heritage of the field. He stimulated the formation of a new subgroup of persons who identified themselves as clinical child psychologists. Psychologist Nicholas Hobbs was one of these and was particularly effective in suggesting how support and treatment could be arranged for disturbed children, coordinating work with them individually, with their families, and with the schools (Hobbs, 1966).

In 1962, a conference on behavior modification was held in Charlottesville, Virginia. It was instigated by psychiatrist Joseph Wolpe, behavior therapist Andrew Salter, and psychologist Leo Reyna, who published the proceedings as a book (Wolpe, Salter, & Reyna, 1964). The American Association of Behavior Therapies (AABT) was founded relatively soon after this landmark conference, in 1966 (in 1969 the "Therapies" in the organization's name was changed to "Therapy"). Despite the interdisciplinary origins of this field, most behavior therapists were psychologists. Until 1972 the AABT met with the American Psychological Association; after that, it met independently.

Like other professional groups (including psychoanalysts), behaviorally-oriented practitioners showed some tendency toward fission. Applied behavior analysis since its beginning tended to isolate itself from behavior therapy with separate professional organizations, journals, and so on (Krantz, 1971). This branch of behavioral psychology dates its own history from a 1947 research conference in Bloomington, Indiana under the aegis of B. F. Skinner. Unlike other behaviorists, the followers of Skinner avoided using dispositional concepts, preferred single-subject experiments to those with group designs, and emphasized the "three-term contingency" involving setting events, response, and environmental consequences.

In 1963, a new component of the National Institutes of Health was created, called the National Institute of Child Health and Human Development (for which the usual acronym is NICHD). Part of the impetus for this new research institute was the interest of President John F. Kennedy and his family in the problem of mental retardation. Research grants related to this problem were shifted from the NIMH to the NICHD, which also had the practical effect of divorcing psychiatry from the medical aspects of mental retardation, which were henceforth considered to be more in the domain of the pediatrician. The creation of this agency also propelled psychology into the role of being the primary discipline concerned with psychopathology associated with mental retardation.

The California School of Professional Psychology was founded in 1969, with multiple campuses not affiliated with any university. The founder was psychologist Nicholas Cummings, who had the backing of the state psychological association in California. From the standpoint of those who founded this and other free-standing schools of professional psychology that soon grew up all over the United States, the existing graduate schools' programs were too elitist and not sufficiently concerned about producing practitioners. A conference of psychologists at Vail, Colorado in 1973 urged the acceptance of a practitioner model of training alongside the now traditional scientist–practitioner model (Korman, 1974). The Vail recommendations also endorsed the trend toward schools of professional psychology.

With the growing dominance of the US health-care system by insurance and HMOs, psychologists were having trouble being reimbursed by third parties for their professional services. The insurance companies complained about the fact that psychology licensing was usually generic and did not distinguish between those qualified to deliver "health care services" and others (such as industrial psychologists) who might not have such training. For this reason, in 1975 psychologists established a National Register of Health Service Providers in Psychology. Each psychologist on this register had to have not only a valid state license but also training in providing health services, such as a formal clinical internship.

In 1988 the American Psychological Society (APS), consisting mostly of research and academic psychologists, split off from the American Psychological Association, which was increasingly dominated by professional practitioners. Many research and academic clinical psychologists felt torn in their loyalties between these two competing organizations.

1.01.9.4 Social Work

In 1917, the first National Conference on Social Work was held in the United States. In 1917 Mary Richmond (1861–1928), a prominent figure in social work, published a book, *Social diagnosis*, which attempted to define the nature of social case work. The existing 15 schools of social work in the United States joined a new association that was established in 1919. At the time, nine of them were linked to universities and the other six to operating social agencies instead, which resembled the long-standing pattern for hospital-based schools of nursing (Leiby, 1978). Schools of social work agreed among themselves in 1939 that the standard educational pattern in the field would be two years of graduate study, resulting in a Master's degree in Social Work (MSW).

By 1945, an organization of social workers in the United States already had 2000 members. In 1955, the National Association of Social Workers (NASW) was established as a combination of several groups. According to Shorter (1997, p. 293), the NASW increased to 80 000 members by 1990, about one-quarter of whom were in private practice.

1.01.9.5 Religious Counseling

Many mental health workers have tended to distance themselves from religion, although they might admit, if pressed, that their work was not unrelated to some of the counseling done by members of the clergy or of religious orders. One twentieth century exception to the attempt to keep mental health and religion separate was the Emmanual Movement. The Rev. Elwood Worcester (1863–1940), who had a philosophy Ph.D. from Leipzig in 1889, in 1904 became rector of the Emmanual Church in Boston. In 1905, he and his colleagues attempted to integrate a mental health clinic into the activities of this church, attracting wide publicity in the press. The medical community of Boston reacted somewhat adversely and uncooperatively. However, historian Eugene Taylor notes that the Emmanual Movement was but "chapter one" of the Clinical Pastoral Education Movement.

Much more recently, science-fiction writer L. Ron Hubbard (1950) launched "dianetics" as an alternative to traditional psychotherapies. In some respects it seems to have been modeled after psychoanalysis. Hubbard went on in 1954 to establish Scientology as a church, at least in terms of its legal status. Shorter (1997) points out that Hubbard opposed ECT and that he appointed Thomas Szasz as Scientology's "consulting psychiatrist." Scientologists mounted campaigns against the use of the drug Ritalin for children with attention deficit disorder. They are equally against the use of Prozac.

1.01.9.6 Interdisciplinary Activities and Advocacy

Mental health professionals cannot justify their existence simply as a way to make a living. They must demonstrate that their activites benefit their patients and society at large. One of the earliest effective consumer advocates was Clifford Whittington Beers (1876–1943), who published the book, *A mind that found itself* in 1908, about his own experiences as a mental patient. A graduate of Yale, Beers had a mental breakdown at age 24 in which he attempted suicide and was hospitalized at the Hartford Retreat. Subsequently, he experienced a manic episode and was an inmate at the Connecticut State Hospital. The book he wrote later described these experiences, not without some bitterness about the way he had been treated at the time, and suggested the need for better practices in this field. His advocacy was endorsed by such influential figures as William James at Harvard and Adolf Meyer at Johns Hopkins, and in 1909 Beers became one of the founders of the National Committee on Mental Hygiene, an organization that still exists. The Rockefeller Brothers supplied a grant to assist the fledgling movement.

A child guidance clinic, attached to the local juvenile court, was a new innovation begun in Chicago in 1908 by physician William Healy (1869–1963) and psychologist Grace Fernald (1879–1950). The Commonwealth Fund (established by the Harkness family) in 1919 made a grant to the National Committee on Mental Hygiene that supported the founding of a series of child guidance clinics around the United States and in other countries. The Commonwealth Fund also provided support to several child guidance clinics in England beginning in the 1920s (Wooldridge, 1994). The typical staffing of a child guidance clinic was an interdisciplinary team consisting of a psychiatrist, psychologist, and social worker. The usual role of the psychiatrist was diagnostic interviewing and, when feasible, psychotherapy. That of the psychologist was mental testing, and that of the social worker was dealing with family members and other community agencies.

As an outgrowth of the child guidance movement in the United States, a new organization was formed in 1924 called the American Orthopsychiatric Association. The Greek word "ortho" means straight, and juvenile courts and child guidance clinics were intended to help youth "go straight" and avoid crime. The first president, not surprisingly, was William Healy. Membership was at first restricted to psychiatrists but in 1926 was broadened to include all mental health professionals. For many years "Ortho" has served as an interdisciplinary meeting place for practicing psychiatrists, psychologists, and social workers. It has also served as an organization advocating for the needs of those, especially the disadvantaged, who were in need of mental health services.

After World War II, the National Institute of Mental Health in the United States, created in 1946 and up and running in 1949, began a program of training grants for all mental health fields (Rubinstein, 1992) that lasted into the Reagan era of the 1980s (Benjamin, 1986). These funds were responsible for huge increases in the number of clinical psychologists and psychiatrists who were trained. Similar increases in funding were also seen in Britain, the Scandinavian countries, and elsewhere in Western Europe as a part of various national health systems. The success of the Soviet Union in putting the satellite Sputnik in orbit in 1957 spurred the US government to even greater efforts in funding research and training in all scientific fields.

In London in 1947, the National Society for Mentally Handicapped Children was founded. It was the first of many groups founded after World War II in various countries primarily by parents of children with mental retardation, cerebral palsy, and many other handicapping conditions. These groups proved to be quite effective in advocating that services be provided for their children, particularly in the area of education. They also advocated that research be funded, including research on prevention. One of the most effective of the parent-advocates for persons with mental retardation was Elizabeth M. Boggs (1915–1996). Besides being the parent of a handicapped son, she was a Cambridge-trained Ph.D. in chemistry and readily trained herself to follow the intricacies of government legislation related to services and research. She was therefore highly effective as a lobbyist for several decades.

In 1963 the Kennedy administration sponsored a bill to establish a nationwide network of interdisciplinary Community Mental Health Centers (CMHCs), which indeed came into being and still exists. The federal government did not plan to continue the funding of these centers indefinitely and relatively quickly shifted their support to states and local communities. It is by now widely acknowledged that these centers were not able to provide as much community support as was hoped for by sponsors of the legislation. Many of the mental patients being discharged from mental hospitals around the country were thus thrown back on their own resources and those of their families. The administration of a CMHC can be in the hands of persons trained in any of several different mental health-related disciplines.

The "antipsychiatry" movement that began in the 1960s was critical not only of psychiatry proper but of the mental health establishment in general. The novel by Kesey (1962), *One flew over the cuckoo's nest*, criticized mental hospital treatment, focusing specifically on ECT and lobotomy. The psychiatric nurse was portrayed simply as part of an oppressive hospital team. The novel was made into a movie in 1975, which no doubt reached an even larger audience. At about the same time, sociologist Erving Goffman published his influential critique of the mental hospital as a "total institution" in which patients suffered numerous abasements, degradations, humiliations, and profanations (Goffman, 1961). Goffman's book was based on field work at St. Elizabeth's Hospital, a government-supported institution in Washington, DC at a time when about 6000 mental patients were housed there.

In 1968 Bengt Nirje, director of the Swedish Parents Association for Mentally Retarded Children, developed the concept of normalization, which was first applied in the field of mental retardation. The idea meant that the goal of the service system is to assist the individual to live in as normal a way as possible, in spite of the presence of any handicapping condition: living at home, going to an ordinary school, having some kind of job or responsible role in the household, shopping at regular stores in the community, being treated by the usual primary physicians, etc. Along with this concept went the idea of "self-advocacy" that individuals with physical or mental disabilities should themselves try to better their situation and not wait for some other group such as parents or professionals to do this for them (Shapiro, 1993).

Advocacy is not just the business of parent groups or of those with mental problems themselves. It is also the business of lawyers, and the mental health bar became very active in the latter half of the twentieth century. One important case was the suit of Wyatt v. Stickney, originating in Alabama in 1971. In this case, the court declared that mental patients had a right to treatment, and that they could not be confined in a state mental hospital without proper treatment for their condition. One side effect of such cases was simply to speed up the course of de-institutionalization. This unfortunately meant that many mental patients not only did not get much treatment, they were sometimes released and had to fend for themselves as best they could, on the streets (Isaac & Armat, 1990).

Behaviorally-oriented token economies also sometimes ran into legal trouble. In a token economy, points are assigned to patients or clients contingent on certain predefined appropriate behaviors, and these points (or tokens) can be exchanged for preferred items or privileges. Kazdin (1978) describes one such situation in a Missouri program sponsored by the Federal Bureau of Prisons in 1972:

> [This project] was terminated because of pending litigation that challenged the involuntary assignment of prisoners to the program without due process, deprivation of privileges normally provided in prisons, aversive conditions used as part of treatment, and other aspects of contingencies. (p. 292)

In a related US Supreme Court decision on the case of Doe v. Bolton (1973), Justice William O. Douglas spoke at length about the right of an individual to be let alone, plan one's own affairs, shape one's own life, do what one pleases, be free from bodily restraint or compulsion, and be free to work, stoll, or loaf. Such words bring back memories of the Bicetre Hospital in Paris in 1789.

In 1975, federal legislation was passed in the United States guaranteeing all children a free appropriate education in the public schools, regardless of handicapping condition. This law was of tremendous importance especially to parents whose handicapped children had been excluded from the public schools or who had not been provided with educational accommodations to their special needs. As an unanticipated side effect, this legislation fueled a tremendous amount of attention to assessing children to determine their eligibility for special education programs (Milofsky, 1989). Also, such special programs at times could seem very expensive to school boards, giving them reason to set up procedural barriers against such placements. This law was both a meal ticket and a major source of frustration to clinicians working with children and to those working within the schools. These laws were revised as recently as 1997 under the heading of the Individuals with Disabilities Education Act, which had much more emphasis on mainstreaming and inclusion. It is too early to see what effects and side effects these changes will have.

The National Alliance for the Mentally Ill (NAMI) was founded in 1979 as a Washington-based national advocacy group, consisting largely of the parents and friends of people with major mental illness such as schizophrenia, depression, or bipolar disorder. As of 1996 it had about 130 000 members. The NAMI view of mental illness as a type of "no fault brain disorder" successfully influenced Congress and the NIMH to give greater emphasis to biologically-oriented research in psychiatry and psychology.

1.01.10 RESEARCH ON PSYCHOPATHOLOGY IN THE TWENTIETH CENTURY

1.01.10.1 General Paralysis of the Insane

In 1913, bacteriologist Hideo Noguchi (1876–1928), working at the Rockefeller Institute, discovered *Treponema pallidum* (the organism causing syphilis) in the brains of paretics and patients with tabes dorsalis, confirming his hypothesis concerning the cause of the disturbance (Clark, 1959). While such discoveries were welcomed, they did have the effect of making the diagnosis and treatment of syphilis and of general paralysis of the insane (GPI) the province of internal medicine rather than that of psychiatry. The psychosocial aspects of syphilis, like those of other sexually transmitted diseases including the AIDS virus, remain of interest to the behavioral sciences under the new guises of behavioral medicine and health psychology.

1.01.10.2 Dementia

In 1901 the physician Alois Alzheimer, working at a Frankfurt hospital, encountered a patient with a dementing condition who upon autopsy had characteristic tangles and neuritic plaques in the brain. This condition is now known to us as Alzheimer's disease (O'Brien, 1996). According to Shorter (1997, p. 103), in 1995 a gene was discovered on chromosome 14 and another one on chromosome 21 that were linked to Alzheimer's disease.

The current focus within the behavioral sciences is probably more on the difficult situation of family caregivers of Alzheimers patients than on the dementing condition itself. Theirs is a high stress situation, and much remains to be discovered about how to offer optimal support to them.

1.01.10.3 Dementia Praecox/Schizophrenia

The eminent German jurist Daniel Schreber (1903) wrote an autobiographical account of his nervous illness which has usually been thought to be schizophrenia. Schreber's story was used by Freud (1911/1959) to outline his concepts of the role of homosexuality in the genesis of paranoia (with statements such as "I love him" being transformed mentally into their opposites, such as "he hates me"). Shorter, a biologically oriented historian of psychiatry, reinterpreted the evidence to argue that Schreber may have actually suffered from neurosyphilis (Shorter, 1997, p. 80).

The name schizophrenia was suggested as a substitute for dementia praecox by Eugen Bleuler (1908, 1911), who was professor of psychiatry at Zurich and director of the Burgholzli hospital there. In contrast to Kraepelin, Bleuler preferred to define this group of disorders in terms of their symptoms rather than relying so much on the chronic course of the psychosis. He listed the core symptoms as including loose associations, flat affect, autistic thinking, and ambivalence, with the symptom of loose associations being the most influential concept. Psychiatrist Carl Jung was originally trained by Bleuler at the Burgholzli and early in his career participated in research on word associations in schizophrenia.

Kurt Schneider (1887–1967), a professor of psychiatry at Heidelberg, came up with an influential list of "first rank symptoms" which he considered to be characteristic of schizophrenia. An example of such a symptom would be mentally hearing one's voice spoken aloud.

One of the first genetic studies of schizophrenia was carried out by Hans Luxenburger in 1928. He was a psychiatrist who worked at Kraepelin's Institute at Munich. He searched through 16 000 birth records in Bavaria and compared them with a patient list, finding 211 patients who were part of a twin pair. Of those who were monozygotic, the concordance rate for schizophrenia was higher than that for dizygotic twins (Shorter, 1997, p. 241).

In contrast to this approach based on genetics, the German refugee psychoanalyst Frieda Fromm-Reichman (1899–1957) emphasized the role of experiences provided the developing child by the "schizophrenogenic mother." Fromm-Reichman was noteworthy for her skill as a psychotherapist for patients with schizophrenia. She began her work at Chestnut Lodge in Maryland in 1935 and was portrayed in the novel *I never promised you a rose garden* as "Dr. Fried." She was part of a group of neo-Freudians in the United States that included psychiatrist Harry Stack Sullivan (1892–1949).

Psychiatrist Franz J. Kallmann (1897–1966), who was trained at Kraepelin's Institute, fled Hitler in 1935 to work at the New York State Psychiatric Institute, associated with Columbia University. He continued his genetic research, finding higher concordance rates for schizophrenia in monozygotic than dizygotic twins, with the corresponding percentages being approximately 85% vs. 15% (Kallmann, 1938, 1946, 1953).

Seymour S. Kety and his associates were involved in a Danish adoption study that was more refined in its methods than previous research, in part because of the quality of public record keeping in Denmark. For Copenhagen adoptees with schizophrenia, this disorder was found to be 10 times more common in adoptees' biological relatives than in their adoptive families (Kety et al., 1968). These findings were essentially replicated in the rest of Denmark (Kety, Wender, Jacpbsen, Ingraham, et al., 1994).

Psychologist Paul Meehl (1962) presented an integrative theory of how schizophrenia might come about, under the title of "schizotaxia, schizotypy, and schizophrenia." Schizotaxia refers to the hypothetical genetic predisposition that is considered necessary but not sufficient to produce the disorder. Schizotypy is the behavioral phenotype produced by this genotype—not necessarily abnormal, but a condition of vulnerability to the disorder. Perhaps schizotypes find interpersonal relationships a bit more aversive than others do, or manifest some subclinical cognitive slippage on occasion, but nothing symptomatic. Full blown schizophrenia is what is seen when a schizotype is subjected to certain stressful physical or psychological conditions.

Psychiatrist Manfred Bleuler, the son of Eugen Bleuler (who had coined the term "schizophrenia"), was like his father a professor in Zurich. M. Bleuler published a study of 208 patients with schizophrenia whom he had followed up for over 20 years (M. Bleuler, 1972). This study confirmed the belief that the disorder was quite variable in its course and outcome. He was not, however, able to realize his ambition of uncovering a Mendelian etiology for schizophrenia, which continues to be elusive.

Psychologist Philip Holzman and his colleagues showed that patients with schizophrenia and their first-degree relatives were more likely to have abnormal smooth pursuit eye movements than control subjects (Holzman et al., 1973). In this procedure, patients are simply asked to try to follow a moving pendulum with their eyes. The abnormality consists of a jagged pattern of eye tracking rather than a smooth one. The finding has been replicated by other investigators and turns out to be a "trait marker" that does not change when an individual decompensates into psychosis or recovers from schizophrenia. The research provides further support for some kind of genetic hypothesis concerning schizophrenia.

For a time, the "dopamine hypothesis" concerning schizophrenia was a serious scientific contender (see e.g., Snyder, 1976). The appeal of this notion was that virtually all of the effective neuroleptic medications seemed to have in common the property of blocking dopamine receptors in the brain. When other equally effective antipsychotic drugs such as clozapine were discovered that did not block dopamine, the idea lost some of its attractiveness.

1.01.10.4 Schizoaffective Disorder

This term was coined by Russian psychiatrist Jacob Kasanin in 1933 to describe a condition with symptoms of both schizophrenia and depression. Its prognosis tended to be intermediate, that is, better than that of schizophrenia but worse than that of depression.

1.01.10.5 Depression

It was thought at first that phenothiazine medications might be rather specific to schizophrenia rather than depression, supporting the validity of the distinction between these two groups of disorders. However, it turned out that depression also responds well to such medications (Klein & Fink, 1962).

Psychiatrist A. T. Beck was one of the first to elaborate a theory explaining how the way a person thinks could contribute to depression, for example, negative thoughts about oneself, the world, and the future, and cognitive distortions such as making internal, global, and pervasive attributions for negative events ("this happened because of me," "I am that way about everything," "that is just the way I am") (A. T. Beck, 1963).

Psychologist Martin Seligman linked up his research on experimentally produced learned helplessness in dogs with similar phenomena seen in humans and moved into research on helplessness and human depression (Seligman, 1975).

1.01.10.6 Neurosis, Including Anxiety

Janet and Raymond (1903) in France wrote about obsessions and psychasthenia (which literally translates as "mental weakness"). Janet considered psychasthenia to be characterized by pathological feelings of inadequacy.

One interesting new development was the discovery of the phenomenon of "experimental neurosis." According to Kazdin (1978), in 1912, Mariya N. Yerofeyeva (1867–1925) was doing research in the St. Petersburg laboratory of Ivan Pavlov (1849–1936), conditioning a dog. As usual, the unconditioned stimulus was food-powder and the unconditioned response was salivation. In this case the conditioned stimulus was a mild electric shock on the dog's skin, which might be expected to elicit defensive reactions competing with salivation. Still, the conditioning was going well up until the point when the shock was administered to one new location after another on the animal's skin. When this was done, conditioning broke down, and the animal developed a lasting disturbance of its overall behavior. In 1913 another of Pavlov's co-workers made a similar discovery. Nataliya R. Shenger-Krestikovnikova (1875–1947) was conditioning a dog to discriminate between a circle and an ellipse. The dog was able to do so at first, but conditioning broke down when the discrimination became too difficult and was replaced by enduring disturbed behavior (Pavlov, 1927).

John Watson (1878–1958) and Rosalie Raynor (1898–1935) published in 1920 their study of "little Albert" that is by now well known (Watson & Raynor, 1920). The child was repeatedly presented with a furry animal in the presence of a loud noise and evidently learned to fear the animal and other objects similar to it as a result of this experience. The authors presented the study as a demonstration of how phobias might originate. Watson (1913) had presented behaviorism as a "purely objective experimental branch of natural science. Its theoretical goal is the prediction and control of behavior" (Watson, 1913, p. 158).

Behaviorists were not the only ones contributing to the study of the phenomena underlying neurotic behavior in this era. Anna Freud (1895–1982) published a well-received book on *The ego and the mechanisms of defense* in 1937. Neurotic symptoms were interpreted as functioning as defensive maneuvers to ward off anxiety.

The psychoanalyst J. H. Masserman (1943) did experiments with cats, putting them into approach-avoidance conflict situations. This work contributed to the thinking of subsequent behavioral researchers concerned with neurotic behaviors who, however, did not necessarily accept some of Masserman's psychoanalytic speculations about his findings. Indeed, both psychoanalysis and behaviorism remain for the most part mutually isolated systems of thought to this day.

Howard S. Lidell (1895–1962) at Cornell University in 1956 reported research producing "experimental neurosis" in dogs, sheep, goats, and rabbits. This showed that the phenomenon was not specific to dogs nor to the particular procedures used in Pavlov's laboratory.

After World War II, the psychoanalytic approach to neurosis took a turn toward "ego psychology," recognizing the existence of a relatively conflict-free sphere of functioning that was little affected by neurotic disorders (Hartman, Kris, & Loewenstein, 1946).

In 1969 British psychiatrist Isaac M. Marks published an influential article suggesting that there might be important genetic factors in fear and anxiety disorders including agoraphobia, panic disorder, and obsessive-compulsive disorder.

Psychiatrist Donald F. Klein found in his research on psychopharmacology that anxiety patients (unlike those who were depressed) did not get better on chlorpromazine. Panic disorder, as a condition distinct from depression, also did not respond to chlorpromazine (Klein, 1964; Klein & Fink, 1962).

1.01.10.7 Shell Shock/Post-traumatic Stress Disorder

W. H. R. Rivers (1864–1922), a British physician, experimental psychologist, and anthropologist, became involved in treating the war neuroses referred to as "shell shock" during World War I and soon developed the view that they were psychologically based rather than being simply a mechanical result of combat trauma (Ackerknecht, 1942). This view is preserved in the present concept of post-traumatic stress disorder (PTSD).

1.01.10.8 Stress

Stress is not in itself considered a type of psychopathology but is considered a risk factor not only for psychopathology but also for various types of physical illness. One of the more influential researchers in this field is psychologist Richard S. Lazarus (1966). His writings define stress as a set of demands that exceed the adaptive capacity of the person. It is important to know not only the physical characteristics of the stressful situation but how these are appraised by the individual, and what coping strategies are adopted to try to deal with the stress.

1.01.10.9 Psychosomatic Disease

Physician Flanders Dunbar (1902–1959) wrote a book on the personality factors she considered to underlie ulcers, colitis, asthma, and coronary heart disease (Dunbar, 1938). The theory that certain "psychosomatic" illnesses exist (e.g., the concept of "the ulcer personality") was subsequently rejected by many. It would be more accurate to say that this notion was broadened into the idea that behavioral and psychosocial factors probably have some importance in most medical conditions. In a later era, Drotar (1981) summarized some of the new approaches taken by "pediatric psychologists" in studying and helping children with chronic medical illness.

1.01.10.10 Anorexia Nervosa and Bulimia

In the 1960s and 1970s there seemed to be a marked increase in the prevalence of self-starvation and other eating disorders, particularly among young women from socially advantaged backgrounds in the United States and Europe. Scholars unearthed a case history of this kind of anorexia nervosa dating back to the seventeenth century, written by Richard Morton (1637–1698), physician to King James II. An eating disorder known as bulimia nervosa also was characterized during the 1970s, consisting of "binging" episodes in which large quantities of food were eaten, followed by "purging" or self-induced vomiting (Russell, 1979). The abuse of laxatives is also sometimes a part of bulimia.

1.01.10.11 Borderline Personality Disorder

In former times the phrase "borderline" referred to an individual with difficulties resembling schizophrenia but not being quite severe enough to deserve that label. Subsequently, borderline came to refer to a personality trait in which an individual had an impaired sense of identity and problems with the boundaries of self. Suicide threats were observed to be common in this condition. Psychoanalyst Otto Kernberg (1967, 1975) was one of the first influential writers on what he called "borderline personality organization." In the *Diagnostic and statistical manual of mental disorders* (American Psychiatric Association, 1980), the condition was labeled as borderline personality disorder (BPD).

1.01.10.12 Antisocial Behavior

An early example of basic research on aggressive behavior was that of John Dollard (1900–1980), Leonard Doob, Neal E. Miller, O. H. Mowrer (1907–1982), and R. R. Sears, reported in their 1939 book, *Frustration and aggression*.

On the clinical side, psychiatrist Hervey M. Cleckley published in 1941 an influential book, *The mask of sanity*, trying to define the characteristics of a psychopath. Cleckley felt that one could not reliably diagnose this condition on the basis of an interview because many psychopaths were so skilled at deception. Instead, one had to look at the way the person behaved in life, disregarding rules and exploiting other people.

A classic study on the stability of antisocial behavior from childhood to adulthood was published in 1966 by sociologist Lee Robins under the title, *Deviant children grown up*. She followed up, approximately 30 years later, a cohort of children originally seen for diagnostic purposes in a St. Louis child guidance clinic. A relatively crude classification was made on the basis of the presenting complaint, so that there were those who presented with antisocial behavior, and others. Also included in the study were subjects who had been children in St. Louis at the same time but were not referred to such a clinic. The follow-up was exceedingly thorough and involved a personal interview using standardized questions and a coding system that later served as a model for the Diagnostic Interview Schedule (DIS) to be used with *DSM-III*. Also, data from the subjects' military service records, Social Security earnings records, the county sherriff's office, and many other such databases were accessed. The main finding of the study was that antisocial children tend to become antisocial adults, many of them qualifying for classification as psychopaths (*DSM* Antisocial Personality Disorder) in adult life.

The stability of aggressive behavior from childhood to adulthood was confirmed in a prospective longitudinal study using peer-ratings of third-grade school children as a means of selecting the subjects. This was carried out by psychologist Leonard Eron and his colleagues (e.g., Huesmann et al., 1984).

1.01.10.13 Substance Abuse

Some drugs have a dual history of being introduced as a treatment or as a means of studying psychopathology, and then later to become abused substances. Such was the fate of cocaine, which Freud to his lasting regret at one time recommended to a friend as a helpful treatment. The friend became addicted and died from an overdose (Gay, 1988). A more recent example would be lysergic acid diethylamide (LSD), discovered in 1943 and once considered promising as a method of gaining insight into the nature of psychotic hallucinations. As Shorter (1997, p. 265) comments, LSD had "no clinical payoffs" and became a "street drug of abuse."

1.01.10.14 Sexual Dysfunction

In 1966, gynecologist W. H. Masters and psychologist Virginia Johnson published their book on human sexual experience, including reports of direct observation of sexual activity and physiological recordings during such events. Their work elucidated aspects of normal sexual response that had previously not been so well understood, for example, female vaginal lubrication during sexual arousal. It also provided new understanding and treatment recommendations for various types of sexual dysfunction, for example, dyspareunia (Masters & Johnson, 1966).

1.01.10.15 Mental Retardation

H. H. Goddard (1912) wrote a book about a family he chose to call the Kallikaks (combining the Greek words for good and bad). The pedigrees of two branches of the family of "Martin Kallikak" were presented, one with considerable mental retardation among its members, the other without this pattern. The book has been derided as an example of a lack of rigor in its attempted analysis of genetic patterns ("look and say Mendelism"). The charge was even made that some of the published photographs were altered to reveal their physical stigmata more clearly. It is also possible that the changes in the photographs were made, in all innocence, to correct the fuzziness of some of the pictures (Fancher, 1987).

Eugenic approaches to mental retardation, considered anathema today, were widely endorsed by professionals and the public earlier in this century. In a famous case, Buck v. Bell (1927), the US Supreme Court even endorsed the use of involuntary sterilization for this purpose. Oliver Wendell Holmes, the chief justice, was quoted as saying "three generations of imbeciles is enough."

Torsten Sjogren, a Swedish geneticist, established in 1931 that Tay Sachs disease is a Mendelian recessive condition (Sjogren, 1931). This disorder involves blindness, deteriorating mental functioning, and death commonly by three years of age. The British geneticist Lionel S. Penrose (1898–1972) was responsible for summarizing for his readers material on both biometric and molecular genetics as related to mental retardation (Penrose, 1951). Philip Levine (1900–1987) and Alexander S. Wiener (1907–1976) elucidated the mechanisms by which Rh factor incompatibilities in the blood between the mother and fetus can produce fetal death or mental retardation.

Psychologist Edgar A. Doll and his colleagues published research implicating birth injuries as being responsible for a certain proportion of the cases of mental retardation (Doll, Phelps, & Melcher, 1932).

In the 1930s, psychologist Harold M. Skeels (1901–1970) and his colleagues published the well-known "Iowa studies" in which infants from an orphanage experienced large positive changes in IQ after being transferred to an environment where they received more attention and stimulation. These were not controlled studies and elicited strong criticism from Terman and others in the hereditarian camp. Nevertheless, the studies were the beginning of a long tradition of research on the effects of early experience on intellectual development and retardation. Skeels later published long-term follow-up data on some of these children, which indicated that they tended to retain their intellectual gains in adult life.

Psychiatrist George Tarjan (1912–1991) and his colleagues at Pacific State Hospital in California, an institution for persons with mental retardation, collaborated with colleagues at University of California at Los Angeles and the University of California at Riverside to develop a federally funded research program on the social epidemiology of mental retardation. This work continues today well after the end of Tarjan's life.

In support of the role of experience in long-term intellectual development, British psychologists A. D. B. Clarke and Ann M. Clarke published some interesting findings in their 1958 book. Their work suggested that individuals from highly deprived backgrounds may be quite retarded in their early development but have the potential for considerable gradual improvement when placed in more favorable surroundings (Clarke & Clarke, 1958).

The publication of psychologist Norman R. Ellis's 1963 edited book, *Handbook of mental deficiency: Psychological theory and research*, catalyzed the development of new research in the behavioral sciences relevant to mental retardation. In 1968, Ellis organized the annual Gatlinburg Conference as a meeting place, mainly for research psychologists interested in mental retardation and developmental disabilities.

In 1959, physician Jerome Lejeune (1926–1994) reported his discovery that Down's syndrome was generally characterized by an extra copy of chromosome number 21 (Hecht, 1994). This was but the first of a long list of chromosomal anomalies to be reported over the next few years, many of which were associated with mental retardation.

A team of researchers at the University of Washington in Seattle established an association between a particular syndrome of malformation in the offspring of alcoholic women (Claren & Smith, 1978; Jones, Smith, Ulleland, & Streissguth, 1973). The syndrome, which was dose-related in severity, is known as fetal alcohol syndrome (FAS) or (when milder) fetal alcohol effect (FAE). FAS is associated with mental retardation and other learning problems as well as behavior disorders in childhood and later life.

Anthropologist Robert Edgerton in 1967 published an ethnographic study of a cohort of persons released after institutionalization for diagnosed mental retardation. Virtually none of them accepted such a self-label. Their method of coping with the stigma of this diagnosis involved drawing around themselves a protective "cloak of competence" and enlisting what Edgerton called "benefactors" in the community to help them do so.

The book, *The bell curve*, written by Herrnstein and Murray (1994), repeated Jensen's arguments concerning the genetic basis of racial differences in intelligence test scores. Once more, the wrath of both professionals and journalists was elicited. The general consensus in the field is that such conclusions are not warranted by available data.

1.01.10.16 Autism

The child psychiatrist Leo Kanner (1894–1981) was the first to describe a syndrome he originally called infantile autism, characterized by extreme social withdrawal and impairments in cognition and language. His basis for doing so was a series of 11 case histories of children he had seen in the Harriet Lane Home of Johns Hopkins University Medical School, where he was a professor (Kanner, 1943). Kanner and Eisenberg (1956) followed up, nine years later, 63 children who had been diagnosed as having early infantile autism. They found the child's language at age five to have great prognostic significance. Those who had no language at age five rarely improved, while those who were able to speak at age five did much better at outcome.

Psychologist Bruno Bettelheim (1903–1990) viewed infantile autism in psychoanalytic terms and tended to "blame" the parents for causing it, even comparing the autistic child's early experiences to those of concentration camp victims (e.g., Bettelheim, 1950, 1967). Such views are part of the explanation for his advocacy of separating these children from their parents in order to facilitate their treatment.

1.01.10.17 "Minimal Brain Dysfunction" (MBD)/ADHD

Neurologist Alfred A. Strauss (1897–1957), who was associated with a residential treatment center for children at Northville, Michigan, carried out research in the 1930s and 1940s on what was originally labeled as endogenous vs. exogenous mental retardation. The term "endogenous" referred to children who functioned at a delayed level but appeared to be physically normal, whereas "exogenous" meant that there was some kind of evidence suggesting neurological dysfunction. Subsequently, the emphasis changed toward children without mental retardation but who had subtle difficulties that were presumed to be of neurological origin. They were first referred to as "brain injured" (Strauss & Lehtinen, 1947), but later were said to suffer from minimal brain dysfunction (MBD).

In 1957 psychiatrist Max Laufer and pediatrician Eric Denhoff redefined a part of what had been considered to be MBD as a hyperkinetic behavior syndrome in children (Laufer & Denhoff, 1957). They advocated the treatment of this behavior disorder with amphetamines. An early description of such behavior problems by British physician G. F. Still (1902) was soon rediscovered. Psychologist Virginia Douglas at McGill University in Canada was soon to argue convincingly in the 1960s and 1970s that the main problem for these children was one of inattention rather than hyperactivity. Thus, it came to be that the current terminology for this common childhood problem is attention deficit hyperactivity disorder (ADHD) (Barkley, 1990).

1.01.10.18 Learning Disabilities

In England, physician James Hinshelwood (1900) published a case of a child with what he called "congenital word blindness," interpreted as a developmental variation of the adult neurological disorder of alexia. Psychologist Edmund B. Huey (1908) wrote a classic monograph on the psychology and pedagogy of reading (reprinted by the MIT Press). Samuel T. Orton (1928, 1937), a psychiatrist and neuropathologist at the University of Iowa, reviewed similar cases he had seen of children with "specific reading disability," which he hypothesized to result from "strephosymbolia," a lack of proper lateralization in the brain that leads the child to confuse mirror images such as b and d with each other. Orton also noted the association of reading, writing, and speech and language problems in children.

Educator Samual A. Kirk (1904–1996) was the person who coined the term "learning disabilities" which served as a rallying point for parent groups as well as for researchers and practitioners since the 1960s. He was also among the first to do experimental research in special education (Kirk, 1952).

1.01.10.19 Child Abuse and Neglect

The problem of abuse and neglect of children, including sexual abuse and psychological abuse, is ancient. It had long been the special concern of child welfare social workers, who are responsible for attending to the needs of family preservation and foster home placement when abuse is detected. The "battered baby" syndrome was rediscovered by the medical community in the early 1960s, and many state laws were soon passed requiring professionals to report suspected abuse. Eventually other professionals such as psychologists and psychiatrists were drawn into professional activities related to abuse. One influential and scholarly book in this area was that written by sociologist David Finkelhor (1984).

1.01.11 RESEARCH ON ASSESSMENT IN THE TWENTIETH CENTURY

1.01.11.1 Intelligence

James McKeen Cattell was a key figure in the early development of psychological assessment in the twentieth century. A disciple of Francis Galton, he originally tried to promote brief measures of basic psychological processes: sensory, motor, reaction time, and the like. His project to demonstrate the utility of such measures in college admissions at Columbia University failed, however.

In trying to demonstrate the presence of a general factor (g) underlying academic and intellectual abilities, Charles E. Spearman (1863–1945) originated the mathematical procedure that came to be known as factor analysis (Spearman, 1904). Godfrey Thomson (1881–1995) effectively critiqued Spearman's mathematical procedures by showing that one could obtain similar evidence for a *g* factor based on a matrix of random numbers (Thomson, 1919). Although Spearman's particular approach to factor analysis was soon superseded by others, his concept of *g* remains somewhat influential.

The psychologist Alfred Binet (1857–1911) and his psychiatrist colleague Theodore Simon (1873–1961) developed an intelligence test (Binet & Simon, 1905) that successfully distinguished children with mental retardation from others whose mental development was normal. They did so by using tasks that resembled what children actually do in school, rather than simple sensory-motor processes. Binet and Simon originated the practice of representing a particular test score in terms of "mental age."

Psychologist Henry H. Goddard (1866–1957), who had been appointed research director at the Vineland Training School in New Jersey in 1906, learned of the Binet–Simon test on a European trip and by 1908 had it translated into English (Doll, 1988). He readily confirmed the ability of the test to discriminate the performance of persons with mental retardation from that of others.

The use of intelligence tests on a broad scale first began with the mass testing of military recruits upon entry of the United States into World War I. Robert M. Yerkes (1876–1956) and his colleagues developed the Army Alpha and Beta tests for this purpose and tested 1.7 million men. After the war, the industrial and educational uses of mass testing were developed, and testing became big business. Reasoning that mental age scores would not be useful for evaluating adults, Yerkes and other colleagues had earlier developed a "point scale" in which each item passed was simply given unit weight (Yerkes, Hardwick, & Bridges, 1915).

Lewis M. Terman (1877–1956), a psychology professor at Stanford University, developed a more standardized version of the Binet–Simon test and collected normative data with it. Terman's revision, known as the Stanford–Binet (Terman, 1916), made use of ratio IQ scores derived by dividing mental age by chronological age and multiplying by 100. The use of such a mental ratio score had been suggested by German psychologist Wilhelm Stern in 1912 (English translation, Stern, 1914). Terman soon began his "genetic studies of genius" (Terman, 1925) in which he followed up children with high Stanford–Binet scores. While few of them really turned out to be geniuses, these children indeed had high levels of academic, vocational, and social accomplishments later in life, validating the test as a predictor of such long-term outcomes (Terman & Oden, 1947).

In England, the mathematician R. A. Fisher and others working in agriculture went beyond descriptive statistics to develop the inferential approach, which eventually revolutionized both psychometric testing and experimental psychology.

The issue of test scores and race, class, and national origin became a major one after World War I, with IQ being interpreted by many as being more inherited than learned. The US government restricted immigration, and many states passed "eugenic" sterilization laws. In 1936 the Soviet government reacted in the opposite direction and put forth a decree banning the use of intelligence tests and all such "pedological" measures. One objection to the use of such measures was that they appeared to discriminate against individuals from proletarian backgrounds as compared to children of bourgeois parents and against members of various ethnic minorities within Soviet society. Oddly enough, among the researchers whose writings were banned at this time was psychologist Lev S. Vygotsky (1896–1934). Vygotsky had done research on the influence of cultural factors on intellectual development and spoke of the "zone of proximal development" as the area for educational focus with a child. When his complete writings were later published, Vygotsky became one of the more influential scientists arguing for environmental influences on intelligence (Kozulin, 1990).

It was ultimately psychologist David Wechsler (1896–1981) who developed the intelligence test into the standardized form best accepted in the United States and other countries in the West (Wechsler, 1939). Wechsler agreed with Yerkes and his associates that mental age scores were not useful for adults.

Psychologist Raymond B. Cattell (1905–1998) (not to be confused with J. McK. Cattell) attempted to develop "culture free" intelligence tests. His theory distinguished between "fluid" and "crystallized" ability. Fluid ability refers to one's performance on relatively novel tasks such as the Raven's Progessive Matrices, which has been found to fall off more steeply with age than crystallized ability, which refers to more practiced skills such as vocabulary.

In 1969, Arthur R. Jensen, a professor of psychology at Berkeley, published an article in the *Harvard Educational Review* that was widely reviled both by many of his professional colleagues and by the public (Jensen, 1969). Essentially, Jensen argued that efforts at compensatory education had failed. He implied that differences in IQ scores among different racial and ethnic groups might be hereditary in nature and difficult to change. Such interpretations of the data on IQ and race were not uncommon in an earlier generation of behavioral scientists but certainly had become politically unacceptable by the 1960s.

Psychologist L. J. Kamin (1974) counter-attacked and charged psychologists of the past such as Henry Goddard and L. M. Terman and present day psychologists such as Arthur Jensen and Cyril Burt with holding racist attitudes. Kamin also charged Cyril Burt with fabricating data in his widely cited research on intelligence in twins reared apart. Careful examination of Burt's data by others suggested that these charges had merit.

1.01.11.2 Categorical Diagnosis of Mental Disorders

Adolf Meyer (1866–1950) was probably the most influential American psychiatrist of his day. Trained in Switzerland in neuropathology and psychiatry, he began work in 1893 as a staff pathologist at Illinois Eastern Hospital in Kankakee (Winters, 1966). He was frustrated in his task from the beginning because of the poor quality of the case records. It was no use trying to establish the correlation of brain autopsy findings with clinical symptoms if the accounts of what the patient was like in life were missing or flawed. Thus, Meyer began a lifelong concern with developing adequate methods for the psychiatric assessment and classification of patients. He continued this preoccupation in subsequent positions at the Worcester Asylum (now Worcester State Hospital) in Massachusetts and ultimately as professor of psychiatry at the Johns Hopkins University. Meyer's standardized descriptions of psychiatric syndromes, or "reaction" patterns as he called them, became well known. Incidentally, Meyer was also interested in therapeutic methods and began applying psychoanalysis within a New York state hospital as early as 1902.

After World War II, many professionals were critical of psychiatric categorical diagnoses because of their low reliability (e.g., Stengel, 1959). Shorter (1997, p. 297) refers to Stengel's article as "the opening cannon in the campaign to revise diagnosis." The group of psychiatrists at Washington University in St. Louis were particularly influential in advocating a more rigorous approach to diagnosis. John P. Feighner, the chief psychiatric resident at Washington University at one point, published the standardized research criteria in use in the Department of Psychiatry there (Feighner et al., 1972). These "Feighner criteria" rapidly came into use by others, especially research psychiatrists and psychologists. An intermediate step between the use of Feighner criteria and the *DSM-III* was constituted by the "research diagnostic criteria" promulgated by Spitzer, Endicott, and Gibbon (1978).

British child psychiatrist Michael Rutter, in his Isle of Wight studies in 1975, demonstrated the applicability of rigorous epidemiological methods in this field. As part of this work he devised standardized psychiatric interviews for parents and children and standardized teacher rating forms. He showed a concern with assessing the reliability and validity of these methods and also was among the first to advocate the use of a multiaxial classification system.

Psychologist David Rosenhan published a striking study using "pseudopatients" who gained admission to mental hospitals on the basis of faked auditory hallucinations (Rosenhan, 1973). They began to act in their usual way, denying any such symptoms, immediately upon admission to a mental hospital. It typically took the hospital a matter of weeks to discharge them, and even so the hospital psychiatric staff tended to render discharge diagnoses such as "schizophrenia in remission." For many, this brought the credibility of the traditional diagnostic system even more into question.

A task force of the American Psychiatric Association soon began work on what was to become in 1980 a thorough revision of its *Diagnostic and statistical manual.* One issue that emerged in the discussion was the question of whether homosexuality should be included in the revised manual as a "mental disorder." Gay advocacy groups were unhappy that homosexuality had been included in the DSM-II in 1968. There was evidence, for example, from the research of psychologist Evelyn Hooker (1907–1996) (e.g., Hooker, 1969), that homosexuality was compatible with the absence of symptoms of any disorder (unless one wished to regard sexual orientation itself as a symptom). To make a long story short, the *DSM-III* taskforce decided to delete homosexuality from the manual, and the American Psychiatric Association affirmed this in an unprecedented referendum in 1974.

Psychologist Theodore Millon lobbied in the *DSM-III* taskforce to include a separate axis for personality disorders, a suggestion that it ultimately implemented, though it did not define many of the personality disorders precisely as Millon may have wished.

According to Reisman (1991), the original *Diagnostic and statistical manual* (*DSM*) of the American Psychiatric Association in 1952 listed 60 disorders. The *DSM-II* in 1968 listed 145, and *DSM-III* in 1980 listed 230 of them. It remains to be seen how well justified this inflation of the number of categories was. It is certainly easier to diagnose the broader categories reliably than it is the narrower ones. Considerable information has been collected on the reliability of the newer *DSM* categories, and they do seem to be more reliable than their counterparts in the earlier versions of the *DSM*. Although the validity of categories of descriptive psychopathology is a perennial issue, it is clear that they can be made more reliable by standardizing the questions asked, the way the answers are coded, as well as the way the information is combined.

1.01.11.3 Interviews

Informal interviewing surely goes back to the dawn of medicine as well as that of philosophy. A Socratic dialogue is after all a type of interview. This is no doubt the most common "assessment" procedure used by mental health professionals and is often the only one used. It was only recently, however, that standardized interviews began to be developed to increase the reliability and validity of clinical judgment. Psychiatrist Max Hamilton (1960) published a standardized, interview-based rating scale for depression that is still in common use. Sociologist Lee Robins and her colleagues developed the Diagnostic Interview Schedule (DIS), to be used by lay interviewers, to generate *DSM-III* diagnoses. The corresponding standardized interview for use by clinicians is called the Structured Clinical Interview for Diagnosis (SCID). These procedures have unquestionably increased the reliability of diagnosis for research purposes, though the question of validity remains a perennial one.

1.01.11.4 Self-report Measures of Personality and Psychopathology

One of the first measures of this kind was the Personal Data Sheet developed by psychologist Robert S. Woodworth (1869–1962), a professor at Columbia University, for military use during World War I (Woodworth, 1917). The items were rationally selected, mainly from psychiatric textbooks, and not initially subjected to empirical validation procedures. Thus, the content of this questionnaire was based on descriptive psychopathology, as developed over the entire preceding century.

The Minnesota Multiphasic Personality Inventory (MMPI), developed by psychologist Starke R. Hathaway (1903–1984) and psychiatrist J. C. McKinley in the 1930s and 1940s (e.g., Hathaway & McKinley, 1940, 1943), was innovative in several ways. For one thing, the clinical scales of the MMPI were empirically validated using psychiatric patients with particular diagnoses as criterion groups and visitors to the University of Minnesota hospitals as a control group. Also, the MMPI contained several innovative "validity scales" that could be used to identify subjects who might be too confused to give reliable answers, who could not read well enough to use the test properly, or who were trying to present themselves to the examiner in an unduly favorable or unfavorable light. The MMPI quickly became one of the most frequently used assessment procedures of its kind. The thousands of empirical studies conducted with it ultimately gave it even more credibility than it acquired from its initial method of validation.

Psychiatrist A. T. Beck published the first article describing the Beck Depression Inventory (BDI) in 1961 (A. T. Beck, Ward, & Mendelsohn, 1961). This theoretically based scale came into extremely wide use and has recently been revised.

Exploiting the conceptual distinction between "state" and "trait" measures, psychologist Charles Spielberger and his colleagues devised and validated separate pencil and paper questionnaires to measure these two different types of anxiety (Spielberger, Gorsuch, & Lushene, 1970). Part of the innovation in these questionnaires is simply in the way the items are worded. For example, one might ask "do you have butterflies in your stomach at this moment?" (state) vs. "do you frequently get butterflies in your stomach?" (trait). A state anxiety measure might be more useful in assessing a person's response to having blood drawn in a medical clinic while a trait measure would be a better choice in assessing neurotic anxiety.

1.01.11.5 Rating Scales

In 1962, psychologist Lester Luborsky published an article on his Mental Health/Sickness Rating Scale. This work was incorporated by Endicott and Spitzer (a psychologist–psychiatrist team) in their Global Assessment Scale (GAS), which was developed in 1976. In this scale, a score of 0 would be given if the patient had died by suicide and a score of 100 if the individual showed superior functioning in every area of life. The current *DSM* uses a further revised version of this called a Global Assessment of Functioning (GAF) scale, which has been shown to be quite reliable.

Psychologist C. Keith Conners developed a teacher rating scale in the 1950s primarily for use in research on the efficacy of stimulant medications with behavior disordered children. This scale soon came into wide use for diagnostic purposes as well as being commonly employed as a treatment outcome measure. It continues to be in use today. Conners also developed parallel scales for use with parents and for reports by the young persons themselves.

Psychologist Thomas Achenbach and his colleagues developed and validated an integrated set of rating scales for assessing child psychopathology, including the Child Behavior Checklist (for parent ratings), a Teacher Report Form, and a Youth Self-report scale. Many of the items of these scales are parallel, and thus, cross informant syndromes can be characterized by combining the scores from all three types of informants. The development of a sizable database using these scales provided a new perspective on the classification of child psychopathology as well as its measurement (Achenbach & Edelbrock, 1978). This approach has become widely influential, especially in the research community concerned with child and adolescent psychopathology.

1.01.11.6 Infant Assessment

Standardized procedures were soon developed for testing infants which superficially resembled those of intelligence tests for older children. Arnold L. Gesell (1880–1961), who was both a psychologist and a physician, developed one of the earliest tests of this kind, with 195 items for infants ranging in age from 3 to 30 months and others for babies up to age 60 months (Gesell, 1929). However, scores on these infant tests proved to have very little long-term stability and thus were mainly of use for research purposes.

The most widely used method for assessing infants, developed by psychologist Nancy Bayley (1899–1994), is known as the Bayley Scales of Infant Development (BSID) (Bayley, 1955, 1969).

1.01.11.7 The Rorschach Test

Swiss psychiatrist Herman Rorschach (1884–1922) published the manual for his famous inkblot test in 1921, the year before his death, and did not have a chance to follow up on his initial work. He had tried out the procedure with patients suffering from various types of psychopathology as well as normal controls. Rorschach originally had a plan of using 15 blots but could only get the publisher to print 10 blots, and the blots unexpectedly were printed including varying shades of gray. Nevertheless, interpreters managed to capitalize on every serendipitous development of this kind. Many of Rorschach's original interpretive hypotheses are still considered valid by many present users of the procedure, most of whom are psychologists of a psychodynamic persuasion.

Samuel J. Beck, who did his Ph.D. dissertation at Columbia on the Rorschach, went on to try to objectify the scoring methods most frequently selected for response and collected norms on the frequency of different perceptions of these areas. It was then possible to score "form quality" (perceptual accuracy, thought to be lowered in schizophrenia) on the basis of these norms (e.g., S. J. Beck, 1937, 1944).

In 1961, psychologist Wayne Holtzman and his colleagues published a new inkblot test that had a more standardized procedure for administration than the Rorschach test. The Holtzman test had clearer rules for scoring and provided norms for various patient groups and for the general population (Holtzman Thorpe, Swartz, & Herron, 1961). Despite its superior psychometric properties, this test was not frequently used in clinical evaluation. It was considered too lengthy and cumbersome.

1.01.11.8 Human Figure Drawings

During the 1920s, psychologist Florence L. Goodenough (1886–1959) at the University of Minnesota developed and validated a standardized way of scoring a child's drawing of a man (Goodenough, 1926). This had some use as a brief cognitive screening procedure, for example, with kindergarten children and in cross-cultural research, but is not often considered to be of much practical clinical value.

In 1948, psychologist Karen A. Machover (1902–1996) wrote a book on the use of human figure drawings as a brief projective test. Unlike Goodenough, Machover did not formally validate her procedure but simply shared with readers her speculations about how such drawings might be interpreted. Human figure drawings of this kind have come into very common use, particularly in the assessment of children. Psychologists Loren and Jean Chapman (1967, 1969) demonstrated that people trying to interpret such human figure drawings are subject to the phenomenon of "illusory correlation" in which they read their biases into what they see. For example, one might judge drawings focusing on the eyes to indicate paranoid trends, when no empirical association between drawing eyes in this way and paranoia exists.

1.01.11.9 Thematic Assessment Methods

In 1935, Christiana Morgan and Henry A. Murray (1893–1988) reported a new procedure, the Thematic Apperception Test, in which the individual was asked to look at reproductions of paintings or photographs and to tell a story about the scene portrayed in each picture. In research in the Harvard Psychological Clinic, the diagnostic team concluded that this was the best procedure in their battery for identifying individuals who were seriously self-deceived concerning their motives (Murray et al., 1938).

1.01.11.10 Sociometric Measures

The "sociogram" based on people's choices of others with whom they would like to share activities or others whom they would like to avoid was developed by psychiatrist Jacob Levy Moreno (1892–1974). Moreno introduced sociometric techniques in a rather unusual book, *Who shall survive*? (Moreno, 1952).

1.01.11.11 Assessment of Social Competence

Psychologist Edgar A. Doll (1889–1969) once served as director of research at the Vineland School and maintained a lifelong interest in the problem of mental retardation. In 1935, Doll published the Vineland Social Maturity Scale, which assessed what an individual characteristically was capable of doing in everyday life, as judged by a knowledgable informant. Such information soon became a part of the way mental retardation was defined.

1.01.11.12 Assessment of Marital and Family Relationships

Psychologist Richard Abidin developed an innovative measure called the Parenting Stress Index (PSI) (Loyd & Abidin, 1985). This work recognized the increasing realization in the field that not only parents have effects on their children. Children's behavior also affects their parents as well, and a difficult child can be a major source of parental stress.

1.01.11.13 Neuropsychological Measures

As neurology developed increasing sophistication, there grew a need for psychological assessment procedures sensitive to the effects of neurological lesions. An early example of such a technique and one that continues to be used frequently is the Bender Gestalt Test, developed by child psychiatrist Lauretta Bender (1938). She used designs taken from the work of Gestalt psychologist Max Wertheimer. This test requires the copying of these nine geometric designs with pencil and paper. A second widely used neuropsychological test, the Visual Retention Test (Benton, 1945), involves memory as well as drawing skills. The Halstead–Reitan battery, which developed out of the work of Ward Halstead (1947), involved a combination of many different subtests and soon came into common use.

A neuropsychologist who became renowned in the field was the Soviet scientist, Alexander R. Luria (1902–1977) who often tended to use qualitative assessment techniques rather than standardized quantitative techniques.

1.01.11.14 Biological Assessment Procedures

A specific test for the presence of syphilitic antibodies in the blood was developed by August von Wasserman in 1906. About 90% of the patients with general paralysis of the insane (GPI) tested positive, confirming the syphilitic etiology of this disease.

The German neuropsychiatrist Hans Berger, after 15 years of trying, succeeded in 1929 in recording electrical activity from the intact human scalp. Thus, the electroencephalograph (EEG) was born. By 1934 the first such recording in the United States had occurred, at Harvard Medical School, and by 1939 the EEG was in general use. It proved to be especially valuable in the diagnosis of seizure disorders and presumably in distinguishing hysterically-based "epilepsy" from the neurologically-based kind.

Physician Murray L. Barr developed a cytogenetic assessment procedure which an individual's cells could be examined under a microscope to determine the number of X chromosomes present in each one. This would ordinarily indicate simply whether the person was male or female (since males have only one X chromosome, while females have two). There are rarer cases of females with only a single X (Turner's syndrome), males with more than one X (Kleinfelter's syndrome), or other persons with multiple supernumerary X chromosomes. Each of these syndromes has its own profile of cognitive and behavioral characteristics (e.g., Money, 1963).

Physician Robert Guthrie (1916–1995) developed a blood test for detecting phenylketonuria (PKU) in newborn babies, making possible a program of early dietary treatment to prevent mental retardation and associated behavior disorders in this Mendelian recessive condition.

Geneticist Herbert A. Lubs (1969) published his findings on a "marker X chromosome," later to be implicated in the discovery of the fragile-X syndrome, one of the most common genetic causes of mental retardation and various associated disabilities (de la Cruz, 1985). It is of interest that fragile-X is a disease that actually does become worse in successive generations as it is passed on because of the increasing number of trinucleotide repeat mutations on the X chromosome. Perhaps this provides at least limited vindication of the nineteenth century psychiatrist B. A. Morel's concept of degeneration.

Endocrinologist Rosalyn S. Yalow developed techniques for the radioimmunoassay of peptide hormones. Among other uses, these procedures allowed the diagnosis and treatment of thyroid disease in newborn infants before any signs of cretinism appeared. She won a Nobel prize for this work in 1977.

New imaging technology is one of the fastest growing areas of assessment. Computerized tomography (CT), for example, is a noninvasive procedure that enables one to view the brain in cross-section and to visualize neuronal structures there as well as tumors, areas of hemorrhage, and so on. Johnstone et al. (1976) used CT relatively soon after it became available. They were able to document that persons with chronic schizophrenia tended to have enlarged cerebral ventricles. They also examined the relation of these to degree of impairment. Besides CT, the new techniques include magnetic resonance imaging (MRI), the positron emission technique (PET) for observing brain metabolic activities, and the study of regional cerebral blood flow (RCBF)(Andreasen, 1988).

1.01.12 RESEARCH ON TREATMENT IN THE TWENTIETH CENTURY

1.01.12.1 The Boston School of Psychotherapy

According to historian Eugene Taylor, the first use of different psychotherapeutic methods began at the Massachusetts General Hospital between 1903 and 1905 and on Ward's Island in New York after 1905. This work, which had its origins in developments in France, was subsequently overshadowed in the United States by developments in psychoanalysis.

1.01.12.2 Psychoanalysis and its Derivatives

Psychoanalysis dominated the mental health scene for more than half of the century and remains of interest today. Psychoanalysis usually involves daily or at least relatively frequent meetings between the therapist and patient, with the patient lying on a couch and engaging in free association ("just say whatever comes to mind"). From the first, this approach to treating neurotic problems achieved good "consumer acceptance" and was sufficiently remunerative that its practitioners could make a living. Admittedly, psychoanalysis was not only a method of treatment but a theory of psychopathology and had important implications for assessment as well. Detailed coverage of all of this is well beyond the scope of this chapter, since there are hundreds of books on this topic.

Freud had broken with his colleague Josef Breuer before the turn of the twentieth century and henceforth referred to psychoanalysis as his own creation. He published books on the interpretation of dreams in 1900, on the psychopathology of everyday life in 1904, on jokes and their connection with the unconscious, on psychosexual development, on sexuality in general in 1905, and many more in later years.

Although psychoanalysis was considered mainly as a treatment for neurotic symptoms, some practitioners tried it with psychotic patients. For example, psychoanalyst Marguerite Sechehaye did psychotherapy with schizophrenic patients at the Burgholzli hospital in Zurich and described her experiences in a 1947 book.

Otto Rank (1884–1939) was encouraged by Freud to obtain a Ph.D. in psychology from the University of Vienna and was for many years a member of the inner circle of psychoanalysis.

Rank's book on the *Trauma of birth* was but the first public example of his departure from strict orthodoxy. Part of Rank's historical importance is as a liaison man between psychoanalysis, social work, and psychology. At a later point in his career, Rank taught at the School of Social Work in Philadelphia and supervised the clinical work of Jessie Taft (Robinson, 1962), who even became his biographer (Taft, 1958). Jessie Taft was, in turn, one of the early supervisors of clinical psychologist Carl Rogers (1902–1987). Thus, there are indirect links between psychoanalysis and Rogers' client-centered or person-centered therapy.

In Germany in 1936, obituaries recorded the death of Bertha Pappenheim (1859–1936), a woman who was only three years younger than Freud. She was a prominent social worker and was perhaps the only one to have her likeness featured on a postage stamp. It eventually became known, via Ernest Jones' biography of Freud, that Bertha Pappenheim was the actual name of the patient in Freud's case history of "Anna O." Breuer and Freud (1895) had credited this patient with inventing the "talking cure." Records subsequently unearthed have documented that Pappenheim was not cured by psychoanalysis after all. She subsequently spent time as a patient in a sanatorium. Incidentally, Pappenheim lived on to become an effective opponent of the use of psychoanalytic methods for children under her care.

The psychoanalyst Eric H. Erickson (1902–1994) was trained by Anna Freud. Although he always presented himself as orthodox in his theoretical beliefs, he nevertheless managed to produce some of the most original and influential contributions. For example, in *Childhood and society*, Erikson (1950) transformed Freud's psychosexual stages into psychosocial ones that many readers found more palatable and added something all of his own about the stages beyond childhood of identity development, intimacy, generativity, and integrity. His writings on identity development have in particular generated much research related to development and the types of psychopathology that attend failure at negotiating it. In terms of treatment activities, Erickson was an acknowledged master at psychoanalytic play therapy.

In the 1950s, psychoanalysts began to be concerned about the absence of research evidence concerning the efficacy of this type of therapy. In 1953 the American Psychoanalytic Association appointed a Central Fact Gathering Committee to address this question. According to Shorter (1997, p. 311; see also Knight, 1953), the discussions of this committee were inconclusive, its report was kept confidential, and the committee was disbanded in 1957.

One of the more recent variants of psychoanalysis is psychoanalytic self-therapy, developed by psychiatrist Heinz Kohut (1913–1981), an emigrant from Vienna to Chicago. Kohut is reported to have been particularly sympathetic to narcissistic and borderline patients, both of whom have notable difficulties with the "self."

With the emergence of managed care in the United States, it became more and more difficult for psychoanalysts and others doing long-term psychotherapy to be reimbursed for their services by third-party payers. Nevertheless, this type of therapy maintains its appeal to some who are willing to pay for it out of their own pockets. Historian Eugene Taylor (personal communication, January 13, 1998) ventured the following prediction:

> psychologists opposed to the medicalization of psychotherapy and the imposition of managed care and its required treatment protocols will split off and develop autonomously, possibly launching an entirely new generation of deregulated philosophical psychologists whose job is to teach people how to live, as psychotherapy itself becomes de-medicalized, de-pathologized, and more oriented toward education for transcendence.

1.01.12.3 Play Therapy for Children

Psychoanalytic methods were first adapted for use with children by Hermine von Hug-Hellmuth (1921). Children generally are not able to lie on a couch and free associate but prefer to play with toys, and thus what became play therapy was a necessary modification. The next developments in the psychoanalytic treatment of children were due to Melanie Klein (1882–1960), who was known for her boldness in making "deep" interpretations of children's play activities (e.g., Klein, 1932). The principal historic importance of Klein's work may lie in its role in the emergence of the "object relations" school of psychoanalysis that was especially influential in Britain. One of Klein's pupils was the pediatrician and psychoanalyst Donald Winnicott (1896–1971). His clinical concepts included the idea that "transitional objects" such as blankets and teddy bears are used by children as mother substitutes.

Child psychiatrist Frederick Allen (1890–1964) had obtained his M.D. from Johns Hopkins University after previously spending three years as a school psychologist. He became Director of the Philadelphia Child Guidance Center and contributed to the literature on play therapy. His work was influenced by that of Otto Rank.

1.01.12.4 Japanese Therapeutic Approaches

The Japanese psychiatrist Shoma Morita (1874–1952) developed psychotherapeutic procedures based on Zen Buddhism that were named after him. Although such approaches are commonly used in Japan, they have not become widely adopted elsewhere (Moriyama, 1991). There continue to be major language and cultural barriers between East and West in the mental health field, as has been illustrated every time this topic has come up in the present chapter.

1.01.12.5 Behavior Therapies

In 1924, psychologist Mary Cover Jones published a study of the case of "Peter" in which she showed how the child could be induced to overcome his fear of a furry animal by gradually increasing exposure to it in the presence of competing activities such as eating. Her supervisor in this research was psychologist John B. Watson, well known as one of the founders of behaviorism and a strong advocate of this theoretical position (Watson, 1913). At the time he worked with Jones, Watson had already been dismissed from his academic position at Johns Hopkins University and had entered the world of business, working for an advertising firm in New York. As Taylor remarks (personal communication, January 13, 1998), "behaviorism met Madison Avenue in the 1920s."

The name of William H. Burnham (1855–1941) is not well known, but Kazdin (1978) says that he was an important pioneer in the development of behavior therapy. Burnham (1924) published a book on *The normal mind: An introduction to mental hygiene and the hygiene of school education.*

Physiologist Edmund Jacobson published an influential procedure for systematically teaching people to relax (Jacobson, 1929). This procedure was later incorporated into behavioral approaches to desensitizing phobias and had many applications in behavioral medicine.

Another behavioral psychologist, Knight Dunlap (1875–1949), a professor at Johns Hopkins University, wrote a book about the use of "negative practice" in order to get rid of unwanted habits such as stuttering or thumb sucking (Dunlap, 1932). To get rid of such a habit, the individual was encouraged to perform it voluntarily over and over, until it became unpleasant to do so.

In 1935, physician Walter L. Voegtlin and his colleagues from the Shadel Sanitarium in Seattle proposed the use of aversion therapy to help alcoholics achieve abstinence. Voegtlin was a gastroenterologist. In the 1950s a follow-up study was published of over 4000 patients who had received this type of treatment. The results were promising though not definitive.

The "bell and pad" method for treating bedwetting in children (still in use today) was first popularized by Mowrer and Mowrer (1938). Their work was based on a method originally developed by pediatrician M. Pfaundler in 1904. A pad containing electronic moisture sensors was placed in the child's bed. If the child urinated during the night, a bell rang to awaken the child, who was then supposed to get up and use the bathroom to finish urinating. Repeated use of this procedure was observed to decrease and then eliminate enuresis, although relapses were a problem. The Mowrers explained how this procedure worked in terms of a Pavlovian conditioning process. Many years later, Azrin and Foxx (1974) developed and field tested a more elaborate version of the Mowrers' behavioral procedures that could even enable parents to toilet train their child in "less than a day."

One of the earliest practitioners to make a living doing behavior therapy was Andrew Salter (1914–1996), from New York City. He had done his undergraduate work at New York University but had no graduate or professional degree. Salter was influenced by Clark Hull's behaviorally oriented writings on hypnosis (Salter, 1941) as well as Pavlov's work on conditioned reflexes (Salter, 1949). His concept of "assertion training" influenced the work of Joseph Wolpe.

According to Kazdin (1978), the psychoanalyst Alexander Herzberg moved to London during World War II and developed a modification of psychoanalysis that he called "graduated tasks" (Herzberg, 1945). Kazdin notes that psychologist Hans J. Eysenck (1916–1997) worked with Herzberg in the 1940s and was influenced by him, thus providing part of the context for Eysenck's later emergence as an advocate for behavior therapy.

Eysenck received his Ph.D. at the University of London in 1940 and in his early days in the field viewed clinical psychology as involving only research and diagnostic testing, not treatment. That he left to psychiatry. Eysenck spent some time in the United States and changed his mind about the role of psychologists as therapists. In 1952 he published a review of the effects of traditional psychotherapy which had the effect of a bombshell on psychotherapists who read it. He concluded that there was no evidence that psychotherapy was effective, because about two-thirds of the patients improved with or without therapy. Eysenck went on to pioneer the field of behavior therapy in Britain.

In 1958, psychiatrist Joseph Wolpe (1915–1997) published an influential book about the use of systematic desensitization for phobias (Wolpe, 1958), claiming a 90% success rate with this type of problem. In its most common variant, Wolpe's treatment involved teaching the patient to relax (using the procedure of Jacobson, 1929). Then the patient learned to associate this state of relaxation with a series of imagined scenes that were arranged in a hierarchy of increasing anxiety related to the object of the patient's phobia. The reasoning was that relaxation and anxiety are incompatible responses, and anxiety can be counter-conditioned in this way. Wolpe had received his M.D. from Witwatersrand in South Africa. In 1956 he came to California to spend his sabbatical at the Center for Advanced Study in the Behavioral Sciences, and it was there that he wrote his 1958 book. In 1962 he took a position at the University of Virginia and then moved to Temple University in 1965, where he spent most of his career.

Systematic desensitization soon attracted the attention of behavioral psychologists who began to carry out empirical research on this topic (e.g., Rachman, 1959; Lang & Lazovik, 1963). Controlled research was supportive of Wolpe's claim that such treatment was effective. However, critics of this research (e.g., Breger & McGaugh, 1965) pointed out, with good reason, the fact that this procedure did not seem to be based purely on behavioral principles. In particular, the role of the hierarchy of imagined scenes in the therapy seemed very cognitive. It is thus not a complete suprise that behavior therapy, as it developed, tended toward a greater and greater rapprochement with cognitive approaches.

1.01.12.6 Behavior Analysis

The basic research of psychologist B. F. Skinner (1904–1990) on the functional analysis of the behavior of rats and pigeons ultimately created a clinical specialty of "applied behavior analysis" that separated itself from other behavioral and certainly from cognitive-behavioral approaches to treatment. Skinner received his Ph.D. from Harvard in 1931 and published his first book in 1938. Psychologist Sidney W. Bijou, who was once a colleague of Skinner's at Indiana University, was one of the first to apply this type of behavior analysis to children and individuals with mental retardation or autism (Bijou, 1996). Others followed suit. For example, using a single subject design, Carl D. Williams (1959) illustrated the use of extinction procedures to eliminate tantrum behaviors in a child at bedtime. Such advice about how to deal with tantrums is now routine in behavioral parent-training programs.

Behavior analysis also proved to be useful with psychotic adult inpatients, as shown by Ayllon and Michael's (1959) paper on "the psychiatric nurse as a behavioral engineer." In 1962, psychologist Montrose Wolf went a step further to design a token economy at Ranier School, an institution for persons with retardation, in Washington. Ayllon and Azrin (1968) did the same thing for all types of state hospital patients. In the 1970s Wolf designed what came to be called the Achievement Place model, a type of family group home for youngsters considered to be at risk for delinquent outcomes. In an Achievement Place home, the live-in foster parents used behavioral principles to encourage various types of prosocial behavior, everything from cleaning up the bathroom, to doing homework, to avoiding aggressive outbursts.

By the 1960s, psychologist O. Ivar Lovaas and his colleagues at University of California at Los Angeles (e.g., Lovaas, 1987) had become engaged in intensive treatment efforts for children with autism, using what came to be called "discrete trial" operant conditioning methods. These efforts eventually attracted great interest and even generated demands from parent groups around the country for their local school systems to fund behavior analysts to work with their children.

Psychiatrist Stuart Agras and his colleagues (Agras, Barlow, & Chapin, 1974) published data showing that anorexia nervosa could be treated using behavior modification. In this and subsequent research it was shown that weight gain could be used as the basis for presentation of contingent reinforcement such as the opportunity to watch television or to see a boyfriend. Anorexic patients exposed to such contingencies gained weight, at least while the contingencies were in effect.

1.01.12.7 Cognitive Therapy

Although not labeled as such at the time, one can see in retrospect that the "personal construct" therapy and the related assessment procedures developed by George A. Kelly (1905–1966) could be considered to be an early variant of cognitive therapy (Kelly, 1955). Kelly thought of every person as a sort of scientist who developed particular ways of construing other people and events as a result of experience. He tried to assess these personal constructs and to help the person modify them so reality would be representated as accurately as possible. Another early example of cognitive therapy was the "rational emotive" treatment developed by psychologist Albert Ellis (1958), a prolific writer on this topic.

In terms of basic research, one could find few more effective advocates of the integration of cognitive and behavioral approaches than Albert Bandura, a professor of psychology at Stanford University (e.g., Bandura, 1971). Early in his career, Bandura published research showing the effects of modeling or imitation on children's behavior, including aggression. Subsequently, Bandura's research incorporated the influence of cognitive variables in such processes. For example, he found that children's behavior would only be influenced by models if they directed their attention toward what the model was doing and engaged in mental rehearsal in order to remember what they saw. Seeing the model rewarded for certain behaviors might not influence the child observer at the time, but would do so later if rewards became available to the child for demonstrating the behaviors later. Bandura's theory also emphasized the importance of "self-efficacy" or the belief that one would be able to accomplish some goal. Even though Bandura did not do treatment outcome research, his theory provided a framework for cognitive-behavior therapy.

Beck's model of cognitive therapy for depression has been influential (Beck, 1976). Among other strategies, it encourages patients to become aware of their negative "automatic thoughts" and to learn to dispute them.

One of the more promising new cognitive-behavioral treatments to emerge is the one developed by Marsha Linehan for borderline personality disorder (Linehan, 1993). It is called "dialectic behavior therapy" in that among other strategies it encourages patients to explore possible alternative explanations for certain distressing events they describe. This treatment has already been shown to be effective in reducing the number of self-destructive acts engaged in by patients with BPD.

1.01.12.8 The Interpersonal Approach

One of America's most innovative psychiatrists was Harry Stack Sullivan (1882–1949). He was especially known for his ability as a therapist working with young male schizophrenic inpatients (Sullivan, 1931). On occasion he encouraged such patients to become intoxicated on alcohol for 3–10 days in order to "open them up" to psychotherapy. Although influenced by psychoanalysis, he defined the field as being primarily concerned with interpersonal relationships (Sullivan, 1953). Sullivan coined the term "significant other" for the person to whom one is most attached and emphasized the importance of "consensual validation" not only by one's therapist but by peers or "chums." Thus, Sullivan thought it important to encourage close, confiding social relationships, especially among youth.

1.01.12.9 Research on Psychotherapy

From the time of Hippocrates down to World War II, one would be hard pressed to come up with many examples of controlled experimental research on psychotherapy or even on many of the medical approaches used to treat mental disorders. It is to the credit of psychologist Carl R. Rogers that he not only developed a new approach to therapy (Rogers, 1951), he also began to use some of the research methods he had learned about in graduate school to study both the process and the outcome of the therapy he had developed. Rogers tape-recorded therapy sessions, had them transcribed, and used coding procedures to try to quantify what was happening. He used control groups, for example, persons who were asked to stay on a waiting list for a time with the promise of receiving the same therapy at a later time. He and his colleagues used standardized measures such as Q-sorts to try to get at psychotherapy outcomes.

Of course, Rogers' therapy tapes and transcripts could also be used by those with theoretical views quite in contrast with his own. Psychologist Charles Truax (1966), for example, managed to demonstrate that Rogerian therapists were not as nondirective as might have been supposed. They were more likely to pay attention, and thus presumably to reinforce, some types of client statements than others.

Psychologist A. E. Bergin is known as both a researcher and as a reviewer of psychotherapy research. One theme he focused on was the fact that psychotherapy can harm as well as help (Bergin, 1966). Researchers and clinicians therefore need to be concerned about "side effects" as well as "effects."

One interesting twist on psychotherapy research presented data suggesting that the provision of psychotherapeutic services in an HMO setting could actually reduce patients' inappropriate utilization of nonpsychiatric medical care (Cummings & Follette, 1967). Such conclusions are rather self-serving to the mental health professions, of course, and more research is needed to confirm or disconfirm these findings.

Another innovation in psychotherapy research was the trend toward using meta-analysis, begun by Smith and Glass (1977). In a meta-analysis, the existing research literature is reviewed quantitatively rather than only in the traditional narrative fashion. Summary statistics such as "effect sizes" are calculated to pool the results of many separate investigations. The best known conclusion of the Smith and Glass (1977) review is simply that psychotherapy "works" as compared to various control conditions. Scholars criticized some authors of meta-analyses, for averaging in flawed studies with those that were well designed. In the familiar language of the computer world, this amounts to "garbage in, garbage out." Meta-analysis was soon applied to child therapy as well (Casey & Berman, 1985; Weisz, Weiss, Alicke, & Klotz, 1987), with much the same conclusions.

The neglected issue of the clinical significance of psychotherapy research was addressed by Jacobson and Revenstorf (1988). It is important in reading a research article in this domain to look not just for "statistically significant" findings, which in the case of studies with large samples may be trivial in size. Instead, one should ask whether the amount of change observed would actually be important to a therapy client. For example, did the client come into therapy with a particular psychological disorder and move into the normal range as a result of treatment?

1.01.12.10 The Psychotherapy Integration Movement

One problem for psychotherapy researchers was the sheer number of "new" therapies being developed. It was widely estimated in the 1980s that there were more than 130 different types of psychological therapies. There was thus a need to try to specify what these diverse approaches might have in common that could account for their effectiveness, as well as to work toward designing specific psychological therapies for specific disorders (Beutler, 1979). A classic work on this topic was psychiatrist Jerome Frank's (1973) book, *Persuasion and healing*. Frank examined not only different types of formal psychotherapy but also traditional religious or shamanistic approaches to healing (for an updated version of this book, see Frank & Frank, 1991).

1.01.12.11 Group Therapy

According to historian Eugene Taylor, group therapy began with the work of Joseph Hersey Pratt in Boston in 1904, with the Class Treatment of patients with tuberculosis. The term "group therapy" was used by psychiatrist J. L. Moreno in 1932. Moreno also developed a version of this approach known as psychodrama and in 1942 established a Psychodramatic Institute.

According to Shorter (1997), psychiatrist Joshua Bierer (1901–1984) introduced the use of group psychotherapy at Runwell Hospital, London, and then at other London hospitals. This eventually took the form of a Social Club, which elected its own members. Bierer founded a Social Psychiatry Center in Hampstead in 1946 and set up the first day hospital in England in 1948.

During the 1960s and 1970s, many new forms of group therapy were developed. Encounter groups became a fad (especially in California). Gestalt Therapy (not to be confused with Gestalt Psychology) and other approaches associated with the human potential movement or humanistic psychology thrived (Buhler, 1971).

1.01.12.12 Self-help Groups

The first and most famous of these groups, Alcoholics Anonymous (AA), was founded in 1935 by "Bill" Wilson and Bob Smith. To this day, persons who attend AA receive social support for their sobriety and turn their attention away from themselves toward helping others. AA was from the first a religiously oriented group, in a sense an offshoot of Frank Buchman's Oxford Group within charismatic, evangelical Christianity. Participants were thus encouraged to confess their allegiance to a "higher power" and their inability to overcome their difficulties without spiritual help. There is, however, a nonreligious alternative approach known as Rational Recovery. Treatment for substance abuse has been a rather separate enterprise from that of traditional psychiatry or psychology. Many of the therapists involved in such work are themselves former alcoholics, and the ideology of AA is strong among them.

1.01.12.13 Marriage and Family Therapy

One pioneer in the development of family therapy was psychologist John Elderkin Bell. According to Bell's own recall of events (personal communication, Iowa City, Iowa, early 1980s), he read in a journal in the 1950s some accounts of the work of colleagues in England who were seeing parents and children in therapy, which he mistakenly interpreted to mean seeing them together, or what is now labeled "conjointly." Thus he began doing this in his own practice, only to discover that he had evidently been the first in the world to do so. He published a book on this topic (Bell, 1961).

The field of marriage therapy grew up as one that was rather separate from the mainstream of either psychiatry or psychology. Many social workers had a special interest in this domain. Gradually, psychologists began to be interested in working with married couples. Jacobson (1989) was a pioneer in carrying out controlled research on behavioral marital therapy.

1.01.12.14 Community Interventions and Prevention

Within both psychiatry and psychology after World War II, some were interested in moving beyond interventions with individuals or small groups. Their desire was to go out into the neighborhoods and the cities to make larger changes. Psychologist Seymour B. Sarason is an example of such an individual. He and his colleagues set up a "psychoeducational clinic" at Yale, not as a place for the evaluation and treatment of individual school children but to exert an impact on the whole school system and on the city. When asked about his motives for such activities, Sarason mentioned the facts that he was Jewish, that he had polio as a child, and that as a youth he had been an adherent of Trotsky. Thus, he was disposed to view problems such as discrimination and economic injustice as community matters rather than simply individual ones.

In England, psychologist Jack Tizard (1919–1979) developed a sort of action research program to move adults from institutions for the retarded out into the community. He and his colleague Neil O'Connor (1917–1997) demonstrated that with suitable instruction these individuals could perform as well as other workers employed in competitive jobs (O'Connor & Tizard, 1956). In part because of research of this kind, deinstitutionalization eventually went from a trickle to a flood. The use of "job coaches" to help handicapped persons adapt to competitive employment is now widespread.

Psychologist George Albee (1968) engaged in extensive research on the cost effectiveness of mental health treatments and the number of personnel required to deliver them. He came to the conclusion that there were simply never going to be enough therapists to do the job. To him, prevention was the only answer. Albee (1970) proceeded to make himself very unpopular with his colleagues by speaking and writing about "the uncertain future of clinical psychology" (he thought the same thing about the future of psychiatry). He suggested on a later occasion that since its clientele mostly consists of relatively well-to-do persons, insurance coverage of psychotherapy might represent a subsidy to the rich from the poor (Albee, 1979).

1.01.12.15 Special Education

The physician and educator Maria Montessori (1870–1952) in Italy began her work by familiarizing herself deeply with the writings of J. M. G. Itard and Edouard Seguin and translating their works into Italian (Kraemer, 1988). She developed educational techniques for young children with delayed development living in the slums of Rome that used structured play-like tasks to engage children's natural attention. As things turned out, her approaches came into common use throughout the world in private preschools for well-to-do children with no particular developmental delays.

Psychiatrist Samuel T. Orton (1937) not only came up with neurological hypotheses about the origins of reading disability in children but also, together with colleagues in education such as Anna Gillingham and Bessie Stillman, devised remedial techniques involving intensive drills on analyzing words into sounds and learning letter–sound associations.

1.01.12.16 Rehabilitation

The psychologist Shepherd Ivory Franz (1874–1933) was presented with an honorary M.D. by medical colleagues at George Washington University for his research contributions. He was also made an honorary member of the American Psychiatric Association. He was an early advocate of rehabilitation for patients with neuropsychological disorders (e.g., speech training for aphasics) as well as for those with severe psychopathology.

Psychologist Marc Gold (1939–1982) was responsible for developing innovative approaches to teaching specific vocational skills to persons with mental handicaps. For example, Gold carried out a task analysis specifying the details of how to put bicycle coaster brakes together and showed that adults with mental retardation could learn to master it. For those of us who as school children had the experience of taking our bicycles apart and unsucessfully trying to put them back together, Gold's accomplishment is most impressive.

1.01.12.17 Drug Therapies

In terms of developing medications for treating mental disorders and understanding how they work, a crucial discovery was of the existence of neurotransmitters, the chemical substances that mediate the synaptic transmission of neural impulses. According to Shorter (1997), the first neurotransmitter to be discovered was acetylcholine by Otto Loewi in 1926. By the mid-1990s, over 40 neurotransmitters had been identified.

The use of neuroleptic drugs for psychotic patients beginning in the early 1950s greatly reduced the census of mental hospitals worldwide. Reisman (1991), for example, reports that the US mental hospital population went down from about 600 000 in the 1940s to about 150 000 in the 1970s. Shorter (1997, p. 248) gives some details of how this all happened, as follows: Henri Laborit, who was a surgeon in the French navy, in 1951 developed the drug chlorpromazine. The chemist who synthesized it was Paul Charpentier. At the time, Laborit was looking for something to potentiate anesthetics. He noted that surgical patients given this drug became rather indifferent to the world. By 1952 the drug began to be used in Paris (Delay, Deniker, & Harl, 1952) and by May 1953 had transformed the disturbed wards there. The same kind of thing then happened in Montreal, London, and Boston later the same year. Psychiatrist Heinz E. Lehmann, who introduced chlorpromazine to Montreal's Verdun Protestant Hospital in 1953 (Lehmann & Hannahan 1954; Lehmann, 1989), of the first to note that the drug had some side effects, some known as extrapyramidal or Parkinsonian, and others subsequently described by the term tardive dyskinesia. The Parkinsonian side effects were temporary and readily treatable with other medications, but this was not so for the dyskinesia, which were sometimes enduring ones.

With regard to the use of lithium in mania, Shorter (1997) gives the following account: in 1949, John Cade, the superintendent of a mental hospital in Australia, injected urine from manic patients and from controls into guinea pigs. The guinea pigs died. To try to make the urine more soluble, Cade mixed lithium with it. On a whim, he decided to inject lithium alone into the guinea pigs. To his surprise, instead of the usual frantic struggling, they lay there placidly. He tried injecting himself with lithium and then injected several mental patients, including 10 with mania, six with schizophrenia, and three with depression. Interestingly, the manic patients tended to improve. It so happened that the *Journal of the American Medical Association* reported in the same year two deaths associated with attempted lithium therapy, so the research community understandably became very cautious about the topic. Nevertheless, Mogens Schou, a Danish psychiatrist, confirmed Cade's results, using a double-blind crossover research design (Schou et al., 1954). It was not until 1970 that the Food and Drug Administration (FDA) in the United States approved the use of lithium, and then only in response to the threat of civil disobedience by an Oregon psychiatrist. Now lithium is in common use with manic patients, though with caution due to its known adverse side effects.

In 1970 psychiatrist Jules Angst (director of the Burgholzli Hospital in Zurich) and his colleagues demonstrated that lithium had similar prophylactic effects in unipolar and bipolar depressed individuals (Angst, Weis, Grof, Baastrup, & Schou, 1970). In other words, lithium has effects on depression as well as on mania.

According to Shorter (1997), meprobamate was first synthesized in 1950 by Bernie Ludwig, an organic chemist for Carter Products, with the consultation of physician Frank Berger. It was tested as an anti-anxiety drug and marketed by Wallace Labs in 1955 under the labels of "Miltown" and "Equanil." It became the first psychiatric drug to become a popular fad. The comic writer S. J. Perelman in 1957 published a book with the title, *The road to Miltown*. Comedian Milton Berle began referring to himself as "Miltown Berle."

Antidepressants were the next medications to emerge, in a somewhat circuitous way. Selikoff, Robitzek, and Ornstein (1952) published a report that iproniazid (Marsilid), a drug used in the treatment of tuberculosis, seemed to have some antidepressant properties. A related drug, named Tofranil, was then developed, tested, and marketed by Geigy (Kuhn, 1957). It was the first of the tricyclic antidepressants, for which 10 million prescriptions were written in 1980 in the United States alone. Swiss psychiatrist Roland Kuhn, who was responsible for this work, had undergone a training analysis under Ludwig Binswanger but later fell away from psychoanalysis (Shorter, 1997, pp. 258–262).

The advent of the benzodiazepines in the 1950s and 1960s is very much a commercially motivated story. Once more, I will depend on the account given by Shorter (pp. 316–319). In 1954, the Hoffman-LaRoche company wished to develop a sedative or anxiolytic drug that would be able to compete with Miltown and Equanil. The chemist involved was Leo Sternbach, and in 1955 he synthesized Librium. Then in 1959 he came up with diazepam (Valium), which was marketed in 1963. Even though these drugs were placed on Schedule IV by the FDA because of their addictive potential, benzodiazepines accounted for about half the prescriptions written for psychiatric office patients by 1990. Later in the 1990s the hottest drug in this group was alprazolam (Xanax). Shorter (1997, p. 320) describes panic disorder (for which this drug was often prescribed) as "the Upjohn illness."

By the 1970s, drug therapy had become quite standard for many types of psychopathology. In 1982, Rafael Osheroff, a 42-year-old physician who had been a patient at Chestnut Lodge in Maryland in 1979, sued the hospital for malpractice because he had been treated only by means of psychotherapy, not medications. The case was settled in his favor by an arbitration panel (Klerman, 1990). This case was noteworthy also because of the fact that Chestnut Lodge was so well known for the use of psychoanalytic treatment for psychosis.

In the 1950s and 1960s child psychiatrist Leon Eisenberg and psychologist C. Keith Conners began to carry out placebo controlled double-blind studies of the effects of stimulant medications on the behavior of disturbed children (e.g., Conners & Eisenberg, 1963). This research supported the efficacy of such medications. The clinical use of Ritalin (methylphenidate) with behavior problem children increased accordingly, at least in the United States. British and other European psychiatrists did not quickly follow suit, however. It is estimated that over two million children in the United States receive Ritalin at present (Shorter, 1997, p. 290).

In France, Lambert (1966) carried out research demonstrating the antimanic properties of valproic acid, which had originally been synthesized as an organic solvent in 1882 and in 1963 discovered, more or less by accident, to act as an anticonvulsant.

In the 1980s, it was discovered that clomipramine (Anafranil) caused improvement in the symptoms of obsessive-compulsive disorder (OCD) in children as well as adults (Rapaport, 1988). This disorder had previously proven quite refractory to treatment by any method, whether psychoanalytic, behavioral, or medical.

A syndrome involving motor and verbal tics, often including curse words, was described by the nineteenth century French physician Charcot and named after his colleague, Georges Gilles de la Tourette (1857–1904), who had been the first to report such a case. Tourette syndrome (as it has come to be called) was rediscovered by psychiatrist E. Shapiro, who found that the tics could be controlled by the drug haloperidol (Haldol), a neuroleptic (Shapiro et al., 1989).

At the time of writing, the biggest story in medications for psychopathology is that of Prozac (fluoxetine) and the general class of drugs of which it is a part, selective serotonin reuptake inhibitors (SSRIs). In the 1990s, Prozac was the drug most commonly prescribed by psychiatrists. By 1992 it was the second most frequently prescribed drug in the world. It is typically employed as an antidepressant. Once more we will depend on the account of Shorter (1997), who even included Prozac in the subtitle of his book on the history of psychiatry. The development of Prozac had its origins in work with 5-hydroxytryptophan (5HT), or serotonin, in the 1950s. In 1957 Bernard Brodie at the National Institutes of Health discovered that reserpine could greatly reduce bodily stores of 5HT. Alec Coppen at the Medical Research Council in Britain in 1963 showed that drugs which were serotonin equivalents could relieve depression. It was Ray Fuller, a senior pharmacologist at Eli Lilly in Indianapolis, who synthesized Prozac (Lilly 110140), working with biochemist David Wong. By 1980 the field testing of this new drug was underway, and in 1987 it was released to the public. Shorter notes that by 1993 almost half of the patient visits to psychiatrists were for mood disorders and typically included a prescription for Prozac.

1.01.12.18 Psychosurgery

The Portuguese physician Egas Moniz (1874–1955) and a surgical colleague developed techniques for lobotomy, in which neural fibers connecting part of the frontal lobes to the rest of the brain were severed (Moniz, 1937). He won a Nobel prize for his work, but this type of surgery is now rarely used as a treatment for psychopathology. In the peak year of 1949, over 5000 lobotomies were performed on mental patients. In the same year, however, controlled research found that the rate of favorable changes did not differ between patients with frontal topectomies and controls (Mettler & Columbi-Greystone Associates, 1949). The frequency of psychosurgery declined greatly thereafter.

1.01.12.19 Electroconvulsive Shock Treatment

Electroconvulsive shock treatment (ECT) was developed by psychiatrists Lucio Bini and Ugo Cerletti (1877–1963) in the 1930s. In initial experiments on dogs, placing the electrodes on the animal's mouth and anus, the investigators found that about half of the subjects died. Later they discovered that if the electrodes were placed on the dog's temples, electric current could be delivered safely (Shorter, 1997, pp. 218–219). Next, they tried the same thing on pigs at a Roman slaughterhouse. Finally, on April 18, 1938, they tried ECT on the first human patient (Cerletti & Bini, 1938). The clinical use of ECT spread rapidly to England, France, and the United States. This therapy is still in use, especially for the treatment of severe depression.

1.01.12.20 Other Biological Treatments and Preventive Measures

In 1910 Paul Ehrlich developed an arsenical compound later called Salvarsan that proved to be an effective treatment of syphilis. This received a mixed reception but by about six years later had been made the basis of a government-supported public health intervention in Britain (Ross & Tomkins, 1997). Surely the most unusual treatment for psychopathology so far developed was that of treating neurosyphilis by infecting the patients with malaria—the fever had the requisite antibiotic effect. This treatment was used by the psychiatrist Julius Wagner-Jauregg (1857–1940), a professor at the University of Vienna. In fact, he won a Nobel prize for this work in 1927. Now there are safer antibiotics available to treat syphilis.

Interestingly, Wagner-Jauregg had also been responsible previously for encouraging the Viennese government to attempt to prevent cretinism and other hyperthyroid conditions by requiring that salt be iodized (Wagner–Jauregg, 1894), an effective public health measure that is now in wide use (Whitrow, 1993).

One of the most curious episodes in the medical treatment of mental disorders was the one involving extracting all the teeth of patients with psychoses, on the theory that focal infections were maintaining their disturbed mental condition. This approach was carried out by Cotton (1922) and was endorsed for a time by prominent figures in the field, including psychiatrist Adolf Meyer. This proved to be nothing more than an unfortunate fad.

Insulin coma was introduced as a treatment for schizophrenia by psychiatrist Manfred Sakel (1900–1957) in Berlin in 1933. The rationale for this treatment was the notion, later considered quite in error, that there was some incompatibility between schizophrenia and seizures. Thus, the reasoning was that if one induced seizures, perhaps these could interrupt the psychosis. The use of insulin coma was discontinued after about two decades in favor of ECT, which was much more reliable in its production of seizures, without so many unwanted side effects, and even ECT was generally used in treating depression, not schizophrenia.

Metrazol (Cardiazol) convulsive therapy for schizophrenia was introduced in 1934 by Ladislaus von Meduna (1896–1964). The rationale for it was similar to that for insulin coma therapy. Meduna was trained as a neuropathologist and had done research comparing the brains of persons with epilepsy to the brains of those with schizophrenia. Thus, he bore major responsibility for the erroneous hypothesis elaborated above about the incompatibility between these two conditions. Meduna published an article on this new treatment in 1935 and a book on it in 1937. Metrazol therapy was also superceded by ECT.

Karl Landsteiner (1868–1948) and A. S. Wiener published their first paper on the Rh factor in human blood in 1940 (Wiener, 1952). This work ultimately led to the ability to prevent fetal malformations and mental retardation due to Rh factor incompatibility in mother and infant (Zimmerman, 1973).

Pediatricians Harry M. Meyer, Jr., chief of the Laboratory of Viral Immunology at the National Institutes of Health (NIH) in Bethesda, Maryland, and Paul D. Parkman developed a rubella vaccine that proved to be an effective preventive measure. It undoubtedly prevented a large number of fetal malformations, stillbirths, and cases of mental retardation, deafness, and other severe disabilities.

Finally, pediatricians Porter W. Anderson, Jr. and David Smith, together with immunologists John Robins and Rachel Schneerson, developed a vaccine against *Hemophilis influenzae* type b (Hib), the agent causing bacterial meningitis and responsible for death, deafness, and mental retardation among children (Holden, 1996).

1.01.13 EPILOGUE

This chapter was begun in search of the forerunners of present day psychiatry and clinical psychology. It was acknowledged in Section 1.01.1 that Hippocrates could not really be considered a psychiatrist. Perhaps he is best regarded as the ancestor of the many physicians who have to deal not only with physical illness but also with the emotional and behavioral problems that commonly present themselves in medical practice. Primary care physicians to this day provide reassurance, dispense advice, and prescribe medications for many such patients and are responsible for the care of those who are not referred to mental health specialists. Such physicians are the mainstay of the managed care system in the United States and of the national health systems of many other countries.

Similarly, Democritus was hardly a clinical psychologist. In his role as a scientist trying to investigate the fundamental causes of mental disorder, he could perhaps be most aptly viewed as the ancestor of persons such as Hideo Noguchi, the bacteriologist working at the Rockefeller Institute. There he discovered syphilis spirochetes in the brains of paretics.

The actual specialty of psychiatry did not develop until the late eighteenth and early nineteenth centuries. A transition took place early in the field in which the role of such physicians changed from that of simply restraining troublesome patients to engaging in "moral treatment" of them. Only a few of the early psychiatrists, such as Esquirol, Griesinger, and Kraepelin, were able to combine their clinical activities with those of a university professor and scientist. During the late nineteenth and early twentieth centuries, psychiatry expanded its scope of practice and for the most part superseded neurology in the care of outpatients with milder functional complaints. Psychiatry was then almost taken over by psychoanalysis, especially in the United States in the era preceding and immediately following World War II. Beginning in the 1950s, psychiatry was revolutionized by the development of psychotropic drugs, first the neuroleptics and subsequently the anxiolytics, antidepressants, and antimanic medications. It is currently undergoing some contraction of its activities due to managed care. Primary care physicians are taking over an increasing amount of the responsibility for the medical management of mental disorders, and other mental health practitioners are carrying out more of the psychotherapy.

Clinical psychology had its origins toward the end of the nineteenth century, with roots in the work of Pierre Janet and others in France and in Witmer's psychology clinic in Philadelphia. Its early development was much influenced by Binet's intelligence test and certainly by psychoanalysis, although psychologists for several decades had many barriers to overcome if they wanted to become psychotherapists. By the time of World War II, clinical psychologists were known primarily as mental testers, and for a time were well accepted into the "clinical team" (e.g., Menninger, 1950). Gradually in the postwar era psychologists became more active as psychotherapists in private practice. They were prominent in the development of behavioral therapy and behavior analysis and were responsible for the origins of formal research on psychotherapy. Along with their colleagues in various academic specialties within psychology, clinical psychologists such as Paul Meehl and Hans Eysenck also contributed significantly to research on psychopathogy. Having fought their battles to become psychotherapists, clinical psychologists then saw much of this role usurped in turn by social workers and master's-level counselors.

At present, it seems that one likely scenario for both psychiatry and clinical psychology might be provided by what happened to neurology at the end of the nineteenth century. Neurology continued to be an essential specialty in terms of research, teaching, and specialized consultation. Indeed, in this "decade of the brain" the neurosciences have become a major scientific frontier, attracting luminaries such as Francis Crick.

But after the 1890s, neurology tended to lose its "market share" in the everyday care of persons with functional nervous ailments. Similarly, psychiatry today remains an essential discipline in terms of its knowledge about areas such as psychopathology and psychopharmacology. Clinical psychology has comparable expertise in fields such as psychometrics and behavioral approaches to treatment. One cannot imagine that either of these fields will disappear, but it is likely for the moment that much of the everyday clinical work will be carried out by professionals with more modest levels of training.

ACKNOWLEDGMENTS

Preparation of this chapter was supported in part by a William Evans Fellowship to the author from the University of Otago, Dunedin, New Zealand. Many persons made helpful comments on early drafts, including John M. Reisman, Marion W. Routh, Eugene Taylor, and C. Eugene Walker.

1.01.14 REFERENCES

Achenbach, T. M., & Edelbrock, C. (1978). The classification of child psychopathology: A review and analysis of empirical efforts. *Psychological Bulletin, 85,* 1275–1301.

Ackerknecht, E. H. (1942). In memory of William H. R. Rivers 1864–1922. *Bulletin of the History of Medicine, 11,* 478–481.

Ackerknecht, E. H. (1967). *Medicine at the Paris hospital 1794–1848.* Baltimore: Johns Hopkins University Press.

Ackerknecht, E. H. (1986). Private institutions in the genesis of psychiatry. *Bulletin of the History of Medicine, 60,* 387–395.

Agras, W. S., Barlow, D. H., & Chapin, H. N. (1974). Behavior modification of anorexia nervosa. *Archives of General Psychiatry, 30,* 279–286.

Albee, G. (1968). Conceptual models and manpower requirements in psychology. *American Psychologist, 23,* 317–320.

Albee, G. (1970). The uncertain future of clinical psychology. *American Psychologist, 25,* 1071–1080.

Albee, G. (1977). Does including psychotherapy in health insurance represent a subsidy to the rich from the poor? *American Psychologist, 32,* 719–721.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

Andreasen, N. C. (1988). Brain imaging applications in psychiatry. *Science, 239,* 1381–1388.

Angst, J., Weis, P., Grof, P., Baastrup, P. C., & Schou, M. (1970). Lithium prophylaxis in recurrent affective disorders. *British Journal of Psychiatry, 116,* 604–614.

Ayllon, T., & Azrin, N. H. (1968). *The token economy.* New York: Appleton-Century-Crofts.

Ayllon, T., & Michael, J. (1959). The psychiatric nurse as a behavioral engineer. *Journal for the Experimental Analysis of Behavior, 2,* 323–334.

Azrin, N. H., & Foxx, R. M. (1974). *Toilet training in less than a day.* New York: Simon & Schuster.

Bandura, A. (1971). Behavior therapy and the models of man. *American Psychologist, 29,* 859–869.

Barkley, R. A. (1990). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment*. New York: Guilford.

Bayley, N. (1955). On the growth of intelligence. *American Psychologist, 10,* 805–818.

Bayley, N. (1969). *Manual for the Bayley Scales of Infant Development*. New York: Psychological Corporation.

Beard, G. (1880). *A practical treatise on nervous exhaustion (neurasthenia)*. New York: Wood.

Beck, A. T. (1963). Thinking and depression. 1. Idiosyncraatic content and cognitive distortion. *Archives of General Psychiatry, 9,* 324–344.

Beck, A. T. (1976). *Cognitive therapy and emotional disorder*. Philadelphia: University of Pennsylvania Press.

Beck, A. T., Ward, C. H., & Mendelsohn, M. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561–571.

Beck, S. J. (1937). *Introduction to the Rorschach method* (Monograph No. 1). New York: American Orthopsychiatric Association.

Beck, S. J. (1944). *Rorschach's test: Vol. 1. Basic processes*. New York: Grune & Stratton.

Beers, C. W. (1908). *A mind that found itself*. New York: Longmans Green.

Bell, J. E. (1961). *Family group therapy* (USPHS Monograph No. 64). Washington, DC: US Government Printing Office.

Bender, L. (1938). *A visual motor gestalt test and its clinical use*. New York: American Orthopsychiatric Association.

Benjamin, L., Jr. (1986). Why don't they understand us? A history of psychology's public image. *American Psychologist, 41,* 941–946.

Benton, A. L. (1945). A visual retention test for clinical use. *Archives of Neurology and Psychiatry, 54,* 212–216.

Bergin, A. E. (1966). Some implications of psychotherapy research for therapeutic practice. *Journal of Abnormal Psychology, 71,* 235–246.

Berrios, G. E. (1996). *The history of mental symptoms: Descriptive psychopathology since the nineteenth century*. Cambridge, UK: Cambridge University Press.

Bertin, C. (1982). *Marie Bonaparte: A life*. New York: Harcourt Brace Jovanovich.

Bettelheim, B. (1950). *Love is not enough*. Glencoe, IL: Free Press.

Bettelheim, B. (1967). *The empty fortress*. New York: Free Press.

Beutler, L. E. (1979). Toward specific psychological therapies for specific conditions. *Journal of Consulting and Clinical Psychology, 47,* 882–897.

Bijou, S. W. (1996). Reflections on some early events related to behavior analysis of child development. *Behavior Analyst, 19,* 49–60.

Binet, A., & Simon, T. (1905). [A new method for the diagnosis of intellectual level of abnormal persons.] *Annee psychologique, 11,* 191–244.

Bleuler, E. (1908). Die Prognose der Dementia praecox (Schizophrenie Gruppe). *Allgemeine Zeitschrift fur Psychiatrie, 65,* 436–437.

Bleuler, E. (1911). *Dementia praecox oder die Gruppe der Schizophrenien*. Tubingen: Diskord (reprinted, 1988).

Bleuler, M. (1972). *Schizophrenic disorders in the light of long-term illness and family history*. Stuttgart: G. Thieme.

Bluestein, B. E. (1992). Percival Bailey and neurology at the University of Chicago, 1928–1939. *Bulletin of the History of Medicine, 66,* 90–113.

Bourgeois, M. L., & Geraud, M. (1997). Jules Baillarger, 1809–1891. *American Journal of Psychiatry, 154,* 262.

Breger, L., & McGaugh, J. L. (1965). Critique and reformulation of "learning-theory" approaches to psychotherapy and neurosis. *Psychological Bulletin, 63,* 338–358.

Breuer, J., & Freud, S. (1986). *Studies in hysteria*. In *The standard ediction of the complete psychological works of Sigmund Freud*. (Vol. 2). London: Hogarth Press and the Institute of Psychoanalysis (originally published, 1895).

Buck v. Bell (1927). 274 U.S. 200.

Buhler, C. (1971). Basic theoretical concepts of humanistic psychology. *American Psychologist, 26,* 378–386.

Burnham, W. H. (1924). *The normal mind: An introduction to mental hygiene and the hygiene of school education*. New York: Appleton.

Burton, R. (1971). *The anatomy of melancholy*. Amsterdam, NY: De Capo Press (facsimile edition) (originally published, 1621).

Bynum, W. F. (1974). Rationale for therapy in British psychiatry: 1780–1835. *Medical History, 18,* 317–334.

Caplan, R. B., & Caplan, G. (1969). *Psychiatry and the community in nineteenth century America*. New York: Basic Books.

Carlson, E. T., & Simpson, M. M. (1969). Models of the nervous system in eighteenth century psychiatry. *Bulletin of the History of Medicine, 43,* 101–115.

Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin, 98,* 388–400.

Cattell, J. M. (1890). Mental tests and measurements. *Mind, 15,* 373–381.

Cerletti, U., & Bini, L. (1938). L'Elettroshock. *Archivio generale di neurologia, 19,* 266–268.

Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology, 72,* 193–204.

Chapman, L. J., & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology, 74,* 271–280.

Claren, S. K., & Smith, D. W. (1978). The fetal alcohol syndrome. *New England Journal of Medicine, 298,* 1063–1067.

Clark, P. F. (1959). Hideyo Noguchi, 1876–1928. *Bulletin of the History of Medicine, 33,* 1–20.

Clarke, A. D. B., & Clarke, A. M. (Eds.) (1958). *Mental deficiency: The changing outlook*. London: Methuen.

Cleckley, H. M. (1941). *The mask of sanity*. St. Louis, MO: Mosby.

Conners, C. K., & Eisenberg, L. (1963). The effects of methylphenidate on symptomatoloty and learning in disturbed children. *American Journal of Psychiatry, 120,* 458–464.

Cotton, H. A. (1922). The etiology and treatment of the so-called functional psychoses. *American Journal of Psychiatry, 2,* 157–210.

Cummings, N., & Fillette, W. T. (1967). Psychiatric services and medical utilization in a prepaid health plan setting. *Medical Care, 5,* 25–35.

Dain, N. (1994). Psychiatry and anti-psychiatry in the United States. In M. S. Micale & R. Porter (Eds.), *Discovering the history of psychiatry* (pp. 415–444). New York: Oxford University Press.

Danek, A., Gudden, W., & Distel, H. (1989). The dream king's psychiatrist Bernard von Gudden (1824–1886): A life committed to rationality. *Archives of Neurology, 46,* 1349–1353.

de la Cruz, F. F. (1985). Fragile X syndrome. *American Journal of Mental Deficiency, 90,* 119–123.

Delay, J., Deniker, R., & Harl, J-M. (1952). Utilisation en therapeutique psychiatrique d'un phenothiazine de'action centrale elective (4560 RP). *Annales Medico-psychologique, 110,* 112–120.

Doe v. Bolton (1973). 410 U.S. 179.

Doll, E. A. (1935). A genetic scale of social maturity. *American Journal of Orthopsychiatry, 5,* 180–190.

Doll, E. A., Phelps, W. M., & Melcher, R.C. (1932). *Mental deficiency due to birth injuries*. New York: Macmillan.

Doll, E. E. (1988). Before the big time: Early history of the Training School at Vineland, 1888–1949. *American Journal on Mental Retardation, 93*, 1–15.

Dollard, J., Doob, L. W., Miller, N. E., Mowrer, O. H., & Sears, R. R. (1939). *Frustration and aggression*. New Haven, CT: Yale University Press.

Dols, M. W. (1987). Insanity and its treatment in Islamic society. *Medical History, 31*, 1–14.

Dowbiggin, I. (1989). French psychiatry and the search for a professional identity: The Societe medico-psychologique, 1840–1870. *Bulletin of the History of Medicine, 63*, 331–355.

Down, J. L. (1866). Observations on the ethnic classification of idiots. *London Hospital Clinic Lecture Reports, 3*, 259–262.

Drotar, D. (1981). Psychological perspectives on chronic childhood illness. *Journal of Pediatric Psychology, 6*, 211–228.

Dunbar, H. F. (1938). *Emotions and bodily changes*. New York: Columbia University Press.

Dunlap, K. (1932). *Habits: Their making and unmaking*. New York: Liveright.

Edgerton, R. (1967). *The cloak of competence*. Berkeley, CA: University of California Press.

Ellenberger, H. (1970). *The discovery of the unconscious: The history and evolution of dynamic psychiatry*. New York: Basic Books.

Ellis, A. (1958). Rational psychotherapy. *Journal of General Psychology, 59*, 35–49.

Ellis, N. R. (Ed.) (1963). *Handbook of mental deficiency: Psychological theory and research*. New York: McGraw-Hill.

Erikson, E. H. (1950). *Childhood and society*. New York: Norton.

Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology, 16*, 319–324.

Fancher, R. E. (1987). Henry Goddard and the Kallikak family photographs: "Conscious skullduggery" or "Whig history"? *American Psychologist, 42*, 585–590.

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Jr., Winokur, G., & Munoz, R. (1972). Diagnostic crtieria for use in psychiatric research. *Archives of General Psychiatry, 26*, 57–63.

Finkelhor, D. (1984). *Child sexual abuse: New theory and research*. New York: Free Press.

Foucault, M. (1961). *Histoire de la Folie a l'age classique*. Paris: Plon (the English translation of this book had the title, *Madness and civilization*).

Frank, J. D. (1973). *Pursuasion and healing: A comparative study of psychotherapy*. Baltimore: Johns Hopkins University Press.

Frank, J. D., & Frank, J. B. (1991). *Persuasion and healing* (Rev. ed.). Baltimore: Johns Hopkins University Press.

Frankl, V. (1962). *Man's search for meaning*. Boston: Beacon Press.

Freud, A. (1937). *The ego and the mechanisms of defense*. Vienna: International Psychoanalytic Verlag.

Freud, S. (1927). *The problem of lay analyses*. New York: Brentano's.

Freud, S. (1959). Psychoanalytic notes upon an autobiographical account of a case of paranoia. *Collected papers* (Vol. 3). New York: Basic Books (originally published, 1911).

Gallaher, D. (1995). *Voice for the mad: The life of Dorothea Dix*. New York: Free Press.

Galton, F. J. (1978). *Hereditary genius: An inquiry into its laws and consequences*. London: Julian Friedmann (originally published, 1869).

Gauld, A. (1992). *A history of hypnotism*. Cambridge, UK: Cambridge University Press.

Gay, P. (1988). *Freud: A life for our time*. New York: Norton.

Gesell, A. (1929). *Infancy and human growth*. New York: Macmillan.

Goddard, H. H. (1912). *The Kallikak family*. New York: Macmillan.

Goffman, E. (1961). *Asylums: Essays on the social situations of mental patients and other inmates*. New York: Doubleday.

Goldstein, J. (1987). *Console and classify: The French psychiatric profession in the nineteenth century*. Cambridge, UK: Cambridge University Press.

Goodenough, F. L. (1926). *Measurement of intelligence by drawings*. Yonkers, NY: World.

Grob, G. N. (1986). Psychiatry and social activism. *Bulletin of the History of Medicine, 60*, 477–501.

Grunbaum, A. (1984). *The foundations of psychoanalysis: A philosophical critique*. Berkeley: University of California Press.

Grunbaum, A. (1993). *Validation in the clinical theory of psychoanalysis: A study in the philosophy of psychoanalysis*. Madison, CT: International Universities Press.

Halstead, W. C. (1947). *Brain and intelligence*. Chicago: University of Chicago Press.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry, 23*, 56–61.

Hanson, F. A. (1993). *Testing testing: Social consequences of the examined life*. Berkeley: University of California Press.

Hartman, H., Kris, E., & Loewenstein, R. M. (1946). Comments on the formation of psychic structure. *Psychoanalytic Study of the Child, 2*, 11–38.

Hathaway, S. R., & McKinley, J. C. (1940). A multiphasic personality inventory schedule. *Journal of Psychology, 10*, 249–254.

Hathaway, S. R., & McKinley, J. C. (1943). *The Minnesota Multiphasic Personality Inventory*. New York: Psychological Corporation.

Hecht, F. (1994). Jerome Lejeune (1926–1994). In memoriam. *American Journal of Human Genetics, 55*, 201–202.

Heiser, K. F. (1945). Certification of psychologists in Connecticut. *Psychological Bulletin, 42*, 624–630.

Hernstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.

Herzberg, A. (1945). *Active psychotherapy*. New York: Grune & Stratton.

Hinshelwood, J. (1900). Congenital word blindness. *Lancet, 1*, 1506.

Hobbs, N. (1966). Helping disturbed children: Psychological and ecological strategies. *American Psychologist, 21*, 1105–1115.

Hoeldtke, R. (1967). The history of associationism and British medical psychology. *Medical History, 11*, 46–65.

Holden, C. (1996). Lasker awards. *Science, 274*, 39.

Hollingworth, L. S. (1918). Tentative suggestions for the certification of practicing psychologists. *Journal of Applied Psychology, 2*, 280–284.

Holtzman, W. H., Thorpe, J. S., Swartz, J. D., & Herron, E. W. (1961). *Inkblot perception and personality*. Austin, TX: University of Texas Press.

Holzman, P. S., Proctor, L. R., & Hughes, D. W. (1973). Eye tracking patterns in schizophrenia. *Science, 181*, 179–181.

Hooker, E. (1969). Parental relations and male homosexuality in patient and non-patient samples. *Journal of Consulting and Clinical Psychology, 33*, 140–142.

Hubbard, L. R. (1950). *Dianetics: The modern science of mental health*. Los Angeles: American St. Hill Organization.

Huesmann, L. R., Eron, L. D., Lefkowitz, M. M., & Walder, L. O. (1984). Stability of aggression over time and generations. *Developmental Psychology, 20*, 1120–1134.

Huey, E. B. (1908). *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press.

Hunter, R., & McAlpine, I. (1963). *Three hundred years of psychiatry, 1535–1860: A history presented in selected English texts*. London: Oxford University Press.

Isaac, R. J., & Armat, V. C. (1990). *Madness in the streets*. New York: Free Press.

Jackson, S. W. (1972). Unusual mental states in medieval Europe. I. Medical syndromes of mental disorder, 400–1100 A.D. *Journal of the History of Medicine and Allied Sciences, 27,* 262–297.

Jackson, S. W. (1978). Melancholia and the waning of the humeral theory. *Journal of the History of Medicine and Allied Sciences, 33,* 367–376.

Jacobson, E. (1929). *Progressive relaxation*. Chicago: University of Chicago Press.

Jacobson, N. (1989). The maintenance of treatment gains following social learning-based marital therapy. *Behavior Therapy, 20,* 325–336.

Jacobson, N., & Revenstorf, D. (1988). Statistics for assessing the clinical significance of psychotherapy techniques. *Behavioral Assessment, 10,* 133–145.

Janet, P., & Raymond, F. (1903). *Obsessions et al psychastehia*. Paris: Alcan.

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review, 39,* 1–123.

Johnstone, E. C., Crow, T. J., Frith, C. D., Husband, J., & Kneel, L. (1976). Cerebral ventricle size and cognitive impairment in chronic schizophrenia. *Lancet, 2,* 924–926.

Jones, K., Smith, D., Ulleland, C., & Streissguth, A. (1973). Patterns of malformation in offspring of chronic alcoholic mothers. *Lancet, 1,* 1267–1271.

Jones, M. C. (1924). The elimination of children's fears. *Journal of Experimental Psychology, 7,* 382–390.

Kallmann, F. J. (1938). *The genetics of schizophrenia: A study of heredity and reproduction in the families of 1,087 schizophrenics*. New York: Augustin.

Kallmann, F. J. (1946). The genetic theory of schizophrenia: An analsis of 691 schizophrenic twin index families. *American Journal of Psychiatry, 103,* 309–322.

Kallmann, F. J. (1953). *Heredity in health and mental disorder: Principles of psychiatric genetics in the light of comparative twin studies*. New York: Norton.

Kamin, L. J. (1974). *The science and politics of IQ*. Potomac, MD: Erlbaum.

Kanner, L. (1943). Autistic disturbance of affective contact. *Nervous Child, 2,* 217–250.

Kanner, L. (1959). Johann Jakob Guggenbuhl and the Abendberg. *Bulletin of the History of Medicine, 33,* 489–502.

Kanner, L., & Eisenberg, L. (1956). Early infantile autism. 1944–55. *American Journal of Orthopsychiatry, 26,* 556–566.

Kazdin, A. E. (1978). *History of behavior modification: Experimental foundations of contemporary research*. Baltimore: University Park Press.

Kelly, G. A. (1955). *The psychology of personal constructs*. New York: Norton.

Kernberg, O. (1967). Borderline personality organization. *Journal of the American Psychoanalytic Association, 15,* 641–685.

Kernberg, O. (1975). *Borderline conditions and pathological narcissism*. New York: Aronson.

Kesey, K. (1962). *One flew over the cuckoo's nest*. New York: Viking.

Kety, S. S. et al. (1968). The types and prevalence of mental illness in the biological and adoptive families of adopted schizophrenics. *Journal of Psychiatric Research, 6* (Suppl. 1), 345–362.

Kety, S. S., Wender, P. H., Jacobsen, B., Ingraham, L. J., Jansson, L., Faber, B., & Kinney, D. K. (1994). Mental illness in the biological and adoptive relatives of schizophrenic adoptees: Replication of the Copenhagen study in the rest of Denmark. *Archives of General Psychiatry, 51,* 442–455.

Kirk, S. A. (1952). Experiments in the early training of the mentally retarded. *American Journal of Mental Deficiency, 56,* 692–700.

Klein, D. F. (1964). Behavioral reaction pattents with phenothiazines. *Archives of General Psychiatry, 7,* 397–408.

Klein, D. F., & Fink, M. (1962). Behavioral reaction patterns with phenothiazines. *Archives of General Psychiatry, 5,* 449–459.

Klein, M. (1932). *The psycho-analysis of children*. New York: Norton.

Klerman, G. L. (1990). The psychiatric patient's right to effective treatment: Implications of Osheroff v. Chestnut Lodge. *American Journal of Psychiatry, 147,* 409–418.

Knight, R. P. (1953). The present status of organized psychoanalysis in the United States. *American Psychoanalytic Association Journal, 1,* 197–221.

Korman, M. (1974). National conference on levels and patterns of professional training in psychology: The major themes. *American Psychologist, 29,* 441–449.

Kozulin, A. (1990). *Vygotsky's psychology: A biography of ideas*. Cambridge, MA: Harvard University Press.

Kraemer, H., & Sprenger, J. (1971). *Malleus malificarum* (M. Summers Trans.). New York: Dover (originally published, 1486.).

Kraemer, R. (1988). *Maria Montessori: A biography*. Reading, MA: Addison-Wesley.

Krantz, D. L. (1971). The separate worlds of operant and non-operant psychology. *Journal of Applied Behavior Analysis, 4,* 61–70.

Kuhn, R. (1957, August 31). Uber die Behandlung depressiver Zustande mit einem Iminodibenzyl derivat (G22355). *Schweitzerische Medicine Wochenschrift, 87,* 1135–1140.

Laing, R. D. (1960). *The divided self: A study of sanity and madness*. London: Tavistock.

Lambert, P. A. (1966). Action neuropsychotropique d'un nouvel anti-epileptique: Le Depamide. *Annales medico-psychologique, 1,* 707–710.

Lang, P. J., & Lazovik, A. D. (1963). Experimental desensitization of a phobia. *Journal of Abnormal and Social Psychology, 66,* 519–525.

Laufer, M. W., & Denhoff, E. (1957). Hyperkinetic behavior syndrome in children. *Pediatrics, 50,* 463–474.

Lazarus, R. S. (1966). *Psychological stress and the coping process*. New York: McGraw-Hill.

Lehmann, H. E. (1989). The introduction of chlorpromazine to North America. *Psychiatric Journal of the University of Ottawa, 14,* 263–265.

Lehmann, H. E., & Hannahan, G. E. (1954). Chlorpromazine: New inhibiting agent for psychomotor excitement and manic states. *AMA Archives of Neurology and Psychiatry, 71,* 227–237.

Leiby, J. (1978). *A history of social welfare and social work in the United States, 1815–1972*. New York: Columbia University Press.

Leigh, D. (1957). Recurrent themes in the history of psychiatry. *Medical History, 1,* 237–248.

Levy, L. (1962). The skew in clinical psychology. *American Psychologist, 17,* 244–249.

Linehan, M. M. (1993). *Cognitive-behavioral therapy of borderline personality disorder*. New York: Guilford.

Locke, J. (1824). *An essay concerning human understanding* (25th ed.). London: William Baynes and Son (originally published, 1639).

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual function in young autistic children. *Journal of Consulting and Clinical Psychology, 55,* 3–9.

Loyd, B. H., & Abidin, R. R. (1985). Revision of the Parenting Stress Index. *Journal of Pediatric Psychology, 10,* 169–177.

Luborsky, L. (1962). Clinicians' judgments of mental health. *Archives of General Psychiatry, 7,* 404–417.

Lubs, H. (1969). A marker X chromosome. *American Journal of Human Genetics, 21,* 231–244.

Machover, K. (1948). *Personality projections in the drawing of the human figure*. Springfield, IL: Charles C. Thomas.

Magner, L. N. (1992). *A history of medicine*. New York: Marcel Dekker.

Magnus-Levy, A. (1944). The heroic age of German medicine. *Bulletin of the History of Medicine, 16,* 331–342.

Marks, I. M. (1969). Genetics of fear and anxiety disorders. *British Journal of Psychiatry, 149,* 406–418.

Marx, O. M. (1994). The beginnings of psychiatric historiography in nineteenth-century Germany. In M. S. Micale & R. Porter (Eds.), *Discovering the history of psychiatry* (pp. 39–52). New York: Oxford University Press.

Masserman, J. H. (1943). *Behavior and neurosis: An experimental psycho-analytic approach to psychobiologic principles*. Chicago: University of Chicago Press.

Masson, J. M. (1984). *The assault on truth: Freud's suppression of the seduction theory*. New York: Farrar Straus Giroux.

Masters, W. H., & Johnson, V. E. (1966). *Human sexual response*. Boston: Little, Brown.

McPherson, F. M. (1986). The professional psychologist in Europe. *American Psychologist, 41,* 302–305.

McReynolds, P. (1997). *Lightner Witmer: His life and times*. Washington, DC: American Psychological Association.

Meehl, P. E. (1962). Schiztaxia, schizotypy, and schizophrenia. *American Psychologist, 17,* 827–838.

Menninger, W. C. (1950). The relationship of clinical psychology and psychiatry. *American Psychologist, 5,* 3–15.

Mettler, F. A., & Columbia-Greystone Associates (Eds.) (1949). *Selective partial ablation of the frontal cortex*. New York: Hoeber.

Milofsky, C. (1989). *Testers and testing: The sociology of school psychology*. New Brunswick, NJ: Rutgers University Press.

Mitchell, S. W. (1894). Address before the 50th Annual Meeting of the American Medico-Psychological Association, held in Philadlephia, May 16, 1894. *Journal of Nervous and Mental Diseases, 21,* 413–438.

Money, J. (1963). Cytogenetics and psychosexual incongruities, with a note on space form blindness. *American Journal of Psychiatry, 119,* 820–827.

Moniz, E. (1937). Prefrontal leucotomy in the treatment of mental disorder. *American Journal of Psychiatry, 93,* 1379–1385.

Mora, G. (Ed.) (1991). *Witches, devils, and doctors in the renaissance*: Edition of Johann Weyer (1583), *De praestigiis daemonum*. Binghamton, NY: Medieval and Renaissance Texts and Studies.

Mora, G. (1994). Early American historials of psychiatry: 1910–1960. In M. S. Micale & R. Porter (Eds.), *Discovering the history of psychiatry* (pp. 53–80). New York: Oxford University Press.

Moreno, J. L. (1952). *Who shall survive*? New York: Beacon House.

Morgan, C. D., & Murray, H. A. (1935). A method for investigating fantasies: The Thematic Apperception Test. *Archives of Neurology and Psychiatry, 34,* 289–306.

Moriyama, N. (1991). Shoma Morita, founder of Morita therapy, and haiku poet Shiki: Origin of Morita therapy. *Japanese Journal of Psychiatry and Neurology, 45,* 787–796.

Mowrer, O. H., & Mowrer, W. M. (1938). Enuresis: A method for its study and treatment. *American Journal of Orthopsychiatry, 8,* 436–459.

Murray, H. A. (1959). *Explorations in personality*. New York: Oxford University Press.

O'Brien, C. (1996). Auguste D. and Alzheimer's disease. *Science, 275,* 28.

O'Connor, N., & Tizard, J. (1956). *The social problem of mental deficiency*. London: Pergamon.

Orton, S. T. (1928). Specific reading disability: Strephosymbolia. *Journal of the American Medical Association, 90,* 1095.

Orton, S. T. (1937). *Reading, writing, and speech problems in children*. London: Chapman and Hall.

Pavlov, I. P. (1927). *Conditioned responses: An investigation of the physiological activity of the cerebral cortex*. G. V. Anrep (trans.). London: Oxford University Press.

Penrose, L. S. (1951). *The biology of mental defect*. London: Sidgwick and Jackson.

Perelman, S. J. (1957). *The road to Miltown; or, Under the spreading atrophy*. New York: Simon and Schuster.

Plato. (1993). *Republic*. R. Waterfield (trans.). Oxford, UK: Oxford University Press.

Porter, R. (1987). *Mind-forg'd manacles: A history of madness in England from the Restoration to the Regency*. Cambridge, MA: Harvard University Press.

Rachman, S. (1959). The treatment of anxiety and phobic reactions by systematic desensitization psychotherapy. *Journal of Abnormal and Social Psychology, 58,* 259–263.

Rapaport, J. L. (1989). *The boy who couldn't stop washing*. New York: Dutton.

Reisman, J. M. (1991). *A history of clinical psychology* (2nd ed.). New York: Hemisphere.

Richmond, M. (1917). *Social diagnosis*. New York: Russell Sage Foundation.

Riese, W. (1949). An outline of a history of ideas in neurology. *Bulletin of the History of Medicine, 23,* 111–136.

Roberts, K. S., & Sacks, N. P. (1954). Dom Duarte and Robert Burton: Two men of melancholy. *Journal of the History of Medicine, 9,* 21–37.

Robinson, V. P. (Ed.) (1962). *Jessie Taft: Therapist and social work educator*. Philadelphia: University of Pennsylvania Press.

Robins, L. N. (1966). *Deviant children grown up*. Baltimore: Williams & Wilkins.

Rogers, C. R. (1951). *Client centered therapy*. Boston: Houghton Mifflin.

Rorschach, H. (1921). *Psychodiagnostic*. Bern: Huber.

Rosen, G. (1964). The mentally ill and the community in western and central Europe during the late middle ages and the renaissance. *Journal of the History of Medicine, 19,* 377–388.

Rosenberg, C. E. (1962). The place of George M. Beard in nineteenth-century psychiatry. *Bulletin of the History of Medicine, 36,* 245–259.

Rosenhan, D. L. (1973). On being sane in insane places. *Science, 179,* 250–258.

Ross, A. O. (1959). *The practice of clinical child psychology*. New York: Grune & Stratton.

Ross, J. E., & Tomkins, S. M. (1997). The British reception of Salvarsan. *Journal of the History of Medicine and Allied Sciences, 52,* 398–423.

Roudinesco, E. (1990). *Jacques Lacan & Co.: A history of psychoanalysis in France, 1925–1985*. J. Mehlman (trans). Chicago: University of Chicago Press.

Rousseau, J. J. (1911). *Emile, or education*. New York: Dutton (originally published, 1762).

Routh, D. K. (1994). *Clinical psychology since 1917: Science, practice, and organization*. New York: Plenum.

Routh, D. K. (1996). Lightner Witmer and the first 100 years of clinical psychology. *American Psychologist, 51,* 244–247.

Routh, D. K. (1997). A history of Division 12 (clinical psychology): Fourscore years. In D. A. Dewsbury (Ed.), *Unification through division: Histories of the divisions of* *the American Psychological Association* (Vol. 2, pp. 55–82). Washington, DC: American Psychological Association.

Rubenstein, E. A. (1992). Robert H. Felix (1904–1990). *American Psychologist, 47,* 589.

Russell, G. F. M. (1979). Bulimia nervosa: An ominous variant of anorexia nervosa. *Psychological Medicine, 9,* 429–448.

Salter, A. (1941). Three techniques of autohypnosis. *Journal of General Psychology, 24,* 423–438.

Salter, A. (1949). *Conditioned reflex therapy*. New York: Creative Age.

Santos, E. H., & Stainbrook, E. (1949). A history of psychiatric nursing in the nineteenth century. Part I. *Journal of the History of Medicine, 4,* 48–60.

Sarton, G. (1954). *Galen of Pergamon*. Lawrence, KS: Univeristy of Kansas Press.

Schlick, T. (1994). Changing disease identities: Cretinism, politics, and surgery (1844–1892). *Medical History, 38,* 421–443.

Schou, M. et al. (1954). The treatment of manic psychoses by the administration of lithium salts. *Journal of Neurology, Neurosurgery, and Psychiatry, 17,* 250–260.

Schreber, D. P. (1903). *Denkwurdigkeiten eines Nervenkranken*. Leipzig: Mutze.

Scull, A. (1993). *The most solitary of afflictions: Madness and society in Britain 1700–1900*. New Haven, CT: Yale University Press.

Scull, A., MacKenzie, C., & Hervey, N. (1996). *Masters of Bedlam: The transformation of the mad-doctoring trade*. Princeton, NJ: Princeton University Press.

Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death*. San Francisco: Freeman.

Selikoff, I. J., Robitzek, E. H., & Ornstein, G. G. (1952). Toxicity of hydrazine derivatives of isonicotinic acid in the chemotherapy of human tuberculosis. *Bulletin of the Sea View Hospital, 13,* 17.

Shapiro, E., Shapiro, A. K., Fulop, G., Hubbard, M., Mandel, J., Nordlie, J., & Phillips, R. A. (1989). Controlled study of haloperidol, pimozide, and placebo for the treatment of Gilles de la Tourette's syndrome. *Archives of General Psychiatry, 46,* 722–730.

Shapiro, J. P. (1993). *No pity: People with disabilities forging a new civil rights movement*. New York: Times Books.

Shorter, E. (1997). *A history of psychiatry*. New York: Wiley.

Sigerist, H. E. (1945). The need for an institute of the history of medicine in India. *Bulletin of the History of Medicine, 17,* 113–126.

Sjogren, T. (1931). Die juvenile amaurotische Idiotie: klinische and erblichkeitsmedicinische Untersuchungen. *Hereditas (Lund), 14,* 197–225.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy studies. *American Psychologist, 32,* 752–760.

Snyder, S. H. (1976). The dopamine hypothesis of schizophrenia: Focus on the dopamine receptor. *American Journal of Psychiatry, 133,* 197–202.

Spearman, C. E. (1904). General intelligence objectively determined and measures. *American Journal of Psychology, 15,* 201–292.

Spielberger, C. D., Gorsuch, R., & Lushene, R. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Spitzer, R. L., Endicott, J., & Gibbon, M. (1978). Research diagnostic criteria: Rationale and reliability. *Archives of General Psychiatry, 35,* 773–782.

Stengel, E. (1959). Classification of mental disorders. *Bulletin of the World Health Organization, 21,* 601–663.

Stern, W. (1914). *The psychological method of testing intelligence*. Baltimore: Warwick & Wood (originally published, 1912).

Still, G. F. (1902). Some abnormal physical conditions in children. *Lancet, 1,* 1008–1012, 1077–1082, 1163–1168.

Strauss, A. A., & Lehtinen, L. E. (1947). *Psychopathology and education of the brain injured child*. New York: Grune & Stratton.

Sullivan, H. S. (1931). The modified psychoanalytic treatment of schizophrenia. *American Journal of Psychiatry, 11,* 519–540.

Sullivan, H. S. (1953). *The interpersonal theory of psychiatry*. New York: Norton. (published posthumously).

Taft, J. (1958). *Otto Rank*. New York: Julian.

Taylor, E. I. (1982). *William James on exceptional mental states: Reconstruction of the 1896 Lowell lectures*. New York: Scribner's Sons.

Temkin, O. (1946). An essay on the usefulness of medical history for medicine. *Bulletin of the History of Medicine, 19,* 9–47.

Temkin, O. (1947). Gall and the phrenological movement. *Bulletin of the History of Medicine, 21,* 275–321.

Temkin, O. (1973). *Galenism: Rise and decline of a medical philosophy*. Ithaca: Cornell University Press.

Temkin, O. (1991). *Hippocrates in a world of pagans and Christians*. Baltimore: Johns Hopkins University Press.

Terman, L. M. (1916). *The measurement of intelligence*. Boston: Houghton Mifflin.

Terman, L. M. (1925). *Genetic studies of genius* (Vol. 1). Stanford, CA: Stanford University Press.

Terman, L. M., & Oden, M. H. (1947). *The gifted child grows up: Twenty-five years' follow-up of a superior group*. Stanford, CA: Stanford University Press.

Thomson, G. (1919). The proof or disproof of the existence of general ability. *British Journal of Psychology, 9,* 321–336.

Truax, C. B. (1966). Reinforcement and nonreeinforcement in Rogerian psychotherapy. *Journal of Abnormal Psychology, 71,* 1–9.

Veith, I. (1955). Psychiatric thought in Chinese medicine. *Journal of the History of Medicine, 10,* 261–268.

Veith, I. (1963). The supernatural in far Eastern concepts of mental disease. *Bulletin of the History of Medicine, 37,* 139–155.

von Hug-Hellmuth, H. (1921). Zur Technik der Kinderanalyse. *Internationale Zeitschrift fur Psychoanalyse, 7.*

Wagner-Jauregg, J. (1894). Thyroid basis of cretinism, use of iodized salt. *Jahrbuch fur Psychiatrie, 12,* 102–107.

Walter, R. D. (1956). What became of the degenerate? A brief history of the concept. *Journal of the History of Medicine, 11,* 422–429.

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review, 23,* 158–179.

Watson, J. B., & Raynor, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3,* 1–14.

Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore: Williams & Wilkins.

Weiner, D. B. (1979). The apprenticeship of Philippe Pinel: A new document, "Observations of Citizen Pussin on the Insane." *American Journal of Psychiatry, 136,* 1128–1134.

Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology, 55,* 542–549.

Whitrow, M. (1993). *Julius Wagner-Jauregg (1857–1940)*. London: SmithGordon.

Wiener, A. S. (1952). History of Rhesus blood types. *Journal of the History of Medicine, 7,* 369–383.

Williams, C. D. (1959). The elimination of tantrum behavior by extinction procedures. *Journal of Abnormal and Social Psychology, 59,* 269.

Winters, E. E. (1966). Adolf Meyer's two and a half years at Kankakee. *Bulletin of the History of Medicine, 40,* 441–458.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press.

Wolpe, J., Salter, A., & Reyna, L. J. (Eds.) (1964). *The conditioning therapies: The challenge in psychotherapy*. New York: Holt, Rinehart, and Winston.

Woods, E. A., & Carlson, E. T. (1961). The psychiatry of Philippe Pinel. *Bulletin of the History of Medicine, 35,* 14–25.

Woodworth, R. S. (1917). *Personal data sheet*. Chicago: Stoelting.

Wooldridge, A. (1994). *Measuring the mind: Education and psychology in England, c. 1860-c. 1900*. Cambridge, UK: Cambridge University Press.

Yerkes, R. M., Hardwick, J. W., & Bridges, R. S. (1915). *A point scale for measuring mental ability*. Baltimore: Warwick & York.

Zimmerman, D. R. (1973). *Rh: The intimate history of a disease and its conquest*. New York: Macmillan.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.02

The Scientific Status of Clinical Psychology

FRANK L. COLLINS, JR.
Oklahoma State University, Stillwater, OK, USA

1.02.1 INTRODUCTION 49
1.02.2 BRIEF HISTORY OF CLINICAL PSYCHOLOGY IN THE UNITED STATES 50
1.02.3 SCIENCE: DEFINITIONS AND MODELS FOR CLINICAL PSYCHOLOGY 51
  1.02.3.1 The Behavior of Scientists 51
  1.02.3.2 Science vs. Art 53
  1.02.3.3 Clinical Scientists and Clinical Practitioners 53
1.02.4 THE CURRENT STATUS OF SCIENCE IN CLINICAL PSYCHOLOGY 55
  1.02.4.1 Diagnosis and Assessment 55
  1.02.4.2 Scientific Status of Clinical Assessment 56
  1.02.4.3 Clinical Interventions 57
  1.02.4.4 Scientific Status of Clinical Interventions 58
1.02.5 TRAINING SCIENCE IN CLINICAL PSYCHOLOGY: THE SCIENTIST–PRACTITIONER MODEL 59
1.02.6 SCIENCE, THE CHANGING HEALTH CARE SYSTEM, AND THE DEMAND FOR PSYCHOLOGY 60
1.02.7 CONCLUSIONS 61
1.02.8 REFERENCES 61

> Psychologists work to develop a valid and reliable body of scientific knowledge based on research. (from the Preamble to the Ethical Principles of Psychologists and Code of Conduct; American Psychological Association, 1992)

1.02.1 INTRODUCTION

Psychology grew out of the fields of philosophy and experimental physiology (Boring, 1950) and is typically described as the scientific study of behavior (Sternberg, 1995; Zimbardo & Gerrig, 1996). This includes both animal and human behavior with the term "behavior" typically so broadly defined as to include physiological acts, overt behaviors, and thoughts. While clinical psychology can be considered as the applied discipline of psychology that focuses on the application of psychological principles to adjustment problems of individuals (Rotter, 1971; Shakow, 1976), the domain of clinical psychology includes "assessment, diagnosis, and treatment of clinical dysfunction; issues in the application of treatment to individuals or society at large; factors that influence adjustment and functioning in everyday life; development across the life span; personality; the family; delivery of psychological services; training of psychologists; and ethical and professional issues" (Kazdin, 1994, p. 1). While science plays a role in each of these areas, this chapter will focus on the scientific status of clinical psychology with respect to assessment and diagnosis, clinical interventions, and clinical training.

From its inception, clinical psychology has been described both as a science and a profession (Peterson, 1976). The flagship journal of clinical psychology, *Clinical Psychology: Science and Practice*, has chosen a name which highlights the dual identity of clinical psychology. Likewise, many of the activities of clinical psychologists do not fit our historical conceptualizations of science. For example, clinical psychologists are less likely to be engaged in laboratory-based "bench" research, rather they are more likely to serve as a consultant to an individual or agency to evaluate and develop a treatment plan relevant to a problem presented by an individual or group of individuals (family, couple, or other system). Depending on the results of the evaluation and plan, the clinical psychologist might also be engaged in the delivery of an intervention (psychotherapy) that has potential for reducing the distress identified by the individual or group being evaluated.

This chapter has several goals. First, it will identify the similarities, while acknowledging the differences, of the science and practice roles in clinical psychology. These differences will be focused on through the discussion of the different roles of clinical psychologists (science and practice) with an emphasis on why training in the science and practice of clinical psychology is critical for these dual roles. To accomplish this goal, a brief review of the history of clinical psychology is presented, followed by a discussion of the philosophy and definition of science. Concepts from behavioral psychology will be presented which will emphasize the role of different types of learning experiences and the importance of the scientist–practitioner training model on the behavior of the clinical psychologist primarily engaged in (i) clinical research—the clinical scientist or (ii) clinical practice—the clinical practitioner. These behavioral concepts will be used to reframe the "scientist" vs. "practitioner" debate in terms of the need to maintain a science-based practice of clinical psychology.

Second, by focusing on the similarities and contributions that clinical scientists and clinical practitioners make to the field of clinical psychology, it will be shown that the important question is not whether or not clinical psychology is a science but, rather, the question is, "What is the current status of science in clinical psychology?" This question will be addressed through a brief discussion of the current scientific status of diagnosis, assessment, and intervention.

Finally, changes in the health care system are likely to influence, or be influenced by, the science-based practice of clinical psychology. Potential issues and change will be discussed in terms of public policy and the science of clinical psychology.

1.02.2 BRIEF HISTORY OF CLINICAL PSYCHOLOGY IN THE UNITED STATES

Historically, the differentiation of psychology from philosophy was the emphasis which psychology placed on science and the scientific method (Sokal, 1992). However, there was little agreement as to which aspects of science were most important. Specifically, two distinct types of science emerged: (i) laboratory-based activities more in line with experimental physiology, and (ii) applied research activities which focused on individual differences and mental measurement (Goodwin, 1983). This division has been identified as a foreshadowing of the science–practice split in psychology through an emphasis of the difference between experimental and applied psychology (Rice, 1997).

Clinical psychology as a specialized area of training began with the establishment of the first psychology clinic by Lightner Witmer in 1986 (Brems, Thevenin, & Routh, 1991; Watson, 1953). Following World War I, many of the psychologists interested in the measurement of individual differences began to consult with organizations outside academia. This led to the first attempt to develop a professional identity complete with accreditation procedures for "consulting" psychologists. Although this movement for accreditation was not supported by the American Psychological Association (APA) Council (Wallin, 1960), consulting was viewed by the APA leadership as a form of applied research with all the components of an emerging professional identity (Rice, 1997).

By 1919 the clinical section of the APA was formed (Edelstein & Brasted, 1983), and several position papers (cf., Poffenberger, 1938; Rogers, 1939; Shakow, 1938) appeared that delineated the distinct requirements necessary to become a "clinical psychologist" (Derner, 1965; Edelstein & Brasted, 1983; Norton-Ford, 1982; Reisman, 1981).

World War II served to provide much of the emphasis for the emergence of the profession of psychology. First, the APA developed new bylaws which gave recognition to the science and profession of psychology (Capshew & Hilgard, 1992). Second, the Veterans Administration sponsored a large program to train clinical psychologists. This program incorporated university-based training in psychological theory with clinical practice training which included therapy in addition to psychological assessment. This model cumulated in the Boulder Conference (Raimy, 1950).

Since that time numerous national conferences have been held to address these issues (e.g., Stanford, Strother, 1956; Miami Beach, Roe, Custard, Moore, Ross, & Skodak, 1959; Hoch, Ross, & Skodak, 1966; Vail, Korman, 1976; West Virginia, Foster, Berler, & Collins, 1982; Gainesville, Belar & Perry, 1992; Ann Arbor, Belar et al., 1993; to name but a few), which resulted in the training guidelines followed today.

The early conferences (Boulder, Stanford, Miami Beach, Chicago, and Vail) focused on clinical training in general, with an emphasis on a broad-based background in psychology. Recommendations resulting from these conferences led to specific evaluation criteria and processes by which the APA accredits clinical training programs. Accreditation by the APA is presumed to be documentation that students receive certain minimum training experiences believed necessary to ensure that they have been appropriately trained as clinical psychologists. Each of these conferences reaffirmed the need for training in both the science and practice of clinical psychology.

The West Virginia Conference focused on the essential ingredients for training psychologists in a single school of thought, behavioral psychology. As an organizer of this meeting, I was sure that this "working conference" would specify objective criteria that would be necessary and sufficient for training in clinical psychology. Much to my surprise there was absolutely no consensus among this group of clinical psychologists, all who were engaged in behavioral training. "Indeed, after 6 hours of discussion among a group of prominent clinical directors, the only critical component of behavioral training that was unanimously endorsed was training in an empirical approach to the study of behavior" (Collins, Foster, & Berler, 1986, p. 304).

This emphasis on the importance of science training was the primary theme of the National Conference on Scientist–Practitioner Education and Training for the Professional Practice of Psychology (Belar & Perry, 1992). Recommendations from this conference will be discussed in greater detail later in this chapter; however, it is important to note that, while science plays a major role in the field of clinical psychology, the practice of clinical psychology has not always been integrated with science.

However, many authors (e.g., Rice, 1997; Schneider, 1990) have expressed concern that the division between the science and practice of psychology may lead to an irreversible division in psychology. The goal of this chapter is to show the importance of science in clinical psychology and the need to adhere to a science-based practice of clinical psychology. First, however, definitions and models will be presented to set the foundation for a discussion of the critical need for a science-based practice of clinical psychology.

1.02.3 SCIENCE: DEFINITIONS AND MODELS FOR CLINICAL PSYCHOLOGY

Defining science is not easy. One approach is to identify activities that we can all agree are *not* science (Sullivan & Collins, 1996). For example, most would agree that religion, art, and mysticism are not scientific fields. Physics, chemistry, physics, and biology are easy examples of science fields, and many of us would agree that sociology, anthropology, and psychology are scientific approaches. However, scholars might well engage in the scientific study of religion (theology), music, painting, etc. Are these individuals engaged in "science?" Does science provide unique perspectives to problems of behavior that are different or distinct from the perspectives provided by religion, art, or mysticism? Are there alternative conceptual models that can be used to describe how scientific behaviors develop? Why do scientists do what they do?

1.02.3.1 The Behavior of Scientists

At the most basic level, scientists engage in problem-solving activities that may include observation, conceptualization, and experimentation. Walker (1963) suggests that "the scientific method is a survival technique that developed during the biological evolution of living things. Any organism or device that includes a suitably connected memory unit can "learn by experience," and this learning by experience contains the basic elements of the scientific method" (p. 15). While the basic elements of the scientific method may account for all learning, contemporary science is a bit more complicated.

Kuhn (1970) offers a somewhat more complex definition of science. For Kuhn, science

> means research firmly based upon one or more past scientific achievements, achievements that some particular scientific community acknowledges for a time as supplying the foundation for its further practice. Today such achievements are recounted, though seldom in their original form, by science textbooks, elementary and advanced. These textbooks expound the body of accepted theory, illustrate many or all of its successful applications, and compare these applications with exemplary observations and experiments. (p. 10)

Thus, science is more than merely learning from experience. A science must be based on information, theory, models, and methods that guide the actions of individuals engaged in science. In addition, knowledge about the behavior of the individual scientist can provide a more complete understanding of science. Sidman (1994) provides an interesting example of the importance of knowing about the scientist and the effect of his or her experiences on the development of theories and models. In the epilogue, he provides excerpts of correspondence between himself and Willard Day. This dialogue provides an interesting perspective of the influence that Day had on Sidman's approach to science.

The behavior of scientists can be broken down into conceptual or decision-making behaviors and experimental method behaviors. The conceptual process used in science is the hypothetico-deductive method (Walker, 1963). This method is a combination of induction, inspiration, and deduction (Walker). Using induction and inspiration, scientists postulate models or theories (Ackermann, 1970). From these models or theories, he or she can then deduce predictions or hypotheses which can then be verified by further observation.

Experimental methods include (i) identification of measures of interest, (ii) developing the methodology to conduct the measurements, and (iii) developing predictive models for analyzing the measures. Identification of measures has historically required specific definitions as to the operations used in measuring the quantity. For example, the weight of an object is the number read from a scale when the object is placed on the scale. The term *operational definition* was coined by P. W. Bridgman, who felt that it was the hallmark of science (Walker, 1963). While it has been argued that concepts which cannot be adequately defined using operations are not in the realm of science (Walker), most philosophers today have abandoned the doctrine of operational or explicit definition (Suppe, 1977).

This is an important point. Technological advances such as the invention of a more powerful microscope can change science by allowing for the operationalization of concepts that were previously seen as "not subject to empirical validations," or for the development of new operationalizations of a concept that may result in different outcomes. Concepts once thought to be unmeasurable become measurable with increased technology. Thus, the operationalization of a concept is quite arbitrary.

This historical focus on operationalization is perhaps one factor that resulted in the perception that science is value-free or a measure of the "truth." Today, there is general agreement that science, and the results of science, are influenced by the beliefs, expectations, and values of the researcher and the research community he or she belongs to for both science in general (Habermas, 1975; Holton, 1973; Polanyi, 1958) and for the science of psychology (Campbell, 1984; Fishman, 1988; Gergen, 1985; Howard, 1985; Koch, 1976, 1981; Krasner & Houts, 1984; Scarr, 1985). At the extreme, science is seen as nothing more than a construction of the human mind. As Scarr put it, "Facts do not have an independent existence" (p. 499). While total discussion of this issue is beyond the scope of this chapter, some general comments are necessary.

The belief that researchers are neutral, value-free truth seekers is probably based on the philosophical principle called logical positivism. Logical positivism is based on two assumptions: "(1) there is an external world independent of human experience, and (2) objective knowledge about this world can be obtained through direct sense experience" (Fishman, 1985, p. 5). Thus, logical positivism suggests that all observations are the ultimate forms of truth since observations are objective (Staats, 1988).

An alternative to logical positivism has been labeled by Gergen (1985) as social constructionism. This model suggests that all knowledge, including scientific knowledge, is a construction of the mind:

> Sensory data are filtered through the knowing apparatus of the human senses and made into perceptions and cognitions. The human mind is also constructed in a social context, and its knowledge is in part created by the social and cultural world. Knowledge of the world is therefore always constructed by the human mind in the working models of reality in the sciences. If this is not evident, consider for a moment the vast differences in our concepts of the world before Galileo, Darwin, Einstein, and Freud. (p. 499)

While there is no doubt that data are a product of the observed and the observer, science can, and will, continue to make major contributions to our understanding (no matter how limited it may be) of behavior.

Hayes (1995) offers an alternative definition of science:

> Science is a human endeavor that has as its purpose the development of increasingly integrated systems of verbal rules that allow us to accomplish analytic goals with precision, scope, and depth, based on verifiable experience ... Science is a special kind of word-producing, rule-making enterprise. (p. 52)

Hayes's definition is consistent with Kuhn (1970) in that both see science as the accumulation of information designed towards a better understanding of the world. Unlike other perspectives, however, Hayes suggests that there are aspects of the world that can be studied and measured which are beyond the scope of rule-governed science.

1.02.3.2 Science vs. Art

As mentioned previously, science must be understood in terms of both the factual theories and method *and* the behavioral experiences of the scientist. Facts are more likely to be learned from reading and studying a subject but experiences are a critical part of any learning endeavor.

It is widely accepted that behavior is influenced by rules and experiences (Hayes, 1989; Haye's, Zettle, & Rosenfarb, 1989). These different influences are often labeled as "rule-governed behavior" (Hayes, 1989, 1995) and "contingency-shaped behavior" (Hayes, 1995). Another way of expressing these ideas might be the difference between what is often thought of as didactic "academic" information vs. practical "experience." In clinical psychology, this translates roughly into what one learns from reading the literature or going to class compared with what one learns from supervised clinical and research opportunities.

There are important differences between those behaviors acquired through rule-governed processes and behavior acquired via contingency-shaping:

> Rule-governed behavior tends to be somewhat more rigid, less modifiable by its direct consequences, more precise in its initial forms, and more subject to arbitrary social contingencies. Contingency-shaped behavior is generally more moldable and modifiable, but it is also more variable, and subject to chance contingencies. (Hayes, 1995, p. 52)

Science focuses primarily on rule-governed behaviors. However, as Hayes (1995) points out, while "science is the best rule-generating institution ever invented, it is wrong to think that all behavior can be *directly* rule-governed . . . Any instance of rule-governed behavior stands ultimately on contingency-shaped behavior, at least to a degree" (pp. 52–53). Thus, experience plays a critical role in science.

An example might be helpful. Articulation of rules that govern behavior is the first step in the development of science and critical to the understanding of the status of a scientific discipline. If we were to develop a science of baseball, we might first want to articulate rules to describe how a person should swing a bat to hit a ball approaching at speeds of nearly 140 kilometers per hour. The rules might include a discussion on stance, placement of the bat, choice of the type of bat, etc. This verbal set of descriptors might be quite elaborate and, combined with some video clips of someone engaged in "batting," we may have the fundamentals for the science of baseball. However, this "science of baseball" will never be an exact replica of baseball experienced by the professional, unless it incorporates the contingency-shaped skills learned through the interaction of the ballplayer and his or her experiences at playing ball.

Likewise, an understanding of the science of clinical psychology requires an understanding of the known rules of clinical psychology, combined with experiences that allow contingency-shaped learning. While the study of science involves both components, it is difficult to document those critical experiential aspects of science training. Most often we focus on issues such as mentoring and directed research as a mechanism to acquire contingency-shaped aspects of science.

There has been a push to evaluate the critical importance of hands-on training in the laboratory for teaching science. For example, Ertepinar and Geban (1996) demonstrated that the addition of a laboratory-based method which focuses on conducting science studies resulted in a greater mastery of science knowledge than merely providing students with science information. Thus, it appears that active involvement in behaviors relevant to science may be an important part of science training.

1.02.3.3 Clinical Scientists and Clinical Practitioners

As mentioned previously, clinical psychologists engage in many different activities with two of the most identifiable being the science and practice roles. Psychologists who emphasize clinical research can be identified as clinical scientists, while individuals who emphasize clinical practice are often described as clinical practitioners. While it is clear that these roles are arbitrary and that science and practice often go hand in hand, it is also clear that some training programs emphasize training for clinical science careers and others emphasize training for clinical practice careers. Likewise, many clinical psychologists engage in both scientific and practice careers simultaneously; however, there is much to be gained from differentiating clinical science training from clinical practice training.

Clinical scientists and clinical practitioners need to be well grounded in the knowledge base of clinical psychology (this will be discussed in greater detail in Section 1.02.4). Such training would allow both types of trainees comparable exposure to the rules of the science of clinical psychology. In addition, basic contingency-shaped experiences in clinical research and clinical practice are critical for both career paths. To become a sound clinical researcher, basic training in clinical practice is critical. Likewise, sound clinical practitioners must have basic experience in clinical research. However, the types of contingency-shaped experiences and the amount of contingency-shaped experiences may be quite different for success as a researcher or practitioner.

The ideal method for gaining contingency-shaped experiences in research would be through supervised research with an experienced mentor who can guide and direct the activity. Such experiences would provide contingency-shaped behaviors critical for science practice in clinical psychology that are unique from the rule-governed science behaviors. These activities are typically formalized in academic training programs in terms of thesis and dissertation requirements; however, successful completion of a dissertation is probably minimal for both clinical scientists and clinical practitioners and more extensive supervised research experiences (such as one or two years of postdoctoral experience with a senior scientist) would enhance the development of contingency-shaped research skills critical for a career as a clinical scientist.

Practice training requires extensive contact with clinical populations. The ideal is supervised clinical work with an experienced professional who can guide and direct the activity. Like research experience, clinical experiences will provide contingency-shaped behaviors critical for the practice of clinical psychology that are unique to the contingency-based research behaviors or the rule-governed aspects of clinical psychology. Academic programs require a minimal amount of clinical practice and a-year of full-time clinical training (internship). However, these experiences are minimal basic training. As with research training, postdoctoral experiences in a clinical setting may be critical for the development of sophisticated contingency-shaped practice skills.

At first glance, this may seem to be merely an attempt to redefine the existing model of clinical psychology training using new jargon. However, this redefinition can greatly influence the manner in which science and practice are taught and may provide a useful model for the integration of science and practice. It should come as no surprise that there will be a great deal of variability with respect to contingency-shaped behaviors. By their definition, these experiences are somewhat unique. Unlike rules, where everyone can be taught the same concepts, contingency-shaped behavior will be much more variable and likely to be influenced by chance contingencies (Hayes, 1995). However, there are certain activities that will enhance training in science or practice. The development of contingency-shaped practice skills can be enhanced through an emphasis on practical training clinical service delivery with a broader range of clients and treatment methods. Science training can be enhanced with an emphasis on practical research training, again with a broader range of methods and problems. While time constraints in a training program will limit the amount of experiential training in both the practice and science aspects of clinical psychology, the time spent engaged in an activity does not guarantee more or better contingency-shaped behavior. It is likely that the quality of the supervised experience (either in research or practice) will have a greater influence on outcomes.

Given that an emphasis on training students for roles as clinical scientists or clinical practitioners may result in unique behavioral repertoires owing to differential rule-governed and contingency-shaped behaviors, it should come as no surprise that both career paths offer unique and important contributions to the science of clinical psychology. A recognition of the different contributions may be critical to our advancement of the science of clinical psychology, but beyond the scope of this chapter. However, several papers (e.g., Clement, 1988, 1996; Davison & Lazarus, 1995; Fensterheim, 1993; Fensterheim & Raw, 1996; Kanfer, 1990; Marten & Heimberg, 1995; Nezu, 1996; Woolfolk & Lazarus, 1979) will provide an overview of this area for interested readers. An understanding of these unique contributions requires better identification of the rule-governed and contingency-shaped aspects of clinical psychology.

Much of the controversy over the science and practice of clinical psychology comes from a failure to recognize the importance of contingency-shaped behaviors and the unique perspective gained from mentored research and mentored clinical training. It has been suggested that if the split between the science and practice of clinical psychology is not resolved, we may see a dissolution between the science and practice of clinical psychology (Rice, 1997). Failure to resolve this split might result in clinical training being removed from academic psychology departments and relegated exclusively to professional schools. In this scenario, Rice has predicted that clinical health psychology would become an allied health profession, rehabilitation psychology would ally itself with occupational therapy, and counseling psychology would evolve into "spiritual counseling." Academic departments would see a decline in graduate students and the blending of traditional psychology departments with other academic units such as neuroscience, human development, computer science, and social science departments.

Rice's (1997) dissolution scenario is by no means the future of psychology, but a probable outcome if science and practice are not emphasized in clinical psychology training. Later in this chapter, Rice's "science-based practice" scenario will be discussed as the ideal outcome for clinical psychology. First, however, the current status of science in clinical psychology is reviewed with an emphasis on identifiable rule-governed and contingency-shaped aspects of the science of clinical psychology.

1.02.4 THE CURRENT STATUS OF SCIENCE IN CLINICAL PSYCHOLOGY

1.02.4.1 Diagnosis and Assessment

The publication of the third edition of the *Diagnostic and statistical manual of mental disorders* (*DSM-III*; American Psychiatric Association, 1980) changed the manner and approach to diagnosis and classification used in clinical psychology. The system was widely adopted (although not uniformly supported, see Kirk & Kutchins, 1992, for a review) by all mental health professions. *DSM-III* (and its successors *DSM-III-R*, American Psychiatric Association, 1987; *DSM-IV*, American Psychiatric Association, 1994) has become the standard classification system for clinical psychology and is represented in almost every major textbook on the topic (Klerman, 1984).

The shift from *DSM-II* (American Psychiatric Association, 1968) to *DSM-III* was profound. *DSM-II* was heavily influenced by psychodynamic theory and used concepts such as neurosis and psychosis as major classifications. *DSM-III* attempted to develop a more empirical, and to some, more scientific classification. This change had great influences on both the science of clinical psychology and psychiatry. Maxmen (1986) described this shift as:

> Psychoanalytic psychiatry bases truth on authority; something is true because Freud said so. Scientific psychiatry bases truth on scientific experimentation ... The old psychiatry derives from theory, the new psychiatry derives from fact. (p. 31)

Changes in the *DSM* (from the third to the fourth edition) have continued to emphasize empiricism and science. Categories were changed (or added) only if there were sufficient data to warrant such a change. Symptoms of disorders likewise were modified, only with considerable evidence that the existing criteria were inadequate and that new criteria were more accurate.

While there is a strong scientific emphasis in the *DSM* system, it is not without its critics (cf., Kaplan, 1983; Kirk & Kutchins, 1992; Miller, Bergstrom, Cross, & Grube, 1981; Schacht, 1985; Singerman, 1981). As with any scientific endeavor, there are some assumptions in which science may have appeared to play little or no role. Spitzer (1985) identified several nonscience aspects involved in the development of the *DSM*, two of which will be summarized here. First, the *DSM* is based on a descriptive approach which emphasized classification on the basis of shared clinical features rather than presumed etiology. Science has not shown that descriptive approaches are superior to other approaches. In fact, there is no compelling reason to believe that a similar disorder would not show up in different individuals as different symptoms. The *DSM* deals with these differences by identification of a list of descriptive criteria, and the clinician merely tallies the number of symptoms presented by the client. A diagnosis is given if the number of symptoms meets or exceeds a predetermined criteria.

Second, Spitzer points out that subjective judgments played a major role in the development of the *DSM-III*. As noted in the introduction to *DSM-III*, "most of the diagnostic criteria are based on clinical judgment, and have not yet been fully validated by data." While the revisions have continued to search for empirical support and less subjective data, clinical judgment continues to play a major role in this diagnostic system.

For the scientist–practitioner and the clinical scientist, the *DSM-III* and its successors provide a significant step towards the development of a strong science. The rules are clearer than other diagnostic systems; however, the system continues to utilize clinical judgments (contingency-shaped behaviors) necessary given the existing knowledge base. As the field develops and more is learned about specific disorders, the *DSM* system allows for correction and changes in the diagnostic rules. It is anticipated that changes will occur and mechanisms have been built in to facilitate and foster change.

The scientific status of psychological assessment is difficult to evaluate. On the one hand, with regard to testing and test development, there are "180 guidelines covering 16 categories pertaining to the activities of test developers, test users, and test takers" (Cone, 1995, p. 204). However, there is considerable debate as to the best use of tests in clinical psychology (cf., Collins & Thompson, 1993; Cone, 1989; Hayes, Nelson, & Jarrett, 1987, 1989; Haynes & Uchigakiuchi, 1993).

With respect to guidelines, the majority focus on (i) technical standards for test construction and evaluation, such as validity, reliability, norming, etc.; (ii) professional standards for test use, such as employment testing, educational testing, etc.; (iii) standards for particular applications, such as testing of individuals for whom English is a second language, visual impairments, etc.; and (iv) standards for administrative procedures, such as scoring and reporting issues, and the rights of test-takers (Cone, 1995).

While the existing standards are important and useful to the science of psychology, Cone (1995) recommends that new standards be developed to improve the utility of psychological assessment methods. Cone's detailed rationale for each of these recommendations will not be presented in this chapter; however, the new standards should:

> (1) Recognize that different subject matters might require modifications in standards.
> (2) Make room for the concept of accuracy, and distinguish it from reliability and validity.
> (3) Distinguish between representational and elaborative validity, and require purveyors of new measures to show the former and some evidence of the latter before publishing the measure.
> (4) Require psychologists to evaluate what they do.
> (5) Develop standards dealing with the process of that evaluation.
> (6) Consider specific standards dealing with functional analysis.
> (7) Include standards focusing on idiographic or local use assessment activities. (Cone, 1995, p. 220)

One of the most important steps for improving the science in clinical assessment is to get practicing psychologists to evaluate what they do. While everyone agrees that evaluation is important, it is difficult to get consensus on the systematic application of psychological methods and practices to clinical assessment (Kazdin, 1996a). Clement (1996) offers a model of practice assessment which focuses on diagnosis, treatment plans, compliance with treatment plans, and assessment of treatment outcome. While this plan may be criticized as relying heavily on self-report measures (Nezu, 1996), the spirit of the plan is an important step for improving the scientific status of assessment in clinical psychology (Hayes, 1996; Lambert & Brown, 1996).

Managed health care may directly influence the use of sound clinical assessment strategies (Nelson-Gray, 1996). The current trend is to develop outcome measures applicable to all clients (Nelson-Gray). While the development of uniform measures may be useful, measures unique to diagnostic categories may be far more helpful because of the differences in success rates with different clinical problems (Nelson-Gray). For example, treatment of social phobia may be best evaluated using the Index of Social Phobia Improvement (Turner, Beidel, & Wolf, 1994); depressed individuals might best be evaluated using the Beck Depression Inventory (Beck, Ward, Mendelsohn, Mock, & Erbaugh, 1961), and anxiety disorders might be evaluated using the State-Trait Anxiety Inventory (Spielberger, Gorusch, & Lushene, 1970). For many situations, however, individualized (or idiographic) assessment methods may be needed (Collins & Thompson, 1993; Nelson-Gray, 1996).

1.02.4.2 Scientific Status of Clinical Assessment

A great deal is known about the psychometric properties of tests (Cone, 1988). Clinical psychology has developed sophisticated measures of intelligence (Kaufman & Harrison, 1991) and personality (Ben-Porath & Butcher, 1991). In general, we have a fairly sophisticated, sound set of rule-governed behaviors as a foundation for the science of clinical psychology. The use of assessment in clinical settings (contingency-shaped experiences) is less clear cut. Since the mid-1970s, there has been a noticeable decline in emphasis on clinical assessment (Garfield & Kurtz, 1973, 1976; Levitt, 1973; Shemberg & Keeley, 1970).

One factor that has contributed to the decreased use of clinical assessment in the field is that clinicians have not always found these data to be useful in treatment (Adams, 1972; Moore, Boblitt, & Wildman, 1968). The behavioral alternative to traditional assessment, behavioral assessment, was developed in part as a reaction to more traditional assessment approaches (Hersen, 1976). More recently, even among behavioral psychologists, reliance on behavioral observations for assessment has given way to more moderate approaches involving the use of psychometrically sound trait measures (Collins & Thompson, 1993; Gross, 1990; Watkins, Campbell, & McGregor, 1990).

Cone (1988) points out that there are two "contrasting models of behavioral assessment" (p. 46): a nomothetic-trait approach and an idiographic-behavioral approach. The nomothetic approach enables the comparison of an individual's score on an assessment measure to a normative reference group. The idiographic-behavioral approach focuses on changes over time for a single individual. Idiographic assessment provides the scientist–practitioner with useful methods for increasing the quality of the science in his or her practice (Barlow, Hayes, & Nelson, 1984).

While the integration of idiographic and nomothetic systems is strongly encouraged today, there is little evidence about how to accomplish this integration. Hersen and Bellack (1988) have advocated an "idiographic approach within the normal nomothetic system" (p. 78). Collins and Thompson (1993) suggest that an equally useful alternative is the use of "nomothetic (personality) data [integrated] *within* the idiographically based behavioral assessment" (p. 69).

Four broad guidelines have been recommended (cf., Collins & Thompson, 1993). First, assessments must use empirically sound measures. There are many unsound measures, frequently tied to particular theoretical constructs or part of "clinical lore." Not only must the assessment measures used have an empirical basis, the "methods must be *administered* and *interpreted* from an empirical data base" (p. 65). Second, assessments derived from specific measures must be confirmed. Most measures are open to multiple interpretations. Third, the behavior of interest must be differentiated from the assessment score. Assessment data are only useful because of their ability to *predict* behavior; however, the assessment data are not the behavior. Finally, assessment data must be collected repeatedly over time. Single measures are rarely useful in clinical settings.

1.02.4.3 Clinical Interventions

Our knowledge about the scientific status of clinical interventions has grown immensely. However, the conclusions drawn from these data are not consistent. On the one hand, there is a lot known about which types of therapy are likely to produce the greatest benefit in some clinical disorders (Kazdin, 1996b); on the other, there is continued evidence that most forms of psychotherapy are useful with little to no differences between the "type" of therapy and success (Seligman, 1995). These somewhat antithetical conclusions have been derived using different scientific methodologies and will be explained in some detail.

By 1980, many psychotherapy researchers (e.g., Bergin & Lambert, 1978; Luborsky, Singer, & Luborsky, 1975; Shapiro & Shapiro, 1982; Smith & Glass, 1977; Smith, Glass, & Miller, 1980; VandenBos & Pino, 1980; Yates & Newman, 1980) reached the general conclusion that "Psychotherapy, as a generic treatment process, [is] demonstrably more effective than no treatment" (VandenBos, 1986, p. 111). The primary reason for this has to do with what has been called "nonspecific factors" thought to be common to all forms of psychotherapy (cf., Arkowitz, 1995; Beutler, 1995; Elkin, 1995; Frank, 1995; Ilardi & Craighead, 1994; Luborsky, 1995; Weinberger, 1995). The years since 1985 have focused on comparative outcome research. That is, the emphasis was no longer on whether or not psychotherapy works, but what type of psychotherapy works with what disorders. In the mid-1980s, VandenBos, in the introduction to the special issues of the *American Psychologist* on Psychotherapy Research, appeared to encourage such research:

> It now appears that single-focus "outcome" (or efficacy) research should be a "thing of the past." Future psychotherapy outcome research should, at a minimum, be "comparative" outcome research—exploring the relative advantages and disadvantages of alternative treatment strategies for patients with different specific psychological and behavioral difficulties. (p. 111)

In 1995, the Task Force on the Promotion and Dissemination of Psychological Procedures published a list which contained "Examples of empirically validated treatments" (Table 3, p. 22). The publication of this list has resulted in immense praise and criticism and, regardless of one's perspective as to the accuracy of the list, the list is important for the evaluation of the scientific status of clinical intervention in psychology.

Many psychotherapy researchers agree with the general conclusions of the Task Force. For example, Seligman (1995) states,

> studies show ... that cognitive therapy, interpersonal therapy, and medications all provide moderate relief from unipolar depressive disorder; that exposure and clomipramine both relieve the symptoms of obsessive-compulsive disorder moderately well but that exposure has more lasting benefits; that cognitive therapy works very well in panic disorder; that systematic desensitization relieves specific phobias; that "applied tension" virtually cures blood and injury phobia; that aversion therapy produces only marginal improvement with sexual offenders; that disulfram (Antibuse) does not provide lasting relief from alcoholism, that flooding plus medication does better in the treatment of agoraphobia than either alone; and that cognitive therapy provides significant relief of bulimia, outperforming medications alone. (pp. 965–966)

Barlow (1996) adds,

> I believe, along with many others, that the evidence is incontrovertible that we have effective psychological interventions for a large number (but not all) of psychological disorders. Numerous studies and subsequent meta-analyses have demonstrated that any number of specific psychotherapeutic approaches, either alone or, in some cases, in combination with pharmacological approaches, are more effective and often longer lasting than credible alternative psychological interventions. (p. 216)

Given these strong endorsements, it may at first seem confusing to learn that some of the more influential psychotherapy researchers (e.g., Garfield, 1996; Havik & VandenBos, 1996; Seligman, 1995; Shapiro, 1996) believe that these research studies are of little importance to the scientific study of psychotherapy.

One of the major criticisms of comparative outcome studies is that the results have little relevance to clinical practice (Garfield, 1996; Havik & VandenBos, 1996; Munoz, Hollon, McGrath, Rehm, & VandenBos, 1994; Seligman, 1995). Participants seen in comparative outcome studies give informed consent to be randomly assigned to different treatment conditions, whereas patients seen in clinical practice do not. Comparative outcome studies frequently use manualized treatment protocols that place guidelines on therapist behaviors and the number of treatment sessions provided, whereas therapists in clinical practice rarely adhere to the type of manualized treatments used in comparative outcome studies (Persons, 1995) and almost never routinely limit the number of sessions.

Seligman (1995) argues that "effectiveness" psychotherapy studies, not comparative outcomes studies, are the best way of "finding out what treatments actually work in the field" (p. 966). To conduct a sound effectiveness study requires allowing a great deal of flexibility. For example, Seligman identifies five "properties that . . . characterize psychotherapy as it is done in the field" (p. 966). These factors will be discussed in some detail as they are critical for evaluation of the scientific status of clinical psychology.

First, psychotherapy cannot be defined in terms of number of sessions. In the field, it continues until there is some improvement or until the patient decides he or she should quit. This is extremely important since the most elaborate effectiveness study conducted to date, the *Consumer Reports* study (1995; Seligman, 1995), found duration of therapy to be positively related to improvement with the greatest improvement reported by individuals in therapy for two or more years.

Second, psychotherapy allows for correction. That is, if a technique or approach is not working for a client, then another approach is typically tried. This issue of self-correction is similar to medical treatments that are discontinued and/or replaced if the desired short-term outcome is not forthcoming. Comparative outcome studies are rarely designed with this in mind. Blanchard and his colleagues (Blanchard et al., 1996) used a self-correcting design in the study of biofeedback and hypertension. Treatment failures were then randomly assigned to an alternative therapy. However, I am not aware of similar designs in the psychotherapy literature.

Third, in the field, patients are often engaged in "active shopping." They learn about different forms of therapy and actively seek out someone who can provide that type or modality of therapy. In controlled outcome studies, random assignment results in a great deal of passivity on the part of the participant (Howard, Orlinsky, & Lueger, 1994).

Fourth, patients seen in the field often have more than one problem, or the problems that they present with do not fit diagnostic criteria. Many times the symptoms of their problems result in multiple diagnoses. See Lilienfeld, Waldman, and Israel (1994), Rutter (1994), Widiger and Ford-Black (1994) for a discussion of the problems with comorbid diagnosis in *DSM*. Comparative outcome studies, however, tend to focus on individuals with fairly specific inclusion criteria. While this is reasonable in terms of the specific data analytic models used in comparative outcome studies, it still makes generalization to field studies more difficult.

Finally, the end goal of field studies is "with *improvement in the general functioning* of patients, as well as amelioration of a disorder and relief of specific, presenting symptoms" (Seligman 1995, p. 967). Comparative outcome studies focus on specific symptom reduction.

1.02.4.4 Scientific Status of Clinical Interventions

Clearly there is a great deal of information available regarding rule-governed aspects of psychotherapy. Elaborate manuals have been developed to help therapists administer a variety of interventions ranging from relaxation training (e.g., Bernstein & Borkovec, 1973) to cognitive therapy (e.g., Beck, Rush, Shaw & Emery, 1979). Therapist manuals are available to help in the treatment of many disorders (e.g. *Panic and anxiety*, Barlow & Craske, 1994; Craske, Meadows, & Barlow, 1994; *Borderline personality disorder*, Linehan, 1993).

The development of manualized therapy has had a significant impact on the science of clinical psychology (Eifert, Schulte, Zvolensky, Lejuez, & Lau, 1997). The operationalization of specific aspects of therapy has contributed significantly to clinical outcome research (Barlow & Craske, 1994; Chambless, 1996; Seligman, 1995); however, the existence of a manual does not guarantee that all therapists will be equally effective, or that they will present the material in the same manner (Chambless). Likewise, manualized therapy does not guarantee that the behaviors identified in the therapy manual are identical to those used in the practice of clinical psychology (Collins & Thompson, 1988).

Most therapy manuals provide some caution that successful implementation of manualized treatment requires skilled therapists. For example, Barlow and Craske (1994) state:

> The first revolution in the development of effective psychosocial treatments during the past decade has been the manualization of these treatments. Because these are structured programs for specific disorders, they can be written in sufficient detail to allow trained therapists to administer them in roughly the same manner that they were proven effective. *This does not imply that therapeutic skills are no longer needed.* In fact, in the training tapes...there are numerous examples of the importance of highly trained psychotherapeutic skills as one proceeds with the program. (Chap. 2, p. 6, italics added)

Chambless (1996) concurs but expands on the usefulness of manualized therapy:

> I agree that manuals do not completely mirror clinical practice but argue that they have much to teach us if the treatment they depict is efficacious. I find that students learn treatment approaches much more quickly from their systematic depiction in manuals...than through supervision alone. (p. 231)

Using the model of rule-governed/contingency-shaped behaviors as a guide, it is clear that the professional use of treatment manuals will foster the acquisition of rule-governed behaviors. Likewise, clinical supervised practice using efficacious treatment manuals sets the occasion for continued contingency-shaped learning. Several writers (Iwamasa & Orsillo, 1997; Persons, 1995) have suggested that incorporating efficacious treatment manuals into a clinical training program is critical. Persons suggests that the main reason most psychotherapists do not conduct therapy consistent with the outcome literature is that they have received little or no training in the use of efficacious training manuals. Such training will enhance the scientific status of clinical interventions by directly shaping therapist behavior.

Flexibility in the use of treatment manuals is less clear. Allowing therapists to "freelance" or stray too far from the procedures outlined in the manuals may not be desired (Eifert et al., 1997). For example, Frank, Kupfer, Wagner, McEachran, and Cornes (1991) demonstrated a strong relationship between adherence and positive outcomes. More research is needed, and it is possible that some manualized treatments need to be closely adhered to, while others may allow for substantial therapist flexibility.

1.02.5 TRAINING SCIENCE IN CLINICAL PSYCHOLOGY: THE SCIENTIST–PRACTITIONER MODEL

The scientist–practitioner model was first articulated at the Boulder Conference (Raimy, 1950) and continues to be one of the most prominent, if not the primary, model for training clinical psychologists. Rice (1997) sees a "science-based practice" as a likely scenario for psychology as a profession, and in my opinion, the most optimistic scenario for psychology. The Gainesville Conference (Belar & Perry, 1992) articulated the components of this model quite well and a brief summary follows.

First, delegates at the Gainesville Conference "asserted that the scientist–practitioner model was essential for the ever-changing discipline of psychology" (Belar & Perry, 1992, p. 71). While there appears to be room for alternative models of training, this model is critical to the continued scientific development of clinical psychology.

Second, the hallmark of this model is the integrated training of science and practice, not the job title or position of the graduate. This difference is important. The usefulness of this training model cannot be evaluated by counting the number of graduates who publish scholarly work or who are employed as clinical scientists. Rather, training in the scientist–practitioner model should produce professionals who use the scientific knowledge base of clinical psychology in their practice and who approach the problems and questions of clinical practice in a scientific manner.

Third, the conference identified the importance of both didactic (rule-governed) and practical (contingency-shaped) training in the science of psychology and the strategies and tactics of applied work. Didactic science training includes competence in the major scientific areas of psychology identified as the biological bases of behavior, cognitive-affective bases of behavior, social bases of behavior, and individual differences. Didactic practice training emphasizes "the application of scientific thinking and behavior to problem-solving and hypothesis-testing in practice" (p. 73). Practical training must focus on the integration of science and practice and should include both predissertation research as well as the completion of a dissertation. The doctoral dissertation should provide an opportunity for the trainee to demonstrate "competence to carry out an original, independent scientific investigation that furthers psychological knowledge" (p. 74).

Clearly, the scientist–practitioner training model articulated by Belar and Perry (1992) is consistent with the model of science discussed throughout this chapter. Scientist–practitioner training requires contemporary didactic training in the knowledge base of clinical psychology. Such training should strengthen the acquisition of rule-governed behaviors critical for the science of clinical psychology. Likewise, practical experience is needed. However, such experiences should continue to integrate science knowledge in all aspects of practice. This includes research and clinical work. Thus, research training in clinical psychology must integrate the acquired knowledge of the science of psychology, with supervised experience conducting clinical research. Likewise, clinical practice must integrate the acquired knowledge of the science of psychology, with supervised experience in clinical assessment and psychotherapy.

While the scientist–practitioner model is typically applied to clinical psychology, Rice (1997) sees a future where all professional psychology is science based:

> In this scenario, all psychological practitioners, including counseling, school, and industrial/organizational (I/O) psychologists, are trained as science-based practitioners to work in a variety of settings. Most of the persons preparing for the practice attend professional schools, which are typically attached to universities and maintain close relations with academic psychology departments. These professional schools usually offer the PhD rather than the PsyD degree, though there is variation with respect to this.
>
> The professional school curriculum is rigorously and narrowly structured during the first two years; students take course work in cognitive psychology, neuropsychology, psychometic theory, developmental and social psychology, psychopharmacology, and research methods. During the next two years, students take more specialized courses to prepare them for work in various organizational settings. A two-year specialty internship completes the training. A research dissertation, usually of an applied nature, is required for all completing the PhD. Fewer students are admitted to these programs because of the training, which is rigorous and expensive due to the large number of full-time faculty who teach these basic courses and who engage in a range of applied psychology research ventures.
>
> Academic psychology departments continue the tradition of preparing teachers and researchers. Teaching positions are fairly abundant in both psychology departments and professional schools. Research positions are found outside academia in commercial and governmental settings. The traditional psychology fields such as social, experimental, developmental, and I/O psychology remain, but most departments offer training in health psychology, experimental psychopathology, or both. (pp. 1177–1178)

While some will cringe at the notion that training in applied psychology should occur in a professional school, this scenario is more consistent with current training goals and in no way minimizes the importance of science training. Likewise, much of the conflict between the science and practice of psychology can be minimized through this scenario.

1.02.6 SCIENCE, THE CHANGING HEALTH CARE SYSTEM, AND THE DEMAND FOR PSYCHOLOGY

Changes in health care appear inevitable (VandenBos, 1993). Two factors will influence the final outcome of this process: (i) cost—all else being equal, less expensive forms of treatment will be preferred; and (ii) effectiveness—for problems with known effective treatment, those treatments shown to be effective will be supported (Barlow, 1994). It is critical, therefore, to develop a strong case for scientific evaluation and implementation of clinical psychology services.

Detailed guidelines are already being developed. For example, the Agency for Health Care Policy and Research, which was commissioned by the federal government to determine the effectiveness of treatments for specific disorders, has published a two-volume document which provides recommendations for treating major depressive disorders (Barlow, 1994). The summarized findings of this evaluation was that medications are the first-line treatment for almost all cases of major depressive disorder. Exceptions to medication are when the depression is mild or when the patient specifically asks for psychosocial treatments (Barlow).

While clinical psychology continues to debate the usefulness of controlled outcome studies, it is clear that these types of studies are providing the major source for the analysis and ultimate development of public policy. Students in clinical psychology must be trained in the science of clinical psychology and the methods that currently appear most efficacious. It is expected that the specific methods will change (Iwamasa & Orsillo, 1997); however, sound professional science training will allow clinical psychology to adapt and develop.

One often overlooked aspect of clinical psychology is the extent to which practice activities are the exclusive domain of clinical psychology, or also in the domain of other health care providers. Clinical psychology as a discipline requires the doctoral degree for entry into the profession; however, many of the professional activities historically identified with clinical psychology are today provided by other professionals, often with a terminal master's degree. The professionals are often licensed by states as counselors, social workers, marriage and family therapists, and in a few states as psychologists.

Given the change in health care that emphasizes less expensive forms of treatment as preferred, it is not surprising that many state agencies and health organizations frequently employ master's training professionals to provide counseling and psychotherapy services. This change, combined with the large number of doctoral clinical psychologists graduating each year, has led to a concern with the current supply and future demand for clinical psychologists (cf., Pion, 1991; Robiner, 1991a, 1991b; VandenBos, DeLeon, & Belar, 1991). While it is not clear how "many" psychologists are needed, it is clear that psychology must identify those aspects of training that make psychologists unique from master's level practitioners. Rigorous training in science that includes rule-governed and contingency-based training may prove to be the critical difference.

Clinical scientists will be expected to show leadership in research on effective clinical assessment and intervention approaches. Clinical practitioners must incorporate the knowledge base of clinical psychology in their practice and identify strengths and weaknesses of proposed efficacious treatments. This dialog between clinical scientists and clinical practitioners is critical for the continued development of science in clinical psychology.

1.02.7 CONCLUSIONS

The hallmark of clinical psychology is the emphasis on science. The scientific status of clinical psychology is strong. Models of science are consistent with the primary training model of clinical psychology (the scientist–practitioner model) which is flexible enough to allow clinical psychologists to focus on the science (clinical scientists) and practice (clinical practitioners) of clinical psychology.

Clinical assessment has a strong didactic knowledge base. While clinical psychologists are not using clinical assessment strategies as frequently as in the past, changes in the way psychological assessment is conceptualized and used appear promising. However, it is likely that clinical assessment and clinical intervention will become more integrated and the need to differentiate between assessment and intervention may no longer be critical.

Much is known about the most effective treatments for some psychological problems. While the development of effective treatments is contested by some, the proliferation of treatment manuals and potential changes in public policy should serve to strengthen the science base of clinical intervention.

In summary, clinical psychology retains a strong commitment to science and science models predominate all aspects of clinical psychology. There is room for growth and new developments in the field; however, the current status of clinical psychology will serve as a strong foundation.

1.02.8 REFERENCES

Ackermann, R. (1970). *Philosophy of science: An introduction*. New York: Pegasus.

Adams, J. (1972). The contribution of the psychological evaluation to psychiatric diagnosis. *Journal of Personality Assessment, 36*, 561–566.

American Psychiatric Association (1968). *Diagnostic and statistical manual of mental disorders* (2nd ed.). Washington, DC: Author.

American Psychiatric Association (1980). Diagnostic and statistical manual of mental disorders (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., Rev.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychological Association (1992). *Ethical principles of psychologists and code of conduct*. Washington, DC: Author.

Arkowitz, H. (1995). Common factors or processes of change in psychotherapy? *Clinical Psychology: Science and Practice, 2*, 94–100.

Barlow, D. H. (1994). Psychological interventions in the era of managed competition. *Clinical Psychology: Science and Practice, 1*, 109–122.

Barlow, D. H., & Craske, M. G. (1994). *Mastery of your anxiety and panic: II*. Albany, NY: Graywind Publications.

Barlow, D. H. (1996). The effectiveness of psychotherapy: Science and policy. *Clinical Psychology: Science and Practice, 3*, 236–240.

Barlow, D. H., Hayes, S. C., & Nelson, R. O. (1984). *The scientist practitioner: Research and accountability in clinical and educational settings*. New York: Pergamon.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression: A treatment manual*. New York: Guilford Press.

Beck, A. T., Ward, C. H., Mendelsohn, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561–571.

Belar, C. D., Bieliauskas, L. A., Klepac, R. K., Larsen, K. G., Stigall, T. T., & Zimet, C. N. (1993). National conference on postdoctoral training in professional psychology. *American Psychologist, 48*, 1284–1289.

Belar, C. D., & Perry, N. W. (1992). National conference on scientist–practitioner education and training for the professional practice of psychology. *American Psychologist, 47*, 71–75.

Ben-Porath, Y. S., & Butcher, J. N. (1991). The historical development of personality assessment. In C. E. Walker (Ed.), *Clinical psychology: Historical and research foundations* (pp. 121–156). New York: Plenum.

Bergin, A. E., & Lambert, M. J. (1978). The evaluation of therapeutic outcomes. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis* (2nd ed., pp. 139–190). New York: Wiley.

Bernstein, D. A., & Borkovec, T. D. (1973). *Progressive relaxation training: A manual for the helping professions*. Champaign, IL: Research Press.

Beutler, L. E. (1995). Common factors and specific effects. *Clinical Psychology: Science and Practice, 2*, 79–82.

Blanchard, E. B., Eisele, G., Vollmer, A., Payne, A., et al. (1996). Controlled evaluation of thermal biofeedback in treatment of elevated blood pressure in unmedicated mild hypertension. *Biofeedback and Self-Regulation, 21*, 167–190.

Boring, E.G. (1950). *A history of experimental psychology* (2nd ed.). New York: Appleton-Century-Crofts.

Brems, C., Thevenin, D. M., & Routh, D. K. (1991). The history of clinical psychology. In C. E. Walker (Ed.), *Clinical psychology: Historical and research foundations* (pp. 3–35). New York: Plenum.

Campbell, D. T. (1984). Can we be scientific in applied science? In R. F. Conner, D. G. Altman, & C. Jackson (Eds.), *Evaluation studies review annual* (Vol. 9, pp. 26–48). Beverly Hills, CA: Sage.

Capshew, J. H., & Hilgard, E. R. (1992). The power of service: World War II and professional reform in the American Psychological Association. In R. B. Evans, V. Staudt-Sexton, & T. C. Cadwallader (Eds.), *The American Psychological Association: A historical perspective* (pp. 149–176). Washington, DC: American Psychological Association.

Chambless, D. L. (1996). In defense of dissemination of empirically supported psychological interventions. *Clinical Psychology: Science and Practice, 3*, 230–235.

Clement, P. W. (1988). Professional role modeling by faculty: The neglected child in doctoral training. *Professional Psychology: Research and Practice, 19*, 253–254.

Clement, P. W. (1996). Evaluation in private practice. *Clinical Psychology: Science and Practice, 3*, 146–159.

Collins, F. L., Foster, S. L., & Berler, E. S. (1986). Clinical training issues for behavioral psychology. *Professional Psychology: Research and Practice, 17*, 301–307.

Collins, F. L., & Thompson, J. K. (1988). On the use of symbolic labels in psychotherapy outcome research: Comment on Wills, Faitler, and Snyder. *Journal of Consulting and Clinical Psychology, 56*, 932–933.

Collins, F. L., & Thompson, J. K. (1993). The integration of empirically-derived personality assessment data into a behavioral conceptualization and treatment plan: Rationale, guidelines, and caveats. *Behavior Modification, 17*, 58–71.

Cone, J. D. (1988). Psychometric considerations and the multiple models of behavioral assessment. In A. S. Bellack & M. Hersen (Eds.), *Behavioral Assessment: A practical handbook* (3rd ed., pp. 42–66). New York: Pergamon.

Cone, J. D. (1989). Is there utility for treatment utility? *American Psychologist, 44*, 1241–1242.

Cone, J. D. (1995). Assessment practice standards. In S. C. Hayes, V. M. Follette, R. M. Dawes, & K. E. Grady (Eds.), *Scientific standards of psychological practice: Issues and recommendations* (pp. 201–224). Reno, NV: Context Press.

Consumer Reports (1995, November). Mental health: Does therapy help? Vol. 60, pp. 734–739.

Craske, M. G., Meadows, E. A., & Barlow, D. H. (1994). *Therapist's guide for the mastery of your anxiety and panic II: Agoraphobia supplement*. Albany, NY: Graywind Publications.

Davison, G. C., & Lazarus, A. A. (1995). The dialectics of science and practice. In S. C. Hayes, V. M. Follette, R. M. Dawes, & K. E. Grady (Eds.), *Scientific standards of psychological practice: Issues and recommendations* (pp. 95–120). Reno, NV: Context Press.

Derner, G. F. (1965). Graduate education in clinical psychology. In B. B. Wolman (Ed.), *Handbook of clinical psychology* (pp. 1403–1414). New York: McGraw-Hill.

Edelstein, B. A., & Brasted, W. (1983). Clinical training. In M. Hersen, A. E. Kazdin, & A. S. Bellack (Eds.), *Clinical psychology handbook* (pp. 35–56). New York: Pergamon.

Eifert, G. H., Schulte, D., Zvolensky, M. J., Lejuez, C. W., & Lau, A. W. (1997). Manualized behavior therapy: Merits and challenges. *Behavior Therapy, 28,* 499–509.

Elkin, I. (1995). Further differentiation of common factors. *Clinical Psychology: Science and Practice, 2*, 75–78.

Ertepinar, H., & Geban, O. (1996). Effect of instruction supplied with the investigative-oriented laboratory approach on achievement in a science course. *Educational Research, 38*, 333–341.

Fensterheim, H. (1993). Letter to the editor. *The Behavior Therapist, 16*, 149.

Fensterheim, H., & Raw, S. D. (1996). Psychotherapy research is not psychotherapy practice. *Clinical Psychology: Science and Practice, 3*, 168–171.

Fishman, D. B. (1985). Where the underlying boundaries are: Organizing psychology by paradigm analysis. *International Newsletter of Uninomic Psychology, 2*, 4–9.

Fishman, D. B. (1988). Pragmatic behaviorism: Saving and nurturing the baby. In D. B. Fishman, F. Rotgers, & C. M. Franks (Eds.), *Paradigms in behavior therapy: Present and promise* (pp. 254–293). New York: Springer.

Foster, S. L., Berler, E. S., & Collins, F. L. (1982). Conference report: The West Virginia Working Conference on Behavioral Training in Clinical Psychology. *The Behavior Therapist, 5*, 184.

Frank, E., Kupfer, D. J., Wagner, E. F., McEachran, A. B., & Cornes, C. (1991). Efficacy of interpersonal psychotherapy as a maintenance treatment of recurrent depression: Contributing factors. *Archives of General Psychiatry, 48*, 1053–1059.

Frank, J. D. (1995). Psychotherapy as rhetoric: Some implications. *Clinical Psychology: Science and Practice, 2*, 90–93.

Garfield, S. L. (1996). Some problems associated with "validated" forms of psychotherapy. *Clinical Psychology: Science and Practice, 3*, 218–229.

Garfield, S. L., & Kurtz, R. (1973). Attitudes toward training in diagnostic testing: A survey of directors on internship training. *Journal of Consulting and Clinical Psychology, 40*, 350–355.

Garfield, S. L., & Kurtz, R. (1976). Clinical psychologists in the 1970s. *American Psychologist, 31*, 1–9.

Gergen, K. J. (1985). The social constructionist movement in modern psychology. *American Psychologist, 40*, 266–275.

Goodwin, C. J. (1983). On the origins of Titchener's Experimentalists. *Journal of the History of the Behavioral Sciences, 21*, 383–389.

Gross, A. M. (1990). An analysis of measures and design strategies in research in *Behavior Therapy*: Is it still behavioral? *The Behavior Therapist, 13*, 203–209.

Habermas, J. (1975). *Legitimization crisis*. Boston: Beacon Press.

Havik, O. E., & VandenBos, G. R. (1996). Limitations of manualized psychotherapy for everyday clinical practice. *Clinical Psychology: Science and Practice, 3*, 264–267.

Hayes, S. C. (Ed.) (1989). *Rule-governed behavior: Cognition: contingencies, and instructional control*. New York: Plenum.

Hayes, S. C. (1995). What do we want from scientific standards of psychological practice? In S. C. Hayes, V. M. Follette, R. M. Dawes, & K. E. Grady (Eds.), *Scientific standards of psychological practice: Issues and recommendations* (pp. 49–66). Reno, NV: Context Press.

Hayes, S. C. (1996). Creating the empirical clinician. *Clinical Psychology: Science and Practice, 3*, 179–181.

Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1987). The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist, 42*, 963–974.

Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1989). The applicability of treatment utility. *American Psychologist, 44*, 1242–1243.

Hayes, S. C., Zettle, R. D., & Rosenfarb, I. (1989). Rule following. In S. C. Hayes (Ed.) *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 191–220). New York: Plenum.

Haynes, S. N., & Uchigakiuchi, P. (1993). Incorporating personality trait measures in behavioral assessment: Nuts in a fruitcake or raisins in a Mai Tai? *Behavior Modification, 17*, 72–92.

Hersen, M. (1976). Historical perspectives in behavioral assessment. In M. Hersen & A. S. Bellack (Eds.), *Behavioral assessment: A practical handbook* (pp. 3–22). New York: Pergamon.

Hersen, M., & Bellack, A. S. (1988). *DSM-III* and behavioral assessment. In A. S. Bellack & M. Hersen (Eds.), *Behavioral assessment: A practical handbook* (3rd ed., pp. 67–84). New York: Pergamon.

Hoch, E. L., Ross, A. O., & Skodak, M. (Eds.) (1966). *Professional preparation of clinical psychologists*. Washington, DC: American Psychological Association.

Holton, G. (1973). *Thematic origins of scientific thought: Kepler to Einstein*. Cambridge, MA: Harvard University Press.

Howard, G. S. (1985). The role of values in the science of psychology. *American Psychologist, 40*, 255–265.

Howard, K., Orlinsky, D., & Lueger, R. (1994). Clinically relevant outcome research in individual psychotherapy. *British Journal of Psychiatry, 165*, 4–8.

Ilardi, S. S., & Craighead, W. E. (1994). The role of nonspecific factors in cognitive-behavior therapy for depression. *Clinical Psychology: Science and Practice, 1*, 138–156.

Iwamasa, G. Y. & Orsillo, S. M. (1997). Individualizing treatment manuals as a challenge for the next generation. *Behavior Therapy, 28, 499–509.*

Kanfer, F. H. (1990). The scientist–practitioner connection: A bridge in need of constant attention. *Professional Psychology: Research and Practice*, *21*, 264–270.

Kaplan, M. (1983). A woman's view of *DSM-III*. *American Psychologist*, *38*, 786–792.

Kaufman, A. S., & Harrison, P. L. (1991). Individual intellectual assessment. In C. E. Walker (Ed.), *Clinical psychology: Historical and research foundations* (pp. 91–119). New York: Plenum.

Kazdin, A. E. (1994). Clinical psychology: Science and practice. *Clinical Psychology: Science and Practice*, *1*, 1–2.

Kazdin, A. E. (1996a). Evaluation in clinical practice: Introduction to the series. *Clinical Psychology: Science and Practice*, *3*, 146–159.

Kazdin, A. E. (1996b). Validated treatments: Introduction to the series. *Clinical Psychology: Science and Practice*, *3*, 216–217.

Kirk, S. A., & Kutchins, H. (1992). *The selling of DSM: The rhetoric of science in psychiatry*. New York: Aldine de Gruyter.

Klerman, G. L. (1984). The advantages of *DSM-III*. *American Journal of Psychiatry*, *141*, 539–542.

Koch, S. (1976). Language communities, search cells, and the psychological studies. In J. K. Cole & W. J. Arnold (Eds.), *Nebraska symposium on motivation 1975: Conceptual foundations of psychology* (pp. 477–559). Lincoln, NE: University of Nebraska Press.

Koch, S. (1981). The nature and limits of psychological knowledge: Lessons of a century qua "science." *American Psychologist*, *36*, 257–269.

Korman, M. (Ed.), (1976). *Levels and patterns of professional training in psychology*. Washington, DC: American Psychological Association.

Krasner, L., & Houts, A. C. (1984). A study of the "value" systems of behavioral scientists. *American Psychologist*, *39*, 840–850.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

Lambert, M. J., & Brown, G. S. (1996). Data-based management for tracking outcome in private practice. *Clinical Psychology: Science and Practice*, *3*, 172–178.

Levitt, E. E. (1973). Internship versus campus: 1964 and 1971. *Professional Psychology*, *4*, 129–132.

Lilienfeld, S. O., Waldman, I. D., & Israel, A. C. (1994). A critical examination of the use of the term and concept of comorbidity in psychopathology research. *Clinical Psychology: Science and Practice*, *1*, 71–83.

Linehan, M. M. (1993). *Skills training manual for treating borderline personality disorder*. New York: Guilford Press.

Luborsky, L. (1995). Are common factors across different psychotherapies the main explanation for the Dodo bird verdict that "Everyone has won so all shall have prizes"? *Clinical Psychology: Science and Practice, 2*, 106–109.

Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies. *Archives of General Psychiatry, 3*, 995–1008.

Marten, P. A., & Heimberg, R. G. (1995). Toward an integration of independent practice and clinical research. *Professional Psychology: Research and Practice, 26*, 48–53.

Maxmen, J. (1986). *The new psychiatry*. New York: New American Library.

Miller, L. S., Bergstrom, D. A., Cross, H. J., & Grube, J. W. (1981). Opinions and use of the DSM system by practicing psychologists. *Professional Psychology, 12*, 385–390.

Moore, G. H., Bobitt, W. E., & Wildman, R. W. (1968). Psychiatric impressions of psychological reports. *Journal of Clinical Psychology, 24*, 373–376.

Munoz, R. F., Hollon, S. D., McGrath, E., Rehm, L., & VandenBos, G. R. (1994). On the AHCPR depression in primary care guidelines. *American Psychologist, 49*, 42–61.

Nelson-Gray,R. O. (1996). Treatment outcome measures: Nomothetic or idiographic? *Clinical Psychology: Science and Practice, 3*, 164–167.

Nezu, A. M. (1996). What are we doing to our patients and should we care if anyone else knows? *Clinical Psychology: Science and Practice, 3*, 160–163.

Norton-Ford, J. D. (1982). Therapy outcome research methods. In P. C. Kendall & J. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 429–460). New York: Wiley.

Persons, J. B. (1995). Why practicing psychologists are slow to adopt empirically-validated treatments. In S. C. Hayes, V. M. Follette, R. M. Dawes, & K. E. Grady (Eds.), *Scientific standards of psychological practice: Issues and recommendations* (pp. 141–157). Reno, NV: Context Press.

Peterson, D. R. (1976). Is psychology a profession? *American Psychologist, 31*, 572–581.

Pion, G. M. (1991). A national human resources agenda for psychology: The need for a broader perspective. *Professional Psychology: Research and Practice, 22*, 449–455.

Poffenberger, A. T. (1938). The training of a clinical psychologist. *Journal of Consulting Psychology, 1*, 1–6.

Polanyi, M. (1958). *Personal knowledge: Towards a post-critical philosophy*. Chicago: University of Chicago Press.

Raimy, J. C. (Ed.) (1950). *Training in clinical psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Reisman, J. (1981). History and current trends in clinical psychology. In C. E. Walker (Ed.), *Clinical practice of psychology* (pp. 1–32). New York: Pergamon.

Rice, C. E. (1997). Scenarios: The scientist–practitioner split and the future of psychology. *American Psychologist, 52*, 1173–1181.

Robiner, W. N. (1991a). How many psychologists are needed? A call for a national psychology human resource agenda. *Professional Psychology: Research and Practice, 22*, 427–440.

Robiner, W. N. (1991b). Dialogue on a human resource agenda for psychology: A welcome and a response. *Professional Psychology: Research and Practice, 22*, 461–463

Roe, A., Custard, J. W., Moore, B. J., Ross, S., & Skodak, M. (Eds.) (1959). *Graduate education in psychology*. Washington, DC: American Psychological Association.

Rogers, C. R. (1939). Needed emphasis in the training of clinical psychologists. *Journal of Consulting Psychology, 3*, 141–143.

Rotter, J. B. (1971). On the evaluation of methods of intervening in other people's lives. *The Clinical Psychologist, 24*, 1–2.

Rutter, M. (1994). Comorbidity: Meanings and mechanisms. *Clinical Psychology: Science and Practice, 1*, 100–103.

Scarr, S. (1985). Constructing psychology: Making facts and fables for our times. *American Psychologist, 40*, 499–512.

Schacht, T. E. (1985). *DSM-III* and the politics of truth. *American Psychologist, 40*, 513–521.

Schneider, S. F. (1990). Psychology at a crossroads. *American Psychologist, 45*, 521–529.

Seligman, M. E. P. (1995). The effectiveness of psychotherapy: The *Consumer Reports* study. *American Psychologist, 50*, 965–974.

Shakow, D. (1938). An internship year for psychologists (with special reference to psychiatric hospitals). *Journal of Consulting Psychology, 2*, 73–76.

Shakow, D. (1976). What is clinical psychology? *American Psychologist, 31*, 553–560.

Shapiro, D. A. (1996). "Validated" treatments and evidence-based psychological services. *Clinical Psychology: Science and Practice, 3*, 256–259.

Shapiro, D. A., & Shapiro, D. (1982). Meta-analysis of comparative therapy outcome studies: A replication and refinement. *Psychological Bulletin, 92*, 581–604.

Shemberg, K., & Keeley, S. (1970). Psychodiagnostic training in the psychiatric setting: Past and present. *Journal of Consulting and Clinical Psychology, 34*, 205–211.

Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston, MA: Authors Cooperative.

Singerman, B. (1981). *DSM-III*: Historical antecedents and present significance. *Journal of Clinical Psychiatry, 42*, 409–410.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32*, 752–760.

Smith, M. L., Glass, G. V., & Miller, T. I. (1980). *The benefits of psychotherapy*. Baltimore: Johns Hopkins University Press.

Sokal, M. M. (1992). Origins and early years of the American Psychological Association: 1890 to 1906. In R. B. Evans, V. Staudt-Sexton, & T. C. Cadwallader (Eds.), *The American Psychological Association: A historical perspective* (pp. 43–71). Washington, DC: American Psychological Association.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologist Press.

Spitzer, R. L. (1985). *DSM-III* and the politics–science dichotomy syndrome: A response to Thomas E. Schacht's "*DSM-III* and the politics of truth." *American Psychologist, 40*, 522–526.

Staats, A.-W. (1988). Paradigmatic behaviorism, unified positivism, and paradigmatic behavior therapy. In D. B. Fishman, F. Rotgers, & C. M. Franks (Eds.), *Paradigms in behavior therapy: Present and promise* (pp. 211–253). New York: Springer.

Sternberg, R. J. (1995). *In search of the human mind*. Fort Worth, TX: Harcourt Brace.

Strother, C. R. (Ed.) (1956). *Psychology and mental health*. Washington, DC: American Psychological Association.

Sullivan, M. A., & Collins, F. L. (1996). *Instructor's resource manual for Wilson, Nathan, O'Leary, & Clark, Abnormal Psychology: Integrating perspectives*. Boston: Allyn & Bacon.

Suppe, F. (1977). *The structure of scientific theories* (2nd ed.) Urbana, IL: University of Illinois Press.

Task Force on the Promotion and Dissemination of Psychological Procedures (1995). Training in and dissemination of empirically validated psychological treatments: Report and recommendations. *The Clinical Psychologist, 48*, 3–23.

Turner, S. M., Beidel, D. C., & Wolff, P. L. (1994). A composite measure to determine improvement following treatment for social phobia: The Index of Social Phobia Improvement. *Behaviour Research and Therapy, 32*, 471–476.

VandenBos, G. R. (1986). Psychotherapy research: A special issue. *American Psychologist, 41*, 111–112.

VandenBos, G. R. (1993). U.S. mental health policy-Proactive evolution in the midst of health care reform. *American Psychologist, 48*, 283–290.

VandenBos, G. R., DeLeon, P. H., & Belar, C. D. (1991). How many psychological practitioners are needed? It's too early to know! *Professional Psychology: Research and Practice, 22*, 441–448.

VandenBos, G. R., & Pino, C. D. (1980). Research in the outcome of psychotherapy. In G. R. VandenBos (Ed.), *Psychotherapy: Practice, research, policy* (pp. 23–69). Beverly Hills, CA: Sage.

Walker, M. (1963). *The nature of scientific thought*. Englewood Cliffs, NJ: Prentice-Hall.

Wallin, J. E. W. (1960). History of the struggles within the American Psychological Association to attain membership requirements, test standardization, certification of psychological practitioners, and professionalization. *Journal of General Psychology, 63*, 287–308.

Watkins, C. E., Campbell, V. L., & McGregor, P. (1990). What types of psychological tests do behavioral (and other) counseling psychologists use? *The Behavior Therapist, 13*, 115–117.

Watson, R. I. (1953). A brief history of clinical psychology. *Psychological Bulletin, 50*, 321–346.

Weinberger, J. (1995). Common factors aren't so common: The common factors dilemma. *Clinical Psychology: Science and Practice, 2*, 45–69.

Widiger, T. A., & Ford-Black, M. M. (1994). Diagnoses and disorders. *Clinical Psychology: Science and Practice, 1*, 84–87.

Woolfolk, R. L., & Lazarus, A. A. (1979). Between laboratory and clinic: Paving the two-way street. *Cognitive Therapy and Research, 3*, 239–244.

Yates, B. J., & Newman, F. L. (1980). Findings of cost-effectiveness and cost–benefit analyses of psychotherapy. In G. R. VandenBos (Ed.), *Psychotherapy: Practice, research, policy* (pp. 163–186). Beverly Hills, CA: Sage.

Zimbardo, P. G., & Gerrig, R. J. (1996). *Psychology and life* (14th ed.). New York: HarperCollins.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.03
Informational Sources in Clinical Psychology

KRISTI ALEXANDER
United States International University, San Diego, CA, USA
and
MICHAEL C. ROBERTS
University of Kansas, Lawrence, KS, USA

1.03.1 INTRODUCTION 68

1.03.2 JOURNALS AND BOOKS 68
  1.03.2.1 Journals 68
    1.03.2.1.1 Psychopathology and abnormal psychology 69
    1.03.2.1.2 General clinical psychology 70
    1.03.2.1.3 Clinical psychology subspecialties and topics 70
    1.03.2.1.4 Therapeutic techniques 71
    1.03.2.1.5 Professional issues 71
    1.03.2.1.6 On-line journals 71
  1.03.2.2 Books 72
  1.03.2.3 Conclusions 72

1.03.3 ABSTRACTING SERVICES 72
  1.03.3.1 PsycINFO 73
    1.03.3.1.1 Psychological Abstracts 74
    1.03.3.1.2 PsycLIT 74
    1.03.3.1.3 PsycSCANs 75
  1.03.3.2 Educational Resources Information Center 75
  1.03.3.3 Index Medicus 76
  1.03.3.4 Focused Abstracting Services 76
    1.03.3.4.1 Dissertation Abstracts International 76
    1.03.3.4.2 Child Development Abstracts and Bibliography 76
    1.03.3.4.3 Clinician's Research Digest 77
  1.03.3.5 Indexing Systems 77
    1.03.3.5.1 Social Science Citation Index 77
    1.03.3.5.2 Current Contents/Social and Behavioral Sciences 77
    1.03.3.5.3 UnCover 78

1.03.4 ELECTRONIC MEDIA: THE INTERNET 79
  1.03.4.1 Electronic Mail 79
  1.03.4.2 Listservers 79
  1.03.4.3 Usenet Newsgroups 80
  1.03.4.4 File Transfer Protocols 80
  1.03.4.5 World Wide Web 80
    1.03.4.5.1 Accessing the World Wide Web 82
    1.03.4.5.2 Collections of resources on the World Wide Web 84

1.03.4.6 Limitations 84
1.03.5 EVALUATING THE RESOURCES 85
1.03.6 REFERENCES 85

1.03.1 INTRODUCTION

> Knowledge is of two kinds. We know a subject ourselves, or we know where we can find information upon it. (Boswell, *Life of Samuel Johnson,* 1775)

A new client presents to your clinical practice with pica. How do you determine the best course of treatment? A colleague asks you to lecture to the family therapy class on innovations in the treatment of marital discord. Where do you find a comprehensive review of the recent literature? You are interested in learning about the new research on post-traumatic stress disorder. Is the library your only source of information?

The field of clinical psychology is diverse and informational material continually becomes available in a multitude of places. Because the half-life of psychological knowledge is brief, the need for updating and accessing resources is nearly constant for both researchers and practitioners alike. The purpose of this chapter is to indicate where sources of information and the vast array of resources are available. Understanding the variety of reference tools that are useful and how to employ them effectively is critical to "exploit the richness of the materials contained within them" (Mendelsohn, 1987, p. 1). In this first section, we outline where to find psychologically relevant material and some of the "nuts and bolts" of accessing the available information. In the second section, we highlight the types of journals and books that may be most useful to the clinical psychologist. Third, we address the various systems and indexes that summarize the psychological literature. Finally, we hope to bring the reader "on-line" with information regarding the use of the Internet for the novice web surfer. It is our hope that this chapter will serve as a road map to the wide world of information available in clinical psychology.

Although it has been speculated that these days are numbered, the search for informational resources in clinical psychology still begins at the library. Those fortunate to have access to a comprehensive, research university library, with all the advantages of interlibrary loan and on-line search capabilities, will be in the best position to retrieve the psychological literature. As would be expected, the reference section of the library is the researcher's first stop. In this area, the holdings of the library are cataloged and access to the bibliographic tools is usually available. Depending on the size and fiscal solvency of the library, the tools may include card catalogs, computer terminals connected to library-specific holdings catalogs, CD-ROM readers for systems such as PsycLIT and Educational Resources Information Center (ERIC), and a variety of journal abstracts and references. Many of the materials described in the third section of this chapter will be located in this section where they are available to users.

Regardless of how this information is obtained, call letters and numbers will be used to find the material in the library. The two systems of classification, the Library of Congress and the Dewey Decimal System, are most frequently used in libraries in the United States and psychological materials are cataloged under a variety of call letters and numbers. The search for books is relatively straightforward but may allow for some aerobic exercise benefits, depending upon the layout of the library. Journals, in contrast, are often more difficult to locate. Most libraries separate, often by several floors, the current periodicals from the bound, archived volumes. Although this would appear to present a minor inconvenience, the consistency with which this procedure is employed is often variable. Thus, a recent edition of *Journal of Consulting and Clinical Psychology* may be unbound and on the shelf with other current issues; bound and in the stacks; unbound in the stacks with some or all of the other issues in that volume; or finally, off to the bindery.

In light of the declining financial resources of university libraries and the often substantial increases in the cost for books and periodical subscriptions, libraries are becoming more selective in their holdings. In the future, the easy and immediate access that many have grown accustomed to may become atypical. Fortunately, other resources are becoming increasingly available to aid the clinical psychologist.

1.03.2 JOURNALS AND BOOKS

1.03.2.1 Journals

The scientific–practitioner field of clinical psychology has relied on scholarly journals as the prime source of professional information. The foundations of today's general and specialty clinical psychology journals may be found in the *American Journal of Psychology* (founded by G. Stanley Hall in 1887), the *Journal of Abnormal Psychology* (founded by Morton Prince in 1906), the *Psychological Clinic* (founded by Lightner Witmer in 1912), the *Journal of Consulting Psychology* (founded by Johnnie Symonds for the American Association of Applied Psychologists; later to become the *Journal of Consulting and Clinical Psychology*), and the *Journal of Clinical Psychology* (founded by Frederick C. Thorne in 1945) (Routh, 1996). Clinical psychology and related applied psychology specialties can now boast literally hundreds of serial publications. Similarly, there has been a significant rise over the years in the number of books serving the field. In this section, we will describe the range of journals and books with a view to the distinctive contributions each can make. The PsycINFO database (of the American Psychological Association [APA]) abstracts around 1500 journals and publications each year (a list of these is available on the APA homepage). Of course, not all of these publications contain information applicable or useful to the clinical psychologist, but many do. Consequently, sorting through the information sources remains a challenging task.

Many professional and scholarly organizations sponsor journals. For example, in addition to the major organizations of APA and American Psychological Society (APS), other groups supporting journals include the American Orthopsychiatric Association, the Society for Research in Child Development, the American Association for Applied and Preventive Psychology, the American Association on Mental Deficiency, the Association for Advancement of Behavior Therapy, and the Association for Humanistic Psychology. Many of these contract with commercial publishers or self-publish their journals.

In addition, numerous commercial publishers have also developed journal publications through contracts with scholarly editors. Some of the major publishers of privately developed journals include Plenum Publishing Corporation, Clinical Psychology Publishing Company, Lawrence Erlbaum Associates, Sage Publications, Cambridge University Press, Human Sciences Press, Elsevier Science/Pergamon Press, and Guilford Publications. These publishers also produce journals sponsored by organizations; many are noted later. All of these journals are abstracted or indexed by the major services as will be described in Section 1.03.3.

Although clinical psychology as an area of practice and research can be divided and subdivided into varieties of categories, in the following sections, we will illustrate the topical coverage of these archival and serial publications through the following subsections: psychopathology and abnormal psychology (Section 1.03.2.1.1), general clinical psychology (Section 1.03.2.1.2), clinical psychology subspecialties and topics (Section 1.03.2.1.3), therapeutic techniques (Section 1.03.2.1.4), and professional issues (Section 1.03.2.1.5).

1.03.2.1.1 Psychopathology and abnormal psychology

Much of the basic research forming the foundation of clinical psychology is conducted under the rubric of psychopathology or abnormal psychology. The oldest journal in this area, the *Journal of Abnormal Psychology*, published by APA, generally covers articles on the etiology of psychopathology and abnormal behavior. Similarly, the *Journal of Abnormal Child Psychology* publishes studies of behavioral pathology in children and adolescents, but also includes etiology, assessment, and some treatment aspects. *Psychological Assessment* branched from the *Journal of Consulting and Clinical Psychology* to focus on assessment theory and methodology employed in clinical psychology research and practice. The *Journal of Psychoeducational Assessment* and the *Journal of Psychopathology and Behavioral Assessment* also attend to the theoretical and methodological aspects of psychological assessment.

Other journals are published to provide scholarly and clinical coverage for a particular diagnostic category or psychological disorder. For example, *Mental Retardation*, *American Journal of Mental Deficiency*, *Journal of Autism and Developmental Disorders*, and *Research in Developmental Disabilities* provide an arena for research in mental retardation and developmental disabilities. The *Journal of Affective Disorders*, *Journal of Anxiety Disorders*, and *Anxiety, Stress, and Coping* focus on internalizing disorders, while *The Schizophrenia Bulletin* and *Schizophrenia Research* publish research dealing with the etiology, diagnosis, and treatment of serious emotional disorders. The *Journal of Substance Abuse* and *Journal of Drug and Alcohol Abuse* disperses scholarly literature on addictive behaviors. Many other journals publish applied and basic research related to specific psychological problems (e.g., *Journal of Traumatic Stress*, *Child Abuse and Neglect*, *Journal of Emotional and Behavioral Disorders*, *Journal of Family Violence*).

1.03.2.1.2 General clinical psychology

The field of clinical psychology is diverse in the range of problems and interventions that are within its domain. This diversity may not have been a major problem in the early years of development; many clinical psychologists considered themselves generalists with the ability to research and intervene on a wide range of problems and psychotherapeutic techniques. Increasingly, however, specialties have emerged within clinical psychology. Nonetheless, some journals remain diverse in their topical catchment and do not restrict their view on the types of research or professional articles they publish. For example, two empirical journals, *Journal of Consulting and Clinical Psychology* and *Journal of Clinical Psychology*, place no restrictions on subject population, theoretical, or methodological approaches. Similarly, two review and conceptual journals, *Clinical Psychology: Science and Practice* and *Clinical Psychology Review*, are open to any relevant content. Within their pages, these journals may intermingle articles reflecting different ages or populations, the types of problems studied, methodological approaches, and theoretical conceptualizations, although some special issues might focus on a specific topic.

1.03.2.1.3 Clinical psychology subspecialties and topics

Increasingly, the field has become an umbrella for research and applications targeted to special populations and different theoretical or methodological approaches. Thus, different ages or psychopathological conditions are some of the divisions used to categorize publications.

Clinical child and pediatric psychology are perhaps the oldest subspecialties in clinical psychology. Clinical psychologists have conducted research and made therapeutic applications for children and adolescents dating back to the inception of the field. Two child-oriented sections within the APA Division of Clinical Psychology (Section on Clinical Child Psychology, Society of Pediatric Psychology) sponsor scholarly journals. The *Journal of Clinical Child Psychology* publishes articles on child and adolescent psychopathology research and interventions, services and programs for children and families, and professional issues, such as training and ethics. The *Journal of Pediatric Psychology* publishes empirical, review, and case study articles related to applications of psychology to pediatric problems and issues. Another APA Division, Child, Youth, and Family Services, recently decided to sponsor a journal, *Children's Services: Social Policy, Research, and Practice*, to begin publication in 1998.

Non-APA journals in this area include *Children's Health Care*, an interdisciplinary journal published by the Association for the Care of Children's Health, with editorial and article contributions by pediatric and clinical child psychologists as well as nurses, child life specialists, and pediatricians. Another interdisciplinary journal, *American Journal of Orthopsychiatry*, is published by the American Orthopsychiatric Association with contributions by psychologists, psychiatrists, and social workers. An international, interdisciplinary journal, *Clinical Child Psychology and Psychiatry*, began publication in 1996. Finally, sometimes the child and family focus gets paired with a particular theoretical approach, as in *Child and Family Behavior Therapy* and *Psychoanalysis of the Child*. Both of these journals publish articles reflecting a specific orientation and target population.

Geropsychology, at the other end of the developmental continuum, focuses attention on adult development and aging, including the development of psychological problems. As a primary resource, APA publishes the journal, *Psychology and Aging*. Other psychology and interdisciplinary publications that cover this age group for clinical psychology interests include *The Gerontologist*, *Clinical Gerontologist*, *Journal of Geropsychology*, *Journal of Aging and Health*, and *Journal of Applied Gerontology*.

Health psychology/behavioral medicine has been a fast growing field allied with clinical psychology, although it is not always conceptualized as a clinical subspecialty (Bernard & Krupat, 1994). Journals that publish articles at the interface of behavioral sciences, health, and illness include the *Journal of Pediatric Psychology*, *Health Psychology* (sponsored by the APA Division of Health Psychology), *Psychology and Health*, *Journal of Health Psychology*, and *Journal of Behavioral Medicine*.

Community psychology is often thought of as distinct from clinical psychology, however, some aspects of its publications bear on clinical psychology interests. For example, the *American Journal of Community Psychology* (sponsored by APA Division 27) and the *Journal of Community Psychology* report applied and basic research as well as evaluations of intervention programs employing clinically important techniques.

Neuropsychology and rehabilitation psychology have several publications of interest and utility to clinical psychologists in different employment settings. These range from basic research on brain and nervous system functioning, assessment and diagnosis, to treatment, rehabilitation, and prevention. Journals in this and related fields include: *International Journal of Clinical Neuropsychology*, *Psychosocial Rehabilitation Journal*, *Rehabilitation Psychology*, and *Journal of Applied Rehabilitation Counseling*.

1.03.2.1.4 Therapeutic techniques

In practice, clinical psychologists use a broad range of psychological/behavioral intervention techniques to improve the functioning of those presenting with problems. Research and practice publications in this area may describe clinical analogues, controlled experiments, and field evaluations. Some journals take a broad view of psychotherapeutic techniques appropriate for publication; others take a more focused view. *Psychotherapy: Theory, Research, and Practice* (sponsored by the APA Division of Psychotherapy) and the *American Journal of Psychotherapy* are among the more generally oriented journals.

In contrast, there are many periodicals that reflect a specific theoretical orientation. Journals that evidence psychoanalytic or psychodynamic viewpoints include *Psychoanalytic Psychology* (sponsored by the APA Division of Psychoanalysis), *Psychoanalytic Review* (sponsored by the National Psychological Association for Psychoanalysis), and *Contemporary Psychoanalysis*.

An orientation to family-based interventions and family and marital therapy is provided through publications such as *Journal of Family Psychology* (sponsored by the APA Division of Family Psychology), *Family Therapy*, *Family Process*, and *The Family Therapy Networker*.

Cognitive-behavioral interventions, as an amalgamation of cognitive therapy approaches and behaviorally-oriented treatments, are represented by such journals as *Cognitive and Behavioral Practice*, *Cognitive Therapy and Research*, and *Journal of Rational-Emotive and Cognitive Behavior Therapy*. Behavioral therapy and behavior analysis publications on clinical interventions derived from learning theory are found in a number of scholarly journals including *Journal of Applied Behavior Analysis, Behavior Therapy, Behavior Modification*, *Journal of Behavior Therapy and Experimental Psychiatry*, *Behavior Research and Therapy*, and the new, on-line journal, *Journal of Behavior Analysis and Therapy*.

Group psychotherapy has been served by several interdisciplinary publications, including the *International Journal of Group Psychotherapy*, *Small Group Behavior*, and *Journal of Child and Adolescent Group Therapy*.

Prevention of mental and physical health problems is a frequently mentioned goal of clinical psychology interventions. Although several journals in the field publish articles on prevention activities, a primary focus on this topic is provided by such journals as *Applied and Preventive Psychology*, *Journal of Primary Prevention*, and *Journal of Prevention and Intervention in the Community*.

Finally, psychopharmacological treatments represent a substantial and growing specialty interest within clinical psychology. Although some treatment and/or psychopathology oriented journals provide coverage of these issues, the *Journal of Psychopharmacology*, *Research Communications in Psychology, Psychiatry, and Behavior*, *Journal of Child and Adolescent Psychopharmacology*, and *Experimental and Clinical Psychopharmacology* focus exclusively on medication and medication-related issues.

1.03.2.1.5 Professional issues

Within the profession of clinical psychology, issues such as (i) ethics for clinicians, researchers, and educators, (ii) training for professional competence, and (iii) professional practice and functioning (including employment and reimbursement considerations) are important topics. Such professional issues are regularly covered by *Professional Psychology: Research and Practice*, *Ethics and Behavior* and *Psychology, Public Policy, and Law*. Practitioner oriented publications have provided clinicians with information on changes in public sector and private practice concerns, insurance reimbursement and managed care issues, such as in *Journal of Mental Health Administration*, *The National Psychologist*, *Professional Practice of Psychology*, *Journal of Clinical Psychology in Medical Settings*, *Psychotherapy in Private Practice*, and *Behavioral Healthcare Tomorrow*.

The range of journals for the field of clinical psychology is extensive; this sampling cannot do justice to the wealth of serial publications archivable for the clinician and researcher. Additional sources can be ascertained through the APA publication, *Journals in Psychology: A Resource Listing for Authors* (APA, 1993) and the list of journals abstracted by PsycINFO (see APA homepage on the World Wide Web).

1.03.2.1.6 On-line journals

In addition to the familiar print form, some journals have also been added to the Internet. The quality of these publications varies, although they provide an additional source of information for the clinical psychologist. Table 1 lists some of the publications currently available on-line.

Table 1 On-line journals for clinical psychology.

Homepage Association	*URL address*
American Journal of Psychotherapy	http://www.apj.com/
Canadian Journal of Behavioural Science	http://www.cycor.ca/Psych/ac-main.html
Gestalt!	http://rdz.stjohns.edu/gestalt!/
International Journal of Psychopathology, Psychopharmacology, and Psychotherapy	http://www.psycom.net/ijpppp.html
Journal of Behavior Analysis and Therapy (jBAT)	http://sage.und.nodak.edu/org/jBAT/jbatinfo.html
PsychNews International	http://www.cmhc.com/pni/ or http://www.mhnet.org/pni/
Psychology	http://cogsci.ecs.soton. ac.uk/harnad/psyc.html

1.03.2.2 Books

In addition to journal publications, published volumes have been the *sine qua non* for the science and clinical practice of psychology. Authored books and edited books with individually authored chapters have summarized the knowledge and progress of this field since G. Stanley Hall.

Currently, a vast number of books, from a variety of publishing companies too many to list here, serve the field. Many book series act as collections of various titles in clinical psychology, such as *Advances in Pediatric Psychology* (Guilford), *Sage Human Services Guides* (Sage), *Wiley Series in Clinical Psychology* (Wiley), *Wiley Series in Personality Processes* (Wiley), *Developmental Clinical Psychology and Psychiatry* (Sage), *Applied Clinical Psychology* (Plenum), *Clinical Child Psychology Library* (Plenum), *Issues in Clinical Child Psychology* (Plenum), and *Stress and Coping* (Plenum).

As with other forms of information, the prodigious number of published volumes makes it difficult to determine their utility. Decisions about quality are often made by considering the source, the qualifications and reputation of the author/book editors, series editor(s), and the publishing company. Fortunately, there is some additional guidance from *Contemporary Psychology*. This journal is devoted to publishing scholarly and critical reviews of new book publications. Other journals also publish a few book reviews (e.g., *Journal of Pediatric Psychology*, *Clinical Psychology Review*). Unfortunately, the time lag for the reviews is often too long to be helpful to the reader. Unlike most journals that employ peer review, books frequently do not undergo adequate scrutiny before publication. This deficit is particularly unfortunate as descriptions of treatments of unproven quality may be widely disseminated with potential adverse impact on clients.

1.03.2.3 Conclusions

The nature of journals and books is changing with the rise of computer and Internet capabilities to transmit information. Although many predict the ultimate demise of the printed paper format for journals and books, others have not been so quick to accept this prediction. Given the seemingly chaotic nature of the Internet development, the need to sort through and apply some standards of accuracy and utility will remain. Thus, it is likely that the fundamental and time-tested aspects of publishing will hold up, that is, the tradition of the editor as gatekeeper, and the peer review process.

1.03.3 ABSTRACTING SERVICES

As was suggested in Section 1.03.2.3, the sheer number of individual references for any given topic may often be overwhelming. Fortunately, there exist several good sources of abstracts, indexes, and other retrieval mechanisms to facilitate the identification of relevant literature. Information systems such as PsycINFO (from the APA) attempt to provide the resources to access the world's literature in psychology and related disciplines. Journals, technical reports from university and government sources, dissertation abstracts, monographic series, and similar materials are the fodder for these systems. The systems do, however, vary considerably. Some abstracting products provide nonevaluative reports of any reference while others abstract the material presented within a selected set of journals. Further, some indexes provide summaries of the literature, while others offer only bibliographic information. In this section we will describe the major systems and ways to access the information they contain. Although we have provided the most current (as of the late 1990s) information regarding abstracting sources in psychology, because of the rapidly evolving technology, databases, particularly CD-ROM systems, change frequently (Reed & Baxter, 1991). Thus, in order to minimize frustration and maximize successful searches, the reader is advised to monitor new developments in the field.

1.03.3.1 PsycINFO

The major source of information regarding the psychological literature is PsycINFO. In addition to maintaining the comprehensive PsycINFO database, this department of the APA produces several other resources from their database such as *Psychological Abstracts*, PsycLIT, and PsycSCAN. The PsycINFO electronic database includes all the bibliographic references in psychology since 1967, a collection of more than one million citations of psychologically relevant material including journal articles, book chapters and books, dissertations, as well as technical and governmental reports (APA, 1996). References prior to 1967 have been digitized and indexed, and will be available as an option to institutions that lease the database (Hayward, 1996). PsycINFO is updated monthly and all references are presented in English, although not all of the literature abstracted is from English-language publications.

PsycINFO is an abstracting system for psychology and related disciplines that is not bound by geographic or linguistic constraints, but rather evaluates the entire body of relevant psychological information. The database includes research reports and other scholarly works, literature reviews, surveys, case studies, theoretical papers, dissertations, books, chapters, conference and case law reports, and, if the original work was included in the database, comments and replies. Excluded are editorials, letters to the editors, news items, interviews, and poems. Unlike some other abstracting services, PsycINFO does not maintain a core list of journals but rather evaluates the psychological relevance of any individual entrant regardless of the source (e.g., journal, book chapter, government monograph, etc.). By providing a compendium of the psychological literature, PsycINFO permits scholars to "stay current" in their field without surveying all the primary sources.

The PsycINFO database, and the other PsycINFO reference systems, supply three fundamental types of information: bibliographic information, an abstract (or summary), and standardized keywords. The database, however, may only be accessed directly through institutions that have an annual lease for its use and through commercial on-line services, such as Knight-Ridder (formerly DIALOG) and CompuServe.

Before using any PsycINFO products, an understanding of the language and logic of the search systems is necessary. As is the case with most searchable databases, Boolean logic is used. Boolean modifiers such as "AND," "OR," and "NOT" may be used to alter the results of a search. By using these modifiers, a large search can be narrowed to a more relevant topic, while restricted topics can be expanded. For example, a search using anxiety "AND" adolescents will generate fewer matches than anxiety "OR" adolescents. If no modifier is specified, an "OR" is assumed, which may result in fewer relevant references.

Familiarity with the *Thesaurus of Psychological Index Terms* (APA, 1997), the key to the "controlled vocabulary" used to index entries in the PsycINFO database, is essential. Because a variety of words exist to describe psychological concepts, the *Thesaurus* provides a standardized wording, thereby decreasing the frustration that may result when confronted with the idiosyncrasies of the language. By using the *Thesaurus* one is redirected from unused words to the keywords used by PsycINFO (and the corresponding five digit subject codes). For example, searching "birth parents," an unused word, retrieved only 23 records, but by substituting the keyword recommended in the *Thesaurus*, "biological family," 125 records were obtained.

In addition, this reference provides related terms, both broader and narrower, that may be used to further define a literature search. In general, use of the narrower descriptor is preferred because PsycINFO records are indexed with the most specific descriptors possible (APA, 1996). Now in its eighth edition, the *Thesaurus* is updated every three years to reflect changes in the vocabulary of psychology and related disciplines.

Although the *Thesaurus* strives for completeness, there may be times when searching the "free-text" fields may be more fruitful. The free-text fields are entries such as the title (TI), abstract (AB), key phrase (KP), and author (AU) that may be searched for a word or phrase. A complete list of the possible "free-text" search fields is available in the *PsycINFO User Manual* (Walker, 1992) and through on-line "help." Searching the free-text fields may be necessary when a concept does not have a corresponding descriptor. When this occurs, searching the title, the key phrase and abstract of journal articles, may result in a successful "match." For example, if a psychologist was interested in finding articles that address the topic "corporate culture" at the FIND prompt, one would enter "corporate culture in TI (title) or corporate culture in KP (key phrase)." For books and chapters, the table of contents may also be searched.

A free-text field search would be appropriate when references with a specific word or phrase are of interest to the researcher. For example, if interested in studies evaluating the effectiveness of the antidepressant "Zoloft," a search for that specific term in the free-text fields would be made. Likewise, a hunt for a specific person or assessment device may be conducted by this method.

Finally, a free-text search may be needed when searching for studies prior to the development of a new descriptor. For example, to retrieve articles on "AIDS prevention" prior to 1994 (the year it was entered into the database as a descriptor), a free-text search would be necessary. The *Thesaurus* indicates when a particular descriptor was added and in some cases provides alternative descriptors to use when searching the older literature. For example, it suggests that to access the research on post-traumatic stress disorder published between 1973 and 1984, one may use "stress reactions" or "traumatic neurosis" as the index words.

Two other "tricks" that may be used to maximize the results of a "free-text" search are the use of the truncation symbol (*) and the wildcard symbol (?). The truncation symbol may be used to retrieve variants of the specified word or phrase. For example, entering "sex* in TI" retrieved 7636 records from the journal database (January 1990 to June 1996) with a broad range of related words in the titles such as sexualized, sexual harassment, sex offenders, sexually abused, and sexual orientation. Consequently, to obtain the most relevant information, it is important to restrict use of the truncation symbol by including as many of the letters in a word as possible. In the above example, "had sexu*" been used, then fewer records would have been produced.

Use of the wildcard symbol (?) in a word will permit a letter (or letters) to replace it. If the "?" is used at the end of a word, it may be used to represent an "s" for plural forms of the word. Another use of the wildcard symbol is to obtain references with alternate spellings of a word. Thus, to retrieve records with both American and British spellings, "color" and "colour," in the title, "colo?r in TI" must be entered.

1.03.3.1.1 Psychological Abstracts

The "work horse" of the PsycINFO stable is *Psychological Abstracts*, the well-used reference for nearly 70 years. This reference, published monthly, is available only in printed form and provides abstracts of journal articles, technical reports, book chapters, and books. Each volume contains approximately 47 000 records. By including coverage from 1927 to the present, scholars can learn of advances in the field as well as the historical contributions to clinical psychology. Because of the volume of psychological literature worldwide and the cost of producing a printed publication, the inclusion of foreign language entries in *Psychological Abstracts* was discontinued in 1988, although the PsycINFO database still retains these references (Reed & Baxter, 1992).

For many years, *Psychological Abstracts* was the only system available to researchers and clinicians who wished to find literature on any given topic in psychology. However, with the development of PsycINFO and related resources, gone are the days of thumbing through the cumbersome subject references to find the page numbers of corresponding abstracts. PsycLIT replaces, for many, the printed abstracts and allows the user to access the same material from a CD-ROM database. Nevertheless, *Psychological Abstracts*, available only through subscription from the APA, remains a staple reference in most libraries. Each issue features subject, book title, and author indexes, and at the end of the year cumulative indexes of authors and subjects are published in a separate volume.

1.03.3.1.2 PsycLIT

PsycLIT is the database with which most individuals are familiar. Although PsycLIT and PsycINFO are frequently used interchangeably, there are significant differences between the two. As previously described, PsycINFO is an on-line database; in contrast, PsycLIT is a CD-ROM database derived from the PsycINFO database, that contains references for journals, book chapters, and books only. Contained within two compact discs, this reference, updated quarterly, comprises three separate databases: (i) journal references from 1974 through 1989, (ii) journal references from 1990 to the present, and (iii) books and book chapters from 1987 to the present. Each entrant includes a bibliographic citation, an abstract, and, for journal articles, index words. The indexing is included to allow similar articles to be selected from the database. Book and chapter records include comparable information, with the addition of the table of contents for books.

To find references on this CD-ROM version of the *Psychological Abstracts*, keywords are entered. As previously discussed, the use of Boolean modifiers is necessary. There are several support systems that make PsycLIT very user-friendly including an "on-disc" Thesaurus, a tutorial, a videotape entitled "How to Search PsycLIT on CD-ROM," several user's manuals from PsycINFO, and a toll free number to the PsycINFO Actionline ([800] 374-2722) (APA, 1996).

Similarly, the ClinPSYC CD-ROM database offers a reduced set of abstracts specifically chosen for their relevance to practitioners and interested others in clinical psychology, behavioral medicine, and related fields. Because of the reduced volume of entries, the ClinPSYC database is much less expensive and may be useful for smaller institutions and organizations.

1.03.3.1.3 PsycSCANs

Finally, PsycINFO publishes field-specific bibliographic and abstract journals including *Psychoanalytic Abstracts* and PsycSCANs. The increasing specificity of these publications allows individuals to keep up to date with current literature in a circumscribed area of interest. The PsycINFO database of over 1300 journals in psychology and related disciplines is searched to provide the citations and abstracts for these journals. For example, *PsycSCAN: Behavior Analysis and Therapy* provides concise summaries of journal articles on basic research and applied topics accompanied by complete citation to assist readers in locating the full-text article. PsycSCANs, published quarterly, are currently available for the following areas of interest: applied psychology, behavior analysis and therapy, clinical psychology, developmental psychology, and learning disabilities/mental retardation.

Psychoanalytic Abstracts and *Neuropsychology Abstracts* are two similar PsycINFO products that encompass a specific area of psychology. The quarterly *Psychoanalytic Abstracts* replaced the *PsycSCAN: Psychoanalysis* in 1993 and features abstracts of journal articles, as well as summaries of book chapters and books relevant to the speciality area of psychoanalysis. A core set of approximately 40 psychoanalytic journals, in addition to the 1300 general psychology journals, are abstracted. *Neuropsychology Abstracts* fulfills a similar function for the field of neuropsychology.

Both the Abstracts and the PsycSCANs can be very useful in that they provide a manageable number of references to peruse. Nonetheless, scholars interested in a wider scope of information may be frustrated by the circumscribed range covered by these journals. For example, clinical psychologists interested in pediatric phenomena and applications would not have access to all relevant research in this area by using the *PsycSCAN: Clinical Psychology* exclusively.

In summary, the PsycINFO products are excellent resources and allow the scholar to access a wide variety of materials through several user-friendly mechanisms. By developing several different kinds of products, PsycINFO has provided multiple levels of access to their database that vary in terms of scope and cost. Thus, while PsycLIT may be too expensive for a small group practice, the lower cost of the PsycSCANs may allow clinical psychologists entry to the research literature.

1.03.3.2 Educational Resources Information Center

In addition to PsycINFO, another large database that produces a variety of products that may be relevant to clinical psychologists is the ERIC. This national information system, established in 1966, provides access to the large body of education-related resources. ERIC is supported and operated by the National Institute of Education (Gay, 1992).

As with PsycINFO, ERIC offers a range of services, although a primary function is the maintenance of a national bibliographic database of over 775 periodicals in education and related fields. This database is compiled from 16 subject-specific "clearinghouses" that abstract and index the education literature. The majority of these clearinghouses, however, will not be of interest to clinical psychologists as they primarily provide teaching-related materials such as curricula and classroom resources. While there is considerable overlap with the material covered by PsycINFO, ERIC does include references for unpublished research that is not usually available from other sources, such as papers presented at professional conferences and studies conducted by school districts. The ERIC database can be accessed through the Internet (URL: http://ericir.syr.edu) and by CD-ROMs purchased from vendors such as Knight-Ridder, Oryx Press, and SilverPlatter Information, Inc.

In addition to the database, two other products available in both microfiche and printed versions are offered by ERIC: *Resources* *in Education* and *Current Index to Journals in Education*. As with the *Thesaurus of Psychological Index Terms* produced by PsycINFO, ERIC has a companion guide, *Thesaurus of ERIC Descriptors* (Houston, 1990) that should be consulted before beginning a search. Although indexed similarly, the monthly *Resources in Education* includes educational research, conference proceedings, position papers, curriculum guides, books, and dissertations, while the *Current Index to Journals in Education* covers only periodicals.

Clinical psychologists interested in references for child development, education, testing and evaluation research will find references within the ERIC database. Notwithstanding, for many other clinical topics, such as psychological disorders and treatment, PsycINFO products would be a preferable resource. For example, while the search with the keyword "pica" resulted in 50 "hits" from the PsycLIT CD-ROM database, the same search performed with the ERIC CD-ROM generated no relevant references.

1.03.3.3 *Index Medicus*

Index Medicus, published by the National Institutes of Health (NIH), is an abstracting reference for all medical journals, including those areas of psychology related to the health sciences. Approximately 62 psychology and 80+ psychiatry journals are indexed, although many of these are "selectively" searched, and not all their references are included (NIH, 1996). Unlike PsycINFO or ERIC, this reference does not provide abstracts or summaries of the material, only bibliographic citations organized by subject and author.

On-line access to *Index Medicus* is available from MEDLINE, which also includes the *Index to Dental Literature* and *International Nursing Index*. MEDLINE may be accessed through the World Wide Web (http://www.healthgate.com/ or http://www.medscape.com/), by institutional lease, and on CD-ROM, through vendors such as Ovid and Knight-Ridder. The *Medical Subject Headings*, the thesaurus for use with MEDLINE and *Index Medicus*, is updated and published annually.

Many topics of interest to psychologists, such as conduct disorders, "bystander intervention," and family therapy, are found on MEDLINE. Most MEDLINE searches, however, will necessitate some culling to obtain the most useful subset of references as both psychosocially oriented articles and physical medicine articles are retrieved. For psychologists interested in health and pediatric psychology, *Index Medicus* and MEDLINE are invaluable references, providing coverage of medical journals not included in PsycINFO. For example, a search for "pica" resulted in the retrieval of 710 references, in contrast to the 50 references obtained through the PsycLIT search.

One of the major limitations of *Index Medicus*, however, is the limited information provided. Because only the title is provided in the index, there is no indication as to the content of the article, making it difficult to identify critical references from this resource. In contrast, MEDLINE searches may include abstracts when they are available, thus overcoming this deficiency.

1.03.3.4 Focused Abstracting Services

In addition to the three major resources previously described, PsycINFO, ERIC, and *Index Medicus*, a number of other resources provide reference sources and informational services relevant to clinical psychology. Although most of the psychological literature will be accessible through PsycINFO and the other "powerhouses," these additional resources may be of interest to clinical psychologists. In this section we will describe these products, with an eye to their unique contributions.

1.03.3.4.1 Dissertation Abstracts International

There are two parts of the *Dissertation Abstracts International* (*DAI*), Part A: Humanities and social sciences, and Part B: Sciences and engineering. Within this resource, dissertations are listed under subject headings, with interdisciplinary dissertations listed under more than one category. The majority of psychology dissertations will be found in Part B, although Part A contains educational psychology, tests and measurements, and gerontology (Reed & Baxter, 1991). Published monthly by the University Microfilms International at the University of Michigan, the *DAI* is available on microfilm, CD-ROM, and in print. Because dissertations often represent the "cutting edge" of research, this reference may be very useful to individuals interested in new developments in the field. In addition, this reference may be one of the few ways to access this body of scholarly works since dissertations are no longer included in PsycLIT or *Psychological Abstracts*.

1.03.3.4.2 Child Development Abstracts and Bibliography

This publication of the Society for Research in Child Development (SRCD) provides abstracts and bibliographies for books and journal articles related to the growth and development of children. Unlike the PsycINFO products, *Child Development Abstracts and Bibliography* maintains a list of journals regularly searched and solicits contributions of abstracts from journals not included in this list (SRCD, 1996). This database reference is available in print and free through the Internet (http://www.journals.uchicago.edu/CDAB/).

1.03.3.4.3 Clinician's Research Digest

Designed specifically for clinicians, the monthly *Clinician's Research Digest* (*CRD*) is published by the APA. This six-page newsletter covers the highlights gleaned from the review of over 50 clinically relevant journals. Each issue includes a summary of recent research and complete bibliographic information. The *CRD* claims to be more "user friendly" than similar publications (APA, 1996), although the restricted number of articles reported and the absence of PsycINFO descriptors and keywords would appear to limit its utility. Nonetheless, for individuals seeking a brief, informal review of clinical research, the *CRD* may be a very approachable resource.

Similar reference information for clinical psychologists may be found in print publications such as the *Behavioral Medicine Abstracts* (published by the Society for Behavioral Medicine), the *Human Resources Abstracts* (published by Sage), and the *Social Work Research and Abstracts* (produced by the National Association of Social Workers). These references are limited to a specific content area and are currently not available electronically. In addition, because of their cost and redundancy with information abstracted by the other larger abstract services, they may not be carried at many libraries.

In contrast, there are several on-line databases that provide abstracts for specific content areas that are only available electronically, with no corresponding print equivalents. Among these are *Child Abuse and Neglect* (produced by the National Center on Child Abuse and Neglect), DRUGINFO which provides information regarding the psychological aspects of drug and alcohol abuse (produced by Drug Information Services at the University of Minnesota), the *Mental Health Abstracts* (published by IFI/Plenum Data Company), and REHABDATA (from the National Rehabilitation Information Center).

The utility and availability of these abstract services varies considerably. The smaller resources reference a much more circumscribed range of journals and are therefore redundant with the larger systems. Conversely, for some psychologists, the reduced number of references may be more manageable, and consequently, useful. With regard to availability, discipline-specific libraries, such as those found at large research universities, may carry the appropriate references. Nevertheless, because of financial exigencies, many institutions have discontinued their subscriptions to these abstract services.

1.03.3.5 Indexing Systems

In contrast to abstracts, indexes provide only bibliographic information, yet they may be useful reference tools for some scholars. Given the emergence of the "computer age," many of these resources are available electronically as well as in print form.

1.03.3.5.1 Social Science Citation Index

This quarterly publication from the H. W. Wilson Company provides only author and subject indexes from 1400 international English-language periodicals and selective coverage of 3300 additional journals. The coverage of the *Social Science Citation Index* (*SSCI*) is very comprehensive; journal articles, review articles, scholarly replies to the literature, interviews, book reviews, meeting notes, corrections, obituaries, and biographies are indexed (Garfield, 1995). In addition to a print edition, the index is also available on CD-ROM and Social SCISEARCH is the corresponding on-line database from the Institute for Scientific Information. While the majority of psychology journals are covered in the *SSCI*, there is some overlap with the companion publication, the *Science Citation Index* (*SSCI*). Both the *SSCI* and the *SCI* are very expensive and therefore may not be available in smaller libraries.

The *SSCI* allows the identification of research that has been cited by others. If an author's work has been cited by another it will be indexed here, allowing interested individuals to find more recent work based on older research. This information may be very useful for scholars seeking to understand the historical trends in research or the impact of a specific empirical finding. A distinct advantage of these indexes is the use of the Permuterm Subject Index. This system indexes every substantive word in an article's title, thus overcoming some of the obstacles of other indexing systems that require the use of specific descriptors.

1.03.3.5.2 Current Contents/Social and Behavioral Sciences

This is one of the seven multidisciplinary editions available from the Current Contents database, a product of the Institute for Scientific

Information (ISI). This database is updated weekly and is available in both print and electronic forms. For the Social and Behavioral Sciences edition, 1421 journals from psychology, psychiatry, education, public health, social work, and education, as well as other social science areas, are indexed (ISI, 1996). All forms of the publication (e.g., print, disk, CD-ROM, on-line) provide the tables of contents from all the journals reviewed, along with author and publisher addresses. The provision of addresses is particularly handy for ordering reprints. In some formats of Current Contents, abstracts and keywords are also included. As with PsycLIT, on-line access to Current Contents/ Social and Behavioral Sciences is available from commercial services such as Knight-Ridder and Ovid Technologies.

Although there is considerable overlap with the PsycINFO database, Current Contents does provide a unique reference tool. For scholars interested in a wide range of psychology-related materials, such as sociology, law, and social work, Current Contents is unparalleled in its scope. The wide range of coverage is not inexpensive, however; the price of a one-year individual subscription on disk being $895 in 1997. Moreover, because it is updated weekly, continual review of the contents is necessary to keep up with the issues, a task some individuals may find difficult.

Another drawback to Current Contents is the division of the database into multiple issues (e.g., Social and Behavioral Sciences, Life Sciences, etc.). Thus, individuals may find that one edition does not include all the articles in their area of interest. Finally, the index terms included in the back of each issue are very broad (e.g., "personality," "child") and therefore provide only minimal guidance in searching for references on specific topics.

1.03.3.5.3 UnCover

Libraries have long maintained interactive systems to share their resources (e.g., through interlibrary loan, faxes) and have developed computerized systems of informational retrieval. One particularly important and useful electronic development has been the establishment of electronic databases for access to scholarly publications. As a result of the widespread use of these information sources, library groups and private companies have established document notification and delivery services. The UnCover Company (a Knight-Ridder Information Company) is one such service (http://www.carl.org/uncover/unchome.html; telephone: 1-800-787-7979; e-mail: uncover@ carl.org). It provides Internet access to almost 17 000 journals and over seven million articles, The literature available is not just in psychology, but from the totality of science, social science, and the humanities. Journals that are related to clinical psychology (as listed in this chapter) are included in the database.

There are several ways to access the UnCover database: at libraries with an institutional lease, by telnet (database.carl.org), by modem ([303] 756–3600), and through the World Wide Web at the address provided. There is no cost to search the database and individuals may search by author's name, keyword, or journal title. Journal titles allow branching to tables of contents for issues published over a period of years. The results of a search are limited to bibliographic information. Once articles are identified in the search, UnCover can be requested to fax the article to a specified fax machine. The average service charge for fax delivery of the article, including the copyright release fee, is $10.25.

An UnCover topic keyword search for the topic of "managed care" identified 2303 entries. By adding the phrase "mental health" to "managed care," the number of entries was reduced to 115 articles in a variety of journals. All entries seemed relevant, although it is not discernible how many potential entries were missed in the search. Examination of the entries will eliminate some from interest; others may be marked for fax delivery. An UnCover keyword search for "depression" found 7687 items and "depression in the elderly" netted 247 entries. Large numbers of entries are not sorted by the service. The user must scroll through the entries to find useful references.

In addition to the search capacities, another service of this electronic data bank is UnCover Complete which allows the ordering and faxing of any articles in journals not indexed in the data bank or from pre-1989 issues. Entry of the bibliographic information obtained from other sources (e.g., PsycLIT) or reference lists of other articles is required.

A particularly useful service available from the UnCover company is UnCover Reveal. UnCover Reveal allows a subscriber to select up to 50 journals from the database for an "awareness service" where the table of contents from the selected journals are sent to the subscriber's e-mail account as the journal issues are published. Once the desired articles are identified, subscribers may then order them to be sent to their facsimile machine within 24 hours. UnCover Reveal also permits the subscriber to identify 25 keyword and/or author names to be searched weekly as new articles are entered into the database and the results delivered to his or her e-mail address. This service is particularly useful for individuals who may not have access, either personally or through their library, to all of the journals in their area of interest or for scholars who want the most current information as soon as it becomes available. A one-year subscription to this service is $25.00.

In general, UnCover appears to be a potentially useful database service for individuals with limited library access. The ability to order articles on-line and receive them within a brief time period is very convenient, albeit costly. The charge for this service, however, may be equivalent to interlibrary loan fees, with a more expedient delivery time.

1.03.4 ELECTRONIC MEDIA: THE INTERNET

The rise of the electronic media in its multitude of forms has created a vast array of new resources for clinical psychologists, including both researchers and practitioners. At the same time, however, the rapid rise and the enormous wealth of resources have made sifting through and evaluating the utility of those resources more difficult. In this section, we will outline the types of resources available in the electronic media, some of the information sites and their contents, how to access that information, and how to evaluate their use. We want to emphasize that, perhaps different from other referential sources of information in clinical psychology heretofore, any list of materials available through the Internet is, upon compilation, almost immediately outdated. That is, informational homepages, professional and topical newsgroups, discussion networks, Internet-based journals, newsletters, publications, and other forums are being created daily (and others cease to exist). Most importantly, we emphasize that the mere availability of abundant information does not ensure its accuracy or utility for the clinical psychologist.

Before describing some useful Internet resources, we want to clarify and define some of the terms. Although it is often described as a single entity, the Internet is the network of networks, (Kehoe, 1995). Individuals who are connected to another computer in some way, either through telephone lines and a modem, or through direct wiring, are part of the Internet and join approximately 20 million other users in 50 countries on any given day. It is through this network of computers that electronic mail (e-mail), file transfer protocols (FTP), mailing lists (listservers), on-line discussions (bulletin boards and newsgroups), and the World Wide Web (WWW) are possible.

Access to the Internet, and the many services it supports, may be made through commercial companies (e.g., America Online, Compuserve, Prodigy, and many local network providers) as well as educational institutions and governmental agencies. Generally, the type of connection to the Internet does not affect the types of services available, but will impact the speed at which information may be accessed.

In the last few years, consumer-oriented, user-friendly, services and tools have developed to make it quite easy for the novice computer user (a "newby" in Internet lingo) to gain access to the multitude of informational sites (as well as entertainment and other diversions) on the Internet. The software for sending out electronic mail or searching the Internet for information via the larger commercial companies are fairly easy to use (although the financial charges for use and time can accumulate fairly rapidly). For those accessing the Internet through other means (e.g., regional or local networks, or .edu or .gov computer accounts), special e-mail software and/or use of a "web browser" may be provided. Web browsers are generally point-and-click tools for searching and accessing information; we have found the easiest ones to use are Netscape and MS Internet Explorer. These programs are free for academic and research users and provide a hypertext interface to the World Wide Web. Easily installed, they can be quickly activated for conducting searches and acquiring information. These browser programs generally provide a range of options for accessing information. We will describe some of the informational search options later in this section.

1.03.4.1 Electronic Mail

Although directly e-mailing other professionals is a useful way of requesting, obtaining, and exchanging information for the clinical psychologist, we will not deal with this medium in detail. Electronic mail addresses for individuals are now frequently included in membership directories for professional organizations such as the APA and the APS.

1.03.4.2 Listservers

Collections of e-mail addresses of those people who have actively subscribed to a central computer list are called listservers (or listservs). These groups are typically centered around a common interest or activity such as members of the Division of Clinical Psychology, or researchers and practitioners with interests in traumatic stress. Members of the listserver may post inquiries for information, posit a thought, initiate dialogues and discussions over professional issues, and provide information via e-mail to the central computer and all members of the group receive copies of the message in their e-mailboxes. Some listserver groups are moderated by a professional; others are freewheeling and supervised only by peer pressure.

As an example of how a listserver group might assist the clinical psychologist, an inquiry to a clinical psychology network from a practitioner asked for valid and reliable ways to assess a particular disorder suspected in a client. Replies from the group included several suggestions of how to handle the case, references to written material, and summaries or abstracts of articles on the topic. Often the replies illuminate and educate, although sometimes the discussion turns into an argument or meaningless chatter. (A word of warning here: There are those who enter a newsgroup apparently to stir up trouble deliberately by posting erroneous information or making statements calculated to provoke the other members of the newsgroup.) To enroll in a listserver usually requires only that one sends a "subscribe" message by e-mail, although a few others require checking against a master list or review of credentials. Some listservers of relevance to clinical psychology are listed in Table 2.

1.03.4.3 Usenet Newsgroups

Usenet newsgroups differ from e-mail listservers in that messages or articles are not transferred directly to a user, but are posted in a central location accessible by individuals" computers through the Internet. Usenet newsgroups typically form around a topic or interest. Members of the newsgroup post messages on the "bulletin board" of the group, and other members may react by posting replies or new positions/statements to the bulletin board for all to see or send personal replies to those who have posted before. There is no Usenet organization or authority, although some newsgroups are "moderated" by an individual who reads and approves a message before posting it in the newsgroup board. Supportive contacts with others facing a particular problem are often facilitated in these groups.

The lengthy list of the thousands of Usenet newsgroups can be accessed through Netscape (http://sunsite.unc.edu/usenet-i/hier-s/top.html). Each newsgroup address is preceded by a category prefix (e.g., sci., soc., alt., etc.) that represents a broad topic area (e.g., science, social, alternative, etc.). Subsequent words serve to add specificity to the particular newsgroup. For example, many of the groups interested in psychology begin with the "alt.psychology" followed by another, more specific term such as "personality." The DejaNews research service (http://www.dejanews.com) may be used to locate newsgroups addressing specific topics. Some examples of Usenet newsgroups that may be of interest to clinical psychologists include:

alt.psychology	(.adlerian, .jung, .help)
alt.support	(.attn-deficit, .depression, .schizophrenia)
alt.recover	(.compulsive eating, .sexual addiction, .aa)
sci.psychology	(.theory, .psychotherapy, .personality)
alt.sexual.abuse.recovery	
soc.support.loneliness	

1.03.4.4 File Transfer Protocols

Through file transfer protocols (FTP), professionals may obtain and exchange information in computer files from one system to another. A variety of information resources including journals, books, and other documents can be downloaded to the clinical psychologist's computer through this mechanism. Searches for information obtainable through FTP may be made through the "Gopher" system which searches for documents and retrieves them or a Wide-Area Information System (WAIS).

1.03.4.5 World Wide Web

The collective term for the global network of interlinked computers is the Web or World Wide Web (WWW). The vast array of information contained in all those computers is available for access through special, easy to use tools. The Web is both a "treasure trove" and a "chaos morass." Indeed, a new user accessing the WWW for the first time will be amazed and perplexed at the variety and range of offerings. After initial shock and fascination in "surfing" around the Net, the Internet user will approach the Web and Internet activities with a more calculated examination of what is useful to him or her.

All Web pages are found by their addresses known as uniform resource locators (URLs). These typically start with http://www and the entire address must be exactly entered. WWW homepages often contain sound, color, and graphics. Consequently, many pages may take some time to download into one's computer, thus a fast modem is recommended (2400 baud).

Individuals, organizations, and companies all have a presence on the Web. Many individuals

Table 2 Listservers for clinical psychology

Organization	*E-mail address*
Behavior Analysis	behvioraanalsys@mankato.msus.edu behav-an@listlistserv.nodak.edu
Division 12 Net (for members of the Division of Clinical Psychology, APA)	div12@listserv.nodak.edu
Society for a Science of Clinical Psychology (for members of the Section III of the Division 12, APA)	sscpnet@bailey.psych.nwu.edu
Group-Psychotherapy	Majordomo@freud.apa.org (in body of message, type: SUBSCRIBE GROUP-PSYCHOTHERAPY)
Society for Community and Research Action (Division 27 of APA)	SCRA-L@UICVM.CC.UIC.EDU
PsyUSA: Practice-Oriented discussions (managed by John Roraback: jmroraback@aol.com)	PsyNetUSA@aol.com
Autinet (Autism)	autinet-request@iol.ie LISTSERV@sjuvm.STJOHNS.EDU
Behavioral Disorders in Children	LISTSERV@ASUVM.INRE.ASU.EDU (in body of message type: Subscribe BEHAVIOR username)
Dual Diagnosis (for professionals with interests in persons with dual diagnoses of mental disorders, drug and alcohol addiction)	MIDAA-dual-diagnosis@idealist.com
NetPsy: Internet Psychotherapy (issues related to psychotherapy services provided via Internet)	listserv@sjuvm.stjohns.edu
PsychNews International	LISTSERV@LISTSERV.NODAK.EDU (in body of message type: SUBSCRIBE PSYCHNEWS Firstname Lastname)
GrantsNet	listserv@list.nih.gov (in the body of the message type: subscribe GNET-L FirstName LastName)

have set up their own homepages containing information about themselves and their interests, businesses, productions, and professional activities. Specific to topics within psychology, professionals and lay people with common interests have established homepages with information about biopsychosocial phenomena (e.g., on dual diagnoses, depression, anxiety disorders, aggression, drug and alcohol addiction). Table 3 provides the Internet addresses for special interest groups.

Similarly, companies offering services or information have also created homepages with book or product catalogs. (We do suggest, until better safeguards are in place, placing any orders for materials through the mail or telephone for credit card number security.) Many colleges and universities also have homepages with links to other sites within their institution, including descriptions of course offerings, faculty, their expertise and research, and other useful information. Other Web sites contain encyclopedic-type information (e.g., ARGUS/University of Michigan Clearinghouse and The Electric Library). Most importantly, governmental agencies, scholars and practitioners, and scientific and professional organizations have created sites of information containing a great variety of resources in their homepages (and interlinked them with others of similar topics). Table 4 provides the Internet addresses for several organization homepages. These are some illustrative samples, not a comprehensive list, of organizations related to clinical psychology accessible on the Net.

Table 3 Special interest and discussion groups on the Web.

Homepage association
URL address

The Obsessive-Compulsive Foundation (Resource Center)
http://www.ocdresource.com
Society for Research in Adolescence
(special interest group in peer relations)
http://www.darling.hhdev.psy.edu/peesig.htm
Group Psychotherapy
http://freud.tau.ac.il/~haimw/group2.html
Behavior OnLine
(discussions and interviews with practitioners and researchers)
http://www.behavior.net/
Dual Diagnosis
(for professionals with interests in persons with dual diagnoses of mental disorders, drug and alcohol addiction)
http://www.erols.com/ksciacca/
The Psychoanalytic Connection
http://marge.infohouse.com/psacnct/

1.03.4.5.1 Accessing the World Wide Web

If the reader is not familiar with accessing the Web, we suggest consulting the variety of books on surfing the Net, although the best way to learn the Net resources, capabilities, and limitations is to take some time to explore it for oneself. We will give here only a brief introduction to accessing the WWW with a focus on clinical psychology resources.

Most novices (and likely more experts) access the Internet for information using one of the popular Web browsers (e.g., Netscape, MS Internet Explorer). There are numerous Web indexes available providing some organization (e.g., the Yahoo index site itself lists and links to over 400 index sites). Programs (called crawlers, robots, spiders, and search engines) allow the user to search through the Web and collect addresses of the information sought through a keyword search; the user can then access the identified sites through pointing-and-clicking (e.g., through Webcrawler, Lycos). These search programs are easily accessed through the Netscape or Mosaic programs. The newest search vehicle is AltaVista, reported to index and search over 30 million documents on the Web out of 225 000 Web servers around the world (Heacock, 1996).

The various sites are, thus, interlinked and permit exploration and searching. The search capacities of the programs or search engines allow the user to enter a keyword or phrase for a search through all the Internet for the concept or exact word/phrase. Often a large number of hits on the topical term will be found. In a search via Yahoo (through Netscape 2.0), we obtained 24 hits or sites with information on the topic of "clinical psychology." Searches can be fine-tuned, specified, or narrowed by the inclusion of additional concepts central to the individual's desired search (e.g., add "psychopathology" or "elderly").

To demonstrate the wealth of information available, we conducted a search via the bigger programs in Netscape on the keyword of "depression" and to refine the concept, added the phrase "in the elderly." The results of the various searches include:

	Depression	*Depression in the elderly*
Yahoo	82 matches	0 matches
Magellan	6640 "results"	12 651 "results"
Excite	over 100 documents	over 100 documents
Infoseek	5681 "best"	13 061 "best"
(searched through 51 191 696 unique URL sites)		
Altavista	7000 matches	7000 matches

The browser programs organize the site hits using an algorithm, and the results are often ranked by "relevance" or confidence in the "fit to the concept." Given the ambiguity of the term "depression" and the range of topics in "elderly," these searches generated quite a few unrelated information sites for such things as "depression era artists," "depression glass," "The Great Depression," and data on a tropical depression off Acapulco. Fortunately, many documents or pages were found for psychological depression. These sites included health center information, on-line depression screening tests, governmental information pages, support groups, professional organizations, consensus

Table 4 Organization homepages relevant to clinical psychology.

Homepage Association	*URL address*
Psychology organizations	
American Psychological Association	http://www.apa.org
American Psychological Society	http://www.hanover.edu/psych/APS/aps.html
British Psychological Society	http://www.bps.org.uk/
Canadian Psychological Association	http://www.cpa.ca/
Society for a Science of Clinical Psychology	http://www.sscp.psych.ndsu.nodak.edu
Clinical child/pediatric psychology organizations	
Society of Pediatric Psychology	http://129.171.43.143/SPP/index.html
University of Minnesota: Pediatric Centers, Policies, and Societies	http://www.peds.umn.edu/centers
Children's Defense Fund	http://www.tmn.com/cdf/index.html
Legal organizations	
ABA Network: American Bar Association	http://www.abanet.org/home.html
LAWlink: ABA Legal Research Selected Starting Points	http://www.abanet.org/lawlink/home.html
The ABA Center on Children and the Law	http://www.abanet.org/child/
Governmental information sites	
National Institute of Mental Health	gopher://gopher.nimh.nih.gov
GrantsNet: Department of Health and Human Services	gopher://gopher.os.dhhs.gov:70/1/Topics/grantsnet
FedWorld	http://www.fedworld.gov
National Institute on Drug Abuse	http://www.nida.nih.gov

conference statements on treatment, a number of testimonials (e.g., autobiographies, poetry, photos) of people describing their own depressive episodes and treatment, and a large number of adverts offering psychological services. Some of the information offered treatment approaches or information of questionable accuracy, efficacy, and taste (and relatedly, ethics).

Once the search service (e.g., Lycos, Altavista) presents the results, the user points-and-clicks on each interesting entry and that site is accessed. Because one site is usually linked to several others of similar topics, one may then follow from one piece of information to another. Alternatively, one may return to the list of search results to explore the next entry. Information in the sites can be read at the time, downloaded for later use, or printed into hardcopy. In our search, many of the same sites were uncovered by the various services (but some search by different mechanisms or in different sources). Going through the multitude of sites identified in the search on "depression" and "depression in the elderly" (or whatever keyword) is eye-straining, if not mind-boggling.

In other keyword searches for clinical psychology topics, we found a similar bewildering, yet potentially helpful, array of information. A search on "panic disorders" turned up the NIH consensus statement and the NIMH patient information pamphlet. A search on "traumatic stress" obtained the homepage for the VAMC in Phoenix with diagnostic and treatment information. "Attention Deficit Disorders" generated a large number of sites on treatments, professionals offering diagnosis and therapy, educational and behavioral assessment, and support services. It is unclear how all of these reflect on the call from some professionals in clinical psychology for validated assessment and treatment information, but the information is all "out there" for public consumption.

1.03.4.5.2 Collections of resources on the World Wide Web

Professional groups and individual psychologists have developed a number of WWW sites in which many other homepages are listed for the users to follow up. (Special icons or indicators are used, such as push-pins, highlighting, underlining, or different colored words, to note that a linkage to another set of information or page may be accessed by pointing at the icon and clicking.) A central site listing homepages is a useful place to start in finding information related to psychology in general and clinical psychology, in particular. Following from one site to another is the best way to explore the Web. Once the user has identified some sites as more useful than others, the URLs can be retained as bookmarks for easy access. Table 5 provides the Internet addresses for several resource links. Each of the sites listed includes links to many other sites related to clinical psychology.

Using the Web browser (and various search engines such as Yahoo, Lycos, Magellan, WebCrawler, AltaVista), one can also enter the name of an organization or group having related interests on a topic and find its resource pages of position papers, scientific or professional articles, and other informative documents. The URLs for the APA links to pages of its Public Policy Office and Practice Directorate, while the APS has several psychology links.

Other organizations' homepages, such as the American Bar Association (ABA), maintain interesting and useful Internet resource linkages as well. The ABA entity, for example, links 2200 sites including the Commission on Mental and Physical Disability Law, the Center on Children and the Law, and the Commission on Domestic Violence, as well as numerous publications and position papers. Similarly, the American Medical Association (http://www.ama-assn.org/) maintains links to the *Journal of the AMA* and other medical information sites. Other useful homepages include the American Academy of Pediatrics (http://www.aap.org/), the American Psychiatric Association (http://www.psych.org/), and the American Academy of Child and Adolescent Psychiatry (http://www.psych.med.umich.edu/web/aacap)/).

Mental health and human services organizations such as The National Alliance for the Mentally Ill (http://www.cais.com/vikings/nami.index.html), the National Mental Health Association (http://www.worldcorp.com/dc-online/nmha), and HandsNet (30 000 human service and public interest organizations: http://www.igc.apc.org/handsnet) also maintain informational sites.

Governmental agencies are often a font of information through Web homepages (URLs) or Gopher sites. For example, the Center for Substance Abuse Prevention (in the Department of Health and Human Services) maintains the informational service of the National Clearinghouse for Alcohol and Drug Information accessible through the Web (http://www.health.org/). Service and research grant information from the Federal government is obtainable through a Gopher search (gopher://gopher.os.dhhs.gov:70/1/Topics/grantsnet). The National Technical Information Services, also known as FedWorld, provides electronic access for health-related information (http://www.fedworld.gov), while information about US House and Senate legislation is accessible through the Library of Congress (http://thomas.loc.gov/).

1.03.4.6 Limitations

The rapid rise of the Internet led to great expectations and excitement over its motherlode of information. Increasingly, however, the Internet has been criticized for a number of limitations. A primary criticism is that there is no overview or organization of the material placed on the Web, for example, and that almost anything can be and is out on the Net for public and professional access. In its very foundation and its chaotic development, "management" has been avoided. While this may be seen as an asset, the explosion of information and the huge number of sites has made it difficult to make sense of it all. In contrast to scientific and professional publications that typically have editors and editorial boards for evaluating the credibility of the information prior to publication, most of the material on the Web is not subject to any overview, except in certain instances (such as for some Internet journals).

As an example of how bad information can be disseminated widely, we have received e-mail messages via a large listserver group with repeated claims that there are federal laws requiring that a report of child abuse be made anytime a child client presents with a change in behavior. (This claim is untrue, and several attempts to correct this information were met with harsh and personal attacks from the vociferous poster of the misinformation.) Similarly, Web sites may offer psychotherapy through the Internet or services and products of

Table 5 Resource links for clinical psychology.

Homepage association	*URL address*
Psychology resources	
Psychology Web Pages	http://www-psy.ucsd.edu/otherpsy.html
Psych Web	http://www.gasou.edu/psychweb
Psyc Site	http://www.unipissing.ca/psyc/psycsite.htm
Hanover College Psychology Department	http://psych.hanover.edu
Mental Health Net	http://www.cmhc.com
PsychScapes Worldwide	http://www.mental-health.com
Informational and related resource links	
Clearinghouse for subject-oriented Internet resource guides	http://www.lib.umich.edu/chhome.html
THOMAS: Legislative information on the Internet	http://www.thomas.loc.gov/
Neuroscience mailing lists	http://www.lm.com/~nab/neurolist.html
MedWeb	http://www.emory.edu/WHSCL/medweb.ejs.html

uncertain ethical or probative value. Indeed, there is no requirement that any Web site contain truthful or useful information. Thus, in our view, the warning of *caveat emptor* (buyer beware) extends to the free access to the Web information sites.

How does one sort through this range of information? As with other professional literature, we suggest the reader consider the source and apply one's professional training of evaluation. Heacock (1996) suggests that many people "have been led to believe that the Internet (and more specifically, the World Wide Web) is some kind of exhaustive library of human knowledge, that anything you want to know is out there, and that there is some simple, logical way to find it. Nothing could be further from the truth" (p. 2D). Thus, Web searching may result in an overabundance of undigested and untested information. The old saying about the "need to separate the wheat from the chaff" is quite applicable here.

1.03.5 EVALUATING THE RESOURCES

Evaluating the vast array of reference materials for clinical psychology may be overwhelming, perhaps similar to "being asked to climb Mount Everest without a Sherpa guide" (Rosnow & Rosnow, 1995, p. 19). It is our conclusion that most of the informational resources available to clinical psychologists are good. The difficulty is, then, in assessing the quality of unknown information and locating appropriate references.

With respect to quality, the peer-review process provides one mechanism by which the caliber of research literature may be kept high. "Refereed" journals, where others knowledgeable about the topic review each article prior to publication, are common, although not universal. Unfortunately, this system also results in many potentially important research reports not being published as the average rejection rate for journals is about 75% (Salkind, 1994). In contrast, most books do not undergo a similar pre-publication evaluation. Thus, it behooves the scholar to keep a skeptical eye open when evaluating the utility of any given resource, be it journal, book, or Usenet newsgroup.

1.03.6 REFERENCES

American Psychological Association (1993). *Journals in psychology: A resource listing for authors* (4th ed.). Washington, DC: Author.

American Psychological Association (1997). *Thesaurus of psychological index terms*. Washington, DC: Author.

American Psychological Association (1996). *PsycINFO: Information services in psychology*. [On-line], Available: http://www.apa.org/psycinfo/psycinfo.html

Bernard, L. C., & Krupat, E. (1994). *Health psychology: Biopsychosocial factors in health and illness*. Fort Worth, TX: Harcourt Brace.

Gay, L. R. (1992). *Educational Research: Competencies for* *analysis and application* (4th ed.). New York: Macmillan.

Garfield, E. (1984). How to search the Social Citations Index. *Current Contents, 27,* 3–11.

Hayward, W. C. (1996, Spring). *PsycINFO News, 16*(1). Washington, DC: American Psychological Association.

Heacock, D. (1996, June 24). Search engines clear Web dumping ground. *Lawrence Journal-World*, 2D.

Houston, J. E. (Ed.) (1990). *Thesaurus of ERIC Descriptors* (12th ed.). Phoenix, AZ: Oryx.

Institute for Scientific Information (1996). *Current Contents Journal Coverage*. Philadelphia: Institute for Scientific Information.

Kehoe, B. P. (1995). *Zen and the art of the internet* (4th ed.). Upper Saddle River, NJ: Prentice-Hall.

Mendelsohn, H. N. (1987). *A guide to the information sources for social work and the human services*. Phoenix, AZ: Oryx.

National Institutes of Health (1996). *Index Medicus*. Bethesda, MD: Author.

Reed, J. G., & Baxter, P. M. (1992). *Library use: Handbook for psychology* (2nd ed.). Washington, DC: American Psychological Association.

Rosnow, R. L., & Rosnow, M. (1995). *Writing papers in psychology* (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Routh, D. K. (1996). Lightner Witmer and the first 100 years of clinical psychology. *American Psychologist, 51*, 244–247.

Salkind, N. J. (1994). *Exploring research* (2nd ed.) New York: Macmillan.

Society for Research in Child Development (1996). *Child development abstracts and bibliography, 70*. Chicago: University of Chicago Press.

Walker, A. (Ed.) (1992). *PsycINFO user manual*. Washington, DC: American Psychological Association.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.04
Contributions of Behavioral Genetics Research to Clinical Psychology

THOMAS G. O'CONNOR and KIRBY DEATER-DECKARD
Institute of Psychiatry, University of London, UK
and
ROBERT PLOMIN
PaSU, University Park, PA, USA

1.04.1 INTRODUCTION 88

1.04.2 HISTORICAL CONTEXTS AND CHANGING CONCEPTUALIZATIONS OF BEHAVIORAL GENETICS AND PSYCHOLOGY 89
- *1.04.2.1 The Changing Dynamics of Behavioral Genetics Research* 89
- *1.04.2.2 Key Studies in Behavioral Genetics Research* 90
  - *1.04.2.2.1 Genetics and major mental illness* 90
  - *1.04.2.2.2 Genetics and development* 90
  - *1.04.2.2.3 Genetics and the environment: sibling differences in adjustment* 90
  - *1.04.2.2.4 Genetics and the environment: environments are not independent of genetic "effects"* 91
  - *1.04.2.2.5 Progress in animal research* 91
  - *1.04.2.2.6 Psychometric developments in behavioral genetics* 91
- *1.04.2.3 Controversies Old and New* 92
- *1.04.2.4 Summary* 92

1.04.3 A SELECTIVE OVERVIEW OF METHODOLOGICAL ISSUES IN BEHAVIORAL GENETICS RESEARCH 92
- *1.04.3.1 A Précis of Quantitative Genetics* 93
- *1.04.3.2 An Overview of Research Designs to Detect Genetic and Environmental Influence* 93
  - *1.04.3.2.1 Twin studies* 93
  - *1.04.3.2.2 Adoption studies* 94
  - *1.04.3.2.3 Family history studies* 94
- *1.04.3.3 Alternative Research Designs* 94
- *1.04.3.4 Methodological Criticisms of Behavioral Genetics Methods* 95
- *1.04.3.5 Methodological Recommendations* 96
  - *1.04.3.5.1 Multiple measures* 96
  - *1.04.3.5.2 Specification of individual-specific environmental risks* 96
- *1.04.3.6 Summary* 97

1.04.4 THE CONTRIBUTIONS OF BEHAVIORAL GENETICS TO RESEARCH ON CLINICAL PHENOTYPES 97
- *1.04.4.1 A Selective Review of Genetic and Environmental Etiology in Psychopathology* 97
- *1.04.4.2 Continua and Categories in Models of Psychopathology* 98
  - *1.04.4.2.1 Distinguishing between the underlying assumptions and quantitative implications of categorical*

and dimensional models 98
1.04.4.2.2 Are shared environmental influences more salient at the extremes? 99
1.04.4.2.3 Continuum, boundaries, and spectra 99
1.04.4.3 Comorbidity of Psychopathologies 99
1.04.4.3.1 Comorbidity of psychopathologies: environmental evidence 101
1.04.4.4 Genetic Heterogeneity 101
1.04.4.4.1 Age of onset and other features 102
1.04.4.5 Genetic Influence on Change and Stability of Psychopathology 102
1.04.4.6 Summary 103
1.04.5 MODELS OF GENETIC AND ENVIRONMENTAL RISK MECHANISMS 103
1.04.5.1 Conceptualizing Genetic Influence 103
1.04.5.2 Conceptualizing the Environment 104
1.04.5.3 Genotype–Environment (G–E) Correlations 104
1.04.5.3.1 Passive genotype–environment correlation 104
1.04.5.3.2 Evocative and active genotype–environment correlations 104
1.04.5.3.3 Hypothesized age changes in genotype–environment correlations 105
1.04.5.4 Genotype–Environment Interactions 105
1.04.5.5 Summary 106
1.04.6 MOLECULAR GENETIC RESEARCH ON PERSONALITY AND PSYCHOPATHOLOGY 107
1.04.6.1 Methods and Questions in Molecular Genetics Research on Psychopathology 107
1.04.6.2 Contributions of Molecular Genetics Research to Models of Psychopathology 107
1.04.6.3 Summary and Future Directions 108
1.04.7 CLINICAL PSYCHOLOGY IMPLICATIONS OF BEHAVIORAL GENETICS RESEARCH FINDINGS 108
1.04.7.1 Clarifying Genetic Risks in Assessment 108
1.04.7.2 Genetic and Environmental Considerations for Treatment and Prevention 108
1.04.7.3 Summary 109
1.04.8 GENERAL SUMMARY AND CONCLUSION 109
1.04.9 REFERENCES 110

1.04.1 INTRODUCTION

There have been numerous significant advances in our understanding of the nature of mental disorder in recent years, and considerable progress derives from the contributions made by behavioral genetic research. Research findings from behavioral genetics have substantially altered how we understand: the risks for major mental illnesses such as schizophrenia, depression, and bipolar illness; personality disorders and dimensions; disorders of childhood and adolescence; stability and change in psychopathology; and the nature of psychosocial risks for psychopathology. No longer a peripheral issue, behavioral genetics has made lasting contributions to the field of clinical psychology. This is perhaps most obviously indicated by the inclusion of a chapter on behavioral genetics in the "Foundations" volume.

Writing a foundations chapter in behavioral genetics is complicated for many reasons. First of all, the field of behavioral genetics is, by design, multidisciplinary and complex. At a very basic level, behavioral genetics provides a conceptual and methodological bridge between the biological and the behavioral sciences. Conceptually, behavioral genetics hypotheses test whether individual differences in a behavioral phenotype can be attributed to genetic and environmental influences. Methodologically, behavioral genetics is the research tool that specifies the samples and methods necessary to answer conceptual questions. A second difficulty in writing this chapter is that there are already a number of excellent texts that introduce many of the key methodological and conceptual issues. In particular, the third edition of *Behavioral genetics* (Plomin, DeFries, McClearn, & Rutter, 1997) introduces many central concepts and methods in behavioral genetics research and reviews of many of the more challenging findings. In addition, other well-received sources focus on more specific areas within behavioral genetics, such as the role of genetic factors in major mental illness (Gottesman, 1991) and childhood psychopathology (McGuffin, Owen, O'Donovan, Thapar, & Gottesman, 1994; Rutter et al., 1990). Still other volumes provide an overview of the recent and ongoing questions of behavioral genetics in psychology (Loehlin, 1992; Plomin, 1994; Plomin & McClearn, 1993). Technical texts on statistical genetics (e.g., Neale & Cardon, 1992) and molecular genetics are also available. In this chapter, we incorporate lessons from each of these perspectives, provide an update on more recent findings with major implications for research on psychopathology, and mold it in a form that is pertinent to clinical psychology.

An additional major challenge for a review chapter is the amazing speed with which the field of behavioral genetics is progressing. Research reviews on the genetics of one area or another published just a few years ago are already outdated. Nowhere is this more apparent than in the field of molecular genetics and the interface between molecular genetics and psychiatry. Accordingly, our primary focus is on the conceptual underpinnings of behavioral genetics. An effort is made to review key findings, but we avoid an encyclopedic approach in which the genetics of all disorders is discussed. Chapters that focus on specific disorders in this series will address genetic influences.

A further orienting remark concerns the perspectives from which we review the relevance of behavioral genetics for clinical psychology. We view behavioral genetics research as relevant not only to our understanding of genetic influences on development, but also to our understanding of psychosocial factors. A theme emphasized throughout the chapter is that behavioral genetic studies provide one of the best handles on environmental influences. A second perspective is that we adopt a developmental approach. The reasons for this are rooted not only in our background as developmental psychopathologists, but also in the fact that empirical studies (behavioral genetic studies included among them) repeatedly emphasize that much can be learned about the behavioral phenotype, risk and protective factors, and mechanisms underlying the development of disorder when clinical phenomena are observed longitudinally.

This chapter is divided into six sections. First, in order to place the current set of findings and methods in context, we briefly highlight the history of behavioral genetics research. Classic papers and exemplary studies that challenged basic notions of child and adult development are emphasized. The central theme in this section is that the pendulum that has swung back and forth between nature and nurture—the terms coined by Francis Galton over a century ago—appears to have settled to a happy equilibrium. Second, we review some of the core conceptual and methodological issues in behavioral genetics. Some of the central issues are included here for readers who may not have come across these ideas previously. In particular, we seek to clarify some of the popular controversies and misunderstandings and differentiate those debates that have been largely resolved from those that require additional attention. Third, we examine how behavioral genetics research has contributed to nosological debate and a number of related issues (e.g., comorbidity) and whether psychopathology is qualitatively different from the normal range.

Fourth, the ways in which behavioral genetics research has contributed to our understanding of development, and to developmental psychopathology in particular, are delineated. The penultimate section introduces the burgeoning field of molecular genetics and its links with clinical psychology and psychiatry. Such is the influence and momentum of molecular genetics that a chapter on the foundations of behavioral genetics and clinical psychology would surely be incomplete without a general discussion of the methods and emergent findings. Finally, the last section discusses some of the implications of behavioral genetics methods and findings for clinical psychology.

1.04.2 HISTORICAL CONTEXTS AND CHANGING CONCEPTUALIZATIONS OF BEHAVIORAL GENETICS AND PSYCHOLOGY

1.04.2.1 The Changing Dynamics of Behavioral Genetics Research

The current prevalence of behavioral genetics research in a very wide array of journals is impressive. Indeed, it is difficult to avoid reading behavioral genetics studies in mainstream clinical, personality and developmental psychology journals such as *Journal of Abnormal Psychology, Journal of Personality and Social Psychology, Developmental Psychology*, or in journals of disciplines that have traditionally been eager to criticize and slow to adopt this approach (e.g., *Journal of Educational Psychology*). The current state of affairs belies the difficult history of behavioral genetics research as applied to psychology and psychiatry.

Curiously, historical assessments of the receptiveness of the general and scientific audience to genetics research in psychology reveal not only outright dismissal, but also over-eager acceptance. The pendulum metaphor as applied to the last half-century, with different swings punctuated by key studies, has been discussed (Plomin & McClearn, 1993). In addition, Kimble (1993) provides an overview of some of the key epistemological issues underlying the history of nature–nurture in psychology. In recent years a critical mass of findings has accumulated to permit a reasoned debate of the conceptual and methodological implications of behavioral genetics research. Many of these debates have been published (e.g., Detterman, 1996; Plomin & Bergeman, 1991, and commentaries; Scarr, 1992). It is indeed encouraging that behavioral genetics research has attracted attention and spurred discussions from psychologists from diverse perspectives (Plomin & McClearn, 1993).

Fortunately, the most recent iteration in the nature–nurture debate has moved beyond discussing *whether* genetic influences are significant in behavioral development. Instead, questions now focus on *how* genetic factors exert their influence. The question of "how" to integrate genetic and environmental factors in developmental theory is not new. Anastasi's (1958) recently exhumed paper raised this issue most directly, and subsequent papers also emphasized the need for a détente or rapprochement between competing genetic and environmental models of development and psychopathology (e.g., Gottesman, 1974; Scarr & McCartney, 1983; Turkheimer & Gottesman, 1991). However, it is only relatively recently that we have been able to come to grips with the empirical analyses needed to answer this question. Some of the important inroads are discussed in Section 1.04.5.

1.04.2.2 Key Studies in Behavioral Genetics Research

Central to the pendulum metaphor noted above are key papers, conferences or books that substantially changed the way in which the behavioral genetics link to clinical psychology was understood. A full discussion of this issue is not within the scope of this chapter, but it is instructive to cite a few examples.

1.04.2.2.1 Genetics and major mental illness

Although genetic influence on schizophrenia (and most major mental disorders) has been widely accepted for many years, this was not always the case. The critical studies to demonstrate the role of genetic factors appeared in the late 1960s and early 1970s (see Gottesman, 1991). As Gottesman (1991) notes, researchers studying schizophrenia were "converted" by the remarkable findings reported by Heston (1966) and others in a ground-breaking conference in 1968.

1.04.2.2.2 Genetics and development

Several key papers ushered behavioral genetics research into "mainstream" psychology and psychopathology research. Behavioral genetics had received some attention in developmental psychology up to the early 1980s, but the influence was largely limited to certain areas such as temperament and cognitive abilities. A special issue of *Child Development* in 1983 (Plomin, 1983) was devoted to a relatively wide range of topics in which behavioral genetics findings offered important insights into developmental processes underlying change and stability and models that sought to integrate genetic and environmental factors from childhood through adolescence. Subsequent publications dealt more extensively with how development could be studied from a behavioral genetic perspective (e.g., Plomin, 1986).

1.04.2.2.3 Genetics and the environment: sibling differences in adjustment

One of the most significant advances in behavioral genetics research concerns not genetics but the environment. Few findings illustrate this point better than the remarkable finding that simply growing up together in the same home does not make siblings (very) similar. Correlations between siblings on a range of personality and other indicators in individual differences research (with the notable exception of cognition) are remarkably weak, and usually less than 0.3. Leaving aside methodological issues such as measurement error, this low correlation suggests that siblings are quite different, and this has lead to the search for factors that make siblings different from one another, that is "nonshared environment." A target article and attached commentaries to Plomin and Daniels's (1987) paper published a decade ago still provides a useful introduction to this issue, and continues to be widely cited in research on family influences, sibling relations and children's psychosocial development.

It is now accepted that nonshared environment, and particularly differential parental treatment, has provided critical insight into the socialization influences within families. Parents do treat their children differently, because of age differences, triangulated family alliances (Vogel & Bell, 1960), or other reasons. Phenotypically, this differential treatment of sibling children has been linked to problems in adjustment. Not surprisingly, a child who has a more negative relationship with his or her parent compared with a sibling, is more likely to have higher levels of behavioral and emotional problems (Dunn, Stocker, & Plomin, 1991; Plomin & Daniels, 1987; Rodgers, Rowe, & Li, 1994). The extent to which the effects of differential treatment are moderated by siblings' perceptions and family context has also been examined (McHale & Pawletko, 1992; Quittner & Opipari, 1994).

Although not usually explicitly tested, the prevailing assumption in social development research was that there was no need to include more than one child per family. Implicitly, the hypothesis was that one child was an adequate "proxy" for processes affecting other children in the family. For example, in research on marital conflict and children's disruptive behavior only one child per family needed to be studied because the focus was on the between-family effects—the association between individual differences in marital conflict and individual differences in children's disruptive behavior across families. However, the implication of nonshared environment is that simply "sharing" parental conflict does not make siblings similar to one another. In addition to encouraging a search for familial and nonfamilial influences that would make siblings different from one another, the possibility was raised that siblings adjusted differently to the same stressor. Several studies now support the role of nonshared environment (Plomin, 1994), and there is a suggestion that siblings are differentially affected by family stresses hitherto thought to be "shared" (O'Connor, Hetherington, & Reiss, in press). Accordingly, it is no longer sufficient to examine just one child per family; assessing the impact of how siblings are similarly or differentially affected by stresses can inform our understanding of how psychosocial risks operate.

1.04.2.2.4 Genetics and the environment: environments are not independent of genetic "effects"

A second major insight concerned the way in which genetic influences were conceptualized: so-called environmental influences are not independent of genetics (Plomin & Bergeman, 1991). That is, environments do not just "happen." Instead, the way in which individuals actively select and respond to their environment is correlated with personality, temperamental, and intellectual abilities that are genetically influenced. In other words, environments are not randomly distributed but correlated with genetic factors (Rutter, 1991).

The ways in which genetic influences shape how individuals approach their environments has been discussed at a conceptual level (Plomin, DeFries, & Loehlin, 1977; Scarr & McCartney, 1983), but there were relatively few empirical examples that directly assessed the hypothesis (Lytten, 1977; Plomin & Rowe, 1979) until the 1980s (Plomin & Bergeman, 1991; Rowe, 1981). Numerous studies published subsequently provide support for their hypothesis (Kendler, 1995, 1996; Plomin, Reiss, Hetherington, & Howe, 1994). Lessons about how genetic factors influence experience are discussed in detail in Section 1.04.5.

1.04.2.2.5 Progress in animal research

The history of behavioral genetics research needs also to include advances made in research using animal models, in which both genotype and environment can be manipulated. The development of animal models has long been considered an important step in studying disorders and diseases such as depression, Alzheimer's disease, alcoholism, and schizophrenia. Genetic research using animal models found evidence for genetic influence in areas such as learning, behavioral effects of drugs, and social behavior, long before similar conclusions were reached in human research (Plomin et al., 1997; Scott, 1977). Animal model research is currently leading the way towards the integration of molecular genetics and psychology, especially in the area of psychopharmacology (Crabbe, Belknap, & Buck, 1994; McClearn et al., 1997), but also in other areas such as fearfulness (Flint et al., 1995).

1.04.2.2.6 Psychometric developments in behavioral genetics

A final area that deserves historical attention is the psychometric progression underlying behavioral genetics research. A brief overview of the key papers and iterations is provided here because behavioral genetics is intensely quantitative, and improved ability to test genetic hypotheses came about only following quantitative advances.

Partialling variance in a phenotypic trait into genetic and environmental factors, the hallmark of behavioral genetics research and an analytic approach that leads to statements regarding heritability, was laid out by Fisher over a half century ago (see Jinks & Fulker, 1970). Falconer (1960) compiled available evidence and distilled the basic elements of research design that formed the basis for quantitative genetics (see Loehlin, 1992). Subsequent contributions that had a major impact on behavioral genetics research were both computational and conceptual in origin. Most notably, Jinks and Fulker (1970) provided the quantitative foundations for contemporary approaches to regression analysis and model-fitting. More recently, DeFries and Fulker (1985) developed a way of examining whether genetic influences are greater for extreme (clinical) groups than for groups within the normal range.

Additionally, because behavioral genetics analyses are computationally intense, the development of software packages has also had much to do with the accessibility of behavioral genetics research approaches (Joreskog & Sorbom, 1989). Moreover, criticisms of the basic assumptions of the twin and other research designs needed to be empirically assessed, such as influences of assortative mating, the equal environments assumption in twin studies, selective placement in adoption studies, and differentiating nonshared environment from error of measurement (McArdle & Goldsmith, 1990).

Conceptual questions regarding longitudinal analyses, analyses of latent traits, and multivariate genetic analyses were also asked, and similarly required new analytic models. Fortunately, the methods to model these sources of influence are now readily available in Neale and Cardon's (1992) comprehensive book.

Finally, molecular genetic research on complex dimensions and disorders required new analytic paradigms from single-gene approaches (Plomin, Owen, & McGuffin, 1994). The essence of these methods is to identify genes in multiple-gene systems that are of modest effect size, which have been called quantitative trait loci (QTL) or susceptibility genes. These advances include QTL linkage approaches that attempt to locate general regions of a chromosome where a QTL resides (Fulker & Cherny, 1996) as well as association analyses that correlate candidate genes with complex traits (e.g., Risch & Merikangas, 1996).

1.04.2.3 Controversies Old and New

Misunderstandings of what research findings from behavioral genetics do and do not imply have led to several controversies, some of which continue to linger. Rutter (1991) outlines many of the misconceptualizations of behavioral genetics research in psychology and psychiatry. One of the most frequent, and perhaps most important, areas of misunderstanding concerns what behavior genetic research examines: individual differences. That is, why do individuals vary, for instance, in their likelihood of approaching and conversing with strangers, performing well on ability tests, or experiencing depressive symptoms.

The magnitude of genetic influence on individual differences does not readily translate into a second paradigm in psychological research: group differences, or why two groups of individuals may differ in, for example, their susceptibility to alcoholism. Despite efforts to clarify the marked contrasts between these approaches, examples of misunderstanding of the distinction between research on individual differences and group differences continue to appear in the scientific and popular press. A notable misunderstanding of genetic methods is made in the attempts to link ethnic differences on ability tests and the findings that scores of ability tests are partly genetically influenced to conclude that group differences in scores are genetically determined. Although our focus is on an individual differences approach to development and psychopathology, it is not the only way that genetic influences operate. Several authors have attempted to integrate research on individual differences with models of species-typical behavior and processes (Buss, 1991).

A second but related source of confusion is heritability, the degree of genetic influence on individual differences. Heritability is a population parameter derived from a particular sample at a particular time. Although estimating heritability is often a "bottom line" of the majority of behavioral genetics research, there is no "true" heritability estimate, just as there is no "true" estimate of the effects of any descriptive statistic such as the correlation between life events and depression. Like any other statistic, heritability can vary depending on sample characteristics, methodology, and historical context.

A further fundamental aspect of understanding heritability is that genetic influences are not fixed, nor is it the case that strongly heritable forms of psychopathology will not respond to environmental input. Thus, there is no necessary relationship between genetic influence and psychosocial susceptibility. Numerous examples of secular changes in aspects of development and psychopathology help illustrate this point (e.g., see Rutter & Smith, 1995).

1.04.2.4 Summary

Not surprisingly, controversy remains regarding the conceptualization of genetic influences in the development of psychopathology. Fortunately, however, there is now good reason to suspect that the dialogue between behavioral geneticists, developmental psychologists, and clinical psychologists and psychiatrists is increasing and has reached a critical mass (Reiss, 1995). The kinds of conceptual questions now asked in behavioral genetics research are more sophisticated than those asked a decade ago, and there is far greater interest in combining analytic strategies to test competing theories of psychopathological processes (Rutter et al., 1997). Moreover, there are a number of investigations currently underway that are likely to provide important clues of how genetic factors are involved in the development of psychopathology (Reiss et al., 1995). In the remainder of the chapter we examine these advances and highlight new research directions.

1.04.3 A SELECTIVE OVERVIEW OF METHODOLOGICAL ISSUES IN BEHAVIORAL GENETICS RESEARCH

As in every other branch of psychological research, behavioral genetics investigations are limited by methods and design. The mechanics of the designs used in behavioral genetics research have been discussed widely (Loehlin, 1992; Plomin et al., 1997; Rutter et al., 1990), and will therefore receive only limited attention here. Rather than discuss the actual methods in detail, we instead emphasize the hitherto largely underappreciated relative strengths and weaknesses of certain designs in light of current models of genetic and environmental influences on psychopathology.

1.04.3.1 A Précis of Quantitative Genetics

The first point to make is that genetic influences, or the role of individual differences in genotype, are estimated based on phenotypic or observed patterns of similarity in genetically informative designs that vary genetic or environmental relatedness. Given the frequent misinterpretation of behavioral genetics findings, we briefly discuss how empirical findings allow us to make inferences regarding the role of genes and the environment.

Behavioral genetics decomposes individual differences in the variation of behavior into genetic, shared environment, and nonshared environmental sources of influence. Genetic influence refers to inherited genetic differences that contribute to observed (phenotypic) differences among individuals. For identical twins, these genetic factors contribute entirely to their phenotypic similarity, while for first-degree relatives (parents, offspring, and siblings), who are 50% similar genetically, additive genetic factors contribute to phenotypic similarity and differences in equal measure.

One way to determine genetic influence is to compare identical (monozygotic or MZ) and fraternal (dizygotic or DZ) twins. If identical twins are more similar than nonidentical twins, it follows that they may be more similar because they are more similar genetically. The extent to which genetic influences are important can be quantified because we know that MZ twins share all of their genes and DZ twins are, on average, 50% similar genetically. Doubling the difference between MZ and DZ correlations estimates heritability, or h^2, the proportion of phenotypic variance that is explained by genotypic variance. Comparing adopted-apart MZ and DZ twins (that is, pairs that were raised in different homes) provides a rare but even more powerful way of estimating genetic influences, that "controls" for shared family environment.

Shared environmental influences can also be straightforwardly assessed. The magnitude of shared environment is most easily seen by comparing unrelated siblings, as in the adoption design. The adoption design is particularly powerful because it "controls" for genetic similarity (i.e., adopted siblings are genetically unrelated). Because unrelated siblings do not share genes, any similarity must be attributed to common environment. Finally, the magnitude of nonshared environment is most easily seen in the differences within pairs of identical twins. Because MZ twins are genetically identical, any differences must be attributed to different experiences (and measurement error).

As a general point, it is noteworthy that whereas it is possible to examine the magnitude of the three parameters assessed in behavioral genetics research (genetic, shared and nonshared environment), different designs have different strengths and weaknesses and are differentially affected by, for instance, gene–gene interactions (which are included in the similarity of MZ twins because they are genetically identical for all genes) and gene–environment correlations (which are differentially assessed in adoption and twin studies). This helps explain why different results are sometimes obtained from twin and adoption designs and why no design should be taken as the absolute estimate of these influences.

1.04.3.2 An Overview of Research Designs to Detect Genetic and Environmental Influence

1.04.3.2.1 Twin studies

The most frequent form of behavioral genetics design is the twin study. Historically, twin studies have provided critical initial findings that have fundamentally altered debates regarding genetic influence on schizophrenia (Gottesman, 1991) and autism (Folstein & Rutter, 1977), among other disorders. To date, twin studies provide data on a wide range of psychiatric and psychological disorders and dimensions virtually across the lifespan (Bouchard & Propping, 1993; Eaves et al., 1997; McGuffin et al., 1994). As noted above, studying twins is perhaps the best way to assess the role of nonshared environment independent of genetic effects. Pike, McGuire, Reiss, Hetherington, and Plomin (1996) exploited this issue directly in their search of nonshared environmental influences that underlie MZ differences in behavioral adjustment in adolescence. They found that differential parental negativity and warmth could explain differences between MZ twins in rate of problem behavior.

Further examples of exploiting discordant MZ twin pairs has provided important insights into which biological substrates associated with psychopathology might (or might not) be under genetic control. For example, Torrey, Bowler, Taylor, and Gottesman (1994) reported that identical twins discordant for schizophrenia also differed in brain structures implicated in its neuropathology. This finding suggests that genetic influences on schizophrenia are unlikely to be mediated through genetic effects on brain pathology.

1.04.3.2.2 Adoption studies

Adoption studies have an equally revered place in behavioral genetics research on psychopathology. For instance, it was the finding that children of schizophrenic mothers who were adopted away (and living with nonschizophrenic families) were at increased risk of developing schizophrenia that lead to a major shift in how the causes of schizophrenia was conceptualized (Heston 1966; see Gottesman, 1991). Furthermore, adoption studies based on registries in Denmark and Sweden have proved indispensable in our understanding of genetic influences on alcoholism and antisocial behavior, among other disorders (Cloninger, Bohman, & Sigvardsson, 1981). Adoption designs are also well suited to assess genotype–environment correlations and interactions, a topic discussed below.

1.04.3.2.3 Family history studies

Family history designs, which also have a long history in psychiatric research, examine the extent to which family members, usually first degree relatives of a proband, or the "target" patient, are affected by a disorder. The central drawback of the family history method is that it is impossible to explain why, in terms of genetic and environmental influences, family members may be similarly affected. Depression, for instance, may run in families because family members share genes or because they share the psychosocial stresses associated with depression (low socioeconomic status, family conflict). Nonetheless, family history studies have made several important contributions to our understanding of genetic influences on psychiatric conditions, for example in the area of heterogeneity (see Coryell, 1997) and age of onset (e.g., Weissman et al., 1984) among others. A final advantage of the family history method is that it allows for the test of genetic hypothesis without requiring that investigators ascertain a sample of twins or adoptive families. It is therefore possible that a genetic component could be added to existing studies of environmental risk for psychopathology. Psychopathology in first-degree relatives can be used as a proxy for genetic risk (although, of course, it also indexes environmental risk), and it would then be possible to assess the extent to which there is overlap between genetic and environmental risk factors for psychopathology.

1.04.3.3 Alternative Research Designs

The designs noted above are the most common, but no means the only way of assessing genetic influence. A recent addition to the armory of behavioral genetics designs is the half-sibling or stepsibling design. The viability of the stepsibling model follows from two major demographic changes affecting families since the mid 1970s. First, the number of adoptions has steadily decreased, and it is especially difficult to ascertain representative samples of adoptive children and adoptive families. The second demographic trend is that the number of divorces and remarriages and, consequently, the number of half- and unrelated siblings in reconstituted families has increased dramatically (see Reiss et al., 1994). The stepsibling design therefore has the unique advantage of uniting the diverse perspectives of behavioral geneticists and family researchers.

From a quantitative genetic perspective, the half- and stepsibling design is as informative as a twin or adoption design because it is possible to estimate the degree to which sibling similarity is explained by known genetic similarity (full siblings are 50% similar genetically, half-siblings are 25% similar genetically). There is evidence that this approach may be a reasonable addition to twin and adoption studies. The most encouraging evidence comes from research on intelligence. When similarity of relatives' intellectual ability is assessed across numerous types of consanguinity, results from half-siblings and cousins complements and does not detract from the overall pattern of results (Bouchard & McGue, 1981). Qualified encouragement for the half/unrelated sibling design comes from the Nonshared Environment and Adolescent Development project (NEAD) (Reiss et al., 1994), the first study to combine the twin and stepsiblings in the same design (Neiderhiser, et al., in press). Other studies employing the half-sibling design have endorsed this strategy for other behaviors (e.g., Van den Oord & Rowe, 1997).

The stepsibling design provides a further parallel to the adoption design in that it is possible to ascertain the magnitude of genetic effects by comparing biological parent–child correlations with stepparent–child correlations. In quantitative genetic terms, the adoptive and stepparent–offspring correlations are identical in one respect, i.e., the correlation can be explained only by shared environmental influence.

There are, of course, a number of limitations that complicate interpretation of the stepsibling model results, including potential assortative mating differences in first and second marriages, variation in length of time sharing a household, and elevated stress in stepfamilies. However, there are clear benefits of this strategy, such as examining how environmental factors (e.g., family context) modify genetic influences.

The final point regarding research design is that studies of biological families can tell us surprisingly little about mechanisms and processes in development (Scarr, 1992). The reason is that, as reviewed in Section 1.04.5, findings from biological families cannot distinguish between genetic and environmental processes, nor can they identify the ways in which these influences correlate and interact. Thus, behavioral geneticists and psychosocially minded investigators need to be equally concerned about the use of genetically informative designs to answer questions of causality (Plomin, 1983).

1.04.3.4 Methodological Criticisms of Behavioral Genetics Methods

Criticisms have been raised, and are continuously raised, about the methods used to make conclusions about the impact of genetic factors on development. We briefly review these criticisms and outline the implications for clinical research.

One criticism of the twin method concerns the "equal environments" assumption, that is MZ twins are assumed to experience environments that are no more similar than DZ twins. We now know that MZ twins do in fact receive a more similar familial and extrafamilial environment than DZ twins (Plomin, 1994). In this sense, then, there is not equal similarity in the environments experienced by MZ and DZ twins. But the issue is not whether this is the case, but whether parental differences in treatment are the cause or the effect of twins' differences in behavior (Lytton, 1977). Only if the former is true is the equal environments assumption violated.

A separate issue is whether there are features about being a twin that are fundamentally different from being a singleton. If so, then again we would be concerned about generalizing findings from twin studies. Relatively few factors distinguish twins from nontwins. For example, there is a suggestion that twins may be more likely to exhibit language difficulties and experience perinatal complications (Rutter & Redshaw, 1991), and these factors may have implications for the development of psychopathology. Some studies also suggest that twins appear to be at slightly elevated risk for conduct or other behavioral problems compared with singletons, but other studies suggest the opposite (see Simonoff, 1992). Additionally, mothers of twins are more likely to experience depression than parents of closely spaced children (Thorpe, Golding, MacGilliuray, & Greenwood, 1991), and there may be twin-specific processes that increase the risk for antisocial behavior, such as imitation (Carey, 1992). This is an important finding given the association between maternal depression and children's adjustment problems. To be sure, the assumption that findings from twins can be generalized to the larger population should not be accepted uncritically, and further research is warranted (Rutter, Simonoff, & Silberg, 1993), but there are good reasons for accepting twin study results in research on psychopathology. A more recent methodological issue in twin studies is chorionic status, as MZ twins can be either monochorionic or dichorionic but DZ twins are always dichorionic.

A number of criticisms have been raised regarding adoption study design as well. Perhaps the most important of these concerns the association between the biological and adoptive family characteristics, or selective placement. Selective placement in which children from high (or low) genetic risk are placed into families at high (or low) environmental risk confounds these sources of risk that adoption designs are designed to disentangle.

Assortative mating, the correlation between spouses, is an issue relevant to both twin and adoption studies. The net effect of assortative mating for genetic studies is that it will deflate heritability estimates in twin studies (because DZ twin correlations will be increased but MZ correlations will be unaffected) and increase parent–offspring correlations between one biological parent and offspring. Although assortative mating is not usually found for personality (see Loehlin, 1992), it is common for educational attainment. For psychopathology, there is an elevated concordance between husbands' and wives' levels of psychopathology for depression and antisocial behavior (e.g., Kreuger, Moffitt, Caspi, & Bleske, in press), but it is important to consider that assortative mating does not necessarily imply assortment on the same form of psychopathology (see Mednick, 1978).

A further methodological concern is that different research designs often produce different estimates for genetic and environmental influences. Of course, there is no reason to expect that different designs would necessarily produce the same results. Nonetheless, attempts to include data from adoption and twin studies are relatively rare (Bouchard & McGue, 1981; Miles & Carey, 1997). Where differences are found among research designs, the general pattern is for twin designs to produce the largest genetic effects. Of course, it may be that the greater genetic effects derived from twin studies is a real finding, attributed to non-additive genetic influences and epistasis that are best assessed by studying twins. That there is variation among different genetically-informative designs, and that this variation may be both substantive and methodological in nature, cautions against over-reliance on a single research design and the need to incorporate multiple designs in the same research program.

1.04.3.5 Methodological Recommendations

Just as behavioral genetics research findings provide clues to the role of both genetic and environmental factors, findings from these studies also suggest methodological recommendations for future studies of mechanisms in development and psychopathology.

1.04.3.5.1 Multiple measures

Methodological problems such as rater bias and other forms of measurement error in behavioral genetics research complicate the interpretation of findings. This is particularly true in research on children and adolescents, in which it is customary to gather information from multiple sources. Were multiple sources of data routinely collected in research on adult psychopathology, the same problems would inevitably result. Rater effects have predictable influences on the relative magnitude of genetic and shared and nonshared environmental estimates. Specifically, compared with child self-report data, there is a general tendency for parent ratings to indicate more shared environment and less nonshared environmental effects; the differences in the magnitude of the genetic parameter among raters is somewhat less predictable. In a recent meta-analysis of behavioral genetics studies of aggression, Miles and Carey (1997) found that methodological differences across studies was as important a contributor to the overall pattern of results—almost as large as the genetic parameters! In other words, there is almost as much variation attributed to methods as to genetic influences (although this was true more for the shared environmental parameter than the genetic parameter). A somewhat exaggerated example of rater effects is given by Thapar and McGuffin (1995). They reported that the heritability of anxiety symptoms in children was 59% according to parent reports but 0% according to children's self-reports! The low agreement invariably found among different reporters in research on psychopathology in childhood is not just an issue of rater bias, because data from different reporters may lead to different substantive findings.

However, even when multiple sources of data are simultaneously assessed—an attempt to analyze an error-free latent construct defined by what is common among raters—the results are not necessarily any clearer. An example of this approach is provided by Simonoff, Pickles, Hewitt, and Silberg (1995), who examined adolescent self-reports and mother and father reports of adolescent behavioral problems. In addition to assessing the genetic–environmental pattern for the common variance among reporters (i.e., the latent trait), these investigators also examined the genetic–environmental pattern *specific* to each rater. Thus, these authors hypothesized that genetic (and environmental) influences on children's behavioral problems could be detected by mothers but not by children or fathers. It is not clear, however, how to interpret mother-specific genetic effects on children's behavioral problems. To date, no satisfactory way of handling method effects has been proposed. What is clear is that single-reporter designs may suggest misleading substantive conclusions, particularly in research on childhood psychopathology. Consequently, multiple-reporter designs should be considered and results based on different designs compared.

1.04.3.5.2 Specification of individual-specific environmental risks

A wealth of research demonstrates that environmental risks that are shared by siblings, or what is termed "shared environment", play little role in the development of psychopathology. The one robust exception to this rule is conduct disorder and aggression, for which there is substantial evidence that simply growing up in the same home (and the concomitant sharing of psychosocial risks such as poverty and neighborhood violence) makes siblings similar beyond the effects of genetics (McGuffin et al., 1994; O'Connor, McGuire, Reiss, Hetherington, & Plomin, 1998; Plomin et al., 1997). In general, however, environmental influences that are salient for development appear to make siblings different from one another.

A central implication of this finding for research is that risk needs to be conceptualized as individual-specific. That is, there is a need to assess the actual impact of the putative risk factor on the individual's construction of events. For example, rather than define parenting risk in terms of global parent attitudes, it is important to assess the parent's attitudes towards the specific child, how the parent acts toward the specific child, how parenting to the target child differs from parenting experienced by siblings. The bottom line of much of the research on nonshared environment is that risk processes occur on an individual-by-individual basis and not on a family-wide level (Plomin, 1995). A corollary is that it is not sufficient to equate exposure to risk with the effect of the risk.

1.04.3.6 Summary

A number of strategies are available to estimating the effects of genes and environment in behavioral development. Previous behavioral genetics research has underscored the methodological strengths and weaknesses of each approach and has illustrated a number of lessons for researchers seeking to integrate genetic and environmental mechanisms in the development of psychopathology.

1.04.4 THE CONTRIBUTIONS OF BEHAVIORAL GENETICS TO RESEARCH ON CLINICAL PHENOTYPES

Having discussed the historical change in behavioral genetics research and some of the central methodological issues in interpreting behavioral genetics findings, we now turn to an overview of the ways in which behavioral genetics research has contributed to our understanding of clinical phenotypes. Our review focuses on five areas: etiology, the continuity/discontinuity debate, comorbidity, heterogeneity, and stability and change.

1.04.4.1 A Selective Review of Genetic and Environmental Etiology in Psychopathology

Genetic influences appear to play an important role in the development of most forms of psychopathology. For example, research consistently shows higher concordance of schizophrenia among MZ than among DZ twins and elevated rates of psychopathology among children adopted away from disordered biological parents, to name just two examples. As we discuss in more detail below, the major questions for genetic research in psychopathology are no longer whether or not (or how much) a disorder is genetic, but what features of a disorder are genetic (e.g., onset and course, heterogeneous subtypes, response to treatment), how does genetic research inform nosology and treatment, what is the nature of genetic risks (e.g., do they influence disorder, core symptoms, or other risks), and how do genetic and environmental risks jointly lead to maladjustment. These questions shape the discussion in this and the following section of this chapter.

Rather than discuss the forms of psychopathology that show genetic influence (this would take an inordinate length of space, and is covered in chapters on specific manifestations of psychopathology), we highlight the more recent and more challenging findings. Several texts and papers describe in detail genetic influences on a range of disorders in childhood and adolescence (for depression, Tsuang & Faraone, 1990; for schizophrenia, Gottesman, 1991; Karayiorgou & Gogos, 1997; for aggression, Bohman, 1996). General overviews are provided by Plomin et al. (1997), and more specific volumes are also available (Gershon & Cloninger, 1994). What follows is an overview of the key conceptual questions for understanding psychopathology for which behavioral genetics research provides unique insights.

It may not be a surprise to learn that many forms of psychopathology are genetically influenced, but it may be surprising to learn how early in development behavioral problems can be attributed to genetic etiology. Several studies identify relatively robust genetic influence on externalizing and internalizing symptoms from age 2–3 (e.g., Van den Oord, Verhulst, & Boomsma, 1996). It is not yet known how (or if) genetic influence on behavioral problems can be differentiated from the genetic influences on temperament at this age, to what extent the results may reflect strong parental reporter effects (i.e., self-reported symptoms of psychopathology are not reliable before late childhood), or the degree to which the picture might be different if extreme cases were included (the above studies were based on normal risk groups and included few "cases"). Nonetheless, the replicated findings that early behavior problems show genetic influence challenges the notion that early forms of problem behavior are entirely determined by family environmental factors. These reports provide a good example of how research that simply shows genetic influence can challenge prevailing views of the nature of psychopathology.

Nor is it the case that genetic influences disappear in older age. Perhaps the best example is provided by research on depression. McGue and Christensen's (1997) study of Danish twins 75 years or older indicated strong genetic influence on both affective and somatic symptoms. These findings are noteworthy because they challenge the notion that somatic symptoms of depression index physical impairments in old age rather than depression *per se*. These findings also contradict the hypothesis that—for a variety of reasons, including the loss of friends and loved ones from death, increasing social isolation, and difficulty adjusting to physical limitations—depression in older age is more psychosocially determined.

1.04.4.2 Continua and Categories in Models of Psychopathology

Although diagnostic or categorical approaches have dominated phenotypic research on psychopathology, there is increasing interest in comparing dimensional and categorical models of disorder. Continuous phenotypic dimensions or syndromes have been demonstrated for a wide variety of medical disorders in addition to psychiatric forms of disorder, including personality disorders (Livesley, Jang, Jackson, & Vernon, 1993; Nigg & Goldsmith, 1994) and depression (Kendler, Neale, Kessler, Heath, & Eaves, 1992a).

The debate about dimensions vs. diagnoses has attracted the attention of genetic investigators in recent years because at the core of this debate lies a fundamental question of etiology: do the same genetic and environmental factors that operate at the extremes also affect the entire continuum in the population? Given the novel findings that behavioral genetics research is contributing to this debate and the important clinical implications derived from these results, we provide a detailed overview of the conceptual and methodological issues in this area.

1.04.4.2.1 Distinguishing between the underlying assumptions and quantitative implications of categorical and dimensional models

The dimensional method assesses quantitative variation throughout a continuous distribution of symptoms. By definition, individual differences one standard deviation below the mean are just as informative as 1, 2, or 3 standard deviations above the mean. The diagnostic approach, on the other hand, emphasizes average differences between an "extreme" diagnosed group and a nondisordered group.

The two approaches require different statistical methods. A quantitative genetic analysis of individual differences in dimensions of psychopathology, which is by far the more straightforward of the two, is based on correlations. In the example of a twin design, one would hypothesize that identical twins would be more similar than fraternal twins. Alternatively, the diagnostic approach is based on twin siblings' concordance for disorder. The genetic effect size or heritability is not estimated from concordances alone, but also takes into consideration base rates in the population. The concordance is translated into *liability* correlations that assume a continuous distribution of risk. This procedure assumes, paradoxically, a continuum of liability despite the categorical measurement. This liability-threshold model assumes that the genetic and environmental contributions to the disorder are continuously distributed, but that disorder only occurs when a certain threshold of liability is exceeded.

An additional statistical approach, referred to as DF extremes analysis (from *D*eFries & *F*ulker, 1985), integrates quantitative genetic research on dimensions and disorders. This approach requires continuous measurement of behaviors that are relevant to the disorder of interest among the cotwins of selected, or diagnosed, probands. Familial resemblance is estimated as the extent to which the mean cotwins' score on this quantitative measure regresses back to the population mean. If the cotwin mean regresses all the way back to the population mean, the *group* correlation is 0; if the cotwin mean is the same as the proband mean, the group correlation is 1. Group correlations for identical and fraternal twins can then be compared to estimate *group heritability*, or the degree to which genetic influences are associated with being in the extreme group.

There is a fundamental difference between the DF extremes method and the liability-threshold method already described. In contrast to the liability-threshold model, group heritability is not based on the assumption that the disorder is manifested only after a certain threshold is reached. On the contrary, disorder-relevant symptoms are assumed to increase continuously from the normal to the abnormal "levels" of the disorder. Group heritability assesses the contribution of genetic and common environmental factors to the average quantitative difference between the selected probands and the population.

Although the liability-threshold and DF extremes models are based on different underlying assumptions, they do not necessarily yield different empirical results. Indeed, if the assumption of the liability-threshold model is correct (i.e., there is an underlying continuum of liability for the disorder), and if the quantitative dimension measured in the DF extremes model assesses the actual continuum underlying the disorder, then the heritability of liability in the liability-threshold model will be very similar to the group heritability in the DF model.

Determining whether genetic influences are more salient at the extreme is complex. For instance, most approaches require extremely large samples (of mild and severe impairment) in order that there is sufficient power to detect even large differences. Additionally, these approaches determine only if the magnitude of genetic influence varies across different levels of impairment; it is not possible to determine from the quantitative methods whether *different* genes are involved in the normal and more extreme levels of impairment; molecular genetic studies are needed for this.

To date, there is only limited evidence that directly addresses the question of whether genetic influence is different at the extremes; most reports do not find differences. For example, using DF analyses, Deater-Deckard, Reiss, Hetherington, and Plomin (1996) found that genetic influence at the extremes was similar to genetic influence on normal variation in parent- or adolescent-reported behavior problems. Gjone et al. (Gjone, Stevenson, & Sundet, 1996; Gjone, Stevenson, Sundet, & Eilertsen, 1996) reported some evidence for greater heritability on more extreme levels of behavioral problems, but this effect depended on how the scale scores were analyzed. Other studies have compared heritability across different definitions of disorder. Kendler et al. (1992a) reported similar heritabilities across different definitions of depression in women, and Slutske et al. (1997) reported identical genetic and environmental influences using broad and more severe definitions of conduct disorder in their study based on retrospective reports from the Australian twin registry. In general, then, the genetic evidence available so far indicates that there is a continuity of genetic influence between the normal and the abnormal (Plomin, Nitz, & Rowe, 1990).

1.04.4.2.2 Are shared environmental influences more salient at the extremes?

A recurring theme throughout this chapter is that genetic studies provide unique clues to understanding the mechanisms through which environmental risks operate. This is also the case in the context of the debates on whether genetic influences affect the continuum of psychopathology. Using the same approaches defined above (e.g., DF analysis), it is possible to assess the extent to which shared environmental influences are greater in extreme, diagnosed groups compared with nondiagnosed, less impaired groups. As in the case of research on genetic influences at the extremes, there are few studies that examine whether environmental effects differ at the extremes compared with the normal range, and available data suggest that the differences are minimal (Deater-Deckard et al., 1996; Eaves et al., 1993).

1.04.4.2.3 Continuum, boundaries, and spectra

A further nosological question related to the debate of categorical and continuous measurements concerns the boundaries around a disorder (this also relates closely to the topic of heterogeneity, which is reviewed next). Behavioral genetics research findings indicate that even with major mental disorders, the phenotype extends more broadly than the traditional diagnostic boundaries. For example, it is clear that first-degree relatives of schizophrenic individuals are not only more likely to exhibit schizophrenia but also schizophrenic-related symptoms such as neurological impairments and schizoid symptoms (Kendler, Gruenberg, & Kinney, 1994). A parallel pattern is found among first-degree relatives of autistic individuals, who exhibit elevated levels of pervasive developmental disorder (Bailey, Phillips, & Rutter, 1996; Bolton et al., 1994; Rutter, Bailey, Simonoff, & Pickles, in press).

An additional example, which complicates matters even further, is that the genetically influenced liability to psychopathology may not concern the disorder *per se*, but associated personality features—as in the case of neuroticism and depression (Kendler, Neale, Kessler, Heath, & Eaves, 1993). In this instance, it is not clear whether associated personality features constitute: a core, less acute condition of the disorder; a distinct feature that increases the risk for the disorder; or a consequence of experiencing a disorder. Genetic research findings may play an important role in nosology, but it is equally true that confusion regarding the clinical phenotypes may lead to misinterpretations of behavioral genetics findings.

1.04.4.3 Comorbidity of Psychopathologies

One of the most robust findings in psychiatric research is that disorders covary at rates far higher than would be expected by chance. This is illustrated by numerous epidemiological reports, such as the National Comorbidity Study (NCS) (Kessler et al., 1994). It is generally found that a minority of individuals account for the majority of disorders. The degree of overlap between and among dimensions of psychopathology is found regardless of whether diagnostic or dimensional indicators are examined. Several explanations have been proposed to explain remarkably–elevated rates of cooccurring disorders or dimensions (Caron & Rutter, 1991). Higher than expected rates of comorbidity or symptom cooccurrence may be explained by: diagnostic overlap of individual symptoms (i.e., a particular symptom is included under more than one diagnosis); the possibility that one disorder may be a risk factor for developing a second disorder; different disorders may be overlapping because of common psychosocial risks; diagnostic schemas may favor multiple diagnoses (*DSM-IV*) over identifying a central diagnosis (*ICD-10*). Other explanation, however, are also possible.

The degree to which two (or more) disorders or dimensions may be correlated because of a shared genetic liability has been examined only recently. Virtually all available evidence indicates that genetic influences are not diagnosis- or dimension-specific. Rather, the nearly universal phenotypic correlations between different dimensions of psychopathology can be explained, in part, by shared genetic factors, called genetic correlations.

A genetic correlation underlying two (or more) syndromes is possible if each syndrome is genetically influenced and the syndromes are correlated phenotypically. The quantitative basis for analyzing genetic correlations is described in Neale and Cardon (1992). In simple terms, genetic influence on the correlation between two disorders is based on the cross-twin cross-correlation. In the case of antisocial and depressive symptoms, genetic influence on the correlation between these two dimensions is suggested if the correlation between twin A's antisocial behavior and twin B's depressive symptoms is greater in MZ than in DZ pairs (or other genetically informative comparison).

In a series of studies on adult female twins in the Virginia twin registry, Kendler and colleagues reported a significant genetic correlation between depression and Generalized Anxiety Disorder (Kendler, Neale, Kessler, Heath, & Eaves, 1992d), alcoholism (Kendler, Heath, Neale, Kessler, & Eaves, 1993), and smoking (Kendler, Neale, MacLean, et al., 1993). In the case of major depression and generalized anxiety disorder, Kendler and colleagues reported a genetic correlation of 1.0 between major depression and generalized anxiety disorder. In other words, the same genes make one vulnerable to depressive and anxious symptoms. The finding of a genetic correlation of 1 is, however, an exception to the general rule that the genetic correlation between correlated dimensions is usually more moderate.

Research findings from psychopathology in childhood and adolescence lead to similar conclusions. For example, the robust phenotypic correlation between depressive symptoms and antisocial behavior in children and adolescents is well known, and there is increasing evidence that both dimensions are under some genetic influences (for depression, see Silberg et al., 1994; for antisocial behavior, see Gottesman & Goldsmith, 1994). A genetic correlation between these dimensions was recently found in a large national study of 720 families using a multi-method measurement design (O'Connor, McGuire et al., in press). Approximately 50% of the phenotypic correlation between these dimensions could be explained by shared genetic influences. A follow-up study conducted approximately two years later found roughly the same results (O'Connor, Neiderhiser, Reiss, Hetherington, & Plomin, 1998). It is important to note that the study also found genetic influences unique to each dimension. Interestingly, the same pattern of genetic and environmental influences was found for boys and girls despite marked mean differences in these dimensions of maladjustment. Data from a large study of juvenile twins in Virginia examined the genetic correlation underlying hyperactivity and oppositional/conduct problems (Silberg et al., 1996). Results of that study indicated that the correlation between these two dimensions was explained largely from the genetic influences common to both syndromes, although the pattern varied slightly by age and gender. Family history data provide an additional source of support for genetic factors underlying different disorders. In this approach, evidence for genetic influences underlying cooccurring disorders is suggested if first-degree relatives of depressed probands are at elevated risk for depression, and for a second disorder, e.g., alcoholism, compared with first-degree relatives of nondepressed probands. Such data are taken as evidence for cosegregation of depression and alcoholism. Cosegregation among disorders is relatively common (see, e.g., duFort, Kovess, & Boivin, 1994; Feng & Baker, 1994; Mednick, 1978) and has also been demonstrated in research on child probands (Puig-Antich et al., 1989). It is important to note, however, that some studies that suggest cosegregation of disorders do not always control for the fact that probands may exhibit multiple disorders.

However, just as there is mounting evidence of nonspecificity of genetic risks for psychopathology, there is also evidence, in some cases, for specificity. Most of the available evidence suggesting specificity of genetic risks derives from family history studies. Among the better known examples is the finding that first-degree relatives of schizophrenic individuals are at greater risk for schizophrenia and schizophrenia-related conditions but not at greater risk for bipolar disorder. This supports the hypothesis, which is also suggested from available molecular genetic studies (see below), that schizophrenia and bipolar disorder are not influenced by the same genes (see, e.g., Kendler et al., 1994; Tsuang & Faraone, 1990). Similar methods also provide convincing evidence that the genetic influences underlying liability to bipolar disorder are not linked to those associated with unipolar depression (see Tsuang & Faraone, 1990; Weissman et al., 1984), suggesting that the same genetic influences are not shared by all affective disorders.

Unfortunately, too few studies have systematically examined cooccurrence among multiple (i.e., more than two) disorders, so it is not yet known whether the same genetic influences may help explain that, among affected individuals, the average number of diagnoses per individual is greater than two (Kessler et al., 1994).

1.04.4.3.1 Comorbidity of psychopathologies: environmental evidence

There is little specificity in the relationship between environmental risks such as interpersonal stress or absence of social support and particular manifestations of psychopathology (Wrate, Rothery, McCabe, Aspin, & Bryce, 1995). Behavioral genetic research has contributed to this line of research by also demonstrating that, controlling for genetic factors, there is evidence that two (or more) disorders or dimensions may cooccur because of common shared and nonshared environmental risks. In the study examining antisocial behavior and depression noted above (O'Connor, McGuire, et al., 1998), there was a small but significant effect of shared environment that contributed to siblings' similarity in both syndromes. That is, by virtue of sharing the same household, siblings share familial experiences that function to make them similar to each other in their levels of both behavioral dimensions (cf. Downey & Coyne, 1990). Of course, studies that rely on only model-fitting cannot identify which shared environmental factors underlie the correlation between behavioral dimensions, but the key advantage is that they are able to identify environmental influence *independent* of genetics.

Similarly, individual-specific environmental risk factors will contribute to the covariation between disorders in multivariate behavioral genetics analyses to the extent that siblings who differ in their experience of environmental risks also differ on both dimensions of psychopathology. An example is provided by the study of antisocial behavior and depressive symptoms in adolescence noted above (O'Connor, McGuire, et al., 1998).

1.04.4.4 Genetic Heterogeneity

The concept of heterogeneity of disorder is a focus of increasing clinical research attention. Heterogeneity is an important topic to study because a clear understanding of the diverse ways in which a disorder may manifest is central to optimally effective treatment and prevention. Notwithstanding the significance of the concept, there has been little progress in how it should be defined. Recent papers on the topic define heterogeneity in diverse ways including presumed etiology, biological correlates, response to treatment, longitudinal course, and association with other disorders. These different definitions likely lead to different conclusions regarding heterogeneity. Moreover, there are few attempts to examine multiple alternative definitions within the same study.

Some of the most important data that focuses on heterogeneity are based on the work of Winokur, Coryell, and colleagues. The "Iowa 500" family study has provided evidence that familial influences are associated with certain subtypes of depressive disorder (see Coryell, 1997 and references therein). Unfortunately, there have been few attempts to replicate these subtypes, and it is not known whether psychopharmacological treatments are differentially effective for different subtypes (but see Sigvardsson, Bohman, & Cloninger, 1996, for the case of alcoholism). A further example of using genetic data to identify subgroups of affected individuals is provided by DiLalla & Gottesman (1989), who hypothesized that whereas adolescence-limited and late-onset delinquency may derive from largely psychosocial factors, antisocial behavior exhibited by individuals who exhibit a stable antisocial trajectory may be genetically mediated (Moffitt, 1993).

Behavioral genetics methods can be used to assess whether different symptom clusters of the same disorder are influenced by the same genes, and whether subgroups of disordered individuals can be identified by degree of genetic risk. In a recent study of depression in Danish twins 75 years or older, McGue & Christensen (1997) tested the hypothesis that two factors of depression in the elderly, somatic and affective symptoms, were influenced by the same genetic factors. They found that despite being clearly distinguishable at a phenotypic level, the two factors of depression were nonetheless explained by the same genetic factors. In other words, there were no genes that were specific to somatic symptoms or affective symptoms. In this way, these authors provide a model for further research to examine whether the heterogeneous manifestations of a dimension of psychopathology have different genetic and environmental origins.

An opposite example is provided by reports on aggression and delinquency, two aspects of conduct disorder. These two behavioral syndromes are distinct, but nonetheless are highly correlated. However, despite the considerable overlap, two studies reported that aggression is strongly genetically influenced but delinquency is under minimal influence (Edelbrock, Rende, Plomin, & Thompson, 1995; Eley, Lichenstein, & Stevenson, in press).

1.04.4.4.1 Age of onset and other features

Behavioral genetics methods have also been used to examine whether certain features of a disorder are under genetic control. Family history data have suggested that younger age of onset is a proxy for stronger genetic influence (Todd, Neuman, Geller, Fox, & Hickok, 1993; Weissman et al., 1984). In general, probands with an earlier age of onset tend to have more first-degree relatives with a similar disorder than those with a late onset. However, although many studies have found that age of onset is correlated with genetic risk, not all studies replicate this pattern, or they report that the pattern holds for adolescent but not childhood onset probands (cf. Harrington et al., 1993).

More convincing support implicating genetic influences on the age of onset is reported by Kendler, Neale, Kessler, Heath and Eaves (1992). In their large sample of adult female twins, they were able to identify a large number of identical and fraternal twins concordant for major depression. When they examined how these concordant, affected cotwins manifested depression, they found that depressed identical twins were no more similar in the likelihood of exhibiting particular symptoms of depression. However, concordant depressed identical twins were more similar than concordant fraternal twins in the age of onset of disorder. A conclusion drawn from these data is that genetic influences underlie the likelihood of experiencing a disorder and perhaps some features such as age of onset, but genetic influences appear not to directly influence the particular pattern of symptoms the disorder takes.

1.04.4.5 Genetic Influence on Change and Stability of Psychopathology

A critical developmental question that awaits firm answers is the extent to which genetic influences mediate the marked stability of disorder from childhood through adulthood. A related question is whether genetic influences on psychopathology change with age, pubertal status, or other markers of development. In other words, to what extent are genetic influences on psychopathology in adolescence and adulthood different from those detected in childhood. It is not that an individuals' genes change, but that genes can have different effects at different points in development.

Consider, for example, the finding that bipolar and schizophrenic disorders are rare before puberty, and may first emerge during adulthood and even into older age. These disorders are known to be under genetic influence, so it seems likely that the genetic mechanisms are not apparent in early childhood (although there is little research addressing this issue) but come "on line" only later in development. Further examples of genetic influences explaining discontinuity or marked change in development include such disorders such as Huntington's disease and life-course changes such as pubertal development. This is an important lesson for molecular genetic studies (see below) that seek to identify the presence of a single or a set of genes associated with the presence of a disorder. The developmental genetic perspective suggests that individuals may have the genetic risk but may not exhibit signs of the disorder until late in development.

The analytic models used to assess genetic influences underlying stability and change in psychopathology are similar to those used to examine genetic correlations underlying comorbid disorders (Plomin, 1986). However, although the analytic models are straightforward, there are few examples of longitudinal behavioral genetics analyses of psychopathology (McGue, Bacon, & Lykken, 1993).

Two studies found that both genetic and environmental factors contribute to stability in adolescent adjustment. For example, Van den Oord and Rowe (1997) reported that genetic and shared environmental influences contribute to the stability of antisocial, anxious/depressed, and hyperactive symptoms. O'Connor, Neiderhiser et al. (1998) found that genetic influences explained approximately half of the stability in antisocial and depressive symptoms, shared environmental factors explained stability of antisocial but not depressive symptoms, and nonshared influences also contributed to stability for both syndromes.

Genetic influences have also been reported in research on adult psychopathology. Data from the Virginia sample studied by Kendler and colleagues suggests that, for example, stability of depressive symptoms is mediated genetically (Kendler, Neale, et al., 1993). Data from a twin registry composed of individuals who served in the military during the Vietnam War indicated that phenotypic stability between juvenile and adult forms of antisocial behavior was due in roughly equal measure to genetic, shared environment, and nonshared environment (Lyons et al., 1995). Although the data are derived from retrospective reports, this is one of the few studies that has examined the well-known phenotypic connection between conduct problems in adolescence and antisocial behavior in adulthood from a genetic perspective.

Family history data are also consistent with the hypothesis that genetic influences underlie stability in psychopathology. As noted, genetic mediation of stability is suggested by parent–offspring concordance rather than age-to-age genetic correlations. In one sense the family history design is a strict test of genetic stability because it requires that the genes associated with the disorder in childhood or adolescence are the same as those associated with its manifestation in adulthood, and that the phenotypic expression of the disorder is similar. Examples of analyses of this type for depression are given by Harrington et al. (1993).

1.04.4.6 Summary

In this section we examined five distinct ways in which behavioral genetics research has informed and challenged our notion of the clinical phenotype. In particular, we provided evidence that genetic etiology for psychopathology is apparent, and in some cases pervasive, throughout the life-span. Second, consistent with the phenotypic research findings, behavioral genetics studies support a continuous model of psychopathology. There is remarkably little evidence that individuals exhibiting clinical levels of disorder are qualitatively distinct from individuals expressing milder forms, in terms of genetic and environmental risks. Third, behavioral genetics findings have helped to explain the pervasive cooccurrence of disorders in childhood through adulthood. The nosological implications of these findings require further attention. Fourth, the concepts of heterogeneity and disorder spectrum have also been shaped by behavioral genetics findings. Finally, although the evidence is limited, there appears to be growing support for the impact of genetic influences on the stability of psychopathology. These findings raise important questions regarding the turning "on and off" of genes and gene mechanisms in development.

In the next section we expand these lessons by examining what behavioral genetics research suggests about the processes associated with the development of psychopathology.

1.04.5 MODELS OF GENETIC AND ENVIRONMENTAL RISK MECHANISMS

The range of behaviors apparently under genetic control is remarkable. It is now well accepted that intelligence, personality and psychopathology are influenced by genetic factors, but recent studies also suggest genetic influence on divorce (McGue & Lykken, 1992), life events (Plomin, Lichtenstein, Pedersen, McClearn, & Nesselroade, 1990), sexual orientation in males (Bailey & Pillard, 1991) and females (Bailey, Pillard, Neale, & Agyei, 1993), social support (Bergeman, Plomin, Pederson, McClearn, & Nesselroade, 1990; Kendler et al., 1995), peer affiliation (Manke, McGuire, Reiss, Hetherington, & Plomin, 1995), personal devotion to religion (Kendler, Gardner & Prescott, 1997), smoking (Kendler, Neale, MacLean, et al., 1993), television viewing (Plomin et al., 1990), behavior in problem-solving interactions (O'Connor, Hetherington, Reiss, & Plomin, 1995), work behavior and a remarkably wide range of attitudes and interests (Plomin et al., 1997), among others. Indeed, some authors have commented that the null hypothesis of no genetic influence is no longer justifiable (see commentaries on Plomin & Bergeman, 1991).

These findings lead to the inevitable conclusion, already stated in the chapter, that the main thrust of research in behavioral genetics is no longer whether a behavior or trait is genetically influenced. Instead, the focus of behavioral genetics research in psychopathology is how to conceptualize the mechanisms through which genetic factors impart their influence, and how genetic and psychosocial risks jointly operate to lead to maladjustment.

1.04.5.1 Conceptualizing Genetic Influence

Many of the models cited for understanding how genes and environment operate together have little relevance to psychopathology. An example of this is PKU, a disorder in which affected individuals are unable to synthesize phenylalanine. This condition, which leads to mental impairment, can be effectively treated if dietary restrictions are followed. Although PKU is a good example of how genetic causes can have environmental cures, it is an inadequate model for understanding genes and environment as applied to psychopathology. In this example, neither the genetic model (single gene) nor the environmental model (environmental risk exposure leads directly and unfailingly to the manifestation of the disorder) parallel what is known about psychopathology.

1.04.5.2 Conceptualizing the Environment

Much of the controversy regarding family influence from behavioral genetics research concerns the finding that shared environmental parameters are usually modest or even zero. Does this finding suggest that there is no effect of the family environment? Absolutely not. Instead, what this suggests is that simply growing up in the same home appears to matter little in shaping children's similarity in behavioral traits. It is worth restating here that estimates of shared familial environment are based on *effects* rather than *measures*. Accordingly, shared environmental influences arise only to the extent to which environments have functionally equivalent effects on siblings.

As noted previously, one of the most important lessons for environmental research has been the importance of nonshared environment. In order to identify environmental influences that play salient roles in development and psychopathology it is critical to assess risks specific to the individual because risks shared by family members (e.g., socioeconomic disadvantage, maternal depression) may nonetheless have nonshared effects.

1.04.5.3 Genotype–Environment (G–E) Correlations

A major advance in integrating genetics and the environment in research has been the discovery of genotype–environment correlations. Some of the lessons learned from these efforts are described below.

1.04.5.3.1 Passive genotype–environment correlation

Genotype–environment correlations come in a variety of forms (Plomin et al., 1977; Scarr & McCartney, 1983). The most obvious example is passive genotype–environment correlation. Passive correlations arise because related parents and children share both genes and environments. That is, children share half of their parents' genes and are exposed to environments provided by them. The correlation is "passive" in that children themselves play no direct role in creating the overlap of genetic and environmental factors. The only way of directly estimating the influence of passive genotype–environment correlations is to compare parents and their offspring in biological and adoptive families.

Passive genotype–environment correlations make a number of studies of children of psychiatric parents difficult to interpret because these children are not only more likely to inherit genetic risks but also to be exposed to dysfunctional environments that psychiatric parents are disproportionately likely to provide (Rutter & Quinton, 1984). Studying the offspring of parents diagnosed with depression, alcoholism, or other disorders was a mainstay of high-risk research, but the presence of passive genotype–environment correlations makes interpretations of the mechanisms leading to behavior problems difficult to determine.

1.04.5.3.2 Evocative and active genotype–environment correlations

The principle that individuals are active participants in and create and select their own environments is widely acknowledged. Individual differences in relationship styles, hyperactivity, conduct disorder, and sociability are sustained through the dynamic interchange between personal characteristics and social experiences (Caspi & Moffitt, 1995; Rutter, Champion, Quinton, Maughan, & Pickles, 1995). There are many illustrations of this principle from phenotypic research, including the finding that aggressive children evoke negative responses from strangers (Anderson, Lytton, & Romney, 1986) and aggressive individuals are more likely to select aggressive partners (Krueger et al., in press; see Caspi & Moffitt, 1995).

Even more striking is the finding that the way in which individuals evoke, select, and react to their environments is not independent of genetic influence. In this way, there is a correlation between the genetic predispositions of individuals and the environments they experience (Plomin, 1994). Comparisons of identical and nonidentical twins (and biological and adopted siblings) consistently demonstrate that the former are more similar than the latter in, for example, their selection of peer groups (Manke et al., 1995) and manner of interacting with others, notably parents (O'Connor et al., 1995). These findings illustrate how there may be genetic influences on environmental measures (Plomin & Bergeman, 1991).

Two recent adoption studies provide parallel examples of genotype–environment evocative correlation as regards antisocial behavior. In a high-risk adoption study, Ge et al. (1996) reported that adopted adolescents at genetic risk for antisocial behavior (by virtue of having a biological parent with diagnosed disorder) were more likely to evoke aggressive, hostile parenting and less likely to evoke supportive parenting from their adoptive mothers and fathers compared with adoptees not at genetic risk. The study also found that the more negative parenting of at-risk adoptees was mediated by their own disruptive, aggressive behavior toward the parent.

A similar study based on the Colorado Adoption Project (CAP) also found that individuals with a biological parent with a history of antisocial behavior were more likely to evoke negative, coercive parenting from their adoptive parent relative to adoptees not at genetic risk for antisocial behavior (O'Connor, Deater-Deckard, Fulker, Rutter, & Plomin, in press). The CAP study found that evocative genotype–environment correlations were modest in magnitude from late childhood to early adolescence. As in the Ge et al. (1996) study, O'Connor, Deater-Deckard et al. (in press) found that the effect was mediated by children's aggressive behavior, but the mediation effect was far less dramatic in the latter study. Both studies provide compelling evidence that genetic influences are not limited to trait-like behaviors, but are equally relevant to socialization processes and family relations.

1.04.5.3.3 Hypothesized age changes in genotype–environment correlations

Scarr and McCartney (1983) advanced the provocative hypothesis that genotype–environment correlations increase with age and change in type. Specifically, they suggested that passive genotype–environment correlations would predominate in infancy and early childhood and that, with increasing age and ability to select one's own environments, there would be a gradual shift and evocative/active genotype–environment correlations would predominate.

There is a natural developmental appeal in this hypothesis, but few data are available to test the idea. What data are available are not consistent with the proposal. For example, rather than wane from infancy to adolescence, a recent adoption study highlighted that passive genotype–environment correlations remain strong in adolescence (McGue, Sharma, & Benson, 1996). Additionally, Lytton (1977), among others, has found that evocative genotype–environment correlations are robust in early childhood. Moreover, one of the few studies of evocative genotype–environment correlations found no sign of developmental change from late childhood through early adolescence (O'Connor, Deater-Deckard et al., in press).

Of course, the hypothesis that there are developmental changes in the type and magnitude of genotype–environment correlations is different from the more straightforward hypothesis that there are age-based changes in the effects of genetic influences and, alternatively, shared and environmental influences. As noted, there is now considerable evidence that genetic influences seem to increase with age and that shared environmental influences decrease steadily from early childhood through adulthood (McCartney, Harris & Bernier, 1990).

One implication of genotype–environment correlations is that the psychosocial processes associated with the development of psychopathology may not be entirely environmentally mediated. Consider the process of coercive parenting and child disruptive behavior outlined by Patterson (1982). It is possible to view both sides of this transactional process as genetically influenced. Children's aggressive behavior is partly genetically influenced; aggressive antisocial parenting may itself be associated with genes affecting the parent's antisocial tendencies or may be evoked by genetically influenced antisocial behavior in the children.

Consistent with the processes suggested above, a handful of studies demonstrate that relations between parenting and child adjustment—the hallmark of developmental research on psychopathology—are mediated by genetic influences. For example, in a twin–sibling study, Pike et al. (1996) attributed the association between mothers' and fathers' negative parenting and adolescent symptoms of depression and antisocial behavior largely to genetic factors. That is, the correlation between, for example, coercive, hostile parental behavior and children's externalizing symptoms arises because genetically influenced aggressive and externalizing behaviors evoke negative parenting.

1.04.5.4 Genotype–Environment Interactions

It is possible to identify several types of genotype–environment interactions. The most common way in which genotype–environment interactions are conceptualized is that there are genetic influences on the sensitivity to environmental risks (Kendler, 1995). Stated differently, individuals at genetic risk are more likely than individuals not at genetic risk to react maladaptively to psychosocial stresses.

This hypothesis is a simple extension of a basic tenet in developmental psychopathology, namely, that there are individual differences in vulnerability to psychopathology. The extension is that a source of these individual differences in vulnerability is genetics.

Despite the general acknowledgement that there are likely to be genotype–environment interactions in development, there are remarkably few examples. Wahlsten (1990; see also commentaries) suggests that the reason may be partly methodological. He rightly points out that many human studies lack the power to detect interaction effects after the main effects are removed. This is particularly relevant to studies that compare correlations, such as between MZ and DZ twins, and particularly if comparisons are made between patterns under high and low stress. Wahlsten's assumption is that genotype–environment interactions are common but go undetected. One reason for believing this argument is that genotype–environment interactions are numerous in animal research in which it is relatively easy to control and constrain both genetic factors (through breeding or using identified strains) and environmental conditions (through experimental manipulation that would not be ethical in human research).

There are several additional reasons why detecting genotype–environment interactions has been difficult. First, there are few examples of behavioral genetics analyses across a wide range of stress conditions. Most authors assume that genotype–environment interactions would be most easily detected in the context of extreme adversity because genetic effects are thought to be stronger in high-risk environments than in normal and low-risk environments—although one could make the opposite argument that high-risk environments might overwhelm individual differences in vulnerability (e.g. Scarr, 1992).

These challenges notwithstanding, there are a number of illustrations of genotype–environment interactions for a range of psychopathological conditions. The Swedish adoption study reported by Bohman (1996) provides one of the best examples. Bohman (1996) reported that children with an antisocial biological parent who were also raised in their biological family (i.e., high genetic risk and environmental risk) were considerably more likely to be antisocial than children of antisocial biological parents who were raised by adoptive parents (i.e., high genetic risk but low-risk environment), children exposed to environmental risk in the absence of genetic risk, and children from the normal population. The implication of the pattern of findings is that children at biological risk are relatively more vulnerable to environmental adversity than children not at genetic risk.

In a separate adoption study of antisocial behavior, Cadoret, Yates, Troughton, Woodworth, & Stewart (1995) reported an association between adverse adoptive home environment and adoptee aggressive, antisocial outcome only in the presence of genetic risk (see also Cloninger, Sigvardsson, Bohman, & von Knorring, 1982). In their study of depression in adult female twins, Kendler et al. (1995) found that the difference in the likelihood of experiencing depressive symptoms between those at high and low genetic risk was substantially magnified following severe life events. Cadoret et al. (1996) also reported genotype–environment interaction for depression in females (but not males), but in that study genetic diathesis was defined not by depression but by alcoholism. Nonetheless, the interaction is the same: environmental adversity is associated with increased likelihood of depression only in the presence of genetic risk.

Finally, Wahlberg et al. (1997) reported a genotype–environment interaction for schizophrenia-related symptoms. Although neither genetic risk (having a biological parent with schizophrenia) nor environmental risk (defined as communicative deviance in the adoptive family) alone resulted in thought disorder among adoptees, the joint effect of both variables was associated with a substantial increase in thought disorder (see also Tienari et al., 1994; see True et al., 1994, for an example as applied to post-traumatic stress disorder).

A slightly different model of genotype–environment interaction was reported by Slutske et al. (1997). These authors found that genetic influence on conduct disorder based on retrospective reports in adulthood existed only if the twins had the same friends; this was true only for boys. In this instance, it is not genetic control over sensitivity to environmental stress, but rather similarity in environmental circumstances that accentuated genetic influences.

1.04.5.5 Summary

Several developmental models need to be considered in elucidating genetic influences in the development of psychopathology (e.g., Bronfenbrenner & Ceci, 1993; Gottesman & Goldsmith, 1994; Kendler & Eaves, 1986; Plomin, 1994; Reiss, 1995; Rutter et al., 1997). Among the possibilities are: (i) genetic and environmental effects operating in an additive manner; (ii) genetically influenced characteristics altering the exposure to environmental risks, or genotype–environment correlations; (iii) genetic influences mediating the association between putative risk and adjustment; and (iv) genetic influences controlling the sensitivity to environmental risks, or genotype–environment interactions. It is important to note that additive, correlated, mediational, and interactive genotype–environment models are not mutually exclusive.

Further improvements and refinements of our understanding of these mechanisms will depend on future research in which genetic risks, like environmental risks, are specified. This, of course, requires the contributions of molecular genetics research.

1.04.6 MOLECULAR GENETIC RESEARCH ON PERSONALITY AND PSYCHOPATHOLOGY

Behavioral genetics refers to the genetic analysis of behavior. Quantitative genetic analyses such as twin and adoption studies are best known, but molecular genetic analyses that attempt to identify specific genes and their functions are becoming increasingly important. Although complex traits like behavior are likely to be influenced by multiple genes, identifying some of these genes will provide the opportunity for more focused analyses of questions raised by quantitative genetics using direct assessment of measured genotypes (Plomin & Rutter, in press).

1.04.6.1 Methods and Questions in Molecular Genetics Research on Psychopathology

A detailed discussion of the relative merits of association and linkage studies and the statistical methods devised to detect quantitative trait loci, among other statistical procedures, is beyond the scope of this chapter (see Baron; 1997; Plomin et al., 1997). Suffice it to say here that just as nonbehavioral genetics investigators are increasingly familiar with the quantitative models and assumptions in that field, the methods and assumptions of molecular genetics will also become increasingly familiar.

Although we avoid a detailed methodological overview, several points deserve attention. First, it is important to recall that molecular genetics studies of psychopathology should not be equated with simplistic reductionistic models. There are a number of reasons why this is the case. The most obvious is that most authors would agree that specific genes will account for a very small percentage of the total variance in psychopathology. Moreover, simply identifying the genes provides no answers to the processes through which genetic and environmental influences lead to psychopathology. Accordingly, results from molecular genetics studies will likely heighten the need for intensive research on environmental processes. There are, to date, few attempts to assess both genetic and psychosocial sources of risk in these studies, and this remains an important direction for future research.

It is important to distinguish research that attempts to identify genetic loci associated with disorder from research that uses previously identified loci (Plomin & Rutter, in press). Although it is difficult and expensive to identify genes associated with behavior, it is relatively easy and inexpensive to use previously identified genes as measured genetic risk factors in any type of research on description, prediction, prevention, and intervention.

1.04.6.2 Contributions of Molecular Genetics Research to Models of Psychopathology

Despite a decade of attempts to identify genes for schizophrenia and bipolar disorder (e.g., Peltonen, 1995), success in identifying genes seems to be coming more quickly for personality dimensions (Cloninger, 1997; Ebstein, Nemanov, Klotz, Gritsenko, & Belmaker, 1997) and more common disorders such as reading disability (Cardon et al., 1994; Grigorenko et al., 1996; Schulte-Korne et al., 1997), attention deficit hyperactivity disorder (ADHD) (LaHoste et al., 1996; Palmer et al., 1997; Sunohara et al., 1997) and opioid dependence (Ebstein & Belmaker, 1997). For schizophrenia, there have been several reports suggesting genes on chromosomes 3, 5, 6, 8, and 22 (see Schwab et al., 1997).

A further example of how molecular genetics has contributed to clinical psychology and psychiatry is Alzheimer's disease. A familial component of Alzheimer's disease, at least in a minority of cases, had been suspected for several years, and genetic influence was implicated in twin studies (Breitner et al., 1993). Molecular genetic studies indicated that late-onset Alzheimer's disease (onset after age 65) that shows a familial pattern is linked to a gene on chromosome 19 (apolipoprotein E4 or Apo-E4) (Corder et al., 1993). Apo-E4 operates as a susceptibility gene, quadrupling risk for Alzheimer's disease, but it is neither necessary nor sufficient for the development of the disorder. In addition, a much rarer form of Alzheimer's disease, early onset, can be caused by a gene on chromosome 14 (Sherrington et al., 1995); other even rarer single genes for early-onset Alzheimer's disease have been found.

1.04.6.3 Summary and Future Directions

The end goal of the "gene chase" in behavioral science (McClearn et al., 1991) is not simply to detect an association between a specific locus (or loci) and behavior. Neither a significant association between an allele and disorder nor a significant heritability estimate indicate how genetic factors operate in development. Once genes are identified, we need to explore the biological and social pathways through which genetic vulnerabilities, expressed through specified biochemical markers, respond to social and psychological events to produce psychopathology. The identification of specific genetic vulnerability markers may lead to advances such as the identification of which individuals may be more sensitive to certain psychosocial stresses and a better understanding of how genetic vulnerabilities covary with psychosocial adversity, as well as progress in defining genetic heterogeneity and genetic correlation between disorders.

1.04.7 CLINICAL PSYCHOLOGY IMPLICATIONS OF BEHAVIORAL GENETICS RESEARCH FINDINGS

There have already been major clinical implications for medical treatment, but what about clinical psychology? In many cases the clinical implications of behavioral genetics findings for psychology are not yet well understood. This is partly because the field is still relatively young as regards psychopathology and because the mechanisms of genetic influences have not yet been clearly delineated.

A further reason why lessons from behavioral genetics studies have only slowly begun to shape clinical practice in psychology is that there continue to be misplaced biases against behavioral genetics findings. This appears to be so despite excellent papers dispelling various myths (Rutter & Plomin, 1997). As the section on genotype–environment models emphasized, there is no "zero-sum" relationship between the role of genetics and the role of environmental influences (which include clinical interventions). Indeed, it is precisely because genetic influences do not operate in isolation from psychosocial stresses that the finding of genetic influences should encourage the development of effective prevention and intervention models.

1.04.7.1 Clarifying Genetic Risks in Assessment

It is already common practice to collect information on family psychiatric history. Whether there is a history of depression or substance use in the family may, most would agree, contribute to understanding the difficulties experienced by the patient and the likelihood of an individual developing a similar disorder. Whether such information would alter how a patient's diagnosis is conceptualized is a slightly more complicated matter. Although behavioral genetics research findings have made important contributions to diagnostic debates, the translation of these findings into clinical practice as applied to an individual is not direct and often fraught with difficulty.

Consequently, it may be that knowledge of genetic risks (e.g., based on family history) contributes most directly not to the conceptualizing diagnosis, but rather to conceptualizing the sources of risk. Research on ADHD provides a good example. ADHD can be managed by psychostimulant medication and parenting practices (Barkley, 1990). The finding that first-degree relatives, notably parents, of ADHD probands are also at elevated risk of ADHD (Biederman et al., 1992) suggests that it would make good clinical sense to assess symptoms of poor concentration and inattention in the parents, even if they are not the identified patient. Assessing ADHD symptoms in the parents may lead to relief in the parent(s) and facilitate parent-training approaches and other forms of parent–child and family treatment.

1.04.7.2 Genetic and Environmental Considerations for Treatment and Prevention

Perhaps the most important lesson for treatment is that there is no necessary connection between causes and cures—specifically, between genetic influence and type of intervention. That is, psychopathology that is strongly influenced by genetic factors may nevertheless be affected by intervention. Thus, finding genetic effects in no way contradicts the benefits of, e.g., behavioral interventions to improve the parenting of antisocial children (Patterson, 1982; Webster-Stratton, Hollinsworth, Kolpacoff, 1989). In this context it is also interesting to note the high rate of schizophrenia in people from the Caribbean living in the UK but not in those living in the Caribbean area (Van Os, Castle, Takei, Der, & Murray, 1996). In addition, it is worth remembering that the factors influencing individual differences in psychopathology are not necessarily the same as those underlying differences in the rate of psychopathology between groups or over time. Consider, for example, the marked secular changes in crime and suicide in young people over the last 50 years and the reported increase in rates of depression and other mental disorders (Rutter & Smith, 1995).

Although somewhat speculative, it is possible to derive treatment and prevention implications from research on genotype–environment correlations and interactions. To take one example, antisocial children often have a "double" risk: inheriting genetic risk from parents and being exposed to coercive interactions and poor monitoring (that may be a result of the parents' antisocial behavior). Consequently, it is possible to conceptualize parent–child interventions as interfering with an otherwise "naturally occurring" passive and evocative genotype–environment correlation. A complementary explanation is also possible based on genotype–environment interactions, namely, intervention and prevention efforts reduce genetic risk by decreasing environmental risks to which these individuals might be particularly susceptible. The number of studies designed to prevent the development of disorder in children of depressed parents is a second example (e.g., Beardslee, Wright, Salt, & Drezner, 1997).

More clinical research is needed that specifically identifies genetic risk status and elucidates treatment processes. In fact, clinical research is one of the best lines of investigation to document genotype–environment interactions and correlations, basic notions of how genes and environment jointly contribute to the development of psychopathology. Moreover, intervention research that explicitly incorporates genetic risk status and manipulates psychosocial experiences may help to identify subgroups of individuals who respond best to particular treatment approaches. This could be accomplished, for instance, by including genetic risk status (based on presence of disorder in first-degree relatives or perhaps even genetic markers) as a predictor of treatment response.

1.04.7.3 Summary

Necessarily, the clinical implications of available behavioral genetics findings are, at this time, largely speculative. Although it is possible to derive some general lessons regarding treatment and prevention from nonclinical studies, there is an unfortunate absence of research that directly fuses psychological intervention and genetic risk models. What available findings do suggest is that the mechanisms of intervention may be far more complex than initially proposed. Psychological interventions may be effective partly because they alter genetically-based susceptibility to environmental risk or eliminate otherwise pervasive genotype–environment correlations that increase the likelihood of psychopathology. The second point which we wish to emphasize here is that intervention and prevention studies are powerful designs in which to test some of the most critical but yet unanswered questions concerning the mechanisms leading to psychopathology.

1.04.8 GENERAL SUMMARY AND CONCLUSION

The remarkable progress in our understanding of genetic influences on a wide range of medical conditions and behavior has been noted by numerous commentators in the scientific and popular press. Fortunately, previous conceptions of genetic influence as immutable and contrary to psychological interventions and socialization theories are now dispelled. Moreover, contemporary concerns regarding ethical and moral issues for genetic research are based, rather, on informed dialogue (e.g., see Gottesman & Bertelson, 1996). No longer a topic of largely political weight, the so-called nature–nurture debate has spawned a major research paradigm that has yielded some of the most important clues to our understanding of the development, etiology, and conceptualization of psychopathology.

We reiterate three general conceptual advances derived from behavioral genetics research. Perhaps the most important is that the conceptualization of genetic and environmental factors has advanced beyond the question of whether genetic influence is relevant (cf. Rutter et al., 1997). Questions that need to be addressed focus instead on how genetic vulnerabilities lead to maladjustment and how to incorporate genetic vulnerability into the already well-established models of risk and resilience. A second, fundamental conclusion is that although authors differ in their optimism regarding how genetic studies can shed light on environmental mechanisms of risk, all would agree that studying psychosocial risks must cooccur with analyses of biological and genetic factors (Reiss, Plomin, & Hetherington, 1993). This is as true for molecular genetics as it is for twin and adoption designs. Finally, one of the most important directions of behavioral genetics research has been to encourage molecular geneticists to search for specific genes that may help answer questions about how genetic influences create vulnerability to psychopathologies. Clearly, progress in understanding genetic and environmental sources of risk for psychopathology thus appears to be following a path of dialectic integration rather than pendulum-like swings that characterized research just a decade ago.

1.04.9 REFERENCES

Anastasi, A. (1958). Heredity, environment, and the question "How?". *Psychological Review, 75,* 81–95.

Anderson, K. E., Lytton, H., & Romney, D. M. (1986). Mothers' interactions with normal and conduct disordered boys: Who affects whom? *Developmental Psychology, 22,* 604–609.

Bailey, A., Phillips, W., & Rutter, M. (1996). Autism: Towards an integration of clinical, genetic, neuropsychological and neurobiological perspectives. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 37,* 89–126.

Bailey, J. M., & Pillard, R. C. (1991). A genetic study of male sexual orientation. *Archives of General Psychiatry, 48,* 1089–1096.

Bailey, J. M., Pillard, R. C., Neale, M. C., Agyei, Y. (1993). Heritable factors influence sexual orientation in women. *Archives of General Psychiatry, 50,* 217–223.

Baron, M. (1997). Genetic linkage and bipolar affective disorder: Progress and pitfalls. *Molecular Psychiatry, 2,* 200–210.

Barkley, R. A. (1990). *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment.* New York: Guilford Press.

Beardslee, W. R., Wright, E. J., Salt, P., & Drezner, K. (1997). Examination of children's responses to two preventive intervention strategies over time. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 196–204.

Bergeman, C. S., Plomin, R., Pederson, N. L., McClearn, G. E., & Nesselroade, J. R. (1990). Genetic and environmental influences on social support: The Swedish Adoption/Twin Study of Aging (SATSA). *Journal of Gerontology, 45,* 101–106.

Biederman, J., Faraone, S. V., Keenan, K., Benjamin, J., Krifcher, B., Moore, C., Buckminster, S., Ugaglia, K., Jellinek, M., Steingard, R., Spencer, T., Norman, D., Kolodny, R., Kraus, I., Perrin, J., Keller, M., & Tsuang, M. (1992). Further evidence for family-genetic risk factors in attention deficit hyperactivity disorder. *Archives of General Psychiatry, 49,* 728–738.

Bohman, M. (1996). Predisposition to criminality: Swedish adoption studies in retrospect. In G. R. Bock & J. A. Goode (Eds.), *Genetics of criminal and antisocial behavior* (pp. 99–114). Chichester, UK: Wiley.

Bolton, P., MacDonald, H., Pickles, A., Rios, P., Goode, S., Crowson, M., Bailey, A., & Rutter, M. (1994). A case-control family history study of autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 35,* 877–900.

Bouchard, T., & McGue, M. (1981). Familial studies of intelligence: A review. *Science, 212,* 1055–1059.

Bouchard, T., & Propping, P. (1993). *Twins as a tool of behavioral genetics.* Chichester, UK: Wiley.

Breitner, J., Gatz, M., Bergern, A., Christian, J., Mortimer, J., McClearn, G. E., Heston, L. L., Welsh, K., Anthony, J., & Folstein, M. F. (1993). Use of twin cohorts for research in Alzheimer's disease. *Neurology, 43,* 261–267.

Bronfenbrenner, U., & Ceci, S. J. (1993). Heredity, environment, and the question "how?"—a first approximation. In R. Plomin & G. E. McClearn (Eds.), *Nature, nurture, and psychology* (pp. 313–324). Washington, DC: APA Books.

Buss, D. M. (1991). Evolutionary personality psychology. *Annual Review of Psychology, 42,* 459–491.

Cadoret, R. J., Winokur, G., Langbehn, D., Troughton, E., Yates, W. & Stewart, M. A. (1996). Depression spectrum disease, I: The role of gene–environment interaction. *American Journal of Psychiatry, 153,* 892–899.

Cadoret, R. J., Yates, W. R., Troughton, E., Woodworth, G., & Stewart, M. A. (1995). Genetic–environmental interactions in the genesis of aggressivity and conduct disorders. *Archives of General Psychiatry, 52,* 916–924.

Cardon, L., Smith, S. D., Fulker, D. W., Kimberling, W. J., Pennington, B. F., & DeFries, J. C. (1994). Quantitative trait locus for reading disability on chromosome 6. *Science, 266,* 276–279.

Carey, G. (1992). Twin imitation for antisocial behavior: Implications for genetic and family environment research. *Journal of Abnormal Psychology, 101,* 18–25.

Caron, C. & Rutter, M. (1991). Comorbidity in child psychopathology: Concepts, issues and research strategies. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 32,* 1063–1080.

Caspi, A., & Moffitt, T. E. (1995). The continuity of maladaptive behaviour: From description to understanding in the study of antisocial behaviour. In D. Cicchetti & D. Cohen (Eds.), *Developmental psychopathology* (Vol. 2, pp. 472–511). New York: Wiley.

Cloninger, C. R. (1997). Quantitative trait loci for personality in families of alcoholics [abstract]. *American Journal of Medical Genetics (Neuropsychiatric Genetics), 74,* 579.

Cloninger, C. R., Bohman, M., Sigvardsson, S. (1981). Inheritance of alcohol abuse: Cross-fostering analysis of adopted men. *Archives of General Psychiatry, 38,* 861–868.

Cloninger, C. R., Sigvardsson, M., Bohman, M., & von Knorring, A. L. (1982). Predisposition to petty criminality in Swedish adoptees, II: Cross fostering analysis of gene–environment interaction. *Archives of General Psychiatry, 39,* 1242–1247.

Corder, E. H., Saunders, A. M., Strittmatter, W., Shmechel, D. E., Gaskell, P., Small, G., Roses, A. D., Haines, J. L., & Pericak-Vance, M. A. (1993). Gene dose of apolipoprotein and type 4 allele and the risk of Alzheimer's disease in late-onset families. *Science, 261,* 921–923.

Coryell, W. (1997). Do psychotic, minor and intermittent depressive disorders exist on a continuum? *Journal of Affective Disorders, 45,* 75–83.

Crabbe, J. C., Belknap, J. K., & Buck, K. J. (1994). Genetic animal models of alcohol and drug use. *Science, 264,* 1715–1723.

Deater-Deckard, K., Reiss, D., Hetherington, E. M., & Plomin, R. (1996). Dimensions and disorders of adolescent adjustment: A quantitative genetic analysis of unselected samples and selected extremes. *Journal of Child Psychology and Psychiatry, 38,* 515–525.

DeFries, J. C., & Fulker, D. W. (1985). Multiple regression analysis of twin data. *Behavior Genetics, 15,* 467–473.

Detterman, D. (1996). *Current topics in human intelligence, Vol. 5: The environment.* Norwood, NJ: Ablex.

DiLalla, L. F., & Gottesman, I. I. (1989). Heterogeneity of causes for delinquency and criminality: Lifespan perspectives. *Development and Psychopathology, 1,* 339–349.

Downey, G., & Coyne, J. C. (1990). Children of depressed parents: An integrative review. *Psychological Bulletin, 108,* 50–76.

du Fort, G., Kovess, V., & Boivin, J. F. (1994). Spouse similarity for psychological distress and well-being: A population study. *Psychological Medicine, 24,* 431–447.

Dunn, J., Stocker, C., & Plomin, R. (1990). Nonshared experiences within the family: Correlates of behavioral problems in middle childhood. *Development and Psychopathology, 2,* 113–126.

Eaves, L. J., Silberg, J. L., Hewitt, J. K., Rutter, M., Meyer, J. M., Neale, M. C., & Pickles, A. (1993). Analyzing twin resemblance in multi system data: Genetic applications of a latent class model for symptoms of conduct disorder in juvenile boys. *Behavior Genetics, 23,* 5–19.

Eaves, L. J., Silberg, J., Meyer, J., Maes, H., Simonoff, E., Pickles, A., Rutter, M., Neale, M., Reynolds, C., Erikson, M., Heath, A., Loeber, R., Truett, K., &

Hewitt, J. (1997). Genetics and developmental psychopathology, 2: The main effects of genes and environment on behavioral problems in the Virginia Twin Study of Adolescent Behavioral Development. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 38,* 965–980.

Ebstein, R. P., & Belmaker, R. H. (1997). Saga of an adventure: Novelty seeking, substance abuse and the dopamine D4 receptor (DRD4) exon III repeat polymorphism. *Molecular Psychiatry, 2,* 381–384.

Ebstein, R. P., Nemanov, L., Klotz, I., Gritsenko, I., & Belmaker, R. H. (1997). Additional evidence for an association between the dopamine D4 receptor (DRD4) exon III repeat polymorphism and the human personality trait of novelty seeking. *Molecular Psychiatry, 2,* 472–477.

Edelbrock, C., Rende, R., Plomin, R., & Thompson, L. A. (1995). A twin study of competence and problems behaviours in childhood and early adolescence. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 36,* 775–785.

Eley, T., Lichenstein, P., & Stevenson, J. (in press). Sex differences in the etiology of aggressive and nonagressive antisocial behavior: Results from twin studies. *Child Development.*

Falconer, D. S. (1960). *Introduction to quantitative genetics.* New York: Ronald Press.

Feng, D., & Baker, L. (1994). Spouse similarity in attitudes, personality, and psychological well-being. *Behavior Genetics, 24,* 357–364.

Flint, J., Corley, R., DeFries, J. C., Fulker, D. W., Gray, J. A., Miller, S., & Collins, A. C. (1995). A simple genetic basis for complex psychological trait in laboratory mice. *Science, 269,* 1432–1435.

Folstein, S., & Rutter, M. (1977). Genetic influences and infantile autism. *Nature, 265,* 726–728.

Fulker, D. W., & Cherny, S. S. (1996). An improved multipoint sib-pair analysis of quantitative traits. *Behavior Genetic, 26,* 527–532.

Ge, X., Conger, R. D., Cadoret, R. J., Neiderhiser, J. M., Yates, W., Troughton, E., & Stewart, M. A. (1996). The developmental interface between nature and nurture: A mutual influence model of child antisocial behavior and parent behaviors. *Developmental Psychology, 32,* 574–589.

Gershon, E. S., & Cloninger, C. R. (1994). *Genetic approaches to mental disorders.* Washington, DC: American Psychiatric Press.

Gjone, H., Stevenson, J., & Sundet, J. M. (1996). Genetic influence on parent-reported attention-related problems in a Norwegian general population twin sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 588–596.

Gjone, H., Stevenson, J., Sundet, J. M., & Eilertsen, D. E. (1996). Changes in heritability across increasing levels of behavior problems in young twins. *Behavior Genetics, 26,* 419–426.

Gottesman, I. I. (1974). Developmental genetics and ontogenetic psychology: Overdue detente and propositions from a matchmaker. In A. D. Pick (Ed.), *Minnesota symposium on child psychology* (pp. 55–80). Minneapolis, MN: University of Minnesota Press.

Gottesman, I. I. (1991). *Schizophrenia genesis: The origins of madness.* New York: Freeman.

Gottesman, I. I., & Bertelson, A. (1996). Legacy of German psychiatric genetics: Hindsight is always 20/20. *American Journal of Medical Genetics, 67,* 317–322.

Gottesman, I. I., & Goldsmith, H. H. (1994). Developmental psychopathology of antisocial behavior: Inserting genes into its ontogenesis and epigenesis. In C. A. Nelson (Ed.), *Threats to optimal development: Minnesota symposium on child psychology* (Vol. 27, pp. 20–44). Hillsdale, NJ: Erlbaum.

Grigorenko, E. L., Wood, F. B., Meyer, M. S., Hart, L. A., Speed, W. C., Shuster, A., & Pauls, D. L. (1996). Susceptibility loci for distinct components of dyslexia on chromosome 6 and 15. *American Journal of Human Genetics, 60,* 27–39.

Harrington, R. C., Fudge, H., Rutter, M. L., Bredenkamp, D., Groothues, C., & Pridham, J. (1993). Childhood and adult depression: A test of continuities with data from a family study. *British Journal of Psychiatry, 162,* 627–633.

Heston, L. L. (1966). Psychiatric disorders in foster home reared children of schizophrenic mothers. *British Journal of Psychiatry, 112,* 819–825.

Jinks, J. J., & Fulker, D. W. (1970). Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behavior. *Psychological Bulletin, 75,* 311–349.

Joreskog, K. G., & Sorbom, D. (1989). *LISREL VII: User's guide* (2nd ed.). Mooresville, IN: Scientific Software.

Karayiorgou, M., & Gogos, J. (1997). Dissecting the genetic complexity of schizophrenia. *Molecular Psychiatry, 2,* 211–223.

Kendler, K. S. (1995). Genetic epidemiology in psychiatry: Taking both genes and environment seriously. *Archives of General Psychiatry, 52,* 895–899.

Kendler, K. S. (1996). Parenting: A genetic–epidemiological perspective. *American Journal of Psychiatry, 153,* 11–20.

Kendler, K. S., & Eaves, L. J. (1986). Models for the joint effect of genotype and environment on liability to psychiatric illness. *American Journal of Psychiatry, 143,* 279–289.

Kendler, K. S., Gardner, C. O., & Prescott, C. A. (1997). Religion, psychopathology, and substance use and abuse: A multimeasure, genetic–epidemiologic study. *American Journal of Psychiatry, 154,* 322–329.

Kendler, K. S., Gruenberg, A. M., & Kinney, D. K. (1994). Independent diagnoses of adoptees and relatives as defined by *DSM-III* in the provincial and national samples of the Danish Adoption Study of schizophrenia. *Archives of General Psychiatry, 51,* 456–468.

Kendler, K. S., Heath, A. C., Neale, M. C., Kessler, R. C. & Eaves, L. J. (1993). Alcoholism and major depression in women: A twin study of the causes of comorbidity. *Archives of General Psychiatry, 50,* 690–698.

Kendler, K. S., Kessler, R. C., Walters, E. E., MacLean, C., Neale, M. C., Heath, A. C., & Eaves, L. J. (1995). Stressful life events, genetic liability, and onset of an episode of major depression in women. *American Journal of Psychiatry, 152,* 833–842.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992b). Childhood parental loss and adult psychopathology in women: A twin study perspective. *Archives of General Psychiatry, 49,* 109–116.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992d). Major depression and generalized anxiety disorder: Same genes, (partly) different environments? *Archives of General Psychiatry, 49,* 716–722.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992a). A population-based twin study of major depression in women: The impact of varying definitions of illness. *Archives of General Psychiatry, 49,* 257–266.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992c). Familial influences on the clinical characteristics of major depression: A twin study. *Acta Psychiatrica Scandinavica, 86,* 371–378.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1993). A longitudinal twin study of 1-year prevalence of major depression in women. *Archives of General Psychiatry, 50,* 843–852.

Kendler, K. S., Neale, M. C., MacLean, C. J., Heath, A. C., Eaves, L. J., & Kessler, R. C. (1993). Smoking and major depression: A causal analysis. *Archives of General Psychiatry, 50,* 36–43.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H.-U., Kendler, K. S. (1994). Lifetime and 12-month prevalence of *DSM-III-R* psychiatric disorders in the United States. *Archives of General Psychiatry, 51,* 8–19.

Kimble, G. A. (1993). Evolution of the nature–nuture issue in the history of psychology. In R. Plomin & G. E. McClearn (Eds.), *Nature, nuture and psychology*. Washington, DC: American Psychological Association.

Krueger, R., Moffitt, T. E., Caspi, A., & Bleske, A. (in press). Assortative mating for antisocial behavior: Developmental and methodological considerations. *Behavior Genetics*.

Livesley, W. J., Jang, K. L., Jackson, D. N., & Vernon, P. A. (1993). Genetic and environmental contributions to dimensions of personality disorder. *American Journal of Psychiatry, 150,* 1826–1831.

Loehlin, J. C. (1992). *Genes and environment in personality development*. Newbury Park, CA: Sage.

Lyons, M. J., True, W. R., Eisen, S. A., Goldberg, J., Meyer, J. M., Faraone, S. V., Eaves, L. J., & Tsuang, M. T. (1995). Differential heritability of adult and juvenile antisocial traits. *Archives of General Psychiatry, 52,* 906–915.

Lytten, H. (1977). Do parents create, or respond to, differences in twins? *Developmental Psychology, 13,* 456–459.

Manke, B., McGuire, S., Reiss, D., Hetherington, E. M., & Plomin, R. (1995). Genetic contributions to adolescent's extrafamilial social interactions: Teachers, best friends, and peers. *Social Development, 4,* 238–256.

McArdle, J. J., & Goldsmith, H. H. (1990). Alternative common factor models for multivariate biometric analyses. *Behavior Genetics, 20,* 569–608.

McCartney, K., Harris, M. J., & Bernieri, F. (1990). Growing up and growing apart: A developmental meta-analysis of twin studies. *Psychological Bulletin, 107,* 226–237.

McClearn, G. E., Plomin, R., Gora-Maslak, G., & Crabbe, J. C. (1991). The gene chase in behavioral science. *Psychological Science, 2,* 222–229.

McClearn, G. E., Taranino, L. M., Rodriguez, L., Jones, B., Blizard, D., & Plomin, R. (1997). Genotypic selection provides experimental confirmation for an alcoholic consumption quantitative trait locus in mouse. *Molecular Psychiatry, 2,* 486–489.

McGue, M., Bacon, S., & Lykken, D. T. (1993). Personality stability and change in early adulthood: A behavioral genetic analysis. *Developmental Psychology, 29,* 96–109.

McGue, M., & Christensen, K. (1997). Genetic and environmental contributions to depression symptomatology: Evidence from Danish twins 75 years of age and older. *Journal of Abnormal Psychology, 106,* 439–448.

McGue, M., & Lykken, D. T. (1992). Genetic influence on risk for divorce. *Psychological Science, 3,* 368–373.

McGue, M., Sharma, A., & Benson, P. (1996). The effect of common rearing on adolescent adjustment: Evidence from a US adoption cohort. *Developmental Psychology, 32,* 604–613.

McGuffin, P., Owen, M. J., O'Donovan, M. C., Thapar, A., & Gottesman, I. I. (1994). *Seminars in psychiatric genetics*. London: Gaskell Press.

McHale, S. M., & Pawletko, T. M. (1992). Differential treatment of siblings in two family contexts. *Child Development, 63,* 68–81.

Mednick, S. A. (1978). Berkson's fallacy and high-risk research. In L. C. Wynne, R. L. Cromwell, & S. S. Matthysse (Eds.), *The nature of schizophrenia: New approaches to research and treatment* (pp. 442–452). New York: Wiley.

Miles, D. R., & Carey, G. (1997). Genetic and environmental architecture to human aggression. *Journal of Personality and Social Psychology, 72,* 207–217.

Moffitt, T. (1993). Adolescence-limited and life-course-persistent antisocial behaviour: A developmental taxonomy. *Psychological Review, 100,* 674–701.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Boston: Kluwer.

Neiderhiser, J. M., O'Connor, T. G., Chipuer, H., Reiss, D., Hetherington, E. M., & Plomin, R. (in press). *Adding stepfamilies to genetic designs*. Manuscript submitted for publication.

Nigg, J. T., & Goldsmith, H. H. (1994). Genetics of personality disorders: Perspectives from personality and psychopathology research. *Psychological Bulletin, 115,* 346–380.

O'Connor, T. G., Deater-Deckard, K., Fulker, D., Rutter, M., & Plomin, R. (in press). Genotype–environment correlations in late childhood and early adolescence: Antisocial behavioral problems and coercive parenting. *Developmental Psychology*.

O'Connor, T. G., Hetherington, E. M., & Reiss, D. (in press). Family systems and adolescent development: Shared and nonshared risk and protective factors in nondivorced and remarried families. *Development and Psychopathology*.

O'Connor, T. G., Hetherington, E. M., Reiss, D., & Plomin, R. (1995). A twin–sibling study of observed parent–adolescent interactions. *Child Development, 66,* 812–829.

O'Connor, T. G., McGuire, S., Reiss, D., Hetherington, E. M., & Plomin, R. (1998). Co-occurrence of depressive symptoms and antisocial behaviour in adolescence: A common genetic liability. *Journal of Abnormal Psychology, 107,* 27–37.

O'Connor, T. G., Neiderhiser, J. M., Reiss, D., Hetherington, E. M., & Plomin, R. (1998). Genetic contributions to continuity, change and co-occurrence of antisocial and depressive symptoms in adolescence. *Journal of Child Psychology and Psychiatry, 39,* 323–326.

Palmer, C. G., Bailey, J. N., Ramey, C., Cantwell, D., De'Homme, M., McGough, J., Woodward, J., Saranow, R., Asarnow, J., Smalley, S., & Nelson, S. (1997). Possible evidence of DATI gene by sex interaction in susceptibility to dopamine transporter [abstract]. *American Journal of Medical Genetics (Neuropsychiatric Genetics), 74,* 630.

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Catalia.

Peltonen, L. (1995). All out for chromosome six. *Nature, 378,* 665–666.

Pike, A., McGuire, S., Reiss, D., Hetherington, E. M., & Plomin, R. (1996). Family environment and adolescent depressive symptoms and antisocial behaviour: A multivariate genetic analysis. *Developmental Psychology, 32,* 574–589.

Plomin, R. (1983). Developmental behavior genetics. *Child Development, 54,* 253–259.

Plomin, R. (1986). *Development, genetics, and psychology*. Hillsdale, NJ: Erlbaum.

Plomin, R. (1994). *Genetics and experience*. Newbury Park, CA: Sage.

Plomin, R. (1995). Genetics and children's experiences in the family. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 36,* 33–68.

Plomin, R., & Bergeman, C. S. (1991). The nature of nurture: Genetic influences on "environmental" measures. *Behavioral and Brain Sciences, 14,* 373–427.

Plomin, R., Corley, R. DeFries, J. C., & Fulker, D. W. (1990). Individual differences in television viewing in early childhood: Nature as well as nurture. *Psychological Science, 1,* 371–377.

Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from one another? *Behavioral and Brain Sciences, 10,* 1–16.

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype–environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin, 84,* 309–322.

Plomin, R., DeFries, J. C., McClearn, G. E., & Rutter, M. (1997). *Behavioral genetics* (3rd ed.). New York: Freeman.

Plomin, R., Lichtenstein, P., Pedersen, N. L., McClearn, G. E., & Nesselroade, J. R. (1990). Genetic influences on life events during the last half of the life span. *Psychology and Aging, 5,* 25–30.

Plomin, R., & McClearn, G. E. (1993). *Nature, nurture, and psychology*. Washington, DC: American Psychological Association.

Plomin, R., Nitz, K., & Rowe, D. C. (1990). Behavioral genetics and aggressive behavior in childhood. In M. Lewis & S. M. Miller (Eds.), *Handbook of developmental psychopathology* (pp. 186–205). New York: Plenum.

Plomin, R., Owen, M. J., & McGuffin, P. (1994). The genetic basis of complex human behviors. *Science, 264,* 1733–1739.

Plomin, R., Reiss, D., Hetherington, E. M., & Howe, G. (1994). Nature and nurture: Genetic contributions to measures of the family environment. *Developmental Psychology, 30,* 32–43.

Plomin, R., & Rowe, D. C. (1979). Genetic and environmental etiology of social behavior in infancy. *Developmental Psychology, 15,* 62–72.

Plomin, R., & Rutter, M. (in press). Child development and molecular genetics: What do we do with the genes once they are found? *Child Development*.

Puig-Antich, J., Goetz, D., Davies, M., Kaplan, T., Davies, S., Ostrow, L., Asnis, L., Twomey, J., Iyengar, S., & Ryan, N. D. (1989). A controlled family history study of prepubertal major depressive disorder. *Archives of General Psychiatry, 46*, 406–418.

Quittner, A. L., & Opipari, L. C. (1994). Differential treatment of siblings: Interview and diary analyses comparing two family contexts. *Child Development, 65*, 800–814.

Reiss, D. (1995). Genetic influence on family systems: Implications for development. *Journal of Marriage and the Family, 57*, 543–560.

Reiss, D, Hetherington, E. M., Plomin, R., Howe, G. W., Simmens, S. J., Henderson, S. H., O'Connor, T. G., Bussell, D. A., Anderson, E. R., & Law, T. C. (1995). Genetic questions for environmental studies: Differential parenting and psychopathology in adolescence. *Archives of General Psychiatry, 52*, 925–936.

Reiss, D., Plomin, R., & Hetherington, E. M. (1993). Genetics and psychiatry: An unheralded window on the environment. *American Journal of Psychiatry, 148*, 283–291.

Reiss, D., Plomin, R., & Hetherington, E. M., Howe, G., Rovine, M., Tryon, A., & Stanley, M. (1994). The separate worlds of teenage siblings: An introduction to the study of the nonshared environment and adolescent development. In E. M. Hetherington, D. Reiss, & R. Plomin (Eds.), *Separate social worlds of siblings: Importance of nonshared environment on development* (pp. 63–109). Hillsdale, NJ: Erlbaum.

Risch, N., & Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science, 273,* 1516–1517.

Rodgers, J. L., Rowe, D. C., & Li, C. (1994). Beyond nature versus nurture: DF analysis of nonshared influences on problem behaviors. *Developmental Psychology, 30*, 374–384.

Rowe, D. (1981). Environmental and genetic influences on perceived parenting: A twin study. *Developmental Psychology, 17*, 203–208.

Rutter, M. (1991). Nature, nurture, and psychopathology: A new look at an old topic. *Development and Psychopathology, 3*, 125–136.

Rutter, M., Bailey, A., Simonoff, E., & Pickles, A. (in press). Genetic influences and autism. In F. Volkmer & D. Cohen (Eds.), *Handbook of autism*. New York: Wiley.

Rutter, M., Champion, L., Quinton, D., Maughan, B., & Pickles, A. (1995). Understanding individual differences in environmental risk exposure. In P. Moen, G. H. Elder, & K. Luscher (Eds.), *Examining lives in context: Perspectives on the ecology of human development* (pp. 61–93). Washington, DC: American Psychological Association.

Rutter, M., Dunn, J., Plomin, R., Simonoff, E., Pickles, A., Maughan, B., Ormel, H., Meyer, J., & Eaves, L. (1997). Integrating nature and nurture: Implications of person–environment correlations and interactions for developmental psychopathology. *Development and Psychopathology, 9*, 335–364.

Rutter, M., MacDonald, H., Le Couteur, A., Harrington, R., Bolton, P., & Bailey, A. (1990). Genetic factors in child psychiatric disorders, II: Empirical findings. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 31,* 39–83.

Rutter, M., & Plomin, R. (1997). Opportunities for psychiatry from genetic findings. *British Journal of Psychiatry, 171*, 209–219.

Rutter, M., & Quinton, D. (1984). Parental psychiatric disorder: Effects on children. *Psychological Medicine, 14*, 853–880.

Rutter, M., & Redshaw, J. (1991). Annotation: Growing up as a twin: Twin–singleton differences in psychological development. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 32*, 885–895.

Rutter, M., Simonoff, E., & Silberg, J. (1993). How informative are twin studies of child psychopathology? In T. J. Bouchard & P. Propping (Eds.), *Twins as a tool of behavioural genetics* (pp. 433–458). New York: Wiley.

Rutter, M., & Smith, D. J. (Eds.) (1995). *Psychosocial disorders in young people: Time trends and their causes.* Chichester, UK: Wiley.

Scarr, S. (1992). Developmental theories for the 1990s: Development and individual differences. *Child Development, 63*, 1–19.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of gene environment effects. *Child Development, 54*, 424–435.

Schulte-Korne, G., Grimm, T., Nothen, M. M., Muller-Myhsok, B., Propping, P., & Remschmidt, H. (1997). Evidence for linkage of spelling disability to chromosome 15 [abstract]. *American Journal of Medical Genetics (Neuropsychiatric Genetics), 74*, 661.

Schwab, S. G., Eckstein, G., Hallmayer, J., Lerer, B., Albus, M., Borrman, M., Lichterman, D., Ertl, M., Maier, W., & Wildenauer, D. (1997). Evidence suggestive of a locus on chromosome 5q31 contributing to susceptibility for schizophrenia in German and Israeli families by multipoint affected sibpair linkage analysis. *Molecular Psychiatry, 2*, 156–160.

Scott, J. P. (1977). Social genetics. *Behavior Genetics, 7,* 327–346.

Sherrington, R., Rogaev, E., Liang, Y., Rogaev, E. A., Levesque, G., Ikeda, M., Chi, H., Lin, C., Li, G., & Holman, K. (1995). Cloning a gene bearing mis-sense mutations in early-onset familial Alzheimer's disease. *Nature, 375,* 754–760.

Sigvardsson, S., Bohman, M., & Cloninger, C. R. (1996). Replication of Stockholm adoption study of alcoholism. *Archives of General Psychiatry, 53,* 681–687.

Silberg, J. L., Erickson, M. T., Meyer, J. M., Eaves, L. J., Rutter, M. L., & Hewitt, J. K. (1994). The application of structural equation modelling to maternal ratings of twins' behavioral and emotional problems. *Journal of Consulting and Clinical Psychology, 62,* 510–521.

Silberg, J. L., Rutter, M., Meyer, J., Maes, H., Hewitt, J., Simonoff, E., Pickles, A., Loeber, R., & Eaves, L. (1996). Genetic and environmental influences on the covariation between hyperactivity and conduct disturbance in juvenile twins. *Journal of Child Psychology and Psychiatry, 37,* 803–816.

Simonoff, E. (1992). A comparison of twins and singletons with child psychiatric disorders: An item sheet study. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 33,* 1319–1332.

Simonoff, E., Pickles, A., Hewitt, J., Silberg, J. (1995). Multiple raters of disruptive child behavior: Using a genetic strategy to examine shared views and bias. *Behavior Genetics, 25,* 311–326.

Slutske, W., Heath, A. C., Dinwiddie, S., Madden, P., Bucholz, K., Dunne, M., Statham, D., & Martin, N. G. (1997). Modeling genetic and environmental influences in the etiology of conduct disorder: A study of 2,682 adult twin pairs. *Journal of Abnormal Psychology, 106,* 266–279.

Sunohara, G., Bar, C. L., Jin, U., Scahcher, R., Roberts, W., Tannock, R., Malone, M., & Kennedy, J. L. (1997). Is the dopamine D4 receptor gene associated with children and adults with attention-deficit-hyperactivity disorder? [abstract]. *American Journal of Medical Genetics (Neuropsychiatric Genetics), 74,* 629.

Thapar, A., & McGuffin, P. (1995). Are anxiety symptoms in childhood heritable? *Journal of Child Psychology and Psychiatry and Allied Disciplines, 36,* 439–447.

Tienari, P. Wynne, L. C., Moring, J., Lahti, I., Naarala, M., Sorri, A., Wahlberg, K.-E., Saarento, O., Seitmaa, M., Kaleva, M., & Laksy, K. (1994). The Finnish adoption study of schizophrenia: Implications for family research. *British Journal of Psychiatry, 164* (Suppl. 23), 20–26.

Todd, R. D., Neuman, R., Geller, B., Fox, L. W., & Hickok, J. (1993). Genetic studies of affective disorders: Should we be starting with childhood onset probands? *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 1164–1171.

Torrey, E. F., Bowler, A. E., Taylor, E. H., & Gottesman, I. I. (1994). *Schizophrenia and manic-depressive disorder.* New York: Basic Books.

True, W. R., Rice, J., & Heath, A. C. (1994). Genetic and environmental influences of twins in post-traumatic stress. *Archives of General Psychiatry, 51,* 838–839.

Turkheimer, E., & Gottesman, I. I. (1991). Individual differences and the canalization of human behavior. *Developmental Psychology, 27,* 18–22.

Tsuang, M., & Faraone, S. D. (1990). *The genetics of mood disorders.* Baltimore: Johns Hopkins University Press.

Van den Oord, E. J., Koot, H. M., Boomsma, D., & Verhulst, F. C. (1995). A twin–singleton comparison of problem behavior in 2–3-year-olds. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 36,* 449–458.

Van den Oord, E. J., & Rowe, D. C. (1997). Continuity and change in children's social maladjustment: A developmental behavior genetic study. *Developmental Psychology, 33,* 319–332.

Van den Oord, E. J., Verhulst, F. C., & Boomsma, D. (1996). A genetic study of maternal and paternal ratings of problem behaviors in 3-year-old twins. *Journal of Abnormal Psychology, 105,* 349–357.

Van Os, J., Castle, D. J., Takei, N., Der, G., & Murray, R. M. (1996). Psychotic illness in ethnic minorities: Clarification from the 1991 census. *Psychological Medicine, 26,* 203–208.

Vogel, E. F., & Bell, N. W. (1960). The emotionally disturbed child as the family scapegoat. In N. W. Bell & E. F. Vogel (Eds.), *A modern introduction to the family* (pp. 382–397). Glencoe, IL: Free Press.

Wahlberg, K. E., Wynne, L. C., Oja, H., Keskitalo, P., Pykalainen, L., Lahti, I., Moring, J., Naarala, M., Sorri, A., Seitmaa, M., Laksy, K., Kolassa, J., & Tienari, P. (1997). Gene–environment interaction in vulnerability to schizophrenia: Findings from the Finnish adoptive family study of schizophrenia. *American Journal of Psychiatry, 154,* 355–362.

Wahlsten, D. (1990). Insensitivity of the analysis of variance to heredity–environment interaction. *Behavior and Brain Sciences, 13,* 109–161.

Webster-Stratton, C., Hollinsworth, T., & Kolpacoff, M. (1989). The long-term effectiveness and clinical significance of three cost-effective training programs for families with conduct-problem children. *Journal of Consulting and Clinical Psychology, 57,* 550–553.

Weissman, M. M., Wickramaratne, P., Merikangas, K. R., Leckman, J. F., Prusoff, B. A., Caruso, K. A., Kidd, K. K., & Gammon, G. D. (1984). Onset of major depression in early adulthood: Increased familial loading and specificity. *Archives of General Psychiatry, 41,* 1136–1143.

Wrate, R. M., Rothery, D. J., McCabe, R. J. R., Aspin, J., & Bryce, G. (1995). Psychosocial adversity in an adolescent psychiatry inpatient cohort: WHO *ICD-10* findings from a multicentre prospective study. *Social Psychiatry and Psychiatric Epidemiology, 30,* 5–13.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.05
Psychobiology

NIELS BIRBAUMER
University of Tübingen, Germany
and
HERTA FLOR
Humboldt-University, Berlin, Germany

1.05.1 PSYCHOBIOLOGY AND CLINICAL PSYCHOLOGY 117
  1.05.1.1 Terminology 117
  1.05.1.2 The Three-systems Approach 117
1.05.2 GENETICS AND CLINICAL PSYCHOLOGY 118
1.05.3 PSYCHONEUROENDOCRINOLOGY AND PSYCHONEUROIMMUNOLOGY 118
  1.05.3.1 Introduction and Definitions 118
  1.05.3.2 Hormones, Bodily Rhythms, and Behavior 119
  1.05.3.3 Hormones and Emotions 119
  1.05.3.4 Relationships Between Nervous and Immune System 120
  1.05.3.5 Stress and the Immune System 120
  1.05.3.6 Learning and the Immune System 120
1.05.4 NEUROCHEMISTRY AND NEUROPHARMACOLOGY 121
  1.05.4.1 Psychopharmacology and Clinical Psychology 121
  1.05.4.2 Mechanisms of Action of Psychoactive Substances 122
    1.05.4.2.1 Neuronal mechanisms 122
    1.05.4.2.2 Receptors for neurochemicals 122
  1.05.4.3 Transmitter Systems 123
    1.05.4.3.1 Amino acids 123
    1.05.4.3.2 Amines 123
    1.05.4.3.3 Opioids 123
    1.05.4.3.4 Neuropeptides 124
    1.05.4.3.5 Acetylcholine 124
1.05.5 PSYCHOPHYSIOLOGY 124
  1.05.5.1 Introduction 124
  1.05.5.2 Concepts in Psychophysiology 124
    1.05.5.2.1 The three-systems approach 124
    1.05.5.2.2 The law of initial values 125
    1.05.5.2.3 Activation, arousal, and directional fractionation 125
    1.05.5.2.4 Stimulus–response specificity 125
    1.05.5.2.5 Individual response stereotopies 126
    1.05.5.2.6 Orienting, habituation, and defense 126
    1.05.5.2.7 Social psychophysiology 126
  1.05.5.3 Methods in Psychophysiology 127
    1.05.5.3.1 Electroencephalography 127
    1.05.5.3.2 Magnetoencephalography (MEG) 127
    1.05.5.3.3 Functional magnetic resonance imaging (fMRI) 127
    1.05.5.3.4 Positron emission tomography (PET) 129
    1.05.5.3.5 Electromyography (EMG) 129

1.05.5.3.6 Electrodermal activity (EDA) 129
1.05.5.3.7 Heart activity and blood pressure 131
1.05.5.4 Applications of Psychophysiology 131
1.05.5.4.1 Biofeedback 131
1.05.5.4.2 Biofeedback of slow cortical potentials 132
1.05.5.4.3 EMG biofeedback 132
1.05.5.4.4 Detection of deception 132
1.05.5.4.5 Clinical psychophysiology 133

1.05.6 CONSCIOUSNESS AND ATTENTION 133

1.05.6.1 Introduction 133
1.05.6.2 Psychophysiology of Consciousness 133
1.05.6.3 The Ascending Reticular Activating System 134
1.05.6.3.1 Neurophysiology of tonic and phasic activation 134
1.05.6.3.2 Transmitter systems involved in attention 134
1.05.6.3.3 Gating of attention 135
1.05.6.4 The Psychophysiology of Selective Attention 135
1.05.6.5 Event-related Potentials as Indicators of Attentional Processes 136
1.05.6.5.1 N1/P1 136
1.05.6.5.2 N2 136
1.05.6.5.3 P3 137

1.05.7 SLEEP, DREAM, CIRCADIAN RHYTHMS, AND SLEEP DISORDERS 137

1.05.7.1 Introduction 137
1.05.7.2 Circadian Clocks 138
1.05.7.3 Sleep and Dream 139
1.05.7.4 The Neurobiology of Sleep 140
1.05.7.5 Psychophysiology of Sleep Stages 140
1.05.7.6 Sleep Disorders 141

1.05.8 LEARNING, PLASTICITY, AND DISORDERS OF LEARNING AND MEMORY 141

1.05.8.1 Introduction 141
1.05.8.2 Mechanisms of Plasticity 142
1.05.8.3 The Formation of Memory Traces 142
1.05.8.4 Neuropsychology of Memory 143
1.05.8.5 Cellular Mechanisms of Learning 144
1.05.8.6 Neuronal Bases of Learning and Memory 145
1.05.8.7 Neurochemistry of Memory 145

1.05.9 MOTIVATION AND MOTIVATION-RELATED DISORDERS 147

1.05.9.1 Introduction 147
1.05.9.2 Hunger and the Eating Disorders 147
1.05.9.3 Sexual Function and the Sexual Disorders 148
1.05.9.3.1 The sexual response 148
1.05.9.3.2 Sexual differentiation and sex hormones 149
1.05.9.3.3 Hormonal basis of sexual behavior 149
1.05.9.3.4 Neuronal mechanisms of sexual behavior 150
1.05.9.3.5 Influence of sex hormones on the brain 150
1.05.9.3.6 Brain processes and homosexuality 151
1.05.9.4 Acquired Motivation and Substance-induced Disorders 151
1.05.9.4.1 The neurobiology of addiction 152
1.05.9.4.2 Learned motivation and addiction 152

1.05.10 EMOTION AND THE EMOTIONAL DISORDERS 154

1.05.10.1 Introduction 154
1.05.10.2 Fear and Anxiety and the Anxiety Disorders 155
1.05.10.2.1 Avoidance, fear, and anxiety 155
1.05.10.2.2 Startle reflex modulation and fear 155
1.05.10.2.3 Psychopharmacology of fear and anxiety 156
1.05.10.3 Sadness, Depression, and the Affective Disorders 156
1.05.10.3.1 Sadness and depression 156
1.05.10.3.2 Neurochemistry of depression 157
1.05.10.4 Aggression and Aggression-related Disorders 158
1.05.10.4.1 Development of aggression 158
1.05.10.4.2 Neuronal basis of aggression 158
1.05.10.4.3 Hormonal basis of aggression 158
1.05.10.4.4 Psychopathy 159
1.05.10.5 The Neocortical Hemispheres and Emotion 159
1.05.10.6 Behavioral Medicine: Application of the Psychophysiology of Emotion to Disease 159

1.05.11 COGNITIVE PROCESSES AND COGNITIVE DISORDERS 160

1.05.11.1 Introduction 160
1.05.11.2 Functions of the Cerebral Hemispheres 160
1.05.11.3 Evolution and Neurophysiology of Language 162
1.05.11.3.1 Language disorders 164

1.05.11.4 The Association Areas of the Neocortex 164
1.05.11.4.1 Parietal cortex 164
1.05.11.4.2 The frontal lobe 166
1.05.11.4.3 The temporal lobe 168
1.05.11.5 Thought Disorders 168
1.05.12 SUMMARY 170
1.05.13 REFERENCES 170

1.05.1 PSYCHOBIOLOGY AND CLINICAL PSYCHOLOGY

When I carefully consider the curious habits of dogs I am compelled to conclude that man is the superior animal. When I consider the curious habits of man I confess, my friend, I am puzzled. (Ezra Pound)

1.05.1.1 Terminology

Psychobiology describes the relationship between psychological and behavioral processes and the underlying physiological mechanisms. It is not only brain functions that are the scope of psychobiology but also the interplay of all bodily activities with behavior and cognition. The term biological psychology is often used synonymously. While physiological psychology and behavioral neuroscience deal with the CNS (central nervous system)–behavior interaction, psychobiology encompasses both peripheral–physiological and CNS functions. Neuropsychology applies a comparable strategy but uses naturally occurring lesions in humans.

1.05.1.2 The Three-systems Approach

All disorders of behavior and thought have a physiological basis, and virtually all physical disorders are strongly related to deviant behavior. Therefore, any separation between mind and body, or brain and behavior, is unjustified. For each disorder, the specific psychological and physical variance must be specified. Any principal distinction between the two levels of observation compromises progress toward the understanding of the etiology of diseases. The psychobiology of disease is an integral part of clinical psychology; any attempt to ignore its significance will lead to misleading conclusions.

Behavior in general can be measured in humans at the level of *verbal reports*, often called subjective level, the level of *physiological responses* (including hormonal, endocrine, and immunological), and the level of *overt (motor) responses*. The three response levels are usually elicited by stimuli in the environment. Their impact on the three systems is defined by their *meaning* or *significance*. As we shall see in Section 1.05.11.3, the construction of meaning is based on a well-described associative learning mechanism in the CNS. The creation of meaning consists of a relatively fast *automatic* process (for simple stimuli it lasts up to 200 ms) which in most instances does not need conscious, *controlled* information processing.

One of the main reasons which makes the *simultaneous* observation and measurement of all three response levels mandatory in clinical psychology is their generally *low correlation* in a given stimulus context. Exceptions, such as extremely intensive emotions (e.g., terror), are rare. The problem becomes particularly obvious if we refer from a given response channel (e.g., an angry outburst) to an *underlying construct* such as an emotion. The different measures of the given response would be expected to be highly correlated with each other because they would be indicators of the same construct.

However, as shown by Öhman (1987), it is quite complicated to decide what a "high correlation" means in this instance. Since behavior, physiology, and verbal reports represent very diverse data domains, most of the variance in each of the measures would be attributable to factors that would not contribute to the variance in the others. For example, a peripheral–physiological measure such as heart rate is primarily constrained by the design and demands of the cardiovascular system, and the variance left for psychological factors is necessarily small. Similarly, verbal reports are shaped by cultural conventions including linguistic limitations which may not affect heart rate at all. Since a high correlation between two measures requires that their variance is determined by the same sources in approximately the same proportions, high correlations between, for example, heart rate and verbal reports would be unlikely. In fact, using the theorems of factor analysis for a theoretical treatment of this problem, Öhman has shown that the maximal between-response channel correlation one could realistically expect in the present case would be 0.30 rather than 0.70. Indeed, even within one of the response domains, for example, within autonomic reponses, covariations between response measures and various emotional stimuli often barely reach statistical significance, typically ranging between 0.2 and 0.3 (Fahrenberg, Walschburger, & Foerster, 1979). Given this limitation on the expected covariation, one would be likely to encounter frequent cases where, for example, behavior and physiology would provide discordant bases for inferring an emotion (for a more extended discussion of the covariation problem see Birbaumer & Öhman, 1993).

In stressing the necessity of physiological definitions and measurements in clinical psychology, particularly neurophysiological and neurochemical studies, we emphasize that physiology cannot substitute for, but rather is dependent on, the other levels. Even with perfect knowledge about the neurophysiological mechanisms, the definition of a disordered behavior requires, in addition, the contextual–situational, the subjective–verbal, and the behavioral components in order to be complete.

The covariation or a lack of covariation between the three response levels frequently becomes the main source of variance in the etiology or maintenance of a disorder. In panic disorder, for example, the aberrant subjective perception and evaluation of cardiovascular responses constitutes the main etiological factor involved (lack of correlation between response levels). On the other hand, in phobics, Lang, Melamed, and Hart (1970) have convincingly demonstrated that a high correlation between the physiological response (cardiovascular in this case), expressive motor behavior (avoidance), and verbal response (fear) provides the best predictor for fast extinction of the response. In some forms of schizophrenia, the discordance and variability between some of the response systems is often regarded as the core of the disorder (see Section 1.05.11.5).

Since physiological mechanisms play such a pivotal role in the etiology of most disorders of behavior, their *assessment* is necessary for planning and documenting *psychological treatment*. Most treatment outcome studies in clinical psychology and psychiatry are satisfied with the documentation of changes in a verbal report or, in the more advanced publications, include systematic behavioral observations. Psychophysiological measurements are more expensive and time-consuming and require appropriate training. Therefore, many clinical psychologists and psychiatrists avoid their application. As will become apparent in the following sections, a lack of acceptance and credibility of psychological treatments and high rates of relapse related to insufficient treatment outcome measures constitute the negative consequences of neglecting the psychobiological response system.

1.05.2 GENETICS AND CLINICAL PSYCHOLOGY

The knowledge of the principles of genetics is critical for the clinical psychologist for several reasons. The genetic make-up of an individual not only determines the limits of systematic behavior modification and treatment but in many disorders allows the guidance of interventional strategies toward the plastic gene-behavior chains and the investment of modification efforts in already fixed and rigid behavioral expressions of genetic activity.

Modern genetics teaches us that the genetic apparatus is a highly flexible system which allows environmental (learning) factors to influence its functions at most if not all levels of synthesis and expression. A nature—nurture polarization, still popular among behavioral and molecular scientists and the public, is *incompatible* with our present state of knowledge. Clinical psychology should incorporate these new scientific advances to target interventions according to the existing biological limits and possibilities. A detailed discussion of genetic foundations of clinical psychology can be found in Chapter 1.04, this volume.

1.05.3 PSYCHONEUROENDOCRINOLOGY AND PSYCHONEUROIMMUNOLOGY

1.05.3.1 Introduction and Definitions

Psychoneuroendocrinology and psychoneuroimmunology deal with the study of the relationship between the hormonal and immune system and behavior, particularly in humans. Behavioral endocrinology and behavioral immunology use mostly animals for the same purpose. Several excellent texts are available for an in-depth review of both fields (cf., Ader, Felten, & Cohen, 1990; Becker, Breedlove, & Crews, 1992; Nelson, 1994). Here we provide a short overview of the main topics of both fields without a repetition of basic immunology or endocrinology. Both can be found in the textbooks on physiology and the above-mentioned introductions.

There are intricate relationships between the nervous system, the immune system, and the endocrine system. Behavior (on all three levels: motor, cognitive, and physiological) is not directly influenced by the two systems but hormones and immune factors enter the nervous system and change behavior through that system. It is important to note that there are reciprocal interactions between the nervous system and behavior: behavior is not only affected by the hormonal and immune system but together with environmental stimuli and the consequences of a particular behavior, behavior itself affects again the two "slow" bodily systems.

Hormones or immune factors can have *organizing* or *activating* effects. An organizing effect would, for example, be the specific formation of hypothalamic nuclei during intrauterine development which later in life cause heterosexual or homosexual behavior (see Section 1.05.9.3.6). An activating effect consists, for example, of the unspecific secretion of ACTH after stressful stimuli. Hormones influence sensory systems, the CNS, as well as the effectors. The hormonal and the immune system are phylogenetically very old systems (from insect to humans with similar structure and function) and act relatively slowly on behaviorally relevant structures. Slow means that the direct actions of the nervous system on muscles are usually much faster (in the millisecond range), whereas the "wet" bloodstream-dependent hormones and immune structures exert their effects within seconds, hours, days, or months. Both fields become increasingly important for clinical psychology and behavior modification: many clinical psychological and psychiatric disorders are caused or at least in some aspects influenced by endocrine and/or immunological agents. Additionally, the tailoring of psychological intervention strategies to endocrine and immunological changes in patients is extremely important, as can be seen from the intricate interplay of circadian hormonal rhythms and depression, sex hormones and sexual deviations, and in the modification of antisocial behavior, treatment of obesity and eating disorders, and stress-related disorders.

1.05.3.2 Hormones, Bodily Rhythms, and Behavior

Nearly all hormones are synthesized and/or secreted in a rhythmic manner, triggered by endogenous oscillators in the CNS which can be *entrained* to exogenous (mainly light) or psychological (mostly social) stimuli within certain limits.

1.05.3.3 Hormones and Emotions

The hormonal reaction to separation and loss constitutes the most intensively studied emotional reaction because of its obvious relevance to the pharmacotherapy of depression. Disruption of the circadian rhythm of GH, ACTH, and cortisol is accompanied by an initial increase in some endogenous opioids in the periventricular gray and a depletion of noradrenaline (NA) and serotonin (5-HT). A productive research direction to understand antidepressive agents might be the insight into the inverse of separation, namely social attachment and love.

Attachment and love certainly contain strong elements of psychological dependency and in the case of separation, withdrawal reactions in the form of sadness or jealousy. Therefore, the considerations in Section 1.05.9.4 concerning drug dependency and the dopamine system are also relevant for love and attachment. All substances which stimulate the limbic dopamine system (opiates, benzodiazepines, and alcohol) reduce the symptoms of separation withdrawal. Particularly effective are direct injections of opiates in the ventral tegmental dopamine system; injection of naloxone, an opiate antagonist, produces cries of separation in rats and monkeys (Wise, 1988).

Affiliative behavior is, however, *heterogeneous* (pair bonding, parental care, mutual defense, sexual interaction, and altruistic–helpful behavior), each behavior will correlate with a specific neurohormonal profile (*response stereotopy*). Situational factors such as touch, smell, eye contact, and temperature play a central role in the formation of those emotion-specific hormonal profiles (*stimulus stereotopy*). The lack of physical contact between many developing mammals including humans and a caregiver prevents slow wave sleep and GH secretion and causes psychosocial dwarfism and autistic social behavior. For example, tactile stimulation in rat pups causes synthesis of the enzyme ornithine decarboxylase (ODC); separation stops the synthesis and brush strokes analogous to the mother's licking restore ODC production. ODC is an enzyme essential for tissue growth and development.

The peptide hormone oxytocin was for a long time known to be the central hormone in the initiation of milk ejection and uterine contractions for viviparity. However, extrahypothalamic hypophyseal oxytocin plays a totally different role from hypothalamic and brain stem-produced oxytocin. The latter is universally important for initiation of maternal behavior, sexual attraction, grooming, and monogamy. It seems as if a certain critical level of oxytocin has to be present in the hypothalamus in order to make social interaction rewarding (Schulkin, 1993). Central β-endorphin and oxytocin both act in an agonistic fashion on the central dopamine psychomotor reward system (see Section 1.05.9.4.2). The behavioral effects of oxytocin and endorphins depend more upon the location of their synthesis in the brain and less on their biochemical or physiological actions on nerve cells or glands. Both are strongly influenced by the circadian rhythm and suppressed during negative emotions such as stress, anxiety, and depression. Soft tactile stimulation on nipples and genitals in social interaction increase their production. Stress, anxiety, and depression are discussed in Section 1.05.10.

1.05.3.4 Relationships Between Nervous and Immune System

The nervous system has various direct and indirect (mostly) endocrine connections with different (but not all) parts of the immune system and vice versa: immune factors regulate and modify several neuronal responses mainly through endocrine effects on the CNS but also within the nervous system through direct activation of immune cells by psychological and physical stimuli.

Several substances in the nervous system (NS) constitute signals for the immune system: these are the main neurotransmitters such as NA, acetylcholine (Ach), and excitatory amino acids, such as glutamate. Neuropeptides modulate immune function through neurotransmitters or directly. Most important are substance P, VIP (vasoactive intestinal peptide), and endogenous opioids. Endocrine growth factors and cytokines (interleukines and interferons) also act as neuroactive messengers.

The immune system, on the other hand, sends signals to the nerve cells through cytokines, neuropeptides, and neurotrophins (such as GH). Immune cells possess receptors for neurotransmitters and peptides, whereas nerve cells are affected mostly indirectly through receptors for neuropetides, histamine, and neurotrophins. Nerve cell metabolism is, however, regulated by glia cells which contain cytokine receptors in addition to all other types of membrane receptors (cf., Schedlowski & Tewes, 1996).

1.05.3.5 Stress and the Immune System

There is agreement that some but not all types of stressors have an immunosuppressive effect which may later contribute to initiation or maintenance of illness and "sickness behavior" during infections or other immune disorders. Depending upon the type of stress, very different and highly specific responses of immune cells result. The main effects of stress on the immune system are mediated by the ACTH-glucocorticoid release of the anterior pituitary–adreno–cortical axis. Glucocorticoids have immunosuppressive effects on some but not all cytokines (i.e., cytotoxic T lymphocytes, $CD8^+$). The immunosuppressive effect may be desirable or adverse for the organism. Glucocorticoids show their immunosuppressive effect mainly on nonactivated immune cells. If they have already been activated by antigen presentation, their effect is negligible. Therefore, in humans, stress increases the probability for newly acquired infections but rarely affects already active immune diseases.

Learned helplessness stress (see Section 1.05.10.3.2) dramatically reduces $CD4^+$ T helper cells through glucocorticoid increase and leaves the organism more vulnerable to new infections. This may explain why in older or already weak populations, loss of the partner results in a significant reduction of survival because of an increased susceptibility to disease (cf., Schedlowski & Tewes, 1996).

On the other hand, an increased plasma level of glucocorticoids after physical stress, for example, blood poisoning, *protects* the organism. Adrenalectomized animals, who are deficient in glucocorticoid release, die from severe stress, whereas intact animals survive severe stress because an excessive immune reaction to external toxins is prevented by the release of glucocorticoids. In addition, some cytokines stimulate the adreno–cortical–pituitary axis and suppress the growth of tumor cells with reactive stimulation of ACTH and cortisol. Therefore, use of the term, "stress-induced immunosuppression" as an indicator of an adverse effect on health and homoeostasis is too general to be useful. It has to be specified which psychological variable leads to which specific immunological change. These changes can be protective or harmful for the organism.

Several studies have shown that *behavioral treatment and social support* have a protective effect on various immune parameters and diseases. In human immunodeficiency virus (HIV)-infected persons, the destruction of $CD4^+$ helper cells and other lymphocytes can be significantly retarded. Survival rates of aging populations can be prolonged for years with appropriate behavioral and social interventions targeted towards improving self-control and self-efficacy. These studies document an increase in lymphocytes and long-term reduction of cortisol levels in the intervention groups (Schedlowski & Tewes, 1996).

1.05.3.6 Learning and the Immune System

Early in the twentieth century, several publications reported the classical conditioning of immune reactions. Most, if not all, of the reports remained unnoticed until 1975, when Ader and Cohen published their seminal paper on classically conditioned immune suppression by using a conditioned taste aversion paradigm in rats (cf., Ader & Cohen, 1985; Ader et al., 1990). The unconditioned stimulus (US) cyclophosphamide (CY) causes severe gastrointestinal disorders and suppression of a host of immune factors. The rats in the experimental group received the neutral saccharin as a conditioned stimulus (CS) and shortly afterwards an injection of CY as US. Three days after the conditioning trial only the CS was given and sheep erythocytes were injected as an antigen-stimulating agent. After nine days the animals were killed and the amount of antibodies against the antigen was counted. A profound reduction of proliferation of T lymphocytes and natural killer (NK) cells was found. The same effect can be induced by the operant procedure of learned helplessness (see Section 1.05.10.3.2): animals in a yoked group that could not escape repetitive aversive shocks showed lasting immune suppression in comparison with the escape group with the same number of shocks presented. Adrenalectomy prevented the effect, demonstrating that the pituitary–adreno–cortical axis is involved in the operant effect but not in the classical conditioning effect; the latter remains after adrenalectomy.

Not only immunosuppression but also activation of immune cells can be classically conditioned, for example, by use of the injection of the cytokine agonist ovalbumine protein as a US. Clinically relevant for human transplantation surgery are experiments demonstrating that skin transplants are much less sensitive to *graft* vs. *host reactions* if the immunosuppressive CS is presented again after transplantation (cf., Ader et al., 1990).

Conditioned immunosuppression and immunoinduction has many potential clinical applications. In autoimmune diseases such as lupus erythematodes, the weekly presentation of an immunosuppressive CS significantly reduced mortality in rats. In humans, conditioned nausea during radiation therapy for cancer can be successfully eliminated by a systematic desensitization and extinction procedure (Heninger, 1995). Moderate physical (pleasurable) exercise leads to a dramatic increase in lymphocytes and NK cells; the antimicrobacterial activity of macrophages is strengthened and depressive mood, which strongly exacerbates immune suppression, is reduced. Since cytokines such as interleukin-1 themselves produce changes in corticosteroid release and monoamine turnover in the brain, the immune system provides negative and positive feedback loops on emotional brain systems.

1.05.4 NEUROCHEMISTRY AND NEUROPHARMACOLOGY

1.05.4.1 Psychopharmacology and Clinical Psychology

Neurochemistry, psychopharmacology, and the scientific disciplines studying the molecular mechanisms of neurons and synapses are the most rapid growing fields in the biobehavioral sciences. The enormous interest and financial investment is primed by the factual and anticipated economic benefits of drug treatments and the relative ease and time-saving application for clinical problems. A representative overview of the rapid growth is provided by the 2000-page volume *Psychopharmacology. The fourth generation of progress*, edited by Bloom and Kupfer (1994). An introduction to basic neurochemistry can be found in Siegel, Agranuff, Albers, and Molinoff (1994).

The relationship between psychopharmacology and clinical psychology is characterized by an attitude of mutual exclusiveness on both sides. At best, biological psychiatrists or psychopharmacologists construe a psychological treatment as an adjunct to the "real" pharmacological treatment. Clinical psychologists, on the other hand, sometimes see the biologically oriented treatments as "unfair competition" to their time-consuming and often nonspecific socially oriented treatment approaches. This is a very unproductive state of affairs since the potential profit of integration could be substantial for both disciplines. If we conceptualize the brain–behavior–cognition relation as a unitarian identity principle or at least as a tightly coupled parallelism, as most neopositivistic natural scientists would prefer, some strategic consequences for the application of psychological, neurochemical, and neurophysiological modifications are obvious. Psychobiologists should, for example, obtain the legal right to prescribe some types of medication.

Psychological or behavioral treatments are often effective for "purely" physical diseases without a psychological etiology, while physical or pharmacological therapies are frequently indicated for complex psychological problems with no or less clear-cut pathophysiology. Behavioral medicine provides many examples of the physical effectiveness of behavioral treatments, while psychiatrists often treat psychological problems (i.e., marital conflict, addictions) effectively with drugs. For example, relapse after drug abuse can be prevented by drugs chemically antagonistic to the one abused (see Section 1.05.9.4, despite the conditioned origin of relapse).

The construction of psychological intervention strategies can and should be modeled after well-known and effective neurophysiological or neurochemical treatment modalities while avoiding the inescapable negative side effects of medical procedures. Along the same lines, drugs or medical strategies should be matched to specific behavioral and psychopathological mechanisms of a particular disorder.

The biological boundaries inherent in some disturbed behaviors or diseases can only be transgressed with physical approaches, while the social boundaries of disorders (e.g., the addictions) construct the limits of biological intervention.

Presently we are more distant than ever from an interdisciplinary interaction and respectful understanding and incorporation of the other discipline in one's own views. The widening gap between biological and social or psychological sciences has many reasons; since human beings have only very limited information-processing capacities, the overflow of scientific information in both fields causes an increasing protective tendency to exclude and ignore deviant and new principles. The narrower and the more rigid the individual's intellect, the more readily other positions are discarded. A still prevalent dichotomy of the mind–brain relationship, particularly in the psychological realm, undermines any fruitful interaction with the questionable argument of the infinite complexity and the intellectual inaccessibility of "subject–object" relations (cf., Lycan, 1996, for a philosophical analysis).

1.05.4.2 Mechanisms of Action of Psychoactive Substances

1.05.4.2.1 Neuronal mechanisms

Behavior and mental activity are represented in the CNS as *cell assemblies* which were formed by associative strengthening of synapses in simultaneously activated neurons (Hebb, 1949). Mental processes are coded as multidimensional *vectors* of excitatory electrical activity in strongly coupled assemblies or so-called "*synfire chains*" (Abeles, 1991). These electrical patterns of synchronous activity tend to *oscillate* in various rhythms because the plastic cell assemblies and brain regions all have recurrent networks forming a virtually endless reticulum of feedback and feedforward connections. The structural, physiological, and molecular processes which determine the electrical activity of nerve cells and neuronal assemblies constitute the biophysical basis of behavioral pharmacology, which attempts to influence the electrical vectors through modification of the neurochemical determinants at the cellular and synaptic levels.

The electrical properties of a cell depend primarily on the structure of its membranes. Those in turn are built by the proteins which are synthesized by the genetic apparatus. Transcription and translation can occur "spontaneously" as part of the intracellular metabolisms or are induced and modified at various steps by incoming external synaptic or hormonal signals.

Plasticity and behavior change therefore depend ultimately upon synaptic changes which alter themselves to dendritic and somatic extra- and intracellular responsivity. During the resting state the membrane receptors are inactivated and an imbalance of electrically charged ions is maintained so that the intracellular membrane is more negatively charged ($\approx 70\,\mu V$) relative to its external part. With the arrival of an action potential which consists of a rapidly progressing wave of depolarization, the voltage gated Ca^{2+} channels are opened and Ca^{2+} enters the synaptic terminal. Ca^{2+} prompts the binding of synaptic vesicles containing the transmitter substance onto the presynaptic membrane terminal which causes their release into the synaptic cleft. The transmitter molecules fuse (bind) with specific receptor molecules at the postsynaptic membrane if they match allosterically the shape of the transmitter. In the case of a match between receptor and transmitter, the postsynaptic membrane channels open and positively charged ions (mostly Na^+) are allowed to follow their electrochemical gradient and other forces quickly (less than milliseconds) and *depolarize* the postsynaptic membrane. Depolarization consists of a change in electrical polarization from an electrically positive charged extracellular membrane to a negatively charged potential (20–100 μ V). Only if many depolarizations at multiple synaptic sites in one cell occur simultaneously will an action potential be created at the axon hillock and information transmitted to other cells.

1.05.4.2.2 Receptors for neurochemicals

Receptors are the constituents of a cell that have the ability to recognize a drug, a hormone, or a transmitter. Membrane-bound receptors such as those for transmitters cause rapid ion flows as described above and activate the so-called *second messenger systems*. This rapid mechanism of opening membrane channels for depolarization happens within one to several milliseconds and is called *classical neurotransmission*.

Steroid hormones enter the cell membrane and bind to steroid receptors which enter the membrane of the nucleus and act directly on DNA–RNA transcription. The second messengers, among the most important Ca^{2+}, cylic adenosine monophosphate (cAMP), and inositoltriphosphate (IP_3), cause a cascade of intracellular chemical reactions which ends with the *phosphorylation* of a protein by the protein kinases. Phosphorylation consists of the binding of a phosphate molecule with a protein that activates its enzymatic properties, and only through this mechanism do proteins become functional and modify a cellular response.

Neuromodulation is slower and uses the binding of an agonist to a nonchannel receptor. Instead of involving an ion channel, these receptors are proteins that change their cytoplasmatic (intracellular) "tail" in order to bind with a so-called G-protein (a guanine nucleotide). The activated G-protein remains in an activated state for a much longer period (10–100 ms) than ion channels, again binding to adenylyl cyclase which activates the second messenger cAMP, repeating the above-described intracellular cascade. cAMP then enters the nucleus and selectively increases the expression of genes. The later phosphorylated ion channel or G-protein-coupled receptor may change the excitability of the membrane for a particular neurotransmitter.

1.05.4.3 Transmitter Systems

The description of behaviorally relevant transmitter and neuromodulator systems followed the discovery and isolation of four psychoactive drugs in the 1950s: reserpine and chlorpromazine in the treatment of schizophrenia, and iproniazide and imipramine for depression. Binding arrays of receptors and cloning of receptors finally elucidated their distribution and action in the CNS.

Three main classes of neurotransmitter systems and some new transmitters can be differentiated: amino acids, amines, and peptides. New transmitter systems whose behavioral functions are less researched include arachidonic acid, nitric oxide (NO), proto-oncogenes, and purinoceptors.

1.05.4.3.1 Amino acids

Excitatory amino acids, particularly L-glutamate, constitute the main excitatory transmitters in the CNS. There is virtually no behavioral function, including memory, learning, and higher cognitive functions, that does not involve glutamate. Three classes of glutamate receptors have been distinguished: *N*-methyl-D-aspartate (NMDA), kainic acid (*KA*), and α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptor. The functional physiological range of glutamate and glutamate receptors is extremely narrow and even slight deviations cause a host of neurological and psychological disorders through *excitotoxicity*. Excitotoxicity of glutamate seems to be involved particularly in schizophrenia and Alzheimer's disease. The three receptor types are distributed at different locations within the brain. They consist of several families of receptor subunits, each with different functional properties. The highest concentrations are found in the cortex and hippocampus, which indicates their role in excitatory cell-assembly formation for associative memory.

1.05.4.3.2 Amines

The catecholamines (adrenaline, noradrenaline, dopamine, tyrosine, and others), serotonin, and histamine are the most important CNS amines. Each of them matches several receptors at the pre- and postsynaptic membranes. Their overall quantity in the CNS is low (approximately 3–5%); their functional role in behavior, however, is enormous. Dopamine of the mesolimbic system constitutes the main transmitter system in the organization of positive reinforcement and *incentive* motivation. Its overactivity, probably caused by dysregulation of cortical and hippocampal glutamate systems, is involved in the etiology and maintenance of schizophrenia; the blockade of dopamine receptors through neuroleptics reduces active hallucinatory symptoms but at the same time causes *anhedonia* and—through reduction of dopamine in the motor nigrostriatal system—*Parkinson's disease* and tardive dyskinesia. Overactivity of both the central and peripheral noradrenaline system is related to *anxiety* states and underavailability of NA to *depression*.

The distribution of *serotonin* (5-hydroxytryptamine, 5-HT) is also widespread in the brain. Tryptamine and melatonin are chemically related hormones. Again, despite its small quantity in the CNS, the significance of serotonin for behavior, mood, and biological rhythms is significant: *eating, sleep, sexual behavior, circadian rhythm, mood*, and *aggression* are more or less regulated by the diverse 5-HT transmitter-receptor systems.

1.05.4.3.3 Opioids

Opium has been known to humans for thousands of years as an analgesic and a euphoria- and obstipation-inducing substance. However, only during the past years has its full importance and structure been elucidated. The fact that our peripheral and central nervous system contain receptors for plant opioid alkaloids was one of the most significant discoveries of pharmacology. In the CNS, opioid receptors are found as in the periphery, centrally, close to pain structures such as the periaqueductal gray, thalamus, subcortically near reflex centers, and in the hypothalamic–pituitary system. The main subtypes of *mu, delta*, and *kappa* receptors are differentially distributed in the CNS and subserve different behavioral functions: all three are involved in pain modulation, others are critical for salt–water balance, cardiovascular functions, sexual behavior, grooming, memory, and the immune system. The addictive properties of opioids are probably related to their affinity with the positive reinforcement system (see Section 1.05.9.4) and with dopamine.

1.05.4.3.4 Neuropeptides

Short-chain amino acids are called peptides and were originally believed to serve mainly hormonal-neuromodulatory roles. We know today that they can function as neuromodulators, hormones, and neurotransmitters depending on their location, molecular substructure, and receptor types. For most of them their role in behavioral functions is unknown. For some, a well-defined psychological “profile” has been developed such as ACTH and stress, oxytocin and the formation of social bonds, substance P and pain, growth hormone and development, sleep and immune function, and the above-described opioids, enkephalins, and endorphins and their role in pain modulation.

1.05.4.3.5 Acetylcholine

Acetylcholine (Ach) is not only the transmitter of the somatic motor system and the parasympathetic system but also a central nervous system agent of great importance. Its functional effects depend on pre- and postsynaptic receptors, nicotinic and muscarinic, which can be classified into several subtypes. Most Ach receptors in the CNS cause excitatory activity (depolarization) at the postsynaptic membranes.

Ach has a modulatory role in the CNS in that its fiber system lacks specific point-to-point projection but exerts a more diffuse, unspecific function on its target cells. Therefore, Ach is usually identified as the main transmitter in arousal and attention. Most neurons originate in the basal forebrain of the *nucleus basalis Meynert* which projects to limbic areas (amygdala) and the entire cerebral cortex. Many intracortical connections are cholinergic and depolarize large groups of assemblies as a threshold regulation mechanism in learning and attention (Birbaumer, Elbert, Canavan, & Rockstroh, 1990; Mesulam, 1995). The *unspecific activation systems* (ascending reticular activating system, ARAS) of the midbrain reticular formation and the thalamus and parts of the basal ganglia are all cholinergic. Learning increases Ach activity, leads to an enlargement of cortical spines on plastic neurons, and induces sprouting in dendritic growth. In Alzheimer’s disease, most basal forebrain neurons die. Electroencephalogram (EEG) desynchronization and negative slow brain potentials depend on an intact Ach system. Cholinomimetic drugs may prove useful for many psychological disorders; to date their side effects are substantial.

1.05.5 PSYCHOPHYSIOLOGY

1.05.5.1 Introduction

The interest of humankind in body–soul or body–behavior relationships probably dates back to Stone-age populations about 250 000 years ago, where trepanations of the skull were obviously performed to modify supposed head diseases. Egyptian and Greek pre-Socratic philosophy and medicine correctly attributed mental functions to the brain and emotional ones to “humors” (Hippocrates) and autonomic functions, particularly the heart. However, Aristotle’s (384–322 BC) view of the heart as the seat of nervous and mental control was maintained until early Renaissance around 1450. Only Arab and Chinese traditions provided exceptions to the strict Aristotelian–Christian rule during the early Middle Ages, particularly in southern Italy and Spain (e.g., Avicenna, whose real name was Ibn Sina).

Psychophysiology together with the rest of medicine had to wait until the mid-nineteenth century for its first experimental investigations, which were largely dependent upon the construction of electrophysiological recording devices after Galvani’s first demonstration of bioelectricity in 1791. The first animal EEG was recorded in 1875 by Richard Caton, and the first human EEG in 1924 by Hans Berger (for a history see Clarke & O’Malley, 1995).

The modern development of psychophysiology is tightly coupled to the construction and availability of the *polygraph*, where several physiological measures can be recorded simultaneously. The foundation of the Society for Psychophysiological Research in 1961, which grew from a small US-only society into a large international organization, marked the worldwide establishment of relatively cheap and universal recording devices.

1.05.5.2 Concepts in Psychophysiology

1.05.5.2.1 The three-systems approach

In order to understand human behavior, in most instances physiological, psychological (cognitive subjective), and motor-behavioral variables have to be assessed in their particular environments. Origin and consequences of this three-systems approach to human behavior as a "whole" were extensively described earlier in this chapter.

1.05.5.2.2 The law of initial values

The law of initial values (LIV) states that the higher the initial value of a particular physiological function, the smaller the increase of this physiological measure in response to a particular stimulus or psychological event. Conversely, the higher the initial, baseline value, the larger the decrease to a stimulus or event which usually produces a decrease in the particular physiological system.

The LIV was found to be roughly true for cardiovascular variables, pupil dilation, vascular variables, but not for skin conductance or skin temperature. In cortical blood flow such as measured by PET or fMRI (position emission tomography or functional magnetic reasonance imaging, see below) and with electrocortical measures, the validity of the LIV remains uncertain.

1.05.5.2.3 Activation, arousal, and directional fractionation

Every behavior, sensory experience, and physiological response varies on a continuum of intensity. However the relationships between the intensity of a particular variable and other variables are rarely linear. The term "activation" is usually applied to the more long-lasting *tonic* (hours, days, months) changes in intensity, mainly within the sleep–wakefulness continuum, while "arousal" signifies the rather short-lasting *phasic* (milleseconds to minutes) changes; the border between both is, however, not fixed. Arousal therefore "rides" on top of a particular activation level being subject to the LIV.

Early in the twentieth century, Yerkes and Dodson independently discovered an inverted U-shaped relationship between muscle tension (activation) and performance in various tasks; later called the Yerkes–Dodson Law. After discovery of the ascending reticular activation system ARAS by Lindsley, Magoun, and Moruzzi (cf., Birbaumer & Schmidt, 1996), a neurophysiological basis of the inverted U-function seems to be at hand, because many neurons in this system showed firing rates compatible with the "law." In addition, moderate electrical stimulation of the ARAS improves performance, but vigorous or rhythmic low-frequency electric stimulation has deleterious effects on behavioral output in animals (Andreassi, 1995). In sum, most studies confirmed the inverted U-shaped curve; however, for some physiological variables and for some types of performance the "law" does not apply.

In order to explain those instances, Lacey and Lacey (1970) formulated their concept of *directional fractionation* of physiological and motor responses. According to this concept, the different physiological response systems are activated or simultaneously deactivated depending upon the physiological and behavioral requirements of the particular situation. Directional fractionation is partially overlapping with another term coined by Lacey and Lacey, stimulus–response (SR) specificity.

1.05.5.2.4 Stimulus–response specificity

Different stimulus situations may call for different response requirements, causing different and situation-specific physiological adaptations. Psychophysiology provided many examples for the correctness of this statement, despite the fact that most situation-specific adaptations occur in body systems that are difficult to record from the intact human without invasive recording, such as colon movements in anticipation of food.

One of the best-studied examples of stimulus-specificity is heart rate slowing in nonthreatening sensory information *intake* and heart rate increase to potentially aversive information *rejection*, the first demonstration of this effect dating back to Lacey and Lacey (cf., Dworkin et al., 1994). The physiological mechanism for this phenomenon was elucidated and confirmed in a series of laboratory studies by Rau and Birbaumer (1993). Increase in pressure of the carotid sinus and other large arteries causes the baroreceptors to fire as a response to blood pressure increase. The afferent fibers to the brain stem end in inhibitory centers of the reticular formation, causing heart rate deceleration and short-lasting cortical inhibition. The neuronal disfacilitation is strong enough to block the impact of aversive overarousing stimulation but improves some form of low-intensity stimulus processing. Cortical slow waves, for example, are dramatically reduced by direct stimulation of carotid baroreceptors through a neck cuff stimulation device (Rau & Birbaumer, 1993). The stress- and pain-reducing impact of phasic blood pressure increase is "used" nonconsciously by genetically prone persons as a stress coping strategy; the rewarding consequences of this strategy increase blood pressure up to the point of essential hypertension.

Another example of situation- or stimulus-specificity are changes in muscle tension: electromyographic recording of facial muscles or micromovements of peripheral muscles

(Birbaumer & Öhman, 1993; Clynes, 1989) can reliably differentiate between the basic emotions (sadness, pleasure, interest, rage) even in the absence of visible changes in facial expressions. Also, imagery of emotional scenes evokes the same peripheral but partially different CNS changes (Schupp, Lutzenberger, Birbaumer, Braun, & Miltner, 1994) as actual perception of these events. These stimulus-specific physiologies interact and overlap with individual response stereotypies.

1.05.5.2.5 Individual response stereotopies

These stereotypies characterize tonic, in most instances long-lasting "preferences" of a person to react and, in the case of pathology, overreact with a preferred bodily response system in different situations. Individual response idiosyncrasies originate from genetic, constitutional, or learned causes. The concept of response specificity, as it is also called, is extremely important for the explanation of psychosomatic, psychiatric, and somatic disorders.

Apart from the development of essential hypertension, individual response stereotopies were found to be responsible for musculoskeletal *chronic pain* related to the lower back, face, and head (Birbaumer, Flor, Lutzenberger, & Elbert, 1995). While in healthy subjects individual response stereotopies are unreliable and short-lasting, they cause pathology in subjects with a high genetic risk and they are triggered under conditions of long-lasting (particularly instrumental or classical conditioning) circumstances such as intensive stress in vulnerable periods of ontogenetic development.

1.05.5.2.6 Orienting, habituation, and defense

Pavlov was the first to examine these prepared response patterns to novel, old, or potentially harmful stimuli in dogs. The principles he described for their occurrence and psychophysiological meaning are as valid today as they were 100 years ago (Pavlov, 1926). Habituation to repeated stimulus presentation is regarded as the most primitive form of learning, learning to cease responding to biologically irrelevant stimuli. In mollusks and other simple animals, the molecular basis of habituation involves the gradual reduction of intracellular cascades which eventually lead to a decreased binding of Ca^{2+} on to presynaptic terminals. The reduced release of transmitters as a consequence weakens the strength of synaptic connections between previously functionally connected cells (cf., Birbaumer & Schmidt, 1996, for a review).

In higher animals and humans, an orienting or defensive response depends on a short-term memory comparison process between actually present and stored similar stimulus configurations. In the case of a mismatch of biologically relevant stimuli, orienting occurs; in cases of potentially threatening stimuli, a slowly habituating defensive "reflex" is recorded. In the CNS, the amplitude of a positive evoked brain potential 300–500 ms after stimulus presentation, the so-called P300, is a good indicator of that mismatch process. In the periphery, heart rate decrease is mostly related to orienting, while heart rate increase is related to defensive responses. Very short-latency muscular reflexes, such as the *startle reflex*, measurable as an eye blink 40 ms after an intensive stimulus, act as an "interrupt" to threatening stimuli and are part of a whole-body escape response disposition (cf., Birbaumer & Öhman, 1993). Modulation of the startle reflex amplitude through emotionally positive or negative background material is one of the most reliable indicators of the emotional condition of the CNS state: startle inhibition through emotionally positive and startle potentiation through emotionally negative material is easy to measure and provides an excellent diagnostic tool in psychopathology (Lang, Bradley, & Cuthbert, 1990).

1.05.5.2.7 Social psychophysiology

All human behavior and subjective experience evolved out of social interaction and communication. The laboratory situation therefore creates an artificial isolation of the individual which may produce response patterns not representative of "real" life. The observation and measurement of motor and verbal behavior, cognition, and emotion simultaneously with the physiology is the most promising but methodologically difficult area of psychophysiology. During the last 20 years, a considerable amount of data has emerged which has shed light on the enormous impact of social factors on physiological response systems (Cacioppo & Petty, 1983; Hauser, 1996). Flor, Elbert, et al. (1995) recently demonstrated a dramatic effect of the presence of a solicitous spouse on the processing of painful stimuli in chronic back pain patients. The mere physical presence of the partner not only induced a more than threefold increase in pain sensitivity but brain electrical potentials to painful stimuli were increased at a very early stage (100 ms) of information processing in patients and high-risk groups. Social factors such as reinforcement history with a spouse obviously modify CNS stimulus processing, already at an early preconscious level not accessible to conscious voluntary control.

1.05.5.3 Methods in Psychophysiology

1.05.5.3.1 Electroencephalography

EEG frequency spectra, averaged evoked responses combined with computer-aided bioelectric imaging of many electrodes on the scalp, are still the most inexpensive CNS measures with an excellent time resolution. In addition, source localization of early evoked potentials allows in some cases a precise anatomical localization of the cortical origins of these potentials. Figure 1 shows localization of electrical sources of tactile evoked primary brain potentials components with a latency of 40–70 ms at the postcentral gyrus in patients with phantom limb pain due to amputation of one upper limb. The index finger of the intact arm and corner of the lip at both sides were stimulated with 1000 stimulus repetitions and EEG responses averaged in order to enhance the signal-to-noise ratio. The same stimulation was repeated after local anesthesia of the stump and shoulder.

The electrical sources of these potentials were calculated and superimposed on each subject's individual magnetic resonance brain image (see below). It can be seen that a precise anatomical localization of the finger and lip region is possible in accordance with the known homuncular organization of the somatosensory cortex. After anesthesia and pain relief, the pain-related pathological reorganization of the brain representations of the phantom limb returns to normal (Birbaumer, Lutzenberger, et al., 1997).

1.05.5.3.2 Magnetoencephalography (MEG)

The development of low-temperature superconducting interference devices (SQUIDs) allows recording of the magnetic fields of neural structures in the brain, which are in the femtotesla range (equivalent to 10^{-15} tesla), the earth's magnetic field being in the range of 10^{-4} T. Since magnetic fields cross all biological tissues not affected by electric resistance, they can be recorded undistorted above the head. MEG has an unprecedented local resolution of 2 mm up to a brain depth of about 5 cm. Its time resolution is like that of EEG—excellent and in the millisecond range. The limitations of MEG technology are the still high costs, the necessity of a mu-metal isolation chamber, and the biophysical fact that only horizontally located dipoles can be recorded from the brain. However, together with EEG imaging, MEG provides the best time and space resolution of all available noninvasive recordings of brain activity.

1.05.5.3.3 Functional magnetic resonance imaging (fMRI)

One disadvantage of EEG/MEG imaging is its low spatial resolution 5 cm or more below the cortical surface. In order to visualize local subcortical activity changes and for an anatomically correct representation of the individual brain, MRI has to be applied. While *structural* MRI was introduced 15 years ago as a noninvasive technique to visualize gray and white matter of the human brain, *functional* MRI (fMRI) is relatively recent and in the few years of its existence has become an extremely successful measure of regional changes in brain blood flow.

The basic physical principles of fMRI are not different from structural MRI, except that with the introduction of efficient power supplies, superconducting magnets, and rapid analog to digital converters, MR images can be taken in the range of tenths of milliseconds. In order to acquire MR images, a 1–3 tesla magnetic field is applied through the head. The magnetic field forces hydrogen proton nuclei into an aligned direction of the magnetic field. The protons exhibit the property of "*spin*": they rotate around their axis. In the magnetic field, the spins of the protons will all be oriented and rotating in the same direction (quantization).

Of these rotating protons in the magnetic field, more of the protons are in a lower energy state which gives the proton ensemble a net magnetization. The time to magnetization after application of the magnetic field is T_1 (for water, the main constituent of brain tissue, it is 2 s). Fat molecules magnetize much more slowly. Therefore, gray and white substances magnetize differently. The uniform field gives the exceedingly small magnetic dipoles because of their alignment a uniform magnetization and a uniform rotating frequency (Lamor frequency). The size of this combined magnetization of the spins is the magnetic resonance signal. With time the spins lose phase coherence and the signal decays, depending upon tissue properties (as T_1). This time is called T_2. At the borders of tissues, a large difference of magnetization may exist over small distances and the proton spins at those borders will be out of phase rapidly. The macroscopic field varies as a consequence of this; the macroscopic decay is T_2. In order to cause the angular orientations of the protons to align, an electromagnetic field of about 63 MHz (radio band with the Lamor frequency) is added to the magnetic field, the protons' rotation is forced away from their longitudinal axis by this

Figure 1 Cortical representation of the digits and the lower lip before and after brachial plexus blockade in a unilateral upper extremity amputee with phantom limb pain who experienced complete pain relief due to the regional anesthesia. The preanesthesia location of the representation of the fifth digit (D%) of the intact hand is indicated by the illed square, and the locations of the representations of the lip of the intact side and the amputated side by the filled squares below. The mirror images (15,18) of the intact digit and the intact lip projected onto the amputation-side hemisphere are marked by open squares. These locations refer to the preanesthesia period. The location of the representation of the amputation-side lip postanesthesia is indicated by a filled square. Note that before brachial plexus blockade, the lip in the amputated side had shifted into the region occupied by the representation of the fingers on the intact side (mirror D5). Twenty minutes after amputation stump anesthesia, the phantom pain was almost eliminated; at the same time there was a dramatic shift of the amputation-side back toward the postition occupied by the lip representation on the intact side (mirror lip, open square). The white dots are the vitamin E capsules marking the electrode positions ("Effects of regional anesthesia on phantom limb pain are mirrored in change in cortical reorganization," by N. Birbaumer, W. Lutzenberger, et al., 1997, *Journal of Neuroscience, 17,* pp. 5503–5508. Copyright 1997 by Oxford University Press. Reprinted with permission).

field; however, they all now rotate in "resonance." This field is large enough to be detected, because it induces an electric voltage at a nearby antenna. At tissue boundaries, the detected frequencies will be different and can be transformed into color contrasts.

Figure 2 shows the result of an fMRI experiment where phobic patients and healthy subjects were exposed to neutral faces and aversive odors. An increase in blood flow (because of increased net magnetization) in the amygdala (but not in the thalamus) was observed for phobics but not the healthy controls during presentation of faces. This confirms animal experiments showing that the amygdala is the critical structure for the emotional valence of fear objects and that the thalamus does not participate in the emotional "coloring" of stimuli, but only in their sensory processing.

1.05.5.3.4 Positron emission tomography (PET)

When a radioactive ligand is injected or inhaled, it is taken up by the cerebral bloodstream and binds to specific cells and regions that are metabolically active. Collision of the protons of the radioactive ligand with the electrons of the particular atom (e.g., ^{15}O) causes annihilation of the two particles and the emission of a photon. Annihilation photons carry high energy which can be detected by a ring of many radiation detectors around the head. These detectors of γ-radiation count the density of radiation and reconstruct their spatial sources in a color-coded brain map.

PET has a less precise spatial and temporal resolution than fMRI or EEG/MEG imaging. However, its capacity to localize specific sources of increased or decreased metabolic activity of specific transmitters and substances such as water, glucose, fluoride, dopamine, and GABA is indispensable for understanding the link between molecular neurochemical processes and the global systemic changes recorded with all brain imaging systems. In general, PET allows the reconstruction of brain areas with increased or decreased *blood flow* relative to a specified baseline condition. However, the target mental activity has to be present continuously for several seconds to minutes and is averaged over several trials before a clear enough signal-to-noise ratio appears. The high cost and its invasive nature, however, limits its value for psychophysiology.

1.05.5.3.5 Electromyography (EMG)

The measurement of action potentials from the motor units of the skeletal muscles is one of the oldest and most useful psychophysiological assessments. DuBois-Reymond recorded for the first time the electrical activity of muscles in 1849. Surface EMGs with electrodes at the skin above a muscle group record Ach-caused depolarization of the motor endplate shortly before the actual contraction, but the correlation of muscular effort and EMG amplitude is very good. The spike frequency from a motor unit ranges from 20 to 1000 Hz; the amplitudes from 50 to 1000 μV. The main frequencies used in psychophysiological experiments range from 70 to 1000 Hz. Today, integration techniques give a measure of total EMG output over time.

EMG recording has proven to be extremely useful for studying mental load, muscular-mental tension, and emotions, particularly facial expressions: with the use of only three facial muscles, the main basic emotions (disgust, anger, pleasure, joy, interest, sadness) can be differentiated even in the absence of visible changes in facial expression. Biofeedback of EMG from the back muscles is the most effective treatment for chronic low back pain; the same is true for trapezius or neck muscles and tension headache (cf., Birbaumer & Öhman, 1993; Flor & Birbaumer, 1993).

1.05.5.3.6 Electrodermal activity (EDA)

DuBois-Reymond also discovered in 1849 that an electric current applied between both hands changed its voltage with psychological events. Obviously, the electrical resistance between superficial tissue is modified by neuronal outflow to the skin. The eccrine sweat glands are innervated by the sympathetic branch of the autonomic nervous system (ANS) and use Ach instead of noradrenaline as a transmitter. The palms of the hands and the soles of the feet contain most of the body's eccrine sweat glands. During situations of emergency, ANS activity increases and the sweat glands are filled with salt-containing sweat which reduces the electrical resistance of the skin. It is unknown which brain regions contribute to this reaction. The debate whether skin conductance reflects more aversive emotional or "pure" arousal responses is still undecided. In experiments using emotional slides, EDA varied only on the arousal dimension and had no relationship with emotional valence (Birbaumer & Öhman, 1993).

EDA has been used in all areas of psychology, neurology, and psychiatry: its application for the detection of deception (see below) is highly controversial. As an indicator of the orienting response (OR), it shows distinct changes with psychopathology: lack of OR habituation in high-risk subjects predicts later schizophrenia, a

Figure 2 Local blood flow changes in the amygdala of healthy (left) and phobic (right) subjects during presentation of a fear conditioned stimulus (below) and an unconditioned stimulus (an odor, above). Increasing dark dots indicates flow (from Birbaumer et al., in press).

high frequency of spontaneous fluctuations is found in generalized anxiety, and nonresponder depressives and schizophrenics exhibit a more chronic course of their disease (Andreassi, 1995; Davison & Neale, 1996). The absence of EDA responses to fearful stimuli is a reliable diagnostic tool for the detection of antisocial personalities and sociopathy.

1.05.5.3.7 Heart activity and blood pressure

Heart rate (HR) and blood pressure are the most frequently used psychophysiological indicators. HR is used because it is easily recorded from any two electrodes affixed at the more right or left side of the body allowing the heart to be positioned between them (i.e., both hands). Blood pressure has become more popular during recent years when noninvasive continuous measurement was realized with pulse-wave velocity devices. The obvious physiological role of the cardiovascular system explains the high correlations of HR with all kinds of behaviors: motor, mental, perception, attention, and orienting; stress, emotion, and motivation; personality, social stimuli, brain interactions, and conditioning (cf., Andreassi, 1995, for a review). The dual sympathetic and parasympathetic nervous control allows separation of the two branches of the ANS: slowing of the heartbeat indicates the chronotopic parasympathic, speeding the increased excitability and contraction force of the sympathetic branch.

For decades psychophysiology was obsessed with discussions of the psychological significance of HR changes: while Obrist (1981) and his co-workers interpreted HR changes as variations of mobilization of blood supply for the somatic musculature (*cardiac–somatic coupling*), Lacey and Lacey (1970)proposed a close relationship with information processing in the brain: HR deceleration and negative going slow cortical potentials should accompany *information intake*, acceleration *information rejection*, and positive going slow cortical potentials. The deceleration in a signaled foreperiod reaction paradigm is caused by phasic blood pressure increase which fires the baroreceptors in the carotis sinus. This leads to parasympathetic slowing of the heart and increased cortical activity because the nucleus of the vagus located in the reticular formation fires the ARAS. Tonic stimulation of the baroreceptors through a specially designed neck cuff, however, produces heart rate slowing and positive going cortical slow waves which causes inhibition of cortical activity and information rejection (Rau, Pauli, Brody, Elbert, & Birbaumer, 1993). The accompanying stress reduction negatively reinforces the preceding blood pressure increase and on a long-term basis causes essential hypertension in genetically prone individuals. Neither Obrist's cardiac–somatic coupling nor Lacey and Lacey's information processing hypothesis was unanimously confirmed. HR slowing clearly appears in *orienting*, while HR increase after intensive or dangerous stimuli constitutes part of the *defensive reaction*.

The interaction of stress and blood pressure increase was also confirmed in different personality patterns such as type A and hostility: type A subjects (competitive, driven, impatient) with hostile attitudes toward others, demonstrate large increases in blood pressure during stressful situations and exhibit a high risk for essential hypertension and stroke. Early behavioral intervention effectively prevents this vicious circle of blood pressure increase, stress reduction through cortical inhibition, and aggressive muscular mobilization.

1.05.5.4 Applications of Psychophysiology

1.05.5.4.1 Biofeedback

The most successful and clinically relevant application of psychophysiology became the instrumental conditioning of physiological variables mostly realized through feedback of the relevant physiological variable on a computer screen; subjects are instructed to change the variable in a particular direction and are positively reinforced for successful modifications or receive immediate feedback for it, which also functions as a reward. The scientific and clinical history of biofeedback is severely hampered by exaggerated claims of the clinical efficacy and by a very early failure to replicate the pioneering animal studies of Miller (1969) and collaborators. Meanwhile, several successful clinical indications have emerged, while the majority of controlled clinical studies revealed that most of the early expressed hopes could not be confirmed. The possibility to condition autonomic and brain functions instrumentally without mediation of the skeletal muscles as postulated by Miller was clearly confirmed in studies with totally paralyzed, artificially respirated subjects who had no voluntary control of their muscles because all motor neurons had been destroyed (Birbaumer, Rockstroh, et al., 1994): after extensive training to self-regulate slow cortical potentials (SCP), a patient with advanced amyotrophic lateral sclerosis (ALS) achieved 100% control of his brain waves and could use his learned ability to communicate by affecting a computer system with his brain waves (Birbaumer, 1998).

Well-established indications for biofeedback treatment are chronic pain states, neuromuscular rehabilitation after CNS damage, epilepsy, encopresis and enuresis, Raynaud's disease, and scoliosis and kyphosis. Doubtful but frequent applications relate to hyperactivity, asthma, essential hypertension, cardiac arrhythmias, anxiety and depression, tics and stuttering, and athletic performance. No lasting changes of biofeedback treatment were reported in well-controlled studies for these conditions.

1.05.5.4.2 Biofeedback of slow cortical potentials

The most extensively researched area in biofeedback is self-regulation of slow cortical potentials. These EEG potentials are recorded with very long time constants (5 s to d.c.) and reflect depolarization (negative potentials) and disfacilitation (positivity) of the upper dendritic cortical layers (Birbaumer et al., 1990). Subjects watch their own SCPs on a computer screen for 2–10 s. The SCPs appear as a stylized rocket on a video monitor moving horizontally or/and vertically. Subjects have to move the rocket to one side of the screen. If positivity of the brain is required (or right and left hemispheric negativity), an "A" appears on the screen; with required cortical negativity, a "B" appears. If the correct movement is performed, immediate reinforcement (e.g., points that later change into money or a "goal" with a smiling face) appears on the screen. After 20–60 trials with feedback, subjects have to produce the required SCP amplitude without feedback, "out of their mind." No specific instructions are given because highly individual cognitive strategies are used for brain self-regulation.

Healthy subjects are able to control their SCPs after 2 h training from vertex electrodes. More localized SCP changes such as the production of cortical negativity in the left precentral region and simultaneous right-hemispheric positivity need several sessions of training (cf., Elbert, Rockstroh, Lutzenberger, & Birbaumer, 1984). After training, substantial effects on behavior were noted: vertex negativity improves performance on many cognitive and motor tasks, hemispherically localized negativity improves performance on the contralateral hand only, positivity leads to deteriorating performance. Several clinical applications were tested in controlled studies but only two showed lasting improvements superior to other established treatments: epilepsy and brain/computer communication in total paralysis (Birbaumer, 1998).

1.05.5.4.3 EMG biofeedback

Changes of the electromyographic activity of different body muscles are amplified and patients watch them on a screen or listen to different tones reflecting the actual amplitude of the particular muscle. Patients are instructed to reduce (or increase) muscle tension in specific situations (e.g., stress confrontation) and receive immediate feedback of their attempts. In chronic low back pain, facial pain, and tension headache, EMG biofeedback of the painful muscle region has proven to be the most effective treatment approach compared with traditional medical treatments and cognitive-behavioral treatments (Flor & Birbaumer, 1993).

EMG biofeedback is also extremely useful in neuromuscular rehabilitation after stroke and other nervous system lesions: foot drop or spastic tension is effectively removed through observation of EMG activity of extensor and flexor muscles at the paralyzed or spastic region. Insertion of EMG electrodes in the anal canal and feedback of internal and external sphincter contraction became the treatment of choice for fecal incontinence and urinary incontinence after neurological disorders such as spina bifida, partial paralysis in paraplegics, old age, and incontinence after surgery. Observation of external sphincter contraction increases sensitivity and perception of anal or bladder filling and leads to sphincter control after only a few sessions.

1.05.5.4.4 Detection of deception

Detection of deception is the most controversial application of psychophysiological methods. The issue has mainly touched the USA because in Europe the "polygraph" is banned from criminal investigation and personnel selection. The understanding of the lawmakers in Europe is that as long as the guilty knowledge test (GKT) or the control question test (CQT) have not proven 100% accuracy in detecting the innocent or the guilty, it should not be used in court. The GKT informs the suspect that several items will be named by the interrogator and the suspect has to repeat them. Only one of the items is true, and the other control items are irrelevant. For example, a robber has used a gun in robbing a bank in order to (i) to buy a house, (ii) to buy a new car, (iii) to pay the doctor's bill, etc. In the CQT several questions are asked, where only one is correct: "Did you use caliber X?," "Did you use caliber Y?," etc. Skin conductance (EDA), blood volume, and respiration are usually recorded. EDA is the most valid measure, but even in laboratory settings only 90% accuracy can be achieved. If the subjects use countermeasures such as distraction strategies, the accuracy falls to 70% at best.

The situation may change with the introduction of event-related brain potentials (ERPs, see Section 1.05.6.5) as measures to detect deception. However, so far only laboratory procedures with faked crimes and subjects playing guilty have been published. (Here, some of them reach 100% accuracy.) In a promising procedure of Farwell and Donchin (1991), three types of items are presented: probes, targets, and irrelevants. Probes are related to the crime scenario, targets could only be known by the suspect. Only larger P300 components of the ERP to targets are counted as evidence. Since ERPs are difficult to control or distort by countermeasures and a lack of P300 components in a subject is already visible in the reactions to probes, this procedure might also reach 100% accuracy in field tests.

1.05.5.4.5 Clinical psychophysiology

A volume by Carlson, Seifert, and Birbaumer called *Clinical applied psychophysiology* (1994) summarized many of the possibilities of psychophysiological measures in the clinic, workforce, and personnel applications. Any psychological, psychiatric, or neurological disorder that is characterized by deficits in information processing causes changes in EEG/MEG recordings and, in the case of emotional disorders, also modifications of autonomic variables and subcortical alterations in PET and fMRI. Psychophysiological measures alone rarely allow diagnosis of a particular disorder but have to be used in conjunction with other parameters (behavioral, cognitive, biochemical, and radiological) in order to characterize a disease state fully. In the case of psychiatric diagnosis where the definition of the disease itself is often unreliable, psychophysiology cannot increase the validity of the diagnosis further because validity depends on reliability.

1.05.6 CONSCIOUSNESS AND ATTENTION

1.05.6.1 Introduction

The problem of consciousness has been at the center of scientific and philosophical interest for millennia. Biological psychology and the neurosciences have made major discoveries about the processes that underlie consciousness and attention. This research has shown that there is no unified concept of consciousness, but rather there are several heterogeneous neuronal processes that underlie these and related functions such as selective attention and short-term memory. The common physiological characteristic of the many heterogeneous forms of consciousness is an increase in widespread arousal that is mirrored in the transition from automatic to controlled processing on a psychological level. Consciousness is closely tied to the selective activation of certain brain regions beyond a defined level of activity as well as the balance between activating and inhibiting neuronal mechanisms. Cognitive psychology has given very detailed descriptions of the various forms of attentional processes (Matlin, 1983; Näätänen, 1990); the focus of this chapter is the biological basis of attention and consciousness.

1.05.6.2 Psychophysiology of Consciousness

Patients with commissurectomies (where the two hemispheres of the brain have been separated) and split brain experiments in animals have shown that split brains lead to split consciousness (Chalmers, Crawley, & Rose, 1971; Sperry, 1964). Contents that have been learned by one hemisphere must be actively transported to the other hemisphere. Without the commissures, the left hand no longer participates in the experiences of the right hand, the visual worlds of both hemispheres are completely separate, the actions of the left side of the body can no longer be described, and verbal–syntactic tasks can only be solved if they are presented to the left hemisphere.

Controlled attention is related to changes in SCPs that develop on the cortical surface during expectancy of a relevant stimulus. In a typical experiment involving SCPs, a warning stimulus (S1) precedes a required motor response that is signaled by a second stimulus (imperative stimulus, S2) several seconds after S1. The ensuing cortical negativity shows two negative peaks: the first peak, which is especially visible in frontal brain regions, represents the expectancy and resource mobilization of the sensory input channels and stimulus processing and is thus primarily related to working memory; the second negative peak reflects primarily motor mobilization and voluntary effort directed towards the response (Birbaumer et al., 1990). The provision of resources is thus proportional to the negativity, and the consumption of resources is proportional to the positivity of the slow cortical potential. The amplitude of the SCP is usually determined by both processes.

The allocation of resources and thus the allocation of cortical activation is, however, site-specific: preactivation and lowering of cortical thresholds for responses is provided in the area where a present or future activity originates. Sudden loss of control and helplessness lead to mobilization (negativity) of resources for analysis of the new situation and related decisions. Not only the analysis of sensory input, but also the preparation for motor acts is characterized by slow cortical negativities which are low in automatic and high in voluntary, conscious movements. Voluntary movements are also preceded by (nonconscious) cortical negativities that occur about 350 ms before the consciousness of the action. Experiments on the self-regulation of these cortical negativities have shown that performance increases with higher negativity and decreases with increasing positivity, thus confirming the role of these potentials as a basis of attentional processes.

In PET studies, which allow for the analysis of both subcortical and cortical activations, the electric and magnetic data on slow potentials have been confirmed. In addition, activation in the basal ganglia and the anterior cingulate has been described during controlled processing. The cingulate gyrus and the frontal cortex seem to be especially active when executive functions of attention (such as the formation of decisions) are required, and passive attention to a stimulus does not lead to activation of the anterior cingulate (Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1991).

Studies using evoked potentials have shown that attention rather than lack of attention in a task can be differentiated in the averaged EEG trace. Obviously, the regulation of attention needs subcortical regulation mechanisms; the regulation itself occurs, however, via thalamocortical and cortical thresholds of activation. This cortical regulation of attention ensures that every stimulus is analyzed by the neocortex (even if it never reaches consciousness) which then inhibits well-known irrelevant stimuli on a cortical level.

1.05.6.3 The Ascending Reticular Activating System

The regulation of attention and consciousness is based on the activity of many widespread brain systems. Attentive behavior and preparatory planning depend on the integrated activity of multisensory and motor systems (Lindsley, Bowden, & Magoun, 1949). Whereas the widespread activation of the neocortex and the maintenance of an optimal tonic level of activation are a function of the reticular formation (RF) of the brain stem, phasic local mobilization is related to the integrated activity of the RF, the basal ganglia, the thalamus, the cingulate gyrus, the parietal region, and the frontal cortex.

The anatomical and physiological basis of waking consciousness is the mesencephalic reticular formation (MRF). This system (also called the ascending reticular activating system, ARAS) fulfills three important functions: (i) generation of wakefulness, (ii) maintenance of muscle tone, and (iii) facilitation and inhibition of the intake and conduction of sensory and motor impulses. EEG desynchronization and behavioral activation are initiated by ascending afferent fibers from the reticular formation to higher centers. The reticular formation is activated by collateral fibers from the specific ascending tracts. During phasic waking responses the impulses of the RF interfere with the thalamic pacemakers which leads to EEG desynchronization in the cortex. In addition, the reticular formation maintains the tonic waking level by constant maintenance of a diffuse arousal of the higher brain centers above a certain critical level. Sleep occurs when the RF is actively inhibited and the afferent input is reduced (see Section 1.05.7).

1.05.6.3.1 Neurophysiology of tonic and phasic activation

Tonic activation is primarily a function of the MRF whereas phasic activation related to selective attention is primarily a function of the nucleus reticularis of the thalamus in conjoint action with the MRF, the frontal and parietal cortex. The intralaminar nuclei of the thalamus can also activate the cortex without involvement of the MRF; this activation is, however, only phasic. Without the MRF the general waking level is disturbed (Steriade & McCarley, 1990). On a cellular level, the activating function of the MRF (there are also some inhibitory actions) relies on a tonic level of depolarized readiness in thalamic and cortical networks which is the precondition for phasic activation, waking, and paradoxical sleep. Thalamocortical activation is expressed by elevated firing rates in neurons with long axons, elevated excitability of the thalamocortical cells, suppression of secondary excitation and inhibition related to spindles and α-waves in the EEG, narrowing and focusing of inhibitory processes, and depolarization of apical dendrites in the neocortex.

1.05.6.3.2 Transmitter systems involved in attention

Despite many pharmacologically activating substances such as the amphetamines or caffeine, the transmitters and neuromodulators of the ARAS are not well established. Although most activating substances influence the synthesis of noradrenaline and dopamine, the catecholaminergic systems probably do not have a direct influence on the MRF-thalamocortical activating structures. The origin of the monoaminergic systems in the midbrain and their extensive branching into the superior regions turns them, however, into candidate structures for activating systems. Most attempts at an exact categorization have, however, so far led to controversial results. Electrical and chemical activation of the origins of the noradrenergic system in the nucleus coeruleus and the serotonergic system in the nucleus raphé lead to hyperpolarization and inhibition of activation in the respective neocortical projection areas. Destruction of the *noradrenergic* system has no effect on EEG activation and tonic wakefulness. Only *cholinergic* stimulation in the midbrain activates cortical regions. It seems, however, to be clear that the monoaminergic systems improve the signal-to-noise ratio in cortical regions by the inhibition of specific areas and by local increases in activation.

Dopamine is, like noradrenaline, involved in many psychological functions. Whereas dopamine is not necessary for the maintenance of tonic wakefulness, dopaminergic systems regulate the construction of goal hierarchies and selective motor attention. Amphetamine in small doses improves attentional performance primarily by an increased presynaptic dopamine release and the inhibition of dopamine re-uptake from the synaptic cleft. In high doses, this leads to motor stereotypies and psychotic behavior. The regulation of attention is especially controlled by the mesolimbic *dopamine* system. Destruction of this system by 6-hydroxydopamine causes irreversible disorganization of attentional processes: animals do not habituate, they are extremely irritable, and they do not learn well; similar disturbances are seen in attention deficit disorders. Cholinergic systems seem to have an activating influence on the cortex, whereas the direction of activation seems to be determined by monoaminergic systems and other neurotransmitters and modulators.

1.05.6.3.3 Gating of attention

The ability of neuroanatomical networks to select part of the incoming information and to exclude the rest from continued processing is called *gating*. The thalamus seems to have an important role in this gating process. Within the thalamus, the *nucleus reticularis thalami* seems to be the gate to the cortex. The nucleus reticularis thalami surrounds the thalamus like a shell and shows an internal structure that is ideal for the selection of incoming sensory activation: the cells in the nucleus reticularis thalami are characterized by widespread dendrites within the nucleus and multipolar axons with many collaterals to the specific thalamic nuclei. These long multipolar axons communicate with the rest of the thalamus and the midbrain but not with the neocortical structures. The nucleus reticularis thalami is organized in a somatotopic and a visuotopic fashion: afferents from the various regions can be differentiated on the basis of their functional significance. Depending on the origin of the afferents, only the specific gate that corresponds to the respective afferent (sensory modality) is opened by the nucleus reticularis thalami.

In addition to the nucleus reticularis, the *pulvinar*, a large nucleus in the posterior thalamus, plays an important role in visual attention. It is involved in the increase of activation that occurs in the posterior parietal cortex when attention is focused on an object. The pulvinar is also closely connected with the lateral prefrontal cortex. If visual stimuli are presented with other potentially distracting stimuli, cells in the pulvinar active and activate the striatal and extrastriatal areas of the visual cortex. In the visual cortex, a signal that has been attended to leads to an elevated evoked potential after only 60 ms. This elevated evoked potential suggests that there is a very early preconscious selection by the thalamus (Posner & Raichle, 1995). The existence of multisensory comparisons on a neocortical level before active gating occurs is based on a series of studies. For example, lesions of the right parietal lobe distort the analysis and comparison of visual and tactile and also partially acoustic stimuli. The consequence is that the contralateral side of the body is ignored as well as stimuli that are presented in that direction in spite of intact perception and motor function (neglect). Attention remains fixated on the contralateral body and face. Every shift of attention leads to an increase of blood flow in the prefrontal and inferior-parietal cortex which is further indicative of the important role of this region for multisensory comparisons. Cells in the lower parietal cortex where visual, acoustic, and tactile information enters fire only when the animal is attentive to the source of the stimulus; axons of these attention cells lead to the frontal cortex as well as the thalamus and the basal ganglia. This further underlines the role of the parietal lobe for multisensory comparison.

1.05.6.4 The Psychophysiology of Selective Attention

Two important parts of the limited capacity control system are the mesencephalic reticular formation as a provider of energy and the nucleus reticularis thalami as a gate for the distribution of activation. There is an additional "decision maker" that opens the cortical areas of the thalamic (visual or acoustic) gates before the actual conscious reduction of the threshold (increase of activation) and which gates the flow of activation to the region where it is needed for further stimulus analysis or preparation of movement. This latter function is provided by the *prefrontal cortex* and the *cingulate gyrus* which receive information about the incoming activity and the result of the nonconscious comparison process from all areas of the neighboring cortex, especially (right) inferior parietal association cortex. They also receive information from the limbic system about the motivational significance of a stimulus. The second system involves the *basal ganglia* and their connection to the nucleus reticularis thalami. This system informs the nucleus reticularis thalami about the current distribution of activation in the neocortex and prevents the continued activation of already activated areas (to a level of epileptic overactivation). This system closes the thalamic gates by increasing inhibition when activation is increased in the respective cortical module and reaches a critical level.

Efferents from the frontal cortex and the mesencephalic reticular formation converge on the reticular nucleus that distributes the thalamo-cortical activation. Stimulation of the mesencephalic formation opens the gate in a nonspecific manner, that is, the amplitudes of the event-related potentials increase, and the EEG is desynchronized. This results in a general readiness for the uptake of information and orienting. Activation of the frontal cortex closes the thalamic gates. In contrast to mesencephalic reticular formation, the afferents and efferents of the prefrontal cortex are anatomically selective in that only a portion of the connections from frontal cortex to the reticular nucleus are activated, another part is silent, and thus a gate (e.g., the medial geniculate for acoustic information) is opened; all other gates remain closed. The cingulate gyrus may activate frontal areas of working memory if difficult decisions have to be made. Stimulation of the caudate nucleus leads to inhibition (positivity, lower activation) in a prescribed cortical area. All neocortical regions project to the striatum, and stimulation of this region leads to behavioral and neuronal inhibition (Birbaumer & Schmidt, 1996). This basal system prevents an increase of activation in the cortico–thalamic feedback circuits: they increase the threshold for activation when activation of the cortical modules exceeds a critical level. The higher the neocortical activation, the stronger the neuronal influx into the basal ganglia and the stronger the closing of the gates. The frontal cortex modulates these neocortical–striatal inhibitory circuits. Disruptions of this system are always related to disturbances in consciousness and attention. Destruction of the basal ganglia leads, for example, to unconsciousness. Incomplete lesions lead to reduction of the readiness potential and disturbances of attention. If the frontal cortex is lesioned, selectivity is disturbed, the person's behavior is guided from all the stimuli that are momentarily available, and too many thalamic gates are opened.

1.05.6.5 Event-related Potentials as Indicators of Attentional Processes

Evoked cortical potentials indicate the time course (latency) and the strength (amplitude) of a certain level of information processing. There seems to be a close correspondence between these electrophysiological substrates and the psychological levels of attention (Hillyard et al., 1995). Figure 3 shows the typical form of an endogenous evoked cortical potential to attended acoustic stimuli. The more automated the attention (the more frequently the stimuli are presented without consequence), the lower the amplitude, especially of the components around 20–50 ms (positive components), of N100 and all ensuing negative components. An increase of amplitudes after 50 ms which occurs during selective attention is very site-specific. The early components are only visible in the primary projection areas. The later components (after N100), however, extend to wide areas of the cortex.

1.05.6.5.1 N1/P1

N1/P1 is first component of the endogenous evoked cortical potential that responds to manipulations of attention and is measured about 100 ms after stimulus onset. N1 is recorded over the primary sensory projection areas. The temporal N1 component, for example, is very sensitive to shifts of attention: if attention is shifted to the right ear, the left temporal amplitude increases and vice versa (see Figure 3). Lesions of the respective projection system abolish N1. N1 increases with stimulus intensity but responds primarily to changes in the temporal sequence of a given sequence of stimuli.

1.05.6.5.2 N2

N2 can be induced by the omission of a stimulus in a sequence of stimuli. It has therefore been termed *mismatch negativity*. In particular, the first part—N2a—which is

Figure 3 Schematic of a typical event-related potential for the time up to 400 ms after the presentation of a simple, but task-relevant stimulus. A dichotic listening task is used where tones (120–320 ms) are presented in rapid succession to the right (5000 Hz tone) or left (3400 Hz tone) ear (see upper portion of the figure). The crosses mark slightly deviant tones (targets), to which the subject has to respond with a key press. Below the averaged evoked potentials to the attended (continuous line) and the unattended (dotted line) stimuli are displayed. After 20–30 ms an elevated potential to the attended stimuli is present (based on Hillyard et al., 1995).

modality-specific seems to be involved in neuronal comparison processes of the incoming stimulus in comparison to the previously stored stimuli of the same modality. Changes of stimulus intensity also lead to N2. N2 is not sensitive to changes in voluntary focused attention and reflects, like N1/P1, automatic attention.

1.05.6.5.3 P3

P3 is not related to any sensory system but occurs in a nonspecific manner, whenever an expectancy has not been fulfilled. It occurs after all target stimuli that are incongruent with a previously formed attention and it can last for seconds, depending on the amount of deviation of expectancy. P3 is a correlate of a sensory process that makes comparisons between stimuli. The stimulus must have been identified as deviant from expectancy before P300 occurs. Thus P300 may reflect an extinction process of the content in short-term memory when an expectancy has to be corrected. P300 may thus be a correlate of reflectory inhibitory processes in short-term memory. Its location is primarily central at the vertex and parietal.

1.05.7 SLEEP, DREAM, CIRCADIAN RHYTHMS, AND SLEEP DISORDERS

1.05.7.1 Introduction

The rotation of the earth around its axis leads to an approximate 24-hour light and temperature rhythm in plants as well as animal organisms influences almost all physiological and psychological expressions. These circadian rhythms are based on the activity of oscillators that originate in the organism and provide well-defined oscillatory periods that are in general synchronized by *Zeitgeber* in the environment. In addition to this circadian periodicity, there are a number of endogenous oscillators with short periodicities, such as EEG, respiration, and some vegetative rhythms. Rhythms which are longer-lasting than the circadian rhythms are called infradian rhythms, the shorter ones are termed ultradian rhythms. When isolated from the *Zeitgeber* (i.e., light) of the environment, most mammals and humans continue to show a circadian rhythm. The periodicity of those freely running rhythms is, however, somewhat longer or shorter than 24 hours, in humans about 25 hours. The endogenous oscillators seem to be innate providers of rhythm. Some adaptation to external *Zeitgebers* is possible: in humans, the maximum time to synchronize is between 23 hours and 27 hours for body temperature, and for motor activity between 20 and 32 hours. Within these limits, the circadian rhythm adapts to a *Zeitgeber*. Outside this region, the rhythms are disturbed and desynchronizations between various rhythms occur (Wever, 1979).

In humans, bright light functions as the most important *Zeitgeber*. In addition, social cues are important. Circadian rhythms also determine motivational behavior. For example, the effectiveness of feedback signals and the ensuing learning and memory performance also depend on the momentary phase of circadian rhythms. The provision of water as a feedback stimulus in a thirsty organism is much more effective during the night than during daytime hours. The closeness of the coupling of endogenous oscillators and the influence of *Zeitgeber* depends on the joined physiological function of the oscillators. Every oscillator is correlated with a specific environmental cycle, that is, the oscillators must have different sensors that convey the specific *Zeitgeber* signal on the internal clock. The sensors for light–dark cycles are probably located in the retina and for eat–fast cycles there are probably sensors in the hypothalamus. True oscillators are pacemakers which measure time independent of external cues. The circadian pacemakers are located in the CNS. Peripheral pacemakers and secondary oscillators synchronize the respective organ system.

The nucleus suprachiasmaticus has been identified as the central circadian pacemaker in mammals. Lesions of the suprachiasmatic nucleii of the ventral hypothalamus lead to complete and long-lasting loss of rhythmicity. Numerous control lesions in and outside the CNS do not have any comparable effect on rhythmicity. The nucleus suprachiasmaticus forces its endogenous rhythm on other structures by the pulsating release of hormones and the rhythmic depolarization of its neurons. In primates, bilateral lesions of the SCN lead to a complete loss in activity rhythms, including the drinking rhythm without effect on the absolute amount of fluid that is ingested (no homeostatic function of circadian rhythms). Tumors in the anterior part of the third ventricle of the chiasma lead in patients to irregular sleeping and difficulty in waking. Similar changes have been observed in the sleep–waking cycle of animals after lesions without changes in the absolute time of sleep and wakefulness.

1.05.7.2 Circadian Clocks

In humans and other animals, sleep–wake cycles are the most important expressions of central pacemakers: the daily change of awareness to non-awareness or partially conscious experience dominates our life like no other bodily rhythm. Another important circadian rhythm is that of ingestion of food and fluid. The intake of food and activity cycles are highly synchronized in mammals. The rhythm is determined by anticipation of the availability of food and thus depends primarily on previous learning. Synchronization of the rhythmicity of other body functions, for example, the production of saliva before the intake of food, depends on an intact nucleus suprachiasmaticus. The anticipatory rhythm related to the search of food is, however, not dependent on the SCN but on the ventromedial nucleus of the hypothalamus. Rhythms for drinking are usually correlated with the rhythms for food intake; they can, however, also be separated from them. The drinking rhythm is strongly related to exogenous factors as SCN lesions destroy the rhythmicity of drinking. The rhythm of body temperature is similar in most mammals: after 6 p.m. the temperature reaches a maximum (Aschoff, Daan, & Groos, 1982). In a phase of inactivation it is continuously lowered and before waking up it shows an anticipatory increase. There are also specific endocrine rhythms: for example, plasma corticosteroids are synchronized with light–dark or sleep–wake rhythms. Like temperature rhythms they are difficult to influence by external factors. The rhythms of psychological variables are coupled to circadian oscillators. For example, pain perception, reaction time, vigilance, and arithmetic speed as well as immediate memory covary with circadian rhythms (Folkard & Monk, 1986).

If circadian rhythms are disturbed, such as in shift-work or during "jet-lag," lasting consequences for physiological and psychological functions ensue. Additional disturbances of circadian periodicity are related to sleep disorders, depression, and epilepsy. The main characteristic of long disturbances in rhythm is the desynchronization of usually highly correlated physiological and psychological variables or extreme synchronization of usually uncorrelated variables. Night-work and shift-work lead to long-lasting disturbances of periodicity and the connected physiological systems. Even experienced shift-workers maintain a low in performance after midnight; additional consequences are gastrointestinal disorders, sleep disorders, respiratory problems, pain, and immunological effects. Jet-lag leads to similar problems such as sleep disorders, gastrointestinal disorders, reduction of vigilance, feelings of nausea, and reduced immune function. The more time zones a flight covers, the more intense are the disorders. The change of only one hour from summer-time to winter-time and vice versa leads to measurable alterations, especially in older persons.

1.05.7.3 Sleep and Dream

Although humans spend about 30% of their time asleep, sleep became a topic of research only in the 1950s. The discovery of rapid eye movement sleep (REMS) by Aserinsky and Kleitman (1953) revealed that the brain is not passive during sleep (as was previously assumed) but that one phase of the sleep cycle (REMS) is very similar to alert wakefulness (Hobson & Stickgold, 1995). The various stages of the sleep cycle are as follows: (i) transition from wakefulness to sleep with grouped occurrences of α-waves; (ii) stage 1 of the sleep cycle characterized by the loss of α, low-amplitude, high-frequency β activity and low amplitude θ; (iii) stage 2 with low-amplitude fast activity with sleep spindles and K-complexes; (iv) stage 3 sleep with about 10–50% of δ; and (v) stage 4 sleep with δ ($>100\,\mu$V, <3 Hz) occurring for more than 50% of the time. In the REM stage, low-amplitude EEG with low θ-waves, so-called see-saw waves, occurs, otherwise the EEG is similar to an attentive waking state. K-complexes might be the correlates of a strong internal discharge of sensory systems, and spindles signal the inhibition of sensorimotor areas. From stage 1 to 4, EMG activity decreases to a complete muscular atonia during REMS.

There is considerable variability across nights and persons in the course of the sleep stages of an individual person. The first NREM stage usually lasts about an hour. The duration of the REM phases is about 104 min overall. In mammals, *slow wave sleep* (SWS: stages 3 and 4) always precedes REMS. A REM–NREM cycle lasts about 90 min, with the first cycles being shorter, and the late cycles being longer. Stage 2 covers about 50% of the entire sleep and extends in time over the course of the night, SWS is very rare in the last cycles. The duration of an individual REM phase extends from 5–10 min at the beginning to 22 min in the last phase. The approximately 90 min long REM–NREM phase has also been termed the basic rest–activity cycle (BRAC) since it may continue during wakefulness. A number of rhythms (such as eating, drinking, heart rate) seem to be synchronized to it (Jovanovic, 1971).

REMS has been called paradoxical sleep because an active brain is coupled with a more or less "paralyzed" body. Rapid eye movements occur in groups over the period of several records (up to 23 s) with variable intervals (200 ms to 23 s) between them. About 20–120 ms before EEG desynchronization during REMS, *ponto-geniculo-occipital waves* (PGOs) become active and remain visible throughout the entire REM phase. They have especially been studied in the cat and nonhuman primates, but their role in human sleep is not yet clear. They can be recorded from widespread areas in the brain including the pons, the lateral geniculate body, and the occipital cortex. The PGO waves of REMS can also be observed during the waking state and seem to be indicators of the activation of neural mechanisms indicative of orienting. They seem to originate in the dorsal pons near the midbrain and are generated by cholinergic neurons between the pons and the midbrain (McCarley, 1990).

REMS is also accompanied by brief phasic muscle activity (myocloni), whereas muscle tone ceases to exist (atonia). This atonia is related to hyperpolarization of α-motoneurons, and the myocloni are related to descending excitatory bursts from the brainstem. In REMS presynaptic inhibition of group I primary afferents and thus suppression of monosynaptic reflexes is found in addition to descending inhibition of polysynaptic reflexes. Concurrent with sensory inhibition in the periphery, there is increased sensory activity in the brain (increased neuronal activity in the specific thalamic nuclei) which seems to be relevant for perceptual stability during REMS and the development of sensory structures in the growing organism (McCormick & Bal, 1997). During phasic REMS the threshold for waking up is highest, whereas in phases of paradoxical low-voltage, high-frequency sleep without REM there is slightly improved response. The ability to respond to internal stimuli decreases continuously from stage 1 to stage 4. During REMS, activity of the sympathetic nervous system is increased. Whereas the regulatory function of the hypothalamus continues during NREMS, it is diminished during REMS. Penile erection and vaginal lubrication during REMS can be used to differentiate organic from psychological factors in erectile dysfunction and lack of vaginal lubrication.

A look at the evolutionary aspects of sleep suggests that small and short-lived animals with a relatively high metabolic rate and a short BRAC sleep longer than mammals who spend a large part of the day searching for food. The duration of sleep seems to be inversely related to the length of the sleep cycle. Large animals with high predator vulnerability seem, however, to have less REMS. In younger people, SWS may constitute up to 40% of total sleep time. Older persons sleep less, wake up more often, show shorter REM phases and generally reduced SWS. Infants show maximum REM activity in the pre- and postnatal phase. REMS thus seems to be correlated with the maturation of the CNS: the high rate of REM in babies might compensate for the lack of structured sensory input from the not yet fully developed sensory system and activates the formation of synaptic connections.

1.05.7.4 The Neurobiology of Sleep

Multiple brain systems are involved in sleep regulation. Sleep-regulating structures seem to be located in the caudal region of the midbrain down to the medulla. SWS seems to be generated by the basal forebrain, the lateral preoptic region of the hypothalamus and the nucleus tractus solitarius which inhibits the activating reticular formation. Whereas the nucleus raphé seems to be involved more in the regulation of NREMS, the nucleus coeruleus seems to be important for REMS coordination. The raphé nucleus synthesizes serotonin, the nucleus coeruleus noradrenaline, and portions of the reticular formation acetylcholine. But SWS is not exclusively dependent on the nucleus raphé, nor is REMS dependent on the nucleus coeruleus. The dorsal serotonergic raphé nuclei lower their firing rate from the waking state to SWS, and during REMS they are completely silent. Nonadrenergic synapses from the nucleus coeruleus probably mediate the phasic activation of the dorsal serotonergic raphé nuclei to external and internal stimuli which might be related to waking up during SWS and REMS. During immobilization in various sleep stages, the dorsal raphé nuclei are inhibited. Immobilization is caused by the activity of nuclei in the pontine tegmentum that activate the nucleus magnocellularis of the reticular formation which in turn inhibits the spinal motorneurons.

Cholinergic reticular structures (reticular formation, pontine and midbrain tegmentum, posterior thalamus, and hypothalamus) are hyperactive during REMS. In contrast, it is assumed that during the very similar waking state, noradrenergic and serotonergic systems are active while the cholinergic influence may be dampened. During REMS, the aminergic systems may be inhibited and the cholinergic systems are dominant. The state of consciousness depends on the reciprocal interaction of cholinergic and aminergic systems. During dreams, the visual input no longer originates in the retina but in the brain stem, which is coupled with reduced serotonergic and noradrenergic activity. The ensuing disinhibition of cholinergic neurons leads to PGO waves which create the subjective images of dreams (Hobson & Stickgold, 1995).

Endogenous sleep factors also seem to play a role in sleep generation. They include several substances such as factor S, sleep-promoting substance, 8-sleep-inducing peptide, and vasoactive peptide. Whereas some of these substances seem to induce sleep, others are produced and released during sleep.

1.05.7.5 Psychophysiology of Sleep Stages

The sleep stages are important regulatory physiological processes. Sleep during the first three hours of the night seems to have an important function for endocrine and immunological responses. Long-term deprivation leads to considerable consequences for the organism: REM and SWS episodes occur during the day, optic and acoustic illusions and hallucinations occur, and after four days paranoid ideation may be reported. The metabolic and behavioral changes are less severe. REM rebound usually occurs when selective REM deprivation is instituted; in general sleep deprivation REM is increased only after about four waking days. Barbiturates and some benzodiazepines suppress SWS and/or REM sleep (both are experienced as aversive), lead to REM rebound in dreaming or wakefulness nights, and therefore have an addictive potential. Dreams occur during REM and NREM sleep but seem to be experienced differently. Dream reports following REMS seem to be more vivid, more visual, and emotional than those experienced in NREMS where the dreams seem to be more abstract. Whereas dreams in the first half of the night seem to be more reality-oriented, those of the second half are more unusual and unreal.

Sleep seems to have an important function in the consolidation of learned material into long-term memory. REM deprivation leads to deficiencies in retrieval of learned material. This reduced consolidation seems to be related to reduced RNS synthesis when REM sleep is deprived. Hippocampal θ, which occurs during the transition from short- to long-term memory, is also present during REMS. In addition, elevated DNS concentrations in the brain have been found subsequent to learning and are also seen subsequent to REM episodes.

Borbely's (1984) theory of sleep and dream regulation suggests that two processes, S (SWS pressure) and C (REM pressure), are active at different points in time. Process S seems to depend on the duration of the waking state and increases in an exponential fashion. Process C is related to the reciprocal value of the temperature oscillator. The entire "sleep pressure" is related to the sum of S and C. This model makes exact predictions about the effects of sleep deprivation on depression, implies the existence of active sleep substances as the basis of S, and views the circadian oscillator of the nucleus suprachiasmaticus as the neuroanatomical basis of C.

1.05.7.6 Sleep Disorders

Although psychological variables are most important in most sleep disorders, biological factors contribute significantly to delayed sleep phase insomnia (DSPI), drug insomnia, sleep apnea, narcolepsy, somnambulism, and sleep disorders in depression. In DSPI persons go to bed late, but often need an hour or more to fall asleep and wake up very early. The cause of DSPI is probably a lack of flexibility of the SWS oscillator to alter the circadian period forward, that is, to induce sleep before the beginning of the circadian period. These patients can in general adapt their endogenous rhythm of 25 hours to a 24-hour cycle, but are unable to correct phase delays due to external factors. Chronotherapy has been used in these patients: based on the assumption of a *Zeitgeber* period of about 27 hours, the patients were instructed to delay sleep by three additional hours per day.

Drug insomnia is related to alterations in the sleep profile by barbiturates and benzodiazepines: REM sleep and SWS are reduced, stages 1 and 2 and β-activity are increased. Cessation of the hypnotic agents leads to REM rebound with nightmares, which are suppressed by renewed drug intake. Rebound insomnia occurs after three days of hypnotics intake with no effect on the original insomnia (Borbely, 1984).

Narcolepsy is characterized by frequent sleep attacks during the day that are related to loss of muscle tone and may cause sudden loss of posture of the person affected by them. They are characterized by the intrusion of REM episodes into the waking state. Narcoleptics enter immediately upon sleep onset into an REM phase which suggests that inhibition of REM-activating structures is reduced. Nightmares with sleep paralysis (the inability to move the limbs) are usually REM episodes with extreme motor inhibition and concurrent awakening. Sleep is distorted in several psychiatric conditions such as depression and post-traumatic stress disorder. In *major depression*, REMS often appears very early and rapid eye movements may be more frequent. This has been related to enhanced cholinergic activity and/or a deficit in noradrenergic and serotonergic activity.

1.05.8 LEARNING, PLASTICITY, AND DISORDERS OF LEARNING AND MEMORY

1.05.8.1 Introduction

The simplest form of learning is *nonassociative learning* such as *habituation* and *sensitization*. These refer to a decrease or an increase in a behavioral response dependent on the number and intensity of stimulus presentations. Associative learning such as classical and instrumental conditioning involves the pairing of two stimuli or a response and a subsequent stimulus in a time-contingent manner. *Classical conditioning* is characterized by the (usually) frequent pairing of a neutral stimulus and a biologically relevant unconditioned stimulus (US) which elicits an unconditioned response (UR). After several trials, the formerly neutral, conditioned stimulus (CS) comes to elicit the conditioned response (CR) which may be agonistic or antagonistic to the UR. Extinction occurs when the CS is presented without the US: the CR slowly ceases to be. CRs may occur towards stimuli that are similar although not identical to the CS: this phenomenon is called generalization. Modern theories of learning (Rescorla, 1988) emphasize that stimulus associations rather than stimulus–response associations are learned: the signal value of a stimulus determines the speed of acquisition. In a latent inhibition design, for example, the CS is presented several times without the US before the onset of a conditioning phase. This retards the later acquisition of a CR to the CS. Imprinting is a special form of conditioning. It denotes a sensitivity for the acquisition of certain S–S or S–R connections in a certain developmental stage.

Operant conditioning is characterized by the learning of a response which is acquired by providing positive or negative reinforcers subsequent to its execution. Although operant and classical conditioning show similarities, it is questionable whether they have a common physiological basis. For example, instrumental learning of autonomic responses is difficult to achieve, whereas classical conditioning of visceral responses is easily obtained and vice versa: classical conditioning of muscular responses needs more learning trials than operant conditioning.

1.05.8.2 Mechanisms of Plasticity

Although memory and learning are closely related, memory research has for a long time focused on the encoding of abstract materials such as numbers and words rather than on the acquisition of skills. The specificity of a memory trace seems to be determined by (i) the location of the neuronal circuit underlying the memory and (ii) the anatomical and physiological characteristics of the cell assembly and the subsequent synaptic modifications. Plastic changes in the development of the organism are based on the branching and growth of dendrites and their dendritic spines. This growth is not random but occurs along "guiding" structures with common chemical affinity: specific chemical growth gradients of the glial cells probably guide axons and dendrite growth (Gazzaniga, 1995). The strong increase in weight of the brain in the first years of life is related to an increase in the number of synapses, myelinization/enlargement of the cells, the dendrites, and dendritic spines, and an increase in glial cells and of the capillary net of the brain. Adequate sensory and motor or emotional stimulation is required for this development.

Sensory or motor deprivation lead to reduced growth and atrophy which is more severe and long lasting the more pervasive and earlier in development it occurred (Rosenzweig & Leimann, 1982). Enriched environmental conditions compared with impoverished environmental conditions have profound effects on dendritic growth and other markers of brain growth which are related to the long-term effects of behavioral learning (only active interaction with the environment causes these changes). Learning also occurs after completion of developmental processes and is probably mediated by sprouting as well as disinhibition of silent connection and changes at spines and their synaptic connections (Cotman & Nadler, 1978).

Sprouting or reactive synaptogenesis has been described as a consequence of nervous system lesions. It refers to the development of new synaptic connections and has been related to recovery of function. The activation of previously inhibited or silent synaptic connections may happen within minutes after a lesion. Recent animal work (Kaas, 1995) has shown that neuronal plasticity of the primary sensory and motor areas is present in the adult organism. Both nervous system lesions as well as behavioral training alter the architecture of sensorimotor maps in the cortex. The functional significance of these cortical changes has recently also been demonstrated in humans. In human amputees, reorganization of the primary somatosensory cortex is highly positively correlated with the magnitude of phantom limb pain. In chronic back pain patients, chronicity of the pain and cortical reorganization show a high positive association. To what extent these reorganizational changes cause these pain phenomena needs to be determined (Birbaumer et al., 1995; Flor, Elbert, et al., 1995).

1.05.8.3 The Formation of Memory Traces

Engrams refer to the electrochemical processes that underlie specific memory contents. Hebb (1949) proposed that reverberating *cell assemblies* are the basis of these memories. These reverberatory circuits are characterized by interconnected excitatory synapses that maintain the neuronal activation. A cell assembly is usually formed by several of these reverberatory circuits that must exceed a certain level of activation for a cell assembly to "fire." The activity in reverberatory circuits may be viewed as the neurophysiological basis of the consolidation of memory traces. Continuing reverberation leads to structural synaptic and cellular changes that represent long-term memory. A certain time of undisturbed reverberatory activity is needed for consolidation, otherwise interference with memory formation occurs. Weak synaptic connections between neurons increase in strength during the consolidation phase if neurons are activated in close temporal proximity or simultaneously in an "associative" manner. These temporally closely coordinated synaptic activations are the basis of conditioning and learning. Once the synaptic connection has been formed, activation of a part of the cell assembly will activate the entire cell assembly. High-frequency synchronous oscillations in the γ-band range seem to be an expression of the mechanism of formation of cell assemblies (Singer & Gray, 1995).

A fundamental principle of neuronal plasticity is based on the Hebb rule (1949): if the axon of neuron A excites neuron B and causes repeatedly or constantly the firing of neuron B (i.e., the suprathreshold excitation of neuron B), then the efficacy of neuron A in exciting neuron B increases either through a growth process or a metabolic change in one or both neurons. Realization of the Hebb rule usually depends on two presynaptic (synapses 1 and 2) and one postsynaptic cell. Assume that synapse 1 is excited by a neutral tone which by itself is not sufficient to excite the postsynaptic cell on which converge synapses 1 and 2. Synapse 2 (which may, for example, be excited by a tactile receptor in the eye) is excited together or shortly after synapse 1 by an air puff to the eye which elicits activation of the blink reflex in the postsynaptic cell. Firing of the postsynaptic cell (caused by synapse 2) now reinforces the activity of all synapses that were simultaneously active at the postsynaptic cell, thus it also increases the excitability of the weak synapse 1. After several pairings of the two stimuli, the tone alone will excite the postsynaptic cell sufficiently to cause it to fire and will thus elicit the blink reflex. Thus, classical conditioning of the blink reflex has been established.

According to Hebb's theory, the excitatory configuration must reverberate in the same cell assembly for some time before structural long-term memory changes occur. The information the CNS needs for memory storage is not the activity of one or a few cells, but the *concurrent* coherent firing pattern of an entire cell assembly. Coherence is formed by the simultaneous firing of an excitatory pattern in the cell assembly where structure and frequency are specific for a memory content. A measure of the coherence is the amplitude of the evoked potential of the EEG or the evoked field of the MEG above a certain cell assembly or the cross-correlation function of the various firing patterns between several cell assemblies. The higher the amplitude of the evoked potential, the more cells must fire synchronously, the more complex the learned materials, the larger the cell assemblies involved, and the longer the time the reverberatory activity must be active in order to attain a critical shif of metabolic activity in the cells.

Every cortical cell is connected to about 4000–10 000 other cortical cells. If there were only fixed connections for the representation of memories, then we would quickly run out of connections to store all memory traces. Since one cell can participate in the representation of many different contents, that is, participate in many different cell assemblies, we have a nearly unlimited number of combinations of functional connections that build the basis of our implicit learning capabilities. This summation of synchronous input to an assembly is necessary because the convergence of only a few synapses would not reach the critical firing threshold. If information is connected its individual elements appear in temporal synchrony and thus allow for the formation of Hebb synapses. The larger the assembly and the more complex the context, the slower the reverberatory activity in the cell assembly. Slow EEG oscillations in the frequency band of 0–15 Hz depend primarily on thalamocortical excitatory circuits and their nonspecific input from reticular subcortical structures. They determine the excitatory thresholds of cortical tissue in the waking and sleep state. Since most of our memories are cortical, cortico-cortical assemblies are necessary for their formation. Calculation of conduction velocities in cortical assemblies suggests that optimal distances (also for widely separated cell assemblies) of two synchronous depolarizations are in the vicinity of 20–60 Hz. It is just the oscillatory frequency in the γ-band range which may be detected in the EEG and MEG when *gestalts*, meanings and words or dreams are perceived. They also appear when new concepts are formed during learning.

1.05.8.4 Neuropsychology of Memory

In 1880, the Russian neurologist Korsakoff first described the *amnesic syndrome* which is characterized by anterograde amnesia, retrograde amnesia, and confabulation. The neurophysiological basis of this syndrome is a destruction of diencephalic and limbic structures and often also the mamillary bodies and the dorsomedial nucleus of the thalamus (for example in alcoholism, by lack of thiamine). The main lesion in Korsakoff's syndrome is, however, located in the hippocampus. The lesions in Korsakoff's syndrome are too diffuse to allow for exact determination of the relevance of temporal structures for memory. Patients with surgical removal of the temporal lobe (e.g., due to epilepsy, as in the famous case H.M.) make more exact examinations possible. The data from these patients suggest that the medial temporal lobe system is essential for declarative memory.

The *medial temporal lobe system* plays an important role in *declarative learning*. The hippocampus receives information from all association areas of the neocortex via the entorhinal cortex as well as from limbic areas, especially the cingulate gyrus, the orbitofrontal cortex, and various regions of the temporal cortex. All these connections are reciprocal in that the hippocampus also has efferent connections to the association cortices where the actual long-term changes of memory storage must take place. The medial-temporal lobe system must be active during the presentation and repetition of memory-related materials so that associative connections can be formed between the various stimuli that are present during encoding. The hippocampus and the adjacent entorhinal cortex connect the various representations of the entire environment that is present during learning with respect to temporal and spatial aspects. Creation of such a context is especially necessary when new situations and new learning materials have to be encoded since new perceptions and new thoughts that have so far not been associatively connected have to be associated. As soon as these contents are associatively learned, a small aspect or a part of the situation is sufficient to reproduce the entire situation. The hippocampal system thus connects the cortical representations of certain situations with each other so that they build one general memory content ("*binding*").

Whereas declarative learning is heavily dependent on an intact medial-temporal lobe system, nondeclarative learning does not need an intact medial-temporal lobe. General cortical processes are less important for the formation of nondeclarative memories although the acquisition of habits and motor skills involves the motor and prefrontal cortical areas of the cortex. Specific subcortical regions have been identified for specific nondeclarative or implicit learning processes. For example, the already described classical conditioning of the nictating membrane of the eye in the rabbit is dependent on an intact cerebellum and the conditioned emotional responses to aversive stimuli in the rat are dependent on an intact amygdala. In general, the learning process takes place in the location where the two sensory informations that are to be associated meet. For example, if the conditioned stimulus is a tone and the unconditioned stimulus an aversive tactile stimulus, the associative connection in the rat is formed in the medial section of the medial geniculate nucleus where both channels of information converge.

1.05.8.5 Cellular Mechanisms of Learning

The cellular basis of learning has been primarily studied in simple organisms such as the sea snail *Aplysia* whose nervous system consists only of 20 000 neurons, and the common fruit fly *Drosophila melanogasta*. Sensitization, habituation, and operant and respondent (Pavlovian) conditioning have been described.

Associative learning in complex organisms may follow similar principles as learning in simple organisms such as *Aplysia*. Since the 1970s it has been shown that habituation, sensitization, and classical conditioning are all related to specific pre- and postsynaptic modifications in neuronal systems with few cellular connections. In *Aplysia*, stimulation of the siphon or the mantle shelf leads to contraction of the siphon, the mantle shelf, and the gill (unconditioned defensive reflex). About 10 stimulations of the siphon lead to habituation of the defensive reflex. This response reduction is related to the reduction of transmitter release in the sensory neuron. Every new action potential reduces the Ca^{2+} influx into the sensory synapse. Long-term habituation (weeks, months) is related to a reduction in the number of vesicles that release transmitter in the synapse. When sensitization occurs, increased transmitter release of sensory interneurons at the motor neuron occurs (Hawkins & Kandel, 1984).

Classical conditioning in Aplysia entails the following processes:

(i) Stimulation of the tail activates a group of facilitator neurons that lead to transmitter release at the synapses of the sensory neuron (presynaptic facilitation).

(ii) The transmitter (serotonin) activates the enzyme adenylyl cyclase which increases cyclic AMP in the synapses of the sensory neurons.

(iii) Increase of cAMP activates a second enzyme, a cAMP-dependent protein kinase.

(iv) The protein kinase closes a specific type of K^+ channel in the membrane and reduces the number of open K^+ channels during the action potential. The protein kinase also activates the synaptic vesicles.

(v) The decrease of the K^+ outflow leads to a broadening, that is, a temporal extension of the next incoming action potential, which causes more Ca^{2+} to enter the synaptic terminal and more transmitter release since the binding of the transmitter to Ca^{2+} is a prerequisite of its release.

(vi) The increased Ca^{2+} binds partially to calmodulin which attaches to the adenylyl cyclase and increases the potential for cAMP production.

In the cortex and hippocampus where glutamate is the excitatory transmitter, NMDA receptors most likely initiate these processes (Linden & Connor, 1995). In hippocampal and cortical dendrites, *long-term potentiation* (LTP) has been described as a potential physiological mechanism of the encoding and retrieval of information in short-term memory. *Homosynaptic LTP* occurs when a cell is stimulated for about 1 s at a 100 Hz frequency. The cell increases its firing frequency and also its synaptic strength when a test stimulus is later presented. The incoming stimuli lead to depolarization and Ca^{2+} influx into the cell because the ensuing Mg^{2+} blockade opens the NMDA (or AMPA) receptor.

Associative LTP occurs when slow stimulation (5 Hz) of an axon is coupled with high frequency activation of an adjacent axon. The connection of the slow-frequency stimulated axon synapse with the postsynaptic membrane leads to a long-lasting increase in excitation and a strengthening of the synaptic connection upon renewed stimulation. *Long-term depression* (LTD) protects the synapse from extreme (epileptoform) LTPs and facilitates the reduction of LTP, thus increasing the possibility for renewed excitation. In its *heterosynaptic form*, one synapse is stimulated tetanically while the adjacent synapse is not preactivated. *Associative LTD* occurs during asynchronous (anti-Hebb) stimulation of two synapses. Low-frequency stimulation (1 Hz for 10 min) leads to *homosynaptic LTD*.

LTD is probably the mechanism responsible for forgetting and affects lateral inhibition in LTP. Whereas LTP leads to increased intracellular Ca^{2+}, the activation of protein kinase (G-protein), and phosphorylization of proteins in the cell and subsequent gene expression with the synthesis of new proteins, LTD leads to dephosphorylization of the receptors and reduced conductivity.

1.05.8.6 Neuronal Bases of Learning and Memory

Simple classical conditioning does not require an intact cortex. Thompson and Krupa (1994) showed the development of an engram in the hippocampus and the nucleus interpositus of the cerebellum during classical eyeblink conditioning in the rat, rabbit, and cat. In the rabbit, closure of the nictating membrane of the eye was used as UR, a tone as CS, and an air puff to the eye as US. Cells in the CA1 layer of the hippocampus become active in the CS–US interval several trials prior to the expression of the CR nictating membrane closure to the tone. The mean firing frequency of these cells exactly predicts the form of the later CR. A similar firing frequency was found in cells of the nucleus interpositus of the cerebellum from which efferents of the cerebellum to other brain regions originate. If these cells are destroyed, the CR is abolished and a new CR (but not the UR) can no longer be acquired. The learning mechanism is the previously described LTD of Purkinje cells in the cerebellar cortex. The tone CS is transported via the auditory tract to the cerebellar cortex and the nucleus interpositus, and the air puff US reaches these structures about 100–300 ms later via the nucleus trigeminus and the olive. The motor UR and CR are relayed from the nucleus interpositus (which is inhibited from the cerebellar cortex and has an excitatory connection to it) to the motor nuclei of the brainstem. Learning occurs in the plastic Purkinje cells: LTD is initiated by the shortly temporally spaced firing of the parallel mossy fibers (CS) and the climbing fibers (US) at the synapses of the parallel fibers of the Purkinje cells which have AMPA receptors. Thus, cerebellar cortical LTD disinhibits the nucleus interpositus (which is usually inhibited by the cerebellar cortex) and the CS can elicit the CR. This engram is thus stored in the cerebellar cortex, not the afferent or efferent structures.

In humans, destruction of the cerebellar cortex likewise leads to disruption or elimination of classical eye lid conditioning, with intact responses to the CS and US (Daum et al., 1993). The cerebellum is, however, only involved in motor learning. The learning of emotional responses or declarative learning are unaffected by cerebellar lesions.

1.05.8.7 Neurochemistry of Memory

The biosynthesis of proteins seems to play an important role in transferring information from short-term into long-term memory. It is unlikely that qualitative changes in the amino acid sequences of the RNS are the underlying mechanism for long-term storage, since these changes only seem to reflect unspecific activation or stress (Davis & Roitblat, 1984). Alterations in gene expression may be viewed as the basis of learning-induced physiological changes.

Learning and other environmental influences modify the binding of proteins that activate transcription of the genetic code via regulatory proteins. Every gene has three regions: a code region which is transcribed by mRNS and translated into a specific protein, and a regulatory region which consists of the promoter region, and an enhancer region which contains binding locations for regulatory proteins which ensure tissue-specific gene expression for every specific type of cell. The promoter region consists of a brief sequence of nucleotides which must first be activated by a regulatory protein before RNS-polymerase can transcribe the structural gene. Phosphorylation of the regulatory proteins is essential for their function as a key in the lock of the regulatory region.

Various antibiotics inhibit cerebral protein biosynthesis during the translation of tRNS into the respective amino acids at the ribosome. About 80–90% of cerebral protein synthesis can be blocked temporarily without severe deficits in behaviors other than memory. Encoding and the input of information is not disturbed if the training time does not overlap too much with the time it takes for antibiotics to exert their effect. Even weeks after the completion of training, the retrieval of memory contents is interfered with. The strongest amnesia occurs when protein synthesis is inhibited shortly before the beginning of training, that is, if protein synthesis is disturbed during training. The retrieval of memory contents is not influenced by protein biosynthesis inhibition because injections of the respective antibiotics at the time of retrieval does not show an effect on learned materials. This means that the protein biosynthesis is only needed for a critical phase of consolidation during and after training. Interestingly, short-term memory is not affected by the inhibition of protein biosynthesis.

Both long-term potentiation and long-term depression lead to a modification of early and late gene expression. Excitation as well as damage to nerve axons (e.g., after deafferentation during amputation) leads to long-lasting LTP or LTD (based on the cell type). Subsequent to injury, hyperexcitability, synaptic facilitation, and growth occur in the deafferented neurons. These may occur within hours and explain the quick and lasting chronic pain that may occur subsequent to amputation. Although this is not an associative learning mechanism, the molecular changes are probably similar to those of associative learning. There is a critical time period during which the content of a memory can be moved from short-term to long-term memory. This critical time period is determined by the duration of the change in gene expression. After LTP, the so-called immediate early genes are altered and subsequently late genes are activated that lead to a permanent synaptic modification by transport of the newly synthesized gene products from the nucleus to the cell membrane. This could be the structural basis of long-term memory.

Except for protein biosynthesis, no single substance or class of substances has been shown to have a causal influence on memory. This suggests that memory can be coded in a number of synaptic and cellular processes that all lead to the same final result: long-term changes in the firing rate of a cell assembly. Whereas central catecholamines seem to have a minor role in learning and memory, peripheral catecholamines seem to play an important role in the consolidation of memories. For example, lesions of the medulla of the adrenal gland lead to a severe reduction of peripheral noradrenaline and concurrently to severe amnesias. The activation of β-adrenergic receptors on the central or peripheral level by emotional stimuli seems to have a strong influence on retrieval from long-term memory. If the β-adrenergic receptors are blocked (e.g., by propanolol), retrieval of emotional but not neutral events from memory is severely disturbed at a later point in time (Cahill, Prins, Weber, & McGaugh, 1994).

Acetylcholine seems to play an important role in learning. There are several cholinergic systems in the mammalian brain and precursors of acetylcholine and metabolites participate in almost all behaviors directly (as transmitters) or indirectly (as neuromodulators). It is therefore especially difficult to show the role of cholinergic systems in associative learning. Scopolamine, which blocks most cholinergic receptors, leads to memory loss for birth incidents if it is given to mothers during delivery. It has sometimes been reported that postsynaptic cholinergic stimulators improve retrieval from memory in patients with Alzheimer's disease. In patients with morbus Alzheimer's disease where neuronal aging processes seem to occur at a very rapid rate, reduced acetylcholine is found, along with a lower density of muscarine receptors and a lower number of cholinergic neurons. The low efficacy of cholinergic stimulation alone suggests that cholinergic and noradrenergic systems have to converge in layers 1 and 2 of the cortex in order to make normal memory and consolidation possible, although they may not be involved in the encoding and storage process. In Alzheimer patients, degeneration of nucleus coeruleus cells has been observed. In animal experiments long-lasting stimulation of the nucleus coeruleus in aging mice led to an inhibition of the forgetting of a shock avoidance response, which would otherwise occur because of the aging process. These results suggest that memory processes depend on the interplay of several transmitters in specific brain regions, primarily in the cortex, hippocampus, and limbic regions.

Just like the neurotransmitters, centrally as well as peripherally injected peptides modulate memory functions. They all seem to have an indirect influence; none of the peptides that have been examined so far are necessary preconditions for learning and memory processes: vasopressin, ACTH, and somatostatin enhance learning and encoding, endorphines and oxytocine have negative influences. Angiotensin, cholecystokinin, and substance P interfere with encoding if they are directly injected into the central nervous system.

Opioid peptides and low doses of morphine-containing substances lead to forgetfulness; naloxone and opioid antagonists improve encoding in various tasks. This memory-enhancing effect may be related to the inhibition of catecholaminergic receptors by the opioids and the blockade of the facilitatory effect of noradrenaline, dopamine, and acetylcholine. Naloxone also prevents retrograde amnesia related to electric shock which is accompanied by a massive output of beta-endorphine and methionin-enkephaline.

The high concentration of glutamate in the hippocampus and neocortex suggests that this amino acid has an important role in consolidation processes. It is not clear if the increase of glutamate receptors and the correlating growth of dendritic spines following long-term potentiation of the hippocampus plays a causal and specific role for memory processes. The use of glutamate to increase learning and memory has so far not shown consistent results. Although Alzheimer's disease is related to a marked reduction of CNS glutamate (60%), treatment efforts that influence glutamate metabolism have so far not shown consistent results.

1.05.9 MOTIVATION AND MOTIVATION-RELATED DISORDERS

1.05.9.1 Introduction

The term motivation is used to describe an *internal state of the organism* that modulates the frequency and intensity of behavioral responses. Motivational states may be based on *innate* drives or may be induced by *acquired* drives. Some drives follow *homeostatic* principles: they are less dependent on environmental influences and the learning history of the individual but are rather elicited by deviations from the internal homeostasis of the body (e.g., hunger, thirst). They show stable levels of attainment that motivate a stereotypical sequence of behaviors if they are exceeded or not attained. Non-homeostatic drives (such as sexuality, exploration) show variable levels of attainment and deprivation times that are highly determined by learning processes and other environmental variables (such as availability, incentive). Behavior is determined by *drives* as well as *reinforcement*. Whereas drives provide the "energy" for a certain behavior, stimuli that increase the probability of the occurrence of a response are termed reinforcers, and stimuli that reduce the probability of the occurrence of a response are termed punishing stimuli. Reinforcement is the increase in response probability in the presence of a discriminative situation by the delivery of reinforcing stimuli subsequent to a response. Drives and reinforcements are based on different neuronal substrates.

In addition to drive reduction, incentive motivation plays an important role in the instigation of goal-directed behavior: after some reinforcing trials not only the drive but also the reinforcement will move the behavior in a specific direction. *Incentive motivation* is determined by (i) the neuronal substrate for reinforcement that is activated by behavioral consequences, (ii) the emotion related to reinforcement which is associated with reinforcement-related stimuli via classical conditioning, and (iii) incentive salience, objects and behaviors that were associated with reinforcement become more prominent ("salient") than other stimuli, capture more attention, and motivate approach or avoidance behavior.

1.05.9.2 Hunger and the Eating Disorders

Hunger is one of the homeostatic drives. Under physiological conditions, a reduction in blood glucose concentration serves as signal for hunger. The glucosensors are located in hypothalamus, brainstem, and liver. Upon the ingestion of food, hunger ceases fairly rapidly (*short-term presorptive satiety*), whereas the ingestion of food is terminated only when the blood glucose levels have again reached their predetermined level (*long-term resorptive satiety*). Classical conditioning usually plays an important role in motivating food intake. Social and environmental stimuli such as dinner time, taste and appearance of food, or persons present at dinner determine the time and the amount of food ingestion much more than physiological factors. Taste and odor cues serve an important function in the selection and ingestion of food (e.g., taste cues increase appetite despite complete reduction of hunger (Carlson, 1991; Legg & Booth, 1984). In addition to *glucostatic* factors, *thermostatic* (food intake proportional to changes in temperature) and *lipostatic* mechanisms (food intake proportional to the metabolism of lipoproteins) have been discussed (Carlson, 1991).

Presorptive satiety is determined by the activation of odor, taste, and mechanoreceptors in the otolaryngeal cavity and trachea as well as by chewing. Additional factors are pressure-sensitive receptors in the stomach and the colon as well as chemoreceptive vagal afferents from the stomach and upper colon that respond to the glucose and amino acid content of food. Resorptive satiety is determined by chemoreceptors of the digestive system as well as all enteroreceptive sensory processes related to hunger feelings. The enhanced glucose availability, enhanced heat production as well as changes in lipid metabolism, activate central receptors that signal satiety. Gastrointestinal hormones are also important in the generation of long-term satiety. For example, cholecystokinin, a neuropeptide, leads to satiety probably mediated via cholecystokinin receptors in the ventromedial hypothalamus. Based on the glucostatic model of hunger and satiety, the activity of the lateral hypothalamus (LH) is viewed as a signal for hunger and the ventromedial hypothalamus (VMH) is viewed as important for satiety. Both centers are assumed to be reciprocally inhibitory. The lateral hypothalamus seems to serve the function of a glucostat and inducer of hunger. An increase in glucose levels inhibits LH activity, activates VMH, and initiates satiety. Destruction of the VMH leads to overeating and obesity. The VMH and its connections to the paraventricular nuclei of the hypothalamus and the efferent vagal nuclei as well as to the nucleus tractus solitarius seem to code taste as nutritionally adaptive and maladaptive. The limbic system and associated cortical structures as well as the motor system are also involved in the planning of food intake, probably via catecholaminergic fiber systems that connect the brain stem, cerebellum, basal ganglia, and cortex. The search for food is probably instigated by neurons in the lateral hypothalamus that are connected to motor regions by dopaminergic fiber bundles.

Eating disorders such as anorexia or bulimia are primarily caused by cultural and psychological variables (see Volume 6). Biological variables are, however, important in the course and maintenance of these disorders since dieting is followed by serious biological sequelae. Ongoing dieting leads to major disruptions of the endocrine system, especially the pituitary–adreno–cortical axis. The regulation of sexual and reproductive functions is disturbed. Reversible reductions of brain mass have also occasionally been reported which are associated with negative consequences such as psychological disorders and ongoing weight problems by 30% of patients. Biological and hereditary factors of metabolic rate seem to be much more important in obesity than in other eating disorders. However, dieting also has negative consequences in the case of obesity where the lost weight is regained and maintained at a higher than prediet level (Carlson, 1991), referred to as "cycling."

1.05.9.3 Sexual Function and the Sexual Disorders

1.05.9.3.1 The sexual response

Four phases of sexual activity can be differentiated: sexual attraction, appetitive behavior, copulatory behavior, and postcopulatory behavior. They have to be synchronized during sexual activity between the involved partners and they are based on clearly differentiable neuronal and hormonal mechanisms.

Sexual attraction and all other phases of sexual behavior are influenced by the *androgen* levels of the male and the *estrogen* levels of the female in most species. Odors, posture, and color changes contribute to sexual attraction. Appetitive behavior involves "invitation" to approach and to mount, changes in posture, erection, and the emission of sounds. Appetitive behavior is disorganized when the cortex is removed, whereas copulatory behavior (intromission, orgasm) does not depend on the neocortex. Orgasm involves ejaculation (male) and contractions of the pelvic muscles and vagina (female) or the penis and the pelvic musculature (male).

In humans, the sexual responses of both sexes are fairly similar. Women usually need a longer appetitive (plateau) phase to attain orgasm and most men have an absolute postcopulatory (refractory) period: subsequent erection and ejaculation is only possible after a certain recreational phase.

The sexual response of the male is initiated by erection which is caused by dilatation of arteries to and in the corpora cavernosa, the corpus spongiosum urethrae, and the sinusoids of the corpora cavernosa via the activation of postganglionic parasympathetic neurons. These neurons are activated by afferents from the penis and surrounding tissue and, in addition, by supraspinal (cortical) structures via psychological mechanisms. In addition to vasodilation, the venous reflex from the corpora cavernosa is reduced by venous constriction; both mechanisms lead to vasocongestion. The sacral spinal cord is the locus of the erectile reflex. Strong activation of afferents from the sexual organs leads to the activation of sympathetic afferents in the lower thoracic and lumbar spine. This sympathetic activation then provokes the release of semen and fluid into the internal urethra. Following this emission, ejaculation is initiated by afferent activation from the prostate and the internal urethra in the pelvic nerves. Both sympathetic and parasympathetic neurons to the sexual organs show maximum activation during ejaculation. The sympathetic afferents cause tonic–clonic contractions of the pelvic musculature and erectile tissue, causing fluids and semen to be expelled via the external urethra.

The female sexual response is characterized by a venous blood occlusion and vascular dilation of the labia majora and minora. Glans and corpus clitoris swell and increase in size and length due to parasympathetically induced vasocongestion. Just as in the male, both afferents from the genital region (especially the clitoris) and supraspinal inputs lead to these changes. Within 6–30 s after stimulation (afferent or supraspinal), lubrication of the vagina ensues in sexually mature females and formation of an "orgiastic cuff" in the outer third of the vagina. The "orgiastic cuff" contracts during orgasm and this activity can be compared to emission and ejaculation in the male. The uterus changes to an erect and enlarged position and contracts during orgasm.

Some persons with complete spinal cord lesions show erection, lubrication, and orgasm when the sexual organs are stimulated but all sexual sensation from the genital region is absent. These patients report, however, "phantom sensations" that are independent of peripheral stimulation.

1.05.9.3.2 Sexual differentiation and sex hormones

The development of a male or female organism is determined by the male reproductive cells in sperms: if they contain a Y chromosome the organism will be male, if they contain an X chromosome, the organism will be female (all ova have X chromosomes). Up to the eighth week of pregnancy the organism is dimorphic; only after this period the sex hormones determine if the organism will have male or female sex organs. In humans, weeks 12–22 of pregnancy and the first six weeks after birth are periods that are sensitive for androgens. During prenatal development, the androgens also act on the CNS and create the gender-specific differences in the hypothalamus and the limbic system, among others. This androgen action on the brain also seems to determine later sexual preferences (hetero-homo-bisexualism).

If an XY chromosome is present, precursors of the testes form in the seventh and eighth week which produce androgens (e.g., testosterone) that are essential for the formation of a male organism. If androgens are lacking, a female organism will develop. Secondary gender characteristics (such as a beard in the male, or breasts in the female) develop in puberty when the hypothalamus begins to produce precursors of the gonadotropic hormones that excite the release of estrogens and testosterone via luteinizing and follicle stimulating hormones of the pituitary gland. These secondary changes can be partially altered in later life if sex hormones are withdrawn or added. *Hermaphroditism* denotes a disorder where the internal or external sexual organs are malformed and consequently no clear gender assignment can be made.

1.05.9.3.3 Hormonal basis of sexual behavior

Regulation of the sex hormones depends on the activity of hypothalamic cells that excrete peptides into the local circulation and through neural pathways or via the pituitary gland into the systemic circulation. These peptides either act as hormones and bind to receptors often very distant from the hypothalamus or they act like transmitters at the synaptic cleft where they exert a tonic modulating influence on neuronal excitability. All sex hormones are regulated by luteinizing hormone releasing hormone (LHRL), also called gonadotropin releasing hormone (GnRH), which is excreted by several hypothalamic and extrahypothalamic cell systems. LHRH stimulates the excretion of luteinizing hormone (LH) and follicle stimulating hormone (FSH) from the anterior pituitary. Whereas dopamine and serotonin have an inhibitory influence, noradrenaline has an excitatory influence on the LHRL cells in hypothalamus. Estradiol and progesterone reaching the CNS via the blood system increase and testosterone decreases its release.

LHRH is bound to the cells that release LH, FSH, and prolactin. Only *rhythmic pulsative* LHRH activity will lead to hormone release (every 3–4 h except for the first phase of the menstrual cycle, where women release hormones every 90 m). LH and FSH stimulate the growth and transformation of the follicle n the ovaries and the production of estrogen; in the male they maintain spermatogenesis and the production of testosterone. In the male organism, they are released at a constant level: testosterone inhibits LHRH (negative feedback). Environmental stimuli (e.g., anticipation of sexual activity) influence FSH and LH release.

In the female organism, LH and FSH are released in a cyclic pattern: increase of estrogen leads to increase of LH and FSH release in the first phase of the menstrual cycle. When the follicle grows in the first days of the cycle, estrogen and progesterone slowly increase, on day 12 on average LH production steeply increases (because of a positive feedback circle of estradiol on LH and FSH). This burst causes ovulation (about 24 h later). The follicle then releases the ovum and the remaining follicle cells increase progesterone synthesis. The follicle turns into the corpus luteum and releases more progesterone and estradiol. This positive feedback circle is followed by negative feedback: the increase of progesterone leads to LH and FSH inhibition, estrogen and progesterone release is terminated around day 24 when LH and FSH release decrease further. Menstruation follows when the mucosa of the uterus loses vitality owing to progesterone reduction.

The influence of the sex hormones on behavior is less powerful in humans than in other mammals. Appetitive and copulatory behavior is maintained after ovarectomy and castration if prior learning of sexual behaviors has taken place. Castration after puberty has only delayed effects on sexual activity (slow decline) and is reduced if prior sexual experience is present. Contraceptives that suppress the rhythm of hypothalamic, pituitary, and ovarial hormones do not influence appetitive and copulatory behavior, whereas testosterone levels have a stronger influence on sexual appetite in the female than estrogens.

Amenorrhea (cessation of menstruation) is often related to psychological factors that suppress LH increase but may also be related to physical factors (e.g., lack of nourishment in anorexia) or diseases such as tumors of the pituitary. Intake of LHRH suppresses LH and FSH release and induces reversible castration in the male (e.g., for treatment of tumors of the prostate or in sexual delinquency). In the male, the normal level of available testosterone is usually unrelated to sexual behavior. However, if a certain level of testosterone ($\geqslant$ 350 ng/l) is lacking, impotence may occur and may be corrected by substituting testosterone. Helplessness and depression reduce, and anticipation of sexual activity increases, testosterone production.

Pheromones—substances that are secreted by an organism and influence the other organism's behavior via the sense of smell—have important consequences on partner selection, sexual behavior, and pregnancy in rodents. Their role in human sexual behavior is less clear.

1.05.9.3.4 Neuronal mechanisms of sexual behavior

Most research on neuronal mechanisms of sexual behavior is based on research in rodents. The applicability of these findings to human sexual behavior is questionable. Within the same organism two different neuronal networks for male and female sexual behavior exist: this *sexual dimorphism* seems to be a universal characteristic of all organisms. The apparent behavior of the organism results from the inhibition and/or excitation of the respective networks. Thus both sexes are "represented" in the brain (Swaab & Hofman, 1995).

In both sexes, the hypothalamus is the central regulating structure controlling sexual behavior in concert with sensory and motor reflexes in the genitals, the autonomic extraspinal ganglia and fibers, and spinal reflexes. Cortex and limbic system exert a modulating influence, primarily on the hypothalamus. The hypothalamus functions as both a neuronal control center and secretory organ. The limbic system, the hypothalamus, and their most important afferents and efferents are all structures with a high content of sex hormones. This system has also been referred to as the *paracrine heart of the neuraxis* since it has a central function in the regulation of homeostatic and nonhomeostatic drives, excitability of moto neurons, the fight–flight response, reproduction, aggression, and territoriality. The anterior hypothalamus has been identified as the region that integrates the various sex-related reflexes into a goal-oriented behavior pattern.

In the male, cutaneous mechanoreceptors of the penis have an important role in the induction of copulatory behavior. If both destruction of olfaction and social isolation are present, male sexual behavior is suppressed in the rat. In the monkey, the medial preoptic region is essential for copulatory behavior. Additional important modulators of male sexual behavior are the frontal and somatosensory cortex and the basolateral nucleus of the amygdala. The so-called *Klüver–Bucy syndrome* in the monkey (tameness and hypersexuality) which occurs after removal of the temporal lobe and the amygdala is probably related to the loss of inhibitory input of the amygdala on the medial preoptic region of the hypothalamus.

Female sexual behavior is likewise coordinated by hypothalamic structures in concert with lower reflexes: the medial preoptic region of the hypothalamus inhibits and the ventromedial nucleus facilitates the lordosis reflex in the female rat. The hypothalamic efferents and the somatosensory afferents converge in the central gray of the midbrain and the dorsolateral reticular formation. The basic hormonal level (primarily estrogens) determines to a large extent the strength of the hypothalamic influence. It thus has a *priming function* on the entire chain of behavior, whereas the individual motor and vegetative elements depend on spinal reflexes.

1.05.9.3.5 Influence of sex hormones on the brain

Androgen- and estrogen-sensitive cell systems have been identified in the hypothalamus and the limbic system. Their growth depends on the amount of peripherally available hormone levels. Axons and dendrites of these cells are modified by androgens and this leads to anatomical differences of these structures in men and women. This might be related to the often reported advantages of verbal vs. spatial abilities in the female vs. male sex and the increased plasticity with respect to verbal defects in the female. Creative musical talent has been associated with profiles of sex hormones that resemble those of the opposite sex. Despite these prenatal sex hormone influences on gender differences, cultural and social factors are paramount.

The interaction of hormonal influences and sexual behavior in humans is best studied by the analysis of disorders in the endocrine glands that occur pre- or postnatally. The *androgenital syndrome* is based on excessive release of male hormones in the fetal stage of development. As a consequence, the external female organs become male (penis but not testes develops) even if the organism is chromosomally female, based on the sex chromosomes, and has internal female organs. These genetically female children have been raised as males and have developed attraction to females in the postpubertal phase because of brain masculinization. In some cases (especially if the external genitals were not well formed) these women were treated postnatally to reduce virilization and their external genitals were surgically corrected; their later sexual orientation remained, however, "male": 48% were reported to be bisexual and 17% homosexual (compared with 2–5% among females in general). Prenatal masculinization of the brain is obviously sufficient to induce lifelong changes in sexual orientation.

In the *androgen-insensitivity syndrome*, a genetically determined insensitivity of the androgen receptors at the androgen effector cells, leads to external female sexual organs in a genetic male. These children are usually raised as females and reportedly show average heterosexual behaviors (i.e., they prefer males as sex partners). In *male hermaphroditism* a lack of androgen production in a genetic male leads to undefined external sexual organs. These children are, therefore, often raised as males or females, independent of their genetic gender. If raised as girls, these genetic males accept men as sexual partners.

In *5-α-reductase-deficit disorder*, a genetic defect leads to lack of the enzyme 5-α-reductase which transforms testosterone to dihydrotestosterone, which is responsible for development of the external male sexual characteristics. This defect does not, however, influence brain development. Thus, a "male brain" is found in a "female body" in a chromosomally male organism. In the first generation, the children were raised as female, when it became clear that masculinization set in during puberty, the next generation of children was raised as males with less behavior problems than the first generation "macho" and often lesbian females (Money, 1987). Thus, brain development seems to have a more substantial influence on sexual orientation than environmental influences and reinforcements of sexual behavior.

1.05.9.3.6 Brain processes and homosexuality

Homosexual behavior seems to be primarily dependent on hormonal changes in the brain. This statement refers to primary (exclusive) homosexuality which seems to have a prevalence of about 5% in men and 2–3% in females across cultures.

In animals (e.g., rats), sexual preference has been modified in the pre- and postnatal phase by altering the action of male or female hormones. For example, extreme stress of the mother often leads to androgen suppression in the male fetus (Bailey, Willerman, & Parks, 1991). Although these rats develop external male sexual organs, they show female copulatory behavior. If female rats are prenatally exposed to androgens, the reverse behaviors occur. A genetic component also seems to be relevant for homosexuality: the X chromosome of some homosexual males seems to have specific markers for homosexuality in the 928 region (Le Vay et al., 1995). This assumption is substantiated by the observation that children who are raised by homosexual couples do not show an increased prevalence of homosexuality, thus rendering the role of environmental factors unlikely. Defense reactions of the maternal immune system against androgens of the male fetus have also been implicated in the development of "female brains" in male bodies.

The common final path of these influences seems to be the hypothalamus. Male and female brains show the most pronounced differences in this structure. It is, however, not clear if these alterations are really at the basis of differences in sexual orientation. Dimorphic regions of the hypothalamus are not fully developed by the postnatal period; rather, the sexually dimorphic nucleus of the preoptic region can only be differentiated between the sexes at age four and there are substantial anatomic and neurochemical alterations in this structure in puberty and in later age. It is therefore likely that some gender differences in the brain may not relate to different brain development but to differential cell death of brain regions in the sexes. The vasopressin-containing subnucleus of the nucleus suprachiasmaticus was found to have twice the size in homosexual as compared with heterosexual males, whereas no reduction (demasculinization) of the sexually dimorphic nucleus of the preoptic region was observed. In transsexuals, similar deviations have been observed (LeVay, 1996).

1.05.9.4 Acquired Motivation and Substance-induced Disorders

The substance-induced disorders with their accompanying addictive behaviors are a "model" for acquired motivation which yielded important information about the neuronal mechanisms of drives and incentives. Although a genetic risk has been proposed for the addictions, they reflect primarily learned behavior patterns that are determined by psychological and biological factors. Although psychosocial factors may play the most important role in the development of the addictions (see Volume 6), their maintenance is strongly determined by learning and central nervous system variables.

The positive and negative reinforcing properties of substances and the time course of their intake are important determinants of addiction. Whereas previous definitions of substance-induced disorders focused mainly on the phenomenon of tolerance (reduction of the efficacy of a drug over time related to pharmacokinetic (metabolic), pharmacodynamic (receptor-related) mechanisms, and withdrawal symptoms), current formulations (American Psychiatric Association [*DSM-IV*], 1994) focus more on compulsive drug seeking and craving as determinants of drug abuse (Robinson & Berridge, 1993).

1.05.9.4.1 The neurobiology of addiction

The psychobiological basis of drug abuse and drug dependence has been well established. Important—although not necessary—characteristics of dependence are the development of tolerance and withdrawal symptoms with repeated drug usage. Tolerance refers to the decrease in efficacy of a drug with repeated administration and has been related to both pharmacokinetic and pharmacodynamic processes. Pharmacokinetic tolerance is based on the increased availability and efficacy of drug-metabolizing enzymes that lead to enhanced absorption of the drug into the blood stream. Pharmacokinetic tolerance is related to changes at the receptor level, for example, dopamine receptors of limbic cells. Both tolerance and the development of withdrawal symptoms are related to neuroadaptive processes whose molecular bases have been described to some extent (Di Chiara, 1995; Self & Nestler, 1995). A structure of special relevance for the development of addictive behaviors is the mesolimbic dopamine system. This system seems to be the final common path of the positive reinforcement system of the brain that was discovered by Olds and Milner (1954) in their experiments on intracranial self-stimulation.

1.05.9.4.2 Learned motivation and addiction

In 1954, Olds and Milner first showed the phenomenon of intracranial self-stimulation (ICSS) (Olds & Milner, 1954). The implantation of electrodes in the septum led rats to self-stimulate at a very high rate up to complete exhaustion. Stimulation of the lower structures in the midbrain (the periventricular system) had an opposite effect: the animals attempted to prevent any kind of electrical stimulation. Olds therefore called those centers pleasure or punishment centers. Persistent ICSS can be elicited from many subcortical and cortical regions: the optimal regions in the rat are the descending medial forebrain bundle, the lateral hypothalamus, and the frontal cortex. Preferred sites of self-stimulation are those fibers of the medial forebrain bundle that lead from the lateral hypothalamus to the posterior and ventral tegmentum. Those sites coincide to some extent with the dopamine system. The dopaminergic fibers from which intracranial self-stimulation can be elicited are mainly ascending in the direction of the forebrain. This includes the part of the medial forebrain bundle that leads to the nucleus accumbens as well as the ventral tegmental bundle, the posterior medial forebrain bundle, and the capsula interna, which have connections to the caudate nucleus, the putamen, and the amygdala. Thus the dopamine system seems to be the common final path of the descending medial forebrain bundle–lateral hypothalamus system and of endogenous opiate cells. Neuroleptics block this final common path of the positive incentive system and lead to anhedonic behavior by reducing incentive salience of reinforcing aspects of the environment.

In humans, rather than complete anhedonia, reduction and blunting of positive affect and reduced incentive motivation seem to be the consequence of the intake of neuroleptics. Whereas dopaminergic neurons seem to be related to the energizing incentive motivational aspects of positively motivated behavior, the endogenous opiates seem to be important for the positive affective tone of reinforcing stimuli. This effect of the endogenous opiate system is probably primarily related to the inhibition of noxious inputs. Opiate receptors are primarily located close to the pain-processing centers in the central nervous system. They modulate pain perception in the dorsal horn of the spinal cord over the periaquaductal gray (where descending pain inhibition can be elicited) to the thalamic relay stations for painful stimuli in the amygdala, the temporal cortex, as well as the striatum, where they end in ascending dopaminergic neurons. The high addictive potential of the opiates is probably based on this aversion-reducing effect.

The development of tolerance and the increase of incentive motivation are based on intracellular sensitization and neuroadaptation. It is assumed that the increase of incentive motivation and incentive salience after the intake of drugs is related to a sensitization process of the mesolimbic dopamine system and that tolerance and development of dependence are related to an intracellular process of neuroadaptation. The positive reinforcing value seems to be elicited by the binding of dopamine to D_2-dopamine receptors or opiate receptors (especially μ- and δ-receptors). They activate G-proteins that lead to reduced adenylyl cyclase activity, reduced cAMP, and reduced cAMP protein kinase. G-proteins directly regulate K^+ and Ca^{2+} channels. The reduced cAMP-dependent protein kinase activity leads to reduced phosphorylization of proteins that regulate neuroexcitability such as ion channels, enzymes, and receptors. The infusion of substances that activate protein kinase block positive reinforcing effects (Sp-cAMPS); inhibition of protein kinase increases reinforcement by blocking phosphorylization of the proteins that are necessary for excitation of the nerve.

In the course of chronic drug intake, the cAMP system is adjusted at a higher level in a compensatory fashion in order to compensate for cAMP reduction by the drugs and to guarantee normal excitability of the cells. This leads to systematic changes in gene expression of the proteins that are important for excitability and could explain the long-term effects of incentive motivation, while cAMP reduction is an immediate effect of the direct (with ICSS or amphetamin) or indirect (e.g. with food) positive reinforcement. It is assumed that persons (or animals) who have a genetically higher risk for the abuse of substances show a molecular structure of the dopaminergic cells which is similar to that of persons who are already dependent. In those persons, externally ingested substances have a weaker effect and have to be ingested at a higher level. Tyroxine hydroxilase (TH) synthesizes dopamine (DA) and binds to dopamine receptors that are coupled to G-proteins and neurofilaments (NF). The neurofilaments transport gene products from the axon to the synapse. The cAMP-dependent protein kinases (PKA) modify the ion channels and stimulate the transcription factors (e.g., CREB, c-fos). The nucleus accumbens (NAc) that projects back to the VTA also has connections to the basal ganglia, especially the ventral pallidum (VP), and the hippocampus (HP) and receives projections from the amygdala (AMYG), the olfactory and gustatory system (Olf), and the cortex. After repeated intake of cocaine or morphine in an animal, a dopaminergic cell is extremely easily conditionable to drugs: now the VTA cells and its axon have shrunk, less dopamine is synthesized and released. Adenylyl cyclase and cAMP show a compensatory elevation and PKA leads to excessive transcription and channel activation. The dopaminergic system becomes overreactive and this leads to abstinence symptoms (Robinson & Berridge, 1993). In general, the increase of positive reinforcement (euphoria) and incentive salience after repeated drug intake and the activity of the dopaminergic system are parallel processes (American Psychiatric Association, 1994). The courses of wanting and liking in the development of an addiction are not necessarily parallel. Whereas drug wanting may increase with repeated drug intake, the euphoria and positive effects of the drug may slowly decrease. Figure 4 shows an integration of the most important systems and processes for the development of liking and wanting, based on the theory of Robinson and Berridge).

The contiguous pairing of CS and US leads to their associative connection (primarily in cortex); the strength of the associative connection also depends on comparison with similar stored contents. The direction of the attention towards a certain stimulus is guided by this memory process. Activation of the dopaminergic incentive system and the positive reinforcement system is also determined by current drive states. Drugs can stimulate each of the two systems, the positive reinforcement system and the incentive system directly and independently, which means that wanting and liking can be separated. Drugs do have joint effects (cross-tolerance), specific psychophysiological effects, that also determine the course of the dependency. The term cross-tolerance denotes the fact that the adaptation to certain substances (e.g., morphine) also leads to tolerance towards another substances (e.g., alcohol), although one does not have ever to have used this drug.

Some of the endogenous opiate systems are in a close functional relationship with the dopaminergic positive reinforcement structures. The euphorizing (reinforcing) effect of exogenous and endogenous opiates seems to be related to the sensitivity and increased excitability of positive reinforcing structures rather than to tolerance and withdrawal symptoms. The identification of neurochemical mechanisms for reinforcement, tolerance, and withdrawal is complicated by the fact that we have at least three endogenous opiate systems with at least four types of receptor classes. Each of these three systems and receptors could be related to different behaviors that occur during addiction. It is, however, clear that cells show compensatory intracellular processes when they bind their receptors with externally added opiates. Withdrawal leads to a sudden loss of these compensatory effects and counter-regulation leads to an overshoot of the compensation. In alcoholism, the unconditioned dependence

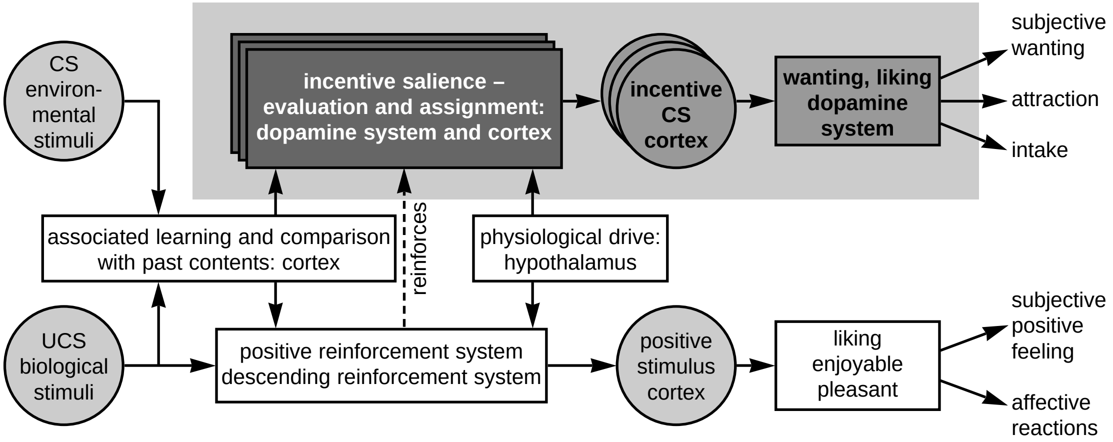

CS environ-mental stimuli
incentive salience – evaluation and assignment: dopamine system and cortex
incentive CS cortex
wanting, liking dopamine system
subjective wanting
attraction
intake
associated learning and comparison with past contents: cortex
reinforces
physiological drive: hypothalamus
UCS biological stimuli
positive reinforcement system descending reinforcement system
positive stimulus cortex
liking enjoyable pleasant
subjective positive feeling
affective reactions

Figure 4 Comprehensive model of the development of addiction. Bottom: development of pleasure; top: development of craving. *Biologische Psychologie*, by N. Birbaumer and R. F. Schmidt, 1996, New York: Springer. Copyright 1996 by Springer. Reprinted with permission.

(neuroadaptation) seems to be at least partially mediated by opiate receptor binding: during oxidation of alcohol in the liver by the enzymes alcohol dehydrogenase and aldehyde hydrogenase, aldehyde metabolites reach the central nervous system where they create condensation products together with central monoamines which form alkaloids that have a morphine-like structure. These products (tetrahydro-isoquinolone, TIQ) act as "false" transmitters and bind to the opiate receptors. This distorts the opiate neuroadaptation that has been described above.

Nicotine seems to stimulate nicotinic cholinergic receptors at low levels; at higher levels the cholinergic receptors are blocked and relaxation occurs. Nicotine also activates the production of catecholamines and serotonin in the brain which probably leads to increased attentional levels. The blockade of nicotinic cholinergic receptors by mecamylamine reduces smoking, but leads, however, to very significant side effects. These neurobiological principles should be taken into account when addictions are treated.

1.05.10 EMOTION AND THE EMOTIONAL DISORDERS

1.05.10.1 Introduction

Emotions are behavioral responses to either aversive or appetitive stimuli that are characterized by responses on the verbal-subjective, the motor-behavioral, and the physiological level. *Emotions* vary along the two basic dimensions of valence (positive–negative) and arousal (arousing vs. calm). Several primary emotions (fear, anger, surprise, disgust, happiness–joy, sadness) have been identified that seem to be cross-culturally invariant (Ekman & Davidson, 1984). *Moods* are longer-lasting response tendencies towards certain emotional states. Emotional states or feelings are usually short-lived, rarely exceed more than a few seconds, and are characterized by specific muscular responses of the facial muscles with a high communicative value as well as specific approach–avoidance patterns. Moods, by contrast, tend to influence cognitive processes rather than elicit motor responses.

Based on the work of Lange, William James proposed that the conscious experience of an emotion is dependent on the physiological changes we perceive from the periphery of our body. Thus the James–Lange theory of emotion proposes that we "feel sorry because we cry, angry because we strike, afraid because we tremble" (James, 1890). In contrast, the Cannon–Bard theory (Cannon, 1929) suggested that emotions were not characterized by a specific peripheral physiological response pattern and could therefore not be detected based solely on the peripheral response. This theory based the source of the emotion in the brain and considered it as a result of the perception and interpretation of emotion-inducing events in the higher nervous system. This latter view is correct: the stimulation of specific brain regions leads to the experience of specific primary emotions. It is, however, likely that peripheral input also contributes to the experience of an emotion. The work of Ekman and Lang and their colleagues suggests that very specific peripheral physiological patterns are related to the primary emotions (Ekman, 1984). Related to this theoretical controversy is the question, to what extent does cognition precede emotion or to what extent emotions are primary and are only followed by cognitions? The great emphasis on cognitive processes in clinical psychological models of emotional disorders in the 1970s and 1980s made this question very pressing (Lazarus, Kanner, & Folkman, 1980; Zajonc, 1980). Schachter and Singer (1962) showed in a number of experiments that emotional states that were artificially induced (e.g., by adrenaline) and could not be explained by the subjects were readily attributed to the situational cues that were presented to the subjects. These results were taken as evidence that cognitive variables (attributions) are primary and the specific peripheral activation pattern is secondary in the interpretation of an emotion. There is, however, evidence that emotions may develop irrespective of conscious evaluations of the situation. For example, subliminal perception experiments or experiments with masked stimuli suggest that emotions may be evoked by these stimuli although they are not consciously perceived (LeDoux, 1995; Öhman & Birbaumer, 1993). Aversion to odors or tastes can also be conditioned in an anesthesized state.

1.05.10.2 Fear and Anxiety and the Anxiety Disorders

1.05.10.2.1 Avoidance, fear, and anxiety

The learning of anxiety and fear responses has been well described in psychology textbooks (Birbaumer & Öhman, 1993) and will not be reviewed here. In accordance with modern theories of learning (Rescorla, 1977), it is important to note that fear conditioning is stimulus, not response, conditioning: the CS acquires the ability to drive brain systems that activate a defense response that is well established in cerebral networks. Animal experiments have shown that the cortex is not needed in the learning of a conditioned emotional response: obviously the thalamo–amygdaloid connection is sufficient to establish a conditioned fear response. Destruction of the amygdala will eliminate the conditioned emotional response (Mackintosh, 1974).

The cortex is, however, needed for differential conditioning where higher order interpretation of the stimulus is important and for the extinction of a conditioned response. Both the primary sensory and the medial prefrontal cortex, which is closely connected to the amygdala, are needed for extinction. The amygdala thus plays an important role in the acquisition of fear. The instrumental maintenance of fear seems, however, to be mediated by the septo-hippocampal system which regulates the expectancy of safety signals and selective motor avoidance responses. The hippocampus is necessary for the interpretation of contextual cues, whereas the lateral nucleus of the amygdala governs the conditioned response. If the central nucleus is destroyed, motor, autonomic, and endocrine CRs are abolished.

Lesions of the septum and hippocampus lead to reduced exploratory behavior in novel situations, reduce orienting to expected (but not unexpected) aversive stimuli, facilitate avoidance and escape behavior, and disturb passive avoidance. Both peripheral sympathetic and central noradrenergic systems act in concert when intense stressors have to be dealt with. Noradrenaline seems to have a nonspecific activating function on the neocortex and increases the signal-to-noise ratio, thus leading to increased attention to intensive stimuli (Cameron, 1994). The noradrenergic fibers are organized into a dorsal bundle (periventricular, originating in the locus coeruleus), which runs to the cortex and hippocampus, and a ventral bundle (tegmental, not originating in locus coeruleus) which runs to the amygdala and the septum. Destruction of the dorsal fibers does not affect learning but selectively prolongs extinction in animals. In this case the filtering of irrelevant materials is disturbed and passive avoidance behavior is maintained. In humans, the findings on regulatory dysfunctions of the noradrenergic system are much more controversial. Instead of a general hyper-reactivity of adrenergic systems, a specific dysregulation seems to be present that is specific for the various anxiety disorders. For example, in *post-traumatic stress disorder*, marked increases in peripheral autonomic responses (e.g., heart rate, startle response) have been found in general, and especially during confrontation with the traumatic event (Lang, Bradley, & Cuthbert, 1993). In *panic disorder*, clonidine, a noradrenergic agonist, leads to a paradoxical inhibition of the release of growth hormone (GH) from the pituitary gland, whereas in healthy controls noradrenaline leads to an increase in GH release. It is assumed that long-lasting NA release has led to excessive levels of NA in the hypothalamus, which have reduced the sensitivity or number (downregulation) of postsynaptic α_2-adrenoreceptors. The subjective improvement of the anxiety attack that follows clonidine is in accordance with the hypothesis of reduced sensitivity of central receptors.

1.05.10.2.2 Startle reflex modulation and fear

The startle reflex is a fast, protective, reflex response of the muscular system to loud noise or other intensive surprising stimuli. In humans, activation of the muscularis orbicularis oculi can be measured by electromyographic recordings about 30–50 ms after a 95–110 dB loud noise was presented. The amplitude of the startle response is influenced by the emotional valence of stimuli that are presented concurrently with the startling noise: fear increases, positive emotions decrease the response (Lang et al., 1993). This affective modulation of the startle response occurs independently of attention and activation but is related to the response disposition of the organism: if the organism is geared toward approach, the reflex is inhibited, avoidance and escape facilitate it since motor responses need to be interrupted when new stimuli are presented which require immediate action. Reflexes that have an approach character such as the closure reflex of the hand or the foot show reverse modulation (Davis, Hitchcock, & Rosen, 1991; Öhman & Birbaumer, 1993).

The brain structures and neurotransmitters involved in modulation of the startle response have been well studied in the rat. Whereas the amygdala is essential for the potentiation of fear-induced startle response in rats, the nucleus accumbens, an incentive and reinforcement structure, shows an inhibitory influence on fear-potentiated startle (Koch & Bubser, 1994). The medial prefrontal and the orbitofrontal cortex primarily inhibit fear during extinction and mediate delay of reinforcement based on expectancies. The medial septum and the hippocampus are activated during stimulus discrimination and passive avoidance of fear stimuli.

Measurement of the startle response has proven to be an important instrument in differentiation of the emotional valence of emotional stimuli. The *valence* of the conditioned stimuli (positive, negative, neutral) determines the potentiation of the startle response, whereas the *arousal* of a stimulus determines changes in skin conductance level (Lang et al., 1993).

1.05.10.2.3 Psychopharmacology of fear and anxiety

Benzodiazepines such as diazepam (Valium) are effective inhibitors of anxiety and bind to GABA (gamma amino butyric acid) receptors prevalent in the limbic system. Benzodiazepines inhibit anxiety by binding to the benzodiazepine receptor which regulates (together with the $\mathrm{GABA_a}$ receptor) $\mathrm{Cl^-}$ influx into the cell. The $\mathrm{Cl^-}$ influx into the cell is enhanced when the benzodiazepine receptor is activated and this increases the hyperpolarization of the cell. This explains why benzodiazepines are not only generally relaxing and soothing but why they have anti-epileptic effects. Since the benzodiazepines do not have a specific effect on anxiety-related neuronal networks, they cannot effectively reduce anxiety and sleep disorders on a long-term basis, although they do have beneficial short-term effects. In addition, they show a substantial addictive potential. Barbiturates, benzodiazepines, and alcohol seem to inhibit the behavioral inhibition system selectively and may thus selectively influence conditioned but not unconditioned fear responses of the passive but not of the active avoidance type. They do not influence approach behaviors. This is in accordance with the finding that phobias respond well to benzodiazepines, whereas obsessive-compulsive disorders respond badly to them. The necessity of cortical involvement in the extinction of fear suggests that pharmacological reduction of fear may not be indicated when long-term reduction of fear and anxiety, for example by behavioral treatments such as confrontation, is desired.

1.05.10.3 Sadness, Depression, and the Affective Disorders

1.05.10.3.1 Sadness and depression

Sadness is a basic negative feeling that accompanies separation and the loss of attachment that is interculturally relatively invariant. *Depression* is a complex mixture of feelings that contains sadness, but also disgust, anger, rage, hostility, fear, guilt, and shame. Thus, depression is always related to a multitude of psychological, social, and biological factors. The focus of this section will be on the biological preconditions of depression.

Sadness and depression are clearly distinct from other feelings not only in subjective experience but also in electromyographic and electrodermal responses. In addition, the expression of different feelings is different in healthy and depressed persons. In electroencephalographic recordings, the induction of negative or depressive moods has been found to correlate with a right frontal activation, but PET studies suggest that the left amygdala and the left prefrontal cortex show increased metabolism. This increase in blood flow in fronto-limbic areas could result from the lack of activity in the positive reinforcing dopaminergic mesocortical connections that are usually inhibitory and could have caused overactivity in the frontal cortex and the amygdala. If one assumes that the left hemisphere and the amygdala show more dopaminergic activity during depressed emotions and more activation, this could lead to an increased inhibition on the left side and an elevated activity in the electrocortical measures in the right frontal lobe (Davidson, 1993).

The psychological factors that are important in depression and bipolar disorders are reviewed in Volume 6. As in other chronic disorders, genetic predisposition may play an important role especially for bipolar affective disorders. Twin studies as well as adoption studies suggest increased rates of concordance in biological relatives of depressed patients. Studies of circadian rhythms in depressed patients suggest that the endogenous clock may be accelerated. Days and weeks before a depressive episode, desynchronizations in temperature periodicity are found. The circadian temperature curve is flattened or completely irregular. The release of growth hormone is reduced. The amount of REM sleep increases, the amount of slow wave sleep decreases, the entire sleep period becomes shorter, frequent awakening, especially in the morning, occurs. REM latency is shortened, probably due to a low slow wave sleep pressure (see Section 1.05.7). Sleep deprivation tends to improve depression the following day.

1.05.10.3.2 Neurochemistry of depression

The effectiveness of antidepressant medications has been based on their ability to increase the availability of noradrenaline and/or serotonin in the synaptic cleft. The monoamine theory of depression has, however, not been substantiated: substances that do not directly affect the noradrenergic and serotonergic system are also effective in treating depression (e.g., lithium). Post mortem analyses of depressive persons show no changes in noradrenaline or serotonin content, the precursor of noradrenaline synthesis, the enzyme dopamine-β-hydroxylase is unchanged, and the metabolites of central noradrenaline (MHPG) and serotonin (5-HIAA) are also not reduced in depressed patients. Thus, either mechanisms other than the increase of the aminergic transmitters in the synaptic cleft must be assumed or neurotransmitters other than noradrenaline and serotonin cause changes in these substances as a secondary effect.

Antidepressives lead to a loss of noradrenergic β_2 and α_2 receptors. The clinical effect of improvement in depression seems to be associated with the speed of reduction of receptor binding and not with the increased availability of noradrenaline and serotonin. Depressions could, therefore, also be caused by too much availability of central noradrenaline and only the destruction of postsynaptic noradrenergic receptors by the drugs (this lasts usually days or weeks) improves the depression. On the other hand, reduction of the number of receptors is associated with a compensatory increase of activity of the remaining β-receptors.

Tricyclic antidepressants lead to a reduction of the synthesis and the metabolism of noradrenaline and serotonin by increasing their availability in the synaptic cleft. The noradrenergic cells fire at a lower level and MHPG in blood and urine is lower. During mania, however, a clear increase of noradrenergic activity has been found that is reduced by lithium. In depressed patients, but also in obsessive-compulsive patients, a lack of sensitivity of the presynaptic α_2 autoreceptors seems to be present. These suppress noradrenaline release when noradrenergic cells are active. When the autoreceptors are blocked, more transmitter is being released. If the receptors are underactive, less noradrenaline is released. The reduction and subsensitivity of the α_2 autoreceptors could, however, also be the consequence of increased noradrenaline and serotonin activity in depression.

Persistent stress seems to increase central as well as peripheral noradrenaline. Long-lasting stress and helplessness lead to a central loss of noradrenaline (Peterson, Maier, & Seligman, 1993). This would coincide with the monoamine theory of depression, but not, however, with an overabundance of noradrenaline. Some of these controversies could be resolved if one considers the dynamic course of action of antidepressants and the dynamic course of coping with stress. In depression and after uncontrollable stress, the activity of cells with β-adrenergic receptors in the brains of animals seems to be low. If demands are addressed to the individual that need energy-mobilizing coping behavior, networks that are connected to noradrenergic systems cannot sufficiently respond (learned helplessness). In contrast, successful coping with stress (immunization) or treatment with antidepressants leads to increased activity of postsynaptic cells with noradrenergic receptors on their postsynaptic membranes. Since β-adrenergic receptors are widely distributed in the brain and can coexist with other receptors on the same neuron, the chronic intake of antidepressants or successful stress management can increase the reactivity of serotonergic and dopaminergic cell systems. The reduced number of β-adrenergic receptors after antidepressant treatment or after stress management with a concurrent increase in the activity of β-adrenergic cells is probably related to a desensitization process of the β-adrenergic membrane. An initially increased noradrenergic activity due to stress or antidepressant intake leads to reduction in the number of β-receptors. The receptors that remain, however, increase their efficacy manifold, which improves the information processing and distribution of these cells.

1.05.10.4 Aggression and Aggression-related Disorders

1.05.10.4.1 Development of aggression

In animals, several subtypes of aggression such as between-male, between-female, fear-induced, maternal, and sexual aggression have been identified. All subclasses of aggression can be reduced to two environmental events: (i) the presentation of aversive, painful stimuli and (ii) the withdrawal of positive stimuli (frustration). If aggression or fear/flight are elicited by certain stimulus configurations, this depends on evaluation of the stimulus configuration; neuronal structures that are involved in these emotions are overlapping.

Aggressive behavior is not a homeostatic drive but primarily a learned behavior whose probability of occurrence is also influenced by constitutional and hormonal factors. The heredity of aggressive behaviors of male mice, for example, is about 0.3–0.5 (with a maximum of 1) based on selective breeding experiments. The variance related to environmental factors varies therefore between 50 and 70%. The genetic transmission for aggressive behavior in humans is not known: the concordance rates of mono- and dicygotic twins seems not to be different, which suggests that the hereditary component is not important. The idea of localization of an aggression gene on the male Y chromosome has not been supported, nor has the relationship between crime and the existence of an XYY gene configuration in men been substantiated (Rose, 1995).

1.05.10.4.2 Neuronal basis of aggression

The localization of aggressive behavior in one or a few brain structures is not possible owing to the heterogeneity of aggressive behavior within and between the species. Some important brain areas for aggression have, however, been identified. Stimulation of the amygdala usually leads to fearful behavior in both humans and animals, but frequently also to aggression. The lateral and medial hypothalamus seem to be integrating structures for aggressive behavior. Stimulation of the lateral hypothalamus leads to prey aggression and stimulation of the medial hypothalamus to affective aggression. Stimulation of the dorsal hypothalamus leads to flight or fear aggression. Experimental lesions in the cortico-medial part of the amygdala can lead to extremely aggressive attacks on living and inanimate objects. It is assumed that the amygdalae regulate the hypothalamic aggression structure via the stria terminalis. The so-called *Klüver–Bucy syndrome* where the destruction of the anterior temporal lobe leads to tameness and hypersexuality in the rhesus monkey seems to be related to destruction of the amygdala and loss of the temporal lobe.

The integrity of normal social interaction in primates and humans seems to be dependent on intact structures and connections of the amygdala, temporal lobe, and prefrontal cortex. These three structures are anatomically connected by extra- and intrathalamic connections. The anterior cingulate gyrus seems to play a central role. The amygdala, the posterior orbitofrontal cortex, and the anterior cingulate gyrus receive input from all higher sensory cortex areas and project themselves again to the higher motor and premotor areas. Their connections with the reinforcing structures therefore predestine them for the analysis and adequate response to the social meaning of a situation.

The anterior cingulate gyrus seems to be necessary for the generation and maintenance of energy for attention. Its destruction leads in humans to *akinetic mutism*, where patients no longer speak and if asked indicate that they have no energy since nothing is of any meaning to them. Lesions of the orbitofrontal cortex frequently cause *pseudopsychopathy* and a loss of social responsibility. The effects of lesions of the amygdala, the cingulate gyrus, and frontal cortex show that the associative connection of peripheral physiological input and the evaluation of a situation is of central significance for emotional experience since the association is formed in this high-level social vegetative muscular regulation system. Persons with epilepsy of the temporal lobe frequently show emotional disturbances and psychiatric complications especially if the right anteromedial temporal regions that are closely connected to limbic structures are impaired.

1.05.10.4.3 Hormonal basis of aggression

Hormones can have an activating or organizing influence on aggressive behavior. Adult animals who have already been involved in fights no longer need testosterone to maintain their rank. Without the presence of androgens in the fetus and in the postnatal phase the neuronal connections that are needed to form aggressive behavior do not develop. In the hypothalamus and septum of certain strains of mice, androgen- and estrogen-sensitive neurons have been found that are correlated with various types of aggressive behavior.

Castration during puberty seems to inhibit postpubertal increases of aggression. In general, testosterone seems to influence the development of intermale aggression; other types of aggression seem to be less sensitive to variations in testosterone levels. Aggressive behavior seems to be also related to cholinergic activity: anticholinergic drugs suppress aggressive attacks in mice almost completely. Long-lasting learned helplessness leads to endorphine-mediated analgesia and immune suppression, as well as inhibition of the production of androgens. Animals that lose in fights also show opiate-induced analgesia.

1.05.10.4.4 Psychopathy

Psychopaths are persons who repeatedly commit antisocial aggressive acts without being impressed by punishment or negative consequences. They seem not to show any remorse or guilt following antisocial activities, however, they are usually intellectually capable of understanding their behavior as well as its consequences. In general, psychopaths seem to be characterized by reduced autonomic responding in situations that usually elicit fear. In passive avoidance situations, psychopaths are often impaired since they are unable to suppress behaviors that have previously been punished. Rather they continue to execute punished sequences, indicating an inhibitory behavior deficit (lack of fear). Consequently, adrenaline and amphetamine seem to improve the behavior of sociopathic persons, and barbiturates and alcohol can increase psychopathic behavior. In a series of studies, Patrick (1994) has shown that fear potentiation of the startle reflex is absent in psychopaths. Lesions in the orbitofrontal cortex lead to a pseudopsychopathic state that is related to a lack of delay of reinforcement and an inability to learn from punishment or the lack of reinforcement. Therefore, psychopathy might be related to deficits in both the amygdala and the orbitofrontal cortex.

1.05.10.5 The Neocortical Hemispheres and Emotion

Differences between the right and the left hemispheres have mostly been discovered in split-brain patients or in patients with specific lesions of the right or the left hemisphere during presentation of tasks that force the unilateral processing of information (e.g., dichotic listening, fixation). Usually, there is continuous exchange between the hemispheres with the left hemisphere playing an important role in the interpretation and attribution of causes. Lesions of the right hemisphere often lead to emotional indifference or euphoric disinhibition, lesions of the left hemisphere may lead to catastrophic reactions with severe depression. In the interpretation of these results it is important to consider that lesions of one hemisphere may result in a disinhibition and thus overactivation of the other hemisphere. Emotional expression is impaired after lesions in right frontal cortex, emotional recognition and discrimination after right posterior lesions. In right parietal lesions the existence and consequences of disease and/or emotional contents are often denied (sensory and emotional neglect), and emotional expression fades or is inadequate. Electroconvulsive treatment for depression is significantly more effective when applied to the right rather than the left hemisphere. By contrast, a left-sided Wada test leads to a depressed state. For the Wada test a sedative drug is injected in the right or left main brain artery, putting the respective hemisphere to sleep. The sensitivity of the right hemisphere for negative feelings is also supported by the fact that aversive stimuli such as pain or unpleasant odors primarily activate the right and positive stimuli primarily activate the left hemisphere.

In depressive disorders, increased right-frontal activation has been reported, and for mania, increased left-frontal activation. Since motor activity and control of right-handers are regulated predominantly by the left hemisphere, right-hemisphere overactivation leads to difficulty in controlling verbal and motor behaviors. This is supported by the fact that left-handers and ambidextrous persons often show emotional, verbal, and psychosomatic disorders. Dyslexias and allergies as well as hyperactivity and irritable bowel syndrome are more frequent in left-handed persons who have a larger corpus callosum with more fibers and a larger right hemisphere. This has been associated with a larger influence of testosterone during development which facilitates growth of the right hemisphere and inhibits growth of the thymus gland (immune competence) in animal experiments.

1.05.10.6 Behavioral Medicine: Application of the Psychophysiology of Emotion to Disease

Behavioral medicine is an interdisciplinary field concerned with the integration of behavioral and medical knowledge in the prevention and treatment of disease with a strong focus on behavioral intervention methods. One of the most important intervention methods in behavioral medicine is *biofeedback* treatment where a biological signal is recorded, amplified, and subsequently converted to a sound or visual signal which can be perceived by the patient. Thus the control over physiological processes that function normally outside our conscious awareness becomes possible. Biofeedback treatment has been used to successfully influence chronic pain, epilepsy, scoliosis and kyphosis, and Raynaud's disease, among others (Schwartz, 1995). Biofeedback follows the principles of operant learning. It is so far not known to what extent control over autonomic function is possible without involvement of the muscular system which is under voluntary control. The curarization experiments by Miller (1978) and his colleagues designed to address this issue yielded controversial results, but supported in general the notion that operant conditioning of autonomic function is possible. The self-regulation of brain potentials without the alteration of peripheral physiological variables also favors this interpretation.

1.05.11 COGNITIVE PROCESSES AND COGNITIVE DISORDERS

1.05.11.1 Introduction

Cognitive functions are all conscious and nonconscious processes that occur during the processing of information from within or outside the organism, for example, encoding, comparison with stored information, distribution of information, decoding, and language. The focus of this section will be on the neurobiological foundations of these processes.

1.05.11.2 Functions of the Cerebral Hemispheres

The term cerebral asymmetry denotes the finding that the functioning of both neocortical hemispheres is an important prerequisite for the regulation of various behaviors and mental functions. Although the right and the left hemispheres usually have a synergistic action in most higher cognitive functions, there is hardly any response where one of the two hemispheres of the brain would not have a certain advantage compared with the other. Although the idea that the two hemispheres subserve different cognitive functions originated in antiquity, it was Paul Broca who provided empirical data to substantiate the different functions of the two cortical hemispheres. Broca found that persons with aphasias all had lesions in the left frontal hemisphere. Shortly after, Karl Wernicke published his observations on patients with left posterior lesions—especially the superior temporal gyrus—which caused a so-called sensory or receptive aphasia; patients could speak but had difficulty in understanding what was told them. Wernicke developed the concept of the regulation of language that is still valid today: the connections of the primary auditory area to the upper posterior temporal lobe (Wernicke's area) and from there to the lower posterior frontal lobe (Broca's area). He also introduced the term "disconnection syndrome" by postulating specific deficits if the connections between these three areas were interrupted. His student Hugo Liepman described a right-handed patient who showed an apraxia of the left hand (inability for voluntary movement) subsequent to a lesion of the corpus callosum. The patient also showed agraphia on the left side (inability to write). In post-mortem studies it was shown that the corpus callosum had been destroyed. Liepman concluded from this case that (i) commands for left-sided movements had to be transferred by the corpus callosum to the right hemisphere and (ii) the left hemisphere was not only relevant for language but also for complex learned movement. This second postulate is supported by the fact that apraxias occur predominantly after left-hemispheric lesions. Subsequent to these lesions, complex voluntary movements cannot be correctly performed even after much training. These apraxias are also called ideomotor apraxias. It is, however, known today that planning, initiating, and execution of voluntary movements can also be initiated in the right hemisphere. The so-called constructive apraxias occur after lesions of the right hemisphere: the deficits that occur here relate, for example, to the construction of puzzles, drawings of clockfaces that can no longer be performed, and drawings and model buildings that can no longer be constructed. Apraxias can, however, also occur after lesions of the basal ganglia and the thalamus.

In the twentieth century, a counter localization position was strongly formulated by Karl Lashley who introduced the principle of equipotentiality. Lashley postulated that specific brain centers were not important for learning and memory but the mass action of several areas of the brain was important. The larger the lesion, the stronger the defect. The position that is taken today is a combination of the two views: certain cognitive processes such as language or motor commands are relatively dominant in one of the hemispheres, other cognitive processes are not specialized to certain brain regions.

Sperry's work on split-brain preparations in cats and monkeys and later a split-brain analysis in patients with complete resection of the corpus callosum have made important contributions to this question (Sperry, 1952, 1964). A number of influences on sensory processing have been revealed in split-brain patients. If the anterior commissure and the corpus callosum are separated, odors from the right nasal mucosa can no longer be named since the olfactory tract is uncrossed. The right hemisphere can, however, use the left hand to choose between objects of different odors. The right hemisphere is more strongly activated by odors from the right nostril than from the left.

Although the auditory tract is only partially crossed, dichotic presentation of words leads to a preference of the right ear. Sounds presented to the right ear are no longer moved to the left hemisphere in split-brain patients, but simple instructions can also be understood from the right hemisphere. Objects that are not viewed in the right visual field cannot be described, although the right hemisphere can correctly identify them (disconnection agnosia).

Tactile discrimination of the hands cannot be transformed from the secondary somatosensory cortex of the contralateral to the ipsilateral side. This means that learned tactile discriminations remain localized in the hand of the contralateral hemisphere. Separate presentations of visual contents to the right and left visual fields have shown that the right hemisphere is dominant in the analysis of visual spatial patterns and manipulative spatial tasks. Although the right hemisphere is in almost all patients expressively aphasic, drawings are only copied correctly by the left hand. Geometric figures are recognized better when they are projected to the left visual field and geometric and tactile tasks are also solved better by the left hand. The right hemisphere is dominant in face recognition, but not, however, in tasks where the faces are semantically categorized. The presentation of chimeric stimuli leads to separate completions of the entire picture in each hemisphere, although all patients report seeing one unified picture.

The sequential lesion of various fiber bundles including the posterior corpus callosum and the anterior commissure in monkeys showed all the brain structures that are involved in visual discrimination. Visual information leaves area 17 via areas 18 and 19 (visual memory), then crosses to the contralateral side and runs on both sides of the lower temporal lobe and from there to the amygdala and the orbitofrontal cortex (emotional and motivational significance). Here again a transfer between both sides takes place. This tract is therefore called the *what-pathway* compared with the upper parietal connection which has been termed the *where-pathway* (Mishkin & Ungerleider, 1982). Unilateral lesions of these systems have few consequences in the monkey; however, disconnection of efferents from areas 17, 18, or 19 in the corpus callosum with additional disruption of connections to the temporal areas leads to blindness, despite preserved optic systems. In addition to apraxia of the left side of the body, agraphia of the left hand to verbal commands and acopea (an inability to copy writing and drawing) of the right hand, tasks that require a cooperation of fine motor movements of both hands are impaired. The few-split brain patients who have expressive language in the right hemisphere suggest that the left hemisphere is rather sequential, analytic, and causal, and that the right is more holistically oriented with rather parallel and intuitive processing.

In new-born babies a preference of left hemispheric processing of language (e.g., right ear) can be detected in dichotic listening tests and in elevated amplitudes for evoked potentials for language on the left side. However, left hemispheric lesions lead to compensatory uptake of language functions by the right hemisphere up to the 10th year of life. Complete compensation seems only to be possible in the first years of an individual's life. Obviously there is an innate predisposition for the localization of language on the left side. Newborn babies already have a larger left-sided planum temporale which is the region that lies within the Sylvian fissure behind the auditory cortex. In adults, the left planum temporale is about 2 cm longer than that of the right side. The better a language has been learned, the smaller the cortical area that is needed to produce a certain language performance. This holds at least for cortical blood flow in PET and fMRI studies: the second less well-learned language has significantly larger increases of blood flow on the right and the left hemisphere. Women have less activation of blood flow during language processing than men.

In addition to the genetically determined anatomical preconditions for laterality, auditory experience during development seems to be a central influencing factor for its development: language-deprived children, that is, children who had very little external stimulation to speak, and are therefore also impaired in speaking, show less left lateralization in dichotic listening tests similar to persons who are born deaf. The perception and motor regulation of sign language seem, however, to be primarily left-dominant which depends on the superiority of the left hemisphere for sequential information processing as well as its dominance for complex and learned movements.

The Wada test has been used as an assessment instrument for the analysis of dominance. This test is used before neurosurgical operations to diagnose the dominance of one hemisphere.

Usually a narcotic agent (sodium amytal) is injected into the carotid artery. This leads some minutes later to ipsilateral anesthesia of the entire hemisphere. For the left hemisphere there is usually complete right-sided paresis and global aphasia. Some 96% of right-handed and 70% of left-handed persons have language localized on the left side and 4% of right-handed persons have language localized on the right side; 15% of left-handed persons have a bilateral language localization based on the Wada test.

The relationship between laterality and handedness is positive but far from perfect. The localization of language is in humans usually a much better predictor of an anatomic organization than hand preference. Left-handedness is usually not related to an improved ability of the right hemisphere for expressive and receptive language; however, left-handed people often show less language problems after left hemispherical lesions. Although the causes for left-handedness are not clear, there seem to be two groups of left-handers: one group with a strong genetic component and a second group where left-handedness may be the consequence of pre- or postnatal brain damage of the left hemisphere with a subsequent compensation of the right hemisphere. The first group is not intellectually different from right-handers, except for a certain increase of musical and artistic talents among left-handers.

Tactile tasks that require the recognition of forms and figures seem to be solved better by the left hand in right-handers if the material is presented in a static manner. In the case of sequential presentation the right hand makes less mistakes. The asymmetry of information processing is complemented by an asymmetry of the planning and execution of movement: although the left hemisphere is dominant in right-handers in learned skill tests (thus apraxias are often correlated with aphasias), we move the head and the eyes more to the right side when we solve verbal problems, when we solve spatial problems we move more to the left. Simultaneous speaking interferes more with activities of the right hand, the humming of melodies interferes more with the left hand. There seems to be an interaction of left dominance for motor activity and preferred cognitive processing style.

There is some convergent opinion that the preference of the right hand in about 75% of humans has to do with the upright gait of humans. The preference for the right side of the body is already present at birth. Hand preference develops later than the superior ability of the right hemisphere to process visuospatial tasks. The lateralization of visuospatial functions in the right hemisphere might be caused by the preferred activation of the left fetal vestibular organs and thus the right hemisphere during pregnancy. The hypothesis of preferred stimulation of the left vestibular organ and the right ear during pregnancy attempts to explain a number of differences in lateralization: for example, that women have a superior ability in verbal fluency (left hemispheric function) but have on the other side a less well-developed ability for spatiogeometric tasks which are solved better by men. Lesions of the left hemisphere lead equally often to aphasias in men and women, although within the hemisphere women tend to be aphasic and apractic rather after anterior, men rather after posterior lesions. In most studies on healthy persons, women were less lateralized in all tasks than men, which suggests that a strong lateralization does not necessarily mean better performance. All the differences between the performance of the left and right hemisphere we have named could be based on a joint anatomical difference, the more variable and therefore flexible intracortical communication of the left hemisphere.

Lateralization of cognitive functions may depend on anatomical differences between the two hemispheres. Left–right differences have not only been found in different areas of the cortex, for example, Broca's and Wernicke's regions, but also in subcortical areas, for example, in the thalamus. These differences are not only on a macroscopic level, but are also visible on a microscopic level in the neuroanatomy of individual neurons, for example, the size of the soma of a pyramidal cell or the structure of their dendritic branches, implying that neuroanatomical differences lead to functional differences (Kolb & Wishaw, 1995; Springer & Deutsch, 1993).

1.05.11.3 Evolution and Neurophysiology of Language

For thousands of years language has been considered as a typically human achievement that separates humans from animals, especially the primates. However, it has been shown that differences between human and animal language are quantitative rather than qualitative. The increase in speed and the heightened variability of the exchange of neuronal communication in the human cortex seems to be a precondition for the achievements of human language as a means of communication.

Vocal language seems to have begun only about 60 000 years ago, and written symbols have been found from up to 30 000 years ago. Language could have developed out of nonverbal communication: gestures as well as vocal sounds are generated primarily in the left hemisphere, and both functions are lost when lesions of the left parietal cortex occur. Emotional sounds and expressions that can, for example, be observed in primates, are less likely to be the beginning of language: emotional sounds are difficult to be conditioned in order to be used as signs; gestures, however, can also be conditioned in primates. Emotional expression seems to be primarily a frontal and basal ganglia phenomenon in human as well as in primates. Nevertheless, the development of language has begun from gestures and the increasingly perfect control of the facial musculature. Gross movements of the body could have been replaced by subtle movements of the lips and the tongue (Crosson, 1992).

The acquisition and use of language can be explained by associative connections of phoneme, word, and sentence assemblies of cortical and subcortical neuronal assemblies. The neocortex may be—especially in its associative areas—viewed as a large associative storage area. The storage of words, sentences, and syntax is part of the functions of both hemispheres. Like the visual cortex, the auditory cortex also has hierarchically built interconnections of neurons that range from simple to hypercomplex and that respond selectively to various aspects of sound: individual cells answer preferentially to certain frequencies, the beginning and the end of sounds and phonetic characteristics of syllables (ba, pa) and consonants (b, g). Many neurons that represent such simple aspects are interconnected to higher-order assemblies if they are frequently excited together. With respect to the language lexicon, therefore, assemblies of phonemes are built that are characterized by different properties. These coupled phonetic networks are connected during development by further associative connections to syllables and word forms. For example, certain cells in the upper temporal cortex show a very characteristic frequency when a word is acoustically presented, since the same cell is part of a different cell assembly based on the type of word that was presented.

Neonates are already more sensitive to syllables of the mother tongue than to control syllables in the pre- and postnatal phase. Between the sixth and the 12th month of life, syllables that are heard and last about 200 ms (auditory cortex) are frequently repeated (inferior frontal cortex) and are thus associatively interconnected to transcortical syllables and later to phonetic assemblies. These articulations lead, of course, also to proprioceptive stimuli from the articulatory musculature that activate an inferior parietal assembly and are thus connected to an entire assembly in the perisylvic region.

When a child learns that certain word forms always occur in certain contexts (at about the age of 2–4 years), this leads to the simultaneous activation of cell assemblies in many cortex areas (e.g., with respect to a drinking glass: visual for glass, tactile for touching the glass, gustatory for taste, etc.) Cell groups are activated that are then the assembly for the content word that represents an object or an action (nouns, verbs, and adjectives). Depending on the associative context of the word, certain brain areas are activated in an assembly that can be assessed by δ-band EEG activity. For example, verbs elicit more motor association and substantives elicit more visual association. This is visible in γ-band oscillations that are more pronounced in the occipital cortex for the nouns and more pronounced in the frontal cortex for the verbs.

The neuronal basis of function words (how, it, is, etc.) that have syntactic functions and are not associated with specific environmental stimuli seems to be the left perisylvan region. When semantic errors occur in a sentence, an event-related brain potential is detected—usually a strong negativity (around 400 ms)—whose location in the brain seems to vary with the location of storage of the respective word. During syntactic errors, late positivities in the left perisylvic regions (around 600–800 ms) seem to occur or late frontal negativity.

It is, however, not clear to what extent those event-related potentials are language specific or to what extent they reflect the excitatory thresholds in semantic or syntactic cell assemblies as has been described before. N400 and other negativities after surprising events seem to indicate a new readiness state of the brain region with new search processes for strategies to solve problems. It is, however, possible to see from the context of a language task based on the event-related potential whether a word or part of a sentence has been detected as right or wrong, when this happens in the brain, and at what location. This can be used for lie detection: the cortical lie detector is a higher positive wave (P300) to words or objects compared with similar control stimuli only the perpetrator can know. If the sequence of the crime is known and test stimuli are well presented, the event-related potentials can lead to very high lie detection quota of 90–100% which are much higher than lie detection based on skin conductance readings.

During imagery of an object, all those brain areas are activated that are also active during its perception. The sequence of the activation seems, however, to be reversed: where the real visual stimulus first activates area 17 (primary visual area) and then the extrastriatal area, especially area 18, imagery first activates area 18 and only afterwards area 17. The time course is, however, comparable to real perception: evoked potential changes in area 18 are already visible about 200 ms after the beginning of an image. Blood flow changes as measured in PET and MRI are in general in the same region as the perception occurred; however, the retrieval of visual contents from memory leads to PET activation, especially in the left temporal parietal, right parietal, and bilateral frontal cortex which occurs in addition to the activation that occurs during perception (Farah, 1995). The frontal activation which is also visible in EEG is probably related to the activation of working memory. The left perisylvic brain region is always necessary during imagery if it is language related.

1.05.11.3.1 Language disorders

We will not describe language disorders in detail but rather discuss their neurobiological basis. *Aphasias* are language disorders that occur in humans who have already learned a language. The cause is usually an ischemic or hemorrhage-related cerebrovascular insult, a tumor, encephalitis, or trauma. All language modalities are usually disturbed. Although lesions of the left hemisphere usually lead to aphasias, right-hemispheric processes are also involved in language processing. Aphasias also occur in subcortical lesions of the white matter, the basal ganglia, or the thalamus. These subcortical aphasias with initial mutism are usually quickly reversible. The cortical aphasia-causing lesions are usually situated in areas close to the Sylvian fissure: Broca's region (Brodmann's areas 44 and 45) and Wernicke's region (area 22) can be differentiated.

All aphasias include disturbances in the naming of objects, the production and understanding of sentences, reading, and writing. The currently most widely used aphasia test, the so-called Token Test, examines to what extent a patient is able to perform manual manipulations with a number of colored sheets. Almost all aphasics show deficits in this test.

The following aphasia syndromes are clinically significant:

(i) Broca's aphasia, where language production problems are predominant;

(ii) Wernicke's aphasia, where language perception problems are predominant;

(iii) global aphasia, where the entire perisylvic region is lesioned and where both language production and language understanding are severely incapacitated;

(iv) amnesic aphasia, which is usually based on lesions of the gyrus singularis or other areas that are close to the left perisylvic regions, sometimes also lesions in the right hemisphere—this aphasia is characterized by minor language deficits, deficits that are characterized by semantic paraphrases and disorders in naming objects;

(v) transcortical aphasias that are characterized by deficits in language production and language understanding or both capabilities and are based on lesions in most of the left area perisylvic area;

(vi) conduction aphasias which show a strong disorder in the ability to repeat words with lesions in the fasciculus arcuatus;

(vii) subcortical aphasias which show an initial mutism and subsequent paraphasias.

An acquired inability to understand written language is called *alexia*. Alexias occur with and without *agraphia*, based on localization of the lesion in the brain tissue. Alexia without agraphia and aphasia seems to be a disconnection syndrome where the connections from the right visual association cortex through the corresponding language area in the left gyrus angularis are interrupted. Alexias with agraphias usually show an isolated lesion in the left gyrus angularis. Agraphias occur very frequently together with aphasias. They can, however, also be assessed independently which points towards partially separated brain structures with regulation of both functions. Lesions in several, often widely different brain regions can cause disorders of writing, which is understandable with respect to the complexity of writing which involves semantic, visual, spatial, and motor functions.

1.05.11.4 The Association Areas of the Neocortex

In a simplifying manner, the three large associative areas of the cortex could be connected to the three main psychological functions: whereas the temporal lobe with its limbic connections seems to subserve primarily memory functions, the frontal cortex seems to regulate motor and motivational behaviors, and the parietal association cortex seems to be the basis of sensory and cognitive function. The posterior striatum is, however, also an important part of parietal function because it provides an indirect subcortical connection specifically to prefrontal regions.

1.05.11.4.1 Parietal cortex

The parietal cortex with its multisensory integration function therefore has an important role as a command center for structuring movements that are directed towards a goal of motivational meaning. The large number of cognitive disorders related to lesions of the parietal region is based on its central anatomical role between the three modalities of vision, hearing, and somatic sense which provide the parietal lobe with information. The posterior parietal lobe with the gyrus angularis, the gyrus supramarginalis, and the upper parietal lobe is disproportionally large on the right hemisphere in humans. This seems to be related to spatial information processing and the regulation of goal-directed movement in space. The efferents of the posterior parietal cortex project to the frontal and temporal association areas, thalamus, striatum, midbrain, and spinal cord. In addition to afferents from the three primary and secondary projection areas, the parietal lobe also receives input from the lateral and posterior thalamus and hypothalamus.

Engrams for reafferent motor and visuokinetic spatial functions are localized in the parietal cortex. The copying of movements and gestures is disturbed after left parietal lesions, and spatial activity related to drawing is disturbed after lesions of the right hemisphere. A number of lesions lead to disturbances, which inhibit the initiation, planning, and execution of complex and goal-directed movement sequences in left-hemispheric ideomotor *apraxia*. The sequential spatial functions of the left parietal region need to be separated from the perceptive functions of the right parietal region. Visual-perceptual skills are, for example, the recognition of visual objects, the synthesis and comparison of visual objects, line orientation, drawing of lines, and recognition of weights. Visuospatial skills are the location of objects in space, judgment of the direction and distance–topographic orientation in space, and the localization of one's own body or body parts in space. Lesions of the parietal lobe therefore frequently lead to topographic *agnosia* and amnesia. Objects, landmarks, and one's own position with respect to them and orientation are disturbed or are not remembered. For example, the orientation of hand movements to a visible goal is not possible. In parietal-occipital lesions, the cell assemblies for individual characteristics of a stimulus, for example, color, form, and location, can no longer be associatively connected. The consequences of this are completely separate experiences of individual elements of the environment and the loss of orientation, for example, the color of letters can no longer be recognized and the location and size of geometric objects can no longer be compared.

Especially remarkable are disorders of face recognition (*prosopagnosia*): the lack of discrimination of unknown faces occurs during lesions of the right parietal lobe, the disorder of the recognition of known faces including one's own in the mirror, however, depends on bilateral occipital–parietal lesions. Contralateral neglect occurs especially after lesions of the lower right parietal lobe. The person no longer responds to visual tactile and acoustic stimuli contralateral to the lesion (usually the left side of the body). The person also reports no contents of this side and never orients towards this side if new stimuli are represented. *Neglect* can be viewed as a spatial perceptual disorder or an attention disorder. That the perception is disorganized is suggested by the neglect of the left side independent of attention. If the patients, for example, are to view the square in front of the dome of Milano, they only report about the right side of the dome during perception as well as during imagery. If a patient, however, changes his or her location so that the left visual field now receives the previously neglected side of the dome, the patient now reports in reality as well as in imagery the previously ignored part of the dome. The representation of the external object is obviously lost depending on the local frame of reference of the person.

The *parietal cortex* seems to be part of a widely distributed cortical–subcortical attentional system. The multimodal, parietal, and superior–temporal association areas have the task of comparing incoming excitatory patterns with stored ones and to extract from them the significance of the pattern. Whereas temporally the "what" of objects seems to be analyzed, the parietal region seems to analyze the special location "where." Motivational and activating structures in the limbic system, the basal ganglia, the reticular formation, thalamus, and the frontal cortex play an important role in this attentional system. Unilateral lesion of each of the subsystems can cause unilateral neglect. The frontal and temporal regions modulate the thalamic filter system of the nucleus reticularis by inhibitory connections. Reticular formation activates all cortical regions, but inhibits the nucleus reticularis, which leads to selective gating of the thalamus for incoming regions by its inhibitory influence on all specific afferents. The right hemisphere is dominant in the regulation of these processes during spatial attention. Neglect can therefore occur after every lesion in any of these regions, especially, however, on the right side. The patient can no longer learn to concentrate on the contralateral side. Lesions of the parietal lobe also lead to disturbances of short-term memory and therefore also long-term encoding, especially of visuospatial material. This does not mean that short-term memory is an exclusive function of the parietal cortex, but only that important steps in the processing of information in declarative short-term memory depend on multisensory integration of the parietal region.

1.05.11.4.2 The frontal lobe

Even extensive lesions of the *frontal cortex* which forms about 30% of the neocortex usually create no sensory or motor dysfunction. They have, however, severe psychological consequences. The frontal cortex, especially the dorsolateral frontal and prefrontal cortex, has an important role in the creation of stable contingencies between responses and their consequences especially if those follow with delay; associations between a cue, the following motor response, and the biological–social consequences lead to the formation of stable expectancies. The frontal cortex plays an important role in those expectancies: if the frontal cortex is lesioned behavior loses its future orientation, behavior is difficult to predict, is irregular and extremely stereotypical and perseverative.

In general, three subsystems of the frontal cortex can be differentiated: (i) the motor and premotor regions including Broca's area and the frontal eye fields (areas 4, 6, 44), (ii) the dorsolateral frontal cortex with areas 8, 9, 10, 44, 45, and 46, which is called the prefrontal cortex together with the orbitofrontal cortex, and (iii) the orbitofrontal cortex which includes areas 11, 12, and 47. The prefrontal cortex has often been called the granular frontal cortex because it has an especially well-developed layer IV of stellate cells and spherically organized short dendrites. Areas 4, 6, 8, and 44 have primarily motor functions, the dorsolateral frontal cortex is related to attentional functions, working memory, and the formation of expectancies; the orbital system is related to motivational functions. The most important afferents to the prefrontal cortex come from the medio-dorsal nucleus of the thalamus. In addition, the anterior nucleus of the thalamus, the hypothalamus, the amygdala, the limbic cortex, and the cingular gyrus as well as the mesencephalic reticular formation and the nonprimary sensory and motor association areas project to the prefrontal cortex. All connections are reciprocal, the prefrontal cortex projects especially to area 7 (somatic), area 22 (auditory), and area 21 (visual). The orbitofrontal cortex also has olfactory connections. The efferents of the frontal cortex go to the basal ganglia, the hippocampi, and the limbic region. The connections of the basal ganglia are especially important in the regulation of attentional behavior.

Lesions of the prefrontal and orbitofrontal cortex lead to disorders of the temporal sequence of behavior and therefore to secondary social disturbances. Lesions of the prefrontal cortex lead to typical changes in behavior that have been described in many textbooks using the example of Phineas Gage, whose brain was damaged by an iron rod as a consequence of premature detonation of explosives. Lesions of the frontal cortex damage the ability to plan ahead and to select the right behavioral sequences, to ignore distracting stimuli, and to continue with behavior that has begun, and to remember what has already been done previously. This temporal organization of behavior by the prefrontal cortex needs detailed information about sensory stimuli and the context of the situation from the parietal and temporal cortex as well as the motivational value of a specific given and stored situation. In animal experiments lesions of the dorsolateral frontal cortex, especially of the stria principalis, lead to disorders that include attention as well as learning functions. Tasks that require delayed responses are especially disturbed. Delayed matching to sample (DMS) tasks are also affected: in these tasks a visual stimulus is usually presented and with some delay the same stimulus is presented together with other stimuli and the animal must choose which stimulus has previously been seen; the right choice is reinforced. The target stimulus is changed on a trial-to-trial basis. DMS tasks are also affected after temporal lesions, since they require visual discrimination. Table 1 gives an overview of the various deficits after lesions of the frontal cortex including its motor and premotor areas.

In expectancy situations, animals as well as humans show a characteristic slow cortical potential. Between a warning stimulus and a second imperative stimulus which requires motor or cognitive responses, two negativities or components of the EEG are formed: one after the first stimulus and one before the second stimulus. The first negativity usually has a prefrontal localization and the second one can vary topographically over various cortical areas depending on the response that has to be executed. This first component is reduced over the frontal cortex when motor acts become automatic, when stimuli habituate, and when stimuli have no signal value, informative stimuli increase this component. The first frontal component of the SCPs seems to be related to preparatory activation of memory contents after presentation of the first warning stimulus. The expectancy of the stimulus is the subjective correlate of this preparatory activation. In patients with bifrontal lesions this first component is absent and explains the disturbance of

Table 1 Overview of the main symptoms after lesions of the frontal lobe.

Symptom	*Site of lesion*
Disorders of movement	
Loss of fine motor skills	Area 4
Loss of force	4,5, dorsolateral
Faulty planning of movement	Premotor, dorsolateral
Voluntary fixation of eyes	Frontal eye fields
Disturbed corollary discharge	Dorsolateral, premotor
Broca's aphasia	44
Loss of divergent thinking	Orbital
Reduced spontaneity	
Disorders in behavioral strategies	Dorsolateral, orbital
Stimulus control of behavior	
Deficient inhibition of responses	Dorsolateral
Search for risk and loss of rules	Prefrontal
Disorders of associative learning	Dorsolateral
Problems in memory for time	
Disturbed perception of frequency	Dorsolateral
Disturbed retrieval of sequences	Dorsolateral
Delayed responses	Dorsolateral
Disordered social behavior	Orbital, dorsolateral
Altered sexual behavior	Orbital
Disordered olfactory discrimination	Orbital

Source: Kolb and Whishaw (1995).

expectancy processes with longer time intervals between two stimuli as found in delayed response tasks. The basic learning of the association between a response and its consequences or the association between two stimuli seems to depend considerably on prefrontal structures. Between CS and UCS or response and reinforcement, repetitive temporal pairing results in an electrophysiological connection that is mirrored in surface negativity. The prefrontal cortex has an important role in the temporal regulation of information and motor acts for other neocortical regions. This active mechanism takes its energy from limbic and subcortical structures. If this source of energy is absent or distributed temporally independent of learned cues, the temporal sequence of thought and behaviors becomes irregular and the probability for goal-directed behavior is reduced. This has been confirmed for cases of bilateral lesions of the prefrontal cortex.

The prefrontal cortex seems to be activated in all situations when a stimulus requires a temporal discrimination. This is the case with all warning stimuli and conditioned stimuli: the organism estimates the probability of occurrence of a certain stimulus and mobilizes the respective sensory and motor systems in a preparatory manner. If the temporal estimation of future events is faulty or destroyed, the time of sensory and motor mobilization is not correctly chosen or it may be suppressed. The consequence is the symptomatology of the frontal patient, which has been described as difficult to predict, with irregular behavior, especially in waiting situations and with extreme distractibility. In lesions of the dorsolateral frontal cortex it has been noted that verbally formulated action programs loose their influence on behavior. The disorganization of behavior plans is accompanied by distractibility, the inability to give up response strategies that have once been chosen (perseveration). The inflexibility and perseveration is especially caused by lesions of left area 9, the distractibility occurs in prefrontal regions in general. In many types of disorders (e.g., anxiety, schizophrenia, tics, obsessive behavior, and criminal behavior), the frontal lobe used to be destroyed in order to reduce symptomatology. Bilateral lesions lead to a pseudopsychopathic state, especially if the right frontal cortex is lesioned and a pseudodepressive state occurs after left frontal lobe lesions. Although there is similarity in those behavioral deficits to psychopathic behaviors, there is less stability in behavior in frontal lesions: within a few minutes usually incompatible behaviors can occur in a fast sequence. Self-control requires a series of cognitive operations that require intact prefrontal regions.

1.05.11.4.3 The temporal lobe

The *temporal lobe* has acoustic, visual, and memory functions that are reflected in its anatomical structure. Based on its anatomic substructures, the temporal lobe has sensory functions (primary and secondary auditory system) and portions of the tertiary visual system (superior and inferior portions); the medial and limbic portion by way of contrast has memory functions and is responsible for affective tone. Whereas the parietal cortex is related to the location of function, the temporal cortex is related to the differentiation of objects. The temporal lobe comprises neocortical regions 20, 21, 22, 37, 38, 41, and 42, which are also called archicortex, and the medial phylogenetically older, tripartite portions of paleocortex: gyrus ambiens, parahippocampal gyrus, and uncus, the entorhinal (area 28) and perihinal cortex (areas 35 and 36) belong to the medial temporal memory system—the hippocampus and amygdala are closely associated with the paleocortex.

Table 2 shows the most important functional disorders after lesions of various temporal regions, based on Kolb and Whishaw (1995). The role of the inferior gyrus temporalis for visual discrimination has been demonstrated in animal as well as human experiments. Bilateral lesions of the auditory cortex do not lead to cortical deafness as is the case in the primary visual field. However, tone discrimination is impaired, especially the minimal time that has to pass in order to discriminate two tones or two vocal sounds. The minimal stimulus duration for tone discrimination is 50 ms, but after lesions it can increase substantially which leads to an inability to perceive language if the sounds and words are not presented for a very long time. This is especially true for the left superior posterior temporal lobe. This sound discrimination seems to be impaired in *dyslexias* and in disorders of language comprehension in children. Extensive training of the discrimination ability can restitute the cortical capability in these children. In monkeys it has been shown that extensive training of acoustic discrimination in the superior temporal lobe and training of visual discrimination in the lower temporal lobe leads to an increase in frequency of action potentials to tone sequence and visual patterns of more than 30% compared with the time before training. As in the somatosensory cortex there is an expansion of complex receptive fields related to learning. Musical talent seems also to be related to the interplay between the two temporal cortices, for example, musicians with and without absolute pitch differ in the size of the left planum temporale (larger in persons with absolute pitch) (Schlang, Jancke, Huang, & Steinmetz, 1995).

1.05.11.5 Thought Disorders

One type of thought disorder with extensive loss of intellectual function and memory are the *dementias*. More than half of all persons with dementias suffer from *morbus Alzheimer*. Very little is known about the causes of Alzheimer's disease. The early occurring severe forms of Alzheimer's might be a dominantly transmitted anomaly in cell metabolism. The mutation of a gene (*s182*) on chromosome 21 may be involved. This gene codes for the precursor protein of β-amyloid, which is responsible for the destruction of cells and can be found in large amounts in the brain of these patients (Levy-Lahad, Wijsman, & Nemens, 1995). Other researchers have also found mutations of chromosome 1, the function of which seems to be similar to that of chromosome 21. Typical for Alzheimer's disease is the atrophy of the gray matter, especially of the association cortices, atrophy of the dendritic branches, and the presence of histologically very well visible senile plaques and neurofibrillary tangles. Senile plaques are deposits of dead nerve cells with a high content of neurotransmitters and proteins. Neurofibrillary tangles consist of excessive multiplication of normal neurofibrils that usually have supporting function for the cells.

In *Parkinson's disease* a loss of dopamine in the nigro-striatal system has been found. The causes of idiopathic Parkinson's disease are unknown; the uptake of pesticides or drugs that destroy dopamine cells might be one cause which may, however, not be relevant for all patients. Parkinson's symptoms may also occur as a consequence of chronic intake of antipsychotic medication with antidopaminergic effects. Since Parkinson's disease is characterized by a loss of dopaminergic cells in the substantia nigra and the adjacent ventral tegmentum, and since the substantia nigra is part of a complex system with motor, cognitive, and motivational functions, motor deficits alone are improbable in Parkinson's disease. For example, Parkinson patients show disorders which are comparable to those in patients with frontal lesions.

The *schizophrenias* are a heterogeneous group of disorders of attention, perception, and language. The neurobiological causes of schizophrenia seem to be related to an overactivity of the mesolimbic dopamine system and a disorder of the left mediotemporal prefrontal cortex region. The neurobiological basis of type I and type II schizophrenia seem to be different: type

Table 2 Disorders after lesions of various temporal regions.

Symptom	*Site of lesion*
1. Disorders of acoustic perception	Areas 22, 41, 42
2. Disorders of the selection of visual and acoustic stimuli	Areas 20, 21, 22, 37, 38
3. Disorders of visual perception	Areas 20, 21
4. Disorders of acoustic perception	Areas 41, 42, 22
5. Disorders of organization and categorization	Areas 21, 38 left
6. Disorders of context memory	Hippocampus
7. Disorders of language perception	Area 22 left
8. Disorders in long-term memory	Area 21 (hippocampus and surrounding tissue)
9. Changes in personality and affect	Areas 21, 38 and amygdala
10. Changes in sexual behavior	Amygdala and temporal?

Source: Kolb and Whishaw (1995).

II patients show clear signs of a degeneration of the brain and have expanded ventricles correlating with intellectual decay. The genetic vulnerability of type II schizophrenia for a virus that might later lead to the schizophrenic disorder may be larger than in type I schizophrenia. Similarities between the type II group and patients with multiple sclerosis suggest an immunological involvement of the CNS in the disorder. In the cerebrospinal fluid of a subgroup of schizophrenics, antibodies against a viral infection similar to the herpes virus have been identified.

However, type I schizophrenics also show neurophysiological changes. Blood flow in the frontal hemispheres, especially the left hemisphere, compared with the occipital region, seems to be reduced, and the frontal glucose uptake in PET also seems to be lower. Event-related brain potentials show a decrease rather than an increase in negativity when attention is directed to a stimulus which might be an indicator of the attentional deficits that have been shown in schizophrenia. Structural changes have also been observed in the temporal hippocampal region. The hippocampus of a schizophrenic shows a disorganized orientation of cells. The growth of the cells in this region takes place in the second and third trimesters of pregnancy and this provides an additional argument for the embryonal cause of some types of schizophrenias. Potentially androgens have a triggering function in the brain of schizophrenics since schizophrenia usually occurs after puberty in men and after the age-related reduction of estrogens in women.

The *dopamine hypothesis of schizophrenia* was based on findings that medication which influences the dopaminergic system has a clear antipsychotic effect in schizophrenic patients. A substance seems to be effective if it quickly binds to dopamine receptors in the cell membrane of the mesolimbic system and if it inactivates the receptor. This suggests that an overactivity of the mesolimbic dopamine system may be a major cause of the positive symptoms in schizophrenia. It is assumed that the so-called D_2-receptors in schizophrenics are either increased in number or overly sensitive. Substances that block the D_2-receptors have the best therapeutic effect, whereas blockade of the D_1-receptors only has small effects on the symptomatology. A general overactivity of the dopaminergic neurons has not been shown. D_2-receptors are activated by apomorphine and are blocked by neuroleptic substances. Dopamine agonists such as amphetamine and cocaine can induce schizophrenia-like symptoms in healthy persons and may increase schizophrenic symptoms in schizophrenics. Whereas the mesolimbic dopamine system seems to be overactive, the frontal cortex seems to have lower metabolic activity as shown in PET studies. This could be related either to a compensatory dopamine suppression in mesocortical dopamine fibers or to the inhibitory effect of dopamine in the prefrontal cortex.

The use of dopamine antagonists leads to dyskinesias and a loss of quality of life as important side effects. The most frequent effects are Parkinson-like disorders in movement such as loss of facial expression and tremor. About 10% of the patients also develop irreversible tardive dyskinesias subsequent to long-term medication: tics related to the face and the tongue, incomprehensible gestures, problems in speaking, gross movement of the arms, etc. The receptors seem to develop a denervation supersensitivity: the ongoing inhibition of the receptors by the drugs causes compensatory irreversible supersensitivity to the blocked transmitter (Wise, 1982). Blockade of the overactive mesolimbic dopamine system by neuroleptica not only leads to a suppression of thought disorders but also to anhedonia. As previously described, the dopaminergic system plays an important role in intracranial self-stimulation: its inhibition reduces the efficacy of positive reinforcement and therefore leads to a general sedation that is medication-induced. The overactivity of the mesolimbic dopamine system as an incentive system could explain some of the symptoms of schizophrenia: since this system creates positive reinforcement and incentive motivation, overactivity leads to reinforcement of a number of stimuli and responses. Unimportant and not connected stimuli become suddenly important and relevant for behavior. The selection of attention breaks down. These are the most important symptoms of the schizophrenic disorder.

1.05.12 SUMMARY

This chapter has given an overview of the psychobiological foundations of clinical psychology. The basic fields of psychobiology, such as psychophysiology, psychopharmacology, psychoendocrinology, psychoimmunology, and behavioral genetics, were defined and important principles of these fields were described. The neurobiology of basic psychological functions such as emotion, motivation, memory, cognition, and language was then described and the psychobiological basis of their dysfunction in clinical–psychological disorders was delineated.

1.05.13 REFERENCES

Abeles, M. (1991). *Corticonics. Neural circuits of the cerebral cortex*. Cambridge, UK: Cambridge University Press.

Ader, R., & Cohen, N. (1985) CNS–immune system interactions: Conditioning phenomena. *The Behavioral and Brain Sciences, 8,* 379–426.

Ader, R., Felten, D., & Cohen, N. (Eds.) (1990). *Psychoneuroimmunology* (2nd ed.). New York: Academic Press.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Andreassi, J. L. (1995). *Psychophysiology* (3rd ed.). Hillsdale, NJ: Erlbaum.

Aschoff, J., Daan, S., & Groos, G. A. (1982). *Vertebrate circadian systems. Structure and physiology*. New York: Springer.

Aserinsky, E., & Kleitman, N. (1953). Regulary occurring periods of eye motility, and concomitant phenomena during sleep. *Science, 118,* 273–274.

Bailey, J. M., Willerman, L., & Parks, C. (1991). A test of the maternal stress theory of human male homosexuality. *Archives of Sexual Behavior, 20,* 277–293.

Becker, J. B., Breedlove, S., & Crews, D. (Eds.) (1992). *Behavioral endocrinology*. Cambridge, MA: MIT Press.

Birbaumer, N., Elbert, T., Canavan, A., & Rockstroh, B. (1990). Slow potentials of the cerebral cortex and behavior. *Physiological Reviews, 70,* 1–41.

Birbaumer, N., Flor, H., Lutzenberger, W., & Elbert, T. (1995). The corticalization of chronic pain. In B. Bromm & J. E. Desmedt (Eds.), *Pain and the brain. From nociception to cognition. Advances in pain research and therapy* (Vol. 22, pp. 331–344). New York: Lippincott-Raven.

Birbaumer, N., Grodd, W., Diedrich, O., Schneider, F., & Flor, H. (1997). *Functional magnetic resonance imaging reveals substantial activation of the amygdala during the processing of social stimuli in social phobics*. Submitted for publication.

Birbaumer, N. (1998). Self-regulation of slow cortical potentials. *The Neuroscientist.*

Birbaumer, N., Lutzenberger, W., Montoya, P., Larbig, W., Unertl, K., Töpfner, S., Grodd, W., Taub, E., & Flor, H. (1997). Effects of regional anesthesia on phantom limb pain are mirrored in changes in cortical reorganization. *Journal of Neuroscience, 17,* 5503–5508.

Birbaumer, N., Lutzenberger, W., & Schleichert, H. (1994). Think to move—and move! Self-regulation of movement-related brain potentials. *Society for Neuroscience Abstracts.*

Birbaumer, N., & Öhman, A. (Eds.) (1993). *The structure of emotion: Psychophysiological, cognitive, and clinical aspects*. Seattle: Hogrefe & Huber.

Birbaumer, N., Rockstroh, B., Elbert, T., Wolf, P., Düchting-Röth, A., Reker, M., Daum, I., Lutzenberger, W., & Dichgans, J. (1994). Biofeedback of slow cortical potentials in epilepsy. In J. Carlson, R. Seifert, & N. Birbaumer (Eds.), *Clinical applied psychophysiology* (pp. 29–42). New York: Plenum.

Birbaumer, N., & Schmidt, R. F. (1996). *Biologische psychologie* (3rd ed.). New York: Springer.

Bloom, F. E., & Kupfer, D. J. (Eds.) (1994). *Psychopharmacology. The fourth generation of progress*. New York: Raven Press.

Borbely, A. A. (1984). Sleep regulation: outline of a model and its implications for depression. In A. Borbely & J. L. Valatx (Eds.), *Sleep mechanisms*. New York: Springer.

Cacioppo, J. T., & Petty, R. E. (1983). *Social psychophysiology: A sourcebook*. New York: Guilford Press.

Cahill, L., Prins, B., Weber, M., & McGaugh, J. (1994). β-adrenergic activation and memory for emotional events. *Nature, 371,* 702–704.

Cameron, O. G. (Ed.) (1994). *Adrenergic dysfunction and psychobiology*. Washington, DC: American Psychiatric Press.

Cannon, W. B. (1929). *Bodily changes in pain, hunger, fear and rage*. New York: Appleton Century.

Carlson, J., Seifert, R., & Birbaumer, N. (Eds.) (1994). *Clinical applied psychophysiology*. New York: Plenum.

Carlson, N. R. (1991) *Physiology of behavior* (4th ed.). Boston: Allyn & Bacon.

Chalmers, N., Crawley, R., & Rose, S. P. R. (Eds.) (1971). *The biological bases of behavior*. London: Harper.

Clarke, E., & O'Mally, C. D. (1995). *The human brain and spinal cord.* (2nd ed.). San Francisco: Norman.

Clynes, M. (1989). Communication and generation of emotion through sentic form. In L. Levi (Ed.), *Emotion.* New York: Raven Press.

Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1991). *Selective attention modulates extrastriate visual regions in humans during visual feature discrimination and recognition. Exploring brain functional anatomy with position tomography*. Ciba Foundation Symposium. New York: Wiley.

Cotman, W. M., & Nadler, L. G. (1978). Reactive synaptogenesis in the hippocampus. In C. W. Cotman (Ed.), *Neuronal plasticity*. New York: Raven Press.

Crosson, B. (1992). *Subcortical functions in language and memory*. New York: Guilford Press.

Daum, I., Ackermann, H., Schugens, M. M., Reimold, C., Dichgans, J., & Birbaumer, N. (1993). The cerebellum and cognitive functions in humans. *Behavioral Neuroscience, 107,* 411–419.

Davidson, R. J. (1993). Cerebral asymmetry and emotion:

Conceptual and methodological conundrums. Special Issue: Neuropsychological perspectives on emotion. *Cognition and Emotion, 7,* 115–138.

Davis, H. P., & Roitblat, H. L. (1984). Cholinergic pharmacology, behavior and age-related memory decline. In L. Squire & N. Butters (Eds.), *Neuropsychology of memory*. New York: Guilford Press.

Davis, M., Hitchcock, J. M., & Rosen, J. B. (1991). Neural mechanisms of fear conditioning measured with the acoustic startle reflex. In J. Madden (Ed.), *Neurobiology of learning, emotion and affect*. New York: Raven Press.

Davison, G., & Neale, J. M. (1996). *Abnormal psychology* (6th ed.). New York: Wiley.

Di Chiara, G. (1995). The role of dopamine in drug abuse viewed from the perspective of its role in motivation. *Drug and Alcohol Dependence, 38,* 95–137.

Dworkin, B. R., Elbert, T., Rau, H., Birbaumer, N., Pauli, P., Droste, C., & Brunia, C. H. M. (1994). Central effects of baroreceptor activation in humans: Attenuation of skeletal reflexes and pain perception. *Proceedings of the National Academy of Science, 91,* 6329–6333.

Ekman, P. (1984). Expression and the nature of emotion. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion*. Hillsdale, NJ: Lawrence.

Ekman, R. & Davidson, R. A. (Eds.) (1984). *The nature of emotion: Fundamental questions*. Oxford, UK: Oxford University Press.

Elbert, T., Rockstroh, B., Lutzenberger, W., & Birbaumer, N. (Eds.) (1984). *Self-regulation of the brain and behavior*. New York: Springer.

Fahrenberg, J., Walschburger, P., & Foerster, F. (1979). *Psychophysiologische Aktivierungsforschung*. München: Minerva.

Farah, M. (1995). The neural bases of mental imagery. In M. Gazzaniga (Ed.), *The cognitive neurosciences*. Cambridge, MA: MIT Press.

Farwell, L., & Donchin, E. (1991). The truth will out: Interrogative polygraphy ("lie detection") with event related brain potentials. *Psychophysiology, 28,* 531–547.

Flor, H., & Birbaumer, N. (1993). Comparison of the efficacy of EMG biofeedback, cognitive–behavioral therapy, and conservative medical interventions in the treatment of chronic musculoskeletal pain. *Journal of Consulting & Clinical Psychology, 61,* 653–658.

Flor, H., Elbert, T., Knecht, S., Wienbruch, C., Pantev, C., Birbaumer, N., Larbig, W., & Taub, E. (1995). Phantom-limb pain as a perceptual correlate of cortical reorganization following arm amputation. *Nature, 375,* 482–484.

Folkard, S., & Monk, T. M. (1986). Chronopsychology: Circadian rhythms and human performance. In A. Gale & J. A. Edwards (Eds.), *Physiological correlates of human behavior* (Vol. 2). London: Academic Press.

Gazzaniga, M. (Ed.) (1995). *The cognitive neurosciences*. Cambridge, MA: MIT Press.

Hauser, E. (1996). *The evolution of communication*. Cambridge, MA: MIT Press.

Hawkins, R. D., & Kandel, E. R. (1984). Is there a cell-biological alphabet for simple forms of learning? *Psychological Review, 91,* 375–391.

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley

Heninger, G. R. (1995). *Neuroimmunology of stress*. In M. J. Friedman et al. (Eds.), *Neurobiological and clinical consequences of stress*. Philadelphia: Lippincott-Raven.

Hillyard, S., Mangun, F. R., Woldorff, M., & Luck, S. (1995). Neural systems mediating selective attention. In M. Gazzaniga (Ed.), *The cognitive neurosciences*. Cambridge, MA: MIT Press.

Hobson, A., & Stickgold, R. (1995). The conscious state paradigm: A neurocognitive approach to waking, sleeping and dreaming. In M. Gazzaniga (Ed.), *The cognitive neurosciences*. Cambridge, MA: MIT Press.

James, W. (1890). *The principles of psychology*. New York: Holt; reprint: Cambridge, MA: Harvard University Press (1983).

Jovanovic, U. J. (1971). *Normal sleep in man*. Stuttgart, Germany: Hippokrates.

Kaas, J. H. (1995). The reorganization of sensory and motor maps in adult mammals. In M. Gazzaniga (Ed.), *The cognitive neurosciences*. Cambridge, MA: MIT Press.

Koch, M., & Bubser, M. (1994). Deficient sensorimotor gating after 6-hydroxydopamine lesion of the rat medial prefrontal cortex is reversed by haloperidol. *European Journal of Neurosciences, 6,* 1837–1845.

Kolb, B., & Whishaw, I. Q. (1995). *Fundamentals of human neuropsychology* (4th ed.). New York: Freeman.

Lacey, J. I., & Lacey, B. C. (1970). Some autonomic-central nervous system interrelationships. In P. Black (Ed.), *Physiological correlates of emotion*. New York: Academic Press.

Lang, P., Melamed, B., & Hart, J. D. (1970). A psychophysiological analysis of fear modification using an automated desensitization procedure. *Journal of Abnormal Psychology, 76*(2), 220–234.

Lang, P., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review, 97,* 377–385.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1993). Emotion, arousal, valence and the startle reflex. In N. Birbaumer & A. Öhman (Eds.), *The structure of emotion*. Toronto, Canada: Hogrefe.

Lazarus, R. S., Kanner, A. D., & Folkman, S. (1980). Emotions: A cognitive phenomenological analysis. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research and experience* (Vol. 1). New York: Academic Press.

LeDoux, J. (1995). Emotion: Clues from the brain. *Annual Review of Psychology, 46,* 209–235.

Legg, C. R., & Booth, D. A. (Eds.) (1984). *Appetite: Neural and behavioral bases*. Oxford, UK: Oxford University Press.

LeVay, S. (1996). *Queer science: The use and abuse of research on homosexuality*. Cambridge, MA: MIT Press.

LeVay, S., Pattatucci, A. M. L., Patterson, C., Lili, D. W., Fulker, S. S., Cherny, S. S., Kruglyak, L., & Hamer, D. H. (1995). Linkage between sexual orientation and chromosome X928 in males but not in females. *Nature Genetics, 11,* 248–256.

Levy-Lahad, E., Wijsman, E., & Nemens, E. (1995). A familial Alzheimer's disease locus on chromosome 1. *Science, 269,* 970–973.

Linden, D. J., & Connor, J. A. (1995). Long-term synaptic depression. *Annual Review of Neuroscience, 18,* 319–357.

Lindsley, D. B., Bowden, J. W., & Magoun, H. W. (1949). Effect upon the EEG of acute injury to the brainstem activating system. *Electroencephalography and Clinical Neurophysiology, 1,* 475–486.

Lycan, W. (1996). *Consciousness and experience*. Cambridge, MA: MIT Press

Mackintosh, N. S. (1974). *Psychology of animal learning*. New York: Academic Press.

Matlin, M. (1983). *Cognition*. New York: Harcourt Brace.

McCarley, R. W. (1990). Brainstem cholinergic systems and models of REM sleep production. In J. Montplaisir & R. Godbout (Eds.), *Sleep and biological rhythms*. Oxford, UK: Oxford University Press.

McCormick, D. A., & Bal, T. (1997). Sleep and arousal: Thalamocortical mechanisms. *Annual Review of Neuroscience, 20,* 185–215.

Mesulam, M. M. (1995). Structure and function of cholinergic pathways. In F. Bloom & D. Kupfer (Eds.), *Psychopharmacology. The fourth generation*. New York: Raven Press.

Miller, N. E. (1969). Learning of visceral and glandular responses. *Science, 163,* 434–445.

Miller, N. E. (1978). Biofeedback and visceral learning. *Annual Review of Psychology, 29,* 373–404.

Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behavioural Brain Research, 6,* 57–77.

Money, J. (1987). Sin, sickness, or status. Homosexual gender identity and psychoneuroendocrinology. *American Psychologist, 42,* 384–399.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences, 13,* 201–288.

Obrist, P. A. (1981). *Cardiovascular psychophysiology: A perspective.* New York: Plenum.

Öhman, A. (1987). The psychophysiology of emotion: An evolutionary–cognitive perspective. In P. K. Ackles, J. R. Jennings & M. G. H. Coles (Eds.), *Advances in psychophysiology* (Vol. 2). Greenwich, CT:JAI Press.

Öhman, A., & Birbaumer, N. (1993). Psychophysiological and cognitive–clinical perspectives on emotion: Introduction and overview. In N. Birbaumer & A. Öhman (Eds.), *The structure of emotion.* Toronto, Canada: Hogrefe & Hubert.

Olds, J., & Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology, 47,* 419–427.

Patrick, C. J. (1994). Emotion and psychopathy: Startling new insights. *Psychophysiology, 31,* 319–330.

Pavlov, I. (1926). *Die höchste Nerventätigkeit (das Verhalten) von Tieren.* Munich: J. Bergmann.

Peterson, C., Maier, S. F., & Seligman, M. E. P. (1993). *Learned helplessness: A theory for the age of personal control.* New York: Oxford University Press.

Posner, M. I., & Raichle, M. E. (1995). *Images of mind.* New York: W. H. Freeman.

Rau, H., Pauli, P., Brody, S., Elbert, T., & Birbaumer, N. (1993). Baroreceptor stimulation alters cortical activity. *Psychophysiology, 30,* 322–325.

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist, 43,* 151–160.

Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craying: an incentive-sensitisation theory of addiction. *Brain Research Reviews, 18,* 247–291.

Rose, R. J. (1995). Genes and human behavior. *Annual Review of Psychology, 46,* 625–654.

Rosenzweig, M. R., & Leimann. A. L. (1982). *Physiological psychology.* Boston: Houghton Mifflin.

Schachter, S., & Singer, S. (1962). Cognitive, social and physiological determinates of emotional state. *Psychological Review, 69,* 379–399.

Schedlowski, M., & Tewes, U. (Eds.) (1996). *Psychoneuroimmunologie.* Heidelberg, Germany: Spektrum.

Schlaug, G., Jäncke, L., Huang, Y. & Steinmetz, H. (1995). *In vivo* evidence of structural brain asymmetry in musicians. *Science, 267,* 699–701.

Schulkin, R. (Ed.) (1993). *Hormonally induced changes in mind and brain.* New York: Academic Press.

Schupp, H., Lutzenberger, W., Birbaumer, N., Braun, C., & Miltner, W. (1994). Neurophysiological differences between perception and imagery. *Cognitive Brain Research, 2,* 77–86.

Schwartz, M. (Ed.) (1995). *Biofeedback* (2nd ed.). New York: Guilford Press.

Self, D. W., & Nestler, E. J. (1995). Molecular mechanisms of drug reinforcement and addiction. *Annual Review of Neuroscience, 18,* 463–495.

Siegel, G., Agranoff, B., Albers, W., & Molinoff, P. (Eds.) (1994). *Basic neurochemistry* (5th ed.). New York: Lippincott-Raven.

Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience, 18,* 555–586.

Sperry, R. W. (1952). The great cerebral commissure. *American Scientist, 40,* 291–312.

Sperry, R. W. (1964). The great cerebral commissure. *Scientific American, January.*

Springer, S. P., & Deutsch, G. (1993). *Left brain, right brain* (3rd ed.). New York: W. H. Freeman.

Steriade, M., & McCarley, R. W. (1990). *Brainstem control of Wake–Sleep states.* New York: Plenum.

Swaab, D. F., & Hofman, M. A. (1995). Sexual differentiation of the human hypothalamus in relation to gender and sexual orientation. *Trends in Neurosciences, 18,* 264–270.

Thompson, R. F., & Krupa, D. J. (1994). Organization of memory traces in the mammalian brain. *Annual Review of Neuroscience, 17,* 519–549.

Watson, J. D., & Crick, F. M. C. (1953). Molecular structure of nucleic acids. *Nature, 171,* 737–738.

Wever, R. A. (1979). *The circadian system of man.* New York: Springer.

Wise, R. (1988). The neurobiology of craving. *Journal of Abnormal Psychology, 97,* 118–132.

Wise, R. A. (1982). Neuroleptics and operant behavior: The anhedonia hypothesis. *Behavioral Brain Sciences, 5,* 39–87.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35,* 151–175.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.06

Fundamental Psychopharmacology

FRANK A. HOLLOWAY
University of Oklahoma Health Sciences Center,
Oklahoma City, OK, USA
and
JESSICA M. PEIRCE
Brookline, MA, USA

1.06.1 INTRODUCTION AND SCOPE 175

1.06.2 CNS NEUROPHARMACOLOGY 175
- *1.06.2.1 Neuroanatomy and Neurophysiology of Neurons* 175
- *1.06.2.2 Features of Neuronal Activity and Synaptic Neurotransmission* 176
  - *1.06.2.2.1 Membrane potentials and neural activation* 176
  - *1.06.2.2.2 Presynaptic events and modulation* 176
  - *1.06.2.2.3 Postsynaptic events and modulation* 177
- *1.06.2.3 Neural Receptors* 177

1.06.3 NEUROTRANSMITTERS AND NEUROMODULATORS 177
- *1.06.3.1 Acetylcholine (ACh)* 178
  - *1.06.3.1.1 Cholinergic neurotransmission and receptors* 178
  - *1.06.3.1.2 Cholinergic neuropharmacology* 178
  - *1.06.3.1.3 Cholinergic pathways* 178
- *1.06.3.2 Norepinephrine (NE) and Epinephrine (Epi)* 179
  - *1.06.3.2.1 Catecholaminergic neurotransmission and receptors* 179
  - *1.06.3.2.2 NE neuropharmacology* 179
  - *1.06.3.2.3 Noradrenergic pathways* 179
- *1.06.3.3 Dopamine (DA)* 180
  - *1.06.3.3.1 Dopaminergic neurotransmission* 180
  - *1.06.3.3.2 Dopamine pathways* 180
  - *1.06.3.3.3 Dopamine receptors* 180
  - *1.06.3.3.4 Dopamine neuropharmacology* 181
- *1.06.3.4 Serotonin (5-hydroxytryptamine, 5-HT)* 181
  - *1.06.3.4.1 Serotonin neurotransmission* 181
  - *1.06.3.4.2 5-HT neuropharmacology* 182
  - *1.06.3.4.3 5-HT pathways* 182
- *1.06.3.5 Amino Acid NTs* 182
  - *1.06.3.5.1 γ-Aminobutyric acid (GABA)* 182
  - *1.06.3.5.2 Glycine* 183
  - *1.06.3.5.3 Glutamate and aspartate* 183
- *1.06.3.6 Neuroactive Peptides and Hormones* 183
  - *1.06.3.6.1 Other neuromodulators* 184
- *1.06.3.7 The Neuropharmacology of Normal Behavioral Functions* 184

1.06.4 BASIC PHARMACOLOGY 184
- *1.06.4.1 Pharmacokinetics* 184

1.06.4.1.1 Drug administration and absorption 184
1.06.4.1.2 Drug distribution 185
1.06.4.1.3 Elimination 185
1.06.4.2 Pharmacodynamics 186
1.06.4.2.1 Receptor theory 186
1.06.4.2.2 Drug receptor properties 186
1.06.4.2.3 Types of drug-receptor interactions 186

1.06.5 PRINCIPLES OF BEHAVIORAL PHARMACOLOGY AND PSYCHOPHARMACOLOGY 186

1.06.5.1 Behavioral Drug Properties 186
1.06.5.1.1 Dose–effect curves (DEC) 186
1.06.5.1.2 Drug sensitivity and effectiveness 187
1.06.5.2 Drug Tolerance 187
1.06.5.2.1 Definitions 187
1.06.5.2.2 Kinds of tolerance 187
1.06.5.3 Sensitization 188
1.06.5.4 Physical Dependence 188
1.06.5.5 Basic Psychopharmacology Principles 188
1.06.5.5.1 Drug–behavior interactions 188
1.06.5.5.2 Baseline effects/rate dependency 188
1.06.5.5.3 Discriminative and state-dependent drug effects 188
1.06.5.5.4 Reward and reinforcing effects of drugs 189
1.06.5.6 Drug Classes 189
1.06.5.7 Drug Side-effects, Interactions, and Overdose 189
1.06.5.8 Drug Regulation 189

1.06.6 THE PSYCHOPHARMACOLOGY OF MENTAL DISORDERS 189

1.06.6.1 Mood Disorders: Depressive Disorders 190
1.06.6.1.1 Characteristics of depression disorders 190
1.06.6.1.2 The catecholamine (CA) hypotheses of depression 190
1.06.6.1.3 The serotonin (5-HT) hypothesis of depression 191
1.06.6.1.4 The cortisol hypothesis of depression 191
1.06.6.1.5 Status of depression theories 191
1.06.6.1.6 Drug therapies for depression 192
1.06.6.2 Mood Disorders: Bipolar Disorder 193
1.06.6.3 Anxiety Disorders: Generalized Anxiety Disorder 193
1.06.6.4 Anxiety Disorders: Panic Disorder 194
1.06.6.4.1 Phenomenology and etiology of panic disorder 194
1.06.6.4.2 Neuropharmacology of panic disorder 194
1.06.6.4.3 Drug therapies for panic disorder 194
1.06.6.5 Anxiety Disorders: Obsessive-compulsive Disorder 194
1.06.6.6 Anxiety Disorders: Post-traumatic Stress Disorder 195
1.06.6.7 Anxiety Disorders: Social Phobia 195
1.06.6.8 Psychotic Disorders: Schizophrenia 196
1.06.6.8.1 Characteristics of psychosis 196
1.06.6.8.2 The phenomenology and biological basis of schizophrenia 196
1.06.6.8.3 The DA hypothesis of schizophrenia 196
1.06.6.8.4 Antipsychotic medications 197
1.06.6.8.5 Antipsychotic drug side-effects 197
1.06.6.8.6 Side-effect profile of newer antipsychotics 197
1.06.6.9 Substance Abuse and Dependence Disorders 198
1.06.6.9.1 Definitions and diagnostic criteria 198
1.06.6.9.2 Models of addiction 198
1.06.6.9.3 Psychopharmacology of and pharmacotherapy for drugs of abuse 198
1.06.6.9.4 Alcoholism 199
1.06.6.9.5 Cocaine and amphetamine dependence 200
1.06.6.9.6 Nicotine and caffeine dependence 200
1.06.6.9.7 Opiate (narcotic) dependence 201
1.06.6.9.8 Sedative-hypnotic and anxiolytic dependence 201
1.06.6.9.9 Marijuana abuse 202
1.06.6.9.10 Abuse of hallucinogens and related drugs 202
1.06.6.9.11 Over-the-counter and prescribed drugs 203
1.06.6.10 Eating Disorders 203
1.06.6.11 Neurological Disorders 204
1.06.6.11.1 Parkinson's disease 204
1.06.6.11.2 Alzheimer's disease and other dementias 204
1.06.6.11.3 Huntington's disease (HD) 205
1.06.6.12 Developmental Disorders Originating in Childhood 205
1.06.6.12.1 Attention-deficit hyperactivity disorder (ADHD) 205
1.06.6.12.2 Other disorders 205

1.06.7 REFERENCES 205

1.06.1 INTRODUCTION AND SCOPE

This chapter is intended to give the well-educated reader an overview of how psychotropic drugs work to produce changes in behavior, mood, or perception. Psychotropic drugs fall into three classes, psychotomimetic (drugs inducing a psychosis), psychedelic (drugs producing a change in one's state of consciousness or mood), and psychotherapeutic (drugs relieving some biopsychological problem). The major psychotherapeutic drug classes are antipsychotics, anxiolytics (anxiety reducing), antidepressants, and stimulants. Accordingly, we will describe how and why health professionals use specific drugs to treat certain neuropsychiatric disorders. It is not our intention to give a comprehensive account of all the relevant topics. However, we will provide citations of current reference works where more technical information can be found. We will introduce the reader to neurophysiology, briefly describing how the targets of psychoactive drugs in the central nervous system (CNS) normally function. Next, we will examine the neuropharmacology of the major classes of neurotransmitters, whose different functions specific drugs can target, and then address the applied biology of drug action (pharmacology). Necessarily, we will discuss some principles of how psychoactive drugs affect behavior (behavioral pharmacology or psychopharmacology). Finally, we will review the psychobiology of neuropsychiatric disorders and briefly describe what medications have been found useful in the treatment of these disorders (clinical psychopharmacology).

To understand how drugs work in the CNS, it is necessary first to understand something about the structure (neuroanatomy) and function (neurophysiology/neuropharmacology) of neurons. The reader can refer to any of three excellent references for more in-depth coverage of topics in these areas (Cooper, Bloom, & Roth, 1996; Feldman, Meyer, & Quenzer, 1997; Kandel, Schwartz, & Jessell, 1991). Carlson (1994) also provides a brief, but excellent, glossary of neurobiological terminology.

The later sources will introduce the reader to key definitions, but a brief overview may be in order here. The key target of drugs is the nervous system, whose basic units are cells called neurons. As described below, neurons talk to one another chemically. The principal chemical signal is the neurotransmitter (NT), a molecule released by one neuron (say Cell "A"). In the simplest scheme, another neuron (say Cell "B") has receptors (again particular types of molecules) designed to detect and respond to the NT from Cell "A." As described below, a host of endogenous (other brain chemicals) and exogenous (external compounds like drugs) factors can influence the communication between neurons. Further, when the chemical messages are too strong or too weak, or when the reception is too sensitive or too poor, the normal neural, physiological, and behavioral functions may be disrupted. Here is where clinical psychopharmacology looks for drugs that will normalize the disrupted function of brain chemistry.

1.06.2 CNS NEUROPHARMACOLOGY

1.06.2.1 Neuroanatomy and Neurophysiology of Neurons

The CNS generally contains two types of cells: neurons and glial cells. Glial cells support the integrity and function of neurons, which serve the brain's sensory or input, integration, and motor or output functions. The principal targets of drugs in the CNS are the neurons, whose components are the cell body or soma, dendrites, and axons. The supporting glial cells are found around all parts of the neuron (astrocytes among axons and dendrites, oligodendrites near cell bodies, and Schwann cells surrounding peripheral axons).

Neurons have five basic features. First, neuron cell bodies contain the genetic material and apparatus necessary to synthesize or make new molecules (amino acids, proteins, peptides, etc.). Cell bodies also play a critical role in initiating a traveling wave of electrical signals (the action potential). Second, neurons generally talk with each other via special extensions from the cell body, called axons and dendrites. Axons form output or efferent paths from the cell body that convey rapid, spiked signals called the action potentials. Dendrites (sometimes called the dendritic tree) provide multibranched afferent or input paths to the cell body, conveying slower, graded potentials. Third, at the end or terminal of axons lies the synaptic junction or cleft, that is, the point where neurons make contact. An axon may synapse on dendrites, the cell body, other axons, or itself. The presynaptic side of the junction produces specialized molecules called neurotransmitters, packages them in vesicles, and usually releases them into the synaptic junction when the action potential of the presynaptic neuron arrives at the axon terminal. Fourth, on both the presynaptic and postsynaptic sides of the synaptic junction are receptors designed to detect and bind the neurotransmitter (NT), in a lock and key fashion. NT receptors are proteins (i.e., long chains of amino acids), which may reside outside the cell, within the cell membrane, and inside the cell. Receptor activation by an NT then leads to certain events or signals that ultimately result in some change in the presynaptic or postsynaptic neuron, for example, initiating its own action potential. Thus, communication between neurons is basically chemical, not electrical. The final feature is that all parts of the neuron, the cell body, axons, and dendrites, have special double-layered membranes separating the inside from the outside of the cell. In addition, the membrane is semipermeable. That is, it contains ion channels through which some ions (charged atoms, e.g., calcium, potassium, chloride) can pass passively in and out of the cell. Other ions or molecules (e.g., sodium) may require an active transport process involving a chemical energy process called an ion pump.

A neuron's structure or neuroanatomy also determines the ways it can communicate with other neurons. For example, monopolar neurons usually have only an axon (e.g., sensory dorsal root ganglion cells), while bipolar neurons have both dendrites and axons with a simple structure (e.g., granule cells of the retina, olfactory mucosa, and auditory nerve). On the other hand, multipolar neurons have large dendritic fields with one major axonal output. The latter can be short or long and straight or highly branched, contacting other neurons (including itself). The latter arrangements mean that even single neurons may participate in at least three very complex functions. For example, they may receive and integrate input or information from many sources. Next, they may produce similar effects in many different neuronal groups simultaneously (e.g., activation functions during arousal or alerting conditions). Finally, they may produce an organized array of different effects in different neuronal groups (e.g., functions related to induction of "states," like rapid eye movement sleep).

1.06.2.2 Features of Neuronal Activity and Synaptic Neurotransmission

Psychoactive drugs produce behavioral changes by altering the ways neurons normally function. We now turn to a brief description of bioelectric features of neuronal activity and transmission, namely, how neurons produce and detect the electrical or chemical signals just described.

1.06.2.2.1 Membrane potentials and neural activation

First, to generate neuronal activity some membrane potential must exist. The neural membrane is semipermeable, meaning that it contains ion channels through which some ions (charged atoms, e.g., calcium, potassium, chloride) can pass passively in and out of the cell. At resting baseline conditions, the neuron is polarized, that is, is more negatively charged inside the cell (low in sodium (Na^+) and high in potassium (K^+) ions) than it is outside the cell (high in Na^+ and low in K^+ ions). An active transport process (the Na^+ pump) maintains this steady state by extruding Na^+ from inside the cell through ion channels in the membrane. This resting state of affairs is called polarization. Certain neurotoxins work by altering the way ion channels normally function, for example, tetrodotoxin can block and tetraethylammonium can activate sodium channels.

Certain events (e.g., synaptic transmission) can produce changes in the membrane potentials of the postsynaptic neuron. In the discussions below, we arbitrarily designate Cell "A" as the presynaptic neuron relative to Cell "B," the postsynaptic neuron. If events started by postsynaptic NT–receptor interaction lead to depolarization of the membrane of Cell "A" (i.e., influx of sodium into the cell and a shift toward intracellular positivity), then excitatory postsynaptic potentials (EPSPs) are produced. If, however, the latter events lead to hyperpolarization of the membrane of Cell "A" (i.e., influx of chloride ions and/or outflow of potassium ions and a further intracellular shift toward negativity), then inhibitory postsynaptic potentials (IPSPs) are produced. The membrane potential of the postsynaptic neuron (Cell "A") results from the summation of EPSPs and IPSPs. Once the postsynaptic membrane reaches a threshold level of depolarization (usually at the junction of the cell body and axon), Cell "A" will "fire" (excitation). The latter event involves the propagation of an all-or-nothing action potential (spike) of marked depolarization through the axon to the neuron terminals. Ultimately, the arrival of this action potential at the terminal could result in synaptic transmitter release, which in turn can act on postsynaptic receptors of Cell "B." Again, however, if Cell "B" becomes hyperpolarized, then the generation of an action potential requires much more excitatory input.

1.06.2.2.2 Presynaptic events and modulation

Small spherical vesicles and other organelles involved in making or storing NTs are found in the presynaptic nerve terminals. Only some synaptic vesicles are bound to the presynaptic membrane. The arrival of the action potential at the presynaptic terminal, plus the copresence of calcium ions (Ca^{2+}), causes some vesicles to fuse with the membrane and releases packets of NT into the synaptic cleft. Generally, how much

NT gets released depends on the relative size of the presynaptic action potential and/or the presynaptic neuron's firing rate. The latter depend on the presynaptic membrane's baseline potential and activation of presynaptic receptors.

Two general kinds of presynaptic receptors exist. Autoreceptors are found on the axon terminal, the cell body, or the dendritic region of the same neuron that releases the NT. NTs from a different neuron activate heteroreceptors. Excitatory synaptic input to the presynaptic terminal produces presynaptic inhibition. Inhibition occurs because a depolarized presynaptic membrane potential will result in a smaller amplitude action potential and less NT release. Inhibitory synaptic input to the presynaptic terminal produces presynaptic facilitation. Here, the hyperpolarized presynaptic membrane potential will result in a larger amplitude action potential (and, perhaps, opening of calcium channels) and therefore, enhanced NT release. Again, modulatory factors include the presynaptic membrane's resting potential, adequate synthesis and storage of the NT, the copresence of neuromodulators (e.g., vasopressin) and CA^{2+}, the proportion of membrane-bound vesicles, and the active transport of the NT back into the neuron or into synaptic vesicles. Psychoactive drugs can act on any of these presynaptic processes, for example, calcium-channel blockers can lower NT release by reducing the availability of this essential ion.

The next issue is whether the NT will effectively activate the postsynaptic receptor. The persistence of NT in the synaptic cleft depends on the three ways of NT removal: its metabolism by extracellular enzymes, its reuptake into the presynaptic terminal (where the NT can undergo intracellular metabolism or be taken up into the vesicles), and its attachment of NT to a postsynaptic receptor. Again, drugs can produce changes in any of the latter processes.

1.06.2.2.3 Postsynaptic events and modulation

Events started by the interaction between Cell "A"'s presynaptic NT and the postsynaptic receptor of Cell "B" can lead to depolarization of the Cell "B" membrane, producing excitatory postsynaptic potentials (EPSPs). However, the latter events also may lead to hyperpolarization of the cell membrane, producing inhibitory postsynaptic potentials (IPSPs). The membrane potential of the postsynaptic neuron results from the summation of EPSPs and IPSPs and must reach a threshold of depolarization for that cell to fire.

Sources of modulation of postsynaptic receptors are more limited. Removal of one presynaptic input to the postsynaptic cell is insignificant, since thousands of neurons may synapse on a given postsynaptic cell. On the other hand, any direct modulation of the postsynaptic cell will alter all subsequent signal communication in that pathway. The postsynaptic cell's responsiveness to an NT can be changed in two principal ways. First, the number of postsynaptic receptors may chronically change. Down-regulation refers to a decrease in receptor sites with prolonged use and/or exposure to an agonist. Up-regulation refers to an increase in receptor sites with prolonged disuse and/or exposure to an antagonist. Here, the agonist or antagonist agent can be an endogenous substance or an exogenous drug. Second, changes in how tightly an NT binds to the receptor (i.e., affinity) can influence the effectiveness of the NT–receptor interaction. NTs, neuropeptides, neuromodulators, or drugs can alter affinity (see Section 1.06.2.3).

1.06.2.3 Neural Receptors

NT receptors, obvious targets for psychoactive drugs, reside outside the cell, within the cell membrane, and inside the cell. These receptors can be found anywhere on presynaptic and postsynaptic neurons. A receptor is a protein membrane site to which a molecule, the ligand, or the NT, binds. Proteins are long peptides (two or more amino acids linked by peptide bonds) that serve in a structural capacity or as enzymes. When the ligand normally activates the receptor, we call it an agonist. An antagonist refers to a ligand that blocks or prevents the receptor's normal activation by an agonist. Ligands described as inverse agonists have receptor effects opposite to those produced by agonists, for example, closing rather than opening an ion channel. Partial agonists and partial inverse agonists have the same but weaker effects as full agonists or inverse agonists. NTs may produce depolarization (postsynaptic excitatory and presynaptic inhibitory synapses) or hyperpolarization (postsynaptic inhibitory or presynaptic excitatory synapses). Here, whether the postsynaptic neuron is more or less likely to be activated defines excitation and inhibition. Research has identified two major receptor classes. Iontotropic receptors are ion-gated and rapid acting, while metabotropic receptors depend on second-messenger processes and are slower acting.

1.06.3 NEUROTRANSMITTERS AND NEUROMODULATORS

Besides the reference works cited above, the interested reader can also refer to three other resources (Bloom & Kupfert, 1995; Hardman & Limbird, 1996; Schatzberg & Nemeroff, 1995). The classical definition of an NT typically involves the sequences of events described in the previous section. That is, the transmitter may be synthesized in the cell body and transported to the axon terminal or may be synthesized in the axon terminal; in either case, it is stored in the terminal. Presynaptic nerve firing releases the transmitter into the synaptic cleft, where the transmitter can produce some variety of synaptic membrane potential change in presynaptic or postsynaptic receptors. Finally, the transmitter is removed from the synaptic junction by some mechanism. Neuromodulation refers to how each step in the synaptic transmission process can be affected by endogenous substances or drugs. This section will review the major classes of CNS NTs. Where relevant, we will mention the peripheral nervous system (PNS), which also may use a particular NT that is also found in the CNS. For each major NT, we briefly introduce the reader to the NT's functions and then provide a short description of its synthesis and metabolism, its neuropharmacological action (i.e., depolarization or hyperpolarization), the effects of agonist or antagonist drugs on its functioning, and the localization/function of neural pathways that use the NT.

1.06.3.1 Acetylcholine (ACh)

ACh is an NT found in brain, spinal cord, nerve endings in muscle, ganglia of the autonomic nervous system (ANS), and postganglionic terminals of the parasympathetic division of the ANS. As described below, this transmitter plays a major CNS role in arousal, memory, and other functions. In the periphery, besides its role in parasympathetic function, it is the major transmitter innervating somatic muscles. Not surprisingly, peripheral dysfunction produces motor impairment and central dysfunction can produce both alterations in states of consciousness (e.g., sleep) and memory function (see discussion in this section and in Section 1.06.6.11).

1.06.3.1.1 Cholinergic neurotransmission and receptors

Cholinergic neurons release ACh from axon terminals. ACh is synthesized from choline, which is transported into neurons. The enzyme acetylcholinesterase metabolizes ACh Importantly, anticholinesterases, such as those found in insecticides (e.g., physostigmine-reversible) or nerve gases (soman-irreversible), result in increased ACh levels. Two major classes of ACh receptors exist, nicotinic and muscarinic. The muscarinic receptors are metabotropic and can lead to either depolarization or hyperpolarization (for subtypes, see Cooper et al., 1996).

1.06.3.1.2 Cholinergic neuropharmacology

Pharmacological intervention can occur at most of the steps in cholinergic neurotransmission. The neurotoxin NVP blocks ACh synthesis and vesamicol, its transport into synaptic vesicles. Black widow spider venom promotes ACh release, while botulinum toxin and other substances inhibit its release. Rabies and curare block nicotinic receptors, and atropine and pirenzepine block muscarinic receptors. Cholinergic neurotransmission sometimes involves negative feedback onto presynaptic muscarinic receptors which, when blocked with atropine, results in increased release of ACh.

1.06.3.1.3 Cholinergic pathways

ACh also serves as the NT at the neuromuscular junction and at synapses in spinal cord autonomic ganglia and the parasympathetic autonomic nervous system. Myasthenia gravis, an autoimmune disease associated with progressive loss of muscle tone, results from damage and antagonism of the ACh receptor. Within the brain, cholinergic neurons may be short regional interneurons (e.g., nucleus accumbens, caudate-putamen) or long projection neurons connecting several brain areas. The most common degenerative neurological disorders involve some loss of normal cholinergic function, for example, Huntington's chorea (striatum) and Alzheimer's disease (nucleus basalis, neocortex, and hippocampus). Unfortunately, attempts to promote ACh activity have not worked well (choline, anticholinesterase, ACh-receptor agonists).

The two main cholinergic projection pathways are the basal forebrain system and the pons-midbrain system. Cholinergic neurons with cell bodies in the basal forebrain send axons to most of the nonstriatal telencephalon. The important higher cognitive functions include attention and memory. Axons of cholinergic neurons with cell bodies in the midbrain and pons (the reticular activating system) bifurcate. They send projections both rostral (up) and caudal (down) to affect wide areas of the brain. These pathways serve both activating and integrative functions, promoting arousal, waking, and even rapid eye movement sleep.

1.06.3.2 Norepinephrine (NE) and Epinephrine (Epi)

NE is a major NT in the brain and in the postganglionic terminal endings of the sympathetic division of the ANS (notably the cardiovascular system). NE, along with Epi, also serves as a hormone (release from the adrenal medulla) involved in a host of peripheral ANS functions (notable cardiovascular activity). In the CNS, as described below, NE is the NT for cells in the brainstem which serve arousal functions involved in attention, mood, and waking. Accordingly, dysfunctions in these CNS systems have been implicated in mood and anxiety disorders (see Section 1.06.6).

1.06.3.2.1 Catecholaminergic neurotransmission and receptors

Three NTs make up the group called the catecholamines (CA)—NE, epinephrine (Epi), and dopamine (DA)—so-called because of the common catechol chemical structure. Adrenergic NTs include just NE and Epi and biogenic amine NTs include CA-NTs plus indole amine NTs (e.g., serotonin). The primary synthesis pathway for all three CA-NTs starts with the amino acid tyrosine that undergoes enzymatic conversion to DOPA (3,4-dihydroxyphenylalanine). From DOPA, the synthesis path leads first to DA, then to NE (via DA-β-hydroxylase), and finally to Epi. Monoamine oxidase (MAO) primarily degrades CAs inside the presynaptic terminal, while catechol-*O*-methyltransferase (COMT) metabolizes CAs in the synaptic junction. However, most of the CA-NT is taken up into the presynaptic cell and restored in vesicles. Many factors affect the rate of CA synthesis, including the concentration of the NT itself (end-product inhibition). The rate and pattern of CA neuron firing can increase its synthesis and turnover or metabolism. A major metabolite of central NE is 3-methoxy-4-hydroxy-phenethyleneglycol (MHPG), measured in blood or urine as an indicator of central NE production.

In mammals, NE is the major NT for the peripheral sympathetic nervous system (PNS) and the CNS, with lesser concentrations of Epi found in both CNS and PNS. Noradrenergic, adrenergic, and dopaminergic neurons respectively use NE, Epi, or DA as their NT. NE and Epi appear to act at one or more of four kinds of adrenergic receptors, located presynaptically and/or postsynaptically: α_1 (a postsynaptic receptor in the brain and in the periphery, where activation contracts smooth muscle), α_2 (located in the brain and in the periphery, where it is a presynaptic receptor modulating NE release), β_1 (located in the brain and in the periphery, where activation produces cardiac stimulation) and β_2 (located in the brain and in the periphery, where activation relaxes smooth muscles). Adrenergic receptors in the periphery (adrenal gland, heart, blood vessels, etc.) are called adrenoceptors. NE and Epi are equally effective at β_1-adrenoceptors, but Epi is more effective than NE at β_2-adrenoceptors. CA receptors on non-CA neurons are generally postsynaptic receptors, with presynaptic autoreceptors located in the CA neuron itself.

CA-NTs acting on autoreceptors regulate the release of all CAs. Activation of presynaptic α_2-adrenergic, muscarinic, opiate and DA autoreceptors produces inhibition of NE release, while presynaptic β_2-adrenergic, nicotinic, and angiotensin II receptor activation results in facilitation of NE release. Further, prolonged exposure to any of the CA or agonist-like drugs (e.g., reuptake blockers) can lead to a subsensitivity or desensitization of the receptor and even a decrease in the number of receptors (down-regulation). This reduction in responsiveness to stimuli, inputs, or agents can be nonspecific (many agents) or specific (one agent). On the other hand, prolonged reduced exposure to any of the CA receptors to the NT can lead to a supersensitivity of the CA receptor to agonist agents and even an increase in the number of receptors (up-regulation). The latter situation could arise from destruction of CA neurons or from a reduced synthesis or increased metabolism of CA-NTs. Essentially, drugs acting on the CA systems can affect both cardiovascular and a host of other neural activities.

1.06.3.2.2 NE neuropharmacology

In PNS and CNS NE neurons, drugs can affect adrenergic activity in at least four general ways. Drugs can act: (i) presynaptically to block NE synthesis (α-methyltyrosine, Fla-63), storage (reserpine, tetrabenazine), release (amphetamine), and reuptake (amphetamine, cocaine, desipramine); (ii) postsynaptically to block (phenoxybenzamine, phentolamine) or mimic (clonidine) the action of the NT on postsynaptic receptors; (iii) on autoreceptors to indirectly influence catecholamine activity (α_2-agonist—clonidine; α_2-antagonists—yohimbine, piperoxan); and (iv) to inhibit the metabolism of NE in the synaptic cleft (pargyline) or presynaptic terminal (tropolone).

1.06.3.2.3 Noradrenergic pathways

Since late 1960s, fluorescence histochemical and other methods have permitted a quite detailed picture of catecholamine pathways. The *locus coeruleus* (LC) comprises the main brainstem area from which axons of NE neurons arise. These neurons act to inhibit cells on which they synapse by way of the slower metabotropic-receptor hypopolarization process. Cell bodies of the large LC neurons are found in the caudal pontine central grey. LC axons typically form long and extensively branched projections over five major pathways to widely separated areas of the brain. Their targets include all the cortices (probably β-adrenergic receptors) and all the specific thalamic and hypothalamic nuclei. The projection patterns of central adrenergic neurons suggest some general action on most of the forebrain and on autonomic activity. Empirical data showing increased LC responsiveness to novel events argues for some role in vigilance, while other data suggest some role in stress-related processes. The potential functions for LC adrenergic neurons are large, for example, sleep and waking, reinforcement, memory formation, affective functions and dysfunctions, and anxiety–pain interactions (see Bloom & Kupfer, 1995).

1.06.3.3 Dopamine (DA)

Dopamine, also a catecholamine, is an NT involved in three major systems associated with activation of motoric, motivational (reward), and learning processes. These systems are described below. Dysfunction or alteration of any one of these systems can have profound consequences, including integrated movements and cognitions, including persistent and abnormal ideations, movements, and behaviors (see Sections 1.06.6.8, 1.06.6.9, and 1.06.6.11.

1.06.3.3.1 Dopaminergic neurotransmission

Neurons using DA as their NT have both similarities (synthesis and metabolism) and differences (organizational complexity) relative to adrenergic neurons. It has already been stated that the synthesis or production of DA is just like that for NE and Epi. What is different in the DA neuron is the absence of dopamine β-hydroxylase, which would convert DA to NE. In the presynaptic cell, MAO converts DA to dihydroxyphenylacetic acid (DOPAC). After release from the presynaptic terminal, COMT converts DA to 3-MT (3-methoxytyramine) and MAO converts 3-MT to homovanillic acid (HVA). In humans, HVA, the primary DA metabolite, provides a good indicator of DA activity when measured in cerebral spinal fluid. The regulation of DA levels is similar to that for NE. DA release depends on the presence of calcium ions presynaptically and on the rate and pattern of firing in the presynaptic cell, with burst-firing producing enhanced DA release. Since DA neurons have presynaptic inhibitory autoreceptors, DA agonists inhibit, while DA antagonists enhance DA release. Like NE and Epi, the major way DA is removed from the synaptic cleft is by reuptake into the presynaptic neuron. This process involves a DA transporter. We now know that several DA-specific compounds can selectively block this reuptake process to enhance DA activity. Cocaine and amphetamine are examples of nonspecific reuptake blockers.

1.06.3.3.2 Dopamine pathways

Dopamine neuron cell bodies, largely found in the midbrain and hypothalamus, are three to four times more numerous than NE neurons. Middle-length DA pathways connect areas of the hypothalamus with the pituitary (involved in prolactin secretion), other areas of the hypothalamus, and limbic system and areas of the medulla with other brainstem regions. The long-length DA neurons arise from cell bodies in the ventral tegmental (VTA) and substantia nigra (SN, nigrostriatal system) areas of the midbrain. SN-DA neurons project to the basal ganglia (neostriatum: caudate and putamen areas) and serve to activate and organize motor functions (damage and/or dysfunction of these cells is found in Parkinson's disease). Like most DA projection systems, the nigrostriatal system contains interactive, reciprocal connections, so that innervation of striatal neurons results in a decrease in SN-DA activity. VTA-DA neurons form two projection systems, the mesocortical system projecting to limbic cortical areas and the mesolimbic system projecting to deeper limbic areas (e.g., nucleus accumbens, and the amygdala). Generally, the mesocortical and mesolimbic systems serve affective and motivational functions. These brain areas appear responsive to positive and negative events or states (stress, conditioned fear, pain and pleasure). Further, they modulate the consequences of such events (affective states like anxiety and "cognitive" processes like reinforcement; see Bloom & Kupfer, 1995; Wise, 1990). The DA projections largely function to inhibit their targets.

1.06.3.3.3 Dopamine receptors

Postsynaptic targets use both of the two major DA receptor subtypes, D_1 and D_2. DA-autoreceptors (located on all parts of the DA neuron) typically are D_2 receptors. Activation of autoreceptors on the cell body or dendrites slow the firing rate of the DA cell, while those at the axon terminal can inhibit DA release or synthesis. Not surprisingly, autoreceptors located far from the cell body synthesis apparatus (e.g., prefrontal or cingulate cortices) only possess the inhibition of release property. Both D_1 and D_2 receptors are metabotropic, involving the second messenger adenylate cyclase, but D_1 activation stimulates while D_2 activation inhibits adenylate cyclase. Further, D_2 receptors are five to ten times more sensitive than D_1 receptors to DA (or DA agonists like apomorphine). Thus, low doses of DA agonists will first affect these autoreceptors, resulting in a net reduction in DA function. Higher doses, of course, will override this effect and stimulate postsynaptic receptors. Nevertheless, postsynaptically, a maximal DA effect on behavior or brain activity may require activation of both D_1 and D_2 receptors. All DA receptors can develop supersensitivity and desensitization with changes in the chronic level of DA activity. So in Parkinson's disease, the loss of SN input produces an increase in striatal D_1 and D_2 receptor density. Increases in D_2 receptor density also are found in the brains of postmortem schizophrenics. However, the presynaptic autoreceptors appear more likely to show desensitization. This means that chronic DA-mediated drug therapy may fluctuate in its efficacy.

Two other DA receptors are subtypes of the D_2 receptor. D_3 receptors are densest in limbic brain areas and have a much higher affinity for specific DA agonists. In primates, the density of a D_4 receptor subtype is highest in the frontal cortex, midbrain, amygdala, and medulla and lowest in the basal ganglia. Of interest, the D_3 and D_4 subtypes have a higher affinity for the atypical antipsychotic Clozapine, which has few basal ganglia side-effects.

1.06.3.3.4 Dopamine neuropharmacology

The same drugs affect NE and DA synthesis, storage, release, reuptake, and metabolism. All CAs produce end-product inhibition of tyrosine hydroxylase (TH) which converts tyrosine to DOPA. Increased neuronal firing changes TH from a low-affinity to a high-affinity form and increases speed of this conversion. α-methyltyrosine inhibits TH, thereby reducing the production of DOPA. DOPA is very rapidly metabolized, so administering L-DOPA dramatically increases the production of DA. DA end-product inhibition of DA synthesis and release also occurs when DA binds with presynaptic receptors. DA agonists (apomorphine) can inhibit synthesis and release of DA, and DA antagonists (antipsychotics like haloperidol) can enhance them. However, in certain pathways (SN), but not others (mesolimbic and mesocortical), tolerance to antipsychotics occurs. A return to baseline levels of DA activity, release, and metabolism after long-term treatment illustrates the latter effect.

Various antipsychotics, such as phenothiazenes (e.g., chlorpromazine), thioxanthines, and butyrophenones act as DA antagonists. The resulting reduction of DA function produces extrapyramidal (i.e., found outside the primary pyramidal motor system) motor side-effects and the alleviation of psychotic symptoms. However, these antipsychotics also can increase DA turnover, presumably by blocking autoreceptors, which would normally decrease DA function. On the other hand, several drugs can enhance DA activity. Amphetamine and cocaine block DA reuptake and amphetamine may induce the release of DA. At high doses, amphetamine and cocaine will produce a psychotic state that is reversible with antipsychotics (this has suggested a pharmacological model for schizophrenia). Nomifensine and benztropine, an anticholinergic drug, also block DA reuptake. Pargyline enhances DA function by MAO inhibition (MAOI). MAO would normally degrade intraneuronal DA. All antidepressive drugs that are MAOIs will increase the amount of DA available for release. Tropolone, which inhibits COMT, can block the degradation of extraneuronal DA.

1.06.3.4 Serotonin (5-hydroxytryptamine, 5-HT)

Serotonin, an indole amine like melatonin, functions as an NT for cells residing in the brainstem. These cells project down into the spinal cord and upwards into many areas of the brain. 5-HT also functions as an NT in the periphery (gut and blood vessels). 5-HT, as described below, apparently has sometimes subtle, regulatory functions for mood, sleep/waking, and motivated behaviors. Clinically, drugs that enhance 5-HT levels have antidepressive effect and may stabilize excessive behavior patterns (see Sections 1.06.6.1 and 1.06.6.5).

1.06.3.4.1 Serotonin neurotransmission

Serotonin (5-HT) synthesis starts with the essential amino acid tryptophan. Tryptophan is converted to 5-hydroxytryptophan, which then is converted into 5-HT. As an interesting side note, 5-HT in the pineal body is converted to melatonin. MAO breaks 5-HT down into 5-hydroxyindoleacetaldehyde, which is further reduced to 5-hydroxyindoleacetic acid (5-HIAA).

Unlike the CAs, 5-HT and 5-HIAA do not inhibit the synthesis of 5-HT and its metabolites. However, blocking their catabolism results in linear increases in both. Several avenues can regulate the production and release of serotonin, including diet.

1.06.3.4.2 5-HT neuropharmacology

5-HT is primarily found in peripheral tissues, with only 1–2% found in the brain. Like the CAs, 5-HT binds to many different receptors, but unlike the CAs, different 5-HT receptors mediate very different functions. We know of four 5-HT$_1$ receptors, four 5-HT$_2$ receptors, and five 5-HT$_{3-7}$ receptors. The most well-researched receptor is the 5-HT$_{1a}$ receptor. Some psychiatrists use buspirone, a possible selective antagonist for this receptor, to treat many anxiety disorders. Ondansetron and granisetron, treatments for chemotherapy-induced nausea and vomiting, antagonize the 5-HT$_3$ receptor. The atypical antipsychotic Clozapine antagonizes 5-HT$_6$ and 5-HT$_7$ receptors, and amitryptaline, a tricyclic antidepressive, also antagonizes the 5-HT$_7$ receptor. 5-HT functions primarily as an inhibitory NT, causing a decrease in tonic firing. 5-HT activity is highest in states of physiological arousal, lowest in sleep, and completely absent in REM sleep. Therefore, an association between increased motor activity and increased 5-HT may exist. This apparently paradoxical action may occur by increased inhibition of inhibitory neurons.

The administration of *p*-chlorophenylalanine hampers enzymatic conversion, competing with tryptophan for access to the enzyme. After synthesis, 5-HT is stored in and released from presynaptic vesicles, just like the CAs. As a result, reserpine will interfere with the storage mechanism of 5-HT as well. The recreational drug lysergic acid diethylamide (LSD) is a partial agonist on presynaptic and postsynaptic 5-HT receptors. It functions to reduce 5-HT release. As mentioned above, buspirone is a 5-HT$_{1a}$ antagonist, which causes reduced binding of 5-HT. MAO inhibitors (MAOI) slow the degradation of 5-HT in the cell, allowing more storage of 5-HT in vesicles. Since the 1980s, fluoxetine, fluvoxamine, sertraline, and paroxetine have been found to act as selective serotonin reuptake inhibitors (SSRI). They serve to increase the amount of 5-HT in the synaptic junction.

1.06.3.4.3 5-HT pathways

5-HT-containing neurons cluster in the raphe nuclei, near the midline of the upper brainstem. Some nuclei project caudally to the medulla and spinal cord, and others project rostrally to the limbic system, striatum, thalamus, and cerebral cortex. Research has discovered two different types of neurons: fat beaded axons from the median raphe and fine axons from the dorsal raphe. Research reveals that the fine axons are more sensitive to MDMA, taken recreationally under the name "Ecstasy." MDMA produces a profound loss of these axons that lasts for at least months, if not forever. MDMA users have a measurable loss of function of 5-HT metabolism and catabolism.

1.06.3.5 Amino Acid NTs

Certain amino acids have also been found in the brain and appear to act as NTs. Strictly speaking, however, many amino acid NTs have not yet met all the criteria for being called NTs. Despite this, they are the most widespread of all the NTs. CAs and acetylcholine NTs are better researched, but account for only a small part of neurotransmission in the brain. γ-aminobutyric acid (GABA) and glycine are the primary, inhibitory NTs and glutamate and aspartate are the excitatory NTs. Both excitatory (NMDA) and inhibitory amino acid NT mechanisms appear to function in one or more components of learning and memory processes.

1.06.3.5.1 γ-Aminobutyric acid (GABA)

Of the two inhibitory amino acid NTs, GABA has been the subject of more research, possibly because of its primary location in the CNS. GABA's particular interest lies in the benzodiazepine (BZD) receptor, which is part of the receptor complex. Activation of this receptor produces anxiolytic (anxiety-reducing) effects (see Section 1.06.6 on anxiety disorders). Also of interest is the fact that the compounds which block activation of GABA or BZD receptors can produce proconvulsant effects (i.e., increased likelihood of seizures).

The synthesis of GABA begins with glutamine, which is then converted to glutamate (an excitatory NT, also known as glutamic acid). Glutamate is then converted to GABA by glutamic acid decarboxylase (GAD). GABA-transaminase (GABA-T) metabolizes GABA. Both GAD and GABA-T require pyridoxine (a form of vitamin B$_6$) as a coenzyme, although in differing affinities. As a result, pyridoxine deficiency causes an increased susceptibility to seizures due to the reduced production of GABA. Pyridoxine deficiency occurs most commonly in infants with a vitamin B$_6$ deficiency, and the resulting seizures can be fatal.

Research identifies two classes of GABA receptors: GABA$_A$ and GABA$_B$. GABA$_A$ has a gated chloride (Cl^-) channel, with binding sites for GABA, BZDs, barbiturates, steroids, and picrotoxin. BZDs increase the frequency of opening of the Cl^- channel, and barbiturates increase the duration of channel opening, thus potentiating GABA's inhibition of neuronal activity. Both ethanol and certain neurosteroids also appear to increase GABA activity. Inverse agonists like beta-carbolines also act at the BZD receptor site but close the Cl^- channel. The latter effect is similar to that produced by proconvulsants such as pentylenetetrazol (metrazol), which act at the picrotoxin site. GABA$_B$ receptors are less common than GABA$_A$ receptors in the CNS and are not linked to Cl^- channels or affected by BZDs or barbiturates. Baclofen is a well-known agonist for GABA$_B$ receptors.

As mentioned above, BZDs and barbiturates potentiate the effects of GABA at postsynaptic receptors, which may explain their common anxiolytic action. Flumazenil, a benzodiazepine antagonist, can block this effect. The GABA agonist muscimol will activate GABA$_A$ postsynaptic receptors, and bicuculline and picrotoxin will block postsynaptic GABA receptors. One can also inhibit degradation of GABA with gabaculine, which is an inhibitor of GABA-T. GABA travels in two long pathways and in short interneurons of the cortex, hippocampus, and spinal cord. GABA is the primary NT in striatal neurons that project to the substantia nigra, which is the brain site with the highest concentration of GABA.

1.06.3.5.2 Glycine

Glycine is very widespread throughout the human body, in part because it is a necessary component of many metabolic pathways. However, research has discovered evidence that it also functions as an NT in inhibitory interneurons of the spinal cord, lower brain stem, and possibly the retina. Strychnine appears to block the inhibitory effects of glycine.

1.06.3.5.3 Glutamate and aspartate

Glutamate and aspartate occur in high concentrations in the brain and exert powerful excitation on neurons. Thus, it is likely that these compounds act as excitatory NTs. Part of their function may involve the establishment of new functional connections between neurons, but dysregulation of their receptors also may produce hyperexcitability, including seizures and even cell death. Each compound also serves in metabolic pathways, and glutamate is involved in GABA synthesis. Thus, differentiating between these functions and any NT activity the compounds may have is difficult. However, research has discovered five different types of excitatory amino acid receptors, the best-known of which is the NMDA receptor. The NMDA receptor has binding sites for both glutamate and a particular kind of glycine molecule. Phencyclidine (PCP) is an NMDA channel blocker and a drug of abuse. NMDA receptors appear to affect memory and developmental plasticity, having high densities in hippocampal and cortical areas.

1.06.3.6 Neuroactive Peptides and Hormones

Neuroactive peptides are protein fragments, cleaved from larger compounds. Peptides, known mostly to coexist with other NTs, enrich the effect of the classical NT. For example, a peptide may alter the rate of production of another NT or strengthen or prolong the effect of an NT. Hormones are substances that are released from one location and travel to another location to act on receptors there. These compounds may act both on CNS and peripheral organ systems.

Opioid peptides, also known as endorphins, mimic endogenous morphine-like substances that have antinociceptive (i.e., antipain) properties. The class of opioid peptides includes proopiomelanocortin (POMC) peptides (e.g., β-endorphin, α-MSH, ACTH), proenkephalin peptides (e.g., met- & leu-enkephalin), and prodynorphin peptides (e.g., dynorphin A & B). The enkephalin peptides may act as NTS (Carlson, 1994). The four major kinds of opiate receptors are mu, kappa, delta, and sigma. Kappa receptors, largely located in the dorsal horn of the spinal cord, induce analgesia by depressing the initial site of pain transmission. The mu receptors, primarily found in the medial thalamus and brainstem, mediate morphine-induced analgesia and morphine's sedative, respiratory depressant, miotic, euphoric, and physical dependence properties. Delta receptors largely overlap both mu receptor location and, to a lesser extent, its effects. Activation of the opiate sigma receptors, primarily found in the limbic system (hippocampus and amygdala) can result in psychotomimetic, hallucinatory, and dysphoric responses.

Several other neuroactive peptides are equally well-researched. Two, released from the posterior pituitary, vasopressin (AVP, antidiuretic hormone) and oxytocin, respectively, causes the kidneys to retain more water and cause contractions of smooth muscles of milk ducts and the male and female sexual systems. Substance P, a tachykinin, is thought to act as a sensory (pain) NT or modulator. Neurotensin, colocalized in some DA and NE neurons, can intensify the actions of those NTS. Glucagon-related peptides, like vasoactive intestinal polypeptide (VIP), colocalize with ACh. in parasympathetic systems, particularly those of the gut. Pancreatic-related polypeptides, like neuropeptide Y, colocalize with the NTs, NE and Epi, affecting the action of these NTs in the brain (e.g., limbic system) and smooth muscle in the periphery. Some hormones are peptides, like AVP, oxytocin, β-endorphin, and ACTH, but others are steroids (small fat soluble molecules), like cortisol, corticotropin-releasing hormone (CRH), and sex hormones. Neuroactive peptides and hormones are occasionally, but rarely, administered exogenously. These peptides may mediate psychological or neurological disorders or modulate CNS drug actions (see Cooper et al., 1996).

1.06.3.6.1 Other neuromodulators

Clearly, there are a host of other neuropeptides and neuromodulatory compounds not described in the above discussion. Neuropeptides include those involved in regulatory processes related to water balance and thirst (e.g., angiotensin II), to nutritional balance, to growth, and/or to gastrointestinal function (e.g., somatostatin, cholecystokinin). Other neuromodulators include adenosine (which is found at most synapses) and a large class of second messengers. The second messenger compounds (including cyclic AMP and GMP, calcium/calmodulin, phosphoinositide, and eicosanoids) serve to mediate the initial reaction produced by NT-receptor binding by activating intracellular enzymes called protein kinases.

1.06.3.7 The Neuropharmacology of Normal Behavioral Functions

Many books devote considerable space to the description of the neuropharmacology of psychological or behaviorally related physiological functions (see Carlson, 1994). Chapter 1.05 also discusses some of this information. However, a very brief mention of one function should serve to illustrate the multiple involvement of the various NT and neuromodulatory systems. Reciprocal inhibitory interactions among cholinergic (giant-celled, lateral reticular area of the pons), serotonergic (raphe nuclei in brainstem), and noradrenergic (locus coeruleus) areas of the brainstem have been implicated in the regulation of sleep and waking and in the various sleep disorder syndromes. Further, the neuromodulator adenosine also may act to inhibit cholinergic activation of the frontal cortex by the midbrain reticular formation. The latter reticular effect thought to be important in arousal functions and in the present context is interesting since caffeine and other xanthene compounds inhibit adenosine's actions.

1.06.4 BASIC PHARMACOLOGY

This section reviews three topics relevant to a general understanding of psychopharmacology. Pharmacokinetics refer to the principles and processes governing how the body absorbs, distributes, and eliminates drugs. Pharmacodynamics define the ways in which targets of drugs (i.e., drug receptors) behave. The subsection on interaction between drugs and receptors begins to introduce some functional ways in which drugs can affect the CNS. Pertinent reference works for the latter two topics include those by Smith and Darlington (1996), Hardman and Limbird (1996), Dyskstra (1992), Johanson (1992), and Julien (1997).

1.06.4.1 Pharmacokinetics

1.06.4.1.1 Drug administration and absorption

Bioavailability refers to that fraction of unchanged drug reaching systemic circulation, whatever the route of administration. Drug absorption depends on the drug dose, the drug formulation and the route of administration. Selection of a drug dose, usually calculated in drug weight per body weight, depends on many factors. The initial therapeutic window defines the drug concentration in the body that produces the desired effect without intolerable side-effects. Development of tolerance (requirement of a higher dose for the same effect) or sensitization (requirement of a lower dose for the same effect), respectively, might call for the upward or downward adjustment of the dose. Incomplete absorption or reabsorption of the drug from the bloodstream into the body tissue can occur before or after first-pass metabolism by the liver, thereby reducing the drug concentration in the blood. The time that the drug concentration in the blood is above the minimum effective concentration determines the intensity duration of the drug effect.

Drug formulation or preparation (i.e., capsule, tablet, liquid, injection, vapor) depends on the route of administration and absorption. The latter include administering the drug by oral intake through the mouth, by insertion into the rectum, by injection (parenterally through the skin), by inhalation through the lungs, or by application and absorption through mucous membranes (nose, mouth, etc.).

Oral drugs must be soluble and stable in the stomach and permeable across stomach or intestinal lining to enter the bloodstream (factors include stomach/intestinal acidity or pH and the drug's lipid or water-soluble properties (see below).

Parenteral administration requires that the drug (*drogue*, French for "a dry powder") be dissolved or suspended in a vehicle, usually physiological or normal saline (0.9% NaCl, table salt in water). Potential problems arising with drug injections include rapid absorption, the necessity of sterile technique, and lack of "recallability." Absorption is not a factor with intravenous (IV, into the vein) injections, where the effect can be immediate. Licensed health professionals find this method useful for emergencies, dose control, and large volumes and/or diluted irritating agents. However, the IV route involves potential increased risks for untoward effects, requires slower injection rates, and cannot be used for oily or insoluble agents.

Intramuscular (IM, into the muscle) and subcutaneous (SC, under the skin) injections also can produce an immediate response with aqueous substances, with sustained drug blood levels. IM injections can prove useful for the administration of moderate volumes and of oily and some irritating agents. However, they cannot be used with anticoagulant drugs and can interfere with some medical assays. SC injections can be useful for insoluble agents or pellets, but cannot be used for large volumes and can produce pain and necrosis. Intraperitoneal (IP, into abdominal or peritoneal cavity) injections are rarely used in humans.

Rectal drug administration may be expedient but absorption is variable. Drugs absorbed through mucus membranes (nose, mouth, rectum) generally enter the blood steam directly (e.g., nicotine gum, nasal decongestants). Inhalation of gaseous drugs produces rapid absorption, but little is known about inhalation of particulates (e.g., smoke).

1.06.4.1.2 Drug distribution

The volume of drug distribution refers to the quantity of the drug in the body tissue relative to blood or plasma. This measure depends on the degree to which the drug exists in an ionized or a nonionized, lipid (fat) soluble state, and binds to proteins in blood limits its entry into tissue. These factors determines what tissue a drug can enter. Finally, drugs must be transported across membranes (usually in capillaries). Drugs that are lipid-soluble or small water-soluble molecules can passively diffuse or be transported across membranes. Drugs that are large water-soluble molecules require active transport across membranes. The blood–brain barrier (a capillary-glial cell layer) admits nonionized, lipid-soluble drugs, but excludes ionized, water-soluble drugs. Whether the drug reaches the CNS depends on the extent to which it stays in the blood, bound to proteins and how well it passes the blood–brain barrier. Similarly, lipid-soluble drugs can readily pass the placental barrier. Drug risk for the fetus is greatest both early (teratogenesis) and late (respiratory depression) in pregnancy. The circulatory system rapidly distributes drugs through the body, unless the drug is bound to blood-borne proteins. However, not all water-soluble drugs are bound to plasma, for example, the lithium ion.

1.06.4.1.3 Elimination

Clearance refers to the rate of elimination of the drug in urine relative to its concentration in blood (or any tissue). Some maintenance concentrations will achieve a steady-state drug level when rate of drug administration equals the total rate of clearance. A drug's half-life is the time required for a drug to reach 50% of steady-state levels during administration or to decay 50% from steady-state levels after administration ceases. Time-course versus drug-concentration curves map both absorption and elimination processes.

Elimination mechanisms involve the metabolism (breakdown) of drugs by one of four major metabolic routes: oxidation, reduction, hydrolysis, and conjugation. Liver and blood-borne enzymes convert drugs into ionized and excretable metabolites. First-pass metabolism (FPM) refers to the initial transport of the drug to the liver via the arterial portal system. Here, much of it is broken down by one or more of the several metabolic pathways. The rate of metabolism varies with drug and route of administration, for example, FPM is greater with oral than intravenous routes of administration. Drugs and their metabolites are excreted though the kidney, lungs, or bile, but also in sweat, saliva, and breast milk. The rate of metabolism may depend on several factors. Enzyme induction promotes increased enzyme activity (see Section 1.06.5.2), depression of enzymes produced by drug accumulation (e.g., disulfiram blocks metabolism of the alcohol metabolite acetaldehyde). Other metabolic factors include age, gender, circadian rhythms, nutrition, disease state, and genetics (including ethnicity). Most drug interactions also depend on metabolic factors. (For a description of metabolic pathways in the CNS, see Cooper et al., 1996). Clearly, metabolism also affects the therapeutic window.

1.06.4.2 Pharmacodynamics

1.06.4.2.1 Receptor theory

A receptor is a binding site (a protein membrane component) to which a ligand, the drug, can bind. This event produces an intracellular action that results in some drug effect. Yet drugs can bind to many cellular sites other than NT receptors. For example, the D_2 receptor is the primary target site for chlorpromazine, an antipsychotic, and for endogenous dopamine. However, the enzyme inositol-1-phosphatase may be the receptor for lithium, and a calcium channel the receptor for verapamil.

1.06.4.2.2 Drug receptor properties

Drug receptors have several distinctive properties. Generally, a limited number of receptors exist, and the extent to which the drug molecules occupy these sites varies with its concentration. Saturability defines the size of this effect with maximum saturation occurring with occupation of all the sites. Receptor selectivity refers to the extent to which a receptor binds with a particular drug rather than other molecules. Selectivity depends both on the receptor and on the size, shape, and bioelectrical charge of the drug molecule. Reversibility of drug/receptor binding occurs, since the latter interaction is generally not permanent.

How avidly or "tightly" a receptor binds a given drug molecule is described as its affinity. Receptors with high affinity for a drug require a lower drug concentration for full saturation. For example, agonists and antagonists could bind to the same receptor but differ in their affinity. A high affinity agonist and low affinity antagonist could lead to an insurmountable drug effect.

Efficacy refers to the size of the intracellular or drug effect when the drug and receptor interact. Zero efficacy drugs include receptor antagonists, which generally have no effect other than preventing the receptor from being activated by an agonist drug. A full agonist drug has high efficacy and can produce the maximum effect on receptors at a sufficient concentration. Partial agonist or inverse agonist drugs have a lower efficacy and cannot produce the maximal effect at any drug concentration level.

Allosteric regulation of receptors with multiple binding sites can occur. Allosteric ("other side") binding sites are usually sites other than that for the endogenous agonist. Activation of these sites can change the shape of the NT receptor, thereby positively or negatively regulating its affinity for the primary endogenous ligand. For example, at the $GABA_A$ receptor complex, BZD agonists enhance and inverse agonists retard activation of $GABA_A$ receptor activation.

1.06.4.2.3 Types of drug-receptor interactions

These interactions determine how drugs affect cell function (see Smith & Darlington, 1996). Activation of extracellular receptors can open an ion channel, for example, the action of BZDs on the $GABA_A$ receptor Cl^- channel. Stimulation of extracellular receptors coupled to an intermediary G-protein can produce intracellular enzyme changes (tetrahydrocannabinoids). Triggering of extracellular receptors on a transmembrane protein can effect intracellular enzyme changes. Finally, drugs can be directly transported across membranes to affect receptors in the cell nucleus, for example, steroids.

At least three types of receptor antagonism exist. Competitive antagonists bind to the same site as the agonist and compete with the agonist for that site. Noncompetitive antagonists bind to a site different from that to which the agonist binds but block receptor activation by the agonist. Finally, partial antagonists, or mixed agonists and antagonists have agonist properties with submaximal efficacy.

1.06.5 PRINCIPLES OF BEHAVIORAL PHARMACOLOGY AND PSYCHOPHARMACOLOGY

Much of the information described below also applies to most nonbehavioral physiological measures of drug effects, for example, heart rate and endocrine effects. Behavioral pharmacology (usually preclinical, animal studies) or psychopharmacology (usually human, clinical studies) defines the special principles or factors one should address when considering drug actions on behavioral or mental processes. Some references relevant to this topic include those by Bloom and Kupfer (1995), Carlton (1983), Leonard (1992), McKim (1986), and the volumes edited by Goldberg and Stolerman (1986), by Grabowski and VandenBos (1992) and by Thompson, Dews, and Barratt (1987).

1.06.5.1 Behavioral Drug Properties

1.06.5.1.1 Dose–effect curves (DEC)

Plotting drug effect against drug dose or concentration conveniently illustrates a variety of drug properties, individual differences in drug action, and changes in drug action over time. Several standard types of dose–response relationships have been identified. Lawful and constant rates of change in response with increases in the dose describe linear DECs (rarely demonstrated across all doses). With sigmoidal DECs, the typical dose–response relation signifies some threshold dose at which a change in the response begins to occur and an asymptotic dose at which higher doses produce no further response changes. Inverted U-shaped dose–effect functions resemble sigmoidal DECs, except that at doses higher than those producing the maximum response, the size of the response change decreases with further dose increases.

1.06.5.1.2 Drug sensitivity and effectiveness

Potency refers to the dose or concentration of the drug required to produce some effect, for example, different drugs can have different potencies for a given effect. The ED_{50} (effective dose) indexes the dose that produces half the maximal drug effect. For example, barbiturates are more potent (require a lower dose in mg/bodyweight) in producing sedation than alcohol. Similarly, haloperidol produces more potent antipsychotic effects than chlorpromazine. Conjoint treatment with a second drug can alter the behavioral potency of drugs (Dwoskin, Neal, & Sparber, 1987). For example, pretreatment with an antagonist will shift the agonist DEC to the right, resulting in a higher ED_{50}, indicating a decreased sensitivity or potency for the agonist drug. Likewise, pretreatment with a similar agonist (#2) could shift the DEC for agonist #1 to the left, resulting in a lower ED_{50} and possibly indicating some synergy between the drugs. A related DEC metric, the toxicity index (TD_{50}) defines the dose at which 50% of patients experience toxic effects. Finally, the therapeutic index, defined as the ratio of the median toxic dose (TD_{50}) to the median effective dose (ED_{50}), provides a relative measure of a drug's toxicity or safety. Drugs with a very low therapeutic index, like lithium and Clozapine, require monitoring of serum drug levels.

Efficacy provides a measure of the maximum desired effect, given a sufficient drug dose. In the behavioral economics of psychotropic medications, knowing the cost/benefits of different drugs is important. Drugs must have sufficient efficacy with the fewest side-effects and lowest costs. Note that idiosyncratic drug responses can occur, for example, some patients become quite agitated when given BZDs such as diazepam or atypical anxiolytics such as buspirone. While some individual differences in response to drugs are unpredictable, certain factors may predictably alter drug potency and/or efficacy, for example, age, gender, ethnicity, prior drug history, and current psychiatric status.

1.06.5.2 Drug Tolerance

1.06.5.2.1 Definitions

This phenomenon is most easily shown with sedative/hypnotic drugs. Two definitions of tolerance differentiate drug history versus individual, predrug history. Baseline tolerance defines those initial individual differences in the potency of a drug to produce some effect in a drug-naive subject. That is, drug-naive individuals could be hyporeactive, normoreactive, or hyperreactive to a given drug. The source of this potency difference could stem from pharmacokinetic or pharmacodynamic factors or from factors like genetics, age, gender, disease state, emotional state, environment, and so on. Acquired drug tolerance defines the situation where prior exposure to a drug decreases the potency of that drug to produce a given effect. Stated differently, to produce the same, initial drug effect requires a higher dose (i.e., the dose-effects curve shifts to the right producing a higher ED_{50}).

1.06.5.2.2 Kinds of tolerance

Physiological tolerance describes the body's adaptation to repeated drug exposure. Metabolic or pharmacokinetic tolerance could result from an increase in the rate of drug metabolism after repeated drug exposures. For example, the induction of metabolic enzymes for the drug could produce metabolic tolerance. Functional or pharmacodynamic tolerance results from a decrease in the way CNS receptors respond to the drug after repeated drug exposures. A down-regulation of receptors (i.e., fewer) or a reduction in receptor affinity or efficacy could produce a functional tolerance effect. Finally, cross-tolerance may occur, where the development of tolerance to one drug conveys tolerance to drugs within the same class or acting on similar CNS mechanisms.

Behavioral tolerance belongs to a special class of functional tolerance, involving conditioned or learned changes rather than drug receptor changes (see Goudie & Demellweek, 1986). The opponent-process theory of drug tolerance (Siegel, 1990) relies on the Pavlovian conditioning model. According to this theory, most drugs appear to have immediate alpha effects (e.g., a decrease in activity), followed by delayed-rebound beta effects. The latter are opposite in direction to that of the alpha effect (e.g., an increase in activity). Tolerance results from conditioning of the beta opponent process. With repeated drug exposures in the same context, the beta effect becomes conditioned to the contextual stimuli. It gradually moves forward in time, reducing the initial alpha drug effect. This kind of learned tolerance is specific to the drug exposure context. The compensatory-response model of behavioral tolerance involves instrumental or operant conditioning (Holloway & King, 1989; Schuster, 1978). According to this model, drug effects on behavior result in a loss of reinforcement. This prompts the acquisition of some strategies to compensate for the drug's effects and restore access to reinforcements. This kind of learned tolerance is specific to the drug's behavioral effects and to the behavioral-reinforcement demand characteristics of the task.

1.06.5.3 Sensitization

Sensitization refers to an increased potency of a drug with repeated exposures to that drug (i.e., decrease in the ED_{50}). The principal theory for sensitization points to pharmacodynamic changes. According to this viewpoint, an increase in the way CNS receptors respond to the drug occurs after repeated drug exposures. The mechanism of these changes is incompletely understood. Nevertheless, the obvious candidates are an up-regulation of receptors (i.e., more) or an increase in receptor affinity or efficacy. This phenomenon is most readily shown with stimulant drugs like amphetamine or cocaine. Some evidence points to the possibility of behavioral or learned sensitization. A related phenomenon, reverse tolerance, refers to an increase in a drug's potency with repeated exposure due to toxic reactions, disease state, or organ damage. For example, chronic alcoholics with cirrhosis may experience increased alcohol effects due to the loss of the chief organ of alcohol metabolism.

1.06.5.4 Physical Dependence

Physical dependence refers to a general physiological adaptation to chronic drug exposure, usually manifested during withdrawal from the drug (see Henningfield, Schuh, & Jarvik, 1995). The chronic withdrawal or abstinence syndrome involves physiological, subjective, and cognitive symptoms. The character of the withdrawal syndrome depends on the drug class. For example, sedative/hypnotics, like alcohol, BZDs, and morphine, generally produce an excitable withdrawal profile. On the other hand, stimulants like cocaine or amphetamine generally produce depressive or lethargic withdrawal symptoms. These data suggest an opponent-process, rebound pattern of withdrawal symptoms. However, both drug classes can produce subjective effects of anxiety. Acute withdrawal after even a single drug exposure also can occur, for example, alcohol hangover.

1.06.5.5 Basic Psychopharmacology Principles

1.06.5.5.1 Drug–behavior interactions

The preceding Sections on tolerance and sensitization clearly illustrate the importance of drug history in understanding the current drug effects on behavior. However, other biobehavioral factors related to drug history also can influence current drug effects. For example, experience with a drug's subjective and/or behavioral effect probably removes novelty *per se* as a determinant of drug action. Further, prior use of related drugs may activate the same enzyme or brain systems as the primary drug (cross-tolerance?). Further, one's subjective or behavioral history can influence the potency and/or efficacy of a drug's behavioral effects: For example, prior practice on a behavioral task may reduce the size of the drug's effects on task performance. Also, a drug can affect newly acquired behaviors or skills (e.g., young drivers) more adversely. Finally one must distinguish between behavioral antagonism versus masking by drugs. Sometimes a drug may decrease performance or behavior, not by antagonizing some endogenous receptor, but by producing other effects that interfere with that behavior (Dwoskin et al., 1987).

1.06.5.5.2 Baseline effects/rate dependency

The intensity or frequency of behavior can influence both the size and direction of a drug's effect on that behavior. For example, with high rates of responding, even a drug like amphetamine will reduce responding (such as methylphenidate's effects in ADHD). Similarly, with low rates of responding, even a drug like alcohol or a benzodiazepine can increase that behavior. In this regard, related baseline factors can modulate a drug's behavioral action. For example, acute or chronic state variables can influence the size of the drug effect. The latter states include an individual's normal states (e.g, arousal level) and any psychiatric state or trait factors (e.g., anxiety or depression).

1.06.5.5.3 Discriminative and state-dependent drug effects

Nearly all psychotropic drugs produce interoceptive stimulus cues that subjects/patients can learn to discriminate or detect (see Colpaert, 1986). These drug stimuli can become conditioned occasion setters for other behaviors. A related phenomenon, state-dependent retrieval, shows that information acquired under a given drug state is most efficiently recalled only when that drug state is present again.

1.06.5.5.4 Reward and reinforcing effects of drugs

Drugs produce positive reinforcement when the antecedent behavior ("self-administration") increases in its frequency of occurrence (see Dworkin & Smith, 1987; Henningfield, Lukas, & Bigelow, 1986; and Young & Herling, 1986, for descriptions of animal and human studies respectively). Drugs also act as negative reinforcers, that is, they may be self-administered to alleviate some aversive condition (e.g., taking aspirin for arthritis or consuming alcohol to avoid withdrawal or hangover). This latter situation describes the "self-medication" hypothesis of drug seeking. In addition, drugs can produce reward or aversive effects (i.e., positive or negative hedonic states), as indicated by an attraction to or an avoidance of places associated with such states. Strictly speaking, reward and positive reinforcement are not the same, but the same neural systems may mediate both, that is, the dopaminergic mesolimbic brain areas (n. accumbens and VTA; Wise, 1990).

1.06.5.6 Drug Classes

Multiple classification schemes for psychotropic drugs exist. Sometimes such drugs are grouped by their mechanism of action (see Cooper et al., 1996), by the pattern of their behavioral effects (e.g., drugs of abuse; see Balfour, 1990; Smith & Darlington, 1996), by the symptoms or disease for which they are used to treat (see Stahl, 1996), or by a mixed set of criteria (see Hardman & Limbird, 1996; Julien, 1998). This chapter will not focus on such classification schemes, but the interested reader can refer to the references just cited. The major problem with such classification schemes is that many psychotropic drugs have profiles showing effects on multiple NTS, on multiple aspects of behavior, and on multiple types of psychiatric symptoms.

1.06.5.7 Drug Side-effects, Interactions, and Overdose

The particular array of side-effects depends on which NT receptors are blocked, i.e., muscarinic (blurred vision, dry mouth, urinary retention, constipation, sinus tachycardia, and cognitive dysfunction), α_1-adrenergic (blockade of some antihypertensive drug effects, postural hypertension, drowsiness, dizziness, and reflex tachycardia), α_2-adrenergic (blockade of some antihypertensive drug effects), D_2-dopaminergic (extrapyramidal movement disorders), and H_1-histaminergic (sedation, weight gain). Drugs can have undesired interactions through pharmacokinetic or pharmacodynamic mechanisms. These interactions can vary widely in their potential to cause serious problems. Tricyclic antidepressive drugs can produce potential hazardous interactions when combined with MAOIs, NE, Epi, or phenothiazines. Drug overdosing could result from certain classes of drug interactions but more commonly results from the patient's deliberate attempt at suicide. See Kaplan and Sadock (1996), Hughes and Pierattini (1992), or Gelenberg and Bassuck (1997) for the diagnosis and treatment guidelines for these drug effects.

1.06.5.8 Drug Regulation

The federal Food and Drug Administration (FDA) in the United States was given the authority under the federal Food, Drug, and Cosmetic Act (FDC) to approve only those new drugs that were both safe and effective. The sale of such drugs is conditional on truthful labeling and all the pertinent information needed for the drug's safe and effective use. Prescription of a drug for unapproved use is not regulated but this does not exempt the prescriber from liability. The Drug Reinforcement Administration classifies drugs according to various levels of their abuse potential, for example, Schedule I drugs like heroin are deemed to have high abuse potential and low usefulness, while Schedule V drugs like lomotil have low abuse potential and high usefulness. Scheduled drugs with higher abuse potential require increased caution when prescribing.

1.06.6 THE PSYCHOPHARMACOLOGY OF MENTAL DISORDERS

Drugs used to treat mental disorders are given for a variety of reasons. Some disorders are better understood than others, and psychopharmacological treatments are based on hypotheses about the biological causes or influences of the disorder. Unfortunately, our understanding of many more disorders is limited. Psychopharmacological treatments for these disorders may be based on theories about the biological systems involved. Stahl (1996) describes at least eight ways that a compromised CNS could affect the manifestation or expression of behavior or mental disease: (i) changes in molecular neurobiology (e.g., the inheritance of an abnormal gene); (ii) loss of synaptic plasticity (i.e., the branching, pruning, growing, and dying of axons and dendrites); (iii) excitotoxicity (e.g., the overexcitation of a cell by glutamate, resulting in cell death); (iv) no neurotransmission (e.g., the loss of DA input to the basal ganglia in Parkinson's disease); (v) too much neurotransmission (see discussion below on the DA hypothesis for schizophrenia); (vi) an imbalance among NTS; (vii) the wrong rate of neurotransmission, and (viii) the wrong neuronal wiring during development. The interested reader is encouraged to refer to any of a number of reference works for more complete descriptions of theories, models, and data on the psychobiology of neuropsychiatric illness (Bloom & Kupfer, 1995; Kaplan & Sadock, 1996).

Frequently, however, treatment drugs that alleviate symptoms of a disorder are discovered both by serendipitous observations and by systematic screening of candidate drugs. Therefore, we will discuss the current theories about the systems involved in many mental disorders and the drugs used to treat the disorders separately. Where research has not differentiated between the biological system involvement of disorders, disorders will be grouped into classes (e.g., depressive disorders). We will not include mental disorders that have no apparent psychopharmacological treatments (e.g., gender identity disorders). Further, because of space limitations, this section will not address all neuropsychiatric disorders (e.g., somatoform disorders, sleep disorders, seizure disorders, and personality disorders).

Before reviewing the clinical psychopharmacology of psychiatric disorder, a few prefatory remarks must be considered. First, currently these disorders are defined based on symptom profiles with inclusionary and exclusionary criteria, and based on some known biological cause (see American Psychiatric Association, 1994; World Health Organization, 1993). However, rapid progress had been made in developing brain chemistry models of these disorders. Improvements in the clinical psychopharmacological study of therapeutic agents have paralleled this trend (e.g., double-blind techniques, preclinical screening with animal models, blind raters, "active placebos," and sophisticated analyses). All of the earlier sections on basic psychopharmacology are relevant here plus two additional concepts, those of rebound and relapse. Relapse refers to the reemergence of the original disease process after discontinuation of the drug, for example, panic attacks return after stopping benzodiazepine medication. Rebound refers to the recurrence of a more severe version of the disease, for example, more severe and frequent panic attacks.

1.06.6.1 Mood Disorders: Depressive Disorders

1.06.6.1.1 Characteristics of depression disorders

The cluster of symptoms in depression involves abnormalities of mood, its quality, its direction of change from normal (down indicates depression, up indicates mania), and its frequency and duration (Stahl, 1996). However, other dimensions of depression involve symptoms that may be vegetative (reductions in sleep, appetite, weight, and sex drive), cognitive (lowered attention span, memory function, frustration tolerance, and self-esteem), impulse control (suicide and homicide attempts), behavioral (lowered motivation, pleasure, interest, stamina), and somatic (increased headaches, stomach aches, and muscle tension). The possible presence of other distinguishable symptom clusters such as anxiety (comorbid) further complicates the symptom picture of depression. Finally, depression may result from or accompany any number of organic conditions, including neurological disorders, metabolic or endocrine disorders, infectious diseases, tumors, drugs, medications, or poisons.

1.06.6.1.2 The catecholamine (CA) hypotheses of depression

Many theories of how depressive disorders (e.g., major depressive disorder, dysthymic disorder) develop have been proposed. The most widely known is the CA hypothesis, first postulated by Schildkraut (1965). That certain drugs that alter CA function also alter depressive symptoms provides the strongest evidence for this hypothesis. For example, the effective antidepressives, MAO inhibitors and tricyclics, work to increase CAs in the synaptic cleft. Additionally, the drug reserpine interferes with the CA storage mechanism and causes massive release of CAs, followed by lower than normal CA levels. Reserpine can induce serious depression that is indistinguishable from endogenous depression and administering L-DOPA, the precursor to DA, reverses reserpine-induced depression. A second area of evidence favoring the CA hypothesis is the connection between depression and the norepinephrine (NE) metabolite, 3-methoxy-4-hydroxyphenylglycol (MHPG). Early studies found urinary levels of the NE metabolite were lower in depressed patients compared with controls. More recently, researchers have found that unipolar patients show a wide range of MHPG levels, and we can probably identify the best treatment based on this measure. Low-MHPG patients generally respond well to NE-active antidepressants (imipramine) or fluoxetine, and high-MHPG patients respond poorly to NE tricyclic antidepressants and fluoxetine. Thus, high-MHPG patients may have a depression not associated with impaired CA function.

Some unanswered questions about the CA hypothesis persist. First, why do drugs such as cocaine (a CA reuptake inhibitor), that act to increase CAs, fail to work as antidepressants? Secondly, why do other drugs, such as iprandole, work as antidepressants yet do not affect any CA system? Thirdly, why do antidepressants require at least two weeks before they produce a measurable clinical effect, although the pharmacological action is essentially immediate? The fourth question is why antidepressants continue to work in spite of postsynaptic receptor desensitization, occurring after chronic antidepressant administration. It seems clear, then, that the CA hypothesis cannot explain all the phenomena associated with depressive disorders, and other or additional theories are necessary.

1.06.6.1.3 The serotonin (5-HT) hypothesis of depression

The inability of the CA hypothesis completely to explain some facets of depression led to the development of other theories. The 5-HT hypothesis was originally suggested by European researchers who were encouraged by findings that depressed patients with low levels of the 5-HT metabolite 5-HIAA were at greater risk for violent suicide attempts. More recent research has found that low 5-HT turnover is related to impulsive violent behavior, although it is not specific to depression.

However, the 5-HT hypothesis led to the development of the currently popular CA and 5-HT hypothesis. This theory relies on the fact that NE and 5-HT systems are known to interact with each other. Selective lesions of 5-HT neurons increase NE receptor density in some regions of the brain, and NE receptors cannot be down-regulated without an intact 5-HT system. The strongest evidence for the CA and 5-HT hypothesis rests on two discoveries: first that well-known antidepressants like the tricyclics affect both NE and 5-HT systems, and second that the 5-HT-specific reuptake inhibitors (SSRIs) have strong antidepressive effects but primarily affect 5-HT systems. The suggestion is that depression disrupts a delicate balance between CA and 5-HT systems that can be restored with medication.

1.06.6.1.4 The cortisol hypothesis of depression

The last theory to be discussed does not involve CA or 5-HT systems at all. It is the hypothalamic-pituitary-adrenocortical (HPA) axis dysregulation hypothesis. This theory is supported by the finding that most people with Cushing's disease have symptoms of depression. Cushing's disease is a disorder of the HPA axis characterized by excessive cortisol secretion. The HPA axis dysregulation hypothesis is based primarily on the usefulness of two endocrine challenge tests, the Dexamethasone Suppression Test (DST) and the Corticotropin Releasing Hormone (CRH) Stimulation Test. Dexamethasone is a glucocorticoid receptor agonist that will drive the negative feedback system for cortisol secretion in a person with a functional HPA axis. In nondepressed people, administering dexamethasone causes a precipitous drop in cortisol production. In approximately 50% of depressed patients, cortisol production drops only a little. These individuals are known as nonsuppressors. Interestingly, the rate of cortisol nonsuppression increases with the severity of the depression (e.g., most people with depression with psychotic features are nonsuppressors). From these findings, it is inferred that the glucocorticoid receptors in depressed people are down-regulated or desensitized, so their cortisol levels would regularly be too high.

Many people with major depression also have high levels of CRH in their corticospinal fluid, which is believed to come from chronic hypersecretion of CRH. In a functional system, CRH is released by the hypothalamus and causes the release of ACTH, which itself causes release of cortisol. Administering CRH to a healthy individual induces a dramatic increase in ACTH and, subsequently, cortisol. This challenge is called the CRH stimulation test. Some depressed patients fail to show as large an increase in ACTH and cortisol, which is thought to be caused by down-regulated or desensitized CRH receptors. If a single depressed patient is given the DST and the CRH stimulation test and age is taken into account, the specificity of diagnosis of endogenous depression is more than 90% (Heuser, Yasouridis, & Holsboer, 1994). However, it is not yet clear whether the high levels of cortisol and CRH are a cause or result of, or even unrelated to, development of depression. In addition, no medication has been found to interact specifically with the HPA axis, although traditional antidepressants will reduce levels of cortisol and CRH.

1.06.6.1.5 Status of depression theories

Clearly, no single NT theory accounts for all the data. This situation results from the fact that depression is not one disorder and that the various NT systems interact. Also, dysfunction of any of several NT systems may be sufficient to produce depression. For example, cholinergic neurons activate muscarinic receptors on the NE neurons of the locus coeruleus. Further, cholinergic agonists are known to induce depressive symptoms and chronically depressed patients are hypersensitive to cholinergic agonists.

1.06.6.1.6 Drug therapies for depression

Clinical psychopharmacology of depression is very effective. Approximately 60–70% of people with depression respond to pharmacotherapy, and rates go up to 80% if limited to mild or moderate major depressive disorder. Three types of antidepressants are currently used: MAO inhibitors, tricyclics and tetracyclics, and SSRIs. MAO-inhibiting drugs were the first antidepressants, and include phenylzine, tranylcypromine, isocarboxazid, and seligiline. MAO inhibitors inhibit the action of MAO, which metabolizes CAs. The result is that more CA is available in the cleft. Both MAO_A and MAO_B inhibitors are useful, as MAO_A breaks down NE, 5HT, and DA, and MAO_B breaks down DA. The three drugs listed inhibit the action of MAO irreversibly, which requires that new enzymes be synthesized. European health professionals use reversible MAO inhibitors that are not currently available in the USA.

However, potentially serious side-effects occur with many MAO inhibitors. The most famous side-effect is the “cheese reaction.” Foods containing tyramine can induce hypertension in patients on MAO inhibitors. MAO also metabolizes tyramine, and its inhibition in the gastrointestinal tract allows tyramine to pass into the blood where it causes a sympathomimetic response. Patients on MAO inhibitors should avoid a range of foods, including certain cheeses, some types of alcohol, and some cured meats and sausages. Fortunately, the MAO_B inhibitor selegiline does not produce the cheese reaction. In addition, because MAO is involved with many systems in the body, patients should not use multiple medications, including many over-the-counter cold or flu remedies. MAO inhibitors should also not be combined with SSRIs or tryptophan treatment. The latter combination poses a great risk of the 5-HT syndrome, which can include tremor, hypertonicity, myoclonus, autonomic symptoms, hallucinations, and fever. MAO inhibitors also commonly induce insomnia. In fact, the alerting properties of MAO inhibitors can be used to advantage in hypersomnic people. A related effect is afternoon fatigue, which can be severe. In part due to the serious side-effects, MAO inhibitors are generally prescribed for treatment of refractory depression, particularly depression with atypical features (hypersomnia, weight gain), combined depressive and anxious symptoms, or personality disorder symptoms.

Tricyclic and tetracyclic antidepressants, such as imipramine, amitriptyline, clomimipramine, doxepin, and maprotiline, are named for their molecular structure. They were first found to alleviate depressive symptoms in the 1950s. Their primary mechanism of action is to block reuptake of NE, 5-HT, and some affect DA reuptake. Although the result is to increase the amount of NE and 5-HT in the synapse, the clinical effect is thought to occur through receptor changes. The prolonged presence of increased, synaptic NE results in desensitization of postsynaptic β-receptors. This state of affairs correlates with the desired clinical effect. In addition, the increased NE also acts to desensitize α_2-autoreceptors, which results in a further increase of the amount of NE released. Increased 5HT in the synapse seems to increase the sensitivity but reduce the density of postsynaptic receptors. The sensitization time course is also correlated to clinical effect. Tricyclics also affect second-messenger systems and act as cholinergic antagonists, either of which may result in a clinical effect. Tetracyclics may also antagonize postsynaptic 5HT and DA receptors.

Side-effects of tricyclics and tetracyclics are generally due to the anticholinergic and antihistaminergic effects. Common side-effects include dry mouth, constipation, blurry vision, and tachycardia, which are all caused by peripheral anticholinergic properties. Central anticholinergic effects can include memory loss, confusion, and delirium. Sedation is another side-effect and it is due to histamine antagonism. Some people also experience sexual dysfunction and weight gain. More important, the margin of safety for these drugs is relatively small. Doses equivalent to two to three week prescriptions can be lethal. Death usually occurs from cardiac toxicity or seizures. Consequently, tricyclics and tetracyclics are used with extreme care in individuals who are or may become suicidal. Finally, patients must wait two to six weeks to see any clinical effect on symptoms of depression.

The 5-HT selective reuptake inhibitors (SSRIs) currently available in the USA are fluoxetine, paroxetine, and sertraline. Fluoxetine was discovered in the 1970s and was introduced for clinical use in 1988. Since then, it has become the most widely prescribed (and widely researched) antidepressant in the USA. As one might surmise by their name, SSRIs inhibit reuptake of 5-HT. Because they have little to no interaction with cholinergic, histaminergic, or noradrenergic systems, they produce fewer side-effects. Fluoxetine has a long half-life and an active metabolite, which also has a long half-life. In clinical trials, fluoxetine has performed as well as or better than tricyclic antidepressants in relieving symptoms of depression.

Although fluoxetine has a better side-effect profile, it does produce more nervousness, sleep disturbances, and nausea than tricyclics. Interestingly, it also produces weight loss, which is opposite from the weight gain often seen with tricyclic use and has been useful in treating obesity. Fluoxetine is also sedating, at rates similar to the tricyclic antidepressants. In spite of lay reports, large clinical trials show no greater risk of suicide or suicide attempts over that seen with other antidepressants. Some have suggested that the activating effects (restlessness and agitation) may cause some susceptible people with a history of suicidal ideation to attempt suicide. Therefore, any patient who exhibits increased activity and feelings of restlessness should be monitored carefully. In contrast to the tricyclics, higher doses of SSRIs do not appear to alleviate symptoms of depression better. In fact, some clinical trials suggest that patients may do less well on higher doses. Other antidepressives include mianserin and iprindole. Newer types of antidepressive act to alter the linkage between the adrenergic and/or serotonergic NT and second-messenger systems, for example, Rolipram, which inhibits cAMP phosphodiesterase.

1.06.6.2 Mood Disorders: Bipolar Disorders

Bipolar disorders (both Bipolar I and Bipolar II) have been effectively treated with medication for nearly 50 years. People with bipolar disorder experience alternating periods of depression and mania. Lithium is the most common psychopharmacological treatment, but other drugs have also been found useful. Lithium is effective in approximately 70–80% of patients presenting with acute mania. The mechanism by which lithium exerts its therapeutic effect is not yet clear, but possibly it enhances the activity of the 5-HT system. Some evidence shows that lithium increases the release of NE but also reduces its postsynaptic receptor sensitivity. However, since lithium affects most NT systems and signal transduction pathways, conclusions cannot yet be drawn as to its mechanism of action. Lithium appears to block an enzyme of the intracellular second-messenger system (inositol phosphate), resulting in a reduced responsiveness in those neurons that depend on such second-messenger linkages (Kandel et al., 1991).

Lithium's side-effects include confusion and memory problems and loss of some motor coordination and development of a slight tremor. It can also induce hypothyroidism. It must be noted that lithium has a small therapeutic window, with high serum levels causing neurotoxicity. Symptoms of lithium's neurotoxic effects include delirium, ataxia, and seizures. Toxicity can be fatal if not treated immediately. Lithium is also teratogenic, so should be used in pregnant women with extreme care.

Lithium is used to treat acute manic episodes, but its effect is not fully realized for five to ten days. As a result, acute episodes are often treated with BZDs or neuroleptics as short-term adjuncts to lithium. Lithium's primary usefulness, however, is in prophylaxis of manic and depressive episodes. Patients who experience mixed episodes or rapid cycling or have comorbid substance abuse are less likely to respond well to lithium. Patients with bipolar disorder also respond well to treatment with the anticonvulsant drugs, valproate and carbamazepine. MAO inhibitors, tricyclics, and SSRIs have also been found useful, particularly to treat the depressive episodes. Antidepressants are used most often as adjunctive treatments or with patients who have not responded well to lithium alone. Further, lithium can produce many side-effects (see Gelenberg & Bassuk, 1997).

1.06.6.3 Anxiety Disorders: Generalized Anxiety Disorder

While depression and its treatment are separated from anxiety and its treatment, most forms of anxiety disorders used to be lumped together under the heading of generalized anxiety disorder (GAD). Anxiety was seen as a normal emotion to some threatening situation (real or imagined). However, since the 1960s, research has principally focused on the anxiety subtypes described below. As happens with social phobia (see below), a clear research focus on the biology of GAD has yet to develop. For many years, GAD has been treated with BZDs, which continue to be the first drug of choice for these patients. However, recent concern about long-term treatment with BZDs has prompted a search for other psychopharmacological interventions. As a result, clinicians welcomed the introduction of buspirone, a non-BZD anxiolytic. Buspirone does not produce dependence and can be withdrawn with relative ease. Although buspirone is effective in treating GAD symptoms, it has a slow onset of action and may not be accepted by patients who want immediate symptom relief. Some studies also show that tricyclic antidepressants and SSRIs may be helpful for patients with GAD. As with depression, many organic conditions can cause anxiety states, including endocrine disorders (e.g., hyperthyroidism), metabolic disorders (e.g., hypoglycemia), and drugs or medications (e.g., caffeine or withdrawal from addictive drugs).

1.06.6.4 Anxiety Disorders: Panic Disorder

1.06.6.4.1 Phenomenology and etiology of panic disorder

Panic disorder (PD) is the best researched anxiety disorder. Twin studies have found higher concordance in monozygotic twins than dizygotic twins, suggesting that PD is heritable. In truth, we have only limited understanding of PD and its biological basis. One current theory is the respiratory hypothesis, the formulation of which stemmed from the discovery that administration of sodium lactate could induce panic attacks that were identical or very similar to spontaneously occurring panic attacks. Researchers found that lactate induces panic attacks in 50–70% of people with PD, but only 10% of people without PD. Later studies have not found any direct CNS effect of lactate, which is primarily a respiratory stimulant. Also, a correlation exists between hyperventilation produced by lactate and the induction of panic attacks.

Common symptoms of panic attack include shortness of breath, choking feelings, and trouble breathing. Voluntary hyperventilation induces panic attacks in 30–50% of people with PD, and inhalation of carbon dioxide-enriched air induces panic attacks in 50–80% of people with PD. The respiratory hypothesis suggests that PDs stem primarily from the increased sensitivity of the central CO_2 receptors. Activation of these receptors leads to a "false suffocation alarm," and triggers the psychological symptoms of anxiety. This hypothesis would suggest a purely psychological intervention of breathing retraining, physiological symptom exposure, and cognitive evaluation of the false suffocation alarm. In fact, cognitive-behavioral therapy reduces vulnerability to lactate-induced panic and appears to have short-term and long-term effects on the reduction of spontaneous panic attacks.

1.06.6.4.2 Neuropharmacology of panic disorder

However, some NT systems have been implicated in PD. Some researchers have found that patients with PD are hyperreactive to yohimbine (an α_2 NE receptor antagonist), while others have found patients to be hyporeactive to clonidine (an α_2 NE receptor agonist). These findings have led to the NE "dysregulation" hypothesis. Researchers have found reduced α_2 NE receptor binding sites and a blunted growth hormone response to clonidine, which suggests α_2 NE receptor downregulation. 5-HT has also been suggested as a possible mechanism for PD. Administration of m-chlorophenylpiperazine (m-CPP), which is a 5-HT agonist, in low doses showed a greater rate of panic induction in PD patients than in nonpatients. This might suggest a postsynaptic receptor supersensitivity. When this discovery is coupled with the suggestion that benzodiazepine receptor sensitivity is reduced, 5-HT is thought by some to be the primary route of PD development. However, the topic clearly requires more research.

1.06.6.4.3 Drug therapies for panic disorder

Psychopharmacological treatment of PD is common and often combined with psychotherapy. High-potency BZDs such as alprazolam and clonazepam have been used with good success. Although they were originally prescribed simply to relieve anticipatory anxiety, later research has suggested that they may function to reduce the number and intensity of the panic attacks themselves. BZDs generally produce symptom relief within a week, a benefit not shared with other medications. Antidepressants also used include the tricyclic imipramine, the tetracyclic clomipramine, and MAO inhibitors. The SSRI fluvoxamine has been found effective, although it is not currently available in the USA. Antidepressants are thought to exert their action by their 5-HT reuptake inhibition, although research has not yet borne this out. In contrast to the BZDs, antidepressants must be administered for three to six weeks to see any improvement. No psychopharmacological treatment to date seems helpful in alleviating symptoms of agoraphobia.

1.06.6.5 Anxiety Disorders: Obsessive-compulsive Disorder

For many years, clinicians were frustrated by the lack of effective psychopharmacological treatments for patients with obsessive-compulsive disorder (OCD). Then scientists discovered that clomipramine could decrease OCD symptoms, and it has become the drug of choice. Until then, clinicians had few clues for how OCD developed or what NT systems were involved. Twin and family studies have not supported the heritability theory of OCD. With the discovery of clomipramine's effectiveness, researchers developed the 5-HT hypothesis. This theory is based on the positive results of 5-HT-enhancing drugs, but additional research has found that OCD patients have increased turnover of 5-HT and its metabolites (5-HIAA).

In addition, the 5-HT agonist m-CPP exacerbates symptoms of OCD. Some researchers have found that pretreatment with clomipramine blocks the m-CPP-induced exacerbation, although others have not always been able to replicate this phenomenon.

The 5-HT system dysregulation also supports a second theory that suggests that patients with OCD suffer striatal damage. Clinically, OCD shares certain characteristics with tic disorders (i.e., uncontrollable repetitive behaviors), and tic disorders are believed to be caused by striatal dysfunction. OCD has also developed in patients with Huntington's, which involves lesions in the striatum. Interestingly, the pathway from the striatum to the thalamus and cortex is heavily dependent on 5-HT, which could explain the effectiveness of 5-HT-enhancing drugs. As mentioned above, 5-HT is primarily inhibitory and the neurons in the striatum appear to inhibit activity downstream after the thalamus. The striatal damage hypothesis postulates that the loss of striatal neurons would increase activity in the thalamocortical pathways. In fact, positron-emission tomography (PET) scans show increased glucose utilization in the caudate and orbitofrontal regions of OCD patients. Other studies have shown a reduction in metabolism in those areas after treatment with clomipramine, fluoxetine, or behavior therapy. In severe cases of OCD, psychosurgery lesioning of pathways from the thalamus to the orbitofrontal areas appears to alleviate symptoms. Thus, the evidence suggests a strong influence of 5-HT in the genesis of OCD.

As mentioned above, psychopharmacological treatment of OCD relies heavily on the use of 5-HT-enhancing drugs such as clomipramine and SSRIs. Over the years, clinicians have tried several other medications, but the 5-HT-enhancing drugs are the only ones currently recommended for use with people who have OCD. Doses higher than those used for depression are usually necessary. In a few cases, the disorder does not respond to any treatment. For these individuals, psychosurgery may be recommended to alleviate their symptoms. The most common surgery involves making lesions in the cingulate cortex on both sides of the brain.

1.06.6.6 Anxiety Disorders: Post-traumatic Stress Disorder

Research on post-traumatic stress disorder (PTSD) has primarily been with individuals who have combat-related PTSD, although other traumas do induce the disorder. The current theory of PTSD states that patients must show autonomic hyperresponsivity. This hyperresponsivity manifests itself in three ways. First, patients respond to reminders of their trauma (e.g., imagery, sounds, videos, or pictures) with increased blood pressure, heart rate, muscle tension, and other physiological indicators of stress. Second, patients also respond to trauma reminders with increased neuroendocrine (epinephrine and norepinephrine) activity. Third, patients have demonstrable HPA axis dysfunction. Acute stress increases cortisol release, and the chronic stress experienced in PTSD appears to result in high cortisol levels for a long period. High cortisol levels appear to damage the hippocampus, which would suggest that PTSD patients might experience neuroanatomical changes. People with PTSD have increased density of glucocorticoid receptors, which results in exaggerated negative feedback of ACTH and cortisol. PTSD patients also supersuppress in response to the DST, which further suggests an exaggerated negative feedback system. All the evidence to date points to PTSD being a disorder of hyperresponsivity.

Unfortunately, researchers have not yet been able to find a psychopharmacological treatment to correct for the autonomic hyperresponsivity. As a result, clinicians continue to rely on medications to treat the emotional and subjective symptoms of PTSD. The most common drugs prescribed are the tricyclic antidepressants, especially amitriptyline and imipramine. Researchers have had mixed results with these drugs, but they appear to have some effect beyond their antidepressive and anti-anxiety effects. Other drugs, such as the MAO inhibitor phenelzine, SSRIs, and carbamazepine have been used with varying success. Overall, research suggests that the positive symptoms of PTSD (e.g., increased arousal, re-experiencing) often respond to psychopharmacological treatment while negative symptoms (e.g., avoidance) do not.

1.06.6.7 Anxiety Disorders: Social Phobia

We know little about the biological basis of social phobia (SP). Some researchers have suggested that the sympathetic nervous system may be dysregulated, because people with SP report autonomic activation symptoms, and β-blockers such as propranolol sometimes help. In the late 1990s, however, a sustained research focus in the biology of SP does not exist. Administration of phenelzine, a MAO inhibitor, alleviates nervousness and fear of criticism for some people with SP. BZDs have also been helpful. Nevertheless, in psychopharmacological treatment, as in the etiology of SP, none of the choices are clearly superior. In fact, some research suggests that cognitive-behavioral treatment may be the best treatment choice.

1.06.6.8 Psychotic Disorders: Schizophrenia

1.06.6.8.1 Characteristics of psychosis

Psychosis is one of the oldest psychiatric syndromes (mixture of symptoms) but is not a specific diagnostic category in the *Diagnostic and statistical manual of mental disorders* (4th ed., *DSM-IV*) or *International classification of diseases* (10th ed., *IDC-10*) (Stahl, 1996). Psychotic symptoms include perceptual distortions (delusions, hallucinations), disorganized speech and behavior, severe distortions in testing reality, and motor disturbances (peculiar postures and stereotyped movements and acts). Mania, depression, substance abuse, and cognitive disorders like Alzheimer's dementia may present with psychotic features. Finally, psychoses can be described as paranoid (delusional beliefs and projections, hostile belligerence, and/or grandiose expansiveness), disorganized or excited, and/or depressive (retardation, apathy, self-blame). Finally, one should recognize that many organic conditions may produce or be associated with psychotic symptoms, for example, space-occupying CNS lesions, cerebral hypoxia, neurological disorders like temporal lobe epilepsy, cerebral vascular disorders, CNS infections like meningitis, metabolic and endocrine disorders like thyroid disease, nutritional deficiencies (thiamine in Wernicke–Korsakoff syndrome), and drug, medications or toxic agents like amphetamines or corticosteroids.

Medications are generally administered to any person who exhibits psychotic symptoms, including people with schizophrenia, schizophreniform disorder, schizoaffective disorder, delusional disorder, brief psychotic disorder, and severe depressive disorders with psychotic symptoms. Some newer research examines etiology and treatment of nonschizophrenic psychotic disorders, but little is known yet.

1.06.6.8.2 The phenomenology and biological basis of schizophrenia

Schizophrenia was first described in the nineteenth century by Benedict-Augustin Morel, who termed it *demence précoce* (Kaplan & Sadock, 1996). Schizophrenia is the most common (1% of the population) and, fortunately, the most well-understood psychotic disorder, but studies suggest that other psychotic disorders have similar etiologies. The symptom profile of schizophrenia really does have two faces. On the one side is the array of positive symptoms suggesting an exacerbation of normal functions (delusions, hallucinations, disorganized speech and behaviors, agitation, and bizarre behaviors). On the other side is the cluster of negative features suggesting a loss of normal function (flattened affect, alogia or reduced fluency of thought and speech, anhedonia, avolition or reduced goal-directed behavior, and poor attention). Finally, schizophrenia is characterized by a significant genetic heritability. Twin studies find monozygotic concordance rates of 30–80%, the highest found in any psychiatric disorder.

Fortunately, research on schizophrenia and theories as to its cause abound. One such theory proposes that schizophrenia may be related to neuroanatomical lesions common to people with the disorder. Neuropsychological tests have revealed deficits in prefrontal cortex and medial temporal lobe function, showing a lack of ability to plan and organize. Many schizophrenics also have more generalized deficits, including enlarged lateral and third ventricles and deeper sulci than people without the disorder. The ventricular enlargement is correlated with greater impairment in functioning and a poorer prognosis, although some have disputed this conclusion.

1.06.6.8.3 The DA hypothesis of schizophrenia

By far the most popular theory of schizophrenia and selected psychotic disorders is the dopaminergic hypothesis. This theory states that schizophrenia is caused by excessive central DA activity. In some people with the disorder, the striatum and nucleus accumbens, which are heavily dependent on DA, exhibit normal volume and area but a reduced neuronal diameter. Further, both areas show increased binding of the DA antagonists haloperidol and spiperone. The substantia nigra does manifest reduced volume, and postmortem studies reveal increased density of D_2 receptors in the caudate and putamen, while the density of the DA transporter and D_1 receptors remain normal. Some evidence suggests the existence of some structural abnormalities in DA receptors and an abnormal linkage between D_1 and D_2 receptors, but these findings are not well-replicated.

Antipsychotics block DA and reduce its function, which supports the dopaminergic hypothesis, but it is not certain that this is the mechanism by which the drugs relieve symptoms. Unmedicated schizophrenics do not show higher levels of HVA, the major metabolite of DA in cerebrospinal fluid, when compared with nonschizophrenics. Medicated schizophrenics do have higher HVA levels, which suggests that antipsychotics actually do work on DA neurons.

Long-term treatment with antipsychotics produces tolerance in nigrostriatal pathways, but not in mesocortical or mesolimbic pathways. In postmortems of schizophrenics on chronic antipsychotics, increased HVA was found in cingulate and frontal cortices (mesocortical and mesolimbic areas), but not in the putamen and nucleus accumbens (nigrostriatal system). Although many studies have found increased numbers of DA receptors in postmortem schizophrenic brains, it is not clear whether these are from the disease or the treatment. Some have suggested that an abnormality in the DA systems projecting to the nucleus accumbens may mediate symptoms like anhedonia, while DA systems projecting to the central nucleus of the amygdala may participate in the generation of paranoid ideation.

1.06.6.8.4 Antipsychotic medications

The primary psychopharmacological treatment for schizophrenia and other psychotic disorders is the use of DA antagonists, otherwise known as antipsychotics. The older typical antipsychotics and the newer "atypical" antipsychotics differ based on their pharmacological mechanism and on when they were found effective. Chlorpromazine, the first antipsychotic, was discovered in 1952. The class of conventional or typical antipsychotics is butyrophenones and phenothiazines, but this category also includes thioxanthenes, dihydroindolones, and dibenzoxazepines. Their DA antagonist properties were discovered when they were found to reduce the behaviors caused by administering dopamine agonists to both humans and animals. The clinical efficacy of typical antipsychotics is directly correlated to the drugs' affinity for the D_2 receptor. Later studies show that for an antipsychotic to work, it must effectively bind the receptor and produce a subsequent decrease in DA activity. Nevertheless, both butyrophenones and phenothiazines affect NE and DA systems. Also, both drugs increase DA turnover, presumably because their blockade of both postsynaptic DA receptors and presynaptic DA autoreceptors produces a feedback activation of these neurons.

Antipsychotic drugs are usually administered orally, but can also be given by depot (slow-release) injection. Medications can be given to treat a psychotic episode, but are more commonly used as maintenance therapy to prevent relapses. Antipsychotics are strong medications and can induce a range of side-effects. Medication effects can include sedation and a lack of interest in one's environment (ataraxia). Antipsychotics also increase prolactin secretion, which can result in breast development in men and milk secretion and menstrual abnormalities in women.

1.06.6.8.5 Antipsychotic drug side-effects

All typical antipsychotics also produce extrapyramidal symptoms (EPS) as a side-effect. Extrapyramidal refers to the localization of the motor dysfunction in the spinal tracts. EPS are motor problems that include Parkinsonian-like symptoms (stiffness, tremor, shuffling gait), acute dystonia (abrupt spasms of head and neck), and akathesia (physical restlessness). Unfortunately, the extent to which the drug works as a DA antagonist correlates with the extent of EPS. So anti-Parkinsonian drugs are often administered to help reduce these symptoms. Tardive dyskinesia (TD) develops in 10–20% of patients on antipsychotics for more than one year. TD is a movement disorder characterized by mouth, tongue, and face movements, besides slow, writhing movements of the limbs and trunk. The elderly and those with affective disorders are particularly vulnerable to developing TD. It was once thought that TD was progressive and irreversible, but this is no longer thought to be true.

Neuroleptic malignant syndrome (NMS) is an uncommon, but potentially fatal reaction to antipsychotics. It is characterized by severe muscular rigidity, autonomic instability, high temperature, and changing levels of consciousness. It happens more often when high-potency antipsychotics are given in high doses and escalated rapidly, and is more common in males and younger patients.

1.06.6.8.6 Side-effect profile of newer antipsychotics

Clozapine is the most common atypical antipsychotic currently used in the United States. Risperidone is also approved and is gaining popularity, and raclopride is still in the clinical trial phase of FDA approval. Atypical antipsychotics are so named because they appear to involve DA receptors other than D_2. Clozapine seems to affect neurons only in the ventral tegmental area and not in the substantia nigra, which is thought to be the source of the EPS and TD symptoms. This drug also has an affinity for 5HT neurons, which may be involved in its function. In fact, Clozapine has a "cocktail" of NT actions, affecting up to nine different receptors (Stahl, 1996).

Clozapine is noted for the rarity of patients developing EPS and TD while on it. NMS also occurs at lower rates than with typical antipsychotics. However, 1–2% of patients develop agranulocytosis (a loss of a particular group of white blood cells), usually within the first six months of use. This can be a fatal condition and, to monitor it, white blood cell counts must be checked weekly. Seizures are another significant side-effect. Because of the potentially serious side-effects, Clozapine is currently approved for use in treatment-refractory schizophrenia, patients with unmanageable EPS, and patients with TD.

Other atypical agents (Julien, 1997) include molindone (Moban), which resembles 5-HT and appears effective for schizophrenia with only moderate sedative and EPS profiles. Loxapine, structurally related to Clozaril, has antipsychotic properties but moderate sedative and EPS profiles and lowers convulsion thresholds. Carbamazapine (Tegretol) is not effective as an antischizophrenia drug. The interested reader should refer to Stahl's summary of some current discovery strategies for new antipsychotics (Stahl, 1996).

1.06.6.9 Substance Abuse and Dependence Disorders

1.06.6.9.1 Definitions and diagnostic criteria

Stahl (1996) outlines the major *DSM-IV* criteria for substance intoxication, abuse, and dependence. Intoxication refers to a reversible pattern of behavioral or psychological changes produced by the drugs' action on the CNS (e.g., mood, belligerence, impaired cognition, judgment, motor skill performance, and social functioning). Substance abuse refers to a maladaptive pattern of using any drug in "culturally disapproved manner" that leads to adverse consequences. The latter include failure to meet obligations, placing self and others in harmful situations, legal difficulties related to drug taking, and social or interpersonal problems. Substance dependence refers to a physiological state of adaptation produced by a prolonged maladaptive pattern of substance use. The *DSM-IV* definition of substance dependence includes three or more of the following characteristics: the development of tolerance (the need for more drug to get the same effect), a drug withdrawal syndrome (usually a drug rebound state opposite to the general effects of the drug), inability to control the amount consumed or to stop the consumption, the subordination of other life activities to drug taking, and/or the persistence of drug taking despite recurring adverse consequences (see Section 1.06.5.4). Relapse refers to recurrence of drug taking after some effective treatment regimen. Other concepts introduced earlier are relevant to this topic, for example, cross-tolerance and cross-dependence (substituting a second drug to substitute for or block withdrawal from the first one). The latter effects may in part account for the phenomena of polysubstance abuse. The reader should take note that addiction itself is not defined in the *DSM-IV*. We like the following one: "a behavioral pattern of drug abuse characterized by an overwhelming involvement with the compulsive use of a drug, the securing of its supply, and a high tendency to relapse after its discontinuation" (Stahl, 1996, p. 336).

1.06.6.9.2 Models of addiction

Of all the psychiatric disorders, those in this category are as much socially as physiologically defined. Marlatt (1992) in fact describes four models based on whether individuals are or are not responsible for developing the addiction and for changing their addictive behaviors. The "medical" or "addiction as a disease" model most readily lends itself to a discussion of psychopharmacological interventions. Most addiction researchers suggest that a drug's potential for abuse rests on its potential to act as a reinforcer and/or to induce euphoria (see Section 1.06.5.5.4 on reinforcing and reward properties of drugs). The latter effects derive from the drug's action on certain mesolimbic brain systems (specifically, the DA neurons from the ventral tegmental area [VTA] innervating the nucleus accumbens) that are involved in normal motivational processes (see Kalivas & Samson, 1992; Wise 1990). Proponents of this view cite the multiple NT systems that participate in this brain region (e.g., $GABA_A$, nicotinic cholinergic, serotonergic, and opioid systems). A key concept in this model is that it is a psychopharmacological model, that is, both the drug and the drug self-administration are equally important in sustaining the larger pattern of drug-seeking behaviors. Some biological theories of addiction suggest that genetic differences in sensitivity to drugs may reflect the under- or overexpression of the latter NT receptors, among others (see Crabbe & Harris 1991). For example, underexpression of one or more of the key NTs involved in the mesolimbic reward system might promote excessive use of drugs that affect such NTs (self-medication?). This idea is the same as Blum's "reward deficiency syndrome" (Blum, Cull, Braverman, & Comings, 1996).

1.06.6.9.3 Psychopharmacology of and pharmacotherapy for drugs of abuse

The following sections will describe some attempts to treat some kinds of drug abuse with drugs. Of course, a key to this discussion is what is being treated and what neuropharmacological systems are involved. The treatment of acute drug intoxication and drug withdrawal reactions is straightforward. The approaches to treating the drug-seeking behavior or drug craving may be a little less obvious. A classic strategy would be to arrange for the drug-taking behavior to be punished, for example, disulfiram or Antabuse (alcohol consumption with Antabuse makes one sick). Another strategy involves pharmacologically blocking the drug's CNS actions, for example, using naltraxone (an opiate receptor antagonist) to counteract heroin overdoses. A final strategy would be to substitute another drug or another form of the abused drug, for example, methadone (a synthetic opiate agonist) maintenance for heroin addicts or nicotine gum for addicted smokers. Apart from treatment of the addiction, medications may be used to treat acute intoxication or overdose, withdrawal reactions, or drug-induced psychosis. The interested reader should examine Schuckit's fine clinical guide to the treatment of drug and alcohol abuse (Schuckit, 1995) or the *Practitioner's guide to psychoactive drugs* (Gelenberg & Bassuk, 1997). Both volumes also review the psychopharmacological profile for all the major drugs of abuse.

1.06.6.9.4 Alcoholism

About 70% of the US adult population consumes alcohol. Problem drinkers number about 10 million and half of these may be physically dependent. Risks for developing alcoholism include sociodemographic factors (lower socioeconomic status, ethnicity, being divorced, nationality), genetic factors (being male, family history of alcoholism), and individual history factors (younger age, antisocial personality, school and conduct problems). Chronic heavy drinking, of course, can lead to certain forms of CNS damage, including Wernicke's disease, Korsakoff's psychosis, and Wernicke–Korsakoff syndrome. Postmortem studies of the latter patients show lesions in the mammillary bodies of the hypothalamus. These disorders appear to be produced by a B_1 vitamin deficiency and are treated with thiamine. Further, alcohol can produce profound CNS or other organ systems during development, for example, the fetal alcohol syndrome.

Alcohol (ethyl alcohol or ethanol), when absorbed, uniformly reached all tissues in the body, including the CNS. Besides the CNS damage, many other organ systems may be compromised by chronic alcohol, for example, the liver and heart. While an alcohol receptor may not exist, several NT systems may mediate some direct acute or chronic effects of alcohol and its acute (hangover) or chronic withdrawal syndrome. Alcohol enhances inhibitory $GABA_A$ neurotransmission, promotes the release of DA in mesolimbic pathways associated with reinforcement and pleasure, decreases how much 5-HT is stored in neurons of the dorsal raphe nucleus, and appears to reduce excitatory neurotransmission at the NMDA form of the glutamate receptor. The liver converts alcohol to acetaldehyde (enzyme-alcohol dehydrogenase), which may promote the release of and interaction with CAs from stores in the adrenals and in the brain. Apparently acetaldehyde and biogenic amines chemically interact to produce opiate-like substances (tetrahydroisoquinolines, TIQs, and beta-carbolines). TIQs appear to promote DA release in the nucleus accumbens (Robert Myers, personal communication) and the beta-carbolines may act as inverse agonists at the benzodiazepine site on $GABA_A$ receptors. Such an inverse agonist action would reduce the impact of drugs that enhance $GABA_A$ activity, for example, BZDs, barbiturates, and alcohol. The latter effect is interesting since animal research suggests that the $GABA_A$ system of animals more likely to self-administer alcohol is less sensitive to the effects of alcohol (Crabbe & Harris, 1991).

Given the opponent-process model of drug rebound and the array of direct CNS effects of alcohol, it is not surprising that alcohol withdrawal signs include excitation, anxiety, and discomfort (note these signs are opposite from alcohol's direct sedative, calming, and euphoric effects). Other characteristics of the prolonged alcohol withdrawal syndrome in physically dependent individuals include insomnia, tremulousness, muscle tension, facial flushing, and sweating, followed by nausea, vomiting, abdominal cramps, hyperreflexia, and anorexia. Later during withdrawal, tremor, seizures, and hallucinations may develop, followed sometimes by delirium tremens (confusion, hallucinations, disorientation, agitation, and delirium). Sedative-hypnotics, like barbiturates and BZDs, have been traditionally used to treat the withdrawal syndrome, and more recently alpha- and beta-adrenergic blockers. The persistent hallucinations after prolonged drinking bouts probably represent an organic condition that can be treated with antipsychotic drugs like the phenothiazines or butyrophenones.

Until recently, pharmacological treatment for alcoholism itself has been limited to disulfiram (Antabuse). That drug blocks the metabolism of acetaldehyde, producing a buildup of acetaldehyde and a plethora of unpleasant autonomic, somatic, subjective and CNS effects. This treatment is generally ineffective because of poor compliance (about 20% remain alcohol-free for a year). In cases of dual-diagnosed alcoholics, treatment of the psychiatric component may improve the substance as well, for example, lithium in bipolar alcoholics. Newer developments include the use of naltraxone to reduce alcohol "craving." The latter is thought to work through the endogenous opiate system interactions with the mesolimbic DA reward system. Similarly, drugs that boost brain 5-HT levels (e.g., SSRIs) have a modest effect in reducing alcohol consumption. Alcoholics receiving buspirone, a partial 5-HT$_{1A}$ agonist, show reduced consumption and anxiety, but further research is needed. Other compounds that may be useful in reducing drinking and/or craving include (Bloom & Kupfer, 1995): acamprosate (a direct GABA receptor agonist), bromocriptine (a nonspecific dopamine-2 receptor agonist), tiapride (a dopamine-2 receptor antagonist), and γ-hydroxybutyrate (an endogenous sedative that increases DA levels).

1.06.6.9.5 Cocaine and amphetamine dependence

Besides marijuana, cocaine and amphetamines represent the most widely abused illegal drugs in the USA. The interested reader can consult any number of references for more details (Gelenberg & Bassuk, 1997; Schuckit, 1995). Both are CNS and cardiovascular stimulants (vasoconstriction), acting in different ways to promote the increase of DA levels, and, to a lesser extent, NE and 5-HT levels. Although cocaine is a local anesthetic, its stimulatory effects derive from its blockade of DA and NE reuptake into the presynaptic neuron. Amphetamine acts primarily to release DA or NE from the presynaptic neuron, but also has reuptake blocking properties. Their abuse potential lies in their activation of the brain reward system, that is, the DA receptors of the mesolimbic nucleus accumbens. Both agents at low to moderate doses produce a positive symptom profile: including euphoria; increased wakefulness, concentration, and alertness; and decreases in fatigue and appetite. However, higher doses can produce a host of negative effects, including tremor, irritability, emotional lability, anxiety, motor tremors, paranoia, panic, and repetitive stereotypes in movements and speech. Overdosing can result in stroke, heart failure, and seizures. The withdrawal profile of these drugs also is similar. Intense craving, agitation, and anxiety are followed by fatigue, by increased sleeping and eating, and by depression, then by listlessness, anhedonia (inability to experience pleasure), decreased interest, and increased drug craving. Longtime cocaine addicts also may show sensitization or reverse tolerance where even low doses can induce bizarre behavioral reactions that appear similar to those of a paranoid psychosis.

The clinical psychopharmacology of amphetamine and cocaine also is similar (see Schatzberg & Nemeroff, 1995). Acute intoxication with either agent can be treated with DA-blocking drugs like chlorpromazine or haloperidol. Both drugs continue to have legitimate medical uses, for example, cocaine is used in ophthalmology and amphetamines are used to treat narcolepsy, ADHD, and exogenous obesity. However, most of the clinical research on agents to counteract stimulant addiction has focused, with uneven success, on cocaine. One "quasi-substitution" strategy has been to use antidepressive drugs like desipramine, which also block DA reuptake. Only modest reductions in the cocaine relapse rates were found. A similar approach using bromocriptine, a DA agonist, looks promising, but double-blind clinical trials are needed. Finally, carbamazepine, an antiseizure medication, has been tried since cocaine appeared to increase "kindled" seizure activity in animal studies. Unfortunately, double-blind studies have yet to show any reduction in relapses back to cocaine use.

1.06.6.9.6 Nicotine and caffeine dependence

The nicotine delivered through cigarette smoking acts directly on nicotinic cholinergic receptors in the CNS (Schuckit, 1995). The reinforcing and other actions of nicotine (mood elevation, enhancement of cognitions, psychomotor stimulation, and decreased appetite) are similar to those of cocaine and amphetamine in that all activate mesolimbic DA neurons. However, the "high" produced by smoking is shorter acting since nicotine apparently briefly shuts down the nicotinic receptor shortly after binding to it. Smokers may up-regulate their nicotine receptors to compensate for the frequent "shutdowns." Further, this latter neuroadaptation may be the basis for nicotine tolerance, dependence, and withdrawal (intense craving, agitation, increased appetite) phenomena. Nicotine also appears to increase NE activity, but to decrease 5-HT levels and turnover. At high doses, nicotine can produce toxic effects. The best current pharmacological treatment of smoking addiction is a "tapering-off" substitution regimen using nicotine gum, transdermal patches, and the recently developed nasal spray (the latter is not currently available).

The caffeine found in coffee, tea, cocoa, colas and over-the-counter medication is a member of the class of compounds called xanthines, which also include theophylline (tea, respiratory medications) and theobromine (chocolate).

Caffeine acts to produce stimulatory effects in multiple organ systems, for example, cardiovascular, respiratory, kidney, and CNS. In the brain, caffeine may block GABA and adenosine receptors (activation of these normally acts to inhibit other neurons) and may increase turnover in CA and 5-HT systems (Shuckit, 1995). "Caffeinism," produced by a daily consumption greater than 500–600 mg, presents a set of symptoms resembling panic attacks. Caffeine dependence is evidenced by its withdrawal effects, which include headaches, increased muscle tension, fatigue, irritability, anxiety, and poor concentration. No specific pharmacological treatment is available other than gradually tapering-off caffeinated products.

1.06.6.9.7 Opiate (narcotic) dependence

Section 1.06.2 described the neuropharmacology of opiate receptors. Different opiate narcotic drugs have different profiles of agonist and/or antagonist actions at mu, kappa, delta, and sigma receptors (see Julien, 1997). The four general categories of abused opiate drugs are: (i) natural alkaloids of opium (opium, morphine, and codeine), (ii) semisynthetic derivatives of morphine (heroin, hydromorphone or Dilaudid, and oxycodone or Percodan), (iii) nonmorphine-derived synthetics (meperidine or Demerol, methadone, and propoxyphene or Darvon, and (iv) opioid-containing preparations (usually with codeine or paregoric). Most of the latter compounds produce their psychotropic effects by acting on mu opiate receptors, which in turn act on other brainstem systems (e.g., the VTA DA reward pathway). All, to a lesser or greater extent, produce analgesia, euphoria, sedation, respiratory depression, cough suppression, pupillary constriction, nausea and vomiting, and a host of other gastrointestinal symptoms.

Nonmedical opioid use of any kind in the USA ranges from 9% in women to 15% in men (for trying heroin, 0.5% and 2%). The acute effects, described as an orgasmic-like rush, last about a minute, followed by a tranquil state lasting for several hours (drowsiness, apathy, motor retardation, mental clouding, depressed respiration, and mood lability). The pupillary constriction and constipation effects of heroin are the most resistant to tolerance development. Commonly, overdosing and death result from the respiratory depression (naloxone and naltrexone, opiate receptor antagonists, are used to treat overdosing). Of course, the opioid withdrawal syndrome, peaking between 48 and 72 hours after last dose, is opposite in character, that is, anxiety, dilated pupils, diarrhea, insomnia, agitation, runny nose and eyes, sweating, and so forth.

Alpha adrenergic agonists like clonidine are sometimes used to treat the overactivation of the sympathetic nervous system during opiate withdrawal. However, more commonly, another opiate agonist like methadone, Darvon, or Demerol or a mixed agonist–antagonist like buprenorphine, is employed to attenuate the withdrawal symptoms. Pharmacological treatment of the opiate addition itself employs one of three strategies. First, opiate antagonists, like naloxone or naltrexone, have proven effective in the well-motivated addict but generally these agents present problems in compliance. Substitution therapies involve the maintenance on some prescribed opiate agonists like methadone or LAAM (a long-acting agonist). The rationale for this approach is that the regimen blunts the addict's craving, reducing criminal and other risky behaviors (prostitution, HIV-infected needles). The rate of relapse one year after heroin or methadone detoxification can run as high as 80–90%. A newer approach has been to use the mixed agonist–antagonist, buprenorphine. The rationale here is that the agonist effects will blunt the craving and drug-seeking behaviors and the antagonist effects will blunt the heroin high produced by any relapses. Clinical trials of this compound are underway.

1.06.6.9.8 Sedative-hypnotic and anxiolytic dependence

These drug abuse problems involve benzodiazepine (BZD) and non-BZD compounds that, with varying potencies, produce CNS depressant effects, ranging from lethargy and sleepiness to anesthesia, and to death from respiratory and heart depression (see Schuckit, 1995). BZDs, commonly used to treat anxiety and/or withdrawal from other drugs, include chlordiazepoxide (Librium), diazepam (Valium), alprazolam (Xanax), and triazolam (Halcion).

Non-BZD CNS depressants include short-acting barbiturates (used as general anesthetics and rarely abused), longer-acting barbiturates like pentobarbital (used as hypnotics to induce sleep and often abused), and long-acting barbiturates like phenobarbital (used to treat epilepsy and other seizure disorders and not usually abused). Other barbiturate-like drugs (e.g., methaqualone or Quããlude) were introduced as hypnotics that lacked the sleep disturbances produced with barbiturate hypnotics. Unfortunately, some, like Quããlude, are widely abused. Other non-BZD CNS depressants include chloral hydrate (used as a short-acting sedative-hypnotic) and meprobamate (used to treat anxiety), and both have high abuse potential.

The specific mechanisms by which all these drugs exert their CNS depressive actions are not completely clear, but most are known to suppress all excitable tissue. The mechanism by which some of these drugs exert their anxiolytic, muscle relaxant, and anticonvulsant actions lies in their affinity for BZD or barbiturate receptors on the $GABA_A$ receptor complex. There they act as allosteric modulators of the chloride ion channel and enhanced inhibitory, GABAergic transmission. Via GABAergic input to the nucleus accumbens, the latter drug actions may produce the reinforcing properties of euphoria or tranquility that lead to abuse of these compounds. Pharmacodynamic tolerance to the latter effects develops quickly with both barbiturates and BZDs, probably through some neuroadaptation involving desensitization of the receptors.

With continuous usage of BZDs or barbiturates and then abrupt cessation, a withdrawal profile characterized as rebound hyperexcitability emerges (see Schuckit, 1995). Barbiturate withdrawal frequently produces seizure, while the BZD withdrawal pattern resembles that of alcohol withdrawal. Not surprising, the treatment of depressant withdrawal symptoms is usually a drug from the same class of drug. Of all the drugs of abuse, depressant drugs pose the most serious problems because of their many medical uses, their potential for abuse, and their potential for overdosing or drug interactions. CNS depressive drugs interact with each other, with alcohol, and with other psychiatric drugs (e.g., tricyclic antidepressives) to produce sedation. Thus, these medications are frequently used in suicide attempts. Fortunately, at least for BZDs, a receptor blocker exists (flumazanil) and can be used for acute BZD intoxication or overdose. Since no effective blocker for non-BZDs exits, other medical means must be considered for barbiturate overdoses (e.g., hemodialysis).

Treatment of depressant, and particularly BZD, addiction is complicated by the fact that the drug is often prescribed for anxiety. Thus, besides BZD withdrawal, the possibility of rebound-anxiety exists. An alternative strategy would be to taper off BZDs, slowly and in small steps. However, the latter approach may be problematic when other factors are present, for example, underlying anxiety disorder, any other drug use (including caffeine, nicotine, and alcohol), premenstrual syndrome, life stressors, or other medical conditions (Stahl, 1996).

1.06.6.9.9 Marijuana abuse

Marijuana and hashish (from the flowering tops of the marijuana plant) derive their effects from the plant's most active ingredient, THC (delta-9-tetrahydrocannabinol). Use of THC dates back almost three millennia and today it is the most frequently used illegal drug. In the 1970s and 1980s, more than 50% of high school seniors used marijuana at least once. Marijuana usually has both sedative and stimulatory acute effects. These include euphoria (including a sense of well-being, relaxation, and friendliness), changes in one's level of consciousness without hallucinations (loss of temporal awareness), feelings of heightened sexual arousal and hunger, mental slowing and sleepiness, and short-term memory loss. However, at high doses, it can produce panic, delirium, and sometimes psychosis. Chronic marijuana use is associated with an amotivational syndrome (decreased drive and ambition), impairment of social and/or performance skills (i.e., poor communication, ineffective interpersonal interactions, and introversion, and/or poor judgment, inattention, and distractibility), and a decline in personal habits and insight (Schuckit, 1995; Stahl, 1996).

Tolerance may develop to THC, and sudden cessation of THC after chronic use may produce withdrawal symptoms, including: a hyperarousal profile (insomnia, anxiety, restlessness, photophobia, and irritability), somatic and autonomic symptoms (myalgia, chills, sweating, diarrhea, yawning), and a mixed depressive profile (depression, craving, anergy, and mental confusion) (Gelenberg & Bassuk, 1997).

THC can be administered any number of ways (though usually smoked). However, its complete elimination may be slow (up to a week or more) since both THC and its metabolites are sequestered in fat deposits for later release. The mechanism by which THC produces its effects is still poorly understood, but some data suggest activation of the septal area of the brain's limbic system (known for its involvement in emotion). The acute memory impairment and long-lasting EEG changes also suggest clear CNS actions. Some evidence suggests that the brain contains THC receptors and perhaps endogenous THC-like ligands. In any event, no clear medication treatment has been tried for THC abuse, other than detoxification and long-term abstinence. Treatment of symptoms includes anxiolytics for anxiety, panic, or flashbacks and antipsychotics for the occasional toxic delirium, psychotic reactions, or hallucinations (Schuckit, 1995).

1.06.6.9.10 Abuse of hallucinogens and related drugs

Hallucinogens include a group of drugs that produce intoxicating effects, described as "highs," characterized by enhanced sensory and perceptual experiences (both exteroceptive and interoceptive) (Stahl, 1996). Sometimes, these effects result in visual illusions and hallucinations. The hallucinations may be psychedelic (i.e., enhanced sensory awareness, "mind expansion," and feelings of oneness with a larger "other") or psychotomimetic (i.e., superficially resembling a psychosis). These drugs produce visual hallucinations and strong emotional reactions, whereas schizophrenia is more typically associated with auditory hallucinations and flat affect (Schuckit, 1995). Initially, the latter symptoms can occur without consciousness changes or confusion. However, the intoxication can progress from a confusional state (disorientation and agitation) to anxiety, depression, and panic, and to toxic delirium and the psychotic symptoms of delusions and paranoia. The panic-like state can include physiological and subjective effects like anxiety and feelings of mind loss. Even after cessation of use, flashbacks may occur.

Schuckit (1995) described four general categories of hallucinogenic agents. Indolealklamines like LSD or lysergic acid diethylamide (a synthetic drug) and Psilocybin (an extract of a mushroom plant) appear to act as partial agonists at 5-HT receptors, particularly 5-HT_2. Phenylethylamine drugs, like mescaline or peyote, and phenylisopropylamine drugs (so called "designer drugs"), like methylene dioxymethamphetamine (MDMA or "ecstasy"), have structures and actions more similar to biogenic amines and amphetamine. Finally, a group of related drugs also can produce hallucinations; these include phencyclidine (PCP or "angel dust") and inhalants (e.g., amyl or butyl nitrites). PCP appears to have a specific mechanism of action, that is, it acts allosterically to block the NMDA type of glutamate receptor by reducing calcium ion influx into the cell. The latter action disrupts normal and necessary excitatory neurotransmission, thereby producing skeletomuscular effects (besides other effects described above) that include catatonia (excitement alternating with catalepsy and stupor), motor seizures, and muscle breakdown. Other commonly abused inhalants include glues, solvents, aerosols, cleaning solutions, and gasoline. While these substances can produce euphoria and lightheadedness, overdosing can produce loss of normal cardiac function, hypoxia, seizures, and death. Tolerance to the effects of inhalants develops rapidly, and hallucinations can occur with abrupt discontinuation after prolonged use. Chronic inhalant abuse is thought to produce brain damage associated with impairment of higher cognitive functions.

Medical management of hallucinations is generally the same as for psychosis, that is, use of antipsychotics. None of the drugs in this class have any current medical use. Some patterns of abuse may be short-lived because of the unpleasant, dysphoric effects of the drugs. However, some drugs, like PCP, are associated with a very treatment-resistant pattern of drug use.

1.06.6.9.11 Over-the-counter and prescribed drugs

We conclude this section by indicating that many legal drugs can be abused. These include analgesics, antihistamines, cold and cough remedies, sleep medications, and caffeinated products. Some research shows that certain combinations of these products produced subjective effects similar to those produced by cocaine, amphetamine, and morphine.

1.06.6.10 Eating Disorders

There are many theories to explain the etiology of anorexia nervosa and bulimia nervosa, ranging from cultural to psychological to biological, but none have emerged as dominant in the field. Most of the biologically based theories hypothesize that people with eating disorders lack certain mechanisms related to feeding, such as appetite, satiety, and desire for balanced nutritional intake. Because depression is so often associated with eating disorders, however, the first medications used were antidepressants.

Antidepressants continue to be the psychopharmacological treatment of choice. Both the traditional tricyclics, such as imipramine and desipramine, and the SSRIs, particularly fluoxetine, have been useful. Recent studies suggest that both types of antidepressants relieve depressive symptoms and eating disorder symptoms in both anorexia and bulimia. The primary outcome measure in people with anorexia is weight gain. Although tricyclics may cause weight gain in other patients and SSRIs may cause weight loss, anorectics taking either class of drug gain weight at equivalent rates. Bulimics also respond well to either drug, although as measured by different outcomes, such as binging and purging behaviors. It is not clear how two classes of drugs so dissimilar in their mechanism of action can produce such similar results.

Studies of other medications are underway. Some researchers who postulate that eating disorders are similar to opiate dependence have used naltrexone with some success. Others are investigating endocrine responses in eating disordered patients, with plans to develop and test new medications in the future.

1.06.6.11 Neurological Disorders

Because of limitations of space, not all neurological disorders can be covered in this section. However, the reader should be aware that many disorders are age-related. Further, medication in elderly individuals presents a unique set of problems and concerns. The reader is directed to a chapter on geriatric psychopharmacology (Gelenberg & Bassuk, 1997).

1.06.6.11.1 Parkinson's disease

Parkinson's disease, a neurodegenerative disorder characterized by resting tremor, muscular rigidity, and difficulty initiating motor activity, is the most common movement disorder, occurring in approximately 1% of the population. The best current theory for PD rests on evidence of the degeneration of pigmented DA cells in the substantia nigra area of the brainstem (see Kandel et al., 1991). These DA neurons normally project to the striatum of the basal ganglia (part of the extrapyramidal motor control system), primarily in the nigrostriatal pathway. This loss or injury to DA neurons results in an associated loss of striatal DA. In Parkinson patients, striatal DA depletion can range from 70 to more than 90% with lesser losses in other DA and non-DA systems. The loss of midbrain DA is associated with the loss of DA transporters in the striatum. Thus, any DA remaining in the synapse is metabolized rather than being taken back up into the terminal. Persistent lowered levels of DA results in an increase in D_1 and D_2 receptor density (upregulation) in the caudate and putamen areas of unmedicated Parkinson patients. Treatment with L-DOPA (see below) returns receptor densities to normal levels. Other data are consistent with this DA loss account of Parkinson's disease. In 1983, heroin addicts, attempting to synthesize their own drug, accidentally made MPTP. This compound, which selectively destroys nigrostriatal DA neurons, produced a Parkinson-like syndrome in the addicts and provided researchers with excellent support of the DA theory.

The best current pharmacotherapy for Parkinson's disease is the DA precursor L-DOPA, usually coadministered with an inhibitor of enzymes that could destroy L-DOPA. Unlike DA itself, L-DOPA can pass the blood–brain barrier and can be converted into DA by the few remaining neurons. In particularly severe cases, DA agonists may prove effective, most commonly, bromocriptine, a D_2 agonist. Other drugs that act to increase DA levels (MAOIs and DA uptake inhibitors) also have been tried. One novel approach uses the drug selegiline, which is thought to work not through inhibition of DA breakdown but as an antioxidant to prevent neuronal death. Other, more experimental approaches, involve the transplantation of fetal adrenal cells directly into the basal ganglia (see Kandel et al., 1991).

1.06.6.11.2 Alzheimer's disease and other dementias

Alzheimer's disease is characterized by severe dysregulation of the acetylcholine pathways in the brain and development of neuritic plaques and neurofibrillary tangles. As with most other neuropsychiatric disorders, its etiology is not clear. Although the plaques and tangles are diagnostic, the cognitive deficits may be caused by the loss of cholinergic function, particularly that generated in the nucleus basalis. This disease causes degeneration of cholinergic neurons in the nucleus basalis, which is the primary ACh input to the cortex. People with the disorder also show reduced ACh activity in the hippocampus, which is thought to be the cause of the memory deficits. Researchers have also found reduced levels of somatostatin, neuropeptide Y, and corticotropin-releasing factor and reduced numbers of NMDA receptors and locus coeruleus neurons. However, the significance of these changes is not yet known. Alzheimer's is associated with Down's syndrome, as most individuals with Down's who live to the upper reaches of the age range develop Alzheimer's. The meaning of this correlation is also unclear. Research is also underway into the cause and meaning of the excessive amounts of β-amyloid protein found in the neurofibrillary tangles of Alzheimer's patients.

Psychopharmacological treatments have focused on the ACh deficit. Researchers have tried giving choline, physostigmine (to inhibit acetylcholinesterase), and the muscarinic agonist arecoline, trying to boost the production or effectiveness of the patient's natural ACh. In the late 1990s, the treatment of choice is tacrine, an anticholinesterase. Tacrine inhibits the breakdown of ACh that increases the amount available for use. However, this only works if enough ACh is being released. Currently, no treatment to stop the degradation of ACh neurons exists. Researchers outside the USA have tried to transplant fetal neural tissue to replace the ACh neurons destroyed by Alzheimer's. Although it is not a psychopharmacological treatment, it is mentioned here because it is designed to treat the disorder directly. However, results have been equivocal and not adequately replicated. Fetal neural transplants are still very controversial, despite the US government's decision to allow fetal tissue use in research.

The interested reader is referred to an excellent chapter by Stahl (1996), which summarizes strategies and data regarding cognitive enhancers and neuroprotective agents.

Vascular dementia is differentiated from other dementias by its sudden onset and stepwise deterioration. Treatment is similar to that for a single stroke. Persons with the disorder may be prescribed cognition enhancers (e.g., free-radical scavengers like vitamin E) and anticoagulants (e.g., aspirin) to relieve symptoms and reduce the chance of another stroke.

1.06.6.11.3 Huntington's disease (HD)

HD, formerly known as Huntington's chorea, is included here primarily because, although defined as a movement disorder, it entails progressive degeneration of cognition and emotion. In some people, the disorder may manifest itself first as a psychiatric disorder or dementia. It is a genetic disorder, transmitted by a gene on the short arm of chromosome 4. The disorder causes a particular type of cholinergic interneuron in the striatum to degenerate. Although researchers do not know why these neurons die, some have suggested that they die of the excitotoxic effects of glutamate or one of the other excitatory amino acid NTs. This results in the disinhibition of the DA neurons in the substantia nigra, which causes the involuntary movements known as chorea.

Unfortunately, a curative treatment for Huntington's disease is not available, but progress should be more rapid now that the gene locus has been identified. Psychopharmacological treatments typically attempt to correct the striatal system imbalance. Treatments include medications such as physostigmine, which is an acetylcholinesterase inhibitor, to enhance the ACh available for use. In addition, DA antagonists like the typical antipsychotics can help to inhibit the overactive DA systems and partially relieve the symptoms. To protect the cholinergic interneurons, NMDA receptor antagonists (to block glutamate) like amantadine and calcium-channel blockers like nimodipine have also been tried, although they are still quite experimental. As with the other movement disorders, fetal striatal transplants have been suggested to replace the neurons lost to the disorder.

1.06.6.12 Developmental Disorders Originating in Childhood

1.06.6.12.1 Attention-deficit hyperactivity disorder (ADHD)

ADHD occurs in approximately 5% of the school-age population and is characterized by inattention, impulsivity/distractibility, and hyperactivity. Researchers do not have a clear theory of its etiology. The most popular current theory is that children with ADHD have some deficits in right hemispheric function, particularly in the prefrontal-striatal circuitry (Heilman, Voeller, & Nadeau, 1991). Several neuropsychological and imaging studies have supported this hypothesis.

However, psychopharmacological treatments of ADHD do not address the hypothesized right hemisphere dysfunction directly. Amphetamine was first used in 1937 to treat ADHD, and it and amphetamine-like drugs continue to be the most popular medications. Medications used in the late 1990s include dextroamphetamine (Dexedrine), methylphenidate (Ritalin), and pemoline (Cylert), all of which improve attention and reduce hyperactivity in children with ADHD. Side-effects include decreased appetite, increased heart rate and blood pressure, insomnia, and tics. ADHD is often comorbid with Tourette's syndrome, and stimulants may worsen or reveal Tourette's-like vocal and motor tics. Pemoline can cause liver dysfunction, and a child on pemoline should have liver function tests every few months to rule it out. Children with ADHD who do not respond to stimulants may need a trial of tricyclic antidepressants such as imipramine and desipramine. Antidepressants are also recommended for children with a comorbid anxiety or depressive disorder.

1.06.6.12.2 Other disorders

For reasons of space and lack of a clear-cut psychopharmacological strategy, most other developmental disorders will not be addressed in this chapter (e.g., mental retardation, autism, Rett's disorder, childhood disintegrative disorder, and Asperger's disorder). However, the reader should be aware that data on pediatric psychopharmacology is limited relative to that for adults (see Gelenberg & Bassuk, 1997, for a review of pediatric psychopharmacology).

1.06.7 REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.) Washington, DC: Author.

Balfour, J. K. (Ed.) (1990). *Psychotropic drugs of abuse.* New York: Pergamon.

Bloom, F. E., & Kupfer, D. J. (Eds.) (1995). *Psychopharmacology: The fourth generation of progress.* New York: Raven Press.

Blum, K., Cull, J. G., Braverman, E. R., & Comings, D. E. (1996). Reward deficiency syndrome. *American Scientist, 84,* 132–145.

Carlson, N. R. (1994). *Physiology of behavior* (5th ed.) Boston: Allyn and Bacon.

Carlton, P. L. (1983). *A primer of behavioral pharmacology*. New York: W. H. Freeman.

Colpaert, F. C. (1986). Drug discrimination: Behavioral, pharmacological, and molecular mechanisms of discriminative drug effects. In S. R. Goldberg & I. P. Stollerman (Eds.), *Behavioral analysis of drug dependence* (pp. 161–193). Orlando, FL: Academic Press.

Cooper, J. R., Bloom, F. E., & Roth, R. H. (1996). *The biochemical basis of neuropharmacology* (7th ed.). New York: Oxford University Press.

Crabbe, J. C., & Harris, R. A. (Eds.) (1991). *The genetic basis of alcohol and drug actions*. New York: Plenum.

Dworkin, S. I., & Smith, J. E. (1987). Neurobiological aspects of drug-seeking behaviors. In T. Thompson, P. B. Dews, & J. E. Barrett (Eds.), *Advances in behavioral pharmacology: Neurobehavioral pharmacology* (vol. 6, pp. 1–43). Hillsdale, NJ: Erlbaum.

Dwoskin, L. P., Neal, B. S., & Sparber, S. B. (1987). A behavioral assay for receptor action(s) of drugs. In T. Thompson, P. B. Dews, & J. E. Barrett (Eds.), *Advances in behavioral pharmacology: Neurobehavioral pharmacology* (Vol. 6, pp. 45–81). Hillsdale, NJ: Erlbaum.

Dyskstra, L. (1992). Drug action. In J. Grabowski & G. R. VandenBos (Eds.), *Psychopharmacology: Basic mechanisms and applied interventions* (pp. 59–96). Washington, DC: American Psychological Association.

Feldman, R. S., Meyer, J. S., & Quenzer, L. F. (1997). *Principles of neuropsychopharmacology*. Sunderland, MA.: Sinauer Associates.

Gelenberg, A. J., & Bassuck, E. L. (1997). *The practicioner's guide to psychoactive drugs* (4th ed.). New York: Plenum.

Goldberg, S. R., & Stolerman, I. P. (Eds.) (1986). *Behavioral analysis of drug dependence*. Orlando, FL: Academic Press.

Goudie, A. J., & Demellweek, C. (1986). Conditioning factors in drug tolerance. In S. R. Goldberg & I. P. Stollerman (Eds.), *Behavioral analysis of drug dependence* (pp. 225–285). Orlando, FL: Academic Press.

Grabowski, J., & VandenBos, G. R. (Eds.) (1992). *Psychopharmacology: Basic mechanisms and applied interventions*. Washington, DC: American Psychological Association.

Hardman, J. G., & Limbird, L. E. (Eds.) (1996). *Goodman & Gilman's the pharmacological basis of therapeutics* (9th ed.) New York: McGraw-Hill.

Heilman, K. M., Voeller, K. K. S., & Nadeau, S. E. (1991). A possible pathophysiologic substrate of attention deficit hyperactivity disorder. *Journal of Child Neurology, 6*(Suppl.), 576–581.

Henningfield, J. E., Lukas, S. E., & Bigelow, G. E. (1986). Human studies of drugs as reinforcers. In S. R. Goldberg & I. P. Stollerman (Eds.), *Behavioral analysis of drug dependence* (pp. 69–122). Orlando, FL: Academic Press.

Henningfield, J. E., Schuh, L. M., & Jarvik, M. E. (1995). Pathophysiology of tobacco dependence. In F. E. Bloom & D. J. Kupfer (Eds.), *Psychopharmacology: The fourth generation of progress* (pp. 1715–1729). New York: Raven Press.

Heuser, I., Yassouridis, A., & Holsboer, F. (1994). The combined dexanethasone/CRH test: A refined laboratory test for psychiatric disorders. *Journal of Psychiatric Research, 28*(4), 341–356.

Holloway, F. A., & King, D. A. (1989). Parallel development of ethanol tolerance and compensatory behaviors. *Pharmacology, Biochemistry & Behavior, 34,* 855–861.

Hughes, J. R., & Pierattini, R. A. (1992). In J. Grabowski & G. R. VandenBos (Eds.), *Psychopharmacology: Basic mechanisms and applied interventions* (pp. 97–125). Washington, DC: American Psychological Association.

Johanson, C.-E. (1992). Biochemical mechanisms of pharmacological principles of drug action. In J. Grabowski & G. R. VandenBos (Eds.), *Psychopharmacology: basic mechanisms and applied interventions* (pp. 11–58). Washington, DC: American Psychological Association.

Julien, R. M. (1998) *A Primer of Drug Action* (8th ed.) New York: W. H. Freeman.

Kalivas, P. W., & Samson, H. H. (1992). The *neurobiology of drug and alcohol addiction*. New York: The New York Academy of Sciences.

Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (Eds.) (1991). *Principles of neural science* (3rd ed.) New York: Elsevier.

Kaplan, H. I., & Sadock, B. J. (1996). *Comprehensive textbook of psychiatry* (6th ed.). Baltimore: Williams and Wilkins.

Leonard, B. E. (1992). *Fundamentals of psychopharmacology*. New York: Wiley.

Marlatt, A. G. (1992). Substance abuse: Implications of a biopsychosocial model for prevention, treatment, and relapse prevention. In J. Grabowski & G. R. VandenBos (Eds.), *Psychopharmacology: Basic mechanisms and applied interventions* (pp. 127–162). Washington, DC: American Psychological Association.

McKim, W. A. (1986). *Drugs and behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Schatzberg, A. F., & Nemeroff, C. B. (Eds.) (1995). *Textbook of psychopharmacology*. Washington, DC: American Psychiatric Press.

Schildkraut, J. J. (1965). The catecholamine hypothesis of affective disorders. *American Journal Psychiatry, 122,* 509–522.

Schuckit, M. A. (1995). *Drug and alcohol abuse*: *A clinical guide to diagnosis and treatment* (4th ed.). New York: Plenum.

Schuster, C. R. (1978). Theoretical basis of behavioral tolerance: Implications for problems of drug abuse. In N. A. Krasnegor (Ed.), *Behavioral tolerance*: *Research and treatment implications*. (NIDA Research Monograph No. 18). Washington DC: US Government Printing Office.

Siegel, S. (1990). Classical conditioning and opiate tolerance and withdrawal. In J. K. Balfour (Ed.), *Psychotropic drugs of abuse*. (pp. 59–86). New York: Pergamon.

Smith, P. F., & Darlington, C. L. (1996). *Clinical psychopharmacology: A primer*. Mahwah, NJ: Erlbaum.

Stahl, S. M. (1996). *Essential psychopharmacology*. Cambridge, UK: Cambridge University Press.

Thompson, T., Dews, P. B., & Barratt, J. E. (Eds.) (1987). *Advances in behavioral pharmacology: Neurobehavioral pharmacology* (Vol. 6). Hillsdale, NJ: Erlbaum.

Wise, R. A. (1990). The role of reward pathways in the development of drug dependence. In J. K. Balfour (Ed.), *Psychotropic drugs of abuse* (pp. 23–58). New York: Pergamon.

World Health Organization. (1993) *International classification of diseases: Classification of mental and behavioral disorders: Clinical descriptions and diagnostic guidelines* (10th ed.) Geneva, Switzerland: Author.

Young, A. M., & Herling, S. (1986). Drugs as reinforcers: Studies in animals. In S. R. Goldberg & I. P. Stollerman (Eds.), *Behavioral analysis of drug dependence* (pp. 9–67). Orlando, FL: Academic Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.07

Animal Models of Psychopathology: Depression, Anxiety, Schizophrenia, Substance Abuse

PAUL WILLNER
University of Wales, Swansea, UK

1.07.1 INTRODUCTION 207
1.07.1.1 Validation of Animal Models 208
1.07.2 ANIMAL MODELS OF DEPRESSION 210
1.07.2.1 Clinical Considerations 210
1.07.2.2 Stress Models 210
1.07.2.3 Other Approaches 212
1.07.3 ANIMAL MODELS OF ANXIETY 213
1.07.3.1 Clinical Considerations 213
1.07.3.2 Conflict Models 214
1.07.3.3 Ethological Models 215
1.07.4 ANIMAL MODELS OF SCHIZOPHRENIA 216
1.07.4.1 Clinical Considerations 216
1.07.4.2 Dopamine Hyperactivity Models 217
1.07.4.3 Attentional Models 218
1.07.5 ANIMAL MODELS OF SUBSTANCE ABUSE 219
1.07.5.1 Drug Self-administration 219
1.07.5.2 Drugs as Positive and Negative Reinforcers 221
1.07.5.3 Limitations 222
1.07.6 OVERVIEW: VALIDITY AND UTILITY OF ANIMAL MODELS 223
1.07.7 SUMMARY 225
1.07.8 REFERENCES 226

1.07.1 INTRODUCTION

The concept of modeling psychopathology in animals is as old as the use of animals in psychological investigations; its roots may be found in the work of Pavlov, Watson, and Rayner (1920) and even earlier (Abramson & Seligman, 1977; Weiner, 1973). However, for many years, practical attempts to devise animal models were sporadic, *ad hoc*, and unconvincing (see Abramson & Seligman, 1977; Keehn, 1979). As a result, animal models of psychopathology were until recently viewed with justified suspicion. This situation has now changed, with the recognition that animal models can provide a means of investigating the neurobiological mechanisms underlying psychopathology: indeed, given the limitations of the investigational techniques currently available for use in human subjects, animal models represent the only means of asking many of the important questions. Animal models can also be of great value in the process of psychotropic drug development, and again, frequently represent the only viable method of predicting novel therapeutic actions. The development and acceptance of animal models may thus be seen as an adjunct to the concurrent growth of psychopharmacology and biological psychiatry.

For the purposes of this chapter, animal models may be defined as representing

> experimental preparations developed in one species for the purpose of studying phenomena occurring in another species. In the case of animal models of human psychopathology one seeks to develop syndromes in animals which resemble those in humans in certain ways, in order to study selected aspects of human psychopathology. (McKinney, 1984)

However, it should be noted in passing that this definition presents a rather limited view of the way in which animal models of psychopathology are actually used. Although the earlier animal models were developed primarily to elucidate psychological processes, animal models are used increasingly, and in some cases, exclusively, within the context of psychopharmacology. Within this context, animal models are used for three distinct purposes (Willner, 1990). In one class of model, which may be termed a behavioral bioassay, behavioral tests are used as a convenient means of measuring the status of an underlying brain mechanism: for example, the bizarre set of drug-induced behaviors known as the serotonin syndrome may be used to measure the responsiveness of serotonin receptors to agonist drugs. For the purposes of this chapter, we simply note the existence of behavioral biossays, but consider them no further. Behavioral models are also used as predictive screening tests within the pharmaceutical industry. In this context, pharmacological considerations are paramount, and the behavioral features of the model are of less importance. The present review will, out of necessity, make some reference to pharmacological studies, but they will not be emphasized. Finally, animal models are used as experimental tools to investigate or illuminate aspects of psychopathology. It is to this use of animal models, in which an attempt is made to simulate the disorder, that the above definition applies. This chapter is primarily concerned with simulation models, and presents an overview of animal models in the four areas most relevant to clinical psychopharmacology: depression, anxiety, schizophrenia, and substance abuse. First, we consider some general issues pertaining to the validity of animal models.

1.07.1.1 Validation of Animal Models

The systematic validation of an animal model is no different in principle from that of any other psychological device, such as a psychometric test (Vernon, 1963) or a psychiatric diagnosis (Carroll, 1989), and the same general approaches to validation are applicable: predictive validity means that performance in the test predicts performance in the condition being modelled; face validity means that there are phenomenological similarities between the two; and construct validity means that the model has a sound theoretical rationale (Willner, 1986, 1991). The systematic assessment of animal models against each of these sets of criteria allows an unbiased comparison of their strengths and weaknesses. In the paper that first considered the problem of validating animal models, and so formed a starting point for subsequent research, McKinney and Bunney (1969) suggested that to be valid, a model should resemble the disorder modeled in four respects: etiology, symptomatology, treatment, and physiological basis. The first two of these domains refer primarily to face validity and the third to predictive validity, while the fourth reflects aspects of construct validity. However, any of these substantive domains could in principle contribute to an evaluation of validity on each of the three dimensions. Before reviewing animal models of psychopathology, it is appropriate to discuss briefly the general principles underlying the concepts of face, predictive, and construct validity as applied to animal models.

The simple comparison of the features of the model with the features of the disorder is the starting point for an evaluation of face validity. Ideally, a model should resemble the disorder in a variety of respects, and the *DSM* checklist approach to psychiatric diagnosis (American Psychiatric Association, 1994) provides one useful starting point for enumerating the areas of potential comparison. If several points of similarity are demonstrable, then it is necessary to ask whether the cluster of symptoms identified forms a coherent grouping that might realistically be seen in a single patient, or whether they are drawn from a variety of diagnostic subgroups. Frequently, animal models focus on a single behavior. In that case, it is essential to assess the importance in the total clinical picture of the symptom modeled. If the behavior in the model consists simply of a change in locomotor activity, for example, this is likely to be of peripheral relevance. Similarly, the face validity of the model is less strongly supported if the symptom modeled is common to several different psychiatric disorders. A further requirement for face validity is that the similarity between behavior in the model and the clinical symptom modeled be demonstrated rather than assumed (Abramson & Seligman, 1977). As Abramson and Seligman observed, animals subjected to uncontrollable electric shocks are undoubtedly stressed, but if severe shock is the only stressor that can be shown to generate a particular set of behavioral or neurochemical changes, then the model has little relevance to the concept of stress as used in human contexts. The demonstration of similarity requires a thorough experimental analysis, which, sadly, is often lacking.

The concept of predictive validity implies that manipulations known to influence the pathological state should have similar effects in the model: thus, manipulations known to precipitate or exacerbate the disorder should precipitate or exacerbate the abnormalities displayed in the model, while manipulations known to relieve the disorder should normalize behavior in the model. In principle, therefore, questions of predictive validity can be addressed to a number of features of simulations, including their etiology and physiological basis. In practice, the predictive validity of animal models relevant to psychopathology is determined largely by their response to therapeutic drugs. In this context, a valid test should be sensitive and specific, it should respond to effective therapeutic agents, and should fail to respond to ineffective agents. Positive responses should occur at sensible doses, and should be demonstrable with a range of structurally diverse compounds, and where applicable, to nonpharmacological treatment modalities; it is sometimes possible to demonstrate a correlation between potency in the model and the clinical dose. Negative responses should be demonstrable with agents that cause behavioral changes similar to the therapeutic effect but by nonspecific actions (e.g., by changing locomotor activity).

A number of other factors that need to be considered in relation to drug effects in animal models are discussed elsewhere (Willner, 1991). However, one important fact that is sometimes overlooked, with misleading consequences, is that the clinical classification of drugs as active or inactive may sometimes be incorrect. Drugs thought to be active on the basis of early open trials are frequently found to be inactive in later well-controlled tests; conversely, a drug may appear to be inactive because the emergence of side effects prevents its administration at adequate dosages, a problem that is less likely to arise in an animal model. It follows that the failure of an animal model to predict accurately will tend to weigh against the model, but may sometimes call instead for a re-evaluation of the clinical wisdom.

In order to evaluate the theoretical rationale of an animal model (construct validity), we require a theoretical account of the disordered behavior in the model, a theoretical account of the disorder itself, and a means of bringing the two theories into alignment. This can only be done if the clinical theory occupies an appropriate framework, which uses terms and concepts applicable also to subhuman species. Clearly, the subjective dimension of psychopathology cannot be central to such a theory, since subjective phenomena in animals are for most practical purposes outside the realm of scientific discourse (Willer, 1991). However, at the level of the cognitive processes underlying psychopathology (Brewin, 1988), and the physiological and biochemical mechanisms that underly those cognitive processes, the possibility exists of constructing parallel theories (Willner, 1986, 1991). It follows from this analysis that the assessment of construct validity involves a number of relatively independent steps. On the one hand, the theoretical account of behavior in the animal model requires evaluation. For example, if an animal model of depression is conceptualized as a decreased ability to respond to rewards, then at the very least, it must be convincingly demonstrated that the decrease in rewarded behavior cannot be explained by, for example, nonspecific sedative effects. Similarly, an animal model of dementia must demonstrate that performance failures result from a disorder of learning or memory, rather than from nonspecific causes, and further work should seek to characterize the specific memory processes involved. Equally important, however, is the evaluation of the account offered of the clinical disorder, both at the level of cognitive analysis and at the level of the relationship between cognitive changes and subjective experience (Brewin, 1988; Willner, 1991). It will be clear that a detailed consideration of the human disorder forms an essential step in the evaluation of animal models, and that the relatively poor state of theoretical understanding of most psychopathologies places an upper limit on construct validity.

1.07.2 ANIMAL MODELS OF DEPRESSION

1.07.2.1 Clinical Considerations

Depression is a heterogeneous disorder, which presents with varying patterns of symptomatology, a fact often overlooked by clinical psychologists. Perhaps the most important distinction is between endogenous and nonendogenous depressions, which differ in their responsiveness to psychological intervention (not, as frequently assumed, in the presence or absence of psychological precipitants: see e.g., Lewinsohn, Zeiss, Zeiss, & Haller, 1977; Paykel, 1979). Endogenous depression, which corresponds to the *DSM-IV* diagnosis of melancholia, is defined by an inability to experience pleasure (anhedonia), and may be conceptualized as a decrease in the sensitivity of brain reward systems (Klein, 1974). Melancholia is characterized by psychomotor retardation and a number of biological markers, responds well to antidepressants, and is a particular target for animal models.

Inevitably, animal models must focus on behavioral and physiological symptoms. Thus, anhedonia (decreased response to rewards), loss of appetite, and sex drive are readily modeled, and loss of interest may be modeled as a decrease in motivation, but such symptoms as feelings of worthlessness and guilt, or thoughts of suicide, fall outside the scope of animal models. Although psychomotor change, which is the most characteristic symptom of severe depression (Nelson & Charney, 1981), is typically modeled as a decrease (retardation) or increase (agitation) in locomotor activity, the clinical phenomena are actually considerably more complex, and there must be serious doubt about the face validity of models that rely exclusively on gross changes in locomotor activity. Depression is effectively treated by a wide variety of chemically and functionally distinct drugs, and it is therefore inevitable that pharmacological considerations also play an important role in the validation of animal models of depression. In addition to questions of generality and specificity of drug effects in an animal model, two features of antidepressant drug action are particularly important: that these drugs have a slow onset of action, typically requiring 2–5 weeks of treatment, and that they fail to elevate mood in nondepressed individuals.

1.07.2.2 Stress Models

Animal models of depression have been reviewed by Katz (1981), Jesberger and Richardson (1985), and Willner (1984, 1990). The largest group of animal models of depression consists of paradigms based on the application of stress. In a simple and widely-used model, sometimes known as the "behavioral despair" test, rats or mice are forced to swim in a confined space; antidepressants are found to increase the duration of active struggling, or conversely, to decrease the duration of immobility which develops later in the test (Porsolt, 1981; Porsolt, LePichon, & Jalfre, 1977). This model has been used extensively in drug development, and functions with reasonable accuracy as an antidepressant screening test (Porsolt, 1981; Borsini & Meli, 1988). However, many workers prefer to use the theoretically-neutral term "forced swim," rather than the theoretically-loaded term "behavioral despair." This more restrained position is endorsed by studies suggesting that immobility during forced swimming may be better characterized as a successful energy-saving coping response (West, 1990).

The importance of proper behavioral analysis is nowhere more evident than in relation to the original stress model of depression, learned helplessness. This procedure was based on the observation that animals exposed to uncontrollable stress (usually electric shocks) were subsequently impaired in learning to escape from shock, a deficit not seen in animals exposed to controllable shock (Seligman, 1975); these learning difficulties are reversed with reasonable selectivity by subchronic (3–7 days) treatment with antidepressant drugs (Martin, Laporte, Soubrie, El Mestikawy, & Hamon, 1989; Sherman, Saquitne, & Petty, 1982). It was proposed that exposure to uncontrollable (but not controllable) stress provides the basis, in animals as in people, for learning that stress is uncontrollable (helplessness), and that this learning has a number of debilitating consequences, including depression (Seligman, 1975). This hypothesis has been immensely influential within clinical psychology, probably more so than any other hypothesis derived from animal models. However, interpretation of the "learned helplessness" effect has been the subject of considerable controversy in both the human (Abramson, Metalsky, & Alloy, 1989; Abramson, Seligman, & Teasdale, 1978) and animal literature.

The term "learned helplessness" implies that the animals perform poorly because they have learned that their responses are ineffective in controlling their environment (Seligman, 1975). However, inescapable shock has a variety of other simpler effects that could also explain many of the behavioral impairments such as decreased locomotor activity (Glazer & Weiss, 1976) and analgesia (Lewis, Cannon, & Liebeskind, 1980). In order to demonstrate that inescapable shock does, additionally, cause "cognitive" impairments, performance accuracy was assessed using a maze task, in which performance would be independent of factors influencing motor speed. As predicted, accuracy was reduced in animals previously subjected to unavoidable shock, confirming the presence of a "cognitive" component to the pattern of impairment (Jackson, Maier, & Rapoport, 1978). However, subsequent work showed that this "cognitive" impairment arose from an increase in distractiblity rather than from a learning disability (Minor, Jackson, & Maier 1984). So inescapable shock does cause "cognitive" impairment, but at the level of attentional processes rather than "helplessness." Furthermore, in the typical experimental paradigm, inescapable shock is not only uncontrollable but also unpredictable in onset and/or offset. It was demonstrated many years ago that simply providing a feedback signal to accompany shock offset conferred protection against the ulcerogenic effect of inescapable (and therefore, uncontrollable) shock (Weiss, 1970). More recently, signals at the onset or offset of uncontrollable shock have been shown to protect against subsequent behavioral suppression. It was hypothesized that unpredictability generates high levels of fear, which are maintained within manageable limits by signals denoting the presence of danger or safety (Jackson & Minor, 1988). This reconceptualization explains some hitherto puzzling observations, such as the fact that inescapably shocked animals are easily distractable (Minor et al., 1984), which is exactly as would be expected in a state of fearful hypervigilance, but is difficult to explain in terms of loss of control (though even this is not the whole story: Maier, 1994). From this perspective, learned helplessness may be more valid as a model of anxiety rather than depression. However, while stress-induced response suppression may be prevented by anxiolytic drugs, they fail to reverse an already-established effect (in contrast to antidepressants, which are effective both before and after stress) (Drugan, Ryan, Minor, & Maier, 1984). This suggests that the relevance of these paradigms may be to a form of anxiety that does not respond to benzodiazepines. Phobia is one possibility, particularly if the conditioned fear interpretation of learned helplessness effects is correct: phobias do not respond to treatment with conventional anxiolytic drugs, but there is evidence that some phobic states do respond to antidepressants (Liebowitz, 1992).

Another problem in attempting to relate the learned helplessness model to depression is that the "core symptom" of the learned helplessness paradigm—an impairment of shock avoidance learning—does not reflect in any obvious way the symptomatology of depression, as captured, for example, by the *DSM-IV* diagnostic system (American Psychiatric Association, 1994). However, the learning impairment is only one among a wide range of consequences of uncontrollable shock, and in fact, the breadth of syptomatic parallels to severe depression has led to the suggestion that rodents subjected to uncontrollable shock could meet the *DSM* criteria for major depression (Weiss et al., 1982)! Among these behavioral abnormalities is a poor performance of rewarded behavior, which, as noted above, may be relevant to melancholia. One manifestation of this effect is a long-lasting decrease in responding for brain stimulation reward, which is specific to certain electrode placements, and therefore suggests a subsensitivity within part of the brain mechanism of reward, rather than, for example, a motor impairment (Zacharko, Bowers, Kokkinidis, & Anisman, 1983; Zacharko, Lalonde, Kasian, & Anisman, 1987). However, this subsensitivity to reward is long-lasting only if the animals are tested in the immediate aftermath of stress; otherwise the effect dissipates rapidly (Zacharko et al., 1983). A related finding is that following an initial exposure to severe stress behavioral deficits may be reinstated by mild stressors that are without effect in normal animals (Anisman & Zacharko, 1982). These studies suggest that it may be possible to develop conditioning models to explain how the risk of depression is elevated for several months in the aftermath of a stressful life event: the predisposing influence of life events in depression is well established, but is less well understood (Brown, 1993). Another important observation is that uncontrollable electric shock has variable behavioral effects (most of which are antidepressant-reversible) in different inbred mouse strains. To take an extreme example, in the C57BL/67 strain, uncontrollable shock severely impaired subsequent learning to escape shock, but had no effect on responding for brain stimulation reward, while the DBA/2J strain showed exactly the opposite pattern of deficits (Shanks & Anisman, 1988; Zacharko et al., 1987). These studies may provide a starting point for investigation of the physiological mechanisms underlying individual differences in responses to stress.

Subsensitivity to reward, an operational definition of anhedonia, the defining symptom of the melancholic subtype of major depression (Klein, 1974), has been most extensively characterized and studied in the chronic mild stress procedure. In this model, chronic, sequential exposure of rats or mice to a variety of mild stressors causes a decrease in their responsiveness to rewards, which is typically monitored as a decrease in the consumption of dilute sucrose solutions. This and related effects may be maintained over a period of weeks or months by continued application of the stress regime (Willner, Towell, Sampson, Muscat, & Sophokleous, 1987; Willner, Muscat, & Papp, 1992; Willner & Papp, 1997). The chronic mild stress paradigm has been extensively validated as a model of anhedonia. Initial studies showed that, in contrast to the decreased intake of dilute solutions of sucrose or saccharin, chronic mild stress did not decrease intake of plain water, food pellets, or concentrated sucrose solutions; thus the effects are not simply nonspecific changes in consummatory behavior (Muscat & Willner, 1992). Neither are the effects on consumption of sweet fluids secondary to loss of body weight, which is another consequence of chronic mild stress (and itself a symptom of depression): weight loss alone does not alter sweet fluid consumption, and in rats or mice subjected to chronic mild stress, decreases in sweet fluid consumption are still present after "correcting" for changes in body weight (Willner, Moreau, Nielsen, Papp, & Sluzewska, 1996).

Convergent evidence to support the hypothesis that changes in sweet fluids reflect an underlying change in sensitivity to rewards comes from studies using place conditioning to measure rewarding effects. In this paradigm, animals display a preference for a distinctive environment in which they have previously received rewards (Carr, Phillips, & Fibiger, 1989). Chronic mild stress attenuated or abolished place preferences established using a variety of natural or drug rewards. By contrast, drug-induced place aversions were unaffected by chronic mild stress, indicating that the effects on place preference conditioning are unlikely to result from a nonspecific motivational impairment or a failure of associative learning (Willner et al., 1992). Finally, chronic mild stress has also been shown to cause an increase in the threshold for brain stimulation reward (Moreau, Jenck, Martin, Mortas, & Haefely, 1992). The convergent evidence from all of these techniques supports the hypothesis that chronic mild stress causes anhedonia: while any one of these effects is susceptible of a variety of interpretations, anhedonia, a generalized decrease in sensitivity to rewards, is the most parsimonious hypothesis to account for all of the data.

Like the learned helplessness procedure, chronic mild stress causes a range of behavioral and physiological changes relevant to depression, in addition to the "core symptom" of the model, in this case, anhedonia. For example, chronic mild stress causes decreases in affective (e.g., sexual) and locomotor behaviors, and changes in the sleeping electroencephalogram characteristic of depression, such as a decrease in the latency to enter the first period of rapid eye movement sleep. Indeed, all of the behavioral symptoms of major depressive disorders listed in *DSM-IV* have now been demonstrated in animals exposed to chronic mild stress, along with most of the biological symptoms characteristic of severe depressions (Willner et al., 1992; Willner & Papp, 1997). The relevance of this model to depression is further supported by the evidence that the decreases in sensitivity to reward are reversed by all of the major classes of antidepressant drugs, and by electroconvulsive shock, but not by a variety of nonantidepressant drugs. Furthermore, the effects of antidepressants in this paradigm resemble the clinical situation, in two important respects: in both cases, treatment typically requires chronic (2–5 weeks) drug administration; and the effects are specific to the stressed/depressed state, since antidepressant treatment does not increase hedonic behavior either in nonstressed control animals or in nondepressed human volunteers (Willner et al., 1987, 1992; Willner & Papp, 1997). In view of these parallels with clinic, the chronic mild stress model has been used extensively to study the mechanisms of action of antidepressant drugs (Willner et al., 1992; Willner & Papp, 1997).

1.07.2.3 Other Approaches

The presumed etiological role in depression of loss events, and particularly, loss of a loved one, has led to the development of a number of animal models of depression based on separation phenomena. The most familiar of these models involves nonhuman primates, either infants isolated from their parents or juveniles isolated from their peer group. These studies originated from the work of Harlow (1958). Harlow's initial studies were designed to investigate mothering, and showed that infant rhesus monkeys reared with surrogate mothers preferred soft, cuddly surrogates to wire models containing a feeding bottle, demonstrating the importance of sensory contact. However, protest responses in infant monkeys separated from real or surrogate mothers were also noted, and these became the main focus of attention. The separation response consists of an initial stage of "protest," characterized by agitation, sleeplessness, and distress calls, followed by "despair," characterized by a decrease in activity, appetite, play and social interaction, and the assumption of a hunched posture and "sad" facial expression (Henn & McKinney, 1987; Suomi, 1976). These symptoms are strikingly similar to those of "anaclitic depression" in institutionalized children (Robertson & Bowlby, 1956). The relationship to adult depression remains uncertain, though recent life event research has confirmed that "loss events" such as bereavement are particularly likely to precipitate depression (Brown & Harris, 1989). Although the evolutionary proximity of primates has led some authors to consider these models to be of particular importance (e.g., Everitt & Keverne, 1979), their contribution has been rather modest. Because of the expense of using primates, the size of experimental groups in most studies has been too small to provide reliable data, and the few pharmacological studies using these models have not been impressive (Willner, 1990).

Separation phenomena of "protest" followed by "despair" are present to some extent in many other species, including cats, dogs, rodent, and precocial birds (Katz, 1981; McKinney & Bunney, 1969), and several of these phenomena have also been used as the basis for the development of animal models of depression. One of these, the reactivation of distress calling in one-week-old chicks appears to perform relatively well as an antidepressant screening test (Lehr, 1989). Chronic (4–6 weeks) isolation of adult rats has been found to cause a disruption of cooperative social behavior (Berger & Schuster, 1982) reminiscent of the poor social performance of depressed people (Lewinsohn, 1974), and this impairment of social cooperation is reversed by chronic antidepressant treatment (Willner et al., 1989). Thus, the investigation of complex phenomena associated with social isolation does not rely on the use of primates.

There are many other animal models of depression, which will not be discussed in detail because they have been studied largely for their pharmacology, and have received little psychological characterization. These include a number of inbred rat or mouse strains that display certain depression-like behaviors (Overstreet, 1993), and could in principle be used to investigate individual differences in vulnerability to depression, if validated. There is also a quite widely-used model based on the destruction of the olfactory bulb in rats, which differs from most other models in that the animals are hyperactive (Cairncross, Cox, Forster, & Wren, 1979; Van Riezen & Leonard, 1990). Thus, this model may reflect different aspects of depression, and may be relevant to agitated states.

This example highlights a significant limitation of animal models of depression, which have tended to focus either on an undifferentiated depressive state or on melancholia, while largely ignoring some other well-defined subtypes of depression. Delusional depression, for example, is differentiated behaviorally from nondelusional depression only by a greater association with psychomotor agitation (Nelson & Charney, 1981), but is pharmacologically distinct, being unresponsive to tricyclic antidepressants but responsive to electroconvulsive therapy (ECT) or to tricyclic/neuroleptic combinations (Nelson, 1987). Bipolar disorder is another well-defined diagnostic category for which there are no animal models. In this case, the depressive episodes appear to be identical to those of unipolar endogenous depression, but episodes of mania are interspersed. While there are a number of animal models of mania (Lyon, 1991; Robbins & Sahakian, 1980), the alternation of depressive-like and manic-like behaviors in an animal model has not yet been systematically addressed. Indeed, the episodic nature of unipolar depression (Angst et al., 1973) is itself another area that has not been explored in animal models.

1.07.3 ANIMAL MODELS OF ANXIETY

1.07.3.1 Clinical Considerations

Animal models of anxiety, which have been reviewed by Gray (1982), Treit (1985), and Green and Hodges (1991), have a long history, deriving from the early work of Pavlov and Masserman on "experimental neurosis." In these early studies, regressive, abnormal behaviors were produced by exposing animals to conflict: in Pavlov's studies, difficult discrimination problems in a conditioning procedure, and in Masserman's studies, approach-avoidance conflict (Abramson & Seligman, 1977). This set the stage for later work, most of which is based squarely on the assumption that conflict procedures are, almost by definition, animal models of anxiety.

A second early precursor to present-day animal models of anxiety is found in Watson's description of conditioned fear in the child "Little Albert" (Watson & Rayner, 1920). This observation, together with the formal demonstration by Miller (1960) that fear could be conditioned in laboratory animals, and suppressed by counterconditioning procedures, led directly to the development of systematic desensitization procedures for phobic anxiety (Wolpe, 1958), which for many years remained the jewel in the crown of behavioral psychotherapies (albeit that we now know that conditioning principles provide a rather incomplete account of the processes underlying systematic desensitization: Wilkins, 1971). Another early contribution of fear conditioning research to the treatment of phobias came from the observation that avoidance responses, once learned, are extremely long-lasting. Though initially perplexing, it soon emerged that this resistance to extinction occurs because an animal that makes a successful avoidance response has no opportunity to learn that the response is no longer necessary. If the animal was prevented from making the avoidance response, and so forcibly exposed to the information that a failure to respond did not lead to adverse consequences, the avoidance response extinguished rapidly. The treatment of phobic anxiety by flooding is based upon this observation (Solomon, Kamin, & Wynne, 1953). Although studies of fear conditioning led directly to the development of two successful treatments, it was later realized that this approach had some serious limitations, notably in failing to address the specificity of phobias to certain classes of objects (e.g., animal phobias are common, knife phobias are not). This feature is found in certain types of animal learning, which, it has been claimed, are preprogrammed by evolution to be selective and long-lasting—one much-discussed example is the conditioned taste aversion, in which novel foods are later avoided if their ingestion is followed by illness (Seligman, 1971). However, while conditioned taste aversions may display some of the characteristics of phobic anxiety, a more recent social learning model, in which monkeys acquire a fear of snakes by observation of already fearful animals (Mineka & Cook, 1986; Mineka, Davidson, Cook, & Keir, 1984), is perhaps more persuasive.

Alongside the concept of "conflict," the second major principle underlying the construction of animal models of anxiety has been the assumption that a model of anxiety should respond specifically to benzodiazepines and functionally related drugs. However, while benzodiazepines are effective in generalized anxiety disorder, they are actually ineffective in most forms of anxiety, such as phobia, obsessive-compulsive disorder and panic (Rickels, 1985; Rickels & Schweizer, 1987). Thus, to the extent that animal models of anxiety have relied on a response to benzodiazepines, they are primarily models of generalized anxiety disorder. Unfortunately, of all forms of anxiety, the symptomatology of generalized anxiety disorder is the least amenable to behavioral modeling, as it consists primarily of autonomic hyperactivity and a feeling of apprehension. In many animal models of anxiety, the behavior expressed appears to have face validity, in the sense that it appears reasonable that an anxious animal might behave in that way. However, the similarities to generalized anxiety disorder are more intuitive than demonstrable. There is also the problem that "virtually any test devised by experimental psychopharmacologists can induce mild or severe fear" (Green & Hodges, 1991). Thus, the appeal to the construct of fear to validate an animal model of anxiety is not particularly persuasive.

Furthermore, benzodiazepines are not necessarily the drugs of choice for anxiety. Antidepressants are commonly prescribed for anxiety—they are clearly more effective than benzodiazepines in panic disorders (Kahn, McNair, & Frankenthaler, 1986; Robinson, Nies, Ravaris, Ives, & Bartlett, 1978), and it is unclear whether there are any conditions, including generalized anxiety disorder, in which benzodiazepines are demonstrably superior to antidepressants in their clinical efficacy. With certain limited exceptions, antidepressants are not effective in traditional animal models of anxiety, and neither are a number of recently developed novel anxiolytics. This has led recently to a certain disillusionment with the traditional conflict models.

1.07.3.2 Conflict Models

Conflict models are based on both unconditioned and conditioned behaviors, and vary considerably in their behavioral sophistication. The simplest test is the examination of general behavior in a novel arena: locomotor activity is usually increased by anxiolytic drugs (e.g., Blumstein & Crawley, 1983). However, this test is of limited value, given the multitude of factors that can influence open-field activity the difficulty of subjecting open-field behavior to behavioral analysis. Some derivatives of the open field such as the hole-board test (File & Wardill, 1975) and the staircase test (Simiand, Keane, & Morre, 1984) are more complex, but share the same limitations. Emergence tests, from a presumed "safe" to a presumed "dangerous" area, are somewhat more sophisticated. A simple test of this type involves measuring crossings between the light and the dark half of a two-compartment chamber (Crawley, Skolnick, & Paul, 1984). Again, however, it is difficult to distinguish conflict behavior from general locomotor activity (Carey & Fry, 1988). The best known model of this group is the elevated plus-maze, a plus-shaped set of runways which is raised off the ground and has two enclosed and two open and exposed arms. Most, but not all, anxiolytic drugs tend to increase the time spent on the open arms, while other classes of drugs, such as stimulants, neuroleptics, and antidepressants, do not. The validity of this test is supported by the demonstration that levels of defaecation and freezing, as well as blood corticosterone are higher in animals confined in the open arms (Critchley & Handley, 1987; Pellow, 1986; Pellow, Chopin, File, & Briley, 1985). A number of animal models of anxiety are based on social interactions of various kinds, though there seems little reason to see any of these tests as being particularly relevant to social phobias. The most popular is a test in which social interaction between pairs of male rats is suppressed by using a novel environment and bright lighting; anxiolytics increase interaction levels under these conditions. Almost alone among animal models of anxiety, the social interaction test requires chronic drug administration resulting in tolerance to sedative effects, which are apparent as a decrease in locomotor activity on acute administration (File & Hyde, 1978).

In conflict models based on conditioned behavior, conflict is explicitly programmed into the experimental paradigm by punishing the animal for performing a response that it has been trained to emit in order to obtain a reward. A simple, and rapidly learned procedure of this kind is the Vogel water-lick conflict test where thirsty animals are first trained to lick from a drinking tube and electric shocks are then delivered through the tube or the floor when the animal licks, causing a suppression of drinking. Nonspecific effects are assessed either by including nonpunished drinking periods or by measuring drug effects on home-cage drinking (Vogel, Beer, & Clody, 1971). A more complex model involves training rats to consume a very sweet (32%) sucrose solution, then switching to a weak (4%) solution. When the switch occurs, the rate of licking falls, and only gradually, over the course of several days, returns to that of a rat maintained throughout on the weaker solution. Benzodiazepine treatment speeds the return to normal behavior (Flaherty, Grigson, Demetrikopoulos, & Demetrikopoulos, 1987; Flaherty, Grigson, & Rowan, 1986; see also Green & Hodges, 1991). Unlike other conflict models, this model is based on frustration rather than fear. Its relevance to anxiety is thus less obvious, though there is considerable evidence for a commonality between the brain mechanisms of fear and frustration (Gray, 1982).

The classic and most widely used conflict model is the Geller–Seifter operant conflict procedure. In this model, animals (rats or pigeons) respond in an operant chamber for food rewards; interspersed into the experimental session are periods (typically signaled by a light or tone) during which responding results in both an increased availability of food and a mild electric shock to the feet. Anxiolytic drugs increase responding during the conflict period, and response rates during the nonconflict period serve as a control for nonspecific drug influences. If necessary, the contingencies may be manipulated so as to ensure equivalent rates of responding in the two components; alternatively, the response rate during the conflict period may be set low to maximize the chance of detecting an increase, or high, to allow the detection of anxiogenic effects (Geller & Seifter, 1960; Sepinwall & Cook, 1978; see also Green & Hodges, 1991). Clearly, this model has manifest advantages over the unconditioned procedures described above in terms of the degree to which the behavioral parameters are under experimental control. However, the procedure is extremely time-consuming, and may require parametric changes during an experiment to maintain stable response rates. The Geller–Seifter procedure appears to have reasonably good face validity. However, the situation is not entirely straightforward. Following the prolonged training necessary to establish stable baselines of responding, animals may no longer display physical signs of fear or autonomic arousal (e.g., Stephens & Andrews, 1991). It is therefore possible that benzodiazepines may increase punished responding by mechanisms other than anxiety reduction, and in fact, there is evidence that they act at least in part by reducing the animal's ability to discriminate between punished and nonpunished components (Hodges, Baum, Taylor, & Green, 1986). It should be noted, however, that this critique of the Geller–Seifter procedure arises precisely because it is amenable to experimental analysis. Other simpler procedures, in which hypotheses concerning behavioral mechanisms are less readily testable, thereby escape criticism.

The construct validity of conflict models, in which behavior is suppressed by "fear" and reinstated by anxiolytic drug treatment, derives in large part from the hypothesis that anxiety results from activity in a forebrain "behavioral inhibition system" (Gray, 1982). This hypothesis, and with it most of the traditional anxiety models, has received a serious challenge from the development of nonbenzodiazepine anxiolytics. The anxiolytic action of buspirone, and related compounds acting at the 1A subtype of serotonin (5HT) receptors, is now well established; however, these drugs are ineffective, or minimally effective, in traditional anxiety models (Griebel, 1995).

1.07.3.3 Ethological Models

An alternative approach views anxiety as arising from threatening situations which elicit defensive behaviors. Defensive behaviors, involving characteristic submissive postures, are seen in an intruder rat or mouse introduced into the home cage of a resident conspecific; these defensive behaviors are suppressed by anxiolytic drugs (Krsiak et al. 1984; Shepherd & Rodgers, 1989). Another model elicits defensive behavior by shocking a rat or mouse through an electric prod; subsequently, the animal will bury the prod in sawdust if given the opportunity to do so. Again, defensive burying is suppressed with reasonable specificity by anxiolytics (Treit, 1985). However, the basic dependent variable in this model, the depth of sawdust piled up, is difficult to quantify.

A more complex approach, derived from studies that were initially conducted largely in wild rats, is based on an ethological analysis of behavior in the presence of, and following exposure to, a predator (a cat). These studies led to the description of a pattern of "risk assessment" behaviors, including visual "scanning" of the area in which the predator was encountered and "stretch approach" alternating with retreat (Blanchard, Rodgers, & Blanchard, 1994; Blanchard, Yudko, Rodgers, & Blanchard, 1993). It has been argued that risk assessment behaviors bear a striking resemblance to the behavioral features of general anxiety disorder: apprehensive expectation, vigilance, hyperattentiveness, and scanning (Blanchard et al., 1994). A number of simpler tests have been developed for evaluating drug effects on risk assessment, including a test in which rats are exposed to cat odor (Blanchard et al., 1993), and observational techniques applied to the elevated plus maze (Rodgers & Cole, 1993). The potentiated startle model is a very simple model of hypervigilance. In this procedure, the startle response to a loud noise, which is very readily quantifiable, is increased by the simultaneous presentation of a light that has previously been paired with electric shock. The potentiation of startle by the fear-inducing signal is suppressed by anxiolytics, which do not, however, affect the basal (unconditioned) response to the loud noise (Davis 1979).

In contrast to the traditional conflict models, in these ethologically-derived tests, both benzodiazepines and 5HT1A agonists are anxiolytic (Blanchard et al., 1993, 1994). The effects of benzodiazepines and 5HT1A agonists on defensive behavior are not, however, identical. In response to an approaching predator, animals display a characteristic repertoire of defensive behaviors. Benzodiazepines increase flight responses in wild rats to contact by a human experimenter; however, the 5HT1A agonists buspirone and gepirone decrease flight responses to such an extent that the rats allow themselves to be touched (Blanchard, Hori, Rodgers, Hendrie, & Blanchard, 1989; Blanchard, Rodgers, Hendrie, & Hori, 1988). Flight has been proposed as a potential model of panic (Deakin & Graeff, 1991), which unlike general anxiety disorder, does not respond to benzodiazepines. While buspirone appears to be ineffective in panic disorder (Sheehan, Raj, Trehan, & Knapp, 1993), positive effects of another 5HT1A agonist, gepirone, have been reported (Pecknold, Luthe, Scott-Fleury, & Jenkins, 1993), and the efficacy of specific serotonin uptake inhibitors is reasonably well established (Sheehan et al., 1993). Thus, different psychopathologies may reflect different components of the defensive behavioral repertoire. It is important to add that one of the defining features of psychopathology is that the behaviors displayed are disproportionate, and this to some extent limits the validity of models that relate anxiety and panic disorders to defensive behaviors resulting from the recent or actual presence of a predator. Nevertheless, these models serve as a rich source of hypotheses concerning the nature of, and the relationship between, anxiety and panic.

1.07.4 ANIMAL MODELS OF SCHIZOPHRENIA

1.07.4.1 Clinical Considerations

As in the case of animal models of anxiety, models of schizophrenia are dominated by its pharmacotherapy. The evidence that neuroleptic drugs act by antagonizing dopamine receptors (Creese, Burt, & Snyder, 1976) is so strong as to be indisputable, and despite many attempts to break out of the dopamine trap, there is as yet no clear evidence that drugs acting at other sites are effective (Rifkin & Siris, 1987). Nevertheless, a substantial proportion of schizophrenic patients do not benefit from neuroleptics, particularly, but not exclusively, those patients with predominantly negative symptoms, such as poverty of speech or social withdrawal. This fact not only stimulates the continuing search for other modes of treatment, but also casts doubt on the necessity of neuroleptic responsiveness in an animal model of schizophrenia. It is important to note that neuroleptics cause significant motor impairments (via their action on dopamine receptors in the caudate nucleus), and that many of the behavioral actions in animals are probably related more closely to these side effects than to their therapeutic action (see Iversen, 1987). The newer "atypical" neuroleptics are thought to minimize side effects by virtue of additional actions on other neurotransmitter systems (Tamminga & Gerlach, 1987) and/or specific actions within "limbic" areas (White & Wang, 1983). The development of these drugs has not, however, challenged the basic premise that neuroleptics act by antagonizing dopamine transmission.

A number of reviews have argued in principle against animal models of schizophrenia on the grounds that the symptoms of this disorder are uniquely human (e.g., Kumar, 1976). In fact, this problem is common to all areas of psychopathology, and arguably is not especially more serious in this case. Most of the symptoms of schizophrenia could in principle have counterparts in animals, particularly so if attention is paid to the structure of behavior rather than its content (Lyon, 1991). Hallucinatory behavior, for example, has been defined operationally as "sequences of well defined behaviors that could apparently only be accounted for by the reaction of the animal to non-existent stimuli" (Nielsen, Lyon, & Ellison, 1983), though it should be added that behavior meeting this criterion in animals reflects visual phenomena rather than the auditory hallucinations that are far more prevalent in schizophrenia. Lyon (1991) has suggested that for the purposes of animal models, the *DSM-III* criteria for schizophrenia should be augmented by two additional symptoms that featured prominently in Bleuler's (1950) account of dementia praecox, excessive alternation or switching between behaviors, and repetitive stereotyped responding. With these additions, Lyon (1991) has argued that the symptoms of schizophrenia may be reduced to four classes of abnormal behaviors: response switching, focusing, fragmentation, and stereotypy, which appear in sequence as the disorder progresses. These behaviors are clearly more relevant to the positive symptoms of schizophrenia (hallucinations, delusions, thought disorder), rather than the negative symptoms (affective flattening, poverty of speech, loss of drive, anhedonia), and this reflects the state of the field, which consists largely of models of positive symptomatology.

1.07.4.2 Dopamine Hyperactivity Models

Animal models of schizophrenia have been reviewed by McKinney and Moran (1981), Iversen (1987), and Lyon (1991). As expected from the predominance of dopamine in theoretical accounts of schizophrenia, the majority of animal models involve pharmacological treatments that increase dopamine receptor stimulation, primarily by using directly acting dopamine agonists, such as apomorphine, or by using amphetamine to release dopamine from presynaptic stores. This strategy is greatly encouraged by the close phenomenological similarities between paranoid schizophrenia and the psychosis that develops in people after prolonged amphetamine intoxication (Connell, 1958). Single, high doses of amphetamine can also elicit the amphetamine psychosis (Griffith, Cavanaugh, Held, & Oates, 1972). In animals high doses of amphetamine cause a species-characteristic pattern of repetitive stereotyped behavior (Randrup & Munkvad, 1971), and a number of paradigms have been described in which amphetamine elicits a repetitive perseveration either of responses, or more interestingly, of complex programs of behavior, such as switching between two responses, which are continued irrespective of their outcome. These effects are antagonized by neuroleptics (Evenden & Robbins, 1983; Ridley & Baker, 1983). Similar patterns of stereotyped switching have been observed in schizophrenic patients (Frith & Dorie, 1983; Lyon, Mejsholm, & Lyon, 1986).

It is frequently difficult to decide whether phenomena of this kind are more relevant to schizophrenia or to mania, which in many respects displays a superficially similar pattern of symptoms. However, manic behavior, while increased in rate, and frequently stereotyped and repetitive, usually retains a goal-directed organization, unlike the fragmentation of behavior characteristic of schizophrenia. Behavior on this criterion is more obviously related to schizophrenia, rather than mania, and is seen following chronic administration of dopamine agonists to rats, cats, or monkeys. This has been achieved either by repeated daily injections of amphetamine or cocaine, typically for a period of many weeks (Ellinwood & Kilbey, 1977, 1980), or by the use of an implanted slow-release amphetamine pellet, which lasts for one or two weeks (Ellison & Eison, 1983; Nielsen et al., 1983). Repeated injection regimes cause, eventually, a wide range of behavioral abnormalities, including hyper-reactivity, abortive or fragmented behaviors, and a large number of motor disorders, including dystonias, ataxia, akithisia, and facial dyskinesias. Motor disturbances are not uncommon in schizophrenics. While these are often ascribed to neuroleptic therapy, postural and oculomotor disturbances were described long before the neuroleptic era (Kraepelin, 1919), and evidence suggests that tardive dyskinesias may also be an intrinsic part of the underlying disorder (Barnes, 1988). An important feature of the chronic stimulant syndrome, particularly apparent in the continuous slow-release model, is the presence of apparent hallucinatory behavior. Nielsen et al. (1983), using stringent operational criteria, described four types of hallucinatory behavior in monkeys, all directed at nonexistent objects, resulting in flight, attack, eating, or other behaviors such as threat responses. These behaviors, as well as motor stereotypies, were suppressed by the neuroleptic haloperidol (Lyon, 1991).

The release of dopamine in subcortical structures is increased in rats reared in social isolation (Jones, Hernandez, Marsden, & Robbins, 1988). Thus, the pathological effects of isolation rearing may to some extent be mediated by dopamine hyperactivity, though clearly, social isolation also has many other effects. Goosen (1981) has described a "social deprivation syndrome" in female rhesus monkeys, consisting of stereotyped behaviors of various kinds, and bizarre and/or self-directed behaviors. These abnormalities resemble some of the behaviors observed following chronic stimulant drug treatment, but the more florid hallucinatory behaviors are far less conspicuous. As some of the effects of social isolation are reversed by antidepressants, isolation rearing provides at best an approximation to a model of schizophrenia. However, the demonstration that prolonged changes in dopamine activity can be elicited environmentally is encouraging.

In addition to causing motor stereotypies that resemble positive symptoms of schizophrenia, a dramatic loss of social interaction is seen following chronic stimulant treatment in rats (Nielsen & Lyon, 1982) or monkeys (Ridley, Baker, & Scragg, 1979), or acute treatment with high doses of amphetamine in monkeys (Ellenbroek & Cools, 1990). These effects in monkeys are seen as an increase in interindividual distance, a decrease in the number of interactions initiated, and an increase in the number of interactions terminated, all leading to an increase in social isolation, which may represent a model for negative symptoms (Ellenbroek & Cools, 1990).

1.07.4.3 Attentional Models

Returning to positive symptoms, it has long been known that schizophrenic patients suffer attentional disorders which have been considered to underly their cognitive deficits (Frith, 1987). This defect is usually characterized as a broadening of attention, leading to an increased response to irrelevant stimuli that would normally be filtered out; a more specific formulation has suggested that the abnormality could be considered to represent a reduction in the influence of past regularities on current perception (Hemsley, 1987). These considerations have led to a number of attempts to model attentional dysfunction; the two most widely used paradigms are the prepulse inhibition and latent inhibition models.

Prepulse inhibition refers to the ability of a warning signal to decrease the amplitude of the startle response to a loud noise. Prepulse inhibition is impaired in schizophrenic patients (Braff, Stone, Callaway, Geyer, & Ball 1978), and is also impaired by dopamine agonist drugs (amphetamine or apomorphine) in animals (Mansbach, Geyer, & Braff, 1988), and by isolation rearing (Wilkinson et al., 1994). While a rather simple behavior, the prepulse inhibition paradigm is of interest because the neural mechanisms underlying the acoustic startle reflex are very well understood, and appear to be very similar in rats and humans (Davis, 1984). This model is used increasingly within the pharmaceutical industry, largely on account of its simplicity. Latent inhibition is a more complex form of habituation, in which preexposure to a stimulus reduces the ability to learn about its consequences when it is later paired with a significant event. This ability is impaired in acute schizophrenics, though not in chronic schizophrenics (Baruch, Hemsley, & Gray, 1988). In rats, where the significant event is an electric shock, latent inhibition is disrupted by amphetamine and potentiated by neuroleptics (Joseph & Jones, 1991). An attractive feature of this procedure is that a conditioned fear paradigm is used. Thus, the conditioned stimulus suppresses operant behavior, latent inhibition lifts the suppression, and amphetamine reverses latent inhibition to reinstate the suppression. In other words, amphetamine causes a decrease in responding, so the effect on latent inhibition cannot be an artefact of psychostimulation. However, the significance of this model is uncertain, as an increased reactivity to preexposed stimuli could reflect a nonspecific state of arousal. This interpretation is supported by a study suggesting that the effect of amphetamine in this model results from an increase in the perceived intensity of the electric shock rather than an altered response to the conditioned stimulus (Killcross et al., 1994).

While animal models of schizophrenia are dominated by dopaminergic drugs, there are a number of models based on the use of drugs interacting with other neurotransmitter systems. The most notable of these are lysergic acid diethylamide (LSD) and phencyclidine (PCP). Jacobs, Trulson, and Stern (1977) described a range of bizarre behaviors in LSD-treated cats that resemble the effects of chronic psychostimulant treatment, though at low doses, limb flicks, which are not seen either in the stimulant models or in schizophrenia, were the most prominent effect of LSD. Despite the popularity in the 1960s of the LSD experience as a model of schizophrenia in humans, it is clear that both clinicians and patients perceive the effects of

LSD as different from schizophrenia (Hollister, 1962), unlike amphetamine, which appears to exacerbate existing symptomatology. PCP, on the other hand, does appear to induce the full range of positive and negative symptoms, including thought disorder, in volunteers and in schizophrenics (Davies & Beech, 1960; Greifenstein, Yoshitake, De Valut, & Gajewski, 1958). In both animal and human subjects, PCP produces stereotyped behaviors as well as deficits in social behavior that are comparable to the social withdrawal and isolation that form an important part of the negative symptom complex. Both sets of effects could be reversed in rats by clozapine which is almost alone among neuroleptic drugs in its activity against negative symptoms; in contrast, the typical neuroleptic haloperidol suppressed PCP-induced stereotypies, but failed to normalize social behavior (Sams-Dodd, 1996). It is interesting that PCP also disrupts prepulse inhibition, and the characteristics of this response actually resemble the schizophrenic deficit more closely than the effects of amphetamine (Mansbach & Geyer, 1989).

Finally, there have been a number of attempts to model schizophrenia using brain lesions. Lesions of the hippocampus have been shown to impair performance in a foraging task, using a holeboard in which food was located below four of the 16 holes, which has been suggested to provide a possible model of thought disorder (Oades & Isaacson, 1978). The effects of hippocampal lesions are similar in many respects to those of amphetamine, and may in general be antagonized by neuroleptics (Schmajuk, 1987), though the neuroleptic reversal of the holeboard impairment was only partial (Oades & Isaacson, 1978). More recently, hippocampal lesions have been shown to cause an increase in sensitivity to the disruptive effect of amphetamine in the prepulse inhibition model (Swerdlow et al., 1995). Significantly, neuropathological abnormalities have been described postmortem in the hippocampus of schizophrenic patients (Kovelman & Scheibel, 1984), suggesting that hippocampal damage may be at least in part responsible for an apparent dopamine hyperactivity in schizophrenia. However, while the effects of hippocampal damage do resemble those of acute amphetamine treatment, the hallucinatory and bizarre behaviors that emerge following chronic amphetamine administration are absent. A particularly interesting variant of this model is based on a neonatal hippocampal lesion. Animals prepared in this way appear normal as juveniles, but when adults are hyperactive, hyper-reactive to amphetamine, and show defective prepulse inhibition (Lipska et al., 1995; Lipska, Jaskiw, & Weinberger, 1993). This is the only model to capture the developmental aspect of schizophrenia, which typically first appears in late adolescence.

1.07.5 ANIMAL MODELS OF SUBSTANCE ABUSE

1.07.5.1 Drug Self-administration

It is clear from the earlier discussion of depression, anxiety, and schizophrenia that the validity of animal models can often be difficult to determine, in part because of the considerable differences in the methodologies used to study animal behavior on the one hand, and human psychopathology, on the other. However, in the case of animal models of addiction, at least superficially, these difficulties appear less acute: just like humans, animals voluntarily self-administer drugs (Bozarth, 1987; Goudie, 1991; Koob, 1995). Before discussing the self-administration model, a second parallel should be mentioned in passing: animals are also able to report on their perceptions of a drug. In drug discrimination procedures, animals receive food rewards for pressing one lever if they have been administered the drug and a different lever if they have received saline. After training, their response to a novel agent reflects the extent to which it is perceived as similar to the training drug. This model is used extensively in mechanistic studies of psychotropic drug action (Stolerman, Samele, Kamien, Mariathasan, & Hague, 1995), and is considered to represent a model of the subjective effects of drugs (Preston & Bigelow, 1991).

In drug self-administration experiments, the animal, typically a rat or a monkey, is implanted with a chronic in-dwelling intravenous catheter, which is connected through plastic tubing to a reservoir located outside the experimental chamber, which places minimal restraints on the animal's freedom of movement. Provided suitable precautions are taken against infection and blood clots, catheters can remain patent for periods of up to several months, allowing for experiments involving complex and lengthy training procedures and/or chronic drug administration. Drug delivery is made contingent on the performance of an operant response, typically lever pressing, but sometimes a less motorically demanding response, such as nose poking. If access to the drug is unlimited, a pattern of self-administration may develop, consisting of periods of drug intake alternating with periods of abstinence, during which other activities such as eating, drinking, and sleeping occur. This binge pattern of drug consumption has been described in animals self-administering a wide variety of psychomotor stimulants and with alcohol; binges are also characteristic of problematic use of both of these drug classes in humans. Opioid self-administration is more constant over time, as it is in human intravenous opiate users (Brady, Griffiths, Hienz, Ator, Lukas, & Lamb, 1987; Yokel, 1987). However, unlimited access to self-administered drugs can lead to fatalities, and therefore drugs are rarely made available for self-administration on an unlimited basis (Yokel, 1987). When drug availability is limited by scheduling daily sessions of several hours duration, responding typically follows an inverted U-shaped dose–response function. Above a minimum dose, responding and drug intake rise with drug dose; this ascending portion of the curve may be relatively steep and difficult to study. Above a certain dose, responding falls with further increases in dose, and this has the effect of maintaining the blood concentration of the drug at a relatively—in some cases, remarkably—constant level, which remains stable over prolonged periods of time (Koob, 1995; Yokel, 1987).

In general there is close agreement between the drugs intravenously self-administered by animals and those which are abused by people. Thus, animals readily learn to self-administer opioids (e.g., morphine, heroin), psychostimulants (e.g., amphetamine, cocaine), barbiturates, dissociative anesthetics (e.g., phencyclidine), and alcohol. Also self-administered, though less reliably, are nicotine, caffeine, and benzodiazepines. Conversely, various drugs that are not abused do not typically support self-administration; these include antidepressants, neuroleptics, and serotonin agonists (e.g., fenfluramine, buspirone). However, there are some anomalies. The most obvious is that animals do not reliably self-administer either hallucinogens (e.g., LSD) or cannabinoids; on the other hand, they do reliably self-administer certain drugs which are not abused, such as kappa opioid agonists and local anesthetics (Yokel, 1987; Young & Herling, 1986). Thus, the agreement between self-administered drugs and abused drugs is not perfect; nevertheless, drug self-administration procedures are routinely used within the pharmaceutical industry to predict whether novel compounds have "abuse liability."

Self-administration procedures can also estimate the relative reinforcing efficacy (relative abuse liability) of different agents. This issue must be approached with caution, since under many circumstances (e.g., cocaine self-administration under fixed ratio schedules of reinforcement and limited daily access), an increase in unit dose of the drug results in a compensatory decrease in response rates, and vice versa. Thus, a higher response rate may not denote a higher efficacy. This problem has been addressed in a number of ways, the most popular being the progressive ratio schedule. In progressive ratio schedules the response requirement progressively increases, during the session or across session, up to the point at which the subject stops responding; at this "breakpoint," it is assumed that the work requirement exceeds the reinforcing effect of the drug. Thus, different drugs can be compared directly according to the work that animals are prepared to perform to earn them. Studies using this procedure have shown that cocaine supports higher breakpoints than a variety of other agents, supporting the popular view of cocaine as a highly addictive drug (Katz 1990; Young & Herling, 1986). Motivation to work for drug, as revealed by performance in a progressive ratio schedule, has also been proposed as a method of measuring drug craving in animals (Markou et al., 1993). This approach is supported by a study of human volunteers reinforced under a progressive ratio schedule by puffs on a cigarette, in which breakpoints correlated significantly with a questionnaire measure of cigarette craving (Willner, Harding, & Eaton, 1995).

The importance of measuring motivation, rather than simple consumption of a drug, is well illustrated by studies of alcohol self-administration. While rats or monkeys readily self-administer alcohol by intravenous injection, the majority of animals will not drink alcohol in sufficient quantities to produce significant pharmacological effects, and certainly insufficient to cause physical dependence. The need to model oral alcohol consumption in animals has stimulated selective breeding programs that have succeeded in producing several rat strains characterized by high alcohol consumption, a preference for alcohol over water in two-bottle tests, and in some cases, the development of physical dependence (Crabbe & Li, 1995). However, when alcohol-drinking rats are required to work for alcohol, differences between the strains emerge. Only one strain, the P (alcohol-preferring) line, works harder under a progressive ratio schedule than its nonpreferring (NP) control strain; other high alcohol-drinking strains (HAD, AA) do not (Ritz, Garcia, Protz, Rael, & George, 1995). While the greater proclivity of P rats to consume alcohol suggests that the P and NP rat strains may prove useful in investigating possible genetic bases for alcoholism, it is equally important to note that the difference between the strains is not absolute. Moderate or high levels of alcohol intake can be induced in outbred animals by introducing alcohol in a sweetened solution and gradually fading out the sucrose component—a progression reminiscent of the process by which many teenagers acquire the taste for alcohol. After initiation of alcohol drinking by means of the sucrose fading technique, alcohol consumption in NP rats is comparable to that of P rats (Samson, Tolliver, Lumeng, & Li, 1989).

1.07.5.2 Drugs as Positive and Negative Reinforcers

The most striking feature of the early literature utilizing drug self-administration models is that drugs function as reinforcers, in operant procedures, in a manner highly reminiscent of conventional reinforcers. In particular, the characteristic patterns of behavior associated with different schedules of reinforcement are maintained as effectively by drugs as by conventional reinforcers, such as food delivery or shock termination (Schuster & Johanson, 1981). These similarities raise the question of whether drug reinforcement has unique characteristics. One potential difference is that with conventional reinforcers, increasing the magnitude of the reinforcer leads to increased responding, whereas with drug reinforcers, the dose–response function has an inverted U-shape, such that increases in dose cause compensatory decreases in response rate (Wise, 1987). However, the same effect can, in fact, be observed with conventional reinforcement: variations in the sweetness of a food reinforcer also lead to changes in responding that follow an inverted U-shaped function (Willner, Phillips, & Muscat, 1991). Another question is whether drugs are more efficaceous than conventional reinforcers. Early studies suggested that cocaine might be more reinforcing than food, insofar as cocaine maintained higher response rates than food in second-order schedules of reinforcement (in which long chains of behavior are required to achieve delivery of a reinforcer). However, this difference in responding is only seen when reinforcers are presented at intervals during the session: if the session terminates upon delivery of the first reinforcer, cocaine and food maintain equivalent rates of responding. Similarly if a session run under these conditions is preceded by an injection of cocaine, this results in faster responding in both cocaine and food reinforcement conditions. These studies suggest strongly that cocaine is not uniquely reinforcing; rather, the apparent differences between cocaine- and food-maintained behavior result from stimulant effects of the cocaine delivered during experimental sessions (Katz & Goldberg, 1987).

Studies have begun to examine the relationship between drug-reinforced and conventionally-reinforced behaviors in more complex settings, where both types of reinforcer are concurrently available. It is clear from studies using progressive ratio procedures that drug-reinforced responding is "price sensitive" (i.e., drug intake is suppressed when the work requirement is too high). This leads to variations in drug taking as a function of the relative price and availability of drugs and alternative reinforcers, which can be described with some precision using concepts derived from economics, such as elasticity of demand (Bickel, Hughes, DeGrandpre, Higgins, & Rizzuto, 1992; Hursh, 1980). For example, in an early study, two baboons were found to respond equally to heroin and food when both were in plentiful supply, but as resources became scarcer, heroin showed greater elasticity: food intake was maintained but heroin intake decreased (Elsmore, Fletcher, Conrad, & Sodetz, 1980). Behavioral economics may provide a useful framework for analyzing human drug taking (Heyman, 1996).

The extensive parallels between drug reinforcement and conventional reinforcement in nondependent animals, together with the fact that the construct of dependence is of very limited value for understanding human drug taking (Sanger, 1991), has led some workers to argue that self-administered drugs function primarily as incentives, which are sought for their rewarding properties (Stewart, de Wit, & Eikelboom, 1984). The incentive properties of drugs of abuse are demonstrated by the so-called "priming" effect: in animals trained to self-administer a drug, and then extinguished so that responding decreases substantially, a single drug administration is sufficient to reinstate responding, in a dose-related fashion, despite the fact that responding leads to no further drug injections. Priming effects have been consistently demonstrated with intravenously self-administered stimulants and opiates (Stewart & de Wit, 1987), as well as oral alcohol (Chiamulera, Valerio, & Tesari, 1995), and can be seen following multiple nonreinforced extinction sessions (Shaham, Rodaros, & Stewart, 1994; Chiamulera et al., 1995). Priming effects are seen not only with the training drug, but are also elicited by other drugs, which are effective in reinstating responding to the extent that they resemble the training drug in drug discrimination assays (Stewart & de Wit, 1987).

Furthermore, priming effects can also be elicited by environmental stimuli previously associated with drug infusions (Stewart et al., 1984). Consistent with these findings, in human drug users, stimuli associated with cocaine have been found to elicit cocaine craving in cocaine users, but were without effect in opiate users or in nondrug users (Ehrman, Robbins, Childress, & O'Brien, 1992). These studies are interpreted as meaning that stimuli associated with abused drugs can acquire conditioned incentive properties. For example, following conditioning sessions in which neutral stimuli are paired with drug infusions, nondependent animals will learn a new response that is reinforced by presentation of the conditioned stimulus (Davis & Smith, 1987). Conditioned reinforcing effects have been studied most extensively in the place-conditioning model (Carr et al., 1989). In this paradigm, animals are typically exposed to one of two distinctive environments in the presence of a drug, and to a different environment in its absence. In a subsequent drug-free choice trial, animals display a preference for environments associated with self-administered drugs, with few exceptions (notably, alcohol), while avoiding environments paired with drugs that are presumed to have aversive effects, such as lithium chloride or the opiate antagonist, naloxone (Goudie, 1991). A major methodological advantage of these conditioning procedures is that reinforcing effects are studied in the absence of the drug, which avoids problems associated with motor or other nonspecific drug effects.

In contrast to this focus on the euphoric properties of abused drugs, the traditional view of drug dependence has emphasized the importance of drug tolerance, physical dependence, and the avoidance of withdrawal distress as a major fact in the maintenance of continued drug use. It is well established that the induction of physical dependence leads to an increase in drug self-administration in animal and human subjects (Cappell et al., 1987; Griffiths, Bigelow, & Henningfield, 1980; Griffiths, Bigelow, & Liebson, 1986). There is also some evidence from animal models that the induction of physical dependence increases not only the quantity of drug consumed, but also the reinforcing efficacy of drug reinforcers in self-administration paradigms. However, the evidence is equivocal. For example, in studies using a progressive ratio schedule to assess relative reinforcing efficacy, induction of physical dependence increased responding maintained by codeine or morphine, but not by another opioid, loperamide, which has minimal abuse potential in humans, or by alcohol (Yanagita, 1987).

Other effects of withdrawal can be readily demonstrated within the self-administration paradigm. For example, in a classic study, rats were maintained in a procedure in which periods of morphine availability alternated with periods in which lever pressing resulted in food availability or in avoidance of electric shock. The withdrawal of morphine led to pronounced decreases in both food deliveries and successful avoidances (Thompson & Schuster, 1964). This phenomenon, known as behavioral dependence, is a sensitive index of withdrawal, since the disruption of operant behavior maintained by conventional reinforcers can be observed following withdrawal from treatment regimes, with a variety of drugs, that produce few if any overt signs of distress (Balster, 1985). Indeed, experiments of this kind provide strong evidence for a brief cocaine withdrawal syndrome, which is still unrecognized by many clinicians. For example, the threshold for brain-stimulation reward is markedly elevated, indicating anhedonia, following withdrawal from as little as 12 hours exposure to cocaine self-administration (Markou & Koob, 1991).

A number of lines of evidence confirm the aversive nature of drug-withdrawal states. The most direct evidence is that morphine-dependent monkeys learn to respond to a conditioned stimulus so as to avoid or delay infusions of an opiate antagonist (Goldberg, Hoffmeister, Schlichting, & Wuttke, 1971). Indeed, stimuli associated with drug withdrawal in dependent subjects, which, typically, have been repeatedly precipitated, in the presence of the conditioned stimulus, by administration of a pharmacological antagonist, can acquire a powerful control over behavior. For example, stimuli or places paired with opiate antagonists can elicit signs of physical withdrawal in morphine-dependent rats (Wikler, 1973), monkeys (Goldberg, 1976), and humans (O'Brien, 1975), which can be seen many weeks after drug withdrawal, and long after offset of the withdrawal period. Like withdrawal itself, such stimuli are aversive, as demonstrated, for example, in place conditioning experiments (Hand, Koob, Stinus, & leMoal, 1988).

1.07.5.3 Limitations

It is clear from this brief review that there have been significant advances in understanding the factors that control the initiation and maintenance of drug self-administration, and many of these factors are of proven relevance to human drug taking. However, a significant limitation on the face validity of animal models of addiction arises from the impoverished environments in which animals typically self-administer drugs. There is evidence that drug taking is decreased by testing animals in social groups (Alexander, Beyerstein, Hadaway, & Coambs, 1981), and as discussed above, it is clear that another potential source of enrichment, the availability, and conditions of availability, of alternative sources of reinforcement, can markedly influence drug taking. Human drug taking typically involves a substantial element of social ritual and cultural conventions, and a high degree of choice (Hartnoll, 1991). These factors may be relatively unimportant in studies of the biological bases of drug action, but may be crucial if the object is to understand the environmental controls on drug-taking behavior. It is tempting to draw the comparison between the environmental impoverishment typically imposed on experimental animals and the conditions of social deprivation that characterize many communities in which drug-taking is rife; however, these parallels require detailed evaluation, and should be viewed with caution.

Another area that has received little attention is the process of change in patterns of drug self-administration. For good and obvious reasons, drug self-administration is usually studied under steady-state conditions, in which drug taking is constant over sessions, and well-regulated within sessions. Much of the literature discussed above describes factors that influence drug self-administration under these conditions. Much human drug taking is also controlled and well-regulated, and one of the major challenges is to understand the conditions in which regulation breaks down, leading to uncontrolled, heavy, and typically, problematic substance use. As discussed above, uncontrolled use, including a characteristic human pattern in which drug or alcohol binges alternate with periods of abstinence, can also be observed in animal models, under appropriate experimental conditions, and studies carried out under these conditions are most encouraging (Koob, 1995). However, there has been little attempt to use animal models to study the evolution of controlled into uncontrolled use, and the factors that control those changes.

As noted earlier, the evaluation of construct validity requires a good theoretical understanding of both the model and the condition modeled (Willner, 1991). Unfortunately, both sides of this equation are rarely met; in the present case, drug intake by animals is now well understood (at least under steady-state conditions), but a coherent theoretical understanding of substance abuse by people remains elusive. The motivational framework outlined above provides an excellent, if limited, account of many aspects of drug taking in animal models. This framework is of great heuristic value in guiding the rapid development of an understanding of the neurobiological substrates mediating the rewarding effects of drugs and the aversive effects of withdrawal (Altman et al., 1996; Koob & Goeders, 1989). However, the importance of these constructs for understanding human substance abuse, and the relevance of the neurobiological insights arising from animal models of drug self-administration, remains to be confirmed

1.07.6 OVERVIEW: VALIDITY AND UTILITY OF ANIMAL MODELS

Having surveyed animal models relevant to four areas of psychopathology, it is of interest now to reconsider the general issue of the validity of animal models in relation to the criteria outlined at the start of this chapter, and to summarize their uses. McKinney and Bunney (1969) suggested that an animal model should resemble the clinical disorder in its etiology, treatment, symptomatology, and physiological basis, and these categories provide a convenient framework within which to examine the contribution of animal models to our understanding of the clinical disorders.

As noted earlier, the orientation of animal models towards psychopharmacology has meant that the predictive validity of the models is determined largely by their pharmacological responsiveness, which has not been discussed in detail in this chapter. In general, however, predictive validity is reasonably high for each class of model, though there are exceptions to this general rule at the level of individual models. Animal models of depression can usually be relied upon to differentiate established antidepressants from known inactive agents, though there are some well-known exceptions which are difficult to interpret (e.g., anticholinergic drugs often appear antidepressant-like: see Willner, 1990). However, a multiplicity of novel agents are currently undergoing clinical trials in depression, and it will be some years before the predictive validity of animal models of depression can be assessed in relation to these novel agents. By contrast, animal models of anxiety, traditionally the jewel in the crown among animal models when assessed by their response to benzodiazepines, are currently in the unfavorable position of coming to terms with the anxiolytic and putatively anxiolytic actions of novel serotonergic agents and, rather belatedly, with the anxiolytic actions of antidepressants.

Animal models of schizophrenia remain for the moment unchallenged as regards their predictive validity, as all drug development in this area has involved variations on the theme of dopamine blockade. However, a comparison with the position of anxiety models a few short years ago suggests caution: we do not know how well the current models would cope with the discovery of novel neuroleptics acting through mechanisms unrelated to dopamine. Animal models of substance abuse perform well in discriminating abused from nonabused substances, but less well when asked to rank different substances for their relative abuse liability.

Given the intrinsic limitation that animal models must be based on behavioral rather than subjective symptomatology, in those cases where efforts have been made to establish face validity, the position is again reasonably encouraging. Core symptoms of the disorder are demonstrable not only in models of depression (e.g., subsensitivity to rewards), anxiety (responses to aversive stimuli), and substance abuse (drug self-administration), but also in models of schizophrenia (bizarre and hallucinatory behaviors). However, a recognition of the diagnostic heterogeneity of the clinical disorders is now becoming a high priority. In the case of schizophrenia, the need for an acceptable model of negative symptomatology has long been explicitly recognized. However, the heterogeneity of anxiety disorders has until recently had little impact on animal models: in comparison to the multitude of tests addressing generalized anxiety disorder, very little thought has been given to the development of animal models of phobic anxiety or panic, and virtually none to obsessive-compulsive disorder. In part, this narrow view reflects the degree to which the field of animal modeling is drug-driven: it is difficult to see any other reason for the scarcity of work on models of phobic anxiety, which is not drug-responsive but lends itself superbly well to the development of models for which face validity may be established. However, domination by drugs is not the whole story: the pharmacological distinction between delusional and nondelusional depressions is well established clinically but has been overlooked in the development of animal models of depression. Rather, the preoccupation with generic disorders of depression and anxiety reflects an unwillingness to look outside the ordered confines of the laboratory at the chaotic richness of the clinical world. This attitude is understandable, but perhaps, short-sighted.

Animal models are at their weakest in the area of construct validity. However, this reflects to a considerable extent the inadequacies of current clinical theorizing. Theories of psychopathology tend to be either strongly psychological or strongly biological in nature, with little attempt to unify the two approaches. Psychologically based theories are well developed in relation to aspects of depression and anxiety, though rudimentary in their account of schizophrenia (see Brewin, 1988). These theories are frequently couched in language difficult to transport into animal models, though where the possibility exists (as, for example in the learned helplessness model of depression), the response has been enthusiastic. Irrespective of the extent to which the terminology of psychological theories of depression or anxiety may or may not be applied to animals, such theories must be capable of explaining the efficacy of antidepressant and anxiolytic drugs. The fact that cognitive theories of psychopathology cannot explain drug effects is a measure of their current limitations. Biologically based theories of psychopathology take the actions of therapeutic agents as one of their starting points, but beyond this, have made relatively little progress. Few biological abnormalities have been determined with any certainty in the brains of psychiatric patients, and as a result, biologically based theories of depression (Willner, 1985), anxiety (Gray, 1982) or schizophrenia (Frith, 1987) are derived in large part from research using animal models (an uncomfortable fact of life that clinicians tend to overlook). While there is an unacceptable degree of circularity in appealing to these theories as support for the construct validity of animal models, it is certainly arguable that the construct validity of animal models in all three areas approaches the limits set by clinical theory. Much the same can be said of substance abuse: the animal literature provides a clear and powerful theoretical framework for understanding drug self-administration, but the human literature is marked by an equally striking lack of theoretical consensus (e.g., Altman et al., 1996).

Perhaps surprisingly, animal models have little to say about the etiology of anxiety. Potentially important insights about the etiology of phobias arise from the studies of observational conditioning of fear in monkeys (Mineka & Cook, 1986), but as far as generalized anxiety disorder and panic are concerned, animal models of anxiety provide no insights into why some people respond normally in situations that generate pathological anxiety in others. Models of depression are more promising, and provide strong support for the view that stress and social isolation are of etiological significance, as well as some pointers to the mechanisms mediating these effects. It may be important to reflect that this research provides clear evidence against the distinction between "psychological" and "biological" depressions, in demonstrating that environmental precipitants can cause profound physiological consequences. The major contribution of animal models to the etiology of schizophrenia has been the demonstration that neuropathological features of the schizophrenic brain, such as hippocampal abnormalities or ventricular widening, could indeed be of etiological importance. Animal models of substance abuse have focused to a certain extent on factors determining individual differences in susceptibility to drug or alcohol self-administration, and have identified clear genetic influences and hormonal variables that could be relevant to human drug taking.

It seems likely that animal models may play an important role in the elucidation of relationships between etiology and symptomatology. In particular, the demonstration that stress can cause insensitivity to rewards may be of great significance for understanding melancholia; there is little in the clinical literature to predict this relationship, which clearly deserves further investigation. The generation of positive, but not negative, schizophrenic-like symptoms by hippocampal lesions similarly suggests that specific etiological factors may relate to specific symptom patterns. In addition to their use in the experimental investigation of problems related to etiology, animal models may also be of value in defining diagnostic boundaries. The current tendency to obscure the border between depression and anxiety has arisen primarily from clinical observations (for example, that antidepressants are also anxiolytic), but has long been prefigured in the involvement of serotonin in animal models of both disorders. Work with animal models has led to a unifying concept that serotonergic systems function to suppress the generation of impulsive behavior (Soubrie, 1986; Willner, 1989), which is of major importance for understanding a spectrum of disorders including depression, anxiety, bulimia, alcoholism, and obesity, in which serotonergic transmission is abnormal and specific serotonin reuptake inhibitors are therapeutically effective.

Animal models were initially introduced, during the behaviorist era, as a means of developing behavioral therapies for psychological disorders, and this remains the arena in which animal models are best known to many clinical psychologists. However, this use of animal models is very little evident in recent or current work. Indeed, it is difficult to see any contribution to behavioral psychotherapy arising out of the current models surveyed in this chapter in the areas of depression, anxiety or schizophrenia. The exception is in the area of addiction, where the demonstration that stimuli conditioned by association with abused substances can acquire motivational properties in their own right has led to attempts to treat addictions by cue exposure therapy, which attempts to extinguish responses to drug-associated cues (Drummond, Tiffany, Remington, & Glautier, 1995). Nevertheless, animal models do contribute in major ways to the treatment of psychopathology. These contributions arise from their use in drug screening programs and their role in determining the mechanisms of action of psychotherapeutic drugs. This does, however, represent a significant change of direction: animal models were initially located within mainstream psychology, but now serve primarily as one of the major foundations of psychopharmacology.

In the future, animal models will have their greatest impact in relation to the physiological mechanisms underlying psychopathology. This is a rich seam of research, which derives not simply from the models themselves but also from the ways in which research in animal models relates to the broad spectrum of research in behavioral neuroscience. In effect, animal models form a critical interface between psychiatry and basic research in behavioral neuroscience; they are the channel through which developments at the basic level are brought into clinical perspective. Animal models serve both as theories of psychopathology and as practical tools for its investigation. Their development is an iterative process: the investigation of an animal model provides hypotheses concerning the clinical condition, and the outcome of the clinical investigations provides the basis for a more refined assessment of the validity of the model. Viewed in this light, the major limitation of animal models, their preoccupation with the objective aspects of psychopathology rather than the subjective, may actually be something of an advantage; it forces the development of a linguistic framework within which the psychological and biological dimensions of psychopathology may ultimately be reconciled.

1.07.7 SUMMARY

This chapter reviews animal models of depression, anxiety, schizophrenia, and substance abuse. Although much of the review, of necessity, deals with drug effects in animal models, the focus throughout is on animal models as simulations of psychopathology, rather than on their utility as drug development tools. Even in this psychological context, the models in current use relate primarily to psychopharmacology rather than to mainstream psychology, and their principal use is in investigations of the physiological mechanisms underlying psychopathology. In contrast to their historical roots, animal models of psychopathology now make relatively little contribution to the development of psychotherapies. For each of the disorders reviewed, models, or classes of model, are assessed for the extent to which they meet the criteria of predictive validity (ability to predict from the model to the clinic), face validity (extent of phenomenological similarity), and construct validity (theoretical rationale). In general, the results of this assessment are quite encouraging.

1.07.8 REFERENCES

Abramson, L. Y., & Seligman, M. E. P. (1977). Modeling psychopathology in the laboratory: History and rationale. In J. D. Maser & M. E. P. Seligman (Eds.), *Psychopathology: Experimental models* (pp. 1–26). San Francisco: Freeman.

Abramson, L. Y., Metalsky, G., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Reviews, 96,* 358–372.

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87,* 49–74.

Alexander, B. K., Beyerstein, B. L., Hadaway, P. F., & Coambs, R. B. (1981). Effect of early and later colony housing on oral ingestion of morphine in rats. *Pharmacology, Biochemistry and Behaviour, 15,* 571–576.

Altman, J., Everitt, B. J., Glautier, S., Markou, A., Nutt, D., Oretti, R., Phillips, G. D., & Robbins, T. W. (1996). The biological, social and clinical bases of drug addiction: Debate and review. *Psychopharmacology, 125,* 285–345.

American Psychiatric Association (1994). *Diagnostic and statistical manual of psychiatric disorders* (4th ed.). Washington DC: Author.

Angst, J., Baastrup, P., Grof, P., Hippius, H., Poldinger, W., & Weis, P. (1973). The course of monopolar depression and bipolar psychoses. *Psychiatrica Neurologica et Neurochirurgica, 76,* 489–500.

Anisman, H. A., & Zacharko, R. M. (1982). Depression: The predisposing influence of stress. *Behavioral and Brain Sciences, 5,* 89–137.

Balster, R. L. (1985) Behavioral studies of tolerance and dependence. In L. S. Seiden, & R. L. Balster (Eds.), *Behavioral pharmacology: The current status* (pp. 403–418) New York: A. R. Liss.

Barnes, T. R. E. (1988) Tardive dyskinesia: Risk factors, pathophysiology and treatment. In Granville-Grossman, K. (Eds.), *Recent advances in clinical psychiatry, No. 6* (pp. 185–205). London: Churchill Livingstone.

Baruch, I., Hemsley, D. R., & Gray, J. A. (1988). Differential performance of acute and chronic schizophrenics in a latent inhibition task. *Journal of Nervous and Mental Disease, 176,* 598–606.

Berger, B. D., & Schuster, R. (1982). An animal model of social interaction: Implications for the analysis of drug action. In M. Y. Spiegelstein & A. Levy (Eds.), *Behavioral models and the analysis of drug action* (pp. 415–428). Amsterdam: Elsevier.

Bickel, W. K., Hughes, J. R., De-Grandpre, R. J., Higgins, S. T., & Rizzuto, P. (1992). Behavioral economics of drug-self-administration. IV. The effects of response requirements on the consumption of and interaction between concurrently available coffee and cigarettes. *Psychopharmacology, 107,* 211–216.

Blanchard, D. C., Hori, K., Rodgers, R. J., Hendrie, C. A., & Blanchard, R. J. (1989). Attenuation of defensive threat and attack in wild rats (*Rattus rattus*) by benzodiazepines. *Psychopharmacology, 97,* 392–401.

Blanchard, D. C., Rodgers, R. J., Hendrie, C. A., & Hori, K. (1988). Taming of wild rats (*Rattus rattus*) by 5HT1A agonists buspirone and gepirone. *Pharmacology and Biochemical Behavior, 31,* 269–278.

Blanchard, R. J., Yudko, E. B., Rodgers, R. J., & Blanchard, D. C. (1993). Defense system psychopharmacology: An ethological approach to the pharmacology of fear and anxiety. *Behavioral Brain Research, 58,* 155–165.

Blanchard, R. J., Rodgers, R. J., & Blanchard, D. C. (1994). Ethological approaches to the preclinical psychopharmacology of anxiety. In T. Palomo & T. Archer (Eds.), *Strategies for studying brain disorders, Vol 1: Depressive, anxiety, and drug abuse disorders* (pp. 1145–1160). London: Farrand Press.

Bleuler, E. (1950). *Dementia Praecox or the group of schizophrenias.* New York: International Universities Press.

Blumstein, L. K., & Crawley, J. N. (1983). Further characterization of a simple automated exploratory model for the anxiolytic effects of benzodiazepines. *Pharmacology, Biochemistry and Behavior, 18,* 37–40.

Borsini, F., & Meli, A. (1988). Is the forced swimming test a suitable test for revealing antidepressant activity? *Psychopharmacology, 94,* 147–160.

Bozarth, M. A. (1987). *Methods of assessing the reinforcing properties of abused drugs.* New York: Springer.

Brady, J. V., Griffiths, R. R., Hienz, R. D., Ator, N. A., Lukas, S. E., & Lamb, R. J. (1987). Assessing drugs for abuse liability and dependence potential in laboratory primates. In M. A. Bozarth (Ed.), *Methods of assessing the reinforcing properties of abused drugs* (pp. 45–85). New York: Springer.

Braff, D. L., Stone, C., Callaway, E., Geyer, M., Glick, I., & Ball, L. (1978). Prestimulus effects on human startle reflex in normals and schizophrenics. *Psychophysiology, 15,* 339–343.

Brewin, C. R. (1988). *Cognitive foundations of clinical psychology.* Hove, UK: Erlbaum.

Brown, G. W. (1993). The role of life events in the aetiology of depressive and anxiety disorders. In S. C. Stanford & P. Salmon (Eds.), *Stress: From synapse to syndrome* (pp. 23–50) London: Academic Press.

Brown, G. W., & Harris, T. O. (1989). *Life events and illness.* New York: Guilford Press.

Cairncross, K. D., Cox, B., Forster, C., & Wren, A. F. (1979). Olfactory projections systems, drugs and behaviour: A review. *Psychoneuroendocrinology, 4,* 253–272.

Cappell, H., Busto, V., Kay, G., Naranjo, C. A., Sellers, E. M., & Sanchez-Craig, M. (1987). Drug deprivation and reinforcement by diazepam in a dependent population. *Psychopharmacology, 91,* 154–160.

Carey, M. P., & Fry, J. P. (1988). An evaluation of light aversion as a measure of anxiolytic benzodiazepine activity in the mouse. *European Journal of Neuroscience,* (Suppl 1) 343.

Carr, G. D., Phillips, A. G., & Fibiger, H. C. (1989). Conditioned place preference as a measure of drug reward. In S. J. Cooper & J. Liebman (Eds.), *The neuropharmacological basis of reward* (pp. 264–319). Oxford, UK: Oxford University Press.

Carroll, B. J. (1989). Diagnostic validity and laboratory studies: Rules of the game. In L. N. Robins & J. E. Barrett (Eds.), *The validity of psychiatric diagnosis* (pp. 229–244). New York: Raven.

Chiamulera C., Valerio, E., & Tessari, M. (1995). Resumption of ethanol-seeking behavior in rats. *Behavioural Pharmacology, 6,* 32–39.

Connell, P. H. (1958). Amphetamine psychosis. *Maudsley monograph No. 5.* London: Oxford University Press.

Crabbe, J. C., & Li, T.-K. (1995). Genetic strategies in preclinical substance abuse research. In F. E. Bloom & D. J. Kupfer (Eds.), *Psychopharmacology: The fourth generation of progress* (pp. 799–811). New York: Raven.

Crawley, J. N., Skolnick, P., & Paul, S. M. (1984). Absence of intrinsic antagonistic actions of benzodiazepine antagonists on an exploratory model of anxiety in the mouse. *Neuropharmacology, 23*, 531–537

Creese, I. N. R., Burt, D. R., & Snyder, S. H. (1976). Dopamine receptor binding predicts clinical and pharmacological potencies of antischizophrenic drugs. *Science, 192*, 481–483.

Critchley, M. A. E., & Handley, S. L. (1987). Effects in the X-maze anxiety model of agents acting at 5-HT1 and 5-HT2 receptors. *Psychopharmacology, 93*, 502–506.

Davies, B. M., & Beech, H. R. (1960). The effects of L-arylcyclohexylamine (sernyl) on twelve normal volunteers. *Journal of Mental Science, 106*, 912–914.

Davis, M. (1979). Diazepam and flurazepam: Effects on conditioned fear as measured with the potentiated startle paradigm. *Psychopharmacology, 62*, 1–7.

Davis, M. (1984). The mammalian startle response. In R. C. Eaton (Ed.), *Neural mechanisms of startle behavior* (pp. 287–351). New York: Plenum.

Davis, W. M., & Smith, S. G. (1987). Conditioned reinforcement as a measure of the rewarding properties of drugs. In M. A. Bozarth (Ed.), *Methods of assessing the reinforcing properties of abused drugs* (pp. 199–210). New York: Springer.

Deakin, J. F. W., & Graeff, F. G. (1991). 5-HT and mechanisms of defence. *Journal of Psychopharmacology, 5*, 305–315.

Drugan, R., Ryan, S. M., Minor, T. R., & Maier, S. F. (1984). Librium prevents the analgesia and shuttle-box escape deficits typically observed following inescapable shock. *Pharmacology, Biochemistry and Behavior, 21*, 749–754.

Drummond, C., Tiffany, S., Remington, R., & Glautier, S. (1995) (Eds.). *Addiction: Cue exposure theory and practice*. London: Wiley.

Ehrman, R. N., Robbins, S. J., Childress, A. R., & O'Brien, C. P. (1992). Conditioned responses to cocaine-related stimuli in cocaine abuse patients. *Psychopharmacology, 107*, 523–529.

Ellenbroek, B. A., & Cools, A. R. (1990). Animal models with construct validity for schizophrenia. *Behavioural Pharmacology, 1*, 469–490.

Ellinwood, E. H. & Kilbey, M. M. (1977). Chronic stimulant intoxication models of psychosis. In I. Hanin & E. Usdin (Eds.), *Animal models in psychiatry and neurology* (pp. 61–74) Oxford, UK: Pergamon.

Ellinwood, E. H., & Kilbey, M. M. (1980). Fundamental mechanisms underlying altered behavior following chronic administration of psychomotor stimulants. *Biological Psychiatry, 15*, 749–757.

Ellison, G., & Eison, M. S. (1983). Continuous amphetamine intoxication: An animal model of the acute psychotic episode. *Psychological Medicine, 13*, 751–761.

Elsmore, T. F., Fletcher, G. V., Conrad, D. G., & Sodetz, F. J. (1980). Reduction of heroin intake in baboons by an economic constraint. *Pharmacology, Biochemistry and Behaviour, 13,* 729–731.

Evenden, J. L., & Robbins, T. W. (1983). Increased response switching, perseveration and perseverative switching following d-amphetamine in the rat. *Psychopharmacology, 80,* 67–73.

Everitt, B. J., & Keverne, E. B. (1979), Models of depression based on behavioural observations of experimental animals. In E. S., Paykel & A. Coppen (Eds.), *Psychopharmacology of affective disorders* (pp. 41–59). Oxford, UK: Oxford University Press.

File, S. E., & Hyde, J. R. G. (1978). Can social interaction be used to measure anxiety. *British Journal of Pharmacology, 62,* 19–24.

File, S. E., & Wardill, A. G. (1975). Validity of head-dipping as a measure of exploration in a modified holeboard. *Psychopharmacologia, 44,* 53–59.

Flaherty, C. F., Grigson, P. S., Demetrikopoulos, P. S., & Demetrikopoulos, M. K. (1987). Effect of clonidine on consummatory negative contrast and on novelty-induced stress. *Pharmacology, Biochemistry and Behavior, 27,* 659–664.

Flaherty, C. F., Grigson, P. S., & Rowan, G. A. (1986). Chlordiazepoxide and the determinants of negative contrast. *Animal Learning and Behavior, 14,* 315–321.

Frith, C. D. (1987). The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. *Psychological Medicine, 17,* 631–648.

Frith, C. D., & Done, J. (1983). Stereotyped responding by schizophrenic patients on a two-choice guessing task. *Psychological Medicine, 13,* 779–786.

Geller, I., & Seifter, J. (1960). The effects of meprobamate, barbiturates, d-amphetamine and promazine on experimentally induced conflict in the rat. *Psychopharmacologia, 9,* 482–492.

Glazer, H. I., & Weiss, J. M. (1976). Long-term interference effect: an alternative to learned helplessness. *Journal of Experimental Psychology and Behavior Processes, 2,* 202–213.

Goldberg, S. R. (1976). Stimuli associated with drug injections as events that control behavior. *Pharmacological Reviews, 27,* 325–340.

Goldberg, S. R., Hoffmeister, F., Schlichting, U., & Wuttke, W. (1971). Aversive properties of nalorphine and naloxone in morphine-dependent rhesus monkeys. *Journal of Pharmacological and Experimental Therapeutics, 179,* 268–276.

Goosen, C. (1981). Abnormal behavior patterns in rhesus monkeys: Symptoms of mental disease? *Biological Psychiatry, 16,* 697–716.

Goudie, A. J. (1991). Animal models of drug abuse and dependence. In P. Willner (Ed.), *Behavioural models in psychopharmacology: Theoretical, industrial and clinical perspectives* (pp. 453–484). Cambridge, UK: Cambridge University Press.

Gray, J. A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. Oxford, UK: Oxford University Press.

Green, S., & Hodges, H. (1991). Animal models of anxiety. In P. Willner (Ed.), *Behavioural models in psychopharmacology: Theoretical, industrial and clinical perspectives* (pp. 21–49). Cambridge, UK: Cambridge University Press.

Greifenstein, F. E., Yoshitake, J., DeValut, M., & Gajewski, J. E. (1958). A study of L-arylcyclohexylamine for anesthesia. *Anesthesia and Analgesia, 37,* 283–294.

Griebel, G. (1995). 5-Hydroxytryptamine-interacting drugs in animal models of anxiety disorders. More than 30 years research. *Pharmacology and Therapeutics, 65,* 319–395.

Griffith, J. D., Cavanaugh, J., Held, J., & Oates, J. A. (1972). Dextroamphetamine: Evaluation of psychotomimetic properties in man. *Archives of General Psychiatry, 26,* 97–100.

Griffiths, R. R., Bigelow, G. E., & Henningfield, J. E. (1980). *Similarities* in animal and human drug taking. In N. K. Mello (Ed.), *Advances in substance abuse* (Vol. 1, pp. 1–90). Greenwich, CT: JAI Press.

Griffiths, R. R., Bigelow, G. E., & Liebson, I. A. (1986). Human coffee drinking: Reinforcing and physical dependence producing effects of caffeine. *Journal of Pharmacology and Experimental Therapeutics, 239,* 416–425.

Hand, T. H., Koob, G. F., Stinus, L., & le Moal, M. (1988). Aversive properties of opiate receptor blockade are centrally mediated and are potentiated by previous exposure to opiates. *Brain Research, 474,* 364–368.

Harlow, H. F. (1958). The nature of love. *American Psychologist, 13,* 673–685.

Hartnoll, R. L. (1991). The relevance of behavioural models of drug abuse and dependence liabilities to the understanding of drug misuse in humans. In P. Willner (Ed.), *Behavioural models in psychopharmacology: Theoretical, industrial and clinical perspectives* (pp. 503–519). Cambridge, UK: Cambridge University Press.

Hemsley, D. R. (1987). An experimental psychological model for schizophrenia. In H. Häfne, W. F. Gattaz, & W. Janzarik (Eds.). *Search for the causes of schizophrenia* (pp. 179–188). Berlin: Springer.

Henn, F. A., & McKinney, W. T. (1987). Animal models in psychiatry. In H. Y. Meltzer (Ed.). *Psychopharmacology: the third generation of progress* (pp. 687–696). New York: Raven.

Heyman, G. M. (1996). Resolving the contradictions of addiction. *Behavioral and Brain Sciences, 19,* 561–610.

Hodges, H., Baum, S., Taylor, P., & Green, S. (1986). Behavioural and pharmacological dissociation of chlordiazepoxide effects on discrimination and punished responding. *Psychopharmacology, 59,* 155–161.

Hollister, L. E. (1962). Drug-induced psychoses and schizophrenic reactions: A critical comparison. *Annals of the New York Academy of Sciences, 96,* 80–88.

Hursh, S. R. (1980). Economic concepts for the analysis of behavior. *Experimental Analysis of Behavior, 34,* 219–238.

Iversen, S. D. (1987). Is it possible to model psychotic states in animals? *Journal of Psychopharmacology, 1,* 154–176.

Jackson, R. L., & Minor, T. R. (1988). Effects of signaling inescapable shock on subsequent escape learning: Implications for theories of coping and "learned helplessness." *Journal of Experimental Psychology: Animal Behavior Processes, 14,* 390–400.

Jackson, R. L., Maier, S. F., & Rapoport, P. M. (1978). Exposure to inescapable shock produces both activity and associative deficits in rats. *Learning and Motivation, 9,* 69–98.

Jacobs, B. L., Trulson, M. E., & Stern, W. C. (1977). Behavioral effects of LSD in the cat: Proposal of an animal behavior model for studying the actions of hallucinogenic drugs. *Brain Research, 132,* 301–314.

Jesberger, J. A., & Richardson, J. S. (1985). Animal models of depression: Parallels and correlates to severe depression in humans. *Biological Psychiatry, 20,* 764–784.

Jones, G. H., Hernandez, T. D., Marsden, C. A., & Robbins, T. W. (1988). Enhanced striatal response to d-amphetamine as revealed by intracerebral dialysis following social isolation in rats. *British Journal of Pharmacology, 94,* 349P.

Joseph, M. H., & Jones, S. H. (1991). Latent inhibition and blocking: Further consideration of their construct validity as animal models of schizophrenia. *Behavioural Pharmacology, 2,* 521–526.

Kahn, R. J., McNair, D. M., & Frankenthaler, L. M. (1987). Tricyclic treatment of generalized anxiety disorder. *Journal of Affective Disorders, 13,* 145–151.

Katz, R. J. (1981). Animal models and human depressive disorders. *Neuroscience and Biobehavioral Reviews, 5,* 231–246.

Katz, J. L. (1990). Models of relative reinforcing efficacy of drugs and their predictive validity. *Behavioural Pharmacology, 1,* 283–301.

Katz, J. L., & Goldberg, S. R. (1987). Second-order schedules of drug injection. In M. A. Bozarth (Ed.), *Methods of assessing the reinforcing properties of abused drugs* (pp. 105–115) New York: Springer.

Keehn, J. D. (1979). *Psychopathology in animals: Research and clinical implications*. New York: Academic Press.

Killcross, A. S., Dickinson, A., & Robbins, T. W. (1994). Amphetamine-induced disruptions of latent inhibition are reinforcer mediated: Implications for animal models of schizophrenic attentional dysfunction. *Psychopharmacology, 115,* 185–195.

Klein, D. F. (1974). Endogenomorphic depression: A conceptual and terminological revision. *Archives of General Psychiatry, 31,* 447–454.

Koob, G. F. (1995). Animal models of drug addiction. In F. E. Bloom & D. J. Kupfer (Eds.), *Psychopharmacology: The fourth generation of progress* (pp. 759–772). New York: Raven.

Koob, G. F., & Goeders, N. E. (1989). Neuroanatomical substrates of drug self-administration. In J. M. Liebman & S. J. Cooper (Eds.), *The neuropharmacological basis of reward* (pp. 214–263). Oxford, UK: Oxford University Press.

Kovelman, J. A., & Scheibel, A. B. (1984). A neurohistological correlate of schizophrenia. *Biological Psychiatry, 19,* 1601–1621.

Kraepelin, E. (1919). *Dementia praecox and paraphrenia.* Edinburgh, UK: Livingstone.

Krsiak, M., Sulcova, A., Donat, P., Tomasikova, Z., Dlohozkova, N., Kosar, E., & Masek, K. (1984). Can social and agonistic interactions be used to detect anxiolytic activity of drugs? In K. A. Miczek, M. R. Kruk, & B. Olivier (Eds.), *Ethopharmacological aggression research* (pp. 93–114). New York: Liss.

Kumar, R. (1976). Animal models for screening new agents: A behavioural view. *British Journal of Clinical Pharmacology, 4,* (Suppl. 1), 13–17.

Lehr, E. (1989). Distress-call reactivation in isolated chicks: A behavioural indicator with high selectivity for antidepressants. *Psychopharmacology, 97,* 145–146.

Lewinsohn, P. M., Zeiss, A. M., Zeiss, R. A., & Haller, R. (1977), Endogenicity and reactivity as orthogonal dimensions in depression. *Journal of Nervous and Mental Disorders, 164,* 327–332.

Lewis, J. W., Cannon, J. T., & Liebeskind, J. C. (1980). Opioid and non-opioid mechanisms of stress-induced analgesia. *Science, 208,* 623–625.

Liebowitz, M. R. (1992). Reversible MAO inhibitors in social phobia, bulimia, and other disorders. *Clinical Neuropharmacology, 15* (Suppl. 1A), 434–435.

Lipska, B. K., Jaskiw, G. E., & Weinberger, D. R. (1993). Delayed effects of neonatal excitotoxic hippocampal damage in the rat: A potential model of schizophrenia. *European Neuropsychopharmacology, 3,* 398.

Lipska, B. K., Swerdlow, N. R., Geyer, M. A., Jaskiw, G. E., Braff, D. L., & Weinberger, D. R. (1995). Neonatal excitotoxic hippocampal damage in rats causes post-pubertal changes in prepulse inhibition of startle and its disruption by apomorphine. *Psychopharmacology, 122,* 35–43.

Lyon, M. (1991). Animal models of mania and schizophrenia. In P. Willner (Ed.), *Behavioural models in psychopharmacology: Theoretical industrial and clinical perspectives* (pp. 253–310). Cambridge, UK: Cambridge University Press.

Lyon, N., Mejsholm, B., & Lyon, M. (1986). Stereotyped responding in schizophrenic outpatients: Cross-cultural confirmation of perseverative switching on a two-choice task. *Journal of Psychiatric Research, 20,* 137–150.

Maier, S. F. (1994). Learned helplessness: Relationships with fear and anxiety. In S. F. Stanford, & P. Salmon (Eds.), *Stress: From synapse to syndrome* (pp. 207–243). London: Academic Press.

Mansbach, R., & Geyer, M., (1989). Effects of phencyclidine and phencyclidine analogs on sensorimotor gating in the rat. *Neuropsychopharmacology, 2,* 299–308.

Mansbach, R., Geyer, M., & Braff, D. (1988). Dopaminergic stimulation disrupts sensorimotor gating in the rat. *Psychopharmacology, 94,* 507–514.

Markou, A., & Koob, G. F. (1991). Postcocaine anhedonia: An animal model of cocaine withdrawal. *Neuropsychopharmacology, 4,* 17–26.

Markou, A., Weiss, F., Gold, L. H., Caine, B., Schulteis, G., & Koob, G. F. (1993). Animal models of drug craving. *Psychopharmacology, 112,* 163–182.

Martin, P., Laporte, A. M., Soubrie, P., El Mestikawy, S., & Hamon, S. (1989) Reversal of helpless behaviour by serotonin reuptake inhibitors. In P. Bevan, A. R. Cools, & T. Archer (Eds.), *Behavioural pharmacology of 5HT* (pp. 231–234). New York: Erlbaum.

McKinney, W. T. (1984). Animal models of depression: An overview. *Psychiatric Developments, 2,* 77–96.

McKinney, W. T., & Bunney, W. E. (1969). Animal model of depression: Review of evidence and implications for research. *Archives of General Psychiatry, 21,* 240–248.

McKinney, W. T., & Moran, E. C. (1981). Animal models of schizophrenia. *American Journal of Psychiatry, 138,* 478–483.

Miller, N. E. (1960). Learning resistance to pain and fear: Effects of over-learning, exposure and rewarded exposure in context. *Journal of Experimental Psychology, 60,* 137–145.

Mineka, S., & Cook, M. (1986). Immunization against the observational conditioning of snake fear in monkeys. *Journal of Abnormal Psychology, 95,* 307–318.

Mineka, S., Davidson, M., Cook, M., & Keir, R. (1984). Observational conditioning of snake fear in rhesus monkeys. *Journal of Abnormal Psychology, 93,* 355–372.

Minor, T. R., Jackson, R. L., & Maier, S. F. (1984). Effects of task-irrelevant cues and reinforcement delay on choice-escape learning following inescapable shock: Evidence for a deficit in selective attention. *Journal of Experimental Psychology: Animal Behavior Processes, 10,* 543–556.

Moreau, J-L., Jenck, F., Martin, J. R., Mortas, P., & Haefely, W. E. (1992). Antidepressant treatment prevents chronic mild stress-induced anhedonia as assessed by ventral tegmentum self-stimulation behaviour in rats. *European Journal of Neuropsychopharmacology, 2,* 43–49.

Muscat, R., & Willner, P. (1992). Suppression of sucrose drinking by chronic mild unpredictable stress: A methodological analysis. *Neuroscience and Biobehavioral Reviews, 16,* 507–517.

Nelson, J. C. (1987). The use of antipsychotic drugs in the treatment of depression. In J. Zohar & R. H. Belmaker (Eds.) *Treating resistant depression* (pp. 131–146). New York: PMA.

Nelson, J. C., & Charney, D. S. (1981). The symptoms of major depression illness. *American Journal of Psychiatry, 138,* 1–13.

Nielsen, E. B., & Lyon, M. (1982). Behavioral alterations during prolonged low level continuous amphetamine administration in a monkey family group (*Cercopithecus aethiops*). *Biological Psychiatry, 17,* 423–434.

Nielsen, E. B., Lyon, M., & Ellison, G. (1983). Apparent hallucinations in monkeys during round-the-clock amphetamine for seven to fourteen days: Possible relevance to amphetamine psychosis. *Journal of Nervous and Mental Disease, 171,* 222–233.

Oades, R. D., & Isaacson, R. L. (1978). The development of food search behavior by rats: The effects of hippocampal damage and haloperidol. *Behavioral and Neural Biology, 24,* 327–337.

O'Brien, C. P. (1975). Experimental analysis of conditioning factors in human narcotic addiction. *Pharmacological Reviews, 27,* 533–543.

Overstreet, D. H. (1993). The Flinders sensitive line rats: A genetic animal model of depression. *Neuroscience and Biobehavioral Reviews, 17,* 51–68.

Paykel, E. S. (1979). Recent life events in the development of the depressive disorders. In R. A. Depue (Ed.), *The psychobiology of depressive disorders: Implications for the effects of stress* (pp. 245–262). New York: Academic Press.

Pecknold, J. C., Luthe, L., Scott-Fleury, M. H., & Jenkins, S. (1993). Gepirone and the treatment of panic disorder: An open study. *Journal of Clinical Psychopharmacology, 13,* 145–149.

Pellow, S. (1986). Anxiolytic and anxiogenic drug effects in a novel test of anxiety: Are exploratory models of anxiety in rodents valid? *Methods and Findings in Experimental Clinical Pharmacology, 8,* 557–565.

Pellow, S., Chopin, P., File, S. E., & Briley, M. (1985). Validation of open:closed arm entries in an elevated plus-maze as a measure of anxiety in the rat. *Journal of Neuroscience Methods, 14,* 149–167.

Porsolt, R. D. (1981). Behavioural despair. In S. J. Enna, J. B. Malick, & E. Richelson (Eds.), *Antidepressants: Neurochemical, behavioral and clinical perspectives* (pp. 121–139). New York: Raven Press.

Porsolt, R. D., LePichon, M., & Jalfre, M. (1977). Depression: A new animal model sensitive to antidepressant treatment. *Nature, 266,* 731–732.

Preston, K. L., & Bigelow, G. E. (1991). Subjective and discriminative effects of drugs. *Behavioural Pharmacology, 2,* 293–313.

Randrup, A., & Munkvad, I. (1971). Behavioural toxicity of amphetamines: Studies in animal experiments. In S. B. de Baker (Ed.). *The correlation of adverse effects in man with observations in animals.* (pp. 6–17). Amsterdam: Excerpta Medica.

Rickels, K. (1985). Benzodiazepines in emotional disorders. In D. E. Smith & D. R. Wesson (Eds.), *The benzodiazepines: Current standards for medical practice* (pp. 97–109). Lancaster: MTP Press.

Rickels, K., & Schweizer, E. E. (1987). Current pharmacotherapy of anxiety and panic. In H. Y. Meltzer (Ed.), *Psychopharmacology: The third generation of progress* (pp. 1193–1203). New York: Raven.

Ridley, R. M., & Baker, H. F. (1983). Is there a relationship between social isolation, cognitive flexibility, and behavioural stereotypy? An analysis of the effects of amphetamine in marmosets. In K. A. Miczek (Ed.), *Ethopharmacology: Primate models of neuropsychiatric disorders* (pp. 101–135). New York: Liss.

Ridley, R. M., Baker, H. F., & Scragg, P. R. (1979). The time course of the behavioural effects of amphetamine and their reversal by haloperidol in a primate species. *Biological Psychiatry, 14,* 753–765.

Rifkin, A., & Siris, S. (1987). Drug treatment of acute schizophrenia. In H. Y. Meltzer (Ed.), *Psychopharmacology: The third generation of progress* (pp. 1095–1101). New York: Raven.

Ritz, M. C., Garcia, J. M., Protz, D., Rael, A.-M., & George, F. R. (1995). Ethanol-reinforced behavior in P, NP, HAD and LAD rats: Differential regulation of reinforcement and motivation. *Behavioural Pharmacology, 5,* 521–531.

Robbins, T. W., & Sahakian, B. J. (1980). Animal models of mania. In R. H. Belmaker & H. M. van Praag (Eds.), *Mania: An evolving concept* (pp. 143–216). New York: Jamaica, Spectrum.

Robinson, D. S., Nies, A., Ravaris, C. L., Ives, J. O., & Bartlett, D. (1978). Clinical pharmacology of phenelzine. *Archives of General Psychiatry, 35,* 629–635.

Robertson, J., & Bowlby, J. (1956). Responses of young children to separation from their mothers. *Cour du Centre Internationale de L'Enfance, 2,* 131–142.

Rodgers, R. J., & Cole, J. C. (1993). Anxiety enhancement in the murine elevated plus-maze by immediate prior exposure to social stressors. *Physiology and Behavior, 53,* 383–388.

Sams-Dodd, F. (1996). Phencyclidine-induced stereotyped behaviour and social isolation in rats: A possible animal model of schizophrenia. *Behavioural Pharmacology, 7,* 3–23.

Samson, H. H., Tolliver, G. A., Lumeng, L., & Li, T.-K. (1989). Ethanol reinforcement in the alcohol nonpreferring rat: Initiation using behavioural techniques without food restriction. *Alcoholism: Clinical and Experiment Research, 13,* 378–385.

Sanger, D. J. (1991). Screening for abuse and dependence liability. In P. Willner (Ed.), *Behavioural models in psychopharmacology: Theoretical, industrial and clinical perspectives* (pp. 486–502). Cambridge, UK: Cambridge University Press.

Schmajuk, N. A. (1987). Animal models for schizophrenia: The hippocampally lesioned animal. *Schizophrenia Bulletin, 13,* 317–327.

Schuster, C. R., & Johanson, C. E. (1981). An analysis of drug-seeking behavior in animals. *Neuroscience and Biobehavioral Reviews, 5,* 315–323.

Seligman, M. E. P. (1971). Phobias and preparedness. *Behavior Therapy, 2,* 307–320.

Seligman, M. E. P. (1975). *Helplessness: On depression, development and death.* San Francisco: Freeman.

Sepinwall, J., & Cook, L. (1978). Behavioral pharmacology of anti-anxiety drugs. In L. L. Iversen, S. D. Iversen, & S. H. Snyder (Eds.), *Handbook of psychopharmacology* (Vol. 13, pp. 345–393). New York: Plenum.

Shaham, Y., Rodaros, D., & Stewart, J. (1994). Reinstatement of heroin-reinforced behavior following long-term extinction: Implications for the treatment of relapse to drug taking. *Behavioural Pharmacology, 5,* 360–364.

Shanks, N., & Anisman, H. (1988). Stressor-provoked behavioral changes in six strains of mice. *Behavioral Neuroscience, 102,* 894–905.

Sheehan, D. V., Raj, B. A., Trehan, R. R., & Knapp, E. L. (1993). Serotonin in panic disorder and social phobia. *International Clinical Psychopharmacology, 8* (Suppl. 2), 63–77.

Shepherd, J. K., & Rodgers, R. J. (1989). Acute and chronic effects of the triazolobenzodiazepine, alprazolam, on defeat and analgesia evoked by conspecific attack in male mice. *Behavioural Pharmacology, 1,* 75–83.

Sherman, A. D., Saquitne, J. L., & Petty, F. (1982). Specificity of the learned helplessness model of depression. *Pharmacology, Biochemistry and Behavior, 16,* 449–454.

Simiand, J., Keane, P. E., & Morre, M. (1984). The staircase test in mice: A simple and efficient procedure for primary screening of anxiolytic agents. *Psychopharmacology, 84,* 48–53.

Solomon, R. L., Kamin, L. J., & Wynne, L. C. (1953). Traumatic avoidance learning: The outcomes of several extinction procedures with dogs. *Journal of Abnormal and Social Psychology, 48,* 291–302.

Soubrie, P. (1986). Reconciling the role of central serotonin neurons in human and animal behaviour. *Behavioral and Brain Sciences, 9,* 319–365.

Stephens, D. N., & Andrews, J. S. (1991). Screening for anxiolytic drugs. In P. Willner (Ed.), *Behavioural models in psychopharmacology: Theoretical, industrial and clinical perspectives* (pp. 50–75). Cambridge, UK: Cambridge University Press.

Stewart, J., & de Wit. H. (1987). In M. A. Bozarth (Ed.), *Methods of assessing the reinforcing properties of abused drugs* (pp. 211–227). New York: Springer.

Stewart, J., de Wit, H., & Eikelboom, R. (1984). Role of unconditioned and conditioned drug effects in the self-administration of opiates and stimulants. *Psychological Reviews, 91,* 251–268.

Stolerman, I. P., Samele, C., Kamien, J. B., Mariathasan, E. A., & Hague, D. S. (1995). A bibliography of drug discrimination research, 1992–1994. *Behavioural Pharmacology, 7,* 643–668.

Suomi, S. J. (1976). Factors affecting responses to social separation in rhesus monkeys. In G. Serban & A. Kling (Eds.), *Animal models in human psychobiology* (pp. 9–26). New York: Plenum.

Swerdlow, N. R., Lipska, B. K., Weinberger, D. R., Braff, D. L., Jaskiw, G. E., & Geyer, M. A. (1995). Increased sensitivity to the sensorimotor gating-disruptive effects of apomorphine after lesions of medial prefrontal cortex or ventral hippocampus in adult rats. *Psychopharmacology, 122,* 27–34.

Tamminga, C. A., & Gerlach, J. (1987). New neuroleptics and experimental antipsychotics in schizophrenia. In H. Y. Meltzer (Ed.), *Psychopharmacology: The third generation of progress* (pp. 1129–1140). New York: Raven Press.

Thompson, T., & Schuster, C. R. (1964). Morphine self-administration, food-reinforced and avoidance behaviors in rhesus monkeys. *Psychopharmacologia, 5,* 87–94.

Treit, D. (1985). Animal models for the study of anti-anxiety agents: A review. *Neuroscience and Biobehavioral Reviews, 9,* 203–222.

van Riezen, H., & Leonard, B. E. (1990). Effects of psychotropic drugs on the behavior and neurochemistry of old factory bulbectomized rats. *Journal of Pharmacology and Experimental Therapeutics, 47,* 21–34.

Vernon, R. E. (1963). *Personality assessment: A critical survey.* London: Methuen.

Vogel, J. R., Beer, B., & Clody, D. E. (1971). A simple and reliable conflict procedure for testing anti-anxiety agents. *Psychopharmacologia, 21,* 1–7.

Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3,* 1–14.

Weiner, H. (1973). Some thoughts on behavioral approaches to therapy. *Psychological Record, 23,* 441–450.

Weiss, J. M. (1970). Somatic effects of predictable and unpredictable shock. *Psychosomatic Medicine, 32,* 397–408.

Weiss, J. M., Bailey, W. H., Goodman, P. A., Hoffman, L. J., Ambrose, M. J., Salman, S., & Charry, J. M. (1982). A model for the neurochemical study of depression. In M. Y. Spiegelstein & A. Levy (Eds.), *Behavioral models and the analysis of drug action* (pp. 195–223). Amsterdam: Elsevier.

West, A. P. (1990). Neurobehavioral studies of forced swimming: The role of learning and memory in the forced swim test. *Progress in Neuropsychopharmacology and Biological Psychiatry, 14,* 863–877.

White, F. J., & Wang, R. Y. (1983). Differential effects of classical and atypical antipsychotic drugs on A9 and A10 dopamine neurons. *Science, 221,* 1054–1057.

Wikler, A. (1973). Dynamics of drug dependence: Implications of a conditioning theory of research and treatment. *Archives of General Psychiatry, 28,* 611–616.

Wilkins, W. (1971). Desensitization: Social and cognitive factors underlying the effectiveness of Wolpe's procedure. *Psychological Bulletin, 76,* 311–317.

Wilkinson, L. S., Killcross, S. S., Humby, T., Hall, F. S., Geyer, M. A., & Robbins, T. W. (1994). Social isolation in the rat produces developmentally specific deficits in prepulse inhibition of the acoustic startle response without disrupting latent inhibition. *Neuropsychopharmacology, 10,* 61–72.

Willner, P. (1984). The validity of animal models of depression. *Psychophamacology, 83,* 1–16.

Willner, P. (1985). *Depression: A psychobiological synthesis.* New York: Wiley.

Willner, P. (1986). Validation criteria for animal models of human mental disorders: Learned helplessness as a paradigm case. *Progress in Neuropsychopharmacology and Biological Psychiatry, 10,* 677–690.

Willner, P. (1989). Towards a theory of serotonergic dysfunction in depression. In P. Bevan, A. R. Cools, & T. Archer (Eds.), *Behavioural pharmacology of 5HT* (pp. 157–178). New York: Erlbaum.

Willner, P. (1990). Animal models of depression: An overview. *Pharmacology and Therapeutics, 45,* 425–455.

Willner, P. (1991). Behavioural models in psychopharmacology. In P. Willner (Ed.), *Behavioural models in psychopharmacology: Theoretical, industrial and clinical perspectives* (pp. 3–18). Cambridge: Cambridge University Press.

Willner, P., Harding, S., & Eaton, G. (1995). Subjective and behavioural evaluation of cigarette cravings. *Psychopharmacology, 118,* 171–177.

Willner, P., Moreau, J.-M., Nielsen, C., Papp, M., & Sluzewska, A. (1996). Decreased hedonic responsiveness following chronic mild stress is not secondary to loss of body weight. *Physiology and Behavior, 60,* 129–134.

Willner, P., Muscat, R., & Papp. M. (1992). Chronic mild stress-induced anhedonia: A realistic animal model of depression. *Neuroscience and Biobehavioural Reviews, 16,* 525–534.

Willner, P. & Papp, M. (1997). Animal models to detect antidepressants: Are new models necessary to detect new agents. In P. Skolnick (Ed.), *Antidepressants: New pharmacological strategies* (pp. 213–234). Totowa, NJ: Humana.

Willner, P., Phillips, G., & Muscat, R., (1991). Suppression of rewarded behaviour by neuroleptic drugs: Can't or won't, and why? In P. Willner (Ed.), *The mesolimbic dopamine system: From motivation to action* (pp. 251–271). New York: Wiley.

Willner, P., Sampson, D., Phillips, G., Fichera, R., Foxlow, P., & Muscat, R. (1989). Effects of isolated housing and chronic antidepressant treatment on cooperative social behaviour in rats. *Behavioural Pharmacology, 1,* 85–90.

Willner, P., Towell, A., Sampson, D., Muscat, R., & Sophokleous, S. (1987). Reduction of sucrose preference by chronic mild stress and its restoration by a tricyclic antidepressant. *Psychopharmacology, 93,* 358–364.

Wise, R. (1987). Intravenous drug self-administration: A special case of positive reinforcement. In M. A. Bozarth (Ed.), *Methods of assessing the reinforcing properties of abused drugs* (pp. 117–141). New York: Springer.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition.* Stanford, CA: Stanford University Press.

Yanagita, T. (1987). Prediction of drug abuse liability from animal studies. In M. A. Bozarth (Ed.), *Methods of assessing the reinforcing properties of abused drugs* (pp. 189–198). New York: Springer.

Yokel, R. A. (1987). Intravenous self-administration: Response rates, the effects of pharmacological challenges, and drug preferences. In M. A. Bozarth (Ed.), *Methods of assessing the reinforcing properties of abused drugs* (pp. 1–33). New York: Springer.

Young, A. M. & Herling, S. (1986). Drugs as reinforcers: Studies in laboratory animals. In S. R. Goldberg & I. P. Stolerman (Eds.), *Behavioral analysis of drug dependence* (pp. 9–67). London: Academic Press.

Zacharko, R. M., Bowers, W. J., Kokkinidis, L., & Anisman, H. (1983). Region-specific reductions of intracranial self-stimulation after uncontrollable stress: Possible effects on reward processes. *Behavioural Brain Research, 9,* 129–141.

Zacharko, R. M., Lalonde, G. T., Kasian, M., & Anisman, H. (1987). Strain specific effects of inescapable shock on intracranial self-stimulation from the nucleus accumbens. *Brain Research, 426,* 164–168.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.08

Cultural Dimensions

JAN BRØGGER
*Norwegian University of Science and Technology,
Trondheim, Norway*

1.08.1 INTRODUCTION 234
1.08.1.1 Anthropology and Psychology 234
1.08.1.2 The Birth of Clinical Psychology 234
1.08.1.3 The Construction of Social Reality 234
1.08.2 THEORIES OF ILLNESS AND DESTINY 235
1.08.2.1 Theories of Natural Causation 235
1.08.2.2 Theories of Supernatural Causation 236
1.08.2.3 Types of Mystical Causation 236
1.08.2.3.1 Fate 236
1.08.2.3.2 Ominous sensations 236
1.08.2.3.3 Contagion 236
1.08.2.3.4 Mystical retribution 237
1.08.2.4 Theories of Animistic Causation 237
1.08.2.4.1 Soul loss 237
1.08.2.4.2 Spirit aggression 238
1.08.2.4.3 Magical causation 238
1.08.2.4.4 Witchcraft 238
1.08.3 SOCIAL SYSTEMS AND TYPES OF EXPLANATION 239
1.08.3.1 Non-Western Systems of Understanding 240
1.08.3.1.1 Complex societies 240
1.08.3.1.2 Early Greek medicine 240
1.08.3.1.3 The Indian medical system 242
1.08.3.1.4 Arabic medicine 243
1.08.3.1.5 Chinese medicine 246
1.08.3.1.6 American medicine 247
1.08.3.2 Small-scaled Societies 248
1.08.3.2.1 Africa 248
1.08.3.2.2 New Guinea 250
1.08.4 THE GROWTH OF THE WESTERN SYSTEM OF UNDERSTANDING 251
1.08.4.1 The Role of the Psychiatric Hospital 251
1.08.4.2 The Cross-cultural Issue and Human Nature 251
1.08.4.3 The Interface Between Psychology and Anthropology 252
1.08.4.4 The Impact of Anthropological Fieldwork 254
1.08.4.5 Culture and Personality 254
1.08.4.6 The Paradigm of Ethnic Identity 255
1.08.4.7 The Cross-cultural Correlational Approach 256
1.08.5 SOCIAL SYSTEMS AND THE INDIVIDUAL 257
1.08.5.1 The Dominant Kinship Dyad 257
1.08.5.2 The Iatrogenic Approach 258
1.08.6 REFERENCES 258

1.08.1 INTRODUCTION

1.08.1.1 Anthropology and Psychology

The collaboration between anthropology and psychology is a recent chapter in the brief history of the behavioral sciences. The interest of anthropologists in the problem of mental illness originated in their concern to understand the relationship between culture and personality. In the early 1930s Margaret Mead, Edward Sapir, and Ruth Benedict were raising questions of how anthroplogy could inform the understanding of the genesis of the human personality.

From 1936 to 1940 the psychiatrist Abraham Kardiner with a group of anthropologists, the most prominent of whom were Cora DuBois, Ralph Linton, Edward Sapir, and Ruth Benedict, were conducting seminars in culture and personality. A spinoff of this seminar was the first comprehensive cross-cultural research program conducted on the interface between anthropology and psychology.

The investigation was carried out on the Indonesian island of Alor under the leadership of DuBois. "We had talked ourselves out, and only field work could test the procedure," she wrote in her introduction to *The people of Alor* (DuBois, 1944), a landmark in cross-cultural psychology. The scope of this investigation was, in her own words, to find out if there was a "demonstrable" relationship between the personalities of adults within a group and the sociocultural milieu in which they lived. If such a relationship were found to exist, its explanation was presumed to lie in the consistency of life experiences ranging from the earliest child-rearing practices and relationships to the reinforcing effects of adult institutions and social roles (DuBois, 1944).

The interest of these pioneers in cross-cultural anthropology, however, was not primarily clinical. In the school of thought initiated by DuBois and her collaborators which flourished from the 1930s and well into the 1950s, the question of the "normal" and "abnormal" and the definition of illness in various societies was not the main issue. The Holy Grail of their investigations was "basic personality structure"—a term coined by Abraham Kardiner or "modal personality" in the language of DuBois. Yet it is these early investigations that have been the main source of inspiration for the current interest in the relationship.

Although the term "clinical psychology" was introduced by Lightner Witmer in 1896, several decades had to pass before the role of the clinical psychologist gradually emerged. The participation of psychologists in mental institutions and clinical work was restricted to psychometrics and mental tests in a team which, in addition to the psychologist, consisted of a psychiatrist and a social worker. The psychologist had an academic training which in the beginning did not include training in psychotherapy and clinical methods, although the Department of Psychology at Columbia University established a curriculum for clinical psychology in 1936. But formal, systematic training of clinical psychologists in the USA did not get off the ground before the 1940s when the psychological damage caused by wartime experiences primarily among soldiers became a major social problem among the more than 16 million World War II veterans which were added to the 4 million from the previous war.

1.08.1.2 The Birth of Clinical Psychology

The time was ripe for clinical psychology, but not yet for a closer collaboration between anthropology and clinical psychology. In the landmark event in the history of clinical psychology which took place in Boulder, Colorado, in the summer of 1949 where the "Boulder model" of clinical training was formulated, neither anthropology nor cross-cultural studies were mentioned in the core curriculum proposed by the conference.

1.08.1.3 The Construction of Social Reality

The relevance of anthropology for clinical psychology thus has only recently been recognized by the community of scholars, and a consistent training program has yet to be formulated. The most obvious potential of cultural understanding to the field of clinical knowledge is in the realm of *social reality*. A few epistemological comments are called for in order to clarify this concept which is essential to the postmodern understanding of culture which is at the cutting edge of anthropological theory today.

There is an unresolved disagreement between proponents of the idea that all phenomena can be studied by the same scientific methods and scholars who maintain that there are phenomena of different orders which call for special methods. The first group subscribes to a theory of unity of science and is sometimes, in colloquial discourse, referred to as positivists or logical empiricists. In the social sciences in general and in psychology in particular the positivist approach is referred to as behaviorism.

The second group maintains that there are certain phenomena which cannot be studied by positivistic methods but must be approached through a special form of understanding colloquially referred to as hermeneutics. Psychoanalysis is an example of a system of understanding based on the hermeneutic method.

The idea of social reality is based on the observation that social life is not of the same order as the physical universe which has an existence independent of human socializing. Social life on the other hand is created or acted out in a social process. In the words of Berger and Luckman (1969) it is "constructed." Social reality is constructed in the process of everyday life through role play following an intricate system of rules. A pioneering study of human interaction was carried out by Goffman (1959).

The anthropological understanding of culture has been strongly influenced by the constructionist approach. The concept of culture developed by the Columbia group under the leadership, of among others, Margareth Mead and Ruth Benedict reflected taxonomic ideas developed in the study of species. These taxonomic ideas are based on the fact that, in the fauna and flora, distinct species can be recognized as isolates with clear boundaries. Although the classification of cultures was not explicitly organized according to the taxonomic paradigm of species, cultures were to a certain degree, regarded as isolates with their distinct and recognizable patterns.

This taxonomic thinking strongly influenced the culture and personality school of thought which flourished in the 1940s in the study of national character. These studies clearly demonstrated that the social sciences are involved in political agendas. The national character studies started as part of the war effort and were designed to understand, and to a certain degree, predict the behavior of the main adversaries—the Germans and the Japanese, and later during the cold war, the Russians. In spite of the political agenda, some of the studies testified to a high quality of scientific craftsmanship. An example of this is Ruth Benedict's *The chrysanthemum and the sword* (1974) in which she presents a study of Japanese culture and personality.

Although cultures are no longer regarded as species, there are definite relationships between culture and personality. Distinct cultures can be recognized, but the focus is now on the nature of their boundaries. The boundaries between species are genetic, a fact which makes them clear isolates. The boundaries between cultures are maintained by entirely different processes. The boundaries are not natural entities, but constructions maintained by complex symbolic efforts of self-ascription, rituals, and social sanctions. The focus is now on ethnic groups and boundaries rather than cultures. This approach was launched with the publication of the influential work *Ethnic groups and boundaries* by Fredrik Barth (1969). The question of ethnic groups articulates with the problems of personal identity which is an important issue in clinical psychology. One of the important interfaces between anthropology and clinical psychology is located in the symbolic space of identity management.

1.08.2 THEORIES OF ILLNESS AND DESTINY

Every human group that has been studied by anthropologists has developed its own form of knowledge to cope with the everpresent adversaries of our existence: disease and misfortune. Two of the pioneers of medical anthropology George M. Foster and Barbara Gallatin Anderson classified theories of illness in technologically simple societies into "personalistic" and "naturalistic" theories (Foster & Anderson, 1978). This classification corresponds to the classification later suggested by George Peter Murdock in his last comprehensive cross-cultural investigation, *Theories of illness. A world survey* (1980). Murdock's investigation is based on the human relations area files at Yale University. On the basis of these files Murdock surveyed theories of illness in 186 human groups or cultures in the language of the 1980s.

1.08.2.1 Theories of Natural Causation

Murdock identified four main theories of natural causation:

(i) *Infection*. Defined as the invasion of the victim's body by noxious microorganisms. It refers particularly to the germ theory of disease which seems to be unique to Western medicine. But some groups have knowledge and understanding of the fact that small organisms like worms and insects may cause disease.

(ii) *Stress*. From the point of view of clinical psychology it is interesting to observe that exposure of victims to psychic strain such as worry, fear, and emotional distress is recognized as a major cause of illness in a large number of groups. Among the Javanese and the Pawnee anxiety is recognized as the major threat to health.

(iii) *Theories of deterioration*. Murdock observes, not without irony, that physical deterioration through old age is not generally recognized as a major cause of illness among human beings. According to Murdock, most of mankind considers itself potentially immortal and is unable to conceive of the infirmities of old or middle age as being caused by the process of aging itself rather than the interference of some hostile agent or force.

(iv) *Accident*. The fact that accidents are rarely regarded as the causes of illness and misfortune indicates that the very notion of accident is a recent philosophical idea in human understanding. The general tendency is to personalize the causation of what we in the modern Western world term accident and to attribute it to supernatural intervention.

1.08.2.2 Theories of Supernatural Causation

Through his comparative analysis Murdock came up with three basic groups of supernatural causation: theories of mystical causation, theories of animistic causation, and theories of magical causation. Of the three, the theories of mystical causation may be regarded as the most sophisticated, entailing as they do impersonal agents as the culprits. In certain ways they are closest to modern theories; particularly the belief in contagion is very close to our notion of infection as if the mystic theories by some sort of intuition forecast the breakthrough of modern medicine. Since the mystical theories are less concerned with social forces than the other explanations, they are at first glance, of less interest to the anthropologist and have received much less attention and are also comparatively rare—but not of less theoretical interest as Mary Douglas has shown in her fascinating study *Purity and danger* (1966).

More widespread, however, are the theories of animistic causation, which imply the existence of personalized supernatural entities. In accordance with this kind of explanation, disease may be caused by the occasional departure of a person's soul, as in the cases of soul loss, but much more frequently through spirit-aggression. According to Murdock's investigation, this is the most frequently encountered type of explanation.

Theories of magical causation are also found in a large number of different cultures and are manifest in two main forms: witchcraft and sorcery. For several reasons, not all of them strictly professional, these twin phenomena have been the favorite of anthropology. One reason is possibly that witchcraft and sorcery loom with sordid prominence in more recent Western history. The topic appeals to the peculiar kind of black fascination which human tragedy and catastrophe are singularly capable of arousing. It provides its audience with the combined advantages of participation and immunity, which never fails to rally the crowds also in our age. It gives modern people that sense of intellectual achievement which sustains the interest in the follies of the recent past. Clearly supernatural ideas also have a special appeal to anthropological field-workers because it was a prominent theme in the heroic age of discovery at a time when anthropologists regarded technologically simple communities as primitive. Although "primitive" has developed a semantic of its own, and today is regarded as a pejorative term, its lexical origin is the Latin term *primus*, which means first. The term primitive indicated that the scholars in the days when words like primitive and savages did not raise an eyebrow even among the most enlightened, were of the opinion that they were studying humankind in its pristine or original state of being. In a way the opinion was that they were investigating our common past.

1.08.2.3 Types of Mystical Causation

1.08.2.3.1 Fate

In the realm of mystical causation fate is of particular interest. It was regarded as the major determinant of illness among the ancient Romans. Fate may be regarded as the precursor of our modern idea of accident or chance. In Roman times fate was understood as the influence of astrological influence, individual predestination, or ill luck. In our modern idea of accident the connection to astrology and occult forces has been suppressed although it is still alive in many folk beliefs which can be ascertained by anyone who cares to consult the horoscopes of popular magazines. It is a sophisticated idea which is primarily found in comparatively complex societies such as Turkey, Burma, Thailand, and Java.

1.08.2.3.2 Ominous sensations

This type of theory is of particular interest because it gives prominence to the impact of potent dreams which has been recognized by modern psychoanalysis. In addition to dreams, impressive sights and sounds may cause illness. Dreaming of a recently deceased relative is an ominous sign among the Massa. Depending on the circumstances such dreams may be important events during psychoanalytic treatment and are another example of the intuitive wisdom sometimes encountered in knowledge systems of technologically simple societies.

1.08.2.3.3 Contagion

The word suggests an incipient understanding of the microbiological transfer of disease. But the mystical concept of contagion which is fairly widespread and attested to by 49 societies in the Murdock sample is of a different order altogether. The mystical contagion refers to fear of the ritually impure, like menstrual blood and corpses. Although the fear of corpses could have a rational dimension in terms of microbiology, the same does not hold for menstrual blood. It is a question of pollution rather than infection and is cognitively related to the purportedly polluting categories of people, such as the "untouchables" of India.

This widespread theory of disease is of considerable interest to clinical psychology because it testifies to a human, cognitive disposition which has been explored by Douglas (1966).

Douglas observes that the impure in cognitive terms is matter out of place. In the physical world order it is recognized by everything in its proper place; in the moral order, impurity is equal to the lack of observance and respect for categories. The classical example is the taboo on pigs' meat which according to Douglas has nothing to do with an assumed archaic knowledge of microbiological hazards, which is the common folk explanation, but with the categorization of animals into ruminating and nonruminating types. This is written quite explicitly in the Old Testament where it is maintained that the pig is unclean because it has split hooves but does not ruminate. This is an important distinction to people of archaic animal husbandry.

Another example is the fear of homosexuality in societies with rigid sex-role segregation. To regard the homosexuals as unclean is not in accordance with modern rationality, but is of the same order as the impurity of untouchables.

Although Westerners, at first sight, regard the contagion theories of other societies as "superstitious" and "irrational," an understanding of the cognitive disposition revealed by Douglas demonstrates that some of our Western theories of illness are merely posing as rational. The fact that homosexuality is no longer regarded as a disease or pathological condition in the USA demonstrates that some of our own theories of illness, at least until quite recently, were based on ideas of irrational impurity.

1.08.2.3.4 Mystical retribution

The idea that disease is a retribution for the violation of moral injunction or taboos is fairly widespread. The most common of the taboos are food, sex, as in the case of incest, and verbal.

Again it is possible to observe that ideas of mystical retribution are not entirely absent in Western society. It is, for example, being maintained by some religious leaders that AIDS is a divine retribution for the violation of sexual taboos.

Food taboos are common among certain ethnic groups in modern Western societies. The origin of these taboos is generally not known and poorly understood. Yet they are observed with care by people who subscribe to these taboos and great discomfort is to be expected in case of violations. It is important for therapists to understand the power of taboos and beliefs. These beliefs and their effects clearly demonstrate that social reality is a construction and reminds us of the fact that human beings are under the spell of different constructions.

Verbal taboos are recognized in our rules against swearing. But, the fear of swearing has lost most of its mystical force and among modern people is regarded merely as bad manners whereas in some societies the breach of verbal taboos may be regarded as the cause of illness.

1.08.2.4 Theories of Animistic Causation

1.08.2.4.1 Soul loss

A common idea among human beings is the existence of an insubstantial double or soul which inhabits the body, but is independent of the physical and territorial restrictions which apply to the physical body. This soul may depart from the body during sleep and permanently at death to lead a separate existence in afterlife. If the soul departs from the body on more than minor excursions during sleep, it may cause illness.

The belief in soul loss is often combined with the institution of shamanism, which has recently attracted some attention among psychotherapists as pointed out by Strathern (1995). Strathern points out that in recent developments in psychotherapy altered states of consciousness have become a main topic of interest. He observes that:

> Trance facilitates such a reframing (of experience), as does spirit possession. At one end we may be dealing with a new image presented to a patient, at the other with a whole new persona, an alter ego juxtaposed to the self. Either a metaphor or a historical figure may be in focus: in either case the result is an altered state of consciousness which whether trance induced or not conduces towards a patient's recovery, a process in which health and identity are brought into consonance with each other. (Strathern, 1995, p. 128)

The term shaman is of Siberian Tungus origin, but the shaman is a well-known character in "anthropological" societies. He or she is a curer who has a special relationship with the spirit world. During consultations, the shamans fall into trance in which they communicate with their tutelary spirits in order to understand the cause of the illness.

The recent interest in shamanism and trance among anthropologists is caused by a major paradigmatic change in the field of cultural understanding. In the classical tradition in the study and interpretation of the personality, the personality was regarded as a reflection in the individual mind of the society and its institution, or in other words: its culture. With the constructionist approach the self was introduced as a concept. Whereas the personality is imposed on the individual by society, the self is an active agent reacting to and sometimes even rebelling against the demands of society. The concept is difficult to operationalize and seems rather nebulous to ardent empiricists. Nevertheless it has acquired some respectability in mainstream anthropology.

1.08.2.4.2 Spirit aggression

This is the most widespread of all the theories of supernatural causation. The major types of supernatural aggressors are nature spirits, disease demons, or lesser divinities. Among the aggressors are also found departed ancestors or kinsmen, ghosts and in some instances deities or gods. The thought is not alien to Christian civilization as is testified in the Bible by the book of Job.

1.08.2.4.3 Magical causation

As pointed out above, witchcraft and sorcery looms with sordid prominence among the theories of illness and destiny in comparatively recent Western history. In the case of sorcery an active human agent uses various techniques to hurt another human being. The main techniques applied by sorcerers are supposed to be verbal spells, prayers, and curses; object intrusions; ritual manipulation on exuvias and belongings of the intended victim; and administration of poisons.

The distinction between rational procedures known to be effective within a scientific paradigm of understanding and irrational procedures is not always clearcut. This is particularly the case with the application of poison. Sorcery has an instrumental dimension which may be reconciled with the modern world view. The difference between the biochemical effects of certain medicines and the occult effect of the medicines of sorcerers is not always obvious to people with no scientific training. Beliefs in sorcery may therefore be quite resilient.

1.08.2.4.4 Witchcraft

In witchcraft beliefs it is assumed that some human beings possess an evil force that may cause disease and misfortune. The force is supposed to be activated automatically, sometimes even without the conscious knowledge of the witch, as demonstrated by Evans-Prichard in his classical monograph *Witchcraft, oracles and magic among the Azande* (1937).

Sorcery on the other hand is a technology of evil. A sorcerer can apply an assortment of procedures based on what we may refer to as irrational knowledge to harm people. A sorcerer is a specialist and may learn his craft.

Witchcraft and sorcery are often combined and, one is tempted to say, confused with one another. In strict anthropological terms, the witches of popular belief in Europe were really sorcerers, who sometimes were supposed to have acquired their evil powers from the Devil himself.

The belief in witchcraft is practically universal in the Circum-Mediterranean region, but is comparatively rare elsewhere in the world except Africa. The prominence of witchcraft in European and American history and thinking is due to the fact that witch craze is a phenomenon primarily of Western culture.

Belief in witchcraft does not always produce witch crazes. In many African societies, witchcraft is regarded as a normal phenomenon, as among the Azande, and does not produce the kinds of moral panic which is characteristic of early modern Europe.

Although popular opinion has it that witchcraze was a social phenomenon of the Middle Ages, its heyday was in the Renaissance. As the British historian Trevor-Roper has pointed out, belief in witchcraft and sorcery in the sixteenth and seventeenth centuries, a period which we imagine as the age of reason, was not

> lingering ancient superstition, only waiting to dissolve. It was a new explosive force, constant and fearfully expanding with the passage of time. It was forewarned by the cultivated popes of the Renaissance, by the great Protestant reformers, by the saints of the Counter-Reformation, by the scholars, lawyers, and churchmen of the age of Scalinger and Lipsius, Bacon Grotius, Beralle and Pascal. (Trevor-Roper, 1970, p. 121)

The craze also hit the USA in Salem, Massachusetts, in the seventeenth century and many completely innocent people were executed.

Witchcraze is of particular interest to clinical psychology as an example of a collective paranoia. Contrary to popular opinion the days of moral panics are not over and it is important that psychologists are aware of the background and dynamics of moral panics because they are sometimes called to the assistance of victims of irrational or pseudorational persecutions.

Although we have not yet found an uncontroversial understanding of moral panics and the rise of the witchcraze in the Renaissance, we have gained some insight into the relationship between social systems and types of explanation of disease and misfortune.

1.08.3 SOCIAL SYSTEMS AND TYPES OF EXPLANATION

Lewis has pointed out that it may sometimes be more illuminating to discard the culturally grounded expressions "spirit-possession" and "witchcraft" altogether and think in terms of oblique and direct mystical attacks (Lewis, 1971). It has been demonstrated by Brogger (1992) that the nature of oblique and direct mystical attacks clearly relate to the nature of dependency between an individual and the community. The crucial dimension seems to be ascribed vs. acquired relationships and statuses on the one hand and dyadic vs. corporate relationships on the other. Ascribed relationships, of course, are those an individual acquires at birth and as a member of his or her community. In small, technologically simple societies, there is little scope for choice and contractual relationships are few. The dominant mode of the traditional village is the ascribed relationship, that of the urban society the acquired form. Dyadic relationships are limited to *one* other person, corporate relationships to "corporations" such as a lineage, an age grade in technologically simple societies, or in modern societies to trade unions or political parties. These dimensions are represented in Table 1.

Acquired dyadic relationship is characteristic of the modern urban situation. In societies where the essential necessities of life are acquired through ascribed, dyadic relationships, witchcraft beliefs will tend to prevail. This does not exclude sorcery, which may be found as a residual category. Where ascribed, corporate relationships dominate, spirit aggression seems to prevail. In cases where acquired corporate relationships dominate, (as in the modern, bureaucratic state), political demonology may develop as part of the belief-system.

Table 1 Corporate and dyadic relationships

	Dyadic	*Corporate*
Ascribed	witchcraft	spirit-aggression
Contractual	sorcery	political demonology

Although the table does not cover all the possible explanations of illness and misfortune its focus on the experiential dimension in exchange relationships adds a comparatively unexplored dimension to the role of economics in the patterning of culture and society. It directs our attention to the cognitive repercussions of the rise of bureaucracy and the introduction of money and markets.

These institutions may have been a prerequisite for a rational understanding in the field of human relationships and therefore the main force behind the breakthrough of scientific thinking in general and of a rational theory of illness in particular.

In systems in which the majority of transactions are sustained within a framework of clearly personal relations, the human intellect easily jumps to the conclusion that people or clearly recognized groups of people not only have economic power, but power over destiny in general and thus over diseases and indeed over life and death. On the basis of a general experience of social life under these conditions, ideas of witchcraft, sorcery, and even spirits do not seem far-fetched.

The general breakthrough of the scientific theory of illness therefore owes as much to the development of market economy and bureaucratic management as to knowledge itself. The fact that many, even when exposed to the ideas of modern medicine, do not resign their beliefs in witchcraft and sorcery clearly demonstrates that knowledge which does not articulate with the experiences of everyday life remains esoteric (Brogger, 1971).

Dependency on the depersonalized forces of bureaucracy, money, and markets, however, does not irreversibly guarantee rational explanation of disease and misfortune. In spite of many impressive achievements, it is as true as ever that when struck by misfortunes and disease, human beings look for someone to blame. As in former times, one's perspective is determined by the nature of one's dependencies. The age of demonology thus is not behind us, only its vocabulary has changed.

Although the modern human being's freedom of fancy with regard to health is restrained by the insights of modern medicine, one is as free as ever to speculate on the social causes of destiny. In tune with the nature of one's dependency, demonology does not operate on the matrix of face-to-face relationships, but on the level of institutions and social categories like race and class. The demonological qualities of national socialism (Nazism) are universally recognized, but this is not the only example. Whether modern witchhunts are based on the assumed conspiracies of "world Jewry" or the "bourgeois capitalist" the effect on its victims is the same: suppression, persecution, and often extermination. Although an anthropology of knowledge cannot promise to free the world from the savagery performed in the name of irrational ideologies, it may provide knowledge which, biding its time, may deprive aggressive ideologies of some of their demonic allure.

These anthropological insights have an important message to clinical psychology because they are a reminder of the forces of irrationality in mental life. It is of particular importance to appreciate that modern rationality is a world view or *Weltanschauung* which has developed under particular social circumstances. Although the majority of people in the Western world regard the scientific world view as true in an absolute sense, anthropology reminds us of the great variety of world views developed in different societies as demonstrated in the varieties of explanations of destiny and illness.

At the same time social science demonstrates that even scientific rationality is merely "skin deep." Behind a front of rationality, human passions often subvert the conquests of scientific reason. It is often the task of the therapist to clarify how egocentric irrationality is the cause of psychological disturbance, and it is sometimes the duty of social scientists to represent the voice of reason in the public domain in situations where moral panics threaten justice and human dignity. Anthropological insights may also threaten our somewhat ethnocentric convictions of having found the master key to truth by raising the question of whether our world view is merely a belief system among many others developed by human civilizations.

1.08.3.1 Non-Western Systems of Understanding

1.08.3.1.1 Complex societies

The various theories of illness presented in the global overview of Murdock testify that all human groups develop disease theories no matter how simple. In the evolution of knowledge, which in an anthropological perspective cannot be understood without attention to the development of political systems, more complex systems of understanding have developed and gained acceptance and influence in wide areas.

The interdependence between knowledge and social institutions is often overlooked as if knowledge has an independent existence outside the sphere of human organization. The fact is that the crystallization of medical knowledge into comprehensive systems has followed in the wake of political empire building. Medical knowledge of course is only a special case, knowledge in general is dependent upon the power and communication structures of societies.

No matter how unpopular empires are in the political thinking of everyday life in the Western world, their importance for the development and management of knowledge is enormous.

The consequences of the Arab expansion after the seventh century AD is today common knowledge and therefore a convenient example. Through the medium of the Arab language and lines of communications created by the Arab rulers, the fund of knowledge accumulated in three of the great civilizations of the Occident and the Near East—The Byzantine, the Greco-Roman, and the Persian—were brought together. In the wake of these events was written an important chapter in the evolution of Western medicine. The knowledge reached Europe primarily through Moorish Spain. At the universities in Europe the medical text, *Canon of medicine*, by the Persian scholar Ibn Sina (980–1037), known by his latinized name Avicenna, was part of the curriculum for several hundred years and helped to set the stage for the development of modern Western medicine.

The great medical traditions of Greek, Indian, Arab, Chinese, and American societies will be explored in the following pages.

1.08.3.1.2 Early Greek medicine

One of the basic ideas in Greek medicine is the role of the bodily "humors." The theory of illness based on the so-called humoral pathology is known to us through the writings of Hippocrates (born around 460 BC), the father of Western medicine. According to Hippocrates, health is a state of equilibrium between the four humors with their associated qualities: blood (hot and moist); phlegm (cold and moist); black bile, also known as *melancholy* (cold and dry), and yellow bile or *choler* (hot and dry).

In the text attributed to Hippocrates we are told that

> The human body contains blood, phlegm, yellow bile and black bile. These are the things that make up its constitution and cause its pains and health. Health is primarily the state in which these constituent substances are in the correct proportion to each other, both in strength and in quantity, and are well mixed. Pain occurs when one of the substances presents either a deficiency or an exess, or is separated in the body and not mixed with the others ... The four humors have specific and different names because there are essential differences in their appearance. They are dissimilar in their qualities of heat, cold, dryness

> and moisture ... Phlegm increases in the winter because as the coldest humor, it is most in keeping with the winter. During the spring the quantity of blood increases, taking charge of the body during the summer and autumn. The hot and cold summer weather is conductive to yellow bile, but, as the cool and dry autumn comes on the bile is cooled, and the black bile preponderates ... The physician must bear in mind that each disease is prominent during the season most in keeping with its nature. (Chadwick & Mann, 1950, pp. 204, 206, 207, 208)

Hippocrates also specifies cures.

> Diseases caused by over-eating are cured by fasting; those caused by starvation are cured by feeding up. Diseases caused by exertion are cured by rest; those caused by indolence are cured by exertion. To put it briefly, the physician should treat disease by the *principle of opposition* to the cause of the disease according to its form its seasonal and age incidence, countering tenseness by relaxation and vice versa. This will bring the patient most relief and seems to me the principle of healing. (Chadwick & Mann, 1950, p. 208)

The most important organs of the body according to early Greek medicine were the heart, the brain, and the liver. These organs were respectively dry and hot, moist and cold, and hot and moist—the healthy body was thought to have an excess of heat and moisture. But each person had his or her particular temperament according to the personal prominence of bodily humors: The "sanguine," cheerful and optimistic; the "phlegmatic," calm and somewhat slow and sluggish; The "bilious," ill-tempered and indeed choleric; and the "melancholic," sad with a tendency for depression and in fact melancholic.

The resilience of the Hippocratic medicine is immediately obvious to us as we recognize that the term "temperament" relates to the notion of temperature in the theory of humor pathology and the categories of temperament are still maintained in everyday language in the West.

The therapies were also logically derived from the theory and were directed against a re-establisment of an humoral balance by means of diet, purging, vomiting, bleeding, and cupping—therapies which are still part of folk medicine in the Mediterranean.

The theory of humor pathology foreshadowed or may be influenced the direction of thought in Western medicine where today the biochemistry of the body is a promising field of investigation and therapy.

In the field of psychiatry, the German psychiatrist Ernst Kretschmer developed a typology of temperaments where somatic characteristics of the body take the place of the Greek humors. Kretschmer recognized three main physical types: pyknic, athletic, and asthenic (leptosome) and added a fourth "mixed" type, the dysplastic. The pyknic type is round in form and has a greater quantity of fatty tissue compared to the leptosome who tend to be relatively thin and of linear build. The term athletic speaks for itself. Kretschmer identified two main types of psychotic dispositions: the schizoid (expressed in schizophrenia) and the cycloid (expressed in manic-depressive psychosis). In the normal population he identified nonpathological dispositions which he referred to as schizothymia, characterized by introversion, sensitivity, and seriousness; and cyclothymia with a more extrovert disposition, affectivity, and a tendency towards mood swings. According to Kretschmer, pyknic individuals tend to be cyclothymic, whereas asthenic, athlethic and dysplastic individuals tend to be schizothymic (Kretschmer, 1925).

In the USA William Sheldon developed a typology according to the same paradigm, although he discarded the strict categorizations of Kretschmer and introduced a taxonomic system based on a systematic study of thousands of college students. He concluded that human beings could be classified under three primary components of body build: endomorphy, mesomorphy, and ectomorphy. The endomorph is reminiscent of the pyknic type characterized by softness, roundness, and a large digestive system. The mesomorph has strong bones and muscles, in other words is athletic. The ectomorph is thin and has a large skin area and nervous system. He recognized three essential dimensions of temperament: viscerotonic, somatotonic and cerebrotonic, assuming that temperament was determined by the somatic disposition (Sheldon & Stevens, 1942).

In this context we are not primarily concerned with the scientific merit of the theories of temperament and personality but in demonstrating the historical background of scientific thought.

A crucial institution for the influence of Greek medicine was the library of Alexandria which was founded in the third century BC. With the manuscripts of this library at his disposal, Galen (130–200 AD) gave the theory of humors its final shape. Although he mainly practiced his medical knowledge in Rome, Galen was an authentic Greek and through his writings Greek medicine conquered oriental Christianity and Islam.

Through the East-Roman orbit, Byzantium, the teachings of Galen reached eastern Christianity and Islam. A crucial role in this transmission of knowledge was played by the

Nestorian sect of Christianity in Edessa in Mesopotamia where Galen's work was translated into Arabic. Near the end of the fifth century the Nestorians had to flee Mesopotamia for political reasons and settled in Gund-Shapur in Persia, present day Iran. A center of medical knowledge was established at Gund-Shapur which strongly influenced Islamic medicine and the Greek theory of humor pathology became part of the Arabic medical tradition. Of particular interest is Gund-Shapur as a mediator between Greek and Indian medicine, the Ayurvedic school.

1.08.3.1.3 The Indian medical system

As one of the great civilizations India has developed its own medical system which is part of the health care of modern India. The roots of Indian medicine can be traced to the Vedic writings from the latter part of the second millennium BC. In the *Rig Veda* we encounter the term *bhisaj* a word which later became synonymous with *vaydya* which still today is the Indian term for the traditional curer. In the early period disease was believed to be due to punishing gods and evil demons. We recognize the interpretation of disease in Murdock's theory of *spirit aggression*. The god Varuna was associated with moral ideas and punished those who transgressed his commands.

The idea of healing is associated with the divine twin gods, the Asvins. In the *Atharva Veda* demons are presented as the responsible agents and part of the text is in fact made up of spells to cure disease and promote longevity.

As pointed out by Basham (1976) in the centuries following the *Atharva Veda*, the last six centuries BC, the traditional Indian medical system developed into something like its present day form. What is today known as *Ayurvedic* medicine has come down from Sanskrit documents from the first century BC. The most important of these is *Caraka Samitha*. Caraka is the supposed author of the text, and he may be regarded as an Indian Hippocrates.

By the first century AD a naturalistic system had developed in India. Although beliefs in personalistic causation prevail until this day on the sub-continent, Ayurvedic medicine is a system with scientific and empirical merit. The term may be translated as "the science of living to a ripe age."

According to Ayurvedic theory, the universe is composed of four elements: earth, water, fire, and air. The human body has three fluids or humors, in Indian *dosha*: phlegm, bile and gall, and wind. In a healthy individual the three doshas are in balance or equilibrium, this notion is consequently referred to as the *tridosha* theory. The doshas relate to the age and the seasons: phlegm with youth and the growing season, bile with middle age and the rainy season, and the wind with old age and winter.

The ideas of temperature are also found in the Ayurvedic circle of ideas. Many foods are thought to have heating and cooling qualities. Hot food, *garam*, include eggs, meat, milk, honey, and sugar; cold foods, *tonda*, include fruit juices, yogurt, rice, and water. These notions are reflected in the theory in which diet has an important part to play.

It is interesting to appreciate that during the period when Ayurvedic medicine reached its mature state, a strict system of socio-religious taboos regulating the menus of the Hindus evolved. There is no indication to the effect that these taboos, which support the hierarchy of castes according to degrees of ritual purity, were developed with the aim of promoting health or avoiding disease. But as Basham (1976) points out, it is surprising how many of the instructions in the *Atharva Veda* would tend to minimize the dangers of infection and food poisoning. "Indian society seems unconsciously to have found a means of remaining healthy as far as possible in a subtropical climate, in its efforts to preserve its ritual purity" (Basham, 1976, p. 19).

This statement recalls the discussion of the religious food taboos in the Western orbit. In our age of reason, we tend to project our rational understanding back to the times when the taboos were codified, times when there was no rational understanding to sustain the dietary habits. The results of the dietary and also hygienic taboos sometimes appear to have had practical effects as if they were based on rational understanding of germs and infection. It is tempting to hypothesize that there is an intuitive wisdom in religious taboos, an intuition which antedates the reflexive understanding of modern science.

The ideas of purity and danger explored by Mary Douglas (1966) demonstrate that the dangers of pollution emerge when essential categories are insulted. The orderliness which these notions promote have obvious prophylactic effects, as is the case when the dead body and things affiliated with it are shunned. The Jewish taboos and hygienic rules definitely served to promote health. Although not the explicit intension of the taboos, they helped to increase the hygienic standards. The Jews shunned stagnant waters, like the wells in the medieval cities, which were sometimes infected. The fact that the Jews made a special effort to secure running water was from our point of view sensible enough. But it cost them dearly because some Christians spread the rumor that the wells had been poisoned by the Jews.

Although too simplistic, ideas of evolution are regarded with suspicion by cosmopolitan anthropology. It is a historical fact that the development of the medical systems both of the Greeks and the Indians followed a trend from the personalistic to naturalistic types of explanation. As we shall see, Chinese and Arabic medicine followed the same developmental trend. It is an empirical fact that naturalist and rational explanations emerged with the development of states and institutions. This does not in itself demonstrate a causal connection between development of complex societies and rationality, but invites a hypothesis to that effect. The secret is most likely hidden in the social organization of knowledge and the rationalization of economy and administration.

The great medical systems of the Orient stopped short of developing the reflexive empirically based medical science of the West. They did, however, develop prophylactic procedures, curing techniques, and mental training programs which are now gradually being recognized as valuable by Western medical science.

Of particular interest is yoga which does not strictly belong to the Ayurvedic curriculum. The term "yoga" has two Sanskrit roots, the first means "to meditate," the second "to join". By joining is understood the connection of the human nature with cosmic sources. The development of yoga as a technique of psychological healing may be regarded as an evolution of early magico-religious beliefs and practices into a rational system. Seen in the light of a theory of knowledge this development is of the same nature as the development of empirical, medical theories on the basis of early, personalistic ideas exemplified by Greek and Indian humoral theories.

Yoga has definitely developed on the basis of religious ideas (Hindu and Buddhist) but is not in itself a religion and is not based on faith, dogma, and sacred texts. Yoga represents a system of mental training developed over milennia by practicing yogis.

The yoga philosophy clearly testifies to its religious parentage. It is based on a rejection of wordly life, an idea shared by Christianity. Its aim is an altered state of consciousness. In the West a general scepticism to yoga has prevailed until quite recently. But, empirical studies suggest that relatively subtle effects produced by different yoga practices can be detected by brain-wave studies (Emerson, 1972).

Underlying yogic practice is the idea of the "subtle body," which consists of a series of interconnected centers called "chacras." The chacras have the ability to process energy from the environment and refine it to a higher level. Seven major centers are recognized. The third-eye center, for example, is located an inch above the meeting of the eyebrows and relates to insight and understanding. These centers are normally closed and it requires conscious effort on the part of the meditator to open them.

The yogic terminology and the sometimes extravagant claims made on its behalf has made it too easy for Western science to disregard the knowledge as obscure. Its promotors in the West also sometimes have proved to be authentic charlatans who have exploited the existential crises of modern men and women to make money. This, and a tendency towards cultural arrogance, has made it too easy for Western science to disregard insights from Asian psychology. But recently, mental health professionals have started an exploration of the Asian techniques on a personal basis. Studies of altered states of consciousness have revealed the complexity of consciousness which for a long time was ignored by behaviorist psychology in spite of the influence of, for example, William James. He demonstrated an open-minded approach:

> Our normal waking consciousness is but one special type of consciousness, whilst all about it, parted from it by the flimsiest of screens, there lie potential forms of consciousness entirely different. We may go through life without suspecting their existence, but apply the requisite stimulus, and at a touch they are there in all their completeness. No account of the universe in its totality can be final which leaves these other forms of consciousness quite disregarded. How to regard them is the question. At any rate they forbid a premature closing of our accounts with reality. (James, 1914, p. 298)

Recent advances in Western science seem to facilitate a more open-minded attitude to Asian psychology and medical practice and theory (Walsh & Vaughan, 1980).

1.08.3.1.4 Arabic medicine

Arabic medicine has two main sources. One is Greek which the Arabs incorporated into their rational system of knowledge during their most enlightened period from the tenth to the twelfth century. The other source is the medical tradition of the bedouins which gradually merged with the teachings of the prophet Muhammed, and is crystallized under the term "prophetic medicine."

The two bodies of medical knowledge have been in conflict with each other primarily because of the prestige of the prophet and the theocratic influence of the clergy. The Greek medical tradition was regarded as alien by the rank and file of the *ulama* the theocratic scholars who, until this day, dominate much of the intellectual life in Islamic countries.

Seen from the perspective of the tenth century AD it is inconceivable that the great advances in medicine should emerge in the Christian West. After the decline of the Roman Empire and with it the Greco-Roman civilization, elite Greek medicine was rescued by the Arabs. During the first three or four centuries after the Hijra, Arabic became the common language of the Islamic world and the language of science. Greek medicine was taken over wholesale by the Arabs who, to the degree that scientific medicine survived, stuck to it until they were subjected to the impact of European science in the nineteenth century. The translation of the Greek medical sources was performed by a comparatively small group of scholars, primarily in the ninth century. Hundreds of Greek works were translated into Arabic and became part of the scientific legacy during the golden age of Arab dominion. The Hippocratic and Galenic systems were saved for world civilization by these early scholars.

Arab practitioners added to the body of knowledge developed by the Greeks the subjects of pharmacology and surgery, which flourished in Andalusia; and ophthalmology, which flourished in Egypt and Mesopotamia where eye diseases for climatic and ecological reasons were the scourge of the people.

This body of knowledge was, as pointed out above, systematized in a treatise on medicine by Ibn Sina, Avicenna, and partly through his famous canon of medicine (*al-qanum fit- t-tibb*). Greek medicine was restored in the West. It was part of the curriculum in Salerno for centuries. The Greeks and the Arabs also made contributions to psychiatry, which of course was an unknown term before the nineteenth century.

The high standard of medical practice in the Arab domain was to a large degree promoted by the medical center in Gund-Shapur. This pre-Islamic center not only promoted an exchange between Indian and Greek medicine, but set a standard for medical craftsmanship which made the physician a respected personality in cosmopolitan Arab culture, although the school was in its heyday run by Nestorian scholars and physicians. The academy was probably founded in the first half of the sixth century after the Academy of Athens was closed in 529.

It is known that when the founder of the Abbasid Caliphate, al-Mansur, came down with a stomach disease and his regular court physicians were helpless, the chief physician at Gund-Shapur was called to Baghdad. When he succesfully cured the ruler, the prestige of Gund-Shapur was established and with that its high standards and medical ethics with its roots in Hippocratic traditions. It is probably also due to the influence of this academy that Arabic medicine by the Hindu and Urdu speakers is known as *Unani*. Unani is the English spelling of the Arabic term *Yunani* which means Greek, that is, Ionian.

Of particular interest is Greek and Arab psychiatry. The Greek physician Rufus of Ephesus wrote a treatise on melancholia which was translated and supplemented by Ishaq ibn Imran, a court physician in Kairouan, at the beginning of the tenth century. Melancholy was treated with medicine and psychotherapy which included music.

Considerable insight in mental illness is testified by the treatments recommended for fixed ideas. The main point is that the ideas were taken seriously and not dismissed, as they sometimes are by modern physicians. The treatment was based on the authority of the physician. It seems that Rufus of Ephesus had reflected on the power of authority of the formation of cognitive states which independently of the Greek tradition, was introduced in European clinics by Franz Anton Mesmer (1734–1815) in his so-called magnetic treatment based on the power of suggestion. Although Mesmer's method was dismissed as fraudulent, he was the first to introduce a systematic application of the authority of the practitioner as a cure of hypochondriacs and hysterics. A monument of his achievement has been raised in the English lexicon in the form of the term "mesmeric."

Profound psychosomatic understanding appears between the lines in Arab medical texts. This is particularly the case with regard to the demeanor and lifestyle demanded of the physician. He had to be a respectable person of high ethical standards and the bedside manners were an essential part of the cure.

Some of the great practitioners are surrounded by a heroic lore which has survived until the present day. Of particular interest is the story of the great Razi who healed the emir of Rhagae (in ancient Persia). This ruler had a paralyzed leg and Razi understood that the ailment was a hysterical condition. Razi ordered drinks and a hot bath "to ripen the humors." But suddenly he drew his knife as if he was prepared to kill the emir, who sprang to his feet, cured of his hysterical ailment. Razi, of course, did not use the term hysteria, but the story testifies to a remarkable psychological insight (Bürgel, 1976, p. 51).

The great Arab scholar Ibn Khaldun (1332–1406) wrote a famous treatise on the decline of Arab civilization which had flourished from the tenth to the twelfth century.

After that the development of knowledge, including medicine, did not just come to a standstill, but fell into a deplorable decline. The reason is complicated and the explanations that are offered are controversial. Some maintain that the reasons are political, as did Ibn Khaldun; others assume that the reasons are found in the economic developments. What we do know, however, is that one of the effects of the decline was a resurgence of personalistic explanations in the realm of medicine. Bürgel (1976) gives the following explanation of this turn of events:

> If we want to know which spiritual forces were most potent in paralyzing the scientific impetus of the golden age, we have to ask what the essentials of science are. I think that one essential is imminent to science, its very soul and and life-blood, and this is rational argumentation propelled by an insatiable curiosity. The other thing essential to science belongs to the history of economy as well as to the history of science, and it is the interest in, the demand for, and the consumption of scientific products. What forces were harmful to these essentials? As for the economic element, a strong and ever-increasing influence was exercised by Islamic mysticism, which turned people to the inner instead of the outer world so that they restricted their natural demands to a minimum. As to curiosity, I have noted the prevailing concept of science as a constant stock of knowledge. Rational thought had several renowned enemies, some of whom could trace their origins to antiquity. I refer to astrology, alchemy, magic and, finally, of Islamic origin, the so called prophetic medicine. These four were looked upon as sciences by the great majority, and even by most of the scholars. Nevertheless, they were hothouses of irrationalism, the rational disguise making them even more harmful. (p. 54)

We recognize in the resurgence of personalistic theories of causation the question of destiny. In monotheistic Islam the basic question with regard to illness was its religious meaning. Pious people regarded illness as the will of God and wondered whether rational cures could be regarded as an obstruction of their ordained destiny. Illness and suffering were regarded as an atonement for sin and it was common opinion that people suffering from, for example, blindness or epilepsy were assured a position in Paradise because of their patient suffering.

Suffering in itself came to be regarded as meritorious sublimely expressed in a cult of martyrs. "He who dies on a sickbed, dies the death of a martyr and is secure against the inquisition of the tomb." This saying is attributed to the Prophet according to *Hadith* the body of lore which is used to give prophetic legitimacy to opinions and procedures not explicitly sanctioned by the Koran. According to the tradition, the Prophet claimed that "Healing resides in three things: a draught of honey, a cut by the cuppingglass (scarification), and the branding by fire (cauterization). But as for branding I forbid it in my people" (Bürgel, 1976, p. 57).

The belief coded in this quotation refer to honey as a medicine. It is still regarded as such at the margins of Islamic culture, for example, among the Sidama of Ethiopia. The cupping is still practiced in the Mediterranean where it has developed a special form through the application of glass. A cup is heated and applied to the surface of the skin. When it cools, a vacuum-effect sucks the skin into the cup. It may be regarded as a milder form of therapeutic infliction of wounds. Cauterization is still practiced as a folk remedy in the periphery of the Islamic orbit. A common practice is to drill a piece of wood, as in the old technique of fire-making. When the stick almost catches fire through the energy of drilling, it is applied to the skin. A number of burns are believed to cure minor sickness. Through prophetic medicine folk remedies were given medical legitimacy, among which was the bedouin idea that all diseases were caused by disorders of the stomach.

Magic healing practices became widespread, and indeed are still widespread in the Islamic orbit. Of particular prominence is the magical protection of talismans. Most visitors to Muslim countries have noticed that children have small bundles around their necks with scriptural talismans. Words from the Koran or the Prophet written on small pieces of cloth or paper are regarded as a necessary protection. A widespread practice in former times was to write holy texts on a washable material, rinse it and give the rinsings to the patient as a medicine. In the popular lore, the Prophet is reported to have performed ritual washing at the bedside and afterwards to have administered the water to the sick as a drug. The lore and some of the Hadith should of course be taken with a grain of salt. Charlatans have always tried to exploit the reputation of prophets and sages to their economic benefit.

The history of the decline of scientific medicine in the Islamic orbit after the twelfth century is only one example of the vulnerability of rational knowledge. Although it cannot be proved, it is probable that empirical science is dependent upon a division of labor through which economic and political interests are divorced from the management and application of knowledge. Seen from the point of view of history, the worst possible situation in the management of knowledge is theocratic monopolies. Theocracies have often been passionate opponents of rational knowledge and have not hesitated to oppress and even execute voices of reason and knowledge. Criticism of prophetic medicine therefore was practically nonexistent. One of the few critics, like Ibn Khaldun, pointed out that it was not part of the Prophet's mission to teach medicine. Another was the Andalucian poet Ibn al-Khatib who repudiated the explanation of plagues given by prophetic medicine pointing out that the validity of the traditions should not be questioned if they were incompatible with the senses (Ullmann, 1970).

The Islamic world today is struggling to reconcile Western science with their traditional *Weltanschauung* and, as Bürgel points out,

> The Islamic world has for over a century entered upon a new period of acculturation. From its very beginning, Islam has revealed a peculiar skill in accommodating and amalgamating foreign cultural influences without losing its own identity. It is just now proving that it has lost nothing of this marvelous faculty. But if there is no chance for the Galenic system as a whole to be revived in our day, one may think of reviving particular achievements—for example, certain well-tried therapies and old drugs—which are in fact now being scientifically tried out and reproduced on a large scale by Unani pharmaceutical companies and state pharmacies in Pakistan, India and Sri Lanka. (Bürgel, 1976, pp. 60–61)

1.08.3.1.5 Chinese medicine

Chinese medicine has evolved as a system of knowledge almost completely independent of the other great civilizations—the Greco-Roman, the Arab, and the Indian. Whereas these three systems have, as we have seen, strongly influenced each other, contact with China before the age of discovery was negligible. Also, in the development of Chinese medicine we can recognize the shift from personalistic to naturalistic explanation which is the *sine qua non* (absolute necessity) of rational knowledge. The traditional Chinese medicine is best known through the classic work *Huang Ti Nei Ching Su Wen* which is claimed to be from the third century BC. However, recent historical research (Veith, 1972) has demonstrated that it most likely was compiled during the Han dynasty (202 BC–221 AD). This proves that naturalistic explanations emerged in China at about the same time as in Greece and India. It is notable that this happened during a period when economic and political developments had reached a degree of maturity calculated on the basis of comparative studies as with the establishment of money, markets, and orderly government.

The basic concepts of Chinese medicine are today common knowledge in the West. The two principles of life, "yin" and "yang," are operating in the human body as everywhere else in the Chinese cosmos. Health is dependent upon the right balance between these two principles. Yang represents heaven, sun, fire, heat, dryness, light, the male principle, the exterior, the right side, life, high, noble, good, beautiful, virtue, order, joy, and wealth—all the positive elements of being. Yin represents the opposite: earth, moon, water, cold, dampness, darkness, the female principle, the interior, the left side, death, low, ignoble, bad, ugly, vice, confusion, and poverty (Foster & Anderson, 1978). Excessive yang causes fever, yin produces chills.

The hot–cold dichotomy in traditional Chinese medicine is reminicent of the Ayurvedic tradition, but probably developed independently in China. It was probably developed as far back as 180 BC. But the major work on the hot–cold dichotomy is from the fourteenth century. In Chia Ming's work *Essential knowledge for eating and drinking* (*Yin-shih-hsü-chih*). (AD 1368), he describes

> 43 kinds of fire and water, 50 kinds of grains, 87 kinds of vegetables, 63 kinds of fruits and nuts, 33 "flavorings" and condiments, 68 varieties of fish, 34 kinds of fowl, and 42 kinds of meat! The entry for each of these 460 entries tells to which of the five flavor categories it belongs, its "character" (specified degrees of hotness or coolness) and the other foods that should be eaten with it. The character of natural rainwater is cold, while that of water from a stalactite cavern which is warm; the flavor of both is sweet. Glutinous rice is said to be warm, and eaten in excess causes fever. Soybeans and fragrant leeks are warm, and eaten in excess causes fever. (Foster & Anderson, 1978, p. 64).

Acupuncture is regarded as a "cold" operation especially suited to diseases caused by an excess of yang.

Diseases are classified according to causation into external and internal diseases. Diseases caused by external forces are yang diseases, yin diseases by internal forces. In Chinese thinking yin and yang have been understood as a single entity combining in any manifestation in the cosmos both positive and negative elements.

Chinese cosmology recognizes five elements: water, fire, metal, wood, and earth. This idea of basic elements which was also part of Greek cosmology developed independently in China. Although this idea from our position of knowledge seems primitive and erronous, it cannot be denied that it put the human mind on the right track towards modern physics and chemistry. In the great hide and seek of knowledge, the Chinese were at the portal of rational, empirical science, but stopped short of entering.

Manfred Porkert has pointed out that Chinese thinking remained in the inductive mode:

> "The scientist who uses the inductive and synthetic mode of cognizance will observe first, and then speculate on his observation; the scientist who supplies the causal and analytic mode will first speculate and act, and after that he will observe. My hypothesis is that, at the dawn of systematic speculation, philosophers of all civilizations first adopted the inductive and synthetic mode of cognition because of its directness and simplicity. Causal analytic thought apparently did not enter the scene before the 4th century BC. In China it only constituted a brief interlude in the works of the mohists and of Hsün-tzu; in the West, the opus of Aristotle marked the beginning of a tradition that was to continue into our days. (Porkert, 1976, p. 63)

Although its concepts seem foreign to the Western mind, it cannot be regarded as inferior to the Hippocratic tradition, the concept and ideas of which were not entirely different from the Chinese with regard to the theory of basic elements. A respectable body of knowledge was accumulated in the Chinese tradition and the efficiency of some of the cures, therapies, and drugs are no longer questioned.

The Chinese physicians never reached the prestige accorded to modern Western physicians or their Arab colleagues in the golden age. According to Confucian philosophy, which is above all concerned with social ethics and the relation of the individual to society, the top of the social hierachy was the literatus. The tiller of the soil was given the second rank. At the bottom were the soldiers and the merchants. Physicians were considered to be technicians (*shu-shi*) and were assigned the middle ranks of society together with engineers, craftsmen, and artists.

Chinese medicine confirms Benedict's ideas on patterns of culture (1934). Although her ideas are no longer part of mainstream anthropology, Chinese culture's meticulous concern for detail and the patient execution of craftsmanship whether in cooking, art, or medicine demonstrates a consistent pattern which never fails to impress the observer. Their belief in their own tradition of knowledge, also in medicine, is demonstrated by the persistence of the hot–cold dichotomy in the medical traditions of modern Hong Kong. Acupuncture is about to become part of Western medicine although its mode of operation in the human body is still unknown.

1.08.3.1.6 American medicine

The situation in America is unique and different from the other civilizations we have been discussing. The reason for this is of course that the civilizations of America more or less disintegrated under the impact of European conquest. The knowledge systems of the Aztec, the Inca, and the Maya have not been adequately brought forward to our times because the social organization which sustained them broke down and no adequate writing system was developed to codify the knowledge in texts. Although Arab medicine went into a period of decline, medical centers like Gund-Shapur served as a guarantor against total decay and the traditions had been safely recorded in a medical canon. The lack of a basic canon is the main reason why pre-Columbian medical knowledge has only survived at the so-called folk level.

Although the terms folk medicine and folk level may offend some egalitarian sensibilities, it is a meaningful term from the point of view of the organization of knowledge. The vast bodies of knowledge accumulated by the complex civilizations need specialized institutions to be stored, retrieved, and above all applied. A pedagogic system of sorts is necessary to transfer knowlege from one generation to the next. When this system breaks down, people everywhere have struggled to save some of the knowledge through their own management. Also, in the civilizations which maintained their political integrity, rational medical knowledge survived, although outside the legitimate orbit defined by clerics as in some of the Islamic countries, varieties of folk medicine flourished. When people have little access to expert knowledge, they resort to curers and try the best they can to retrieve the necessary knowledge from the lore of their families and the local community and in the case of European immigrants to the USA, from books.

The folk medicine of America has been primarily studied by folklorists, but in recent years anthropologists have joined forces with the folklorists in the development of ethnomedicine.

In the USA a plurality of folk medicines exists. Each immigrant group has to some extent developed its own folk medicine. But it is customary to recognize three distinct types: Euro-, Black, and Spanish-American.

Euro-American folk medicine is unique because most of the immigrants were able to read and thus were not merely dependent upon oral traditions. But in the early years of frontier life, particularly in the nineteenth century, doctors were limited and not always elite experts. However, Indian experts were highly respected as "powwow" doctors. Pickard and Buley claim that "in some western communities in the early years there were native American doctors who were held in quite as high repute as regular white doctors" (Pickard & Buley, 1945, p. 36).

A part of American culture in general is the do-it-yourself tradition which has created a special literary genre. Richard Carter's *Valuable vegetable medical prescriptions for the cure of all nervous and putrid disorders* (1815) was widely distributed.

The native American knowledge came to its right in books like S. H. Selman's *The Indian guide to health* (1836), and William Daily's *The Indian doctor's practice of medicine* (1848). An impressive number of home remedy books in the medical mainstream tradition were also printed, the best known is possibly John Gunn's *Domestic medicine* or *Poor man's friend, in the house of affliction, pain and sickness*, first published in 1830. By 1885 it had gone through 213 editions.

The high level of literacy among many of the immigrant settlers gave the folk medicine a professional touch which is unique and may be regarded as an intermediate type of knowledge from the point of view of social organization. Through literacy, the line of communication from specialist centers to the people was shortened and knowledge was not subjected to the same kind of distortions we know from oral traditions.

The best known of the "authentic" folk medicine of the Euro-American type is that of the Pennsylvania Dutch (Amish). As Gebhard (1976) points out, ideas still alive among the Amish, for example, in the saying "Feed a cold and starve a fever," date back to Celsus (50 AD) and the use of cobwebs to cure bleeding goes back to Galen.

The black folk medicine represents a purely oral tradition. The best known variety is the so-called "voodoo" or "conjure" which developed in the early nineteenth century around New Orleans. Black folk medicine has three main sources: Catholicism, seventeenth and eighteenth century European occultism, and traditional African medicine and belief.

Snow classifies the black explanation of disease as "natural" and "unnatural": "The first includes deficiency in protecting the body against bad weather, and divine punishment for sin. The second has to do with the individual's position as a member of society" (Snow, 1973, p. 272). As "unnatural" Snow refers specifically to a belief in witchcraft, a theory of illness which Foster and Anderson (1978) characterize as personalistic. "Nothing else in the fieldwork elicited so much emotional response as the question about witchcraft. Whether answers were negative or affirmative, the vehemence was often startling" (Foster & Anderson, 1978, p. 274).

The belief in witchcraft is probably of European, not African, origin. African witch belief has never reached the degree of hysteria as in Europe and later in the USA, and witch craze is an authentic Euro-American phenomenon. The beliefs, therefore, have probably been taken over from the European settlers. The idea of protecting the body against bad weather has a ring of humoral pathology which suggests Greek influences. What is of authentic African origin is the herbal lore and practical cures and remedies including a most capable midwifeship. African elements are certainly also part of voodoo rituals.

Spanish-American folk medicine represents a more coherent system than the others and is based on an equilibrium model of health based on a theory of balance between hot and cold *calidades*. Illness is the result of an excess of heat or cold, and the origin of these notions is fairly obvious. Spanish–American folk medicine is based on the Greek pattern. Since Spain, particularly in Moorish times, was the center of medical knowledge in Europe, it is not surprising that the Greek–Arab articulation of humoral pathology dominated the medical scene in the homeland of the *conquistadores*. It was brought to the New World with the European settlers and remained the major medical paradigm well into the eighteenth century. From the elite centers of medicine in Spanish-speaking America it found its way to the mestizo peasants and acculturated native American groups. To a large degree it replaced preconquest medicine. However, it has been suggested that, for example, the hot–cold dichotomy in Mexican folk medicine has its roots in Aztec beliefs (Lopez-Austin, 1974).

1.08.3.2 Small-scaled Societies

1.08.3.2.1 Africa

In Murdock's (1980) comparative study we noticed the great variety of explanation of disease and misfortune. A typical example of spirit aggression is recorded from the Sidama people of southwestern Ethiopia and may serve as an example both of theory and practice in this "medical genre." It will also give a portrait of a prominent character in the medical practice in small-scaled societies, the shaman (Brogger, 1986).

Disease among the Sidama is commonly explained as the result of spirit-aggression. The Sidamas are firmly convinced that they are surrounded by an army of spirits.

Some of the spirits live in large trees, others near the river. They are invisible, but are capable of invading the bodies of men and talking through the mouth of the person they choose as their host. The spirits are capable of striking men and animals with disease and misfortune, but can also be of help if they are treated properly.

The spirits rarely strike only once. They prefer to establish a lasting relationship with a person. When such a relationship is established, the spirit demands regular attention through spirit sessions, *hayatas*, in which communication with the spirits is based on a state of dissociation or trance in the host. This state is reached through drumming and singing during the spirit session.

When a relationship with a spirit has been established, it cannot be dissolved and lasts until the person dies. When a man dies, it is very often claimed that he was killed by his spirits. To get some systematic knowledge of the spirit cult, a sample of 29 cases was selected from one local curer. Although the spirits are regarded as heavy burdens, most people were afraid of losing them. A spirit usually heralds its choice of host with a sickness. Of the 29 cases recorded, 20 started with some sort of sickness. The individual histories of acquisition have the same general pattern, but with some variations of details. The following case is fairly typical, although richer in details than most. This man was 65 years old at the time of the fieldwork, and had entertained the relationship with the spirits for more than 30 years.

> I hated the sheitanna in the beginning, I was insulting them. But one day when I came from the market I asked my wife to make coffee for me. As I drank this coffee, I suddenly began to shake, and I did not know what was happening to me, and I asked my wife. She told me that I became very strange. At that time I was the student of Abba (an Italian priest). I explained to him what had happened to me. Abba told me that spirits were dangerous, and he gave me a kind of paper with the image of Jesus Christ and told me to keep it under my pillow and to pray. This cured me for some time. But when the Italians left Ethiopia, Abba left with them, and the spirits caught me again. I got pain all over my body and I asked people what I should do, and they told me to make a hayata and pay gabo, gifts, to the spirits. I gave him a sheep and honey, but although the spirit made me shake, it did not talk. Therefore I brought a kalicha to the hayata, and he made the spirit talk. I had two spirits, Dillo and Warro. These were the masters, but other spirits would come at the same time. These spirits can kill and make a person sick, but they can also be of help. They gave me children, and I built a galma for them. In the early days I had hayatas very frequently, sometimes every two months. That protected me and my family from sickness. My father also had spirits. But his spirits were different. My spirits came from Dulanna, all spirits come from big people like Guto and Dulanna. (Brogger, 1986, p. 66)

Usually the spirit will after some time demand its own house, galma. The galma is a smaller replica of the sidama house, but constructed for the sole purpose of serving the hayata sessions. It is built close to the main hut as a kind of chapel for the households of the compound. Guto and Dulanna are the names of particularly powerful shamans who often introduce new spirits to the sidama "pantheon."

The attacks of the *sheitanna* are in many cases interpreted by a *kalicha*. But in a number of cases the presence of the spirits is so obvious that no consultation is regarded as necessary. When contact has been established by a kalicha and his spirit or spirits, professional assistance will rarely be necessary. Most of the people who served as hosts for spirits understood when the spirits demanded a tribute. This did not happen only through sickness—a spirit could sometimes let its demands be known through spells of instant possession. In a number of cases the spirit would approach a future host through a dream.

A 55-year-old and very active hayata organizer with his own *galma* got his first *sheitani* in the following way: "It first started through a dream. I saw it like a picture of a human being and it told me to make hayata. But I had no disease and I refused. It then reappeared in a dream." (Brogger, 1986, p. 67)

But he still hesitated. The same dream was repeated several times, and gradually his mental state changed during the day. He could no longer concentrate on his work and he sometimes had hallucinations. Sometimes he was talking to himself, and one day he pointed at an imaginary sacrifice and said: "These intestines of the sheep we must bring to the kalicha." Finally, he could no longer resist, and he made his first hayata and became completely normal. From then on he gave hayatas regularly.

Not every Sidama had a spirit; an enlightened guess is that about 10% of the adult population has a long-standing personal relationship with one or more spirits. But those who have no spirit of their own may participate in the hayatas of close relatives and give gifts on these occasions. Thus, most of the community of grown-ups are in one way or another affiliated with one or more spirit. When an ardent spirit-worshipper dies, the spirit is often inherited by a close relative. In fact 10 of the above 29 cases inherited their spirits from their mother, father, or spouse.

The local curer is technically in anthropological terminology, a shaman. Although he is formally concerned with healing, his main service is to articulate social conflicts in terms of the local system of interpretation of disease and misfortune. As the chief administrator of the spirit cult and its main source of authority, he is crucial to our understanding of shamanism.

The political leaders among the Sidama are the elders. The stateliness of the elders contrasts sharply with the demeanor of the shaman. If stateliness is the hallmark of the elders, eccentricity is that of the shaman. They surround themselves with a set of symbols which give them a certain flamboyance, particularly their jewelry and hairstyle. They will often let their hair grow long and appear in robes reminiscent of Arab traders. They adorn themselves with conspicuous rings of twisted copper, bracelets, and necklaces. They show little care for dignity but rather inspire an aura of mystery bordering on the sinister.

The elders are respected, shamans are feared. On the basis of their functions in the local communities the elders could rightly be called priests. They bear the responsibility for communal sacrifices, they may promote fertility through their blessings, and they mediate the tribal communication with the Sky-God, Magano. They are the guarantors of the respectable and the moral order.

The shamans, on the other hand, negotiate the relationship with a number of capricious spirits which haunt the local communities in Sidama. We have seen how the spirits impose themselves on ordinary people, and how communication with them was established through the spirit session, the hayata. Most of the worshippers of spirits merely attend to the problems of their immediate families.

There is no formal process of recruitment of shamans. A shaman establishes himself by his own choice as a public practitioner in the art of spirit-management. In order to project the image of a shaman he has to change his way of dressing and general appearance. He will either build a special hut or arrange his house in a way that facilitates the accommodation of clients. Among ordinary men the spirits are regarded as jealous and demanding, and unnecessary contact with them should be avoided. But when someone is struck by misfortune and sickness, the presence of the spirits is actively solicited so they can be persuaded to withdraw their influence.

The cures of the shamans are efficient within their context of understanding, as in cases where the illness or disturbance belongs to the moral order and has its roots in disturbed social relationships.

The healing traditions of shamans have attracted the attention of modern Western medicine and psychology recently and they cannot be dismissed as charlatans, as demonstrated by Andrew Strathern (1995).

1.08.3.2.2 New Guinea

Among the Baktaman of New Guinea according to Barth (1975) sickness is thought to be caused by both spirits and sorcerers, but also by breaches of taboos. But, maintains Barth,

> Baktaman also operate with a concept that seems close to a Western pre-scientific notion of "disease"—a condition, often communicated or initiated by a particular incident, causing discomfort of recognized kinds and even death. Thus, for example, one man who developed a cold chastised another for having been spitting and blowing mucus all over the place when he had been sick; this had been eaten by the pigs, stored in their fat, and communicated to the patient when he ate the pork. Pressed to explain the nature of what was communicated an informant fell back on the idiom of spirit; but it makes sense for him none the less to declare that a particular condition of being sick is caused not by sorcerers or spirit agents, but "just sickness." (p. 137)

Sickness then can be brought about by the following causes: spirit attacks, male sorcery, female sorcery, breach of taboo, menstrual pollution, and plain sickness.

It is not unreasonable to suspect that the category "plain sickness" has been added to the Baktaman taxonomy recently under the influence of Western medicine, small elements of which were brought to the attention of the tribe by anthropologists.

The Baktaman have developed a "technique" of disease prevention with an interesting and familiar ring: the amulet. Baktaman children, like children in the Mediterranean, wear small bundles around their necks, usually a selection from dog femurs, water-lizard mandibles, cassowary claws, and ginger wood. The spirits from the animals are supposed to guard against spirit attacks. Grown-ups also wear talismans, but their amulets are secret and hidden in minute bags suspended from the neck to hang over the breastbone near the seat of one's spirit consciousness.

Various cures are available, but none of them seem to have a rational basis. They have developed medicines, but their mode of action is magical. For example, spice bark is chewed and blown over sick persons with a combined whistling/blowing gesture normally used in conversations signifying "get lost." Infections are transferred to the taro-plant by injecting a small specimen of pus from the sore under the skin of the taro. Bleeding of the infected area is also practiced. If these cures fail, the ultimate remedy for sickness is pig sacrifice with an appeal to the ancestors with a mountain spirit as an intermediary. The sacrifice is administered by a shaman. It is not reported that the shaman induces trance in the curing process. Trance, however, is an essential prerequisite of a shaman as among the Sidama.

1.08.4 THE GROWTH OF THE WESTERN SYSTEM OF UNDERSTANDING

1.08.4.1 The Role of the Psychiatric Hospital

Until the eighteen century, personalistic explanation of disease, and particularly mental disease, was common in Europe. In the clinical literature it is referred to as animistic explanations (Brogger, 1986). In cognitive terms there is no major difference between the belief in spirit aggression in eighteenth century Europe and the personalistic explanations we have encountered in the ethnographic literature. Among the Sidama the spirits are managed by the shaman, in the European tradition by the priest. Although the shaman and the priest both have the responsibility of relieving the distress caused by assumed spirits, the priest should not be classified with the shaman. The priest never resorts to trance, but holds an office based on some sort of expertise and legitimacy which is dependent upon specialized, bureaucratic institutions. The shaman has to sustain his position through his own charisma and role play. He or she is dominating in small-scaled societies. Priesthood is an institution of complex societies. In the European tradition possessions were treated by exorcism, and the church has codified rituals of exorcism.

The belief that mental distress is caused by evil spirits is surviving, to a degree, even in a modern industrialized society like the USA (Goffman, 1961). But today legitimate treatment is based on the scientific *Weltanschauung* with the psychiatric hospital as one of the core institutions. The psychiatric hospital, however, does not have an honorable pedigree. It evolved from institutions that were created to house the poor, the uprooted, the unemployed, and the insane—in other words the outcasts emerging in the wake of the disruption of the agrarian society at the dawn of industrialism in the seventeenth century. The outcasts in Paris, London, or Philadelphia in the eighteenth century were confined respectively to La Salptriere, Bedlam, and Pennsylvania Hospital.

The insane were given worse treatment than the other inmates. Foucault (1965) give the following description of their situation:

> when the waters of the Seine rose, those cells situated at the level of the sewers became not only more unhealthy, but worse still, a refuge for a swarm of huge rats, which during the night attacked the unfortunates confined there and bit them wherever they could reach them; madwomen have been found with feet, hands, and face torn by bites which are often dangerous and from which several have died. (pp. 70–71)

But by the end of the eighteenth century more humane treatment of the insane was gradually introduced. It is significant that these changes followed in the wake of modernization in the full sense of the word both within the realm of economics and production which, it may be argued, paved the way for the age of reason in Europe and the USA.

The first hospital to remove the chains from the patients was St. Boniface in Florence, Italy. But the most famous event is Philippe Pinel's removal of the chains in his capacity of the newly appointed director of La Bictre in Paris during the French Revolution in 1792.

This was not done without opposition. A famous exchange between Pinel and one of the radical leaders, Couthon, is interesting because it forecasts an understanding of certain psychiatric symptoms which were not generally understood before the middle of the twentieth century. After having been insulted by one of the patients, Couthon turned to Pinel and said: "Now citizen, are you mad yourself to seek to unchain such beasts?" Pinel replied calmly: "I am convinced that these madmen are so intractable only because they have been deprived of air and liberty (Rosenhan & Seligman, 1995, p. 40). By these words Pinel was ahead of his time. The iatrogenic aspect of certain psychiatric symptoms was not generally understood before the American anthropologist Goffman published his work *Asylums* (1961).

1.08.4.2 The Cross-cultural Issue and Human Nature

Two systems of understanding can be recognized at our present stage of knowledge with regard to the cross-cultural issue. First, the universalist approach has as an underlying paradigm that one system of understanding may cover the human family as a whole. The second system, the particularistic, is based on the assumption that cultures must be understood on their own terms and cannot be unified in one master theory. The Indian *Weltanschauung* may, for example, appear as strange and incomprehensible to us, and therefore in terms of Western folk belief regarded as "wrong." If we try to translate this *Weltanschauung* into the terms of our own and, for example, insist that the cows are holy simply because they are economically important, we act as universalists. A particularist will respect the Indian view and try to understand and even acquire it as a system of knowledge. It is not a question of an absolute dichotomy, and a taxonomic scheme may be a little premature.

During a symposium at the University of California, San Diego, in 1980, a group of leading scholars in the field of cultural theory met to discuss the state of their art. One of their objectives was to establish contact with the psychological sciences, for in the words of Shweder and LeVine (1984):

> The symbol-and-meaning approach to culture seems to be an unusually well kept secret in the psychological sciences, most notable among developmental researchers working on such culture-saturated phenomena as personhood and self, morality and convention, social cognition and interpersonal relationships, self regulation and emotional response. It was the best bet of the planning group that this was the right time to bring recent developments in culture theory "out of the closet."

It is relevant at this stage to point out that American anthropology made great contributions when a group of scholars met and established some sort of consensus and joined forces under a common paradigm of understanding. The Columbia group of the 1930s is a case in point.

The group met in San Diego in 1980 and was following in the wake of an established tradition of the organization and management of knowledge in the USA, and may serve as an example of a cultural pattern in itself. Although no clear consensus came out of their efforts, they set the agenda for cultural studies for the next decade or so and some of their key participants, Richard A. Shweder, Robert A. LeVine, Roy G. D'Andrade, and others, with Clifford Geerz as an inspiring pioneer of cultural studies, have taken over the torch from the Columbia group.

During the next decade the field of cultural psychology was developed. Stigler, Shweder, and Herdt (1990) summarize its scope with the following statement:

> Cultural psychology is the study of the way cultural traditions and social practices regulate, express, transform, and permute the human psyche, resulting less in psychic unity for humankind than in ethnic divergences in mind self, and emotion. Cultural psychology is the study of the ways subject and object, self and other, psyche and culture, person and context, figure and ground, practitioner and practice live together, require each other, and dynamically, dialectically, and jointly make each other up. (p. 1)

A comprehensive curriculum has yet to be designed.

1.08.4.3 The Interface Between Psychology and Anthropology

One of the great pioneers in psychology, the German Wilhelm Wundt, was one of the first to use ethnographic material to inform psychological problems. He wrote in the genre which today is referred to as culture and personality. But although his influence in the growth of modern psychology is enormous, his writings on cultures never caught the imagination of his colleagues.

It was above all Sigmund Freud who became the pioneer in the exploration of the interface between psychology and anthropology. In a number of works he tried to explain how psychoanalysis could be used to explain the origins and functioning of social institutions. Freud assumed that his model of the structure of the mind was valid cross-culturally. He assumed that the phenomena of repression, by which painful experiences are kept out of the consciousness by unconscious forces, was universal. He regarded rituals and taboos he studied in the ethnographic literature as caused by processes similar to those he studied in his consulting-room in Vienna, that is, to ward off anxiety. His views have been summarized by Bock (1988, p. 32) in the following illuminating phrase: "Culture is to society as neuroses is to the individual."

Psychoanalytic anthropology initiated by Sigmund Freud caught the imagination of several generations of psychologists and anthropologists, particularly in the USA where it has served as a major source of inspiration until the present.

The pioneering work in psychoanalytic anthropology is *Totem and taboo* (Freud, 1950). In this ground-breaking work, Freud approaches these two classical anthropological problems. The term taboo is today part of the language of everyday life with only a slightly extended semantic sphere of reference compared to its strict anthropological meaning. In anthropology today it covers all types of prohibitions from the incest taboo to Jewish and Indian dietary rules. In Freud's thesis, the Polynesian taboo served as a point of departure, and the term is linguistically Polynesian.

Freud approached his patients in Vienna not only as a doctor within the Western clinical tradition, but as a natural historian. Many of his neurotic patients had restricted their lives with self-imposed prohibitions almost indistinguishable from the authentic taboos from the South Seas. In his explanation, he offered a new and comprehensive understanding of social custom in psychological terms. His basic assumption is that just as neurotic behavior is the result of a compromise between unconscious erotic and hostile impulses, taboos are expression of ambivalent emotions.

The taboo surrounding Polynesian rulers are cumbersome and deprive the ruler of normal and gratifying social intercourse. In the realm of the conscious sphere of discourse, the taboo is an expression of the highest esteem. But in the hidden agenda of the unconscious mind it is a punishment for their exaltation, a revenge taken on them by their subjects.

The taboo on physical contact with the dead is common in preliterate societies. Freud maintains that "the taboo has grown up on the basis of an ambivalent emotional attitude...the contrast between conscious pain and unconscious satisfaction over the death that has occurred" (Freud, 1950, p. 61).

Although Freud's interpretation is controversial and not generally accepted today, it nevertheless is an expression of the psychoanalytic insight in human cognition which today is part of our cultural system of understanding generally. The key concept is "rationalization," a mental process through which deeper and hidden motives are presented to the world as well as to the conscious ego in an acceptable and morally legitimate fashion.

Freud articulated in his theory of the human soul one of the most important insights from nineteenth century humanism: the human capacity to sustain illusions. Intellectuals and philosophers have known for millennia that fraud is one of the great problems of human existence and have developed logic to aid their fellow men in the search for truth. Freud, however, went beyond the syllogisms and realm of pure cognition into the hidden hinterland at the border of consciousness and beyond.

This thinking was in the spirit of the nineteenth century which in philosophical terms was explored by Nietzsche, particularly in his work *Jenseits von Gut und Bose* (*Beyond good and evil* [1935]). The decline in the belief in the power of reason is heralded by Nietzsche in the following words: "The German soul has corridors and hidden, connected channels, and it has caves, and secret hiding-places and dungeons." This first rumbling of early post-modernism inspired Freud to an attack on the hidden abysses in the human mind by reason, which is the main project of psychoanalysis.

Before Sigmund Freud, Karl Marx had presented his own views on the human capacity for illusions in his thesis on "false consciousness." He applied the term to the realm of political consciousness and maintained that the workers were living in a situation of false conciousness as long as they did not realize that they were oppressed. The false consciousness—with religion, the opium of the people, as the main culprit—was, however, the result of capitalism and therefore not an unavoidable situation.

Freud was of the opposite view and was convinced of the fact that rationalization was part of human cognition everywhere. Ibsen, who also elaborated on the great insight of the nineteenth century, gave the idea a literary articulation in *The wild duck* where he made it clear that human beings could not live happily without the lies of life, the cherished illusion people sustain in everyday life.

Freud's treatment of totemism also set the stage for an investigation of the relationship between parents and children in the human family and it was part of his theory of the developmental stages of early childhood. The mythological core in the exploration of totemism is that of King Oedipus who unknowingly killed his father and later married his mother. Upon discovering this, he blinded himself and went into exile.

Freud had discovered during the analysis of his patients the problems caused by the fact that children's erotic and hostile impulses are frequently directed toward family members. He was particularly interested in the relationship between mother and son and the sometimes explicit wish on the part of a small boy to eliminate the father to have the mother's undivided attention. But since small boys also love their fathers, the classic ambivalent situation emerges with great psychological costs, the psychological distress caused by the oedipus complex.

In presenting his insights on these sensitive issues Freud created his own myth inspired by the prevailing Darwinian idea of the primal horde, the assumed prefamily groups before the dawn of human civilization, dominated by a violent and jealeous father who keeps all the females for himself:

> One day the brothers who had been driven out came together, killed and devoured their father

> and so made an end of the patriarchal horde....The violent primal father had been the feared and envied model of each one of the company of brothers: and in the act of devouring him they accomplished their identification with him ... The totem meal would thus be a repetition and a commemoration of this criminal deed, which was the beginning of social organization, of moral restrictions and of religion....Society was now based on complicity in the common crime; religion was based on the sense of guilt and remorse attaching to it, while morality was based partly on the exigencies of this society (fraternal society) and partly on the penance demanded by the sense of guilt (exogamy and attempts at atonement) (Freud, 1950, p. 141–146)

Although a pioneer in the application of ethnographic knowledge in the understanding of the human soul, Sigmund Freud like James Frazer relied solely on written sources and never did any fieldwork himself. Those were the days before professional anthropological fieldwork brought the social life of distant cultures within the reach of the intellectual community.

1.08.4.4 The Impact of Anthropological Fieldwork

The anthropologist who set the standard for anthropological fieldwork was Bronislaw Malinowski. His works from the South Seas could partly confirm and partly correct some of Freud's assumptions. With regard to repression of sexuality Malinowski could confirm Freud's opinion regarding its psychological costs (Malinowski, 1927, 1932). With regard to the universality of the oedipus complex, however, Malinowski pointed out that among the matrilineal Trobrianders in the South Seas, it was the mother's brother rather than the father who was the source of discipline in the life of a small boy, and therefore the primary object of childish aggression.

Comparative research through modern anthropological fieldwork not only demanded a revision of Freud's theories. Also, ideas of so-called primitive mentality, promoted above all by Levy-Bruhl (1923), which maintained that individuals in preliterate societies think in a manner not only quantitatively, but also qualitatively different from individuals from complex societies, had to be revised. Cognitive anthropologists who have studied the system of knowledge in preliterate societies with the regard to the mastery of their environment, have found rational systems which demonstrate that no generalized primitivity of thought can be observed (Price-Williams, 1980).

Of particular interest in this connection are investigations of how people in complex societies make judgements under uncertainty. As Kahneman, Slovic, and Tversky (1982) have demonstrated, people under certain kinds of stress regress into "primitive mentality" which then should be regarded as a mode of thought depending upon circumstances and not on culture.

Cognitive studies of non-Euro-American cultures, however, may inform our understanding of modes of thought in general as demonstrated by Bateson (1958) in his study of dichotomous thinking and the reversal of opposites in the Naven ritual of New Guinea. Gardner, in his book on structuralism (1981), suggests that cognitive anthropologists like Levi-Strauss may broaden and deepen cognitive psychology in the same way as have linguistics and developmental psychology.

1.08.4.5 Culture and Personality

The influence of anthropological fieldwork was greatly enhanced by the works of Mead who started publishing her studies from the South Seas in 1928. Her works influenced not only her own field of anthropology, but also psychology and above all public opinion. Her books were widely read (Mead, 1935, 1949), and she became a celebrity and a most respected anthropologist, although some of her conclusions were met with some reserve. After her death a virulent criticism of her work and methods has been raised by the Australian anthropologist Freeman (1983) who claimed that she had been misled by her informants and her claim that Samoan adolescence was relatively free of conflict was wrong. After the professional discussion on her work had run its course, the general agreement is that Freeman overstated his case, although some of his criticism had to be accepted.

When Freeman published his work, however, the culture and personality school of thought was already history. But Mead had left a lasting impact both on American anthropology and public opinion.

Her work grew out of the Columbia group under the leadership of Franz Boas. One of Boas' main concerns was the rampant racism both in Europe and in the USA in the first decades of the twentieth century. His main assumption was that it is the culture not the genes which determine human behavior and institutions.

The school of thought which emerged from the group at Columbia University was focused on the configurations of themes found in different cultures and their mode of thought was influenced by Gestalt psychology. The

Gestalt psychologists had pointed out that our perception is self-organizing and that patterns emerge on the basis of certain principles like similarity and proximity. The insight of the Gestalt psychologists that wholes have properties not found in the constituent parts of a pattern were applied to culture and reversed the trend towards atomism in American anthropology.

Distribution of traits, elements of customs, material culture, or kinship were, at the beginning of the twentieth century, used to inform rather speculative evolutionary theories and theories of cultural diffusion. The trait investigations almost disrupted culture systems, turning them into a hodgepodge of traits.

Benedict's *Patterns of culture* (1934) is the most elegant articulation of the configuration approach of the Columbia group which crystallized into the culture and personality episode in American anthropology. This school of thought was strongly influenced by psychoanalysis which is obvious in the following words of Benedict: "Tradition is as neurotic as any patient, its overgrown fear of deviation from its fortuitous standards conforms to all the usual definitions of the psychopathic" (p. 252).

The Columbia group was also inspired by ideas of tolerance, the roots of which can be traced to John Stuart Mill's ideas of freedom. Through her books from her various fieldworks, Mead (1928, 1935, 1949) not only convinced the anthropological community, but people in general, that human beings in different cultures are different, not because of any difference in their genetic make-up, but because the human personality is molded by culture into different configurations of basic human themes. In *Sex and temperament in three primitive societies* (1935), Mead demonstrated that ideas of masculinity and femininity are determined by culture and not by nature, and thus made a considerable effort to rid modern society of the yoke of traditional ideas of social roles. Her influence was not merely scientific, but also ideological. She published her major works in a period when parts of Europe were in the grips of a bigot and sinister racism with roots partly in misinformed and partly misread anthropology and biology.

Through Mead, the didactic role of anthropology came into focus. But this role is not entirely unproblematic. When the center of the anthropological stage is given to great humanists like Mead, the cause of reason has little to fear. But with the potential ideological impact of anthropology in mind, it is essential that the highest standard of method is observed when it comes to questions of values. The culture and personality school suffered a set back in the wake of its strategic use first in World War II and later in the Cold War in the study of national character. The portraits painted by the scholars of the national character of an assortment of other nations were clearly influenced by political concerns and the geopolitical mission of the USA. But even if the problems of culture and personality are not in the focus at the present time, there is an enduring interest in the relationship between the individual person and his culture both in psychology and anthropology.

1.08.4.6 The Paradigm of Ethnic Identity

The problems formerly treated under the headline of culture and personality are today being addressed under the headline of ethnicity. An ethnic group is normally considered to be a subgroup within a larger community or nation. It is of particular relevance in the USA where many different cultures meet within the framework of a judicially defined citizenship. The idea of the big melting pot gave way, particularly during the 1960s, to a resurgence of interest in ethnic identity. Formerly subdued and sometimes harassed groups, who used to be ashamed of their culture, changed their strategy and demanded respect for their traditions and style of life. This was a worldwide phenomenon indicating a shift of paradigm in cultural understanding. The Euro-American culture, which held a hegemonic place in the cosmopolitan scale of values as described by the Italian philosopher Gramsci was challenged by dominated ethnic groups of people all over the world: The native American, the Scandinavian Lapps or Saami, the Maoris of New Zealand joined forces in an ethnic uprising demanding respect for their unique cultures.

This great movement of self-respect focuses on identity and is related to nationalism as it developed in Europe in the nineteenth century with the emergence of new national languages and new nations. A hegemonic culture like the German was challenged by a number of ethnic groups within the framework of the Austrian–Hungarian Empire: the Czech, the Slovaks, the Hungarians, and a mosaic of people in the Balkans struggled to construct their own national identity based primarily on shared beliefs, shared language, and mythological origin. As Anderson (1983) has demonstrated, the feeling of identity and togetherness on the basis of national ideas is based on an imagined community dependent upon modern media of communication.

Many of the ethnic groups in the USA derive from the various European nations who brought their national identities with them across the ocean. For good reason, Barth's work on ethnicity (1969) became the most quoted anthropological text in the 1970s and 1980s.

The psychological effect of identity problems has been explored by Erik Homburger Erikson and very sensitively described in *Childhood and society* (1950). Although Erikson does not present his description and analysis of the Sioux Indians of the Great Plains and other groups under the label of ethnicity, he demonstrates how the Sioux identity, as it had developed through history, was based on their tradition of being mounted warriors and bison hunters. He shows in detail how their identity in Sioux pedagogics is developed in intimate articulation with the stages of development, which he regards as pan-human, to establish an ethnic self-confidence elegantly adapted to their ecological situation. But through military defeat, economic dependency, and other features of reservation life, the self-confidence of the Sioux has been broken. Some of their psychological reactions are reminiscent of the old-prisoner-syndrome of concentration camps with the characteristic Freudian identification with the aggressor.

The field of ethnicity is a particularly promising common ground for a collaboration between anthropologists and clinical psychologists. The relevance of ethnic dimensions in psychotherapy has been explored by Abel and Metraux (1974). The blacks from Brazil, Africa, or the Caribbean represent different cultural configurations and should not be confused with the blacks in the USA.

The ethnic differences are not in principle different from the socioeconomic differences which have come to the attention of scholars of the patients of Freud and Adler. As Ansbacher (1959) points out, three-fourths of Freud's patients were rich, while most of Adler's patients were poor. For good reason Adler coined the term "inferiority complex" and regarded the will to power as the dominant theme of life, whereas Freud maintained that the basic theme was "eros." The clinical picture of their clients was different and can only be understood if the relationship between culture and personality is taken into account.

When cultural differences are of a greater order than among not too distantly related ethnic groups, investigators have recorded *culture bound disorders*. Whether these are specific to ethnocultural settings or variants of disorders found under different schemes of classification in the West is not clear. Kiev (1972) has tried to show the parallels in the Western diagnostic system and classify the Central and South American *susto*, soul loss, and the Chinese *koro*, the fear of the penis or the labia withdrawing into the body, as anxiety states. The Southeast Asian *amok*, the symptoms of which are eruptive violence followed by amnesia and exhaustion is classified as a dissociative state. He classifies these culture bound ailments as a subgroup of Western neurotic disorders. Yap (1969) on the other hand classifies culture bound disorders as reactive psychoses.

The ethnocentric bias of Western psychiatry is a challenge to anthropology as demonstrated by Marsella and White (1982) which hopefully can be remedied through an open-minded exchange of ideas and rigorous cross-cultural research.

1.08.4.7 The Cross-cultural Correlational Approach

A promising venture into cross-cultural research is associated with another group of scholars at Yale University: John Dollard, George Peter Murdock, and John Whiting. Jointly they developed the cross-cultural correlational approach at Yale Institute of Human Relations. We have already taken advantage of Murdock's encyclopedic work on theories of illness. One of the most prominent studies in the cross-cultural correlation approach is that of Whiting and Child (1953). They criticized the studies of national character and the configurations school because of its lack of scientific method and embarked upon an ambitious program of correlation studies. In his summary of Whiting and Child, Bock (1988) selects the following statement of their method:

> In the correlation testing method the supposed antecedent condition is looked for as it occurs or fails to occur in the natural course of events in a number of cases. Instances are collected of its presence and its absence, or of its presence in various degrees. The supposed consequent condition is also looked for in each of these cases and determined to be present or absent in a given degree. It is then possible to determine whether there is a consistent relation between the two, and thus whether the hypothesis is confirmed or negated. Statistical techniques may be applied and the evidence of connection may be based on objective procedures which are repeatable. (p. 108)

The hypotheses to be tested are derived from psychoanalytic theory. Two kinds of customs were to be tested: childtraining practices and beliefs concerning the causes of illness. Whiting and Child identified five so-called systems of behavior based on the Freudian theories of child development: oral, anal, sexual, dependence, and aggression. The children are assumed to develop socialization anxiety in each of these systems depending on the degree of frustration or indulgence they experience. High oral indulgence is noted when a child is nursed whenever it pleases. Severe weaning customs will cause "oral socialization anxiety," and the hypothesis assumes that high oral anxiety will produce theories of illness based on oral behavior: you are sick because of something you said or ate. This hypothesis is mentioned because it produced the most convincing correlation in the whole study.

Whiting and Child (1953) also explored the problem of male initiations which in many societies are extravagantly painful with scarification, circumcision, whipping, and hazing. They assumed that the cruelty of the initiators (i.e., the fathers), had something to do with the sleeping arrangements in the homes and the postpartum sex taboos. They made the assumption that in societies where mother and infant share the same quarters for an extended period of time and at the same time the postpartum taboo prevented the father from having sex with the mother for more than one year, cruel initiation ceremonies were to be expected. Correlation analysis confirmed the hypothesis.

We recognize the Freudian touch of the hypothesis in the element of rationalization in the staging of the ritual where the fathers under false, but unconscious excuses, seize the occasion to give their aggression and even sadism a free rein in the good name of pedagogics.

This Freudian insight is one of the best allies in the protracted struggle to reduce the suffering and abuse in human life because it rips the mask off moralists who exploit occasions to indulge their cruelty in the name of saintliness. The problem with psychoanalytic insight, however, is that it is hard to prove by acceptable and rigorous scientific methods. It is to the great merit of Whiting and Child that they brought their Freudian hypothesis to an irrefutable test.

Although Whiting and Child have been criticized for, among other things, overlooking the dramatization of male solidarity in puberty rituals, the correlation method still holds promise although it no longer occupies center stage as in the 1950s. One of the reasons is the great cost of statistical investigation of such a scale. To this can be added the new look on culture which tends to regard the approach of Whiting and Child, focusing on traits, as old-fashioned.

Culture is a symbolic system which is similar to a language and calls for its own methodology as pointed out by LeVine. In his words of a speech community:

> Their capacity for mutual understanding is accompanied by a remarkable consensus about rules of pronunciation and grammar such that linguists can discover the rules prevailing in the whole community not from an extensive survey, but from a small number of informants. This redundancy of rules in a speech community is theoretically significant in illustrating how language functions as social communication, it is methodologically significant in permitting a mode of systematic inquiry distinct from the sample survey. (LeVine, 1984, p. 68)

1.08.5 SOCIAL SYSTEMS AND THE INDIVIDUAL

1.08.5.1 The Dominant Kinship Dyad

The individual tends to be somewhat distant in the analysis of social systems. A set of hypotheses with regard to the dominant kinship dyad was presented by Hsu (1965). In Hsu's work the individual does not take center stage, but his approach is in tune with the analysis of the social process which was to change the course of anthropology in the late 1950s and 1960s. According to Hsu's hypothesis, human societies everywhere give priority to one kinship dyad from the core dyad of procreation and the intrinsic qualities of this relationship are assumed to influence several cultural parameters. From the core dyad of mother and father other essential dyads can be traced: father–son, mother–daughter, brother–brother, and so on.

Each dyad had intrinsic qualities. The mother–father or husband–wife dyad have the following intrinsic qualities: discontinuity, equality, sexuality. The father–son dyad has the quality of continuity because a son becomes a father himself and through the links between father and son a lineage can be traced. Authority is intrinsic because of the differences in age and command of resources. Even in egalitarian societies, where the fathers try to be buddies with their sons, authority cannot be denied, only under-communicated. In other words, the oedipal situation cannot be avoided. There is hardly an element of sexuality in the relationship. The husband–wife dyad on the other hand is discontinous, one marriage does not beget another, there are no lineages of marriages. The relationship is contractual even if it lasts a lifetime, but it always comes to an end. Since an intimate relationship cannot be sustained if the authority dimension dominates, the relationship is intrinsically egalitarian. Its very *raison d'être* is sexual.

According to Hsu's hypothesis, cultures always give priority to one of the possible dyads, and the intrinsic qualities of the dyad which becomes dominant will be reflected in the cultural system. Societies where the father–son dyad is dominant will respect traditions and authority and sexual matters will not be culturally dominant. If on the other hand the husband–wife dyad dominates, respect for traditions will be less and authority will be under-communicated if not resented. Sexual matters will be culturally dominant.

The effects of the father–son dyad have been more or less thoroughly explored by Hsu (1971) and the evidence is in favor of the hypothesis. In Arab and traditional Chinese culture, the father–son dyad is unquestionably dominant and the respect for tradition and authority is obvious. Likewise sexuality is not a prominent theme in the public sphere of literature, media, and discourse. In the USA on the other hand, the husband–wife dyad is clearly dominant. Although the respect for the constitution is unquestionable, the image of North Americans as progressives and egalitarian is not unfounded. The question of sexuality is a dominant theme both in the private and the public sphere.

Hsu's hypothesis is not uncontroversial. This, however, is the fate of culture theories generally. So far other dominant dyads, like, for example, mother–daughter, have not been explored.

1.08.5.2 The Iatrogenic Approach

In the Durkheimian tradition of anthropology and sociology, the systems have dominated at the cost of the individual in the English speaking world in the first part of the twentieth century. The person has more or less been regarded as a puppet in the strings of the social system. In the 1950s, however, the works of Max Weber reached the centers of anthropological discourse and the person as an actor entered the stage.

At the London School of Economics, Leach radically changed the systemic approach which until then had dominated British social anthropology. In his analysis of political systems (Leach, 1954), he demonstrated how human beings as decision-makers interacted with the social system. That changed the image of small-scaled societies as timeless depositories of unchangeable traditions. This image had been created from lack of information. Professional field descriptions before 1950 were more or less snapshots freezing moments in time.

At the University of Chicago, Erving Goffman demonstrated how the anthropological method of participant observation could be applied to cultures beyond the small-scaled societies. In *The presentation of self in everyday life* (1959), Goffman analysed the strategies of personal behavior with the theater as a metaphor. He demonstrated how human beings played their roles, not in the make-believe way of the stage, but with the same objective: to impress the audience. His thick description of "impression management" codified in language processes which formerly were not part of conscious discourse. He described how social reality is created through role-play, how a select part of a person's role repertory is sustained in a situation through strategic interaction. Life after all is composed of situations, most of them trivial. Goffman made the situation the main object of his study and demonstrated how social reality is created on the spot and thus may be regarded as a kind of illusion. After Goffman it is no longer possible to regard culture as a thing, it is a process sustained by strategic interaction. The insights of Goffman connect with psychoanalysis on the inner stage of consciousness. Human beings also venture to impress themselves, the ego is an actor who creates the person from the raw material of basic drives.

In his study of psychiatric hospitals we recall the observation of Pinel during the French Revolution when he pointed out that the behavior of "lunatics" had something to do with the way they were treated (Foucault, 1965). Goffman demonstrated that many of the symptoms recorded by psychiatrists were samples of strategic interaction. One of the strategies is situational withdrawal. A perfectly comprehensive reaction to an impossible situation is codified by the psychiatrists as regression. Goffman opened up the Pandora's box of iatrogenic symptoms, symptoms created by the medical profession. His insight is valid beyond the psychiatric hospital, it applies to total institutions everywhere and he raised a problem which remains to be explored in depth: how psychological disorders may be imposed on the individual by society and by the medical and psychological professions. If these professions decide that a condition is pathological, it becomes a social reality, and a psychiatric disorder is magically established. This is sometimes an overlooked part of the construction of social reality. Anthropology may through the insights of how social systems and institutions mold the individual, and the person as a category, and vice versa, cut some of the strings which bind the individual to social systems.

1.08.6 REFERENCES

Abel, T. M., & Metraux, R. (1974). *Culture and psychotherapy*. New Haven, CT: College and University Press.

Anderson, B. (1983). *Imagined communities. Reflections on the origin and spread of nationalism.* London: Verso.

Ansbacher, H. L. (1959). The significance of the socio-economic status of the patients of Freud and Adler. *American Journal of Psychotherapy, 13,* 376–382.

Barth, F. (Ed.) (1969). *Ethnic groups and boundaries: The social organization of cultural difference.* London: Allen and Unwin.

Barth, F. (1975). *Ritual and knowledge among the Baktaman of New Guinea.* Oslo, Norway: Universitetsforlaget/ Yale University Press.

Basham, A. L. (1976). The Practice of Medicine in Ancient and Medieval India. In C. Leslie (Ed.), *Asian medical systems: A comparative study.* Berkeley, CA: University of California Press.

Bateson, G. (1958). *Naven* (2nd ed.). Stanford, CA: Stanford University Press.

Benedict, R. (1934). *Patterns of culture.* New York: Houghton Mifflin.

Benedict, R. (1974). *The chrysanthemum and the sword. Patterns of Japanese culture.* New York: New American Library.

Berger, P., & Luckman, T. (1969). *The social construction of reality: A treatise in the sociology of knowledge.* Harmondsworth, UK: Penguin.

Bock, P. K. (1988). *Rethinking psychological anthropology. Continuity and change in the study of human action.* New York: Freeman.

Brogger, J. (1971). *Montevarese: A study of peasant society and culture in southern Italy.* Bergen, Norway: Universitetsforlaget.

Brogger, J. (1986). *Belief and experience among the Sidamo. A case study towards an anthropology of knowledge.* Oslo, Norway: Norwegian University Press.

Brogger, J. (1992). *Nazare. Women and men in a prebureucratic Portuguese fishing village.* Orlando, FL: Harcourt Brace Jovanovich.

Bürgel, J. C. (1976). Secular and religious features of Medieval Arabic medicine. In C. Leslie (Ed.), *Asian medical systems: A comparative study.* Berkeley, CA: University of California Press.

Chadwick, J., & Mann, W. N. (Trans. and Eds.) (1950). *The medical works of Hippocrates: A new translation from the original Greek made especially for English readers.* Oxford, UK: Blackwell.

Douglas, M. (1966). *Purity and danger.* London: Routledge & Kegan Paul.

DuBois, C. (1944). *The people of Alor: A social-psychological study of an east Indian island* (Vol. 1). New York: Harper Torchbooks.

Emerson, V. F. (1972). Can belief systems influence neurophysiology? Some implications of research on meditation. *Newsletter Review, 5.*

Erikson, H. E. (1950). *Childhood and society.* New York: Norton.

Evans-Pritchard, E. E. (1937). *Witchcraft, oracles and magic among the Azande.* Oxford, UK: Clarendon Press.

Foster G. M., & Anderson, B. G. (1978). *Medical Anthropology.* New York: Wiley.

Foucault, M. (1965). *Madness and civilization: A history of insanity in the age of reason.* New York: Pantheon.

Freeman, D. (1983). *Margaret Mead and Samoa: The making and unmaking of an anthropological myth.* Cambridge, MA: Harvard University Press.

Freud, S. (1950). *Totem and taboo.* New York: Norton. (First German ed. 1913).

Gardner, H. (1981). *The quest for the mind: Piaget, Levi-Strauss, and the structuralist movement.* Chicago: University of Chicago Press.

Gebhard, B. (1976). The interrelationship of scientific and folk medicine in the United States of America since 1850. In W. D. Hand (Ed.), *American folk medicine: A symposium.* Berkeley, CA: University of California Press.

Goffman, E. (1959). *The presentation of self in everyday life.* New York: Doubleday.

Goffman, E. (1961). *Asylums: Essays on the social situation of mental patients and other inmates.* New York: Doubleday.

Hsu, F. L. K. (1965). The effect of dominant kinship relationship and non-kin behavior. *American Anthropologist, 67,* 638–661.

Hsu, F. L. K. (1971). A Hypothesis on kinship and culture. In F. L. K. Hsu (Ed.), *Kinship and culture* (pp. 3–30). Chicago: Aldine.

James, W. (1914). *The varieties of religious experience: A study in human nature.* New York: Longmans, Green. (Original work published 1902).

Kahneman, D., Slovic, P., & Tversky, A. (Eds.) (1982). *Judgement under uncertainty: Heuristics and biases.* Cambridge, MA: Cambridge University Press

Kiev, A. (1972). *Transcultural psychiatry.* New York: Free Press.

Kretschmer, E. (1925). *Physique and character: An investigation of the nature of constitution and of the theory of temperament.* London: Paul, Trench, & Trubner.

Leach, E. R. (1954). *Political systems of Highland Burma.* New York: Humanities Press.

LeVine, R. (1984). Properties of culture: An ethnographic view. In R. A. LeVine & R. A. Shweder (Eds.), *Culture theory. Essays on mind, self, and emotion.* Cambridge, MA: Cambridge University Press.

Levy-Bruhl, L. (1923). *Primitive mentality.* London: Allen & Unwin.

Lewis, I. M. (1971). *Ecstatic religion: An anthropological study of spirit possession and Shamanism.* Harmondsworth, UK: Penguin.

Lopez-Austin, A. (1974). Sahagun's work and the medicine of the ancient Nahuas: Possibilities for study. In M. S. Edmonson (Ed.), *In sixteenth century Mexico: The work of Sahagun.* Albuquerque, NM: University of New Mexico Press.

Malinowski, B. (1927). *Sex and repression in savage society.* New York: Harcourt Brace.

Malinowski, B. (1932). *The sexual life of savages in northwestern Melanesia.* London: Routledge & Kegan Paul.

Marsella, A. J., & White, G. (Eds.) (1982). *Cultural conceptions of mental health and therapy.* Boston: Reidel.

Mead, M. (1928). *Coming of age in Samoa.* New York: Morrow.

Mead, M. (1935). *Sex and temperament in three primitive societies.* New York: Morrow.

Mead, M. (1949). *Male and female: A study of the sexes in a changing world.* New York: Morrow.

Murdock. G. P. (1980). *Theories of illness, A world survey.* Pittsburgh, PA: University of Pittsburgh Press.

Nietzsche, F. (1935). *Beyond good and evil* Chicago: Regenery (German ed. 1887).

Pickard, M. E. & Buley, R. C. (1945). *The midwest pioneer: His ills, cures, and doctors.* Crawfordsville, IN: Banta.

Porkert, M. (1976). The intellectual and social impulses behind the evolution of traditional Chinese medicine. In C. Leslie (Ed.), *Asian medical systems: A comparative study.* Berkeley, CA. University of California Press.

Price-Williams, D. R. (1980). Anthropological approaches to cognition and their relevance to psychology. In H. C. Triandis & W. Lonner (Eds.), *Handbook of cross-cultural psychology, Vol 3: Basic processes.* Boston: Allyn & Bacon.

Rosenhan, D. L., & Seligman, M. E. P. (1995). *Abnormal psychology.* (3rd ed.). New York: Norton.

Sheldon, W. H., & Stevens, S. S. (1942). *The varieties of* *human temperament: A psychology of constitional differences*. New York: Harper.

Shweder, R. A., & LeVine, R. A. (Eds.) (1984). *Culture theory. Essays on mind self, and emotion*. Cambridge, MA: Cambridge University Press.

Snow, L. F. (1973). I was borne just exactly with the gift: An interview with a Voodoo practitioner. *Journal of American Folklore, 86,* 272–281.

Stigler, J. W., Shweder, R. A., & Herdt, G. (Eds.) (1990). *Cultural psychology. Essays on comparative human development*. Cambridge, UK: Cambridge University Press.

Strathern, A. (1995). Trance and the theory of healing. Sociogenic and psychogenic components of consciousness. In A. P. Cohen & N. Rappaport (Eds.), *Questions of consciousness*. London: Routledge.

Trevor-Roper, H. R. (1970). The European witch-craze. In M. Marwick (Ed.), *Witchcraft and society*. Harmondsworth, UK: Penguin.

Ullmann, M. (1970). Die medizin in Islam. *Handbuch der Orientalistik*. Erste Abteilung, Ergäzungsband VI, Erster Abschnitt. Leiden, The Netherlands: Bill.

Veith, I. (1972). *The yellow emperor's classic of internal medicine*. Berkeley, CA: University of Callifornia Press.

Walsh, R. N., & Vaughan, F. E. (Eds.) (1980). *Beyond ego: Transpersonal dimensions in psychology*. Los Angeles: Tarcher.

Whiting, J. W. M., & Child, I. L. (1953). *Child training and personality: A cross-cultural study*. New Haven, CT: Yale University Press.

Yap, P. M. (1969). The culture-bound reactive syndrome. In W. Candill & T. Lin (Eds.), *Mental health research in Asia and the Pacific*. Honolulu, HI: University Press of Hawaii.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.09

Life-span Developmental Theories

GISELA LABOUVIE-VIEF
Wayne State University, Detroit, MI, USA

and

MANFRED DIEHL
Colorado State University, Colorado Springs, CO, USA

1.09.1 INTRODUCTION 262

1.09.2 ISSUES IN LIFE-SPAN DEVELOPMENT 263
  1.09.2.1 The Postmodern Turn 263
  1.09.2.2 Development in the Postmodern Era 263
    1.09.2.2.1 The "end point" of development 263
    1.09.2.2.2 The role of context 264
    1.09.2.2.3 Individualism vs. relatedness 264
    1.09.2.2.4 Coherence of the individual 265
    1.09.2.2.5 Modes of thought 265
    1.09.2.2.6 Growth vs. decline 265
    1.09.2.2.7 Stability vs. plasticity 266

1.09.3 COGNITIVE DEVELOPMENT 266
  1.09.3.1 Psychometric Intelligence 266
    1.09.3.1.1 Differential decline 267
    1.09.3.1.2 The role of context 267
    1.09.3.1.3 Plasticity of intelligence 269
  1.09.3.2 Everyday Intelligence 270
  1.09.3.3 Cognitive-developmental Theories 271
    1.09.3.3.1 Piaget's legacy 271
    1.09.3.3.2 Mature reasoning structures 272
    1.09.3.3.3 Cognitive development and the self 273
    1.09.3.3.4 Emotional development 275
    1.09.3.3.5 Wisdom 277

1.09.4 PERSONALITY DEVELOPMENT 277
  1.09.4.1 Developmental Level Approaches 279
    1.09.4.1.1 Freud's legacy 279
    1.09.4.1.2 Beyond the ego 279
    1.09.4.1.3 Ego and culture 280
    1.09.4.1.4 The duality of the self 281
  1.09.4.2 The Trait Approach 282
    1.09.4.2.1 Traits as dimensions of stability 282
    1.09.4.2.2 Context as moderator of stability 284
  1.09.4.3 Empirical Research on Self and Personality 284
    1.09.4.3.1 Stability and change 284
    1.09.4.3.2 Midlife crisis 285

1.09.5 CONTEXT IN PERSONALITY DEVELOPMENT 286
1.09.5.1 Gender, Coping, and Personality 289
1.09.6 CONCLUSIONS 290
1.09.7 REFERENCES 290

1.09.1 INTRODUCTION

In her book *Culture and commitment*, Margaret Mead (1970) proposed that modern life has profoundly changed our experience of the life course. In relatively simple societies, Mead suggests, the life structure for individuals was clear and predictable. Children could look at the lives of their grandparents, and take those lives as models for their own future. In such a society, "change is so slow and imperceptible that grandparents, holding newborn grandchildren in their arms, cannot conceive of any other future for the children than their own past lives. The past of the adults is the future of each new generation: their lives provide the ground plan" (pp. 13–14), a plan often believed to be supported by divine and mythic processes (Labouvie-Vief, 1994).

In contrast, suggests Mead, contemporary society involves profound differences in how the life course is experienced. An example of such a society is the USA during and after the beginning of the Industrial Revolution, when streams of immigrants sought new opportunities on this continent. Parents no longer were the experts in how to live, and children were as likely to learn from their peers as their parents. Thus, the past no longer was a guide to the future, but rather the future became shaped by the present. Learning and developing are no longer so one-sided, but elders are likely to learn from the young, just as the young learn from the old.

Along with Mead, many authors (e.g., Hareven, 1978) have suggested that social changes such as the Industrial Revolution and the information revolution have altered the way we experience life. No longer do we live in communities or countries that are relatively isolated, and homogeneous, but we are becoming citizens of a global community constituted by individuals from diverse cultural, ethnic, and racial backgrounds. Further, technological changes have produced increases in the life-span as a result of the elimination of early mortality. This "greying" of the population has dramatically increased the number of vital individuals of advanced age.

All of these unprecedented historical changes have changed the way we look at development through life. This is true in several basic ways. First, in the past, emphasis was on socializing individuals into stable adult roles defined by cultures that were relatively homogeneous. Thus many theories described adulthood in terms of stability, while childhood was described in terms of change and transformation (e.g., Baltes, 1987; Baltes, Lindenberger, & Staudinger, 1998; Labouvie-Vief & Chandler, 1978). However, contemporary life highlights change and transformation along the adult portion of the life-span as well. As a result, there has been a wide-ranging re-evaluation of what constitutes "maturity," as well as the pathways by which maturity is reached. In fact, this new emphasis on continued growth in adulthood could be referred to as the "birth of adulthood," borrowing from Philippe Aries' (1962) famous statement that earlier historical periods brought a similar "birth of childhood" as a distinct developmental period.

A second consequence of historical changes has been that many cultures no longer are based on groups of individuals homogeneous in racial and ethnic features, but rather represent a "melting pot" of groups characterized by diversity (see Cavanaugh, 1997). This focus on diversity also has affected the way we view adulthood. While in the past, adulthood often was defined by a single outcome (such as a strong ego or abstract reasoning), now we are realizing that outcomes can vary widely depending upon the context in which individuals grow up. Thus, not only has development been extended into adulthood, but the pathways of development also have taken on more diversity. This "multilinearity" (Baltes, 1987; Fisher & Biddell, 1992; Labouvie-Vief & Chandler, 1978) is another feature important in recent writings about the life-span.

Finally, all of these historical changes also have had profound implications for the ways in which we look at the process of aging. Since aging often is defined relative to some theoretical apex of optimal functioning, redefining that apex has direct implications for views of aging. Often defined as a regressive deviation from some youthful standard of optimal functioning (e.g., Labouvie-Vief, 1985), aging is now understood as a period that has its own developmental challenges and opportunities, and that needs to be understood within the context of the total life course (see Erikson, 1982; Labouvie-Vief, 1980; Staudinger, Marsiske, & Baltes, 1995).

In this chapter, we highlight these recent re-evaluations in theoretical outlook. First, we will briefly highlight some of the general theoretical changes that have shaped current views of development in general, and of life-span development and aging in particular. Second, we will turn to a summary of how these changes are expressed in the evolution about thinking about intellectual functioning, while in the following section we will discuss similar changes in the domain of self and personality. We will discuss several theoretical movements that represent blends of these areas that have traditionally been isolated.

1.09.2 ISSUES IN LIFE-SPAN DEVELOPMENT

Although the main concern in this chapter is with psychological theories of life-span changes, it is useful to place these theories within a broader context of intellectual–historical changes. Concepts of human nature and the meaning of mature adaptive competence are embedded into broad historical patterns that cut across many disciplines, from philosophy through literature to art and religion. All of these domains have seen pervasive changes in how the ideal human is defined, for example, what blend of reason or faith, emotion and intuition, of independence or interdependence, of "masculinity" or "femininity," and of similar polarities should characterize that ideal individual that crowns our theories of development, and what standards, as a result, should guide how to measure the resulting characteristics. In this section, we will outline some of these changes. In doing so, we will highlight their pervasive and general nature and merely anticipate the themes they set for future psychological questions.

1.09.2.1 The Postmodern Turn

If one were to name a single characterization that Western intellectual tradition has held up as applying to the kind of ideal individual we envision when we think of advanced levels of development, probably it would be that of rational thought and conduct. Yet, while the basic definition of the nature of rationality has remained more or less constant for many centuries, the last hundred years or so have witnessed a wide-ranging transformation of the concepts of "reason" and "rationality." Many disciplines (from philosophy, mathematics, and physics to art, literature, and religion) have participated in criticizing and deconstructing the traditional concept of reason. Following Plato (see Labouvie-Vief, 1994), the essential features of reason were understood as its independence of the senses, the emotions, and other organismic processes. Reason was thought to transcend variations in context and to refer to abstract and conscious decision-making capacities located inside the individual mind or brain. It described our mental natures rather than our participation in a bodily, material universe, and it referred to an active "grasping" rather than a receptivity. In short, what was emphasized was reason's objective nature as opposed to its subjectivity.

The rationalist model has stood as an ideal of maturity throughout the ages, and has served as the basis for the most important early theories of development, including those of Freud and Piaget (Labouvie-Vief, 1994). However, as the twentieth century began to witness broad global transformations in the nature of health, information, and community—the kind of changes, indeed, that Mead refers to—dissatisfaction with the old models of mind and development has arisen in many disciplines. As this faith in the power of logos is being criticized, many contemporary views of the mind have begun to return to the dimensions of mind and self that were abandoned by the classical view. Emotion and intuition, time and process, variability and the role of context, all of these elements that previously were devalued have come to the forefront in discussions. This contemporary confrontation with the other side of rationality is often referred to as "postmodern" attitude (Gergen, 1991; Lyotard, 1987). The postmodern attitude asks that we transform the dogma of the inseparability of reason from other, traditionally less valued functions of thought. Thus, a new view is emerging that holds that reason has inherent elements that are located in the domain traditionally relegated to the subjective: of the discourses unique to a culture, of conventions that have ritualized what we imagine of the rhetorical and practical, of the mythic and symbolic, and of the organic and emotional (see Labouvie-Vief, 1994).

1.09.2.2 Development in the Postmodern Era

As far as psychological conceptions of individuals and their development are concerned, the postmodern revolution has had several effects on conceptions of the life course.

1.09.2.2.1 The "end point" of development

A first profound change has been in the nature of what we call mature, ideal functioning. In traditional theories, such as those of Freud (e.g., 1925/1963) and Piaget (e.g., Inhelder & Piaget, 1975), the end point of development usually coincides with early adolescence and the achievement of sexual and physical maturity. Along with this early apex, theories stressed such attributes as self-control and the ability to adopt cultural roles. However, in more recent theories, the notions of biological and psychologic maturity or generativity have been specifically decoupled; hence concepts of maturity, following Mead's (1970) suggestion, extend beyond the faithful, stability-guaranteeing adherence to societal norms and emphasize continued capacity for critique, openness, and transformation of individuals and society. As Jung (1933) stated, the emphasis becomes less on socialization and more on individuation. In general, we value the individual's attempts at meaning making, while valuing less the propensity to adopt the meanings that are given by conventional social roles and settings.

A good example of this new emphasis is given by Baumeister's (1986) discussions on the nature of the self. In medieval times, there was little emphasis on inner self and its struggles. For example, there was almost no autobiographical writing. Interest in portraying life stories rather was in exemplary and ideal life, as in the stories surrounding saints. Similarly, in painting, the focus was not on the portraits of ordinary individuals. Rather, the emphasis was on biblical and religious topics and images, which were depicted in a very formal and idealized way, such as icons. Nor did the individual author or artist matter: for example, early medieval icons or manuscripts were not signed; the role of artists was not to impress their own individuality on the work, but rather to provide a faithful copy of an original. In the sixteenth century, in contrast, a focus on the inner self emerged. The notion became popular that the "self" is a hidden, inner space that is different from one's outer appearance, and writers and artists' now attempted to characterize this new inner self.

1.09.2.2.2 The role of context

One of the consequences of the new model of persons was the vindication of the role of context (see Labouvie-Vief & Chandler, 1978). Thus, rather than thinking of individuals as mere replications of the same universal law, individuality and the contextual features that give shape to it became significant as essential features of developmental theories. A new theoretical emphasis on context was influenced by trends early in the twentieth century, when there was a flurry of philosophical and scientific writings that dealt a profound blow to the notion that what is right and true could forever be defined in an objective, universal, timeless way (see Labouvie-Vief, 1994, for discussion). Instead, many thinkers began to propose that truth (at least as far as it concerns relatively complex states of affairs) is not a static and universal thing, but that it is best understood as a dynamic process, something that evolves, that is being constructed, and that is being continually redefined. Truth has, therefore, an explicit historical and developmental perspective, and is moved out of the realm of axioms and ideals into the very context of human practice (see Labouvie-Vief, 1994).

These contextualist assumptions have modified recent theoretical attitudes regarding psychological development and aging, as well as the fulcrum that is to balance the two, namely maturity. Early theories tended to adopt either of two dualistic views: organicism (see Reese & Overton, 1970), with its emphasis on transcultural and transhistoric processes, or mechanism with views that focus on environmental contingencies that shape behavioral repertories. In contrast, many current views combine assumptions of these two models. They assume that, on the one hand, sequential and fairly pervasive processes of growth and decline are a part of life-span changes; yet on the other hand, these changes are modified by historical or cultural contexts, as well as by specific individual experiences. Indeed, a widely held view in the late 1990s is that basic developmental pathways may be shared by many groups of individuals; yet at the same time, the specific up- and downturns or even the specific directions such pathways may take for any one individual, are based on the experiences they encounter along the journey of their development. Thus from a core pattern, one may envision different pathways that partially divide in fan-shaped fashion (see Baltes, 1987; Bronfenbrenner, 1979).

1.09.2.2.3 Individualism vs. relatedness

With the rediscovery of context also came the realization that the mind is not just an intrapersonal entity of separate selves. Although awareness of self and mind as relatively separate is essential for functioning in a complex society (see Elias, 1982; Labouvie-Vief, 1994), it is also true that all individual minds or selves participate in the interrelationships that constitute a collective. Thus, classical models of mind and self have idealized the individual in terms of such characteristics and values as individuality, autonomy, independence, achievement motivation, and identity; in contrast, such values as relatedness, compassion, and dependence were seen as less important and indeed as less mature (see Bakan, 1966; Gilligan, 1982; Guisinger & Blatt, 1994).

In contrast, many theories have begun to account for the profound ways in which humans are interrelated (e.g., Bowlby, 1969, 1988; Erikson, 1982). Indeed, such interrelatedness must constitute a basic, biologically based propensity in a species that is neotenous and has a long period of extrauterine development. Failure to be in relation to others, in turn, is associated not only with a host of psychological disorders, but also biological disorders such as depressed immunocompetence (Bakan, 1966; Weiss, 1987). As a consequence, a number of more recent models of development have emphasized the need for a balance between autonomy and dependence of relatedness as attributes of mature adults (e.g., Erikson, 1982; Gilligan, 1982; Guisinger & Blatt, 1994; Kohut, 1977).

1.09.2.2.4 Coherence of the individual

The roles of context and relation also have supported how we look at the single individual, not just the individual in relationship to other individuals. An assumption often made in the past—in fact explicit in Plato's theory of mind—was that characteristics emanate from some inner structures such as an ego or some other central processor that gives coherence and unity to self and experience. However, many recent theories also emphasize that different, and sometimes vastly contradictory, experiences and attributes may be activated by different contexts or relationships. Thus individuals are no longer seen as having a single coherent mind, self, or personality, but are seen to combine in themselves contradictions and conflicts in ways that define the unique individuality of a person. Thus concepts of the mind or self are becoming more "modular," with specific modules activated in different contexts and relationships (e.g., Ashmore & Ogilvie, 1992).

1.09.2.2.5 Modes of thought

Another aspect of the mind that been deconstructed is the notion that the mind can be described by a single mode of thinking. Traditional theories of development have placed primary emphasis on rational thinking and conscious decision-making. In turn, such functions of thought as are expressed in artistic, literary, or religious activities were considered to form a less valued aspect. Indeed, they were considered to be less mature or more childlike (see the discussions on Piaget and Freud below).

Labouvie-Vief (1994) recently has addressed this issue. In her view, rationality primarily has been associated with progressive trends in development, while emotions, intuition, and symbolic thinking were identified with regressive trends. In actuality, however, rational views of development are usually undergirded by a consistent language of images and symbols. For example, we typically think of rational functions as masculine while those related to the emotions are stereotyped as feminine. Similarly, we tend to associate the rational with light, airy heights, and golden hues, while associating the nonrational with darkness, earthy and watery depths, and silvery or even black shades. Such symbolism is pervasive in art, literature, and religion, but it also permeates the theoretical discourse about development. Thus an emerging way to think about the mind is to propose the parallel development of two functions. Accordingly, many theorists (e.g., Bruner, 1986; Epstein, 1990; Labouvie-Vief, 1994) now suggest that more narrative, organismic, and experiential modes of thinking form modes of thought that can develop in parallel to rational thought; these modes, then, can show progressive features and occur even at very advances levels of complexity.

1.09.2.2.6 Growth vs. decline

A further shift that has occurred in our views of human nature is a deconstruction of the notion of perfection that was basic to past views of rationality. As Gergen (1991; see also Sampson, 1977) suggested, the traditional view was that rationality assured progress and success, that it ultimately had as its goal perfection. This belief in the perfectibility of human nature also has guided many past theories of development. For example, in Piaget's theory, development reaches its apex as individuals understand logical laws; thereby, Piaget believed, ideas of necessity and perfection could come to guide behavioral regulation.

This "psychology of more" has posed a conceptual problem for researchers in the field of adulthood: if youth is construed as perfection, it is difficult to reconcile the notion of adulthood growth with some of the obvious declines that are a part of adult life. Yet, as suggested by Labouvie-Vief (1980, 1981, 1982a, 1982b; see also Baltes et al., 1998), losses are not unique to later life. Indeed, the whole process of development is interwoven with losses, and often gains in development are related to losses in a direct gain–loss relationship. Thus often what is called decline or regressions can result from the operation of developmental processes that are usually related to progressive development, or growth.

Several principles demonstrate this trade-off between growth and decline. For example, as development proceeds, old structures often become displaced or function with less efficiency. Thus it is well demonstrated in the child development literature (for review, see Labouvie-Vief & Schell, 1982) that as individuals move on to more complex levels that are based on more abstract rules and principles, lower-order performance actually becomes less efficient. Another principle is that of selective depletion, a principle widely important in evolutionary processes. Much as evolution is based on an oversupply of organisms out of which only the "fittest" are selected, so a similar mechanism operates in developing systems. For example, the progressive maturation of the brain during the first stages of life actually involves profound losses as well. Originally, neurons are supplied in abundance, but as workable networks are established, many neurons are "weeded out" through a process of selective death. Finally, development typically involves increasing degrees of specialization, or the trading off of potential for realized structure. Thus, smooth functioning at the system or collective level can be traded off for specialization at the individual level.

In Labouvie-Vief's work, these trade-off processes are proactive and progressive. In addition, Baltes (1987) has added to these progressive trade-off processes reactive ones that may reflect processes of aging proper. Thus, as in the process of growing older, psychological and social resources decline, individuals actually may evolve highly adaptive and resilient compensatory strategies. In addition, individuals who experience a loss of capacity may compensate by selecting one area and focusing an increased amount of attention and energy on it. Such selective optimization at times is associated with extraordinary and highly evolved skills.

1.09.2.2.7 Stability vs. plasticity

The above discussion further highlights that even in later stages of development, a considerable capacity exists for positive change. Earlier stability models, such as Freud's psychoanalytic theory, assumed that individuals were fairly fixed in later life, and indeed that interventions such as psychotherapy were not to be recommended for adults older than about 40 years of age. In contrast, following the lead of Bloom (1964), more recent positions criticized this stability position. Instead, there now is a tendency to explore the potentials for change and plasticity even in late life.

In sum, the twentieth century has been witness to a dramatic revision in notions of development and concepts of what constitutes advanced development. Where original concepts often emphasized a tightly regulated developmental course that primarily reflected biological regulation, that was culturally and historically universal, and that had as its goal a coherent and conscious individual, in the late 1990s views are highlighting characteristics that appear more appropriate to an older and worldwide community.

1.09.3 COGNITIVE DEVELOPMENT

It is proper to start a discussion of theories of life-span development with a review of cognitive development. As already stated, early theoretical views of development were very much influenced by a prevalent rationalism, and academic psychology definitely saw as its primary task to expand on and refine early notions of the rational person, which stood as a prototype of successful, adaptive development. In fact, as Riegel (1976) noted, until the middle of the century, most life-span discussions centered around issues of cognitive development which virtually dominated the literature.

Following historical trends, we begin this section with psychometric research on intelligence, which served as a sample case in which many questions about life-span development were raised for the first time. Thus, psychometric research on intellectual functions formed the arena in which questions about life course development were raised and methodologically elaborated. This section is followed by discussing the more recent influence of cognitive-developmental tradition, which has added to these earlier discussions new ones about the structure and function of intelligence and cognition—concepts which had been developed in the context of predicting educational outcomes in the young (see Baltes & Labouvie, 1973; Labouvie-Vief, 1985)—at later stages of the life-span.

1.09.3.1 Psychometric Intelligence

As Gardner and Clark (1992) have described it, the "psychometric approach to intellectual development seeks to define and quantify dimensions of intelligence, primarily through the collection of individual differences data and through the construction of reliable and valid measurement scales" (p. 16). Thus, researchers who adopt the psychometric approach address questions such as: How can intellectual development be quantified? How can such quantifications be used to predict achievements later in life? How can the intelligence of individuals be meaningfully compared? What factors make up intelligence, and do the factors change with age? (Siegler & Richards, 1982).

Early research on the life course of intelligence was strongly guided by the stability assumptions of the existing theoretical frameworks, which considered intelligence usually as a basic biological endowment with a strong genetic base. In fact, Riegel (1976) noted that early research on intelligence neglected contextual differences, and cultural background was not even believed to affect levels of intelligence. In contrast, debates in intelligence research in the late 1990s focus primarily on issues related to the contributions of genetic/hereditary and sociocultural/contextual influences to individuals' intellectual performance, on the different developmental trajectories of intellectual abilities across the life-span, and on secular trends in intellectual performance (Flynn, 1984, 1987; Schaie, 1994, 1996). In particular, research on the development of intellectual abilities in adulthood and old age has shown different patterns of aging for different mental abilities and has documented much more intra-individual plasticity, and interindividual and intercohort differences than was previously expected (Baltes, 1987; Labouvie-Vief, 1985; Schaie, 1996). Thus, in the following we will review the major findings with regard to the development of intellectual abilities across the life-span, placing particular emphasis on intelligence development during the adult years.

Perhaps the most interesting research in terms of psychometric intelligence since the late 1950s has been conducted with adults of different ages, plotting the developmental trajectories of different mental abilities across the adult life-span and into old age. This research has greatly enriched our knowledge about intellectual development across the life-span and has mostly revised the decrement model of intellectual aging that persisted in the 1940s, 1950s, and 1960s (Labouvie-Vief, 1985). Although early research suggested dramatic "declines" in intelligence beginning as early as 20 or 30 years of age, more recent findings with regard to intellectual functioning demonstrate that these age differences are confounded with cultural-historical change, and that genuine normative declines set in much later, probably in the sixth life decade. In addition, researchers now emphasize the potential of gains and continued growth along with losses (Baltes, 1987; Labouvie-Vief, 1977) as well as the capacity for plasticity during all phases of adulthood (Baltes & Baltes, 1990; Dixon & Bäckman, 1995). In particular, three developments have greatly influenced the research on intellectual functioning in adulthood.

1.09.3.1.1 Differential decline

First, the adoption of Horn and Cattell's (1966) theory of *fluid* (G_f) and *crystallized* (G_c) *intelligence* as an organizing theoretical framework alerted researchers to the possibility of ability-specific age trajectories and the multidirectionality of intellectual development during the adult years (Baltes, 1987). Horn and Cattell's (1966) proposition that abilities which primarily reflect a person's knowledge inculcated through the process of acculturation (i.e., crystallized intelligence, G_c) continue to increase into the sixties and seventies, whereas abstract reasoning abilities (i.e., fluid intelligence, G_f) start to decline in middle adulthood, is now supported by a large body of research (Horn & Hofer, 1992). Horn and Hofer's (1992) explanation that the observed declines in G_f seem to be primarily related to the loss of ability to maintain close attention and to divide attention is consistent with findings obtained from studies conducted within the information-processing paradigm. Numerous experimental studies have shown that changes in abstract reasoning abilities show strong associations with changes in working memory and reduced speed of processing (Salthouse, 1991). As several researchers have suggested, these age-related changes in working memory and speed of processing may reflect neurological changes in the aging person and may explain that even under most optimal training conditions, older adults do not outperform younger adults on complex cognitive tasks such as a mnemonic skill (Baltes & Kliegl, 1992; Kliegl, Smith, & Baltes, 1989).

1.09.3.1.2 The role of context

The second development that has contributed to a revised view on intellectual development during adulthood is related to findings from a number of longitudinal studies of aging (see Schaie, 1983). Such studies tend to suggest much less dramatic declines than cross-sectional research had indicated. In particular, the work of Schaie (1996), working with the Seattle Longitudinal Study (SLS), has shown that these patterns point to very systematic and theoretically important differences between cross-sectional and longitudinal research.

The SLS is unique because it was designed as a cohort-sequential study (Schaie, 1965, 1986; see Figure 1). This research design consists of a coordinated series of longitudinal and cross-sectional studies and permits the joint study of individual development and the effects of cultural-historical change. Thus, this design overcomes the shortcomings of the cross-sectional and longitudinal methods traditionally used in developmental psychology and allows generalizations about the nature and dynamics of age-related change, including the variability created by sociocultural conditions. Moreover, the SLS has followed its aging participants over a 35-year period, thus providing the most comprehensive data set on intellectual development in adulthood (see Schaie, 1994, 1996).

What are the major findings from the SLS? Although there is a plethora of findings from the SLS (for a comprehensive review, see Schaie, 1996), only the major findings will be highlighted here. First, Schaie and his colleagues found that there is no uniform pattern of age-related changes in adulthood across all intellectual abilities. Cross-sectional analyses, for example, show that some primary abilities (i.e., spatial orientation, inductive reasoning) peak in young adulthood, whereas others peak in midlife (i.e., verbal meaning and number; see Figure 2). In contrast, longitudinal analyses showed at least modest gains for all abilities from young adulthood to early middle age (see Figure 3).

Second, longitudinal analyses also showed that, except for word fluency, average age decrements in psychometric abilities cannot be reliably documented prior to age 60. However, reliable average decrement was indeed found for all abilities by age 67 (Schaie, 1994). Third, results from the SLS have also documented substantial generational (cohort) differences in psychometric abilities, that is, different generations show different patterns of intellectual aging. For example, more recent cohorts have been shown to perform at a higher level on the abilities of verbal meaning, inductive reasoning, and spatial orientation, whereas older cohorts outperform younger cohorts on number and word fluency.

Similar patterns of generational differences in intellectual functioning have also been found in other Western countries and have been termed the "Flynn effect" (Flynn, 1987). The Flynn effect, in essence, refers to the well-documented rise in intelligence test performance in subsequent generations. This steady increase seems to reflect the combined effect of a variety of environmental influences, such as increased complexity of life, improvements in education, health care, and nutrition and other factors more difficult to pinpoint (Flynn, 1987).

The SLS also provides a great amount of information with regard to contextual antecedents of individual differences in age-related change, thus emphasizing the sociocultural/contextual dependence of patterns of intellectual aging. Schaie and his colleagues, for example, showed that the absence of cardiovascular and other chronic diseases (Hertzog, Schaie, & Gribbin, 1978), the presence of favorable environmental circumstances, the

Study Waves					
1956	**1963**	**1970**	**1977**	**1984**	**1991**
S_1T_1 (*N* = 500)	S_1T_2 (*N* = 303)	S_1T_3 (*N* = 162)	S_1T_4 (*N* = 130)	S_1T_5 (*N* = 92)	S_1T_6 (*N* = 71)
	S_2T_2 (*N* = 997)	S_2T_3 (*N* = 420)	S_2T_4 (*N* = 337)	S_2T_5 (*N* = 204)	S_2T_6 (*N* = 161)
		S_3T_3 (*N* = 705)	S_3T_4 (*N* = 340)	S_3T_5 (*N* = 225)	S_3T_6 (*N* = 175)
			S_4T_4 (*N* = 612)	S_4T_5 (*N* = 294)	S_4T_6 (*N* = 201)
				S_5T_5 (*N* = 628)	S_5T_6 (*N* = 428)
					S_6T_6 (*N* = 690)

S = Sample; T = Time of measurement

Figure 1 Design of the Seattle Longitudinal Study. "The course of adult intellectual development," by K. W. Schaie, 1994, *American Psychologist, 49,* p. 305. Copyright 1994 by the American Psychological Association. Reprinted with permission.

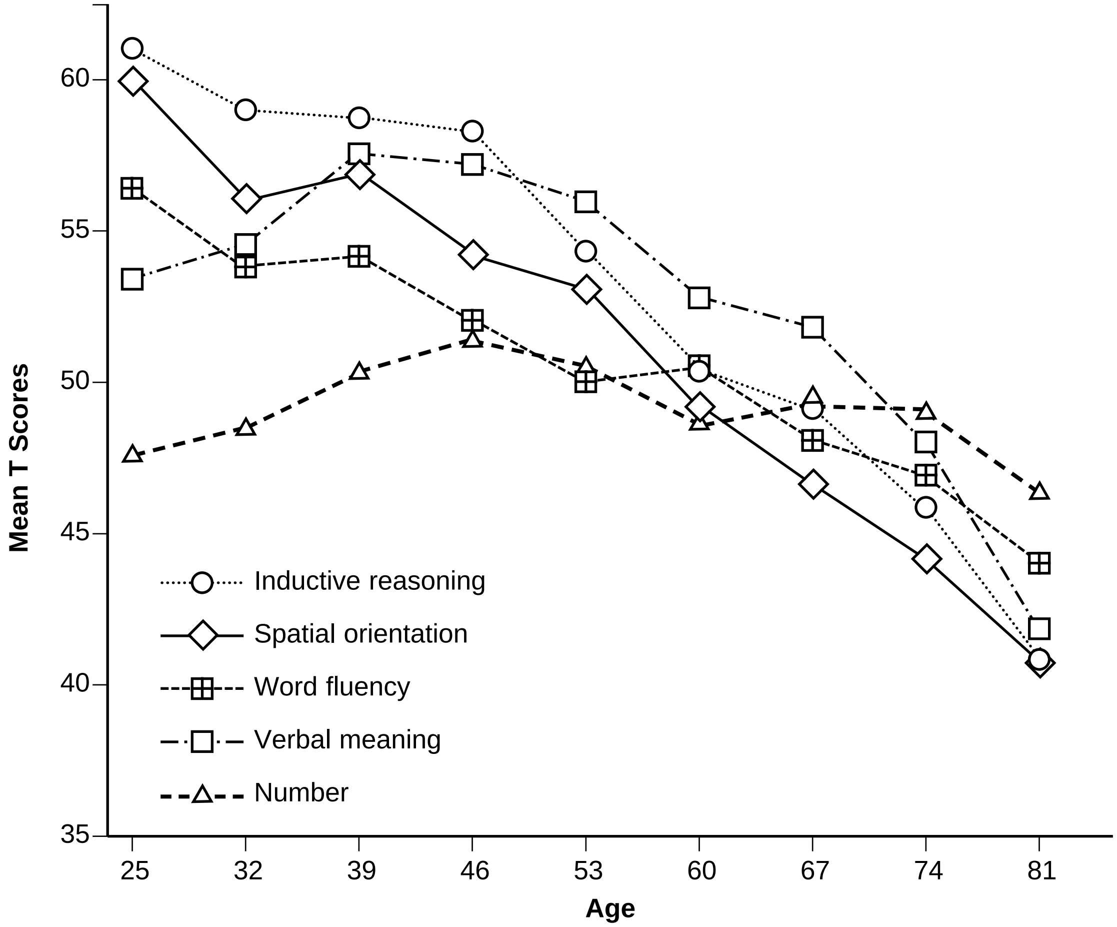

Figure 2 Cross-sectional mean T score for single markers of the primary mental abilities (1991 data). "The course of adult intellectual development," by K. W. Schaie, 1994, *American Psychologist, 49,* p. 306. Copyright 1994 by the American Psychological Association. Reprinted with permission.

involvement in intellectually stimulating activities, the maintenance of a high level of perceptual processing speed, a flexible personality style at midlife, as well as satisfaction with one's life's accomplishments (Schaie, 1984) all are important contributors to successful intellectual aging. The role of complex living arrangements, including a person's work environment, for the maintenance of high levels of intellectual functioning has also been documented by Kohn and Schooler (1973; Schooler, 1987) and has been documented in animal models for a long time (for a review, see Labouvie-Vief, 1985).

1.09.3.1.3 Plasticity of intelligence

Finally, Schaie and Willis (1986) provided the most compelling evidence for the intraindividual plasticity (i.e., intraindividual modifiability) of intellectual functioning when they showed that reliable decline in the fluid abilities of spatial orientation and inductive reasoning was reversible through specific training efforts. This finding, by the way, is consistent with the increasing research documenting the plasticity of the aging human brain (for a review, see Labouvie-Vief, 1985; Scheibel, 1996).

This latter finding can also be seen as part of the third major development that contributed to a changed view on adult intellectual development. In the early 1970s, Baltes and Labouvie (1973) suggested that researchers should systematically examine the modifiability of older adults' intellectual abilities. During the first phase of cognitive intervention research with older adults, these suggestions resulted in a number of studies showing that older adults' performance on intelligence tests could be improved through simple practice (Hofland, Willis, & Baltes, 1981) as well as through special training programs (Baltes, Dittmann-Kohli, & Kliegl, 1986; Baltes & Willis, 1982; Hayslip, 1989). Later, Baltes and his colleagues (Baltes & Kliegl, 1992; Kliegl, Smith, & Baltes, 1989), however, also showed that older adults' range of plasticity is more limited than younger adults' range of plasticity. Using the "testing-the-limits

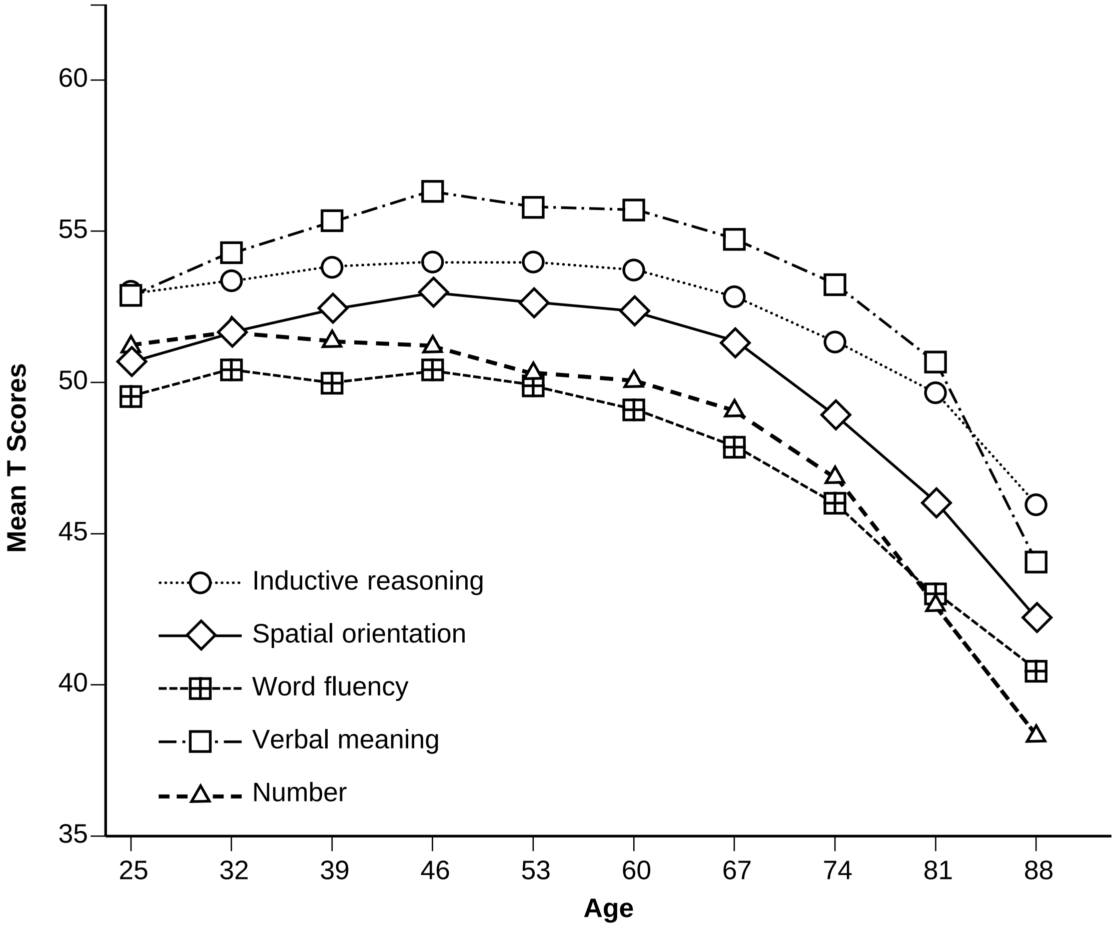

Figure 3 Longitudinal estimates of mean T scores for single markers of primary mental abilities (from 7-year within-subject data)."The course of adult intellectual development," by K. W. Schaie, 1994, *American Psychologist, 49,* p. 306. Copyright 1994 by the American Psychological Association. Reprinted with permission.

paradigm," Baltes and Kliegl (1992), for example, trained older and younger adults in the mnemonic strategy of the method of loci and showed that in no instance were the trained older adults able to outperform the trained younger adults. Instead, training actually increased the separation between younger and older adults, providing evidence that older adults' reserve capacity seems to be more limited compared to younger adults' reserve capacity, an argument that had previously been made by Botwinick (1977).

An interesting contribution of cognitive intervention research in adulthood is that it also highlights the interface between cognitive functioning, personality, and affect. In particular, several researchers have shown that improvements in older adults' intellectual performance can even be achieved, through interventions that focus on personality-related processes such as self-efficacy beliefs (Labouvie-Vief & Gonda, 1976; Lachman, Weaver, Bandura, Elliott, & Lewkowicz, 1992) or anxiety reduction (Hayslip, 1989; Hayslip, Maloy, & Kohl, 1995) rather than on direct cognitive strategies training. These studies have provided substantial evidence for a rather strong connection between personality and cognitive functioning in adulthood, an issue to be further discussed in a later section.

1.09.3.2 Everyday Intelligence

In recent years, researchers' interest in adults' intellectual performance has also shifted to some new areas of inquiry. In her 1985 review of the literature on adult intelligence and cognition, Labouvie-Vief (1985), for example, argued for a "re-examination of the concept of intelligence as an adaptive capacity" (p. 501) and presented findings with regard to the adaptive function of cognition in adulthood. This notion was echoed by Salthouse (1990) when he stated that "one of the greatest challenges in the field of psychology and aging is to account for the discrepancy between the inferred cognitive status of older adults based on their performance in psychometric testing situations and cognitive laboratories and that derived from observations of their successful functioning in everyday situations" (p. 310).

Thus, in recent years older adults' performance with regard to practical or everyday problems has become a major focus of cognitive aging research (Poon, Rubin, & Wilson, 1989; Puckett & Reese, 1993). Several researchers, for example, have examined older adults' intellectual performance with regard to tasks that are designed to simulate situations of everyday life (for review see Diehl, Willis, & Schaie, 1995). Overall, this research has shown that older adults do not solve everyday problems as effectively as middle-aged individuals but that they usually outperform young adults (Denney & Palmer, 1981; Denney & Pearce, 1989). That older adults' practical problem solving ability is not completely independent of their functioning on psychometric tests was shown by Diehl et al. (1995). These authors found that older adults' performance on a set of observed tasks of daily living showed substantial relations to the abilities of perceptual speed, memory, fluid intelligence, and crystallized intelligence. Indeed, fluid intelligence was found to be the strongest predictor of older adults' performance on everyday problem-solving tasks (see also, Willis & Schaie, 1986). Overall, these findings provide support in favor of a hierarchical perspective of everyday cognition. The main assumptions of this perspective have been outlined by Willis and Schaie (1993) and can be summarized in three postulates. First, problems encountered in everyday life require the activation and application of multiple mental abilities and processes for their solution. Second, different types of everyday problems require the application of different mental abilities. Third, competence with regard to basic cognitive abilities and processes is a necessary but not sufficient condition for the successful solution of everyday problems; domain-specific knowledge is likely to be required as well for successful everyday problem solving.

In summary, theory and research on the development of psychometric intelligence has always been at the forefront of life-span developmental psychology (see Baltes, 1987) and has greatly influenced researchers' thinking in other domains. For example, concepts such as multidimensionality, multidirectionality, plasticity, or sociocultural influences on development are among those that are now increasingly incorporated in research on personality development and the self. Moreover, perspectives on positive development in adulthood and aging were first explored in the context of intellectual and cognitive development and are now translated into other areas of human development.

1.09.3.3 Cognitive-developmental Theories

As compared to the psychometric approach, cognitive-developmental approaches are relative latecomers to the field of life-span changes in cognitive functioning. Yet their origins come out of the same concerns about validity that gave rise to much contemporary research within the psychometric tradition, namely issues relating to the validity of tests when applied outside of the settings for which they were originally developed. Thus, similar to Piaget's original criticism that quantitative differences in intellectual achievements reflect core differences in the organization of thinking, researchers in the field of adulthood wondered if thought structures in adulthood and later life might not reflect coherent reorganizations of thinking (see Labouvie-Vief, 1980, 1982a, 1982b; Riegel, 1973).

1.09.3.3.1 Piaget's legacy

When Piaget's cognitive developmental view arrived on the US scene, it created no less than a revolution in the field of developmental psychology. His proposal that intelligence and cognition evolved in the same way that other biological structures do, in terms of a series of qualitatively different and sequentially organized stages, proved an extremely powerful umbrella for a host of developmental changes from infancy to early adulthood. In the late 1990s, it is widely accepted that the evolution of structures akin to those he proposed underlies a large variety of cognitive competencies, from those that are rather formal and concern themselves with scientific thinking, to those that are more focused on everyday kinds of issues such as the understanding and regulation of emotions and representations of the self (Labouvie-Vief, DeVoe, & Bulka, 1989).

Nevertheless, the theory recently has been widely criticized. In keeping with the issues discussed in the first section of this chapter, Piaget's theory evolved in an intellectual context that emphasized the priority of rational functioning, which formed the endpoint of Piaget's theory. Similarly, the theory proposed a single, universal pathway to this endpoint. It further emphasized coherence of functioning at each stage of development, such that stages were assumed to provide broad integrative frames for all of the individual's thought and behavior.

This objectivism in Piaget's theory emerges as somewhat of a paradox, since we have noted that the theory was based on the basic assumption—shared in many ways with Freud—that mental life could be explained as a form of biological life. Thus, his work also marks a watershed in a broad trend to re-embed the mind in the organismic and biological world. In the abstract, that assumption is evident in many aspects of the theory: in its empirical implementation, however, Piaget's theory is successively slanted towards a form of mentalism that becomes dissociated from its organic base.

Riegel (1973) has commented on this slant and argued that in Piaget's theory, development is represented as a successive abstraction of thought from a rich organismic base. While in the theory we are confronted with children who are emoting, playing, symbolizing, and struggling with issues of authority, Piaget's adolescent is described almost exclusively in terms of idealized structures of mathematical and scientific thinking. These structures begin, according to Piaget, in cognition's close ties to sensorimotor processes, but gradually they become more autonomous and less dependent on immediate sensory experience. Thus sensorimotor intelligence moves into the symbolic processes of preoperational intelligence, and then begins to concern itself with more abstract relations and operations.

In "concrete operations," these still are closely tied to the concrete realities of the perceived world, but in "formal operations" they become the ability to operate on purely abstract processes, relations, and thoughts. Thus, in Piaget's theory, meaning systems that originate in the organismic, the sensorimotor, the figurative, the dynamic, and the personal gradually are displaced by ones that are abstract, conceptual, stable, conventional, and impersonal. But in general, the theory does not offer an integration of the two, except in the most abstract discussions of the dynamics of development.

These general criticisms have led, since the late 1970s, to two important expansions of a Piagetian view. One pathway has been to develop neo-Piagetian theories that expand Piaget's views to other domains and contexts (see Case, 1991; Fischer, 1980). In general, these theories hold to the notion that there is an ordered sequence of levels of complexity. Yet they make a clearer differentiation between this logical sequence of levels of complexity, and how these levels actually are translated into thought and behavior. Thus issues such as context and domains of familiarity all become significant variables to be added to a theoretical account of development. Further, another significant trend has been that the levels are generalized beyond the domain of scientific thinking to domains that are more strongly related to emotions, self, values, and relationships (Case, 1991; Damon & Hart, 1982; Fischer & Ayoub, 1994).

Another core approach has primarily focused on adults, and on re-evaluating notions of maturity that may be better applicable to adulthood. As argued by Labouvie-Vief (1994), the role integrating formal and organismic aspects of thinking may become especially important in adulthood, where individuals become more active participants in the construction of knowledge. To an extent, the dynamics of early development may require that the individual put aside private and inner forms of meaning, since the process of socialization requires an outward turn. However, in later development that outward turn must be balanced with an inward orientation: a higher degree of introspection, understanding of intrapersonal dynamics, and historical and psychological analysis.

1.09.3.3.2 Mature reasoning structures

Riegel's (1973) notion of the emergence of dialectical operations in adulthood has inspired several authors to describe such dialectical, multilevel modes of thinking. Riegel's work has remained rather abstract, but many empirical approaches took their impetus from Perry's (1970) influential study on how college men rediscover the relationship of reasoning to subjective and interpersonal variables. A similar study has since been done with women (Clinchy & Zimmerman, 1982; see also Belenky, Clinchy, Goldberger, & Tarule, 1986). Both of these studies argued that formal thinking makes the individual uniquely vulnerable to a dualistic emphasis on product rather than process, stasis rather than change. This rigidity was loosened as the youth moved into a position of contextual relativism. Now the individual rejected the notion of any one certain truth and asserted instead that truth is inherently relative to a context. Eventually, this fairly radical relativism is integrated in a final position of "commitment in relativism": in the midst of multiple logical choices, the individual now accepts the need for an integration that is ultimately subjective.

The notion that a stage of contextual relativism follows one of the certainty of formal operations has influenced most subsequent work, but most authors have sided with Kohlberg in searching for a final stage that goes beyond a merely subjective synthesis and strives for some higher order integration. Just how to describe this most mature level of cognitive functioning has varied, however, from author to author. One approach is the work of Commons and collaborators (e.g., Commons, Armon, Richards, & Schrader, 1989; Commons, Richards, & Kuhn, 1982). These authors suggested that, although formal thinkers are able to abstract order relations in a systematic manner, the entire system is not regarded as a single entity that can be analyzed in terms of characteristics and relationships and thus systematically compared to other systems. In contrast, more advanced thinkers are able to describe features of different systems in terms of an underlying common language which yields transformations across systems. This work is echoed by that of other authors who also note that the hallmark of mature adult thinking is the ability to move from intrasystem thinking to the ability to construct co-ordinations and transformative relationships between abstract systems (e.g. Labouvie-Vief, 1982a; Fisher, Hand, & Russell, 1984; Sinnott, 1989).

Other authors have specifically suggested that adulthood brings an enhanced ability to think dialectically (e.g., Basseches, 1984; Kramer, 1989; Labouvie-Vief, 1982a, 1982b, 1992; Reich, 1990; Sinnott, 1989); thus, adults are better able to analyze systems and concepts in terms of history, context, subjective processes, and change and movement. Kramer (1989), for example, proposes that thinking from adolescence to mature adulthood develops from a dualistic and absolutistic, through a relativistic, to a dialectical level. For example, in a comparison of young, middle-aged and elderly adults, Kramer and Woodruff (1986) found older adults to display the highest level of relativistic and dialectical thinking. There also was some evidence that Piagetian formal thinking was prerequisite but not sufficient for dialectical thinking, but that relativistic thinking actually was necessary but not sufficient for formal operations.

Much of the work above has concentrated on relatively abstract features of the coordination between systems. Just what are the kinds of systems and system features that mature individuals are able to analyze? Several studies have focused on how individuals can coordinate such abstract concepts as truth and certainty on the one hand, and subjectivity and interpretation on the other. For example, Kitchener and King (1981; see also Kitchener & Brenner, 1990; Kitchener, King, Wood, & Davison, 1989) have offered a detailed account of how individuals come to coordinate such categories as objective knowledge and subjective belief. These authors also carry further Perry's contention that the mark of mature rationality is a return of a new, if more educated, form of subjectivism. Rather, they suggest that from this more abstract form of subjectivism, the individual again begins to search for a set of "objective" criteria. However, these criteria are more procedural and dynamic: even though one may not, in principle, know what is good evidence, one has a general procedure for distinguishing solid evidence from less trustworthy evidence by examining the process through which knowledge has been gained.

This work has many implications for modern discussions about the nature of mature reason. Beyond these implications, it may also help highlight the cognitive roots of some major psychological differences between adolescents and adults. Labouvie-Vief (1982a), for example, has argued that the adolescent's understanding of truth as certain and computable may be a significant factor in her or his defense structure. Since that structure is based on a dualism between objective and subjective processes, it does not involve a mechanism by which such defensive cognitive distortions can be analyzed and, therefore, corrected. In turn, the movement to a model of knowledge that is more historically situated may bring a more integrated structure that, by analyzing its own subjectivity, provides a more powerful device for self-regulation and correction (Blanchard-Fields, 1986; Pascual-Leone, 1984; Wood, 1983).

1.09.3.3.3 Cognitive development and the self

The research discussed above suggests that as individuals move into mature adulthood, they are better able to relate the process of thinking back to a set of subjective processes (intentions, values, historical conditions, etc). Thus not surprisingly, these lines of thinking also have been extended to individuals' thinking about self and emotions. A major influence on this work has been Kohlberg's (1984) theory of moral development. Kohlberg (e.g., 1984) has addressed himself most explicitly to morality, but his theory has a much broader sweep and in fact involves an implicit theory of evolving self-structures around which values and emotions are organized. These implications have been worked out by a group of subsequent researchers both in the childhood (e.g., Damon & Hart, 1982; Selman, 1980) and adulthood (see below) areas. For example, Kegan (1982) has presented a theory of self-development from adolescence to mature adulthood. According to Kegan, the youthful and/or conventional self remains fused with an interpersonal and then institutional matrix, and the ability to experience distinctness and an individuated self remains limited. That capacity to maintain a more autonomous sense of selfhood emerges at the final stage, when self and other can be understood as entities that transcend interpersonal and institutional meanings. Thus, both a more authentic sense of selfhood and a deeper capacity for intimacy can result.

Armon (1984) has applied Kohlberg's notions to conceptions of what constitutes a good or ethical life and work. In her study, preconventional thought is focused around merely individual interests and desires, while conventional thinking becomes organized around such goals as socially beneficial work, positive interpersonal experiences, financial security, and social utility. For the postconventional individual, the potential conflict between social commitment and self-fulfillment is heightened and the individual aims at balancing responsibility to self with that to society.

Fowler (1981) has also applied Kohlbergian (as well as Eriksonian) notions to another domain, that of faith and religious values. Fowler argues that the preconventional individual has a magical and concrete conception of God. At the conventional level, the orientation is based on acceptance of cultural values with only little critical evaluation: religion now is oriented towards maintaining dogma and interpersonal validation. In contrast, the postconventional individual becomes both critical of conventional religion, and evolves a perspective in which conventional religious frameworks form separate though equivalent paths towards more universal spiritual goals.

Another theory proposing a broad reorganization of the self in relation to the social collective is Loevinger's (1976; Loevinger & Wessler, 1970) model of ego development. Loevinger proposed that the impulsiveness of the child gives way to a conventional mode in which one's own subjective inner life is suppressed and subordinated to social norms. The mature individual, however, evolves a more flexible language in which the conflicts between impulse and norm, self and society, or inner and outer, are first acknowledged and then integrated within more self-chosen standards. Cognitive complexity thus replaces a youthful language of self-regulation preoccupied with physical, mental, or emotional control, good–bad dichotomies, and little or no tolerance of intrapersonal or interpersonal conflict. The developmental significance of Loevinger's conception and measure has been demonstrated in age-related trends in both cross-sectional and longitudinal samples (e.g., Cook-Greuter, in press; Hauser, 1976; Redmore & Loevinger, 1979). Ego development is also a good index of such dimensions of developmental complexity as moral development and empathy (see Loevinger, 1976b) and intellectual development (see below).

Labouvie-Vief and collaborators (Labouvie-Vief, Chiodo, Goguen, Diehl, & Orwoll, 1995; Labouvie-Vief, Diehl, & Coyle, 1996) developed a framework within which to examine such changes. In this framework, statements about self and others (in particular, parents) are classified into five developmental levels that are ordered in terms of increasing complexity (see Table 1). In this research, younger or less mature individuals framed self and others in terms of a conventional perspective: self and others were described in terms of an organized, codified, and abstract set of role expectations. At a more advanced level, the institutional values become susceptible to doubt and criticism: for example, such values can be "carried too far." Instead, a dynamic perspective evolves in which descriptions of self and others convey in vivid language the unique and evolving experience of individuals within the context of their particular life histories. Lives now are understood in the context of multiple frames, cultural, social, and psychological, for example. There is keen insight into the psychological dynamics that are at the root of human diversity, yet an understanding that such diversity appears to be regulated by a common human heritage.

Just as the self becomes viewed more from the perspective of historical patterns and general emotional transformation, so do the parents. Younger individuals and those over the age of 60 primarily describe their parents in the interpersonal context of their roles as providers of emotional and financial support to the self, or, to a lesser extent, in the institutional context of their societal position. Few youthful individuals represent their parents as more autonomous individuals in their own right. In contrast, around midlife there is a peak of responses that are appreciative of the unique individuality of parents: Participants describe their parents not just as carriers of parental and other social roles, but show an awareness of the conditions that shaped the parents and made them become the persons they are or were.

These results are consistent with views which suggest that a restructuring of representations of one's parents are part and parcel of the reorganizations in self often associated with middle adulthood (Jung, 1933; Kohut, 1977;

Table 1 Levels of self- and other-differentiation.

Level	*Description*	*Example*[a]
0 Concrete/ presystemic	The language used is simple and concrete. Characteristics and physical traits are seen as global. Events are detailed in simple seriation. Action-oriented behaviors describe activities. No references to goals or psychological processes occur.	I am nice. My dad is very nice because he buys me cool stuff. She is the nicest mom in the world. I'm tall.
1 Interpersonal/ protosystemic	Simple evaluations are made that reflect the values of the immediate social group. Traits described are nondifferentiated. Individuals are described in terms of relationships (simple descriptors) and social networks. An emphasis is placed on features of the self or others that make for in-group acceptance.	I like to fool around in class and make my friends laugh. My mom is young that is why she understands when I get into trouble. My father is very protective of me, and he cares about me.
2 Institutional/ intrasystemic	Interpersonal descriptors indicate a clearer sense of the individual within the social group. Traits at this level indicate a more self-directed and goal-directed individual whose evaluations are guided by achievement-oriented and conventional goals, values, and roles. Achievement of these goals and values is a frequent theme.	I am an empathic and committed friend. I work hard to support my children and really love them. A caring person, she will give her last dime to her children or anyone else she thinks is in need or something. He was a strict disciplinarian, bit of a drinking problem but never missed a days work.
3 Contextual/ intrasystemic	Descriptions are critical of convention, involve an awareness of how traits change, and give a sense of individuals with their own value system. Institutional goals are re-examined and put into historical or psychological perspective. Descriptions involve references to processes and contrasts over time.	I am a singer, an actress and a writer and want to use these talents more creatively than I do now. Accumulates material things and educational & professional credentials to prop herself up. At 83 my father is slowing down quite a bit and is not the same as he was most of his life.
4 Dynamic/ intersubjective	Roles and traits are described at a complex psychological level and reflect awareness of underlying, often unconscious, motivation and reciprocal interaction. Activities and goals are seen as subject to continual revision as one gains knowledge of oneself and others. Reference is to multiple dimensions of life history, and an emphasis on process, becoming, and emergence.	I work for profit now rather than for satisfaction, partly because of my (guilty) need to continue to support my family. A fine balance of stregnth & vulnerability in may aspects, however a marked inability to discern male competence and motive.

[a] Errors in grammar and spelling reflect actual responses.

Labouvie-Vief, 1994; Levinson, Darrow, Klein, Levinson, & McKee, 1978). Overall, however, it is also notable that relatively few individuals display the higher levels of parent representations. This finding that higher levels of representation are quite rare in the population and that they tend to be concentrated in the middle age period is also replicated by many other studies (see Labouvie-Vief, Chiodo et al., 1995).

1.09.3.3.4 Emotional development

Advances in cognitive complexity, such as the ones discussed, also have implications for the understanding and regulation of emotions (Labouvie-Vief, DeVoe, & Bulka, 1989). Following suggestions by Piaget (1980), a number of researchers have shown that as individuals re-evaluate and reinterpret aspects of reality, their understanding of the nature and causes of emotions changes as well. For example, from childhood to adolescence, several changes in emotional understanding and regulation are apparent (see Fisher & Ayoub, 1994; Harter & Monsour, 1992; Labouvie-Vief, in press; Labouvie-Vief, Hakim-Larson, DeVoe, & Schoeberlein, 1989). First, the emotional vocabulary becomes more differentiated, referring increasingly to complex blends of emotions and to emotions of contrasting valence. Second, the standards by which emotions are regulated become more abstract as the need for direct dyadic supervision is replaced by more internal and normative controls. Third, emotions also become understood as more internal processes and are described in terms of thoughts, wishes, and rules rather than direct physical actions and consequences.

Important as these skills are in the process of early socialization, they are particularly significant in the context of development at early life stages when individuals need to define their emotional lives in accord with external criteria: they begin to master a set of rules that permit them to regulate behavior in accordance with cultural dictates. In contrast to this outward movement, adulthood may bring a compensatory movement inward. A focus on inner dynamics, on private experience, and on rich organismic experience and emotive content now comes to the fore, a process Gutmann (1987) refers to as the "greening" of the mature individual.

This notion was supported in research by Labouvie-Vief and colleagues (Labouvie-Vief, DeVoe, & Bulka, 1989), showing that from adolescence to middle adulthood, emotional understanding and regulation show systematic differences. The language of emotions of younger individuals primarily emphasized their inner and mental nature. It was static; terms were used as descriptors and emphasized how one *should* feel; external rules and standards of conduct rather than the felt experience characterized individuals' expression of emotions. In contrast, mature adults—those around middle adulthood—gave evidence of a significantly reorganized emotional language. Feelings were described in terms of a vivid felt process, dealing not with static states, but with process and transformation. At the same time, individuals began to differentiate an inner realm of emotional experience from an outer realm of convention. The conflict between these realms was acknowledged, and the individual was concerned with accepting impulses and thoughts that previously seemed too overwhelming to accept.

This research suggests that in individuals who are older/more mature, emotions are represented in less dualistic and polarized ways. This conclusion is also supported by research of Blanchard-Fields (e.g., 1986; Blanchard-Fields, 1996) who showed that older adults' thinking about emotional conflicts is better integrated. For example, older individuals are better able to differentiate their (emotionally influenced) interpretations from a body of data they evaluate; they also are less likely to view others' behavior in static terms, but more likely to explain it in terms of contextual factors.

Evidence such as this may suggest that older adults may become experts at dealing with emotionally relevant information, a conclusion that is congruent with research examining how individuals of different ages interpret narratives such as stories that have a highly emotional and symbolic content (Adams, 1986; Adams, Labouvie-Vief, Hobart, & Dorosz, 1990; Jepson & Labouvie-Vief, 1992). That research suggests that adolescents and college students in reading text primarily focus on literal features, as they attend to the structure of actions and events depicted in the text. However, for mature and older adults the primary interest is not in this literal action–event structure, but rather in what it reveals about underlying emotional and motivational patterns of the human condition. To that extent, the mature adults' interest in text becomes more abstract and symbolic: A narrative does not refer to the concrete here and now of protagonists and their actions, but rather is taken as indicative of human actions in general (see also Jepson & Labouvie-Vief, 1992).

It is important to note that this focus on emotions is not always the result of genuine integration. For example, later adulthood may bring a more general bias for emotion-based responses, as noted by Blanchard-Fields (1994) and Carstensen (1992; Carstensen & Turk-Charles, 1994). In the research by Labouvie-Vief, and in Adams' work too, it was noted that the psychologizing approach appeared to occur with high frequency in older individuals. However, it did so in styles that differed in complexity and elaboration. Some adults would give summaries that involved complex back-and-forth references between the text base and some psychological process they thought the text symbolized; in fact, in Adams' (1986) study, such responses peaked in middle-aged adults. Other responses were rather global and undifferentiated, with affectively laden responses such as "The trouble is that nobody believes in God anymore," and it was these responses that were most frequent in the older group.

The observation that emotion-based language sometimes is rather global and undifferentiated suggests that even though older adults may use a more emotion-based language, this emotional language may not necessarily imply the ability to reason about emotions in complex and coordinated ways. Thus other mechanisms than emotional understanding may play a role. For example, Brandstädter (e.g., Brandstädter & Greve, 1994) and Carstensen (1996) both suggest that awareness of time and mortality are a major factor by which an increasing acceptance of emotions is ushered in, even though this awareness may be stimulated by events, such as illness, that are not related to age. However, although such realization may effect a switch in the preferred processing system as implied by Carstenson's work, such a simple switch must be distinguished from the kinds of processes of cognitive-affective integration we are concerned with. Further, it is also possible that certain cognitive restrictions, particularly in the domain of fluid intelligence, may limit some older individuals' ability to modulate complex affects (see Labouvie-Vief, Chiodo et al., 1995, Labouvie-Vief & Diehl., 1996; Labouvie-Vief, Diehl et al, 1995). Thus conceivably, an increase in emotion-based language might reflect *less* complex emotion-regulation strategies.

These findings on emotion language may be related to research on coping and defense processes. A number of studies (for review see Aldwin, 1994; Diehl, Coyle, & Labouvie-Vief, 1996) on coping/defense processes across the life-span have demonstrated that older individuals, when compared to younger age groups, show a pattern of more positive coping and less defensive coping. This pattern continues at least into the sixth decade. For example, in the research by Diehl and collaborators, older individuals scored lower

One current tendency in the literature is to propose that such findings indicate a flexible coping strategy by which older individuals adjust to the realities of later life by the adoption of "secondary" or "emotion-focused" coping strategies (see Blanchard-Fields, 1996; Brandtstädter & Greve, 1994; Carstensen, 1996; Schultz, 1996). However, it is not clear at this time whether coping strategies actually may represent a mixture of strategies. For example, data of the Diehl et al. (1996) study suggest that secondary strategies of meaning transformation (e.g., through the rescaling of goals and aspirations) may not invariably imply genuine transformation of affect, but also can indicate its denial or repression. Thus it is possible that some individuals maintain positive adaptation and self-organization under conditions of relatively high-level functioning; however, others may rely on coping strategies that involve denial, repression, and gating out of disturbing affect. Even though such more fragmenting techniques may serve to maintain emotional equilibrium in the face of declining resources, they may do so at the cost of an overall picture of more norm dependence and rigidity.

1.09.3.3.5 Wisdom

All of the above proposals indicate that rational processes on the one hand, and processes related to self and emotions on the other, are profoundly interconnected. Even though this is true of all of the life-span, it appears to be only at relatively advanced levels that individuals are able to represent and understand these relationships and integrate them within single, nonconflicting systems. This ability to bridge the tensions between the universal and the contextual, the theoretical and the pragmatic, and the rational and emotional is often referred to as wisdom (Baltes & Staudinger, 1993; Clayton & Birren, 1980; Labouvie-Vief, 1990; Staudinger & Baltes, 1994, 1996).

While all of the work discussed above related to processes of wisdom (see Sternberg, 1990), the most detailed research project thus far has attempted to provide a specific operational definition of "wisdom." According to Baltes and his colleagues (Baltes, Smith, & Staudinger, 1992; Baltes & Staudinger, 1993), wisdom, defined as expert knowledge with regard to important but uncertain matters of life, can be characterized by several components as outlined in Table 2. Individuals receive ratings on each of these components, and an overall "wisdom score" is derived. As is true of the studies reported in this section, Baltes and his collaborators found that although age in itself was not a sufficient condition for the development of wisdom, older adults were more likely to be nominated as wise, and those nominated in fact produced the highest wisdom scores. Baltes et al. (1992) also found that personality measures, in addition to standard tests of crystallized and fluid intelligence, predicted wisdom scores, underscoring the interface between personality and cognition (see also Staudinger & Baltes, 1996).

1.09.4 PERSONALITY DEVELOPMENT

As Riegel (1977) noted in his work on the history of life-span development, the area of personality and social relations was a relative newcomer to the field. Early theorists such as Freud had proposed that personality was set rather early, and that it was unlikely to change (except in the direction of regression) in the second half of life. In contrast, the post-World

Table 2 Use of the wisdom-related criteria to evaluate discourse about life matters.

Criterion	*Instantiations*
Factual knowledge	Who, when, where? Examples of possible different situations Multiple options (forms of love and marriage)
Procedural knowledge	Strategies of information search, decision-making, and advice giving Timing of advice Monitoring of emotional reactions Cost–benefit analysis: scenarios Means–ends analysis
Life-span contextualism	Age-graded contexts (e.g., issues of adolescence) Culturally graded contexts (e.g., change in norms) Idiosyncratic contexts across time and life domains (e.g., terminal illness) Interrelations, tensions, priorities of life domains
Relativism	Religious and personal preferences Current vs. future values Historical period Cultural relativism
Uncertainty	No perfect solution Optimization of gains vs. loss Future not fully predictable Backup solutions
	Examples of responses (abbreviated)
Low score	A 15-year-old girl wants to get married? No, no way. Marrying at age 15 would be utterly wrong. One has to tell the girl that marriage is not possible. [After further probing] It would be irresponsible to support such an idea. No, this is just a crazy idea.
High score	Well, on the surface, this seems like an easy problem. On average, marriage for 15-year-old girls is not a good thing. I guess many girls might think about it when they fall in love for the first time. And, then, there are situations where the average case does not fit. Perhaps in this instance, special life circumstances are involved, such that the girl has a terminal illness. Or this girl may not be from this country. Perhaps she lives in another country and historical period. Before I offer a final evaluation I would need more information.

"The search for a psychology of wisdom," by P. B. Baltes and V. M. Staudinger, 1993, *Current Directions in Psychological Science, 2,* p. 78. Copyright 1993 by the American Psychological Society. Reprinted with permission.

War II era has seen a dramatic rise in inquiries into change and stability in personality and self in adulthood and later life.

The question about change and stability in personality will be answered according to the investigators' theoretical position (Block, 1995a, 1995b; Costa & McCrae, 1980; Helson, 1993; McCrae & Costa, 1990) as to how personality has been defined and measured, and on how stability and change have been assessed. To date, two major theoretical approaches have dominated the field of personality psychology and have influenced research on personality development. Developmental level approaches have been strongly influenced by the psychodynamic tradition. This tradition historically was the first to elaborate on personality organization, and has proposed a set of rich and detailed hypotheses about developmental stages of life, emphasizing qualitative changes in personality adaptation at different life periods. The trait-theoretical approach, in contrast, has been much more methodologically oriented, and emphasized the basic continuity and stability of personality across the life-span. However, both approaches in recent years have begun to draw on each other, the psychodynamic tradition providing important theoretical questions, and the psychometric tradition offering a strong set of methods to examine some of these questions. Both of these approaches, moreover, have begun to embrace contextual propositions that more and more explore the ways in which self and personality respond to the unique settings in which individuals develop and age.

1.09.4.1 Developmental Level Approaches

Level approaches to adult personality are strongly influenced by psychoanalytic thinking, along with its more recent expansions into object relations and self theory (see Greenberg & Mitchell, 1983). Psychoanalysis itself reflected a significant movement in redefinitions of mind and self, pointing out that emotions and relationships with others are not derived processes, but the very ground on which cognition is built. Psychoanalysis thus followed a broad trend to take the study of the mind out of the transcendent and to make it part of the natural world.

1.09.4.1.1 Freud's legacy

Freud's theory pioneered the proposition that rational processes are built on a layer of desires and needs or instincts. Yet paradoxically, even though he helped link back the mind with natural processes, Freud—along with many other theoreticians, such as Piaget (see previous section)—retained an objectivist bias, since as in traditional philosophy, the superiority of rational secondary processes over the "irrational" functioning of the primary processes was maintained. Thus development in large part becomes a matter of the victory of conscious, ego-oriented thought, while thought related to the emotion- and symbol-based secondary process was considered of less value.

This dualism is reflected in the terms Freud chose to refer to the psychic structures that were primarily related to each of these thought processes: the id and ego, respectively. It is interesting here to reflect on the historical origin of these terms. As Bettelheim (1983) pointed out, when Freud's work was translated from the original German into English, the meaning of these terms were significantly changed, as well. Freud's translators chose such abstract, Latin, scientific sounding words as "id" and "ego," in contrast to Freud's choice of such simple, common language words "I" (the German *Ich*) for ego, and "it" (The German *es*) for "id." The word "I," of course, indicates an immediate, direct personal identification with part of ourselves. And it is interesting that for Freud, that direct place of self-identification is the structure that mediates decision making, planning, and conscious thought. Thus, the self is, in effect, equated with logos functioning. This is where we are assumed to feel ourselves most "at home," where our sense of reality is most secure. The word "it" in contrast connotes a distanced, rejecting, and critical attitude toward that other part of ourselves.

This devaluation of nonrational processes makes Freud's theory unsuited to deal with many issues important across the life-span. To begin with, the theory is individualistic, since it puts values of independence, good ego boundaries, and so forth at the center (see Gilligan, 1982; Labouvie-Vief, 1994). The focus is on the separate individual, even from the start: desires and instincts are processes first located in the individual and then cathected onto another (Greenberg & Mitchell, 1983). That other is not treated as a separate self, but merely as a projection screen for one's own wishes. This lack of genuine relatedness is particularly important in the area of gender relations and gender development, since masculinity is favored sexually as a primary, more valuable state. By metaphorically extending notions of the superiority of the phallic principle into the mental domain, ego, consciousness, moral restraint are lined up with masculinity *and* maturity, while the feminine is considered a less developed state. As a result, and following millennia of historical tradition (Labouvie-Vief, 1994), in Freud's theory women are not really represented as being different from the self, as having an independent existence, their own desires, their own subjectivity. Rather, they are merely represented as others, beings inferior to men, thus making them objects to the male action, mind and gaze (see Benjamin, 1988; Labouvie-Vief, 1994).

Another important limit of Freud's theory is that it does not deal with forms of thinking more closely related to the primary process domain in a progressive and positive fashion. For example, religious processes, group processes, and love all are treated as a regressive return to infantile wishes for merging. Such wishes are to be suspended from the rule of ego-oriented thought. Thus, Freud's theory does not distinguish between mature and immature forms of these feelings.

Finally, Freud does not account for cultural construction of this hierarchical model. He assumes it is inherent in universal processes of gender dimorphism and ego formation rather than a cultural model. Ego and superego structures thus are viewed as stable structures that do not require further change. Indeed, further change tends to be pathologized as "regressive"; and Freud believed that psychotherapy cannot be successful in the second half of life.

1.09.4.1.2 Beyond the ego

Some of these limitations of Freud's theory were addressed by two individuals who themselves were intimately familiar with Freud's ideas. The first, Carl Gustav Jung, was one of Freud's collaborators and his hoped-for successor. When he began work with Freud, Jung had already done important work on the nature of the unconscious. However, as he continued to work with Freud, his position gradually changed from that of his mentor, and eventually he broke from Freud in a bitter confrontation that has become famous.

Jung was critical of several basic tenets in Freud's theory. He believed that Freud had too negative a view of the nature of the unconscious, and that he put too much emphasis on the importance of ego control. Instead, Jung believed that the basic organizing principle, though it included the ego, was not primarily conscious but rather unconscious. Jung termed this general organizing principle the self, anticipating many modern efforts to delineate the nature of the self.

Jung also believed that Freud's view of the unconscious was too much based on the individual's personal experience: the unconscious in Freud's theory consisted mostly of contents that have become repressed in the individual's personal experience. Jung argued that this interpretation did not allow positive accounts of an individual's symbolic life as it is expressed in religion, art, dreams, and so forth. He believed that the unconscious also contained many experiences that are part of our heritage as a species, and that express important emotional experiences related to love and hate, sexuality and the generational flow. These experiences are expressed in the form of story telling and myth, ritual and visual symbol (see Labouvie-Vief, 1994).

Jung also argued that one result of Freud's idealization of the rational was an overvaluation of the masculine and a devaluation of the feminine. Unlike Freud's theory, which held that healthy development requires that men suppress the feminine and women the masculine, Jung believed that successful development requires that the person integrate the two. Thus he anticipated much recent research on the nature of androgyny.

As a result of these conceptual shifts, Jung and some of his students suggested that Freud's view of development was more pertinent for the first part of life, when in the interest of cultural adaptation and socialization the individual must suppress aspects of the self, including contrasexual tendencies. All of these movements (together with the cognitive limitations discussed in the previous section) encouraged the formation of fairly rigid dualisms between mind and body, thinking and emotion, conscious and unconscious, outer and inner, good and evil, masculine and feminine, and so forth.

In contrast to the socializing emphasis of early life, Jung (1933) suggested that around the middle of life, development shifts to individuation. Individuals now have an opportunity to accept unconscious motives, to integrate contrasexual aspects and tendencies, and in general to integrate the dualisms formed in early development. In this way, they begin to transcend a narrow self-identification with ego and superego and become guided by a more genuine dialectic relationship between ego processes and the deeper patterns of human experience expressed by the self. Jung felt that these patterns indicate deep longings and desires that are often unconscious. Yet they are expressed in stories and symbols worldwide, and they express universal aspects of the human experience. These midlife tasks, Jung believed, often were resolved in a crisis. Jung in fact coined the term "midlife crisis," and he suggested that such a crisis was a necessary and positive aspect of later life development. He thought that as a result of that crisis, individuals were able to experience an upsurge of creativity and well-being.

1.09.4.1.3 Ego and culture

Like Jung, Erik Erikson (1982) felt that Freud's model of development was framed too narrowly, resulting in an overemphasis on autonomy, a failure to account for the diverse role of culture, and a regressive view of the role of religion and spirituality in the lives of individuals. One of the important steps Erikson took was to expand Freud's psychosexual view of development to a psychosocial one. He did so by first rooting the development of autonomy in the basic fact of human relatedness. Second, he generalized modes of relating beyond adolescence to mature adulthood and later life, where individuals were better able to give up narrow self-concerns about their own survival and participate in the caring for and management of culture.

One significant feature of Erikson's theory is that it implicitly proposes a generalization of the Freudian notion of "identification." For Freud, processes of identification took place in the "phallic" period as a result of the internalization of parental attributes and rules. In contrast, Erikson proposes that the self defines itself through successively widening circles of relations. Thus, at mature adult levels, the self becomes invested in culture and even humankind and the human condition, in general.

Another significant feature of Erikson's theory is that he assumed that each stage of human development was defined by a genuine dialectical tension (he called this tension "crisis") between the opposing tendencies of autonomy and relatedness. Thus, the function of the first important relation, child–mother, is to build a secure base of trust, a foundation from which later autonomy could be built. However, each stage is defined in terms of the degree to which these oppositions can be integrated, and failures to achieve complete integration are the norm rather than exceptions. Thus, most individuals are able to develop a core self-sector that permits the experience of trust and hope, yet also may retain more problematic self-sectors in which basic mistrust and withdrawal prevail. At successive later levels, these core qualities evolve into broader, more generalized ones.

One of the core crises of adulthood, that of generativity vs. stagnation, revolves around the task of surrendering the youthful need for self-idealization (Gutmann, 1994). As individuals grow older they usually are confronted with the limits of their creativity, their physical beauty, or their sexual attractiveness. Finding these sources of admiration and of pleasure dwindling, individuals may respond with an increased need for admiration and self-affirmation. Yet they also are able to step back from this need for admiration and, in experiencing its painful hold on the self, let go of it and transmute it into a broader understanding of the human situation. And, out of this experience of loss, the individual can recover a sense of admiring and idealizing others. Altruistic forms of passing on one's knowledge or financial resources, such as in forms of mentoring or philanthropy, are a good example of such admiration of others. Ideally, such generative support of others is no longer subordinated to one's own need to find advancement of admiration (see also McAdams, 1994).

While according to Erikson the experience of generativity is necessary to ward off a pervasive sense of stagnation in middle adulthood, true generativity may not come easily to most adults. Ultimately, the breadth and depth of generative concerns also are related to the ability to face and integrate negative experience (Erikson, 1982). As Jung (1933; see also Whitmont, 1969) in particular has pointed out, in growing up individuals usually learn to split off negative aspects of reality and idealize its positive side. Yet mature adulthood is the time when these neglected aspects need to be integrated to allow a more balanced view of reality. To face warded off negative experience, yet not be defeated by it, is a difficult challenge that not all individuals face successfully. Accepting destructive impulses in oneself and others—confronting hate, envy, evil and corruption—can be devastating; yet moving to maturity requires that we be able to do so without exploitation or masochism.

Many individuals retreat from that crisis. Why? Erikson thinks that ultimately, this is rooted in our need to defend ourselves, our inability to trust. Jaques (1965) also suggests that the midlife re-examination is not necessarily resolved successfully by all. Instead, adults may retreat from the crisis:

> the compulsive attempts, in many men and women reaching middle age, to remain young, the hypochondriacal concern over health and appearance, the emergence of sexual promiscuity in order to prove youth and potency, the hollowness and lack of genuine enjoyment of life ... are familiar patterns. These are attempts at a race against time. And in addition to the impoverishment of emotional life contained in the foregoing activities, real character deterioration is always possible ... Increase in arrogance, and ruthlessness concealing pangs of envy—or self-effacing humbleness and weakness concealing fantasies of omnipotence—are symptomatic of such change ... These defensive fantasies are equally as persecuting, however, as the chaotic and hopeless internal situation they are meant to mitigate. (Jaques, 1965, p. 511)

Erikson also suggested that after the crisis of generativity vs. stagnation, the individual faces a last crisis, that of integrity vs. despair. At the close of life, individuals need to accept their own life cycle and the people in it as something that had to be and that permitted of no substitutions. Alternatively, individuals may not be able to reach such acceptance. Rather than accepting life as their own responsibility, they may feel that time is too short, that alternate roads are blocked; they may harbor feelings of disgust, misanthropy, and a chronic displeasure with particular institutions and particular people.

1.09.4.1.4 The duality of the self

Although both Jung and Erikson pushed the boundaries of Freud's theory, many recent criticisms suggest that one of their core concepts—that of the basic duality of the self—was not sufficiently developed, and both maintained a bias towards autonomy, individuality, and masculinity. That concept of duality should be seen as the full conjoining of two fundamental developmental lines which Bakan (1966) called agency versus communion and others have called an orientation to separateness versus connection (e.g., Gilligan, 1982; Miller, 1976), power and achievement versus intimacy and love (McAdams, 1985), or individuation and attachment (Franz & White, 1985). These orientations influence experiential forms of motivation, adaptation, and strategies of coping and defense. In mature development, these lines are integrated in configurations that include a consolidated self which combines autonomy and productivity with healthy relatedness.

One example of such a theory is Kohut's theory of the self (Kohut, 1977). Development, according to Kohut involves the dialectic of two tendencies, one related to our need to be admired by others, the other the need to admire others. The first tendency is related to narcissism/grandiosity, which in their immature forms reflect a defensively inflated sense of self in excessive need of admiration. Yet in its more positive forms, it gives rise to mature creativity. On the other hand, the ability to renounce one's narcissism gives rise to the ability to mirror and idealize others, a core ingredient of one's ability to love. These tendencies may occur in balanced relationship; but in less ideal cases, one may be inflated at the expense of the other. Kohut also thinks that midlife can be crisis-like, as individuals realize that early adaptations are no longer useful.

Many scholars have noted, however, that each tendency is maladaptive when present in exaggerated or unmitigated forms (e.g., Bakan, 1966). Thus when normal development is disrupted, distorted modes of adaptation occur along one developmental line or another. For example, Benjamin (1988) suggests that one immature form of these two lines occurs in the interlocking of roles of domination and submission, where the dominator combines self-inflation with the contempt and maltreatment of the other, while the victim or subordinate surrenders healthy narcissistic needs and develops excessive needs for mirroring. Specific psychopathologies associated with each form have been outlined by Blatt and Shichman (1983). The self-defining, "introjective" axis involves a complex of affects and behaviors associated with excessive concerns about identity, self-control, and self-worth. The introjective character uses counteractive defense mechanisms to manage aggression directed toward the self and others, and to fend off experiences of loss of individuation. The interpersonal or "anaclitic" axis involves a constellation of issues concerned with disrupted relatedness and intimacy. Avoidant defenses (denial, repression) are used against feelings of loneliness and alienation. Anaclitic pathologies involve issues of dependency, unity, and belonging.

Theoretical frameworks such as these have become extremely important in theories of development. One theoretical framework that has been highly successful in working with this polarity is attachment theory as developed by Bowlby. Attachment theory also is an outgrowth of psychoanalytic theory, but was reworked by Bowlby into an ethological theory including principles of control theory. Attachment theory is specifically built on a conjoining of needs for autonomy and attachment, as shown in Figure 4 (Hazan & Shaver, 1994). Ideally, individuals are able to shift from exploration and the development of autonomy to security- and closeness-seeking behaviors. However, if early relationships do not encourage flexible shifting between both tendencies, individuals are likely to develop different forms of less adaptive attachment styles. Further, these styles are assumed to have a stable influence throughout the life course.

1.09.4.2 The Trait Approach

On the surface, the trait or individual differences approach to personality and personality development appears not ideally suited for the study of change and transformation across the life-span (Allport, 1937; Cattell, 1979; Costa & McCrae, 1994; Eysenck, 1982; McCrae & Costa, 1990). Trait theorists believe that an individual's personality can be defined as a limited number of basic behavioral tendencies that the individual has and consistently displays. These basic tendencies may be inherited or acquired and may or may not be malleable over the course of development. Behavioral tendencies and temperament traits that are usually examined by trait theorists include extraversion/introversion, emotional lability (i.e., neuroticism), shyness, conscientiousness, rigidity/flexibility, hostility, and others. Trait theorists assume that over the course of the life-span, these basic tendencies interact with external influences to produce characteristic adaptations which show a high degree of stability (Costa & McCrae, 1994; McCrae & Costa, 1990). An important consequence of this assumption of stability/continuity is that the structure of individuals' personality can be assessed using self-report questionnaires or objective ratings from others.

1.09.4.2.1 Traits as dimensions of stability

Historically, the trait-theoretical approach has centered around three major questions: How many traits are there?; How many traits are necessary to describe individuals' personality in a comprehensive manner?; and How stable are personality traits across the human life-span? In order to provide an answer to the

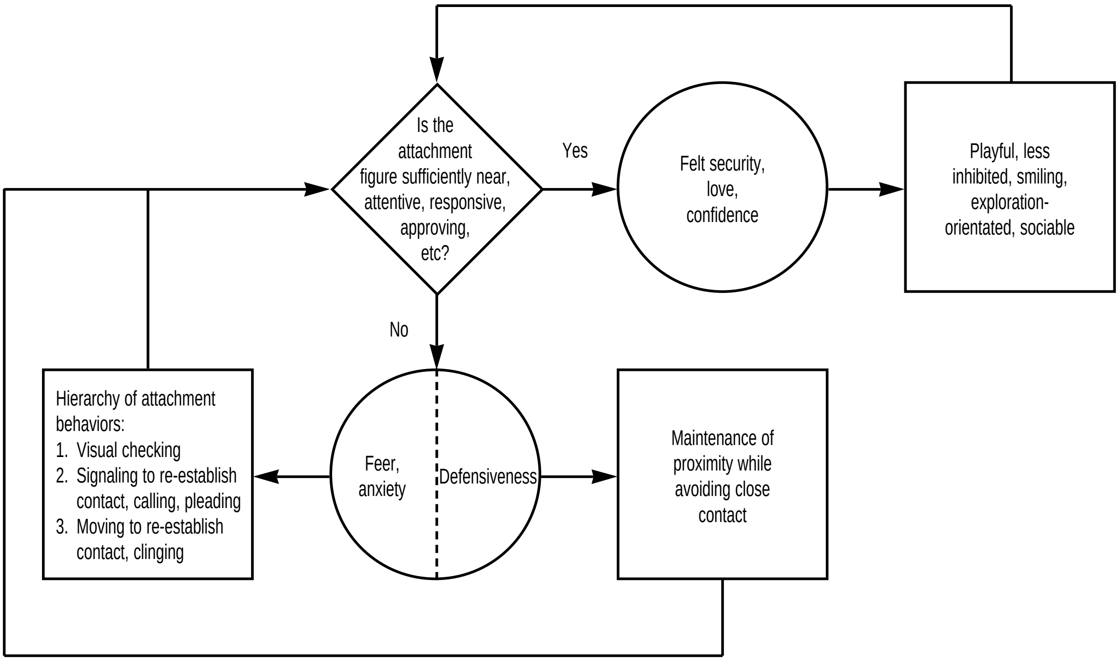
Is the attachment figure sufficiently near, attentive, responsive, approving, etc?
Yes
Felt security, love, confidence
Playful, less inhibited, smiling, exploration-orientated, sociable
No
Hierarchy of attachment behaviors:
1. Visual checking
2. Signaling to re-establish contact, calling, pleading
3. Moving to re-establish contact, clinging
Feer, anxiety
Defensiveness
Maintenance of proximity while avoiding close contact

Figure 4 The Attachment Behavioral System. "Attachmemt as an organizational framework for research on close relationships." by C. Hazan, & P. R. Shaver, 1994, *Psychological Inquiry, 5,* p. 3. Copyright 1994 by Lawrence Erlbaun Associates. Reprinted with permission.

first question, Allport and Odbert (1936) argued that all trait descriptions should basically be contained in common language, a proposition that was echoed by Cattell (1943a, 1943b), and resulted in the lexical strategy to personality assessment (Allport & Odbert, 1936). Other researchers (Cattell, 1947; Goldberg, 1981; Norman, 1963; Tupes & Christal, 1961/1992) built on Allport and Odbert's (1936) initial work and derived lists of trait names (for a detailed description, see Block, 1995a) which they analyzed in a variety of ways. In particular, the use of the mathematical technique of factor analysis resulted in various taxonomies of personality factors, ranging from 5 to 16 factors, that were considered sufficient for the assessment and description of personality. These endeavors based on the lexical approach were also paralleled by endeavors to construct questionnaires for the assessment of personality structures. Specifically, Costa and McCrae's (1992a) development of, and research with, the NEO and the NEO-PI-R has increasingly created the impression that five general personality factors (i.e., Neuroticism—N, Extraversion—E, Openness to Experience—O, Agreeableness—A, and Conscientiousness—C) describe individuals' personalities in a comprehensive fashion. Moreover, Costa and McCrae (1988, 1994) have provided considerable evidence showing that these five personality factors display a great deal of stability across the adult life-span.

Recent critics of the "Big Five" model of personality (see Block, 1995a, 1995b), however, have pointed out several limitations of this approach. First, Block has meticulously documented inconsistencies and subjectivity in the use of the lexical approach by different researchers (Block, 1995a). Second, Block has also questioned the inconsistent and often very subjective use of factor analysis, specific rotation methods, and the labeling of the derived factors to describe particular factor structures (Block, 1995a). Taken together, Block (1995b) has drawn two major conclusions based on his review of the relevant literature. First, he concluded that from the beginning the trait perspective has been inherently biased in its measurement approach toward the stability of traits, thus neglecting contextual influences and age-related factors that may contribute to variability in personality development. Second, he also concluded that advocates of the "Big Five" factor model present conclusions in support of the five-factor approach (FFA) far stronger than seems to be warranted (Block, 1995b). In contrast to the proponents of the FFA, Block (1995a, 1995b) has argued in favor of a perspective in personality research that incorporates both the sociocultural context in which personality development occurs and the notion of age-period specific variability across individuals' life course (see also Caspi, 1987). Thus, Block (1995a, 1995b) has argued in favor of a perspective that focuses on change as well as on stability in personality development. Block's position and concerns regarding the FFA have been shared by other personality theorists (Emmons, 1995; McAdams, 1992). Winter (1996), for example, has come to the conclusion that although "the five-factor theory is a promising development, it is hardly the final word even in trait psychology' (p. 471).

1.09.4.2.2 Context as moderator of stability

The contextualist orientation to personality and personality development rests on several basic assumptions. First, contextualists usually focus on subtypes of persons or personality syndromes and compare the developmental patterns of the identified subtypes over time, because they assume that personality development is *a priori* characterized by interindividual differences in the intra-individual continuity of personality characteristics (Block, 1981; Caspi, 1987). Second, contextualists place great emphasis on studying personality "the long way" (Block, 1993) in order to examine the effects of historical events (e.g., Great Depression, World War II, Vietnam War) and social changes (e.g., women's liberation movement, changes in child-rearing practices) as well as the effects of individual life events (e.g., parenthood, change in occupation, divorce) on individuals' personality development (Caspi, Bem, & Elder, 1989). Third, contextualists study personality from a transactional perspective (Magnusson & Törestad, 1993). That is, they examine how the motivations, goals and resources of individuals interact with the surrounding social world; and when and why the organization of motivations and resources changes as a consequence of changes in the social world (Helson, 1993). In summary, the contextual approach to personality development situates individuals in their social world and examines how the reciprocal interactions between person and social context affect the emergence, maintenance, and transformation of distinct personality characteristics. Thus, although contextualists are, to some extent, interested in questions of stability and change of personality across the life course, their primary focus concerns the issues of differential change patterns for different personality types. The assumption that both stability and change are part of personality development across the life-span is central to this perspective.

1.09.4.3 Empirical Research on Self and Personality

1.09.4.3.1 Stability and change

The diffrent perspectives in personality research differ greatly in their assumptions about stability and change. Researchers who subscribe to a trait perspective or any other kind of stability-oriented model (i.e., the attachment framework; Bowlby, 1988) assume a great amount of stability, whereas researchers who subscribe to a contextual perspective are more inclined to expect changes in personality. Given this controversial situation, a careful review of the respective literatures seems to be warranted.

What do studies in the individual difference tradition tell us about continuity or change in personality development? As Conley (1984) has pointed out, some of the most impressive evidence with regard to the longitudinal consistency of personality comes from studies of individual differences in children's emotional health. For example, ratings of children's emotional health have shown moderate consistencies over long time periods (Cox, 1970; Symonds, 1961; Vaillant, 1974) and have shown good predictive validity with regard to mental health outcomes (Johns, Mednick, & Shulsinger, 1982). Moreover, the temperamental qualities of three year olds have shown considerable predictive validity with regard to personality traits in young adulthood (Caspi & Silva, 1995; Kagan & Moss, 1962; Thomas & Chess, 1986).

These findings are complemented by research on personality development in adulthood. Using data from the Kelly Longitudinal Study, Conley (1984) showed for men and women moderate mean correlations for neuroticism indicators and for indicators of social introversion–extraversion across a 45-year period (see also Leon, Gillum, Gillum, & Gouze, 1979). In a subsequent study, Conley (1985) found that individual differences on the traits neuroticism, social extraversion, and impulse control were rather stable over a period of 20 years even when measured via different methods. Other longitudinal studies that have documented a great deal of stability in personality functioning are the Duke University Studies of Normal Aging (Siegler, George, & Okun, 1979), the Bonn Longitudinal Study of Aging (Schmitz-Scherzer & Thomae, 1983), and the Seattle Longitudinal Study (Schaie, 1996; Schaie & Willis, 1991). Although different measures of personality functioning were used in these studies, all three research programs showed remarkable stability (i.e., very limited change over time) in participants' self-reported personality descriptions.

The most vocal advocates of personality stability across the adult life-span have been Costa and McCrae (1980, 1988; McCrae & Costa, 1990) who have argued that after the age of 30 individuals' personality is "set like plaster" (Costa & McCrae, 1994). Based on data from the Boston Normative Aging Study (NAS) and the Baltimore Longitudinal Study of Aging (BLSA), Costa and McCrae (1988, 1994) have provided a considerable evidence in support of their position. For example, they have reported moderate to high 30-year stability coefficients (Costa & McCrae, 1994). Moreover, since the early 1980s, Costa and McCrae (1992b, 1994) have systematically elaborated the five-factor model of personality (John, 1990) and have devised a self-report questionnaire, the NEO-PI-R (Costa & McCrae, 1992a), that assesses the "Big Five." In terms of mean level differences, Costa and McCrae (1992a) have shown in cross-sectional studies that the "Big Five" show a fairly consistent pattern of age differences across the adult life-span: small declines are found for N, E, and O, whereas A and C show small increases. Although cross-sectional studies cannot rule out that the observed age differences reflect generational rather than maturational processes, findings from a six-year longitudinal study of self-reports and spouse ratings on the NEO-PI (Costa & McCrae, 1988) showed a longitudinal decline in self-reported N, but also a longitudinal increase in spouse-rated N. No significant time-related changes were found for either E or O. The most consistent longitudinal change found in adults over age 30 has been a decline in activity level (Costa & McCrae, 1988; Douglas & Arenberg, 1978; Field & Millsap; 1991).

Although this review of the literature shows that adult personality can be characterized by a fair amount of stability, high retest correlations by no means preclude the potential for change. For example, it is important to note that, in general, the test–retest correlations account for only 35–50% of the variance, leaving half of the variance unaccounted for. Moreover, this unaccounted variability is probably only in part due to lack of reliability of the measures, but is most likely related to true shifts in the ranking of people on the personality characteristic in question. Thus, overall it seems to be most appropriate to think about personality as an entity with relative stability depending on the life circumstances and the life experiences of each individual. We will return to this issue below.

1.09.4.3.2 Midlife crisis

As suggested above, whether or not stability remains high depends on the life period under consideration. One of the life periods that in particular has been associated with the potential for change and transformation is midlife. The notion that there are major reorganizations around midlife has been most strongly argued by theoreticians, most notably Jung (1932), Jaques (1965), and Levinson and collaborators (Levinson et al., 1978). All of these individuals have suggested that a reorganization of personality is a normative phenomenon, at least in an ideal sense. Thus, from a Jungian perspective, midlife is said to bring the opportunity to re-examine our constructions of good and bad, birth and death, dependence and independence, or masculinity and femininity; Erikson suggests a broadening of our ability to invest self in culture; and Jaques argues for a change in creative style from detailed and polarized to integrative and tragic. Further, Vaillant (1993) has suggested that around midlife there is a maturing of strategies of emotion regulation and defense (see also Labouvie-Vief, Hakim-Larson, & Hobart, 1987; Diehl, Coyle, & Labouvie-Vief, 1996).

Research related to the notion of the notion of a midlife crisis has not been extremely supportive of the notion, however. For example, there is little evidence that upheavals and dramatic change are related to a unique period of adulthood. Although some individuals do experience a crisis, those individuals may suffer from general problems of psychopathology (e.g., Rosenberg & Farrell, 1976). Thus, some researchers in the field have tended to conclude that the notion of a midlife reorganization has been overstated.

However, there is some fairly good evidence that many individuals experience some reorganization of self and values across the adult life-span. Why, then, should it be difficult to find clear evidence? There are two major reasons. One is the issue raised by Block (1995a; see also Labouvie-Vief, DeVoe, & Bulka, 1989) discussed earlier, that is the general bias in many current assessments towards stability. For example, our review of cognitive theories of adulthood suggests that the meaning individuals at different life-stages impart on test items may change quite profoundly with age or cognitive complexity. However, to the extent that tests rely on yes–no answers or simple scaling of agreement/disagreement with items, no room is left for different constructions of meaning.

This issue was raised by Labouvie-Vief et al. (1989) in the context of examining coping strategies across the life-span. While some researchers (e.g., Lazarus, 1991) have argued against generalized age differences in coping strategies, others (e.g., Diehl et al., 1996) have reported consistent patterns of strategies of coping and defense. One reason for this ambiguity is that current codings of coping strategies confound strategies of different levels of complexity and cognition–affect integration within a single category. In earlier research on self-regulation strategies, we (Labouvie-Vief et al., 1987, 1989; see also Block, 1995a) already noted that the same strategy can have very different meanings depending upon the developmental level at which it appears. For example, turning to others may imply that individuals attempt to dissipate anxiety through aligning with others "in the same boat"—a less mature strategy—or that they seek out others in an attempt to consider alternative choices and raise their level of objectivity a more mature strategy.

Another major issue to be considered is that the major dynamic driving changes associated with midlife may not be so much period-dependent, but rather may follow general cognitive changes. We have already reviewed a considerable body of neo-Piagetian research that suggests that individuals around the middle of adulthood show the most complex understanding of self, emotions, motivations, in general, attributes, that fall under the rubric of "personality." At the same time, this literature suggests that more significantly than age, independent measures of cognitive complexity are the strongest predictors of higher levels of complexity. Thus, phenomena usually associated with a midlife crisis may be the result of general gains in cognitive complexity from early to middle adulthood.

Connections between adults' personality organization and their cognitive complexity have increasingly been emphasized by Labouvie-Vief and her colleagues (Labouvie-Vief, 1994; Labouvie-Vief, DeVoe, & Bulka, 1989). Labouvie-Vief et al. (1989), for example, found in a study with adults, that age, verbal ability, and level of ego development were systematically related to a more reflexive understanding of emotions and to the use of more flexible coping strategies. Similarly, in a study on adults' self-representations, Labouvie-Vief, Chiodo, et al. (1995a) showed that in older adults a combination of intellectual and personality variables were the significant predictors of participants' complexity of self-representations, whereas younger adults' complexity of self-representations was predicted by cognitive variables only. Taken together, these studies are the result of a more positive perspective on the aging self and represent first attempts to address the interfaces between intelligence and personality in later life.

While, in general, these data suggest that so-called midlife changes are the result of rather gradual transformations, somewhat more dramatic crises may, however, be experienced by individuals whose early adaptations turn out to be problematic, and who as a consequence may encounter a relatively sudden and dramatic need to regroup their adaptive efforts and to reorganize their life structures. Stewart (1996) recently has proposed such a contextual view. Thus she argues that for most of the women studied, midlife may be a time of re-examination of their lives. If life patterns are found wanting, these women may engage in a midlife adjustment, correcting their patterns by such changes as moving from family to work orientation. This group of women subsequently shows a positive turn in their patterns of adjustment. However, not all women who are dissatisfied with their lives apply such an adjustment, and failure to do so is related to negative consequences for their subsequent life course.

1.09.5 CONTEXT IN PERSONALITY DEVELOPMENT

As described earlier, contextualists are more inclined than trait theorists to assume changes in personality development and to examine stability and change in response to changes in context (see Bloom, 1964; Helson, 1993). Bloom (1964), for example, reviewed a number of studies of personality stability and also examined the stability of other variables, including physical characteristics and intellectual abilities. Based on this review, he postulated that the change gradient of most traits, when plotted against age, is a positively sloped function with decreasing acceleration. Specifically, he rejected the idea that personality is fully developed by early adulthood.

That stability and change are characteristic of personality development during adulthood has been shown by several studies. For example, using data which covered a time span of 50 years, Haan, Millsap, and Hartka (1986) found that consistency indexes were largest for adjacent age periods but were considerably smaller when examined over the entire time span. In addition, consistency coefficients varied by age period, suggesting that personality stability may be considerably reduced during periods of the life course that are characterized by extensive transitions and role changes (e.g., the transition to parenthood in young adulthood). A similar pattern of findings has been reported by Skolnick (1986) with regard to attachment relationships. In brief, Skolnick's study provided a modest degree of support for attachment theory, showing some significant correlations between early relationship experiences with the mother and later social behavior.

The most striking finding, however, was the variability of individual relationship careers across the life-span, indicating the "developmental openness" of early relationship experiences to influences in later parts of the life course.

That both stability and change characterize personality development even in advanced old age has been shown by Field and Millsap (1991) for the Berkeley Older Generation Study. Although Field and Millsap (1991) found moderate rank order stability for the traits Satisfaction, Extraversion, Agreeableness, and Intellect, they also showed significant mean level changes in some of these traits over a 14-year period as study participants grew into old age. For example, more than one-third of the participants increased significantly over time in Agreeableness, and both men and women and old-old and oldest-old participants declined significantly in Extraversion. Field and Millsap (1991) concluded that these findings do not support the common stereotype that personality "rigidifies" in old age.

Taken together, these studies contribute in two major ways to the literature on personality development. First, in combination with the long-term findings reported by Conley (1984, 1985), these results document that the consistency coefficients tend to decline as the testing interval increases, resulting in only a modest degree of stability in personality components when individuals are examined over long periods of time. Second, these studies also suggest that periods that are characterized by life course transitions and changes in important social roles (e.g., work and family) are accompanied by considerably less personality stability than is found elsewhere in the life-span (see also Kogan, 1990). Thus, these studies provide evidence for the hypothesis that interindividual stability in personality varies at different portions of the life-span (Moss & Susman, 1980).

Bloom's (1964) argument that patterns of stability and change may reflect the influence of contextual factors has been addressed in an exemplary fashion in the context of several longitudinal studies of women's personality development during adulthood. Most notably, Helson and her colleagues (Helson, Mitchell, & Moane, 1984; Helson & Moane, 1987; Helson, Stewart, & Ostrove, 1995) have adopted a process approach and have examined the interplay between social context and personality development for several samples of adult women. Helson, Mitchell, and Moane (1984) used data from the Mills Longitudinal Study (Helson, 1967) to investigate the ways in which adherence or nonadherence to social clock projects in adulthood was related to women's changes in personality. Three social clock projects were identified in this study: the feminine social clock (FSC) for women who had adopted traditional feminine roles; the masculine occupational clock (MOC) for women who had chosen a career; and no social clock (NSC) project for those women who were neither committed to an FSC nor an MOC.

Using this typology, Helson et al. (1984) showed that women who adhered to the FSC scored significantly higher on personality variables indicative of conformity and norm orientation compared to MOC and NSC. During early adulthood, these women's personality development was dominated by the adaptation to the roles of wife and mother, which was frequently accompanied by a withdrawal from social life, the suppression of impulse and spontaneity, a more negative self-image, and decreased feelings of competence. Twenty percent of the women who adhered to the FSC relinquished this life structure and divorced between the age of 28 and 35. In contrast, women who had chosen the MOC at age 28 were less respectful of norms and more rebellious toward what they experienced as constrictive pressures. Although these women did not score lower on femininity or on well-being, they were more independent and self-assertive than their FSC counterparts (Helson et al., 1984). Long-term follow-up of the MOC women showed that those who continued to stay on the MOC into middle adulthood showed greater confidence, initiative, forcefulness, and intellectual independence than women who did not continue the MOC. Finally, women who were not engaged in any social clock project (NSC) also showed a distinct personality structure. In particular, they showed lower well-being and self-acceptance, were less independent and more norm-oriented. Virtually all NSC women were dissatisfied with their jobs, and the single women reported being lonely. By their early forties, some NSC women were better adjusted than in early adulthood and were working or had established a family. In summary, by taking an individual differences approach to patterns of adherence to the social clock, this study showed for a sample of adult women that distinct patterns of personality characteristics were related to different social clock projects and to different life outcomes across the first 20 years of adulthood.

In another study, Helson, Stewart, and Ostrove (1995) examined how different ego-identity patterns were related to women's personality characteristics and life outcomes in three different longitudinal samples. In particular, Helson et al. (1995) distinguished four ego-identity groups depending on whether individuals had an integrated or unintegrated identity and whether they had actively searched for an identity or had accepted a foreclosed identity. Across the three samples, results were consistent in showing that unintegrated accepters showed less initiative than other women, that unintegrated searchers had less impulse control, and that integrated accepters scored higher on support of norms and traditional values. These relations showed high consistency across time. Moreover, in the three cohorts, identity status showed different relations to life outcomes such as marital status, family status, and work. Overall, Helson and her colleagues (see also Helson & Moane, 1987; Helson & Wink, 1992) have shown that women's personality structures changed not only in systematic and normative ways in early and middle adulthood (Helson, 1993; Helson & Stewart, 1994), but that the observed changes were often related to specific changes in social roles and transitions in social contexts, thus creating distinctly different life paths for individuals with different personalities (see also York & John, 1992).

A great deal of work based on the contextual approach has also been conducted by Block and his research group (Block, 1971, 1993). In 1968, Block and Block (1980) initiated a longitudinal study of personality development in which they enrolled a heterogeneous sample of 128 children from two nursery schools in Berkeley, CA, which has been followed into early adulthood. Block and Robins (1993), for example, examined the longitudinal consistency and change in self-concept and self-esteem from early adolescence to early adulthood. They found that the self-esteem of girls showed a fair amount of continuity from ages 14 through 23, whereas the self-esteem of boys showed marked restructuring from age 14 to age 18, and reasonable consistency from ages 18 to 23. For both sexes, self-esteem was not related to intelligence or to social class. Young women and young men who scored high on self-esteem were independently characterized by observers as resilient, assertive rather than submissive, undiscouraged by adversity, without fluctuating moods, decisive, having a sense of personal meaning, responsive to humor, and unpreoccupied by ruminative fantasy (Block & Robins, 1993). However, there were also a variety of sex differences suggesting that self-esteem is embedded into very different characterological contexts for the two sexes. Young women who scored high on self-esteem seemed happy, warmly extraverted, and deeply concerned about interpersonal relationships, whereas young men with high self-esteem seemed self-focused and defensively critical, uneasy, and reluctant to make connections with others.

In another study, Block, Gjerde, and Block (1991) examined the personality antecedents of depressive tendencies in 18-year-olds (see also Gjerde, 1995). They found that as early as age seven, boys who subsequently reported feelings of depression were aggressive, self-aggrandizing, and undercontrolled, whereas girls with later depressive tendencies were shy and reserved, oversocialized, intropunitive, and overcontrolling. Similar gender differences were observed in early adolescence. At age 14, boys with later depressive tendencies were more likely to use marijuana and harder drugs, whereas girls with later depressive tendencies showed no tendency to use marijuana but did show a marked tendency to experiment with hard drugs. These girls also had a low self-esteem. Interestingly, in girls, preschool IQ correlated positively with depressive tendencies, whereas in boys preschool IQ was negatively related to depressive symptoms.

Finally, Block (1993) has reported data on the rank order stability of ego resiliency and ego control over a 20-year period. For the boys and young men, the correlations for ego resiliency were consistently positive throughout the years, suggesting that individual differences in ego resiliency are identifiable from an early age and continue over the next 20 years. The ego resiliency correlations for the girls, however, showed considerable discontinuity from childhood to adolescence and early adulthood. That means that for girls, being resilient during the childhood years had no implication for being ego resilient in adolescence or beyond. With regard to ego control, the cross-occasion correlations were consistently positive for both sexes, suggesting that from an early age on individual differences in the level of ego control are recognizable and continue to distinguish individuals for at least the next 20 years and, from the evidence of other studies (see Block, 1971), even beyond. Interestingly, however, the longitudinal relations between ego resiliency and ego control differed markedly again for girls and boys. Block (1993) explains the sex-specific pattern of relationships with the differential socialization of the sexes which makes girls grow up in a more structured and directive world than boys. In summary, Block and his colleagues (Block, 1993) have provided a great deal of evidence for the longitudinal plasticity and consistency of personality from childhood to early adulthood. In particular, they have shown that personality development proceeds differently for men and women, that consistency and plasticity are age-period specific and time-limited, and that contextual factors, such as parental child-rearing styles (Roberts, Block, & Block, 1984) and social class membership, influence personality development (see also Caspi, Bem, & Elder, 1989; Caspi, Elder, & Bem, 1987, 1988).

Support for a contextual explanation of personality development across the adult life-span has also been provided by Jones and Meredith (1996). In this study, latent curve analysis was used to examine six major aspects of personality for a sample of 211 individuals across either a 30- or 40-year time span. Jones and Meredith (1996) found that a systematic cross-time pattern existed for five of the six personality characteristics: four of the five curves indicated personality change, and one curve indicated no change. However, even with a unifying basic shape to individuals' cross-time patterns, Jones and Meredith (1996) found individual differences in the degree and direction of change, with some people showing increases, others showing decreases, and yet others showing little change in personality across the time period examined. Moreover, changes in the personality characteristics were in most instances time-limited and age-period specific, suggesting that individuals responded to life-stage specific challenges related to changes in social roles or age-graded social transitions (see Caspi, 1987).

In summary, studies which have examined personality development from a contextual perspective complement findings from trait-oriented studies. Overall, these studies show that both consistency and change characterize personality development across the life-span (Helson, 1993; Helson & Stewart, 1994). Moreover, contextually oriented studies have shown that men's and women's personality development often proceeds differently and is related to different problems of maladaptation and different manifestations of mental health problems (see Nolen-Hoeksema & Girgus, 1994).

1.09.5.1 Gender, Coping, and Personality

Empirical evidence that males and females undergo somewhat different development was provided by Cramer (1991). Throughout the school age period, males progressively externalized conflict, relying on protection and/or direct aggression as defenses, while females increasingly used defenses that internalized conflict, directing aggression inward. Similarly, Labouvie-Vief reports that women's predominant coping strategies are based on self-doubt and turning against the self, while those of men are more likely to use strategies of externalizing and dissociating (Diehl, Coyle, & Labouvie-Vief, 1996; Labouvie-Vief, Hakim-Larson, & Hobart, 1987). Perhaps as a consequence of these patterns, males and females evidence different patterns of vulnerability/resilience for psychological disorders with age. Special interest has been directed to gender differences in aggression, which increase from childhood to adulthood, while female prevalence of unipolar depression persists into adulthood and is generally twice that of males.

With respect to depression, women of diverse social roles and ethnic backgrounds appear to display similar risk factors associated with perceptions of personal control, power structures, and gender stereotyping; even in a relatively advantaged, professional group, women continued to struggle with vulnerability to depression associated with achievement and affiliation needs in often hostile environments (McGrath, Keita, Strickland, & Russo, 1990). Nolen-Hoeksema and Girgus (1994) examined the literature on sex differences in unipolar depression and found the most consistent explanation regarding the depressive experience related to people's responses, or coping strategy, to depressive episodes. While men tended to favor active responses to negative moods, such as thinking of other things, ignoring their problems, or participating in physical exercise, women became more ruminative and introspective, which matches a more passive and emotionally focused feminine style.

Profound differences also exist in how men and women subjectively experience their successes in the academic and intellectual domains. For example, many studies have demonstrated gender deviations in the relationship of actual intellectual performance and one's self-concept about those performances. Men and boys are more likely to adopt a self-enhancing strategy in evaluative situations (Dweck, Davidson, Nelson & Enna, 1978; Martin & Nivens, 1987; Roberts, 1991; Roberts & Nolen-Hoeksema, 1989). That is, men tend initially to overestimate their performance and remain unaffected by negative evaluations, while women underestimate their performance and are more reactive to others' feedback, both positive and negative. Such splits are supported by the feedback boys and girls receive from parents and teachers (American Association of University Women [AAUW], 1992; Yee & Eccles, 1988).

Contemporary research suggests that such patterns continue into the late 1990s. Despite few objective differences in intellectual status between males and females, women on average continue to attain lower levels of achievement than men (e.g., Kaufman & Richardson, 1982; Tomlinson-Keasey & Blurton, 1992). Although in early childhood, girls equal or surpass boys on most measures of school achievement, by early adolescence, they show a steady decline in self-esteem as they go through school. By high school, girls are found to retreat from achievement-related challenges, particularly in math and science (AAUW, 1992; Kerr, 1985; Terman & Oden, 1959; Yee & Eccles, 1988). Especially in such gender-stereotyped areas as math, this results in girls' disclaiming a sense of agency, even though they may outperform boys (Byrne, Shavelson, & Marsh, 1992).

The work of Gilligan (Gilligan, 1982, 1990) and others also shows that girls who, at the dawning of adolescence were self-assured and outspoken, become increasingly tentative and silent about what they know. In part, the girls' silence mirrors the silence of adult women who act as their models. "At the edge of adolescence, eleven- and twelve-year old girls observe where and when women speak and when they are silent" (Gilligan, 1990, p. 25). In part, they come to doubt that what girls and women know and experience has value in the adult world, because it too often "brings a message of exclusion—stay out; because it brings a message of subordination—stay under; because it brings a message of objectification—become an object of another's worship or desire, see yourself as you have been seen over centuries through a male gaze ... keep quiet and notice the absence of women and say nothing" (Gilligan, 1990, p. 26).

As noted earlier, however, these different orientations may be particularly characteristic of the early half of the life-span, while the second half brings a relaxation of the demands for "sex-appropriate" behaviors (Jung, 1932; Labouvie-Vief, 1994). As reviewed by Labouvie-Vief (1994), some individuals around the middle of adulthood do appear to experience such integrative changes. Accordingly, some women around midlife claim domains they have avoided thus far, moving from a stance of interiority to one of openness, assertiveness, and power. In turn, some men integrate their "feminine" aspects, accepting their vulnerability and nurturance. However, many studies have not carefully ruled out alternative interpretations. For example, Gutmann's discussions suggest that some men, rather than integrating notions of feminine power, become threatened and enfeebled by a sense of owerpowering, "masculine" women. Further, such postulated midlife changes have not yet been differentiated from cultural/cohort changes that may imply a move towards less gender role polarization.

1.09.6 CONCLUSIONS

In this chapter, we have reviewed some recent trends in the domain of life-span development. We have suggested that with the broad worldwide reorganization in population dynamics, age distributions, and information systems has come a new way to look at development across the life-span. Not only has the concept of development been pushed from early adulthood into later adult stages, but the kinds of characteristics supposed to indicate "mature" functioning also have been broadened. In general, there is a strong emphasis on complexity and introspection, on the integration of emotion and cognition, and on the ability to locate oneself in a complex network of contextual, historical, and interpersonal regularities. In addition, in contrast to traditional views, recent developmental approaches also have systematically highlighted the role of context in shaping individualized patterns of development.

This shift is visible both in the area of cognitive and of personality development. In both of these domains, research has increasingly departed from old stability assumptions and pointed out the wide range to which dimensions of cognition and personality change over individual and historical time. We have also placed emphasis on research that suggests that despite this diversity of developmental patterns, there may be several relatively uniform change patterns that characterize different stages of adulthood. Thus, research on transitions from early to late adulthood indicates that there may be a host of reorganizations that lend substance to older notions of a "midlife crisis," with increases in complexity of thinking about change and psychological process in self and others.

ACKNOWLEDGMENT

Preparation of this chapter was supported by grant 1RO1-AG09203 to the first author.

1.09.7 REFERENCES

American Association of University Women (1992). *The AAUW report: How schools shortchange girls.* Washington, DC: AAUW Educational Foundation.

Adams, C. (1986). *Qualitative changes in text memory from adolescence to mature adulthood.* Unpublished doctoral dissertation, Wayne State University, Detroit, MI.

Adams, C., Labouvie-Vief, G., Hobart C. J., & Dorosz M. (1990). Adult age group differences in story recall style. *Journal of Gerontology, 45,* P17–P27.

Aldwin, C. M. (1994). *Stress, coping, and development: An integrative perspective.* New York: Guilford Press.

Allport, G. W. (1937). *Personality: A psychological interpretation.* New York: Holt.

Allport, G. W., & Odbert, H. S. (1936). Trait names: A psycho-lexical study. *Psychological Monographs, 47* (1, Whole Issue No. 211).

Aries, P. (1962). *Centuries of childhood: A social history of* *family life* (R. Baldick, Trans.) New York: Vintage Books.

Armon, C. (1984). Ideals of the good life and moral judgment: Ethical reasoning. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), *Beyond formal operations: Late adolescent and adult cognitive development* (pp. 357–380). New York: Praeger.

Ashmore, R. D., & Ogilvie, D. M. (1992). He's such a nice boy ... when he's with grandma: Gender and evaluation in self-with-other representations. In T. M. Brinthaupt & R. P. Lipka (Eds.), *The self: Definitional and methodological issues* (pp. 236–290). Albany, NY: State University of New York Press.

Bakan, D. (1966). *The duality of human existence*. Chicago: Rand McNally.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology, 23,* 611–626.

Baltes, P. B., & Baltes, M. M. (1990). Psychological perspectives on successful aging: The model of selective optimization with compensation. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 1–34). Cambridge, UK: Cambridge University Press.

Baltes, P. B., Dittmann-Kohli, F., & Kliegl, R. (1986). Reserve capacity of the elderly in aging-sensitive tests of fluid intelligence: Replication and extension. *Psychology and Aging, 1,* 172–177.

Baltes, P. B., & Kliegl, R. (1992). Further testing of limits in cognitive plasticity in old age: Negative age differences in a mnemonic skill are robust. *Developmental Psychology, 28,* 121–125.

Baltes, P. B., & Labouvie, G. V. (1973). Adult development of intellectual performance: Description, explanation, and modification. In C. Eisdorfer & M. P. Lawton (Eds.), *The psychology of adult development and aging* (pp. 157–219). Washington, DC: American Psychological Association.

Baltes, P. B., Lindenberger, U., & Staudinger, U. M. (1998). Life-span theory in developmental psychology. In W. Damon & R. M. Lerner (Eds.), *Theoretical models of human development. Handbook of child psychology* (5th ed., Vol. 1) New York: Wiley.

Baltes, P. B., Smith, J., & Staudinger, U. M. (1992). Wisdom and successful aging. In T. B. Sonderegger (Ed.), *Nebraska Symposium on Motivation 1991: Vol. 39. Psychology and aging* (pp. 123–167). Lincoln, NE: University of Nebraska Press.

Baltes, P. B., & Staudinger, U. M. (1993). The search for a psychology of wisdom. *Current Directions in Psychological Science, 2,* 75–80.

Baltes, P. B., & Willis, S. L. (1982). Enhancement (plasticity) of intellectual functioning in old age: Penn State's Adult Development and Enrichment Project (ADEPT). In F. I. M. Craik & S. E. Trehub (Eds.), *Aging and cognitive processes* (pp. 353–389). New York: Plenum.

Basseches, M. A. (1984). Dialectical thinking as metasystematic form of cognitive organization. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), *Beyond formal operations: Late adolescent and adult cognitive development* (pp. 216–238). New York: Praeger.

Baumeister, R. F. (1986). *Cultural change and the struggle for self.* New York: Oxford University Press.

Belenky, M. F., Clinchy, B. M., Goldberger, N. R., & Tarule, J. M. (1986). *Women's ways of knowing*. New York: Basic Books.

Benjamin, J. (1988). *Bonds of love: Psychoanalysis, feminism, and the problem of domination*. New York: Pantheon Books.

Bettelheim, B. (1983). *Freud and man's soul*. New York: Knopf.

Biddel, T. R., & Fischer, K. W. (1992). Beyond the stage debate: Action, structure, and variability in Piagetian theory and research. In R. J. Sternberg & C. A. Berg (Eds.), *Intellectual development* (pp. 100–140). New York: Cambridge University Press.

Blanchard-Fields, F. (1986). Reasoning on social dilemmas varying in emotional saliency: An adult developmental perspective. *Psychology and Aging, 1*, 325–332.

Blanchard-Fields, F. (1996). The role of emotion in social cognition across the adult life-span. In K. W. Schaie (Chair), *Cognitive and affective aspects of emotion.* Symposium conducted at the 1996 Emotion and Adult Development Conference, University Park, PA.

Blatt, S. J. & Shichman, S. (1983). Two primary configurations of psychopathology. *Psychoanalysis and Contemporary Thought, 6*, 187–254.

Block, J. (1971). *Lives through time*. Berkeley, CA: Bancroft.

Block, J. (1981). Some enduring and consequential structures of personality. In A. I. Rabin, J. Aronoff, A. M. Barclay, & R. A. Zucker (Eds.), *Further explorations in personality* (pp. 27–43). New York: Wiley.

Block, J. (1993). Studying personality the long way. In D. C. Funder, R. D. Parke, C. Tomlinson-Keasey, & K. Widaman (Eds.), *Studying lives through time: Personality and development* (pp. 9–41). Washington, DC: American Psychological Association.

Block, J. (1995a). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin, 117,* 187–215.

Block, J. (1995b). Going beyond the five factors given: Rejoinder to Costa and McCrae (1995) and Goldberg and Saucier (1995). *Psychological Bulletin, 117*, 226–229.

Block, J., Gjerde, P. F., & Block, J. H. (1991). Personality antecedents of depressive tendencies in 18-year-olds: A prospective study. *Journal of Personality and Social Psychology, 60*, 726–738.

Block, J., & Robins, R. W. (1993). A longitudinal study of consistency and change in self-esteem from early adolescence to early adulthood. *Child Development, 64*, 909–923.

Block, J. H., & Block, J. (1980). The role of ego-control and ego-resiliency in the organization of behavior. In W. A. Collins (Ed.), *Minnesota Symposium on Child Psychology: Vol. 13. Development of cognition, affect, and social relations* (pp. 39–101). Hillsdale, NJ: Erlbaum.

Bloom, B. S. (1964). *Stability and change in human characteristics*. New York: Wiley.

Botwinick, J. (1977). Intellectual abilities. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 580–605). New York: Van Nostrand Reinhold.

Bowlby, J. (1969). *Attachment and loss: Attachment* (Vol. 1). New York: Basic Press.

Bowlby, J. (1988). *A secure base: Parent–child attachment and healthy human development*. New York: Basic Books.

Brandstädter, J. & Greve, W. (1994). The aging self: Stabilizing and protective processes. *Developmental Review, 14*, 52–80.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Cambridge, MA: Harvard University Press.

Bruner, J. (1986). *Actual minds, possible worlds.* Cambridge, MA: Harvard University Press.

Byrne, B. M., Shavelson, R. J., & Marsh, H. W. (1992). Multigroup comparisons in self-concept research: Reexamining the assumptions of equivalent structure and measurement. In T. M. Brinhaupt & R. P. Lipka (Eds.), *The self: Definitional and methodological issues* (pp. 172–203). Albany, NY: State University of New York Press.

Carstensen, L. L. (1992). Social and emotional patterns in adulthood: Support for socioemotional selectivity theory. *Psychology and Aging, 7*, 331–338.

Carstensen, L. L. (1996). Social context of emotional experience. In M. P. Lawton (Chair), *Social aspects of emotion*. Symposium conducted at the 1996 Emotion and Adult Development Conference, University Park, PA.

Carstensen, L. L. & Turk-Charles, S. (1994). The salience of emotion across the adult life-span. *Psychology and Aging, 9,* 259–264.

Case, R. (1991). Stages in the development of the young child's first sense of self. *Developmental Review, 11,* 210–230

Case, R. (1996). The role of psychological defenses in the representation and regulation of close personal relationships across the life-span. In G. Noam & K. Fischer (Eds.), *Development and vulnerability in close relations* (pp. 59–88). Hillsdale, NJ: Erlbaum.

Caspi, A. (1987). Personality in the life course. *Journal of Personality and Social Psychology, 53,* 1203–1213.

Caspi, A., Bem, D. J., & Elder, G. H., Jr. (1989). Continuities and consequences of interactional styles across the life course. *Journal of Personality, 57,* 375–406.

Caspi, A., Elder, G. H., Jr., & Bem, D. J. (1987). Moving against the world: Life-course patterns of explosive children. *Developmental Psychology, 23,* 308–313.

Caspi, A., Elder, G. H., Jr., & Bem, D. J. (1988). Moving away from the world: Life-course patterns of shy children. *Developmental Psychology, 24,* 824–831.

Caspi, A., & Silva, P. A. (1995). Temperamental qualities at age three predict personality traits in young adulthood: Longitudinal evidence from a birth cohort. *Child Development, 66,* 486–498.

Cattell, R. B. (1943a). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology, 38,* 476–506.

Cattell, R. B. (1943b). The description of personality: 1. Foundations of trait measurement. *Psychological Review, 50,* 559–594.

Cattell, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychometrika, 12,* 197–220.

Cattell, R. B. (1979). *Personality and learning theory: Vol. 1. The structure of personality in its environment*. New York: Springer.

Cavanaugh, J. C. (1997). *Adult development and aging* (3rd ed.). Pacific Grove/CA: Brooks Cole.

Clayton, V. P., & Birren, J. E. (1980). The development of wisdom across the life-span: A reexamination of an ancient topic. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (Vol. 3, pp. 103–135). New York: Academic Press.

Clinchy, B., & Zimmerman C. (1982). Epistemology and agency in the development of undergraduate women. In P. Perun (Ed.), *The undergraduate woman* (pp. 161–181). Lexington, MA: Lexington Book.

Commons, M. L., Richards, F. A. & Kuhn D. (1982). Systematic and metasystematic reasoning: A case for levels of reasoning beyond Piaget's stage of formal operations. *Child Development, 53,* 1058–1069.

Conley, J. J. (1984). Longitudinal consistency of adult personality: Self-reported psychological characteristics across 45 years. *Journal of Personality and Social Psychology, 47,* 1325–1333.

Conley, J. J. (1985). Longitudinal stability of personality traits: A multitrait-multimethod-multioccasion analysis. *Journal of Personality and Social Psychology, 49,* 1266–1282.

Costa, P. T., Jr., & McCrae, R. R. (1980). Still stable after all these years: Personality as a key to some issues in adulthood and old age. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (Vol. 3, pp. 65–102). New York: Academic Press.

Costa, P. T., Jr., & McCrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology, 54,* 853–863.

Costa, P. T., Jr., & McCrae, R. R. (1992a). *Professional manual: Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI)*. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., & McCrae, R. R. (1992b). Trait psychology comes of age. In T. B. Sonderegger (Ed.), *Nebraska Symposium on Motivation 1991: Vol. 39. Psychology and aging* (pp. 169–204). Lincoln, NE: University of Nebraska Press.

Costa, P. T., Jr., & McCrae, R. R. (1994). Set like plaster? Evidence for the stability of adult personality. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change*? (pp. 21–40). Washington, DC: American Psychological Association.

Cox, R. D. (1970). *Youth into maturity*. New York: Mental Health Materials Center.

Cramer, P. (1979). Defense mechanisms in adolescence. *Developmental Psychology, 15,* 477–478.

Cramer, P. (1991). *The development of defense mechanisms: Theory, research, and assessment*. New York: Springer-Verlag.

Damon, W., & Hart, D. (1982). The development of self-understanding from childhood through adolescence. *Child Development, 53,* 841–864.

Denney, N. W., & Palmer, A. M. (1981). Adult age differences on traditional and practical problem-solving measures. *Journal of Gerontology, 36,* 323–328.

Denney, N. W., & Pearce, K. A. (1989). A developmental study of practical problem solving in adults. *Psychology and Aging, 4,* 438–442.

Diehl, M., Coyle, N., & Labouvie-Vief, G. (1996). Age and sex differences in strategies of coping and defense across the life-span. *Psychology and Aging, 11,* 127–139.

Diehl, M., Willis, S. L., & Schaie, K. W. (1995). Everyday problem solving in older adults: Observational assessment and cognitive correlates. *Psychology and Aging, 10,* 478–491.

Dixon, R. A., & Bäckman, L. (Eds.) (1995). *Compensating for psychological deficits and declines: Managing losses and promoting gains*. Mahwah, NJ: Erlbaum.

Douglas, K., & Arenberg, D. (1978). Age changes, cohort differences, and cultural change on the Guilford–Zimmerman Temperament Survey. *Journal of Gerontology, 33,* 737–747.

Dweck, C. S., Davidson, W., Nelson, S., & Enna, B. (1978). Sex differences in learned helplessness: II. The contingencies of evaluation feedback in the classroom; III. An experimental analysis. *Developmental Psychology, 14,* 268–276.

Elias, N. (1982). *The civilizing process: Vol. 2: Power and civility*. New York: Pantheon.

Emmons, R. A. (1995). Levels and domains in personality: An introduction. *Journal of Personality, 63,* 341–364.

Epstein, S. (1990). Cognitive-experiential self-theory. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research*. New York: Guilford Press.

Erikson, E. H. (1982). *The life cycle completed*. New York: Norton.

Eysenck, H. J. (1982). *Personality, genetics and behavior*. New York: Praeger.

Field, D., & Millsap, R. E. (1991). Personality in advanced old age: Continuity or change? *Journal of Gerontology: Psychological Science, 46,* P299–P308.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review, 87,* 477–531.

Fisher, K. W., & Ayoub, C. (1994). Affective splitting and dissociation in normal and maltreated children: Developmental pathways for self in relationships. In D. Cicchetti & S. L. Toth (Eds.), *Rochester Symposium on*

Developmental Psychopathology: Vol 5. Disorders and dysfunctions of the self (pp. 147–222). Rochester, NY: Rochester University Press.

Fischer, K. W., Hand, H. H., & Russell, S. (1984). The development of abstractions in adolescence and adulthood. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), *Beyond formal operations: Late adolescent and adult cognitive development* (pp. 43–72). New York: Praeger.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95,* 29–51.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101,* 171–191.

Fordham, F. (1966). *An introduction to Jung's psychology*. New York: Penguin.

Fowler, J. W. (1981). *Stages of faith: The psychology of human development and the quest for meaning*. San Francisco: Harper & Row.

Franz, C. E., & White, K. M. (1985). Individuation and attachment in personality development: Extending Erikson's theory. *Journal of Personality, 53,* 224–256.

Freud, S. (1925/1963). Some psychological consequences of the anatomical differences between the sexes. In P. Rieff (Ed.), *Sexuality and the psychology of love* (pp. 183–193). New York: Macmillan.

Gardner, M. K., & Clark, E. (1992). The psychometric perspective on intellectual development in childhood and adolescence. In J. R. Sternberg & C. A. Berg (Eds.), *Intellectual development* (pp. 16–43). New York: Cambridge University Press.

Gergen, K. J. (1991). *The saturated self*. New York: Basic Books.

Gilligan, C. (1982). *In a different voice*. Cambridge, MA: Harvard University Press.

Gilligan, C. (1990, Fall). Joining the resistance: Psychology, politics, girls and women. *Michigan Quarterly Review,* 501–536.

Gjerde, P. F. (1995). Alternative pathways to chronic depressive symptoms in young adults: Gender differences in developmental trajectories. *Child Development, 66,* 1277–1300.

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of personality and social psychology* (Vol. 2, pp. 203–234). Beverly Hills, CA: Sage.

Greenberg, J. R. & Mitchell, S. A. (1983). *Object relations in a psychoanalytic theory*. Cambridge, MA: Havard University Press.

Guisinger, J. S., & Blatt, S. J. (1994). Individuality and relatedness. *American Psychologist, 49,* 104–111.

Gutmann, D. (1987). *Reclaimed powers: Toward a new psychology of men and women in later life*. New York: Basic Books.

Haan, N., Millsap, R., & Hartka, E. (1986). As time goes by: Change and stability in personality over fifty years. *Psychology and Aging, 1,* 220–232.

Hareven, T. K. (1978). The last stage: Historical adulthood and old age. In E. H. Erikson (Ed.), *Adulthood* (pp. 201–215) New York: Norton.

Harter, S., & Monsour, A. (1992). Developmental analysis of conflict caused by opposing attributes in the adoloescent self-portrait. *Developmental Psychology, 28* (2), 251–260.

Hauser, S. T. (1976). Loevinger's model and measure of ego development: A critical review. *Psychological Bulletin, 83,* 928–955.

Hayslip, B., Jr. (1989). Alternative mechanisms for improvements in fluid ability performance among older adults. *Psychology and Aging, 4,* 122–124.

Hayslip, B., Jr., Maloy, R. M., & Kohl, R. (1995). Long-term efficacy of fluid ability interventions with older adults. *Journal of Gerontology: Psychological Sciences, 50B,* P141–P149.

Hazan, C., & Shaver, P. (1994). Attachment as an organizational framework for research on close relationships. *Psychological Inquiry, 5,* 1–22.

Helson, R. (1967). Personality characteristics and developmental history of creative college women. *Genetic Psychology Monographs, 76,* 205–256.

Helson, R. (1993). Comparing longitudinal studies of adult development: Toward a paradigm of tension between stability and change. In D. C. Funder, R. D. Parke, C. Tomlinson-Keasey, & K. Widaman (Eds.), *Studying lives through time: Personality and development* (pp. 93–119). Washington, DC: American Psychological Association.

Helson, R., Mitchell, V., & Moane, G. (1984). Personality and patterns of adherence and nonadherence to the social clock. *Journal of Personality and Social Psychology, 46,* 1079–1096.

Helson, R., & Moane, G. (1987). Personality change in women from college to midlife. *Journal of Personality and Social Psychology, 53,* 176–186.

Helson, R., & Stewart, A. J. (1994). Personality change in adulthood. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change*? (pp. 201–225). Washington, DC: American Psychological Association.

Helson, R., Stewart, A. J., & Ostrove, J. (1995). Identity in three cohorts of midlife women. *Journal of Personality and Social Psychology, 69,* 544–557.

Helson, R., & Wink, P. (1992). Personality change in women from the early 40s to the early 50s. *Psychology and Aging, 7,* 46–55.

Hertzog, C., Schaie, K. W., & Gribbin, K. (1978). Cardiovascular disease and changes in intellectual functioning from middle to old age. *Journal of Gerontology, 33,* 872–883.

Hofland, B. F., Willis, S. L., & Baltes, P. B. (1981). Fluid intelligence performance in the elderly: Intraindividual variability and conditions of assessment. *Journal of Educational Psychology, 73,* 573–586.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology, 57,* 253–270.

Horn, J. L., & Hofer, S. M. (1992). Major abilities and development in the adult period. In R. J. Sternberg & C. A. Berg (Eds.), *Intellectual development* (pp. 44–99). New York: Cambridge University Press.

Inhelder, B. & Piaget J. (1975). *The growth of logical thinking: From childhood to adolescence*. New York: Basic Books.

Jaques, E. (1965). Death and the mid-life crisis. *International Journal of Psychoanalysis, 46,* 502–514.

Jepson, K. & Labouvie-Vief, G. (1992). Symbolic processing in youth and elders. In R. West & J. Sinnott (Eds.), *Everyday memory and aging: Current research and methodology*. New York: Springer.

John, O. P. (1990). The "Big-Five' factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66–100). New York: Guilford Press.

Johns, R. S., Mednick, S. A., & Shulsinger, F. (1982). Teacher reports as a prediction of schizophrenia and borderline schizophrenia: A Bayesian decision analysis. *Journal of Abnormal Psychology, 91,* 399–413.

Jones, C. J., & Meredith, W. (1996). Patterns of personality change across the life-span. *Psychology and Aging, 11,* 57–65.

Jung, C. G. (1933). *Modern man in search of a soul* (Trans. W. S. Dell & C. F. Baynes). New York: Harcourt Brace & World.

Kagan, J., & Moss, H. A. (1962). *Birth to maturity*. New York: Wiley.

Kaufman, D. R., & Richardson, B. L. (1982). *Achievement* *and women: Challenging the assumptions*. New York: Free Press.

Kegan, R. (1982). *The evolving self.* Cambridge, MA: Harvard University Press.

Kernberg, O. F. (1980). *Internal world and external reality*. New York: Aronson.

Kerr, B. A. (1985). *Smart girls, gifted women*. Columbus, OH: Ohio Publishing.

Kitchener, K. S. & Brenner, H. G. (1990). Wisdom and reflective judgement: Knowing in the face of uncertainty. In R. Sternberg (Ed.), *Wisdom: Its nature, origins and development* (pp. 212–229). Cambridge, UK: Cambridge University Press.

Kitchener, K. S., & King, P. M. (1981). Reflective judgment: Concepts of justification and their relationship to age and education. *Journal of Applied Developmental Psychology, 2,* 89–116.

Kitchener, K. S., King, P. M., Wood, P. K., & Davison, M. L. (1989). Sequentiality and consistency in the development of reflective judgment: A six-year longitudinal study. *Journal of Applied Developmental Psychology, 10,* 73–95.

Kliegl, R., Smith, J., & Baltes, P. B. (1989). Testing-the-limits and the study of adult age differences in cognitive plasticity of a mnemonic skill. *Developmental Psychology, 25,* 247–256.

Kogan, N. (1990). Personality and aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 330–346). San Diego, CA: Academic Press.

Kohlberg, L. (1984). *Essays on moral development: Vol. 2. The psychology of moral development*. San Francisco, CA: Harper & Row.

Kohn, M. L., & Schooler, C. (1973). Occupational experience and psychological functioning: An assessment of reciprocal effects. *American Sociological Review, 38,* 97–118.

Kohut, H. (1977). *The restoration of the self.* New York: International Universities Press.

Kramer, D. A. (1989). Development of an awareness of contradiction across the life-span and the question of post-formal operations. In M. L. Commons, J.-D. Sinnott, F. A. Richards, & C. Armon (Eds.), *Adult development: Vol. 1. Comparisons and applications of developmental models* (pp. 133–160). New York: Praeger.

Kramer, D. A., & Woodruff, D. S. (1986). Relativistic and dialectical thought in three adult age groups. *Human Development, 29,* 280–290.

Labouvie-Vief, G. (1977). Adult cognitive development: In search of alternative interpretations. *Merrill-Palmer Quarterly, 23*, 227–263.

Labouvie-Vief, G. (1980). Beyond formal operations: Uses and limits of pure logic in life-span development. *Human Development, 23*, 141–161.

Labouvie-Vief, G. (1981). Re-active and pro-active aspects of constructivism: A life-span model. In R. M. Lerner & N. A. Busch-Rossnagel (Eds.), *Individuals as producers of their development: A life-span perspective* (pp. 197–320). New York: Academic Press.

Labouvie-Vief, G. (1982a). Dynamic development and mature autonomy: A theoretical prologue. *Human Development, 25*, 161–191.

Labouvie-Vief, G. (1982b). Growth and aging in life-span perspective. *Human Development, 25*, 65–79.

Labouvie-Vief, G. (1985). Intelligence and cognition. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 500–530). New York: Van Nostrand Reinhold.

Labouvie-Vief, G. (1990). Wisdom as integrated thought: Historical and developmental perspectives. In R. J. Sternberg (Ed.), *Wisdom: Its nature, origins, and development* (pp. 52–83). Cambridge, UK: Cambridge University Press.

Labouvie-Vief, G. (1992). A neo-Piagetian perspective on adult cognitive development. In R. L. Sternberg & C. A Berg (Eds.), *Intellectual development* (pp. 197–228). New York: Cambridge University Press.

Labouvie-Vief, G. (1994). *Psyche and Eros: Mind and gender in the life course*. New York: Cambridge University Press.

Labouvie-Vief, G. (in press). Emotional integration in later life. In P. M. Lawton & K. W. Schaie (Eds.), *Emotions in adult development*. New York: Springer.

Labouvie-Vief, G., & Chandler, M. J. (1978). Cognitive development and life-span developmental theory: Idealistic vs. contextual perspectives. In P. B. Baltes (Ed.), *Life-span development and behavior*. New York: Academic Press.

Labouvie-Vief, G., Chiodo, L. M., Goguen, L. A., Diehl, M., & Orwoll, L. (1995). Representations of self across the life-span. *Psychology and Aging, 10*, 404–415.

Labouvie-Vief, G., DeVoe, M., & Bulka, D. (1989). Speaking about feelings: Conceptions of emotion across the life-span. *Psychology and Aging, 4*, 425–437.

Labouvie-Vief, G, & Diehl, M. (1996). *Cognitive complexity and maturity of coping and defense: Two dimensions of cognitive-emotional development*. Manuscript submitted for publication.

Labouvie-Vief, G., Diehl, M., Chiodo, L. M., & Coyle, N. (1995). Representations of self and parents across the life-span. *Journal of Adult Development, 2*, 207–222.

Labouvie-Vief, G., & Gonda, J. N. (1976). Cognitive strategy training and intellectual performance in the elderly. *Journal of Gerontology, 31*, 327–332.

Labouvie-Vief, G., Hakim-Larson, J., DeVoe, M., & Scholeberlein, S. (1989). Emotions and self-regulation: A life-span view. *Human Development, 32*, 279–299.

Labouvie-Vief, G., Hakim-Larson, J., & Hobart, C. J. (1987). Age, ego level, and the life-span development of coping and defense processes. *Psychology and Aging, 2*, 286–293.

Labouvie-Vief, G., & Schell, D. (1982). Learning and memory in later life: A developmental perspective. In B. Wolman & G. Stricker (Eds.), *Handbook of developmental psychology* (pp. 829–846). Englewood Cliffs, NJ: Prentice-Hall.

Lachman, M. E., Weaver, S. L., Bandura, M., Elliott, E., & Lewkowicz, C. J. (1992). Improving memory and control beliefs through cognitive restructuring and self-generated strategies. *Journal of Gerontology: Psychological Sciences, 47*, P293–P299.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.

Leon, G. R., Gillum, B., Gillum, R., & Gouze, M. (1979). Personality stability and change over a 30-year period: Middle to old age. *Journal of Consulting and Clinical Psychology, 47*, 517–524.

Levinson, D. J., Darrow, C. N., Klein, E. B., Levinson, M. H., & McKee, B. (1978). *The seasons of a man's life*. New York: Ballantine.

Loevinger, J. (1976). *Ego development: Conceptions and theories.* San Francisco, CA: Jossey-Bass.

Loevinger, J. & Wessler, R. (1970). *Measuring ego development*. San Francisco, CA: Jossey-Bass.

Lyotard, J. F. (1987). The postmodern condition. In K. Baynes, J. Bohman, & T. McCarthy (Eds.), *After philosophy: End or transformation* (pp. 73–94). Cambridge, MA: MIT Press.

Magnusson, D., & Törestad, B. (1993). A holistic view of personality: A model revisited. *Annual Review of Psychology, 44,* 427–452.

Martin, V., & Nivens, M. K. (1987). The attritional response of males and females to noncontigent feedback. *Sex Roles, 16,* 453–462.

McAdams, D. P. (1985). *Power, intimacy, and the life story: Personological inquiries into identity*. New York: Dorsey Press.

McAdams, D. P. (1992). The five-factor model in personality: A critical appraisal. *Journal of Personality, 60,* 329–361.

McAdams, D. P. (1994). Can personality change? Levels of stability and growth in personality across the life-span. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change*? (pp. 299–313). Washington, DC: American Psychological Association.

McCrae, R. R., & Costa, P. T., Jr. (1990). *Personality in adulthood.* New York: Guilford Press.

McGrath, E., Keita, G. P., Strickland, B. R., & Russo, N. F. (1990). *Women and depression: Risk factors and treatment issues*. Washington, DC: American Psychological Association.

Mead, M. (1970). *Culture and committment*. Garden City, NJ: Doubleday.

Moss, H. A., & Susman, E. J. (1980). Longitudinal study of personality development. In O. G. Brim, Jr., & J. Kagan (Eds.), *Constancy and change in human development* (pp. 530–595). Cambridge, MA: Harvard University Press.

Neumann, E. (1954). *The origins and history of consciousness*. Princeton, NJ: Princeton University Press.

Nolen-Hoeksema, S., & Girgus, J. S. (1994). The emergence of gender differences in depression during adolescence. *Psychological Bulletin, 115,* 424–443.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality rating. *Journal of Abnormal and Social Psychology, 66,* 574–583.

Pascual-Leone, J. (1984). Attention, dialectic, and mental effort: Toward an organismic theory of life stages. In M. L. Commons, F. A. Richards, & C. Armon (Eds.). *Beyond formal operations: Late adolescent and adult cognitive development*. New York: Praeger.

Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years*. New York: Holt, Rinehart & Winston.

Piaget, J. (1980). *Experiments in contradiction* (Trans. D. Coleman). Chicago: University of Chicago Press.

Poon, L. W., Rubin, D. C., & Wilson, B. A. (Eds.). (1989). *Everyday cognition in adulthood and late life*. New York: Cambridge University Press.

Puckett, J. M., & Reese, H. W. (Eds.). (1993). *Mechanisms of everyday cognition*. Hillsdale, NJ: Erlbaum.

Redmore, C. D., & Loevinger, J. (1979). Ego development in adolescence: Longitudinal studies. *Journal of Youth and Adolescence, 8,* 1–20.

Reese, H., & Overton, W. (1970). Models of development and theories of development. In L. R. Goulet & P. B. Baltes (Eds.), *Life-span development psychology* (pp. 115–145). New York: Academic Press.

Reich, K. H. (1990). *Commonalities and differences of Piagetian operations and complementary reasoning: A conceptual model and its support by empirical data*. Paper presented at the meeting of the Jean Piaget Society, Philadelphia.

Riegel, K. F. (1973). Dialectical operations: The final period of cognitive development. *Human Development, 16,* 346–370.

Riegel, K. F. (1976). The dialectics of human development. *American Psychologist, 13,* 689–700.

Roberts, T. (1991). Gender and influence of evaluations on self-assessments in achievement settings. *Sex Roles, 32,* 297–308.

Roberts, G., Block, J. H., & Block, J. (1984). Continuity and change in parents' child-rearing practices. *Child Development, 55,* 586–597.

Roberts, T., & Nolen-Hoeksema, S. (1989). Sex differences in reactions to evaluative feedback. *Sex Roles, 21,* 725–747.

Rosenberg, S. D., & Farrell, M. P. (1976). Identity and crisis in middle-aged men. *International Journal of Aging and Human Development, 7,* 153–170.

Salthouse, T. A. (1990). Cognitive competence and expertise in aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 310–319). San Diego, CA: Academic Press.

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Erlbaum.

Sampson, E. E. (1977). Psychology and the American ideal. *Journal of Personality and Social Psychology, 35,* 767–782.

Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin, 64,* 91–107.

Schaie, K. W. (1983). What can we learn from the longitudinal study of adult psychological development? In K. W. Schaie (Ed.), *Longitudinal studies of adult psychological development* (pp. 1–19). New York: Guilford Press.

Schaie, K. W. (1984). Midlife influences upon intellectual functioning in old age. *International Journal of Behavioral Development, 7,* 463–478.

Schaie, K. W. (1986). Beyond calendar definitions of age, time, and cohort: The general developmental model revisited. *Developmental Review, 6,* 252–277.

Schaie, K. W. (1994). The course of adult intellectual development. *American Psychologist, 49,* 304–313.

Schaie, K. W. (1996). *Intellectual development in adulthood: The Seattle Longitudinal Study*. New York: Cambridge University Press.

Schaie, K. W., & Willis, S. L. (1986). Can intellectual decline in the elderly be reversed? *Developmental Psychology, 22,* 223–232.

Schaie, K. W., & Willis, S. L. (1991). Adult personality and psychomotor performance: Cross-sectional and longitudinal analyses. *Journal of Gerontology: Psychological Sciences, 46,* P275–P284.

Scheibel, A. B. (1996). Structural and functional changes in the aging brain. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 105–128). San Diego, CA: Academic Press.

Schmitz-Scherzer, R., & Thomae, H. (1983). Constancy and change of behavior in old age: Findings from the Bonn Longitudinal Study on Aging. In K. W. Schaie (Ed.), *Longitudinal studies of adult psychological development* (pp. 191–221). New York: Guilford Press.

Schooler, C. (1987). Psychological effects of complex environments during the life-span: A review and theory. In C. Schooler & K. W. Schaie (Eds.), *Cognitive functioning and social structure over the life course* (pp. 24–49). Norwood, NJ: Ablex.

Schultz, R. (1996). Experiences of affects and emotions over the life course. In K. W. Schaie (Chair), *Cognitive and affective aspects of emotion*. Symposium conducted at the 1996 Emotion and Adult Development Conference, University Park, PA.

Selman, R. L. (1980). *The growth of interpersonal understanding: Developmental and clinical analyses*. New York: Academic Press.

Siegler, I. C., George, L. K., & Okun, M. A. (1979). Cross-sequential analysis of adult personality. *Developmental Psychology, 15,* 350–351.

Siegler, R. S., & Richards, D. D. (1982). The development of intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 897–971). New York: Cambridge University Press.

Sinnott, J. D. (1989). Life-span relativistic postformal thought: Methodology and data from everyday problem solving studies. In M. L. Commons, J. D. Sinnott, F. A. Richards, & C. Armon (Eds.), *Adult development: Vol. 1. Comparisons and applications of developmental models* (pp. 239–278). New York: Praeger.

Skolnick, A. (1986). Early attachment and personal relationships across the life course. In P. B. Baltes, D. Featherman, & R. M. Lerner (Eds.), *Life-span develop-*

ment and behavior (Vol. 7, pp. 173–206). Hillsdale, NJ: Erlbaum.

Staudinger, U. M., & Baltes, P. B. (1994). The psychology of wisdom. In R. J. Sternberg (Ed.), *Encyclopedia of intelligence* (pp. 1143–1152). New York: Macmillan.

Staudinger, U. M., & Baltes, P. B. (1996). Interactive minds: A facilitative setting for wisdom-related performance? *Journal of Personality and Social Psychology, 71,* 746–762.

Staudinger, U. M., Marsiske, M., & Baltes, P. B. (1995). Resilience and reserve capacity in later adulthood: Potentials and limits of development across the life-span. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Vol. 2. Risk, disorder, and adaptation* (pp. 801–847). New York: Wiley.

Sternberg, R. J. (1990), *Wisdom: Its nature, origins, and development*. Cambridge, UK: Cambridge University Press.

Stewart, A. J. (1996, August). *Personality in middle age: Gender, history and mid-course corrections*. Murray Award Lecture presented at the American Psychological Association, Toronto, Canada.

Symonds, P. M. (1961). *From adolescent to adult*. New York: Columbia University Press.

Terman, L. M., & Oden, M. (1959). *Genetic Studies of Genius: Vol. 5. The gifted group at mid-life*. Stanford, CA: University of California Press.

Thomas, A., & Chess, S. (1986). The New York Longitudinal Study: From infancy to early adult life. In R. Plomin & J. Dunn (Eds.), *The study of temperament: Changes, continuities, and challenges* (pp. 39–52). Hillsdale, NJ: Erlbaum.

Tomlinson-Keasey, C., & Blurton, E. U. (1992). Gifted women's lives: Aspirations, achievements, and personal adjustment. In J. Carlson (Ed.), *Cognition and educational practice: An international perspective* (pp. 151–176). Greenwich, CT: JAI Press.

Tupes, E. C., & Christal, R. E. (1992). Recurrent personality factors based on trait ratings. *Journal of Personality, 60,* 225–251. (Reprinted from USAF ASD Tech. Rep. 61–97, 1961, Lackland, Air Force Base, TX: US Air Force.)

Vaillant, G. E. (1974). Natural history of male psychological health: II. Some antecedents of healthy adult adjustment. *Archives of General Psychiatry, 31,* 15–22.

Vaillant, G. E. (1993). *The wisdom of the ego*. Cambridge MA: Harvard University Press.

Whitmont, E. C. (1969). *The symbolic quest: Basic concepts of analytical psychology*. New York: Harper & Row.

Willis, S. L., & Schaie, K. W. (1986). Practical intelligence in later adulthood. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 236–268). New York: Cambridge University Press.

Willis, S. L., & Schaie, K. W. (1993). Everyday cognition: Taxonomic and methodological considerations. In J. M Puckett & H. W. Reese (Eds.), *Mechanisms of everyday cognition* (pp. 33–53). Hillsdale, NJ: Erlbaum.

Winter, D. G. (1996). *Personality: Analysis and interpretation of lives*. New York: McGraw-Hill.

Wood, P. K. (1983). Inquiring systems and problem structure: Implications for cognitive development. *Human Development, 26,* 249–265.

Yee, D. K., & Eccles, J. S. (1988). Parent perceptions and attributions for children's math achievement. *Sex Roles, 19,* 317–333.

York, K. L., & John, O. P. (1992). The four faces of Eve: A typological analysis of women's personality at midlife. *Journal of Personality and Social Psychology, 63,* 494–508.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.10
Social Psychological Foundations of Clinical Psychology

JOHN H. HARVEY and JULIA OMARZU
University of Iowa, Iowa City, IA, USA

1.10.1 INTRODUCTION: THE HISTORY OF THE SOCIAL–CLINICAL PSYCHOLOGY INTERFACE 298

1.10.2 SOCIAL PERCEPTION AND ATTRIBUTION 298
- *1.10.2.1 An Essential Base* 298
- *1.10.2.2 Basic Ideas and Theories* 299
  - *1.10.2.2.1 Heider* 299
  - *1.10.2.2.2 Jones and Davis* 300
  - *1.10.2.2.3 Actor–observer hypothesis and "fundamental attribution error"* 300
  - *1.10.2.2.4 Social perception and prejudice* 301
  - *1.10.2.2.5 Kelley* 301
  - *1.10.2.2.6 Bem and Schachter* 301
- *1.10.2.3 More Recent Applications of Attribution Theory* 302
  - *1.10.2.3.1 Learned helplessness and depression* 302
  - *1.10.2.3.2 Depression and correspondent inference* 302
  - *1.10.2.3.3 "Sane in insane places"* 303

1.10.3 SOCIAL INFLUENCE AND PERSUASION 304
- *1.10.3.1 Theories of Social Influence* 304
  - *1.10.3.1.1 Conformity* 304
  - *1.10.3.1.2 Group influences* 305
  - *1.10.3.1.3 Reducing the effects of negative social influence* 306
- *1.10.3.2 Theories of Persuasion* 306
  - *1.10.3.2.1 Cognitive dissonance* 306
  - *1.10.3.2.2 Elaboration likelihood* 307
- *1.10.3.3 Resisting Persuasion* 308

1.10.4 INTERPERSONAL ATTRACTION AND CLOSE RELATIONSHIPS 309
- *1.10.4.1 Forming Relationships* 309
- *1.10.4.2 Maintaining Close Relationships* 310
- *1.10.4.3 Dissolving Close Relationships* 311
- *1.10.4.4 Grieving and Recovering from the Loss of Close Relationships* 312
  - *1.10.4.4.1 The account* 312
  - *1.10.4.4.2 Shattered assumptions* 313

1.10.5 ALTRUISM AND HELPING BEHAVIOR 314
- *1.10.5.1 Psychology's Role in Promoting Human Selfishness* 314
- *1.10.5.2 General Principles of Altruism and Helping Behavior* 315

1.10.6 CONCLUSION: SYNTHESIS OF THE "SOCIAL ANIMAL IN A COMPLEX WORLD" 317

1.10.7 REFERENCES 318

1.10.1 INTRODUCTION: THE HISTORY OF THE SOCIAL–CLINICAL PSYCHOLOGY INTERFACE

> One achieves mental health to the extent that one becomes aware of one's interpersonal relations. (Harry Stack Sullivan)

In this chapter, we will describe areas of basic research in social psychology that provide a foundation for applications in clinical psychology. First, it is important to define social psychology. We prefer the definition provided by pioneering social psychologist Gordon Allport who defined social psychology as "an attempt to understand and explain how the thought, feeling, and behavior of individuals are influenced by the actual, imagined, or implied presence of others" (Allport, 1985, p. 3). The breadth of this definition is readily apparent when one tries to imagine human situations not involving one or more of these elements. Even in the realm of biological phenomena, social processes have been found to interact to influence behavior. From the early part of the twentieth century, this breadth of inquiry that represents social psychology has been appealing to clinical psychologists interested in understanding the bases of abnormal behavior. Social and clinical psychology have a long history of interaction that we will trace very briefly here.

In 1921, Morton Prince and Floyd Allport changed the title of the *Journal of Abnormal Psychology* that Prince had begun in 1906 to the *Journal of Abnormal and Social Psychology* (*JASP*). They began this forum to reflect the key role of interpersonal processes in psychopathology and psychotherapy. These scholars had observed the pivotal role of interpersonal relations in various aspects of psychopathology observed during World War I and also believed that the study of social behavior benefited from research on abnormal mental processes.

While *JASP* was split into the *Journal of Abnormal Psychology* and the *Journal of Personality and Social Psychology* in 1965, explicit work at the interface of clinical and social psychology has been vigorous since the late 1920s. Many scholars deepened our understanding of the role of social processes in clinical psychology during this period. Various neo-Freudians including Harry Stack Sullivan and Karen Horney in the 1920s and 1930s emphasized social processes in phenomena which psychoanalysts previously had interpreted without much reference to interpersonal relations. In the 1930s and 1940s, Kurt Lewin (see Patnoe, 1988), known as the founder of experimental social psychology, promoted his ideas about an action psychology that involved the application of theory to the solution of interpersonal and group system problems. Frank (1965) wrote about persuasion and healing. A host of workers authored or edited books that analyzed and stimulated integrations of social and clinical psychology, including: Goldstein, Heller, and Sechrest (1966), Brehm (1976), Weary and Mirels (1982), Strong and Claiborn (1982), and Leary and Miller (1986).

This high degree of interdisciplinary activity between social and clinical psychology is most vividly displayed in the 1990s via the research published in many scholarly forums. Furthermore, a journal, the *Journal of Social and Clinical Psychology*, explicitly representing this interface, was created in the 1980s (Harvey, 1983).

In the following sections, we will review broad aspects of social psychology's contributions to the foundations of clinical psychology. The following topics will be considered: social perception and attribution, including discussion of prejudice; persuasion and social influence; interpersonal attraction and close relationships; and altruism and helping behavior. We should stress that our review will be selective. There now is a vast landscape of scholarly interaction between social and clinical psychology, with work in each field influencing work in the other. For a handbook-length treatment of this landscape, the reader is referred to Snyder and Forsyth (1991).

1.10.2 SOCIAL PERCEPTION AND ATTRIBUTION

1.10.2.1 An Essential Base

The topics of social perception and attribution occupy a central role in research and practice in clinical psychology. As will be shown in a later discussion of Rosenhan's famous study "On being sane in insane places" (Rosenhan, 1973), social perception, or how we perceive other people, is central to the understanding of labeling behavior as abnormal. Rosenhan essentially showed that even educated professionals, too quickly and without cautious scrutiny of the evidence, label others and then direct behavior toward them that corresponds with the label. In so doing, he revealed a major feature of social perception, namely that it often is biased in terms of its rational, careful consideration of available information. Beyond Rosenhan's demonstration, it has been recognized since the mid-twentieth century that social perception and attribution, which refers to inferences about causality and responsibility, are pivotal in determining how people interact with one another. These processes, hence, are critical not only in the unfolding of a person's social interactions in all areas of life, but also in therapist–client interactions during therapy.

In social psychology, social perception often is treated as the general area that subsumes a number of subareas, including attribution, social cognition, impression formation, person perception, and self-perception (Weary, Stanley, & Harvey, 1989). But since the late 1950s, each of these subtopics has spawned a large, specialized literature. It has been a long time since scholars such as Heider (1958) and Kelley (1967) attempted to integrate the diverse ideas and findings. Nevertheless, their earlier analyses still have great reach regarding the development of theory and research pertaining to people's perception of others.

The concept of self is a general topic that traditionally was considered separately from social perception (James, 1890). However, self-perception, which refers to the inferences we make about ourselves, was conceived by Bem (1972) as involving a process in which people learn about themselves and their attitudes and dispositions by engaging in an attribution process not unlike what they do when they make inferences about others. In addition, Jones (1964) built on Goffman's (1959) ideas about self-presentation and developed a set of penetrating analyses of how people tactically present themselves to others in order to try to fulfill various goals of social interaction. For example, people may engage in ingratiation, self-promotion, self-martyring, and so on with the goal of social or material gain.

Early work on social perception demonstrated the remarkable degree to which people's own self-biases and motivations affect their perceptions of others. A classic example of this tendency toward bias in perception was provided by Hastorf and Cantril (1954). They showed students at Dartmouth and Princeton Universities film from a football game played in 1951 between those two teams. The game was rough, with lots of penalties especially against Dartmouth, and Princeton's star All-American selection Dick Kazmaier was injured in the rough play. Princeton won but commentary in the local papers continued for some time about the disgusting display of poor sportsmanship.

Hastorf and Cantril's procedure involved showing students at both schools the same portions of the game film and asking them to indicate on a questionnaire the rough play infractions they saw. Hastorf and Cantril found that Princeton students showed a very strong tendency to judge Dartmouth players as most responsible for the rough play. On the other hand, Dartmouth students de-emphasized their players' rough play and judged the teams as generally equally guilty of infractions. Thus, there is historical evidence for the idea that people can observe the same stimuli and come to different judgments based on their motivations or self-biasing positions. Such evidence is important both to therapists who must make judgments about the dispositions of patients/clients and also to therapists and clinical scholars interested in analyzing the distortions in people's perceptions of others in their social environment.

1.10.2.2 Basic Ideas and Theories

In describing how social perception and attribution help us understand psychopathology, labeling of mental illness, and therapeutic interactions, the early foundation analyses will be described. Heider (1958) provided the most general statement of social perception and attribution in describing his naive or common sense psychology. This psychology was concerned with how the "person on the street" systematizes information to make inferences about the social world. It articulated general schemas that people seem to learn and have in mind in inferring how personal forces such as traits and skills interact with environmental forces such as luck in determining behavior. It included analysis of how people infer: the internal–external dichotomy in much social perception, with causal inferences focused either on people's traits and states of mind (internal attribution) or on environmental forces (external attribution); the quest to find dispositional properties behind others' actions because such properties help the perceiver feel more confident about inferences than do external attributions; the causes of success and failure of their own or others' behavior in achievement contexts.

1.10.2.2.1 Heider

Heider (1976) believed that people perceive causal relations as part of their basic perception of the environment. He discussed how when the back door of his home slammed shut, he automatically began to develop inferences, starting first with the attribution that his wife had just come home from grocery shopping. In Heider's analysis, people are seen as trying to develop organized, meaningful perspectives about the numerous events that they observe practically every moment of every day, for it is only "by referring transient and variable behavior and events to relatively unchanging underlying conditions" (1958, p. 79) that individuals can predict and control their environment. Hence, Heider's answer to why he believed people are almost always, if only implicitly, engaged in attributional activity is to enhance their understanding, prediction, and control over their personal and social environments.

Heider's 1958 book *The psychology of interpersonal relations* provided germs of insights that were developed in later syntheses of attribution concepts by Jones and Davis (1965), Kelley (1967), Jones et al. (1972), and Bem (1972). We will trace these basic ideas as they relate to the foundations of clinical science in the 1990s.

1.10.2.2.2 Jones and Davis

Jones and Davis (1965) developed correspondent inference theory to trace how people search for dispositional causes behind others' actions. According to this model, people use the outcome and nature of actions (including whether the actions were seen as freely taken) to infer intention and disposition on the part of the actor. For example, when members of a jury are making a decision about the guilt of a driver involved in a traffic accident, they will examine the circumstances surrounding the accident. If the defendant-driver was witnessed to be speeding, they will consider why he or she was speeding. Freedom of action might be assessed in the context of the driver's perceived medical emergency. If such an emergency were perceived by the driver and deemed to be reasonable by the jury, this may be seen as reducing the driver's freedom to drive more responsibly; hence, the driver may be judged as less culpable. Did the driver have a history of speeding or receiving driving tickets (such evidence pertains to the driver's dispositions)? If the driver did have such a history, it may lead to a stronger judgment that he or she is culpable. Thus, in accord with this type of deductive reasoning, the jury may be construed as making a correspondent inference. A correspondent inference would be that the driver's actions were consistent with his or her dispositions.

Factors that Jones and Davis (1965) identified as affecting correspondent inferences included the extremity of an outcome and the social desirability of an action (with extreme outcomes and low social desirability both implicating correspondence regarding an actor's intentions or traits). In the traffic accident example above, if someone were killed because of the driver's speeding (extreme outcome), even the medical emergency might not ameliorate a strong judgment against the driver.

1.10.2.2.3 Actor–observer hypothesis and "fundamental attribution error"

Jones and colleagues' work on attributional processes led to many influential ideas and findings. Jones and Nisbett (1972) proposed the actor–observer (A–O) hypothesis, which predicts that actors will attribute their behavior to situational forces; whereas an observer of the same action will attribute the behavior to dispositions held by the actor. Jones and Nisbett's reasoning for this proposed A–O difference was that observers focus on actors as the figure part of their figure-ground perception of the social world. Also, observers may not have experience observing a particular actor previously. Actors, however, do have experience self-observing how they react differently according to different situations. In addition, actors perceive the situation as figural vs. the situation.

While factors such as empathy (Regan & Totten, 1975) have been found to reduce the divergence between actors and observers, this hypothesis has had a long and prosperous life in the attribution literature. In fact, qualifying conditions for the actor-observer effect continue to be studied in the 1990s (e.g., Robins, Spranca, & Mendelsohn, 1996). In general, Jones and colleagues' work has had a rippling effect through clinical psychological research dealing with the inference of traits by individuals and therapists (Harvey & Galvin, 1984). This latter work cautions therapists about the possible mistake of interring a person's behavior in terms of the person's dispositions, without careful examination of the situational context of the behavior.

Jones and colleagues' work also spurred development of the so-called "fundamental attribution error," which is a vital concept for clinical psychology and now forms the base for cognitive approaches to prejudice (Ross, 1977). According to the "fundamental attribution error," people have a tendency to emphasize dispositional attribution versus situational attribution in viewing others' behavior. Across varying types of situations, people tend to over-attribute actions to dispositions (Funder, 1987). However, does this tendency represent an "error" in the strict sense of an error (e.g., perceiving that a 200 pound person weighs 175 pounds)? No, as several scholars (e.g., Harvey, Town, & Yarkin, 1981) have argued, errors in social perception cannot be as readily established as they can in object perception (when pounds or inches, for example, can be measured with their meaning given ordinal values). Rather, what is occurring in the research on the "fundamental attribution error" is establishing what more accurately could be called a "fundamental bias toward dispositionality in attribution."

1.10.2.2.4 Social perception and prejudice

Jones and colleagues' work also has formed a base for contemporary cognitive approaches to prejudice. These investigators have analyzed correspondence in attribution mainly in situations in which people are asked to infer others' attitudes based on their behavior. A ready application to prejudice is the hypothesis that prejudiced people will make negative dispositional attributions regarding a target group's behavior (e.g., "They are on welfare because they are lazy"). Research in this vein, however, has shown more subtle prejudice effects. Hart and Morry (in press) found that white observers made different dispositional inferences about African-American speakers than they did about white speakers as a function of the speaker's nonverbal behavior, with the attributions about African-American speakers being more polarized in negative as well as positive directions. Relatedly, in an attitude attribution paradigm in which a speaker takes a certain position, Linville and Jones (1980) found that social evaluations of outgroup members (opposite sex) were more extreme than evaluations of ingroup members (same sex) in the direction being taken by the speaker. Such a finding may be readily linked to the evidence of bias in social perception demonstrated by Hastorf and Cantril (1954).

More generally, a defining characteristic of findings in cognitive social applications of attribution theory is the attribution of extreme, less complex and varied, qualities to outgroup versus ingroup members (Linville, 1982). As noted by Devine (1989), low prejudice individuals appear to be able to replace stereotype-congruent thoughts with thoughts reflecting equality, whereas high prejudice individuals do not appear to be able to inhibit stereotype-consistent thoughts (or have the motivation to do so).

1.10.2.2.5 Kelley

Kelley(1967, 1972) introduced the idea that sometimes people engage in a detailed attributional analysis not unlike the analysis of variance in statistics, and on other occasions they engage in quick deliberations about causality using priming cognitive sets that he referred to as causal schemata. His analysis was broader than Jones and Davis's model, referring both to social and self-perception as well as attribution. His conception suggested a number of hypotheses including that extreme outcomes will necessitate multiple necessary causal schemata (i.e., when something extreme happens, people tend quickly to embrace the idea that there were multiple causes for the event). Kelley's treatment also systematized many of Heider's ideas, such as the concept that behavior sometimes engulfs the field, meaning that people's behavior is so salient that it swamps out perception of situational impacts on their behavior. This idea is consistent with the "fundamental attribution error" and with Bem's (1972) integration of self-perception into attribution theory. Another major contribution by Kelley (1979) was the introduction of attribution theoretical reasoning to understanding close relationships. As shown, for example, by Orvis, Kelley, and Butler (1976), couples often diverge in their attributions about the causes of conflict in their relationships. This divergence, and whether or not it can be affected by therapy, may play a significant role in whether or not their relationship can be sustained over time.

1.10.2.2.6 Bem and Schachter

Bem (1972) and Schachter (1964) were not explicit attribution theorists. Yet their ideas have been assimilated into the attribution perspective, and each has presented concepts that have had currency in clinical science. Bem argued that people use their behavior, and the context in which it occurs, to infer their internal states, including attitudes. He suggested that they use the same information they would use in making inferences about others' internal states to make inferences about their own. This description of a backward-type attribution/perception process has been influential in showing that people do not always have clearly in mind their attitudes about social issues, nor readily understand why they are acting as they are (Nisbett & Ross, 1980). The power of looking back and seeing ourselves engaged in some line of action (e.g., a fight with our spouse) may be great in determining what we think, our next emotional state, and subsequent behavior. Although Bem's analysis is viewed as extreme in its rendition of how people make inferences about their own attitudes and feelings (Weary et al., 1989), it has had an impact on contemporary social perception work and on practitioners' emphasis on getting people to act first as a way of constructively changing attitudes and behavior (as in stopping smoking before a stop-smoking attitude is firmly in place).

Schachter's (1964) theory of emotion has been strongly linked to clinical issues such as phobias. This theory argues that perceptible physiological arousal and labeling of this arousal in accord with situational or cognitive factors are necessary for the subjective experience of emotional stress. For example, suppose a person stutters rather badly in various social interactions. In a particular situation involving public speaking, this person may perceive that he or she is aroused and attribute the arousal to fear of public-speaking. If there is no plausible, outside cause of the arousal, the individual may become even more anxious. As suggested by Schachter's students Valins and Nisbett (1972), the person may exacerbate the stuttering in this situation when no other source exists for the anxious arousal. However, if this person were induced to attribute the arousal to some other source (such as a drug taken earlier in the day), it is possible that the real arousal associated with public-speaking could be lessened by virtue of the misattribution of anxiety to the drug.

In an indirect way, this hypothesis was confirmed in a finding by Storms and Nisbett (1970) who used placebos in a study with insomniacs to induce some participants to believe that the drug would relax them, while others were led to believe that the drug would arouse them. The investigators reasoned that the participants who were naturally aroused at bedtime but who could attribute the arousal to the drug would get to sleep more quickly than would those who believed that the drug was supposed to relax them, yet as usual were aroused. This reverse placebo effect held sway in research concerned with application of attribution theory until the 1980s. From that point, however, there have been few further demonstrations. The reason probably is that there is always an ethical question associated with giving people who suffer from some ailment false information about what a supposed drug will do for them. Beyond that, the area never established the boundaries very well regarding when indirect versus direct placebo, or suggestion, effects would occur (Kellogg & Baron, 1975).

1.10.2.3 More Recent Applications of Attribution Theory

1.10.2.3.1 Learned helplessness and depression

In a very general way, attribution theory continues to be one of the most frequently applied approaches to clinical phenomena. The application is "general" because unfortunately the work sometimes does not involve much bridging between early, basic theory and contemporary theory. An example is the attributional analysis of learned helplessness and depression. Abramson, Seligman, and Teasdale (1978) took Seligman's (1975) idea of learned helplessness in dogs and added attributional components to suggest that people who are highly depressed are prone to helpless behavior because of a particular attribution style they have developed for most important areas of action. In this depressed attributional style, it is hypothesized that individuals do not attribute positive outcomes to their own actions. These outcomes are perceived as external, global, and chance-oriented factors. If outcomes are negative, however, these individuals perceive themselves as responsible for the outcomes. In effect, they are showing helplessness and the accompanying sense of hopelessness, not unlike what Seligman showed with dogs who could not escape electric shock in a shock box situation.

There is general evidence to support the value of attributional style in understanding depression. As noted originally, however, no attempt has been made to link this work to its implicit base in Heider's (1958) analysis of the conditions of action, nor to the derivative theory and research by Weiner (1974) that posited and showed evidence of how people attribute causes for success and failure in achievement contexts.

1.10.2.3.2 Depression and correspondent inference

Depression is the psychopathology that contemporary cognitive social psychologists most often have tried to understand. Weary and colleagues' program of work is illustrative. They argue that moderately depressed people (usually college students, since severely depressed or hospitalized persons have seldom been studied) are particularly sensitive to information about others that might render their worlds more understandable, predictable, and controllable. Such information has included social comparison feedback (Weary, Elbin, & Hill, 1987) and another's negative or unexpected behaviors (Weary, Jordan, & Hill, 1985). Weary and colleagues (1989) further contend that moderately depressed perceivers use a style of information processing that may be characterized as effortful, vigilant, and complex. Accordingly, Yost and Weary (1996) showed that moderately depressed persons were less likely than nondepressed persons to make correspondent inferences and more likely to process the available social information when they were not preoccupied with a cognitive overload task.

Going back to the extrapolations of Jones's ideas about correspondence to the "fundamental attribution error" and to our understanding of prejudice, Yost and Weary's research implies that given the opportunity, moderately depressed people will be less likely to jump to correspondent inferences about others. Rather, they will engage in the effortful task of analyzing situations to determine how they may have affected behavior. It is not clear what severely depressed persons would do in this situation. They may not even care. As for nondepressed persons and the negative implication of this work for their ready correspondent inferences, the major question may be whether they are motivated to be concerned about how the situation affected another's behavior, or even to be empathic with another regarding the outcome.

1.10.2.3.3 "Sane in insane places"

This was the title of a well-known study by Rosenhan (1973). In this study, eight "normal" persons gained entry to 12 mental hospitals, many of which were considered to be excellent teaching hospitals, that were located in different states on the east and west coasts of the USA. The "Patients" ranged in age from early twenties to middle age and held occupations such as psychologist, pediatrician, psychiatrist, painter, and housewife. They had no history of mental illness. To be admitted, they indicated that they had been hearing vague voices and that they felt their lives were empty and hollow. These symptoms were chosen because there was no indication at that time in the literature of what might be called existential psychoses. The pseudopatients also reported typical life courses involving both losses and joys and a variety of experiences that were fairly common for persons at their age. They then were admitted to a mental hospital with no knowledge of when they would be released. Their behavior in the hospitals was described as cooperative and friendly.

Rosenhan found that during the public "show" of sanity, when pseudopatients' histories were taken, there was no instance of detection by the admitting person(s). Nor was their possible sane condition questioned in any of the admitting reports. The length of stay ranged from seven to 52 days, with an average of 19 days. The pseudopatients were not carefully observed in any instance of hospitalization. Interestingly, 35 out of 118 patients on the admission ward did indicate suspicions about the sanity of the pseudopatients.

Rosenhan offered a number of conclusions and speculations regarding this evidence. First, he acknowledged that the tendency to call the healthy sick may be most pronounced during the early admissions stage, since there had been no opportunity to observe the pseudopatients prior to this time. He also linked the data to work of Asch (1946) in the impression formation field, which showed that people often assimilate perceptions of others around strong, central traits ("warm," "cold," "crazy") that they impute to the others. Perhaps that is what the attending physicians, nurses, and attendants were doing with the pseudopatients: assimilating an interpretation of who they were around certain labels such as "someone who hears voices and who feels empty."

Beyond the implication of the "stickiness of the psychodiagnostic label," it was clear from Rosenhan's study that patients are not very carefully observed, at least not back in those years of mental healthcare. Why? Physicians may have been too busy. Or they may have simply assumed that certain behavioral patterns would be observed if, indeed, careful observation had been done. Yet, they did not even question the pseudopatients very thoroughly. "Patient engages in writing behavior" was the daily nursing comment on one of the pseudopatients, who indicated that he never was asked what he was writing or why he was writing it. Consistent with the "fundamental attribution error," behaviors that were consistent with diagnosis, such as nervousness, were attributed to the pseudopatient's personality. Rarely did the analysis of patient behavior look toward the complex stimuli surrounding the patient. Consistent with what has been referred to as the self-fulfilling prophecy in social perception (Darley & Fazio, 1980), it was evident to Rosenhan that the typical patient received similar treatment from the staff and sometimes reacted by showing a sense of powerlessness and resignation. Rosenhan also suggested that the hierarchical, often frenzied, atmosphere of the mental hospital encouraged a depersonalization of patients by the staff.

Related to Rosenhan's study is the work of sociologists on labeling theory. Sociological social psychologist Goffman (1959) wrote cogently about the presentation of self in everyday life. In this seminal analysis, Goffman posited the similarity between human social behavior and the production of a theatrical work, with tactical stage presentations, entrances and exits, and as much going on "behind the scenes" as was apparent in the presentations to the audience. Goffman (1961) also produced an influential work entitled *Asylums* that addressed the social organization of deviance based on his observations from playing the role of an inmate in an asylum (Goffman, 1961). Central to Goffman's thinking was the view that people readily apply labels to one another, and even to themselves, and then their behavior often is consistent with those labels.

The sociologist Thomas Scheff (1966) went further than Goffman in his labeling theory as it is applied to psychiatric diagnoses. He argued that psychiatric diagnoses were merely convenient labels attached to individuals who violated conventional behavioral norms. He suggested that the stigmatizing process sometimes produced disturbed behavior. Concepts of mental illness, according to Scheff, were part of an ideology embedded in the historical and cultural present of the white middle class of Western societies. Scheff believed that the functioning of psychiatric labels was to reify and legitimate the existing social order.

A lot of progress in diagnosis and treatment has been made since the early writings on labeling theory and Rosenhan's work were published. Still, this work shows the potential insidious nature of prejudicial social perception processes in influencing mental health diagnosis and treatment. The same type of process may be found in the education and prison systems, and in any institution, wherein people are expected to progress from one level to another, but are judged as unequally able or prepared for such progress. The reader also is referred to Rosenthal & Jacobson's (1968) demonstration of the effect of labeling students according to pseudo IQ scores on their "educational progress," as a related piece of evidence about the impact of labeling.

1.10.3 SOCIAL INFLUENCE AND PERSUASION

1.10.3.1 Theories of Social Influence

Why and how are people influenced by the actions of others? An understanding of how social interactions affect subsequent behaviors can be invaluable to any clinician. Not only does research in this area help in analyzing troubling situations clients may face, but it also adds a depth of understanding to the social interaction inherent in any therapy session.

Social influence is triggered by the desire to be socially accepted and to fit into a group. Festinger (1954) proposed that people continually compare their attitudes and performance to those around them in order to verify their actions and monitor their relative social standing. Using this social comparison process, people adjust accordingly in order to gain approval or status. Social influence can alter behavior in both ambiguous situations with no clear objective answer, and in those where individuals are striving to be accurate or factually correct.

Social psychologist Turner (1991) proposes a social categorization model of social influence. Individuals identify themselves by social groups or categories, and strive to exemplify the characteristics of these categories while also maintaining individual distinctiveness. Thus we are driven to emulate others whom we feel belong to our categories, and to differentiate ourselves from those whom we see as outsiders. Because social surroundings are in constant flux, so are the social influences being brought to bear at any time. As allegiances and the saliencies of particular self-categories shift, so do the relative influential strengths of the individuals surrounding us.

There has been a strong negative connotation associated with the power of social influence. Possibly this stems from the pattern of research in this area which has studied influence primarily by investigating how social forces can increase individuals' errors in judgment or opinion. It is also possible that a Western cultural emphasis on the strength of the individual produces a distrust of any social situation which seemingly diminishes that strength. Social influences, however, are not inherently negative. While they may at times make us more susceptible to destructive norms or behaviors, they also provide the support, affirmation, and security that we, as social animals, need.

1.10.3.1.1 Conformity

Conformity consists of changing one's behavior to match a group's social norm. Sherif (1937) and Asch (1955) performed classic experimental work demonstrating that individuals confronted with a group norm will often conform to that norm, even if it is obviously incorrect. Sherif's studies utilized the optical illusion of the autokinetic effect, asking participants to judge the illusory movement of an actually stationary light. The participants' judgments converged over time, creating a group norm.

Asch, convinced that Sherif's effect was due to the ambiguous nature of the stimulus, attempted to demonstrate this using judgments of line length. One naive subject was placed in a group of confederates. Three lines of varying length were displayed to the group. The researcher then presented a fourth line which clearly matched one of the preceding three in length. The task was simple and unambiguous: choose the matching line. Asch expected to see Sherif's conformity effect disappear.

The researcher requested verbal responses, with confederates always answering prior to the naive participant. The participant thus heard all of the group's answers and knew they would hear his. On a certain percentage of the trials, the confederates all gave the same, incorrect answers. Rather than defy the group, a third of the subjects conformed to the erring confederates, also giving the incorrect responses (Asch, 1955). In postexperimental interviews, these subjects often reported being confused about the group's behavior but not wanting to "ruin" the experiment by being wrong.

Asch and his colleagues conducted many variations of this paradigm, including changing the size of the confederate group and the magnitude of the confederate "error." There were always those who conformed. The variation that did succeed in virtually eliminating conformity was the inclusion of an "ally" among the confederates who gave a nonconforming response. Just one other rebel apparently gave the subjects courage to state their true judgments.

In an interesting sociological extension of the conformity work, Jacobs and Campbell (1961) used Sherif's paradigm, cycling new naive subjects into the experiment. They provided a convincing demonstration of how an erroneous social norm can be "passed down" through generations, as the new subjects conformed to the exaggerated responses of the "older generation" of previously influenced subjects.

These studies are especially revealing in that no overt social influence was aimed at the participants. There were no attempts to persuade subjects to "go along" with the group and no explicit threats of any kind associated with disagreement. The pressure to conform, to be part of the whole, is thus a powerful one, even when a situation is straightforward. This pressure only increases when events are uncertain or confusing, as most real world situations are.

Sociologist Phillips (1974) has studied imitative suicides. Phillips's research indicates that the suicide rate increases significantly following a news story on suicide, only in the area over which the news of the first suicide has been publicized. Presumably, these additional suicides are conforming to the example of the first. Incidences of single-car accidents and plane crashes which result in fatalities also appear to increase following the reporting of a suicide. Phillips's contention is that, these, too, are sometimes "copycat" suicides masquerading as accidents (Phillips, 1979). People who are already experiencing confusion or questions about social identities or situations may be especially vulnerable to conformity pressures and react with self-destructive behavior.

1.10.3.1.2 Group influences

Group influences may be more important to individual behavior than is overtly evident. People naturally categorize themselves as belonging to one group or another, even when the groups are not clearly defined or even particularly meaningful. Assigning people to groups with meaningless labels, such as "overestimators" and "underestimators" causes individuals to align themselves as part of their "group" (Brewer & Weber, 1994). Even when subjects are told they are not yet part of the group, but are being considered for later membership, they respond to tasks as if they were already group members (Mackie & Cooper, 1984).

Once individuals identify themselves as members of a group, perceptions of self vs. others change. "Ingroup bias" develops, such that what one's own group does is seen as better, more important, and preferable to other groups' actions. Members of a majority group are perceived to be more heterogeneous in opinion and action than members of a minority group, who are all perceived to be much alike (Mullen & Hu, 1989). As the size of the minority decreases in respect to the majority, this perception of homogeneity increases, and vice versa. This appears to be true even of the perceptions of the minority group members; they themselves may perceive their group to be more alike in attitudes and behaviors than they do the members of the majority (Mullen, 1991).

These effects of group membership may partially explain what we view as "prejudiced" behavior, including the belief that members of an ethnic group are "all alike" or cannot be "told apart." This ingroup–outgroup phenomenon may also account for some of the anger that minority group members might sometimes feel toward others of "their" group who assimilate or appear to identify more closely with the majority group.

Group membership also impacts on individual decisions and attitudes. Group polarization is the tendency for the average individual group member's opinions to become more extreme after exposure to the opinions of the group as a whole. Once thought to be a unidirectional movement in favor of higher risk (and thus termed "risky shift"), psychologists now agree that polarization can occur in either direction, depending on the initial tendencies of the group. If the average member begins somewhat in favor of an idea, polarization will occur in the direction of extreme positivity. If the average member begins somewhat against, polarization will push the group in the direction of extreme negativity. There is considerable debate over whether this polarization effect is caused by persuasive arguments of the group majority, normative influence, or the desire to exemplify the group norm proposed by social categorization theory, (Turner, 1991).

Clinicians working with groups, either in a therapeutic situation or otherwise, should be aware of these group influence effects. Individuals may behave quite differently in a group than in an individual counseling session. Opinions and emotions may become more extreme in a group setting. Publicly acknowledging group membership has been used quite effectively in treatments for alcoholism or other addictions, in which admitting the problem is considered to be an important step in therapy. This strategy reminds clients dramatically of their membership. In other therapeutic contexts this may not be appropriate. Care should also be taken in labeling groups. A "cancer survivors" group is probably a better name than "cancer patient" or "cancer victim" support group.

The effects of group membership are quite powerful. A cohesive social group that promotes positive norms can be a valuable source of social support and strength for an individual. An equally cohesive group which focuses on less positive behaviors can be a destructive trap from which it can be difficult to escape.

1.10.3.1.3 Reducing the effects of negative social influence

There are methods of reducing the effect of social influence which is contributing to negative or dysfunctional behavior. Social impact theory, conceived by Bibb Latané, attempts to predict the intensity of social influence effects. The intensity, number, and proximity of sources of influence predict the amount of influence directed at the target (Latané, 1981). Thus, to reduce the amount of influence, it is necessary to increase the distance (physical or psychological) from the sources, or diffuse their effect over a larger number of targets by enlisting allies. Even one partner in dissent reduced the Asch conformity effect to almost nil (Asch, 1955).

Another strategy for combating social influence is to make prior commitments before being exposed to the sources of influence. Individuals who write down their judgments or share them with others prior to an influence manipulation are much less likely to be influenced (Deutsch & Gerard, 1955). Deciding ahead of time how to cope with social pressure gives one leverage against it.

In group situations, a minority can effect change and override some powerful norms, if key conditions exist. A minority of one is generally ineffective in changing majority opinion; it is too easy to dismiss one person. A minority of two or more, however, can sometimes convince others to defect from the majority stance, beginning a "snowball" of change (Moscovici, Lage, & Naffrechoux, 1969). The minority is most effective if it begins by being flexible, not rigid, and overdemanding. The minority group also needs to be perceived, in some way, as being "part" of the majority; it must be perceived as a new viewpoint within the majority group, rather than an outside force fighting for inclusion (Turner, 1991). Finally, and perhaps most importantly, minorities must be consistent and confident in their dissenting opinions (Moscovici et al., 1969).

The work in minority influence highlights some of the difficulties in changing attitudes. Persuading others to alter opinions or to comply with certain behaviors is especially complex when competing social influences are strong. There are facets of human social responses that can be utilized in order to facilitate the process of persuasion.

1.10.3.2 Theories of Persuasion

1.10.3.2.1 Cognitive dissonance

Consistency in behavior is something humans also strive for (Festinger, 1957). In his work on cognitive dissonance, Leon Festinger demonstrated that when there is a conflict between attitude and behavior, individuals are driven to resolve that conflict. The "dissonance" or conflict between what we may believe and what we have actually done creates great tension. Our behavior appears inconsistent with our beliefs. When a behavior has been performed and can no longer be changed, we resolve any dissonance by altering our belief about the behavior. Thus, a boring task is afterwards described as interesting, if it was done for little reward.

We feel obligated to live up to prior commitments in order to preserve that feeling of consistency in our own behavior. Once we have made a small concession to a cause or have conditionally agreed with some form of behavior, we may feel compelled to comply with future, more extreme, demands that are consistent with this previous action. Cialdini (1993) calls this the "foot-in-the-door" technique of behavioral compliance.

Commitment is an important determinant both of cognitive dissonance and of self-perception (see earlier discussion of Bem's 1972 conception). As an illustration, a friend may suggest to a person that it would be helpful if the person helps with the political campaign of a candidate in a local election. The person may make a commitment to engage in the campaigning despite having previously not liked the candidate that much. After engaging in the campaigning, the person may experience dissonance in connection with having publicly supported a candidate whom he or she did not previous like very much and try to reduce the dissonance by deciding that he or she really likes the candidate. From a self-perception position, the person simply "looks back" and sees that he or she decided to do the campaigning without much pressure. Hence, the person may conclude that he or she "must really like the candidate." Each of these conceptions may be viable in various arenas when people take actions that are counterattitudinal in nature.

The formal application of cognitive dissonance theory to clinical intervention has been proposed by Cooper and colleagues (Cooper & Axsom, 1982). Dissonance was theorized by Festinger (1957) as an aversive motivation that will stimulate actions designed to re-establish homeostasis in the organism. In dissonance theory, a person's expenditure of effort on a task may cause dissonance, which then may be reduced by enhancing cognitions supporting the value of the effort expenditure. Cooper and Axsom (1982) suggest that effort justification is a common ingredient in psychotherapy. Their thesis is that when a person expends major effort and spends quite a bit of money on therapy, therapeutic success will be facilitated.

The reasons for this enhancement include increased motivation on the part of the client and a more positive view about the likelihood of therapeutic success. Feelings of personal choice and responsibility are at the core of these reasons, and they represent essential conditions for dissonance to occur; people must feel that they had choice and were responsible for the action that led to the dissonance. Cooper and Axsom report research using experimental designs with college students interested in weight control that reveal support for this dissonance explanation of therapeutic outcome. Only participants in a condition involving high dissonance (expenditure of effort via exercise and choice in this type of effort) showed weight improvement relative to comparison conditions involving different approaches to weight control.

As noted by Cooper and Axsom, their dissonance analysis bears considerable similarity to other theoretical conceptions including Bandura's (1977) view that psychotherapy will be successful to the extent that it enhances people's feelings of personal efficacy. Also, this dissonance perspective is similar to Frank's (1974) proposal that the primary function of all psychotherapies is to restore a person's sense of mastery. Such conceptions are integrally related because of their common emphasis on the importance of perceived personal choice and responsibility in mediating thought processes over a variety of behavioral contexts.

There have been other applications of dissonance to the practice of psychotherapy. Strong (1968) argued that when a therapist provides opinions that are discrepant from the client's, the client experiences dissonance. The client may reduce this dissonance by derogating and discrediting the therapist. When the therapist is perceived as an expert, this reduces the client's ability to discredit the advice given and enhances the likelihood of attitude change in direction of the therapist's message.

1.10.3.2.2 Elaboration likelihood

The elaboration likelihood model is one of the most influential approaches in the history of work on attitude change and persuasion. This model was proposed by Petty and Cacioppo (1981, 1986), and it has greatly reinvigorated work on attitude change since the early 1980s. As we will see, this model also has been applied to applications of clinical and counseling practice.

According to the elaboration likelihood theory, there are two distinct routes to persuasion. One, called the central route, views attitude change as resulting from a person's diligent consideration of information that is central to what people feel are the true merits of an advocated position. The second, or peripheral route, views attitude change as occurring when an advocated position has been associated either with positive or negative cues, or when the person uses a simple decision rule to evaluate a persuasive message. These cues and decision rules may shape attitudes or allow a person to decide what attitudinal position to adopt without the need for engaging in extensive issue-relevant thinking. Petty and Cacioppo also admit that sometimes attitude change processes may involve a combination of central and peripheral processing.

This approach to attitude change helps us understand why people may be influenced linearly by the attractiveness of the communicator, or by the communicator's status and presumed knowledge. It also helps us understand better why people sometimes resist persuasion attempts (see discussion below), since some factor such as forewarning of persuasion has activated their thoughtful analysis of the positions being advocated. In its generality and neutrality on whether people are driven by motives such as the aversive motivation presumably involved in dissonance, this approach can subsume processes such as dissonance and self-perception since either may be involved in a particular activation of central or peripheral processing.

Since factors such as level of perceived expertise of a therapist generally are believed to be critical variables in the outcome of psychotherapy (Strupp, 1981), the elaboration likelihood model is quite relevant to the therapy process. If the person's motivation and ability to think are very low, then increasing the therapist's expertise should enhance compliance with the therapy by serving as a peripheral cue. If a person's motivation and ability to think are very high, however, then expertise should not affect attitudes or compliance since they will be based primarily on a careful evaluation of the issue-relevant arguments presented. Therapy analog studies by Heesaker, Petty, and Cacioppo (1983) have provided support for this general line of reasoning.

In addition, research that did not involve an elaboration likelihood explanation has been reinterpreted as revealing the role of this process. For example, in an analog research design, Heppner and Dixon (1978) exposed participants to an expert or an inexpert counselor who attempted to change their opinions about their problem-solving abilities. The expert counselors produced greater change on an immediate assessment and an assessment taken one week later. In a delayed expertise finding, Strong and Schmidt (1970) found that although expert and inexpert counselors did not differ in the amount of attitude change induced immediately after the very brief influence session, one week later the expert counselor had produced more change than had the inexpert counselor. The result suggests that participants may have been motivated to reconsider the information provided by the expert after the therapy session.

The elaboration likelihood model has been extensively imported into extant work on psychotherapy and counseling (e.g., Heesaker, Conner, & Prichard, 1995). It is an important model for therapy considerations in part because of the value it accords central processing. This type of processing presumably leads to more persistent change and to behavior that is more consistent with the changed attitude.

1.10.3.3 Resisting Persuasion

While other scholars were studying attitude change in the 1960s, William McGuire and his colleagues began an important research program on factors that contribute to resisting persuasion. This work is important for parents and practitioners alike, since it is critical that people learn early in life to withstand pressures to conform and/or act in ways that contradict their values. McGuire's interest in this topic was partially stimulated by the study of how some US prisoners of war had succumbed to the influence attempts of their North Korean captors during the Korean War.

More generally, McQuire's ideas relate to what threatens people and their attitudes and values. As has been shown by Brehm (1966) in his theory of psychological reactance, people cherish their freedoms—including their freedom to hold their attitudes and beliefs—and will react against attempts to limit those freedoms. But what are some of the most enduring ways that people can resist persuasion attempts?

In an early demonstration, McGuire and Papageorgis (1962) showed that forewarning an audience of an upcoming discrepant communication on an involving topic produces resistance to persuasion by stimulating counterarguments in anticipation of the message. In terms of the previously described elaboration likelihood model, such forewarning activates central processing. McGuire believed that attacks on people's attitudes and beliefs for which a person was not prepared would be the most debilitating to maintaining those attitudes and beliefs.

To test his ideas, McGuire (1964) extended his earlier line of work on forewarning by proposing a biological analogy for resistance to persuasion. Just as people can be made more resistant to a disease by giving them a small dose of the germ so that they can develop antibodies, perhaps their attitudes can be strengthened by giving them a mild form of attack of their position along with arguments that refute the attack. Presumably, this inoculation can then make them more resistant to stronger attacks on their position. From the position of elaboration likelihood, the attack on an important topic engages the central processing mode, and individuals are stimulated to think of counterarguments on their own.

McGuire (1964) also addressed the question of whether active or passive participation in an inoculation treatment is more effective when preparing people to defend their attitudes or beliefs. He found that passive defenses appear to be superior when the attacking message occurs very shortly after the defense (in which case people may respond in a rote fashion with the defenses they have just learned). However, active defenses are more effective when the attack occurs at a later point in time (e.g., after one week). As time goes on, people are motivated to develop their own case and to think up stronger arguments to support the case.

Overall, McGuire and his colleagues reported impressive evidence to support the inoculation model of resistance to persuasion. The research paradigm in this work involved attacking cultural truisms such as: "It is important to brush one's teeth after every meal." An attack on this truism might be that frequent brushing causes gum irritation and eventually gum disease. Individuals presumably have not had practice defending cultural truisms. Hence, such an attack may cause them uncertainty about the value of regular teeth brushing. A refutational defense for an attack on teeth brushing might be to admit that some gum irritation may be produced by poor brushing habits, but that on the whole brushing is essential to preventing gum disease. Unfortunately, there has been little application of McGuire's logic to attitudes and beliefs that are very important to research participants (e.g., regarding the merits of income taxes). This is because these topics do not lend themselves to investigation in an experimental situation. People often have given such topics considerable thought and hence are not readily susceptible to attack. Nevertheless, the basic principles of inoculation theory, especially when coupled with the reasoning of the elaboration likelihood model, appear to represent a valuable set of tools for practitioners and all persons involved in the process of socialization of young people. A fuller discussion of inoculation theory as it is integrated with the elaboration likelihood model can be found in Eagly and Chaiken (1993).

1.10.4 INTERPERSONAL ATTRACTION AND CLOSE RELATIONSHIPS

The area of interpersonal attraction is a traditional one in social psychology, having been the topic of early work in the 1940s and 1950s and enjoying prosperous development until about the late 1970s. This research focused on factors involved in interpersonal attraction and very early dating behavior. These factors included physical attractiveness, similarity in attitudes and personalities, and physical proximity. Interpersonal attraction as a mainline interest in social psychology gave way to work on close relationships around the late 1970s. This latter focus subsumed interpersonal attraction but emphasized intimate, more long-term relationships. It also included how people maintain, and why they sometimes end, close relationships and the consequences of relationship breakdown (or divorce in the case of married couples). The close relationships area is burgeoning in popularity and involves a melding of fields including social and clinical psychology, along with such disparate additional fields as counseling psychology, family studies, sociology, communication studies, religion, philosophy, and traumatology.

Most of the work that will be discussed in this section has derived from research focusing on heterosexual, close relationships. There has been a small amount of work comparing couples with different sexual orientations (e.g., Blumstein & Schwartz's 1983 "American Couples" study) that will be noted in the discussion. What will not be covered in this discussion are the literatures on friendship and on family relationships.

1.10.4.1 Forming Relationships

Three main factors have been identified as involved in contributing to the formation of interpersonal attraction and then possibly later close relationships between two people. These factors are physical proximity, physical attractiveness, and similarity in attitudes and personality, each of which can be described via a classic study (see Hatfield & Rapson, 1993, for a review).

Festinger, Schachter, and Back (1950) investigated proximity by studying the pattern of how people formed friendships in a new housing development. Festinger et al. found that functional proximity was critical in determining the growth of relationships. For example, people became friends with others whom they encountered doing their normal chores such as putting out the trash. Monge and Kirste (1980) suggested that what is most important about proximity is whether or not there is an opportunity to communicate.

The classic study for physical attractiveness was conducted by Walster (Hatfield), Aronson, Abrahams, and Rottman (1966). They held a college computer dance and paired men and women randomly. Each individual's personality, intelligence, social skills, and physical attractiveness were measured (coders judged the individuals' attractiveness when they showed up at the dance). Walster et al. found that only attractiveness mattered in terms of partners liking one another. Of course, this situation was that of strangers meeting strangers. Still, this finding reveals the potency early on of physical attraction in the liking and dating process.

Hatfield and Sprecher (1986) review a large body of work that shows how differently people are treated when they are perceived as attractive than when they are perceived as unattractive. There appears to be a beauty stereotype in Western societies that has changed little over time. Actress Candice Bergen said of looks:

> My looks grease the palm of life, but I resent that they're so important in our society. I resent that people are excluded ... But I'm certainly not talking about myself, because it's clear I've gotten the long end of the stick. (interview in *The Kansas City Times*, April 14, 1984, C-2)

Physical attractiveness has also been implicated in the self-fulfilling prophecy phenomenon that was discussed in the section on social perception. In a provocative demonstration, Snyder, Tanke, and Berscheid (1977) recruited men and women for a study on "the acquaintance process." Men then were given a photo and biographical information about their partners. The photo was a fake. It depicted either "a beautiful" or "an unattractive" woman. Men were asked their first impressions of the woman they saw and about whom they read. Those who believed they had been assigned a beautiful partner expected her to be sociable, poised, humorous, and socially skilled. Those who thought they had been assigned to an unattractive partner expected her to be unsociable, awkward, serious, and socially inept.

In the next phase of the study, men were asked to get acquainted with their partners via the telephone, without actually seeing them. Male expectations had a dramatic impact on the ways they talked to their partners in the telephone calls. That telephone approach, in turn, created a correspondingly major impact on the response of the women. Essentially, the women became what the men expected them to be. If the women had been perceived to be attractive, they indeed talked on the phone as if they were poised and so on. If the women had been perceived to be unattractive, their conversation showed the awkwardness with which the men already had labeled them. By studying the phone recordings, it was found that the males had elicited these divergent lines of behavior either by showing a warm, accepting tone (for the woman perceived to be attractive), or by a less inviting tone (for the woman perceived to be unattractive).

Newcomb (1961) studied the role of similarity in the development of relationships. He studied a group of college men who were given free housing in return for providing ratings of their friendship choices and attitudes. After the men had become acquainted, Newcomb found that liking was greatest among those who had similar attitudes and beliefs. This emphasis on similarity in attitudes and beliefs also was found in a series of laboratory studies by Byrne (1971). In comparison to other kinds of similarity (e.g., in personalities), similarity in attitudes and beliefs appears to be a fairly strong "glue" for continued development of a relationship. Why?

One possibility is that people like other people who are similar to them because that similarity validates their own attitudes and beliefs. It feels good to encounter others who have come to lines of logic about the world that are similar to our own. Further, people who have such similarity often have similar socioeconomic backgrounds, which tends to make them more comfortable than less similar backgrounds or which tends to elicit less conflict over time.

What value does the foregoing evidence have for the practitioner? Factors such as similarity and physical attractiveness operate pervasively in our social relations. To be ignorant of them is to be ignorant of part of the basis of our achieving or not achieving satisfying relationships with others. In addition, one of the most common concerns on the part of single persons in dealing with dilemmas in their lives is how they can meet others for closeness. The practitioner needs to be aware of these basic findings, which have relevance to strategies for meeting others. Increasingly, people are becoming activists in developing their own networks of single persons who are similar in age and who have similar interests and affiliate in order to extend their opportunities for interaction and possible closeness (Harvey, 1995).

1.10.4.2 Maintaining Close Relationships

Therapists are centrally involved in how people maintain close relationships over time. Berscheid (1994) noted in a major review of the literature on close, personal relationships that despite this burgeoning literature, we still have little evidence about the processes that are involved over time in determining the progression of close relationships. We simply do not have strong information on questions such as: What determines how close and satisfied couples will become, and what determines whether they will maintain high levels of satisfaction and closeness? This gap in our knowledge contrasts with the extensive theoretical and empirical work both on the determinants of interpersonal relationships and the determinants and effects of breakdown in such relationships.

One relevant topic is that of courtship stages and progression toward marriage (Cate, Huston, & Nesselroade, 1986; Huston, McHale, & Crouter, 1986). This work has involved following a sample of college students (who were matriculating at Pennsylvania State University in the mid- to late 1970s) from the time of becoming engaged to several years into their marriage, or after they had divorced. The investigators inquiried about these young people's feelings of satisfaction, and factors involved in their level of satisfaction. Often respondents were first asked to chart their progression toward marriage on graphs. Among the findings deriving from this program are: Significant conflict in courtship is strongly, positively correlated with major conflict developing around the one to two year mark in later marriage; for example, major early periods of ambivalence and conflict often were related to postmarriage conflict that threatened the relationship. The progression and possible deterioration of relationships are affected by numerous factors including chance and individual participants' networks of friends who either approve or disapprove of the relationship.

Harvey and Omarzu (1997) have developed a process-oriented theoretical analysis of relationship closeness and satisfaction that incorporates other major approaches to the development of these states. They refer to this analysis as "minding the close relationship" to emphasize the integral role of each partner's mind in the flow of interaction in a close relationship. Harvey and Omarzu theorize that to achieve and maintain a high level of satisfaction and closeness, a couple must mutually engage in a set of continuous activities involving: seeking knowledge about their partner and the relationship; self-disclosure; attributions of positive intentions and motivations to their partners for the relationship-relevant actions they observe their partner taking; acceptance of, and respect for, what they learn about their partners over time; and other forms of behavior that are meant to be constructive in their impact on the relationship. These last forms of behavior can include both small acts such as feet rubs and large acts such as ongoing problem solving in parenting and coordination of home and work activities. Critical to the minding process is the meaning imputed to relevant acts by the partners. Within a strong, growing relationship, most acts will be construed to reflect positively upon the partners. Further, such positivity will be based at least to a considerable degree upon sharing of information, perspective, and careful listening and analysis of information.

One of the benefits of Harvey and Omarzu's analysis is that it integrates earlier theoretical conceptions on social penetration which is based on self-disclosure (e.g., Altman & Taylor, 1973; Derlega, Metts, Petronio, & Margulis, 1993), intimacy processes (e.g., Prager, 1995, Reis & Shaver, 1988), love (e.g., Hendrick & Hendrick, 1992), the expansion of self in love (Aron & Aron, 1996), empathy (e.g., Ickes, 1993), and positivity and other illusions in close relationships (e.g., Murray, Holmes, & Griffin, 1996; Swann, 1996). We believe that this minding the close relationship approach has considerable implications for scholars and practitioners interested in relationship progression issues. Nonetheless, this approach still must be subjected to careful study that will better reveal its conceptual and applied value.

1.10.4.3 Dissolving Close Relationships

Fincham and Bradbury (1991) have conducted important research on marital distress that is based on attribution theory in social psychology. According to their contextual model of marital interaction, when one spouse behaves, the partner attends to and perceives that behavior, assigns some meaning to it, and exhibits a response. Between the spouse's behavior and the partner's response is the processing stage. Bradbury and Fincham (1992) argue that the processing stage involves reactions to the stimulus, momentary or transient thoughts, and personality factors such as mood state and attitudes or beliefs about relationships. They also argue that this approach expands the behavioral model of marital interaction beyond its strong emphasis upon behavior and satisfaction by incorporating various cognitive, affective, and personality mediators.

One of the most consistent findings from Fincham and Bradbury's research program is that distressed couples tend to make attributions that blame their partner and exonerate themselves for negative outcomes (such as major embarrassments regarding the state of the relationships occurring in public). On the other hand, nondistressed couples tend to make attributions that do not blame the other for negative outcomes, or that emphasize the joint responsibility for such outcomes (Fincham & Bradbury, 1993). These researchers have showed that to the extent that spouses make nonbenign attributions for negative outcomes at one point in time, their marital satisfaction is lower a year later.

This latter finding is important in making stronger a cause–effect role for attributions in the process of close relationship dissolution. Attributions may serve as a valuable index of relationship distress, but are they causal elements in the chain of factors involved in the ultimate breakdown of relationships? Fincham and Bradbury's (1993) results suggest that attributions at least may become contributory causal factors during the period of breakdown, if not involved from the inception of the process.

Other researchers also have implicated attributions in relationship distress and dissolution. Holtzworth-Munroe and Jacobson (1985) reported that wome tend to make attributions about the state of their relationships throughout its course. Men, however, tend to make attributions about the state of their relationships mainly when trouble has arisen. Presumably, women may be more socialized to engage in relationship caretaking than men, and hence women may be more alert and active in inferring the social psychological dynamics of relationship events. This argument is consistent with other findings and general theorizing on gender differences in dissolution (Bernard, 1982; Vaughan, 1986). It may be true, however, that as people increasingly have learned from therapy and social science's work on relationship problems, many men may have become more active processors of their relationship dynamics (Harvey, 1995).

The idea of attributional divergence has been prominent in the application of attribution theory to close relationship distress and dissolution. Orvis et al. (1976) and Harvey, Wells, and Alvarez (1978) reported that couples experiencing major distress not only made different attributions about the causes of their distress, but also they could not predict their partners' attributions. Such findings suggest that distressed couples often have distanced themselves from dialogue about relationship problems to the point that they are not cognizant of each other's explanations for conflict events. Hill, Rubin, and Peplau (1976) reported a similar kind of attributional divergence among college dating couples who had terminated their relationships. These investigators found that there was disagreement about who left whom, with each respondent claiming responsibility for being the leaver. Hill et al. also noted that other evidence suggested it most often was the woman who had been the leaver (a finding replicated by Vaughan, 1986).

1.10.4.4 Grieving and Recovering from the Loss of Close Relationships

1.10.4.4.1 The account

In a book aimed at people who are grieving the ending of their close relationships, the sociologist Weiss (1975) suggested that people use accounts to help them cope with loss in marital separation and divorce. Although "the account" was only generally defined by Weiss, a definition consistent with his statement is: a storylike construction of why the relationship breakdown occurred, who was responsible, key events, and consequences. He argued for the importance of the account in a time of high distress for the individual as follows:

> The account is of major psychological importance to the separated, not only because it settles the issue of who was responsible for what, but also because it imposes on the confused marital events a beginning, middle, and end and so organizes the events in a conceptually manageable unity. Once understood in this way, the events can be dealt with: They can be seen as outcomes of identifiable causes and, eventually, can be seen as past, over and external to the individual's present self. (Weiss, 1975, p. 15).

Weiss used narrative evidence collected from a group of persons attending "Parents without Partners" meetings to evaluate the role of the account in adjustment to marital dissolution. He found that those persons who had well-developed accounts managed the often convoluted and debilitating period of separation prior to formal divorce much better than did those who did not have very well-developed accounts. The latter took much longer to recognize that the relationship had ended and that they needed to grieve and move on. Weiss said about this group of persons who do not readily begin account making, "Those who cannot construct accounts sometimes feel that their perplexity keeps them from detaching themselves from the distressing experiences. They say, 'If only I knew what happened, if only I could understand why'" (p. 15).

Weiss' work has considerable applicability for therapists concerned with helping persons in the throes of separation develop their understandings or stories. As Weiss suggested, an account is not necessarily an objective and impartial interpretation and description. In fact, it may be highly biased and distorted. Nevertheless, these constructions may have powerful psychological truth for the individual.

Weiss' seminal observations had considerable influence on the development of theory and research about how people cope with and recover from major losses including close relationship dissolution. Harvey, Weber, Galvin, Huszti, and Garnick (1986) refined Weiss' logic of the account to define account as a mental construction containing explanations and descriptive material for events and states of being of self, other, and the world. In this conception, account making involves both private and public activities. Harvey and colleagues argued that there were several dynamics of account making, including: search for understanding and meaning associated with a loss, achieving completion regarding a loss, achieving catharsis in venting about the loss, presenting themselves in a certain light relative to responsibility for the loss, self-esteem enhancement, and persuading others about the nature of and reasons for the loss. Several of these functions are inherently social activities (e.g., persuasion of others), and later work focused on the importance of the social interaction component of account making.

This extension of account making and accounts logic was amplified by Weber, Harvey, and Stanley (1987), Harvey, Weber, and Orbuch

(1990), and Weber and Harvey (1994) into a model of the role of account making and accounts in people's reactions to major stressors and loss events in general. This model also represented a revision of Horowitz's (1986) stage model of how people react to stress. In this model, account making is the process of working on an account, which is the product of account making. It starts to occur somewhere in a sequence that often involves these phases (that may occur in this order, or in a less linear fashion): a feeling of being numbed by the loss; an outcry about the loss; denial; intrusive imagery; eventually private work at constructing a story of what happened and why; public confiding in a close other (or others) who listens and possibly offers reactions; a feeling of working through the account, acceptance of the loss, and completion of the major grieving; learning from the loss; and a major element of identity change associated with the loss and this processing (e.g., the person no longer sees herself as "John Smith's wife," but now as "John Smith's widow who will go on alone and make a new life for herself"). In studies based on this model, the empathic response of a dependable confidant to a person's loss has been found to be vital in the recovery process (e.g., Harvey, Orbuch, Chwalisz, & Garwood, 1991).

Appreciation of work on the role of account making in relationship conflict and dissolution may be beneficial to practitioners in several ways. It is important to know what is the accounts-narrative truth to a client. While recognizing that a client's story may represent only one of many interpretations of an event or situation, a practitioner can appreciate that the client still may be acting on the basis of this interpretation. Further, the interpretation may need to be challenged. People may be open to challenge of their story and may learn from the feedback and questions posed by a therapist. A practitioner also can help a client recognize the value of developing a story and the functions of story development and the confiding of them to close others. As psychiatrist Robert Coles (1989) has cogently argued, a person's narrative and willingness to share it should be respected as a universal gift.

1.10.4.4.2 Shattered assumptions

Janoff-Bulman's (1992) work on shattered assumptions provides an invaluable perspective on adaptation to loss. She argues that the psychological disequilibrium that ensues from major losses such as divorce or death of a spouse or loved one often stems from the shattering of our fundamental assumptions about the world. These assumptions include that our world is secure and benevolent. Of course, over the course of living many humans discover that such assumptions are not tenable. Like Frankl (1956), Janoff-Bulman suggests that even more basic is people's quest for and assumption of meaning and personal control in their lives. Janoff-Bulman contends that the victim's seemingly paradoxical and maladaptive response such as self-blame for a traumatic event is in fact a strategy, possibly unconscious, to restore the illusions of comprehension and control.

Janoff-Bulman also argues that the wide differences in how people respond to traumas such as death of a loved one are due in part to the nature of the event, and which assumptions are most threatened. Survivors of a human-induced trauma, for example, are posited to view themselves more negatively than are survivors of a natural disease process. The idea is that intentional harm carried out by a human perpetrator has a negative impact on the self-worth of survivors, even if this reasoning is irrational. It is an amplification on the "Why me?" question survivors often pose, but with special pertinence to the human-caused trauma.

One of the major contributions for practitioners deriving from Janoff-Bulman's work is the analysis of major tasks with which survivors are confronted in order to adapt to their loss and begin to heal. Like the account-making logic, Janoff-Bulman emphasizes the reframing of interpretations that is required in the recovery process. Included is a rebuilding of a victim's basic assumptions about the world and self. Somehow the survivor needs to integrate the loss into new schemas of understanding that make the loss event less frightening (which may be associated with intrusive, debilitating images and thoughts, as was suggested in the account-making model). The integration may involve reprioritizing of values and goals in living—lessons learned from the loss—and trying to find redeeming value in the suffering (as in the generative act of giving back to others who may suffer similar plights).

Beyond cognition, Janoff-Bulman contends that victims must not be passive, but active in confronting loss. Simple activity such as gardening, caring for friends and pets, and work on hobbies may help. Support groups and close, loving friends may help people "get out of the house" and begin telling their story to others, or if that is too difficult, listening to others' stories. Also, similar to the account-making model, Janoff-Bulman emphasizes the role of close, caring others in a victim's recovery. She indicates that there is considerable evidence that people's adaptation to loss is most often facilitated rather than impeded by the presence of close others in their lives. Finally, she warns of others' tendencies to blame the victim. Lerner (1980) has theorized a "just world hypothesis" that suggests people believe others deserve their suffering and eventually get what they deserve in life. The tendency to blame victims is central to the phenomenon depicted by this hypothesis.

As Janoff-Bulman notes, in the stress and trauma area the emphasis on people's interpretations and assumptions devolves principally from Lazarus's (1966) pioneering conceptual work and research on how people's interpretations, or appraisals, mediate their reactions to stress. Inherent in this reasoning is that the person has the capacity after a major stressor to reassess interpretations and assumptions. Some losses are so profound that they severely test this capacity.

Examples of losses that may be too overwhelming for any ready cognitive adaptation include: wartime combat veterans' experience of personal terror and the deaths of close friends (e.g., see Shay's discussion of Vietnam veterans' reactions in *Achilles in Vietnam*, 1994); individuals' reactions to the untimely and deeply, personally devastating loss of a spouse (e.g., Lewis' powerful book of grieving "notes" in *A grief observed*, 1961); and the horrific experiences of World War II concentration camp survivors who lost close loved ones to brutality and lost self-respect in bearing degrading, inhuman conditions (e.g., see the powerful stories told in Langer's *Holocaust testimonies*, 1991).

These works help the practitioner define the boundaries of possibility for intervention by interpretation, story development and telling, and reframing of assumptions. The therapeutic relevance of stories of loss and trauma also is realized in family therapy approaches that follow Frankl's (1956) ideas and embrace the empowering value of people's search for meaning in times of great stress (e.g., Lantz, 1993). Increasingly, therapists are advocating the value of people telling their own stories of loss as a means of adaptation (e.g., Howard, 1991).

How do we know when recovery from a major loss has been achieved? Weiss (1988) offers the following set of criteria:

(i) ability to give energy to everyday life;

(ii) psychological comfort, as demonstrated by freedom from pain and distress;

(iii) ability to experience gratification: to feel pleasure when desirable, hoped-for, or enriching events occur;

(iv) hopefulness regarding the future, being able to plan and care about plans; and

(v) ability to function with reasonable adequacy in social roles as spouse, parent, and member of the community.

This last criterion encompasses the idea of generativity that has been emphasized by many scholars, most notably Erikson (1963), which may involve a human's most important work both on behalf of others and oneself.

1.10.5 ALTRUISM AND HELPING BEHAVIOR

1.10.5.1 Psychology's Role in Promoting Human Selfishness

Before discussing contemporary social psychological work on altruism and helping behavior, we will consider a basic question for the field of psychology that has been raised in this area: What role has psychology played in promoting excessive self-absorption and egoism as compared to concern for others in society? Personality-social psychologists Michael and Lise Wallach (1983) wrote an important book on people's nature, that has much relevance for the area of altruism, entitled *Psychology's sanction for selfishness*. In this, the Wallachs present arguments to the effect that many of the most influential workers in the history of personality, social, and clinical psychology have committed the "error of egoism" in the positions they have advocated.

The Wallachs are suggesting that major schools of thought in psychology offer a particular view of the human that embodies an egoistic concern for self and too little concern for other people. What is this selfishness and "error of egoism" that the Wallachs describe? They do not offer precise definitions, but do offer suggestions and examples. Selfishness is equated with an "emphasis ... on narcissism, self-concern and preoccupation with 'me'" (Wallach & Wallach, 1983, p. ix), and "with achieving satisfaction for our needs and desires" (p. 18).

The Wallachs suggest that the basis for this portrait of the human as guided by self-centered motivation is found in Freud's emphasis on the importance of satisfaction of bodily needs. The Wallachs point out that within Freud's position, direct gratification of such needs can be most optimally achieved when individuals are free of prescription, restraint, and determination from the outside. Beyond Freud and neo-Freudians such as Sullivan and Horney, the Wallachs also suggest that humanistic psychologists such as Maslow and Rogers play into the selfishness theme. "When we reach Maslow and Rogers, much of society and our relations with other people is seen as obstructing our growth and the development of our potentialities. Self-determination becomes a matter of overriding significance" (p. 152).

Wallach and Wallach take on the influential social psychological analysis of interpersonal relations by Kelley and Thibaut (1978). Kelly and Thibaut's interdependence analysis is concerned with the structures of ongoing situations. The structures are defined in terms of patterns of positive and negative outcomes controlled by the joint actions of interdependent people (i.e., people such as couples in close relationships whose behavior, thoughts, and feelings are intertwined in significant ways over time). Wallach and Wallach suggest that this approach implicitly prescribes mainly a self-concern. They say about the Kelley and Thibaut logic, "Life with others is viewed as a negotiation process, using the carrots and sticks at one's disposal in pursuit of reciprocity ... Commitment to another person ... becomes a gambit or move that is tendered to induce reciprocation" (pp. 177–178).

Overall, the Wallachs believe that work such as the foregoing has furthered a cynicism about the nature of humans, with little weight given to the importance of people's commitment to others as an end in itself, rather than as a means to a personal goal. The theoretical work on altruism by social psychologists that will be described below would be viewed by the Wallachs as also embracing self-interest as the basis of most altruism, or helping of others in time of need. Readers who are interested in counterarguments about the Wallachs' position are referred to the *Journal of Social and Clinical Psychology* special issue on this topic (#1, 1985). A major theme by critics of the Wallachs' position in that set of papers is that their analysis often was too black and white and simplistic in representing complex analyses by other scholars. Regarding the Wallachs' critique of Kelley and Thibaut, Kelley and Thibaut (1985) disputed that their theory was prescriptive of a "win at all costs" rationale. They also argued:

> it will be apparent that we see no incompatibility between deriving altruism from self-interest, on the one hand, and the existence of genuine or authentic altruism, on the other ... the Wallachs make the mistake of putting our theory into the heads of the prosocial actors ... we view altruism ... as adaptations to problems of social interdependence ... it seems clear that the Wallachs regard altruism and any logic or sign of self-interest as incompatible ... we do not ... Historically and developmentally, self-interest underlies altruism and other interpersonal tendencies. Contemporaneously, prosocial behavior is sometimes promoted by thoughts of self-interest, but it also often reflects direct responsiveness to and concern for other persons' interests. (Kelly & Thibaut, 1985, pp. 29–31).

Beyond such rejoinders, many of the responding analysts indicated a belief that the Wallachs had performed a major service in focusing on psychological views of human behavior and the extent to which behavior is motivated by selfishness. Further, the Wallachs' theme about cynicism regarding psychology's role in modern society was seconded and embellished by an interdisciplinary team of scholars, Bellah, Madsen, Sullivan, Swidler, and Tipton (1985), in a well-known book entitled *Habits of the heart: Individualism and commitment in American life*.

1.10.5.2 General Principles of Altruism and Helping Behavior

There is an extensive literature on altruism and helping or prosocial behavior in social psychology. Among the questions of interest to this area are: What are the conditions that stimulate and impede people's coming to the aid of others in distress? Is there an altruistic personality, and is this personality operative across diverse situations in which helping others may bring great danger to oneself? We will deal with these questions in this section.

There are two major reasons why this literature should be of considerable interest to practicing psychologists. First, it is in the interest of society that people be socialized to help one another and to be genuinely committed to rendering help in others' times of distress when they can be of assistance. Second, as has been cogently argued in many classical psychological analyses, people can gain perspective, courage, and strength to face their own dilemmas when they extend themselves to others who are in distress. Frankl (1956), for example, used his own experience in a Nazi concentration camp to argue that people often create or find meaning in their lives to the extent that they focus on the needs of others and are sincerely concerned about their fellow humans. The term "caretaker" increasingly has been a part of psychologists' vocabulary. In general, it has been found that acting in this role makes people feel good (Williams & Clark, 1989). The general questions here concern when and how people can be effective caretakers of others, without reducing such others' own capacity for self-accomplishment or -empowerment (Steele, 1990).

Altruism theory and research were given a major stimulus by the publication of Latané and Darley's (1970) *The unresponsive bystander: Why doesn't he help?* This book won awards for its contribution to the field of psychology and has become a classic regarding social conditions for helping behavior. In the book Latané and

Darley discussed the case of Kitty Genovese who in 1964, while touring New York, was stabbed to death on a New York City sidewalk (Rosenthal, 1964). The murder was carried out over a 45-minute period and was witnessed by 38 people who heard Genovese's screams, and even observed some of the struggle, from their windows in the nearby apartments. Yet, amazingly, no one tried to intervene or otherwise help Genovese. Why such indifference?

Two principal insights were tested by Latané and Darley. One was that there was pluralistic ignorance: "No one else is reacting; it must not be an emergency after all." Two, there could have been a diffusion of responsibility: "I'm sure someone else already called the police." Latané and Darley tested these ideas in the lab and field by creating fires, accidents, thefts, fights, and seizures. For example, in one procedure, they let smoke fill the room where subject participants, either alone or with others, were completing irrelevant tasks. Would the participants take some action to notify authorities that a fire might be occurring? Or in another situation, a person in a nearby room was overheard by participants alone or in groups apparently having a seizure. Would the participants go to this person's assistance?

As with the work of Milgram (1974, see Section 1.10.6), at the present time it is not clear that such procedures would stand the test of "human subjects review committees" who evaluate research proposals for potential harm to participants. But the results were fascinating and generally have met the test of time and other workers' examination of similar questions.

Latané and Darley repeatedly found that as the number of bystanders in a situation increased, the percentage of participant subjects who helped the experimental accomplice decreased. Among those who did respond, an increase in the number of bystanders led to increased delay in taking action. Thus, Latané and Darley concluded that the "bystander effect" was a powerful phenomenon in influencing helping behavior in emergency situations. These investigators and their colleagues subsequently elaborated on what they viewed as the cognitive/decision-making steps that bystanders experience in an emergency situation. These included: recognizing and attending to the emergency, checking the interpretation that an emergency exists, deciding whether personally to assume responsibility, deciding whether to take some line of action to help, and if so, deciding what line of action is feasible.

An extensive literature on conditions of helping behavior and personality types associated with helping behavior has developed in social psychology since the late 1970s. We will only highlight some of the major directions. One area of work concerns the impact of momentary emotional states on helping behavior. Surprisingly, it has been found that depressed people are especially likely to help others in need if what they have to do to help is interesting or pleasant (Cunningham, Shaffer, Barbee, Wolff, & Kelley, 1990). This general finding has been linked to the negative state relief model (Cialdini, Kenrick, & Bauman, 1982). This model suggests that helping makes people feel good, and those in a negative mood are motivated to do anything to make themselves feel better. As for positive emotions, it often has been found that people who are made to feel momentarily happy are inclined to help strangers in need. The state of happiness has been created in a variety of ways such as by receiving cookies (Isen & Levin, 1972), by listening to a comedy album (Wilson, 1981), or even going out on a day when the sun is shining (Cunningham, 1979).

Batson and his colleagues have conducted notable research on what has been referred to as the empathy-altruism hypothesis (e.g., Batson, Duncan, Ackerman, Buckley, & Birch, 1981). This hypothesis states that when people perceive an emergency in which they have the potential to render assistance, empathy is aroused and then helping occurs. If empathy is not aroused and if people can avoid the situation, helping will not occur; if it is difficult for people to avoid the situation, helping may occur in the absence of empathy.

To test the empathy-altruism hypothesis, Batson, O'Quinn, Fultz, Vanderplas, and Isen (1983) devised an experimental procedure in which they manipulated empathic concern for the victim and ease of escape from the situation. Participants were informed that they would either have to perform a task while receiving random electric shocks while an observer watched on closed-circuit television, or be the observer while another participant received the shocks. The real subject participant always served as the observer. As the observation began, the accomplice pretended to be in pain when shock was given. But she announced that despite having had a traumatic experience with electricity as a child, she could go on with the study. The experimenter then asked the subject participant if he or she was willing to take the victim's place. Empathic concern was manipulated by the investigator informing the subject participants that the victim had values that were similar to their own (high empathy) or very dissimilar (low empathy). The investigator also manipulated the ease with which the subject participants could escape from the situation by informing them either that they could leave after watching only two shock trials (easy escape) or that they had to watch 10 such trials (difficult escape).

Batson et al. (1983) found that when empathy was low and escape was easy, subject participants tended to leave rather than engage in prosocial behavior by taking the place of the victim. When empathy was high, subject participants volunteered to take the victim's place, regardless of how easy it was to escape. It also was found that highly empathic subject participants volunteered even when they believed that neither the investigator nor the victim would ever know that they had declined to help. Hence, their helping was not due to evaluative apprehension in the situation, but rather appears to have been mainly influenced by empathy for the victim.

The above genre of research on altruism and helping behavior has focused mainly on factors affecting helping in artificial experimental contexts and in interactions among college students who were not acquainted with one another. In the real world, the opportunity to help others in need often occurs in situations in which people know one another. Is there a personality type associated with helping? This is a question that has been tested via focusing on those who have engaged in helping behavior in situations involving real emergencies or great danger to the helpers.

For example, Oliner and Oliner (1988) obtained personality evidence on people from various European countries who actively intervened to try to rescue Jews during World War II. Oliner and Oliner found considerable similarity among the helpers. They had an internal locus of control (i.e., they believed that their own actions made a difference in changing their environments). They also believed in a just world, which meant that they believed that people who do good deeds will be rewarded for them. They felt socially responsible and believed that people should do their best to help others who are in need. They felt empathy for persons who were in distress, and were not egocentric in general in their approach to other people.

Another powerful presentation of humans' caring spirit is contained in Eva Fogelman's book *Conscience and courage* (1994). She, too, interviewed persons who had rescued Jews during the Holocaust. Her treatment highlighted the small events that often shaped long-term, complex sequences of helping. One of the people she interviewed was a small businessman in Warsaw, Poland. When the Nazis conquered Poland and created the Jewish ghetto, this man said that he was curious about what was happening in the ghetto and found a way to visit it. He said that one transforming moment in that visit motivated him to spend the next four years risking his life and the lives of his family to help Jews hide from the Nazis. The moment was the sight of flies swarming over the bodies of dead and dying Jewish children who had been killed by the Nazi's brutal treatment of them.

1.10.6 CONCLUSION: SYNTHESIS OF THE "SOCIAL ANIMAL IN A COMPLEX WORLD"

Elliot Aronson wrote a popular introductory social psychology textbook entitled *The social animal* (1995). In this book, Aronson described why humans may be conceived as social animals, and began this thesis with Aristotle's statement:

> Man is by nature a social animal; an individual who is unsocial naturally and not accidentally is either beneath our notice or more than human. Society is something in nature that precedes the individual. Anyone who either cannot lead the common life or is so self-sufficient as not to need to, and therefore does not partake of society, is either a beast or a god. (*Politics*, c. 328 BC, reprinted Aronson, 1995, p. xvix)

While we may not be beasts, if we are reclusive we still are influenced by our society. The reclusive "Unibomber" was an example of a person who was greatly influenced by outside events and who sought to influence them in his own deadly way, yet he also was extremely reclusive in his behavior. Aronson contended that humans are more rationalizing animals than rational in their decision making (Aronson argued from a cognitive dissonance position). He cited many examples from reckless behavior in driving, smoking, and use of weapons, to warfare and genocide. He also argued that "People who do crazy things are not necessarily crazy" (p. 9). By this, Aronson meant that social conditions such as those studied by Milgram (1974) in his work on obedience may exert a powerful influence on human behavior.

In studies that were designed to examine how people might be induced to obey authority figures, somewhat similar to what might have occurred in Nazi Germany in the 1930s and 1940s, Milgram developed an experimental situation at Yale University that was advertised as involving the study of associative learning. If typical members of the community show up at the Yale psychology laboratories and are told that the study they are volunteering to do concerns learning behavior, why should they doubt the investigator? That was the type of situation set up by Milgram. The researchers were "official-sounding and -looking" in order to try to determine if ordinary people would obey an authority and do something that they thought might be harming another person. In the course of the experiments, these individuals were "required" to give what they thought were increasingly intense shocks to another person (actually an experimental accomplice). Many of the participants did so, even to the point of the accomplice feigning a heart attack. Milgram's methods were questionable in terms of potential harm to the participants and could not be carried out today, but they show the great power of social situations.

Situations such as those studied in the obedience work are only illustrative of the diverse social situations that people encounter on almost a daily basis. These situations may involve close others, acquaintanceship groups, work associates, or strangers, but whatever their nature they may totally engulf a person's field of thought and action. People may be induced to act in a way that is opposed to their personality and moral position because of the nature and influence of the situation.

We believe that a broadly construed social psychological point of view in the last part of the twentieth century represents a complex view of human life that is appropriate to the intricacies of human behavior, behavioral disorders, and psychopathology. This approach points toward many of the types of social variables and situations that influence people's behavior. As was shown by the 1996 Olympic Park bombing deaths in Atlanta, USA, in many ways we live in a terrifying world. Terrorism and other kinds of daily violence show the propensity of people to kill others in an instant with little guilt. Why? Inevitably the best available answers to such questions take us on an odyssey into work on social psychological findings and theories. These findings and theories are insufficient, in and of themselves, to answer such global questions. But they do sensitize us to the many nuances of understanding that may be tenable in considering causation. Coupled with the work from related fields in psychology, this body of work frames an understanding of some of the major dynamics behind these seemingly inane acts.

Contemporary social psychology is well integrated with the fields of personality, health psychology, and research in clinical psychology, all of which form an indispensable base for the practice of clinical psychology. This melding together of fields in the pursuit of knowledge and relevant application is illustrated in all of the sections of this chapter. It leads to a synergy in research that is quite positive and exciting for progress in understanding better the "social animal."

1.10.7 REFERENCES

Abramson, L. Y., Seligman, M. E., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87*(1), 49–74.

Allport, G. W. (1985). The historical background of social psychology. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 1, pp. 1–46). New York: Random House.

Altman, I., & Taylor, D. A. (1973). *Social penetration: The development of interpersonal relationships*. New York: Holt, Rinehart, and Winston.

Aron, A., & Aron, E. N. (1996). Self and self-expansion in relationships. In G. J. O. Fletcher & J. Fitness (Eds.), *Knowledge structures in close relationships: A social psychological approach* (pp. 325–344). Mahwah, NJ: Erlbaum.

Aronson, E. (1995). *The social animal* (7th ed.). New York: Freeman.

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41,* 258–290.

Asch, S. E. (1955). Opinions and social pressure. *Scientific American, 193,* 31–35.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215.

Batson, C. D., Duncan, B. D., Ackerman, P., Buckley, T., & Birch, K. (1981). Is empathic emotion a source of altruistic motivation? *Journal of Personality and Social Psychology, 40,* 290–302.

Batson, C. D., O'Quinn, K., Fulz, J., Vanderplas, M., & Isen, A. M. (1983). Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help. *Journal of Personality and Social Psychology, 45,* 706–718.

Bellah, R. N., Madsen, R., Sullivan, W. M., Swidler, A., & Tipton, S. M. (1985). *Habits of the heart: Individualism and commitment in American life*. New York: Harper and Row.

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1–62). New York: Academic Press.

Bernard, J. (1982). *The future of marriage* (2nd ed.). New Haven, CT: Yale University Press.

Berscheid, E. (1994). Interpersonal relationships. *Annual Review of Psychology, 45,* 79–129.

Blumstein, P., & Schwartz, P. (1983). *American couples*. New York: William Mobrow.

Bradbury, T. N., & Fincham, F. D. (1992). Attributions and behavior in marital interaction. *Journal of Personality and Social Psychology, 63*(4), 613–628.

Brehm, J. W. (1966). *A theory of psychological reactance*. New York: Academic Press.

Brehm, S. S. (1976). *The application of social psychology to clinical practice*. Washington, DC: Hemisphere.

Brewer, M. B. & Weber, J. G. (1994). Self-evaluation effects of interpersonal versus intergroup social comparison. *Journal of Personality and Social Psychology, 66,* 268–275.

Byrne, D. (1971). *The attraction paradigm*. New York: Academic Press.

Cate, R. M., Huston, T. L., & Nesselroade, J. R. (1986). Premarital relationships: Toward the identification of alternative pathways to marriage. *Journal of Social and Clinical Psychology, 4*(1), 3–22.

Cialdini, R. B. (1993). *Influence: Science and practice* (3rd ed.). New York: HarperCollins.

Cialdini, R. B., Kenrick, D. T., & Bauman, D. J. (1982). Effects of mood on prosocial behavior in children and adults. In N. Eisenberg (Ed.), *The development of prosocial behavior* (pp. 151–188). New York: Academic Press.

Coles, R. (1989). *The call of stories: Teaching and the moral imagination*. Boston: Houghton Mifflin.

Cooper, J., & Axsom, D. (1982). Effort justification in psychotherapy. In G. Weary & H. L. Mirels (Eds.), *Integrations of clinical and social psychology* (pp. 214–230). New York: Oxford University Press.

Cunningham, M. R. (1979). Weather, mood, and helping behavior: Quasi-experiments with the sunshine samaritan. *Journal of Personality and Social Psychology, 37*(11), 1947–1956.

Cunningham, M. R., Shaffer, D. R., Barbee, A. P., Wolff, P. L., & Kelley, D. L. (1990). Separate processes in the relation of elation and depression to helping: Social versus personal concerns. *Journal of Experimental Social Psychology, 26*(1), 13–33.

Darley, J. M., & Fazio, R. H. (1980). Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist, 35*(10), 867–881.

Derlega, V. J., Metts, S., Petronio, S., & Margulis, S. T. (1993). *Self-disclosure*. Newbury Park, CA: Sage.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influence upon individual judgment. *Journal of Abnormal and Social Psychology, 51,* 629–636.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*(1), 5–18.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.

Erikson, E. H. (1963). *Childhood and society* (2nd ed.). New York: Norton.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7,* 117–140.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Festinger, L., Schachter, S., & Back, K. W. (1930). *Social pressures in informal groups: A study of human factors in housing*. New York: Harper.

Fincham, F. D., & Bradbury, T. N. (1991). Cognition in marriage: A program of research on attribution. In W. H. Jones & D. Perlman (Eds.), *Advances in personal relationships: A Research annual* (Vol. 2, pp. 159–203). London: Jessica Kingsley.

Fincham, F. D., & Bradbury, T. N. (1993). Marital satisfaction, depression, and attributions: A longitudinal analysis. *Journal of Personality and Social Psychology, 64*(3), 442–452.

Fogelman, E. (1994). *Conscience and courage: Rescuers of Jews during the Holocaust*. New York: Doubleday/Anchor.

Frank, J. (1965). *Persuasion and healing*. Baltimore: Johns Hopkins University Press.

Frank, J. (1974). Psychotherapy: The restoration of morale. *American Journal of Psychiatry, 131,* 271–274.

Frankl, V. E. (1956). *Man's search for meaning*. New York: Washington Square.

Funder, D. C. (1987). Errors and mistakes: Evaluating the accuracy of social judgment. *Psychological Bulletin, 101*(1), 75–90.

Goffman, E. (1959). *The presentation of self in everyday life*. Garden City, NY: Doubleday.

Goffman, E. (1961). *Asylums: Essays on the social situation of mental patients and other immates*. Garden City, NY: Anchor.

Goldstein, A. P., Heller, K., & Sechrest, L. B. (1966). *Psychotherapy and the psychology of behavior change*. New York: Wiley.

Hart, A. J., & Morry, M. M. (in press). Nonverbal behavior, race, and attitude attributions. *Journal of Experimental Social Psychology*.

Harvey, J. H. (1983). The founding of the *Journal of Social and Clinical Psychology. Journal of Social and Clinical Psychology, 1,* 1–3.

Harvey, J. H. (1995). *Odyssey of the heart: The search for closeness, intimacy, and love*. Salt Lake City, UT: Freeman.

Harvey, J. H., & Galvin, K. S. (1984). Clinical implications of attribution theory and Research. *Clinical Psychological Review, 4*(1), 15–33.

Harvey, J. H., & Omarzu, J. (1997). Minding the close relationship. *Personality and Social Psychology Review, 1,* 223–239.

Harvey, J. H., Orbuch, T. L., Chwalisz, K. D., & Garwood, G. (1991). Coping with sexual assault: The role of account-making and confiding. *Journal of Traumatic Stress, 4*(4), 515–531.

Harvey, J. H., Town, J. P., & Yarkin, K. L. (1981). How fundamental is "the fundamental attribution error?" *Journal of Personality and Social Psychology, 40*(2), 346–349.

Harvey, J. H., Weber, A. L., Galvin, D. S., Huszti, H. D., & Garnick, N. N. (1986). Attribution and the termination of close relationships: A special focus on the account. In R. Gilmour & S. Duck (Eds.), *The emerging field of personal relationships* (pp. 189–201). Hillsdale, NJ: Erlbaum.

Harvey, J. H., Weber, A. L., & Orbuch, T. L. (1990). *Interpersonal Accounts: A social psychological perspective*. Oxford, UK: Blackwell.

Harvey, J. H., Wells, G. L., & Alvarez, M. D. (1978). Attribution in the context of conflict and separation in close relationships. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), *New directions in attribution research* (Vol. 2, pp. 235–260). Hillsdale, NJ: Erlbaum.

Hastorf, A. H., & Cantril, H. (1954). They saw a game: A case study. *Journal of Abnormal and Social Psychology, 49*(1), 129–134.

Hatfield, E., & Rapson, R. L. (1993). *Love, sex, and intimacy: Their psychology, biology, and history*. New York: HarperCollins.

Hatfield, E., & Sprecher, S. (1986). *Mirror, mirror: The importance of looks in everyday life*. Albany, NY: SUNY Press.

Heesacker, M., Conner, K., & Prichard, S. (1995). Individual counseling and psychotherapy: Applications from the social psychology of attitude change. *Counseling Psychologist, 23*(4), 611–632.

Heesacker, M., Petty, R. E., & Cacioppo, J. T. (1983). Field dependence and attitude change: Source credibility can alter persuasion by affecting message-relevant thinking. *Journal of Personality, 51*(4), 653–666.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley (reprinted by Erlbaum).

Heider, F. (1976). A conversation with Fritz Heider. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), *New directions in attribution theory* (Vol. 1, pp. 3–18). Hillsdale, NJ: Erlbaum.

Hendrick, S. S., & Hendrick, C. (1992). *Romantic love*. Newbury Park, CA: Sage.

Heppner, P. P., & Dixon, D. N. (1978). Effects of client perceived need and counselor role on clients' behaviors. *Journal of Counseling Psychology, 25*(6), 514–519.

Hill, C. T., Rubin, Z., & Peplau, L. A. (1976). Breakups before marriage: The end of 103 affairs. *Journal of Social Issues, 32*(1), 147–168.

Holtzworth-Munroe, A., & Jacobson, N. S. (1985). Causal attributions of marital couples: When do they search for causes? *Journal of Personality and Social Psychology, 48*(6), 1398–1412.

Horowitz, M. J. (1986). *Stress response syndromes* (2nd ed.). Northvale, NJ: Jason Aronson.

Howard, G. S. (1991). Cultural tales: A narrative approach to thinking, cross-cultural psychology, and psychotherapy. *American Psychologist, 46,* 187–197.

Huston, T. L., McHale, S. M., & Crouter, A. C. (1986). When the honeymoon's over: Changes in the marriage relationship over the first year. In R. Gilmour & S. W. Duck (Eds.), *The emerging field of personal relationships* (pp. 109–132). Hillsdale, NJ: Erlbaum.

Ickes, W. (1993). Empathic accuracy. Special Issue: Viewpoints on personality: Consensus, self–other agreement, and accuracy in personality judgement. *Journal of Personality, 61*(4), 587–610.

Isen, A. M., & Levin, P. A. (1972). Effect of feeling good on helping: Cookies and kindness. *Journal of Personality and Social Psychology, 21,* 384–388.

Jacobs, R. C., & Campbell, D. T. (1961). The perpetuation of an arbitrary norm tradition through several generations of a laboratory microculture. *Journal of Abnormal and Social Psychology, 62,* 649–658.

James, W. (1890). *The principles of psychology*. New York: Holt.

Janoff-Bulman, R. (1992). *Shattered assumptions: Toward a new psychology of trauma*. New York: Free Press.

Jones, E. E. (1964). *Ingratiation: A social psychological analysis*. New York: Appleton-Century-Crofts.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219–263). New York: Academic Press.

Jones, E. E., Kanouse, D., Kelley, H. H., Nisbett, R. E., Valins, S., & Weiner, B. (Eds.) (1972). *Attribution: Perceiving the causes of behavior*. Morristown, NJ: General Learning Press.

Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79–94). Morristown, NJ: General Learning Press.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation* (Vol. 15, pp. 192–240). Lincoln, NE: University of Nebraska Press.

Kelley, H. H. (1972). Causal schemata and the attribution process. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 151–174). Morristown, NJ: General Learning Press.

Kelley, H. H. (1979). *Personal relationships: Their structures and processes*. Hillsdale, NJ: Erlbaum.

Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations: A theory of interdependence*. New York: Wiley.

Kelley, H. H., & Thibaut, J. W. (1985). Self-interest, science, and cynicism. *Journal of Social and Clinical Psychology, 3,* 26–32.

Kellogg, R., & Baron, R. S. (1975). Attribution theory, insomnia, and the reverse placebo effect: A reversal of Storms and Nisbett's findings. *Journal of Personality and Social Psychology, 32*(2), 231–236.

Langer, L. L. (1991). *Holocaust testimonies: The Ruins of Memory*. New Haven: Yale University Press.

Lantz, J. E. (1993). *Existential family therapy: Using the concepts of Viktor Franck*. Northvale, NJ: Jason Aronson.

Latané, B. (1981). The psychology of social impact. *American Psychologist, 36*(4), 343–356.

Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* New York: Appleton-Century-Crofts.

Lazarus, R. S. (1966). *Psychological stress and the coping process*. New York: McGraw-Hill.

Leary, M. R., & Miller, R. S. (1986). *Social psychology and dysfunctional behavior: Origins, diagnosis, and treatment*. New York: Springer-Verlag.

Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion* New York: Plenum.

Lewis, C. S. (1961). *A grief observed*. New York: Bantam.

Linville, P. W. (1982). The complexity-extremity effect and age-based stereotyping. *Journal of Personality and Social Psychology, 42,* 193–211.

Linville, P. W., & Jones, E. E. (1980). Polarized appraisals of outgroup members. *Journal of Personality and Social Psychology, 38*(5), 689–703.

Mackie, D., & Cooper, J. (1984). Attitude polarization: Effects of group membership. *Journal of Personality and Social Psychology, 46*(3), 575–585.

McGuire, W. J. (1964). Inducing resistance to persuasion: Some contemporary approaches. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 1). New York: Academic Press.

McGuire, W. J., & Papageorgis, D. (1962). Effectiveness of forewarning in developing resistance to persuasion. *Public Opinion Quarterly, 26,* 24–34.

Milgram, S. (1974). *Obedience to authority: An experimental view*. New York: Harper and Row.

Monge, P. R., & Kirste, K. K. (1980). Measuring proximity in human organization. *Social Psychology Quarterly, 43*(1), 110–115.

Moscovici, S., Lage, S., & Naffrechoux, M. (1969). Influence of a consistent minority on the responses of a majority in a color perception task. *Sociometry, 32,* 365–380.

Mullen, B. (1991). Group composition, salience and cognitive representations: The phenomenology of being in a group. *Journal of Experimental Social Psychology, 27,* 297–323.

Mullen, B., & Hu, L. (1989). Perceptions of ingroup and outgroup variability: A meta-analytic integration. *Basic and Applied Social Psychology, 10*(3), 233–252.

Murray, S. L., Holmes, J. G., & Griffin, D. W. (1996). The benefits of positive illusions: Idealization and the construction of satisfaction in close relationships. *Journal of Personality and Social Psychology, 70*(1), 79–98.

Newcomb, T. M. (1961). *The acquaintance process*. New York: Holt, Rinehart and Winston.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Oliner, S. P., & Oliner, P. M. (1988). *The altruistic personality: Rescues of Jews in Nazi Europe*. New York: Free Press.

Orvis, B. R., Kelley, H. H., & Butler, D. (1976). Attributional conflict in young couples. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), *New directions in attribution research* (Vol. 1, pp. 353–386). Hillsdale, NJ: Erlbaum.

Patnoe, S. (1988). *A narrative history of experimental social psychology: The Lewin tradition*. New York: Springer-Verlag.

Petty, R. E., & Cacioppo, J. T. (1981). *Attitudes and persuasion: Classic and contemporary approaches*. Dubuque, IA: Brown.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: peripheral routes to attitude change*. New York: Springer-Verlag.

Phillips, D. P. (1974). The influence of suggestion on suicide: substantive and theoretical implications of the Werther effect. *American Sociological Review, 39,* 340–354.

Phillips, D. P. (1979). Suicide, motor vehicle fatalities and the mass media: Evidence toward a theory of suggestion. *American Journal of Sociology, 84,* 1150–1174.

Prager, K. J. (1995). *The psychology of latimacy*. New York: Guilford.

Regan, D. T., & Totten, J. (1975). Empathy and attribution: Turning observers into actors. *Journal of Personality and Social Psychology, 32*(5), 850–856.

Reis, H. T., & Shaver, P. (1988). Intimacy as an interpersonal process. In S. W. Duck (Ed.), *Handbook* *of personal relationships* (pp. 367–389). Chichester, UK: Wiley.

Robins, R. W., Spranca, M. D., & Mendelsohn, G. A. (1996). The actor–observer effect revisited: Effects of individual differences and repeated social interactions on actor and observer attributions. *Journal of Personality and Social Psychology, 71*(2), 375–389.

Rosenhan, D. L. (1973). On being sane in insane places. *Science, 179,* 250–258.

Rosenthal. A. M. (1964). *Thirty-eight witnesses*. New York: McGraw-Hill.

Rosenthal, R., & Jacobson, L (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York: Holt, Rinehart and Winston.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173–220). New York: Academic Press.

Schachter, S. (1964). The interaction of cognitive and physiological determinants of emotional state. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 1, pp. 49–80). New York: Academic Press.

Scheff, T. J. (1966). *Being mental ill: A sociological theory*. Chicago: University of Chicago Press.

Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death*. San Francisco: Freeman.

Shay, J. (1994). *Achilles in Vietnam: Contact trauma and the undoing character*. New York: Atheneum/Macmillan.

Sherif, M. (1937). An experimental approach to the study of attitudes. *Sociometry, 1,* 90–98.

Snyder, C. R., & Forsyth, D. R. (Eds.) (1991). *Handbook of Social and clinical psychology*. New York: Pergamon.

Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology, 35,* 656–666.

Steele, S. (1990). *The content of our character*. New York: HarperCollins.

Storms, M. D., & Nisbett, R. E. (1970). Insomnia and the attribution process. *Journal of Personality and Social Psychology, 16,* 319–328.

Strong, S. R. (1968). Counseling: An interpersonal influence process. *Journal of Counseling Psychology, 15,* 215–224.

Strong, S. R., & Claiborn, C. D. (1982). *Change through interaction: Social psychological processes of counseling and psychotherapy*. New York: Wiley-Interscience.

Strong, S. R., & Schmidt, L. D. (1970). Expertness and influence in counseling. *Journal of Counseling Psychology, 17,* 81–87.

Strupp, H. (1981). The outcome problem in psychotherapy: Contemporary perspectives. In J. H. Harvey & M. M. Parks (Eds.), *Psychotherapy research and behavior change* (pp. 41–71). Washington, DC: American Psychological Association.

Swann, W. B., Jr. (1996). *Self-traps: The elusive quest for higher self-esteem*. New York: Freeman.

Turner, J. C. (1991). *Social influence*. Pacific Grove, CA: Brooks/Cole.

Valins, S., & Nisbett, R. E. (1972). Attribution processes in the development and treatment of emotional disorders. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 137–150). Morristown, NJ: General Learning Press.

Vaughan, D. (1986). *Uncoupling: Turning points in intimate relationships*. New York: Oxford University Press.

Wallach, M. A., & Wallach, L. (1983). *Psychology's sanction for selfishness The error of egoism in theory and therapy*. San Francisco: Freeman.

Walster (Hatfield), E., Aronson, V., Abrahams, D., & Rottman, L. (1966). The importance of physical attractiveness in dating behavior. *Journal of Personality and Social Psychology, 4,* 508–516.

Weary, G., Elbin, S. D., & Hill, M. G. (1987). Attribution and social comparison processes in depression. *Journal of Personality and Social Psychology, 52,* 605–610.

Weary, G., Jordan, J. S., & Hill, M. G. (1985). The attributional norm of internality and depressive sensitivity to social information. *Journal of Personality and Social Psychology, 49*(5), 1283–1293.

Weary, G. & Mirels, H. L. (Eds.) (1982). *Integrations of clinical and social psychology*. New York: Oxford University Press.

Weary, G., Stanley, M. A., & Harvey, J. H. (1989). *Attribution*. New York: Springer-Verlag.

Weber, A. L., & Harvey, J. H. (1994). Accounts in coping with relationship loss. In A. L. Weber & J. H. Harvey (Eds.), *Perspectives on close relationships* (pp. 285–306). Boston: Allyn & Bacon.

Weber, A. L., Harvey, J. H., & Stanley, M. A. (1987). The nature and motivations of accounts for failed relationships. In R. Burnett, P. McGhee, & D. C. Clarke (Eds.), *Accounting for relationships* (pp. 114–135). London: Methuen.

Weiner, B. (1974). *Achievement motivation and attribution theory*. Morristown, NJ: General Learning Press.

Weiss, R. S. (1975). *Marital separation*. New York: Basic Books.

Weiss, R. S. (1988). Loss and recovery. *Journal of Social Issues, 44,* 37–52.

Williamson, G. M., & Clark, M. S. (1989). Providing help and desired relationship type as determinants of changes in moods and self-evaluations. *Journal of Personality and Social Psychology, 56*(5), 722–734.

Wilson, D. W. (1981). Is helping a laughing matter? *Psychology, 18,* 6–9.

Yost, J. H., & Weary, G. (1996). Depression and the correspondent inference bias: Evidence for more effortful cognitive processing. *Personality and Social Psychology Bulletin, 22*(2), 192–200.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.11

Cognitive Psychology

RICHARD HOFFMAN and MUSTAFA al'ABSI
University of Minnesota, Duluth School of Medicine, Duluth, MN, USA

1.11.1 INTRODUCTION 323

1.11.1.1 Cognition and Clinical Psychology 323
1.11.1.2 Models of Cognition, Learning, and Memory 324
1.11.1.3 Conscious and Unconscious Processing 326
1.11.1.4 Neural Network Models and Classical Conditioning 329
1.11.1.5 Multiple-level Neural Network Models 330

1.11.2 MOOD AND MEMORY 331

1.11.2.1 Effects of Affective States 331
1.11.2.2 State-dependent Learning 332
1.11.2.2.1 The Network Theory of Affect 332
1.11.2.3 Depressed Mood and Cognitive Performance 333

1.11.3 COGNITION–EMOTION INTERACTIONS 335

1.11.3.1 Cognitive Factors 335
1.11.3.2 Cognitive Appraisal 336
1.11.3.3 Attribution and Affect 337
1.11.3.4 Psychobiological Substrates 337

1.11.4 COGNITIVE MODELS OF ANXIETY 338

1.11.4.1 General Cognitive Models of Anxiety 338
1.11.4.2 Specific Phobias 339
1.11.4.3 Panic Disorder and Agoraphobia 339
1.11.4.4 Social Phobias 340
1.11.4.5 Post-traumatic Stress Disorder 341
1.11.4.6 Generalized Anxiety Disorder 341

1.11.5 COGNITIVE MODELS OF DEPRESSION 342

1.11.5.1 Cognitive Distortion Models 342
1.11.5.2 Differential Activation Model 344
1.11.5.3 Cognitive Appraisal Model 345
1.11.5.4 Learned Helplessness Model 346
1.11.5.4.1 Revised helplessness model 346
1.11.5.4.2 Hopelessness theory 347

1.11.6 CONCLUSIONS 348

1.11.7 REFERENCES 348

1.11.1 INTRODUCTION

1.11.1.1 Cognition and Clinical Psychology

Cognitive psychology is that specialty area of psychology concerned with the examination of how knowledge is acquired, how it is internally represented, how it interacts and modifies or is modified by previous experience and existing information, and how acquired knowledge affects feelings and behavior. The subject matter encompasses not only the process of thinking (attention, concentration, consciousness, memory), but also an examination of emotions and motivation in relation to cognition. There has been a dramatic shift in the theoretical base of clinical psychology since the 1940s from psychoanalytic and behavioral models to cognitive models as the task of clinical psychology has changed to encompass more than simply disturbed behavior and maladaptive habits. Clinical psychologists now concern themselves routinely with such aspects of human behavior as the emotional state of their patient, their patient's ability to cope with difficult situations, their interpersonal problems, their sense of well-being or self-esteem, their memory and concentration, their thoughts and thinking process, and their physiological health status (Brewin, 1988).

In many respects this is a golden age for cognitive psychology. The rapid development of noninvasive neuroimaging technologies in the mid-1980s has allowed researchers the unprecedented opportunity to confirm some hypotheses of cognitive functioning derived from laboratory learning experiments by looking at neural function at the time that the subject performs the experiment. The use of positron emission tomography (PET), functional magnetic resonance imaging (fMRI), event-related potentials (ERPs), electroencephalography (EEG) and magnetoencephalography (MEG) technologies allows researchers to look at real-time "pictures" of the human brain in the process of cognition through measurements of blood flow and brain electrical activity. The reader is referred to Posner & Raichle (1994) for an overview of this fascinating area and Posner & Raichle (1995) for further commentary and critique. Noninvasive neuroimaging has led to increased collaboration between cognitive psychologists, neurologists, neuropsychologists, clinical psychologists, psychiatrists, mathematicians, physicists, and neuroscientists of all sorts. The point is being rapidly reached that brain, mind, and behavior need not be studied in isolation, and emerging electrophysiological technologies such as multiple single unit recordings and ERPs will only add to the available tools at the disposal of future cognitive scientists. Models of cognitive processes, neural networks, and mathematical models of cognitive function have also become highly developed since the late 1980s (Taylor, 1997), there have been enormous strides in our understanding of memory functioning from lesion studies (Zeki, 1993), and hundreds of studies have been published in the areas of mood, motivation, and cognitive social psychology. Within this same time span, cognitive behavior therapy and rational emotive behavior therapy have become highly refined therapeutic modalities that address psychological problems ranging from anxiety and depression to panic disorder, anger problems, obsessive-compulsive disorder, and suicidal behavior (Ellis & Dryden, 1997; Freeman & Reinecke, 1993; Yankura & Dryden, 1997).

This chapter will provide an overview of some of the current literature in cognitive psychology that is particularly germane to the practice of clinical psychology in general and the practice of cognitive behavioral therapies in particular, with the clear understanding that this is a highly condensed and highly selective sampling of a rapidly expanding area. A major focus will be an examination of the origin and modification of emotional states, particularly anxiety and depression, due to their central role in contemporary clinical psychology.

1.11.1.2 Models of Cognition, Learning, and Memory

Since the 1970s, a central tenet in cognitive psychology has been that much, if not all of cognition can be conceptualized as information processing, and a great deal of work has been done since that time to identify component cognitive processes that underlie cognition. Early models of cognition were mainly derived inferentially by systematically varying stimulus properties and carefully measuring reponses. The availability of high-speed computers allowed the later development of more sophisticated discrete stage linear mathematical models which, at least in a very limited sense, were capable of mimicking some aspects of cognitive function. The later use of more sophisticated computer systems in the 1980s dovetailed with the growing awareness that functional brain systems and complex neural cicuits rather than linear modules were involved in higher brain function. Since that time, computer simulations of cognitive processes have been built that allow hypothesis testing of complex interactive systems and model the possible effect of localized brain damage on subsequent brain function. Computational models have been built which examine acquired disorders of reading (Hinton & Shallice, 1991; Mozer & Behrmann, 1990; Patterson, Seidenberg, & McClelland, 1990; Plaut & Shallice, 1993) as well as disorders of perception, attention, memory, frontal lobe function, and language (Feinberg & Farah, 1997).

Computational models of cognition have influenced the field of cognitive psychology since the 1970s by providing meaningful heuristics to link environmental events and information stored in memory to subsequent behaviors. The earliest such models were based primarily on digital computer processing methodology. These models used processing algorithms based upon the premise that environmental events trigger the retrieval of symbolically coded events from a discrete physical memory location. These events are operated upon by some central processor according to previously stored instructions, and information about the subsequent behavior is fed back to memory stores by this same central processor (Newell & Simons, 1972). More recently, computational models of cognition have been developed that more closely model brain function and rely upon parallel distributed processing (PDP) architecture. These have provided the methodology for artificial neural networks as cognitive processing analogues (Anderson, 1995; Rumelhart & McClelland, 1986). Because the human brain is a functionally and anatomically distributed and highly interactive system, PDP modeling has provided a conceptual framework and concrete tools to begin to explain the effects of local lesions on cognitive function as well (Farah, 1997).

Briefly stated, PDP systems consist of large numbers of neuron-like units multiply connected to each other by weighted connections that determine the amount of activation or neural transmission that flows from one unit to another. Each unit or collection of units functions locally and in parallel with other units, with no true central processor governing the behavior of the network. An event is represented as a pattern of activation distributed across units. PDP networks can accomodate multiple types of learning depending upon the computational properties of the network. These computational properties in general are a function of the activation rule of that network (e.g., linear vs. nonlinear activation), the type of connectivity within the network (e.g., unidirectional vs. bidirectional), and the learning rule imposed by the architecture of input units vs. output units. Long-term memory is represented by the encoding of a pattern of connection strengths distributed within a population of units or neurons.

A basic premise of most PDP models of memory is that because every part of the brain or all collections of neural units possess the ability to learn, memory cannot therefore be housed in a specific structure or structures of the brain as was the case in classical theories of memory, but rather expressed as memory processes or circuits which are related to those areas processing that information. This is consistent with the observation that areas of the brain which process auditory information house auditory memory and areas that process visual information also house visual memory, for example, and is consistent with the observation that functionally similar areas of the brain may contain multiple memories of any specific event based upon their specific function (color, form, size, motion, etc.), all or some of which may be recalled when a memory of the event is recalled (Kolb & Whishaw, 1996).

The process of memory recall can be divided into two broad classes or types with different underlying neural structures. Explicit memory refers to the conscious recollection of previously experienced events, whereas implicit memory is a process that is automatic, occurs at a level outside of consciousness, and is typically inferred from the effect that experience or practice has on the performance of a particular task (Moscovitch, 1984). Implicit memory is encoded in much the same way that it is perceived, whereas there is some active reorganization of data that occurs in explicit memory.

Explicit memory can be further subdivided into two subtypes, strategic and associative/cue-dependent, based upon those types of tests used to elicit explicit memory. Associative/cue-dependent explicit memory is a relatively automatic response that is elicited following a cue that is sufficient to bring that memory to mind. Strategic memory is that in which cuing does not elicit a specific response but rather initiates a memory "search" not unlike problem-solving (Moscovitch, 1992). Petri and Mishkin (1994) have identified four major limbic structures involved in the explicit memory system: the rhinal cortex, the amygdala, the hippocampus, and the prefrontal cortex. These structures have reciprocal connections with the medial thalamus, the basal forebrain and orbital frontal cortex, and sensory areas of the neocortex. Experimental results suggest that the rhinal cortex is involved in object memory, the hippocampus in spatial memory, and the amygdala in emotional memory (Kolb & Whishaw, 1996). Evidence both from animal models and lesion studies in humans has demonstrated sparing of old memories and implicit memories following damage to medial temporal structures, but severe impairments in the formation of new explicit memories.

Implicit memory can be further subdivided into two subtypes, procedural and item-specific, based upon those types of tests used to elicit implicit memory. Procedural memory refers to the acquisition and retention of general skills, procedures, or rules including learned motor tasks. Item-specific tests, in contrast, are concerned with the acquisition and retention of a particular type of information and the increased accuracy or speed with which a previously seen item identified on repetition is a function of the repetition priming effect (Moscovitch, 1992).

Petri and Mishkin (1994) have identified the caudate nucleus and putamen (basal ganglia) as those brain structures centrally involved in the implicit memory system. The basal ganglia receive projections from cells in the substantia nigra and all regions of the neocortex and send projections via the globus pallidus and ventral thalamus to the premotor cortex. Animals with damage in basal ganglia circuits demonstrate preserved recognition memory, spatial memory, and emotional memory, but demonstrate impairments in learning motor skills, learning to make appropriate responses to cues, and demonstrate poor performance on association tasks (Petri & Mishkin, 1994).

Finally, fundamental to the understanding of conscious and unconscious processing is the inter-relationship between networks of attention and short-term memory or "working memory." Working memory refers to the form of memory used to hold words, names, digits, or other items in memory for a brief period of time (Baddely, 1986). Working memory appears to be a function of the prefrontal cortex which allows the maintenance of information in short-term memory that is later used to guide action (Goldman-Rakic, 1987; Moscovitch & Umilta, 1991). Short-term or working memory has a different structural basis from long-term memory. The parietal-frontal spatial system, which includes Brodmann's frontal cortex area 8 and projections from the parietal cortex, is involved in short-term memory for spatial location of objects and in searching for an object when a stimulus is presented. Frontal cortex areas 9 and 46, with projections from the inferior temporal cortex (the inferior-temporal dorsolateral frontal system), are involved in the short-term memory for visual objects and remembering objects that are identified sequentially (Petrides, Alivisatos, Evans, & Meyer, 1993; Wilson, Scalaidhe, & Goldman-Rakic, 1993). The available evidence suggests that working memory and long-term memory are parallel mechanisms in which material is processed separately and simultaneously, although information can be pulled from long-term memory and used in working memory for short-term problem-solving.

Electrophysiological evidence from primate studies suggests that there are at least three different types of attentional mechanisms in the mammalian brain (Moran & Desimone, 1985; Petersen, Robinson, & Morris, 1987; Spitzer, Desimone, & Moran, 1988). There is evidence for an attentional mechanism in the parietal lobe that enhances spatial attention, one in the inferior temporal region that selects objects, and one in the visual and posterior temporal cortex that selects object features. In addition, cells in the frontal eye fields have been demonstrated to select movement. PET and ERP data with human subjects are consistent with these findings (Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1991; Corbetta, Miezin, Shulman, & Petersen, 1993; Naatanen, 1992).

Regions of the frontal lobes are also activated in tasks requiring response selection. For example, the anterior cingulate cortex is activated in divided attention tasks as well as the dorsolateral prefrontal cortex (Corbetta et al., 1991), and verb-generation tasks activate the inferior frontal cortex (Posner & Raichle, 1994). Posner and Petersen (1990) have proposed that the frontal lobe therefore contains a second attentional system or executive attentional system which plays a major role in the activation of selective attention systems of the posterior cortex and which is closely related to the working memory functions of the frontal lobe. In this model, the frontal lobe is responsible for the programming of mental operations, and the content of consciousness consists of information within working memory that is operated on by the executive attentional system.

Clinically, it is not always easy to separate attention and working memory. Damage to prefrontal areas of the brain, if extensive enough, causes deficits in both attentional circuits and working memory. Also in some instances, observed deficits which appear to represent functional disorders of working memory may instead be residual or secondary to damage to atentional mechanisms. An assessment model of attention has been proposed by Mateer and Mapou (1996) that at least partially addresses this by carefully delineating the boundaries of attention and separating it into two component parts, deployment and capacity/encoding. Deployment is the ability to channel and focus attention, which includes the assessment of arousal, focused attention, and sustained attention. Arousal is assessed by observation, focused attention is assessed by tasks that require the rapid scanning and identification of targets, and sustained attention is assessed by continuous performance tests. Capacity/encoding is the ability to hold information and process it even if distracted or required to divide attention, and includes tasks that measure span of attention, resistance to interference, and mental manipulation.

1.11.1.3 Conscious and Unconscious Processing

An understanding and appreciation of the unconscious has been one of the central themes of clinical psychology historically, and certainly a key concept in psychoanalysis and psychoanalytic psychotherapy. A companion theme is the emphasis placed upon gaining an awareness of the internal state of the patient, an awareness both on the part of the patient and on the part of the therapist. Patients and nonpatients alike respond to situations and other people in ways that seem disconnected to the current situation and have day-to-day experiences that seem to be neither desired nor clearly volitional. Is this a result of experiences or learning that are below the threshold of awareness or in some way processed differently than those experiences of which we are fully aware?

Early observations of symptom remission in hysterical patients under hypnosis led Freud, Breuer, and later Janet to develop a model of repressed emotions and cognitive content to explain the seemingly miraculous disappearance and reappearance of all manner of dissociative symptoms (blindness, paralysis, anesthesias) in response to hypnosis and suggestion (Brewin, 1988). It was thought that such hysterical symptoms could be the result of traumatic experiences that could not be readily recalled by patients due to the protective psychological mechanism they labeled repression but which, under special circumstances, could be reexperienced or reexamined to the patient's benefit. Working through the patient's resistance to experiencing painful affect and unacceptable thoughts from unconscious to preconscious to conscious awareness, then becomes a central tenet of the therapeutic process. It was felt that the content of dreams and parapraxes or slips of the tongue were further "evidence" of unconscious processes. This model of repressed traumatic emotional events leading to symptoms and seemingly being responsible for day-to-day behavior requires some examination of a mechanism or mechanisms to explain information processing that occurs out of conscious awareness. There has been a great deal of evidence from the laboratories of both cognitive psychologists and cognitive neuropsychologists, particularly since the late 1980s, to confirm a central role of the automatic processing of information without conscious representation.

Early evidence of cognitive performance without awareness can be found in studies in the neuropsychological literature describing patients with amnesic syndrome and so-called "blindsight," reported in the 1960s and 1970s, although there are anecdotal reports dating back at least to the late nineteenth century. Although it is beyond the scope of this chapter to take an exhaustive look at this fascinating area, it would be remiss not to report some examples in detail which illustrate not only preserved discrete cognitive functioning without the awareness of the patient that this capacity remains, but also preserved discrete cognitive functioning that occurs in the total absence of any awareness on the patient's part that they are even responding in any meaningful way to external events, events that they seemingly cannot discriminate or detect. For a thorough review of this rich area and the relationship of such syndromes to consciousness, the reader is referred to the publication by Weiskrantz (1997).

One of the first well-documented cases of learning in the absence of awareness was the famous patient H.M., who was tested extensively over the course of several years by multiple investigators (Corkin, 1968). H.M. became severely amnesic following the removal of structures in the medial portions of both of his temporal lobes to treat intractable epilepsy. Although H.M. had severe memory difficulties following this neurosurgical procedure, he was nonetheless subsequently able to to master pursuit rotor tasks and learn mirror drawing tasks and retain them from session to session, despite the fact that he had no memory of the experimental sessions or memory of ever having seen or used the experimental apparatus. This same sort of motor skill learning and retention of training without any conscious awareness of the learning event has been documented in other densely amnesic patients by multiple authors (Brooks & Baddeley, 1976; Schacter, 1994; Schacter, McAndrews, & Moscovitch, 1988).

Classical conditioning has also been demonstrated in the absence of awareness. In a series of experiments reported by Weiskrantz and Warrington (1979), a severely amnesic Korsakoff patient was conditioned for anticipatory eye closure with a neutral light and tone preceding an air puff to the eye. Despite the fact that he reported no recall between conditioning trials of the air puff or what the apparatus that was in front of him was being used for, the conditioned stimulus reliably elicted the conditioned eyeblink.

Perhaps more surprising has been the observation that densely amnesic patients are able to learn and retain verbal information without explicit recognition or recall. In a series of clever experiments, severely amnesic patients were shown a series of words and tested for recognition after several minutes (Weiskrantz & Warrington, 1970). If asked to respond yes or no to whether they had seen specific words, their performance was at a chance level. If instead they were provided with a very fragmented drawing of a target word and asked to guess the identity of the word, they were much more able to identify those fragmented words to which they had been previously exposed than those to which they had not been exposed, despite having no recollection whatsoever of having seen any of the words.

This sort of "priming" effect to enhance recall in the absence of awareness has also been demonstrated by providing densely amnesic subjects with the first few letters of target words, which increased the facilitation of retention induced by the previous exposure, the so-called stem completion task. Retention intervals as long as four months have been reported following priming in severely amnesic patients such as H.M. (Milner, Corkin, & Teuber, 1968), demonstrating that priming is capable of revealing implicit memory (implicit processing) in the absence of explicit memory.

Experiments with cortically blind patients have revealed that the phenomenon of information processing without awareness is not simply confined to memory systems. It has been demonstrated by several investigators that patients who have sustained damage to the striate cortex of the brain at Brodmann area 17 (the V1 projection area) and who are cortically blind can nonetheless demonstrate the ability to detect the presence of stimuli, to locate them in space, to discriminate direction of movement, to discriminate the orientation of lines, and to discriminate between different wavelengths of light, despite having no awareness that there is any stimulus present in their blind hemifield (Blythe, Kennard, & Ruddock, 1987; Cowey & Stoerig, 1991; Weiskrantz, 1986, 1995).

These seemingly impossible feats of visual perception are explainable due to multiple cortical connections to the human optic nerve and retina. The major input from the optic nerve goes to the dorsal lateral geniculate nucleus and ultimately to V1, and if this circuit is intact patients cannot only "see" objects in the visual field but also have the awareness of "seeing." About 10–15% of the output from the optic nerve goes to nine other visual pathways or connections (superior colliculus, ventral lateral geniculate nucleus, pre-tectum, nucleus of the optic tract, dorsal terminal accessory optic nucleus, lateral terminal accessory optic nucleus, medial terminal accessory optic nucleus, inferior pulvinar, and suprachiasmatic nucleus) which affords the possibility of retaining some aspects of the ability to "see" but without the awareness of seeing (Sahraie et al., 1997; Weiskrantz, 1997).

If the phenomena of "blindsight" and memory retention in severely amnesic patients are to have relevance to the everyday practice of clinical psychology, similar processing of information without awareness would need to be documented in nonbrain injured individuals as well. Although an area not without considerable controversy, investigations in the area of subliminal perception and subliminal semantic activation provide additional evidence that processing of information can and does occur in the absence of conscious awareness, a level of processing that has been referred to as unconscious cognition (Greenwald & Draine, 1997).

Merickle and Joordens (1997) have used a version of the Jacoby exclusion task (Debner & Jacoby, 1994; Jacoby, 1991) with short stimulus onset asynchronies to determine the critical stimulus duration at which the unconscious influence of priming exceeded the conscious influence of priming in a word stem completion task. This same sort of priming effect (very similar to the priming effect observed in amnesic patients) has been demonstrated in the Stroop color word task (Logan, Zbrodoff, & Williamson, 1984; Merikle & Cheesman, 1987), in both visual and auditory word recognition tasks (Groeger, 1988; Marcel, 1983), and in visual search tasks (Schneider & Shiffrin, 1977). Multiple neural pathways for processing these types of information are the likely neuroanatomical substrate to account for these types of processing independent of awareness.

Several lines of evidence from numerous investigators in cognitive psychology suggest that there are two general types of information processing: conscious or control processes which appear to be clearly under the control of the individual and automatic processes which are triggered by a stimulus and operate outside of awareness and are therefore not under the control of the individual (Posner & Snyder, 1975). Control processes by definition would be limited by the attention span of the individual, could be characterized as deliberate, effortful, adaptable, and to some extent highly flexible, but processes that could be easily disrupted by stimuli competing for that individual's attention. Control processes would be more likely to be elicited by novel situations which require more effortful cognitive processing (Brewin, 1988), but in highly complex or ambiguous situations the limitations imposed by attention and attention span may lead to decisions that are made by simple rules or *a priori* expectations or hypotheses. Automatic processes, in contrast, require minimal attention, may be activated without intention or awareness, are likely to be rapid and inflexible, and are more likely to be activated in familiar situations in which the individual has had considerable experience. Since automatic processes run in the background below the level of awareness, they are subject to misattributions of causality by linking two events together that occur close together in time and by judging more salient stimuli as causally prior. Automatic processing has an obvious correlate in what has been referred to by investigators as implicit memory in which prior experience influences later processing without conscious or deliberate recollection of the experience (Schacter, 1987), whereas the correlative memory analogue for conscious processing is explicit memory. There is at least some evidence to suggest that conscious processing of information inhibits automatic processing of that and contiguous stimuli (Posner & Snyder, 1975), suggesting some inhibitory effect of conscious processing and inhibitory effect of explicit memory.

Cognitive models of conscious processing and explicit memory vs. automatic processing and implicit memory have a direct relationship to psychopathology and patient behavior. The excessive rumination of depressed patients may be a manifestation of the inhibitory effect of conscious processing such that a patient's narrow focus on perceived deficiencies or perceived slights by other people inhibits dissonant information that might be contradictory and of a more positive nature and excludes it from subjective experience (Beck, Rush, Shaw, & Emery, 1979; Brewin, 1988; Dixon, 1981). Negative self-talk and negative core assumptions can therefore be seen as not only tied to the content of information processing or cognition but also the particular process of cognition. There is likely also a companion narrowing of focus in major depressive disorder from processing external information and maintaining alertness to processing internally generated thoughts and feelings. PET scan studies of unipolar depressed patients reported by Posner and Raichle (1995) have documented striking increases in cerebral blood flow in the frontal cortex, particularly in the left frontal lobe and left prefrontal cortex, and decreases in blood flow in the parietal and posterior temporal lobes compared with matched control subjects. Even more striking, when normal subjects are asked to contemplate sad thoughts or memories to induce a sad mood, the same left prefrontal blood flow increases are seen.

One central focus in the cognitive-behavioral treatment of depression is to understand and uncover unarticulated core assumptions that guide patients' attitudes and beliefs about themselves and act as a filter or template to assign values to everyday experiences, core assumptions that typically patients are not aware or are only minimally aware of having (Beck et al., 1979). These core assumptions may be formed in large part by the automatic processing of information, and would therefore be expected to exist without the awareness of the patient. Selective attention or selective abstraction is seen in depressed and anxious patients in which a single detail taken out of context becomes the conceptual frame for the entire experience and other sometimes more salient information is ignored. This phenomenon has a corollary in the laboratory in the automatic attention responses reported by Schneider and Shiffrin (1977) in which extensive experience with a particular stimulus in a visual search task led to subjects immediately attending to that stimulus irrespective of the complexity of the target display or the memory load of previous learning trials.

There have been a number of cognitive models that have been proposed which have elaborated upon or modified the simple, dichotomous conscious/automatic processing model (Schacter, 1989; Schacter et al., 1988) and taken into account the relationship between phenomenal consciousness, procedural memory and procedural systems, declarative/episodic memory, and executive functions, including models that embrace multiple levels and types of consciousness (Block, 1995). Since the late 1980s, numerous cognitive psychologists, neurophysiologists, mathematicians, and computer scientists have been actively constructing and elaborating a wide variety of computational, neural network models to account for the properties of consciousness, attention, memory, and higher cognitive processes (Alavi & Taylor, 1993; Gray, Buhusi, & Schmajuk, 1997; Rolls, 1997; Sun, 1997; Taylor, 1997) which may ultimately have a profound affect on the way clinical psychologists conceptualize psychopathology.

1.11.1.4 Neural Network Models and Classical Conditioning

The cognitive abnormalities characteristic of the acute phase of schizophrenic psychosis (hallucinations, delusions, fragmentation, loss of selective attention) have been conceptualized by Hemsley (1987) as reflecting a breakdown in the normal control of current information processing by disruption of the normal ability to use stored regularities as a guide to current information processing. Information that is typically processed in an automatic, nonconscious mode is then directly experienced and processed in a controlled or conscious mode (Gray, Feldon, Rawlins, Hemsley, & Smith, 1991; Schneider & Shiffrin, 1977). There is a body of evidence to suggest that a disruption in latent inhibition may account for this phenomenon.

Latent inhibition involves a classical conditioning paradigm which embraces elements o cognitive processing (Dunn, Atwater, & Kilts, 1993; Ellenbroek & Cools, 1990; Killcross, Dickinson, & Robbins, 1994; Weiner, 1990). Briefly stated, latent inhibition is the difference in efficacy of conditioning between conditioned responses evoked by pre-exposed conditioned stimuli and conditioned responses evoked by nonpre-exposed conditioned stimuli (Gray et al., 1997). Latent inhibition is decreased or blocked by amphetamine treatment in animal models (Solomon & Staton, 1982; Weiner, Lubow, & Feldon, 1984) and in human subjects (Gray, Pickering, Hemsley, Dawling, & Gray, 1992; Thornton et al., 1996). This block is reversible by injection of haloperidol (Gray et al., 1997).

A mathematical neural network model for latent inhibition has been developed (Buhusi & Schmajuk, 1996; Gray et al., 1997; Schmajuk, Lam, & Gray, 1996) which explains the cognitive symptoms seen in acute schizophrenia as the result of a blockade of latent inhibition, which is thought to be related to disruption of dopaminergic pathways beginning in the nucleus accumbens and cascading through the nucleus reticularis thalami and thalamocortical sensory relay nuclei. This model mathematically describes elements of the neural network theory of latent inhibition reported by Schmajuk (1997), which involves interaction between trace short-term memory for conditioned stimuli, attentional long-term memory for conditioned stimuli, associative long-term memory for conditioned and unconditioned stimuli, and intermediate-term memories for conditioned and unconditioned stimuli and their predictions.

1.11.1.5 Multiple-level Neural Network Models

Taylor (1997) has modeled a two-stage neural network that accounts for blindsight and neglect which is very similar conceptually to previously described conscious/automatic processing models, and has proposed a third stage to this model (the ACTION net) roughly corresponding to those cortical areas involved in executive functioning, planning, reasoning, and working memory (Brodmann areas 46 and 9). This third network is of particular interest for clinical psychologists because it involves neuroanatomic structures related to affect as well as memory and executive functioning. This model is based upon the assumption of the basal ganglia (striatum and globus pallidus) acting in a disinhibitory manner on thalamocortical recurrent loops. The ACTION net encompasses five networks (corresponding to the five frontal loops identified by Alexander, DeLong, & Strick, 1986) involving motor, premotor, and supplementary motor cortex (action sequences), limbic (affect), orbitofrontal (social), frontal eye fields (eye movements), and dorsolateral prefrontal cortex (cognition and attention). Although such a model can account for emotional tone as global activation from the limbic system and does allow for some speculation about the relationship of awareness, attention, consciousness, memory, and mood, this model does not include true integration of limbic structures, hippocampus, and related areas. Of great value, however, is the compatibility of this model with neural net models of the amygdala (Monchi & Taylor, 1995, 1997), a cortical structure that appears to have a central role in depression.

PET scan studies reported by Posner and Raichle (1995) have confirmed increases in blood flow in depressed patients in a group of cells in the amygdala and the medial dorsal nucleus of the thalamus and decreases in the caudate nucleus. Examination of PET data from a second study of patients with previous episodes of severe depression now in remission (Drevets et al., 1992) revealed blood flow increases in the amygdala with normal flow in the thalamus and the caudate, and depressed patients in remission who were also actively treated with antidepressant medication had normal blood flow in the amygdala. One role of the amygdala in cognition appears to be the affective labeling or assignment of emotional significance to experiences in long-term memory. Posner and Raichle (1995) have concluded that if the increased blood flow in the amygdala seen in depressed patients represents a disruption in normal amygdala function, this dysfunction could result in a negative emotional state assigned to all experiences resulting in depressed mood or could result in anhedonia due to the prevention of positive emotional labels assigned to events, thus tying together neural substrate, memory, cognition, and mood.

As noted earlier, left prefrontal blood flow increases have been noted on PET scans during bouts of depression and also during induced sadness (Drevets et al., 1992). If the same dysfunction hypothesis is applied to this finding, there is again the implication that cognition, memory, and mood are inter-related. It is well known that patients with structural damage to the left prefrontal area demonstrate problems with set shifting and rigidity of thinking, and describe having difficulty shifting problem-solving strategies even when they know that a given strategy is incorrect. Posner and Raichle (1995) have suggested that disruption in left prefrontal blood flow in depressed patients could lead to this same sort of difficulty in the set shifting of emotions and thoughts, leading to depressive rumination. Memory and mood are again intertwined in this process. Induced negative mood increases in left prefrontal blood flow occur coincident with the conscious association of thoughts or emotion with information held in long-term memory. Depressive rumination and the experience of incessant negative thoughts also likely involve making conscious associations in working memory with information retrieved from long-term memory. There is increasing evidence from a number of laboratories that memory and mood are inter-related.

1.11.2 MOOD AND MEMORY

1.11.2.1 Effects of Affective States

Cognitive models of negative mood (e.g., Blaney, 1986; Carver & Scheier, 1990) assume that there is a two-way interaction in cognitive processing involving memory and the regulation of emotions, that is, (i) the way an individual reacts to an affective stimulus is influenced by previously acquired memories, and (ii) the way we remember previous events is influenced by affective states that prevailed when we encountered those events.

Accumulated evidence suggests that memory processes are influenced by affective states during which memories were acquired. Since the 1970s interest in the study of memory performance in the context of negative mood and depression has been extended to laboratory settings by both cognitive and clinical psychologists. Laboratory studies usually involve experimentally inducing a depressed, elated, or neutral mood to observe its effect on memory performance.

The most frequent method used to induce depressed mood is the Velten Mood Induction procedure (Velten, 1968). In this method, mood is induced by having subjects read statements that are related to a depressed or elated mood. Other methods, such as hypnosis, imagery, or music, have also been used in the experimental induction of mood.

The effect of a depressed mood on recall has been demonstrated in experimental settings by several researchers. Leight and Ellis (1981) conducted two such experiments. In the first they investigated the effect of depressed mood on recall and transfer of training. Subjects received depressed, elated, or neutral mood induction. Then, they were trained on a letter sequence task. Twenty-four hours after the training session, subjects received the same training plus a transfer task in which they were presented with similar letters, but different spatial groupings. Overall, subjects' recall improved over trials. However, subjects in the depressed condition showed a slower improvement rate than subjects in the elation condition. In the transfer phase, subjects who previously received depressed mood induction recalled significantly fewer letters than subjects who never received depressed mood induction. Subjects who received neutral mood induction were able to recall more letters than subjects who received depressed mood induction. Depressed mood during learning in other experiments led to a lower rate of recall, a lower rate of learning of letter sequences, and lower number of letters chunked when compared with neutral mood. These findings were interpreted by Leight and Ellis as an indication of deficits in strategy learning (Leight & Ellis, 1981).

The effect of depressed mood induction may also be dependent on the nature of tasks being performed, for example, difficulty levels. In a series of experiments, Ellis, Thomas, and Rodriguez (1984) found poorer performance in free recall among subjects who received depressed mood induction vs. subjects who received neutral mood induction. A sentence completion task was used. In this task subjects were asked to complete a sentence with a suitable word. The sentences were either easy or difficult. It was found that depressed mood induction was associated with worse performance relative to performance during a neutral mood condition when subjects were asked to recall words used with difficult sentences. These authors also found that when subjects had the opportunity to process the information semantically using strategies such as elaboration, they were able to show better performance, supporting a resource-allocation model (Ellis et al., 1984). The model suggests that depressed mood has debilitating effects when the encoding processes required are demanding, difficult, and complex.

Similar conclusions were reported by Hasher, Rose, Zacks, Sauft, and Doren (1985), who indicated that memory deficits in depression may be attributed to the ineffective use of processes that require cognitive effort, such as organization of information, imagination, rehearsal, and other mnemonic devices. According to these authors, encoding operations require attentional capacity which could vary, depending on whether these operations are automatic or effortful. Depression is one of the variables thought to limit the availability of attentional capacity to perform effortful operation. Effects of depressed mood would therefore be evidenced on memory tasks that require effortful processing.

Hertel and Hardin (1990) proposed that memory deficit in depression may be attributed to the lack of cognitive initiative in depressed individuals, due to less initiative in using strategies that would help them remember. When these subjects are directed to use such strategies, memory performance should improve. To test these hypotheses, Hertel and Hardin (1990) conducted three experiments. They compared the performance of college students who received depressed mood-induction with that of subjects who received neutral mood induction. The task used in these experiments involved presenting subjects with two lists of homophones in the context of questions that reflected the less typical meaning of the homophones, and with added questions to disguise the target words. Spelling and recognition lists were constructed from the two lists of homophones. After receiving the mood induction procedures, subjects were asked to spell new and old homophones to test their memory of the old homophones. When subjects who received depressed mood induction were provided with spelling strategies, they performed as well as those who were in the neutral mood condition. In addition to their argument about the role of cognitive initiative in memory, Hertel and Hardin (1990) suggested that motivational differences between groups may exist. Subjects who received depressed mood induction may have had an excuse not to try hard (Hertel & Harden, 1990).

Failure to find memory deficits among subjects who received depressed mood induction was reported by other authors, however (e.g., Hasher et al., 1985; see also Ellis, 1985).

In summary, three main hypotheses have attempted to address the effect of depressed mood on memory. The first is the hypothesis that poor recall of information learned when subjects have a depressed mood can be attributed to poor learning strategies and lack of flexibility (Leight & Ellis, 1981). The second hypothesis indicates that people who are in a depressed mood have less resources to process information, that is, according to the resource allocation model (Ellis et al., 1984), when the task requires more processing effort, depressed subjects show poorer performance than if the task is easy. Third, Hertel and Hardin (1990) suggested that low level of cognitive initiative is what causes poor memory performance in depressed individuals.

Although the experimental literature seems to support a detrimental effect of depressed mood on memory, some questions still need to be addressed. For example, how can we differentiate between memory deficits in depression and state-dependent effect? Also, to what extent could a state-dependent effect account for the effect of depressed mood on memory? These questions seem to be important when we take into consideration the fact that studies have assessed memory performance in states that may have been different from the state when the learning took place. That is, subjects who learned materials while they were depressed and then were tested in a neutral mood may have shown a poor memory performance as a result between the mismatch of the mood state under which they recalled and the mood state under which they learned these materials.

1.11.2.2 State-dependent Learning

In this section hypotheses are discussed that address the effect of mood on memory in the context of mood state-dependent learning (e.g., Bower, 1981). One hypothesis that has been introduced to account for poor memory performance in depression is the mood-congruence hypothesis. This hypothesis has received intense attention from scientists interested in both cognitive processing and emotions (see Blaney, 1986; Matt, Vazques, & Campbell, 1992; Singer & Salovey, 1988). In its basic form, the hypothesis states that learning and remembering stimuli will be easier when there is a match between affective tone of the learned materials with the receiver's current mood state than when there is a "mismatch." That is, performance is more efficient when the recall occurs in a mood state that is similar to the mood state when the learning process took place (Singer & Salovey, 1988), possibly due to efficiency in mnemonic processing (Blaney, 1986).

1.11.2.2.1 The Network Theory of Affect

The Network Theory of Affect, introduced by Bower (1981), attempts to account for the mood-congruency effect. One of the salient factors of this theory is the importance of mood state during the learning process. The mood state is thought to serve as a cue when recalling this information and included in associations with coincident events. When this emotion unit is activated, it is used as a cue to retrieve its associated events.

According to this theory, emotions may be viewed as nodes providing points of contacts within a semantic network. Nodes related to a particular affective state are connected to other nodes of related moods that provide representations of events or connections that were formed during learning. In this manner, nodes related to anxious state, for example, are more likely to be connected with anxiety-related concepts and experiences (e.g., fear, heightened autonomic arousal, and muscular activity, etc.) During a particular prevailing mood, nodes that tap these mood states are activated, and they in turn activate other related nodes. Information congruent with prevailing mood states acquires stronger processing than discongruent information according to this theory. This leads to elaborate processing at both encoding and retrieval phases, and this in turn contributes to a better ability to recall mood-congruent than mood-incongruent information (Gilligan & Bower, 1984; Singer & Salovey, 1988). Encoding information that is congruent with current mood state supposedly occurs through nodes that connect a widespread network, which leads to a dense and elaborate representation of incoming information. In contrast, information that is incongruent with current affective state taps a relatively less elaborate and narrower scope of network representation. A similar pattern of processes occurs during retrieval, where retrieval of mood-congruent information is facilitated by biased search of relevant activated nodes, which leads to increased accessibility of mood-congruent materials. Incongruent materials therefore are relatively less accessible for search and retrieval (Bower, 1981; Riskind, 1989).

Studies conducted to evaluate this hypothesis have observed effects of mood-congruency on retrieving learned information (cf. Riskind, 1989; Singer & Salovy, 1988). These studies were done both in normal participants exposed to different types of mood induction procedures and in individuals who were classified as depressed (cf. Blaney, 1986).

Bower (1981) conducted a related study with college students. Either a depressed or a happy mood was induced using imagination guided by hypnosis suggestions. The main finding of this study supported the mood state-dependent hypothesis. Subjects who learned a word list in the depressed mood were able to recall more words if they were tested in the same depressed mood than if they were tested in the happy mood. A similar pattern of effect was found with subjects who learned the list in a happy mood.

Eich and Metcalfe (1989) went further to suggest that mood state-dependency influences one's memory for internal more than external events. Memory for external events is defined as a memory based on sensory stimuli that are brought into awareness from external sources through perception. Memory for internal events, on the other hand, is defined as those events created through mental processes such as imagination, reasoning, thoughts, etc. (Eich & Metcalfe, 1989). They conducted an experiment with college students in which they asked subjects to generate or read 16 items while in either a happy or a sad mood. The mood-induction technique used in that experiment was continuous music. Subjects were tested two days later. The mood during the testing session was either matched or not matched with the mood during learning. The findings showed a significant impairment in remembering generated items when the mood was shifted. The effect of the mood shift was significantly less when recalling read items. The authors indicated that relying on internal rather than external resources in creating and processing cues for retrieval makes memory more mood dependent.

Mathews and Bradley (1983) used Velten mood induction and music to induce either a depressed or a neutral mood. Positivity of personal memories was used as a dependent variable in that study. Subjects who received depressed-mood induction recalled more negative memories than subjects with a neutral mood. Similar results were found by Forgas, Bower, and Krantz (1984) who used hypnosis to induce either an elated or a depressed mood. Subjects were asked to remember an interaction task they had recently completed. Subjects who received a depressed mood induction recalled more unpleasant and stressful aspects of that experience than subjects who received an elated mood induction. On the other hand, subjects in the elation condition recalled more memories related to the friendly and relaxed interaction they had during that task.

Riskind, Rholes, and Eggers (1982) asked 52 undergraduates to recall four pleasant and four unpleasant life experiences. Then mood induction procedures were administered using the Velten mood induction technique. Subjects received either depressed or elated mood induction. The latencies for recalling positive and negative experiences were compared across groups. Subjects in the elation condition were significantly faster in recalling positive experiences than subjects in a depression condition.

While many studies have provided data to support the mood state-dependent hypothesis, some studies failed to provide such support (e.g., Schare, Lisman, & Spear, 1984; see Blaney, 1986). Schare et al. (1984) conducted two experiments using mood-induction procedures to induce depressed and neutral mood in a sample of college students. They used a word list learning task. Subjects were tested in either a matched or mismatched mood state compared to the original state in which they learned the lists. None of the experiments demonstrated a state-dependent effect.

1.11.2.3 Depressed Mood and Cognitive Performance

Studies conducted both in experimental settings with normal subjects and in clinical settings with depressed patients have provided evidence to support a negative effect of depressed mood on memory performance. Several theories have attempted to explain the cause of this effect. The main themes of these theories indicate that memory deficits in depression could be attributed to the mood state-dependent effect and the severity of the psychopathology of depression (Johnson & Magaro, 1987). The state-dependent theory attributes the poor performance in depressed individuals to the incongruency of the mood state under which subjects recall materials that may have been learned in a different mood state. Other theories propose that poor learning strategy and lack of cognitive initiatives may be the reason for the poor memory performance. The mood state-dependent effect is explained by pointing out the importance of contextual cues that could be provided in a particular mood state. These cues help in activating particular memory associations that are congruent with the current mood state. It is also believed that the depressed mood state activates a negative self-schema that results in selective encoding and retrieval of materials that match the mood state. Depression may affect memory through disrupting the "normal" pattern of organizing and processing information, the low level of effort exerted when encoding and storing information, or may be a product of a conservative response style and a lack of motivation to learn and process information efficiently.

The state-dependent effect may contribute to memory deficits in clinically diagnosed depression. Weingartner, Miller, and Murphy (1977) used a free association task with eight manic-depressive patients to investigate state-dependent memory in depressed patients. Subjects were asked to recall associations in either the same or a different mood state. Patients showed better performance when they were asked to recall the associations in a similar state to their learning state. According to these authors, mood state did not only affect the way these associations were encoded but also provided help to determine how subjects could retrieve the relevant cues of that state.

The mood-congruence hypothesis (discussed earlier) has also been tested in studies conducted with depressed patients (e.g., Clark & Teasdale, 1982; Gotlib, 1981; Matt, Vazquez, & Campbell, 1992; Tham et al., 1997). In some of these studies, depressed individuals were provided with different types of feedback about their performance. Subjects given negative feedback were able to remember these tasks better than when tasks were followed by positive feedback (Calev, 1996; Channon & Baker, 1996; Singer & Salovey, 1986). Johnson, Petzel, Hartney, and Morgan (1983) asked depressed and nondepressed students to recall the content of tasks on which they had been successful or unsuccessful. Both depressed and nondepressed groups had a similar amount of success or failure, since the success or lack of success was under the experimenters' control. The authors found that depressed subjects were more able to recall failure tasks than nondepressed subjects. Nondepressed subjects compared with depressed subjects remembered more "successful tasks."

Another type of material used to investigate the mood-congruence hypothesis is self-reference statements or positive/negative materials. Clark and Teasdale (1982) conducted a study with depressed patients who were tested twice on different occasions to tap the diurnal variation in their mood. Materials with a depressing tone remembered while subjects were more depressed were compared with similar materials remembered while patients were in a less intense depressed mood. The results showed that more "unhappy" and less "happy" materials were remembered while patients were in the more depressed time. When patients were in the less depressed occasion, the results were reversed with the more happy and less unhappy materials remembered by patients.

Why do depressed patients show memory deficits? Weingartner, Cohen, Murphy, Marello, and Gerdt (1981) investigated the strategies depressed individuals use to process information. In one experiment they found that depressed patients did not use elaborative strategies to encode information. Based on this and other experiments, these authors concluded that depressed patients were unable to benefit from encoding operations that would help them organize material and remember it. This conclusion indicates that a processing deficit may exist in depressed individuals. This deficit may be attributed to ineffective organization in encoding (Johnson & Magaro, 1987). This conclusion is similar to the conclusion reported by researchers who have investigated this issue in laboratory settings (Hasher et al., 1985).

Some authors have argued that short-term memory deficit may be the cause of overall memory deficits in depression (e.g., see Koh & Wolpert, 1983). This argument was based on the finding that deficits were found in short- but not long-term memory in depressed patients compared with nonpsychiatric patients (see Johnson & Magaro, 1987). However, inconsistent results were also reported. For example, Henry, Weingarten, and Murphy (1973) found long- but not short-term memory deficits among a group of bipolar and unipolar depressed patients when they performed tasks while they were depressed compared with their performance when they were not depressed.

Other factors not related to cognitive processes were also suggested to explain memory deficits among depressed individuals. One of these factors is the apparent lack of high motivation in at least some depressed individuals (Cohen, Weingartner, Smallberg, Pickar, & Murphy, (1982). Effort is necessary to perform memory tasks that require elaboration and active rehearsal (Johnson & Magaro, 1987), but depressed individuals seem to lack the ability to maintain a high level of persistency (Miller, 1975) which may affect their performance. To investigate the relationship between memory deficit in depression and motivation, Cohen et al. (1982) conducted a study with 11 depressed patients and five normal controls. They demonstrated a negative correlation between memory performance and severity of depression and also between severity of depression and the tendency to exert effort to perform this task. A positive correlation was found between memory performance and the effort spent to perform the task.

Another factor that has been suggested to account for memory deficits among depressed individuals is the tendency to be cautious in giving responses when they are examined. Johnson and Magaro (1987) argued that this is not a cognitive deficit but rather a tendency to respond with a certain level of confidence that may make depressives conservative in their response to testing situations. According to this argument, memory may be intact and the information needed may be available in the memory store. The problem is instead related to the restrictions depressed people impose on what they should report, and how sure they are regarding the "accuracy" of the information they report (see Johnson & Magaro, 1987).

Theories that focus on the state-dependent effects and the mood-congruency effects assume that learned materials are available in memory. The failure to remember these materials could be attributed to the lack of sufficient retrieval cues that subjects used when they encoded these materials. When subjects are in a mood state that is similar to the mood state in which they learned materials, they may benefit from this mood by getting better access to associations in memory. Similarly, materials that are congruent with a particular mood may be processed more effectively because of the availability of many cues that could enhance the encoding process and retrieval. One could also infer that subjects who learn materials which are congruent with their mood would tend to have a lower threshold for encoding these materials and for chunking them in units that are easily accessible.

Theories that emphasize the effect of poor encoding strategies and low effort to process information effectively assume that materials cannot be remembered because they were not learned or they are not available in memory. Therefore, a depressed individual who could not recall particular events from a previous day may not have this information in their memory any more, or may not have stored this information in the first place. On the other hand, a state-dependent theorist may argue that this individual may have been in a different mood when they learned about these events, and that they would more likely to remember them if they were in a similar mood state. All these inferences still need to be directly tested.

In summary, to understand the effect of depressed mood on memory, it may be useful to consider both the state-dependent hypothesis and theories that focus on encoding deficits. It is also important to consider the severity of depressive symptoms as a contributing variable. That is, while the mood state-dependent phenomenon may explain some of the findings when using mood induction with normal subjects, memory deficits among severely depressed individuals involve more complex mechanisms. These mechanisms may include poor learning strategies, less effort exerted in encoding information, lack of cognitive initiative to use strategies, and a conservative response style, all of which are associated with the level of severity of depression.

1.11.3 COGNITION–EMOTION INTERACTIONS

1.11.3.1 Cognitive Factors

The cognitive model of behavior emphasizes the contribution of cognitive functioning to any emotional distress. Cognitive processes such as perception, reasoning, and thinking, represent substrates for mental health. Faulty processes, such as misperception of social situations, negative thinking, evaluation of oneself without sufficient data, and patterns of attributing failures to oneself, contribute to negative affect and increase the risk for psychopathology.

In addressing cognitive contributions to emotional behavior, it is important to distinguish between various terms used to describe cognition, including such terms as cognitive processes, cognitive structure, cognitive content, and cognitive products (Ingram & Kendall, 1986). Cognitive processes are the operations performed to receive, store, transform, and coordinate information. Cognitive structures are the organization of information. Cognitive structures influence our perception of surrounding situations and events. The cognitive content is the actual information that we store and process, including what we tell ourselves via self-talk. Cognitive products are those outputs that result from the various cognitive processes, including conclusions one may reach (Kendall & Hammen, 1995).

Relating a certain style of processing to a mood state may be better accomplished with the distinction between these terms in mind. For example, the way one explains an event (i.e., attribution) is a cognitive product that may result from various cognitive processes. The way these cognitive processes are conducted would depend on the content and the cognitive structure for that person. The attribution therefore may vary in persons who are anxious, depressed, or in a happy mood.

This pattern of interaction between cognition and emotion is well accepted. However, the question of whether cognition and emotion should be considered as separate systems is still strongly debated. Gray (1990) articulated a model in which he described neurobiological systems that may underlie both cognition and emotion. He argues that the overlap between the two systems is very strong, to the extent that it may be impossible to distinguish between substrates of cognition and substrates of emotion.

Other emotion researchers, on the other hand, argue that separate pathways can be delineated to account for cognition and emotion, respectively (Izard, 1992; LeDoux, 1992; Panksepp, 1992). This school of thought also acknowledges the close interaction between these two brain systems. Indeed, the separation between the two tends to be blurred in certain conditions (LeDoux, 1989, 1992). One example of such conditions is when rapid unconscious computations of a situation that is survival-related have to take place. Such a situation, while it involves the emotional system, is also accepted as a type of cognition. The distinction is more pronounced when the process involves rational, analytic processing not related to basic survival efforts.

Most of the actively researched theories in clinical psychology deal with both cognitive and emotional aspects of behavior in a complementary fashion. To this end, several cognitive theories of emotion have been proposed and have enjoyed significant attention. These theories generally agree that a stimulus when perceived is linked or elaborated on by previous experiences, and based on this elaboration an emotional response ensues. An emotional response can be discrete or a mix of multiple emotions, with the most widely accepted emotions being anger, happiness, sadness, fear, and disgust (Ekman, 1973).

1.11.3.2 Cognitive Appraisal

Lazarus (1991) discussed the generation of emotions as a result of cognitive appraisals. Lazarus' model of emotion proposes that events we encounter are thought over, and responses to each event are selected consciously from a range of options. Lazarus' model of emotion emphasizes perceived control over the environment as a critical determinant of the psychological impact of events.

Lazarus and Folkman (1984) have described how interactions with the environment generate emotions, noting the continuous evaluation process of events that people encounter. This evaluation, termed appraisal, is classified as familiar or unfamiliar and as threatening or nonthreatening. Various cognitive and behavioral strategies to deal with these challenges are also developed. Emotions experienced in these situations provide information about whether these strategies are successful and motivate additional behaviors.

According to Lazarus and Folkman (1984), there are two stages of appraising situations. The first, called primary appraisal, involves evaluating the threatening aspects of the situation and is geared towards developing a plan to deal with the situation. The second stage, called secondary appraisal, involves evaluation of the options available for dealing with the challenging situation. This appraisal determines the cognitive, behavioral, emotional, and physiological responses to environmental events.

Factors that influence primary appraisal of an event include one's beliefs about how the world should work and one's own commitments to given courses of action. Events that are discongruent with beliefs and commitments in life are perceived as threatening events. When an event is appraised as threatening, it invokes a simultaneous emotional reaction. This reaction signals alarm and motivates various behavioral and psychological responses along with physiological activation to enable one to undertake planned coping behaviors (Lovallo, 1997). On the other hand, the process of secondary appraisal focuses on efforts that might be used to "cope" with situations at hand. The efforts may include coping strategies and coping behaviors, and can be overt or covert.

Lazarus and Folkman (1984) classify coping responses into two types: problem-focused and emotion-focused strategies. Problem-focused strategies are directed at the source of the problem itself, with efforts geared towards gaining information, changing events, and modifying beliefs and commitments. Problem-focused strategies enhance one's awareness and knowledge about the situation, and therefore enhance one's cognitive and behavioral coping resources. These resources help in reducing the threat value of the event. On the other hand, emotion-focused strategies direct efforts towards changing the psychological reactions to an event. This change is geared towards limiting the degree of emotional distress that may be caused by a stressful event.

1.11.3.3 Attribution and Affect

For a subjective emotional state to take place, both peripheral and central inputs contribute. In the early 1960s Schacter and Singer (1962) described how these sources of physiological change contribute to emotion. According to these authors, an emotional response is a product of two factors: perception of peripheral physiological arousal and the label given to this arousal. The specific emotional response is then based on the cognitive appraisal of the source of arousal, whether it is attributed to external or internal stimuli.

In their original study, Schacter and Singer (1962) independently manipulated epinephrine and autonomic activation on the one hand and manipulated attribution of the source of arousal on the other. Subjects were injected with either epinephrine or placebo to induce physiological arousal. Subjects who were injected with epinephrine received no information or were misinformed. Subjects were not told that they would feel either happy or angry. The prediction was that subjects who were aroused but were not able to attribute this arousal to epinephrine would express stronger motions. The findings supported the model, although not definitively.

In the late 1970s, subsequent studies showed that increased sympathetic arousal contributes to evaluating situations in a negative manner (Marshall & Zimbardo, 1979). However, the role of Schachter and Singer's (1962) two factors in predicting interactions between situations and perception of autonomic state has also been criticized (Erdmann, 1983). One area of criticism has been the supposed overestimation of the role of arousal in the interaction between arousal and attribution. The role of arousal seems to be more accepted when it is perceived as providing feedback which intensifies the experience of emotion (Reisenzein, 1983).

1.11.3.4 Psychobiological Substrates

The study of cognition–emotion interaction may be further advanced by understanding the biological underpinning of such interactions. Recent efforts in cognitive neuroscience research have generated excitement about the value of a multidisciplinary approach to study emotion and cognitive processes. Indeed, most of the prominent theoretical accounts of emotion–cognition interaction acknowledge the importance of incorporating physiological knowledge in understanding this interaction and propose neurobiological substrates for it. The next section includes a brief discussion of an account of this interaction.

The way cognition influences our emotion, or emotion influences our cognition, is influenced by processing that takes place at multiple physiological and anatomical levels. Gray (1990) discusses the prominent role of the hippocampal and amygdalary systems. He proposes the presence of three basic emotional systems that include a behavioral approach system, a fight-or-flight system, and a behavioral inhibition system. Each system is influenced by a specific neuroanatomical structure. The hippocampal system (particularly the septal area) is the biological substrate for behavioral inhibition, while the amygdalary system is the substrate of the fight-or-flight system. Both biological systems, according to Gray (1990), may be involved in the behavioral approach emotion. Other systems, such as brainstem and neocortical systems, may also be involved in organizing approach behavior.

The anatomical delineation of these hypothesized interactions can be further appreciated by reviewing the interconnections of the hippocampus, the amygdala, and other emotion-related brain systems. The hippocampus has various connections with cortical and brainstem structures. The hippocampus is also a primary structure involved in the conscious recollection of past information (Squire et al., 1992), and in processing spatial, contextual, and configural information (Eichenbaum, Otto, & Cohen, 1992; Rudy & Sutherland, 1995). The hippocampus provides inputs to structures involved in emotional processes, particularly the amygdala (LeDoux, 1993). The hippocampus interacts extensively with the amygdala and other stress effectors, especially the hypothalamic–pituitary–adrenocortical (HPA) system (Chrousos & Gold, 1992). Projections from the forebrain to the hippocampal formation also indicate involvement of the latter in attention processes.

Advancing this information to explain abnormal behaviors, Gray (1990) notes that lesions in the septal area produce behavioral effects that resemble those seen as a result of antianxiety drugs. Based on this finding, it was proposed that anxiety may be related to functional changes in the septohippocampal system (Gray, 1990). Additional cited support for this hypothesis includes the heightened sensitivity of this system for signals predicting punishment and innate fear cues. This sensitivity results in inhibition of behavior and enhancement of autonomic activity.

The integration of emotion and cognition also involves the amygdala. The amygdala is viewed as the primary brain structure involved in emotional behavior and expression (Aggelton, 1992). It is a key structure in the memory of emotionally-relevant information (LeDoux, 1993). It plays a significant role in providing emotional meaning to a stimulus. Research with brain damaged humans has demonstrated the importance of an intact connection of the amygdala with other parts of the limbic system for the perception and expression of emotion (Bechara et al., 1995). For example, bilateral lesions in the amygdala prevent conditioned autonomic responses to aversive stimuli, in spite of an intact memory of declarative information about the presented stimuli (Bechara et al.).

Like the hippocampus, the amygdala receives projections from and provides inputs to a wide range of sensory and cortical structures (Derryberry & Tucker, 1992). The amygdala exerts a facilitatory influence on HPA activity (Feldman, Conforti, & Weidenfeld, 1995) and on the autonomic and behavioral apparatus associated with stress (Weiss et al., 1994). It projects output to various brainstem structures including the pontine reticular formation and the brainstem aminergic nuclei. The pontine reticular formation facilitates the behavioral and biological response. The locus coeruleus and raphe nuclei in the brainstem send their aminergic fibers to the cortex, limbic structures, and to the spinal cord. This influences the level of cortical activation and the interaction of descending cognitive output, emotional experience, and autonomic and motor outflow.

Cortical structures also influence both cognition and expression of emotion. An important area of research that has received increasing attention is the effect of cortical hemispheric lateralization on emotional perception and expression. The right hemisphere is generally thought to play a more prominent role in nonverbal, spatial, and affective processes than the left hemisphere. The latter is more prominent in verbal, sequential, and logical reasoning (Borod, 1992). The dominance of the hemisphere in emotional processing seems to be specific to negative affect, although it has been suggested that this hemisphere is also involved in emotional perception and expression, regardless of the specific affective tone (Davidson, 1993). Relative to the left hemisphere, the right hemisphere has more direct connections with subcortical structures involved in emotional expression (Gainotti, Caltagirone, & Zoccolotti, 1993).

A further delineation of the role of right vs. left hemisphere assumes that the left hemisphere is more involved in the expression and experience of positive affect, whereas the left hemisphere is more involved in the perception of positive affect and both the perception and expression of negative affect (Borod, 1992). Related to this, the left hemisphere may be responsible for affect-related approach behavior (Fox, 1991). This specialized dominance seems to occur particularly in the anterior area of the cortex, whereas the posterior region seems to be associated with cognitive-perceptual processes (Davidson, 1993). Consistent with this hypothesis, studies of right hemisphere lesions have documented a steeper decline in sympathetic responses to stimuli compared with those noted in studies of left-hemisphere damage (Tucker & Williamson, 1984).

This line of work has implications for understanding the association of mood states, cognitive performance, psychopathology, and frontal lobe activation. Intra- and interhemispheric differences in EEG activity have been found to be related to depressed mood disposition and withdrawal in clinical and nonclinical populations (Davidson, 1993; Fox, 1991).

1.11.4 COGNITIVE MODELS OF ANXIETY

1.11.4.1 General Cognitive Models of Anxiety

Behavioral techniques such as desensitization, flooding, and response prevention have been historically effective treatments for anxiety disorders, but recent cognitive-behavioral therapy methods, in combination with behavioral techniques, have demonstrated superior efficacy (Brewin, 1996). Purely cognitive techniques, however, may not be superior to purely behavioral methods (Rachman, 1993). There are some commonalities in the mechanisms underlying the effectiveness of cognitive-behavioral therapy for anxiety.

Rachman (1990) noted several ways of acquiring fear, either by direct exposure to aversive or traumatic stimuli, by observing others displaying fear, or by hearing or reading verbal or written information about fear and danger. Similarly, catastrophic misinterpretation of symptoms based on ignorance or incorrect beliefs can lead to panic disorder (Clark, 1988) and prolonged post-traumatic stress disorder (PTSD; Ehlers & Steil, in press). In these latter cases, the representations in memory giving rise to feelings of fear may be verbally accessible and can therefore be altered by new verbal information. Many contemporary clinical approaches are based upon appraisal theories of emotion such as those of Lazarus (1991) and Weiner (1985), which emphasize the role played by conscious assessments of the meaning and cause of internal and external stimuli and of the coping resources available to deal with any prospective threat.

Verbally accessible knowledge may also be subject to nonconscious automatic processing, particularly when it is primed by a previous stimulus. As noted by Clark (1988), catastrophic thoughts may come to mind so rapidly that patients are unaware of the interpretive process, which may in part explain panic attacks triggered in sleep. Also, almost by definition, automatic processing does not differentiate between verbally and situationally accessible knowledge (Brewin, 1996).

1.11.4.2 Specific Phobias

Specific phobias are among the most common circumscribed anxiety disorders. Phobias may be based on unconscious representations, independent conscious representations, or some combination of the two. The particular pattern of response to different treatment approaches depends on the nature of the memories.

Most cognitive-behavioral treatment approaches to phobia include an emphasis on helping individuals undermine habits of cognitive or behavioral avoidance and to achieve habituation to a feared stimulus. More sophisticated cognitive approaches emphasize the importance of factors such as attention and memory (Dickinson, 1987) and conscious revaluation of the unconditioned stimulus (Davey, 1992, 1995).

Lang (1979) proposed that the representations underlying phobias, which he termed fear memories, contain three kinds of information: details of the location and physical characteristics of the feared situation (stimulus elements); details of the verbal, physiological, and behavioral responses that occurred in the situation (response elements); and an interpretation of the stimulus and response elements and their significance for the individual (meaning elements). Foa and Kozak (1986) propose that cognitive therapy for phobias involves changing information in the fear memory by first activating the memory and then incorporating new experiences arranged by the therapist into that memory (such as within- and between-session habituation to the feared situation).

Brewin's (1989) dual representation theory proposes that in therapy new memories are laid down as "situationally accessible knowledge." Patients in treatment are encouraged to expose themselves to as good an example as possible of what they fear and to experience the same sensations. In therapy, however, this experience culminates in mastery and the habituation of fear rather than in avoidance. Patients exposed to these feared situations are subsequently more likely to access this new memory than they are to access the original memory containing more fearful response elements and a more threatening meaning (Brewin, 1996).

1.11.4.3 Panic Disorder and Agoraphobia

According to Clark (1988), panic arises from specific catastrophic misinterpretations of sensations such as tachycardia or dizziness. Treatment therefore involves a variety of techniques aimed at eliciting and challenging these misinterpretations. In contrast, Wolpe and Rowan (1989) suggest that at least in some patients there are no misinterpretations and that panic is instead the result of conditioning.

Panic provocation studies using CO_2 inhalation provide an analogue model for naturally occurring panic attacks (Gorman et al., 1988; Sanderson & Beck, 1989) and demonstrate the importance of cognition in mediating the emotional response to physiological sensations secondary to hyperventilation. CO_2 inhalation studies suggest that panic attacks are more than mere conditioned responses to a conditioned stimulus. Panic disorder patients who panic in this situation experience cognitive symptoms (e.g., fear of dying or losing control) in addition to somatic symptoms, whereas nonpanicking panic disorder patients and controls experience only somatic symptoms (Sanderson & Beck, 1989). Direct manipulation of cognitive variables also affects the likelihood of a CO_2 provoked panic attack. Rapee, Mattick, and Murrell (1986) found that panic disorder patients who were told which somatic symptoms to expect prior to the inhalation of CO_2 reported significantly less intense somatic symptoms and catastrophic thoughts compared to those who were provided no such explanation. Sanderson, Rapee, and Barlow (1989) found that panic disorder patients who believed they could control the flow of CO_2 (although in reality they could not) experienced significantly less anxiety and panic during the inhalation compared to those who did not believe this.

Rachman, Levitt, and Lopatka (1987) provide further evidence for the importance of cognition in the production of panic attacks. "Threat-relevant" cognitions accompanied the majority (74.6%) of panic episodes reported by panic disorder patients. When combinations of two or three symptoms were examined, Rachman et al. (1987) found that "no panics were reported unless there was an associated cognition." On the other hand, when the same combination of symptoms were accompanied by a fearful cognition, the episode almost invariably resulted in panic.

The biopsychosocial model of panic disorder (Barlow, 1988) conceptualizes the initial panic attack as a misfiring of the fear system under stressful life circumstances in physiologically and psychologically vulnerable individuals. Physiological vulnerability is defined as an overly reactive or labile autonomic nervous system which may in part be determined genetically, consistent with the higher concordance rates of panic disorder seen within first-degree relatives (Crowe, Noyes, Pauls, & Slymen, 1983; Moran & Andrews, 1985; Torgersen, 1983). Psychological vulnerability is conceptualized as a set of danger-laden beliefs about bodily sensations and about the world in general. The concepts of uncontrollability (Barlow, 1988) and anxiety sensitivity (Reiss, Peterson, Gursky, & McNally, 1986) are central in this model of panic disorder and panic disorder with agoraphobia. Persons who panic tend to have stronger beliefs and fears of physical and mental harm arising from specific bodily sensations (Chambless, Caputo, Bright, & Gallagher, 1984; Clark et al., 1988; Holt & Andrews, 1989).

Barlow (1988) and Wolpe and Rowan (1989) emphasize the traumatic nature of the first panic attack and the consequent learning that takes place. The traumatic nature of initial panic attacks is apparent in the relatively high frequency of individuals seeking emergency medical help due to their misappraisal of the panic attack as signaling impending death. Consequently, fearful associations are likely to develop with the situational context and the physical sensations present at the time of the first panic. The concept of learning fearfulness of bodily sensations is based on Razran's (1961) account of interoceptive conditioning. This is a form of conditioning that is relatively resistant to extinction, and interoceptively conditioned fear responses are not dependent on conscious awareness of triggering cues. Panic attacks may seem to be uncued when in fact they are triggered by benign and subtle fluctuations in physical state (Barlow, 1988). For example, a slight change in heart rate might cue or elicit fearfulness in some patients in anticipation of experiencing a panic attack. Such internal fear cues can trigger more abrupt, and less predictable fear, which in turn leads to greater anticipatory anxiety about the recurrence of fear (Craske, 1991).

Anxious anticipation of panic increases the likelihood of its occurrence, since anxious arousal increases the availability of arousal sensations that have become conditioned cues for panic, in addition to increasing the degree of attentional vigilance for such cues (Barlow, 1988). In this manner, a cycle is established between panic and anxious apprehension about panic. It is hypothesized that sex-role behaviors and expectations and associated situational demands influence the degree to which agoraphobia emerges as a behavioral style in response to panic anticipation (Craske & Barlow, 1988).

Craske and Barlow (1993) describe a 15-session cognitive-behavioral treatment protocol that is specific for panic disorder and agoraphobia involving self-monitoring, cognitive restructuring, breathing retraining, applied relaxation, interoceptive exposure, and *in vivo* situational exposure.

1.11.4.4 Social Phobias

Beck, Emery, and Greenberg (1985) describe a cognitive model of anxiety and phobias, including social phobias. The schema, the basic cognitive structure that guides the processing of information, serves as the core concept in this cognitive model. Schemas are sets of rules that classify, prioritize, and interpret incoming information to the person as well as facilitate the retrieval of relevant information from memory. Schemas are grouped into modes that create a cognitive set or processing bias that operates across situations. Anxiety-disordered individuals, according to Beck et al. (1985), typically function in the vulnerability mode, that is, the individual sees the world as a dangerous place in which they must constantly be vigilant to potential threat. As a result, neutral or mildly positive cues are misinterpreted negatively, while positive or safety cues are discounted or ignored. Memories of past successes or available coping resources are also underestimated or overlooked.

Social phobics are hypervigilant to cues that denote the possibility of negative evaluation by others. These cues may be situational, interpersonal, or internal. Social phobics devote excessive attentional resources to the detection of potential social cue threats (Hope, Rapee, Heimberg, & Dombeck, 1990; Mattia, Heimberg, & Hope, 1993; Smith, Ingram, & Brehm, 1983).

The self-presentational model of social anxiety (Leary, 1988; Schlenker & Leary, 1982) states that social anxiety occurs when an individual desires to make a particular impression on others but doubts they will be successful in doing so. How one might appear to others must be of importance and there must be apprehension about one's ability to engage in sufficient impression management. Individuals may be excessively motivated to create a particular impression because of a high need for approval, and several factors may affect a person's expectations about meeting impression management goals, including perceived or actual deficits in social skills, low self-esteem, and low outcome expectancies.

Hope and Heimberg (1993) describe a treatment protocol for cognitive-behavioral group therapy of social phobia which consists of simulated exposures to feared situations, cognitive restructuring, and homework assignments for *in vivo* exposure.

1.11.4.5 Post-traumatic Stress Disorder

Creamer, Burgess, and Pattison (1992) have developed a cognitive processing model for reactions to trauma that includes a feedback loop among intrusions, avoidance, and symptom levels. The model views the successful processing or integrating of the trauma as central to successful recovery. This model sees the cognitive processing mechanisms involved in recovery over time as occurring in five stages. Stage 1 is objective exposure. The major fact at this stage is the severity of the traumatic stressor. Stage 2 is network formation. This is determined primarily by subjective perceptions and meaning attached to the experience. Stage 3 is labeled intrusion. During this stage, the memory network is activated in an attempt to process and resolve the trauma-related memories. Stage 4, avoidance, is characterized by the use of escape and avoidance as coping strategies in response to intrusions. During stage 5, which is labeled outcome, recovery is achieved through network resolution processing. Such factors as pretrauma functioning and biological processes are not incorporated into this model. Chemtob, Roitblat, Hamada, Carlson, and Twentyman (1988) have developed a similar information processing model of PTSD.

The Jones and Barlow (1992) model considers PTSD as a variant of other anxiety disorders, particularly panic disorders, with anxious apprehension modifying or amplifying a predisposition to respond to stress with chronic autonomic overarousal or noradrenal lability. Anxious apprehension involves distorted processing of information along with extremely negative affect. The individual perceives the traumatic event and subsequent reexperiencing as unpredictable, uncontrollable, aversive events and reacts with chronic overarousal, hypervigilance, and narrowing of attention. This sets up a feedback loop in which hyperarousal, hypervigilance, and narrowing of attentional focus increase intrusive thoughts and reexperiencing.

Foa, Steketee, and Rothbaum (1989) and Rothbaum and Foa (1992) have articulated a cognitive-behavioral treatment approach for PTSD based on Lang's (1979) model of the emotional processing of fear, proposing that PTSD results from inadequate processing of the trauma stimuli, responses, and the meaning associated with them. Foa et al. (1989) suggested that traumatic events create especially large and complex fear networks that are activated readily because of the large number of interconnections formed through conditioning and generalization. Associations that were once considered neutral or safe may now be connected with fear. This leads to a sense of unpredictability and uncontrollability that is important in the development and maintenance of PTSD. Treatment requires activation of the fear memory and incorporation of new information incompatible with the current fear structure, so that new memories are formed. The PTSD victim is asked to recall the initial trauma in detail and helped to process the memory until it is no longer intensely painful. This is combined with *in vivo* exposure to feared but objectively safe stimuli. Cognitive processing therapy (CPT) is a therapy model developed to treat the specific symptoms of post-traumatic anxiety and stress disorder in victims of sexual assault (Calhoun & Resick, 1993; Resick & Schnicke, 1992), which combines cognitive restructuring with exposure-based therapy.

1.11.4.6 Generalized Anxiety Disorder

In generalized anxiety disorder (GAD), there are often complex sets of negative beliefs about the self or the external world that are activated in a variety of situations, and the underlying memory representations are thought to contain more abstract information than the specific memories underlying phobias. These representations are thought to consist of summaries of numerous aversive experiences produced by complex unconscious computations that abstract common meanings from repeated experiences (Brewin, 1996). In contrast to other anxiety disorders, evidence suggests that GAD has a more characterological presentation, although fluctuations in the course of GAD are often noted corresponding to the presence or absence of life stressors.

Patients with GAD often present with a lifelong history of generalized anxiety. For example, several studies have found that many patients with GAD cannot report a clear age of onset or report an onset dating back to childhood (Anderson, Noyes, & Crowe, 1984; Barlow, Blanchard, Vermilyea, Vermilyea, & Dinardo, 1986; Butler, Fennel, Robson, & Gelder, 1991; Cameron, Thyer, Nesse, & Curtis, 1986; Noyes, Clarkson, Crowe, Yates, & McChesney, 1987; Noyes et al., 1992; Rapee, 1985; Sanderson & Barlow, 1990).

Two types of cognitive distortions involved in excessive anxiety and ruminative worry are probability overestimation and catastrophic thinking (Brown, O'Leary, & Barlow, 1993). Cognitions involving probability overestimation are defined as those in which a person overestimates the likelihood of the occurrence of a negative event (that is actually unlikely to occur). Catastrophic thinking is defined as the tendency to view an event as intolerable, unmanageable, or beyond one's ability to cope with successfully. Also in the category of catastrophic thinking are thoughts that involve drawing extreme conclusions or consequences to minor, unimportant events. Cognitions reflecting a strong need for perfection or personal responsibility (and of drawing extreme negative conclusions of the consequences of not being perfect or responsible) fall under this category as well. Worry has been conceptualized as a negative reinforcer that serves to dampen physiological reactivity to emotional processing (Borkovec & Hu, 1990). In a sense, worry may serve to hinder complete processing of more disturbing thoughts or images.

Two principal targets of treatment intervention for GAD are excessive, uncontrollable worry and its accompanying persistent overarousal and central nervous system symptoms. These cognitive and somatic features have been addressed most frequently with cognitive therapy and some form of relaxation treatment. Targeting worry directly via an exposure-based paradigm has emerged as a potentially effective treatment component (Craske, Barlow, & O'Leary, 1992). Brown et al. (1993) outline a 12- to 15-session protocol for GAD involving self-monitoring, progressive muscle relaxation, cognitive restructuring (including worry exposure), worry behavior prevention, problem-solving, and time management.

1.11.5 COGNITIVE MODELS OF DEPRESSION

1.11.5.1 Cognitive Distortion Models

The cognitive distortion model of depression is based upon the assumption that erroneous beliefs and maladaptive information processing play a key role in the onset and maintenance of depression (Beck, 1963, 1967, 1976). Individuals at risk for becoming depressed are said to possess a latent cognitive schema that becomes active under conditions of stress. Once activated, this schema influences the way information is processed and affects the nature of specific beliefs about the self, one's experiences in the world, and one's future (Kovacs & Beck, 1978).

In this model, a negatively biased cognitive set constitutes the core process (Beck, 1976; Beck et al., 1979). When faced with stressful life events, depression-prone individuals experience negative thoughts. These thoughts typically consist of negative views of the self, the world, and the future which precede the experience of depression. To such individuals, the environment presents overwhelming obstacles that guarantee personal failure. These depressed patients view the environment as overwhelming, as presenting insuperable obstacles that cannot be overcome, and as continually resulting in failure or loss. They view the future as hopeless and they believe that their own efforts will be insufficient to change the unsatisfying course of their life.

Three major concepts in this model are the cognitive triad, negative schemas, and cognitive errors. The cognitive triad consists of three patterns of negative ideas and attitudes about the self, world, and future that are said to characterize people who are depressed.

The depressive view of the self includes negative evaluations of abilities and worth as a person. In comparing themselves to others, depressed individuals are conceptualized as seeing themselves as defective and inadequate. This negative self-view pervades virtually all aspects of life and results in an overwhelming sense of worthlessness. Beck and his associates (Beck, 1963; Beck & Hurvich, 1959; Beck & Ward, 1961) observed that clinical interviews with depressed patients were dominated by themes of failure and personal inadequacy. Hollon and Kendall (1980) have documented the association of depression with self-reports of frequent automatic thoughts of personal inadequacy and maladjustment. In laboratory tasks, depressed individuals consistently expected their performance to be worse than average, and when asked to provide evaluative self-ratings, people who were depressed selected trait descriptors with significantly negative connotations.

Second, Beck (1963, 1967) observed that depressed persons perceived their daily experiences as permeated by themes of loss and stress. Their view of the world appeared to be that life is an overwhelming burden filled with excessive demands and daily defeats, and consequently depressed individuals experienced a pervasive sense of helplessness. Depressed patients rate even the most common interpersonal problems as significantly more difficult to deal with than people who are not depressed (Funabiki & Calhoun, 1979), and individuals who have apparently recovered from depressive episodes continue to dream about problems and losses more often than nondepressed people (Hauri, 1976).

The third component of the cognitive triad is the depressive view that the future is hopeless and that the current unpleasant condition will continue without any possibility of improvement. Researchers have consistently found a strong association between depression and measures of pessimism and hopelessness (Erickson, Post, & Paige, 1975; Gottschalk, 1974; Hollon & Beck, 1979). In addition, depressed persons tend to dwell on past failures rather than to look toward future possibilities (Miller, 1975; Shaw, 1979).

According to this model, most symptoms of depression may be attributed to the negative cognitive triad. Sad affect and depressed mood are assumed to be a direct consequence of negative cognitions. In support of this, Velten (1968), Frost, Graff, and Becker (1979), and Teasdale and Fogarty (1979) have demonstrated that having subjects concentrate on depressive thoughts can induce a depressed mood.

Other symptoms of depression are also viewed as a consequence of the depressed person's cognitions in this model. Decreased activity levels, for example, are thought to be the direct result of the pessimism and hopelessness that characterize a depressed person's thoughts.

A major component in Beck's model of depression is the concept of schemas. Schemas are stable, long-standing thought patterns representing a person's generalizations about past experiences. According to this model, schemas serve to organize information from past circumstances that appears to be relevant to a current situation and direct attention selectively to particular aspects of a situation. Depression-prone individuals develop schemas consisting of stable but negative views of themselves and their experience. The schemas that produce depression often involve the perception of a personal loss or damage to one's self-worth. Depression-prone individuals tend to respond to life circumstances in a fixed, negative manner, independent of what is occurring in their environment due to the negative nature of their internalized schemas.

Once learned, a schema is usually out of awareness until the person encounters circumstances reminiscent of the conditions in which the schema was learned. At that point, the person may employ the schema to organize and process information about the situation. Depressive schemas predispose the individual to distort events so as to maintain a negative view of the self, the environment, and the future.

Derry and Kuiper (1981) compared the memories of depressed and nondepressed patients for adjectives that were either meaningful to them or descriptive of them. Depressed patients recalled more adjectives with depressive content that they had judged descriptive of themselves than nondepressed patients or normal controls, which lends support to the hypothesis that depressed individuals utilize a negative schema for processing and retaining personal information. Abramson, Alloy, and Rosoff (1981) examined the hypothesis that negative self-schema may prevent depressed persons from generating hypotheses about the contingencies between their behavior and possible outcomes, and found that when asked to develop generalizations for exerting control, depressed subjects underestimated the potential amount of control available to them.

In this model, stressful life events activate depressive schemas. These schemas, in turn, are responsible for distortions in the way depression-prone individuals perceive and interpret experiences. Although interpretations based on depressive schemas often result in conclusions that are logically inaccurate, depression-prone individuals maintain depressive schemas even in the face of evidence disproving their validity. Beck et al. (1979) contend that the maintenance of depressive schemas is a consequence of a faulty system of information processing in which the individual draws illogical conclusions due to basic cognitive distortions. According to Beck, depressed individuals make these errors in logic when evaluating experiences. Consequently, their thoughts are characterized by extreme, negative, categorical, absolute, and judgmental cognitions. Their distorted perceptions serve to maintain negative views of themselves, the world, and their future.

Beck (1967, 1976) has described a number of specific types of cognitive distortions that tend to be present in the thoughts of individuals with emotional disorders. "Arbitrary inference" refers to drawing a specific conclusion in the absence of evidence to support the conclusion. In "absolutistic, dichotomous thinking," only one or both of two extreme alternatives are considered. "Overgeneralizing" refers to a single instance being taken as representative of a broader class of situations or characteristics. In "selective abstraction," some aspects of a situation, typically negative, are attended to or remembered at the expense of other more positive aspects. In "mind reading," the attitudes or future actions of others are assumed without evidence. "Personalizing" refers to the assumption that an action is directed toward (or occurs because of) oneself rather than some other aspect of the situation. "Should statements" refer to absolute imperatives that are expressed regarding an individual's or others' behavior. In "catastrophizing," extreme negative outcomes are anticipated without substantial evidence. In "minimizing," the significance of positive outcomes is downplayed, just the reverse of "catastrophizing."

Recent developments in this cognitive model suggest that the presence of very early schemas is an important predisposing factor for many patients with depression (Beck et al., 1990; Stein & Young, 1992; Young, 1990). Young identified a subset of schemas which he has labeled "early maladaptive schemas" (Young, Beck, & Weinberger, 1993). "Early Maladaptive Schemas refer to extremely stable and enduring themes that develop during childhood and are elaborated upon throughout the individual's lifetime" (Young, 1990, p. 9). Young has identified 16 of these early maladaptive schemas in the six following domains: instability and disconnection, impaired autonomy, undesirability, restricted self-expression, restricted gratification, and impaired limits.

According to Young, children learn to construct reality through early experiences with the environment, especially with significant others. Sometimes these early experiences lead children to accept attitudes and beliefs that will later prove maladaptive. These schemas are usually out of awareness, and may remain dormant until a life event stimulates the schema. Once this schema is activated, the patient categorizes, selects, and encodes information in such a way that the maladaptive schema is maintained. In this way, early maladaptive schemas predispose depressed patients to distort events in a characteristic fashion.

Early maladaptive schemas are defined as (i) *a priori* truths about oneself and/or the environment; (ii) self-perpetuating and resistent to change; (iii) dysfunctional; (iv) often triggered by some environmental change; (v) tied to high levels of emotion when activated; and (vi) usually are the result of an initial interaction between the child's innate temperament and dysfunctional developmental experiences with family members or caretakers (Young, 1990).

Cognitive-behavior therapy of depression is based on the premise that individuals can learn to recognize and modify their negative beliefs and maladaptive information-processing tendencies to prevent or alleviate depression (Beck, 1964, 1970; Beck et al., 1979). This approach is typically operationalized as a structured, didactic intervention in which clients are encouraged to test the accuracy of their own beliefs, often by means of modifying their behavior in a systematic fashion. Treatment is usually provided in a skills-training format, with the goal being to help clients acquire a frame of reference and a set of procedures that they themselves can apply to forestall future depressions.

Brewin (1989) suggested that rather than create new memories, cognitive therapy for depression attempts instead to limit the ease with which these memories are activated by the current environment. One common feature of depression is that negative mood changes are elicited by a wide range of stimuli. Patients respond to many relatively harmless situations as though these situations contained enormous potential for psychological or physical threat.

By drawing attention to the patient's apparent assumptions and challenging these with the use of logic and behavioral experimentation, the patient can develop new rules for discriminating between situations that are truly threatening and those that merely elicit the feeling of being threatened, similar to the process of construct elaboration described by Kelly (1955). Practice in making these discriminations then changes the content of verbally accessible knowledge so that previously threatening situations are reclassified and automatic activation of the unconscious representations is decreased (Brewin, 1989).

In his activation-based model of cognitive therapy for depression (Brewin, 1996), relevant situationally accessible memories are not overwritten and remain available to be reactivated, which explains the high likelihood of relapse in depression. The more varied the negative experiences contributing to the formation of the memories, the fewer distinctive stimulus features the memories contain. This promotes overgeneralization and impedes the identification of relevant discriminations. According to Brewin, relevant cues associated with any event may activate or deactivate memories. Negative life events are therefore strongly associated with the onset of depression where there is prior low self-esteem (Brown, Andrews, Harris, Adler, & Bridge, 1986), and positive events are strongly associated with recovery from depression (Brown, Lemyre, & Bifulco, 1992).

1.11.5.2 Differential Activation Model

The differential activation model of depression (Clark & Teasdale, 1985; Teasdale, 1983; Teasdale & Dent, 1987) attempts to account for vulnerability both to onset and persistence of depression. The differential activation hypothesis assumes that in addition to any differences in cognitive organization that may be apparent in the nondepressed state and in addition to idiosyncratic cognitive schemas that may be activated, vulnerability to severe and persistent depression depends upon differences in patterns of thinking that are activated while in the depressed state. It is suggested that the patterns of thinking activated in the initial depressed state will determine whether that state remains mild or transient, or whether it becomes more severe (onset vulnerability). Differences in the patterns activated in more severe states will determine whether those states show remission or become persistent and chronic (persistence vulnerability).

According to this hypothesis, the original source of the depression may not matter too much. The crucial factor that determines whether the initial depression will intensify and persist is the pattern of thinking that exists when the patient is depressed. The way in which experiences are interpreted will be a joint function of the nature of the experience and the state of the information processing system that interprets them. The latter will determine the type of information that is selected for attention and the nature of the interpretative categories that are highly primed and most likely to be used to interpret experiences (Bargh, Bond, Lombardi, & Tota, 1986). Similarly, the type of information from memory that is processed will be a function both of what is available in memory and what is most accessible at a given time. Interpretative biases may also affect the interpretation of this material.

Teasdale (1983) has elaborated upon Bower's associative network theory (Bower, 1981) to explain the effect of mood on cognitive processing. Teasdale (1985) has proposed that faulty beliefs about symptoms (depression about depression) lead to the exacerbation of depression. Depressed mood increases the accessibility of representations of depressing experiences and the accessibility of negative interpretative categories and constructs. In depressed mood, there will not only be an increased likelihood that unhappy memories will come to mind, but there will also be a negative bias in the way situations are perceived and interpreted, and in the way in which inferences and predictions are made using information from the environment and from memory. As a result of these effects of mood on cognitive processing, once a person is in a depressed state any environmental input that can be interpreted in ways that will produce further depression will be more likely to be interpreted in that way. Similarly, the memories that come most easily to mind in the depressed state are those that are likely to perpetuate the depression. This reciprocal relationship between depression and cognitive processing creates the potential for a positive feedback loop. Depressed mood and cognitive processing can act to reinforce each other and to set up a vicious cycle that will tend to intensify and maintain depression.

Barnard and Teasdale (1991) and Teasdale and Barnard (1993) have elaborated upon the original differential activation hypothesis within the framework of a comprehensive systemic model of the organization and function of resources underlying human cognition, the interacting cognitive subsystems (ICS) model. Their ICS model, an extension of Barnard's original ICS model (Barnard, 1985), proposes a framework of nine peripheral and central cognitive subsystems: the acoustic, morphonolexical, articulatory, propositional, implicational, body state, visual, object, and limb subsystems. Each subsystem is hypothesized to contribute either directly or indirectly to the production or experience of emotion.

The ICS model characterizes persistent depression as a product of the establishment of self-maintaining, reverberatory patterns of cognitive activity within the implicational subsystem, with streams of negative thoughts that are experienced by depressed patients as markers of that activity. This theory suggests that the function of therapy is to disrupt the repeated synthesis of high-level schematic representations containing generic meanings prototypical of previous depressing situations (so-called depressive interlock). These representations contain idiosyncratic sensory, proprioceptive, and meaning information synthesized from past experiences of depression that cannot be fully captured by a verbal description. Interlock involves numerous cognitive subsystems, which include those responsible for processing sensory and proprioceptive data as well as those that extract meaning at propositional and higher-order (implicational) levels. Disruption of interlock may be achieved in a number of ways. For example, if the representations were continually being activated by ongoing stress, problem-solving training aimed at resolving that stress might effectively deactivate them. Consistent with this view, distraction procedures that interrupt the flow of negative thoughts may also lead to an alleviation of patients" depression (Fennell, Teasdale, Jones, & Damle, 1987).

1.11.5.3 Cognitive Appraisal Model

As noted earlier, the work of Lazarus has long been associated with cognitive appraisal processes mediating emotional responses to stressful stimuli (Lazarus, Averill, & Opton, 1970). This has provided a general framework for explaining depressive reactions to personally stressful events and circumstances (Coyne, Aldwin, & Lazarus, 1981; Folkman & Lazarus, 1986; Lazarus, Averill, & Opton, 1970; Lazarus & Folkman, 1984).

The central component of the cognitive appraisal model of depression is an emphasis on person–environment reciprocal influences, cognitive appraisal processes, and coping. This transactional perspective emphasizes relational factors that are thought to "transcend the separate sets of person and environment variables of which they are comprised" (Lazarus & Launier, 1978). In Lazarus' model, both environmental and personal factors are seen to be involved in an ongoing reciprocal relationship, which can be studied as person–environment units. Transactions between person and environment are mediated by two processes: appraisals and coping.

According to Lazarus and Launier (1978), "cognitive appraisal can be simply understood as the mental process of placing any event in one of a series of evaluative categories related either to its significance for the person's well-being (primary appraisal) or to the available coping resources and options (secondary appraisal)." The emotional responses of sadness and depression are therefore a direct outcome of appraisals of the meaning of an event in relation to that person's well-being. The emotional response may be immediate and nonreflective, so that thoughts and feelings are simultaneous, but appraisal is an essential component of emotion (Lazarus, 1982).

Primary appraisals may lead to judgments that a circumstance is irrelevant, benign or positive, or stressful. A stressful appraisal can typically be classified as one of three types: harm/loss (the damage has already occurred), threat (anticipated harm or loss), and challenge. The outcome of this primary appraisal, in this case the experience of depression, will be determined by the transaction between personal experiences, predisposition, and environmental factors. Coping appraisals or secondary appraisals of perceived coping resources and options reciprocally interact with the primary appraisal of the nature and magnitude of the stressor and its effect on the person's well-being.

Similar to other cognitive models of depression, the cognitive appraisal model emphasizes the central role of cognitive processes in provoking depressive reactions to stressful events. The transactional emphasis, however, differs from perspectives that emphasize either the qualities of events (Paykel, 1979; Brown & Harris, 1978) or personal characteristics (Abramson, Seligman, & Teasdale, 1978; Beck, 1967). The Lazarus model views the individual as someone in a continuing state of appraising and negotiating stressful occurrences and everyday circumstances. Although threatening or stressful acute events or ongoing chronic stressors are seen as important antecedents of emotional reactions, everyday "hassles," or more minor events can also be associated with depressive symptomatology (Kanner, Coyne, Schaefer, & Lazarus, 1981).

1.11.5.4 Learned Helplessness Model

The learned helplessness model was first described in the 1960s by Overmier and Seligman (1967) and Seligman and Maier (1967). They reported that dogs which experienced inescapable electric shock demonstrated significant motivational, learning, and emotional deficits and failed to initiate behaviors to terminate the shock. Seligman suggested that these deficits were a consequence of the animals learning that their behavior had no impact on the outcome of whether or not they received a shock. After experience with uncontrollable outcomes, they appeared to develop low expectancies for exerting control over later outcomes which could be controlled. These low expectancies appeared to produce motivational, affective, and behavioral deficits similar to depression observed in humans.

Seligman (1975) compared the responses of people experiencing feelings of helplessness as a result of experimental manipulations in the laboratory with the behavior of individuals suffering from naturally occurring depression. Both groups demonstrated similar response patterns including passivity, slowed learning, lowered aggression, loss of appetite, negative expectations, and feelings of helplessness, hopelessness, and powerlessness.

The original helplessness model did not specify which factors account for differences in the severity and chronicity of depression, however, nor did it address subtypes of clinical depression or explain such symptoms as guilt and blame (Abramson & Sackheim, 1977). In some instances, uncontrollability of outcome also not only failed to produce helplessness but actually facilitated subjects' subsequent performance (Wortman & Brehm, 1975).

1.11.5.4.1 Revised helplessness model

In response to the need for a more complex model of human depression, Abramson, et al. (1978) reformulated the helplessness theory, placing an emphasis on individual causal interpretation of uncontrollable events. They proposed that an individual's attribution for loss of control serves as a mediator between the absence of control and the emergence of helplessness-related deficits and depression and that helplessness and depression are most likely to occur following either the inability to produce a highly desired outcome or the inability to prevent a highly aversive outcome. When confronted by such circumstances, individuals attempt to explain why the circumstances have occurred and these explanations determine how they will respond to events.

In this revised model, it is assumed that attributions are the primary determinants of an individual's actions and affect. According to Abramson et al. (1978), the cause may be attributed either to the person or to the situation (an internal vs. an external explanation). The cause may also be viewed as either transient or persistent across time (an unstable vs. a stable explanation), and the cause may be perceived to have an impact on a variety of outcomes or may be limited just to that particular event (a global vs. a specific explanation). The revised helplessness model proposes that particular consequences are associated with attributions made across each of these three dimensions. For example, an individual who attributes a negative event to an internal cause is likely to experience a loss of self-esteem. The stability of causal beliefs is hypothesized to affect the persistence or chronicity of depression following negative events. If a negative event is explainable by a persistent cause, then depressive reactions are likely to be chronic. If an individual believes that a global factor has caused the negative event, then deficits associated with helplessness and depression are likely to occur in a variety of situations.

The revised helplessness model assumes that individuals demonstrate consistency in attribution and tend to invoke the same sorts of causal explanations for different negative events in their lives. An individual who typically invokes internal, stable, and global causal beliefs will therefore be at a high risk for developing depression when faced with uncontrollable negative events. Brewin (1985) and Sweeney, Anderson, and Bailey (1986) have reviewed the literature examining this relationship between measures of attributional style and depression.

According the revised model, when an individual perceives negative events as beyond their control, the expectation of future uncontrollability is likely to occur. In such circumstances, the specific attributions that develop are influenced by the reality of the particular circumstances and the individual's explanatory style. The expectation of future uncontrollability is thought to be sufficient to produce most of the symptoms of depression, including a lowered rate of activity, cognitive deficits such as impaired learning, emotional responses of sadness and anxiety, and a lowering of appetitive drives. In addition, the expectation of uncontrollability is hypothesized to bring about physiological consequences such as neurochemical changes and an increase in susceptibility to disease.

1.11.5.4.2 Hopelessness theory

Abramson, Metalsky, and Alloy (1988) present a further elaboration of the revised theory of human helplessness and depression which is relabeled as the hopelessness theory of depression. The hopelessness theory of depression specifies hopelessness as a proximal sufficient and necessary cause of depression beginning with the occurrence of negative life events and ending with the symptoms of a type of depression called hopelessness depression. It has been speculated (Seligman, 1978) that hopelessness depression may not directly correspond to current diagnostic categories of depression, and may even include psychological phenomena not currently contained in psychiatric nosologies. The hopelessness subtype of depression is hypothesized to be characterized by three primary symptoms and one secondary symptom. The major symptoms are described as the retarded initiation of voluntary responses (a motivational symptom), a difficulty in seeing that one's responses control outcomes related or similar to the outcome about which one feels hopeless (a cognitive symptom), and sad affect (an emotional symptom). The secondary symptom of lowered self-esteem is hypothesized to be related to the expectation that other people can attain the outcome that the patient feels hopeless to attain.

The causal pathway hypothesized by Abramson et al. (1988) which culminates in hopelessness suggests that at least two distinct variants of hopelessness depression may exist. The first theoretical variant of hopelessness depression is called attributional style-event hopelessness depression. This variant of depression would develop in an individual who exhibits a depressogenic attributional style and is confronted with a negative life event which by itself may not be sufficient to trigger hopelessness in most people, but when interpreted or explained by an individual with this attributional style is sufficient to do so. A second variant of hopelessness depression is labeled as event hopelessness depression. This theoretical variant of hopelessness depression would develop in an individual who does not exhibit the hypothesized depressogenic attributional style or other depressogenic cognitive styles, but who is confronted with an event sufficient to engender hopelessness in most people. It is thought that event hopelessness depression may overlap with the diagnostic category of acute dysphoria (Klein, 11974) and that attributional style-event hopelessness depression may correspond to chronic over-reactive dysphoria.

Hopelessness depression is conceptualized as the culmination of a series of events beginning with a negative life stress and including depressogenic attributional styles, situational cues, stable and global attributions for the negative life event, and other contributory causal pathways such as deficits in social support networks. Intervention to reverse hopelessness depression or to reduce vulnerability to depressive episodes and depression-proneness can by definition occur at any point in the causal chain leading to hopelessness, and possible intervention strategies are reviewed by Alloy, Clements, and Kolden (1985), Beach, Abramson, and Levine (1981), and Halberstadt, Andrews, Metalsky, and Abramson (1984).

1.11.6 CONCLUSIONS

In addressing the cognitive basis of clinical psychology, several developments have been considered in the field of cognitive psychology, and how they may impact the development, presentation, and treatment of psychopathology. The focus was limited to salient topics that embody multifaceted influences of other scientific endeavors, including those of information networking research, emotion, and abnormal psychology. In doing so, other issues of importance to the field of cognitive psychology have not been addressed comprehensively. It is hoped that this approach, while relatively less inclusive, provides the reader with a flavor of how information processing can impact, or at minimum interact with, other factors in the development of psychopathology.

Broadly defined, cognitive approaches to emotional response focus on evaluation of information processing patterns that contribute to directing and organizing a behavioral/emotional response. The role of cognitive processes such as appraisal determine how we react at multiple behavioral and physiological levels, and how these reactions may contribute to maladaptive behaviors. In this vein, the role of cognitive distortions as important processing dispositions that may predispose an individual to emotional problems have been addressed.

The discussion of anxiety and depression as examples of emotional disorders directly impacted by cognitive factors provides an illustration of the nature of cognitive contributions in the cascade of events leading to the development of psychopathology. This discussion was preceded by an overview of how cognition and emotion interact in a dynamic fashion. The approach chosen treats cognitive and emotional change, not only as potential mediators for each other, but also as consequences of each other. Such may be the case when addressing cognitive deficits observed in depressed individuals.

A trend that has been noted here is of multiple disciplines addressing cognition-based questions. The increased volume of these interdisciplinary investigations is being nurtured by the growth in cognitive neuroscience, psychophysiology, and neurobiology. Methodological and conceptual advancements in these areas are being brought to bear to increase understanding of how cognitive processing affects behavior at the level of multiple systems. Some of these issues have been brought to the forefront in discussion of the interaction of cognition and emotion and the viability of integrating research in this area with research in neurobiology and neuroanatomy.

The challenge that lies ahead is in bringing about a coherent integration of these disciplines that is theoretically sound and empirically supported. Such a framework not only would allow efforts in these areas to be well focused and well guided, but would also foster further cross-disciplinary research on constituent questions within each area to address how cognitive factors contribute with other factors in precipitating psychopathology.

1.11.7 REFERENCES

Abramson, L. Y., Alloy, L. B., & Rosoff, R. (1981). Depression and the generation of complex hypotheses in the judgment of contingency. *Behavior Research and Therapy, 19,* 35–45.

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1988). The hopelessness theory of depression: Does the research test the theory? In L. Y. Abramson (Ed.), *Social cognition and clinical psychology: A synthesis* (pp. 33–65). New York: Guilford Press.

Abramson, L. Y., & Sackheim, H. A. (1977). A paradox in depression: Uncontrollability and self-blame. *Psychological Bulletin, 84,* 835–851.

Abramson, L., Seligman, M., & Teasdale, J. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87,* 49–74.

Aggleton, J. P. (Ed.) (1992). *The amygdala: Neurobiological aspects of emotion, memory and mental dysfunction.* New York: Wiley-Liss.

Alavi, F., & Taylor, J. G. (1993). Mathematical analysis of a global network for attention. In J. G. Taylor (Ed.), *Mathematical approaches to neural networks.* Oxford, UK: Elsevier.

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organisation of functionally segregated circuits linking basal ganglia and cortex. *Annual Review Neuroscience, 9,* 357–381.

Alloy, L. B., Clements, C., & Kolden, G. (1985). The cognitive diathesis-stress theories of depression: Therapeutic implications. In S. Reiss & R. Bootzin (Eds.), *Theoretical issues in behavior therapy.* New York: Academic Press.

Anderson, D. J., Noyes, R., & Crowe, R. R. (1984). A comparison of panic disorder and generalized anxiety disorder. *American Journal of Psychiatry, 141,* 572–575.

Anderson, J. A. (1995). *An introduction to neural networks.* Cambridge, MA: MIT Press.

Baddely, A. (1986) *Working memory.* London: Oxford University Press.

Bargh, J. A., Bond, R. N., Lombardi, W. J., & Tota, M. E.

(1986). The additive nature of chronic and temporary sources of construct accessibility. *Journal of Personality and Social Psychology, 50,* 869–878.

Barlow, D. H. (1988). *Anxiety and its disorders: The nature and treatment of anxiety and panic*. New York: Guilford Press.

Barlow, D. H., Blanchard, E. B., Vermilyea, J. A., Vermilyea, B. B., & DiNardo, P. A. (1986). Generalized anxiety and generalized anxiety disorder: Description and reconceptualization. *American Journal of Psychiatry, 143,* 40–44.

Barnard, P. J. (1985). Interacting cognitive subsystems: A psycholinguistic approach to short-term memory. In A. Ellis (Ed.), *Progress in the psychology of language* (Vol. 2, pp. 197–258). London: Erlbaum.

Barnard, P. J., & Teasdale, J. D. (1991). Interacting cognitive subsystems: A systematic approach to cognitive-affective interaction and change. *Cognition and Emotion, 5,* 1–39.

Beach, S. R. H., Abramson, L. Y., & Levine, F. M. (1981). Attributional reformulation of learned helplessness and depression: Therapeutic implications. In J. F. Clarkin & H. I. Glazer (Eds.), *Depression: Behavioral and directive intervention strategies*. New York: Garland.

Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., & Damasio, A. R. (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science, 269,* 1115–1118.

Beck, A. T. (1963). Thinking and depression. *Archives of General Psychiatry, 9,* 324–333.

Beck, A. T. (1964). Thinking and depression: 2. Theory and therapy. *Archives of General Psychiatry, 10,* 561–571.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York: Harper & Row.

Beck, A. T. (1970). Cognitive therapy: Nature and relation to behavior therapy. *Behavior Therapy, 1,* 184–200.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.

Beck, A. T., Emery, G., & Greenberg, R. L. (1985). *Anxiety disorders and phobias: A cognitive perspective*. New York: Basic Books.

Beck, A. T., Freeman, A., Pretzer, J., Davis, D. D., Fleming, B., Ottaviani, R., Beck, J., Simon, K. M., Padesky, C., Meyer, J., & Trexler, L. (1990). *Cognitive therapy of personality disorders*. New York: Guilford Press.

Beck, A. T., & Hurvich, M. S. (1959). Psychological correlates of depression: Frequency of "masochistic" dream content in a private practice sample. *Psychosomatic Medicine, 21,* 50–55.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Wiley.

Beck, A. T., & Ward, C. H. (1961). Dreams of depressed patients: characteristic themes in manifest content. *Archives of General Psychiatry, 5,* 562–571.

Blaney, P. H. (1986). Affect and memory: A review. *Psychological Bulletin, 99,* 229–246.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences, 18,* 227–287.

Blythe, I. M., Kennard, C., & Ruddock, K. H. (1987). Residual vision in patients with retrogeniculate lesions of the visual pathways. *Brain, 110,* 887–905.

Borkovec, T. D., & Hu, S. (1990). The effect of worry on cardiovascular response to phobic imagery. *Behavior Research and Therapy, 28,* 69–73.

Borod, J. C. (1992). Interhemispheric and intrahemispheric control of emotion: A focus on unilateral brain damage. *Journal of Consulting and Clinical Psychology, 60,* 339–348.

Bower, G. H. (1981). Mood and Memory. *American Psychologist, 36,* 129–148.

Bower, G. H., Gilligan, S. G., & Monteiro, K. P. (1981). Selectivity of learning caused by affective states. *Journal of Experimental Psychology: General, 110,* 451–473.

Brewin, C. R. (1985). Depression and causal attributions: What is their relation? *Psychological Bulletin, 98,* 297–309.

Brewin, C. R. (1988). *Cognitive foundations of clinical psychology*. London: Erlbaum.

Brewin, C. R. (1989). Cognitive change processes in psychotherapy. *Psychological Review, 96,* 379–394.

Brewin, C. R. (1996). Theoretical foundations of cognitive-behavioral therapy for anxiety and depression. *Annual Review of Psychology, 47,* 33–57.

Brooks, D. N., & Baddeley, A. D. (1976). What can amnesic patients learn? *Neuropsychologia, 14,* 111–122.

Brown, G. W., Andrews, B., Harris, T. O., Adler, Z., & Bridge, L. (1986). Social support, self-esteem and depression. *Psychological Medicine, 16,* 813–831.

Brown, G. W., & Harris, T. O. (1978). *Social origins of depression: A study of psychiatric disorders in women*. New York: Free Press.

Brown, G. W., Lemyre, L., & Bifulco, A. (1992). Social factors and recovery from anxiety and depressive disorders: A test of specificity. *British Journal of Psychiatry, 161,* 44–54.

Brown, T. A., O'Leary, T. A., & Barlow, D. H. (1993). Generalized anxiety disorder. In D. H. Barlow (Ed.), *Clinical handbook of psychological disorders: A step-by-step treatment manual* (2nd ed., pp. 137–188). New York: Guilford Press.

Buhusi, C., & Schmajuk, N. A. (1996). Attention, configuration and hippocampal function. *Hippocampus, 6,* 621–642.

Butler, G., Fennell, M., Robson, P., & Gelder, M. (1991). Comparison of behavior therapy and cognitive-behavior therapy in the treatment of generalized anxiety disorder. *Journal of Consulting Clinical Psychology, 59,* 167–175.

Calev, A. (1996). Affect and memory in depression: Evidence of better delayed recall of positive than negative affect words. *Psychopathology, 29,* 71–76.

Calhoun, K. S., & Resick, P. A. (1993). Post-traumatic stress disorder. In D. H. Barlow (Ed.), *Clinical handbook of psychological disorders: A step-by-step treatment manual* (2nd ed., pp. 48–98). New York: Guilford Press.

Cameron, O. G., Thyer, B. A., Nesse, R. M., & Curtis, G. C. (1986). Symptom profiles of patients with *DSM-III* anxiety disorders. *American Journal of Psychiatry, 143,* 1132–1137.

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review, 97,* 19–35.

Chambless, D. L., Caputo, G., Bright, P., & Gallagher, R. (1984). Assessment of fear in agoraphobics: The Body Sensations Questionnnaire and the Agoraphobic Cognitions Questionnnaire. *Journal of Consulting and Clinical Psychology, 52,* 1090–1097.

Channon, S., & Baker, J. E. (1996). Depression and problem-solving performance on a fault-diagnosis task. *Applied Cognitive Psychology, 10,* 327–336.

Chemtob, C., Roitblat, H., Hamada, R., Carlson, J., & Twentyman, C. (1988). A cognitive action theory of posttraumatic stress disorder. *Journal of Anxiety Disorders, 2,* 253–275.

Chrousos, G. P., & Gold, P. W. (1992). The concepts of stress and stress system disorders: Overview of physical and behavioral homeostasis. *Journal of the American Medical Association, 267,* 1244–1252.

Clark, D. M. (1988). A cognitive model of panic attacks. In S. Rachman & J. Maser (Eds.), *Panic: Psychological perspectives* (pp. 71–89). Hillsdale, NJ: Erlbaum.

Clark, D.M., Salkovskis, P., Gelder, M., Koehler, C., Martin, M., Anastasiades, P., Hackmann, A., Middleton, H., & Jeavons, A. (1988). Tests of a cognitive theory of panic. In I. Hand & H. Wittchen (Eds.), *Panic and phobias II*. Berlin: Springer.

Clark, D. M., & Teasdale, J. D. (1982). Diurnal variation in clinical depression and accessibility of memories of positive and negative experiences. *Journal of Abnormal Psychology, 91,* 87–95.

Clark, D. M., & Teasdale, J. D. (1985). Constraints on the effects of mood on memory. *Journal of Personality and Social Psychology, 48,* 1595–1608.

Cohen, R. M., Weingartner, H., Smallberg, S. A., Pickar, D., & Murphy, D. L. (1982). Effort and cognition in depression. *Archives of General Psychiatry, 39,* 593–597.

Corbetta, M., Miezin, F. M., Dobmeyer, G. L., Shulman, G. L., & Petersen, S. E. (1991). Selective and divided attention during visual discrimination of shape, color, and speed: Functional anatomy by positron emission tomography. *Journal of Neuroscience, 11,* 2383–2402.

Corbetta, M., Miezin, F. M., Shulman, G. L., & Petersen, S. E. (1993). A PET study of visuospatial attention. *Journal of Neuroscience, 13,* 1202–1226.

Corkin, S. (1968). Acquisition of motor skill after bilateral medial temporal lobe excision. *Neuropsychologia, 6,* 255–265.

Cowey, A., & Stoerig, P. (1991). The neurobiology of blindsight. *Trends in Neuroscience, 29,* 65–80.

Coyne, J., Aldwin, C., & Lazarus, R. (1981). Depression and coping in stressful situations. *Journal of Abnormal Psychology, 90,* 439–441.

Craske, M. G. (1991). Phobic fear and panic attacks: The same emotional state triggered by different cues? *Clinical Psychological Review, 11,* 599–620.

Craske, M. G., & Barlow, D. H. (1988). A review of the relationship between panic and avoidance. *Clinical Psychological Review, 8,* 667–685.

Craske, M. G., & Barlow, D. H. (1993). Panic disorder and agoraphobia. In D. H. Barlow (Ed.), *Clinical handbook of psychological disorders: A step-by-step treatment manual* (2nd ed., pp. 1–47). New York: Guilford Press.

Craske, M. G., Barlow, D. H., & O'Leary, T. A. (1992). *Mastery of your anxiety and worry*. Albany, NY: Graywind Publications.

Creamer, M., Burgess, P., & Pattison, P. (1992). Reaction to trauma: A cognitive processing model. *Journal of Abnormal Psychology, 101,* 453–459.

Crowe, R. R., Noyes, R., Pauls, D. L., & Slymen, D. J. (1983). A family study of panic disorder. *Archives of General Psychiatry, 40,* 1065–1069.

Davey, G. C. L. (1992). Classical conditioning and the acquisition of human fears and phobias: A review and synthesis of the literature. *Advances in Behavior Research and Therapy, 14,* 29–66.

Davey, G. C. L. (1995). Preparedness and phobias: Specific evolved associations or a generalized expectancy bias? *Behavioral and Brain Sciences, 18,* 289–325.

Davidson, R. J. (1992). Emotion and affective style: Hemispheric substrates. *Psychological Science, 3,* 39–43.

Debner, J. A., & Jacoby, L. L. (1994). Unconscious perception: Attention, awareness, and control. *Journal of Experimental Psychology Learning, Memory, and Cognition, 20,* 304–317.

Derry, P. A., & Kuiper, N. A. (1981). Schematic processing and self-reference in clinical depression. *Journal of Abnormal Psychology, 90,* 286–297.

Derryberry, D., & Tucker, D. M. (1992). Neural mechanisms of emotion. *Journal of Consulting and Clinical Psychology, 60,* 329–338.

Dickinson, A. (1987). Animal conditioning and learning theory. In H. Eysenck & I. Martin (Eds.), *Theoretical foundations of behavior therapy* (pp. 57–79). New York: Plenum.

Dixon, N. F. (1981). *Preconscious processes.* Chichester, UK: Wiley.

Drevets, W. C., Videen, T. O., Price, J. L., Preskorn, S. H., Carmichael, S. T., & Raichle, M. E. (1992). A functional anatomical study of unipolar depression. *Journal of Neuroscience, 12,* 3628–3641.

Dunn, L. A., Atwater, G. E., & Kilts, C. D. (1993). Effects of anti-psychotic drugs on latent inhibition: Sensitivity and specificity of an animal behavioural model of clinical drug action. *Psychopharmacology, 112,* 315–323.

Ehlers, A., & Steil, R. (in press). Maintenance of intrusive memories in posttraumatic stress disorder: A cognitive approach. *Behav Cognitive Psychotherapy.*

Eich, E., & Metcalfe, J. (1989). Mood dependent memory for internal versus external events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 443–455.

Eichenbaum, H., Otto, T., & Cohen, N. J. (1992). The hippocampus—what does it do? *Behavioral and Neural Biology, 57,* 2–36.

Ekman, P. (1973). Cross-cultural studies of facial expression. In P. Ekman (Ed.), *Darwin and facial expression: A century of research in review*. New York: Academic Press.

Ellenbroek, B. A., & Cools, A. (1990). Animal models with construct validity for schizophrenia. *Behavioural Pharmacology, 1,* 469–490.

Ellis, A., & Dryden, W. (1997). *The practice of rational emotive therapy* (2nd ed.). New York: Springer.

Ellis, H. C. (1985). On the importance of mood intensity and encoding demands in memory: Commentary on Hasher, Rose, Zacks, Sanft, and Doren. *Journal of Experimental Psychology: General, 114,* 392–395.

Ellis, H. C., Thomas, R. L., & Rodriguez, I. A. (1984). Emotional mood state and memory: Elaborative encoding, semantic processing, and cognitive effort. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 470–482.

Erdmann, G. (1983). Autonomic drugs as tolls in differentiating psychopharmacology. In W. Janke (Ed.), *Response variability to psychotropic drugs* (pp. 275–292). New York: Pergamon.

Erickson, R. C., Post, R. D., & Paige, A. B. (1975). Hope as a psychiatric variable. *Journal of Clinical Psychology, 31,* 324–330.

Farah, M. J. (1997). Computational modeling in behavioral neurology and neuropsychology. In T. E. Feinberg & M. J. Farah (Eds.), *Behavioral neurology and neuropsychology* (pp. 121–130). New York: McGraw-Hill.

Feinberg, T. E., & Farah, M. J. (1997). The development of modern behavioral neurology and neuropsychology. In T. E. Feinberg & M. J. Farah (Eds.), *Behavioral neurology and neuropsychology* (pp. 3–23). New York: McGraw-Hill.

Feldman, S., Conforti, N., & Weidenfeld, J. (1995) Limbic pathways and hypothalamic neurotransmitters mediating adrenocortical responses to neural stimuli. *Neuroscience Biobehavioral Reviews, 19,* 235–240.

Fennell, M. J. V., Teasdale, J. D., Jones, S., & Damle, A. (1987). Distraction in neurotic and endogenous depression: An investigation of negative thinking in major depressive disorder. *Psychological Medicine, 17,* 441–452.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: exposure to corrective information. *Psychological Bulletin, 99,* 20–35.

Foa, E. B., Steketee, G. S., & Rothbaum, B. O. (1989). Behavioral/cognitive conceptualizations of post-traumatic stress disorder. *Behavioral Therapy, 20,* 155–176.

Folkman, S., & Lazarus, R. (1986). Stress processes and depressive symptomatology. *Journal of Abnormal Psychology, 95,* 107–113.

Forgas, J. P., Bower, G. H., & Krantz, S. E. (1984). The influence of mood on perception of social interactions. *Journal of Experimental Social Psychology, 20,* 497–513.

Fox, N. A. (1991). If it's not left, it's right. *American Psychologist, 46,* 863–872.

Freeman, A., & Reinecke, M. (1993). *Cognitive therapy of* *suicidal behavior: A manual for treatment*. New York: Springer.

Frost, R. O., Graf, M., & Becker, J. (1979). Self-devaluation and depressed moods. *Journal of Consulting and Clinical Psychology, 47,* 958–962.

Funabiki, D., & Calhoun, J. (1979). Use of a behavioral-analytic procedure in evaluating two models of depression. *Journal of Consulting and Clinical Psychology, 47,* 183–185.

Gainotti, G., Caltagirone, C., & Zoccolotti, P. (1993). Left/right and cortical/subcortical dichotomies in the neuropsychological study of human emotions. *Cognition and Emotion, 7,* 71–93.

Gilligan, S. G., & Bower, G. H. (1984). Cognitive consequences of emotional arousal. In C. E. Izard, J. Kagan, & R. Zajonc (Eds.), *Emotion, Cognition, and Behavior* (pp. 547–588). New York: Cambridge University Press.

Goldman-Rakic, P. S. (1992). Working memory and the mind. *Scientific American, 267*(3), 111–117.

Gorman, J., Fyer, M., Goetz, R., Askanazi, J., Liebowitz, M., Fyer, A., Kinney, J., & Klein, D. (1988). Ventilatory physiology of patients with panic disorder. *Archives of General Psychiatry, 45,* 31–39.

Gotlib, I. H. (1981). Self-reinforcement and recall: Differential deficits in depressed and nondepressed psychiatric inpatients. *Journal of Abnormal Psychology, 90,* 521–530.

Gottschalk, L. A. (1974). A hope scale applicable to verbal samples. *Archives of General Psychiatry, 30,* 779–785.

Gray, J. A. (1982). *The neuropsychology of anxiety*. New York: Oxford University Press.

Gray, J. A. (1990). Brain systems that mediate both emotion and cognition. *Cognition and Emotion, 4,* 269–288.

Gray, J. A., Buhusi, C. V., & Schmajuk, N. (1997). The transition from automatic to controlled processing. *Neural Networks, 10,* 1257–1268.

Gray, J. A., Feldon, J., Rawlins, J. N. P., Hemsley, D. R., & Smith, A. D. (1991). The neuropsychology of schizophrenia. *Behavioral and Brain Sciences, 14,* 1–84.

Gray, N. S., Pickering, A. D., Hemsley, D. R., Dawling, S., & Gray, J. A. (1992). Abolition of latent inhibition by a single 5 mg dose of d-amphetamine in man. *Psychopharmacology, 107,* 425–430.

Greenwald, A. G., & Draine, S. C. (1997). Do subliminal stimuli enter the mind unnoticed? Tests with a new method. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 83–108). Mahwah, NJ: Erlbaum.

Groeger, J. A. (1988). Qualitatively different effects of undetected and unidentified auditory primes. *Quarterly Journal of Experimental Psychology, 40,* 323–329.

Halberstadt, L. J., Andrews, D., Metalsky, G. I., & Abramson, L. Y. (1984). Helplessness, hopelessness, and depression: A review of progress and future directions. In N. S. Endler & J. Hunt (Eds.), *Personality and behavior disorders*. New York: Wiley.

Hasher, L., Rose, K. C., Zacks, R. T., Sanft, H., & Doren, B. (1985). Mood, recall, and selectivity effects in normal college students. *Journal of Experimental Psychology: General, 114,* 104–118.

Hauri, P. (1976). Dreams in patients remitted from reactive depression. *Journal of Abnormal Psychology, 85,* 1–10.

Hemsley, D. R. (1987). An experimental psychological model for schizophrenia. In H. Hafner, W. F. Gattaz, & W. Janzavik (Eds.), *Search for the causes of schizophrenia*. Berlin: Springer.

Henry, G. M., Weingartner, H., Murphy, D. L. (1973). Influence of affective states and psychoactive drugs on verbal learning and memory. *American Journal of Psychiatry, 130,* 966–971.

Hertel, P. T., & Hardin, T. S. (1990). Remembering with and without awareness in a depressed mood: Evidence of deficits in initiative. *Journal of Experimental Psychology: General, 119,* 45–59.

Hinton, G. E., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychology Review, 98,* 74–95.

Hollon, S. D., & Beck, A. T. (1979). Cognitive therapy of depression. In P. Kendall & S. D. Hollon (Eds.), *Cognitive behavioral interventions: Theory, research, and procedures*. New York: Academic Press.

Hollon, S. D., & Kendall, P. C. (1980). Cognitive self-statements in depression: Development of an automatic thoughts questionnaire. *Cognitive Therapy Research, 4,* 383–395.

Holt, P., & Andrews, G. (1989). Hyperventilation and anxiety in panic disorder, social phobia, generalized anxiety disorder, and normal controls. *Behavior Research and Therapy, 27,* 453–460.

Hope, D. A., & Heimberg, R. G. (1993). Social phobia and social anxiety. In D. H. Barlow (Ed.), *Clinical handbook of psychological disorders: A step-by-step treatment manual* (2nd ed., pp. 99–136). New York: Guilford Press.

Hope, D. A., Rappe, R. M., Heimberg, R. G., & Dombeck, M. S. (1990). Representations of the self in social phobia: Vulnerability to social threat. *Cognitive Therapy and Research, 14,* 177–189.

Ingram, R. E., & Kendall, P. C. (1986). Cognitive clinical psychology: Implications of an information processing perspective. In R. E. Ingram (Ed.), *Information processing approaches to clinical psychology*. New York: Academic Press.

Izard, C. E. (1992). Basic emotions, relations among emotions, and emotion-cognitive relations. *Psychological Review, 99,* 561–565.

Jacoby, L. L. (1991). A process disociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513–541.

Johnson, J. E., Petzel, T. P., Hartney, L. M., & Morgan, L. M. (1983). Recall and importance ratings of completed and uncompleted tasks as a function of depression. *Cognitive Therapy and Research, 7,* 51–56.

Johnson, M. H., & Magaro, P. A. (1987). Effects of mood and severity on depression and mania. *Psychological Bulletin, 101,* 28–40.

Jones, J. C., & Barlow, D. H. (1992). A new model of posttraumatic stress disorder: Implications for the future. In P. A. Saigh (Ed.), *Posttraumatic stress disorder* (pp. 147–165). New York: Macmillan.

Kanner, A. D., Coyne, J. C., Schaefer, C., & Lazarus, R. (1981). Comparison of two modes of stress measurement: Daily hassles and uplifts versus major life events. *Journal of Behavioral Medicine, 4,* 1–39.

Kelly, G. A. (1955). *The psychology of personal constructs* (Vol. 1). New York: Norton Press.

Kendall, P. C., & Hammen, C. (1995). *Abnormal psychology*. Boston: Houghton Mifflin.

Killcross, A. S., Dickinson, A., & Robbins, T. W. (1994). Amphetamine-induced disruption of latent inhibition are reinforcer medicated: Implications for animal models of schizophrenic attentional dysfunction. *Psychopharmacology, 115,* 185–195.

Klein, D. F. (1974). Endogenomorphic depression: Conceptual and terminological revision. *Archives of General Psychiatry, 31,* 447–454.

Koh, S. D., & Wolpert, E. A. (1983). Memory scanning and retrieval in affective disorders. *Psychiatry Research, 8,* 289–297.

Kolb, B., & Whishaw, I. Q. (1996). *Fundamentals of human neuropsychology*. New York: W. H. Freeman & Company.

Kovacs, M., & Beck, A. T. (1978). Maladaptive cognitive structures in depression. *American Journal of Psychiatry, 135,* 525–533.

Lang, P. J. (1979). A bio-informational theory of emotional imagery. *Psychophysiology, 16,* 495–512.

Lazarus, R. S. (1982). Thoughts on the relations between emotions and cognition. *American Psychologist, 37,* 1019–1024.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.

Lazarus, R. S., Averill, J. R., & Opton, E. M. (1970). Towards a cognitive theory of emotion. In M. Arnold (Ed.), *Feelings and emotions* (pp. 207–232). New York: Academic Press.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.

Lazarus, R. S., & Launier, R. (1978). Stress-related transactions between person and environment. In L. Pervin & M. Lewis (Eds.), *Perspectives in interactional psychology*. New York: Plenum.

Leary, M. R. (1988). A comprehensive approach to the treatment of social anxieties: The self-presentational model. *Phobia Practice Research Journal, 1,* 48–57.

LeDoux, J. E. (1989). Cognitive-emotional interactions in the brain. *Cognition and Emotion, 3,* 267–289.

LeDoux, J. E. (1992). Brain mechanisms of emotion and emotional learning. *Current Opinion in Neurobiology, 2,* 191–197.

LeDoux, J. E. (1993). Emotional memory systems in the brain. *Behavioral Brain Research, 58,* 69–79.

Leight, K. A., & Ellis, H. A. (1981). Emotional mood states, strategies, and state-dependency in memory. *Journal of Verbal Learning and Verbal Behavior, 20,* 251–275.

Logan, G. D., Zbrodoff, N. J., & Williamson, J. (1984). Strategies in the color-word Stroop task. *Bulletin of the Psychonomic Society, 22,* 135–138.

Lovallo, W. R. (1997). *Stress and health: Biological and psychological interactions*. Thousand Oaks, CA: Sage.

Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15,* 197–237.

Marshall, G. D., & Zimbardo, P. G. (1979). Affective consequences of inadequately explained physiological arousal. *Journal of Personality and Social Psychology, 37,* 70–88.

Mateer, C. A., & Mapou, R. (1996). Understanding, evaluating and managing attention disorders following traumatic brain injury. *Journal of Head Trauma Rehabilitation, 11,* 10–16.

Mathews, A., & Bradley, B. (1983). Mood and the self-reference bias in recall. *Behavior Research and Therapy, 21,* 233–239.

Matt, G. E., Vazquez, C., & Campbell, W. K. (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review, 12,* 227–255.

Mattia, J. I., Heimberg, R. G., & Hope, D. A. (1993). The revised Stroop color-naming task in social phobics: Diagnostic and treatment outcome implications. *Behavior Research and Therapy, 31,* 305–313.

Merikle, P. M., & Cheesman, J. (1987). Current status of research on subliminal perception. In M. Wallendorf & P. F. Anderson (Eds.), *Advances in consumer research* (Vol. XIV, pp. 298–302). Provo, UT: Association for Consumer Research.

Merikle, P. M., & Joordens, S. (1997). Measuring unconscious influences. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 109–123). Mahwah, NJ: Erlbaum.

Miller, W. R. (1975). Psychological deficit in depression. *Psychological Bulletin, 82,* 238–260.

Milner, A. D., Corkin, S., & Teuber, H. L. (1968). Further analysis of the hippocampal amnesic syndrome: 14-year follow-up study of H. M. *Neuropsychologia, 6,* 215–235.

Monchi, O., & Taylor, J. G. (1995). A model of the prefrontal loop that includes the basal ganglia in solving the recency task. *Proceedings of WCNN95*. Washington, DC: INNS Press.

Monchi, O., & Taylor, J. G. (1997). A model of coupled frontal working memories for delayed response tasks and the Wisconsin Card Sorting Task. In P. P. Wang (Ed.), *International Joint Conference in Information Sciences* (2nd IICIN'97, pp. 34–37). Amsterdam: Elsevier.

Moran, C., & Andrews, G. (1985). The familial occurrence of agoraphobia. *British Journal of Psychiatry, 146,* 262–267.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229,* 782–784.

Moscovitch, M. (1984). The sufficient conditions for demonstrating preserved memory in amnesia: A task analysis. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (1st ed., pp. 104–114). New York: Guilford Press.

Moscovitch, M. (1992). A neuropsychological model of memory and consciousness. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (2nd ed., pp. 5–22). New York: Guilford Press.

Moscovitch, M., & Umilta, C. (1991). Conscious and nonconscious aspects of memory: A neuropsychological framework of modules and central systems. In R. G. Lister & H. J. Weingartner (Eds.), *Perspectives on cognitive neuroscience*. Oxford, UK: Oxford University Press.

Mozer, M. C., & Behrmann, M. (1990). On the interaction of selective attention and lexical knowledge: A connectionist account of neglect dyslexia. *Journal of Cognitive Neuroscience, 5,* 89–117.

Naatanen, T. (1992). *Attention and brain function*. Hillsdale, NJ: Erlbaum.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Noyes, R., Clarkson, C., Crowe, R. R., Yates, W. R., & McChesney, C. M. (1987). A family study of generalized anxiety disorder. *American Journal of Psychiatry, 144,* 1019–1024.

Noyes, R., Woodman, C., Garvey, M. J., Cook, B. L., Suelzer, M., Clancy, J., & Anderson, D. J. (1992). Generalized anxiety disorder vs. panic disorder: Distinguishing characteristics and patterns of co-morbidity. *Journal of Nervous and Mental Disease, 180,* 369–379.

Overmier, J. B., & Seligman, M. E. P. (1967). Effects of inescapable shock upon subsequent escape and avoidance learning. *Journal of Comparative Physiology and Psychology, 63,* 23–33.

Panksepp, J. (1992). A critical role of "affective neuroscience" in resolving what is basic about basic emotions. *Psychological Review, 99,* 554–560.

Patterson, K. E., Seidenberg, M. S., & McClelland, J. L. (1990). Connections and disconnections: Acquired dyslexia in a computational model of reading processes. In R. G. M. Morris (Ed.), *Parallel distributed processing: Implications for psychology and neuroscience*. London: Oxford University Press.

Paykel, E. S. (1979). Recent life events in the development of the depressive disorder. In R. A. Depue (Ed.), *The psychobiology of the depressive disorders: Implications for the effects of stress* (pp. 245–253). New York: Academic Press.

Petersen, S. E., Robinson, D. L., & Morris, J. D. (1987). Contributions of the pulvinar to visual spatial orientation. *Neuropsychologia, 25,* 97–106.

Petri, H. L., & Mishkin, M. (1994). Behaviorism, cognitivism and the neuropsychology of memory. *American Scientist, 82,* 30–37.

Petrides, M., Alivisatos, B., Evans, A. C., & Meyer, E. (1993). Dissociation of human mid-dorsolateral from posterior dorsolateral frontal cortex in memory processing. *Proceedings of the National Academy of Science USA, 90,* 873–877.

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology, 10,* 377–400.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the brain. *Annual Review of Neuroscience, 13,* 25–42.

Posner, M. I., & Raichle, M. E. (1994). *Images of mind.* Scientific American Library, NY: Freeman.

Posner, M. I., & Raichle, M. E. (1995). Precis of images of mind. *Behavioral and Brain Sciences, 18,* 327–383.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing in cognition: The Loyola Symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Rachman, S. (1990) The determinants and treatment of simple phobias. *Advances in Behavior Research and Therapy, 12,* 1–30.

Rachman, S. (1993). A critique of cognitive therapy for anxiety disorders. *Journal of Behavioral Therapy and Experimental Psychiatry, 24,* 279–288.

Rachman, S., Levitt, K., & Lopatka, C. (1987). Panic: The links between cognitions and bodily symptoms—I. *Behavior Research and Therapy, 25,* 411–423.

Rapee, R. M. (1985). The distinction between panic disorder and generalized anxiety disorder: Clinical presentation. *Australia and New Zealand Journal of Psychiatry, 19,* 227–232.

Rapee, R. M., Mattick, R., & Murrell, E. (1986). Cognitive mediation in the affective component of spontaneous panic attacks. *Journal of Behavioral Therapy and Experimental Psychiatry, 17,* 245–253.

Razran, G. (1961). The observable unconscious and the inferable conscious in current soviet psychophysiology: Interoceptive conditioning, semantic conditioning, and the orienting reflex. *Psychological Review, 68,* 81–147.

Reisenzein, R. (1983). The Schachter theory of emotion: Two decades later. *Psychological Bulletin, 94,* 239–264.

Reiss, S., Peterson, R., Gursky, D., & McNally, R. (1986). Anxiety sensitivity, anxiety frequency, and the prediction of fearfulness. *Behavior Research and Therapy, 24,* 1–8.

Resick, P. A., & Schnicke, M. K. (1992). *Cognitive processing therapy for rape victims: A treatment manual.* Newbury Park, CA: Sage.

Riskind, J. H. (1989). The mediating mechanisms in mood and memory: A cognitive-priming formulation. *Journal of Social Behavior and Personality, 4,* 173–184.

Riskind, J. H., Rholes, W. S., & Eggers, J. (1982). The Velten mood induction procedure: Effects on mood and memory. *Journal of Consulting and Clinical Psychology, 50,* 146–147.

Rolls, E. (1997). Consciousness in neural networks? *Neural Networks, 10,* 1227–1240.

Rothbaum, B. O., & Foa, E. B. (1992). Exposure therapy for rape victims with post-traumatic stress disorder. *Behavioral Therapist, 15,* 219–222.

Rudy, J. W., & Sutherland, R. J. (1995). Configural association theory and the hippocampal formation: an appraisal and reconfiguration. *Hippocampus, 5,* 375–89.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition.* Cambridge, MA: MIT Press.

Sahraie, A., Weiskrantz, L., Barbur, J. L., Simmons, A., Williams, S. C. R., & Brammer, M. J. (1997). Pattern of neuronal activity associated with conscious and unconscious processing of visual signals. *Proceedings of the National Academy of Science USA, 94,* 9406–9411.

Sanderson, W. C., & Barlow, D. H. (1990). A description of patients diagnosed with *DSM-III-R* generalized anxiety disorder. *Journal of Nervous and Mental Disease, 178,* 588–591.

Sanderson, W. C., & Beck, A. T. (1989). Classical conditioning model of panic disorder: Response to Wolpe and Rowan. *Behavior Research and Therapy, 27,* 581–582.

Sanderson, W. C., Rapee, R., & Barlow, D. (1989). The influence of an illusion of control on panic attacks induced via inhalation of 5.5% carbon dioxide enriched air. *Archives of General Psychiatry, 46,* 157–162.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 501–518.

Schacter, D. L. (1989). On the relation between memoy and consciousness: Dissociable interactions and conscious experience. In H. Roediger & F. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving*. Hillsdale, NJ: Erlbaum.

Schacter, D. L. (1994). Priming and multiple memory systems: Perceptual mechanisms of implicit memory. In D. L. Schacter & D. Tulving (Eds.), *Memory systems 1994*. Cambridge, MA: MIT Press.

Schacter, D. L., McAndrews, M. P., & Moscovitch, M. (1988). Access to consciousness: dissociations between implicit and explicit knowledge in neuropsychological syndromes. In L. Weiskrantz (Ed.), *Thought without language* (pp. 242–278). Oxford, UK: Oxford University Press.

Schacter, S., & Singer, J. E. (1962). Cognitive, social and physiological determinants of emotional state. *Psychological Review, 69,* 379–399.

Schare, M. L., Lisman, S. A., & Spear, N. E. (1984). The effects of mood variation on state dependent retention. *Cognitive Therapy and Research, 8,* 387–408.

Schlenker, B. R., & Leary, M. R. (1982). Social anxiety and self-presentation: A conceptualization and model. *Psychological Bulletin, 92,* 641–669.

Schmajuk, N. A. (1997). *Animal learning and cognition.* Cambridge: Cambridge University Press.

Schmajuk, N. A., Lam, Y. W., & Gray, J. A. (1996). Latent inhibition: a neural network approach. *Journal of Experimental Psychology (Animal Behaviour Processes), 22,* 321–349.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84,* 1–66.

Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death.* San Francisco: Freeman.

Seligman, M. E. P. (1978). Comment and integration. *Journal Abnormal Psychology, 87,* 165–179.

Seligman, M. E. P., & Maier, S. F. (1967). Failure to escape traumatic shock. *Journal Experimental Psychology, 74,* 1–9.

Shaw, B. F. (1979). The theoretical and experimental foundations of a cognitive model for depression. In P. Pliner, K. R. Blankenstein, & I. M. Sigel (Eds.), *Advances in the study of communication and affect* (Vol. 5, pp. 81–119). New York: Plenum.

Singer, J. A., & Salovey, P. (1988). Mood and memory: Evaluating the Network Theory of Affect. *Clinical Psychology Review, 8,* 211–251.

Smith, T. W., Ingram, R. E., & Brehm, S. S. (1983). Social anxiety, anxious self-preoccupation, and recall of self-relevant information. *Journal of Personality and Social Psychology, 44,* 1276–1283.

Solomon, P. R., & Staton, D. M. (1982). Differential effects of microinjections of d-amphetamine into the nucleus accumbens or the caudate-putamen on the rat's ability to ignore an irrelevant stimulus. *Biological Psychiatry, 17,* 743–756.

Spitzer, H., Desimone, R., & Moran, J. (1988). Increased attention enhances both behavioral and neuronal performance. *Science, 240,* 338–340.

Squire, L. R., Ojemann, J. G., Miezin, F. M., Petersen, S. E., Videen, T. O., & Raichle, M. E. (1992). Activation of the hippocampus in normal humans: a functional anatomical study of memory. *Proceedings of the National Academy of Science USA, 89,* 1837–1841.

Stein, D. J., & Young, J. E. (1992). Schema approach to personality disorders. In D. J. Stein & J. E. Young (Eds.), *Cognitive science and clinical disorders* (pp. 271–288). San Diego, CA: Academic Press.

Sun, R. (1997). Learning, action and consciousness: A hybrid approach toward modelling consciousness. *Neural Networks, 10,* 1317–1331.

Sweeney, P. D., Anderson, K., & Bailey, S. (1986). Attributional style in depression: A meta-analytic review. *Journal of Personality and Social Psychology, 50,* 974–991.

Taylor, J. G. (1997). Neural networks for consciousness. *Neural Networks, 10*(7), 1207–1225.

Teasdale, J. D. (1983). Negative thinking in depression: Cause, effect, or reciprocal relationship? *Advances in Behavioral Research Therapy, 5,* 3–25.

Teasdale, J. D. (1985). Psychological treatments for depression: How do they work? *Behavioral Research Therapy, 23,* 157–165.

Teasdale, J. D., & Barnard, P. J. (1993). *Cognition, affect, and change: Re-modeling depressive thought.* Hove, UK: Erlbaum.

Teasdale, J. D., & Dent, J. (1987). Cognitive vulnerability to depression: An investigation of two hypotheses. *British Journal of Clinical Psychology, 26,* 113–126.

Teasdale, J. D., & Fogarty, S. J. (1979). Differential effects of induced mood on retrieval of pleasant and unpleasant events from episodic memory. *Journal of Abnormal Psychology, 88,* 248–257.

Tham, A., Engelbrektson, K., Mathe, A. A., Johnson, L., Olsson, E., & AbergWistedt, A (1997). Impaired neuropsychological performance in euthymic patients with recurring mood disorders. *Journal of Clinical Psychiatry, 58,* 26–29.

Thornton, J. C., Dawe, S., Lee, C., Capstick, C., Corr, P. J., Cotter, P., Frangou, S., Gray, N. S., Russell, M. A., & Gray, J. A. (1996). Effects of nicotine and amphetamine on latent inhibition in human subjects. *Psychopharmacology, 127,* 164–173.

Torgerson, S. (1983). Genetic factors in anxiety disorders. *Archives of General Psychiatry, 40,* 1085–1089.

Tucker, D. M., & Williamson, P. A. (1984). Asymmetric neural control systems in human self-regulation. *Psychological Review, 91,* 185–215.

Velten, E. (1968). A laboratory task for induction of mood states. *Behavior Therapy and Research, 6,* 473–482.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92,* 548–573.

Weiner, I. (1990). Neural substrates of latent inhibition: The switching model. *Psychological Bulletin, 108,* 442–461.

Weiner, I., Lubow, R.E., & Feldon, J. (1984). Abolition of the expression but not the acquisition of latent inhibition by chronic amphetamine in rats. *Psychopharmacology, 83,* 194–199.

Weingartner, H., Cohen, R. M., Murphy, D. L., Marello, J., & Gerdt, C. (1981). Cognitive processes in depression. *Archives of General Psychiatry, 38,* 42–47.

Weingartner, H., Miller, H., & Murphy, D. L. (1977). Mood-state dependent retrieval of verbal association. *Journal of Abnormal Psychology, 86,* 276–284.

Weiskrantz, L. (1986). *Blindsight. A case study and implications.* Oxford, UK: Oxford University Press.

Weiskrantz, L. (1995). Blindsight: Not an island unto itself. *Current Directions in Psychological Science, 4,* 146–151.

Weiskrantz, L. (1997). *Consciousness lost and found: A neuropsychological exploration.* Oxford, UK: Oxford University Press.

Weiskrantz, L., & Warrington, E. K. (1970). Verbal learning and retention by amnesic patients using partial information. *Psychonomic Science, 20,* 210–211.

Weiskrantz, L., & Warrington, E. K. (1979). Conditioning in amnesic patients. *Neuropsychologia, 17,* 187–194.

Weiss, J. M., Stout, J. C., Aaron, M. F., Quan, N., Owens, M. J., Butler, P. D., & Nemeroff, C. B. (1994). Depression and anxiety: role of the locus coeruleus and corticotropin-releasing factor. *Brain Research Bulletin, 35,* 561–572.

Wilson, F. A., Scalaidhe, S. P., & Goldman-Rakic, P. S. (1993). Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science, 260,* 1955–1958.

Wolpe, J., & Rowan, V. (1989). Panic disorder: A product of classical conditioning. *Behavior Research and Therapy, 27,* 583–585.

Wortman, C. B., & Brehm, J. W. (1975). Responses to uncontrollable outcomes: An integration of reactance theory and the learned helplessness model. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 279–315). New York: Academic Press.

Yankura, J., & Dryden, W. (1997). *Using REBT with common psychological problems.* New York: Springer.

Young, J. E. (1990). *Cognitive therapy for personality disorders: A schema-focused approach.* Sarasota, FL: Professional Resource Exchange.

Young, J. E., Beck, A. T., & Weinberger, A. (1993). Depression. In D. H. Barlow (Ed.), *Clinical handbook of psychological disorders: A step-by-step treatment manual* (pp. 240–277). New York: Guilford Press.

Zeki, S. (1993). *A vision of the brain.* London: Blackwell.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.12

Epidemiology and its Rubrics

JAMES C. ANTHONY and MICHELLE L. VAN ETTEN
Johns Hopkins University, Baltimore, MD, USA

1.12.1 OVERVIEW 355
1.12.2 BACKGROUND 358
1.12.3 THE RUBRICS OF EPIDEMIOLOGY 361
1.12.3.1 Rubric 1: How Many? A Question of Quantity 361
1.12.3.2 Rubric 2: Where? A Question of Location 365
1.12.3.3 Rubric 3: Why? A Question of Causes 369
1.12.3.4 Rubric 4: How? A Question of Mechanisms 376
1.12.3.5 Rubric 5: What Can Be Done? A Question of Prevention and Intervention 381
1.12.4 CONCLUSION 385
1.12.5 REFERENCES 386

1.12.1 OVERVIEW

With a heavy reliance upon quantitative methods, epidemiology functions as a lens to aid clinicians who wish to know the answers to five important questions about what affects their patients. The five questions are:

(i) In the community at large, how many are affected?

(ii) Where are the affected cases more likely to be found?

(iii) What accounts for some people in the community becoming cases while others do not?

(iv) What linkages of states and processes influence who becomes and remains a case?

(v) What can be done to prevent and intervene?

These questions form the rubrics for epidemiology, that is, its major headings.

Using epidemiology as a lens, a clinician can see along dimensions that stretch beyond the individual patient, and beyond the threshold of the office practice, clinic, and hospital, toward the population from which the patient has surfaced. Viewed through this lens, the experience of individual patients is seen against a background of the affected cases, frequently numerous, who never seek help or receive care. Epidemiology also provides evidence about the especially informative background experience of mentally and behaviorally healthy individuals in the community who might have become cases but did not. As discussed in this chapter, epidemiology provides a sharp focus and greater understanding with respect to any patient's clinical condition. More than any other scientific discipline, epidemiology involves population-level studies of each condition of health or ill-health. Epidemiology focuses upon the nature and extent of the occurrence of each condition, its distribution in the community, the causes and mechanisms of its development, and the well-established or hotly debated means for intervention, especially in terms of health promotion and disease prevention.

Among the most remarkable accomplishments of epidemiology has been demonstration of a counterintuitive idea. Namely, effective prevention and intervention strategies are often devised at an early stage of hypothesis testing, well before there is definitive and thorough evidence on the specific causes and on causal or protective mechanisms (Wynder, 1994). This chapter draws attention to numerous examples of effective health promotion and preventive interventions before causes and causal mechanisms were known. The chapter also provides illustrations that effective prevention can elude us even when we have firm knowledge about specific causes.

As a discipline, epidemiology started to intersect with the content areas of clinical psychology and psychiatry in the early nineteenth century. Valuable field surveys of mental morbidity in circumscribed national and local populations were underway during the 1830s, as an extension of growing concern about how many people in the population were mentally ill, and whether the mentally ill were mistreated. These epidemiological field surveys on mental health were completed a decade before the European father of epidemiology, John Snow, launched his studies of cholera, and about 80 years before Joseph Goldberger embarked on pellagra research and became the discipline's American father.

In the mid-nineteenth century, while Snow was defining cholera epidemiology, his senior colleague in London's Epidemiological Society, William Farr, started a range of investigations on mental disorders as part of his general studies of mortality from all causes. For instance, using numerical methods he had learned during French medical training under Pierre Charles Alexander Louis, Farr compared mental hospital patients admitted to different asylums with respect to their chance of being discharged in a state of recovery vs. the risk of dying in the institution.

The transition of epidemiology from a prenumerate phase into numeracy actually might be seen best in relation to the study of causes of mental health and disturbances of mental life and behavior. Some of the nineteenth-century questions provoking this transition to numeracy were the following: (i) which nations and systems of government produced the mentally healthiest populations, (ii) was it true that slavery actually improved the mental health of the Africans transported to America, and (3) can't we reduce the number of cases of mental disorder by finding the families that produce these cases and doing something about them?

In numerate form, epidemiology has become a scientific discipline with specific guiding principles and concepts. It also has a unique methodology and associated jargon. Most theories of epidemiology can be expressed in a generic framework of interaction between agent, host, and environment (D.E. Lilienfeld & Stolley, 1994). Nevertheless, these theories often remain specific to an object of study, in that it makes no more sense to speak of an "epidemiology of all mental illness" than it does to speak of an "epidemiology of all infectious diseases" or an "epidemiology of all cancer" (Anthony, Eaton, & Henderson, 1995). We have an epidemiology of cholera, which has some similarities but is not the same as our epidemiology of AIDS, although both are conditions of infectious origin. We have an epidemiology of brain infarction and hemorrhage (stroke), which is not the same as the epidemiology of neurosyphilis or the epidemiology of late-life dementia, even though all of these conditions involve the brain. The epidemiology of suicide attempts is not one and the same as the epidemiology of suicide (Borges, Anthony, & Garrison, 1995; Monk, 1987). The epidemiology of suicide can be distinguished from the epidemiology of major depression, although many cases of major depression become suicide victims (Moscicki, 1995).

There is an epidemiology for "mental health," an epidemiology for individual behaviors (e.g., suicide attempts), and for each of the categories of mental disorders listed in psychopathology's classifications and nomenclatures. For example, we have an epidemiology of schizophrenia, within which hypotheses of genetic causation contend with those of prenatal infections and insults, as well as hypotheses involving later conditions of environment or gene–environment interaction (Jablensky, 1995). We have an epidemiology of drug dependence, also with a new prominence of genetic hypotheses, but with undeniable demands for gene–environment interactions. For example, without encountering an opportunity to ingest cocaine, no one can become cocaine dependent. There now is evidence that genes may determine personality traits that, in turn, influence who seeks out novel opportunities, such as the opportunity to use cocaine. Other evidence points to genetic influences on who becomes cocaine dependent once cocaine is ingested (Uhl, Blum, Noble, & Smith, 1993). Cocaine, in turn, can influence the production of messenger RNA and other gene products of potential later importance to the development of cocaine dependence (Hyman & Nestler, 1996).

Given its orientation to substantive theory, epidemiology has a position in relation to clinical psychology and psychiatry that is analogous to the position of experimental pharmacology and therapeutics or human genetics. There is no unifying theory to master as there might be in physics or sociology, but there is a benefit for the practicing clinician to be familiar with concepts developed or refined within epidemiology, such as "epidemic," "risk factor," "odds ratio," and "exposure opportunity." These epidemiological concepts serve clinicians as well as do pharmacological concepts such as "the dose–response relationship," "tolerance," and "withdrawal," or concepts from genetics such as "heritability," the human genome, and genomic imprinting. In addition, to be able to use or simply to interpret scientific findings that emerge from epidemiology, the clinician should become familiar with its research designs such as the epidemiologic case–control study, and the pitfalls or threats to validity of results in epidemiological research. This familiarity complements what clinicians must learn about placebo-controlled double-blind randomized trials of experimental therapeutics, or comparisons of proband-wise vs. pair-wise concordance for monozygotic vs. dizygotic twin pairs in genetics research.

There are also important opportunities for collaborations between epidemiologists, clinical psychologists, and psychiatrists, and for borrowing aspects of methodology from epidemiology for the purposes of clinical research. As described in this chapter, there is a long tradition for these collaborations, and also for the borrowing of epidemiological methods by psychologists and psychiatrists, who have had no other way to test the psychopathological hypotheses they have framed.

Despite the early misgivings of some leaders in epidemiology (e.g., Frost, 1927; Levin, 1950), and unhampered by psychology's internal debate about its idiographic and nomothetic traditions, there now is much traffic at the intersection where epidemiology meets with psychology and psychiatry. Until about 1980, two years of serious library work was more than sufficient for a scholar to master the history of work at this intersection. This could be accomplished by first developing an acquaintance with every noteworthy book and article on the topic published up to that date, and then by evaluating every new publication on this topic, year by year.

This kind of comprehensive view of the field no longer is within easy reach. Recent upward trends in the number of published articles, shown in Figure 1, have placed a comprehensive view outside the range of all but the most serious full-time scholars in the history of psychiatric epidemiology and epidemiological psychology. Adapting themselves, clinicians interested in epidemiology now generally approach the field as subspecialists. For example, many focus on one or two domains such as the epidemiology of depression, schizophrenia, or drug dependence; likewise for those who take graduate training in public health and epidemiology, but are also interested in psychopathology. Regrettably, subspecialization in any field is associated with a narrowing of perspective on what has been learned. Important threads of historical origin can be lost. Early concepts and principles can be overlooked. This is what we trade in return for benefits that come with the focused attention of subspecialization.

This chapter is intended as a partial antidote to overfocused subspecialization, providing an overview of epidemiology and its origins while also introducing some selected examples of the intersection of epidemiology with clinical psychology and psychiatry. It is not a comprehensive review of all progress in this domain of epidemiology. Instead, the chapter introduces and reviews a useful selection of key organizing features, concepts, and principles in this field as they have been applied to the study of mental health and to categorical disturbances of mental life and behavior. The history of mental morbidity surveys is covered briefly, including longitudinal and prospective studies to estimate the incidence of mental disorders. It seems that these field studies retain a central position, and they are likely to remain central until still-controversial methodological issues have been resolved. By sorting out these methodological issues, greater attention can be given to what should prove to be a more productive debate on substantive hypotheses about causes, causal and protective mechanisms, and intervention strategies.

The final section of the chapter outlines future directions for research and scholarship in this field. It stresses the possibilities for collaboration between clinical scientists and epidemiologists, with recognition that some investigators will wear both hats. Whereas this chapter provides an orientation to background, concepts, and principles, a companion chapter in this collection provides a more detailed introduction to some specific methods of epidemiology as applied to the study of psychological and psychiatric disturbances (see Chapter 3.06, Volume 3).

All of these topics are covered more thoroughly in other works. There are textbooks of epidemiology, including Major Greenwood's early classic *Epidemics and crowd-diseases: An introduction to the study of epidemiology* (Greenwood, 1935), as well as those more recently published by Morris (1957), A. M. Lilienfeld (1980), D. E. Lilienfeld & Stolley (1994), and Gordis (1996). There are parallel developments at the intersection of epidemiology, psychiatry, and psychology, including useful texts (e.g., Cooper & Morgan, 1973; D. Goldberg & Huxley, 1992; Tsuang, Tohen, & Zahner, 1995). Short review articles and useful brief monographs also exist (e.g., Bland, 1988;

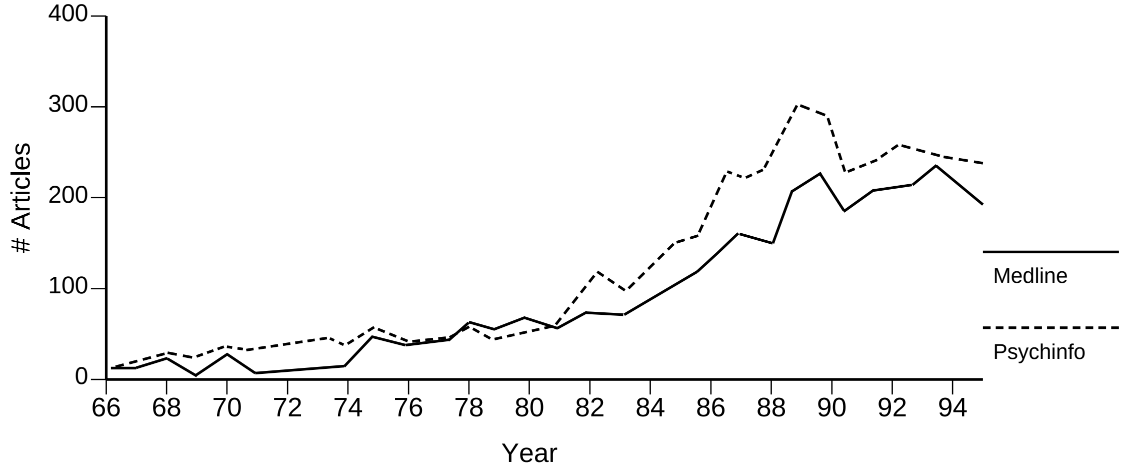

Figure 1 Time trend for number of journal articles published on the topic of epidemiological psychology/psychiatry during the period 1966 through 1995: results of Medline and Pychinfo searches. (Note: Boolean searches were used with the following text words: Epidemiolog$ and Psychiatr$ or Psycholog$ or Mental.)

Cooper, 1993; Morrison, 1959; Reid, 1960; Robins, 1978; Rutter, 1981; Shepherd, 1985; Shepherd & Cooper, 1964).

When seeking examples to illustrate the rubrics of epidemiology, we often turned to the *Annotated bibliography of psychiatric epidemiology*, edited and published in 1992 by Michele Tansella, Giovanni de Girolamo, and Norman Sartorius. This useful resource represents a recent effort to compile articles and books reporting achievements in this field which 44 experts from 21 countries had ranked as "outstanding or of special relevance" (Tansella et al., 1992).

1.12.2 BACKGROUND

Epidemiology and its rubrics came to the surface during the nineteenth century, but our field has a long history in a prenumerate form. For thousands of years, physicians and merchants have written about the coming and going of human diseases in relation to months and seasons of the year, according to conditions of the land (e.g., swampy vs. arid), and across tribes or families of people. The most famous of these ancient notes on the history, medical geography, and anthropology of human suffering is a collection pulled together under the name of one person: *Of airs, waters, and places*, attributed to Hippocrates.

A physician of the late Renaissance, Girolamo Fracastoro, followed in these footsteps and observed that variations in occurrence of disease might be due to invisible but identifiable agents in air and water. In the late seventeenth century, Ramazzini differentiated diseases and causes of death found to be more common in some occupations, less common in others. For example, the phrase "mad as a hatter" and Lewis Carroll's "Mad Hatter" reflect an appreciation of the toxic effects of the mercury used to prepare felt for hats, which includes an intoxication characterized by prominent disturbances of mental life and behavior. As exemplified by this linkage between felt making and mercury poisoning, a common thread running through these early origins of epidemiology was an attempt to find understandable patterns in the occurrence of disease and death, expressed in relation to conditions of time, place, and person.

Epidemiology's modern era and more quantitative orientation emerged from seventeenth-century advances in mathematics, including an early "life table" method developed by Edward Halley of comet fame, as well as more basic foundations of probability and statistics. While Europe passed through the Age of Enlightenment and the Industrial Revolution, these advances were refined and then applied in work to promote the public's health and well-being. The inspiration for this work was drawn in part from the social philosophy of French mathematician Condorcet, eighteenth-century founder of the social sciences. The momentum carried over from mathematics and social philosophy into medicine—as expressed by the use of numerical methods by French physician Pierre Charles Alexander Louis and his students, in order to improve patient care and the health of the public (Farr, 1885; A. M. Lilienfeld & Lilienfeld, 1979). One of Louis's students was William Farr of London, who defined the field of "vital statistics" during a long career of studying Great Britain's administrative statistics on births, deaths, and disease.

In early nineteenth-century Great Britain, waves of Asiatic cholera started to sweep through London and other European communities, regularly killing thousands. Responding to the cholera epidemics, Farr used Halley's life table and numerical methods taught by Louis to study cholera's patterned aftermath in mortality statistics. Based on this work, Farr helped to frame the majority position of London's Epidemiological Society: that breathing miasmatic vapors from unclean water or soil caused cholera. Anaesthesiologist John Snow thought otherwise, and chose to work against common opinion. Aided by Farr, Snow assembled the available data on cholera deaths, gathered his own data, and used numerical methods to marshal evidence that cholera was due to drinking contaminated water. Snow's thesis about cholera as a waterborne disease at first was rejected by his peers, including Farr. Nevertheless, in time, Snow's evidence became overwhelming; the miasma theory was overturned. England and the other industrializing countries began to prevent cholera epidemics by improving water sanitation—years before bacteriologist Robert Koch finally discovered the mechanism of action and identity of the cholera vibrio, which we now trace as the disease's specific cause. For this victory of research on cholera, John Snow is now known as a father of epidemiology.

The numerical methods of Louis also crossed the Atlantic to the Americas (Osler, 1897), where nineteenth-century diseases such as yellow fever and pellagra compromised the public's health and prompted major investigations, just as AIDS has done in our own time. The history of epidemiological research on these diseases teaches one of the most important lessons learned by Snow in his work on cholera: Effective prevention does not require thorough knowledge of specific causes nor causal mechanisms. Methods of prevention can be identified and put into place well before definitive etiological evidence is available. Furthermore, firm evidence about etiology and causal mechanism does not always translate into effective prevention.

In the case of yellow fever, many nineteenth-century authorities believed that the disease was spread by inanimate particles ("fomites") in the clothing, bedding, and other possessions of infected persons. In the name of prevention, millions of dollars worth of these possessions were tossed away or burned, until epidemiological and clinical experiments under the direction of Walter Reed's Yellow Fever Commission produced conclusive evidence that yellow fever was mosquito borne. With this knowledge, the commission began an effective program of yellow fever prevention by draining puddles and swamps, installing nets and screens, and otherwise creating effective barriers between mosquitoes and their human hosts. Many years passed before the yellow fever virus was identified and an effective preventive immunization was created. Thus, effective yellow fever prevention was underway well before it became possible to trace the causal mechanisms and pathways by which the virus passes from mosquitoes to humans.

Pellagra is known as the disease of the three Ds: dermatitis, diarrhea, and dementia (e.g., see Macleod, 1978). Although pellagra was not mentioned by the ancients, it aroused attention during the nineteenth century, especially in the southern United States, Latin America, and other places where maize had become a staple cereal and where other nutrients were in short supply. At the dawn of the twentieth century, the connection with maize was overshadowed by the fact that many of the observed patterns of pellagra occurrence resembled those of an infectious disease with respect to matters such as seasonal variation and clustering of cases within vulnerable subgroups (e.g., the poor). Based on success with cholera, rabies, and other infectious diseases, there was a bacteriological optimism about pellagra, with hope for an early discovery of the responsible bacterium. Nonetheless, from the perspective of Joseph Goldberger, the American father of epidemiology (Goldberger, 1964), there was a different way to interpret the data. In Goldberger's mind, the evidence on pellagra as an infectious disease was just as readily interpretable as evidence favouring a nutritional deficiency—perhaps a deficiency of one of the "vitamines" named by Casimir Fund, a pioneer in the new science of nutrition. Goldberger wrote:

> In view of the great uncertainty that exists as to the true cause of pellagra, it may not be amiss to suggest that pending the final solution of this problem it may be well to attempt to prevent the disease by improving the dietary of those among whom it seems most prevalent. In this direction I would urge the reduction in cereals, vegetables, and canned foods that enter to so large an extent into the dietary of the people of the South and an increase in the fresh animal food component, such as fresh meats, eggs, and milk (Goldberger, 1914a).

A year later, illustrating how initial epidemiological findings about the location of observed cases can lead toward investigations into prevention and causes, Goldberger explained his reasoning (Goldberger, Waring, & Willets, 1915):

> In a paper published [last year] . . . , attention was called to certain epidemiological observations relating to pellagra which appeared inexplicable on any theory of communicability. These observations showed that, at certain institutions at which pellagra was either epidemic or had long been endemic among the inmates, the nurses and attendants, drawn from the class economically and socially identical with that most afflicted in the population at large, appeared uniformly to be immune, although living in the same environment and under the same conditions as did the inmates. Neither "contact" nor insect transmission seemed capable of explaining such a phenomenon. It was suggested that the explanation was to be found in a difference, which was believed to exist, in the diet of the two groups of residents. . . . It was suggested, therefore, that it might be well to attempt to prevent the disease, . . . by providing those subject to pellagra with a diet such as that enjoyed by well-to-do people, who as a group are practically free from the disease.

Thereafter, Goldberger conducted this very type of experiment, as well as other key studies. In so doing, he produced evidence necessary to confirm that pellagra and the pellagra psychoses are not communicable diseases, and were preventable by appropriate changes in diet such as he had described in 1914. Of course, for many, the evidence from his diet experiment was among the most compelling evidence in favor of his dietary hypothesis. Subsequently, prevention of pellagra became a topic of public activity and public policy, well before 1926, when the pellagra preventive factor was narrowed down to one of the B group of vitamins. Goldberger died in January 1929. Almost a decade passed before a research team at the University of Wisconsin identified niacin as the specific protective factor in relation to the etiology of pellagra.

In the meantime, pellagra has essentially disappeared from the United States, Canada, Western Europe, and other places where either high standards of living or other mechanisms promote niacin-containing diets. For the time being, it has become difficult for US-based psychopathologists to observe cases of pellagra psychoses except when the clinical picture is secondary to chronic alcohol dependence. To see primary pellagra psychoses, psychopathologists typically must travel abroad to countries where various constraints thwart good nutrition, and with much lower standards of living than can be found at home.

A more contemporary history lesson on our capacity to start effective prevention campaigns before full knowledge of causes or causal mechanisms is provided by the recent epidemic of human immunodeficiency virus (HIV) infections and the acquired immunodeficiency syndrome (AIDS). Well before HIV had been isolated as an infectious agent, and before general concepts or diagnostic criteria for AIDS could be specified, epidemiologists and clinicians at the Centers for Disease Control were able to spot an unexpected occurrence of *Pneumocystitis carinii* pneumonia. They narrowed its predominate location to young homosexual men, some being consumers of amyl nitrite ("poppers") and other drugs, but also positive for exposure to a cytomegalovirus (Centers for Disease Control, 1981a, 1981b, 1981c; Goedert et al., 1982). As the story unfolded, the epidemiologists gathered enough information to shift gears from a search for locations (where are cases most likely to be found?) to a more focused search for causes (why are some affected while others are not?). In this process, they were able to identify inhalant and other drug use as a correlated behavior or cofactor lacking singular causal importance. While clarifying that drug use by itself was not a cause of HIV infection, these studies noted the high-risk status of unprotected receptive anal intercourse with multiple male sex partners, often in conjunction with the use of poppers and other inhalant drugs, or with crack or other psychostimulant drugs. Prevention campaigns to reduce the risk of AIDS in the homosexual community were underway before the specific viral etiology was understood, and before the risk to heterosexuals was appreciated fully. Because the prevention campaigns target the proximal developmental antecedents of high-risk HIV-related behaviors, these early campaigns retain utility for the prevention of HIV infection and AIDS even today, years after isolation of HIV infection as the necessary causal factor for AIDS (National Institute of Allergy and Infectious Diseases [NIAID], 1995).

Drawn from the history of epidemiology (Wynder, 1994), the lesson that effective prevention can precede firm knowledge of causes, causal mechanisms, and developmental sequences is of special importance in relation to disturbances of mental life and behavior. With few exceptions, it is unlikely that definitive and thorough evidence on the causal or protective mechanisms and developmental processes of these disturbances will be available in the near future. But there is no need for research on prevention of these disturbances to await such definitive and thorough evidence, even though it must be acknowledged that every important advance in understanding basic causal mechanisms and developmental processes typically reveals new targets for prevention and intervention. The value of these advances in public health terms is that each advance can lead to discovery of more efficacious interventions with fewer unwanted side effects and at less cost. In the meantime, preventive trials can move forward and become part of the arsenal of scientific tools used to discover basic causal and protective mechanisms (Mrazek & Haggerty, 1994).

As it happens, we still do not have completely definitive and thorough evidence on the causes and causal mechanisms for any of the infectious or nutritional diseases mentioned above, even though we know how to prevent them, at least in theory. That is, the identification of specific etiological agents for cholera, pellagra, and AIDS did not constitute final brushstrokes in the total picture of disease etiology. In Farr's day, for example, the cholera death rate was lower for the well-to-do than for the impoverished people of London. Apparently, this in part was because wealth was protective: it coincided with access to clean water supplies before there was a complete appreciation of the link between water sanitation and cholera. The association with wealth also seems to be present in the recent upsurge of cholera epidemics during the late twentieth century. Nonetheless, it would be imprudent to delay our efforts to prevent cholera until we understand all the different ways that wealth might promote access to pure water supplies, or until we understand the determinants of the distribution of wealth in society as it might relate to health status. In public health work, the details of these causal mechanisms are of secondary importance when water sanitation, vaccines, and other effective preventive maneuvers are already known.

In Goldberger's day, the well-to-do rarely developed pellagra, as noted in his article from 1915, quoted above. Attempting to develop a more complete view of the epidemiology of pellagra, Goldberger used social science methods designed by a collaborator (Edgar Sydenstricker, an economist). In so doing, his research group was able to document a socioeconomic substrate which was one step back from dietary deficiency in the causal chain leading toward pellagra. In later years, Sydenstricker turned his attention to medical insurance plans and other societal-level means of promoting public health (Sydenstricker, 1964). Needless to say, this aspect of the causal network for diseases associated with socioeconomic status (SES) continues to be underinvestigated.

Turning attention from the more sociological aspects of causal mechanisms for disease to a fundamental biological aspect, we still are many years away from definitive evidence on the chromosomal locations of those genes responsible for our species' widely shared vulnerability to *Cholera vibrio*, or to niacin deficiency. These genes, combinations of genes, and their gene products represent yet another unexplored part of the causal mechanisms and developmental processes leading toward diseases for which we now have at least some effective preventive interventions.

Finally, in our own time, except for a vocal minority in the scientific world, most scientists are confident that the underlying cause of AIDS is a virus, and not the inhalants, the drinking, cocaine use, or sex *per se*. But the person-to-person spread of this virus, by routes including unsafe sex practices and drug injection practices, is determined at least in part by drug-taking behavior and behavioral disturbances; by alcohol and cocaine use, as well as the sex-associated behaviors of inhaling poppers, or taking methamphetamine via inhalation of smoke ("ice") or injection. These behavioral characteristics by which HIV infection is spread are in turn determined by a broad array of influences that range from the domain of genetics to that of social policies toward drug use and treatment of drug dependence. Hence, even though HIV is confirmed as a necessary cause of infections that lead to AIDS, this evidence did not complete the picture of causal mechanisms for HIV infection and AIDS. In the present instance, knowledge of a necessary cause has not yet produced an effective means to block the now global epidemic spread of this infection (NIAID, 1995).

1.12.3 THE RUBRICS OF EPIDEMIOLOGY

Whereas the search for causes, causal mechanisms, and methods of disease prevention and health promotion are central to epidemiology, many in this field choose to concentrate on the other rubrics listed in Table 1. Indeed, epidemiology's first intersection with the subject matter of clinical psychiatry and psychology addressed pragmatic questions. How many persons suffer from mental maladies? Where in the community population are these cases to be found? Are they being treated badly or well?

1.12.3.1 Rubric 1: How Many? A Question of Quantity

Many clinicians in the United States are interested in the mental and behavioral disturbances that affect adolescents and adults in the age range 15–54 years; fewer are acquainted with epidemiologic estimates on the prevalence and incidence of these disturbances (e.g., Kessler et al., 1994; Regier et al., 1993). Many do not appreciate that more Americans in this age range have been affected by a history of

Table 1 The five main rubrics, subject matter, and questions of epidemiology.

Main rubrics	*Subject matter*	*Questions asked by epidemiologists*
(i) How many?	Quantity	How many in the population are affected as cases, or are becoming cases?
(ii) Where?	Location	Where in the population are affected cases more likely to be found, with respect to characteristics, conditions, and processes of person, place, and time?
(iii) Why?	Causes	What accounts for some people becoming affected while others are spared?
(iv) How?	Mechanisms	What sequences or linkages of states and processes influence who becomes and remains a case and who does not?
(v) What can be done?	Prevention and intervention	What can be done to prevent or delay the onset of suffering, to shorten its duration, or otherwise to reduce the burden of affected cases, their families, and society in general?

dependence on psychoactive drugs than by other psychiatric disturbances now accorded higher priority in mental health service delivery systems, prevention, and sponsored research programs. For example, the most recent nationally representative sample survey of 15–54-year-olds in the USA, indicates that cannabis (marijuana) dependence has affected about 4.2% in this population. By comparison, this survey, conducted between 1990 and 1992, found only 3.5% to have had a history of panic disorder.

> Cocaine dependence was almost as prevalent as antisocial personality disorder (2.7% versus 3.5%) and was about 70% more common than bipolar disorder (1.6% prevalence). Alcohol dependence (14.1%) was somewhat more common than simple phobia (11.3%) and social phobia (13.3%), whereas tobacco dependence at 24.1% was more prevalent than major depressive disorder at 17.1%. (Anthony, Warner, & Kessler, 1994)

The prominence of alcohol and other drug problems in the United States can also be seen when epidemiology seeks to quantify how many persons are currently active cases of mental and behavioral disturbances. This is true whether the case assessment method involves computerized diagnoses based on information from standardized interviews conducted by lay persons (e.g., the Diagnostic Interview Schedule, or DIS) or standardized psychiatric diagnoses made by clinicians after a formal psychiatric assessment with opportunities for cross-checking and probing freely. For example, in the Baltimore (Maryland) arm of the National Institute of Mental Health (NIMH) Epidemiologic Catchment Area community surveys carried out during 1981–1982, the DIS method found 3.6% of the study population with a history of recently active *DSM-III* alcohol use disorders (i.e., with symptoms present during the month prior to assessment). An estimated 1.1% had active drug use disorders involving marijuana and other internationally regulated substances. The DIS estimate for one-month prevalence of *DSM-III* Major Depression was intermediate, at 2.3%. By comparison, the corresponding one-month prevalence estimates based on standardized psychiatric examinations of the study subjects were 6.9% for alcohol use disorders, 3.1% for drug use disorders, and 1.1% for major depression. The most commonly found mental disorder by both methods of case assessment was phobic disorder: according to the DIS method, 11.2% of the Baltimore study population had phobic disorder with symptoms present in the month prior to assessment. The corresponding estimate based upon the psychiatric examination was 21.3% (Anthony et al., 1985). Other one-month prevalence estimates from the Baltimore study are presented in Table 2.

These recent estimates from epidemiologic surveys of community populations are not definitive for all persons, places, or times; variation in epidemiological estimates proves to be as important as constancy. Nevertheless, point estimates from community surveys of mental and behavioral disturbances have been indispensable when policy makers and planners have faced hard decisions about allocation of scarce health resources, and when society has to

Table 2 One-month prevalence of selected *DSM-III* conditions, by two methods of ascertainment. Data from the Eastern Baltimore Mental Health Survey conducted as part of the NIMH Epidemiologic Catchment Area Program, 1981.

DSM-III category	*Prevalence estimate based upon standardized clinical diagnosis by study psychiatrist* (%)	*Prevalence estimate based upon Diagnostic Interview Schedule administered by lay interviewer* (%)
Phobic disorders	21.3 (1.9)	11.2 (1.2)
Alcohol use disorders	6.9 (1.1)	3.6 (0.8)
Drug use disorders	3.1 (1.0)	1.1 (0.5)
Major depressive episode	1.1 (0.4)	2.3 (0.4)
Schizophrenia	0.5 (0.1)	0.7 (0.2)
Manic episode	0.4 (1.2)	0.4 (0.9)
Obsessive-compulsive disorder	0.3 (0.1)	1.3 (0.3)
Panic disorder	<0.1 (0.07)	0.8 (0.2)

These results are based upon multistage sequenced sampling of adult household residents living in eastern Baltimore at the time of the survey. The diagnoses were from independently conducted standardized psychiatric examinations and Diagnostic Interview Schedule assessments of the same subjects, who were selected by probability sampling from the community and also by probability sampling for the diagnostic assessments. These estimates take the varying probabilities of selection into account, as reported in Anthony et al., 1985, from which this table is adapted. Numbers in parentheses are the standard errors of the observed prevalence estimates.

weigh the relative public health importance of one health or social problem vs. another. For many years the allocation of resources to the mental health field has been constrained by an overemphasis on the study of death certificates, the causes of death, and prevailing mortality rates. During the twentieth century, there has been an increasing appreciation that conquest of communicable diseases yields a worldwide increase in the prevalence of mental and behavioral disturbances that disable but do not necessarily kill (e.g., see Gruenberg, 1977; Kramer, 1983; Levav, Lima, Somoza, Kramer, & Gonzales, 1989). Recent evaluations of the global burden of diseases take both prevalence and disability into account, drawing estimates of prevalence and levels of disability from epidemiologic surveys. These evaluations show major depression and alcohol use disorders to be at the forefront of the major disabling health conditions that face countries worldwide, now that more humans are living into adulthood when they otherwise would have died during infancy and childhood (Murray & Lopez, 1996). The same process of reduced mortality with increased survivorship has given Alzheimer's disease and the dementias of late life a much more prominent aspect than human society has ever faced before (Gruenberg, 1977).

The history of epidemiologic surveys of mental and behavioral disturbances dates back to the early nineteenth century. A famous Professor Holst of Norway is cited as having been responsible for the first formal community survey of mental disorders. This survey was completed in 1835 in order to assess the number of citizens who were suffering from mental disturbances and how they were being attended. The method involved two stages of case assessment. In the first stage, the clergy were asked to review each roster of parishioners, which included all citizens of the area served by the church without respect to religious preference, and to make notes about the mental and behavioral characteristics of each listed person. In a second stage, these notes were compiled and reviewed by a royal commission of alienists (psychiatrists of the time), who sorted the disturbances into categories of idiotia, dementia, melancholia, and mania. Based on these assignments, the survey team then produced estimates for the prevalence of each of these conditions in the urban and rural areas of Norway, and for men and women (Table 3; Holst, 1835, translated by Massey, 1852).

In the United States, the first formal epidemiologic surveys of these disturbances were the responsibility of the Census Bureau. In the 1840 census, the census takers were asked to identify cases of insanity as they went from community to community, household to household, and institution to institution, with case assessment either directly observed or reported by household, community, or institutional leaders. Edward Jarvis, who was William Farr's American counterpart, compiled many of the results of the survey, with both good and bad results. Jarvis, a physician, was an early leader in the American Statistical Association, and his name should be remembered in this context for two important reasons.

First, Jarvis studied patterns of admission to mental institutions, and found that the probability of being admitted to such an institution

Table 3 Prevalence estimates from an early community survey of mental disorders in 1835 in Norway.[a]

Prevalence (cases per 10 000)	*Mania*	*Melancholia*	*Dementia*	*Idiotia*
All citizens	6.05	5.31	4.35	14.21
Males	6.20	5.19	4.46	15.12
Females	5.91	5.43	4.25	13.34
Citizens of towns				
Males	9.27	5.69	5.60	7.97
Females	9.03	6.66	3.85	5.03
Rural districts				
Males	5.84	5.13	4.31	15.96
Females	5.52	5.28	4.30	14.38

[a] These prevalence estimates (cases per 10 000 population) are from a report written by Professor Holst of Norway, 1835, translated by Massey, 1852. As described in the text, the clergy of each parish in the country made notes about the mental life and behavior of each area resident. These notes were reviewed by a panel of specialists, who sought to identify the mentally ill, and to sort their disturbances into the categories listed above.

varied inversely with the distance of one's residence from the institution (Jarvis, 1865–66). This relationship, now known as "Jarvis's Law" still held when it was checked in the mid-twentieth century (e.g., Kramer, Goldstein, Israel, & Johnson, 1955). Second, Jarvis was proud of the results of the 1840 census, and praised them as a major advance in the understanding of insanity in the United States. Illustrating his point, he used the census data to calculate the prevalence of insanity among "colored" persons living in the slave states of the South vs. the corresponding prevalence among those living in the free states of the North. In so doing, he found and reported that the prevalence of insanity in this group was lower in the South than in the North, only to discover, after his publication of the results, that serious errors had produced an undercounting of mentally disordered slaves in the South. Despite repeated retractions, and to his great dismay, Jarvis was unable to stop use of this statistical evidence in the proslavery arguments that ultimately led to the great Civil War between the North and the South (Deutsch, 1944; Jarvis, 1843a, 1843b).

Another neglected classic survey was conducted during 1916 by Aaron J. Rosanoff, a psychiatrist who worked in Nassau County, New York. Rosanoff's two-stage community survey involved a sophisticated first-stage assessment designed to screen for mental and behavioral disturbances via brief tests and item sets (e.g., cognitive tests of orientation and memory, and proverb interpretation). In the second stage, Rosanoff or another psychiatrist colleague examined every one of the individuals suspected of having some mental or behavioral disturbance. After this direct examination, they sorted the disturbances into discrete categories from the classification of neuropsychiatric conditions of the era, and calculated prevalence estimates for each of the categorized disturbances. Though the results of this early sophisticated community survey of mental disorders were used to help plan for mental health services in the county, the survey had an unfortunate origin in the eugenics movement. One of its stated goals was to identify families who were responsible for bearing offspring with mental and behavioral disturbances, and doing something about these families (Rosanoff, 1917).

This chapter is too brief to provide a complete account of the history of community surveys of mental and behavioral disturbances. Many surveys of this type have been completed during the twentieth century, both in the US and abroad (e.g., see Cooper & Morgan, 1973; Dohrenwend & Dohrenwend, 1974; Lemkau, Tietze, & Cooper, 1942; Robins & Regier, 1991; Schwab & Schwab-Stone, 1978; Tansella et al., 1992). For the present purposes, it might suffice to say that the impulse to count the number of persons affected by mental and behavioral disturbances in the community was first expressed in scientific studies during the first half of the nineteenth century. This impulse grew to full strength during the twentieth century and continues unabated. First oriented to counting how many persons in the population are affected by each type of mental and behavioral disturbance (i.e., the prevalence of each disorder), this epidemiologic research has expanded to other domains. For example, via prospective and longitudinal research designs, epidemiologists now seek estimates for the risk of becoming a case for the first time (i.e., the incidence of each disorder), and also seek to gauge the size of an "iceberg" phenomenon (i.e., the relative proportion of active cases who have come to the attention of health authorities vs. the proportion who remain without attention or care). Figure 2 illustrates the iceberg phenomenon, which might help convey how it is that epidemiology helps clinicians see beyond the horizons of clinical practice (Eaton et al., 1989).

Of course, once questions about treated and untreated cases of mental disorder are raised, it becomes necessary to account for why some cases are recognized and treated, while others are not. The transition from being a case in the community to being a recognized case and receiving clinical attention has been characterized in relation to filters that promote or impede these transitions. As identified by Goldberg and Huxley (1980, 1992), the progress of a case from the community toward psychiatric attention appears to be influenced strongly by the type and severity of the mental disorder (see Gallo, Marino, Ford, & Anthony, 1995; Marino, Gallo, Ford, & Anthony, 1995). In this expanded research mission, epidemiology turns to its second rubric.

1.12.3.2 Rubric 2: Where? A Question of Location

Either before or along with the capability to estimate how many persons in a population are dying or are falling ill, epidemiology can provide a means for making separate estimates for geographic locations or political jurisdictions, for different chronological ages or time periods, and for shared or individual circumstances such as social class and sex. In many circumstances, these estimates are presented in relation to dimensions of location. For example, one such dimension was the distance of one's residence from a mental hospital, as expressed in Jarvis's Law (Jarvis, 1865–66). Dimensions of socioeconomic status also have been prominent in epidemiological studies of mental disorders. For example, in one recent study, a research group led by psychologist Bruce P. Dohrenwend presented estimates for the occurrence of mental disorders in relation to a dimension of social status defined according to educational attainment. Separate estimates were presented for persons located at different positions along that dimension (Dohrenwend, et al., 1992).

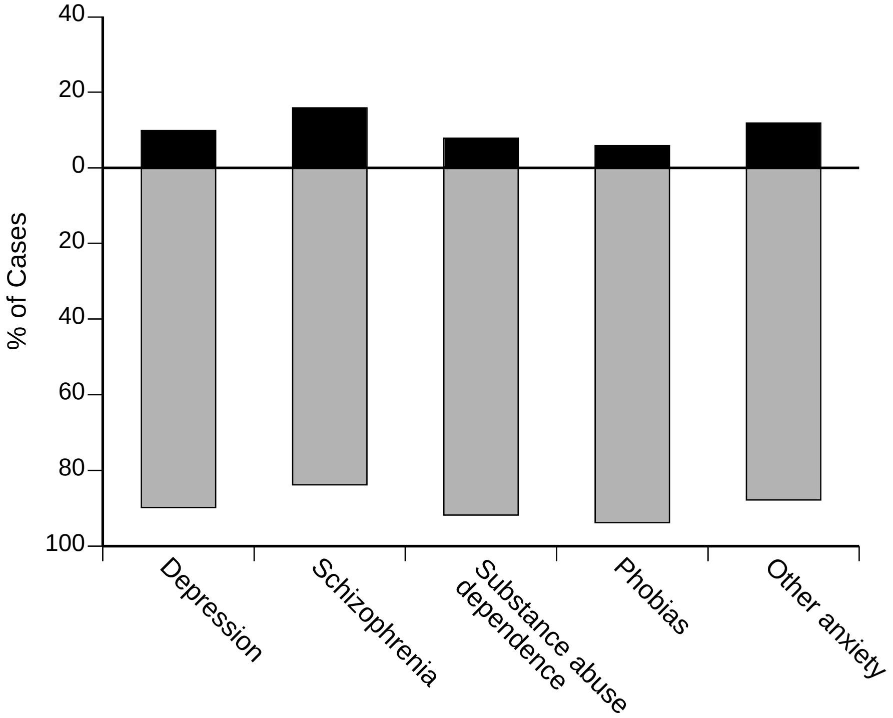

Figure 2 The "iceberg" phenomenon in the epidemiology of mental disorders. Data from the National Institute of Mental Health Epidemiologic Catchment Area (ECA) surveys, 1979–1985. The total area of each bar represents 100% of the active cases, in each of the listed categories of *DSM-III* disorders, found in the ECA surveys. The solid area above the x-axis represents the proportion of cases who had been seen by a mental health specialist (the "tip of the iceberg"), while the hatched area below the x-axis shows the proportion of active cases not seen by mental health specialists (Marino et al., 1995).

Often, the locational dimensions of time, chronological age, and stages of development have been of central importance in epidemiological studies. For centuries, epidemics have been observed to come and go, either varying in relation to the season of the year (e.g., influenza) or with much longer between-epidemic intervals (e.g., smallpox, cholera). Many communicable diseases such as measles, chickenpox, and mumps were discovered to affect mainly the young and rarely the old: a manifestation of durable immunity after an initial attack. While these epidemics raged among older children, newborn infants might be spared for a few months, due either to immunity acquired from the mother, or because of being swaddled or otherwise protected from contact with the infective agent. Those who managed to escape relatively mild disease during childhood might be faced with much more serious medical complications once the years of puberty were reached, as is the case with the mumps.

On occasion, the dimension of time serves as a marker for preventive actions, and early epidemiologists discovered that studies of disease occurrence over time might help them to understand the preventive value of these actions. For example, decades before Farr was able to rely upon Great Britain's national vital statistics data for calculation of mortality rates, Robert Watt had studied the burial registers of Glasgow, Scotland, to learn the annual numbers of smallpox deaths during the period 1783–1812. Watt observed generally declining numbers of smallpox deaths; proportionately, smallpox deaths became much less prominent among all deaths, despite generally increasing population size in Glasgow. Watt attributed the observed decline in the occurrence of smallpox deaths to the initiation of smallpox vaccination in that area (Watt, 1813, cited in Farr, 1885).

Nonetheless, in epidemiology, the question of where (and when) we are more likely to find cases occurring often runs ahead of questions about causes, mechanisms, and prevention or intervention, especially when there are no strong theories to guide more probing investigations. This has had some advantages, because intervention and prevention-oriented services can be mobilized toward the locations of incident and active cases, once there is some firm evidence about location, and even before there is good theory to guide causal investigations. This was true in the history of the HIV and AIDS epidemics, as noted above: interventions were directed toward the gay community, and especially toward those in the gay community who were engaged in high-risk sexual encounters tending to facilitate the spread of sexually transmitted diseases. The nature of the connection with sexual encounters was not understood at all, but the capacity to start developing interventions directed toward apparently risky sexual activities was created early in the epidemic.

Sometimes we need to sustain basic descriptive efforts to learn where cases are more likely to be found, even when advances are being made in tests of competing theories about causes and causal mechanisms. This certainly is true in the instance of the epidemiology of suicide, which has been studied carefully for about 160 years, starting with William Farr's discovery of evidence both conforming with and contradicting Par M. Bronc's "social law" that "suicide is most common where education is most diffused: suicides and scholars increase in the same ratio." The supportive evidence was based upon a comparison of four jurisdictions in which it was possible to review marriage records in order to determine the proportion of persons who were able to write their names at the time of being married. Based on this indication of education in these four areas, it was found that the number of suicides per 100 000 population was 2.2 in the district with the lowest level of education and 10.9 in the district with the highest level of education. Of course, this type of epidemiological evidence was liable to the "ecological fallacy" in that it is operating at a unit of analysis that is more aggregate than individual. There was no assurance that it was the more educated persons in these districts who were committing suicide. Pursuing this line of reasoning, Farr found the contradictory evidence: namely, in a comparison of the number of suicides per 10 000 males, Farr found a value of 6.7 suicides per 10 000 among servants and coachmen, vs. a lower value of 4.9 per 10 000 among capitalists, professionals, and other educated men (Farr, 1838, reprinted 1885).

Later in the nineteenth century, Durkheim (1897) tested more complex theories of suicide in relation to social structural characteristics and social constructs such as anomie, using vital statistics data on suicide mortality rates. Nonetheless, we still have not come to a firm understanding of the causes and causal mechanisms of suicide (Monk, 1987). Moreover, in recent years, by plotting the age-specific suicide mortality rates over time, in the United States and in many other countries, it has been possible to spot an increasing occurrence of suicide deaths among young people and in late life, especially among elderly men in the United States (Moscicki, 1995). This work illustrates useful epidemiological efforts to identify the location of incident cases in terms of the increasing trends over time and in terms of these two age subgroups within the population.

With firm evidence that these are real and persisting increases, and not insignificant and transient phenomena, we have all the more reason to probe more deeply, to frame, and to test theories about causes, causal mechanisms, and the means of prevention.

Farr maintained his interest in suicide mortality throughout his career. Toward the end of his life, he completed an international comparison of suicide deaths and other deaths from violent means (e.g., homicides). In so doing, he found substantially greater numbers of suicides per million living persons in Switzerland (196), Prussia (134), Austria (113), and Bavaria (103), as compared to Ireland (21), Finland (34), Italy (37), and Scotland (37) (Farr, 1885). Farr was well aware that differences such as these might be explained by variation in the death certification process (Farr, 1885).

Complementing these nineteenth-century analyses to investigate where suicide mortality was more likely to be occurring, at that time there also were analyses on the location of active cases of mental disorders. We already have mentioned Jarvis's troubles in connection with his analyses of the 1840 US Census data on insanity among free and enslaved African-Americans in the various states of this country. The early nineteenth-century Norwegian survey included prevalence estimates for specific mental disorders, cross-classified for urban vs. rural parts of the country, and also by sex and age (Holst, 1852, translated by Massey), as presented in Table 3. In Wine's report on the 1880 US Census, detailed tables were presented to show both the numbers and proportions of US residents affected by insanity and idiocy in each state (i.e., the state-specific prevalence values), as well as the age of onset of these conditions, with cross-classification analyses according to the sexes, various ages, and social groups. The report also gave consideration to the excess frequency of insanity among foreign-born residents (3881 per million inhabitants) vs. native-born inhabitants (1509 per million), and to the hypothesis that foreign nations were shipping their insane citizens to the shores of the United States. Addressing this hypothesis, Wines reanalyzed the census data, showing that the excess was much reduced when a correction was made by subtracting persons under the age of 16 from the denominators of these prevalence values. This correction was motivated by his consideration of the age of first onset of insanity, which was observed quite infrequently prior to age 16 (Wines, 1888).

In the nineteenth-century literature on psychological medicine, there also were fairly sophisticated discussions of the methodological problems associated with making international comparisons of mental morbidity statistics, as well as comparisons between the sexes and other variations in the liability to mental disorders (e.g., with reference to marital status). For example, in Tuke's *Dictionary of psychological medicine*, there is a serious critique of previous methods of calculating the relative liability to insanity of different communities, with an emphasis on the problem of whether the numbers of insane persons were counted in a methodologically comparable way across countries. In addition, considering whether there might be increases in the liability to become insane over time, Tuke (1892) noted:

> A moment's consideration will show that the only proper test of the increase of mental disease is the proportion of first attacks to the population during different periods. First admissions [to mental institutions] are clearly not identical with first attacks, seeing that a patient may be admitted into an asylum for the first time, and yet have had one or more previous attacks of insanity.

From Rosanoff's survey in 1916 onward through the twentieth century, it has been uncommon for epidemiologists to estimate how many individuals are becoming affected by mental disorders, or how many individuals have become recently active cases, without also analyzing the variation to the prevalence or incidence of these disorders in relation to location in time, from place to place, or across circumstances. For example, in addition to reporting an overall morbidity risk estimate to age 56 for both sexes, in a nonconcurrent prospective study conducted on Bornholm Island (Denmark), Fremming (1951) estimated the morbidity risk of alcoholism for males as 3.41%, but was unable to do so for females because only two of the women in the study population had developed alcoholism.

In a comparison of this type, even when the goal simply is to locate where the cases are more likely to be occurring or more likely to be found, it is possible to take differences or to form ratios. Taking differences in Fremming's data on morbid risk of alcoholism for males vs. females, we can compute an excess morbidity risk of 3.0% or greater, given that the morbidity risk for males was 3.41% and that for women truly was too small to calculate. Alternatively, we can compute the ratio of the two morbidity risk estimates, which in this instance would prove to be far greater than a threefold excess in morbidity risk for males vs. females. In that the female morbidity risk value is close to zero, an estimate based upon the actual ratio of male morbidity risk to female morbidity risk would tend towards infinity.

In addition to these risk difference and risk ratio computations that are used to quantify the magnitude of variation from one subgroup to another, from place to place, and time to time, etc., we also can transform each morbidity risk estimate into an odds estimate, and then take the ratio of the odds. The term "odds," by conventional definition, refers to a probability divided by the complement of that probability. To the extent that we can ignore the fact that incidence rates are expressed in relation to units of time, if these units are kept the same for the groups being compared, we can convert the standard form of the incidence rate ("probability of becoming a case for the first time") into an odds value by treating its numerical value as a proportion, and by expressing the odds as that proportion divided by the complement of the proportion. Thereafter, we can divide the odds of becoming a case for one group (or time or place) by the odds of becoming a case for another group (or time or place). The resulting "odds ratio" also serves well as a statistical measure when studying variation in the location of ocrrence of cases. Indeed, under some circumstances, the odds ratio approximates the risk ratio defined above (e.g., see Breslow & Day, 1980).

If the proportions under study are prevalence values from cross-sectional survey data, then it has become customary to use the terms "prevalence difference," "prevalence ratio," and "prevalence odds ratio" instead of "risk difference" or "incidence difference," "risk ratio," or "incidence ratio," and "odds ratio." Indeed, if the risk values are cumulative probabilities over some specified interval of time, then there is good reason to use the terms "cumulative risk difference," "cumulative risk ratio," and "cumulative odds ratio" in order to distinguish these statistical measures of variation from the corresponding measures based upon incidence rates.

The same comparative statistics to identify variation in the location of cases (e.g., estimates of prevalence odds ratios and cumulative risk ratios) can be derived from epidemiologic case–control studies, as well as from other comparative study designs related to the case–control paradigm. The case–control research design can be thought of as a special case of the sequenced two-stage (or multistage) survey sample design: in case–control studies we always sample on the outcome variable in the first stage of assessment. In a second stage of assessment, we measure locational grouping variables of interest in the cases and also in sampled controls (noncases). That is, we seek a sample of people who are becoming cases, and who seek a comparison sample of people we have not become cases.

In a case–control study, the values "observed" for the cases can be compared to the "expected values" observed for the noncases, much as a clinician might measure some characteristic of a case, and then make a comparison of this observed value to a table of reference values or "norms" for a reference population. If we are interested in making comparisons about the incidence of a specific condition, we see cases who have just become cases ("incident cases") and we compare them to individuals who have no past history of having become a case ("controls") or we compare them to the reference population out of which the incident cases have surfaced ("a sample of the study base"). In this instance, the magnitude of association between case status and the locational characteristic of interest is typically estimated using the odds ratio. However, it can be useful to think of the odds in a special way. For example, to conduct a case–control study of the association between sex (being female) and major depression, we might recruit a sample of incident cases of major depression as well as noncases with no history of major depression. Among cases, we estimate the odds that an incident case of depression is female, and among the noncases, we estimate the odds ratio that a noncase is female. Properly designed, the case–control study, with its sample based on the outcome variable (major depression), yields a ratio of these two odds that serves as a good approximation to the incidence odds ratio that would be derived from a prospective study in which the sample has been assembled with reference to the locational grouping variable of interest (here, "being female"), and with subsequent follow-up assessments to evaluate whether one group is more likely than another to become a new case of major depression. This is an important aspect of case–control study design that merits reflection: the odds ratio estimate from a case–control study often serves well as an estimate for the incidence odds ratio from a prospective study (Breslow & Day, 1980; Schlesselman, 1982).

Two central advantages of a case–control research design over a prospective study design are that the case–control study can often be completed with less time and effort, and a case–control study also allows testing of multiple hypotheses with a single sample. For this reason, the analysis of case–control study data is often organized as a test of one or more central hypotheses of greatest interest, followed by a more exploratory search for associations of less central interest. That is, once key hypotheses have been tested, analyses of case–control data can shift toward a more exploratory mode of investigation in which the aim is to screen for new, previously unestablished patterns of associations that might serve as leads to causal hypothesis testing in future studies.

Nonetheless, the most powerful uses of the case–control study design have been observed when there is a need for rapid detection of leads for more probing causal investigations, when the supply of available cases is limited (i.e., sample size is small), and when there is no viable opportunity to use a prospective study design or a randomized controlled trial to test alternative hypotheses. This situation can arise when an epidemic outbreak of a disease occurs and there is an urgent need for a means of prevention or other public health action before a deliberate and long-term search for causes can be undertaken. In some instances, randomized trials are ethically impossible (e.g., to test whether fetal exposure to cigarette smoke via maternal smoking might increase a child's risk of later drug dependence). In other instances, the planning of prospective or longitudinal studies requires preliminary information best gained by relatively inexpensive and rapidly conducted case–control study methods.

Under circumstances such as these, and until effective means of public health action are well in hand, it makes good sense to be very enthusiastic about the prospect of completing multiple case–control studies that show where the new cases are located. At a minimum, these studies can yield new information and expertise which then can be harnessed within the framework of prospective studies to test causal hypotheses. The recent history of the international research program on Alzheimer's disease provides a useful example of this linkage from informative case–control studies toward more probing prospective investigations. Specifically, multiple case–control studies on senile dementia of the Alzheimer's type (SDAT) were conducted between 1975 and the late 1980s, and virtually all of these studies were based on cases referred to specialty clinics of neurology or neuropsychiatry. Estimated odds ratios from almost all of these studies pinpointed a past history of head trauma with loss of consciousness as a locational characteristic with excess occurrence of SDAT (Anthony & Aboraya, 1992; Breteler, Claus, van Duijn, Launer, & Hofman, 1992). Given the biological plausibility that head trauma at one stage of human development might cause an increased risk of SDAT at a later stage, and despite the clear possibility that referred cases might be a biased sample of all cases, investigators who learned the ropes of SDAT research via case–control studies now are conducting prospective and longitudinal studies, including longitudinal investigations of twins who are discordant for head trauma, in order to probe the possible causal significance of this observed association. With the human species, these observational studies are the best we can do: a randomized controlled trial of head trauma as a cause of SDAT makes for an interesting thought experiment, but is out of the question.

1.12.3.3 Rubric 3: Why? A Question of Causes

The history of epidemiology's search for causes of mental and behavioral disturbances dates back to the prenumerate period when Ramazzini made observations about occupation-related diseases such as the ailments and premature mortality affecting hatters, medical workers, and others who handled mercury incautiously. As mentioned earlier in this chapter, in the middle of the nineteenth century, William Farr also turned his attention to the subject matter of psychology and psychiatry when he considered links between education and suicide, but stopped short of a probing causal analysis.

Nevertheless, by the middle of the 1800s, John Snow and others had taken Louis's numerical methods for the clinical study of patients and had adapted them for causal analyses of the probabilities and rates of diseases that affect populations. As practiced today, these adapted numerical methods encompass study design and analysis tools such as randomization, stratification, and matching. In the main, all of these refinements of numerical methods are used to create balance or comparability when we search for the causes of human suffering or the means to prevent or reduce this suffering.

For example, by randomly allocating patients to different treatment interventions, we seek to constrain the influence of possibly distorting influences such as sex or age in order to allow the effects of the various interventions to shine through. When we are careless in our work, we simply assume that randomization works to constrain the influence of these variables, even though randomization in any given trial actually might fail, and often does so. When we are careful, we check to make sure that randomization has worked to our benefit, looking to see that the distributions of potentially influential variables are balanced across our various intervention groups. Theory or prior evidence guides us when we choose which pre-intervention variables might be distorting influences on treatment outcome, and the means of assessing these variables. Checking to see that the distributions of these variables are balanced means that we look at statistical indicators for each group in relation to the central tendency of the variables (e.g., mean, median), their dispersion (e.g., variance, standard deviation, interquartile range), skewness, and kurtosis, allowing comparisons within and between the groups to be made.

When we find that the hypothesized influential variables are not in balance across the groups to be compared (i.e., when the groups are "nonequivalent"), we have a number of options at our disposal. For example, we can rerandomize until balanced distributions are achieved. In this fashion, group equivalence can be achieved on all pertinent variables except for the interventions to be compared. Or, when rerandomization is not possible, we have the option of conducting a stratified analysis that will "hold constant" the influential variables when we seek to know whether the intervention had its desired effect. For example, if we think educational level might introduce a distorting influence on the trial results, and if we find out that one randomly allocated intervention group has proportionately too many college-educated patients, we can stratify in our analyses, making sure the distribution of the education variable is balanced within strata before making a contrast of intervention effects. Two special cases of stratification involve (i) matching, which is a kind of "fine-grained" stratification which we use to make sure that the matching variable has the same distribution within the groups to be compared, and (ii) exclusion, which amounts to stratifying and then eliminating from consideration any instances of a powerful distorting variable. For example, after completing all baseline assessments for a trial of a new psychotherapeutic intervention, we might find that, somehow, three individuals already taking a serotonin reuptake inhibitor slipped into one group, and two individuals taking this same antidepressant drug slipped into another group. In this instance, we might well decide to eliminate these individuals from the analytical sample, restricting our study and our analyses to the stratum of individuals not taking such antidepressant medicines.

All of these numerical methods for constraining the influence of possibly distorting variables were available to early epidemiologists and were used by them in their efforts to search for the causes of disease in populations. For example, once Farr had come to agree with Snow's rejection of the miasma theory about the cholera epidemics, and to focus on water supply, he made a before-and-after comparison that sought to clarify the cause of an observed drop in cholera mortality rates, namely, a change in water supply between 1849 and 1854. In his reasoning, he illustrated methods of bringing into balance and holding constant an area's elevation above sea level (unchanged before and after), as well as the presence of poverty, crowding, and other potentially causal factors. As described in Table 4, unless taken into account, these factors (labeled **a**, **b**, **c**, **d**, **e**, and **z**) might have distorted Farr's before-and-after comparison.

The twentieth century has added refinements to this application of numerical methods in epidemiological research, but it is hard to improve upon the logic of Farr's original reasoning. Included among the more recent refinements are: a greater appreciation of the odds ratio as an indicator of how strongly one variable might influence another; widespread use of a multiple logistic regression model, that can be used to express the odds of disease occurrence in relation to set of potentially influential variables (corresponding to Farr's **a**, **b**, **c**, **d**, **e**, and **z** in Table 4); improvements in the estimation of the odds ratio under conditions of matched samples; advances in statistical methods needed to estimate the odds ratio when multiple distorting influences must be kept in check; and software programming plus computational advances which now place these statistical innovations within reach of each individual investigator who has access to a personal or office computer, whereas less than 20 years ago large mainframe computers were required.

In concept, the multiple regression model serves a function analogous to that of stratification or matching in more elementary numerical methods. Namely, the idea is to apply multiple regression models in order to hold constant or adjust for the potentially distorting influence of other variables. Application of this method is not unique to epidemiology as a scientific enterprise, and some investigators in psychology have taken the multiple regression approach far beyond anything to be seen in the public health branches of epidemiology (e.g., see Bentler & Stein, 1992; Newcomb & Felix-Ortiz, 1992). Nonetheless, more than most fields, epidemiology now makes use of the multiple logistic regression model, fostered mainly by graduate training that teaches how to derive and interpret the odds ratio from fourfold cross-classification tables and to extract the same from logistic regression (e.g, see Breslow & Day, 1980; Hosmer & Lemeshow, 1989).

In its search for causes, epidemiology also has benefited from twentieth-century advances in the design of experiments, contributed by Sir Ronald A. Fisher and others. In this respect, epidemiology is no different from other scientific fields that make use of both experimental and observational research designs. However,

Table 4 Epidemiologic reasoning in the nineteenth century.

Farr's reasoning illustrated early methods of bringing into balance and holding constant an area's elevation above sea level (unchanged before and after), the presence of poverty, crowding, and other potentially causal factors which otherwise might have a distorting influence on the comparison. As expressed here in a reorganized outline, and with correction of two minor typographical errors, Farr summarized his reasoning as follows (excerpt from Farr, 1885, pp. 375–377):

(i) The excess of mortality (**x**) by cholera in the Southwark and Lambeth water-fields in 1849 and 1854 was produced by one or other of all the possible causes existing in 1849, repersented by **a**, **b**, **c**, **d**, **e**, **z**.
(ii) In the second period (1854), as well as the first (1849), all the possible causes remained unchanged, except the possible cause **z**, which varied.
(iii) With it varied **x** so as to diminish as **z** diminished. Therefore **z** was the cause of **x**.
(iv) If the enumeration of the possible causes is complete, and given the assumption that the forces of **a**, **b**, **c**, **d**, and **e** are substantially unchanged, it is difficult to resist the conclusion that **z** was the cause of **x**.
(v) The obvious possible causes are exposure to infection (**a**), bad drains (**b**), crowding (**c**), poverty (**d**), imperfect medical relief (**e**), impure water (**z**).
(vi) None of these, as far as it is known, differed, as far as Lambeth is concerned, in 1849 and in 1854, except the water (**z**), which was very impure in 1849, and much purer in 1854.

the epidemiological case–control study design has been refined by epidemiologists through years of usage in this century, first in applications involving research on lung cancer, then on other cancers, mental retardation and other neuropsychiatric disorders of children (A. M. Lilienfeld, Pasamanick, & Rogers, 1955) and, more recently, on a broad array of diseases and health conditions, including schizophrenia, panic attacks, and major depression (Anthony, Tien, & Petronis, 1989; Dohrenwend et al., 1992; Link, Dohrenwend, & Skodol, 1986). Because it tends to be underappreciated outside of epidemiology, this particular research design merits some special attention.

As implied in our introduction to this topic, if the epidemiological case–control study design did not already exist, we would have to invent it. This is not mere boasting on behalf of epidemiologists and biostatisticians, such as Morton Levin, Abraham Lilienfeld, Nathan Mantel, Norman Breslow, and others, who have labored to develop and refine case–control methods. The case–control research design poses some of the most impressive challenges in epidemiology, but epidemiology could not respond to pressing public health problems without this design. An example might help drive home this point. The following illustration of the value of case–control research draws upon the experience of clinicians and epidemiologists who faced a recent epidemic of eosinophilia-myalgia syndrome (EMS), described in detail by Kilbourne (1992) and summarized here.

First identified as an outgrowth of the work of astute clinicians who compared notes on three patients in October 1989, EMS was soon discovered to be affecting more than 1500 people in the United States alone, and with more cases abroad. The main laboratory finding on these cases was striking eosinophilia, increased eosinophils, and eosinophil precursors in bone marrow, but the main presenting ailment of affected cases was severe muscle pain, which generally was severe enough to constrain the patient's usual activities and often was disabling. A great majority of cases also complained of headaches, feelings of numbness or muscle weakness, and fatigue. As the case histories progressed, cognitive deficits associated with EMS were identified, as well as neuropathy due to axonal loss and documented by electromyography and nerve conduction studies. In some cases, the neuropathy was severe and progressive to the point of quadriplegia; some cases required mechanical ventilation to compensate for respiratory insufficiency secondary to the neuropathy. Over the course of a relatively brief epidemic period from 1989 to 1990, a small number of deaths attributed to EMS accumulated, but the epidemic was curbed and its apparent cause was identified in time to keep the EMS-associated mortality to a minimum. This rapid action is attributable to a series of very small case–control studies, all conducted within one month of clinical reports on the first EMS cases; that is, far more rapidly than it would have taken to mount standard longitudinal, prospective, or experimental investigations.

The clinicians reporting the first three EMS cases thought that the appearance of this syndrome had something to do with the patients' ingestion of oral L-tryptophan, an amino acid precursor to serotonin, taken by them as a dietary supplement (e.g., with the intention of improving mood or reducing sleep problems). The reported syndrome brought other possibilities to the minds of epidemiologists at the US Centers for Disease Control (CDC). For example, an epidemic form of a very similar syndrome had developed in Spain during the early 1980s, ultimately affecting 20 000 cases and accounting for 300 deaths. Traced by case–control and clinical methods, the Spanish epidemic ultimately had been linked to ingestion of rapeseed oil, though the specific causal mechanism or contaminant still remains unknown. Another suspected cause was the nematode *Trichinella spiralis*, which causes trichinosis when ingested via undercooked and contaminated pork. Indeed, a similar syndrome might appear as a consequence of neoplasms, and when no more than a handful of cases had been reported, this possibility could not be dismissed entirely.

Once the CDC investigators set up a system to collect clinical reports on cases of this syndrome, the known number of affected cases mounted. By July 1990, confirmed reports of EMS had been received from every state of the USA, and more than 1500 cases had been reported. Prevalence, calculated as the number of reported cases per million (cpm) population, was greatest in New Mexico and Oregon (>20 cpm population), but also was high in South Carolina (15.1 cpm) Minnesota (14.5 cpm), Colorado and Arizona (13 cpm), California, Idaho, and Washington State (9–10 cpm). The lowest value was from Louisiana, with only 0.2 cases per million population.

Two case–control studies, one in New Mexico and one in Minnesota, proved to be pivotal in establishing an empirical link between EMS and oral L-tryptophan, despite a background of observed associations with other factors, and suspected causes such as trichinosis and cancer. For example, from the early case reports, a female excess was apparent: 84% of the reported EMS cases in the US were women, 97% were Whites not of Hispanic origin, the median age of cases was 48 years, and few cases were under 25 years of age. In the New Mexico case–control investigation, the history of ingestion of dietary L-trytophan and 32 other suspected etiologic factors was ascertained for 11 identified cases, along with 22 neighbors of the cases of the same sex and age groups; in the Minnesota case–control study, the characteristics of 12 identified cases were compared with 12 noncases matched on sex, age, and telephone exchange. In both studies, the associations between ingestion of tryptophan-containing nutritional supplements were expressed in relation to huge odds ratios: all of the cases in both studies had ingested tryptophan-containing products; fewer than 10% of the noncases had done so. Furthermore, only one of the 85 cases reported to the CDC by early November 1989 was found to have no prior history of ingesting tryptophan-containing products.

On the basis of this evidence, because other suspected causes had been ruled out, and because no biases appeared to account for the observed EMS–tryptophan associations, the US Food and Drug Administration (FDA) asked manufacturers and suppliers for a voluntary recall of tryptophan-containing products, effective November 17, 1989—less than six weeks after the original report on the syndrome! Thereafter, the number of reported EMS cases declined sharply.

Although there were more than 1000 manufacturers of tryptophan at that time, it was discovered that six Japanese companies were responsible for essentially all of the dietary tryptophan consumed in the United States during 1989. Furthermore, adaptations of the case–control strategy narrowed the search to a single Japanese manufacturer and to retail lots of dietary tryptophan that had been produced from this manufacturer's supplies between mid-1988 and mid-1989. Whereas other manufacturers might have produced tryptophan that caused EMS, the risk of EMS was shown to be markedly greater among users of a single manufacturer's retail lines.

Securing the cooperation of the Japanese company, the epidemiologists sorted batches or lots of tryptophan in relation to whether they were found to be linked to one or more cases of EMS or to no cases of EMS. Tryptophan batches linked to the EMS cases were more likely to have involved changes in standard fermentation and purification processes. Two of several changes were implicated: less powdered activated carbon had been used in product purification, and in addition, a new strain of *Bacillus amyloliquefaciens* had been used for fermentation. Subsequent chromatographic studies were used to identify a possible contaminant, apparently the result of using the new strain of bacterium for fermentation, and this contaminant (EBT) is now regarded as a candidate for the etiologic agent in this epidemic of EMS. (However, subsequent to the end of the epidemic, randomized controlled trials of EBT exposure in Lewis rats showed that EBT can cause some but not all of the histopathological changes associated with EMS.)

Although some postepidemic cases of EMS have been reported, and it is claimed that there are other causes of EMS besides badly manufactured dietary tryptophan (Spitzer et al., 1995), the 1989–1990 epidemic subsided to negligible levels within months of the FDA regulatory action (Kilbourne, 1992). Hence, dramatic reductions in the incidence of EMS following the FDA action constitute additional important evidence that the root causes of this epidemic involved dietary tryptophan in some way.

As exemplified by the story of the eosinophilia-myalgia syndrome, there often are public health emergencies that cannot be addressed quickly without case–control methods. Because they seem to be developmentally backward and sample on the outcome, epidemiologic case–control study methods are often criticized by behavioral scientists who seem not to appreciate how this tool can be used complementarily with other research strategies. However, case–control studies have a demonstrated value in public health emergencies such as the EMS outbreak and have also helped to propel etiologic research on Alzheimer's disease and other nonemergency public health problems in a manner that has facilitated the design of subsequent prospective and longitudinal investigations. For a fraction of the cost of other investigations, the epidemiological case–control study can yield invaluable information needed to plan prospective and longitudinal studies (e.g., estimates for the effect sizes needed to determine sample size values in prospective studies, or empirical evidence on plausible but possibly unimportant confounding variables). Under these circumstances, it often is imprudent to initiate fieldwork in expensive longitudinal or prospective studies until after epidemiologic case–control research has been completed.

Three emerging lines of case–control research on mental and behavioral disturbances deserve special note because they are likely to appear with increasing frequency in the research literature of interest to clinicians in the professions of psychology and psychiatry. First, in essence, studies of discordant twin pairs are sometimes special cases of case–control research: in these instances, a twin who has been affected by a disorder is found, and subsequently there is an effort to determine whether the index twin's co-twin has also been affected. These twin studies hold constant inherited characteristics (in the case of monozygotic twins) or allow them to vary (in the case of dizygotic twins) in order to identify aspects of the environment that might modify inherited vulnerabilities or that might affect occurrence of mental and behavioral disturbances independent of inheritance. The same logic applies to studies of discordant sibling sets in general. What makes them epidemiological (if they are at all) is that there is a deliberate search beyond treated cases to assess who has become sick and who remains well, and to determine what might be causing this variation in the occurrence of sickness. What makes them examples of the case–control study design is that they sample on the outcome by starting with a sample of affected twins.

There now are many such twin studies of mental disorders and a growing number of discordant sibling studies. For many years, twin researchers were preoccupied with establishing whether inheritance had anything to do with the etiology of mental and behavioral disturbances, and the focus of attention was on a question of genes vs. environment or on estimation of the heritability of a disorder or a trait. More recently, the questions have looked beyond the more basic issue of whether inheritance matters. Instead, the hypotheses are framed in terms of which genes might account for observed inheritance of mental and behavioral disturbances or traits, and how genes and environment might combine in dynamic interplay to influence the development of these disturbances and traits. Simply by way of illustration, the twin studies of Breitner et al. (1995) provide a starting point in relation to the epidemiology of Alzheimer's disease. The work of Pickens, Svikis, McGue, Lykken, and Heston (1991) and Johnson, van den Bree, Uhl, and Pickens (1996) illustrate the case–control variety of twin study research as applied to alcohol and other drug dependence.

In a second line of case–control research, inter-related with twin studies, are recent tests of etiologic hypotheses about specific genetic polymorphisms. These experiments are especially well-served by the epidemiological case–control study design because the nature of the suspected causal associations has a temporal ordering that finesses some of the interpretive challenges sometimes faced in case–control studies. Namely, it is hard to confuse the temporal sequence that leads from genes and gene expression toward mental and behavioral disturbances, even when the underlying theory specifies complicated interactions between genes and environment. Once they have been used to establish a well-replicated linkage between specifically mapped genes or gene markers and subsequent mental or behavioral disturbances, adaptations of the case–control design can be used to test targeted hypotheses about how environmental conditions or processes might modify or interact with these genetic influences. This can be accomplished in the same way that case–control methods had to be adapted to evaluate which manufacturing processes might have changed the risk of EMS and caused that epidemic to occur.

Again, providing a starting point to understand this line of research, the first studies to investigate polymorphisms associated with the dopamine D2 receptor gene and their suspected causal linkage with severe alcoholism and cocaine dependence represent applications of the epidemiological case–control study design (Blum et al., 1990; Uhl et al., 1993). The continuing debate over these studies mainly has to do with the controversial selection of their noncase series ("controls"). Applying epidemiological strategies within the framework of family-genetic studies, and via case–control analyses of the resulting study data, Strittmatter et al. (1993) at Duke University and Henderson et al. (1995) at the Australian National University in Canberra have made advances in our understanding of polymorphisms linked to a gene for apolipoprotein epsilon-4 and their suspected causal or protective associations with Alzheimer's disease.

In a third line of research with importance for clinical psychology and psychiatry, the epidemiological case–control study design has been used with increasing frequency to analyze the evidence from large-sample epidemiological studies in a manner that holds constant shared aspects of local neighborhood environment in order to test hypotheses about how personal and individual characteristics might account for excess risk of mental and behavioral disturbances. The opportunity to use the case–control study design in this context emerges from a feature of the multistage area probability sampling plans that are used to draw these study samples. That is, these plans first sample from lists of primary sampling units (PSUs) which sometimes are as large as states or large metropolitan areas. From within the PSUs, there is another stage of probability sampling from blocks, block groups, census tracts, or neighborhoods within which there are multiple households or dwelling units; these "neighborhood"-level aggregations then are sampled. From within the sampled "neighborhoods" there is an additional stage of probability sampling of listed households or dwelling units, and then from within the households, there is a listing of all eligible study participants, and a random sample of these is taken in order to designate a participant for the study. Once sampling and recruitment has occurred, the study's assessment procedures are undertaken, either cross-sectionally or prospectively, and when mental disorders are under study, it is possible to identify those who have become cases as well as those who have not. Followed prospectively over time, and reassessed, those who have not become cases at the time of the baseline assessment are at risk of becoming incident cases.

Typically, over the span of a standard follow-up interval in these prospective studies, very few candidates become incident cases during the follow-up interval. In consequence, within any given sampled neighborhood, it is possible to compare and contrast the candidates who remain noncases with those who become incident cases. Thus, by using the conditional form of the logistic regression model or other numerical methods that treat the neighborhood-level variable as a matched characteristic, an investigator is allowed to hold constant shared aspects of living in a particular neighborhood (e.g., social disorganization, disadvantage, street-level availability of drugs, police presence). At the same time, it is possible to estimate the risk of becoming an incident case in relation to a suspected causal factor that has been measured at the personal or behavioral level.

This form of epidemiological case–control study design has been used to test a variety of hypotheses about suspected causal factors for mental disorders and behavioral disturbances:

(i) Whether an individual's use of cocaine might increase risk of panic attacks (Anthony et al., 1989), hallucinatory or delusion-like experiences (Tiens & Anthony, 1990), and obsessive-compulsive behavior (Crum & Anthony, 1993).

(ii) Whether failing to achieve social expectations for education and schooling increase the risk of developing *DSM-III* alcohol dependence or alcohol abuse in adulthood (Crum, Bucholz, Helzer, & Anthony, 1992), or the risk of developing major depression in middle age or later life (Gallo, Royall, & Anthony, 1993).

(iii) Whether jobs with high levels of physical or psychological demands and low levels of personal control and autonomy might increase risk of drug dependence in adulthood (Muntaner, Anthony, Crum, & Eaton, 1995).

Even the data from a cross-sectional survey can be organized within the framework of a case–control study and used to investigate questions about causation if an investigator will take care to gather the data carefully and analyze the results with attention to possible distorting influences. For example, one of the most informative recent studies on the issue of social causation of mental disorders was completed, in essence, as a cross-sectional study, but with case–control analyses of the resulting study data. The hypotheses to be tested in this study have long occupied the attention of epidemiologists, who were able to notice, even in mental hospital statistics, that individuals with mental disorders often seemed more socioeconomically disadvantaged than one might expect (e.g., see Faris & Dunham, 1939). A challenging question has been whether it is the disadvantaged social conditions that increase the risk of mental disorders or whether it is more that the mental disorders cause individuals to be held back or even to lose social status. For example, lifelong social disadvantage would tend to hold people back from acquiring prestigious social status (e.g., as reflected in years of schooling, advanced degrees, and educational attainment in general). Accordingly, any socially disadvantaged person who managed to achieve high social status such as a good education might be an especially healthy and adaptive person, with life chances relatively unencumbered by mental disorders. Contrariwise, any socially advantaged person who failed to achieve high social status might more likely be one whose life chances had been encumbered by mental disorders. Evaluated at each point along a gradient of social status (e.g., educational attainment), social causation might result in higher risk of mental disorders among persons with lifelong social disadvantage due to discrimination based on race or ethnicity when compared to persons without such disadvantage. Social selection might result in higher risk of mental disorders among persons with lifelong social advantage when compared to persons without such disadvantage (Dohrenwend et al., 1992).

Presenting this analysis of the selection–causation issue, Dohrenwend et al. (1992) studied Jews living in Israel, cross-classifying them in relation to their membership in groups whose African family heritage, physical appearance, and ethnicity might place them at a lifelong social disadvantage. In addition to making standardized assessments of their lifetime histories of mental disorders, the research team evaluated each study participant's social status in relation to educational attainment, and also measured other suspected influences on mental life and behavior (e.g., age, sex, marital status). In a careful analysis of occurrence of different types of mental disorders in groups cross-classified by disadvantage attributed to ethnicity and socioeconomic status, these investigators found patterns of association consistent with a more prominent influence of social selection in relation to disorders such as schizophrenia. However, the evidence was consistent with a more prominent influence of social causation in relation to disorders such as the substance use disorders. That is, the evidence on schizophrenia and related disorders indicated social selection mechanisms at work: socially disadvantaged persons with high educational attainment actually were less likely to have developed schizophrenia than socially advantaged persons with high educational attainment. In contrast, all along the gradient of educational attainment, the estimated risk of having developed substance use disorders was higher than expected for the socially disadvantaged when compared to socially advantaged persons. Case–control analyses based on the multiple logistic regression model were completed in order to adjust for other suspected distorting influences such as marital status, but the results from such analyses did not alter these inferences from the study (Dohrenwend et al., 1992). This important investigation serves as a good illustration of how a cross-sectional epidemiological study can do more than estimate how many people are affected, and where within the population the affected cases are more likely to be found. With careful reasoning, assessment, and analyses, this study shed new light on an important enduring issue about the causes of mental disorders, and identified some important new issues that can be probed more thoroughly in future prospective and longitudinal studies.

An often overlooked feature of the case–control study design is its capacity to make efficient and rapid use of study resources. In concrete terms, many prospective studies of mental disorders have involved samples of literally thousands of participants. Upon follow up, no more than a small fraction of these participants has developed into an incident case of any specific mental disorder. In the resulting analyses of all the study data, there are hundreds of noncase control participants for every incident case of mental disorder. This can represent an unnecessary and inefficient allocation of study resources in that statistical power in case–control hypothesis testing reaches near-maximum levels once the ratio of noncases to cases is 10:1 (Schlesselman, 1982).

Rather than computerizing and analyzing the study data on the thousands of participants, it would be justifiable and efficient to apply case–control reasoning, namely to identify the incident cases plus 10 noncases for comparison with each incident case, and to computerize and analyze the study data on each set of 11 participants. If useful, the noncases could be matched to the cases on possible distorting variables such as age, sex, or neighborhood, in order to hold these characteristics constant and focus attention on suspected causal factors at the personal or behavioral level.

The utility of this form of "nested" case–control analysis becomes especially clear in targeted research when the hypotheses involve laboratory work for genetic studies, biomarkers, or other bioassays, or when each participant's record includes data that must be reviewed in detail and coded. For example, in large-sample epidemiological research, it has become quite common to ask each of the thousands of participants for permission to draw a blood sample or to take a buccal swab sample of tissue, which can be frozen and archived, and from which sufficient DNA can be extracted to test for homozygosity or heterozygosity with respect to specific genes or genetic polymorphisms. Although less common, there also are epidemiological studies with large samples in which videotapes of interviews or family interaction patterns are made and archived, in an effort to characterize facial expressions of each participant, or to evaluate dynamic interactional processes such as coercive interactions between family members. Along the lines of case–control reasoning, a case–control study can be nested within these large-sample studies. When cases from these studies are compared with noncases, it is possible to obtain efficient and unbiased study estimates by drawing case material from the stored archives (e.g., the frozen blood sample, or the archived videotape), along with material on 3–10 noncases for each case, and to conduct the detailed and intensive assays only on this subset of the thousands of participants. In studies of bioassays or videotape coding of facial expression or family interaction, this might mean no more than 100–500 assays, vs. the thousands that otherwise would be required in order to obtain the assay data on all the participants. The "nested" case–control design and associated statistical issues have been described by Checkoway and Demers (1994).

As implied throughout this section, we have concentrated on the epidemiological case–control study design because its advantages for etiological research are often overlooked by investigators whose own graduate research training has led them to focus on either cross-sectional surveys, prospective and longitudinal research, or experimental design. Nonetheless, as shown in examples of cross-sectional survey research, what distinguishes an epidemiological investigation that answers questions about causation is not to be found in the investigator's choice of a specific design, but rather in the combination of research design with numerical methods used to analyze study data and to interpret the resulting evidence. In an investigation oriented toward the first two rubrics of epidemiology—quantity and location—there is no need for detailed attention to the issue of whether the study subgroups, or the cases and noncases, might be balanced with respect to potentially distorting influences. There is no reason to master or apply techniques of randomization, stratification, matching, exclusion, or statistical adjustment that can be used to bring these influences into balance. What characterizes epidemiological research on the question of causation is attention to this issue of balance, and to the appropriate use of techniques that can bring distorting influences into balance, so that we might gain a clearer view of what accounts for some individuals becoming cases while other individuals are spared.

1.12.3.4 Rubric 4: How? A Question of Mechanisms

Epidemiologists turn to the wishy-washy concept of a risk factor whenever there is a constraint on our understanding about the causes and causal mechanisms that influence the risk of ill-health or disease. For example, on occasion, the available evidence highlights locations in space or time, or within subgroups of populations, where cases are more likely to be found. Depending upon the focus on prevalence or incidence, the observed associations will be described in terms of prevalence or incidence correlates. With these circumstances, nothing more than a correlation is implied; causation clearly is not established.

As this evidence is worked up, it sometimes becomes possible to point toward characteristics, conditions, or processes that seem to have some influence on the risk of becoming a case, but uncertainty remains. There is an unwillingness to finger a specific cause or causal mechanism, but the accumulated evidence implies an association that has some causal significance, at least in theory. Under these circumstances of still limited understanding, epidemiologists would like to say that a cause or a causal factor has been identified, but instead they will pull the punch and say that they have identified a "risk factor."

The late lamented Abraham Lilienfeld, who served at Johns Hopkins as Professor and Chair of Epidemiology for many years, and later as Professor and Chair Pro Tem of the Mental Hygiene Department, introduced one of the present authors to this idea that the term "risk factor" reflects a limited understanding of causation and causal mechanisms. When asked about the origins of the term, his memory led him back to the 1950s and the Framingham Heart Study team's original observation that the risk of heart disease was found to increase with increasing age as that study's participants were followed forward in time. At the time, he said, the investigators did not wish to claim that age caused heart disease, as if heart disease were an inevitable consequence of growing old and there was nothing to do about it. At the same time, they inferred that increasing age was an important marker of age-related causal characteristics, processes, or events that might be altered in order to reduce the risk of heart disease, and they did not wish to throw away the idea that some modifiable aspect of the aging process might become an appropriate target for successful intervention and prevention. In this respect, the term risk factor represents a felicitous choice of words because, by etymology, a "factor" can be taken to represent a force that might turn out to make a difference: it is possible that it has causal significance, either as an independent cause, a contributing cause or cofactor that works conjointly with other causes, or as an inert marker that runs along with some other important causal factor in a complex causal chain.

For present purposes, the origins and etymology of risk factor are less important than the array of distinctions to be drawn between prevalence correlates and incidence correlates on the one hand, and between risk factor and "cause" or "causal mechanism" on the other. The distinction between prevalence correlate and incidence correlate refers to variation in the epidemiological indicator of disease frequency or occurrence. If the observed association involves something that is associated with being a case or having become a case, then it involves a prevalence correlate. If the observed association involves something that is associated with becoming a case, then it involves an incidence correlate. Some prevalence correlates might also prove to be incidence correlates, but many prevalence correlates are neither risk factors nor incidence correlates. Rather, they are consequences of having become a case, or they are associated with the duration of remaining a case. For example, receiving clinical services from a mental health specialist is associated with being a case of a mental disorder: active cases are more likely to receive specialty services than noncases (Marino et al., 1995). If we are to make sense of the terms incidence correlate and risk factor then attention to the issue of temporal sequencing is required: the risk factor appears first, and becoming a case comes later.

Within epidemiology, the "incubation period" or "induction period" that precedes occurrence of a case serves as a useful heuristic and reminder of the temporal sequencing between the risk-promoting characteristic, event, or process, and the later occurrence of clinically apparent signs and symptoms. The domains of clinical psychology and psychiatry hold some direct analogies for the incubation period of communicable diseases. For example, with respect to post-traumatic stress disorder (PTSD), the signs and symptoms of the disorder can be observed shortly after exposure to the traumatic event, or the induction period for the clinical syndrome might last for a long period, as observed among wartime veterans in relation to "delayed-onset" PTSD occurring months or even years after exposure to combat-related trauma. With respect to dependence upon various drugs, there are measurable induction periods from the initial "exposure opportunity" until the first use of a drug, for instance, as measured in terms of the first time a youth has a real opportunity to smoke a marijuana cigarette until the very first smoking experience (Anthony & Helzer, 1995; Van Etten, Neumark, & Anthony, 1997). Afterwards, there is a measurable induction period from the time of first drug use until the appearance of the initial drug problem (e.g., see Anthony & Petronis, 1995) or until the appearance of a fully developed clinical syndrome of drug dependence.

Within the set of "incidence correlates" we have "risk factors" and we have "causes" and "causal mechanisms." The distinction between "risk factor" and "cause" or "causal mechanism" refers to an investigator's judgment about whether the available evidence is sufficient to warrant to say more than "it seems like this might be making a difference in who becomes a case and who does not." When an investigator is willing to say no more than this, then the term "risk factor" is warranted. When the evidence is sufficient to allow investigators to make a causal inference, to point a finger at a specific characteristic, condition, event, or process, and to name it as a "cause," then there is no need to appeal to the wishy-washy concept of a "risk factor"; they might as well call it a "cause" or a "causal factor." When they have marshaled suitable evidence to allow them to point out specific sequences or linkages of states and processes that influence who becomes and remains a case and who does not, then they have clarified our understanding of at least some aspect of "causal mechanism," which remains one of the highest and most esteemed goals of the scientific enterprise of public health work.

Epidemiology's history of conducting empirical investigations that probe causal mechanisms dates back at least to the 1840s, when Peter Ludwig Panum, a Danish pathologist, studied an outbreak of measles on the Faroe Islands of the North Atlantic. Though the successive eras of bacteriology and virology had not yet begun, and no one had any idea that measles was caused by a virus, Panum recorded the details of how measles swept through the Faroe Islands, recording notes about the spread from person to person, from place to place, and over time. Coupled with his close observations and careful reasoning, what allowed Panum to shed new light on the epidemic spread of measles was the fact that this was essentially a "virgin" population, in which some 65 years had passed since the last recorded measles epidemic and which therefore had no immunity to speak of. In contrast to prior observers, who had watched measles epidemics where many population members already had acquired immunity from past infection and therefore did not become cases, Panum was able to see that more than 70% of a population would contract measles when periods as long as 65 years passed between outbreaks. He appreciated that measles was spread by person-to-person contact, but also hypothesized a mechanism that now is considered to be uncommon but not impossible: namely, measles is spread when noncases come into contact with clothing worn by a case (Panum, 1846, translated by Hatcher & Dimont, 1940).

Of course, later in the nineteenth century, the scientific contributions of Louis Pasteur and Robert Koch created a bacteriological revolution that helped to clarify the causal mechanisms by which measles and other communicable diseases might be spread, from person to person or by other means. Even so, many years passed from the time that specific infectious agents were identified until it became possible to use antibody tests in order to clarify which individuals had come into effective contact with an infective agent (sufficient to produce an antibody response), but had not developed manifestations of disease and were clinically inapparent cases. Analogously, Mendel's advances in genetics were not widely appreciated until after the chromatids had been identified as specific mechanisms of genetic inheritance late in the nineteenth century; once this aspect of inheritance was illuminated, additional work in genetics led to increasing advances (McKusick, 1996), but also to efforts at social control as reflected in the eugenics movement.

As described early in this chapter, it was before the bacteriological revolution and later discovery of filterable viruses that John Snow and other epidemiologists were able to begin to distinguish diseases whose spread was primarily person-to-person, such as smallpox, from those emanating from a common source. The prototypical common source was the infamous Broad Street well and pump in London, which John Snow identified as a source of cholera during the London epidemic of August 1854. Snow learned that nearly all of the deaths from cholera in south London had occurred among residents living within a short distance of that shallow well, and he had linked drinking of the well's water to cholera mortality as far away as Hampstead, where a former resident had bottled water from the pump conveyed to her each day. Although the handle of the pump was removed in early September, the epidemic had already started to die down, and the pump was back in operation soon after the epidemic subsided—perhaps because no one had a clear understanding of what was contaminating the well water and how it had become contaminated. This aspect of causal mechanism was clarified later by one of Snow's colleagues, who learned that the cholera epidemic had started soon after the death of a local baby from diarrheal disease. Helped by the residents, he located a cesspool less than three feet from the well, into which the baby's discharges had been thrown. Upon inspection, the cesspool and well were discovered to be constructed so poorly as to allow the cesspool's contents to flow through the well's faulty brickwork into the water supply. This prompted repair of the well, but it was not until the threat of a new cholera epidemic in 1866 that the Broad Street pump and shallow well were finally disabled (Chave, 1958).

Attempts to identify the more microscopic aspects of the causal mechanism of cholera started around 1849, when Snow first articulated his theory of waterborne infection. These efforts included early experiments in which rodents were fed fecal matter and other material from cholera victims. Nonetheless, this early "animal model" of cholera failed, and it was not until 1876 that Koch isolated *Cholera vibrio* and identified it as the specific microscopic causal agent for cholera. Of course, even after this success, aspects of the causal mechanism eluded Koch, including knowledge of what might account for some individuals being exposed to *C. vibrio* without ever developing cholera, as well as the information that oral rehydration therapy is sufficient to prevent many of the deaths from cholera once it occurs.

Several of the tools used by early epidemiologists to distinguish between common source epidemics and those spread by person-to-person contact can be traced directly back to the numerical methods of data analysis and careful study of clinical cases that had been taught to William Farr and others by Pierre Charles Alexandre Louis. One of these tools involved plotting the epidemic curve in relation to the day-to-day and hour-to-hour occurrence of new cases; that is, a frequency distribution of cases by the time of onset. As Farr discovered, diseases such as smallpox that are mainly spread by either direct or indirect person-to-person contact have an epidemic curve that grows slowly toward a peak and then gradually subsides, in accord with equations that could be replicated from epidemic period to epidemic period (Farr, 1885). In contrast, outbreaks emanating from a common vehicle or source, without any elaboration via person-to-person spread, typically rise quickly to a peak and then drop off quickly as well because there are no secondary cases following exposure to each primary case. The epidemic curve of eosinophilia-myalgia syndrome followed this kind of trace, consistent with a common source in the contaminated dietary tryptophan products which were quickly removed from the market, although not showing the more explosive rise and fall of the epidemic curve that can be seen when there is common-vehicle spread of disease after a single exposure at one point in time. Within the domain of clinical psychology and psychiatry, post-traumatic stress syndromes after a mass disaster such as an earthquake often follow this "common source" type of epidemic curve, although some of the occurrence of these syndromes is delayed in onset, which tends to extend the peak of the epidemic curve.

Epidemiologists soon learned of complications to be faced when attempting to understand causal mechanisms of disease spread by studying epidemic curves. In some instances, these complications now can be seen as manifestations of immunity and a mechanism of antibody response that had been carried over from exposure to an infective agent during a previous outbreak or epidemic. In the case of measles, for example, mechanisms of acquired immunity prevent individuals with past exposure from developing measles, and this can lead to a lower peak and more rapid decline of the epidemic curve when the between-outbreak interval was short; the susceptible individuals either were surrounded by persons with acquired immunity who were not exposing them to the virus, or were becoming cases rapidly. As evidence that the epidemic had "burned out" all the susceptibles in a community's households, the epidemic would end.

In other instances, the interpretation of epidemic curves can be complicated when the outbreak starts from a common vector, but subsequently there is person-to-person spread. The result is often an epidemic curve that rises to a peak but then falls slowly in conjunction with the elaboration of cases secondary to person-to-person spread. For example, salmonella might be introduced into an institutional setting via food (e.g., bulk eggs or poultry, unpasteurized milk products), and then spread among patients or inmates via person-to-person contact and sometimes by fomites that have come into contact with an infected person and convey the bacteria to another.

Shortly after the turn of the twentieth century, epidemiologists like Henry Value Chapin developed new tools for studying the mechanisms that account for outbreaks and epidemics of disease. Chapin recognized that it was important to take the temporal sequencing of incident cases into account when studying the spread of diseases within families or other groups of people. His observations led to specification of an epidemiological tool known as the "secondary attack rate," which has been extremely useful in epidemiological studies aimed at identifying familial and other groupwise aggregations of disease, and at investigating mechanisms that might account for the observed aggregation.

In operational terms, investigating the familial secondary attack rate requires an attempt to identify the first or "index" case that surfaces within each family, and to identify all of the remaining susceptible individuals in the family in order to determine which of them becomes a subsequent or "secondary" case (after the index case). The timing of the occurrence of the secondary cases relative to the index case and relative to other secondary cases is also studied. Among those who do not become cases, it is necessary to attempt a distinction between individuals who have a previous history of being affected by the condition (e.g., on the basis of past history of symptoms and signs) and those who have no apparent history but remain susceptible to future occurrences. Once these details have been investigated, the secondary attack rate for each family can be estimated as a conditional probability by dividing the number of cases who are secondary to the index case in a given family by the total number of susceptible persons in that family, setting aside the experience of the index case and of all persons with a past history of the condition. If necessary, the time of onset of each case can be used as a stratifying variable in order to express the probability of becoming a case in relation to units of time. Thus, the secondary attack rate is related to the conditional probability of becoming a case during some specified time span, given that an index case has occurred and that there is no prior history of the condition. Especially when an infection leaves few inapparent cases (i.e., all infected persons develop some form of clinical expression of the infection), the secondary attack rate can be especially useful. The magnitude of a secondary attack rate helps to clarify whether familial aggregation is occurring at all, and also can be used to narrow the search for specific causal mechanisms and agents of disease.

As might be surmised, years of accumulated experience with various diseases have left epidemiologists in a position to recognize the specific "signatures" of various diseases that are disclosed in their epidemic curves and in indicators such as incubation periods and secondary attack rates. These signatures allow epidemiologists to start narrowing the search for causal agents and mechanisms, even before it is possible to complete laboratory studies that might be used to isolate factors such as bacteria, viruses, and contaminants of diet or poisons. Part of the signature consists of the clinical features in the form of symptoms and signs that can be observed without aid of laboratory tests, but in addition to these features observed by studying individual patients, the signature includes what can be learned or inferred about incubation periods from the time of suspected exposure to the appearance of the first clinical manifestations, the distribution of these incubation periods as summarized for a group of cases, the shape of the epidemic curve, and the size of secondary attack rates. As graduate students of epidemiology learn in their first laboratory exercises on outbreaks of foodborne illnesses, these signatures are sometimes essential for the discovery of causal mechanisms and can help public health officials decide, for instance, whether an outbreak of gastrointestinal illness might be the result of eating egg salad vs. tuna salad. When laboratory evidence on specific agents cannot be obtained, as often is the case, the study of these clinical and epidemiological details often provides the only evidence available for determining what might be causing the outbreak, and what are the indications for reducing or preventing the occurrence of future cases.

As for the application of these principles of epidemiology to disturbances of mental life and behavior, we have already mentioned PTSD and drug dependence. In addition, casual mechanisms for outbreaks of somatoform illness and epidemic hysteria can be studied using the concepts and principles of epidemiology (e.g., E. L. Goldberg, 1973). The epidemiological concepts of exposure opportunity, effective contact, inapparent cases, incubation periods, and secondary attack rates can also be used when studying normal and abnormal responses to events such as death by suicide of a classmate or co-worker, and when studying possibly neurotoxic and psychotoxic responses to industrial accidents or wartime exposures (e.g., the destruction of nerve gases or other biological weapons during the Persian Gulf War). In these instances, it is often useful to adapt the concepts and methods of epidemiology without forcing them to apply where the adaptation makes no sense. For example, in lieu of a focus on familial aggregation when estimating secondary attack rates, it is possible to substitute a focus on a classroom, a school's student body, a military platoon, or the workers in each shift of factory work.

The application of family study research methods in epidemiology has occurred in parallel with developments in other fields of inquiry. A "genealogical random test method," criticized by Fremming in his report on the Bornholm Island study, evolved into a variety of family-genetic research designs which work from an identified index case, with a backward look toward the history of mental disorders in deceased forebears and other older relatives, and a concurrent or even prospective assessment of mental disorders in surviving relatives, spouses, and progeny (e.g., see Merikangas, Risch, & Weissman, 1994; Weissman, 1993). Studies of especially informative pedigrees, and family study designs to detect founder effects, fit into this tradition of family-oriented epidemiological research, as do the nonconcurrent and concurrent prospective designs of twin and sibling research (Risch & Zhang, 1996).

Nonconcurrent twin studies involve recruitment and assessment of the past history of mental disorders in twin pairs, both monozygotic and dizygotic, most often reared together but sometimes separated at birth or in later years, and reared apart. The resulting study data can be analyzed within the case–control framework, but it is often more sensible to analyze the data as if it had been gathered prospectively, treating each twin pair as a homogeneous risk set observed from conception or birth until the time of follow-up. When the twins have been separated at birth or later, there are particular opportunities to tease apart the independent and combined effects of genes and the shared and nonshared aspects of their environments and experiences (e.g., see Breitner et al., 1993; DiLalla, Carey, Gottesman, & Bouchard, 1996; Kaprio, Koskenvuo, & Langinvainio, 1984; Kendler, Eaves, Walters, Neale, Heath, & Kessler, 1996; Kendler, Neale, Heath, Kessler, & Eaves, 1994; Langinvainio, Kaprio, Koskenvuo, & Lonnqvist, 1984; Lyons, 1996; Lyons et al., 1996; Lyons et al., 1995; Scherrer et al., 1996).

There is an increasing number of concurrent prospective studies of twin pairs who are recruited during fetal life or soon after birth, and followed prospectively to observe their disturbances of mental life and behavior. Closely related epidemiological research designs for siblings in general have been developed, one of the most interesting of which is based on a late twentieth-century upturn in the frequency of sequential monogamy and an associated increase in prevalence of half-sibs in the general population. That is, epidemiologists now have a greatly increased opportunity to recruit families in which there has been a change in spouse or partner during the reproductive years. As described by Olsen et al. (1997):

> Like traditional square dances, modern life offers a large selection of new partners, only some of whom remain with one dance partner for more than a

limited period of time; a large number of individuals, therefore, have half-brothers or half-sisters, thus providing an opportunity to study the effect of a change in partner on reproductive outcome.

Olsen and colleagues coin the term "computerized square dance design" for epidemiological research in which computerized population registries are used to identify couples who change partners or social conditions within the reproductive time period, and to use the resulting risk sets for research on genetic and environmental contributions to adverse health outcomes. Despite possible distortions from developmentally important influences such as genomic imprinting, a variety of concurrent prospective designs with twins and half-siblings can be used to clarify genetic and environmental contributions to the occurrence of disease, and to study the interplay of these influences (Hall, 1992, Plomin, 1990). In their own application of this research design, Olsen et al. (1997) have studied congenital abnormalities, most recently focusing upon a cohort of men who fathered a child with a facial cleft defect and who had at least one additional child during the study period. As with the parallel concurrent prospective studies of twins reared together and apart, these studies of half-sibs can help tease apart genetic and environmental aspects of disease risk, with the advantage of prospective assessments from the fetal period or birth.

Notwithstanding their limitations, such as incomplete control over theoretically important distorting influences such as genomic imprinting (Hall, 1992), these "natural experiments" are among our most powerful opportunities for epidemiological research into the mechanisms underlying the risk of mental and behavioral disturbances. Harnessed with gene-mapping and biomarking tools of molecular biology and genetics, these concurrent and prospective research designs open new doors for epidemiological studies of these disorders. These research designs will help to shift the focus of attention away from the less fruitful study of the transmission of complex "inherited traits" to the more productive study of the mechanisms by which genes mapped on chromosomes are transmitted from parent to offspring, the phenotypic effects of these genes, and the influence of environment on patterns of gene expression during different developmental periods of human life (e.g., see Hulka & Margolin, 1992; Hyman & Nestler, 1996; Kosofsky, Genova, & Hyman, 1994).

In addition to these applications of family study research designs and longitudinal research designs in the study of causal mechanisms, there are some new developments in epidemiological research on inherited traits. As we seek to discover causal mechanisms that lead toward disease, including mental and behavioral disorders, it is possible to think of large-scale epidemiological studies in which cases of mental disorders and samples of noncases are asked to provide biological samples and allow bioassays to clarify how their genes might respond to specific environmental challenges.

In a direction opposite to that of the microscopic study of genes responding to environmental challenges, epidemiology is also benefiting from an increased capacity to examine how characteristics of the household, neighborhood, and community at large can have an influence on associations observed at the level of the individual organism. For example, using an adaptation of the generalized estimating equations developed by Scott Zeger and K. Y. Liang, Katz et al. (1993) have been able to study how individual-level, household-level, and village-level characteristics work together to influence rates of diarrheal disease in developing countries. The statistical methodology has been refined to allow estimation of the strength of a possibly causal association between the handwashing behavior of individuals and their risk of diarrheal disease, and then to tease apart how much of that association might be influenced by household-level characteristics (e.g., whether there is running water in the household), and how much might be influenced by village-level characteristics (e.g., whether the village practice is to fertilize crops with night soil, that is, human fecal matter). Together with related developments in hierarchical modeling (e.g., Bryk & Raudenbush, 1992), these innovations can soon become part of epidemiological research designs for inspecting causal mechanisms at various levels of organization: from the nation-state down to the community, family, and individual levels, and onward to the level of tissues and cells (e.g., see Sampson, Raudenbush, & Earls, 1997). As described by some observers (e.g., Susser & Susser, 1996a, 1996b), the results of these innovations will amount to a change in orientation for epidemiology, characterized (we think) by an increasing focus upon mechanisms that cause human suffering and the mechanisms for preventing human suffering, as distinct from "risk factors" that might or might not have causal significance.

1.12.3.5 Rubric 5: What Can Be Done? A Question of Prevention and Intervention

The theme of epidemiology in the service of prevention and intervention runs through this entire chapter and its stories of cholera and pellagra. Indeed, sustained by both predisposition and training, epidemiologists aspire to make a difference in the world by discovering or devising effective ways to prevent human disease and suffering. In epidemiology, this practical goal generally is given a value even greater than those of discovering causes of disease, causal mechanisms, or effective treatment of active cases.

The degree of commitment of epidemiologists to the goal of prevention can sometimes be measured in relation to personal risk, suffering and even death. For instance, soon after the Spanish–American War, Major Jesse W. Lazear joined Walter Reed's Yellow Fever Board in Cuba, which first sought to identify the specific cause of yellow fever by testing the idea that yellow fever was caused by the hog-cholera *Bacillus icteroides*. Unable to find evidence in support of the hypothesis about bacteria or any other specific causal agent, the Board decided that it might be better to investigate alternative mechanisms by which yellow fever was spread.

They turned to a "formite" hypothesis, according to which yellow fever was spread by contact with clothing and bedding soiled by the excrement of yellow fever sufferers, and also to physician Carlos Juan Finlay's theory about spread by mosquitoes. As the research team leader responsible for breeding and tending the experimental mosquitoes, Lazear came to be bitten by a mosquito, contracted yellow fever, and died. Others tested the fomite theory by sleeping on bed sheets and pillows that had been fouled by the blood and vomit of actively suffering yellow fever victims: none developed the disease. More than 20 volunteers participated in life-risking experiments on yellow fever in this manner before the Board accumulated sufficient evidence to confirm Finlay's theory, and to guide Colonel William Crawford Gorgas toward an effective means of prevention. Namely, Gorgas saw that the mosquitoes generally remained near humans for a nutritional blood supply, and sought breeding places, to lay eggs, in nearby standing water. He organized teams to drain the standing water near homes; the density of mosquitoes near human habitations became much reduced, and the incidence of yellow fever declined as cases were prevented. (While yellow fever was being prevented by Gorgas' methods, 30 years passed before Max Theiler identified a specific viral cause for yellow fever, from which an effective yellow fever vaccine was produced.)

The commitment of epidemiologists to prevention can also be seen in recent episodes of the history of smallpox, a communicable disease characterized by rash and fever, and with a case fatality ratio of up to 40%. Smallpox was recognized by the ancients, but is with us no more. The virus that causes smallpox was eradicated from all but research laboratories and eliminated as a scourge of humankind in 1977. As background, partially effective means of inoculation against smallpox were introduced to Great Britain before 1750; Edward Jenner discovered a more effective vaccination method in 1796. The specific orthopoxvirus was identified by the middle of the twentieth century, while the vaccine and methods for its cryopreservation required refinement even beyond that date. However, the last smallpox case in the United States was recorded in 1949. In more tropical and developing countries, smallpox epidemics continued for an additional two decades, despite widespread mass vaccination with a refined and stable vaccine. Finally, epidemiologists working through the World Health Organization (WHO) devised a two-pronged strategy, described by epidemiologist D. A. Henderson (1980) as follows:

> (1) a 2- to 3-year program of systematic vaccination to reach at least 80 percent of the population and so diminish transmission, and (2) the development during this period of a reporting network and surveillance system that would permit prompt detection of cases so that outbreaks could be contained and transmission finally disrupted.

This global epidemiological surveillance–containment strategy was initiated in 1967. By 1976, WHO had reduced the number of smallpox-affected countries to Ethiopia and Somalia. Mobilizing for a final phase of the surveillance–containment method, WHO teams contained the last smallpox epidemic in early 1977. The last known case developed smallpox rash on October 26, 1977 in Somalia (D. A. Henderson, 1980).

Complementing its attention to smallpox and other diseases of infectious origin, more than 30 years ago the American Public Health Association recognized discoveries and accomplishments in the prevention and control of mental disorders by publishing a handbook analogous to its compendium on communicable diseases (American Public Health Association [APHA], 1962, 1994). Entitled *Mental disorders: A guide to prevention and control*, this APHA handbook continues to serve as a remarkable statement of what we already know about how to prevent mental disorders. Now an out-of-print classic, it is required reading for any serious prevention researcher in the field of alcohol, drug, or mental disorders, or epidemiologist interested in these disorders.

The introductory chapters of the APHA handbook convey the best ideas from one of the most challenging prevention scientists of the twentieth century, Ernest M. Gruenberg, who has already been mentioned in this chapter for his key article *The failures of success* (Gruenberg, 1977). In the handbook, Gruenberg and his APHA colleagues first focus attention on the prevention of a "social breakdown syndrome," which is a breakdown of social functioning concurrent with anger, hostility and sometimes social withdrawal, often observed when mental disorders are accompanied by changes in affected individuals' social relationships after the onset of mental disorders—and often aggravated when these patients remain institutionalized for overlong periods of time.

Tackling the problem of how to prevent this type of syndrome, Gruenberg and his colleagues sought to address what is still a pressing question for primary care practitioners and specialists: namely, how can we reorganize the delivery of health and mental health services so as to reduce the suffering and secondary disabilities faced by victims of mental disorders? Seeking to answer this question, and to reduce the secondary adverse consequences of exposure to the "total institution" of a mental hospital, Gruenberg and his colleagues devised a population-oriented and patient-centered system of linkages between family, primary care provider, mental hospital, and other specialty sector resources. Focusing upon the social breakdown syndrome, his team used their system in an effort to prevent these unintended adverse consequences of hospitalization. It is important to acknowledge that Gruenberg's own attempt to evaluate his system of organizing mental health services was promising but inconclusive (Gruenberg, 1970). Nonetheless, the ideas retain pertinence, especially in the current era of clinical care for mentally disordered patients, during which concern about the mental disorders and secondary disabilities of sometimes angry, hostile, and sometimes socially withdrawn homeless persons has replaced a previous concern about the adverse consequences of excessively prolonged institutionalization (Gruenberg, 1977).

After introducing concepts and tasks for prevention in relation to the social breakdown syndrome, the APHA handbook turns to the task of cataloging all of the now recognized mental disorders and the techniques to prevent these disorders, much as the APHA communicable diseases handbook does for infectious diseases. With effectively compiled tables and narrative text, the handbook reviews the causes of specific mental and behavioral disturbances as they were known in 1962, as well as the means of prevention when the causes and causal mechanisms remained unclear. Thus, for example, there are detailed tables of physical agents that can give rise to mental and behavioral disturbances, including the intoxication and delirium secondary to preventable exposures to chemicals such as mercury, carbon disulfide, and cocaine, as well as description of the types of effects and the antidotes, if known. There are also tables and text covering the infections that can produce mental disorders, such as St. Louis encephalitis, rubella (measles), and toxoplasmosis, which remain active as causes of mental disorders to this day. We are not aware of any more comprehensive catalog of the currently known techniques for preventing mental and behavioral disturbances.

A more recent attempt to summarize the means of prevention of mental and behavioral disturbances was published by the United States Institute of Medicine (IOM) in 1994, but this effort has much less focus upon the more firmly established means of prevention, for instance, increasing rubella immunization to reduce measles-related neuropsychiatric diseases (IOM report; Mrazek & Haggerty, 1994). Instead, the focus of the report is upon conditions for which the means of prevention are still quite uncertain (e.g., schizophrenia, major depression, conduct disorders).

The IOM report is particularly valuable for its description of a cyclical prevention research process by which emerging evidence on the suspected risk factors and causal mechanisms is translated into prevention program evaluations. To the present authors' knowledge, this is the most thorough review and summary of evidence from recent experimental field trials concerned with the prevention or reduced occurrence of mental and behavioral disturbances or with the promotion of mental health.

One useful illustration in the IOM report involves prevention of conduct problems or misbehavior, such as breaking rules about underage tobacco use and drinking. Some prevention scientists have approached these problems by means of comprehensive community-based programming efforts, and have conducted evaluations by comparing communities with and without a broad profile of political, school, and family-based interventions (e.g., Pentz et al., 1989; Perry et al., 1996). Others have oriented their prevention efforts to more focused health education initiatives within schools and classrooms, relying more heavily upon what can be accomplished within classrooms than upon a more broadly based approach (e.g., Botvin, Baker, Filazzola, & Botvin, 1990; Ellickson & Bell, 1990).

A different point of departure can be seen in the work of research teams formed by J. David Hawkins and by Sheppard G. Kellam. Orienting themselves with regard to a long tradition of clinical observations about how drug users seen in clinical practice seemed to have an excess of childhood conduct problems, Hawkins and Kellam also noted Lee Robins' observations about the continuity of childhood conduct problems in relation to adult antisocial behavior. This was exemplified in a classic nonconcurrent prospective study over a period of more than 25 years, during which childhood rule-breaking and conduct problems were found to account for adult drug problems (Robins, 1966). Kellam and his Woodlawn, Chicago research team (Kellam, Brown, Rubin, & Ensminger, 1983) extended this line of research in their own prospective study of youths, who were first recruited and measured at school entry in 1966 and then followed up at age 16–17 in the mid-1970s. The Woodlawn research group found that first-graders who had been rated by teachers as rule breakers were about three times more likely to have become frequent tobacco, alcohol, and marijuana users by their teen years. The longitudinal association was particularly strong among boys, less so among girls.

Both Hawkins and Kellam were able to perceive that inferences about the causal significance of these observed associations depended upon completion of comparative trials. That is, the available evidence was not sufficient to sort out competing explanations for the observed data. On the one hand, it might be said that prevention of early misbehavior could reduce the risk of later drug involvement. On the other hand, it could be said that the apparent predictive linkage from early misbehavior to later drug involvement was no more than a spurious artifact of some unmeasured or unknown causal determinant of both early conduct problems and later drug use. The evidence to unravel this complex research issue would remain incomplete without preventive trials to study the impact of early childhood interventions. That is, within the framework of a preventive trial, there could be an attempt to alter early childhood misbehavior. If the early intervention proved to be effective, and if there is a causal link between early misbehavior and later drug involvement, then the children exposed to the intervention program should be less likely to initiate drug use.

The trials to test these causal and preventive hypotheses were initiated in the 1980s, with recruitment of children as they entered early elementary school. Early results are promising, both in the Seattle trial organized by Hawkins' research group and in the Baltimore trial organized by Kellam's research group (Hawkins, Von Cleve, & Catalano, 1991; Kellam & Anthony, in press; Kellam, Rebok, Ialongo, & Mayer, 1994; Kellam et al., 1991; O'Donnell, Hawkins, Catalano, Abbott, & Day, 1995). Nonetheless, even if these trials end with inconclusive results, they clarify a new direction for prevention research in relation to mental and behavioral disturbances. Namely, analogous to the communicable disease epidemiologists who sought either to identify specific causes or to identify causal mechanisms, we have psychologically and psychiatrically oriented epidemiologists who seek to identify specific causes or causal mechanisms for mental and behavioral disturbances. As these investigations yield evidence of suspected associations of causal significance, the observed findings can be translated into research questions for prevention researchers, who must take the next step of using experimental evaluation methods to organize field trials that can probe more deeply into the suspected causal significance of the observed associations (Mrazek & Haggerty, 1994).

Several additional preventive studies conducted by teams led by or including epidemiologists deserve mention in this chapter. One of the most interesting of these involved an analysis of data that was gathered over years for the National Institute on Drug Abuse's *Monitoring the future* survey of high school seniors. Literally thousands of these high school seniors have been surveyed each year since 1976 in order to produce a trace of epidemiological trends in the prevalence of drug use. O'Malley and Wagenaar (1991) took advantage of these survey data to evaluate the impact of state-level restrictions on underage alcohol consumption and the occurrence of fatal car crashes. This evaluation of state-level policy making was possible because over a relatively short span of years the high school senior survey covered multiple states with different drinking regulations, some allowing legal alcohol consumption at age 18, and some prohibiting it until age 21. This is, of course, an important public health and public policy question, and one that would be difficult to test using an experimental design at the state level or at the local level. Notwithstanding some difficulties in study implementation and interpretation (e.g., relating to underage drinkers' capacities to cross state lines in order to become legal drinkers), the authors' analyses provided confirmatory evidence of modest but tangible beneficial effects of state laws to postpone legal drinking until age 21. Exploratory subanalyses indicated that the beneficial effects might be conveyed by factors such as reducing the amount of time that underage drinkers spend in bars and taverns.

Another pair of epidemiological studies investigated how handgun regulations might be influencing the rates of homicides and suicides in North America, a topic of considerable public health and political interest. These investigations were made possible by making a direct comparison between the homicide and suicide rates of Seattle, Washington, and Vancouver, British Columbia. These two metropolitan areas of the Pacific Northwest are comparable in many respects, but they have substantially different handgun regulations. In Seattle, the handgun regulations are more lax, whereas in Vancouver, they are quite restrictive. Studying homicide rates, and making comparisons via analyses that were able to adjust for some of the apparent differences between Seattle and Vancouver, the research team found evidence of substantially lower homicide rates in Vancouver, and drew the inference that these lower rates could be attributed to the greater restrictions on handguns. Though an important epidemiological investigation, the study lacked a capacity to hold constant some of the potentially distorting influences on homicide rates, such as variation in the street drug trade in each city, Seattle's more prominent contrasts between wealthy and poor neighborhoods in close proximity, and the existence of a more comprehensive social "safety net" in Vancouver, which might reduce feelings of frustration and despair, with a resulting impact on aggressive behavior (Sloan et al., 1988).

The investigation on the possible preventive impact of handgun regulations in relation to suicide mortality rates, in general, was negative. Using the same methodology used in the homicide study, the research team compared Seattle and Vancouver, and found insufficient evidence to support the claim that more restrictive handgun laws can help reduce suicide death rates. Nonetheless, in exploratory subgroup analyses, the research team found some evidence that certain subgroups of the population might have lower suicide mortality rates due to restrictive handgun legislation. Namely, there was a tendency for the suicide rates to be lower for Native American young people living in Vancouver when compared to counterparts in Seattle (Sloan et al., 1990). As noted with respect to the study of homicides, interpretation of these comparative data is complicated by some between-city imbalances in possibly distorting variables, but this takes nothing away from an important epidemiological study of a significant issue for prevention.

One of the recent innovations that can help in the process of translating new scientific information on causal mechanisms into prevention efforts is the Cochrane Collaboration, which is an international network of individuals and groups who prepare, maintain, and disseminate systematic reviews of the effects of health care (Bero & Rennie, 1992; Sheldon & Chalmers, 1994) Clive Adams in Great Britain serves as first editor of a Cochrane Collaborative Review Group concerned with mental health, with a focus on therapeutic interventions. The development of a collaborative review group concerned with prevention of alcohol, drug, and mental disorders was the subject of planning meetings started by C. Hendricks Brown, Pat Mrazek, and James C. Anthony in 1996. Publications based on the work on this review group should begin to appear in 1998–2000.

There are also some distinctively epidemiological research designs that have been brought into service for the evaluation of prevention programs, but which have not yet been used in the domains of psychology and psychiatry. For example, Comstock (1994) has described how the epidemiologic case–control design has been harnessed to evaluate the effectiveness of new vaccines.

There are some clear advantages to using case–control methods to evaluate prevention efforts, as well as some unusual challenges that can block proper inference. Nonetheless, the case–control study method may be especially helpful in understanding variation in response to preventive or therapeutic efforts. Unless a program evaluation has been pre-planned for measurement of all pertinent response-affecting variables, there might prove to be no alternative to case–control methods when we seek to account for why some individuals responded to the intervention while others did not.

1.12.4 CONCLUSION

Epidemiological research continues to make exciting discoveries about the prevention of human disease and suffering, not only in relation to the long-time scourges for epidemiologists such as cholera, for which a new and more effective vaccine has just been developed, but also in relation to mental and behavioral disturbances (Mrazek & Haggerty, 1994). Nonetheless, there still remain important areas for new collaborations between epidemiologists, clinical psychologists, and psychiatrists. Future progress depends upon a narrowing of the gap between epidemiological psychology and psychiatry on one side and psychiatric or psychological epidemiology on the other .

As described in earlier sections of this chapter, clinical psychologists and psychiatrists have already harnessed tools of epidemiology when studying causal hypotheses and program evaluation issues. The works of Dohrenwend et al. (1992) and Kellam already have been cited (Kellam et al., 1983), but there are many other clinicians who have worked in this arena. The cumulative epidemiological studies undertaken by Michael Rutter's research group in Great Britain serve as an especially important example of how to build a research program to answer questions of central importance in public mental health (e.g., see Rutter, 1989; Rutter, Tizard, Yule, Graham, & Whitmore 1976).

In the United States, Lloyd Johnston's and Jerrold Bachman's *Monitoring the future* surveys, published each year since the mid-1970s, serve as an annual reminder of the power of epidemiological methods in answering questions such as "How many high school seniors are taking drugs?" The field studies of Bentler, Newcomb, Brook, Muthen, and others have drawn heavily upon epidemiological methods and strategies while remaining psychological in focus (e.g., see Bentler & Stein, 1992; Brook & Newcomb, 1995; Muthen, 1995, 1996; Newcomb & Bentler, 1988). These investigations represent important progress at the interface of epidemiology, psychology, and psychometrics. Nonetheless, there still are some areas of differences between epidemiology and the clinical sciences of psychology and psychiatry that can be addressed in greater detail.

Epidemiology, originating with the study of a discrete categorical event such as death or frank disease, tends to emphasize the study of manifest indicators of human suffering. In addition, epidemiology's strong links to clinical medicine reinforce this tradition. In terms of measurement, this often means a focus on discrete categories and transitions between categories. With regard to the study of causal hypotheses, this emphasis often implies a reluctance to depend heavily upon statistical modeling when basic associations cannot be perceived without the assumptions of these models. From psychology, we often have the opposite tendencies, with a much heavier focus upon latent variables, and with considerable reliance upon structural equation modeling, which is sometimes so complicated that the basic relationships cannot be perceived except from within the framework of these models. As epidemiologists collaborate with clinical psychologists and psychiatrists, there is a need for greater dialog and truly interdisciplinary interaction on important matters such as these. Through dialog and interaction, we can hope for continuing progress under each of the rubrics of epidemiology as they relate to the disciplines of clinical psychology and psychiatry.

1.12.5 REFERENCES

American Public Health Association (1962). *Mental disorders: A guide to prevention and control.* Washington, DC: Author.

American Public Health Association (1994). *Control of communicable diseases in man.* Washington, DC: Author.

Anthony, J. C., & Aboraya, A. (1992). The epidemiology of selected mental disorders in later life. In J. E. Birren, R. B. Sloane, & G. D. Cohen (Eds.), *Handbook of mental health and aging* (2nd ed., pp. 27–73). Orlando, FL: Academic Press.

Anthony, J. C., Eaton, W. W., & Henderson, A. S. (1995). Psychiatric epidemiology. *Epidemiologic Reviews, 17*(1), 1–8.

Anthony, J. C., Folstein, M. F., Romanoski, A. J., et al. (1985). Comparison of the lay Diagnostic Interview Schedule and a standardized psychiatric diagnosis: Experience in eastern Baltimore. *Archives of General Psychiatry, 42,* 667–675.

Anthony, J. C., & Helzer, J. E. (1995). Epidemiology of drug dependence. In M. T. Tsuang, M. Tohen, & G. E. P. Zahner (Eds.), *Textbook in psychiatric epidemiology* (pp. 361–406). New York: Wiley-Liss.

Anthony J. C., & Petronis, K. R. (1991). Suspected risk factors for depression among adults 18–44 years old. *Epidemiology, 2,* 123–132.

Anthony, J. C., & Petronis, K. R. (1995). Early-onset drug use and risk of later drug problems. *Drug and Alcohol Dependency, 40,* 9–15.

Anthony, J. C., Tien, A. Y., & Petronis, K. R. (1989). Epidemiologic evidence on cocaine use and panic attacks. *American Journal of Epidemiology, 129,* 543–549.

Anthony, J. C., Warner, L. A., & Kessler, R. C. (1994). Comparative epidemiology of dependence on tobacco, alcohol, controlled substances, and inhalants: Basic findings from the National Comorbidity Survey. *Experimental and Clinical Psychopharmacology, 2*(3), 244–268.

Bentler, P. M., & Stein, J. A. (1992). Structural equation models in medical research. *Statistical Methods in Medical Research, 1*(2), 159–181.

Bero, L., & Rennie, D. (1992). The Cochrane Collaboration. Preparing, maintaining, and disseminating systematic reviews of the effects of health care. *Journal of the American Medical Association, 274*(24), 1935–1938.

Bland, R. C. (1988). Psychiatric epidemiology. *Canadian Journal of Psychiatry, 33*(7), 618–625.

Blum, K., Noble, E. P., Sheridan, P. J., Montgomery, A., Ritchie, T., Jagadeeswaran, P., Nogami, H., Briggs, A. H., & Cohn, J. B. (1990). Allelic association of human dopamine D2 receptor gene in alcoholism. *Journal of the American Medical Association, 263*(15), 2055–2060.

Borges, G., Anthony, J. C., & Garrison, C. Z. (1995). Methodological issues relevant to epidemiologic investigations of suicidal behaviors of adolescents. *Epidemiologic Reviews, 17*(1), 228–239.

Botvin, G. J., Baker, E., Filazzola, A. D., & Botvin, E. M. (1990). A cognitive-behavioral approach to substance abuse prevention: one-year follow-up. *Addictive Behavior, 15*(1), 47–63.

Breitner, J. C., Gatz, M., Bergem, A. L., Christian, J. C., Mortimer, J. A., McClearn, G. E., Heston, L. L., Welsh, K. A., Anthony, J. C., Folstein, M. F. et al. (1993). Use of twin cohorts for research in Alzheimer's disease. *Neurology, 43*(2), 261–267.

Breitner, J. C., Welsh, K. A., Gau, B. A., McDonald, W. M., Steffens, D. C., Saunders, A. M., Magruder, K. M., Helms, M. J., Plassman, B. L., Folstein, M. F., et al. (1995). Alzheimer's disease in the National Academy of Sciences–National Research Council Registry of Aging Twin Veterans. III. Detection of cases, longitudinal results, and observations on twin concordance. *Archives of Neurology, 52*(8), 763–771.

Breslow, N. E., & Day, N. E. (1980). Statistical methods in cancer research. Vol. I—The analysis of case–control studies. (IARC Scientific Publications No. 32). Lyon, France: International Agency for Research on Cancer.

Breteler, M. M., Claus, J. J., van Duijn, C. M., Launer, L. J., & Hofman, A. (1992). Epidemiology of Alzheimer's disease. *Epidemiologic Review, 14,* 59–82.

Brook, J. S., & Newcomb, M. D. (1995). Childhood aggression and unconventionality: Impact on later academic achievement, drug use, and workforce involvement. *Journal of Genetic Psychology, 156*(4), 393–410.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models.* Newbury Park, CA: Sage.

Centers for Disease Control (1981a) Pneumocystis pneumonia—Los Angeles. *Morbidity and Mortality Weekly Report, 30,* 250–252.

Centers for Disease Control (1981b). Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men—New York City and California. *Morbidity and Mortality Weekly Report, 30,* 305–308.

Centers for Disease Control (1981c). Follow-up on Kaposi's sarcoma and Pneumocystis pneumonia. *Morbidity and Mortality Weekly Report, 30,* 409–410.

Chave, S. P. W. (1958). John Snow, the Broad Street pump and after. *The Medical Officer, 99,* 347–349.

Checkoway, H., & Demers, P. A. (1994). Occupational case–control studies. *Epidemiologic Review, 16*(1), 151–162.

Comstock, G. W. (1994). Evaluating vaccine effectiveness and vaccine efficacy by means of case–control studies. *Epidemiologic Review, 16*(1), 77–89.

Cooper, B. (1993). Single spies and battalions: The clinical epidemiology of mental disorders. *Psychological Medicine, 23*(4), 891–907.

Cooper, B., & Morgan, H. G. (1973). *Epidemiological psychiatry.* Springfield, IL: Thomas.

Crum, R. M., & Anthony, J. C. (1993). Cocaine use and other suspected risk factors for obsessive-compulsive disorder: A prospective study with data from the Epidemiologic Catchment Area surveys. *Drug and Alcohol Dependency, 31,* 281–295.

Crum, R. M., Bucholz, K. K., Helzer, J. E., & Anthony, J. C. (1992). The risk of alcohol abuse and dependence in adulthood: The association with educational level. *American Journal of Epidemiology, 135,* 989–999.

Deutsch, A. (1944). The history of mental hygiene. In *American Psychiatric Association: One hundred years of American psychiatry.* New York: Columbia University Press.

DiLalla, D. L., Carey, G., Gottesman, I. I., & Bouchard, T. J., Jr. (1996). Heritability of MMPI personality indicators of psychopathology in twins reared apart. *Journal of Abnormal Psychology, 105*(4), 491–499.

Dohrenwend, B. P., & Dohrenwend, B. S. (1974). Social and cultural influences on psychopathology. *Annual Review of Psychology, 25,* 417–452.

Dohrenwend, B. P., Levav, I., Shrout, P. E., Schwartz, S., Naveh, G., Link, B. G., Skodol, A. E., & Stueve, A. (1992). Socioeconomic status and psychiatric disorders: The causation–selection issue. *Science, 255,* 946–952.

Durkheim, E. (1897). *Le suicide: Etude de sociologie.* Paris: F. Alcan.

Eaton, W. W., Kramer, M., Anthony, J. C. et al. (1989). The incidence of specific DIS/*DSM-III* mental disorders: data from the NIMH Epidemiologic Catchment area Program. *Acta Psychiatrica Scandinavica, 79,* 163–178.

Ellickson, P. L., & Bell, R. M. (1990). Drug prevention in junior high: a multi-site longitudinal test. *Science, 247* (4948), 1299–1305.

Faris, R. E. L., & Dunham, H. W. (1939). *Mental disorders in urban areas.* Chicago: University of Chicago Press.

Farr, W. (1885). *Vital statistics: A memorial volume of selections from the reports and writings of William Farr.* N. A. Humphreys (Ed.), For the Sanitary Institute of Great Britain. London: The Sanitary Institute of Great Britain.

Fremming, K. H. (1951). *The expectation of mental infirmity in a sample of the Danish population. Occasional papers on eugenics, no. 7.* London: Cassell.

Frost, W. H. (1927). *Epidemiology.* Nelson Loose-Leaf System. Public Health-Preventive Medicine. Vol. 2, Chap. 7, pp. 163–190. New York: Thomas Nelson. Also available in K. F. Maxcy (ed.) (1941). *Papers of Wade Hampton Frost, MD: Contribution to epidemiological method* (pp. 439–542) New York: The Commonwealth Fund.

Gallo, J. J., Marino, S., Ford, D., & Anthony, J. C. (1995). Filters on the pathway to mental health care, II. Sociodemographic factors. *Psychological Medicine, 25*(6), 1149–1160.

Gallo, J. J., Royall, D. R., & Anthony, J. C. (1993). Risk factors for the onset of depression in middle age and later life. *Journal of the Society of Psychiatry and Psychiatric Epidemiology*. *28,* 101–108.

Goedert, J. J., Neuland, C. Y., Wallen, W. C., Greene, M. H., Mann, D. L., Murray, C., Strong, D. M., Fraumeni, J. F., Jr., & Blattner, W. A. (1982). Amyl nitrite may alter T lymphocytes in homosexual men. *Lancet 1* (8269), 412–416.

Goldberg, D. P., & Huxley, P. (1980). *Mental illness in the community.* London: Tavistock.

Goldberg, D., & Huxley, P. (1992). *Common mental disorders: A bio-social model.* London: Tavistock/Routledge.

Goldberg, E. L. (1973). Crowd hysteria in a junior high school. *Journal of School Health, 43*(6), 362–366.

Goldberger, J. (1914a). The etiology of pellagra: The significance of certain epidemiological observations with respect thereto. *Public Health Reports, 29,* 1683–1686.

Goldberger, J. (1914b). The cause and prevention of pellagra. *Public Health Reports, 29,* 2354–2357.

Goldberger, J. (1964). *Goldberger on pellagra.* Terris, M. (ed.). Baton Rouge, LA: Louisiana State University Press.

Goldberger, J., Waring, C. H., & Willets, D. G. (1915). The prevention of pellagra. A test of diet among institutional imates. *Public Health Reports, 30*(43), 3117–3131.

Gordis, L. (1996). *Epidemiology.* Philadelphia: Saunders.

Greenwood, M. (1935). Epidemics and crowd diseases: An introduction to the study of epidemiology. New York: Macmillan.

Gruenberg, E. M. (1970). Mental health services can be organized to prevent chronic disability. *Community Mental Health Journal, 6*(6), 431–436.

Gruenberg, E. M. (1977). The failures of success. *Milbank Mem Fund Q Health Soc. 55,* 3–24.

Hall, J. G. (1992). Genomic imprinting and its clinical implications. *New England Journal of Medicine, 326*(12), 827–829.

Hawkins, J. D., Von Cleve, E., & Catalano, R. F., Jr. (1991). Reducing early childhood aggression: results of a primary prevention program. *Journal of the American Academy of Child and Adolescent Psychiatry, 30*(2), 208–217.

Henderson, A. S., Easteal, S., Jorm, A. F., Mackinnon, A. J., Korten, A. E., Christensen, H., Croft, L., & Jacomb, P. A. (1995). Apolipoprotein E allele epsilon 4, dementia, and cognitive decline in a population sample. *Lancet, 346*(8987), 1387–1390.

Henderson, D. A. (1980). Diseases transmitted from person to person. The eradication of smallpox. In J. M. Last (Ed.), *Public health and preventive medicine* (11th ed., pp. 95–110). New York: Appleton-Century-Crofts.

Holst (1852). On the statistics of the insane, blind, deaf and dumb, and lepers of Norway. (Trans. Massey, A. S. O.). *Royal Statistical Society of London Journal, 15,* 250–256.

Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression.* New York: Wiley.

Hulka, B. S., & Margolin, B. H. (1992). Methodological issues in epidemiologic studies using biologic markers.

American Journal of Epidemiology, 135(2), 200–209.

Hyman, S. E., & Nestler, E. J. (1996). Initiation and adaptation: a paradigm for understanding psychotropic drug action. *American Journal of Psychiatry, 153*(2), 151–162.

Jablensky, A. (1995). Schizophrenia: Recent epidemiologic issues. *Epidemiologic Reviews, 17*(1), 10–20.

Jarvis, E. (1843a). Statistics of insanity in the United States. *Boston Medical and Surgical Journal, 27,* 116–121.

Jarvis, E. (1843b). Statistics of insanity in the United States (Part II). *Boston Medical and Surgical Journal, 27,* 281–282.

Jarvis, E. P. (1865–66). Influence of distance from and nearness to an insane hospital on its use by the people. *American Journal of Insanity, 22.*

Johnson, E. O., van den Bree, M. B., Uhl, G. R., & Pickens, R. W. (1996). Indicators of genetic and environmental influences in drug abusing individuals. *Drug and Alcohol Dependency, 41*(1), 17–23.

Kaprio, J., Koskenvuo, M., & Langinvainio, H. (1984). Finnish twins reared apart. IV: Smoking and drinking habits. A preliminary analysis of the effect of heredity and environment. *Acta Genet Med Gemellol (Roma), 33*(3), 425–433.

Katz, J., Carey, V. J., Zeger, S. L., et al. (1993). Estimation of design effects and diarrhea clustering within households and villages. *American Journal of Epidemiology, 138*(11), 994–1006.

Kellam, S. G., & Anthony, J. C. (in press). Targeting early antecedents to prevent tobacco smoking: Findings from an epidemiologically-based randomized field trial.

Kellam, S. G., Brown, C. H., Rubin, B. R., & Ensminger, M. E. (1983). Paths leading to teenage psychiatric symptoms and substance use: Developmental epidemiologic studies in Woodlawn. In S. B. Guze, F. J. Earls, & J. E. Barrett (Eds.), *Childhood psychopathology and development* (pp. 17–51). New York: Raven Press.

Kellam, S. G. Rebok, G. W., Ialongo, N., & Mayer, L. S. (1994). The course and malleability of aggressive behavior from early first grade into middle school: results of a developmental epidemiologically-based preventive trial. *Journal of Child Psychology and Psychiatry, 35*(2), 259–291.

Kellam, S. G., Werthamer-Larsson, L., Dolan, L. J., Brown, C. H., Mayer, L. S., Rebok, G. W., Anthony, J. C., Laudolff, J., & Edelsohn, G. (1991). Developmental epidemiologically based preventive trials: baseline modeling of early target behaviors and depressive symptoms. *American Journal of Community Psychology, 19*(4), 563–584.

Kendler, K. S., Eaves, L. J., Walters, E. E., Neale, M. C., Heath, A. C., & Kessler, R. C. (1996). The identification and validation of distinct depressive syndromes in a population-based sample of female twins. *Archives of General Psychiatry, 53*(5), 391–399.

Kendler, K. S., Neale, M. C., Heath, A. C., Kessler, R. C., & Eaves, L. J. (1994). A twin-family study of alcoholism in women. *American Journal of Psychiatry, 151*(5), 707–715.

Kessler, R. C., McGonagle, K. A., Shanyang, Z., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H.-U., & Kendler, K. S. (1994). Lifetime and 12-month prevalence of *DSM-III-R* psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Archives of General Psychiatry, 51,* 8–19.

Kilbourne, E. M. (1992). Eosinophilia–myalgia syndrome: Coming to grips with a new illness. *Epidemiologic Reviews, 14,* 16–36.

Kosofsky, B. E., Genova, L. M., & Hyman, S. E. (1994). Postnatal age defines specificity of immediate early gene induction by cocaine in developing rat brain. *Journal of Comparative Neurology, 351*(1), 27–40.

Kramer, M. (1983). The increasing prevalence of mental disorders: a pandemic threat. *Psychiatric Quarterly, 55*(2–3), 115–143.

Kramer, M., Goldstein, H., Israel, R. H., & Johnson, N. A. (1955). *An historical study of the disposition of first admissions to a state mental hospital. Experience of the Warren State Hospital during the period 1916–1950.* Public Health Service Publication No. 445, Public Health Monograph No. 32. Washington, DC: Government Printing Office.

Langinvainio, H., Kaprio, J., Koskenvuo, M., & Lonnqvist, J. (1984). Finnish twins reared apart. III: Personality factors. *Acta Genet Med Gemellol (Roma), 22*(2), 259–264.

Lemkau, P. Tietze, C., & Cooper, M. (1942). A survey of statistical studies on the prevalence and incidence of mental disorder in sample populations. *Public Health Reports, 58,* 1909–1921.

Levav, I., Lima, B. R., Somoza, L. M., Kramer, M., & Gonzales, R. (1989). Salud mental para todos en America Latina y el Caribe. Bases epidemiologicas para la accion. *Boletin de la Oficina Sanitaria Panamericana, 107,* 196–219.

Levin, M. L. (1950). Discussion of a paper entitled "Evaluation of this material." In *Epidemiology of mental disorder: papers presented at a round table at the 1949 Annual Conference of the Milbank Memorial Fund.* November 16–17, 1949, New York, (p. 96). New York: Milbank Memorial Fund.

Lilienfeld, A. M. (1980). *Foundations of epidemiology*. New York: Oxford University Press.

Lilienfeld, A. M., & Lilienfeld, D. E. (1979). A century of case–control studies: Progress? *Journal of Chronic Disease, 32*(1–2), 5–13.

Lilienfeld, A. M., Pasamanick, B., & Rogers, M. (1955). Relationship between pregnancy experience and the development of certain neuropsychiatric disorders in childhood. *American Journal of Public Health, 45,* 637–641.

Lilienfeld, D. E., & Stolley, P. D. (1994). *Foundations of epidemiology* (3rd ed.). New York: Oxford University Press.

Link, B. G., Dohrenwend, B. P., & Skodol, A. E. (1986). Socio-economic status and schizophrenia: Noisome occupational characteristics as a risk factor. *American Sociological Review, 51*(2), 242–258.

Lyons, M. J. (1996). A twin study of self-reported criminal behaviour. *Ciba Foundation symposium, 194,* 61–70.

Lyons, M. J., Nong, L., Eisen, S., Goldberg, J., True, W., & Tsuang, M. T. (1996). Genetic and environmental influences on psychiatric pathology over time. Paper presented at the 35th Annual Meeting of the American College of Neuropsychopharmacology, San Juan, Puerto Rico.

Lyons, M. I., True, W. R., Eisen, S. A., Goldberg, J., Meyer, J. M., Faraone, S. V., Eaves, L. J., & Tsuang, M. T. (1995). Differential heritability of adult and juvenile antisocial traits. *Archives of General Psychiatry, 52*(11), 906–915.

Macleod, J. (1978). *Davidson's principles and practices of medicine* (12th ed.). Edinburgh, UK: Churchill Livingstone.

Marino, S., Gallo, J. I., Ford, D., & Anthony, J. C. (1995). Filters on the pathway to mental health care, I. Incident mental disorders. *Psychological Medicine, 25*(6), 1135–1148.

McKusick, V. A. (1996). Book review: Genetics and medicine in the United States, 1800–1922 by A. R. Rushton. *Bulletin of the History of Medicine, 70*(4), 731–733.

Merikangas, K. R., Risch, N. J., Weissman, M. M. (1994). Comorbidity and co-transmission of alcoholism, anxiety and depression. *Psychological Medicine, 24*(1), 69–80.

Monk, M. (1987). Epidemiology of suicide. *Epidemiologic*

Reviews, 9, 51–69.

Morris, J. N. (1957). *Uses of epidemiology*. London: Livingstone.

Morrison, S. L. (1959). Principles and methods of epidemiological research and their application to psychiatric illness. *Journal of Mental Science, 105*(441), 999–1011.

Moscicki, E. K. (1995). Epidemiology of suicidal behavior. *Suicide and Life Threatening Behaviour, 25*(1), 22–35.

Mrazek, P. J., & Haggerty, R. J. (Eds.) (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington, DC: National Academy Press.

Muntaner, C., Anthony, J. C., Crum, R., & Eaton, W. W. (1995). Psychosocial dimensions of work and the risk of drug dependence among adults. *American Journal of Epidemiology, 142*(2), 183–190.

Murray, C. J. L., & Lopez, A. D., (Eds.) (1996) *The global burden of disease*. Geneva, Switzerland: World Health Organization.

Muthen, B. O. (1995). Factor analysis of alcohol abuse and dependence symptom items in the 1988 National Health Interview Survey. *Addiction, 90*(5), 637–645.

Muthen, B. O. (1996). Psychometric evaluation of diagnostic criteria: application to a two-dimensional model of alcohol abuse and dependence. *Drug and Alcohol Dependency, 41*(2), 101–112.

National Institute of Allergy and Infectious Diseases (1995). The relationship between the human immunodeficiency virus and the acquired immunodeficiency syndrome [on-line]. Available: http:www.niaid.nih.gov/publications/hivaids/all.htm. Bethesda, MD: Author.

Newcomb, M. D., & Bentler, P. M. (1988). The impact of family context, deviant attitudes, and emotional distress on adolescent drug use: Longitudinal latent-variable analyses of mothers and their children. *Journal of Research in Personality, 22*(2), 154–176.

Newcomb, M. D., & Felix-Ortiz, M. (1992). Multiple protective and risk factors for drug use and abuse: Cross-sectional and prospective findings. *J Pers Society Psychology, 63*(2), 280–296 .

O'Donnell, J., Hawkins, J. D., Catalano, R. F., Abbott, R. D., & Day, L. E. (1995). Preventing school failure, drug use, and delinquency among low-income children: Long-term intervention in elementary schools. *American Journal of Orthopsychiatry, 65*(1), 87–100.

Olsen, J., Schmidt, M. M., & Christensen, K. (1997). Evaluation of nature–nurture impact on reproductive health using half-siblings. *Epidemiology, 8*(1), 6–11.

O'Malley, P. M., Wagenaar, A. C. (1991). Effects of minimum drinking age laws on alcohol use, related behaviors and traffic crash involvement among American youth: 1976–1987. *Journal of Studies on Alcohol, 52*(5), 478–491.

Osler, W. (1897). Influence of Louis on American medicine. *Bulletin of the Johns Hopkins Hospital, 8*, 161–164.

Panum, P. L. (1940). Observations made during the epidemic of measles on the Faroe Islands in the Year 1846. J. A. Doull (Ed.), Trans. A. Hatcher, & J. Dimont. New York: Delta Omega Society.

Pentz, M. A., Dwyer, J. H., MacKinnon, D. P., Flay, B. R., Hansen, W. B., Wang, E. Y., & Johnson, C. A. (1989). A multicommunity trial for primary prevention of adolescent drug abuse. Effects on drug use prevalence. *Journal of the American Medical Association, 261*(22), 3259–3266.

Perry, C. L. (1987). Results of prevention programs with adolescents. *Drug and Alcohol Dependency, 20*(1), 13–19.

Perry, C. L., Williams, C. L., Veblen-Mortenson, S., Toomey, T. L., Komro, K. A., Anstine, P. S., McGovern, P. G., Finnegan, J. R., Forster, J. L., Wagenaar, A. C., & Wolfson, M. (1996). Project Northland: Outcomes of a communitywide alcohol use prevention program during early adolescence. *American Journal of Public Health, 86*(7), 956–965.

Pickens, R. W., Svikis, D. S., McGue, M., Lykken, D. T., Heston, L. L., & Clayton, P. J. (1991). Heterogeneity in the inheritance of alcoholism: A study of male and female twins. *Archives of General Psychiatry, 48*(1), 19–28.

Plomin, R. (1990). The role of inheritance in behavior. *Science, 248*(4952), 183–188.

Regier, D. A., Farmer, M. E., Rae, D. S., Myers, J. K., Kramer, M., Robins, L. N., George, L. K., Karno, M., & Locke, B. Z. (1993). One-month prevalence of mental disorders in the United States and sociodemographic characteristics: The Epidemiologic Catchment Area study. *Acta Psychiatrica Scandinavica 88*(1), 35–47.

Reid, D. D. (1960). *Epidemiological methods in the study of mental disorders. Public Health Paper No. 2*. Geneva: World Health Organization.

Risch, N. J., & Zhang, H. (1996). Mapping quantitative trait loci with extreme discordant sib pairs: sampling considerations. *American Journal of Human Genetics, 58*(4), 836–843.

Robins, L. N. (1966). *Deviant children grown up; A sociological and psychiatric study of sociopathic personality*. Baltimore: Williams and Wilkins.

Robins, L. N. (1978). Psychiatric epidemiology. *Archives of General Psychiatry, 35*(6), 697–702.

Robins, L. N., & Regier, D. A. (Eds.) (1991). *Psychiatric disorders of America*. New York: Free Press.

Rosanoff, A. J. (1917). Survey of mental disorders: In Nassau County, New York, July–Oct., 1916. *Psychiatric Bulletin, 2*(2), 1–73.

Rutter, M. (1981). Epidemiological/longitudinal strategies and causal research in child psychiatry. *Journal of the American Academy of Child Psychiatry, 20*(3), 513–544.

Rutter, M. (1989). Isle of Wight revisited: twenty-five years of child psychiatric epidemiology. *Journal of the American Academy of Child and Adolescent Psychiatry*, *28*(5), 633–653.

Rutter, M., Tizard, J., Yule, W., Graham, P., & Whitmore, K. (1976). Research report: Isle of Wight Studies, 1964–1974. *Psychological Medicine 6*(2), 313–332.

Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science*, *277* (5328), 918–924.

Scherrer, J. F., Lin, N., Eisen, S. A., Goldberg, J., True, W. R., Lyons, M. J., & Tsuang, M. T. (1996). The association of antisocial personality symptoms with marijuana abuse/dependence. A monozygotic co-twin control study. *Journal of Nervous Mental Disorders, 184*(10), 611–615.

Schlesselman, J. (1982). *Case-control studies*. New York: Oxford University Press.

Schwab, J. J., & Schwab-Stone, M. E. (1978). *Sociocultural roots of mental illness: An epidemiologic survey*. New York: Plenum.

Sheldon, T., & Chalmers, I. (1994). The UK Cochrane Centre and the NHS Centre for reviews and dissemination: Respective roles within the information systems strategy of the NHS R & D programme, coordination and principles underlying collaboration. *Health Economics*, *3*(3), 201–203.

Shepherd, M. (1985). Psychiatric epidemiology and epidemiological psychiatry. *American Journal of Public Health*, *75*(3), 275–276.

Shepherd, M., & Cooper, B. (1964). Epidemiology and mental disorder: A review. *Journal of Neurology, Neurosurgery, and Psychiatry*, *27*, 277–1283.

Sloan, J. H., Kellermann, A. L., Reay, D. T., Ferris, J. A., Koepsell, T., Rivara, F. P., Rice, C., Gray, L., & LoGerfo, J. (1988). Handgun regulations, crime, assaults, and homicide. A tale of two cities. *New England*

Journal of Medicine, *319*(19), 1256–1262.

Sloan, J. H., Rivara, F. P., Reay, D. T., Ferris, J. A., & Kellerman, A. L. (1990). Firearm regulations and rates of suicide: A comparison of two metropolitan areas. *New England Journal of Medicine*, *322*(6), 369–373.

Snow, J. (1855). *On the mode of communication of cholera*. London: John Churchill. Reprinted in W. H. Frost (Ed.) (1936). *Snow on cholera*. New York: Commonwealth Fund.

Spitzer, W. O., Haggerty, J. L., Berkson, L., Davis, W., Palmer, W., Tamblyn, R., Laprise, R., Faith, J. M., Elmore, J. G., & Horwitz, R. L. (1995). Continuing occurrence of eosinophilia myalgia syndrome in Canada. *British Journal of Rheumatology*, *34*(3), 246–251.

Strittmatter, W. J., Saunders, A. M., Schemechel, D., Pericak-Vance, M., Enghild, J., Salvesen, G. S., & Roses, A. D. (1993). Apolipoprotein E: High-affinity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proceedings of the National Academy of Sciences USA*, *90*(5), 1977–1981.

Susser, M., & Adelstein, A. (1975). An introduction to the work of William Farr. *American Journal of Epidemiology*, *101*(6), 469–476.

Susser, M., & Susser, E. (1996a). Choosing a future for epidemiology: II. From black box to Chinese boxes and eco-epidemiology. *American Journal of Public Health*, *86*(5), 674–677.

Susser, M., & Susser, E. (1996b). Choosing a future for epidemiology: I. Eras and paradigms. *American Journal of Public Health*, *86*(5), 668–673.

Sydenstricker, E. (1974). *The challenge of facts*. R. V. Kasius (Ed.). New York: Prodist.

Tansella, M., de Girolamo, G., & Sartorius, N. (1992). *Annotated bibliography of psychiatric epidemiology*. London: Gaskell.

Tien, A. Y., & Anthony, J. C. (1990). Epidemiological analysis of alcohol and drug use as risk factors for psychotic experiences. *Journal of Nervous and Mental Diseases, 178,* 473–480.

Tsuang, M. T., Tohen, M., & Zahner, G. E. P. (Eds.) (1995). *Textbook in psychiatric epidemiology*. New York: Wiley-Liss.

Tuke, D. H. (Ed.) (1892). A dictionary of psychological medicine. 2 vol. Reprinted by Arno Press, New York.

Uhl, G., Blum, K., Noble, E., & Smith, S. (1993). Substance abuse vulnerability and DRD2 receptor genes. *Trends in Neuroscience 16*(3), 83.

Van Etten, M. L., Neumark, Y. D., & Anthony, J. C. (1997). Initial opportunity to use marijuana and the transition to first use: United States, 1979–1994. *Drug and Alcohol Dependence, 49,* 1–7.

Watt, R. (1813). Appendix to treatise on Chincough (pp. 375–379). As cited in Farr, William (1885). *Vital statistics: A memorial volume of selections from the reports and writings of William Farr*. Edited for the Sanitary Institute of Great Britain, by Noel A. Humphreys. London: The Sanitary Institute of Great britain.

Weissman, M. M. (1993). Family genetic studies of panic disorder. *Journal of Psychiatric Research, 27* (Suppl. 1) 69–78.

Wines, F. H. (1888). *Report on the defective, dependent, and delinquent classes of the population of the United States as returned at the Tenth Census (June 1, 1880)*. Washington, DC: Government Printing Office.

Wynder, E. L. (1994). Invited commentary: Studies in mechanism and prevention. Striking a proper balance. *American Journal of Epidemiology, 139*(6), 547–549.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.13

Learning Theory

GRAHAM C. L. DAVEY
University of Sussex, UK

1.13.1 INTRODUCTION 391

1.13.2 BASIC DEFINITIONS, CONCEPTS, AND PROCESSES OF CONDITIONING 393
- *1.13.2.1 Classical Conditioning* 393
  - *1.13.2.1.1 Pavlov's prototypical conditioning procedure* 393
  - *1.13.2.1.2 The nature of the conditioned response* 393
  - *1.13.2.1.3 Control procedures* 393
  - *1.13.2.1.4 Basic phenomena of classical conditioning* 394
  - *1.13.2.1.5 Special features of human classical conditioning* 396
- *1.13.2.2 Operant Conditioning* 398
  - *1.13.2.2.1 Basic concepts in operant conditioning* 398
  - *1.13.2.2.2 Basic phenomena of operant conditioning* 399
  - *1.13.2.2.3 Special features of operant conditioning in humans* 400

1.13.3 HISTORICAL ASPECTS OF LEARNING THEORY AND CLINICAL PHENOMENA 401
- *1.13.3.1 Classical Conditioning* 401
  - *1.13.3.1.1 The "Little Albert" study and attempts at its replication* 401
  - *1.13.3.1.2 Pavlov and experimental neuroses* 402
  - *1.13.3.1.3 Two-factor theory of neuroses* 403
  - *1.13.3.1.4 Eysenck's theory of incubation* 403
  - *1.13.3.1.5 The origins of behavior therapy* 404
- *1.13.3.2 Operant Conditioning* 404
  - *1.13.3.2.1 Principles of the experimental analysis of behavior and their application to psychopathology* 404
  - *1.13.3.2.2 Operant reinforcement and psychopathology* 405

1.13.4 CURRENT APPLICATIONS OF LEARNING THEORY IN CLINICAL AND EXPERIMENTAL PSYCHOPATHLOGY 406
- *1.13.4.1 Classical Conditioning* 406
  - *1.13.4.1.1 Traditional criticisms of classical conditioning models of phobias* 406
  - *1.13.4.1.2 Contemporary cognitive models of human classical conditioning* 407
  - *1.13.4.1.3 How contemporary conditioning models of phobias account for the traditional criticisms* 407
  - *1.13.4.1.4 Further applications of classical conditioning to anxiety-based disorders* 410
  - *1.13.4.1.5 Classical conditioning in behavior therapy* 410
- *1.13.4.2 Operant Conditioning* 412
  - *1.13.4.2.1 Functional analysis* 412
  - *1.13.4.2.2 Token economy schemes* 413
  - *1.13.4.2.3 Response shaping* 414
  - *1.13.4.2.4 Multifaceted behavioral self-control programs* 415

1.13.5 SUMMARY 416

1.13.6 REFERENCES 416

1.13.1 INTRODUCTION

Ever since the first controlled experiments were systematically carried out on learning and conditioning in the early part of the twentieth century, the study and understanding of the processes of learning have contributed significantly to all aspects of clinical psychology. The classical conditioning paradigm established by Pavlov very soon led to the investigation of psychopathology in terms of learning processes, with Pavlov's own studies of experimental neuroses in animals leading to a conceptualization of many aspects of psychopathology not in terms of dysfunction but in terms of the animal's experience of controlled replicable environmental contingencies. This process began the demystification of psychopathology, and enabled clinicians to remove much of the stigma that had been associated with psychopathology—psychological disorders were not so much the result of dysfunctional psychological systems, but the result of learning experiences generated by perfectly normal learning mechanisms.

Pavlov's work on classical conditioning led rather naturally to the use of that conditioning paradigm in the explanation of a variety of acquired psychological disorders. The most famous of these was the use of the classical conditioning paradigm in the explanation of specific phobias by J. B. Watson and his coworker Rayner in the early 1920s. Their study of the classical conditioning of a specific phobia in eight-month old "Little Albert" (Watson & Rayner, 1920) has entered psychological folklore as the prototypical conditioning study of acquired fears. It is still quoted in nearly all clinical psychology textbooks that deal with anxiety, even though most texts criticize the ability of this study to explain many of the important features of acquired fears (see Sections 1.13.3.1.1 and 1.13.4.1.1). Even though it has had a significant impact on contemporary models of acquired fears (Davey, 1992a), the original Watson and Rayner study is poorly reported and frequently misdescribed in secondary texts (Harris, 1979). Nevertheless, this should not underestimate the importance of the conditioning paradigm in contributing to the understanding of a range of acquired fears, including specific phobias, panic disorder and, more recently post-traumatic stress disorder (PTSD).

While the studies of classical conditioning processes were providing useful insights into the etiology of anxiety-based disorders, it became clear that if conditioning theory could provide an explanation of these problems, then they should also provide principles relevant to the development of successful therapeutic procedures. This was exactly the thinking of Joseph Wolpe and Hans Eysenck in the 1950s. Disillusionment with the failure of psychodynamic therapies to fit comfortably into emerging objective and scientific conceptualizations of psychological processes led to the development of behavior therapies. These therapies were based on objectively derived principles of conditioning and were consistent with the kinds of processes that conditioning theories assumed were responsible for the generation of anxiety and emotion-based disorders. Put in its simplest form, if anxiety-based disorders were manifestations of inappropriately learned conditioned responses, then those conditioned responses could be eliminated by adopting procedures which would extinguish those responses. This, as we shall see later, was the rationale for the development of a variety of highly successful behavior therapies, including flooding, counterconditioning, and systematic desensitization (see Section 1.13.4.1.5).

Apart from the use of classical conditioning procedures in both the explanation and amelioration of psychological disorders, the development of operant conditioning paradigms offered completely new and different applications to clinical psychology. While there were clearly some psychological problems whose etiology could be explained by reference to the nature of the consequences of behavior (e.g., through conducting a functional analysis of the contingencies which were maintaining the problem), what operant conditioning also offered was a valuable tool for changing behavior. This has generated a wide range of behavior therapy and behavior modification programs that not only help to maintain psychologically healthy and desirable behaviors (e.g., *The token economy*; Ayllon & Azrin, 1968), but also enable the therapist to develop new behavior repertoires where none existed before (e.g., in establishing basic self-help and communication behaviors in individuals with learning disabilities). What is so valuable about the paradigm offered by operant conditioning is that it enables the therapist to examine the functional relationships between problem behavior and environmental contingencies, and in many cases allows the therapist to provide the client with a treatment program that the client can establish and execute on his or her own. Such behavioral self-control programs have become a valuable addition to many therapeutic programs (e.g., diet control, treatment of addictions, etc., see Section 1.13.4.2.4).

This chapter will provide a structured overview of the contribution of learning theory to clinical psychology and will begin with a description and explanation of the basic processes of conditioning. This will be followed by an assessment of the seminal historical contributions from learning theorists. The chapter will conclude with an overview of the contribution of contemporary conditioning theory to the etiology and treatment of psychological disorders.

1.13.2 BASIC DEFINITIONS, CONCEPTS, AND PROCESSES OF CONDITIONING

1.13.2.1 Classical Conditioning

1.13.2.1.1 Pavlov's prototypical conditioning procedure

In classical conditioning a conditioned stimulus (CS) is paired in close temporal contiguity with an unconditioned stimulus (UCS). After a number of CS-UCS pairings, the CS comes to elicit a conditioned response (CR) relevant in some way to the UCS. This process is displayed schematically in Figure 1. The UCS is usually some biologically important event which itself evokes an unlearned, reflexive reaction known as the unconditioned response (UCR). In Pavlov's prototypical conditioning experiment, a hungry dog was presented with pairings of a bell or metronome (the CS) and a bowl of food (the UCS), after a number of pairings of CS and UCS, the dog began to salivate to the CS (the CR).

UCSs used to generate conditioned responding are typically either appetitive or aversive. Appetitive UCSs are such things as food or water to a hungry or thirsty animal. Aversive UCSs are stimuli which the animal finds either painful or distressing (such as a brief mild electric shock or a loud noise).

1.13.2.1.2 The nature of the conditioned response

In most traditional classical conditioning studies (such as Pavlov's traditional salivary conditioning procedure), the CR is often what looks like a faithful facsimile of the UCR. In Pavlov's study the UCR is salivation and the learned CR to the CS is also salivation. However, the learned CR can often be rather different from the type of response elicited by the UCS. For example, in rats an aversive UCS such as electric shock will elicit a UCR which consists of heart-rate increases, jumping, and squealing, whereas the CS which is paired with electric shock often results in the acquisition of "emotional" anticipatory responses in the presence of the CS—such responses include heart-rate decreases, freezing defecating, etc. (Black, 1971; Borgealt, Donahoe, & Weinstein, 1972). Some traditional explanations of classical conditioning have argued that what occurs during classical conditioning is a form of "reflex transfer" in which the CS becomes a "substitute" for the UCS and thus elicits the same responses as the UCS (cf. Davey, 1989b). This is clearly not the whole story. The CS does elicit a range of responses which are relevant to the UCS, but they are not necessarily the responses that are elicited by the UCS. This has important implications for the acquisition of specific fears in humans. The dental phobic does not wince and cry out in pain while sitting in the dentist's waiting room; what is elicited by the dentist's waiting room is an anticipatory fear rather than responses that would normally be reactions to the dentist's drill. What the reader should be aware of is that classical conditioning is not necessarily the nonconscious, reflexive form of learning that is portrayed in traditional textbooks. It is a highly cognitive form of associative learning which permits a wide range of conscious and nonconscious reactions to be emitted to the CS as a result of learning about the relationship between the CS and UCS (Davey, 1992a; see Sections 1.13.2.1.4 (vi) and 1.13.4.1.2).

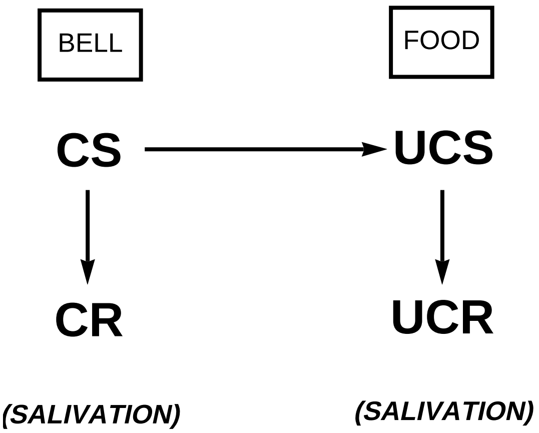

Figure 1 Schematic representation of the typical classical conditioning procedure. Pairing a conditioned stimulus, CS (e.g., a bell), with an unconditioned stimulus, UCS (e.g., food), eventually evokes the conditioned response, CR, to the CS (in this case salivation).

1.13.2.1.3 Control procedures

The important experimental manipulation in classical conditioning is the paired relationship between the CS and the UCS. Therefore, in order to conclude that classical conditioning has occurred, it is important to be able to determine that any responses which are elicited by the CS are the result of the CS being paired with the UCS. Clearly, there are other learning processes, such as habituation and sensitization, that could produce spurious behavior changes in a classical conditioning procedure. For example, a study by Grether (1938) first made monkeys fearful by presenting a number of powder flashes. Following this, the monkeys were given a series of presentations of a bell, all of which elicited reliable fright reactions which they had not done prior to the powder flashes. In this example, the powder flash was a UCS and the bell a CS, yet a fearful response to the CS had been established without any pairing between CS and UCS.

Because the important manipulation in classical conditioning is the arrangement of the CS and UCS in a consistent temporal or predictive relationship, any classical conditioning experiment should possess controls which retain all the features of the conditioning procedure *except* the predictive relationship between CS and UCS. The most appropriate control procedure is known as the truly random control (Rescorla, 1967). In this control procedure, subjects are given a similar number of CS and UCS presentations to experimental subjects, but CS and UCS are programmed independently of each other. Thus, control subjects receive exactly the same number of CS and UCS presentations, but not the predictive relationship between them. This distinction is important particularly in relation to the acquisition of fearful or anxious responding, because such responding can be acquired through nonassociative processes such as sensitization as demonstrated in the Grether experiment (above). Models of psychopathology based on classical conditioning emphasize the importance of the learning of an *association* between the CS and UCS (e.g., Davey, 1992a, 1997), and it is this association that generates many of the features of psychopathological responding (see Section 1.13.4.1).

1.13.2.1.4 Basic phenomena of classical conditioning

(i) Predictive value of the CS

Traditionally it has been assumed that the important feature of classical conditioning which generated the CR was the simple pairing of CS and UCS. That is, it was assumed that the number of pairings of the CS with a temporally contiguous UCS was a prime determinant of the strength of the CR (e.g., Kimble, 1961). Nevertheless, while temporal contiguity is an important contributor to learning, it is not necessarily the number of CS–UCS pairings that is important, but the predictive significance of the CS as measured by the correlation between the CS and UCS. When we come to discuss the features of classical conditioning in humans, it will become clear that it is not the number of pairings between CS and UCS that determines either the speed of acquisition of the CR or the strength of the CR. Reliable conditioning can occur after a single pairing of CS and UCS as long as the subject was immediately aware of the predictive significance of the CS (cf. Davey, 1992a; Dawson & Schell, 1987;). Indeed, if human subjects are pre-experimentally informed of the predictive relationship between the CS and the UCS, a full-blown CR can be exhibited on the first presentation of the CS before *any* CS–UCS pairings have been experienced (Dawson & Grings, 1968; Deane, 1969; Wilson, 1968).

(ii) Extinction

If, after conditioning has been acquired, a CS is subsequently presented alone, the strength of the CR to that CS subsequently declines over trials until the CS no longer elicits a CR. This is known as extinction. This process has been used extensively in the application of conditioning models to psychopathology, largely in the development of behavior treatments for emotional disorders (e.g., flooding, systematic desensitization). These therapies attempt to ensure that the emotion-eliciting stimulus (the CS) is not presented in conjunction with an aversive UCS (such as fear), and thus the learned relationship between the CS and trauma is extinguished. There is some doubt about whether extinction represents the unlearning of previously acquired CS–UCS associations, because presenting a novel stimulus or an additional single CS–UCS pairing can often fully reinstate a previously extinguished CR (Pavlov, 1927). This suggests that extinction may represent either the inhibition of a learned CR or the learning of new CS–UCS associations which override the original ones (cf. Davey, 1989c). In either case, extinguished CRs do not have to be totally relearned in order to be reinstated and their spontaneous recovery may explain at least some instances of relapse following treatment of emotional disorders by therapies based on extinction principles.

(iii) Latent inhibition

When a CS is presented alone on a number of occasions prior to being paired with a UCS, it subsequently takes significantly longer to generate a CR to that CS than when no preexposure has taken place. This is known as latent inhibition and has been widely demonstrated in a variety of classical conditioning preparations in both animal and human subjects (Lubow, 1973; Siddle & Remington, 1987). This feature of classical conditioning is potentially important in a number of conditioning theories of psychopathology because it predicts that if a previously neutral CS (e.g., a car) is paired with an aversive UCS (e.g., a traumatic car accident), then whether the individual acquires a conditioned fear response to cars will be inversely proportional to the amount of prior nonaversive experience they have had with cars. This principle, for example, can potentially help to explain why some people acquire a phobia following pairing of a stimulus or event with a traumatic UCS and some do not (cf. Davey, 1989a).

(iv) Blocking

If a UCS is already reliably predicted by a CS (call it A), then subsequently signaling that UCS with a compound CS consisting of the original CS, A, plus a new component (call it B), results in little or no learning about the new component B (Kamin 1968, 1969). This is known as blocking, and it is a robust phenomenon in animal conditioning (cf. Mackintosh, 1983). Blocking is normally considered to be an associative phenomenon in that organisms do not appear to learn about the new component B because the UCS is already reliably signaled by component A (Dickinson & Mackintosh, 1979; Rescorla & Wagner, 1975). Although this is an important phenomenon for theories of associative learning in animals (Pearce & Hall, 1980; Rescorla & Wagner, 1972), it has rarely been applied in the context of conditioning models of psychopathology. This may be because it has been difficult to experimentally isolate and demonstrate examples of blocking in human subjects (Davey, 1992a; Davey & Singh, 1988).

(v) Higher-order conditioning

Once a CS has been associated with a UCS and is capable of eliciting a reliable CR, that CS can then be used to reinforce other potential CSs. For instance, second-order conditioning can be demonstrated using the following procedure: a CS1 (e.g., a light) is paired with a UCS (e.g., food); then CS2 (e.g., a tone) is paired with CS1 (the light). This will usually result in a CR relevant to the original UCS (food) being evoked by CS2, even though CS2 has never been directly paired with food (e.g., Rescorla, 1980; Rizley & Rescorla, 1972). This phenomenon also has potential importance for conditioning models of psychopathology because it implies that emotional reactions can be acquired through higher-order conditioning in which the potentially phobic CS has never been paired directly with a traumatic UCS.

(vi) What is learnt during classical conditioning?

Although we know that pairing a CS with a UCS will generate a CR to the CS, we cannot easily explain how this happens. Most probably the organism has learnt to associate certain contiguous or predictive events together and it is these learned associations which mediate the CR. However, the theoretical problem is that we cannot directly observe what associations have been formed but have to infer them from the experimental manipulations that are carried out. We can ask very generally what kinds of things the organism might have learnt during classical conditioning which led it to emit a CR during the CS. For example, Pavlov's dog may have learned that the bell (CS) predicts food; that is, after some pairings of the bell (CS) with food (UCS), the bell comes to elicit a memory or internal representation of the food that is to follow, and it is this evoked representation that stimulates salivation. Alternatively, the dog may be much less "cognitive" in its learning. Instead of becoming associated with food, the bell may have become associated with the act of salivation which is also associated with food. Thus, a direct link may have been formed between centers representing the bell and the reflex arc which controls salivation. In order to differentiate between these two possibilities, animal learning theorists have developed inferential techniques which enable them to ascertain exactly what kinds of learned associations are acquired during classical conditioning.

One simple technique is known as postconditioning stimulus revaluation (Dickinson, 1980; Rescorla, 1980). This procedure involves attempting to revalue the UCS for the subject, and normally consists of three stages: (a) subjects are given pairings of CS and UCS until a CR is established (b) subjects are then given off-the-baseline training with the UCS alone in which the UCS is revalued (e.g., a food UCS may be paired with gastric poisoning to make it less attractive and less palatable); and (c) the animal is then given test presentations of the CS. The logic here is that, if the CR is mediated by a CS–UCS association, then revaluation of the UCS will also affect the CR because the CR is mediated via the UCS representation. If, however, the CR is mediated by a more reflexive association between the CS and the UCR reflex, then the CR will not be affected by changes in the evaluation of the UCS. Most studies which have used this technique have demonstrated that first-order classical conditioning (both appetitive and aversive) appears to be mediated by CS–UCS associations (cf. Davey, 1989c; Dickinson, 1980; Rescorla, 1980) and not by reflexive associations formed between the CS and the response to the UCS. These findings have a number of important implications for our conceptions of the classical conditioning process, and for the application of classical conditioning to psychopathological phenomena. First, it implies that classical conditioning is not a simple, mechanistic or reflexive form of learning. Instead, it involves (a) the learning of associations between contiguous and predictive events, and (b) mediation of the CR by representations of the UCS and the knowledge of the UCS contained in that representation. Thus, classical conditioning is a complex cognitive process which involves the formation of representations of events as well as the establishment of associations. Second, since CRs generally appear to be mediated by information about the UCS contained in a UCS representation, then the nature or strength of the CR can be affected by changes in the information contained in the UCS representation. This has extremely important implications for conditioning models of emotional disorders, and has been a central feature of contemporary conditioning models of phobias (Davey, 1992a, 1997; see Sections 1.13.4.1.2 and 1.13.4.1.3).

1.13.2.1.5 Special features of human classical conditioning

(i) Similarities between animal and human conditioning

Davey (1992a) has argued that while there are some important differences in classical conditioning between human and nonhuman animals (see below), the basic conditioning phenomena tend to be common to both. For example, human subjects exhibit blocking (Dickinson, Shanks, & Evenden, 1984) and latent inhibition (Siddle & Remington, 1987; Siddle, Remington, & Churchill, 1985). Higher-order conditioning can be demonstrated in humans (Davey & Arulampalam, 1981, 1982; Davey & McKenna, 1983), and CR strength in humans can be shown to be a function of the predictive significance of the CS as measured by the correlation between the CS and UCS (Alloy & Tabachnik, 1984; Prokasy & Kumpfer, 1969, Prokasy & Williams, 1979). In terms of the associations that are learned during classical conditioning, human subjects also appear to learn to associate CS and UCS so that the CR is mediated by the subject's internal representation of the UCS (Davey & McKenna, 1983; White & Davey, 1989).

(ii) Conscious awareness of contingencies

A large number of studies have suggested that human subjects only exhibit a differential CR when they are able to verbalize the CS–UCS contingency (see Dawson & Schell, 1987 for a review). Only subjects who can verbalize the correct CS–UCS relationship in postexperimental interviews exhibit differential CRs (Baer & Fuhrer, 1968, 1970); when awareness of contingencies is measured on a trial-by-trial basis, differential conditioning appears only after the appearance of contingency awareness (Dawson & Biferno, 1973; Dawson, Schell, & Tweddle-Banis, 1986), and subjects regularly fail to exhibit differential conditioning in studies which deliberately attempt to "mask" the relationship between CS and UCS by employing distracter tasks (Dawson, 1973; Dawson, Catania, Schell, & Grings, 1979). These findings have led some theorists to suggest that the classical conditioning of autonomic responses (such as those concerned with emotional responding) is only possible when individuals are consciously aware of the CS–UCS relationship, and that this can only occur as a result of strategic processing of the contingencies (e.g., Dawson & Schell, 1987; Dawson, Catania, Schell, & Grings, 1979). However, there are other forms of human classical conditioning that appear to occur in the absence of awareness of the CS–UCS contingency. There are (a) the evaluative conditioning procedure, where a subjectively neutral stimulus (CS) is paired with a liked or disliked stimulus (UCS) which results in the CS acquiring positive or negative valence depending on the valency of the UCS (Baeyens, Eelen, Van den Bergh, & Crombez, 1989, 1990), and (b) human aversive conditioning studies in which the CS is a picture of a fear-relevant stimulus (e.g., a snake) which is presented subliminally (Esteves, Dimberg, Parra, & Öhman, 1994; Öhman, 1993). Nevertheless, there is still much theoretical argument about whether these latter two paradigms represent conditioning without awareness (Davey, 1994, 1995), and a definitive answer to the question of whether conscious processing of contingencies is necessary for classical conditioning in humans is still awaited.

(iii) Re-evaluation of the UCS

Studies using the postconditioning stimulus reevaluation procedure (see Section 1.13.2.1.4 (vi)) have shown that by independently manipulating the animal's evaluation of the UCS, the strength or nature of the CR can also be changed. This implies that there is a nonassociative process that can affect the strength of the CR which is independent of the strength of the association between the CS and UCS. For example, in aversive conditioning, if the animal's evaluation of the UCS is changed so that it now perceives the UCS as being less aversive than it originally was, then the next presentation of the CS will evoke a significantly weaker CR.

Davey (1989c, 1992a) has argued that the process of UCS reevaluation is significantly more important in modulating the strength of the CR in humans than it is in nonhuman animals, and that because humans are social animals that have developed sophisticated communication skills, there will be many ways in which experiences and communicated information can alter an individual's perception and evaluation of a UCS. Thus, because learnt CRs are mediated via a representation of the UCS, then arguably the most important contributor to the strength of a CR in human classical conditioning is not necessarily the strength of the association between CS and UCS, but how the individual perceives and evaluates the UCS. This is critically important when it comes to considering conditioning models of psychopathology, and will be discussed more fully in Section 1.13.4.1.3.

However, it is worth listing the kinds of processes that can contribute to changes in evaluation of the UCS in human subjects.

(a) Experiences with the UCS alone. Some postconditioning experiences with the UCS in the absence of the CS can lead the individual to revalue the UCS. For example, an individual may reassess a UCS more favorably if they experience a number of UCS-alone trials which allow their fear to habituate (Davey & McKenna, 1983). Similarly, the perceived aversiveness of the UCS may be inflated by experiences with a similar UCS of greater intensity (e.g., White & Davey, 1989). Both of these processes will have a subsequent effect on the strength of the CR when the CS is next presented.

(b) Socially/verbally transmitted information about the UCS. In a laboratory conditioning experiment, subjects can simply be told that on future presentations the UCS will be more or less intense than before. If the subject believes this information, then the evaluation of the UCS is changed, and this affects the strength of the CR to subsequent CS presentations (Davey & McKenna, 1983).

(c) Interpretation of interoceptive cues. Many anxious people attend to their own bodily sensations and use these stimuli as a means of assessing the aversive nature of potentially threatening consequences (cf. Davey, 1988; Parkinson, 1985). As a result, the individual's reaction to either the CS or UCS can act as an important source of information for evaluating the UCS. For example, if they believe they are emitting a strong fear response to a CS (even though this may not be the case physiologically), then they will often attribute this to being fearful of the UCS, and this inflation of the aversive evaluation of the UCS in turn does produce an actual increase in the strength of the CR (Davey & Matchett, 1996; Russell & Davey, 1991).

(d) Cognitive rehearsal of the UCS. It is clear that individuals who suffer anxiety disorders have a tendency to focus on and rehearse the possible aversive outcomes of phobic encounters (Marks, 1987), and this ruminative tendency may act to inflate their aversive evaluation of the potential aversive outcome (UCS). A laboratory conditioning study by Davey and Matchett (1994) showed this to be the case, but only in subjects who were already anxious. When asked to experimentally rehearse the UCS after conditioning had been completed, subsequent presentations of the CS elicited a greater magnitude fear CR than had occurred prior to rehearsal, but this effect was found only in subjects with high levels of trait anxiety or who had undergone a procedure which induced an acute anxious state. Davey and Hatchett argued that this increase in the strength of the CR following rehearsal of the UCS in anxious subjects was the result of rehearsal inflating the aversive evaluation of the UCS in these subjects.

(e) Coping strategies which neutralize the UCS. Many individuals who have aversive-conditioning experiences during their daily lives (e.g., traumatic car accidents, being bitten by a dog, painful or traumatic dental experiences) subsequently fail to develop a learned anxiety or fear reaction to these stimuli. One reason for this failure may be the ability of some individuals to devalue or neutralize the trauma immediately following the experience, and they may do this by adopting appraisal strategies which allow them to effectively devalue the stressful meaning of the trauma. In a recent study Davey, McDonald et al. (1997) identified seven factorially independent reappraisal strategies all of which contribute to trauma or UCS devaluation. These are downward comparison (e.g., "other people are worse off than me") (Wills, 1981), positive reappraisal (e.g., "in every problem there is something good") (Davey, 1993), cognitive disengagement (e.g., "the problems involved in this situation simply aren't important enough to get upset about"), optimism (e.g., "everything will work itself out in the end") (Scheier & Carver, 1992), faith in social support (e.g., "I have others who can help me through this"), denial (e.g., "I refuse to believe this is happening") (Breznitz, 1983), and life perspective (e.g., "I can put up with these problems as long as everything else in my life is okay"). Davey et al. (1997) found that use of these strategies (except for denial) was positively correlated with measures of psychological health and inversely correlated with a variety of measures of psychopathology. In addition,

Davey, Burgess, and Rashes (1995) found that both simple phobics and panic disorder patients differed from normal controls by reporting reduced use of appraisal strategies which devalue the stressful meaning of a trauma. Such coping strategies may therefore play an important role in insulating individuals from the acquisition of phobic responding following traumatic conditioning experiences.

(iv) UCS expectancy biases

It was mentioned in Section 1.13.2.1.4 (i) that human subjects will emit a fear CR to the first presentation of a CS if they have previously been informed that the CS is to be followed by an aversive UCS. Thus, a UCS "expectancy" can be generated simply through verbal transmission of the contingencies. Life consists of regular discussion and transmission of contingencies relevant to daily living, so it is not unreasonable to suppose that individuals will develop beliefs and expectancies about what will happen when they encounter a conditioning episode. As a result, when they enter a conditioning episode, individuals are not *tabula rasa* but will hold beliefs and expectancies about what is likely to happen.

An example of the associative biases that influence the formation of CS–UCS associations can be found in covariation assessment studies. Studies of covariation have pointed out that assessing whether two stimuli covary appears to be influenced by both situational information (i.e., current information about the contingency) and prior expectations or beliefs about the covariation (e.g., Alloy & Tabachnik, 1984; Crocker, 1981). There are a variety of circumstances in which the combination of situational information (e.g., the experienced CS–UCS contingency) and prior expectancies (e.g., the belief that a CS is to be followed by a particular kind of UCS) give rise to what is called a covariation bias which generates a distorted perception of the covariation—usually in the direction of the prior expectation (Alloy & Tabachnik, 1984).

Expectancy or covariation biases can be found quite frequently in human classical conditioning, especially in circumstances where subjects will have developed a belief that a particular type of CS may be followed by a particular type of UCS (even though that is not necessarily the case). For example, when pictures of fear-relevant stimuli (i.e., stimuli that are potentially dangerous such as a snake or a gun) are used as CSs in aversive conditioning studies, subjects acquire the CR more rapidly than if the CS were a picture of a fear-irrelevant stimulus (e.g., a flower). The CR in such circumstances also shows a significant resistance to extinction (Hugdahl & Johnsen, 1989; Öhman, 1979). A series of experiments by Davey (1992b) showed that this tendency for fear-relevant stimuli to show stronger conditioning effects than fear-irrelevant stimuli was probably the result of a pre-experimental expectancy bias. In this study, Davey used a "threat" conditioning procedure (where subjects are warned that they might receive aversive UCSs following some CSs, but in fact receive none) and found that subjects began the experiment with a significantly higher expectancy of aversive UCSs following fear-relevant stimuli than fear-irrelevant stimuli. These studies indicate that individuals enter a conditioning episode believing that fear-relevant stimuli are more likely to have aversive consequences than fear-irrelevant stimuli. This speeds up the learning of associations between fear-relevant CSs and aversive UCSs, and also makes them more resistant to extinction. There has been considerable debate about how such expectancy biases are derived, with some arguing that they have been biologically determined as a result of evolutionary pressures (e.g., Cook, & Mineka 1990; Öhman, Dimberg, & Öst, 1985), and others that they are the result of culturally transmitted information about the stimulus (e.g., Davey, 1995; Merkelbach & de Jong, 1997). This discussion has particular relevance to the application of conditioning models to phobic responding, and we will return to it in Section 1.13.4.1.3.

1.13.2.2 Operant Conditioning

1.13.2.2.1 Basic concepts in operant conditioning

Whereas classical conditioning involves a contingency relationship between two environmental stimuli (the CS and UCS), operant conditioning involves a contingency between some aspect of the organism's behavior and some aspect of the environment. Thus, whereas classical conditioning involves *stimulus–stimulus contingencies,* operant conditioning involves *response–stimulus contingencies*. Operant conditioning acts to alter the frequency of the response involved in the contingency relationship in a predictable way, depending on the nature of the consequence of the response. Those consequences which increase the future probability of the response are known as *reinforcers*, and those that decrease the future probability of the response are called *punishers*. A common example of a reinforcer is food for a hungry animal, and a frequently used punisher in experimental studies of operant conditioning is mild electric shock (cf. Davey, 1989b). Reinforcers and punishers can be either positive or negative, depending on the kind of change the response exerts on the environment. Positive means that the response adds something to the situation (e.g., it delivers a food pellet or electric shock), and negative means that the response removes or avoids some environmental consequence (e.g., it terminates or avoids an electric shock). Thus, reinforcer/punisher refers to the effect of the consequence on the frequency of the response, and the adjective positive/negative refers to the nature of the response consequence. This terminology is illustrated in Figure 2.

A further term that is used frequently in operant conditioning is *discriminative stimulus*. When a particular response is reinforced only in the presence of a specific stimulus (e.g., a light), the organism will come to make that response only in the presence of that stimulus and not when the stimulus is absent. In this case the light is known as a discriminative stimulus because it "sets the occasion" for the organism to respond and for the response to be reinforced. The process of learning to respond in the presence of a discriminative stimulus is called *discrimination learning*.

In an organism's normal environment many responses are maintained by reinforcement contingencies. Analyzing an organism's interactions with its environment in order to discover the nature and frequency of the reinforcement contingencies that are maintaining behaviors is know as *functional analysis*, and functional analysis is a process that has been regularly utilized in clinical psychology to discover the environmental contingencies that might be maintaining problematic or dysfunctional behavior.

1.13.2.2.2 Basic phenomena of operant conditioning

(i) Superstitious reinforcement

It is unclear what aspect of the relationship between response and reinforcer is important for the operant conditioning process. It could be that a causal relationship between response and reinforcer is necessary for operant reinforcement to occur; alternatively, it could merely be temporal contiguity between a response and a reinforcer that is necessary. Skinner (1948) argued that organisms are sensitive to contiguity between response and reinforcer in such a way that mere accidental correlation between response and reinforcer will result in the organism learning about the relationship and modifying its behavior accordingly. For example, winning money on a fruit machine may accidentally coincide with a sequence of behaviors emitted by the player (e.g., pressing buttons in a particular sequence with a particular speed, etc.). This sequence of

	REINFORCER	**PUNISHER**
POSITIVE	RECEIVING A FOOD PELLET FOR PRESSING A LEVER	RECEIVING AN ELECTRIC SHOCK FOR PRESSING A LEVER
POSITIVE	PRESSING A LEVER SERVES TO AVOID OR ESCAPE ELECTRIC SHOCK	PRESSING A LEVER DELAYS OR PREVENTS THE DELIVERY OF FOOD

Figure 2 Reinforcement and punishment as procedures in operant conditioning. Each cell contains an example of positive and negative reinforcement and punishment.

behaviors has no causal relationship to winning money, but may be perceived as doing so by the player. Hence the player emits these sequences more frequently, and they are then more likely to be accidentally associated with winning in the future, thus strengthening the ritual. This process is known as *superstitious reinforcement*, and is a common feature of operant conditioning. In clinical contexts, superstitious reinforcement may be responsible for a variety of phenomena, including inappropriate rituals and the accidental reinforcement of disruptive or dysfunctional behaviors.

(ii) Response shaping

From the basic principles of operant conditioning, it is obvious that a particular behavior must first occur for some unspecified reason before it can be reinforced. When a behavior is commonplace in an organism's repertoire the problem of this initial occurrence is not an important one. However, when one wishes to reinforce a behavior pattern which occurs either very rarely or is particularly complex, the problem of the first occurrence of the behavior is a salient one. Instead of waiting indefinitely for the first occurrence of a rare behavior, the way around this is to reinforce *successive approximations* to the desired behavior. For instance, in order to speed up the acquisition of lever-pressing in a rat, approximations to this behavior can be successively reinforced. The experimenter may first choose to reinforce going to the side of the apparatus where the lever is situated (a response that is likely to occur quite frequently). When this is well-established, the response requirement is made stricter by reinforcing the animal only when it touches the lever, and so on. This process of response shaping is an important one in behavior modification programs based on operant conditioning, because it allows new and complex behavior repertoires to be constructed where none had existed before (e.g., in building up self-help and communication skills in individuals with learning disabilities) (see Section 1.13.4.2.3).

(iii) Conditioned reinforcement

Some operant reinforcers are considered to have intrinsic and immediate biological importance (e.g., food to a hungry animal). However, effective reinforcers can be established which do not have this intrinsic biological value. *Conditioned reinforcers* are those which acquire their reinforcing properties by being paired with biologically important stimuli. If, for instance, a hungry rat is trained to press a lever to obtain food, and food delivery is accompanied by a brief tone stimulus, that tone stimulus is likely to acquire reinforcing properties. This can be demonstrated by showing that rats which have received such training will continue to press the lever merely to deliver the tone (e.g., Pierrel & Sherman, 1963). Conditioned reinforcement is a valuable contributor to many behavior modification programs in that it enables clinicians to develop reinforcers which are convenient to deliver, will reinforce all members of a group, and which can be exchanged for a variety of biologically important stimuli. The most well-known of these procedures is the *Token Economy* (see Section 1.13.4.2.2).

(iv) Schedules of reinforcement

One of the important features of operant conditioning is that not every instance of a response needs to be reinforced in order to increase the frequency of the response. Intermittent reinforcement of this kind can be programmed on what are known as *schedules of reinforcement*, with responses being reinforced on the basis of time (e.g., time since the last reinforcement) or number (e.g., every *n*th response is reinforced). Basic schedules of reinforcement generate characteristic patterns of behavior which it is not necessary to elaborate on here (see Davey, 1989b; Ferster & Skinner, 1957).

However, intermittent reinforcement of the kind used in schedules of reinforcement does have one important quality—it produces robust responding that is significantly more resistant to extinction than when continuous reinforcement is used (cf. Ferster & Skinner, 1957). This is an important consideration when behavior modifiers are attempting to develop patterns and repertoires of behavior that will survive beyond the therapeutic conditions in which they are initially reinforced.

1.13.2.2.3 Special features of operant conditioning in humans

While it is clear that human subjects do learn about the relationship between their behavior and its consequences in simple operant conditioning procedures (e.g., Alloy & Tabachnik, 1984; Dickinson, Shanks, & Evenden, 1984), the variables that influence human operant performance appear to be more involved than those already implicated in the performance of nonhuman animals. For example, human operant performance is often rigid and insensitive to changes in the reinforcement contingencies. Some experimenters have encountered great difficulty in extinguishing a simple learned response in human subjects under experimental conditions (Bijou & Baer, 1966; Buchwald, 1959). Others have reported that human responding on schedules of reinforcement frequently entails erratic rates often interspersed with long and unpredictable pauses (Davey, 1989b; Lowe, 1979; Weiner, 1969). Lowe (1983) has argued that these apparently maladaptive aspects of human operant performance can to a large extent be explained by the fact that, unlike nonhuman animal performance, much human operant performance becomes rule-governed rather than contingency-governed. That is, when faced with any learning situation, human subjects will attempt to formulate hypotheses about the relationships between behavior and reinforcement. Very often, these initial hypotheses are wrong, especially in the artificial and simplistic environment of the operant conditioning laboratory. Lowe (1983) suggests that "it is the human subject's capacity to formulate their own descriptions of reinforcement contingencies and to use these descriptions to formulate rules to govern their behavior which results in human operant behavior being so different from that of lower animals" (p. 77). The conclusion, therefore, is that while humans are clearly able to learn about the relationship between their behavior and its consequences, the factors which translate this learning into performance are complex, and frequently involve processes not obviously observed in nonhuman performance.

1.13.3 HISTORICAL ASPECTS OF LEARNING THEORY AND CLINICAL PHENOMENA

1.13.3.1 Classical Conditioning

1.13.3.1.1 The "Little Albert" study and attempts at its replication

Attempts to explain human psychopathology in terms of classical conditioning processes date back to the "Little Albert" study reported by Watson and Rayner (1920). Watson and Rayner attempted to condition in 11-month-old Albert a fear of his pet white rat. They did this by pairing the pet rat (the CS) with a loud noise produced by striking an iron bar (the UCS). After several pairings of CS and UCS, Little Albert would begin to cry when the rat was introduced into the room (the CR) (see Figure 3). In the subsequent 80 years this study has drifted into psychological folklore, being quoted in most psychology textbooks (frequently with inaccurate details) as the prototypical example of a learned fear (Harris, 1979). Yet while it appeared to constitute a demonstration of a learned fear, Watson never attempted to formulate a theory of phobias based on classical conditioning (cf. Eysenck, 1979)—mainly because at the time there was nothing available to him that resembled a cohesive body of knowledge on classical conditioning itself. When critics subsequently came to analyse conditioning accounts of acquired fears, the Watson–Rayner approach could only be judged by evaluating the adequacy of the procedure, and this—not surprisingly—was nearly always found wanting.

Less well publicized than the Little Albert study are the subsequent attempts to replicate it. These studies almost invariably failed to reproduce the straightforward learning reported by Watson and Rayner and often reported other curious side effects.

English (1929) reported attempting to condition a fear in a 14-month-old girl. A wooden toy duck was used as the CS which was paired with a hammer blow to a large metal bar (UCS). Even after 50 pairings, English reported no evidence of a CR to the wooden duck. However, these negative results might be attributed to a weak UCS: English found that the UCS itself failed to evoke fear in the girl. In another child—who was reported to be frightened of the noise UCS—English does report conditioned fear to a stuffed black cat.

In another study, Bregman, a student of Thorndike (1935), paired a variety of objects (wooden shapes and colored clothes) with an electric bell (UCS). None of the objects used as CSs became a focus of fear, although the electric bell UCS was genuinely frightening to the infants used as subjects.

During these attempts to replicate the original Watson and Rayner study, Valentine (1930) observed that it seemed to be easier to condition fears to furry or leathery objects rather than more common household objects. English had reported that the child in his original study—although not showing fear to a wooden duck—did subsequently learn through conditioning to fear a pair of patent leather boots she had never seen before.

More recently, writers have revived these attempted replications as evidence against simple contiguity-based conditioning accounts of specific fears (e.g., Emmelkamp, 1982; Eysenck, 1979; Marks, 1987). Even the original Little Albert study has not gone uncriticized. Some reports suggest that Little Albert only showed fear to the rat when his comforting thumb was pulled out of his mouth, while others imply that Watson and Rayner never reported a replication although they must undoubtedly have attempted one.

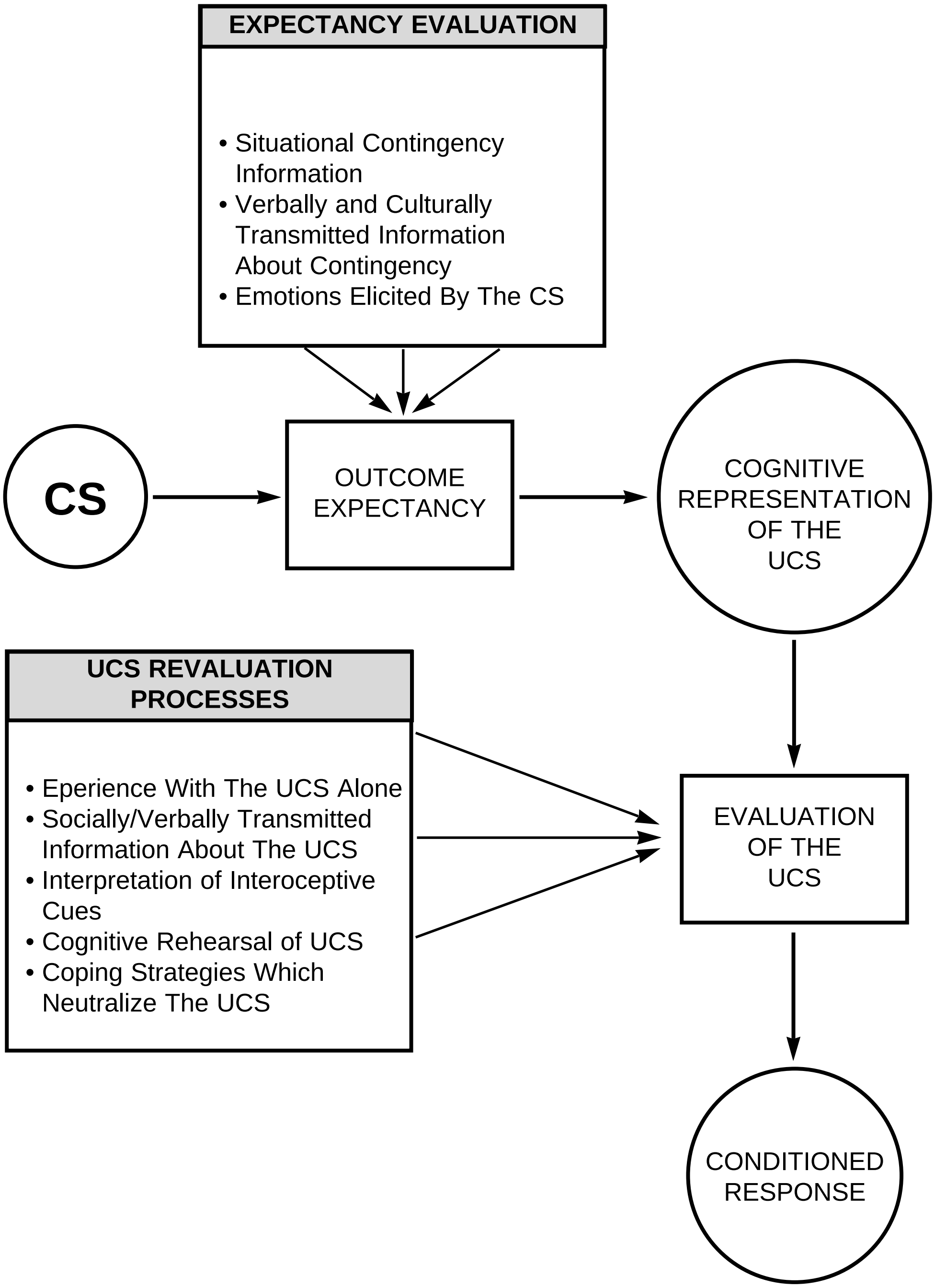
EXPECTANCY EVALUATION
• Situational Contingency Information
• Verbally and Culturally Transmitted Information About Contingency
• Emotions Elicited By The CS
CS
OUTCOME EXPECTANCY
COGNITIVE REPRESENTATION OF THE UCS
UCS REVALUATION PROCESSES
• Eperience With The UCS Alone
• Socially/Verbally Transmitted Information About The UCS
• Interpretation of Interoceptive Cues
• Cognitive Rehearsal of UCS
• Coping Strategies Which Neutralize The UCS
EVALUATION OF THE UCS
CONDITIONED RESPONSE

Figure 3 A schematic representation of a contemporary model of human classical conditioning (see text for further explanation) (*Phobias: Handbood of theory, research and treatment*, by G. C. L. Davey, 1997, Chichester, UK: Wiley. Copyright 1997 by Wiley. Reprinted with permission).

Nevertheless, despite criticisms of the Watson and Rayner study, it was a ground-breaking attempt to apply principles derived from experimental psychology to psychopathology. Many of the failures to replicate this seminal study have themselves been criticized on methodological grounds (e.g., Delprato, 1980), and the Watson and Rayner study is still the influential forerunner of contemporary conditioning models of fears and phobias (cf. Davey, 1997).

1.13.3.1.2 Pavlov and experimental neuroses

One of the most single important premises of the learning theory approach to psychopathology and therapy is that much of psychopathology can be considered as resulting from perfectly normal learning processes: that is, if an organism encounters a specific set of environmental contingencies then this will result in the learning of emotional reactions or dysfunctional responses. Thus, psychopathology was to be explained, not in terms of the dysfunction of any psychological or physiological process, but in the interaction between perfectly normally functioning learning processes and particular environmental contingencies.

Perhaps the oldest application of this view derives from Pavlov's own studies on experimental neuroses. In studies of discrimination learning he and his colleagues found that certain types of experimental procedures induced signs of anxiety and irritability in previously friendly and cooperative dogs (Pavlov, 1927, p. 29). In fact, the symptoms resembled those of behavior labeled as "neurotic" in humans. One particular procedure involved training the dog to salivate when a circle was presented on a screen. Following this, the subject was taught a discrimination between the circle (CS+) and an ellipse (CS−), with an initial ratio between the semi-axes of 2:1. The discrimination was learnt quite quickly and as training progressed the shape of the ellipse was changed until it was almost circle-like (a ratio between the semi-axes of 9:8). Kimble (1961) describes the subsequent change in the dog's behavior: "The hitherto quiet dog began to squeal in its stand, kept wriggling about, tore off with its teeth the apparatus for mechanical stimulation of the skin and bit through the tubes leading from the animal's room to the observer's ... On being taken into the experimental room the dog now barked violently ... In short, it presented symptoms of a condition which, in human beings, we would call neurosis" (p. 441).

Numerous other studies have described similar behavioral effects under conditions which require a very difficult discrimination (Masserman, 1943, 1950; Schneiderman et al., 1971). The effects are usually quite predictable and the phenomenon extremely persistent (Anderson & Parmenter, 1941).

This type of experiment served as a forerunner of animal conditioning studies which attempted to discover the environmental conditions underlying psychopathological phenomena. Influential others included the approach-avoidance conflict paradigm reported by Masserman (1943), and the learned helplessness procedure reported by Seligman (1976). Both described the systematic nature of environmental contingencies which generated behavior resembling human psychopathology, and both were influential in determining theoretical conceptions of neurosis and depression.

1.13.3.1.3 Two-factor theory of neuroses

The first attempt to construct a genuine conditioning theory of acquired fears was made by Mowrer (1947). This was based on Mowrer's two-factor theory of learning, and attempted to combine both classical and operant conditioning into a single model. Mowrer hypothesized that on encountering pairings between a CS and an aversive UCS, fear or anxiety became conditioned to the CS. Subsequent escape from, or avoidance of, the CS acted to reduce this anxiety or fear, and this process of fear reduction acted to operantly reinforce the avoidance or escape response and maintain the phobia. This model was particularly attractive because it appealed to specific accreditable learning principles, and also helped to explain what was known as the "neurotic paradox"—that is, why phobic avoidance was so persistent even in the absence of continued CS–UCS pairings.

While Mowrer's account was influential as the first integrated conditioning model of acquired fears (Eysenck & Rachman, 1965), it has generally failed to stand up to a number of theoretical and practical considerations. First, it does not explain why fear evoked by the CS does not extinguish rapidly with continued avoidance of the UCS. Second, there is little evidence for the hypothetical state of fear postulated to occur during the CS presentation and which is reduced as a result of avoidance. "Fear" does not appear to bear any simple relationship to its measures (McAllister & McAllister, 1971) and, indeed, both nonhuman animals and humans frequently show no physiological signs of fear during the presentation of an aversively conditioned CS (Brady, Kelly, & Plumlee, 1969; Davey & Arulampalam, 1982)—especially when an avoidance response is available (cf. Seligman, 1976). Nevertheless, attempts to understand phobic responding have still needed to explain what had been known as the "neurotic paradox," and more recent cognitive-based models of anxiety-based disorders have used Mowrer's two-factor theory as a basis for evolving more sophisticated models of both phobias (cf. Davey, 1997) and panic disorder (e.g., Salkovskis, 1991).

1.13.3.1.4 Eysenck's theory of incubation

Over a number of years Eysenck has developed a conditioning model of neuroses designed specifically to account for the fact that phobic reactions are often observed to increase in intensity with nonreinforced presentations of the CS (Eysenck, 1976, 1979; Eysenck & Kelley, 1987). This phenomenon is known as *incubation*, and while it is a common clinical phenomenon, it is clearly inconsistent with the conditioning principle of extinction which predicts a reduction in the intensity of the CR with successive nonreinforced presentations of the CS.

Eysenck attempted to resolve this paradox by suggesting that the CS comes to act as a partial substitute for the UCS. That is, once paired with an aversive UCS, the CS comes to evoke fear; on subsequent occasions when the CS is presented alone, the fear evoked by the CS can act as a potent reinforcer, thus generating an increasingly stronger CR over a series of CS-only presentations. Eventually, because of this process, the CR can become even stronger and more intense than the original UCR.

At the time, Eysenck's model was an attempt to explain one of the features of phobic responding which clearly did not fit comfortably into existing models of classical conditioning, and during the 1960s and 1970s represented a series of *ad hoc* models designed to explain individual anomalies in conditioning accounts of psychopathology. These individual *ad hoc* accounts have not proved to be enduring, but have been superseded by integrated conditioning models which account for a whole range of traditional criticisms leveled at conditioning models (cf. Davey, 1992, 1997; see Section 1.13.4.1.2).

1.13.3.1.5 The origins of behavior therapy

Behavior therapies based on conditioning principles have been shown to provide some of the most effective forms of treatment for a whole range of psychological problems (cf. Rachman & Wilson, 1980), but this begs the question of why conditioning principles were chosen to form the basis of an approach to applied human psychology. In this respect, it is instructive to trace the origins of behavior therapy and to pinpoint some of the events from which it developed. Krasner (1971) has outlined a number of these historical factors:

(i) the early conditioning studies of J. B. Watson and his subsequent influence on experimental psychology;

(ii) the field of operant conditioning, and in particular a seminal report by Lindsley, Skinner, and Solomon (1953) on conditioned bar-pulling in psychotics;

(iii) Wolpe's development of reciprocal inhibition techniques based on the classical conditioning research of Pavlov and Hull;

(iv) the work of Eysenck and colleagues at the Maudsley Hospital, UK, working within a Hullian framework;

(v) Dollard and Miller's (1950) attempts to interpret psychoanalysis into learning theory terms;

(vi) the development of behavioral and social learning alternatives to the traditional medical model of psychopathology; and

(vii) the failure of psychodynamic and psychoanalytic psychotherapies as indicated by internal dissatisfaction and widespread external critique.

In effect, these historical influences can be distilled to form just three good reasons why learning theory was to provide a basis for new psychotherapies. First, the writings of behaviorists J. B. Watson and later B. F. Skinner (1953) were to stress a belief in the continuity of psychological mechanisms between nonhuman animals and humans. If conditioning principles had been established with animals then—since humans were only quantitatively rather than qualitatively different from animals in learning abilities—these principles should also apply to humans. Second, there was a growing dissatisfaction in the late 1940s and 1950s with the medical or disease model of psychopathology. This was gradually being replaced by a "faulty learning" account which stressed the need to treat symptoms as *bona fide* behavioral problems rather than mere symptoms of underlying causes. Third, the growing belief in the greater efficiency and usefulness of concepts derived from experimental psychology. In particular, that branch of experimental psychology which rested on the rubric of behaviorism was most favored, primarily because the principles of behavior derived from behavioristic analyses readily suggested therapeutic action and had greater predictive power than their contemporary psychodynamic counterparts.

1.13.3.2 Operant Conditioning

1.13.3.2.1 Principles of the experimental analysis of behavior and their application to psychopathology

From its beginnings as the study of learning processes in animals (Ferster & Skinner, 1957; Skinner, 1938), operant conditioning developed as the basis for an entire philosophy of the understanding of human behavior. Pioneered by B.F. Skinner, operant reinforcement principles provided the basis for an entire analysis of human behavior, covering aspects of human behavior such as social behavior and human achievement, problem-solving, language development (Skinner, 1953, 1957, 1969, 1974), and eventually being touted as a form of behavioral technology which could shape and determine the nature of whole cultures (Skinner, 1971). This form of behaviorism has come to be known alternatively as radical behaviorism, the experimental analysis of behavior or simply behavior analysis, and the principles surrounding this approach to the explanation of human behavior are clearly ones which would have relevance to understanding and treating human psychopathology.

The radical behaviorist approach to psychopathology which gave rise to the treatments called behavior modification techniques (see Section 1.13.4.2) is based on a number of inter related assumptions:

(i) that human behavior is governed by the principles of learning which are embodied in operant and classical conditioning principles;

(ii) many human behaviors (including many of those which are diagnosable as psychological disorders) are acquired, maintained, and modified by these principles of learning;

(iii) since this approach emphasizes that the causes of behavior are located in the environment (in the form of reinforcement contingencies) and not "within" the individual (in the form of "motives," "intentions," etc.), these causes can be readily accessed and manipulated;

(iv) if the causes of behavior can be manipulated, then the behavior of the individual can also be manipulated in a controlled and predictable fashion.

In these assumptions were the beginnings of a technology of behavior, and this technology was first applied in the late 1950s and early 1960s to behavioral problems in the clinical setting (e.g., Aylion & Azrin, 1968). From these beginnings developed the broader field of behavior modification and techniques then evolved from single-client therapies based on classical conditioning principles to group therapy and management procedures based on operant conditioning. The introduction of operant principles into the field of therapy was an important step because operant psychology offered a wider range of therapeutic possibilities. First, it provided a method for changing complex and integrated behavior patterns in a way that classical conditioning procedures could not. Second, it allowed the introduction of group management procedures where therapy could be conducted in large groups and in settings which resembled those found in everyday life (e.g., Token Economy procedures, Kazdin, 1975, 1981; see Section 1.13.4.2.2). Third, it provided a framework in which those variables controlling behavior could be analyzed, and as a consequence of this analysis, more appropriate therapy could be devised. Sandler and Davidson (1973) emphasize the importance of a functional analysis of this kind: "The views expressed by Skinner and other operant theorists suggest that a better understanding of pathological conditioning can be accomplished by analysing the interactions between (a) the variables involved in an individual's behavior history, and (b) those determinants currently impinging upon the organism. With the knowledge of the former we can better predict how the latter will influence behavior" (Sandler & Davidson, 1973, pp. 63–64).

Since behavior analysts assume that psychological problems are acquired via normal learning processes, there is thus nothing pathological about them, and the removal of the problematic behavior thus removes the problem. This aspect of the traditional behaviorist approach has been controversial since it assumes that hypothetical underlying causes either do not exist or are irrelevant to successful therapy. Opponents to this view have argued that simply dealing with behavioral symptoms without addressing the basic underlying cause of the problem will result in "symptom substitution" (i.e., that the successfully treated symptom will simply be replaced by another). However, the evidence on this is equivocal. Some studies have found little evidence for symptom substitution following therapies based on learning principles (e.g., Kazdin, 1975; Rachman & Wilson, 1980), while some others have (e.g., Willems, 1974).

1.13.3.2.2 Operant reinforcement and psychopathology

Of all the applications of learning principles to psychopathology, perhaps the most obvious one is to suggest that some features of psychopathological behavior are developed and maintained because they have reinforcing consequences. They may either reduce anxiety (as in phobic avoidance), or they may help the individual to acquire certain things which are valuable or important to them (such as attention or approval).

The way in which attention, for example, can act as a reinforcer is well illustrated in a study reported by Williams (1959). This study reported the modification of tantrum behavior in a 21-month-old infant. The child had been seriously ill in very early life and one parent had always spent time at the child's bedside waiting for him to fall asleep. When a parent was not present the child cried. This crying behavior was virtually extinguished by allowing the child to cry without reinforcing it with attention. To underline the fact that this crying behavior was controlled by attentive consequences, Williams also reports that crying was reinstated by an aunt who unwittingly reinforced the behavior by re-entering the bedroom when the child cried.

A further study by Ayllon, Haughton, and Hughes (1965) compared explanations of bizarre psychotic behavior patterns given by learning and psychodynamic theories of psychopatholgy. They reinforced a female schizophrenic resident in a psychiatric hospital for carrying a broom. Whenever she was observed holding the broom a nurse would approach her, offer her a cigarette, or give her a token which could be exchanged for a cigarette. Eventually, when this behavior was established, it was transferred from a continuous to an intermittent reinforcement schedule until the patient was carrying the broom around for a considerable part of the day. It was at this point that Ayllon et al. called in two psychodynamic therapists (who were unaware of the reinforcement schedule) to give their opinions on the nature of this behavior. One of them gave the following reply: "Her constant and compulsive pacing, holding a broom in the manner she does, could be seen as a ritualistic procedure, a magical action ... Her broom would be then: (1) a child that gives her love and she gives him in return her devotion, (2) a phallic symbol, (3) the sceptre of an omnipotent queen ... this is a magical procedure in which the patient carries out her wishes, expressed in a way that is far beyond our solid, rational and conventional way of thinking and acting" (Ayllon et al., 1965, p. 3).

There are two points to be made here. First, the description given by the psychodynamic therapist may well represent what goes on in the patient's head—we do not know for sure—but it does not in any way reflect the process by which the behavior was acquired. Second, although Ayllon et al. have systematized the acquisition process in this example, it seems reasonable to suppose that contingencies of this kind could be unwittingly set up quite frequently. Although this study does caricature the process of acquisition, it certainly implies that operant reinforcement can play a powerful role in establishing maladaptive, inappropriate, or bizarre behavior patterns.

1.13.4 CURRENT APPLICATIONS OF LEARNING THEORY IN CLINICAL AND EXPERIMENTAL PSYCHOPATHLOGY

1.13.4.1 Classical Conditioning

1.13.4.1.1 Traditional criticisms of classical conditioning models of phobias

The discussion of the Little Albert experiment of Watson and Rayner (1920) described in Section 1.13.3.1.1 provides a basic framework for a possible model of fears and phobias based on classical conditioning processes. Watson and Rayner never explicitly articulated the principles of this conditioning model, although the implication was that it was a model based on findings from animal learning studies, and in particular the acquisition of conditioned fear to a CS based on the contiguous pairing of that CS with an aversive UCS. In the subsequent 60 years there have been many criticisms of the classical conditioning model of phobias, and they are ones which primarily address a contiguity-based model of classical conditioning (e.g., Emmelkamp, 1982; Rachman, 1977). It is worth detailing these criticisms, so that they can be assessed in the context of more recently developed contemporary conditioning models of fears and phobias.

First, many phobics appear unable to recall any trauma or aversive conditioning experiences at the time of the onset of their phobia (Emmelkamp, 1982; Rachman, 1977). This appears to be particularly true of some animal phobics such as snake or spider phobics (Davey, 1992c; Murray & Foote, 1979), and also height phobics and water phobics (Menzies & Clarke, 1993a, 1993b).

Second, not all people who experience pain or trauma (UCS) paired with a situation (CS) develop a phobia. For example, not everyone who has a traumatic experience undergoing dental treatment acquires a dental phobia (Lautch, 1971), not everyone who experiences a violent thunderstorm acquires a thunderstorm phobia (Liddell & Lyons, 1978), and not all fliers who experience a traumatic flying accident express a subsequent anxiety of flying (Aitken, Lister, & Main, 1981; Goorney, 1970). A simple contiguity-based conditioning model does not appear to have the power to predict when an individual will acquire a phobia and when they will not.

Third, a simple incremental–decremental model of conditioning (where pairing a CS with an aversive UCS produces an increment in fear to the CS, and an unreinforced presentation of the CS produces a decrement in fear) does not appear to account for the common clinical phenomenon of incubation of fear (see Section 1.13.3.1.4). While incubation is a common clinical phenomenon (Eysenck, 1979), it is rarely found in laboratory analogue studies (cf. Richards & Martin, 1990). Incremental–decremental models of phobias would predict a successive decrease in fear with successive nonreinforced CS presentations (extinction), not an increase in fear as found with incubation.

Fourth, simple conditioning models treat all stimuli as equally likely to enter into association with aversive consequences, yet fears and phobias are not evenly distributed across stimuli and experiences. People appear to develop phobias of animals (snakes, spiders), heights, water, death, thunder, and fire more readily than fear of hammers, electricity outlets, knives, guns, and so on, even though the latter group of stimuli seem to have a high likelihood of being associated with trauma (Agras, Sylvester, & Oliveau, 1969; Kirkpatrick, 1984; Seligman, 1971). This uneven distribution of fears appears to violate the Pavlovian principle of equipotentiality, which states that all stimuli should be equally capable of entering into associative relationships with a consequential UCS.

Fifth, many surveys suggest that a substantial percentage of phobics appear to acquire their fear through observational learning rather than direct experience with trauma (Menzies & Clarke, 1993a, 1993b; Öst & Hugdahl, 1981; Rachman, 1977). This has generally been conceived of as a route to phobias that is an alternative to direct conditioning.

The next section argues that while these criticisms are valid in terms of traditional contiguity-based classical conditioning models, they are significantly less damaging for more recently developed contemporary cognitive conditioning models of phobias.

1.13.4.1.2 Contemporary cognitive models of human classical conditioning

Contemporary models of human classical conditioning differ from their contiguity-based predecessors in a number of important ways.

First, unlike their behaviorist forerunners, modern learning theories are happy to incorporate cognitive conceptualizations of the conditioning process. Examples of such kinds of incorporations include the willingness to attempt to discover the exact nature of the associations that are formed during conditioning—even though they cannot be directly observed (see Section 1.13.2.1.4 (vi)).

Second, it is now well accepted that many factors other than the experienced pairing of the CS and UCS can affect the strength of the association between the CS and UCS. In the case of humans, these include verbally and culturally transmitted information about the CS–UCS contingency (e.g., Dawson & Grings, 1968; Wilson, 1968), existing beliefs and expectancies about the possible consequences associated with a particular CS (Davey, 1992; Honeybourne, Matchett, & Davey, 1993), and emotional reactions currently associated with the CS (e.g., Davey & Dixon, 1996) (see Section 1.13.2.1.5 (iv)).

Third, and arguably of most importance in this context, is the finding that the strength of a CR can be radically influenced, not just by the strength of the CS–UCS association, but also by the way in which the individual evaluates the UCS (Davey, 1989c, 1992a). In humans, there are a variety of processes that can influence the evaluation of the UCS and some of these have been described in Section 1.13.2.1.5 (iii).

Figure 3 provides a schematic representation of a contemporary cognitive model of human classical conditioning. This illustrates the kinds of factors that may influence the strength of an association between CS and UCS (Expectancy Evaluations), and also how the UCS representation's evocation of a fear CR will be influenced by how the UCS has been evaluated or revalued (UCS Revaluation Processes).

1.13.4.1.3 How contemporary conditioning models of phobias account for the traditional criticisms

(i) Inability to recall trauma at the onset of the phobia

Many phobics appear unable to recall any trauma at the time of the first appearance of their phobia. Nevertheless, the contemporary conditioning view conceives of acquisition of the CS–UCS association and modulation of the aversiveness of the UCS as relatively independent processes (Figure 3). Thus, it is quite conceivable for an individual to learn an association between a CS and UCS when the UCS is relatively nonaversive, and then subsequently (through the process of UCS inflation—see Section 1.13.2.1.5 (iii)) to have the aversiveness of the UCS inflated. This process would eventually generate a fear response to the CS without that CS ever having been explicitly paired with an *aversive* UCS.

White and Davey (1989) demonstrated this process in a laboratory analogue, and then described a putative "real-world" analogue of the following kind. An individual may witness an unknown person die of a heart attack on a bus or a train; on future occasions, riding on public transport may evoke memories of this incident, but no anxiety. Subsequently, however, that individual may be present when a close friend or relative dies of a heart attack, thus inflating the aversive properties of heart attacks. This may then give rise to acute anxiety when riding on public transport. Davey, De Jong, and Tallis (1993) report a number of actual case histories in which this type of scenario fits the etiology of a variety of anxiety disorders, including phobias, panic disorder, and obsessive-compulsive disorder. What these findings suggest is that a conditioning account of anxiety disorders is no longer bound by the need to discover contiguous stimulus-trauma experiences in the etiologies of individual cases.

(ii) Not all people who experience pain or trauma develop a phobia

There are a number of processes by which contemporary conditioning theory would predict that encounters with traumatic stimuli would not generate acquired fears.

First, the process of latent inhibition (see Section 1.13.2.1.4 (iii)) allows for the failure to develop a CR following pairings of a CS and UCS. Davey (1989a) tested this possibility by investigating the role of latent inhibition in the acquisition of dental phobias. He found that individuals who reported having a painful dental experience but did not acquire a dental phobia reported having their first painful experience significantly later in the dental treatment careers than individuals who did acquire a dental phobia. Thus, consistent with the operation of latent inhibition, individuals who did not acquire a phobia had the equivalent of a series of CS alone trials prior to the first pairing of the CS with a traumatic UCS.

Second, the process of UCS revaluation permits the aversiveness of any experienced UCS to be devalued following an aversive conditioning episode. There are a number of processes that have been identified as contributing to UCS reevaluation (see Section 1.13.2.1.5 (iii)), and effective devaluation of the UCS will result in the failure of the CS to evoke a substantial fear CR. Some empirical support is lent to this hypothesis by the fact that individuals with clinically diagnosed specific phobias and panic disorder are significantly less likely to report using trauma devaluation strategies (see Section 1.13.2.1.5 (iii)) than a sample of normal control subjects (Davey, Burgess, & Rashes, 1995).

(iii) Conditioning models do not appear to account for the clinical phenomenon of incubation

When a CS is presented alone on several occasions following conditioning, extinction should occur (see Section 1.13.2.1.4 (ii)). However, in many phobic circumstances, presentation of the CS alone leads paradoxically to incubation of the fear (Eysenck, 1979). Nevertheless, contemporary conditioning theories are able to describe some circumstances in which CS-alone presentations following conditioning may lead to increases rather than decreases in the strength of the CR. This has again been linked to the operation of UCS revaluation processes (Davey, 1997; Davey & Matchett, 1993). Davey and Matchett found that when subjects in a conditioning experiment were asked to cognitively rehearse the aversive UCS for a period following acquisition trials and prior to a series of extinction CS-alone trials, subjects who were either high in trait anxiety levels or who had an induced anxious mood subsequently showed *increased* fear CRs to CS presentations during extinction. Davey and Matchett argued that this effect was the result of UCS rehearsal in anxious individuals inflating the aversive evaluation of the UCS, and thus causing an increased strength CR on future CS presentations. This process is analogous to the rumination about the consequences of contact with phobic stimuli that is often reported in individuals with anxiety disorders, and may provide at least a partial explanation of clinical incubation effects.

(iv) Fears and phobias are not evenly distributed across stimuli and experiences

There appears to be an uneven distribution of fears, with some stimuli being very common foci for phobias (e.g., snakes, spiders, heights, water), and other—equally potentially dangerous stimuli—only rarely being reported as phobic stimuli (e.g., guns, hammers, electricity outlets).

The traditional explanation of this uneven distribution of fears has been a variant of conditioning theory called *preparedness theory* (Seligman, 1971). This account hypothesizes that stimuli which tend to become the focus for fears and phobias are those which have been hazardous for our pretechnological ancestors (e.g., poisonous snakes and spiders, dangerous heights, etc.). This has resulted in the evolution of phylogenetically-based predispositions to associate biologically "prepared" stimuli with aversive outcomes.

Support for the preparedness hypothesis has come from:

(a) Laboratory conditioning studies which have paired "prepared" CSs (such as pictures of snakes and spiders) with aversive UCSs such as electric shock (e.g., Öhman, Dimberg, & Öst, 1985; Öhman, Frederickson, & Hugdahl, 1978). These studies have demonstrated a greater resistance to extinction in "prepared" stimuli than "unprepared" stimuli (McNally, 1987).

(b) Studies of the observational learning of fear of snakes in primates. These studies have indicated that monkeys who are not initially afraid of snakes will rapidly acquire an intense fear when they have watched a wild-reared monkey behaving fearfully in response to a toy snake, but will regularly fail to acquire fear towards artificial flowers or a toy rabbit using the same paradigm (Cook & Mineka, 1987, 1989, 1990; Cook, Mineka, Wolkenstein, & Laitsch, 1985; Mineka, Davidson, Cook, & Weir, 1984).

However, Davey (1995) has critically reviewed much of the evidence that is often used to support the biological preparedness hypothesis. Among the problems identified with this evidence are:

(a) that the evidence from clinical studies largely fails to support predictions from preparedness theory (de Silva, 1988; de Silva, Rachman, & Seligman, 1977); clinical phobias identified as "prepared" rarely exhibit the resistance to extinction expected of such fears,

(b) the failure of the primate studies of Mineka and colleagues to use a balanced design; as a result these studies are open to explanations other than biological preparedness, and

(c) the finding that ontogenetic fear-relevant stimuli (such as guns, electricity outlets, etc.) often display similar conditioning characteristics to so-called "biologically prepared" stimuli (such as snakes and spiders) (Davey & Dixon, 1996; Honeybourne, Matchett, & Davey, 1993; McNally & Heatherton, 1993).

As an alternative to biological preparedness, Davey (1992b, 1995) has argued that in laboratory preparedness experiments, human subjects begin the conditioning procedure with an inflated estimate of the probability of fear-relevant stimuli being followed by aversive UCSs (Davey, 1992b). Details of some of the evidence supporting this view are presented in Section 1.13.2.1.5 (iv). Subsequent studies have identified some of the factors that give rise to this UCS expectancy bias. These include estimates of the "dangerousness" of the CS (Davey & Craigie, 1997; Davey & Dixon, 1996), judgements about the semiotic similarity of the CS to the potential aversive UCS (Davey & Dixon, 1996; Hamm, Viatl, & Lang, 1989), and existing prior fear to the CS (Diamond, Matchett, & Davey, 1995; Davey, & Dixon 1996). Interestingly, nearly all of these studies have failed to find any significant differences between phylogenetic and ontogenetic fear-relevant stimuli in the determinants of UCS expectancy bias. This suggests that the same underlying process may mediate expectancy bias to both types of stimuli.

To summarize, there are at least two contrasting conditioning-based theories which attempt to explain the uneven distribution of fears. Preparedness theory is essentially an evolutionary-based account which argues that many phylogenetic fear-relevant stimuli are biologically "prepared" to enter into associations with aversive outcomes (Seligman, 1971). Alternatively, the UCS expectancy bias account says that the selective conditioning of certain stimuli results from biases in the way that individuals process information concerning fear-relevant stimuli; these biases create a heightened expectancy of aversive outcomes following such stimuli, and hence such stimuli readily enter into association with aversive outcomes (Davey, 1995).

(v) Indirect associative routes to phobias

Some writers have suggested that, rather than being caused by specific direct conditioning experiences, the majority of phobias are caused by vicarious learning in the form of either observational learning (watching someone else being frightened in the presence of the phobic stimulus) or verbally transmitted information about the phobic stimulus and its potentially threatening consequences (Öst & Hugdahl, 1981; cf. Rachman, 1977). However, contemporary conditioning models make it clear that associations can be formed between a stimulus (CS) and its outcome (UCS) in a variety of ways—not just by direct conditioning experiences (see Section 1.13.2.1.5 (iv)). An implication of this is that an association between a CS and UCS can be learned as a result of information about the contingency, or through observing someone else experiencing the contingency. In terms of contemporary conditioning models, this is still conditioning. Admittedly, there is very little information at present on the strength and persistence of associations that are learned in these vicarious manners. They may be weaker and less resistant to extinction than associations acquired through direct experience, but they do appear to be mediated by the same associative mechanism that mediates conditioning through direct experience (Mineka & Cook, 1993). However, evidence that is available does suggest that direct conditioning experiences are more memorable than vicarious learning experiences, and that a majority of strong fears are attributed by individuals to direct conditioning rather than vicarious learning (Merckelbach, de Ruiter, van den Hout, & Hoekstra, 1989; Withers & Deane, 1995).

(vi) Summary

Section 1.13.4.1.3 has described how contemporary models of conditioning have contributed to our understanding of anxiety-based clinical disorders. Contemporary conditioning models differ from their more traditional counterparts in that they explicitly contain both an associative element and an element which allows the anxiety or fear CR to be modulated through revaluation of the UCS. Such conditioning models provide a systematic framework for the understanding of many features of anxiety-based responding.

1.13.4.1.4 *Further applications of classical conditioning to anxiety-based disorders*

While classical conditioning models have primarily been applied to the understanding of phobic responding, conditioning models are increasingly being applied to an understanding of the etiology of other anxiety disorders, and, indeed, are being integrated into broader cognitive accounts of psychological disorders. At a specific level, classical conditioning models have been postulated to account for some of the important features of panic disorder (Goldstein & Chambers, 1978; Wolpe & Rowan, 1988) and post-traumatic stress disorder (PTSD) (Fairbank & Brown, 1987; Keane, Zimering, & Caddell, 1985; Kilpatrick, Veronen, & Best, 1985). However, aspects of associative learning are becoming increasingly integrated into broader cognitive models of psychopathology—especially those models that postulate the existence of associative networks underlying the expression and maintenance of fear and anxiety (e.g., Brewin, 1988, 1989; Foa & Kozak, 1986), and the revaluation of information related to threat (e.g., Brewin, Dalgleish, & Joseph, 1996). Thus, while conditioning models continue to be valuable at the level of specific phenomena, the processes contained in contemporary models of conditioning are themselves having increasing influence on the conceptualizations of broader-based cognitive models of psychopathology.

1.13.4.1.5 *Classical conditioning in behavior therapy*

Ever since Wolpe (1958) published what was arguably the seminal text on behavior therapy, a variety of therapeutic techniques have been developed which have their origins in classical conditioning principles. Many of these methods have since been modified and supplemented, but almost all owe their origins to the Pavlovian principle of extinction. The assumption was that if emotional disorders were learned through a process of classical conditioning, then those disordered responses could be eliminated through applying procedures which effectively disrupted the learned CS–UCS association and extinguished the learned emotional response. The most famous of such techniques are flooding, counterconditioning, and systematic desensitization. Wolpe's main contribution was to introduce the notion of *reciprocal inhibition*, by which the emotional response is eliminated not just by extinction, but also by attaching a response to the emotion-eliciting stimulus which is incompatible with the pathological emotion.

There has been a tendency to assume that such techniques are applicable only to the treatment of anxiety and phobic disorders (e.g., Bandura, 1969; Paul & Bernstein, 1973), but they have in fact been applied successfully to a wider range of disorders including sexual dysfunction (e.g., Kockott, Dittmar, & Nusselt, 1975; Mathews et al., 1976), marital conflict (e.g., Jacobson & Weiss, 1978), and addictive disorders (e.g., Lichtenstein & Danaher, 1976; O'Leary & Wilson, 1975).

Such techniques have, over the years, been subjected to rigorous tests of their internal and external validity (cf. Rachman & Wilson, 1980), with results convincingly demonstrating that such techniques are successful because of the principles on which they are based, and superior in outcome to many other differing types of psychotherapy.

(i) Flooding

Flooding is an extinction-based therapy procedure involving therapist-directed prolonged exposure to the anxiety-eliciting stimulus or situation. Flooding can be conducted *in vivo* or *in vitro*, the latter by asking the client to imagine extended contact with their phobic stimulus. It is not necessary for high levels of fear to be elicited during the flooding experience, and mere exposure appears to be the sufficient condition (Hafner & Marks, 1976; Mathews, 1978). This would be consistent with the extinction rationale on which the method is based, and inconsistent with more psychoanalytically-oriented versions of exposure therapy (such as implosion therapy) which demand that high levels of fear are elicited during exposure in order for treatment to be effective.

Studies have demonstrated that flooding is significantly more effective than placebo treatments, indicating that its success is not simply the result of client expectations (Gelder et al., 1973).

(ii) Counterconditioning

Counterconditioning is also a procedure based on Pavlovian extinction principles, but as well as extinguishing the emotional CR it develops an acceptable alternative CR. The procedure is illustrated well in an animal analogue study by Klein (1969). Initially rats were trained on a discriminated avoidance task to run between two compartments in a shuttlebox. This response was subsequently extinguished, but during the extinction procedure the rats were divided into three groups: one group was confined to the compartment where shock had been delivered and was also given food in this compartment (the counterconditioning group); animals in the second group were simply confined in the compartment without food (flooding); and members of the third group were not confined at all (extinction). The results suggested that the counterconditioning procedure was the most effective in eliminating the conditioned response.

A variation of the counterconditioning procedure is known as systematic desensitization by reciprocal inhibition (Wolpe, 1958).

(iii) Systematic desensitization

In this variant of counterconditioning, the therapist constructs a ranked list of events or stimuli to which the client reacts with increasing fear or anxiety (the fear hierarchy). The client is then trained to relax while at the same time being exposed to the stimulus condition at the bottom of the fear hierarchy. When the client feels quite relaxed in this situation, they then progress on to the next most fear-inducing event on the hierarchy (Wolpe & Lazarus, 1966). This technique involves extinction of the anxiety response through graduated exposure to the anxiety-eliciting stimuli, and counterconditioning of a response incompatible with anxiety (relaxation) to these stimuli (the principle of reciprocal inhibition).

Systematic desensitization has arguably been one of the most durable and successful of all the behavior therapy procedures. Studies have shown that there is no methodologically acceptable evidence to prove that the therapeutic benefits derived from systematic desensitization can be attributed to nonspecific treatment effects such as placebo influences and expectations of therapeutic improvement (Gelder et al., 1973; Kazdin & Wilcoxon, 1976; Wilson, 1973), and in comparative outcome studies systematic desensitization has generally been shown to be more effective than both no treatment conditions and almost every other psychotherapy variant with which it has been compared (cf. Leitenberg, 1976; Rachman & Wilson, 1980).

(iv) Aversion therapy

A rather different use of classical conditioning principles is found in aversion therapy. This is where the conditioning paradigm is used to condition an aversive response to a formally attractive stimulus. Aversion therapy is most widely used in the treatment of addictive behaviors such as alcoholism, and aversive UCSs that have been used include electric shock and drugs (such as emetine) that induce unpleasant physiological reactions (cf. Rachman & Teasdale, 1969; Wilson, 1978). Some original studies by Voegtlin and Lemere (1942) and Lemere and Voegtlin (1950) serve as examples of this method with alcoholics. In their procedure alcoholic patients were given injections of emetine or apomorphine, which quickly elicit both nausea and vomiting (UCS). Immediately prior to vomiting the patient is given a drink of their favorite alcoholic beverage (CS). The success of this procedure depends on whether the CS (alcohol) comes to elicit a CR of nausea. Some early studies of this type of conditioning therapy report relatively successful treatment of addictive disorders such as alcoholism. For example, Wiens, Montague, Manaugh, and English (1976) reported that 12 months after receiving aversion therapy, 63% of 261 treated patients were still abstaining. Nevertheless, treatment programs such as aversion therapy do appear to rely for their long-term success on related support programs such as community reinforcement (e.g., Azrin, 1976), and there has been little success in demonstrating that aversion therapy using electric shock UCSs is any more effective than placebo control conditions (Hedberg & Campbell, 1974; Wilson, 1978; Wilson, Leaf, & Nathan, 1975). However, certain types of aversion therapy do have a role in the treatment of specific disorders, especially addictive disorders, but it is a role that is best embedded in a multifaceted approach to therapy.

(v) Recent developments in the treatment of anxiety-based disorders

While classical conditioning principles have contributed substantially to the development of exposure-based therapies, the last 10–15 years have seen rapid and important changes in the nature of therapies for anxiety-based disorders. Arguably the two most important aspects of change have been (a) the emergence of cognitive therapies, and (b) the development of treatment packages addressed to specific disorders. These changes have meant that while incorporating some elements of traditional conditioning principles in their procedures, contemporary treatment packages are considerably more multifaceted than their behavior therapy predecessors, and their primary objective is to change aspects of the cognitions maintaining the disorder (e.g., Beck, 1976; Öst, 1997).

Nevertheless, many of these contemporary procedures retain exposure as the central element in the treatment (e.g., Booth & Rachman, 1992; Öst, 1989), although what may have changed is the rationale for retaining exposure. Traditional behavior therapies such as flooding and systematic desensitization included exposure because of its role in facilitating extinction of any associative link between the phobic stimulus (CS) and fearful outcomes (UCS). However, the perceived role of exposure appears to have changed somewhat to one that either disconfirms dysfunctional beliefs about the phobic stimulus (e.g., Öst, 1997; Salkovskis, 1981) or facilitates the acquisition of behaviors designed to help the individual cope with contact with the phobic stimulus or situation (e.g., "guided mastery," Williams, Turner, & Peer, 1985).

Thus, while traditional classical conditioning principles can be seen as the seminal influence in the development of the basic behavior therapy treatments for emotional disorders, their direct influence in contemporary therapies is limited. However, what has yet to be exploited is the therapeutic value of contemporary conditioning models (see Section 1.13.4.1.2). These recent models have cognitive components (such as UCS revaluation processes) which are compatible with modern cognitive conceptions of therapy and which seem highly suited to therapeutic application; however, this potential has yet to be explored.

1.13.4.2 Operant Conditioning

Operant conditioning has contributed extensively to modern-day clinical psychology, both at the level of behavioral assessment and specific techniques. This section covers some of the more influential of these applications.

1.13.4.2.1 Functional analysis

One of the important by-products of the study of operant conditioning was the discovery that behavior could be understood in terms of the consistent relationships between behavior and the environment. For example, rats in a Skinner box would learn to press a lever because that act had a consistent relationship with the consequential delivery of food. This identification of consistent relationships between behavior and environmental events led to the development of the behavioral assessment technique known as functional analysis. This is currently an assessment technique that is widely used in clinical psychology to understand aberrant behavior and also to facilitate effective therapy.

While functional analysis has its roots in operant conditioning, it has now developed into an assessment technique that is considered to be theoretically neutral (Owens & Ashcroft, 1982). As a result there is often dispute about how functional analysis should be defined. A theoretically neutral definition would be one that emphasizes the functional relationships between behavior and environment. Haynes and O'Brien (1990), for instance, define functional analysis as "The identification of important, controllable, causal functional relationships applicable to a specified set of target behaviors for an individual client" (p. 654). This definition emphasizes that the relationships identified should be controllable—otherwise they cannot be manipulated during the course of treatment, and that functional analysis is idiographic (addressing causal relationships for behavior problems of individual clients) rather than nomothetic (addressing causal relationships for a behavior problem across clients).

Functional analysis has two specific roles in the treatment process. First, it allows the therapist to identify reinforcers that may be maintaining aberrant behavior. As a result the therapist may then be able to manipulate those reinforcement contingencies in such a way that will eradicate the aberrant behavior. For example, self-injurious behaviors or challenging behaviors may be maintained by any number of reinforcing consequences such as seeking attention, being left alone, or sensory stimulation. Identifying exactly the type of consequence that reinforces such behavior allows the therapist to disrupt the contingency and reduce the frequency of the behavior through extinction (e.g., Mazaleski, Iwata, Vollmer, Zarcone, & Smith, 1993; Wacker et al., 1990), or even through presenting reinforcers noncontingently (e.g., Hagopian, Fisher, & Legacy, 1994; Vollmer, Iwata, Zarcone, Smith, & Mazaleski, 1993). Second, a functional analysis is important because it improves the therapist's ability to develop an effective behavioral intervention. For example, attempting to eradicate aberrant behaviors by reinforcing appropriate ones is less likely to be successful if the reinforcement contingencies maintaining the aberrant behaviors have not been identified and are still operating (Vollmer & Smith, 1996). This effectively means that existing reinforcement contingencies are competing with the reinforcement contingencies used to develop the appropriate new responses.

Functional analysis has been adopted across a range of clinical settings as an aid to effective treatment. These applications include aggressive/challenging behavior (O'Reilly, 1995; Samson & McDonnell, 1990), tantrums (Darby et al., 1992), pica (Mace & Knight, 1986), stereotypy (Mace, Browder, & Lyn, 1987), attention deficit hyperactivity disorder (Northrup, Broussard, Jones, George, Vollmer, & Herring, 1995), depression (Ferster, 1985), anorexia nervosa

(Slade, 1982), and self-injurious behavior (Iwata, Dorsey, Slifer, Bauman, & Richman, 1985).

1.13.4.2.2 Token economy schemes

One great advantage of operant conditioning principles is that they are more readily adaptable to group therapy and group management situations than principles derived from classical conditioning, and perhaps the first group management procedure of this kind was the token economy (Ayllon & Azrin, 1968). In this type of program the participants receive tokens (a generalized reinforcer) when they have engaged in appropriate behavior and they can, at some later time, exchange these tokens for a variety of desired items. In the psychiatric setting, these might include access to the hospital grounds, preferred consumables, etc. The token acts as a generalized conditioned reinforcer and can be delivered with the minimum of delay after the required response is emitted. Its primary use in clinical psychology is as a therapeutic program with psychiatric patients (usually inpatients, but it has increasingly been used in community programs, cf. Corrigan, 1991). Behaviors normally fostered in these programs are prosocial or self-care ones (e.g., combing hair, bathing, brushing teeth, etc.) (Ayllon & Azrin, 1965). With more withdrawn patients, tokens can be administered simply to encourage socialization. Such programs help to increase the sociability of institutionalized individuals, strengthen behaviors that are likely to be needed in life outside the institution, and help to keep the individual receptive to other, more specialized, therapies. As a technique for behavior modification, the token economy has a number of advantages: (i) it does not necessarily need professionally trained therapists to appropriately dispense tokens; (ii) such programs prevent the deterioration of normal social and self-care behaviors on admission to an institution such as a psychiatric hospital (Dickerson, Ringel, & Parente, 1994; Kazdin & Bootzin, 1972; Lippman & Motta, 1993); and (iii) they can be used in any group situation, either as a therapeutic or a management program (Kazdin, 1975, 1981).

A number of studies have demonstrated that token economies can have significant therapeutic gains. For example, Gripp and Magro (1971) showed that schizophrenic patients in a token economy ward improved significantly more than patients in a traditional ward. Herson, Eisler, Alford, and Agras (1973) found that depressed patients exhibited less depressive behavior while participating in a token economy. Gershone, Errickson, Mitchell and Paulson (1977) found that patients in a token economy scheme were better groomed, spent more time in activities and less time in bed, and made fewer disturbing comments than patients on a traditional ward. Patients on token economy schemes also earn discharge significantly sooner than patients who are not on such a scheme (Hofmeister, Scheckenbach, & Clayton, 1979). A large-scale study by Paul and Lentz (1977) compared token economy programs with milieu therapy and traditional custodial care approaches. The results of this comprehensive four-and-a-half year study showed that participation in both the token economy and milieu therapy groups was followed by major improvements at the end of treatment. However, the token economy group produced significantly greater reductions in bizarre behavior and increases in adaptive behavior such as self-care and interpersonal social skills than the milieu therapy. Assessment of overall functioning on a number of standardized scales showed that the token economy program resulted in significantly greater improvement than the milieu therapy at every six-month evaluation.

Nevertheless, despite these positive findings, recent surveys indicate that the use of token economies in clinical settings is in serious decline (Boudewyns, Fry, & Nightengale, 1986; Corrigan, 1995; Hall & Baker, 1973). The reasons for this are not entirely clear given the demonstrated efficacy of the technique as a therapeutic process. However, a number of reasons have been put forward for this decline.

First, there are legal and ethical issues which need to be considered. This is especially so when decisions have to be made about who will participate in token economies, for how long, and what will be made available as positive reinforcers. Legislation over the past 25 years has sought to protect patients" rights, and treatment staff are severely constrained with regard to the use of more basic items as reinforcers (Glynn, 1990)—especially when patients now have a legal right to their own personal property, humane treatment including comfortable bed, chair, bedside table, nutritious meals, cheerful furnishings, and so on.

Second, one of the major challenges for token economies has been maintenance and generalization of therapeutic effects. To the extent that patients can obtain reinforcers outside the program and avoid punishment by exiting from the program, the therapeutic benefit of token economies becomes less useful (Glynn, 1990). It is true that some studies have shown that behaviors targeted for improvement in a token economy scheme return to low baseline levels outside of the program (e.g., Ayllon & Azrin, 1968; Walker & Buckley, 1968). However, there are other studies that have shown positive effects of maintenance and generalization (Banzett, Liberman, & Moore, 1984). Nevertheless, it should be pointed out that generalization is not a passive process, and clinicians must actively build into the program strategies that transfer positive effects to settings outside the treatment scheme (Stokes & Baer, 1977; Stokes & Osnes, 1988).

Third, some other proponents of the token economy have argued that its decline has been the result of unfounded misconceptions about the nature and efficacy of such programs (e.g., Corrigan, 1995). These include such misconceptions as token economies not being therapeutically effective, their benefits do not generalize, they do not provide individualized treatment, they are abusive and cohersive, and they are not practical to implement in the context of present-day attempts to treat patients in the community. Corrigan (1995) argues that these are all unfounded, and that the token economy remains an important and valuable tool for the management of patients and staff in treatment settings.

1.13.4.2.3 Response shaping

As noted in Section 1.13.2.2.2 (ii), response shaping is a useful procedure for strengthening rarely-occurring behaviors or building up complex response repertoires, and this method is utilized regularly in behavior modification programs. An early study by Isaacs, Thomas, and Goldiamond (1960) serves to illustrate this method. They attempted to reinstate verbal behavior in a psychiatric inpatient who had been mute for over 19 years. In this example, the target behavior occurs relatively infrequently, and thus has to be approached via the reinforcement of successive approximations to the behavior. They discovered that although the patient was withdrawn, he did appear to respond to chewing gum, which they considered would act as an effective reinforcer. They then broke down the target behavior so that it could be reached by reinforcing a series of approximations to verbal behavior. The first responses to be reinforced were fairly simple, discrete responses whose baseline levels were high enough for them to occur spontaneously within a training session. The shaping program went as follows.

(i) When the patient moved his eyes towards the chewing gum, he was reinforced by being given the gum; after 2 weeks the probability of this response was relatively high.

(ii) The experimenters then only gave the patient gum when he moved his mouth and lips; by the end of week 3, these behaviors were relatively frequent.

(iii) The experimenters then withheld gum until the patient made vocalizations of some sort; by the end of the fourth week the patient was moving his eyes and lips and making audible "croaking" noises.

(iv) During weeks 4 and 5, the experimenter asked the patient to "say gum," repeating this each time the patient vocalized. At the end of week 6, the patient spontaneously said "gum please."

(v) In later sessions the patient verbally responded to questions from the experimenters, but only in the therapeutic situation.

(vi) To enable verbal behavior to generalize beyond the experimental setting, the patient was placed back on the ward and the nursing staff were asked to attend to his needs—but only if he verbalized them.

This example demonstrates a number of features of the response shaping procedure in clinical settings. First, it provides an example of how response shaping can be a powerful and effective means of establishing complex response repertoires relatively quickly. Second, it also illustrates the distinction between "arbitrary" and "natural" reinforcers in behavior modification. In this case, chewing gum was an effective reinforcer for the behaviors being shaped—but it is an "arbitrary" one in that it is not a normal reinforcer for verbal behavior. Thus, while chewing gum may have acted as an effective reinforcer during the shaping process, in order to be maintained in any way, verbalizations need to be transferred to a more "natural" reinforcer for those behaviors. This was the aim of stage (vi) in the study where the patient's needs were met only if he verbalized them.

Finally, there was no apparent follow-up analysis of the gains achieved in this study, and one suspects that, once back in the unstructured setting of the ward, the patient in the Isaacs et al. (1960) study would have reverted to his previous mute state. However, this study does still emphasize two things. First, as is the case with token economy studies, behavior change has to be subsequently supported by stable and structured changes to the individual's environment which will maintain the therapeutic gains achieved in the behavior modification program. Second, whether the patient in this study did revert to a mute state still does not deny the usefulness of response shaping procedures in swiftly developing relatively complex behavior repertoires—the problem of response maintenance, however, usually requires other considerations (cf. Glynn, 1990; Stokes & Baer, 1977; Stokes & Osnes, 1988).

1.13.4.2.4 Multifaceted behavioral self-control programs

Operant conditioning offers a wide range of principles that the clinician can utilize for therapeutic purposes, but so far we have discussed those that center on individual principles and which are administered in relatively traditional therapist–client settings. However, operant conditioning principles can be used as the basis for devising behavior change programs that individuals can apply themselves in a way which structures and controls the interactions between their behavior and their environment. These were traditionally known as behavioral self-control programs (e.g., Thoresen & Mahoney, 1974), but have since been adapted into multifaceted behavioral programs to deal with a variety of personal problems which include addictions, habits, obsessions, and other behavioral problems (e.g., Lutzker & Martin, 1981; Stuart & Davies, 1972). The resulting program is one that the individual can manage and assess without the need for constant, external therapeutic supervision.

Operant principles that have been incorporated into behavioral self-control programs include stimulus control (environmental planning), reinforcement/punishment, the principle of response-reinforcement contiguity, response shaping and response discrimination (response feedback), among others.

(i) Stimulus control (environmental planning)

This essentially involves fading out or extinguishing undesirable responses by decreasing the number of stimuli which will elicit them. For example, in the treatment of obesity, eating responses will occur in the presence of very many environmental cues that all gain control over the response. To regulate overeating, therefore, it is important to begin by reducing the number of these cues. This can be achieved by eating only in specific environments (such as a kitchen), or even eating from one particular, distinctively patterned plate, and by never engaging in reinforcing activities while eating (for example, watching TV or reading an interesting magazine) (Ferster, Nurnberger, & Levitt, 1962; Stuart, 1967).

(ii) Reinforcement/punishment

In contrast to controlling the stimuli that elicit target behaviors, the individual can also control the consequences of these behaviors by ensuring that desirable behaviors are reinforced and undesirable behaviors go unreinforced or are punished. The control of the consequences of behavior in this way has frequently been used in weight loss programs. Typical reinforcers include refunding of portions of a deposit or fee for meeting behavioral goals (e.g., Harris & Bruner, 1971; Harris & Hallbauer, 1973), and the return or loss of prewritten cheques (perhaps sent to some organization that the individual particularly dislikes) or valuables (Mann, 1972). While such contingencies have proven relatively effective in helping clients to meet behavioral targets, they are often inadequate in maintaining any therapeutic gains. One solution to this is to facilitate generalization of gains by using contingency contracting procedures which foster mutual reinforcement of desirable behaviors between family members (e.g., J. R. Lutzker & S. Z. Lutzker, 1977).

(iii) Response shaping and the setting of attainable therapeutic targets

All behavioral programs set attainment targets of some kind, and it is extremely important that any subgoals in the programs are attainable. For example, in weight-loss programs it is critical that weekly targets are attainable: if they are not met because they are overambitious, then this is tantamount to punishing the effort that was expended in attempting to meet the goal. The reinforcement of modest gains on a regular basis is advisable, and is consistent with the principles inherent in response shaping (see Section 1.13.4.2.3).

(iv) Response discrimination/feedback

Many habitual or addictive behaviors are characterized by the fact that the individual is not always aware of their occurrence, has poor recall of the frequency of the behavior, or, has little structured knowledge of the circumstances in which the behavior is emitted. This being the case, it is perhaps not surprising that such individuals have difficulty controlling these behaviors. One way in which this can be overcome is by including in the program a period of self-observation, where the individual records or charts information relevant to controlling the response.

This may take the form of a diary in which the person notes the time of occurrence of the target behavior, the circumstances in which the behavior is emitted, and any potentially reinforcing or punishing consequences of the behavior (Mahoney, 1974; Stuart, 1971). While this form of self-observation may be beneficial in enabling the individual to identify the frequency of target behaviors and to make readjustments to existing contingencies (i.e, it is a form of functional analysis, see Section

1.13.4.2.1), it is normally the case that self-monitoring alone is not sufficient to maintain any therapeutic gains, but it is best utilized in the context of more multifaceted programs (Bellack, Schwartz, & Rosensky, 1974; Mahoney, 1974).

(v) Response-reinforcer contiguity/contingency

While many people claim to be aware of the principle of operant reinforcement, most rarely apply it consistently. For instance, a student may decide to spend two hours in the library writing a paper and then reward oneself for this effort by going and having a coffee and a chat with friends. However, after two hours the student finds that he or she is working so well that he or she continues writing until concentration and motivation begins to wane, and then goes for coffee. With all the good intentions in the world, what has happened is that the student has inadvertently reinforced behaviors consistent with falling levels of concentration and motivation rather than the two hours focused work that preceded this. Thus, the essence of behavioral programming in self-control procedures is to ensure that the correct behaviors are reinforced, and that sloppy management of contingencies does not inadvertently reinforce inappropriate behaviors.

A good example of a multifaceted behavioral self-control program based on these rules is a behavioral control program aimed at weight loss devised by Stuart (1967). The main elements of this program were: (a) recording of time and quantity of food consumption (self-observation), (b) weighing in before each meal and before bedtime (aiding discrimination of the consequences of eating), (c) removal of food from all places in the house except the kitchen (environmental planning), (d) pairing eating with no other activity that might be reinforcing (such as watching TV) (behavioral programming), (e) setting a weight-loss goal of 1–2 pounds per week (setting attainable goals), (f) slowing down the pace of eating (defining appropriate responses), and (g) substituting other activities for between-meal eating (programming acceptable competing responses).

1.13.5 SUMMARY

This chapter has traced the history of the influence of learning theory on clinical theory and practice. It is clear that the principles of both classical and operant conditioning have played prominent roles in the development of therapeutic techniques and understanding the etiology of psychopathology. Originally, in an attempt to put psychotherapy on a more scientific footing, it was learning theory that provided the basis for a new experimental analysis of psychopathology and a radical alternative to contemporary psychodynamic interpretations. Learning theory is still evolving—especially that area of learning theory that is concerned with human conditioning—and as a result, it still has much to offer clinical psychology in terms of its contribution to the understanding of psychopathology and to the development of new and effective therapies.

1.13.6 REFERENCES

Agras, S., Sylvester, D., & Oliveau, D. (1969). The epidemiology of common fears and phobias. *Comprehensive Psychiatry, 10,* 151–156.

Aitken, R. C. B., Lister, J. A., & Main, C. J. (1981). Identification of features associated with flying phobias in aircrew. *British Journal of Psychiatry, 139,* 38–42.

Alloy, L. B., & Tabachnik, N. (1984). Assessment of variation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review, 91,* 441–485.

Anderson, D. O., & Parmenter, R. (1941). A long-term study of the experimental neurosis in sheep and in the dog. *Psychosomatic Medicine Monographs, 2, Nos. 3 & 4.*

Ayllon, T., & Azrin, N. (1965). The measurement and reinforcement of behavior of psychotics. *Journal of the Experimental Analysis of Behavior, 8,* 357–383.

Ayllon, T., & Azrin, N. H. (1968). *The token economy: A motivational system for therapy and rehabilitation.* New York: Appleton-Century-Crofts.

Ayllon, T., Haughton, E., & Hughes, H. B. (1965). Interpretation of symptoms: Fact or fiction. *Behaviour Research & Therapy, 3,* 1–7.

Azrin, N. H. (1976). Improvements in the community-reinforcement approach to alcoholism. *Behaviour Research & Therapy, 14,* 339–348.

Baer, P. E., & Fuhrer, M. J. (1968). Cognitive processes during differential trace and delayed conditioning of the GSR. *Journal of Experimental Psychology, 78,* 81–88.

Baeyens, F., Eelen, P., Van den Bergh, O., & Crombez, G. (1989). Acquired affective evaluative value: Conservative but not unchangeable. *Behaviour Research & Therapy, 27,* 279–287.

Baeyens, F., Eelen, P, & Van den Bergh, O. (1990). Contingency awareness in evaluative conditioning: A case for unaware affective-evaluative learning. *Cognition & Emotion, 4,* 3–18.

Bandura, A. (1969). *Principles of behavior modification.* New York: Holt.

Banzett, L. K., Liberman, R. P., & Moore, J. W. (1984). Long-term follow-up of the effects of behavior therapy. *Hospital & Community Psychiatry, 35,* 277–279.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders.* New York: New American Library.

Bellack, A. S., Schwartz, J., & Rozensky, R. H. (1974). The contribution of external control to self-control in a weight reduction program. *Journal of Behavior Therapy & Experimental Psychiatry, 5,* 245–249.

Bijou, S. W., & Baer, D. M. (1966). Operant methods in child behavior and development. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 718–789). New York: Appleton-Century-Crofts.

Black, A. H. (1971). Autonomic aversive conditioning in infrahuman subjects. In F. R. Brush (Ed.), *Aversive* *conditioning and learning*. New York: Academic Press.

Borgealt, A. J., Donahoe, J. W., & Weinstein, A. (1972). Effects of delayed and trace components of a compound CS on conditioned suppression and heart rate. *Psychonomic Science, 26,* 13–15.

Booth, R., & Rachman, S. (1992). The reduction of claustrophobia—I. *Behaviour, Research & Therapy, 30,* 207–221.

Boudewyns, P. A., Fry, T. J., & Nightengale, E. (1986). Token economy programs in VA medical centers: Where are they today? *Behavior Therapist, 9,* 126–127.

Brady, J. V., Kelly, D., & Plumlee, L. (1969). Autonomic and behavioral responses of the rhesus monkey to emotional conditioning. *Annals of the New York Academy of Science, 159,* 959–975.

Brewin, C. R., (1988). *Cognitive foundations of clinical psychology*. London: Erlbaum.

Brewin, C. R. (1989). Cognitive change processes in psychotherapy. *Psychological Review, 96,* 379–394.

Brewin, C. R., Dalgleish, T., & Joseph, S. (1996). A dual representation theory of post-traumatic stress disorder. *Psychological Review, 103,* 670–686.

Breznitz, S. (Ed.) (1983). *The denial of stress*. New York: International Universities Press.

Buchwald, A. M. (1959). Extinction after acquisition under different verbal reinforcement combinations. *Journal of Experimental Psychology, 57,* 43–48.

Cook, M., & Mineka, S. (1987). Second order conditioning and overshadowing in the observational conditioning of fear in monkeys. *Behaviour Research & Therapy, 25,* 349–364.

Cook, M., & Mineka, S. (1989). Observational conditioning of fear to fear-relevant versus fear-irrelevant stimuli in rhesus monkeys. *Journal of Abnormal Psychology, 98,* 448–459.

Cook, M., & Mineka, S. (1990). Selective associations in the observational conditioning of fear in rhesus monkeys. *Journal of Experimental Psychology: Animal Behavior Processes, 16,* 372–389.

Cook, M., Mineka, S., Wolkenstein, B., & Laitsch, K. (1985). Observational conditioning of snake fear in unrelated rhesus monkeys. *Journal of Abnormal Psychology, 94,* 591–610.

Corrigan, P. W. (1991). Strategies that overcome barriers to token economies in community programs for severely mentally ill adults. *Community Mental Health Journal, 27,* 7–30.

Corrigan, P. W. (1995). Use of a token economy with seriously mentally ill patients: Criticisms and misconceptions. *Psychiatric Services, 46,* 1258–1263.

Crocker, J. (1981). Judgements of covariation by social perceivers. *Psychological Bulletin, 90,* 272–292.

Davey, G. C. L. (1988). Pavlovian conditioning in humans: UCS revaluation and self-observation of conditioned responding. *Medical Science Research, 16,* 957–961.

Davey, G. C. L. (1989a). Dental phobias and anxieties: Evidence for conditioning processes in the acquisition and modulation of a learned fear. *Behaviour Research & Therapy, 27,* 51–58.

Davey, G. C. L. (1989b). *Ecological learning theory*. London: Routledge.

Davey, G. C. L. (1989c). UCS revaluation and conditioning models of acquired fears. *Behaviour Research & Therapy, 27,* 521–528.

Davey, G. C. L. (1992a). Classical conditioning and the acquisition of human fears and phobias: A review and synthesis of the literature. *Advances in Behaviour Research & Therapy, 14,* 29–66.

Davey, G. C. L. (1992b). An expectancy model of laboratory preparedness effects. *Journal of Experimental Psychology: General, 121,* 24–40.

Davey, G. C. L. (1992c). Characteristics of individuals with fear of spiders. *Anxiety Research, 4,* 299–314.

Davey, G. C. L. (1993). A comparison of three cognitive appraisal strategies: The role of threat devaluation in problem-focused coping. *Personality & Individual Differences, 14,* 535–546.

Davey, G. C. L. (1994). Is evaluative conditioning a qualitatively distinct form of classical conditioning? *Behaviour Research & Therapy, 32,* 291–299.

Davey, G. C. L. (1995). Preparedness and phobias: Specific evolved associations or a generalized expectancy bias. *Behavioral & Brain Sciences, 18,* 289–325.

Davey, G. C. L. (1997). A conditioning model of phobias. In G. C. L Davey (Ed.), *Phobias: Handbook of theory, research and treatment*. Chichester, UK: Wiley.

Davey, G. C. L., & Arulampalam, T. (1981). Second-order electrodermal conditioning in humans. *IRCS Medical Science, 9,* 567–568.

Davey, G. C. L., & Arulampalam, T. (1982). Second-order "fear" conditioning in humans: Persistence of CR2 following extinction of CR1. *Behaviour Research & Therapy, 20,* 391–396.

Davey, G. C. L., Burgess I., & Rashes, R. (1995). Coping strategies and phobias: The relationship between fears, phobias and methods of coping with stressors. *British Journal of Clinical Psychology, 34,* 423–434.

Davey, G. C. L., & Craigie, P. (1997). Manipulation of dangerousness judgements to fear relevant stimuli: Effects on *a priori* UCS expectancy and a *posteriori* covariation assessment. *Behaviour Research & Therapy, 35,* 607–617.

Davey, G. C. L., De Jong, P. J., & Tallis, F. (1993). UCS inflation in the aetiology of a variety of anxiety disorders: Some case histories. *Behaviour Research & Therapy, 31,* 495–498.

Davey, G. C. L., & Dixon, A. (1996). The expectancy bias model of selective associations: The relationship of judgements of CS dangerousness, CS–UCS similarity and prior fear to *a priori* and *a posteriori* covariation assessments. *Behaviour Research & Therapy, 34,* 235–252.

Davey, G. C. L., & Matchett, G. (1993). UCS rehearsal and the retention and enhancement of differential "fear" conditioning: Effects of trait and state anxiety. *Journal of Abnormal Psychology, 103,* 708–718.

Davey, G. C. L., & Matchett, G. (1996). The effects of response feedback on conditioned responding during extinction: Implications for the role of interoception in anxiety-based disorders. *Journal of Psychophysiology, 10,* 291–302.

Davey, G. C. L., McDonald, A. S., Ferguson, C. E., O'Neill, A.-M., Shepherd, J., & Band, D. (1997). Cognitive neutralising strategies, coping and psychological health. Manuscript submitted for publication.

Davey, G. C. L., & McKenna, I. (1983). The effects of postconditioning revaluation of CS1 and UCS following Pavlovian second-order electrodermal conditioning in humans. *Quarterly Journal of Experimental Psychology, 35B,* 125–133.

Davey, G. C. L., & Singh, J. (1988). The Kamin "blocking" effect and electrodermal conditioning in humans. *Journal of Psychophysiology, 2,* 17–25.

Dawson, M. E. (1973). Can classical conditioning occur without contingency learning? A review and evaluation of the evidence. *Psychophysiology, 10,* 82–86.

Dawson, M. E., & Biferno, M. A. (1973). Concurrent measurement of awareness and electrodermal classical conditioning. *Journal of Experimental Psychology, 101,* 55–62.

Dawson, M. E., Catania, J. J., Schell, A. M., & Grings, W. W. (1979). Autonomic classical conditioning as a function of awareness of stimulus contingencies. *Biological Psychology, 9,* 23–40.

Dawson, M. E., & Grings, W. W. (1968). Comparison of classical conditioning and relational learning. *Journal of Experimental Psychology, 76,* 227–231.

Dawson, M. E., & Schell, A. M. (1987). The role of "controlled" and "automatic" cognitive processes in human autonomic classical conditioning. In G. C. L. Davey (Ed.), *Cognitive processes and Pavlovian conditioning in humans*. Chichester, UK: Wiley.

Dawson, M. E., Schell, A. M., & Tweddle-Banis, H. T. (1986). Greater resistance to extinction of electrodermal responses conditioned to potentially phobic CSs: A noncognitive process? *Psychophysiology, 23,* 552–561.

de Silva, P. (1988). Phobias and preparedness: Replication and extension. *Behaviour Research & Therapy, 26,* 97–98.

de Silva, P., Rachman, S. J., & Seligman, M. E. P. (1977). Prepared phobias and obsessions: Therapeutic outcome. *Behaviour Research & Therapy, 15,* 65–77.

Deane, G. E. (1969). Cardiac activity during experimentally induced anxiety. *Psychophysiology, 6,* 17–30.

Delprato, D. J. (1980). Hereditary determinants of fears and phobias: A critical review. *Behavior Therapy, 11,* 79–103.

Derby, K. M., Wacker, D., Sasso, G., Steege, M., Northrup, J., Cigrand, K., & Asmus, J. (1992). Brief functional assessment techniques to evaluate aberrant behavior in an outpatient setting: A summary of 79 cases. *Journal of Applied Behavior Analysis, 25,* 713–721.

Diamond, D., Matchett, G., & Davey, G. C. L. (1995). The effect of prior fear levels on UCS-expectancy ratings to a fear-relevant stimulus. *Quarterly Journal of Experimental Psychology, 48A,* 237–247.

Dickerson, F., Ringel, N., & Parente, F. (1994). Seclusion and restraint, assaultiveness, and patient performance in a token economy. *Hospital & Community Psychiatry, 45,* 168–170.

Dickinson, A. (1980). *Contemporary animal learning theory*. Cambridge, UK: Cambridge University Press.

Dickinson, A., & Mackintosh, N. J. (1979). Reinforcer specificity in the enhancement of conditioning by post-trial surprise. *Journal of Experimental Psychology: Animal Behavior Processes, 5,* 162–177.

Dickinson, A., Shanks, D., & Evenden, J. (1984). Judgement of act-outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology, 36A,* 29–50.

Dollard, J., & Miller, N. E. (1950). *Personality and psychotherapy: An analysis in terms of learning, thinking and culture*. New York: McGraw-Hill.

Emmelkamp, P. M. G. (1982). *Phobic and obsessive-compulsive disorders: Theory, research and practice*. New York: Plenum.

English, H. B. (1929). Three cases of the "conditioned fear response." *Journal of Abnormal & Social Psychology, 34,* 221–225.

Esteves, F., Dimberg, U., Parra, C., & Öhman, A. (1994). Nonconscious associative learning: Pavlovian conditioning of skin conductance responses to masked fear-relevant facial stimuli. *Psychophysiology, 31,* 375–385.

Eysenck, H. J. (1976). The learning theory model of neurosis—a new approach. *Behaviour Research & Therapy, 14,* 251–268.

Eysenck, H. J., & Kelley, M. J. (1987). The interaction of neurohormones with Pavlovian A and Pavlovian B conditioning in the causation of neurosis, extinction, and incubation of anxiety. In G. C. L. Davey (Ed.), *Cognitive processes and Pavlovian conditioning in humans* (pp. 251–286). Chichester, UK: Wiley.

Eysenck, H. J., & Rachman, S. J. (1965). *Causes and cures of neuroses*. London: Routledge.

Eysenck, H. J. (1979). The conditioning model of neurosis. *Behavioral & Brain Sciences, 2,* 155–199.

Fairbank, J. A., & Brown, T. A. (1987). Current behavioral approaches to the treatment of post-traumatic stress disorder. *The Behavior Therapist, 3,* 57–64.

Ferster, C. B., Nurnberger, J. I., & Levitt, E. B. (1962). The control of eating. *Journal of Mathetics, 1,* 87–109.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York: Appleton-Century-Crofts.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99,* 20–35.

Gelder, M., Bancroft, J. H. J., Gath, D., Johnston, D. W., Mathews A. M., & Shaw P. M. (1973). Specific and non-specific factors in behaviour therapy. *British Journal of Psychiatry, 123,* 445–462.

Gershone, J. R., Errickson, E. A., Mitchell, J. E., & Paulson, D. A. (1977). Behavioral comparison of a token economy and a standard psychiatric treatment ward. *Journal of Behavior Therapy & Experimental Psychiatry, 8,* 381–385.

Glynn, S. M. (1990). Token economy approaches for psychiatric patients: Progress and pitfalls over 25 years. *Behavior Modification, 14,* 383–407.

Goldstein, A. J., & Chambers, D. L. (1978). A reanalysis of agoraphobia. *Behavior Therapy, 9,* 47–59.

Goorney, A. B. (1970). Treatment of aviation phobias by behaviour therapy. *British Journal of Psychiatry, 117,* 535–544.

Grether, W. F. (1938). Pseudo-conditioning without paired stimulation encountered in attempted backward conditioning. *Journal of Comparative Psychology, 25,* 91–96.

Gripp, R. F., & Magaro, P. A. (1971). A token economy program evaluated with untreated control ward comparisons. *Behaviour Research & Therapy, 9,* 137–149.

Hafner, J., & Marks, I. (1976). Exposure *in vivo* of agoraphobics: Contributions of diazipam, group exposure and anxiety evocation. *Psychological Medicine, 6,* 71–88.

Hagopian, L. P., Fisher, W. W., & Legacy, S. M. (1994). Schedule effects of noncontingent reinforcement on attention-maintained destructive behavior in identical quadruplets. *Journal of Applied Behavior Analysis, 27,* 317–325.

Hall, J., & Baker, R. (1973). Token economy systems: Breakdown and control. *Behaviour Research & Therapy, 11,* 253–263.

Hamm, A. O., Viatl, D., & Lang, P. J. (1989). Fear conditioning, meaning and belongingness: A selective analysis. *Journal of Abnormal Psychology, 98,* 395–406.

Harris, B. (1979). Whatever happened to Little Albert? *American Psychologist, 34,* 151–160.

Harris, M. B., & Bruner, C. G. (1971). A comparison of a self-control and a contract procedure for weight control. *Behaviour Research & Therapy, 9,* 347–354.

Harris, M. B., & Hallbauer, E. S. (1973). Self-directed weight control through eating and exercise. *Behaviour Research & Therapy, 11,* 523–529.

Haynes, S. N., & O'Brien, W. H. (1990). Functional analysis in behavior therapy. *Clinical Psychology Review, 10,* 649–668.

Hedberg, A. G., & Campbell L., III (1974). A comparison of four behavioral treatments of alcoholism. *Journal of Behavior Therapy and Experimental Psychiatry, 5,* 251–256.

Hersen, M., Eisler, R. M., Alford, G., & Agras, S. (1973). Effects of token economy on neurotic depression: An experimental analysis. *Behavior Therapy, 4,* 392–397.

Hofmeister, J. F., Schneckenbach, A. F., & Clayton, S. H. (1979). A behavioral program for the treatment of chronic patients. *American Journal of Psychiatry, 136,* 396–400.

Honeybourne, C., Matchett, G., & Davey, G. C. L. (1993). Expectancy models of laboratory preparedness effects: A UCS-expectancy bias in phylogenetic and ontogenetic fear-relevant stimuli. *Behavior Therapy, 24,* 253–264.

Hugdahl, K., & Johnsen, B. H. (1989). Preparedness and electrodermal fear conditioning: Ontogenetic vs. phylogenetic explanations. *Behaviour Research & Therapy, 27,* 269–278.

Isaacs, W., Thomas, J., & Goldiamond, I. (1960). Application of operant conditioning to reinstate verbal behavior in psychotics. *Journal of Speech & Hearing Disorders, 25,* 8–12.

Iwata, B. A, Dorsey, M. F., Slifer, K. J., Bauman K. E., & Richman, G. S. (1985). Toward a functional analysis of self-injury. In G. Murphy & B. Wilson (Eds.), *Self-injurious behavior*. Kidderminster, UK: BIHM.

Jacobson, N., & Weiss, R. L. (1978). Behavioral marriage therapy III. The contents of Gurman et al. may be hazardous to our health. *Family Process, 17,* 149–163.

Kamin, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Eds.), *Miami symposium on the prediction of behavior: Aversive stimulation*. Miami, FL: University of Miami Press.

Kamin, L. J. (1969). Predictability, surprise, attention and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior*. New York: Appleton-Century-Crofts.

Kazdin, A. E. (1975). *Behavior modification in applied settings*. Homewood, IL: Dorsey Press.

Kazdin, A. E., & Bootzin, R. R. (1972). The token economy: An evaluative review. *Journal of Applied Behavior Analysis, 5,* 343–372.

Kazdin, A. E., & Wilcoxon, L. A. (1976). Systematic desensitization and non-specific treatment effects: A methodological evaluation. *Psychological Bulletin, 83, 729–758*.

Kazdin, A. E. (1981). The token economy. In G.C.L.Davey (Eds.), *Applications of conditioning theory*. London: Methuen.

Keane, T. M., Zimering, R. T., & Caddell, J. M. (1985). A behavioral formulation of post-traumatic stress disorder in Vietnam veterans. *The Behavior Therapist, 8,* 9–12.

Kilpatrick, D. G., Veronen L. J., & Best C. L. (1985). Factors predicting psychological distress among rape victims. In C. R. Figley (Ed.), *Trauma and its wake* (pp. 113–141). New York: Brunner/Mazel.

Kimble, G. A. (1961). *Hilgard and Marquis" Conditioning and learning*. London: Methuen.

Kirkpatrick, D. R. (1984). Age, gender and patterns of common intense fears among adults. *Behavior Research & Therapy, 22,* 141–150.

Klein, S. B. (1969). Counterconditioning and fear reduction in the rat. *Psychonomic Science, 77,* 150–151.

Kockott, G., Dittmar, F., & Nusselt, L. (1975). Systematic desensitization of erectile impotence: A controlled study. *Archives of Sexual Behavior, 4,* 493–500.

Krasner, L. (1971). Behavior therapy. *Annual Review of Psychology, 22,* 483–532.

Lautch, H. (1971). Dental phobia. *British Journal of Psychiatry, 119,* 151–158.

Leitenberg, H. (1976). *Handbook of behavior modification and behavior therapy*. Englewood Cliffs, NJ: Prentice-Hall.

Lemere, F., & Voegtlin, W. L. (1950). An evaluation of aversive treatment of alcoholism. *Quarterly Journal of the Study of Alcoholism, 11,* 199–204.

Lichtenstein, E., & Danaher, B. (1976). Modification of smoking behavior: A critical analysis of theory, research and practice. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification* (Vol. 3). New York: Academic Press.

Liddell, A., & Lyons, M. (1978). Thunderstorm phobias. *Behavior Research & Therapy, 16,* 306–308.

Lindsley, O. R., Skinner, B. F., & Solomon, H. C. (1953). *Studies in behavior therapy: Status report I.* Waltham, MA: Metropolitan State Hospital.

Lippman, M. R., & Motta, R. W. (1993). Effects of positive and negative reinforcement on daily living skills in chronic psychiatric patients in community residences. *Journal of Clinical Psychology, 49,* 654–662.

Lowe, C. F. (1979). Determinants of human operant behavior. In M. D. Zeiler & P. Harzem (Eds.), *Advances in analysis of behavior* (Vol 1, pp. 159–192). Chichester, UK: Wiley.

Lowe, C. F. (1983). Radical behaviorism and human psychology. In G. C. L. Davey (Ed.), *Animal models of human behavior*. Chichester, UK: Wiley.

Lubow, R. E. (1973). Latent inhibition. *Psychological Bulletin, 79,* 398–407.

Lutzker, J. R., & Lutzker, S. Z. (1977). A two-dimensional contract: Weight loss and household responsibility performance. In E. E. Abramson (Ed.), *Behavioral approaches to weight control* (pp. 114–122). New York: Springer.

Lutzker, J. R., & Martin, J. A. (1981). *Behavior change*. Monterey, CA: Brooks.

Mace, F. C., Browder, D., & Lyn, Y. (1987). Analysis of demand conditions associated with stereotypy. *Journal of Behavior Therapy & Experimental Psychiatry, 18,* 25–31.

Mace, F. C., & Knight, D. (1986). Functional analysis and treatment of severe pica. *Journal of Applied Behavior Analysis, 19,* 411–416.

Mackintosh, N. J. (1983). *Conditioning and associative learning*. Oxford, UK: Oxford University Press.

Mahoney, M. J. (1974). Self-reward and self-monitoring techniques for weight control. *Behavior Therapy, 5,* 48–57.

Mann, R. A. (1972). The behavior-therapeutic use of contingency contracting to control an adult weight problem: Weight control. *Journal of Applied Behavior Analysis, 5,* 99–109.

Marks, I. M. (1987). *Fears, phobias and rituals*. New York: Oxford University Press.

Masserman, J. H. (1943). *Behavior and neurosis*. Chicago: University of Chicago Press.

Masserman, J. H. (1950). Experimental neurosis. *Scientific American, 182,* 38–50.

Mathews, A. (1978). Fear-reduction research and clinical phobias. *Psychological Bulletin, 85,* 390–406.

Mathews, A. M., Bancroft, J., Whitehead, A., Hackmann, A., Julier, D., Bancroft, J., Gath, D., & Shaw, P. (1976). The behavioral treatment of sexual inadequacy: A comparative study. *Behaviour Research & Therapy, 14,* 427–436.

Mazaleski, J. L., Iwata, B. A., Vollmer, T. R., Zarcone, J. R., & Smith, R. G. (1993). Analysis of the reinforcement and extinction components in DRO contingencies with self-injury. *Journal of Applied Behavior Analysis, 26,* 143–156.

McAllister, W. R., & McAllister, D. E. (1971). Behavioral measurement of conditioned fear. In F. R. Brush (Ed.), *Aversive conditioning and learning*. New York: Academic Press.

McNally, R. J. (1987). Preparedness and phobias: A review. *Psychological Bulletin, 101,* 283–303.

McNally, R. J., & Heatherton, T. F. (1993). Are covariation biases attributable to *a priori* expectancy biases. *Behaviour Research & Therapy, 31,* 653–658.

Menzies, R. G., & Clarke, J. C. (1993a). The aetiology of childhood water phobia. *Behaviour Research & Therapy, 31,* 499–501.

Menzies, R. G., & Clarke, J. C. (1993b). The etiolology of fear of heights and its relationship to severity and individual response patterns. *Behaviour Research & Therapy, 31,* 355–365.

Merckelbach, H., & de Jong, P. J. (1997). Evolutionary models of phobias. In G. C. L. Davey (Ed.), *Phobias: Handbook of theory, research and treatment*. Chichester, UK: Wiley.

Merckelbach, H., de Ruiter, C., van den Hout, M. A., & Hoekstra, R. (1989). Conditioning experiences and phobias. *Behaviour Research & Therapy, 27,* 657–662.

Mineka, S., & Cook, M. (1993). Mechanisms involved in the observational conditioning of fear. *Journal of*

Experimental Psychology: General, 122, 23–38.

Mineka, S., Davidson, M., Cook, M., & Weir, R. (1984). Observational conditioning of snake fear in rhesus monkeys. *Journal of Abnormal Psychology, 93,* 355–372.

Mowrer, O. H. (1947). On the dual nature of learning—a reinterpretation of "conditioning" and "problem-solving." *Harvard Educational Review, 17,* 102–148.

Murray, E., & Foote, F. (1979). The origins of fear of snakes. *Behaviour Research & Therapy, 17,* 489–493.

Northup, J., Broussard, C., Jones, K., George, T., Vollmer, T. R., & Herring, M. (1995). The differential effects of teacher and peer attention on the disruptive classroom behavior of three children with a diagnosis of attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis, 28,* 277–288.

O'Leary, K. D., & Wilson, G. T. (1975). *Behavior therapy: Application and outcome*. Englewood Cliffs, NJ: Prentice-Hall.

O'Reilly, M. F. (1995). Functional analysis and treatment of escape-maintained aggression correlated with sleep deprivation. *Journal of Applied Behavior Analysis, 28,* 225–226.

Öhman, A. (1979). Fear relevance, autonomic conditioning and phobias: A laboratory model. In P. O. Sjoden, S. Bates, & W. S. Dockens III (Eds.), *Trends in behaviour therapy*. New York: Academic Press.

Öhman, A. (1993). Fear and anxiety as emotional phenomena: Clinical phenomenology, evolutionary perspectives, and information processing mechanisms. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions*. New York: Guilford Press.

Öhman, A., Dimberg, U., & Öst, L-G (1985). Animal and social phobias: Biological constraints on learned fear responses. In S. Reiss & R. R. Bootzin (Eds.), *Theoretical issues in behavior therapy*. New York: Academic Press.

Öhman, A., Fredrikson, M. L., & Hugdahl, K. (1978). Towards an experimental model of simple phobic reactions. *Behavioural Analysis & Modification, 2,* 97–114.

Öst, L-G. (1989). A one-session group treatment for specific phobias. *Behaviour Research & Therapy, 27,* 1–7.

Öst, L-G. (1997). Rapid treatment of specific phobias. In G. C. L. Davey (Ed.), *Phobias: Handbook of theory, research and treatment*. Chichester, UK: Wiley.

Öst, L-G., & Hugdahl, K. (1981). Acquisition of phobias and anxiety response patterns in clinical patients. *Behaviour Research & Therapy, 19,* 439–447.

Owens, R. G., & Ashcroft, J. E. (1982). Functional analysis in applied psychology. *British Journal of Clinical Psychology, 21,* 181–189.

Parkinson, B. (1985). Emotional effects of false autonomic feedback. *Psychological Bulletin, 98,* 471–494.

Paul, G. L., & Bernstein, D. (1973). *Anxiety and clinical problems: Systematic desensitization and related techniques*. Morristown, NJ: General Learning Press.

Paul, G. L., & Lentz, R. J. (1977). *Psychological treatment of chronic mental patients: Milieu versus social-learning programs*. Cambridge, MA: Harvard University Press.

Pavlov, I. P. (1927). *Conditioned reflexes*. Oxford, UK: Oxford University Press.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review, 87,* 532-552.

Pierrel, R., & Sherman, J. G. (1963). Barnabus, the rat with the college training. *Brown Alumni Monthly, February*.

Prokasy, W. F., & Kumpfer, K. L. (1969). Conditioning, probability of reinforcement and sequential behavior in human conditioning with intermittent schedules. *Psychonomic Science, 14,* 49–50.

Prokasy, W. F., & Williams, W. C. (1979). Information processing and the decremental effect of intermittent reinforcement schedules in human conditioning. *Bulletin of the Psychonomic Society, 14,* 57–60.

Rachman, S. (1977). The conditioning theory of fear acquisition: A critical examination. *Behaviour Research & Therapy, 15,* 375–387.

Rachman, S. J., & Teasdale J. (1969). *Aversion therapy and behavior disorders: An analysis*. Coral Gables, FL: University of Miami Press.

Rachman, S. J., & Wilson, G. T. (1980). *The effects of psychological therapy*. London: Pergamon.

Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review, 74,* 71–80.

Rescorla, R. A. (1980). *Pavlovian second-order conditioning: Studies in associative learning*. Hillsdale, NJ: Erlbaum.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory*. New York: Appleton-Century-Crofts.

Richards, M., & Martin, I. (1990). Eysenck's incubation of fear hypothesis: An experimental test. *Behaviour Research & Therapy, 28,* 373–384.

Rizley, R. C., & Rescorla, R. A. (1972). Associations in second-order conditioning and sensory preconditioning. *Journal of Comparative & Physiological Psychology, 81,* 1–11.

Russell, C., & Davey, G. C. L. (1991). The effects of false response feedback on human "fear" conditioning. *Behaviour Research & Therapy, 29,* 19–196.

Salkovskis, P. M. (1991). The importance of behaviour in the maintenance of anxiety and panic: A cognitive account. *Behavioural Psychotherapy, 19,* 6–19.

Samson, D. M., & McDonnell, A. A. (1990). Functional analysis and challenging behaviours. *Behavioural Psychotherapy, 18,* 259–271.

Sandler, S., & Davidson, R. S. (1973). *Psychopathology: Learning theory, research and applications*. New York: Harper Row.

Scheier, M. F., & Carver, C. S. (1992). Effects of optimism on psychological and physical well-being: Theoretical overview and empirical update. *Cognitive Therapy & Research, 16,* 201–228.

Schneiderman, N. E., Pearl, L., Wilson, W., Metcalf, F., Moore, J. W., & Swadlow, H. A. (1971). Stimulus control in rabbits (*Oryctolagus cuniculus*) as a function of different intensities of intracranial stimulation. *Journal of Comparative & Physiological Psychology, 76,* 175–186.

Seligman, M. E. P. (1971). Phobias and preparedness. *Behavior Therapy, 2,* 307–320.

Seligman, M. E. P. (1976). *Helplessness*. San Francisco: Freeman.

Siddle, D. A. T., & Remington, B. (1987). Latent inhibition and human Pavlovian conditioning: Research and relevance. In G. C. L. Davey (Ed.), *Cognitive processes and Pavlovian conditioning in humans* (pp. 115–146). Chichester, UK: Wiley.

Siddle, D. A. T., Remington, B., & Churchill, M. (1985). Effects of conditioned stimulus pre-exposure on human electrodermal conditioning. *Biological Psychology, 20,* 113–127.

Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1948). "Superstition" in the pigeon. *Journal of Experimental Psychology, 38,* 168–172.

Skinner, B. F. (1953). *Science and human behavior*. New York: MacMillan.

Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1969). *Contigencies of reinforcement: A theoretical analysis*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1971). *Beyond freedom and dignity*. New York: Alfred A. Knopf.

Skinner, B. F. (1974). *About behaviorism*. London: Jonathon Cape.

Slade, P. (1982). Towards a functional analysis of Anorexia Nervosa and Bulimia Nervosa. *British Journal of Clinical Psychology, 21,* 167–179.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10,* 349–369.

Stokes, T. F., & Osnes, P. G. (1988). The developing applied technology of generalization and maintenance. In R. H. Homer, G. Dunlap, & R. L. Koegel (Eds.), *Generalization and maintenance: Life style changes in applied settings*. Baltimore: Brookes.

Stuart, R. B. (1967). Behavioral control of overeating. *Behaviour Research & Therapy, 5,* 357–365.

Stuart, R. B. (1971). A three-dimensional program for the treatment of obesity. *Behaviour Research & Therapy, 9,* 177–186.

Stuart, R. B., & Davis, B. (1972). *Slim chance in a fat world: Behavioral control of obesity*. Champaign, IL.: Research Press.

Thoresen, C. E., & Mahoney, M. J. (1974). *Behavioral self-control*. New York: Holt, Rinehart & Winston.

Thorndike, E. L. (1935). *The psychology of wants, interests and attitudes*. London: Appleton-Century-Crofts.

Valentine, C. W. (1930). The innate bases of fear. *Journal of Genetical Psychology, 37,* 394–420.

Voegtlin, W. L., & Lemere, F. (1942). The treatment of alcohol addiction: A review of the literature. *Quarterly Journal of the Study of Alcohol, 2,* 717–803.

Vollmer, T. R., Iwata, B. A., Zarcone, J. R., Smith, R. G., & Mazaleski, J. L. (1993). The role of attention in the treatment of attention-maintained self-injurious behavior: Noncontingent reinforcement (NCR) and differential reinforcement of other behavior (DRO). *Journal of Applied Behavior Analysis, 26,* 9–21.

Vollmer, T. R., & Smith, R. G. (1996). Some current themes in functional analysis research. *Research in Development Disabilities, 17,* 229–249.

Wacker, D. P., Steege, M. W., Northup, J., Sasso, G., Berg, W., Reimers, T., Cooper, L., Cigrand, K., & Donn, L. (1990). A component analysis of functional communication training across three topographies of severe behavior problems. *Journal of Applied Behavior Analysis, 23,* 417–429.

Walker, H., & Buckley, N. (1968). The use of positive reinforcement in conditioning attending behavior. *Journal of Applied Behavior Analysis, 1,* 242–250.

Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3,* 1–14.

Weiner, H. (1969). Controlling human fixed-interval performance. *Journal of the Experimental Analysis of Behavior, 12,* 349–373.

White, K., & Davey, G. C. L. (1989). Sensory preconditioning and UCS inflation in human "fear" conditioning. *Behaviour Research & Therapy, 27,* 161–166.

Wiens, A. N., Montage, J. R., Manaugh, T. S., & English, C. J. (1976). Pharmacological aversive counterconditioning to alcohol in a private hospital: One year follow-up. *Journal of Studies on Alcohol, 37,* 1320–1324.

Willems, E. P. (1974). Behavioral technology and behavioral ecology. *Journal of Applied Behavior Analysis, 7,* 151–165.

Williams, C. D. (1959). The elimination of tantrum behaviors by extinction procedures. *Journal of Abnormal and Social Psychology, 59,* 269.

Williams, S. L., Turner, S. M., & Peer, D. F. (1985). Guided mastery and performance desensitization treatments for severe acrophobia. *Journal of Consulting & Clinical Psychology, 53,* 237–247.

Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin, 90,* 245–271.

Wilson, G. T. (1968). Reversal of differential GSR conditioning by instructions. *Journal of Experimental Psychology, 76,* 491–493.

Wilson, G. T. (1973). Effects of false feedback of avoidance behavior: "Cognitive" desensitization revisited. *Journal of Personality & Social Psychology, 28,* 115–123.

Wilson, G. T. (1978). Aversion therapy for alcoholism: Issues, ethics, and evidence. In G. A. Marlatt & P. E. Nathan (Eds.), *Behavioral approaches to alcoholism*. New Brunswick, NJ: Center for Alcohol Studies.

Wilson, G. T., Leaf, R. C., & Nathan, P. E. (1975). The aversive control of excessive alcohol consumption by chronic alcoholics in the laboratory setting. *Journal of Applied Behaviour Analysis, 8,* 13–26.

Withers, R. D. & Deane, F. P. (1995). Origins of common fears: Effects of severity, anxiety responses and memories of onset. *Behaviour Research & Therapy, 33,* 903–915.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press.

Wolpe, J., & Lazarus, A. A. (1966). *Behavior therapy techniques: A guide to the treatment of neuroses*. New York: Pergamon.

Wolpe. J., & Rowan, V. C. (1988). Panic disorder: A product of classical conditioning. *Behavior Research & Therapy, 26,* 441–450.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.14

Psychodynamic Theory

PETER FONAGY
University College London, UK

1.14.1 INTRODUCTION 424
1.14.1.1 The Basic Psychodynamic Model 424
1.14.1.2 The Plurality of Psychodynamic Theories and the Homogeneity of Psychodynamic Frames of Reference 424
1.14.1.3 Problems of Validity of Psychodynamic Theories 425
1.14.1.4 Psychic Reality, Physical Reality, Fantasy, and Reconstruction 425
1.14.1.5 Further Controversial Issues 426
1.14.2 FREUD'S MULTIPLE MODELS OF THE MIND 426
1.14.2.1 The Affect-trauma and Topographic Frames of Reference 426
1.14.2.2 The Structural Frame of Reference 427
1.14.2.3 Some Limitations of Freud's Models 427
1.14.3 THE STRUCTURAL APPROACH 427
1.14.3.1 American Ego Psychology 427
1.14.3.1.1 Hartmann's classical model 427
1.14.3.1.2 Modern structural theory 428
1.14.3.1.3 Extensions of the ego psychology model 428
1.14.3.1.4 The classical model of the neurosis 429
1.14.3.1.5 The classical model of severe psychopathology 429
1.14.3.2 The Structural Developmental Tradition 430
1.14.3.2.1 Anna Freud 430
1.14.3.2.2 Margaret Mahler 431
1.14.3.2.3 Sandler's reframing of the structural model 432
1.14.3.3 Strengths and Weaknesses of the Model 432
1.14.4 THE OBJECT RELATIONS APPROACH 432
1.14.4.1 Introduction to the Object Relations Approach 432
1.14.4.2 The Klein–Bion Model 433
1.14.4.2.1 Some essential Kleinian concepts 433
1.14.4.2.2 Kleinian approach to neurosis 435
1.14.4.2.3 Rosenfeld's developmental model of narcissism 435
1.14.4.2.4 Kleinian approach to severe psychopathology 435
1.14.4.2.5 Strengths and weaknesses of the Klein–Bion model 435
1.14.4.3 The Fairbairn–Winnicott Model 436
1.14.4.3.1 Major innovations of the "Independent" tradition 436
1.14.4.3.2 The Independent view of neurotic processes and severe psychopathology 437
1.14.4.3.3 Strengths and weaknesses of the approach 437
1.14.4.4 Kohut's Self-psychology 438
1.14.4.4.1 Innovative aspects of Kohut's theory 438
1.14.4.4.2 Self-psychological model of mental disorder 438
1.14.4.4.3 Strengths and weaknesses of Kohut's model 438
1.14.4.5 Kernberg's Integrated Theory 439
1.14.4.5.1 Internalization and the construction of self, object, and affect triads 439
1.14.4.5.2 Levels of pathology within Kernberg's system 439
1.14.4.5.3 Borderline personality disorder 439
1.14.4.5.4 Strengths and weaknesses of Kernberg's model 440
1.14.4.6 Bowlby's Attachment Theory Model 440
1.14.4.6.1 Innovative aspects of Bowlby's approach 440

1.14.4.6.2 Attachment and psychopathology 441
1.14.4.6.3 Strengths and weaknesses of Bowlby's theory 441
1.14.5 EVIDENCE FOR PSYCHODYNAMIC THEORIES 441
1.14.6 REFERENCES 443

1.14.1 INTRODUCTION

1.14.1.1 The Basic Psychodynamic Model

This chapter provides a review of psychodynamic theories, including classical and contemporary structural theories, developments of ego psychological models, and British and US object relational approaches. The discussion of each of these psychoanalytical schools is organized in terms of illustrating the contributions each may be seen to provide to clinical psychology in terms of etiological treatment and empirical considerations.

Psychodynamic psychology, Freud's discovery and invention (Cooper, 1985), has enjoyed considerable success as an explanatory framework for understanding psychopathology and clinical work because its few basic assumptions and propositions are open to endless elaboration, revision, and refinement. Most specific psychodynamic propositions are data dependent, that is, they may be revised or discarded without fundamentally affecting the integrity of the psychoanalytic approach. There may be some assumptions which are essential to the theory (Sandler & Joffe, 1969). These include:

(i) *psychic determinism*, namely that cognitive, emotional and behavioral aspects of pathology may be most conveniently studied in terms of psychological causes (rather then physical causality or random biological events);

(ii) *the pleasure–unpleasure principle*, namely that behavior may be seen as an adaptive effort at minimizing psychic pain and maximizing psychic pleasure and a sense of intrapsychic safety;

(iii) *the biological nature* of the organism drives its psychological adaptation;

(iv) *the dynamic unconscious*, namely that mental forces contend for control over access to consciousness and to actions; and

(v) *the genetic–developmental* proposition which states that behavior is explicable in terms of earlier (or even earliest infantile) events.

1.14.1.2 The Plurality of Psychodynamic Theories and the Homogeneity of Psychodynamic Frames of Reference

Psychodynamic theories are diverse in terms of focus but share a common frame of reference. Undoubtedly because of the flexibility of its assumptions, psychodynamic theory has been in a constant state of change since its inception. Historically, it has evolved from an early concentration on the role of instincts (sexual and aggressive) in psychopathology via a focus on the functions of the ego, to the current interest in the early mother–infant dyad and its long-term effect upon interpersonal relationships and their internal representation (object relations). In the meantime a psychology of the self has become part of most psychoanalytic theories, and the move away from a physical science metaphor for psychopathology has resulted in a clinical theory where both pathology and its treatment are viewed in terms of mental representations.

Psychodynamic theory has changed literally out of all recognition over the past 30 years. Yet mental health scientists, if they think about psychodynamic theories at all, are still likely to think of these as some parody of the libidinal developmental phases of oral, anal, phallic, and genital sexuality: excessive indulgence of any of these pleasures, as well as their unreasonable frustration, causes fixations which psychoanalysts are supposed to regard as the roots of adverse developmental consequences.

Although now clearly outdated, there was a remarkably prescient aspect to Freud's (1905b) psychosexual theory. His general approach, seeing adult pathology as the outcome of deviations from normal developmental processes, set a developmental frame of reference for psychoanalytic thought which has remained important to the present day and currently links psychodynamic thought to the vibrant discipline of developmental psychopathology (Cicchetti & Cohen, 1995). The continuity of personality from childhood through adolescence to adulthood, and the role of childhood experience in setting these trajectories, lies at the heart of the psychodynamic approach and serves to unify it.

Early theories, however, continue to exist side-by-side with more recent formulations. Some writers even suggest that a number of frameworks are necessary to provide a comprehensive theory of psychopathology and therapy (Pine, 1985). In practice, each theory has its "range of convenience," particular clinical phenomena which it addresses in a particularly helpful way. For example, dreams are still best explained in the context of Freud's topographical frame of reference (Freud, 1900). By contrast, neurotic disorders are most easily explained psychodynamically as part of a psychology of conflict (Brenner, 1982). Axis II disorders, particularly personality disorders in the Dramatic Cluster, are most productively explored in the context of theories which have evolved as repudiations of the conflict-based models (Kohut, 1971).

1.14.1.3 Problems of Validity of Psychodynamic Theories

A major shortcoming of psychodynamic models lies in the fact that hypotheses about the nature of development, psychopathology, or treatment techniques are rarely based on direct evidence. To accept clinical accounts as validating psychodynamic hypotheses runs into justifiable opposition from philosophers of science who regard them as irretrievably confounded by suggestion (e.g., Grünbaum, 1984). Psychoanalytic metapsychology is at best loosely coupled to clinical observations (Fonagy, 1982; Holzman, 1985; G. S. Klein, 1976b; Ricoeur, 1977; Schafer, 1976) and clinical reports cannot therefore provide an independent confirmation for psychodynamic theory.

Their clinical roots explain why psychoanalytic ideas tend to reflect the population of patients which preoccupied particular theoreticians. For example, Winnicott (1965c) conceived of inauthenticity and the false self as a core problem and focused on its cause as the failure of good enough mothering and the inadequate provision of a holding environment. Kohut's (1971, 1977) central clinical puzzle was the question of how one develops an enfeebled self, and he oriented his interests towards the mother's capacity for empathic responsiveness. It is rarely clear whether each psychodynamic approach is associated with a particular category of clinical cases or, as is more likely the case, theoreticians reconstruct their patients in ways that fit the theory.

There is a further reason why psychodynamic ideas are hard to integrate into the network of behavioral science knowledge even at the level of clinical description. Psychodynamic diagnostic descriptions map poorly onto current psychiatric nomenclature (Jacobson & Cooper, 1993; Shapiro, 1989). Many psychodynamically oriented clinicians claim that the phenomenological approach of modern psychiatry is inherently alien to the etiological models of psychoanalysis (A. Freud, 1965), and the intention of the framers of the *Diagnostic and statistical manual of mental disorders*, 4th edition (*DSM-IV*), is to exclude all psychodynamic data and inference as basic to clinical assessment (Frances, Pincus, Widiger, Davis, & First, 1990). The behavior-based, descriptive approach of *DSM* minimizes clinical inference, is rooted in logical positivism and thus favors behavioral and biological orientations (M. A. Schwartz, 1991) over dynamic and other potentially useful orientations to psychopathology (M. A. Schwartz & Wiggins, 1988).

Psychodynamic accounts have a further pervasive limitation: that of gender bias. Since the work of Freud (1900), masculine development has invariably been more coherently described than its feminine counterpart (Horney, 1939). In contrast, psychodynamic models far more often implicate the mother in pathological processes than they do the father, with remarkably few exceptions (e.g., Limentani, 1989).

1.14.1.4 Psychic Reality, Physical Reality, Fantasy, and Reconstruction

A specific issue concerning the status of clinical evidence has emerged recently and acquired the status of a major pseudoscientific debate: that of the validity of reconstructions of childhood traumatic experiences in therapy (Allen, 1995; Sandler & Fonagy, 1997). The classical psychoanalytic view emphasized the intrapsychic experience of the individual and was relatively uninterested in the "real" world. There was a silent assumption that the maturational stages of drives were of greater importance than so-called "accidents" of the environment. In, contrast many more recent theories, based on the study of adult pathology, posit that the actual behaviors of the caregivers towards the young child, as recovered in therapy, are of crucial significance to pathogenesis (e.g., Bowlby, 1958; Kohut, 1971; Sullivan, 1953; Winnicott, 1965a). Do such reconstructions have a truth value?

There is controversy in psychoanalysis (reflecting a culture-wide debate) concerning the "knowability" of early experience. Shengold (1989) links the controversy to the eighteenth-century debate initiated by George Berkley concerning the knowability of reality beyond the mind and its ideas. Masson (1984) fueled the current debate by his oversimplified *Assault on truth* in which he chastised Freud for having defensively abandoned and deliberately withheld evidence supporting the seduction theory of neurosis. Fredrick Crews (1995), in a series of vicious attacks, laid the blame on false memories of childhood sexual abuse at the door of psychodynamic approaches.

Faced with the unpleasant choice between risks of fabrication and denial, many psychodynamic clinicians adopted a hermeneutic approach to psychodynamic theory (e.g., Ricoeur, 1977; Spence, 1982). They repudiate a therapeutic search for the "real" past and embrace the criterion of internal coherence as the sole appropriate test of "truth." For example, Spence (1987) insists that psychoanalysis cannot claim to have a privileged position with regard to what was (Freud's archaeological past), and encountering the past in the therapeutic context is an act of creation of a "plausible" coherent narrative of our patient's life. Spence (1982) gives a critical warning: "Once stated, it (the narrative truth) becomes partially true, as it is repeated and extended, it becomes familiar; and as its familiarity adds to its plausibility, it becomes completely true" (p. 177).

There is no adequate resolution to this debate. In most of the cases the quality of the patient's recall leaves little room for doubt as to whether abuse actually happened. In others, there is room for doubt, and while the search for meaning is a ubiquitous aspect of human personality, the therapist must resist the temptation to give false meaning to current misery, anguish, and dejection by "discovering" a spurious historical account of early deprivation. There is no special expertise concerning the client's past which the therapist has access to beyond that which emerges as part of the relationship that the client creates with him or her in the therapeutic situation. This experience is unique, but whether it is an enactment of an actual past experience or the product of a defensively distorted representation of a self–other relationship, fantasized or real, is, at the current state of knowledge, impossible to determine.

1.14.1.5 Further Controversial Issues

The integration of psychodynamic theories into a coherent singular model has been undermined by a number of unresolved issues which include the following.

(i) The relative importance for pathogenesis of early childhood (infantile) experiences in contrast to later oedipal (age 4) and middle-childhood events (see Cooper, 1983).

(ii) The interpersonal vs. intrapsychic nature of human development: is the self an individual construction as conceived by Freud or is it an intersubjective unit derived from the interaction of mother and infant (Winnicott, 1956) or infant and his or her self-objects (Kohut, 1971)?

(iii) Is the basic constitution of human nature in conflict with its environment which must tame and inhibit it as Freud (1930) and Klein (1957) conceive, or is it potentially consonant with the external world provided that the latter does not thwart its desire for self-actualization, through malevolence and unresponsiveness (Winnicott, 1965b)?

(iv) Can stage theory used to describe child development be extended to the adult, as for example Erikson (1950a) suggests, or is development better seen as the repetition of characteristic positions throughout an individual's life course (Klein, 1952)?

(v) If development continues throughout life, is such development merely a shift of concerns (e.g., Tyson & Tyson, 1990), or does the structure of the mind remain flexible, open to drastic alteration and the generation of new structures, more or less, throughout the lifespan (e.g., Emde, 1985)?

1.14.2 FREUD'S MULTIPLE MODELS OF THE MIND

1.14.2.1 The Affect-trauma and Topographic Frames of Reference

His work with hypnosis led Freud (Freud & Breuer, 1895) to believe that he had discovered the etiology of neurosis in the actual event of childhood seduction. In this conception the interpersonal event of the early trauma was represented in a distorted form in the neurotic symptom. For example, a child of eight with hysterical blindness may have achieved relative internal safety by "shutting his eyes" to the memory of having witnessed his mother's rape. In this model he assumed little by way of a mental apparatus, simply the physical conversion of emotional energy generated and damned up by trauma. The therapeutic intervention was appropriately principally cathartic.

The discovery of free association and dream interpretation as clinical-research tools prompted Freud to seek a more elaborate psychological model of mental disorder. The three-layered topography of the mind—the systems unconscious, preconscious, and conscious—motivated by sexual (and later aggressive) instincts gave adequate accounts of the therapeutic value of the recovery of repressed, unconscious sexual fantasies of early childhood. The turning away from his seduction hypothesis in favor of his second model, however, discredited psychoanalytic theory as a social theory of psychopathology and therapy. It led Freud (1905b) to attempt to explain all actions in terms of the failure of the child's mental apparatus to deal adequately with the pressures of a maturationally predetermined sequence of drive states. Adult psychopathologies, as well as dreaming, jokes, and parapraxes, were seen as the revisiting of unresolved childhood conflicts over sexuality (Freud, 1900, 1901, 1905a). He now saw anxiety as arising from the failure of repression of unacceptable sexual wishes (Freud, 1905b, p. 224).

1.14.2.2 The Structural Frame of Reference

The fundamental influence of the social environment again found a preeminent place with the third major shift in Freud's thinking (Freud, 1920, 1923, 1926). Freud's model now recognized three mental structures: the id (sexual and aggressive instincts), the ego (the self as agent) and the superego (a personalized source of internal and social moral injunction). This new structural theory was to survive long after Freud because of the compelling fit with clinical observational data. In his dual instinct theory, he describes the child's struggle with innate destructive and self-destructive forces as well as sexual conflicts (Freud, 1920). Freud (1926) also revised his theory of anxiety from one caused by inhibited biological drives, to a psychological state linked to the perception of internal (instinctual or moral) or external danger. The danger situation was specified as the fear of helplessness resulting from loss (loss of the mother, her esteem, loss of a body-part, or loss of self-regard).

This revision achieved two goals: (i) it restored adaptation to the external world as an essential part of the psychoanalytic account, and (ii) it recast the theory into cognitive terms, making way for a representational account of the mind (see below). Freud's theory of pathology was now of inadequately resolved conflict (Freud, 1938). The hypothesis that conflicts within the human mind are essentially of three kinds (wish vs. moral injunction, wish vs. reality, and internal reality vs. external reality) has had extraordinary explanatory power. In particular, the ego's capacity to create defenses which organize characterological and symptomatic constructions as part of an adaptive process, became the cornerstone of psychodynamic theorization and clinical work in the USA (Hartmann, Kris, & Loewenstein, 1946) and Britain (A. Freud, 1946) for 30–40 years.

1.14.2.3 Some Limitations of Freud's Models

The limitations of Freud's models are manifold, and many of these were pointed out by Freud's contemporaries who moved away from organized psychoanalysis. For example, Jung's controversial rejection of libido theory drew attention away from the undoubted advances he made in the understanding of narcissism and his development of a theory of the self throughout the life-cycle (Jung, 1916, 1923). Other important omissions were:

(i) the cultural and social context, including interpersonal aspects of development;

(ii) the significance of the experiences of infancy;

(iii) the developmental significance of the real behavior of the real parents, including transactional processes;

(iv) the role of attachment and safety in development alongside the role of instinctual drives;

(v) the synthesizing function of the self;

(vi) the importance of the nonconflictual aspects of development;

(vii) the role of the experiences of adolescence and adulthood in shaping normal and pathological development.

1.14.3 THE STRUCTURAL APPROACH

1.14.3.1 American Ego Psychology

1.14.3.1.1 Hartmann's classical model

Freud's third, structural, model was refined and advanced in the ego psychology of Heinz Hartmann and his colleagues. Hartmann et al. (1946) postulated an initial undifferentiated matrix which contains the individual's endowment and from which both the id and the ego originate. They also introduced the concept of an "average expectable environment," which affirmed the importance of the parental contribution to development, and outlined a scheme for the phase-specific maturation of autonomous, conflict-free ego functions, accommodating both environmental and maturational influences upon personality development.

Development of the ego is driven by a maturational pull, whereby independently emerging components and functions come to be linked, forming a coherently functioning organization (the ego) which is more complex than the sum of its parts (Hartmann, 1958, 1964c). Stages of ego development represent nodal points at which "fixation" may occur and to which, under the stress of conflict, the individual may return. For example, obsessive-compulsive disorder is seen by structural theorists (Arlow & Brenner, 1964) as a regression to the phase of ego functioning characteristic of the two-year-old (magical phenomenalism, repetitive, ritualistic behaviors).

Hartmann (1958) accurately claimed that psychoanalysts frequently used the developmental point of view in a reductionist way. His concept of the "change of function" (Hartmann, 1958, p. 25) and secondary autonomy (Hartmann, 1964a) pointed to how the same behavior in the adult may serve quite different functions and is likely to be functionally independent from the childhood wish from which it may have originated. The persistence of dependent behavior in adulthood cannot be treated as if it were a simple repetition of the individual's early relationship with the mother. Adult behaviors should be seen as having multiple functions (Brenner, 1979). The failure to recognize this has been termed the "genetic fallacy" (Hartmann, 1964b, p. 221). It is not uncommon for primitive modes of mental functioning (e.g., splitting or identity diffusion) in severe personality-disordered individuals to be treated as evidence for the persistence or regressive recurrence of early pathogenic developmental experiences. Yet, their reemergence in adult mental functioning may be linked to later or persistent trauma (see Tyson & Tyson, 1990).

1.14.3.1.2 Modern structural theory

Partly in response to the challenge of object relations theories (see below), there has been a revival of the structural theory in psychoanalysis. Modern structural theory (see, e.g., Boesky, 1988) retains the tripartite model of id, ego, and superego, but dispenses with concepts of psychic energy and other problematic notions. The theory takes as its central premise the ubiquitous nature of internal psychic conflict (see Brenner, 1982). Brenner suggests that all mental contents (thoughts, actions, plans, fantasies, and symptoms) are compromise formations which are best conceived of as multiply determined by components of conflict:

(i) a drive derivative, conceived of in this context as an intense, personal childhood wish for gratification;

(ii) unpleasure in the form of anxiety, or depressive affect, and their ideational contents of object loss, loss of love, or castration, associated with the drive derivative;

(iii) defense, which functions to minimize unpleasure; and

(iv) manifestations of superego functioning such as guilt, self-punishment, remorse, and atonement.

Self and object representations, in this scheme, are the result of compromise formation which in their turn effect further compromises between the tendencies above.

1.14.3.1.3 Extensions of the ego psychology model

(i) Contribution of René Spitz

Spitz (1959) was one of the first "empiricists" of the psychoanalytic tradition. He proposed that major developmental shifts in psychological organization, marked by the emergence of new behaviors and new forms of affective expression (e.g., social smiling), occurred when functions are brought into new relation with one another and are linked into a coherent unit. He drew attention to the meaning of new forms of emotional expression ("psychic organizers") such as the smiling response (2–3 months), marking the initial differentiation of self and object, eight month anxiety which indicates differentiation amongst objects, especially of the "libidinal object proper" and the assertion of self in the "no" gesture between 10 and 18 months. The way in which these organizers herald dramatic changes in interpersonal interactions was elaborated in a highly influential series of papers by Robert Emde (1980a, 1980b, 1980c).

Spitz (1965) was also a pioneer in seeing the infant's and child's human partner as "quickening" the development of the child's innate abilities and mediating all perception, behavior, and knowledge. Spitz (1957) saw self-regulation as arising out of the mother–infant relationship. Psychoanalytic observational studies repeatedly showed the ways in which constitutional, early environmental, and interactional factors contribute to the structuring of the self-regulatory process leading to adaptation or maladaptation (Spitz, 1959). The mother's emotional expression at first serves a "soothing" or "containing" function which facilitates the restoration of homeostasis and emotional equilibrium. Later, the infant uses the mother's emotional response as a signaling device to indicate safety. Later still, the infant internalizes the affective response and uses his or her own emotional reaction as a signal of safety or danger (Emde, 1981). These concepts were to become a core part of attachment theory (see below).

(ii) The work of Erikson

An important extension to the classical model was proposed by Erik Erikson (1950b). His concern was the interaction of social norms and biological drives in generating self and identity. He described eight developmental stages spanning the whole of life (e.g., the "identity crisis" of adolescence) determined by biologically caused life events that were thought to disturb the equilibrium between drives and social adjustment. Personality would be arrested if the developmental challenge was not mastered through the evolution of new skills and attitudes, compromising later developmental stages. Erikson was remarkable among psychoanalysts for his attention to cultural and family factors and his extension of the developmental model to the entire life-cycle. His theory introduced plasticity to the psychoanalytical model, as well as attributing critical importance to the need for a coherent self concept fulfilled in a coherent, supportive social milieu.

(iii) The work of Edith Jacobson

Edith Jacobson (1964) advanced the idea that the infant acquires self and object images with good (libidinal) or bad (aggressive) valences, depending on experiences of gratification or frustration with the caretaker. In order to clarify and distinguish the concepts of ego, self, and self-representation, she used the term "self-representation" to stress the notion of the self and object as they were experienced, as distinguished from external objects. She stated that the ego was a structure in contrast to the self, which is the totality of the bodily and psychic person, and defined self-representation as "the unconscious, preconscious, conscious, intra psychic representation of the bodily and mental self in the system ego" (p. 19). She assumed that introjections and identificatory processes replaced the state of primitive fusion, and, through these, traits and actions of objects became internalized parts of self-images. She was particularly concerned with superego formation, which she saw as initially polarized between pleasure and unpleasure, then by issues of strength and weakness, and finally as the internalization of ethical considerations which regulated self-esteem as well as behavior.

1.14.3.1.4 The classical model of the neurosis

The classical model of the neurosis is well known and will not be elaborated in detail here. Childhood sexual wishes are presumed to arouse conscious repugnance when experienced in adulthood because of their aggressive or incestuous content. They can reach awareness only when disguised. The neurotic compromise represents a disguised (id) derivative of childhood sexuality, the ego's defense, and signal anxiety marking the ego's experience of internal danger. It unifies the wish and the reaction against it in a part of the personality that is experienced as separate (ego dystonic). This is characterized by a subjective experience of punishment, suffering, and irritation which originates from, and is designed to placate, the superego. The sequence of events associated with neurotic reaction are: (i) frustration, (ii) regression, (iii) internal incompatibility, (iv) signal anxiety, (v) defense by regression, (vi) return of the repressed, (vii) compromise formation and symptomatic disorders.

Specific neurotic reactions are thought to reflect regression to particular developmental fixations and characteristic modes of compromise formation. Regression occurs because the compromise formations established to resolve the oedipal dilemma (the so-called ubiquitous infantile neurosis) was challenged by internal or external stressors. In conversion hysteria, the compromise achieves dramatic representation in somatic form and reflects an oral or phallic fixation. In obsessional neurosis, it is assumed that the ego binds anal sadistic and aggressive drive derivatives into forms of secondary process thinking (e.g., ruminations, obsessional doubts, etc.), but it is developmentally unable to neutralize these drive derivatives and therefore aggression and anal concerns will be transparent and arouse massive anxieties (see e.g., Fenichel 1945; Glover 1949). In phobias, the fear is externalized and is wholly psychological in its presentation, but in unconscious content, may reflect quite similar developmental concerns. The process remains largely an internal one; the neurotic compromise that results in obsessions is located in the thought processes themselves.

Psychoanalytic writers have noted problems with this concept of neurosis. As the focus shifted in the psychoanalytic literature to preoedipal development—which had been the primary focus for the Kleinian group—questions were raised about the obligatory link between adult neurosis and oedipal issues. Outside of psychoanalysis, the concept of neurosis has been assaulted with much greater ferocity. The concept of neurosis has been discredited as the trend toward descriptive clarity and reliability in diagnosis gains the upper hand, together with the push toward atheoretical, operational definitions of psychiatric disorders and biological explanations of pathogenesis. Derided as vague, unreliable, impossible to verify empirically, overinclusive, and tied to an obsolete theory, neurosis has been excluded from psychiatry's official diagnostic classifications. Despite this slight, neurosis refused to disappear.

1.14.3.1.5 The classical model of severe psychopathology

Whereas notions of neurotic pathology have, on the whole, evolved little since the structural theory of Freud, models of personality disorder have become "paradigmatic" of various psychoanalytic models. As subsequent sections will illustrate, extremes of personality types, now embodied in the psychiatric diagnostic schemes as the second axis of psychiatric diagnosis (American Psychiatric Association, 1994), are formulated radically differently across different theoretical models.

Structural theory distinguishes those character disorders which resemble neurosis in terms of dynamic considerations and those which reflect a non-neurotic pattern based on structural deficit (see Waelder, 1960). The so-called character neurosis (a concept introduced by Alexander (1930)) of the former category is assumed to be dynamically similar to neurosis with the exception that compromise formations are not split off from the ego and thus the symptoms are not experienced as ego alien or ego dystonic. More severe personality disorders—for example, narcissistic personality disorder—are regarded as a consequence of a developmental arrest, deviation, or disharmony. The structural view of such cases tends to be in terms of faulty ego development (see Gitelson, 1955; Rangell, 1955). Important ego functions such as reality testing, anxiety tolerance, and stable defenses are impaired, while others appear to retain their integrity, thus giving the patient a semblance of normality.

(i) Structural model of borderline personality disorder

Knight (1953) was the first to propose a comprehensive model of personality disorder in terms of ego functions impaired by traumatic development. Among the ego functions he considered were: "integration, concept formation, judgment, realistic planning, and defending against eruption into conscious thinking of id impulses and their fantasy elaborations" (p. 6). Erikson (1959a, 1959b) in his epigenetic sequence of identity formation described the syndrome of identity diffusion, which he saw as reflecting deficiencies in a sustained sense of self-sameness, temporal continuity of self experience, and a feeling of affiliation with a social group of reference. Jacobson (1964) drew attention to how these individuals, at times, experience their mental functions and bodily organs not as belonging to them, but as objects which they wish to expel. They may also attach their mental and body self to external objects. She saw them as retaining an "adolescent fluidity of moods" (Jacobson, 1964, p. 159).

(ii) Structural theory of antisocial personality disorder

Aichhorn (1935) was the first psychoanalyst to seriously work with delinquent individuals. In his influential formulation he posited a failure of progression from the pleasure principle to the reality principle in conjunction with a malformation of the superego. Reich (1933) suggested that the ego kept the superego at a distance, causing it to be isolated and therefore unable to prevent the individual from yielding to an impulse.

Johnson and Szurek (1952) suggested that superego lacunae (lack of superego in certain circumscribed areas) was the nature of the superego pathology. Such gaps in the superego (which is regarded as an internalization of parental morality) were thought to occur because of the parents" unconscious wish to act out forbidden impulses; the child is unconsciously encouraged by the parents to act in amoral ways, but is consciously discouraged from doing this. Lampl-de-Groot (1949) suggested that the balance of the superego and the ego ideal explained why certain individuals became neurotically depressed, whilst others became antisocial. The former corresponds to a severe superego and strong ego ideal, whereas the latter is a consequence of a menacing superego and a weak ego ideal.

These ideas exemplify the strengths and weaknesses of structural theory. While the suggestions fit well with clinical observation and phenomenology, their explanatory value is sustained by the assumption of the *actuality* (concrete existence) of the psychic structures proposed by Freud, yet Freud was clear that his metapsychology was, in essence, metaphoric (Schafer, 1976).

1.14.3.2 The Structural Developmental Tradition

1.14.3.2.1 Anna Freud

Anna Freud represented, and to some degree continues to represent, the structural psychodynamic model, although her work embodies a unique psychodynamic approach to both psychopathology and clinical work. Anna Freud (1965) was one of the first coherently to adopt a developmental perspective on psychopathology. She provided a comprehensive developmental theory using the metaphor of developmental lines to stress the continuity and cumulative character and transactional nature of childhood development. For example, aspects of the child's relationship to the mother may be described as a line moving from "dependency to emotional self-reliance to adult object relationships," "from irresponsibility to responsibility in body management." Unevenness of development may be regarded as a risk factor for psychiatric disturbance, and thus developmental lines have etiological significance. A child's problem may be understood in terms of an arrest or regression in terms of a particular line of development (A. Freud, 1965).

The clinical implication of her formulation is that in addressing disturbance the psychodynamic clinician should focus not only on the determinants of symptomatic aspects of the disorder, but also on offering "developmental help" to the child and restore him/her to the "path of normal development" (A. Freud, 1981).

(i) Developmental approach to anxiety

The developmental approach is helpful to distinguish emotional experiences at different levels of maturity. Yorke, Wiseberg, and Freeman (1989), following in Anna Freud's tradition, differentiate a nameless terror, a principally somatic experience developmentally rooted within the undifferentiated mother–baby unit, psychic panic dating back to a time when the infant is capable of mental experience but not yet capable of affect regulation, pervasive anxiety where there is fear of helplessness but already some capacity to control affect and a signal anxiety which functions to circumvent over-arousal by activating defenses. While this theory is compelling, there is surprisingly little observational data to support the notion that early stages of development are associated with more pervasive and intense emotional experience (Harris, 1994; Stern, 1985).

(ii) Anna Freud's approach to severe psychopathology

Anna Freud agrees with structural theorists in regarding severe personality disorders as reflecting structural deficits such as defects in reality testing, the dominance of primitive defenses, limited capacities for anxiety tolerance, poor superego development, etc. She explains these as developmental disturbances (deviations or disharmonies). For example, Yorke et al. (1989) suggested that the inadequate response by the mother to an infant's instinctual needs creates dangers and external conflict. Such disharmony of need and external environment will be most intensely felt when structuralization is not yet ready to sustain the pressures caused by the internal and external stresses thus created. Ego development will suffer because the internalizing and identificatory processes will be specifically threatened. A constant representation of the other, for example, may not develop if the early relationship with the mother is disrupted by trauma. The failure to achieve structured compromise produces the labile character of borderline and other personality disturbances. Narcissistic character disorder is seen as rooted in early emotional deprivation which compromises the process by which objects (representations of people) are invested with instinctual energy. The individual attempts to identify with the frustrating and disappointing object, providing a focus for libidinal cathexis that heightens narcissism and cathexis of the self (egocentrism).

1.14.3.2.2 Margaret Mahler

Margaret Mahler (Mahler, 1968; Mahler, Pine, & Bergman, 1975), using a naturalistic observational strategy of infants and young children, evolved an influential psychodynamic model. Of particular relevance was her description of the separation–individuation process beginning at 4–5 months and ending in the third year of life. Overall, the process is one of moving away from a symbiotic unity with the caregiver to a consolidation of individuality. A critical subphase is the 6–10 month period commencing in the latter half of the second year of life when the child's greater awareness of separateness intensifies separation anxiety at the same time as promoting a greater need to be separate from her. The caregiver's handling of this subphase ("rapprochement") is thought to be critical, as the mother must combine emotional availability with a gentle push towards independence. If the balance of availability and push toward independence is weighted too much on either side, the infant may become desperately dependent and clingy, experience great difficulty in investing his or her environment with sufficient interest, and his or her pleasure and confidence in his or her own functioning will be impaired (Settlage, 1977).

Masterson (1972, 1976) believes that borderline patients experience a deep conflict between the wish for independence and the threat of loss of love and thus search for a clinging tie with a mother substitute. Such a tie will temporarily ensure a feeling of safety, but any wish for self-assertiveness will present him with the terror of abandonment. A lifelong and vicious cycle of brief blissful unions, ruptures and emptiness and depression will ensue.

He discusses "abandonment depression" as the consequence of the borderline child's quest for separation from the withdrawing or aggressive maternal object who in turn, for pathological reasons of her own, wishes to keep the child in a symbiotic relationship with her. The patient develops a fear that "his very existence is dependent ultimately upon the presence of need gratifying and life sustaining others" (R. Klein, 1989, p. 36). The withdrawing and rewarding object representations are kept rigidly separate to maintain the possibility of symbiotic union with the rewarding object and to ward off abandonment depression. The borderline individual's dramatic response to actual separation is thus explained by his incomplete separation from his objects, with the psychological experience of separation becoming equivalent to a loss of a part of the self. Borderline patients" common vigorous pursuing of their therapists at home, in their holidays, or in other professional activity can be understood in this way.

There is strong evidence, particularly from retrospective studies, that borderline patients appear to have parents with mental illness, personality disorder, and drug abuse (e.g., Ogata, Silk, & Goodrich, 1990a) and frequently report having experienced family violence, physical and sexual abuse and other trauma (Ogata et al., 1990b; Paris, 1992; Paris & Zweig-Frank, 1992). Evidence is only indirectly supportive of Masterson's view because data are retrospective and because experiences of maltreatment usually pertain to teenage years rather than early development. More recent developmental evidence tends to disconfirm Mahler's assumption of an early self-object merger postulated by Mahler for the first half of the first year of life (see Gergely, 1991). Nevertheless, Mahler's emphasis on the second year of life has been helpful in identifying what may be a critical period for the emergence of awareness of mental states in both self and others, in other words an appreciation of human individuality (Dunn, 1996; Fonagy et al., 1995a).

1.14.3.2.3 Sandler's reframing of the structural model

Bringing a unique blend of developmental and psychological sophistication to the classical structural model, Joseph Sandler can be said to have achieved the most coherent integration between object relations theory and the structural model. Of greatest clinical relevance is Sandler's meticulously systematic work to move psychoanalysis to the level of representations and affects (Sandler & Rosenblatt, 1962). Sandler reframed the psychodynamic theory of motivation, highlighting the role of affects rather than drive state (Sandler, 1985). A key concept is that of safety, an aspect of phenomenal experience which is intensely sought throughout development (Sandler, 1987). Past patterns of relationships and associated emotional states are thought to be actively recreated by the individual because of the associated sense of safety, even when these patterns appear maladaptive (Sandler & Sandler, 1987). An important aspect of this formulation is the distinction between past and present unconscious (Sandler & Sandler, 1987). According to Sandler, the etiological and appropriate therapeutic focus of psychodynamic approaches is not early childhood experience and primitive fantasy, but rather the current derivatives of these experiences—the ways in which unconscious wishes and fantasies of the past create dynamic conflicts and elicit defensive maneuvers in the present.

The direct influence of one individual on another within Sandler's scheme (Sandler, 1976) is accounted for by the evocation of particular roles in the mind of the other person who is being influenced. The behavioral role of the influencing person is seen as crucial in eliciting a complementary response from the other participant. In this way, within the individual's present unconscious, infantile or childhood patterns of relationships may be enacted or actualized in adult interpersonal interaction.

1.14.3.3 Strengths and Weaknesses of the Model

The quasiphysiological character of the original model has been the subject of intense criticism (e.g., G. Klein, 1976a). Many of the classical explanations of psychopathology appear to reify the hypothetical constructs of ego, superego, and id in order to produce credible explanations. To talk of lacunae in the superego as a cause of antisocial behavior is only plausible if we assume that such a structure actually exists in the mind rather than conceiving of it as a metaphor evocatively summarizing a certain range of behaviors. Stripped of reification, many structural accounts are revealed as circular. For example, an absence of the ego function of anxiety tolerance, stripped of metapsychological jargon, means no more than such an individual gets very upset in situations that induce anxiety.

This is not to say that either the clinical or the metapsychological observations are without value. In reading many of these classical papers, even with the sophistication freely provided by hindsight, it is hard not to be impressed by clinical sensitivity and the remarkable integrative abilities of many of these pioneers who, in many instances, decades before other professionals had highlighted key characteristics of important disorders.

1.14.4 THE OBJECT RELATIONS APPROACH

1.14.4.1 Introduction to the Object Relations Approach

Greenberg and Mitchell (1983), in their definitive review of theoretical work on object relations theory to date, use the term to denote theories "concerned with exploring the relationship between real, external people and the internal images and residues of relations with them and the significance of these residues for psychic functioning" (p. 14).

The rise of object relations theories signaled a change of focus in psychodynamic theories. Intrapsychic conflict, particularly conflicts relating to the sexual and aggressive drives, and the central organization of oedipal compromises and the complementary influences of biological and experiential forces in development are no longer the cornerstones of psychodynamic theory, a change which has been regretted by some (see, e.g., Spruiell, 1988). Regardless of particular theoretical models, psychodynamic thinking seems to have moved in the 1980s towards a phenomenologically based perspective which emphasizes the individual's experience of being with others and with the therapist during clinical work (see, e.g., Loewald, 1986). The clinical emphasis upon experience inevitably drives theory away from a structural mechanistic model towards what Mitchell (1988) broadly terms "relational theory." Patients in treatment express themselves in terms of relationships (Modell, 1990), and the move towards object-relations-based psychodynamic theories may thus be seen as led by an increasing demand on clinicians to explore clinical phenomena from the point of view of the patient.

There are several assumptions which object relations theories share (see Fonagy, Target, Steele, & Gerber, 1995b). These include:

(i) that severe pathology has pre-oedipal origins (i.e., the first three years of life);

(ii) that the pattern of relationships with objects becomes increasingly complex with development;

(iii) that the stages of this development represent a maturational sequence which exists across cultures but which nevertheless may be distorted by pathological personal experiences;

(iv) that early patterns of object relations are repeated, are in some sense fixed and reproduced throughout life;

(v) that the developmental continuum of these relationships is isomorphic with the continuum of pathology;

(vi) that patients" clinical reactions to their therapist provide a window for examining healthy and pathological aspects of early relationship patterns.

There exist, however, considerable differences between psychoanalytic theories in terms of the rigor with which the problem of object relationships is tackled. Friedman (1988) differentiates between hard and soft object relations theories. Hard theorists, in which he includes Melanie Klein, Fairbairn, and Kernberg, see much hate, anger, and destruction, and dwell on obstacles, illness, and confrontation, whereas soft object relations theorists (Balint, Winnicott, and Kohut) deal with love, innocence, growth needs, fulfillment, and progressive unfolding.

1.14.4.2 The Klein–Bion Model

Melanie Klein's (1964, 1975a, 1975c) model combines the structural model with an interpersonal, object-relations model of development. It was with the contributions of Bion (1962) that the theory became a fully relational object relations model. Until 1935 Klein was basically working with the theoretical framework of Freud and Karl Abraham. The 1935 (1975a) and 1940 (1984) papers on the depressive position, the 1946 (1952) paper on the paranoid-schizoid position, and the 1957 book *Envy and gratitude* established her as the leader of an original psychoanalytic tradition.

1.14.4.2.1 Some essential Kleinian concepts

(i) The two basic positions

In the Kleinian model the human psyche has two basic positions: the "paranoid-schizoid" and the "depressive" (Klein, 1952). In the paranoid-schizoid position, the relationship to the object (the caregiver) is to a part object, split into a persecutory and idealized relationship, and similarly the ego (the self) is split. In the depressive position, the relation is to an integrated image, both loved and hated, and the ego is more integrated. The paranoid-schizoid superego is split between the excessively idealized ego ideal characterized by the experience of narcissistic omnipotence and the extremely persecutory superego of paranoid states. In the depressive position the superego is a hurt love object with human features.

The paranoid-schizoid position is the infant's earliest relationship with the external world and is dominated by innate internal representations (Klein, 1975c, p. 248). The infant's initial attempt at organizing internal and external perceptions is dominated by splitting. In this way he attributes all goodness, love, and pleasure to an ideal object, and all pain, distress, and badness to a persecutory one. All good feelings of affection and desire are aimed at the idealized good object which the infant wishes to possess, take inside (introject) and experience as himself (identify with). Negative affect (hatred, disgust, etc.) is directed to the persecutory object and projected (externalized) onto it since the infant wishes to rid himself of everything that is experienced as bad and disruptive. The infant's mental life is envisaged as extremely labile; good rapidly turns into bad, the bad gets worse, and the good gets increasingly idealized. Each external object has at least one good and one bad representation, but both representations are partial and not the whole person.

The depressive position is marked by the infant's capacity to perceive the mother as the whole person who accounts for both his good and bad experiences and is seen by Klein (1975a, p. 310) as the central achievement or process in the child's development. At this moment the infant becomes aware of his or her own capacity to love and hate the parent. The discovery of this ambivalence and the growing capacity to recognize absence and potential loss of the attacked object, opens the child to the experience of guilt over hostility to a loved object. This is what Klein calls "depressive anxiety" as distinct from the persecutory anxieties of the earlier paranoid-schizoid position. Working through the characteristic experiences of the depressive position brings with it reparative feelings. The psychic pain associated with the integration is so great that it can lead to defenses characteristic of this position, including manic or obsessional reparation, total denial of damage or contempt. Segal (1957) links the capacities for symbolization and sublimation to depressive reparation. Once the object is perceived as a mentally independent entity, it is seen as having desires, wishes, loyalties and attachments of his/her own, and oedipal concerns about the feeling about the "third" may arise (Britton, 1989).

(ii) Projective identification

For Melanie Klein (1957), projective identification is an unconscious infantile phantasy by which the infant is able to relocate his persecutory experiences by separating (splitting) them from his self representation and making them part of his image of a particular object. Disowned unconscious feelings of rage or shame are firmly believed by the infant to exist within the mother. By acting in subtle but influential ways, he may achieve a confirming reaction of criticism or even persecution. The phantasy of a magical control over an object may be achieved in this way. Projective identification is not a truly internal process and involves the object who may experience it as manipulation, seduction or a myriad of other forms of psychic influence. Spillius (1992) suggests the use of the term "evocatory projective identification" to designate instances where the recipient of projective identification is put under pressure to have the feelings appropriate to the projector's phantasy.

Bion (1959) pointed to the general necessity for projective identification in infancy, a time when the individual is ill-equipped to absorb its impressions of the world. By projecting these elements into another human mind (a container) that has the capability to accept, absorb and transform them into meanings, his mental survival was ensured. The absence of a suitable container makes projective identification a pathogenic process of evacuation. Bion (1962) discussed the significance of the mother's capacity mentally to "contain" the baby and respond to the infant emotionally and in terms of physical care in a manner that modulates unmanageable feelings. Capable caregivers are likely to experience and transform these feelings into a tolerable form which probably combines mirroring of intolerable affect in combination with emotional signals, indicating that the affect is "contained," that is, under control (Bion's "alpha function"). The infant can cope with, accept and re-internalize what was projected, thus creating a representation of these emotional moments of interaction with the caregiver which is tolerable, in place of his original experience which was not. In time he internalizes the function of transformation and will have the capacity to contain or regulate his own negative affective states. The nonverbal nature of this process implies that physical proximity of the caregiver is essential.

(iii) Envy and gratitude

Klein (1957) suggests that early, primitive envy represents a particularly malignant form of innate aggression. This is because, unlike other forms which are turned against bad objects, already seen as persecutory, envy is hatred directed to the good object and arouses a premature expression of depressive anxiety about damage to the good object. The child may resent the inevitable limitations of maternal care, cannot tolerate the mother's control over it and would prefer to destroy it rather than experience the frustration. This interferes with the primal differentiation of "good" and "bad." Excessive envy is seen as interfering with working through the paranoid-schizoid position, and ultimately it becomes the developmental precursor of many forms of confusional states (Rosenfeld, 1950).

(iv) Primitive transferences

In therapy, Kleinians prefer to work exclusively with interpretations, primary transference interpretations aimed at the patient's current anxieties. They work psychodynamically with relatively severe disorders, and stress early interpretation of negative transferences derived from the paranoid-schizoid position. Kleinians have contributed enormously to our understanding and use of projective and introjective aspects of the countertransference (e.g., Racker, 1968, chap. 6.5). A pioneer in this respect was Bion (v1962), for whom transference and countertransference are about the transfer of intolerable mental pain by projective identification, originally from infant to mother, and in the treatment situation from patient to therapist.

1.14.4.2.2 Kleinian approach to neurosis

Psychological illness reflects the predominance of the paranoid-schizoid position whereas health implies the stabilization of the depressive framework, promoting development and maturity. In neurotic states, the transition between schizo-paranoid and depressive positions is perceived as partial; the superego contains both paranoid and depressive features, manifesting as persecutory guilt. The most common anxiety is the fear of guilt and the possibility of the loss of the loved object. If the depressive position is not approached, the anxiety will be about fragmentation, annihilation, and persecution, and the reality sense will be grossly distorted by projections. This picture is more fitting of patients with severe personality disorders, such as borderline or narcissistic personalities.

Bion (1962) outlined the processes that can lead to pathology in the paranoid-schizoid position. He names two factors: (i) deficiencies in the mother's capacity for receiving the child's communications through projective identification ("reverie") (see also the formulation of primary maternal preoccupation by Winnicott, 1962), and (ii) overwhelming envy in the infant.

1.14.4.2.3 Rosenfeld's developmental model of narcissism

Rosenfeld (1964, 1988) views narcissistic states as characterized by omnipotent object relations and defenses which deny the separateness and identity of the object. He stresses the destructiveness of the narcissist's relationship with others, his ruthless use of people, his denial of their value to him and yet need for them. To recognize the object would mean recognizing the object's control over "goodness" and his vulnerability to separation from it. By introjective identification the narcissistic individual lays claim to the good part of the object and in fantasy owns it. Projective identificatory processes help him deposit his own perceived inadequacies in others whom he can then denigrate and devalue. To deal with their envy, narcissists devalue their objects (their therapists, their spouses). Denigration serves the function of avoiding recognizing goodness in others, which the person finds threatening to his or her own delusional state of self-idealization as in possession of the "good breast."

1.14.4.2.4 Kleinian approach to severe psychopathology

Klein's (1975b) formulations were vital to the understanding of borderline personality conditions. The paranoid-schizoid condition is the template for borderline personality functioning:

(i) In object relationships, splitting predominates over repression, and others are seen as either idealized or denigrated. There is no real knowledge of the other and the inner world is populated by parts (or caricatures) of the object.

(ii) Since the depressive position is avoided and all badness is pushed into the object, there can be no genuine sadness, mourning, or guilt.

(iii) Projective identification predominates; communication cannot be meaningful, and the object is manipulated by being forced to experience unacceptable aspects of the borderline individual's personality.

Splitting is both a cause and a consequence of the borderline person's difficulty in maintaining an ambivalent, balanced view of both self and object, which in Kleinian theory would require the acknowledgment of his experientially overwhelming destructive potential. With more than one object at their disposal, borderline individuals may externalize their incapacity to integrate good and bad objects, by polarizing people working with them and constantly attacking the links between them (Main, 1957).

1.14.4.2.5 Strengths and weaknesses of the Klein–Bion model

Critics of Kleinian formulations focused on the unwarranted assumption of the early development of higher-order cognitive and perceptual capacities (see Hayman, 1989; Yorke, 1971). For example, the mechanism of projection in the paranoid-schizoid position assumes a differentiated sense of self and other, since were there not yet a differentiation of inner and outer, self and other, it would seem impossible to displace the experienced source of negative affect from the self onto an external object. By now there is detailed documentation of remarkably abstract and complex cognitive capacities of the human infant. In particular there exists compelling evidence that the infant differentiates between the self and other; for example, at five months a baby differentiates a video image of his or her own legs moving from those of another baby's legs (Bahrick & Watson, 1985).

Kleinian assumptions concerning infantile mental capacities extend beyond the representation of self and other and assume causal reasoning on the part of the infant. Studies of the perception of causality in infancy (Bower, 1989) and causal reasoning (Golinkoff, Hardig, Carlson, & Sexton, 1984) suggest that the human mind is innately predisposed to impose a causal structure on perceptual experience. If the Kleinian model of development is inaccurate, it is so because it under- rather than overestimates the child's cognitive capacities.

More pertinent current criticisms address the "fuzziness" (Greenberg & Mitchell, 1983, pp. 148–149) of Klein's descriptions of the development of mental structure. The emphasis upon "phantasy" as the building block of mental structure confounds the experiential and nonexperiential aspects of mental functioning (see Sandler & Joffe, 1969). Klein and post-Kleinians have moved mental structuralization into the experiential realm, which carries with it the advantage of experience-nearness for clinicians, and rids theory of much reified pseudoscientific terminology. However, it bypasses essential questions concerning the nature of the mechanisms underpinning the organizations of mental functions.

1.14.4.3 The Fairbairn–Winnicott Model

1.14.4.3.1 Major innovations of the "Independent" tradition

The term "Independent" is not a misnomer; unlike other psychoanalytic schools, this tradition is the work of a number of analysts working independently, without a single leader or reference point. Not surprisingly, the approach lacks the internal coherence of a more unified group. Fairbairn (1954) was the systematic theory builder, but major contributions came from Winnicott (1958a) and Balint (1959, 1968).

The focus of the Independents on early development led them away from a libidinally driven structural model to develop a "self–object" theory. Parts of the self are seen as in dynamic interaction with each other and with complementary internal and external objects. Fairbairn (1954, 1963) envisages the self as a crucial agent of motivation; there is no emotion without the self and no self without emotion (Rayner, 1991). Fairbairn (1952) wrote "The libido is primarily object-seeking (rather than pleasure seeking as in classical theory)" (p. 82). Winnicott (1958a) described the powerful desire to develop a sense of self and an equally strong impulse to hide or falsify it. The loss of optimal intimacy with the primary object will give rise to "splitting" in the self (the ego). Conflicting multiple ego–object systems are seen as the developmental roots of psychopathology.

Winnicott (1965b) saw the child as evolving from a unity of infant and mother. The mother holds the infant, both actually and figuratively, and so gives cohesion to his or her sensorimotor elements. Holding is based on "comprehension," holding in mind, of the infant's mental state. Her primary maternal preoccupation (a partial withdrawal from activities other than the baby and a state of heightened sensitivity to her own self, her body, and the baby) helps the mother achieve a state whereby the baby is provided with the "illusion" that the mother responds accurately to his or her gesture because she is his/her own creation (a part of him/her).

Winnicott (1971a) saw object relating as arising from the experience of magical omnipotence. The breast in phantasy is the infant's creation. When self–object differentiation is incomplete, object representations are best designated self-objects. The infant's muscular "attacks" on the mother, and her survival of them, facilitate the development of the self and the mother's release from omnipotent control. The infant can perceive her as a real or separate other who can be used properly and not just omnipotently. Winnicott (1956) suggests that optimal development of self-esteem depends on the mother's capacity for affective "mirroring" of the infant. Inevitable failures and frustration are essential for ultimate adaptation in that they facilitate the breakaway from infantile omnipotence and give an opportunity to the mother to *repair* the inevitable hurt by permitting *regression* to complete fusion.

The idea of "transitional phenomena" shows how the infant uses the mother to facilitate independent functioning (Winnicott, 1971b). A favorite blanket may help to soothe the infant because it is grasped in the moment when the infant fantasizes about the breast-feeding situation, and it is associated with calling the mother (and the breast) to mind in her absence. The physical object is both the infant (the "me" aspect) and the mother (the "not-me" aspect); it is transitional in facilitating the move from the omnipotence that must occur in relating to a subjectively created object, to relating to the "real" mother who is seen as part of external reality.

The mother has to be "good-enough," but her failure is expectable and is the motivator of growth. Winnicott (1958b) stresses that the baby must not be challenged too soon about the mother's "realness" (her independent existence) and asked to negotiate the "me and the not-me." The baby's omnipotence gives rise to the ego nuclei which will in time become integrated in the real experience of the *I* (the true self). Serious trauma at the developmental level prior to the stable differentiation of self and objects, creates a basic fault in the structure of the psyche, which Balint (1968) envisages not as a split or fracture, but as a profound misordering. A person manifesting a basic fault has an underlying feeling that something is not right about him or her; he or she is not resentful about this but invariably seeks a solution in the environment. The basic fault is seen as the developmental root of personality disorder.

1.14.4.3.2 The Independent view of neurotic processes and severe psychopathology

(i) Splitting of the self

The key contribution by Fairbairn (1944) is the proposition that early trauma of great severity is stored in memories which are "frozen" or dissociated from a person's central ego or functional self. This conception steps beyond the classical psychoanalytic notion of repression in developmental accounts of psychopathology. The classical model of pathogenesis (conflict → repression → reactivation of conflict → neurotic compromise) is still seen to apply to conflicts which reach the oedipal (3–4 year) level. The Independent model concerns disorders of the self, thought to arise out of traumatic events before that age. Anxiety, as all pathological states, is regarded as rooted in the conflicts over infantile dependence. The regressive wish to be dependent carries with it the threat of loss of identity. The progressive goal of separation generates anxiety over feeling isolated, alone, and unsupported. Whilst a retreat to the home base might offer short-term relief, eventually it creates anxiety over engulfment and loss of identity, thus reinitiating the conflict.

(ii) Schizoid personality

Schizoid personality (Fairbairn, 1952) arises out of the infantile experience that love is destructive for the mother and therefore has to be inhibited along with all intimacy. In schizoid states the ego is so split that the individual may be mystified about himself and is transiently disturbed about reality (finding the familiar in the unfamiliar and vice versa). These individuals resist perceiving others as whole persons; they hide their love, and, to protect themselves from others" love, they erect defenses designed to distance others, seeming indifferent, or even hateful, giving themselves over to the pleasures of hating and destruction.

(iii) The false self

If trauma occurs at the stage of absolute dependency, an adaptation in the form of a "caretaker self" may develop (Winnicott, 1965a, 1971a). If the mother is unable to provide comprehension of the infant through her gestures, the infant will be forced into a situation of compliance, which has mimicry but which is alien to his or her true self. The infant, and then child, will have the capacity to "go through the motions" of interpersonal relationships, but these encounters will be with the false self which will serve to hide the true self (Winnicott, 1965b). At the extreme, it is only functional outside of intimate relationships, and when called upon within such relationships (by, for example, an intensive psychotherapeutic encounter) it may "break down" and leave unprotected an infantile and poorly developed sense of "true self."

1.14.4.3.3 Strengths and weaknesses of the approach

Post-war British psychoanalysis, particularly as exemplified in the work of Winnicott, has been immensely influential in bringing about the sea change in psychodynamic thinking which we have witnessed over the last 20 years. Winnicott, and other Independent object-relations theorists, became a source of inspiration to US object-relations theorists such as Modell (1984) and Ogden (1986). At the same time they have influenced the work of major French psychoanalytic thinkers such as Andre Green (1975). In addition, its impact on Kohutian thinking was considerable although not often acknowledged.

The major weakness of the British object-relations school is its naive reconstruction of infancy in the adult mind. Although infant research confirms some speculations and informal observations, it cannot sustain the developmental argument of a linear evolution from infancy to adulthood. Human development is far too complex for infantile experiences to have direct links to adult pathology. In fact, to the extent that such research is available, longitudinal studies of infancy suggest that personality organization is subject to reorganization throughout development, based on significant positive and negative influences (Block, Block, & Gjerde, 1988; Emde, 1988a, 1988b). The infantile experiences described by Winnicott and others are thus no less metaphorical and reductionistic than the ego psychological formulations they aimed to replace.

1.14.4.4 Kohut's Self-psychology

1.14.4.4.1 Innovative aspects of Kohut's theory

Kohut (1971, 1977, 1984) proposed that the main developmental attainment for any individual is the achievement of a cohesive self. Self-cohesion is the primary motivational property guiding human behavior, and is derived from inevitable disappointments of grandiosity and exhibitionistic needs (Kohut, 1971).

Self-development is the internalization of the self-object. A self-object is defined as a person in the environment who performs particular functions for the self; these functions evoke the experience of selfhood (Wolf, 1988). To begin with, a mirroring selfobject (assumed to be the mother) allows the unfolding of exhibitionism and grandiosity. Frustration, when phase appropriate and optimal in degree, permits a gradual modulation of infantile omnipotence through "a transmuting internalization" of this mirroring function. Empathic responses from the mirroring selfobject facilitate the unfolding of the infant's grandiosity, exhibitionism, feelings of perfection and omnipotence, and this enables him or her to build an idealized image of the parent with whom he or she wishes to merge. The idealization of self-objects, also through internalization, leads to the development of ideals. The internalizing of the mirroring function and idealized self-object leads to the emergence of a "bipolar self" with its ambitions and ideals and the natural talents available to it. "Transmuting internalization" of the self-object leads gradually to consolidation of the *nuclear self* (Kohut & Wolf, 1978, p. 83). Kohut (1977) suggested that it is the "enfeebled self" which turns defensively towards pleasure aims (drives). Drives are breakdown products of disappointments to the self, usually involving failures in emotional attunement of selfobjects.

1.14.4.4.2 Self-psychological model of mental disorder

When parents consistently fail to provide for the child's narcissistic needs, Kohut suggests that the archaic grandiose self and the idealized parent imago may become hardened and fail to be integrated into later structures. Akhtar and Anderson, however, (1982) summarize some of the behavioral features of narcissistic personality disorder included in Kohut's writing (see Kohut, 1971): rage as a reaction to threats to self-esteem, the need for revenge to deal with narcissistic injury, a lack of capacity to form and maintain relationships, perverse sexuality or a lack of sexual interest, lack of empathy with others, pathological lying, a limited capacity for taking a playful or humorous stance towards the self, hypomanic states of exaltation, and overconcern with the bodily self. There is a defect in the self, predisposing the individual to experiences of threatened fragmentation and empty depression. These experiences are disguised by defensive and compensatory behavior such as the seeking of adulation and intense excitement.

The grandiose self persists in an unneutralized way because the child is not met with appropriate mirroring responses. The idealized parental imago will remain, and the child will be left with an unattainable, unrealistic, or partial system of values and ideas. Patients in Kohutian treatment are allowed to express both their idealization of the therapist and their own grandiosity without being met by confrontation or interpretation. The empathic stance taken by the therapist is believed to reactivate the developmental process, and, through the inevitable and gradual disappointment of the patient, the neutralization of both the grandiose self and the idealization of the caregiving figures will be resumed.

1.14.4.4.3 Strengths and weaknesses of Kohut's model

Kohut's clinical approach has strong following amongst psychologists. This is principally because of the clinical relevance of his ideas and their usefulness in psychotherapeutic work. There are clear and readily implementable technical recommendations arising out of Kohut's approach based around the need to "restore the self" using therapeutic methods which perform the developmental task of transmuting internalization.

There are numerous critical reviews of Kohut's theory from within structural psychoanalytic models (see, e.g., Slap & Trunnel, 1987). Schwartz (1978), for example, finds Kohut's descriptions suffer from overinclusiveness. Fragmentations of the self are believed to include depressions, depersonalizations, and disorganizing anxieties, as well as temporary or encapsulated psychotic states.

A major problem in Kohut's more recent formulations is the implicit confusion between self and self-representation. The self is presented by Kohut in representational terms, yet Kohut ascribes to it motivational properties and tendencies such as goals, plans, and self-esteem motivation (see Kohut, 1971). In this way, the self comes to denote most, if not all, of the personality and therefore becomes a superfluous term, much as the concept of ego was overextended by ego psychologists (see, e.g., Schafer, 1976).

1.14.4.5 Kernberg's Integrated Theory

1.14.4.5.1 Internalization and the construction of self, object, and affect triads

In Otto Kernberg's model, affects serve as the primary motivational system (Kernberg, 1983). He suggests that combinations of a self-representation, an object-representation, and an affect state linking them are the essential units of psychic structure. He terms these self–object–affect triads. His model is based on reconstructions from the treatment of severely disturbed adults, which are strongly influenced by Kleinian theory. It is less concerned with the child's real experience and focuses on the force of introjects and fantasies.

Kernberg (1976) described self-image as a product of the process of internalization. For Kernberg (1976) internalization entails three processes:

(i) *Introjection*, the most basic, involves the reproduction of an interaction with the environment by means of the clustering of memory traces attached to the self- or object-image and the interactions of the two in their affective context.

(ii) *Identification* presumes the child's cognitive ability to recognize the variety of role dimensions that exist in interactions with others, and involves the capacity of the self to model itself after the object influenced by fantasy and affect. The individual's experiences of gratification and frustration influence affective states and determine the degree to which self-representation is flexible, true, and complex.

(iii) *Ego identity*, borrowed from Erikson (1959b), as the overall organization of introjections and identifications under the synthesizing influence of the ego.

1.14.4.5.2 Levels of pathology within Kernberg's system

Kernberg suggested that all levels of disturbance are more complex in severe personality disturbance but exist across the entire spectrum of psychopathology. Kernberg (1984) sees neurotic pathology as regression to a relatively integrated, although repressed, infantile self, connected to relatively integrated although also unconscious representations of the parental object. Patients with a neurotic, rather than borderline, level of personality organization are able to integrate positive and negative representations of self and others. They have passed through developmental phases where good and bad representations of self and others were combined across affective valances and a complex presentation containing both loving and hostile elements evolved.

Even such relatively integrated internal representations, however, contain dyadic units which reflect either a defensive or an impulsive aspect of psychic conflict. An individual is highly susceptible to anxiety when configurations of self- and object-representations are highly charged affectively, and are poorly differentiated. For example, a representation of the self as being weak and vulnerable may be coupled with an object representation of ruthless domination with a violent affective tone. When this configuration is activated in therapy or elsewhere, the individual may become highly anxious.

In more severe psychopathology, identification with self- or object-representations may rapidly alternate. For example, an individual may feel criticized, but very quickly the criticism can shift from the self onto the other; and the critic is now seen as the self who is hurt and mistreated, and the individual identifies with the critical stance. This oscillation of self and other accounts for many instances where impulses appear to change into their opposite (active into passive, good into bad). Very severe levels of character pathology are typified by primitive dissociation or splitting of internalized object relations which entail a lack of integration of self- and object-representations, projections of primitive superego nuclei, splitting, impulsivity, lack of empathy, and the unmodulated expression of libido and aggression (Kernberg, 1984). Splitting is a key feature of borderline personality organization, antisocial personalities, "as if" characters, patients with multiple sexual deviations in narcissistic personalities, addictions, and even in analytically approachable psychosis. The highly unrealistic, sharply idealized or (through aggression) highly persecutory self- and object-representations characteristic of this group do not correspond to any real relationship. As object relations are poorly integrated, the reversals of the enactment of self- and other-representations may be particularly rapid. This can make relationships with such individuals confusing and even chaotic. For example, love and hate may exist in a dissociated way side by side; several object relations may be condensed into single images.

1.14.4.5.3 Borderline personality disorder

It should be noted that, for Kernberg, borderline is a level of psychic organization rather than a nosological entity. His criteria for the disorder include: (i) nonspecific manifestations of ego weakness (poor affect tolerance, impulse control and sublimatory capacity), (ii) primitive defenses, (iii) identity diffusion, (iv) intact reality testing but a propensity to shift towards dream-like thinking. There is some empirical evidence in support of Kernberg's criteria (see Clarkin, Foelsch, & Kernberg, 1996; Kernberg, Goldstein, & Carr, 1981).

For Kernberg (1977), the root cause of borderline states is the intensity of destructive and aggressive impulses, and the relative weakness of ego structures available to handle them. Kernberg sees the borderline individual as using the developmentally early defenses (e.g., splitting) in an attempt to separate contradictory images of self and others to protect positive images from being overwhelmed by negative and hostile ones. Primitive idealization, also a consequence of splitting, protects the individual from the "all bad" objects through creating an omnipotent object in fantasy which is the container of grandiose identifications. Splitting also results in a "diffuse sense of identity" which is characterized by a confused internal representation of the "real" object, and an unintegrated primitive superego which sets unattainable ideals and internalized persecutory images.

Kernberg (1987) illustrates how the self-destructiveness, self-mutilating behavior, and suicidal gestures tend to coincide with intense attacks of rage towards the object. They can serve to re-establish control over the environment by evoking guilt feelings, or express unconscious guilt over the success of a deepening relationship. In some patients self-destructiveness occurs because their self-image becomes "infiltrated" with aggression, so that they experience increased self-esteem and a confirmation of their grandiosity in self-mutilation or masochistic sexual perversions. The caring professions can respond only with despair to these patients" obvious sense of triumph in their victory over pain and death. Their efforts seem futile to the patient, who at an unconscious level experiences a sense of being in control over death. Self-mutilation, such as cutting, may also protect from the identity diffusion (derealization) which is a constant threat to the fragmented internal world of the borderline individual.

1.14.4.5.4 Strengths and weaknesses of Kernberg's model

Kernberg's is perhaps the most systematically elaborated of all the psychodynamic models, yet it shares many weaknesses with all the other models we have considered. These can be briefly summarized as follows.

(i) It is logically erroneous to assume that just because some features of severe psychopathology have characteristics in common with children's thinking that the functions served in the adult mind are identical in the two contexts. Thus the developmental metaphor cannot be used for explanatory purposes.

(ii) A related weakness is the emphasis on early development as the root of severe psychological disturbance. Many significant environmental influences do not make themselves felt until late childhood, adolescence, or even adulthood.

(iii) There is a lack of specificity in Kernberg's model which is apparent perhaps only because he, more than other writers, attempts actively to distinguish between subtypes of disorders. Yet it is not clear from his theory why, for example, borderline personality disordered individuals spontaneously improve with time (McGlashan, 1986; Stone, 1990).

1.14.4.6 Bowlby's Attachment Theory Model

1.14.4.6.1 Innovative aspects of Bowlby's approach

The infant comes into the world predisposed to participate in social interactions. The British psychoanalyst, John Bowlby (1969, 1973, 1980) was the first to give central place to the child's biological proclivity to form attachments, to initiate, maintain, and terminate interactions with the caregiver and use him/her as a "secure base" for exploration and self-enhancement. Bowlby's (1969) critical contribution was his focus on the infant's need for unbroken (secure) early attachment to the mother. The child who does not have such provision is likely to show signs of partial deprivation: excessive need for love or for revenge, gross guilt and depression; or complete deprivation: listlessness, quiet unresponsiveness, and retardation of development. Later there are signs of superficiality, want of real feeling, lack of concentration, deceit and compulsive thieving. Later Bowlby (1973) placed these reactions into a framework of reactions to separation: protest → despair → detachment.

Bowlby's attachment theory is unlike most other psychoanalytic formulations in that it is, for the most part, prospective (Bowlby, 1969). Laboratory investigations such as those of Brazleton and colleagues provided important support for Bowlby in demonstrating the innate social disposition of the infant, and the adverse consequences if expectations of social responsiveness from the caretaker are not met (see Tronick, Als, Adamson, Wise, & Brazelton, 1978). He is also most bold in claiming that the infantile roots of pathology lie in *actual* realistically based fears.

Following Bowlby (1973), the attachment of infants to their parents is recognized across the social science disciplines as a fundamental psychological process affecting human development across the lifespan (see Goldberg, Muir, & Kerr, 1995). Secure (safe and stable) vs. insecure (anxious or ambivalent) attachment of the child to its parents has been identified as a primary influence upon the child's evolving adaptation to the environment. Quality of attachment can be successfully assessed in infancy using a simple laboratory separation between a 12–18 months-old child and the caregiver, developed by John Bowlby's closest colleague, Mary Ainsworth (Ainsworth, Blehar, Waters, & Wall, 1978). Upon reunion, secure infants actively seek contact with and are soothed by the caregiver. Insecurely attached infants either avoid the caregiver or resist the caregiver's attempt to comfort them. The pervasive influence of the quality of infant–parent bonds upon subsequent social, cognitive, and emotional development has been intensively studied (Carlson & Sroufe, 1995). A number of longitudinal studies show that attachment in infancy strongly influences many aspects of psychological adaptation, including social behavior (Skolnick, 1986), affect regulation (Erickson, Sroufe, & Egeland, 1985), cognitive resourcefulness (Grossmann & Grossmann, 1991), and psychological disturbance (Sroufe, 1989).

Infant patterns of attachment, however, do not invariably determine subsequent attachment relationships (Lamb, 1987), and this variability is probably due to qualities of the internal representation of attachment patterns.

1.14.4.6.2 Attachment and psychopathology

Bowlby (1969, 1973) suggests that disruption in the functioning of the attachment system will interfere with the child's developing capacities for regulating his behavior, emotions, and arousal. He argues that since children have many of their first experiences of emotional states (intense anger and anxiety, as well as love and happiness) in the context of their early attachment relationships, the quality of these relationships will determine their capacity for self-regulation at times of high stress. Insecurely attached children should therefore be more vulnerable to emotional and behavioral disregulation, and have fewer opportunities to elaborate the capacity to regulate emotional experiences, than secure ones (see also Ainsworth et al., 1978). Further, Bowlby (1973, 1980) maintained that secure attachment will generate internal working models of relationships characterized by an expectation of emotional as well as physical support, leading to positive self-concept and confidence in the availability and responsiveness of the other. Insensitive parenting will give rise to insecure models of relationships, characterized by lack of trust in the other and a self-representation as unworthy and undeserving of love and affection.

Broadly speaking, Bowlby's prediction that insecure attachments are associated with various later difficulties has been borne out by empirical research. Insecurely attached children appear to be more likely to experience fluctuating and unpredictable affective states, including intensely negative emotions such as excessive sadness and anger (Cassidy, 1994). Insecure attachments are associated with maladaptive functioning in other contexts, with problems of emotional disregulation, heightened sensitivity to stress, pervasive anxiety and distress problems in interpersonal relationships, internalizing and externalizing disorders (Carlson & Sroufe, 1995).

1.14.4.6.3 Strengths and weaknesses of Bowlby's theory

Bowlby's approach is a general systems theory and, as such, has many strong features. It has close ties to empirical data. His approach is also consistent with the current interpersonal or relational emphasis in psychodynamic therapy (Greenberg, 1991). The notion of internal working model is a useful, if somewhat vague, metaphor (Dunn, 1996) which encompasses the three components of human goal-oriented interpersonal behavior: the self, an object, and an interactional field within which the two establish a specific pattern of relationship (Greenberg, 1991). Bowlby's theory is strong with regard to identifying "maintaining factors" in pathological transactions without placing an unduly strong emphasis on past experience.

The weakness of his theory is in failing to specify how change may be most readily achieved, or indeed what might constitute change. For interpersonalists, change is often seen as deriving from the development of a new relationship with the attachment figure of the therapist (Holmes, 1993). Traditional psychodynamic notions of aggression and conflict also are difficult to integrate with Bowlby's ideas.

1.14.5 EVIDENCE FOR PSYCHODYNAMIC THEORIES

We have already touched on the speculative nature of many psychodynamic constructs. The tendency for psychodynamic theorists to identify evidence from their clinical work consistent with their theories is overwhelming and is most likely responsible for the numerous overlapping formulations offered to fit comparable data sets. In this section we shall consider some of the areas of active research in which such interplay could have taken place more intensively given interest from either the research or the psychodynamic clinical community.

As has been emphasized throughout this chapter, the framework perhaps most consistently made use of is the developmental one. The psychodynamic assumption that childhood experience has complex causal relationships with the characteristics of adult adaptation has been supported by evidence from a number of fields of study.

There are numerous longitudinal studies which demonstrate that problems of adaptation are predictable on the basis of childhood characteristics. The longitudinal study of Block and Block (1980) demonstrated that both drug abuse (Shedler & Block, 1990) and depression (Block, Gjerde, & Block, 1991) can be predicted from childhood characteristics. Such findings are naturally equally consistent with genetic predispositions; however, many findings contain subtle specific associations which are more readily accounted for in psychodynamic terms than in terms of behavior genetics. For example, in the Block et al. (1991) report, men with depressive features were observed to have been impulse-ridden in childhood, whereas women with the same diagnosis had been overcontrolled.

Research using a wide range of methodologies uniformly reports that poor parenting practices are a major contributing factor to the development of psychiatric disorders. The balance of evidence now favors the case for the pervasive and apparently irreversible impact of inadequate early social relationships (e.g., Hodges & Tizard, 1989a, 1989b). It is also recognized that, in particular during early months, when growth and maturation are most rapid, the degree and quality of caregiver assistance may be critical for the long-term biological organization of the child (e.g., Greenough & Black, 1992). The accessibility of attachment patterns to empirical research (Ainsworth et al., 1978) has prompted numerous longitudinal studies. These have demonstrated that particularly secure mother–infant attachment predicts predominantly good long-term outcome (Carlson & Sroufe, 1995) and disorganized patterns of attachment are most likely to be associated with pathological patterns of adaptation (Lyons-Ruth, 1996). In particular, there is unassailable evidence to suggest that, at the extreme end, childhood maltreatment is a powerful cause of maladaptive behavior (Cicchetti & Toth, 1995).

As we have seen, one of the core assumptions of modern psychodynamic approaches concerns the representational processes which are assumed to underpin maladaptive behavior. Object relations approaches to psychopathology (Greenberg & Mitchell, 1983) assume that disorders are rooted in cognitive structures which bias the representation of interpersonal relationships. Object relations formulations of psychodynamic theory have much in common with approaches rooted in social cognition, as has been shown by Western (1991). Distortions or pervasive biases of social cognition are generally accepted as central to understanding manifestations of severe psychopathology, whatever the causes of such cognitive distortions may be (e.g., Mineka & Sutton, 1992). Blatt and his colleagues (e.g., Blatt et al. 1994) demonstrated cognitive distortions in the narratives of depressed patients, particularly marked deficits in the description of the self. Those patients whose experience of depression was one of emptiness, loneliness, and labile, diffuse negative affectivity are more likely to be diagnosed with borderline personality disorder. Such associations may be relatively easy to account for in terms of differences in self and object representations in borderline and nonborderline individuals. The quality of depressive experience reflects the sparsity of the internal world of the borderline individual.

Modern cognitive science has increasingly drawn on notions of unconscious cognitive processes to account for performance in the areas of memory (e.g., Schacter, 1992), thought (e.g., Jacoby & Kelly, 1992), and emotion (e.g., Murphy & Zajonc, 1993). That much cognitive processing occurs outside of awareness is an inevitable conclusion given our enhanced understanding of the complexity of these processes. Somewhat more controversial, but equally unassailable, is the idea that conscious and unconscious processing of the same theme may produce qualitatively different outputs. A study by Fazio and his colleagues (Fazio, Jackson, Dunton, & Williams, 1995) illustrates this point. They administered a measure of conscious racial attitudes to subjects at the same time as assessing unconscious or implicit attitudes by seeing the extent to which subjects" responses to adjectives was slowed down or speeded up by the presentation of black or white faces. The subjects" task was simply to categorize adjectives as good or bad. An implicit negative attitude was indicated by the face of a black person speeding up responses to negative words, and a white person's face speeding up responses to positive adjectives. The subjects then engaged in conversation with a black confederate who rated them in terms of their friendliness as she discussed racial issues with them. Interestingly the priming task, but not the conscious attitude ratings, predicted the blind independent ratings of the confederate.

This necessarily brief survey has illustrated that psychodynamic notions are not "outlandish" from the point of view of empirical psychology. None of these studies can "prove" psychodynamic ideas, as the two pertain to independent realms of human functioning: the laboratory and the consulting room. However, taken as a whole, they point to the validity of underlying psychological processes implicitly posited by psychodynamic theoreticians when offering their model of human development and the pathological outcomes that may arise from it. Psychodynamic theory is unlikely ever to be comprehensively validated in empirical studies. All that we can hope for is that the general model of a person, painstakingly assembled on the basis of neuropsychological, social and cognitive-psychological investigations, will not be too dissimilar from the vision of human functioning which has emerged from the clinical endeavors of psychodynamically oriented practitioners.

1.14.6 REFERENCES

Aichhorn, A. (1935). *Wayward youth*. New York: Viking. (Original work published 1925).

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the Strange Situation*. Hillsdale, NJ: Erlbaum.

Akhtar, S., & Anderson, T. J. (1982). Overview: Narcissistic personality disorder. *American Journal of Psychiatry, 139,* 12–20.

Alexander, F. (1930). The neurotic character. *International Journal of Psycho-Analysis, 11,* 292–311.

Allen, J. (1995). The spectrum of accuracy in memories of childhood trauma. *Harvard Review of Psychiatry, 3,* 84–95.

American Psychiatric Association. (1994). *DSM-IV draft criteria—Task force on DSM-IV*. Washington, DC: Author.

Arlow, J. A., & Brenner, C. (1964). *Psychoanalytic concepts and the structural theory*. New York: International Universities Press.

Bahrick, L. R., & Watson, J. S. (1985). Detection of intermodal proprioceptive-visual contingency as a potential basis of self-perception in infancy. *Developmental Psychology, 21,* 963–973.

Balint, M. (1959). *Thrills and regressions*. London: Hogarth.

Balint, M. (1968). *The basic fault*. London: Tavistock.

Bion, W. R. (1959). Attacks on linking. *International Journal of Psycho-Analysis, 40,* 308–315.

Bion, W. R. (1962). *Learning from experience*. London: Heinemann.

Blatt, S., Ford, R. Q., Berman, W. H., Cook, B., Cramer, P., & Robins, C. E. (1994). *Therapeutic change: An object relations approach*. New York: Plenum.

Block, J., & Block, J. (1980). The role of ego-control and ego-resiliency in the organization of behavior. In W. A. Collins (Ed.), *Minnesota symposium on child psychology: Development of cognitive affect and social relations* (Vol. 13, pp. 39–101). Hillsdale, NJ: Erlbaum.

Block, J., Block, J. H., & Gjerde, P. F. (1988). Parental functioning and the home environment in families of divorce: Prospective and concurrent analyses. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 207–213.

Block, J. H., Gjerde, P., & Block, J. (1991). Personality antecedents of depressive tendencies in 18-year-olds: A prospective study. *Journal of Personality and Social Psychology, 60,* 726–738.

Boesky, D. (1988). A discussion of evidential criteria for therapeutic change. In A. Rothstein (Ed.), *How does treatment help? On the modes of therapeutic action of psychoanalytic therapy* (pp. 171–180). Madison, CT: International Universities Press.

Bower, T. R. (1989). *The rational infant: Learning in infancy*. New York: Freeman.

Bowlby, J. (1958). The nature of the child's tie to his mother. *International Journal of Psycho-Analysis, 39,* 350–373.

Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment*. London: Hogarth Press and the Institute of Psycho-Analysis.

Bowlby, J. (1973). *Attachment and loss: Vol. 2. Separation: Anxiety and anger*. London: Hogarth Press and Institute of Psycho-Analysis.

Bowlby, J. (1980). *Attachment and loss: Vol. 3. Loss: Sadness and depression*. London: Hogarth Press and Institute of Psycho-Analysis.

Brenner, C. (1979). The components of psychic conflict and its consequences in mental life. *Psychoanalytic Quarterly, 48,* 547–567.

Brenner, C. (1982). *The mind in conflict*. New York: International Universities Press.

Britton, R. (1989). The missing link: Parental sexuality in the Oedipus complex. In J. Steiner (Ed.), *The Oedipus complex today* (pp. 83–102). London: Karnac Books.

Carlson, E., & Sroufe, L. A. (1995). Contribution of attachment theory to developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Vol. 1. Theory and methods* (pp. 581–617). New York: Wiley.

Cassidy, J. (1994). Emotion regulation: influences of attachment relationships. In N. A. Fox (Ed.), *The development of emotion regulation. Monograph of the society for research in child development* (Serial No. 240) (pp. 228–249). Chicago: University of Chicago Press.

Cicchetti, D., & Cohen, D. J. (1995). Perspectives on developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Manual of developmental psychopathology* (pp. 3–23). New York: Wiley.

Cicchetti, D., & Toth, S. L. (1995). A developmental psychopathology perspective on child abuse and neglect. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 541–565.

Clarkin, J., Foelsch, P. A., & Kernberg, O. F. (1996). *Manual for the Inventory of Personality Organization (IPO)*. New York: Cornell University Medical College.

Cooper, A. M. (1983). Psychoanalytic inquiry and new knowledge. In J. D. Lichtenberg & S. Kaplan (Eds.), *Reflections on self psychology* (pp. 19–34). Hillsdale, NJ: Analytic Press.

Cooper, A. M. (1985). A historical review of psychoanalytic paradigms. In A. Rothstein (Ed.), *Models of the mind: Their relationships to clinical work* (pp. 5–20). New York: International Universities Press.

Crews, F. (1995). *The memory wars: Freud's legacy in dispute*. London: Granta Books.

Dunn, J. (1996). The Emanuel Miller Memorial Lecture 1995: Children's relationships: Bridging the divide between cognitive and social development. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 37,* 507–518.

Emde, R. N. (1980a). Emotional availability: A reciprocal reward system for infants and parents with implications for prevention of psychosocial disorders. In P. M. Taylor & F. Orlando (Eds.), *Parent–infant relationships* (pp. 87–115). New York: Grune & Stratton.

Emde, R. N. (1980b). Toward a psychoanalytic theory of affect: Part 1. The organizational model and its propositions. In S. I. Greenspan & G. H. Pollock (Eds.), *The course of life: Vol. 1. infancy and early childhood* (pp. 63–83). Washington, DC: DHSS.

Emde, R. N. (1980c). Toward a psychoanalytic theory of affect: Part II. Emerging models of emotional development in infancy. In S. I. Greenspan & G. H. Pollock (Eds.), *The course of life: Vol. 1. Infancy and early childhood* (pp. 85–112). Washington, DC: DHSS.

Emde, R. N. (1981). Changing models of infancy and the nature of early development: Remodelling the foundation. *Journal of the American Psychoanalytic Association, 29,* 179–219.

Emde, R. N. (1985). From adolescence to midlife: Remodeling the structure of adult development. *Journal of the American Psychoanalytic Association, 33,* 165–192.

Emde, R. N. (1988a). Development terminable and interminable. I. Innate and motivational factors from infancy. *International Journal of Psycho-Analysis, 69,* 23–42.

Emde, R. N. (1988b). Development terminable and interminable. II. Recent psychoanalytic theory and therapeutic considerations. *International Journal of Psycho-Analysis, 69,* 283–296.

Erickson, M. F., Sroufe, L. A., & Egeland, B. (1985). The relationship between quality of attachment and behavior problems in preschool in a high-risk sample. *Monographs of the Society for Research in Child Development, 50,* 147–166.

Erikson, E. H. (1950a). *Childhood and society*. New York: Norton.

Erikson, E. H. (1950b). Growth and crises of the "healthy personality." *The healthy personality*. Caldwell, NJ: Progress Associates.

Erikson, E. H. (1959a). *Identity and the life cycle*. New York: International Universities Press.

Erikson, E. H. (1959b). The problem of ego identity. In *Identity and the life cycle* (pp. 104–164). New York: International Universities Press. (Original work published 1956.)

Fairbairn, W. R. D. (1944). Endopspychic structure considered in terms of object-relationships. *International Journal of Psycho-Analysis, 25,* 60–93.

Fairbairn, W. R. D. (1952). *An object-relations theory of the personality*. New York: Basic Books.

Fairbairn, W. R. D. (1954). Observations on the nature of hysterical states. *British Journal of Medical Psychology, 29,* 112–127.

Fairbairn, W. R. D. (1963). Synopsis of an object-relations theory of the personality. *International Journal of Psycho-Analysis, 44,* 224–225.

Fazio, R., Jackson, J. R., Dunton, B., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69,* 1013–1027.

Fenichel, O. (1945). *The psychoanalytic theory of neurosis*. New York: Norton and Routledge.

Fonagy, P. (1982). Psychoanalysis and empirical science. *International Review of Psychoanalysis, 9,* 125–145.

Fonagy, P., Steele, M., Steele, H., Leigh, T., Kennedy, R., Mattoon, G., & Target, M. (1995a). Attachment, the reflective self, and borderline states: The predictive specificity of the Adult Attachment Interview and pathological emotional development. In S. Goldberg, R. Muir, & J. Kerr (Eds.), *Attachment theory: Social developmental and clinical perspectives* (pp. 233–278). Hillsdale, NJ: Analytic Press.

Fonagy, P., Target, M., Steele, M., & Gerber, A. (1995b). Psychoanalytic perspectives on developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Vol. 1. Theory and methods* (pp. 504–554). New York: Wiley.

Frances, A., Pincus, H., Widiger, T., Davis, W., & First, M. (1990). DSM-IV: Work in progress. *American Journal of Psychiatry, 147,* 1439–1448.

Freud, A. (1946). *The ego and the mechanisms of defence*. New York: International Universities Press. (Original work published 1936.)

Freud, A. (1965). *Normality and pathology in childhood.* Harmondsworth, UK: Penguin.

Freud, A. (1981). *Changes in psychoanalytic practice and experience*. New York: International Universities Press.

Freud, S. (1900). The interpretation of dreams. In J. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vols. 4, 5, pp. 1–715). London: Hogarth.

Freud, S. (1901). The psychopathology of everyday life. In J. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 6, pp. 1–190). London: Hogarth.

Freud, S. (1905a). Jokes and their relation to the unconscious. In J. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 8, pp. 1–236). London: Hogarth.

Freud, S. (1905b). Three essays on the theory of sexuality. In J. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 7, pp. 123–230). London: Hogarth.

Freud, S. (1920). Beyond the pleasure principle. In J. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 18, pp. 1–64). London: Hogarth.

Freud, S. (1923). The ego and the id. In J. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 19, pp. 1–59). London: Hogarth.

Freud, S. (1926). Inhibitions, symptoms and anxiety. In J. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 20, pp. 77–172). London: Hogarth.

Freud, S. (1930). Civilization and its discontents. In J. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 21, pp. 57–146). London: Hogarth.

Freud, S. (1938). An outline of psychoanalysis. In J. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 23, pp. 139–208). London: Hogarth.

Freud, S., & Breuer, J. (1895). Studies on hysteria. In J. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 2, pp. 1–305). London: Hogarth.

Friedman, L. (1988). The clinical popularity of object relations concepts. *Psychoanalytic Quarterly, LVII,* 667–691.

Gergely, G. (1991). Developmental reconstructions: Infancy from the point of view of psychoanalysis and developmental psychology. *Psychoanalysis and Contemporary Thought, 14,* 3–55.

Gitelson, M. (1955). Contribution to scientific proceedings, reported by L. Rangell. *Journal of the American Psychoanalytic Association, 3,* 294–295.

Glover, E. (1949). *Psycho-Analysis*. New York: Staples.

Goldberg, S., Muir, R., & Kerr, J. (1995). *Attachment theory: Social, developmental and clinical perspectives*. Hillsdale, NJ: Analytic Press.

Golinkoff, R. M., Hardig, C. B., Carlson, V., & Sexton, M. E. (1984). The infant's perception of causal events: The distinction between animate and inanimate objects. In L. P. Lipsitt & C. Rovee-Collier (Eds.), *Advances in infancy* *research* (pp. 125–165). Norwood, NJ: Ablex.

Green, A. (1975). The analyst, symbolization and absence in the analytic setting: On changes in analytic practice and analytic experience. *International Journal of Psycho-Analysis, 56,* 1–22.

Greenberg, J. (1991). *Oedipus and beyond.* Cambridge, MA: Harvard University Press.

Greenberg, J. R., & Mitchell, S. A. (1983). *Object relations in psychoanalytic theory*. Cambridge, MA: Harvard University Press.

Greenough, W., & Black, J. (1992). Induction of brain structure by experience: Substrates for cognitive development. In M. R. Gunnar & C. A. Nelson (Eds.), *Minnesota symposia on child psychology* (Vol. 24, pp. 155–200). Hillsdale, NJ: Erlbaum.

Grossmann, K. E., & Grossmann, K. (1991). Attachment quality as an organizer of emotional and behavioural responses in a longitudinal perspective. In C. M. Parkes, J. Stevenson-Hinde, & J. Marris (Eds.), *Attachment across the life cycle* (pp. 93–114). London: Routledge.

Grünbaum, A. (1984). *The foundations of psychoanalysis: A philosophical critique*. Berkeley, CA: University of California Press.

Harris, P. L. (1994). The child's understanding of emotion: Developmental change and the family environment. *Journal of Child Psychology and Psychiatry, 35,* 3–28.

Hartmann, H. (1958). *Ego psychology and the problem of adaptation.* New York: International Universities Press: (Original work published in 1939.)

Hartmann, H. (1964a). *Comments on the psychoanalytic theory of the ego.* New York: International Universities Press. (Original work published 1950.)

Hartmann, H. (1964b). Notes on the theory of sublimation. In *Essays on ego psychology* (pp. 215–240). New York: International Universities Press. (Original work published in 1955.)

Hartmann, H. (1964c). The mutual influences in the development of ego and id. In *Essays on ego psychology* (pp. 155–182). New York: International Universities Press. (Original work published in 1952.)

Hartmann, H., Kris, E., & Loewenstein, R. (1946). Comments on the formation of psychic structure. *The Psychoanalytic Study of the Child*, *2,* 11–38.

Hayman, A. (1989). What do we mean by "phantasy"? *International Journal of Psycho-Analysis*, *70,* 105–114.

Hodges, J., & Tizard, B. (1989a). IQ and behavioral adjustment of ex-institutional adolescents. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *30,* 53–75.

Hodges, J., & Tizard, B. (1989b). Social and family relationships of ex-institutional adolescents. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *30,* 77–97.

Holmes, J. (1993). *John Bowlby and attachment theory*. London: Routledge.

Holzman, P. S. (1985). Psychoanalysis: Is the therapy destroying the science? *Journal of the American Psychoanalytic Association*, *33,* 725–770.

Horney, K. (1939). *New ways in psychoanalysis.* New York: Norton.

Jacobson, E. (1964). *The self and the object world.* New York: International Universities Press.

Jacobson, W., & Cooper, A. M. (1993). Psychodynamic diagnosis in the era of the current DSM's. In N. E. Miller, L. Luborsky, J. P. Barber, & J. P. Docherty (Eds.), *Psychodynamic treatment research: A handbook for clinical practice* (pp. 109–126). New York: Basic Books.

Jacoby, L., & Kelley, C. M. (1992). A process-dissociation framework for investigating unconscious influences: Freudian slips, projective tests, subliminal perception, and signal detection theory. *Current Directions in Psychological Science*, *1,* 174–179.

Johnson, A. M., & Szurek, S. A. (1952). The genesis of antisocial acting out in children and adults. *Psychoanalytic Quarterly*, *21,* 323–343.

Jung, C. G. (1916). *Psychology of the unconscious*. London: Routledge & Kegan Paul.

Jung, C. G. (1923). *Psychological types*. London: Routledge & Kegan Paul.

Kernberg, O. F. (1976). *Object relations theory and clinical psychoanalysis*. New York: Aronson.

Kernberg, O. F. (1977). The structural diagnosis of borderline personality organization. In P. Hartocollis (Ed.), *Borderline personality disorders: The concept, the syndrome, the patient* (pp. 87–121). New York: International Universities Press.

Kernberg, O. F. (1983). Object relations theory and character analysis. *Journal of the American Psychoanalytic Association*, *31,* 247–271.

Kernberg, O. F. (1984). *Severe personality disorders: Psychotherapeutic strategies*. New Haven, CT: Yale University Press.

Kernberg, O. F. (1987). Borderline personality disorder: A psychodynamic approach. *Journal of Personality Disorders*, *1,* 344–346.

Kernberg, O. F., Goldstein, E., & Carr, A. (1981). Diagnosing borderline personality. *Journal of Nervous and Mental Disease*, *169,* 225–231.

Klein, G. (1976a). *Psychoanalytic theory: An exploration of essentials.* New York: International Universities Press.

Klein, G. S. (1976b). Freud's two theories of sexuality. *Psychological Issues, Monograph*, *36,* 14–70.

Klein, M. (1952). Notes on some schizoid mechanisms. In M. Klein, P. Heimann, S. Issacs, & J. Riviere (Eds.), *Developments in psychoanalysis* (pp. 292–320). London: Hogarth.

Klein, M. (1957). *Envy and gratitude*. New York: Basic Books.

Klein, M. (1964). *Contributions to psychoanalysis, 1921–1945*. New York: McGraw-Hill. (Original work published 1934.)

Klein, M. (1975a). A contribution to the psychogenesis of manic-depressive states. In *The writings of Melanie Klein* (pp. 236–289). London: Hogarth (Original work published 1935.)

Klein, M. (1975b). On the theory of anxiety and guilt. In *Envy and gratitude and other works, 1946–1963*. New York: Delacorte Press. (Original work published 1948.)

Klein, M. (1975c). Our adult world and its roots in infancy. In *Envy and gratitude and other works, 1946–1963*. New York: Delacorte Press. (Original work published 1959.)

Klein, M. (1984). Mourning and its relation to manic-depressive states. In *Love, guilt and reparation and other works 1921–1945* (pp. 344–369). New York: Macmillan. (Original work published 1940.)

Klein, R. (1989). Introduction to the disorders of the self. In J. F. Masterson & R. Klein (Eds.), *Psychotherapy of the disorders of the self* (pp. 30–46). New York: Brunner/Mazel.

Knight, R. P. (1953). Borderline states. *Bulletin of the Menninger Clinic, 17,* 1–12.

Kohut, H. (1971). *The analysis of the self.* New York: International Universities Press.

Kohut, H. (1977). *The restoration of the self.* New York: International Universities Press.

Kohut, H. (1984). *How does analysis cure?* Chicago: University of Chicago Press.

Kohut, H., & Wolf, E. S. (1978). The disorders of the self and their treatment: An outline. *International Journal of Psychoanalysis, 59*, 413–425.

Lamb, M. (1987). Predictive implications of individual differences in attachment. *Journal of Consulting Clinical Psychology, 55*, 817–824.

Lampl-de-Groot, J. (1949). Neurotics, delinquents and ideal formation. In K. R. Eissler (Ed.), *Searchlights on* *delinquency* (pp. 225–245). New York: International Universities Press.

Limentani, A. (1989). *Between Freud and Klein*. London: Free Association.

Loewald, H. (1986). Transference–countertransference. *Journal of the American Psychoanalytic Association, 34*, 275–288.

Lyons-Ruth, K. (1996). Attachment relationships among children with aggressive behavior problems: The role of disorganized early attachment patterns. *Journal of Consulting and Clinical Psychology, 64*, 32–40.

Mahler, M. (1968). *On human symbiosis and the vicissitudes of individuation*. New York: International Universities Press.

Mahler, M. S., Pine, F., & Bergman, A. (1975). *The psychological birth of the human infant: Symbiosis and individuation*. New York: Basic Books.

Main, T. (1957). The ailment. *British Journal of Medical Psychology, 30*, 129–145.

Masson, J. (1984). *The assault on truth: Freud's suppression of the seduction theory*. New York: Farrar, Straus & Giroux.

Masterson, J. F. (1972). *Treatment of the borderline adolescent: A developmental approach*. New York: Wiley.

Masterson, J. F. (1976). *Psychotherapy of the borderline adult: A developmental approach*. New York: Brunner/Mazel.

McGlashan, T. (1986). Long-term outcome of borderline patients. *Archives of General Psychiatry, 40*, 20–30.

Mineka, S., & Sutton, S. K. (1992). Cognitive biases and the emotional disorders. *Psychological Science, 3*, 65–69.

Mitchell, S. A. (1988). *Relational concepts in psychoanalysis: An integration*. Cambridge, MA: Harvard University Press.

Modell, A. (1984). *Psychoanalysis in a new context*. New York: International Universities Press.

Modell, A. (1990). *Other times, other realities*. Cambridge, MA: Harvard University Press.

Murphy, S. T., & Zajonc, R. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64*, 723–739.

Ogata, S. N., Silk, K. R., & Goodrich, S. (1990a). The childhood experience of the borderline patient. In P. Links (Ed.), *Family environment and borderline personality disorder* (pp. 87–103). Washington, DC: American Psychiatric Press.

Ogata, S. N., Silk, K. R., Goodrich, S., Lohr, N. E., Westen, D., & Hill, E. (1990b). Childhood sexual and physical abuse in adult patients with borderline personality disorder. *American Journal of Psychiatry, 147*, 1008–1013.

Ogden, T. (1986). *The matrix of the mind: Object relations and the psychoanalytic dialogue*. New York: Aronson.

Paris, J. (1992). Social risk factors for borderline personality disorder: A review and hypothesis. *Canadian Journal of Psychiatry, 37*, 510–515.

Paris, J., & Zweig-Frank, H. (1992). A critical review of the role of childhood sexual abuse in the etiology of borderline personality disorder. *Canadian Journal of Psychiatry, 37*, 125–128.

Pine, F. (1985). *Developmental theory and clinical process*. New Haven, CT: Yale University Press.

Racker, H. (1968). *Transference and countertransference*. New York: International Universities Press.

Rangell, L. (1955). The borderline case. *Journal of the American Psychoanalytic Association, 3*, 285–298.

Rayner, E. (1991). *The independent mind in British psychoanalysis*. London: Free Association Books.

Reich, W. (1933). *Character analysis* (3rd ed.) (V. R. Carfagno, Trans.). New York: Farrar, Straus & Giroux.

Ricoeur, P. (1977). The question of proof in Freud's psychoanalytic writings. *Journal of the American Psychoanalytic Association, 25*, 835–871.

Rosenfeld, H. (1950). Notes on the psychopathology of confusional states in chronic schizophrenias. *International Journal of Psycho-Analysis, 31*, 132–137.

Rosenfeld, H. (1964). On the psychopathology of narcissism: A clinical approach. *International Journal of Psycho-Analysis, 45*, 332–337.

Rosenfeld, H. (1988). Contribution to the psychopathology of psychotic states: The importance of projective identificatioin in the ego structure and object relations of the psychotic patient. In E. B. Spillius (Ed.), *Melanie Klein today* (pp. 117–137). London: Routledge. (Original work published 1971.)

Sandler, J. (1976). Countertransference and role-responsiveness. *International Review of Psycho-Analysis, 3*, 43–47.

Sandler, J. (1987). The background of safety. In *From safety to superego: Selected papers of Joseph Sandler* (pp. 1–8). London: Karnac. (Original work published 1960).

Sandler, J. (1985). Towards a reconsideration of the psychoanalytic theory of motivation. *Bulletin of the Anna Freud Centre, 8*, 223–243.

Sandler, J., & Fonagy, P. (Eds.) (1997). *Recovered memories of abuse: True or false?* London: Karnac Books.

Sandler, J., & Joffe, W. G. (1969). Towards a basic psychoanalytic model. *International Journal of Psycho-Analysis, 50*, 79–90.

Sandler, J., & Rosenblatt, B. (1962). The concept of the representational world. *The Psychoanalytic Study of the Child, 17*, 128–145.

Sandler, J., & Sandler, A.-M. (1987). The past unconscious, the present unconscious and the vicissitudes of guilt. *International Journal of Psycho-Analysis, 68*, 331–341.

Schacter, D. L. (1992). Understanding implicit memory: A cognitive neuroscience approach. *American Psychologist, 47*, 559–569.

Schafer, R. (1976). *A new language for psychoanalysis*. New Haven, CT: Yale University Press.

Schwartz, L. (1978). Review of "The Restoration of the Self" by Heinz Kohut. *Psychoanalytic Quarterly, XLVII*, 436–443.

Schwartz, M. A. (1991). The nature and classification of the personality disorders: A reexamination of basic premises. *Journal of Personality Disorders, 5*, 25–30.

Schwartz, M. A., & Wiggins, O. P. (1988). Perspectivism and the methods of psychiatry. *Comprehensive Psychiatry, 29*, 237–251.

Segal, H. (1957). Notes on symbol formation. *International Journal of Psycho-Analysis, 38*, 391–397.

Settlage, C. F. (1977). The psychoanalytic understanding of narcissistic and borderline personality disorders: Advances in developmental theory. *Journal of the American Psychoanalytic Association, 25*, 805–833.

Shapiro, T. (1989). Psychoanalytic classification and empiricism with borderline personality disorder as a model. *Journal of Consulting and Clinical Psychology, 57*, 187–194.

Shedler, J., & Block, J. (1990). Adolescent drug use and psychological health: A longitudinal inquiry. *American Psychologist, 45*, 612–630.

Shengold, L. (1989). *Soul murder: The effects of childhood abuse and deprivation*. New York: Ballantine Books.

Skolnick, A. (1986). Early attachment and personal relationships across the life course. In D. F. P. Baltes, D. L. Featherman, & R. Lerner (Eds.), *Lifespan development and behaviour* (Vol. 7, pp. 173–206). Hillsdale, NJ: Erlbaum.

Slap, J., & Trunnel, E. E. (1987). Reflections on the self state dream. *Psychoanalytic Quarterly, 61*, 251–262.

Spence, D. (1987). *The Freudian metaphor*. New York: Norton.

Spence, D. P. (1982). *Narrative truth and historical truth: Meaning and interpretation in psychoanalysis*. New York: Norton.

Spillius, E. B. (1992). Clinical experiences of projective identification. In R. Anderson (Ed.), *Clinical lectures on Klein and Bion* (pp. 59–73). London: Routledge.

Spitz, R. (1965). *The first year of life*. New York: International Universities Press.

Spitz, R. A. (1957). *No and yes: On the genesis of human communication*. New York: International Universities Press.

Spitz, R. A. (1959). *A genetic field theory of ego formation: Its implications for pathology*. New York: International Universities Press.

Spruiell, V. (1988). The indivisibility of Freudian object relations and drive theories. *Psychoanalytic Quarterly, LVII*, 597–625.

Sroufe, L. A. (1989). Pathways to adaptation and maladaptation: Psychopathology as a developmental deviation. In D. Cicchetti (Ed.), *Rochester symposium on developmental psychopathology: The emergence of a discipline* (pp. 13–40). Hillsdale, NJ: Erlbaum.

Stern, D. N. (1985). *The interpersonal world of the infant: A view from psychoanalysis and developmental psychology*. New York: Basic Books.

Stone, M. H. (1990). *The fate of borderline patients: Successful outcome and psychiatric practice*. New York: Guilford Press.

Sullivan, H. S. (1953). *The interpersonal theory of psychiatry*. New York: Norton.

Tronick, E., Als, H., Adamson, L., Wise, S., & Brazelton, T. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *Journal of Child Psychiatry*, *17,* 1–13.

Tyson, P., & Tyson, R. L. (1990). *Psychoanalytic theories of development: An integration*. New Haven, CT: Yale University Press.

Waelder, R. (1960). *Basic theory of psychoanalysis*. New York: International Universities Press.

Westen, D. (1991). Social cognition and object relations. *Psychological Bulletin*, *109,* 429–455.

Winnicott, D. W. (1956). Mirror role of mother and family in child development. In D. W. Winnicott (Ed.), *Playing and reality* (pp. 111–118). London: Tavistock.

Winnicott, D. W. (1958a). *Collected papers: Through paediatrics to psycho-analysis*. London: Tavistock.

Winnicott, D. W. (1958b). Primary maternal preoccupation. In D. W. Winnicott (Ed.), *Collected papers: Through paediatrics to psycho-analysis* (pp. 300–305). London: Tavistock. (Original work published 1954.)

Winnicott, D. W. (1962). The theory of the parent–infant relationship—further remarks. *International Journal of Psycho-Analysis*, *43,* 238–245.

Winnicott, D. W. (1965a). Ego distortion in terms of true and false. In *The maturational process and the facilitating environment* (pp. 140–152). New York: International Universities Press. (Original work published 1960.)

Winnicott, D. W. (1965b). *The maturational process and the facilitating environment*. London: Hogarth.

Winnicott, D. W. (1965c). The theory of the parent–infant relationship. In *The maturational process and the facilitating environment* (pp. 37–55). New York: International Universities Press. (Original work published 1960.)

Winnicott, D. W. (1971a). *Playing and reality*. London: Tavistock.

Winnicott, D. W. (1971b). Transitional objects and transitional phenomena. In D. W. Winnicott (Ed.), *Playing and reality* (pp. 1–25). London: Tavistock.

Wolf, E. S. (1988). Case discussion and position statement. *Psychoanalytic Inquiry*, *8,* 546–551.

Yorke, C. (1971). Some suggestions for a critique of Kleinian psychology. *The Psychoanalytic Study of the Child*, *26,* 129–155.

Yorke, C., Wiseberg, S., & Freeman, T. (1989). *Development and psychopathology: Studies in psychoanalytic psychiatry*. New Haven, CT: Yale University Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.15

Phenomenological, Existential, and Humanistic Foundations for Psychology as a Human Science

CONSTANCE T. FISCHER
Duquesne University, Pittsburgh, PA, USA

1.15.1 INTRODUCTION 449
1.15.2 PHENOMENOLOGY 451
1.15.3 EXISTENTIALISM 453
1.15.4 PHENOMENOLOGICAL–EXISTENTIAL PSYCHIATRY 455
1.15.5 HUMANISTIC PSYCHOLOGY 457
1.15.6 HUMAN SCIENCE PSYCHOLOGY 458
1.15.7 EMPIRICAL PHENOMENOLOGICAL RESEARCH METHOD 460
  1.15.7.1 Procedures 461
    1.15.7.1.1 Collecting instances 462
    1.15.7.1.2 Organizing the data 462
    1.15.7.1.3 Data analysis and forms of results 462
  1.15.7.2 Clarifications 463
  1.15.7.3 Example: Being in Privacy 464
  1.15.7.4 Example: Being Impatient 464
  1.15.7.5 Comments 464
1.15.8 HUMAN SCIENCE PSYCHOLOGICAL ASSESSMENT 465
  1.15.8.1 Developing and Contextualizing Issues for the Assessment 465
  1.15.8.2 Using Tests 466
  1.15.8.3 Collaborating and Intervening 467
  1.15.8.4 Writing Reports 468
  1.15.8.5 Judging Validity 469
  1.15.8.6 Comments 469
1.15.9 CONCLUDING REMARKS 469
1.15.10 REFERENCES 470

1.15.1 INTRODUCTION

This chapter highlights the historical foundations of psychology as a human science, a rigorous, empirically-based study of humans *as* humans. That is, its approach to humans takes into account human characteristics that include but go beyond those of the objects, materials, and processes that are addressed by the natural sciences. These human characteristics include our experiencing and acting in accordance with meanings, such as anticipated futures, pasts that are alive in the present, and the world that we relate to personally.

In contrast, for the most part psychology has achieved its rightful status as a science by adopting the assumptions and methods of early twentieth century natural sciences, particularly physics and physiology. As this *Comprehensive Clinical Psychology* series attests, that adoption has been beneficial for clinical psychology as well as for our discipline's other specialties. However, there has been a cost: the content of traditional scientific psychology has been limited to subject matter that is amenable to the experimental and statistical methods of natural science: "If we can't count it, it doesn't count as psychology." As practicing clinicians we often have to figure out how to integrate the science we were taught with what was left out—human striving, hope, moral sensitivity, and the like—which are not reducible to biology and independent determinants. Although contemporary cognitive psychology allows much more of life into psychology, it both explains away some of it and does not quite address how uniquely human content is to be accounted for scientifically. Even as researchers and theoreticians, we find that we often have relied on aspects of life not addressed by our science in order to develop hypotheses and links among theoretical themes.

Phenomenological and existential movements have contributed philosophical foundations for a human science psychology. Humanistic psychology has promoted positive values in theory and practice. Briefly, phenomenological philosophy notes that the starting point of any science is the direct appearance (phenomenon) of subject matter to human observation. Analysis of appearances allows us to discover how humans construe meanings—the basis of all knowing, including that of the natural sciences. From a human science perspective, when we study biological (including neurological and physiological) aspects of being human, it is appropriate to use the procedures developed by those disciplines; but we should be mindful that these procedures, observations, and sense-making are all human undertakings and construals. Moreover, the studied biological aspects are always of a person; each implies and influences the other. Hence, we ought not automatically impose experimental designs and explanations based on theories and methods developed for studying nonhumans. Again, from a human science perspective, when we study humans as such, we ought to be faithful to human reality, and to develop research methods explicitly suited to that purpose.

Existential philosophy addresses the universal human condition of each individual having repeatedly to make personal decisions and sense of his or her existence, with no guarantees except that everything human carries co-authored meanings and that life ends. Human science psychologists hold that when studying human situations as such, at some point researchers ought to attend to whatever existential themes become apparent, rather than "controlling" them out for the sake of "objective science."

The founding phenomenological and existential authors were European, primarily German and French. Humanistic psychology is a uniquely North American movement, arising in the 1960s both as a popular protest against dehumanizing trends in the culture at large, and as a course championed by certain psychologists. In its early years, it was characterized as the "human potential movement," reflecting its insistence that society in general and psychology in particular ought to address not only dysfunctions, limits, deficits, and their supposed causes, but individuals' potential for growth and for positive experiences such as joy, love, creativity, and community. Contemporary humanistic psychology continues to remind us that research, theories, and practices bring consequences to society and to individuals, and that we ought to be mindful of values as we practice psychology.

Applied human science psychology occurs within an existential–phenomenological or related hermeneutic frame. "Hermeneutic" here implies an approach which acknowledges that understandings necessarily are developed through perspectival meanings—interpretations. Human science practitioners are developing clinical and research methods, and bodies of understanding, appropriate to humans as beings who are always in relation to their world(s). In that these beings are always involved in interpersonal and societal, and hence, moral undertakings, humanistic psychology's concerns are ever present. Although a human science approach to psychology respects, and indeed in some instances counts on, the work of natural science psychology, it understands that work within a broader frame.

This chapter will now go on to address phenomenological, existential, and humanistic movements in more detail. Each section is written for readers not particularly familiar with its approach rather than for psychologists already working within this orientation. Intraorientation disputes are bypassed. The chapter will then turn to human science psychology, which draws on the above traditions. The connections of human science psychology with hermeneutics, the "posts" (postmodernism, postenlightenment, postpositivism), and constructionism will be mentioned. Finally, because most readers are not familiar with the implications for the practice of a human science psychology, separate sections are devoted to its mainstay research method, empirical phenomenological investigation, and to human science psychological assessment. These sections describe practical differences that flow from this alternative philosophy of science. Psychotherapy is not addressed in a separate section because it appears earlier in Section 1.15.4, and because a chapter on existential psychotherapy appears in a later volume of this series.

What follows is written in the United States, to psychologists, in the last couple of years of the twentieth century. This rendition is colored by my long affiliation with the Psychology Department of Duquesne University, which began to develop its existential–phenomenological approach to psychology as a human science some 30 years ago. Other routes to a human science psychology also could be written. Certainly hundreds of additional references could have been included.

1.15.2 PHENOMENOLOGY

The phenomenology addressed in this chapter is only remotely related to two common uses of the term in North America. The first refers simply to taking experience seriously, in its own right, as Carl Rogers did, or as a journalist does when summarizing reported experiences. In the present context these practices would be characterized as *phenomenal* rather than as phenomenological. That is, they refer to experience, but not to the further analysis of that experience which would develop its structure and implications for understanding the human construction of situations. "Phenomenology" in this chapter refers to just such study of ("ology") how an event or object appears ("phenomenon"). The second usage common in North America is a medical one, in which "phenomenology" refers to similar outward appearances despite diverse etiologies. For example, we might encounter an article on the phenomenology of schizophrenia, only to find that it dealt with common externally identifiable symptoms rather than with schizophrenic experience.

So-called beginnings always have their predecessors. Still, as we look back from contemporary phenomenology, its foundational philosopher is Edmund Husserl (1859–1938). Martin Heidegger (1889–1976), who addressed existential themes in his own phenomenology, has been more influential in contemporary applied psychology. Their critiques of psychology's philosophy of science continue to challenge our assumptions about humans, about ways of knowing, and about the nature of science and of psychology.

Husserl earned his doctorate in mathematics, and attempted in his first book, *Philsophie der Arithmatik*, published in 1891, to demonstrate that the basis of logic and mathematics was psychological, that is, that they were based only on the brain's operations. He later retracted and refuted this position, known as "psychologism," which was in large part an effort to establish absolute foundations for knowledge. In his critique, Husserl noted that psychological laws are generalizations from empirical observations; that is, logic is used by psychologists to gather data and validate findings. Logic thus turns out to be more reasonably viewed as a human means of checking on thinking, rather than being reducible to brain functions. Biology is necessary for thought, but not sufficient to account for human reasoning and self-reflection.

Three centuries earlier Descartes, too, had sought to establish a foundation for absolute truth. His approach was to systematically doubt all assertions until he reached the indubitable: "I think, therefore I am." The cost of this proof, however, was a radical separation of mind and matter, a separation with which psychology still contends. Husserl initiated phenomenology's resolution to that artificial distinction. In a very significant sense, Husserl *un*separated mind and material. In contrast to Descartes' method of systematically doubting the existence of objects until ultimate certainty was discovered in the "I who thinks and doubts," Husserl's method was to return respectfully to things in the world, considering them in their own right. By "things" he meant any things of which we are conscious—*phenomena*. Experience thereby was expanded beyond sense perception, another shift of importance for psychology. Moreover, Husserl contended that consciousness is not isolated, separate from the world; consciousness is always of something, and that "something" is necessarily laden with meaning. The term "intentionality" in philosophy often refers not to purposiveness but to this inevitable attending to, intending toward of consciousness. Although we of course can distinguish our thought processes or perception (a "noetic" focus) from what we are conscious of (a "noematic" focus), the two imply each other.

Husserl referred to the everyday ways in which people are attuned to the world as the "natural attitude," a prephilosophical attitude. When we are being scientific, we adopt a much narrower stance. Within that specialized frame, we can accomplish positive advances in conceptions and technology. However, Husserl's critique remains salient for today's Western psychology: as we go about our lives as striving, hoping, creative, disciplined scientists, we have forgotten that we are more fully human than the specialized attitude takes into account. Similarly, we are quick to explain and treat depression in terms of serotonin levels, while being forgetful that even in instances of biological predisposition, a person's struggles, losses, and defeats are also essential for the development of depression. In other words, we too readily explain all human activity in terms of behavior, sensation, neurology, or cognition, thereby ignoring or explaining away the perceiving, striving person. Working in terms of restrictive concepts can be decidedly useful for particular purposes, but our findings too are restricted when we fail to return to the experiencing person.

Husserl's philosophical method avoids severing the world (e.g., in the form of scientific categories and measurements) from the person who is in relation to that world. The phenomenological epoché involves questioning and for the moment putting aside (bracketing) one's presuppositions related to both theory and the natural attitude. As we then examine a phenomenon in its full variety, we become aware of surprises and of instances that do not fit our anticipations; we discover our assumptions, which we bracket to look anew. Husserl acknowledged, however, that although we can bracket even the question of whether the world is real, we cannot bracket our own consciousness which is an integral aspect of whatever we examine.

Husserl intended that phenomenology become a science of essences—of what was essential for any phenomenon to be that particular phenomenon. The *eidos* (essence) was known through remaining actively present to the object of consciousness and its variations, through forming language about what appeared, and through comparing these appearances to related phenomena. "Consciousness" thus is our active presence to whatever we are attending to. Husserl anticipated that various disciplines would pursue different phenomena and essences. He meant for this project to retain the unity of consciousness and world, in particular to bypass both subjectivism and empiricism. However, many of his contemporaries and many of today's philosophers and psychologists have regarded this eidetic phenomenology as moving into idealism, that is, as emphasizing thought at the expense of the world.

Nevertheless, Husserl's spelling out that consciousness is necessary for any experience, and that it is necessary for unifying our fleeting impressions into coherent ones, has had an enduring positive impact. In particular, in human science contexts, consciousness is regarded as an active contribution to our experience; nevermore is consciousness only a passive recipient of input.

In part responding to criticism of his abstract, detached concern with essences, Husserl in his later years expanded his earlier view of the importance of the *Lebenswelt*—the lived world. The lived world is just that, the world as we live it prior to theorizing. The hyphens in expressions such as "being-in-the-world" are intended to convey that humans are not in the world in the same manner as coffee is in a cup, but rather that humans are always in a unitary relation with the world. That relation is the (necessary) context or horizon within or against which we perceive. In turn, as we study our perceptions of objects we also draw out themes of, and better understand, that lived background, world. Husserl's most influential books on this continent have been *Ideas: General introduction to pure phenomenology* (1913/1962) and *The crisis of European sciences and transcendental phenomenology* (1935–37/1970).

Heidegger studied with Husserl, was his assistant, and later was awarded what had been Husserl's chair at Freiburg University. However, Heidegger developed phenomenology in a decidedly existential direction. That is, he left behind Husserl's rather sterile focus on consciousness and its objects, and instead centered his study on being-in-the-world. Heidegger's academic background and the cultural context of his early adult years differed significantly from those of Husserl. Heidegger's basic graduate training was not in mathematics or science, but in philosophy and the history of theology. His historical context was a Europe coping with World War I, concerned about the effects of the industrial revolution, questioning society and its notions of "progress," and seeing both nationalism and finished philosophical systems as stultifying and dangerous. Soren Kierkegaard's (1813–1855) anguished and earnest writings, which came to be known as Christian existential works, were republished in this era. The educated public consumed Dostoyevsky's novels, with their vivid accounts of complex motivations, people caught in circumstances, introspection about responsibility, and life's uncertainties and agonies.

Within that context, Heidegger turned his attention to questions of human existence beyond Husserl's concern with consciousness. Heidegger explicitly eschewed metaphysical philosophy with its various claims of ultimate knowledge of the nature and organization of matter and causes. Unfortunately, his ontological philosophy is not as accessible to most readers as are the writings of phenomenological and existential psychologists and psychiatrists. In addition, Heidegger's penchant for relying on arcane etymological derivations for his distinctive use of German words has made his works difficult to read even for Germans. Moreover, translators, faced with the dilemma of having no comparable terms in English, often have retained his original German. Heidegger's major, most influential, text is *Being and time* (1927/1962).

Heidegger set himself the task of developing an ontology—an explication of the meaning of Being. The capitalization indicates concern with the *nature of* existence and being, in contradistinction to concern with actual beings. Heidegger's term for human Being, "*Dasein*" ("da" = there; "sein" = being), emphasizes that human reality is always situated, always in relation to a surrounding "there," never without environment. However, *Dasein* does not refer to individual humans, as has sometimes been misassumed (understandably, and usually productively). Rather, for Heidegger *Dasein* refers to the character of existence as being both open to the world's possibilities and as always finding itself already in the world. We do, however, explore human Being through examining actual individuals' basic modes of being-in-the-world (*existenitiala*).

Care is one such mode. We find that events and things matter to us; we cannot help but to care. But we care in authentic and in inauthentic ways. When we are open to the depth, newness, and possibilities of what we are attending to, and when we heed our responsibility for choosing that openness, we are living authentically. When we turn other beings into fixed objects or adopt a closed, unreflectively technological, or judgmental attitude, we are living inauthentically. Of course it is impossible to go about life being continuously or fully authentic. We are stuck with always being at least partially inauthentic; the challenge is to strive for authenticity. Our finitude, our always being-toward-death, is another of the basic modes of being-in-the world. Our finitude reminds us of our responsibility to make our choices within our limited time. We are always caught within temporal moments, coming from and going toward. Heidegger's distinguishing clock time from temporality is a particularly influential theme for psychology. As human-science researchers and clinicians we are keenly aware that lived time differs meaningfully from measured time (see Slife, 1993).

Heidegger extended his analysis of finitude and temporality to include our being inevitably historical—caught in historical times and moments. Our language and understandings are historically formed, hence we cannot arrive at an absolute or final interpretation of Being. However, *Dasein* is openness to future as well as to past. This basic human way of being (toward the future) has been ignored both by Husserl's eidetic phenomenology and by contemporary mainstream psychology. In addressing Being beyond consciousness, Heidegger established a hermeneutics ("interpretation") of existence; his philosophical hermeneutics also addressed the assumptions of social sciences' methods.

While Heidegger focused on interpretation of Being, other philosophers with similar interests have focused on particular topics such as language (e.g., Gadamer) and symbolism in psychoanalysis and religious experience (Ricouer). Psychologists have employed the hermeneutic philosophical tradition (see Hoy, 1978) both to examine individuals' ways of taking up their lives, and to research phenomena—to explore a particular experience and its lived-world ground across instances. The experiences as reported by individuals are referred to as phenomenal instances, instances of unexamined living of a phenomenon (such as being anxious, or bored, or whatever). A phenomenological comprehension is yielded by (hermeneutically) examining instances, asking what becomes evident about how phenomenal appearances reveal and instantiate general human foundations of meaning-making, such as caring and temporality, in the case of a particular phenomenon. See Packer and Addison (1989) for a broader range of hermeneutic psychological studies, where "hermeneutic" refers to the interpretive (meaning-making) character of the research, and where researchers acknowledge perspective, ambiguity, and the always unfinished aspect of understanding. That understanding begins in the life world, where the researcher reflects on possible understandings of the subject matter, then has a dialog with his or her own and others' prior understandings, and circles back to the life-world phenomenon to deepen and revise the earlier understanding. Hermeneutic interpretation does not explain the phenomenon in terms of external or underlying, or any other, variables; it remains with what the researcher is present to in the life world. This interpretation is *not* a translation into established systems of knowledge.

Readings in regard to Husserl, Heidegger, and other phenomenological philosophers include: Edie (1987), Kockelmans (1985), Natanson (1966), and Spiegelberg (1965). Also see *The encyclopedia of phenomenology* (Embree et al., 1997).

1.15.3 EXISTENTIALISM

Jean-Paul Sartre (1905–1980), probably the name North Americans associate most readily with existential philosophy, wrote in close dialog with the works of Husserl and Heidegger. In his own way, he appropriated consciousness as a constituting, construing character of being human. He took up the notions of humans being temporal and situated—always in relation. His *Being and nothingness* (1943/1956) addresses consciousness, Being, and their relation. For Sartre, consciousness is forever striving either to be at one with Being or to become a concrete something. Sartre's famous sentence, "Existence precedes essence" refers to this necessity of constructing one's meanings, because they are not given in our merely existing. Hence also we are "condemned to freedom" to perpetually choose our meanings, especially through our actions.

In much of his work, Sartre emphasized the emptiness of our future being a "not yet" and of our past being a "no longer." For him, the basic character of consciousness is a perpetual dialectic between Being and Nothingness. Sartre certainly is not a philosopher of joy! Not all existential writers stress the anguish of dealing with the possibility of the void, but all do, in one way or another, emphasize the bad faith of not acknowledging both our situatedness and the availability of at least some degree of choice. Most existentially oriented philosophers and psychologists are sympathetic with Sartre's amplification of Heidegger's concerns about our falling into a technological attitude, characterizing others as "they," and living as though we too are objects—already complete, determinate. Sartre wrote primarily of such failures in our efforts to be open to our own and others' subjectivity, freedom, humanity. Others, especially Martin Buber in his *I and thou* (1958), have emphasized that it is through recognizing another person's humanity that we discover our own. Our social context renders us radically intersubjective. Authentic community is possible only through the efforts to affirm our own and others' openness to perspective and possibility.

This standpoint was at odds with psychoanalysis, of which Sartre was a serious reader and critic. He developed an alternative existential psychoanalysis, not as a system of treatment but as a study of motivation, conflict, and of what others identified as neurosis. He deciphered the dynamics of behavior in terms of a person's earlier fundamental choices and of his or her continuing existential projects. He published compelling existential analyses of Genet, Flaubert, and Baudelaire.

Sartre's impact on psychiatry and clinical psychology, as well as on the educated public, has been significant. Existential–phenomenological clinicians have emphasized that often what we diagnose as pathological depression is the state of a person struggling with choices that he or she sees as determining who he or she will be. In seeming paradox, it can be liberating to acknowledge that one is struggling with, "condemned to," an existential dilemma, rather than with "having a depression." (As will be seen in the following brief discussion of Merleau-Ponty's work, the body's chemistry can be seen as participating in the depressed state, rather than as originating it.) Even a person caught in a brain-damaged body, or living out a genetic predisposition to schizophrenia, nevertheless still participates in forming the meanings of his or her circumstance, in making continuing choices, and in shaping his or her brain-damaged or schizophrenic life.

Sartre's (1948) analysis of emotion as a "bad faith" attempt to magically escape from the meaning of one's situation has been useful in ordinary life and in clinical practice. The author's own empirical phenomenological research on becoming angry, for example, echoes Sartre in finding that the angry person's seemingly powerful protest against whatever is blocking his or her way is a self-deceptive effort not to attend to a sense of being made helpless or of feeling demeaned (Fischer, 1998).

Maurice Merleau-Ponty (1908–1961) was a contemporary of Sartre's, but from his position in psychology at the Sorbonne he addressed different issues, in particular those of then contemporary psychology: perceptual theory, behaviorism, and neuropsychology. In *The phenomenology of perception* (1945/1962), Merleau-Ponty bypassed the old mind–body split by describing persons as being embodied consciousness. We are our bodies even as we transcend their physicality. Our bodily being is an opening onto things; our body and things imply each other, for example, as one reaches for a cup of coffee. Moreover, we perceive only against horizons, against physical settings and biographical contexts. As horizons shift in accordance with our movement and our efforts toward clarity, perception too shifts. In this sense what we know through perception is always ambiguous, never once-and-for-all; likewise knowledge. This state of affairs is not to be lamented, but rather to be acknowledged against claims of fixed meanings and of absolute knowledge.

In the above work, and in his *The structure of behavior* (1942/1963), Merleau-Ponty described three interpenetrating orders of being. As bodily beings, we are part of the *physical order*, just like rocks, and other objects that "obey the laws of nature." We also participate in the *biological order*, just like other organic, animate beings. Finally, we participate in the *human order*, which is grounded in and limited by the other orders but also surpasses them. The human order affects the others: a strained life can occasion a heart attack or clinical depression. Reciprocally, meditation or a human-made pill can alleviate biological strain. By now, as we enter the twenty-first century, these ideas do not seem the least bit controversial, but mainstream psychology has not adopted a philosophy of science that can accommodate such observations.

Let me mention a few other major existential authors before moving on to address existential psychiatry and psychology. Victor Frankl's (1905–1997) short, readable *Man's search for meaning: From death camp to existentialism* (1946/1962) has been widely read. Its compelling account of the psychiatrist's survival in a Nazi concentration camp gave credence to the importance of actively creating personal meaning. Rollo May, as first editor of *Existence: A new dimension in psychiatry and psychology* (1958), although writing for a professional audience, opened the way for English-speaking readers to explore what was happening in continental philosophy and psychiatry. Although in later years, May wrote from a more humanistic concern to wider audiences, the introduction he wrote for this book served as a solid introduction to existentialism and phenomenology for many North Americans. Adrian van Kaam's (1966) *Existential foundations of psychology* is a classic text. In more recent times, Irving Yalom (1980) has written a widely read version of *Existential psychotherapy* that presents as its guiding existential themes the necessity of dealing with death, freedom, isolation, and meaninglessness. The author's own, extended, version of existential themes includes finitude (limitation), choice and meaning, intersubjectivity, and lived body (see Leder, 1990, in addition to Merleau-Ponty, in regard to the latter; see Fischer, 1991, for an overview of phenomenological–existential psychotherapy).

1.15.4 PHENOMENOLOGICAL–EXISTENTIAL PSYCHIATRY

From the 1920s to the 1960s, while North American psychiatry imported psychoanalytic theory and method, and North American psychology developed behaviorism and diagnostic assessment instruments, a significant segment of European psychiatry created an approach that addressed the work of both Freud and Heidegger (see Spiegelberg's (1972) *Phenomenology in psychology and psychiatry: A historical introduction*). The two major figures were Swiss.

Ludwig Binswanger (1881–1966) and Medard Boss (1903–1990) both described their work as *daseins*-analysis in contrast to psychoanalysis. In both cases Heidegger's *Dasein*—humans' always being in relation—was taken up at an existential rather than at an ontological level. Binswanger later wrote of his productive error in applying Heidegger's notions directly to actual beings rather than understanding his characterizations as being about the nature of human Being. It was indeed productive. Binswanger's early training was with Bleuler and Jung at a clinic in Zurich. He opted out of an academic appointment, and instead served as director of a Swiss sanatorium (Bellevue), where over the years he hosted such foundational scholars as Husserl, Heidegger, Freud, Pfander, Scheler, Cassirer, and Buber.

Binswanger remained steadfast in his concern that psychiatry should be properly, rigorously, scientific; that is, it should be open to person-in-world rather than arbitrarily narrow its focus to presumed internal states and historical causes. His applied phenomenological method included "bracketing," putting aside theory and philosophy while allowing more and more of a person's being-in-the-world to become apparent through a respectful, caring presence to that person. He collated his findings across instances of those ways of being-in-the-world we call schizophrenia, mania, and melancholia. Perhaps his best known individual analyses are "The Case of Ellen West," found in May *et al.*'s (1958) *Existence*, and "The Case of Lola Voss," found in Needleman's (1963) edited volume, *Being-in-the-world.* Binswanger never regarded his efforts as settled accomplishments nor even as building toward a complete system, but rather as explorations beyond the artificial constraints of a mechanistic psychiatry. Many of his adaptations of Heidegger's notions, however, have provided readers with an alternative access to ordinary living as well as to restricted or disordered existence.

Binswanger emphasized the we-relationship, a we-hood, between the clinician and the other person, and he studied not "personality" but rather the patient's worlds—his or her multiple relationships. Broadly, those worlds may be addressed in terms of an *umwelt*—one's relation to his or her nonpersonal environment, an *eigenwelt*—one's private world or relations with self, and a *mitwelt*—one's social relations. Always he attended to temporal and spatial unfoldings of personal topographies, such as Ellen West's ethereal, tomb, and action worlds. Binswanger described various "failures of *Dasein*," notably those of losing one's way, finding one's meanings mixed up, and settling for a life of mannerisms.

Binswanger did not attempt to develop a system of therapy, but he did agree that his *Daseinsanalyse* carried implications for helping persons whose ways of being in the world had become problematic. He advised allowing the patient to explore how he or she had lost his or her way, and he encouraged the patient's re-exploring our shared world, both in and beyond therapy sessions. He regarded therapy as an interpersonal encounter, and transference as a present encounter between the participants. He regarded dreams as reflecting one's relations, that is, one's lived worlds, and as indicating openings to future choices.

Medard Boss's work is more systematic and more widely published in English than that of Binswanger (see Boss, 1949, 1958, 1963; Moss, 1978). He respectfully bridged Freud's psychoanalytic method into phenomenological psychiatry. He received his medical education at the University of Zurich, worked with Bleuler, received Freudian psychoanalytic training, and continued his psychoanalytic studies in London and Berlin with Ernest Jones, Horney, and Fenichel. After a long collaboration with Jung, Boss ultimately found himself dissatisfied with the deterministic assumptions of both Freud's and Jung's theories. During World War II Boss studied Binswanger's work, delved into Heidegger's writing, and then consulted with Heidegger directly. Boss held seminars in Zurich at which Heidegger spoke to psychiatrists and other physicians.

Boss noted common references in Freud and Heidegger to openness and freedom. He regarded his own phenomenologically inspired *Daseinsanalysis* as remaining truer to Freud's insights than was Freud himself when Freud converted these insignts into the language of the prevailing natural sciences. However, Boss found psychoanalytic *technique* to be highly compatible with Heidegger's writings. For example, following the rule of free association returns the patient to an openness to experience and meanings, through which he or she loosens self-imposed constrictions.

Free association and dream analysis uncover not just the past, but previously resisted futures, which now become newly possible in the context of an authentic, caring relationship. Freud's injunction to maintain "evenly hovering attention" is affirmed, but is accompanied by an injunction against translating the patient's reports into reductive psychoanalytic concepts. Reductive interpretations too often transfer the patient out of his or her restricted understandings into the restrictive understandings of the therapist.

Like Binswanger, Boss consistently attends to the patient's matrices of meaning, or lived worlds, as the place where the patient can experience invitations for new possibilities. He adopted Heidegger's distinction between "intervening care" and "anticipatory care," that is, a distinction between intervening by making decisions, providing interpretations, giving advice, or administering medical treatment, and intervening by alerting the patient to his or her emerging possibilities. For example, a clinician might inquire, "So might this dream be an instance of your already relating to your family as possibly becoming accepting of you?" The clinician also guides the patient by noting resistance to his or own desires. Note that the goals are other than symptom removal or conflict resolution.

Consonantly, Boss distinguishes between neurotic guilt which originates in childhood transgressions and existential guilt which is a failure to remain open to one's possibilities. Both can be terribly constraining, as one avoids looking at continuing choice, lives in a past, and dreads the future. Remaining open, as Heidegger wrote, is not an achievement, but rather is a feature of being human; we are less fully human to the extent that we turn away from possibility amidst the conditions in which we find ourselves. Both attempting to override the limits of our situations and settling for closed objectifications of self or others are inauthentic. Patients are not so much encouraged to form relationships with others or to self-disclose, but rather are encouraged to allow fundamental "world openness" to occur. Boss replaced the psychoanalytic "why?" (does the patient behave in self-defeating ways, recapitulate earlier trauma, etc.) with an existential "why not?" (take your longings seriously, move on from stultifying stances, relate to others more intimately).

The above understandings contrast markedly with psychoanalytic notions of conflicts occurring internally. Rather than seeking internal resolution, Boss encourages revised and opened relationships (with self, other people, things, places, the past and future). Boss regards "acting out" not as resistance but rather as trying out possibilities in a new setting. Likewise, he regards "transference" as the patient's bringing old ways and old perceptions to the relationship going on in therapy now. For a clinician to deny his or her additional part in the therapy encounter is inauthentic and contrary to the project of the client trying out revised ways of relating. Boss, regards *Daseins analysis* as especially demanding of the clinician, who makes use of his or her experience with the patient to encourage greater openness. He regards therapy as a genuine encounter in which the clinician is relatively selfless and restrained and especially respectful of the patient's individuality.

There are, of course, many other contributors from this period in which Europeans sought to develop psychological approaches that respected humans' relational meaning-making. Among these are Johannes Buytendijk, Henri Ey, Karl Jaspers, Eugen Minkowski, Paul Schilder, Erwin Straus, J. H. Van den Berg, and Victor von

Gebsattel. Their work has been of interest to like-minded psychiatrists and psychologists around the world. However, historically there has been little communication between these authors and mainstream psychiatry and psychology, although the former group generally has been traditionally trained and makes use of a range of theory and data. Indeed, my characterization of my own orientation is probably accurate for most phenomenological–existential practitioners: our philosophical orientation is phenomenological–existential; we make pluralistic use of developmental theories (such as those of Erikson, Sullivan, Kohut, etc.); we integrate research studies into our ongoing understandings and questions; and within this frame we make eclectic use of a range of techniques. Unlike other major approaches, no basic therapeutic method is identified with phenomenological–existential clinical practice, as the case is for psychoanalysis, behaviorism, and cognitive psychology. Several psychotherapists have developed particular techniques, useful within many frameworks; examples are Frankl's (1946/1962) paradoxical intentionism and Bugental's (1965) "What's stopping you?"

In closing this section, I offer several corrections of common misunderstandings. As reflected in the preceding sections, existential philosophy and practice do not posit unfettered free will. Nor do they emphasize dread and anguish. A second point: Europe is no longer a fountainhead of phenomenological–existential psychology and psychiatry. North American behaviorism and research design were quickly imported by academicians in Europe, Great Britain, and Scandanavia. Now, interest in human science alternatives for psychology is developing in all continents just as it is in North America.

1.15.5 HUMANISTIC PSYCHOLOGY

Humanistic psychology emerged as a full-blown professional movement in the 1960s. The California-centered countercultural "happening," which protested society's suppressions, is similarly named and is sometimes conflated with a reflectively planned movement within psychology (cf. Smith, 1990). The happening, in which assorted citizens, professionals, and many psychologists participated, also stressed the importance of aspects of life neglected by education and the social sciences: joy, creativity, love, self-affirmation, and spontaneous expression of affect and belief. Overlap between the popular and professional movements did occur. American Psychological Association (APA) annual conferences in those days were attended by sandal-footed, knap-sacked, openly breast-feeding members of the Association of Humanistic Psychology (AHP), whose meeting overlapped with the APA conference. The establishment types were dismayed as some members of the other organization interrupted symposium speakers from the audience with proclamations such as, "I just have to tell you that I love your soul!" One psychologist who earlier had published in staid journals propounded nude group therapy in swimming pools.

Unfortunately, for those psychologists who have not kept up with the literature of humanistic psychology, those earlier dramatic times still color the movement. To the contrary, the AHP's journal, *The Journal of Humanistic Psychology*, has consistently been a major source for well-considered work that explores and supports central human phenomena that are still underaddressed by mainstream education and psychology: our spiritual, creative, spontaneous, caring, transcendent, and transpersonal moments (see Bynum, 1994, in regard to transpersonal psychology). Even while acknowledging the painful and darker sides of existence, the AHP continues to champion human potential for positivity and to promote responsibility for self and to others. Other positive legacies of the mixed beginnings of humanistic psychology have been its contributions to the growth of self-help organizations, and to education and psychology at least nodding to notions of the "whole person"—the experiencing, thinking, valuing, emotional, bodily, spiritual, choosing, interpersonal, behaving person.

Indeed, dissatisfactions with the 1950s partitive, deductive, and reductive experimental research model, with sterile and deterministic behaviorism, and with pathology-oriented psychoanalytic theory, led Abraham Maslow to form a mailing group of kindred professionals. He referred to their efforts as the "third force" in psychology's evolution, following psychoanalysis and behaviorism. Their exchange of writings eventually grew into the *Journal of Humanistic Psychology* (founded in 1961). Among the first board of editors for the journal were Andras Angyal, Charlotte Buhler, Kurt Goldstein, Dorothy Lee, Rollo May, Lewis Mumford, and David Riesman. In 1963 the founding meeting of the AHP took place in Philadelphia, where nearly 100 like-minded professionals confirmed the importance of inserting the place of values into psychology's subject matter and activities. Gordon Allport, Jacques Barzun, George Kelly, Clark Moustakas, Gardner Murphy, Henry Murray, and Carl Rogers were among nationally known figures who attended the Old Saybrook (CT) follow-up conference in 1964. International meetings followed. The first psychology departments dedicated to a humanistic approach were Sonoma State College (CA) in 1966, West Georgia College in 1969, and the Humanistic Psychology Institute, HPI (San Francisco) in 1970. The HPI has since become Saybrook Institute (see deCarvalho, 1994, for an organizational history of this period).

In the 1950s and 1960s, as psychology moved increasingly out of academia and into practice, Carl Rogers' *Client-centered counseling* (1951) and *On being a person* (1961) were welcomed. He emphasized persons' capacity for growth through self-understanding, enhanced through the empathic, authentic presence of the counselor. He referred to "counseling" rather than to psychotherapy or analysis in large part because those terms belonged to the legally protected domain of medicine. Rogers' work was in striking contrast with psychiatry's emphasis on psychoanalytic theory and psychopathology, and with psychology's emphasis on learning theory and behavior modification. In academia, determinism dominated psychology, which identified itself as a science that could "predict and control" behavior. Maslow's *Toward a psychology of being* (1962) and his *I-thou knowledge in a psychology of science* (1966) were welcomed as alternative frameworks for thinking about ourselves and our clients, see also Bugental (1963).

In 1971, Division 32 (Humanistic Psychology) of the APA was established, largely through the efforts of AHP members. Its purpose was to bring the concepts, theories, and philosophy of humanistic psychology to the research, education, and professional applications of scientific psychology. Among its nationally known first board members were David Bakan, Leonard Blank, Albert Ellis, Carman Harari, James Klee, and Everett Shostrom. Among the presidents of the APA who have also been leaders within humanistic psychology are Carl Rogers, Abraham Maslow, and Brewster Smith. The Division's journal, *The Humanistic Psychologist*, publishes articles on theoretical and philosophical issues, and on methodological advances in human science research. Many members are also actively involved in the APA's Division of Theoretical and Philosophical Psychology.

Accounts of the founders of this movement, their writings, and the issues they dealt with can be found in deCarvalho (1991b), Royce and Mos (1981), and Welch, Tate, and Richards (1978). Division 32's contribution to APA's centennial publications on the divisions' histories and contributions may be found in Wertz (1994).

This brief overview is concluded with a backward- and a forward-looking note. By now, the humanism of earlier centuries has been a touch point for humanistic psychology, but in fact it was not explicitly raised as a source of inspiration in the founding years of the latter movement. The phrase "humanistic psychology" was adopted for the AHP's journal only after rejection of options such as "person psychology," "self psychology," and "orthopsychology" (Greening, 1985; Sutich, 1962); Gordon Allport had first used the term with its present meaning in 1930 (deCarvalho, 1991a). From the Renaissance onward, "humanism" has referred to humans' unique values, creativity, and reason, which exceed both Church doctrine and our material nature (see Bullock, 1985). Humanistic psychology has helped to put psychology in touch with both historical humanism and contemporary arts. Humanistically informed psychologists continue to remind us both of the place of values in our personal and work worlds, and of the importance of examining the assumptions and implications of our scientific practices.

1.15.6 HUMAN SCIENCE PSYCHOLOGY

Human science psychology is an approach to psychology's subject matter. It is not a philosophy or a theory. The approach is intended to respect humans as a subject matter that differs from nonhuman material. A thorough human science approach would encourage study of, and work with, humans' physical, biological, and meaning-making character all at once. Many clinicians attempt to practice in just that way, but when they are not familiar with an explicit foundational framework, their efforts are not as consistent as they could be.

Many researchers could work within a human science frame while continuing to pursue their psychological studies through experimental and other statistical designs, which are efficient for dealing with large samples and with the biological (and chemical, neurological, etc.) aspects of being human. However, these studies would take into account that variables do not explain behavior or experience, but rather are superimposed grids for organizing the orderliness of psychological matters. An analogy is the field of economics, where explanations and predictions are made in terms of unemployment rates, gross national product, and so on, although the human order underlies those created and imposed measures. Ultimately the measures reflect the actions of individuals living within various circumstances and working, enjoying, competing, making decisions, spending, investing, striving, risking, and so on. In both psychology and economics, the imposed grids are definitely useful, but they are most useful when we take into account that we made them up, and that we understand and serve people best when we return from our constructions to reflect on individuals going about their lives. These reflections allow us to revise our constructions.

At this point human science psychology is a critique and a call. It is a critique of mainstream assumptions and practices for their being forgetful of the fuller human context, and it is a call to develop a thorough alternative. Adopting a human science framework for psychology's ongoing work would eventually obviate the need for the term. Of course by then other critiques would emerge, and psychology would accordingly advance in new directions. Positivism similarly was adopted as a corrective to nineteenth and early twentieth century armchair philosophizing, and much of the achievements of North American psychology are the product of that adoption. By now newer generations of psychologists ask "What's positivism?," not because it is no longer around, but because it is not being taught as such. Even though many theorists say that positivism served its purpose, and now has few strict adherents, in the absence of widespread alternative views, implicit positivism determines the research designs that are supported by grants, publications, and so on. Positivism has become the invisible approach of mainstream academic psychology. Because psychology no longer has to develop and prove itself as a science, philosophy of science courses are rarely taught. Slife and Williams (1997) have called for specialty training in philosophical/theoretical psychology so that departments will have such expertise available and, once again even within experimental programs, will engage in discussions of assumptions and their implications.

Toward that end, let me briefly rehash the historically-based approach of a human science psychology, and some contemporary trends with which it is joined. William Dilthey (1833–1911) urged that psychology become a rigorous and systematic science, a model for all the Geisteswissenschaften (the disciplines of mind, spirituality, and human affairs). He argued against the new discipline modeling itself on the highly successful disciplines of time, mathematics, and the natural sciences (see Dilthey, 1894/1977; Hodges, 1944). Nevertheless, even though Wilhelm Wundt also wrote a social psychology that was closer to Dilthey's proposed human science, and William James wrote extensively on consciousness, it was their fledgling laboratories in Germany and the United States, respectively, that were taken up by the new discipline. Their philosophical essays and books were ignored in favor of hands-on, technically-oriented practice. Nevertheless, a sampling of books and essays in developmental, clinical, and social psychology through the 1940s shows that most authors were more holistic, and more historically and philosophically minded, than is the case today.

Although North American psychology has always housed a broad mix of theorists and practitioners, by the 1950s "scientific" psychology was equated with experimental psychology. By the 1960s, many of us who were not satisfied with that situation discovered the translations of existential phenomenology that were finding their way from Europe. Within this country, many critiques and calls for returning persons as such to our studies were published, for example, those by Gordon Allport, David Bakan, Roger Barker, Joseph Lyons, Robert McLeod, and Nevitt Sanford. In 1970 Amedeo Giorgi published *Psychology as a human science: A phenomenologically based approach.* Trained in experimental psychology, he had encountered its limitations as a natural science, and turned to Dilthey and phenomenological writers for foundations appropriate to studying humans as humans. His book has been a touch point for diverse readers who have looked for support and inspiration for their own kindred reflections.

Institutionally, Duquesne University's Psychology Department, shaped in the 1960s by Adrian van Kaam and Giorgi, has been a center for the systematic development of psychology as a human science. Its work, however, has been foundational and corrective and has not yet embodied a full human science research program, which would include our cultural, biological, etc., dimensions. Nevertheless, at Duquesne alone, nearly 200 empirical phenomenological research dissertations have been completed; gradually substantive content area comprehensions are taking shape. Moreover, as qualitative research methods are increasingly practiced across disciplines and continents, their practitioners are increasingly identifying themselves with a human science tradition.

Saybrook Institute, which also offers a doctorate, stressing humanistic psychology, offers substantive work in a human science tradition. The masters degree psychology programs at West Georgia College and Seattle University are dedicated to humanistic/human science psychology. Other institutions include clusters of colleagues pursuing this approach, and of course there are hundreds of individual faculty members doing likewise along with thousands of nonacademics. Many individuals do not use the "human science" title, and many are working within overlapping interest areas. A small but representative sampling would include constructivist psychotherapists, who have advanced George Kelley's (1955) analysis of how we construct our worlds and can be encouraged to construct them differently; narrative therapists, who help clients to see how they have storied their lives and might co-author them differently; and action researchers, who not only acknowledge but plan and track their influence on their field subjects. Related areas of scholarship include social constructionism, which explores how science as well as society and individuals are inevitably shaped by our ideas and interactions, and which finds that we have no access to truth aside from these constructions; linguistic studies, which explore the inherent role of language in our sense of reality and possibility for action; and feminist studies, which in a variety of ways highlight the culturally situated social construction of gendered possibility. (Leading sources on social constructionism include Berger & Luckmann, 1967, and Gergen, 1985, 1994; sources on personal construct therapy include Epting, 1984 and Leitner, 1985; linguistic and discourse studies include Barnard, 1998, and Gavey, 1997; feminist works include Harding, 1987; and Hawkensworth, 1989.)

Many other approaches and practices that nowadays carry one of the "post" labels also dovetail with human science psychology. "Post-enlightenment" refers to orientations that emerged in response to disenchantment with the enlightenment era's assumption that God had created a mathematically ordered world which mortals could uncover through empiricism and logic. "Postmodern" refers similarly to contemporary movements across the arts and sciences that have ceased to look for absolute Truth. Reality is always known only from perspectives, which are necessarily local (not universal). Contrary to some protests, this position is not one of "anything goes" relativism, but rather one that respects subgroup realities, and calls for accountability through specifying the motives and history through which particular understandings emerge (see Aanstoos, 1990). Finally, "postpositivism" refers to having superseded the notions that science requires that its knowledge be based on sense data that have been processed through mathematics based on Aristotelian logic.

To review, human science psychology differentiates itself from positivistic traditions by drawing on broad hermeneutic traditions and/or phenomenological descriptions of persons as always being in relation to the world, always temporally on the way from and toward, and on existential extensions of those descriptions of persons as always being in relation to lived body and to the social world, and as inevitably participating in making meaning and choices no matter how constrained. The mind–matter split is undone. The moral dimension of human matters is more clearly proper to psychology. The old goals of predict and control become those of describe, understand, and influence. Data developed using natural science methods, including neurological and biochemical data, are comprehended within this frame. To the extent that a body participates in the human order, its biological order is part of that structure; even sciences of the body are human constructions. Description of human perception, experience, and action is narrative rather than reductive, and it respects ambiguity as sometimes inherent to the subject matter rather than as a deficient account. Accounts are structural—retaining the mutuality of differentiated aspects of whatever is described. Understanding and influencing require respecting persons as choosing in accordance with both contingencies and personal meanings. Validity and objectivity become matters of specifying means and perspectives, and of sharing data, comprehensions, and practical outcomes, toward consensual agreement. Progress evolves as new perspectives are continually brought to bear.

Below are concrete illustrations of approaches to research and to psychological assessment that flow from a human science orientation. Representative psychotherapeutic practices have been described in Section 1.15.4 on existential psychiatry and psychology. I have drawn on work with which I have been intimately connected, and which has been developed within the Duquesne community for over 30 years, in order to provide examples of what the foregoing discussion might come down to. A wide range of nonphenomenological substantive work can be characterized as human scientific. Human science psychology will develop and contribute to the extent that variations and innovations are undertaken and discussed.

Further examples of human science, broadly conceived, can be found in the *Duquesne Studies* series edited by Giorgi and colleagues (1971, 1975, 1979, 1983) and in the series edited by Valle and colleagues (1978, 1989, 1997).

1.15.7 EMPIRICAL PHENOMENOLOGICAL RESEARCH METHOD

The term "empirical" in "empirical phenomenological research" refers to directly observed, pretheoretical experience (one's own or another's) in contrast to data derived from the use of predetermined categories and through measurement. "Empirical" also refers to the availability to other researchers of data and analytic steps, so that they can check the extent to which they come to similar findings. The basic method, developed and elaborated for over 30 years at Duquesne University, is a systematic means of addressing psychological phenomena in their own terms, that is, in terms of how they were lived. Empirical phenomenological analyses of "being in privacy" and "being impatient" will be reported below. "Phenomenological" here refers both to the more immediately accessible "what it's like" for individuals (to be impatient, etc.) and also to what researchers learn when they ask what else these phenomenal descriptions can reveal about the process of being human in that situation. In the latter regard researchers repeatedly rediscover their subjects as being temporal, always in relation and on the way, restricted and choosing, and so on. At Duquesne we have referred to a cohesive representation of these themes as they appear in, and shape, the phenomenon as its phenomenological structure.

Empirical phenomenological research is existential in that psychologists have departed from a purer phenomenological philosophy, to address people in their necessarily existential situations. "The method," moreover, is a great variety of variations and innovations, as von Eckartsberg (1986) helpfully described in *Life–world experience: Existential–phenomenological research approaches in psychology*. For other examples and discussions of variations of the method, see Aanstoos, 1987; Colaizzi, 1973, 1978; Giorgi, 1985; Richer, 1978; Walsh, 1995; and Wertz, 1983.

Qualitative research in general addresses the experience and action of persons as they engage in particular situations. Findings are in terms of the life world rather than transformations into categories and measurements. Traditional quantitative findings of course provide ground for reflection in concert with qualitative description. Many nonphenomenological qualitative research methods are also appropriate for a human science psychology. These include ethnography, linguistic and conversation analysis, case studies, ethnomethodology, and some forms of grounded theory. Psychology has been slow to become involved with qualitative research methods in that they do not fit our received, rather narrow, notions of science. In contrast, sociology, education, gender studies, human development, nursing, and counseling all have established or are building major qualitative research traditions. About a third of the papers presented at the American Educational Research Association conferences are qualitative studies. At last count, five nursing journals publish significant qualitative research. About a third of the research presentations at the Society for Psychotherapy Research conferences is qualitative. Well over half of the research papers presented at the Association for Women in Psychology Conferences use qualitative methods.

Many variations on the EP (empirical phenomenological) research procedures described here have been, and could be, undertaken. Guy Kashgarian (1997) employed a variety of linguistic analysis as part of his phenomenological analysis of instances of "being criticized" reported by persons diagnosed as narcissistic personality disorder and those not diagnosed. Martin Packer (e.g., Packer & Scott, 1992) encourages EP analyses based in part on ethnographic data.

A thorough human science research program would accommodate biological, cultural, and psychological dimensions of any studied circumstance. It is to be hoped that before long psychologists will undertake studies of the mutuality of biological and psychological depression, for example. In the meantime, this chapter focuses on Duquesne's EP research method because it offers a crucial approach to understanding human phenomena otherwise missing from psychology's efforts. EP research is the appropriate method when the question is "What is such-and-such a phenomenon?," as in "What is privacy?" or "What is being impatient?" When our questions have to do with absolute amounts and with amounts of change, traditional measurement and experimental methods are appropriate. Unfortunately, historically psychology typically has not asked the What question, and has instead resorted to operational definitions (defining something in terms of measurement criteria, as in hunger = percentage of body weight loss, or of food reduction).

1.15.7.1 Procedures

Long before beginning a formal study, one has noted instances of the phenomenon in one's own life, talked with colleagues and friends about instances in their lives, and noted occurrences in newspapers, novels, and so on. Research and theoretical literatures have been consulted for their accounts. As in all research, formal participants for the study are chosen on the combined grounds of practicality/availability and of being appropriate for the issues and literatures one hopes to address. Because EP analyses are so arduous and time-intensive, a dissertation study might involve only five or six formal participants. One funded project with multiple researchers included 50 formal subjects (Fischer & Wertz, 1979). Regardless of the number of formal participants, the study's preliminary findings are compared with other instances, both to reconsider one's work, and to make comparisons with regard to differences among participants in different situations.

1.15.7.1.1 Collecting instances

Although EP research need not be constrained to analysis of verbal reports, which is the tradition that developed at Duquesne, through faculty members' projects, more than 200 doctoral dissertations, the work of graduates, and the work of others who have based their research on the Duquesne tradition. Requests for descriptions from pilot participants result in a refined formulation of the research request that is presented to the formal participants. An example: "Please recall a situation in which you became angry. Please write out what was going on before that situation, what happened as the situation evolved, and what happened then. Include details that will help us to understand what it was like to be you throughout the situation. Thank you." Participants have agreed in advance to provide a description, and hence have been recalling long before they sit down to write. Some researchers have recorded interviews rather than requesting written accounts; usually those accounts are briefer on the one hand or rambling on the other. Written reports allow participants to organize and elaborate their recollections.

The researcher types any recorded or handwritten accounts, and reads and rereads them before returning to the participant for a follow-up interview, which is taped and later transcribed. Typically, the researcher asks if there are additions to the account that the participant might have thought of since providing the description. Then he or she is given a copy of the transcript, from which the researcher reads aloud any sections that seem to require clarification. Open-ended requests for elaboration are made, in the form of, "Could you tell me more about that?" Only at the end of the interview might the researcher ask direct questions such as "Did that occur on the same day?" or "Would you say that this instance of being impatient is typical for you?"

1.15.7.1.2 Organizing the data

The researcher typically recasts each account into chronological order, inserting interview elaborations (in a contrasting font) into the relevant text. At this point the researcher decides on a system for breaking the text into workable form. Usually this involves numbering "meaning units"—segments of text that seem to be cohesive and that are followed by a shift in the reporter's account. Some researchers number phrases, some entire paragraphs. The point is both to have a way to check that one has included all the text in the analysis, and to later be able to cite sections in every report to illustrate each finding.

1.15.7.1.3 Data analysis and forms of results

Some researchers have met regularly as a group, exchanging their own descriptions, or those that they have collected, working together to formulate their findings. Others have met regularly with the providers of accounts, to collaboratively revise their findings in discussion with the providers. Most researchers working on dissertations have worked singly, in consultation with a faculty committee. Of course one can combine these approaches.

Although variously formulated, most of the work at Duquesne includes the following processes. One immerses oneself in the descriptions, "dwelling" in them, as we say, becoming familiar with their internal cohesion and their nuances at a prearticulate level. Of course at the same time, one is noting recurrent themes, and jotting down phrases and metaphors that hint at what seems to be present in all the accounts. The method section of the research report itemizes the formal steps that were conducted and documented, and that a reader may follow along to see whether the reported findings are evident, and whether the procedures seem to have been followed. However, the formal steps do not characterize all that goes on as one becomes more deeply immersed in apprehending and giving expression to the phenomenon under study. For example, while supervising a student psychotherapist, I realize that a phrase she has just uttered is perfectly suited to characterize an aspect of what my research participants have reported. While doing my morning physical exercises, a just-right metaphor seems to come out of nowhere. Upon rereading a transcription for the umpteenth time, I realize this person's wording fits all my participants' instances. While reviewing my notes for a lecture on object relations, I somehow realize that I have been forcing a distinction on my data.

One is not discovering external, independent, features of an object, but rather is making a phenomenon visible and differentiable from related phenomena. The process requires four kinds of rigor: diligence in describing and following formal procedures, faithful presence to data, disciplined and systematic reflection on what is appearing in instances of the topic, and finally, creativity in expressing—representing—making visible—the unity and differentiations of the phenomenon. We struggle with the limitations and possibilities posed by language and by our being historically and culturally situated.

The formal steps usually include uncluttering each participant's report by dropping out information not essential to that person's experience for it to be an example of the particular phenomenon. For example, in Ms. Smith's account of being criminally victimized, we might drop out the names of her children and the color of her handbag. We might condense references to having done three loads of washing, made all the beds, and vacuumed to "after a morning of heavy housework ... " The resulting condensation might be referred to as an individual or situated description, and can serve as one form of the study's results.

Next, each condensation is analyzed by asking what is being said about where this person was going in his or her life, how he or she was in relation to self, others, and the world, and how he or she was living a past and the future. The researcher might then attempt to represent this understanding by writing a narrative version that stays close to the participant's own words but draws out experienced meanings more explicitly and in relation to the research question (e.g., "What is the experience of being criminally victimized?"). Here, the researcher might say, in part, "Having completed her planned morning of heavy housework, Ms Smith gathered her children into the family car, anticipating a successful shopping trip." This level of analysis can be presented to the participant as a check on having rendered her explicit and unarticulated experience faithfully. This version of the findings again is both one form of findings and a step toward further analysis. This step, which might be a half-page to two or three pages in length, is referred to as an individual phenomenological description or structure. By "structure" we mean that the descriptive account is cohesive in its explication of differentiable but essential aspects, and that none is made more basic than any other; each aspect implies the others.

Finally, the researcher examines each individual phenomenological structure, identifies the themes that are shared by all, and formulates a single phenomenological structure that expresses what was necessary and essential for all the accounts to be instances of the phenomenon under study. This level of finding is of course more abstract than the preceding ones. The researcher's implicit and explicit questions to the data are evident in the findings. For example, my own societal and clinical concerns become evident to me at some point as I reflect further on my findings. Levels of abstraction vary among research authors, as does style of representation. Some writers put greater stress on evoking readers' lived sense of a phenomenon, and others place greater emphasis on highlighting themes that lend themselves to dialog with philosophy. In any case, the findings must be demonstrable in every instance of formal data, and in all instances encountered outside of the study. No settling for probability levels here! Quotations from the participants' original accounts document and illustrate the findings.

1.15.7.2 Clarifications

When undertaking a study, researchers first put aside theoretical positions and what they already know about their biases, and then specify what they know about their guiding interests (e.g., making a rape victim's experience available to law enforcement, medical, and court personnel). However, throughout these steps and in the findings, the researcher repeatedly discovers his or her preconceptions (for example, that being impatient is a negative stance, or that privacy is privacy-from). He or she makes formal note of the discovery for later discussion, and in the meantime tries to hold the assumptions in abeyance. We recognize that we cannot thoroughly separate ourselves from our historical, cultural, biographical contexts nor from motives for undertaking a study, but we hold ourselves accountable for bracketing all that we can, and specifying all that we identify about the perspectives from which we have viewed our data.

Researchers follow different steps and different exercises within those steps depending on the subject matter, their own route into phenomenology, and so on. In one way or another, all follow what can be called a "dwelling and hermeneutic presence" to their data. Analysis does not occur in a finished step by finished step fashion. Rather, the researcher repeatedly immerses him- or herself in the original descriptions and in the forms of findings developed so far. Sometimes this delving is a roaming around, noticing nuances. Other times it is a reading of one account with another account in mind, allowing synergistic meanings to emerge. When a new insight or understanding emerges, the researcher returns to the earlier understandings, and reworks them in the light of the newer perspective. This systematic circling back to revise one's interpretation, the hermeneutic movement, is difficult to represent in a procedures section.

1.15.7.3 Example: Being in Privacy

The following structure of being in privacy has been useful in distinguishing privacy from secrecy. It also makes a case for the importance of solitude, even in public settings, for the development of wonder, reflection, and so on.

Instances from which the analysis was developed included being immersed in a novel until interrupted by telephone solicitor, considering a painting in an art gallery until a tour group came through, working on a math problem until the teacher looked over one's shoulder, enjoying a flower garden until chimes reminded one of obligations.

These characterizations are paraphrased from Fischer (1971) (pp. 153–154):

> As we look back, it seems that in privacy one's sense of self falls away, and one flows with whatever is being attended to. One seems nearly, but not quite, merged with that something. This attending is recalled as including an aura of familiarity, at-homeness, of things fitting even when they do so unpleasantly or newly. In privacy one's relation to subject matter may occur through reflection, action, imagination, whatever. The relation includes relative openness to and wonder at the object's possibilities, facets, realities, even when the person is in a problem-solving mode or is reflecting on an unfortunate aspect of one's life.
>
> One's attention eventually turns, seemingly in its own course, to other matters, usually things to be done. However, when privacy is disrupted prematurely, attention is torn from its prior focus, and shifts repeatedly among intruder, self as caught by the intrusion, the peripheral world, and the lost focus. One now contends with contingencies—time, space, demands. Affect varies with desire and felt capacity to deal with the intrusion, and to recapture the disrupted presence to the previously evolving subject.

1.15.7.4 Example: Being Impatient

The following Condensed General Structure is taken from Donna Coufal's (1997) dissertation. Among other themes, the dissertation discusses the surprising discovery of a moral dimension, and compares being impatient with phenomenological analyses of being angry, frustrated, and anxious. Instances provided by her participants (Ps) included being stuck in an automobile bank teller line, coping with a client who seemed not to be hearing the counselor, and being lost on a country road while trying to get to a social event.

> The awareness of time as protracted but imminently fleeting is central to the experience of being impatient. What was initially taken for granted, that time is sufficient to meet P's goals and that the future will unfold as P anticipates, is no longer certain. P's sense of time feels distorted. An earlier absence of concern about the future gives way to intensifying affect in the face of waiting. In spite of obstacles in the way, P is poised to move forward, resisting the temptation to become emotional.
>
> As P's focus narrows, the complexities of the situation go unnoticed, particularly as P becomes tired and frustrated. P tries to force the situation to evolve as planned and cannot imagine alternative courses of action. The situation is one that demands something new of P who nevertheless grasps at the familiar in the face of the unknown. P overlooks evidence that would help in navigating the situation more efficiently. In attempting to avoid waiting, P creates further delays and finds him- or herself frustrated or defeated.
>
> In denying limitations of a habitual style, P is confronted with the annihilation of previously unchallenged assumptions about him- or herself. P becomes vigilantly focused on preserving a shred of his or her identity in failing to meet his or her own expectations. P admonishes self and others while waiting. Innocent delays are experienced as moral affronts. P casts about for targets of blame, alternatively believing delays to be purposive on the part of others or reflecting inadequacies in P's own participation.
>
> In spite of the awareness that being impatient worsens the situation, any attempts to relax are defeated by P's rigid alertness. P tries to hurry in order to feel less vulnerable, but remains tense or becomes angry. Although P has restrained him- or herself from becoming emotional, when looking back on the situation, P identifies his or her responses as disproportionate to the demands of the situation.
>
> Having failed, P reflects on the situation and becomes open to its particular meanings and possibilities. Now the present and each small step P takes become more figural, while larger, future oriented goals become horizonal. Nevertheless, P arrives at destination too deflated or exhausted to enjoy any sense of accomplishment except for the modest satisfaction of having modified his or her initial plan.
>
> In contrast, when patient, P experiences openness to unexpected events. Anticipation is tolerable, even pleasurable, and P responds imaginatively when action is not yet possible, appreciating the limitations imposed by the situation, but not assuming a catastrophic outcome.

1.15.7.5 Comments

Another researcher or reader may suggest a more evocative, accurate, felicitous, or otherwise preferable word or phrase, may identify additional assumptions on the researcher's part, may point out relations to other phenomena, and so on. That is how bodies of understanding evolve. Note that I have referred throughout to "understandings" rather than to "knowledge" in order to emphasize that what we know, even though through our natural science or human science methods it is always an interpretation.

What is phenomenological research good for? It provides a means of addressing experience as such. It enhances consideration of the human order of events and how we participate in them. Within such descriptions, we discover choice points; we are freer than our psychology's literature on determinants would lead us to believe. Phenomenological research also helps us to see the relations among phenomena, and how one can turn into another, which again helps to guide our choices. This type of descriptive psychology puts us in touch with the humanities, art, literature, and drama, making them available to our psychological theories.

Adrian van Kaam's dissertation (1958) at the University of Chicago, on the experience of feeling really understood in psychotherapy, was a forerunner of the research method developed at Duquesne, where van Kaam brought together the group who initially began the department. Amedeo Giorgi took the lead in developing conceptual and procedural foundations for existential–phenomenological research. Other faculty members and graduates have contributed to and furthered their development, for example, C. Aanstoos, S. Churchill, P. Collaizzi, C. Fischer, W. Fischer, S. Halling, and P. Richer. Among phenomenologically oriented researchers who developed their research approaches independently of Duquesne are Howard Pollio at the University of Tennessee and Joseph deRivera at Clark University. Further illustrations of EP research and writings on method can be found in the *Duquesne studies in phenomenological psychology* series (Giorgi et al., 1971, 1975, 1979, 1983), in the *Journal of Phenomenological Psychology*, and in the journal *Methods: A Journal for Human Science*.

1.15.8 HUMAN SCIENCE PSYCHOLOGICAL ASSESSMENT

The author developed the following approach and practices while teaching psychological assessment within the human science psychology program at Duquesne University. One does not have to adopt the orientation of this chapter to incorporate the practices into assessment. I do think that accommodating this chapter to one's own assumptions about the nature of being human can allow one to engage a client more fully and consistently as a co-assessor, and to participate in developing tailored, viable changes in his or her life. Stephen Finn (1996) and his colleagues at the Center for Therapeutic Assessment in Austin and Leonard Handler at the University of Tennessee have developed similar approaches from orientations that overlap mine. Finn and others have published research reports on the effectiveness of, and client satisfaction with, assessment collaboration and feedback. Slowly, my textbook (Fischer, 1994), *Individualizing psychological assessment*, originally published in 1985, is increasingly referenced or adopted in graduate courses. It seems to me and my colleagues that our work in collaborative assessment validates the private inclinations of practitioners, whose university training, in contrast, had followed the laboratory tradition—in which one gathers data and draws inferences, without consulting with the subject. Availability of an explicit human science framework encourages efforts to involve clients as co-investigators.

1.15.8.1 Developing and Contextualizing Issues for the Assessment

If an assessment referral has been made by a third party, I ask that person to tell me more concretely what the issues are. Often the referral remains abstract until I ask what decisions might be influenced by the assessment. For example, in one instance, a director of an adolescent program requested an assessment of a youngster's "personality structure." Upon discussion he initially revised his request to be "evaluate Axis II diagnoses," and then after further discussion, to "assess capacity for empathic relations." When asked about what decisions he was faced with in regard to the young woman, he said that an immediate one was whether she could be placed in a therapy group that required mutual support among the members and an ability to introspect. He agreed with the assessor that he also was trying to determine how psychologically developed the client was so as to anticipate her level of involvement with the overall program and to determine whether several months of participation might or might not make a difference. The assessor then asked for information and observations that had been gathered on both sides of the issues. It turned out that the client's family had been unstable in many ways all her life, and that she had been truant from school regularly (and was failing most classes), had a reputation for defending herself from insults through physical retaliation, was suspected of being responsible for items missing from the school lunch room, and was seen as a loner at school. On the other hand, she had earned a "B+" from an English teacher she seemed to like, had taught a younger brother to read, and had impressed the Program's social worker as being sad rather than as defiant and as having a definite, albeit dry, sense of humor.

Often, this "getting down to earth" process suggests next steps that enable bypassing formal assessment. On the other hand, just a few consultations, such as that above, with referring staff members, results in assessment requests that are already accompanied by concrete events from a client's life and by a clear notion of how the assessment might address the client's situation. The assessor encourages the referring parties to discuss their dilemmas directly with the client, and to ask for experiences that might weigh on one side or the other. Inevitably, staff members discover that clients are more capable of entering these discussions than was assumed. In addition, explaining and discussing issues with clients, without professional jargon, clarifies staff notions while revealing the clients' understandings and concerns. This process is constructive in itself.

When a formal assessment is arranged, whether on referral or from a direct request from the individual, I continue the above procedure. I ask the client for his or her understanding of the referral, and I clarify whatever I know about it. Later, after we have developed ways of working together, I may ask for the above clarifications again. At that point I am sometimes told that a psychiatrist or therapist wants to know whether the client is really crazy, suicidal, or hopeless. Almost needless to say, it is imperative that I have previously obtained the referring party's permission to discuss the presented issues with the client, so that I can be honest, reassuring, and exploratory, as called for.

I ask what the client would like to learn from the assessment, and I ask what the client already understands about his or her, and others', concerns. As we talk, I try to use his or her terms in order to keep us both grounded in the client's life rather than in abstractions. I help the client to "deconstruct" any personality constructs, and to provide examples of the "when-nots" as well as the "whens" of problematic reactions and actions. For example, a woman tells me that she shares her therapist's concern that she might be pushing beyond her "natural ability," that she might not have "enough intelligence" to pursue a promotion to a managerial position. We discussed what the managerial job requirements would be, and agreed that we would look into her styles of approaching problems and would relate our observations to the managerial position.

Before we began testing, we had discovered that Fran has gotten along well with co-workers and subordinates, and has been regarded as efficient and conscientious. But she has backed away from projects that seemed to require "jumping in and taking the ball" into "unknown territory." A when-not, however, had occurred when she decided to have an addition built on her home. She had written numerous notes and questions in a binder, consulted several contractors, then consulted friends who had built additions, and checked in regularly with the contractor and subcontractors. We agreed that with the freedom to consult she had indeed taken the ball into previously unknown territory.

Fran offered another issue for assessment, namely whether she was a "dependent personality." We de"construct"ed that notion and wound up discussing times that she had and had not deferred to others' judgments when in fact she could have acted on her own. From there we went on to the Minnesota Multiphasic Personality Inventory (MMPI) and Rorschach to see what else we might find in regard to her questions. I reminded Fran that I would summarize our findings after our next meeting, and then would send a copy of my written report to both her and her therapist. The beginning of the assessment process has already been collaborative, individualized, and oriented toward the life world.

1.15.8.2 Using Tests

Any assessor should be knowledgeable in the realms of test selection, profiles, and norms, and in their relations to theories of, and research on, personality patterns, psychopathology diagnostic systems, styles of coping and thinking, and levels of effectiveness. This expertise allows the assessor to form impressions, be surprised, and form alternative understandings. To be most useful, however, the assessor must also know how to use this expertise to gain access to a client's life. Within a human-science approach, our primary data are life events. As above, after locating referral issues in life situations, we then turn to assessment tools, to theories, and to research to gain access to and to explore those life events. If written assessments end instead with scores, categories, and diagnoses, I would say that they are unfinished, that they have reported our tools rather than findings pertinent to an individual's life. The life world should be our point of return from testing, as well as our point of departure into testing.

Test scores do not reveal anything more basic or explanatory than we could find by following a person around through various contexts and observing his or her reactions and actions. Of course the latter "option" is not efficient or usually even feasible. At any rate, the assessor's explorations with the client culminate in a description of how the client has contributed to his or her successes and failings, and how the client might recognize alternative routes to revised goals. Our understandings weave a story, a narrative account, in contrast to assessments that explain in terms of "underlying variables."

Examples of using test patterns in discussion, and then of using individual tests interventionally, follow. I had been asked to conduct an assessment of an electrical engineer whose MMPI-2 had been too defensive to allow immediate clearance into restricted areas of a nuclear plant. He and I agreed that he had tried to look as unproblematic as possible in order to obtain clearance, but also that he has always seen himself as "a good citizen," that he does not look very deeply into personal matters, and that a retest would probably result again in what would look defensive to others. I remarked that although he had not flagged in filling out several pages of background forms I had given him, he struck me as being tired. He said no, that it had just been a long day. "Well," I asked, "how about this possibility?: Although you aren't depressed [he nods], this asterisk here [I point to the subtle depression score on a computer printout] often happens when a person has been kind of down, not as alert or enthusiastic as usual, even though people at work and at home haven't noticed." Mr. Kern looked up at me with surprise, nodded quizzically, but said he did not know much more about it. I then mentioned that on the Rorschach he had been attuned to a lot of things "just happening" [inanimate m], like a leaf that was being blown in the wind, and a lit candle melting. I reported to him that when people are able to see that kind of "just happening," that their lives are often in a holding pattern, that they know where they would like to be going, but it is as though they just have to tread water. To this, Mr. Kern nodded decisively, and when I suggested that most people know what that holding pattern is all about, and that they also find themselves tossing in their sleep about it, he explained to me that his entire work unit had been waiting for eight months to hear whether the corporation was going to close them out of existence. He volunteered that during the same time he and his wife were waiting to see whether they would be able send their son to the college he preferred. Mr. Kern explained then that perhaps he was feeling "worn down" but that he was not "tired," which meant to him beginning to lose hope or not being able to pay attention. Although Mr. Kern had not wanted to examine his current life at all, he found that he had done so in a way that affirmed his previously nonconscious sense of his circumstances, and that unburdened him a bit.

Individual tests and techniques allow both client and assessor to observe the person going about an activity. Similar past situations become available to the client in ways that talk alone would not have afforded. For example, Ms. Rose had wanted to know if we might discover how it was that her boss gave her low performance ratings for efficiency, although she gets more done than anyone else in the office by the end of the day. As I watched her copy the Bender designs, I wondered if my reaction was similar to that of her boss. Ms. Rose looked at each design for a good while, stretched, then held her pencil lightly with one hand and anchored the paper with just two fingers of the other. She drew parallel lines for the square, then half the circle, then finished the box, then the circle. She continued in her unorthodox manner through the nine designs. She never looked at me, and never looked back at the stimulus card after first observing it. When she had finished the designs, I was astounded to see that they were geometrically precise and neat, and that she had finished in considerably less than the average time. We laughed together as she realized that she similarly has created her own world in the busy office, where others cannot see what she is up to as she does things out of order, stretches, seems oblivious to the clock, and so on; others had not noticed the nicely completed projects among the part projects arranged on her desk. Through our work, Ms. Rose was already finding that she could make sense of her circumstance; she had discovered how she had been contributing to others' perceptions of her. We then role-played doing the Bender again with her making comments that allowed me to know what she was doing and that she was indeed following directions. She remarked later that she was now "liberated" both to be herself and to win the approval of her boss.

1.15.8.3 Collaborating and Intervening

Finn has pointed out that this approach to assessment can be undertaken as a therapeutic activity in its own right. Note the therapeutic process in the following examples of intervening into the client's usual course. On the Rorschach, Harriett had hesitatingly reported a fight between two bears who were both injured. On the next card, instead of the usual humans cooperating on ordinary tasks, she reported "warriors from another planet; they make me laugh—they're too stupid to know that they're hermaphrodites." I pointed out a similar sequence with other cards, and asked her what she already knew about a pattern of running from being afraid or from being in touch with the danger of being injured, and then trying to hide through being silly, voicing put-downs, and so on. Harriett readily provided examples within her family and her personal relationships. We used life instances and the Rorschach instances to help her to recognize landmarks which could indicate that she had entered frightening territory. She could then look for turn-off points in order to change course if she found herself veering into silliness or offensiveness. One landmark she identified was an inclination to convert her fearfully tightened lips into a sneer. We agreed that she might then purposively transform the coming sneer into an expression of determination to stay her course while checking out whether she was indeed in dangerous territory.

On other occasions an assessor might intervene during a client's telling of a Thematic Apperception Test (TAT) story. For example: "Geez, John, is this going to end dismally like all the rest? Can you tell it so the fellow figures out a plan?" This was John's first recognition that he had been co-authoring his own dreary circumstances outside the assessment.

Note that this collaborative, interventional approach affirms the client as an active agent, sometimes self-defeating, but also capable of changing course. It focuses on actual situations, whether during assessment or the client's larger life, rather than on drives, traits, cognitions, determinants, and so on. The client is seen as always in relation to goals, to pasts, to obstacles, to invitations. We characterize the person in terms of his or her construing situations and of being formed through choices in those situations. Although not illustrated in this chapter, assessor and client often agree to disagree, or agree that they cannot spell out what it is they are in touch with. In that process, they become at least peripherally aware of both the necessity for, and the inadequacy of, language. Shared humanity and understandings between the participants move the assessment along, despite personal differences. The client knows he or she has been met by someone attempting to deal authentically with limits and possibilities, even when the assessor comes to conclusions not in the client's favor, such as findings of neurological impairment, or of unsuitability for unsupervised visits with a child.

1.15.8.4 Writing Reports

In human science assessment reports, "findings" are both clarifications of past life events and identification of personally viable options for the client. So reports are crafted to remind readers of the client's agency—his or her power to initiate shifts in course. Constructs are eschewed as implying that external factors account for the person's actions. Any necessary jargon is explained in everyday terms. When psychologists cannot write their findings in terms of everyday events, they literally do not know "what in the world" they are talking about. Reports are written in first or third person (Mr. Jones, I, we) rather than referring to "the examiner" or to "this patient." We write with verbs rather than nouns whenever possible, in order to move from abstractions to actual behavior. For example, instead of "John's low self-esteem keeps him from entering the contest," we might write, "Fearing that he would be made fun of, John has not entered the contest." Yes, it does take extra effort to locate and contextualize the actual instances of what has been abstracted into nouns, and it usually takes more writing to describe them than to name constructs. But the worthwhile outcome is understandings of actual events shared by all participants (client, family, teachers therapists, etc.). We also write in the past tense rather than in the present tense, to indicate situatedness and the possibility of change. For example, instead of "Janine sees authority figures as demeaning," we might say, "Janine has seen her bosses and teachers as likely to demean her." We try to own our perceptions rather than attributing all conclusions to test data. For example: "As Ms Gertz glared at me when I asked personal questions, I thought of the MMPI's references to 'nondisclosive' and 'hostile,' which I see as a defensive stance on her part."

Some impressions are only implied or evoked rather than spelled out, not reflecting deficient understanding so much as respect for the ambiguity inherent to the life world. A report that is faithfully descriptive does not attempt false clarity. For example: "Something about the contrast between Mr. Hersh's stiff upper body and his causally stretched out legs struck me as being as disjunctive as was his account of how he had been arrested."

We may add a technical appendix to reports, with commentary on assorted test features, for psychologist readers. This information is regarded as a tool, data for further reflection, not as findings. Again, "findings" are clarifications of life events. We explain any diagnoses contained in the report, in the client's language. We often provide the appropriate pages of the *DSM* (*Diagnostic and statistical manual of mental disorders*) for the client to read and then discuss.

The report is regarded by all as a progress report, a report of progress to date in understanding the client's situations, and in developing concrete suggestions. After addressing any referring party's concerns, the report ends with a section itemizing suggestions that client and assessor developed. For example, "(2) When you've arrived at a solution instantly, remember to slow down to show others how you got there, or how they might get there. You'll recall that we talked about this in regard to my not seeing your "battlefield" on the Rorschach until you guided me through it. We then related that to times your co-workers and supervisor have thought that you were just shooting from the hip, until you helped them to find their way to your insight."

We include a section for the client to write commentary about the report. Often we find that the client has developed further understandings since our last meeting (usually two meetings occur prior to reading the written report). Even when the client does not want to write much on the report, the invitation affirms the client as a collaborator in the process. Knowing that the client is going to read the formal report also encourages the author to find expressions that are true to the client's life.

1.15.8.5 Judging Validity

In that the report deals with direct observations and reports of actual events, we are not faced with the task of justifying inferences. Instead, the criterion is that client and readers must recognize in the report's descriptions touch points with their interpretations or experiences of this client or similar clients, and therein a consensual validity is formed. Ambiguity of the life world is respected when certain findings are implied or evoked rather than spelled out. Differences in perspective are respected as inevitable, and it is understood that further observation and reflection will result in continuing clarifications and revisions. No portrait, description, or other interpretation can be final. In the meantime, the practical utility of the concrete, individualized, suggestions is a strong form of validity.

1.15.8.6 Comments

Working within a human science frame does not bypass the critical contributions of neuropsychology, medicine, cognitive psychology, object relations theory, diagnostic systems, and so on. We take these contributions most seriously but do not adopt the reductive philosophies within which many of them were developed and/or frequently are practiced. For example, although it is definitely helpful to know that a person matches criteria for borderline personality disorder, that match does not explain the person. Although various knowledge domains do provide information about a person's present restrictions and possibilities, we still must explore that person's ways of going about life, so we can locate when-nots, help that person to develop personally viable pivot points, and so on.

Even when intending to work within a human science orientation, we find ourselves falling into our culture's and our discipline's reductive attitude all too easily. Hence, in my assessment courses I include numerous exercises to remind us of the primacy of the life world, and that test data are always about a life that we are responsible for getting back to. On the blackboard I draw the first Bender design, but with the circle and square overlapping instead of adjacent. The class offers ideas about how an overlapping rendition might occur. The ensuing range reminds us not to rush to interpretation from manuals, but rather to be open to a range of possibilities: being in a rush, attending to just getting through the task rather than trying to be precise, resenting having to go through an assessment, leaning on others for comfort or support or to avoid the possibility of not connecting, being unmindful of boundaries. In another exercise, Rorschach students write descriptions of their own situations in "Rorschachese": "There I was, M^aing toward Fd, when a large H AG'd into the line in front of me; I found myself Hxing (H) images, and imagining him MOR." To further ground Rorschach ratios in life, we act out introversive and extratensive styles of entering the classroom late. In still another exercise, we observe slides of the American painter Andrew Wyeth's representational art, and talk about how we too re-present our subjects. We do so without "artist's license" to change features, but with awareness that we too are creating a portrait, from our perspectives and with our lives as points of access.

1.15.9 CONCLUDING REMARKS

Through emulation of the natural sciences, psychology rightly has achieved status as a rigorous discipline. From that secure place, however, we now note that the experimental method, when based on a residual positivistic philosophy of science, is not fully adequate for studying human affairs. An approach that studies humans in terms of their being subjects (agents) as well as objects is human science psychology. This approach respects data formed through natural science methods, but looks more broadly to lived-world contexts to reflect on their significance.

As fewer psychologists identify with radical behaviorism or a strict logical positivism, and more hope for psychology to do greater justice to *human* phenomena, psychologists increasingly are adopting the term "human science" to characterize their own leanings. Simultaneously, qualitative research methods are finding their way from other disciplines into psychology. Through our national organization, psychology advocates respect for cultural and individual differences, and urges that we become familiar with what it is like to be in other people's situations. In these and many other ways we find ourselves increasingly heeding persons' unitary relations with their worlds. I have written this chapter as an invitation to revisit the human science literature, which affords promising possibilities of unifying our discipline. Although human science psychology has been outside the mainstream, its history is as long as that of prevailing psychology, and its grounding philosophy is at least as formidable.

Studies exploring the mutuality of biology and psychology, integrative research methods, and transdiscipline cooperation are all still in their beginning stages. However, their emergence is now accelerating through the work of diverse communities of scholars and practitioners. The second edition of *Comprehensive Clinical Psychology* will probably include in its foundations volume the growing literatures of hermeneutics, linguistics, feminist studies, constructionism, and other endeavors that acknowledge human participation in what we take to be reality. Indeed, perhaps the second edition will explicitly summarize work exploring how it is that our clinical practices necessarily draw on our life worlds. These are exciting, promising times.

1.15.10 REFERENCES

Aanstoos, C. M. (1987). A comparative survey of human science psychologies. *Methods, 1*, 1–36.

Aanstoos, C. M. (Ed.) (1990). Psychology and postmodernity. *The Humanistic Psychologist, 18*(1), 1–132.

Barnard, S. (1998). Socializing psycholinguistic discourse: Language as praxis in Lacan. In K. Malone & S. Friedlander (Eds.), *Looking at Lacan: A Lacanian reader for psychologists*. Albany, NY: State University of New York Press.

Berger, P., & Luckmann, T. (1967). *The social construction of reality*. London: Penguin.

Boss, M. (1949). *Meaning and content of sexual perversions*. New York: Grune & Stratton.

Boss, M. (1958). *The analysis of dreams*. New York: Philosophical Library.

Boss, M. (1963). *Psychoanalysis and daseinsanalysis* (L. B. Lefebre, Trans.). New York: Basic Books.

Buber, M. (1958). *I and thou*. New York: Scribner's.

Bugental, J. F. T. (1963). Humanistic psychology: A new break-through. *American Psychologist, 18*, 563–567.

Bugental, J. F. T. (1965). *The search for authenticity: An existential–analytic approach to psychotherpy*. New York: Holt, Rinehart, & Winston.

Bullock, A. (1985). *The humanist tradition in the West*. New York: Norton.

Bynum, E. B. (1994). Brief overview of transpersonal psychology. In F. J. Wertz (Ed.), *The humanistic movement: Recovering the person in psychology*, pp. 186–191. Lake Worth, FL: Gardner Press.

Colaizzi, P. (1973). *Reflection and research in psychology: A phenomenological study of learning*. Dubuque, IA: Kendall Hunt.

Colaizzi, P. (1978). Psychological research as the phenomenologist views it. In R. Valle & M. King (Eds.), *Existential–phenomenological alternatives for psychology*. New York: Oxford University Press.

Coufal, D. (1997). *Being impatient: An existential phenomenological study*. Unpublished doctoral dissertation. Pittsburgh: Duquesne University.

deCarvalho, R. J. (1991a). Gordon Allport and humanistic psychology. *Journal of Humanistic Psychology, 31*, 8–13.

deCarvalho, R. J. (1991b). *The founders of humanistic psychology*. New York: Praeger.

deCarvalho, R. J. (1994). The institutionalization of humanistic psychology. In F. J. Wertz (Ed.), *The humanistic movement: Recovering the person in psychology*. Lake Worth, FL: Gardner Press.

Dilthey, W. (1977). *Descriptive psychology and historical understanding* (R. Zaner & K. Heigas Trans.). The Hague; The Netherlands Nijhoff. (Original work published 1894).

Edie, J. M. (1987). *Edmund Husserl's phenomenology: A critical commentary*. Bloomington: Indiana University Press.

Embree, L., Behne, E. A., Carr, D., Evan, J. C., Huertas-Jourda, J., Kockelmans, J. J., McKenna, W., Mickunas, A., Mohanty, J. N., Seebohm, T. M., & Zaner, R. M. (Eds.) (1997), *Encyclopedia of phenomenology*. Boston: Kluwer.

Epting, F. R. (1984). *Personal construct counseling and psychotherapy*. New York: Wiley.

Finn, S. E. (1996). *A manual for using the MMPI-2 as a therapeutic intervention*. Minneapolis, MN: University of Minnesota Press.

Fischer, C. T. (1971). Toward the structure of privacy: Implications for psychological assessment. In A. Giorgi, W. F. Fischer, & R. von Eckartsberg (Eds.), *Duquesne studies in phenomenological psychology* (Vol. 1). Pittsburg, PA: Duquesne University Press.

Fischer, C. T. (1991). Phenomenological-existential psychotherapy. In M. Hersen, A. E. Kazdin, & A. S. Bellack (Eds.). *The clinical psychology handbook* (2nd ed. pp. 534–550). New York: Pergamon.

Fischer, C. T. (1994). *Individualizing psychological assessment*. Hillsdale, NJ: Erlbaum. (Original work published 1985.)

Fischer, C. T. & Wertz, F. J. (1979). Phenomenological analyses of the experience of being criminally victimized. In A, Giorgi, R. Knowles, & D. Smith (Eds.). *Duquesne studies in phenomenological psychology* (Vol. 3 pp. 135–158). Pittsburgh: Duquesne University Press.

Fischer, C. T. (1998). Being angry revealed as self-deceptive protest: An empirical phenomenological analysis. In R. Valle (Ed.), *Phenomenological inquiry in psychology: Existential and transpersonal dimensions* (pp. 111–122). New York: Plenum.

Frankl, V. (1946/1962). *Man's search for meaning: From death camp to existentialism* (I. Lasch, Trans.) Boston: Beacon Press. (Original work published 1946.)

Frankl, V. (1986). *The doctor and the soul: From psychotherapy to logotherapy* (2nd ed.) (R. & C. Winston, Trans.). New York: Vintage Books. (Original work published 1946.)

Gavey, N. (1997). Feminist poststructuralism and discourse analysis. In M. Gergen & S. Davis (Eds.), *Toward a new psychology of gender*. New York: Routledge.

Gergen, K. J. (1985). The social constructionist movement in modern psychology. *American Psychologist, 40*, 266–275.

Gergen, K. J. (1994). *Realities and relationships*. Cambridge, MA: Harvard University Press.

Giorgi, A. (1970). *Psychology as a human science: A phenomenologically based approach*. New York: Harper & Row.

Giorgi, A. (Ed.) (1985). *Phenomenology and psychological research*. Pittsburgh: Duquesne University Press.

Giorgi, A., Fischer, W. F., & von Eckartsberg, R. (Eds.) (1971). *Duquesne studies in phenomenological psychology*. Pittsburgh: Duquesne University Press.

Giorgi, A., Fischer, C. T., & Murray, E. L. (Eds.) (1975). *Duquesne studies in phenomenological psychology (Vol. 2)*. Pittsburgh: Duquesne University Press.

Giorgi, A., Knowles, R., & Smith, D. L. (Eds.). (1979). *Duquesne studies in phenomenological psychology (Vol. 3)*. Pittsburgh: Duquesne University Press.

Giorgi, A., Barton, T., & Maes, C. (Eds.). (1983). *Duquesne studies in phenomenological psychology (Vol. 4)*. Pittsburgh: Duquesne University Press.

Greening, T. (1985). The origins of the *Journal of Humanistic Psychology* and the Association of Humanistic Psychology. *Journal of Humanistic Psychology, 25*, 7–11.

Harding, S. (1987). Introduction: Is there a feminist method? In S. Arding (Ed.), *Feminism and methodology: social sciences issues* (pp. 1–13). Bloomington, IN: Indiana University Press.

Hawkesworth, M. (1989). Knowers, knowing, known: Feminist theory and claims of truth. *Signs: Journal of Women in Culture and Society, 14*, 533–557.

Heidegger, M. (1962). *Being and time* (J. Macquarrie & E. Robinson, Trans.). London: SMC Press. (Original work published 1927.)

Hodges, H. A. (Ed.) (1944). *Wilhelm Dilthey: An introduction*. London: Routledge.

Hoy, D. C. (1978). *The critical circle: Literature, history, and philosophical hermeneutics*. Berkeley: University of California Press.

Husserl, E. (1962). *Ideas: General introduction to pure phenomenology* (W. R. Boyce Gibson, Trans.). New York: Collier—MacMillan. (Original work published 1913.)

Husserl, E. (1970). *The crisis of European sciences and transcendental phenomenology* (D. Carr, Trans.). Evanston, IL: Northwestern University Press. (Original work published 1935–37).

Kashgarian, G. T. (1997). *Narcissitic injury: A phenomenological–linguistic study of the experience of feeling criticized*. Unpublished doctoral dissertation. Pittsburgh: Duquesne University.

Kelley, G. A. (1955). *The psychology of personal constructs*. New York: Routledge.

Kockelmans, J. J. (1985). *Heidegger and science*. Washington, DC: University Press of America.

Leder, D. (1990). *The absent body*. Chicago: University of Chicago Press.

Leitner, L. M. (1985). The terrors of cognition: On the experiential validity of personal construct theory. In D. Bannister (Ed.), *Issues and approaches in personal construct theory* (pp. 83–103). London: Academic Press.

Maslow, A. H. (1962). *Toward a psychology of being*. Princeton, NJ: Van Nostrand.

Maslow, A. H. (1966). *I-Thou knowledge in a psychology of science*. New York: Harper & Row.

May, R., Angel, E., & Ellenberger, H. F. (Eds.) (1958). *Existence: A new dimension in psychiatry and psychology*. New York: Basic Books.

Merleau-Ponty, M. (1962) *The phenomenology of perception* (C. Smith, Trans.). New York: Humanities Press. (Original work published 1945.)

Merleau-Ponty, M. (1963). *The structure of behavior*. (A. L. Fisher, Trans.). Boston: Beacon Press. (Original work published 1942.)

Moss, D. (1978). Medard Boss. In R. S. Valle & M. King (Eds.), *Existential–Phenomenological alternatives for psychology* (pp. 308–322). New York: Oxford University Press.

Natanson, M. (1966). *Essays in phenomenology*. The Hague, The Netherlands: Martinus Nijhoff.

Needleman, J. (Ed.) (1963). *Being-in-the-world: Selected papers of Ludwig Binswanger*. New York: Basic Books.

Packer, M. J., & Addison, R. B. (Eds.) (1989). *Entering the circle: Hermeneutic investigation in psychology*. Albany NY: State University of New York Press.

Packer, M. J., & Scott, B. (1992). The hermeneutic investigation of peer relations. In T. Winegar & J. Valsinar (Eds.), *Children's development within social contexts: Research and methodology* (Vol. 2, pp. 75–111). Hillsdale: NJ: Erlbaum.

Richer, P. (1978). A phenomenological analysis of the perception of geometric illusions. *Journal of Phenomenological Psychology, 8*, 123–135.

Rogers, C. R. (1951). *Client-centered counseling: Its current practice, implications, and theory*. Boston: Houghton-Mifflin.

Rogers, C. R. (1961). *On becoming a person: A therapist's view of psychotherapy*. Boston: Houghton-Mifflin.

Royce, J., & Mos, L. (Eds.) (1981). *Humanistic psychology: Concepts and criticisms*. New York: Plenum.

Sartre, J. P. (1948). *The emotions: Outline of a theory* (B. Frechtman, Trans.). New York: Philosophical Library. (Original work published 1939.)

Sartre, J. P. (1956). *Being and nothingness: An essay on phenomenological ontology* (H. Barnes, Trans.). New York: Philosophical Library. (Original work published 1943.)

Slife, B. D. (1993). *Time and psychological explanation*. Albany, NY: State University of New York Press.

Slife, B. D., & Williams, R. N. (1997). Toward a theoretical psychology: Should a subdiscipline be formally recognized? *American Psychologist, 52*, 117–129.

Smith, M. B. (1990). Humanistic psychology. *Journal of Humanistic Psychology*, *30*, 6–21.

Spiegelberg, H. (1965). *The phenomenological movement: A historical introduction* (2nd ed., Vols. 1 and 2). The Hague, The Netherlands: Nijhoff.

Spiegelberg, H. (1972) *Phenomenology in psychology and psychiatry: A historical introduction*. Evanston, IL: Northwestern University Press.

Sutich, A. (1962). American Association of Humanistic Psychology: Articles of the Association. *Journal of Humanistic Psychology, 2*, 96–97.

Valle, R. S. (Ed.) (1997). *Phenomenological inquiry in psychology: Existential and transpersonal dimensions*. New York: Plenum.

Valle, R. S., & Halling, S. (Eds.) (1989). *Existential–phenomenological perspectives in psychology: Exploring the breadth of human experience*. New York: Plenum.

Valle, R. S., & King, M. (Eds.) (1978). *Existential–phenomenological alternatives for psychology*. New York: Oxford University Press.

Van Kaam, A. (1958). *The experience of feeling really understood by a person*. Unpublished doctoral dissertation, Western Reserve University.

Van Kaam, A. (1966). *Existential foundations of psychology*. Pittsburgh, PA: Duquesne University Press.

Von Eckartsberg, R. (1986). *Life–world experience: Existential–phenomenological research approaches in psychology*. Washington, DC: Center for Advanced Research in Phenomenology & University Press of America.

Walsh, R. (1995). The study of values in psychotherapy: A critique and call for an alternative method. *Psychotherapy Research, 5*, 313–326.

Welch, I. D., Tate, G., & Richards, F. (Eds.) (1978). *Humanistic psychology: A source book*. Buffalo: Prometheus Books.

Wertz, F. (1983). From everyday to psychological description. *Journal of Phenomenological Psychology, 14*, 97–241.

Wertz, F. J. (Ed.) (1994). *The humanistic movement: Recovering the person in psychology*. Lake Worth, FL: Gardner Press.

Yalom, I. (1980). *Existential psychotherapy*. New York: Basic Books.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

1.16
Family Systems and Family Psychology

JAY L. LEBOW
Chicago Center for Family Health and University of Chicago, IL, USA
and
ALAN S. GURMAN
University of Wisconsin Medical School, Madison, WI, USA

1.16.1 INTRODUCTION 474

1.16.2 DEFINING THE PURVIEW OF FAMILY THERAPY 474

1.16.3 THE VARIETY OF FAMILY THERAPIES 475

1.16.4 A BRIEF HISTORY OF COUPLE, FAMILY, AND SYSTEMS THERAPIES 475

1.16.5 CORE SYSTEMIC CONCEPTS 477

1.16.6 MODELS OF FAMILY THERAPY 478
- *1.16.6.1 Structural Family Therapy* 479
- *1.16.6.2 Strategic Approaches* 479
  - *1.16.6.2.1 The Mental Research Institute* 480
  - *1.16.6.2.2 Haley's problem-solving therapy* 480
  - *1.16.6.2.3 Milan systemic therapy* 481
  - *1.16.6.2.4 Solution-focused therapy* 482
- *1.16.6.3 Cognitive-behavioral Approaches* 482
- *1.16.6.4 Psychoeducation* 484
- *1.16.6.5 Bowen Therapy and Other Intergenerational Approaches* 485
- *1.16.6.6 Psychodynamic Approaches* 486
- *1.16.6.7 Experiential Approaches* 487
- *1.16.6.8 Narrative Approaches* 488
- *1.16.6.9 Integrative Approaches* 488

1.16.7 THE EFFICACY OF COUPLE AND FAMILY THERAPY 489

1.16.8 TRENDS IN COUPLE AND FAMILY THERAPY 490

1.16.9 THE DEVELOPMENT OF FAMILY PSYCHOLOGY 491

1.16.10 SUMMARY 492

1.16.11 REFERENCES 492

1.16.1 INTRODUCTION

The place of family systems theories and couples and family therapies within psychology has vastly expanded in recent years. Once seen as a radical departure from the more traditional focus on the individual that has typified the mental health disciplines, family systems viewpoints have now been with us for over 40 years and gained wide acceptance. Many systemic theories and therapies have been developed, and couple and family therapies are now among the most widely practiced. Several prominent guild organizations supporting the practice of family therapy have blossomed, including The American Association for Marriage and Family Therapy, the American Family Therapy Academy, and the Division of Family Psychology of the American Psychological Association, as have a number of prominent journals including *Family Process*, *Journal of Marital and Family Therapy*, and *Family Therapy Networker*. The Division of Family Psychology within the American Psychological Association now has over 6000 members, while the American Association for Marriage and Family Therapy has over 25 000 members.

The essence of what has driven all this attention lies in the emergence of a broad recognition of the importance of the family in the life of the individuals within it and the society made up of families. Whether we consider the impact of a depressed parent on a child, the role of poor parenting practices as a risk factor for conduct disorder in children, or the impact spouses have on one another, both clinical experience and much research point to the enormous influence of the family. Gurin, Veroff, and Feld (1960) found that 42% of all people who had sought professional help for psychological problems viewed their problems as related to a marital problem, and another 17% viewed their problems as pertaining to family relationships.

However, family therapy is truly unified only in the shared belief that relationships are of at least as much importance in the behavior and experience of people as are internal processes within individuals or broader social forces. As Gurman, Kniskern, and Pinsof (1986) have suggested, family therapy includes therapists from many professions, has no unified theory, and few techniques that are specific to it. There are many distinct systemic therapies, which differ enormously from one another. Some are directed to the treatment of families, some to subsystems within the family (e.g., couples), and others abandon the specific focus on the family entirely, aiming at the broader social nexus. In this chapter, we overview basic systems concepts and the field of family therapy, highlighting both the common threads underlying these methods and the differences that have emerged across the schools of practice.

1.16.2 DEFINING THE PURVIEW OF FAMILY THERAPY

Delineating what is meant by family therapy is far more complex than might be thought. The simplest way of labeling a therapy as a "family" therapy is to look at who is seen in treatment. From this vantage point, family therapy occurs when more than one member of a family are seen together in psychotherapy. This definition has considerable advantage in its parsimony. Accepting this definition, it is clear when family therapy is occurring and when it is not, an appraisal that can be made through a simple head count. From this viewpoint, the meeting of a couple, a parent and son, or a multigenerational family with a therapist, all constitute family therapy, whereas meetings of therapists with one individual or with unrelated individuals in groups do not. In most such usage, the term "family therapy" is restricted to instances when family members are seen together at one time (also more specifically termed conjoint family therapy), although at times the definition has been expanded to include therapies in which various members of the family are seen in different sessions, termed concurrent treatment (e.g., when a child is seen in some meetings and parents in another).

Despite the attractiveness of such a simple straightforward definition, most family therapists have not found this way of conceptualizing family therapy either satisfying or sufficient. Instead, most family therapists view the essence of family therapy as centered on maintaining a focus on the system rather than who is seen in treatment. They point to therapies in which multiple members of families are seen but in which the focus remains on an individual (e.g., as when the focus is exclusively on changing the symptomatic client) as failing to meet the criterion for family therapy, and to other therapies that directly involve only one client but focus on the social system (e.g., Bowen Therapy) as family therapies. It has been the presence or absence of a systemic focus of treatment that has emerged as most important to the majority of those who practice family therapy. Based on such a notion, Gurman et al. (1986) offered the following classic definition of family therapy: "Family therapy may be defined as any psychotherapeutic endeavor that explicitly focuses on altering the interactions between or among family members and seeks to improve the functioning of the family as a unit, or its subsystems, and/or the functioning of individual members of the family." This definition best fits with the zeitgeist of family therapy, but leaves some room for debate about whether a particular treatment is "family therapy."

A related issue in demarcating the territory of family therapy concerns the place of couple therapy, specifically whether couple therapy represents a subset of family therapy or a separate endeavor. Some have argued that because couples therapy calls for a distinct set of skills, it should be considered separately (Alexander, Holtzworth-Monroe, & Jameson, 1994). More typically, couples therapy has been regarded as a subset of family therapy (Lebow & Gurman, 1995). Most prominently, those who emphasize the importance of a systemic viewpoint as the essence of family therapy view the treatment of couples as simply work with one family subsystem.

It also should be mentioned that in the late 1990s "marital and family therapy" has become "couple and family therapy." Respect for the diverse forms of family in our society (Walsh, 1993) has led to this fundamental change in nomenclature.

1.16.3 THE VARIETY OF FAMILY THERAPIES

Family therapy is actually a number of different activities linked by a few common understandings. Systemic concepts merely set a frame for possibilities. They do not limit the range of intervention. Both the mediating and ultimate goals of treatment may vary, as may the theoretical frame, the treatment strategies, and the specific interventions employed (Gurman, 1978). Therefore, it makes conceptual sense to subdivide family therapies into ones that share common characteristics.

Family therapies can be classified best along two axes: the first consisting of who is seen in treatment (e.g., individual, couple, nuclear family, extended family), and the second, the theoretical perspective on which the therapy centers (e.g., structural, strategic, object relations, or a type of integration). Both who is seen and the theory of the approach contribute to the differences between therapies at the level of operations.

It is also crucial in examining this literature to keep in focus the importance of treatment goals (Gurman, 1978). In couple and family therapy, change in family process is always both a mediating and an ultimate goal. Most couple and family therapy also has an additional goal, be it a change in the behavior of an individual or in a broader dimension of relational life, for example, marital satisfaction. In examining family therapy, we must maintain a focus on the impact on both family process and other treatment goals.

1.16.4 A BRIEF HISTORY OF COUPLE, FAMILY, AND SYSTEMS THERAPIES

To understand the enormous variation among family therapies, it is important to grasp the diverse origins of the field. The earliest couple and family therapy was framed as a direct extension of existent models of individual therapy. Although much of this work was conjoint, the essence of these methods consisted of helping to uncover individual patterns to be altered. In these therapies, some of the core techniques for working with more than one client in the room were developed (e.g., early variants of communication training), but these therapies were largely considered adjunctive to what was viewed as the more important work of individual therapy.

Family therapy leapt to prominence through the work in the 1950s and 1960s of such figures as Nathan Ackerman, John Bell, Ivan Boszormenyi-Nagy, Murray Bowen, James Framo, Jay Haley, Donald Jackson, Salvador Minuchin, Virginia Satir, Carl Whitaker, and Lyman Wynne, who shared a common belief in the core importance of the family system as well as a great deal of personal charisma. In contrast to earlier couple and family therapies, this work actively questioned and argued against the traditional individual oriented view of problem development and treatment. The early work of these pioneers included a wide range of family-focused intervention, crossing the boundaries of schools of family therapy that have subsequently emerged. Each developed theories and methods of intervention, and several ultimately shaped their methods into the foundation of a school of treatment. Each incorporated aspects of systems theory into their work, most prominently the view that causality is best conceived of as a circular process in which behavior is seen as interdependent and subject to mutual influence. Within the zeitgeist of the time, the behavior of "identified" patients (i.e., those with symptoms who were labeled as having the problems) were seen as a reflection of underlying family process, that is, the family was viewed as the principal locus of the problem, central in its development, and the most appropriate context for treatment. The first generation of family therapists emphasized systemic concepts with the kind of fervor that goes with those who have discovered an as yet undiscovered truth, and were highly critical of traditional methods of mental health intervention in which individuals were seen out of their natural context in the social system.

The influences on this generation of pioneers themselves were enormously diverse. Some had backgrounds in psychoanalysis (e.g., Ackerman, Framo) that they brought to bear. Some had backgrounds in fields other than the mental health professions, such as anthropology (e.g., Bateson, Weakland) and communication (e.g., Haley). In general, new and exciting ideas were welcome and sought out. Thus, for example, Haley became quite interested in the work of Milton Erikson, bringing core techniques of his variant of hypnotherapy, emphasizing paradoxical intervention, into the mainstream of methods of practice in family therapy. Observation of the importance of difficulties in communication in troubled families, based on the work of the double bind project (Bateson, Jackson, Haley, & Weakland, 1954) and the parallel work of Wynne, Ryckoff, Day, and Hirsch (1958) and Lidz, Cornelison, Fleck, and Terry (1957), also exerted an enormous impact, focusing emphasis on changing these deviant communication patterns.

Through the 1970s and 1980s, the field of family therapy passed from the shared excitement about the core importance of the family system to emphasizing differences across the many distinct schools that emerged. Some of these schools accepted or transmuted a range of concepts from individual therapy, creating schools that are psychoanalytic (Ackerman, 1958; Scharff & Scharff, 1987), experiential (Greenberg & Johnson, 1988; Whitaker & Keith, 1981), and behavioral (Patterson, 1982). Other schools rejected virtually all aspects of individual models of treatment, remaining exclusively focused on aspects of the social system such as structure (Minuchin, 1974) overcoming family homeostasis (Watzlawick, Beavin, & Fisch, 1974), or intergenerational process (Bowen, 1978). Across these schools, a vision developed of a powerful therapist (sometimes literally referred to as a wizard), jousting or performing some version verbal judo to free up the family from its patterns. Through this time, family therapy grew enormously in popularity and began to enter into the mainstream of practice.

In the 1980s, voices began to emerge within family therapy that were highly critical of aspects of the schools of practice that had developed. Most prominently, the feminist critique pointed to the numerous male assumptions that were endemic to most models in family therapy (e.g., that fathers should hold the major executive position in the family). Feminists called for a more egalitarian family therapy (Goldner, 1985; Hare-Mustin & Marecek, 1988). Other criticisms focused on the basically homeostatic vision of family life in family therapy, in which families were seen as highly resistant to change and readily relapsing into dysfunctional patterns. Alternative visions developed emphasizing family resilience and the natural tendency in families toward change (Walsh, 1982, 1993).

Other questions were raised about whether the therapist needed to be the powerful enactor of change depicted in many of the earlier models of treatment. Theoretical precepts of second-order cybernetics were advanced (Hoffman, 1981) that emphasized the view that there were no "objective" observers (Von Glaserfeld, 1984), and that therapists, like all others, become part of the system, being influenced by it as well as influencing it. Following Gergen (1981), social constructivism further argued that knowing and knowledge are socially constructed through language and discourse. Models of family therapy developed that emphasized collaboration (Anderson & Goolishian, 1988) and the personal construction of narrative (White & Epston, 1990).

Still others became critical of the exclusive focus on the family as the locus of change in systemic models. Movement has clearly been away from the earlier simplistic notion that the family was the sole etiologic agent in the development of difficulties, and that family therapy was the sole preferred method of intervention across all difficulties. An integrative viewpoint is emerging that includes not only the concepts from various family methods of intervention, but also interventions at the level of the individual (Lebow, 1984, 1987a) and larger system (Breunlin, Schwartz, & Karrer, 1992). Some have even called for a basic redefinition of systemic therapy, moving from a specific focus on the family to a broader vision of consultation with social systems, including but not limited to families (Wynne, McDaniel, & Weber, 1986).

Family therapy has also been examined through the lens of several other pertinent and powerful vantage points including culture (Boyd-Franklin, 1989; McGoldrick, Pierce, & Giordano, 1988), life cycle development (Carter & McGoldrick, 1988), and postmodernism (Hare-Mustin & Maracek, 1988). This has been a field boiling over with ideas and concepts, and ways of examining the family and how to have an impact on it. It has been a continually developing field, in which treatments have emerged and been refined, theory has undergone considerable revision, and assumptions have been continually examined within the emerging vantage points about family within the broader society.

1.16.5 CORE SYSTEMIC CONCEPTS

A few core assumptions underlie all family therapy, the most central being the core importance assigned to the social system in influencing individual behavior. Families are seen as having a powerful effect in shaping and maintaining individual patterns of behavior. There have been a range of ways of conceptualizing this influence represented in the various models of family therapy.

One set of core concepts that have been prominent within a wide range of models derives from general systems theory, developed by Von Bertalanffy (1969) as a way of understanding all systems, animate and inanimate. The central tenet of general systems theory is that the whole is more than the sum of its parts. Therefore, to understand any part (e.g., an individual), one must grasp its relation to the whole of which it is a part (e.g., the family).

Within general systems theory, humans are viewed as part of what is termed an "open" system, one in which there is ongoing exchange with those outside, be they other individuals, families, or other systems. Open systems remain subject to influences from outside. Systems (e.g., families) are also made up of subsystems (e.g., a couple, children), which affect one another, and sum to more than their parts. General systems theory focuses on the most global principles of how systems evolve, be they particles in space or human families.

A principle that assumes great importance in general systems theory is equifinality. Equifinality suggests that there are many ways of reaching particular configurations within the system and that the particular pathway by which a configuration has been reached does not matter. End states reached through different pathways are equivalent. Applied to family systems, an emphasis on equifinality focuses concern on the state the family is presently in and not how the family reached that state. To the extent equifinality is stressed, history and individual motivation assumes lesser importance, while the topography of how the system is presently organized assumes more.

In general systems theory, how behavior is understood is viewed as a function of the context in which it is conceived. In a classic example cited by Watzlawick, Beavin, and Jackson (1967), the meaning of seeing a man quacking at ducks is significantly altered by the knowledge that this man is Konrad Lorenz, engaged in experiments about imprinting. Without that piece of information, the behavior looks eccentric or psychotic, but sense can be made of this behavior in the appropriate context. Early in the history of family therapy, a focus on context became the cornerstone of the belief that the behavior of all family members would make sense if only the meaning of the behavior in the appropriate context could be deciphered. In particular, severe mental illness was seen as the product of behavior that made sense in a particular context (e.g, within the family process), although it appeared to make little sense when seen outside of that context.

A common extension of these concepts early in the history of family therapy was to label the behavior of the family members displaying psychopathology or other problematic behavior as "identified patients." These "identified patients" were typically seen as the victims of labeling of their behavior when considered out of the appropriate family context. Within this viewpoint, the real patient and bearer of the problem is the family, and therefore, family therapy is the most appropriate method of bringing about change. A frequently encountered correlate of this set of beliefs was dismissal of the biological basis for severe mental disorder (e.g., Haley, 1963).

General systems theory also emphasizes circular paths of causality. Rather than focusing on linear pathways of actions followed by reaction, attention centers on recursive patterns of mutual interaction and influence. If the behavior of one person affects that of another (e.g., a father punishes his child), attention still must focus on how the response of the second person affects the first (e.g., the child's aggressive behavior leads to the parent's punishment, the parent's punishment leads to the child temporarily stopping the behavior, followed by further aggressive behavior). From the perspective of general systems theory, the system, not a single person, is responsible for the behavior that is maintained through such circular pathways.

Some of the critiques and reappraisals of systems theory have been directed at these concepts. One line of criticism has centered on the notion of "identified patients." Much research has shown very real disabilities in those with mental illness. Many now question the use of this term when applied to situations in which there is psychopathology. In particular, the psychoeducational movement has called attention to some of the costs of this approach in alienating families from the treatment they need, because of the sense of shared family causation that accompanies usage of the term "identified patient."

Others have argued the limits of circular notions of causality (Dell, 1986), pointing to the dangers inherent in the argument that all parties are equally responsible for sequences of behavior. In particular, the example of family violence has been cited as an instance where individual responsibility and lineal arcs of causality need to be highlighted, lest the observer be left with an inappropriate sense that batterer and victim are coequal in responsibility for violent behavior (Dell, 1986, Goldner, Penn, Sheinberg, & Walter, 1990). There is considerable agreement among family therapists that both lineal and circular pathways of causality and problem maintenance need to be considered in assessing family systems.

Cybernetics, the science of communication and control in man and machine, developed by Wiener (1961) and others, added to the understandings developed through general systems theory among family therapists. In cybernetics, systems are viewed as self-correcting, influenced in an ongoing way by feedback. Feedback is the process by which a system gains information to self-correct to maintain a steady state or move toward a goal. Positive feedback describes input that increases deviations from the steady state, negative feedback describes input that reduces such deviations. Homeostasis is the powerful force moving the system toward a steady state. Early family therapy was profoundly influenced by the idea that human systems were homeostatic, that is, moving toward the reduction of change. As a result, most first-generation family therapies were based in the notion that powerful interventions needed to be created to reorganize the family, overcoming the homeostatic forces.

Systemic thinking has given more weight to morphogenesis, the natural force moving the system toward change. A systemic emphasis on morphogenesis creates very different implications for psychotherapy than a homeostatic emphasis. Models emphasizing morphogenesis suggest that initiating the process of change is likely to kick off a positive chain reaction, very unlike the minimization of change thought to be active in a perspective emphasizing homeostasis. Thus, in morphogenetic models, launching small changes is seen as likely to be productive in creating movement within the system, while in homeostatic models, only the most powerful of interventions are seen as likely to produce change. In a similar vein, Maturana and Verela (1980) highlight the process they termed autopoesis, whereby the internal structure of the living system determines its behavior. Following the notion of autopoesis, the therapist can only perturb the system to make changes that the client system itself produces.

Another important emphasis in systemic theory has been on communication processes. The point of entry for these ideas in family therapy was the double bind theory of schizophrenia developed by Bateson et al. (1954), which suggested that the psychotic process was the product of disturbed communication. In the double bind, two or more parties are involved in an important relationship that is ongoing. A primary injunction is given, such as "show me your feelings." A second injunction is given that conflicts with the first, such as "The feelings you have are unacceptable and should not be verbalized." Given that the recipient of the communication cannot leave the field, he or she is in a double bind that induces anxiety. The double bind theory suggested that the repeated exposure to such binds results in responses that break out of the bind through psychotic processes.

Although empirical research never was able to demonstrate high frequencies of such double binding on the part of parents of schizophrenics (indeed, in the research, the presence of double bind communication could not even be reliably rated as present by observers), this work has remained highly influential in family therapy in focusing interest on deviant communication processes in families. Beginning with the premise that "you can't not communicate" (Watzlawick et al., 1967), this work sought to understand the communication which was occurring, and alter patterns of communication so that the system could use this communication to reorganize in a more functional manner.

1.16.6 MODELS OF FAMILY THERAPY

Many specific couple and family therapies have been developed. Some of these models have been directed to specific difficulties (e.g., Kaplan's treatment for sexual disorders), but the majority have been directed toward a broad range of problems. Some family therapies are aimed at resolving difficulties that are explicitly labeled as being about family relationships (e.g., couple therapy aimed at marital dissatisfaction; family therapy aimed at overcoming differences between parents and their grown children), whereas other family therapies utilize a family systems perspective for intervention with problems that manifest themselves in the behavior of an individual, and which from another perspective might be seen as "individual" problems (e.g., depression). The numerous couple and family therapies can be divided into a few distinct categories on the basis of emphasis. These are structural, strategic, cognitive-behavioral, psychoeducational, intergenerational, psychodynamic, experiential, narrative, and integrative schools of family therapy.

1.16.6.1 Structural Family Therapy

Structural family therapy, developed by Salvador Minuchin, is an example of a family therapy largely drawn purely from systemic concepts. Structural family therapy emphasizes the power of the social system, as manifested through family structure. Health and dysfunction are directly viewed as products of the effectiveness of the family structure. By family structure, Minuchin means the regulating codes as manifested in the operational patterns through which people relate to one another in order to carry out functions. The three primary dimensions of structure are boundary, alliance, and power.

Boundaries are the rules defining who participates and how, who is in and who is out of an operation, regulating contact. The strength of boundaries vary, ranging from rigid, resulting in disengagement, to very permeable, resulting in enmeshment. At the disengaged end of the spectrum, families act like they have little to do with each other, leaving the individuals substantially disconnected. At the enmeshed end, there are violations of function boundaries, in which family members intrude into functions that are the domain of other family members. Structural family therapy aims to move families away from the extremes of enmeshment and disengagement.

Alliances are the joining or opposition of one member of a system to another in carrying out an operation. Alignments are inevitable. They become dysfunctional when they become fixed and unchanging (stable coalitions) or when they are primarily cross-generation. Triangulation describes the process of two people demanding that a third join with them against the other. Structural family therapy aims to create alliances that are functional (e.g., parents with one another), while at the same time not becoming rigid.

Power describes the relative influence of each family member on the outcome of an activity. Power can be functionally distributed with the primary locus in the older generation or can become rigidly held by one individual or coalition, or there can be insufficient executive authority (weak executive function). Power is seen as best held in the hands of an executive (e.g., a parental coalition), but in such a fashion to leave everyone with some degree of power.

Treatment in structural family therapy primarily consists of efforts to change these elements of family structure. Given the homeostatic vision of systems at the base of structural therapy, therapists seek to create powerful in-session experiences to work to alter the family's organization. Frequently, a transaction is created by the therapist that promotes in-session family members" habitual patterns of relating (called enactment). There is much emphasis on joining with the family through such interventions as tracking (adopting the symbols of the family's life), accommodation (relating to the family in congruence with the family's patterns), and mimesis (joining with the family by becoming like the family in manner or content). Ultimately, interventions are directed to fully restructuring the system. Symptomatic change in "identified" patients is assigned a much less important role than change in structure of the system.

The structural approach remains the most influential school within family therapy. Some of its ideas, such as the importance of boundary, alliance, and power in family systems have come to be broadly accepted by most family therapists. However, there has also been considerable criticism of some of the assumptions of structural therapy. Most emphatically, the highly gender-based nature of some of the assumptions about the roles of men and women and boys and girls within structural family therapy has resulted in considerable criticism, particularly from feminists. The emphasis of this approach on equifinality and, thus limited concern with history or the internal process of individuals, has also been attacked. Minuchin (1996) has recently moved to a more gender aware version of the structural approach, and acknowledged a greater appreciation for the importance of history.

1.16.6.2 Strategic Approaches

Strategic approaches are the most purely systemic of the family therapies. These models simply ask, "What is the most expeditious route to promote systemic change?" Strategic approaches also share an orientation toward brief focused intervention. From a strategic viewpoint, change is a discontinuous process. The goal is to intervene, find a new way of functioning that works better, and promptly end the treatment.

Strategic models have also been closely identified with several specific methods. One common thread is the use of "paradoxical" interventions, in which directives are offered which if acted on would move the family in the opposite direction from that which is desired. Although direct interventions are also common in strategic methods, the rapier-like effort to find the simplest and most expedient pathway to change is best represented by the use of paradox. Strategic methods have also been closely associated with the use of team approaches, utilizing observers behind one-way mirrors as part of the intervention process, typically offering commentary or directives to the therapist and family. Strategic methods have been further associated with cool detached stances on the part of the therapist (e.g., the MRI model), although there have been significant exceptions to this trend. Strategic therapies center on altering feedback cycles within the family, but do not seek to enable insight within the family about such cycles. Change, not learning about the change process, is clearly the center of attention. The strategies of change invoked vary significantly with the specific strategic model.

1.16.6.2.1 The Mental Research Institute

The first systems-based strategic model within family therapy was the "Mental Research Institute" (MRI) or "Palo Alto" model developed by Jackson, Watzlawick, and Weakland, and their colleagues (actually the first MRI model, as there have been subsequent other MRI models). The MRI model derived from a mix of systems theory, cybernetics, and the study of communication processes.

In the original MRI model, problems are viewed as a natural part of family life that families regularly deal with on their own. The need for intervention is seen as stemming not from the problem itself, but from how family members treat the problem. When families become stuck in systemic patterns and in efforts to solve problems ("more of the same"), changes in behavior, termed "first-order change," are seen as unlikely to resolve the problem, and possibly making the problem even worse. Therefore, therapy focuses on the creation of "second-order change," an alteration in the rules of the system that govern interactions. Treatment consists of identifying the ways problems are maintained by the behavior within the system, examining the rules that lie beneath these behaviors, and then changing the rules. Reframing and paradoxical interventions are viewed as the most powerful tools for initiating second-order change. Reframing consists of active efforts by the therapist to create a new and different understanding of old events that has a more benign meaning. For example, in recasting behavior that has been seen as bad as being out of the control of the individual, a different reality is created. Reframing has become perhaps the most common intervention in couple and family therapy, employed by therapists regardless of their orientation.

Paradoxical directives capitalize on the force in social systems to move against the direction toward which it is directed. For example, the therapist may list reasons why change is not likely to be productive or even harmful. Although much more controversial than reframing, paradox has become common within family therapies, especially strategic ones.

Treatment within the MRI model always remains brief and focused. The therapist's cool detached stance is not designed to make for long-term attachment, and termination is encouraged by the therapist as soon as problem resolution has been substantially initiated.

The MRI model stands as the exemplar of a strategic family therapy, and has been highly influential among all other strategic therapies. Several of its core concepts (reframing, "more of the same," first- and second-order change) have been adopted within a broad range of family approaches. Nonetheless, although highly influential, this approach today is rarely seen in practice in its pure form. The paradoxical bent of the intervention strategy and the detached stance of the therapist have caused this approach to receive a great deal of attention, but have also led many who have tried this approach to move on to other models that emphasize a more collaborative approach between clients and therapist. In addition, little empirical support is available for this approach.

1.16.6.2.2 Haley's problem-solving therapy

Haley's (1976, 1980) problem-solving therapy and the closely related work of his colleague Madanes (1981) combines a strategic use of paradoxical techniques with goals that typify structural family therapy. Problem-solving therapy strongly emphasizes grasping and working with the function that behaviors serve within the system. Most often, this function is conceptualized as a struggle for power and control.

An initial stage of the therapist and client becoming acquainted is followed by a stage in which each person's perspective about the problems which exist is elicited. Specific observation by the therapist focuses on triangles (who supports whom in interaction) and hierarchy (who has what power), but these ideas are not directly shared with the clients. Directives to the family follow from a consideration of solutions that have been attempted, and aim to engage the family in new and different behaviors. Many of the techniques utilized by Haley and Madanes derive from the hypnotic work of Milton Erickson, aimed at increasing suggestibility and openness to change. One example is the pretend technique

(Madanes, 1981), in which the family is directed to have children pretend to have symptoms and parents pretend to help them, a paradoxical technique that suggests the possibility of overt control over patterns thought to be out of conscious control. Other commonly employed techniques more directly aim at establishing a coalition between parents to help adult children leave home.

Haley remains highly controversial in his adherence to some of the earliest systemic conceptualizations of family therapists, for example, the view that "identified patients" carry symptoms entirely due to the function these symptoms serve within the family and a denial of the existence of mental illness. For Haley, psychopathology is always the product of a dysfunctional social system, not due to biology or individual psychology. These ideas, once a welcome contrast to the determinism of biological and psychoanalytic formulation, now appear rigid and stale in the wake of the development of the considerable literature delineating the biological and psychological basis for severe disorder, and the emergence of highly effective psychoeducational treatments that demonstrate how an approach can be family based and yet consistent with the best data about schizophrenia and other mental illness.

1.16.6.2.3 Milan systemic therapy

A number of strategic therapeutic approaches have been developed in Milan, Italy, by Selvini-Palazzoli, Boscolo, Cecchin, Prata and their colleagues in various combinations. Versions of these models have varied enormously, although all have maintained a strategic focus.

In the classic Milan therapy that brought worldwide recognition to this group (Selvini-Palazzoli, Boscolo, Cecchin, & Prata, 1978), sessions are held approximately once per month, almost always involving a team situated behind a one-way mirror. The team forms a hypothesis about the family, to be modified over the course of treatment. During a break in each session, the team formulates an intervention. The therapist then brings the intervention to the family, most often through positive connotation or the prescription of a ritual.

Positive connotation consists of reframing behavior in a positive light, most frequently through suggesting how the behavior serves the goals of the system. Positive connotation aims to change the family view of dysfunctional behavior, while also decreasing resistance by allowing each family member to emerge with a positive view of their own behavior.

The rituals prescribed in Milan systemic therapy move to exaggerate or move against rigid patterns in the family. Most of the rituals have an ironic quality and engender confusion, although some (e.g., one called "odd and even days," in which control is given to each parent on alternating days) merely serve to call attention to patterns in the family, and thereby move the family to see their ability to impact on the situation and resolve their difficulties. The early Milan approach strongly stressed the importance of therapist neutrality in delivering these interventions.

In the most influential variant of the Milan methods, Boscolo and Cecchin (Cecchin, 1987) moved away from directives toward the use of what they termed "circular questions." Circular questions are questions used to learn about differences in the family that might provide clues to recursive family patterns. Circular questions include ones about differences in the perception of relationships (who is closer?), differences between before and after something else happened (were you more depressed before or after the birth of the baby?), and hypothetical differences (if you had not married, how would your life be different?). Curiosity is the essential ingredient in circular questioning. The aim is not to move the family toward a specific goal, but to initiate thought and conversation in order to create greater understanding of how the present situation and the family's behavior in it came about, what the systemic patterns are that help keep the family from resolving their difficulties, and what are the most productive pathways toward change. Work in this model becomes much more collaborative than in the earlier version of Milan therapy.

Selvini-Palazzoli (1986) added another variant of the Milan model. Selvini-Palazzoli came to believe that disturbed patients were inevitably caught up in what she termed the "dirty game," a power struggle between parents in which patient's symptoms help support one parent. Her response was what she called "the invariant prescription," applied to all families. In the invariant prescription, the therapist suggests to parents that they tell family members that they have a secret and go out together mysteriously without warning other family members, and that they then observe the family's reaction. The "invariant prescription" aims to help strengthen the alliance between the parents and enable understanding of dysfunctional patterns in the family. Although rooted in some of the same observations about triangulation in disturbed families as those made by Haley, Bowen, and Minuchin, this approach has acquired little support because of its highly pathological view of family processes, its ignoring of the mounting evidence demonstrating the importance of expressed emotion to recidivism, and its failure to respond to differences among families.

More generally, the Milan approaches have proven to be highly influential. Although only a small number of family therapists practice any of the variants of Milan therapy, the attitude of curiosity and prompting of circular questions have come to serve as the base of investigation for many family therapists. There has been very little empirical testing of the Milan approaches.

1.16.6.2.4 Solution-focused therapy

Emerging into prominence since the mid-1980s have been a number of approaches that take positive connotation and reframing a step further, attempting to move discussion fully to thinking in terms of solutions rather than problems. Among the best known of this solution-focused set of approaches is the work of DeShazer (1985, 1988), Berg (1993), and O'Hanlon and Weiner-Davis (1989). The solution-focused approaches assume that clients want to change and reject the notion of deeply ingrained pathology. Instead, these approaches seek to introduce ways of thinking about and facing difficulties that are different, and can kindle the family's own process of resolving their difficulties. One favorite technique is to look for exceptions, times when problems have not been present or overcome. Another has been to nurture and help clients notice small changes from which they can build larger ones. For example, De Shazer asks clients to observe what happens in their lives that they want to continue. DeShazer and colleagues also employ the "miracle question": "Suppose one night, while you were asleep there was a miracle and this problem was solved. How would you know? What would be different?" All these techniques are designed to help clients begin to think in terms of solutions and the ability to resolve difficulties rather than in terms of problems and one's difficulty in resolving them.

Solution-focused approaches have been among the most widely influential in the 1990s, especially in the context of the mandate of managed care for brief therapy. The positive focus and optimistic frame of these models has proven most welcome to families and therapists alike. Criticism has focused on the repetitive use of the same few interventions (e.g., the miracle question), and on vary simplistic notion of problem development and resolution implicit in the model. Ultimately, whether such parsimony and positive ideas are sufficient is an empirical question. Unfortunately, given their promise, solution-focused approaches remain among the least studied of the family therapies.

1.16.6.3 Cognitive-behavioral Approaches

Cognitive-behavioral models extend behavioral principles to the treatment of family systems. These models have primarily been utilized in work with child behavior problems (especially conduct disorder and delinquency) and with difficulties encountered by couples (especially marital dissatisfaction). Cognitive-behavioral methods begin with the assumption that thoughts and behavior are central to all aspects of functioning and that the most efficacious pathways to change directly address dysfunctional thoughts and behavioral patterns.

Classical and operant conditioning are the central mechanisms for shaping behavior within a behavioral paradigm. In behavioral family therapy, operant conditioning has assumed particularly great importance.

Central to operant conditioning is reinforcement: presenting some event or behavior that increases the rate of a particular response. Humans are seen as inevitably affected by the reinforcements they receive. However, behavioral family therapy is not fully the product of classical learning theory, but instead its application in the social context, called social learning theory. Here, learning is not simply the product of primary reinforcers such as food, but of social reinforcers, such as approval. Social learning also occurs both directly from experiences that reinforce or punish, and indirectly through processes such as modeling, in which learning occurs through observation of contingencies. This array of processes shape social behavior.

Social exchange theory (Thibaut & Kelley, 1959) also has had a prominent place in cognitive-behavioral approaches. Social exchange theory suggests that individuals strive to maximize their outcomes, to increase the rewards they receive, and decrease the costs. Behavior from one person is viewed as likely to be met with reciprocity to behavior from another, so that positive behavior will lead to positive behavior, and punishment to punishment on the part of the other. In particular, couples are regarded as likely to develop social exchanges that can become mutually supportive (each emitting positives to the other) or coersive (each emitting punishing behaviors). Problem behavior is viewed as primarily the product of either skill deficits that stem from a lack of knowledge, or from the establishment of coersive exchange. Skill training is aimed at providing the knowledge and experience needed to engage in appropriate social behaviors, be they as a spouse or a parent. Positive exchange is altered directly by helping clients become more aware of patterns of exchange, and by negotiation of a more satisfying *quid pro quo*.

The cognitive theories that make up the cognitive part of cognitive-behavioral emphasize the development and maintenance of dysfunctional or "irrational" thought processes, and direct efforts to alter these cognitions through learning in therapy. Cognitive interventions examine the ideas that lie behind behavior and emotion for the presence of core distortions. The emphasis lies in being able to understand the importance of the thought that lies between an experience and the resultant feeling. Cognitive interventions principally help clients to understand and alter the tendency to overgeneralize, personalize, or be overly negative about events that are occurring. Homework is essential in tracking and assessing beliefs, just as it is essential in accomplishing behavioral goals.

Early in the history of behavioral couple and family therapy, there was little that was systemic in the approach. Systems theory was not part of the theoretical base. Instead, the emphasis was on ways of promoting individual learning to change family patterns. It was not unusual for behavioral family therapists to meet exclusively with parents to train them in better parent practices to shape the behavior of children. Indeed, much of this work was called behavioral parent training rather than family therapy.

More recent work by cognitive-behavioral therapists has been much more systemic. For example, Patterson and Chamberlain (1992) clearly explicate the reciprocal influence of child and parent in conduct disorder, and emphasize how therapists if they do not remain sensitive to client needs can engage in too frequent teaching behaviors that promote noncompliance with therapeutic tasks. Alexander and Parsons (1982) have added the systemic notion of function to behavior analysis. In their functional family therapy, attention centers on identifying the function of behavior, much in the manner of the Haley or MRI models. Once identified, behavioral interventions are introduced to accomplish this function in a less damaging way.

Cognitive-behavioral therapies are more similar to one another than other groupings of family therapies, such as strategic or intergenerational. It is a strength of these models that the work of each investigator builds on that of others. Even if models have different names and slightly different components, they are essentially similar in how to treat specific problems. Couples therapies are like one another, as are treatments for children and adolescent problems.

Weiss (1978), Jacobson and Margolin (1979), and Stuart (1980) have all articulated similar approaches to couples therapy. Each approach begins with a behavioral assessment that includes the use of instruments to assess general levels of relationship satisfaction such as the dyadic adjustment scale. However, the primary focus of the assessment is on delineating problematic exchanges, specific target behaviors, and themes in the relationship that require change, evaluated through client recording of these behaviors between sessions, therapist observation of typical interactions, and clients completing self-report forms. The results of the assessment are directly shared with the couple, highlighting the areas in their relationship that require attention, leading to the development of a blueprint for change. Much of couple dissatisfaction is viewed as the product of the low level of positive reinforcement and high level of coersive exchange within the relationship, an often replicated finding in maritally distressed couples.

Many interventions are brought to bear to change these interactions. The monitoring of behavior, through tracking and sometimes including the use of videotape feedback, help couples to objectify their behavior, to see it from the perspective of an outsider. Where specific skills are lacking, skills training is employed to develop competencies. Typical skills training includes the development of communication skills such as attending, reflecting, listening, and speaking, and of problem-solving skills such as the abilities to define problems, generate alternative solutions, and reach naturally satisfying outcomes. Behavior exchange is specifically addressed through the development of contracts between the parties about these exchanges, most based in a *quid pro quo*, in which the behavior of one party is directly exchanged for the behavior of the other.

Behavioral couples treatments have a particularly strong record of demonstrating success in empirical studies, at least in short-term effectiveness (Lebow & Gurman, 1995). The great conundrum for behavioral couples therapy lies in addressing the aspects of the relationship that are not simply about behavior, those that involve feeling states, particularly love and caring. Behavioral couple therapists have therefore stretched the model to accommodate the obvious importance of this aspect of the relationship. In early formulations, some form of noncontingent loving behavior was prescribed, called "caring days" or "love days." Jacobson and Christenson (1996) have emphasized the importance of developing accepting behaviors in addition to other skills. Other criticisms of behavioral couples approaches have more recently emphasized the lack of attention to history, meaning, and internal dynamics in the approach (Gurman & Knudson, 1978). The addition of cognitive theory and interventions have mitigated some of this criticism, but not all.

In sex therapy, a range of specific techniques for dealing with sexual problems are added to couple therapy. At times, interventions other than behavioral ones are included, as in Kaplan's (1979) widely circulated integrative model, but sex therapy always retains what primarily is a behavioral core. Much of this core, developed by Masters and Johnson (1970), derives from the well-demonstrated insight that anxiety is antithetical to sexual response and, that through classical conditioning, relaxation can replace anxiety. Sex therapy invariably includes the use of what are termed "sensate focus" techniques to induce relaxation. Other specific techniques are added which are specifically tailored to each sexual dysfunction. Sex therapy numbers among the most effective therapies in outcome studies, although the rates of success reported by LoPiccolo and LoPiccolo in 1978 are considerably lower than those reported by Masters and Johnson.

Much of the treatment of child problems in behavioral family therapy has exclusively focused on the parents in behavioral parent training. Given the theoretical orientation emphasizing reinforcement as crucial in behavior, and the large body of data available suggesting that the parents of problematic children help shape their dysfunctional behavior and respond poorly to it, many behavior therapists have concluded that time in therapy is best spent with the parents who control the reinforcers rather than with the children, especially when children are small. As in behavioral couple therapy, parent training begins with an assessment phase in which patterns of thought and behavior are recorded and connected to the target behavior of concern. This leads to a functional analysis of the problematic behavior, from which a plan is formed specifying the skills that need to be mastered and changes in contingencies that need to occur for the problem to be eradicated. Focus may center on caring behaviors as well as on establishing control. If the problematic behavior on the part of the child is restricted to a single area of concern, specific contingencies may be created in response to that behavior. For example, a program may be constructed of reward for school work. When problems are encountered in a number of areas, more comprehensive contingency programs are developed. Home token economies and point systems provide ways for credit to accrue for positive behavior and debits for problematic behaviors, with rewards dispensed for overall performance. In all programs, the preference for positive reward over punishment in shaping behavior is emphasized.

Although behavioral parenting programs have been very successful in helping parents to deal with a wide range of problems, they have not been sufficient to deal with more difficult problems. Here, the developers of these treatments have refocused their attention on the need to find ways to overcome resistance. Patterson and his colleagues (Patterson & Chamberlain, 1992) have highlighted situations in which therapist directives merely increase resistance, and sought ways to intervene that respond to client reactance. In a similar vein, Alexander and Parsons (1982) have directed functional family therapy for adolescent delinquent behavior toward ways of examining function and using that knowledge to enable engagement and cooperation in treatment. Further, in the work of groups led by Patterson, Alexander, and Henggeler, the behavior of the child or adolescent and the impact of peer groups and other relevant systems is accorded equal footing with the intervention with parents. These models assume a systemic perspective and resemble typical family therapy far more than parent training.

Behavioral family approaches to child and adolescent problems are among the most researched and validated of psychotherapies. Although the early variants of these approaches could readily be criticized for their simplistic analysis and failure to consider a range of important variables (e.g., history and meaning), recent work has been auspicious, leaving these among the most sophisticated as well as validated psychotherapies.

1.16.6.4 Psychoeducation

Psychoeducational approaches to the treatment of disorders are based on the notion that syndromes that seriously impair functioning have a biopsychosocial basis, and that illness models of these disorders are to be taken seriously. Sometimes illness models are fully incorporated as part of the approach, while in other variants they are examined with families as possible explanations for the disorder. Flowing directly from these ideas is the corollary that affected individuals and their families can gain a great deal from developing their understanding of the relevant problem and acquiring the requisite skills that best fit with that particular problem. The goal is to establish a collaborative partnership with families who suffer from mental disorders, providing them with the kind of information and skills most needed. Information is viewed as a major route to coping. Beyond this constant, psychoeducational treatments include an eclectic mix of interventions derived from individual and family therapy that have particular relevance to the particular syndrome, as well as psychopharmacological interventions.

Psychoeducational family treatments were developed in the context of schizophrenia, where the treatments have been most extensively refined and tested, and extended to other disorders, most notably manic-depressive disorder. In the context of schizophrenia, a group at Western Psychiatric Institute in Pittsburgh (Anderson, Reiss, & Hogerty, 1986) and another at UCLA (Falloon, Boyd, & McGill, 1984) developed related, although somewhat different, psychoeducational methods. Each featured medication for the person with schizophrenia, along with education for the family. Each approach centered substantially around the notion that expressed emotion in families must be reduced. These approaches followed the extensive studies by Leff and Vaughn (1985), among others, demonstrating that when schizophrenics are exposed to much expressed emotion in families, consisting of criticism and high negative affect, high rates of recidivism and symptomology ensue.

Anderson et al.'s unique contribution lies in what they termed "survival skills" workshops, which over a full day, present the current state of knowledge about schizophrenia to families. These workshops seek to impart information, increase the sense of social support, and reverse the negative interaction families of disturbed individuals often have with mental health providers. Families are regarded as full collaborators and taught in these workshops both what is known and what is speculative about schizophrenia. The Anderson et al. (1986) family model also accented work in therapy designed to alter dysfunctional aspects of family structure, and a minimalist approach to intervention in sessions that included the schizophrenic, with one constant goal being to keep expressed emotion to a minimum.

The methods of Falloon, Boyd, and McGill (1984) also seek to involve family and reduce expressed emotion, but place greater emphasis on behavioral skills training. There also is a greater emphasis in this model on crisis management when the inevitable crises develop in the lives of these families. Both the UCLA and the Pittsburgh groups reported remarkable levels of outcome in sophisticated clinical trials, which have been followed up with replication. With both methods, expressed emotion has been decreased in these families, compliance with procedures increased, and recidivism and symptoms decreased in the index patient. As such, this research suggests these treatments offer the most consistent cost-effective treatment for schizophrenia. This work has been followed up with similar procedures for families dealing with manic depression (Miklowitz & Goldstein, 1990).

One striking aspect of these models has been the inclusion of family in treatment in a way that has proved highly acceptable to these families, in contrast to earlier methods of dealing with these families that left many families blamed and highly dissatisfied. Much of this earlier work questioned the very existence of mental illness or even of disturbed internal processes in the schizophrenic. Some (e.g., Haley, 1976) even suggested that medication for the schizophrenic is harmful since it further establishes the patient in the sick role, and obscures what were regarded as the inevitable systemic issues. The strong evidence for the efficacy of psychoeducation, coupled with the lack of evidence that family therapy works in these samples without the use of medication, suggests that the psychoeducational form of family therapy is clearly superior to the earlier variety. Indeed, it appears clear that the highly stimulating family therapies of old frequently seen in the tapes of many masters of family therapy in treating schizophrenia provide exactly what is not needed: a highly stimulating environment likely to be difficult for the patient and an environment in which families are likely to feel blamed for the problem. All told, for schizophrenia and other severe mental disorders, psychoeducational treatments number among the most successful family therapies, and are becoming widely disseminated as the standard for care.

1.16.6.5 Bowen Therapy and Other Intergenerational Approaches

Bowen (1978) developed a prominent form of family therapy, now called Bowen family systems therapy, that incorporates systems theory, along with a unique language for considering the relationship of the individual with family. The crux of the Bowen approach lies in the concept of differentiation of self, which essentially amounts to the ability to distinguish thoughts and feelings. For Bowen, psychological and systemic health is a direct function of level of differentiation. When individuals differentiate themselves from family processes, they are viewed as less susceptible to the pathology-inducing aspects of the system. Differentiation is clearly distinguished from cutoff, the establishment of rigid boundaries that minimize contact with family, which is viewed in this system as innately problematic.

In Bowen's theory, individual development is largely shaped by the family system. Bowen envisioned what he termed an "undifferentiated family ego mass," of beliefs and feelings within families that are transmitted through what he termed a "family projection process" across generations. The position of the individual in the family, in part determined by birth order and in part by other factors, is also viewed as of key importance in shaping the individual. A key element of family processes lies in the creation of triangles, in which the interaction between two individuals are affected by the presence of a third. Triangles are viewed as inevitable but also treacherous for individual development.

In Bowen family systems therapy, each member of the family involved is helped to increase their ability to manage their own anxiety. In individual differentiation lies the basis for better family relationships. Therefore, much of the work focuses on the relationships adult clients have with their families and, most specifically, their families of origin. Family of origin is typically not seen directly in treatment, but examined through forays outside sessions in which the client learns about family processes, experiences them, and finds new ways to cope with them. In this part of the therapy, termed coaching, the therapist helps the client develop a plan for investigation and examine the thoughts and feelings that emerge. Exploration involves both direct contacts with living relatives and efforts to learn about and experience feelings in relation to deceased family. Genograms, diagrams of the multigenerational family systems of participants, are employed to help in this examination, to shape exploration and set goals. Because this is the essential process of treatment, much of Bowen family systems therapy is conducted with only a single client in the office, although the work is principally centered on their family relationships.

Bowen family systems therapy represents a bridge between individual and family therapy. Although couched in systemic terms, many of Bowen's ideas about differentiation resonate with the concepts of object relations and cognitive models of therapy. Although much of the unique terminology developed by Bowen adds little (with a few clear highly salient exceptions such as "coaching"), Bowen developed a method that has enabled an exciting and moving voyage of exploration of family processes by innumerable clients. Although the clinical experience of many family therapists suggests that this is a highly satisfying and effective treatment for clients, there has unfortunately been a paucity of research investigation of this approach.

There have been numerous other approaches that have centered on the kind of examination of intergenerational process developed by Bowen. Most prominent are approaches that create family rituals that serve as cathartic events for negotiating the emotional turmoil resulting from multi-generational legacies (Imber-Black & Roberts, 1993). Others have focused on explicating and adapting the family myths that have evolved.

1.16.6.6 Psychodynamic Approaches

Although psychoanalytic formulations have often served as the foil in expositions of several family therapies against which the value of focused systemic therapies could be highlighted, there has also been a long standing tradition of family therapies that have incorporated psychodynamic concepts (Reiner, 1997). Early in the history of family therapy, Ackerman, Boszormenyi-Nagy, Framo, Sager, Steirlin, and others created treatments that blended systems concepts with specific psychodynamic theories. More recently, the refinement of object relations concepts within psychoanalysis has led to the emergence of several family therapies that take object relations one step further, to consider those dynamics directly in the context of the family in treatment (e.g., D. Scharff & J. S. Scharff, 1987; Slipp, 1993).

Psychodynamic family approaches share a number of common characteristics, despite there being considerable variation with the particular psychodynamic formulation included. These approaches have roots in psychoanalytic concepts, but in all cases there is a significant transformation in the family context.

Most basic to psychodynamic formulations in couple and family therapy is the notion of an active dynamic internal process within individuals. Psychodynamic approaches share the belief that unconscious mental processes are extremely important and that early experience has a crucial influence on later behavior and experience. Psychodynamic family therapy has retained from its individual therapy counterpart an emphasis on the processes that occur within individuals.

Psychodynamic therapists also emphasize maintaining the frame of treatment, the formal arrangements such as frequency, time, and length of sessions. Creating an appropriate frame is viewed as leading to the development of a holding environment (D. Scharff & J. S. Scharff, 1987), in which the therapist tolerates clients" anxieties and tensions while remaining empathic with their emotional experiences. The creation of a holding environment also allows for observations of deviations that arise from the frame. Another important route into unconscious process is the understanding of transference: the client's displacement or projections onto others of feelings, impulses, defenses, and fantasies from important past relationships or conflicts. These projections help recapitulate important aspects of clients" earlier relationships in therapy or in the family relationships. In psychodynamic family therapy, transferences are observed as much in relation to other family members, particularly spouses, as in relation to the therapist.

Psychodynamic couple and family therapies also accent the therapist's awareness of their own feelings in the therapy process. Countertransference, the therapist's reactions to the client based on client transferences or on the therapist's own personal experience, are viewed as important sources of information about client process. Most importantly, in projective identification, the therapist may be induced to feel or behave as the client has in the past or as others have behaved and felt toward the client. In most psychodynamic couple and family therapy, the understanding and owning of projective identification on the part of the client assumes an especially important place in the process.

Therapists offer interpretations that provide meaning to behavior by explicating unconscious processes. Change is seen as the product of working through one's issues over time, often as the product of these interpretations. Understanding resistance, the process, often rooted in anxiety, that moves against therapeutic goals, is also important in enabling change.

In many instances, the language and interventions of psychodynamic family therapy merely extends those of individual psychodynamic therapy and, most especially, object relations theory (e.g. D. Scharff & J. S. Scharff, 1987). Other family therapies begin with similar concepts, but draw on quite different methods, and language. For example, Boszormenyi-Nagy and Spark (1973) examine much of the same territory as an existential psychodynamic psychotherapy, but look at relationships in terms of what they term "invisible loyalties" rather than using traditional language in their contextual therapy. Their work aims at exploring multi-generational processes in families with an eye to what they term "ledgers," the balance of what has been given and received by each individual. The central tenet of the therapy lies in helping clients deal with and balance the ledger they bring from their families of origin. The stance of the therapist toward the family, termed by Boszormenyi-Nagy and Spark "multidirected partiality," is also much like the hovering attention basic in more recent forms of psychoanalytic approaches, but carried over to the family context with new language and intervention strategies.

The work of Framo (1981) also draws heavily on the psychodynamic tradition, but adapted to the context of the family therapy. Framo developed what he termed family-of-origin sessions as part of couple or family therapy, in which the adults in the therapy would meet for a few sessions with members of their own family of origin, in order to understand better and resolve the outstanding issues that derive from that experience.

Although pure-form psychoanalytic therapies are relatively infrequently encountered in couple and family therapy, psychodynamic principles are central in the practice of many family therapists. These concepts have been highly influential, especially in their important place in a variety of integrative therapies. Many of the intergenerational and experiential family therapies also include a number of psychodynamic concepts. Psychodynamic therapies have rarely been evaluated through research. The demonstrated effects of a variant of pychodynamic therapy, insight-oriented couples therapy (Snyder & Wills, 1989), suggests the likelihood of a promising future for these treatments in research, should this research ever be carried out.

1.16.6.7 Experiential Approaches

Prominent experiential couple and family therapies have been developed by Whitaker (Whitaker & Keith, 1981), Satir (1983), and Greenberg and Johnson (1988). Each of these approaches places the emphasis on felt experience, that is, in the therapy restoring liveliness and connection. In each, the primary instrument is the therapist, who uses self as an instrument toward change.

Each experiential family therapy employs different intervention strategies. Whitaker utilized a wide array of techniques, ranging from direct commentary to physically wrestling with clients, all aimed to fight emotional deadness. Satir employed exercises from the human potential movement with which she was so long associated, and Greenberg and Johnson utilize methods of focused interaction between couples derived from Gestalt therapy. Although there are relatively few family therapists trained in these specific methods, this work remains enormously influential in highlighting the importance of the person of the therapist, and the need to maintain liveliness and authenticity in the work of couple and family therapy.

1.16.6.8 Narrative Approaches

At the time of the publication of this volume, narrative concepts have captured the greatest recent attention among family therapies. Michael White (White & Epston, 1990) has emerged as the major figure in the narrative movement. Other prominent figures include Anderson and Goolishian (1988), Combs and Freedman (1990), and Hare-Mustin (Hare-Mustin & Marecek, 1988). Their approaches vary in specifics, but all have roots in social constructivism (Gergen, 1981, 1991), the notion that knowing is socially constructed through language and discourse, and depends on the context of the observer. Each places a strong emphasis on thought processes and beliefs, and in each difficulties are envisioned as the product of stories that have been socially created and can be reconstructed. Problem-oriented descriptions are replaced by stories of accomplishment.

White emphasizes externalizing problems, that is, seeing them as separate entities from the individuals involved. Much like solution-oriented and MRI therapists, White also emphasizes the outcomes that occur when individuals have been successful in overcoming problems. Anderson and Goolishian (1988) offer the ultimate extension of this type of approach, fully replacing the notion of the expert therapist with the idea of therapist and clients as equal partners in conversation.

Rather than merely opening discourse, these approaches also accent the freeing of repressed voices. For White, following Foucoult, the dialogue needs to be as much about overcoming societal oppression as about family processes.

Although sometimes these therapies do involve seeing families together, these therapies move far from the family emphasis of early family therapy. Much of the work is done with individuals, with only passing references to the family system. As yet, we also have little in the way of outcome research testing the effectiveness of these models. However, narrative models have already gained many proponents, and have influenced many more family therapists, in creating a frame for treatment that emphasizes a co-equal collaborative conversational style that de-emphasizes the therapist's role as expert, in emphasizing the importance of the client's voice, and in directly working with client stories about their lives to help create more workable realities.

1.16.6.9 Integrative Approaches

Integrative methods have become commonplace in couple and family therapy. Not only has a considerable literature emerged concerned with integration (Lebow 1984, 1987a, 1987b; Liddle, 1984; Grunebaum, 1988; Moultrop, 1981, 1989), and numerous integrative models have been developed and widely disseminated (e.g. Feldman, 1985, 1990; Gurman, 1981; L'Abate, 1986; Pinsof, 1995), but the move to integration has become so much part of the fabric of our work that it largely goes unrecognized.

Integrative models merge the raw material of the various approaches. This merger occurs at three distinct levels: theory, strategy, and intervention. Because there are numerous therapies to merge, and several levels along which to merge them, integrative models vary enormously in content. Some integrative approaches accent each therapist's building of a personal method, while others offer highly prescriptive delineations of therapeutic ingredients and a specific map for when to do what. Moultrop's work (1981, 1989), describing the therapist's personal odyssey, and Lebow's (1987a), suggesting guidelines for each therapist's model development, offer examples of conceptions focused on the therapist as an agent of integration. Jacobson's (1992) couple therapy and Liddle, Dakof, and Diamond's (1992) treatment for adolescent drug abusers and their families offer examples of highly specific prescribed routes to intervention, which can be specified in treatment manuals. Other models, such as Pinsof's (1983) problem centered therapy or Gurman's (1981) integrative marital therapy, bridge this chasm through prescribing ingredients, but allowing varying levels of room for improvisation, especially for more advanced practitioners.

Most integrative efforts combine behavioral notions of learning, with a systemic understanding of the family process, and the individual psychodynamics that are brought to bear in these patterns (Feldman, 1985, 1990; Gurman, 1981; Kirschner & Kirschner, 1986; Nichols, 1987, 1995; Pinsof, 1983, 1995; Sander, 1979; Wachtel, 1992, 1995; Wachtel & Wachtel, 1985). Pinsof's (1983) "Problem centered therapy" offers a highly refined version of this type of model, in which self-psychology is the internal system. Gurman (1981, 1992) has developed a combination of object relations, behavioral, and systemic procedures for working with couples. Feldman (1990) adds an integration of similar ingredients envisioned as linked in a multilevel circular process.

Much of the creative edge in integration has been concerned with the development of specific treatments for specific populations. Goldner et al. (1990) have merged feminist, narrative, systemic, and psychodynamic concepts in the treatment of abuse within couples. Liddle et al. (1992) and Piercy and Frankel (1989) have brought structural, systems, and behavioral principles together along with a developmental perspective in the treatment of adolescent chemical dependency. Similarly, Kaplan (1974) has brought an integrative approach to sex therapy; Addis and Jacobson (1991) to adult depression, Rolland (1994), Wright and Leahy (1994) and Wood (1994) to families with physical illness; Wachtel (1995) to families with young children; Steinglass (1992) to alcoholism; Alexander and Parsons (1982) to adolescent acting out; Markman (1979) to premarital couples; Trepper and Barrett (1989) to sexual abuse, and Harkaway (1989) to eating disorders. Each of these methods provides a hearty integration, that includes theory, methods of intervention, and particular values that guide treatment.

Feminists (Goldner, 1985; Hare-Mustin & Marecek, 1988, 1990) and those who offer treatment in diverse cultures (Aponte, 1985; Boyd-Franklin, 1989, McGoldrick et al., 1988) have focused attention on the obvious importance of race, class, and gender, and on the value of therapists" shaping treatment in relation to these factors. This has resulted in the development of several integrative treatments, specifically designed for particular cultural groups or gender-related issues (Boyd-Franklin, 1989; Goldner, 1985; Szapocznik & Kurtines, 1989; Watts-Jones, 1992). These models move beyond the notion of one method for all to a better understanding of which methods work best in what combination with various populations. Culture and gender have also been incorporated as anchors in some broader efforts at integration (Breunlin et al., 1992). The feminist and cultural perspectives have also helped elucidate the underlying assumptions about gender and culture within treatment models, leading to a more informed discussion of what is being integrated.

It should be noted here that even though feminist family therapy presumes an integration of feminist ideas and methods of family therapy, feminist work has seldom been labeled as integrative. Nevertheless, feminists have provided some of the richest integrative frameworks; many models include a wide range of precepts and interventions. Similarly, the integrative nature of the models of those working with specific cultural groups often goes overlooked in considerations of integrative treatment models.

As yet, research has only assessed a small number of these integrative treatment models. Nonetheless, the results of the existent research are very promising. The integrative methods studied have proven highly effective at dealing with a wide range of problems.

1.16.7 THE EFFICACY OF COUPLE AND FAMILY THERAPY

Three decades of research have confirmed and reconfirmed the effectiveness of couple and family therapy. Reviews of the literature conclude that the outcomes achieved by treatment groups have exceeded those of control groups (Gurman & Kniskern, 1978, 1981; Gurman et al., 1986; Alexander et al., 1994; Jacobson & Addis, 1993; Baucom & Hoffman, 1986). With more and better research emerging the evidence for overall effectiveness has become unequivocal (e.g., Jacobson, Dobson, Fruzzetti, Schmaling, & Salusky, 1991; Snyder & Wills, 1989). Although the majority of studies have focused on behavioral treatments, there is now also a considerable base of nonbehavioral treatment studies that point to treatment efficacy. Further, the earlier stated concerns about possible confounds of treatment method that might mitigate this conclusion (Beach & O'Leary, 1985; Gurman & Kniskern 1978; Bednar, Burlingame, & Masters, 1988; Lebow, 1981) can now be put to rest. A range of excellent methodological studies confirms the general findings of efficacy found in less rigorous research.

Nonetheless, there is clearly a differential amount of evidence supporting the efficacy of the various couple and family therapies (Lebow & Gurman, 1995). Behavioral approaches largely have extensive bodies of research support, particularly in treating childhood and adolescent conduct disorder, marital dissatisfaction, and adolescent acting out. Structural approaches also have considerable support, particularly in child and adolescent conduct disorder. An experiential approach to couple therapy, emotionally focused therapy (Greenberg, Ford, Alden, & Johnson, 1993), and a psychodynamic approach, insight-oriented (Snyder & Wills, 1989), also have garnered research support. A considerable body of research evidence points to the efficacy of treatments that integrate individual and conjoint treatments, including psychoeducational treatments of schizophrenia (Hogarty, Anderson, Reiss, Knornblith, & Greenwald, 1986), cognitive-behavioral treatment of depression (O'Leary & Beach, 1990), multisystemic treatment (Henggeler, 1993), and functional family therapy for adolescent delinquent behavior (Alexander & Parsons, 1982), and multidimensional family therapy for adolescent drug abuse (Liddle, Dakof, & Diamond, 1992). There exists very little research support for a variety of widely practiced couple and family therapies including Bowen, narrative, humanistic, strategic, and solution-focused approaches. Differences in knowledge about models are even more pronounced than earlier (Gurman et al., 1986).

Looked at from the perspective of the presenting problem, couple and family therapies have been demonstrated to have considerable value in treating depression, anxiety disorder, panic disorder, schizophrenia, alcoholism and marital maladjustment in adults, conduct disorder, autism, and drug abuse in children and adolescents (Lebow & Gurman, 1995; Pinsof & Wynne, 1995). Most of these demonstrations of effectiveness include a combination of couple and family therapy with other interventions. Offered alone, couple and family therapies emerge as the only demonstrated effective means for impacting on couple and family issues (e.g., couple conflict and marital maladjustment). Frequently it appears that even a small amount of family involvement adds immeasurably to treatment effectiveness and increases acceptability and participation in treatment (Lebow & Gurman, 1995). There are indications that treatments diminish in their effectiveness over time. Although outcomes are quite impressive in the short term, effects often dissipate (Lebow & Gurman, 1995).

An outstanding body of research is now also available associated with family process and family development (Cox & Paley, 1997). This work implies many hypotheses for treatment research. For example, Gottman and associates have carried out several studies that have added immeasurably to our knowledge of patterns and sequences of dysfunction in marriage (Buehlman, Goldman, & Katz, 1992; Gottman, 1991, 1992, 1993; Gottman & Krofkoff, 1989), such as showing particular patterns in couples conflict that directly lead to decreasing levels of marital satisfaction and ultimately to divorce. Similarly, powerful bodies of research concerned with such issues as patterns in divorcing and remarried families (Bray & Hetherington, 1993), family transitions around the birth of children (Cowan & Cowan, 1992), and patterns in the alcoholic family (Steinglass, Bennett, Wolin, & Reiss 1987) have clear clinical implications.

A few other trends in the research that have emerged are particularly noteworthy. We are seeing a trend toward more clinical trials research comparing treatments. Paradoxically, research also shows the impact the intense study of process can focus on a few cases, particularly when the cases are selected by outcome status and the methods for assessing process clearly focus on change events rather than engaging in a hunting expedition (Greenberg, 1991). We also are beginning to see a great deal of research on treatments that transcend the labels "individual," "couple," or "family" therapy, just as clinical methods are moving to transcend these boundaries (Lebow, 1987a, 1987b). We are also seeing more efforts to be conscious of gender and culture in research. No longer is the assumption made that findings are necessarily generalizable across gender or culture. Perhaps most promising, family therapists and family researchers have begun to engage in dialogue, suggesting that the notable gap between research and practice may narrow.

1.16.8 TRENDS IN COUPLE AND FAMILY THERAPY

A number of trends are clearly emerging in family therapy. These include:

(i) Most family therapies have moved away from radical systemic positions, which focused solely on interpersonal processes and viewed the etiology of disorder, treatment, and the assessment of change exclusively in terms of the properties of systems (e.g., circular causality and homeostasis). These radical systemic viewpoints showed little respect for any linear epistemology, including the traditional methods of empirical research, and focused attention on a complex set of variables that were extremely difficult to measure. A shift has occurred in which interpersonal processes, feedback loops, and the system's role in problem generation and resolution are still emphasized; but attention broadens to include individual functioning, the larger social system, and other levels of analysis.

(ii) The boundaries among individual, couple, and family therapy and between schools of family therapy are blurring. There are more and more integrative therapies that include couple and family treatment. Many therapies that include significant aspects of systems theory and interventions with couples and family do not label themselves "family therapy" (Pinsof & Wynne, 1995).

(iii) Methods of practice in family therapy are becoming more generic. Therapists practice in ways that are more integrative, and utilize strategies and interventions that cross a range of schools of practice. Concepts like therapeutic alliance (Pinsof & Catherall, 1986), loss (McGoldrick & Walsh, 1983), life cycle (Carter & McGoldrick, 1988), cohesion (Olson, 1986), isomorphism (Liddle, 1984); macrosystem

(Imber-Black, 1988), and differentiation (Bowen, 1978); interventions such as enactment (Minuchin & Fishman, 1981), ritual (Imber-Black & Roberts, 1993), and genogram (McGoldrick & Gerson, 1985); and tactics for dealing with such processes as engagement (Minuchin & Fishman, 1981), resistance (Anderson & Stewart, 1983), and termination (Lebow, 1995) are used by a wide range of family therapists, transcending the approach in which they were first utilized. We have also seen the development of a common language that transcends approach (Simon, Stierlin, & Wynne, 1985), and the beginnings of catalogs of interventions that transcend orientation (Figley & Nelson, 1989; Minuchin & Fishman, 1981).

(iv) Biology is clearly emerging in the literature as an important variable in family life, in such disparate contexts as conflict in marriage (Gottman & Levenson, 1992) and processes around schizophrenia in the family (Leff & Vaughn, 1985). Family reactions to crises in health (Rolland, 1994) and, more broadly, the field of family systems medicine have come to occupy much attention.

(v) Many models now feature a very strong cognitive emphasis on intervention strategy. Behavioral models have added cognitive theory and intervention, strategic models a heavy emphasis on reframing, and narrative models fully focus on the creation of new stories about events.

(vi) There is a growing tendency to develop specific variants of treatment to respond to the specific issues in particular kinds of cases, especially different diagnostic groups. Such treatments as psychoeducational intervention for schizophrenia (Hogarty et al., 1986), cognitive-behavioral couple treatment of depression (O'Leary & Beach, 1990), multisystemic treatment (Henggeler, 1993) and functional family therapy for adolescent delinquent behavior (Alexander & Parsons, 1982), and multidimensional family therapy for adolescent drug abuse (Liddle et al., 1992); all are fully structured and adapted to the specific population in focus.

(vii) Much attention has been focused on the meaning and methods for intervening in family violence (Avis, 1992; Goldner et al., 1990). In attempting to deal with family violence, many assumptions of systems theory have required modification, for example, the absolute centrality of circular causal explanations.

(viii) As previously noted, gender and culture have emerged as vital considerations in shaping practice. No longer is the assumption made that the same treatment will be appropriate across varying groups (McGoldrick et al., 1988). Patterns of behavior and expectations vary and generalizations about family process must always be viewed through the lenses of gender and culture.

This focus on gender and culture has taken two forms. Some have offered critiques of the concepts of family therapy, explicating the ethnocentric and gendered aspects of the most widely taught models of treatment. For example, Green and Werner (1996) have focused on the tendency to view closely connected families, frequently found in many minority cultures as pathologically enmeshed. Green and Werner argue that closeness can be and often is a positive value. Similarly, Goldner (1985) has described the gender assumptions implicit in most schools of family therapy. The second tact has been to explicate specific methods that fit with the culture of particular groups of individuals (McGoldrick et al., 1988) or powerful gendered issues (Carter, 1996; Goldner et al., 1990).

(ix) Others have become increasingly concerned with the impact and relationship of systems larger than the family. Early family therapy focused on the family and its dynamics. More recent thinking has viewed the family as only one of multiple systems impinging on the individual. The family system is viewed in relation to a number of other influences, such as schools, the legal system, poverty, and peer groups. Multisystemic treatments targeting multiple systems are beginning to emerge (e.g., Henggeler, 1993; Liddle et al., 1992). Some family therapists (Wynne et al., 1986) now even refer to themselves as system consultants.

(x) An increasing emphasis in family therapy has been on a life cycle perspective (Carter & McGoldrick, 1988). It has become typical to consider a family in relation to where they are within the developmental process, and shape interventions accordingly.

(xi) There is much more emphasis than earlier in family therapy on health, normality, and resilience, as opposed to homeostasis, resistance, and pathology (Walsh, 1993). Not only do the specific approaches that have captured the most recent attention feature these assumptions (e.g., solution focused and narrative), but most family therapy now emphasizes resilience and coping.

1.16.9 THE DEVELOPMENT OF FAMILY PSYCHOLOGY

From its beginning, couple and family therapy has been a multidisciplinary endeavor. Perhaps more than in the development of any other psychotherapy, psychologists, psychiatrists, social workers, psychiatric nurses, and others (even including anthropologists and sociologists) have worked together in harmony. Indeed, many of the central developments in the field have been carried out in family institute settings, away from the organized centers of training in specific disciplines.

The downside of this pattern of development has been that it has taken considerable time for family therapy to occupy a central place within the specific disciplines, including psychology. Well after the emergence of family therapy as a much practiced endeavor, few training programs within the mental health disciplines offered training in systems therapy. For example, Stanton (1976) reported that as late as the mid-1970s, only 10 psychology programs offered family therapy training. However, since the mid-1970s, this pattern has substantially changed, as structures have evolved in training and practice that recognize the importance of the family system in psychology.

Foremost has been the creation of a subdiscipline, family psychology. The American Psychological Association has recognized the importance of family psychology in creating the Division of Family Psychology (Division 43) and a journal to represent scholarship in this discipline, *Journal of Family Psychology*. In surveys, 63% of psychology programs with special treatment services had family therapy clinics (Sayette & Mayne, 1990), 19% of all faculty had systemic theoretical orientation (Mayne, Norcross, & Sayette, 1994), and family research was the second most represented area of research across departments (Sayette & Mayne, 1990). Eighty-five percent of psychology internships reported some marital and family training (Soloman, Ott, & Roach, 1986). In many departments, a specific specialization in family psychology has been created (Green, 1996). In addition, many psychologists continue to receive training in postdegree programs in family institute settings.

As this is being written, discussion is underway about whether family psychology should be regarded as a specialization on a par with clinical and counseling psychology or whether it should be regarded as a proficiency lying within clinical psychology. Green (1996) has argued cogently for the former, suggesting that "there is nothing about clinical or counseling psychology that makes them inherently more generalist than a family psychology program would be." Green further suggests the creation of a full predoctoral specialization in family psychology as well as a postdegree specialization. Whatever the outcome of this debate, it is clear that family psychology has moved to a central place within psychology, just as systemic therapies have become among the most widely practiced therapies, and as couple and family intervention has come to be seen as essential in the treatment of a wide array of problems.

1.16.10 SUMMARY

Systems formulations of family process and family systems orientations to psychotherapy have been prominent methods of conceptualization and intervention since the early 1960s. A variety of distinct forms of couple and family therapy have developed and become widely practiced, including structural, strategic, experiential, inter-generational, cognitive-behavioral, and pyscho-educational methods. Approaches to couple and family therapy are increasingly integrative, incorporating a range of concepts and intervention strategies, and including a focus on the individuals within the family as well as on the broader social system in which the family resides. In the future, we are likely to see the further development of such therapies that feature a systemic orientation, but which are less distinctly "family" therapies, in which therapists move their focus across levels of social system: from family, to couple, to individual, to the larger system.

Knowledge derived from research about the family system and family therapy also is growing. Much of this research supports the tenets of systems theory as an explanatory set of concepts (Cox & Paley, 1997). This research is also helping to build a base of greater understanding of family processes. Research assessing the efficacy of couple and family therapy has consistently demonstrated its efficacy. Specifically, couple and family therapy appears to have a unique impact in alleviating relationship difficulties, to contribute in important ways in the multimodal treatment of psychopathology, and to help engage many clients who cannot be engaged in other forms of treatment.

Family psychology is emerging as a prominent field within psychology. In the twenty-first century, we are likely to see it occupy an even more central place within the discipline, with greater formal recognition of the specific skill base needed to work with couples and families.

1.16.11 REFERENCES

Ackerman, N. W. (1958). *The psychodynamics of family life*. New York: Basic Book.

Alexander, J. F. Holzworth-Monroe, A., & Jameson (1994). The process and outcome of marital and family therapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed.). New York: Wiley

Addis, M. A., & Jacobson, N. S. (1991). Integration of cognitive therapy and behavioral marital therapy for depression. *Journal of Psychotherapy Integration, 1,* 249–264.

Alexander J. F., & Parsons, B. (1982). *Functional family therapy: Principles and procedures.* Carmel, CA: Brooks/Cole.

Anderson, H., & Goolishian, H. (1988). Human systems as linguistic systems. *Family Process, 27*, 371–393.

Anderson, C. M., Reiss, D., & Hogarty, B. (1986). *Schizophrenia and the family.* New York: Guilford Press.

Anderson, C. M., & Stewart, S. (1983). *Mastering resistance: A practical guide to family therapy.* New York: Guilford Press.

Aponte, H. (1985), The negotiation of values in therapy. *Family Process, 24*, 323–338.

Avis, J. M. (1992). Where are all the family therapists? Abuse and violence within families and family therapy's response. *Journal of Marital Family Therapy, 18,* 223–230.

Barton C., & Alexander J. F. (1981). Functional family therapy. In A. S. Gurman & D. P. Kniskern (Eds.), *Handbook of family therapy.* New York: Brunner/Mazel.

Bateson, G., Jackson, D., Haley, J., and Weakland, J. (1954). Toward a theory of schizophrenia. *Behavioral Science, 1*, 251–264.

Baucom D. H., & Epstein N. (1990) *Cognitive-behavioral marital therapy.* New York: Brunner/Mazel.

Baucom, D. H., & Hoffman, J. A. (1986). The effectiveness of marital therapy: Current status and application to the clinical setting. In N. Jacobson & A. Gurman (Eds.), *Clinical handbook of marital therapy* (pp. 597–620). New York: Guilford Press.

Beach, S. R. H., & O'Leary, K. D. (1992). Treating depression in the context of marital discord: Outcome and predictors of response of marital therapy versus cognitive therapy. *Behavior Therapy, 23,* 507–528.

Bednar, R. L., Burlingame, G. M., & Masters, K. S. (1988). Sysems of family treatment: Substance or semantics? *Annual Review of Psychology, 39,* 401–434.

Berg, I. K. (1993). *Family based services.* New York: Norton.

Boscolo L., Cecchin G., Hoffman L., & Penn P. (1987). *Milan systemic therapy.* New York: Basic Books.

Boszormenyi-Nagy, I., & Spark, G. (1973). *Invisible loyalties.* Hagerstown, MD: Harper and Row.

Bowen, M. (1978). *Family therapy in clinical practice.* Northvale, NY: Jason Aronson.

Boyd-Franklin, N. (1989). *Black families in family therapy: A multi-systems approach.* New York: Guilford Press.

Buehlman, K. T., Gottman, J. M., & Katz, L. F. (1992). How a couple views their past predicts their future: Predicting divorce from an oral history interview. *Journal of Family Psychology, 5*(3, 4), 295–318.

Bray, J., & Hetherington, E. M. (1993). Families in transition: Introduction and overview. *Journal of Family Psychology, 7,* 3–8.

Breunlin D., Schwartz, R., & Karrer, B. (1992). *Metaframeworks: Transcending the models of family therapy.* San Francisco: Jossey-Bass.

Carter, B. (1996). *Love, honor, & negotiate.* New York: Pocket Books.

Carter, B., and McGoldrick, M. (1988). *The changing family life cycle: A framework for family therapy* (2nd ed.) Boston: Allyn & Bacon.

Cecchin, G. (1987). Hypothesizing, circularity and neutrality revisited: An invitation to curiosity. *Family Process, 26*, 405–413.

Combs, G., & Freedman, J. (1990). *Symbol, story, and ceremony.* New York: Norton.

Cowan, C. P., & Cowan, P. A. (1992). *When partners become parents.* New York: Basic Books.

Cox, M. J., & Paley, B. (1997). Families as systems. *Annual review of psychology, 48*, 243–267.

Dell, P. (1986). In defense of lineal causality. *Family Process, 25*, 515–524.

De Shazer, S. (1985). *Keys to solution in brief therapy.* New York: Norton.

De Shazer S. 1988. *Clues: Investigating solutions in brief therapy.* New York: Norton.

Falloon I, Boyd J., & McGill C. (1984). *Family care of schizophrenia: A problem-solving approach to the treatment of the mentally ill.* New York: Guilford Press.

Falloon I. R. H., Boyd J. L., McGill C. W., Williamson M., & Razani J., (1985). Family management in the prevention of morbidity of schizophrenia: Clinical outcome of a two year longitudinal study. *Archives of General Psychiatry, 42*, 887–896.

Feldman, L. B. (1985). Integrative multi-level therapy: A comprehensive interpersonal and intrapsychic approach. *Journal of Marital and Family Therapy, 11,* 357–372.

Feldman, L. B. (1990). *Multi-dimensional family therapy.* New York: Guilford Press.

Feldman, L. B., & Pinsof, W. M. (1982). Problem maintenance in family systems: An integrative model. *Journal of Marriage and Family Therapy, 8*, 295–308.

Figley, C., & Nelson, T. S. (1989). Basic family therapy skills 1: Conceptualization and initial findings. *Journal of Marital and Family Therapy, 15*, 349–365.

Framo, J. L. (1981). The integration of marital therapy with sessions with family of origin. In A. S. Gurman & D. P. Kniskern (Eds.), *Handbook of family therapy.* New York: Brunner/Mazel.

Gergen, K. J. (1981). The social constructionist movement in modern psychology. *American Psychologist, 40,* 266–275.

Gergen, K. (1991). *The saturated self.* New York: Basic Books.

Glaserfeld, E. Von. (1984). An introduction to radical constructivism. In P. Watzlawick (Ed.), *The invented reality. How do we know what we believe we know?* (pp. 17–40). New York: Norton.

Goldner, V. (1985). Feminism and family therapy. *Family Process, 24*, 31–48.

Goldner, V., Penn, P., Sheinberg, M., & Walter, G. (1990). Love and violence: Gender paradoxes in volatile attachments. *Family Process, 29*, 343–364.

Gottman, J. M. (1991). Predicting the longitudinal course of marriages. *Journal of Marital and Family Therapy, 17*(1), 3–7.

Gottman, J. (1992). The roles of conflict engagement, escalation, and avoidance in marital interaction: A longitudinal view of five types of couples. *Journal of Consulting and Clinical Psychology, 61,* 6–15.

Gottman, J. (1993). A theory of marital dissolution and stability. *Journal of Family Psychology, 7,* 57–75.

Gottman, J. M., & Krokoff, L. (1989). Marital interaction and marital satisfaction. A longitudinal view. *Journal of Consulting and Clinical Psychology, 57,* 47–52.

Gottman, J. M., & Levenson, R. W. (1992). Marital processes predictive of later dissolution: Behavior, physiology, and health. *Journal of Personality and Social Psychology, 63*(2), 221–233.

Gottman, J., Markman, H., & Notarius, C. (1977). The topography of marital conflict: A sequential analysis of verbal and nonverbal behavior. *Journal of Marriage and the Family, 39*, 461–477.

Green, R. J. (1986). Education in family psychology. *Family Psychologist, Winter*, 10–15.

Green, R. J., & Werner, R. (1996). Enmeshment and disengagement. *Family Process, 35,* 112–132.

Greenberg, L. S. (1991). Research on the process of change. *Psychotherapy Research, 1*(1), 3–16.

Greenberg, L. S., Ford, C. O., Alden, L. S., & Johnson, S. M. (1993). In-session change in emotionally focused therapy. *Journal of Consulting and Clinical Psychology, 61*, 78–84.

Greenberg, L. S., & Johnson, S. M. (1988). *Emotionally focused therapy for couples*. New York: Guilford Press.

Grunebaum, H. (1988). The relationship of family theory to family therapy. *Journal of Marital & Family Therapy, 14*, 1–14.

Guerney, B. G. (1977). *Relationship enhancement: Skill training programs for therapy, problem prevention and enrichment*. San Francisco: Jossey-Bass.

Gurin, E., Veroff, J., & Feld, S. (1960). *Americans view their mental health*. New York, Basic Books.

Gurman, A. S. (1978). Contemporary marital therapies. In T. Paolino & B. McCrady, *Marriage and marital therapy: Psychoanalytic, behavioral, and system theory perspectives*. New York: Brunner-Mazel,

Gurman, A. S. (1981). Integrative marital therapy: Toward the development of an interpersonal approach. In S. Budman (Ed.) *Forms of brief therapy*. New York: Guilford Press.

Gurman, A. S., & Kniskern, D. P. (1978). Research on marital and family therapy: progress, perspective, and prospect. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis* (2nd ed., pp. 817–901). New York: Wiley.

Gurman, A. S., & Kniskern, D. P. (1978). Technolatry, methodolatry, and the results of family therapy. *Family Process, 17*, 275–281.

Gurman, A. S., & Kniskern, D. P. (1981). *Handbook of family therapy*. New York: Brunner/Mazel.

Gurman, A. S., Kniskern, D. P., & Pinsof W. M. (1986). Research on marital and family therapies. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed., pp. 565–624). New York: Wiley.

Gurman, A. S. & Knudson, R (1978). Behavioral marriage therapy: A psychodynamic-systems analysis and critique. *Family Process, 17*, 121–138.

Haley, J. (1963). *Strategies of psychotherapy*. New York: Grune & Stratton.

Haley, J. (1976). *Problem-solving therapy*. San Francisco: Jossey-Bass.

Haley, J. (1980). *Leaving home*. New York: McGraw-Hill.

Hare-Mustin, R. T., & Marecek, J. (1988). The meaning of difference: Gender theory, postmodernism and psychology. *American Psychologist, 43*, 455–464.

Harkaway, J. E. (1989). Family intervention in the treatment of childhood and adolescent obesity. In J. E. Harkaway (Ed.), *Eating disorders*. Rockville, MD: Aspen.

Henggeler S. W. (1993). Multi-systemic treatment of serious juvenile offenders: implications for the treatment of substance-abusing youths. In L. S. Onkine, J. D. Blaine, & J. J. Borne (Eds.), *Behavioral treatments for drug abuse and dependence* (pp. 181–199). Rockville, MD: National Institute on Drug Abuse Research Monograph 137.

Henggeler, S. W., Melton, G. B., & Smith, L. A. (1992). Family preservation using multisystemic therapy: An effective alternative to incarcerating serious juvenile offenders. *Journal of Consulting and Clinical Psychology, 60*(6), 953–961.

Hoffman, L. (1981). *Foundations of family therapy: A conceptual framework for system change*. New York: Basic Books.

Hogarty, G., Anderson, C. M., Reiss, D. J., Knornblith, S. J., & Greenwald, D. P., (1986). Family psychoeducation, social skills training and maintenance chemotherapy in the aftercare treatment of schizophrenia: I. One-year effects of a controlled study on relapse and expressed emotion. *Archives of General Psychiatry, 43*, 633–642.

Imber-Black, E. (1988). *Families and larger systems: A family therapist's guide through the labyrinth*. New York: Guilford Press.

Imber-Black, E., & Roberts, J. (1993). *Rituals from our times: Celebrating, healing, and changing our lives and our relationships*. New York: Harper Collins.

Jacobson, N. S. (1992). Behavioral couple therapy: A new beginning. *Behavior Therapy, 23*, 493–505.

Jacobson, N. S., & Addis, M. E. (1993). Research on couple therapy: What do we know? where are we going? *Journal of Consulting and Clinical Psychology, 61*(1), 85–93.

Jacobson, N. S., & Christenson A. (1996). *Acceptance and change in couple therapy*. New York: Norton.

Jacobson, N. S., Dobson, K., Fruzzetti, A., Schmaling, K. B., & Salusky, S. (1991). Marital therapy as a treatment for depression. *Journal of Consulting and Clinical Psychology, 59*(4), 547–557.

Jacobson, N. S., & Margolin, G. (1979). *Marital therapy: Strategies based on social learning and behavior exchange principles*. New York: Brunner/Mazel.

Kaplan, H. S. (1979). *The new sex therapy: active treatment of sexual dysfunctions*. New York: Brunner/Mazel.

Kirschner, D., & Kirschner, S. (1986). *Comprehensive family therapy: An integration of systemic and psychodynamic treatment models*. New York: Brunner/Mazel.

L'Abate, L. (1986). *Systemic family therapy*. New York: Brunnel-Mazel.

Lebow, J. (1981). Issues in the assessment of outcome in family therapy. *Family Process, 20*, 167–188.

Lebow, J. L. (1987a). Developing a personal integration in family therapy: Principles for model construction and practice. *Journal of Marital and Family Therapy, 13*, 1–14.

Lebow, J. L. (1987b). Integrative family therapy: An overview of major issues. *Psychotherapy, 40*, 584–594.

Lebow, J. L. (1984). On the value of integrating approaches to family therapy. *Journal of Marital and Family Therapy, 10*, 127–138.

Lebow, J. L. (1995). Termination in marital and family therapy. In R. H. Mikesell, D. D. Lusterman, & S. H. McDaniel (Eds.), *Integrating family therapy handbook of family psychology and systems theory* (1st ed.). Washington, DC: American Psychological Association.

Lebow, J., & Gurman, A. S. (1995). Research assessing couple and family therapy. *Annual Review of Psychology, 46*, 27–57.

Leff, J., & Vaughn, C. (1985). *Expressed emotion in families: Its significance for mental illness*. New York: Guilford Press.

Liddle, H. A. (1984). Towards a dialectical-contextual-coevolutionarey translation of structural-strategic family therapy. *Journal of Strategic and Systemic Therapies, 3*, 57–66.

Liddle, H. A., Dakof, G. A., & Diamond, G. (1992). Adolescent substance abuse: Multidimensional family therapy in action. In E. Kaufman & P. Kaufmann (Eds.), *Family therapy of drug and alcohol abuse* (2nd ed.) Boston: Allyn & Bacon.

Lidz T., Cornelison, A., Fleck, S., & Terry, D. (1957). Intrafamilial environment of the schizophrenic patient. *Psychiatry, 20* 329–342.

LoPiccolo, J., & LoPiccolo, L. (1978). *Handbook of sex therapy*. New York: Plenum.

Madanes, C. (1981). *Strategic family therapy*. New York: Jossey-Bass.

Markman, H. J. (1979). Application of a behavioral model of marriage in predicting relationship satisfaction of couples planning marriage. *Journal of Consulting and Clinical Psychology, 47*, 743–749.

Markman, H. J., Floyd, F. J., Stanley, S. W. & Lewis, H. C. (1986). Prevention. In N. S. Jacobson & A. S. Gurman (Eds.), *Clinical handbook of marital therapy*. New York: Guilford Press.

Masters, W. H., & Johnson, V. S. (1970). *Human sexual inadequacy*. Boston: Little-Brown.

Maturana, H. R., & Varela, F. J. (1980). *Autopoesis and cognition*. Boston: D. Riedel

Maturana, H. R., & Varela, F. J. (1987). *The tree of knowledge*: *The biological roots of human understanding*. Boston: New Science Library.

Mayne, T. J., Norcross, J. C., & Sayette, M. A. (1994). Admission requirements, acceptance rates, and financial assistance in clinical psychology programs. *American Psychologist, 49*, 806–811.

McGoldrick, M., & Gerson, R. (1985). *Genograms in family assessment*. New York: Norton.

McGoldrick, M., Pierce, C., & Giordano, J. (1988). *Ethnicity and family therapy*. New York: Guilford Press.

McGoldrick, M., & Walsh, F. (1983). A systemic view of family history and loss. In L. R. Wolberg & M. L. Aronson (Eds.), *Group and family therapy*. New York: Brunner-Mazel.

Miklowitz, D. J., & Goldstein, M. J. (1990). Behavioral family treatment for patients with bipolar affective disorder. *Behavior Modification, 14*, 457–489.

Minuchin, S. (1974). *Families and family therapy*. Cambridge, MA: Harvard University Press.

Minuchin, S. (1996). Interview. *Family Therapy Networker, 20*(November/December), 50–57.

Minuchin, S. & Fishman, C. (1981). *Family therapy techniques*. Cambridge, MA: Harvard University Press.

Moultrop, D. (1981). Towards an integrated model of family therapy. *Clinical Social Work Journal, 9,* 111–125.

Moultrop, D. J. (1989). Integration: A coming of age. *Contemporary Family Therapy, 8*, 159–167.

Nichols, M. P. (1987). *The self in the system: Expanding the limits of psychotherapy*. New York: Brunner-Mazel.

Nichols, W. C. (1995). *Treating people in families*. New York: Guilford Press.

O'Hanlon W., & Weiner-Davis, M. (1989). *In search of solutions*. New York: Norton.

O'Leary, K. D., & Beach, R. H. (1990). Marital therapy: A viable treatment for depression and marital discord. *American Journal of Psychiatry, 147*, 183–186.

Olson, D. H. (1986). Circumplex model VII: Validation studies and FACES III. *Family Process, 25,* 337–351.

Patterson G. R. (1982). *Coercive family process*. Eugene, OR: Castalia.

Patterson, G. R., & Chamberlain, P. (1992). A functional analysis of resistance (A neobehavioral perspective). In H. Arkowitz (Ed.), *Why don't people change? New perspectives on resistance and noncomplicance*. New York: Guilford Press.

Piercy, F. P., & Frankel, B. R. (1989). The evolution of an integrative family therapy for substance abusing adolescents. *Journal of Family Psychology, 3,* 5–21.

Pinsof, W. M. (1983). Integrative problem-centered therapy: Toward the synthesis of family and individual psychotherapies. *Journal of Marital and Family Therapy, 9*, 19–35.

Pinsof, W. & Catherall, D. (1986). The integrative psychotherapy alliance. *Journal of Marital and Family Therapy, 12*, 137–152.

Pinsof, W. (1995). *Integrative problem centered therapy*. New York: Basic Books.

Pinsof, W., & Wynne, L. (1995). Special Issue: The effectiveness of marital and family therapy. *Journal of Marital and Family Therapy, 21,* 339–610.

Reiner, P. (1997). Psychoanalytic approaches to supervising marriage and family therapy. In C. Storm & T. Todd (Eds.), *The complete systemic supervisor*. Boston: Allyn & Bacon.

Rolland, J. S. (1994). *Families, illness and disability: An integrative treatment model*. New York: Basic Books.

Sander, F. (1979). *Individual and family therapy*. New York: Jason Aronson.

Satir, V. (1983). *Conjoint family therapy*. Palo Alto, CA: Science and Behavior Books.

Sayette, M. A., & Mayne, T. J. (1990). Survey of current clinical and research trends in clinical psychology. *American Psychologist, 45,* 1263–1266.

Scharff, D., & Scharff, J. S. (1987). *Object relations family therapy*. New York: Jason Aronson.

Selvini-Palazzoli, M. (1986). Toward a general model of psychotic games. *Journal of Marital and Family Therapy, 12,* 339–349.

Selvini-Palazzoli, M., Boscolo, L., Cecchin, G., & Prata, G. (1978). *Paradox and counterparadox*. New York: Jason Aronson

Simon, F., Stierlin, H., & Wynne, L. (1985). *The language of family therapy: A systemic vocabulary and source book*. New York: Family Process Press.

Slipp, S. (1993). *Object relations: A dynamic bridge between individual and family treatment*. New York: Jason Aronson.

Snyder, D. K. & Wills, R. M. (1989). Behavioral versus insight-oriented marital therapy: Effects on individual and interspousal functioning. *Journal of Consulting and Clinical Psychology, 57*, 39–46.

Soloman, J., Ott, J., & Roach, A. (1986). A survey of predoctoral psychology interns in marriage and family therapy. *Journal of Marital and Family Therapy, 12,* 269–280.

Stanton, M. D. (1975). Family therapy training: Academic and internship opportunities for psychologists. *Family Process, 14*, 433–439.

Steinglass, P. (1992). Treatment of alcoholism. American Association of Marital and Family Therapy Annual Meeting, Miami, FL.

Steinglass, P., Bennett, L. A., Wolin, S. J., & Reiss, D. (1987). *The alcoholic family*. New York: Basic Books.

Stuart, R. (1980). *Helping couples change*. New York: Guilford Press.

Szapocznik, J., & Kurtines, W. M. (1989). *Breakthroughs in family treatment*. New York: Springer

Thibaut, J., & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley

Trepper, T. S., & Barrett, M. J. (1989). *Systemic treatment of incest: A therapeutic handbook*. New York: Brunner/Mazel.

Von Bertalanffy, L. (1969). *General systems theory: Essays on its foundation and development*. New York: George Braziller.

Wachtel, E. (1992). An integrative approach to working with troubled children and their families. *Journal of Psychotherapy Integration, 2,* 207–224.

Wachtel, E. F. (1995). *Treating troubled children and their families*. New York: Guilford Press.

Wachtel E. F., & Wachtel, P. L. (1995). *Family dynamics in individual psychotherapy*. New York: Guilford Press.

Walsh, F. (Ed.) (1982). *Normal Family Processes; A guide to clinical strategies*. New York: Guilford Press.

Walsh, F. (Ed.) (1993) *Normal family processes* (2nd ed.). New York: Guilford Press.

Watts-Jones, D. (1992). Cultural and integraive therapy issues in the treatment of a Jamaican woman with panic disorder. *Family Process, 31,* 105–113.

Watzlawick, P., Weakland, J., & Fisch, R. (1974). *Change: Principles of problem formation and problem resolution*. New York: Norton.

Watzlawick, P., Beavin, P. and Jackson, D. (1967). *Pragmatics of human communication*. New York: Norton.

Whitaker, C. A. & Keith, D. V. (1981) Symbolic-experiential family therapy. In A. S. Gurman & D. P. Kniskern (Eds.), *Handbook of family therapy*. New York: Brunner/Mazel.

Wiener, N. (1991). *Cybernetics or control and communication in the animal and in the machine*. Cambridge, MA: MIT Press.

Weiss, R. (1978). The conceptualization of marriage from a behavioral perspective. In T. Paolino & B. McCrady. (Eds.), *Marriage and marital therapy: Psychoanalytic, behavioral, and systems theory perspectives*. New York: Brunner/Mazel.

White, M., & Epston, D. (1990). *Narrative means to therapeutic ends*. New York: Norton.

Wood, B. L. Beyond the psychosomatic family: A biobehavioral model of pediatric illness. *Family Process, 32*, 261–278.

Wright, L. M., & Leahey, M. (1994). Calgary family intervention model: One way to think about change. *Journal of Marital and Family Therapy, 20,* 381–396.

Wynne, L., McDaniel, S., & Weber T. (1986). Systems consultation: A new perspective for family therapy. New York: Guilford Press.

Wynne, L. C., Ryckoff, I., Day, J., & Hirsch, S. I. (1958). Pseudomutuality in the family relationships of schizophrenics. *Psychiatry, 20,* 205–220.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.01

History and Emerging Trends in Education and Training for Clinical Psychology in the European Union

INGRID LUNT
University of London, UK

2.01.1 INTRODUCTION 2

2.01.2 EMERGENCE OF CLINICAL PSYCHOLOGY 2
  2.01.2.1 Scientist–Practitoner Model 3

2.01.3 PATTERNS OF UNIVERSITY EDUCATION AND TRAINING 4
  2.01.3.1 The Continuous Specialist Model 4
  2.01.3.2 The Continuous and Generic Education Model 4
  2.01.3.3 Discontinuous and Specialist Education or the Three-plus-three Model 4

2.01.4 MOBILITY WITHIN THE EUROPEAN UNION 5

2.01.5 THE ROLE OF THE EUROPEAN FEDERATION OF PROFESSIONAL PSYCHOLOGISTS ASSOCIATIONS 6

2.01.6 TRAINING ROUTES IN THE DIFFERENT MEMBER STATES OF THE EUROPEAN UNION 7
  2.01.6.1 The Nordic Countries (Denmark, Finland, Sweden) 7
  2.01.6.2 The United Kingdom 7
  2.01.6.3 Ireland 8
  2.01.6.4 Greece 8
  2.01.6.5 France 8
  2.01.6.6 Spain 8

2.01.7 DEMOGRAPHIC FEATURES OF CLINICAL PSYCHOLOGISTS 8

2.01.8 FUTURE ISSUES 8

2.01.9 SUMMARY 9

2.01.10 APPENDIX 1: EFPPA OPTIMAL STANDARDS FOR PROFESSIONAL TRAINING 9

2.01.11 APPENDIX 2: MEMBER ASSOCIATIONS OF EFPPA (1997) 10

2.01.12 APPENDIX 3: BIBLIOGRAPHY 10

2.01.13 REFERENCES 11

2.01.1 INTRODUCTION

This chapter covers some of the history and emerging trends in education and training for clinical psychology mainly, but not only, in the European Union (EU). The EU is developing and changing rapidly. In 1997 there are 15 member states, covering most of the western European countries (but not Iceland, Norway, and Switzerland). However, many of the central and eastern European countries have applied for membership, and are working towards criteria which will enable them to join. This will mean an expanded and changed Europe in the not too distant future.

The countries of Europe have very different histories, which are reflected in the development of their professions and, of particular relevance here, in the development of clinical psychology as a profession and the training for the profession. Clearly, increased mobility between students, practitioners, and researchers, and some of the moves within the EU, are having the effect of bringing together some of the trends and achieving greater "convergence," a term used in the EU. Nevertheless, in the field of psychology and its education and training, the European countries remain strikingly different, at least between regional groupings. This chapter will attempt to consider European clinical psychology education and training in its similarities and its differences.

As mentioned, the region of Europe consists both of the EU and nonmember states, many of the latter in eastern Europe and formerly belonging to the Soviet Union. The countries of Europe constitute a very diverse group which may nevertheless be grouped into broad regions with some commonality in education and training patterns. For these purposes, at a very broad level it is possible to identify a Nordic group of countries (the Scandinavian countries), with considerable commonalities in their education and training; the UK, which has much in common with education and training in the USA and other parts of the English-speaking world; the German-speaking countries; and a southern European group. Of course, within these regions, what characterizes Europe is its diversity and the individuality of its countries.

This chapter will, to some extent, use regional groupings in its consideration of clinical psychology education and training. There are a number of models or definitions of the field of clinical psychology, a field which emerged in European countries substantially after World War II. This issue is particularly important as there are moves towards greater mobility following the Treaty of Rome and the European Directives (in particular Directive EC 89/48). Education and training in the different countries follow very different patterns, which are related to the traditions and history of university education in the different countries. Nevertheless, it is possible to distinguish three or four broad patterns or models of education and training. All countries are moving towards increasing specialization and postqualification training.

The chapter will draw substantially on work carried out within the European Federation of Professional Psychologists' Associations (EFPPA), a federation of national psychology associations with 28 member associations (1997) representing all countries of the EU, all other countries in western Europe, and a growing number of member associations from central and eastern Europe. This federation provides a unique opportunity for comparison between the practices of different European countries, and a forum for discussion and debate of important issues.

2.01.2 EMERGENCE OF CLINICAL PSYCHOLOGY

Although there are differences between countries, it is possible to provide a general definition of the field which would be agreeable to all European countries. Clinical psychologists may be defined as psychologists who:

> apply psychology in a clinical context, usually a hospital, medical or community setting, with people (patients or staff) who consider themselves to be in need of a psychological perspective on their lives. In practice, the majority of clinical psychologists contribute to the assessment and treatment of people who see themselves as having psychological problems, such as those with mental health difficulties, but they also work with the handicapped, families, those with learning difficulties, and, more widely, with staff and organisation. (Llewelyn, 1994, p. 115)

In a survey carried out by an EFPPA Task Force on Clinical Psychology, all countries except France agreed with the APA definition of this field.

Clinical psychology emerged as a recognizable profession at different times in different countries. In both the UK and the USA, the two world wars provided a significant impetus to the emergence of this profession, initially through the need to develop psychological tests to recruit suitable personnel. In the UK, the formation of the National Health Service in 1948 provided a considerable impetus to the development of this new profession. In these early days, the role was largely one of laboratory technician administering psychometric and other tests, usually for medical practitioners. However, in the UK the development of behavior therapy led to clinical psychologists developing a therapeutic role, and by the 1960s they had become clinical practitioners in their own right.

Again, in the Nordic countries, clinical psychology emerged substantially after World War II, in the 1940s and early 1950s. The Norwegian Psychological Association was founded in 1934, the Danish Psychological Association in 1947, the Swedish Psychological Association in 1955, and the Finnish Union of Psychologists in 1957. The formation of these psychological associations reflected the emergence in these countries of a profession of psychology, mainly clinical psychology, which has progressed through similar broad major phases as in the UK: a phase focusing on diagnostic examination and testing of individual patients, a phase focusing on therapeutic work involving mainly psychotherapy, a phase focusing on indirect work with other professional groups through techniques such as consultation and training.

In Spain, on the other hand, clinical psychology has a shorter history (Belloch & Olabarria, 1994) and is said to have emerged in the 1970s, along with fundamental changes in Spanish society. Similarly in Italy, clinical psychology as a professional application emerged substantially in the 1960s (Spaltro, 1993) and there remain some tensions over the autonomy and role of clinical psychologists and their relationship with medical practitioners, especially psychiatrists. The Association of Greek Psychologists, founded in 1963, reflects the existence of professional psychology in Greece, although at that time, and until very recently, all clinical psychologists received their training overseas, and the profession was very much dominated by the medical profession (Velli, 1995).

In all European countries, clinical psychologists represent the largest group of psychologists, and have experienced a rapid growth in their number between the 1960s and the 1980s with the expansion of mental health provision and the growing awareness of the contribution of clinical psychology to a wide and diverse range of areas of work (EFPPA, 1997). However, in many European countries the emergence of the specialty of health psychology, with a focus on prevention rather than treatment, and the promotion of health rather than a more therapeutic function, has led to attempts to define and differentiate fields of psychological activity within the mental health field. For example, three EFPPA task forces, focusing on clinical psychology, health psychology, and psychotherapy, have drawn up a model defining overlap and separate areas of activity within the health field, with corresponding commonalities and differences in education and training (EFPPA, 1997).

2.01.2.1 Scientist–Practitoner Model

In the UK, there is a strong commitment to a scientist practitioner model: "the clinical psychologist is first and foremost an 'applied scientist' or 'scientist–practitioner' who seeks to use scientific knowledge to a beneficial end" (Marzillier & Hall, 1990, p. 9). This commitment also characterizes the Nordic countries, which have been influenced substantially by the USA and also by the UK. These countries would sign up to the definition that:

> clinical psychologists share several common attributes. They are psychologists because they have been trained to use the guidelines of knowledge of psychology in their professional work. They are clinicians because they attempt to understand people in their natural complexity and in their continuous adaptive transformations ... they are scientists because they utilise the scientific methods to achieve objectivity and precision in their professional work. Finally they are professionals because they render important human services by helping individuals, social groups and communities to solve psychological problems and improve the quality of life. (Kendall & Norton-Ford, 1982, p. 4)

This model is also espoused by Spanish clinical psychology where Belloch and Olabarria (1994, p. 144) state that clinical psychology training is "very similar to that proposed in the 1940s, in the famous Boulder Conference, organized by the APA." In France, where there is a strong clinical psychoanalytic tradition, there is less of a commitment to the "Boulder" model.

However, some of the problems inherent in the scientist practitioner model, such as bridging "the gap between ... formal training and scientific thinking on the one hand, and the demands of practice on the other" (Raimy, 1950, quoted in Barlow, Hayes, & Nelson, 1984, p. 9) are also evident in some European countries (see Pilgrim & Treacher, 1992, who write, "British training programmes have consistently espoused the scientist–professional model and in so doing have glossed over the difficulties that are inherent to the model," p. 80). The three phases of professional training for psychologists, described by Peterson (1991) as the preprofessional, the scientist professional, and the professional phase appear to be replicated to some extent in European countries, although in the late 1990s all professional training for clinical psychologists is based in universities rather than professional schools, and there are several countries where the education and training is predominantly academic, and is then followed by a separate internship or registration/licenseship phase, usually based in a hospital or clinic.

2.01.3 PATTERNS OF UNIVERSITY EDUCATION AND TRAINING

Europe is characterized by the diversity of its university education systems. For example, the time taken to complete training in the different countries differs, and the degree awarded also varies. Some countries award a diploma or license, others award a degree such as Masters or Doctorate, on completion of professional training as a (clinical) psychologist. Some counries award a general qualification as a professional psychologist, while others award a more specialized qualification, for example, in the field of clinical psychology. However, it is possible (and see below) to state that most countries in Europe require a period of at least five to six years education and training in order to be able to practice as a clinical psychologist, and increasingly they are developing postgraduate qualifications and specialization.

There have been a number of surveys considering patterns of training, carried out mainly by EFPPA task groups (see Section 2.01.5). An early task force on education and training for psychology agreed in 1990 a document entitled *Optimal standards for professional training in psychology* (EFPPA, 1990) (see Appendix 1). Agreed policy in this area, therefore, is a six-year length of education and training, based in a university or other institution of higher education, consisting of a theoretical foundation with some internship, and containing a research element. A more recent task force, which focuses on clinical psychology, compares patterns of education and training in its interim report (EFPPA, 1997). Lunt (1993, 1994) and Newstead and Makinen (1997) have carried out some further comparisons.

As mentioned, most of the countries of western Europe require a similar overall length of time (i.e., between four and six years) for professional qualification as a (clinical) psychologist. The countries vary considerably in issues such as whether there is any practicum or internship within the training or whether this becomes a postgraduate requirement for registration or licensing, and how specialized the training is.

However, it is possible to contrast three different approaches in the EU, and this is particularly the case in relation to how specialized or generic the professional training is. We can distinguish between three general forms of organization in professional training: the "continuous specialist," the "continuous generic," and the "discontinuous specialist" organization; there are, of course, some "mixed mode" (Lunt, 1993; see also Newstead, 1994). Newstead and Makinen (1997) refer to these three forms of organization as the "five year specialist," "five year generalist," and the "three-plus-three" models.

2.01.3.1 The Continuous Specialist Model

In this model, students enroll at the outset on a psychology course which specializes in a single area such as clinical, educational, or organizational psychology. In the EFPPA survey, this model was found in Sweden and Belgium (Newstead, 1994). To an extent, this is the "mixed mode," referred to by Lunt (1993). In a few countries such as The Netherlands with a more "mixed mode," a four to six year program leads to the degree of *Doctorandus* at the same time as the professional title *psycholoog*. This entails a one-year general introduction then a three-year specialization in clinical psychology.

2.01.3.2 The Continuous and Generic Education Model

This is the "five year generalist" model referred to by Newstead and Makinen (1997). Those countries (e.g., the Nordic countries) which have a more continuous and generic training provide a general degree or program in psychology which leads to a qualification to practice as a psychologist (unspecialized). Later specialization and qualification occur on the job. In the EFPPA survey (Newstead, 1994), this is the most frequently found model, found in Austria, Spain, Hungary, Italy, Germany, Switzerland, Denmark, and Norway.

2.01.3.3 Discontinuous and Specialist Education or the Three-plus-three Model

Some countries have a "discontinuous" and "specialist" organization in which students study for three or four years and receive a general qualification in psychology, following which they enroll on a new course (often at a different institution) where they specialize in clinical psychology at Master's or, increasingly, Doctorate level. This model is followed by the UK, Ireland, Malta, and Greece.

These different ways of organizing training in psychology may lead to problems in comparing qualifications across Europe, since a training leading to a qualification to practice as a clinical psychologist, for example, in one country may be either too general or too specific for another European country. In those countries in Europe where the qualification is more generic and where all professional psychology graduates are able to practice in the range of settings as educational, clinical, organizational/industrial psychologists with later specialization occurring on the job and in the specific context, the qualification is likely be too general in comparison with those countries where the basic professional education and training is specialized, and evaluation of equivalence of qualifications is enormously complex.

2.01.4 MOBILITY WITHIN THE EUROPEAN UNION

In the early days of the European Community (EC), a single "common" market within the EC was envisaged in the Treaty of Rome which established the European Economic Community (EEC) in 1957: "Freedom to work anywhere in the European Community is one of the basic rights laid down by the Treaty of Rome." Article 48 of the Treaty provides for the free movement of labor to be achieved by the removal of discrimination, based on nationality, in respect of employment. Article 57 allows for the mutual recognition and coordination of professional qualifications.

In the early stages of the implementation of the Treaty, attempts were made to harmonize qualifications across member countries. Some professions were made subject to their own special or sectoral directive containing agreements on the harmonization (or standardization) across all the member countries of the content of courses, minimum length of study and training, and skills that a graduate in the field should possess. Sectoral directives were agreed for seven professions (doctors, dentists, nurses, midwives, vetinarians, pharmacists, and architects). However, it soon became clear that attempts to harmonize qualifications across member states were immensely complex and time-consuming and the task of agreeing a common curriculum for all EC professionals even in a single field began to look impossible. As McPherson states:

> The slow progress was partly of course due to the technical problems of harmonising what were often quite heterogeneous systems of training and practice. However, possibly the main reason was that free movement and "right of establishment" was a political goal for which few professions expressed much enthusiasm. Many of the negotiations showed a lack of urgency. (McPherson, 1988, p. 354)

With a growing awareness that harmonization of qualifications was an almost impossible task, as every country was required to agree to common content, length, and nature of training, in 1985 the European Commission introduced a new approach to cover other professions to which access is in some way restricted (or regulated) by the state, either by law or through a professional organization and which require at least three years university level training or equivalent. This approach was to be based on equivalence rather than harmonization and on the mutual recognition of qualifications. As Haaksman (1990) put it:

> in comparing study courses and professional skills, equivalence is more important than equality. In other words credentials evaluation should be based on mutual trust and respect rather than suspicion and pedantry ... the acquired skills should be comparable. The thing that matters is: equivalence in its most literal meaning. (Haaksman, 1990, p. 25)

This new approach took the form of the Directive 89/48/EC, entitled Mutual Recognition of Higher Education Diplomas, which applies to any regulated profession for which at least three years higher education level training is required and enables a professional from one member state to become a member of the equivalent profession in another member state without having to requalify. It does this by the principle of mutual recognition which enables suitably qualified professionals to have their qualifications recognized and to use the professional title. The Directive is intended to encourage the free movement of skilled labor around the EU by specifying the basis on which one member state will recognize, or at least evaluate, the professional qualifications awarded by another. It acknowledges that standards and content differ between countries and seeks to establish some equivalence between those trained in the countries of the EU.

This means that a psychologist qualified and recognized in one member state should be able to move and practice in another member state (subject to certain provisos). Implementation of the Directive is in most countries the responsibility of a government department, though unusually, in the UK this responsibility is given to the professional association, the British Psychological Society (BPS), which is the "competent authority" in terms of the Directive, and which therefore has responsibility for evaluating the equivalence of qualifications obtained in other countries.

2.01.5 THE ROLE OF THE EUROPEAN FEDERATION OF PROFESSIONAL PSYCHOLOGISTS" ASSOCIATIONS

EFPPA is a federation of 28 (1997) national European psychology associations which, among other things, has a central role in the development of standards in clinical psychology, in considerations of the implementation of the European Directives and in liaison work with Brussels. EFPPA was formed in 1981 at a time when the matters of mobility and mutual recognition were becoming more pressing, and psychologists and psychology associations of member states realized that a federation would provide a professionally and politically useful way to move forward and to begin to develop common policies in this area. Clinical psychologists, in particular, were faced with the growing prospects of mobility between countries in the EU, and the implications of the Treaty of Rome (see above).

The Federation plays an important role within the EU, and provides an important forum for exchanging information on patterns of education and training, legislation and regulation of psychologists, negotiating with politicians and policy makers, and influencing policy at a European level. It achieves much of its work through its task forces, which are small groups drawn from many European countries and usually focused on a particular topic or subfield within psychology (e.g., the task forces on clinical psychology, education and training, European legal matters). The member associations of EFPPA differ considerably in their nature: some are both scientific and professional organizations, others are traditional "learned societies," yet others operate effectively as trade unions, others are themselves federations of smaller associations within the same country. However, in 1997 all the national member associations include clinical psychologists. The fact that EFPPA is a federation of professional psychology associations means that much of its efforts are spent on clinical psychology and clinical psychologists, and is also the largest group of psychologists within Europe.

As mentioned above, EFPPA has had two task forces on education and training; the Task Force on Education and Training (which reported in 1990) produced the document entitled *Optimal standards for professional training for psychologists* (Appendix 1). In its report, the Task Force stated:

> EFPPA ... fully recognises that the content and form of training in professional psychology in each country will inevitably reflect its own legal, educational and professional traditions. However, to encourage international co-operation among professional psychologists and to assist member associations in their efforts to develop professional training and practice in their own countries, EFPPA is publishing the following statement which describes the optimal standards for training required for independent professional practice. (EFPPA, 1990, p. 1)

This was an attempt to produce a very broad framework with some elements, such as length of training, location in university base, theory and practical application, with which all member associations could agree, or at least would aspire towards. This document provided a very important beginning, although it has no legal status and is nothing more (nor less) than a rather broad goal. As may be seen (Appendix 1), the *Optimal standards* document states that preparation for autonomous practice in psychology should consist of at least two parts (theoretical and skills-based), lasting at least six years, based within a university and containing a research element. If it were possible to ensure these basic standards and level, recognition of qualifications between different countries would become more straightforward. However, in formulating these standards as "optimal" rather than "minimal," EFPPA recognizes that they are aspirational rather than required standards, and that EFPPA has a role in promoting the quality of professional education in clinical psychology across its member associations.

A further Task Force on the Evaluation of Curricula, formed in 1991, is carrying out a study of curriculum documentation in psychology and detailed interviews in different universities across Europe, in order to compare the different curricula in psychology of some universities in a number of European countries so as to increase an awareness and understanding of common and different areas in psychology education and training. The aims of the study are:

> to study the way in which psychology is taught in a range of institutions in different countries; to determine commonalities in training; to investigate the feasibility of laying down minimal standards for the accreditation of psychology curricula across Europe; to provide a basis for determining equivalence of psychology qualifications in European countries. (EFPPA, 1990, p. 1)

Newstead (1994) made a preliminary report of the work of this task force. As Newstead and Makinen (1997) have pointed out, there is considerable similarity and agreement between different countries over basic psychological theory. The differences appear particularly in the more professional domain, and in particular over how specialized the basic education is and what model of practice is adopted.

A third task force focuses specifically on clinical psychology (EFPPA, 1997) and aims to compare training and practice in clinical psychology across EFPPA member countries. This task force has carried out a questionnaire survey of clinical psychology practice and training across European countries, and has collaborated with task forces in the fields of health psychology and psychotherapy in an attempt to define the field.

2.01.6 TRAINING ROUTES IN THE DIFFERENT MEMBER STATES OF THE EUROPEAN UNION

In order to compare training routes in the different countries in Europe, and by definition therefore all the member states of the European Union, EFPPA set up a Task Force in Clinical Psychology in 1995. This Task Force has collected a substantial amount of data through a questionnaire exploring education and training routes and clinical psychology practice.

Most of the countries in the EU have considerable overlap in the content of the psychology curriculum, at least for the first three years of study, when universities in most countries provide a foundation in mainly theoretical areas of psychology. The Curriculum Task Force study (Newstead, 1994) identifies the percentage of time spent on different areas of psychology in the first three years of study (psychonomics, biological, social, developmental, individual differences, theory, methodology, applied psychology, project/thesis, professional placement). There was considerable agreement in the percentages of time spent on these different areas of the curriculum, except for one difference in that a few countries included a professional placement during this phase, while the majority left this until a later stage in the education and training. Some examples of different types of education and training routes will illustrate differences in patterns. It is in the second phase (the second two or three years) of professional qualification that there are substantial differences between European countries, specifically both in how specialized or general is the training, and the model of practice developed in the training period. There are also considerable differences in the nature and mode of professional training of clinical skills, and the nature of the internship or clinical placements undertaken.

2.01.6.1 The Nordic Countries (Denmark, Finland, Sweden)

Denmark has two universities offering education and training in clinical psychology, Finland has six, while Sweden has five. In these countries, the training period is at least five years, and the degree awarded is the degree of Candidate of Psychology, which in turn qualifies for postgraduate studies aiming at the degree of Licentiate and Doctor of Philosophy. In Finland every year about 180 students begin studying psychology as a main subject, making a total of around 1200 students of psychology in Finland (Saari, 1995). In all the Nordic countries postgraduate specialization is being developed, and increasingly required in order to practice as a clinical psychologist. In Finland, for example, specialist qualifications have existed since 1992 at the licentiate degree level in the areas of psychotherapy, health psychology, educational psychology, neuropsychology, and the psychology of work and organizations.

2.01.6.2 The United Kingdom

In the UK there are three phases of professional training: undergraduate (Bachelors), postgraduate (Masters and Doctorate) and postqualification. Clinical psychologists are qualified by obtaining first a Bachelors degree in psychology (of the breadth and standard determined by the BPS which accredits all courses); then a Doctorate in Clinical Psychology (DClinPsy), a professional doctorate taking three years and based in a university, which includes substantial internship or supervised practice in at least the three prescribed areas of clinical practice: adult mental health, child mental health, learning difficulties. The BPS has considerable control in determining the content and nature of psychology that qualifies for professional training. It defines the theoretical foundation, and also the professional training, through controlling the criteria for entry to the Register of Chartered Psychologists. This is a voluntary register which guarantees the qualifications of clinical psychologists and their undertaking to abide by the Code of Conduct and disciplinary procedures of the BPS.

There are about 18 universities where it is possible to study the final three years (doctoral level) to become a clinical psychologist, although there are over 100 universities where the undergraduate (Bachelors) degree in psychology is offered. It should be noted that in the UK, where psychology is one of the most popular subjects of study for school leavers aged

18, only 20% of those who achieve a Bachelors degree in psychology continue to become professional psychologists, and the number of training places in clinical psychology is determined by the National Health Service and is very restricted.

2.01.6.3 Ireland

Ireland follows the same pattern of training as the UK, though it awards two-year Masters degrees as the final qualification for clinical psychologists; there are two universities offering clinical psychology training in Ireland. Both in the UK and in Ireland there has been a model of qualification provided by the BPS which existed as a form of inservice training, studying for the qualification while working as an assistant clinical psychologist on a form of internship. This has been largely superseded by university-based courses.

2.01.6.4 Greece

In Greece, four universities offer degrees in psychology, and are developing two-year Masters programs for clinical psychology.

2.01.6.5 France

Training for clinical psychologists takes five years, progressing through the general Diploma which is awarded after two years, the License after the third year, and the DESS after the fifth year.

2.01.6.6 Spain

In Spain there are 12 faculties of psychology and a somewhat larger number of departments of psychology. Clinical psychology education consists of two cycles, a core general program lasting at least two years, and an advanced more specialist program lasting at least two years. This leads to a qualification to practice (*licendiado en psicologo*), although the license through the Colegio Oficial de Psicologos, which was formed in 1979, requires the five year degree in psychology and a period of supervised practice or internship for the title of *psicologo*. However, postgraduate training is being developed at both Magister and Doctoral level, with specializations which include clinical psychology (Belloch & Ollabaria, 1994; Prieto, 1991).

2.01.7 DEMOGRAPHIC FEATURES OF CLINICAL PSYCHOLOGISTS

In many countries psychology has become one of, if not, the most popular subjects to study at university. The majority of students studying psychology aspire to become clinical psychologists, and for this reason, in many European countries there is an oversupply of qualified practitioners. Some countries operate so-called *numerus clausus*, either at the start of the psychology study, or during the study. This controls the numbers in training. In the UK where specialist training in clinical psychology is funded by the National Health Service, where the vast majority of clinical psychologists work, the number of "trainee" posts is strictly limited and is planned according to staffing needs in the different regions of the country. In other countries where there is a tradition of predominantly private practice in clinical psychology, there are large numbers of qualified psychologists unable to find work. In all European countries, the ratio of female to male students is between 3:1 and 6:1; leaving the profession in danger of becoming an almost feminized profession in the future (Schorr & Saari, 1995). There has also been some difficulty in many countries in recruiting students from the range of ethnic groups represented in Europe's increasingly multiethnic population. This clearly has implications for the clinical treatment of different client groups.

2.01.8 FUTURE ISSUES

The main issue continues to be mobility of clinical psychologists and evaluation of equivalence. In the late 1990s each country retains its own distinctive pattern of qualification. There are fears that moves to a more common pattern of qualification could lead to a reduction in standards or length of study period; all countries have succeeded in developing the level and length of training of clinical psychology qualification, and moves which threaten this are keenly resisted. On the other hand, evaluation of equivalence is a complex and time-consuming task, and one which may ultimately militate against the spirit of the EU and its directives on free movement. The EU and its Commission has a policy on free movement, and therefore encourages any moves to facilitate this. Some professions have developed an informal "common platform" which is a standard of qualification that is agreed across all EU member states. This is under discussion within EFPPA, although any detailed standard, for example, a European diploma, will be extremely complex and controversial. It remains the case that clinical psychologists in some countries are almost a separate profession with separate education and training (from other psychologists) whereas in other countries they have a specialization from a generic psychology qualification. This major difference shows no sign of reducing.

A further issue concerns the teaching of professional issues and ethical issues. For some countries the qualification in clinical psychology is mainly academic, with little internship or professional experience, while in others the education required clearly specifies supervised practice through internship. As ethical issues become increasingly important, EFPPA has been concerned to develop education in this area, both through the use of exemplar ethical dilemmas, and through a common agreement among European countries that a certain coverage in this area is essential. This is difficult for those countries with a more purely academic approach to the professional training.

As new fields of practice in the health field evolve, in particular health psychology, there is a difficulty in defining the scope of the field of clinical psychology and its training. For example, in some countries it has been said that there may no longer be a field of clinical psychology, since there are strong moves towards health psychology and a focus on preventive work. These new areas, in particular health psychology, have worked out their own curriculum and pattern of education (and EFPPA's task force in this area has played a key role in this); however, the countries differ in the extent and nature of their specialisms within the health field. In one respect psychologists working within the health system could be said to be becoming more generic, while on the other hand there are increasing specializations within this field of work.

2.01.9 SUMMARY

This chapter has presented some broad commonalities and differences in education and training for clinical psychologists in the EU, and highlighted some of the implications for mobility and evaluation of equivalence. It has described some of the work and functions of EFPPA in this respect, and attempted to identify some present and future issues for clinical psychology educators in Europe. As the number of clinical psychologists increases, and the question of mobility becomes more pressing, there will be increasing attempts to develop more common frameworks and standards for education in this field; the challenge is to achieve a balance between allowing individual countries their own autonomy (subsidiarity, as it is called), and developing more common agreed frameworks within which individual countries develop their own patterns of education and training.

2.01.10 APPENDIX 1: EFPPA OPTIMAL STANDARDS FOR PROFESSIONAL TRAINING

EFPPA is an organization of national associations of psychology, the great majority of which represent both the scientific and the professional community.

At its General Assembly held in Luxembourg, September 14–16, 1990, EFPPA fully recognised that the content and form of training in professional psychology in each country will inevitably reflect its own legal, educational, and professional traditions.

However, to encourage international cooperation among professional psychologists and to assist member associations in their efforts to develop professional training and practice in their countries, EFPPA is publishing the following statement which describes the optimal standards for training required for autonomous professional practice.

Member associations which subscribe to these optimal standards are encouraged to work towards their achievement.

1. Preparation for autonomous practice of psychology comprises at least two components: a core programme and an advanced professional training in psychology.

The core programme is concerned with the knowledge and skills relating to psychology as a scientific discipline and is common to all branches and specialisms within psychology.

In the advanced component the student will acquire the knowledge and skills which are necessary for autonomous practice in a chosen field of professional psychology.

At all levels, one should aim at an optimal integration of the core program and the professional training.

2. Both components should be provided within a university or equivalent institute of higher education. In some countries with an established tradition of doing so, professional training may be provided by an affiliated professional school or training programme.

3. Together the two components should last at least six years, with the distribution between the core programme and the advanced training determined by each country, according to national circumstances. However, at least half the time should be devoted to the core programme.

4. All training should be accredited in ways acceptable to the national association.

5. Entry to independent or autonomous professional practice should be restricted to those who have completed both components.

6. The core programme in psychology should provide a broad introduction to psychology. It

should include the traditional subdisciplines in psychology, ranging from the biologically oriented approaches to those that are cognitively and socially oriented and including developmental, methodological, philosophical and ethical issues.

7. Professional training should cover the theoretical knowledge, skills, competencies, and research abilities required in the applied field of psychology as set out below:

Theoretical knowledge

(a) A variety of different theoretical models should be taught because no single model is able to cope satisfactorily with the range of problems that confront the professional psychologist.

(b) Theoretical models must be considered critically so that students are fully aware of their limitations as well as their advantages.

(c) Theoretical teaching needs to be integrated with practice.

Skills and competencies

(a) Substantial practical training is essential.

(b) Practical training should include experience in a variety of settings, methods and approaches. It should include work with individuals, groups, and organizations as well as practice in assessment, program development, and evaluation.

(c) Ethical considerations should be taken into account in training.

(d) Practical training needs to cover communication skills as well as the transmission of psychological skills to others through teaching, supervision and consultation.

(e) Like other aspects of training, practical skills need to be tested and examined.

Research training

(a) Because of the importance of evaluating practices as well as the need to develop new models, techniques, and intervention programs, an appreciation of the methods of applied research is essential.

(b) Students should have the experience of conducting an original and independent research project as part of their training (thesis).

8. National associations have an obligation to ensure that the education and training provided is consistent with their codes of practice and ethical standards.

9. It is recommended that national associations should encourage and promote, keep a record of, and if possible accredit, courses for the training of professional psychologists.

This document was produced by a Task Force on Education and Training in Psychology, consisting of representatives from among the member associations of EFPPA.

Psychological associations in the following countries are member associations of EFPPA (July 1990): Austria, Belgium, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Italy, Liechtenstein, Luxembourg, The Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland, UK.

Dated July 1990.

2.01.11 APPENDIX 2: MEMBER ASSOCIATIONS OF EFPPA (1997)

Austria, Belgium, Croatia, Cyprus, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Liechtenstein, Luxembourg, Malta, The Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, UK.

2.01.12 APPENDIX 3: BIBLIOGRAPHY

There follows a list of short articles with descriptions of psychology in countries of the European Union.

Austria

Guttmann, G. (1992). Austria. In V. S. Sexton & J. D. Hogan (Eds.), *International psychology: Views from around the world.* Lincoln, NE: University of Nebraska Press.

Belgium

d'Ydewalle, G. (1992). Belgium. In V. S. Sexton & J. D. Hogan (Eds.) *International psychology: Views from around the world.* Lincoln, NE: University of Nebraska Press.

Janssen, P. (1992). The psychological profession in Belgium. *News from EFPPA, 6*(2), 15–19.

Denmark

Foltved, P. (1992). Profile of Danish Psychological Association. *News from EFPPA, 6*(1), 16–18.

Foltved, P. (1995). Professional psychology in Denmark. In A. Schorr & S. Saari (Eds.), *Psychology in Europe* (pp. 15–22). Gottingen, Germany: Hogrefe.

Finland

Legislation on a psychologist's practice: Aims and achievements (1991). *News from EFPPA, 5*(4), 15–17.

Saari, S. (1991). Psychology training in Finland *News from EFPPA, 5*(1), 6–8.

Saari, S. (1992). Profile of Finnish psychology and the Union of Finnish Psychologists. *News from EFPPA, 6*(4), 22–26.

Saari, S. (1995a). The development of psychology in Finland. *News from EFPPA, 9*(2), 4–9.

Saari, S. (1995b). The legalisation of Finnish psychology. In A. Schorr & S. Saari (Eds.), *Psychology in Europe* (pp. 23–34). Gottingen, Germany: Hogrefe.

France

Castro, D., & Denis, M. (1996). Société

Française de Psychologie. *European Psychologist, 1*(2), 149.

Castro, D., & Engelhart, D. (1994). De la formation à la professionalisation. Processus de professionalisation de la psychologie appliquée en France. *European Review of Applied Psychology, 43*(2), 101–112.

Sanches, A. A. (1992). France. In V. S. Sexton & J. D. Hogan (Eds.), *International psychology: Views from around the world.* Lincoln, NE: University of Nebraska Press.

Trognon, A. (1987). France. In A. R. Gilgen & C. K. Gilgen (Eds.), *International handbook of psychology*. Westport, CT: Greenwood Press.

Germany

Graumann, C. F., & Metraux, A. (1992). Federal Republic of Germany. In V. S. Sexton & J. D. Hogan (Eds.), *International psychology: Views from around the world.* Lincoln, NE: University of Nebraska Press.

Schorr, A. (1995). German psychology after reunification. In A. Schorr & S. Saari (Eds.), *Psychology in Europe* (pp. 35–58). Gottingen, Germany: Hogrefe.

Greece

Georgas, J. (1995). Psychology in Greece. In A. Schorr & S. Saari (Eds.), *Psychology in Europe* (pp. 59–76). Gottingen, Germany: Hogrefe.

Velli, T. (1995). Profile of Greece. *News from EFPPA, 9*(1), 20.

Ireland

Brady, T., & McLoone, S. (1992). Ireland. In V. S. Sexton & J. D. Hogan (Eds.), *International psychology: Views from around the world.* Lincoln, NE: University of Nebraska Press.

Italy

Avallone, F. (1994). Training and professional spheres of psychology in France. *European Review of Applied Psychology, 43*(2), 113–121.

Comunian, A.L. (1992). Italy. In V. S. Sexton & J. D. Hogan (Eds.), *International psychology: Views from around the world.* Lincoln, NE: University of Nebraska Press.

Spaltro, E. (1993). Conflicts and developments of Italian psychology. *News from EFPPA, 7*(1), 17–21.

Netherlands

Dijkhuis, J. (1991). Psychologists in The Netherlands: Their training and professional regulations. *News from EFPPA, 5*(2 & 3), 20–22.

Jonkergouw, T. (1991). Profile of the NIP. *News from EFPPA, 5*(4), 23–25.

van Drunen, P. (1995). Professional psychology in the Netherlands. In A. Schorr & S. Saari (Eds.), *Psychology in Europe* (pp. 77–96). Gottingen, Germany: Hogrefe.

van der Ploeg, H. M. (1992). The Netherlands. In V. S. Sexton & J. D. Hogan (Eds.), *International psychology: Views from around the world.* Lincoln, NE: University of Nebraska Press.

Portugal

Pereira, F. (1994). Psychology in Portugal. *News from EFPPA, 8*(3), 22–24.

Pereira, F. (1995). Psychology in Portugal. In A. Schorr & S. Saari (Eds.), *Psychology in Europe* (pp. 97–102). Gottingen, German: Hogrefe.

Spain

Belloch, A., & Olabarria, B. (1994). Clinical psychology: Current status and future prospects. *Applied Psychology: An International Review, 43*(2), 193–211.

Carpentero, H. (1996). Psychology in Spain. *European Psychologist, 1*(1), 71–73.

Prieto, J. M. (1991). Studying psychology in Spain. *News from EFPPA, 5*(4), 18–22.

Tortosa, F. et al. Professional issues in Spanish psychology: Historical profile and present situation. *European Review of Applied Psychology, 43*(2), 123–141.

Yela, M. (1987). Spain. In A. R. Gilgen & C. K. Gilgen (Eds.), *International handbook of psychology*. Westport, CT: Greenwood Press.

Sweden

Persson, H. (1995). Psychology in Sweden. In A. Schorr & S. Saari (Eds.), *Psychology in Europe*. Gottingen, Germany: Hogrefe.

The Swedish Psychological Association (1992). *News from EFPPA, 6*(2), 23–26.

United Kingdom

BPS (1992). The undergraduate curriculum in psychology in the UK. *News from EFPPA, 6*(2), 10–14.

Lindsay, G., & Lunt, I. (1993). Developments in professional psychology in the UK. *News from EFPPA, 7*(2), 3–7.

Lunt, I. (1995a). The training and regulation of psychologists in the United Kingdom. *News from EFPPA, 9*(1), 17–19.

Lunt, I. (1995b). Demographic trends in professional psychology in the United Kingdom. In A. Schorr & S. Saari (Eds.), *Psychology in Europe* (pp. 149–162). Gottingen, Germany: Hogrefe.

Lunt, I., & Lindsay, G. (1993). Professional psychologists in the United Kingdom. *European Review of Applied Psychology, 43*, 91–98.

Newman, C. (1991). Profile of the BPS. *News from EFPPA, 5*(2 & 3), 23.

2.01.13 REFERENCES

Belloch, A., & Olabarria, B. (1994) Clinical psychology: Current status and future prospects. *Applied Psychology: An International Review, 43*(2), 193–211.

EFPPA (1990). *Optimal standards.* Stockholm, Sweden: EFPPA Secretariat

EFPPA (1991). *Task force on evaluation of curriculi: Terms of reference.* Stockholm, Sweden: EFPPA Secretariat.

EFPPA (1997). *Interim report to the General Assembly of EFPPA of Clinical Psychology Task Force.* Stockholm, Sweden: EFPPA Secretariat.

Haaksman, D. (1990). The European Directive and its implications for psychologists, their associations and EFPPA. *News from EFPPA, 4*(1), 24–27.

Kendall, P. C., & Norton-Ford, J. D. (1982). *Clinical psychology. Scientific and professional dimensions.* New York: Wiley.

Llewelyn, S. P. (1994). Assessment and therapy in clinical psychology. In P. Spurgeon, R. Davies, & A. Chapman (Eds.), *Elements of applied psychology* (pp. 115–132). Chur, Switzerland: Harwood.

Lunt, I. (1993). Educational psychology. The European dimension. *Newsletter of the BPS Scottish Division of Educational and Child Psychology, 1,* 1–6.

Lunt, I. (1994). Training and registration of professional psychologists in Europe. Keynote address at First Congress of Psychology, Malta. *News from EFPPA, 8* (3), 10–18.

Marzillier, J. & Hall, J. (1990). *What is clinical psychology?* (2nd ed.). Oxford, UK: Oxford University Press.

McPherson, F. (1988). Psychologists and the EEC. *The Psychologist, 9,* 353–355.

Newstead, S. (1994). The psychology curriculum and the training of psychologists in Europe. *News from EFPPA, 8*(4), 11–14.

Newstead, S., & Makinen, S. (1997). Psychology teaching in Europe. *European Psychologist, 2*(1), 3–10.

Peterson, D. R. (1991). Connection and disconnection of research and practice in the education of professional psychologists. *American Psychologist, 46*(4), 422–429.

Pilgrim, D., & Treacher, A. (1992). *Clinical psychology observed.* London: Routledge.

Prieto, J. M. (1991). Studying psychology in Spain. *News from EFPPA, 5*(4), 18–22.

Raimy, V. C. (Ed.) (1950). *Training in clinical psychology.* New York: Prentice-Hall.

Saari, S. (1995). The development of psychology in Finland. *News from EFPPA, 9*(2), 4–9.

Schorr, A., & Saari, S. (Eds.) (1995). *Psychology in Europe.* Gottingen, Germany: Hogrefe.

Spaltro, E. (1993). Conflicts and developments of Italian psychology. *News from EFPPA, 7*(1), 17–21.

Velli, T. (1995). Profile of Greece. Theoni Velli. *News from EFPPA, 9*(1), 20.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.02

History and Emerging Trends in Education and Training for Clinical Psychology in the United States

LYNN P. REHM

University of Houston, TX, USA

2.02.1 INTRODUCTION 13

2.02.2 THE BEGINNING YEARS 14

2.02.2.1 Science and Practice 14

2.02.2.2 Early Curricula 15

2.02.3 THE AFTERMATH OF WORLD WAR I 15

2.02.3.1 Science and Practice 15

2.02.4 WORLD WAR II AND ENTRY INTO THE MODERN ERA OF CLINICAL TRAINING 17

2.02.4.1 Boulder Conference 17

2.02.4.2 Professional Developments 18

2.02.4.3 Education and Training 19

2.02.4.4 The Vail Conference 20

2.02.5 CURRENT THEMES IN CLINICAL TRAINING 21

2.02.6 REFERENCES 23

2.02.1 INTRODUCTION

Clinical psychology celebrated its centennial in 1996. The evolution of the field, over this 100 years plus, has reflected the historical events in each era. In particular, World Wars I and II had major influences on its growth and directions. The stresses and strains that have occurred within psychology generally have steered the profession in one direction or another. Theoretical debates, professional associations, conferences, and the initiatives of individual psychologists have been some of the pushes and pulls that have altered the course of clinical psychology.

Many of the tensions in clinical psychology today can be traced back through history. The position of psychology among academic disciplines and among professional rivals has been a continuing battle of borders and territory. The relationship between clinical psychology and other areas of psychology has similarly been one of growth, interchange, overlap, and jostling for position. The scope of clinical psychology practice has expanded in competition with neighbors outside and inside psychology. The identity of clinical psychology has evolved from these conflicts and debates. The results of all of these influences are manifest in the themes of education and training in clinical psychology today.

In this chapter, I will attempt to review some of the major events influencing clinical psychology in different eras of development (Reisman, 1991)—its beginning years, between the world wars, growth years following World War II and the modern years of professional expansion. I will try to extract some of the themes and issues of education and training that intertwine with the events of those eras.

2.02.2 THE BEGINNING YEARS

The history of scientific psychology is traditionally dated to 1879 when Wilhelm Wundt began a doctoral program to train students in psychology at the University of Leipzig. Wundt had opened his psychology laboratory in 1875, the same year William James established a psychology demonstration laboratory at Harvard. Wundt and James trained many of the early pioneers of psychology in the USA.

The 1880s and 1890s were a period of openings and beginnings for psychology as it became established in American universities. G. Stanley Hall, a student of James, took a position at Johns Hopkins in 1881 and founded the second psychological laboratory there in 1883. In 1887 he founded the *American Journal of Psychology*. Hall became President of Clark University in 1889 and in 1892 he founded the American Psychological Association there and became its first president. James McKeen Cattell, a student of Wundt, took a position at the University of Pennsylvania in 1888 and founded the third psychological laboratory. In the same year, Joseph Jastrow, with a doctorate from Johns Hopkins, took a position at the University of Wisconsin. Cattel and Jastrow were the first psychologists to have the title Professor of Psychology.

The beginning of clinical psychology is traditionally dated to 1896 when Lightner Witmer opened a psychology clinic at the University of Pennsylvania. Witmer obtained his doctorate with Wundt in 1892 and in that year took over the psychological laboratory from Cattel. With the opening of the clinic, Witmer began a training program in the field he came to call "clinical psychology," although he preferred the term "psychological orthogenics." He advocated the "clinical method" of teaching and in 1897 began a summer course in the clinic based on learning from individual clinical cases, the first practicum in clinical psychology. Witmer also founded the journal *The Psychological Clinic* in 1907. Witmer's clinic was devoted to helping children with intellectual and learning difficulties and can be thought of as the first child guidance clinic.

The first decades of the twentieth century saw new service and teaching clinics and new applied training programs open. Carl Seashore and R. L. Sylvester opened a Psychological Clinic at the University of Iowa in 1908 and a similar clinic was opened at Clark University in 1909. Courses in clinical psychology were begun in the same year at the universities of Minnesota and Washington where clinics soon followed.

In 1909 the Juvenile Psychopathic Institute was formed in Chicago to aid the courts in dealing with delinquent children. It was staffed by William Healy, a physician, and Grace Fernald, a psychologist. This was the first true child guidance clinic and clinic outside of a university setting. Henry H. Goddard working as the director of a psychological research laboratory at the Institute for Backward Children in Vineland, New Jersey, began the first internship in psychology at Vineland in 1908 (Reisman, 1991).

Psychological application was largely limited to "mental testing." In the early years this consisted of simple tests of reaction time, and speed and accuracy measures on simple tasks. A number of batteries of such tests were proposed with great optimism that they could be used to predict everything from separating psychiatric cases from retarded cases, to selecting children for school programs, and to guiding people in the choice of vocations. These hopes were soon dashed by the results of studies that found little in the way of reliable differences between groups or of practical predictions that could be made with these measures. At the same time, there was a new direction in measurement as intellectual assessment was taking hold of the USA in the form of various translations and revisions of the French Binet–Simon tests. Intelligence testing became the primary focus of psychological application. While much was going on in the field of psychiatry, such as Emile Kraeplin's first psychiatric text in 1883, and Freud and Brewer's *Studies in Hysteria* in 1895, psychology expressed little interest in psychopathology or psychotherapy. Intellectual testing and remedial teaching were its assessments and interventions.

2.02.2.1 Science and Practice

The opening of Witmer's clinic was not widely celebrated in psychology. Many psychologists felt that application was premature and that psychology should be solely a basic science. Wundt had founded psychology as the scientific study of consciousness. He described the interests of his student Cattell in individual differences as *ganz amerikanishe* (wholly American) as Wundt was interested in universals of experience (Fowler, 1996). The American Psychological Association (APA) was founded to further psychological science, and in 1904 it tightened its membership criteria to require evidence of scientific contributions by new members. This focus on science derived from several concerns. The public image of psychology was mixed with the popular interest in mediums, clairvoyance, mind-reading, and other supernatural phenomena. Joseph Jastrow lamented in his 1901 APA presidential address that in the eyes of the public, "psychologist" was the equivalent of "spook chaser." He devoted great energy to educating the public about the science of psychology, for example, he took a battery of 15 early mental tests to a booth at the Chicago World's Fair in 1893 and showed the public the tests while collecting data on mental and sensory capacities.

William James' highly popular 1890 book, *Principles of Psychology*, recommended practicing psychologically healthy habits. The book was coolly received by academic psychology for the reasons that it was too applied, went beyond empirical fact, and contained too much philosophy and religion. In the academic world, where nearly all psychologists were employed, psychology was trying to establish itself as a separate science apart from philosophy, religion, education, or physiology. At Harvard, William James was Professor of Physiology, and then Professor of Philosophy before he briefly became a Professor of Psychology. While there was much disagreement within psychology about what type of science it should be, it was generally agreed that it was a science that should have its own separate identity. Those who tended toward application were seen as tarnishing this scientific image.

2.02.2.2 Early Curricula

In its early period then, psychology education overlapped with other fields, such as physiology, medicine, philosophy, and pedagogy. From England, Francis Galton added statistical methods such as the correlation coefficient to the scientific methodology of psychology. Graduate education in clinical psychology similarly borrowed from other fields. Witmer recognized an overlap of clinical psychology with other professions. In the 1904–1905 academic year, Witmer arranged for his students to take courses in psychiatry and neuropathology. Witmer's clinic and others like it used the consultation of physicians, social workers, educators, and others. It is noteworthy that psychology in its very beginnings included a doctoral graduate degree. In the USA, universities had only recently been bestowing doctoral degrees at all. Yale was the first to do this in 1861, followed by Harvard in 1872. Johns Hopkins opened in 1876 entirely as a graduate school.

Clinical psychology began to develop its own curricular identity. By 1908, five psychology courses in clinical psychology were offered at the University of Pennsylvania—three in developmental psychology, one in abnormal psychology, and one in mental and physical deficits in school children. From the beginning Witmer saw the clinic as central to the clinical method of teaching. Students were to observe and participate in a scientific study of cases. The identity of clinical psychology was further aided by such events as the publication in 1910 of *The Manual of Mental Tests* by Guy M. Whipple—the first standard, widely-used text on psychological testing. Clinical psychologists were largely mental testers, but with a research point of view. By and large they were employed in universities doing teaching and research, but perhaps also participating in clinics. A few psychologists began getting jobs in research laboratories in various institutions for the retarded or mentally ill, and a few were employed in schools. Psychology, as a new science, and clinical psychology, as a new applied component of psychology, were established and ready to take on new challenges.

2.02.3 THE AFTERMATH OF WORLD WAR I

In World War I the US army adopted psychological testing to process and place troops. The Army Alpha and Beta tests were developed and psychology began to invest itself in the field of mental testing. World War I suddenly and greatly expanded the need for testing and for trained personnel to administer tests. At about the same time, psychological assessment entered the schools on a larger scale with another increased demand for "mental testers." New tests were developed in the intellectual domain and testing expanded into the areas of personality assessment, physiological assessment of emotions, and vocational selection. Psychology struggled with who should give mental tests, as they were given by people with doctorates and masters degrees in psychology, but also by bachelor level psychometrists, teachers, and other professionals.

2.02.3.1 Science and Practice

Training for assessment became a contentious issue within psychology. In World War I, Robert Yerkes, President of the APA, headed up psychology's war efforts in the area of psychological testing. He was criticized by J. E. Wallace Wallin for recruiting experimental psychologists with little training in mental testing (Routh, 1994). Already in 1915, Whipple had proposed in a resolution adopted by APA "that this Association discourages the use of mental tests by those unqualified" (Reisman, 1991, p 115). The dispute about qualifications for testing erupted at the December 1917, APA convention at Carnegie Institute in Pittsburgh. The APA formed a committee on qualifications for testing experts, but it was not enough to please a group of clinical psychologists who met on December 29, 1917 and formed the American Association of Clinical Psychologists with Wallin as president, Leta Hollingsworth as secretary, and six members. This first association of clinical psychologists voted itself out of existence two years later when, with the efforts of Yerkes, they rejoined APA as the Clinical Section. APA agreed to attend to the interests of applied psychology.

The issues of clinical qualifications and APA's attention to practice continued to be sources of friction within APA. In 1921, APA reaffirmed its purpose to advance science and its requirement for research contributions for membership. It recommended certification for "consulting" (applied) psychologists and elimination of the Clinical Section. The certification program was not very successful because of its high standards and the Clinical Section was not eliminated. In 1924 the Clinical Section became the Clinical Division of APA and recommended standards for clinical education including a Ph.D. from an approved graduate school and four years of professional experience including one year in diagnosis. In 1926 APA opened its membership to associate members with a Ph.D. in psychology and employment in the field, but without the scientific publications necessary for full voting members. Within three years, associate members outnumbered full members.

In the 1930s an APA Committee on Standards of Training recommended standards of training for clinical psychology including a Ph.D. and one year of supervised training or a master's degree and one year of training for "assistant psychologists." Pleas were made to APA to set standards for psychological testing. State and local associations of clinical psychologists and consulting psychologists (those in private practice) arose, and in 1937 the Clinical Division of APA dissolved itself and recommended that its members join the new American Association of Applied Psychology with its sections in Clinical Psychology, Consulting Psychology, Educational Psychology, and a few others.

The curriculum of clinical psychology expanded in these decades as new theoretical approaches were introduced. The theories of Freud and his followers grew in influence in the United States at the same time that John D. Watson disseminated his ideas about behaviorism. Watson reinforced the notion that the study of animal learning could lead to applications of principles to humans. Gestalt and phenomenological approaches were being introduced as well. A 1930 textbook in abnormal psychology by H. L. Hollingsworth opened with a series of seven introductory chapters describing theoretical approaches to the topic before addressing any disorders. Psychiatry was going through a new biological era with the introduction of psychosurgery, narcosis therapy, and chemical inductions of seizures, and a parallel interest in biological assessments was popular in psychology. Witmer wrote of "psychochemistry" and offered a course in "Behavior and Metabolism." The nature vs. nurture debate raged with a focus on mental retardation. On the psychology side, nurture gained ground during this era. Psychologists argued about the structure of intelligence, and the debate between Spearman's general factor model and Thorndike's specific factor model was animated.

Psychology expanded in the settings of training, in research, in publication, and in areas of practice as well. In 1915 the first department of applied psychology was established at Carnegie Institute of Technology with Walter Van Dyke Bingham as its chair. Similar programs opened at other universities although the Carnegie Tech department was disbanded in 1924. In 1927 the Harvard Psychology Clinic was opened by psychiatrist Morton Prince who gave the *Journal of Abnormal Psychology* to APA, adding to APA's new publication initiatives. James McKeen Cattell founded the Psychological Corporation in 1921 to publish psychological tests. Child guidance clinics opened around the country with a growing concern about juvenile delinquency. Practice in schools, clinics, institutions, and individual consultation increased in the latter part of this era.

Psychology fought with organized psychiatry over whether psychologists could diagnose mental retardation. Intervention became a larger part of psychological practice, but conflict with psychiatry was largely averted by the convention that psychological interventions were educational in nature and not psychotherapy. Psychological intervention increased with play therapy, group therapy, and specific learning theory based interventions and psychotherapy in general gained credence in the eye of the public. Psychoanalysis was limited in the USA to physicians. Psychology was poised for a new leap into the future with World War II.

2.02.4 WORLD WAR II AND ENTRY INTO THE MODERN ERA OF CLINICAL TRAINING

Just as World War I, World War II had a major impact on clinical psychology. Again testing was employed in the military on a large scale and psychologists were called on to develop, validate, and administer the testing programs. Both the government and the public were coming to view psychotherapy and counseling favorably and these views moved psychology in the direction of training for intervention. In 1941 the National Research Council urged psychiatry and psychology to cooperate for the good of national defense. This action aided the recognition of psychology in the military and muted some of psychiatry's objection to psychologists entering into the psychotherapy role. Estimates of the need for psychotherapists and counselors in the military, veterans services, and state mental health programs led to government funding of graduate education in psychology. By the end of the war, the Veterans Administration (VA) was offering half-time salaries to graduate students in training to work in VA facilities, and the Public Health Service was providing grants to graduate training programs for staff, facilities improvement, and student fellowships in the mental health fields. New training programs were started to meet the need and take advantage of these support programs.

2.02.4.1 Boulder Conference

In 1945 the VA and National Institute of Mental Health (NIMH) asked APA to provide a list of graduate programs that provided appropriate clinical training to justify supporting their students. The APA responded by appointing a Committee on Training in Clinical Psychology chaired by David Shakow. In 1947 this committee issued what became known as the Shakow Report (see Shakow, 1969). The report was expanded in 1949 with some elaborations on suggested curricula for clinical training. The path of clinical psychology now led to one of the major milestones in the history of clinical psychology, the Boulder Conference.

In August 1949 the NIMH sponsored a week-long Conference on Training in Clinical Psychology in Boulder, Colorado. Victor Raimy, a faculty member of the host Department of Psychology at the University of Colorado, reported on the conference in the form of a book (Raimy, 1950). The conference was filled with discussion, debate, and disagreement. Most of the actions taken at the conference were to ratify portions of the Shakow Report. The articulation of a series of consensus positions for clinical psychology was a great accomplishment, even though positions adopted at the conference would continue to be debated for decades to come.

The conference report is best known for setting forth the scientist–practitioner model of training in which students are trained both in research and practice. Raimy (1950) reported that agreement was reached on a statement that "research competence is acquired only by doing research" (p. 88). There were, however, dissenters on both the experimentalist and clinician sides of the issue. The report is very explicit about the idea that the scientist–practitioner model is a model of training, and not a model of practice. It was anticipated that clinical psychologists would work in a variety of settings demanding a variety of roles. Private practice was discouraged by the conferees. Though some schools had dropped the master's degree, the conference recommended research requirements at both a predissertation and dissertation level. Strong training in statistics was recommended on the science side. On the practice side, training in psychotherapy was endorsed, though some lamented the lack of empirical support for psychotherapy. Raimy (p. 93) quotes one anonymous wag as saying "Psychotherapy is an unspecified technique applied to unspecified problems of unpredictable outcome. For this technique we recommend rigorous training." The ethical issue of offering unvalidated techniques was raised.

The doctoral degree was asserted as the level for entry into practice in psychology, affirming earlier recommendations by various groups. Training in testing was specifically discouraged for undergraduates. No decision was made on appropriate training at the master's level, a problem that psychology has never resolved. A four-year program was envisioned with an internship in the third year during which the dissertation topic would be formulated. Dissertations in any area of psychology were deemed acceptable, although research in personality was recommended. After some debate, the group also decided that multiple case studies would be acceptable as a dissertation, although there was no agreement whether a single case study would be appropriate.

Concern was expressed over the relationship between the graduate programs and the federal agencies that were funding them. All agreed that quality assessment and control was the role of the universities and that the agencies should not dictate curriculum. Some noted that the funding of students had some negative effects. The large time commitments that were required of students on VA stipends were making it difficult for students to finish their programs in a timely manner. In addition, the emphasis on VA work was reducing the opportunities to see women and children in practicum or internship. Indeed, psychology's traditional focus on children was in the process of change.

The conference participants agreed that a minimum curriculum should include four core areas:

(i) general psychology, including systematic, experimental, and social psychology;

(ii) clinical psychology and mental hygiene, including broad training in theory, method, and techniques;

(iii) field work with a variety of clinical problems at different levels of responsibility; and

(iv) research and methods of investigation.

However, beyond this minimum the conference favored a diversity of approaches among university departments to prevent the field from stagnating. The conference urged the field to move forward in the areas of accrediting of graduate programs, developing standards for internships, adopting an ethical code, and certification or licensing of individuals. Progress was soon made on these goals.

2.02.4.2 Professional Developments

Accreditation began in 1947 to accommodate the request of the VA to identify high-quality training programs. Forty-two programs that had achieved some level of initial approval were invited to the Boulder Conference. The growth of accredited programs can be seen by the numbers in Table 1. Provisional accreditation was given for a number of years to new programs that did not as yet have track records for their products. In the 1950s internships began to be accredited as well, though the VA was slow to acknowledge the APA accreditation system and for many years VA internships were largely considered outside the accreditation system.

APA adopted its first ethics code for clinical and consulting psychologists in 1951. It was expanded in 1952 to cover all professional psychology and published in 1953. Periodic revisions and amendments to the code occurred in 1958, 1963, 1968, 1977, 1979, 1981, and 1990. The current *Ethical Principles of Psychologists* was finalized in December 1992. The current code covers a wide range of activities in expanded areas of practice such as forensic activities and issues in education and research.

Connecticut was the first state to pass a certification law for psychologists in 1945. The law required a Ph.D. in psychology and one year of experience. Virginia followed in 1946 with a law requiring a Ph.D. and five years of experience. The first licensing law was passed in Georgia in 1951. APA officially favored certification over licensure. Within psychology, some felt that licensure was premature. Outside of psychology licensure was resisted by psychiatry which opposed independent practice by psychologists. It was not until 1977 that Missouri completed the cycle by becoming the last state to pass a psychology licensing law. Because the process was so prolonged, the American Board of Examiners in Professional Psychology (ABEPP) was formed in 1947 to provide a certificate or diploma to clinical, counseling, or industrial psychologists with a doctorate, and five years of experience, who passed a written and an oral examination.

Clinical psychology returned to APA in 1945 when APA expanded its mission statement to advance psychology both as a science and as a profession. With the change, the modern division structure was adopted. In the original plan for 19 divisions, there was to have been a Division 11, the Division of Abnormal Psychology and Psychotherapy, and a Division 12, the Division of Clinical Psychology. The former was to be for the researchers and the latter was to have been a direct continuation of the Clinical Section of the American Association of Applied Psychology. The members reconsidered, however, and formed only one division, 12, with the name the Division of Clinical and Abnormal Psychology. In 1955 it took its present name of the Division of Clinical Psychology. Other organizations pertinent to clinical psychology were formed in the ensuing years. During the early 1950s directors of clinical training programs began meeting at APA conventions and evolved into the Council of University Directors of Clinical Psychology, an organization that advances training by and for clinical directors. The Association of Psychology Internship Centers (APIC) was formed to aid and advance internship programs. APIC later became the Association of Psychology Postdoctoral and Internship Centers (APPIC).

The factions of science and practice continued their disputes in and out of APA in the modern eras. Other organizations were founded because of dissatisfaction with APA. For example, the Psychonomic Society was formed in 1959 as a home for basic researchers. Divisions proliferated in APA to represent ever-expanding and differentiating interests. Some practice divisions

Table 1 Number of APA fully and provisionally accredited doctoral programs by decade.

Year	*Clinical*	*Counseling*	*School*	*Professional/ scientific*
1950–Full	35			
1960–Full	56	26		
1970–Full/ Provisional	73/1	21/1		
1980–Full/ Provisional	112/5	21/6	11/5	3
1990–Full	163	59	38	3
1996–Full	188	68	44	7

formed to serve interests within the clinical psychology community, as with the Psychotherapy Division, 29, and later the Division of Independent Practice, 42. As practitioners gained dominance in APA, science and practice in APA explored division into separate "assemblies" in the late 1980s, but the constitutional amendment that would have split the organization was defeated by the members. APA central office was reorganized around science, practice, public interest, and education constituencies. Practice members were assessed an additional sum to advance professional interests. Another association, the American Psychological Society, was formed with an emphasis on research only. With these changes, relative peace has been maintained in APA in recent years.

In the 1960s, 1970s and 1980s the practice of clinical psychology continued to expand into new areas and to become more differentiated into specialties. Psychological practice became better recognized in law, and better funded by third-party sources. The establishment of community mental health centers in the 1960s opened another source of employment for psychologists and was a sign of the country's acceptance of psychiatric and psychological outpatient treatment. With licensing established, psychologists gained other forms of recognition through state legislation, such as freedom of choice for the public to choose psychologists under insurance plans, and hospital practice in some states. Psychology was recognized for coverage under federal health plans like Champus and Medicaid, and in most private mental health benefit plans.

2.02.4.3 Education and Training

After Boulder there were a number of other training conferences held to keep up with the changing world of practice. Among the most prominent were the following:

(i) Conference on Psychology and Mental Health. Stanford, CA, 1955 (Strother 1956).

(ii) Conference on Graduate Education in Psychology. Miami Beach, FL, 1958 (Roe, 1959).

(iii) Conference on the Professional Preparation of Clinical Psychologists. Chicago, IL 1965 (Hoch, Ross, & Winder, 1966).

(iv) Conference on Levels and Patterns of Professional Training in Psychology. Vail, CO, 1973 (Korman, 1974).

(v) National Conference and Self-Study: Quality in Professional Psychology Training. La Jolla, CA, 1981 (Callan, Peterson, & Stricker, 1986).

(vi) National Conference on the Future of Graduate Education and Training in Professional Psychology. Mission Bay San Diego, CA, 1986 (Bourg et al., 1987).

(vii) Conference on Graduate Training in Psychology, Salt Lake City, UT, 1986 (Bickman, 1987).

(viii) National Conference on Internship Training in Psychology. Gainesville, FL, 1987 (Belar et al., 1987).

(ix) The National Conference on Scientist-Practitioner Education and Training for the Professional Practice of Psychology. Gainesville, FL, 1990 (Assembly of Scientist-Practitioner Psychologists & Department of Clinical and Health Psychology, University of Florida, 1990).

(x) National Conference on Postdoctoral Training in Professional Psychology. Ann Arbor, MI, 1992 (Larsen et al., 1993).

(xi) National Conference on Postdoctoral Education and Training in Psychology. Norman, OK, 1994 (Education Directorate, APA, 1995).

Comparison among the recommendations of these conferences highlights the controversial issues in clinical psychology education and underscores our inability to reach consensus. Whereas the 1949 Boulder Conference had discouraged teaching testing to undergraduates, the 1965 Chicago Conference recommended that practical training for undergraduates including psychological testing and interviewing would be appropriate. In 1986 the Utah Conference reiterated the Boulder idea that undergraduate education was not for practice in psychology. Master's training has been the topic of recommendations by several conferences. The Chicago Conference took a stand against master's training in clinical psychology, but recognized that specific practical skills might be taught to persons who would employ nonpsychological titles at the master's level. The Vail Conference in 1973 was much more favorable towards master's training for practice under supervision. A core curriculum was recommended in outline by the Boulder Conference, and the importance of a core was reiterated by the early conferences thereafter. The 1965 Chicago Conference opined that "the notion of a core curriculum is no longer possible" (Hoch et al., 1966, p. 47). Utah favored a core, and cited the accreditation standards list in 1986, but wanted to leave specification up to departments. The National Council of Schools of Psychology focused their 1986 Mission Bay Conference on issues of a core, and developed a detailed outline of what it should be. Some of the bolder recommendations of the conferences were not acted on by the field. For example, the 1987 Gainesville Internship Conference recommended a two-year internship, and the 1986 Utah Conference recommended that by 1995 free-standing professional schools would no longer be eligible for APA accreditation until they affiliated with a regionally accredited university.

2.02.4.4 The Vail Conference

The Vail Conference deserves special comment because it was a clear milestone for clinical psychology in marking the acceptance of professional schools. The 1973 Conference on Levels and Patterns of Professional Training in Psychology was convened because of growing concerns that the Boulder model of scientist–practitioner training was not meeting the needs of a changing professional world, of society, and of groups of students underrepresented in psychology. By this time, several professional schools and programs in psychology had been initiated. The program at Adelphi founded by Gordon Derner in 1951 was the precursor of this movement. His "scholar–practitioner" program was intended to provide students with a knowledge base that would let them critically read research, but not train them in actually doing research. This program was accredited by the APA in 1957 on appeal. In 1968 the University of Illinois initiated a Psy.D. track in their clinical program under the direction of Donald Peterson. Students in this track took the same course work as the Ph.D. students, but did more in-depth clinical work rather than the usual thesis and dissertation research. In 1969, the California School of Professional Psychology (CSPP) was founded with free-standing campuses in major cities of California. The CSPP, offering a Ph.D., was founded by members of the California Psychological Association who felt the existing programs in the state were not meeting the needs of the public. By the time of the Vail Conference, the National Council of Schools of Professional Psychology (NCSPP) (later to become the National Council of Schools and Programs of Professional Psychology, NCSPPP) had been formed to advance the interests of the professional schools.

The Vail Conference endorsed the new trend in professional training while acknowledging that it should expand upon and not replace scientist–practitioner training and should not exclude the science base of the profession. The conference urged all training programs to become more attuned to the public needs that they are serving, and to incorporate diversity in their training curricula, practicum, target populations, students, and faculty. Affirmative action for women and minorities was urged. The conference also took a strong stand on professional training at multiple levels—Associate of Arts, bachelor's, master's, and doctoral. The master's degree was posed at the journeyman level of practice with the doctorate representing the master level of practice. The conference urged APA to recognize persons with master's degrees as entitled to the title of "psychologist" and to grant full membership to them. The recommendations of the conference fell short of endorsing licensure and independent practice at the master's level. Vail illustrates the fact that training conferences do better at endorsing trends already underway in the field (e.g., professional schools) and do less well at initiating new trends (e.g., official recognition of master's level psychologists).

Finally, the Vail conference envisioned a career ladder or lattice. Psychologists at all levels were to participate in continuing educational development to expand and update their skills and to move up the career ladder. Training programs were encouraged to offer these professional development curricula in the form of flexible, small units of education in ways that would count for career and degree advancement. They envisioned people accumulating continuing education that would lead to the next professional degree up the ladder as well as to new skills across the lattice.

2.02.5 CURRENT THEMES IN CLINICAL TRAINING

Clinical psychology's century of history has bought us a long way from the laboratories, summer practica, and applied teaching clinics of our early days. In 1996 there were about 84 000 professional psychologists licensed in the USA. According to the recent practice analysis conducted by the Association of State and Provincial Psychology Boards, about half of these are clinical psychologists (Greenberg, Smith, & Muenzen, 1995). The 188 accredited doctoral programs in clinical psychology are varied in location (department of psychology, school of education, department of psychiatry, free-standing professional school, professional school within a university, etc.), degree offered (Ph.D., Ed.D., or Psy.D.), size (number of students, faculty, and their ratio), curriculum (required and elective courses, core content, theories and methods emphasized), orientation to science and practice (research and practice experience required), facilities (classrooms, laboratories, and clinics), faculty (orientation, training, areas of expertise, clinical and research activities), and students (admission criteria, goals and orientation). These programs exist in a changing world of public perceptions and needs, market conditions, and the rapid evolution of the health care system. A number of themes go into this mix.

Defining a core curriculum has always been a problem in clinical psychology. We have always been reluctant to be too specific in identifying a core for fear of impinging on the academic freedom and innovation of individual programs. The result, however, is that we are not able to assure the public that a clinical psychologist can be guaranteed to have some minimal competencies. Programs often go in the other direction by trying to be overly inclusive in what they require. The Boulder Conference envisioned a four-year program. Today's students are taking six to seven years or more to complete a doctorate. There are constant pressures to add to the curriculum as new topics become increasingly relevant. How will we include psychopharmacology, mental health law, neuropsychological assessment, human sexuality, cultural diversity, health economics, health administration, and other worthy topics in an already crowded curriculum? In addition, there are pressures to extend training in other ways. Should the internship be one or two years? Most states require two years of experience for licensure. Should not both of these be organized and accredited training experiences? Should the internship come before or after the awarding of the doctorate? Perhaps training programs should maintain quality control over the complete educational experience of their students, including the internship. Perhaps internships should be autonomous in their quality control. An argument can be made that, to parallel other mental health professions, we should award the academic degree and follow it with a residency. The economic consideration that a degree facilitates third-party reimbursement cannot be entirely ignored.

Is the length of training, both academic and internship, justified by the economics of the cost, job market, and income potential? Some believe we are overproducing psychologists. Increasingly clinical psychology faces competition from other mental health providers. Master's level providers are pouring in to the field, providing the same functions as clinical psychologists at lower cost. On the other hand, new markets seem to be opening to clinical psychologists. Health psychology has moved into virtually every medical specialty area. Clinical neuropsychology, sports psychology, the psychology of health promotion, forensic and correctional psychology, outcomes evaluation in managed care, etc., are new and expanding areas of clinical psychological practice. Will the new markets make demand equal to supply?

Psychology in the USA has never been able to solve the problem of multiple levels of training. The traditional position of the APA is that in the US, the doctorate is the entry level for independent practice and people trained at the master's level need supervision. This position has been asserted back to our earliest years, though the reality of education is that we have trained at multiple levels for practice. For the most part, the USA is alone in the doctoral level policy. In most countries of the world, the master's level is the entry practice level for psychology and the doctorate is a specialized teaching and research degree. Even our neighbor Canada seems to be recognizing master's level registration for practice in more of its provinces. Master's level practitioners are flooding the mental health job market under new titles such as "Professional Counselor," "Marriage and Family Therapist," "Substance Abuse Counselor," etc. We in psychology, and particularly in clinical psychology, are providing much of the training for master's level practice, but then we give away regulation and control to newly formed and newly defined disciplines with separate licensing boards and professional organizations. The problem has been recognized for decades and many recommendations have been made (e.g., some bold suggestions by the Vail Conference), but no official position on integrating the field seems to be attainable. Increasingly psychology is also training people at the bachelor's and even the AA level for positions in mental health, such as psychiatric technicians and case managers. No training model has appeared on our horizon to integrate our increasingly multiple levels of training.

How does training in clinical psychology fit into the increasingly large list of specialty and proficiency designations we see in psychology? Where do we fit *vis-à-vis* health psychology—is it a subset of clinical, a subspecialty, a neighbor specialty, or a number of specialties among which clinical is one? Do the traditional specialty designations continue to make sense? Do clinical and counseling psychology overlap more than they differ? How is the public served if we offer services to children under the labels child clinical psychologists, child psychologists, applied developmental psychologists, pediatric psychologists, family psychologists, school psychologists, and clinical school psychologists? Currently within APA, the Council for the Recognition of Specialties and Proficiencies in Professional Psychology is charged with sorting out this terminological mess. The Division of Clinical Psychology has asserted that clinical is a broad general practice specialty or a field, and other health service specialties may be tracks within and/or postdoctoral extensions of doctoral training in clinical psychology. Professional psychology must face this problem with a taxonomy that makes sense for training programs and for the public's ability to know who we are and what we do. Marking out the domains and boundaries of specialties would also help to define core curricula for the specialties.

A dramatic change is underway in our health delivery system in the USA. The industrialization of health care is an apt description for the reorganization of providers, the management of care, the imposition of guidelines, the review of utilization and quality, the measurement of outcomes in care delivery, and the other changes that are occurring daily. "Behavioral health" is being offered by large corporations on a capitated basis to a large number of industries, buying for large numbers of consumers. How is education in clinical psychology responding to these changes? We need to educate ourselves on the changes taking place. We need to train our students in the new systems of care delivery and offer practica in the new health care settings. Health care delivery is becoming more integrated into primary care settings with specialists on call for "one stop shopping." Mental health care is, therefore, returning to a more interdisciplinary practice, and we need to train our students to work in interprofessional settings on teams with a variety of professionals. We need to draw on our measurement skills to train students to implement new methods of outcomes assessment for our services and the services of others. We need to look at our intervention skill training and train students in the delivery of empirically effective methods in cost-efficient modalities. The Division of Clinical Psychology has initiated a project for identifying "Empirically Supported Treatments" (Task Force on the Promotion and Dissemination of Psychological Procedures, 1995). The project has been much debated (for example, a special series of articles is in press in the *Journal of Consulting and Clinical Psychology*), but was intended to call attention within psychology to the need to establish and present to various audiences an evidence base for our interventions. We need to move more of our interventions research out into the field to include evaluations of the effectiveness of treatment in actual care delivery settings. We should give more attention to studying different levels of intervention for problems of different severities. Prevention services should become more important in large-scale coordinated health care delivery systems.

Psychiatry has entered a new era of "medicalization" or "biological psychiatry." The "decade of the brain" has focused research on the biological substrata of behavior and behavior disorders. Psychology education is paralleling this trend. Our students need to be taught the basics of brain and nervous system structure, physiology, chemistry, and function. "Biological bases of behavior," as an element of our curriculum, takes on larger proportions in our core curriculum. Psychopharmacology is now recommended as a content area that should be included in all health service provider curricula in psychology.

A possible extension of this biological trend is the much debated (e.g., special feature of seven articles in the March, 1996 *American Psychologist*, Current Issues: Prescription Privileges) push for prescription privileges for psychologists. On the side favoring prescription privileges is the argument that psychologists should be able to provide full services to consumers and not have to rely on referals for medication. Psychologists may be better able to integrate pharmacotherapy into a psychological perspective and may choose more often not to medicate. Underserved populations may be better served if psychologists can expand their service options. Training seems to be feasible in a relatively short period of time. An increased market share for psychologists is also an important motivation for the move in this direction. On the side opposing prescription privileges for psychologists is the argument that training for prescription privileges fundamentally distorts the educational process away from our traditional expertise in behavior, affect, and cognition. Our traditional expertise is complementary to psychiatry's biological expertise as we have taken over preeminence in the field of psychotherapy. Why compete for pharmacotherapy where we will never have the equivalent basic and applied training to physicians? The debate is evolving in our journals and conventions and in our training programs. It will be resolved in the state legislatures.

The combining of science and practice in the education of clinical psychologists has been a central theme of the last 100 years. Today we recognize a broad spectrum of positions on the issue, with "scholar–practitioner" programs at one end and "clinical scientist" programs at the other. We like to boast that our training in science makes us distinct from other mental health professionals, but empirical demonstration of the difference is difficult. The best mix of science and practice in our training programs will depend on our goals, our preferences, and the evidence of our success. Hopefully our science will continue to inform our practice and help to resolve the many dilemmas we will face in our second century.

2.02.6 REFERENCES

Assembly of Scientist–Practitioner Psychologists & Department of Clinical and Health Psychology, University of Florida (1990). *The National Conference on Scientist-Practitioner Education and Training for the Professional Practice of Psychology*. Gainesville, FL: Author.

Belar, C. D., Bieliauskas, L. A., Larsen, K. G., Mensh, I. N., Poey, K., & Roehlke, H. J. (1987). *Proceedings of the national conference on internship training in psychology*. Baton Rouge, LA: Land & Land Printers.

Bickman, L. (1987). Proceedings of the National Conference on Graduate Education in Psychology, University of Utah, June 13–19, 1987 [special issue]. *American Psychologist*, *42*(12), 1041–1047.

Bourg, E. F., Bent, R. J., Callan, J. E., Jones, N. F., McHolland, J., & Stricker, G. (1987). *Standards and evaluation in the education and training of professional psychologists: Knowledge, attitudes, and skills*. Norman, OK: Transcript Press.

Callan, J. E., Peterson, D. R., & Stricker, G. (1986). *Quality in professional psychology training: National conference and self-study*. LaJolla, CA: National Council of Schools of Professional Psychology.

Education Directorate, American Psychological Association (1995). *Education and training beyond the doctoral degree: Proceedings of the American Psychological Association national conference on postdoctoral education and training in psychology*. Madison, CT: International Universities Press.

Fowler, R. D. (1996, August). *APA and clinical psychology: Past, present and future*. Invited address at the meeting of the American Psychological Association, Toronto, ON.

Greenberg, S., Smith, I. L., & Muenzen, P. M. (1995). *Study of the practice of licensed psychologists in the United States and Canada*. New York: Professional Examination Service.

Hoch, E. L., Ross, A. O., & Winder, C. L. (Eds.) (1966). *Professional preparation of clinical psychologists*. Washington, DC: American Psychological Association.

Hollingsworth, H. L. (1930). *Abnormal psychology*. New York: Ronald Press Company.

Korman, M. (1974). National conference on levels and patterns of professional training in psychology: The major themes. *American Psychologist*, *29*, 441–449.

Larsen, K. G., Belar, C. D., Bieliauskas, L. A., Klepac, R. K., Stigall, T. T., & Zimet, C. N. (1993). *Proceedings: National conference on postdoctoral training in professional psychology*. Baton Rouge, LA: Land & Land Printers.

Raimy, V. (1950). *Training in clinical psychology*. New York: Prentice Hall.

Reisman, J. M. (1991). *A history of clinical psychology* (2nd ed.). New York: Hemisphere Publications.

Report of the Committee on Training in Clinical Psychology of the American Psychological Association Submitted at the Detroit Meeting of the American Psychological Association, September 9–13 (1947). Recommended graduate training program in clinical psychology. *American Psychologist*, *2*, 539–558.

Roe, A. (Ed.) (1959). *Graduate education in psychology* (Report of the Conference on Graduate Education in Psychology sponsored by the Education and Training Board of APA, held at Miami Beach, FL, November 29—December 7, 1958). Washington, DC: American Psychological Association.

Routh, D. K. (1994). *Clinical psychology since 1917: Science, practice and organization*. New York: Plenum Press.

Shakow, D. (1969). *Clinical psychology as a science and profession: A 40-year odyssey*. Chicago, IL: Aldine Publishing.

Strother, C. R. (1956). *Psychology and mental health*. Washington, DC: American Psychological Association.

Task Force on the Promotion and Dissemination of Psychological Procedures (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist*, *48*, 3–23.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.03

Internship Education and Training in Professional Psychology

SANFORD. L. PEDERSON
Wright State University, Dayton, OH, USA

2.03.1 INTRODUCTION 25
2.03.2 CONCEPTUAL ORIGINS OF THE NEED FOR INTERNSHIP TRAINING 27
2.03.3 CONFERENCES ADDRESSING INTERNSHIP EDUCATION AND TRAINING 28
2.03.3.1 Boulder 28
2.03.3.2 Sanford 28
2.03.3.3 Miami 28
2.03.3.4 Gainesville 29
2.03.3.5 Ann Arbor 29
2.03.3.6 Orlando 30
2.03.4 RESEARCH ON INTERNSHIP EDUCATION AND TRAINING IN CLINICAL PSYCHOLOGY IN THE LATE 1990s 31
2.03.5 FOUNDING AND DEVELOPMENT OF THE ASSOCIATION OF PSYCHOLOGY POSTDOCTORAL AND INTERNSHIP CENTERS 31
2.03.5.1 APPIC Membership Criteria 32
2.03.6 DATA ON NUMBERS OF PROGRAMS AND INTERNS IN THE USA AND CANADA IN 1997 33
2.03.7 CONTINUING ISSUES UNDER DISCUSSION 33
2.03.7.1 Prerequisites for Internship Appointment 33
2.03.7.2 Uniform Requirement and Uniform Applications 33
2.03.7.3 Uniform Notification Day and Computer Matching 34
2.03.8 FUNDING FOR INTERNSHIP EDUCATION AND TRAINING 34
2.03.8.1 Present and Future Prospects 34
2.03.9 FUTURE DIRECTIONS 34
2.03.10 SUMMARY 35
2.03.11 REFERENCES 35

2.03.1 INTRODUCTION

The internship is the final required formally organized training experience before the awarding of the doctoral degree. It is intended to be a clinical experience that is fully integrated with the doctoral program experience. It is the culmination of three to six years of postgraduate education, and is intended to be the time when all that has been learned is integrated into a new understanding and a new sophistication in clinical thinking and clinical skills.

The internship in clinical psychology has generally been considered the "capstone" of professional training. It is the training experience where the science and practice of psychology, and the learning that has occurred in both the classroom and the clinic, are integrated in an intensive and extensive experience that has as its goal the completion of the fully qualified clinical psychologist. The internship in clinical psychology is typically a full-time experience where the student provides a variety of psychological services under the supervision of a licensed psychologist during the final year of training prior to the awarding of the doctoral degree.

The idea of an "internship" is almost as old as clinical psychology itself. H. H. Goddard developed the first formal internship program at the Vineland Training School in 1908, only 12 years after Lightner Witmer opened his psychology clinic at the University of Pennsylvania in 1896. However, internship training as a formal part of training in clinical psychology did not begin for many years, when formal training programs in clinical psychology were first described (American Psychological Association [APA], 1945). The APA, through the leadership of David Shakow, began the development of criteria for internships in clinical psychology at the same time that it began developing criteria for doctoral programs in psychology (APA, 1947). The APA began formally accrediting independent internships in clinical psychology in 1955. Today there are over 500 internship programs in the USA and Canada (Hall & Cantrell, 1996).

It is generally accepted that the internship is a prerequisite for the professional practice of clinical psychology. The internship is that part of the training of the clinical psychologist where the professional identity is finalized and the new clinician begins to develop a stronger sense of career goals. Given the multitude of training and developmental tasks to be accomplished during the internship, one year is actually quite a brief period of time. In fact, many newer models suggest that an additional postdoctoral training year is either desirable or necessary (Belar et al., 1989; Larsen et al., 1993; Strother, 1956).

Internships have been developed in a wide variety of settings, ranging from the traditional VA hospital and medical school to the university counseling center and the community mental health center. There are even some unusual internships in places such as employee assistance programs. Many newly developing internships are consortia, with several different agencies, often of different types, banding together to pool resources to provide an internship where none of the agencies could do so individually. Consortia also have an advantage of providing a wide range of types of placements for interns, the opportunity to be exposed to a variety of clinical environments. However, consortia are sometimes difficult to coordinate, can suffer from an uneven quality of supervision, can lack a consistent training model and focus, provide fewer opportunities for interactions with other interns, and increase the possibility of an intern "getting lost" in a large system where the training director does not have routine weekly or more frequent contact with interns.

Blom (1995) developed a taxonomy of types of internship settings classified by the range of patients served, the continuum of care provided, and the variety of interventions offered. The first type he describes, class I—comprehensive care settings, serves a wide range of patients across a full continuum of care in a variety of contexts with a variety of interventions. These include veterans affairs medical centers, military medical centers, medical school teaching hospitals, large public and private multispecialty medical centers, and certain consortia that include the full range of types of care settings. The second type, class II—limited care settings, does not offer the full range of opportunities across the three dimensions used. These include psychiatric hospitals, comprehensive community mental health centers, private general hospitals, small public and private hospitals, and children's hospitals. The third type, class III—specialty care settings, focus on a particular class of patients or interventions. These include outpatient mental health clinics, head injury rehabilitation centers, substance abuse programs, and university counseling centers. This taxonomy can be useful when a program of study for an individual is being considered, in light of specific career goals.

While the one year full-time internship remains the most common internship experience, a number of alternative models have also been developed. As new models of doctoral education have been developed, some of the traditional internship concepts have been modified to meet the needs of changing doctoral training programs. Some of the forces that have driven the changes in internship models include the changing demographics of psychology doctoral students and the changing economics of graduate and professional training in psychology, training models that call for greater integration of the classroom and clinic throughout graduate training, and the continuing development of specialty areas in clinical psychology requiring specialized training beyond the internship. These forces have taken what was once a relatively unitary model of internship training and fractionated it into a multiplicity of models.

The half-time two-year internship has been considered an acceptable alternative to the one-year full-time internship since the APA began accrediting independent internships in 1955. It is often seen as a less desirable alternative selected only because of specific practical constraints, such as child care and other family issues, or the lack of availability of full-time internships in cases where students do not see themselves as having geographic mobility. The negative aspects of the half-time internship include the lack of a full-time immersion experience in clinical work, the inability to be present for half of the work of the multidisciplinary team, and the lesser degree of professional socialization and identification.

Additionally, many of the available (but not APA-accredited) half-time internships are without stipends. A wide range of groups and conferences have endorsed the concept that all interns should be compensated at a level appropriate to their training. Kohut and Cooney (1993) found that only 54% of part-time interns received stipends, while 98% of full-time interns received stipends.

In spite of these disadvantages, there is a continuing call for greater numbers of half-time internships in clinical psychology. This is likely due to the need for larger numbers of internship opportunities in specific geographic areas which allow the intern to have other responsibilities at the same time as the internship. However, for internship sites there are few advantages, and indeed a number of disadvantages, to offering half-time positions as opposed to full-time positions.

Some professional schools of psychology have adopted a model of two years of two-thirds time internship integrated with the final year of coursework and the dissertation (Dorken & Cummings, 1977). This model does allow for greater integration of classroom work and clinic work, but also suffers from many of the same disadvantages of the two-year half-time internship model.

Captive internships, internships that are only for students enrolled in a particular doctoral program and are integrated with that doctoral program, may take any of the forms described earlier. The advantages and disadvantages of captive internships have been extensively discussed in the literature. In summary, while the captive internship allows for a consistent educational philosophy and set of quality control standards (Abrams & Brabender, 1993), it also does not provide interns with exposure to other approaches and alternative role models (Fraser & Bent, 1994).

While the primary source of information about internships available in the USA and Canada is the *APPIC Directory* (Hall & Cantrell, 1996), there are a number of specialty listings also available. Cripe (1995) lists doctoral programs, internships, and postdoctoral fellowships in clinical neuropsychology, while Blanchard (1994) lists programs with significant research opportunities for clinical psychology interns. Other specialized societies also offer directories of internship opportunities in their specialty.

There is a wide variety of resources available to aid students in their search for an internship in clinical psychology. Megargee (1997) is currently the most comprehensive guide for the prospective intern, covering academic, practical, and strategic aspects of applying to, interviewing for, and obtaining an internship in clinical psychology. Another text is Levinger and Schefres (1995). Numerous journal articles have been written to advise the potential intern, such as Lopez, Oehlert, and Moberly (1996), Mellott, Arden, and Cho (1996), Plante (1996), and Stewart and Stewart (1996a, 1996b). Workshops are also frequently available at major psychological organization meetings, such as Pederson (1996).

2.03.2 CONCEPTUAL ORIGINS OF THE NEED FOR INTERNSHIP TRAINING

While internships in clinical psychology had existed for many years, the first clear articulation of the need for internship training as a routine part of professional preparation in clinical psychology was in the Report of the Subcommittee on Graduate Internship Training in Psychology (APA, 1945). This committee, chaired by David Shakow, described the characteristics of the clinical psychology intern and the clinical psychology internship program, as well as its relationship to the doctoral degree program.

Early in their report the committee observed that, "It is impossible to consider the problem of internships separately from the other aspects of preparation for clinical psychological practice" (p. 245). So from the earliest beginnings of formalized internship programs in clinical psychology, we see an emphasis on the integration of the work done in the classroom and laboratory with the work done in the clinic. This early model of the internship placed it in the third year of graduate education, after a one-year practicum experience in the second year, and prior to a return to the university to complete the dissertation, partake in cross-disciplinary seminars, and complete any additional needed courses. It should be noted that use of the term "residence appointment" referred to the expectation that the intern would live on the grounds of the internship institution. This was felt to be important, since

"It permits twenty-four hour residency in an atmosphere permeated with a psychology of a living kind and appears to have more potential for the maturation of the personality than almost any other experience ordinarily available to the psychologist" (p. 260). The report closes with a repeated discussion of the need for and ways in which the internship and the academic program be integrated, and some specific recommendations such as the establishment of a standardized internship certificate, and the development of standards for internship centers. It is of historical interest to note that this early report raised the question of whether the Ph.D. or the Psy.D. was the more appropriate degree for the clinical psychologist.

2.03.3 CONFERENCES ADDRESSING INTERNSHIP EDUCATION AND TRAINING

While many conferences have been held that examined the education and training of the clinical psychologist, not all have focused clearly on issues of internship training. The conferences at Chicago (Hoch, Ross, & Winder, 1966), Vail (Korman, 1976), and Utah (Bickman & Ellis, 1990) have focused primarily on issues of curriculum and design of the doctoral degree program, with relatively less emphasis on the internship aspect of clinical psychology education and training. A brief review of those conferences having a significant focus on internship training in clinical psychology follows.

2.03.3.1 Boulder

Much of what may be considered the traditional model of internship training originated at the Boulder Conference. The conference held at Boulder in 1949 (Raimy, 1950) was the first time that a large group of psychologists formally gathered together to answer questions about how clinical psychologists should be trained. The growing need for mental health services, particularly that prompted by the return of soldiers and sailors from World War II, prompted the USA Veterans Administration and the APA to begin to develop formal models for the education and training of clinical psychologists.

The internship was originally conceptualized as a predoctoral experience where the intern was immersed in the activities of a clinical service setting for a full year. It was intended to differ from a clerkship or practicum in several ways. First and foremost, it was a full-time year-long immersion in clinical service activities. It would take place in a multidisciplinary setting, and would be at a training site that was not a part of the student's doctoral degree program. It would be focused on the development of the skill necessary for semi-independent practice through a gradual increase in responsibility and case management. Finally, the internship was intended to be a socialization in the values of the professional practice of clinical psychology.

The traditional model of internship training, as an integral part of the Boulder model of doctoral education, consists of full-time students in a program with primarily full-time faculty. In the Boulder plan, the internship was to be taken in the third year of study, with the student returning to the university in the fourth year to complete the dissertation. This aspect, the third-year internship, has changed over time to where most internships are taken in the fourth or fifth year of graduate study, frequently after dissertation data have at least been collected (Davies, 1987).

The Boulder conference strongly expressed the opinion that science and practice should be integrated throughout the curriculum, and that the academic experience and the internship experience also be closely integrated with each other. It was one of the first of many (frequently unheeded) calls for close working relationship between academic departments and internship agencies.

2.03.3.2 Stanford

The Boulder Conference recommended a follow-up conference be held in five years. In 1955, delegates representing 37 universities and 14 field training agencies met at Stanford University (Strother, 1956). While the focus was less on specific recommendations than the Boulder Conference, the Stanford Conference did call for a four-year academic program followed by a two-year postdoctoral internship. There was also a recommendation that universities and field training agencies provide joint appointments for faculty and staff, so as to enhance integration of the science and practice of clinical psychology. While the structural changes recommended by the Stanford Conference were generally unheeded, there has been movement toward some types of joint appointments in close relationships between universities and field training agencies, including internship agencies.

2.03.3.3 Miami

While the conference held in Miami in 1958 (Roe, Gustad, Moore, Ross, & Skodad, 1959) was concerned with graduate education in psychology in general, not just in clinical, it nonetheless had some specific concerns about the internship in clinical psychology. It was at this conference that some of the earliest conflicts between internship agencies and academic programs began to arise, in addition to an articulation of the frustration involved in attempting to have the internship serve a true integrative function. This concern over integration of science and practice, which the early models suggested was so important, is a theme which can be seen to have continued in a variety of forms to the present day.

2.03.3.4 Gainesville

The internship plays a critical role in the education and training of clinical psychologists. Yet, as a separate educational experience, it has received relatively little attention until the past 10–15 years. While more general conferences on the education and training of clinical psychologists all affirmed the need for an internship that was in some fashion (usually unspecified) integrated with coursework and practicum experiences, it was not until 1987 that there was a national conference devoted exclusively to considering the internship in professional psychology.

The Association of Psychology Postdoctoral and Internship Centers (APPIC) convened a conference in Gainesville, Florida, in 1987 to explore the purpose of the internship and develop models for quality internship training (Belar et al., 1989). To date, this is the only national conference to focus exclusively on internship training. Forty-eight invited delegates representing the full range of internship settings met for four days to consider a number of questions, including (i) the purpose of the internship, (ii) which psychologists need an internship, (iii) when an internship should occur, (iv), the core requirements of internship, (v), the ideal characteristics on an internship program, and (vi) outcome criteria for evaluating internship training programs.

The delegates to this conference articulated a model of internship as a two-year experience, one year predoctoral and one year postdoctoral. The predoctoral year was described as the generalist training year, with the postdoctoral year where specialty training might begin. This more closely approximates the model used in training physicians, one generalist year followed by specialty training. The two-year model is also consistent with most state and provincial licensing requirements (Association of State and Provincial Psychology Boards, ASPPB, 1996). In this, the two-year internship was seen as having the goal of producing an autonomous professional. In many ways, the Gainesville Conference can be seen as setting the stage for the later APPIC-sponsored Ann Arbor Conference on Postdoctoral Training in Professional Psychology.

The Gainesville Conference, while not explicitly affirming the Boulder model of doctoral education, is best seen as a continuation of the development of the traditional internship model. The Gainesville Conference concluded that the internship should occur after completion of all academic program requirements, including the dissertation.

The Gainesville Conference also affirmed that all interns should be paid a stipend, repeated the call for close and cooperative working relationships between the academic department and the internship agency, and agreed upon a core set of experiences for the predoctoral internship year: (i) a variety of methods of assessment and diagnosis; (ii) a variety of interventions/treatments across a variety of problems and diverse populations; (iii) experience with culturally and ethnically diverse populations, but not limited to delivery of direct services; (iv) research and its practical applications; (v) application of empirical skill and critical thinking to professional practice; (vi) professional, legal, and ethical issues; and (vii) introduction to supervision and management of psychological services.

2.03.3.5 Ann Arbor

While focusing particularly on postdoctoral education and training, in many ways the Ann Arbor Conference (Larsen et al., 1993) can be seen as a logical extension of the Gainesville internship conference. Not only were both conferences primary sponsored by APPIC, but the articulation at Ann Arbor of the need for postdoctoral training for all psychologists, not just those choosing a recognized specialty, is the logical extension of the Gainesville two-year internship model.

Both the Gainesville and Ann Arbor conferences clearly stated that all internship training should occur in sites accredited by the APA and/or the Canadian Psychological Association. The delegates to both Gainesville and Ann Arbor (there was considerable overlap of delegates between the two conferences) clearly felt that all training in professional psychology should take place within the context of programs that had the commitment to quality education and training to submit their programs as candidates for accreditation by the appropriate body. The statement that completion of an accredited internship should be a prerequisite for accredited postdoctoral training was an outgrowth of that philosophy.

2.03.3.6 Orlando

In 1997, APPIC and APA jointly prepared a major national conference to examine issues of supply and demand throughout the professional lifespan of psychologists. What follows is a summary of the reasons for this conference and the plans for its execution, drawn from Pederson (1997).

It has been becoming increasingly clear that there are larger numbers of students seeking professional psychology internships than there are available internship positions. While the number of professional psychology internship positions continues to grow, the number of students seeking internships appears to be increasing even more rapidly. In the late 1990s there is no reason to expect a reversal of this trend to occur soon.

The positive employment outlook for new doctorates in professional psychology has waned. Other areas of psychology have been suffering similar shrinkage in traditional employment for some time. While reliable data sources often lag somewhat behind the rapid change of current events, a multitude of anecdotal and trend evidence suggests that the new doctoral psychologist has a more difficult time finding employment than was previously so, and that salaries are decreasing. In some areas this shrinkage has been dramatic.

There have been numerous efforts to examine parts of the changing practice marketplace picture, such as the APA Presidential Task Force on Training for Managed Care Settings, the APA Division 42 project on Training and Education for Work with Managed Care, the work on Managed Care and Training currently being done by the Center for Mental Health Services, and the APA Board of Educational Affairs Task Force in Education and Training in Professional Psychology. There have been, of course, numerous conferences on training models for doctoral education, internship training, and postdoctoral training and education, such as the Boulder, Vail, Gainesville, Ann Arbor, and Norman Conferences, to name only a few. However, the work of all these separate groups needs to be brought together and placed in the context of other, broader questions, to take a comprehensive and future-oriented look at where psychology is going as a discipline and as a profession.

Psychology continues to be the most popular undergraduate major. This trend will probably continue to increase. This means even greater numbers of students seeking admission to doctoral training programs in professional psychology. We are seeing ever-increasing numbers of new professional psychology doctoral programs opening in an effort to fill this market demand for training in professional psychology. This, too, is likely to continue.

As the size of government at all levels continues to decrease, the amount of public money dedicated to psychology will continue to dwindle. This decrease has been and will continue to be seen in education, research, and professional services, indeed for the entire discipline of psychology, not just for professional psychology and not just for those involved in the education and training of professional psychologists. The effects are being felt first in the education and training community and in the practice community, but those effects are, and will continue to effect, the entire career span of psychologists. Even if we could somehow magically create enough internships for all those seeking them today, we would still have a shortfall as the number of students in training continues to increase. If that increase is continually met, we will soon be overpopulated with psychologists who are competing not only with each other but with the larger and ever increasing numbers of master's level practitioners in a variety of disciplines, including our own.

We are entering a crucial phase in the development of psychology as a discipline and as a profession. Psychology as a discipline is less than 150 years old, and just 100 years old as a profession. We are now seeing the ending of the first major expansion of the discipline and profession that occurred in concert with a wide gamut of social changes following World War II. Before psychology collapses under its own weight, we must take a critical look at where we are going, and what changes need to be made. Models of education and training, models of scientific employment, and models of practice that have served the field so well are beginning to no longer serve the needs either of psychology or of society effectively. We can continue to allow the situation to develop in a fashion that is controlled by forces largely external to psychology, or we can have a hand in determining our own destiny.

As market forces drive more and more of the discipline and profession of psychology, we must look at what we are doing educationally to prepare the next generation of psychologists for the twenty-first century and beyond. The demand for training is increasing while the competition to provide services of all kinds is also increasing. Steps must be taken to correct the many imbalances of supply and demand that exist throughout the career span of the psychologist, before these many imbalances overwhelm us.

Clearly, the answers to these questions cannot be developed in a single conference. In fact, it is unlikely that even all the necessary questions can be formulated at a single conference. Rather, this conference should be seen as an initial stage in a process with which we may be occupied for many years. While the many training conferences that have been held have looked at training models as they "should be," we must begin to look at training and employment as they are driven by a market economy.

Approximately 50 delegates representing 17 national psychology organizations from the USA and Canada and 50 delegates at-large convened in Orlando, Florida, on November 12–16, 1997. The delegates developed answers to a set of questions concerning what psychology's response should be to the changing nature of clinical psychological practice, with the increasing influence of economic and marketing factors. The questions fall into four basic groups, those concerned with entry into professional training and the early preparation of the professional psychologist, those concerned with the transition to internship, those concerned with the early career experiences, and those concerned with later career and exit issues. The responses to these questions included a thorough response, concrete and specific recommendations, and a plan for implementation which will include a time frame, recommended responsible parties, and challenges to implementation. It is hoped that this structure will provide an answer to the problem observed by Laughlin and Worley (1991) that many psychology conferences take place and offer visionary solutions, but those solutions are infrequently actually implemented. The full conference proceedings will be published in 1998.

2.03.4 RESEARCH ON INTERNSHIP EDUCATION AND TRAINING IN CLINICAL PSYCHOLOGY IN THE LATE 1990s

Stedman (in press) provides an excellent review of the literature to date on internship training in clinical psychology. Stedman divided the current literature into several categories:

> The Gainesville Conference, general characteristics of internship programs, interns preparation for internship, intern selection process, supervision of interns, evaluation of interns, management of impaired interns, reflections of interns on their internship experience, and finally, consideration of the internship director. (p. 3)

Much of the work reviewed by Stedman is either of a purely theoretical nature, or of restricted applicability to a particular setting or type of setting. Stedman calls for more empirical research on the nature of internship training, the numbers of students and sites, and alternative models of internship training.

While there has always been generalized agreement within the clinical psychology training community about the need for an internship, the term is surprisingly absent from most state and provincial licensing laws. Only 10 states (no Canadian provinces) make reference to an "internship" requirement. While all such laws require "supervised experience," most do not require an "internship" (ASPPB, 1996). There is a clear difference between these two types of experiences, although it may be a difficult distinction for lawmakers to grasp. That distinction is that an internship is a planned programmatic series of clinical and classroom experiences under supervision that is designed to develop the skills and abilities of the individual intern, whereas supervised experience is just that, experience that is supervised, without the planned, programmatic, or didactic aspects of an internship program.

2.03.5 FOUNDING AND DEVELOPMENT OF THE ASSOCIATION OF PSYCHOLOGY POSTDOCTORAL AND INTERNSHIP CENTERS

In 1968, a group of medical school psychologists consisting of Elton Ash, Ronald Fox, David Kemp, Robert Martin, and Arthur Wiens met to form what was then called the Association of Psychology Internship Centers (APIC). APIC was initially developed to attempt to bring some order to the chaotic way in which students and internships were matched to each other. The first set of "rules" was developed to begin to provide common dates of notification and negotiation.

In 1972 APIC published its first *Directory* of internship programs, listing 90 programs (Fox, 1972). Growth was rapid. By the 5th Edition of the *Directory* there were 261 programs listed. In 1989 APIC was the first national organization to develop standards for postdoctoral programs in professional psychology. In 1991, APIC changed its name to APPIC (Association of Psychology Postdoctoral and Internship Centers) to reflect its growing involvement in postdoctoral training. Today there are 534 internship programs and 77 postdoctoral programs who are members of APPIC and listed in the *APPIC Directory* (Hall & Cantrell, 1996).

It is generally accepted that APPIC Membership Criteria (see Section 2.03.5.1) define what is (and therefore what is not) an internship in psychology. It can therefore be fairly stated that the *APPIC Directory* lists almost all legitimate internship programs in clinical psychology. Note that there exist one or two APA-accredited doctoral programs that include captive internship programs meeting APPIC membership criteria that are not actually members of APPIC.

Probably the most influential and controversial aspect of APPIC policy has been the Uniform Notification Day Policies (formally known as "Policies on Offers and Acceptances"). It is this procedure that has governed the matching of intern applicants and intern positions in the USA and Canada since the mid-1970s.

Simply stated, the policies state that internships may not offer internship positions nor request, accept, or use information from applicants about their relative rankings of internship sites until the second Monday in February (Uniform Notification Day, UND). As currently construed, UND lasts for four hours, from 9 a.m. Central Time to 1 p.m. Central Time. This latest iteration is one of a string of policy changes shortening the time from its original week, down to one and a half days for many years, now to one-half day. Of course the advent of computer matching of applicants and internships may reduce this time to a few nanoseconds.

In 1986 APIC began the development of a body to enforce its policies, and the APPIC Standards and Review Committee was born. Since this time, APPIC has issued numerous private sanctions, and has publicly censured and placed on probation three of its member programs for violation of its Policy on Offers and Acceptances. Increasingly as the number of applicants has continued to outstrip the number of available positions, there have been increasing informal complaints of widespread violations of APPIC policy. The bulk of the reputed violations concern the APPIC policy that states that internship sites may not ask applicants where the internship site ranks for them, the so-called "first choice" issue. In spite of the increase in informal secondhand, thirdhand (and more remote) complaints, the number of complaints actually filed with the APPIC Standards and Review Committee has decreased. Some attribute this to a fear of later retaliation against either the student or the student's doctoral program, whether real or only feared.

Since the mid-1980s, APPIC has increasingly moved toward a much more active advocacy role. APPIC is now seen as the organization that speaks to the interests of and standards for experiential professional psychology education and training at both the doctoral and postdoctoral levels. It has addressed issues of internship funding, the critical importance of the internship in the preparation of clinical psychologists, and is beginning to address broad issues of supply and demand not only for training in clinical psychology but also markets for psychological services. Many of the future directions discussed later in this chapter are arising out of specific APPIC initiatives. A more complete history of APPIC may be found in Stedman (1989) and Pederson (1993a).

2.03.5.1 APPIC Membership Criteria

These membership criteria for doctoral psychology internship programs were revised in 1994; internships that are accredited by the APA or the Canadian Psychological Association are recognized as meeting APPIC doctoral membership criteria. All others must meet all of the following criteria (i.e., 1 through 14 below) and are reviewed for adherence to the criteria every three years.

1. A psychology internship is an organized training program that, in contrast to supervised experience or on-the-job training, is designed to provide the intern with a planned, programmed sequence of training experiences. The primary focus and purpose is assuring breadth and quality of training.
2. The internship agency has a clearly designated doctoral level staff psychologist who is responsible for the integrity and quality of the training program, actively licensed (certified or registered) by the State Board of Examiners in the jurisdiction where the program exists, and present at the training facility a minimum of 20 hours a week.
3. The internship training agency staff consists of at least two full time equivalent doctoral level psychologists who serve as primary supervisors, who are actively licensed (certified or registered) by the State Board of Examiners in the jurisdiction where the program exists.
4. Intern supervision is provided by staff members of the internship agency or by qualified affiliates of that agency who carry clinical responsibility for the cases being supervised. At least two hours per week of regularly scheduled individual supervision are provided by one or more doctoral level licensed psychologists (regardless of whether the internship is completed in one year or two). Supervision is provided with the specific intent of dealing with psychological services rendered directly by the intern.
5. The internship provides training in a range of psychological assessment and intervention activities conducted directly with the recipients of psychological services.
6. At least 25% of the trainee's time is in face-to-face psychological services to patients/clients.

7. The internship must provide at least two hours per week in didactic activities such as case conferences, seminars, inservice training, or grand rounds.

8. Internship training is at post-clerkship, post-practicum, and post-externship level, and precedes the granting of the doctoral degree.

9. The internship agency has a minimum of two full time equivalent interns at the internship level of training during any period of training. These interns must be on site and in training at the time of initial application for APPIC membership.

10. The internship level psychology trainees have a title such as "Intern," "Resident," "Fellow," or other designation of trainee status.

11. The internship agency has a written statement or brochure which provides a clear description of the nature of the training program, including the goals and content of the internship and clear expectations for quantity and quality of the trainee's work, and is made available to prospective interns.

12. Internship programs have documented due process procedures for interns which are given to interns at the beginning of the training period.

13. The internship experience (minimum 1500 hours) must be completed in no less than nine months and no more than 24 months.

14. APPIC member programs are required to issue a certificate of internship completion which includes the word "Psychology" to all interns who have successfully completed the program.

(*APPIC Membership Criteria*. Copyright 1996, Association of Psychology Postdoctoral and Internship Centers. Reprinted with permission.)

2.03.6 DATA ON NUMBERS OF PROGRAMS AND INTERNS IN THE USA AND CANADA IN 1997

There were, during the 1996–1997 training year, a total of 534 agencies providing 2431 full-time internship positions in APPIC-member programs. Of these, 411 agencies were accredited by either the APA or Canadian Psychological Association, and those agencies provided 2070 full-time positions. There were an additional 65 half-time and 76 unpaid positions at these agencies, making a total of 2572 positions available in the USA and Canada (Hall & Cantrell, 1996). The number of internship positions has been rising or held steady for the past eight years (Pederson, 1995).

One major problem is that there is no data source in 1997 on the total number of students enrolled in doctoral programs in clinical, counseling, and school psychology in the USA and Canada. While there is some suggestion that the number of graduate students seeking internships has begun to outpace the growth in the number of internship positions, the magnitude of the problem is unknown (Pederson, 1995, 1997). As the data suggesting that the imbalance between the number of internship positions and the number of students increases, there is an increasing call for some national action on this issue (Pederson, 1997).

2.03.7 CONTINUING ISSUES UNDER DISCUSSION

2.03.7.1 Prerequisites for Internship Appointment

Prior to 1993, there was much confusion in the field of clinical psychology as to what constituted a "practicum hour." Building on the work of several other groups and with the cooperation of the Council of Chairs of Training Councils, APPIC published the first uniform definition of a practicum hour and a standard form on which to report those hours. The definition essentially revolves around the clock hour as the basic unit, and the separation of direct, face-to-face provision of services from indirect service and other training activities. It also defines practicum hours as those receiving academic credit, not work experience prior to or concurrent with doctoral training in clinical psychology.

There remains no universal agreement on the ideal number of practicum hours; statements of 400–2000 hours have been heard in various quarters. Normative data on the actual numbers of practicum hours obtained by clinical psychology graduate students prior to internship is conspicuously lacking.

Given the tremendous diversity of internship settings, it is highly unlikely that any additional standardization of prerequisites for internship will occur at any time soon. Some sites require completion of two years of graduate study, some require three years, some require a completed dissertation, some require a completed dissertation proposal. Some internship sites may require preparation in specific assessment and intervention techniques, while others may be much more flexible.

2.03.7.2 Uniform Requirement and Uniform Applications

For many years graduate students and graduate degree programs have been asking that the process of applying to internships be simplified, in particular by a uniform application form. APPIC began studying the possible adoption of a uniform application in 1991. The US Department of Veterans Affairs (VA), began such a project in 1994 for use by its internship programs. APPIC built upon the initial structure of the VA application and developed an APPIC Uniform Internship Application in 1997. Use of this application is restricted to APPIC-member internships, and its use is optional for APPIC members. The APPIC Uniform Internship Application is available from the APPIC World Wide Web Site at http://www.appic.org.

2.03.7.3 Uniform Notification Day and Computer Matching

Since its inception in the early 1970s, the APPIC Uniform Notification Day for Internship Offers and Acceptances has been controversial. A variety of different methods and different amounts of time have been used, with full satisfaction of all parties concerned still an elusive goal. Stedman (1989) provides a historical review of these various procedures and their outcomes.

Computerized matching of internships and interns has been considered since the 1970s (Pederson, 1993b). The first APPIC vote on computer matching was on the proposal of Briggs (1984), which was defeated. Later, Pederson and Blom (1988) presented another proposal, this time with a somewhat different matching algorithm which allowed for greater flexibility on the part of internship sites in ranking their choices. In 1989–1990, APPIC conducted the Pilot Study of a Computerized Internship Matching Program (Zimet, Blom, Pederson, & Klepac, 1991). Due to the voluntary nature of participation in the Pilot Study, a complete data set was not obtained, which led to many positions and applicants for whom matches could not be determined. Due to the state-of-the-art of computer systems at the time, and the volunteer staffing of the project, there were also a variety of procedural difficulties. This system was also rejected in voting by the APPIC membership.

More recently, APPIC surveys have suggested that a majority of internship directors now favor a computerized matching system, and the APPIC Board of Directors has appointed a committee to obtain proposals for a professional operated system based on the algorithm of Pederson and Blom (1988) to present to the membership for a vote as early as 1998. Although funding of the project remains an important issue, it is anticipated that a computerized matching system will likely be in place and in operation by 2000 at the latest.

2.03.8 FUNDING FOR INTERNSHIP EDUCATION AND TRAINING

2.03.8.1 Present and Future Prospects

Funding for internship programs in clinical psychology has traditionally come from one type or another of government funding either directly or indirectly. Indeed, the earliest formal paid internships in clinical psychology were offered by the US Veterans Administration, which played a key role in the development of standards for clinical psychology education and training (Laughlin & Worley, 1991). A variety of other federal, state, and local government agencies have housed many internship programs. As the size of government in the USA continues to decrease at all levels, this pattern is rapidly changing. Clinical psychology internship training is being caught in the industrialization of health care, and beginning to suffer as a result.

Many long-established government funded internship programs have closed entirely, some due to the closing of their host institutions. More privately funded internships are opening, but these are often under strong pressure to demonstrate their cost-effectiveness. As many managed care companies are beginning to refuse to pay for services provided by nonlicensed personnel, psychology internship programs are seeing a significant decrease in revenue in many locales. This has led to a variety of possible systemic strategies on the part of the training community, including the possibility of seeking changes in state and provincial licensing laws to provide for a limited licensure for interns. Some believe that this would be an easier process if the internship were to occur after receipt of the doctoral degree. Others fear that moving the internship to the postdoctoral year without a simultaneous change in licensing laws requiring an internship would lead to most psychologists not obtaining this additional training if it was not required by either doctoral programs or licensing laws.

2.03.9 FUTURE DIRECTIONS

While much of the future of clinical psychology internship training may well be established at the November 1997 Orlando conference, there are a number of directions upon which it is safe to speculate. Let us begin with a few relatively obvious ones. The managed care movement will continue to grow in both size and influence. The days of one-person private practice will be over, and almost all psychologists will be employed by some kind of company or agency. Some of these companies may be larger group partnerships, but the cottage industry days will be far behind. The impact of this on training will be slight at first, but will grow larger as financial issues begin to outweigh clinical concerns. Training directors will be forced to ever-increasing efforts to show how their internships and postdoctoral programs are cost-effective for the organization. Many programs will close when they are unable to convince their business-oriented companies that training pays off financially. Public sector training such as VA, state, and county hospitals will become even more financially pressed, and downsizing will continue. Postdoctoral training will expand to provide specialty training to fill niche markets. Specialization will also begin to occur earlier and earlier in the training sequence, as we are already seeing in some fields such as health psychology and neuropsychology.

Short-term and highly marketable treatment programs focused on acute problems will predominate in both practice and training. Training programs without training in several of these types of marketable skills will be hard-pressed to recruit quality interns unless they are research-based programs training faculty rather than practitioners. Marketable will not necessarily imply empirically validated effectiveness, but rather documented "customer satisfaction." This will force clinical psychologists, in their training and practice, to attend more to what patients *say* they want rather than what we think they need.

The impact of technology, of distance learning, and telehealth is difficult to estimate at this time. While distance learning is growing, there is a good deal of informal discussion about its appropriateness for clinical training. The increasing focus on serving the needs of rural populations may foster greater development of psychological service delivery via telecommunications, and hence the supervision of psychological services and the training of clinical psychologists via telecommunications.

One thing that will not change, and probably never will, is the human interaction between student and teacher. Since ancient times when a school was a log with a teacher at one end and a student at the other, the importance of the relationship between those with knowledge to share and those seeking that knowledge has thankfully remained little changed.

2.03.10 SUMMARY

Internship training in clinical psychology is defined and discussed, and the history and conceptual origins of the need for internship training in clinical psychology are explored. The major conferences devoted in whole or part to questions of internship training are reviewed, including a conference to be held in 1997 that may well have a major impact on the future of internship training in clinical psychology. The relevant literature is reviewed, and it is noted that there are many critical areas where empirical data is lacking. The history of the APPIC is presented, as are the APPIC membership criteria and data on the current numbers of interns and internship positions in the USA and Canada. Several current issues are explored, including the development by APPIC of a standardized internship application, and the possible adoption of a computerized internship matching system. The current and future funding of internship training in clinical psychology is discussed. Finally, possible future directions and issues in internship training are examined.

2.03.11 REFERENCES

Abrams, J. C., & Brabender, V. (1993, January). *A captive-consortium internship model in a Psy.D. program.* Paper presented at the Conference of the National Association of Schools of Professional Psychology, La Jolla, CA.

American Psychological Association (1945). Subcommittee report on graduate internship training in psychology. *Journal of Consulting Psychology, 9,* 243–266.

American Psychological Association (1947). Recommended graduate training program in clinical psychology. *American Psychologist, 2,* 539–558.

Association of State and Provincial Psychology Boards (1996). *Handbook of licensing and certification requirements for psychologists in North America.* Montgomery, AL: Author.

Belar, C. D., Bieliauskas, L. A., Larsen, K. G., Mensh, I. N., Poey, K., & Roehlke, H. J. (Eds.) (1989). *Proceedings of the National Conference on Internship Training in Psychology.* Washington, DC: Association of Psychology Internship Centers.

Bickman, L., & Ellis, H. (Eds.) (1990). *Preparing psychologists for the 21st century.* Proceedings of the National Conference on Graduate Education in Psychology. Hillsdale, NJ: Erlbaum.

Blanchard, J. J. (1994). *Directory of research opportunities for clinical psychology interns.* Albuquerque, NM: Society for a Science of Clinical Psychology.

Blom, B. E. (1995). Internship Settings. In G. K. Zammit & J. W. Hull (Eds.), *Guidebook for clinical psychology interns* (pp. 27–52). New York: Plenum.

Briggs, D. (1984) Proposal for a centralized computer internship matching system. *APIC Newsletter, 9,* 11–15.

Cripe, L. (1995). Listing of training programs in clinical neuropsychology. *The Clinical Neuropsychologist, 9,* 327–398.

Davies, R. M. (1987). Demise of the third year internship: The changing role of the internship in graduate training. *Professional Psychology: Research & Practice, 18,* 481–484.

Dorken, H., & Cummings, N. A. (1977). A school of psychology as innovation in professional education: The California School of Professional Psychology. *Professional Psychology, 8,* 129–148.

Fox, R. (1972). *Directory of internship programs in psychology.* Los Angeles: Association of Psychology Internship Centers.

Fraser, J. S., & Bent, R. J. (1994, January). *Reshaping predoctoral internships in the age of professional school training*. Paper presented at the Conference of the National Council of Schools of Professional Psychology, La Jolla, CA.

Hall, R. G., & Cantrell, P. J. (1996). *APPIC directory: Internship and postdoctoral programs in professional psychology* (25th ed.). Washington, DC: Association of Psychology Postdoctoral and Internship Centers.

Hoch, E. L., Ross, A. O., & Winder, C. L. (Eds.) (1969). *Professional preparation of clinical psychologists*. Washington, DC: American Psychological Association.

Kohut, J., & Cooney, B. (1993). *Internship data*. Paper presented at the Conference of the National Council of Schools of Professional Psychology, New Orleans, LA.

Korman, M. (1976). *Levels and patterns of professional training in psychology*. Washington, DC: American Psychological Association.

Larsen, K. G., Belar, C. D., Bieliauskas, L. A., Klepac, R. K., Stigall, T. T., & Zimet, C. N. (1993). *Proceedings of the National Conference on Postdoctoral Training in Professional Psychology*. Washington, DC: Association of Psychology Postdoctoral and Internship Centers.

Laughlin, P. R., & Worley, J. L. (1991). Roles of the American Psychological Association and the Veterans Administration in the development of internships in psychology. *American Psychologist, 46,* 430–436.

Levinger, C., & Schefres, I. (1995). *Everything you need to know to get a psychology internship*. Los Angeles: Internship Publishers.

Lopez, S. J., Oehlert, M. E., & Moberly, R. L. (1996). Selection criteria for American Psychological Association-Accredited internship programs: A survey of training directors. *Professional Psychology: Research and Practice, 27,* 518–520.

Megargee, E. I. (1997) *Megargee's guide to obtaining a psychology internship* (3rd ed.) Muncie, IN: Accelerated Development.

Mellott, R. N., Arden, I. A., & Cho, M. E. (1996). Preparing for the internship: Tips for the prospective applicant. *Professional Psychology: Research and Practice, 27,* 190–196.

Pederson, S. L. (Ed.) (1993a). 25th Anniversary Special Issue (Special Issue). *APPIC Newsletter, 19*(1).

Pederson, S. L. (1993b). The history of computerized internship matching in APPIC. *APPIC Newsletter, 19,* 56–58.

Pederson, S. L. (1995). *Supply and demand for internships in professional psychology: A preliminary report*. Unpublished manuscript prepared for the Council of Chairs of Training Councils.

Pederson, S. L. (1996, August). Overview of the internship application, interviewing, and selection process. In S. L. Pederson (Chair), *Applying and interviewing for professional psychology internships*. Symposium conducted at the meeting of the American Psychological Association, Toronto, Ontario, Canada.

Pederson, S. L. (1997). Supply and demand: Training and employment opportunities in professional psychology. *APPIC Newsletter, 22*(5), 25–27.

Pederson, S. L., & Blom, B. E. (1988). A proposal for a computerized APIC National Intern Matching Program. *APIC Newsletter, 13,* 53–64.

Plante, T. G. (1996). Ten principles of success for psychology trainees embarking on their careers. *Professional Psychology: Research and Practice, 27,* 304–307.

Raimy, V. C. (Ed.) (1950). *Training in clinical psychology*. New York: Prentice-Hall

Roe, A., Gustad, J. W., Moore, B. V., Ross, S., & Skodak, M. (Eds.) (1959). *Graduate education in psychology*. Washington, DC: American Psychological Association.

Stedman, J. M. (1989). The history of the APIC selection process. *APIC Newsletter, 14,* 35–43.

Stedman, J. M. (in press). What we know about predoctoral internship training: A review. *Professional Psychology: Research and Practice*.

Stewart, A. E., & Stewart, E. A. (1996a). Personal and practical considerations in selecting a psychology internship. *Professional Psychology: Research and Practice, 27,* 295–303.

Stewart, A. E., & Stewart, E. A. (1996b). A decision-making technique for choosing a psychology internship. *Professional Psychology: Research and Practice, 27,* 521–526.

Strother, C. R. (1956) *Psychology and mental health*. Washington, DC: American Psychological Association.

Zimet, C. Z., Blom, B. E., Pederson, S. L., & Klepac, R. K. (1990). *The Results of the APIC Pilot Study of the Computerized Internship Matching Program*. Symposium conducted at the meeting of the American Psychological Association, Boston.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.04

Postdoctoral Education and Training in Clinical Psychology

KERRY de S. HAMSHER
Neuropsychology Clinic, Milwaukie, WI, USA

2.04.1 INTRODUCTION 37
2.04.1.1 Models for Education and Training Beyond the Doctoral Degree 37
2.04.2 FIFTY YEARS OF EVOLUTION 38
2.04.3 GROWTH OF FORMAL POSTDOCTORAL PROGRAMS 40
2.04.4 ORGANIZING FOR EDUCATION, CREDENTIALING, AND ACCREDITATION: THE CLINICAL NEUROPSYCHOLOGY EXAMPLE 40
2.04.5 FUTURE 43
2.04.6 REFERENCES 45

2.04.1 INTRODUCTION

In its broadest definition, postdoctoral education and training in clinical psychology can refer to all education and training beyond the doctoral degree. This was, in fact, the theme of a national conference convened in 1994, to develop a broad scope of issues related to postgraduate education and training for the discipline of psychology in the twenty-first century (American Psychological Association [APA], 1995). Delegates to the conference focused on six general issues: (i) models for education and training beyond the doctoral degree; (ii) societal needs, changing demographics, and national policies; (iii) expanding knowledge and skills base; (iv) impact of changing technology; (v) evaluation and lifelong learning; and (vi) funding.

2.04.1.1 Models for Education and Training Beyond the Doctoral Degree

The following types of postdoctoral education and training were presented as an initial concept for consideration:

(i) *Formal organized program.* These programs are comprised of specific goals, specified duration of training, specified methods, and have stated criteria for evaluation.

(ii) *Individualized mentorship program.* This program, also known as the apprenticeship model, is one in which the postdoctoral trainee works for one to three years under the direction of an active, established psychologist (mentor) in an adequate learning environment and carries out an organized program of a focused nature. Such mentorships are typically found in research and occasionally in practice and teaching programs.

(iii) *Individualized supervision program.* In this program the emphasis is on individualized supervision of the trainee. It is often used to provide advanced training in a given area of practice such as interpersonal therapy. Individual supervision is required; reading and/or didactic course work may be required and training may be for one or two years. A second type of individual supervision of the trainee's experience by a licensed psychologist may be required for state licensure purposes.

(iv) *Sabbatical/paid leave*. This is institution-sponsored faculty development time, usually involving a range of educational activities designed and proposed by the recipient and ranging from one or two months to one year.

(v) *Continuing education*. Continuing education is the hallmark of lifelong learning. It provides the opportunity for updating the psychologist's knowledge in a given area and expanding it in light of new knowledge derived from other sources (e.g., other disciplines). It can be both formal and informal.

(vi) *Faculty/staff development, peer consultation, and self-study*. This is usually more informal and less systematic than the other forms described above.

This taxonomy has subsequently been adopted by the APA Council of Representatives for the field of psychology (Table 1).

2.04.2 FIFTY YEARS OF EVOLUTION

Early generations of clinical psychologists were considered educated and trained in the specialty of clinical psychology upon receiving their doctoral degrees. Over the years, the need for postdoctoral education for professional practice gradually became more apparent when it was clear that new opportunities for service required specialized training and that critical social needs required trained specialists. Recommendations for supervised training beyond the doctoral degree have been made for at least 50 years.

(i) *Boulder*. In August 1949, a conference supported by the National Institute of Mental Health (NIMH) convened at Boulder, Colorado, to consider the training of clinical psychologists. Seventy-one representatives from training universities, mental health agencies, and allied profession participated (Raimy, 1950). The participants encouraged broad professional training in clinical psychology, leaving more specialization to postdoctoral experience.

(ii) *Stanford*. In August 1955, again with NIMH support, a conference on Education and Training for Psychological Contributions to Mental Health was held at Stanford University. The deciding factor leading to the conference was the rapid growth within the previous three or four years of the mental hygiene movement and the realization that this movement was going to have a far-reaching effect on psychology. The institute strongly favored efforts to define a common core of professional training that would constitute basic preparation for various special fields. The participants suggested that some special training could be included in a predoctoral program, but full competence in any specialty would undoubtedly require some postdoctoral training (Strother, 1956).

(iii) *Miami*. The Miami conference participants (Roe, Gustad, Moore, Ross, & Skodak, 1959) discussed issues of specialization in psychological practice and talked about the gradual development of a two-year internship to follow the doctoral degree. The model of a two-year, supervised period of field training before subspecialty or general practice is not new and it has stood the test of time. It is today the modal preparation period before licensure and practice.

(iv) *Chicago*. Participants in the Chicago conference (Hoch, Ross, & Winder, 1966) differentiated postgraduate and postdoctoral training. Postgraduate training was thought necessary for individuals with doctorates in other specialties who wished to become clinical psychologists. After much debate about postdoctoral training, the conference endorsed the statement that: "Although postdoctoral training is desirable for all clinical specialists, it is deemed essential for those anticipating independent professional practice and also for those who are to teach and supervise in clinical specialty courses" (Hoch et al., 1966, p. 47). It was also agreed that the postdoctoral program should never serve to compensate for deficiencies in training at the predoctoral level. The conferees wished that professional preparation be a continuous, integrated sequence with postdoctoral training, building on predoctoral preparation.

(v) *Menninger Postdoctoral Conference*. An interest group in postdoctoral clinical training was formed in 1958 to provide a forum for the exchange of ideas among psychologists interested in seeking or offering programs of supervised clinical experience beyond the doctoral degree. The efforts of this group to mount a national conference came to fruition on May 18–21 1972, when 42 psychologists representing postdoctoral training facilities, university departments of psychology, the Veterans Administration, APA, NIMH, and doctoral and postdoctoral trainees met at the Menninger Foundation to share their thoughts and experiences (Weiner, 1973). Attendees shared information and raised concerns about current and future postdoctoral education; they developed few policy statements.

(vi) *Gainesville Internship Conference*. The National Conference on Internship Training in Psychology (Belar et al., 1989) proposed that the traditional one year internship be expanded to two year's experience: the first year at doctoral level and the second at postdoctoral level. Two years of supervised

Table 1 Taxonomy to describe postdoctoral education (for practice, research, and teaching in psychology).

	Models					
Variables	*Formal programs*	*Individual mentorship*	*Individual supervision*	*Continuing education*	*Sabbatical/paid leave*	*Faculty development, peer consultation, and self-study*
Orientation/target population	Practice, research, teaching	Research	Practice	Practice, research, teaching	Research, teaching	Practice, research, teaching
Purpose	Specialization, respecialization	Specialized research	Advanced practice competencies	Update/renew knowledge	New knowledge, consolidation of knowledge	Update, confirm, and share knowledge
Content	Clinical competence, advanced research, teaching competence	Individual focus	Clinical care readings	Selected topic areas	Individually tailored	Selected topic areas
Duration	1–2 years	1–3 years	1–5 years	1 day–2 weeks	3–12 months	*ad hoc*
Funding	Institutional, governmental	Institutional, governmental	Institutional, self	Self	Institutional, self	Self
Evaluation/competency	Stated criteria	Stated criteria	Established criteria	Examination	Self-evaluation products	Self-evaluation

experience was deemed necessary to produce an autonomously functioning professional psychologist. In general, the Gainesville conference resulted in a raising of standards for the training of professional psychologists. It codified the contemporary status of the postdoctoral year and asserted the need for organized quality control over this aspect of professional training.

(vii) *NIMH*. The NIMH has received policy recommendations from the National Conference on Implementing Public–Academic Linkages for Clinical Training in Psychology (Callan, Johnson, Leon, Magrab, & Meyers, 1990). Regarding postdoctoral education and training, the conferees suggested that a variety of postdoctoral fellowships should be developed for training focused on priority populations, one to two years for clinical service and administrative training, and two to three years for combined clinical service and research training. Continuity of training experience between the university, internship and postdoctoral facilities was also suggested to ensure that training focused on the actual needs of priority populations and on high-quality services.

(viii) *Ann Arbor*. The National Conference on Postdoctoral Training in Psychology was convened in Ann Arbor in 1992 by the APA, the Association of State and Provincial Psychology Boards (ASPPB), and the Association of Psychology Postdoctoral and Internship Centers (APPIC) (Belar et al., 1993). Its purpose was to develop specific postdoctoral training criteria for students, training programs, and credentialing authorities. Postdoctoral training consisting of one year full-time (or two year half-time) was recommended for psychologists who wanted to teach, conduct research, or to practice. Various structural, supervisory, context, content, and evaluation issues were considered and the conference report was essentially the first detailed APA policy statement regarding postdoctoral training. These conference criteria were subsequently modified by the Inter-organizational Council for Accreditation of Postdoctoral Programs in Psychology (IOC) (see below) and served as the background material for the drafting of the APA's Committee on Accreditation (CoA) accreditation criteria for postdoctoral programs.

(ix) *Norman, Oklahoma*. The 1994 APA National Conference on Postdoctoral Education and Training in Psychology has already been alluded to above (APA, 1995). This was a conference that addressed education and training beyond the doctoral degree very broadly. It was designed for teachers, researchers, and practitioners in psychology.

2.04.3 GROWTH OF FORMAL POSTDOCTORAL PROGRAMS

Alexander (1995) attempted to identify postdoctoral programs in existence in 1960–1962. He reported that, excluding VA installations offering postdoctoral internships, perhaps 22 institutions were offering continuing one or two year programs of study, with an additional 10 that had trained students sporadically over the years as the occasion arose. Eight programs were located in medial schools and five others in private clinical facilities. Most of the remainder were found in joint hospital, clinical, or medical school settings. One program was founded in a school for exceptional children, and one in a university department of psychology. Irving Weiner (1968) identified 46 postdoctoral programs in clinical psychology in 1967 that were (i) designed for psychologists with a degree from an APA-approved doctoral program, (ii) administered and directed principally by psychologists, and (iii) focused primarily on education for advanced professional competence rather than on in-service or on-the-job training.

Wiens (1993) reported on a tabulation of postdoctoral training programs that were listed in various sources in 1990–1991. The largest number was gleaned from *The APA Monitor* from January to December 1991. Other sources of information were the *APPIC Directory*, the *Division of Health Psychology Directory*, and the *Clinical Neuropsychology Directory*. The combined listings totaled 388 postdoctoral training sites which were accommodating 694 trainees. Tabulation of additional listings for 1991 yielded a total of approximately 500 sites.

The listing of postdoctoral training sites was an effort to create greater awareness of the number and variety of postdoctoral education and training opportunities. It was also an effort to encourage discussion of establishing mechanisms for credentialing or accrediting programs that meet established guidelines for training in one or another specialty area.

2.04.4 ORGANIZING FOR EDUCATION, CREDENTIALING, AND ACCREDITATION: THE CLINICAL NEUROPSYCHOLOGY EXAMPLE

In the 1980s, the APA Task Force on Scope and Criteria of Accreditation recommended that the accreditation of education and training programs should be expanded to cover postdoctoral residency programs beginning in the four traditional fields: clinical, counseling, school, and industrial–organizational psychology. Mention was made of eventually turning attention to emerging specialties and, in a footnote to this, clinical neuropsychology was mentioned. By this time a postdoctoral residency in clinical neuropsychology had become the traditional preparation within the specialty. There had been a history of such going back to at least the 1960s. Clinical neuropsychology had obtained a division within APA in 1980. This Division 40 cooperated with the International Neuropsychological Society in its Task Force on Education, Accreditation, and Credentialing. The first of the INS-Division 40 Joint Task Force reports (Meier, 1981) called for the development of board certification in clinical neuropsychology. Just as the APA had sponsored and initially funded the development of the American Board of Professional Psychology (ABPP) to provide board certification for all of psychology beginning in 1947, the APA Division of Clinical Neuropsychology followed suit in sponsoring the American Board of Clinical Neuropsychology which was incorporated in 1981. Board certification examinations under the aegis, scrutiny, and supervision of the ABPP were accomplished in 1984. Subsequently, the INS-Division 40 Task Force reissued guidelines for the doctoral, predoctoral internship, and postdoctoral residency levels of education and training within the specialty. Listings of such programs which self-asserted compliance with the respective guidelines were published (Cripe, 1988).

In 1988, a special meeting was called and held in Iowa City, in conjunction with a meeting of the Midwest Neuropsychology Group. The premise presented was that clinical neuropsychology now had in place all the prerequisite bodies and published guidelines that needed to exist prior to the initiation of accreditation, or so it was thought at that time. At this point the APA had taken no steps to accredit residency programs in the traditional fields, let alone even recognize clinical neuropsychology as a specialty. Directors of residency programs in clinical neuropsychology within contiguous Midwestern states (Iowa, Michigan, Minnesota, Wisconsin, and eventually Illinois and Ohio), agreed to form an organization the purpose of which was to develop mechanisms for the accreditation of residency programs in the specialty of clinical neuropsychology. The Midwest Neuropsychology Consortium of Postdoctoral Programs in Clinical Neuropsychology (MNC) considered itself a working group which would be kept small until accreditation mechanisms were in place, at which time it was planned to expand so as to accept programs within North America. Building on the published guidelines (INS-Division 40 Task Force on Education, Accreditation, and Credentialing, 1986, 1987), the MNC began work on elaborating standards for a residency program, the curriculum, standardization of application materials, deadlines and notification dates and policies. Early informal inquiries had failed to find an independent organization interested in providing oversight of the anticipated accreditation activities. However, after the MNC field tested its policies, procedures, and materials for accreditation, and was about to commence official accreditation site visits in 1991, Manfred Meier (1981) captured the interest of the ABPP in this endeavor.

The ABPP agreed to become involved in accreditation but it would not do so alone. The ABPP called for the development of the Interorganizational Council for Accreditation of Postdoctoral Programs in Psychology (IOC) and an inaugural meeting was held in Minneapolis, Minnesota, in September 1991. Representatives came from the Education Directorate, the Board of Educational Affairs, the Accreditation Office and the Executive Counsel of the APA, along with the Canadian Psychological Association, the National Register of Health Service Providers in Psychology (NRHSPP), the Canadian Register of Health Service Providers in Psychology, the ASPPB, the APPIC, and the National Council of Schools of Professional Psychology (NCSPP). Among these major organizations, all but the NCSPP maintained its membership on the IOC. It was determined that the specialties would be represented by the association of directors of residency programs within the specialty, if such existed; otherwise, a representative of the ABPP member board for the specialty would fulfill this role. This occasioned the MNC to enact its planned expansion and it evolved into the Association of Postdoctoral Programs in Clinical Neuropsychology (APPCN). The APPCN financially supported the representative of clinical neuropsychology to the IOC. The IOC met two to three times a year. Organizational conflicts and debates over concepts dominated the early agenda of the IOC. For example, although the APA asserted proprietary rights concerning accreditation, at that time only the specialty of clinical neuropsychology was positioned to have its residency programs evaluated for accreditation and the APA did not recognize clinical neuropsychology as a specialty (nor a number of other putative specialties which already had been recognized by ABPP). Another issue of key importance to the specialties in professional psychology, and the subject of spirited debate, was the procedure by which standards for education and training in the specialties would be recognized and adopted so that the corollary accreditation criteria could be established.

The APA eventually began a process to address this problem. Authorization was obtained to explore mechanisms by which the APA could come to officially recognize new specialties. The outcome was the formation of the Commission for the Recognition of Specialties and Proficiencies in Professional Psychology (CRSPPP). The CRSPPP worked through considerable developmental issues, along with defining its policies and procedures. After the successful completion of a demonstration project, the CRSPPP was fully commissioned by the APA Council of Representatives in 1995. The Division of Clinical Neuropsychology then officially petitioned the CRSPPP for specialty recognition and presented their application for review. In 1996, the Council of Representatives approved the CRSPPP recommendation, and clinical neuropsychology became the first new specialty to be recognized in professional psychology.

Meanwhile, the IOC decided to postpone consideration of specialty issues in accreditation and proceeded with the generation of guidelines for the accreditation of generic postdoctoral residencies. Generic programs refer to education and training programs that occur after completion of the doctoral program and the predoctoral internship. They are intended to advance the competencies of the residents in a general fashion but do not lead to qualification for the practice of any specialty within professional psychology. Stimulating this process was the National Conference on Postdoctoral Training in Professional Psychology, held in Ann Arbor in 1992 (Belar et al., 1993), which was described above. Building on the APA's guidelines for the accreditation of internship programs, the conference prepared a draft of guidelines for generic residencies. Later, this document was presented to the IOC and APA's Committee on Accreditation (CoA). The Conference on Education and Training Beyond the Doctoral Degree (APA, 1995) held in Norman, Oklahoma in 1994, laid down a foundation of thought about postdoctoral education, thereby providing a broad context for further work on these issues. This conference assembled data and considered possible ramifications and implications from the growth and expansion of residency education and training programs and other changes in professional psychology. One of the recommendations to come from this conference was to move the predoctoral internship to a postdoctoral experience. This policy proposal was endorsed by Council of Graduate Departments of Psychology (COGDOP) in 1997.

Upon completion of the Joint Task Force with Division 40, the INS divorced itself from further involvement in professional matters. That is, the INS was an international, multidisciplinary scientific society and it was realized that its composition was not wellsuited for such matters. Although the INS would remain generally interested in, and supportive of, education in neuropsychology, it also recognized it did not have the resources to assist all of its represented disciplines in all the countries of its membership.

Like other scientific disciplines and practice specialties, the history of neuropsychology in the USA was in part imprinted with a growing number of organizations (Benton, 1987). The INS helped to give birth to Division 40, Division 40 helped to give birth to the ABCN. In accordance with Division 40 guidelines, and in parallel with what became the standard in medicine, the APPCN required that the directors of residency programs be board certified within the specialty. In addition to these kindred, psychologist groups, a multidisciplinary interest group arose which was especially concerned with continuing education and practice matters, that is, the National Academy of Neuropsychology (NAN).

Based on the work of the IOC, which incorporated many of the points raised by the National Conference on Postdoctoral Training in Professional Psychology (Belar et al., 1993), the Conference on Education and Training Beyond the Doctoral Degree (APA, 1995), and ensuing collaboration, the Committee on Accreditation in 1996 adopted the Guidelines and Principles for Accreditation and developed the Accreditation Operating Procedures for Postdoctoral Programs (Committee on Accreditation, 1997; Office of Program Consultation and Accreditation, 1997).

By 1993, clinical neuropsychology's participation in the IOC led to two stirring realizations. First, in order to secure the extension of accreditation to its specialty residency programs, the APPCN would have to seek and obtain general acceptance of their guidelines within the specialty profession. Second, the specialty profession had become so highly differentiated as to reach the point when some form of reintegration appeared necessary in order to achieve the first realization. To this end, proposals calling for an integration of organizations were unanimously approved by the boards of directors of both ABCN and APPCN, as well as by the executive committee of the Division of Clinical Neuropsychology. However, two years elapsed with no concrete action having been taken. To precipitate events, the chief officers of the ABCN, APPCN, Division 40, NAN, and the INS were invited to discuss the proposal to which most had agreed. All agreed to take up the cause of this orphaned brainchild, that is, that a metaorganization of the major organizations within clinical neuropsychology should be established. The title of the Clinical Neuropsychology Synarchy (CNS) was adopted. This erstwhile rare word, synarchy, meaning governance through joint sovereignty, was deemed apropos in its renascent application.

With unanimous accord, the CNS articulated two fundamental positions. These were that clinical neuropsychology was only a specialty and not a proficiency and that there should be but one board certification agency within the specialty. Upon the formation of the CNS it was immediately apparent there were important organizational gaps within the specialty which would become glaring when consideration turned to reissuing integrated guidelines for education and training in clinical neuropsychology. In consequence, the CNS issued a call for two organizing meetings, both of which were held in conjunction with the North American meetings of the INS. In 1996, this led to the formation of the Association for Doctoral Education in Clinical Neuropsychology (ADECN); in 1997, the Association for Internship Centers in Clinical Neuropsychology (AITCN) was brought into existence. Starting before the formation of the latter group, a planning committee of the CNS undertook preparations to bring about the Houston Conference on Specialty Education and Training in Clinical Neuropsychology. The Houston conference was successfully concluded in September of 1997. From this conference it was established that, from this time onward, a residency would become a necessary component in the education and training of a clinical neuropsychologist, and that the beginning of this specialty path began in doctoral programs in professional psychology.

In August of 1996, the declaration of the APA Council of Representatives affirming the CRSPPP recommendation for APA to recognize the specialty of clinical neuropsychology removed a major obstacle to accreditation in clinical neuropsychology and other specialties. Consequently, completion of the formative work for the recognition and representation of the specialties in the accreditation process was the paramount issue before the IOC at its November, 1996 meeting In that meeting the major regulatory organizations proposed the formation of a Council of Specialties (CoS) and charged it with various roles and tasks including that of recognizing, standardizing and interfacing specialty standards for education and training with the accreditation criteria for use by the CoA. In earlier sessions of the IOC, the specialty representatives persuasively argued that the form and content of the specialty residency would essentially define the nature of the specialist. Therefore, it must be made practicable for the psychological specialties to exert their sovereignty and carry out their responsibility to design and refine their own fates on a continuing basis. In support of the CoS, each specialty seeking accreditation of their education and training programs was instructed to form their own synarchy or specialty council, to which at least the following specialty-specific organizations were to be invited to participate: the ABPP member board (i.e., the board certifying agency within the specialty), the corresponding ABPP academy (to represent publically recognized qualified practitioners of the specialty), the APA division or divisions relating to the specialty, and representation of the educators, be they at the doctoral, internship, or residency levels of education and training (as befit the individual specialty). When formed, these specialty councils will become responsible for selecting their representatives to the CoS. The CoS also was charged with the responsibility to facilitate the resolution of disputes that may develop about interspecialty boundaries. The potential for such becomes particularly germane to the CoS when they may affect the accreditation of postdoctoral specialty programs in professional psychology, the recognition of specialties, or the certification of specialists.

It was through many years of preparation and organization that clinical neuropsychology earned the privilege (albeit with tribulations) to blaze a trail, clearing the way for other specialties of psychology to continue their maturation. Of course, this was not without capitalizing upon the developments and experiences of these other specialties in psychology as well as those in medicine. Nor could these advancements have taken place were it not for an audience of prepared minds in professional psychology. It is for this we must be especially grateful for the series of national conferences, dating back to that seminal meeting in Boulder, Colorado, which framed many of the questions, and the subsequent chorus of responses delivered by the collectives of wisdom convened over the past 50 years.

2.04.5 FUTURE

Uncertainties about how the structure of services will unfold in the health care and mental health arenas and how the education of doctoral-level health care professionals will be financially supported in the future place a haze on the horizon of professional psychology. Be that as it may, at least some imperatives can be seen clearly in the foreground. The first of these, of course, is funding. In order to provide quality residency programs in professional psychology, these programs must obtain stable and sufficient funding to support the educational activities of these programs, their residents and faculty. In the current state of flux, a number of deleterious factors are in play. Among these is the pressure from the managed care industry to take over the direction of health care services. In some instances, this has entailed the pounding of the square peg of professional health care services into the round hole of a manufacturing model often favored by some of the business-minded interlopers. The restriction or denial of services for the purposes of short-term financial gains by some third party insurers, their difficulty in considering and assessing the financial impact of services that lie beyond the year's end, and ways to factor-in the quality of life of the insured into the balance sheet are problems that eventually must be overcome.

The organizational structure of accreditation for professional psychology programs is overdue for a re-evaluation and redeployment. Presently it is housed within the membership organization of the APA. Has this become an historical anachronism? The CA is operating on an old model which some may feel is outdated. For example, the final decision on accreditation is principally decided by the committee as a whole on which there is no specific representation of the specialties at the postdoctoral level. Furthermore, in the methods by which the committee's decision to deny the accreditation of a seemingly inadequate program may be overturned, there is no representation of the principal stakeholders. The field of medicine that has probably worked on these problems longer than any other health care profession. In medicine, the overarching authority for accreditation is based on an interorganizational model. This is a model that psychology must consider if other goals, as described below, are to be achieved. A seedling for such a model may exist in the goals and aspirations of the Council of Credentialing Organizations in Professional Psychology (CCOPP). One goal of the CCOPP is to seek an integration of credentialing, accreditation, and other regulatory functions within professional psychology. Furthermore, at some point in the future, consideration of those aspects of the North American Free Trade Agreement (NAFTA) pertaining to professional credentialing and recognition will come into play.

As advanced and specialty education and training at the postdoctoral level unfolds, there will need to be a greater specification and integration of education and training programs at the doctoral, internship and postdoctoral residency levels. Also, if the internship is moved to the postdoctoral stage in the educational sequence, then coordination between internship and residency programs shall be required. In addition, provision must be made for subspecialization through joint residency or postdoctoral fellowship programs in the forthcoming conceptual model for education and training in professional psychology, and the accreditation thereof.

It would be short-sighted if the planning of the new model were to stop short at the point of a resident's departure from postgraduate education. The expected products of these education and training programs, i.e., the knowledge and competencies envisaged to be gained through such, must be coordinated with board certification standards for the individual practitioner in the specialties. At the same time, work must begin on the further considerations that will emerge down stream in the life of a practitioner in professional psychology. These include the timely application of recertification standards within the board-certified specialties and possibly in the relicensing of the general practitioner. To achieve this, mechanisms to plan for, entrain, and steer continuing educational programs in specified directions for the generalist as well as the specialist must be explicated. How the provision for, and approval of, the continuing education of the specialist within their specialty is yet to be proposed. Whatever this may be, the continuing education of the specialist and the periodic board recertification standards to be applied to them must dovetail. Obviously, there must be some parallel between the available curriculum for continuing education, the contemporary standards to be applied to the accreditation of doctoral, internship and residency programs, and the contemporary board certification standards applied to these programs' graduates.

Advancements in the science and profession of psychology, changes in the health care arena, and their interactions call for built-in adaptive capacities in this new model. Such flexibility, however, can no longer be provided through the absence of standards and their enforcement in professional psychology. Instead, mechanisms for prospective planning, rapid and coordinated communication of such changes, and their incorporation into the regulatory activities of professional psychology will need to become the modus operandi as we enter the twenty-first century. It is difficult to foresee how these necessary evolutionary milestones can be achieved without the acceptance and development of a synarchy drawn from the principal constituent organizational components of professional psychology as a whole.

2.04.6 REFERENCES

Alexander, I. E. (1995). Postdoctoral training in clinical psychology. In B. B. Wolman (Ed.), *Handbook of clinical psychology* (pp. 1415–1426). New York: McGraw-Hill.

American Psychological Association (1995). *Education and training beyond the doctoral degree: Proceedings of the American Psychological Association National Conference on Postdoctoral Education and Training in Psychology*. Washington, DC: Author.

Belar, C. D., Bieliauskas, L. A., Larsen, K. G., Mensh, I. N., Poey, K., & Roehlke, H. J. (1989). National conference on internship training in psychology. *American Psychologist, 44,* 60–65.

Belar, C. D., Bieliauskas, L. A., Klepac, R. K., Larsen, K. G., Stigall, T. T., & Zimet, C. N. (1993). National conference on postdoctoral training in professional psychology. *American Psychologist, 48,* 1284–1289.

Benton, A. (1987). Evolution of a clinical specialty. *The Clinical Neuropsychologist, 1,* 5–8.

Callan, J. E., Johnson, D. L., Leon, G. R., Magrab, P. R., & Myers, H. F. (1990). *Policy recommendations from the national conference on implementing public–academic linkages for clinical training in psychology*. Washington, DC: National Institute of Mental Health.

Committee on Accreditation (1997). *The annual report of the committee on accreditation 1996: In the spirit of the new era*. Washington, DC: American Psychological Association.

Cripe, L. (1988). Listing of training programs in clinical neuropsychology 1987. *The Clinical Neuropsychologist, 2,* 13–24.

Hoch, E. L., Ross, A. O., & Winder, C. L. (1966). *Professional preparation of clinical psychologists*. Washington, DC: American Psychological Association.

INS-Division 40 Task Force on Education, Accreditation, and Credentialing (1986). Guidelines for postdoctoral training in clinical neuropsychology. *Newsletter 40*(4), 4.

INS-Division 40 Task Force on Education, Accreditation, and Credentialing (1987). Reports of the INS-Division 40 Task Force on Education, Accreditation, and Credentialing: Guidelines for doctoral training programs in clinical neuropsychology. *The Clinical Neuropsychologist, 1,* 29–34.

Meier, M. J. (1981, September). Report of the Task-Force on Education, Accreditation, and Credentialing of the International Neuropsychological Society. *The INS Bulletin,* 5–10.

Office of Program Consultation and Accreditation (1997). *Book 1: Guidelines and principles for accreditation of programs in professional psychology; Book 2: Accreditation operating procedures of the Committee on Accreditation*. Washington, DC: American Psychological Association.

Raimy, V. C. (1950). *Training in clinical psychology*. New York: Prentice-Hall.

Roe, A., Gustad, J. W., Moore, B. V., Ross, S., & Skodak, M. (1950). *Graduate education in psychology*. Washington, DC: American Psychological Association.

Strother, C. R. (1956). *Psychology and mental health.* Washington, DC: American Psychological Association.

Weiner, I. B. (1968). Postdoctoral training in clinical psychology. *American Psychologist, 23,* 374–377.

Weiner, I. B. (1973). *Postdoctoral education in clinical psychology*. Topeka, KS: Menninger Foundation.

Wiens, A. N. (1993). Postdoctoral education–training for specialty practice: Long anticipated, finally realized. *American Psychologist, 48,* 415–422.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.05

Accreditation of Education and Training Programs: Its Role in Shaping a Profession

PAUL D. NELSON
American Psychological Association, Washington, DC, USA

2.05.1 THE PROFESSIONS, ACCREDITATION, AND PUBLIC ACCOUNTABILITY 47
2.05.2 THE CONTEXT FOR ACCREDITATION IN THE PROFESSIONS 48
2.05.3 ACCREDITATION IN PSYCHOLOGY: THE ADVENT OF A PROFESSION 49
2.05.4 GOALS AND PRINCIPLES OF PROFESSIONAL EDUCATION AND TRAINING 51
2.05.5 MODELS OR PHILOSOPHIES OF PROFESSIONAL EDUCATION AND TRAINING 52
2.05.6 ACCREDITATION STANDARDS: CURRICULUM AND COMPETENCE CONTROVERSY 53
2.05.7 STORM WARNINGS IN ACCREDITATION: TIME FOR CHANGE OF HEADING 55
2.05.8 OUTCOMES OF ACCREDITATION: ISSUES OF PUBLIC ACCOUNTABILITY 56
2.05.9 REFERENCES 58

2.05.1 THE PROFESSIONS, ACCREDITATION, AND PUBLIC ACCOUNTABILITY

In his annual report to the American Psychological Association (APA), writing as the Association's Executive Secretary, Sanford (1951) spoke about the qualities of a "good profession," suggesting their application to the field of psychology. A most fundamental quality of the professions, Sanford advanced, is a sense of social responsibility. A quarter of a century later, Peterson (1976a) invited our attention again to the characteristics of a profession, challenging us to think about the extent to which they are manifest in psychology, advancing towards that quality of an applied discipline. On all accounts, most will agree that psychology, in addition to being a scientific discipline, is a profession.

Professions are characterized by a differentiated and dynamic body of knowledge; a system of disciplined methods of inquiry and practice through which that body of knowledge is challenged, expanded or modified, and applied; and a core of shared values and attitudes that together speak to a sense of integrity and social responsibility. Professions are characterized as well by a sense of self-discipline and self-regulation, governed by standards of practice and ethical principles in accordance with which their members are subject to peer review.

The foundation of any profession is the requisite education and training that leads to entry level of practice. Consistent with the disciplinary principles just stated, it is the responsibility of a profession to establish education, training, credentialing, and practice standards or guidelines (Drum & Hall, 1993;

Merton, 1969). It must also ensure the validity of those standards or guidelines through regular, periodic evaluation (Kane, 1982; Koocher, 1979; Sechrest & Chatel, 1987; Stevenson & Norcross, 1987).

Typically, it is through accreditation that professions set the standards for education and training leading to their entry level of practice. As a self-regulatory, peer review process, accreditation serves to assess and publicly certify the quality of professional education and training relative to those standards. In so doing, accreditation models the self-critical and peer review qualities expected of professionals in the review of their practice. Thus, it might be argued, accreditation is a link between professional training and practice (Nelson & Aletky, 1987) beyond its value in certifying the quality of education and training for entry level into the profession. In these ways, accreditation and the professions are linked inextricably by a common orientation to public accountability, a sense of social responsibility.

The emphasis on foundational education and training in the professions is not intended to suggest that early inservice and lifelong career experiences in professional service, or their attendant continuing professional education and training programs are of little consequence in shaping the profession. It is at the foundational education and training level in any profession, however, that the qualities characteristic of a competent professional take root. It is at this level that students must be prepared in values and attitude as well as in knowledge and skills for lifelong learning in the discipline.

This chapter reflects on the developmental history of psychology as a scientific discipline and how that development in response to public need led to accreditation activities and the early seeds of a profession by the mid-twentieth century. In addressing these issues, the author intends to illustrate how accreditation has functioned as a forum in some instances, and a lightning rod in others, for national debate of professional education and training issues among and between psychologists within and outside the academy. Through this process, the accreditation standards reflect the thinking of psychologists about the goals and principles of professional education and training, the models or philosophies of training through which those ends are to be achieved, answers to the questions "what constitutes an appropriate curriculum?" and "what competencies do we expect of our graduates?" In the final analysis, the process of accreditation serves a public accountability function for the profession and the higher education community.

2.05.2 THE CONTEXT FOR ACCREDITATION IN THE PROFESSIONS

The early development of clinical psychology training in the scientific laboratory of a research university was not an accident of history, but rather an event consistent with changes that were taking place in leading research universities in the US at the same time. With the chartering of land-grant colleges and universities (Christy & Williamson, 1992), the turn of the nineteenth century was a period during which even the private research universities were conscious of their mission not only to create and disseminate knowledge within the community of scholars but to apply that knowledge to the benefit of society, a public service mission ... an element of mission on which the professions are founded.

Indeed, professional education and training through the universities was a natural conduit for this element of mission to society, perhaps exemplified best at the time of Lightner Witmer's laboratory clinic (Brotemarkle, 1947) by the transformation that was underway in American medical education (Starr, 1982). Drawing upon the scientific traditions of the great German universities, the seeds of modern, science-based medical education were sown in the likes of Johns Hopkins University, Harvard University, and the University of Pennsylvania in the late nineteenth century.

It was not only the university's role in bringing science to bear on medical education that was significant, however, but equally important was the university's alignment with teaching hospitals that served indigent populations of the community and those even from outside the local community whose health was at risk of being impaired by uncommon disease, the latter being the focus of medical research. Thus, in theory at least, alignment of the classroom, laboratory, and clinic afforded medical students first-hand clinical experience and involvement in medical research under the supervision of qualified faculty and clinical scientists. Prior to that time, and even subsequent thereto for some years, medical education was virtually without national standards. Entrepreneurial apprenticeships in the community were common as a way to receive clinical training, and medical schools were comparably without a standard curriculum or metric by which to gauge quality.

This was the background against which the medical profession carried out with support of the Carnegie Foundation a major assessment of its education and training practices, leading to an overhaul of American medical education and training in the early years of the twentieth century (Flexner, 1925). In addition to setting national standards, the profession developed procedures by which to review periodically each medical school, resulting in what was the earliest form of specialized accreditation applied to professional education and training. Within the first few decades of the twentieth century, accreditation practices were also applied to the professions of dentistry, law, and engineering. Now, there is no profession of note in American society without an accreditation process applied to its foundational, entry-level education, and training programs. Accreditation, despite its frailties, has had a profound influence on professional education and training and therefore upon the professions themselves (Young, Chambers, & Kells, 1983).

For reason of developmental differences between medicine and clinical psychology as science-based professional disciplines, accreditation in psychology was not to be implemented, let alone conceived, for another 50 years from the time of Witmer's initial training clinic. At the turn of the nineteenth century, psychology had no identity as a profession. It was a relatively new science, in search of an identity distinct from its parent bodies of philosophy and the physiological sciences. Although its applications began early in the likes of Witmer's clinic at the University of Pennsylvania and another applied laboratory run by Munsterberg and his students at Harvard focused on industrial work problems (Stern, 1917), most developments in the discipline of psychology at that time were of theoretical and scientific significance apart from practical applications.

Applications of psychology outside the university laboratory were stimulated by the involvement of academic psychologists in the nation's service during World War I (Yerkes, 1918). While the decades between the two world wars witnessed more applications of psychology to a broader array of social problems, even advancing major theoretical and methodological schools of thought as reflected in the work of Kurt Lewin and his colleagues (Marrow, 1969), the vast majority of psychologists in the US remained employed primarily in colleges and universities as teachers and researchers, devoted to advancing the science of psychology. Indeed, feeling disfranchised by their academic colleagues whose leadership and values pervaded the APA, the minority of psychologists whose primary employment was outside the academy during those years, psychologists providing primarily clinical and consulting services, formed their own national association in the 1930s, the American Association of Applied Psychology (AAAP).

Although the AAAP and APA were reunified by events of World War II, there remained a sense of skepticism that such unity would last (Hunt, 1975). It didn't, entirely. Nearly 50 years subsequent to the AAAP development, the American Psychological Society (APS) was formed to allow those whose commitments were primarily to science in the university, by then the minority of APA members, to have their own national association (though many remained members of APA too). Throughout it all, the APA remains the largest national association of psychologists with the broadest array of interests and identities across the science and profession of psychology.

In the context of this chapter, the point of these observations is to document how the tensions resulting from disparate value orientations and perspectives between those whose commitment first and foremost has been to advance psychology as a science and those whose comparable priorities have been to advance psychology as a profession, referred to by some as "the two cultures of psychology," have been a central force operative in this discipline in the twentieth century (Albee, 1970; Chein, 1966; Fernberger, 1932; Kimble, 1984; Shakow, 1978; Sheridan, 1990; Tryon, 1963). It is a force that has played a significant role in the foundational education and training for psychology as a profession and in the accreditation process which was to come, following World War II.

2.05.3 ACCREDITATION IN PSYCHOLOGY: THE ADVENT OF A PROFESSION

It was for reasons of public service and accountability that accreditation practices were applied initially to graduate education and training in psychology. With the return of veterans following World War II, and the shortage of doctoral-level clinical psychologists to provide the psychological services needed in Veterans Administration hospitals and other government agencies, the Veterans Administration, followed shortly by the US Public Health Service, requested of the APA Board of Directors in December 1945 a list of education institutions possessing facilities adequate for training clinical psychologists. The idea behind this request was that public funds would be directed in support of those programs that were appropriately qualified to carry out such training. The APA accepted this responsibility and assigned it to the Committee on Graduate and Professional Training which had been gathering data on graduate training in psychology (Sears, 1946).

This was a significant development in the history of psychology. It represented recognition by the public of the worth and need for professional services of psychologists trained at the doctoral level, based upon their demonstrated competence in a myriad of roles of public service during World War II. Furthermore, it represented a public invitation for the profession to accept responsibility for evaluating its education and training programs preparatory for entry-level practice. It was the dawn of a new era in psychology, one in which psychology would flourish not only as a science but also as a profession. In the decade prior to World War II, two-thirds or more of the new doctoral graduates in psychology each year found primary employment in academic settings (Clark, 1957). In the years following World War II, this was to change. Now, in the late 1990s, approximately the same proportion of doctoral graduates in psychology that once found primary employment in academic settings find primary employment in applied, professional service settings outside the academy (APA Research Office, 1997a).

The result of the APA's action in response to the Veterans Administration and US Public Health Service was a report one year later on clinical training facilities in psychology departments at 40 universities, the first published account of quality in education and training for practice in clinical psychology. The Committee on Graduate and Professional Training cautioned readers that the criteria it developed to evaluate quality of training in these academic departments were not necessarily the only possible criteria and that their application to diverse institutions could be problematic. Then it added:

> If accreditation continues to be desirable for the profession, new criteria will gradually develop. ... The present criteria must be viewed only as steps along the way; they are useful for measuring one stage in a type of professional training that is changing and growing every year. (Sears, 1947, p. 204)

The project was based on the application of published criteria for professional education and training through the process of peer review. With the additional involvement of the graduate departments, university administrations, and clinical program faculties in an analysis (self-study) of their capabilities and needs relative to the criteria, the project was the harbinger of accreditation in psychology.

It was important enough a development, moreover, for the APA Board of Directors to authorize appointment of a special committee on Training in Clinical Psychology in March 1947, the tasks of which would be to: (i) develop formal recommendations for education and training standards in clinical psychology for academic institutions and field service agencies that provide internship and practicum training; (ii) to study and visit the institutions that provide this training, preparing a report on each; and (iii) to maintain communication with other national associations or committees having related missions or interests.

The new committee was chaired by David Shakow and presented its recommendations on clinical training in psychology at the annual meeting of the APA in September 1947. Approved by the APA Council of Representatives, the recommendations became the first official accreditation standards and criteria for psychology (APA Committee on Training in Clinical Psychology, 1947). A year later, the committee published its first report of the outcomes of its review of clinical psychology programs in operation at that time, a process that included site visits by members of the Committee and others who assisted (APA Committee on Training in Clinical Psychology, 1948). Clinical psychology training programs in 36 universities were listed as the initial accredited programs under the new criteria.

From the beginning, it was clear that the committee was sensitive to the importance of the university's involvement in the accreditation review process, and that the procedures for such review should be valid for the purposes intended without being unduly burdensome on either the universities or the Committee. The Committee also understood that the profession of psychology was in its infancy and that education and training for the profession must be afforded opportunity to develop along different lines, even perhaps different philosophies in time. To wit, the committee set forth its stated purpose as follows:

> Our aims are ... to achieve general agreement on the goals of training and encourage experimentation on methods of achieving these goals and to suggest ways of establishing high standards in a setting of flexibility and reasonable freedom. We also hold that the goals should not be determined by special situations and special demands, but should be oriented toward the question of what is the best training in clinical psychology. (APA Committee on Training in Clinical Psychology, 1947, p. 543)

Even with this broad goal, however, in its first year or two of visiting and evaluating university programs in clinical psychology, the committee discovered that any enthusiasm for accreditation that may have emanated from the APA leadership during that time was not shared universally among university department chairs and others of the academic institutions responsible for the training of psychologists. Shakow and his committee colleagues shared their experience with the accreditation process by reflecting:

> During our visits to universities, we have found attitudes towards the committee's work which varied all the way from those of hearty acceptance, confidence, and appreciation, to those of honest conviction that the committee is a usurper of power and authority and, that by the use of its power and its methods, does damage to the work of graduate departments.
>
> We are convinced of the importance of setting standards and evaluating performance in clinical training—in fact we see this process as inevitable if clinical psychology is to establish itself soundly and be a credit to psychology as a whole. We have set up criteria and administered procedures in ways that seemed best; we are neither satisfied that they are perfect nor persuaded that they are dangerous or ineffective. It is natural that there should be differences of opinion in the Association regarding the goals and techniques we have accepted. Both these differences and our own convictions are still matters of opinion rather than demonstrated fact. (APA Committee on Training in Clinical Psychology, 1949, pp. 340–341)

In sharing this early experience and in their prophesy of the future, the Shakow committee correctly anticipated tensions and challenges that were to characterize the accreditation of professional education and training programs in the decades ahead. As the Sears committee before it had partially forecast, Shakow's committee also envisioned that the process of accrediting programs in professional (clinical) psychology would change over time with regard to the criteria and procedures developed. They knew, in addition, that without a foundation of broad goals and principles to guide the accreditation process, the criteria and procedures in effect at any time could be misapplied, misunderstood, or otherwise impotent. Most important, being of a scientific discipline themselves, they knew how important it would be to evaluate outcomes of graduate professional education programs and of the accreditation process itself. That task was their legacy to future generations of accreditors.

2.05.4 GOALS AND PRINCIPLES OF PROFESSIONAL EDUCATION AND TRAINING

Given the history of this discipline and its state of development at the time of the Shakow committee's work, it is understandable that their approach to developing foundational education and training standards for clinical psychology was guided by the following considerations:

> We are cognizant of the great difficulties which the shift from an academic to a professional program involves in a university setting. We recognize that this change must take much effort and time and that even were it possible to set up a fairly fixed schedule of training, such a step would at present be both premature and ill-advised because of the great need for experimentation in ways of implementing a sound program. We are therefore emphasizing the goals and principles of what we consider a desirable program rather than attempting to lay out a detailed blueprint. (APA Committee on Training in Clinical Psychology, 1947, pp. 539–540)

In an account of the history of graduate professional education and training in psychology (Sheridan, Matarazzo, & Nelson, 1995), two major concerns were noted as the principal focus of debates on accreditation over the past 50 years. One is whether and how education and training in the science and clinical practice of psychology can coexist in one program. The other salient issue has been whether accreditation, with its externally-imposed standards and criteria, inherently violates or otherwise compromises the spirit of innovation and free inquiry that is *sine qua non* of the academy and, in the case of psychology, its science faculty and students.

Certainly the goals set forth by Shakow and his colleagues were sufficiently broad, recommending high standards but with flexibility of perspective to accommodate different ways by which those standards could be achieved. Still, the committee needed a bridge or transfer function by which to link the broad goals of accreditation to the development of standards for professional education and training and the process by which the attainment of those standards could be measured. The bridge they found in the earlier work of Flexner (1925, p. 176):

> ... the medical school cannot expect to produce fully trained doctors; it can at most hope to equip students with a limited amount of knowledge, to train them in the method and spirit of scientific medicine and to launch them with a momentum that will make them active learners-observers, readers, thinkers, and experimenters—for years to come ...

to which Shakow and his colleagues added:

> If we substitute clinical psychology for medicine, this statement expresses the essential point which we wish to make in this report. (APA Committee on training in Clinical Psychology, 1947, p. 540)

From this base the Shakow committee set forth the following (paraphrased) principles to serve as standards of professional education and training in psychology and to guide the development of accreditation criteria by which the attainment of those standards could be measured:

(i) A clinical psychologist must first and foremost be a psychologist ... having a point of view and core of knowledge and training which is common to all psychologists.

(ii) Education for the doctorate in clinical psychology should be as rigorous and extensive as that for the traditional doctorate ... combining academic and clinical training.

(iii) Preparation should be broad; it should be directed to research and professional goals, not to narrow technical goals ... with psychologists as a research worker emphasized.

(iv) There should be training in: (a) general psychology; (b) psychodynamics of behavior; (c) diagnostic methods; (d) research methods; (e) related disciplines; and (f) therapy.

(v) Programs should be concerned mainly with basic courses and principles rather than multiple courses in technique, with advanced courses requiring knowledge from earlier courses.

(vi) Courses should be scrutinized for their content rather than judged by their titles. Equally important is ... the quality of the teaching and the integration of courses.

(vii) Instruction should be organized around a careful integration of theory and practice, of academic and field work, by persons representing both aspects (of training).

(viii) Throughout all years of graduate work the student should have contact, both direct and indirect, with clinical (and normal) material ... allowing a range of experience.

(ix) The atmosphere of training should encourage the increase of student maturity, a sense of responsibility for patients and clients, and excitement about psychological problems.

(x) There should be opportunity to study with students of other disciplines ... to acquire a sense of modesty and to learn the value of teamwork in research and service.

(xi) Throughout training there should be an emphasis on the research implications of phenomena being addressed ... with the set to ask "how," "why," "what is the evidence?"

(xii) In addition to the research implications of data, students need to develop a sensitivity to the social implications ... not only to the individual patient or client but to society. (APA Committee on Training in Clinical Psychology, 1947, pp. 543–545).

In the criteria developed for review of programs, these principles were embellished with extensive guidelines and precautions about academic and clinical training, including the year of internship required for completion of doctoral training in clinical psychology. The most important feature of the committee's work, however, was establishing the foundation of broad goals and guiding principles for professional education and training in psychology at the doctoral level. These goals and principles have stood the test of time and remain in effect, essentially, as the foundation upon which the current accreditation guidelines and related program review procedures are based (Committee on Accreditation, 1996).

These principles of education and training in professional psychology reflect more than the thinking of small committees. They have been endorsed and reaffirmed, perhaps in modified form to reflect the development of the field, in each of the past five decades by national conferences of psychologists, educators, and practitioners, convened for different purposes at different times. The first of those was the 1949 Boulder Conference with its focus on clinical training (Raimy, 1950); the most recent on graduate education broadly was the 1987 Utah Conference (Bickman & Ellis, 1990).

2.05.5 MODELS OR PHILOSOPHIES OF PROFESSIONAL EDUCATION AND TRAINING

Conceptually, it is through a model or philosophy of education and training that educators translate their broad, more abstract goals and principles of education and training to a pedagogical plan by which the learning process and attainment of education and training goals can be realized. While there are many ways in which this can be done in programs of professional psychology, perhaps the most salient parameter of conceptualization historically has been that which ties the "model of training" to psychology's two cultures, science and practice. This is not surprising in the context of psychology's history.

While many resolutions about graduate education in clinical psychology were articulated at the Boulder Conference, and concerns were aired about the new accreditation process, most remembered from that conference over the years was its endorsement of the Shakow committee principles and what became known as the "scientist–practitioner" or "scientist–professional" model of education and training in clinical psychology. The spirit of the "scientist–practitioner" philosophy or model, however, had been expressed a few years earlier. An APA committee chaired by Robert Yerkes, in planning for the future of the association, had professed "that the development of psychology as a science and practice should proceed in close relation; and that the scientists of the profession should welcome technological developments and aid and encourage them" (Boring et al., 1942, p. 624).

Subsequent to the Boulder conference, the "scientist–practitioner" model was reinforced at other national conferences on professional education (Hoch, Ross, & Winder, 1966) and graduate education (Roe, Gustad, Moore, Ross, & Skodak, 1959), as well as by national conferences focused on the preparation of school psychologists (Cutts, 1955) and counseling psychologists (Thompson & Super 1964). The model itself was the focus of a conference (Belar & Perry, 1991). How, if at all, science and practice inform each other has been a fundamental issue in the development of psychology as a science-based profession (Barlow, 1981; Belar, 1990) and remains a question of philosophical and pragmatic importance (Stricker, 1997).

The "scientist–practitioner" model has been implemented in different ways among doctoral programs over the years (Henry, 1984). Among accredited programs, there are some that truly achieve an integration of research and practice in the thinking and activities of students and faculty. In others espousing the same model, some with more emphasis on practice, some with more emphasis on science, and some with equal emphasis on both, there is at times bifurcation of the two domains and their respective faculties. It remains, nonetheless, the most commonly espoused model of training among accredited programs today. Indeed, some argue that there is no need for alternative training models (Perry, 1979).

Not all have argued that point, however (Fox, Barclay, & Rodgers, 1982; Peterson, 1976b, 1991; Sechrest, 1990), nor did the delegates of the 1973 Vail Conference (Korman, 1973). Although several of the issues raised at that conference had been discussed at earlier conferences or had been the focus of APA committee work during the 1960s, the Vail Conference clearly endorsed the idea of a professional degree in psychology, the Psy.D. degree, and the value of professional schools of psychology, in addition to the traditional academic Ph.D. degree granted by universities in recognition of scholarly achievement in a discipline.

The model more in keeping with this orientation to professional education and training has been most commonly labeled "professional-scholar" or "practitioner-scholar," connoting an emphasis on preparation for professional practice with the scholarly perspective and habits of an inquiring mind appropriate to the clinical practitioner. The orientations of the practitioner of a profession, and therefore fundamental to professional education and training, it is argued, are different from the orientations of science and scientists (Peterson & Peterson, 1997). Comparative analyses of accredited programs across the spectrum of models reveal differences in admission requirements, acceptance rates, financial assistance, and theoretical orientation (Mayne, Norcross, & Sayette, 1994), activity patterns among students and faculty, as well as time to degree (Gaddy, Charlot-Swilley, Nelson, & Reich, 1995), and differences in professional activity and perspective about their work among graduates (Conway, 1988; Peterson, 1985).

A second theme advanced by the Vail Conference was the need for professional psychologists to be educated and trained with a sensitivity to and understanding of individual and cultural diversity in human lives, a theme consistent with the growing awareness and demographic projections of the ethnic, racial, and cultural diversity of American society and the importance of that fact for psychological research and services. That theme was to become even more pronounced during the 1980s, with considerations of human diversity expanded as well to include sensitivity to issues of gender, sexual orientation, disabilities, and religion. These developments were an expression of the profession's increasing sense of social responsibility in response to events of society during one of the most turbulent decades in US history.

Among the concerns was the perceived public need for many more psychologists to provide mental health services than could be prepared through traditional university academic departments, the corresponding need for their training to be focused on professional issues of practice as well as on the foundational principles set forth earlier by the Shakow committee, and the need for psychologists to have a personal and professional sensitivity to the qualities of individual and cultural diversity that characterize populations requiring psychological services. The accreditation criteria rewritten and approved in the late 1970s recognized and included these two major Vail Conference outcomes, expressing them in the form of goals and principles (American Psychological Association, 1979).

2.05.6 ACCREDITATION STANDARDS: CURRICULUM AND COMPETENCE CONTROVERSY

Another national working conference during the 1970s that influenced the way in which the

1979 accreditation criteria were written was one focused on education and credentialing in psychology (Wellner, 1978). The issue of particular concern at the conference was how psychology as a discipline can be distinguished from other disciplines often confused with psychology by the public. Of particular concern to some was the observation that those having a doctoral degree in psychology do not achieve that academic status through a common curriculum.

Historically, most graduate students in doctoral programs of psychology had a common core of learning experiences including basic textbooks in their first year or two of study (Matarazzo, 1987). If this was the case at least through the 1950s or 1960s, however, it was due to the fact that it made sense to individual faculties over the years and not because graduate departments of psychology endorsed as policy a nationally standardized core curriculum as such. As Ericksen (1958, p. 58) put it, following a survey of graduate department faculties on the value of a core curriculum: "It is quite apparent that we value variability, experimentation, and educational freedom far more than we value conformity and standardization."

While this value remains strong among ranks of the academy, it was to pose serious problems by the 1970s for licensing boards of examiners in psychology, especially in clarifying formal distinctions between the educational qualifications of psychologists and those of licensure applicants having graduated from other than psychology departments or programs. By this time, in clinical psychology, the numbers of programs of different models were increasing and, with an expanding body of psychological knowledge, there were varied curriculum options for students even at times in the same program. Although psychology had always taken pride in its scientific base and the corollary research training expected of those who would be recipients of the doctoral degree, this argument of psychology's distinction from seemingly related disciplines failed to convince the critics and in some instances could not be substantiated in graduate education and training for all licensure candidates.

The result of the 1976 conference on education and credentialing in psychology was the development of a foundational curriculum for the scientific discipline of psychology, the presence of which in a graduate program of study would allegedly designate that program as one in psychology. In did not specify what courses, textbooks, or other media of learning were to be required, but rather the substantive areas of the scientific discipline in which all students of professional psychology at least must demonstrate competence at a graduate level of understanding. The curriculum areas to serve as a foundation for understanding psychology as a scientific discipline were: history of the discipline; research methods and statistics; ethics and professional issues; and each of biological, cognitive-affective, social, and individual bases of behavior. They were incorporated in the accreditation criteria of the late 1970s, with the additional requirement of the core applied areas of competency, such as psychodiagnostics, psychological assessment, intervention procedures, consultation, and program evaluation (American Psychological Association, 1979).

These were not new curriculum foundations for graduate professional education in psychology, of course; they had been articulated in one form or another in the earlier criteria developed by the Shakow committee and in the criteria revised in the early 1970s at which time emphasis was placed on diversity of program models and the evaluation of program quality in the context of a program's chosen model (American Psychological Association, 1973). What seemed most important, however, was the consensus about curriculum areas achieved among those representing the perspectives of accreditation and credentialing in professional psychology. As a consequence, the core curriculum areas included in the accreditation criteria (American Psychological Association, 1979) were those also adopted as "designation criteria" by the Association of State and Provincial Psychology Boards and the National Register of Health Service Providers in Psychology in carrying out their individual psychologist credentialing review functions. The notion was that these curriculum criteria would clarify operationally whether or not an applicant for licensure or other certification as a psychologist had in fact completed a doctoral program in psychology, as distinct from another discipline (Council for the National Register of Health Service Providers in Psychology, 1997). It was clearly the expectation that the accreditation process would ensure at least some standardization in the curriculum areas in which all students in accredited professional psychology programs would be expected to demonstrate competence, if not identical coursework or other experience. Was that objective achieved? The answer to this question resides in the mind of the beholder, to be sure.

Certainly the implementation of accreditation criteria revised in the late 1970s did not result in programs having a standard curriculum at the course and textbook level or in the form of a standard lock-step curriculum plan, nor was that intended. On the other hand, graduates of accredited programs performed better than graduates of unaccredited programs on examination for licensure as a psychologist, much of the examination being based on these foundational areas of curriculum (McGaha & Minder, 1993; Ross, Holzman, Handal, & Gilner, 1991). Moreover, in a practice analysis survey of licensed psychologists conducted by the Professional Examination Service on behalf of the Association of State and Provincial Psychology Boards, most practitioners rated the foundational curriculum areas cited in the accreditation criteria as being at least moderately important and useful in their practice (Greenberg, Smith, & Muenzen, 1996).

While not altogether surprising, a disappointing exception to that finding was the observation among licensed practitioners that their training in "research and evaluation" seemed to be less important and used less frequently in practice than other substantive areas of the discipline's scientific foundations. When considered in the context of criticism previously targeted to graduate training in research methods and measurement (Aiken et al., 1990; Lambert, 1991), this finding raises significant questions about the one area of competence in which psychology traditionally took pride and differentiated its graduates from the professional service providers of other disciplines.

2.05.7 STORM WARNINGS IN ACCREDITATION: TIME FOR CHANGE OF HEADING

The emphasis on core curriculum issues of the 1970s and 1980s was one over which the accreditors were challenged on the one hand by their academic colleagues and on the other by their colleagues in practice, the former predominantly advocating against such a requirement imposed by the accreditors and the latter predominantly in favor of such. Academicians charged that the accreditation process was characterized by what some referred to as a "checklist mentality" having nothing to do with quality of professional education and training. This allegation within psychology was not unlike the charge against other accrediting bodies of the time that their concerns were focused mostly on resources and structures of education, not on the quality of outcomes.

Other concerns about the quality of training in clinical psychology, particularly, were raised with regard to the conceptual orientations (Albee, 1970, 1983; Levy, 1984; Sarason, 1981). At yet another level of criticism, there were those calling for something akin to the "Flexner report" to revamp the curriculum in professional psychology (Strickland, 1985) over concern that by allowing so many different models and programs of psychology to develop there was no standard quality of outcome. A set of core curriculum requirements and clinical competencies shared by all who complete professional education and training, some argued, is the hallmark of a profession (Fox, 1982, 1994; Fox, Kovacs, & Graham, 1985).

Together, the academicians and the practitioners each wanted a more active voice in the accreditation process. There were overtures made about alternative accrediting bodies for different types of programs and questions raised as in decades before about the necessity of accreditation. Whatever concerns about professional education and training were at issue in public debate at meetings of psychologists, the problems cited could most always be laid at the feet of the accreditation process. It was not a quiet time! This period of turbulence for accreditation in psychology had historical roots at least 50 years in the making. If accreditation applied to its professional education and training programs was to have even a chance of addressing constructively the various forces of tension that have been a part of the discipline's history, clearly needed was a different form of organization and operating procedures for the accrediting body itself.

The national context for change in accreditation was also right at the time, for psychology had not been alone in its dilemma by the mid-1980s. Criticism of accreditors and their procedures of quality assessment was gaining momentum in most professions and was being directed at the regional accreditors of postsecondary education institutions as well. The issue of "who controls the accreditors" and related matters of public accountability in accreditation and higher education in general provided a national context in the late 1980s for considerable change in American accreditation circles (Millard, 1991). Higher education institutions themselves were not exempt from public criticism during the same period over concerns for public accountability (Bok, 1990; Ewell, 1994; Wilshire, 1990).

In an effort to bring psychology educators and practitioners together around these issues, a significant development was the convening in 1988 of an *ad hoc* group, the Joint Council on Professional Education in Psychology (JCPEP). The report produced two years later by that group was an integrative summary of the thinking and recommendations generated by different national conferences in the 1980s, national organizations of psychology educators and practitioners, and the mid-1980s work of the APA Task Force on Scope and Criteria for

Accreditation (Stigall et al., 1990). Perhaps most important in developing the report and its recommendations was the strong degree of consensus that had been reached by constituencies that theretofore had often been at variance in their views about accreditation and its role in professional education and training. It was in effect a precursor of what was to be a new accreditation body, the members of which were to represent much the same diversity of constituencies as had the task force.

Thus, following three years of negotiation, a newly constituted Committee on Accreditation was commissioned, a multiorganizational body that would represent the different perspectives of educators, practitioners, and publics served by professional education and training in psychology and the process of accreditation by which the quality of those programs is assessed and certified (Sheridan, Matarazze, & Nelson, 1995). The Committee on Accreditation would be responsible not only for the quality assesment of programs in professional education and training, but also for the formulation and promulgation of accreditation policies and procedures (American Psychological Association, 1991).

The new Committee on Accreditation set to work in the fall of 1992 on a major self-study of its own. The outcome three years later, following more than a year of iterative public comment on earlier drafts of proposed changes, was the unanimous approval by the committee members of a new set of guidelines and procedures for the accreditation of programs in professional psychology (Committee on Accreditation, 1996). Although compromises were necessary along the way, and anyone could find some aspect of the guidelines and procedures with which to quarrel, all major national organizations representing psychologist educators and practitioners signed on, having had ample opportunity to participate in the process that led to the final guidelines.

In setting about its task of designing new accreditation guidelines, the Committee on Accreditation surveyed all communities affected by accreditation to assess their concerns about accreditation and ideas of how it could be improved. The committee also learned from the experiences of other groups convened during the 1980s to propose solutions to some of the problems that had been experienced in accreditation. In addition to national conferences on graduate and professional education and training, significant among those efforts had been the meta self-study process initiated by the National Council of Schools and Programs in Professional Psychology (NCSPP) whose annual conferences and resulting resolutions about professional education and training have been summarized (Peterson, Peterson, Abrams, & Stricker, 1997).

2.05.8 OUTCOMES OF ACCREDITATION: ISSUES OF PUBLIC ACCOUNTABILITY

Most important, reflecting the wisdom of the Shakow committee nearly 50 years earlier, the new accreditation guidelines were framed by a set of guiding principles for professional education and training and a set of goals that recognized simultaneously the needs and values of the academy in its role and those of the profession in being publicly accountable. As its predecessor accrediting body 20 years earlier had seen fit to do (American Psychological Association, 1973), the new committee developed guidelines that would assess program quality in the context of a program's publicly stated education and training goals and objectives. Programs are to demonstrate how they expect to achieve those ends in terms of the resources available to them, their chosen curriculum plan (including clinical training), and the types of competencies they expect of their program graduates.

These curriculum and competency outcomes are framed at a general level in terms of foundational areas of training in psychology as a science and applied (health/human service) discipline. A particularly critical addition to these constructs, however, is the objective of developing in students "attitudes essential for life-long learning, scholarly inquiry, and professional problem-solving as psychologists in the context of an evolving body of scientific and professional knowledge" (Committee on Accreditation, 1996, p. 6).

The most significant departure from earlier accreditation standards and criteria for doctoral and internship programs, however, is the public accountability standard that "the program demonstrates a commitment to excellence through self-study, which assures that its goals and objectives are met, enhances the quality of professional education and training obtained by its students, and contributes to the fulfillment of its sponsor institution's mission" (Committee on Accreditation, 1996, pp. 9, 16).

Moreover, in demonstrating that they meet this standard, accredited programs must demonstrate through this self-study analysis effectiveness in achieving their goals and objectives, and the procedures they have in place to make program adjustments in their goals, resources, and education and training structures that may be deemed necessary by the outcomes analyses conducted. In addition, doctoral programs are expected to evaluate the appropriateness of their goals, objectives, training model, and curriculum in the context of: (i) the sponsor institution's mission and goals; (ii) local, regional, and national needs for psychological services; (iii) national standards of practice; (iv) the evolving scientific and professional knowledge basis of practice; and (v) its graduates' job placements and career paths. Internship programs are accountable in analogous ways. These are aspects of public accountability, the need for which has never been greater.

Accountability to students preparing for careers in psychology is a matter to which graduate programs must be sensitive, given the cost of education and training and the changing employment market for graduates. This matter has become a focus of considerable concern in fields of science (Tobias, Chubin, & Aylesworth, 1995). For science careers in psychology, this concern prompted the APA to form a task force in 1996, the purpose of which was to develop information about career opportunities for research psychologists in other than traditional academic departments, an update on the work of a comparable task force appointed a decade earlier (Klatzky, Alluisi, Cook, Forehand, & Howell, 1985).

There is likewise a changing market of employment in the health and human services areas for psychologists, reflected in the trend away from independent practice and what has been historically a predominant identification with mental health services towards employment in multidisciplinary settings responsible for primary and tertiary care in the full range of health services (APA Research Office, 1997; Glueckauf, Frank, Bond, & McGrew, 1996; Stone et al., 1987). The requirement that psychologists be able to function effectively with professionals of other disciplines on clinical or other forms of problem-solving teams is not a new employment outcome for psychologists of course. It was perceived indeed to be the form of employment most useful for clinical psychologists at the time of the Shakow committee's work. They referred to the professional person as one who "learns the techniques and importance of the group attack on problems: how best to work with other professional groups for the benefit of the individual client and the frequent necessity of identifying himself with a group even broader than his own professional group, namely, the 'team'" (APA Committee on Training in Clinical Psychology, 1947, pp. 556–557).

Beyond the ability to function as part of a multidisciplinary team, psychologists employed in healthcare services in the 1990s also need to understand their role and the potential contributions that they can make in an environment characterized by organized care systems that are operated on a cost-control basis (Friedman, Sobel, Myers, Caudill, & Benson, 1995; Groth-Marnat & Edkins, 1996). This requires some additional professional knowledge and skills, but also adaptation of the foundational knowledge and skills traditional to broadly-trained clinical psychologists to a different service delivery model from that since the 1960s. Psychologists have considerably more to contribute than they have in the maintenance of health and prevention of illness and injury, a major theme of the managed-care era. In that same vein, psychologists have much to contribute to public health services, through the applications of psychological knowledge to shaping and evaluating public policies on education, health, and other human services (Lorion, Iscoe, DeLeon, & VandenBos, 1996).

All things need to be considered in a continuously changing employment market; however, graduates best prepared will be those whose education and training exemplifies the principles of breadth and foundational competencies that Shakow and his committee first articulated, principles that remain in effect in the 1990s. These competencies must include the critical-thinking, problem-solving skills that psychologists bring to bear on problems of human behavior, the types of skills to which psychology's adaptiveness to changing employment markets of the past have been attributed. Indeed, despite the concerns of students and others about the future employment market for psychologists, the graduates of accredited professional education and training programs in psychology are employed in a myriad of settings appropriate to their professional training (Gaddy, Charlot-Swilley, Nelson, & Reich, 1995), also reflected in the diversity of settings in which clinical psychology graduate students take internship training (Sheridan et al., 1995). Thus, the significant increase in numbers of clinical psychologists since the 1960s attributable to the increase in professional schools of psychology (Strickland, 1985) has been accompanied by an increase in the types of settings and roles in which psychologists function. The question at hand is whether this pattern will continue.

Although the answer to this question lies in the future, there are early career employment indicators among graduates of doctoral programs in psychology that bear close monitoring in the years immediately ahead. Within a year of graduation, for example, while unemployment rates remain low, they have increased slightly and graduates are requiring somewhat more time to find initial employment than their counterparts a few years earlier. The proportion of graduates entering postdoctoral residency or fellowship positions is increasing, a pattern that bears closer examination, and proportionately more graduates report that their knowledge and skills are somewhat underutilized in their primary employment settings than was true for earlier years (APA Research Office, 1997a, 1997b; Kohout & Wicherski, 1993; Wicherski & Kohout, 1995).

Much of this undoubtedly reflects the changing market for all. Nonetheless, they should not go unheeded, especially by those who prepare students for the future, the educators. The question "Are graduate departments responsible for the marketability of their graduates?" was one posed of participants at the national conference on graduate education in psychology in the late 1980s (Grosslight, 1990), a question just as timely in the late 1990s. Through the accreditation process, professional education and training programs are expected to evaluate the effectiveness of their graduates. The test for the accreditors is the extent to which they hold accredited programs accountable to such expectations and what effect that has on professional education and training in psychology and the outcomes thereof.

"The real answer to the question 'Shall we evaluate?' lies somewhere in the future," the Shakow committee stated (APA Committee on Training in Clinical Psychology, 1949, p. 341). That future is now! Professional programs of education and training in psychology, their host institutions, the profession, the students and graduates of the programs, and the other publics served by all of the preceding shall be the evaluators. All of these vital perspectives are ensured formal representation on the accrediting body in psychology, for the first time in the 50 years of psychology's experience with accreditation. Thus, the accreditation process serves as a forum through which outcomes of professional education and training, and those of the accreditation process itself, can be assessed and evaluated against standards of public accountability—the hallmark of a profession.

2.05.9 REFERENCES

Aiken, L. S., West, S. G., Sechrest, L., Reno, R. R., Roediger, H. L. III, Scarr, S., Kazdin, A. E., & Sherman, S. J. (1990). Graduate training in statistics, methodology, and measurement in psychology. *American Psychologist, 45,* 721–734.

Albee, G. (1970). The uncertain future of clinical psychology. *American Psychologist, 25,* 1071–1080.

Albee, G. (1983). Preventing psychopathology and promoting human potential. *American Psychologist, 37,* 1043–1060.

American Psychological Association (1973). *Criteria for accreditation.* Washington, DC: APA.

American Psychological Association (1979). *Criteria for Accreditation.* Washington, DC: APA.

American Psychological Association (1991). *Policies for accreditation governance.* Washington, DC: APA.

APA Committee on Training in Clinical Psychology (1947). Recommended graduate training programs in clinical psychology. *American Psychologist, 2,* 539–558.

APA Committee on Training in Clinical Psychology (1948). Clinical training facilities: 1948. *American Psychologist, 3,* 317–318.

APA Committee on Training in Clinical Psychology (1949). Doctoral training programs in clinical psychology: 1949. *American Psychologist, 4,* 331–341.

APA Research Office (1997a). Preliminary results: 1995 doctorate employment survey. *Personal communication.* Washington, DC: APA.

APA Research Office (1997b). Full-time employment settings in 1995 for APA members awarded Ph.D.s in clinical psychology in 1985 and 1990. *Personal communication.* Washington, DC: APA.

Barlow, D. H. (1981). On the relation of clinical research to clinical practice: current issues, new directions. *Journal of Consulting and Clinical Psychology, 49,* 147–155.

Belar, C. D. (1990). Continued integration of scientific and practitioner graduate education in psychology. In L. Bickman & H. Ellis (Eds.), *Preparing psychologists for the 21st century: proceedings of the national conference on graduate education in psychology* (pp. 285–299). Hillsdale, NJ: Erlbaum.

Belar, C. D., & Perry, N. W., Jr. (Eds.) (1991). *Proceedings of the national conference on scientist-practitioner education and training for the professional practice of psychology.* Sarasota, FL: Professional Resource Press.

Bickman, L. J., & Ellis, H. (Eds.) (1990). *Preparing Psychologists for the 21st century: proceedings of the national conference on graduate education in psychology.* Hillsdale, NJ: Erlbaum.

Bok, D. (1990). *Universities and the future of America.* Durham, NC: Duke University Press.

Boring, E. G., Bryan, A. I., Doll, E. A., Elliot, R. M., Hilgard, C., Stone, P., & Yerkes, M. (1942). First report of the subcommittee on survey and planning for psychology. *Psychological Bulletin, 39,* 619–630.

Brotemarkle, R. A. (1947). Fifty years of clinical psychology: 1896–1946. *Journal of Consulting Psychology, 11,* 1–4.

Chein, I. (1966). Some sources of divisiveness among psychologists. *American Psychologist, 21,* 333–342.

Christy, R. D., & Williamson, L. (Eds.) (1992). *A century of service: land-grant colleges and universities, 1890–1990.* New Brunswick, NJ: Transaction Publishers.

Clark, K. E. (1957). *America's psychologists: a survey of a growing profession.* Washington, DC: APA.

Committee on Accreditation (1996). *Guidelines and principles for accreditation of programs in professional psychology.* Washington, DC: APA.

Conway, J. B. (1988). Differences among clinical psychologists: scientists, practitioners, and scientist-practitioners. *Professional Psychology: Research and Practice, 19,* 642–655.

Council for the National Register of Health Service Providers in Psychology (1997). *National Register of Health Service Providers in Psychology* (12th ed., pp. 3–4). Washington, DC: CNRHSP.

Cutts, N. (Ed.) (1955). *School psychologists at mid-century: a report of the Thayer Conference on the functions, qualifications, and training of school psychologists.* Washington, DC: APA.

Drum, D. J., & Hall, J. E. (1993). Psychology's self-regulation and the setting of professional standards. *Applied & Preventive Psychology, 2,* 151–161.

Ericksen, S. C. (1958). The core curriculum is a dependent variable. *American Psychologist, 13,* 56–58.

Ewell, P. T. (1994). A matter of integrity: accountability and the future of self-regulation. *Change, 26,* 24–29.

Fernberger, S. W. (1932). History of the American Psychological Association. *Psychological Bulletin, 29,* 1–89.

Flexner, A. (1925). *Medical education.* New York: Macmillan.

Fox, R. E. (1982). The need for reorientation of clinical psychology. *American Psychologist, 37,* 1051–1057.

Fox, R. E. (1994). Training psychologists for the twenty-first century. *American Psychologist, 49,* 200–206.

Fox, R. E., Barclay, A. G., & Rodgers, D. A. (1982). The foundations of professional psychology. *American Psychologist, 37,* 306–312.

Fox, R. E., Kovacs, A. L., & Graham, S. R. (1985). Proposals for a revolution in the preparation and regulation of professional psychologists. *American Psychologist, 40,* 1042–1050.

Friedman, R., Sobel, D., Myers, P., Caudill, M., & Benson, H. (1995). Behavioral medicine, clinical health psychology, and cost offset. *Health Psychology, 14,* 509–518.

Gaddy, C. D., Charlot-Swilley, D., Nelson, P. D., & Reich, J. N. (1995). Selected characteristics of accredited programs in psychology: a review of outcomes. *Professional Psychology: Research and Practice, 26,* 507–513.

Glueckauf, R. L., Frank, R. G., Bond, G. R., & McGrew, J. H. (Eds.) (1996). *Psychological practice in a changing health care system.* New York: Springer.

Greenberg, S., Smith, I. L., & Muenzen, P. M. (1996). *Study of the practice of licensed psychologists in the United States and Canada.* New York: Professional Examination Service.

Grosslight, J. (Ed.) (1990) Are graduate departments responsible for the marketability of their graduates? In L. Bickman and H. Ellis (Eds.), *Preparing psychologists for the 21st century* (pp. 163–173). Hillsdale, NJ: Erlbaum.

Groth-Marnat, G., & Edkins, G. (1996). Professional psychologists in general health care settings: a review of the financial efficacy of direct treatment interventions. *Professional Psychology: Research and Practice, 27,* 161–174.

Henry, B. (1984). The future of clinical training: forward into the past. *The Clinical Psychologist, Winter,* 25–26.

Hoch, E. L., Ross, A. O., & Winder, C. L. (Eds.) (1966). *Professional preparation of clinical psychologists.* Washington, DC: APA.

Hunt, W. A. (1975). Clinical psychology in 1944–45. *Journal of Clinical Psychology, 31,* 173–178.

Kane, M. T. (1982). The validity of licensure examinations. *American Psychologist, 37,* 911–918.

Kimble, G. A. (1984). Psychology's two cultures. *American Psychologist, 39,* 833–839.

Klatzky, R. L., Alluisi, E. A., Cook, W. A., Forehand, G. A., & Howell, W. C. (1985). Experimental psychologists in industry: perspectives of employers, employees, and educators. *American Psychologist, 40,* 1031–1037.

Kohout, J., & Wicherski, M. (1993). *1991 doctorate employment survey.* Washington, DC: APA.

Koocher, G. P. (1979). Credentialing in psychology: close encounters with competence? *American Psychologist, 34,* 696–702.

Korman, M. (Ed.) (1973). *Levels and patterns of professional training in psychology.* Washington, DC: American Psychological Association.

Lambert, N. M. (1991). The crisis in measurement literacy in psychology and education. *Educational Psychologist, 26,* 23–35.

Levy, L. H. (1984). The metamorphosis of clinical psychology. *American Psychologist, 39,* 486–494.

Lorion, R. P., Iscoe, I., DeLeon, P. H., & VandenBos, G. R. (Eds.) (1996). *Psychology and public policy: balancing public service and professional need.* Washington, DC: APA.

Marrow, A. J. (1969). *The practical theorist: the life and works of Kurt Lewin.* New York: Basic Books.

Matarazzo, J. D. (1987). There is only one psychology, no specialties, but many applications. *American Psychologist, 42,* 893–903.

Mayne, T. J., Norcross, J. C., & Sayette, M. A. (1994). Admission requirements, acceptance rates, and financial assistance in clinical psychology programs. *American Psychologist, 49,* 806–811.

McGaha, S., & Minder, C. (1993). Factors influencing performance on the Examination for Professional Practice in Psychology (EPPP). *Professional Psychology: Research and Practice, 24,* 107–109.

Merton, R. K. (1969). The functions of the professional association. *American Journal of Hospital Pharmacy, 26,* 636–641.

Millard, R. M. (1991). *Today's myths and tomorrow's realities.* San Francisco: Jossey-Bass.

Nelson, P. D., & Aletky, P. J. (1987). Accreditation: a link between training and practice. In B. A. Edelstein & E. S. Berler (Eds.), *Evaluation and accountability in clinical training* (pp. 231–252). New York: Plenum.

Perry, N. W., Jr. (1979). Why clinical psychology does not need alternative training models. *American Psychologist, 34,* 603–611.

Peterson, D. R. (1976a). Is psychology a profession? *American Psychologist, 31,* 572–583.

Peterson, D. R. (1976b). Need for the Doctor of Psychology degree in professional psychology. *American Psychologist, 31,* 792–798.

Peterson, D. R. (1985). Twenty years of practitioner training in psychology. *American Psychologist, 40,* 441–451.

Peterson, D. R. (1991). Connection and disconnection of research and practice in the education of professional psychologists. *American Psychologist, 46,* 422–429.

Peterson, D. R., & Peterson, R. L. (1997). Ways of knowing in a profession: towards an epistemology for the education of professional psychologists. In D. R. Peterson (Ed.), *Education professional psychologists: history and guiding conceptions.* Washington, DC: APA.

Peterson, R. L., Peterson, D. R., Abrams, J. C., & Stricker, G. (1997). The National Council of Schools and Programs of Professional Psychology education model. *Professional Psychology: Research and Practice, 28,* 373–386.

Raimy, V. C. (Ed.) (1950). *Training in clinical psychology.* New York: Prentice-Hall.

Roe, A., Gustad, J. W., Moore, B. V., Ross, S., & Skodak, M. (Eds.) (1959). *Graduate education in psychology.* Washington, DC: APA.

Ross, M. J., Holzman, L. A., Handal, P. J., & Gilner, F. H. (1991). Performance on the Examination for Professional Practice in Psychology as a function of specialty, degree, administrative housing, and accreditation status. *Professional Psychology: Research and Practice, 22,* 347–350.

Sanford, F. H. (1951). Annual report of the executive secretary: 1951. *American Psychologist, 6,* 664–670.

Sarason, S. B. (1981). An asocial psychology and a misdirected clinical psychology. *American Psychologist, 36,* 827–836.

Sears, R. R. (1946). Graduate training facilities. *American Psychologist, 1,* 135–150.

Sears, R. R. (1947). Clinical training facilities: 1947. *American Psychologist, 2,* 199–205.

Sechrest, L. (1990). A case for separate-but-equal in clinical training. In L. Bickman & H. Ellis (Eds.), *Preparing psychologists for the 21st century: proceedings of the national conference on graduate education in psychology* (pp. 69–75). Hillsdale, NJ: Erlbaum.

Sechrest, L., & Chatel, D. M. (1987). Evaluation and accountability in training for professional psychology: an overview. In B. A. Edelstein & E. S. Berler (Eds.), *Evaluation and accountability in clinical psychology* (pp. 1–7). New York: Plenum.

Shakow, D. (1978). Clinical psychology seen some 50 years later. *American Psychologist, 33,* 148–158.

Sheridan, E. P. (1990). Science versus practice or science and practice? In L. Bickman & H. Ellis (Eds.), *Preparing psychologists for the 21st century: proceedings of the national conference on graduate education in psychology* (pp. 225–230). Hillsdale, NJ: Erlbaum.

Sheridan, E. P., Matarazzo, J. D., & Nelson, P. D. (1995). Accreditation of psychology's graduate professional education and training programs: an historical perspective. *Professional Psychology: Research and Practice, 26,* 386–392.

Starr, P. (1982). *The social transformation of American medicine*. New York: Basic Books.

Stern, W. (1917). Hugo Munsterberg: in memorium. *Journal of Applied Psychology, 1,* 186–188.

Stevenson, J. F. & Norcross, J. C. (1987). Current status to training evaluation in clinical psychology. In B. A. Edelstein & E. S. Berler (Eds.), *Evaluation and accountability in clinical training* (pp. 77–115). New York: Plenum.

Stigall, T. T., Bourg, E. F., Bricklin, P. M., Kovacs, A. L., Larsen, K. G., Lorion, R. P., Nelson, P. D., Nurse, A. R., Pugh, R. W., & Wiens, A. N. (Eds.) (1990). *Report of the Joint Council on Professional Education in Psychology*. Baton Rouge, LA: Land and Land.

Stone, G. C., Weiss, S. M., Matarazzo, J. D., Miller, N. E., Rodin, J., Belar, C. D., Follick, M. J., & Singer, J. E. (Eds.) (1987). *Health psychology: a discipline and a profession*. Chicago: University of Chicago Press.

Stricker, G. (1997) Are science and practice commensurable? *American Psychologist, 52,* 442–448.

Strickland, B. R. (1985). Over the boulder(s) and through the vail. *The Clinical Psychologist, Summer,* 52–56.

Thompson, A. S., & Super, D. E. (Eds.) (1964). *The professional preparation of counseling psychologists*. New York: Columbia University Teachers College Bureau of Publications.

Tobias, S., Chubin, D. E., & Aylesworth, K. (1995). *Rethinking science as a career*. Tucson, AZ: Research Corporation.

Tryon, R. C. (1963). Psychology in flux: the academic-professional bipolarity. *American Psychologist, 18,* 134–143.

Wellner, A. M. (Ed.) (1978). *Education and credentialing in psychology: a proposal for a national commission*. Washington, DC: APA.

Wicherski, M. & Kohout, J. (1995). *1993 doctorate employment survey*. Washington, DC: APA.

Wilshire, B. (1990). *The moral collapse of the university: professionalism, purity, and alienation*. Albany, NY: State University of New York Press.

Yerkes, R. M. (1918). Psychology in relation to the war. *Psychological Review, 25,* 85–115.

Young, K. E., Chambers, C. M., & Kells, H. R. (Eds.) (1983). *Understanding accreditation*. San Francisco: Jossey-Bass.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.06

The Development of Standards and the Regulation of the Practice of Clinical Psychology in Europe

FRANK M. McPHERSON
Tayside Area Clinical Psychology Department, Dundee, UK

2.06.1 INTRODUCTION 62

2.06.2 THE DEVELOPMENT OF REGULATION 62
  2.06.2.1 Historical Overview 62
  2.06.2.2 Influences on Regulation 62
    2.06.2.2.1 Political influences 62
    2.06.2.2.2 Employment arrangements 63
    2.06.2.2.3 Influence of the profession 63

2.06.3 CURRENT STATE OF LEGAL REGULATION 63
  2.06.3.1 Nature of the Laws Regulating Clinical Psychology 63
  2.06.3.2 Scope of the Laws 64
  2.06.3.3 Regulatory Bodies 64
  2.06.3.4 Laws Protecting a Title 64
  2.06.3.5 Laws Regulating the Practice of Clinical Psychology 65

2.06.4 EDUCATIONAL AND TRAINING REQUIREMENTS NECESSARY FOR LEGAL RECOGNITION 66
  2.06.4.1 Structure of Training 66
  2.06.4.2 Length of Training 66
  2.06.4.3 Academic Content of Training 66
    2.06.4.3.1 The science of psychology 66
    2.06.4.3.2 The academic bases of clinical psychology 67
  2.06.4.4 Practical Training and Experience 67
  2.06.4.5 Postqualification Training 68

2.06.5 DISCIPLINARY ARRANGEMENTS 69

2.06.6 INTERNATIONAL REGULATION 69
  2.06.6.1 The European Union 69
  2.06.6.2 General Directive 89/48/EEC 70
  2.06.6.3 Effects of the Directive 70

2.06.7 FUTURE DEVELOPMENTS 71
  2.06.7.1 National Regulation 71
  2.06.7.2 International Regulation 71

2.06.8 REFERENCES 71

2.06.1 INTRODUCTION

In this chapter, the term "practice" of clinical psychology is used to refer both to independent practice and to salaried employment; "licensing" and "legal regulation" are used interchangeably.

Any attempt to describe the legal status of clinical psychology in Europe faces three major difficulties. First, the situation is developing rapidly. In the three years since 1993, the law in six European nation states (hereafter "nations") has changed and some measure of international regulation has been introduced. New laws are anticipated in several nations before the end of the decade. Second, detailed accounts of the regulations which apply to clinical psychology are available for only a few European nations. Whenever possible, published accounts have been supplemented by unpublished papers and personal communications.

Finally, and probably most important, are the difficulties which result from the major differences which exist between the European nations, in the social role of legal recognition, in the ways in which the law protects clinical psychology, and in the extent of that protection. These differences can be understood only in the context of the formal legal systems and of the informal custom and practice which determine how these laws are applied, as well as of the educational, health care, and employment arrangements of each nation. This context is often difficult for non-nationals to grasp and it is easy to be misled by laws and their provisions which, although superficially similar, have different implications within their nations of origin and, conversely, by arrangements which are less different than, at first sight, they appear.

2.06.2 THE DEVELOPMENT OF REGULATION

2.06.2.1 Historical Overview

Ironically, a nation which no longer exists, the German Democratic Republic, was, in 1967, the first European nation to regulate clinical psychology (McPherson, 1986). By a decade later, seven more had done so, namely, Hungary, Iceland, The Netherlands, Norway, Portugal, Spain, and Sweden (Kjølstad, 1983). By 1996, to that number had been added Austria (Friedlmayer & Rössler, 1995), Belgium (Baneke, 1995), Denmark (Føltved, 1995), Finland (Saari, 1995), France (Matefi & Häring, 1993), Germany (Matefi & Häring, 1993), Greece (Georgas, 1995), Italy (Bonal, 1990; Lombardo, 1990), and the UK (Lunt, 1995).

Although together these account for only about 40% of European nations, they include most of those in which clinical psychology exists as a well-developed profession. The main exceptions are the Czech Republic (Brožek & Hoscovec, 1995), Ireland (Brady & McLoone, 1992), Poland (Toeplitz-Winiewka, 1992), Slovakia (Kováč, 1995), Switzerland (Peel, 1992), and Turkey (Sahin, 1995).

2.06.2.2 Influences on Regulation

In addition to these differences in whether and, if so, when laws were introduced, even more striking differences exist in how regulation has been implemented, as the remainder of the chapter will show. Although there is little documentary evidence, three sets of influences appear to have contributed to this variability.

2.06.2.2.1 Political influences

Although differences in regulation might seem unsurprising, in view of the political divisions which have existed between European nations throughout much of the history of professional psychology (McPherson, 1991), the general political ideology of a nation appears to have had little effect. Thus, the earliest European nations with psychology laws included some which, at the time, had both left and right wing authoritarian regimes and others which then, as now, were parliamentary democracies. Moreover, there are even major differences between nations which politically are often grouped together.

The Nordic nations have been cooperating for many years, as have their psychologists who, in the 1950s, made joint proposals to the Nordic Council (of Governments) for the regulation of psychology. Although these were accepted in 1958 and were implemented by Norway in 1973, by Iceland in 1976 and, partly, by Sweden in 1978, it took until 1993 for Finland (Saari, 1995) and 1994 for Denmark (Føltved, 1995) to regulate psychology and 1994 for Sweden to achieve full title protection (Nevalainen, 1995; Persson, 1995).

Of the three Benelux nations, The Netherlands had a law protecting the title "psychologist" as long ago as 1971, whereas in Belgium there was no such law until 1993, the same year in which "psychologist" lost its protected status in The Netherlands. Luxembourg has no psychology laws (Baneke, 1995; van Drunen, 1995).

More specific political attitudes have undoubtedly affected the process of regulation.

Attempts to introduce laws were often opposed by some psychologists and political groups, especially in the 1960s, on the grounds that they threatened individual rights and freedoms (Saari, 1995). More recently, legislation in some nations has been influenced by the determination of governments to promote consumer protection and free competition. Although this has encouraged legislation, the resulting laws have tended only to protect the title, since this allows the consumer to identify qualified and ethical practitioners. The licensing of specific activities has been opposed, since this would restrict competition (Baneke, 1995).

2.06.2.2.2 *Employment arrangements*

Not unexpectedly, the employment and reimbursement arrangements in the health care system of a nation appear to have a significant influence on regulation. Legislation, for example, in 1993 in The Netherlands and Finland, has often been in the context of reforms in the health care system and, as will be discussed later, the influence of employers and insurance companies on the development of postqualification training has been significant.

In the UK, the "social recognition" of clinical psychology provided by the National Health Service (NHS) has been a major reason why there is as yet no full legal protection of any relevant title (McPherson, 1983). Traditionally, the great majority of UK clinical psychologists have been salaried employees of the NHS and, until recently, there has been very little independent practice. The British Psychological Society (BPS) influenced NHS disciplinary procedures and the entry requirements which, in turn, determined the form, duration, and content of UK clinical psychology training. For as long as the NHS remained the monopoly employer and accepted BPS advice, there was little perceived need for additional, legal protection, either for clinical psychologists themselves, or for their patients. The recent renewal of interest in full statutory legislation follows a period of health care reform, which has seen the possibility of greater diversity in employment arrangements within the NHS, along with an increased role for the private and independent provision of services.

The NHS has also influenced the breadth of clinical training obtained in the UK, which is greater than elsewhere in Europe. Although willing to set entry standards, the NHS has been reluctant to approve specialty credentialing, so that a clinical psychologist employed by the NHS is eligible to work with any category of patient and clinical problem. This, in turn, has caused the BPS to ensure that trainee clinical psychologists obtain experience with all the main NHS client groups.

2.06.2.2.3 *Influence of the profession*

The role of the national associations of psychology often appears to have been crucial. Historical accounts of the progress towards legal regulation in some nations emphasize the importance of the political lobbying carried out by the national associations over many years. These accounts also note their role, once political agreement had been reached, in influencing the technical aspects of regulation, such as training requirements and practice arrangements. Where there have been several, competing psychology associations within a nation, or where there has been disagreement about training or practice, legal regulation often appears to have been delayed.

2.06.3 CURRENT STATE OF LEGAL REGULATION

2.06.3.1 Nature of the Laws Regulating Clinical Psychology

There are very wide variations in the legal means by which psychology is regulated in the different European nations. Laws relating to psychology are of four, general types.

First, the unauthorized use of a university degree and title is probably illegal under the criminal or civil legal codes of all European nations; in some, these degree titles enable the user to be identified as having studied psychology.

Second, there are laws which are specific to the title or practice of psychology. Nations with such laws include Austria (1990), Belgium (1993), Denmark (1993), France (1985), Germany (1985), Greece (1979), Hungary (1974), Iceland (1976), Italy (1989), Norway (1973), Portugal (1972), and Sweden (1978 and 1994).

Third, in some nations, the regulation of psychology is included in more general provisions. For example, the 1979 law in Spain was enacted under an older law applying to the regulation of academic disciplines (Prieto & Avila, 1994) and, in 1993 in Finland and The Netherlands, the protection of titles resulted from laws applying to all health care professionals (Saari, 1995; Baneke, 1995).

Finally, in the absence of regulation obtained by Government-sponsored legislation, the national associations of psychologists of several nations—including Germany, The

Netherlands, and the UK—confer titles on their members which they protect by the use of civil laws.

2.06.3.2 Scope of the Laws

Although, for most nations, the laws protecting psychology apply to the whole range of professional psychology activities, for some nations the protection is limited in scope. In Sweden, "psychologist" (Persson, 1995) and in Austria, "clinical psychologist" and "health psychologist" (Friedlmayer & Rössler, 1995) are protected only within the health system, as will be "clinical psychologist" in The Netherlands when the 1993 law is fully implemented (Baneke, 1995; van Drunen, 1995). The Finnish law is interesting: it defines *all* registered psychologists as health care professionals, thus extending to other branches of the profession the protection of title which exists within the health system (Saari, 1995).

2.06.3.3 Regulatory Bodies

Probably in every nation in which there are laws, other than France (Matefi and Häring, 1993), there is a regulatory body for psychology. However, no two European nations have the same arrangements.

One way in which they differ is the extent to which the body is specific to psychology. In Spain and the UK, authority to regulate the profession has been delegated to the national associations of psychologists. Thus, in Spain a Colegia Oficial formed by Parliament with the agreement of the Crown, exists to organize a group of professional experts who share the same academic credentials. It provides a self-governing institution, to which the state delegates power to regulate the work of its members. The Colegia Oficial de Psicólogos (COP) was formed in 1979 to undertake these responsibilities for psychology (Prieto & Avila, 1994). Likewise, in the UK the BPS was given authority by the Crown to set up a voluntary register of its own members and to award and protect the title "chartered psychologist" and several specialist variations (Lunt, 1995).

By contrast, several nations have a psychology board established by the relevant law. Thus, in Belgium, the Psychologists Commission was set up by the Crown (Baneke, 1995). (The term "by the Crown" should be explained. In several European nations with constitutional monarchies, Parliament passes the laws necessary for regulating a profession, but the regulatory body that is set up acts on behalf of the Crown—king or queen—partly to protect the profession from direct political influence.) In Denmark, a nine member board includes representatives of the relevant government departments and the municipalities (Føltved, 1995). In Italy, regional councils are responsible for holding registers of recognized psychologists and for monitoring their professional activities, while a national council is responsible for developing codes of conduct, accrediting private training institutions and recommending scales of fees, for example, to insurance companies (Bonal, 1990; Lombardo, 1990).

In other countries, some or all aspects of regulation are the responsibility of a body with a remit for the wider health care system, such as a government ministry.

A second difference concerns the two main functions of a regulatory board, to grant licenses to individual applicants and to enforce a disciplinary procedure. For some nations, these are done by the same body, although usually with some internal division of function. Elsewhere, in particular where licenses are awarded by a body with a remit wider than psychology, a different agency is in charge of disciplinary matters. In Norway, the licensing of psychologists is done by the Ministry of Health, while the Central Health Authority (CHA) has the power to revoke a psychologist's license, in accordance with the Psychologists Law (Kjølstad, 1995). In Sweden, the National Board of Health and Welfare authorizes the title and the Medical Responsibility Board is concerned with the regulation of professional conduct (Persson, 1995). In Finland, these are the responsibilities, respectively, of the Ministry of Social Affairs and Health and of the National Board of Medicolegal Affairs (Saari, 1995).

2.06.3.4 Laws Protecting a Title

The most common form of legal recognition in Europe is that which protects the use of a designated title. Four different arrangements can be identified.

First, there is the protection of the university degree. Those used by psychology graduates in some European nations are common to several disciplines, for example, *doctoraandus* (*drs*) in The Netherlands. Often, however, they are specific to psychology, for example, *diplom-psychologe* in Germany, *lic. psych.* in the French-speaking parts of Switzerland (although the more general *lic. phil.* is used in the German-speaking cantons) and *cand. psych.* in Denmark. These laws provide limited title protection, in that they indicate that the user has obtained a university degree in psychology.

Second, in all but one of the nations with psychology laws—Portugal being the exception—a professional title is given some measure of legal protection. In most, the protection is of the generic "psychologist," or some equivalent, which does not identify particular specialties within professional psychology. This is the case in Finland, Germany, Iceland, Italy, Norway, and Spain, among other nations. Although "licensed psychologist" was protected in Sweden for many years, there were no restrictions on the use of "psychologist" until 1994 (Persson, 1995). In The Netherlands, the protection which had existed for *psycholoog* since 1971 was abolished by the 1993 law which, however, introduced a protected specialist title (Baneke, 1995; van Drunen, 1995).

There are variations within some nations. Although, in Switzerland, "psychologist" is not protected at federal level (Peel, 1992), it seems that protection exists in some of the component cantons (Matelfi & Häring, 1993). Under the recent law in bilingual Belgium, there are two protected titles, *psycholoog* in the Dutch-speaking parts and *psychologue* in the French-speaking regions (Baneke, 1995).

Third, in several nations, protection is given to a specialist title, usually in addition to the generic "psychologist," although in Austria only "clinical psychologist" and "health psychologist" are protected. In France, *psychologue* can be used alone, but also with a qualifying adjective, for example, indicating clinical specialization. Denmark protects *psykolog,* but those with an additional two years of supervised clinical experience may also apply to use *autoriseret psykolog* (Føltved, 1995). In The Netherlands, the 1993 law created a fully protected title for clinical psychologists, *klinisch psycholoog* (Baneke, 1995). Hungary specifies different levels of training for psychologists and for specialist, clinical psychologists (Pléh, 1991) and a similar law, which has existed in Greece since 1979, has recently been implemented (Georgas, 1995).

Finally, some national associations of psychology confer titles which attract limited protection. Thus, in 1966, the Nederlands Instituut van Psychologen (NIP) introduced the title *NIP klinisch psycholoog* for use by those members who had undergone a period of clinical training and supervised experience specified by NIP. This will continue to be used until the 1993 law is fully in force. To compensate for the loss of protection for *psycholoog,* NIP recently introduced the title *NIP psycholoog* for those of its members who have the relevant *drs* degree and further experience. These NIP titles are protected by patent and trademark (van Drunen, 1995). The German association for professional psychologists, the Berufsverband Deutscher Psychologen (BDP), has a similar arrangement which enables members who have obtained the *diplom-psycholog* and further training and experience to refer to themselves as "BDP clinical psychologists." The Norsk Psykologforening in Norway, in addition to the legally-protected generic title, authorizes the use of a more restricted title by clinical psychologists who have met their additional criteria (Kløve, 1992).

2.06.3.5 Laws Regulating the Practice of Clinical Psychology

To practice any branch of psychology in Spain, it is legally necessary to be a member of the COP (Rodriguez Sutil, 1995) and, in Portugal, to obtain a license from the Ministry of Work, although there appear to be no sanctions for noncompliance (Pereira, 1995).

Entry into the field of health care appears, in some nations, to be restricted to licensed practitioners, although the scope of these restrictions is not always clear. Elsewhere, although there may be no formal restrictions, informal agreements ensure that only licensed psychologists are employed. Where there are restrictions, they tend always to apply, wholly or mainly, to the field of public health. Where they affect independent practice, it is normally in connection with Government-supported insurance companies which, in some nations, will reimburse the fees only of psychologists with the appropriate, protected title. The nations which, by whatever means, most stringently restrict entry into health care to licensed psychologists, are Austria, Finland, Hungary, Norway, and Sweden.

Within the field of health care, there are few examples of laws which explicitly permit or prohibit clinical psychologists from performing specific activities. The Austrian law defines the activities associated with clinical psychology (Friedlmeyer & Rössler, 1995). In Norway, when the 1973 Psychologists Law was passed, other laws were amended so that psychologists were excluded from regulations to combat charlatans and could use hypnosis in the treatment of illness (Kløve, 1992). The recent law in Finland states that "only a licensed professional is entitled to perform the tasks belonging to the said profession and to use the title in question," but it is not yet clear to what extent the application of the law will result in any specific activities being held to be exclusive to psychologists (Saari, 1995).

Psychotherapy is outside the scope of this chapter. However, the situation in brief is that, in some nations, psychotherapy is regarded as a profession in its own right and is regulated independently. In some of these, clinical psychology training allows registration also as a psychotherapist whereas, in others, additional training is required. In several other nations, in which psychotherapy is regarded as an activity to be carried out by members of more generic professions, special practising licenses are required. In most of these, clinical psychologists are eligible for these licenses but, in at least one, psychotherapy can be practised only under the (usually nominal) supervision of a medical practitioner. In some nations, including the UK, there are no regulations.

2.06.4 EDUCATIONAL AND TRAINING REQUIREMENTS NECESSARY FOR LEGAL RECOGNITION

2.06.4.1 Structure of Training

In Greece, Hungary, and the UK, the structure of clinical psychology training is somewhat similar to that in North America, with the university phase being discontinuous, consisting of two separate degree courses. The first of these is concerned with the scientific bases of psychology, while the second is a program of academic, research, and practical studies in the clinical applications of psychology. Hitherto, these latter programs have led to masters degrees, the term "doctoral level" training having no meaning in Europe, where the Ph.D. and its equivalents are awarded only for research. An exception is the UK, where, in the last few years, universities have begun awarding doctoral degrees (usually *doctor of clinical psychology*) on completion of the second degree.

Elsewhere in Europe, the arrangements are quite different. Potential clinical psychologists take a single degree, which normally makes them eligible for licensing as a qualified professional psychologist, although in some nations a further period of practical training and experience is necessary (McPherson, 1991, 1992).

2.06.4.2 Length of Training

The minimum length of training necessary for legal registration, as defined in the relevant laws of each nation, often underestimates the actual time spent by many clinical psychology students before becoming licensed. In many European nations, individual students are given considerable freedom to determine the length of their university studies and, although governments are making strenuous efforts to curtail this freedom, the average time taken may be two or more years longer than the period specified in the relevant curriculum or law.

In most of those nations which legally protect the generic title "psychologist," the minimum period of training prior to licensing is five years: this is the case in Belgium, France, Spain, and Portugal, although in Sweden it is six years—five of study, followed by one year of supervised practice—and in Norway it is between six and seven years (McPherson, 1991, 1992).

Where the protection is of "clinical psychologist," or some other specialist title, the prescribed minimum time is usually longer. Thus, in Denmark, although a five year training leads to the protected *psykolog,* a further two years of supervised practice are required to become an *autorisiret psykolog*. (Føltved, 1995). In Austria, seven years are necessary, in a single degree, as they are in Hungary, although in a five-year first degree, followed by a two-year clinical training period (László & Pléh, 1992). In The Netherlands, eligibility for *klinisch psycholoog* requires four years of inservice training and experience, subsequent to the four year *drs* (Baneke, 1995; van Drunen, 1995). The UK title *chartered clinical psychologist* is awarded after a minimum of six years, three spent on the first (*bachelors*) degree and three on a graduate (*postgraduate* in UK parlance) clinical training course (Lunt, 1995).

2.06.4.3 Academic Content of Training

There are difficulties in comparing the educational requirements necessary to obtain a protected title. Published curricula are potentially misleading, because of wide variations between nations in the relative time typically devoted to formal instruction and to private study. Internationally, there is an absence of agreed definitions of terms such as "supervised experience," "credit hours," and "year of study" and of a system of on-site inspection and peer review, to establish the relationship between what is described and what is done in practice.

In many nations, because of decrees by the national government, or the influence of the profession, there is considerable similarity between different universities in the content of psychology degrees. In some, however, the traditional autonomy of universities remains, so that big differences exist in the curricula which lead to the protected title, thus allowing few generalizations to be made about the "typical" content of training.

2.06.4.3.1 The science of psychology

In those nations in which clinical psychology training involves two separate degree courses, the first (e.g., leading in the UK to the *bachelors* degree) is wholly or mainly devoted to studies in the scientific discipline of psychology. Hence, it covers all or most of the following topics: the biological bases of behavior; perception, learning, and memory; thinking and language; personality, intelligence, and individual differences; social psychology; developmental psychology; research design and quantitative methods; and specialist areas. A research dissertation will also normally be required (McPherson, 1991).

In those nations in which potential clinical psychologists study for a single degree, these topics are usually included during the first semesters, lasting two or three years (McPherson, 1991, 1992).

2.06.4.3.2 The academic bases of clinical psychology

Where training is discontinuous, the academic teaching in clinical psychology and related topics takes place during the course of studies leading to the second, specialist degree in clinical psychology, which is devoted almost exclusively to these subjects.

Where there is a single degree, all potential clinical psychologists take academic courses relevant to clinical psychology during the later semesters. However, there are major differences between nations in the extent to which specialization is possible at this stage.

In particular, in the Nordic nations, all students receive an education in the academic bases, not only of clinical psychology, but also of other branches of professional, applied psychology. Thus, in Norway, the 1987 revision of the standards for the *cand. psychol.* degree require that at least four semesters should be devoted to studying five applied disciplines: namely, school and educational psychology, organizational and occupational psychology, medical and health psychology, community psychology, and clinical psychology. In Denmark, after five "scientific bases" semesters, students spend four semesters studying organizational, educational, and clinical psychology, among other topics. In Finland and Sweden there is also no specialization, with all students undertaking more or less the same courses. Some other nations also expose students to a broad education in professional psychology; thus, in Spanish universities, the second cycle of studies, of between two and three years, includes clinical and educational psychology and the psychology of groups and organizations.

In some other nations, a much greater degree of specialization is possible. In Germany, students are often able to spend their last two or three years studying only one branch of professional, applied psychology—usually clinical, educational, or organizational psychology. Roughly similar arrangements also exist in The Netherlands, Belgium, Italy, and France, although the range and content of the optional specializations vary greatly.

2.06.4.4 Practical Training and Experience

There are major differences between nations in the amount and in the scope of clinical training and experience required of potential clinical psychologists, before they are eligible for the protected title. There are also differences in the stage of the training process at which the experience is obtained. There are three main types of arrangement.

First, in most of the nations in which training is continuous, all the practical training and experience necessary for licensing is carried out during the program leading to the single degree. In some of these, in particular where the profession appears to have had a significant influence on the university curricula, this practicum period is substantial. Elsewhere, however, universities have historically been seen as providing the educational bases for professional training, rather than the training itself, which is thus normally carried out after graduation and hence also after licensing (McPherson, 1992). The minimum practicum period therefore varies greatly. Norway requires four months of practical training and eight months of practical experience. In Finland, there is a supervised practicum of six months and, in Belgium, one which occupies four days weekly for at least seven months. In Sweden, in addition to undertaking a practicum, all students must receive personal therapy. Elsewhere, practical training can be much shorter. In some German universities, only two, six-week periods are required and, in some Spanish universities, only approximately 90 hours.

Second, the main period of supervised training and experience can occur after completion of the single degree. In Denmark, at least two years of such supervised experience are necessary before *autoriseret psykolog* status is granted and Italy and Sweden require a one year internship before the trainee is eligible for licensing.

Finally, where training is discontinuous, the bulk of the program for the second degree is given over to clinical training and supervised practice. The greatest breadth of experience is provided in the UK where, to become a *chartered clinical psychologist,* trainees must have undertaken five supervised and assessed clinical placements, each of at least 65 days: with adults with psychological disorders, with children and adolescents, with people with a learning disability (mental handicap), with older adults, and in another, more specialized area of the trainee's own choice. In the other nations with this structure, trainees obtain experience with fewer client groups—seldom, for example, with people with a learning disability—but often for longer periods.

2.06.4.5 Postqualification Training

Postqualification training and experience, that is, obtained after a clinical psychologist has been licensed, is very much more important in some European nations than, for example, in North America. As described above, the protected generic title can often be obtained without the clinical psychologist having received a significant amount of specialized, clinically relevant education, or practical clinical training. Some years ago, this began to be recognized as a problem. Psychologists who had been licensed and who wished to work as clinicians, found that they required substantial, additional specialist training, both to be acceptable to employers in the health sector and to have marketable skill in independent practice. Also, potential employers and clients, insurance companies concerned with the reimbursement of fees for clinical psychology assessment and treatment, and others who wished to identify psychologists with appropriate clinical skills, found that the generic titles were not helpful. Several solutions to this problem have been attempted.

Possibly the most satisfactory solution has been to introduce protected specialist titles. In The Netherlands, "psychologist" is being replaced with "clinical psychologist" as the legally protected title. Another solution, adopted by several nations, is to have, in addition to the generic title, a legally protected specialist title, available after programs of advanced training. The example of Denmark has already been noted and both Finland and Sweden are implementing lengthy programs, on topics such as psychotherapy, health psychology, and neuropsychology, which will form the basis of postregistration specialty licensing. It was an awareness of the limitations of generic licensing which persuaded Austria to protect two specialist titles, "health psychologist" and "clinical psychologist."

In the absence of protected specialist titles, several national associations arranged their own, advanced programs of study in clinical psychology, leading to a qualification approved by themselves. As noted above, such qualifications have achieved quasilegal status: that is, the title can usually be protected in the civil courts and—usually as the result of lengthy negotiation by the association—is used by employers and by insurance companies to determine eligibility for employment or reimbursement of fees. One of the most extensive programs is that leading to the title "specialist clinical psychologist," awarded by the Norwegian Psychological Association. This requires some five years of postlicensing experience, in at least two settings, including work with both adults and children and at least two years spent as a full-time employee in one setting; a minimum of two hours of supervision per week for two years is specified, along with attendance at courses and workshops. This title qualifies clinical psychologists to be appointed to certain senior posts in the health service and has implications for reimbursement of insurance (Kløve, 1992). Two other examples include: (i) the postqualification title "NIP clinical psychologist," awarded in The Netherlands since 1966, but which will be replaced by the fully protected "clinical psychologist" established by the 1993 law, although the training requirements will continue to be determined largely by NIP (Baneke, 1995); and (ii) the specialist title "BDP clinical psychologist," awarded by the German association, which also has implications for employment and reimbursement. Both require between four and six years of part-time study.

Elsewhere, employers have organized their own training programs, to supplement the training of licensed psychologists. The former German Democratic Republic had an elaborate system of training and regulation, which enabled clinical psychologists to have equal legal status to medical doctors (Kossakowski, 1992). In Spain, a system of internship, Psicólogo Interno Residente (PIR), has been developed by the Ministries of Education and Health, analogous to that for medical graduates. Psychologists who are already registered with the COP spend three years on supervised placements in the public health system (Belloch & Olabarría, 1994).

In several nations, the need was met by private organizations, which developed very many training programs, dedicated to teaching specific clinical skills, usually psychotherapeutic procedures. Some of these are very extensive, lasting several years and involving academic teaching as well as supervised practice. Clinically relevant courses with a more academic orientation are offered by universities, in particular in France and Spain. Often, this additional training does no more than increase the marketability of clinical psychologists in independent practice. However, in Germany, the major health insurance companies have agreed to reimburse the fees only of those clinical psychologists who have undergone specified training, for example, in psychoanalytic or behavioral psychotherapy.

Because many of these programs take place after the clinical psychologist has been licensed, they are not usually subject to scrutiny or control, either by the official regulatory body, or by the relevant national association of psychologists. This has led to anxieties about the quality of training. Some national associations have therefore developed an informal system of accrediting the courses run by private organizations and universities. In nations in which there is no other protection of title or practice, for example, Turkey (Sahin, 1995), this accreditation is seen as a first step towards full legal regulation.

2.06.5 DISCIPLINARY ARRANGEMENTS

The disciplinary arrangements which apply in most of Europe are broadly similar to those in North America, in that licensed clinical psychologists are subject to the criminal and civil law, to the codes of ethics and conduct of their professional bodies, and to the regulations of the relevant licensing authorities. However, the detailed procedures vary greatly.

The range of different bodies involved in regulating conduct in the various European nations has previously been outlined. In those in which clinical psychologists are regulated as health care professionals, both the legal requirements and the penalties imposed appear to be more stringent than elsewhere. Thus, in Finland, the clinical psychologist is *legally* bound, among other things: (i) to apply generally approved and empirically tested procedures in accordance with his/her training; (ii) to maintain and improve the skills required for the practice of the profession (although employers are also legally required to allow him/her to attend any necessary training); (iii) to be aware of the regulations and statutes covering professional conduct; (iv) to create and keep confidential patient records; and (v) to provide assistance to anyone in need of urgent treatment. The penalty for failing to maintain the confidentiality prescribed by law can be six months imprisonment (Saari, 1995).

In most of the nations which legally regulate clinical psychology, the national association has a formal code of ethics which usually appears to have at least some role in disciplinary proceedings, although in some, for example, Norway, the relevant law also includes statements about prohibited conduct (Kjølstad, 1995). However, there appear to be no codes of conduct in use which are as specific and detailed as that produced by the Association of State and Provincial Psychology Boards (see Chapter 24, this volume), and anecdotal evidence suggests that some existing codes are not helpful in disciplinary proceedings.

2.06.6 INTERNATIONAL REGULATION

2.06.6.1 The European Union

The international influences on the legal regulation of clinical psychology in Europe derive from the European Union (EU), formerly the European Economic Community (EEC). Since 1994, EU provisions for the mutual recognition of professional qualifications have also applied to some of the member nations of the European Free Trade Association (EFTA). In 1996, the EU and EFTA nations—collectively referred to as the European Economic Association (EEA)—to which these provisions applied were: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, The Netherlands, Norway, Portugal, Spain, Sweden, and the UK.

One aim of the EU is to create a "common market" of professions and occupations, by removing barriers to free movement. The legislative means of achieving this is the "Directive," which is binding on member nations in relation to the required results, although the means of achieving them maybe decided by each nation (McPherson, 1989).

The initial approach was "sectorial," with a separate Directive for each profession. This required member nations to guarantee, for that profession, the mutual recognition of qualifications and the "right of establishment," in other words, that an EU national, trained in one member nation, would have his/her qualifications accepted in any other member nation in which he/she wished to work. Directives were issued only after a significant amount of "harmonization" of training had been achieved, so that each sectorial Directive was able to specify, for the EU as a whole, a common set of minimum training requirements for entry into the profession.

However, progress was very slow and, although work was begun on over 50 Directives, by the early 1980s only seven—mainly concerning health professions—had been implemented. Although this was partly due to the technical problems of harmonizing what were often quite heterogeneous training arrangements, delays were mainly due to a lack of urgency: free movement was a political goal, for which few professions expressed enthusiasm (Orzark, 1983; McPherson, 1989).

The EU lost patience with these delays and adopted a “horizontal” approach, with a single Directive which applied to all professions and occupations which required at least three years of university training. This “general system for the recognition of higher education diplomas” was published in draft form in 1985 and took effect, as Directive 89/48/EEC, in January 1993.

2.06.6.2 General Directive 89/48/EEC

Psychology had not been subject to a sectorial Directive because, in the words of one EU official to the author, “psychology training is so diverse that free movement should not be encouraged.” The horizontal Directive applies only to those professional activities which are regulated and, although no definition is given, the clear intention is that the term “regulation” should be defined very broadly. Hence, in many EEA member nations, the Directive covers all clinical psychology practice, both salaried and independent; in all others, at least some aspects are affected.

The Directive is concerned with members of a profession (“migrants”) who move from one EEA member nation to another (the “host”). Its main provision is to make illegal any attempt to prevent a migrant clinical psychologist from practicing, merely because they had trained in another member nation; it extends not only to direct discrimination, but also to the setting of licensing or employment requirements which are in excess of those demanded of clinical psychologists of the host nation. It permits an aggrieved migrant to use the civil courts of the host nation to take action against an employer, or a regulatory body and, ultimately, to take the case to the European Court of Justice, which has powers which are normally binding on EU member nations.

Unlike its sectorial predecessors, this Directive was not preceded by a period of harmonization and does not specify training requirements. Therefore, it applies regardless of any differences in the training of a migrant and of clinical psychologists of the host nation. Moreover, EU officials emphasize that, when training courses and professional skills are being compared, the *equivalence* of qualifications is more important than their exact equality. However, where differences in training are significant in length or content, after any postqualification experience has been taken into consideration, the Directive does allow the regulatory body of the host to require the migrant, either to complete an “adaptation period” under some measure of supervision, or to take an examination on those academic or professional topics which they have not previously covered. There are strict limits to what can be required (McPherson, 1990a).

2.06.6.3 Effects of the Directive

Thus far, the Directive is having little effect, because few clinical psychologists migrate within the EEA, due to an absence of work opportunities and to language difficulties. There is some movement between the Nordic nations, but the Directive has added little to the common market for psychologists which they have operated for some years.

The exception is the UK, to which there is a significant amount of migration, because there are vacant clinical psychology posts and many EEA clinical psychologists have English as their second language. The least enthusiastic member of the European family of nations has thus done most to implement the Directive and the BPS has established a committee to scrutinize the qualifications of migrants who have applied for registration as *chartered clinical psychologists.* Most are required to undergo additional training and supervised practice, mainly because of their lack of clinical experience with people with learning disabilities (mental handicap) and with older adults.

One area of uncertainty has not yet been addressed. The Directive applies only to mobility within the same profession but, since some nations protect generic and others specialist titles, it is not clear whether the profession is “psychology,” “clinical psychology” (and other specialist variants), or both. It has yet to be determined whether a host nation which protects a specialist title can legally decide that, because a particular migrant was, by title and training, a “psychologist,” they belong to a different profession, and so the Directive does not apply.

Similarly, the distinction between “clinical psychology” and “psychotherapy” is causing difficulties. Some migrants, although “clinical psychologists” in their own nation, have been refused right of establishment because the host considered that they were actually “psychotherapists,” because their training did not contain sufficient psychology. France has criticised the Directive on the grounds that, even with the maximum permitted adaptation period, migrants could practice in France after several years training less than that required of French clinical psychologists. However, this training gap derives not from the five years necessary to obtain the protected title, but from the very long postqualification training which many French clinical psychologists undertake to become psychotherapists.

If these ambiguities are seen by the EU to obstruct the working of the Directive, psychologists in Europe might be required, either by the EU itself or by a court, to clarify their professional structure and to define "core curricula," which would enable clinical psychology to be discriminated objectively, both from other branches of psychology and from psychotherapy.

2.06.7 FUTURE DEVELOPMENTS

2.06.7.1 National Regulation

Over the next few years, some additional nations are likely to enact laws regulating psychology. The national associations of the Czech Republic, Slovakia, Switzerland, Poland, Turkey, Romania, and the UK, among others, have all indicated that laws are being prepared, for which government support is being sought.

As noted above, it appears that the protection provided by generic titles is increasingly being seen as inadequate, in particular in the health field. This might influence which title these new laws will protect and might also lead to existing laws being changed, as they were in The Netherlands, so that specialist titles, obtained after longer, fully regulated training and experience, will be introduced to replace or supplement the generic "psychologist." However, there is no reason to expect that there will be a change in the preference of most governments for protecting the title, rather than the practice, of clinical psychology.

2.06.7.2 International Regulation

There may well be EU Directives, for example, on the application of Value Added Tax to clinical services and on aspects of consumer protection, which will affect clinical psychology. However, in the field of regulation, the main concern of the EU continues to be to remove barriers to mobility and to the right of establishment. It is clear that, in psychology, the barriers are created less by restrictive national laws and regulations than by language and the absence of work opportunities. Thus, any sectorial Directive for psychology which is proposed by the profession would probably be seen by the EU to be unnecessary. Instead, collaboration between universities and practitioners is being encouraged as a means of promoting harmonization in the training and practice of psychologists.

If further legislation is ever thought to be necessary to enable clinical psychologists to become more mobile, a sectorial Directive, to establish the equivalence of national qualifications, is only one possibility. Another approach was pioneered in the 1980s by FEANI, the European federation of national associations of engineers. According to this model, EU nations would set up a European Register, with its own regulatory body and entry requirements and would agree to grant right of establishment to psychologists on the Register (McPherson, 1990b).

2.06.8 REFERENCES

Baneke, R. F. (1995). Benelux, a bundle of contrasts. *News from EFPPA, 9*(1), 10–12.

Belloch, A. & Olabarría, B. (1994). Clinical psychology: Current status and future prospects. *Applied Psychology: An International Review, 43,* 193–211.

Bonal, I. (1990). L'ordre à l'italienne. *Le Journal des Psychologues, 80,* 55–56.

Brady, T. & McLoone, J. (1992). Ireland. In V. S. Sexton & J. D. Hogan (Eds.), *International psychology: Views from around the world.* Lincoln, NE: University of Nebraska Press.

Brožek, J. & Hoscovec, J. (1995). Psychology in the Czech Republic. In A. Schorr and S. Saari (Eds.), *Psychology in Europe. Facts, figures, realities* (pp. 3–14). Seattle, WA: Hogrefe & Huber.

Føltved, P. (1995). Professional psychology in Denmark. In A. Schorr & S. Saari (Eds.), *Psychology in Europe. Facts, figures, realities* (pp. 15–22). Seattle, WA: Hogrefe & Huber.

Friedlmayer, S. & Rössler, E. (1995). Professional identity and public image of Austrian psychologists. In A. Schorr & S. Saari (Eds.), *Psychology in Europe. Facts, figures, realities* (pp. 165–179). Seattle, WA: Hogrefe & Huber.

Georgas, J. (1995). Psychology in Greece. In A. Schorr & S. Saari (Eds.), *Psychology in Europe. Facts, figures, realities* (pp. 59–75). Seattle, WA: Hogrefe & Huber.

Kjølstad, H. (1983). Professional psychologists associations in Europe 1982. *Tidsskrift for Norsk Psykologforening, 20,* 544–549.

Kjølstad, H. (1995). *Disciplining psychologists in Norway.* Paper presented at the First International Congress on Licensure, Certification, and Credentialing, New Orleans, LA.

Kløve, H. (1992). Norway. In V. S. Sexton & J. D. Hogan (Eds.), *International psychology: Views from around the world.* Lincoln, NE: University of Nebraska Press.

Kossakowski, A. (1992). German Democratic Republic. In V. S. Sexton & J. D. Hogan (Eds.), *International psychology: Views from around the world.* Lincoln, NE: University of Nebraska Press.

Kováč, D. (1995). Psychology in Slovakia—Leaving the Iron Curtain behind. In A. Schorr & S. Saari (Eds.), *Psychology in Europe. Facts, figures, realities* (pp. 103–110). Seattle, WA: Hogrefe & Huber.

László, J. & Pléh, C. (1992). Hungary. In V. S. Sexton & J. D. Hogan (Eds.), *International psychology: Views from around the world.* Lincoln, NE: University of Nebraska Press.

Lombardo, G. P. (1990). Formation et rôle du psychologue clinicien en Italie. *Le Journal des Psychologues, 80,* 57–58.

Lunt, I. (1995). The training and regulation of psychologists in the United Kingdom. *News from EFPPA, 9*(1), 17–19.

Matefi, G. & Häring, E. (1993). *Inventory of regulations in the field of psychology in European countries*. Bonn, Germany: BDP.

McPherson, F. M. (1983). The United Kingdom. *International Journal of Psychology, 6,* 27–34.

McPherson, F. M. (1986). The professional psychologist in Europe. *American Psychologist, 41,* 302– 305.

McPherson, F. M. (1989). Psychologists and the European Economic Community. *The Psychologist, 1,* 353–355.

McPherson, F. M. (1990a). Psychologists and the EEC (II). *The Psychologist, 2,* 382–383.

McPherson, F. M. (1990b, November). The regulation of a European profession of clinical psychology. *Clinical Psychology Forum,* 17–19.

McPherson, F. M. (1991). Professional psychology training in Western and Nordic Europe. *Educational and Child Psychology, 8,* 56–63.

McPherson, F. M. (1992). Clinical psychology training in Europe. *British Journal of Clinical Psychology, 31,* 419–428.

Nevalainen, V. (1995). Legal matters in the Nordic countries. *News from EFPPA, 9*(1), 13–14.

Orzark, L. H. (1983). International authority and national regulation. *Law and Human Behavior, 7,* 251–264.

Peel, R. (1992). Suisse: Une psychologie "fédérale"? *Psychologie Europe, 6,* 2–3.

Pereira, F. (1995). Psychology in Portugal. In A. Schorr & S. Saari (Eds.), *Psychology in Europe. Facts, figures, realities* (pp. 97–102). Seattle, WA: Hogrefe & Huber.

Persson, H. (1995). Psychology in Sweden. In A. Schorr & S. Saari (Eds.), *Psychology in Europe. Facts, figures, realities* (pp. 261–272). Seattle, WA: Hogrefe & Huber.

Pléh, C. (1991). Some notes on training psychologists in Hungary. *News from EFPPA, 5*(4), 12–14.

Prieto, J. M. & Avila, A. (1994). Linking certified knowledge to labor markets. *Applied Psychology: An International Review, 43,* 113–130.

Rodriguez Sutil, C. (1995). The situation in Spain and southern European countries. *News from EFPPA, 9(1),* 15–16.

Saari, S. (1995). The legalization of Finnish psychology. In A. Schorr & S. Saari (Eds.), *Psychology in Europe. Facts, figures, realities* (pp. 23–34). Seattle, WA: Hogrefe & Huber.

Sahin, N. H. (1995). Psychology in Turkey—steps towards accreditation. *News from EFPPA, 9(3),* 11–17.

Toeplitz-Winiewka, M. (1992). The Polish Psychological Association. *News from EFPPA, 6*(3), 22–25.

van Drunen, P. (1995). Professional psychology in The Netherlands. In A. Schorr & S. Saari (Eds.), *Psychology in Europe. Facts, figures, realities* (pp. 77–96). Seattle, WA: Hogrefe & Huber.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.07

The Development of Standards and the Regulation of the Practice of Clinical Psychology in China

QICHENG JING
Chinese Academy of Sciences, Beijing, China
and
PEICHENG HU
Beijing Medical University, China

2.07.1 BRIEF INTRODUCTION OF THE DEVELOPMENT OF CLINICAL PSYCHOLOGY IN CHINA 74

2.07.2 THE USE OF THE TERMINOLOGY "CLINICAL PSYCHOLOGY" AND "MEDICAL PSYCHOLOGY" 75

2.07.3 PROFESSIONALS IN CLINICAL PSYCHOLOGY 75

2.07.4 REGULATIONS OF PSYCHOLOGICAL TESTING 77
- *2.07.4.1 Registration of Psychological Tests* 77
- *2.07.4.2 Qualifications of Personnel Using Psychological Tests* 77
- *2.07.4.3 Control and Management of Tests* 77

2.07.5 REGULATIONS OF PSYCHOLOGICAL COUNSELING AND PSYCHOTHERAPY 77
- *2.07.5.1 Qualifications* 79
  - *2.07.5.1.1 Academic credentials* 79
  - *2.07.5.1.2 Knowledge* 79
  - *2.07.5.1.3 Clinical practice* 79
- *2.07.5.2 Responsibility* 79

2.07.6 THE CURRICULUM, TRAINER REQUIREMENTS, AND TRAINING PROGRAM IN CLINICAL PSYCHOLOGY 79

2.07.7 PSYCHOTHERAPY AND CHINESE TRADITIONAL MEDICINE 80

2.07.8 CODE OF CONDUCT AND ETHICAL GUIDELINES FOR PSYCHOLOGICAL TESTING, COUNSELING, AND PSYCHOTHERAPY 80

2.07.9 CONTINUING EDUCATION MANDATES 81

2.07.10 FUTURE DIRECTIONS OF CHINESE CLINICAL PSYCHOLOGY 82

2.07.11 REFERENCES 82

2.07.1 BRIEF INTRODUCTION OF THE DEVELOPMENT OF CLINICAL PSYCHOLOGY IN CHINA

Chinese psychology has a long past but a short history. Some 2000 years ago, in the first Chinese medical book, *The Yellow Emperors' Inner Classic*, there are many descriptions of psychological thoughts. One concerns the "five phase theory" in which it is supposed that diseases can lead to emotions of anger, happiness, pensiveness, sadness, and fear, and that these, in turn, may cause damage to the liver, heart, spleen, lungs, and kidneys. Hence, it is important to have good emotions to maintain mental health. This example is typical of the basic ideas in Chinese clinical psychology.

In 1917, the first psychological laboratory was established at Peking University. This marked the beginning of modern Chinese psychology. In 1920, the first department of psychology was established at Nanking Teachers' College, and it was quickly followed by departments at the Universities of Peking, Tsinghua, and Fudan. The Chinese Psychological Society (CPS) was established in 1921 and, also around this period, the Chinese Association of Psychological Testing, the Society of Psychoanalysis, and the Society of Mental Health were established. In 1922, an academic psychological journal was begun by Professor Zhang Yaohsiang. By this time, a few Chinese tests were available and some psychologists started to translate Western tests (Jing, 1994).

Professor Ding Zan set up the first clinic of mental health in 1936 (Kan, 1989), and from this time, clinical psychology in China developed at a much faster pace. The Western working pattern of psychiatry, that is, the combined working group of psychiatrists, clinical psychologists, and social workers, was soon introduced in Shanghai, Nanking, Peking, and Changsha. By 1945, the first book on mental health, *Comprehensive mental health* (Ding, 1945), was available.

After the founding of the People's Republic of China in 1949, China followed a similar pattern of development in science to that in the Soviet Union. Pavlov's theory of conditioned reflexes was taken as the starting point in all medical and biological sciences, hence the dominant methodology in clinical psychology was the behavioral approach. In the early 1950s Professor Li Xintian treated patients suffering from neurosis by psychotherapy and obtained good results. This was the first application of psychotherapy to mental patients in the new China.

The Chinese Cultural Revolution began in 1966 and ended in 1976. During this period psychology was attacked as a "pseudoscience," a bourgeois discipline imported from the West. The CPS was dissolved and Chinese psychologists were sent to factories and farms in remote parts of the country to do physical labor and for thought reform. Only after the Cultural Revolution was psychology rehabilitated and recognized as a scientific discipline. The ultraleftist trend of thought in denouncing science was then criticized. Clinical psychology was once again given a scientific stance.

In 1979, the CPS was re-established after the chaos of 10 years of Cultural Revolution. In September 1985, the Chinese Mental Health Association (CMHA) was inaugurated and held its first national congress. From the 1980s, Professor Peng Youfang of Fujian province, Professor Chen Peizhang of Shanxi province, and Professor Wang Xiaodao from Beijing began psychological counseling and practiced psychotherapy in hospitals. Several years later, the Beijing Municipal Bureau of Health published the standards in the ratings of hospitals. To achieve the standard of grade 3A (the highest rating), a hospital had to provide psychological counseling and psychotherapy in its related departments. Thus, due to this government mandate, psychological counseling and psychotherapy soon became widely used in Chinese hospitals.

In Chinese universities, the first counseling center was set up in 1982 at the Huzhou Normal College of Zhejiang province. The Psychometrics and Counseling Service Center at the Beijing Normal University was established in 1985. After that, counseling centers were set up in Beijing, Shanghai, and in the provinces of Shandong, Shanxi, and Hubei. These counseling centers were established at different types of institutions: 6.6% in hospitals, 70.0% in university student clinics, and 23.4% in university departments of psychology, philosophy, or social sciences. According to a survey conducted in 1989 on 126 000 students, some 20% of college students have mental health problems.

Since the late 1970s, China has adopted a new policy of economic reform and opening up to the outside world. It now has one of the highest rates of economic growth in the world, with an average Gross Domestic Product increase of 12% for five years running. The introduction of a market economy based on free competition, and the strong influence of Western culture with the increase in scenes of sex, terror, violence, and crime on television and other media, have resulted in a general moral decline of the younger generation, and an increase in the number of behavioral disorders. The need for clinical psychology is strongly felt.

2.07.2 THE USE OF THE TERMINOLOGY "CLINICAL PSYCHOLOGY" AND "MEDICAL PSYCHOLOGY"

What is psychology in the minds of ordinary Chinese people? With the publicity that psychology has gained through the media in recent years, more and more people have begun to understand that psychology is a scientific discipline. Some people, however, still think that it is a profession of mind reading or which guesses others people's thoughts. Nevertheless, since most psychologists work in institutions of higher learning, they are respected by ordinary people, just like professors in other fields.

The term "medical psychology" was introduced from the former Soviet Union in 1950s, and is still used in China. At present, the difference between clinical psychology and medical psychology is very small. Some psychologists have proposed replacing the term "medical psychology" with "clinical psychology," but this was rejected at the 1995 Annual Conference of Medical Psychology. In the formal documents of Chinese State Education Commission, the term "medical psychology" is used as a branch of psychology, and the term "clinical psychology" is not used at all. This is one of the reasons why the term "clinical psychology" is not as popular as it is in Western countries, although the actual work of medical psychologists is the same as Western clinical psychologists.

Generally speaking, the field of clinical psychology can be divided into four parts:

(i) *Medical psychology in medical schools*. Most medical schools have departments, sections, and research centers in medical psychology. Medical students undertake medical psychology as one of the required courses of the curriculum. There are out-patient departments of psychological counseling and psychotherapy in Chinese general hospitals. Some general hospitals have in-patient departments of psychosomatic diseases. Most medical psychologists are trained in medicine and receive additional training in psychology. They have prescription rights in hospitals. Other medical psychologists are trained in psychology. They do not have prescription rights and this greatly restricts their performance.

(ii) *Psychiatry in mental hospitals*. Most psychiatrists in China also consider themselves to be medical psychologists. They treat mental patients with medication and psychotherapy. They also take part in psychological counseling or psychotherapy outside of the hospital. The leadership and the majority of the membership of CMHA are psychiatrists.

(iii) *Psychological counseling in universities and schools*. The mental health of students has become one of the most important problems in Chinese education. Psychological counseling centers have been set up in several universities and schools. Originally, most of the workers in these centers were part-time staff. But now, with the development of psychological counseling, the educational administration has realized its importance and recruited more full-time staff. Regarding the professional background of the staff members in these counseling centers, some came from psychology or medicine, but most came from other fields; usually they are graduates from the same university or school.

(iv) *Clinics of psychological counseling and psychotherapy in society*. Often, medical psychologists, psychiatrists, and applied psychologists work together to set up a clinic to provide service to the public. Most clinics are supported by the local government, a few are supported by private institutions. Some NGOs (nongovernment organizations), such as the Jinglun Family Science Center, offer free psychological counseling to the public and have a good reputation in society.

Ordinary people in China like to call the medical psychologist the "psychological doctor." A magazine named *Psychological Doctor* was published as a popular science journal. After the establishment of the CMHA, two journals were published under its auspices, namely, the *Journal of Chinese Mental Health* and the *Journal of Chinese Clinical Psychology*. The term "clinical psychology" was used for the first time in a title of an official publication and this marked the beginning of its public use in China.

2.07.3 PROFESSIONALS IN CLINICAL PSYCHOLOGY

During the Third Conference of the CMHA held in 1996, the President of the Association Professor Cheng Xueshi, declared that its membership had reached 20 000. The majority of members came from the medical profession, while only a small part came from psychology and other behavioral science disciplines. Professor Gong Yaoxian from Hunan Medical University made a survey of the constituents of clinical psychologists in 30 provinces and cities (Gong, 1993). This showed that 76.3–87.5% came from the medical profession, whereas only 2.3–8.3% came from psychology. Detailed statistics are shown in Table 1.

Table 1 Professionals in clinical psychology in China.

	Education		*Original profession*			*Training*		
Clinical work	*College*	*Middle school*	*Medicine*	*Psychology*	*Other*	*Ph.D.*	*Master*	*Other*
Testing	1017 (73.6%)	364 (26.4%)	1053 (76.3%)	114 (8.3%)	214 (15.5%)	11 (0.9%)	105 (8.7%)	1091 (90.4%)
Psychotherapy	802 (81.4%)	183 (18.6%)	862 (87.5%)	23 (2.3%)	100 (10.2%)	6 (0.7%)	64 (7.8%)	746 (91.4%)
Counseling	967 (86.2%)	155 (13.8%)	891 (79.4%)	81 (7.2%)	150 (13.4%)	10 (1.1%)	98 (10.0%)	816 (88.3%)

2.07.4 REGULATIONS OF PSYCHOLOGICAL TESTING

In 1992, the CPS drew up the Regulations of Psychological Testing. This included regulations for the use of psychological tests, and the control and management of psychological tests (CPS, 1993a). The definition of a psychological test is "a measurement tool for counseling, appraisal, and forecasting the differentiation of intelligence; for teaching students according to their aptitude; to select talented people; for occupational guidance, and clinical diagnosis." After the Cultural Revolution, psychological testing was rehabilitated and soon became popular in China. Western tests were introduced into China and revised using Chinese norms. Most tests are used in hospitals, but they are also applied in universities and colleges. Child health care institutions use intelligence and behavioral tests for children. Some Chinese tests were also developed, mainly for use in the educational field. The most frequently used Western tests are the WISC, EPQ, WAIS-RC, MMPI, and WPSSI (Gong & Li, 1996). Psychological tests used in the medical field are mainly for clinical diagnosis, counseling, and psychotherapy. Nonprojective tests are widely used, as are rating scales for personality assessment. Projective tests are seldom used (Dai, Deng, Ryan, & Paulo, 1993). The regulations for psychological testing pointed out that every member and organization under the auspices of the CPS has the responsibility to promote the healthy development of psychological testing when engaged in diagnosis, research, and the using and selling of psychological tests.

2.07.4.1 Registration of Psychological Tests

(i) All psychological tests, edited, revised, PS published, or sold by individual members or groups of the CPS must be registered at the Psychological Testing Committee of the Society.

(ii) Standardized tests which are scientifically valid can be approved, registered, classified, cataloged, and published in the official journal of the CPS, the *Acta Psychologica Sinica*.

2.07.4.2 Qualifications of Personnel Using Psychological Tests

(i) A graduate from the department of psychology and who has engaged in psychological testing for more than two years under an experienced supervisor can be approved as a qualified tester.

(ii) One who has taken part in a psychological test training workshop registered and approved by the Committee of Psychological Testing of CPS can be qualified as a tester.

(iii) The certificates of qualification are of two kinds: the certificate for a single test and the certificate for multiple tests.

2.07.4.3 Control and Management of Tests

(i) In any psychological test manual, the instruction, procedure, and user qualifications must be described in detail.

(ii) Anyone who has the qualifications of testing can buy materials of psychological tests with their certificate and is obliged to take good care of the testing materials.

(iii) The test user must strictly follow every rule in the test manual. The user must choose the appropriate tests for diagnosis and for decision making. Records must be kept and made available for examination.

(iv) Any member or institution under the CPS who wants to revise or sell psychological tests published by other authors must obtain the consent from the test author or unit. The organization who wants to print, publish, or sell psychological test materials must register at the Committee of Psychological Testing of the CPS. The buyer must also have the qualifications for using psychological tests.

(v) To guarantee the scientific nature and practical value of the tests, the content of the standardized tests cannot be published in nonprofessional magazines.

2.07.5 REGULATIONS OF PSYCHOLOGICAL COUNSELING AND PSYCHOTHERAPY

Since the late 1970s, after implementing a policy of reform and opening up to the outside world with a market economy, there has been a marked increase in the number of mental disorders (Chen & Xu, 1996). A nationwide survey was conducted in 1994, in universities and hospitals, on the number of cases of the 10 most common mental disorders and the psychotherapies used for their treatment. The results are shown in Table 2 (Gong & Li, 1996).

In October 1992, the "Regulations for Workers of Psychological Counseling and Psychotherapy in Health Systems" was

Table 2 Numbers of cases of psychotherapy and mental disorder.

Mental disorder	*Behavioral*	*Cognitive*	*Supportive*	*Psy-analysis*	*Morita*	*Biofeedback*	*Hypnosis*	*Total*
Compulsion	114 (44.4%)	39 (15.2%)	18 (7.0%)	28 (10.9%)	32 (12.5%)	8 (3.1%)	3 (1.2%)	257
Phobia	88 (42.5%)	39 (18.8%)	16 (7.7%)	21 (10.5%)	14 (6.8%)	8 (3.9%)	5 (2.4%)	207
Anxiety	41 (23.7%)	38 (22.0%)	26 (15.0%)	21 (12.1%)	17 (9.8%)	15 (8.7%)	5 (2.9%)	173
Depression	11 (8.3%)	42 (31.6%)	44 (33.1%)	16 (12.0%)	2 (1.5%)	4 (3.0%)	4 (3.0%)	133
Schizophrenia	30 (32.6%)	8 (8.7%)	36 (39.1%)	11 (12.0%)	2 (2.2%)	3 (3.3%)	0	92
Neurasthenia	5 (6.9%)	15 (20.8%)	28 (38.9%)	5 (6.9%)	7 (9.7%)	6 (8.3%)	1 (1.4%)	72
Sexual disorder	34 (47.2%)	9 (12.5%)	5 (6.9%)	13 (18.3%)	2 (2.8%)	1 (1.4%)	2 (2.8%)	72
Delusion	4 (5.6%)	27 (38.0%)	14 (19.7%)	6 (8.5%)	13 (18.3%)	2 (2.8%)	1 (1.4%)	71
Hysteria	8 (12.1%)	11 (16.7%)	9 (13.6%)	6 (9.1%)	1 (1.5%)	2 (3.0%)	25 (37.9%)	66
Neurosis	12 (22.6%)	11 (20.8%)	13 (24.5%)	6 (11.3%)	4 (7.6%)	1 (1.9%)	1 (1.9%)	53

announced jointly by the CPS and the CMHA (1993). The contents of the regulations are as follows:

2.07.5.1 Qualifications

2.07.5.1.1 Academic credentials

Academic credentials for professional positions of counseling and psychotherapy should be one of the following:

(i) A person who has received a bachelor's degree or a higher degree from a university department of psychology and has been engaged in psychological counseling and psychotherapy.

(ii) A person who has received a bachelor's degree or a higher degree from a medical school and has been engaged in psychological counseling and psychotherapy.

(iii) A person who has received academic credential of professional training in a college approved by the government, and has been engaged in psychological counseling and psychotherapy.

(iv) A person who does not have any of the above academic credentials but has a middle or higher level professional post (assistant professorship or higher) and has received training in psychological counseling and psychotherapy in a national or provincial organization and has been engaged in psychological counseling and psychotherapy.

2.07.5.1.2 Knowledge

Practitioners of psychological counseling and psychotherapy should have knowledge of theories of psychotherapy, general psychology, developmental psychology, neurology, abnormal psychology (or psychiatry), personality psychology, techniques of interview, psychological diagnosis, and psychological testing. Practitioners without a medical background should have additional medical training, such as internal medicine or pediatrics.

2.07.5.1.3 Clinical practice

Any person with the above academic qualifications should engage in psychological counseling and psychotherapy under the supervision of experienced psychotherapists for more than half a year before independent practice.

2.07.5.2 Responsibility

The code of conduct for a practitioner of psychological counseling and psychotherapy necessarily entails the following:

(i) An understanding of the nature of psychological counseling and psychotherapy, and the principle of maintaining the secrecy and rights of the client.

(ii) Collecting the information directly related to the client and understanding the client's needs.

(iii) Keeping a neutral relationship with the client. If there is any overstepping of this relationship, the treatment must be stopped.

(iv) Knowing the limits of service. During the diagnosis, testing, and treatment any demands exceeding the worker's capacity cannot be fulfilled.

(v) Applying the principle that the treatment is to direct the patients or clients to help themselves and to make efforts to improve themselves.

2.07.6 THE CURRICULUM, TRAINER REQUIREMENTS, AND TRAINING PROGRAM IN CLINICAL PSYCHOLOGY

There are different levels of training programs in China: doctoral, master, bachelor, college for professional training, advanced study, short term training, and training by correspondence. Since the 1980s, clinical psychology has become a popular profession with an increasing number of people wishing to study it or wanting to move into it from their original profession.

In May 1979, the first department of medical psychology was set up at Beijing Medical College. In 1980, a formal document from the Ministry of Public Health announced that every medical school should offer a course of medical psychology. This promoted the popularization of medical psychology. From 1985, medical psychology became the required course for medical students in many medical schools. By 1995, there were 34 medical schools (comprising 48.6% of all medical schools) in the country having medical psychology as a required course, and 36 medical schools (51.4%) having medical psychology as selected course (Li, 1995).

Among the 70 medical schools investigated, 26 have set up departments of medical psychology and 44 have placed medical psychology as a division in other departments. For example, in 15 medical schools (21.4%), medical psychology is a division under the psychiatry department; in six (8.6%), medical psychology is a division under the neurology department; and in another six (8.6%) medical psychology belongs to the physiology department. In other schools, medical psychology is combined with medical education, medical management, or social medicine.

The curriculum of the department of medical psychology in medical schools consists of general psychology, physiological psychology, neuropsychology, abnormal psychology, health psychology, psychological counseling, psychotherapy, psychological testing, psychosomatic diseases, and nursing psychology. In most medical schools (74.3%), medical psychology is taught as a complete course, but in some it is taught in two parts, the basic course and the clinical course.

The contents and teaching hours are different according to the different schools. The teaching hours of medical psychology range from 45 to 72 hours. Generally speaking, psychiatry and nursing students study more hours than students from other disciplines. A unified textbook of medical psychology for nationwide use was published in 1991 and remains the textbook for medical students.

There are 11 universities and colleges which enroll postgraduate students for master or doctoral degrees (Li, 1995). In 17 medical schools (24.3%), short term (half or one year) programs are available for further study in medical psychology.

Regarding the medical psychology faculty, 49 medical schools (70%) have full-time professors. On average, there are four teachers in each school. The professional backgrounds of the full-time professors are: 83.3% come from psychiatry, neurology, physiology, pediatrics, or Chinese traditional medicine; 12.2% come from psychology; and 4.5% come from education. Most part-time professors have backgrounds in psychiatry or neurology.

One of the great influences in the development of clinical psychology came from medical education. In the beginning of the 1980s, the "medical model" was a favored topic in the media. Today, the trend is a shift from the biomedical model to a bio-psycho–socio-medical model. This has popularized medical psychology with consequent rapid development.

2.07.7 PSYCHOTHERAPY AND CHINESE TRADITIONAL MEDICINE

Distinct from Western medicine, Chinese traditional medicine has its own principles as a special branch of medical science. There are five principles of Chinese traditional medicine applicable to the treatment of mental disorders (Yan, 1995):

(i) *Principle of Integration.* To cure an illness a holistic approach is necessary. This principle involves the utilization of more than one therapeutic method so that the patient may receive a combination of psychotherapy, physical therapy, occupational therapy, acupuncture, and traditional Chinese herbal preparations.

(ii) *Principle of Education.* This principle places emphasis on helping the patient to understand their illness through direct or indirect explanations of the symptoms, the coping mechanisms, and the causes of mental disturbance. The depth of explanation will depend on the patient's receptiveness, level of education, and desire for insight. Generally speaking, unconscious mechanisms are rarely explored directly, but dreams, art, and even calligraphy may provide useful signs of disease for the therapist.

(iii) *Principle of Mobilization.* Chinese traditional medicine lays stress on inner factors of the patient. This principle involves the process of mobilizing the patient's latent initiative and motivation for cure. The patient is helped to overcome passivity and dependence on the therapist. By mobilization, the patient is to gain self-reliance, assume responsibility for their mental health, and to cope in harmony with society.

(iv) *Principle of "Dredging."* For somatic functions, there are two principles in Chinese traditional medicine: improving blood circulation to relieve stress, and pain relief by dredging. These principles, when applied to mental disorders, refer to the process of uncovering factors which have contributed to the mental disturbance. The term "dredging" implies searching for, and bringing to the surface, obstacles which lie at the bottom of the mind.

(v) *Principle of Unity.* Traditional medicine emphasizes unity between the human being and the universe; psychotherapy deals with the unity and harmony of the patient with his family and the social environment at large.

In ancient Chinese psychology, body and mind are integrated into a single theory and a healthy mind is considered just as important as a healthy body. Psychological health is characterized by a psychological state of equilibrium through mental self-regulation, so that behavior, habits, and desires can be kept under control. This theory of integration of mind and body, which values harmony and equilibrium, is being increasingly mastered by Chinese clinical psychologists in their practice.

2.07.8 CODE OF CONDUCT AND ETHICAL GUIDELINES FOR PSYCHOLOGICAL TESTING, COUNSELING, AND PSYCHOTHERAPY

In December 1992, the CPS issued the "Ethical Guidelines for Workers of Psychological Testing" (CPS, 1993b). Its three main principles are as follows: first, that anyone who uses psychological tests in diagnosis, counseling, and for selection should bear the qualifications approved by the Committee of Psychological Testing of the CPS; second, that the user should obey the principles of scientific objectivity when applying tests; and, finally, that the user should not hinder the subject's normal functions. In particular a practitioner of psychological testing should:

(i) Have responsibility and observe scientific principles, be serious, show caution, and modesty.

(ii) Obey the national laws, and the rules and regulations of psychological testing.

(iii) Provide true and accurate information, when introducing the results and effects of psychological tests, and avoid acting impetuously, making false diagnosis, and causing misunderstanding.

(iv) Show respect for the client's personality and maintain all personal information in secrecy, except when the conditions lead to the harm to others or to society, only then can the practitioner reveal the truth to persons or units involved.

(v) Test strictly according to the requirements of the profession or society; never misuse tests for profit

(vi) Sustain the effectiveness of any tests by maintaining secrecy over any content, material, standard, and norm which are not suitable for publication.

(vii) Discuss the results of tests with clients in a correct way and provide useful help and suggestions. In general, the practitioner should only give the explanation of the test results, not the score of each test item.

(viii) Respect other practitioners and organizations of psychological testing with sincerity, learn from one another, and cooperate with one another.

(ix) Consider the contradiction, when editing, revising, and selling the psychological tests, caused by the distribution of benefits; avoid any detrimental effects to the development of psychological testing.

In October 1993, the CPS and the CMHA jointly published the "Code of Conduct for Psychological Counseling and Psychotherapy" (CPS & CMHA, 1993). This document proposed that the practitioner of psychological counseling and psychotherapy should:

(i) Respect their profession, obey the national laws, and the rules and codes of conduct of medical practice.

(ii) Be willing to help others, share empathy, show seriousness, show respect, care, and understanding to the client.

(iii) Have rich professional knowledge and the ability to interact positively with others in society.

(iv) Maintain emotional stability; specifically, to avoid accepting clients when in a bad emotional condition.

2.07.9 CONTINUING EDUCATION MANDATES

In June 1996, a National Continuing Education Committee of the Health System was established. Rules were laid down for continuing education requirements for medical workers and medical psychologists. The rules pointed out that "every health worker after graduation must take part in continuing education and obtain a score of more than 25 points each year" (Ministry of Public Health, 1996). The score of 25 points consists of two parts. The first part is a score of 5 to 10, the second part of 15 to 20. The first score is the combined result from: the national continuing education program; the ministerial or provincial continuing education program; and participation in international and national conferences. The second score is the combined result from self-study, publication of papers and books, and attending lectures and other activities.

Since the founding of the People's Republic of China in 1949, about 100 educational colleges have been set up for the training of health workers. Since 1978, many professors have been sent abroad as visiting scholars or visiting scientists for further study to acquire new knowledge and to gain experience in an international perspective.

The "Regulations of Continuing Education" was issued by the Ministry of Human Affairs in 1991. This greatly promoted continuing education in China. In 1994, the "Training Center for Psychological Counseling and Psychotherapy" was set up in Beijing Medical University. As well as Chinese lecturers, lecturers from other countries were recruited. To date, some 300 students have graduated from this training center; many have played important roles in the development of clinical psychology throughout the country.

Chinese clinical psychologists are interested in understanding and learning Western clinical psychology. Before 1949, Chinese psychology was greatly influenced by psychology from the US, but after 1949, the Russian Pavlovian psychology became the dominant trend. After the Cultural Revolution, particularly after the initiation of China's economic reform, Western psychology, especially American psychology, again had a profound influence.

Psychology, being both a natural science and a social science, is more influenced by a country's cultural background. Facing the stresses and challenges in the new era of rapid social and economic change, Chinese people are demanding more from clinical psychologists than they can offer. Chinese clinical psychologists have been doing more research in their own environment than trying to relate to Western theories and methodology. Increasingly, Chinese clinical psychologists are advancing their own theories and suggestions to solve indigenous clinical problems.

2.07.10 FUTURE DIRECTIONS OF CHINESE CLINICAL PSYCHOLOGY

With the progress of China's economic reform, there have been great changes in the lives of individuals, as well as society at large. The psychological impact and imbalance caused by these changes can be greatly felt. Demand for clinical psychology, psychological counseling, and psychotherapy is on the increase. This is a moving force for the development of Chinese clinical psychology. Equally, the improved quality of life, together with a decreasing quality of environmental influence, means that there is an increasing need for psychological guidance. Under these circumstances we can foresee the rapid development of clinical psychology in China.

Compared with clinical psychology in the developed countries there is much work to do in the near future. Some of the necessary areas of study are as follows:

(i) In clinical psychology, we should insist upon combining theory with practice, popularizing the discipline, and promoting its development. We should investigate the status of the psychological health of the Chinese people throughout the nation in order to provide a better service to the population. Improving the nation's psychological health is of great importance to China's economic construction and the future of the nation. At the same time, the application of clinical psychology can improve the people's capacity for adjustment to the social change taking place in the country.

(ii) We should work together with medical doctors to prevent psychological and psychosomatic diseases, including neurosis, personality disorder, hypertension, coronary heart disease, diabetes, cancer, and other diseases. Such diseases can be treated from a mental health perspective. We should inform people how to deal with psychosocial problems and to promote physical and psychological health.

(iii) The popularization and education of mental health will eventually help to promote the material and spiritual well-being of the people. Popular science education should be addressed to different age groups and people from different walks of life, in ways and by methods which reach the whole of society.

(iv) Central and local governments will be involved in the procedure of approval of qualifications. We can experiment on government involvement in the big cities first, such as Beijing and Shanghai, and gradually spread to the whole country. On the basis of approval of qualifications, license for psychotherapy should be given to clinical psychologists as a legal certification for practice. Meanwhile, strengthening the training programs for different levels of professionals is an urgent necessity.

2.07.11 REFERENCES

Chen, C. H., & Xu, C. L. (1996). Progress of epidemiology of mental disorder in China. *Chinese Journal of Psychiatry, 29* (2), 113–116.

Chinese Psychological Society (1993a). The regulations for psychological testing. *Acta Psychologica Sinica, 25,* 221–222.

Chinese Psychological Society (1993b). Ethical guidelines for workers of psychological testing. *Acta Psychologica Sinica, 25,* 222.

Chinese Psychological Society & Chinese Mental Health Association (1993). The regulations for workers of psychological counselling and psychotherapy in health systems. *Acta Psychologica Sinica, 25,* 223–224.

Dai, X. Y., Deng, L. X., Ryan, J. J., & Paolo, A. M. (1993). A Survey of psychological tests used in Chinese clinical psychology and comparison with data in the United States. *Journal of Chinese Clinical Psychology, 1,* 47–50.

Ding, Z. (1945). *Comprehensive mental health.* Chongqing, China: Commercial Press.

Gong, Y. X. (1993). The history and present situation of clinical psychology. *Journal of Chinese Clinical Psychology, 1, 1–3.*

Gong, Y. X., & Li, Q. Z. (1996). The investigation of current Chinese clinical psychology and its prospect. *Journal of Chinese Clinical Psychology, 4,* 1–9.

Jing, Q. C. (1994). Psychology in China. In R. J. Corsini (Ed.) *Encyclopedia of psychology* (2nd ed., Vol. 3, pp. 179–182). New York: Wiley.

Kan, D. G. (1989). The need for rapid development of medical psychology in China. *Medical Education, 9,* 30–31.

Li, X. T. (1995). *The establishment of medical psychology teaching and research units in Chinese medical colleges. (Monograph)* Beijing, China: Beijing Medical University Press.

Ministry of Public Health (1996). *The trial credit system in continuing medical education.* (Government Document).

Yan, H. Q. (1995). Psychotherapy and Chinese culture. *Journal of Chinese Mental Health, Supplement 61–62.*

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.08

The Development of Standards and the Regulation of the Practice of Clinical Psychology in India

SANTOSH K. VERMA
Postgraduate Institute of Medical Education and Research, Chandigarh, India

2.08.1 INTRODUCTION 83
2.08.2 RECOGNITION AND USE OF THE TITLE “CLINICAL PSYCHOLOGIST” 85
2.08.3 LEVELS, REGISTRATION, CERTIFICATION, LICENSURE 86
2.08.4 NECESSARY EDUCATIONAL AND TRAINING QUALIFICATIONS 86
2.08.5 SCOPE OF PRACTICE AND PRACTICE STANDARDS 87
2.08.6 CODE OF CONDUCT AND ETHICS GUIDELINES 89
2.08.7 CONTINUING EDUCATION MANDATES 90
2.08.8 FUTURE DIRECTIONS 90
2.08.9 SUMMARY 91
2.08.10 REFERENCES 91

2.08.1 INTRODUCTION

It is said that psychology has a long past but short history. It is understandable, therefore, that an applied branch of psychology—clinical psychology—will have an even shorter history, particularly in a developing country like India. Training in clinical psychology in India began in 1951 at Banaras Hindu University, Varanasi in Uttar Pradesh (Northern India), as a one-year diploma course in clinical psychology. But it took off more certainly in 1956 with a two-year diploma course in Medical and Social Psychology (now a postgraduate Master of Philosophy degree in Clinical Psychology, rather than the original diploma course) at the All India Institute of Mental Health (now called the National Institute of Mental Health and Neuro-Sciences) at Bangalore in Karnataka State (Southern India). The Indian Association of Clinical Psychology (IACP) was formed in 1968 and its official organ, the *Indian Journal of Clinical Psychology*, began in 1974 (Verma, 1984; Verma & Puri, 1986). Highlights of the historical developments of clinical psychology in India, are listed in Table 1.

At present, there are over 600 qualified clinical psychologists who received their training at the National Institute of Mental Health and Neuro-Sciences, Bangalore (338), the

Table 1 Brief highlights of development of clinical psychology in India.

Phase	*Year*	*Comment*
I (before 1950)	1921	Indian Psycho-analytic Society formed (Dr. Girija Shanker Bose)
	1925	Indian Psychological Association formed (Dr. N. N. Sen Gupta)
	1926	*Indian Journal of Psychology* started
	1930	Indian Psychoanalytical Institute for establishing psychoanalysis in practice
	1947	*Samiksha* (official publication of the Indian Psycho-analytical Society) started
		India became independent (of British rule)
II (1950–1970)	1951	Formal training in clinical psychology (one-year course) started at Banaras Hindu University, Varanasi (Uttar Pradesh)
	1956	Diploma in Medical and Social Psychology (two-year course) started at the All India Institute of Mental Health, Bangalore
	1962	Diploma in Medical and Social Psychology (two-year course) started at the Hospital for Mental Diseases, Ranchi
	1968	IACP formed (President: Dr. N. N. Sen; and Hon. General Secretary: Dr. G. G. Prabhu)
III (1971–1990)	1972	Diploma in Medical and Social Psychology started at the BM Institute of Mental Health, Ahmedabad
	1973	Ph.D. in Clinical Psychology started at the Postgraduate Institute of Medical Education and Research, Chandigarh
	1974	*Indian Journal of Clinical Psychology* (official organ of the IACP) started (Editor: Dr. S. K. Verma)
		First Professor of Clinical Psychology appointed (Prof. H. N. Murthy at Bangalore)
	1979	Workshop on training in clinical psychology (sponsored by the World Health Organization and the Indian Association of Clinical Psychologists) held at Bangalore
	1985	*Journal of Personality and Clinical Studies* (official organ of the Delhi Association of Clinical Psychologists) started at Delhi (Editor: Dr. Habib Ahmad)
	1986	Consumer Protection Act (Act no. 68)
	1987	Mental Health Act (Act no. 14) replacing the Indian Lunacy Act, 1912
IV (present phase)	1992	Rehabilitation Council of India Act (Act no. 34)
	1993/94	"Code of Conduct" adopted by the IACP during its Silver Jubilee
	1994	Policy statement of ethical considerations by Indian Council of Medical Research
	1995/96	The Persons with Disabilities Act 1995 (Act no. 1 of 1996)

Central Institute of Psychiatry, Ranchi (210), or the BM Institute of Mental Health, Ahmedabad (12). These locations are situated in the southern, northern, and western parts of India, respectively; this information has been obtained directly from the directors of these institutions.

In addition, there are a number of Ph.D.s in clinical psychology: from Bangalore (39), Ranchi (7), and the Postgraduate Institute of Medical Education and Research, Chandigarh (10); furthermore, there are others who have completed short-term courses, for limited purposes, from these and other centers, for example, in psychodiagnostics, counseling, psychotherapy, behavior therapy, neuropsychology, and projective psychology.

Clinical psychology is being taught at postgraduate (masters) level as a compulsory paper in only three universities, and as an optional paper in another 26 universities, while 27 universities have no such paper (Ramalingaswami, 1980). Each paper at this level, includes a theory/written examination with or without practicals in psychodiagnostic techniques. A formal dissertation is optional but where it exists it may be on any topic, including that of clinical psychology.

Thus, the shortage of trained practitioners in this area is understandable and unlikely to be satisfactorily overcome in the near future. On top of this is the continuous drainage due to the "three Ms" (Prabhu, 1983), that is, *matrimony* (females leaving the profession after marriage), *misplacement* (those leaving the field of clinical psychology for other professions), and *migration* (those seeking opportunities abroad, rather than remaining unemployed or underemployed).

It is not surprising, therefore, that only about half of the 600 qualified clinical psychologists are members of the IACP (IACP, 1995b).

2.08.2 RECOGNITION AND USE OF THE TITLE "CLINICAL PSYCHOLOGIST"

The title "clinical psychologist" is at present gaining increasingly wider recognition and acceptance, although some in India still continue to confuse this with "psychiatrist" or tend to call any applied psychologist clinical psychologist. Now, there are qualified clinical psychologists working in mental hospitals, general hospitals (in psychiatric, pediatric, neurological, or rehabilitation units), prisons, industry, schools, community clinics (child guidance centers, drug-deaddiction clinics, counseling centers, etc.), military hospitals, selection boards, and research institutes. Some clinical psychologists are also in private practice and they are well accepted by the community at large (Paintal, 1987; Verma, 1982).

The legal position of the profession of clinical psychologists is not completely clear. The Mental Health Act (1987), describes a "clinical psychologist" as a nonofficial member of the Central or State Mental Health Authority. (The author is a member of Chandigarh Mental Health Authority.) Surprisingly, not all State Mental Health Authorities have clinical psychologists on their boards, although there may be psychologists working in some universities, for example, in the Karnataka State. The use of the services of clinical psychologists is also mentioned in this act as one of the minimum legal facilities required for issuing a license to a psychiatric hospital or nursing home. However, the title is neither defined, nor is the necessary qualification described under this act. So, there remains scope for confusions.

Similarly, professionalism is promoted through the recently passed Rehabilitation Council of India Act (1992). The aim of this act is "to provide for the constitution of the Rehabilitation Council of India (RCI) for regulating the training of rehabilitation professionals and the maintenance of a Central Rehabilitation Register and for matters connected therewith or incidental thereto." Its members are appointed by the central government from the ministries of Welfare, Health, and Finance, and also from national bodies such as the University Grants Commission, the Indian Council of Medical Research, State Welfare departments, as well as from private rehabilitation professionals, medical practitioners, and members of parliament. It makes specific mention of "clinical psychologists" as being one of the experts in the rehabilitation professionals field and, like the Mental Health Act, stops just there, without defining this title or listing the qualifications. This is unlike other rehabilitation professionals (e.g., audiologists, speech therapists, hearing aid technicians) whose positions are well defined and their qualifications are listed along with training courses and institutes.

Law Courts in some parts of India, sometimes refer cases directly for the clinical psychologist's opinion in certain medico–legal cases. Special education schools and institutes require and demand intelligence assessment by a clinical psychologist for admission to their schools. Industries require their opinion and advice for selection or promotion as well as settlement disputes, such as, for disability compensation. Cases are also referred by medical boards, psychiatrists, neurologists, pediatricians, neurosurgeons, and speech therapists for different purposes (e.g., for differential diagnosis, counseling, assessment of cognitive functions, assessment of personality conflicts, suitability for different types of therapies, fitness to begin or continue duties after accidents, illness, disability assessments/benefits, income tax rebates). Such referrals are increasing in frequency. Clinical psychologists are also recognized by and invited to serve as mental health professionals, advisors, or experts on the panels of different national research institutes such as the Indian Council of Social Science Research, the Indian Council of Medical Research, the National Council of Educational Research and Training, the National Institute for Mentally Handicapped (NIMH), the National Institute for Visually Handicapped, and the Indian Institute of Criminology and Forensic Science. In fact the founding director of NIMH at Secunderabad happens to be a clinical psychologist. Many senior clinical psychologists (including the author) have served these research institutes for several years now with distinction as advisors, experts, teachers, or evaluators of research projects, and the scope is increasing.

Regarding the use (or misuse) of the title "clinical psychologist," Prabhu (1983) states:

> In India, there is a tendency among those with a Masters' degree or a Doctorate in Psychology to call themselves a *clinical psychologist* and claim they have *some*, *considerable*, *required* or *extensive*

> training in clinical psychology. One can emphasize with their enthusiasm but it will be dangerous and unethical to let loose all such claimants to *render service* to society. (p. xi)

One can hardly disagree with this statement but the significant point to note here is the *need* for and *importance* of the title "clinical psychologist"; otherwise few, if any, would have liked to use (or misuse) this title, except for legal purposes, when they have to use it and/or defend it in court.

2.08.3 LEVELS, REGISTRATION, CERTIFICATION, LICENSURE

At the community level there are many interested, intelligent laymen, parents, teachers, and other health professionals, including indigenous faith healers, who use commonsense, keen observation, and insight to help certain kinds of clientele, sometimes with good results too. Some of them have also written books on varied applied aspects and are recognized as "clinical psychologists" by the population at large, even though they have no formal degrees awarded by recognized institutes or universities. Such individuals have always existed, even before the recognized training institutes or universities came into existence.

There are also intelligent laymen who had studied psychology at different levels (graduate and postgraduate levels). They have taken theoretical papers such as applied psychology, abnormal psychology, clinical psychology, personality, and criminal psychology, and subsequently pose as clinical psychologists and work in the community. Some of them might have undertaken short courses in psychodiagnostics, behavior therapy, neuropsychology, counseling, yoga therapy, transcendental meditation, vipassana, or hypnosis, for example, run by some of the recognized schools or universities. They, however, are not fully-fledged clinical psychologists, even though they may be working in the overlapping areas of this helping profession, and even though some of them, at least, are rendering valuable service to the ailing human condition in a country beset with a very limited trained workforce and a large population.

All these are functioning at different levels. At one extreme are fully trained clinical psychologists correctly in position. But, at the other, there are certain untrained or inadequately trained, though intelligent, laymen, psychologists or related professionals who often misuse the title clinical psychologist. Additionally, clinical psychologists are often confused by the media with psychiatrists, psychoanalysts, or even psychologists writing on popular topics. Absence of satisfactory registration, certification, and licensure procedures only increases this confusion.

The IACP defines the term "clinical psychology" as an applied branch of psychology which, drawing heavily from the biological, medical, and social sciences, aims to study systematically all kinds of pathological deviations of human behaviors and preventive (including legal) techniques and measures (IACP, 1995a). It has five classes of individual membership, namely, fellows, professional members, associate members, honorary members, and corresponding members (IACP, 1995b). Only the first two categories are fully qualified as clinical psychologists. The IACP has prepared a list (although incomplete) of qualified professional members and has even issued registration numbers and certificates to some of these qualified clinical psychologists (for a token fee) for practicing purposes. This, however, is not a legal sanction, nor can it prevent others from calling themselves clinical psychologists, and even practicing as such.

Only the RCI (RCI Act, 1992), by virtue of its statutory powers, has been authorized by the Indian government to register with it all those who possess the requisite qualifications of clinical psychologists (however, it is not defined in the schedule published by this council), and issue valid certificates to this effect for practice as one of the "rehabilitation personnels." This, however, is a bone of contention with IACP and opinion is sharply divided amongst the members because it refers to only one of the functions of the clinical psychologists, and it clubs clinical psychologists with lesser trained professions including, audiologists, speech therapists, hearing aid and ear mold technicians, and special education teachers. The latter point has implications for the pay structure of professionals. The result is that very few qualified clinical psychologists bother to register with IACP. The controversy and the fight continues, and so does the demand for a single, separate, independent council for clinical psychologists, covering the entire field of their work.

For obvious reasons, therefore, there is currently no separate licensing system in India covering the entire activities of clinical psychologists.

2.08.4 NECESSARY EDUCATIONAL AND TRAINING QUALIFICATIONS

Both the Mental Health Act of 1987 and the RCI Act of 1992 are, more or less, silent on matters of qualifications. Nevertheless, the IACP Secretariat (Gupta, 1992, 1993) had suggested the educational and training qualifications for clinical psychologists (to be included in their schedule) to the RCI. These recommendations are as follows:

(i) *Educational qualifications*. Either a diploma in Medical and Social Psychology or equivalent Indian or foreign diploma or degree, or a Ph.D. in psychology and at least four years experience of full-time research, training, or practice of clinical psychology in a recognized institution, or a master's degree in psychology with at least seven years of experience of full-time research, training, or practice of clinical psychology in a recognized institution.

(ii) *Training programs*. The principal institutions and their courses are listed in Table 2. However, there are other premier institutions offering a diploma in Medical and Social Psychology, such as the BM Institute of Mental Health, Ahmedabad, and institutions providing for the qualifications of Doctorate in Clinical Psychology, such as the All India Institute of Medical Sciences, New Delhi. These are not included in the official schedule under the RCI Act of 1992. Brief descriptions of the towns and districts in which these institutions are situated are given in Table 3.

Steps are afoot to further modify the qualifications and training programs of the above institutions, restricting them to the degree qualifications from Bangalore and Ranchi only (Gupta, 1994). It is these somewhat contradictory recommendations, made at different periods of time, even by IACP officials, which further add to the prevailing confusion. To the author, these moves seem shortsighted and ill-advised, particularly in view of the shortage of trained professionals and the ever increasing population compared to the training resources available in the country.

2.08.5 SCOPE OF PRACTICE AND PRACTICE STANDARDS

The scope of practice and practice standards are not officially stated or clarified anywhere. They are, however, inherent in the very definition of the term clinical psychology as stated, defined, described, and implied in the IACP constitution and writings of clinical psychologists, presidential addresses, editorials, letters to the editors, and the various seminars on the role and importance of a clinical psychologist in society at large (Ghosh 1977; Misra, 1995; Prabhu, 1983; Ramalingaswami, 1976; Sen, 1975; Sen Mazumdar, 1976; Verma & Pershad, 1980; Verma & Puri 1996).

Whereas the diagnostic role has been emphasized by Ramalingaswami (1976) and the therapeutic role by Sen (1975), Ghosh (1977), Sen Mazumdar (1976), and Misra (1995); Prabhu (1983), on the other hand, has emphasized both roles (a diagnostic-cum-therapeutic role), as well as research in both the areas, depending upon what he calls SCA, that is, the "setting" of the work, the "clients" served, and the "activities" demanded of him, which may vary so widely in India.

The scope of practice is widening, taking in, for example, hospitals (mental hospitals, general hospitals, and private nursing homes in the community), community clinics (child guidance clinics, drug-deaddiction centers, family counseling clinics, marriage counseling clinics, reformatory clinics, observation homes, and orphanages), research institutes, prisons, special education schools, and private practice (part- or full-time practice as clinical psychologist). Surprisingly, however, professionals are as yet silent about practice standards and so it seems is the community at large (Verma & Pershad, 1980; Verma & Puri, 1996).

Table 2 Principal institutions offering courses in clinical psychology.

Institutions	*Courses*
National Institute of Mental Health and Neurosciences, Bangalore	Master of Philosophy in Clinical Psychology (two years) Ph.D. in Clinical Psychology (three years)
Central Institute of Psychiatry Ranchi	Master of Philosophy in Clinical Psychology (two years) Ph.D. in Clinical Psychology (three years)
Postgraduate Institute of Medical Education and Research, Chandigarh	Ph.D. in Clinical Psychology (three years)
Banaras Hindu University, Varanasi	Ph.D. in Clinical Psychology (three years)

Table 3 Main centers of clinical psychology in India.

Ahmedabad	It is the capital of the State of Gujrat and is situated in the west. There is a private business management group which runs the hospital (with no inpatients) and schools for mentally retarded and for normal children, besides many other activities. It is known as BM Institute of Mental Health. A two-year diploma course was started but was discontinued after a few years. It is now replaced by a one-year course in psychodiagnostics and counseling, for limited and circumscribed purposes only
Bangalore	One of the five largest cities of India and situated in the south (in the state of Karnataka), with one of the oldest mental hospitals in the country (the All India Institute of Mental Health, now called the National Institute of Mental Health and Neuro-Sciences). It is here that the two-year diploma course in medical and social psychology was started in a hospital setting. It has one of the largest departments of clinical psychology in the country and has to its credit many of the firsts (such as the first Professor of Clinical Psychology) in the area of clinical psychology
Chandigarh	The Union Territory of Chandigarh is also the capital of two neighboring states, Punjab and Haryana: both lay claim to this “city beautiful,” as it is known. It is also a seat of learning with many educational institutions and is situated in northern India. It has a large, general hospital (Nehru Hospital) attached to the Postgraduate Institute of Medical Education and Research (PGIMER) which runs a Ph.D. program (three years) in clinical psychology, besides short-term courses in psychodiagnostics, which is lacking in many north Indian medical colleges and hospitals. It has a drug deaddiction center which runs a continuing education program for the north Indian states. The *Indian Journal of Clinical Psychology* was started here and for its first 10 years the editorial office was here. The PGIMER, together with the All India Institute of Medical Sciences, New Delhi, were formed by an Act of Parliament as “Institutes of Excellence” to serve the ailing population of the country
Delhi	It is the capital city of India and now a State. It is situated in the center of India (earlier it was a Union Territory). It also has a training center for the Ph.D. in clinical psychology, short-term training in neuropsychology is possible and it has a drug deaddiction center within the Department of Psychiatry, All India Institute of Medical Sciences, Delhi
Ranchi	It is the summer capital of the state of Bihar and is situated in north-east India. It has three mental hospitals. It is the second training center where the two-year diploma course in medical and social psychology was started. The Hospital for Mental Diseases (formerly called the “European Mental Hospital” as only European mental patients were admitted under British rule) is now called Central Institute of Psychiatry. Here again, the training was in a mental hospital setting and it is being continued even today
Varanasi	This is a holy city for Hindus and is situated in the north (in the state of Uttar Pradesh). It is one of the oldest cities in India, with three universities and was considered as one of the important seats of learning, even in ancient India. A one-year diploma course in clinical psychology was started, for the first time in India, in a university setting (not hospital setting, although Banaras Hindu University has a medical college and hospital), at a time when psychology was still being taught as part of the curriculum in the Department of Philosophy. The emphasis was then on a psychoanalytical approach. The course was discontinued long before the author joined Banaras Hindu University for a masters degree in psychology in 1960. The start of clinical psychology, thus, was more or less symbolic but indicative of the felt need for such a course

Two related and welcome developments have taken place which are likely to have far-reaching effects in matters of practice standards. One is the Consumer Protection Act (1986), which now also extends over medical services. Rights of patients are now being increasingly recognized and accepted, and the doctor–patient relationship is being seen in a new light. Patients are now considered to be “consumers” of health services. Patients and their relatives can demand and challenge the qualities of the services rendered to them in the Consumer Protection Courts under this act. The second welcome development concerns the “Code of Conduct” recently circulated by the IACP to its members (Gupta, 1993). It is expected that clinical psychologists must possess a minimum degree of competence and should exercise reasonable care in the discharge of their duties. While the first development refers to the legal (external) control, the second is more of an internal and moral one. The first is binding, the second is not, although it is considered desirable. However, as described earlier, only half of trained clinical psychologists are members of the (IACP) and hence it has no legal control over the practice of nonmembers.

The RCI does not cover the entire range of activities of a clinical psychologist, nor has it defined the title clinical psychologist, hence very few clinical psychologists are concerned or registered with it. The Consumer Protection Act is too recent and the schedule of training institutes too incomplete to be commented upon. But it has the potential to have a restraining hand on some of the activities of a clinical psychologist—if considered objectionable—in the future. It has a legal sanction for it. Similarly, the Mental Health Act of 1987 can also impose restrictions but only through the Central or State Mental Health Authorities, which have been formed in some states with clinical psychologists as their nonofficial members. To date, however, they have not been very active but the potential for control is definitely there.

Practice standards for clinical psychologists have not been spelled out very clearly so far. Nevertheless, it cannot be said that there are no provisions for some kind of control over their activities in the light of the two developments described above.

2.08.6 CODE OF CONDUCT AND ETHICS GUIDELINES

During the Silver Jubilee year of IACP (1993/94), a code of conduct (Gupta, 1993) was adopted and circulated among the members. This concerned:

(i) *Professional competence and services.* Interest of the patient is of paramount importance and a clinical psychologist should keep abreast of recent developments in their field.

(ii) *Referrals.* In cases where proper assessment is not possible, the same should be communicated to the referring source and where the need for consultation by a physician psychiatrist or other health professional is felt necessary, the clinical psychologist should do what is necessary at the earliest opportunity.

(iii) *Method of expert opinion.* The clinical psychologist should assume full responsibility of their opinion under all circumstances.

(iv) *Consent for treatment.* Consent should be taken before starting any treatment after explaining to the patient and their available relative, the nature of illness, the method of psychological treatment, and factors associated with its efficacy and the so-called risk factors of the illness.

(v) *Patient welfare.* A clinical psychologist should maintain a high regard for the patient's integrity and welfare and should not take up a case that is not fairly within their competence.

(vi) *Court testimony.* A clinical psychologist should refrain from any bias or prejudice and their comments should be based upon the test findings and their observations.

(vii) *Confidentiality.* The information elicited from the patient and their relatives should not be disclosed to anyone other than concerned coprofessional(s) or, an appropriate authority; the clinical psychologist should not allow test material to be taken out of the clinic or laboratory, except for purposes of teaching and training students.

The IACP constitution suggests that membership may be terminated for any breach of the code of conducts. Nothing, however, is firmly laid down. Currently, the clinical psychologist is legally free to practice their profession even if IACP membership is terminated.

Pershad (1977) talks about professional ethics in psychological testing. The Indian Psychological Association constituted a National Test Commission in 1975 but appointed no clinical psychologist to its board. In addition, some training and research institutes (such as the Postgraduate Institute of Medical Education and Research, Chandigarh or the All India Institute of Medical Sciences) have their own Ethics Committees to screen all research projects, including Masters and Ph.D. theses, originating from that institute and involving human subjects. The Indian Council of Medical Research (ICMR) has put forward certain guidelines for consideration by such research institutes (ICMR circular dated September 1994). These cover such matters as the adequate protection of the rights and welfare of human subjects, whether risks are outweighed by potential benefits, adequate informed consent, the competence of the investigators, and the adequate provision of peer support.

The RCI Act (1992) has provisions for stricter punishments, for example, "imprisonment for a term which may extend to one year or with a fine which may extend to one thousand rupees or both" if individuals act in contravention of any of its provisions. However, so far these provisions exist on paper only. It has not even defined the term rehabilitation, although it mentions which certain professions fall under its provisions. The only disability covered in this Act, that may interest a clinical psychologist is of mental retardation.

However, the Persons with Disabilities Act (1995) which covers equal opportunities, protection of rights, and full participation, has defined disability (including both mental retardation and mental illness), persons with disabilities (with not less than 40% of any disability) and rehabilitation (a process aimed at enabling persons with disabilities to reach and maintain their optimal physical, sensory, intellectual, psychiatric, or sociofunctional level). This act was passed to give effect to the Proclamation on the Full Participation and Equality of the People with Disabilities in the Asian and Pacific Region, for which a meeting was held in Beijing on December 1–5, 1992, and to which India was also a signatory. Unlike the other acts mentioned earlier, there is no direct mention of clinical psychology under this act but the provision is there for rules and standards or required facilities to be provided or maintained in any institutions for persons with disabilities, to be framed by the state government. A certificate by the clinical psychologist to the degree of disability in a person with disabilities is also acceptable under this act.

A clinical psychologist under this act will have some legal control over their activities and, indirectly, shall be required to maintain certain minimum standards.

2.08.7 CONTINUING EDUCATION MANDATES

Although nothing is mandatory regarding continuing education for clinical psychologists, many nevertheless, of their own volition, do join workshops, seminars, or continuing education programs as trainees, participants, observers, or resource persons (Verma & Puri, 1996). For example, continuing education programs are being run at the National Institute of Mental Health and Neuro-Sciences, Bangalore, for those students who qualified before 1974 (Prabhu, 1983). The areas of reeducation include psychodiagnostics, behavior therapy, psychotherapy, stress management, neuropsychological assessments, depression, anxiety, mental retardation, family therapy, projective psychology, and cognitive psychology. These programs are often funded by premier national institutions such as the ICMR, the University Grants Commission, or the Indian Council of Social Science Research and conducted at any training and research institute including universities. Although this is done on a priority basis it is done as a result of an individual's initiative and is never mandatory. Joining such opportunities is easy, as is ignoring them.

Up-to-date knowledge is considered desirable although not mandatory (Verma & Puri, 1996). The IACP code of conduct (Gupta, 1993) refers to the need for a clinical psychologist to keep abreast of the latest developments in the field to enable him to provide efficient and effective services to the sick and needy. However, many such developments are controversial and expert opinion is clearly strongly divided on some of them, with the result that no firm opinion about their requirement has yet been expressed (Verma & Puri, 1996). The desirability of attending at least one of the courses in an area of one's choice, is, however, not questioned. It is the mode of regulation of the rules that has not yet been worked out to general satisfaction.

Here again, it is the absence of proper registration that sets the limits to any effort at improving the quality of services through these programs. There is no authority that can demand it and seek its compliance by all Indian clinical psychologists.

2.08.8 FUTURE DIRECTIONS

Some senior clinical psychologists have expressed their opinions through their writings (mainly in the *Indian Journal of Clinical Psychology*), presidential addresses, and talks in various national forums, seminars, workshops, or symposia on the future of clinical psychology in India (Ghosh, 1977; Mishra, 1995; Prabhu, 1983; Ramalingaswami, 1976; Sen, 1975; Sen Mazumdar, 1976; Verma & Puri, 1996).

Minimum acceptable standards have been suggested for postgraduate level with at least two years full-time diagnostic-cum-therapeutic work in supervised training (preferably an MPhil in clinical psychology) and three years supervised research (Ph.D. in clinical psychology); in short this involves a seven year training program after graduation (Prabhu, 1983). Is it possible to reduce this "wastage" of time, effort, and money by "catching them young" and providing stepwise training in clinical psychology in a well-organized manner?

The need for a separate registration of clinical psychologists, covering the entire practice of this specialty and not only for one or more aspects of the field, has been strongly felt and is being continually advocated at various national and regional forums (Prabhu, 1983; Verma & Puri, 1996). This demand is now gaining ground.

The horizons of the field are expanding and new challenges are arising (Paintal, 1987; Verma, 1982). The emphasis on different types of role for individuals, with different interests and aptitudes, working in different settings, and for different needs of the population served, has often been made (Prabhu, 1983), but not so well organized or accepted (Verma & Puri, 1996). Many clinical psychologists have strong biases in favor of or against some of these roles, which further complicates the issues. Obviously, there cannot be an ideal role model for all such professionals.

Other issues include the numbers leaving the profession, the problems of selection criteria, the point of entry into the training program, the duration and contents of the training program itself (Prabhu, 1983), the need for emphasis on codes of conduct (Gupta, 1993; Verma & Puri, 1996), and on the continuing education mandate (Verma & Puri, 1996).

However, the central and most critical issue remains that of the registration and certification of clinical psychologists by a governmental agency with full statutory powers, such as the Medical Council of India (Verma & Puri, 1996). This seems a distant goal but, in fact, it lies at the root of many problems facing this relatively young profession in India.

2.08.9 SUMMARY

Three points can be made:

(i) Clinical psychology, as an applied branch of psychology, is now being increasingly recognized, accepted, and respected, although it is relatively young and there are not many clinical psychologists in India when considered with respect to the vast population (Prabhu, 1983; Verma & Puri, 1996). There is a limited number of training centres, and there is a continuous loss of trained personnel due to migration, marriage, and misplacement (Prabhu, 1983).

(ii) In the absence of proper registration, the statutory requirements in terms of proper training, the continuous education mandate, an enforceable code of conduct, and practice standards are absent. There is nothing to stop anyone calling himself or herself a clinical psychologist. Although there is no separate, independent council of clinical psychologists with statutory powers, the term is recognized by the Mental Health Act of 1987 and the RCI Act of 1992. Medico–legal cases are sometimes directly referred by courts of law to the clinical psychologists (Verma & Pershad, 1980; Verma & Puri, 1996), and certificates issued by them regarding the degree of mental disability are accepted for legal purposes (Persons with Disabilities Act, 1995). At the same time there are greater legal responsibilities towards patient care and quality of service rendered since the Consumer Protection Act of 1986 come into force.

(iii) Again, on the positive side, there is an active association (the IACP) together with its official organ (the *Indian Journal of Clinical Psychology*). There are active professional members and training programs in addition to regular annual conferences in different parts of the country (Verma, 1984; Verma & Puri, 1996). There is now a greater tolerance of each other's viewpoints, a flexibility in overall approach to mental health problems, and an acceptance of the differing roles that a clinical psychologist can take depending upon the setting, the population served, and the activities required of him (Prabhu, 1983). Opportunities for creative innovations in different settings are there for the young clinical psychologist with enough potential and enthusiasm (Prabhu, 1983). The future is bright indeed for participants in this fast-expanding field.

ACKNOWLEDGMENTS

The author is grateful to the Director and to the Head of the Department of Psychiatry, Postgraduate Institute of Medical Education and Research, Chandigarh, for permission to write this chapter, and to the many senior clinical psychologists of the country, including the office bearers of the Indian Association of Clinical Psychologists, for providing and helping in the collection of the necessary material for this chapter, whose names are too numerous to be included here. Help and information from various training institutes, including the National Institute of Mental Health and Neuro-Sciences, Bangalore, the Central Institute of Psychiatry, Ranchi, the BM Institute of Mental Health, Ahmedabad, and the Rehabilitation Council of India, New Delhi, is gratefully acknowledged. Last, but not least, my thanks to Mr Alok Verma, computer programmer, for preparing this chapter in the prescribed manner.

2.08.10 REFERENCES

The Consumer Protection Act (1986). *Act no. 68 of 1986*. New Delhi: Government of India.

Ghosh, A. (1977). Role of a clinical psychologist: a psychiatrist's point of view (Letter to the editor). *Indian Journal of Clinical Psychology*, *4*, 198–199.

Gupta, S. C. (1992, November). Letter sent to Rehabilitation Council of India, New Delhi and copy circulated among members of IACP. Lucknow, India.

Gupta, S. C. (1993). *Code of conduct* (adopted by IACP, circulated to its members). Lucknow, India: IACP Secretariate.

Gupta, S. C. (1994, October 10). Circular to members of IACP. Lucknow, India: IACP Secretariate.

Indian Association of Clinical Psychologists (1995a). *Amended Constitution of Indian Association of Clinical Psychologist*. Jaipur, India: IACP Secretariate.

Indian Association of Clinical Psychologists (1995b). *Membership list*. Jaipur, India: IACP Secretariate.

Indian Council of Medical Research (1994, September). *Policy statement of ethical considerations involved in research on human subjects*. (Circulated to all the important research institutes in India). New Delhi, India: ICMR.

The Mental Health Act (1987). (*Act no. 14 of 1987*) with State Mental Health Rules (1990) and Central Mental Health Authority Rules (1990). Delhi, India: Law Publishers (India).

Mishra, H. (1995). Clinician clinical psychologist (Editorial). *Indian Journal of Clinical Psychology*, *22*, i–ii.

Paintal, H. K. (1987). Expanding horizons for clinical psychologists in India. *Indian Journal of Clinical Psychology*, *14*, 5–7.

Pershad, D. (1977). Professional ethics in psychological testing. *ISPT Journal of Research*, *1*, 65–68.

Persons with Disabilities Act (1995). (*Act no. 1 of 1996*. Equal opportunities, protection of rights and full participation Act) Published in *The Gazette of India*, Extraordinary, Part II, Section 1, January 1, 1996.

Prabhu, G. G. (1983) Clinical psychology: Then and now. *Indian Journal of Clinical Psychology*, *10*, i–xvi.

Ramalingaswami, P. (1976). Clinical psychology in India: Need for a new perspective. *Indian Journal of Clinical Psychology*, *3*, 53–58.

Ramalingaswami, P. (1980). *Psychology in India: Challenges and opportunities*, Delhi, India: Prachi Publications.

Rehabilitation Council of India Act (1992) *Act no. 34 of 1992*. Published in *The Gazette of India,* Part II, Section 1, no. 56, September 2, 1992.

Sen, N. N. (1975). Clinical psychology in the eighties (Editorial). *Indian Journal of Clinical Psychology*, *2*, i–iv.

Sen Mazumdar, D. P. (1976). A profession in search of its image (Editorial). *Indian Journal of Clinical Psychology*, *3*, 1–4.

Verma, S. K. (1982). The field of clinical psychology in India (Editorial). *Indian Journal of Clinical Psychology*, *9*, i–ii.

Verma, S. K. (1984). Indian Journal of Clinical Psychology: The first ten years. *Indian Journal of Clinical Psychology*, *11*, 11–20.

Verma, S. K., & Pershad, D. (1980). Psychological referrals in a psychiatric unit of a general hospital. *Indian Journal of Clinical Psychology*, *7*, 157–162.

Verma, S. K., & Puri, A. (1996). Clinical psychology in India. In *Proceedings of Asian and Indian Academy of Applied Psychology Conference*, Aligarh Muslim University, India.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.09
Clinical Psychology in Latin America

HÉCTOR C. FERNÁNDEZ-ALVAREZ
AIGLE, Buenos Aires, Argentina

2.09.1 INTRODUCTION 93
2.09.2 EDUCATION: TRAINING AND SPECIALIZATION COURSES 95
2.09.3 LEGAL ASPECTS AND REGULATIONS 97
2.09.4 AREAS OF APPLICATION 98
  2.09.4.1 Psychoanalysis 98
  2.09.4.2 Behavioral Analysis and Change 99
  2.09.4.3 The Humanistic Movement 99
  2.09.4.4 Psychodrama 99
  2.09.4.5 Systemic Therapy 99
  2.09.4.6 Cognitive Therapy 100
  2.09.4.7 Other Approaches 100
2.09.5 EXERCISING THE PROFESSION 100
2.09.6 ETHICAL QUESTIONS 101
2.09.7 PUBLICATIONS 101
2.09.8 FUTURE DIRECTIONS 102
2.09.9 REFERENCES 102

2.09.1 INTRODUCTION

Latin America is a vast geopolitical region, a territory stretching from south of the Rio Grande to the Strait of Magellan, made up of a heterogeneous group of nations forged under the cultural influence of Western European countries with a Latin heritage. Within its boundaries can also be found a few smaller areas which were created under the cultural influence of Danish and British colonizers. Due not only to these circumstances, but also to the varying living conditions in some countries, there are a few lingering doubts regarding exact borders. However, it is generally agreed that Latin America is a subcontinent made up of more than 30 countries and inhabited by some 500 million people. In 1970 the urban population amounted to 58% of the total population. By 1980 it had reached 65% and it is expected to be 77% in the year 2000. Its people are grouped into various regions and subregions: Mexico and Central America, the Caribbean and the Antilles Islands, and South America (which is, in turn divided among Brazil, the Andean nations, and the Southern Cone countries). Some countries, such as Brazil and Mexico, have very large populations (150 and 100 million inhabitants, respectively), while other nations are small both geographically and demographically.

Latin American countries have important common cultural traits, such as similarity in language (with Spanish and Portuguese predominating), religion (mostly Catholic), and the political system of their governments (republican democracies), as well as harsh socioeconomic conditions throughout most of the subcontinent. But there are a few exceptions, such as Cuba, which has had a clearly different social–political form of government since the 1960s. On the other hand, Puerto Rico is an Associate State of the USA, and Mexico is more and more closely linked to the other countries of North America and, consequently, more detached from the countries of Central America to which it is joined by historical and ethnic reasons. For some, Puerto Rico and Mexico are, precisely, the two countries with the closest cultural affinity to the USA (Díaz-Guerrero, 1995), but this is not an opinion unanimously held among psychologists in those countries.

However, there are also great differences within Latin America stemming from the social makeup, ethnic roots, and migratory currents which influenced each region. Since the period of the Conquest (1492), the continent has been exposed to successive waves of cultural invasion. The indigenous people showed greater resistance to acculturation in some areas, such as Mexico and Peru, but they could not prevent being assimilated. One of the results of this process was heavy crossbreeding between white settlers and native inhabitants creating the mestizo or Creole race. This, nevertheless, did not occur without cruel practices of wiping out native populations. In central Mexico, for example, the indigenous population declined from an estimated 25 million people in the year 1519 to just one million in 1605. In other countries, such as Chile and especially Argentina, the aboriginal population was reduced to extremely low levels, so much so that the general population today owes its makeup to successive waves of immigrants from the European continent. In another context, a significant flow of black Africans made their way to Brazil and some islands of the Caribbean where today they account for a large number of the inhabitants of those nations.

During the nineteenth-century, the cultural influence of the USA had a major impact, especially in the central zone and the Caribbean, while European influence predominated in the south. The weight of those cultures made themselves strongly felt in the nations of the continent as a whole, influencing the development of psychology. Therefore, it is impossible to speak of "a single" Latin American psychology. Instead, it can be said that there are various developments in accordance with the different regions and countries. Nevertheless, it cannot be categorically said that there has not been a Latin American psychology. The following are its most common traits:

(i) The development of psychology as an independent discipline is something relatively recent; universities began to provide courses after the Second World War. Therefore, the regulations concerning practicing the various disciplines in general, and clinical psychology in particular, only began to be implemented in recent years. An important consequence of this is, that the average age of Latin American professionals is very young and there is a scarcity of resources with specialized training.

(ii) The evolution of Latin American psychology reflects the strong impact of the psychology coming from other regions. Its development has depended on developments in the USA and Europe. The weight of North American psychology was felt strongest in certain regions like Mexico, Central America, the Caribbean, and northern Andean nations. European influence, mainly French and German, was stronger in the southern part of the continent. European psychology also had a more decisive influence in the clinical area, whereas for years the US influence was greater in other fields. The developments of the discipline in the former Soviet Union had some influence in several countries between the 1960s and 1990s, especially in Cuba, where it contributed both to university study programs and also scientific and professional practices.

(iii) Psychology in general, and clinical psychology in particular, has been mostly oriented according to practical and professional needs. Research has been scarce and has concentrated in the areas of greatest transference and application. In addition, a significant number of psychologists exercise clinical psychology as a private practice, some as their principal activity, others as a complement to another field (Marin, Kennedy, & Boyce, 1987).

(iv) The model of psychology developed in Latin America was oriented not only towards the practical, but also granted significant importance to the social and community aspects in all fields of application. In its search it oriented theory and methodology towards the use of instruments that reflected significant social commitment. (Montero, 1994).

(v) The cultural diversity led to a marked interest in the ethnopsychological developments in the various fields of application. Clinical psychology was not excluded from this process, either in the field of psychopathology or therapy. Studies in this area are perhaps insufficient, but they indicate specific concern among many psychologists on the continent.

Latin American psychologists maintain exchanges through various organizations, each corresponding to different fields of the discipline. But the Interamerican Society of Psychology (SIP) has contributed most to facilitating such exchange. This society was founded in December 1951 in Mexico during the Fourth International Congress on Mental Health. It held its first congress in 1953 in Santo Domingo (Dominican Republic) and has been holding congresses every two years since. The SIP has facilitated access to information and debate on the core aspects of the discipline, while at the same time promoting a dialogue not only among the psychologists of the continent, but also with their colleagues in North America. It has workgroups organized in each specialty, one of which is focused on the area of clinical psychology.

In this Chapter, the following aspects of the practice are dealt with: education, legal matters, fields of application, professional practice, ethical questions, publications, and future directions. Examples are selected from some of the countries which are sufficiently representative of each case. An exhaustive presentation, contemplating the situation in each country, is not possible in the available space.

2.09.2 EDUCATION: TRAINING AND SPECIALIZATION COURSES

During the first half of the twentieth-century, psychological studies were part of medical science or, more frequently, philosophical studies. During this period, drawing on the arrival of numerous European thinkers, developments in psychology were oriented towards two traditional tendencies: psychoanalysis and psychometry (Ardila, 1986; Marin et al. 1987), benefiting from the cultural permeability of the continent.

Teaching psychology in universities as an independent scientific discipline began in some countries in the 1940s, and it took hold throughout the following decade. Guatemala (1946), Colombia (1947), Chile (1948), Brazil (1953), Peru (1955), Argentina (1956), and Venezuela (1956) were the first countries to offer a degree in psychology. By the mid-1970s, those studies had spread throughout the continent. Most courses were offered by state universities, but there were some cases of nonstate university centers, for example, institutes related to religious organizations, such as Universidad de Santo Tomás, Cuba or the Pontificia Universidade Católica de Rio de Janeiro, Brazil.

In 1974, the Latin American Conference on Psychology Training (held in Bogotá, Colombia) was organized for the purpose of discussing the basic course requirements for the career in the continent and to agree on merging university policies on training psychologists (Klappenbach & Paves, 1994; Vilanova, 1993). The conference conclusions pointed to the advantage of a comprehensive basic education and the need for training to take into account both scientific and professional considerations. As will be seen further on, what really happened was somewhat different: orientation in the career and the discipline were mostly directed towards practical and applied aspects. The conference also recommended striving to make psychology an independent field of study within the university, which in fact took place not long afterwards. The original idea to hold similar encounters in the future did not happen. Instead, activities were carried out within the congresses organized by the SIP, especially since the late 1980s. That has been the venue for presenting the efforts undertaken within América Latina Formación Académica, a cooperation program between Europe and Latin America aimed at making studies in psychology at universities in Latin America compatible. In that context, the ideal studies profile for training psychologists was defined as meeting the following principles: (i) a balance between systematization and flexibility, (ii) a balance between commitment to knowledge and commitment to professionalization, (iii) a command of the various research methods and techniques, and (iv) the acquisition of the theoretical and practical capabilities and sufficient skills and techniques to measure, assess, diagnose, and develop assessment, intervention and prevention programs in the various fields of applied psychology.

For many years, very few university centers taught psychology; it was taught exclusively in state or religious institutions. This situation has changed dramatically since 1970. New private and nonreligious universities throughout the continent are now offering courses in psychology. This has helped to diversify the education offered and to create different professional models. Although resisted for many years, especially in the countries most dependent on the European model, new models finally took hold. As a result of this process, the number of centers offering studies in psychology multiplied, reflecting varying degrees of academic and scientific quality. This drew a lot of criticism, but was also a valuable incentive to increase control and regulation mechanisms for the practice, so necessary at the time. In some countries, such as Chile, the number of institutions providing these studies grew exponentially in just a few years. In 1996, both Brazil and Mexico had one hundred universities offering studies in psychology. Such a large number of educational establishments in the field poses a difficult task of coordination and regulation. In Mexico, that task was given to the Consejo Nacional para la Enseñanza e Investigación en Psicología.

As a result of this process, there has been significant growth in the number of students choosing this discipline, most of them women, according to the general trend of professionals registered in the continent over the past several years. Psychology education in Cuba has been quite different from the rest of the continent. University education is public and study plans are closely linked to government education and health policy. For many years, the basic education curriculum in psychology included developing substantial skills in strategic areas and strict knowledge of dialectical materialism (Mitjans Martínez, 1987). Changes in the world political map have produced modifications in that training emphasis but the teaching continues to strive for student training to be heavily oriented towards community needs.

The degrees in psychology given by Latin American universities basically comprise undergraduate course studies which normally last four to five years. At present there is an important trend oriented toward reforming course studies to adjust them to a four-year program in basic studies followed by postgraduate studies toward a master or doctorate as an essential requirement for exercising the profession. Those studies usually include orientations in various fields, the most common being: educational, labor, organizational, forensic, and clinical psychology. A significant percentage of students choose clinical psychology, especially in certain countries like Mexico, Brazil, Uruguay, and Argentina. In Argentina, clinical psychology has been the overwhelming orientation for many years, and psychology came to be considered practically synonymous with clinical psychology.

Clinical courses usually include experience in healthcare centers like psychiatric hospitals, general hospitals, community centers for health prevention and promotion, and other related areas. Students tend to have access to a wide scope of practice during training at government public assistance centers or private centers dedicated to specialized practices. Several universities have their own healthcare centers used for training students. Founded in 1957, the Clinical Psychological Office of the Universidad de San Marcos in Lima, Peru is probably the oldest psychology care center organized by an academic institution for treating its own staff and, at the same time, meeting the needs of the general public. Other interesting examples are the Centro de Servicios Psicológicos of the Universidad Autónoma de Mexico (UNAM), founded in 1981 (Benavidez Torres & Nuñez Obando, 1983), the Centro Asistencial de la Universidad Central de Caracas and the Clinical-Escuela of the Pontificia Universidade Católica de Campinas in Campinas, Brazil. But the model of university-run centers is still in a minority, partly due to budgetary reasons and partly because of the operational difficulties they pose.

In all cases, however, learning technologies has been a priority over fomenting research. Also contributing to that is the fact that many countries have few full-time professors. In fact, professors frequently divide their time between teaching and academic pursuits, and professional activities outside the realm of the university. But this is not the case in all countries. The prevailing practice in Puerto Rico and Brazil, for example, is very different.

The degree universities most frequently grant is Psychologist or Licensure in Psychology. This degree usually authorizes practicing any field of the profession, including clinical practice. Therefore, graduates can make diagnoses, provide psychological treatments, and carry out tasks in community mental health assistance. Despite that, psychologists interested in working in the clinical field usually receive specialized training in order to compete in the labor market. Although not a formal requirement, this is considered necessary in order to obtain the qualifications for social and professional recognition. When it was not duly formalized, this training was provided in public health centers, private training centers, and by professional associations rather than in universities. In general, it reflected a specific theoretical model or focus. Most followed the schools of behavioral change or psychoanalysis. Other important contributions to the education of clinical psychologists came from the fields of psychiatry and sociology. Psychiatrists played a large role in the teaching throughout many years, but this influence has declined considerably since 1980.

Until recently, psychologists interested in postgraduate studies attended courses at universities in the USA and Europe. Mexico was one of the first countries to offer postgraduate studies through the UNAM which, starting in 1958, organized three different programs, including one in the clinical area. The University of Puerto Rico was also one of the first institutions (1965) to offer postgraduate studies, as was the Pontificia Universidade Catolica de Rio de Janeiro, Brazil (1966). After this, postgraduate programs in clinical psychology began to be offered in Guatemala (Universidad Francisco Marroquín, 1975), Argentina (Universidad de Belgrano, 1975), Colombia (Universidad de Santo Tomás, 1977) and Costa Rica (Universidad de Costa Rica, 1979). Brazil also created the Post-Graduate Studies Program in Clinical Psychology at the Pontifícia Universidade Católica de Sao Paulo, which has offered a masters program since 1976 and doctoral studies starting in 1982. The general goals are the study of psychic reality in its component processes, its sociocultural conditions, its crises and transformations. Since 1980 there has been a major shift in this aspect, with numerous postgraduate courses becoming available. Today, most countries offer postgraduate (masters and doctors) programs in psychology.

Part of the education and training of psychologists usually includes supervision of their work. However, until the mid-1970s psychiatrists tended to perform that task. There were two reasons for this: the strength of the medical profession in mental health administration and the newness of the specialty which limited the experience of psychologists. Since the mid-1970s this situation has changed a great deal, and psychologists have become supervisors at all levels of mental health care.

2.09.3 LEGAL ASPECTS AND REGULATIONS

The creation of university level psychology studies did not mean immediate legal recognition. Legislation authorizing the practice of the profession came many years later. Before that, clinical psychologists saw their activity regulated within the laws governing medicine and their role was mainly limited to being assistants to psychiatrists. Brazil is acknowledged as the pioneer in drawing up a legal body which recognized psychology as an independent profession. In 1962 Law 4119 was passed and Article 13 established that: "The use of psychological methods and techniques for the purpose of (i) psychological diagnosis, (ii) professional orientation and selection, (iii) psychopedagogical guidance, and (iv) solution of adjustment problems shall be considered the exclusive functions of psychologists." The same legal body recognized the following as specializations in this discipline: educational psychology, clinical psychology, and labor applied psychology (Conselho Regional de Psicología, 1996).

The situation of legal recognition began to change in the mid-1970s. Legal recognition and qualification of psychologists was approved in several countries: Venezuela (1975), Colombia (1983), Puerto Rico (1983), Argentina (1985), and Bolivia (1987). In Colombia, Law 58, passed in 1983, recognizes psychology as an independent profession, regulating its exercise throughout the country in all its specialties. Article 11 of the law establishes that

> the functions of a person holding a degree in psychology include, among others, the use of psychological methods and techniques in pursuit of the following objectives: basic and applied research, teaching, psychological diagnosis, psychological treatment, vocational and professional orientation and selection, analysis and modification of individual or group behavior, and psychological prophylaxis. (Pérez-Gómez, 1990)

In Puerto Rico that same year, Law 96 regulated the practice of psychology in all its fields, authorizing "the diagnosis and application of principles, methods and procedures in order to understand, predict, influence or change behavior." In most countries, certification of clinical psychologists is granted by the health and/or education ministries. In some cases, such as Costa Rica and Venezuela, it includes registration in social security and social insurance services.

The legal recognition of psychology as a profession led to the establishment of professional organizations which, depending on the country, were called Association, Society, or College of Psychologists. Some societies group professionals in general or by specialty. For example, in Mexico, Rodríguez and Jurado Cárdenas (1986) speak of the existence of eight associations in the field of clinical psychology. In some countries, belonging to such an entity is an essential requirement to be able to practice the profession. Costa Rica saw the birth of the profession when Law 1977 created the College of Professional Psychologists, which also allowed specialization in clinical psychology. In a similar way, the Peruvian decree law of 1980 approved the creation of the College of Psychologists for the purpose of regulating the practice and assuring it is exercised in accordance with the Code of Professional Ethics. The functions of those entities have been mostly related to promoting all aspects of the discipline and, especially defending the professional working rights of their members. Several efforts have been made to set up a continent-wide federation grouping the various national entities. That has not yet been achieved in a consistent way, probably due to existing differences between regions.

As stated above, certification to practice clinical psychology is normally obtained on receiving a university degree. But this situation is undergoing great change at present. Several countries are quickly beginning to regulate new requirements for exercising the profession, which in practice means establishing the obligation of doing postgraduate studies. Following the US model, a new law in Puerto Rico stipulated that as from 1994 it is necessary to have a doctoral degree to practice clinical psychology. The reason for this requirement is grounded on the fact that clinical psychology is closely linked to health (Alvarez & Velez-Pastrana, 1995). In many countries, psychologists' associations and organizations have been promoting the creation of commissions linked to the certification of the clinical specialization with requirements in addition to possessing a degree in psychology. In Chile, for example, in 1995, the Comisión Nacional de Acreditación de Psicólogos Clínicos issued a set of rules which establishes, among the necessary certification criteria, the requirement of having completed postgraduate studies with proper supervision of practice.

Until recently, since clinical psychologists did not have to take courses to keep themselves up-to-date in the specialty, only the pressure coming from the law of supply and demand acted as a motivation for professionals to seek additional training. But in 1970, UNESCO called attention to the need to promote "permanent education" as a fundamental principle in education policy. The growth in scientific exchange through regional organizations and international events was a powerful factor which helped to create a new attitude in this respect. This development could change the conditions for teaching and training clinical psychologists, fostering continuing education, training, specialization, and updating modalities (Miranda, 1987).

2.09.4 AREAS OF APPLICATION

Clinical psychology in Latin America found fertile ground for its development in some areas, especially in the field of psychotherapy and community mental health. Psychotherapy came on the scene very early and throughout the twentieth-century has undergone continuous evolution. Among the developing regions, Latin America is the region where psychotherapy found its most propitious environment. This field has also felt the strong influence of the USA and Europe, as can be clearly seen upon examining the development of the main approaches to treatment.

2.09.4.1 Psychoanalysis

The extraordinary expansion of psychoanalysis in the Southern Cone is well known. The Asociación Psicoanalítica Argentina, founded in 1942, was the venue of some of the most significant theoretical contributions in the field. It came to have a great number of members, among them renowned international figures. Its official publication, the *Revista de Psicoanálisis* journal, begun in 1943, is one of the most prestigious publications in psychoanalysis. A division within that organization, which took place in the 1970s, led to the creation of an alternative body, the Asociacíon Psicoanalítica de Buenos Aires, with its publication *Psicoanálisis*. Both associations, which number some 1000 members each, are recognized by the International Psychoanalytical Association (IPA). After a few years of discord, the two associations have developed a constructive relationship in recent years. As an indication of the international importance of Argentine psychoanalysis, it is worth mentioning that the last president of the IPA was an Argentinian.

Different variants with strong followings arose throughout the evolution of psychoanalysis in Argentina, most significant among them being the followers of the Kleinian model. Since the mid-1970s the prevailing orientation has been that of Jacques Lacan's followers, who introduced his teachings in various places, including hospitals and universities. The phenomenon of Lacanism in Buenos Aires spread not only to the rest of the country but was disseminated throughout Latin America. The huge exodus of Argentine psychoanalysts in the second half of the 1970s contributed decisively in this regard, paving the way for continental meetings where the so-called Lacanamerican movement was furthered. An important base for that movement had been earlier established in Caracas, Venezuela. In the 1990s this school of thought has made headway in several countries of the continent, from countries with large populations like Brazil, to smaller ones such as Ecuador where a Centro de Estudios Freudianos was organized within the Universidad Católica de Guayaquil.

Brazilian tradition in psychoanalysis, taken as a whole, is rich and best seen in Rio de Janeiro. Other centers of note include Sao Paulo, Porto Alegre, and Bahia. Several associations represent this orientation in Brazil, which also has several important publications, such as the prestigious *Journal de Psicoanalise*. The journal published by the influential Asociación Psicoanalítica Uruguaya is also well-known. There too, for many years, psychoanalysis was a synonym for psychological treatment. This orientation has also undergone important developments in the rest of the continent, especially in Colombia, Venezuela, Chile, and Peru, countries which also have duly recognized psychoanalytical associations. In Panama the most relevant orientation was that of the object relations theory. Mexico has a very particular situation, reflecting great changes over time. Very powerful in Mexico in the 1950s under the influence of Frommian thought, psychoanalysis gradually fell from its zenith in the following decades. But it still has an important base there with influential organizations such as the Asociación Psicoanalítica Mexicana and the Asociación Mexicana de Psicoterapia Psicoanalítica. On a continental level, the organization of the psychoanalytical movement is best represented in the Federación de América Latina, founded in 1965 on the basis of its predecessor, Confederación de Psicoanálisis de América Latina.

2.09.4.2 Behavioral Analysis and Change

Behaviorism advanced along two paths: behavioral analysis (related to experimental aspects) and behavioral modification (linked to therapeutic aspects). Its greatest influence was felt in Mexico and in the countries of Central America and the Caribbean. But it also carried powerful weight in Venezuela, Colombia, and the countries along the Pacific. Likewise, it had numerous followers in Brazil, headquarters of a solid organization of professionals: the Sociedad Brasilera de Terapia Comportamental. In 1996 the society organized a meeting at the Universidad de Campinas which was attended by over 1000 therapists.

The importance of its development in Mexico has been seen at various levels: in the education of significant contingents of clinical psychologists, in eminent publications, and in numerous transcendental scientific events. The dominating model was the application of different therapeutic strategies derived from Skinnerian thought. One of the distinguished centers for work has been the Departamento de Psicología y Medicina de Rehabilitación de la Universidad Veracruzana, the venue of the First Annual Symposium on Behavioral Change held in 1971. Later development of this movement in Mexico shows many similarities with the movement which took place in the same theory in the US, where there was a significant movement toward the development of new procedures based on a behavioral–cognitive approach.

Important developments in behavioral change have taken place in other countries of the continent. In the Caribbean, Puerto Rico is one of the countries where most work has been carried out; another is the Dominican Republic. Guatemala and Costa Rica stand out among the nations of Central America. There have also been vigorous developments in Venezuela and Peru, grouped around organizations like the Asociación Venezolana para el Avance de las Ciencias del Comportamiento and the Sociedad de Análisis y Modificación del Comportamiento de Perú. Its evolution in Colombia was doubly significant. Not only was it a center for important activity but it was also where the Asociación Latinoamericana de Análisis y Modificación de Conducta (ALAMOC) was created. ALAMOC's periodical congresses bring together most of the specialists in the region. Its last congress was held in La Paz, Bolivia. This group has also been responsible for publications in this field.

2.09.4.3 The Humanistic Movement

The different orientations that tend to group themselves under this name are well represented in many countries of the continent, with significant representations in Peru, Chile, and Mexico. The client-centered Rogerian approach, has important representatives, highlighting the work carried out by the group at the Universidad Iberoamericana de Mexico. The last Latin American congress of this orientation was held in La Paz, Bolivia. Another approach within this school, the gestalt therapy, has made important advances in countries like Argentina, Brazil, Uruguay, Peru, Mexico, and Chile. The movement has achieved prominent status in Chile, where much of the specialized literature in Spanish is published. The last international congress of this orientation was held in Buenos Aires, Argentina.

2.09.4.4 Psychodrama

The Federación Latinoamericana of this orientation, currently headquartered in Asunción, Paraguay, groups numerous institutions and therapists from several countries throughout the continent. Argentina has been the forerunner of psychodrama but the greatest number of therapists and institutions are found in Brazil, especially in the north. In this country, it is the Federación de Psicodrama, which groups several associations in the country and has around 3000 members. Among the other countries with important developments in this field, special mention should be given to Mexico, which has a very active group in this specialty.

2.09.4.5 Systemic Therapy

A more recent development, dating from the early 1980s, systemic therapy has found strong followings in Argentina, Brazil, Chile, Colombia, Ecuador, Mexico, Puerto Rico, Uruguay, and Venezuela. For the most part, developments in this field are branches of US models, having found great acceptance in society in a short period of time. There are numerous national associations promoting local and regional meetings and producing several publications providing information on advances in the discipline on the continent.

2.09.4.6 Cognitive Therapy

This is probably the newest therapeutic orientation in the field of psychotherapy. It is currently in a phase of full expansion, especially in Argentina, Chile, Brazil, and Mexico. The first regional meeting is being planned and hopes to bring together the experience of several countries. But the best indication of the rapid growth in importance of this approach is the fact that Mexico has been chosen as the venue for the World Congress on Cognitive Therapy in 1998.

2.09.4.7 Other Approaches

In addition to smaller scale developments of other specific approaches, there are three important trends in the panorama of Latin American psychotherapy which merit special mention. One of them, with a long history, is the broad movement of group therapy. Reflecting that importance, the international association in this specialty has chosen Buenos Aires to host its congress. Second, the importance family therapy has acquired should also be noted. The main developments in this field came from, but are not limited to, systemic therapy. A monographic issue of the *Revista Interamericana de Psicología* provides up-to-date information on developments in this field (Bernal & Alvarez, 1989).

Finally, the role of psychotherapy in Latin American is highly significant in its support of the integration movement (Fernández-Alvarez, 1992). The first signs of interest in this area go back to the early 1980s. But this tendency gained greater momentum throughout the 1990s, especially as the result of the activity of some groups of professionals, among the most important being, the Fundación AIGLE headquartered in Buenos Aires, Argentina, and Centro Científico de Desarrollo Psicológico in Santiago, Chile.

Community mental health is a field where a great deal of work and creativity has taken place on the continent as a result of a combination of the pressing needs of the population and the high level of social commitment of many Latin American scientists and professionals (especially psychologists). Literature on this topic emphasizes that numerous and varied pilot programs in community mental health have arisen in Latin America (Levav, 1992). Many of these projects were experimental and almost always implemented in urban areas.

Psychologists joined the interdisciplinary teams which put those programs into practice. Pointed examples include the Mental Health Project in Barrio Santa Eduviges, Tegucigalpa, Honduras; the Centro de Atención Psicosocial de Ciudad Sandino, Nicaragua; and the mental health program in the community of Curundú, Panama. Some of these projects have been collaborations with university institutions, such as the Centro de Asistencia a la Comunidad of the Universidad de Rosario in Argentina, which was created in 1984 with three areas of activity: mental health care, education, and social prevention. Its objectives include the creation of a community health center and a network of social work and health promotion, in addition to providing education and professional training.

The mental health programs started on the continent share some common characteristics. These projects sponsor preventive and health promotion activities, particularly through work in schools and other grassroots community organizations. In most of the projects, special importance is given to solving problems with a marked psychosocial bias. The greatest number of programs involving clinical psychologists are in the areas of mental health care and rehabilitation of persons suffering from severe mental disorders, alcoholism, and drug abuse, as well as broaching problems related to violence and crime. The 1990s have seen a great deal of effort being directed to the fight against AIDS. The available professional resources are usually scarce, requiring fast and intensive training courses. In many cases, as a means of multiplying the preventive effects, emphasis is placed on creating self-help groups.

2.09.5 EXERCISING THE PROFESSION

Practicing clinical psychology is conditioned by three overriding factors: (i) the economic limitations of the continent and of each region, (ii) the power of the physician in general, especially that of the psychiatrist, and (iii) the struggle between the different clinical–theoretical orientations. For years, Latin American psychologists found many job opportunities due to the favorable relation between the number of professionals available and existing job supply. The latter centered round two large sectors: on one hand, work in government institutions, especially those in the area of education and health which proposed the implementation of community programs; and, on the other hand, in public and private practices dedicated to the diagnosis and treatment of mental disorders. The number of psychologists working in state health institutions has increased from the 1970s. Taking Costa as an example, its work there was divided between practices in psychiatric hospitals and general hospitals, planning and programming health policies (including preventive activities) and healthcare activities in urban areas (Campos, 1991).

The work of psychologists in government programs has obviously been dependent on the changing conditions which occur in each country. In brief periods of time the economy of many countries in the continent have undergone great fluctuations which affected their social policies. Nonetheless, for a long time the governments of the region have promoted social service plans which favored setting up community action programs. The intensification of the socioeconomic crisis in Latin America in recent years has had a negative impact on state programs. State reforms and low budgets have made it difficult for governments to carry on significant public assistance and community action plans. This has grown worse due to cutbacks in foreign aid granted by international cooperation agencies.

The activity of psychologists in mental health services has been covered through various sources of financing. Hospital services care for the neediest sectors of the population, which in some countries amount to a very large number of people. Health insurance plans run by trade unions and powerful institutions cover the healthcare needs of their members. Other sectors of the population, belonging to the middle and upper-middle strata, choose private health care. More recently, prepaid health insurance plans have appeared as an alternative for covering some mental health needs. The job panorama for psychologists in this field has been growing worse because of two factors: first, the huge growth in the number of professionals, owing to the rise in university graduates, producing a negative balance between supply and demand. This did not happen in Cuba because the professional training of psychologists has been planned by the government in line with the employment needs of the country. Second, less funds have been available in society because of the economic crisis already mentioned.

The institutions in charge of mental health administration began to intervene heavily in planning services, usually with the primary aim of changing the cost/benefit ratio with an eye toward lowering service costs. They encouraged short-term therapy methods and reduced the number of professionals working in private practice. As a result of this, professionals saw their earnings dwindle. In addition, there has been a steady rise in the unemployment rate. At present, there is no managed mental healthcare policy in Latin America as has been gradually implemented in the USA, but steps being taken seem to be moving in that direction. World globalization makes it possible to suppose that the transfer of policies from dominant societies to dependent societies will take place rapidly. Regardless, such change must necessarily adapt to the peculiarities of the new scenario.

2.09.6 ETHICAL QUESTIONS

As the process of regulating the practice made headway at all levels, a growing number of codes of ethics and ethics tribunals began to work at regulating practice standards. Qualification of the profession is usually accompanied by a code of ethics. Professional associations—psychologists colleges in most countries—are entrusted with the task of ensuring that ethical norms are met. In this sense, practice standards are regulated by peers, based on the legal prescriptions and rules current in each country. The users or consumers of psychological services should resort to ordinary courts when they believe they have been injured by a practicing professional.

There are codes of ethics in all countries with many principles in common, strongly inspired by the practicing standards of the American Psychological Association. As an example, Panama's code, drawn up in 1992, describes ethical principles in the following areas: responsibility, competence, legal and moral norms, public statements, confidentiality, people's welfare, professional relations, assessment techniques, research with human participants, and the care and use of animals in research. In accordance with the 1984 law in Honduras governing the College of Psychologists, the Norms of Professional Ethics in effect are aimed at orienting the exercise of psychology as a profession according to the duties which psychologists must fulfill in relation to their patients, colleagues, and society as a whole.

2.09.7 PUBLICATIONS

Even though great efforts have been made in the continent to provide a flow of biographical material, these efforts have often been frustrated by the lack of continuity in the publications, mostly due to economic difficulties. However, some journals have managed to appear regularly and maintain their level of excellence throughout the years, such as the already mentioned *Revista de Psicoanálisis*. Among the oldest publications, the following are worth mentioning: *Arquivos Brasileiros de Psicologia* (Brazil, founded in 1949); *Acta Psiquiátrica y Psicológica de América Latina* (Argentina, 1954); *Revista del Hospital Psiquiátrico de La Habana* (Cuba, 1959); *Revista Interamericana de Psicologi/a* (Interamerican Society of Psychology, 1967); and *Revista Latinoamericana de Psicología* (Colombia, 1969).

These publications contain articles covering various fields of psychology, including clinical psychology. There are also specialized publications reflecting a significant example of the work of Latin American clinical psychologists, such as: *Avances en Psicología Clínica Latinoamericana* (Colombia); *Revista Argentina de Clínica Psicológica* (Argentina); *Terapia Psicológica* (Chile); and *Salud Mental* (México)

There is also a substantial number of publications on various specific approaches.

2.09.8 FUTURE DIRECTIONS

According to World Bank data (1993), in 1990, Latin America had an average per capita income of US$2190 dollars. But it was also one of the three regions in the world with negative economic growth since 1980. The profound crisis that has been afflicting the region will require psychologists to come up with creative answers to respond to an already very visible paradox: a population with greater needs which must be helped with fewer resources.

An overview of the evolution of psychology in Latin America indicates that its most positive aspects have been the high social significance of its practices and the important theoretical and technological development it has achieved despite only counting on scarce resources. The most negative aspects are related to the weakness of culturally specific proposals, because the models and instruments used were mostly imported from other regions. This was felt most in the rural areas where the population groups with greatest ethnic singularity live. For the upcoming years, scientists foresee an increase in demographic displacement towards urban centers, which will probably be accompanied by the creation of human settlements with difficult living conditions. This leads to expectations of significant growth in substance abuse, violence, and other social pathologies which will require doubling the efforts made in community aid activities.

The most urgent aspects on which the future development of clinical psychology in the region should focus are: (i) emphasizing cultural values at all levels of development; (ii) legalizing regulations governing the profession, with specifications regarding certification, supervision of practices, and requirements for continuing professional education; (iii) consolidating postgraduate studies and specialization training programs; (iv) fomenting research at all levels, without ignoring the importance of developing new technology; and (v) being increasingly accurate in designing community action programs with an eye toward more efficiency.

2.09.9 REFERENCES

Alvarez, V. & Velez-Pastrana, M. C. (1995). La profesión de la psicología en Puerto Rico. *Revista Latinoamericana de Psicología, 27,* 175–185.

Ardila, R. (1986). *La psicología en América Latina. Pasado, presente y futuro.* Méxieco: Siglo XXI.

Benavidez Torres, J. Y., & Nuñez Obando, R. (1983). *Una década de la facultad de psicología: 1973–1983.* México: UNAM.

Bernal, G., & Alvarez, A. I. (1989). Psicología y terapia familiar desde Latinoamérica. *Revista Interamericana de Psicología, 23.*

Campos, G. A. (1991). La psicología clínica en las instituciones de salud costarricense. *Revista Costarricense de Psicología, 9*(18), 59–72.

Conselho Regional De Psicologia (1996). *Manual do CRP—06 1996.* Sao Paulo, Brazil: Conselho Regional de Psicologia.

Díaz-Guerrero, R. (1995). Origins and development of Mexican Etnopsychology. *World Psychology, 1,* 49–67.

Fernández-Alvarez, H. (1992). *Fundamentos de un modelo integrativo de psicoterapia.* Buenos Aires, Argentina: Paidós.

Klappenbach, H. A., & Pavesi, P. (1994). Una historia de la psicología en Latinoamérica. *Revista Latinoamericana de Psicología, 26,* 445–482.

Levav, I. (1992). *Temas de salud mental en la comunidad.* Washington, DC: Organizacíon Panamericana de la Salud.

Marin, G., Kennedy, S., & Boyce, B. C. (1987). *Latin American Psychology: a guide to research and training.* Washington, DC: American Psychological Association.

Miranda, R. (1987). La educacíon continua en la enseñanza superior. *Enseñanza e Investigacíon en Psicología, 13,* 13–23.

Mitjans Martínez, A. (1987). *La formación del psicólogo en Cuba.* La Habana, Cuba: Facultad de Psicología.

Montero, M. (1994). *Psicología social comunitaria.* Guadalajara, México: Universidad de Guadalajara.

Pérez-Gómez, A. (1990). *Hacia el siglo XXI: La psicología como ciencia y como profesión en Colombia.* Bogotá, Colombia: Universidad de Los Andes.

Rodríguez, M. A., & Jurado Cárdenas, S. (1986). Las Asociaciones Profesionales de Psicólogos en México. *Revista Mexicana de Psicología, 3,* 197–202.

Vilanova, A. R. (1993). La formación de psicólogos en Iberoamérica. *Acta Psiquiátrica y Psicológica de América Latina, 39,* 193–205.

World Bank (1993). *Informe sobre el desarrollo mundial 1993.* Washington, DC: World Bank.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.10

Guidelines for the Development, Use, and Interpretation of Credentialing Examinations

CRAIG G. SCHOON
Professional Examination Service, New York, NY, USA

2.10.1 INTRODUCTION 104
2.10.1.1 History of Regulatory Processes in the Professions 104
2.10.1.1.1 The political theory of regulation 104
2.10.1.1.2 A brief history of regulation 105
2.10.1.2 Forms of Mandatory and Voluntary Practice Regulation 105
2.10.1.3 The Impact of the Credentialing Process 106
2.10.1.4 The Role of Standards and Guidelines 106
2.10.2 PRIMARY STAGES IN THE DEVELOPMENT, USE, INTERPRETATION, AND EVALUATION OF CREDENTIALING PROGRAMS 107
2.10.2.1 Conducting a Strategic Credentialing Analysis 108
2.10.2.1.1 Define legal status of credentialing organizations 108
2.10.2.1.2 Define purpose of the credential 108
2.10.2.1.3 Define theory of credentialing 109
2.10.2.1.4 Define key stakeholders of the credentialing organization 110
2.10.2.1.5 Define key values and philosophy of the organization 110
2.10.2.1.6 Define what makes the credential unique within the context of other credentialing activities within the same professions 110
2.10.2.2 Conducting an Occupational Practice Analysis 110
2.10.2.3 Developing Assessment Tools 112
2.10.2.3.1 Develop assessment tools that achieve the credentialing mission of the organization 112
2.10.2.3.2 Describe the credentialing theory in terms of the assessment tools that will be used to implement the theory 112
2.10.2.3.3 Build assessment tools from validated test specifications 113
2.10.2.3.4 Build assessment tools that meet the needs and expectations of one's stakeholders 113
2.10.2.3.5 Set performance standards for the assessment tools 114
2.10.2.4 Conducting Validation Analyses 114
2.10.2.4.1 Conduct validation analyses that address the evidential and consequential basis of text interpretation and test use 114
2.10.2.5 Minimize Construct-irrelevant Variance and Construct-underrepresentation 117
2.10.2.6 Conduct an Ongoing Program of Research, Evaluation, and Policy Appraisal 118
2.10.3 CONCLUSION 118
2.10.4 REFERENCES 119

2.10.1 INTRODUCTION

This chapter presents a discussion of the purpose and function of guidelines and standards for the development, use, and evaluation of licensure and certification programs. A brief history of the regulatory process in the professions is presented, including a discussion of the theory, purpose, and function of credentialing as a regulatory procedure. Then, the need for standards and guidelines in the development, use, and evaluation of licensure and certification examinations is examined, with an emphasis on how guidelines and standards serve as aids in meeting the needs and expectations of stakeholders in the credentialing process.

2.10.1.1 History of Regulatory Processes in the Professions

2.10.1.1.1 The political theory of regulation

One underlying ideal of regulation, which might be termed "libertarian," holds that the public is entitled to maximum liberty in personal decisions regarding the purchase of products and services, including such services as healthcare and those that may affect their financial well-being. The ideal extends the same liberty to professionals wishing to provide services and products, thus maximizing freedom to practice along with freedom to choose. Accordingly, a libertarian ideal would minimize regulation of services and products by government, since, in theory, regulation could restrict the range of products and services that the public may choose from (Friedman, 1962).

A converse "totalitarian" ideal reflects a belief that the state is best suited to make decisions regarding the production of goods and the provision of services. Under such a model, control of the means of production and of the delivery of services is deemed to be in the best interests of the public. The need for this control rests on the public's presumed inability to make correct decisions, either due to lack of sufficient facts upon which to make a rational choice, or because those who produce products and provide services wish to limit the ability to make decisions that may endanger their interests.

In a democracy such as that of the USA, the public's interests are balanced between trust in the public's freedom to make its own choices and the responsibility of the state to protect the health and welfare of its citizens. It is in a democracy that the struggle between the libertarian and the totalitarian ideals takes place. To what extent is the government responsible for limiting the public's choices so that the public's health, welfare, and safety are protected? To what extent should the public be free to choose as they wish, and to experience the consequences of their choices? The issue of regulation of professional practice has been at the center of this debate (Friedman, 1962).

The conflict between consumer freedom and government regulation of professional practice is well summarized in a publication by the Association of State and Provincial Psychology Boards (ASPPB) (1990), the body responsible for the licensing of psychologists in the USA and Canada:

> The laws are intended to protect the public by limiting licensure to those persons who are qualified to practice psychology as defined by state and provincial [Canadian] law. The legal basis for licensure lies in the right of the state to enact legislation to protect its citizens. *Caveat emptor*, or "buyer beware," is felt to be an unsound maxim when the "buyer" of services cannot be sufficiently well informed to beware, and hence states have enacted regulatory boards to license qualified practitioners. A professional board is a state or provincial agency acting to protect the public, not serve the profession.

The regulation of professional practice, then, makes the assumption that the citizens of the USA are not sufficiently able to make decisions regarding the purchase of services that affect their health, safety, and welfare, either because they lack sufficient information or because the government has a responsibility to exert some influence on the selection of qualified practitioners—a benevolent form of *in loco parentis*. From this perspective, the cost of making a wrong decision is considered too great: the loss of life, health, safety, or welfare are decision outcomes against which the public must be protected.

The debate regarding the pros and cons of regulation takes several forms. On the libertarian side, it is argued that regulation severely restricts consumer choice by limiting the supply of practitioners to those who meet regulatory requirements. In a recent issue of the *Wall Street Journal*, licensure of cabdrivers was listed as one of the main causes of the decrease in quality of life for New Yorkers, due to both the limitation of the number of cabs operating and the fare increases resulting from the limited competition.

From the libertarian perspective, regulation of professional practice is viewed as government interference in the market operating to the advantage of those who are licensed, since a licensed professional can charge more for services because of this limitation of competition. The libertarian perspective may also presume the existence of explicit strategies by practitioners to exploit state regulation laws that restrict entry into professional practice, thereby increasing the economic advantage to those who meet regulatory requirements.

On the regulation side of the argument are those who believe that the protective function of regulation is ill-served by the political process. According to this perspective, professions may unduly influence legislators to enact regulatory laws that lack genuine benefit to the public. Practice-regulation laws, to be truly effective, should regulate only those professions where malpractice will harm the public's health, welfare, and safety. Those espousing this model further hold that too many professions have achieved regulated status, including professions characterized by little or no capacity to cause harm to the public. Examples of this from the literature include such statements as, "Some of the more unusual professions regulated in at least one state include: Babcock testers, bankruptcy salespersons, wire rope inspectors, lime vendors, mussel fishers, pheasant club operators, safe mechanics, apprentice scalers, resident and non-resident sea moss rakers, tree injectors, weather modifiers, livestock weighers, lightening rod installers, hemp growers, endless chain agents, and egg brokers" (Brinegar, 1990); and "What individuals in these occupations actually do and why regulation is needed may not always be comprehensible" (Schmitt, 1995, p. 8).

This chapter explores how psychologists and others in the measurement community may contribute to this debate by developing measurement and evaluation standards and guidelines applying to the development, use, and evaluation of the decision goals held by stakeholders of the regulation process. As psychologists and measurement specialists, our goal is to provide a scientific structure to frame the debate over practice regulation. Our assumption is that debate in a democracy is an essential tool by which the citizens of a country evaluate and assess those products, services, and governments which best meet their needs and expectations. Published standards and guidelines help to facilitate this evaluation process, so that the public's interests may be met.

2.10.1.1.2 A brief history of regulation

The medical profession was the first to be regulated by law in the USA, with Virginia being the first state to enact regulations restricting the practice of medicine (in 1639), followed by Massachusetts (1649) and New York (1665). The wording of the first regulation statutes points toward public protection as their purpose. Thus, in Massachusetts, the law indicated that the activities of the following individuals should be regulated: "Chirugeons, Midwives, Physitians or others [who were] imployed at any time about the bodye of men, women or children, for preservation of life, or health" (Shryock, 1967, p. 1).

After passage of these initial laws regulating the practice of medicine, regulation by law of professional practice declined until, by the mid-1850s, virtually no state was engaged in regulation of any profession, including medicine. By that time, it was assumed that those who chose to practice medicine were adequately trained by existing educational institutions, who mounted effective efforts to reinforce and promote such a belief among legislators. The same held true for the legal profession; by 1840, almost anyone who wished to practice law could do so by engaging in an apprenticeship controlled by the profession itself.

The regulation of professional practice reemerged in the latter part of the nineteenth century, spurred by state medical societies who were concerned about low standards of education and training in new medical schools. The lobbying of state medical societies for regulation was effective, with Texas being the first state to establish a state examining board in 1873. Licensure of medical practice then spread throughout all states, with New Hampshire being the last state to license the practice of physicians in 1915.

Today, licensure has spread broadly across many professions. The Council of State Governments reports that states currently regulate over 1000 professions (Brinegar, 1990). The concern has been over the inflation of regulation. As Schmitt reports:

> In the immediate past, the emphasis was on licensure, licensure, and more licensure without the appearance of much regard for whether the laws were necessary to protect the public from harm or quackery. Any profession that could get the support of a senator or representative had an excellent chance of obtaining licensure status. In fact, licensing legislation may have been based not so much on logic, but rather on who introduced the bill, who the lobbyist was, and how much financial backing was available. (Schmitt, 1995, p. 9)

2.10.1.2 Forms of Mandatory and Voluntary Practice Regulation

The three vehicles that are commonly employed to regulate professions are practice control (licensure), government and voluntary title control (certification), and registration. In 1977 the US Department of Health, Education, and Welfare defined licensure as:

> the process by which an agency of government grants permission to an individual to engage in a given occupation upon finding that the applicant has attained the minimal degree of competency necessary to ensure that the public health, safety, and welfare will be reasonably well protected.

This definition remains in effect today, and licensure laws are enforced through the police power of individual states. State laws also specify the prerequisites for obtaining licensure status. An individual proven in a court of law to be practicing medicine without a state license is subject to imprisonment.

Government title control establishes specific legal criteria governing the use of a given occupational or professional title. This practice does not restrict the practice of the profession, but rather restricts the titles a practitioner may use. For example, while a Certified Public Accountant (CPA) is a practitioner who has met standards established by the state, nonCPAs may, nonetheless, practice accounting. Consumers of accounting services may use the title to select a practitioner who has achieved established and explicit standards of education, experience, and achievement on standardized examinations. Government title certification is used, then, to provide useful consumer information to the users of title-protected professions.

Voluntary title control, usually known as certification, may be undertaken by professions or trades that wish to establish standards signifying the attainment of specified levels of education, experience, and achievement. Practitioners who voluntarily meet such standards are certified as having attained them by their profession or trade group. Thus, many physician groups certify areas of specialty practice in medicine.

In this chapter, we will refer to the activities involved in all forms of regulation, both mandatory and voluntary, as credentialing. McCready (1982, p. 74) indicates that:

> Credentialing is a generic term that subsumes licensing, certification, registration, and institutional licensure by the states, as well as standards of competence where no licensure is required and certification by private organizations where it is required for practice by reference in state law.

2.10.1.3 The Impact of the Credentialing Process

Credentialing decisions are "high-stakes" decisions. In addition to receiving substantial economic and social benefits, those who are credentialed may be charged with a legal responsibility to engage in practice that does not harm the public. The cost associated with this liability is considerable, as are the consequences of malpractice or unethical behavior. Likewise, substantial expense has already accrued to an individual who is denied a credential. The financial and time investment in preparation for the profession is forfeited, and the economic, social, and psychological consequences of failure to receive a credential may be very significant. Along with the implied benefit to the public of valid credentialing decisions, the credentialing process entails a public cost associated with restriction of the number of practitioners, with its accompanying increase in the cost of the services provided. All of these consequences indicate the high-stakes nature of credentialing decisions.

Credentialing outcomes exert an impact on a host of stakeholders, including the public, the members of the professions, professional associations, educational institutions, membership associations, legislators, members of state boards, testing organizations, standard setting organizations, and others. Given the significance of this impact, stakeholders desire a credentialing system that meets their needs and expectations, and that may be characterized as reasonable, rational, fair, and scientific.

In the USA, stakeholders of credentialing systems may seek to influence the credentialing process through available political and social channels. Thus, members of a profession will utilize the political process to promote the interests of their profession, seeking introduction of appropriate legislation as well as working through their professional associations. Likewise, other stakeholders of credentialing may also access appropriate channels for pursuing their particular interests in credentialing outcomes. For example, the public may seek, through its political representatives, to enact credentialing systems that genuinely meet their stated public-protection objectives, while public-interest organizations such as Fair Test and the Citizens Advocacy Center may be used to promote public interests and to ensure that needs and expectations are met.

2.10.1.4 The Role of Standards and Guidelines

To the extent that the stakeholders of credentialing call for outcomes that are reasonable, rational, fair, and scientific, they use standards and guidelines as an important means of evaluating credentialing programs. The *Standards for Educational and Psychological Testing*, published by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (AERA, APA, NCME) states the following regarding the purpose of the *Standards*:

> Educational and psychological testing has also been the target of extensive scrutiny, criticism, and debate both outside and within the professional testing community. The most frequent criticisms are that tests play too great a role in the lives of students and employees and that tests are biased and exclusionary. In consideration of these and other criticisms, the *Standards* is intended to provide a basis for evaluating the quality of testing practices as they affect the various parties involved. (AERA, APA, NCME, 1985, p. 1)

In general, then, to the extent that a given credentialing program has a significant impact (economic, physical well-being, status and/or class membership) on the stakeholders affected by that policy (the public, members of the profession, educational institutions, foundations, etc.), stakeholders will call for evidence that the selection policy is reasonable, rational, fair, and based on scientific principles. Moreover, the greater the impact of the credentialing decision on stakeholders, the greater the potential for challenges to the credentialing program's outcomes, and therefore the more rigorous the evidence required in support of the policy. According to Schoon and Smith (1996a):

> The evidence a credentialing organization provides to support its policy decisions is a function of the expectations of its stakeholders. Consequently, if an organization considers an independent standard setting body such as the National Commission for Certifying Agencies (NCCA) as one of its stakeholders, the credentialing organization will try to follow and meet the NCCA standards and guidelines. Also, an organization whose stakeholders include the scientific measurement community will most likely adhere to the *Standards for Educational and Psychological Testing* (AERA et al., 1985). These independent standards require that considerable evidence be provided to support the reasonableness, rationality, fairness, and scientific foundation of a credentialing program. (pp. 6–7)

2.10.2 PRIMARY STAGES IN THE DEVELOPMENT, USE, INTERPRETATION, AND EVALUATION OF CREDENTIALING PROGRAMS

There are several published standards and guidelines relevant to the development, use, interpretation, and evaluation of credentialing programs. These include the *Standards for Educational and Psychological Testing* (1985), jointly issued by AERA, APA, and NCME; the *Guidelines for Computer-Based Tests and Interpretations* (1988), issued by the APA; the *Principles for the Validation and Use of Personnel Selection Procedures* (1987), issued by the APA Division of Industrial–Organizational Psychology; the *Principles of Fairness: An Examining Guide for Credentialing Boards* (1993), issued by the Council on Licensure, Enforcement, and Regulation and the National Organization for Competency Assurance; the *ETS Standards for Quality and Fairness* (1987), issued by the Educational Testing Service; the *Development, Administration, Scoring and Reporting of Credentialing Examinations: Recommendations for Board Members* (1993), issued by the Council on Licensure, Enforcement and Regulation; the *NCCA Guidelines for Certification Approval* (1991), issued by the National Commission for Certifying Agencies; the Americans with Disabilities Act (1990); the *Uniform Guidelines on Employee Selection Procedures* (1978), issued by the Equal Employment Opportunity Commission, Civil Service Commission, US Department of Labor, and US Department of Justice; and the *Guidelines for the Development, Use, and Evaluation of Licensure and Certification Programs* (1995), published by the Professional Examination Service (PES). These standards and guidelines address the means by which reasonable, rational, fair, and scientific credentialing decisions can be made, either explicitly as standards address specifically to credentialing issues, or implicitly as generic principles applicable to all measurement situations.

The PES *Guidelines for the Development, Use, and Evaluation of Licensure and Certification Programs* (the PES *Guidelines*) (1995) surveys the documents mentioned above and incorporates into a single publication the standards, guidelines, and principles as they are applied to credentialing programs. The PES *Guidelines* assume as a guiding principle the perspective offered by Messick (1989) that validity is a unitary concept that can be termed "construct validity," a concept that encompasses the evidential and consequential bases of test interpretation and test use.

Messick was critical of the *Standards for Educational and Psychological Testing* (AERA et al., 1985) from several perspectives. He points out that the *Standards* encourage the perception that, in many situations, only one kind of validity evidence suffices for establishing the validity of the assessment tool (Messick, 1989, p. 92). This has led, for example, to an over-reliance on establishing content validity as the sole means of demonstrating the validity of credentialing examinations (Popham, 1978). Messick also criticized the *Standards* for its lack of emphasis on the unitary nature of validity, and particularly for its failure to address the value implications of test interpretation and the social consequences of test use, as aspects of construct validity. As our summary of the theory and history of credentialing has indicated, credentialing is rife with value implications and social consequences. Messick believes that these issues must be addressed if credentialing outcomes are to be adequately evaluated in relation to stakeholder needs and expectations.

The PES *Guidelines* incorporate Messick's validation perspective as the theoretical foundation for the development, use, and evaluation of credentialing examinations. The perspective taken in the PES *Guidelines* is that the meaning and purpose of a credentialing program should be explicitly stated in what may be termed a statement of credentialing mission. This explicit statement constitutes a "theory of credentialing" whose meaning and consequences may be rationally tested.

The process of evaluating credentialing programs is not only a scientific issue, but a political and social activity as well. Thus, a comprehensive evaluation cannot rest solely on, say, a correlation coefficient, or the judgment of a content expert. A correlation coefficient or an expert judgment arises from, and must be interpreted within, the social, political, and scientific context of our culture. Again, Messick (1995) states that:

> validity, reliability, comparability, and fairness are not just measurement principles, they are social values that have meaning and force outside of measurement whenever evaluative judgments and decisions are made. As a salient social value, validity assumes both a scientific and a political role that can by no means be fulfilled by a simple correlation coefficient between test scores and a purported criterion . . . or by expert judgments that test content is relevant to the proposed test use. (p. 742)

Messick goes on to state that validity constitutes an evaluative summary of both evidential and consequential outcomes of score interpretation and use. The PES *Guidelines* accept this perspective and suggest specific measures that will assist organizations in addressing evidential and consequential validation issues.

2.10.2.1 Conducting a Strategic Credentialing Analysis

According to this perspective, the statement of an organization's credentialing mission is the foundation of the program's validity edifice. The PES *Guidelines* suggest, therefore, that the initial stage consists of a complete "strategic credentialing analysis," leading to the formulation of a statement of the credentialing mission; such an analysis should address the following issues: (i) the legal status of the credentialing organization, especially if the organization is classified as not-for-profit, in which case credentialing activities may be an element in the justification of exemption from taxation; (ii) the purpose of the credential and how credentialing outcomes will protect and/or inform the public or benefit the members of a profession; (iii) the theory of credentialing, which describes how the credentialing program will be conducted to effect credentialing decisions, and its hypothesized outcomes; (iv) the key stakeholders of the organization and how the stated credentialing mission will serve their needs and expectations; (v) the key values and philosophy of the organization; and (vi) what makes the organization's credentialing mission unique within the context of other credentialing activities within the same profession.

2.10.2.1.1 Define legal status of credentialing organizations

The legal status of a credentialing organization has a direct relation to the purpose of a credentialing effort. Thus, if the credential involves licensure, then the credential's purpose is stated by constitutional and state law to involve the protection of the public's health, safety, and welfare. Validation efforts in regard to licensure examinations, then, must be related to the extent to which the health, safety, and welfare of the public is served. As Messick (1995) states:

> Broadly speaking . . . validity is an inductive summary of both the existing evidence for, and the potential consequences of, score interpretation and use. Hence, what is to be validated is not the test or observation device as such but the inferences derived from test scores or other indicators—inferences about score meaning or interpretation and about the implications for action that the interpretation entails. (p. 13)

It is clear, then, that the legal purpose of licensure is to protect the public health, safety, and welfare, and that validation activities must relate to that legal purpose.

2.10.2.1.2 Define purpose of the credential

Similarly, many credentialing organizations are classified legally as nonprofit corporations, and their nonprofit status implies a social purpose, as stated in the laws governing tax-exempt status. As a nonprofit entity, the organization pays no taxes on any excess revenue exceeding operating expenses. Hence, as an entity supported by public tax dollars, the nonprofit organization must adopt a mission that benefits the public, either directly or indirectly. For example, organizations organized as 501(c)(3) corporations, which are termed "charitable" in the tax code, must have an explicitly stated mission to serve the public good. Several testing organizations—including American College Testing, Educational Testing Service, the PES, and the National Board of Medical Examiners—have been granted 501(c)(3) status, and thus are required by law to have a stated public service mission. Holding 501(c)(3) status is also not uncommon among professional organizations.

Most professional certification and membership organizations are organized under a status known as a 501(c)(6) corporation. According to tax law, such organizations are called "business leagues," and their purpose is to serve the interests of the profession. Thus, although they are granted nonprofit status because the promotion of professional interests has been deemed to be in the best interest of the public, credentialing organizations are not legally bound to promote public health, safety, and welfare as an explicit purpose of their certification activities. Some of these organizations may define the purpose of certification as a means of denoting attainment of established performance standards within their profession. The implication of voluntary certification, then, does not necessarily include a benefit to the public; rather, its purpose may be to promote the interests of the profession, a mission that is in accord with the legally stated function of 501(c)(6) organizations (Hopkins, 1992). Alternately, some voluntary credentials denote competence in practice areas that do impact public health or safety; in these cases sponsoring organizations may incorporate such a claim into their statement-of-purpose examples.

As we have seen, an explicit statement of the purpose of a credential is essential to establishing the grounds for evaluating the outcomes of a credentialing program with respect to its stated intention. This explicit statement constitutes a pledge that a credentialing program's purpose is clearly stated and open to public scrutiny and evaluation. For example, it is not uncommon for critics of regulation (Dawes, 1994; Friedman, 1962; Shimberg, Esser, & Kruger 1973) to discredit the assumed purpose of licensure to protect the public. One chapter in Dawes' book, *House of cards: Psychology and psychotherapy built on myth* (1994) is entitled, "LICENSING: The Myth of Protecting the Public." Most authors who are critical of licensure deny that the true purpose of licensure is to protect the public; rather, they believe that its purpose is to promote the interests of the profession.

The function of standards and guidelines in this debate is to suggest that the credentialing body state publicly the purpose of a credential, thereby rendering the program open to evaluation and appraisal in terms of the evidence that can be accumulated in support of the stated purpose. Open declarations of the purpose of a credential contribute to an open and evidence-based debate on the more scientific question of whether or not the credentialing outcomes are in accord with their stated intention.

2.10.2.1.3 Define theory of credentialing

The PES *Guidelines* suggest that the statement of the purpose of the credential be supported by a description of the theory of credentialing that underlies the statement of purpose. By a "theory of credentialing" is meant a description of the assumptions underlying the credentialing program. For example, most credentialing programs require specified amounts or types of education, experience, supervision, performance on examinations, and character verification, and the organization usually assumes (hypothesizes) that the standards applied to each of the "credentialing variables" relates to the desired credentialing outcome. To become a licensed psychologist, for example, a candidate must attain:

> a doctoral degree in psychology from an approved program, or the equivalent as determined by the board. … One or two years supervised experience in a setting approved by the board. … Demonstration of relevant knowledge through passing an objective written examination … citizenship, age, and other requirements, as well as requiring evidence of good moral character. (ASPPB, undated, pp. 2–3)

The hypotheses underlying credentialing criteria such as these are that the required performance standards for each variable relate to the knowledge, skills, abilities, and character necessary for minimum competence at the entry level, in the practice of a profession. The PES *Guidelines* suggest that the requisites for credentialing be stated as testable hypotheses, in the form of a credentialing theory, and that the desired outcomes be hypothesized so that objective evaluation of the credentialing outcomes can be conducted.

Schoon and Smith (1996a, 1996b) contend that each of the variables in a credentialing selection system (Cronbach, 1980) should be associated with a rationally established performance standard. Rather than placing sole emphasis on the setting a rational passing point for a credentialing program's written examination, (rational in that it relates to the purpose of the credential, for example, selecting minimally competent candidates), the establishment of rational and explicit standards should also be extended to requisite education, experience, character verification, and any other relevant criteria. Each of the requirements for these variables should relate specifically to the stated purpose of the credentialing program. In this regard, Friedman (1962) believes that the amount of education required for attainment of a credential is often only marginally related to a licensure program's stated purpose of protecting the public. He argues, in the case of medicine, that the educational-plus-experience model places emphasis on excellence to a far greater extent than would be required to produce physicians capable of serving the public competently. The position of the PES *Guidelines* is that such questions can be framed as empirical issues related to the purpose and theory of the credential, which should be addressed to meet the needs of an organization's stakeholders.

2.10.2.1.4 Define key stakeholders of the credentialing organization

Accordingly, the PES *Guidelines* suggest the conduct of a "stakeholder analysis" is to describe those to whom the credentialing organization has an obligation, or upon whom the credential has a significant impact. The PES *Guidelines* (1.4, p. 21) encourage a stakeholder analysis with the goal that those affected by the outcomes of a program's credentialing decisions should be systematically included in the formulation, design, and implementation of the policy underlying the establishment of the credentialing program.

2.10.2.1.5 Define key values and philosophy of the organization

Messick (1989), as we have seen, believes that, included in the construct validity paradigm, are the value implications and social consequences of test interpretation and test use. In keeping with this concept, the PES *Guidelines* recommend that a credentialing organization "Describe the key values and philosophy of the organization as they relate to the credentialing mission" (1.5, p. 21). Given that credentialing outcomes can have a significant, high-stakes impact on the lives of their stakeholders, relevant values might include accuracy, reasonableness, rationality, fairness, due process, timeliness, honesty, openness, responsiveness, and sensitivity to stakeholder feedback. By publicly subscribing to values that are relevant to the needs and expectations of its stakeholders, a credentialing organization facilitates external evaluation of the extent to which these values have been incorporated into credentialing policy.

Messick (1989) argues that measurement of any kind, and the associated outcomes of that measurement, take place within a given value perspective. Outcomes will be interpreted within that value perspective, and may well serve one value perspective rather than another. Thus, making the organization's particular value perspective explicit is an important element in determining the overall "meaning" of the credentialing program and its intended outcomes, and including effort to include perspectives of all stakeholders will provide the value context within which the outcomes of a credentialing program may be evaluated.

2.10.2.1.6 Define what makes the credential unique within the context of other credentialing activities within the same professions

The PES *Guidelines* recommend, as part of the strategic credentialing analysis, that an organization describes what makes the credential unique within the context of other related credentials. For example, in many professions, there may be competing or closely related credentials. For the sake of making clear to stakeholders the credential's precise purpose, the organization must clearly distinguish it from others. This can be done in terms of the theory of credentialing, by specifying differences in such requirements as education, experience, previous credentials, or types of examinations passed, by delineating any difference in the stated purpose of the credentials. Again, the purpose of credentialing should be to enlighten, not to confuse, those holding a stake in the credential, while achieving the intended result of the credentialing decisions made.

2.10.2.2 Conducting an Occupational Practice Analysis

After the credentialing mission has been established, conducting an occupational practice analysis is the initial step in generating empirical evidence relating to the theory of credentialing stated in the strategic credentialing analysis. As such, the occupational practice analysis enumerates, through traditional scientific means, the *constructs* underlying credential performance, that is, needed levels of education and experience; previous credentials; knowledge, ability, and skill levels; and character traits essential to successfully meeting standards desired to obtain the credential.

With regard to validation of credentialing examinations, it has been stated that, "the problem is not one of evaluating tests, it is one of developing and validating hypotheses" (Guion, 1976, p. 791). This hypothesis-testing approach to credentialing program—embodied initially in the conduct of an occupational practice analysis—holds that:

> The answer with respect to applied hypotheses is the same as for substantive or scientific hypotheses, namely, from the implications of construct theories. In applied instances, of course, these would typically be construct theories of performance domains as well as construct theories of the critical aspects of performance that it would be important to diagnose, certify, or predict in connection with applied decision making. Thus, domain constructs and criterion constructs—that is, the knowledge, skills, cognitive processes, personal attributes, or whatever that are entailed in successful domain performance—provide a rational basis for selecting attributes that are representative of the domain or likely to be predictive of criterion success. Construct theories of each identified attribute then serve to guide test construction and test evaluation. Furthermore, the construct theory of the performance domain also affords a rational basis for appraising the meaning and adequacy of the criterion measures themselves. (Messick, 1989, p. 64)

A critical outcome of this scientific approach is the generation of a "construct-referenced" test, rather than a traditional "criterion-referenced" examination. Here, the goal is to utilize the constructs enumerated in the occupational practice analysis as the basis of examinations which, once administered, generate construct-valid scores. It is this type of result that will: "maximize the meaningfulness of score interpretation and minimize construct-irrelevant test variance. The resulting construct-valid scores then provide empirical components for a rationally defensible prediction system and rational components for empirically informed decision making" (Messick, 1989, p. 65).

As pointed out by Schoon and Smith (1996b), the occupational performance analysis is the foundation of the entire credentialing program, and should be used to establish performance standards for all the variables in the theory of credentialing that relate to the credentialing selection system:

> Individuals who establish performance standards should have a thorough understanding of the conceptual foundations of the credential, and should have as much evidence as possible regarding the behaviors they will be designating as defining credentialed status. One important source of data is the occupational practice analysis, job analysis, or role delineation. Those establishing performance standards should have a thorough understanding of the practice analysis, since its goal is to define performances critical to the definition of credentialed behavior and to delineate the knowledge, abilities, and skills that underlie these performances. The statements in the analysis relate directly to the task of establishing performance standards that are reasonable, rational, fair, and relevant to the behavior to be credentialed. The practice analysis should also address the amount of education and experience necessary to acquire the knowledge and skill required for successful performance at the specified performance level of the credential; the practice analysis is, therefore, relevant to the establishment of performance standards for all the variables in the credentialing selection system. (p. 154)

Although there is no single standard method for conducting an occupational practice analysis, commonly used approaches include evaluations by committees of experts in the profession, reviews of work logs, observations of practice, interviews with practitioners, and surveys of practice (Schoon, 1985). Often, more than one of these approaches is used to generate information, so that the constructs generated from one source can be checked against the constructs generated from another (Knapp & Knapp, 1995). The methodologies used to generate data for the occupational practice analysis should be described in detail to facilitate replication of the procedures used at a later date, or by an independent investigator or agency.

Inclusion of all affected stakeholders in the definition of the constructs to be assessed in a credentialing program will contribute to the validity of the credentialing decisions made, and ensure support and acceptance of the credential.

Generating a representative sample of credential stakeholders is described in the PES *Guidelines*, as follows:

> Describe the rationale for selecting the participants in the data collection process, including committee members, practitioners, educators, or others who are identified as legitimate sources for the data that will comprise the occupational practice analysis. Care should be taken to represent all the major areas of practice within the occupation and to achieve a distribution among all geographical areas and practice settings where practice may differ in scope or emphasis (e.g., urban vs. rural, differing educational levels).

> Describe the source of recommendations for participants in the process, and how and by whom final selections of participants were made, including who was not selected and for what reasons. All sampling procedures used to select participants must be described, along with the goal of the sampling process. (2.3, p. 26)

> The purpose of the validation process is to: (a) generate data on the criticality, importance, and frequency of all elements in the practice analysis; (b) indicate if any performance domains, tasks, or Knowledge-ability-skills (KSAs) were overlooked or incorrectly included; and (c) determine if the analysis accurately portrays those behaviors that are critical to the purpose of the credential, and thereby to protecting and/or informing the public and/or promoting the profession. It is necessary to thoroughly describe the procedures used to collect and analyze the validation data, including sampling techniques, rating scales, open-ended responses, or other means of evaluating the practice analysis data. (2.4, p. 26)

Generation of appropriate validation data is essential in building the measurement and hypothesis-testing structure of a sound credentialing program. The analysis of this data will explicate the logical connections that link the constructs derived from the practice analysis with each assessment tool, and the assessment tools with the theory of how credentialing decisions are made to accomplish the mission of the organization.

Specifically, the validation data must be analyzed: (i) to determine what constructs can and should be assessed to accomplish the objectives of the credentialing program; (ii) to identify the assessment tools required to produce construct-relevant scores that can be used to make valid and rational credentialing decisions; (iii) to generate the test specifications for each assessment instrument; and (iv) to generate the instructions to be used by experts in developing the response formats (e.g., multiple-choice examination or clinical simulations) that will best link the construct-related assessment tools to the constructs enumerated in the occupational practice analysis.

2.10.2.3 Developing Assessment Tools

2.10.2.3.1 Develop assessment tools that achieve the credentialing mission of the organization

In a "selection system" (Cronbach, 1980), all variables used to make selection decisions are included as "tests" or "assessment tools." Most of the focus of standards and guidelines has been on explicitly designated assessments such as written examinations, practical examinations, and clinical simulations; these assessment tools are expected to meet the highest technical and psychometric standards when used in high-stakes testing activities such as credentialing. If we consider credentialing, however, to be a selection system, then each selection variable in that system should receive scrutiny in relation to relevant guidelines or standards. Thus, we have made the case that a credentialing organization should, as a first step in the credentialing process, list all the variables in the credentialing selection system, such as education, experience, supervisors ratings, or character references. These variables comprise, in their relationships and structure, a theory of credentialing. Hence, in discussing the standards and guidelines relevant to the development of assessment tools for a credentialing program, our focus will be on the entire array of variables used to make a credentialing decision. This perspective is in accord with the concept of a coherent credentialing selection system, each element of which should be based on a firm scientific and policy foundation.

The general goal statement for developing credentialing assessment tools, then, can be stated as follows:

> Develop an assessment program that achieves the credentialing mission by utilizing data from the strategic credentialing analysis and the occupational practice analysis. Assessment tools should include all formats for decision-making, including such factors as amount of education required for the credential, amount of experience required, supervision documentation, recommendations, or graduation from accredited schools or training programs. The assessment selection system should follow from the theory of credentialing and provide evidence that can be used in the validation of that theory. (PES *Guidelines*, p. 29)

2.10.2.3.2 Describe the credentialing theory in terms of the assessment tools that will be used to implement the theory

Development of the assessment tools for a credentialing selection system should follow logically from the list of selection variables that constitute the theory of credentialing; this listing will have acquired an empirical foundation via the occupational practice analysis. For example, if the practice analysis indicates that the knowledge, skills, and abilities required to practice at the entry level of a profession require a high school education, then using attainment of a college degree as a selection criterion would be considered overly restrictive, that is, such a tool would yield a selection ratio that exceeds the estimated base rate of entry-level competence required for safe practice.

The theory of credentialing and the occupational practice analysis will also address the predicted outcomes of application of the credentialing selection system, and each selection tool should be rationalized in terms of its place in the credentialing selection system. Only selection variables relevant to desired selection outcomes should be included in the credentialing selection system. In terms of the PES *Guidelines*:

> Describe the theory for the assessments that will constitute the tools to be used to make credentialing decisions. This theory should describe the rationale for the amount of education required for the credential, the amount of experience, supervision, or other achievements, and the reasons for using particular assessment formats, such as multiple-choice examinations, clinical simulations, computer-based examinations, practical examinations, oral interviews, or in-basket exercises. The description should include the rationale that links performance on each assessment tool to the theory of credentialing and to the occupational practice analysis. (3.1.1, p. 29)

2.10.2.3.3 Build assessment tools from validated test specifications

The assessment tools that are hypothesized to achieve the credentialing mission are constructed according to the "test specifications" that are the end product of the occupational practice analysis. The test specifications establish a logical connection between the response formats of the assessment tools and the constructs derived from the occupational practice analysis. For example, in the case of a written examination, each item must be judged for its relevance to producing the desired credentialing outcome. In addition, affirmative response must be generated to questions regarding the linkage between the item and the construct it is to measure. For example, after identifying an item in a psychology licensing examination as associated with the construct Competence in Diagnosing Mental Pathology, we could then ask, "Does this item provide information relevant to the construct in question?"

Generally, the initial construction of assessment tools proceeds from test specifications to format generation based on judgments of the logical connection between response formats and the constructs to be measured. The nature of these judgments will relate to the type of outcome decisions that are to be made with the assessment tool. According to Messick (1989):

> As a starting point, when an existing test is to be evaluated for its applicability to a particular job, judgments by job incumbents, supervisors, or other experts could be obtained for the extent to which the test specifications match the importance dimension for the job. Better yet, at a more detailed level, each test item could be judged for the extent to which the knowledge and skills tapped are pertinent to the marker tasks that define the job-importance dimensions [Rosenfeld, Thornton & Skurnick, 1986]. At a more global or molar level with respect to the job, each test item could be judged as to whether or not it taps knowledge or skill that is essential to overall job performance; more globally with respect to an ability or skill test, each of the specific job tasks could be judged as to whether the tested skill is essential for that task performance [Lawshe, 1975]. (p. 69)

2.10.2.3.4 Build assessment tools that meet the needs and expectations of one's stakeholders

Judgments about the applicability of assessment tools to the purpose of the credential should be made by the stakeholders identified in the stakeholder analysis conducted as part of the strategic credentialing analysis, and all stakeholders significantly affected by the credentialing outcomes should be represented. For example, an organization's theory of credentialing may state that credentialing outcomes will be valid indicators of entry-level competence, and that this outcome is not intended to be biased against any minority group. Such a statement would be supported by evidence that minority representatives judged the test formats relevant to hypothesized construct domains and established that the formats were not biased against minority candidates. The actual outcome of the assessment tool regarding bias would have to be subjected to empirical test.

In general, if the stakeholders of the credential demand that the credentialing program be based on procedures that are rational, reasonable, fair, accurate, and scientific, then all response formats should be judged by stakeholders for such criteria as: (i) adherence to standards for item construction and review; (ii) adherence to guidelines for sensitivity review that seek to eliminate language that is sexist, racist, or discriminatory in any way toward any group of candidates; (iii) adherence to guidelines for clarity, accuracy, and consistency; (iv) assurance that test manuals give instructions that are clear, simple, accurate, and complete; and (v) assurance that candidates are not in any way adversely affected by the typography, layout, format, or response method of the examination (3.3.2, p. 33).

2.10.2.3.5 Set performance standards for the assessment tools

Setting performance standards for assessment tools is the critical means of linking the assessment tool to the theory of credentialing outcomes. The assessment tool in and of itself is simply a means of gathering response information regarding construct domain performance. That information is then accessed in reference to a performance standard which is theorized to produce selection outcomes that are in accord with the organization's theory of credentialing (Schoon & Smith, 1996b).

The method used to establish performance standards for the assessment tools included in the credentialing selection system should be described in detail, including the size and composition of the group of experts used to establish the standards, the hypotheses linking expert judgments with desired credentialing outcomes, the method used to obtain expert judgments that will be in accord with the desired outcomes, and the psychometric theory, if any, underlying the standard setting method applied to the assessment tool.

The stakeholder analysis conducted previously again becomes useful in choosing a representative group of stakeholders to comprise standard setting groups. Omission of an important stakeholder group may place the credibility of the credentialing program at risk. If a group of stakeholders that are significantly affected by the outcomes of the credentialing decision are not included in setting the performance standards, this group can claim, with justification, that their needs and expectations were not met by the credential. If the ultimate stakeholder of licensing examinations, the public, are not included in setting performance standards, such a claim could result in the elimination of the credentialing program. Although a public representative may not be qualified to make expert judgments about an issue such as construct relevance, or the probability that a minimally competent candidate would pass a given item, the representative would be capable of assessing whether the process employed to make such judgments is in keeping with the interests of a consumer of the credentialing outcomes.

It is the position of the PES *Guidelines* that establishing performance standards is as much an issue of credentialing policy as it is a purely psychometric exercise. Because credentialing outcomes take place within the value structure of our culture, they have significant social effects. Accordingly, the expert judges that establish performance standards should be versed in the organization's theory of credentialing, the credentialing mission, and the results of the occupational practice analysis. As stated in the PES *Guidelines*:

> Once the committee members have adequate familiarity with the mission, theory, and occupational practice analysis, they should be presented with a description of the standard setting methodology. The exact manner of presenting the methodology must be described so that it can be presented to an organization's stakeholders, replicated, and evaluated. The mission and goals of the credentialing program must be clearly described in relation to each of the selection variables for which a standard will be established. Questions for discussion should be recorded, along with a summary of the discussion held. These questions, and the summary of the discussion, should be public documents that could be made available to the public and to all relevant stakeholders of the credential. (3.1.3, p. 31)

The above guidelines apply to each variable in the credentialing selection system. For example, in terms of the amount of education required to attain a credential, the type and amount of education needed, and for what purpose; this should be presented in the theory of credentialing. The "performance standard" established for requisite education should be set by the credentialing organization's stakeholders, and these stakeholders should be versed in the theory of credentialing as it applies to education, with the goal that the amount of education required for a credential is neither more nor less than that required for the intended credentialing outcomes.

Critics of credentialing (Friedman, 1962; Pew Report, 1995) have indicated that the amount of education required to be eligible for licensing in many professions is not adequately related to actual entry-level practice requirements, and is therefore not a valid "assessment tool" in the credentialing selection system. Their charge is that, if the outcomes of educational selection were validated against entry-level requirements, then "excellence" would seem to have been the performance standard rather than "minimal competence." To apply scrutiny only to written examinations, then, ignores the validation of all the other assessment tools in the credentialing assessment system.

2.10.2.4 Conducting Validation Analyses

2.10.2.4.1 Conduct validation analyses that address the evidential and consequential basis of test interpretation and test use

The validation of a credentialing program consists of evaluation of the adequacy and appropriateness of the inferences and actions that follow from the program's credentialing decisions. According to Messick (1989):

> Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and action based on test scores or other modes of assessmentelan inductive summary of both the existing evidence for and the potential consequences of score interpretation and use. Hence, what is to be validated is not the test or observation device as such but the inferences derived from test scores or other indicators—inferences about score meaning or interpretation and about the implications for action that the interpretation entails. (p. 13)

(i) Conduct validation analyses that address the evidential basis of test use

The starting point for construct validation of a credentialing selection system is the collection of inductive evaluations of the hypotheses stated in the theory of credentialing. The processes of conducting an occupational analysis, developing test specifications, and developing assessment tools from these specifications have typically been evaluated for adequacy only in terms of their content validity. This approach tends to focus on the assessment tool, in and of itself, as valid, without consideration of the decision outcomes produced by the use of the tool in a credentialing program. This has been particularly true of credentialing, where the role of content validity has been justified as the only validity that, in essence, can be achieved (Popham, 1978).

Messick (1989) criticizes this contention, which underlies some portions of the *Standards for Educational and Psychological Testing* (AERA et al., 1985), stating that:

> The major problem here is that so-called content validity is focused upon test *forms* rather than test *scores*, upon *instruments* rather than measurements [italics added] ... because validity, reliability, comparability, and fairness are not just measurement principles, they are social values that have meaning and force outside of measurement, whenever evaluative judgments and decisions are made. As a salient social value, validity assumes both a scientific and a political role that can by no means be fulfilled by a simple correlation coefficient between test scores and a purported criterion ... or by expert judgments that test content is relevant to the proposed test use. (pp. 41–42)

The problem with many content-validation approaches is that they fail to focus on the nature of the functional relationship being validated. Validation, and most scientific activity, can be seen as the evaluation of observations concerning variables that are hypothesized to be related in some functional way. In the case of content validity, the functionally related variables are the content of the assessment tools and the domain of reference, or construct domains specified in the occupational analysis. Content validation of credentialing examinations, then, focuses on ensuring that an empirically verifiable relationship exists between the content of assessment tools and the constructs of the occupational analysis, as indicated earlier in this chapter. These studies, however, should constitute not the end, but rather the commencement of the validation of a credentialing program.

The PES *Guidelines* suggest that the matrix of correlations amongst the variables comprising the theory of credentialing will provide evaluative information regarding the adequacy and appropriateness of a theory of credentialing. Examples of hypotheses are:

(i) There will be a significant correlation between grades obtained in school and performance on the written credentialing examination.

(ii) There will be a significant positive correlation between grades obtained in the school-based clinical training and performance on a clinical simulation examination required for credentialing.

(iii) There will be a significant correlation between the amount of education received and performance on examinations for credentialing.

(iv) Experience, as measured by supervisor ratings, will correlate positively with both school performance and performance on credentialing examinations.

Correlational findings in regard to these hypotheses will provide evidence regarding the theory of credentialing from which they emerged. For example, if no positive relationship exists between school performance and performance on the credentialing examination, what is the implication for the hypothesis that education provides the necessary skills and abilities necessary to practice the profession? Are the essential KSAs not being taught? Are schools too selective, so that there is restriction of range to the extent that a positive correlation could not occur (indicating that the performance standard established for entry into and graduation from school may not be adequately related to required performance at the entry level).

Evidence in relation to the hypotheses embodied in the theory of credentialing can be obtained readily and inexpensively, and such evidence can have significant implications for both the theory and the policy that stems therefrom. Gathering evidence regarding these assumptions and hypotheses will contribute significantly toward the interpretive meaning of the credentialing assessment tools and, therefore, of the credential and its supporting theory.

The PES *Guidelines* (4.1.7, p. 40) also recommend that concurrent validation studies be undertaken to investigate the relationship between performance on credentialing assessment tools and on outcome measures such as practical examinations, oral examinations, supervisors' ratings, or other outcome measures. Such studies, even though they may not meet the scientific requirements for criterion-related validation, can contribute to the interpretive meaning of the variables comprising the theory of credentialing. For example, state credentialing boards in several professions have constructed practical examinations, which may be administered in addition to a mandated written examination, to achieve enhanced measurement of critical skills and abilities. In some states, for example, veterinary medicine boards have constructed practical examinations involving hands-on skills and abilities in diagnosing and treating animals. These boards have established criteria for minimally acceptable performance for entry-level practice of veterinary medicine.

In cases such as this, where multiple assessment instruments are in use, correlational studies on different types of examinations can be undertaken to good effect. Taking the veterinary medicine example, one could hypothesize that there would be positive significant correlations amongst the written, clinical simulation and practical examinations, but that the correlation between the clinical simulation and the practical examination results should be higher than that between the written examination and the practical examination. The findings of correlational studies such as this have considerable implications for the development, use, and interpretation of the national and state credentialing examinations, and on the credentialing theory for the veterinary profession (Jones & Schoon, 1993; Jones, Schoon, & Lovler, 1994).

In the area of professional licensure, the conduct of genuine outcome studies is impeded by the fact that failing candidates do not enter the profession. Nonetheless, concurrent validation studies of licensure programs can and should be conducted for the evaluative evidence they provide in reference to the adequacy and appropriateness of the theory of credentialing. In the area of voluntary certification, however, the lack of a control group does not, in theory, preclude effective criterion-predictive studies. In this case, were the theory of credentialing to hypothesize that certified practitioners will perform more effectively than their noncertified counterparts, or that certified individuals will produce enhanced outcomes relating to client health, safety, and welfare, these studies can be conducted with full methodological justification. Such hypothesized claims can be scientifically examined, since both certified and noncertified individuals can practice the profession, and their practice outcomes on the same patient populations could be measured for predicted differences. Indeed, such studies should be conducted by certification organizations if their theory of credentialing states that certification benefits the public health, safety, and welfare.

(ii) Conduct validation analyses that address the evidential basis of test interpretation

Because of the high-stakes outcomes of credentialing examinations, and because validation entails the evaluation of the inferences and actions resulting from credentialing decisions, the PES *Guidelines* suggest that an organization:

> Conduct studies that gather evidence regarding the appropriate interpretation of the outcomes of assessment tool use in the credentialing selection system. The evidence gathered should support the stated purpose of the credentialing program, as described in the strategic credentialing analysis and credentialing mission statement. Studies should be conducted for each of the assessment tools in the assessment selection system, with the intention of providing evidence that each contributes to making credentialing decisions consistent with the credentialing mission. (4.2, p. 41)

As we have seen, such studies may be difficult to conduct in the case of licensing examinations. However, these studies may be possible in the case of certification programs. For example, assume that the purpose of a given certification program is to promote the profession by recognizing voluntarily acquired expertise. This program might adopt hypotheses concerning expected correlates to attaining certification, including higher pay, more responsible positions within the same practice setting, higher supervisors' ratings (as measured with blind ratings regarding certified status), and enhanced client satisfaction with practice outcomes. All these hypotheses can and should be examined to provide support or counter-support to the stated purpose of the credential and, therefore, to this organization's theory of credentialing. Studies such as this are recommended in PES *Guidelines* which state:

> Evaluate performance standards in reference to the degree that empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions made as a result of implementation of the standards. The inferences and actions made are credentialing decisions; these decisions should be evaluated in the context of the adopted theory of credentialing, and the predictions and inferences that flow from that theory. (4.2.5, p. 43)

(iii) Conduct validation analyses that address the consequential basis of test use

Messick (1989) makes a very strong case that the value and social consequences of score-based actions must be evaluated as part of the validity evidence provided in support of a measurement theory. The outcomes of a selection system such as credentialing are highly value-laden, and are characterized by substantial social impact. To ignore the value implications of credentialing decisions, and the social outcomes of these decisions, risks the alienation of one's stakeholders and, therefore, support for the program. The measurement issue here is that, to the extent that a credentialing program's stakeholders require evidence that the process is rational, reasonable, fair, and scientific, the sponsoring organization is obligated to demonstrate that the credentialing decisions made are reasonable, rational, fair, and scientific, and it is neither unduly reflective of a given set of values nor biased against minority groups.

Many critics of licensure (Dawes, 1994; Friedman, 1962; Shimburg, 1991) charge that the values of members of the profession are paramount in arriving at licensure decisions, and that licensure programs are used to protect the economic and social interests of professional groups, rather than the interests of the public. The essential charge, then, is that the licensure program addresses the values of one stakeholder group (the members of a profession) to the detriment of other stakeholder groups (the public). The PES *Guidelines* summarize these recommendations:

> Analyze and interpret the value implications of licensing and certification programs. (4.3, p. 44),

and:

> The intended and unintended social consequences of making credentialing decisions should be addressed. The investigation of the consequences and side effects of credentialing decisions is a critical aspect of the validation of the use of tests to make those decisions. Of particular importance is the investigation of any adverse impact that results from the application of an organization's credentialing decision rules; any adverse impact must be examined to see whether it is the result of test invalidity, or of the valid application of the decision rules that constitute the organization's policies. (4.4, p. 45)

2.10.2.5 Minimize Construct-irrelevant Variance and Construct-underrepresentation

The PES *Guidelines* contain a number of specific guidelines aimed at ongoing procedures to eliminate or minimize the two sources of threats to construct validity, construct-irrelevant variance and construct-underrepresentation. On the one hand, construct-irrelevant variance occurs when "the test contains excess reliable variance that is irrelevant to the interpreted construct" (Messick, 1989, p. 34). Construct-underrepresentation, on the other hand, occurs when a test fails to measure enough of the primary construct being measured.

The two major sources of construct-irrelevant variance occur under conditions where tests are either too easy or too difficult in relation to the actual abilities of the candidate population. For example, tests may be too easy for "test-wise" candidates, who can spot the correct answer as a result of improper item construction. Or, a test may be too hard for reasons irrelevant to the construct being measured, perhaps because it requires verbal skills in excess of those required for minimally competent performance, or problem solving skills that exceed those required for performance that would be adequate to protect the public. In either case, to ensure that credentialing decisions are reasonable, rational, accurate, fair, and scientific, we must take steps on an ongoing basis to eliminate or minimize these two sources of error.

In general, procedures must be followed that will ensure that test items have been thoroughly edited to eliminate faulty item construction, inappropriate reading levels, biased language, or language that discriminates against minority groups. In addition, examinations must be printed in print large enough so that candidate visual acuity is not a factor in test scores. Test instructions must be clear and simple, so that verbal intelligence or problem solving skills greater than those required for entry-level competence are not influencing test scores. Tests must be administered under standardized conditions so that certain candidates are neither advantaged or disadvantaged as a result of administration conditions.

Maintaining the security of examination materials is essential to eliminating construct-irrelevant variance: a candidate who obtains a high score because they have access to examination content is obtaining a score that is not relevant to their abilities or skills. Examination instructions and eligibility information should be thorough, accessible, and written at the same reading level required for the examination, in order to prevent score misinterpretation as a result of irrelevant factors.

In conclusion, the score reports generated must not create confusion among users of the scores or the candidates who receive them. Confusion in score interpretation can lead to the establishment of performance standards that are not related to the genuine requirements for minimal competency or to the performance standard established by the credentialing organization. Thus, a construct is interpreted correctly only to the extent that the scores for that construct lead to desired outcomes. Mistakes in scoring or in score interpretation can and do cause errors that affect false-positive and false-negative credentialing decisions, and hence lead to score-based actions that are not in accord with an organization's theory of credentialing.

2.10.2.6 Conduct an Ongoing Program of Research, Evaluation, and Policy Appraisal

Credentialing programs take place through the policies of the organization. In this chapter, we have recommended that credentialing policy be based on a theory of credentialing that results in a statement of credentialing purpose, or credentialing mission. The initial policies of a credentialing organization, or the policies that exist at any point in time for that organization, will reflect the theory of credentialing either implicitly or explicitly. The final section of the PES *Guidelines* recommends that credentialing organizations conduct an ongoing program of research, evaluation, and policy appraisal; that they obtain feedback obtained through such studies for use in reevaluating the credentialing theory and how that theory is realized in credentialing policy; and that they establish ongoing dialogue with stakeholders to share the results of these studies and to obtain their appraisals of the organization's efforts to meet their needs and expectations.

A program of constant evaluation and policy appraisal demonstrates the credentialing organization's commitment to providing a credentialing program that is rational, reasonable, fair, accurate, and scientific. The PES *Guidelines* suggest that this program address such issues as: (i) currentness of the occupational practice analysis; (ii) appropriateness of assessment tools; (iii) implications of already-conducted validity studies, as well as plans for additional studies to test credentialing hypotheses; (iv) maintenance of the constant search for sources of construct-irrelevant variance or construct-underrepresentation; and (v) evaluation and reevaluation of the organization's credentialing policy in light of the finding of this research program.

2.10.3 CONCLUSION

Credentialing, a form of measurement characterized by high-stakes consequences, is situated in the political context of the USA. As such, it embodies the values of our culture and has significant social impact on majority and minority groups. Credentialing has many stakeholders, and each stakeholder group's values and interests exert influence upon the process. The role of standards and guidelines is to frame the dialogue among these groups, such that evidence in support of each perspective is presented clearly and effectively.

Standards and guidelines for credentialing cannot attempt to resolve all issues; rather, they provide the tools whereby evidence can be collected and evaluated, and decisions made, in a rational, reasonable, fair, and scientific manner. Measurement professionals need to bear in mind this role as facilitator of objective, unbiased discussion, evaluation, and appraisal. As Messick (1989) concludes:

> When confronted with media controversy, judicial review, or legal dispute, test specialists are often thrown off balance and frequently filter evidence through their own social values without weighting alternative value perspectives. The impact of social values on the interpretation of evidence is difficult to avoid in any event. However, the effect is less variable or idiosyncratic if the evidence is interpreted in the context not only of personal values and the values of others, but of technical standards to guide professional judgment as well. Because validity is the one essential justification for test interpretation and use, professional judgments about testing should be guided by validity principles in particular and by the broader testing Standards [AERA et al., 1985] in general. However, to be credible, these validity principles and testing standards must be applied sensibly in practice and communicated sensibly to nonprofessional participants in the public debate [Kleinman & Farley, 1985]. This strategy is predicated on the assumption that "scientific argument and political argument differ in degree rather than kind, science having a longer time horizon, more homogeneous participants, and more appeal to formal reasoning [Cronbach, 1988, pp. 6–7]." (p. 91)

2.10.4 REFERENCES

American Association of State Psychology Boards. *Entry requirements for Professional Practice in Psychology: A guide for student and faculty (undated, obtained in 1990)*. Association of State and Provincial Psychology Boards, Montgomery, AL.

American Education Research Association, American Psychological Association, and National Council on Measurement in Education (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

American Psychological Association (1988). *Guidelines for computer-based tests and interpretations*. Washington, DC: Author.

American Psychological Association, Division of Industrial–Organizational Psychology (1987). *Principles for the validation and use of personnel selection procedures* (3rd ed.). College Park, MD: Author.

Americans with Disabilities Act (1990). 42 USCA 12101 *et seq*. (Also: 56 Fed. Reg. 35734, 56 Fed. Reg. 35694, 56 Fed. Reg. 35592, 56 Fed. Reg. 36731).

Brinegar, P. (Ed.) (1990). *Occupational and professional regulations in the states: A comprehensive compilation*. Lexington, KY: The National Clearinghouse on Licensure, Enforcement, and Regulation.

Council on Licensure, Enforcement and Regulations (1993). *Development, administration, scoring and reporting of credentialing examinations: Recommendations for board members*. Lexington, KY: Author.

Council on Licensure, Enforcement, and Regulation and the National Organization for Competency Assurance. (1993). *Principles of fairness: An examining guide for credentialing boards*. Lexington, KY: Author.

Cronbach, L. J. (1980). Validity on parole: How can we go straight? New directions for testing and measurement: Measuring achievement over a decade. *Proceedings of the 1979 ETS Invitational Conference*, 99–108. San Francisco: Jossey-Bass.

Cronbach, L. J. (1988). Five perspectives on validation argument. In H. H. Wainer and H. Braun (Eds.), *Test validity* (pp. 3–17). Hillsdale, NJ: Erlbaum.

Dawes, R. (1994). *House of cards: Psychology and psychotherapy built on myth*. New York: Free Press.

Educational Testing Service (1987). *ETS standards for quality and fairness*. Princeton, NJ: Author.

Equal Employment Opportunity Commission, Civil Service Commission, US Department of Labor and US Department of Justice (1978, August 25). *Uniform guidelines on employee selection procedures*. Fed. Reg., 43, (166), 38290–38315.

Friedman, M. (1962). *Capitalism and freedom*. Chicago: University of Chicago Press.

Guion, R. M. (1976). Recruiting, selection, and job placement. In M. D. Dunnene (Ed.), *Handbook of industrial and organization psychology* (pp. 777–828). Chicago: Rand McNally.

Hopkins, B. R. (1992). *The law of tax-exempt organizations* (6th ed.). New York: Wiley.

Jones, J. P., & Schoon, C. G. (1993, April). *An examination of the predictive value of two objective measures in veterinary medicine*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Jones, J. P., Schoon, C. G., & Lovler, R. L. (1994, April). *An examination of the underlying constructs measured by three credentialing examinations in veterinary medicine*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Kleiman, L. S., & Farley, R. H. (1985). The implications of professional and legal guidelines for court decisions involving criterion-related validity: A review and analysis. *Personnel Psychology, 38,* 803–833.

Knapp, I. E., & Knapp, L. G. (1995). Practice Analysis: Building the foundation for validity. In J. C. Imapara (Ed.), *Licensure testing: Purposes, procedures, and practices* (pp. 93–116). Lincoln, NE: University of Nebraska, Buros Institute of Mental Measurements.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28,* 563–575.

McCready, L. A. (1982, Fall). Emerging healthcare occupations: The system under siege. *Health Care Management Reviews,* 71–76.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: American Council on Education & Macmillan.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50,* 741–749.

National Commission for Certifying Agencies (1991). *NCCA Guidelines for certification approval*. Washington, DC: Author.

Pew Report. Report of the Taskforce on Healthcare Workforce Regulation (1995). *Reforming health care workforce regulations: Policy considerations for the 21st century*. San Francisco: Pew Health Professions Commission.

Popham, W. J. (1978). *Criterion-referenced measurement*. Englewood Cliffs, NJ: Prentice-Hall.

Professional Examination Service (1995). *Guidelines for the development, use, and evaluation of licensure and certification programs*. New York: Author.

Rosenfeld, M., Thornton, R. F., & Skurnick, L. (1986). *Analysis of the professional functions of teachers: Relationships between job functions and the NTE Core Battery*. Princeton, NJ: Educational Testing Service.

Schmidt, K. L. (1995). What is Licensure? In J. C. Imapara (Ed.), *Licensure testing: Purposes, procedures, and practices* (pp. 1–34). Lincoln, NE: University of Nebraska, Buros Institute of Mental Measurements.

Schoon, C. G. (1985). Methods for defining and assessing professional competence [Special Issue: The Measure of Competence]. *Professional Practice of Psychology, 6*(1), 144–155.

Schoon, C. G., & Smith, I. L. (1996a). In A. H. Browning, A. C. Bugbee, & M. A. Mullins (Eds.), *Certification: A NOCA handbook* (pp. 149–190). Washington, DC: National Organization for Competency Assurance.

Schoon, C. G., & Smith, I. L. (1996b). Guidelines for credentialing programs. *Professional Education Research Quarterly, 17* (4), 6–9.

Shimberg, B. (1991). *Regulation in the public interest: Myth or reality?* (Resource Briefs, 91–1). Lexington, KY: The Council on Licensure, Enforcement, and Regulation.

Shimberg, B., Esser, B. F., & Kruger, D. H. (1973). *Occupational licensing: Practices and policies*. Washington, DC: Public Affairs Press.

Shyrock R. H. (1967). *Medical Licensing in America, 1650–1965*. Baltimore: The Johns Hopkins University Press.

US Department of Health, Education and Welfare (1977). *Credentialing health manpower*. Washington, DC: Author.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.11

Defining Constructs Underlying the Practice of Psychology in the United States and Canada

I. LEON SMITH and SANDRA GREENBERG

Professional Examination Service, New York, NY, USA

2.11.1 INTRODUCTION 122

2.11.2 METHOD 123

2.11.2.1 Committee Structure 123
2.11.2.2 Structure and Development of the Draft Delineations of the Practice of Psychology 123
2.11.2.3 Refining the Delineations of the Practice of Psychology 124
2.11.2.4 Development of a Survey of Practice 124
2.11.2.5 Sampling Plan 125
2.11.2.6 Conduct of the Survey 125

2.11.3 SUMMARY OF RESULTS 126

2.11.3.1 Survey Response Rate 126
2.11.3.2 Professional Background and Demographic Descriptions 126
2.11.3.3 Results Related to the Delineations 126
2.11.3.3.1 Process-based delineation 127
2.11.3.3.2 Content-based delineation 127
2.11.3.4 Qualitative Results 128
2.11.3.5 Process- and Content-based Profiles of Practice 129
2.11.3.6 Development of Test Specifications 129

2.11.4 CONCLUSIONS 129

2.11.5 APPENDIX 130

2.11.5.1 Roles and Definitions 130
2.11.5.1.1 Responsibilities related to direct service 130
2.11.5.1.2 Responsibilities related to outreach and consultation 130
2.11.5.1.3 Responsibilities related to academic preparation and professional development 131
2.11.5.1.4 Responsibilities related to research and evaluation 131
2.11.5.2 Content Areas and Definitions 131
2.11.5.2.1 Knowledge statements related to biological bases of behavior 132
2.11.5.2.2 Knowledge statements related to cognitive-affective bases of behavior 132
2.11.5.2.3 Knowledge statements related to social and multicultural bases of behavior 132
2.11.5.2.4 Knowledge statements related to growth and lifespan development 133
2.11.5.2.5 Knowledge statements related to assessment and diagnosis 133
2.11.5.2.6 Knowledge statements related to treatment/intervention 134
2.11.5.2.7 Knowledge statements related to research and evaluation 134
2.11.5.2.8 Knowledge statements related to ethical/legal/professional issues 134

2.11.6 REFERENCES 134

2.11.1 INTRODUCTION

The process of establishing content validity, as described in Chapter 11 of the 1985 revision of the *Standards for Educational and Psychological Testing* (the *Standards*) published by the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME), emphasizes the need to conduct a practice analysis to ensure that the knowledge, skills, or abilities assessed in credentialing initiatives are limited to those required for competent performance and that they serve a public protection function. According to contemporary theorists like Messick (1989,1995), content validity is one aspect of a comprehensive theory of construct validity. From Messick's perspective, the meaning and interpretation of test scores is established on the basis of many different types of empirical investigations, of which a practice analysis is but one example. This view is consistent with the thinking of Schoon and Smith (1996), who suggest that a practice analysis is the second of six stages in the development, use, and evaluation of licensure examination programs, and immediately follows the conduct of a strategic credentialing analysis which is addressed by Schoon in Chapter 10, this volume.

Although there is no standard way to conduct a practice analysis, common strategies used to define the critical performance areas and the underlying knowledge, skills, or abilities include the use of a committee of subject-matter experts, the compilation of worker logs, observations of practice, and interviews with practitioners (Knapp & Knapp, 1995). In general, it is best to collect practice data from more than one source, as a way of confirming or supplementing the data obtained from other sources. The resulting definition of practice can then be validated by collecting ratings from a representative sample of practitioners. The validation process generates data on the criticality, importance, and frequency of the elements in the practice definition which can be used to evaluate the completeness and appropriateness of the performance areas and the associated knowledge, skills, or abilities. Analysis of the validation data also leads to the development of test specifications supporting examination development activities and to the preparation of instructions for developing the test questions. The outcomes of a practice analysis, then, become the primary basis by which a sponsor of a national licensure program such as the Association of State and Provincial Psychology Boards (ASPPB), formerly the American Association of State Psychology Boards, establishes and defends the content validity of its credentialing initiative.

ASPPB maintains a research program on the Examination for Professional Practice in Psychology (EPPP) consistent with the procedures described above and with the content validity emphasis articulated in the *Standards*. For more information on ASPPB's research program, See Rosen (1991). For a brief history of the origins of the psychology licensing program, the reader is referred to Chapter 10, this volume. Finally, information on EPPP test performance, candidate background information, pass/fail statistics, and the impact of test preparation services on candidate performance is provided by Rosen in Chapter 12, this volume.

ASPPB's research includes three steps: (i) the conduct of a study of professional practice, (ii) the development of revised test specifications based on the study of practice, and (iii) the independent verification that the items and/or forms of the EPPP reflect the test specifications in a representative and fair manner and are consistent with the generic philosophy underlying the EPPP.

Since the mid-1970s, ASPPB has conducted two sets of investigations, each associated with these three steps. ASPPB contracted with Professional Examination Service (PES) to initiate the first set of studies (Richman, 1982). PES conducted an investigation aimed at clarifying the content most appropriate for the EPPP. On the basis of the study, ASPPB developed a new EPPP content outline. All items in the ASPPB item bank were recorded to incorporate the clarified information on the role and knowledge statements. An additional content validation study was conducted under contract to PES (Smith, 1984) to check on the clarity of the new content outline, the quality of the items, and the correspondence between the content categories in the test outline and the items.

ASPPB contracted with the Educational Testing Service to initiate the second set of studies (Rosenfield, Shimberg, & Thornton, 1983). Using existing resources and expert judgments, Rosenfeld et al. created an inventory consisting of 59 responsibilities, 61 procedures, techniques, and resources, and 49 knowledge areas, including ethical/legal considerations and defining the work of professional psychologists. Licensed psychologists in the USA and Canada rated the inventory in terms of time spent, importance, level of application, and level of judgment exercised.

Subsequently, revised test specifications were developed under contract to PES (Rosen & Mirone, 1986). The procedures called for an integration of the results of the job analysis conducted by Rosenfeld et al. (1983) with the findings of the earlier study completed by Richman (1982). All items in the ASPPB item bank were reclassified on the basis of the categories in the new test specifications. ASPPB contracted with PES (Hambleton & Smith, 1988) to conduct a content validity study to provide an independent check that the first two forms of the EPPP based on the new test specifications reflected these specifications, and that the items were free of any unintended content bias with regard to ethnicity, gender, and age. Finally, ASPPB funded PES (Smith & Greenberg, 1991) to address selected recommendations made as part of an audit of the EPPP examination program conducted by the State of California (Werner, 1989). The knowledge, skill, and application statements in the test specifications were augmented with specific topics or elements of practice, the interpretation of the responsibilities was clarified, and guidelines were prepared to facilitate development and classification of items.

The practice analysis study reported herein was conducted by PES under contract with ASPPB and initiated the third comprehensive set of investigations to examine what licensed psychologists in the USA and Canada do and know in order to ensure the continued content validity of the EPPP (Greenberg, Smith, & Muenzen, 1996).

2.11.2 METHOD

2.11.2.1 Committee Structure

The practice analysis study was conducted in collaboration with a five-member Practice Analysis Advisory Committee (PAAC) and a 12-member task force of subject-matter experts appointed by the PAAC on the basis of nominations received from ASPPB's Board of Directors, the Examination Development Committee, and the PAAC. The members of the PAAC represented key decision-makers in ASPPB; in part, they were responsible for guiding aspects of the licensure program, including examination development and validation efforts. The members of the task force represented psychologists in different major areas of practice, including new and emerging specialties.

2.11.2.2 Structure and Development of the Draft Delineations of the Practice of Psychology

Process- and content-based approaches were implemented to study the practice of licensed psychologists in the USA and Canada. Process- and content-based approaches to practice analysis are appropriate for professions such as psychology in which the primary professional behaviors are cognitive in nature (Schoon, 1985).

A process-based approach was used because it provides a structure for describing contemporary practice—that is, for identifying what psychologists do—and because it facilitates the development of examination items in a practice-related framework. The process-based approach comprises the delineation of roles and associated responsibilities performed by psychologists.

Roles represent constructs described by major categories of activities. For example, the role of Direct Service was identified and defined as the provision and/or administration of psychological services to clients, patients, and/or organizations in the areas of problem definition, need assessment, and diagnosis; and the design, implementation, and evaluation of interventions.

Responsibilities represent the specific activities that psychologists perform within each role. For example, to observe, interview, and gather information from the patient/client/organization and related sources (e.g., relevant others, written records, referral source) in order to identify the problems/needs and their contexts, is one responsibility performed in connection with the Direct Service role.

A content-based approach was used because it provides a structure for identifying a comprehensive listing of the critical knowledge that psychologists need in order to perform the responsibilities identified as part of professional practice. The content-based approach provides a user-friendly template for giving feedback to candidates and communicating summary results of EPPP performance to universities, professional schools, and training programs. The content-based approach comprises the delineation of content areas and associated knowledge statements.

Content areas represent constructs described by categories of knowledge used by psychologists in the performance of responsibilities. For example, Biological Bases of Behavior a content area encompassing knowledge of neuroscience, the physiological bases of behavior and illness, and psychopharmacology.

Knowledge statements represent organized bodies of information associated with specific content areas and are necessary in the performance of responsibilities. For example, knowledge of basic neuroscience (e.g., neuroanatomy, neurophysiology, neurochemistry) and clinical neuroscience (e.g., brain–behavior relationships, neurological syndromes and their contribution to cognitive and emotional status and behavior), is a knowledge statement that describes information subsumed in the Biological Bases of Behavior content area.

2.11.2.3 Refining the Delineations of the Practice of Psychology

The following data methods were implemented to supplement and evaluate the structure and content of the draft delineations initially developed by the task force: (i) a sample of 25 licensed psychologists completed a practice log designed to elicit a list of professional responsibilities performed by psychologists; (ii) a sample of 40 licensed psychologists participated in a critical incidents interview in which participants were asked to reflect upon incidents in which they felt particularly effective and particularly ineffective in their practice and to identify the knowledge which made a difference to the outcome; (iii) job descriptions and supervisor's rating forms were collected from members of the PAAC, the task force, and the participants in both the practice log recording and the critical incidents interviews; and (iv) a sample of 29 licensed psychologists completed an independent review of the draft delineations.

The task force and the PAAC used the information in preparing the final delineations. Table 1 presents the roles and the number of associated responsibilities in the process-based delineation, and the content areas and the number of associated knowledge statements in the content-based delineation. A complete copy of the delineations is given in the appendix to this chapter (see Section 2.11.5).

2.11.2.4 Development of a Survey of Practice

In order to validate the process- and content-based delineations, a survey of the practice of psychology was developed for dissemination to licensed psychologists in the USA and Canada. Rating scales were designed to collect data on the roles and responsibilities performed by licensed psychologists and the content areas and knowledge required for practice. A two-part pilot test of the survey was conducted to identify question defects such as ambiguous items and cognitive difficulties in forming answers to questions. First, a sample of 38 licensed psychologists completed the survey via a mail review and evaluated each section to identify any unclear elements. Then, a second sample of four licensed psychologists completed a mail review of the survey and were debriefed in an in-depth telephone interview regarding the cognitive distinctions they made among the rating scales and the scale points associated with the rating scales (Bolton, 1993).

The results of the two-part pilot test were used to refine the process- and content-based delineations, the instructions, and the rating scales. The final survey document, entitled the "Survey of the Practice of Psychology" (the Survey), included six sections. Section One used a set of three rating scales to elicit information about the roles performed in the practice of psychology. Survey respondents were asked to rate the roles in regard to three questions:

(i) How important was the role to *your practice* as a psychologist during the past year?

(ii) What percentage of *your work time* was devoted to each role during the past year?

(iii) How critical is performing the role *protecting the patient/client/public* from harm?

Section Two used a similar set of rating scales to elicit information about the responsibilities in the practice of psychology associated with each role. Survey respondents were asked to rate the responsibilities in regard to three questions:

(i) How important was the performance of the responsibility to *your practice* as a psychologist during the past year?

(ii) How *frequently* have *you* performed this responsibility during the past year?

(iii) How critical is performing the responsibility to *protecting the patient/client/public* from harm?

Section Three elicited information about the content areas—or categories of knowledge—that psychologists use in practice. Survey respondents were asked to rate the content areas in regard to three questions:

(i) How important was the content area to *your practice* as a psychologist during the past year?

(ii) How *frequently* have *you* called upon knowledge from the content area in your practice during the past year?

(iii) How critical is the content area to *protecting the patient/client/public* from harm?

Section Four elicited information about the knowledge that is associated with the content areas and is needed to perform the responsibilities. Survey respondents were asked to rate the knowledge statements in regard to three questions:

(i) What level best represents *your use* of this knowledge in your practice?

(ii) At what point should the knowledge be *acquired* by psychologists?

(iii) How critical is possessing the knowledge to *protecting the patient/client/public* from harm?

Section Five elicited information about the demographic and professional background of respondents with respect to their work in

Table 1 Process- and content-based delineations of the practice of psychology.

Role	*Number of responsibilities*
Direct Service	11
Outreach and Consultation	10
Academic Preparation and Professional Development	8
Research and Evaluation	10

Content area	*Number of knowledge statements*
Biological Bases of Behavior	6
Cognitive-Affective Bases of Behavior	6
Social and Multicultural Bases of Behavior	11
Growth and Lifespan Development	8
Assessment and Diagnosis	14
Treatment/Intervention	11
Research and Evaluation	5
Ethical/Legal/Professional Issues	5

psychology, such as primary and secondary employment setting, practice area, theoretical orientation, specialization or respecialization, and specialty certification; and with respect to demographic variables, such as education, sex, disability, and race or ethnicity.

Section Six solicited qualitative comments wherein respondents were requested to describe any professional responsibilities they perform or knowledge they call upon that was not included in the survey. Most importantly, respondents were asked to comment upon what they believe to be the long-range changes occurring in the practice of psychology.

2.11.2.5 Sampling Plan

The sampling plan for the dissemination of the Survey was designed to (i) generate a sample of 7500 licensed psychologists from those jurisdictions in the USA and Canada that employed the EPPP as an element in the credentialing process, and (ii) ensure the representation of licensed psychologists at or near the entry level of the profession. A minimum of 60 licensed psychologists were sampled from each jurisdiction using the EPPP, regardless of the size of the jurisdiction, and recently licensed psychologists, operationally defined as having been licensed in 1989 or later, were oversampled in every jurisdiction.

The sampling plan included three different data sources: (i) ASPPB-member boards, (ii) the APA Membership Database, and (iii) the Canadian Register of Health Service Providers in Psychology (Canadian Register). In general, the majority of the sample of recent licensees was drawn directly from the ASPPB-member boards using a systematic random sampling technique, and the majority of the sample of less-recent licensees was drawn from the remaining two sources using a systematic random sampling technique.

2.11.2.6 Conduct of the Survey

The conduct of the Survey was designed to yield a high rate of return. Accordingly, the plan involved three carefully timed sequential mailings to the survey recipients as well as one telephone follow-up to a sample of nonrespondents (Dillman, 1978).

(i) Potential participants received a letter describing the nature and scope of the project and inviting their participation in the data collection efforts. Recipients who declined to participate or who were no longer active in the profession were requested to call collect so as to be eliminated from the sample. Potential survey recipients in the three Canadian provinces where the French-language version of the EPPP was offered received an English/French bilingual version of the letter and a postage-paid return postcard for requesting a French-language version of the survey.

(ii) Approximately two weeks later, all remaining potential members of the sample received an English- or French-language survey packet, as appropriate. The survey packet contained: a covering letter reviewing the purpose and importance of the data collection, and assuring the recipients about the confidentiality of the responses; a survey and a postage-paid return envelope; and a postage-paid tracking postcard to be returned, indicating return or nonreturn of the completed survey. As a follow-up, each member of the sample received an English- or French-language reminder/thank-you postcard, as appropriate.

(iii) Subsequently, each member of the sample *not* returning a tracking postcard received a second and then a third survey packet. The packets were identical to the first except that the covering letter identified the recipient as not previously returning a completed survey. In addition, the letter informed recipients of ways to obtain the results of the study. The letter in the third packet urged the recipients to respond so that their own practice might be reflected in ASPPB's picture of contemporary practice. Follow-up reminder/thank-you postcards were sent to all members of the sample who received the second and third survey packets.

(iv) Approximately four weeks after the final mailing, a telephone/mail survey of 250 nonresponders was conducted to estimate the potential impact of nonresponse bias.

2.11.3 SUMMARY OF RESULTS

2.11.3.1 Survey Response Rate

The return rate for completed Surveys from eligible participants was 60%, which is considered excellent for surveys of this type (Knapp & Knapp, 1995). Consistent with the sampling plan, every ASPPB-member jurisdiction was represented in the returns, and recently licensed psychologists were overrepresented in proportion to their representation in the population.

2.11.3.2 Professional Background and Demographic Descriptions

(i) More than two-thirds of the US and Canadian respondents indicated that their primary employment setting fell within the general category of human-service settings (including individual independent practice and group psychological practice), whereas about 20% indicated that their primary employment setting was an educational institution/school system.

(ii) US respondents were most likely to have earned a doctoral degree, including the doctorate of psychology degree, as their highest degree in psychology. Canadian respondents were equally likely to have earned either a master's degree or a doctoral degree as their highest degree in psychology.

(iii) Recently licensed US respondents were more likely to have earned a degree from a free-standing school of professional psychology than were less recently licensed respondents. Few Canadian respondents earned degrees from free-standing schools of professional psychology.

(iv) More than one half of the US and Canadian respondents indicated clinical psychology as their major area of training; counseling was reported as the major area of training by about 20% of the US respondents and 15% of the Canadian respondents.

(v) In both the USA and Canada, more respondents indicated clinical child psychology and clinical neuropsychology as their current major area of practice than had trained in those areas.

(vi) In both the USA and Canada, about one-half of respondents described their primary theoretical orientation as cognitive/behavioral or behavioral. In the USA, about one-quarter of the respondents described their primary theoretical orientation as psychodynamic, and no more than 8% identified any one other primary theoretical orientation. In Canada, approximately equal numbers described their primary theoretical orientation as psychodynamic and existential/humanistic (both 18%), and no more than 6% identified any one other primary theoretical orientation.

(vii) About one-quarter of the US respondents and approximately 15% of the Canadian respondents indicated that they had participated in a postdoctoral specialization or respecialization program.

(viii) More than one-half of US and Canadian recently licensed respondents were female, compared to about 40% of the less recently licensed US and Canadian respondents.

(ix) About 93% of the US respondents and 89% of the Canadian respondents indicated they were Caucasian/White. American-Indian or Alaskan Native, Asian or Pacific Islander, Black or African-American, and Hispanic/Latino psychologists were represented among the respondents.

(x) Selected comparisons of professional background and demographic variables indicated that the current sample was representative of the populations of licensed psychologists included in the APA Membership Database and the Canadian Register.

(xi) The results of the respondent/nonrespondent survey revealed few differences in key background and demographic variables.

2.11.3.3 Results Related to the Delineations

Preliminary statistical analyses of the Importance, Time/Frequency, and Criticality ratings indicated that there were few differences in the results of the respondent/nonrespondent survey, or in the ratings of recently vs. less recently licensed respondents and US versus Canadian respondents. Accordingly, the ratings of all the respondents on the roles and responsibilities, and on the content areas and knowledge statements were combined for final data analysis in connection with the development of profiles of practice and revised test specifications for the EPPP. Readers are encouraged to review the appendix to this chapter (see Section 2.11.5) for a complete description of the roles and responsibilities, and the content areas and knowledge statements. These elements comprise the components of the test specifications.

2.11.3.3.1 Process-based delineation

(i) Roles

The Direct Service role was rated most important to the practice of the respondents and most critical to the protection of the patient/client/public from harm. Respondents spent about 70% of their time in that role.

(a) The Direct Service role was rated as moderately to very important to the practice of the respondents; in fact, nearly 90% of the ratings were at scale point three, indicating that the role was very important to the practice of the respondents, and the associated standard deviation was the smallest of any role and in connection with any rating scale. The remaining roles—Outreach and Consultation, Academic/Professional Development, Research and Evaluation—were rated as minimally to moderately important to their practice.

(b) The mean Time ratings indicated that respondents spent two-thirds of their time in the Direct Service role, and one-third of their time in all other roles. However, the magnitude of the standard deviations associated with the mean Time ratings indicated a great deal of variability in the practice patterns of the respondents.

(c) All four delineated roles were rated as moderately to highly critical to the protection of the patient/client/public from harm.

(d) In general, the Importance and Time ratings on the roles varied more than the Criticality ratings; respondents consistently rated the roles as moderately to highly critical to the protection of the patient/client/public, regardless of their own professional practice patterns.

(e) The ratings of the roles by respondents representing various major areas of practice demonstrated some similarities and differences. For example, in terms of the Importance rating scale, the ratings were generally similar in regard to the Direct Service role, but somewhat dissimilar in regard to the Academic Preparation and Professional Development role and the Research and Evaluation role.

(ii) Responsibilities

The Importance, Frequency, and Criticality ratings of the responsibilities were generally consistent with the Importance, Time, and Criticality ratings of the associated roles; thus, the respondents rated the responsibilities associated with the Direct Service role as most important to their practice, most frequently performed, and most critical to the protection of the patient/client/public from harm. In general, the Importance and Frequency ratings on the responsibilities varied more than the Criticality ratings.

(a) Respondents rated virtually all the responsibilities as moderately to highly critical to the protection of the patient/client/public from harm, regardless of their own professional practice patterns.

(b) One responsibility associated with the Direct Service role—to provide direct service in a manner consistent with professional and ethical standards and guidelines, and state/provincial and national laws and regulations—was rated highest of any responsibility on the Importance, Frequency, and Criticality rating scales. Within the Outreach and Consultation role, a similar responsibility—to provide education and consultation in a manner consistent with current professional and ethical standards and guidelines, and state/provincial and national laws and regulations—was rated highest on the same three rating scales.

(c) The ratings of the responsibilities by respondents representing various major practice areas demonstrated similarities and differences. Respondents in various major areas of practice displayed a greater degree of consensus in their Criticality ratings than in their Importance and Frequency ratings.

2.11.3.3.2 Content-based delineation

(i) Content areas

The Treatment/Intervention and Ethical/Legal/Professional Issues content areas were rated most important to the practice of the respondents and most critical to the protection of the patient/client/public from harm; the knowledge base associated with the Treatment/Intervention content area was called upon most frequently by the respondents.

(a) The ratings of the content areas indicated that all areas—except Research and

Evaluation—were moderately to very important to the practice of the respondents and were called upon moderately to very frequently.

(b) All content areas—including Research and Evaluation—were rated as moderately to highly critical to the protection of the patient/client/public from harm.

(c) In general, the Criticality ratings of the respondents were somewhat higher than the Importance ratings.

(d) The ratings of the content areas by respondents in different major areas of practice were generally similar. The respondents in different major areas of practice displayed a great degree of similarity in what content areas they considered critical to the protection of the patient/client/public from harm, even though they varied considerably in their ratings of what content areas are important to and frequently performed in their own practice.

(ii) Knowledge statements

About 85% of the respondents used 51 of the 66 knowledge statements either at the recognition/recall or the apply/interpret/integrate level, and about the same percentage of respondents rated 61 of the 66 knowledge statements as necessary to be acquired at some point in time—either primarily before or after licensure. Finally, the Criticality ratings indicated that 43 of the 66 knowledge statements were rated moderately to highly critical to the protection of the patient/client/public from harm. Other key findings were as follows:

(a) Of the 15 knowledge statements rated by a majority of the respondents as not used or not supported for acquisition, 12 were in two content areas—Assessment and Diagnosis, and Treatment/Intervention.

(b) With few exceptions, the ratings of the knowledge statements by respondents representing various major areas of practice were similar. However, respondents in various major areas of practice were least similar in their ratings of knowledge statements associated with the two content areas, Assessment and Diagnosis and Treatment/Intervention.

2.11.3.4 Qualitative Results

About two-thirds of the respondents identified long-range changes occurring in the practice of the profession, especially those that may affect the knowledge required of licensed psychologists in the future. The following summary characterizes the unique and the common areas of change identified by the US and Canadian respondents:

(i) More than one-quarter of the US respondents commented on the impact of the shift from independent practice to managed care. Respondents described their concerns regarding the constraints of managed care, for example, the focus on brief, solution-oriented therapy; the ethical dilemmas posed around issues such as confidentiality, privacy, the termination of therapy, and the move to outpatient rather than inpatient treatment; the reliance on computer-based assessment and diagnosis; the requirements for quality assurance and outcomes-based assessment; and the focus on the business aspects of practice, including the entrepreneurial management and administration of practice.

(ii) Nearly one-quarter of the US respondents and somewhat fewer of the Canadian respondents commented on the increased emphasis on the biological bases of behavior in clinical practice.

(iii) More than 10% of the US respondents and somewhat fewer of the Canadian respondents commented on the redefinition of the practitioner as a primary healthcare provider or as a case manager supervising other allied health practitioners responsible for direct services.

(iv) Nearly 10% of the US and Canadian respondents commented on one or more of the following shifts toward (a) a more multiculturally diverse patient/client population, (b) a larger geriatric patient/client population, and (c) a more severely disturbed (and violent) society. Each of these trends is associated with a specific knowledge base.

(v) Approximately 10% of the Canadian respondents described the shift from employment within the government-sponsored healthcare system to independent practice and/or employee-assistance programs, in contrast to the move in the US from independent practice to managed care.

Both US and Canadian respondents identified the following types of knowledge and skills as needed in the future:

(i) Knowledge regarding the biological bases of behavior, brief treatments and/or solution-oriented therapeutic techniques, hypnotherapy, forensic psychology, employee-assistance programs, quality assurance procedures, outcomes-based assessment, and specialties in psychology.

(ii) Professional skills associated with consulting and/or working in interdisciplinary collaborative teams, procedures to integrate aspects of spirituality into therapy; sensitivity skills to handle the increasingly diverse patient/client population, and self-assessment skills to recognize one's strengths and limits.

(iii) Generic skills, such as computer skills, for managerial and professional aspects of practice, business skills for developing and administering a practice, and marketing skills for surviving in practice.

2.11.3.5 Process- and Content-based Profiles of Practice

Preliminary statistical analyses revealed that among the Importance, Time/Frequency, and Criticality scales, the latter two ratings contributed maximum independent information to a description or profile of practice. Accordingly, a process-based profile was developed by combining the Criticality and Time/Frequency ratings associated with the roles and responsibilities, and a content-based profile was developed by combining the Criticality and Frequency ratings associated with the content areas, as well as the Criticality and Acquisition ratings associated with the knowledge statements which were rank-ordered from most critical and appropriate to be acquired before licensure, to least critical and not necessary at any point in practice.

In the case of the process-based profile, the largest element related to the Direct Service role (about 70%); the remaining three roles contributed less than 30% to the overall profile of the critical activities performed by licensed psychologists. With respect to the content-based profile, seven of eight content areas each contributed between 11% and 16% to the profile, while one content area—Research and Evaluation—made a smaller contribution (6%) to the profile. The Ethical, Legal, and Professional Issues content area included three of the 10 highest-ranked knowledge statements; the Growth and Lifespan Development, Assessment and Diagnosis, and Treatment/Intervention content areas each included two of the 10 highest-ranked knowledge statements; and the Biological Bases of Behavior content area included one of the 10 highest-ranked knowledge statements.

2.11.3.6 Development of Test Specifications

A systematic review of the results related to the process- and content-based profiles of practice was conducted by a specially constituted ASPPB Test Specifications Panel. The following recommendations were prepared to ensure that the EPPP would reflect the responsibilities performed by licensed psychologists making the greatest contribution to the protection of the patient/client/public from harm; emphasize the knowledge needed by licensed psychologists that serves the public protection function of regulation; and address the issue of developing a generic examination for psychologists.

(i) The use of the content areas and knowledge statements as the primary organizing structure for the test specifications and feedback to the candidates.

(ii) The adoption of percentage weights for the eight content areas.

(iii) The use of a hierarchical weighting for each knowledge statement, indicating the degree to which related questions might be included in each version of the EPPP. In order of priority, the system incorporated the Criticality ratings and the Acquisition ratings for the knowledge statements.

(iv) The revision of eight knowledge statements, including additions, deletions, and modifications of the examples included as part of the knowledge statements; and the elimination of three knowledge statements via the incorporation of the related examples into other knowledge statements.

(v) The use of the roles and responsibilities as an organizing structure for item-writing initiatives and examination assembly.

(vi) The adoption of percentage weights for the four roles and associated responsibilities;

(vii) The use of a classification system by which: (a) test items would be identified by role, responsibility, content area, and knowledge statement; and (b) forms of the EPPP would be constructed so as to reflect the content- and process-based weights in the test specifications.

2.11.4 CONCLUSIONS

The analysis of the practice of psychology reported in this chapter provides a picture of what licensed psychologists in the USA and Canada know and do. However, it is apparent from the descriptions of the short- and long-range changes occurring in the practice of the profession that frequent reviews and updates are needed to ensure that the picture of practice—and the content of the EPPP—continue to reflect contemporary practice. Rapid changes in the delivery of healthcare services mandated by the move to managed care in the USA and to nongovernment-sponsored employment in Canada may require an expanded knowledge base. Similarly, the changing gender base of the profession and the shifts in the client base, along with the move to increased interdisciplinary collaboration may impact the required knowledge base.

Accordingly, future practice analysis initiatives may require: (i) the periodic, but frequent monitoring of specific settings and/or specialties undergoing rapid changes in practice, (ii) the study of professional relationships among intra- and interdisciplinary team members, and (iii) qualitative and quantitative outcome studies of practice as well as investigations of the perceptions of employers and the public regarding the demands of practice and the value and meaning of the licensure credential in psychology. The outcomes of these kinds of studies will be very useful in the continuous process of re-examining and re-evaluating the role and function of regulation in our society (Interprofessional Workgroup on Health Professions Regulation, 1996; Pew Health Professions Commission, 1995).

The analysis of practice provides a database for policy-making in regard to education, certification, and accreditation initiatives in professional psychology. The results of the analysis should be reviewed by those responsible for: college, university, and professional school curricula; training programs; continuing education, in-service education; the development and implementation of specialty certification initiatives; and the accreditation of school, internship, and residency training programs.

Finally, the results of the current analysis should be viewed in the context of practice in other nations and/or political entities. The implementation and/or emergence of the North Atlantic Free Trade Agreement, the General Agreement on Trade and Tariffs, and the European Union indicates the appropriateness of a global perspective and the need to understand the commonalities and differences in practice and in the regulation of psychology around the world that could be provided by the conduct of international studies of practice (ASPPB, 1996).

2.11.5 APPENDIX

2.11.5.1 Roles and Definitions

Direct Service. The provision and/or administration of psychological services to clients, patients, and/or organizations in the areas of problem definition, need assessment, and diagnosis; and the design, implementation, and evaluation of interventions.

Outreach and Consultation. The preparation, presentation, and coordination of educational programs, and/or the dissemination of information or the provision of expertise to a variety of audiences.

Academic Preparation and Professional Development. The development, implementation, and administration of education programs for psychologists, including teaching, supervision, and curricula.

Research and Evaluation. The development and/or participation in any investigation and/or the use of results to expand or refine knowledge or to improve programs and services.

2.11.5.1.1 Responsibilities related to direct service

(i) Make and/or receive referrals.

(ii) Coordinate service delivery with other psychologists and professionals (e.g., health professionals, managed care systems, organizational personnel, schools, community groups, and other outside agencies).

(iii) Observe, interview, and gather information from patient/client/organization and related sources (e.g., relevant others, written records, referral source) to identify the problems/needs and their contexts.

(iv) Develop assessment procedures and/or instruments (e.g., behavioral analyses, structured interviews, work samples, performance tests) for the assessment of relevant characteristics of individuals, groups, jobs, organizations, educational and social institutions, and/or environments

(v) Select, administer, and score norm-referenced, standardized, or other instruments for the assessment of relevant characteristics of individuals, groups, jobs, organizations, educational and social institutions, and/or environments.

(vi) Evaluate and integrate results of information-gathering and assessment processes with scientific/professional knowledge to formulate/reformulate working hypotheses, diagnoses, and intervention recommendations.

(vii) Plan, design, and implement intervention programs (e.g., define goals and objectives, identify appropriate intervention targets and strategies).

(viii) Monitor and evaluate efficacy of interventions/programs, and modify as appropriate.

(ix) Document and/or communicate assessment results, intervention recommendations, progress, and outcomes.

(x) In administering a professional practice—design, implement, and monitor quality assurance, quality control, risk management, and/or other procedures.

(xi) Provide direct service in a manner consistent with current professional and ethical standards and guidelines, and state/provincial and national laws and regulations.

2.11.5.1.2 Responsibilities related to outreach and consultation

(i) Prepare/present/coordinate health promotion programs or workshops (e.g., smoking cessation, parenting, anger control management, informational programs on community psychological services/resources).

(ii) Prepare/present/coordinate prevention and/or early intervention programs for at-risk populations (e.g., substance abuse prevention, HIV–AIDS education, community programs for the elderly).

(iii) Prepare/present/coordinate classes, seminars, or workshops for clients, family and significant others, personnel in school systems, medical and allied healthcare personnel, human resource personnel, and/or the general public.

(iv) Provide expertise to and/or serve on local/state/provincial/federal agencies (e.g., community outreach program, jurisdictional licensing board, legal system).

(v) Provide expertise to and/or serve on local/state/provincial, national or international professional psychology organizations (e.g., APA and Canadian Psychological Association (CPA), American Psychological Society, ASPPBs, Society for Industrial and Organizational Psychology).

(vi) Disseminate knowledge of psychology and its value to the general public.

(vii) Provide consultation regarding design, methodology, statistical analysis, and/or significance of reported data and conclusions of a research or evaluation study.

(viii) Provide expertise to and/or serve on site-specific patient-care, education-related, or research-related committees (e.g., program accreditation, Institutional Review Board, Quality Assurance Committee).

(ix) Establish and maintain intra- and interdisciplinary collaborative relationships, within institutional settings and with other professionals.

(x) Provide education and consultation in a manner consistent with current professional and ethical standards and guidelines, and state/provincial and national laws and regulations.

2.11.5.1.3 Responsibilities related to academic preparation and professional development

(i) Prepare/present/coordinate classes, seminars, or workshops for undergraduates, pre- and postdoctoral students, and professional psychologists.

(ii) Develop/administer/coordinate pre- and postdoctoral practicum, internship, and fellowship programs in human services settings.

(iii) Develop/administer/coordinate undergraduate, pre- and postdoctoral, and continuing education programs in professional psychology.

(iv) Supervise pre- and postdoctoral students and professional practitioners to enhance service delivery.

(v) Supervise and advise undergraduates and pre- and postdoctoral students on research/evaluation (e.g., honors thesis, dissertation).

(vi) Provide mentoring for undergraduates, pre- and postdoctoral students, and/or professional psychologists.

(vii) Provide training in a manner consistent with current professional and ethical standards and guidelines, and state/provincial and national laws and regulations.

(viii) Participate in professional self-development and continuing education designed to enhance professional and personal knowledge and skills.

2.11.5.1.4 Responsibilities related to research and evaluation

(i) Critically review and appraise existing literature with regard to study design, methodology, method of analysis, and generalizability of results and conclusions.

(ii) Use the existing knowledge base to formulate clear research/evaluation questions or to guide intervention or program development.

(iii) Engage in research in a manner that ensures protection of human and/or animal rights, adhering to current professional and ethical standards/guidelines, and jurisdictional and national laws/regulations.

(iv) Formulate research/program evaluation hypotheses, and design appropriate methods to conduct the study.

(v) Collect and analyze data using appropriate methods of analysis (e.g., qualitative, quantitative).

(vi) Report research findings and implications according to professionally accepted standards.

(vii) Submit research findings to peer review for publication and/or presentation.

(viii) Apply research findings in practice, with awareness of strengths and limitations of application.

(ix) Prepare proposals to funding agencies.

(x) Provide expertise to and/or serve in an editorial capacity on professional journals or other refereed publications, or review proposals to funding agencies.

2.11.5.2 Content Areas and Definitions

Biological Bases of Behavior. Knowledge of neuroscience, the physiological bases of behavior and illness, and psychopharmacology.

Cognitive-Affective Bases of Behavior. Knowledge of cognitive science; theories of learning, memory, motivation, and emotion; and factors that influence an individual's cognitive performance and/or emotional experience.

Social and Multicultural Bases of Behavior. Knowledge of social cognition, social interaction processes, and organizational dynamics; theories of personality; and issues in diversity (multiethnic, multicultural, gender, sexual orientation, and disability).

Growth and Lifespan Development. Knowledge of age-appropriate child, adolescent, and adult development; atypical patterns of development; and the protective and risk factors that influence developmental outcomes for individuals.

Assessment and Diagnosis. Knowledge of psychometrics, assessment models, methods for assessment of individuals and organizations/systems, and diagnostic classification systems and issues.

Treatment/Intervention. Knowledge of individual, group, or organizational interventions for specific concerns/disorders; treatment theories; and consultation models and processes.

Research and Evaluation. Knowledge of research design, methodology, and program evaluation; statistical procedures; and criteria for accurate interpretation of research findings.

Ethical/Legal/Professional Issues. Knowledge of the ethical code, professional standards for practice, legal mandates, guidelines for ethical decision-making, and professional training and supervision.

2.11.5.2.1 Knowledge statements related to biological bases of behavior

(i) Basic neuroscience (e.g., neuroanatomy, neurophysiology, neurochemistry) and clinical neuroscience (e.g., brain–behavior relationships, neurological syndromes and their contribution to cognitive and emotional status and behavior).

(ii) Physiological correlates/determinants of behavior and affect (e.g., symptoms of common psychophysiologic reactions and syndromes, such as hyperventilation, anxiety disorders, depressive disorders, stress reactions, headaches, irritable bowel syndrome).

(iii) Biological bases of the behavior and affect associated with acute and chronic illness (e.g., post-stroke depression, diabetes, AIDS, asthma, chemotherapy, fibromyalgia, hypoglycemia, schizophrenia). Includes knowledge of psychoneuroimmunology.

(iv) Basic psychopharmacology (e.g., medication effects, side-effects, and interactions). Includes knowledge of drug metabolism, drug categories (e.g., anxiolytics, antidepressants, antipsychotics, anticonvulsants), and addictive/dependency potential.

(v) Genetic transmission (e.g., the relationship of dominant and recessive genes) and its role in understanding disorders and their behavioral, emotional, and psychosocial manifestations (e.g., Duchenne's muscular dystrophy, Huntington's disease, Down's syndrome).

(vi) Relationship of stress to biological and psychological functioning, with particular reference to lifestyle and lifestyle modification (e.g., cardiac rehabilitation, smoking cessation).

2.11.5.2.2 Knowledge statements related to cognitive-affective bases of behavior

(i) Cognitive science (e.g., sensation and perception, attention, memory, language and spatial skills, intelligence, information processing, problem-solving, strategies for organizing information).

(ii) Theories and principles of learning (e.g., social learning, classical and operant conditioning, primacy/recency effects).

(iii) Theories of motivation (e.g., need/value approaches, cognitive choice approaches, self-regulation).

(iv) Theories of emotions.

(v) Reciprocal interrelationships among cognitions/beliefs, behavior, affect, temperament, and mood (e.g., healthy functioning, performance anxiety, performance enhancement, job satisfaction, depression).

(vi) Influence of psychosocial factors (e.g., sex differences, family styles and characteristics, academic/occupational success) on beliefs/cognitions and behaviors.

2.11.5.2.3 Knowledge statements related to social and multicultural bases of behavior

(i) Social cognition and perception (e.g., attribution theory and biases, information integration, confirmation bias, person perception, development of stereotypes, racism).

(ii) Social interaction (e.g., interpersonal relationships, aggression, altruism, attraction).

(iii) Group dynamics and organizational structures (e.g., school systems, gang behavior, family systems, group thinking, cultural behavior, conformity, compliance, obedience, persuasion) and social influences on individual functioning.

(iv) Environmental/ecological psychology (e.g., person–environment fit, crowding, pollution, noise).

(v) Theories of personality that describe behavior and the etiology of atypical behavior. Includes knowledge of limitations in existing theories for understanding the effect of diversity (e.g., age, ethnicity, gender).

(vi) Multicultural and multiethnic diversity (e.g., racial/ethnic minorities, gender, age, disability, sexual orientation, religious groups, between- and within-group differences).

(vii) Theories of identity development of multicultural/multiethnic groups (e.g., acculturation theories, racial/ethnic identity).

(viii) Role that race, ethnicity, gender, sexual orientation, disability, and other cultural differences play in the psychosocial, political, and economic development of individuals/ groups.

(ix) Sexual orientation issues (e.g., sexual identity, gay/lesbian/bisexual, family issues).

(x) Psychology of gender (e.g., psychology of women, psychology of men, gender identity development).

(xi) Disability and rehabilitation issues (e.g., inclusion, psychological impact of disability).

2.11.5.2.4 Knowledge statements related to growth and lifespan development

(i) Normal growth and development (cognitive, social, personality, moral, emotional, and physical) from conception through old age.

(ii) Normative or age-expected behaviors (e.g., normal age range, individual differences) and how the definition of normative behavior is influenced by culture.

(iii) Risk factors which predict an atypical developmental course (e.g., nutritional deficiencies; healthcare, including prenatal care; availability of social support; adequacy of income and housing; poverty; parental alcohol/drug abuse).

(iv) Interventions to reduce risk factors (e.g., poor healthcare, nutritional deficiencies, violence) and to increase resilience (e.g., protective factors such as caregiving, increased social support) and competence (e.g., skill building) of individuals living in at-risk environments.

(v) Life-event changes that can alter the normal course of development (e.g., injury, trauma, illness, onset of chronic disease or disorder in self or parent, death, divorce).

(vi) Theories of development (e.g., constructivist theory, social learning theory, ecological theory).

(vii) How development is influenced by the organism–environment interaction over time (e.g., understanding the relationship between the behavior of the individual and the social, academic, or work environment).

(viii) Family systems functioning and family stages in life and how these impact on individuals (e.g., family life cycle, parent–adolescent communication, birth of child).

2.11.5.2.5 Knowledge statements related to assessment and diagnosis

(i) Psychometric theory and concepts (e.g., measurement, reliability, validity, item characteristics, test fairness, standardization, norms), and test validation procedures (e.g., criterion, predictive, construct, and content strategies; appropriate measurement standards and legal regulations).

(ii) Assessment models (e.g., psychometric, behavioral, neuropsychological, ecological).

(iii) Tests for the measurement of characteristics of individuals (e.g., social, emotional, and behavioral functioning; cognitive; achievement; aptitude; personality; neuropsychological; vocational interest), and the adaptation of these tests for use with special populations.

(iv) Techniques other than tests (e.g., interviews, surveys, naturalistic and structured behavioral observations, physical status, history/ biographical data) for the measurement of characteristics of individuals.

(v) Instruments and methods for the measurement of characteristics of jobs, organizations, educational and other social institutions (e.g., job analysis, job evaluation, need assessment, organizational diagnosis, ecological assessment).

(vi) Methods for evaluating environmental/ ecological influences on individuals, groups, or organizations (e.g., organizational frameworks, functional analysis of behavior).

(vii) Criteria for selecting assessment devices/approaches (e.g., cultural appropriateness, utility analysis and cost effectiveness, relevance to referral concern).

(viii) Various classification systems (e.g., *DSM-IV* (*Diagnostic and statistical manual of mental disorders, 4th edition*), AAMR (American Academy of Mental Retardation), ICD (*International classification of diseases*)) for diagnosing client functioning.

(ix) Epidemiology of behavioral disorders, base rates of disorders in clinical or demographic populations, comorbidity among behavioral disorders and with medical disorders.

(x) Theory and techniques for the measurement of client changes (e.g., client tracking, formative and summative evaluation, program evaluation).

(xi) Human imaging principles, uses, and clinical implications (e.g., magnetic resonance imaging, computed tomography scanning, positron emission tomography scanning, electroencephalography, SPECT, single photoemission computed tomography).

(xii) Human laboratory principles, uses, and clinical implications (e.g., drug screens, screening for genetic disorders).

(xiii) Use of computers and related technology in implementing tests, surveys, and other forms of assessment.

(xiv) Quality assurance measurement techniques.

2.11.5.2.6 Knowledge statements related to treatment/intervention

(i) Treatment planning process, including differential diagnosis and efficacy and outcome data.

(ii) Theories of treatment (e.g., behavioral, cognitive, and cognitive-behavioral approaches; psychodynamic approaches; systems/ecological approaches; humanistic approaches).

(iii) Treatment techniques/interventions for specific concerns or specific populations (e.g., marital and family, group therapy, crisis intervention, play therapy, feminist therapy, approaches to stress management, psychoeducational, time-limited/brief therapy, compensation strategies, culturally appropriate treatments and interventions).

(iv) System theories and system interventions (e.g., change of environment, school system, community interventions, family, job and equipment design, consultation).

(v) Organizational interventions (e.g., organizational development, organizational change, performance enhancement/management).

(vi) Consultation models (e.g., mental health, behavioral, instructional, organizational) and processes (e.g., stages, communication skills).

(vii) Human resource management interventions (e.g., selection, performance appraisal, training).

(viii) Theories of career development and counseling (e.g., career assessment, career counseling techniques).

(ix) Adjunctive and alternative interventions and appropriate referral (e.g., physicians, 12-step programs, psychopharmacology, inpatient or partial hospitalization, support groups).

(x) Service delivery systems (e.g., education, health, mental health, social services, forensics, business, and industry), including the roles of other professionals.

(xi) Quality assurance measurement techniques.

2.11.5.2.7 Knowledge statements related to research and evaluation

(i) Research methods (e.g., sampling, instrumentation, data collection procedures).

(ii) Research design (e.g., hypothesis generation, experimental, quasi-experimental, naturalistic inquiry, group designs, and single-case research).

(iii) Appropriate analytical methods (e.g., qualitative, quantitative, descriptive, inferential; univariate, bivariate, and multivariate; parametric and nonparametric) and interpretation (e.g., causal versus correlational; degree and nature of generalizability).

(iv) Criteria for critical appraisal and utilization of research (e.g., technical adequacy; limitations to generalizations; threats to internal, external, and construct validity).

(v) Program planning and evaluation strategies and techniques (e.g., need assessment, process/implementation evaluation, outcome evaluation).

2.11.5.2.8 Knowledge statements related to ethical/legal/professional issues

(i) APA Ethical Principles of Psychologists and Code of Conduct and/or Canadian Code of Ethics for Psychologists (e.g., confidentiality, research, dual relationships, limits of competence, advertising practices, informed consent, record-keeping).

(ii) Professional standards and guidelines for the practice of psychology (e.g., APA/CPA Standards for Providers of Psychological Services, AERA/APA/NCME Standards for Educational and Psychological Testing. ASPPB Code of Conduct, ASPPB Model Licensure Act, credentialing requirements for advanced specialties and proficiencies, other published guidelines for special populations such as women and minorities).

(iii) Pertinent federal, state and/or provincial laws/statutes that affect psychological practice (e.g., laws and regulations relating to family and child protection, education, disabilities, discrimination, duty to warn and privileged communication, commitment and least restrictive care, continuing education requirements, practice regulations, licensure regulations).

(iv) Ethical decision-making process (e.g., balancing professionalism with entrepreneurship, integration of ethical principles and legal/regulatory standards).

(v) Models and approaches for training and supervision of self and others (e.g., methods for developing and enhancing knowledge and proficiency, continuing education, professional self-management, clinical supervision, peer consultation and supervision).

2.11.6 REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1985). *Standards for educa-*

tional and psychological testing. Washington, DC: American Psychological Association.

Association of State and Provincial Psychology Boards (1996). *Regulation of professional psychology: An international resource directory*. Montgomery, AL: Author.

Bolton, R. (1993). Pretesting questionnaires: Content analyses of respondents' concurrent verbal protocols. *Marketing Science*, *12*, 280–303.

Dillman, D. A. (1978). *Mail and telephone surveys: The total design method.* New York: Wiley.

Greenberg, S., Smith, I. L., & Muenzen, P. M. (1996). *Study of the practice of licensed psychologist in the United States and Canada*. Montgomery, AL: Association of State and Provincial Psychology Boards.

Hambleton, R. K., & Smith, I. L. (1988). *Content validity and fairness review of the 1987 forms of the Examination for Professional Practice in Psychology (EPPP) (final report)*. New York: Professional Examination Service.

Interprofessional Workgroup on Health Professions Regulation (1996, November). *Views of the licensure and regulation of healthcare professionals*. Unpublished manuscript.

Knapp, J. E., & Knapp, L. G. (1995). Practice analysis: Building the foundation for validity. In J. C. Impara (Ed.), *Licensure testing: Purposes, procedures, and practices* (pp. 93–116). Lincoln, NE: Buros Institute of Mental Measurements.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: Macmillan.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*, 741–749.

Pew Health Professions Commission (1995, December). *Reforming healthcare workforce regulation: Policy considerations for the 21st century*. San Francisco, CA: Author.

Richman, S. (1982). *The role delineation study for the Examination for Professional Practice in Psychology (final report)*. New York: Professional Examination Service.

Rosen, G. A. (1991). *Research digest: the Examination for Professional Practice in Psychology*. Montgomery, AL: American Association of State Psychology Boards.

Rosen, G. A., & Mirone, J. A. (1986). *The test specification revision project for the Examination for Professional Practice in Psychology (final report)*. New York: Professional Examination Service.

Rosenfeld, M., Shimberg, B., & Thornton, R. F. (1983). *Job analysis of licensed psychologists in the United States and Canada*. Princeton, NJ: The Educational Testing Service.

Schoon, C. G. (1985). Methods for defining and assessing professional competence. *Professional Practice of Psychology*, *6*(1), 144–155.

Schoon, C. G., & Smith, I. L. (1996, Summer). Guidelines for credentialing programs. *Professions Education Research Quarterly*, *17*(4), 6–9.

Smith, I. L. (1984). *Content validity study of the EPPP item bank*. New York: Professional Examination Service.

Smith, I. L., & Greenberg, S. (1991). *Final report for the EPPP test specifications enhancement project*. New York: Professional Examination Service.

Werner, E. (1989). *Analysis of the validity of the Examination for Professional Practice of Psychology*. Sacramento, CA: California Department of Consumer Affairs.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.12

An Assessment Tool in Credentialing: The Examination for Professional Practice in Psychology

GERALD A. ROSEN
Sylvan Learning Systems, Philadelphia, PA, USA

2.12.1 INTRODUCTION 138

2.12.2 CONTENT OF THE EPPP 138
  2.12.2.1 The Role Delineation Study of Professional Psychology 138
  2.12.2.2 The Job Analysis of Licensed Psychologists 140
  2.12.2.3 The Test Specifications Revision Project 141
  2.12.2.4 The Test Specifications Enhancement Project 141
  2.12.2.5 The Study of the Practice of Licensed Psychologists in the USA and Canada 141

2.12.3 CONSTRUCTION, ADMINISTRATION, AND SCORING OF THE EPPP 145
  2.12.3.1 Item Development 145
  2.12.3.2 Test Construction 145
  2.12.3.3 Test Administration 145
  2.12.3.4 Scoring 146

2.12.4 CANDIDATE PERFORMANCE ON THE EPPP 146
  2.12.4.1 Background Survey Data 146
  2.12.4.2 The Educational Reporting Service 147

2.12.5 RELIABILITY AND VALIDITY OF THE EPPP 148
  2.12.5.1 Reliability 148
  2.12.5.2 Validity 148
    2.12.5.2.1 Construct validity 148
    2.12.5.2.2 Content validity 148

2.12.6 FUTURE DIRECTIONS FOR THE EPPP 149
  2.12.6.1 Specialty Examinations 149
  2.12.6.2 Overseas Use of the EPPP 149
  2.12.6.3 Computer Administration of the EPPP 149

2.12.7 SUMMARY 150

2.12.8 REFERENCES 150

2.12.1 INTRODUCTION

The Examination for Professional Practice in Psychology (EPPP) is a 200-item multiple-choice examination owned and developed by the Association of State and Provincial Psychology Boards (ASPPB). It is made available by ASPPB to its members for their use in evaluating candidates for licensure or certification as psychologists. The EPPP is administered semiannually in April and October in paper-and-pencil format by the ASPPB member boards. Examinations are scored by the Professional Examination Service (PES), a not-for-profit testing company. PES also provides technical assistance in the development of items and examinations as well as a variety of support services including printing, shipping, and statistical analyses.

According to the ASPPB:

> The EPPP is intended to evaluate the knowledge that should have been acquired by any candidate who has obtained a doctoral degree in psychology followed by one or two years of postdoctoral experience. Such candidates are expected to have acquired a broad base of knowledge of psychology, regardless of individual specialties. This knowledge, and the candidate's ability to apply it, are assessed through the candidate's responses to objective, multiple-choice questions representative of the field at large. Most candidates with the required academic and experiential preparation should be able to pass the test. (ASPPB, 1995a, p. 3)

2.12.2 CONTENT OF THE EPPP

In its early years, the content of the EPPP was based on a list of subjects determined by the members of the Examination Committee, a panel of psychologists appointed by the Board of Directors of ASPPB. For example, Form 6 covered the following subjects: background knowledge; methodology; clinical; other applications; and professional conduct, affairs, and ethics (Carlson, 1978). By today's standards of test development, this list represents a remarkable mix of extremely broad topics. Such informal methods of test construction were common in the 1960s and persisted well into the next decade in psychology and many other professions with national licensure examination programs. However, as the number of participating psychology licensure boards and candidates increased, and as the USA became a more litigious society, attention was drawn to the issue of enhancing the legal defensibility of the EPPP. In response to that need, ASPPB, known at that time as the American Association of State Psychology Boards, funded a series of studies and development projects in order to produce and enhance formal test specifications for the EPPP.

2.12.2.1 The Role Delineation Study of Professional Psychology

In 1978, ASPPB selected PES to conduct a role delineation study of professional psychology in order to help determine the most appropriate content for the EPPP (Richman, 1982). In the first phase of the study, a blue-ribbon panel of 10 psychologists representing clinical, counseling, industrial/organizational, school, and academic psychology described professional psychology in terms of five performance domains, 19 roles, and 49 knowledge statements. Their work was submitted to a group of 80 selected psychologists who evaluated the elements of the role delineation in terms of importance for entry-level practice and to a group of 90 Fellows of the American Psychological Association (APA) who rated the domains and roles for importance and measurability in a multiple-choice test format. Based on these evaluations, the knowledge statements were discarded and the document was trimmed to four domains of practice and eight professional roles. In 1981, a second panel of well-known psychologists added a fifth domain of practice, combined two of the roles, and wrote 58 new knowledge statements linked to the roles and domains. The revised document was submitted to a sample of over 1000 psychologists, of whom 432 rated the elements on six scales, including, for example, importance for entry-level practice and frequency of performance. The final document was submitted to the ASPPB Board of Directors and Examination Committee. With only minor additional changes, the new test specifications were used to develop Forms 16–25 of the EPPP administered from April 1982 to October 1986. Table 1 contains a summary of the test specifications based on the role delineation study.

The importance of the role delineation study in the history of the EPPP cannot be overestimated. For the first time the examination, which, by 1982, was being administered to over 4000 licensure candidates annually, had a systematically derived content outline; one that changed its focus from broad scientific knowledge to knowledge required for safe and effective practice. Perhaps more significantly, it was the first time in the history of the practice of psychology in the USA that there was a consensus as to what constituted performance as a professional psychologist.

Table 1 Summary of EPPP test specifications: April 1982–October 1986.

Domain I: Techniques for appraising and assessing (25–30%)

Role 1: Select, modify, and use psychological assessment techniques/instruments
Requires knowledge of:

1. Widely used psychological tests
2. Limitations and appropriate uses of techniques of behavioral assessment
3. Limitations and appropriate uses of paper-and-pencil tests
4. Standards for test construction
5. Sources of information about psychological assessment techniques
6. Standards of test administration and release of test information
7. Procedures for assessing presenting problems
8. Behavioral effects of disabilities and mental retardation

Role 2: Interpret and report results of assessments
Requires knowledge of:

1. Standard procedures for reporting test scores
2. Modalities for the communication of results
3. Methods of presentation of recommendations for interventions
4. Statistical issues in the interpretation of test scores
5. Relationships between test results and diagnosis
6. Symptom patterns of major diagnostic categories of psychopathology
7. Effects of major psychotropic drugs
8. Relationships between psychological theories and assessment
9. Social determinants of behavior
10. Signs of organizational effectiveness
11. Techniques for matching abilities, skills, and interests to job specifications

Domain II: Design, implementation, and assessment of intervention (25–30%)

Role 3: Design, implement, and evaluate an intervention plan
Requires knowledge of:

1. Components of an intervention plan
2. Major psychological intervention theories
3. Major psychological intervention techniques
4. Psychotherapy research
5. Demographic and environmental assessment techniques
6. Indications and contraindications for matching clients with interventions
7. Community resources
8. Current educational processes/programs
9. Organizational processes
10. Organizational consultation techniques
11. Techniques for identifying measurable goals
12. Record keeping requirements
13. Forms of pre- and postevaluation
14. Process evaluation
15. Peer review techniques
16. Program evaluation
17. Time-line goals
18. Indicators of client acceptance of interventions

Domain III: Uses of the psychological literature (15%)

Role 4: Interpret and apply general psychological literature
Requires knowledge of:

1. Developmental psychology
2. Individual differences and intelligence
3. Information process and thinking
4. Learning and memory
5. Personality and social behavior
6. Physiological basis of behavior
7. Psychological systems and theories
8. Psychopathology
9. Sensation, perception, motivation, and change
10. Specialty areas (e.g., biofeedback, women, forensic)
11. Specialty areas (e.g., clinical, counseling, school, industrial)

Table 1 (continued)

Domain IV: Techniques of research (15%)

Role 5: Design and implement research
Requires knowledge of:

1. Research designs
2. Experimental artifacts
3. Data presentation
4. Test construction
5. Theories of test construction
6. APA and division 14 standards for test construction and validation
7. Elementary computer technology
8. Statistics

Domain V: Professional and ethical issues (18%)

Role 6: Evaluate activities within the limitations of professional, ethical, and legal statutes
Requires knowledge of:

1. APA Code of Ethics
2. Legal issues

2.12.2.2 The Job Analysis of Licensed Psychologists

Role delineation studies are based upon expert judgments of professional practice roles and responsibilities and the knowledge required for their performance. Job analyses are typically based upon direct observations of and/or interviews with job incumbents. In 1981, in an attempt to assemble this type of direct evidence of what constituted the professional practice of psychology, ASPPB selected Educational Testing Service (ETS) to perform a job analysis of licensed psychologists. The ETS job analysis (Rosenfeld, Shimberg, & Thornton, 1983) began with the development of a job analysis inventory based on the duties and responsibilities of psychologists gathered from a literature search, two advisory committees, and interviews with approximately 100 licensed psychologists. A draft job analysis inventory was field-tested on a sample of 150 licensed psychologists. Data received from 80 respondents were used to produce a final inventory containing three broad areas of practice, 59 responsibilities, and 111 knowledge and technique statements. The inventory was mailed to 2994 licensed psychologists in the USA and Canada. Usable responses were received from 1585 individuals, of which 1307 (82%) were clinical, counseling, school, or industrial psychologists, representing the four major practice areas of licensed psychologists.

Factor analyses of the ratings of the 59 responsibilities were used to produce four job dimensions: research and measurement; intervention; organizational applications; and assessment. Of the 111 knowledge and technique statements, 62 were used by at least half of the respondents in the four major practice areas, and 49 of the 62 were rated as important and frequently used. When these 49 knowledge and technique statements were examined for their correlations with the four dimensions, 38 significant linkages were found. The authors concluded that the 38 knowledge and technique statements represented an important body of psychological knowledge frequently applied by licensed psychologists in the four major practice areas.

It should be noted that the job analysis project did not result in the production of new test specifications for the EPPP. It did, however, produce a wealth of new information about the practice of psychology by licensed psychologists. Based on these findings, the authors offered numerous recommendations for the production of a new EPPP content outline including: (i) review of the linkages between the knowledge and technique statements and the dimensions of practice by an independent panel of experts; (ii) further definition of the knowledge and technique statements in operational terms; and (iii) determination of the best assessment mode (e.g., oral vs. written examination) for each of the knowledge and technique statements. Of particular interest was the investigators' statement that:

> The job analysis study has shown that, despite differences on their work settings, client populations, nature of problems, and techniques used, licensed psychologists, in general, share many common responsibilities and use many of the same knowledge and techniques in their work.(p. IX-2)

This statement, and the data upon which it rested, served as the justification for the continued development of the EPPP as an examination measuring broad, general knowledge of psychology. Table 2 contains a summary of the ETS job analysis findings.

2.12.2.3 The Test Specifications Revision Project

In 1985, ASPPB selected PES to undertake a project to produce revised test specifications for the EPPP (Rosen & Mirone, 1986). To accomplish this task, data from the two previous studies, the role delineation and the job analysis, were used. The methodology of the study was similar to the role delineation. A panel of 25 experts representing the four major areas of professional psychology produced an initial document consisting of six dimensions of practice, 58 responsibilities, and 58 knowledge and applications statements. The selection of these particular elements was based on the panel's ratings of professional level, importance, public protection, and most appropriate assessment method. The 58 responsibilities showed considerable overlap and were subsequently reduced to nine by a process of combining and rewording. The 58 knowledge and technique statements were linked to the nine responsibilities. The panel was unable to link nine of the knowledge and applications statements, and thus their number was reduced to 49.

In the final phase of the project, a second panel of experts reduced the number of dimensions from six to five, linked the nine roles to the five dimensions, confirmed the linkages of the knowledge and applications statements to the responsibilities, and, through a process relying on their pooled judgments, provided weights for the dimensions and responsibilities. As a result, the EPPP had a new set of test specifications based on a role delineation, a job analysis, and an extended review by two additional representative panels of experts. Table 3 contains a summary of the revised EPPP test specifications used to develop Forms 26–46 administered from April 1987 to October 1996.

2.12.2.4 The Test Specifications Enhancement Project

In 1990, ASPPB selected PES to conduct a small project, suggested by Werner (1989), to review the EPPP test specifications, examine the accuracy of item classifications, and provide recommendations for improving the effectiveness of the test development process (Smith & Greenberg, 1991). The project used a representative panel of psychologists to review the accuracy of classification of a sample of 176 items, draft recommendations for a set of test construction quality control procedures, and suggest clarifications to the wordings of the dimensions, responsibilities, and knowledge and application statements. Of the 176 items reviewed, 94% were found to be classified correctly. Incorrectly classified items were reclassified or dropped from the item bank. The review of the test specifications resulted in the minor rewording of one dimension and several of the responsibilities and knowledge and application statements. Finally, a set of guidelines was submitted to the Examination Committee for its use in the test development process. While the test specifications enhancement project did not result in any significant changes to the EPPP test specifications or the examination development process, its findings lent significant credibility to both.

2.12.2.5 The Study of the Practice of Licensed Psychologists in the USA and Canada

In 1993, ASPPB requested that PES conduct an analysis of the practice of licensed psychologists in the USA and Canada. Developing a new set of test specifications for the EPPP was only one of the purposes of the study. The others were: (i) to update the list of responsibilities and knowledge required for safe and effective professional practice by licensed psychologists; (ii) to provide a basis for communicating the contemporary role of the licensed psychologist to the public and third-party healthcare payers regarding the scope of professional psychological practice; and (iii) to identify current and emerging practice areas for reviewing educational curricula, schools, training programs, continuing education programs, and specialty recognition initiatives (Greenberg, Smith, & Muenzen, 1995).

The project was conducted in collaboration with a five-member oversight committee and a 12-member representative panel of psychologists. The oversight committee was responsible for policy decisions, selecting the members of the panel, and performing reviews of draft and final survey and practice delineation documents. The panel produced the draft and final practice delineations and reviewed data analyses and results. They also selected the subject-matter experts who took part in the complementary data collection activities.

The practice analysis produced five roles with four groups of underlying responsibilities, and

Table 2 Summary of the findings of the ETS job analysis study.

Dimension I: Research and measurement

Responsibilities:

1. Determine research design for a study
2. Analyze data
3. Observe, record, or collect data
4. Prepare written reports of research
5. Develop methods of assessing critical variables
6. Identify methods of assessing critical variables
7. Define questions to be answered
8. Identify the problem to be studied
9. Develop proposals for research studies
10. Evaluate relevant literature
11. Adhere to guidelines for the conduct of research
12. Assist others in understanding the implications of research findings
13. Keep abreast of professional and scientific developments
14. Serve as an expert witness

Knowledge and techniques:

1. Data collection techniques
2. Test construction
3. Validity concepts
4. Reliability concepts
5. Test fairness
6. Research and experimental design
7. Basic statistics
8. Advanced statistics
9. Sampling theory
10. Criterion measures

Dimension II: Intervention

Responsibilities:

1. Plan intervention strategies
2. Discuss alternative courses of action
3. Set realistic goals
4. Obtain informed consent
5. Provide assistance on personal/organizational problems
6. Recommend services of other professionals
7. Maintain liaison with other service providers
8. Monitor and evaluate effectiveness of interventions
9. Modify intervention strategy

Knowledge and techniques:

1. Interviewing techniques
2. Observational techniques
3. Stress management techniques
4. Conflict management techniques
5. Consultation techniques
6. Human growth and development
7. Learning
8. Personality
9. Individual differences
10. Perception
11. Intelligence
12. Ethical guidelines
13. Specialty guidelines

Dimension III: Organizational applications

Responsibilities:

1. Evaluate the physical and technological environment
2. Evaluate the human resource needs of the organization
3. Conduct cost-benefit analyses
4. Provide assistance in modifying programs or systems

Table 2 (continued)

Knowledge and techniques:
1. Decision-making strategies
2. Learning
3. Motivation
4. Cognition
5. Reliability concepts
6. Organizational morale

Dimension IV: Assessment

Responsibilities:
1. Select instruments, techniques, or procedures
2. Administer and score instruments, techniques, or procedures
3. Interpret and evaluate results
4. Integrate assessment data with other information
5. Prepare a report of findings and diagnosis/problem definition

Knowledge and techniques:
1. Group tests of mental ability
2. Diagnostic tests
3. Tests of social maturity
4. Measures of self-concept
5. Standardized test procedures
6. Personality
7. Motivation
8. Perception
9. Cognition
10. Intelligence
11. Test construction
12. Test scores and norms
13. Strengths and limitations of assessment procedures
14. Validity concepts
15. Reliability concepts
16. Factors affecting test performance
17. Coaching and practice
18. Test fairness

eight content areas each with an underlying group of knowledge statements. The roles included direct service, outreach and consultation, academic preparation and professional development, research and evaluation, and other. The responsibilities were grouped under four of these headings: direct service (e.g., plan, design, and implement intervention programs); outreach and consultation (e.g., disseminate knowledge of psychology and its value to the general public); academic preparation and professional development (e.g., provide mentoring to students and/or professional psychologists); and research and evaluation (e.g., apply research findings to practice). The eight content areas were: assessment and diagnosis; biological bases of behavior; cognitive–affective bases of behavior; ethical/legal/professional issues; growth and lifespan development; research and evaluation; social and multicultural bases of behavior; and treatment/intervention. Only the eight content areas are weighted for test construction purposes. The roles and responsibilities are used primarily as an aid in the item development process (K. Midkiff, PES, personal communication, August 21, 1996). For a complete description of the practice analysis project and the content areas and knowledge statements contained in the test specifications now used to develop the EPPP see Chapter 11, this volume.

Comparisons with previous versions of the test specifications reveal some interesting changes. First, the specifications now contain subject-matter areas (i.e., content areas) from psychology as an academic discipline (e.g., biological bases of behavior, growth and lifespan development). The two earlier versions essentially confined themselves to a strict practice-related format. Practice relatedness is important for legal defensibility because it helps to ensure that an examination will only contain

Table 3 Summary of EPPP test specifications: April 1987–October 1996.

Dimension I: Problem definition/diagnosis (26%)

Responsibility 1: Conduct interviews, observe and gather information from client/patient and relevant others
Requires knowledge of:
1. Techniques of behavioral assessment
2. Theories and principles useful in identifying patient needs and problems
3. Factors affecting behavior
4. Symptoms of common physical diseases and psychophysiological reactions and syndromes
5. Effects of major psychotropic drugs and common prescription drugs on behavior, affect, and cognition

Responsibility 2: Select, administer, and score instruments and techniques and interpret results
Requires knowledge of:
1. Techniques for testing and/or measuring relevant characteristics
2. Concepts relating to tests and measurement

Responsibility 3: Organize and evaluate collected information, plan for additional information as needed, and formulate a working hypothesis
Requires knowledge of:
1. Techniques of data organization
2. The APA's *Diagnostic and statistical manual of mental disorders* (1997) Fourth edition

Dimension II: Design, implementation, and assessment of interventions (26%)

Responsibility 4: Design, conduct, and evaluate interventions and programs to promote effective functioning
Requires knowledge of:
1. Intervention techniques
2. Attitude and value formation and modification techniques

Dimension III: Research (17.5%)

Responsibility 5: Design and implement research and report conclusions and recommendations
Requires knowledge of:
1. Methods of inquiry
2. Research design
3. Experimental artifacts
4. Statistics

Dimension IV: Professional, ethical, and legal issues (16.5%)

Responsibility 6: Take steps to ensure adherence to professional, governmental, and judicial guidelines and regulations for professional and scientific activities
Requires knowledge of:
1. Ethical, professional, and legal issues

Dimension V: Applications to social systems (14%)

Responsibility 7: Develop, conduct, evaluate, and modify intervention strategies and programs designed to promote effective functioning within the social system
Requires knowledge of:
1. Factors affecting the quality of work life
2. Performance evaluation and/or appraisal methods
3. Consultation models and techniques
4. Work sample and other performance tests
5. Decision making strategies
6. Factors affecting the morale of social systems
7. Organizational structures and processes
8. Group dynamics

Responsibility 8: Evaluate the physical, technological, and social environments of the system
Requires knowledge of:
1. Evaluation techniques for physical, technological, and social environments

Responsibility 9: Evaluate the human resource needs of the system
Requires knowledge of:
1. Techniques for the evaluation of human resource needs

items requiring knowledge related to safe and effective practice. However, there is no reason why content areas cannot also be included in practice-related test specifications, as long as the relationship to practice has been demonstrated. Furthermore, the inclusion of content areas helps make the test specifications more meaningful to licensure candidates because these terms are much closer to those used by them as descriptors of what they have learned.

While the new test specifications document continues the ASPPB policy of ensuring that the EPPP is based on current practice, it is not without some potential problems. Although many of the problematic features appear in the unweighted portions of the test specifications, they are still worth noting and are potentially important. For example, the category "other" listed among the five roles seems particularly unfortunate. It is defined as job-related knowledge which is not role specific, but no responsibilities or knowledge statements have been delineated for it. Such "waste-basket" categories can serve as a tempting means of using items that do not properly fit the test specifications. That is, if a item cannot be correctly classified, it can always be put in the "other" category and used anyway. Also, it is arguably the case that the test specifications contain some elements that are not appropriate to entry-level practice. Possibly problematic is the role, academic preparation and professional development, which includes the development, implementation, and administration of educational programs for psychologists. Another element in this category is content area 4, knowledge statement 8, which specifies models and approaches for training and supervision of self and others. Finally, some of the content areas of the test specifications could benefit from clearer wording. For example, "growth and lifespan development" might be better termed "developmental bases of behavior," making it parallel with several other categories.

2.12.3 CONSTRUCTION, ADMINISTRATION, AND SCORING OF THE EPPP

2.12.3.1 Item Development

The EPPP is constructed from a bank of items developed and maintained by PES. Approximately four times per year, PES conducts an item-writing workshop in a different region of the USA or Canada. Licensed psychologists and, occasionally, advanced graduate students are invited to participate. PES supplies instructional materials and item-writing assignments to participants in advance of the workshop. The participants are asked to write a specified number of items prior to the workshop. Most of the time spent during the meeting consists of critical reviews of the previously prepared items. Most items undergo significant revisions and some are discarded. All retained items are checked for accuracy of classification. When these tasks are completed, the remainder of the workshop time is devoted to the writing and review of additional items. Some workshops are specifically targeted to address shortage areas in the item bank.

2.12.3.2 Test Construction

The EPPP is constructed by the Examination Committee with the assistance of PES. The committee meets twice each year in January and June. At the January meeting work is begun on the EPPP form to be administered in October of the same year. The process begins with a draft examination prepared by PES. The draft consists of the previous October's form with approximately 70 items removed, based on a review of item analysis data and item usage information, and between 80 and 90 items added from the EPPP item bank. The committee reviews each item individually and removes or revises items as it sees fit. PES records the instructions of the committee and executes the changes after the meeting. Also, at the January meeting, the committee reviews for a second time the EPPP form to be administered in April of the same year. The April form will have had its first review at the committee's meeting the previous June. Once again, the committee reviews each item individually and makes revisions and deletions. The PES staff record the committee's instructions and keep a tally of the item classifications. After the committee has completed its second review of the April examination form, additional item deletions and substitutions from the item bank are made to bring the examination into conformity with the test specifications. A third and final review of the examination is done by mail after the meeting. The June meeting is identical to January, except that the October form receives a second review and the form for April of the following year is reviewed for the first time.

2.12.3.3 Test Administration

The EPPP is administered by the ASPPB member boards. However, the boards must administer the EPPP in the manner and on the dates specified by ASPPB. This means that all procedures in the Manual of Instructions (PES, 1995a) and all ASPPB policies and procedures must be followed. The examination contains 200, four-option multiple-choice questions and its timed portion is four hours long. The testing session begins with the administration of an optional, nine-item candidate background survey. After the survey has been completed, the test booklets are distributed and the standardized instructions are read to the candidates. The timed portion of the test administration begins immediately thereafter. Candidates with documented disabilities and/or impairments may be tested under nonstandard conditions. Extra testing time is the most common nonstandard condition. However, large-print and Braille examinations are routinely made available. Other special accommodations include readers, signers, and other amanuenses. Some jurisdictions routinely provide extra testing time to candidates who are not native speakers of English. It is the policy of ASPPB to encourage member boards to provide reasonable accommodations to candidates with disabilities or impairments.

2.12.3.4 Scoring

Immediately following test administration, answer sheets are returned to PES for scanning and scoring. Prior to final scoring of the EPPP, PES performs and reviews an item analysis to determine if there are any statistical data that suggest that any of the 200 items may be problematic. This includes miskeyed items, items with no defensible correct answer, and items with more than one defensible correct answer. The results of the item analysis review are discussed with the chairperson and other members of the examination committee. Occasionally, outside experts in a particular subject area are also consulted. Also considered are comments received from examinees. (When tested, candidates are informed that any comments submitted in writing to PES during the two-week period following the examination will be considered in the final scoring process of the EPPP.) Items deemed problematic, if any, are either rekeyed (i.e., have the key corrected) or multiply keyed (i.e., have more than one option scored as correct). EPPP scores are always based on 200 items. Scores are reported by PES to the member boards which, in turn, notify the candidates of their test results and licensure decisions.

It is important to note that all licensure decisions are made by the boards. Neither PES nor ASPPB participate in the licensure decision process. They do, however, provide information that is sometimes used as an aid in the process. For example, along with the examination results, PES supplies a cumulative data table which provides normative information (number, range, mean, and standard deviation), on seven categories of examinees (e.g., all doctoral candidates, doctoral first-time candidates, all masters degree candidates). Many boards continue to use these data for passing score determination. PES also provides an equipercentile equating table that links scores on the just administered form of the EPPP to a base form. The equating table is used by several boards that employ criterion-referenced passing scores. In 1992, ASPPB conducted an investigation into performance on the EPPP, passing score determination procedures of its member boards, and passing scores applied by the boards. As a result of the study, ASPPB formally recommended a uniform passing score of 70% for the EPPP. It was hoped that the use of a uniform passing score by the member boards would enhance the geographical mobility of licensed psychologists by simplifying the process of licensure by endorsement. To date, over 30 boards have adopted the suggested uniform passing score.

2.12.4 CANDIDATE PERFORMANCE ON THE EPPP

Candidate performance on the EPPP has shown remarkable stability over time. For example, EPPP Forms 26–42, administered from April 1987 to April 1995, had an overall mean and standard deviation of 147.9 and 3.0, respectively (PES, 1995b). Scores in the part of the distribution where most passing scores fall, 130–150, are even more stable. It is common for the PES generated equating table to show passing score adjustments of only 0–2 points within this interval of scores.

2.12.4.1 Background Survey Data

ASPPB and PES have graciously made available rarely published data on candidate performance on the EPPP. Specifically, they have provided the numbers, means, and standard deviations of EPPP scores associated with five of the questions on the background survey administered with the EPPP. While most candidates respond to the survey, they are not required to do so. Furthermore, there is no way to verify the accuracy of their responses. Nevertheless, the data, presented in Table 4, clearly provide interesting information on the EPPP and the licensure candidates who sit for it.

For example, graduates of APA accredited programs do better on the EPPP than others. In 1995, 87% of the survey respondents specialized in clinical (57%), counseling (23%), or school (7%) psychology. Indeed, the smallest identified specialization is industrial–organizational psychology (0.5%). This is ironic, given ASPPB's commitment to the representation, some might say overrepresentation, of industrial–organizational psychologists on its expert panels in all its studies of psychological practice. Finally, and it should come as no surprise, the more candidates study for the EPPP the better they tend to do. While there is a slight downturn at the highest level of preparation, the trend is clear. Thus, the best advice for doing well on the EPPP, as with most achievement tests, remains, "study hard!"

2.12.4.2 The Educational Reporting Service

Another source of information on candidate performance on the EPPP is the *Educational Reporting Service: EPPP Performance by Designated Doctoral Programs in Psychology* (ASPPB, 1995b). This listing provides the numbers, means, and standard deviations for

Table 4 1995 EPPP performance data associated with selected background survey response categories.

Response category	*Number*	*Mean*	*SD*
Department awarding highest degree:			
Psychology	4268	144.9	21.1
Educational psychology	997	139.0	21.8
Other	392	137.4	21.4
Graduate of APA accredited program:			
Yes	3000	150.3	19.1
No	770	138.9	21.3
Not applicable	1822	134.0	20.9
Major area of graduate specialization:			
Clinical	3348	146.3	20.9
Counseling	1349	138.2	22.2
School	413	138.5	20.0
Educational	108	137.4	21.7
Developmental	80	145.1	17.3
Community	65	128.5	18.9
Experimental	46	148.0	17.3
Social–personality	34	141.8	20.1
Industrial–organizational	27	143.6	16.0
Other	152	144.0	19.6
Primary professional affiliation:			
Mental health agency	1269	140.9	21.5
Private practice	1016	141.6	21.0
Hospital	830	148.4	20.3
College/university	640	148.2	20.6
School system	501	137.2	21.4
Government	450	141.6	21.7
Medical school/center	266	153.2	17.9
Business	107	133.6	22.7
Research institute	33	150.5	26.6
Other	416	142.8	20.8
Time spent preparing for the EPPP:			
None	47	116.3	26.6
Up to 49 hours	502	129.8	23.9
50–99 hours	1423	142.6	21.6
100–199 hours	1849	147.3	19.7
200 or more hours	1676	144.7	19.4

hundreds of doctoral programs in psychology in the USA and Canada. To be included in the publication a program must meet the requirements for "designation" applied jointly by ASPPB and the Council for the National Register of Health Service Providers in Psychology. The source of the data is candidate self-report at the time of EPPP testing. Tabulations are performed by PES after the October test administration and the publication is revised annually.

2.12.5 RELIABILITY AND VALIDITY OF THE EPPP

2.12.5.1 Reliability

Reliability coefficients (KR_{20}) of EPPP Forms 26–42, administered from April 1987 to April 1995, have ranged from 0.90 to 0.93 (PES, 1995b). Thus, there is little doubt that the EPPP is a reliable instrument. The validity of the EPPP rests primarily upon its content. That is, the EPPP is a valid instrument to the extent that its test specifications are an appropriate representation of the practice of professional psychology and to the extent that test forms accurately represent the test specifications. The above section on EPPP content documents a continuing commitment on the part of ASPPB to study professional practice and to ensure that it is reflected in the test specifications. The mission of the examination committee is to build tests according to the test specifications. The remarkable reliability and stability of candidate performance are indicators of the quality of their work. Other indicators of the validity of the EPPP were seen in Table 4. Where the numbers are large enough to provide stable data, all the major differences in mean scores are in the expected directions.

2.12.5.2 Validity

A digest of EPPP research, including validation research, is published by ASPPB (Rosen, 1991). Included in the document are summaries of criterion-related and content validity studies. The EPPP, like most licensure examinations, has not stimulated a wealth of validation research. Such studies are not easy to conduct. Data are sometimes unavailable for a variety of legal reasons, including confidentiality. Criterion-related validity studies often confirm the obvious, as can, in some cases, be seen in Table 4. Predictive validity studies are essentially impossible to conduct, for practical as well as theoretical reasons, and may be entirely inappropriate for licensure examinations (Rosen, 1986). Content validity studies are of the greatest importance for licensure examinations, but they are relatively costly. Furthermore, because of the need to safeguard the security of examination related materials, they are typically performed by the agency sponsoring or owning the licensure examination, its testing company, or some other carefully selected agent.

2.12.5.2.1 Construct validity

Several researchers have demonstrated that as education in psychology increases, scores on the EPPP increase (Hayes & Mullins, 1978; Hayes & Schreimer, 1978; Terris, 1973). These studies assumed that the Ph.D. represents the highest degree, followed by the other doctorates. While holders of the Ph.D. in psychology continue to outscore others, the gap between them and licensure candidates with Ed.D. and Psy.D. degrees has greatly narrowed in the last decade (PES, 1995b). Other research has demonstrated that among undergraduate majors in psychology, members of honors programs or honorary societies perform better on the EPPP than others (Shrader, 1980). Tipton, Elder, and Ritz (1988) showed that graduates of APA accredited programs outperform graduates of nonaccredited programs. These findings are supported by the background survey data cited above. Kupfersmid and Fiala (1991) demonstrated that EPPP performance of graduates of APA accredited clinical and counseling programs is superior to the performance of graduates of school psychology programs and nonaccredited clinical and counseling psychology programs.

The largest construct validity study of the EPPP was conducted by Hoffman and Aderet (1984) with funding from ASPPB. The authors collected graduate school grades and EPPP scores for 866 psychologists. They discovered that: (i) patterns of mastery of psychology displayed during training were positively correlated with EPPP performance; (ii) candidates from APA accredited programs scored significantly higher than candidates from nonaccredited programs; (iii) EPPP performance was significantly higher for candidates who had been undergraduate psychology majors, particularly among counseling psychology majors; and (iv) candidates with specialties in clinical psychology performed significantly better than candidates from school, counseling, and educational psychology.

2.12.5.2.2 Content validity

As stated above, content-validity studies are the most important part of any validation strategy for licensure examinations. Two such studies of the EPPP are summarized here. Readers desiring more extensive content-validity information on the EPPP are referred to Smith, Hambleton, and Rosen (1988) and Rosen (1991).

Smith (1984) used three panels of experts to review the accuracy of item classifications and rate items for their contribution to the public protection function of the EPPP. The study demonstrated that the EPPP items could be readily classified into the roles in the test outline, although there was less success with knowledge statements. It was noted, however, that knowledge statements were not weighted elements of the test specifications.

Hambleton and Smith (1988) performed a content-validity study of the first two EPPP forms developed with the test specifications adopted in April 1987. Two panels of judges classified 400 items according to the responsibilities and knowledge statements in the test specifications and performed a fairness review of the items. Using an extremely stringent criterion of classification accuracy, the investigators determined that approximately 65% of the items were correctly classified as responsibilities and 52% were correctly classified as knowledge statements. The panelists identified seven items requiring additional review for fairness issues. However, none of these items was found to contain offensive stereotyping. All 400 were deemed appropriate for use on a generic licensure examination.

2.12.6 FUTURE DIRECTIONS FOR THE EPPP

2.12.6.1 Specialty Examinations

While almost all the licensure laws in North America remain relatively general, several states (e.g., Indiana) have adopted new laws which make distinctions between psychologists who are health service providers and others. If this trend continues, one logical consequence will be a growing demand for examinations that also recognize this distinction. It seems reasonable that when such differences are codified into law, different examinations become appropriate for the licensure candidates falling into the different categories. One problem associated with separate examinations has to do with the numbers of candidates who might require specialized examinations. As was shown in Table 4, almost all the EPPP candidates (over 85%) come from clinical, counseling, or school psychology programs. Frankly, there would not be enough candidates to justify the costs involved in developing specialized versions of the EPPP. One possible solution to this problem might be the development of a core examination to be taken by all candidates, with shorter specialty examinations for health service vs. nonhealth service providers in psychology.

2.12.6.2 Overseas Use of the EPPP

ASPPB has maintained a long-standing relationship with the European Federation of Psychology Associations. The two organizations have been exploring the possible use of the EPPP outside North America. In 1994, a version of the EPPP was administered in English in Scandinavia to a small sample of psychologists from the Nordic countries. This effort was part of a pilot project to determine the possible relevance and utility of the EPPP for licensure purposes in Europe. While the outcome of the pilot test has not been made public by ASPPB, it remains an indication of interest in the EPPP. ASPPB routinely receives requests for information about the EPPP from psychological associations and regulatory agencies in other countries. While it is unlikely that forms of the EPPP as administered in North America would be useful in the credentialing of psychologists in other parts of the world, it is clear that ASPPB is the most experienced entity when it comes to the regulation of psychological practice, and has much to share concerning the development and administration of a psychometrically sound licensure examination.

2.12.6.3 Computer Administration of the EPPP

ASPPB has tentative plans to begin computer administration of the EPPP by the year 2001 and has contracted with PES to perform pilot studies to gather information needed to support the decisions that must be made as part of the transition. These decisions bear upon examination equivalence, test security procedures, the psychometric testing model to be employed, and numerous administrative matters that are affected by a conversion from paper-and-pencil to computer-based testing. Moving to computer-administered examinations will remove the burden of test administration from the ASPPB member boards and enhance the standardization of testing conditions across test sites. Depending upon the mode of testing used, candidates may have the opportunity to take examinations by appointment at any of hundreds of locations. Testing may be available year round or during several predetermined testing periods. Scores may be made available to candidates immediately upon completion of the examination or just a few days thereafter. Thus, a major benefit to the candidates will be the increased availability of the examination and faster knowledge of results. This could well expedite hiring, promotions, and salary increases for many psychologists.

2.12.7 SUMMARY

The EPPP is a well-designed and well-constructed examination. Since it was first administered in May 1965 to one licensure candidate in West Virginia, ASPPB has endeavored to effect steady improvement in the EPPP. Those efforts have been most apparent in the years since 1978 when the first comprehensive effort to develop test specifications was undertaken. The improvements to the EPPP test specifications and studies of its content validity have continued on a consistent basis since the first project began. The examination is extremely reliable and has well-documented content validity. The criterion-related validity studies of the EPPP have all produced results that have supported the contention that it is a valid examination. The item development methodology provides an opportunity for psychologists throughout North America to participate in the overall process of test development. The test construction procedures employed by ASPPB, with the support of PES, provide the experts who serve on the ASPPB Examination Committee with the necessary time and opportunity to ensure that each form of the examination is thoughtfully produced and fully representative of the test specifications. In all relevant aspects, the EPPP conforms to the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1985). In short, the EPPP is a reliable and valid instrument that is a highly effective tool for the evaluation of candidates for licensure or certification as psychologists.

2.12.8 REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

Association of State and Provincial Psychology Boards (1995a). *Information for candidates: The Examination for Professional Practice in Psychology* (brochure). Montgomery, AL: Author.

Association of State and Provincial Psychology Boards (1995b). *Educational reporting service: EPPP performance by designated doctoral program in psychology*. New York: Professional Examination Service.

Carlson, H. (1978). The AASPB story: The beginnings and first 16 years of the American Association of State Psychology Boards, 1961–1977. *American Psychologist, 33*, 486–495.

Greenberg, S., Smith, I. L., & Muenzen, P. M. (1995). *Study of the practice of professional psychologists in the United States and Canada* (Executive summary). New York: Professional Examination Service.

Hambleton, R. K. & Smith, I. L. (1988). *Content validity and fairness review of the 1987 forms of the Examination for Professional Practice in Psychology (EPPP)*. New York: Professional Examination Service.

Hayes, J. R. & Mullins, D. (1978). Background characteristics of psychologists taking the Examination for Professional Practice in Psychology. *Texas Psychologist, 30*, 1–6.

Hayes, J. R. & Schreimer, R. (1978). Comparison of degree received and performance on the licensing examination. *Psychological Reports, 40*, 42–49.

Hoffman, P. J. & Aderet, A. (1984). *Empirical validity of the EPPP*. Los Altos, CA: COGITAN.

Kupfersmid, J. & Fiala, M. (1991). Comparison of EPPP scores among graduates of varying psychology programs. *American Psychologist, 46*, 534–535.

Professional Examination Service (1995a). *Manual of instructions for the Examination for Professional Practice in Psychology*. New York: Author

Professional Examination Service (1995b). *Annual report on the licensing examination program of the Association of State and Provincial Psychology Boards*. New York: Author.

Richman, S. (1982). *Final report to the American Association of State Psychology Boards on the role delineation study for the Examination for Professional Practice in Psychology*. New York: Professional Examination Service.

Rosen G. A. (1986). A perspective on predictive validity and licensure examinations. *Professional Practice of Psychology, 7*, 116–123.

Rosen, G. A. (1991). *Research digest: The Examination for Professional Practice in Psychology*. Montgomery, AL: Association of State and Provincial Psychology Boards.

Rosen, G. A., & Mirone, J. A. (1986). *Final report of the test specifications revision project*. New York: Professional Examination Service.

Rosenfeld, M., Shimberg, B., & Thornton, R. F. (1983). *Job analysis of licensed psychologists in the United States and Canada*. Princeton, NJ: Educational Testing Service.

Shrader, R. R. (1980). Validation studies on the Examination for Professional Practice in Psychology. *Professional Practice of Psychology, 1*, 23–30.

Smith, I. L (1984). *Content validity study of the AASPB item bank*. New York: Professional Examination Service.

Smith, I. L. & Greenberg, S. (1991). *Final report for the test specifications enhancement project*. New York: Professional Examination Service.

Smith, I. L., Hambleton, R. K., & Rosen, G. A. (1988). Content validity investigations of the Examination for Professional Practice in Psychology. *Professional Practice of Psychology, 9*, 43–80.

Terris, L. D. (1973). The national licensing examination. *Professional Psychology, 4*, 386–391.

Tipton, R. M., Elder, D. D., & Ritz, S. (1988). Performance on the national licensure examination as a function of program specialty and APA accreditation status. *Professional Practice of Psychology, 9*, 81–89.

Werner, E. (1989). *Analysis of the validity of the Examination for Professional Practice in Psychology*. Sacramento, CA: California Department of Consumer Affairs.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.13
An Assessment Tool in Credentialing: The Oral Examination for Licensure

DIANE S. HILL
Somerville and Company Inc., Denver, CO, USA

2.13.1 INTRODUCTION 151
2.13.2 WHY LICENSING BOARDS USE ORAL EXAMINATIONS 152
2.13.3 WHAT IS AN ORAL EXAMINATION? 152
2.13.4 TYPES OF ORAL EXAMINATIONS 153
2.13.5 RELIABILITY 155
2.13.6 VALIDITY 155
2.13.7 SCORING 157
2.13.8 PROS AND CONS OF ORAL EXAMINATIONS 157
2.13.9 FUTURE DIRECTIONS 159
2.13.10 REFERENCES 160

2.13.1 INTRODUCTION

The modern history of oral examinations used for credentialing purposes began in 1917 with the first administration of a medical specialty examination. A publication by the American Board of Medical Specialties (Mancall & Bashook, 1995) notes that physicians, like psychologists, have long argued over the viability of oral examinations as part of a credentialing process. At the time of writing, 34 of the 61 North American jurisdictions (Association of State and Provincial Psychology Boards [ASPPB], 1996) that license psychologists use an oral examination as part of a multimodal assessment strategy. Since the field of psychometrics is a domain within the discipline of psychology, we might hope and expect that psychology licensure examinations are beyond reproach. Most psychologists who have endured the rigors of a licensing examination would argue that there is room for improvement.

While this state of affairs is perhaps disappointing for the profession and its practitioners, it should not be all that surprising. Licensing boards have to balance an array of competing legal, technical, cultural, and resource considerations when developing their assessment strategies. Boards adopt a variety of procedures to carry out their gate-keeping function in the service of public protection. At the same time, state, provincial, and federal laws obligate boards not to restrict the rights of professionals to practice their trade. In some states and provinces the political and cultural climate favors an open free market that keeps restrictions to a minimum. In other jurisdictions there is acute awareness and concern over the possibility that incompetents and charlatans will visit harm upon the public. All psychology licensing boards believe in developing reliable and valid assessment procedures. Some have more resources than others with which to do so. Within the context of these often competing concerns and commitments, boards continually strive to carry out their gate-keeping responsibilities with ever-increasing skill and efficacy. Some form of oral examination will, in all likelihood, continue to be part of a multimodal assessment strategy in over one-half of the jurisdictions in North America.

The purpose of this chapter is to serve as an introduction to the use of oral examinations for licensure. It outlines the reasons why boards include the oral examination as a licensing requirement; defines three distinct prototypes; discusses the significant methodological issues; lists some of the pros and cons; and describes future directions. The writer has adapted this manuscript from an earlier manual sponsored by the ASPPB (Hill, Edwards, Crawford, & Warner; 1986). The authors wrote that publication as a guide to best practices for licensing boards that were currently using, or hoped to start using, an oral examination. The purpose of this chapter is to provide university faculty, graduate students, interns, and those preparing for licensure with an overview of what to expect in an oral examination. Those who read this chapter will also be in a better position to encourage their particular jurisdiction to use state-of-the-art assessment procedures when evaluating candidates for licensure.

2.13.2 WHY LICENSING BOARDS USE ORAL EXAMINATIONS

Boards use the oral examination to measure job-relevant knowledge, skills, and personal attributes that they cannot readily address using other assessment modalities. Jurisdictions that include the oral examination do so in combination with other assessment instruments and sources of information as part of a multimodal licensing strategy. The oral examination is particularly useful for boards that wish to observe and evaluate a candidate's understanding of the limits of their own competence. Examiners also value the oral format for the opportunity it affords to observe and rate a candidate's ongoing clinical reasoning. Some oral procedures allow examiners to evaluate the degree of consistency between the candidate's stated theoretical framework, and the application of that framework to clinical issues. Finally, some boards use the oral examination to evaluate the impact of a candidate's personal, emotional, and social functioning on their capacity to provide psychological services.

2.13.3 WHAT IS AN ORAL EXAMINATION?

An oral examination involves a process of verbal interaction between a panel of trained examiners and a candidate for a psychology license. It may be general in content or focus on the candidate's area of specialization, depending on the predetermined objectives of the examination. It does not stand alone as an assessment tool. Rather, it provides information that is supplementary to that derived from a written examination such as the Examination for the Professional Practice of Psychology (EPPP). Examiners assign scores on an oral examination based on their ratings of the candidate's verbal responses. They base their ratings on predetermined criteria.

Candidates in an oral examination respond orally to a series of stimuli, in face-to-face interaction with one or more examiners. While the response format tends to be common to all oral examinations, the manner in which different boards choose to present stimuli may vary. If an applicant applies for licensure in a jurisdiction that uses an oral examination, they can expect the examination to feature one or more of the following procedures.

(i) *Oral questions*. This is the simplest and perhaps most common procedure. Each examiner asks the candidate a series of questions. Boards that use oral examinations differ in the extent to which they favor more or less structure. Some predetermine all questions; others use a partially predetermined set of questions. Some boards prefer to develop questions during the examination itself that are consistent with prescribed themes.

(ii) *Written questions*. Examiners may present prepared questions to the candidate in writing, though the candidate responds orally. This procedure has the advantage of assuring consistency across candidates.

(iii) *Work sample*. At a stated time prior to the examination, the board asks each candidate to submit materials from their caseload. Examiners then formulate questions based on these materials.

(iv) *Case presentation*. The candidate describes one or more cases from their own professional experience.

(v) *Standardized vignette*. The candidate observes a scenario illustrating a relevant situation from the practice of psychology. It may depict a sample of a client's behavior during psychotherapy, or another relevant scenario typically encountered by a licensed psychologist. The candidate may encounter the vignette in written, audio, or video form.

(vi) *Live simulation*. Performers may pose as clients exhibiting various behaviors. The candidate may conduct a clinical interview with "the patient" while the examiners observe. Questions from the examiners follow.

2.13.4 TYPES OF ORAL EXAMINATIONS

In developing their examination procedure, all licensing boards face decisions that require them to balance their concern for protecting the public against the importance of preserving fair trade practices. They must also weigh preferences for flexibility and individual tailoring of an examination against requirements for consistency and standardization. As jurisdictions have addressed these issues, they have adopted procedures that fall into one of three broad categories: (i) the structured oral examination, (ii) the semistructured examination, and (iii) the unstructured examination. While this division into categories is useful, it is also arbitrary, because the degree of structure varies along a continuum. However, in order to develop a clearer picture of the possible approaches to the oral examination, it is useful to compare the three prototypes along the following dimensions:

(i) *Content*. The structured examination consists of predetermined specific questions or standardized stimuli (e.g., scenarios), with specified ratings or scoring criteria for answers. For example, if the scoring system employs a five-point rating scale, the examiners have in front of them sample answers worth five points, three points, and one point, to guide their ratings. There is a set examination format and sequence that does not permit deviations.

The semistructured examination comprises a set of predetermined specific situations that are applicable to groups of candidates (e.g., candidates in clinical psychology, counseling psychology, industrial psychology). Boards also prepare initial questions ahead of time. During the examination, however, the examiners may deviate from these situations and questions, or seek clarification of the candidate's responses.

Boards that use unstructured examinations usually tailor them for each candidate. They prepare examination themes, rather than individual questions, for each candidate, taking into account, for example, his or her academic background, work experience, or theoretical orientation.

(ii) *Content consistency across candidates*. Consistency across candidates is highest in the structured examination and lowest in the unstructured examination. It can be high in the semistructured examination, depending on how the board develops and uses situations and questions. It is highest in the structured examination because every candidate must respond to the same situations and questions, regardless of his or her background or theoretical orientation.

(iii) *Tailoring of questions to the candidate*. The structured examination, does not allow for spontaneous or tailored questions, so that all candidates receive the identical examination. In the semistructured examination, examiners develop tailor-made questions based on the examination blueprint. These questions frequently complement an initial uniform set of questions asked of all candidates. The unstructured examination consists of questions that are specific to each candidate's background, experience, and theoretical approach. Examiners specifically formulate questions for that person around preselected common examination themes.

(iv) *Formulation of questions in view of prior responses*. The structured examination does not allow for questions that derive from a candidate's prior responses. Those who train the examiners instruct them to remember that the quality of a candidate's responses to different questions may vary widely. They make every effort not to allow themselves to be influenced by prior answers when recording their ratings. In semistructured examinations, follow-up questions are appropriate and often essential. The opportunity to follow up on a candidate's response is a significant strength of the unstructured approach. Candidate responses typically lead to further probing from the examiner.

(v) *Scoring*. The structured examination includes specific rating scales for each item. Authors of the examination define anchor points on the rating scale using sample answers for each question. The semistructured examination uses a series of predetermined scoring categories, each with specific rating scales. The scoring categories are broader than the dimensions used for the structured examination. Within each scoring category, there may be both mandatory and suggested areas of questioning. Examiners record ratings independently for each scoring category. In the unstructured examination, examiners rate the candidate's performance on the basis of relatively broad, preselected dimensions (for example, the candidate's awareness of his or her own limits of competence).

(vi) *Examination development costs*. Costs tend to be the highest for the structured examination, relatively high for the semistructured, and lowest for the unstructured examination.

(vii) *Duration*. The structured examination is potentially the shortest, since authors of the examination have preselected the questions and specific scoring criteria, and examiners do not engage in spontaneous interaction with the candidate. The semistructured examination may be of intermediate duration, because it permits unplanned interactions but does not require them. The unstructured examination is potentially the longest, as it is most likely to include spontaneous questions and the examiners may delve into areas suggested by the candidate's previous answers.

(viii) *Decision alternatives*. Theoretically, the decision alternatives are not inherent to any examination format. Each of the three formats may lead to any of the decision alternatives described below. However, the following paragraphs describe the decision alternatives typically associated with each of the three oral examination formats.

The structured examination typically yields a pass or fail decision from the panel of examiners. A prototypical semistructured examination might yield one of the following decisions: pass, fail, or re-examine. Boards reserve this last recommendation for split or unclear decisions. In such cases, they often conduct the re-examination immediately using another panel of examiners.

The unstructured examination yields a recommendation of pass, or defer judgment. In other words, examiners recommend that candidates who do not pass sit for re-examination at some future date. The panel's recommendation stems from the combined overall ratings of the individual examiners with or without preliminary discussion of the individual ratings among examiners.

Regardless of the format of the examination, the examining panel communicates its recommendation to the board following the examination. In all cases, the decision authority rests with the board. The outcome of any oral examination, therefore, is a recommendation to the board by the examining panel.

(ix) *Postexamination feedback to candidates*. In theory, the nature of the postexamination feedback provided to the candidate is independent of the examination format. In practice, somewhat different types of feedback accompany prototypical unstructured, semistructured and structured formats. Following a prototypical structured examination, the candidate typically receives written feedback as numerical subscores on particular items. Because proponents of structured, examinations favor treating all candidates equally, this strategy does not usually allow for qualitative feedback concerning performance during the examination.

After a prototypical semistructured oral examination, candidates may receive feedback either orally or in writing. Boards base their feedback on generic scoring categories (each of which subsumes a number of items or questions), and it is likely to be qualitative rather than numerical. The panel provides feedback after it makes its recommendation. In the unstructured situation, the chair of the examining panel frequently provides verbal feedback to the candidate. The feedback typically consists of qualitative statements about the candidate's strengths and weaknesses, and may include remedial suggestions. Examiners do not provide feedback until they have finalized their recommendation.

(x) *Ratings of relevant personal characteristics of a candidate*. Examining panels may observe and rate personal characteristics of a candidate that are relevant to the practice of psychology. Typically, they use separate rating scales and criteria, even when they are not the object of direct questioning. In the unstructured and semistructured formats, information that the board receives on the candidate's behavior during the examination may warrant direct questioning in this area. Jurisdictions vary in the degree to which they believe that rating of relevant personal characteristics is an appropriate activity for a licensing examination. Those who favor ratings of this type argue that such evaluations are the responsibility of the board and essential for public protection. Those who oppose such ratings argue that lack of objectivity in this type of measurement leaves too much room for examiner bias.

The three types of oral examinations outlined here are prototypes or illustrations rather than descriptions of actual examinations that might be encountered on the road to licensure. In reality, the examination procedures adopted by a particular jurisdiction are a function of historical and cultural precedent, financial resources, and the knowledge, skill, and experience available to apply to the task. Each approach has its advantages. High structure increases ease of computing quantitative psychometric properties, but reduces flexibility, the ability to customize, and spontaneous interaction.

2.13.5 RELIABILITY

"Reliability" refers to consistency of measurement. Indices of reliability compare the proportion of error variance with that of true variance. There are a number of ways of estimating reliability, and examining boards typically apply one or more to their oral examination. Standard 11.3 of the revised *Technical standards for educational and psychological testing* (American Psychological Association [APA], 1985) states that boards should provide estimates of the reliability of licensure or certification decisions.

One study (Novy, Kopel, & Swank, 1996) examined two sources of error variance in the June 1992 administration of the Texas oral examination. This board uses a moderately interactive oral examination based on vignettes that are developed for four different subdisciplines of psychology: clinical, counseling, school, and industrial–organizational. These investigators found that indices of inter-rater reliability were quite high. Only five of the 22 coefficients of agreement calculated fell below 0.85. The authors noted, however, that indices of equivalence across vignettes in the same discipline were less than optimal. Generalizability coefficients that took into account error variance attributable to both raters and vignettes were as follows: clinical, 0.70; counseling, 0.55; and school, 0.65. This small but suggestive study points to a need to attend to equivalency measures in addition to measures of inter-rater reliability when developing oral examinations.

The ASPPB and experts in the field of measurement have, however, taken the position that inter-rater reliability is the most appropriate form of reliability when evaluating oral examinations. The principal purpose in all licensing examinations, including orals, is to arrive at a pass or fail decision concerning each candidate. For this reason, reliability procedures typically focus on the extent to which the examiners reach a final recommendation. This ultimate objective for the licensure process has been used to argue that efforts to establish reliability should focus on examiner's ratings rather than on candidate's responses. The *Technical standards for educational and psychological testing* (APA, 1985, p. 5) define inter-rater reliability as the "consistency of judgments made about people or objects among raters or sets of raters." It is particularly important to assess inter-rater agreement when observational data involve subtle discriminations, as is the case in oral examinations.

Differing perceptions on the part of examiners is the source of error variance that has received the most attention from those interested in oral examinations. To the extent that examination authors and administrators can minimize those disagreements, reliability will increase. The following measures will help to minimize examiner error and increase reliability:

(i) It is important that all examiners are clear about the purpose of the examination. Written guidelines on purposes and procedures are essential.

(ii) Examiners should obtain an outline of appropriate themes, vignettes, or items to ensure that the major elements of the examination are clear and receive appropriate coverage.

(iii) Criteria for scoring should be explicit.

(iv) Two or more examiners should be assigned to each candidate.

(v) Board members need to take responsibility for ensuring that examiners receive training.

(vi) Each examining team should have at least one member who has previous examining experience.

(vii) Board members should monitor the examination sessions by either observing selected sessions or reviewing tape recordings.

(viii) Candidates need to have adequate time to respond to the questions and situations posed in the examination.

2.13.6 VALIDITY

The validity of an assessment instrument refers to the extent to which the test measures what it purports to measure. It tells us what we can infer from test scores. Procedures for determining the validity of a test or examination explore the relationships between performance on the test and other independently observable behaviors. Psychometricians who develop assessment procedures have an array of validation techniques from which to choose. These fall into three distinct but related categories: content-validation procedures, indices of criterion-related validity, and construct-validation strategies. Numerous texts on tests and measurement describe these different methods of assessing the validity of a test. The focus of this discussion is on the content-validation procedures that North American boards are using to answer two questions: (i) Does the test cover a representative sample of the skills and knowledge required for entrance to the independent practice of psychology? (ii) Is performance on the licensing examination reasonably free from the influence of irrelevant variables?

Studies published in the early 1980s (Kane, 1982; Shimberg, 1981) argued convincingly that content validity, rather than criterion-related validity is appropriate for licensing examinations. The validation strategies that apply to personnel selection and those that pertain to licensing are different. Employers often validate personnel selection tests that identify the best among qualified candidates using criterion-related approach. Licensing examinations, in contrast, differentiate between candidates who possess the knowledge, skills, and personal attributes necessary for independent practice, and those who do not. Licensing examinations are most appropriately validated by obtaining evidence of content validity.

The steps that have been taken to create a valid, state-of-the-art national multiple choice examination (the EPPP) can prove helpful to boards interested in developing reliable and valid oral examinations. The process for establishing content validity is a well established one. It is also very costly. It consists of six basic steps: (i) conducting a job analysis, (ii) developing test specifications, (iii) developing a content outline, (iv) constructing an examination blueprint, (v) writing the oral examination proper, and (vi) cross-validating the judgments of the oral examination writers. Other contributors to this volume describe in detail the steps required for developing valid credentialing examinations (see Chapters 10–12 and 14, this volume). The procedure will receive a brief review here to give the reader an idea of the time, money, and expertise necessary to create reliable and valid oral examinations.

(i) *The job analysis*. Over the past 15 years, the ASPPB has commissioned three studies that form the basis for the development of valid credentialing examinations. The Professional Examination Service (PES) has conducted two of them (Greenberg, Smith, & Muenzen, 1995; Richman, 1982), and the Educational Testing Service has conducted one (Rosenfeld, Shimberg, & Thornton, 1983). These job analyses are the source of valuable data concerning what psychologists do on the job and the knowledge and skills necessary for the practice of psychology. Boards committed to developing state-of-the art oral examinations should review these studies as they develop their test specifications. The purpose of the job analysis is to provide a comprehensive description of the roles, responsibilities, and knowledge required for the independent practice of psychology.

(ii) *The table of test specifications*. The PES has also developed two sets of test specifications based on these job analyses. Because these test specifications are based on comprehensive studies of what psychologists do on the job, they also provide an appropriate basis for the development and content validation of other licensing examinations in psychology, including the oral examination. To the extent that jurisdictions view the EPPP and oral examinations as instruments intended to provide complementary but relevant information about candidates for licensing, the content outlines for both instruments would ideally originate from a common table of test specifications.

Once the job analysis and test specifications are complete, the next step is to develop a content outline. Because there is no coordinated effort among jurisdictions to develop oral examinations, each board creates their own. Boards vary considerably in the extent to which they have the resources available to complete steps (iii)–(vi).

(iii) *Developing a content outline*. In order to complete this step, boards need to determine which statements from the table of test specifications they can most appropriately measure using an oral format. The content of the oral examination should complement rather than repeat what other instruments or sources of information already address. For example, one study (Richman, 1982) suggests that the multiple-choice format of the EPPP does not particularly suit the measurement of complex inferences and deductions in decision-making. Boards that use oral examinations would therefore do well to consider including this skill in the content outline of the oral examination.

(iv) *The development of an examination blueprint*. The blueprint is a statement of what any one examination will cover. A given examination typically covers a sample or subset of topics included in the content outline. The sampling procedure will vary depending on the board's purpose in using an oral examination. For example, if a jurisdiction is committed to a generic oral examination, it may first identify statements from the content outline that merit inclusion in every oral examination. Next it may identify the remaining statements to be included through a stratified random sampling procedure. The number of statements selected is a function of practical considerations such as intended format and length of the examination.

Some state boards (e.g., Iowa) use the oral to examine a candidate's knowledge of laws relevant to the practice of psychology. In this case, knowledge and skill statements that relate to jurisprudence are the only statements sampled. Other boards have various rationales for tailoring their examinations to individuals or groups of individuals. The Louisiana board members, for example, believe that it is important to identify for the public a psychologist's particular area of specialization. In this case, the board primarily selects knowledge and skill statements that relate to the candidate's specialty.

(v) *Writing the oral examination.* The next step is to write the oral examination for any given administration. This step consists of writing the themes, scenarios, or questions that correspond to the final examination blueprint. What examination authors write depends both on the nature of the statements that comprise the blueprint and on the format of the oral examination (structured, semistructured, or unstructured). This step includes not only the writing of items, situations, or themes, but also the development of appropriate rating or scoring criteria. Because the writing of items, scenarios, or themes is a judgmental rather than an empirical activity, it should involve licensed psychologists in the process. Ideally, each panel of writers would include members of groups (e.g., ethnic minorities) that the jurisdiction has reason to believe might be at risk of being given an unfair or biased oral examination.

(vi) *Cross-validation of examination writers' judgments.* The task of developing items, situations, or themes that reflect the final blueprint is one which requires judgment. Therefore, a panel of independent experts should determine whether the oral examination as written is an accurate reflection of the blueprint. The board should select this panel on the basis of comparable criteria to those used for the selection of examination writers. Particularly when the review panel includes members of relevant minority groups, this step provides an additional opportunity to ensure that the examination does not include inappropriate biases or potentially discriminatory items.

The procedures described above are well recognized in the field of psychometrics. There is nothing particularly mysterious or technologically difficult about the process of creating content-valid oral examinations. The process of creating a reliable, valid, and multimodal examination strategy that includes an oral examination is not an especially arcane one. However, it is costly in terms of time, money, and expertise. Most boards simply do not have the resources to adopt all the above procedures vigorously. Many boards use some or most of them. However, until there is greater coordination of effort and resources among the psychology licensing boards in North America, many oral examinations will remain less than state-of-the-art. This topic is discussed at greater length in Section 2.13.9.

2.13.7 SCORING

Boards should score all oral examinations using well-defined, pre-established rating criteria. Test authors should provide the examiners with rating sheets that list the dimensions on which they are to rate the candidate's performance. Rating sheets need to include rating scales with verbal descriptors of the points on the scale. The dimensions on which examiners rate the candidate's performance vary depending on the format of the examination. For example, examiners rate performance on a structured examination according to specific rating criteria that they apply to predetermined questions. Raters on an unstructured examination assign ratings along a limited number of relatively broad categories. Examiners may rate performance on a semistructured examination according to both broad and specific criteria, as this approach incorporates elements of the structured and unstructured examinations.

Boards need to decide in advance how to combine these ratings into a judgment about whether or not to license the candidate. For example, a jurisdiction that uses five-point Likert scales may wish to add the scale scores so as to derive a total score, and establish a specified cut-off or passing score. Another jurisdiction may assign more weight to some dimensions than to others. Many combining formulas are possible. The essential point is that the formula is pre-established, with a defensible cut-off point. A cut-off point is defensible to the extent that it demonstrably differentiates between candidates who are ready for independent practice and those who are not.

Finally, for the scoring system employed by any jurisdiction to be defensible, boards need to train examiners in the use of the rating criteria and rating sheets before the examinations begin.

2.13.8 PROS AND CONS OF ORAL EXAMINATIONS

Boards favor oral examinations because they believe that the procedure measures factors that contribute unique information to a certification decision. At least two studies, one in our own profession and one in medicine, suggest that this is the case. PES conducted the first study in the mid-1980s (Rosen & Mirone, 1986). In this investigation, the researchers asked a panel of seven raters to review the 1986 test specifications and to indicate which assessment technique (multiple choice, essay, structured oral, interactive oral, or simulation) most appropriately applies when measuring the knowledge application or skill associated with each statement. The main finding of this pilot project was that the multiple-choice format was less well-suited than others, including the oral examinations, to the measurement of complex inferences and deductions, especially as they relate to decision-making.

More specifically, this small but suggestive study indicates that the oral examination is more appropriate than the multiple-choice format for assessing many of the skills required to define and diagnose problems. Raters favored the oral examination as a tool for measuring a candidate's ability to apply appropriate knowledge to the task of selecting, developing, and interpreting certain diagnostic tests. They also saw the oral examination as superior to a multiple-choice procedure in measuring the candidate's ability to use certain skills, such as stress-management techniques, and apply knowledge of learning, individual differences, or motivation in designing, implementing, and assessing interventions. Finally, raters saw the oral examination as the tool of choice for measuring the ability to apply some kinds of ethical and legal knowledge to the conduct of scientific and professional activities.

More recently, the American Board of Medical Specialties sponsored a symposium on using oral examinations to assess clinical learning. The proceeds from that meeting are available (Mancall & Bashook, 1995) and make informative reading for anyone interested in the use of oral examinations for credentialing purposes. All the papers included in the compendium support the thesis that some form of oral examination is superior to other techniques for assessing a professional's ability to reason and make judgments. One author (Reinhart, 1995) describes a predictive validity study that asked the question: Does the oral examination predict actual performance in a diplomate's own emergency department cases? The investigators developed a standardized method to assess emergency physicians' performances based on their own patient charts. Predictive-validity measures indicated that the correlation between oral examination performance and physician performance on the job was 0.45. Performance on the oral examination was more strongly associated with performance in the emergency department than performance in the written examination. The authors concluded that the oral examination does, in fact, predict physician performance, and that it measures some aspects of performance not assessed by the written examination.

A further study suggested that the oral examination added unique information to the certification decision in the arena of clinical competence, composure when managing a patient, ability to solicit clinical data, and use of consultants. Taken together, these two studies offer valuable insight about where to look for the unique contributions of the oral examination in our own profession. They also suggest some methodological possibilities for investigating the validity of an oral examination over and beyond content validity.

In addition to these empirical studies demonstrating the advantages of the oral format, there are a number of widely held beliefs about the value of the oral examination that await confirmation using more formal research techniques. The two basic questions that studies should address are: (i) Is a particular skill or knowledge demonstrably related to a psychologist's ability to practice independently at an entry level? (ii) Is the skill or knowledge best measured in an oral format?

The list of knowledge and skills that require further investigation includes: (i) interactive skills relevant to the practice of psychology; (ii) consistency between a candidate's theoretical framework and the application of this framework to clinical problems; (iii) the candidate's ability to apply the knowledge and skills of the profession to concrete situations; (iv) the candidate's personal functioning in relation to his or her ability to provide sound psychological services; (v) the candidate's ability to reason and make decisions; and (vi) the candidate's knowledge of the limitation of his or her skills and competence.

Other advantages of the oral examination pertain to one or more of the distinct prototypes described. It is also the case that these advantages are more controversial, and while favored by some may be viewed as problematic by others.

In its unstructured or semistructured formats, the oral examination has the advantage of *flexibility*. Not all jurisdictions, however, would view flexibility as advantageous. Jurisdiction that stress the need to examine all candidates in a consistent manner are likely to consider flexibility a disadvantage. Jurisdictions that tailor parts or all of the examination to individual candidates value the potential flexibility of the oral format. The flexibility of unstructured and semistructured oral examinations gives the following advantages:

(i) The candidate may seek clarification of questions from the examiners.

(ii) The examiners may seek clarification or expansion of answers from the candidate.

(iii) Examination themes or questions can be tailored to the specific background and specific career interests of individual candidates.

(iv) Unusual or seemingly inappropriate responses or comments from a candidate may be followed up by the examiners even when not foreseen before the examination.

(v) The examination may include informal comments from the examiners. For example, the examiners may include certain remarks intended to welcome the candidate to the profession of psychology.

In its structured format, the oral examination offers the potential advantage of *consistency* of measurement across candidates. Jurisdictions that favor giving all candidates the opportunity to demonstrate their competence through the same assessment instrument(s) will probably favor a structured format. This format has the following advantages:

(i) There are specific scoring criteria for all answers. Examination authors develop sample answers that define points on a rating scale in advance. This procedure may reduce error variance in ratings and enhance their objectivity.

(ii) Examiner training may be facilitated by the fact that all examiners learn to rate responses to exactly the same questions.

(iii) Because all candidates are given the identical opportunity to demonstrate competence, there may be less likelihood of charges of unequal or discriminatory treatment by failed candidates.

(iv) Although most jurisdictions have adopted a generic approach to licensing, many offer one or more examinations in different specialty areas. Such jurisdictions, if they wish to employ the structured format, would use the same examination for all candidates in any one specialty area (e.g., clinical psychology), but different examinations across specialty areas.

While the oral examination format offers many advantages, including those described above, it is not without potential measurement and practical shortcomings. The following are important potential limitations, of which users should be aware:

(i) *Measurement limitations or shortcomings*:

(a) Reliability may be below acceptable standards unless well-developed rating scales are used by properly trained raters, and boards calculate and continually work to improve the equivalency coefficients for similar forms of the oral examination.

(b) Allegations of subjectivity or examiner bias are probable unless special precautions are taken to ensure that the examinations are developed and administered in a fair and impartial manner.

(c) Failure to establish and maintain rapport may lead to contamination of examination results if this omission causes undue stress for the candidate.

(ii) *Practical limitations or shortcomings*.

(a) Many knowledge and skill items that are relevant to the practice of psychology are more validly and/or reliably measured by other examination modalities than the oral (e.g., by written examinations such as the EPPP, or by written simulations). Licensing boards should therefore use the oral examination in conjunction with other instruments and sources of data.

(b) The oral examination is human resource intensive, costly, time consuming, and relatively difficult to develop, administer, and score.

The manual published by ASPPB (Hill et al., 1986) argues that the oral examination is the preferred modality for assessing integration and problem solving skills, interpersonal skills, theoretical orientation, personal attributes, and verbal skills. The argument rests on the notion that an examination fulfills its potential only when boards use proper construction, administration and scoring procedures. The authors also note that the national multiple-choice examination (EPPP) is complementary to the oral format in measuring the knowledge and many of the skills required for independent practice. Boards may use the EPPP and the oral examination for complementary purposes, thereby enhancing the validity of the overall assessment process.

2.13.9 FUTURE DIRECTIONS

At present, oral examinations for licensing serve a valuable purpose, but they are not as useful as they might be. The major problems associated with the oral examination derive from the limited resources available to individual boards for their development.

Although the use of oral examinations by licensing boards has become increasingly sophisticated, and will no doubt continue to do so, problem areas remain including the following:

(i) *Lack of reciprocity*. Candidates who have passed an oral examination in one jurisdiction and wish to be licensed in another frequently have to take a new oral examination. Boards could facilitate and enhance the quality and cost-effectiveness of their examinations if jurisdictions with common oral examination objectives cooperated in their development.

(ii) *Duplication of efforts*. At present, each jurisdiction develops its own oral examination. This process is necessarily wasteful of human and financial resources. Boards could reduce such duplication of efforts by jointly developing oral examinations with other jurisdictions that share their examination objectives.

(iii) *Limitations imposed by the licensing acts*. The requirements of psychology licensing acts vary across jurisdictions. In some instances, the act places explicit limitations on the number or types of licensing examinations that a board may use. In most instances, however, common oral examinations could be used across jurisdictions, provided that such examinations are developed through a cooperative effort.

(iv) *Lack of item pools.* At present, few jurisdictions have access to adequate pools of oral examination questions. Such pools are important to the effective use of structured and semistructured oral examinations. Sharing of item pools among jurisdictions with common oral examination objectives and formats represents a potential solution to this problem.

When all is said and done, the fact remains that individual jurisdictions are still responsible for determining their candidate's competence to practice. Solutions to these problems, therefore, will depend on voluntary participation by jurisdictions. There are legitimate obstacles to such cooperation and sharing at present. Among these, the need to maintain the confidentiality of the examination content and specific rating criteria is foremost. Cooperation and sharing are desirable, however, and should be increasingly possible in the not too distant future.

Positive strides in this direction are already in progress. Six North American jurisdictions (Iowa, Kentucky, Manitoba, Missouri, Oklahoma, and Ontario) have signed the North American Reciprocity Agreement. Three boards (Kentucky, Missouri, and Oklahoma) have formed a second consortium and signed the ASPPB Reciprocity Agreement. In order to participate in either of these agreements, a jurisdiction must use an oral examination when it licenses candidates. The purpose of these agreements is to facilitate mobility for psychologists residing in member jurisdictions. Psychologists who wish to move from Iowa to Kentucky, for example, are not required to retake the EPPP or an oral examination. They will only be tested on laws relating to the practice of psychology in their new jurisdiction.

If the trend toward forming agreements designed to remove barriers to mobility continues, jurisdictions may begin to see the wisdom of pooling resources to create more reliable and valid oral examinations. The hope for the near future is that boards with similar objectives will find ways to cooperate in executing the costly and time-consuming steps required to upgrade the reliability and validity of their oral examinations. A longer term solution may well be for the ASPPB to commission an appropriate vendor to develop one or more oral examinations that would be purchased and administered by individual jurisdictions.

2.13.10 REFERENCES

American Psychological Association (1985). *Technical standards for educational and psychological testing.* Washington, DC: Author.

Association of State and Provincial Psychology Boards (1996). *Handbook of licensing and certification requirements for psychologists in North America.* Montgomery, AL: Author.

Greenberg, S., Smith, I. L., & Muenzen, P. M. (1995). *Executive summary: Study of the practice of licensed psychologists in the United States and Canada.* Montgomery, AL: Association of State and Provincial Psychology Boards.

Hill, D. S., Edwards, H. P., Crawford, W., & Warner, J. (1986). *Oral examinations for licensing or certification: A manual for psychology boards.* Montgomery, AL: American Association of State Psychology Boards.

Kane, M. T. (1982). The validity of licensure examinations. *American Psychologist, 37,* 911–918.

Mancall, E. L. (Ed.) (1990). *The oral examination: An historical perspective.* Evanston, IL: American Board of Medical Specialties.

Mancall, E. L., & Bashook, P. G. (Eds.) (1995). *Assessing clinical reasoning: The oral examination and alternative methods.* Evanston, IL: American Board of Medical Specialties.

Novy, D. M., Kopel, K. F., & Swank, P. R. (1996). Psychometrics of oral examinations for psychology licensure: The Texas examination as an example. *Professional Psychology: Research and Practice, 27,* 415–417.

Reinhart, M. A. (1995). Advantages to using the oral examination. In E. L. Mancall & P. G. Bashook (Eds.), *Assessing clinical reasoning: The oral examination and alternative methods* (pp. 31–40). Evanston, IL: American Board of Medical Specialties.

Richman, S. (1982). *Final report to the American Association of State Psychology Boards on the role delineation study for the examination for professional practice in psychology.* New York: Professional Examination Service.

Rosen, G. A., & Mirone, J. A. (1986). *Final report of the test specification project for the Examination for Professional Practice in Psychology.* New York: Professional Examination Service.

Rosenfeld, M., Shimberg, B. S., & Thornton, R. (1983). *Job analysis of licensed psychologists in the United States and Canada: A study of responsibilities and requirements.* Princeton, NJ: Educational Testing Service.

Shimberg, B. S. (1981). Testing for licensure and certification. *American Psychologist, 36,* 1138–1146.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.14

Credentialing Issues, Current and Future: United States and Canada

STEPHEN T. DeMERS
University of Kentucky, Lexington, KY, USA

2.14.1 INTRODUCTION 161

2.14.2 HISTORICAL PERSPECTIVE ON PSYCHOLOGY LICENSURE 161

2.14.3 MISSION OF PSYCHOLOGY REGULATORY BOARDS 162
  2.14.3.1 Credentialing Qualified Practitioners 163
  2.14.3.2 Investigating Complaints 164

2.14.4 CURRENT AND FUTURE ISSUES 165
  2.14.4.1 Educational Standards for Independent Practice 165
  2.14.4.2 Scope of Practice Issues 166
  2.14.4.3 Mobility Across Jurisdictions 167

2.14.5 CONCLUSIONS 168

2.14.6 REFERENCES 168

2.14.1 INTRODUCTION

This chapter is intended to increase awareness of both the history and the current and future concerns relative to the professional regulation of psychologists in the USA and Canada. The mechanisms of professional regulation include the development of standards for professional preparation, the promulgation of codes of professional conduct and ethics, the adoption of credentialing mechanisms and, finally, the implementation of disciplinary procedures. In the USA and Canada, professional regulation is accomplished through an interplay between professional societies, such as the American (APA) and Canadian (CPA) Psychological Associations, and legislatively mandated regulatory bodies such as psychology licensing boards. This chapter focuses on the role of psychology licensing boards in the regulation of the profession and the current and future issues that licensing boards, licensed psychologists and the public face with respect to the credentialing and discipline of psychologists.

For purposes of brevity, states and provinces are referred to throughout the text as jurisdictions. The term “regulatory body” is used to refer to the legally appointed, elected, or otherwise constituted agency within a jurisdiction to regulate the profession of psychology. In US jurisdictions, the regulatory body is often called the board of psychology, whereas in Canadian jurisdictions, the regulatory body in psychology may be referred to as a register, college, association, corporation, or board.

2.14.2 HISTORICAL PERSPECTIVE ON PSYCHOLOGY LICENSURE

According to *Webster's New World Dictionary* (1986), a profession is described as a calling or vocation involving specialized knowledge, long and intensive preparation, a commitment to public service, and a commitment to high standards of performance and conduct. Numerous authors (Gross, 1978; Newton, 1988; Weissman, 1984), have described the hallmarks of a profession. These include the presence of professional organizations that standardize and advance the development of the field, recognized academic programs that offer degrees documenting the specialized training necessary to competently practice the profession, codes of ethics, and a process whereby entry to the profession is controlled and monitored.

Members of a profession are usually identifiable by a license or certificate issued by a legislatively mandated professional regulatory body. Such regulatory bodies are typically empowered by state or provincial statutes to identify by training, experience, and examination the competent members of the profession, and to discipline incompetent or negligent practitioners in order to protect the public from harm. One of the hallmarks of psychology's emergence as a recognized profession was the passage of psychology licensing laws in states and provinces in the USA and Canada.

Another aspect of regulation is the process of disciplining members of the profession who violate the public trust. Professional associations and regulatory bodies adopt and promulgate ethical standards and codes of professional conduct in psychology (APA, 1992; Association of State and Provincial Psychology Boards [ASPPB], 1991; CPA, 1991) Both professional associations and psychology regulatory boards adjudicate complaints and impose sanctions on those association members and licensed psychologists found guilty of misconduct. The rationale for this professional self-regulation is based on the belief that the average citizen is not competent to judge appropriate professional behavior, and thus professionals should be regulated by their peers.

Although interest in regulation of professions can be traced to ancient times, professional licensing in the USA is a fairly recent phenomenon. A more complete description of the history of psychology regulation is given by Sinclair, Simon, and Pettifor, (1996). In 1888, a landmark US Supreme Court decision (*Dent v. West Virginia*, 1889) legitimized the policing powers of state government over the professions when these powers are used to protect the health and welfare of the public. By 1912, nearly all states and provinces had medical practice acts. The first psychology licensing act was passed in Connecticut in 1945. As for most early psychology laws, the Connecticut law protected the title of "psychologist" but did not preclude others (particularly physicians) from using psychological methods. Ontario passed the first licensing law for psychology in Canada in 1960. By 1977, all states had passed some form of psychology regulation and, in 1990, Prince Edward Island was the last Canadian province to enact a psychology licensing law. Currently there are 61 psychology regulatory boards throughout the USA and Canada including the 50 states, nine provinces, the District of Columbia, and Guam (ASPPB, 1997).

As more states followed Connecticut's lead, there was a growing need for a common standard of training and a common examination to aid the identification of competent psychological practitioners. This need for a common examination was the impetus for the creation of the American Association of State Psychology Boards in 1961. (The organization's name was changed to the Association of State and Provincial Psychology Boards in 1992.) The organization started with representatives from 21 member jurisdictions with licensing laws for psychologists.

The organization's early focus was on the development of a common examination. In 1964, the Association administered the first version (i.e., Form 1) of the Examination for Professional Practice of Psychology (EPPP) to one candidate in Virginia. By the time Form 2 was completed in 1967, it was administered to 122 candidates from 15 states. The EPPP has always been a multiple choice test with items written by an examination committee of content experts across various specialties in psychology. The original EPPP had 150 items, however, this number gradually increased to 200 items where it has remained since 1969. The EPPP has undergone numerous validation studies. Currently, the EPPP is administered twice each year to approximately 5700 candidates in 61 member jurisdictions across the USA and Canada. It is available in both French and English language versions.

The Association has expanded its mission over the years, beyond the examination program to include the development of model licensing laws and an ASPPB Code of Conduct (ASPPB, 1991), development of a disciplinary data system, and an educational program including an annual convention and midwinter meeting, newsletters, publications and other training materials, and programs for members of regulatory bodies.

2.14.3 MISSION OF PSYCHOLOGY REGULATORY BOARDS

As noted earlier, the main purpose of psychology licensing laws is to protect the public from unqualified, incompetent, impaired, or unethical individuals who seek to provide psychological services to the public. Consumer groups and many psychologists themselves misperceive the mission of psychology regulatory boards as focused on protecting the profession. However, the mission, and the activities pursued by boards to accomplish that mission, are clearly focused on protection of the public. Psychology regulatory boards accomplish their mission through two main activities, namely, issuing credentials to qualified and competent practitioners, and investigating complaints against practitioners who are already credentialed. Each of these main activities involves several aspects that are described in more detail in the sections below.

2.14.3.1 Credentialing Qualified Practitioners

Psychology licensing acts are designed to protect the public by identifying and credentialing only those individuals who are qualified and competent to practice psychology safely and ethically. Psychology regulatory boards accomplish this goal by establishing strict standards for educational preparation and supervised experience, and by administering examinations to measure the professional knowledge and skills of the applicant for licensure. Each jurisdiction establishes specific educational requirements for licensure as a psychologist including degree level and type (e.g., Ph.D., Ed.D., Psy.D., M.S., M.A., M.Ed.), program of study (e.g., must be a doctoral degree in psychology), institutional accreditation (e.g., must be regionally accredited in the USA or recognized by Royal Charter in Canada), and specific course content (e.g., program must include courses in professional ethics, research methods, diagnosis and assessment, treatment interventions, etc.).

Similarly, each jurisdiction adopts specific requirements for supervised professional experiences prior to licensure eligibility. For example, one state may require two years of full-time supervised experience, one year of which must be postdoctoral, whereas another may require only one year which may be the predoctoral internship year (ASPPB, 1997). Finally, each jurisdiction also adopts one or more examinations that applicants for licensure must pass at specified levels. Almost all psychology regulatory boards in the USA and Canada use the Examination for Professional Practice in Psychology (EPPP) published by the Association of State and Provincial Psychology Boards (ASPPB) to assess knowledge deemed essential to practice in psychology. The EPPP has been validated over the years through a series of comprehensive job analyses and validation studies. The passing point on the EPPP varies across jurisdictions with most adopting the ASPPB recommended pass point of 70% of the 200 multiple choice items. Some jurisdictions also give oral or written examinations focused on mental health laws, ethics, practice competence, or some combination of these. These so-called complementary exams are typically locally developed, normed or standardized, and scored.

Although the major professional associations such as the APA, CPA, and ASPPB exert considerable influence over members of the profession who serve on licensing boards, the ultimate decisions about setting the educational, supervised experience or examination standards often rests with the legislative body (e.g., state or provincial government) that created and oversees the psychology regulatory board (Retfalvi & Simon, 1996; Wand, 1993). Regulatory boards have the authority and duty to suggest changes in the standards they use, but any party can seek to amend the standards. The final decision about changes to these standards is typically made by a political entity such as a state or provincial legislature and subject to the political forces at work in any locality.

Hence, there is no consistently applied standard of education, supervised experience or examination that is uniform across the USA or Canada. Most psychology licensing acts in the USA require a doctoral degree in psychology for the independent practice of psychology (ASPPB, 1997). A few allow some form of independent practice to persons with a masters degree in psychology and nearly half of the 61 jurisdictions in the USA and Canada allow individuals with masters degrees in psychology to practice as psychological associates under supervision of a licensed psychologist with a doctoral degree (ASPPB, 1997). In Canada, some provinces require standards equivalent to the doctoral standard in most USA states, whereas others grant full independent practice to individuals with masters degrees (Wand, 1993). Also, these educational, supervisory, and examination standards are not static—they change over time as a result of professional association initiatives to raise standards, political opposition to existing standards by some disenfranchized group, or economic forces.

Such diverse and fluctuating requirements for entry into the profession of psychology, both within and across jurisdictions, has resulted in great hardship for licensed psychologists who seek to relocate and find that differences in licensing laws across jurisdictions make them suddenly unacceptable for practice. This same patchwork quilt of licensure standards in psychology has seemed to confuse the public about the training psychologists receive prior to being credentialed to practice. One of the long standing missions of the ASPPB has been to bring greater uniformity across jurisdictions in the standards for licensure so that the public is less confused, and so that practitioners can move more freely across jurisdictions without fear of being denied a license in the new jurisdiction simply because of idiosyncrasies in the new jurisdiction's licensing law.

2.14.3.2 Investigating Complaints

The second major activity that psychology regulatory boards conduct to fulfill their mission is investigating complaints of incompetent or unethical practice against currently licensed practitioners. Each psychology licensing act specifies the grounds for which someone licensed as a psychologist may be disciplined, and the process by which such a charge would be investigated and adjudicated (Reaves, 1996).

Because a license to practice one's profession is clearly a property right of some value and consequence, a governmentally sanctioned process to deny, restrict, or rescind such a license must be conducted with adequate attention to the individual's rights of due process and appeal (Reaves, 1996). In the USA, a psychology regulatory board must provide the following constitutionally specified protections to any individual who is charged with a violation of the licensing act:

(i) adequate notice of the charges against them;
(ii) an opportunity for a fair and impartial hearing;
(iii) an opportunity to confront witnesses against them;
(iv) a decision that is based on the evidence;
(v) a record of the proceedings; and
(vi) some form of appeal or judicial review.

Similar legal protections and processes exist in federal and provincial law in Canada (Reaves, 1996).

Although some complaints are dismissed and others are settled through a negotiated agreement of stipulated conditions of remedy (e.g., mandatory supervision, additional training in ethics, etc.), complaints involving charges of significant misconduct or those charges that are contested are resolved through a hearing process. Before a hearing is scheduled, most jurisdictions in the USA and Canada require that the psychology board must have received a signed, written complaint, conducted an initial investigation of the merits of the charges, notified the psychologist of the specific charges, identified witnesses and subpoenaed documents, and taken any necessary depositions. The hearing itself is conducted in a variety of formats across jurisdictions, depending on local law and the resources available to the regulatory board. In some jurisdictions the hearing is conducted before the regulatory board with the board serving as a panel of inquiry and a decision-making body. In other arrangements, the board is assisted in this process by a hearing officer who serves to rule on motions and arbitrate disputes. In still other situations, the case is heard by the hearing officer alone and a recommendation and transcript of the proceeding is made available to the board which later renders its decision. Following an adverse decision, a psychologist may appeal to a separate legal tribunal (e.g., a circuit court in a state) if they believe they did not receive a fair and impartial hearing of the complaint.

As a result of the hearing, the regulatory board issues its findings of fact and conclusions of law, that is, its findings of what it believes happened and whether such facts constitute a violation of the licensing law. Where a violation is found, the regulatory board may impose an appropriate sanction ranging from a reprimand to a permanent revocation of one's license to practice. Other common sanctions imposed either alone or in combination include suspension of the license, restriction in the type of clients or problem areas addressed (e.g., no children or adolescents because of insufficient expertise with this population, or no forensic evaluations), a return to supervised practice, additional training or education, therapy, restitution of assets, reimbursement of fees, and administrative fines (Reaves, 1996).

According to Peterson (1996), complaints against psychologists have increased substantially since the late 1980s. This may be the result of greater consumer awareness, more vigilant response to complaints from licensing boards, or just better reporting of the same number of complaints that have typically occurred. Since 1983, the ASPPB has maintained a disciplinary data system (DDS) that collects and disseminates summaries of disciplinary actions from its member psychology regulatory boards throughout the USA and Canada. Until 1996, the summary reports of disciplinary actions were disseminated in printed form, first annually, and then quarterly as the number of reports grew. Beginning in 1996, the DDS was placed on-line electronically, whereby jurisdictions could report and retrieve information about disciplinary actions against a specific individual via secured electronic mail, giving regulatory boards immediate and timely access to all disciplinary actions contained in the system.

In a paper (Reaves, 1997) presented at the annual convention of the APA, Randolph Reaves, Executive Director and General

Counsel of the ASPPB, listed the following top 10 reasons for disciplinary action against psychologists in the USA and Canada for the period 1983–1997:

(i) sexual or inappropriate dual relationship with client;

(ii) unprofessional, unethical, or negligent practice;

(iii) fraudulent acts;

(iv) conviction of crime;

(v) inadequate or improper supervision;

(vi) impairment;

(vii) breach of confidentialty;

(viii) improper or inadequate record keeping;

(ix) fraud in application for license; and

(x) failure to comply with continuing education requirements.

According to Reaves (1997), the increase in disciplinary action can be noted through changes in the rate of reporting for several of these categories. For example, in the five years from 1983 to 1988, only 45 psychologists were disciplined for sexual or dual relationships with clients, but in the next 10 years this category increased to a total of 556 actions. Similarly, unprofessional, unethical, or negligent practice jumped from 29 in 1988 to 458 by 1997.

The consequences of being found in violation of the psychology licensing act are not limited to the sanctions imposed by the regulatory board. Reports from the DDS are routinely distributed to other relevant psychology-credentialing or professional associations as well as all regulatory boards (Reaves, 1997). The current distribution list of the DDS includes the ethics committees of the APA, the CPA, the American Board of Professional Psychology (ABPP), the National and Canadian Registers of Health Service Providers in Psychology, and the US Department of Health and Human Services. Given recent government initiatives, media attention and consumer activism, it is quite possible that information from the DDS may soon be made available to the general public through the ASPPB web site as a means for boards to discharge their responsibility to protect the public using modern methods of information dissemination. As a result of such broad distribution, individuals sanctioned by their home board of psychology may find sanctions imposed by other regulatory boards where they hold additional licenses, professional association ethics committees, other credentialing bodies such as the ABPP, or even insurance companies and managed health care providers.

As noted earlier, because psychology regulatory boards are composed of psychologists as well as public advocacy members, many psychologists notified of charges lodged against them expect or hope to receive lenient treatment from their psychologist peers on the board. However, psychologists serving on regulatory boards learn quickly that the role given to them by the legislative act creating the board of psychology requires them to protect the public and not show blind allegiance to the profession.

Furthermore, the rules established to protect the rights of individuals accused of violations necessitates a formal and complete investigation of all complaints, making informal resolution unlikely. However, because the boards of psychology typically decide the issue raised in a complaint, a psychologist unjustly accused can be assured that the panel of individuals who will decide the merits of the complaint is likely to be composed of other psychologists who will understand the conduct of the professional probably better than the average citizen who serves on a civil or criminal jury panel.

2.14.4 CURRENT AND FUTURE ISSUES

A number of challenges to psychology regulation lie ahead, both in the immediate future and the years to come. These challenges include disagreements over the educational standards required to call oneself a psychologist and to practice independently, attempts to redefine the scope of practice regulated by psychology boards (including both expansion of scope of practice to include new areas such as telehealth and prescriptive authority for psychologists, and resistance to encroachments by other professionals into the conduct psychological assessments), and finally, professional and governmental attempts to facilitate the mobility of practitioners both within and between national borders.

2.14.4.1 Educational Standards for Independent Practice

Since the Boulder Conference in 1949 (Raimy, 1950), organized psychology in the USA has espoused the doctoral degree as the entry level for independent practice in psychology. Most licensing laws for psychology passed in both the USA and Canada since Connecticut's first law adopted this doctoral standard in the required education of psychologists for independent practice (Retfalvi & Simon, 1996). A handful of states and provinces adopted the masters degree in psychology, sometimes due to the lack of availability of doctoral level psychologists to fill positions, sometimes due to political forces at work, and sometimes due to concerns about cost, accessibility in rural areas, or other factors. Many colleges and universities continued to prepare individuals at the masters degree level and many early licensing laws contained exemptions from the doctoral standard for state and local government, allowing publicly funded mental health agencies to employ individuals who did not meet the educational requirements for independent practice in that jurisdiction. Also, departments of education in the various American states began credentialing school psychologists for practice in that nation's public schools, and most of these state departments of education adopted a nondoctoral educational standard for independent practice (Fagan & Wise, 1993). All of these factors created a situation where controversy was likely to emerge over the standard for independent practice in psychology.

Such controversy has emerged in recent years in several different ways. Exemptions from licensure requirements for publicly employed psychologists has been challenged by patient advocates, and by professional associations, resulting in the virtual elimination of such exemptions. Individuals with master's degrees credentialed as psychological associates and working under mandatory supervision in jurisdictions allowing for such practice are now seeking to extricate themselves from the supervision requirement by changing licensing laws to allow independent practice with a master's degree. Also, school psychologists with nondoctoral credentials from departments of education are now seeking the right to practice outside the purview of the school without meeting the doctoral standard for independent practice required for such nonschool work in most states. In Canada, several provinces which previously held a doctoral standard similar to the typical state in the USA have rewritten their educational requirements to allow some form of independent practice for individuals trained at the master's-degree level.

Consequently, unlike medicine and law, the profession of psychology in the USA and Canada has never successfully advanced a universally accepted standard of educational preparation for its practitioners. The APA has promulgated standards of accreditation of training programs and model licensing laws, all of which adhere to the doctoral standard (APA, 1995). It remains to be seen if the APA will succeed in maintaining this standard in the years to come.

2.14.4.2 Scope of Practice Issues

Another recent challenge facing psychology has been attempts to expand psychologists' scope of practice into areas such as authority to write pharmaceutical prescriptions and telehealth. In 1997, bills were submitted in several state legislatures in the USA that were designed to amend the psychology licensing act to allow psychologists to write pharmaceutical prescriptions dealing with routine management of mental and behavioral disorders. Early suggestions, that psychologists should seek prescriptive authority, met with considerable resistance within psychology itself. However, recent surveys within the APA have indicated a considerable shift with a majority of practicing psychologists now supporting such authority. Some psychologists who are still opposed cite concerns about practitioners abandoning verbal psychotherapy in favor of writing prescriptions, but most proponents argue that psychologists' extensive training in verbal therapeutic interventions is most likely to result in a balanced and comprehensive treatment regimen should psychologists obtain the ability to prescribe psychoactive medicines as well (Bass, 1996; Deleon & Wiggins, 1996).

For psychology licensing boards, such proposals for change in scope of practice mean opening up licensing laws for amendments to the definition of the practice of psychology. Also, in the legislative arena, a professional group cannot control where the issue will end, resulting in a risk that psychology licensing laws could be altered in unfavorable as well as favorable ways whenever changes in licensing are attempted.

Another aspect of changes in scope of practice involves concerns over the regulation of the practice of psychology and other professions in the age of telecommunications. Licensing laws are premised on the fact that psychological services are delivered in a particular jurisdiction by an individual licensed to practice in that jurisdiction. What happens to regulation when psychological services are delivered to clients via the Internet or other telecommunications system across state, provincial or even national boundaries? If a client feels harmed, to whom do they complain—the board of psychology where they reside, which may have no jurisdiction over an out-of-jurisdiction provider, or the jurisdiction where the provider resides which may feel the activity did not occur in their jurisdiction?

Beyond the complexities of filing complaints, another issue being raised concerns the limits of what constitutes acceptable practice in this medium. Although one can easily think of examples of national or international experts who could and should be available for consultation about rare or specialized cases or topics, should more routine psychotherapy be conducted via electronic bulletin boards for the convenience or benefit of the practitioner or the client. In the absence of any system of national or international licensing, the regulation of psychological services delivered via electronic means is tenuous at best and perhaps nonexistent.

2.14.4.3 Mobility Across Jurisdictions

Given the inconsistency in standards for education and training in psychology across the various states and provinces, and the tendency for most jurisdictions to raise standards for education and supervised experience over time, psychologists licensed in one jurisdiction, particularly those trained perhaps decades earlier, often have difficulty obtaining licenses in new jurisdictions should they choose to relocate. This is true regardless of the fact that the psychologist met all applicable standards at the time they were originally licensed and may have practiced without any disciplinary actions against them for many years. For example, most jurisdictions who are members of ASPPB require 2 years of supervised experience, 1 year of which must be postdoctoral, in order to obtain the initial license to practice. However, such postdoctoral experience was virtually nonexistent as a licensure requirement 20 years previously. Consider, for example, an individual trained and initially licensed 20 years ago, who now wishes to relocate and obtain a license in a jurisdiction that requires a year of postdoctoral supervision.

Most jurisdictions have no mechanism to waive current requirements in such situations because their provisions for reciprocity typically state that the applicant must hold credentials essentially equivalent to the current licensure standards. Regulatory bodies often feel compelled to follow their requirements and procedures rigidly or face challenges from applicants without prior licenses or experience who are also lacking some aspect of qualification and want the board to grant them an exemption or waiver of standards. Obviously, the difference in these situations is that the previously licensed individual has a history of acceptable practice, but bureaucracies often have trouble with recognizing important but perhaps subtle distinctions.

For many years the ASPPB has been pursuing programs designed to facilitate mobility for the benefit of both licensed psychologists and regulatory bodies who are members of the ASPPB. For example, the ASPPB has supported efforts by member jurisdictions to enter into reciprocity agreements. Under the reciprocity agreement approach to mobility, all licensed psychologists in one jurisdiction are eligible for licensure in a cooperating jurisdiction based on evidence of comparable standards in current licensure requirements across jurisdictions. Reciprocity agreements are voluntary arrangements between jurisdictions and, although they are based on comparability in current standards, such agreements require acceptance of all licensed psychologists from a participating jurisdiction regardless of what standards were in place at the time any individual psychologist was originally licensed.

The ASPPB agreement of reciprocity has been in place for several years. However, despite considerable effort and promotion, only eight of 61 member jurisdictions participate. Many jurisdictions do not qualify for the agreement because they do not administer an oral exam, or they do not require a year of postdoctoral supervision. Some jurisdictions do not grant regulatory boards the legal authority to enter into such reciprocity agreements. Finally, despite the fact that many jurisdictions are ineligible or unable to participate in the reciprocity agreement, thousands of individual psychologists in those nonparticipating states and provinces do have the required education, experience and examination performance to meet the standards contained in the reciprocity agreement, but their mobility is not facilitated.

Since 1990, the ASPPB has been working on another more individualized approach to mobility, known as the Certificate of Professional Qualification in Psychology (CPQ) program. This individual certification and credentials bank approach to mobility would offer a repository for licensure-related data and an ASPPB-issued endorsement or certificate to individually licensed psychologists meeting standards for training, examination and professional experience adopted by the ASPPB. ASPPB member jurisdictions would be encouraged to accept holders of the ASPPB certificate as meeting the education, experience and examination requirements for licensure. The individual certification approach to mobility avoids many of the pitfalls that have plagued the widespread adoption of the reciprocity agreement, since it is based on an individual voluntarily demonstrating compliance with ASPPB-recommended standards for mobility.

A central aspect of the CPQ program is creation of an ASPPB credentials bank where individual psychologists can submit evidence of their education, training, experience and examination performance that can be maintained by the ASPPB and accessed as needed. Such a credentials bank will reduce the burdens placed on psychologists and regulatory boards when an experienced psychologist seeks to retrieve and submit evidence of meeting licensure requirements, perhaps long after their training and initial licensure. The CPQ program will be augmented by two other ASPPB services supplying crucial information about applicants for the CPQ, namely the score reporting service and the DDS, each of which offer the most complete repository of information on exam performance and disciplinary data available.

Thus, the CPQ program offers the potential for individual psychologists to deposit their credentials, submit evidence of appropriate education, experience, prior licensure, exam performance, and history of professional discipline to a receiving jurisdiction and avoid the complications and pitfalls of trying to document compliance with current licensure criteria through the traditional initial application procedure. The program is scheduled to begin accepting applications from individual psychologists in August of 1998.

In Canada, pressures from the federal government for free trade across provincial boundaries led to passage of the Agreement on Internal Trade (AIT). Among other things, AIT requires that provinces seek mechanisms and solutions to the problem of inconsistent standards in the regulation of professions across the provinces. The law requires that provinces identify procedures that will allow a professional recognized to practice a profession in one province to be recognized and credentialed for similar work in any other province.

Given the diversity of educational standards for licensing psychologists across the provinces, AIT is forcing Canadian psychology regulatory bodies to revisit their educational standards and make changes. For example, the province of Ontario has held to a doctoral standard for independent practice in psychology, similar to US standards, for many years. In response to AIT, Ontario recently amended their psychology licensing law to allow independent practice of psychological associates at the masters degree level. Thus, Ontario appears to have dealt with AIT by creating a credential for individuals with masters degrees in psychology to practice independently as psychological associates while the title "psychologist" is still reserved for individuals trained at the doctoral level. The full impact of AIT on the regulation of psychology in Canada remains to be seen.

Similar governmental initiatives facilitating free trade, such as the North American Free Trade Agreement (NAFTA) between the USA, Canada, and Mexico, could create even more complex requirements for homogenizing licensing standards for education, supervised experience, and examination across these countries. So far, NAFTA has not trampled on the psychology-credentialing systems that each country has worked hard to develop to meet the needs of its citizens. However, it is not certain how free trade initiatives in the future will affect the regulation of professionals around the world.

2.14.5 CONCLUSIONS

Regulation of individuals to practice psychology in the USA and Canada is a complex and multifaceted process. It involves the interplay of professional associations, legislatively created regulatory bodies, state and provincial governments, and the public. The sole purpose for the passage and continuation of psychology licensing laws is to protect the public from incompetent, or unethical practitioners. When the public (or at least their elected public representatives) lose confidence in the ability of the profession to regulate itself, then psychology licensing laws are either modified or rescinded. Although it may be tempting for the profession to feel that it knows best how to regulate the practice of psychology, the profession must be mindful to educate the public of the reasons underlying its preferred standards and policies, or risk having them rejected by those representing the public's interests.

As modern society changes through technological advances, marketplace forces, and government initiatives, it may be necessary to modify the standards and processes that have been created to regulate the profession of psychology to date. The USA and Canada have been world leaders in the development of credentialing mechanisms and fair and impartial disciplinary proceedings for psychologists. Despite the challenges outlined above, hopefully that leadership can continue for the benefit of the consumers of psychological services and the profession itself.

2.14.6 REFERENCES

American Psychological Association (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47*, 1597–1611.

American Psychological Association (1995). *Guidelines and principles for accreditation of professional programs in psychology*. Washington, DC: Author.

Association of State and Provincial Psychology Boards (1991). *ASPPB code of conduct*. Montgomery, AL: Author.

Association of State and Provincial Psychology Boards (1997). *Handbook of licensure and certification requirements in the United States and Canada*. Montgomery, AL: Author.

Bass, L. J. (1996). Future trends. In L. J. Bass, S. T. DeMers, J. R. P. Ogloff, C. Peterson, J. Pettifor, R. P. Reaves, T. Retfalvi, N. P. Simon, C. Sinclair, & R. Tipton (Eds.), *Professional conduct and discipline in psychology* (pp. 143–155). Washington, DC: American

Psychological Association and Montgomery, AL: Association of State and Provincial Psychology Boards.

Canadian Psychological Association (1991). *Canadian code of ethics for psychologists. Revised.* Ottawa, Canada: Author.

Deleon, P. H., & Wiggins, J. G. (1996). Prescription privileges for psychologists. *American Psychologist, 51*(3), 225–229.

Dent v. West Virginia, 129 U.S. 114, 9 St. Ct. 231 (1889).

Fagan, T., & Wise, P. S. (1993). *School psychology: Past, present and future*. New York: Longman.

Gross, S. J. (1978). The myth of professional licensing. *American Psychologist, 11*, 1009–1016.

Newton, L. H. (1988) Lawgiving for professional life: Reflections on the place of the professional code. In A. Flores (Ed.), *Professional ideals*. Belmont, CA: Wadsworth.

Peterson, C. (1996) Common problem areas and their causes resulting in disciplinary actions. In L. J. Bass, S. T. DeMers, J. R. P. Ogloff, C. Peterson, J. Pettifor, R. P. Reaves, T. Retfalvi, N. P. Simon, C. Sinclair, & R. Tipton (Eds.), *Professional conduct and discipline in psychology* (pp. 71–89). Washington, DC: American Psychological Association and Montgomery, AL: Association of State and Provincial Psychology Boards.

Raimy, V. C. (1950). *Training in clinical psychology*. New York: Prentice-Hall.

Reaves, R. P. (1996). Enforcement of codes of conduct by regulatory boards and professional associations. In L. J. Bass, S. T. DeMers, J. R. P. Ogloff, C. Peterson, J. Pettifor, R. P. Reaves, T. Retfalvi, N. P. Simon, C. Sinclair, & R. Tipton (Eds.), *Professional conduct and discipline in psychology* (pp. 101–116). Washington, DC: American Psychological Association and Montgomery, AL: Association of State and Provincial Psychology Boards.

Reaves, R. P. (1997). *Disciplinary actions involving psychologists in the United States and Canada: Top ten reasons*. Paper presented at the annual convention of the American Psychological Association, August, Chicago, IL.

Retfalvi, T., & Simon, N. P. (1996). Licensing, certification, registration, chartering and credentialing. In L. J. Bass, S. T. DeMers, J. R. P. Ogloff, C. Peterson, J. Pettifor, R. P. Reaves, T. Retfalvi, N. P. Simon, C. Sinclair, & R. Tipton (Eds.), *Professional conduct and discipline in psychology* (pp. 39–51). Washington, DC: American Psychological Association and Montgomery, AL: Association of State and Provincial Psychology Boards.

Sinclair, C., Simon, N. P., & Pettifor, J. (1996). History of ethical codes and licensure. In L. J. Bass, S. T. DeMers, J. R. P. Ogloff, C. Peterson, J. Pettifor, R. P. Reaves, T. Retfalvi, N. P. Simon, C. Sinclair, & R. Tipton (Eds.), *Professional conduct and discipline in psychology* (pp. 1–15). Washington, DC: American Psychological Association and Montgomery, AL: Association of State and Provincial Psychology Boards.

Wand, B. (1993). The nature of regulation and entrance criteria. In K. Dobson & D. Dobson (Eds.), *Professional psychology in Canada*(pp. 149–165). Toronto, Canada: Hoegrefe & Huber.

Webster's new world dictionary of American English (1986). New York: Webster's New World.

Weissman, H. N. (1984). Professional standards from the perspective of the sociology of professions. *Professional Psychology: Research and Practice, 15*(4), 471–472.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.15

Credentialing Issues, Current and Future: Scandinavian Countries

STEFAN JERN
University of Linköping, Sweden

2.15.1 INTRODUCTION 171
2.15.2 SCANDINAVIAN PROFESSIONAL PSYCHOLOGY: AN OVERVIEW 172
2.15.3 ACADEMIC TRAINING OF PSYCHOLOGISTS IN SCANDINAVIA 172
  2.15.3.1 Content and Forms of Academic Training in Sweden 173
  2.15.3.2 Specialist Training and Certification in Scandinavia 174
2.15.4 CERTIFICATION SYSTEMS IN SCANDINAVIA 175
  2.15.4.1 Norway 175
  2.15.4.2 Sweden 175
    2.15.4.2.1 Certification of psychotherapists in Sweden 176
  2.15.4.3 Denmark 176
  2.15.4.4 Finland 176
  2.15.4.5 Iceland 177
2.15.5 ETHICS SYSTEMS IN SCANDINAVIA 177
2.15.6 VIOLATIONS OF PRACTICE STANDARDS 177
2.15.7 MOBILITY OF PSYCHOLOGISTS AND RECIPROCITY OF LICENSURE 178
  2.15.7.1 Inter-Scandinavian Agreements and Mobility 178
  2.15.7.2 EU Regulations 178
2.15.8 FUTURE DIRECTIONS 178
  2.15.8.1 Continuing Harmonization 179
  2.15.8.2 Deregulation: Threat or Possibility? 179
  2.15.8.3 Remaining Issues 179
    2.15.8.3.1 Access to the national health insurance systems 179
    2.15.8.3.2 Subordination to the medical systems 179
2.15.9 SUMMARY 180
2.15.10 REFERENCES 180

2.15.1 INTRODUCTION

The Scandinavian countries Denmark, Finland, Iceland, Norway, and Sweden make up a cultural unity of 23.5 million people. They have been members of the Nordic Council since 1955, and, with the exception of Iceland and Norway, are member states of the European Union.

The inhabitants of Scandinavia share a common linguistic and cultural background, and their history, political experiences, and geographic closeness have created among them a spirit of Scandinavian cooperation and togetherness. Cooperation in the area of legislation began in the nineteenth century and has resulted in a high degree of harmonization of the legal systems of the region. Examples of early cooperation are the Treaty on Exemption from the Requirements of a Passport (1954) and the Treaty on a Common Labor Market (1954), as well as agreements on cultural cooperation (1975), social security (1982), and the right to practice for licensed professionals in the health systems (1984, not including Iceland).

The Scandinavian countries are developed welfare states with a shared belief in centralism and a high reliance on state control and regulation of, for example, health, social security, educational, and professional systems. As a consequence, since the Second World War Scandinavian professional psychology has developed in a context where training and practice have been influenced to a great degree by government regulations. However, professional organizations have had an important influence on decisions taken by the respective governments. The emergence of standards for training and credentialing systems have in the main followed a common path in the five countries, even though the pace in each may have been different. This chapter outlines the credentialing systems in the Nordic countries. Most examples and specific details given in this chapter are for Swedish conditions, as, in the main, these may be viewed as typical for most of Scandinavia. The number of psychologists in Scandinavia is approximately 15 000 (1996, including students and interns).

As no clear and agreed-upon definitions of licensure or certification exist in Europe, the term "authorization" will be used throughout to cover the means by which credentialing is achieved.

2.15.2 SCANDINAVIAN PROFESSIONAL PSYCHOLOGY: AN OVERVIEW

Compared to the other Scandinavian countries, the emergence of psychology, both as an academic discipline and as a profession, was relatively late in Sweden and Finland. Psychology in Sweden essentially developed after the Second World War, being influenced initially by German Gestalt psychology and later by American psychology as an academic discipline. Psychology as a separate academic discipline and as a professional practice emerged in Denmark and Norway in the 1930s, but did not flourish until the 1960s.

Even though the first professional psychologists in Scandinavia were active in the area of occupational psychology, clinical psychology in Sweden is discussed here as an example to illustrate general developmental trends.

Clinical psychology in Sweden, as in many other European countries, has developed in the context of a national health system run and financed by local and regional authorities. The health-insurance system is comprehensive, general, and has a universal outreach. For this reason the issue of statutory independence of medical supervision and management responsibility has probably become more of a key issue in the development of the profession in Europe than was the case in the USA. There is, however, a small, and now increasing, private sector as well (Jern, 1995).

Clinical psychology in Sweden grew out of clinical practice, initially virtually immune to the influence of academic departments of psychology and psychiatry (Nilsson, 1981).

The first Swedish clinical psychologists with an academic training appeared in the late 1930s in child guidance centers, which later evolved into child and adolescent psychiatric clinics. Within the framework of child psychiatric teamwork, the numbers of these pioneers multiplied to quite an impressive workforce by the early 1950s.

During the 1940s, general psychiatry turned to psychological university institutions for cooperation in diagnostics. Researchers then emerged as visiting consultants in a few university clinics. The clinical psychologist was looked upon as a "learned scholar."

The first positions for clinical psychologists at psychiatric hospitals and clinics opened up in the period between 1950 and 1955, and were mainly seen as the basis of clinical diagnostic laboratories performing psychometric assessment upon a psychiatrist's request. The role was that of "the paramedical assistant."

In the 1960s their duties began to include psychological treatment as well, and the 1970s saw the clinical psychologist enter the psychiatric team as a professional with a full range of skills, including psychotherapy, supervision, program development, and leadership tasks. This is the point where the "clinician–professional" enters the field.

Child and adolescent clinicians in general have been a decade ahead of adult clinicians in all areas of development (Jern, 1979).

2.15.3 ACADEMIC TRAINING OF PSYCHOLOGISTS IN SCANDINAVIA

Academic training of psychologists in Scandinavia is offered at state universities, of which there are three in Denmark, six in Finland, one in Iceland (undergraduate studies only), six in Sweden, and three in Norway. Universities are separate entities with unique charters, but are funded by public means (state taxes). Usually, objectives and curricula for professional training programs in, for example, medicine, law, and psychology, leading to authorization (certification or licensure), have by tradition been strictly regulated by central state authorities, although some local variation has been allowed. The professional associations have set their own standards for training since the mid-1950s, and usually this has had a strong impact on the curricula decided by the central university authorities in all countries. No accreditation systems are in effect as this is alien to the northern European academic tradition. The widespread Boulder model, as a general standard of training was succeeded by the Vail model in the mid-1970s. Today there is a strong move towards professional schools and a pure professional model with some research training, particularly at the newer universities.

Since 1993, each Swedish university has been free to set its own standards, but for psychology a voluntary agreement has been reached to maintain a uniform objective and a flexible system of 11 courses. Entry requirements are set by central authorities, but may vary locally depending on demand and supply. The University of Linköping, which was the last to start a program for the training of psychologists (1995), has developed a different model (problem-based learning), and has local selection and admission procedures.

Training programs are offered within faculties of social science, except for the University of Bergen, Norway, which has a faculty of psychology.

University tuition in Scandinavia is free and most students manage to support themselves on the state student loan system. Since the late 1970s the output in Sweden has been around 300 students per annum, reducing to 150 in the early 1990s. This last figure will more than double during the second half of the 1990s through a doubling of the intake and the addition of one university. These measures have been taken due to an anticipated reduction, by almost half, of the number of psychologists by the year 2010 following retirement of the core body, which comprises psychologists born during and after the Second World War.

In no Scandinavian country is a doctorate needed to become a psychologist. This is in line with the European tradition, where the doctorate degree is considered the first step in a research career. The standards and requirements for attaining a doctorate are generally viewed as too high for professional training, and are in many respects inadequate as the basis for professional practice. The length of training and the academic titles used in the different countries are given in Table 1.

The general model in the Nordic countries involves 4.5–5.5 years of academic training together with a subsequent or partially incorporated internship of 1–2 years" duration, leading to a generic certificate in psychology. The exception to this is Denmark, where a smaller proportion of school psychologists follow an older curriculum. On the northern European continent the same duration and scope of the academic and practical training is usually the rule, but more recent specialization of studies and the subsequent internship has led to the creation of specialized licences or diplomas. In July 1990, the European Federation of Professional Psychologists Association (EFPPA) agreed on optimal standards for professional training in psychology (EFPPA, 1990; Matefi & Häring, 1993), which encompass the generalist as well as the specialized model.

2.15.3.1 Content and Forms of Academic Training in Sweden

Admission to the five-year training program is granted after 12 years of general schooling with a sufficient amount of mathematics, Swedish, and English. This training program is extremely popular, and competition for entrance is very strong, with 16–40 applicants per opening. There is also a requirement of one year's general experience of working life before entrance (Jern, 1995).

The five-year academic training is generic and divided into 11 main courses as shown in Table 2.

Parker (1990), among others, has pointed out the split between academic and applied psychology in Sweden. Certainly this scenario is not unique to Sweden, but it has created specific problems in the training system over the years. The first chair in psychology was established as late as 1948, and the first professor in clinical psychology was appointed as recently as 1983. However, the academic side of Swedish clinical psychology is internationally well known and recognized. With regard to the lack of interaction between academic and applied psychology, Sjöberg (1986) has offered the explanation that, although by the late 1970s research had reached a high standard of methodological refinement, it had little to offer in the way of a metatheory for applied clinical psychology.

Through the establishment of chairs and departments in applied and clinical psychology in the 1970s, the beginning of a remedy was instituted. As the number of experienced professional psychologists has increased, a steady flow of applicants for the doctoral programs has been noted by the universities.

Table 1 Comparison of study time and the academic titles used in Scandinavian countries.

Country	*Training time* (years)	*Title*[a]
Denmark	4.5–5	Cand. psych., Cand. paed. psych.
Finland	6–7	Psykologian Kandidaatti
Iceland	Scandinavian standard[b]	
Norway	6–7	Psych. Cand.
Sweden	5	Psykologexamen[c]

[a] Direct comparisons cannot be made as the longer programs often include at least one year of internship, which in shorter programs is completed after the examination. However, most programs contain one or more periods of practical experience ("practicum") ranging from 8 to 20 weeks of full-time study. [b] The University of Reykjavik offers a three-year undergraduate program, which has to be completed abroad, formerly usually in Denmark, Norway, or Sweden, but now more often in the UK or the USA. [c] Literally "psychologists exam" equivalent to, for example, the "physician's exam" or "dentist's exam." Unfortunately the official English translation on the Swedish certificate reads: "Master of psychology," which might give a false impression if the European Masters degree is set equal to the American MA, which is usually of a more limited scope. The correct translation would probably be Psy.D.

Table 2 General curriculum: psychologist program, Sweden.

Course	*Points*[a] *Lund University*	*Min./max. in six universities*
Introduction	5	5–7
Psychological functions and their biological bases	16	15–20
Humankind in different developmental perspectives	19	17–23
Society, analysis of society, and social psychology	12	10–15
Personality and psychopathology	19	15–20
Assessment and intervention (individuals)	16	16–22
Practicum	15	15–15
Assessment and intervention (groups, organizations, and environment)	38	30–40
Psychotherapy and psychological treatment	25	22–28
Psychological research methodology	15	10–15
Thesis	20	20–20
Individual choice	0	0–10

[a] One point is equivalent to one week's full-time study in the 40-week academic year.

The stipulations for the diplomas in specialist areas of psychology have also intensified the research activities among practitioners, and the split between the academic and applied fields seems to be diminishing.

2.15.3.2 Specialist Training and Certification in Scandinavia

All the Nordic countries, except Iceland, have some type of specialist training following the completion of the authorization. The duration of the training, which is made up of supervision, course work, and a thesis, is usually around five years and is combined with practice in the chosen area of specialization. The certificate of specialization is awarded by the psychological associations (in Norway since 1985, in Denmark since 1994, and in Sweden since 1995). The Finnish system requires the candidate to take a licentiate's degree after having passed a professional program of 80 weeks duration. The degree, is not regulated by the state or by the association. Areas of specialization differ somewhat between countries, but cover clinical psychology, organizational psychology, community psychology, and school psychology, with some subspecializations. The certificate as a specialist is recognized and minimally protected only by the authorities in Norway.

2.15.4 CERTIFICATION SYSTEMS IN SCANDINAVIA

In general, since the 1970s the credentialing systems have moved from professional organizations to state authorities, and they are now a public responsibility in all countries. This applies only to the primary licensure level, however, as systems of specialization have developed within the associations and remain with them. Supervisory bodies controlling the conduct of psychologists with regard to adherence to the principles of sound science and reliable experience are also a part of the public sector, whereas ethical issues are supervised, judged, and decided on by bodies within the ethics boards of the professional organizations, or equivalent professional bodies.

2.15.4.1 Norway

Since 1973, the law has stipulated that no one may present himself or herself as a psychologist unless he or she has acquired the degree of Psykologie Kandidatexamen; can give evidence of an internship of a specified length; and can provide proof of leading a respectable life. The government will appoint a Psychologists Council to counsel the Department of Social Issues in professional matters and issue provisions regulating the practice of psychology. Psychological methods can be (and are) protected by Royal Ordinance. The department can withdraw its authorization of a psychologist if advanced age, a handicapping disease, or chemical abuse make him or her unfit to practice. Withdrawal of authorization can also occur if a psychologist neglects his or her duties or refuses to follow provisions imposed by the public authorities. The decision is not arrived at in a law court but through an administrative procedure. The confidentiality of psychologists is regulated by law.

2.15.4.2 Sweden

From the mid-1950s until 1978, authorization as a psychologist was granted by the Swedish Psychological Association for persons holding an equivalent of the US Ph.D. (filosofie licentiatexamen) in psychology or education and a 1.5-year internship or, for assistant psychologists a BA in psychology, educational psychology, or sociology and six months of practical training under supervision. Persons fulfilling these requirements were awarded a diploma stating that he or she was authorized by the association to practice either as an independent psychologist or as an assistant psychologist.

After a rather short political campaign, and opposed only by parts of the medical and legal professions, in 1978, psychology was included in the Law on Authorization to Practice a Profession within the Health and Medical Services. Licensure requirements involve taking the psychologist's exam and the completion of a one-year internship under supervision (practical training for psychologists [PTP]), with experience from individual and organizational work. Until 1996 the internship had to be completed within the public services, but it can now be done in a private practice setting as well.

The law originally protected the titles "authorized psychologist" and "authorized psychotherapist" and included a provision that the authorized psychotherapist should also state his or her basic training. In 1992 this law was amended with a penal protection for the professional designations "psychologist" and "psychotherapist" when offering healthcare services. The use of these designations outside the healthcare area is not regulated by law, where the protection is of the title "authorized psychologist." Directives and general advice on authorization of psychologists are issued by the Swedish National Board of Health and Welfare, which is also the agency with the right to inspect and control all medical professions. There is no psychologists law, as in Norway, Finland, and Denmark, and no separate psychologists council. With the exception of hypnosis and the practice of certain forms of psychotherapy with some defined groups of patients, there are no regulated or protected psychological fields.

The authorization can be withdrawn on essentially the same grounds as in Norway. This is executed by the Board of Responsibility of the Health Care System (HSAN), which since 1980 has been the supervisory authority for all authorized and unauthorized personnel in the health system. The task of the body is to examine mistakes and neglect in the exercise of medical professions. This is done with special regard to the principles of sound science and reliable experience, whereas ethical considerations usually play a minor role. HSAN can also issue cautions and warnings as administrative decisions. HSAN is a free-standing government authority of eight members appointed by the government for three-year terms. Two members have backgrounds in the medical professions, whereas the others represent various labor unions and the parliament. The chairperson is a judge. Complaints may be filed by patients, relatives of patients, the National Board of Health and Welfare, the Parliamentary Ombudsman, and by the Attorney General (Langton, 1984).

The number of authorized psychologists in Sweden was 6200 in 1994, among a population of 8.8 million inhabitants. This figure includes clinical (2150), educational (800), community, occupational, and organizational psychologists. Of the clinical psychologists, approximately 1000 are in positions within general psychiatry, 800 within child and adolescence, and approximately 200–300 are in neuropsychology/neurology, medical rehabilitation, pediatrics, gynecology/obstetrics, internal medicine, transplantation surgery, addiction units, and pain control centers. Some 100 are self-employed full-time private practitioners. Of the 2150 clinical psychologists, I would estimate that around 1000 run a part-time private practice in addition to their salaried positions, with a mean workload of about 5–10 hours per week. I have not included in these numbers the significant number of clinically related colleagues working in antenatal clinics, nurseries and playschools, or in the areas of mental handicap and social services. There are no clinical psychologists in school psychology.

2.15.4.2.1 Certification of psychotherapists in Sweden

A special issue in Sweden is the certification of psychotherapists; a system which has been in operation since the early 1980s. Originally intended as postgraduate specialization training for psychologists and psychiatrists, the university examination for "psychotherapist" evolved as the basis for a free-standing profession which could also be entered by, for example, social workers, nurses, dentists, clergymen, and others, provided they had completed a 1.5-year postgraduate preparatory course in psychotherapy. The number of authorized psychotherapists has now reached 2600 as shown in Table 3.

Table 3 Number of authorised psychotherapists according to basic training: Sweden (January 1, 1994).

Basic training	*No.*	*%*
Authorized psychologist	1651	63
Social worker	432	17
Authorized physician (psychiatrist)	274	10
Authorized nurse	42	2
Others	200	8
Total	2599	100

The orientation of university training leading to the title "authorized psychotherapist" should be eclectic, but has in practice been overwhelmingly psychoanalytically oriented. Training in group psychotherapy, behavior modification, cognitive therapy, and family therapy can also lead to authorization. Training in these areas is offered by different private institutes, often run by the appropriate professional organization (Jern, 1995).

2.15.4.3 Denmark

Since 1993, Denmark has had a psychologists law protecting the title "psychologist" (restricted to those having earned either of the two university degrees mentioned below) and regulating the entrance to the profession by stipulating the degrees of Candidatus psychologiae or Candidatus paedagogiae et psychologiae supplemented by two-year internship as requirements for authorization.

The Ministry of Social Affairs appoints a Psychologists Council to act as the authorizing and supervisory body of psychologists to serve for a period of four years. Of the nine members of the council, two are appointed by the Danish Psychological Association. Complaints regarding an authorized psychologist should be submitted to the Patients Committee, which is the Danish equivalent of the Swedish HSAN. The committee is authorized to issue cautions and warnings. Withdrawal of authorization as a psychologist is decided in civil courts upon referral from the Psychologists Council. In addition, a psychologist who has shown grave negligence or severe lack of judgement in his or her practice of the profession can be sentenced to a fine or to imprisonment.

2.15.4.4 Finland

A law established in 1994 regulates professions in the health and medical services and gives the conditions for all authorized professionals, including psychologists, which are defined as belonging to a medical profession, irrespective of the area in which they practice. The title "psychologist" is protected, as are the tasks of the profession. The law regulates training criteria, general ethics issues, confidentiality, and continuing education. The supervision of professional conduct is delegated to the Centre for Legal Protection of Patients in The Health Service, which has a standing similar to that of HSAN in Sweden. The sanctions are essentially the same as those of HSAN, but the consequences for a psychologist of being convicted for a felony in a civil court seem to be graver than in other Scandinavian countries.

2.15.4.5 Iceland

Iceland is not part of the Agreement on the Right to Practice for Licensed Professionals in the Health Systems, but protects the title "psychologist" by law. The approximately 130 psychologists in Iceland are supervised by the Ministry of Education.

2.15.5 ETHICS SYSTEMS IN SCANDINAVIA

All Scandinavian psychological associations abide by the Ethical Principles for Scandinavian Psychologists, which were adopted during the period 1987–1991. The associations generally have boards of ethics for supervision of the ethical standard of individual psychologists and standing task forces for policy matters in the ethics area. Final decisions regarding suspension and expulsion are taken by the board of directors and minor sanctions, such as cautions and warnings, are decided on by the board of ethics. In some countries (e.g., Sweden) appeals can be raised to higher levels within the association, such as the board of appeals or the general assembly.

2.15.6 VIOLATIONS OF PRACTICE STANDARDS

The national supervisory boards do not keep open records of filed complaints or decisions, but do in some cases publish yearbooks. The establishment of a national disciplinary databank has been discussed in Sweden. Opposition from the associations and societies of medical doctors has been strong, and no action has been taken by the authorities. As the relevant laws are fairly new in all countries, except Norway and Sweden, data concerning psychologists are scarce at present.

In Sweden early reports (Langton, 1984) indicate that HSAN, files approximately 800 complaints a year, of which approximately 500 concern medical doctors, with around 50 each for nurses and dentists. Approximately 10 psychologists' cases have been filed every year, and of these about seven have been subject to disciplinary measures. Since 1981 not more than five psychologists have had their authorizations withdrawn, which should be compared to a yearly total for all medical professions of between 30 and 40.

The most frequent types of violation are difficult to narrow down from the figures available from HSAN, but a general observation has been that cases concerning child custody, sexual abuse, and psychotherapy increased during the late 1980s and early 1990s. A clearer picture may be arrived at by studying cases filed with and adjudicated by the Swedish Board of Ethics (see Table 4). In the ethics field no disciplinary data banks are kept. According to the statutes of the associations, settled cases should be published regularly in the associations' journals, but for reasons unknown, this rarely occurs.

Violations of practice standards seem to occur mainly in the sensitive fields of social investigations, forensic psychology, and clinical psychology/psychotherapy. The majority of

Table 4 Distribution of filed and adjudicated cases, Swedish Board of Ethics 1983–1995.

	Cases by year (filed/adjudicated)				
Category	*1983–84*	*1985–86*	*1987–88*	*1989–91*	*1992–95*
Social investigations and incest	9/1[a]	17/5[a]	12/5[a]	14/8[a] 10[a]	15/15[a] 13[a]
Clinical psychology and psychotherapy	9/2[a]	15/2[a]	17/2[a]	24/3[c]	20/3[b]
School psychology	4	2	0	4/2[a]	0
Public statements	7	8/1[a]	3	1	1/1[a]
Personal fitness	2	1	0	1	0
Educational activities	0	4	4	1	2/2[a]
Professional relations	0	4	1	4	0
Organizational	1	0	2/1[a]	4/4[a]	1/1[a]
Research	3/1[a]	3/1[a]	0	0	0
Other	0	0	2	0	0
Total	35/4[a]	54/9[a]	41/8[a]	24[a] 53/3[c]	32[a] 39/3[b]

[a]Caution. [b]Warning. [c]Expulsion recommended to the Board of Directors.

cases have been in areas where psychologists act as experts to courts in legal procedures on child custody and sexual abuse, and in clinical settings where psychologists have been accused of improper sexual advances towards clients.

The picture thus overlaps with what has been reported by HSAN, and is probably typical for most of Scandinavia. It has also been noted that, while in the 1980s public concern about psychologists' inappropriate professional conduct was centered around child custody and sexual abuse, in the 1990s this shifted slightly towards an increased concern about psychotherapy including memory recovering techniques, repressed memories, and similar phenomena.

2.15.7 MOBILITY OF PSYCHOLOGISTS AND RECIPROCITY OF LICENSURE

In essence, two different systems apply to the Scandinavian countries: inter-Scandinavian agreements and European Union (EU) regulations.

2.15.7.1 Inter-Scandinavian Agreements and Mobility

The Treaty on a Common Labor Market (1954) and the Agreement on the Right to Practice for Licensed Professionals in the Health Systems (1984) give the rights of free movement and practice to all psychologists in Scandinavia, irrespective of whether or not the psychologist is a citizen of an EU state or not. Every citizen in the Nordic countries is free to reside and work in any of the countries in accordance with the treaty of 1954.

When licensing laws and the regulation of psychology were introduced, the situation became somewhat more complicated and the right for professionals to practice was agreed upon by all Scandinavian countries, except Iceland. This also means that a license, a certification, or an authorization given to a psychologist in one Scandinavian country will be acknowledged by the receiving country as equivalent to its own authorization. Minor requirements for supplementary examination in, for example, social security legislation and language tests are allowed for.

The mobility under this law has been rather slight over the years, with the main movement being from Denmark, which suffers from unemployment problems for psychologists, to Norway, which has a steady shortage of psychologists, especially in the subarctic region. For language reasons, movements to Finland have been negligible, whereas Sweden has had some exchanges with Norway, especially in the border districts. Movement from Denmark to Sweden, and vice versa, is rare. It should be noted that, through the Nordic treaties, Norway became an equal of the other Scandinavian EU member states, in this respect.

2.15.7.2 EU Regulations

The EU Directive 89/48, which covers the free movement of workers between member states of the EU, gives all citizens the right to have their training acknowledged by all member states, provided it has been approved in their own countries.

The directive regulates "vertically" the validity of diplomas and examinations of professionals, with a minimum of three years academic training. The directive covers psychologists, even though EFPPA has striven for a special or "horizontal" directive (like those for medicine and law) covering psychology exclusively. As psychology is not a regulated profession in all member states this is not possible to achieve. The emphasis on issuing directives has diminished under heavy criticism from many member states, and it is now highly unlikely that the vertical directive will be enlarged to encompass psychologists.

The movement of psychologists between the Scandinavian states has not been affected to any great degree by the directive. Its main implication for psychology in Scandinavia at the moment is that it may harmonize the procedures for judging non-EU-trained psychologists' diplomas and licenses. The influx of non-Scandinavian psychologists is low, probably due to language barriers and labor-market requirements.

As the possibilities under the special directive failed to materialize, EFPPA has changed policy, and it is now striving to formulate standards for a "European psychologist" diploma that could be accepted by member states as a "professional passport."

2.15.8 FUTURE DIRECTIONS

The more than 20-year process that has brought about national authorization systems for psychologists in Scandinavia was coordinated initially through a close cooperation of the associations and more recently through EFPPA. The development in Norway has been the guiding principle for this work. The ultimate goal has been national licensing laws, that protect the title and the activities of psychologists, these being administered through a psychologists council under the national health authorities. In this struggle, attempts have been made to create a Nordic psychologists council within the Nordic Council, so far without success. Norway and Denmark have succeeded in creating their own professional bodies within the state administration, whereas resistance has been strong on the political level in Sweden.

2.15.8.1 Continuing Harmonization

The present situation presents a picture of legislation that is both united and divided. With the exception of Iceland, all states protect the title "psychologist" and regulate the training standards for authorization. Finland and Sweden have no psychologist law or psychologists councils and incorporate psychologists in a more general supervisory system for all medical professions. Sweden is also the only country that refuses to protect the activities of psychologists by law. The work on harmonization of legislation will probably continue, and may lead to a further harmonization of the credentialing and supervisory systems in the Nordic countries.

2.15.8.2 Deregulation: Threat or Possibility?

The beginning of an opposing force can be seen in the emergence of national deregulation regarding the professions. For example, in Sweden the attempts to have the control of specialization transferred to the Swedish Board of Health and Welfare may be thwarted by this move. The specialization system of medical doctors was initially a purely professional concern, but in the 1950s it was taken over by the national board. Lately, some aspects of the system have been returned to the professional bodies, and future development in this direction may affect the strivings of psychologists to have their specialities recognized by other professions, the general public, and society at large. It should not be taken for granted, however, that this trend will be entirely negative to the development of professional independence. It may be that the close cooperation between the profession and the public administration can be broken in favor of a more independent standing. If this is to be the case, the issues of a European harmonization of training and credentialing standards through EFPPA will be even more important.

2.15.8.3 Remaining Issues

As a rather young profession, psychology in general has been involved in professional competition to gain its place in society. In clinical psychology this has implied competition with psychiatrists and psychiatric social workers and nurses; in school psychology competition with remedial teachers, welfare officers, and head teachers; and so on in different areas. The battle for certification has been won in Scandinavia, and authorized psychologists are now regarded as belonging to an independent profession, both legally and in the eyes of the public. The relationship with the medical profession, on both an organizational and a daily level, could generally be described as reasonable, and is characterized by mutual respect. However, some problem areas remain, and I will point out a few central ones.

2.15.8.3.1 Access to the national health insurance systems

Clinical psychologists in independent practice can be reimbursed by national health insurance only in Norway. In Sweden this can be done only if the psychologist is also an authorized psychotherapist and the patient has been referred by a physician. This option is open only to the select few who have managed to negotiate a contract with the local health authority. In the other Scandinavian countries no reimbursement is possible. This severely limits free competition with other professions, and the issue will have to be approached on an international level, either within EFPPA or through purely Scandinavian efforts.

2.15.8.3.2 Subordination to the medical systems

As competition within the healthcare market is not free, in order to practice, most clinical psychologists are dependent on positions within the general health system, mostly within psychiatry. This system is legally regulated by the medical profession through the respective National Boards of Health and Welfare, where influential posts and bodies are held by physicians. This has led, for example, to a situation where the medical profession together with administrators has managed to abolish psychological units, departments, and clinics within the hospital system, with the exception of a few university hospitals. Instead, psychologists are incorporated into existing multidisciplinary teams under medical leadership. The phenomenon of being "a guest in somebody else's house" is rather typical for most branches of psychology in Scandinavia. The lack of its own organizational and economic basis within the system seriously hampers development, and perhaps even applied research. However, it must be emphasized that the cooperation with other professions on a daily basis in general carries very few elements of conflict. A possible development that might emerge as a consequence of a common credentialing system and increased autonomy for the profession in Europe is a new emphasis on private and independent practice.

2.15.9 SUMMARY

Since the Second World War, professional psychology in the five Scandinavian countries has evolved in a culturally and linguistically close-knit region, which has been distinguished by remarkable economic and social progress. This has led to a welfare-state framework where a belief in strong, centralized, governmental control of health, social security, and educational systems has been shared by most. All professions have striven to make alliances with governments and supervisory state boards.

Since the 1970s psychology has been regulated by laws and decrees in all Scandinavian countries, except Iceland. The psychological associations have had a strong influence on the training standards and credentialing systems and it might be said that there exists a common Scandinavian training model for psychologists. The professional title is protected in the four major countries and professional activities are protected in all but Sweden. The licensure has full reciprocity within Scandinavia as a result of similar training and credentialing systems, inter-Scandinavian agreements on a common labor market, and the acknowledgement of Nordic licenses in the health field. Directives from the EU have also furthered the free movement of psychologists, but mobility between countries is still low.

The system of control and supervision varies, with two main models having been established: the Danish/Norwegian model, which implies a psychologists law and a psychologists council; and the Swedish/Finnish model, where psychologists are credentialed and supervised by the same body as other medical professions. Goals for the future include creating a common Nordic psychologists council within the Nordic Council in order to further harmonize legislation. It is not clear at the moment if a rising tide of deregulation will lead to a return to authorization and standard setting by the professional associations. Whatever course the development will take, the joint forces of psychologists in Scandinavia and Europe will be needed to defend what has been achieved, and to further future integration of the profession in the region.

2.15.10 REFERENCES

European Federation of Professional Psychologists Association (1990). *EFPPA optimal standards for professional training in psychology* (EFPPA Booklet 3). Bonn, Germany: Berufsverband Deutscher Psychologen.

Jern, S. (1979). The development of clinical psychology: from learned science to professional treatment (in Swedish). *Psykolognytt, 16,* 8–10.

Jern, S. (1995). *Clinical psychology in Sweden today: Problems and prospects.* Paper presented March 18, 1995, at the International Symposium of Clinical Psychology, Nagoya, Japan.

Langton, B. (1984). The law of supervision—the Board of Responsibility of the Health Care System (in Swedish). *Proceedings of the Swedish Society of Physicians, 93*(5), 91–99.

Matefi, G., & Häring, E. (1993). *Inventory of regulations in the field of psychology in European countries.* Bonn, Germany: Berufsverband Deutscher Psychologen.

Nilsson, I. (1981). The history of psychology in Sweden: Laboratories, testing and clinical psychology (in Swedish). *Kritisk Psykologi, 1,* 4–25.

Parker, A. (1990). Towards an international profession of clinical psychology: An Anglo–Swedish perspective. *International Journal of Psychology, 25*(3), 355–370.

Sjöberg, L. (1986). Research directions in psychology (in Swedish). *Swedish research of today.* Stockholm, Sweden: HSFR.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.16

Recredentialing: Periodic Reassessment of Practice Expertise

DAVID A. RODGERS
Cleveland Clinic Foundation, OH, USA

2.16.1 INTRODUCTION 182

2.16.2 ALTERNATIVE METHODS OF ADDRESSING PROFESSIONAL INCOMPETENCE 182
  2.16.2.1 Economic 182
  2.16.2.2 Regulatory 183
  2.16.2.3 Litigation 183
  2.16.2.4 Collegial 183

2.16.3 PROACTIVE VS. RETROACTIVE ACCOUNTABILITY 183

2.16.4 WHO RECREDENTIALS, FOR WHAT PURPOSE? 183
  2.16.4.1 Licensure 184
  2.16.4.2 Specialty 184
  2.16.4.3 Job 184

2.16.5 RECREDENTIALING TACTICS 185
  2.16.5.1 Procedural Requirements 185
  2.16.5.2 Assessments of Knowledge 186
  2.16.5.3 Skills Demonstration 186
  2.16.5.4 Outcome Evidence 188

2.16.6 WHAT CONSTITUTES PROFESSIONAL COMPETENCE? 188

2.16.7 PRACTITIONER WORK 189
  2.16.7.1 Addressing a Public-need Domain 189
  2.16.7.2 Diagnosis 189
  2.16.7.3 Prescription 189
  2.16.7.4 Client Supervision 189

2.16.8 MATURITY OF THE PROFESSION AND RECREDENTIALING 190

2.16.9 ADDRESSING THE UNKNOWN 190
  2.16.9.1 Initial Credentialing 190
  2.16.9.2 Practitioner Solution 191
  2.16.9.3 Practitioner Recredentialing 191

2.16.10 SUMMARY 192

2.16.11 REFERENCES 192

2.16.1 INTRODUCTION

There is growing discussion of and pressure for formal recredentialing in health services fields (Allred & Steiner, 1994; Bennett & Langdon, 1995; Conn, 1994; Department of Health, Education, & Welfare, 1977; Johanson, 1995; McClennan & Herlihy, 1995; Norcini, 1994; Saunders & Paget, 1994; Wiens, 1983; Wilson, 1993). In psychology, as early as 1973, the American Board of Professional Psychology (ABPP) was considering a recredentialing requirement (Mayfield, 1987; Ross, 1973). In the 1990s, the American Psychological Association (APA) College of Professional Psychology has considered requiring a periodic recredentialing examination for renewal of its proficiency certifications. In medicine, the American Board of Family Practice embraced recredentialing when the board was first formed in 1969, and, by 1993, 16 of the remaining 18 US medical specialty boards had adopted it (Leigh, Young, & Haley, 1993). So have medical specialty groups in other countries (Clunie, 1993; Delahunt, Beer, & Taylor, 1992; Egerton, 1994; Goto, 1994; Phillips & Willis, 1994), nursing specialties (Eagle, 1992; Haag, 1992; Roessler, 1994; Simon, 1992), dental specialties (Allen, 1994; Nelson, 1994), and other health related groups (Chambless et al., 1993; Grim & Grim, 1995; Hasbargen & Culclasure, 1994; Pfister, 1995; Zarowitz et al., 1993). The phenomenon is nevertheless a relatively new one and is not without controversy. Medicine, for example, existed for at least 23 centuries with a defined ethics code, with little or no use of competency-based recredentialing requirements.

The concept of periodic professional recredentialing based on reassessment of practice expertise is highly attractive. Some professionals do not age well. Incompetent professional help is hazardous to the public and damaging to the reputation of all professionals. If credentialing insures competency, then recredentialing should weed out the professionally senescent or irresponsible or lazy who do not maintain their professional expertise over time. Many psychologists would agree with the observation of Sechrest and Chatel that, "Accountability for a profession entails the ability to vouch for the competencies and other desirable characteristics of members of the profession" (1987, p. 2). Periodic reassessment of professional expertise as a requirement for continued credentialing should therefore be a protection both for the public and for the profession.

The foregoing argument seems unassailable except for one critical flaw, the assumption that credentialing, and recredentialing, can insure competency. Even while strongly arguing for recredentialing in their own medical specialty, Bennett and Langdon state a conclusion that almost all who examine the issues carefully eventually reach: "Does [the recredentialing process] measure the ability of any individual to practice medicine? No, it does not" (Bennett & Langdon, 1995, p. 556). While the concept may be sound, the possibility of validly preassessing professional competency may still be remote (Wiens & Dörken, 1986). Without valid assessment procedures, recredentialing based on assessments of uncertain validity tends to be a sham (e.g., Koocher, 1979).

Such complexities warrant a more detailed look at the nature of professional competency and the mechanisms for protecting the public from incompetent practice, as these relate to the issue of recredentialing of professional expertise.

2.16.2 ALTERNATIVE METHODS OF ADDRESSING PROFESSIONAL INCOMPETENCE

The ultimate purpose of recredentialing is to identify professionals who were formerly presumed competent but who may have become incompetent since the last credentialing. Although "recredentialing" by that name may be a relatively new development (see D'Costa, 1986, p. 138), the culture has always had devices for protecting itself from incompetent professionals. That being so, then the question to be asked of the present recredentialing movement is whether or not it improves upon those traditional protections. A brief review of some of the traditional mechanisms is useful, for perspective.

Four ways that performance measures enter into evaluation of professional competency, ways on which the culture has primarily relied in the past, are through: (i) economic factors, (ii) disciplinary regulatory factors, (iii) litigation factors, and (iv) collegial controls by guild structures.

2.16.2.1 Economic

On the economic side, market forces are driven by performance in its most comprehensive sense, such that some degree of professional competence is necessary for professional reputation and economic survival. Maintaining a job or maintaining clients is therefore one form of continuous recredentialing. Indeed, economists might argue that the ultimate criterion of professional expertise *is* economic survival, and the more that economic forces are allowed to operate unencumbered, the more effectively they will serve the public (see Gross, 1978).

2.16.2.2 Regulatory

On the disciplinary regulatory side, professional service that falls below the floor of formally adopted codes of conduct can expose the professional to loss of licensure and privilege of practice. The APA Ethics Code and the Association of State and Provincial Psychology Boards (ASPPB) Code of Conduct are two such measures against which many regulatory groups judge adequacy of professional performance.

2.16.2.3 Litigation

On the litigation side, failure to practice at the level of competence expected of the profession, even though that level is poorly defined and may be capriciously determined by both litigants and jury, can result in malpractice suits. With its impact on both economic status and professional reputation, such litigation not only economically redresses past woes but also to some degree certifies against future professional incompetence and practice.

2.16.2.4 Collegial

On the collegial control side, since each professional's well-being depends heavily on the reputation of his or her profession, collegial groups have a strong interest in preventing colleagues from sullying the profession's image. Since they also have intimate knowledge of the workings of the profession, they are often in an optimal position to exert effective corrective pressure on colleagues who would deviate from acceptable standards. Paradoxically, therefore, self-regulation of a profession can be, and has often proved to be, relatively effective in protecting the public interest. The Hippocratic Oath, originating in the fourth century BC, taken upon joining the ranks of physicians, is often cited as the real beginning of the profession of medicine in the western world, and one of the primary bases for its prospering as a service profession.

Although often effective, professional self-regulation by collegial groups does introduce the dilemma of "the foxes guarding the chicken coop," such that professional protection can be, and often is, placed ahead of public good (Gross, 1978). Especially in the modern era, with an explosion of information resources that tend to undercut the "secret society mystique" that has allowed professionals to claim immunity from public scrutiny and public regulation in the past (Gross, 1978), guild controls have been increasingly replaced with public structures. However, it is worth noting that these public structures are often little more than a giving of legal sanction to the same professionals, appointed to governmental regulatory boards, that would otherwise have sat on professional society ethics boards (see Gross, 1978). This tendency to shift guild control from the collegial organization to the statutory regulatory body has appropriately been called "incestuous" by Matarazzo (1977), but would seem to be a pragmatic first step in the transition to more public accountability.

2.16.3 PROACTIVE VS. RETROACTIVE ACCOUNTABILITY

Neither economic, nor malpractice, nor disciplinary regulatory controls involve a formal demonstration of professional expertise, and all invoke correction only after public injury has occurred. Some form of proactive periodic evaluation, "proactive accountability," that does not depend on prior public injury, would therefore seem to be desirable (Department of Health, Education, & Welfare, 1977; Sechrest & Chatel, 1987). It is, of course, just such proactive accountability that initial credentialing and periodic recredentialing are intended to accomplish, whether the credentialing is in the form of licensure or of special group certification or of guild membership.

Even the requirement for periodic renewal of credential, a requirement that is characteristic of most licensures or certifications or professional group membership, is a process for reviewing the professional's record, to see if specified minimal standards have been maintained. Such renewal is usually based on no more than payment of fees, updating of address, and absence of clear evidence of malfunctioning, such as absence of a felony conviction. In this form, recredentialing has long been a part of professional regulation. In the remainder of this chapter, however, the concept of recredentialing will be reserved for renewal of credential based on some formal assessment of potential for future competency, in the absence of a history of incompetence. It will be differentiated from controls that are predicated on prior occurrence of lack of competency, which are the common bases for disciplinary actions or malpractice suits. The goal of recredentialing is to protect the public from future harm, not to redress past harm.

2.16.4 WHO RECREDENTIALS, FOR WHAT PURPOSE?

Since recredentialing is intended to protect the public by eliminating unqualified applicants who currently hold a qualifying credential and who therefore have not yet demonstrated disqualifying lack of competence, and since failure of the recredentialing process becomes the *de facto* definition of "unqualified," the process can be neither benign nor nonthreatening to the professional (Holmes, 1986). The threat varies considerably, of course, depending on who does the recredentialing and for what purpose.Three such categories of recredentialing can be specified: (i) licensure, (ii) membership of a specialty group, and (iii) retention of a job position. There is a tendency, although not a perfect correlation, for the potential for public protection to increase as threat to the professional increases. Therefore, protection of the public and protection of the professional are often in conflict (Gross, 1978), a fact that greatly influences almost all recredentialing considerations. Both can be profoundly affected by the specific evaluation procedures or standards used, and balancing out these two factors constitutes much of the "art" of developing recredentialing programs.

2.16.4.1 Licensure

Recredentialing for licensure is the most threatening to the professional, because it jeopardizes a major investment of time, effort, and money in a professional career. On the other hand, it is potentially most protective of the public. Without removal of the professional license, some segment of the public will probably continue to be exposed to the services of a presumably incompetent professional. Responsibility for such licensure recredentialing nearly always resides ultimately with a governmental authority rather than with a collegial group. Few professions, and no US or Canadian psychology jurisdiction, now require recredentialing for licensure, except for a mandatory continuing education feature, which may or may not qualify as a true recredentialing measure.

2.16.4.2 Specialty

The issue of what constitutes a recognized psychological specialty, an issue that is much in debate in the APA, is not intended to be raised by this discussion. In this chapter, the term "speciality" is used in its more generic sense to refer to any subgroup of a profession that distinguishes itself by inclusive and exclusive criteria from other members of the profession, and that uses some form of credentialing to identify membership in the subgroup.

Specialty credentialing, and recredentialing when it is undertaken, is usually done by a professional group rather than by a governmental agency. When such is the case, loss of specialty credentialing does not carry loss of career implications to the extent that loss of licensure would. Nevertheless, membership in a specialty board can allow access to job situations or remuneration structures or other professional perquisites that are denied non members, to a degree that can markedly affect professional welfare. Even governmental regulations can reference specialty recognition, giving such recognition the force of law, as though the credentialing were done by a governmental entity, and thereby increasing its professional importance. The threat to the professional of such recredentialing can therefore be high. The fact that the process is usually controlled by the profession itself, and can therefore be more specifically attuned to the concerns of the professionals, can reduce the threat somewhat but may at the same time reduce the assurance that the public interest will be protected (Gross, 1978). For example, it can be used more to assure orthodoxy than to assure competency, or to imply competency without actually assessing professional skills.

Recredentialing for membership in medical specialty boards is becoming common, as has already been noted. While it is still uncommon in psychology specialties, it is being considered in several. The ABPP diplomas and the APA College of Professional Psychology proficiency certificates are two examples already cited.

2.16.4.3 Job

Even though it is less threatening than loss of licensure, recredentialing for job retention poses considerable risk to the professional, representing loss of job security and potential loss of income. It may or may not offer protection to the public, depending largely on the nature of the recredentialing requirements. In the airline industry and some other jobs of high risk potential that require strong and measurable technological skills, recredentialing for retention of job is common practice and relatively effective when based on apparently valid procedures addressing issues of valid public and professional concern. In such situations, the professional's job insecurity is considered secondary to the public interest, and is handled in other ways. For example, airplane pilots can actually purchase loss-of-license insurance.

In psychology, job recredentialing is just beginning to be used. The newest player in the healthcare recredentialing picture is managed care, which may be developing its own recredentialing norms for panel members or other providers with whom it contracts. Such recredentialing may be as much performance based, in terms of history of meeting the demands of management, as it is competency based in terms of reassessment of usually recognized professional skills. As noted, such recredentialing may or may not protect the public. For example, dropping a psychologist from a managed care panel for having too high an average number of therapy sessions per patient, as judged by excessively money-driven norms, might actually harm the public rather than benefit it. Conversely, a managed care structure and most employer oversight situations offer opportunities for objective assessment of professional performance, and therefore could provide more valid assessments of professional competency than can most alternative assessment methods. Except to note that this trend may be the most influential of all factors in the future, data on both the nature and the direction of managed care and other job-retention recredentialing in psychology is sufficiently new that those issues are not reviewed separately in this chapter.

2.16.5 RECREDENTIALING TACTICS

Each of the foregoing categories of recredentialing may have a somewhat different primary emphasis. Licensure primarily protects the public, specialty credentialing primarily protects the integrity and image of the specialty, and job recredentialing primarily selects for whatever skills are deemed critical for the particular job. They nevertheless share a common concern of predicting future professional competency. They therefore tend to draw on a common body of approaches or tactics. The tactics used do much to determine the success of the recredentialing process, the satisfaction of the professionals, and the protection of the public. Four common classes of tactics can be identified: (i) procedural requirements, (ii) knowledge evaluations, (iii) skill demonstration, and (iv) outcome evidence.

2.16.5.1 Procedural Requirements

Recredentialing can be contingent on completion of certain procedures, such as completion of certain educational requirements (e.g. mandatory continuing education), performance of a certain number of critical professional tasks (e.g., log-book record of core procedures done by internal medicine residents (Wigton, 1992)), or consistent conformance to certain standards of practice (e.g., an average of no more than a specified number of therapy sessions per client).

The most common of these procedural requirements in psychology is completion of a defined number of continuing professional education (CPE) credits. The intent is to insure that the professional remains current, by exposure to an adequate amount of the newly emerging information in the field. It is doubtful whether mandatory CPE as currently implemented in most psychology jurisdictions represents an assessment of potential for future competency, however, because almost anyone of almost any level of competence or incompetence could complete the requirements, by simply sitting through the required number of hours of lecture (see Brennan, 1992; Conn, 1994). While mandatory CPE for relicensure is usually not considered to be "recredentialing" in psychology, it is considered to be an adequate basis for recredentialing by others (see Delahunt et al., 1992; Conn, 1994) and is the primary current basis for recredentialing of its certificate holders by the APA College of Professional Psychology.

The literature on mandatory CPE is complex and is beyond the scope of this chapter, except for two relevant observations. First, mandatory CPE requirements that are carefully tailored to the specific demands of professional practice are the most likely to assist in maintenance of professional competency (Conn, 1994; Rodgers, 1989). For example, the requirement to complete an annual "new developments in the field" review course prepared by a select group of specialty experts is likely to be more stimulative of professional competency than are the "catch-as-catch-can" market place driven programs that characterize most mandatory CPE psychology requirements. Such courses would also be more complicated and much more expensive for the credentialing group to prepare (Rodgers, 1989). Second, the primary hazard of mandatory CPE, other than its cost, is that it can give a false sense of security to the public. That outcome may of course please a professional group (Gross, 1978), especially if it relieves the professionals of the anxiety of evaluations that are more uncertain of successful passage.

Procedural requirements other than mandatory CPE are unusual in psychology, but might appropriately be considered, particularly in specialty areas. For example, completion of a certain number of biofeedback assessments and treatment plans might be a useful requirement for a biofeedback group desiring to assure a degree of currency in its techniques. Unless outcomes are monitored, however, no procedural approach that does not contain within it an inherent criterion of competency can assure presence of professional expertise. At worst, it only assures repeated exposure to or experience with bad approaches.

2.16.5.2 Assessments of Knowledge

Assessment of technical knowledge is the most common ingredient of formal assessment of professional skills, at least in psychology (see D'Costa, 1986). This may be so because educational systems have developed so much experience with such assessment. Many of psychology's major credentialing examinations are knowledge based. For example, the Examination for the Professional Practice of Psychology (EPPP), used by all US jurisdictions as a primary basis for professional licensure (Hall, 1987), is an examination of knowledge of the psychological concepts underlying the profession. Knowledge proficiency is also the type of credential offered by the APA College of Professional Psychology. Periodic repassage of a knowledge base evaluation is incorporated into several medical specialty recredentialing programs and is often considered as a possible requirement for recredentialing in psychology areas as well. For example, the APA College of Professional Psychology has debated incorporating such a requirement into at least some phase of its proficiency recredentialing requirements.

For use as reassuring indicators of professional expertise, knowledge base assessments have at least three major limitations. First, knowledge is not synonymous with either skill or competence. Even an elegant textbook knowledge of how to do brain surgery, or how to play basketball, does not safely translate into performance skills. Data supports this conclusion. For an example with even a highly circumscribed activity, a work sample test of anesthesia residents administering a spinal anesthesia correlated only 0.20 (4 % of variance) with a multiple-choice knowledge test of the procedure (Howard, 1987).

Second, demonstration of competent knowledge in one information domain does not assure adequate knowledge in a separate critical but unassessed domain (Howard, 1987; Koocher, 1979; Tonesk, 1987). For example, the EPPP assesses a broad universe of information, but falls far short of assessing the specific knowledge that a psychoanalyst or a client-centered therapist or a behavior-modification therapist must draw on daily in order to function effectively within his or her own professional framework.

Third, performance on knowledge-based exams may depend as much on competency in the techniques of assessing the knowledge and on familiarity with the specific vocabulary of the assessment procedure itself as it does on "basic knowledge." Psychology's primary vocabulary is distressingly imprecise, such that words can and do mean different things to different people. For example, what, precisely, is meant by "medical necessity," by "the unconscious," by "emotion," or by "psychotherapy?" Especially in the more precise nuances of professional expertise, such nonstandardized meanings can render the task of truly accurate assessment of knowledge precarious at best and impossible at worst. This problem is compounded when the examination is compressed into such restricted formats as multiple choice options, in which minor differences of interpretation can make significant differences in outcome. Probably reflecting in large part this impact of increasing unfamiliarity with formal test taking and with the conventional vocabulary of the classroom, the American Board of Family Practice recredentialing experience has found a statistically significant systematic decline in examination scores, of up to two standard deviations overall, upon subsequent administrations of its examinations to its experienced practitioners (Leigh et al., 1993). The alternative interpretation seems unlikely, that the family practitioner fresh out of a residency is a substantially more competent and more knowledgeable physician than is the 18-year veteran in the field.

The general conclusion is warranted that knowledge is a necessary, but not a sufficient, ingredient of professional competency. It is part of the professional psychology licensing examination in the USA, but is not the sole basis for licensure, for good reason. If it becomes the sole basis for recredentialing, then such recredentialing can be only a partial, and often misleading, indicator of competency.

2.16.5.3 Skills Demonstration

Like knowledge, skill is a necessary but not sufficient ingredient of competency. Skill is the ability to implement a decision once it is made or to achieve a goal once it is chosen. It is the surgeon's "good hands," the golfer's ability to place the ball near the hole on a shot out of a sand trap, or the client-centered therapist's ability to recognize and reflect the "true feeling" embedded in a client's verbiage. For assessment purposes, skill tends to have a "built-in" quality, such that it can be demonstrated on demand and cannot easily be faked. A candidate can daydream through a CPE lecture or cheat on knowledge-based exams, but cannot fake 90 words a minute on a typing exam or 10 straight baskets from the free-throw line. Carefully constructed, skills assessments can implicitly incorporate knowledge and judgment assessment as well, by requiring skills to be demonstrated in real or simulation situations that require knowledge and judgment of what skills to apply. Such skill assessment is the essential component of airline pilot recredentialing, in which actual performance in handling a plane or a simulator in a situation calling for flying skill combined with flying judgment is observed and evaluated.

When they can be effectively applied, skills-based assessments are attractive recredentialing options. Simulation of real performance situations may allow close assessment of actual performance behavior. The ABPP professional competency examination (Mayfield, 1987; Ross, 1973) and several of the medical specialty board exams, in which the candidate is observed while assessing a real or hypothetical patient or addressing a real or hypothetical clinical problem, are of the skills-assessment type. Some graduate training programs in psychology are incorporating skills assessment into their formal evaluation processes, and nearly all incorporate such assessments informally. Whenever a student is observed role playing a therapy technique or giving a standardized test or is asked to evaluate a projective instrument product in a class discussion or is observed through a one-way mirror while interacting with a client, a skills assessment is being undertaken.

In psychology, there are three limitations to skills assessment as a basis for recredentialing. First, integrity and consistency of application of the skill in real-life situations may differ from that in the test-taking situation. Second, the critical skills necessary for a particular credential must be identified, and such skills are not often either obvious or agreed upon. Third, the valid basis for determining skill competency and the relationship of the skill to actual performance competency should both be unambiguous, and they seldom are.

The first problem, that test behavior may differ from actual clinical behavior, is primarily a problem for performance variables that have differential payoff for the professional within and without the test situation. The motivation for accurately reading a pathology slide probably changes little from test situation to laboratory situation, for example. Unfortunately, many professional psychology skills can be used exploitatively as well as constructively, with one use being demonstrated in an assessment situation and the other in client situations. For example, the exploitation of the therapist–client relationship for sexual or other favors in fact occurs, and would be a clear basis for decredentialing, but would rarely be displayed or observable in an assessment setting. Indeed, a performance evaluation of integrity or other character traits, which require more absolute than relative consistency and over an extended rather than limited time period, and are therefore usually fakable, is almost impossible in a limited test situation. Such traits are nevertheless a critical factor in professional outcome competency.

The solution to the second problem, of what skills to assess, may be obvious for a typist, and perhaps even for a surgeon, but is not for a psychologist. As Sechrest and Chatel have observed: “A major problem for professional psychology is, of course, that we do not really know how to specify what psychologists ought to be able to do, ought in fact to do, and we know even less about how well they ought to be able to do anything” (1987, p. 5). If the generic field of psychology, as Howard (1987) observes, is not easily describable in terms of specific skills, then perhaps subspecialties within psychology are. A skills approach might be useful, for example, in such areas as biofeedback or hypnosis or behavior modification, even if not in generic recredentialing.

The third problem, of criterion for determining the adequacy of a skill, has been especially troublesome in psychology. Many procedural skills in other fields have relatively unambiguous and objective indicators of adequacy. A hairline fracture on radiograph either is detected or is missed. An airplane is either kept within or deviates from a predetermined glide path range. Similar compelling evidence for the appropriateness of a given therapy response or of specific marital counseling advice is generally lacking (Berven, 1987). This dilemma has often been resolved within psychology by relying on the expert judgment of colleagues. If the skill appears adequate to the observer, it is judged adequate. Unfortunately, when there is ambiguity concerning the relationship of the skill to the professional goal, as there often is in psychological practice (Howard, 1987), such an observer-judgment approach is perhaps a better guarantee of orthodoxy than of competence, and as much a measure of the observer as of the observee.

For many purposes, when specific skills are known to play a critical role in professional competency, then skill assessment can be a useful basis for recredentialing. It is being incorporated into most medical specialty board recertification programs. Psychology has not yet identified its critical skills (Stevenson & Norcross, 1987), except possibly in some specialty areas, but perhaps will as it matures more as a profession. While considerably short of a completely valid measure of overall professional competency, the ability of a candidate to demonstrate certain critical professional procedural skills is nevertheless strongly reassuring concerning his or her potential for competency.

2.16.5.4 Outcome Evidence

In terms of public benefit, outcome effectiveness is the ultimate criterion against which all professional expertise is validated (see Berven, 1987; Koocher, 1979). Before such validations can be made, though, outcome effectiveness must itself be reliably measurable (Howard, 1987; Stevenson & Norcross, 1987). In many fields, the outcome produced by the professional services is unambiguous and objectively identifiable. Repair of a broken arm, for example, should result in a reknitted bone and an arm that can be used essentially in the same manner as the same arm before it was broken. Symptomatic relief, such as relief of pain, may be a secondary goal, but is not equivalent to the repair of the break itself.

In psychology, the outcomes achieved are usually ambiguous, and the precision with which they can be either specified or identified is usually far from adequate. Psychology has generally failed, as a profession, to declare what its goals are and what evidence can be used conclusively either to justify or to condemn its actions. Other than client satisfaction, which has been the basis for most psychology professional services (it is also equally the basis, of course, for the services of the drug dealer or the strip teaser), there are almost no agreed upon outcome measures for the effectiveness of psychological work (Sechrest et al., 1987). Should marriage counseling produce divorce, nondivorce? Is symptom relief (e.g., of depression or anxiety) enough, or should hypothetical underlying causes be identified and removed? Should the "cure" address solely the current episode or should it serve as an inoculation against any and all recurrences? Are the psychologist's goals different from or the same as those of a psychiatrist or a social worker addressing the same problem, or indeed do these different disciplines actually address the same problems even when they use the same *ICD-10* (current version of the *International Classification of Diseases*) codes for diagnosis?

Until such basic questions are answered, good outcome measures of professional competence will be almost impossible to specify. Until such outcome measures can be specified, the concept of recredentialing to establish current professional competency is relatively meaningless, because competency will remain undefined. Ultimately, therefore, consideration of the fundamental requirements of effective competency recredentialing requires consideration of the basic nature, goals, and methods of the profession. In no other dimension is the immaturity of the profession of psychology more strongly reflected than in the lack of clarity of these fundamental issues. These issues warrant further attention later in this chapter.

Fields within psychology, such as psychoanalysis, client-centered therapy, behavior modification, and biofeedback therapy, often have more precise concepts of appropriate outcomes for their own areas of practice than does the profession as a whole (Howard, 1987; Tonesk, 1987). This may mean that outcome measures and recredentialing for specialty or proficiency areas is more appropriate than recredentialing for the entire field. Indeed, orthodoxy itself can be an appropriate goal for some groups, such that peer judgment of competency becomes the defining goal of a desirable outcome. With assured orthodoxy, the public protection comes from knowing what to expect, even if not from assurance that a particular outcome will result.

2.16.6 WHAT CONSTITUTES PROFESSIONAL COMPETENCE?

The public's thirst for reassurance, the profession's marketplace competitiveness to satisfy such thirst, and the vested interest of others such as managed care to control professional practice seem to be driving psychology, like medicine and many other professions (e.g., Norcini, 1994), toward recredentialing arrangements. If such is the case, then it would seem wise for the profession to pursue as expeditiously as possible a clear understanding of the goals of the profession, its methods, and the conditions that will contribute best to its future development. Some consideration of these issues as they relate to the recredentialing dilemma follows.

Professional work can be arbitrarily divided into two broad types, practitioner and technological, depending on how the work domain is defined (Rodgers, 1980, 1986). Practitioner work is defined independently of the techniques it uses, by the area of need it addresses. Practitioners may or may not be able to provide a satisfactory solution to the problem they confront, but accept responsibility for it anyway, because it is in their professional domain. Medicine, for example, is defined as that profession which addresses problems of illness, even those illnesses for which it has no cure. Law, similarly, is defined as that profession that addresses legal problems, whether or not the problems have clear legal solutions. Because they must choose methods to fit the given problems, rather than selecting only problems that fit their methods, practitioner professions generally require "unlimited" licenses of practice.

A technology field, in contrast, is defined by its techniques, and limits its scope of work to those problems that fit the technology. A technology field is not expected to address problems for which it lacks a solution. Airplane piloting, for example, is defined by, and limited to, the ability to fly a plane. It does not address any and all transportation needs.

Psychological practice can be of either practitioner or technological type, depending on its focus. For example, the defined and orthodox therapies, such as client centered, psychoanalytic, or behavior modification, are technologies by this conceptualization, being limited to those conditions that are appropriate to the particular approach. Psychologists who specialize exclusively in one or more of these orthodoxies would appropriately be classed as technologists, as that term is here used. In contrast, the clinical psychologist whose office is open to anyone with a psychological problem is a practitioner, even if his or her most common (but not exclusive) method of handling such problems is a particular technological orthodoxy.

In general, technologies lend themselves well to recredentialing approaches, as the evaluation of competency can focus on either the orthodoxy of the professional, by peer judgment of a work sample, or the expected outcome if that is clearly implied by the procedure, such as the relief of phobic symptomatology by a behavior-modification deconditioning regimen. In contrast, practitioner approaches are difficult to assess with competency measures, because they often necessarily address problems of unknown solution with techniques of unknown validity, and routinely address problems of inadequately known solution with techniques known to be only partially effective. Under such circumstances, it is almost impossible on any given trial to divide the variance of either success or failure between an accurate component representing professional skill and other components not reasonably attributable to the work of the professional.

As most psychologists, even specialists in a particular approach, are usually open to assessment and management of problems that may not exactly fit the particular orthodoxy, it is suggested that practitioner work underlies most licensed clinical psychology practice. Understanding the nature and value of such work may therefore further clarify the complexity of the problem of recredentialing in psychology.

2.16.7 PRACTITIONER WORK

2.16.7.1 Addressing a Public-need Domain

As has already been noted, there is little consensus on the proper domain of professional psychology. It has been argued elsewhere (Rodgers, 1981, 1993, 1994) that it should be defined as any problem in human coping effectiveness or personal efficacy. Precise identification of the public needs addressed is ultimately critical for the profession, but is not essential for the material to follow, and is beyond the scope of this chapter.

2.16.7.2 Diagnosis

In practitioner work, the client typically enters the door with a problem for which the solution has not yet been established. The first practitioner act is therefore typically one of diagnosis, that is, determination of the nature of the presenting problem and its relationship to known corrective procedures. Only in the earliest days of the profession could the concept reasonably be held that there was no need for diagnosis, because a single solution (e.g., psychoanalysis, client-centered therapy, behavior modification, etc.) could fit all problems. Most of the profession has given up the hope that the inherent uncertainty of the task which inevitably confronts the practitioner could be converted into the more comfortable certainty of a technological “one size fits all” approach. Diagnosis is therefore necessary, to identify the presenting problem in terms that allow it to be channeled to relevant corrective approaches.

2.16.7.3 Prescription

After diagnosis, the solution to the presenting problem must be determined and recommended. The second practitioner act is therefore typically one of “prescription” of solution. Depending on the nature and clarity of the problem, the solution might be treatment with an orthodox approach, either by the practitioner or by referral to a technical specialist. Often, it is a practical procedure that is designed to deal with the immediate consequences of the problem more than with the problem itself (Rodgers, 1964).

2.16.7.4 Client Supervision

Unfortunately, psychological solutions are seldom as routine as identification of a problem and referral to an appropriate technologist for inevitably effective treatment. Therefore, a third step is usually necessary in practitioner work, and that is to follow the client’s progress as corrective steps are tried, modify the diagnosis and treatment plans as necessary, and determine when maximum or optimal benefit has been achieved such that treatment should be discontinued. This third step can be called assuming professional responsibility for the patient, or "client supervision" (see Ohio Administrative Code Section 4732-5-01(B)(1.3)).

The critical dimension in these three essential activities of the practitioner is professional judgment in the face of uncertainty, not application of specific techniques. Indeed, a practitioner could utilize technology "extenders" or referral sources, for psychometrics, for deconditioning, for group therapy, even for individual therapy, and concentrate solely on the professional decision-making of diagnosis and prescription and client supervision, and adequately fulfill all the essential professional requirements of a competent practitioner. The task of assessing the professional expertise of such a professional would then become the task of assessing the adequacy of his or her professional judgment, not his or her skill at carrying out specific procedures. To be sure, such exclusive concentration on judgmental tasks seldom occurs, partially because it is not consistent with the training models used in psychology (but is not inconsistent with those used in medicine), and partially because some of the most useful diagnostic and client management information can come from the interaction associated with treatment or other procedural approaches. Nevertheless, such analysis does suggest that professional judgment, not professional technique, constitutes the *essential* nature of psychological practitioner work.

2.16.8 MATURITY OF THE PROFESSION AND RECREDENTIALING

The maturity of a practitioner field can be measured by the proportion of presenting problems that have clearly defined diagnoses and clearly defined optimal treatment approaches. In this regard, medicine at the turn of the twentieth century is astoundingly mature as compared to medicine at the turn of the nineteenth century. Antibiotics, anesthetics, and exploding biochemical knowledge have converted many formerly ambiguous judgment calls into essentially computer-specifiable exact protocols. When computer certainty can substitute for human judgment, then assessment of the competency of the human judgment is relatively easy and objective. In many of its fields, medicine is approaching that stage of certainty, even in its practitioner areas. As a consequence, effective assessment of competency of professional medical judgment is becoming feasible, perhaps for the first time, and periodic recredentialing based on measuring such effectiveness is becoming tenable.

In this respect, psychology is far from a mature practitioner field. It does not have its own catalog of types of presenting problems, continuing to borrow psychiatry's *DSM-IV* nosology. It does not have workable outcome measures in any substantial area of practice, except possibly measures of symptomatic relief. It has entrenched disputes over appropriate approaches to almost all of the problems that it recognizes as being within its domain, usually along orthodoxy lines of preferred therapies. As previously noted, the profession has not even objectively defined or identified some of its most fundamental concepts. For example, what is "unconscious process?" How does psychotherapy differ from "counseling?" Is "marital counseling" counseling or is it therapy? Is not therapy simply a meaningless term used to describe whatever it is a psychologist does to try to address a client's problems? These areas of confusion or lack of precision are evidence of an immature profession. More importantly, they are evidence that there are few clear-cut objectively definable "correct" judgments for the decisions the practitioner psychologist makes daily (see Berven, 1987, p. 223).

Even conformity to conventional thinking may be wrong, in terms of ideal outcome, when the clear solutions to problems are not known. Within this century, conventional wisdom has called for protracted and ultimately debilitating bed rest following a heart attack, and for radical mastectomy for all breast cancers however small. Under such circumstances, objective, defensible, valid assessment of competent professional judgment is at best difficult and at worst impossible. Therefore, recredentialing decisions based on assessing such competency become precarious (Koocher, 1979).

2.16.9 ADDRESSING THE UNKNOWN

The foregoing analysis is not intended to present a hopelessly dismal picture of the profession of psychology. On the contrary, it is intended to highlight a stage that every developing profession must go through, a stage of considerable uncertainty during which the public may long for reassurance concerning the professional's ability, and may unknowingly seek false security through recredentialing processes that can only minimally address the concerns for which they are designed. The positive aspect of the practitioner approach perhaps needs highlighting.

2.16.9.1 Initial Credentialing

Initial credentialing usually assures that practitioners are extensively trained in the best techniques known to the culture to address the problem areas of the professional domain; have demonstrated a capacity in apprentice roles to function at least minimally effectively; have during education and training and internships, survived many years of close scrutiny by mentors and members of the profession and been found not to display "fatal flaws," characterologic or otherwise, for carrying on the work of the profession; have demonstrated extensive technical knowledge of the field; and have been placed, after licensure, under the ongoing scrutiny of a regulatory board that can at any time measure their performance against a code-of-conduct floor below which their professional work must not fall. In short, the licensed psychologist has been selected to be as knowledgeable as possible about the area and to be unlikely to constitute an unusual hazard to the public, while still being able to render reasonable service. Such a select person usually recognizes and appropriately addresses common problems of known solution. It is the problems of unknown or only partially known solution, but of substantial cultural importance, that raise questions.

2.16.9.2 Practitioner Solution

For these important problems of unknown or ambiguous solution, the question must be raised of how the culture is to address them. To require a guarantee of success is to leave them unaddressed, or to turn them over to dishonest pretenders who falsely guarantee more than they can deliver. The practitioner solution has been to restrict those who can attempt to help such problems to a carefully selected licensed group; to try to limit, by means of protracted training and apprenticeship and a legally enforced code of conduct, the harm that such a select group can do; and then to leave that group free to do the best it can. In other words, the practitioner is deliberately given the freedom to function within a cultural veil of ignorance, because the need is there, and to do what makes the most sense in spite of the ignorance, in the hopes that someone will ultimately lift the veil and allow predictable solutions without creating too much harm in the process.

Those practitioner professions should advance the most rapidly that incorporate strong procedures for progressing through ignorance. This means incorporating procedures that encourage individualized approaches that are outcome driven or at least are outcome assessable in ambiguous areas, and discouraging ones that allow for excessive nonoutcome-driven idiosyncrasy. It also means avoiding procedures that either push too much toward unproved orthodoxy or carry more potential risk than any potential benefit can justify.

In this context, young professions such as psychology can concentrate too hard on proving their adequacy and not hard enough on identifying their inadequacy, and the cultural inadequacy, for dealing with critical problems. Medicine advertises its inability to cure cancer and many other illnesses, and stimulates the research funding to find solutions. Practitioner psychology does not even clearly define the problem domain within which it is willing to contend with the limitations of the present stage of knowledge, seldom admits inability to address a problem, seldom advocates for public support to study its areas of ignorance, seldom identifies clear outcome goals, and often seems more comfortable with technological certainty than with practitioner uncertainty, sometimes thereby becoming a valuable technological extender of a medical practitioner, but blurring its own professional identity and own contributions to lifting the cultural veil of ignorance that cloaks today's unsolved psychological problems.

In the ideal world, the practitioners are a guide through the quicksands of cultural ignorance in their own specialized professional domain, far more than a scientific tour guide over a well-established path. Some (e.g., Stern, 1984) have argued that the practitioner's work should become identical with that of the scientist at this stage, a carefully constructed *in situ* experiment directed toward resolving the problem. This view is disputed by Rodgers, (1964), except insofar as clear outcome goals that allow confirmation/disconfirmation of the professional work inevitably make the work scientific. The primary creed of the practitioner in the face of ignorance is to avoid harm, a commitment that often prevents a definitive scientific test. The next commitment is to deal practically with whatever aspects of the problem lend themselves to being addressed. These steps are often far different from "scientific inquiry," but are critical (Rodgers, 1964). The third step is to function within the accepted standards of the profession, which standards may reflect art more than science but at least reflect a consensus judgment that is safer than individual capriciousness. Finally, if experimental procedures are introduced into the treatment, then these should indeed be approached with the same perspective as a scientist, but this final step may be a rare occurrence in routine practice.

2.16.9.3 Practitioner Recredentialing

It is suggested that dealing with the problems of cultural ignorance is not primarily an individual responsibility, but is the responsibility of the profession as a whole. Adequacy in this role must essentially be judged over time and for the profession, in terms of evidence of successes and failures in the real world and progression toward removing the ignorance. In most practitioner fields, including psychology, competency in this basic role of the practitioner, dealing effectively with important problems under a veil of current ignorance, is essentially unmeasurable by examination, at least by currently practical technology. For this role, high standards of preparation for initial licensure, the influences of professional guild membership and collegial pressures, and the controls of the marketplace, of regulatory board codes of conduct, and of citizen resort to malpractice redress may serve both the public and the profession better than can any feasible recredentialing process, at least until the profession and the field of knowledge which underwrites it matures to the point that a large body of its professional judgment can be effectively reduced to rather precise and agreed upon decision-tree protocols. Recredentialing seems better suited to specialty and job situations in an immature profession than to the general practitioner function.

2.16.10 SUMMARY

Credentialing of expertise in psychology raises many questions and many complexities. Recredentialing conceptually raises many more. In practitioner work, initial credentialing may, appropriately, do more to weed out those at high risk for doing harm and low potential for doing good—the poorly educated, the naïve, the obviously irresponsible, and the professionally unsocialized—than it does to assess and assure current professional expertise. Placing reliance on further experience to sharpen the necessary skills, initial credentialing may set an important floor of reasonable professional training and socialization, but still not guarantee that the essentially untried and newly credentialed practitioner will be immediately effective. Recredentialing, in contrast, has little place unless it directly addresses current professional competency. That, it turns out, is a difficult task in professional psychology and especially in practitioner psychology.

No US or Canadian jurisdiction has yet adopted a competency recredentialing process for the generic or clinical psychology license. The American Board of Professional Psychology, the American Psychological Association College of Professional Psychology, and other specialty or proficiency certifiers within the field have explored competency-based recredentialing procedures, but few if any have adopted them. Therefore, practical experience with such procedures in psychology is lacking. The medical field is gaining experience with recredentialing approaches, however, within their specialty boards. The evidence so far gained suggests that the procedures are feasible in professionally mature specialty areas, are generally demanding and expensive if they attempt to address performance issues, are more easily and defensibly adaptable to procedural and technical than to clinical skills, and are probably more appropriate tools for professional specialty groups than for legal jurisdictions. Examination of the tasks of a professionally immature field such as psychology suggests that control processes other than recredentialing may better serve society and the profession as a whole for the regulation of the practitioner, although recredentialing may be an appropriate approach in technological or specialty areas. The profession's ability to provide valid assurance of professional outcomes competency is currently sufficiently limited in almost all areas of psychology that recredentialing may create more false sense of security than protection for the public. Its primary use in psychology may be to assure orthodoxy or competency in procedural skills or an underlying knowledge base in specialty or proficiency areas.

2.16.11 REFERENCES

Allen, D. L. (1994). A report on compulsory continuing dental education requirements for relicensure. *International Dental Journal, 44*(6), 637–640.

Allred, T. J., & Steiner, L. (1994). Alternate-site testing. Consider the analyst. *Clinics in Laboratory Medicine, 14*(3), 569–604.

Baldwin's Ohio administrative code (Vol. 11) (1997). St. Paul, MN: West Group.

Bennett, W. M., & Langdon, L. O. (1995). Recertification: a challenge to those trained in the "good old days." *American Journal of Kidney Diseases, 26*(3), 556.

Berven, N. L. (1987). Improving evaluation in counselor training and credentialing through standardized simulations. In B. A. Edelstein & E. S. Berler (Eds.), *Evaluation and accountability in clinical training* (pp. 203–229). New York: Plenum.

Brennan, A. (1992). Analysing mandatory continuing education. *Nursing Standard, 6*(42), 29–32.

Chambless, L. E., McMahon, R., Finch, A., Sorlie, P., Heiss, G., Lyles, R., & Wu, K. K. (1993). ARIC hemostasis study—III. Quality control. Atherosclerosis risk in communities. *Thrombosis & Haemostasis, 70*(4), 588–594.

Clunie, G. J. (1993). Postgraduate medical education—comparisons between the United Kingdom and Australia. *Journal of the Royal College of Physicians of London, 27*(2), 154–157.

Conn, R. B. (1994). Continuing clinical chemistry education in the United States. *Clinica Chimica Acta, 232*(1–2), 47–51.

D'Costa, A. G. (1986). The validity of credentialing examinations. *Evaluation and the Health Professions, 9*(2), 137–169.

Delahunt, B., Beer, I. D., & Taylor, D. E. (1992). Attitudes to specialist recertification: results of a national survey among pathologists. *New Zealand Medical Journal, 105*(947), 493–495.

Department of Health, Education, & Welfare (1977). *Credentialing health manpower* (DHEW Publication No. PHS 77/5005); (NTIS No. HRP-0024664). Washington, DC: DHEW.

Eagle, C. (1992). Anaesthesia and education. *Canadian Journal of Anaesthesia, 39*(2), 158–165.

Egerton, W. S. (1994). Health care delivery system in Australia and its effect on surgical education and training. *World Journal of Surgery, 18*(5), 662–665.

Goto, Y. (1994). Specialist training and certification process for anaesthesiologist in Japan. *Annals of the Academy of Medicine, Singapore, 23*(4), 630–632.

Grim, C. M., & Grim, C. E. (1995). A curriculum for the training and certification of blood pressure measurement for health care providers. *Canadian Journal of Cardiology, 11* (Suppl. H), 38H–42H.

Gross, S. J. (1978). The myth of professional licensing. *American Psychologist, 33*(11), 1009–1016.

Haag, G. P. (1992). Point: retesting should be required for CRNA recertification. *Nurse Anesthesia, 3*(1), 3–5.

Hall, J. E. (1987). Licensure and certification of psychologists. In B. A. Edelstein & E. S. Berler (Eds.), *Evaluation and accountability in clinical training* (pp. 253–281). New York: Plenum.

Hasbargen, J. A., & Culclasure, T. F. (1994). Special Forces Medical Sergeants (18 Delta) recertification. *Military Medicine, 159*(1), 7–9.

Holmes, S. E. (1986). Comments on this special issue. *Evaluation and the Health Professions, 9*(2), 131–136.

Howard, A. (1987). Work samples and simulations in competency evaluation. In B. A. Edelstein & E. S. Berler (Eds.), *Evaluation and accountability in clinical training* (pp. 55–76). New York: Plenum.

Johanson, W. G. (1995). The ABIM recertification program—nearing liftoff. *Chest, 108*(1), 1–2.

Koocher, G. P. (1979). Credentialing in psychology: close encounters with competence? *American Psychologist, 34*(8), 696–702.

Leigh, T. M., Young, P. R., & Haley, J. V. (1993). Performances of family practice diplomates on successive mandatory recertification examinations. *Academic Medicine, 68*(12), 912–919.

Matarazzo, J. D. (1977). Higher education, professional accreditation, and licensure. *American Psychologist, 32*(10), 856–859.

Mayfield, P. N. (1987). Certification of psychologists by the American Board of Professional Psychology. In B. A. Edelstein & E. S. Berler (Eds.), *Evaluation and accountability in clinical training* (pp. 283–297). New York: Plenum.

McClennan, B. L., & Herlihy, C. S. (1995). The continuing competence needs of physicians: a survey of the medical specialty societies. *American Journal of Roentgenology, 165*(4), 789–795.

Nelson, L. P. (1994). American Board of Pediatric Dentistry recertification section examination. *Pediatric Dentistry, 16*(5), 381–382.

Norcini, J. J. (1994). Recertification in the medical specialties. *Academic Medicine, 69*(10 Suppl.), S90–S94.

Pfister, S. M. (1995). Home oxygen therapy: indications, administration, recertification, and patient education. *Nurse Practitioner, 20*(7), 44, 47–52, 54–56.

Phillips, G. D., & Willis, R. J. (1994). Specialist training and examinations in anaesthesia and intensive care in Australia and New Zealand. *Annals of the Academy of Medicine, Singapore, 23*(4), 610–613.

Rodgers, D. A. (1964). In favor of separation of academic and professional training. *American Psychologist, 19*(8), 675–680.

Rodgers, D. A. (1980). The status of psychologists in hospitals: technicians or professionals. *Clinical Psychologist, 33*(4), 5–7.

Rodgers, D. A. (1981). A proposed model psychology licensing law. *Professional Practice of Psychology, 2*(1), 47–71.

Rodgers, D. A. (1986). Psychologists as practitioners, not technicians. In H. Dörken & Associates (Eds.), *Professional psychology in transition: Meeting today's challenges* (pp. 141–173). San Francisco: Jossey-Bass.

Rodgers, D. A. (1989). Mandatory CE solution with teeth: week's retreat at college campus. *Ohio Psychologist, 35*(6), 16–18.

Rodgers, D. A. (1993). When will psychology "psychologize" instead of "medicalize?" *The National Psychologist, March/April, 15*.

Rodgers, D. A. (1994, April 21). *A suggested Nosological Classification System for Clinical Psychology*. Paper presented at the convention of the Ohio Psychological Association, Columbus, OH.

Roessler, G. (1994). Information on certification and recertification of transcultural nurse specialists. *Journal of Transcultural Nursing, 6*(1), 46.

Ross, A. O. (1973). A rejoinder on behalf of ABPP. *American Psychologist, 28*(1), 88–89.

Saunders, N. A., & Paget, N. S. (1994). Maintenance of professional standards and the recertification of physicians. *Journal of Internal Medicine, 236*(4), 363–365.

Sechrest, L., & Chatel, D. M. (1987). Evaluation and accountability in training for professional psychology: an overview. In B. A. Edelstein & E. S. Berler (Eds.), *Evaluation and accountability in clinical training* (pp. 1–37). New York: Plenum.

Simon, A. (1992). Preparation for recertification. *Journal of Post Anesthesia Nursing, 7*(1), 59–61.

Stern, S. (1984). Professional training and professional competence: a critique of current thinking. *Professional Psychology Research and Practice, 15*(2), 230–243.

Stevenson, J. F., & Norcross, J. C. (1987). Current status of training evaluation in clinical psychology. In B. A. Edelstein & E. S. Berler (Eds.), *Evaluation and accountability in clinical training* (pp. 77–115). New York: Plenum.

Tonesk, X. (1987). Clinical evaluation of medical students in US medical schools. In B. A. Edelstein & E. S. Berler (Eds.), *Evaluation and accountability in clinical training* (pp. 183–202). New York: Plenum Press.

Wiens, A. N. (1983). Toward a conceptualization of competency assurance. *Professional Practice of Psychology, 4*(2), 1–15.

Wiens, A. N., & Dörken, H. (1986). Establishing and enforcing standards to assure professional competency. In H. Dörken & Associates (Eds.), *Professional psychology in transition: Meeting today's challenges* (pp. 174–199). San Francisco: Jossey-Bass.

Wigton, R. S. (1992). Training internists in procedural skills. *Annals of Internal Medicine, 116*(12 Pt. 2), 1091–1093.

Wilson, F. C. (1993). Credentialing in medicine. *Annals of Thoracic Surgery, 55*(5), 1345–1348.

Zarowitz, B. J., Vlasses, P. H., Dukes, G. E., Gal, P., Miller, W. A., Porter, W. C., Rush, D. R., Schneider, P., & Talbert, R. L. (1993). Pharmacotherapy specialty certification examination. IV. 1992 results and process modifications, including recertification. The 1992 Specialty Council on Pharmacotherapy, Board of Pharmaceutical Specialties. *Pharmacotherapy, 13*(3), 262–266.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.17

Continuing Education: Updating of Knowledge and Skills

EDWARD F. BOURG
California School of Professional Psychology, Alameda, CA, USA

2.17.1 INTRODUCTION 196

2.17.2 DEFINITION AND SPECIFICATION OF CONTINUING PROFESSIONAL EDUCATION IN CLINICAL PSYCHOLOGY 196

2.17.3 THE EFFECT OF THE EXPONENTIAL GROWTH IN KNOWLEDGE AND EXPANSION OF SKILLS IN CLINICAL PSYCHOLOGY ON CPE 197

- *2.17.3.1 The Shrinking Half-life of Graduate Professional Education* 197
- *2.17.3.2 Differentiation within the Field of Clinical Psychology* 197
- *2.17.3.3 CPE for Proficiencies* 198
- *2.17.3.4 CPE for Specialties* 198

2.17.4 HEALTH CARE REFORM AND CPE 199

- *2.17.4.1 CPE as a Cost Center* 199
- *2.17.4.2 CPE for Expanding Roles in Clinical Psychology* 199

2.17.5 SHAPING CPE FOR THE ACTIVE PRACTITIONER 200

- *2.17.5.1 More Sophisticated Needs Assessments* 200
- *2.17.5.2 CPE for the "Reflective Practitioner"* 200
- *2.17.5.3 Active-learning and Skills-oriented CPE* 201
- *2.17.5.4 CPE for Diverse Learning Styles* 201

2.17.6 VARIED AND INNOVATIVE VENUES FOR CPE 201

- *2.17.6.1 Informal and Undocumented CPE* 201
- *2.17.6.2 Independent Study CPE* 202
- *2.17.6.3 Teleconference and Videoconference CPE* 202
- *2.17.6.4 Computer-assisted CPE* 202
- *2.17.6.5 In-service CPE* 203

2.17.7 QUALITY CONTROL IN CPE: AN INSTRUMENT FOR ENHANCING QUALITY CONTROL IN PRACTICE 203

- *2.17.7.1 Mandatory CPE* 203
- *2.17.7.2 Beyond Sponsor Approval to Program Review* 204

2.17.8 FUTURE DIRECTIONS: A LONG-TERM RESEARCH AGENDA FOR CPE 204

- *2.17.8.1 A Model: Research in Continuing Medical Education* 205
- *2.17.8.2 Application to a Research Agenda for Clinical Psychology* 205
- *2.17.8.3 Key Components of Effective CPE* 205

2.17.9 SUMMARY 206

2.17.10 REFERENCES 207

2.17.1 INTRODUCTION

An essential element of a profession is that its members embrace the concept of lifelong learning, namely, a commitment to stay knowledgeable about developments in their given field of practice (Peterson, 1976). Continued professional education characterizes mature professions, most importantly, because professions ultimately exist to serve the public welfare and the public has a vested interest in assuring that individual practitioners have up-to-date knowledge and skills. The public has a right to know, for example, that a clinical psychologist treating severe depression is aware of the latest advances in pharmacotherapy for this illness. For almost all practicing psychologists such detailed knowledge could be available only from continuing professional education, since research findings confirming the biological determinants of depression, improved pharmacological treatments for severe depression, and more effective psychosocial treatments of depression are themselves relatively recent.

A commitment to such lifelong learning has been a hallmark of clinical psychology. It is reflected in the proceedings of the various national conferences which have set the standards for education and training in clinical psychology since the late 1940s (Bourg et al., 1987; Korman, 1974; Peterson et al., 1992; Raimy, 1950). The imperative to engage in continued professional education is also reflected in the American Psychological Association (APA) *Ethical Principles of Psychologists and Code of Conduct* (APA, 1992). The enrollment of clinical psychologists in formal continuing education programs, documented in several studies (Brown, Leichtman, Blass, & Fleisher, 1982; McNamara, 1977), is an important method for fulfilling that professional and ethical mandate.

2.17.2 DEFINITION AND SPECIFICATION OF CONTINUING PROFESSIONAL EDUCATION IN CLINICAL PSYCHOLOGY

The term "continuing education" (CE) generally refers to education for the adult learner and currently is used to describe a very large array of adult learning experiences reflected, for example, in the diverse offerings of the "extension division" of major universities. For the professional, an extensive literature around continuing education emphasizing themes such as quality assurance, effectiveness, and methods of teaching has been developed. In the context of a better educated workforce, continuing professional education has been described as follows:

> Education to maintain and improve professional competence is called Continued Professional Education, and is a special subset of the broad field of adult CE. As such it is part of the most rapidly growing segment of American education, one of importance to our society and economy, and one symbolizing profound changes in the way we approach education, work and the relationship between the two. (Hunt, 1992, p. 1)

This chapter about continuing professional education in clinical psychology will consider the major questions within the field for professional psychologists, more specifically clinical psychologists. Although the more general term CE has been the customary nomenclature for such learning experiences, it would enhance public trust for clinical psychology to adopt a more professional referent for these learning experiences. Accordingly, the term "continuing professional education" (CPE) will be used in this chapter except when referring to specific institutional or organizational structures currently designated under the current CE rubric.

CPE has been defined for health care providers by the American Hospital Association (AHA) as follows: " . . . a planned series of activities that enables an individual to acquire the skills, knowledge, and behavior necessary to meet current job requirements or remedy identified deficiencies" (AHA, 1978, p. 2).

Several aspects of this definition are especially relevant for CPE in clinical psychology. First, the emphasis on planning suggests that the individual practitioner should attend CPE programs that fit within the context of a planned self-directed curriculum for development congruent with one's professional life cycle. Also, CPE providers should follow principles of good learning design in developing courses as organized sequences of learning and programs which differentiate the needs of learners. Second, CPE courses should encompass both knowledge-based objectives and skills training; in the field of clinical psychology, as in most professions, insufficient attention has been given to skills training. Finally, the focus on remediation suggests that, for participants and providers alike, considerable attention should be given to cogent, problem-oriented needs assessment.

The concept of lifelong learning for practitioners after the graduate degree includes the whole spectrum of a professional career which now spans 40–50 years of practice. Recently there has been emphasis on specifying a more coherent education and training experience immediately following the doctoral degree (Stewart, Horn, Becker, & Kline, 1993). In a recent survey, a substantial number of postdoctoral programs were identified and described (Wiens & Baum, 1995). Widespread discussion has included a national conference focused on practitioner programs (Association of Psychology Postdoctoral and Internship Centers, 1993); another national conference sponsored by the APA in January 1994, which included discussion and recommendations about postdoctoral education for all practitioner roles (APA, 1995); and the development of criteria and procedures by the APA Committee on Accreditation. In August 1996 the APA Council of Representatives approved criteria for the accreditation of postdoctoral programs for practitioners. "Postdoctoral residency" is the preferred title for these training experiences and in almost all states such postdoctoral training is required for licensure. A professional ethic has clearly evolved in which the basic formal education and training of a professional psychologist in general, and a clinical psychologist in particular, should include the postdoctoral residency. In that sense CPE for clinical psychology really applies to programs of further training after the postdoctoral residency and such programs offered for clinical psychologists in active practice will be the focus of this chapter.

2.17.3 THE EFFECT OF THE EXPONENTIAL GROWTH IN KNOWLEDGE AND EXPANSION OF SKILLS IN CLINICAL PSYCHOLOGY ON CPE

2.17.3.1 The Shrinking Half-life of Graduate Professional Education

The field of clinical psychology flourished in the United States after World War II, partly in response to the need for trained mental health practitioners to deal with the psychological sequelae of combat. The alliance formed among the National Institute of Mental Health, the Veterans Administration, and clinical psychology programs in universities established the basic model for clinical training (Humphreys, 1996). In the years following the war, except for the recent focus on the postdoctoral residency noted above, attention to standards in the education and training of clinical psychologists has been focused primarily on the quality of doctoral programs in clinical psychology.

Yet, given the expansion in knowledge and skills in the field of clinical psychology, the half-life of graduate education is continually shrinking. This concept refers to the estimate that the original body of knowledge learned in a doctoral program will be only one-half as useful in a given period of time, the "half-life," as when the doctorate was originally awarded. In the early 1970s Dubin estimated that the half-life of a doctorate in psychology was 10–12 years (Dubin, 1972). Twenty years later, the participants in the 1994 APA National Conference on Postdoctoral Education and Training in Psychology estimated the half-life of graduate education in psychology to be approximately eight years (APA, 1995). We can expect this trend to continue and lead to greater reliance on CPE across the professional life span of the practitioner. Some authors predict that CPE will eventually become the most important part of the education and training of professionals (Houle, 1980; Hunt, 1992). In the 1993 World Summit on Medical Education, particular emphasis was placed on the importance of continuing education to meet the challenges facing medicine in the twenty-first century: "Continuing Education is the single main process by which medical services can be reorganized, reprogrammed, and made relevant to the contemporary needs of societies" (Walton, 1995, p. 43).

In 1994 the APA sponsored the first national conference broadly focused on postdoctoral education and training in psychology with discussions across the range of roles of psychologists as teachers, researchers, and practitioners. This conference was also inclusive in specifically extending its deliberations to the lifelong education and training of all psychologists. It explicitly included discussion about issues and problems in the CPE of practitioners, including clinical psychologists (APA, 1995) and heralded a shift in emphasis on the paramount "locus" of professional education and training for practitioners. It is easy to imagine that in the not too distant future a relatively senior clinical psychologist will have acquired most of the knowledge and skills used in professional practice through CPE.

2.17.3.2 Differentiation within the Field of Clinical Psychology

A second implication for CPE resulting from the exponential growth in knowledge and the expanding skill base in clinical psychology is increased specialization in professional psychology in general and clinical psychology in particular. Such differentiation is characteristic of well-established professions with an expanding knowledge and skill base, and the demand for innovative opportunities for in-depth CPE programs in which clinical psychologists can learn new specialties and/or subspecialties is a clear consequence. After struggling with the need for standards and procedures to guide this evolution in professional psychology, a consensus has recently been established that further specialization in professional psychology should be guided by a coherent set of policies and procedures. In February 1995 the APA established the Commission for the Recognition of Specialties and Proficiencies in Professional Psychology (CRSPPP) with profound implications for CPE in psychology. One of the CRSPPP standards for seeking recognition as a separate specialty or proficiency is clearly established CPE criteria and programs in a particular field of professional practice. The principles approved for CRSPPP establish three broad parameters for distinguishing among specialties and subspecialties (proficiencies) which have implications for CPE: populations, problems, and theory and techniques.

Not surprisingly, most of the subfields associated with differences in each of these three categories are represented by organized interested groups and/or established divisions within APA. For example, in the area of distinct populations served, applications are likely to be submitted to CRSPPP for specialty or proficiency recognition for women, children, older adults, and ethnic minorities. Many CPE programs are already being offered that directly relate to the specific prevention and treatment needs of each of these populations (Crespi & Rigazio-Digilio, 1992; Myers, Wohlford, Guzman, & Echemendia, 1991; Panayotoff, 1993). As these, and similar areas of practice, continue to evolve and differentiate, more diverse CPE programs will be needed.

Similarly, as psychologically based prevention and treatment programs evolve for both long-standing problems (e.g., serious mental illness) and newer problems (e.g., AIDS), more in-depth CPE programs will be needed (Robiner, Ohnsong, Parker, & Strike, 1994). The APA HOPE program, a "train-the-trainer" CPE effort for the prevention and treatment of AIDS, is a model for an extensive CPE program that can be made available throughout the country for the prevention and treatment of a specific problem. Similarly, a program, jointly sponsored by APA and many state associations, for training a large cadre of psychologists across the country capable of responding in disaster relief teams, is an impressive example of a large-scale CPE program focused on a specific psychological problem of national importance.

Finally, further refinements in more traditional theories (e.g., psychoanalytic, behavioral, or family systems) and newer theoretical subfields (e.g., self-psychology, dialectical behavior therapy, or narrative therapy) will also affect existing CPE programs or spawn new CPE programs for educating and training practitioners. Again, as these or other subfields reach the status of specialties or proficiencies in professional psychology, practitioners will seek CPE programs both to inform themselves as responsible professionals and to seek credentials.

2.17.3.3 CPE for Proficiencies

Both of these motivations have contributed to the recent development of the APA College of Professional Psychology, charged with credentialing individual psychologists in clearly delineated subspecialties or proficiencies. Credentials in these areas are likely to have implications for professional reimbursement in the managed care environment. For example, the College has already established a system for psychologists to qualify for credentials in the field of alcohol and substance abuse. Although the College is itself specifically prohibited from offering CPE programs, CPE providers are already designing programming in this area to assist practitioners in meeting the educational criteria and learning the content specified by the College for both initial certification and periodic recertification. Similar CPE demand and development in other areas in which the College will offer credentialing for proficiencies in professional psychology can be anticipated.

2.17.3.4 CPE for Specialties

Similarly, the various specialties that have *de facto* recognition by the American Boards of Professional Psychology (ABPP) already have elaborate education and training criteria for practitioners who seek credentials in these areas. In the late 1990s ABPP offered diplomate status in nine areas: behavioral psychology, clinical psychology, neuropsychology, counseling psychology, family psychology, forensic psychology, health psychology, industrial organizational psychology, and school psychology. All these specialties and any future specialty area to be recognized by CRSPPP, and any specialty certificate system that recognizes the expertise of individual providers, will need to ensure that CPE programs designed to assist practitioners in such specialty training are provided.

One example of the exponential growth of knowledge and expansion of skills that opens a new avenue for specialized practice in clinical psychology is psychopharmacology. In August 1996, the APA Council of Representatives approved the development of appropriate education and training programs for the authority to prescribe psychotropic medications within an expanded scope of practice for psychologists who are health care providers. Although some general guidelines for such training programs have been promulgated by APA, the development of adequately trained psychologists in such a specialty will call for innovative, carefully sequenced, in-depth CPE programs.

2.17.4 HEALTH CARE REFORM AND CPE

In almost all developed countries the years from the late 1980s have seen a revolution in health care systems that is driven by both political and economic forces. A basic focus in health care reform is the general concern for competent, consistent quality care at a price that immediate payers, such as insurance companies, and intermediate payers, such as employers, are willing to expend (Azzaretto, 1992; Cummings, 1995). The public policy questions surrounding this development are immense. For example, the United States spends more per capita on health care costs than almost any other developed country, and its competitiveness in the evolving global economy is thereby compromised. Also, in an era of fiscal constraint and limited resources, the public must ultimately determine what it is willing to allocate for health care in the context of other pressing economic and social demands.

2.17.4.1 CPE as a Cost Center

These developments impact CPE on a number of levels. First, CPE is itself an integral part of the health care system. As such, in spite of its manifest usefulness in updating health care professionals, CPE must be subject to the same norms of accountability as other health care practices. From the point of view of the managed care company, CPE is a cost center and subject not only to questions of effectiveness in general but cost-effectiveness in particular (Slotnick, Lichtaver, & Raskowski, 1995). With health care reform there is increased emphasis on interdisciplinary practice and therefore a mandate for CPE programs that emphasize such practice. For example, recent studies in the treatment of the seriously mentally ill call for greater interdisciplinary care (Johnson, 1990). Recent breakthroughs in understanding the biological basis for these illnesses and advances in psychopharmacological treatment have led to greater emphasis on collaborative practice between clinical psychologists and psychiatrists as well as primary care physicians, who prescribe the majority of psychotropic medications in the United States (APA, July 1992). Also, the emphasis on case management of clients with a dual diagnosis and the seriously mentally ill calls for closer collaboration with social workers, nurses, and substance abuse counselors.

2.17.4.2 CPE for Expanding Roles in Clinical Psychology

In the evolving health care system, in both the private (including profit and nonprofit organizations) and the public sector, the emphasis on cost-effectiveness associated with managed care encourages the utilization of professionals with less education and training than clinical psychologists to provide direct service. Accordingly, in the context of health care reforms, clinical psychologists are likely to need CPE to broaden their roles beyond the direct provision of mental health services, the activity with which they have become most identified since the late 1960s (Cummings, 1995; Humphreys, 1996).

In order to function in the evolving health care system, many clinical psychologists will need retraining to develop their skills in clinical and administrative supervision, program development and administration, and designing community prevention programs (Broskowski, 1995; Cummings, 1995; Humphreys, 1996). One important role for clinical psychologists in a managed care environment emphasizing cost-effectiveness and accountability will be in the design and implementation of outcome research. Another opportunity to broaden the practice of clinical psychologists may be in the burgeoning field of health psychology, certainly an allied, and many would argue a derivative, field of clinical psychology (Matarazzo, 1987).

Ultimately, clinical psychologists may begin to focus on an even broader spectrum of behaviors and settings which promote health, productivity, quality of life, and a greater sense of community. As this occurs, the need for CPE programs that emphasize the interdisciplinary training of clinical psychologists will expand beyond merely the inclusion of other health care providers. Increasingly those psychologists with specialized training in fields such as neuropsychology and forensic psychology are working with professionals outside the health care field, for example, in the areas of child custody and worker's compensation. Similarly, as clinical psychologists move to more community systems perspectives, they will need to work more collaboratively with other professions such as educators, architects, organizational consultants, and urban planners. CPE programs for clinical psychologists will be required to facilitate such collaboration.

2.17.5 SHAPING CPE FOR THE ACTIVE PRACTITIONER

Based on several important longitudinal studies, psychologists have led the way in clearly differentiating distinct stages in the adult life cycle. Parallel developments have occurred in adult learning theory and application. Providers of CPE for professionals have developed a relatively coherent set of criteria for the adult learner that take account of both adult learning principles and the specific circumstances of the professional work place (Jennett & Swanson, 1994; Mann, 1990; Scott, 1994). For example, for the health care provider relevant criteria include active learning methods, more case specific teaching, and CPE format design which takes into account multiple professional demands (Slotnick et al., 1995).

2.17.5.1 More Sophisticated Needs Assessments

Some principles of adult learning have been incorporated into criteria for CPE programs for clinical psychologists. For example, program offerings based on needs assessments and the specification of learning objectives, both fundamental adult learning practices, have been incorporated in the APA Committee for the Approval of Continuing Education Sponsors (CACES) criteria. Yet, needs assessment for CPE programs for professional psychologists can improve by drawing on both practice and research in CPE in allied health professions. For example, a more complex needs assessment, offering a model for psychology, has recently been conducted by the American Speech, Language, and Hearing Association (L. Phillips and S. Russell, personal communications, August 9, 1996). The goal was to develop CPE courses which clearly distinguished between needs (based on identified problems in practice) and wants (often related to "hot" topics or prominent presenters). This model, utilizing focus groups from the field, recognized that the more traditional method of devising and/or approving CPE programs via a small committee of educators and/or practitioners is inevitably hampered by the scope of their restricted view and decisions based on the limited program submissions that emerge from an "open call" for CPE programs.

2.17.5.2 CPE for the "Reflective Practitioner"

In developing CPE programs for active, especially more senior, practitioners, the concept of the "reflective practitioner" is especially relevant. Schön uses the term to describe a professional making use of a body of knowledge and professional skills in approaching a specific class of society's problems within its domain (Schön, 1983, 1987). This concept has special relevance in clinical psychology in which the major instrument of change is the self and the dictum "healer know thyself" has added meaning.

The reflective practitioner concept is useful in framing a theoretical paradigm for CPE because of the nature of any profession like clinical psychology which is based on science. On the one hand, the basic paradigm of "applied science" is applicable in the sense that applications are based on theory substantiated through research, and practice must inform science by providing the basic clinical datum on which more controlled research studies can be undertaken. On the other hand, given the specific problems individual clients actually present in the complex multifaceted context of the health care system, clinical applications must of necessity "run ahead" of science (Peterson, 1991). Science is pure and allows for absolutes while practice is client-specific, presenting unique issues from case to case. Clinical interventions inevitably involve professional judgment that goes beyond the straightforward application of specific imperatives derived from the science base of psychology. The reflective clinical psychologist is by definition engaged in a continuous professional education paradigm since each client encounter is both a professional service and a subject for continued professional development.

In this context, how does the ongoing learning of the clinical psychology reflective practitioner proceed? Progress in professional knowledge and skills is based on the development of professionals who can reflect across the spectrum of basic values, theory, strategies of intervention, and specific techniques. The first task of the clinical psychologist is always problem setting because they are continually confronted by uniqueness in applying the body of knowledge and skills accumulated. In clinical psychology, the unique features of a case may arise from the larger social, cultural, or familial context which surrounds the client, the specific nature and quality of the presenting problem(s), and/or in the fundamental fact of "individual difference" which every person brings to the clinical encounter; all give rise to conflict and ambiguity with respect to the "correct" intervention.

According to Schön (1987), reflective practitioners move through five stages in their attempt to learn from experience: knowing in action, surprise, reflection in action, the experiment, and reflection on action. Schön developed his ideas with reference to the micro-learning environment of the individual practitioner who is engaged in a continuous process of informal CPE. The concept of the reflective practitioner also has great promise at the macro level of CPE program design and research on the effectiveness of CPE (Fox, 1991; Slotnick et al., 1995). For example, in the context of learning a particular theoretical refinement, intervention strategy, or new technique, reflective clinical psychologists can focus on one or more of the five processes in which they are engaged with specific cases to become more effective in professional judgments and performance: improving one's knowing in action, sharpening one's ability to identify professional surprises, increasing one's effectiveness in reflection in action while caring for the client, designing more effective and appropriate experiments, and improving one's ability to integrate experiences through reflection in action. In designing CPE programs for clinical psychologists, explicit attention can be given to addressing one or more of the five stages in the learning process.

2.17.5.3 Active-learning and Skills-oriented CPE

CPE faculty can also make use of the reflective practitioner paradigm in designing instructional methods to capitalize on this ongoing process by participant learners. While traditional lectures are effective for communicating specific new information, CPE courses are too reliant on traditional knowledge dissemination. Practitioners bring a rich wealth of material from their ongoing reflection process to every CPE course. The literature on CPE across professions is increasingly focusing on a "new paradigm for CPE" which emphasizes more use of active-learning principles and techniques (Moore, Green, Jay, Leist, & Maitland, 1994; Queeney, 1992). Cervero (1988) has suggested a three-step process of allowing the professional first to see, then to try, and then to reflect on the knowledge and skills taught that is more consonant with the practitioner's usual learning style in the practice context. This process facilitates learning through the social construction of professional acumen within a community of practitioners which is both more respectful of and more responsive to the practitioner's felt needs.

Moreover, active-learning instructional methods better capture the complexity involved in any clinical intervention. The practitioner draws from a rich store of basic values, theoretical paradigms, and intervention strategies in making specific technical interventions. Formal CPE programs would be well advised to address more directly the interplay between basic values, theory, intervention strategies, and specific techniques. CPE programs for clinical psychologists are too often focused on theory, whereas most practitioners operate along the more practical end of the spectrum dealing with intervention strategies, specific techniques, and even professional "moves," that is, concrete verbal and nonverbal behaviors (Shapiro & Reiff, 1993). Greater direct relevance to the practitioner's life is gained in courses taught from a problem-based methodological paradigm (Engel et al., 1992; Moore et al., 1994; Slotnick et al., 1995).

2.17.5.4 CPE for Diverse Learning Styles

Finally, in designing programs for active practitioners, more account should be taken of the fact that practitioners bring diverse learning styles and interests to CPE which cut across specific content areas. For example, both gender-based ways of knowing and culturally responsive approaches to learning should be considered in the teaching styles and learning methods utilized (Briscoe, 1990; McGowen & Hart, 1990). Also, clinical services for women and ethnic minority populations need special attention in CPE programs to retrain clinical psychologists in response to changing demographic realities (Lerman & Porter, 1990; Myers et al., 1991).

2.17.6 VARIED AND INNOVATIVE VENUES FOR CPE

2.17.6.1 Informal and Undocumented CPE

It is important to recognize that much important lifelong learning by the active practitioner goes on through informal CPE vehicles. Perhaps the largest informal CPE provider for clinical psychologists in the United States is the APA Office of Publications. Their expansion of computer-assisted learning, through such services as PsycLit and PsycInfo, has greatly eased access to this and related information sources. The APA Office of Publications has launched its first electronic publishing venture with a videotape series teaching psychotherapy skills. Although documentation is scant, clinical psychologists appear to make extensive use of other informal channels of CPE, for example, case supervision provided by senior clinicians, peer supervision groups, in-service training in health care settings, and attendance at professional conferences. The APA CE Programs Office has also provided informal CPE by offering audiotapes of the best-attended sessions at the annual APA Convention.

While some other allied health care professions, for example, pharmacy, provide formal CPE credit for attendance at large conferences and meetings (B. Hammonds, personal communication, August 8, 1996), clinical psychology has exercised more caution with respect to awarding formal CPE credit because of the difficulty in assuring appropriate monitoring. However, many psychologists utilize such forums for informal CPE, thereby gaining access to a broad range of ideas from which they can make more informed decisions regarding future investment in formal CPE study. Smaller, focused conferences are increasingly being deemed appropriate venues for formal CPE credit. For example, clinical psychologists attending the Annual State Leadership Conference, sponsored by the APA Practice Directorate, can gain CPE credit through a carefully planned and monitored system of programs teaching knowledge and skills relevant to contemporary practice.

2.17.6.2 Independent Study CPE

Independent study CPE programs, when carefully planned and monitored, are increasingly available for CPE credit (Woolfort, Lang, Farghaly, Ziemiecki, & Faja, 1991). The APA CE Committee, in conjunction with the APA Office of Publications, has a well-developed independent study program that has grown rapidly since the late 1980s. In 1996, APA's independent study programs expanded to include the Clinician's Research Digest, an abstract of research articles relevant to the practice of clinical psychology. Other CPE providers offer print-based independent study programs, and increasingly, audiotape- and/or videotape-based independent study programs are available. This rapid expansion of independent study programs in clinical psychology will increase exponentially into the twenty-first century with breakthroughs in computer-assisted and distance learning.

2.17.6.3 Teleconference and Videoconference CPE

Some allied health professions have already developed programs based on new technologies (Engel et al., 1992; Hampton, Mazmanian, and Smith, 1994) For example, since the early 1990s the American Speech, Language, and Hearing Association (ASHA) has sponsored a series of national teleconferences, that is, telephone conference calls linking CPE learners across the country. Although this format only allows for limited interactivity in the approximately 40 sites (some accommodating hundreds of participants) selected, it does provide a user-friendly and relatively inexpensive vehicle for disseminating specific new information. ASHA has also pioneered several national videoconferences, concluding that such programming is best done in partnership with established national videoconference vendors in which the association takes primary responsibility for curriculum and faculty, the videoconference expert takes responsibility for all technical logistics, and both share in marketing efforts and profits (S. Russell, personal communication, August 9, 1996).

CPE for other allied health professions is increasingly being provided through preproduced videotapes and over closed circuit television. For example, one for-profit provider has established a national hospital network which delivers ongoing medical education for physicians and nurses in regularly scheduled programming throughout the day. In May 1994 the APA CE Committee and the National Council of Schools and Programs in Professional Psychology pioneered a national CPE videoconference for clinical psychologists on revisions in the newly published *DSM-IV*. This venue proved ideal for delivering such a relatively objective body of information with which all clinical psychologists must be conversant. However, it is not yet clear that clinical psychologists see this vehicle, even with an interactive format, as comparable to face-to-face contact with faculty when dealing with more complex clinical knowledge and skills.

2.17.6.4 Computer-assisted CPE

Another promising new venue for CPE, already being developed in several allied health professions, is the use of computer-assisted learning both through the use of a CD-ROM format or an Internet format that permits regular updating. As noted above, APA has pioneered informal CPE with the CD-ROM format by providing access to a very large data bank in the form of abstracts from a pool of books and articles. With the development of high-end multimedia capability, however, it is clear that we have the capacity to develop sophisticated interactive CPE learning vehicles. In fact, a number of such programs have been developed for undergraduate courses in psychology and for CPE courses in allied health fields (Dillon, 1996). The main obstacle for this venue is that front end production costs are very high and with rapidly changing knowledge and skills, it is difficult to judge whether the expense of such production costs are warranted.

Direct on-line CPE programming through the Internet allows for regular updating of the material presented but production costs, limited access, and lack of sophisticated computer skills by active practitioners is currently problematic. However, the advantages of such a learning vehicle, already being pioneered in some professional certificate programs other than psychology, are clear: (i) no travel time is required; (ii) learning at one's own pace is facilitated; (iii) faculties are capable of both closely monitoring what is being learned by each student and effectively working in both a classroom and tutorial framework simultaneously; (iv) utilization of asynchronous learning with each learner having complete flexibility for presence "in the classroom" when personal and professional availability allow for optimal attention; (v) the possibility of ongoing assessment is maximized when faculties are able to assess and address each individual learner's pace and mastery. Other than planning and production costs, the major disadvantage of such a CPE venue in a field like clinical psychology where subtle verbal and nonverbal cues are so important, is lack of direct face-to-face contact between faculty and learners and among colearners.

2.17.6.5 In-service CPE

Until recently most CPE providers have focused on attracting individual practitioners. With the expansion of managed care in both the private and public sector, doubtless more attention will be given to in-service training (Hargrove, 1992; Marsick, 1990; Presland, 1993). Such programs have the advantage of being both cost-effective and providing the opportunity to attend to individual as well as systemic issues in professional practice. Behavioral health care organizations are now offering such CPE programs for managed care companies who recognize that this training is beneficial for their providers and ultimately their patient subscribers (Lynton, 1992).

2.17.7 QUALITY CONTROL IN CPE: AN INSTRUMENT FOR ENHANCING QUALITY CONTROL IN PRACTICE

CPE will be of increasing importance in clinical psychology because the evolution of the health care system in the developed countries, including the United States, is placing increased emphasis on quality control mechanisms. It is evident that greater accountability will be required of all clinical psychologists (Stricker & Rodriguez, 1988). The three primary methods generally used to ensure competence among professionals are retesting for license renewal, peer review, and mandatory CPE. Of the three methods, the first generally is opposed by professionals and has not been well received in surveys among licensed psychologists (Hellkamp, Imm, & Moll, 1989). Peer review is to some extent incorporated in evolving practice settings, for example, as part of review for hospital privileges for clinical psychologists. CPE, however, is by far the most common method for upgrading the knowledge and skills among health service providers, including clinical psychologists. CPE incorporates peer review, at least in some instructional formats, and provides both an incentive and a structure for maintaining competence in clinical practice.

2.17.7.1 Mandatory CPE

Since the late 1970s the advantages and disadvantages of mandatory continuing professional education (MCPE) have been contested across a broad range of professions. Overall, the practice seems to be clearly gaining acceptance (Garganta, 1989; Phillips 1994b; Queeney & English, 1994) and this has been the trend with licensed psychologists (VandeCreek, Knapp, & Brace, 1990). With a substantial increase in jurisdictions requiring MCPE since the late 1980s, currently 39 of the 61 jurisdictions represented by the Association of State and Provincial Psychology Boards (ASPPB) in the USA and Canada have adopted some form of MCPE and four others are actively considering the practice (ASPPB, 1996). One obvious place to begin assessing quality control in MCPE is to determine what similarities cut across the 39 jurisdictions which have mandated CPE for relicensure; unfortunately there is a wide range of variation among the jurisdictions regarding both the number of hours required for relicensure (10–50 credits per year) and the kind of MCPE activities which various jurisdictions allow under their MCPE relicensure provisions.

A Task Force on Model State and Provincial CE Regulations of the Board of Educational Affairs of the APA, after reviewing the existing statutes, issued the following summary recommendations as a national standard for the USA:

(i) the current modal number of 20 CPE credit hours per year for relicensure is a reasonable guideline for MCPE;

(ii) the APA CACES mechanism for sponsor approval should be recognized by states;

(iii) the current criteria of the APA CACES should be used by jurisdictions which establish independent criteria for CPE approval status;

(iv) the current APA CE Committee criteria (e.g, clearly articulated learning objectives based on adequate needs assessments and assessment of program quality by an independent panel of experts) should be used by jurisdictions which wish to adopt criteria for specific approval of every CPE activity;

(v) a clearinghouse should be established by one or more national organizations responsible for credentialing psychologists in the health care field to foster simplification of tracking compliance with MCPE regulations; and

(vi) participants should maintain their own records of MCPE attendance for some years and regulatory agencies should conduct random audits of such records in assessing compliance with MCPE (APA, October, 1995).

2.17.7.2 Beyond Sponsor Approval to Program Review

Most states accept programs offered by CPE sponsors approved by the APA CACES for licensure renewal. There has been concern, however, during the 1990s about whether a sponsor approval system, however refined the criteria and review process, can sufficiently guarantee the quality of each program offered by approved sponsors. Special concern attends to the practice of cosponsoring activities with institutions which are not part of the sponsor approval system. Although the APA CACES criteria clearly specify involvement in the planning of such activities by the approved cosponsor, it is difficult to monitor compliance. In the context of using MCPE as the major method of assuring quality control among clinical psychology practitioners, this concern has added significance.

One innovative response to this specific concern about reliance on a general sponsor approval system deserves attention and further monitoring. In January 1996, a new MCPE regulation took effect in California requiring that all CPE courses acceptable for relicensure must be individually reviewed and approved by an independent committee established by the Psychology Association Accrediting Agency (Board of Psychology, 1995). The Board of Psychology has revised its regulations to include programs which incorporate an equivalent individual program review process by other professional groups, for example, CE courses sponsored by the APA CE Committee.

One of the major concerns in this new system is the cost and complexity of maintaining such a detailed approval structure. Wisconsin, after maintaining a similar system for some years, moved to the simpler sponsor approval system (A. Pacht, personal communication, August 14, 1996). On the other hand, tighter course by course regulation is under consideration by two other state psychology boards. There is some concern about how well such a system will effectively and fairly control quality across CPE programs. For example, the accrediting agency in California understandably relies heavily on "primary reviewers" of each program and there has been no report about inter-rater reliability.

2.17.8 FUTURE DIRECTIONS: A LONG-TERM RESEARCH AGENDA FOR CPE

One concern raised about MCPE as a measure of assuring quality control is the limited empirical evidence for the effectiveness of CPE in improving the performance of professionals such as clinical psychologists (Phillips, 1994b; Queeney & English, 1994). More importantly, carefully controlled studies have not established that CPE actually improves the health of clients served by clinical psychologists. The effectiveness of CPE in various professions, especially allied health professions, has been a matter of concern over the 1990s. There is, however, growing evidence that CPE does favorably impact performance when CPE programs are well designed according to a sound program development model (Fox 1991; Moore et al., 1994; Phillips, 1994a, 1994b; Raymond, 1991). Moreover, methods for determining program quality are gradually improving as various health care professions work to address the issue of quality assurance in CPE (Newcomer, 1989). A key strategy, albeit an expensive and arduous undertaking, is to simultaneously target various levels in the CPE enterprise in an evaluation process: CPE sponsors, CPE faculty, CPE learning methods, CPE curricula design, learners' acquired knowledge and skills, and improved competence in practice.

In all professions, it is difficult to collect evidence in support of CPE's effect on competence because of the large number of variables affecting performance and the difficulty of designing controlled studies which isolate the impact of specific CPE interventions. Clinical psychology will have to establish a sophisticated research program to determine the effectiveness of CPE. Ultimately the question about CPE's effectiveness will probably have to be addressed in the same way that clinical psychologists addressed the question of psychotherapy's effectiveness. The global question will have to be framed in more specific terms: what kinds of CPE programs are effective for learning what kinds of clinical psychology skills, for what kinds of problems?

2.17.8.1 A Model: Research in Continuing Medical Education

In the health care field, the most impressive research agenda relating to the effectiveness of CPE programs has been undertaken with respect to CPE for physicians (Beaudry, 1989; Davis & Fox, 1994; Davis, Thomson, Oxman, & Haynes, 1995; Fox, Mazamian, & Putnam, 1989; Haynes, Davis, McKibbon, & Tugwell, 1984). A systematic review of the voluminous literature on the effect of Continuing Medical Education (CME), based on a multiyear research project, clearly demonstrated how well advanced the research agenda is on effective CME (Davis et al., 1995). The selection criteria for the study were very strict by current standards in the field of CPE for clinical psychologists: (i) only randomized control trials of CME programs; (ii) clearly replicable educational interventions; (iii) physicians had to be in the majority in the CPE programs included; and (iv) objective measurements of physician performance or health care outcomes had to be present. Using these criteria, the authors selected only 99 studies, which included 160 educational interventions, from an impressive database of over 6000 studies on the broadly defined topic of effective CME programs.

The most effective CME interventions included patient-mediated strategies, peer discussion, audit with feedback, and practice rehearsal. The study concludes that with such carefully controlled and objectively measured CME interventions, physician performance, and to a lesser extent health care outcomes, may be positively altered. The authors emphasize that their selection criteria systematically excluded the results of qualitative research method " ... that would afford additional valuable insights into physician behavioral change" (Davis et al., 1995, p. 703). It should not be surprising, given what clinical psychologists know from learning theory, that the most effective educational interventions emphasized practice-enabling strategies (e.g., role playing), peer discussion, and reinforcing learning methods. Variation in effectiveness was also attributable to motivation, especially to individual readiness to change based on the practitioner's clear perception of a gap between need and current knowledge and skills. In addition, CME programs were more effective in impacting physician performance when barriers to changing behaviors were systematically addressed.

2.17.8.2 Application to a Research Agenda for Clinical Psychology

Probably the most unsettling finding for clinical psychologists from the Davis et al. (1995) study was that relatively short (one day or less) formal CME events generally effected no change. The modal referent for a CPE course in clinical psychology is a one or two day workshop. In fact, the APA CACES criteria for the minimum length of a CE course is one hour, primarily to accommodate CPE programs set in the context of institutional in-service training. However, this negative finding must be interpreted in the context of the strict selection criteria and emphasis on performance enhancement in this specific study. As noted above, many workshops for clinical psychologists are focused on new knowledge and/or theory. Also, an important but limited role of such workshops has been posited in the three-phased learning model developed by Fox et al. (1989): "preparing to change," "making the change," and "solidifying the change." Using this conception, their research in CME found that traditional shorter focused workshops were more useful at the initial and final stage. For actually making performance change, that is, learning new skills, ongoing interaction with peers was essential. This principle is, in fact, incorporated in all specialty training in clinical psychology under the rubric of "supervised professional experience" and widely institutionalized, for example, in the long-standing model for training clinical psychologists in psychoanalysis.

The impressive research agenda for CME programs for physicians described above would need to be adapted to clinical psychology where it may be more difficult to determine such rigorous objective measures of "sound performance." Nevertheless, just as there has been increasing support for empirically validated clinical interventions, a research agenda in CPE for clinical psychology can be envisioned which approximates the standards which have been set by CME program providers and researchers.

2.17.8.3 Key Components of Effective CPE

There does seem to be a clear consensus about criteria which characterize good CPE programs across professions. The International Association for Continuing Education and Training (IACET) has established 10 criteria for the Continuing Education Unit (IACET, 1993). Although these criteria apply across the broad panoply of CE programs sponsored, for example, in large university "extension divisions," the criteria are, in fact, generally reflected in the APA CACES criteria. An eminent consultant on CPE has drawn on these criteria in emphasizing six critical components of effective CPE programs: needs assessment, learning outcomes, qualified CPE faculty, content and instructional methods, demonstration of learning, and program evaluation (Phillips, 1994b). Although these are familiar concepts, the systematic program development model presented relative to these six criteria is very sophisticated and would serve as a helpful guideline for CPE program providers in clinical psychology. For example, in discussing needs assessment, Phillips emphasizes the distinction between need and want. The former is based on clearly identifying the gaps in current knowledge and skills which negatively impact performance. All too often current needs assessment in clinical psychology is simply based on practitioners' "wants," that is, identifying topics in which they are interested. Sophisticated needs assessment is emphasized often as the most important criterion in enhancing effectiveness in CPE (Queeney, 1992; Vanek, Carey, Secic, Jackman, & Fleshler, 1994). In considering qualified CPE faculty, attention is often limited to the qualifications of the faculty with respect to content material but notice should also be taken of the instructor's ability to address the specific needs of adult professional learners (Lynton, 1992; Scott, 1994). Careful attention to these six criteria in controlled studies on CPE would serve as a good starting point for a research agenda on effective CPE programs in clinical psychology.

In an age of cost containment for clinical services one final point is warranted. Those in the educational establishment are understandably committed to the position that more education is better, but for the practitioner CPE can realistically be viewed as a cost center (Slotnick et al., 1995). CPE competes for resources with other aspects of the clinical psychologist's professional and personal life and includes both gains (peer contact, new knowledge and skills, opportunities to gain wisdom from respected practitioners) and costs (enrollment and travel fees, loss of income, unavailability for family contact). More attention must be given to the cost-effectiveness of CPE programs. If the cost of the program is too high for the participant practitioner, whether in financial terms or difficulty of access, there is little likelihood that a CPE program, however well designed, will reach many clinical psychologists. This latter point will be especially important as CPE programs move beyond the mere dissemination of new knowledge in the customary one or two day CPE format to longer term in-depth programs which focus on building skills for more effective clinical practice.

2.17.9 SUMMARY

In this chapter, major future directions in CPE for clinical psychologists have been suggested: continued rapid expansion of knowledge and skills in psychology will demand more CPE as the relative half-life of graduate education continues to diminish; differentiation in the field will continue and specialization and subspecialization will place greater demands on CPE programs that combine the acquisition of knowledge and skills for treating new populations and problems with innovative theories and techniques; the active practitioner will demand greater flexibility and access to CPE and increased utilization by CPE providers of new technology; changes in the health care system focused on greater accountability by clinical psychologists will demand CPE as a predominant method for quality assurance.

Most important, if a chapter on CPE is included in a book on clinical psychology 20 years hence, our field must possess a more coherent body of research on effective CPE programs. That research agenda can build on and complement the work of other health care professions. With thoughtful planning and commitment, it will be difficult but not impossible to disentangle the interaction effects of behavioral change, educational format, practice setting, and the clinical psychology interventions utilized.

In the larger context of changes in the health care system, the industrialization of clinical psychology is here to stay (Cummings, 1995). In a sense, all of the trends noted above are inherently associated with sweeping changes in the delivery of medical services (Walton, 1995). In that context, it will be all the more important to continue the long tradition in clinical psychology which emphasises the need for lifelong learning. There is, however, an important caveat in focusing on the need for CPE exclusively in the context of accountability, an inherently "outer directed" motive. Clinical psychologists also engage in CPE for a combination of psychological factors related to job satisfaction, including sustained interest and excitement, professional affiliation and peer interaction, validation of one's knowledge and skills, and a heightened sense of competence in dealing with troublesome client problems. These psychological factors are some of the major reasons that professionals give for participating in CPE and are the salient variables that many consider in choosing among CPE program options (Slotnick et al., 1995). Even in an industrialized health care system, idealism will be necessary for the health care professional and CPE programs for clinical psychologists will need to continue to appeal to such "inner directed" motives.

2.17.10 REFERENCES

American Hospital Association (1978). *Consideration of legislative mandatory continuing education proposals*. (AHA No. G037 13M–11/78–6474). Chicago: Author.

American Psychological Association (1995). *Education and training beyond the doctoral degree: Proceedings of the American Psychological Association National Conference on Postdoctoral Education and Training in Psychology*. Washington, DC; Author.

American Psychological Association (1992). Ethical principles of psychologists and code of conduct. *American Psychologist*. *47*(12), 1597–1611.

American Psychological Association (October, 1995). *Final report of the Board of Educational Affairs Task Force on Developing Model State and Provincial Continuing Education (CE) Regulations*. Washington, DC: Author.

American Psychological Association (July 1992). *Report of the ad hoc task force on psychopharmacology of the American Psychological Association*. Washington, DC: Author.

Association of Psychology Postdoctoral and Internship Centers (1993). *Proceedings of the National Conference on Postdoctoral Training in Psychology, October 27–31, 1992*. Washington, DC: Author.

Association of State and Provincial Psychology Boards (1996). *The handbook of licensure and certification requirements for psychologists in North America*. Montgomery, AL: Author.

Azzaretto, J. F. (1992). Quality control in continuing professional education: Accountability, effectiveness, and regulation. In E. S. Hunt (Ed.), *Professional workers as learners: The scope, problems, and accountability of continuing professional education in the 1990s* (Report No. OR–92–3006). Washington, DC: Office of Educational Research and Improvement. (ERIC Document Reproduction Service No. 350 945).

Beaudry, J. S. (1989). The effectiveness of continuing medical education: A quantitative synthesis. *The Journal of Continuing Education in the Health Professions, 9*(4), 285–307.

Board of Psychology, State of California Department of Consumer Affairs (1995). *Regulations relating to the practice of psychology, 1397.60–1397.69*. Sacramento, CA: Author.

Bourg, E. F., Bent, R. J., Callan, J. E., Jones, N. F., McHolland, J., & Striker, G. (1987). *Standards and evaluation in the education and training of professional psychologists*. Norman, OK: Transcript Press.

Briscoe, D. B. (1990). Community education: A culturally responsive approach to learning. In J. M. Ross-Gordon, L. G. Martin, & D. B. Briscoe (Eds.), *Serving culturally diverse populations. New directions for adult and continuing education. No. 48, Winter 1990: The Jossey-Bass higher and adult education series* (pp. 81–91). San Francisco: Jossey-Bass.

Broskowski, A. T. (1995). The evolution of health care: Implications for the training and careers of psychologists. *Professional Psychology: Research and Practice, 26*(2), 156–162.

Brown, R., Leichtman, S., Blass, T., & Fleisher, E. (1982). Mandated continuing education: Impact on Maryland psychologists. *Professional Psychology: Research and Practice, 13*, 404–411.

Cervero, R. M. (1988). *Effective continuing education for professionals*. San Francisco: Jossey-Bass.

Crespi, T. D., & Rigazio-Digilio, S. A. (1992). Professional development and continuing education: Implications for the school psychologist. *School Psychology International, 13,* 347–355.

Cummings, N. A. (1995). Impact of managed care on employment and training: A primer for survival. *Professional Psychology: Research and Practice, 26*(1), 10–15.

Davis, D. A., & Fox, R. D. (Eds.) (1994). *The physician as learner: Linking research to practice*. Chicago: American Medical Association Press.

Davis, D. A., Thomson, M. A., Oxman, A. D., & Haynes, R. B. (1992). Evidence for the effectiveness of CME: A review of 50 randomized controlled trials. *Journal of the American Medical Association, 268*(9), 1111–1117.

Davis, D. A., Thomson, M. A., Oxman, A. D., & Haynes, R. B. (1995). Changing physician performance: A systematic review of continuing medical education strategies. *Journal of the American Medical Association, 274*(9), 700–705.

Dillon, C. L. (1996). Distance education research and continuing professional education: Reframing questions for the emerging information infrastructure. *The Journal of Continuing Education in the Health Professions, 16*(1), 5–13.

Dubin, S. (1972). Obsolescence or lifelong education: A choice for the professional. *American Psychologist, 7,* 486–498.

Engel, C. E., Browne, E., Nyarango, P., Akor, S., Khwaja, A., Karim, A. A., & Towle, A. (1992). Problem-based learning in distance education: A first exploration in continuing medical education. *Medical Education, 26*(5), 389–401.

Fox, R. D. (1991). New research agendas for CME: Organizing principles for the study of self-directed curricula for change. *The Journal of Continuing Education in the Health Professions, 11*(3), 155–167.

Fox, R. D., Mazamian, P. E., & Putnam, W. E., (Eds.) (1989). *Changing and learning in the lives of physicians*. New York: Praeger.

Garganta, K. J. (1989). *The question of mandatory continuing education for professionals*. Cambridge, MA: Harvard University. (ERIC Document Reproduction Service No. 356 422).

Hampton, C. L., Mazmanian, P. E., & Smith, T. J. (1994). An interactive videoconference: An effective CME delivery system. *The Journal of Continuing Education in the Health Professions, 14*(2), 83–89.

Hargrove, D. S. (1992). Community mental health center staff development for service to special populations. In S. Cooper & T. Lentner (Eds.), *Innovations in community mental health* (pp. 141–169). Sarasota, FL: Professional Resource Books.

Haynes, R. B., Davis, D. A., McKibbon, A., & Tugwell, P. (1984). A critical appraisal of the efficacy of continuing medical education. *Journal of the American Medical Association, 251*(1), 61–64.

Hellkamp, D. T., Imm. P., & Moll, D. (1989). Mandatory continuing education (MCE): Desirable or undesirable? A survey of executive officers of state psychological associations. *Journal of Training & Practice in Professional Psychology, 3*(2), 33–46.

Houle, C. O. (1980). *Continuing learning in the professions*. San Francisco: Jossey-Bass.

Humphreys, K. (1996). Clinical psychologists as psychotherapists: History, future, and alternatives. *American Psychologist, 51*(3), 190–197.

Hunt, E. S. (1992). Introduction: The professional learner as worker. In E. S. Hunt (Ed.), *Professional workers as learners: The scope, problems and accountability of continuing professional education in the 1990s* (Report No. OR-92–3006). Washington, DC: Office of Educational Research and Improvement. (ERIC Document Reproduction Service No. 350 945).

International Association for Continuing Education and

Training (1993). *The continuing education criteria and guidelines*. Washington, DC: Author.

Jenett, P. A., & Swanson, R. W. (1994). Traditional and new approaches to CME: Perceptions of a variety of CME activities. *The Journal of Continuing Education in the Health Professions, 14*(2), 75–82.

Johnson, D. L. (Ed.) (1990). *Service needs of the seriously mentally ill: Training implications for psychology*. Washington, DC: American Psychological Association.

Korman, M. (1974). National conference on levels and patterns of professional training in psychology: The major themes. *American Psychologist, 34,* 696–702.

Lerman, H., & Porter, N. (1990). *Feminist issues in psychotherapy*. New York: Springer.

Lynton, E. A. (1992). Continuing professional education: Comparative approaches and lessons. In E. S. Hunt (Ed.), *Professional workers as learners: The scope, problems and accountability of continuing professional education in the 1990's* (Report No. OR-92–3006). Washington, DC: Office of Educational Research and Improvement (ERIC Document Reproduction Service No. 350 945).

Mann, K. V. (1990). Enhancing learning: How can learning theory help? *The Journal of Continuing Education in the Health Professions, 10*(4), 177–186.

Marsick, V. J. (1990). Human service organizations as communities of learning. In M. Galbreith (Ed.), *Education through community organizations* (pp. 45–54). San Francisco: Jossey-Bass.

Matarazzo, J. D. (1987). There is only one psychology, no specialties, but many applications. *American Psychologist, 42*(10), 893–903.

McGowen, K. R., & Hart, L. E. (1990). Still different after all these years: Gender differences in professional identity formation. *Professional Psychology: Research and Practice, 21*(2), 118–123.

McNamara, J. R. (1977). Patterns of continuing education for Ohio psychologists: A survey of interests and activities (1972–1974). *Professional Psychology, 8,* 368–376.

Moore, D. E., Green, J. S., Jay, S. J., Leist, J. C., & Maitland, F. M. (1994). Creating a new paradigm for CME: Seizing opportunities within the health care revolution. *The Journal of Continuing Education in the Health Professions, 14*(1), 4–31.

Myers, H. F., Wohlford, P., Guzman, L. P., & Echemendia, R. J. (Eds.) (1991). *Ethnic minority perspectives on clinical training and services in psychology*. Washington, DC: American Psychological Association.

Newcomer, K. E. (1989). Techniques for determining program quality. In J. S. Wholey & K. E. Newcomer (Eds.), *Improving government performance: Evaluating strategies for strengthening public agencies and programs. The Jossey-Bass public administration series and the Jossey-Bass management series* (pp. 126–140). San Francisco: Jossey-Bass.

Panayotoff, K. G. (1993). The impact of continuing education on the health of older adults. *Educational Gerontology, 19,* 9–20.

Peterson, D. R. (1976). Is psychology a profession? *American Psychologist, 31*(8), 572–581.

Peterson, D. R. (1991). Connection and disconnection of research and practice in the education of professional psychologists. *American Psychologist, 46*(4), 422–429.

Peterson, R. L., McHolland, J., Bent, R. J., Davis-Russell, E., Edwall, G. E., Maggidson, E., Polite, K., Singer, D. L., & Striker, G. (Eds.) (1992). *The core curriculum in professional psychology*. Washington, DC: American Psychological Association.

Phillips, L. (1994a). *The continuing education guide: The CEU and other professional development criteria*. Dubuque, IA: Kendall/Hunt.

Phillips, L. (1994b). *A study of mandatory continuing education*. Washington, DC: American Psychological Association.

Presland, J. (1993). Planning for continuing professional development. *Educational Psychology in Practice, 8*(4), 225–233.

Queeney, D. S. (1992). Problems of content and delivery in continuing professional education. In E. S. Hunt (Ed.), *Professional workers as learners: The scope, problems and accountability of continuing professional education in the 1990s* (Report No. OR-92–3006). Washington, DC: Office of Educational Research and Improvement (ERIC Document Reproduction Service No. 350 945).

Queeney, D. S., & English, J. K. (1994). *Mandatory continuing education: A status report. Information series, No. 57*. Columbus, OH: Center on Education and Training for Employment. (ERIC Document Reproduction Service No. 372 306).

Raimy, V. C. (Ed.) (1950). *Training in clinical psychology*. New York: Prentice-Hall.

Raymond, M. R. (1991, April). *The effectiveness of continuing education in the health professions: A reanalysis of the literature*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Robiner, W. N., Ohnsong, T. J., Parker, S. A., & Strike, B. (1994). Psychologists' HIV/AIDS experiences, anxiety, knowledge and continuing education: A preliminary report. *The Journal of Continuing Education in the Health Professions, 14*(4), 196–211.

Schön, D. A. (1983). *The reflective practitioner*. New York: Basic Books.

Schön, D. A. (1987). *Educating the reflective practitioner*. San Francisco: Jossey-Bass.

Scott, C. J. (1994). Applied adult learning theory: Broadening traditional CME programs with self-guided, computer-assisted learning. *The Journal of Continuing Education in the Health Professions, 14*(2), 91–99.

Shapiro, S. B., & Reiff, J. (1993). A framework for reflective inquiry on practice: Beyond intuition and experience. *Psychological Reports, 73,* 1379–1394.

Slotnick, H. B., Lichtauer, D. F., & Raskowski, R. R. (1995). Rethinking continuing medical education. *The Journal of Continuing Education in the Health Professions, 15*(1), 8–22.

Stewart, J. A., Horn, D. L., Becker, J. M., & Kline, J. S. (1993). Postdoctoral training in severe mental illness: A model for trainee development. *Professional Psychology: Research and Practice, 24*(3), 286–292.

Stricker, G., & Rodriguez, A. R. (1988). *Handbook of quality assurance in mental health*. New York: Plenum Press.

VandeCreek, L., Knapp, S., & Brace, K. (1990). Mandatory continuing education for licensed psychologists: Its rationale and current implementation. *Professional Psychology: Research and Practice, 21*(2), 135–140.

Vanek, E. P., Carey, W. D., Secic, M., Jackman, D. M., & Fleshler, B. (1994). Value of needs assessment surveys. *The Journal of Continuing Education in the Health Professions, 14*(4), 224–231.

Walton, H. J. (1995). World summit on medical education and continuing education. *The Journal of Continuing Education in the Health Professions, 15*(1), 40–47.

Wiens, A. N., & Baum, C. G. (1995). Characteristics of current postdoctoral programs. In *Education and training beyond the doctoral degree: Proceedings of the American Psychological Association National Conference on Postdoctoral Education and Training in Psychology*. Washington, DC; American Psychological Association.

Woolfort, M. W., Lang, W. P., Farghaly, M. M., Ziemiecki, T. L., & Faja, B. W. (1991). Varying the format of CDE: Practitioners' perceptions of need and usefulness. *The Journal of Continuing Education in the Health Professions, 11,* 215–224.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.18

Generalist and Specialist Training in Clinical Psychology in Australia

PAUL R. MARTIN
University of New England, Armidale, NSW, Australia

2.18.1 INTRODUCTION 209
2.18.2 HISTORY OF EDUCATION AND TRAINING FOR CLINICAL PSYCHOLOGY IN AUSTRALIA 210
2.18.3 THE AUSTRALIAN PSYCHOLOGICAL SOCIETY AND ACCREDITATION OF CLINICAL TRAINING PROGRAMS 211
2.18.4 STATE AND TERRITORY REGISTRATION ACTS FOR PSYCHOLOGISTS AND SPECIALIST PSYCHOLOGISTS 212
2.18.5 MASTERS DEGREES: GENERALIST AND SPECIALIST TRAINING 213
2.18.6 CLINICAL PH.D.s: GENERALIST AND SPECIALIST TRAINING 213
2.18.7 PROFESSIONAL DOCTORATES: GENERALIST AND SPECIALIST TRAINING 215
2.18.8 THE CONTEXT FOR AUSTRALIAN TRENDS IN CLINICAL TRAINING 216
2.18.9 DISTINCTIVE FEATURES OF CLINICAL TRAINING IN AUSTRALIA 217
2.18.10 FUTURE DIRECTIONS IN CLINICAL TRAINING 218
2.18.11 SUMMARY 219
2.18.12 REFERENCES 220

2.18.1 INTRODUCTION

Postgraduate training in clinical psychology in Australia has gone through a dramatic growth period since the 1980s during which time the number of base level courses has doubled and professional doctorates have been introduced at several universities. The rapid change promises to continue into the early twenty-first century so that Australian universities will have a range of training programs that bear little resemblance to the situation in the 1980s.

In preparing this chapter, a questionnaire on clinical training was sent to the heads of all 38 university departments/schools of psychology in Australia. Ultimately, all heads responded, so that this chapter is based on a comprehensive review of current offerings. Sections 2.18.5–2.18.7 report the questionnaire responses.

A word of caution needs to be emphasized with respect to the term "specialist." When used in the context of the discipline of psychology, clinical is viewed as one specialty along with other professional specialties such as counselling, organizational, and educational. On the other hand, clinical psychology covers a broad domain and specialties within clinical psychology are often recognized such as mental health,

health psychology, and child and adolescent psychology. Although it is tempting to call these specialties subspecialties, the problem is that it is debatable with many fields whether they should be identified as a specialty or subspecialty. Some clinical psychologists see health psychology, for example, as a subspecialty of clinical psychology, while some health psychologists see clinical psychology as a subspecialty of health psychology. In this chapter, the term subspecialty will not be used. The meaning of the term specialty should be clear from the context in which it is used and the examples given.

The chapter will begin by developing a context for a discussion of offerings in clinical training. Hence, a brief historical review of training will be presented followed by discussion of the roles of the Australian Psychological Society (APS) and the State and Territory Psychology Registration Boards. This will be followed by an analysis of the various forms of clinical training available in Australia including a consideration of distinctive features. The chapter will conclude with a section on future directions for clinical training.

2.18.2 HISTORY OF EDUCATION AND TRAINING FOR CLINICAL PSYCHOLOGY IN AUSTRALIA

The first universities in Australia were established in the 1850s. The earliest approximation to courses in psychology were taught in philosophy departments and date back to the latter part of the nineteenth century. The first department of psychology in Australia was founded in 1927 at the University of Sydney. This department was followed by departments at the University of Western Australia (UWA) (1930) and the University of New England (1939), or New England University College as it was known at the time.

With respect to the practice of clinical psychology, one of the pioneers was Stanley Porteous, well known for the Maze Test he developed, who established a clinic for subnormal and maladjusted children in Melbourne in 1913 (Taft, 1982). During the 1920s, State governments appointed psychologists in New South Wales, South Australia, Tasmania, and Western Australia (Birnbrauer, 1996). As no Australian universities offered training in clinical psychology during this period, the early practitioners had diverse backgrounds. Elton Mayo in Queensland, for example, was trained as a philosopher, and Ethel Stoneman in Western Australia completed most of her training in the US and Europe.

The growth of professional services in Australia outpaced the provision of appropriate academic training. For many years, three-year graduates were employed as psychologists and the greater part of learning and skills acquisition took place on the job in study groups (Ivison, 1977). In 1949, UWA offered the first systematic training in clinical psychology by way of a one-year full-time fourth year course. This course included nondirective (client-centered) counseling and play therapy, and teaching methods included case work under supervision utilizing one-way screens, group discussion, and role playing. In 1956, this course expanded into a two year full-time postgraduate Diploma in Clinical Psychology (DCP). The University of Sydney established a similar qualification three years later.

The modern era of clinical training in Australia began in 1966 with the introduction of the first masters program. UWA again led the way converting its DCP into the Master of Psychology (Clinical), followed by the University of New South Wales introducing a Master of Psychology (M.Psych.) in 1969. The number of M.Psych. (Clinical) programs increased steadily during the 1970s and 1980s to around 12, and more than doubled during the 1990s.

Although the number of M.Psych. programs has continued to grow, many academics responsible for training in clinical psychology have considered for some time that the masters programs need upgrading to extended programs that will lead to doctoral degrees (Martin, 1989a). The Australian Ph.D. is modeled on the British rather than the North American Ph.D. in that it is a research degree assessed by thesis. It is possible for departments to require students to complete coursework as part of their Ph.D. studies, but typically this is not a significant component of the degree. Through the 1980s Australian universities were opposed to introducing coursework Ph.D.s or professional doctorates as such developments were seen as devaluing the traditional research Ph.D.s. Hence, the only upgraded programs in clinical psychology that were introduced during this period were clinical Ph.D.s or combined M.Psych./Ph.D.s. UWA, for example, established a combined M.Psych./Ph.D. in 1983 whereby students completed all the requirements of both degrees simultaneously except the M.Psych. thesis requirement was replaced by a requirement to complete additional coursework and field placements. Since this time there has been a steady growth in the number of such programs.

The clinical Ph.D.s and combined M.Psych./Ph.D.s essentially upgraded training in clinical psychology by elevating the requirements for the research component of training. Some academics believed that the real need was for increased education in the methods and knowledge base of psychology, and for expanded on-the-job skills training under supervision. The major changes in Australian higher education that began in the late 1980s (National Board of Employment, Education, and Training [NBEET], 1989) created a new environment that provided opportunities for developing innovative programs that met these needs. In 1991, Murdoch University introduced a doctor of psychology (D.Psych.) program that was essentially an extension of their M.Psych. program. The following year UWA established a D.Psych. modeled on Australian medical education which requires six years at university (bachelor of medicine, bachelor of surgery), followed by two years of in-service training, at which point medical practitioners can apply for training programs in the various medical specialties such as neurology, psychiatry, and oncology. Hence, the UWA D.Psych. required applicants to have completed six years at university (bachelors and masters degrees) and two years post-masters supervised experience. The degree was offered in three specialities in clinical psychology: mental health, health psychology, and child and adolescent psychology (Martin, 1993). Only a limited number of D.Psych. programs were set up until 1997 at which time the number increased dramatically.

2.18.3 THE AUSTRALIAN PSYCHOLOGICAL SOCIETY AND ACCREDITATION OF CLINICAL TRAINING PROGRAMS

In 1945, 28 psychologists voted to become an Overseas Branch of the British Psychological Society (O'Neil, 1987). The APS was established as an independent entity in 1965. The *Australian Journal of Psychology* was founded by the Society in 1949 and was joined by the *Australian Psychologist* in 1966. The first edition of the APS Code of Conduct was published in 1968. The APS now has 33 branches and regional groups, 13 interest groups, and over 11 600 members. It has nine Colleges: Clinical Psychologists, Clinical Neuropsychologists, Community Psychologists, Counseling Psychologists, Educational and Developmental Psychologists, Forensic Psychologists, Organisational Psychologists, Sport Psychologists, and Health Psychologists (formed in 1996).

Eligibility criteria for membership of the APS involves the applicant holding an accredited four-year degree in psychology and having undergone a further two years of postgraduate training, either through an accredited university program (i.e., masters or doctorate) or under the supervision of a member of APS. The APS has changed its membership requirements so that from the year 2000, applicants will not be considered eligible under the two years of postgraduate supervised experience rule but will have to complete an accredited university postgraduate program. Eligibility for membership of the Colleges varies. For the College of Clinical Psychologists, applicants are required to have completed an accredited masters degree in clinical psychology followed by two years supervised experience.

The APS first became involved in accreditation of courses in 1972. The accreditation process has changed markedly, however, with the introduction of a new set of guidelines in August 1995 effective from the beginning of 1996 (APS, 1995). These guidelines set out criteria for departments/schools of psychology on the one hand and programs in psychology on the other. The departmental guidelines pertain mainly to staffing and resources. The program guidelines are divided into four sections; three year programs, fourth year programs, professional fifth and sixth year programs, and research higher degree programs. In addition to these general guidelines, each of the Colleges specify guidelines pertaining to the particular College's area of specialty. Clinical psychology programs, for example, have to satisfy the general guidelines for professional fifth and sixth year programs and the guidelines of the College of Clinical Psychologists.

The APS accreditation guidelines emphasize that training should be based on the scientist–practitioner model and specify the following core areas of study:

(i) the systematic study of the mechanisms and etiology of psychological disorders in adults and children;
(ii) the principles and methods of behavioral, psychometric, and clinical assessment of significant psychological problems;
(iii) the principals, procedures, and techniques of psychological intervention and rehabilitation;
(iv) psychopharmacology;
(v) research methods and evaluation;
(vi) health psychology, behavioral medicine, and rehabilitation; and
(vii) ethics and professional standards.

Additional areas of study listed as options include:

(i) neuropsychology;
(ii) forensic psychology;
(iii) sexual dysfunctions;
(iv) interventions with children and families;

(v) working with special groups such as the elderly;
(vi) advanced research and assessment methodology;
(vii) organizational psychology (health administration);
(viii) educational psychology; and
(ix) advanced study of core areas.

The APS accreditation guidelines require students to spend at least 1000 hours completing practicums including a minimum of 400 hours of practical training face to face with clients. Supervision must be provided by clinical psychologists with at least three years of post-masters experience and total contact with supervisors should be at least 200 hours. Students are required to complete three to five practicums external to the university and the practicums must provide students with the experience of dealing with a wide range of clinical problems (acute and chronic), across varying age ranges (child, adult), settings (inpatient/outpatient, community), and using a variety of clinical skills (assessment, treatment, and professional). Practicums typically involve attending the agency for two to three days per week over a period of two to four months.

The APS accreditation guidelines specify that students are required to complete a project which involves collection of original data. The topic of the research must be of direct relevance to the field of clinical psychology, and supervision should be provided by one or more individuals of whom at least one must be from the university department. Research reports are required to be presented in 10 000–15 000 words; and are usually assessed by two examiners who may be members of the university department or external examiners.

All departments of psychology are reviewed in a five-year cycle. The process involves submitting an application followed by a site visit. The review panel conducting the site visit includes a member or nominee of the local State or Territory Psychology Registration Board and representatives from each of the Colleges for which the department offers specialist courses. Site visits last for a minimum of two days and include a meeting with the Vice-Chancellor or nominee, academic and general staff, and students. Also, teaching and research facilities are inspected and examples of students' work, such as theses, are studied. It should by emphasized that the new accreditation guidelines are quite detailed and specific. Also, site visits only developed in the 1980s and 1990s. Accreditation can serve an important role in improving clinical training (Nelson, 1993), and the new APS guidelines and procedures should help achieve this outcome.

2.18.4 STATE AND TERRITORY REGISTRATION ACTS FOR PSYCHOLOGISTS AND SPECIALIST PSYCHOLOGISTS

In 1965, Victoria was the first state to enact a Psychological Practices Act. The other five states and two territories followed with the last act passed in the Australian Capital Territory in 1995. These acts require all practicing psychologists to be registered, and they regulate psychology at a state or territory level via a board which consists usually of psychologists representing different sectors of the profession (e.g., university, health, private practice), a lawyer, a psychiatrist, and a registrar. Most recently enacted legislation includes a representative of consumer interests and this is likely to be a feature of future amendments to acts. The boards report to the Ministers of Health in each state or territory.

Most state and territory registration boards only recognize the title of psychologist. Victoria and Western Australia, however, offer statutory recognition of a range of specialist titles (Wood, 1996). Victoria recognizes clinical, clinical neuropsychological, counseling, educational and developmental, organizational, forensic, sport, and community. Western Australia recognizes the same specialties with the exception of community. Other states and territories are considering recognizing specialist titles.

Eligibility criteria for registration usually involves the applicant holding an accredited four-year degree in psychology and having undergone a further two years of postgraduate training, either through an accredited university program or under the supervision of an experienced registered psychologist. Where specialist titles exist, eligibility for the specialist title is based on whether applicants are eligible for membership of relevant colleges of the APS.

The decision by the APS to require six years of university training for membership by the year 2000 has not been followed by a parallel shift in eligibility criteria by the state and territory registration boards, except for the Northern Territory Psychology Registration Board which will also change to a six-year educational requirement from the year 2000. A survey of the members of the registration boards in all states and territories, however, revealed that an overall majority of members supported increasing the educational requirements for registration to six years (Healy & Franklin, in press). This survey also indicated that a majority of board members favored the introduction of specialist registration. There is a strong move to bring the registration acts of the different states and territories into line with each other but this may take time to achieve.

2.18.5 MASTERS DEGREES: GENERALIST AND SPECIALIST TRAINING

In 1997 there were 25 masters programs in Australia offering training in clinical psychology (see Table 1). Of these programs, 19 were APS accredited with the remaining six too new to qualify for accreditation. Four new masters programs are due to commence in 1998. The nomenclature for most programs is Master of Psychology (Clinical) or Master of Clinical Psychology but some variations exist including Master of Applied Psychology (Clinical), MA (Clinical Psychology), and M.Sc. (Clinical Psychology). Some programs have added health to clinical.

The APS accreditation guidelines state that entry to these programs should be by way of an honors degree in psychology or equivalent, and this seems to be the norm although it was not so in the past (Martin, 1989b). One program has specified the additional requirement of at least one year's professional experience.

The APS accreditation guidelines indicate that masters programs should be constituted as follows: coursework, 40–50%; research, 20–30%; practical placement, 25–30%. In practice, the coursework component varies from 25–59% with a mean of 45%; the research component varies from 19–50% with a mean of 26%; and the practical placement component varies from 17–50% with a mean of 30%. Hence, the means are in line with the APS guidelines but some courses are outside the guidelines. The variability between programs can be seen by comparing three universities who reported the following ratios of coursework–research–practical placements: university A, 56–19–25; university B, 25–50–25; and university C, 25–25–50.

The majority of masters programs are considered to be generalist rather than specialist in orientation. The only exceptions were four programs that indicated a specialty in health psychology and two programs that indicated a specialty in mental health. It is not clear, however, whether these programs had more content in these areas than other programs that did not claim a specialization. When coordinators of masters programs were asked whether there were features of their programs that differentiated them from other programs, some of the responses suggested emphases in particular areas of psychology such as forensic psychology, multicultural psychology, geropsychology, rural psychology, educational/child psychology, and organizational psychology. Most programs adopt a predominantly cognitive-behavioral approach to assessment and treatment, but two programs suggested that they were different by virtue of covering additional theoretical orientations such as psychodynamic, constructionist, and family systems approach.

An example of a masters program is included in Table 2. The table illustrates that the three components of coursework, practicum, and research are represented in each of the four semesters of the program but the balance changes over time. That is, the early part of the program is coursework dominated principally to prepare students for their first practicum but also to help students get started on their research. As the program progresses, the balance changes with coursework steadily decreasing and more time being devoted to practicum and research commitments.

2.18.6 CLINICAL PH.D.s: GENERALIST AND SPECIALIST TRAINING

It is difficult to summarize the situation with respect to clinical Ph.D.s and combined M.Psych./Ph.D.s because of the diversity of arrangements that have been negotiated in different universities. These programs essentially involve students completing most if not all of the requirements for both degrees, more or less simultaneously. In 1997, 14 departments indicated that such programs were available in their university. The number of such programs is debatable, however, as it is obviously possible to complete both degrees at any university that offers both degrees, so the question is only whether special arrangements have been made for completing the two degrees concurrently.

One of the main variations between programs is whether individuals who complete the program successfully take out both degrees or only the higher degree (i.e., Ph.D.). About half the programs award both degrees with the other half acknowledging the completion of the M.Psych. requirements via adding clinical in brackets after Ph.D. or listing the M.Psych. units on the Ph.D. transcript. Programs that award a Ph.D. only need to seek accreditation from the APS for the clinical component of the program. A number of programs have received such accreditation.

Entry requirements for these programs are variable. Some programs simply require an honors degree or equivalent as for M.Psych. programs. Others have elevated this requirement to holding a first class honors and one program requires an M.Psych. or equivalent.

Table 1 Australian universities offering postgraduate training in clinical psychology.

	Master of Psychology			*Clinical Ph.D. or combined M.Psych/Ph.D.*			*Doctor of Psychology*		
University	*Operating 1997*	*APS accredited*	*Planned 1998*	*Operating 1997*	*APS accredited*	*Planned 1998*	*Operating 1997*	*APS accredited*	*Planned 1998*
Adelaide	√	√							
Australian National	√	√		√	√				
Australian Catholic			√						
Bond	√						√		
Charles Sturt	√								
Curtin	√	√		√	√				
Deakin	√	√					√		
Edith Cowan	√			√					
Flinders	√	√		√	√				
Griffith (Nathan)	√	√		√	√				
Griffith (Gold Coast)	√								
James Cook	√	√							
La Trobe	√	√		√	√		√		
Macquarie	√	√		√					
Melbourne	√	√					√		
Monash									√
Murdoch	√	√		√	√		√	√	
New England			√						√
New South Wales	√	√							
Newcastle	√	√							
Queensland	√	√		√	√				
Royal Melbourne IT	√			√					
South Australia	√						√		
Southern Queensland			√						
Sydney	√	√		√	√				
Tasmania	√	√		√	√				
Victoria	√	√					√		
Western Australia	√	√		√	√				
Western Sydney			√						
Wollongong	√	√		√			√	√	
Total	25	19	4	14	10	0	8	2	2

Another variation is to accept applications only from individuals completing M.Psych. programs. The usual arrangement here is for students to transfer into the clinical Ph.D. or combined M.Psych./Ph.D. following successful completion of the first year of the M.Psych. Some programs allow entry via a number of different routes (e.g., direct entry from honors or transfer from M.Psych.).

Traditionally, universities insist on completion of the full Ph.D. requirements for awarding a Ph.D. That is, they will not accept trade-offs along the lines of reducing the scope of research requirements against clinical coursework and placement requirements. More flexibility is shown sometimes with M.Psych. requirements whereby, for example, most universities seem prepared to waive the M.Psych. research requirements if students are completing a Ph.D. concurrently on a clinical topic. Some universities insist, however, on other requirements being added to compensate for dropping the research requirement (e.g., extra coursework and placements). Given that students have to complete the full requirements of a Ph.D., the requirements for clinical Ph.D.s and

Table 2 Outline of the Master of Psychology (Clinical) Program at The University of Western Australia.

Semester	*Year*
One	First
Coursework	Research, evaluation, and accountability I Psychotherapy I (general introduction to psychotherapy) Clinical problems I (includes psychopathology) Child and adult assessment Neuropsychology Preparation for practicum (includes interviewing and practicum orientation) Workshops (four half-day and one full day)
Clinical practicum	Placement in university clinic I (begins half way through semester)
Research	Develop a research topic and find a supervisor
Two	
Coursework	Research, evaluation, and accountability II Psychotherapy II (cognitive behavioural therapy) Clinical problems II (includes health psychology) Workshops (two half day and ten full day)
Clinical practicum	Placement in university clinic II External practicum I
Research	Submit a research proposal
Three	Second
Coursework	Case presentations I Intellectual disability Interventions with children and adolescents Family therapy Workshops (three half day and one full day)
Clinical practicum	Placement in university clinic III (ends half way through semester) External practicum II
Research	Carry out research project
Four	
Coursework	Case presentations II Workshops (one half day and five full day)
Clinical practicum	External practicum III
Research	Submit a thesis in form of journal manuscript

Source: Martin (1996a).

combined M.Psych./Ph.D.s are typically research dominated (i.e, at least two-thirds research). One exception is a Ph.D. (Psychotherapy) program that has been established with 50% coursework, 25% research, and 25% practical placements.

Most of these programs do not aim to provide specialist training. Some programs do provide specialist training in areas such as health psychology by means of the M.Psych. component and, of course, the research component focuses on a specific topic and hence may provide some degree of specialist training. The specialist training available in the program therefore depends on the expertise of members of the academic and adjunctive staff as the potential supervisors. The Ph.D. (Psychotherapy) can be conceptualized as a specialist degree, and one program provides training towards a career in academia by means of including training in areas such as teaching, manuscript reviewing, and writing research grant applications.

2.18.7 PROFESSIONAL DOCTORATES: GENERALIST AND SPECIALIST TRAINING

In 1997 there were eight D.Psych. programs in Australia offering training in clinical psychology, six operating for the first time in 1997. A ninth university had a D.Psych. program running but removed it from the statute books. Another two universities are planning to introduce D.Psych. degrees and a third is exploring possibilities. Only two D.Psych. programs are accredited by the APS but they are all likely to be accredited as they incorporate all the requirements of M.Psych. programs.

The degree is referred to generally as a Doctor of Psychology and abbreviated to D.Psych. but one program has the title Doctor of Clinical Psychology and one program uses the abbreviation Psy.D. Direct entry is possible to four programs with an upper second or first class honours degree. A fifth program requires good performance in the first year of the M.Psych. and a sixth program requires good performance during both years of the M.Psych. program. One program offered entry by means of any of the three routes described above. The period for completion of a D.Psych. is generally three years of full-time study but one program requires four years of full-time study.

The APS accreditation guidelines indicate that professional doctorates should be constituted as follows: coursework, 20–30%; research, 40–70%, practical placement, 20–30%. In practice, the coursework component varies from 25 to 39% with a mean of 32%; the research component varies from 33 to 50% with a mean of 43%, and the practical placement component varies from 11 to 33% with a mean of 24%.

Four of the programs did not claim to provide specialist training. One program indicated a specialty in child, adolescent, and family psychology. Another program indicated that students had to choose two of three specialist areas in their third year: health, child, and adult.

2.18.8 THE CONTEXT FOR AUSTRALIAN TRENDS IN CLINICAL TRAINING

Masters degrees in clinical psychology, the entry-level qualification for clinical practice in Australia, have doubled in number since the late 1980s with five more planned for 1998. Placing this trend in an international context, masters level training in clinical psychology in the US has been an issue for more than 30 years as many have argued that this level of training is inadequate for producing fully-fledged scientist–practitioners. Nevertheless, the number of masters programs has increased over time rather than decreased (Edelstein & Brasted, 1991). In contrast, the number of masters programs in the UK has decreased during the 1990s as master of psychology programs have been replaced by doctor of psychology programs (Martin, 1996b).

The increase in masters programs in Australia is probably driven by two main factors. First, the number of universities in Australia doubled between the late 1980s and the late 1990s as a result of legislation converting former teacher training colleges and institutes of technology into universities, and the advent of private universities. Second, the APS changed its membership rules to include the requirement of a higher degree, as noted earlier. This move put pressure on universities to offer professional masters programs as a means of helping their students to gain APS membership. In fact, in an early draft of the current APS accreditation guidelines, the APS stated that only departments offering such programs would be accredited. Although this requirement was removed in later drafts, all departments have been left with the clear impression that they are expected to offer at least one professional masters program.

If the increase in masters programs is large, the increase in doctor of psychology programs is even larger with two programs operating in 1996 against eight programs running in 1997 and two more planned for 1998. Of course, doctor of psychology programs in the US have a history dating back approximately 30 years but the reason for the emergence of the Psy.D. in the US was different from the reason for the emergence of the D.Psych. in Australia. The North American Psy.D. had its origins in clinical psychologists concluding that the Ph.D. was too research dominated to meet the needs of aspirant practitioners (Barlow, Hayes, & Nelson, 1984). The advent of the Australian D.Psych. was based on a desire to upgrade the existing masters programs. Hence, while some North American Psy.D.s have eliminated research training from the curriculum, research is a major component of the Australian D.Psych. representing on average 43% of the degree.

The first doctor of clinical psychology degree was introduced in the UK at about the same time as the D.Psych. in Australia and is usually abbreviated to DClinPsy. Both degrees are extended masters programs but the British DClinPsy is somewhat different in structure to the Australian D.Psych. as the average percentage of the programs devoted to coursework, research, and practical placements is 27%, 26%, and 47%, respectively (Martin, 1996b). Hence, compared to the Australian degree it is more focused on practicums and less on research. The number of DClinPsy programs in the UK has proliferated very rapidly from the first program in 1992 to 17 in 1995 with at least another six in a planning or developmental phase.

One point that stands out from Table 1 is the large number of clinical training programs in Australia. Twenty-five universities are offering clinical training and this number is expected to increase to 30 by 1998. By comparison, the UK has 26 clinical training programs despite the fact that the population of the UK is approximately three times that of Australia. In 1997, 19 universities offered doctoral level training in Australia and by 1998 this number will have risen to 21. In 1989 there were 158 doctoral programs in the US (Edelstein & Brasted, 1991), so that there are approximately twice as many doctoral programs in Australia per capita of population as in the US. The development of doctoral programs in Australia is a particularly important step forward given the data indicating the superior competence of graduates from doctoral programs compared to masters programs (Robiner, Arbisi, & Edwall, 1994).

The impetus for the development of the D.Psych. was the changes in education in Australia discussed earlier (NBEET, 1989). It is relevant to point out that these changes gave rise to a range of professional doctorates being introduced in Australia in addition to the professional doctorates in psychology. In 1996, for example, whilst there were only two D.Psych. programs operating, there were 19 doctor of education degrees (Ed.D.) and 28 other professional doctorates (Shanahan, 1996). Legal/juridical science ($n = 8$) and business administration ($n = 5$) were the second and third most common types of professional doctorate.

The number of Australian universities offering clinical Ph.D.s or combined M.Psych./Ph.D. programs has grown steadily since the early 1980s to 14 programs in the late 1990s. Hence, this option is more common than the D.Psych. and is certainly the preferred option by approximately half the universities offering clinical training. Australian combined programs really have no equivalent in the US as they incorporate the full requirements of a British/Australian, research-only Ph.D. in contrast to North American Ph.D.s which have lesser research requirements due to including coursework and practicum components. Isolated examples of combined programs have existed in the UK for many years but have never been a major provider of clinical training.

2.18.9 DISTINCTIVE FEATURES OF CLINICAL TRAINING IN AUSTRALIA

Comparing clinical training in Australia with clinical training in the US and the UK, a number of impressions arise. First, the Australian system is less diverse in some ways and more diverse in other ways. On the one hand, clinical training in Australia is offered only by university departments of psychology. In contrast, clinical training in the UK has traditionally been offered by university departments of psychology, university departments of psychiatry, jointly by university departments of psychology and psychiatry, and regional or district health authorities in combination with the British Psychological Society. Qualifications in clinical psychology are now awarded through the Open University.

In the US, some clinical training is carried out at professional schools which may or may not be attached to universities. On the other hand, the range of options in clinical training in Australia, that is, M.Psych., D.Psych., clinical Ph.D., and combined M.Psych./Ph.D. is almost certainly broader than the range of options available in the UK and possibly broader than the range of options available in the US.

The most distinctive feature of clinical training in Australia is the consistent commitment to the scientist–practitioner model, and the emphasis on research training which comprises on average 26% of the M.Psych., 43% of the D.Psych., and at least 67% of clinical Ph.D.s. These weightings of the research component of clinical training are greater than are found in the US and the UK. It is easy to see why clinical training in Australia has developed in this way. It reflects, for example, the point made above that clinical training in Australia is based exclusively in university departments of psychology which traditionally place a very high value on research. It also reflects the fact that the profession of clinical psychology, which is likely to favor more emphasis on practicums rather than research, has had relatively little impact on training in Australia compared to other countries. This has occurred for various reasons such as the weak accreditation processes that have operated in Australia in the past (a situation which is, hopefully, changing); and the fact that the profession in Australia is weak and divided compared to some other countries. The latter occurs for many reasons but is related to the small number of clinical psychologists (perhaps 1000–2000 in total) spread across a vast continent, employed by several separate agencies in each of six states and two territories.

What effect has the greater emphasis on research training had? On the positive side, graduates of clinical training programs in Australia seem well satisfied with their research training. A survey by Byrne (1982), for example, reported that research was one of the areas of training that attracted higher adequacy ratings than some other aspects of clinical training. A survey of the profession indicated that 73% of respondents endorsed the length of research training as appropriate and only 25% were dissatisfied with the quality of research training (Martin, 1989b).

The scientist–practitioner model has often been criticized on the grounds that it is an ideal rather than a reality because few clinical psychologists engage in research. A much quoted statement, for example, is the report by Barlow (1980) that the modal number of publications of clinical psychologists in the US was zero despite the majority holding Ph.D.s. It would be nice to believe that the greater emphasis on research in Australian clinical training would result in a higher proportion of clinical psychologists engaging in research. Unfortunately, the survey discussed above suggested that only a minority of clinical psychologists in Australia carried out research even when less stringent criteria than publications were used to index research activity (Martin, 1989b). On the other hand, D.Psych.s and even clinical Ph.D.s have been introduced too recently to gauge their impact on research output, so perhaps a survey of the future will produce more positive findings in terms of research contributions from clinical psychologists in Australia.

2.18.10 FUTURE DIRECTIONS IN CLINICAL TRAINING

In commenting on recent trends in clinical training in the US, Edelstein and Brasted (1991) argued that "the most notable change in the field of clinical psychology in the past two decades has been the shift toward specialization" (p. 51). They went on to state that further specialization in the future was inevitable. The survey results reported in Sections 2.18.5–2.18.7 show little evidence of specialization in Australia even at the doctoral level. A few programs indicated a specialization in health psychology, and there were isolated examples of other specializations such as mental health, child and adolescent, forensic, and geropsychology. Perhaps the lack of specialist training is not surprising as the APS accreditation guidelines allow little freedom for developing specialized programs.

It does seem likely that specialist training in Australia will increase in the future. Much specialist training in the US is accomplished at the postdoctoral level, after generalist training has been completed. As doctoral programs replace masters programs in Australia, postdoctoral training is the next logical step. Professional development activities have never been mandatory in Australia but are now required for continuing membership of the APS Colleges (APS, 1997), and achievement of this important milestone will undoubtedly stimulate developments in postqualification training, many of which will be of a specialist nature.

A significant and unfortunate trend that is emerging in clinical training in Australia is the increased costs for the trainees (Martin, 1996a). Unlike the UK, relatively few clinical students in Australia receive financial support in the form of scholarships or traineeships. In 1988, the federal government introduced the Higher Education Contribution Scheme whereby graduates are required to pay a proportion of the costs of their courses either by means of an upfront (discounted) payment or after graduation by means of a tax levied on their income. During the first few years of the scheme, clinical students were exempted from the tax through government scholarships but after a cutback in the number of available scholarships in 1994, clinical students found themselves liable for the tax. In 1993, the federal government decided to allow universities to charge fees for postgraduate courses. Clinical programs were reluctant to capitalize on this opportunity, but decreased government funding has forced some clinical programs to charge fees and probably most programs will charge fees in the very near future.

It is difficult to predict what effect the introduction of fees will have on clinical training. A decreased demand is possible at least from the financially disadvantaged. Fees may encourage more students to take a break between completing their undergraduate and postgraduate degrees so that they can generate the necessary funds for the latter. More students may opt for part-time study to enable concurrent part-time employment. Students may select programs on more financially-orientated criteria such as which programs will facilitate them earning income whilst studying, and which programs will enable them to gain higher incomes on graduation so that they can repay loans more quickly.

Many changes are taking place in the work environment of clinical psychologists that will impact on clinical training. Traditionally, the majority of clinical psychologists in Australia have worked in the public sector but the size of the public sector is shrinking as the federal and state governments cut spending by privatizing or amalgamating public hospitals and contracting out services. In contrast, private practice is a growth area in psychology and other human service professions. These changes will create challenges for clinical training as practicum supervisors have been drawn typically from the public rather than the private sector, and likewise research has been limited to the public sector. Clinical programs will need to reorientate their training to cater for the changed employment opportunities for graduates.

Other significant developments in the work environment include the move towards "multiskilling" which is based on the assumption that there are no important differences between psychologists and other health care professionals such as social workers and occupational therapists (Franklin, 1993). Multiskilling has been introduced by federal and state governments, and implications include advertising positions as open to all healthcare professions rather than tied to a particular profession. Budgeting principals have been changed so that funds are allocated for service provision rather than to professionally based departments resulting in professions having to argue for the relative merits of their services. Changes such as these will force clinical training programs to produce graduates who can compete effectively with other health care professionals. Graduates will need to have a clear sense of professional identity and of the unique contribution that clinical psychology can make to the provision of human services. The cautious, critical approach that characterizes the scientific side of the scientist–practitioner model will have to be counterbalanced by the confidence of an established profession with much to offer. Something akin to marketing and promotional skills should become part of every clinical training program.

Finally, it would be inappropriate to conclude a section on future directions without considering the role of technology, given the rapid process of evolution in this field. Technological developments are likely to impact on clinical training in a number of ways. First, computers are used increasingly for psychological assessment and intervention (Bloom, 1992) so that training programs will need to educate students in the use of the new technologies. Second, computers have a number of features that can be used to advantage in training including presenting instructional materials, monitoring students' understanding of the materials, providing students with feedback on their performance, and allowing students to determine the pace and direction of their learning (Kenardy & Adams, 1993). Two approaches to computer-based clinical training have been reported in the literature (Bloom, 1992). One is computer-based clinical consultation in which expert or knowledge based systems are used to assist in assessment and treatment planning. These systems can provide a teaching function by describing the reasoning process that they invoked in reaching decisions. The other approach is simulation of the encounter between the healthcare provider and the client in which the computer is programmed to take the role of the client.

The third contribution of new technology, and perhaps the most exciting, is the role it can play in enhancing distance education and breaking down barriers between universities, not just within national boundaries but also around the world. If it is accepted that there is substantial overlap between clinical training programs even in different countries then similar material is taught in almost 200 doctoral programs each year just in the US, the UK, and Australia, let alone the rest of the world. This is clearly an inefficient system and with most universities experiencing pressure on their resources, alternative models need to be considered. Developments in telecommunications and fiber optics, and the expansion of the World Wide Web and use of email, create opportunities that were not available in the past. Experts in instructional design increasingly are developing distance education materials that facilitate interaction between students and teachers, and the development of skills, rather than simply transmitting information. Hopefully, the future will see the development of high quality clinical teaching materials and units that will be used internationally to the advantage of both teachers and students.

2.18.11 SUMMARY

Clinical psychology training in Australia is entering an exciting phase of its development. The number of programs has increased dramatically and new types of degrees have burst on to the scene. Australia now has a larger number of programs and diverse offerings that are characterized internationally by a strong emphasis on research training. The APS has developed its accreditation principles and processes to the point where they could have an important influence on clinical training, and the advent of mandatory professional development will impact on postqualification clinical training. State and territory legislation related to registration of psychologists needs revision to keep pace with these developments but there are encouraging signs that this process has begun.

The above developments will lead to many challenges. How will an adequate number of clinical supervisors be found given the contraction of the public sector on the one hand and the increased number of trainees needing supervision on the other? Who will employ the expanded group of graduates from these programs? Will the increased financial pressures on students encourage training programs to make changes that will attract students but have detrimental effects such as minimizing training in areas associated with great social need but low financial rewards for clinical psychologists?

The 1990s have been characterized by marked change for clinical training in Australia. The next decade will be even more interesting.

ACKNOWLEDGMENTS

I would like to express my appreciation to all the individuals who completed the questionnaires on which this chapter is partly based.

2.18.12 REFERENCES

Australian Psychological Society (1995). *Accreditation guidelines*. Melbourne, Australia: Author.

Australian Psychological Society (1997). *Professional development handbook*. Melbourne, Australia: Author.

Barlow, D. H. (1980). Behavior therapy: The next decade. *Behavior Therapy, 11,* 315–328.

Barlow, D. H., Hayes, S. C., & Nelson, R. O. (1984). *The scientist practitioner: Research and accountability in clinical and educational settings*. New York: Pergamon.

Birnbrauer, J. S. (1996). The development of clinical psychology. In P. R. Martin & J. S. Birnbrauer (Eds.), *Clinical psychology: Profession and practice in Australia* (pp. 21–51). Melbourne, Australia: Macmillan.

Bloom, B. L. (1992) Computer-assisted psychological intervention: A review and commentary. *Clinical Psychology Review, 12,* 169–197.

Byrne, D. G. (1982). Clinical psychology in Australia: A survey of activities and aspirations. *Australian Psychologist, 17,* 87–96.

Edelstein, B. A., & Brasted, W. S. (1991). Clinical training. In M. Hersen, A. E. Kazdin, & A. S. Bellack (Eds.), *The clinical psychology handbook* (2nd ed., pp. 45–65). New York: Pergamon.

Franklin, J. (1993) Crisis? *The Bulletin of the Australian Psychological Society, 15,* 6–8.

Healy, M., & Franklin, J. (in press). Psychologist registration boards: Views on four and six year education. *Australian Psychologist.*

Ivison, D. J. (1977). Clinical psychology in public service. In M. Nixon & R. Taft (Eds.), *Psychology in Australia: Achievements and prospects* (pp. 175–183). Rushcutters Bay, Australia: Pergamon.

Kenardy, J., & Adams, C. (1993). Computers in cognitive-behaviour therapy. *Australian Psychologist, 28,* 189–194.

Martin, P. R. (1989a). “Specialist” clinical psychologists: Upgrading training in clinical psychology. *Australian Psychologist, 24,* 3–11.

Martin, P. R. (1989b). The scientist–practitioner model and clinical psychology: Time for change? *Australian Psychologist, 24,* 71–92.

Martin, P. R. (1993). The Doctor of Psychology (Clinical) degree at UWA and related developments overseas. *The Bulletin of the Australian Psychological Society, 15,* 12–13.

Martin, P. R. (1996a). Training in clinical psychology. In P. R. Martin & J. S. Bimbrauer (Eds.), *Clinical psychology: Profession and practise in Australia* (pp. 52–76). Melbourne, Australia: Macmillan.

Martin, P. R. (1996b). A survey of doctoral degrees in clinical psychology in the United Kingdom and Australia. *Australian Psychologist, 31,* 191–193.

National Board of Employment, Education, and Training (1989). *Australian graduate studies and higher degrees. Initial report by the Higher Education Council.* Canberra, Australia: Australian Government Publishing Service.

Nelson, P. D. (1993). Improving the clinical training of psychologists: A role for accreditation? In P. Wohlford, H. F. Meyers, & T. Callan (Eds.), *Serving the seriously mentally ill: Public academic linkages in services* (pp. 207–211). Washington, DC: American Psychological Association.

O’Neil, W. (1987). *A century of psychology in Australia.* Sydney, Australia: Sydney University Press.

Robiner, W. N., Arbisi, P., & Edwall, G. E. (1994). The basis of the doctoral degree for psychology licensure. *Clinical Psychology Review, 14,* 227–254.

Shanahan, P. J. (1996). Professional doctorates other than the doctor of education in Australian Universities: Some comparative data. In T. W. Maxwell & P. J. Shanahan (Eds.), *Which way professional doctorates? Context and cases* (pp. 13–28). Armidale, Australia: Faculty of Education, Health, and Professional Studies, University of New England.

Taft, R. (1982). Psychology and its history in Australia. *Australian Psychologist, 17,* 31–39.

Wood, M. M. (1996). Professional affiliations, registration and ethical issues. In P. R. Martin & J. S. Birnbrauer (Eds.), *Clinical psychology: Profession and practice in Australia* (pp. 77–100). Melbourne, Australia: Macmillan.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.19

Commission for the Recognition of Specialties and Proficiencies in Professional Psychology

TOMMY T. STIGALL
The Psychology Group, Baton Rouge, LA, USA

2.19.1 INTRODUCTION 221

2.19.2 NEW SPECIALTIES AND PROFICIENCIES 222

  2.19.2.1 Clinical Neuropsychology 223
  2.19.2.2 Clinical Health Psychology 223
  2.19.2.3 Biofeedback: Applied Psychophysiology 223

2.19.3 OWNERSHIP OF SPECIALTIES AND PROFICIENCIES 224

2.19.4 GENERAL PRACTICE SPECIALTIES AND SUBSPECIALTIES 224

2.19.5 RELATIONSHIP OF THE COMMISSION TO OTHER CREDENTIALING BODIES 225

  2.19.5.1 Marketplace Influences 225
  2.19.5.2 State Licensing Boards 225
  2.19.5.3 American Board of Professional Psychology 226
  2.19.5.4 National Register of Health Service Providers in Psychology 226
  2.19.5.5 APA College of Professional Psychology 226
  2.19.5.6 APA Committee on Accreditation 227
  2.19.5.7 Council of Credentialing Organizations in Professional Psychology 227

2.19.6 TOWARD A TAXONOMY OF SPECIALTIES AND PROFICIENCIES 228

2.19.7 REFERENCES 229

2.19.1 INTRODUCTION

The Commission for the Recognition of Specialties and Proficiencies in Professional Psychology (CRSPPP) was established by the American Psychological Association (APA) Council of Representatives in February 1995. Implicit in the Council's action was a recognition of the need to bring some order into the conceptualization of specialties and to communicate more effectively with the general public about the nature and scope of psychological practice. In creating this new entity, the APA declared itself open for business in the matter of formally recognizing both practice specialties and proficiencies.

> There shall be a Commission for the Recognition of Specialties and Proficiencies in Profession Psychology. The Commission (a) shall review petitions from petitioning organizations requesting the Association's recognition of a professional specialty or proficiency, and (b) shall establish mechanisms for the periodic evaluation and renewal of such recognition. The Commission shall select from among recognition petitions it reviews those deemed appropriate to recommend to the Council of Representatives as meriting formal

> confirmation. Council shall confirm the recognition of a specialty or proficiency in professional psychology only when such recommendation for recognition shall have been placed before it by the Commission following the completion of an evaluation conducted according to the Commission's procedures. (American Psychological Association, 1996a, Association Rule 90–8)

Thus, the Council delegated responsibility to the Commission for receiving, reviewing and recommending meritorious petitions, while reserving unto itself the final authority to approve any newly proposed specialty or proficiency. At the same time, the Council approved specific criteria and principles for the recognition of specialties and proficiencies. Four historical specialties, clinical, counseling, school, and industrial/organizational psychology, were recognized *de facto*, subject to eventual review and reconfirmation by the normal petitioning process.

By virtue of the fact that the APA College of Professional Psychology previously had been granted authority to examine applicants and certify qualified psychologists in the Treatment of Alcohol and Other Psychoactive Substance Use Disorders, this became the first proficiency of professional psychology to be so recognized. Thereafter, newly proposed proficiencies would have to conform to the CRSPPP review process and be recognized by the Council before they could be considered by the College for the awarding of certificates.

According to the procedures that the Commission would follow, each new or existing specialty and proficiency would be required to submit to review and reconfirmation not less than every seven years. Failure to do so could subject the specialty or proficiency to withdrawal of recognition by APA.

The Commission, as initially constituted, was composed of 13 members, all of whom except for the public member had been appointed to serve on the predecessor Joint Interim Committee for the Identification and Recognition of Specialties and Proficiencies which had been charged with the responsibility to develop recommendations regarding the specialty and proficiency recognition process. (The first commissioners were Cynthia D. Belar, Bruce E. Bennett, Thomas J. Boll [co-chair], Patricia M. Bricklin, Arthur L. Kovacs, Nadine M. Lambert, William P. MacLeod [public member], Joseph D. Matarazzo, Jack K. Plummer, Edward P. Sheridan, Tommy T. Stigall, Arthur N. Wiens, and Jack G. Wiggins, Jr. [co-chair].) Beginning in February 1997, the CRSPPP would be composed of nine commissioners, including one public member who is not to be a psychologist. The remaining eight commissioners must be licensed psychologists and members of the APA. The public member is appointed by the APA Board of Directors. Others are elected by the Council to serve three-year, staggered terms from slates prepared by the following APA governance groups: Board of Educational Affairs, Board of Professional Affairs, Board for the Advancement of Psychology in the Public Interest, and the Committee for the Advancement of Professional Practice.

In addition to pursuing its work of reviewing petitions in light of the definitional criteria and operating procedures under which the Commission operates, the CRSPPP must also contend with important conceptual and philosophical issues. Among these are questions about the nature and purpose of professional specialization and an evolving quest for a conceptual model—a taxonomy for specialties and proficiencies.

2.19.2 NEW SPECIALTIES AND PROFICIENCIES

Specialization can be thought of either as the division of human labor for increased efficiency, or in terms of the organization of knowledge. Both concepts are relevant to an understanding of specialties and proficiencies, as these terms are employed by the CRSPPP, and both are embedded in the following definition, which is incorporated in the APA policy document *Principles for the Recognition of Specialties in Professional Psychology* (American Psychological Association, 1995a):

> A specialty is a defined area of psychological practice which requires advanced knowledge and skills acquired through an organized sequence of education and training. The advanced knowledge and skill specific to a specialty are obtained subsequent to the acquisition of core scientific and professional foundations in psychology. (p. 2)

In addition to the concepts of an organized body of knowledge and the application of this knowledge to practice, the definition emphasizes that the knowledge and skill for specialty practice must be at an advanced level and that it must be acquired in an organized sequence of education and training. A specialty rests on and grows out of the core knowledge and skills of the discipline. It does not supplant or supersede it, but it is distinctive and different from the generic core.

The principles specify 12 criteria that apply in the evaluation of petitions seeking APA recognition of new specialties. Among other things, the criteria stipulate that there must be an identifiable public need for the specialty and that the new specialty must be distinctive, though not necessarily unique, with regard to the three parameters of specialty practice: populations served, problems addressed, and procedures and techniques utilized. For a positive recommendation to be made by the CRSPPP, the petition and supporting documents must be complete and must provide evidence to satisfy all the criteria.

2.19.2.1 Clinical Neuropsychology

In August 1996, the Council of Representatives voted for the first time to recognize a new specialty on the basis of a petition submitted by the APA Division of Clinical Neuropsychology (Division 40). Clinical neuropsychology thus became the fifth specialty in professional psychology officially recognized by the APA (DeLeon, 1997, p. 856). Central to the recognition process is the requirement for a clear and concise definition of the practice of any approved specialty or proficiency. The definition is maintained as a part of the archival record and serves as the basis for communication with the public about the specialized professional services offered by psychologists who are associated with the specialty or proficiency in question. For this purpose, the following definition of clinical neuropsychology was adopted:

> Clinical Neuropsychology is a specialty that applies principles of assessment and intervention based upon the scientific study of human behavior as it relates to normal and abnormal functioning of the central nervous system. The specialty is dedicated to enhancing the understanding of brain–behavior relationships and the application of such knowledge to human problems. (American Psychological Association, 1996, p. 155)

2.19.2.2 Clinical Health Psychology

One year later, in August 1997, the specialty of clinical health psychology was officially recognized by the Council, based on a petition review and favorable recommendation by the CRSPPP. The petition had been prepared by the APA Division of Health Psychology (Division 38). In approving clinical health psychology as the sixth officially recognized specialty in professional psychology, the Council accepted the following definition of this specialty:

> The specialty of Clinical Health Psychology applies scientific knowledge of the interrelationships among behavioral, emotional, cognitive, social and biological components in health and disease to the promotion and maintenance of health; the prevention, treatment and rehabilitation of illness and disability; and the improvement of the health care system. The distinct focus of Clinical Health Psychology is on physical health problems. The specialty is dedicated to the development of high quality services to individuals, families, and health care systems. (American Psychological Association, 1997, p. 158)

In the deliberations of the Council, a controversy arose about the use of the modifier “clinical” in the title of emerging specialties and proficiencies. Proponents argued that the use of “clinical” is necessary to differentiate training programs preparing psychologists for practice from training programs preparing clinical scientists. An opposing point of view held that psychologists who had not neen trained specifically in a clinical psychology program (e.g., counseling psychologists or school psychologists) would be disadvantaged by such terminology. While approving “Clinical Health Psychology” as the title of the new specialty, the Council imposed a moratorium on the recognition of future specialties and proficiencies incorporating “clinical” in the title until “the confusing and problematic meanings surrounding the generic use of the term “clinical” have been addressed and resolved by the APA to the satisfaction of Council” (Council of Representatives, 1997). A special task force was to be appointed by the APA President to address the complex issues involved and report its recommendations to the Council. The task force was to coordinate its activities with the CRSPPP.

2.19.2.3 Biofeedback: Applied Psychophysiology

Biofeedback: Applied Psychophysiology was recognized as a proficiency in professional psychology also at the August 1997 meeting of the Council. But, unlike the petitions for clinical neuropsychology and health psychology, both of which had come from APA divisions, the petition for this new proficiency was submitted from an interdisciplinary organization, The Biofeedback Certification Institute of America. Since 1981, the Institute has certified individual practitioners from various disciplines as proficient in biofeedback. The CRSPPP assumes that proficiencies are not necessarily subsumed by any particular specialty of psychology and they may, in fact, be shared with other disciplines.

> A proficiency is a circumscribed activity in the general practice of professional psychology or one or more of its specialties. The relationship between

> a body of knowledge and a set of skills related to the parameters of practice ... represents the most critical aspect of the definition of a proficiency. (American Psychological Association, 1995b)

Unlike specialties, which must demonstrate their saliency with respect to all three of the essential parameters of practice, proficiencies are required to show that they are distinctive in at least one of the parameters. Emphasis on the procedural and technical aspects of the proficiency can be seen in the official definition of biofeedback: applied psychophysiology:

> Biofeedback refers to a group of therapeutic procedures that use electronic instruments to record and display to the patient information about the ongoing activity of various body processes of which the person is usually unaware. The goal of biofeedback therapy is to help the patient achieve voluntary control over body processes that are normally involuntary or that have become involuntary through accident or disease. Biofeedback is used in the treatment of many medical and psychological disorders. (American Psychological Association, 1997, p. 160)

It is expected that the petition review process will lead to other recommendations for recognition of new specialties and proficiencies. Strong interest has been shown on the part of clinical, counseling, and school psychology in an early submission of petitions for continued recognition on the basis of a formal review by the CRSPPP. Not all petitions reviewed are recommended for recognition. Those that are incomplete or fail to satisfy all the required criteria are returned to the petitioning organization with an explanation of the identified deficiencies. It is left to the petitioning organization whether to revise and resubmit the petition at a later date.

2.19.3 OWNERSHIP OF SPECIALTIES AND PROFICIENCIES

One of the conceptual issues with which the CRSPPP has struggled is that of ownership of specialties and proficiencies. Initial thinking had been that there must be a single administrative organization that would take responsibility for the specialty or proficiency and would continue to oversee the development of this specialized aspect of professional practice. On the advice of legal counsel, and after further deliberation, the CRSPPP has taken a different position on this important question.

The Commission now takes the position that specialties and proficiencies are not owned by any particular membership group or organization, although one or more professional organizations may indeed be closely associated with a particular specialty or proficiency. In spite of the fact that some organization must take responsibility to petition for official APA recognition, it does not follow that APA recognition confers any special status on the petitioning body. For purposes of the APA recognition process, ownership of a specialty or proficiency is assigned to the profession of psychology at large.

It is assumed that there will be at least one organization which will continue to encourage the development of a specialty or proficiency in order to assure its viability. In the event an officially recognized specialty or proficiency should become obsolete or no longer serve a public need, the option for sunsetting comes into play.

2.19.4 GENERAL PRACTICE SPECIALTIES AND SUBSPECIALTIES

The notion of general practice specialties is well established in healthcare and is closely related to the concept of primary care provider. In medicine, internists, family practice physicians, obstetricians, gynecologists, and pediatricians, are often identified in this way. Practice audits in psychology (Greenberg, Smith, & Muenzen, 1996; Rosenfeld, Shimberg, & Thornton, 1983) suggest that the specialties of clinical and counseling psychology, and perhaps school psychology as well, may deserve to be thought of as general practice specialties.

A petition prepared by the APA Division of Clinical Psychology (Division 12) for continued recognition of clinical psychology as a specialty in professional psychology describes the specialty as "the largest general practice specialty providing psychological services in the US and around the world" (Division of Clinical Psychology, 1995, p. 1). The proposed definition of clinical psychology contained in the petition relies on very broad language of the sort that typically is found in the definition of the practice of psychology in state licensing laws. The petition declares, "As a general practice specialty, Clinical Psychology is essentially the equivalent of 'health service provider' in psychology" (p. 2). The petition further asserts that "other health service provider specialties, by and large, have their origins in Clinical Psychology" (p. 2).

The Commission has been receptive to the concept of general practice specialties in professional psychology, including the notion that there could be several such general practice specialties. Thus far, however, the CRSPPP has not addressed the more complex matter of subspecialties or the possible hierarchical relationship among existing and emerging specialties. The development of a formal taxonomy for specialties and proficiencies may eventually prove helpful in establishing a consensus about these matters within the professional community.

2.19.5 RELATIONSHIP OF THE COMMISSION TO OTHER CREDENTIALING BODIES

It is important to understand that the Commission does not credential individual psychologists in practice specialties or proficiencies, those functions being reserved for other appropriate bodies such as state licensing boards, the APA College of Professional Psychology, and the American Board of Professional Psychology (ABPP). The CRSPPP does expect that its decisions and the official recognition of specialties and proficiencies by the APA will inform the activities of credentialing authorities, accreditation bodies, and training institutions, as well as employers and consumers of psychological services.

2.19.5.1 Marketplace Influences

Identification of specialists by appropriate credentialing bodies is a service to the public. The unfettered proliferation of self-declared competencies and the abandonment of definitional criteria to market forces alone is not an acceptable solution to the need for reliable and useful information that will help to guide those persons in need of a psychologist's services.

If one refers to the yellow pages of any metropolitan telephone directory, it is possible to see listed any number of practice areas in association with the advertising of individual psychologists. In the pages of professional journals and trade publications, one can find announcements for all manner of entrepreneurial specialty credentials or certifications; just send in your application and fee. Psychologists seeking to apply for managed care panels are frequently confronted with checklists that they are expected to complete in order to document their areas of specialization. One large managed care company has announced its intention to credential empaneled behavioral health care providers, including psychologists, according to its own list of 13 "specialties" for which specified criteria have been developed. (Merit Behavioral Care Corporation, 1996). The 13 listed specialties are abuse (physical and sexual), addictions, AIDS/HIV, biofeedback, child and adolescent therapy, geriatrics, group therapy, eating disorders, marriage and family therapy, neuropsychological assessment, psychological testing, psychopharmacology, and sexual dysfunction.

It is little wonder that surveys of consumer opinions reveal widespread confusion and lack of knowledge regarding psychology and psychological services (Farberman, 1997). In a free-market economy and free society, it is probably neither possible nor desirable to restrain the free exchange of information about professional and commercial services. But a more uniform system of terminology and definition of specialties and proficiencies in professional psychology should help to improve public understanding about psychological services and increase public confidence in the specialized services that psychologists have to offer.

2.19.5.2 State Licensing Boards

A license to practice psychology issued by the state licensing board is an important, but insufficient, credential to help the consumer identify which kind of psychologist is best qualified to deal with a particular problem. The licensing laws in most jurisdictions are generic, and boards of examiners have relied heavily on ethical guidelines that call upon psychologists to practice in accordance with their documented professional education, training and experience (American Psychological Association, 1992). Beyond this custom, however, a number of states also attempt to regulate specialty practice directly.

A review of state licensing laws (Stigall & Hall, 1998) reveals that the most frequent practice involves separate licensing or certification for school psychologists, typically by the state education agency. The next most frequent occurrence is certification of psychologists as health service providers on the basis of additional qualifications beyond those required for initial licensing. In attempting to identify for the public those psychologists qualified to render health services, licensing boards are following a definitional concept and standard established by the Council for the National Register of Health Service Providers in Psychology (1996).

A few states have sought to license clinical psychologists. In still other states, boards of examiners have taken it upon themselves to develop standards that can be used on a local basis to identify psychologists with special qualifications or areas of practice. Typically, this involves documentation of specified training and supervised experience relevant to a particular specialty or proficiency. The psychologist's special status is registered with the board or, in some instances, a special certificate is issued.

While intended to protect the public, efforts by state licensing boards have been hampered by the absence of clear professional standards regarding specialties and proficiencies in psychology. The result has been largely a patchwork approach, based upon variable standards and interpretations by local authorities. More uniform professional standards applied to specialties and proficiencies should prove useful to officials responsible for administering the licensing and regulatory provisions of state laws. A principal function of the Association of State and Provincial Psychology Boards (ASPPB) has been to promote greater administrative uniformity and consistency across jurisdictions.

2.19.5.3 American Board of Professional Psychology

The organization best known for credentialing specialists in psychology is the American Board of Professional Psychology (ABPP). The ABPP and its predecessor organization, the American Board of Examiners in Professional Psychology, has been examining applicants seeking the ABPP Diploma since 1947. Over the years the number of specialties recognized by the ABPP has grown, so that by 1997 there were 10 separate diplomate categories: behavioral psychology, clinical neuropsychology, clinical psychology, counseling psychology, family psychology, forensic psychology, health psychology, industrial/organizational psychology, psychoanalysis in psychology, and school psychology (ABPP, 1997).

The ABPP diplomate status is intended to signify advanced competence in an area of specialty practice. Eligibility for the ABPP requires licensure as a psychologist and successful performance on the examination conducted by one of the ABPP specialty boards. Additional requirements may be imposed for acceptance by a given specialty.

In the absence of any official external review process for identifying and recognizing specialties, the ABPP has been in the unenviable position of having to make decisions regarding both the recognition of specialties and the credentialing of individual psychologists. Specialty-regulating activities of professional—as opposed to governmental—organizations carry with them an element of antitrust risk, which appears to be directly related to the consolidation of power in a single organization to influence the market for specialty services. For an analysis of the antitrust implications of specialty credentialing and regulation, see Krauss, Ratner, and Sales (1997). With the advent of the CRSPPP, a decision has been reached to concentrate on the ABPP's core function of credentialing specialists, while relying on the APA to exercise the prior function of formally recognizing specialties in professional psychology (J. G. Matarazzo, personal communication, April 25, 1997).

2.19.5.4 National Register of Health Service Providers in Psychology

A voluntary listing of over 15 000 psychologists who have been found to meet the definition of a health service provider in psychology is provided by the *National Register of Health Service Providers in Psychology*. To be listed in the publication, an individual must be licensed for the independent practice of psychology, hold a doctoral degree in psychology, and have completed two years of supervised experience in health services in psychology. One year of experience must be in an organized health service training program or internship, and one year must be postdoctoral. Specific definitional criteria apply to the doctoral program, the internship and the postdoctoral supervised experience. A health service provider in psychology is considered to be "trained and experienced in the delivery of direct, preventive, assessment and therapeutic intervention services to individuals whose growth, adjustment, or functioning is impaired or to individuals who otherwise seek services" (Council for the National Register of Health Service Providers in Pschology, 1996).

As a service to subscribers, the National Register lists up to five areas of expertise for each registrant, along with other identifying information, theoretical orientation, ages or groups targeted, and languages spoken. Registrants may select from a listing of 51 areas of expertise, ranging from "adjustment disorders and relationship problems" to "terminally ill." In order to claim an area of expertise, registrants must document the particular doctoral and postdoctoral education, training, and experience relevant to the claimed expertise. A careful inspection of the listing suggests that most would conform more closely to the concept of proficiency than specialty, as these terms are used by the CRSPPP.

2.19.5.5 APA College of Professional Psychology

With the ascendance of managed care over the past decade, large numbers of psychologists have been confronted with the need to document their qualifications in specific areas of practice in order to meet the requirements imposed by managed care companies and serve on provider panels. This problem was especially acute for those practitioners desiring to provide services to the alcohol and substance abuse population and for whom a generic license to practice psychology was deemed insufficient by the managed care industry. The College of Professional Psychology was therefore established by the APA in 1994 to certify those psychologists who demonstrate that they have mastered the specific knowledge and skill in one or more recognized practice proficiencies.

Since the CRSPPP and the College of Professional Psychology are both integral parts of the APA, formal policies of the organization govern the relationship between the two bodies. Association Rule 130.5 specifies that the College "shall have the authority to issue certificates of proficiency ... from among *recognized* [italics added] practice proficiencies ... subject to ratification by the Council" (American Psychological Association, 1996a, p. 44). The effect of this policy is to restrict the issuance of certificates by the College to that pool of proficiencies already recognized by the Council, based upon prior review and recommendation by the CRSPPP.

Considering the number of petitions for proficiencies that have been submitted to CRSPPP, it is very likely that additional proficiencies will be recognized in the future. At the time of writing, however, only two proficiencies, treatment of alcohol and other psychoactive substance use disorders and biofeedback: applied psychophysiology, had been officially recognized by the APA, and only the former had been developed as a certificate offering by the College.

2.19.5.6 APA Committee on Accreditation

The APA Committee on Accreditation (CoA) is recognized by the US Department of Education as an official accrediting body for professional training programs in psychology. For more than half a century, the APA has implicitly recognized professional specialties through its accreditation of doctoral training programs in clinical, counseling, and school psychology. Accreditation is a voluntary process whereby academic and internship training programs may achieve the APA's imprimatur of approval, based upon a process of self-study and external review. In carrying out this important function, the CoA enjoys unusual autonomy and self-determination in keeping with the requirements of the US Department of Education and APA policy.

More recently, the APA Council has taken steps to expand the scope of accreditation at both the doctoral and postdoctoral levels of training (DeLeon, 1997, p. 857). These changes in accreditation policy and procedures, which have major implications for the development and recognition of specialties in professional psychology, had been anticipated in a report issued some years earlier by the Joint Council on Professional Education in Psychology (JCPEP). Among the resolutions contained in the JCPEP report were the following:

> The JCPEP supports a change in APA policy that would make application for APA accreditation available to all doctoral programs in psychology that publicly declare their intent to educate and train psychologists for practice.
>
> The JCPEP urges the APA to approve accreditation criteria for postdoctoral residency training that, in addition to providing advanced specialty training, could satisfy the statutory requirements in most jurisdictions for the supervised experience prerequisite to licensure. (Stigall et al., 1990, p. 25)

As new and emerging specialties become recognized by the APA, it is likely that formal postdoctoral education and training will become an integral component of preparation for specialty practice. The CoA will play a key role in developing standards for accreditation of postdoctoral residency training, as well as new doctoral programs in professional psychology. The CRSPPP acknowledges that specialty preparation may incorporate elements of both doctoral and postdoctoral education and training. The criteria for approval of any petition for recognition of a specialty state that "a specialty requires advanced, specialty-specific scientific knowledge," as well as "advanced didactic and experiential preparation that provide the basis for services with respect to the essential parameters of practice" (American Psychological Association, 1995a, p. 4).

2.19.5.7 Council of Credentialing Organizations in Professional Psychology

Individuals contemplating a career in the practice of psychology are confronted with an often bewildering array of educational and credentialing requirements: doctoral education, internship training, postdoctoral supervised experience and licensure as a psychologist, perhaps additional postdoctoral education and training in an area of specialization, specialty examination and certification, and continuing professional education, which may be for the purpose of renewal of license or recertification in one or more practice proficiencies. Each of these domains of professional education, training, or credentialing is overseen by one or more separate administrative bodies. In some instances, these bodies are statutorily based, as in the case of state licensing boards; or they may be part of a professional membership organization, as is true of CRSPPP; or they may be separately incorporated entities such as the National Register and the ABPP.

The need for an interorganizational forum to deliberate and promote more consistent standards and procedures in accreditation and credentialing was recognized by the Joint Interim Committee which had developed the blueprint for the CRSPPP and recommended its creation by the APA. The Council of Credentialing Organizations in Professional Psychology (CCOPP) was established in response to this need. Since 1995, several meetings of representatives from the participating organizations have been held, and plans have been formulated for the future activities of CCOPP.

The mission of the CCOPP is to "provide interorganizational leadership in the integration of policies and procedures for quality and process improvement in designation, accreditation, and certification of training programs, individuals, specialties, and proficiencies in professional psychology" (Council of Credentialing Organizations in Professional Psychology, 1996, p. 1). Participation in the CCOPP is by invitation and in 1996 included representatives from the ABPP, the ASPPB, the National Register, the Canadian Register of Health Service Providers in Psychology, the Canadian Psychological Association, the Association of Psychology Postdoctoral and Internship Centers (APPIC), and from the APA's CRSPPP, the College of Professional Psychology, and the CoA.

An overall goal articulated by the CCOPP is to establish "a seamless credentialing process for professional psychology, one that affords a coherent relationship between individual credentials obtained at different stages and for different purposes in one's professional career" (Council of Credentialing Organizations in Professional Psychology, 1996, p. 2). Other issues with which the CCOPP has been involved are: the development of a database related to credentialing in professional psychology; encouraging the formation of a council of specialties, which would serve in an advisory capacity to the CoA regarding postdoctoral accreditation; and endorsement of a certification proposal advanced by the ASPPB, which is intended to facilitate the interstate mobility of psychologists. The CCOPP has expressed a strong interest in having an opportunity to comment on petitions for the recognition of specialties and proficiencies that come before the CRSPPP.

2.19.6 TOWARD A TAXONOMY OF SPECIALTIES AND PROFICIENCIES

It should be apparent that the specialty and proficiency recognition process followed by the Commission is empirically based. Petitions are initiated by organizations external to the CRSPPP and are evaluated in accordance with established criteria and procedures. Those that are found to meet the criteria are recommended for formal recognition by the APA. At the same time, an element of judgment is always a part of the process, and there are implicit assumptions and biases that cannot altogether be eliminated.

If a formal taxonomy for describing and ordering specialties and proficiencies in professional psychology were available, it could help to make the recognition process more explicit and public. In an effort to move forward with the development of such a taxonomy, the CRSPPP has articulated a number of basic assumptions:

> (i) The purpose of identification and recognition of specialties in professional psychology is to promote communication with the general public about psychological services. Our primary audience is not other psychologists.
> (ii) Information about specialties in professional psychology must be communicated in language that is easily understood by the general public.
> (iii) Specialties are recognized as they are developed by the profession, based upon the advancement of psychological knowledge and identified public need.
> (iv) Specialties in professional psychology have the following characteristics: (a) A shared core of scientific and professional knowledge, skill, and attitudes; (b) A distinctive body of knowledge and professional education and training; (c) An education and training sequence consisting of doctoral study, internship, and residency; (d) Assessment, consultation, and treatment or other interventions with individuals, groups, or organizations, tailored for each specialty; (e) Practice proficiencies, which may be subsumed by a given specialty or which may be common to one or more specialties. (Commission for the Recognition of Specialties and Proficiencies in Professional Psychology, 1996)

Based upon these assumptions, the CRSPPP has devised the following rationale and schema for differentiating between general practice specialties and other specialties and proficiencies in professional psychology:

The public will continue to need the services of general practice specialties, such as those of clinical, counseling, school, and industrial/organizational psychology. The general practice specialties include both the common core of knowledge as well as the advanced core of knowledge particular to each specialty. The emergence of new specialties to provide needed psychological services must also be recognized and validated.

1. *General Practice Specialties*
CRSPPP assumes that the majority of professional psychologists are in the general practice of psychology (clinical, counseling, school, and industrial/organizational psychology).

2. *Other Specialties*
The defining criteria for specialties and proficiencies are found in the CRSPPP policy documents. These include problems addressed, populations served, and procedures or technologies employed. All three criteria are required for specialty recognition, and one or more are required for proficiency recognition.

Specialties of professional psychology may evolve from education and training common to one or more general practice specialties, extending that practice with more specialized psychological education, training, and practice competencies, while integrating these elements with relevant scientific knowledge from related disciplines. By comparison, subspecialties can evolve from more intense or focused training and practice within a recognized specialty.

The following groupings are illustrative and not definitive or exhaustive. Other combinations can be conceptualized.

a. Psychology and Neurosciences
- Clinical Neuropsychology
- Pediatric Neuropsychology
- Geriatric Psychology

b. Psychology and Health Sciences
- Clinical Health Psychology
- Pediatric Psychology

c. Psychology and Forensic Sciences
- Correctional Psychology
- Police Psychology

Note: Bulleted titles could emerge as recognized proficiencies, subspecialties, or specialties, depending on their meeting the criteria for specialty or proficiency recognition. (Commission for the Recognition of Specialties and Proficiencies in Professional Psychology, 1966)

A viable taxonomy will require further revision and elaboration of these preliminary elements. Since the evolution of specialties and proficiencies is a dynamic process, any taxonomy must also remain open to change. It is likely that future models will continue to reflect the interplay of empirical and conceptual components. But such robust concepts as general practice specialties, subspecialties, and the grouping of related categories are likely to remain useful in the development of an acceptable taxonomy.

As the knowledge base in the discipline of psychology and related fields continues to expand, it is incumbent on the profession itself to identify for the public both the officially recognized areas of specialized psychological practice and those practitioners qualified to represent themselves as specialists. The specialty and proficiency recognition process established by the APA and administered by the CRSPPP is one indication of the maturing of a responsible profession.

2.19.7 REFERENCES

American Board of Professional Psychology (1997). *Directory of Diplomates*. Columbia, MO: Author.

American Psychological Association (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47,* 1597–1611.

American Psychological Association (1995a). *Principles for the recognition of specialties in professional psychology*. Washington, DC: Author.

American Psychological Association (1995b). *Principles for the recognition of proficiencies in professional psychology*. Washington, DC: Author.

American Psychological Association (1996a). *Association rules*. Washington, DC: Author.

American Psychological Association (1996b). *CRSPPP recommendation to recognize clinical neuropsychology as a specialty in professional psychology*. (Council of Representatives Agenda Item No. 13, August 8 and 11, 1996). Washington, DC: Author.

American Psychological Association (1997). *CRSPPP recommendation to recognize specialties and proficiencies in professional psychology* (Council of Representatives Agenda Item No. 11, August 14 and 17, 1997). Washington, DC: Author.

Commission for the Recognition of Specialties and Proficiencies in Professional Psychology (1996). *Minutes of meeting, October 13, 1996*. Washington, DC: American Psychological Association.

Council for the National Register of Health Service Providers in Psychology (1996). *National Register of health service providers in psychology*. Washington, DC: Author.

Council of Credentialing Organizations in Professional Psychology (1996). *Meeting summary, Santa Fe, NM, November 17, 1996*. Washington, DC: American Psychological Association.

Council of Representatives (1997). *Minutes of meeting August 14 and 17, 1997*. Washington, DC: American Psychological Association.

DeLeon, P. H. (1997). Proceedings of the American Psychological Association, Incorporated, for the year 1996: Minutes of the annual meeting of the Council of Representatives, August 8 and 11, 1996, Toronto, Ontario, Canada, and February 21–23, 1997, Washington, DC. *American Psychologist, 52,* 813–868.

Division of Clinical Psychology, American Psychological Association (1995). *Petition for the recognition of a specialty in professional psychology*. Washington, DC: Author.

Farberman, R. K. (1997). Public attitudes about psychologists and mental health care: Research to guide the American Psychological Association Public Education Campaign. *Professional Psychology: Research and Practice, 28,* 128–136.

Greenberg, S., Smith, I. L., & Muenzen, P. M. (1996). *Study of the practice of licensed psychologists in the United States and Canada.* Montgomery, AL: Association of State and Provincial Psychology Boards.

Krauss, D. A., Ratner, J. R., & Sales, B. D. (1997). The antitrust, discrimination, and malpractice implications of specialization. *Applied & Preventive Psychology, 6,* 15–33.

Merit Behavioral Care Corporation (1996). Specialty verification project offers objective criteria for specialist designation. *MBC Spectrum, 2,* 5.

Rosenfeld, M., Shimberg, B., & Thornton, R. F. (1983). *Job analysis of licensed psychologists in the United States and Canada: a study of responsibilities and requirements.* Princeton, NJ: Educational Testing Service.

Stigall, T. T., Bourg, E. F., Bricklin, P. M., Kovacs, A. L., Larsen, K. G., Lorion, R. P., Nelson, P. D., Nurse, A. R., Pugh, R. W., & Wiens, A. N. (Eds.) (1990). *Report of the Joint Council on Professional Education in Psychology.* Baton Rouge, LA: Joint Council on Professional Education in Psychology.

Stigall, T. T., & Hall, J. E. (1998). Licensure and specialization. In B. D. Sales & J. E. Hall (Eds.), *Specialization in psychology*. New York: Plenum.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.20

Individual Psychologist: Specialty Board Certification

WALTER B. PRYZWANSKY
University of North Carolina at Chapel Hill, NC, USA

2.20.1 INTRODUCTION 231
2.20.1.1 ABPP's Role in Specialty Board Certification 232
2.20.1.1.1 History 232
2.20.1.2 ABPP Specialty Recognition 233
2.20.1.3 Accreditation of Postdoctoral Training Programs 235
2.20.2 CURRENT ABPP BOARD CERTIFICATION REQUIREMENTS 236
2.20.2.1 Eligibility 236
2.20.2.2 The ABPP Examination Process 236
2.20.3 THE ABPP PROFESSIONAL PRACTICE APPRAISAL: SOME EXAMPLES 237
2.20.3.1 Work Samples 237
2.20.3.2 Oral Examination 237
2.20.4 PSYCHOLOGISTS AS "HEALTH SERVICES PROVIDERS" 238
2.20.5 THE IMPACT OF BOARD CERTIFICATION ON THE PROFESSION 238
2.20.6 FUTURE DIRECTIONS 239
2.20.7 REFERENCES 239

2.20.1 INTRODUCTION

The American Board of Professional Psychology (ABPP) is a voluntary professional credentialing organization. As the major, unitary organization in the US certifying specialty practitioners in psychology, ABPP serves the public need by awarding diplomas that certify psychologists who practice in various specialty areas of psychology. Board certification assures the public that the specialists designated by the ABPP have successfully completed an approved doctoral program, and successfully completed an examination process designed to assess the knowledge, skills, and experience required to provide high-quality psychological services in a particular specialty. The goals of ABPP are:

(i) to serve the public welfare by establishing and overseeing standards of specialty credentialing in psychology and, through its Members Boards, to implement examination procedures which lead to the systematic certification of specialists who meet those standards;

(ii) to serve the public by maintaining a registry of those holding diplomas in various specialties recognized by the Board and by providing public information relevant to the quality and access of specialized psychological services;

(iii) to function as a unitary organization with the public, psychology, and other professionals, external groups, regulatory agencies, and its Member Boards;

(iv) to support specialty boards in the development, validation, and administration of examinations to identify qualified practitioners in recognized psychology specialty areas; and

(v) to promote high standards of research, training, and the practice of psychology.

2.20.1.1 ABPP's Role in Specialty Board Certification

2.20.1.1.1 History

Interest in the credentialing of psychologists can be traced back to 1917 when the American Psychological Association (APA) established a committee to certify qualified APA members based upon written and oral examinations. At that time the committee distinguished between two types of practitioners: the *consulting psychologists*, who are competent in administering and interpreting psychological tests, and *clinical psychologists*, who are broadly knowledgeable about psychopathology. Their recommendation was to certify only the former types of psychologists. However, in spite of the establishment of a certification process, very few APA members became certified over the next 10 years, leading APA to abandon this effort (Goodstein, 1988). However, the need for a psychology practice credential continued to be expressed, so that following World War II, when applied psychology services were experiencing rapid growth, the time was ripe for the reconsideration of a system to certify psychologists. For instance, by 1945 only one state licensing law had been adopted and progress on that front did not appear to be imminent. Consequently, the APA facilitated the development of the American Board of Examiners in Professional Psychology (ABEPP), now the ABPP, as a separate and independent credentialing organization in 1947. The purpose of ABPP was to grant certificates of competence to psychologists who passed both generic and specialty examinations in the fields of professional psychology in which they practiced (ABEPP, 1947). The certificates were issued initially for the following specialties: clinical, counseling, and industrial-organizational psychology. Thus, the history of ABPP typically can be read as a reflector of the certification movement in psychology and specialization.

The ABPP national certificate, then, was seen to have the potential to serve as a uniform base for establishing standards for state licensure. But, rather than offer an entry level credential, the ABPP decided to certify only specialists at the highest level of competence. No doubt this decision reflected the potential of the movement for state licensure of psychologists. By 1948, there were three states with such licensing acts and, subsequently, 32 states by 1960. Instead, the ABPP diploma, awarded upon successful completion of an examination, came to represent a standard of excellence and the highest recognition of competence in professional psychology. Applications were accepted for review after the psychologist documented five years of successful practice in professional psychology.

From its inception, ABPP has been involved with credentialing specialties and promoting national standards of training and practice. Consequently, it should be no surprise that nearly from the time of its incorporation ABPP has been petitioned by groups of psychologists claiming to represent a specialty practice area, but for a variety of reasons, those proposals never became fully developed arguments. Meanwhile, the original three existing specialties received some additional validation as practice areas when the process of APA doctoral training program accreditation was instituted. It was not until 20 years later (1967) that a fourth specialty area was added, and then only in conjunction with APA's decision to make available accreditation of doctoral *school psychology* programs.

The four specialties (sometimes referred to as "original," "traditional," or "defacto" specialties in the literature) proceeded to credential psychologists utilizing similar criteria for determining eligibility and for conducting examination processes (Mayfield, 1987). Once the applicant met the eligibility criteria he/she was admitted into candidacy. The next step involved the preparation of a professional work sample which reflected a typical professional service provided by the candidate. The work sample was submitted with a transcript and/or audio/video tape of the session. Following acceptance of the work sample, an oral examination was then scheduled which included a discussion of the work sample, an *in vivo* observation of the candidate with a client, and discussion of ethical and substantive issues related to the "practice" samples and the candidate's specialty area as well as work activity. The basic dimensions of competency that were evaluated were as follows: (i) the effectiveness of the candidates efforts toward constructive intervention based on realistic assessment of the problem presented; (ii) awareness of the relevance of research and theory; and (iii) sensitivity to the ethical implications of professional practice. The examination process was conducted by three diplomates in one of six regions of the nation; the examination usually included one certified psychologist from a specialty other than the declared specialty of the candidate.

The above description covers what has been seen as ABPP's first two historical phases, and it's role in the specialization of the psychology field (Drum & Pryzwansky, in press). The first phase, "identification and clarification of initial specialty practice areas in psychology through intradiscipline consensus," was largely a consensus-seeking process, without specific criteria to differentiate emerging specialties from existing ones. The second phase involved the managing of the certification of specialists within the identified specialties. Thus, the procedures for the examination and certification of advanced competence received the attention of the credentialing organization.

By the late 1970s, the demands for particular specialty recognition began to intensify so that ABPP shifted its focus to this challenge. While committees within APA struggled with the development of principles and criteria for recognition, no formal policy was adopted. Drum and Pryzwansky (in press) note that this period represented a significant phase for the ABPP organization wherein recognition of new and emerging specialties, and incorporation of organizations supporting those specialties, was begun. ABPP explored the development of specialty recognition criteria which included the insurance of competence-based specialty-oriented examinations of clinical judgment. As a result this experience led to two related projects, specialty recognition and accreditation of postdoctoral training programs.

2.20.1.2 ABPP Specialty Recognition

The expansion of the psychologist's services to new work settings, along with the development of knowledge and interventions, resulted in continued and increased petitions to ABPP to facilitate the formal acknowledgment of this development in the field. APA's Board of Professional Affairs committees and task forces had laid the foundation for such a venture by fostering the development of principles and criteria for recognition of specialties in psychology, but no organizational mechanism was forthcoming to operationalize a process of recognition. However, ABPP was able to utilize the work of APA's Subcommittee on Specialization and in 1979 committed itself to establish "specialty examinations in any area of professional psychology where a critical mass of a group of psychologists:

(i) petition for such an area of specialization;

(ii) establish criteria for educational qualification;

(iii) establish criteria for minimum supervised practice;

(iv) establish criteria for examination processes and procedures; and

(v) state willingness to serve as examiners" (Pryzwansky, 1993).

A specialty was defined as an area in the practice of psychology that connotes special competency acquired through an organized sequence of formal education, training, and experience.

Thus, the unique element of ABPP recognition involvement is that the specialty was required to demonstrate its maturity by documenting the fact that a national group of psychologists existed who practiced in the claimed specialty and that financial viability existed to maintain a competency-based examination. What eventually evolved from those initial experiences and the use of existing APA documents was a refinement of guidelines and procedures for the development of an intensive self-study and review by ABPP affirming that standards of affiliation have been met.

The first stage of specialty recognition consists of the documentation of conceptual and practical viability of the applicant specialty. A detailed document is prepared for initial review by the ABPP Standards Committee before full review for recognition is entertained by the Board of Trustees. This document includes six major areas of information. In addition to identification of a credentialing board, information is sought regarding fiscal and long-term stability of the professional specialty. Congruence with ABPP's basic requirements for application must be demonstrated. A separate section of the self-study document includes demographic estimates of the practitioners, their geographic distribution, plus the names and description of predoctoral and/or postdoctoral training programs in the specialty which then completes this section. The "parameters of practices" constitute the third section. Included in this section are descriptions of the typical characteristics of clients and their problems or issues along with the listing of the primary practice settings. A most important section covers the operational definition of the specialty area along with the primary rationale for its recognition and affiliation with ABPP. As noted in the manual, the procedures that the applying Board will use in all their credentialing activities should relate directly to the professional competencies and activities that are defined. Applicant specialty areas are required to define up to 10 basic practice activities in the following professional competency areas: assessment, intervention, consultation, supervision, research and inquiry, consumer protection, and professional development.

Furthermore, the defined practice activities are judged in two ways, that is (i) essential or important but not essential, and (ii) unique vs. shared wherein the practice of other specialties are concerned. Separate sections of the document address how the specialty area has developed education and training standards or guidelines that assure essential activities are learned, as well as the practice standards or guidelines of the specialty. The specialty is required to present a literature review summarizing published research of the past decade on the efficacy of services funded by practitioners in the specialty. Finally, the credentialing activities which will be used in evaluating individual psychologists are solicited. Successful completion of this stage leads to acceptance into the two monitoring phases of affiliation. An implementation plan including final versions of all examination materials, timetable, etc. are then prepared. The specialty board must demonstrate the capacity to administer at least 30 examinations per year. The final phase involves monitoring of all aspects of the specialty boards credentialing process for at least the next two years. This phase includes an ABPP observer at the examination. During this time the specialty area will have full participating rights on the BOT. Successful completion of this phase will result in full affiliation of the Board and the recognition of the psychologists successfully passing these examinations as ABPP Diplomates.

Drum and Pryzwansky (in press) have noted that the recognition system which evolved has led to the identification of two types of specialties. The first type can be referred to as "unbounded" in the sense that some of the original specialties recognized by Psychology and ABPP were broad enough to be considered a professional field. The second type ("bounded") had more circumscribed boundaries with a dependence on the initial specialties in the sense that they have roots in one or more of those traditional specialties. Furthermore, the fact that many of the newer specialties developed and/or are operating within interdisciplinary contexts challenged the training criterion in a psychology specialty definition, and led to some of the founders of the interdisciplinary organizations unable to meet eligibility criteria for this credential. Also, in a related sense, many of the newer specialties have training programs in nonuniversity settings, particularly the postdoctoral areas. Consequently, the designation and/or accreditation system has yet to be fully developed in the psychology field, leaving specialty board organizations as the key to encouraging standards and practice setting guidelines.

As a result of the above described criteria and mechanism for recognition, two specialties were given full recognition and affiliated with ABPP in the mid-1980s: Clinical Neuropsychology (1984) and Forensic Psychology (1985). An updated criteria and procedures recognition and affiliation document evolved which described a three-stage process typically covering three years, *Application Manual for Specialty Affiliation* (ABPP, 1996). In the 1990s four additional specialties were recognized: Family Psychology (1991), Health Psychology (1991), Behavior Psychology (1993), and Psychoanalytic Psychology (1996). It should be noted that the practice activities in any specialty seldom are exclusive to that specialty and that most practice activities are shared with the general practice of professional psychology. The pattern of practice activities, including limiting the scope of practice, and focusing upon more complex or unique problems or technologies, is more relevant in defining a specialty together with advanced education, training, and experience (ABPP, 1996). Brief definitions of each of the 10 specialty boards are as follows:

(i) The American Board of Behavioral Psychology—Applied behavior analysis, cognitive behavior therapy, and cognitive therapy.

(ii) The American Board of Clinical Psychology—Assessment and treatment of mental, physical, emotional, and behavioral disorders.

(iii) The American Board of Clinical Neuropsychology—Evaluation of brain–behavior relationships and the treatment of cognitive, attentional, learning, and memory disorders.

(iv) The American Board of Counseling Psychology—Individual, group, and community interventions for emotional, behavioral, vocational, and mental health problems using preventive, developmental, and remedial approaches, and in the assessment, diagnosis, and treatment of psychopathology.

(v) The American Board of Family Psychology—Intervention and assessment in Family Psychology embracing the total family system including individuals, couples, families, and the intergenerational system.

(vi) The American Board of Forensic Psychology—The practice of psychology as related to law and the legal system.

(vii) The American Board of Health Psychology—The practice of psychology related to health, including the prevention, treatment, and rehabilitation of illness.

(viii) The American Board of Industrial/Organization Psychology—The application of psychology to problems of organizations and individuals/groups in organizational settings.

(ix) The American Board of Psychoanalysis in Psychology—A body of knowledge, a research method, and a treatment technique based on a description of organized unconscious, dynamic mental processes to aid in understanding emotions, attitudes, and belief systems to facilitate more satisfying, secure, and productive living.

(x) The American Board of School Psychology—The facilitation of learning and the promotion of mental health in schools and educational settings.

In addition to these 10 Specialty Boards, two specialty boards are in the second phase of the three-phase process of affiliation with the ABPP. These boards are developing examination materials and credentialing procedures consistent with required standards. The boards should be examining board members and candidates in the near future. The two boards are: The American Board of Group Psychology—The application of general psychological and specific group principles to change, modify, or adapt individual behavior, as well as patterning or modifying the behavior of groups; and The American Board of Rehabilitation Psychology—Assisting the individual with an injury or illness to achieve optional physical, psychological, and interpersonal functioning.

It is interesting to note that while the original four specialties developed at the doctoral level, some of the new or emerging specialties which have been ABPP recognized are clearly postdoctoral in the sense that they build on the doctoral training experience of the "traditional" specialties. Since APA recently established a Commission for the Recognition of Specialties and Proficiencies (DeLeon, 1995), ABPP has suspended its recognition mission to ascertain whether the Commission, speaking for the psychology field, will fill the original need of professional psychologists which prompted ABPP's actions since the late 1980s. Future specialties recognized by the Commission still will be required to validate the fiduciary, conceptual, and practice maturity involved in competency-based credentialing to become ABPP affiliated.

As a result of the ABPP growth in credentialing, structured changes have followed. The examination of each specialty is now developed, monitored, and revised by its respective specialty board. The credential, therefore, has evolved to a specialty board certification recognition of competence. The ABPP maintains minimum standards for applicant eligibility and examination beyond which the Specialty Board may add and/or insure relevance for the specialty. A nonprofit professional credentialing organization, ABPP, currently serves the public need by providing oversight of the certifying of psychologists competent to deliver high-quality services in various specialty areas of psychology.

2.20.1.3 Accreditation of Postdoctoral Training Programs

The emergence of ABPP postdoctoral specialties was a clear indication of the need for a mechanism to acknowledge the education and training standards for the development of competencies in those areas. Consequently, in 1991, ABPP organized the Conference on the Accreditation of Postdoctoral Programs in Psychology which resulted in the formation of the Interorganizational Council for Accreditation of Postdoctoral Programs in Professional Psychology (IOC). Organizations that made up the IOC included ABPP, National Register of Health Service Providers in Psychology, Association of Professional Psychology Internship Council, the Association of State and Provincial Psychology Boards, APA and later its Committee on Accreditation (COA), and representatives from the ABPP Specialty Boards.

The IOC first developed postdoctoral accreditation criteria as well as a Self-Study Document, both of which have subsequently been merged into the APA's COA *Guidelines and Principles for Accreditation of Programs in Professional Psychology*. A joint IOC–COA task force, in addition to addressing merging of documents, has been considering ways for continuing involvement of the IOC representation. Of interest is the requirement that "where postdoctoral residency programs offer education and training in a specialty, the program meets the additional standards promulgated by recognized specialty groups" (IOC, 1994). As a result, when a postdoctoral residency program applying for APA accreditation offers specialty training and requests to be so acknowledged, the entrance requirements must include criteria promulgated by recognized specialty groups.

A Council of Specialties is being formed with the broad aim of developing a general quality assurance process to provide guidance in the development of specialty-specific accreditation standards and self-study documents. Furthermore, such a council could conceivably serve as a mechanism for validating the existence of training standards for a specific specialty, as well as assist the COA in implementing accreditation procedures such as identifying review panels and site visitors for formal involvement in the COA process. Consequently, while much remains to be done in terms of postdoctoral specialty accreditation, significant steps toward the end have been realized.

2.20.2 CURRENT ABPP BOARD CERTIFICATION REQUIREMENTS

2.20.2.1 Eligibility

The purpose of a Specialty Board is to evaluate the qualifications of individual applicants' preparation, experience, knowledge, and competence in a specialty (ABPP, 1995). National in scope, the Specialty Board uniformly implements its policy and procedures in examining psychologists who volunteer to engage in this process. Independently incorporated specialty boards are stable organizations affiliated with ABPP which reflect the current development of the specialty. The individual Boards meet minimum eligibility standards set by the ABPP Board of Trustees, the latter entity serving as a quality control agent in their oversight of the Boards. What follows then is a general description of those minimum standards.

In terms of *education* the applicant psychologist must document that they hold a doctorate in psychology from a regionally accredited institution in the US. In addition the applicant must have had two years of *supervision* in the practice of his or her specialty (one of which may be at the predoctoral level) and three years of *experience*. The experience requirement can be satisfied in two ways. First, the applicant may have completed an accredited postdoctoral in the respective specialty plus one year of pre- or postdoctoral experience. Alternatively, the applicant can document three years of experience, one of which may be predoctoral. Finally, it is required that applicants who provide psychological services of a health care nature will hold a license to practice psychology in the state where he or she practices. The specialty may include additional education, training, and specific experiential criteria in determining eligibility, such as evidence of identification with the specialty, successful completion of a recognized postdoctoral training program, and specific types of practice experience at the postdoctoral level. It is interesting to note the shift from a five-year experience requirement which originally characterized the ABPP, and was a criterion consistent with the concept that certification recognized advanced competency among experienced practitioners in psychology. The addition of six specialties, several of which consider themselves postdoctoral specialties in the sense that specialist training takes place after completion of training in the "traditional" specialties at the doctoral level, has led to an evolution in the role of this credential. While advanced competency is axiomatic in the case of a postdoctoral specialist, *specialty board certification* is a more accurate descriptor of the credential and communicates the essence of the credential clearly both within the profession as well as to other professions and the public.

2.20.2.2 The ABPP Examination Process

Once eligibility has been determined, there are two stages in the examination process. First, the preparation of an acceptable work sample is required which will constitute the basis for one part of the face-to-face oral examination. Specialty Boards have specific guidelines for the development of these materials which should reflect the current practice of the candidate. Appropriate documentation is required. For example, in addition to assessment reports, therapy and/or supervision transcripts, and program evaluation models, the candidate is required to provide an analysis which includes appropriate background data, goals and/or objectives of the activity, relevant theoretical and research scheme, and an evaluative summation including a critique and plans for subsequent follow-up. Audio and/or video tapes of the activity are often included along with other relevant material. The examination proper, then, involves examination of the work sample(s), and may also include an on-site practice sample wherein the examining team can observe the candidate interact under *in vivo* conditions with a live client, or respond to specially prepared videotaped observations of an actual client or an interview conducted with an actual client. The variability noted here reflects the uniqueness of each specialty and the particular avenues of opportunities to assess the clinical practice of the specialty. Furthermore, specialties utilize one or more of several examination models ranging from the assessment center approach, wherein candidates move from station to station, to the more traditional *in vivo* examination approach. All specialty examinations also have an ethics component, again with varying assessment approaches, as well as a component which deals with the substantive content of the specialty.

The evolution from regional examination centers, staffed by representatives from the four traditional specialties, to a national specialty board examination model has facilitated two outcomes. A national board promotes reliability, validity, and consistency objectives in assessment by reducing variability factors which can emerge due to regional emphases. Second, the relevance of content and practice situations is enhanced for the candidate when control of the examination is maintained by the Specialty Board. In turn, the ABPP Board of Trustees serves an oversight function, insuring that fair, sound, and reasonable examinations are conducted, and the examinations themselves are evaluated and updated. This oversight function is assigned to the ABPP Standards Committee which maintains minimum standards and expectations for each of its Specialty Boards, promotes development of the examination process, and facilitates communication and sharing among these Boards with the objective of increasing their relevance for the individual psychologist as well as the profession. The value of any credential, and in particular a competency oriented one, depends on the integrity of the examination.

Finally, ABPP Diplomates, as those holding specialty board certification are called, also are members in their respective specialty Academy. The academies serve as membership organizations promoting continuing education, public service, and advocacy functions for the specialty.

2.20.3 THE ABPP PROFESSIONAL PRACTICE APPRAISAL: SOME EXAMPLES

As noted above, all of the 10 ABPP specialties require candidates to submit examples of typical, relatively current, professional practice activities in the form of a work sample. Following the acceptance of the work sample, an interactive oral examination conducted by three trained psychologist examiners from the candidate's specialty is then conducted. The latter is usually individually administered, although at this time two specialties utilize an assessment center type model of examination. Since state licensure is required of all candidates in practice, those applicants have passed the additional objective examination developed and administered by the Association for State and Provincial Psychology Boards. Only one ABPP specialty, clinical neuropsychology, requires successful completion of an objective examination covering basic content in that specialty. This 100-item written examination contains content covering the lifespan from prenatal development through senescence. The five content categories are (i) neuropsychological assessment; (ii) clinical neuropsychology; (iii) basic and clinical neuroscience; (iv) behavioral (clinical) neurology; and (v) general clinical psychology. The majority of the test items fall within the middle three categories.

What follows are some illustrative descriptions of the process utilized by the three specialties of clinical psychology, clinical neuropsychology, and counseling psychology. Obviously, reference should be made to the credentialing materials of each specialty for a complete and official description of the process.

2.20.3.1 Work Samples

The work sample generally is considered as an example of the candidate's professional practice style and competence in carrying out activities reflective of his/her practice. The Counseling Psychology Board (CP) requires one work sample which must consist of two parts, one of which is focused on assessment and the other on intervention, supplemented by audio and/or video tapes of the sample along with an extensive written documentation. The Clinical Neuropsychology (CN) work sample consists of two cases, each of which must include the neuropsychological evaluation basis and the conceptual basis for the interventions employed, as well as the steps taken to assure an adequate test of the efficacy of the interventions. The case supporting materials are outlined. The Clinical Psychology work sample consists of two videotaped samples of approximately 50 minutes each, one depicting an unrehearsed psychological assessment effort, the other depicting an unrehearsed intervention effort, drawn from recent typical clinical practice. Again, documenting material is specified in each instance.

2.20.3.2 Oral Examination

In terms of the oral examination proper, the same substantive areas are involved and differences exist only in the assessment strategy which is utilized. Counseling Psychology begins the examination with a prerecorded interview with a client constituting an initial intake session (the observed, taped interview must be prepared within 60 days prior to the scheduled oral examination). Following viewing by the examination committee, the candidate will be questioned regarding his/her conclusions to ascertain such information as the understanding of the client and the presenting problem. The work sample, earlier described, is compared and contrasted then with the case. Two additional parts of the examination remain. The candidate is next presented with two written vignettes from which one is chosen, and one or more ways of addressing the presenting problem with appropriate interventions other than individual counseling are presented. Finally, the candidate is asked to discuss the ethical and legal implications of a presented written vignette.

Clinical Neuropsychology begins their oral examination with a fact-finding exercise in which initial neuropsychological consultation in a case is expected. Both a child and adult case are presented for this segment of the examination and the candidate chooses the case and begins to request information on which eventual conclusions are to be made. The work sample is then used as a point of departure for examiners' questions to allow the candidate to demonstrate a reasonable, rational, and defensible approach to patient evaluation, treatment, and report writing. Finally, the candidate is presented with a neuropsychologically related vignette and asked to identify issues. The candidate is asked questions about his/her own professional practices, professional involvement, and research activities.

The Clinical Psychology specialty begins with an examination of the two work sample materials and the professional statement. This segment of the examination is then followed by ethical and legal issues. The candidate is presented with an ethical vignette to review and discuss. This vignette, and the ethical dilemma identified from his/her own practice and presented as part of the candidate's professional statement, will form the context for this segment of the examination.

The above abbreviated examples of examinations from three of the 10 ABPP specialties provide a sample of the competency-based examinations used as part of the ABPP Board certification process. More detailed and current information regarding these three specialties as well as the other ABPP specialties may be obtained by writing to the individual boards at ABPP's Central Office at 2100 East Broadway, Suite 313, Columbia, MO 65201-6082.

2.20.4 PSYCHOLOGISTS AS "HEALTH SERVICES PROVIDERS"

An interesting ABPP development occurred in the early 1970s in response to discussions taking place then regarding national health insurance plans. It was noted that psychologists usually were not included in many of the proposed plans in part due to the fact that insurance carriers had no formal mechanism for determining which credentialed psychologists were qualified to offer health services. Although ABPP credentialed psychologists with a specialty practice, some of those did not offer such services, most notable the large percentage of industrial/organizational psychologists. More relevant to this need to identify psychologists in health services was the fact that state licensure was generic so that it provided no information regarding this question. In large measure as a result of ABPP's credentialing history, the APA Board of Directors specifically requested in 1974 that ABPP establish a National Council for the Register of Health Service Providers in Psychology. The ABPP central office organized such a meeting for the Council in 1974 and funded the organization of the National Register. Further discussions considered the development of "junior diplomate" and the original interest in simply establishing a system of identifying health service providers (HSP) from among all licensed psychologists. The National Register continued its affiliation with ABPP during its start up, but given the demand for a Registry system (7500 applicants in a three-month period), the Council soon evolved into a completely separate organization with the purpose of validating psychologists' credentials as meeting criteria for listing as a health service provider in psychology. Subsequently, some states have added a similar validation from the licensees separate from the license itself.

Most of the ABPP specialties include the basic HSP requirements within the board certification process, and as the psychology credentialing bodies move toward an integrated system for credentialing psychologists, this relationship may become even more apparent. Such systems would include the doctoral internship, postdoctoral training, licensure, listing in the National Register, and ABPP board certification in a specialty; a connection with doctoral and postdoctoral accreditation would complete the system.

2.20.5 THE IMPACT OF BOARD CERTIFICATION ON THE PROFESSION

As the professional has matured so has the ABPP Diplomate evolved as a credential for psychologists. Originally conceived as the practice credential in the absence of state licensure, the ABPP credential then took on the aura of a validation of expertise associated with senior status in a professional and since the late 1980s increasingly is conceptualized as an examination-based endorsement of specialty practice competencies. Clearly, several of the professional psychology specialties (e.g., clinical neuropsychology and forensic psychology) are well along the path leading to the expectation among employers and consumers that attainment of board certification is necessary for practice. Similarly, as the APA continues with the formal process of recognition of specialties (and proficiencies), the natural outcome will require a system for the validation of training and practice in the specialty. Thus, the current ABPP Boards can be viewed as anticipating, if not influencing, the specialization movement within professional psychology. ABPP is the only unitary organization recognized by the profession as certifying specialty practitioners in psychology.

The two aforementioned specialties are increasingly recognized within their practice settings, that is, medically-oriented systems and the judicial system. In addition, the US Department of Defense and the US Public Health Service are committed to implementing pay increments tied to employment longevity within their policies, but implementation has not been realized. At this time two states (MO and OR) recognize psychologists with ABPP board certification employed in the Department of Mental Health system. Over 40 states have a reciprocity type stipulation for psychologists holding the ABPP written into their psychology licensing regulations. Finally, until revision of the APA doctoral program accreditation criteria (1996), which has become "outcome" oriented with less attention to the structure and process of the training program, the qualifications of the directors of the training program ironically have been de-emphasized in a sense so that evidence such as the ABPP credential, or APA Fellow Status, are not mentioned, and thereby no longer encouraged through that mechanism. Discussions regarding accreditation of postdoctoral training in a specialty have noted the importance of the requirement of board certification for the director of the training program.

Another influence of the ABPP certificate subtle as it is, but nevertheless important, concerns the professional development concept in psychological practice. Preparation for the ABPP peer examination has many intangible rewards for psychologists undergoing the activity. The preparation of a work sample leads to a unique structured self-reflection exercise; coupled with the examination process itself the individual psychologist will have experienced activities that contribute to the practitioner's growth and development. The attainment of an independent, rigorous verification that the standards and competencies required in a specialty area were met is an important personal as well as professional reward. Likewise, the continuing upgrading of the examination by the individual boards can lead to and/or reinforce training developments in the specialty. At some future date, when coordination among training programs and accrediting as well as credentialing bodies are better organized, documenting the fact of board certification for graduates can serve as important feedback to the training program and lead to improvement in the quality of professional psychology education and practice.

2.20.6 FUTURE DIRECTIONS

Clearly, activity in the development of tools assessing professional practice will continue and enhance the relevance of the examinations in the respective specialties. The use of the still preferred *in vivo* examinations will no doubt continue to be used as a yardstick against which more cost-effective, increasingly standardized simulated approaches will be tested, approaches which give examiners and candidates more flexibility in dealing with problem-solving scenarios. Also, it is anticipated that computerized assessment alternatives will begin to drop in cost and receive renewed attention. The end result should be even more viable assessment procedures capable of promoting professional development while serving as a measure of competency in practice.

By contrast, the future recognition of specialties will most likely slow in terms of the immediate future, partly to give the applied psychology arena a chance to catch its breath, and partly for the APA process headed by CRSPP to interface with the movements within the credentialing community. At the same time, subspecialities and proficiencies will more than likely continue to grow. Finally, the next decade should bring the realization of a seamless education, training, and accreditation–credentialing system wherein all components are fitted together. The result should be greater protection for the public along with improved service stemming from stronger training programs.

2.20.7 REFERENCES

American Board of Examiners in Professional Psychology (1947). Report of the American Board of Examiners in Professional Psychology. *American Psychologist, 2*(502), 512–517.

American Board of Professional Psychology (1995). *Standards for Boards of the American Board of Professional Psychology*. ABPP.

American Board of Professional Psychology (1996). *Application manual for specialty affiliation*. ABPP.

DeLeon, P. H. (1995). Proceedings of the American Psychological Association, Incorporated, for the Year 1994: Educational Affairs. *American Psychologist, 50*(8), 665–666.

Drum, D. J., & Pryzwansky, W. B. (in press). ABPP and the development of specialization. In B. A. Sales & J. E. Hall (Eds.), *Specialization in psychology*. New York: Plenum.

Goodstein, L. D. (1988). The role of the American Psychological Association in legitimizing the practice of psychology: A historical perspective. *Professional Practice of Psychology, 9*(1), 1–9.

Mayfield, P. N. (1987). Certification of psychologists by the American Board of Professional Psychology. In B. A. Edelstein & E. S. Berler (Eds.), *Evaluation and accountability in clinical training* (pp. 283–297). New York: Plenum.

Pryzwansky, W. B. (1993). Credentialing in an era of specialties: The ABPP Diplomate. *The Diplomate, 13*(1), 5–7.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.21
Hospital Staff Membership: United States

KRIS R. LUDWIGSEN
Kaiser Foundation Hospital, Martinez-Walnut Creek, CA, USA

2.21.1 INTRODUCTION 241
2.21.2 PSYCHOLOGISTS' ROLES IN HOSPITALS 241
2.21.3 OBTAINING INDEPENDENT PRIVILEGES 242
2.21.4 HOSPITAL BYLAWS 243
2.21.5 REQUIREMENTS FOR MEDICAL STAFF MEMBERSHIP 243
2.21.6 CREDENTIALING AND PRIVILEGING 244
  2.21.6.1 Privileging by Patient Categories 244
  2.21.6.2 Privileging by Assessment and Treatment Procedures 244
  2.21.6.3 Privileging by Intensity of Containment 245
2.21.7 ETHICAL ISSUES 245
2.21.8 MANAGED CARE IN HOSPITAL PRACTICE 246
2.21.9 FUTURE DIRECTIONS 247
2.21.10 SUMMARY 247
2.21.11 REFERENCES 248

2.21.1 INTRODUCTION

One of the primary issues which has marked the coming of age of clinical psychology as an independent health care profession is the attainment of hospital practice privileges to include independent admission, treatment, and discharge privileges of the most acute and disturbed patients. The clinical role of psychologists in hospitals began in World War II, concurrent with the birth of clinical psychology as a profession. Psychologists initially served as psychodiagnostic consultants evaluating the cognitive, intellectual, and neuropsychological impact of war injuries in Veterans Administration (VA) hospitals with limited responsibilities and expertise. Since then US psychologists have been practicing in state mental hospitals, military hospitals, mental retardation facilities, medical surgical hospitals, chemical dependency facilities, rehabilitation and geriatric facilities, private psychiatric hospitals, VA hospitals, health maintenance organization hospitals, and day treatment centers with some degree of authority. For 50 years, psychologists have been demonstrating their expertise in inpatient care and their value to the hospital as an institution.

2.21.2 PSYCHOLOGISTS' ROLES IN HOSPITALS

Psychologists' roles within hospitals have extended from salaried staff to independent contractors to attending or consulting staff. Psychologists have held the salaried positions of hospital administrator and assistant administrator, state hospital superintendent, and hospital owner. As clinical services director, the psychologist is responsible for overseeing the range of clinical programs and services offered. As director of psychological services, the psychologist determines policies for the psychology staff, oversees training of psychology interns and determines what services salaried psychologists provide within the facility. Psychologists have served as program directors for the adolescent unit, the dual diagnosis program, the day treatment and partial hospitalization programs, as well as quality assurance coordinators. As independent contractors, psychologists have contracted with hospitals to offer, for example, biofeedback, neuropsychological assessment, assertiveness training and anger management groups, groups to manage stress and chronic pain, eating disorders, sexual abuse and incest survivors groups on an ongoing or time-limited basis.

Where consistent with state law, psychologists have held designated medical staff privileges with formal credentialing by the hospital. The Joint Commission for Accreditation of Healthcare Organizations (JCAHO) identifies three levels of medical staff privileges: active staff, courtesy staff, and consulting staff. Both active and courtesy staff may admit patients to the hospital and serve as the primary attending clinician. However, active staff are more involved with the hospital since they have voting privileges within the hospital governance, are required to serve on hospital committees, and are on the call roster to receive emergency referrals. Courtesy staff have the same level of clinical privileges but may not admit more than a specific number of patients per year, are not on the hospital call roster, do not serve on hospital committees, and do not have voting privileges. Consulting staff provide specialty diagnostic or treatment assistance to the attending clinician, such as neurological or neuropsychological assessment, pain management, or biofeedback. Medical staff privileges may be held concurrently with a part-time salaried or independent contractor position.

2.21.3 OBTAINING INDEPENDENT PRIVILEGES

Although it has been a protracted and arduous struggle, psychologists in the USA could not have attained their present extent of hospital privileges prior to passing independent licensure and certification laws which established a certain parity with physicians on an outpatient basis. The hospital practice movement began with several psychologists in various states, military hospitals, and rural areas over the last 30 years, exercising informal privileges in treating patients admitted to the hospital from their own outpatient practices. Collaborating psychiatrists and physicians signed orders and provided informal, collegial support until the former Joint Commission for Accreditation of Hospitals (JCAH) reviewed these procedures and requested that hospitals rescind informal privileges for psychologists. Faced with the possible threat of losing JCAH accreditation, hospitals complied. The JCAH was authorized by the US Department of Health, Education and Welfare to regulate hospitals, and was thereby empowered to promulgate and enforce rules. Founded in 1951 by various physician groups to establish consistent standards for quality of care, for example, how frequently physicians had to see their hospitalized patients, the JCAH was the primary organization accrediting hospitals and health care facilities. The JCAH standards explicitly required physicians to "take medical responsibility for" patients admitted to the hospital with the support of federal statutes and state legislation (Wiggins & Ludwigsen, 1991). The American Psychological Association (APA) and the Association for the Advancement of Psychology (AAP) had petitioned the JCAH for formal status on hospital medical staffs and representation within the JCAH (Zaro, Batchelor, Ginsberg, & Pallak, 1982). However, the decade of struggle by organized psychology did not bear fruit until the AAP notified the Federal Trade Commission (FTC) of the implications of restraint of trade within the health care sector and the FTC determined that professions as well as businesses were subject to antitrust action. Subsequently, hospitals became aware of the antitrust implications of excluding psychologists as a health care profession from hospitals (Bersoff, 1983).

In 1978, after three years of legislative efforts, California was the first state to grant statutory permission for hospitals to include psychologists on the medical staff and to accord them admitting and attending privileges (Enright, Resnick, Ludwigsen, & DeLeon, 1993). Since then, several other states have obtained hospital practice privileges for psychologists. Currently, psychologists have full privileges in various medical schools and are on the medical staff of VA hospitals on a facility-by-facility basis where permitted by state law (Resnick & Morris, 1997).

As psychologists became recognized as health care professionals within hospitals, they began to request guidelines from the APA on hospital practice for psychologists. In response, the APA commissioned *A hospital practice primer for psychologists* (Committee on Professional Practice of the Board of Professional Affairs, 1985) and *Hospital practice advocacy issues* (Committee on Professional Practice of the Board of Professional Affairs, 1988) and adopted the *Guidelines on hospital practice privileges: Credentialing and bylaws* (Resnick, Enright, & Thompson, 1991). Other APA standards relevant to hospital practice are discussed in *Professional liability and risk management* (Bennett, Bryant, VandenBos, & Greenwood, 1990) and *The ethical principles of psychologists and code of conduct* (APA, 1992). Finally, *The psychologist's legal handbook* (Stromberg et al., 1988), discussions of hospital politics (Rozensky, 1991), and of training psychologists for hospital practice (Ludwigsen & Albright, 1994) offer useful guidance to psychologists pursuing independent privileges in hospitals.

Despite the passage of state legislation explicitly permitting hospital privileges for psychologists, the Department of Health Services in California promulgated regulations stating that psychiatrists alone could be responsible for the patient's diagnosis and treatment. In response, a group of independent psychologists filed suit to challenge these regulations as inconsistent with state law. The *CAPP v. Rank* suit was originally upheld by the California Superior Court in 1985, but was appealed with a series of rulings favoring the medical establishment in 1988. The final decision by the California Supreme Court stated that in diagnosis and treatment, psychology and psychiatry are on equal footing where neither is subject to constraints from which the other is free (CAPP v. Rank, 1990).

Faced with the possibility for an antitrust suit asserting discrimination against psychologists as a profession, the JCAH modified its statement on the composition of the medical staff in 1985 to state that the organized medical staff "includes fully licensed physicians and may include other licensed individuals permitted by law and by the hospital to provide patient care independently" (JCAHO, 1992, p.55), and added an APA representative to its oversight committee for mental health care standards in 1988. However, while state psychological associations were mounting campaigns to obtain hospital privileges for psychologists, the JCAH removed staff restrictions, and the *CAPP v. Rank* suit was decided in favor of psychologists practicing independently in hospitals, another barrier to hospital practice emerged. The federal Social Security Administration required physician supervision of nonphysician professionals for hospitals receiving federal reimbursement through Medicare and Medicaid funds (Social Security Administration, 1993). Through the advocacy efforts of the APA the Social Security Acts Amendments of 1994 now guarantees licensed psychologists independent practice and reimbursement for services rendered in a hospital where consistent with state law under the Psychology Services in Hospitals section (Social Security Acts Amendments, 1994).

2.21.4 HOSPITAL BYLAWS

Where state law is permissive, enabling, or silent regarding hospital privileges for psychologists, the next step is to review the current hospital bylaws to determine the appropriate medical staff status of psychologists within the facility. Hospital bylaws are public documents available through the board of directors, the medical staff coordinator or hospital administrator, which define the standards for professional staff membership. Bylaws specify the staff categories available to applicants, and the clinical privileges available within each category (Rozensky, 1997). Modifications are proposed in the Bylaws Committee and then voted upon by the active staff. Changes proposed should be consistent with the wording and principles of practice espoused by the particular hospital's bylaws. Extending the bylaws to enable psychologists to become active, courtesy, or consulting staff members involves a certain degree of research and preparation, for example, on psychologists' current competencies, the role of psychiatrists and other clinicians in the facility, the state nurse practice act, and psychologists' legal scope of practice. Discussions with medical staff colleagues, the Bylaws Committee, the administrator and the board of directors may also be necessary to effect changes in the bylaws. Rozensky (1991) notes that it is more effective to change the rules from within an organization than to challenge them as a newcomer without a full appreciation of the hospital's history regarding the roles of clinical psychologists.

2.21.5 REQUIREMENTS FOR MEDICAL STAFF MEMBERSHIP

Although hospitals are required to be consistent and nondiscriminatory in their policies regarding staff membership, they are not obligated to accept all applicants. Stromberg et al. (1988) note that the entire staff of a hospital or a particular department may be closed to applicants, with appropriate documentation of rationale, if the inpatient facility is inadequate to treat patients likely to be admitted by the applicant, or if the existing staff is sufficient to meet the hospital's needs. Therefore, in applying for privileges, it is important to consider what expertise (e.g., neuropsychology) or benefits (e.g., patient admissions) the practitioner will be contributing towards developing a mutually beneficial relationship with the hospital.

The specific requirements for hospital staff membership differ from hospital to hospital. Stromberg et al. discuss relevant criteria for staff membership including level of education, clinical internships, fellowships or other training, experience, licensure, and current clinical competencies. The American Psychological Association recommends at a minimum: (i) a doctoral degree from a regionally accredited program training psychologists; (ii) state licensure at the independent practice level; and (iii) one year or more of supervised experience in a health care facility with an inpatient component (Resnick et al., 1991). By comparison, the American Psychiatric Association (1986) recommends that privileges be based on verifiable data such as graduation from an accredited medical school, record of postgraduate or specialty training, licensure, specialty board certification, prior experience including previous hospital appointments, letters of recommendation, and lack of malpractice history.

Stromberg et al. note that hospital staff members may be required to carry a certain level of malpractice insurance due to the higher acuity level of hospitalized patients and that privileges may be denied if there are pending or completed professional liability actions or loss of membership or privileges at other inpatient facilities. Privileges may also be denied if there is evidence of significant health impairment, of difficulty in working with others, or if the practitioner's personal style is documented to jeopardize quality of care. Finally, the hospital may determine residential limits for attending staff so that patients can be treated in a timely way in an emergency.

2.21.6 CREDENTIALING AND PRIVILEGING

According to JCAHO standards, all practitioners who treat patients in a hospital must be credentialed whether they are members of the professional staff, independent contractors, or salaried staff. The process of delineation and renewal of clinical privileges is described in the hospital bylaws. Credentialing refers to providing appropriate documentation for the requirements to become a member of the hospital staff. Although the bylaws specify the staff categories available and the clinical privileges available within each, the granting of specific clinical privileges is a separate process from designating a staff category.

Privileging refers to specifying the clinician's scope of practice within each staff category concerning the various patient assessment and treatment procedures. In addition, temporary privileges may be extended relative to an emergency admission, and provisional privileges may be granted to practitioners seeking attending privileges while undergoing the proctoring process.

Credentialing procedures define psychologists, similar to physicians, dentists, and podiatrists, as independent practitioners within the hospital, and define their scope of practice within each inpatient facility. The *JCAHO Accreditation manual for hospitals, 1992* (JCAHO, 1992), Standard MS.2.15, states that apart from the system for granting, renewal, or revision of clinical privileges, the granting of clinical privileges is hospital specific and based on current competence. The hospital may be as specific or general as is practical and necessary. Any limitations on the practitioner's privileges to admit and treat patients should also be specified, for example, that privileges regarding assessment procedures do not include neuropsychological assessment.

Models of privileging for psychologists can be summarized under three categories: privileging by patient categories, privileging by assessment and treatment procedures, and privileging by intensity of containment (Ludwigsen, 1993).

2.21.6.1 Privileging by Patient Categories

In several hospitals, units and programs are organized around the needs of different patient groups, such as children, adolescents, adults, and geriatric patients, as well as chemically dependent, rehabilitation, mentally retarded, eating disorder, and forensic patients. Privileges to treat specific patient categories can be granted based on the psychologist's demonstrated training, experience, and current expertise, and may be considered an extension of defining the psychologist's outpatient specialties.

2.21.6.2 Privileging by Assessment and Treatment Procedures

The privileging process can be conceptualized according to specialized expertise such as cognitive-behavioral therapy, pain management, neuropsychological assessment, biofeedback, clinical hypnosis, and marital and family psychotherapy, which psychologists will be contributing to the hospital. Specialized credentialing standards may be required for neuropsychology, biofeedback, and clinical hypnosis. The APA Division of Clinical Neuropsychology defined the independent practitioner of neuropsychology as having completed two or more years of appropriate supervised training in delivering neuropsychological services in a clinical setting, including a one-year predoctoral internship, postdoctoral training, and clinical neuropsychology experience ("Definition of a clinical neuropsychologist," 1989). Board certification through the American Board of Clinical Neuropsychology further includes a rigorous examination of credentials and clinical competence. The hospital may require documentation of specialized training in biofeedback and clinical hypnosis, for example, certification by the Biofeedback Certification Institute of America or the American Society of Clinical Hypnosis, as appropriate to the community standard of practice and quality of patient care.

2.21.6.3 Privileging by Intensity of Containment

In psychiatric hospitals and state mental hospitals, privileges can be granted for treating patients on the locked unit, the acute care unit, the open unit, and the day treatment unit, based on the assumption that intensifying levels of acuity require greater clinical expertise, as in identifying the potential for and managing suicidal and assaultive behavior. Psychologists applying to admit and treat patients on the locked unit would probably be expected to show documentation of past training and/or experience with suicidal, assaultive, or violent patients, or to complete a period of proctoring on a specified number of acute patients before assuming independent patient care responsibilities. However, many psychologists pursuing admission privileges at such hospitals are requesting the authority to admit and treat suicidal patients from their own practices within an inpatient setting to provide continuity of care.

Proctoring is the most important part of the privileging process because it involves direct observation and evaluation of the prospective staff member's ability to make appropriate and timely assessment and treatment decisions, and to initiate collaborative implementation with the multidisciplinary hospital team. Quarterly proctoring reports are submitted to the Credentials Committee which recommends any changes in staff status after the provisional period. Peer reviews of patient charts may subsequently be used as a basis for recommending recredentialing.

The American Psychiatric Association (1986) states that the ultimate decision regarding who practices in the hospital and what credentials they need for privileges resides with the individual hospital, with quality of patient care as the guiding principle. Recognizing that reviews of psychologists' applications for privileges were often haphazard and conducted by credentials committees having little familiarity with psychologists' training and practice, and that hospitals and medical staffs could be sued for negligent and inappropriate credentialing standards, the APA asserted that the responsibility to ensure quality of patient care must include an orderly, fair, and systematic review of psychologists' credentials. Psychologists must play the major role in delineating clinical privileges, specifying the necessary criteria, reviewing credentials and recommending specific privileges for psychologists, in addition to specifying criteria for renewal of privileges and participating in that process (Resnick et al., 1991). In establishing credentialing and privileging procedures for their discipline at specific inpatient facilities, psychologists should survey APA guidelines, appraise the character and mission of the specific hospital, and review credentialing and privileging procedures for related health care professions.

2.21.7 ETHICAL ISSUES

Linton (1993) and Pope (1990) identify a number of ethical issues of which psychologists entering hospital practice should be cognizant:

(i) Are the lines of clinical responsibility sufficiently clear so that no patient's needs are inadverently overlooked? Is it clear at all times who is responsible for patient management?

(ii) Is the patient adequately protected in the case of suicidal or assaultive impulses or sexual advances from other patients?

(iii) Are clinical decisions made on the basis of the patient's best interests? Psychologists must be sensitive to their responsibility to avoid premature termination of treatment (i.e., patient abandonment) despite the denial of authorization for further treatment by third-party payers. Pope notes that psychologists must also be aware of the financial and political forces that determine a hospital's decisions regarding admission, resource allocation, length of stay, follow-up upon discharge, and the likelihood of readmission in addition to the clinicians' treatment priorities.

(iv) Is the hospital charging appropriately for psychologists' services? In one hospital psychologists' diagnoses were being changed by the billing department to obtain higher reimbursement.

(v) Are there adequate safeguards to protect the patient's confidentiality? Do patients adequately understand who will have access to information about their treatment? Patients may not realize that the details of their private lives can become public knowledge through inpatient records, team meetings, or informal hallway and nursing station "consults." Inpatient treatment is, by its collaborative nature, considerably less confidential than the one-to-one relationship of the outpatient setting. Psychologists must therefore balance the need to discuss and document patient care with protecting potentially harmful details from disclosure. Protecting confidentiality is more difficult for patients in a rural hospital with their interwoven community relationships than in a large hospital in an urban area. The best defense against potential charges is a clear rationale, documented in advance on the patient's chart, concerning the psychologists' decision to include or exclude specific information (Linton, 1993).

(vi) Does the psychologist have the education, training, and experience to carry out clinical responsibilities with the specific patient population? Psychologists must be careful to define their boundaries and limits regarding the kinds of patients, disorders, assessments, therapeutic procedures, and treatment settings they are qualified, by training and experience, to handle, as reinforced by APA ethical guidelines.

(vii) When psychological consultation is requested, as in a medical surgical hospital, does the patient understand the reason for the consultation and provide an informed consent to it?

The hospital milieu is a more political environment than the outpatient psychotherapy setting. Turf issues and rivalries between professions, tensions between administrators and employees, conflicts between practitioners concerning different theoretical orientations or basic values, and competitive struggles for power, income, or recognition are to be expected. To ameliorate potential ethical problems in hospital practice, Linton recommends thinking issues through, consulting with an ethics partner or committee and documenting in advance. Defending oneself after the fact is much easier if sincere attention to and clear recognition of one's limits can be documented through charts, reports, or oral consultations.

2.21.8 MANAGED CARE IN HOSPITAL PRACTICE

In the 1980s, hospitals began to experience pressures toward fiscal accountability and cost containment. Medical costs were increasing at twice and mental health costs at four times the rate of inflation (Wiggins & Ludwigsen, 1991). Since health care in the USA is largely employer based, many employers and third-party payers, including indemnity plans and state and federal health care plans, began searching for ways to lower their escalating costs. Managed care organizations (MCOs) are designed to reduce utilization, for example, through preauthorization and concurrent utilization review, or to reduce costs, for example, through claims review and shifting to less expensive alternatives to hospitalization. With inpatient care consuming some 65–80% of mental health reimbursement, it was inevitable that managed care would target inpatient treatment for intensive monitoring and controls (Broskowski, 1994). To reduce costs, MCOs focused on decreasing the number of admissions and length of stay, utilizing hospitalization primarily for containment and stabilization of, for example, suicidal, assaultive, or psychotic behavior. The patient is discharged after a brief stay to the least restrictive safe setting, such as a day treatment program.

Continuous quality improvement (CQI) approaches have been adapted from manufacturing industry to health care to focus on demonstrating successful treatment outcomes, cost effectiveness, quality control, and consumer/patient satisfaction. They include the adoption of imposed protocol and practice guidelines based on research (Shueman, Troy, & Mayhugh, 1994), for example, for treatment of major depression. Practitioners and hospitals are being treated as contracting health care providers and are being asked to demonstrate accountability to MCOs through treatment outcomes assessment and patient satisfaction.

The draconian measures instituted by MCOs to control costs have resulted in premature discharge from hospitals in some cases; however, legal challenges to MCO decisions have been successfully raised. In *Wickline v. the State of California,* Marshall and Muszynski (1993) note that third-party payers (MCOs) can be held legally accountable when clinically inappropriate decisions result from cost-containment priorities, for example, when appeals made on a patient's behalf for hospital treatment are arbitrarily ignored, unreasonably disregarded, or overridden. The implication of the court's decision is that the priority of cost containment ends when quality of care begins to be compromised. Although managed care has been a largely unregulated industry undergoing considerable growing pains, the second generation of MCOs is focusing on access, demonstrated outcomes, and quality in addition to cost-control.

2.21.9 FUTURE DIRECTIONS

The future of psychology in hospital practice is related to the interaction of several factors: (i) the development of psychology as a profession, including the outcome of advocacy efforts at the federal, state, and local level; (ii) the development of clinical and scientific research; (iii) and the evolution of hospitals within the health care delivery system, and the forces impacting them, such as managed care directives. In contrast to medicine, nursing, and social work, psychology is a relative newcomer to the health care professions practicing in hospitals, and the problem of developing a consistent professional identity is one with which the profession is still struggling. Matarazzo (1977) notes that professions usually begin with haphazard, inconsistent credentialing and proceed through the loose organization of practitioners in an effort to establish the identifying and unique features of the field. This appears to be the current status of psychologists in hospital practice.

Through organized advocacy efforts at the federal, state, and local level, psychologists have broadened their scope of practice to include practicing independently in hospitals. This has required a sustained effort by committed groups of psychologists to overcome varying degrees of resistance, slow progress, entrenched opposition, and interdisciplinary warfare. The legislative issues with the greatest implications for psychologists' autonomy in hospital practice are those supporting independent admitting privileges and prescription privileges. Such privileges are essential to implementing a psychological model of care in contrast to the more traditional disease model. By its nature, the psychological model of treatment is focused on the psychoeducational approach to health management and problems in living, including the prevention, moderation, and amelioration of psychological stress resulting in physical and psychological disorders. Concurrent with psychologists' advocacy efforts, the knowledge basis of the discipline has expanded dramatically, particularly in the areas of health psychology (Sweet, Rozensky, & Tovian, 1991) and primary care. As the most broadly trained mental health professionals, psychologists also have an emerging role in rural hospitals and emergency rooms.

2.21.10 SUMMARY

Psychologists have entered and considerably expanded their clinical roles and responsibilities in hospitals since World War II. In doing so, they have had to confront established professions such as medicine and nursing in demonstrating their clinical expertise and value to the hospital. In order to gain a firm footing within hospitals, psychologists have had to organize and advocate for legal and regulatory changes at the federal, state, and individual hospital level, while simultaneously developing an appropriate professional identity for inpatient practice. As relative newcomers to the hospital, psychologists need to be aware of the complexity of political and organizational issues in pursuing independent privileges or other clinical roles. They must use this knowledge to extend the hospital's bylaws, credentialing, and privileging procedures appropriately for their discipline. In contrast to outpatient practice, patients treated in a hospital are more acute and more clinically complex, and must be treated in collaboration with other members of the treatment team. These factors raise a host of ethical and legal issues which must be carefully considered in a timely manner.

In the USA, hospitals and practitioners are undergoing radical changes in providing health care services, due to the incursion of managed care directives and to related CQI approaches. These changes constitute a health care revolution reminiscent of a quote regarding an earlier revolution: "it was the best of times, it was the worst of times." It requires considerable ingenuity from psychologists to envisage the direction of developments and how best to position themselves, while maintaining the integrity of the philosophy and values of their profession in their work. Nonetheless, the clinical and research developments of the past 50 years have proved that psychologists are innovators with valuable contributions to make in many different venues. As knowledge of the mind–body interaction develops in the healing professions, the value of the psychological treatment model has become increasingly apparent.

Psychologists' training in behavioral approaches to patient care and their research design skills in developing measurable variables to assess outcomes effectiveness are of great value in the current climate of accountability. However, psychologists must recognize, measure, and demonstrate their value to psychiatric, medical–surgical, chemical dependency, geriatric, rehabilitation, and other facilities to assure their inclusion and their autonomy in evolving health care delivery systems.

2.21.11 REFERENCES

American Psychiatric Association (1986). *Credentials and privileges for psychiatrists in hospital-based services*. (Task Force Report No. 24). Washington, DC: Author.

American Psychological Association (1992). The ethical principles of psychologists and code of conduct. *American Psychologist, 47,* 1597–1611.

Bennett, B. E., Bryant, B.K., VandenBos, G.R., & Greenwood, A. (1990). *Professional liability and risk management*. Washington, DC: American Psychological Association.

Bersoff, D. N. (1983). Hospital privileges and the antitrust laws. *American Psychologist, 38,* 1238–1241.

Broskowski, A. (1994). Current mental health care environments: Why managed care is necessary. In R. L. Lowman & R. J. Resnick (Eds.), *The mental health professional's guide to managed care* (pp. 1–18). Washington, DC: American Psychological Association (pp. 1–18).

CAPP v. Rank (1990). 51 Cal.3d 1, 793 P.2d 2.

Committee on Professional Practice of the Board of Professional Affairs (1985). *A hospital practice primer for psychologists*. Washington, DC: American Psychological Association.

Committee on Professional Practice of the Board of Professional Affairs (1988). *Hospital practice advocacy issues*. Washington, DC: American Psychological Association.

Definition of a clinical neuropsychologist (1989). *The Clinical Neuropsychologist, 3,* 22.

Enright, M. F., Resnick, R. J., Ludwigsen, K. R., & DeLeon, P. H. (1993). Hospital practice: Psychology's call to action. *Professional Psychology: Research and Practice, 24,* 135–141.

Joint Commission for Accreditation of Healthcare Organizations (1992). *Accreditation manual for hospitals 1992*. Chicago: Author.

Linton, J. C (1993). Current ethical issues in hospital settings. *Register Report, 19*(1), 9–10.

Ludwigsen, K. R. (1993). Credentialing and privileging issues in hospital practice. *Register Report, 19*(2), 3–6.

Ludwigsen, K. R., & Albright, D. G. (1994). Training psychologists for hospital practice: A proposal. *Professional Psychology: Research and Practice, 25*(3), 241–246.

Marshall, J. W., & Muszynski, I. L. (1993). Legal briefing: Providers should not fear UR entities or court decisions. *Register Report, 19*(2), 7–8, 11.

Matarazzo, J. D. (1977). Higher education, professional accreditation, and licensure. *American Psychologist, 32,* 856–859.

Pope, K. S. (1990). Ethical and malpractice issues in hospital practice. *American Psychologist, 45,* 1066–1070.

Resnick, R. J., Enright, M. F., & Thompson, R. J. (1991). *Guidelines on hospital practice: Credentialing and bylaws*. Washington, DC: American Psychological Association.

Resnick, R. J., & Morris, J. A., Jr. (1997). The history of rural hospital psychology. In J. A. Morris, Jr. (Ed.), *Practicing psychology in rural settings: Hospital privileges and collaborative care*. Washington, DC: American Psychological Association.

Rozensky, R. H. (1991). Psychologists, politics, and hospitals. In J. J. Sweet, R. H. Rozensky, & S. M. Tovian (Eds.), *Handbook of clinical psychology in medical settings* (pp. 59–79). New York: Plenum.

Rozensky, R. H. (1997). Medical staff membership and participation in rural hospitals. In J. A. Morris, Jr. (Ed.), *Practicing psychology in rural settings: Hospital privileges and collaborative care* (pp. 19–36). Washington, DC: American Psychological Association.

Shueman, S. A., Troy, W. G., & Mayhugh, S. L. (1994). Some questions and answers about managed behavioral health care. *Register Report, 20*(1), 1, 6–9.

Social Security Administration (1993). Sec. 1861(e)(4), 42 U.S.C. 1395x (e)(4). Compilation of the Social Security Laws (Committee Print WMCP:103–5). Washington, DC: US Government Printing Office.

Social Security Acts Amendments (1994). Pub. L. No. 103–432, Sec. 104.

Stromberg, C. D., Haggarty, D. J., Leibenluft, R. F., McMillian, M. H., Mishkin, B., Rubin, B. L., & Trilling, H. R. (1988). *The psychologist's legal handbook*. Washington, DC: The Council for the National Register of Health Service Providers in Psychology.

Sweet, J. J., Rozensky, R. H., & Tovian, S. M. (Eds.) (1991). *Handbook of clinical psychology in medical settings*. New York: Plenum.

Wiggins, J. G., & Ludwigsen, K. R. (1991). Marketing psychological services in hospitals. In J. J. Sweet, R. H. Rozensky, & S. M. Tovian (Eds.), *Handbook of clinical psychology in medical settings* (pp. 127–137). New York: Plenum.

Zaro, J. S., Batchelor, W. F., Ginsberg, M. R., & Pallak, M. S. (1982). Psychology and the JCAH: Reflections on a decade of struggle. *American Psychologist, 37,* 1342–1349.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.22

Individual Psychologists: Professional Registration and Regulatory Procedures in the United Kingdom

JOHN N. HALL
University of Oxford, Warneford Hospital, UK

2.22.1 INTRODUCTION 249
2.22.2 BRITISH PROFESSIONAL REGULATORY PROCEDURES 250
2.22.3 THE PATTERN OF ADMINISTRATIVE AND PROFESSIONAL STRUCTURES IN THE NATIONAL HEALTH SERVICE 251
2.22.4 EMPLOYMENT CONTRACTS, PRACTICE PRIVILEGES, AND INDEMNITY PROCEDURES 251
2.22.5 BRITISH PSYCHOLOGICAL SOCIETY REGISTRATION PROCEDURES 252
2.22.6 RECOGNITION OF FOREIGN QUALIFICATIONS BY THE BRITISH PSYCHOLOGICAL SOCIETY 253
2.22.7 ADDITIONAL REGISTRATION AND TRAINING PROCEDURES 254
2.22.8 REFERENCES 255

2.22.1 INTRODUCTION

The origins of regulation of practice of different professions in the UK go back to the medieval craft guilds, when the guilds of individual cities set standards of practice for individual crafts and trades, and by regulating the numbers of apprentices thereby controlled entry to the different trades. In the City of London particularly, a complex system of guilds grew up which was relatively unchallenged until the early nineteenth century. In England at the turn of the nineteenth century only three professions were recognized: the church (meaning the established Church of England), law, and medicine, the latter two of these both having "superior" branches and "inferior" branches, in medicine illustrated by the privileged Royal College of Physicians, and the lowlier apothecaries who existed to meet the need for some medical expertise in poorer areas. The procedure of practice-privileges is antedated by the procedure in the former English voluntary hospitals in the eighteenth and nineteenth centuries of visiting physicians, where respected doctors gave some of their time to the voluntary institutions, and thereby gained both status and advertisement for their skills, available separately on a private fee-paying basis.

From the middle of the nineteenth century there was a rapid growth in the recognition of independent professions; Larson (1977) indicates that of 13 then contemporary professions, 10 had established national associations between 1825 and 1880. The status and rights of emerging professions were confirmed in Britain either by a Royal Charter to the institution concerned, as with the different engineering institutes, or by a specific Act of Parliament, as for medical practitioners through the Medical Act in 1858. The way in which state registration for general nurses in England was finally achieved in 1919, after a battle between the "good-enough" minimal training approach, and the professional high-standards approach, illustrates the power-politics which lay behind the negotiation of professional practice regulation. These same historical issues and processes of the move from apprenticeship to academically justified training, the establishment of status and titles, attempts to protect from competition, the control of training, and the market consequences of under-supply, widely discussed in the sociology of professionalism, should be acknowledged in any discussion of professional practice regulation.

2.22.2 BRITISH PROFESSIONAL REGULATORY PROCEDURES

The legal basis of the present structure of professional regulatory procedures in the UK still rests mostly on the use of Acts of Parliament and of Royal Charters, which apply to the whole country. The larger professions, and those where there is a substantial degree of potential risk, such as pharmacists, are regulated by Act of Parliament, involving careful scrutiny of the proposed legislation by interested parties, and public debate in Parliament of the issues. An Act of Parliament in 1960, for example, created a Council for "Professions Supplementary to Medicine," which registers, for example, physiotherapists as well as radiographers. This Act stipulates a number of medically qualified members of the Council. Legislation is being considered which would release these professions from such heavy medical involvement.

Other professions may be regulated by Royal Charter, where a professional group wishes to have formal recognition of its public standing, but where scarce Parliamentary time is not needed (or available!) because contentious or major issues of public concern are not thought to arise. Bodies set up in this way have a protected right to be consulted on matters of public concern, and to nominate representatives for other public bodies. Despite the archaic name, the procedure for obtaining a charter involves careful legal scrutiny and consultation with other affected public bodies, but does not involve Parliamentary debate. The British Psychological Society (BPS), equivalent in role to the American Psychological Association, was founded in 1901, but was not "incorporated" by Royal Charter until 1965. The objects of the BPS, as set out in the Charter, include "setting a high standard of professional education and knowledge," and "to maintain ... a register of Chartered Psychologists," so that the standards the BPS sets are the nationally accepted professional standards, and the Register of Chartered Psychologists was set up in 1988.

A third category of regulation is essentially self-regulation, where a body sets itself up as a training body, perhaps also setting standards of practice, and is widely recognized and respected for those standards, which are tacitly accepted. A good example is the British Association for Counselling, the main British professional body for counselors, whose accreditation procedure is widely accepted as the standard required for independent practice as a counselor. Other areas of practice may have a number of training bodies, with widely differing standards, and where no one group has established public confidence. An obvious British example are hypnotists, where the yellow pages of any local telephone directory show about 10 different qualifications quoted, although there are respected bodies regulating the practice of hypnosis of those already qualified in a recognized profession of origin.

Since a UK government report of 1971 into the practice of scientology, which recommended legislation to register psychotherapists, a number of proposals have been made to register psychotherapists, including one draft Parliamentary bill, the 1981 Bright Bill. Given the then difficulty in the various psychotherapy bodies agreeing on common standards, this was not pursued, since the position of the government was that it would not support legislation which would have been contentious in the face of such disagreement. However, in 1993 the United Kingdom Council for Psychotherapy was created, which registers psychotherapists under a number of specialized headings, although this procedure is not supported by either an Act of Parliament or Royal Charter.

A major factor in the practice regulation of all UK healthcare professions is the dominance of the National Health Service (NHS), which is the national system of healthcare provision, established in 1948, which includes primary-care medical services as well as secondary-care services. The NHS is the principal employer of nearly all categories of such professions. The clinical component of training of nearly all healthcare professions takes place in NHS settings, linked in most cases to universities, with funding of training supported at least partially in some way by central government, so there is a degree of integration of training systems with service delivery systems.

The practice assumptions of the NHS, as setter of practice standards, and the career structure it offers, have a powerful influence on all private practice. Private practice tends either to be done on a part-time basis associated with a part-time NHS job, or full-time after obtaining substantial NHS experience. Family doctors, or general practitioners (GPs) who are themselves mostly contracted to work for the NHS, are key gatekeepers to private practice, so being known and respected by local GPs is a valuable contributor to successful private practice, at least initially. For professions other than doctors, the proportion of practitioners working wholly or partly outside the NHS varies; there is high demand for private practice podiatry and physiotherapy, for example, but for clinical psychology probably not more than 5% of the national practitioner resource is available privately.

2.22.3 THE PATTERN OF ADMINISTRATIVE AND PROFESSIONAL STRUCTURES IN THE NATIONAL HEALTH SERVICE

Since its inception in 1948, the NHS has gone through a number of organizational changes. Up to 1974 individual hospitals were run by hospital management committees, but since then by health authorities servicing a named area, or latterly district, covering usually a county or major city, with populations typically about 250 000, emphasizing the importance of focusing services on a served population. The then area and district health authorities continued to be organized according to nationally prescribed standard patterns, and clinical psychology services were typically organized as area services, with one clinical psychology service or department providing services to all client groups, including mental health, child, and learning disability services, as well as neurological and general medical services in larger districts.

From the early 1990s the creation of legally autonomous but still state-funded NHS trusts has meant that services are organized more flexibly. The strategic focus of the NHS is now on a primary-care led service, but most clinical psychologists in Britain work in services administered by secondary-care agencies. There are substantial variations in management arrangements for clinical psychologists: services for a number of trusts may be run together as agencies; there may be one central department for a trust; services may be organized as locality services, serving a population of, say, 60 000 people; or services run as a client group/clinical directorate arrangements.

As part of the older centralized system, conditions of service for the NHS in the UK, with a total population of 56 000 000 nationally, were negotiated centrally via what is known as the Whitley Council. This system involved negotiation between the management side, representing both NHS management and with financial input from the Treasury, and the staff side representing the trade unions and professional associations. The Whitley Council has a number of subcommittees concerned with individual staff groupings, and for this purpose clinical psychologists were grouped together with biochemists and physicists in Whitley Professional and Technical Council A, being seen as scientists. This Whitley PTA Council agreed on national grading structures, which until 1990 offered a standard national career progression.

In 1990, a new more flexible grading system was introduced, but the former grade titles are still often used, with higher grade clinical psychologists often being called consultant, specialist head of psychology services, with one such post usually being appointed or recognized as trust head or professional lead. A senior clinical psychologist is appointed within the national Department of Health, and another senior psychologist is a direct adviser to the Chief Medical Officer of the government. While this Whitley system is somewhat cumbersome, it has created a nationally accepted framework for professional standards, and the appointment and grading of psychologists, with a high degree of comparability across the country as compared to the American position. A report originating from national concern about the supply of clinical psychologists (Manpower Planning Advisory Group, 1990), and an overview of British clinical psychology practice (Marzillier & Hall, 1998) together give a good account of the state of clinical psychology in Britain.

2.22.4 EMPLOYMENT CONTRACTS, PRACTICE PRIVILEGES, AND INDEMNITY PROCEDURES

Within the NHS, the trust itself, as the employer, is responsible for all patient care carried out under its authority. The trust, therefore, has a responsibility to assure itself that any staff who come into patient contact are competent to carry out what is expected of them.

For directly employed staff this requirement will be met by thorough checking of qualifications, of professional registration status (and checking of any required annual registration certificates and numbers), and of experience by interview, and by ongoing professional supervision. On appointment, such staff will receive from their employers information about professional practice standards within their trust, and disciplinary procedures of the employing trust if there is any breach of those standards.

Any other staff who come into patient contact, such as university-employed or research-grant funded staff, trainees from other training courses, or foreign visitors, will be required to have an honorary contract, which is both a recognition of their competence, and also covers their work for insurance purposes by the employer's indemnity. An honorary contract will also give information to the person concerned of standards of practice, and usually must name a senior qualified member of the profession concerned who will supervise clinical work. The professional competence standards for an honorary contract then become similar to those expected if the person were acting in a similar capacity as an employee.

A number of private companies offer healthcare services in the UK, and professional staff directly employed by them will be covered by their employer's insurance, and may also be required to possess their own indemnity insurance. The professional standards adopted will be at least those set by the NHS, so the NHS standards do in effect permeate the whole healthcare system in the UK as a minimum. Private hospitals may also offer practice privileges to practitioners, who thereby may use the facilities of the hospital on payment of appropriate charges. Requests for practice privileges are carefully checked and scrutinized, checking carefully the area of practice experience of the person seeking appointment.

To American eyes, British consumers of healthcare may seem either incredibly trusting or stupid, since the level of health-related litigation is extremely low by US standards. UK common law recognizes the principle of "vicarious liability," whereby an employer is responsible for the acts of employees. Practice in the NHS means that it is the employing trust, and not the employee, who is usually sued for professional negligence, with the exception of medical staff. Accordingly there is no requirement for employed clinical psychologists, who do no private practice, to be individually professionally insured. There is no case known to the author of legal damages being awarded against a British clinical psychologist arising out of malpractice, although clinical psychologists have been dismissed from their posts or struck off the Register for malpractice.

Although indemnity insurance is not mandatory for employed psychologists, many clinical psychologists do take out professional insurance, since if a malpractice case is pursued, the interest of the employing trust is to resolve the dispute as quickly and cheaply as possible, and it may therefore settle out of court without attempting any defence of the individual practitioners who thereby may be seen to be at fault. The purpose of the insurance under these circumstances would be to protect the reputation of the clinician, although they were not the defendant in the actual litigation. Any independent practitioner should have indemnity cover, and it is a requirement of the BPS professional ethical code (BPS, 1995) that private clinical psychologists should be so insured. Indemnity insurance is issued by a number of companies, some specializing in therapists more generally, and covers a range of liabilities. As an example, the BPS has links with an insurance scheme, currently offering £1 000 000 cover for any one claim for a premium of £45. Partnerships of psychologists may also be insured. Legal representation is also available through the scheme. The trade union for psychologists also has a legal advice scheme available, but membership of the union does not provide indemnity cover.

2.22.5 BRITISH PSYCHOLOGICAL SOCIETY REGISTRATION PROCEDURES

As already mentioned the BPS maintains a Register of Chartered Psychologists. Registration requires: (i) possessing a first degree (i.e., a bachelors degree with honours) with psychology as a main subject, and which covers the general areas of psychology required by the BPS, and (ii) since obtaining that degree or equivalent, completing a period of study or practice sufficient for practice without supervision for three years full-time, or an equivalent period part-time. Prior to 1990, most British clinical psychology courses lasted three years, with some being for two years. Since 1990, all British clinical psychology training courses have to be of three years full-time duration to meet Whitley Council standards, all courses now conferring doctorates.

Accordingly, on completion of training a clinical psychologists now fulfills both Whitley

Council criteria for employment as a qualified clinical psychologist in the NHS, and BPS criteria for registration. There is a regular accreditation procedure for training courses, conducted by the BPS, ensuring that all courses meet consistent training criteria, and courses will also meet the somewhat varying academic requirements of the university conferring the doctoral degree, usually a Doctor of Clinical Psychology degree. The degree of Doctor of Philosophy has always been in the UK solely a research degree, and not, as in the USA, sometimes a practitioner qualification.

As a condition of registering, chartered psychologists are also required to hold a practicing certificate, issued annually. Exemption may be granted from holding a practice certificate; this applies to retired psychologists, or those not working, or to teachers or researchers in psychology who undertake not to provide any psychological interventions that are held out as being of benefit to others. A member of the public who believes themselves to be in professional contact with a person who holds themselves out to be a chartered psychologist may therefore expect to find their name on the published register, and may ask to see a current practice certificate. By contrast with American practice, it is very rare for UK health professionals to display their academic and professional certificates and diplomas at their place of work.

The chartering regulations make provision for four other specified categories of chartered psychologist, as well as clinical psychologists. There are Chartered Educational Psychologists, Chartered Occupational Psychologists, Chartered Forensic Psychologists, and Chartered Counselling Psychologists. They may use the term C.Psychol. after their name, with an abbreviation to describe which type of chartered psychologist they are, such as C.Couns. Psychol. Professional training courses in the five different branches of applied psychology recognized by the chartering procedure are entirely separate. However, a lateral transfer scheme has been introduced whereby experienced psychologists in other fields may transfer to be recognized as clinical psychologists. This is achieved on a case-by-case analysis of their training and experience, so an individually defined training program is determined, which may include further academic study, research work, and certainly additional supervised clinical experience. On completing the training requirements, they are issued with a "statement of equivalence" from the Committee for the Scrutiny of Individual Clinical Qualifications of the BPS, which then entitles them to work in the NHS.

The total membership of the BPS in 1996 was 21 200, with another 5800 student subscribers. In 1996 there were 8454 chartered psychologists in total, with 3184 full members of the Division of Clinical Psychology, and 1164 full members of the Division of Counselling Psychology.

The legal position in Britain is accordingly that while anyone can continue to call themselves a psychologist, it is possible to sue for misrepresentation if a person described themselves as a chartered psychologist and was offering psychological services. The BPS has prepared the text of a draft Parliamentary bill, and is seeking a Parliamentary sponsor to progress the bill, the passing of which would further protect the public against those misrepresenting themselves as clinical psychologists.

If anyone in receipt of services from a chartered psychologist wishes to complain of the standard of professional service they have received, then the complaints and disciplinary procedure of the BPS will be followed. This has a two-stage process. Minor breaches of the BPS guidelines will be dealt with by a BPS officer. The complaint may then be first considered by an investigatory committee; if they believe a serious professional misdemeanor may have occurred, an investigatory panel may be appointed, to examine the matter further. If a minor offence has taken place, the member may be rebuked or warned. If the circumstances of the complaint suggest that removal from the register may be appropriate, the investigatory panel make a recommendation back to the investigatory committee, who may then set up a disciplinary committee, with external lay membership in addition to senior "Fellows" of the BPS. The last annual report of the BPS, for 1996–1997, showed an increase of 24% of allegations over the previous year, although this still equates to only 127 allegations received.

2.22.6 RECOGNITION OF FOREIGN QUALIFICATIONS BY THE BRITISH PSYCHOLOGICAL SOCIETY

The differing healthcare systems of countries will demand different skills from clinical psychologists: there is no neat definition of clinical psychology practice which travels well from one culture to another. The distinctive requirement of healthcare systems such as that of the UK, where most clinical psychologists work within a public healthcare system requiring services to be offered to a population, is that clinical psychologists should be able to work generically. Accordingly, clinical training requirements are that trainees must complete during their three years four core placements: in adult mental health, children and young people, older people, and learning disability, with an additional fifth specialty placement chosen by the trainee.

Overseas clinical psychologists who wish to work in the UK will follow a process similar to the lateral transfer process already described. On a case by case basis their training will be examined, and a statement of training needs set out. If an overseas clinical psychologist has not had experience in all the above areas, then they may be required to carry out a supervised clinical placement, usually 65 full working days in each area where needed, before a statement of equivalence can be given. Additional academic or research requirements may also be set. Any overseas clinical psychologist wishing to work in the UK is strongly advised to ascertain their status as early as possible, as they cannot be employed in the NHS without a statement of equivalence.

The Treaty of Rome has meant that all member states of the European Community (EC) must agree on common professional standards to bring about movement of labor within the EC. Movement is becoming more common, so that a considerable number of European trained psychologists, especially from The Netherlands and Germany, are seeking work in Britain. The European Federation of Professional Psychologists Associations (EFPPA) exists to promote collaboration between European psychologists, though not all member states of EFPPA are EC members. Given the historical and personal links between the UK and the commonwealth countries, there is a substantial number of clinical psychologists working in Britain who have qualified in such countries, especially Australia, New Zealand, and South Africa, where educational standards and standards of professional training in clinical psychology are close to the British model, unlike Canada where training conforms to the US standard.

Apart from checking professional qualifications, any foreign national, other than an EC citizen, wanting to work in the UK as a qualified clinical psychologist must obtain a work permit. This is obtained on a case-by-case basis by application to the government Department of Employment. Given the shortage of clinical psychologists in Britain, a special fast-track system exists for clinical psychologists to give them accelerated progress through the work permit process.

2.22.7 ADDITIONAL REGISTRATION AND TRAINING PROCEDURES

In general, there are no national legal requirements for additional registration, beyond the base professional registration. Within mental health practice, for example, the nationally accepted professional standards of the main five mental health professions, that is, clinical psychology, occupational therapy, psychiatric nursing, psychiatry, and mental health social work, are such that a formal qualification and experience in any one of these would be taken as evidence of competence to practice, and accordingly there is little incentive for members of any of these professions to have any supplementary registration. The respected Sainsbury Centre for Mental Health (1997) has published proposals for compulsory reaccreditation or licensing in community mental health for all these groups of staff, on a multidisciplinary, practice-related basis. Given that one of the government health ministers of the Labor government elected in 1997 has a strong background in mental health, this may be taken up.

Although a chartered clinical psychologist has little incentive to acquire further formal certification, many clinical psychologists do acquire further formal training and qualifications, in addition to ongoing professional development through workshops, conferences, and local short courses. These may be either in specific forms of psychological treatment, such as the postbasic Oxford Cognitive Therapy training, approved by the BPS, and validated by the University of Oxford, or in healthcare related management skills, such as an MBA. The BPS has a formal process for recognizing postqualification training, through a standing committee on continuing professional development. There are national associations for family therapy, group therapy, and cognitive-analytic therapy, for example, offering postbasic training to well-accepted standards. For clinical psychologists, many of these further qualifications will lead to a postgraduate diploma or masters level degree. Some clinical psychologists do choose to register with the United Kingdom Council for Psychotherapy, but such registration is voluntary.

Clinical psychologists do not have any specified statutory responsibilities or duties under Acts of Parliament, unlike psychiatrists, who have statutory duties with regard to the compulsory detention of patients in a psychiatric hospital under the Mental Health Act 1983. However, under Section 58 of that act they may be required to act as an independent assessor in cases where a patient, already legally detained under the Act, is deemed to require electroconvulsive therapy, for example, and does not give consent. A small number of clinical psychologists also act as mental health commissioners (alongside commissioners from other professional and legal backgrounds), regularly visiting all psychiatric hospitals in the UK and having a duty to inspect the condition of all patients admitted under the Act. The Education Act 1980 provides for the "statementing" of children with special needs, such as those with profound learning disabilities. Any statement is required to include separate statements from a teacher, a doctor, and an educational psychologist, and locally it may be agreed formally that a child clinical psychologist may prepare a psychological statement acceptable for Act purposes. Clinical psychologists in the UK generally are relieved not to have statutory duties, given the burden of being on on-call rotas and the disruption of regular work patterns they involve.

2.22.8 REFERENCES

British Psychological Society (1995). *Ethical code and code of professional practice*. Leicester, UK. Author.

Larson, M. S. (1977). *The rise of professionalism: a sociological analysis*. Berkeley, CA: University of California Press.

Manpower Planning Advisory Group. (1990). *Manpower Planning Advisory Group Report on Clinical Psychology*. London: HMSO.

Marzillier, J. S., & Hall, J. N. (1998). *What is clinical psychology?* Oxford, UK: Oxford University Press.

Sainsbury Centre for Mental Health (1997). *Review of the roles and training of mental health care staff*. London: Author.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.23
Ethical Principles of Psychologists: United States

CYNTHIA A. STURM
Portland, OR, USA

2.23.1 INTRODUCTION 258

2.23.2 AMERICAN PSYCHOLOGICAL ASSOCIATION ETHICS CODE 258
2.23.2.1 Evolution of the Code 258
2.23.2.2 Ethics Education 258
2.23.2.3 Aspirational Values Underlying the Ethical Principles 259
2.23.2.4 Ethical Principles of Psychologists and Code of Conduct 259
2.23.2.5 APA Practice Standards and Guidelines 260

2.23.3 REGULATION OF THE PRACTICE OF PSYCHOLOGY 260
2.23.3.1 State and National Professional Organizations 260
2.23.3.2 State Licensure 261
2.23.3.3 Informal Peer Monitoring 261

2.23.4 ETHICAL DECISION-MAKING 261
2.23.4.1 Taking a Proactive Approach 261
2.23.4.2 Identification of Ethical Dilemmas 262
2.23.4.3 Applying a Decision-making Model 262

2.23.5 PROFESSIONAL COMPETENCE 263
2.23.5.1 Competence: The Cornerstone of Ethics 263
2.23.5.2 Maintaining and Enhancing Competence 264
2.23.5.3 Professional Impairment 264

2.23.6 PROFESSIONAL RELATIONSHIP 264
2.23.6.1 Client Welfare 264
2.23.6.2 Structuring the Professional Relationship 265
2.23.6.3 Informed Consent 265
2.23.6.4 Interruption and Termination of Services 265

2.23.7 PRIVACY AND CONFIDENTIALITY 266
2.23.7.1 Privacy, Privilege, and Confidentiality 266
2.23.7.2 Limits or Exceptions to Confidentiality 266
2.23.7.3 Third-party Requests for Information 266
2.23.7.4 Couple and Family Therapy 266
2.23.7.5 Documentation and Record Keeping 267

2.23.8 MULTIPLE RELATIONSHIPS 267
2.23.8.1 Types of Dual Relationships 267
2.23.8.2 Professional Boundaries 268
2.23.8.3 Sexual Intimacies in Professional Relationships 268
2.23.8.4 Risk Management and Self-care 269

2.23.9 FORENSIC ACTIVITIES 269
2.23.9.1 Forensic Roles 269
2.23.9.2 Specialty Guidelines for Forensic Psychologists 270

2.23.10 FUTURE DIRECTIONS IN PROFESSIONAL ETHICS 270
2.23.11 SUMMARY 271
2.23.12 REFERENCES 271

2.23.1 INTRODUCTION

Psychologists in the USA bear a broad responsibility to conduct themselves professionally to the highest standard. Ethical behavior represents a personal lifelong commitment, a responsibility to colleagues and the profession of psychology, and a pledge to uphold the public trust. The Ethics Code of the American Psychological Association (APA) has evolved as a living document that embodies the ideals, values, and principles of psychological service. The mandatory standards defined in the *Ethical Principles of Psychologists and Code of Conduct* (APA, 1992) function to identify expectations of professional practice, clarify necessary expertise, and ensure protection of the public consumers of psychological services. The profession of psychology in the USA is autonomous from government oversight and supervision by other professions, due to the capacity for self-monitoring through enforcement of the Ethics Code by licensing boards and ethics committees in national and state professional organizations.

2.23.2 AMERICAN PSYCHOLOGICAL ASSOCIATION ETHICS CODE

2.23.2.1 Evolution of the Code

The growth of applied psychology in the USA after World War II led the American Psychological Association, founded in 1892, to recognize the need for a formal ethics code to define standards of practice, educate members regarding professional values, and provide an avenue to address complaints of unethical conduct. An ethical code represented a key element in the structuring of psychology as an independent profession, along with the establishment of graduate training programs, accreditation procedures, and credentialing and licensing mechanisms.

The development of the initial code *Ethical Standards for Psychologists* (APA, 1953) was based upon more than 1000 critical incidents solicited from APA members, reflecting case situations in which psychologists had made decisions with ethical implications. The rationale for this innovative empirical methodology was to create a code with enough relevance and specificity to provide practical guidance to psychologists (Hobbs, 1948). To date there have been six major revisions of the code. The APA Ethics Committee can initiate a revision process based upon the need for changes or additions to the code.

An Ethics Code Task Force was appointed in November 1996 to review the 1992 code and begin the intensive process of another major revision. Elements of the review will include a critical incident survey, a call for comment on the 1992 code, reviews of relevant articles and guidelines, and a legal and case review (APA Ethics Committee, 1995). Draft revisions are published nationally in *The APA Monitor* for review and comment by the APA membership. The last revision process, which produced the 1992 code, spanned six years from initiation to final publication (Canter, Bennett, Jones, & Nagy, 1994). Thus, the Ethics Code, as a living document, continues to evolve to reflect the most current social issues and emerging areas of concern in the psychologist's professional experience.

2.23.2.2 Ethics Education

Graduate training programs are a primary source of education in professional values. Students learn about ethics through coursework, and then apply these principles in supervised clinical practica and internship experience. The importance of formal ethics training has been supported by Handelsman (1986). He asserts that learning ethics by "osmosis," solely in the context of supervised clinical experience, fails to ensure that general ethical theory and reasoning skills can be applied effectively in future clinical contexts.

Psychologists must renew their knowledge of ethical standards of behavior throughout their careers, as professional values evolve over time. For professionals in the workforce, this usually occurs through attending workshops and reviewing relevant APA publications. Journal articles which summarize the educational and adjudicative activities of the APA Ethics Committee are published annually in the *American Psychologist*. The status of evolving standards and guidelines are discussed in *The APA Monitor*, a monthly newspaper. Illustrative casebooks are usually published by the APA after a major revision of the Ethics Code. A comprehensive commentary on the APA Ethics Code by Canter et al. (1994) provides a thorough interpretative analysis of the 1992 code. Current resources on professional, ethical, and policy issues can also be obtained electronically through PsychNet (SM), APA's home page on the World Wide Web.

2.23.2.3 Aspirational Values Underlying the Ethical Principles

Aspirational values are the overarching moral principles that guide ethical decision-making. As ideals, they form the ceiling of expectations regarding professional behavior and offer general guidance across a broad range of situations. Kitchener (1984) summarized five of the fundamental ethical values reflected in human services work. The principle of autonomy emphasizes the promotion of self-determination, freedom of choice, and responsibility for one's own behavior. Nonmaleficence highlights the ethical imperative to avoid or minimize the risk of harm to clients. Beneficence emphasizes the importance of promoting the welfare of others and, more broadly, to contribute to society. The principle of justice encompasses equity, fairness, and nondiscrimination. Finally, fidelity encourages honesty, reliability, and faithfulness to commitments.

Previous versions of the APA Ethics Code mixed aspirational principles with mandatory standards for professional ethical behavior. The structure of the 1992 code changed markedly. The aspirational, nonenforceable values are now in the preamble and general principles introducing the code, and the 102 enforceable Ethical Standards comprise the primary text (APA, 1992). The general principles are summarized below:

(i) *Competence.* This reflects the professional standard of good practice. Psychologists practice only in areas for which they possess adequate education, training, or experience. They recognize that special populations may require specific competencies and take care to protect the welfare of those receiving services when practicing in new areas lacking written standards.

(ii) *Integrity.* Integrity speaks to honesty and fairness in psychologists' professional activities. Statements informing the public regarding credentials, competencies, policies, or services must be accurate and straightforward. Honest self-examination is likewise critical to avoid potential harm to clients from personal limitations or biases which compromise professional effectiveness or objectivity.

(iii) *Professional and scientific responsibility.* This consideration highlights ethical conduct in the psychologist's interactions with other professionals. They must remain watchful of the ethical behavior of colleagues and utilize consultation proactively for themselves.

(iv) *Respect for people's rights and dignity.* These qualities are promoted through advocacy for human rights, privacy, and self-determination. Psychologists maintain awareness of human differences and avoid discrimination.

(v) *Concern for others' welfare.* Psychologists clarify any role conflicts that carry the potential for compromising the best interests of the client. They are aware of the real and perceived power differential in professional relationships and do not take advantage of this.

(vi) *Social responsibility.* In a larger context, psychologists carry a broad responsibility to contribute to the welfare of society and relieve human suffering.

2.23.2.4 Ethical Principles of Psychologists and Code of Conduct

It has been a general criticism of ethics codes that they are too abstract to guide decision-making and behavior in specific situations. Difficulties with enforcement of the 1981 code due to vague language (White v. North Carolina State Board of Examiners of Practicing Psychologists, 1990) led to the need for a major revision with more explicit language and legally enforceable standards (Bersoff, 1994). The 102 ethical standards in the Code of Conduct define unitary ethical concepts in behaviorally specific terms. These are mandatory minimal standards of professional conduct, the floor of ethical behavior. The standards are organized into eight topics: (i) general standards; (ii) evaluation, assessment, or intervention; (iii) advertising and other public statements; (iv) therapy; (v) privacy and confidentiality; (vi) teaching, training supervision, research, and publishing; (vii) forensic activities; and (viii) resolving ethical issues.

In addition to the new structure of the code, important changes were made in several content areas. The concept of "client" has been broadened to describe the professional duty to students, supervisees, employees, and research participants. More attention is given to the reality of overlapping professional roles and multiple relationships. Decision rules are provided to assist psychologists in managing complex professional relationships in order to maintain objectivity, protect the treatment process, and minimize the risk of harm or exploitation to the client. The code requires sensitivity to issues of cultural diversity, human differences, and unfair discrimination. Due to the serious problem of sexual misconduct in the profession (Pope, 1994), stronger language is evident in prohibitions against sexual intimacies with clients, former clients, supervisees, and students. New sections were included on forensic issues and resolution of ethical dilemmas.

More specific guidance in the structuring of the treatment contract and informed consent is provided. Several standards cover the business aspects of psychological practice, such as fees and financial arrangements, barter, and referral practices. Changes regarding advertising and public statements are a response to concerns from the Federal Trade Commission regarding restraint of trade resulting from previous ethical prohibitions about advertising (Koocher, 1994).

Commentaries on the 1992 code highlight its strengths as: (i) attention to cultural diversity, special populations, and individual differences; (ii) pragmatic guidance in structuring individual therapy; (iii) more explicit guidelines regarding multiple relationships; (iv) inclusion of forensic activities; and (v) greater breadth as well as specificity within the standards (Bersoff, 1994). Payton (1994) has criticized the current code for its increased emphasis on the scientific basis of our professional identity and, conversely, the diminished focus on fundamental human rights in the aspirational language of the general principles. Other weaknesses have been noted in the frequent use of qualified or legalistic language, ambiguity, and remaining problems with enforceability (Bersoff, 1994). Concerns have been expressed that steps taken to operationalize the code leave the impression of being protectionistic of the profession (Payton, 1994; Vasquez, 1994). As a result, psychology has lost an inspirational code which strives to promote social responsibility and "stir one to the highest plane of ethical functioning" (Koocher, 1994, p. 361).

2.23.2.5 APA Practice Standards and Guidelines

Various supplemental guidelines and standards of practice have been adopted by the APA or its divisions to further educate psychologists, professional bodies, and courts regarding professional standards. Although not enforceable or mandatory, these policies are utilized by ethics committees and practitioners for important technical guidance. The *General Guidelines for Providers of Psychological Services* (APA, 1987), and *Specialty Guidelines for the Delivery of Services by Clinical Psychologists, Counseling Psychologists, Industrial/Organizational Psychologists and School Psychologists* (APA, 1981), currently under revision, were developed to "define minimum levels of competent professional practices and procedures" (Keith-Spiegel & Koocher, 1985, p. 224). Other clinically relevant standards include *Record Keeping Guidelines* (APA, 1993c) and *Guidelines for Providers of Psychological Services to Ethnic, Linguistic, and Culturally Diverse Populations* (APA, 1993b).

APA guidelines pertaining to assessment include the *Standards for Educational and Psychological Testing* (APA, 1985), which are currently under revision. Complex or specialized areas of practice, such as forensic psychology, often require more specific clarifications by experts in the field. The *Specialty Guidelines for Forensic Psychologists* (Committee on Ethical Guidelines for Forensic Psychologists, 1991), were developed and adopted by the APA Forensic Psychology Division and the American Psychology-Law Society, but have not been adopted by the APA as a whole. The APA has recently adopted the *Guidelines for Child Custody Evaluations and Divorce Proceedings* (APA, 1994), an area of specialization in which ethical complaints are frequently received.

Important standards regarding research include the *Guidelines for Ethical Conduct in the Care and Use of Animals* (APA, 1993a), and the *Ethical Principles in the Conduct of Research with Human Participants* (APA, 1982), now being revised.

The APA Ethics Committee also provides ongoing ethical guidance on specific topics through various policy statements, such as Military Psychologists and Confidentiality (APA, Ethics Committee, 1994, p. 665), "Take home" tests (APA, Ethics Committee, 1994, p. 665), and Services by Telephone, Teleconferencing, and Internet (APA, Ethics Committee, 1996, p. 1285).

2.23.3 REGULATION OF THE PRACTICE OF PSYCHOLOGY

2.23.3.1 State and National Professional Organizations

The APA and the independent state and local psychological associations promote the profession of psychology through advocacy, education, and regulation of members by peer monitoring. Voluntary membership in these professional organizations carries with it the commitment to adhere to the ethical code of the organization. The APA code is typically adopted by the state psychological associations. Other professional societies, such as the American Society for Clinical Hypnosis, Feminist Therapy Institute, American Association for Marriage and Family Therapy, and Association for the Advancement of Behavior Therapy have also developed their own ethics codes (Bass et al., 1996).

National and state ethics committees serve to educate members and respond to and investigate complaints from consumers or other members regarding unprofessional conduct. Ethics committees' rules and procedures define the process of adjudication of complaints against members and the types of sanctions. For less serious violations, which are not likely to have caused harm, committees may respond educatively or with a letter of reprimand. Remedial coursework or clinical supervision may be required. If unethical conduct involves a serious risk of harm to others, more severe sanctions including censure, stipulated resignation of association membership, or expulsion may be used. Unfortunately, due to increasing concerns about liability, many state association ethics committees have suspended their investigatory and adjudicatory functions and now provide primarily educational and consultative services (Nagy, 1996).

State psychological association Professional Standards Review Committees (PSRCs) offer another type of consumer protection. Often focused more specifically upon the review of professional practices and billing, these committees respond to questions raised by providers, consumers, or third-party payers. PSRCs typically function in an advisory role and direct their efforts towards arbitration of complaints (Keith-Spiegel & Koocher, 1985).

2.23.3.2 State Licensure

State licensing boards serve a regulatory function by defining the entry requirements for psychologists who provide services to the public autonomously and use the title of "psychologist." State boards also have legal jurisdiction over those already licensed to monitor and enforce standards of professional conduct. Boards usually adopt either the APA *Ethical Principles of Psychologists and Code of Conduct* (APA, 1992) or the *ASPPB Code of Conduct* (Association of State and Provincial Psychology Boards [ASPPB], 1991). The ASPPB is an organization that assists regulatory boards in the USA and Canada in developing standardized regulatory procedures.

The disciplinary power of licensing boards varies significantly from that of the state association. Only the licensing board can revoke the license to practice, thereby protecting the public from harmful practitioners. State psychological association ethics committees can refer an ethics complaint to the licensing board when a significant risk of harm to others exists through incompetence, sexual misconduct, or professional impairment. It is important to note that many psychologists, such as educators and researchers, have no requirement to be licensed and are therefore exempt from this type of regulation.

2.23.3.3 Informal Peer Monitoring

The 1992 Ethics Code includes a new section on resolving ethical issues, which delineates psychologists' responsibilities when confronted with possible ethical violations by colleagues. If there is uncertainty about a complex ethical issue, consultation with colleagues who have expertise in ethics, as well as state or local psychological association ethics committees may be quite useful (Canter et al., 1994). Informal attempts to resolve potential violations by colleagues are encouraged if the situation is relatively minor and amenable to an educatory intervention. However, in cases of serious misconduct with a risk of harm to others, referral to state or national ethics committees or licensing boards may be a more appropriate action, assuming confidentiality issues can be resolved.

2.23.4 ETHICAL DECISION-MAKING

2.23.4.1 Taking a Proactive Approach

In their classic text, *Ethics in Psychology: Professional Standards and Cases*, Keith-Spiegel and Koocher (1985) advocate a proactive stance of preventing ethical violations through education, sensitization, and self-awareness. Awareness of ethical issues pertinent to each psychologist's scope of professional practice promotes effective self-monitoring. In a survey of 679 APA members, Pope and Vetter (1992) solicited examples of "critical incidents" representing ethical dilemmas that psychologists commonly encounter in their work. The 10 most frequently cited areas of ethical concern were: (i) confidentiality; (ii) blurred, dual, or conflictual relationships; (iii) payment sources, plans, settings, and methods; (iv) academic settings, teaching dilemmas, and concerns about training; (v) forensic psychology; (vi) research; (vii) conduct of colleagues; (viii) sexual issues; (ix) assessment; and (x) questionable or harmful interventions.

The 1992 Ethics Code incorporates specific standards of conduct aimed at preventing or minimizing harm to consumers (Canter et al., 1994). A good practitioner should carefully structure service agreements at the start of the professional relationship, anticipate dilemmas associated with high risk clients or situations, and maintain personal awareness of conflicts or challenges. Experienced clinicians can often quickly identify warning signs at the outset of treatment which may indicate a hidden agenda, or point to a potentially difficult personality style. For example, a treatment referral for a client involved in litigation may be an indirect request for evaluation, under the guise of treatment. Early difficulties in establishing a therapeutic relationship, such as missed appointments, miscommunications, or challenges to usual office procedures may be indicative of a high risk client. Strong emotional reactions to clients should likewise be observed and monitored by clinicians.

The importance of utilizing a formal decision-making model to analyze ethical dilemmas has been emphasized in the literature (Canter et al., 1994; Hill, Glaser, & Harden, 1995; Keith-Spiegel & Koocher, 1985; Kitchener, 1984; Pettifor, 1996). The Canadian Psychological Association (CPA) has incorporated a decision-making model into their ethical code (CPA, 1991). In practice, psychologists must often make ethical decisions during the process of clinical intervention, with minimal opportunity for reflection about alternatives and weighing of potential short-term and long-range consequences. Use of a decision-making framework provides the practitioner a consistent structure to analyze ethical dilemmas, ensure critical factors are not overlooked, and weigh how decisions might affect the rights of various involved parties.

2.23.4.2 Identification of Ethical Dilemmas

Ethically competent practice requires a knowledge of: (i) the current APA Ethics Code and APA practice guidelines and standards; (ii) applicable local, state, and federal laws, including the licensing law; and (iii) institutional or agency rules, regulations, and policies, or contractual obligations (Canter et al., 1994). Ethical awareness involves recognition of an issue as having ethical dimensions or ramifications. Lack of sensitivity may be due to limited clinical experience, practice outside of usual areas of competency, or personal issues which compromise the psychologist's objectivity. Several recent publications have incorporated self-evaluation questions which assist clinicians in reviewing and sensitizing themselves to ethical conflicts in specific areas of practice (Bennett, Bryant, VandenBos, & Greenwood, 1990; Canter et al., 1994; Peterson, 1996).

Analysis of problematic ethical case scenarios must begin with identification of the central issues. A training format developed by the author clarifies five basic areas to consider:

(i) *Clinical context*. This highlights the unique aspects of each case, and may include the initial clinical presentation of the client, diagnostic features, and/or dynamic formulation. In cases of third-party referrals or multiperson therapies, clarification of "who is the client?" may be important. Aspects of human diversity or cultural context should be considered.

(ii) *Ethical issues*. These may include issues of confidentiality, competence to diagnose and treat, sharing of test results, informed consent, role conflicts, boundary issues, or any other issue defined by the Ethics Code.

(iii) *Legal issues*. These must be differentiated from ethical issues. Common legal issues involve privilege, mandatory child and elder abuse reporting, involuntary commitment, situations involving threats of violence to self or others, and billing practices. When the psychologist's ethical duties conflict with the law, steps must be taken to resolve conflicts in a responsible manner while maintaining the highest commitment to the ethics code (Canter et al., 1994). For example, the legal system may subpoena records which include raw test data, placing the psychologist in an ethical–legal dilemma regarding maintaining appropriate confidentiality of test data or protocols (Committee on Legal Issues, APA, 1996). Contractual obligations, for example, with managed care organizations, also represent legal agreements with which psychologists must weigh any competing ethical demands.

(iv) *Moral issues*. These fall beyond the realm of ethical codes and laws. Social issues such as domestic violence, divorce, abortion, gay rights, or physician-assisted suicide are heavily value laden. It is imperative that psychologists avoid any misuse of power or influence and respect the client's rights to self-determination.

(v) *Pragmatic issues*. These encompass the practical aspects of ethical problem-solving. Examples might include a working knowledge of how the legal system acts upon reports of child abuse in a particular community, or the situation in which culturally competent referral resources for a minority client are not readily available.

2.23.4.3 Applying a Decision-making Model

Several authors have outlined rational-evaluative models of ethical decision-making (Canter et al., 1994; CPA, 1991; Kitchener, 1984). Hill, Glaser, and Harden (1995) have added a feeling-intuitive dimension to their feminist model for ethical decision-making, taking into account the emotional reactions of the clinician and the impact upon the therapeutic relationship. These models follow four general steps:

(i) *Generating alternative solutions.* To encourage consideration of multiple solutions, develop at least three potential courses of action.

(ii) *Analysis of each course of action.* For each solution, identify the short-and long-term effects upon the primary client and other affected parties. How does the solution affect the clinician and the treatment relationship? Rank the potential choices.

(iii) *Choice and implementation.* Despite the likelihood that no single option is ideal, a choice of action must be made to implement an ethical path. Avoidance or inaction may be unethical.

(iv) *Evaluation.* A review of the decision process and its consequences is critical. It may be necessary to reengage in the decision-making process if there is still an uncertain or negative outcome.

2.23.5 PROFESSIONAL COMPETENCE

2.23.5.1 Competence: The Cornerstone of Ethics

Competent practice that promotes human welfare and protects clients from harm is the foundation of professional ethics. The boundaries of professional competence are based upon "education, training, supervised experience, or appropriate professional experience" (1992, p. 1600). Professional psychology programs provide an integration of basic science courses with applied training and theory (Tipton, 1996). In the USA, the doctoral degree is considered the most appropriate level of training for entry into independent practice. The APA currently recognizes five areas of professional specialization: clinical, counseling, school, industrial/organizational, and neuropsychology.

The importance of basing judgments upon scientifically and professionally derived knowledge and using empirically founded treatment techniques and assessment measures is highlighted in the Preamble as well as in the Ethical Standards (APA, 1992). Psychologists carry an ongoing responsibility to maintain expertise through keeping current with the professional and scientific literature in their field of activity. The challenge is that "... competence is not a static concept. Rather, it must be evaluated against a changing context of existing knowledge in the field" (Haas & Malouf, 1989, p. 19). Ideally, psychologists make a lifelong commitment to learning and personal growth.

Theoretically, the public is protected from unqualified, incompetent, or unethical practitioners through the state licensure process (Retfalvi & Simon, 1996). Regulatory boards review entry-level academic credentials, supervised internship, and postdoctoral experience. Written and oral examinations cover general psychological knowledge, competence in the candidate's area of specialization, and ethical reasoning. However, enforcement of competence through licensing boards guarantees only minimal levels of knowledge, education, and experience. Cases of gross incompetence may be detected through complaint procedures, but subtle incompetence often goes undetected (Haas & Malouf, 1989). Farley, F. (1996).

Although there is agreement that competence is a fundamental expectation, defining competence has proven difficult (Claiborn, 1982). Clear definitions of professional competence or knowledge necessary in areas of specialization are often lacking. Ultimately, the boundaries of competence are self-monitored. The psychologist determines their scope of practice according to their general skills, areas of specialization, and populations commonly treated. Some delivery settings may require specific expertise, such as hospital, rural, forensic, or industrial/organizational practice. Psychologists should be familiar with laws and regulations applicable to their employment settings. Specific clinical competencies and ethical awareness may be required for multiperson therapies such as marital, family, or group interventions (Lakin, 1994). Knowledge of personal strengths and weaknesses is essential in order to monitor one's competence.

Haas and Malouf (1989, p. 16) describe the competent clinician as having the necessary *knowledge*, effective *skills* in applying knowledge, and appropriate *judgment*, or knowing "what to do, how to do it, and when to do it." Overholser and Fine (1990) discuss subtle aspects in defining the boundaries of competence. Inadequate knowledge or misapplication of skills to particular disorders or populations can occur. Relationship skills are essential to an effective treatment relationship. Inability to form a rapport or, conversely, overinvolvement with clients, are examples of subtle forms of incompetence. Personality problems of the therapist may compromise objectivity in evaluation and intervention with clients. Transient stressors such as illness, divorce, or death of a family member, may render the therapist temporarily incompetent.

Issues of human differences are relevant to psychologists in their roles as practitioners, evaluators, teachers, supervisors, researchers, and employers. They must develop the necessary competence to address any differences based upon "age, gender, race, ethnicity, national origin, religion, sexual orientation, disability, language, or socioeconomic status"

(APA, 1992, p. 1601). Cultural competence in testing and assessment practice involves cultural sensitivity and knowledge, a systematic approach to understanding the client as a cultural entity, and methods of service delivery which are culture specific (Dana, 1994). As the practitioner encounters special populations or unique clinical issues for which they lack adequate training, appropriate use of consultation or referral to a more qualified psychologist is warranted.

2.23.5.2 Maintaining and Enhancing Competence

Continuing education, through attendance of formal workshops, clinical and research conferences, or study groups, is critical to update knowledge and skills. Many state licensing boards have adopted mandatory standards for continuing education hours to maintain licensure. In some cases, education hours must cover specific content areas, such as ethics, AIDS education, or child abuse.

Preparation for new areas of practice requires an appropriately thorough study of the literature, completion of relevant academic coursework, training, and supervised experience or ongoing consultation with an expert in the area of specialization.

The APA developed a policy in 1976 for doctoral-level psychologists wishing to change their specialty which outlines formal steps for preparation (APA, 1976). Psychologists must meet the necessary requirements of doctoral training in the new specialty with credit given for requirements previously satisfied. Respecialization in areas involving services in clinical, counseling, or school psychology should include appropriate practicum and internship experience through APA-approved training programs.

2.23.5.3 Professional Impairment

An important aspect of professional competence is managing "personal problems and conflicts" (APA, 1992, p. 1601) which can diminish effectiveness and carry a risk of harm to clients, students, or colleagues. Psychologists must take appropriate steps to prevent serious impairment. Appropriate responses to early signs of personal difficulty include efforts to gain additional support, limiting services, seeking additional consultation or pursuing personal therapy. Ongoing consultation or participation in peer supervision may also benefit the professional by providing mutual support and counteracting isolation.

Professional impairment may be evidenced by absenteeism, diminished work quality, heightened conflicts with co-workers or inappropriate professional boundaries, including sexual exploitation of clients. Impairment is frequently associated with alcohol or drug abuse. Denial is a hallmark of many psychological problems, particularly substance abuse, often precluding self-awareness and intervention as required by the Ethics Code. Impaired judgment can also be associated with burnout, a pattern of emotional exhaustion resulting from chronic demands on physical or emotional resources in the work setting (Freudenberger, 1975).

The Ethics Code encourages psychologists to attempt informal resolution of ethical violations, including addressing potential impairment of colleagues. However, psychologists are often reluctant to confront distressed colleagues. Several state psychological associations have responded to concerns about professional impairment by developing psychologist support committees. The function of these programs is to protect the interests of persons served by psychologists and to assist psychologists in preventing or managing distress so that it does not affect their professional duties.

2.23.6 PROFESSIONAL RELATIONSHIP

2.23.6.1 Client Welfare

The APA Code (1992, p. 1599) "has as its primary goal the welfare and protection of the individuals and groups with whom psychologists work." This fiduciary role requires psychologists always to consider what is in the best interests of the client in making treatment decisions, and maintain safe, beneficial, and objective professional roles. Psychologists must avoid exploitation of those over whom they have supervisory or evaluative authority, such as students, supervisees, research participants, and employees.

The 1992 code emphasizes the importance of sensitivity to issues of human diversity, and "appropriate respect to the fundamental rights, dignity and worth of all people" (APA, 1992, p. 1599). Psychologists must avoid any unfair or discriminatory practices based upon "age, gender, race, ethnicity, national origin, religion, sexual orientation, disability, language, and socioeconomic status" (APA, 1992, p. 1601).

Broader application of ethical duties requires psychologists to be aware of their professional and scientific responsibilities to the greater community and the society as a whole. The importance of societal goals may outweigh individual interests in some clinical situations. Exceptions to confidentiality based upon the duty to protect others in the community is one example. The reality of limited health care finances in the USA raises the dilemma of whether to promote distribution of mental health services according to the needs of the population as a whole, or to advocate for individual needs.

2.23.6.2 Structuring the Professional Relationship

Ethical standards require that psychologists take care at the outset of the therapeutic relationship to discuss with clients significant information about their services. Optimally, the client is actively involved in a mutual process of developing the therapeutic contract.

Treatment providers are encouraged to clarify information in oral or written form that is understandable to the client, and respond to any questions or concerns (Koocher, 1994). Use of standard written office policies has been encouraged for purposes of documentation and effective communication. Such forms usually include basic information such as the professional's qualifications or background, goals and anticipated course of treatment, alternatives to treatment, risks and benefits, confidentiality, and cancellation or missed appointment policies.

Financial arrangements are an essential part of the initial treatment contract. Attention must be paid to matching the client's needs with available resources and identifying any dilemmas at the outset, such as where the treatment needs may exceed the insurance benefits or client's ability to pay (Koocher, 1994).

In psychological assessment situations, psychologists must clearly explain their role as an evaluator. To ensure an objective and unbiased evaluative opinion, the professional must avoid any conflicting or multiple roles. In general, evaluation and treatment roles are mutually exclusive. The code requires that an appropriate explanation of assessment results is provided to the client in understandable language. Any limitations to receiving the results must be identified at the outset.

Services are frequently requested by third parties, such as schools or courts. This may raise ethical dilemmas around potential conflicts of interest between the needs of the payer and the client. It is critical to inform all concerned parties about the nature of the services, responsibilities to each party, and whether evaluative information will be shared or kept confidential.

The increasing involvement of insurers and managed care companies as third parties in monitoring the planning and delivery of mental health services has created further ethical complexities in structuring treatment. Clients must be informed at the outset regarding potential restrictions in reimbursement related to the insurer's definition of "medical necessity" of treatment, requirements to reveal diagnostic and other treatment information to their insurer, and any financial incentives that encourage the practitioner to limit services. From the insurer's perspective, the Ethics Code requires that psychologists provide accurate information regarding the nature of services and charges in reports to third-party payers (APA, 1992). Explanation of the limits of confidentiality is key to the practice of psychotherapy, where personal information is shared with the therapist, as well as for evaluation, where results may be shared with other parties.

2.23.6.3 Informed Consent

The 1992 code emphasizes the importance of obtaining informed consent from clients at the beginning of service. Informed consent requires three elements: (i) the capacity to consent, (ii) adequate knowledge or information regarding the decision, and (iii) voluntariness, without "undue influence" of others. In-person discussion and written documentation of the informed consent process is necessary. Some consider this to be an ongoing process during treatment.

In cases where the client is not capable of giving informed consent, such as with young minors, developmentally disabled, or other severely impaired individuals, the legally authorized guardian should give permission. Populations who are mandated to treatment or evaluation should be carefully evaluated to ensure that their voluntariness has not been compromised.

2.23.6.4 Interruption and Termination of Services

The Ethics Code also delineates the professional's responsibility to provide continuity of care. Anticipation and preparation for potential interruptions to service must be considered by the psychologist to avoid abandonment of the client (Canter et al., 1994). In the event of relocation or discontinuation of practice, illness, or death of a psychologist, steps should be taken to provide for transfer of client records and referral to an appropriate provider. Changes in the client's location or financial status may also interrupt treatment (Vasquez, 1994). For example, financially limited clients who exhaust their insurance coverage, yet remain in need of services must be appropriately referred or continued by the therapist until stabilization or conclusion of treatment. New issues occur in the context of managed care and group practices, where practitioners may leave a practice or panel, and are no longer eligible to provide reimbursable services to a particular client. Whatever the reason, termination of services must always be handled with care and sensitivity.

2.23.7 PRIVACY AND CONFIDENTIALITY

2.23.7.1 Privacy, Privilege, and Confidentiality

The concept of privacy is a personal right based in constitutional law which allows an individual to determine to what extent personal information, thoughts, or feelings will be shared with others (Keith-Spiegel & Koocher, 1985). Confidentiality is an ethical commitment by the psychologist to maintain the privacy of information disclosed in treatment or assessment. It is important to distinguish between the ethical principle of confidentiality and the legal concept of privilege. Privilege refers to the legal right of the client to prevent information shared in certain protected relationships, such as psychotherapy, from being disclosed in court.

2.23.7.2 Limits or Exceptions to Confidentiality

An integral part of structuring professional services is the discussion of the limits of confidentiality and the potential uses of the information gathered in treatment or assessment. Conflicts between ethical and legal standards may be most apparent in this area. Common exceptions to confidentiality include dangerousness to self or others (e.g., suicidality or homicidality), and mandated reporting of suspected child or elder abuse. Required reports to employers, courts, or other third parties may also affect confidentiality. In situations where clients file a malpractice suit against the psychologist, or use their mental status as a legal defense, such actions may serve to waive privilege.

2.23.7.3 Third-party Requests for Information

Voluntary releases to disclose information to others may be given by the client for various purposes, such as communication with other providers, family members, or eligibility for other types of services and benefits. To ensure a valid consent, it is essential to inform the client about the contents of the records to be released and the potential consequences of releasing that information. For example, a therapy client may become involved in divorce litigation and request that the therapist provide case records to support a claim for custody, without realizing that the record as a whole may contain clinical information unfavorable to their legal case. The client's written consent to release information should be documented in the records.

The Ethics Code encourages psychologists to minimize intrusions on privacy by limiting permissible disclosures to information relevant to the purpose of the request. Voluntary disclosure of information to support insurance claims is a common example where the client must agree to release their diagnosis to the insurer in order to get reimbursed for treatment costs. Managed mental health care plans often require extensive information in the process of authorizing treatment, such as the full *Diagnostic and statistical manual of mental disorders, 4th edition* (*DSM-IV*) diagnosis, treatment plans, progress reports, and termination summaries (American Psychiatric Association, 1994). In some reimbursement systems, copies of chart notes are required. Involvement of the client in examining the risks and benefits of disclosure is essential to ensure their rights to make the best choices for themselves.

2.23.7.4 Couple and Family Therapy

A positive change in the 1992 code was the recognition of the ethical complexities involved in multiperson therapies. Group and family therapies differ from dyadic treatment models in important ways, requiring the psychologist to have competence in the technical skills of these modalities and anticipate ethical dilemmas unique to them (Lakin, 1994). In group therapy, the therapist asks group members to respect one another's confidentiality, but pragmatically cannot guarantee that information discussed in session will remain private.

In treatment situations where the couple or family system is the focus, the Ethics Code requires the psychologist to clarify (i) which individuals are being treated as clients, (ii) the role the psychologist will have with each party, and (iii) the ways in which information will be shared for the purposes of treatment or kept confidential from some of the parties (APA, 1992). In family therapy, where members of the system may also be seen in couple or individual sessions, disclosures may not always be held as confidential (Lakin, 1994). A frequent ethical dilemma occurs when parents divorce and later seek access to family treatment records for legal purposes. In family therapy situations where children are not able to give direct consent, the psychologist should provide information at a level understandable to the child, seek their assent and maintain an active consideration of their best interests. Treatment with step-families raises challenging issues regarding the involvement of noncustodial parents in the consent to treatment and their rights to information regarding their children.

2.23.7.5 Documentation and Record Keeping

The primary purpose of the treatment record is to benefit the client. APA Record Keeping Guidelines (APA, 1993c) provide specific guidance regarding the appropriate content, control, maintenance, and retention of records. Federal, state, and local laws may also govern the disposition of records. Thorough documentation should include the client's presenting issues, history, diagnostic formulation, and treatment plan. This clinical information facilitates communication and coordination with other professionals, and continuity of care should the psychologist become unavailable. Psychologists must guard against any misuse of records. For these reasons, sensitive information, speculations, or personal opinions are rarely appropriate in the record (Soisson, VandeCreek, & Knapp, 1987).

Details of informed consent, releases of information, and financial agreements must also be documented. Accurate record keeping is part of the professional's risk management, providing evidence of appropriate treatment and documenting significant decisions, such as steps taken to manage suicide risk (Soisson et al., 1987). A complete and current record permits the psychologist to monitor their own services and participate in external quality assurance reviews.

Recent technological advances in electronic record keeping and computerized databases have greatly impacted issues of privacy of records. Coding to disguise the identity of clients' mental health records is essential when others can access the database without client consent. Protection of privacy is also important when client information is transmitted via facsimile, cellular phones, or electronic mail.

2.23.8 MULTIPLE RELATIONSHIPS

2.23.8.1 Types of Dual Relationships

Multiple or dual relationships occur when a psychologist functions in a professional role concurrently or consecutively with another professional or personal role (Pope, 1991). The Ethics Code acknowledges the inherent power differential in clinical, supervisory, or teaching roles where psychologists have clinical knowledge or evaluative power. Examples of personal roles outside of a service setting include friendships, social involvements, or business relationships. The psychologist should refrain from promising or taking on new roles when any prior relationship creates a risk of harm (APA, 1992). Overlapping relationships can be categorized as sexual, nonsexual, or conflicts of interest (Table 1).

The 1992 Ethics Code is more explicit and restrictive about dual relationships than previous codes. Sexual exploitation in professional relationships is potentially very harmful to the client as well as damaging to the credibility of the profession. Sexual intimacies with current clients, students, or supervisees are expressly prohibited by the code. In the case of former psychotherapy clients, the ban has been extended to at least two years after termination of the professional relationship. Accepting former sexual partners as therapy clients is now also unethical. Psychologists must avoid any type of sexually harassing behavior.

Nonsexual dual relationships are problematic to the extent that the secondary role has the potential to interfere with the fulfillment of the professional's responsibility in the primary role. Sensitivity to any harmful consequences for the client must be the foremost ethical consideration. For example, if a psychologist were asked to perform a cognitive evaluation of an elder in the psychologist's religious community, a potential ethical dilemma is that the professional's objectivity may be compromised by the prior relationship. Despite confidentiality, the client might feel unduly exposed in the future when attending religious functions where the psychologist is present. The psychologist's influence and the client's vulnerability in the professional relationship carry over to the secondary relationship (Barnett, 1996).

Conflicts of interest involve current or preexisting financial and business relationships, or acceptance of gifts or services from clients. In these situations, the self-interest of the psychologist can compromise the fiduciary relationship (Smith & Fitzpatrick, 1995). Because of these issues, the change in the recent code permitting barter with clients in circumstances which can be demonstrated as not exploitative or clinically contraindicated has been controversial (Koocher, 1994).

All of this said, the APA Code acknowledges that multiple relationships may be at times unavoidable or unforseeable. Psychologists who work in small rural or ethnic minority communities, for example, encounter a range of situations in which they may incidentally or intentionally interact with clients in other roles or contexts. Graduate training settings which incorporate teaching, supervisory, collegial, research, and social roles, also tend to create overlapping relationships. Dual relationships

Table 1 Types of dual relationships.

Sexual	
	Sexual intimacies with clients
	during therapy
	post-termination
	Sexual intimacies with students, supervisees in training
	Treatment of former sexual partners
	Sexual harassment
Nonsexual	
	Treatment of students, supervisees, employees, friends, relatives
	Research with clients, supervisees, employees, friends, relatives
	Mixing evaluative/supervisory and therapeutic roles
	Therapeutic and collegial or professional relationships
	Post-termination friendship
	Social relationships with clients, students, supervisees, research participants, employees
Conflicts of interest	
	Business/financial relationships with clients
	Acceptance of gifts or services

are not always harmful, however, psychologists must be sensitive to their use of power to avoid any exploitation as a result of their position.

2.23.8.2 Professional Boundaries

Psychologists identify difficulties with professional boundaries as a common ethical dilemma (Pope & Vetter, 1992). In clinical settings, a positive therapeutic relationship is the necessary basis for effective treatment interventions. "Proper boundaries provide a foundation for this relationship by fostering a sense of safety and the belief that the clinician will always act in the client's best interest" (Smith & Fitzpatrick, 1995, p. 500). Attention to the maintenance of professional boundaries is ethically imperative so as to: (i) safeguard the client's trust in the professional relationship, (ii) prevent the loss of clinical or evaluative objectivity, and (iii) avoid any risk of harm or exploitation of the client.

Relevant parameters of the therapeutic framework include a defined professional role, agreed upon times and places for treatment, clear financial agreements, and confidential handling of information (Pope, 1994). Boundary violations are departures from accepted practice which place the client or treatment process at risk (Smith & Fitzpatrick, 1995). Examples might include the clinician's self-disclosure of personal problems or opinions, social involvement outside of treatment, inappropriate or sexualized physical contact (e.g., hugging, touching, kissing), or acceptance of gifts or favors. It is always the responsibility of the professional to maintain sound boundaries, even if the client "tests" the relationship by asking for personal involvement. In most cases, boundary violations represent a gratification of the clinician's needs which overshadow the client's best interests, reflect a misuse of professional power or influence, and risk a violation of the client's trust and dependency (Brown, 1994).

2.23.8.3 Sexual Intimacies in Professional Relationships

A general erosion of treatment boundaries often precedes more serious exploitation of clients. For example, Simon (1991) reported that inappropriate therapist self-disclosure is the most common boundary violation shown to be a precursor to therapist–client sexual intimacy. Although the ethical prohibition of sexual intimacies with clients includes a two year post-termination period (APA, 1992), even after this time period the psychologist must demonstrate that the sexual or romantic relationship with the former client does not involve any exploitation. Interpretation of this standard has emphasized a close examination of several risk factors for the former client's vulnerability to exploitation, such as the nature of the therapy, the client's personal history and mental status, and potential adverse impacts on the client or others, which, in essence, make this an almost never rule (Canter et al., 1994).

Several rationales have been discussed in the literature to support a permanent ban on sexual involvement with clients. Some would argue that because of the power imbalance, the client may not be in a position to clearly evaluate the risks of becoming personally involved with a therapist. The existence of transferential aspects of treatment is widely clinically accepted and some courts have recognized the mishandling of the transference in judgments regarding therapist sexual misconduct (Gabbard, 1994). The internal therapist, a positive nondistorted representation of the therapist, may continue to sustain growth and change for the client up to several years post-termination (Pope, 1994). Treatment models which emphasize the ongoing role of the therapist as a consultant throughout the life span also support the preservation of the treatment relationship.

Although the sexual exploitation of clients is a serious problem in the profession, survey data by Pope, Keith-Spiegel, and Tabachnick (1986) found that sexual attraction to clients is actually a common experience that must be differentiated from therapist–client sexual intimacy. Eighty-seven percent of psychotherapists ($n = 575$) reported that they had experienced sexual attraction to clients on at least one occasion. It is important for psychologists to acknowledge and understand their sexual feelings in order to prevent acting upon them. Recommended avenues of support for clinicians dealing with feelings of attraction that risk undermining the therapeutic relationship include supervision, professional consultation, peer support, or personal psychotherapy. Proactive steps have been recommended to improve clinical training in the complexities of boundary issues in psychotherapy and the sexual dynamics of clinical relationships (Borys & Pope, 1989; Pope et al., 1986).

Once considered only relevant to treatment settings, attention to the role of boundaries in educational and training settings requires psychologists to consider the supervisee's or student's needs first when making decisions that might impact them. The Ethics Code prohibits sexual relationships with students or supervisees in training, where the psychologist has direct evaluative authority (APA, 1992). The pervasiveness of the power differential in graduate training must be appreciated, as faculty may have ongoing evaluative roles for many years. Teaching and supervising faculty must act as appropriate role models of social and sexual behavior (Hammel, Olkin, & Taube, 1996). Research on student–educator sexual contact suggest a potential modeling effect, in that these students themselves later have greater difficulties with professional boundary violations (Pope, Levenson, & Schover, 1979).

2.23.8.4 Risk Management and Self-care

Understanding risk factors for both the client and professional facilitates a proactive stance towards maintaining therapeutic boundaries. Clients with impulsive, dependent, or borderline features, traumatized populations, especially those with a history of sexual abuse, and chronically ill clients may pose particular challenges in maintaining professional boundaries. Rural settings create higher incidences of overlapping professional and personal roles. Risk factors for the psychologist often involve professional or personal isolation, depression, substance abuse, unresolved personal issues, ongoing stressors, or mental impairment. The code encourages attention to conflicts and personal problems which may impair performance, lead to overidentification with the client, or cause the clinician to reinforce client dependency. Because of the risk of bias involved when multiple relationships occur, the use of supervision or consultation with an expert in ethics to facilitate sound decision-making is encouraged (Canter et al., 1994)

A primary function of professional self-care is to reduce risk factors associated with ethical violations, particularly those involving boundary issues. Some theorize that the therapist's well-being is positively related to the therapeutic outcome (Faunce, 1990), and advocate for self-care as an ethical responsibility (Pettifor, 1996). Optimal self-care for professionals involves physical, emotional, mental, and spiritual dimensions (Faunce, 1990).

Awareness of the impact of professional work is imperative. Although emotional stress cannot always be eliminated, it can be understood and managed. Self-knowledge is important regarding (i) motivations and expectations in professional work, (ii) unresolved personal conflicts, (iii) types of clients one works well with or does not work with effectively, (iv) countertransference reactions, (v) acceptance of limitations, and personal reactions to stress.

Use of specific proactive self-care strategies that help balance personal and professional life such as exercise, leisure activities, and meditation or relaxation, can prevent burnout and promote healthy professional functioning. Psychologists can seek support through professional supervision, personal therapy, or ongoing peer support systems.

2.23.9 FORENSIC ACTIVITIES

2.23.9.1 Forensic Roles

Ethical issues in the forensic activities of psychologists are addressed for the first time in the 1992 code. This addition reflects the apparent increase in forensically related activities of psychologists. Forensic mental health experts serve as consultants to attorneys, provide psychological evaluation, and give expert testimony in judicial settings (Perrin & Sales, 1994). General clinicians may be called into court for testimony, receive subpoenas for records, or provide occasional service to the legal system.

Increasingly, the court system and society are turning to mental health records, seeking factual evidence. Therapy records are being subpoenaed, even at a later date, for legal purposes such as worker's compensation claims, divorce and custody proceedings, and personal injury suits. Some treatment populations, such as abuse victims, carry a high potential for forensic involvement (Walker, 1994). Clinicians need to anticipate with their clients the likelihood of intrusion into the privacy of psychotherapy in such situations. It is critical to separate therapeutic and expert evaluative functions, because of the risk of reaching one-sided conclusions when they are derived solely from therapeutically based observations.

Due to the potential impact of their recommendations, psychologists in expert roles must take care in establishing adequate foundations for their conclusions and, conversely, identify any limitations to their data or the scope of their recommendations. The psychologist must remain independent of economic or professional pressure to modify their opinions. Attorneys may attempt to pressure the psychologist to advocate for their client's position (Canter et al., 1994), such as by asking the psychologist to make stronger statements than are supported by the available data. This highlights an important difference between the adversarial nature of the legal system which promotes advocacy, and the scientific model of psychology which values objectivity.

2.23.9.2 Specialty Guidelines for Forensic Psychologists

Because of the complexities and pitfalls involved in forensic practice, two additional documents have been developed to supplement the Ethics Code. The *Specialty Guidelines for Forensic Psychologists* were developed by a joint Committee on Ethical Guidelines of Division 41 and the American Academy of Forensic Psychology to "improve the quality of forensic psychological services offered to individual clients and the legal system" (Committee on Special Guidelines Forensic Psychologists, 1991, p. 655). They have not been adopted by the APA as a whole.

Competency in forensic psychology includes specialized knowledge in an area of forensic expertise, and interdisciplinary knowledge such as familiarity with the legal setting and the legal and professional standards governing psychologists' roles in the judicial system. Expert knowledge of unique populations, for example, criminal populations, divorcing families, or injured workers, is also critical.

The specialty guidelines provide detailed attention to appropriate methods and procedures for conducting forensic evaluations, objective bases for expert opinions, and the role of the expert witness. As with other psychological services, clarification of the professional role, informed consent, careful attention to confidentiality and privilege issues, and thorough documentation are essential. Because of the potential influence of their opinions, particular care must be taken to ensure the civil rights of the parties involved in legal proceedings.

Child custody evaluations, a specialized area of forensic psychology, are among the most complex assessments that psychologists conduct (Oberlander, 1995), and a frequent basis for ethical complaints to licensing boards and ethics committees. The frequency of high conflict divorce cases in custody evaluation and the seriousness of the potential loss to either parent tends to promote contentiousness by the recipient of services. Because psychologists' recommendations have the potential to influence many lives, keen awareness of the needs and rights of multiple parties and maintaining impartiality are critical. The *Guidelines for Child Custody Evaluations in Divorce Proceedings* (APA, 1994) were developed to provide practitioners and ethics committees with more detailed and consistent standards for competence and practice in this difficult area. The guidelines stress the need for balanced assessment based upon interviews with all relevant parties. Psychologists should refrain from making conclusions regarding parties they did not directly evaluate.

2.23.10 FUTURE DIRECTIONS IN PROFESSIONAL ETHICS

The question "How can psychology contribute more to improving the human condition in the future?" (Farley, 1996, p. 773) reflects the aspirational ideal of professional psychology. Various trends are emerging with potential impact on professional psychology and ethical practice. Growing multicultural awareness within psychology is now influencing training programs (Bass, 1996). Practice issues related to the delivery of culturally competent services and access to services for ethnic and minority populations will be a significant challenge in the future.

Psychologists in the USA are actively pursuing hospital staff privileges which give them the authority to manage patient care in hospital settings. A related development is clinical psychopharmacology, a new field of specialization within professional psychology which is beginning to generate relevant ethical practice guidelines (Buelow & Chafetz, 1996). Prescription privileges for psychologists, allowing psychologists to dispense psychotropic medications, is a controversial role which necessitates careful consideration by the profession. Ethical guidelines must address specific issues of competence, scope of practice, and referral (Buelow & Chafetz, 1996), as well as address the relationship between medical and psychological treatment models.

The growth of managed mental health care in the USA is challenging psychologists to develop new service delivery systems and forge contractual relationships with managed care organizations that also preserve professional autonomy. The impact has been divisive among colleagues who may advocate or oppose the emphasis on shorter-term treatment models that promise cost containment, but are seen by others as rationing of treatment. Confidentiality under such models can be easily compromised by the competing need for documentation and quality assurance. Psychologists must increasingly deal with the legal and business aspects of practice, which must be kept in balance with the therapeutic focus on the relationship with the client.

The editorial article *Ethics in cyberspace* encompasses a wide range of issues and dilemmas resulting from the rapid evolution of information technology, office computerization, electronic communication, and use of the Internet (Koocher, 1996, p. 89). The global impact of the information superhighway on psychology and social systems as a whole is yet to be seen.

2.23.11 SUMMARY

This review of ethical issues in professional psychology in the USA has highlighted the concepts of competence, elements of the professional relationship, confidentiality, and multiple relationships, which form the basis of our professional standards. The aspirational values which inspire professional work seek to enhance the personal freedom and self-determination of clients and promote the broad goals of social welfare. Although commonalities in ethics may extend across international boundaries, our knowledge and practice grows out of the context of an industrialized culture. The evolution of information technology will permit greater opportunities for an international dialogue within psychology regarding the use of psychological science and practice to benefit the global society.

2.23.12 REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychological Association (1953). *Ethical standards for psychologists*. Washington, DC: Author.

American Psychological Association (1976). Policy on training for psychologists wishing to change their specialty. Adopted by the Council of Representatives, January 23–25.

American Psychological Association (1981). *Specialty guidelines for the delivery of services by clinical psychologists, counseling psychologists, industrial organizational psychologists, and school psychologists*. Washington, DC: Author.

American Psychological Association (1982). *Ethical principles in the conduct of research with human participants*. Washington, DC: Author.

American Psychological Association (1985). *Standards for Educational and Psychological Testing*. Washington, DC: Author.

American Psychological Association (1987). General guidelines for providers of psychological services. *American Psychologist, 42,* 712–723.

American Psychological Association (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47,* 1597–1611.

American Psychological Association (1993a). *Guidelines for ethical conduct in the care and use of animals*. Washington, DC: Author.

American Psychological Association (1993b). Guidelines for providers of psychological services to ethnic, linguistic, and culturally diverse populations. *American Psychologist, 48,* 45–48.

American Psychological Association (1993c). Record keeping guidelines. *American Psychologist, 48,* 984–986.

American Psychological Association (1994). Guidelines for child custody evaluations in divorce proceedings. *American Psychologist, 49,* 677–680.

American Psychological Association, Ethics Committee (1994). Report of the Ethics Committee, 1993. *American Psychologist, 49,* 659–666.

American Psychological Association, Ethics Committee (1995, November). Ethics Committee plan for next review of ethics code. Unpublished plan.

American Psychological Association, Ethics Committee (1996). Report of the Ethics Committee, 1995. *American Psychologist, 51,* 1279–1286.

Association of State and Provincial Psychology Boards. (1991). *ASPPB Code of Conduct*. Montgomery, AL: Author.

Barnett, J. E. (1996). Boundary issues and dual relationships: Where to draw the line? *The Independent Practitioner, 16,* 138–140.

Bass, L. J. (1996). Future trends. In L. J. Bass et al. (Eds.), *Professional conduct and discipline in psychology* (pp. 143–156). Washington, DC: American Psychological Association.

Bass, L. J., DeMers, S. T., Ogloff, J. R., Peterson, C., Pettifor, J. L., Reaves, R. P., Retfalvi, T., Simon, N. P., Sinclair, C., & Tipton, R. M. (1996). *Professional conduct and discipline in psychology*. Washington, DC: American Psychological Association.

Bennett, B. E., Bryant, B. K., VandenBos, G. R., &

Greenwood, A. (1990). *Professional liability and risk management*. Washington, DC: American Psychological Association.

Bersoff, D. N. (1994). Explicit ambiguity: The 1992 Ethics Code as an oxymoron. *Professional Psychology: Research and Practice, 25,* 382–387.

Borys, D. S., & Pope, K. S. (1989). Dual relationships between therapist and client: A national study of psychologists, psychiatrists, and social workers. *Professional Psychology: Research and Practice, 20,* 283–293.

Brown, L. S. (1994). Boundaries in feminist therapy: A conceptual formulation. *Women and Therapy, 15,* 29–38.

Buelow, G. D., & Chafetz, M. D. (1996). Proposed ethical practice guidelines for clinical pharmacopsychology: Sharpening a new focus in psychology. *Professional Psychology: Research and Practice, 27,* 53–58.

Canadian Psychological Association (1991). *Canadian code of ethics for psychologists. Revised.* Ottawa, ON, Canada: Author.

Canter, M. B., Bennett, B. E., Jones, S. E., & Nagy, T. F. (1994). *Ethics for psychologists: A commentary on the APA Ethics Code*. Washington, DC: American Psychological Association.

Claiborn, W. L. (1982). The problem of professional incompetence. *Professional Psychology, 13,* 153–159.

Committee on Ethical Guidelines for Forensic Psychologists (1991). Specialty guidelines for forensic psychologists. *Law and Human Behavior, 15,* 655–665.

Committee on Legal Issues, American Psychological Association (1996). Strategies for private practitioners coping with subpoenas or compelled testimony for client records or test data. *Professional Psychology: Research and Practice, 27,* 245–251.

Dana, R. (1994). Testing and assessment ethics for all persons: Beginning and agenda. *Professional Psychology: Research and Practice, 25,* 349–354.

Farley, F. (1996). From the heart. *American Psychologist, 51,* 772–776.

Faunce, P. S. (1990). The self-care and wellness of feminist therapists. In H. Lerman & N. Porter (Eds.), *Feminist ethics in psychotherapy* (pp. 185–194). New York: Springer.

Freudenberger, H. J. (1975). The staff burn-out syndrome in alternative institutions. *Psychotherapy: Theory, Research and Practice, 12,* 73–82.

Gabbard, G. (1994). Reconsidering the American Psychological Association's policy on sex with former patients: Is it justifiable? *Professional Psychology: Research and Practice, 25,* 329–335.

Haas, L., & Malouf, J. (1989). *Keeping up the good work: A practitioner's guide to mental health ethics*. Sarasota, FL: Professional Resource Exchange.

Hammel, G. A., Olkin, R., & Taube, D. O. (1996). Student–educator sex in clinical and counseling psychology doctoral training. *Professional Psychology: Research and Practice, 27,* 93–97.

Handelsman, M. K. (1986). Problems with ethics training by "osmosis." *Professional Psychology: Research and Practice, 17,* 371–372.

Hill, M., Glaser, K., & Harden, J. (1995). A feminist model for ethical decision-making. In E. J. Rave & C. C. Larsen (Eds.), *Ethical decision-making in therapy* (pp. 18–37). New York: Guilford Press.

Hobbs, N. (1948). The development of a code of ethical standards for psychology. *American Psychologist, 3,* 80–84.

Keith-Spiegel, P., & Koocher, G. P. (1985). *Ethics in psychology: Professional standards and cases*. New York: Random House.

Kitchener, K. S. (1984). Intuition, critical evaluation and ethical principles: The foundation for ethical decisions in counseling psychology. *The Counseling Psychologist, 12,* 43–55.

Koocher, G. P. (1994). The commerce of professional psychology and the new Ethics Code. *Professional Psychology: Research and Practice, 25,* 355–361.

Koocher, G. P. (1996). Editorial: Ethics in cyberspace. *Ethics and Behavior, 6,* 89.

Lakin, M. (1994). Morality in group and family therapies: Multiperson therapies and the 1992 Ethics Code. *Professional Psychology: Research and Practice, 25,* 344–348.

Nagy, T. F. (1996 January/February). Ethics committees: Investigate or educate? *The National Psychologist,* 15–16.

Oberlander, L. (1995). Ethical responsibilities in child custody evaluations: Implications for evaluation methodology. *Ethics and Behavior, 5,* 311–332.

Overholser, J. C., & Fine, M. A. (1990). Defining the boundaries of professional competence: managing subtle cases of clinical incompetence. *Professional Psychology: Research and Practice, 21,* 462–469.

Payton, C. R. (1994). Implications of the 1992 Ethics Code for diverse groups. *Professional Psychology: Research and Practice, 25,* 317–320.

Perrin, G. I., & Sales, B. D. (1994). Forensic standards in the American Psychological Association's new Ethics Code. *Professional Psychology: Research and Practice, 25,* 376–381.

Peterson, C. (1996). Professional conduct and discipline in psychology self-evaluation. In L. J. Bass et al. (Eds.), *Professional conduct and discipline in psychology* (pp. 277–294). Washington, DC: American Psychological Association.

Pettifor, J. L. (1996). Maintaining professional conduct in daily practice. In L. J. Bass et al. (Eds.), *Professional conduct and discipline in psychology* (pp. 91–100). Washington, DC: American Psychological Association.

Pope, K. S. (1991). Dual relationships in psychotherapy. *Ethics and Behavior, 1,* 22–34.

Pope, K. S. (1994). *Sexual involvement with therapists: Patient assessment subsequent therapy, forensics*. Washington, DC: American Psychological Association.

Pope, K. S., Keith-Spiegel, P., & Tabachnick, B. G. (1986). Sexual attraction to clients: The human therapist and the (sometimes) inhuman training system. *American Psychologist, 41,* 147–158.

Pope, K. S., Levenson, H., & Schover, L. R. (1979). Sexual intimacy in psychology training: Results and implications of a national survey. *American Psychologist, 34,* 682–689.

Pope, K. S., & Vetter, V. A. (1992). Ethical dilemmas encountered by members of the American Psychological Association: A nationwide survey. *American Psychologist, 47,* 397–411.

Retfalvi, T., & Simon, N. P. (1996). Licensing, certification, registration, chartering, and credentialing. In L. J. Bass et al. (Eds.), *Professional conduct and discipline in psychology* (pp. 39–52). Washington, DC: American Psychological Association.

Simon, R. I. (1991). Psychological injury caused by boundary violation precursors to therapist–patient sex. *Psychiatric Annals, 21,* 614–619.

Smith, D., & Fitzpatrick, M. (1995). Patient–therapy boundary issues: An integrative review of theory and research. *Professional Psychology: Research and Practice, 26,* 499–506.

Soisson, E. L., VandeCreek, L., & Knapp, S. (1987). Thorough record keeping: A good defense in a litigious era. *Professional Psychology: Research and Practice, 18,* 498–502.

Tipton, R. M. (1996). Education and training. In L. J. Bass et al. (Eds.), *Professional conduct and discipline in psychology* (pp. 17–38). Washington, DC: American Psychological Association.

Vasquez, M. (1994). Implications for the 1992 Ethics Code for the practice of individual psychotherapy. *Professional Psychology: Research and Practice, 25,* 321–328.

Walker, L. E. (1994). *Abused women and survivor therapy*. Washington, DC: American Psychological Association.

White v. North Carolina State Board of Examiners of Practicing Psychologists; 8810SC1137 North Carolina Court of Appeals, Wake County No. 86-CVS-8131 (Feb. 6, 1990).

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.24

Code of Conduct: Association of State and Provincial Psychology Boards

NORMA P. SIMON
New Hope Guild Training Programs Private Practice, New York, NY, USA

2.24.1 INTRODUCTION 276
2.24.1.1 Committee Formation 276
2.24.2 HISTORY AND PHILOSOPHY OF THE ASPPB CODE OF CONDUCT 276
2.24.2.1 Preparation by Staff 276
2.24.2.2 Disparity of Codes 276
2.24.2.3 Filling the Legal Gap 276
2.24.2.4 Living Document 277
2.24.2.5 New APA Code 277
2.24.3 ASPPB CODE 277
2.24.3.1 Protection for the Public 277
2.24.3.2 Sections of the Code 277
2.24.3.3 Rules of Conduct 277
2.24.4 DISTINCTIVE PROVISIONS 278
2.24.4.1 Competence 278
2.24.4.2 Formal Opinion 278
2.24.4.3 Record Keeping 278
2.24.4.4 Supervisory Records 278
2.24.4.5 Arrangements in an Emergency 278
2.24.4.6 Impaired Objectivity and Dual Relationships 278
2.24.4.7 Therapists Making Deals 279
2.24.4.8 Client Welfare 279
2.24.4.9 Supervisee Welfare 279
2.24.4.10 Major Difference: Law vs. the Code 279
2.24.4.11 Confidentiality in Perpetuity 279
2.24.4.12 Display of License 279
2.24.4.13 Violation of Law 279
2.24.5 AREAS OF AGREEMENT 280
2.24.5.1 Post-therapy Time Limit 280
2.24.5.2 Barter 280
2.24.5.3 Other Comments on Similarities 280
2.24.6 CAUSES OF DISCIPLINARY ACTION AND THE DISCIPLINARY DATA BANK 280
2.24.6.1 Cause for Changes in the Codes 280
2.24.6.2 Disciplinary Data Bank System 281
2.24.7 FUTURE DIRECTIONS 281

2.24.7.1 Code Revisions 281
2.24.7.2 New Developments 282
2.24.8 CONCLUSION 282
2.24.9 REFERENCES 282

2.24.1 INTRODUCTION

2.24.1.1 Committee Formation

In 1987, the Executive Committee of the American Association of State Psychology Boards (AASPB), the former name of the Association of State and Provincial Psychology Boards (ASPPB), under the leadership of then president Patricia Bricklin, formed a new committee, the Model Licensure Committee (MLC), charged with developing a model code of conduct for the ASPPB and a model act for state/provincial licensure. At that time it was abundantly clear to the 59 member Boards (representing 51 jurisdictions in the USA and eight provinces of Canada) that the American Psychological Association (APA) Model Act for State Licensure for Psychologists lacked certain provisions which the Boards considered important. It was also noted by the Boards at that time that the number of charges brought by patients/clients against psychologists was increasing exponentially. At about that time also, the North Carolina courts were hearing the case of *White v. North Carolina State Board of Examiners of Practicing Psychologists* (1990), in which the judiciary eventually ruled that the Ethical Principles of Psychologists set forth in the preamble to the *APA Code of Ethics* could not be used as the basis for disciplinary action due, at least in part, to its lack of specificity. Consequently, the Executive Committee of the AASPB firmly believed that a more definitive and legally precise document was needed to help guide the various state and provincial jurisdictions in refining their own codes of conduct under which their Boards could act.

When the new MLC first met in 1988, it was informed that the Executive Committee placed a high priority on the MLC's development of a new code of conduct prior to drafting its own version of a model act. Accordingly, this was done.

2.24.2 HISTORY AND PHILOSOPHY OF THE ASPPB CODE OF CONDUCT

Each of the 59 jurisdictions already had some form of a code of conduct arrived at via various methods. They either: (i) referenced the *APA Code of Ethics* (or, in Canada, the Canadian Psychological Association (CPA) *Code of Ethics*) by name; (ii) more or less copied the relevant code of ethics without using its name, if the legal requirements of the jurisdiction did not permit legislation to merely adopt by reference a professional association's code of ethics; (iii) utilized a more or less generic code for all professions in the jurisdiction, with the addition of special provisions for psychology; or (iv) developed an individual code on their own, solely to meet the needs of that jurisdiction. Because the various jurisdictions used such different methods to arrive at their codes, uniformity was impossible.

2.24.2.1 Preparation by Staff

The MLC had asked the AASPB staff to provide the Executive Committee with as many codes of conduct of both the US and Canadian jurisdictions as practicable and, when the MLC began its work, it had 27 codes to review and consider.

When the review was finally completed and a new code drafted, portions of codes from 14 of those jurisdictions were incorporated into the new document.

2.24.2.2 Disparity of Codes

One difficulty the MLC faced initially was the obvious disparity among the many codes of conduct already in use: some were general, some were specific, and some were combinations. The MLC debated long and hard about whether to start their draft of the code by using the most detailed or the most generic among those earlier codes as the basis for developing the new *AASPB Code of Conduct*. After lengthy and occasionally heated discussion, it was decided that, since there were so many codes already in existence, it would be far more efficient for the MLC to adapt and supplement the various existing codes into one cohesive document, rather than to draft their new code without utilizing any of that material.

2.24.2.3 Filling the Legal Gap

From the beginning, and by consensus, the MLC also determined that it did not want to develop a code of conduct that differed greatly from ones that already existed, but rather it wanted to develop a code that would fill the problematic legal gaps then existing due to the aspirational nature of the then Canadian and American Psychological Associations' ethics codes. For this reason, the MLC determined that the *AASPB Code of Conduct* should be brief, to the point, and most definitely not aspirational in nature. The MLC wanted their new code to say, as clearly as possible: "This is required of you," or "This is not acceptable." Ambiguity was to be eliminated wherever possible. Also, the code was to be as positive as possible while remaining crystal clear as to permissible behavior for a psychologist.

2.24.2.4 Living Document

In view of the foregoing, the MLC ultimately decided to start with one of the very detailed codes already in existence and reshape it to contain relevant provisions from other codes, as well as new wording designed to meet all of their requirements. That is what was done and what exists today is intended as a living document, designed to change as times and conditions change and to be further clarified if and when any ambiguity is found to exist.

In some major areas the AASPB code was a trail blazer in defining and sharpening the "shalls" and "shall nots" for psychologists. There were other areas, however, where the new code merely restated, in what the MLC believed to be more precise terms, the material already contained in other codes.

2.24.2.5 New APA Code

After the *ASPPB code of conduct* (ASPPB, 1991) was completed and approved by ASPPB member Boards in October 1990, the APA, aware that their *Code of Ethics* no longer met the legal needs of the various jurisdictions, developed its own *Ethical principles of psychologists and code of conduct* (APA, 1992).

2.24.3 ASPPB CODE

The provisions of the ASPPB code were carefully drafted to possess the following characteristics (ASPPB, 1991, p. 2):

(i) They pertain to the process or technicalities of professional relationships, not to the content of the professional's judgment.

(ii) They pertain to the process or mechanics of the professional relationship, not to the content of the professional judgment itself.

(iii) They are intended primarily to protect the public interest.

(iv) They are as nonintrusive as possible.

(v) They are essentially unambiguous.

(vi) They assure that necessary information will be available for the regulatory board to judge compliance.

(vii) They are sufficient unto themselves without explanatory material.

(viii) They are nonoptional and always pertain; they are coercive and nontrivial.

2.24.3.1 Protection for the Public

The ASPPB believes that its code of conduct differs significantly in function and purpose from the ethics codes of professional associations because the primary purpose of the *ASPPB Code of Conduct* is to protect the public, whereas an organization of professionals necessarily also concerns itself with serving the needs and desires of its own often disparate groups of members. Although a professional association's ethics code may state that the protection of the public is the main function of its code, an association's very existence as a "guild" requires that its code of ethics have a different orientation from that of the ASPPB. This is because the latter was promulgated solely as a guide for, or to be incorporated into, legislation authorizing action by a regulatory body whose *only* purpose is to protect the public. To this date, three jurisdictions—Oklahoma, Maine, and South Carolina—have enacted the entire *ASPPB Code of Conduct* into law. Other jurisdictions have adopted parts of the code (R. Reaves, personal communication, 1996, August 6).

2.24.3.2 Sections of the Code

The overall goal of the committee was to make its code of conduct as simple as possible, both to understand and to follow. The code begins with an introduction, expressing its basic philosophy, followed by a section containing definitions intended to provide clarity and specificity in the use of such commonly used terms as "client," "confidential information," "court order," and "professional relationship."

2.24.3.3 Rules of Conduct

The ASPPB Rules of Conduct, which form the main section of the code, although organized around the same major moral values as other broadly descriptive codes of ethics, imply, though do not expressly state, certain underlying basic tenets of ethically minded psychologists. The headings of the ASPPB Rules of Conduct are: Competence, Impaired Objectivity and Dual Relationships, Client Welfare, Welfare of Supervisees and Research Subjects, Protecting Confidentiality of Clients, Representation of Services, Fees, and Statements, Assessment Procedures, Violation of Law, and

Aiding Illegal Practice. Under these headings, the rules themselves contain many subheadings which include explicit requirements.

2.24.4 DISTINCTIVE PROVISIONS

Reported here are examples of the more significant instances which make the ASPPB code distinctive, and a number of important areas where the ASPPB, APA, and CPA codes are congruent. The ASPPB's thinking was that, once a state board examined both the similarities and the differences between the various codes, each jurisdiction would have a sound basis for using the material to develop a code best suited to its own purposes.

2.24.4.1 Competence

Subsection III.A.3 of the ASPPB code mandates certain behavior by a psychologist when developing competency in an area either new to the individual or to the profession. It states, in part:

> The psychologist ... shall engage in ongoing consultation with other psychologists or relevant professionals and shall seek appropriate education and training in the new area. The psychologist shall inform clients of the innovative nature and the known risks associated with the services, so that the client can exercise freedom of choice regarding such services. (p. 8)

Neither of the other codes specifically requires the psychologist to inform the patient of the innovative nature of a treatment procedure.

2.24.4.2 Formal Opinion

Subsection III.A.5 of the ASPPB code states that: "A psychologist rendering a formal professional opinion about a person, for example about the fitness of a parent in a custody hearing, shall not do so without direct and professional contact with, or a formal assessment of, that person" (p. 8).

This is a stronger statement on the subject than those contained in either the APA or CPA codes. If adopted into law, this would have implications for psychologists who work for managed care companies as case managers, and who routinely make decisions without direct contact with, or formal assessment of, a patient.

2.24.4.3 Record Keeping

Subsection III.A.6.a of the ASPPB code sets forth requirements for more detailed record keeping than is mandated by either of the other codes.

The APA has a record keeping guideline, developed by its Committee on Professional Practices and Standards of the Board of Professional Affairs, in addition to the requirement in its ethics code. However, this more detailed guideline is not part of the APA code. The CPA code is quite general and contains no guidelines for the psychologist as to the requisite content of the records.

The ASPPB code (p. 9) states that the psychologist's records shall contain:

(i) the presenting problem(s) or purpose or diagnosis,

(ii) the fee arrangement,

(iii) the date and substance of each billed or service-count contact or service,

(iv) any test results or other evaluative results obtained and any basic test data from which they were derived,

(v) notation and results of formal consults with other providers, and

(vi) a copy of all test or other evaluative reports prepared as part of the professional relationship.

2.24.4.4 Supervisory Records

Subsection III.A.6.d of the ASPPB code contains a provision requiring supervisors to keep records on supervisory sessions for five years. This is predicated on the acknowledgment that supervisors are generally legally responsible for actions taken by their supervisees, and that having accurate records of the supervision could prove to be quite important to both. Neither of the other codes speaks to this issue.

2.24.4.5 Arrangements in an Emergency

Subsection III.A.7 of the ASPPB code requires the psychologist to make arrangements in case of emergency needs of the patient, should the psychologist be unavailable. The APA code also speaks to this point, but the Canadian code does not mention preplanning for such events. The ASPPB code states, in part: "Continuity of care. The psychologist shall make arrangements for another appropriate professional or professionals to deal with emergency needs of his or her clients, as appropriate, during periods of his or her foreseeable absences from professional availability" (p. 10).

2.24.4.6 Impaired Objectivity and Dual Relationships

Subsection III.B.2.b of the ASPPB code states:

> The psychologist, in interacting with a client or former client to whom the psychologist has at any time within the previous 24 months rendered counseling, psychotherapeutic, or other professional psychological services for the treatment or amelioration of emotional distress or behavioral inadequacy, will not . . . [followed by three caveats, two of which are also covered in the other codes and will be noted below under the subject of agreements, but the third of which is singular to the ASPPB Code] . . . enter into a financial or other potentially exploitive relationship with him or her. (p. 11)

The *Canadian code of ethics for psychologists* (CPA, 1991) mentions business interests entered into at the expense of psychologists' clients, which touches on this concern as well.

2.24.4.7 Therapists Making Deals

The MLC had learned of several instances in which therapists had pressured clients to either invest in "deals" involving the therapist or "encouraged" clients to buy items that the therapist was selling. Although these were not as personally harmful as a sexual relationship between therapist and client, the individual client clearly could not make his or her decision as an entirely independent individual. The therapy would be jeopardized because the psychologist would be in a dual relationship with the client and therefore unable to promote the client's welfare in a totally objective manner.

2.24.4.8 Client Welfare

Under client welfare, there are two rules singular to the ASPPB code. Subsection III.C.5 and Subsection 6 of the ASPPB code state: "The psychologist providing services to an individual client shall not induce that client(s) to solicit business on behalf of the psychologist . . . "; and "The psychologist providing services to a client shall make an appropriate referral of the client to another professional when requested to do so by the client" (p. 12).

The former is covered somewhat in the Testimonials section (3.05) of the APA code which states that psychologists: " . . . do not solicit testimonials from current psychotherapy clients or patients or other persons who because of their particular circumstances are vulnerable to undue influence" (p. 8). The CPA code explicitly forbids exploiting a relationship to further the personal, political, or business interests of the psychologist. The specification of the financial prohibition is an important contribution in this code (CPA, 1991).

2.24.4.9 Supervisee Welfare

Again, under Welfare of Supervisees and Research Subjects, Subsection III.D of the ASPPB Code, the inclusion of "financially" as a "shall not" along with "sexually," is new.

2.24.4.10 Major Difference: Law vs. the Code

One of the major differences among the codes appears under Subsection III.E, Protecting Confidentiality of Clients.

Subsection III.E.7 of the ASPPB code states that: "The psychologist may release confidential information upon court order, as defined in Section II of this Code, or to conform with state, federal, or provincial law, rule, or regulation." The code (p. 14) thus clearly states that the law is a higher authority than the *ASPPB Code of Conduct* if provisions of the two are in conflict, whereas both the CPA and APA codes seem to imply that their codes of ethics are above the law. This is a very significant difference and cause for interesting speculation and debate. In this regard, Sinclair (1995) notes how the codes differ as to whether the law or the code shall prevail and that the ASPPB code specifies that the law governs. The other two (in Sinclair's words) "seem to allow for the possibility of some civil disobedience."

2.24.4.11 Confidentiality in Perpetuity

Subsection III.E.12 of the ASPPB code states that: "The psychologist shall continue to treat as confidential information regarding a client after the professional relationship between the psychologist and the client has ceased." This is singular to the ASPPB code (p. 15) and of considerable importance.

2.24.4.12 Display of License

Subsection III.F.1 of the ASPPB code (Representation of Services) is another provision which is contained only in this code, and requires the professional to display a current license. Although this requirement has also been assumed to be a requirement of a number of jurisdictions' codes, it is actually specified in only a few.

2.24.4.13 Violation of Law

Subsection III.I of the ASPPB code (Violations of Law) states that: "The psychologist shall not violate any applicable statute or administrative rule regulating the practice of psychology."

This requirement is specific to this code, although other codes have rules under which this general provision might be assumed.

2.24.5 AREAS OF AGREEMENT

Perhaps the major areas of agreement among all of these codes are those concerning confidentiality in general, dual relationships, and barter. Although the ASPPB code and the 1992 revision of the APA code are more specific than the CPA code, all agree that sexual relationships of any kind during treatment or within two years thereafter (for APA and ASPPB) is not permitted.

2.24.5.1 Post-therapy Time Limit

A great deal of debate occurred at meetings of both the MLC and the APA committees revising the ethics codes in setting time limits. Fourteen jurisdictions require in their administrative codes that two years must elapse before a personal relationship may be initiated. Florida's code states that the psychologist–client relationship is "deemed to continue in perpetuity." Eleven jurisdictions have even made such "sexual misconduct" a criminal offense (R. Reaves, personal communication, 1996, August 6). The ASPPB committee concluded that two years was a legally defensible period of time and so specified. Caveats in both the ASPPB and the APA codes put the burden on the psychologist to prove that the client was not "vulnerable, by reason of emotional or cognitive disorder, to exploitive influence by the psychologist" (ASPPB, 1991, p. 11). Unless the client is proven to not be vulnerable, the prohibition extends indefinitely. The APA code is more detailed than that of the ASPPB in the analysis of what the psychologist would have to prove.

2.24.5.2 Barter

One area under Fees and Statements (ASPPB code Subsection III.F) which is agreed upon by both the ASPPB and the APA codes, but is not mentioned in the CPA code, treats the practice of barter. Although the wording is somewhat different in each of those two codes, both agree that barter can be exploitive and should be avoided. ASPPB's rural psychologists, however, informed the MLC that it is not always possible to avoid barter in their geographical areas, and so have developed ways to safeguard the pride of the client without having the barter become a "dual relationship." One practice was to have the client donate his or her services to a community organization rather than to the psychologist.

2.24.5.3 Other Comments on Similarities

There are many other provisions in which there are slight differences in wording, but substantial agreement in content, among the codes. Sinclair (1996) summarizes the strengths and weaknesses of the codes by stating that:

> the three codes complement one another, serving different purposes and providing both overlapping and singular examples of psychologists' responsibilities. The current versions are attempts by three different bodies of psychologists in Canada and the United States at similar points in time, to define the obligations of psychologists.

2.24.6 CAUSES OF DISCIPLINARY ACTION AND THE DISCIPLINARY DATA BANK

The nature of any code of conduct necessarily flows from the types of complaints that a jurisdiction receives or that an organization finds are being brought against its membership. For example, in earlier APA ethics codes, the references to dual relationships were more general and aspirational than the specific wording regarding sexual relationships with former clients contained in the 1992 *Ethical Principles of Psychologists and Code of Conduct* of the APA as noted below.

Principle 6a of the APA 1989 *Ethical Principles of Psychologists* states:

> Psychologists are continually cognizant of their own needs and of their potentially influential position vis-à-vis persons such as clients, students and subordinates. They avoid exploiting the trust and dependency of such persons. ... Examples of such dual relationships include, but are not limited to, ... sexual intimacies with clients. (p. 393)

The 1992 APA code is far more explicit than its predecessor, as is the ASPPB Code of Conduct.

2.24.6.1 Cause for Changes in the Codes

Changes in the codes of conduct of governmental jurisdictions as well as in the APA and ASPPB Codes, especially regarding dual relationships, were due in large measure to the increasing number of charges brought against psychologists across the country for violating the dual relationship provisions of the existing codes. Because the relevant provision of the code was often phrased in general terms, the ASPPB determined that this essential part of their code needed a definitive clear statement and therefore provided it.

2.24.6.2 Disciplinary Data Bank System

Data collected by the ASPPB Disciplinary Data System in recent years is evidence that these changes were, indeed, needed. Because the ASPPB collects this data, it is able to aid jurisdictions in determining the need to discipline psychologists who violate their codes of conduct in this respect, and the ASPPB also ensures that this information flows across jurisdictional lines so that, if a psychologist has been disciplined in one jurisdiction, all jurisdictions will be informed of this action and be able to react to it. In addition, by having data on the extent to which various areas of the code are being violated, organizations can continue to update their codes to provide more clearly defined rules for the practice of psychology in their jurisdictions.

All of the above depends on maintaining a constantly updated disciplinary data system to house the database for all of the jurisdictions. The ASPPB data bank has been in existence since 1983. The first report on disciplinary actions taken against psychologists by the member Boards was issued on August 15, 1983 (J. Pippin, personal communication, August 21, 1996). However, only since 1995 have the interactive aspects of the database been in place. This allows jurisdictions both to enter data and to check to see if an individual has violated the code in another jurisdiction. The pilot project, which has been in operation serving both small and large jurisdictions and connecting all of them with the ASPPB central office through a secured network, has recently been completed and was a huge success. In the past, two of the problems with the data bank have been the difficulty in collecting the data and convincing the various jurisdictions of the importance of this effort. It was also necessary to convince the jurisdictions of the importance of their supplying the information to the data bank in a timely manner in order to make the system truly useful. With the recent availability of an on-line service, compliance has increased. The ASPPB is now providing access to the system for all member jurisdictions, which was to be completed by October, 1996.

These data, so necessary to help the jurisdictions understand what they may be facing in terms of future disciplinary cases, also inform psychologists where, and what kinds of, education needs to be emphasized in teaching ethics and ethical dilemmas. Table 1 lists the reported disciplinary actions over the 13-year period from August 1983 through July 1996.

2.24.7 FUTURE DIRECTIONS

2.24.7.1 Code Revisions

As the field and practice of psychology change, a number of amendments will need to be made to the *ASPPB code of conduct* and to other organizations' ethics codes. Probably the two most significant developments for the future, and ones that will have huge implications for both professional organizations and licensure boards, are managed care and the Internet. Contracts which mental health providers are being directed to sign with managed care companies are, at least, ethically questionable. The provider is required to hold the company "harmless" and to limit information given to the client/patient, both of which place

Table 1 Reported disciplinary actions for psychologists August 1983–July 1996. Compiled from actions reported to the ASPPB Disciplinary Data System by ASPPB member boards. (Reprinted with the permission of ASPPB.)

Reason for disciplinary action	*Number disciplined*
Sexual/dual relationship with patient	471[a]
Unprofessional/unethical/negligent practice	344[b]
Fraudulent acts	142
Conviction of crimes	105
Inadequate or improper supervision	60[c]
Impairment	52
Breach of confidentiality	45[d]
Improper/inadequate record keeping	41
Fraud in application for license	35
Failure to comply with continuing education requirements	12

[a] When this table was first compiled in January 1988, 45 psychologists had been disciplined for this reason. [b] In 1988 this total was 29. [c] This category doubled between August 1993 and May 1994. [d] There was a 51% increase in this category between August 1993 and May 1994.

the provider in a serious ethical dilemma. This means that organizations and jurisdictions must either change the nature of their codes to accommodate the managed care companies' requirements, or sanction the provider for not adhering more closely to the ethics code or code of conduct of these jurisdictions. New legislation has been introduced in a number of legislatures to remedy some of these problems; however, it will be a long time, if ever, before either the patient or the provider will have the necessary protection once afforded to individuals involved in psychological treatment.

2.24.7.2 New Developments

The Internet, World Wide Web, and other new and exciting communications developments also pose interesting problems for licensing boards. For example, if psychological services are provided by a psychologist on the Internet and a complaint results, which jurisdiction has authority over the provider—the one where the provider lives and works or the one where the client lives; both; or neither? What will happen to the requirement in the ASPPB code mentioned above that the professional has face-to-face contact with the client? What legal changes will be necessary to protect the public while increasing the use and effectiveness of treatment and/or assessment in the new media?

2.24.8 CONCLUSION

All of these codes are in a state of flux (as they should be), so the ways in which they converge and differ will continue to ebb and flow with time. As jurisdictions consider the development or improvement of their laws in this regard, the ASPPB model for a code of conduct should be of considerable assistance in helping each jurisdiction make sound decisions about what it needs to include in a legally defensible document for that jurisdiction. The ASPPB code provides a sound frame of reference for each jurisdiction to consider, because it contains sound, legal language that sets forth clearly and concisely the do's and don'ts governing the appropriate behavior of psychologists necessary to protect the public.

2.24.9 REFERENCES

American Psychological Association (1990). Ethical principles of psychologists (amended June 2, 1989). *American Psychologist, 45*(3). Washington, DC: Author.

American Psychological Association (1992). *Ethical principles of psychologists and code of conduct*. Washington, DC: Author.

Association of State and Provincial Psychology Boards (1991). *Code of conduct*. Montgomery, AL: Author.

Canadian Psychological Association (1991). *Canadian code of ethics for psychologists*. (Revised). Ottawa, ON; Canada: Author.

Sinclair, C. (1995). The Canadian code of ethics for psychologists. *First International Congress of Licensure, Certification, and Credentialling*. Association of State and Provincial Psychology Boards, New Orleans, LA, April.

Sinclair, C. (1996). A comparison of codes of professional conduct and ethics. In Professional Conduct and Discipline in Psychology. American Psychological Association. Washington, DC.

White v. North Carolina State Board of Examiners of Practicing Psychologists, 388 S.E. 2d 148 (N.C. App. 1990).

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.25

Legal Considerations in Practice

RANDOLPH P. REAVES
Association of State and Provincial Psychology Boards, Montgomery, AL, USA

2.25.1 INTRODUCTION 284
2.25.2 DEFINITIONS 284
2.25.3 THE REGULATION OF THE PRACTICE OF PSYCHOLOGY 284
- *2.25.3.1 Statutes Vary Significantly* 284
- *2.25.3.2 Qualifications for Licensure* 285

2.25.4 THE BUSINESS OF PRACTICING PSYCHOLOGY 286
- *2.25.4.1 State Regulation of the Business of Practicing Psychology* 286
- *2.25.4.2 State Laws that Impact the Practice of Psychology* 286
- *2.25.4.3 State Child and Elderly Abuse Reporting Acts* 287
- *2.25.4.4 Federal Laws Impacting the Practice of Psychology* 289
  - *2.25.4.4.1 CHAMPUS* 289
  - *2.25.4.4.2 Medicare* 289
  - *2.25.4.4.3 Medicaid* 289
- *2.25.4.5 Other State and Federal Laws Impacting the Practice of Psychology* 290
  - *2.25.4.5.1 Freedom of choice* 290
  - *2.25.4.5.2 Laws prohibiting certain billing practices* 290
  - *2.25.4.5.3 Laws that involve the legal system* 290
  - *2.25.4.5.4 Commitment laws* 290

2.25.5 CIVIL, CRIMINAL, AND LICENSE RELATED LIABILITY 290
- *2.25.5.1 Civil Liability* 290
  - *2.25.5.1.1 Failure to warn* 290
  - *2.25.5.1.2 Failure to prevent suicide* 291
  - *2.25.5.1.3 Failure to protect* 291
  - *2.25.5.1.4 Failure to properly supervise hospitalized patients* 292
  - *2.25.5.1.5 Negligent release of dangerous patients* 293
  - *2.25.5.1.6 Negligent prescription of contraindicated drugs* 293
  - *2.25.5.1.7 Failure to report suspected child abuse* 294
  - *2.25.5.1.8 Negligence of auxiliaries* 294
  - *2.25.5.1.9 Sexual relations with clients or patients* 295
  - *2.25.5.1.10 Public disclosure of confidential information* 295
  - *2.25.5.1.11 Confidentiality in divorce and child custody proceedings* 296
  - *2.25.5.1.12 Implanting false memories* 297
  - *2.25.5.1.13 Other forms of civil liability* 298
- *2.25.5.2 Criminal Liability* 298
  - *2.25.5.2.1 Sexual abuse of patients* 298
  - *2.25.5.2.2 Fraud and inappropriate billing practices* 299
  - *2.25.5.2.3 Mandatory reporting laws* 300
  - *2.25.5.2.4 License-related liability* 300
  - *2.25.5.2.5 Common reasons for disciplinary action* 300
  - *2.25.5.2.6 Less common reasons for disciplinary action* 301
  - *2.25.5.2.7 Disciplinary actions in other professions* 301
  - *2.25.5.2.8 Mandatory reporting laws* 301

2.25.5.2.9 The complaint process 301
2.25.5.2.10 Constitutional and statutory protections 302
2.25.6 CONCLUSION 302
2.25.7 REFERENCES 302

2.25.1 INTRODUCTION

Many books can be and have been written about psychology and its interaction with the field of law. Just deciding where to begin to discuss psychology and law presents a dilemma of considerable proportions. Since the focus of this chapter is the application of law to the practice of psychology, one logical starting place would be the laws that regulate admission to practice the profession in the various US jurisdictions. However, it would probably benefit the reader to first review the meaning of certain words or terms used throughout the chapter.

2.25.2 DEFINITIONS

There are a number of words or terms that appear regularly throughout this chapter and some explanation is in order.

(i) *Agency*. This refers to an administrative body established by Congress or state legislatures to run the government on a day to day basis.

(ii) *Board*. This is often used to refer to the agency that regulates the practice of psychology in each US jurisdiction.

(iii) *Case law*. This is a term that refers to important provisions of judicial opinions that interpret the meaning of statutes or the common law.

(iv) *Certification*. This is the process by which a governmental or nongovernmental agency or association grants authority to use a specified title to an individual who has met predetermined qualifications.

(v) *Licensure*. This is the process by which an agency of government grants permission to an individual to engage in a given occupation upon finding that the applicant has attained the minimal degree of competency necessary to ensure that the public health, safety, and welfare will be reasonably well protected. Since the law establishing a licensed occupation usually sets forth the "scope of practice" covered by the act, licensing laws are often referred to as "practice acts."

(vi) *Regulations*. This is a term used to refer to the rules (or regulations) passed by federal or state agencies under the rule making authority typically delegated to the agency through its enabling legislation. Often the words rules and regulations are used together.

(vii) *Statute*. This refers to a law passed by Congress or a state legislature. Statutes are also referred to as acts or simply laws.

2.25.3 THE REGULATION OF THE PRACTICE OF PSYCHOLOGY

The first statute that legally recognized the practice of the profession was passed in the state of Connecticut in 1945. Over the course of the next 50 years almost every US jurisdiction passed a statute that either regulates the practice of psychology or dictates which individuals are legally entitled to refer to themselves as psychologists. These include all the states, and the District of Columbia, as well as the territories of Guam and Puerto Rico. In 1994, the US Virgin Islands passed a regulatory law. Given the small populations of the remaining US territories, it is entirely possible that the Virgin Islands is the last jurisdiction that will regulate the profession in this country.

2.25.3.1 Statutes Vary Significantly

Not all the statutes referred to above regulate the actual practice of psychology, but they comprise the clear majority. Those jurisdictions that regulate the practice have enacted laws that include a definition of practice and preclude unqualified and unlicensed persons from performing such functions. The definitions, of course, vary considerably. The definition contained in the Association of State and Provincial Psychology Board (ASPPB) (1992) Model Act reads:

> Practice of psychology is defined as the observation, description, evaluation, interpretation, and/or modification of human behavior by the application of psychological principles, methods, or procedures, for the purpose of preventing or eliminating symptomatic, maladaptive, or undesired behavior and of enhancing interpersonal relationships, work and life adjustment, personal effectiveness, behavioral health and mental health. The practice of psychology includes, but is not limited to, psychological testing and the evaluation or assessment of personal characteristics, such as intelligence, personality, abilities, interests, aptitudes, and neuropsychological functioning; counseling, psychoanalysis, psychotherapy, hypnosis, biofeedback, and behavior analysis and therapy; diagnosis and treatment of mental and emotional

> disorder or disability, alcoholism and substance abuse, disorders of habit or conduct, as well as of the psychological aspects of physical illness, accident, injury, or disability; and psychoeducational evaluation, therapy, remediation, and consultation. Psychological services may be rendered to individuals, families, groups, organizations, institutions and the public. The practice of psychology shall be construed within the meaning of this definition without regard to whether payment is received for services rendered.

In other jurisdictions "title" acts regulate only the title "psychologist" allowing nonpsychologists to perform functions psychologists perform so long as they use other titles such as "therapist" or "psychotherapists." For example, Colorado Revised Statutes 12–43–216 (1992) Title Use Restriction states:

> A psychologist, clinical social worker, marriage and family therapist, or professional counselor may only use the title for which he is licensed under this article. Except as provided in section 12–43–304 (5), no other person shall hold himself out to the public by any title or description of services incorporating the terms "licensed clinical social worker," "clinical social worker," "LCSW," "marriage and family therapist," "LMFT," "professional counselor," "psychologist," "psychology," or "psychological" and he shall not state or imply that he is licensed to practice clinical social work, marriage and family therapy, professional counseling, or psychology. Nothing in this section shall prohibit a person from stating or using the educational degrees which he has obtained.

The viability of "title acts" have been questioned since the decision of the US Court of Appeals for the Eleventh Circuit in *Abrahamson v. Gonzales (1992)*. In this decision this federal court held that denying individuals the right to use the title psychologist while they are legally allowed to perform the functions of a psychologist, violates such practitioners rights to commercial free speech and was unconstitutional. Since the *Abrahamson v. Gonzales (1992)* decision, a number of jurisdictions including Arizona, Florida, Mississippi, and Rhode Island have amended their regulatory statutes in favor of practice acts.

No two psychology regulatory acts are the same, although the pattern of regulation is similar. Most such acts contain definitions, the size and composition of a board of examiners, qualifications for licensure, grounds for disciplinary sanction, renewal requirements and a prohibition against unauthorized practice. Most acts reference some recognized Code of Ethics or Code of Conduct and all include a provision allowing the board to adopt rules and regulations to flesh out the broad parameters of regulatory authority contained in the statute. The boards utilize the rule making process to further define provisions on education, training and other requirements and to detail the everyday operation of the agency.

2.25.3.2 Qualifications for Licensure

One of the major differences among the jurisdictions relates to the level of training required for independent practice. The majority of US jurisdictions require a doctoral degree, although the actual definition of doctoral degree varies among jurisdictions.

Some jurisdictions require a master's degree to qualify for independent practice, however, those jurisdictions comprise a distinct minority that includes North Carolina, Puerto Rico, Vermont, and West Virginia. Many jurisdictions license at both educational levels but require "master's level" psychologists to practice under the supervision of licensed doctoral level psychologists. Titles utilized by these practitioners typically include psychological assistants, associates, and examiners.

Most US jurisdictions issue a generic license entitling the holder of the license to practice psychology. Some states such as Hawaii have chosen to credential only those psychologists that provide health services. Other jurisdictions issue generic licenses and confer health service provider (HSP) designation upon properly qualified practitioners.

There are experience requirements in most US jurisdictions. Typically, two years of supervised experience is required, although the predoctoral year of internship qualifies as a year of experience. Almost every jurisdiction requires passage of the Examination for Professional Practice in Psychology and some form of oral or jurisprudence examination.

At least one jurisdiction, South Carolina, issues licenses in specialty areas, such as clinical, counseling, school, and industrial/organizational psychology. Most jurisdictions, however, recognize individual practitioner's documented areas of competence but do not issue an actual license in the areas or specialties.

School psychology is a recognized specialty in the field and there are many doctoral level school psychologists practicing today. In a few states, which include Ohio and Texas, school psychologists are licensed by the state psychology boards. However, the vast majority of school psychologists do not have doctoral degrees and are not credentialed by boards of psychological examiners, but rather are certified by state boards of education. In most jurisdictions, school psychologists are restricted to practicing within the public school systems and are not allowed to practice independently.

Before closing this section, it should be noted that most psychology regulatory laws contain certain exemptions. These provisions typically exempt individuals in certain practice settings from the application of the statute. It is not unusual for a US jurisdiction to exempt psychologists employed by state or federal governments from meeting the requirements of the regulatory law so long as they only practice psychology within the exempt setting.

It is very important to remember that all psychology boards have rule making powers. These rules can relate to many aspects of practice, from the applicable code of conduct to the length of time records must be retained. Psychologists in every jurisdiction must not only know, but also follow these rules on a regular basis.

2.25.4 THE BUSINESS OF PRACTICING PSYCHOLOGY

Psychologists practicing independently are professionals, but also members of the business community. Like other business people, they are subject to state, federal, and local laws.

2.25.4.1 State Regulation of the Business of Practicing Psychology

Every state, territory, and the District of Columbia have statutes that involve the business practice of psychologists. Some involve reimbursement programs. Others relate to the forms of business practice available to psychologists.

The typical business forms of practice include solo practitioners, partnerships, and professional associations (PAs) or professional corporations (PCs). Many states have passed laws creating a new business entity known as limited liability companies (LLCs). There may be other forms of business practice in some jurisdictions.

Each of these business forms must comply with state and local laws and regulations. Each is required to obtain and renew various business licenses, deduct applicable taxes, and follow all applicable zoning and building laws or codes.

Many states impose restrictions on the corporate practice of professions. It is not unusual for a state statute to forbid stock ownership by nonprofessionals or to prohibit different professionals from holding stock in the same professional corporation. For example,

> **Purpose for which professional corporations may be organized.** Domestic professional corporations may be organized under this article only for the purpose of rendering professional services and services ancillary thereto within a single profession, except that, the same professional corporation or not-for-profit professional corporation may render both medical and dental services, provided that in the case of a professional corporation, at least one shareholder of such professional corporation is a duly licensed medical professional and at least one shareholder is a duly licensed dental professional at the time both services are rendered, and each shareholder is a duly licensed medical or dental professional, or, in the case of a not-for-profit professional corporation, all of the professional services rendered by such corporation are rendered by duly licensed medical professionals and duly licensed dental professionals. (Code of Alabama 1975, Section 10–4–383)

All psychologists that practice independently must choose a form of business practice. It is certainly prudent to consult other knowledgeable professionals such as certified public accountants or attorneys before choosing which business entity to adopt.

2.25.4.2 State Laws that Impact the Practice of Psychology

There are state statutes that impact areas of practice that are not financial in nature. Every jurisdiction has at least one, but probably more, statutes that relate to the confidentiality of patient communications or records. The statutes vary widely and the practitioner must know and understand these legal requirements. For an example of the difference in such confidentiality statutes compare Idaho, Idaho Code §54–2314 (1992) and Oklahoma, Oklahoma Statutes Title 59, §1376 (1993).

> **Privileged communication—confidential relations and communications between psychologist and client.** A person licensed as a psychologist under the provisions of this act cannot, without the written consent of his client, be examined in a civil or criminal action as to any information acquired in the course of his professional services in behalf of the client. The confidential relations and communications between a psychologist and his client are on the same basis as those provided by law between an attorney and client, and nothing in this article shall be construed to require any such privileged communication to be disclosed. (Idaho Code §54–2314(1992))

> **Confidential communications—disclosure—exceptions—threats of patient to self or others.** All communications between a licensed psychologist and

the individual with whom the psychologist engages in the practice of psychology are confidential. At the initiation of the professional relationship the psychologist shall inform the patient of the following limitations to the confidentiality of their communications. No psychologist, colleague, agent or employee of any psychologist, whether professional, clerical, academic or therapeutic, shall disclose any information acquired or revealed in the course of or in connection with the performance of the psychologist's professional services, including the fact, circumstances, findings or records of such services, except under the following circumstances:

1. Pursuant to the provisions of Section 2503 of Title 12 of the Oklahoma Statutes or where otherwise provided by law;
2. Upon express, written consent of the patient;
3. Upon the need to disclose information to protect the rights and safety of self or others if:
   a. the patient presents a clear and present danger to himself and refuses explicitly or by behavior to voluntarily accept further appropriate treatment. In such circumstances, where the psychologist has a reasonable basis to believe that a patient can be committed to a hospital pursuant to Section 5–401 of Title 43A of the Oklahoma Statutes, the psychologist shall have a duty to seek commitment. The psychologist may also contact members of the patient's family, or other individuals if in the opinion of the psychologist, such contact would assist in protecting the safety of the patient.
   b. the patient has communicated to the psychologist an explicit threat to kill or inflict serious bodily injury upon a reasonably identified person and the patient has the apparent intent and ability to carry out the threat. In such circumstances the psychologist shall have a duty to take reasonable precautions. A psychologist shall be deemed to have taken reasonable precautions if the psychologist makes reasonable efforts to take one or more of the following actions:
      (i) communicates a threat of death or serious bodily injury to the reasonably identified person,
      (ii) notifies an appropriate law enforcement agency in the vicinity where the patient or any potential victim resides,
      (iii) arranges for the patient to be hospitalized voluntarily, or
      (iv) takes appropriate steps to initiate proceedings for involuntary hospitalization pursuant to law.
   c. the patient has a history of physical violence which is known to the psychologist and the psychologist has a reasonable basis to believe that there is clear and imminent danger that the patient will attempt to kill or inflict serious bodily injury upon a reasonably identified person. In such circumstances the psychologist shall have a duty to take reasonable precaution. A psychologist shall be deemed to have taken reasonable precautions if the psychologist makes reasonable efforts to take one or more of the following actions:
      (i) communicates a threat of death or serious bodily injury to the reasonably identified person,
      (ii) notifies an appropriate law enforcement agency in the vicinity where the patient or any potential victim resides,
      (iii) arranges for the patient to be hospitalized voluntarily, or
      (iv) takes appropriate steps to initiate proceedings for involuntary hospitalization pursuant to law.
   d. nothing contained in subparagraph (b) of this paragraph shall require a psychologist to take any action which, in the exercise of reasonable professional judgment, would endanger the psychologist or increase the danger to a potential victim or victims, or
   e. the psychologist shall only disclose that information which is essential in order to protect the rights and safety of others;
4. In order to collect amounts owed by the patient for professional services rendered by the psychologist or employees of the psychologist. Provided, the psychologist may only disclose the nature of services provided, the dates of services, the amount due for services and other relevant financial information. If the patient raises as a defense to said action, a substantive assertion concerning the competence of a psychologist or the quality of the services provided, the psychologist may disclose whatever information is necessary to rebut such assertion;
5. In any proceeding brought by the patient against the psychologist and in any malpractice, criminal or license revocation proceeding in which disclosure is necessary or relevant to the claim or defense of the psychologist; or
6. In such situations as shall be defined by the rules and regulations of the Board. (Oklahoma Statutes Title 59, §1376 (1993))

2.25.4.3 State Child and Elderly Abuse Reporting Acts

All states mandate reporting of suspected child abuse. Some states mandate the reporting of suspected abuse of the elderly. Again these statutes vary widely and the competent practitioner must understand the applicable statutes. Compare the wide ranging difference in such laws:

> Mandatory reporting.
>
> (a) All hospitals, clinics, sanitariums, doctors, physicians, surgeons, medical examiners, coroners, dentists, osteopaths, optometrists, chiropractors, podiatrists, nurses, school teachers and officials, peace officers, law enforcement officials, pharmacists, social workers, day care workers or employees, mental health professionals or any

other person called upon to render aid or medical assistance to any child, when such child is known or suspected to be a victim of child abuse or neglect, shall be required to report, or cause a report to be made of the same, orally, either by telephone or direct communication immediately followed by a written report, to a duly constituted authority. (Code of Alabama 1975, Section 26–14–3)

Report, duty; time

(a) Except as provided in subdivision (b), any child care custodian, health practitioner, employee of a child protective agency, child visitation monitor, firefighter, animal control officer, or humane society officer who has knowledge of, or observes a child, in his or her professional capacity or within the scope of his or her employment, whom he or she knows or reasonably suspects has been the victim of child abuse, shall report the known or suspected instance of child abuse to a child protective agency immediately or as soon as practically possible by telephone and shall prepare and send a written report thereof within 36 hours of receiving the information concerning the incident. A child protective agency shall be notified and a report shall be prepared and sent even if the child has expired, regardless of whether or not the possible abuse was a factor contributing to the death, and even if suspected child abuse was discovered during an autopsy. For the purposes of this article, "reasonable suspicion" means that it is objectively reasonable for a person to entertain a suspicion, based upon facts that could cause a reasonable person in a like position, drawing when appropriate on his or her training and experience, to suspect child abuse. For the purpose of this article, the pregnancy of a minor does not, in and of itself, constitute a basis of reasonable suspicion of sexual abuse.

(b) Any child care custodian, health practitioner, employee of a child protective agency, child visitation monitor, firefighter, animal control officer, or humane society officer who has knowledge of or who reasonably suspects that mental suffering has been inflicted upon a child or that his or her emotional well-being is endangered in any other way, may report the known or suspected instance of child abuse to a child protective agency.

(c) Any commercial film and photographic print processor who has knowledge of or observes, within the scope of his or her professional capacity or employment, any film, photograph, videotape, negative, or slide depicting a child under the age of 14 years engaged in an act of sexual conduct, shall report the instance of suspected child abuse to the law enforcement agency having jurisdiction over the case immediately, or as soon as practically possible, by telephone, and shall prepare and send a written report of it with a copy of the film, photograph, videotape, negative, or slide attached within 36 hours of receiving the information concerning the incident. As used in this subdivision, "sexual conduct" means any of the following:

(1) Sexual intercourse, including genital–genital, oral–genital, anal–genital, or oral–anal, whether between persons of the same or opposite sex or between humans and animals.

(2) Penetration of the vagina or rectum by any object.

(3) Masturbation for the purpose of sexual stimulation of the viewer.

(4) Sadomasochistic abuse for the purpose of sexual stimulation of the viewer.

(5) Exhibition of the genitals, pubic, or rectal areas of any person for the purpose of sexual stimulation of the viewer.

(d) Any other person who has knowledge of or observes a child whom he or she knows or reasonably suspects has been a victim of child abuse may report the known or suspected instance of child abuse to a child protective agency.

(e) When two or more persons who are required to report are present and jointly have knowledge of a known or suspected instance of child abuse, and when there is agreement among them, the telephone report may be made by a member of the team selected by mutual agreement and a single report may be made and signed by the selected member of the reporting team. Any member who has knowledge that the member designated to report has failed to do so shall thereafter make the report.

(f) The reporting duties under this section are individual, and no supervisor or administrator may impede or inhibit the reporting duties, and no person making a report shall be subject to any sanction for making the report. However, internal procedures to facilitate reporting and apprise supervisors and administrators of reports may be established provided that they are not inconsistent with this article. The internal procedures shall not require any employee required to make reports pursuant to this article to disclose his or her identity to the employer.

(g) A county probation or welfare department shall immediately, or as soon as practically possible, report by telephone to the law enforcement agency having jurisdiction over the case, to the agency given the responsibility for investigation of cases under Section 300 of the Welfare and Institutions Code, and to the district attorney's office every known or suspected instance of child abuse, as defined in Section 11165.6, except acts or omissions coming within subdivision (b) of Section 11165.2, or reports made pursuant to Section 11165.13 based on risk to a child which relates solely to the inability of the parent to provide the child with regular care due to the parent's substance abuse, which shall be reported only to the county welfare department. A county probation or welfare department also shall send a written report thereof within 36 hours of receiving the information concerning the incident to any agency to which it is required to make a telephone report under this subdivision. A law enforcement agency shall immediately, or as soon as practically possible, report by telephone to the agency given responsibility for investigation of cases under Section 300 of the Welfare and Institutions Code

> and to the district attorney's office every known or suspected instance of child abuse reported to it, except acts or omissions coming within subdivision (b) of Section 11165.2, which shall be reported only to the county welfare department. A law enforcement agency shall report to the county welfare department every known or suspected instance of child abuse reported to it which is alleged to have occurred as a result of the action of a person responsible for the child's welfare, or as the result of the failure of a person responsible for the child's welfare to adequately protect the minor from abuse when the person responsible for the child's welfare knew or reasonably should have known that the minor was in danger of abuse. A law enforcement agency also shall send a written report thereof within 36 hours of receiving the information concerning the incident to any agency to which it is required to make a telephone report under this subdivision. (California Penal Code §11166 (West 1995))

As mentioned, some jurisdictions require psychologists to report suspected abuse of the elderly. An example from Missouri reads:

> Report of elder abuse, penalty—false report, penalty—evidence of prior convictions
>
> 1. When any physician, medical examiner, coroner, dentist, chiropractor, optometrist, podiatrist, resident intern, nurse, hospital and clinic personnel engaged in examination, care or treatment of persons, or other health practitioner, psychologist, mental health professional, social worker, adult day care center worker, nursing home worker, probation or parole officer, Christian Science practitioner, peace officer or law enforcement official, or other person with responsibility for the care of a person sixty years of age or older has reasonable cause to suspect that such a person has been subjected to abuse or neglect or observes such a person being subjected to conditions or circumstances which would reasonably result in abuse or neglect, he shall immediately report or cause a report to be made to the department in accordance with the provisions of sections 660.250 to 660.295, RSMo. Any other person who becomes aware of circumstances which may reasonably be expected to be the result of or result in abuse or neglect may report to the department.
> 2. Any person who knowingly fails to make a report as required in subsection 1 of this section is guilty of a class A misdemeanor.
> 3. Any person who purposely files a false report of elder abuse or neglect shall be guilty of a class A misdemeanor.
> 4. Every person who has been previously convicted of or pled guilty to making a false report to the department and who is subsequently convicted of making a false report under subsection 3 of this section is guilty of a class D felony.
> 5. Evidence of prior convictions of false reporting shall be heard by the court, out of the hearing of the jury, prior to the submission of the case to the jury, and the court shall determine the existence of the prior convictions. (Missouri Revised Statutes §565.188 (1993))

2.25.4.4 Federal Laws Impacting the Practice of Psychology

There are many federal laws that relate to the practice of psychology. These laws can probably be lumped into three groups that would include tax law, reimbursement programs, and statutes that protect the confidentiality of patient records.

The Internal Revenue Code mandates that psychologists adhere to the financial requirements of the business side of practice. Deductions for employees income tax withholdings, Federal Insurance Contributions Act (FICA) and Federal Unemployment Tax Act (FUTA) are required.

A whole book could be written regarding the federal programs that provide reimbursement to psychologists and other health care professionals for services provided to covered individuals. Some examples include the Civilian Health Programs for the Uniformed Services (CHAMPUS), Medicare, and Medicaid.

2.25.4.4.1 CHAMPUS

The Civilian Health and Medical Program for the Uniformed Services provides nonmilitary health benefits to active duty military personnel, retired personnel and their dependents, and to dependents of deceased military personnel. Under CHAMPUS, a psychologist can obtain reimbursement for certain services by meeting qualifications and conditions (Reaves, 1996).

2.25.4.4.2 Medicare

In the United States, Medicare is a federally financed health benefits program primarily for persons aged 65 or older (Medicare Act, 1965). Medicare also covers certain disabled individuals. Medicare is administered by the states; however, the Health Care Finance Administration (HCFA), which is part of the United States Department of Health and Human Services, establishes the relevant rules and regulations. Reimbursement for psychologists' services provided to Medicare beneficiaries is severely restricted, although some recent changes allow for reimbursement on a limited basis.

2.25.4.4.3 Medicaid

Medicaid is a medical-assistance program for certain low-income persons in the United States. The Medicaid program is administered by individual states under broad federal guidelines (Medicaid Act, as amended 1984). Unlike Medicare, which is a federally financed and regulated program, Medicaid is financed by a combination of federal and state funds. The rules governing eligibility, the coverage of treatments, services and procedures, and the method and amount of reimbursement vary widely from jurisdiction to jurisdiction.

The importance of Medicaid to psychologists can vary greatly as well, depending on their practice location and their ability to serve the indigent population (Reaves, 1995).

These statutes contain broad provisions that are further defined by agency rule. Psychologists must adhere to these statutes and their attendant rules or run the risk of serious liability.

2.25.4.5 Other State and Federal Laws Impacting the Practice of Psychology

2.25.4.5.1 Freedom of choice

Approximately 40 states and the District of Columbia have freedom of choice laws that prohibit health insurers from restricting the licensed providers from whom an individual may obtain certain psychological services. These laws are important to the practice of psychology because they assure the patient the right to see nonphysician as well as physician mental health care providers. Some of these laws also are referred to as "direct recognition" laws because they afford the consumer the right to have reimbursed services of a specified professional without a prior physician referral. Freedom of choice and direct recognition are distinct concepts, however, and some state laws provide freedom of choice only after physician referral, that is, without direct recognition of psychologists; at least one federal statute complicates this area. The Employee Retirement Insurance Security Act (ERISA) pre-empts such state statutes and makes them inapplicable to health plans covered by this law.

2.25.4.5.2 Laws prohibiting certain billing practices

There are statutes that prohibit certain types of billing practices. These include federal laws against fraud and abuse in the Medicaid and Medicare programs and state laws against theft, kickbacks, fee splitting, fraud, and similar activities. Health care fraud has been identified as a significant contributor to rising health costs in the United States, and therefore, these fraud and abuse laws have been enacted in many jurisdictions as well as at the federal level.

2.25.4.5.3 Laws that involve the legal system

Every day, the number of psychologists called upon to testify in courtrooms grows larger. Psychological testimony is useful and often critical to decisions regarding competency, child custody, and criminal responsibility. There are hundreds of federal and state laws that could impact such testimony such as laws relating to evidence and confidential communications. An effort to discuss such a lengthy body of law is beyond the scope of this chapter.

2.25.4.5.4 Commitment laws

It should be noted that each jurisdiction has its own unique set of statutes and regulations that concern both voluntary and involuntary commitment of the insane or incapacitated. Psychologists are often called upon to express opinions as to whether individuals should or should not be committed to treatment facilities. These laws are by their very nature complex. Psychologists must know and understand such statutes before attempting to testify as experts in such proceedings.

2.25.5 CIVIL, CRIMINAL, AND LICENSE RELATED LIABILITY

2.25.5.1 Civil Liability

This chapter would not be complete without some explanation of the liability that exists for the negligent, incompetent, or illegal practice of the profession. These lawsuits are based on statutes, common law and rapidly expanding case law.

2.25.5.1.1 Failure to warn

There are probably very few members of the psychological community who are not familiar with the name "Tarasoff." Tatiana Tarasoff was a California teenager killed by a young man (Poddar) who was a patient at Cowell Memorial Hospital at the University of California at Berkeley. The killer was examined prior to the crime by a psychologist who recommended that he be committed, by two psychiatrists who concurred with this evaluation and recommendation, and by another psychiatrist, chief of the department, who countermanded the psychologist's recommendation and directed the staff to take no action to confine him.

Without belaboring the facts, the man carried out his previously expressed threat and killed the unknowing victim. One other important, very discouraging fact, is that the psychologist warned the campus police. Three officers took Poddar into custody, but released him because they considered him rational.

The importance of the Tarasoff case is in the conclusions reached by the Supreme Court of California regarding the liability of the psychologist and psychiatrists. That court held that when such a professional determines, or pursuant to the standards of the profession should determine, that a patient presents a serious danger of violence to another, he incurs an obligation to use reasonable care to protect the intended victim from such danger. Further, the discharge of this duty may require the therapist to take one or more various steps depending on the nature of the case, including warning the victim or others likely to apprise the victim of the danger, notifying the police, or taking whatever steps are reasonably necessary under the circumstances. The duty may also extend to other foreseeable persons who may be injured if the threat is carried out, such as a young child of the intended victim.

In another California case, an appellate court upheld a "dangerous patient" exception to the psychotherapist–patient privilege. The psychologist involved had disclosed to two women his fear that both might be harmed by either of two patients. He advised them of communications made in therapy sessions and reflected on audiotape. This court held that the protective privilege ends where the public's peril begins.

The usefulness of this theory of liability is not limited to California; perhaps as many as six or seven other jurisdictions in different parts of the United States have embraced this theory. These now include Colorado (Perreira v. Colorado, 1989), Georgia (Bradley Center Inc. v. Wessner, 1982), Kansas (Durflinger v. Artiles, 1983), Kentucky (Evans v. Morehead Clinic, 1988), Michigan (Davis v. Lhim, 1983/1988), North Carolina, and Texas (Zezulka v. Thapar, 1996). In other states, such as Wisconsin, significant settlements have been reached based on such theories (Lindsey v. Rousseau, 1993). The "duty to warn" can be found in the Restatement(2d) of Torts. Section 315 states:

> There is no duty so to control the conduct of a third person as to prevent him from causing physical harm to another unless:
>
> (a) a special relation exists between the actor and the third person which imposes a duty upon the actor to control the third person's conduct, or
>
> (b) a special relation exists between the actor and the other which gives to the other a right to protection.

In 1989, the Arizona Supreme Court adopted a "zone of danger" test as the scope of a psychiatrist's duty to protect third parties. In this case, the foreseeability of the victim was the key (Hamman v. County of Maricopa, 1989). However, it is not a theory that will be easily accepted by many courts in our present conservative era. Similar cases have been lost in Connecticut (Fraser v. US, 1996), Indiana (Webb v. Jarvis, 1990), Iowa (Matter of Estate of Votteler, 1982), Minnesota (Cairl v. State, 1982), and South Carolina (Ellis v. US, 1978), and there are a variety of defenses to such suits.

2.25.5.1.2 *Failure to prevent suicide*

Many civil lawsuits are brought against psychologists for failing to prevent patients or clients from taking their own lives. The theory of liability is very similar to that in failure to protect cases. A psychologist with a genuine concern for a patient in such conditions should consider options including: therapeutic management of patients; discussing the matter with the patient; warning potential victims; calling police; seeking involuntary commitment of patient, consistent with state laws; referring the patient to a physician or another provider for prescription of medication; seeking assistance of other providers, clergy or friends of the patient; and trying to involve the patient and family or others in conferring and resolving the threating situation.

Record keeping in such cases is essential. Particular care should be given to recording: precisely what the patient said or did bearing on threats of violence; when the therapist learned what; features of the patient's personal history or clinical status that caused the therapist to believe or disbelieve that violence would occur; and exactly what the therapist did to prevent or protect against the threat.

2.25.5.1.3 *Failure to protect*

Closely akin to "failure to warn" cases are those referred to as "failure to protect" cases. They also are similar to negligent release cases, and both causes of action are typically alleged.

A "failure to protect" case differs from a "failure to warn" case in that the victim in the former is rarely identified or identifiable. Liability is predicated upon the negligence of the provider, which allows the patient access to the victim.

A good example is *Tamsen v. Weber* (1990), in which a patient escaped from the Arizona State Hospital and later abducted a passerby and beat her severely. The litigation later revealed that Traham (the patient) was admitted to the hospital as a danger to himself. While there he attempted suicide. Records that were ultimately produced indicated he was suffering from major depression recurrent with psychotic features and as one expert testified "thinking of killing people."

Nevertheless, the attending psychiatrist granted Traham unsupervised ground privileges. During an unsupervised period on the hospital grounds, he escaped and assaulted Tamsen the next day. The Arizona Court of Appeals held that the psychiatrist owed a duty to protect Tamsen from the violent acts of Traham although Tamsen was not an identifiable victim. Therefore, the alleged negligence in granting unsupervised ground privileges should have gone to a jury on the issue of causation.

An interesting and important case decided in 1994 illustrates how far this liability may extend. In *Almonte v. New York Medical College* (1994), a federal district court refused to dismiss a case against New York Medical College and a professor employed by the college. Both were sued by the parents of a child who was sexually assaulted by a psychiatrist who had trained at the college and had received psychoanalytic training from the individual defendant (Ingram). During the course of his training, the perpetrator (DeMasi) revealed to Ingram that he was a pedophiliac and intended to pursue a career in child psychiatry. The court noted that Ingram, as DeMasi's instructor/analyst, had a measure of control not present in traditional analyst–voluntary patient relationships; he had available to him feasible and not unreasonably burdensome mechanisms for controlling DeMasi. The court also concluded that if no other duty was owed, Ingram had at least the duty to warn or take other steps to protect future patients and therefore the complaint stated a viable claim against Ingram for failure to exercise reasonable care to protect the plaintiffs' son from foreseeable harm.

Some cases that involve harm done by one patient to another might logically fall in this area. A good example is *Halverson v. Pikes Peak Family Counseling and Mental Health Center, Inc.* (1991). Here a patient was sexually assaulted by another patient with a history of violent behavior while she was an inpatient at the facility.

Before leaving this area of liability, it would be prudent to explore the duty to warn or protect foreseeable victims from patients with sexually transmittable diseases. The problem of confidentiality and AIDS poses a dilemma for practitioners. Already a number of physicians have been sued by patients on such theories as negligent infliction of emotional distress when treated by an HIV-positive practitioner who failed to reveal the condition and obtain consent (Faya v. Almaraz, 1993). Many therapists are concerned with the notion that they might be put in the position of having to advise an individual that, should the presence of the HIV virus be revealed to them, they would have no choice but to warn foreseeable victims. That concern is understandable. However, viewing that situation from a purely legal point of view, it easily fits into the failure to warn/protect doctrines.

There is no doubt that an individual who is HIV positive could do considerable harm to an unknowing spouse and/or sexual partner(s). Already there have been numerous successful suits against sexual partners (Kathleen K. v. Robert B., 1984). In many cases, those spouses or partners would be known to the provider and therefore foreseeable victims who could be notified. Arguably, courts will find it easier to uphold a failure to warn (protect) theory in this situation than with a potentially violent person.

A working group on HIV confidentiality at Creighton University Center for Health Policy and Ethics (1991) after conducting a national study on confidentiality in maternal and pediatric HIV, also citing *Tarasoff v. Regents of the University of California* (1976), suggests a formula for decision-making relative to such disclosures:

> When the risk of infection is significant, the identity of the third party is known, warning is likely to be effective in preventing infection, and every reasonable measure to convince the patient to disclose has failed, professionals have a strong ethical obligation to warn those at risk.

At least two courts have agreed. In a 1995 decision, the Alaska Supreme Court found that a physician's disclosure to a husband that his wife had tested positive for the HIV virus was justified in light of the physician's duty to protect third parties from foreseeable exposure to contagious disease. The court described the disclosure as "privileged as a matter of law" (Chizman v. Mackie, 1995). In another 1995 decision, *Reisner v. Regents of the University of California* (1995), the California Court of Appeal held that a third party could sue a physician who failed to advise his girlfriend that she was infected with HIV.

2.25.5.1.4 Failure to properly supervise hospitalized patients

This is another cause of action for negligent conduct. Due to the special relationship which exists between a therapist and a patient, a recognizable duty often arises to protect the patient from himself. Breach of the duty occurs when the therapist does, or more often fails to do something which allows the patient to harm himself. (The reader should note that while many of these cases involve hospitals, invariably the individual providers are also sued.)

A good example of such a supervision case is *Smith v. Rush Presbyterian St. Luke's Medical Centre* (1980), a case in which a young male with "paranoid personality" was admitted to a psychiatric hospital. Following futile attempts to cut his wrists with a pop top, the patient was put on a "suicide risk" list. This classification required observation every 15 min. The patient was also placed in a room with an electrically operated bed and while unobserved, lowered the bed onto his neck and head. The jury verdict was $75 000.

There are a host of such cases and a variety of negligent acts or omissions involved. For instance, there is *Abille v. U.S.* (1980), where a psychiatric patient jumped from an Air Force hospital window. Not only was the patient unattended, but the window had no security device such as a detention screen. In *Pisel v. The Stamford Hospital* (1980), the patient, again unattended, wedged his head into the rails of his bed resulting in brain damage. With *Roesler v. Monora Medical Center* (1978), the patient used a pillowcase to strangle herself and in *Herndobler v. Riveredge Hospital* (1979), an unrestrained patient suffocated herself with a plastic garbage can liner.

There have been a great many cases in which courts have predicated liability on negligent supervision that resulted in injury to other patients. Almost all of these cases involve either rape or other serious physical injuries. Some were perpetrated by other patient, and some involve employees of the various institutions. Of particular concern is *James v. Albergo* (1993), in which an Illinois court held that neither the doctrines of sovereign immunity nor public official immunity barred a lawsuit brought against the professional staff of a state mental institution. In this case, parents alleged that professional personnel at the hospital had misdiagnosed the mental condition of their 5 year old child, recommended his treatment at the hospital, then improperly placed him under the supervision of another minor who eventually committed repeated acts of sexual abuse.

2.25.5.1.5 Negligent release of dangerous patients

Similar to the failure-to-protect type of negligence action is the suit brought for the negligent release of a dangerous patient. Take for instance *Davis v. Lihm* (1983/1988), a Michigan case in which the deceased's 25 year-old son was originally admitted to a state hospital with a diagnosis of paranoid schizophrenia and a history of suicide attempts. Over a 3 year period the patient was admitted and released five more times. His records indicated not only his suicidal tendencies, but the fact that he committed himself to avoid problems at home. Two weeks after his last discharge, his mother was shot and killed by her son while trying to prevent his suicide attempt. The case was tried on both theories, negligent release and failure to warn, and resulted in a jury verdict of $500 000. The case was reversed 5 years later, but on the ground of governmental immunity rather than failure to warn.

A North Carolina case, *Davis v. North Carolina Dept. of Human Resources* (1995), is worth noting. In this negligent release case, the appellate court agreed that liability existed due to the failure of a mental health professional to sufficiently warn a trial judge as to the dangerous nature of a criminal who was committed for such an evaluation.

Negligence in failing to admit suicidal patients or prematurely releasing such patients has also given rise to liability. An example is *Dunn v. Howard University Hospital* (1983), in which a 28 year-old woman was released from an emergency room after a short period of observation. She had been brought to the emergency room by a friend and complained of severe anxiety, depression, and guilt feelings. She also expressed a desire for suicide and self-maiming. The woman had a history of hospitalizations and was previously diagnosed as a manic depressive psychotic. After her release, she jumped to her death. A court later awarded her parents $500 000.

Failure to prevent a patient from escaping from a mental institution can also give rise to liability, particularly if others are injured apprehending the patient. A classic example is *Santangero v. State* (1993), in which two police officers were injured. Evidence in the case revealed that the patient had many previous successful escapes from the institution and that three of those escapes had been accomplished in the same manner. The court agreed that these facts indicated an unreasonable and careless attitude on the part of the state toward the safety of the public as it related to the custody and supervision of the patient. Additionally, the state also violated its regulations by marking the patient as "discharged" 30 days after his escape.

2.25.5.1.6 Negligent prescription of contraindicated drugs

Another cause of action which should be of interest to psychologists, as the movement for prescription privileges grows, is the negligent prescription of contraindicated drugs. This cause of action includes the same elements as other types of professional negligence. It goes without saying that almost every practitioner has a duty to act in a manner that does not cause the patient's condition to worsen. Breach of that duty occurs when the contraindicated drugs, that is, liable to cause the patient harm rather than cure, are prescribed.

A good example of this type of action is *Webb v. Lightburn* (1985), a Colorado suit settled in 1980 for $100 000. In this case the decedent was diagnosed as having a passive–aggressive personality disorder with depressive and hysterical features and a history of drug overdoses. Despite such, and only days after an attempted suicide, the defendant psychiatrist renewed two prescriptions for Seconal. The patient took them all a few days later and died.

Another $100 000 settlement was reached in 1979 in a New York case styled *Haggerty v. New York* (1985). In this case the deceased had been hospitalized previously for suicidal tendencies. On the occasion of his last commitment, Tofranil was prescribed for him, even though his medical records reflected the fact that Tofranil exacerbated his suicidal tendencies. While the ward attendant was asleep, the patient hanged himself with his belt. A similar case is *Greene v. Guarino* (1992). The case involves a substantial settlement for the estate of a man who allegedly committed suicide by ingesting a 30-day supply of an antidepressant. The psychiatrist who prescribed the drug was sued on the theory that he negligently provided a suicidal patient with a toxic amount of the drug.

These are the obvious types of cases, but when dealing with medication of the mentally ill, a number of other areas of potential liability arise. Without spending time outlining them, they include the failure to prescribe adequate antipsychotic medication allowing individuals to cause injury to themselves, an example of which is *Jansen v. University Hospital* (1983).

Another area of concern is the potential for creating a drug dependency. A large verdict resulted in the case of *Badger v. Greenfield* (1981), where a plaintiff produced sufficient proof that his psychiatrist negligently created a drug dependency which later resulted in the loss of significant income when the patient became unable to work. In a more likely scenario giving rise to liability, a physician continued to supply a teenager prescriptions for Preludin and Tuinol allegedly for weight reduction even after it was clear she had become addicted. Despite having promised her parents he would not prescribe for her any more, he did so, and she died of an overdose (Argus v. Scheppegrell, 1985).

Excessive administration of Antabuse has also resulted in liability with large verdicts (Sawyer v. Tauber, 1985), as well as negligent supervision of patients taking such medication (Medi-Stat, Inc. v. Kustunin, 1990). However, the main area of liability, with the largest verdicts, occurs with the improper use of tranquilizers such as Mellaril, Penothiazine, Stelazine, Haldol, and Triavil. A trilogy of cases arising in Iowa (Clites v. State of Iowa, 1982), Colorado (Collins v. Cushner, 1982) and Michigan (Faigenbaum v. Oakland, 1986) indicate that confined patients who develop tardive dyskinesia are the patients who will obtain the most staggering settlements and verdicts. However, for a case involving significant liability when a patient's Haldol dosage was suddenly reduced and damage occurred, see *Leal v. Simon* (1989).

2.25.5.1.7 Failure to report suspected child abuse

Providers should be familiar with mandatory reporting requirements involving suspected child abuse. Occasionally, providers are reluctant to report if the abuse appears to be an isolated event and the abusing parent has sought treatment.

The Missouri Court of Appeals addressed the issue of whether the alleged victim of child abuse could sue two psychologists for failing to comply with the state's Child Abuse Reporting Act. The court held that such a civil suit could not be predicated on the statute since the law did not address the issue of private actions and there was no basis to imply a private right of action since the law did not provide a civil remedy.

However, the court held that the child could bring an action based on common law negligence for failure to warn authorities of the abuse. The court held that the state's public policy favoring protection of children from sexual abuse supports the recognition of a common law duty in this type case. The court limited the right to sue only for failure to warn of specific risks of future harm to readily identifiable victims (Bradley v. Ray, 1995).

2.25.5.1.8 Negligence of auxiliaries

Prior to concluding this section on negligence, it should be noted that very often the person at fault in such cases is not the primary provider, but the staff or clinic auxiliary. The laws of agency and respondent superior require a principal to accept responsibility for the actions of her agents so long as the agent is acting within the scope of her employment.

If the primary provider hires any personnel to assist her in her practice, she must be aware that any negligent acts committed by her assistants will, more than likely, be imputed to her. This situation can be exacerbated if the assistant is not properly qualified, trained or supervised by the primary provider. A classic example is *Andrews v. US* (1984), where a physician's assistant engaged in sexual relations with a female patient under the guise of treatment. Another example is *Doe v. Belmont College* (1992), where a patient sued the counselor's partner, a licensed psychologist, for inadequate supervision and the college that employed both for negligent hiring.

An even better example is *Huntley v. State of New York* (1984), in which a patient at a state psychiatric hospital was permitted to leave the hospital premises and attempted to commit suicide by jumping from the roof of a nearby parking garage. Evidence later revealed that the patient had disclosed her specific suicide plans to a hospital staff member the day before the incident, and the staff member had failed to transmit such information to the staff psychiatrist who controlled the patient's privileges to leave the hospital premises. Auxiliaries should be instructed in writing regarding the confidential nature of patient records. If inappropriate release of such information violates written employee guidelines, liability may not extend to employers.

2.25.5.1.9 Sexual relations with clients or patients

Sexual intimacies with clients or patients are forbidden by all recognized codes of conduct. The harm such conduct does to victims has been studied and clearly documented (*Sexual misconduct in the practice of medicine*). Similar to the reactions of women who have been sexually assaulted, female patients tend to feel angry, abandoned, humiliated, mistreated, or exploited by their therapists. Some have even committed suicide (Koren v. Weihs, 1994). The largest civil judgments have been reserved for those professionals that engage in such heinous conduct.

At least one state's highest court has held that a psychologist who initiates sex with a patient can be liable for negligent infliction of emotional distress. In *Corgan v. Muehling* (1991), the Illinois Supreme Court also held that the patient need not allege or prove physical injury to pursue liability under this cause of action. Consent is no defense to this type of malpractice.

Additionally, some cases find that employment facilities are also liable when such conduct occurs. A good example is *Doe v. Samaritan Counseling Center* (1990). In at least two cases, liability attached when the provider engaged in a sexual relationship with the patient's spouse.

2.25.5.1.10 Public disclosure of confidential information

A very special relationship exists between therapists and their patients. Confidentiality of communication is the base upon which this relationship exists. Public disclosure of confidential communications can give rise to liability if the substance of the facts revealed is offensive, embarrassing, objectionable, or harmful to the patient's reputation (Horne v. Paton, 1973)

Practitioners should understand the differences in certain legal terms so that they can communicate more effectively with patients who know even less about the provider/patient relationship. The three important terms or concepts are privacy, confidentiality, and privilege. Reasonable definitions are:

(i) *Privacy*. Right to be left alone and to make decisions.

(ii) *Confidentiality*. Reasonable expectations that information will not be divulged.

(iii) *Privilege*. Legal protection against being forced to break the expectation of confidentiality.

Any practitioner would be well advised to make certain that the privilege exists before advising a potential client. The circumstances surrounding the manner in which the information was obtained may destroy the privilege (People v. District Court, County of Adams, 1990). Changing statutes and emerging case law make this a real concern. A good example of this dilemma is *State ex ref. Juvenile Dept. of Lincoln County v. Ashley* (1990), in which the Oregon Supreme Court ruled that the psychotherapist–patient privilege did not apply to communications made during drug counseling. Sometimes a supervising provider assumes the privilege attaches to supervisees, when, in fact, it may not. In many states, legislatures have abolished the privilege in cases involving workers' compensation claims (*Code of Alamaba* 1975, Section 25–5–77). However, a West Virginia physician was chagrined to learn that West Virginia statutes had only abolished the privilege for written medical reports. In *Morris v. Consolidation Coal Co.* (1994), an appellate court held that an injured worker could sue for unauthorized *ex parte* oral communication between the physician and an adversarial party that involved providing confidential information unrelated to written medical reports authorized by W.Va. Code 23-4-7. And, providers should not assume that just because the issue of a party's mental condition is raised in a civil proceeding, it automatically means the privilege has been waived.

Patients should obviously be informed of limitations on confidentiality. Suits can arise in the simplest of circumstances. For example, in *Prince v. St. Francis–St. George Hosp., Inc.* (1985), a provider and clinic were sued when a physician sent an insurance form to the client's husband's employer that indicated she was being treated for alcohol abuse. Suits occasionally are brought for wrongful disclosure when the patient becomes a danger to herself or others. However, at least two courts have recognized the provider's responsibility to warn intended victims and have dismissed suits against providers under the "dangerous patient" exception to the privilege (Mendendez v. Superior Court, 1992). It would also be an excellent idea to review some of the other case law in this area. Several courts have ruled favorably on plaintiffs' claims for invasion of privacy and intentional infliction of emotional distress. Psychologists treating patients suffering from acquired immunodeficiency syndrome should be especially careful and know not only case law but also antidiscrimination laws where such exist.

2.25.5.1.11 Confidentiality in divorce and child custody proceedings

In the area of child custody disputes, mental health professionals should act cautiously. Do not evaluate a minor without the custodial parent's consent. Know your state statutes as regards confidentiality, advise the parent(s) correctly, and document the advice carefully. Counsel defending such a claim should review the cases resolved in favor of defendants, such as *Werner v. Kliewer* (1985), *Jordan v. Kelly* (1984), and *Williams v. Congdon* (1979).

Another case of great significance is *Snow v. Koeppl* (1990), which involves the revelation of allegedly confidential information in a family court case. In this case, the Wisconsin Court of Appeals upheld a trial court's decision to dismiss claims for invasion of privacy and breach of confidentiality. The suit was brought against a psychologist who obtained and delivered to the court excerpts from the plaintiff's earlier counseling records pursuant to a court-ordered psychological evaluation of the plaintiff in a family court action. The appellate court agreed that the psychologist was entitled to absolute immunity.

The reviewing court concluded that a judicial order creating official duties, directly and closely connected to court proceedings, should be broadly construed in favor of the person required to act. Thus a directive to perform a psychological evaluation would be construed to protect the court official from liability for the examination and use of reports of earlier psychological counseling or therapy, since such information bore a general frame of reference and relation to the subject matter of the court's order.

A similar case is *Guity v. Kandilakis* (1991), where a psychologist was ordered by a trial court judge to testify about matters arising during joint counseling sessions with a husband and wife. When the husband later sued the psychologist, the court held that he was immune from suit since the court had ordered him to disclose. Just as important is *Howard v. Drapkin* (1990), in which a disgruntled parent sued a psychologist, hired by both parties to evaluate allegations of sexual abuse, for a custody proceeding. The trial and appellate courts agreed that:

(i) the psychologist was entitled to common-law immunity as a quasi-judicial officer participating in the judicial process; and

(ii) the psychologist was entitled to a statutory privilege for publication in a judicial proceeding.

In an important 1994 case, the Texas Supreme Court held that a mental health professional does not owe a duty to a parent not to negligently misdiagnose a child's condition. The court decided that while a risk of harm to a parent accused of sexual abuse was foreseeable, it was outweighed by the need to allow mental health professionals dealing with the sensitive issue of child abuse to exercise their professional judgment in diagnosing abuse "without the judicial imposition of a countervailing duty to third parties" (Bird v. W. C. W., 1994).

However, also note the case of *Awai v. Kotin* (1993), in which the Colorado Court of Appeals held that court-appointed psychologists were not entitled to absolute immunity from claims that they negligently treated a parent referred to them for evaluation in a divorce/child custody proceeding. The court held that absolute immunity for quasi-judicial functions applied only to those functions intimately related and essential to the judicial decision-making process.

Before leaving this area of potential liability, it would be prudent to review case law abrogating the statutory privilege in parental rights termination cases and divorce litigation. In Alabama, for example, there have been three such cases that give cause for much concern. These cases recognize that privileged communication exists, however, the courts have ruled that the welfare of the child outweighs the need for confidentiality. The New York case of *Matter of W.H.* (1993), a child neglect proceeding, is also instructive.

Mental health professionals should consider requiring patients to acknowledge, in writing, the limitations of confidentiality. Counsel dealing with a case involving alleged breach of confidentiality in US government facilities should review the federal Privacy Act.

2.25.5.1.12 Implanting false memories

One of the most controversial areas in psychology practice is the concept of repressed childhood memories. Much of the increase in the acceptance of this concept is credited to the publication in 1988 of *The courage to heal: A guide for women survivors of child sexual abuse* (Bass & Davis, 1988). Since its publication and other similar books, thousands of women have brought forth claims of childhood sexual abuse that they believed had been repressed due to the trauma involved.

One such case was brought in Ohio and resulted in a multimillion-dollar verdict against the uncle of a 33 year-old woman with such a claim.

In the last five years, serious questions have been raised concerning the validity of the concept of repressed memories. There have been studies demonstrating that people are extremely susceptible to suggestion and misinformation regarding their childhood memories and that memories in general are unreliable. A False Memory Syndrome Foundation has been created, which now has many chapters across the country.

Much of the rise and fall of the false memory concept parallels the similar rise and fall of abuse claims involving satanic cults. The rumors of widely spread satanic cults grew in the late 1970s and early 1980s and were popularized by various Christian evangelical groups and the media. In 1989, the book *Michelle remembers* (Smith & Pazder, 1989) was published and alleged a link between such abuse and repressed memories. In the past few years, there have been a number of events that have called into question the existence of such cults, as well as the validity of recovered memories. These include studies by the Federal Bureau of Investigation (Lanning, K. B.: *Ritual abuse: A law enforcement view or perspective,* 1991) and other publications (Wright, L.: *Remember satan,* 1994).

It will be years, perhaps many years, before the mental health community reaches consensus on these issues. However, there are therapists being sued for allegedly implanting false memories in patients' minds. The first such case was tried in California. In *Ramona v. Ramona* (1994), the jury returned a $500 000 verdict against therapists the jury believed had, at least, reinforced memories that were false.

While certainly not as publicized as the Ramona case in California, two other courts have held that parents may pursue claims against mental health professionals for allegedly implanting false memories in their children who were patients of the providers. A case published in the federal reporter is *Sullivan v. Cheshier* (1994), in which a psychologist was sued by a patient's parents who alleged the false memory of sexual abuse by her sibling was implanted while the child was under hypnosis.

In reviewing the case, the federal court held that summary dismissal was inappropriate because a genuine issue of material fact existed. The court noted that the parents had conducted an investigation and found no evidence to corroborate the claims. Additionally, the parents produced a witness who denied the truth of the statements and that prior to the hypnosis sessions, the child had never made such claims. The court also noted that after the patient's statements of abuse, she declined all family contacts unless the family members admitted the truth of the statements.

In a 1995 Pennsylvania case, another federal district court held that parents (Tumans) who had paid for their daughter's (Diane) psychiatric treatment could pursue a breach of contract claim against mental health professionals who allegedly implanted false memories of abuse in the patient's mind. The parents had brought their 20 year-old daughter to a clinic (Genesis) to be treated for an eating disorder.

The clinic, a licensed psychologist, and a licensed social worker were sued. The agents of Genesis agreed to treat the young woman but required the parents to separate from their daughter for 2 years. The parents later alleged their daughter was brainwashed when the therapists implanted false memories that her parents were members of a satanic cult that had murdered her twin and sexually abused her. In group therapy sessions, Genesis counselors allegedly told Diane she had been a victim of a satanic cult, and Diane later made the same statements to the group. According to the Tuman's allegations, counselors solicited money from various sources and gave it to Diane to help her hide from her parents, who, according to the complaint, have not seen or heard from her in more than 2 years.

The court agreed that Diane's parents had adequately stated a breach of contract claim in that they alleged that Genesis counselors breached a contractual promise to provide their daughter with adequate care and to properly treat her eating disorder. The court rejected Genesis' contention that the Tumans' complaint was a negligence claim in disguise and should be dismissed.

The court dismissed a claim for negligence against Genesis, holding the entity did not owe a duty of care to Diane's parents. However, it refused to dismiss negligence claims against the counselors, saying it was reasonably foreseeable that the parents of a patient would be harmed by their negligent conduct. The court also declined to dismiss counts alleging defamation, intentional infliction of emotional distress, and misrepresentation, but did dismiss a claim for interference with filial relations, holding that Pennsylvania courts do not recognize such a cause of action (*Tuman v. Genesis Associates,* 1995).

In a 1994 case, *Farris v. Compton* (1994), the District of Columbia Court of Appeals entered an important ruling on the relevant statute of limitations in such cases. This court held that where the plaintiffs have alleged repression of childhood sexual abuse, their claims do not accrue until they have recovered their memories enough to know that they were injured and why.

The case involved two sisters who sued their brother alleging sexual abuse some 25 years earlier. The trial court dismissed the action, and the federal appellate court certified the statute of limitations question to the D.C. Court of Appeals. That court first noted that since the alleged tort had occurred when the plaintiffs were minors, the statute of limitations did not begin to run until they reached majority. At this point, the court found the plaintiffs had repressed memories of the abuse. And, the court concluded, their claims did not accrue until they discovered or should have discovered all of the essential elements of the cause of action.

The court rejected the defendant's contention that defending himself for actions that were admittedly 25 years old would be an unreasonable burden. It also disagreed with the defendant's assertion that an exception to the running of the limitations statute should only be made by the legislature, noting that it was the court's obligation to construe the word "accrues."

And in July 1995, a state district court jury in St. Paul, MN, awarded $2.4 million to a Minnesota woman who claimed her psychiatrist planted false memories of sexual abuse in her mind. The plaintiff's husband was awarded $210 000 for loss of consortium. According to a report appearing in the Bureau of National Affairs Health Law Reporter (Vol. 4, 1995), the plaintiff alleged that the psychiatrist told her she must have been repeatedly abused by several relatives and had no memory of such because she suffered from "repression and denial" and could not get better unless she not only remembered the abuse but also relived it and confronted her abusers.

Apparently, the plaintiff broke off relations with her immediate family. Additional allegations included the use of mind altering drugs and coercion, as well as failure to advise the plaintiff of the controversial nature of the treatment or alternatives to it. In January 1996, another St. Paul, MN, jury awarded more than $2.5 million to another woman who brought similar claims against the same provider. This particular provider faces six more similar lawsuits from former patients (BNA Health Law Reporter, 1996b).

2.25.5.1.13 Other forms of civil liability

There are any number of other forms of civil liability that could be discussed but are limited by space constraints. These, however, do include what is known as a duty to inform, assault and battery, false imprisonment, fraud, and breach of contract.

2.25.5.2 Criminal Liability

Criminal liability is rarely associated with the practice of the mental health professions. However, there are activities that can result in arrest and incarceration. The two major causes of criminal liability are sex with clients and fraudulent billing practices.

2.25.5.2.1 Sexual abuse of patients

A number of states, including California, Colorado, Florida, Georgia, Idaho, Maine, Minnesota, New Hampshire, New Mexico, North Dakota, Texas, and Wisconsin have enacted statutes that make it a crime for a therapist to engage in sex with a client. Michigan has a statute that allows prosecution for sexual intercourse under the pretext of medical treatment. A number of cases, for such offenses as rape and sexual assault, have been tried to successful conclusions for the prosecution.

Wisconsin was the first state to pass a law to criminalize sex between therapist and client, in 1984. Its original law made such activity a Class A misdemeanor. In 1985, Wisconsin upgraded its statute to make sex between therapist and client a felony. That same year, Minnesota passed a law making such conduct a felony. The Wisconsin statute reads as follows:

> 940.22 Sexual exploitation by therapist; duty to report.
> (2) Sexual contact prohibited. Any person who is or who holds himself or herself out to be a therapist and who intentionally has sexual contact with a patient or client during any ongoing therapist–patient or therapist–client relationship, regardless of whether it occurs during any treatment, con-

> sultation, interview or examination, is guilty of a Class D felony. Consent is not an issue in an action under this subsection. (Wisconsin Statues §940.22 (1985))

The version passed by the Minnesota Legislature reads as follows:

> 609.345 Criminal Sexual Conduct in the Fourth Degree. Subdivision 1. Crime defined. A person who engages in sexual conduct with another person is guilty of criminal sexual conduct in the fourth degree if any of the following circumstances exist:
> ... (h) the actor is a psychotherapist and the complainant is a patient of the psychotherapist and the sexual conduct occurred during the psychotherapy session. Consent by the complainant is not a defense;
> (i) the actor is a psychotherapist and the complainant is a patient or former patient of the psychotherapist and the patient or former patient is emotionally dependent upon the psychotherapist;
> (j) the actor is a psychotherapist and the complainant is a patient or former patient and the sexual contact occurred by means of therapeutic deception. Consent by the complainant is not a defense; ...

According to Schoener, Migrom, Gonsiorek, Luepker, and Conroe (1989) (Minnesota Statutes §609.345 (1985)) in *Psychotherapists' sexual involvement with clients: Invention and prevention,* the first case to be tried in Wisconsin under the 1984 statute involved a doctor of divinity and psychotherapist who allegedly engaged in sex with a female client. The defendant died prior to trial, however, and the case became moot.

However, in 1990, a counselor in Minnesota was convicted of four counts of psychotherapist–patient criminal sexual conduct (State v. Dutton, 1990).

In 1992 the Supreme Court of Colorado considered a constitutional challenge to Section 18–3-405.5(1), 8 B C.R.S. (1988 Supp.), a Colorado statute making sexual penetration of a client by a psychotherapist a class 4 felony. The psychotherapist who was convicted of four counts of aggravated sexual assault under the statute raised a First Amendment challenge.

Colorado's highest court noted that while certain private activities and intimate relationships may qualify for an elevated status of fundamental constitutional rights, the law has never allowed consenting adults simply because they are adults and consent, to engage in any type of sexual behavior of their choice under any circumstances. The court then held that neither a treating psychotherapist nor a client has a fundamental constitutional right to engage in sexual intercourse with the other during the existence of the psychotherapist–client relationship and the statute was therefore not unconstitutional.

The court pointed out the special circumstances inherent in the relationship between psychotherapist and his or her client(s) and held that the state could legitimately protect not only nonconsenting victims of sexual assault but also consenting clients. Due process and equal protection claims were also rejected (Ferguson v. People, 1992).

And, in 1996, a Georgia psychologist was indicted on sexual assault charges for allegedly having sex with a patient while treating her for multiple personality disorder. Several individuals have been successfully prosecuted under the California statute. There are a number of other criminal charges that can result from inappropriate or unlawful sexual contact with patients and colleagues. *Saenz v. Alexander* (1991) is an example of a case in which both criminal and civil liability arose from such conduct.

2.25.5.2.2 Fraud and inappropriate billing practices

Fraud and inappropriate billing practices may also result in criminal liability for mental health professionals. Many have been convicted of fraud or other theft related charges incident to inappropriate or purely illegal billing practices (Marrero v. Dept. of Professional Regulation, 1993).

Virtually every jurisdiction in the US and Canada has laws that prohibit this type of activity. Some are routine criminal statutes that prohibit simple theft, theft by deception, fraud, embezzlement, and the like. Other criminal statutes, enacted to alleviate the abuse within the health care delivery system, are specifically designed for this purpose.

Particular care should be taken when billing the federal government. Often limitations on fees may be found only in obscure federal regulations. Nevertheless, the criminal penalties are serious. In a 1994 prosecution (BNA Health Law Reporter, 1994), a Kansas psychologist was found guilty of soliciting and receiving kickbacks for referring patients and was subject to possible incarceration for 100 years and fines up to $4.75 million. In 1995, a Virginia psychiatric and substance abuse facility pleaded guilty to violating the Medicare and Medicaid anti-kickback statute when it provided improper inducements to prospective patients. Providers' arrangements with the CHAMPUS are now being closely scrutinized. Some cases of inappropriate billing have resulted in serious criminal penalties (US v. Custodio, 1994).

In a 1996 case, a New Mexico psychiatrist learned how serious such matters can be. He was found guilty on 228 counts of submitting false claims to three federal programs. The psychiatrist was not only sentenced to 5 years' probation, but also ordered to repay $700 000 to the federal government and was excluded from the three federal programs (BNA Health Law Reporter, 1996a)

Note that billing activities may involve use of the mail and bring other statutes into play. A number of health care professionals have been convicted of mail fraud for submitting fraudulent reimbursement claims to insurance companies and health plans (US v. Harpster, 1991).

Also note the decision of a New York federal district court, *US v. Willis* (1991). Here the court refused to dismiss a securities fraud indictment against a psychiatrist arising from insider trading on a tip from a patient. Not all such cases result in criminal indictments. In some instances, the federal authorities are willing to forego criminal prosecution, but only after large amounts of money are repaid to the government entities that have been defrauded. In October 1995, a nonprofit drug and alcohol treatment provider in Orlando, FL settled such allegations of falsifying information to the federal government, but only after repaying $500 000.

2.25.5.2.3 Mandatory reporting laws

All mental health professionals should be aware that most mandatory reporting laws include criminal penalties for violators. The possibility of criminal prosecution makes it imperative that every mental health professional familiarizes themself with these statutes and understand the reporting requirements.

2.25.5.2.4 License-related liability

The professional regulatory boards in the various US jurisdictions typically have the legal authority to revoke or suspend professional licenses. In some jurisdictions, that authority rests with a Department of Professional Regulation or some other administrative agency. Regardless of which entity holds the authority, the parameters of such authority will derive from the same statute.

The statute referred to, typically the jurisdiction's "practice act," may contain a reference to a code of conduct or a code of ethics. Violation of the referenced code may be a ground for disciplinary action. For example, the Alabama Psychology Practice Act, *Code of Alabama* 1975, Section 34–26–1 (a)(17) states:

> Violation of the code of ethics adopted in the rules and regulations of the board.

The *Idaho Code* Chapter 23, §54–2305 (1992) provides:

> The board of psychologist examiners shall have the following powers ... (b) To adopt, and, from time to time, revise such rules and regulations not inconsistent with the law as may be necessary to carry into effect the provisions of this act. Such rules and regulations shall include, but not be limited to, (1) a code of ethics for psychologists in the state consistent with the current, and as future amended, ethical standards for psychologists of the American Psychological Association; ...

As these provisions demonstrate, mental health professionals are often legally required to conform their behavior to the dictates of such codes. In many jurisdictions, the relevant code is repeated in full in the board's rules and regulations. However, in others they are simply incorporated by reference. In either case, therapists must know and understand the relevant code as a violation can be grounds for disciplinary action. In some jurisdictions, with some professions, grounds for disciplinary actions will also be found in agency rules and regulations. Woody, in a 1994 study of disciplinary grounds released by the Florida Department of Professional Regulation for the years 1990–1991, 1991–1992, and 1992–1993, found that in that jurisdiction the total number of licensed psychologists averaged 2727 per year. The number of formal complaints averaged 148. Woody's statistics indicate that in the three years studied, 5.3%, 4.6% and 6.7% of licensed psychologists, respectively, had a complaint filed against them (Woody, R. H.: *50 ways to avoid malpractice,* 1988).

2.25.5.2.5 Common reasons for disciplinary action

Over the years, significant data regarding the reasons for license-related liability has been gathered by several organizations, including ASPPB through its Disciplinary Data System. By far the most common cause of disciplinary action is sexual or other dual relationships with patients. However, dual relationships are not the only cause, and in many instances, complaints against psychologists will involve more than one cause for discipline. Table 1 was created by ASPPB in 1995 after careful analysis of its disciplinary data.

During 1993, ASPPB made a concerted effort to convince jurisdictions to release records of disciplinary actions so that an accurate historical database could be created. The jurisdictions

Table 1 Reported disciplinary actions for psychologists: August 1983–July 1996.

Reason for disciplinary action	*Number disciplined*
Sexual/dual relationship with patient	471[a]
Unprofessional/unethical/negligent practice	344[b]
Fraudulent acts	142
Conviction of crimes	105
Inadequate or improper supervision	60[c]
Impairment	52
Breach of confidentiality	45[d]
Fraud in application for license	35
Improper/inadequate record keeping	41
Failure to comply with continuing education requirements	12

[a] When this table was first compiled in January 1988, 45 psychologists had been disciplined for this reason. [b] In 1988, this total was 29. [c] This category more than doubled between August 1993 and January 1996. [d] There was a 50% increase in this category between August 1993 and January 1996.

responded enthusiastically, and during that year, approximately 350 more disciplinary actions were recorded. In 1997, the Disciplinary Data bank holds approximately 1500 names of psychologists whose licenses have been impaired due to negligent, incompetent, unethical, or illegal activities.

There are data banks for other professions. Disciplinary actions against social workers' licenses are being reported to the data bank maintained by the American Association of State Social Work Boards. Actions against psychiatrists' licenses are reported to the National Practitioner Data Bank.

2.25.5.2.6 Less common reasons for disciplinary action

Of the approximately 1500 cases reported to ASPPB the majority fall within the categories listed above. There are, however, many other reasons for disciplinary sanctions and some less common reasons:

(i) failure to make child support payments;
(ii) performing an unrecognized modality (exorcism);
(iii) practicing outside areas of training;
(iv) failure to retain client records;
(v) lack of good moral character;
(vi) assaulting a patient;
(vii) failure to comply with board order;
(viii) lack of professional judgment in communicating personal opinions;
(ix) failure to report suspected child abuse;
(x) use of nude therapy;
(xi) deceptive and misleading advertising of practice;
(xii) misrepresentation of credentials;
(xiii) noncompliance with continuing education requirements.

The number of complaints brought to regulatory boards is rising. Several reasons for the increase exist. Those include the fact that the public is more knowledgeable about the regulatory process, boards are more accessible, and many state ethics committees have abandoned efforts to discipline their memberships (*The National Psychologist,* 1996).

2.25.5.2.7 Disciplinary actions in other professions

Data generated and analyzed by the Federation of State Medical Boards clearly indicate that the number of complaints and disciplinary actions involving physicians and charges of sexual misconduct are rising significantly. It is this author's considered opinion that grounds for disciplinary action against psychiatrists parallel the data generated by the psychology boards. Further, that sexual misconduct, fraudulent billing, and conviction of crimes would clearly be a majority of all cases.

2.25.5.2.8 Mandatory reporting laws

Mental health professionals should be keenly aware of any reporting law relative to knowledge of sexual misconduct by other mental health professionals. A number of states have "whistle blower" statutes, but all are different. Some such laws forbid disclosure without the patient's consent. All such laws have serious repercussions, such as disciplinary action, for failure to follow statutory mandates.

2.25.5.2.9 The complaint process

Virtually all jurisdictions require that complaints concerning psychologists' behavior be in writing. If such a complaint appears valid, an investigation ensues. In some jurisdictions, trained investigators are employed to perform this task. In smaller jurisdictions, a board member may be assigned as an investigating officer.

2.25.5.2.10 Constitutional and statutory protections

Enforcement of these codes involves the application of constitutional and statutory protections. A license to practice is a "property" right that cannot be taken away without due process of law. Therefore, governmental authorities must provide any licensee charged with an infraction at least the following:

(i) adequate notice of the charges;
(ii) an opportunity for a hearing;
(iii) a fair and impartial hearing panel;
(iv) the opportunity to confront and cross-examine adverse witnesses;
(v) a decision based on the evidence;
(vi) a record of the proceedings; and
(vii) some form of judicial review.

These requirements are mandated either by the US Constitution, most state constitutions and/or state statutes. In an effort to standardize procedures among administrative agencies, most have been incorporated in each jurisdictions' administrative procedures act.

2.25.6 CONCLUSION

The interaction of law and psychology is a constant, never ending process. Psychologists are required to know, understand, and follow the requirements imposed by statute and regulation. To be successful they must understand the potential liabilities for unethical, incompetent, or illegal practice. As the foregoing pages illustrate, those requirements and responsibilities are quite significant.

2.25.7 REFERENCES

Abill v. US (1980). 482 F. Supp. 703.

Abrahamson v. Gonzales (1992). 949 F.2d 1567 (11th Cir.).

Almonte v. New York Medical College (1994). 851 F.Supp. 34 (D. Conn.).

Andrews v. US (1984). 732 F.2d 366.

Argus v. Scheppegrell (1985). 472 So.2d 573 (La.).

Association of State and Provincial Psychology Boards (1992). *Model act for licensure of psychologist*. Montgomery, AL: Author.

Association of State and Provincial Psychology Boards (1996). *Handbook of licensing and certification requirements for psychologists in North America*. Montgomery, AL: Author.

Awai v. Kotin (1993). 872 P.2d 1332 (Colo.App.).

Badger v. Greenfield (1981). 24 ATLA L.Rep. 43.

Bass, E., & Davis, L. (1955). *The courage to heal: A guide for women survivors of child sexual abuse*. New York: Harper and Row.

Bird v. W. C. W. (1994). 868 S.W.2d 767 (Tex.).

BNA Health Law Reporter (1994). *3*, 1696.

BNA Health Law Reporter (1995a). *4*, 687.

BNA Health Law Reporter (1995b). *4*, 1266.

BNA Health Law Reporter (1995c). *4*, 1572.

BNA Health Law Reporter (1996a). *5*, 223.

BNA Health Law Reporter (1996b). *5*, 742.

Bradley Center Inc. v. Wessner, Ga., 296 S.E.2d 693.

Bradley v. Ray (1995). 904 S.W.2d 302 (Mo.App. W. D.).

Cairl v. State (1982). 323 N.W. 2d 20 (Minn.).

California Penal Code § 11166 (West 1995).

Chizman v. Mackie, 896 P.2d 196 (Alaska 1995).

Clites v. State of Iowa, Iowa App., 322 N.W.2d 917.

Code of Alabama 1975, Section 10–4–383.

Code of Alabama 1975, Section 25–5–77.

Code of Alabama 1975, Section 26–14–3.

Code of Alabama 1975, Section 34–26–46(a)(17).

Collins v. Cushner (1982, July). 25 ATLA L. Rep. 185.

Colorado Revised Statutes §12–43–216 (1992).

Corgan v. Muehling (1991). 574 N.E. 2d 602 (Ill).

Creighton University Center for Health Policy and Ethics Working Group on HIV Confidentiality (1991). Confidentiality and its limits: ethical guidelines for maternal/pediatric HIV infection. *Creighton Law Review*, *25*, 1439–1460.

Davis v. Lhim (1983/1988). 335 N.W. 2d 481 (Mich. App.) (reversed on the ground of immunity in 1988 by the Michigan Supreme Court, 442 N.W. 2d 688).

Davis v. North Carolina Dept. of Human Resources (1995). 465 S.E.2d 2 (N.C. App.).

Doe v. Belmont College (1992). 35 ATLA L. Rep. 148.

Doe v. Samaritan Counseling Center (1990). 791 P.2d 344 (Alaska).

Dunn v. Howard University Hospital (1983). 26 ATLA L. Rep. 232.

Durflinger v. Artiles (1983). 673 P.2d 86 (Kan.).

Ellis v. US 484 F. Supp. 4 (D.C.S.C. 1978).

Evans v. Morehead Clinic, Ky. App., 749 S.W.2d 696.

Faigenbaum v. Oakland Medical Center (1986). 393 N.W. 2d 847 (Mich).

Farris v. Compton (1994). 652 A.2d 49 (D.C.).

Faya v. Almaraz (1993). 620 A.2d 327 (Md.)

Ferguson v. People (1992). 824 P.2d 803 (Colo.).

Fraser v. US (1996). 674 A.2d 811 (Conn.).

Green v. Guarino (1992). 35 ATLA L. Rep. 29.

Guity v. Kandilakis (1991). 821 S.W.2d 595 (Tenn.App.).

Haggerty v. New York (1985). 23 ATLA L. Rep. 285.

Halverston v. Pikes Peak Family Counseling and Mental Health Center, Inc. (1991). 795 P.2d 1352 (Colo.App.).

Hamman v. County of Maricopa (1989). 775 P.2d 1122 (Ariz.).

Herndobler v. Riveredge Hospital (1979). 23 ATLA L.Rep. 139.

Horne v. Patton (1973). 287 So.2d 824 (Ala.).

Howard v. Drapkin (1990). 271 Cal.Rptr. 893 (Cal.App.2 Dist.).

Huntley v. State of New York (1984). 464 N.E. 2d 467 (N.Y.).

Idaho Code §54–2305 (1992).

Idaho Code §54–2314 (19) (1992).

In re C.P. (1990). 563 N.E. 2d 1275 (Ind.).

James v. Albergo (1993). 626 N.E. 2d 1137 (Ill.App.).

Jansen v. University Hospital (1983). 26 ATLA L.Rep. 138.

Jordan v. Kelly (1984). 728 F.2d 1.

Kathleen K. v. Robert B. (1984). 198 Cal.Rptr. 273 (Cal.App.).

Koren v. Weihs (1994). 607 N.Y.S. 257 (A.D.1 Dept.).

Lanning, K. B. (1991). *Ritual Abuse: A Law Enforcement View or Perspective*, 15 Child Abuse and Neglect, pp. 171–173.

Leal v. Simon (1989). 542 N.Y.S. 2d 328 (A.D.2 Dept.).

Lindsey v. Rousseau (1993). 36 ATLA L. Rep. 220.

Marrero v. Dept. of Professional Regulation (1993). 662 So.2d 1109 (Fla. App. 1 Dist.).

Matter of Estate of Votteler, Iowa, 327 N.W.2d 759.

Matter of W. H. (1993). 602 N.Y.S.2d 70 (Fam. Ct.).

Medi-Stat, Inc. v. Kustunin (1990). 792 S.W.2d 869 (Ark.).

Mendendez v. Superior Court (1992). 834 P.2d 785 (Cal.App.).

Minnesota Statutes §609.345 (1985).

Missouri Revised Statutes §565.188 (1993).

Morris v. Consolidation Coal Co. (1994). 446 S.E. 2d 648 (W.Va.).

Ohio Revised Code Annotated §4732.12 (Baldwin, 1994).

Oklahoma Statutes Title 59, §1376 (1993).

People v. District Court, County of Adams (1990). 797 P.2d 1259 (Colo.).

Perreira v. Colorado (1989). 768 P.2d 1198 (Colo.).

Pisel v. Stamford Hospital (1980). Conn. 430 A.2d 1.

Prince v. St. Francis-St. George Hosp., Inc. (1985). 484 N.E. 2d 265 (Ohio App.).

Ramona v. Ramona (1994). No. 61898, Cal. Superior Ct.

Reaves, R. P. (1995). *Avoiding liability in the mental health professions*. Montgomery, AL: Publications for Professionals.

Reaves, R. P. (1996). Laws that Affect the Practice of Psychology. In American Psychological Association & Association of State and Provincial Boards (Eds.), *Professional conduct and discipline in psychology* (pp. 109–116). Washington, DC: American Psychological Association.

Reisner v. Regents of the University of California (1995). 37 Cal.Rptr.2d 518 (Ct. App.).

Roesler v. Menora Medical Center (1978). 21 ATLA L. Rep. 327.

Saenz v. Alexander (1991). 854 So.2d 1061 (Fla.App. 1 Dist.).

Santangelo v. State (1993). 601 N.Y.S. 2d 305 (A.D.2 Dept.).

Sawyer v. Tauber (1985). 23 ATLA L.Rep. 277.

Schoener, G. R., Migrom, J. H., Gonsiorek, J. C., Luepker, E. T., & Conroe, R. M. (1989). *Psychotherapists' sexual involvement with clients: Intervention and prevention*. Minneapolis, MN: Walk-In Counseling Center.

Sexual misconduct in the practice of medicine. *Journal of the American Medical Association*, *266*(19), 34.

Smith, M., & Pazder, L. (1989). *Michelle remembers* (1991). New York: Pocket Books.

Smith v. Rush Presbyterian St. Luke's Medical Center (1980). 23 ATLA L.Rep. 44.

Snow v. Koeppl (1990). 464 N.W. 2d 215 (Wis.App.).

South Carolina Code Annotated §40–55–180 (Law. Co-op. 1992).

State ex rel. Juvenile Dept. of Lincoln County v. Ashley (1990). 790 P.2d 547 (Or. App.).

State v. Dutton (1990). 450 N.W. 2d 189 (Minn.App.).

Sullivan v. Cheshier (1994). 846 F. Supp. 654 (N.D. Ill.).

Tamsen v. Weber (1990). 802 P.2d 1063 (Ariz. App.).

Tarasoff v. Regents of the University of California (1976). 551 P.2d 334.

The National Psychologist, (1996). *Jan/Feb*, 8.

Tuman v. Genesis Associates (1995). D.C.E. Pa., No.95–0272.

US v. Custodio (1994). 39 F.3d 1121 (C.A. 10 Colo.).

US v. Harpster (1991). 759 F. Supp. 735 (D. Kan.).

US v. Willis (1991). 778 F. Supp. 205 (S.D.N.Y.).

Webb v. Jarvis (1990). 553 N.E. 2d 151 (Ind. App. 1 Dist. 1990).

Webb v. Lightburn (1985). 23 ATLA L.Rep. 285.

Werner v. Kliewer (1985). 710 P.2d 1250 (Kan.).

Williams v. Congdon (1979). 257 S.E.2d 677 (N.C.App.).

Wisconsin Statutes §940.22. (1985).

Woody, R.H. (1988). *50 ways to avoid malpractice*. Sarasota, FL: Professional Resource Exchange.

Wright, L. (1994). *Remember satan*, New York: Knopf.

Zezulka v. Thapar (1996). No. 01–94–01195-CV (Tex. Ct.App., 1st Dist. Jan. 29).

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.26

Responsible Test Use in Clinical Application

MANFRED J. MEIER
University of Minnesota, Twin Cities, MN, USA

2.26.1 INTRODUCTION 305
2.26.1.1 Representative Standards Documents 306
2.26.2 SUMMARY OF THE JCTP FACTORS 307
2.26.2.1 Factor 1: Comprehensive Assessment 307
2.26.2.2 Factor 2: Proper Test Use 307
2.26.2.3 Factor 3: Psychometric Knowledge 307
2.26.2.4 Factor 4: Maintaining Integrity of Test Results 307
2.26.2.5 Factor 5: Accuracy of Scoring 307
2.26.2.6 Factor 6: Appropriate Use of Norms 307
2.26.2.7 Factor 7: Interpretive Feedback—Clients/Patients/Referral Sources 308
2.26.3 IMPLICATIONS OF THE RUST STATEMENT REVISED 308
2.26.3.1 Test Decisions 308
2.26.3.2 Qualifications of Test Users 309
2.26.3.3 Test Selection 309
2.26.3.4 Test Administration 310
2.26.3.5 Test Scoring 310
2.26.3.6 Test Score Interpretation 311
2.26.3.7 Communication of Test Results and Interpretations 311
2.26.4 CONCLUSION 312
2.26.5 REFERENCES 312

2.26.1 INTRODUCTION

This chapter is designed to assist the user of psychological tests in a professionally and ethically responsible manner, consistent with the selected published guidelines and standards. It is intended to identify salient and relevant issues or principles to facilitate the structuring of test use and the interpretation of test data. Therefore, it is hoped to assist graduate students and beginning practitioners in pursuing the goal of maximizing the usefulness of psychological testing and assessment in clinical practice.

Although psychology has been perhaps the foremost discipline in the development and use of formal tests for assessing cognition and behavior, psychological testing and assessment have not always been central to the activities of clinical psychologists. Instead, the importance of psychological testing has varied and for a considerable period of time was even considered to be of relatively little importance in the practice of clinical psychology. After World War II, the practice of psychotherapy became more central, even to the exclusion of psychological testing and assessment for many practicing clinical psychologists. However, over the past 50 years, psychological assessment has assumed greater importance in everyday application, largely as a function of the expansion of assessment activities in clinical neuropsychology and in personality assessment. Despite such cyclical trends, the self-monitoring of responsible practices and the search for effectively applicable standards continues, not only in psychology but in related disciplines as well. Deliberations directed at the development of test standards have been generated by the Joint Committee on Testing Practices which consists of representatives from the American Counseling Association (ACA), the American Psychological Association (APA), the American Speech–Language–Hearing Association (ASLHA), and the National Association of School Psychologists (NASP). Similarly, the Standards for Educational and Psychological Testing (American Psychological Association, 1985) have been revised repeatedly since their inception in 1947 as a result of the cooperative effort of the American Educational Research Association (AERA), the APA, and the National Council on Measurement in Education (NCME). Numerous other organizations have also provided leadership and direction in this extensive domain. It is beyond the scope of this chapter to integrate the many perspectives available in the literature relative to responsible test use. The selective identification of issues and concerns for the present purpose will hopefully orient the interested graduate student or practitioner in clinical psychology to access additional information as needed to establish the responsible practice of clinical psychological assessment.

2.26.1.1 Representative Standards Documents

The Standards for Educational and Psychological Testing (SEPT) reflect understandings reached by consensus on the part of the various organizations that have been involved in their periodic revision since the 1940s. The latest revision is in process and is expected to be made available to the participating organizations some time in 1997 for their ratification. The last previous revision was published in 1985 and was divided into four major sections: Technical Standards for Test Construction and Evaluation; Professional Standards for Test Use; Standards for Particular Applications; and Standards for Administrative Procedures. Of particular relevance for this chapter is Part II, Professional Standards for Test Use, particularly the sections on general principles and clinical testing. The new revision is expected to combine the various specialty chapters into just two chapters: one dealing with educational and the other with psychological testing and assessment issues. The psychological testing component adds material for some of the specialized clinical applications, particularly in relation to neuropsychological and forensic assessments.

Familiarity with this document can be regarded as essential for structuring individual practices in a responsible manner. The many principles promulgated in the standards can be accessed in a more limited and circumscribed manner in the *Responsibilities of Users of Standardized Tests* (RUST) published by the Association for Measurements in Counseling and Education (1991). Such documents provide statements of the collective reflections of many individuals and organizations. Another valuable and systematically organized work is an empirically based product of the Joint Committee on Testing Practices (JCTP, 1988) with representatives of the ACA, the APA, the ASLHA, and the NASP (Eyde et al., 1993). The latter represents the outcome of the work of two committees of the JCTP: the Test User Qualifications Working Group (TUQWG); and the Test User Training Work Group (TUTWG). The resulting study sought to identify the various issues and problems underlying test misuse by means of a database utilizing "critical incidents" detection methods that involved various tests and testing contexts or situations. Such efforts add empirically to the basis for defining, monitoring, and evaluating testing and assessment practice in a given setting or situation. The "critical incident" method isolated a collection of case studies relating to 12 generic test types: educational achievement; employment; group scholastic ability; hearing; individual intelligence; interest; multiple abilities; neuropsychological; personality; projective; readiness; and speech–language. The resulting case book was intended to augment texts and other instructional materials in a graduate curriculum for potential use in the seven assessment contexts from which the case studies were derived: counseling/training; education; employment; general health; mental health; neuropsychology; and speech–language–hearing.

The remainder of this chapter is an attempt to compare the factors yielded by the JCTP study to the test user sequence based on the RUST Statement (Revised 1991). Such a comparison and preliminary integration of those efforts will hopefully provide a foundation for utilizing the new revision of the SEPT 1985 document for the responsible use of psychological and neuropsychological tests in clinical practice.

2.26.2 SUMMARY OF THE JCTP FACTORS

It seems reasonable to assume that the factors derived from the case analyses in the JCTP study will generalize with some adjustments for any dissimilarities across settings and across different purposes of testing. In Eyde et al. (1993), the major factors summarized below also include an account of selected elements that loaded highest on each factor, including those that appear to have the greatest implication for competent test use in general clinical, neuropsychological, and forensic settings.

2.26.2.1 Factor 1: Comprehensive Assessment

This factor reflects the distinction between testing and assessment in applications since it relates to the need to obtain a comprehensive personally relevant database for integration with test score databases to increase the validity of test score-based inferences in the interpretation (Matarazzo, 1990). In addition to medical and psychosocial histories, the comprehensiveness of the assessment requires a working knowledge of the shortcomings of personal qualifications as well as the known limitations of particular tests, consideration of clinical observations during interview and testing, and the possible use of additional tests to confirm or disconfirm inferences that can only be stated as hypotheses (rather than firm conclusions) on the basis of the initial assessment strategy. Comprehensive assessment obviously calls for considerable judgmental elaboration within an empirically sound historical and test score framework.

2.26.2.2 Factor 2: Proper Test Use

This factor appeared to relate to the acceptance of a strong test user responsibility for establishing competence and for the application of a deliberate quality control procedure that reflects an understanding of the multivariate nature of the assessment enterprise. This responsibility includes maintenance of a neutral and objective attitude and the necessary education and training for attaining quality control and the necessary competence for effective assessment practice.

2.26.2.3 Factor 3: Psychometric Knowledge

This factor relates primarily to the knowledge of the technical properties of tests to a depth that permits the user to evaluate the significance of standard errors of measurement, reliability, and validity and to incorporate such knowledge into test interpretation. Considering the vast array of psychological and neuropsychological tests in use, any given clinical practitioner will not ordinarily have a working knowledge of the technical properties of all the available tests. Instead, a typical practitioner will focus on a circumscribed body of tests that are both routinely used and well-understood in terms of applications pertaining to particular referral questions and test purposes. Lacking such knowledge for a particular test would then require additional education and training such as continuing education and formal postdoctoral residency, for example.

2.26.2.4 Factor 4: Maintaining Integrity of Test Results

Although factor 3 also involves the knowledgeable application of psychometric principles, such principles may or may not be effectively applied in the actual interpretation of test results. Thus, this factor relates more to limitations of particular test scores for the purposes being addressed in specific situations or referral contexts. Elements loading highest on this factor include the necessity for avoiding the use of absolute cut-off scores since they do not take into account errors of measurement. Factor 4 reflects the need to correctly apply psychometric knowledge when interpreting a test or test battery in a particular set of circumstances. The necessity for correct application increases as a function of the gravity of any actions that may ultimately be based on test interpretation such as diagnostic, dispositional, and treatment decisions.

2.26.2.5 Factor 5: Accuracy of Scoring

Responsible test use and interpretation obviously depends on the accuracy of the test scores. Psychological test interpretation, accentuated even more in the forensic context, necessitates the establishment of audit procedures in the recording process and periodic checks on score accuracy. Even seemingly minor scoring errors can grossly reduce the validity of inferences derived from individual tests and test battery scores.

2.26.2.6 Factor 6: Appropriate Use of Norms

In a highly diverse society, even norms that have wide applicability need to be understood for their limitations when applied to groups of individuals that were not adequately represented in the normative data for the test or test battery at issue. The effective application of normative data in test interpretation becomes even more crucial as the decisions based on test interpretations, if invalid, may have serious deleterious consequences for the individuals being assessed. Where norms are limited, responsible test use requires extreme caution in the derivation of any conclusions from psychometric data.

2.26.2.7 Factor 7: Interpretive Feedback—Clients/Patients/Referral Sources

Direct interpretive feedback to clients occurs less frequently in clinical settings but is rather routine in counseling situations. The referring professionals (e.g., physicians, attorney) ultimately may interpret the test results to individuals based on the clinical psychological or neuropsychological report. The factor calls for the need for the test interpreter to educate the referring professional to perform this role within the consultation framework or permit the interpreter to assume this role. Interpretive feedback requires the use of language that is readily understood by the recipient and the appreciation of the likely consequences of providing such information to an emotionally disturbed or cognitively impaired individual.

In addition to utilizing the JCTP case book as a means of identifying the more common aspects of test misuse, the case studies for assessing human behavior permit an analysis of the incident for the likely factors involved and the elements that are most likely to be most relevant. Extended to a given clinical practice, problematic incidents, as they are encountered, can be identified by the practitioner and subjected to similar analysis. Such a procedure can also be routinely applied in the context of a psychological test use audit as a periodic check on effectiveness of test use within a practice. Where problematic incidents may occur, peer consultation and review can then be secured in order to avoid such incidents in the future but also to prevent other potentially detrimental incidents from occurring. Deliberate attention to the factors and to their implications for establishing test use guidelines within a given practice should help ensure the effectiveness of the professional in a psychological or neuropsychological assessment role.

2.26.3 IMPLICATIONS OF THE RUST STATEMENT REVISED

It is interesting that the empirically derived factors identified in the JCTP study was anticipated surreptitiously on the basis of the *a priori* deliberations that led to a consensual statement from a committee representing the various divisions and regions of the American Association for Counseling and Development (AACD, 1978). This document was revised subsequently by the Association for Measurements in Counseling and Education (1991). The 1991 statement begins with a consideration of the kinds of decisions made in the counseling situation on the basis, at least in part, of test data. The RUST guidelines emphasize that responsible test use requires a definition of the purposes of testing in determining the information to be obtained and to be interpreted in order to achieve those purposes.

The statement then goes on to address the qualifications of test users, the necessary information for the selection of tests, the conditions favoring the proper test interpretation, and the effective communication of test results by means of interpretative reports to clients, professionals, and other involved parties as appropriate to the counseling situation. Since there is considerable overlap between the issues identified as important to test use in the counseling situation, as compared with the clinical situation, the guidelines may apply to clinical assessments as well. However, the emphasis may vary somewhat in the clinical situation where the purposes may be significantly different, largely as a function of the kinds of referral issues that arise in these respective settings. In the counseling situation, the emphasis is on test use in educational and vocational contexts, although clinical test methods and purposes may also become relevant. In the clinical situation, test use focuses more on referral issues that arise in the mental health, general medical, neurological, and rehabilitation contexts. The remainder of this chapter is an attempt to provide an *a priori* synopsis of the RUST guidelines to be considered, along with the JCTP case book. Such an overview of these documents will hopefully facilitate efforts on the part of clinical practitioners to appreciate and apply the current as well as the anticipated revision of the Standards for Educational and Psychological Testing expected to be published in 1998.

2.26.3.1 Test Decisions

The RUST revision emphasizes the decision issue since the counseling process characteristically is driven by purposes of placement or selection, and the prediction of educational, vocational, some clinical outcomes, and possible negative consequences of decisions. Such purposes are less commonly encountered in the clinical context where description, diagnosis, and the assessment of behavioral change over time as a function of intervention predominate as the more characteristic purposes of testing. Referrals for psychological assessment have traditionally emphasized the establishment of intellectual level with the secondary objectives of providing inferences regarding the individual's ability to participate in psychological and psychiatric interventions, where mental health settings are concerned, and for participating in pre- and postmedical and surgical interventions in nonpsychiatric settings. Similarly, personality assessment is frequently introduced in the generic clinical context as part of a broader strategy to arrive at specific psychiatric diagnoses, treatment plans, and discharge-related dispositional decisions. Clinical neuropsychological assessment constitutes an extension of the use of psychological tests beyond the generic clinical context since the referral issues that define the purposes of the testing require more extensive test batteries and related sophistication in the derivation of valid inferences. The referral issues range from the identification of organic brain change due to a neurogenic process to the differentiation of functional psychiatric from neuropathic and pathophysiologic factors underlying psychological dysfunction. Where organic factors have been implicated, the referral questions may then be extended to prompt the derivation of inferences regarding the probable specific underlying etiological factors and their relative contributions to the pattern of spared and impaired neuropsychological functions that have been identified by means of the test data. The neuropsychological test data are then considered within the broader psychosocial, medical, educational, and vocational data set available at the time of the assessment. Along with etiological considerations, the referral questions may include the need for comment on the likely nature, extent, location, and predicted course of the inferred neurological involvement. Depending upon the ultimate course of the neurological disease process in such instances, the neuropsychological assessment may address the need and a rationale for cognitive and other intervention, including a plan for evaluating the outcome of such interventions and for predicting the long-range course of the individual's psychological and neuropsychological functioning under various assumptions regarding alternatives. Since the tests utilized in generic clinical psychological assessment are characteristically included in a neuropsychological assessment, it is not surprising that competence in neuropsychological assessment requires education and supervised training that includes but also extends well beyond the assessment competence obtained in a traditional, generic clinical predoctoral curriculum (Division 40 Task Force, 1987).

Although clinical decisions are rarely made solely on the basis of the psychological interpretations, such interpretations may weigh heavily within the assessment and intervention framework in the decision process. Even more importantly, psychological test interpretations may play a pivotal role in judicial decisions affecting individuals in both civil and criminal actions.

2.26.3.2 Qualifications of Test Users

It is generally assumed that all psychologists, even at a masters degree level, have had formal course work and supervised training in psychological testing. Programs may vary widely in terms of expectations of graduate students in this domain so that there may not be much uniformity in training opportunities across programs. Even with relatively extensive training, the most expert individuals may not have sufficient knowledge and sets of skills to utilize all instruments available within their practice domain and certainly not for application to all situations in which they may perform test services. The responsibility for judging the relevance and appropriateness of personal, professional qualifications rests fundamentally with the test users as guided by their understanding of professional ethics, personal strengths and limitations, and awareness of the expectations for professional psychologists as proposed by state and national agencies. The latter include credentialing agencies such as state licensure boards, guidelines provided by peers through individual and organizational supportive statements or documents, and commitments to continuing education and advanced study not dictated by statute. When the test users judge that a given assessment challenge exceeds their training and previously supervised experience, they are expected to seek peer consultation and/or initiate referral to another more qualified psychologist. Even technically sound testing, when introduced to address purposes that are not fully understood or that require more expertise for effective application, may result in serious test misuse. Such qualifications include a thorough knowledge of the tests being used, including the selection of particular tests for addressing specific referral questions.

2.26.3.3 Test Selection

After considering the purposes of testing and the conduct of a self-analysis of the necessary knowledge and competence to address the purposes, the professional psychologist is then faced with the selection of tests. This requires a deliberate analysis, based on the test literature, of the most appropriate and valid procedures to be used, any necessary accommodations to meet the needs of individuals who are not represented in the normative population for the test(s) being selected, the evidence favoring psychological constructs underlying the test scores for those tests that are ultimately selected, and the intended use of test interpretation. Thus the appropriateness of the tests selected is a function of the available validity evidence for the use of the test with individuals of the population on which it has been normed, the qualifications of the test user and the likelihood that inferences derived from the test can be stated in probabilistic terms based on the validity evidence for the test. The technical properties of the test should be established and accessible from the test manual and other sources that reflect the technical expectations as achieved by consensus of measurement professionals and as exemplified by the standards for educational and psychological testing and by the assorted codes of fair testing practices.

The selector of tests, therefore, is guided by knowledge of the documented evidence of test score reliability, the adequacy of the norms, the clarity and comprehensibility of the instructions, the test qualities for the population at issue, and the clarity of the scoring criteria and their reliable application. Where individuals of different cultural, racial, ethnic, education, and vocation backgrounds are being compared, the selection of tests should be guided by knowledge of the published norms and any limitations of the norms for application of the test to nonrepresentative individuals and groups. Similarly, bilingualism and language dominance should influence test selection where appropriate. Anyone who has worked in clinical settings will attest to the fact that certain instruments are favored, often (but not always) realistically based on the breadth and depth of the literature pertaining to the use of the test. This also seems to hold true for test batteries in clinical neuropsychological assessment where standard batteries are frequently used. As is the case with widely used individual tests in mental health applications, neuropsychological batteries in widest use tend to have fairly extensive supporting literature. However, as the controversy regarding the use of fixed vs. flexible test batteries continues, there is a growing awareness of the need to justify the particular tests to be used on the basis of relevance for the purposes of testing and the soundness of their psychometric properties for obtaining an adequate data set for drawing valid interpretations.

2.26.3.4 Test Administration

The primary quality issues surrounding test administration involve the necessary environmental circumstances for the encouragement of optimal performances during testing, the importance of administering tests in accord with the standardized instructions and procedures established by the publisher, and the necessity for monitoring examiner as well as examinee behavior to ensure the maintenance of rapport and effective communication. It is essential to inform the examinee with regard to the purpose of the testing and the ultimate use for which the evaluation is being done. Where the professional psychologist is delegating test administration to a trained psychometrist or technician, the responsibility for administration of the tests remains with the supervising psychologist who is also responsible for determining the qualifications of the individual who will actually administer the tests, including previous training and supervised experience, preferably provided under the auspices of an accredited educational institution or training hospital. Where computerized test administration may be concerned, available guidelines should be followed (Green et al. 1984). Whatever the mode of administration, it is necessary to provide a testing environment that is free from distractions and from unpleasant or uncomfortable environmental conditions. Observational data relating to environmental, general health, cognitive or neurological limitations, and emotional variations or deviations should be obtained since such variables may reduce the validity of inferences that will ultimately be derived from the test data. Observations should be systematically recorded and attached to the test score protocol. Interpersonal skills of the test administrator should be applied in a manner that facilitates the performance, within the boundaries set by the test instructions, of the examinee. This consideration underscores the necessity for considerable training and supervised experience of the psychometrist or technician as well as the supervising psychologist.

2.26.3.5 Test Scoring

Minimization of measurement error requires that scoring procedures are accurate and executed in accord with the scoring criteria that have been established for tests in question. This calls for rigorous training of anyone to whom the scoring function may be delegated and the periodic introduction of auditing procedures to establish the reliability of the scorers in a given test use situation. Audit also implies that scored test responses be screened to recognize improbable or grossly distorted (qualitative or quantitative) features of the test data. Such review can also be considered a check on the veracity of the examinee in addition to being part of the test scoring routine. The accuracy of computed raw scores and their conversion to standard scores is fundamental to any immediate or later application.

2.26.3.6 Test Score Interpretation

Test interpretation is a function of the psychologist responsible for deriving inferences from the test results for the purpose of addressing the referral issues that prompted the assessment request. The relevance and meaning of the scores depends upon the interpreter's knowledge of the validity evidence for deriving inferences relating to the construct or constructs underlying the test scores. In turn, the interpretability of the data in light of such knowledge depends upon the reliability of the test scores.

Even when all the assumptions underlying the effective application of a given test are met, such as adequate qualification of the test user, effective test administration, accurate test scoring, and the data are considered sufficiently reliable for subsequent interpretation, it is necessary to acknowledge that all test scores have some degree of error and that no test outcome is likely to be exactly replicable across testings of the same individual or of groups of individuals such as those represented in the available norms for the test. Acknowledgment of error of measurement is essential for moderating the derivation of inferences in test interpretation in a realistic and contextually appropriate manner. Both scorer and test reliabilities need to be considered in the analysis of subtest score differences that are commonly at issue, in general clinical applications, and in neuropsychological assessment.

Utilization of the validity evidence for the test or test battery requires knowledge of the test literature relating to the constructs for the intended use of the test. This places a large burden on the qualifications of the test user as test interpreter and, therefore, requires explicit recognition on the part of the test user of any personal limitations of knowledge and educational/experiential background for the intended use of the test(s) within a particular referral context. The array of inferences derived from tests and test batteries in clinical practice readily become relatively convoluted and complex to a degree that cannot always be justified on the basis of the available literature. Attempts to differentiate tentative inferences (as hypotheses) from firm inferences (as probabilistic conclusions) then become critical considerations in test interpretation. Such distinctions require carefully articulated qualifications in the communication of test interpretations.

It is now widely accepted that the concept of validity is reflected in the inferences derived from tests and is not inherent to the tests *per se* (American Psychological Association, 1985). Test scores are interpreted differently depending upon the purpose for which the test is being applied and the nature and extent of the construct validity evidenced for these inferences within the body of literature that provides the basis for the intended use of the test for a given purpose.

In clinical and clinical neuropsychological applications, tests are used for drawing diagnostic and dispositional inferences, predicting outcomes of various interventions, identifying psychological processes most affected by a psychogenic or a neurogenic condition, monitoring the course of subsequent development after the occurrence of a psychogenic or neurogenic event or process, and for designing interventions that are likely to modify favorably the course of a functional psychiatric or neuropathological process. The many technical and professional issues to be considered in test interpretation are addressed in the SEPT (APA, 1985) that are being revised for publication in 1998.

Accordingly, valid test interpretation depends on the many components of test usage in the formal psychological assessment process. As the ultimate outcome of test usage, communication of the meaning of the test results provides the interface between the professional psychologist and all other users of the information yielded by the assessment process.

2.26.3.7 Communication of Test Results and Interpretations

Tests results and interpretations are ordinarily communicated by means of formal written reports and sometimes by informal oral reports as well. The latter may occur when providing immediate feedback to the referring professional but may also occur upon request either as an incidental or a formal aspect of the referral where the individual being assessed is concerned. Confidentiality of the examinees is characteristically protected by statute so that release of test information must be obtained formally from the examinees or their legal representatives. The consent to release information requires that the examinees be sufficiently informed to make a reasoned judgment with respect to the potential consequences of releasing this information. If examinees are not sufficiently intact to be able to make judgments about the utilization of this information in their best interest, yet remains legally competent, it is the responsibility of the test user to enlist family members and appropriate representatives to facilitate the obtaining of informed consent to be examined and to distribute information regarding the assessment to other individuals.

Assuming the competence to interpret the test results, the professional psychologist is careful to explain concepts imbedded in the inferences derived from test scores, either in the written interpretation or by other reporting means. Inferences are routinely adjusted to incorporate known relationships between various demographic variables and test behaviors. The relative contribution of such variables should be reflected in the inferences or conclusions. All individuals involved in these transactions should be alerted to any limitations of the tests and the assessment as a whole, especially the significance of error of measurement and the potential sources of such errors. Most importantly, where alternative explanations carry even minimal likelihood of being valid, such explanations should be given due consideration in the course of the narrative and weighed in accord with the validity evidence for the test or test battery score/pattern inferences.

The examinee and related interested parties have the right to obtain copies of the materials generated by the assessment process and to challenge the validity of the process itself. It is beyond the purview of this chapter to explore this issue. It arises frequently in the forensic context where considerations relating to the responsible use of tests is of paramount importance and highly subject to formal adversarial challenges.

2.26.4 CONCLUSION

This chapter attempts to identify some developments that have a bearing on the effective and responsible use of psychological and neuropsychological tests in clinical practice. It is intended to orient the professional psychologist to a growing literature that reflects the collective judgment of many of the organizations involved in the development of guidelines for responsible psychological test use. The 1998 version of the Standards for Educational and Psychological Testing is expected to expand significantly on earlier versions and, thereby, to favor even more deliberate application of test standards to all major types of psychological evaluations. The empirical establishment of instances of test misuse using the critical incident technique and the *a priori* deliberation of various organizations has reinforced further the need for the continuously progressive amendation of previous guidelines and standards for psychological test use.

2.26.5 REFERENCES

American Association for Counseling and Development (1978). *Responsibilities of users of standardized tests*. Alexandria, VA: American Association for Counseling and Development.

American Psychological Association (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

Association for Measurements in Counseling and Education (1991). *Responsibilities of users of standardized tests*. Alexandria, VA: Association for Measurements in Counseling and Education.

Division 40 Task Force (1987). Reports of the task force on education, accreditation and credentialing. *The Clinical Neuropsychologist*, *1*, 22–29.

Eyde, L. D., Robertson, G. J., Krug, S. E., Moreland, K. L., Robertson, A. G., Shewan, C. M., Harrison, P. L., Porch, B. E., Hammer, A. L., & Primoff, E. S. (1993). *Responsible test use: Case studies for assessing human behavior*. Washington, DC: American Psychological Association.

Green, B. F., Bock, R. D., Humphries, L. G., Linn, R. L., & Reckase, M. D. (1984). Technical guidelines for assessing computerized adaptive tests. *Journal of Educational Measurement*, *21*(4), 347–360.

Joint Committee on Testing Practices (1988). *Code of fair testing practices in education*. Washington, DC: American Psychological Association.

Matarazzo, J. (1990). Psychological assessment versus psychological testing: Validation from Binet to the school, clinic, and courtroom. *American Psychologist*, *45*(9), 999–1017.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.27
Critical Paths, Practice Guidelines, and Empirically-validated Procedures

PETER E. NATHAN
University of Iowa, Iowa City, IA, USA

2.27.1 INTRODUCTION 313
2.27.2 THE EVOLUTION OF EMPIRICALLY-VALIDATED TREATMENTS 313
2.27.3 PRACTICE GUIDELINES 316
2.27.4 TRAINING IN EMPIRICALLY-VALIDATED TREATMENTS 318
2.27.5 MANUALS FOR EMPIRICALLY-VALIDATED TREATMENTS 319
2.27.6 FUTURE DIRECTIONS 320
2.27.7 SUMMARY 321
2.27.8 REFERENCES 321

2.27.1 INTRODUCTION

The chapter traces the evolution of research on psychosocial treatment outcomes from the 1950s when Hans Eysenck published *The effects of psychotherapy: An evaluation,* which concluded that the psychotherapies of the time were largely ineffective (Eysenck, 1952). In the 1990s, more than 40 years later, advances in outcomes research methodology and the increasing effectiveness of psychosocial treatments have combined to yield empirically-validated psychosocial treatments for a wide range of *DSM-IV* disorders. The chapter also weighs an associated development, the publication of practice guidelines by both the American Psychiatric Association (APA) and the Division of Clinical Psychology of the American Psychological Association. Practice guidelines, like empirically-validated treatments, are controversial, in part because they have the potential to circumscribe the clinical decision-making of mental health professionals. Finally, the chapter considers arguments for and against empirically-validated treatments and practice guidelines, concluding that they represent substantial opportunities for gain for both psychologists and psychiatrists and their patients.

2.27.2 THE EVOLUTION OF EMPIRICALLY-VALIDATED TREATMENTS

Hans Eysenck's 1952 review of psychotherapy effectiveness, updated and extended in 1960 and 1969 (Eysenck, 1960, 1969), was undoubtedly the most important single impetus to the psychosocial treatment research that then unfolded over the ensuing 40 years. Eysenck reached two notable conclusions in his 1952 review. Estimating the spontaneous remission rate of the neuroses, as a benchmark against which to compare the effects of psychotherapy, Eysenck determined that "roughly two-thirds of a group of neurotic patients will recover or improve to a marked extent within about two years of the onset of their illness" (1952, p. 322). On reanalyzing data from Landis (1937) and Denker (1946), Eysenck (1952) also provocatively calculated that "patients treated by means of psychoanalysis improved to the extent of 44%; patients treated eclectically improved to the extent of 64%; patients treated only custodially or by general practitioners improved to the extent of 72%. There thus appears to be an inverse correlation between recovery and psychotherapy" (1952, p. 322). Eysenck's controversial 1952 bottom line, elaborated in 1960 and 1969, was that the psychotherapies in widest use in the 1950s were largely ineffective.

While Eysenck acknowledged that the inadequate methodology of most of the treatment outcome studies of his time required him to qualify his most provocative conclusions, he did not withhold them. Instead, he insisted that his calculations pointed in only one direction: "They fail to prove that psychotherapy, Freudian or otherwise, facilitates the recovery of neurotic patients" (Eysenck, 1952, p. 323).

With a nod to Lewis Carroll's dodo bird, the lead paper at the 1973 annual meeting of the American Psychopathological Association, by Luborsky, Singer, and Luborsky (1976), is titled "Comparative studies of psychotherapies: Is it true that everyone has won and all must have prizes?" Its central premise echoes one Eysenckian theme—that more adequate research methodologies for examining, weighing, and comparing treatment efficacy must be developed before an unencumbered picture of psychotherapy effectiveness is possible. However, contrary to Eysenck and despite the limitations of their methods, these researchers reached a different conclusion about psychotherapy effectiveness:

> Our title implies what I think many of us believe—that all psychotherapies produce some benefits for some patients. What we do not know is whether there are psychotherapies which produce significantly better results and whether certain psychotherapies are especially well suited to certain patients. (Luborsky et al., 1976, p. 3)

Four years after the 1973 meeting of the American Psychopathological Association at which Luborsky and his colleagues shared their conclusions about psychotherapy effectiveness, Smith and Glass (1977) reported on a new statistical procedure called meta-analysis, designed to determine the effectiveness of psychotherapy independently of the shortcomings of the therapy outcome studies of the time. Despite the signal advantages of their new method, Smith and Glass concluded, as had Luborsky and his colleagues, that while psychotherapy is generally effective, it was not possible to differentiate among therapies in effectiveness.

Incorporating almost 400 outcome studies in their meta-analysis, Smith and Glass calculated that "the average study showed a 0.68 standard deviation superiority of the treated group over the control group" (p. 756), which convinced them that psychotherapy works. Of 10 types of therapy evaluated, the average effect size for systematic desensitization was highest, at 0.90, followed by rational-emotive therapy at 0.77, and behavior modification at 0.76. Client-centered therapy achieved an effect size of 0.63, while the effect size for psychodynamic therapy was 0.59. Despite these apparent differences in effect size, however, Smith and Glass could not distinguish among these therapies in effectiveness.

Praised for its substantial sample size and an innovative statistical methodology, Smith and Glass's meta-analysis method was also roundly criticized for the conceptual and methodological spread of the 385 studies included in the analysis, for their agreement with Luborsky et al.'s positive views on the general effects of psychotherapy despite apparently insufficient evidence, and for their unexplained failure to include some of the best-designed behavior therapy outcome studies of the early 1970s in their analysis.

Summarizing the research literature on psychotherapy outcomes published during the decades of the 1950s, 1960s, and 1970s, but emphasizing the voluminous literature of the 1970s, Rachman and Wilson's *The effects of psychological therapy,* published in 1980, reached a distinctly underwhelming assessment of the effectiveness of psychotherapy *ca* 1980:

> Since the First Edition (Rachman, 1971) was published, some slight progress has been made in the attempt to produce evidence to support the claims made on behalf of psychotherapy in general ... Nevertheless, it is our view that modest evidence now supports the claim that psychotherapy is capable of producing some beneficial changes—but the negative results still outnumber the positive findings, and both of these are exceeded by reports that are beyond interpretation. (Rachman & Wilson, 1980, p. 259)

In the first edition of the book, Rachman (1971) agreed with Eysenck's earlier judgment that psychotherapy had not been shown to be effective in the 1950s and 1960s. However, by 1980, the improved methodology of outcome research in the 1970s, coupled with advances in behavioral treatment methods, gave Rachman and Wilson reason for guarded optimism.

Also notable in the 1980 volume is Rachman and Wilson's prescient anticipation of cognitive behavior therapy's development into a valuable addition to the therapeutic armamentarium.

Two articles from a 1986 special issue of the *American Psychologist* on psychotherapy outcome research deserve brief mention. In the first of these papers, Howard, Kopta, Krause, and Orlinsky (1986) reported on a successful effort to quantify the dose–effect relationship in psychotherapy. Selecting the therapy session as the dose unit, and the percentage of patients improved as the effect, Howard and his colleagues undertook a "probit analysis" of 15 diverse therapy outcome studies. They reported that "10% to 18% of patients could be expected to have shown some improvement after the first session of psychotherapy, by eight sessions, 48% to 58% of patients would be expected to have measurably improved, (and) about 75% of patients should have shown measurable improvements by the end of six months of once-weekly psychotherapy" (Howard et al., 1986, p. 162). This quantitative determination of a dose–effect relationship for psychotherapy, previously demonstrated for a variety of psychopharmacological treatments, constituted a strong empirical voice of support for psychotherapy in the 1980s.

The title of Stiles, Shapiro, and Elliott's (1986) *American Psychologist* article, "Are all psychotherapies equivalent?," intentionally echoes Luborsky's 1973 review. Starting from the premise that the outcome data to date still "appear to support the conclusion that outcomes of diverse therapies are generally similar," Stiles and his colleagues proceed to explore some possible determinants of the paradox: "outcome equivalence contrasted with content nonequivalence" (Stiles et al., 1986, p. 165). In other words, in the 1980s, prominent clinical scientists were still debating the continuing failure of efforts to differentiate effective from ineffective treatments for specific disorders, although Rachman and Wilson (1980) had previously marshaled strong evidence pointing to the superiority of behavior therapy for a variety of common syndromes.

In 1994, the 180 000 subscribers to *Consumer Reports* were asked a series of questions about their experiences with mental health professionals, physicians, medications, and self-help groups in the largest survey of psychotherapy outcomes. A mail survey of matters of this sensitivity and complexity, a substantially undefined and essentially undiagnosed group of respondents, the absence of an untreated control group, and outcome questions that focused on generalized "improvement" rather than specific outcomes targeted to symptoms all raised concerns about this survey. Most troubling was the very low response rate to the mailed questionnaire, covered in more detail below (*Consumer Reports* Editors, 1996).

The authors of the survey portrayed its principal findings as follows:

(i) Almost half of the respondents whose emotional state was "very poor" or "fairly poor" reported significant improvement following therapy.

(ii) The longer the psychotherapy lasted, the more it helped.

(iii) Psychotherapy alone worked as well as combined psychotherapy and pharmacotherapy; while most persons who took prescribed medication found it helpful, many reported side effects.

As noted above, the study's principal shortcoming was its low response rate: 4100 respondents (slightly $>$ 2%) of the 180 000 subscribers to the magazine reported seeking professional help or joining groups, while only 2900 (ca. 1.5%) reported actually consulting a mental health professional. These numbers are so low they force us to assume that many subscribers with treatment experience decided not to respond to the survey and to hypothesize that more subscribers who felt good about psychotherapy chose to complete the questionnaire than those who did not, thereby skewing the findings in a positive direction. In the absence of data on the universe of *Consumer Reports* subscribers who sought professional help for their emotional problems, this serious design problem cannot be dismissed (*Consumer Reports* Editors, 1996).

While Martin Seligman, a consultant to the project, acknowledged the gap between its methodology and that of other contemporary psychotherapy outcome studies in a 1995 critique of the study (Seligman, 1995), he nonetheless concluded that the *Consumer Reports* survey "complements the (more traditional) efficacy method, and—the best features of these two methods can be combined into a more ideal method that will best provide empirical validation of psychotherapy" (Seligman, 1995, p. 965). Seligman urged his readers to appreciate the difference between efficacy studies—the traditional "gold standard" for judging psychotherapy outcomes—and effectiveness studies, feasibility and clinical utility in the real world, as epitomized by the *Consumer Reports* survey (*Consumer Reports* Editors, 1996).

If efficacy studies show us only part of psychotherapy as a whole, effectiveness studies show us another part. And while it seems likely that the two together reveal more about psychotherapy outcomes than either does separately, we have yet to figure out how to integrate the two. That lesson will only be learned over time and with data from additional studies, which must address the serious methodological problems that prevent us from according the *Consumer Reports* study more credence.

Treatments that work (Nathan & Gorman, 1997) is a comprehensive survey of empirically validated treatments for the full range of *DSM-IV* disorders. Its more than 30 chapters are written largely by the clinical scientists responsible for much of the most consequential treatment outcome research completed during the 1980s and 1990s. The volume offers impressive testimony to additions made, in many instances in the 1990s, in the number of treatments empirically validated for an ever-broader range of the *DSM-IV* disorders. The book also documents gains over the same period in the methodological rigor of the outcome studies on which these judgments were based.

From a handful of somatic treatments for schizophrenia of questionable efficacy in the 1950s, all with devastating side-effects, the chapters of *Treatments that work* on pharmacological treatments for the psychoses reflect the availability in the 1990s, of safer, more effective drug treatments for the most common and disabling psychoses. Just as noteworthy is that a number of the "neuroses," the conditions for whose effective treatment Eysenck despaired in 1950s, can now be treated, with good to excellent outcomes, by pharmacotherapy, often in combination with psychosocial treatments.

The volume also confirms the even greater changes which characterize psychosocial treatments. In 1952, Eysenck could find no evidence of psychotherapy's effectiveness. Now, psychosocial treatments, many of them cognitive-behavioral, have shown clear efficacy for a range of diverse conditions. Moreover, and in many ways just as significantly, psychosocial treatments have now been linked with pharmacological treatments in combined regimens that often demonstrate greater therapeutic effectiveness than either treatment by itself.

In the mid-1990s, it is becoming easier than ever to specify which therapy delivered by which therapist is most effective for which person in which treatment setting. While much progress remains to be made in both outcome methodology and further elaboration of both efficacy and effectiveness, a great deal is already evident.

2.27.3 PRACTICE GUIDELINES

These substantial advances in both outcomes methodology and numbers of empirically-validated treatments have led to an unanticipated consequence: the controversial development of practice guidelines, whose multiple potential uses are now as hotly debated as the efficacy of psychotherapy was in the 1980s.

Pressures on treatment providers to use empirically validated treatments have grown with the identification of more and more of them in the 1990s; the rapid development of managed care (Pallack, 1995) has added to these pressures. For both reasons, several practice guidelines—all having a common goal, the detailed specification of effective treatments for a wide range of syndromes—have appeared. Guidelines developed by the two leading groups of mental health professionals, the APA and Division 12 of the American Psychological Association, are reviewed below.

Essentially identical prefaces introduce the three (APA) practice guidelines published up to 1995. They include the Practice Guidelines for: "the Treatment of Patients with Major Depressive Disorder in Adults" (APA, 1993), "the Treatment of Patients with Bipolar Disorder" (APA, 1994), and "the Treatment of Patients with Substance Use Disorders: Alcohol, Cocaine, Opioids" (APA, 1995). All three explicitly deny the intention "to be construed or to serve as a standard of medical care," because "standards of medical care are determined on the basis of all clinical data available for an individual case and are subject to change as scientific knowledge and technological advance and patterns evolve" (APA, 1993, p. v; 1994, p. iv; 1995, p. 4).

These denials of the use of these psychiatric guidelines as "standards of medical care" are key to understanding a major difference between them and the guideline developed by clinical psychology (Division 12 Task Force, 1995), discussed below. The APA practice guidelines are written to be suggestive rather than prescriptive, in part because they accord substantial weight to clinicians' judgments, not only to the empirical findings. This ambiguity in the role of practice guidelines—when are they suggestive rather than prescriptive and when, if ever, are they prescriptive rather than suggestive—remains a vexing issue, both for those who consult practice guidelines and those who develop them (Grady, 1995; Persons, 1995)

Each psychiatric guideline briefly details disease definition, epidemiology, and natural history, summarizes treatment principles and reviews treatment alternatives at greater length, and concludes with a synopsis of research directions. The coverage of treatment alternatives in these guidelines can best be described as comprehensive but telegraphic.

Persons, Thase, and Crits-Christoph (1996) have claimed that the 1993 guideline for the treatment of major depressive disorder underemphasizes the empirical data on the positive outcomes of several psychosocial treatments and overstates the value of combined psychosocial–pharmacological interventions, thereby reflecting an undervaluation of empirical findings throughout the guideline in favor of "clinical experience" and "clinical judgment."

To help understand how this situation might have come about, the following brief clarification of the use of standards of proof in the 1993 guideline should be considered:

> Definitive standards are difficult to achieve, except in narrow circumstances where multiple replicated studies and wide clinical opinion dictate certain forms of treatment. In other areas much is left to the clinical judgment and expertise of the clinician. (APA, 1993, p. 1)

This apparent willingness to accord "the clinical judgment and expertise of the clinician" equal weight with the empirical findings on treatment outcomes, distinguishes the three APA practice guidelines from the Division 12 Practice Guidelines reviewed. Was "clinical judgment" used consistently as an excuse to favor a psychopharmacological over a psychosocial treatment in the 1993 psychiatric guidelines—or their successors? While that claim cannot be either confirmed or refuted, it can be said that one excellent reason to adopt and use a consistent set of evaluative criteria, as the Division 12 Task Force did, is to avoid any appearance of guild-based bias.

As has been noted, a Task Force of Division 12 of the American Psychological Association published "Training in and dissemination of empirically-validated psychological treatments: report and recommendations" (Division 12 Task Force, 1995). Although the Task Force was constituted "to consider methods for educating clinical psychologists, third party payers, and the public about effective psychotherapies" (p. 3) rather than to develop a set of practice guidelines, its report and recommendations actually constitute a condensed, preliminary set of empirically validated psychosocial treatment guidelines.

Three categories of treatment efficacy are specified: well-established treatments; probably efficacious treatments; and experimental treatments (treatments that have not yet been established as at least probably efficacious). Tables 1 and 2 reproduce the criteria by which the Task Force made these discriminations.

Table 1 Criteria for empirically-validated treatments: well-established treatments.

I. At least two good group design studies, conducted by different investigators, demonstrating efficacy in one or more of the following ways:
  A. Superior to pill or psychological placebo or to another treatment.
  B. Equivalent to an already established treatment in studies with adequate statistical power.
OR
II. A large series of single case design studies demonstrating efficacy. These studies must have:
  A. Used good experimental designs and
  B. Compared the intervention to another treatment as in I.A.
Further criteria for both I and II:
III. Studies must be conducted with treatment manuals.
IV. Characteristics of the client samples must be clearly specified.
(Division 12 Task Force, 1995, p. 21)

Table 2 Criteria for empirically-validated treatments: probably efficacious treatments.

I. Two studies showing the treatment is more effective than a waiting list control group.
OR
II. Two studies otherwise meeting the well-established treatment criteria I, III, and IV, but both are conducted by the same investigator.
Or one good study demonstrating effectiveness by these same criteria.
OR
III. At least two good studies demonstrating effectiveness but flawed by heterogeneity of the client samples.
OR
IV. A small series of single-case design studies otherwise meeting the well-established treatment criteria II, III, and IV.
(Division 12 Task Force, 1995, p. 22)

Twenty-two "well-established treatments" for 21 different *DSM-IV* syndromes are identified, as are seven "probably efficacious" treatments for the same number of disorders. With the exception of family education programs for schizophrenia and interpersonal therapy for bulimia and depression, all the "well-established" treatments are behavioral. Similarly, all but the brief psychodynamic therapies listed as "probably efficacious" are behavioral.

The restricted number of empirically-validated treatments listed in the Division 12 Task Force report contrasts sharply with the larger number of diverse treatments endorsed by the

APA (1993, 1994, 1995). As noted above, differences in the standards of proof demanded by the two groups appear to account for this marked difference in number of recommended treatments.

Although the three sets of APA practice guidelines employed a rigorous rating system to reflect the methodological adequacy of the treatment outcome studies reviewed, the willingness of the drafters of these guidelines to include considerations of "clinical practice" and "expert clinical judgment" provided latitude for the inclusion of treatments for which consistently positive findings had not been reported. By contrast, the Division 12 practice guideline was shaped by very specific criteria. While these criteria are not overly demanding, they nonetheless require that studies supporting the efficacy of a treatment meet specific methodological criteria that many other therapy outcome studies have been unable to meet.

Why do behavioral approaches to treatment figure so prominently in these two lists of empirically-validated treatments? One reason is that, as Nathan and Gorman (1997) and others (e.g, Craighead, Kazdin, & Mahoney 1994; Goldfried & Davison, 1994) have suggested, an impressive body of outcome research shows that behavioral treatments can be effective for quite a number of disorders. However, these treatments are also designed to induce the discrete behavioral changes that are the target of many treatment outcome measures; they lend themselves especially well to manualized treatment programs; they are products of the behavioral movement's long history of respect for experimental validation and they tend to induce desired changes within a brief enough duration to fit into the typical, time-limited outcome study. A continuing issue will be whether treatments drawn from traditions not so amenable to accepted outcome research will ultimately be empirically validated when the methodology for doing so has been developed—or whether behavioral treatments will continue to predominate because they achieve positive outcomes other treatments do not.

Prominent psychotherapy researchers, including Garfield (1996), have criticized practice guidelines for being premature by implying greater knowledge of the variables that produce positive change than the state of our knowledge warrants. However, Nathan (1996) believes sufficient progress has been made in developing advanced outcomes methodologies and effective treatments to justify taking a chance and redoubling the efforts made to firmly base clinical practices on the best available empirical findings. Terence Wilson has written with eloquence on behalf of the same point:

> The development and implementation of empirically-validated treatments is seen by many mental health professionals as not only desirable but ethically imperative. It has been commonplace for some time now to emphasize patients' right to treatment and their right to refuse treatment. But patients should (also) have a right to safe and effective treatment. (Wilson, 1995, p. 163)

2.27.4 TRAINING IN EMPIRICALLY-VALIDATED TREATMENTS

The authors of the two sets of practice guidelines write of the importance of integrating training in empirically-validated treatments into training programs for family physicians and mental health professionals. However, a report by the Task Force on Promotion and Dissemination of Psychological Procedures (Crits-Christoph, Frank, Chambless, Brody, & Karp, 1995) revealed that clinical psychology training programs vary from 0 to 96% in the number of empirically-validated psychotherapeutic treatments they teach. Moreover, as Wilson (1995) laments, "underscoring the lack of attention to empirically-validated treatments, the report revealed that more than one-fifth of the programs did not teach anything about 75% or more of the treatment methods listed by the Task Force" (p. 165). Although comparable data for training programs in psychiatry and social work are not available, there is no reason to believe they differ in their degree of commitment to training in empirically-validated treatments.

Beyond the compelling arguments in favor of training in the use of empirically validated treatments from both clinical and ethical perspectives, proponents of doing so also point to an economic justification. In the managed care era of the 1990s, mental health practices that do not have strong empirical support—and do not yield the best possible outcomes—are unlikely to continue to be reimbursed (Barlow, 1994, Broskowski, 1995). Empirical validation of psychological treatments is especially important, Barlow (1994) has observed, since both the public and federal and state policy-makers mistakenly believe that few psychosocial treatments are efficacious. However, as the review of the evolution of empirically-validated treatments in this chapter strongly suggests, many psychological treatments both alone and in combination with drugs are now effective for a wide variety of conditions. Only when these empirical findings are more widely disseminated will patients gain fullest possible access to these treatments.

Wilson (1995) and Persons (1995), supporters of efforts to establish and promote empirically-validated treatments, have assembled some of the reasons empirically validated treatments have not been more widely embraced by mental health professionals. One of the most important factors is the failure of programs training mental health professionals to emphasize empirically-validated therapeutic methods and, conversely, the tendency of these programs to continue to provide training in therapies that lack empirical evidence of efficacy. These authors also cite the paucity of mental health professionals familiar with the empirical data on therapeutic outcomes, the persistent view among many clinicians that all psychotherapies are more or less effective, and the fact that most consumers of mental health services, like most mental health professionals, are not well informed about effective treatments, in further explanation of the relative neglect of empirically-validated treatments.

Since the early 1990s, more scientist-clinicians have begun to acknowledge the contributions empirically-based therapy practice guidelines can make to clinical practice. Clinton, McCormick, and Besteman (1994), Ogles, Lambert, and Sawyer (1995), Schooler (1994), and Schulberg and Rush (1994), among others, have described the diverse benefits therapy practice guidelines could bring. With bulimia nervosa, the focus of much of his clinical research, as an example, Wilson (1995) makes an especially strong case for the widest possible dissemination of empirically validated treatments:

> Moving aggressively in the direction of developing and implementing empirically-validated treatment methods would seem imperative in securing the place of psychological therapy in future health care policy and planning. . . . The development and implementation of empirically-validated treatments is seen by many mental health professionals as not only desirable but ethically imperative. It has been commonplace for some time now to emphasize patients' right to treatment and their right to refuse treatment. But patients should (also) have a right to safe and effective treatment. (Wilson, 1995, p. 163)

2.27.5 MANUALS FOR EMPIRICALLY-VALIDATED TREATMENTS

The decision of the drafters of the Division 12 practice guidelines (Division 12 Task Force, 1995) to include among the criteria for "well-established treatments" that the outcome studies demonstrating their efficacy used treatment manuals recognizes the importance of ensuring that the treatment is delivered as intended in these studies. There is also a growing sentiment among authorities on psychosocial treatment (e.g. Wilson, 1995) that treatment manuals have an important place in every clinician's clinical armamentarium, because they ensure that the treatment provided to the clinician's patients conforms to the standards of the empirically-validated treatment.

For this reason, annotated lists of manuals for empirically-validated treatments have begun to appear. One of the first (Sanderson & Woody, 1995) was a project of the Task Force on Psychological Interventions, a companion group to the Division 12 Task Force on Promotion and Dissemination of Psychological Interventions that developed the Division 12 practice guidelines. To compile their list, Sanderson, Woody, and their colleagues "wrote to leading investigators in the respective areas of treatment research, particularly those whose work formed the basis for judging a particular treatment to be efficacious" (Sanderson & Woody, 1995, p. 7).

The result was an annotated list of treatment manuals, complete with supporting references, for 17 treatments for 13 distinct syndromes: bulimia (cognitive behavioral therapy, interpersonal therapy); chronic headache (behavioral treatment); chronic pain (cognitive behavioral treatment); chronic mental illness (token economy programs); depression (cognitive therapy, interpersonal therapy); discordant couples (behavior therapy); enuresis (behavioral treatment); generalized anxiety disorder (anxiety management, cognitive behavior therapy); obsessive compulsive disorder (behavioral treatment); panic disorder (cognitive therapy); post traumatic stress disorder (cognitive behavioral treatment); social phobia (cognitive behavioral group therapy); and specific phobia (systematic desensitization, exposure therapy).

While the correspondence between this list of treatment manuals and the interventions meeting the Division 12 Task Force criteria for "well-established treatments" is predictably high, the two lists are not identical, although both largely include behavioral and cognitive behavioral interventions. A book has also been published (Van Hasselt & Hersen, 1996) which contains 17 detailed treatment manuals, including many of those identified by Sanderson and Woody (1995).

As noted above, Garfield's criticism of efforts to publicize empirically validated treatments (Garfield 1996) includes concerns about the growing preference for the use of treatment manuals in outcome studies. While Garfield recognizes that the inclusion of treatment manuals in outcome research is designed to buttress the integrity of the therapy being evaluated, he believes that requiring manuals to be employed raises a potentially important problem of external validity:

> The issue raised here is: Are the research findings secured with the use of training manuals directly relevant for the clinical practice of psychotherapy as performed by most practitioners? Whereas the psychotherapists in the research setting are trained and monitored to conform to a specific manual, this clearly is not the case with the average or modal practitioner . . . Consequently, research results based on the use of therapy manuals may have somewhat limited direct relevance for clinical practice. (Garfield, 1996, p. 220)

Garfield also cites Kazdin (1991) who laments the incompleteness of many treatment manuals as well as their failure to reflect the complexity of the therapeutic exchange between patient and therapist, as well as Dobson and Shaw (1988), who ask whether the use of treatment manuals renders research on therapist variables and therapeutic process more difficult.

Wilson (1995) takes issue with these criticisms of the use of treatment manuals both in psychotherapy research and in clinical settings. His advocacy for substantially more widespread use of treatment manuals derives from his conviction that empirically-validated treatments have been substantially undertaught, underpromoted, and underutilized; their wider promulgation, presumably, would be enhanced by an extension of the use of the treatment manuals developed for empirically-validated treatments, whose use would increase the likelihood that the treatment employed was the treatment originally assessed as valid.

2.27.6 FUTURE DIRECTIONS

Practice guidelines are understandably controversial. While the best of them reflect an empirical standard by which the adequacy of a treatment can be judged, they also impose constraints on the decision making of clinicians. Moreover, because they presumably reflect best practices, they will be of interest to managed-care practice monitors, just as they are to clinicians; some managed-care monitors will doubtless use them to assess the quality of practice of individual clinicians, with distinct economic consequences and contrary to the expressed intentions of their developers.

As noted above, concerns have also been raised that considerations other than empirical findings influence the drafters of guidelines. As time goes on and the ultimately corrosive potential of this temptation becomes even more clear, we can hope it will recede as a problem. One solution would be the development of a protocol to guide creators of practice guidelines by setting down a consensus statement of general principles applicable to all syndromes and all treatments. A Task Force of the American Psychological Association has proposed a "Template for developing guidelines: interventions for mental disorders and psychosocial aspects of physical disorders" (American Psychological Association Task Force, 1995). To a substantially greater extent than either the Division 12 or APA practice guidelines, the document explicates key aspects of the processes by which standards of proof ought to be selected and utilized. When further developed, refined, and evaluated, the template might well enable the developers of practice guidelines across the mental health professions to agree on important fundamentals. Above all, such an agreement would lead to establishment of a common set of standards of proof, reducing or eliminating unnecessary and confusing differences among guidelines.

Developed "to assure comprehensiveness and consistency" of practice guidelines, the template has two features especially relevant to this discussion. The first is designed to ensure that the efficacy of a treatment is not judged solely from the results of acceptable outcome studies alone, as important as they are; the nature of the treatment comparisons in those studies are to be considered in this judgment as well. Thus, the drafters of the Division 12 practice guidelines emphasized the adequacy of research designs and the quality of outcomes methodology in reaching their conclusions on efficacious treatments. Adding the nature of the outcome comparison—specifically to no treatment, nonspecific variables in treatment, or treatments known to be effective—provides an additional powerful evaluative dimension. A randomized clinical trial that compares an experimental treatment with one or more established, effective treatments, clearly represents a more powerful test than a randomized clinical trial that compares the experimental treatment to nonrobust comparison treatments or no treatment at all. Clinical scientists have known and generally followed this strategy when possible; making it explicit for consumers of treatment research as well as developers of practice guidelines is the next step.

A second notable feature of the template is the following distinction it draws between the efficacy of an intervention and its effectiveness or clinical utility; the same distinction was also drawn by Seligman (1995) in evaluating the results of the *Consumer Reports* (*Consumer Reports* Editors, 1996) survey:

> clinical practice guidelines for behavioral health care (should) be constructed on the basis of two simultaneous considerations or "axes." The first is that guidelines take into consideration a rigorous assessment of scientific evidence with the goal of measuring the efficacy of any given intervention (efficacy). The second axis specifies that guidelines consider the applicability and feasibility of the intervention in the local setting where it is to be proffered (clinical utility). (American Psychological Association Task Force, 1995, p. i)

While a series of well-designed studies might establish the efficacy of an intervention, unless it is effective in real-life clinical settings, it will have little utility.

> The clinical utility axis refers to the ability (and willingness) of practitioners to use, and of patients to accept, the treatment in question, and to the range of applicability of that treatment. It reflects the extent to which the intervention, regardless of the efficacy that may or may not have been demonstrated in the clinical research setting, will be effective in the practice setting in which it is to be applied. Among the factors that will be considered under this rubric are the generalizability of administering the intervention in various settings, the feasibility of the intervention across patients and settings, and the costs and benefits associated with the administration of the intervention. (American Psychological Association Task Force, 1995, p. 13)

Up to 1997, the methodology for efficacy studies has been substantially more advanced than that for effectiveness studies. While occasional studies of the generalizability of a treatment have been undertaken (e.g. Finney & Moos, 1997), the methodology for undertaking feasibility and costs and benefits studies requires substantial further development. An important future direction, then, involves the development of research strategies that will undertake effectiveness studies complementary to efficacy studies of the same treatments for the same syndromes.

2.27.7 SUMMARY

The history of systematic efforts to identify empirically-validated treatments for mental disorders is a short one. Only in the 1950s, coincident with Eysenck's influential article, "The effects of psychotherapy: an evaluation" (Eysenck, 1952), was treatment research first subjected to rigorous methodological scrutiny in the effort to distinguish effective from ineffective mental health treatments. Since then, research designs increasingly capable of discriminating efficacious psychosocial and pharmacological treatments have been developed. With Eysenck's 1952 article as a starting point, this chapter reviews several important efforts to distinguish empirically-validated treatments over the past 50 years.

Practice guidelines, reflecting more potent outcomes methodologies and more effective treatments, have been proposed by both the APA and the Division of Clinical Psychology of the American Psychological Association. The guidelines are controversial, however, both because they are new and untested, and because they have the potential to encroach upon the clinical decision-making role that has traditionally been solely the treating clinician's. Differences in the standards of proof on which the psychiatric and psychological practice guidelines rest are explored; they appear to stem from differences in the methodological criteria used to judge outcome studies as well as the degree to which clinical experience and judgment are accorded emphasis.

This review of the current status of empirically-validated treatments, practice guidelines, and standards of proof enables us to anticipate a series of desirable future directions for the continuation of efforts to promote best therapeutic practices among mental health professionals.

2.27.8 REFERENCES

American Psychiatric Association (1993). Practice guideline for the treatment of major depressive disorder in adults. *American Psychiatric Association, 150* (4 Suppl.), 1–26.

American Psychiatric Association (1994). Practice guideline for the treatment of patients with bipolar disorder. *American Journal of Psychiatry, 151* (12 Suppl.), 1–36.

American Psychiatric Association (1995). Practice guideline for the treatment of patients with substance use disorders: alcohol, cocaine, opioids. *American Journal of Psychiatry, 152* (11 Suppl.), 1–59.

American Psychological Association Task Force (1995). *Template for developing guidelines: Interventions for mental disorders and psychosocial aspects of physical disorders*. Washington, DC: Author.

Barlow, D. H. (1994). Psychological intervention in the era of managed competition. *Clinical Psychology: Science and Practice, 1,* 109–122.

Broskowski, A. T. (1995). The evolution of health care: Implications for the training and careers of psychologists. *Professional Psychology: Research and Practice, 26,* 156–162.

Clinton, J. J., McCormick, K., & Besteman, J. (1994). Enhancing clinical practice: The role of practice guidelines. *American Psychologist, 49,* 30–33.

Consumer Reports Editors (1996). Mental health: Does therapy help? *Consumer Reports, 60,* 734–737.

Craighead, L. W., Craighead, W. E., Kazdin, A. E., & Mahoney, M. J. (1994). *Cognitive and behavioral interventions: An empirical approach to mental health problems*. Boston: Allyn & Bacon.

Crits-Christoph, P., Frank, E., Chambless, D. L., Brody, C., & Karp., J. F. (1995). Training in empirically validated treatments: What are clinical psychology students learning? *Professional Psychology: Research and Practice, 26,* 514–522.

Denker, P. (1946). Results of treatment of psychoneuroses by the G.P. *New York State Journal of Medicine, 46,* 2164–2166.

Division 12 Task Force (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist, 48,* 3–23.

Dobson, K. S., & Shaw, B. F. (1988). The use of treatment manuals in cognitive therapy: Experience and issues. *Journal of Consulting and Clinical Psychology, 56,* 673–680.

Eysenck, H. J. (1952) The effects of psychotherapy: an evaluation. *Journal of Consulting Psychology, 16,* 319–324.

Eysenck, H. J. (1960). *Behavior therapy and the neuroses*. Oxford, UK: Pergamon.

Eysenck, H. J. (1969). *The effects of psychotherapy*. New York: Science House.

Finney, J., & Moos, R. (1997). Treatments for alcohol dependence. In P. E. Nathan & J. M. Gorman (Eds.), *Treatments that work*. New York: Oxford University Press.

Garfield, S. L. (1996). Some problems associated with "validated" forms of psychotherapy. *Clinical Psychology: Science and Practice, 3,* 218–229.

Goldfried, M. R., & Davison, G. C. (1994). *Clinical behavior therapy*. New York: Wiley.

Grady, K. E. (1995). Compliance with standards of care: Evidence from medical research. In S. C. Hayes, V. M. Follette, R. M. Dawes, & K. E. Grady (Eds.), *Scientific standards of psychological practice: Issues and recommendations* (pp. 83–91). Reno, NV: Context.

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–effect relationship in psychotherapy. *American Psychologist, 41,* 159–164.

Kazdin, A. E. (1991). Treatment research: The investigation and evaluation of psychotherapy. In M. Hersen, A. E. Kazdin, & A. S. Bellack (Eds.), *The clinical psychology handbook* (2nd ed., pp. 293–312) New York: Pergamon.

Landis, C. (1937). A statistical evaluation of psychotherapeutic methods. In L. E. Hinsie (Ed.), *Concepts and problems in psychotherapy*. New York: Columbia University Press.

Luborsky, L., Singer, B., & Luborsky, L. (1976). Comparative studies of psychotherapies: Is it true that "everybody has won and all must have prizes"? In R. L. Spitzer & D. F. Klein (Eds.), *Evaluation of psychological therapies* (pp. 3–22). Baltimore: Johns Hopkins University Press.

Nathan, P. E. (1996). Validated forms of psychotherapy may lead to better-validated psychotherapy. *Clinical Psychology: Science and Practice, 3,* 250–254.

Nathan, P. E., & Gorman, J. M. (1997). *Treatments that work*. New York: Oxford University Press.

Ogles, B. M., Lambert, M. J., & Sawyer, J. D. (1995). Clinical significance of the National Institute of Mental Health treatment of depression collaborative research program data. *Journal of Consulting and Clinical Psychology, 63,* 321–326.

Pallack, M. S. (1995). Managed care and outcomes-based standards in the health care revolution. In S. C. Hayes, V. M. Follette, R. M. Dawes, & K. E. Grady (Eds.), *Scientific standards of psychological practice: Issues and recommendations* (pp. 73–77). Reno, NV: Context.

Persons, J. B. (1995). Why practicing psychologists are slow to adopt empirically-validated treatments. In S. C. Hayes, V. M. Follette, R. M. Dawes, & K. E. Grady (Eds.), *Scientific standards of psychological practice: Issues and recommendations* (pp. 141–157). Reno, NV: Context.

Persons, J. B., Thase, M. E., & Crits-Christoph, P. (1996). The role of psychotherapy in the treatment of depression. *Archives of General Psychiatry, 53,* 283–290.

Rachman, S. J. (1971). *The effects of psychological therapy*. Oxford, UK: Pergamon.

Rachman, S. J., & Wilson, G. T. (1980). *The effects of psychological therapy* (2nd enlarged ed.). Oxford, UK: Pergamon.

Sanderson, W. C., & Woody, S. (1995). Manuals for empirically validated treatments. *The Clinical Psychologist, 48,* 7–11.

Schooler, N. R. (1994). Translating treatment research findings into clinical action. *American Journal of Psychiatry, 151,* 1719–1721.

Schulberg, H. C., & Rush, A. J. (1994). Clinical practice guidelines for managing major depression in primary care practice: Implications for psychologists. *American Psychologist, 49,* 34–41.

Seligman, M. E. P. (1995). The effectiveness of psychotherapy: the *Consumer Reports* study. *American Psychologist, 50,* 965–974.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32,* 752–760.

Stiles, W. B., Shapiro, D. A., & Elliott, R. (1986). "Are all psychotherapies equivalent?" *American Psychologist, 41,* 165–180.

Van Hasselt, V. B., & Hersen, M. (Eds.) (1996). *Sourcebook of psychological treatment manuals for adult disorders*. New York: Plenum.

Wilson, G. T. (1995). Empirically validated treatments as a basis for clinical practice: Problems and prospects. In S. C. Hayes, V. M. Follette, R. M. Dawes, & K. E. Grady (Eds.), *Scientific standards of psychological practice: Issues and recommendations* (pp. 163–196). Reno, NV: Context.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.28

Evidence-based Practice in Clinical Psychology

WILLIAM J. KELLEHER
Nova Southeastern University, Fort Lauderdale, FL, USA

2.28.1 INTRODUCTION 323
2.28.1.1 Definition and Components of "Evidence-based Practice" 323
2.28.1.2 Growth of the Evidence-based Practice Model in Other Disciplines 324
2.28.1.3 Organization and Objectives 324
2.28.2 CURRENT FACTORS ENCOURAGING EVIDENCE-BASED PRACTICE 325
2.28.2.1 Major Changes in Funding Sources of Health Care 325
2.28.2.2 Professional Ethics and Commitment to Quality Patient Care 326
2.28.2.3 Empirical Support of Specific Treatment Protocols for Certain Clinical Problems 326
2.28.3 BARRIERS AGAINST EVIDENCE-BASED PRACTICE 327
2.28.3.1 Previous Fee for Service Funding System and Professional Inertia 327
2.28.3.2 Literature Suggesting an Equivalence of All Psychotherapy 327
2.28.3.3 The Gap Between Practitioners and Researchers 328
2.28.3.4 Limited Dissemination and Training of Empirically Supported Therapy 328
2.28.3.5 Proliferation and Nonstandardization of Therapeutic Outcome Measures 329
2.28.3.6 Growth of Multidisciplinary and Paraprofessional Care Settings 329
2.28.3.7 Inconsistent Advocacy of Evidence-based Practice for Some Therapy Orientations 330
2.28.4 INNOVATIONS FOR POTENTIAL EVIDENCE-BASED PRACTICE SUPPORT 331
2.28.4.1 Advances in Computerized Database Technology 331
2.28.4.2 Growth of Provider Networks for Increased Collaboration and Peer Review 332
2.28.4.3 New Applied Research Design and Statistical Techniques 332
2.28.4.4 Development of Clinical Practice Guidelines 333
2.28.4.5 Innovations for Computerized On-line Information Exchange 333
2.28.5 FUTURE DIRECTIONS AND RECOMMENDATIONS 333
2.28.5.1 National Psychological Associations 333
2.28.5.2 Clinical Psychology Providers and Researchers 334
2.28.5.3 Clinical Psychology Doctoral Training Programs 334
2.28.6 SUMMARY 335
2.28.7 REFERENCES 335

2.28.1 INTRODUCTION

2.28.1.1 Definition and Components of "Evidence-based Practice"

> Understanding certain rules of evidence is necessary to correctly interpret literature on causation, prognosis, diagnostic tests, and treatment strategy. It follows that clinicians should regularly consult the original literature (and be able to critically appraise the methods and results sections) in solving clinical problems and providing optimal patient care.

At the same time, systematic attempts to record observations in a reproducible and unbiased

> fashion markedly increase the confidence one can have in knowledge about patient prognosis, the value of diagnostic tests, and the efficacy of treatment. In the absence of systematic observation one must be cautious in the interpretation of information derived from clinical experience and intuition, for it may be misleading. (p. 2421)

Interestingly, the above principles are not taken from a clinical psychology treatise detailing the scientist–practitioner model or from a text in applied clinical psychology research design. Although it appears they just as well could have been from the nature of their content. Rather, they were published in the Journal of the American Medical Association (Evidence-based Medicine Working Group, 1992) as basic assumptions for what is refered to as the new paradigm of "Evidence-based medicine." This medical endorsement of the "evidence-based" concept for the practice of health care delivery seems especially noteworthy and timely as increased focus is being placed today on the relevance of the scientist–practitioner model in Clinical Psychology (Montgomery & Aylon, 1995; O'Sullivan & Quevillon, 1992). What is being touted as new underpinings for the delivery of care by physicians has been advocated as the scientist–practitioner model for the training and practice of clinical psychologists since its pronouncement at the 1949 Boulder Conference (Raimy, 1950). This model's advocacy for the integration of empirical research and clinical psychology practice has been more fully articulated and affirmed in the 1992 National Conference on Scientist–Practitioner Education and Training for the Professional Practice of Psychology (Belar & Perry, 1992). The policy statement from that conference contains principles which are very similar to those detailed above for the model of "evidence-based medicine" and extend these further into specific recommendations for psychology doctoral training programs and the conduct of applied clinical research. What can be taken as a summary statement of that policy statement and the overall scientist–practitioner model follows: "Scientist–practitioner psychologists reflect a research orientation in their practice and a practice relevance in their research" (Belar & Perry, 1992, p. 74).

The concept of evidenced-based practice focuses on the clinical care delivery aspect of the scientist–practitioner model and has two primary components that will be examined in this chapter. In the evidence-based practice of clinical psychology, psychologists' commitment to their profession and to their patients demands an equal commitment to: (i) utilizing evidence-based intervention strategies and protocols that have demonstrated empirical support for the specific clinical problems for which their patients are seeking care, and (ii) employing an ongoing evidence-based patient progress tracking system with empirically supported outcome measures to assess treatment effectiveness. In adhering to both of these components, psychologists will be furthering the growth of their patients and, at the same time, the overall growth of empirical evidence for the practice of psychological intervention.

2.28.1.2 Growth of the Evidence-based Practice Model in Other Disciplines

These two components of evidence-based practice in clinical psychology are very consistent with the assumptions for the medicine counterpart quoted above. The opportunity for evidence-based practice is not unique for one discipline; better that it can be applied across health care disciplines in this age of multi-discipline providers' increasing involvement in the care of the same patient. Within the mental health field, social work has had advocates of an evidence-based model, which has at times been referred to as empirical clinical practice (Jayaratne & Levy, 1979, Witkin, 1991). There have been references in the psychiatric literature of a movement toward evidence-based psychiatry (Goldner & Bilsker, 1995), a concept that might have seemed quite improbable until the 1990s. What explains the growth since the late 1980s in these disciplines of the evidence-based practice concept which clinical psychology has been encouraging within its scientist–practitioner model for years? Why is it in the interest of our patients and our profession that as clinical psychologists we do not decrease our long standing advocacy of evidence-based practice but rather work toward its more consistent and widespread application throughout our discipline? These are two of the major questions to be addressed in this chapter.

2.28.1.3 Organization and Objectives

This review focuses on both components of evidence-based practice in clinical psychology as described above in terms of:

(i) current and projected future key factors encouraging use of the evidence-based practice model;

(ii) historical and current barriers to widespread adoption of the evidence-based practice model;

(iii) professional and technological innovations which have potential for expanding the application of evidence-based practice in mental health care delivery;

(iv) future directions forecasted and recommended for continued development of evidence-based practice; and

(v) a summary of the major supported findings and conclusions.

2.28.2 CURRENT FACTORS ENCOURAGING EVIDENCE-BASED PRACTICE

2.28.2.1 Major Changes in Funding Sources of Health Care

Since the mid-1980s the costs of health care in the United States have been increasing annually four percentage points or more above the rate of inflation (Burlingame, Lambert, Reisinger, Neff, & Mosier, 1995). Over the same period of time the amount spent for mental health care has been increasing at an astoundingly greater rate, with some estimates as high as 30–40% yearly (Giles, 1993). Broskowski (1991) attributes the increase in mental health care utilization to such factors as the growth in numbers of mental health care providers, decreasing stigma of mental health care with popularization in the media, and increased incidence of substance abuse problems. In response to these rapidly and seriously rising costs, private insurance and public agency funding sources for general health and mental health care have made very significant changes in their organization and mechanisms for continuing care coverage.

In the private sector, we have seen the dramatic development and revolutionary growth of managed health care to include employee assistance programs, Health Maintenance Organizations (HMOs), and Preferred Provider Organizations (PPOs). These are the third party payers' answers for dealing with the escalating health and mental health care costs (Foos, Ottens, & Hill, 1991). To illustrate the prevalence of these new forms of care delivery, Montgomery (1988) refers to a 1987 survey which indicated that already 80% of Fortune 500 companies had employee assistance programs for crisis intervention and short-term therapy. Because of this rise in managed care organizations, it has been predicted that by the year 2000 over half the private mental health practitioners will be out of their solo practice and only 5% of the physicians will be practicing independently (Cummings, 1987). Perhaps more to the point of our potential patients' source of funding for their care, it is estimated that in 1992 70% of the US population were receiving their health care by an HMO or other managed care system (Foos et al., 1991).

Given these facts and projected trends, it is necessary for clinical psychologists to be fully aware of the important parameters and to have the professional tools and skills for successful practice in the managed care environment. Much has been documented about the managed care emphases on cost and time efficiency in health care delivery (Giles & Marafiote, 1994; Mirin & Namerow, 1991), professional accountability (Benbenishty & Oyserman, 1995; Ludwigsen & Enright, 1988), utilizing metrics and outcome measures to demonstrate effective treatment (Giles, 1991; Linder, 1991), and preauthorization of a specified treatment plan and duration (Foos et al., 1991). Psychologists practicing from an evidence-based approach are well suited for these requirements as they will offer a detailed and planned treatment protocol which has empirical support and are prepared to work with the patient to monitor their progress regularly with objective measures. The evidence-based concentration on providing the treatment which is known to be successful for what the patient desires is also likely to satisfy the managed care agency's concern for customer satisfaction and keeps the provider very competitive for future referrals. Being as efficient as possible in the time limited therapy format, which tends to characterize empirically supported protocols, allows the evidence-based provider to provide relative savings in cost and time for the managed care agency as well as the patient. This savings potential takes on added significance in the capitation budget system of managed care which will only pay the provider a fixed amount per covered life per month (Giles, 1993).

Since the mid-1980s the public sector has seen dramatic and ongoing cuts in governmental budgets for support of mental health care. This has led to considerable downsizing and mental health providers having "to do more with less." As we have seen, this reduction in funding is occurring during a time of increased requests for services. Similarly, as we have seen for desired evidence-based practice responses to the managed care environment, using empirically supported interventions and system design can hopefully help psychologists and other mental health providers work more efficiently. Providing successful treatment with a time limited course of therapy can provide reinforcement of providers' quality work. This may also indirectly work to offset reportedly increasing community mental health center staff dissatisfaction and stress (Ebben, Bliss, & Perlman, 1991).

2.28.2.2 Professional Ethics and Commitment to Quality Patient Care

As part of their commitment to our patients, psychologists owe it to them to offer and be skilled in those treatments already empirically proven to provide the greatest therapeutic gain in the shortest amount of time. This is a central position of the evidence-based practice model and it touches on what is the ethical thing for psychologists to do when there are known therapy protocols with empirical support for the type of problem their patient is experiencing. Yeaton and Sechrest (1981), refering to what strength of therapy to recommend to a patient, advocate using what they call the "weakest that works" rule of thumb. Eysenck (1994) puts this in perspective perhaps more succinctly, "Where superior, cheaper, and quicker-acting methods of therapy are available, the continued use of expensive, lengthy, and noneffective methods is unethical and contrary to the best traditions of medicine" (p. 491). This scenario applies when there are known intervention strategies that work. As Elliott, Stiles, and Shapiro (1993) point out, when the research evidence is equivocal or not supportive of a particular treatment after controlled study, it is ethical to share that information with a patient who is seeking help for a related problem. Ignoring the positive empirical results supporting a particular treatment should be avoided just as much as making an unjustified blanket recommendation for an unsupported treatment.

As part of their ethical commitment to their patients, psychologists also owe it to them to follow another core principle of evidence-based practice, namely to be objectively tracking the degree of progress they are or are not making in response to treatment (Nelson, 1981). They then know whether to continue, modify, or discontinue treatment procedures and accordingly to do what they can to maintain a high quality of patient service. With the consistent use of outcome measures as the evidence-based practice model recommends, they will then be in the position to also fulfill a commitment to the profession and future potential patients through contributing to the applied clinical psychology research base. This is the scientist–practitioner at work and relates to what David Shakow (1976) believes is a central answer to the question posed in the title of his address to Division 12 of the American Psychological Association, namely, "What is Clinical Psychology?" Another way of capturing this aspect of evidence-based practice was expressed by Benbenishty and Oyserman (1995), "The challenge, therefore, is to collect clinically relevant data systematically, process and analyze it in ways that will inform practice, and to fold these insights back into the daily practice" (p. 312).

2.28.2.3 Empirical Support of Specific Treatment Protocols for Certain Clinical Problems

The scientist–practitioner model has generated numerous controlled studies (including attention/placebo and interorientation comparisons), meta-analyses, case studies, and review articles documenting the effectiveness of certain psychological treatments (Andrews, 1993; Giles, Neims, & Prial, 1993). The American Psychological Association created a task force to examine the research literature to report on the state of empirically supported treatments and to make any recommendations for the further advancement of successful therapy protocols. The findings of this task force have been summarized in a document referred to as the "Chambless Report" (based on the name of the task force chairperson) and it documents treatments found to be effective for many clinical problems, such as anxiety disorders, depression, tension and migraine headaches, marital discord, childhood enuresis, bulimia, tics, and obsessive–compulsive behavior (Task Force on Promotion and Dissemination of Psychological Procedures, 1995).

These empirically proven treatment protocols now make up the core of the current evidence-based practice treatment offerings and without their tested efficacy, psychotherapeutic interventions for their related conditions could not be presented with legitimate and ethical confidence. Many of these have well-developed and detailed protocols including manuals, patient information and workbook materials, monitoring forms, recommended outcome measures, prescribed sequences of therapeutic tasks, and projected program schedules. These supporting materials make it easier for the evidence-based practice provider to be trained in and to duplicate the effective program. However, this report and the treatments found to be effective and worthy of increased application have met with criticism and a lack of widespread distribution (Montgomery & Aylon, 1995). Giles et al. (1993) provide responses to criticisms which have been offered regarding the purported irrelevance of the meta-analyses, methodological deficiencies, and experimenter bias. These are critiques not unique to this body of literature but ones which some believe are actually managed with adequate experimental controls in many of these studies.

2.28.3 BARRIERS AGAINST EVIDENCE-BASED PRACTICE

2.28.3.1 Previous Fee for Service Funding System and Professional Inertia

Historically mental health care has been funded by third party payers according to a fee for service basis (Foos et al., 1991). This had encouraged the practice of treating patients for as long and as intensively as possible, the opposite of the evidence-based practice weakest that works rule of thumb presented earlier. At times, perhaps, this lengthy and intense therapy schedule format was motivated by the therapist's goal to guide the patient to a fully self-actualized state of optimum functioning or in some sense a cure. There is virtually no existing data that any form of therapy has been effective in helping patients achieve such an end-state and that such a therapeutic goal should be the target for intervention. In addition, unfortunately, there has been a lack of financial incentives to change this style of clinical practice but rather strong incentives to keep it going (Cummings, 1987). Bruce Thyer (1995) highlights that not only did this incentive system encourage more lengthy treatments than was necessary but also less concern about tracking progress and adjusting treatment accordingly; irrespective of outcome payment was the same. He observed further,

> Once one has graduated with a professional degree and obtained a state license to practice as a mental health services provider, there are few if any strictures as to what constitutes suitable treatment. Quite literally almost anything goes. (p. 93)

The two components of employing empirically proven protocols and monitoring ongoing measures of patient progress which we have presented as central to the evidence-based practice model were clearly not being encouraged by these fiscal contingencies. With no external pressure or patient outcome data to motivate change, providers were likely to stay with what was familiar to them in their treatment strategies. Again, this is quite inconsistent with evidence-based practice and choosing a therapy protocol not because it feels comfortable for the therapist but because it has been proven to be a right one for the patient. This calls for the therapist to be much more flexible, be open to new and more effective techniques, and to go where the data supports to go.

Unfortunately, the increasing costs generated by this tendency of what Giles (1993) calls the "overprescription of services" contributed to what we have already presented as the managed care revolution in health care delivery. Furthermore, with the discovery of unnecessary or self-serving tests and care being ordered, third party payers developed an increasing loss of faith in health care professionals' judgment, competence, and ethics (Ludwigsen & Enright, 1988). To illustrate some of the consequences of this distrust, we can look at a statement of a medical professional looking in at the mental health care system. John Montgomery (1988), a physician representing Aetna Life & Casualty writes,

> Employers also should be aware of the policies or practices of different treatment facilities. For example, some institutions continue to believe that patients with demonstrated psychological problems usually require a minimum of one full year of treatment. (p. 89)

It appears reasonable to assume that this physician would have a much more positive and trusting response to an evidence-based practice psychologist who was oriented toward offering the treatment that had the best chance of producing successful results for a particular patient in the shortest period of time rather than demanding an excessive minimum amount of time for all patients. Burlingame et al. (1995) propose that given the low ebb in professional trust, it is more critical than ever for providers to document clinical results at the end of treatment as well as by ongoing monitoring of patient progress. This is clearly consistent with the evidence-based practice model.

2.28.3.2 Literature Suggesting an Equivalence of All Psychotherapy

There has been a body of literature which purports an equivalence of all psychotherapeutic approaches to any clinical problem (Bergin & Lambert, 1978; Frank, 1973; Garfield, 1976). Eysenck (1952) wrote one of the first review papers of this type, concluding that psychotherapy has not been proven to have any efficacy. Factors such as the nonspecific placebo effect and spontaneous remission have been offered in defense of Eysenck's and these later reviewers' position. This equivalency conclusion was overlooking the later discovered statistical power and sample size problem and other methodological concerns with interorientation comparisons (Kazdin, 1994). These published conclusions from reviews of the psychotherapy literature could easily have added to the professional inertia described above that experienced providers stay with the therapy techniques for which they had training and felt familiar; after all, it appeared that no other approach would be any more effective for any kind of patient problem.

However, as it can be seen, psychologists now have data to support the efficacy of particular therapy protocols for many different kinds of clinical problems. Even Eysenck (1993, 1994) has modified his conclusions from reviews of the latest literature to note that behavior therapy is effective with a number of conditions. Yet the trend appears to be continuing for therapists to choose empirically undocumented or even contraindicated therapies rather than the empirically proven ones (Montgomery & Ayllon, 1995). Unfortunately, even managed care companies can operate on this outdated equivalency premise and, as a result, fund whoever will provide the equivalent intervention service cheaper, including paraprofessionals with limited training and/or supervision (Giles & Marafiote, 1994). Another misinterpretation possible for managed care system designers from this equivalency assumption was that all mental health patients and problems were equivalent and, therefore, could be successfully managed with the same maximum number of sessions. Given this history and current state of misattributed restrictions on funding, the evidence-base practitioners have their work cut out for them in educating and modifying the managed care community as well as the providers who are still unaware of the advances in empirically supported treatments.

2.28.3.3 The Gap Between Practitioners and Researchers

Clinical researchers (including those demonstrating empirical success with specific treatment protocols) have been criticized for designing their studies or writing their reports without taking into account the realistic clinic caseloads or language to which providers can easily relate (Montgomery & Ayllon, 1995). This contributes to practitioners not consulting or relying on research results for guidance on how to practice (Forsyth & Strong, 1986; Hoshmand & Polkinghorne, 1992). Kordy (1995) presents the conflict between researchers and providers as an interactive one, "Critics blame psychotherapy research for asking the questions which no practitioner is interested in because he/she already knows the answer or doesn't want to know it" (p. 128). Some surveys (Norcross, Prochaska, & Gallagher, 1989) suggest that few providers rely on empirical evidence to guide their practice. Barlow (1981) wrote, "Clinical research has very little or no influence on clinical practice" (p. 147). Needless to say, these observations illustrate the professional habits of a large number of psychologists which were antithetical to the evidence-based practice model. Some (Frank, 1984) went so far as to say that the researcher and the clinician were fundamentally different people and implied that never the twain shall meet. In a related tone, Elliot et al. (1993) refer to the relationship between those who do research and those involved with clinical practice as a dysfunctional one and question whether given this rift any accommodation is possible.

To challenge the notion of separateness and incompatibility of research and practice, Steinmeyer (1994) suggests that scientist–practitioner activity is a much more appropriate term than implying the model is the mark of the person involved. Persons (1993) perceives the reasons for the scientist–practitioner gap to be quite understandable given the different reinforcement contingencies operating on each. Belar and Perry (1992) present in the National Conference of Scientist–Practitioner Education and Training policy statement, "The scientist–practitioner is, by definition, also committed to bridging the gap between scientific foundations and professional practice" (p. 72). They propose that the hyphen between "scientist" and "practitioner" be replaced in future writings by a "~" or an "x" to symbolize integration which they believe is very possible and should be stimulated from the very beginning of psychology doctoral training programs. Related to this, there is further hope of this scientist–practitioner integration generated by the results of such surveys as that of O'Sullivan and Quevillon (1992) which indicate that 98% of doctoral program directors report their programs have always followed, and plan to continue to follow, the Boulder scientist–practitioner training model. The growth of evidence-based practice requires that these graduating psychologists carry this model out into their first clinical positions and beyond. Hopefully, the innovations in technology, applied research methodology, and provider organization, which are reviewed in the next section of this chapter, will help to overcome the barriers discussed in this section and further this growth for self-sustaining progress toward more efficient and effective patient care into the twenty-first century.

2.28.3.4 Limited Dissemination and Training of Empirically Supported Therapy

Many of the more empirically supported treatments are relatively new and have not had widespread dissemination (Andrews, 1993). Many providers, therefore, are not trained in these protocols and have had limited opportunities to pursue this training. This seems to be at least a contributing factor to these effective therapies "not being utilized by even a minority of practitioners" (Giles, 1993) in addition to the barriers for evidence-based practice we have already reviewed. To be more specific, training in these new techniques has had limited availability on a psychology graduate school level as well as in postgraduate or continuing professional education courses. Thyer (1995) comments that snake oil Continuing Education Units are rampant and that, given the limited availability of continuing education credits, the more nonempirically supported treatments are endorsed by professional organizations for Continuing Education credit, the less there will be for the evidence-based therapies.

Professional mental health organizations have been slow to provide official support and impetus for more widespread adoption of the evidence-based practice model and empirically supported treatments. Accreditation standards for academic training and internships tend to omit reference to empirically based treatment. Workshops and training programs in treatments not empirically supported continue to receive support for continuing education credit. Licensing boards are also not addressing the importance of evidence-based practice in their examination and review of professional credentials. All of these organizational behaviors contribute to the contingencies and reinforcement system in which psychology providers serve and as such play a role in the delayed advancement of evidence-based practice.

2.28.3.5 Proliferation and Nonstandardization of Therapeutic Outcome Measures

There has been a proliferation of instruments purported to measure psychological treatment outcome with little effort made to standardize a set of the more valid and reliable metrics for use with particular conditions (Mirin & Namerow, 1991). One review article count totaled over 1400 different measures which have been utilized, with many of these having been designed and employed in a single study (Lambert & Hill, 1994; Lambert, Ogles, & Masters, 1992). This very large number and lack of standardized selection could be contributing to the frequently observed clinical situation of choosing an outcome measure that will not significantly add to the evidence base or, worse yet, will not accurately track the patient's progress. Reportedly, 33% of all US health care institutions are not active in ongoing outcomes research in spite of the staggering number of measures that have been utilized by others (Burlingame et al., 1995).

In addition to this proliferation issue, other factors have been hypothesized as barriers to providers consistently employing outcome measures of their treatment effectiveness. The cost of resources to collect, score, and report the results of these measures in some settings is seen as prohibitive. Some purport that assessing psychotherapy-related change is extremely complex and difficult (Martinovich, Saunders, & Howard, 1996). For example, the selection of measures could well include a variety of change areas and be completed by the patient or significant others. There is widespread disagreement about what types of measures deserve the highest priority with a variety of conceptual schemes and a combination of measures proposed (Elliott, 1992; Lambert et al., 1992; Ryle, 1995). Psychologists committed to evidence-based practice must therefore keep their focus while navigating this maze of metrics with an educated eye for the most relevant, valid, and reliable measures that realistically lend themselves for completion by the patients on a repeating schedule.

In summarizing the current state of affairs in outcome measure selection and the future expansion of the evidence-based practice model, Lambert and Hill (1994) conclude,

> We look forward to the day when researchers can agree on a minimal core outcome battery for measuring change in patients with specific disorders. We are convinced that most of the necessary measures to be included in such batteries already exist.

Indeed, consensus on outcome measures would advance evidence-based practice and the further investigation of which therapies offer the greater potential for patients to achieve their emotional and behavioral goals. In the mean time, it would be wise for providers to make educated choices given the best information available and begin gathering data. The third party payer agencies, as has been presented, want to see the results of such tracking and if the treating psychologists do not make the selection of outcome measures themselves the funding source is likely to fill the void and dictate which metric they expect to see employed. Compared to the third party payer, the psychologist who is functioning as a scientist–practitioner will be in a much better position to choose the instruments sensitive to the most relevant areas of patient change.

2.28.3.6 Growth of Multidisciplinary and Paraprofessional Care Settings

"Perhaps the biggest impediment to collecting useful outcome data is the complex nature of the mental health care delivery system itself"

(Mirin & Namerow, 1991, p. 1009). Since the mid-1980s, there has been a significant increase in demand for mental health care services and a resultant burgeoning of patient caseloads. Managed care has stimulated an increase in patients being treated in settings staffed by practitioners from a variety of health care disciplines and paraprofessionals as well. This type of setting has been touted as offering a better continuum of care than the more traditional single discipline office. However, as a result, it is becoming more difficult pragmatically for providers to monitor each patient's entire record closely and be able to oversee the selection, administration, analysis, and reporting of relevant outcome measures efficiently. This is further complicated by the increasingly complex administrative and fiscal tracking of patients required by their health care insurance companies. Some patients have supplemental insurance plans which need to be addressed in addition to their primary coverage. Therefore, it has become quite difficult for any one provider to organize and manage the tremendous amount of information and data accumulating for his or her patients.

At the same time in these settings, not all the providers, or the health care organization itself, will necessarily share the psychologist's training or inclination to operate within the scientist–practitioner, evidence-based model. Steinmeyer (1994) observes that a lack of differentiation in clinical position requirements among mental health professional disciplines has been developing; what the psychologist, social worker, or psychiatrist can offer is more and more being seen as interchangeable or equivalent. Furthermore, the paraprofessional is being viewed and hired as a cheaper and preferred form of labor who can provide similar or equivalent services as the more expensive psychiatrists, psychologists, and social workers (Cummings, 1987). Thus, the care organization funding the mental health professional could well not be specifically looking for a scientist–practitioner or someone with a psychologist's training and inclination for evidence-based practice. The psychologist working in such a setting will then have to rely much more on his or her own resources and contingencies to pursue and maintain a focus on providing care which is empirically supported and empirically monitored.

This trend away from recognizing the scientific background psychologists and other professionals can bring to the care organization becomes more critical as it is being found that the multidisciplinary continuum of care setting will not always, necessarily, produce effective, or even cost efficient, care. Bickman (1996) reporting on the results of the comprehensive, multidisciplinary Fort Bragg Demonstration Project for children and adolescents states,

> This study raises significant questions about the validity of several widely held beliefs. The cost-effectiveness of the continuum of care has been challenged, as has, more generally, the effectiveness of services as delivered in community settings. (p. 699)

The author further observes that the system of care delivery changes we are seeing with managed care may well improve access and customer satisfaction ratings but "is unlikely to reform clinical outcomes unless it also reforms the actual services delivered. However, there is little research to provide guidance as to what changes in services are required." Therefore, it is concluded that the need for research integrated with practice in the community care delivery setting is much more needed now than ever; unfortunately this is occurring at a time when some organizations are selecting their providers irrespective of their support and training for such evidence-based practice.

This ongoing need for applied clinical research can also be supported by the rapid development observed in brief and packaged therapies in the managed care settings (Foos et al., 1991). As it is being found that "multidisciplinary" does not always lead to more effective and efficient care, so too brief or time limited care may not necessarily result in longer term savings and success. This type of care delivery calls for as much empirical research as has been advocated for the more traditional methods of mental health care. Similarly, an increase in manualized therapy, while it can have advantages for potentially increasing the consistent replication of proven protocols, does not insure positive results and needs to be studied further. Therefore, the opportunities for the scientist–practitioner to be challenged in today's health care environment appear to be increasing but, as has been presented, so too are the potential interferences.

2.28.3.7 Inconsistent Advocacy of Evidence-based Practice for Some Therapy Orientations

Some psychotherapeutic orientations have proponents sending out mixed signals about the relevance and importance of utilizing outcome measures to track treatment effectiveness. For example, consider what is being supported and advocated for the study of psychoanalysis or psychodynamic therapy. From one perspective (Gray, 1996) it is being written that because the desired outcomes of psychoanalysis (i.e., developing a more coherent sense of self, restructuring one's personality, rewriting one's history, or resuming emotional growth) can be unique for the individual patient, "to operationalize these variables for an outcome study may at this time be beyond our technical grasp" (p. 215). This article continues that even writing daily progress notes "violates the core psychoanalytic technique of listening with even-hovering attention." Notes are therefore not likely to be available as a database for review of psychoanalysis. This position certainly does not suggest psychoanalysis and an evidence-based practice orientation are a match even though the title of this article is *Developing Practice Guidelines for Psychoanalysis.* This position, antithetical to evidence-based practice, flies in the face of the work of the Strupp group at Vanderbilt University, the Luborsky group at the University of Pennsylvania, and others who are finding ways to deal with the outcome measure problem in psychoanalysis (Vaughn, Marshall, MacKinnon, & Roose, 1994). Connolly and Strupp (1996) use cluster analysis with patients' writings of reported change to study outcome of psychoanalytic treatment. Luborsky (1995) has changed his position about how viable the integration of research and practice can be. He is now "welcoming the new emphasis on 'grass-roots' research-audit and 'low-tech' investigations aimed at monitoring service delivery and providing information of use to both providers and users of services" (p. 123). This is very close to the research drives practice and practice drives research principle that is underlying evidence-based practice. As consistently or as inconsistently a psychotherapeutic orientation's proponents advocate this position, so may go evidence-based practice for the providers espousing that orientation, whether it be psychoanalytic, behavior, existential, interpersonal, cognitive-behavioral, or other type of therapy.

2.28.4 INNOVATIONS FOR POTENTIAL EVIDENCE-BASED PRACTICE SUPPORT

2.28.4.1 Advances in Computerized Database Technology

There has been a rapid and potentially very valuable growth in the development of clinical information database management systems especially designed for mental health care delivery (Wedding, Topolski, & McGaha, 1995). The better designed of these systems can greatly stimulate the movement toward evidence-based practice through more accurate and efficient computerized processing of the data related to the patient, the provider, the assessment, the treatment, the outcome measures, and many other potentially important variables. As has been presented earlier, the amount of information accumulating and being required to be processed for today's patient in the managed care environment is becoming more than one provider or a group of providers can process, let alone continuously track, analyze, and report in a coherent manner. The computerized information management system can greatly encourage attending to the relevance of a particular treatment for a particular type of patient problem, given that the relevant data can be more easily and clearly visible to all involved. The lessons such a system's utilization allows the providers to learn about effective patient care can be fed back into the system for prescribing such empirically proven therapeutic alternatives with future patients who have similar clinical problems. "The assumption is that sound information could guide program efforts by the results obtained, rather than following practices based on convention or professional ideology" (p. 48). This seems quite in line with the scientist–practitioner concept and a natural for evidence-based practice.

In this age of multidisciplinary provider care delivery settings, the use of a well-designed clinical information management system can also lead to improved communication within the provider network. This in turn could foster more integration of the entire behavioral care community working with a patient and increase overall efficiency and effectiveness (Fessenden & Vogt, 1995). Furthermore, with the emphasis we have seen on fiscal concerns for health care, the cost savings factor of such automated systems should not be ignored. The more comprehensive clinical database package system can also track accounting information and produce billing and other financial reports; and it can do all of this with fewer administrative and accounting errors than noncomputerized systems (Brown, Fraser, & Bendoraitis, 1995). Therefore, what can be good for clinical evidence-based practice can also be very beneficial for the funding system that supports the organization where the care is provided.

Some potential problems, however, need to be addressed and considered with the utilization of a comprehensive clinical information management system. There could be a threat to protecting the patient's confidentiality given that a considerable amount of sensitive information will be residing on a computer hard drive or floppy disk. Obviously, it is critical that appropriate security measures are well-planned and in place from the point of first instaling the system. This will limit the access and processing of all personal information to only those qualified and ethically bound by their profession to preserve the patient's privacy rights. Another important consideration is how the information system is utilized by the organization's administration or personnel office. The clinical value of the system can be undermined if it is used more for managerial control (e.g., personnel utilization review or checking provider productivity) than it is for patient care outcome tracking and reinforcement of staff members' quality work (Binner, 1993).

2.28.4.2 Growth of Provider Networks for Increased Collaboration and Peer Review

As briefly highlighted previously, a variety of different types of provider networks and organizations are being established to maximize efficiency and provide patients with integrated comprehensive care when indicated. These have the potential to allow for easier professional consultation, peer review, and comparison of provider databases to encourage treatment which is responsive to patient progress or lack thereof and is in accord with empirically proven standards of care (Broskowski, 1991). There are descriptions in the literature of such agencies for providers as: psychologists and others in Biodyne (Cummings, 1987), psychiatrists in a professional affiliation group and an HMO virtual group practice (Pomerantz, Liptzin, Carter, & Perlman, 1994, 1995), and psychiatric networks in Canada (Kates, Adsett, Bellissimo, Levine, & Humphrey, 1993). However, as we have seen with the Fort Bragg Demonstration Project results (Bickman, 1996), a network or provider group organization does not by the fact alone of its formation or the breadth of the services it offers insure that effective and efficient treatment will be the results. This would need to be monitored and confirmed through planed empirical evaluation and controled comparisons with other forms of service delivery.

An apparently well-planned and organized model for this type of provider collaboration is that of the Pennsylvania Practice Research Network. This network brings together psychologists who are members of the Pennsylvania Psychological Association (PPA) for several important goals: (i) to design a core assessment battery to measure treatment outcome of all patients treated by PPA psychologists; (ii) to start research based on these assessments as well as patient demographic and diagnostic information to address questions regarding such issues as the nature of the disorders, predictors of outcome, patterns of technique practice, relationships of techniques to outcomes, and overall efficacy of therapy in Pennsylvania, and (iii) to use the network infrastructure "to create other basic and applied therapy research studies among subsets of interested Network member" (T. D. Borkovec, personal communication, May 6, 1996). This network seems to be well on the road to taking the fullest advantage of the significant potential available through a provider organization to further effective patient care. It is the scientist–practitioner model and evidence-based practice in action.

2.28.4.3 New Applied Research Design and Statistical Techniques

We have seen the development of clinical research design and statistical techniques which allow the individual provider or network of providers to contribute to the evidence base for effectiveness of psychotherapy without changing their day-to-day care delivery schedule or caseload (Hoshmand & Polkinghorne, 1992). These techniques and design procedures, if used appropriately, can help narrow the gap and foster the integration of research and practice in clinical psychology and other health care disciplines. Three of the more significant of these developments appear to be: time series methodology (Hayes, 1981), single case study design (Kazdin, 1981), and using measures of clinical significance (Jacobson, Follette, & Revenstorf, 1984; Jacobson & Truax, 1991; Tingey, Lambert, Burlingame, & Hansen, 1996). Time series methodology and case study designs allow for providers to systematically track a single subject or groups of single subjects in a controled study of their response to treatment in comparison to themselves in various conditions of therapy or in a period of no intervention. As Kazdin (1981) notes these designs are not meant to replace experimental research, they "can provide relatively clear information about the impact of treatment" and "probably will continue to contribute greatly to the information available in the field" (p. 191). Measures of clinical significance evaluate the extent to which patients who have received intervention compare with individuals who are functioning well. It could involve analyzing the degree to which the treated patients perform at, or are within, the normative range for a certain type of measure or instrument. Another measure of clinical significance could be to examine whether the individuals who received therapy have made behavioral changes of a certain magnitude and whether they are now significantly improved in their performance rates of this behavior compared to a sample known to be in a dysfunctional group. This type of measure can be seen as getting closer to the bottom line of what patients themselves and third party payers are interested in as outcome results and as such add a significantly important tool to the repertoire of the provider committed to evidence-based practice.

2.28.4.4 Development of Clinical Practice Guidelines

Mental health care disciplines are beginning to follow the lead of medicine and promote the establishment of clinical practice guidelines for particular diagnoses. In medicine, these published guidelines are recommended strategies for the management of patients with a specific clinical condition and are designed to assist the physician with clinical decision-making for choice of medication, clinical tests, prescribed patient activity, and other patient management parameters. In addition, they could indicate the range of what are considered acceptable approaches for diagnosing, managing, and possibly preventing the particular conditions. When applied to mental health care, this practice guideline concept has the potential for clearly highlighting and supporting empirically proven psychotherapeutic protocols for our own profession, referring providers in other disciplines, third party payers, and potential patients. Those programs referred to earlier which have already been found to be effective in empirical research with specific mental health problems would seem to be excellent sources for developing the content of such mental health practice guidelines.

To further illustrate how far this guidelines concept can be stretched and utilized even in the absence of objective empirical evidence consider the proposal of practice guidelines for psychoanalysis previously discussed. Gray (1996) refers to these as "consensus-based practice guidelines" which "codify and integrate existing clinical intelligence and the rich oral tradition in medicine. This makes them better suited than outcome-based parameters for the task of helping the psychiatrist integrate the biospsychosocial aspects of a case" (pp. 213–214). Gray goes on to state that developing guidelines for psychoanalysis on the basis of outcome studies would be "incomplete and difficult to implement." An evidence-based practice orientation would seem to lend itself to finding the development of practice guidelines for conditions with demonstrated treatment effectiveness not as difficult to implement.

2.28.4.5 Innovations for Computerized On-line Information Exchange

Technology has significantly improved for computerized on-line professional literature searches and for staying current with the empirical research literature. Through the use of a computer and a modem, it is no longer necessary to leave the office, travel to a library with adequate medical resources, spend hours going through card catalogs and journal reviews, locating relevant articles, and writing notes on important findings. Now this all can be done very efficiently, accurately, and comprehensively within a few minutes search through the Internet. There is data which suggests that medical students can significantly profit from specific training in how to use these on-line resources, including Medline, in support of the evidence-based medicine model (Evidence-Based Medicine Working Group, 1992). The APA sponsors such an on-line service for searching the mental health literature. Given the potential offered by this technological innovation, it could be said that the evidence-based practitioner has never had it easier.

2.28.5 FUTURE DIRECTIONS AND RECOMMENDATIONS

It appears most likely that the emphases on treatment cost and time efficiency as well as being able to demonstrate clinical effectiveness of therapy will not go away but rather become mandatory for continued mental health care funding. Therefore, the more important and critical becomes the reliance of clinical psychologists on both aspects of the evidence-based practice model, research driving practice and practice driving research. The following recommendations for the different components of the psychological care delivery structure seem worthy of serious consideration and pursuit for more quickly responding to these anticipated requirements and, at the same time, they can contribute to the further growth of evidence-based practice in mental health care. Many of these recommendations are consistent with those from the Chambless Report referred to earlier as well as from the conclusions of similar empirically supported literature reviews such as Thyer (1995).

2.28.5.1 National Psychological Associations

(i) Require training for empirically proven treatments for full accreditation of psychology graduate school and internship programs.

(ii) Publicly endorse treatments which have received empirical support. Actively promote and sponsor workshops in these techniques. Encourage members to become skilled in and utilize these protocols as treatments of choice with their patients who present with related clinical problems.

(iii) Grant continuing education credit only for those workshops offering training in treatments with an established record of empirical support.

(iv) Work with managed care organizations to help them better understand the different parameters (i.e., number and frequency of recommended treatment sessions, other assessment and treatment resources found necessary, etc.) of treatment programs required for different clinical conditions. Educate the administrators of such agencies that not all patients with all mental health conditions have been found to be helped within the same maximum number of allotted sessions.

(v) Organize and sponsor more conferences and symposia similar to the 1993 National Conference on Scientist–Practitioner Education and Training to stimulate more definition and support of the evidence-based practice model.

(vi) Include elements in the organization's code of ethical practice stressing that patients be presented with, as first choice interventions, those strategies that have received empirical support for the type of problem they are seeking help and that providers routinely monitor treatment progress with relevant and validated outcome measures.

(vii) Publish clinical practice guidelines for the conditions that the organization's members have established treatments with empirical support.

2.28.5.2 Clinical Psychology Providers and Researchers

(i) Remain current with applied clinical literature, seeking to be knowledgeable of empirically proven treatments, especially for the type of clinical problems the provider is most commonly requested to treat.

(ii) Obtain training in those empirically proven protocols for clinical problems most frequently encountered. Utilize the protocols, manual, and patient educational or monitoring materials that have already been prepared if they are available for use. This will encourage adherence to the proven techniques and efficiency of effort.

(iii) If not skilled in a proven protocol for the problem with which a patient is presenting, seriously consider referring the patient to a provider (if one is conveniently available) who is trained for such treatment. Explain the circumstances regarding the potential referral to the patient.

(iv) Wherever possible, utilize ongoing objective outcome measures and participate in a clinical information database management system to more easily track ongoing patient progress. Consider reporting illustrative results in a case study or time series format for a professional journal.

(v) Network with other psychologists for participation in a peer review process and possible collaboration in an applied clinical research project. Affiliation with a local hospital may be a mechanism for this type of association.

(vi) Adopt a mind set of therapeutic goals with shorter term sights than the notion of full self-actualization, being oriented toward helping the patient achieve a more adaptive level of functioning and emotional well-being.

(vii) As best as possible, utilize sound clinical judgment in offering treatment for those problems not as yet fully researched. Be systematic in monitoring the results of a treatment plan which has not be adequately tested for that type of clinical condition. Review the data and patient report frequently and be ready to change the direction of therapy as the results are suggesting.

(viii) Research needs to move outside the laboratory and academic settings to study services in the community and clinics where care is more routinely provided. The need has never been as great as it is now for the integration of research and practice.

2.28.5.3 Clinical Psychology Doctoral Training Programs

(i) Follow the guidelines developed at the National Conference on Scientist–Practitioner Education and Training for the Professional Practice of Psychology (Belar & Perry, 1993).

(ii) Recruit faculty who are skilled and experienced in applied clinical research, especially with experience investigating empirical supported treatments, and who support the scientist–practitioner model.

(iii) Offer course work, supervision, and practicum experiences for students to acquire and increase their skills for those treatment protocols with demonstrated treatment effectiveness. Similarly, provide instruction in treatment outcome measure selection and a variety of practical research design techniques, including those for single subject and case study.

(iv) Provide facilities and training for efficiently using the latest technologies in literature reviews and management information systems.

2.28.6 SUMMARY

Evidence-based practice in clinical psychology involves consistently utilizing empirically supported intervention techniques which are proven to be effective for a patient's presenting clinical problem as well as continuously monitoring treatment progress with validated outcome measures. It is consistent with the scientist–practitioner model which since the 1950s has been strongly advocated for clinical psychology training and practice. Given today's managed care environment with its call for cost and time efficiency, accountability, demonstrated results, and customer focus in health care delivery, the need and relevance of the evidence-based practice model have never been greater. This model is also considered very appropriate for today's psychologists to adopt given its consistency with the ethical standards of clinical psychology practice and the significant amount of studies in the professional mental health literature which document the effectiveness of specific treatment protocols for a variety of clinical conditions. These empirically supported treatments are highlighted in the APA Division 12 Task Force Report (1993) commissioned to investigate the effectiveness of psychotherapy literature.

Historically, there have been barriers to the widespread adoption of evidence-based practice and the scientist–practitioner model. These have included: the fee for service funding system, earlier review articles which suggested an equivalence of results for all psychotherapy, a significant gap between those professionals who conducted research and those who delivered patient care, a lack of dissemination and available training for the empirically proven treatments, a tendency for therapists to stay with delivering service in the orientation that felt familiar and not track its real effectiveness for each patient, an overwhelming number of available outcome measures with little standardization of a recommended selection for use, and some inconsistent advocacy of the scientist–practitioner approach within certain therapeutic orientations.

During the 1990s, advances in computer technology, clinical information management systems, and medicine's movement in the evidence-based practice direction are now supporting evidence-based practice in clinical psychology more than ever. National psychological associations, psychology doctoral programs, and individual psychologists can now take full advantage of this prime opportunity to stimulate a more widespread growth of evidence-based practice in clinical psychology by pursuing a variety of recommended actions. Through such action, our profession and our patients can only gain.

2.28.7 REFERENCES

Andrews, G. (1993). The essential psychotherapies. *British Journal of Psychiatry*, *162*, 447–451.

Barlow, D. H. (1981). On the relation of clinical research to clinical practice: Current issues, new directions. *Journal of Consulting and Clinical Psychology*, *49*, 147–155.

Belar, C. D., & Perry, N. W. (1992). National conference on scientist–practitioner education and training for the professional practice of psychology. *American Psychologist*, *47*, 71–75.

Benbenishty, R., & Oyserman, D. (1995). Integrated information systems for human services: A conceptual framework, methodology, and technology. *Computers in Human Services*, *12*, 311–325.

Bergin, A. E., & Lambert, M. J. (1978). The evaluation of therapeutic outcomes. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (2nd ed., pp. 16–37). New York: Wiley.

Bickman, L. (1996). A continuum of care: More is not always better. *American Psychologist*, *51*, 689–701.

Binner, P. R. (1993). Information systems and mental health services: Issues for the 90's. *Computers in Human Services*, 47–57.

Broskowski, A. (1991). Current mental health care environments: Why managed care is necessary. *Professional Psychology: Research and Practice*, *22*, 6–14.

Brown, G. S., Fraser, J. B., & Bendoraitis, T. M. (1995). Transforming the future: The coming impact of CIS. *Behavioral Health Management*, *15*, 8–12.

Burlingame, G. M., Lambert, M. J., Reisinger, C. W, Neff, W. M., & Mosier, J. (1995). Pragmatics of tracking mental health outcomes in a managed care setting. *The Journal of Mental Health Administration*, *22*, 226–236.

Connolly, M. B., & Strupp, H. H. (1996). Cluster analysis of patient reported psychotherapy outcomes. *Psychotherapy Research*, *6*, 30–42.

Cummings, N. A. (1987). The future of psychotherapy: One psychologist's perspective. *American Journal of Psychotherapy*, *41*, 349–360.

Ebben, P. A., Bliss, D., & Perlman, B. (1991). Brief report: Problems and issues in community mental health service delivery in the 1990s. *Community Mental Health Journal*, *27*, 225–229.

Elliott, R. (1992). A conceptual analysis of Lambert, Ogles, and Masters's conceptual scheme for outcome assessment. *Journal of Counseling & Development*, *70*, 535–537.

Elliott, R., Stiles, W. B., & Shapiro, D. A. (1993). Are some psychotherapies more equivalent than others? In T. R. Giles (Ed.), *Handbook of effective psychotherapy* (pp. 455–479). New York: Plenum.

Evidence-based Medicine Working Group (1992). Evidence-based medicine: A new approach to teaching the practice of medicine. *Journal of the American Medical Association 268*, 2420–2425.

Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology*, *16*, 319–324.

Eysenck, H. J. (1993). Forty years on: The outcome problem in psychotherapy revisited. In T. R. Giles (Ed.), *Handbook of effective psychotherapy* (pp. 3–20). New York: Plenum.

Eysenck, H. J. (1994). The outcome problem in psychotherapy: What have we learned? *Behaviour Research and Therapy*, *32*, 477–495.

Fessenden, R., & Vogt, D. (1995). Speaking the same language: A new approach to documentation. *Behavioral Health Management*, *15*, 22–27.

Foos, J. A., Ottens, A. J., & Hill, L. K. (1991). Managed mental health: A primer for counselors. *Journal of Counseling & Development*, *69*, 332–336.

Forsyth, D. R., & Strong, S. R. (1986). The scientific study of counseling and psychotherapy: A unificationist view. *American Psychologist*, *41*, 113–119.

Frank, G. (1984). The Boulder model: History, rationale, and critique. *Professional Psychology: Research & Practice*, *15*, 417–435.

Frank, J. D. (1973). *Persuasion and healing*. Baltimore, MD: Johns Hopkins University Press.

Garfield, S. L. (1976). All roads lead to Rome. *Contemporary Psychology*, *21*, 328–329.

Giles, T. R. (1991). Managed mental health care and effective psychotherapy: A step in the right direction? *Journal of Behavior Therapy and Experimental Psychiatry*, *22*, 83–86.

Giles. T. R. (1993). Consumer advocacy and effective psychotherapy: The managed care alternative. In T. R. Giles (Ed.), *Handbook of effective psychotherapy* (pp. 481–488). New York: Plenum.

Giles, T. R., & Marafiote, R. A. (1994). Managed care and psychotherapy outcome: Has the pendulum swung too far? *The Behavior Therapist*, *17*, 239–244.

Giles, T. R., Neims, D. M., & Prial, E. M. (1993). The relative efficacy of prescriptive techniques. In T. R. Giles (Ed.), *Handbook of effective psychotherapy* (pp. 21–39). New York: Plenum.

Goldner, E. M., & Bilsker, D. (1995). Evidence-based psychiatry. *Canadian Journal of Psychiatry*, *40*, 97–101.

Gray, S. H. (1996). Developing practice guidelines for psychoanalysis. *Journal of Psychotherapy Practice and Research*, *5*, 213–227.

Hayes, S. C. (1981). Single case experimental design and empirical clinical practice. *Journal of Consulting and Clinical Psychology*, *49*, 193–211.

Hoshmand, L. T., & Polkinghorne, D. E. (1992). Redefining the science–practice relationship and professional training. *American Psychologist*, *47*, 55–66.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, *15*, 336–352.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12–19.

Jayaratne, S., & Levy, R. L. (1979). *Empirical clinical practice*. New York: Columbia University Press.

Kates, N., Adsett, A., Bellissimo, T., Levine, L., & Humphrey, B. (1993). The Hamilton (McMaster) psychiatric network: The evolution of an integrated network of psychiatric services. *Canadian Journal of Psychiatry*, *38*, 315–318.

Kazdin, A. E. (1981). Drawing valid inferences from case studies. *Journal of Consulting and Clinical Psychology*, *49*, 183–192.

Kazdin, A. E. (1994). Methodology, design, and evaluation in psychotherapy research. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 19–71). New York: Wiley.

Kordy, H. (1995). Does psychotherapy research answer the questions of practitoners, and should it? *Psychotherapy Research*, *5*, 128–130.

Lambert, M. J., & Hill, C. E. (1994). Assessing psychotherapy outcomes and processes. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 72–113). New York: Wiley.

Lambert, M. J., Ogles, B. M., & Masters, K. S. (1992). Choosing outcome assessment devices: An organizational and conceptual scheme. *Journal of Counseling & Development*, *70*, 527–532.

Linder, J. C. (1991). Outcomes measurement: Compliance tool or strategic initiative? *Health Care Management Review*, *16*, 21–33.

Luborsky, L. (1995). The same and divergent views of "fashions and preoccupations in psychotherapy research." *Psychotherapy Research*, *5*, 118–120.

Ludwigsen, K. R., & Enright, M. F. (1988). The health care revolution: Implications for psychology and hospital practice. *Psychotherapy*, *25*, 424–428.

Martinovich, Z., Saunders, S., & Howard, K. I. (1996). Some comments on "assessing clinical significance". *Psychotherapy Research*, *6*, 124–132.

Mirin, S. M., & Namerow, M. J. (1991). Why study treatment outcome? *Hospital and Community Psychiatry*, *42*, 1007–1013.

Montgomery, J. S. (1988). Shrink mental health care costs. *Personnel Journal*, *67*, 86–91.

Montgomery, R. W., & Ayllon, T. (1995). Matching verbal repertoires: Understanding the contingencies of practice in order to functionally communicate with clinicians. *Journal of Behavior Therapy and Experimental Psychiatry*, *26*, 99–105.

Nelson, R. O. (1981). Realistic dependent measures for clinical use. *Journal of Consulting and Clinical Psychology*, *49*, 168–182.

Norcross, J. C., Prochaska, J., & Gallagher, K. (1989). Clinical psychologists in the 1980s: Theory, research, and practice. *Clinical Psychologist*, *12*, 42–53.

O'Sullivan, J. J., & Quevillon, R. P. (1992). 40 years later is the Boulder model still alive. *American Psychologist*, *47*, 67–70.

Persons, J. (1993). Why are there so few scientist–practitioners? A modest proposal. *The Behavior Therapist*, *16*, 34.

Pomerantz, J. M., Liptzin, B., Carter, A. H., & Perlman, M. S. (1994) The professional affiliation group: A new model for managed mental health care. *Hospital and Community Psychiatry*, *45*, 308–310.

Pomerantz, J. M., Liptzin, B., Carter, A., & Perlman, M. S. (1995). Development and management of a "virtual" group practice: Behavioral clinicians and organizations linked by a capitation contract with an HMO. *Psychiatric Annals*, *25*, 504–508.

Raimy, V. C. (Ed.) (1950). *Training in clinical psychology*. New York: Prentice-Hall.

Ryle, A. (1995). Fashions and preoccupations in psychotherapy research. *Psychotherapy Research*, *5*, 113–117.

Shakow, D. (1976). What is clinical psychology? *American Psychologist*, *31*, 553–560.

Steinmeyer, C. H. (1994). Hypotheses regarding determinants of scientist–practitioner behavior. *The Behavior Therapist*, *17*, 12–13.

Task Force on Promotion and Dissemination of Psychological Procedures (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist*, *48*, 3–23.

Thyer, B. A. (1995). Promoting an empiricist agenda within the human services: An ethical and humanistic imperative. *Journal of Behavior Therapy and Experimental Psychiatry*, *26*, 93–98.

Tingey, R. C., Lambert, M. J., Burlingame, G. M., & Hansen, N. B. (1996). Assessing clinical significance: Proposed extensions to method. *Psychotherapy Research*, *6*, 109–123.

Vaughn, S. C., Marshall, R. D., MacKinnon, R. A. & Roose, S. P. (1994). Current psychotherapy research methodology applied to psychoanalysis: A feasibility study. *Journal of Psychotherapy Practice and Research*, *3*, 334–340.

Wedding, D., Topolski, J., & McGaha, A. (1995). Maintaining the confidentiality of computerized mental health outcome data. *The Journal of Mental Health Administration*, *22*, 237–244.

Witkin, S. L. (1991). Empirical clinical practice: A critical appraisal. *Social Work*, *36*, 158–163.

Yeaton, W. H., & Sechrest, L. (1981). Critical dimensions in the choice of maintenance of successful treatments: Strength, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology*, *49*, 156–167.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.29

Principal Practice Settings for Clinical Psychologists in European Countries

HARRY VAN DER VLUGT
Tilburg University, The Netherlands

2.29.1 INTRODUCTION 339

2.29.2 THE NETHERLANDS 340
  2.29.2.1 General Historical Background 340
  2.29.2.2 Employment 340
  2.29.2.3 Employment Settings for Clinical Psychologists 341
    2.29.2.3.1 Employment opportunities 342
    2.29.2.3.2 Private practice 342
  2.29.2.4 Some Final Statements 342

2.29.3 COUNTRY-RELATED INFORMATION 343
  2.29.3.1 Belgium 344
    2.29.3.1.1 Work setting 344
    2.29.3.1.2 Future 344
  2.29.3.2 Czech Republic 344
  2.29.3.3 France 345
    2.29.3.3.1 Postgraduate training 345
  2.29.3.4 Germany 346
    2.29.3.4.1 Aspects of psychological training 346
  2.29.3.5 Greece 347
  2.29.3.6 Italy 348
  2.29.3.7 Portugal 348
  2.29.3.8 Spain 349
  2.29.3.9 Switzerland 350
    2.29.3.9.1 Clinical psychology in Switzerland 351

2.29.4 CONCLUDING REMARKS 351

2.29.5 REFERENCES 352

2.29.1 INTRODUCTION

In writing this chapter, several unexpected problems arose. The main problem was how to define a clinical psychologist. Apparently the answer should be a psychologist who works in a clinical setting. But this does not solve the problem because:

(i) Some clinical psychologists work as psychotherapists. In many countries psychotherapists are registered in a separate register including all kinds of psychotherapists with different backgrounds such as social work, theology, psychology, or psychiatry. In other (mostly southern European) countries psychotherapy is considered to be a curative activity which is the privilege of the medical profession.

(ii) Not only psychologists who are officially trained as clinical psychologists work as such in the health system. Many other psychologists work in a clinical setting, for example, child psychologists, psychogerontologists, rehabilitation psychologists, neuropsychologists, health psychologists, and even "orthopedagogues" who hold positions comparable with child psychologists.

(iii) Not all clinical psychologists are members of their national psychological society. The percentage of psychologists who are members of the local societies varies strongly from country to country.

(iv) Due to continuous changes with respect to health services in different countries, the educational systems at the university level are changing as well.

(v) Although there is a tendency within the EC to come to a general ruling with regard to quality of training and a common deontology, so far the differences between countries are that large that globalization is still a pious wish.

Because of the above mentioned problems, the situation in The Netherlands is first highlighted. This is relatively straightforward due to two excellent reviews (Van Drunen, 1995; Brunenberg, Neifmeiger, & Hutschemarker, 1995). From there on other countries in Europe are compared. This comparison is often limited because it is impossible to get the necessary basic information.

2.29.2 THE NETHERLANDS

2.29.2.1 General Historical Background

In 1992 the first century of scientific psychology in The Netherlands was celebrated. The year 1892, when Gerard Heymans established the first psychological laboratory in Groningen, is generally considered the year of its birth. However, it was not until 1921 that psychology became a possible specialization for students of philosophy. At the same time, applied psychology started, especially in the areas of personnel selection and vocational guidance. In 1924, Philips was the first private company to hire a full-time professional psychologist. In 1938, a professional organization of psychologists was founded: the *Nederlandsch Instituut van Practizeerende Psychologen* (Dutch Association of Practising Psychologists, NIPP).

Up to 1940 the NIPP only had 28 members, of whom seven had a medical background. In 1941, by academic statute, psychology was recognized as an independent discipline, and after the war all universities started a full program in psychology. In the early 1950s there were about 100 first-year students in psychology, in the early 1960s more than 250, in the 1970s more than 1000 and in the late 1990s about 2250. This led to a rapid increase in the number of psychologists from 80 in 1945 to about 700 in the early 1960s, about 2500 in the early 1970s, 10 000 in the mid 1980s, and at least 20 000 in the late 1990s; at least 1200 psychologists for every 1 000 000 inhabitants.

Another interesting aspect is the increase of women in psychology. In the mid 1970s only 35% of the graduates were women. In the late 1990s more than 70% of the graduates are women. In the late 1990s the number of men and women in psychology is about the same but in the future women will outnumber men. Also of interest is the number of members of the professional organization NIPP. In the mid 1960s about 80% of psychologists were member of the NIPP but now only about 30% of the psychologists are members.

2.29.2.2 Employment

Until the early 1970s every psychologist could find a job easily. In the 1990s about 80% of graduates find a job within a year. The percentage of unemployment varies per subdiscipline with the lowest unemployment rate for industrial psychologists and an unemployment rate of up to 35% for clinical and child psychology.

The title *psycholoog* used to be protected by law but as of September 1, 1993 its formal protection by law was abandoned.

In November 1993 a law with regard to the professions in individual healthcare (*wet Beroepen in de Individuele Gezondheidszorg* [*Wet Big*]) was accepted by the Dutch government. This law contains rules for a number of professions in individual healthcare with regard to registration, title protection, the kind of activities one is allowed to perform, code of conduct, ethics, and disciplinary procedures.

With regard to the clinical psychologist some specifications are made about official registration as a "Healthcare Psychologist" ("Health Psychologist" by this law is defined as a psychologist who works within the individual healthcare system and contains a larger category than the traditional health psychologist), the specialty of healthcare psychologist, and the educational requirements. A special taskforce was asked to develop specialty requirements with regard to special training and skills for the profession of healthcare psychologist. With the Dutch Psychological Society and the Dutch Pedagogic Society it was agreed that both university tracts should offer the option of becoming a registered healthcare psychologist.

In order to obtain an actual and updated view of the requirements, the kind of work, and the number of psychologists that work within this field, an evaluation was done by the Dutch Center of Mental Health Care (*Nederlands centrum Geestelijke Volksgezondheid,* NcGv). In 1994 it was estimated by the Dutch Central Bureau of Statistics (CBS) that there were about 19 000 psychologists in The Netherlands. A large number of them work within the healthcare systems.

In 1942 the first clinical psychologist was appointed at a hospital under the supervision of a psychiatrist mostly dealing with diagnostics and treatment. Over the years diagnostics and treatment (including psychotherapy) became part of the skills of the clinical psychologist. In 1966 psychologists were allowed to become members of the Dutch Psychotherapy Society (*Nederlandse Vereniging voor Psychotherapie*, NVP).

In the 1970s and 1980s a rapid growth, including a larger differentiation, took place. Clinical psychologists organized themselves within the Dutch Psychological Society (*Nederlands Instituut van Psychologen,* NIP) as a Clinical Psychology section. The number of sections increased due to the increase in differentiations within healthcare. Finally, all these sections were assembled again within the section "Psychologists within Healthcare Systems." With the changes in the healthcare system, resulting in changes in the role of the clinical psychologists, the educational system at the universities also changed. After finishing their specialty in clinical psychology at a university level, graduates had to work for several years under supervision in order to be registered as a clinical psychologist by the Dutch Psychological Society.

In 1982 the university system changed into a two-phase system. It was clear right from the beginning that completion of the first four years of the two-phase system would not guarantee a sufficiently qualified clinical psychologist. At the same time a government held register of psychotherapists was instituted. Special registration requirements for the clinical psychologist were formulated and postdoctoral programs (of four years) for clinical psychologists began. In 1995, 1100 psychologists were registered as clinical psychologists.

2.29.2.3 Employment Settings for Clinical Psychologists

In the late 1990s, psychologists are working within all aspects of the healthcare system, inpatient and outpatient clinics, day care centers, general and mental healthcare, institutions for the mentally handicapped, and so on.

In 1991, Beekman and Heijke estimated the number of psychologists working within the healthcare system by combining data from the CBS, the NIP, and the National Hospital Institute (NZI). This resulted in the overview shown in Table 1.

Another estimate comes from Van Son and Van der Staak (1993). Based upon the figures of the NIP, they estimate that of the 7000 members, about 3000 work within the healthcare system. Of those, 1100 are officially registered as clinical psychologists. On the assumption that 50% of professional psychologists are members of the NIP, 6000 psychologists will be working within the healthcare system.

According to the above-mentioned authors, of the 3000 members of the NIP about 1800 psychologists work in outpatient care and 1000 participate in first-echelon care; about 365 work within rehabilitation centers or nursing homes; about 650 work in psychiatric clinics and rest homes. Institutions for the mentally handicapped, forensic institutions, and general hospitals each employ about 200 clinical psychologists. So by extrapolation (50% membership of the NIP), they arrive at the number of 6000 clinical psychologists.

Table 1 Number of clinical psychologist places within the healthcare system in 1990 (in full-time equivalents of 40 hours).

Inpatient healthcare	
General hospitals	390
Psychiatric hospitals	466
Medical child institutions	12
Medical child day care centers	30
Forensic institutions	40
Nursing homes	59
Institutions for mentally handicapped	145
Institutions for sensory handicapped	13
Subtotal	1155
Outpatient healthcare	
Medical school services	23
Outpatient psychiatric care	800
Institutes for drug addicts	53
Others	12
First-echelon services	245
Subtotal	1133
Forensic	
Prisons	33
Child protection	12
Subtotal	45
Total	2333

As one can see, estimates vary from about 2500 to 6000 clinical psychologists. Because of this large variability in estimates a new estimate was made by a taskforce of the Dutch Center of Mental Health (Brunenberg, Neijmeijer, & Hutschemaekers, 1995).

This review concentrates upon psychologists and pedagogues working within general healthcare, mental healthcare, care for the physically handicapped, care for the mentally handicapped, forensic care, and individual healthcare related private practice.

The psychologists had to meet four criteria: (i) a university education as psychologist or pedagogue, (ii) being employed within the healthcare system, (iii) working within a direct healthcare situation and not as manager, prevention specialist, teacher, or researcher, and (iv) not solely working as a psychotherapist.

First they selected from the membership file of the Dutch Psychological Society all members belonging to the Section of Healthcare ($N=3409$), all members of the Section of Child Psychology ($N=292$), and all members of the Dutch Pedagogic Society ($N=2577$). These 6278 members each received a questionnaire. The second source of information included 1177 institutions in the healthcare setting who were asked to give numbers and names of the psychologists and pedagogues working at their institution. About 53% of the institutions responded, indicating that 1613 psychologists and 679 pedagogues were working there. After eliminating all duplicated names, 7600 questionnaires (4595 psychologists and 3005 pedagogues) were mailed out and 62% responded. For the sake of simplicity from now on differentiation between psychologists and pedagogues will not be made and they will be referred to as clinical psychologists.

Of the 4298 respondents 104 had another form of university education, 1154 were working outside the healthcare system, 326 had other (management, research, teaching) functions within the healthcare system, and 352 worked exclusively as psychotherapists.

The following data are based upon 2362 clinical psychologists working in individual healthcare. Based upon extrapolation (correction for the number of respondents per type of institution), the authors come to an estimate of 4813 clinical psychologists working within the individual healthcare system. This does not include the approximate 15% of clinical psychologists working in private practice (this 15% is based upon individual questionnaires). In total about 5645 clinical psychologists are discussed.

In order to get the whole picture, also included are clinical psychologists with other functions such as management, research, teaching (1320), and also those who work solely as psychotherapists (1030), giving a total of 7383 clinical psychologists working in the individual healthcare system, of whom 62% are females. Table 2 gives a breakdown of clinical psychologists' activities within the healthcare system.

2.29.2.3.1 Employment opportunities

The majority of psychologists have a paid job. About 5% of psychologists have a position as postdoc internship in combination with a paid job. Only 2% have a postdoc internship without a job, 3% work as volunteers.

2.29.2.3.2 Private practice

Approximately 25% of psychologists have some kind of private practice. Of this group, 60% have a private practice exclusively (15% of the total group).

2.29.2.4 Some Final Statements

As of June 1994, 1663 pedagogues and 3310 psychologists worked within the healthcare system. This means that according to the above statistics, within the healthcare system, 200 psychologists and 100 pedagogues are employed per 1 000 000 of the population.

On a yearly basis, about 1000 students complete their psychology studies. Of these, 500 specialize in the health-related areas (clinical and developmental psychology). In the late 1990s, there are about 1000 unemployed psychologists in the health-related area. It is expected that by the year 2001 we will have 4000 unemployed psychologists.

With respect to the *Wet Big* every year (with an optimistic scenario) 160 new places are expected. However, this scenario is based upon the employment situation within the healthcare system. It is as if the NIP is content with the situation in the 1990s and is accepting the status quo. One might wonder if the situation in 1997 is acceptable. Many clinical psychologists complain about the heavy workload they have. Data suggest that only one hour is spent on more than 50% of diagnoses; this is not a very acceptable situation.

Based upon the above analysis and the financial constraints in the first part of the second phase of study (the "healthcare psychologist" part) only 84 people are allowed to enter this new program every year. These students need a paid internship place for two years, so in

Table 2 Comparison of the activities of clinical psychologists within different settings.

	General healthcare	*Mental healthcare*	*Mentally handicapped*	*Physically handicapped*	*Private practice*
Average contract					
Hours per week	27.5	25.9	28.3	21.8	21
Additional job (%)	16.0	20.0		3.0	35
Private practice (%)	12.0	24.0		5.0	
Time distribution (%)					
Client oriented activities	75.0	80.0		80.0	100.0
Diagnostics	25.0	16.0		16.0	13.0
Treatment	18.0	28.0		7.0	44.0
Client rounds	10.0	10.0		15.0	
Guidance	10.0	10.0		11.0	
Advising staff				11.0	
Administration	10.0	10.0		11.0	12.0
Diagnostics (%)					
Interview	44.0	40.0		30.0	70.0
Observation	12.0	23.0		25.0	8.0
Tests	42.0	35.0		41.0	21.0
Others	2.0	2.0			
Types of diagnosis (%)					
Somatopsychological	11.0	2.0		0.0	7.0
Neuropsychological	40.0	12.0		11.0	
Intelligence	8.0	11.0		30.0	5.0
Personality	13.0	26.0		12.0	33.0
Psychopathology	11.0	21.0		8.0	20.0
Social functioning	6.0	11.0		12.0	15.0
Developmental				12.0	
Diagnositc contacts					
Time per contact (mins)	93	92		111	100
Number of contacts	3	3		3	3
Treatment contacts					
Time per contact (mins)	60	42		45	53
Frequency	1/2 mths	1 wk		1 wk	1 wk
Patients per year	53	47		107	33
New contacts per year	93	62		25	42
Average treatment time		11 mths		9 mths	8 mths

total, 168 places are needed. The idea is to train these psychologists in such a way that they have a very broad basic background in the area of the individual healthcare. The only differentiation that is foreseen is between adults and children. After the first part of the second phase, the "health psychologist" will be registered in an official government register. Out of this registered group, around 25% will be allowed to continue for the second part of the second phase to be trained to as specialists in clinical psychology (comparable with medical specialists). How specialized they should be is still a matter of discussion. On the one hand, it is suggested that they should be generalists in terms of clinical psychologist/psychotherapist and on the other hand it is suggested that they should be highly specialized clinical child psychologists, neuropsychologists, health psychologists, psychogerontologists with or without "psycho" therapy qualifications.

2.29.3 COUNTRY-RELATED INFORMATION

For information on the UK see Chapter 22, this volume, and for information on the Scandinavian countries see Chapter 15, this volume.

2.29.3.1 Belgium

Most of the information on Belgium has been forwarded by Bob Cools, Chairman of the Flemish Society for Clinical Psychology.

Clinical psychology is a relative young discipline in Belgium. The first program in clinical psychology started in 1967 at the University of Leuven, followed in 1975 by the University of Gent, and in 1976 by the Free University of Brussels. In 1997, psychology can be studied at seven universities and there are about 5000 psychology students.

The university program has two cycles: *de kandidaturen* (two years) and *de licenties* (three years). During these last three years the student can specialize in clinical psychology. After the *licentie* one is allowed to function professionally as a psychologist.

After the *licentie* there are several opportunities inside or outside the university system for further specialization. A questionnaire by the *Vlaamse Vereniging van Klinisch Psychologen* (Flemish Society for Clinical Psychologists) indicated that out of 481 clinical psychologists, 61% followed an additional specialty in psychotherapy. In the late 1990s, there are about 3000 clinical psychologists in Flanders.

From their foundation, the *Belgische Federatie van Psychologen* (BFP) worked towards a formal recognition of the title *psycholoog* by law. Finally in 1993 the university level *licentiaten in Psychologie* was recognized. However, by 1997 the law had not been put into effect. Protection of the title by law could be considered as the first step toward a fully-fledged status for clinical psychology with appropriate by-laws.

Although clinical psychologists work in the health profession, their five year university education is hardly recognized and their postgraduate training is not recognized at all. They are not allowed to work by themselves in a curative way. Even if their psychological assessment is in some way health related, they are not officially permitted to practice, nor will they be reimbursed by the health authorities.

The title "psychotherapist" is not officially protected. Psychotherapy as a curative activity is considered to be part of the medical profession. Violation of the Royal Degree by a clinical psychologist is an offence which will be submitted to a public prosecutor and is liable to a law suit. A Belgian minister of health labeled clinical psychology as a paramedical discipline. However, the BFP refused to accept this position and claims its own independent status as a scientific discipline in its own right. Legally there is no solution in the foreseeable future for this problem.

2.29.3.1.1 Work setting

Of the above-mentioned 481 interviewed clinical psychologists, a vast majority worked within a health setting, 40% in an intramural, residential setting, 30% in an extramural, outpatient setting, and 20% in private practice. Ten percent work in an educational setting, and 10% work in private business or for a governmental agency. Only 2% work in a research setting. Six percent are unemployed. The employment market is becoming increasingly more competitive and it takes students longer to find a tenured job.

The fact that a licenciate clinical psychologist follows an additional academic track (mostly psychotherapy) appears to be a significant factor in obtaining a job. About 61% of Belgian psychologists follow this additional therapy training.

2.29.3.1.2 Future

The traditional job perception of a clinical psychologist working as a psychotherapist within the health system is still unclear in Belgium. The scattering of all kinds of psychologists in all kinds of different small societies prevents the BFP from extending any political pressure or power. Relatively few (clinical) psychologists are members of the BFP. Because of this current dead-end street situation, it is time for clinical psychologists in Belgium and for the Belgian university system to widen the objectives of clinical psychology towards prevention, health, sport, and geriatrics.

2.29.3.2 Czech Republic

Three universities in the Czech Republic (Charles University in Prague, Masaryk University in Brno, and Palacký University in Olomouc) offer a five-year psychology program. All three universities offer a specialization in clinical psychology.

Psychology in the Czech Republic is still changing as a result of events in its history. According to Brožek and Hoskovec (1995), one can distinguish three periods: (i) From the "victorious February" of 1948 to the end of November 1989. The country was called the Czechoslovak Socialist Republic under the Soviet-controlled totalitarian communism. (ii) From the "Velvet Revolution" of 1989 to December 31, 1992, the end of the Federal Republic of Czechoslovakia. (iii) From January 1, 1993, the date of the emergence of the Czech Republic and Slovakia as separate states. The consequences of this split are not clear yet. There still exists the *Journal Československá*

Psychologie, human contacts still continue, and joint meetings are still held alternating the site between the Czech Republic and Slovakia.

All the psychology associations of the Czech Republic are collective members of the Czech–Moravian Psychological Society, associated with the Academy of Science. The society has about 1700 individual members, representing about one-third of the total number of psychologists in the Czech Republic. The total number of psychologists in the Czech Republic is about 5500. Of the six associations, one is called Association of Clinical Psychologists of Bohemia, Moravia, and Silesia. Psychology is becoming more and more important within the healthcare system. Health services, including mental health services, continue to be available free of charge. In the area of counseling, privatization of services is progressing slowly. In the late 1990s only 25% of the counselors in the field of marital and family counseling have a permit for private practice.

2.29.3.3 France

La Société Française de Psychologie (The French Psychological Society) was founded in 1901 as a scientific society. It was only in December 1991 that practitioners were allowed to join the society, and in 1993 there were approximately 2000 members.

Psychology can be studied at 29 public universities and two private universities in France. The program takes five years and is subdivided in three parts. The first part takes two years and results in a *Diplôme d'Études Universitaires Générales* (DEUG). The second part, which also takes two years, results in a *Licence* or a *Maîtrice*. The third part takes one year and leads to the *Diplôme d'Études Supérieures Spécialsées* (DESS) in Psychology and *Diplôme d'Études Approfondies* (DEA) in Psychology.

The three main fields of practice are clinical psychology, work and organizational psychology, and educational psychology and guidance. The three main fields include many other fields such as psychotherapy, childhood and adolescent psychology, expert evaluation, neuropsychology, social psychology, cognitive psychology, economical psychology, and sport psychology.

Since 1985 the title *psychologue* has been protected by law on a national level enacted by parliament and government.

In 1991 Reyns and Verquerre published a study carried out in 1987. They mailed a questionnaire to all psychologists who obtained their DESS or DEA in 1984–1986. Only 895 psychologists (39%) responded. Of these, 659 (73.6%) were clinical psychologists, 151 (16.9%) were social psychologists, work psychologists, or organizational psychologists, and 85 (9.5%) were research psychologists. The major employers of clinical psychologists are national associations (e.g., Red Cross) (44%), the public sector such as child guidance centers (30.2%), hospitals (19.7%), and private practice (14.3%). The activities of the clinical psychologists were 58% clinical psychology, 32.7% institutional work, 31.1% psychotherapy, 26.9% training, 15.6% selection, and only 5.3% research (because many psychologists were involved with more than one activity this adds up to more than 100%).

Of the interviewed clinical psychologists, the following techniques were used: 59.8% used tests, projective techniques, or graphology; 88.3% used interview techniques; 42.9% used group interviews or group conferences; 5.1% used questionnaires or experiments; 22% used specific therapeutic techniques like psychoanalysis, gestalt therapy, or system therapy; 6.4% used a specific theoretical approach like psychoanalysis or system theory; and 3.5% used role-play techniques or transactional analysis. Of the interviewed clinical psychologists, 23.3% were unemployed, 67.4% had a regular monthly income, and 9.3% had a private practice. This relatively low percentage of clinical psychologists in private practice is due to the fact that there are no regulations in France. The independent clinical psychologist does not get paid by the government or a health insurance system. Because the French consider clinical psychology as a medical activity, and because they are not used to paying for medical activities themselves, it is almost impossible to run a private practice in France.

2.29.3.3.1 Postgraduate training

According to Reyns and Verquerre (1991), 57.5% of clinical psychologists follows some kind of postgraduate training. Approximately 20% of clinical psychologists have a secondary diploma from the third part. Apart from the university program there are many specialty training programs which are mainly complementary. There are programs in psychotherapy, psychoanalysis, psychopedagogics, criminology, sexuality, even graphology. There is an ethical code (*deontologie*) for the members of the *Société Française de la Psychologie*.

Payment is insured either by the state or by the *départments*. The psychological services that are covered by this system are health service (clinical psychology and psychotherapy), *départment* services (clinical psychology, psychotherapy), justice (clinical psychology, psychotherapy, expert evaluation), education (educational psychology, guidance), and private services (clinical psychology, educational psychology, psychotherapy, sometimes organizational psychology).

2.29.3.4 Germany

When looking at the situation of clinical psychology in Germany, one should bear in mind the fact that the reunification between East and West Germany only took place in 1990. The current *Berufsverband Deutscher Psychologen*, BDP (German Association of Professional Psychologists) includes members of the former *Gesellschaft für Psychologie der DDR* (Society of Psychology of the GDR). This information is based upon a representative survey of members of the BDP in 1990/1991 and former members of the GDR in 1991/1992. Of the 10 000 members of the BDP, one-third received a detailed questionnaire and 53.8% responded ($N=1630$), and of the 2500 members of the GDR, one-sixth received a detailed, and for the former East Germany, adapted questionnaire and 37.9% responded ($N=154$). This study was done by Schorr (1995). If we look at the subject sample we do not see any differences with regard to age (all are about an average of 43/44 years of age). However, there is a striking gender difference. Results showed that in the former West Germany the male–female distribution among psychologists was more or less fifty/fifty, in the former East Germany almost 70% were female. Another striking difference was the fact that at the time of the survey no psychologist was in private practice in the former East Germany compared with 27% in the former West Germany.

By comparing the areas of specialization one can see that clinical psychology was the largest specialization in what were West (81.5%) and East (70.9%) Germany. This was followed by work and organizational psychology in the West (11.2%) with only 7.1% in the East; while in the former East Germany, school and educational psychology came second with 16.2% with only 4% in the West.

By comparing the applied fields with each other it was first noticed, as was mentioned before, that 27.1% of the former West German psychologists worked in a private practice, whereas there were none in private practice from the former East Germany. Employment in counseling centers was about the same (18.9% in the former West Germany and 17.8% in the former East Germany). In clinics and rehabilitation clinics, one could see a significantly higher percentage of psychologists in the former East Germany (33.4%) contrary to only 14.3% in the former West Germany. In psychiatry the same kind of significant difference could be seen with 19.1% in the former East Germany and 9.7% in the former West Germany. For alcohol and drug therapy centers, the distribution was 4.6% in the former West Germany and 2.8% in the former East Germany.

If we look at the main tasks of a psychologist it can be seen that what were West and East German psychologists mention counseling and psychotherapy first (respectively 84.6% and 88.3%) followed by psychological assessment which was significantly higher for the former East German psychologists (West 51.1% and East 66.9%) and administrative duties which were significantly higher for psychologists from the former West Germany (West 32% and East 10.4%).

In summary, the above shows that the vast majority of psychologists in the former West and East Germany specialized in clinical psychology and that in the former East Germany about 55.3% of the psychologists were employed within a clinical setting, while in the former West Germany only 28.6% were working in a clinical setting. To what extent the percentage of psychologists in private practice are involved in clinical work remains unclear from this data.

2.29.3.4.1 Aspects of psychological training

(i) Basic training

In Germany there are 37 psychology programs at 36 universities. Each year approximately 6000 students apply to study psychology but only 3300 are accepted. In 1996 there were about 40 000 psychologists; around 6000 unemployed, 41% work in a clinical setting and 19% work independently as therapists or counselors. The whole study program takes six to seven years.

The *Hochschulramengesetz* (new frame law) regulates the general program of 10 semesters, without any specialization. However, many universities have not yet implemented this program. The idea is to add a third stage of around three years to the program, allowing about 5–20% of the *Diplom* students to continue to become "clinical psychologists." The consequence might be depreciation of the *Diplom* and an increase in unemployment which already is 20% for psychologists. This amounts to about 6000 psychologists between the age of 25 and 35 years, of which 60% are women. To date, the title "psychologist" is not protected by law. However, there is a *Berufsordnung für Psychologen* (ethical code) for members of the BDP.

(ii) Postgraduate training and supervision

Psychologists possess, to a surprising extent, further qualifications in addition to their degree in psychology. When quoting only academic qualifications as an example, 25.3% of the former West German psychologists and 10.4% of the former East German psychologists possess an additional qualification. When asked about postgraduate qualifications, qualifications in psychotherapy predominate. More than half of all German psychologists possessed a postgraduate qualification at the time of inquiry. Close to two-thirds of all German psychologists stated that they attended professional training regularly.

Because many psychologists possess a postgraduate training in psychotherapy, they tend to work curatively (not only advising as a counselor). This leads to a conflict of interest with physicians who in the strict sense of the word "curative" were the only ones allowed by the law of 1939 (*Heilpraktiker-Gesetz*) to practice a curative profession. Others are allowed to work in a curative way, however, they do not get paid by the health insurances directly. Psychologists only get paid if their service is requested by a physician. Since physicians are having problems in the job market themselves, they have started practicing psychotherapy themselves, and by doing so, have pushed psychologists out of the market.

In the late 1990s, there are close to 40 000 psychologists in Germany, increasing every year by around 1500. Of those, 41% are working in the clinicial area and 19% work independently as therapists or counselors. This means that only 60% of the psychologists who are trained in clinical psychology find a job in the clinical field. About 20% are unemployed, amounting to about 6000 psychologists between the age of 25 and 35 years old of which 60% is female. Of the 40 000 psychologists only 35% are members of the German Association of Professional Psychologists.

2.29.3.5 Greece

According to James Georgas (1995), when thinking about the history of psychology we should realize that twenty five centuries ago, the ancient Greek philosophers defined and explored the subject, making them the precursors of many theories of modern psychology.

For instance, the concept of empirism led to a school of behaviorism that has dominated psychology in the twentieth century. This concept was originally developed by Protagora, Empéricos, the Stoics, Iamblichos, Plato, and Aristotle. The term *charten agraphon* was employed by Aphrodiseas many centuries before Locke employed its Latin synonym *tabula rasa* (Georgas, 1995).

Among all the ancient Hellenic philosophers, Aristotle is acknowledged as the most accomplished in defining and systematically studying the phenomenon of psychology. However, this chapter deals with modern psychology in modern Greece. The establishment of modern psychology is ironically probably delayed because the philosophers at the University of Athens emphasized the importance of the Greek philosophers to psychology. While in the 1950s and the 1960s psychology separated from philosophy at many American and European universities, it took Greece until 1983 to establish divisions of psychology within departments of philosophy, education, and psychology. In 1987 the first autonomous Department of Psychology was created at the University of Crete. All major Greek universities (i.e., Athens, Thessaloniki, Ioannin, and Crete) now have a department of psychology. These departments run four-year graduate programs leading to a degree called *Ptychio* (Masters degree). In addition, programs leading to the Doctor of Philosophy are awarded at all of these universities.

Because psychology was late in becoming a discipline in itself, one should realize that the majority of Greek psychologists are trained abroad. The largest group has been trained in France, followed by a second group trained in the UK, followed by Germany, the USA, and Canada. One could say that in Greece, psychology is a kaleidoscopic assembly of different theoretical schools and practical applications. The majority of psychologists in Greece (around 800) are members of the Association of Greek Psychologists. Members have to have a degree in psychology at the masters level. A second association was established in 1991, the Hellenic Psychological Society, whose members are university faculty members and psychologists in research centers. The minimum criterion for membership is a Ph.D. in psychology.

A law which licenses psychologists to practice was passed in 1979, regulated by the Ministry of Health and Welfare. However, it only came into force, primarily because of the problems in delineating the specific types of "psychological work" of the three main fields: clinical psychology, educational psychology, and organizational psychology.

Psychologists in Greece work in a variety of settings. By far the largest number are clinical psychologists. The first mental health centers in Greece were established in the 1950s and subsequently the Institute for Child Health and several private mental health centers were also established. Most clinical psychologists are employed in mental health centers. Only few of them combine a full-time position in the mental health or hospital system with a private practice (Georgas, 1995).

2.29.3.6 Italy

Most of the following information is based upon an article by Avallone (1993). In order to understand psychology in Italy one needs some historical background. As in other countries, one might say that experimental psychology started in the second half of the nineteenth century. The first theoretical paper *La psicologia come scienza positiva* by Roberto Ardigò was published in 1870. The first Italian psychological journal appeared in 1905, *Rivista Italiana di Psicologia*, and in 1912 *Psiche-Rivista di studi psicologici*. In 1906, the Fifth International Congress of Psychology took place in Rome and the first three academic chairs were created in Rome, Naples, and Torino. Psychology was introduced in schools and institutions. Mental tests were translated and developed, and in 1910 the Italian Psychological Society was founded. During the first world war in 1915 psychological tests were used to select army personnel, but then the development seemed to stop. The strong influence of the philosophical tradition, but even more due to the influence of the neoidealism (=fascism) and the second world war, negatively influenced the development of science in general and psychology in particular. There also existed an almost traditional dichotomy between pure basic research and applied research resulting in a professional practice without a methodology or theory. It was not until 1971 that one could study psychology in Italy resulting in a graduate degree in psychology. Even in the 1990s the majority of professors in psychology have their primary background in philosophy or medicine. As of 1971 there was a four-year program in psychology, divided into experimental, educational and applied psychology. After the need for a better curriculum became apparent, the four-year program was increased to a five-year program in 1985. At the end of a two-year *propadeuse*, the student studies for another three years in one of the following areas: (i) general and experimental psychology (about 8% of the students); (ii) developmental and educational psychology (15%); (iii) clinical and community psychology (62%); and (iv) work and organizational psychology (15%). Like in other European countries, a stereotypical interest in clinical psychology is seen (Carli, 1987), showing that being a psychotherapist is the ultimate goal of many psychology students.

From 1985 to 1996, the number of universities offering a psychology program increased (Torino, Trieste, Padova, Bologna, Cagliari, Rome, Palermo). In 1996, 11 universities offered a psychology program. Apart from the regular curriculum, Italian universities offer *Corsi di Perfezionamento* allowing students to specialize even further during at least one year. The *Scuolo di Specializzazione* offers a limited number of students to specialize during three or four years in, for example, clinical psychology. The *Dottorato di Ricerca* prepares the student for a research career. Apart from these university programs which are available for a limited number of students, many private programs are available in order to become a psychotherapist in almost any thinkable form of therapy.

In 1989 the State intervened by law, regulating the psychological profession; only psychologists with an academic degree were allowed to be registered in the official register.

Psychotherapists are only registered after an additional four years of university or university equivalent (State recognized) training. Because of this, the new generation of psychologists in Italy can claim a solid base in science for their applied activities.

The major applied area is clinical psychology with many psychologists entering private practice after the State required four-year training. However, many clinical psychologists work for the State, province, and local health units in many health-related structures (being called mental hygiene centers for psychiatric assistance, because Italy had banned all psychiatric hospitals as a result of the antipsychiatry movement).

In Italy, the *Societá Italiana di Psicologia* (SIP) approved its deontological code on September 24, 1981.

2.29.3.7 Portugal

If one looks at the development of psychology in Portugal one should look at the time before and after the democratic revolution in April 1974.

Before the revolution, the State-run universities did not acknowledge psychology as a subject in its own right, it was only a subsidiary course in philosophy. However, several organizations, headed by the Catholic church, created a five-year degree course in psychology in the *Instituto Superior de Psicologia Aplicada* (Higher Institute of Applied Psychology, ISPA), which had evolved from the "Pedagogical Higher Institute" which was founded in 1962. Public universities did not start a psychology program until 1977. Consequently, the majority of certified psychologists claiming a degree in psychology between 1969 and 1982 received their training at the ISPA. These psychologists are assembled in the official but small *Sociedade Portuguesa de Psicólogos* (Portuguese Psychological Society, SPP).

Following the revolution, the proposal to create a psychology course at the public universities was made. In 1977 the psychology course was offered within the framework of the Faculties of the Humanities. A little later the course was offered by the newly created Faculties of Psychology and Educational Sciences (FOCE). In the meantime, psychology degrees were still awarded at the ISPA, which by now was detached from the Catholic founding organizations. At the same time a new *Sindicato Nacional dos Psicólogos* (National Union of Psychologists, SNP) was founded in order to assemble all practitioners of psychology together; those with a psychology degree from ISPA, those with a philosophy degree from State universities, or those with a psychology degree from public universities after 1982.

Rivalry between psychologists from public universities and psychologists from the ISPA resulted in a law recognizing the equality of the ISPA degree with the public university degree. In the context of this competition, the *Associaçâo Portuguesa de licenciados em psicologia* (Association of Psychology Graduates from Public Universities) was established in 1982. The *Associaçâo Portuguesa de Psicólogos* (Portuguese Psychological Society, APP) was also founded.

Once the relationships among psychologists started to improve, members of the Association of Psychology Graduates from Public Universities changed their association into *Associaçâo de Psicologia Portuguesa* (Association of Portuguese Psychologists, APPORT). More recently the Portuguese Society for Clinical Psychology was founded. In 1997 work is in progress to establish the *Ordem dos Psicólogos* (Association of Professional Psychologists). In order to create a single representative body, four of the Portuguese psychological organizations joined forces, thus establishing the "National Board of Psychologists." The purpose of this Board is to draw up the statutes of the "Association of Professional Psychologists."

From the above one can see that the situation is quite complex in Portugal. Because of this it is almost impossible to produce realistic figures and data. There is a rumor that Portugal has about 4000 psychologists, but an estimate of 2500 seems more realistic. In 1996 there were about 3000 undergraduates in psychology. After a five-year program about 600 get their degree every year.

Most psychologists are employed in the educational sector and within organizations in the private sector. The situation with regard to clinical activities is unclear.

The title of "psychologist" is semiprotected. In order to be employed as a psychologist, one has to have an official document called the *carteire profissional* (professional certificate) issued by the Ministry of Work. However, illegal practice does not entail any legal consequences. Hopefully the establishment of the Association of Professional Psychologists will improve this situation (Pereira, 1995).

Four universities offer a psychology program: Lisboa, Porto, Coimbra, and Ispa. The program takes five years and there is a fixed number of entrants.

2.29.3.8 Spain

The current trend in psychology is intervention oriented. This intervention is, due to the specific developments in science in Spain, very technically oriented. In order to understand this development one has to focus upon the scientific dimensions of fundamental psychology in Spain. The emergence of scientific technology became the social justification for psychology as a profession. Standard studies about the history of psychology in Spain usually start with Sénèque, followed by the Renaissance, and ends by analyzing the last 100 years. If we concentrate on the period in which the current discipline of psychology is constructed, we can distinguish two periods.

The first period started at the beginning of the twentieth century until the Civil War (1936–1939). The second period started in the late 1940s to early 1950s, and ended with the Franco regime in the second half of the 1970s. During this period Spain was isolated from the scientific world. The transition to democracy coincided with the start of independent psychology departments. In Spain there exists a rich tradition studying psychological phenomena from a philosophical, medical, historical, and even a literary background. This rich tradition has prevented the modern development of psychology for many years.

As a major event the Spanish psychologists mention the XIe International Psychology Conference planned to take place from September 6 to September 12, 1936 in Madrid. Cajal was president at that time. However, due to the Civil War, the conference was canceled and took place in 1938 in Paris.

The creation of the Spanish Psychological Society started in 1950 and the bylaws were published in 1952. In the following years within the Official College of doctors and licences in philosophy a section of professional psychologists started. In the meantime other psychological associations were created. All these different societies and associations formed the *Colegio Oficial de Psicólogos* (COP) in 1979.

It took until May 1979 that by a Décret Royal psychology became an independent discipline within the university system. Psychology is taught in at least 20 universities and institutions. Within 10 years the number of registered students in psychology increased from about 1000 in 1977 to over 5000 in 1988. The number of doctorates increased from 13 in 1977 to 139 in 1987.

In 1980 the university education in psychology, resulting in a licence, was recognized by law (BOE, April 14, 1980). Each university faculty confers the degree of *Licenciatura en Psicologia* and the COP legally confers the protected title of *Psicologia*. There is still an enormous interest in psychology with more than 39 000 students in 1996 compared to initially 100 students. Currently, there are more than 35 000 psychologists in Spain. The same is true for the Spanish Psychological Societies. The COP and the *Colegi de Cataluña* had about 4500 members in 1981 and now have more than 25 000 members. Spain has probably the highest percentage of psychologists per capita (about 575 psychologists per 1 000 000 inhabitants) in the world.

Because psychology is a young discipline the average age of the Spanish psychologist is quite young. About 30% are below the age of 30, while only 15.4% are older than 40 years. Psychology attracts more women than men (63% women). Of the women the majority prefer to work as school-psychologists or pedagogues (41%), and in traffic and safety psychology 84% of psychologists are female. Clinical psychology is preferred by 37% of the students. Rehabilitation psychology attracts about 18%, while organizational psychology, social psychology, and forensic psychology each attract only less than 10%.

The percentage of unemployment among psychologists is not too bad; only 10.5% vs. 17.3% in the active population and 10.5% vs. 13.6% of the academic population. But this situation will change soon. Among the younger generation of psychologists (younger than 25 years of age), the unemployment rate is high (37.5%). Psychologists have to focus upon new applied areas. There is still an overemphasis in educational psychology, pedagogics, and clinical psychology. This might be the reason why 53.7% of the clinical psychologists work in private practice. More attention should be paid to organizational psychology, social services, justice, and traffic psychology, while there is hardly any interest for environmental, sports, forensic, or military psychology. It is only these new areas that still offer job opportunities.

The first step towards a deontology for Spanish psychologists was set during the first National Psychology Congress in Madrid in 1984, were Alfredo Fierro presented a paper called *Deontological Code* (first draft). Many discussions followed and in 1987 a taskforce of 100 experts got together in Madrid, resulting in a proposal that after elaboration by Alejandro Avila (president) was presented to the Government Board of the *Colegio Oficial de Psicólogos*. The board ratified this proposal.

2.29.3.9 Switzerland

The following observations are based upon a periodically conducted investigation into the employment possibilities for Swiss recent graduates by a working group for academic career counseling (Diem, 1995).

The 1988 sample had a reply rate of 60% and among the 4856 responding graduates, 200 of them were psychologists. Although the unemployment rate for psychologists one year after they graduated is only 5.5% for Swiss standards it is almost twice as high as the average unemployment rate of 3%. Especially when there is an economic recession, like in 1985, the unemployment rate among psychologists was as high as 12%. The major employers for psychologists are universities (20%), social services (42.3%), health services (14.1%), private services like banks and insurance companies (7.9%) and schools, as teachers (7.3%). For 35% of the jobs recently occupied by psychologists, no academic degree was required. Also psychologists are more frequently employed on a part-time basis which is related to the high percentage (68%) of women in psychology (in all other academics an average of 31%). Overall, academic women engage in part-time work (up to 40%), while the figure for men reaches 21%. The higher contingent of women can also explain the lower average income of psychologists. A study by the *Arbeitsgemeinschaft für Akademische Berufsberatung* reveals that women earn about 20% less than men in all employment areas.

The Nationalfond study *La réussite professionnelle des économistes, juristes, médecins et psychologues* (Berberat, Goldschmid, Neirynck-Carton-de-Wiart, Nughes, & Ricci, 1990) gives us some information about the labor market for psychologists. The subject sample comprises 332 psychologists (179 women) who completed their studies at a western Swiss university between 1949 and 1989. It regards exclusively members of the Federation of Swiss Psychologists (*Föderation Schweizer Psychologinnen und Psychologen*, FSP). Almost all of them were employed. Twelve percent worked at Universities or Research Institutions, 30% in education psychology or social services, 21% in private practice, 13% in hospitals and clinics, and 14% in the schools. As one can see these figures differ slightly from the general situation in Switzerland mentioned earlier.

2.29.3.9.1 Clinical psychology in Switzerland

Compared to other countries in Europe, a relatively small percentage (13–14%) of the psychologists are working within the health services. This might be due to the fact that in Switzerland the foundation of clinical psychology was laid by medical doctors. Initially psychologists were only allowed to work with aptitude tests. In the 1950s, psychologists were allowed to enter into psychiatric treatment, initially as diagnostic experts, but later on they increasingly applied themselves to psychopathology (originally a medical discipline). More recently psychotherapy was added to their repertoire. Seen as a whole, the relationship between physicians and clinical psychologists is still largely a dependent one and clinical psychologists tend to represent a marginal group in psychiatric institutions.

In private practice the competition is unequal. Medical doctors' work is entirely compensated by health insurance companies, while psychologists' services are often only partially covered or not compensated at all.

In the research area the picture is totally different. They either work at universities, in research departments of psychiatric clinics, or in clinical research funded by industry. Psychologists' thorough training in methodology is highly respected. Another problem in Switzerland is that the title of "psychologist" is not protected by law.

2.29.4 CONCLUDING REMARKS

Looking over past developments in Europe with regard to psychology in general and clinical psychology in particular, we see some trends emerging. First of all we can see that all over Europe the professionalization of psychology is taking place. Most of the time the professionalization is triggered and catalyzed by clinical psychologists interacting with public health organizations. In order to guarantee a professional high standard of psychological expertise, official registers of psychologists are contrived. Although the need for registration is most strongly felt within the health care services in several countries, this format of registration is extended to all psychologists or for several specializations within psychology.

Another trend we see is that the registration, initially coordinated by national psychological societies, is increasingly done by the official governmental agencies mostly in coordination with the national psychological societies. In some countries the registration leads to protection of the title "psychologist," in other countries only some specialty areas within psychology are registered. Usually registration by law of the title "psychologist" is automatically related to quality control mechanisms.

With regard to education we see an increasing tendency all over Europe to shorten the study time (mostly six to seven-year programs revert to four to five-year programs). In countries where they maintain a longer curriculum, a numerus clausus or numerus fixus is instituted. The consequence of this shortening of the study time is that most studies are considered to be inadequate and/or insufficient. Because of that, additional training programs are developed and required in order to become a registered psychologist. In that case there is always a fixed total with regard to the number of potential candidates who are allowed into this additional postdoctoral training program. So far, a system of continuous education is not officially instituted in Europe.

The kind of work psychologists are doing is diversifying rapidly (e.g., traffic psychology in Spain), but still the majority of psychologists work in health-related areas. In many countries psychologists are still an extension of medical doctors but increasingly it is observable that psychology is recognized as a specialty by its own right, and that the psychologist is paid through the health system and health insurances directly (however, many times only after a referral by a medical doctor).

Next to the increase in the health-related area, we see an increase in private practice: sometimes a single-person practice but often a group practice sometimes even with other health-related professionals. This form of privatization is due to the fact that government institutes hire less people on a tenure basis. Often the official governmental institutions prefer to subcontract special orders/requests to a private organization.

The third area of expansion is in research. The healthcare world, but also (pharmaceutical) industries, are becoming aware of the excellent research training many psychologists in many countries get. This is also one of the few areas where within healthcare there is no competition between medical doctors and psychologists. Psychologists are by virtue of training far better researchers.

I should also say something about the change within clinical psychology itself. It used to be a soft discipline with all "funny" kinds of psychotherapy, and it was hard to prove that they did any good to the clients/patients. But due to financial cut-backs in the more hardcore areas of experimental and research psychology, many experimental, cognitive, psychophysiological, personality, and neuropsychologists showed a clinical inclination/interest which had a strong modifying effect.

All of a sudden assessment became more important, diagnosis was allowed and even mandatory later on, treatment effects were evaluated, theoretical backgrounds were studied, rehabilitation programs and treatment programs were based upon scientific theories from neuroscience, cognitive psychology, or personality psychology.

Although psychology grew in the bosom of philosophy and moved from there into social and educational sciences, by now clinical psychology is becoming a β-science with a very close relationship towards neurosciences.

The above mentioned developments are strongly supported and often initiated by the European Federation of Professional Psychologists Associations (EFPPA). In 1993 Member Associations of the EFPPA were: from Austria the *Berufsverband Österreichischer Psychologen*, from Belgium the *Belgische Federatie van Psychologen—Fédération Belge de Psychologues*, from Denmark the *Dansk Psykologforening*, from Estonia the Union of Estonian Psychologists, from Finland the *Suomen Psykologiitto ry*, from France the *Association Nationale des Organisations de Psychologues*, from Germany the *Berufsverband Deutscher Psychologen*, from Greece the Association of Greek Psychologists, from Hungary the *Magyar Pszichológiai Társaság*, from Iceland the Association of Icelandic Psychology, from Ireland the Psychological Society of Ireland, from Italy the Italian Network of Professional Psychologists' Associations, from Liechtenstein the *Berufsverein der Psychologen Liechtensteins*, from Luxemburg the *Association des Psychologues diplomés*, from Malta the Malta Union of Professional Psychologists, from the Netherlands the *Nederlands Instituut van Psychologen*, from Norway the *Norsk Psykologforening*, from Poland the *Polskie Towarzystwo Psychologiczne*, from Portugal the *Sindicato Nacional dos Psicologos*, from Slovenia the *Drustvo Psihologov Slivenije*, from Spain the *Colegio Oficial de Psicológos*, from Sweden the *Sveriges Psykologförbund*, from Switzerland the *Föderation der Schweizer Psychologen*, from Turkey the Turkish Psychologists Association, from the UK the British Psychological Society, and the Affiliate Member, the Israel Psychological Association.

ACKNOWLEDGMENTS

I would like to thank the quoted authors for their indirect contribution. I hope that, due to limited space, I did not do injustice to the refered articles. Any misquotation is my own responsibility.

2.29.5 REFERENCES

Avallone, F. (1993). Training and professional spheres of psychology in Italy. *European Review of Applied Psychology, 43,* 113–119.

Beekman, T. & Heijke, H. (1991). *Prognoses voor 1990–1995. De arbeidsmarkt voor Klinisch Psychologen* [Prognosis for 1990–1995. The employment market for clinical psychologists]. *De Psycholoog, 26,* 187–193.

Berberat, A. L., Goldschmid, M. L., Neirynck-Carton-de Wiart, I., Nulghes, M. A., & Ricci, J. L. (1990). *La réussite professionelle des économistes, juristes, médecins et psychologues.* Lausanne, Switzerland: Chaire de Pedagogue et Didactique, Ecole Polytechnique Fédérale.

Brožek, J., & Hoscovec, J. (1995). Psychology in the Czech Republic. In A. Schorr & S. Saari (Eds.), *Psychology in Europe* (pp. 3–14). Seattle, WA: Hogrefe & Huber.

Brunenberg, W., Neijmeijer, L., & Hutschemaekers, G. (1995). *Beroep: psycholoog/pedagoog* [Profession: psychologist/pedagogue]. Utrecht, The Netherlands: Nederlands centrum Geestelijke volksgezondheid.

Carli, R. (1987). *Psicologia clinica.* Torino, Italy: Utet.

Diem, M. (1995). Swiss psychologists and the labor market. In A. Schorr & S. Saari (Eds.). *Psychology in Europe* (pp. 125–148). Seattle, WA: Hogrefe & Huber.

Georgas, J. (1995). Psychology in Greece. In A. Schorr & S. Saari (Eds.). *Psychology in Europe* (pp. 59–76). Seattle, WA: Hogrefe & Huber.

Network of the Executive Secretaries of the Members of the European Federation of Professional Psychologists Associations (Ed.) (1993). *Inventory of regulations in the field of psychology in European countries.* [Brochure]. Bonn, Germany: Berufsverband Deutscher Psychologen.

Pereira, F. (1995). Psychology in Portugal. In A. Schorr & S. Saari (Eds.), *Psychology in Europe* (pp. 97–102). Seattle, WA: Hogrefe & Huber.

Reyns, S. & Verquerre, R. (1991). L'insertion professionnelle des psychologues. *Psychologues et Psychologies, 99,* 2–28.

Schorr, A. (1995). German psychology after reunification. In A. Schorr & S. Saari (Eds.), *Psychology in Europe* (pp. 35–58). Seattle, WA: Hogrefe & Huber.

Van Drunen, P. (1995). Professional psychology in The Netherlands. In A. Schorr & S. Saari (Eds.), *Psychology in Europe* (pp. 77–96). Seattle, WA: Hogrefe & Huber.

van Son, M. J. M., & van der Staak, C. P. F. (1993) Inleiding Handboek klinische psychologie [Introduction handbook clinical psychology]. In M. van Son & C. van der Staak (Eds.), *Handboek klinische psychologie.* Houten, The Netherlands: Bohn, Stafleu Van Loghum.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.30

Establishing and Maintaining a Clinical Practice in the United States

DONALD J. WEINSTEIN
Beachwood, OH, USA

2.30.1 INTRODUCTION 353

2.30.2 CONSIDERATIONS IN ESTABLISHING A PRACTICE 354
  2.30.2.1 Location(s) 354
  2.30.2.2 Comprehensive Service Base 354
  2.30.2.3 Comfort and Attractiveness for Me and for the Public 354
  2.30.2.4 Hours of Service 354
  2.30.2.5 Overhead Surveillance 354
  2.30.2.6 Responsiveness to Caller 354
  2.30.2.7 Income Consideration 354
  2.30.2.8 Values and Ethics 355

2.30.3 COMPONENTS OF A PRACTICE 355
  2.30.3.1 Division of Clinical Practice 355
  2.30.3.2 Division of Employee Assistance Programs 355
  2.30.3.3 Division of Legal Consultation 355
  2.30.3.4 Division of Mental Retardation and Developmental Disabilities 355
  2.30.3.5 Division of Long-term Care 356
  2.30.3.6 Division of Practice Development 356

2.30.4 GENERAL PHILOSOPHY 356

2.30.5 STRUCTURE/ADVISORS/STAFFING 357
  2.30.5.1 Advisors/Consultants 357
    2.30.5.1.1 Accountants 358
    2.30.5.1.2 Attorneys 359
    2.30.5.1.3 Insurance agent 359
    2.30.5.1.4 Support staff 360
  2.30.5.2 Clinical Staff 361
  2.30.5.3 Continued Reading 362

2.30.6 THE FUTURE 362

2.30.1 INTRODUCTION

Clearly, there is probably more than one way to develop and maintain a practice during these troubled times in the healthcare arena; what follow are ideas, attitudes, and systems that have worked for one group practice in Ohio. The goal of this chapter is to expose the reader to the enormous potential available to a psychologist along with the techniques for the maintenance and survival of an independent practice in the 1990s. To do that, a historical view or developmental perspective, coupled with a biographical sketch, is taken. The evolution of the practice of psychology may be gleaned elsewhere in these volumes, for example, the evolution from freedom of choice legislation to ERISA to the chapter on managed care, but in this chapter, the reader will discover the need to continually "stroke"—never just to "tread water." The concept of the need for a strong working philosophy along with different processes and "products" will be discussed to illuminate the types of projects and attitudes that have been the driving force for this practice.

2.30.2 CONSIDERATIONS IN ESTABLISHING A PRACTICE

When I initially opened the doors of the practice, a number of considerations, attitudes, and preconceptions drove my decisions.

2.30.2.1 Location(s)

I wanted to be within 30 minutes of population groups to make the office(s) easily accessible.

2.30.2.2 Comprehensive Service Base

I wanted to be a full service health psychology practice provider, enabling the referral source to recognize a mental/behavioral health need, having to only think of one provider. To that end, I knew that I needed male and female staff, child and adult therapists, and assessors and therapists with interest/expertise around varying presenting problems.

2.30.2.3 Comfort and Attractiveness for Me and for the Public

After working at a county hospital with a view of an air shaft from a hospital-green cinder block 12 ft × 12 ft room, I knew that I wanted my "suite of offices" to be pleasant for the staff and clients. My expectation was that the staff and I would be spending a lot of time there and it needed to be comfortable. In addition, I wanted the office to be no more than 10 minutes from home. This allowed me to never miss a dinner with my family and never miss participating in any school or extracurricular function of my children. In addition, I found the unattractiveness of the hospital insulting to the clients.

2.30.2.4 Hours of Service

In the current day and age of two working and volunteering parents, I also believed that it was important to have evening and weekend hours to be able to accommodate clients' work schedules (as long as it was not abusive to the therapist).

2.30.2.5 Overhead Surveillance

I have found that most people do not follow their overheads or pay attention to their costs. I wanted to be able to serve people in their locations with attractive offices at convenient hours, but I also wanted to keep a close eye on costs. These are not necessarily contradictory. Systems (to be described below) were put in place from the start to allow that to happen. Multiple locations do not mean duplication of support staff or equipment.

2.30.2.6 Responsiveness to Caller

Whether the caller was a referral source or a client, I wanted the clinical staff to return calls within three hours. To that extent, everybody was required to carry pagers. When cellular phones became available, the traveling and busiest therapists all were provided with them. Each therapist remained responsible (on call) for their own caseload. For the first 15 years of the practice, I took all after-hours calls not associated with a specific staff person. At this point, the calls are handled by the Director of Clinical Practice. Responsiveness also relates to commitment to the practice, and, therefore, I am not an advocate of part-time employees who work elsewhere. (I have an appreciation of part-time employees who parent the remainder of their time. Barring clinical crises, we attempt to respect their childcare commitments.) Those who work elsewhere invariably will be torn between loyalties if a choice needs to be made. If you have a clinician who works part-time for you and is a school psychologist during the day, the client or referral source may need to speak with this clinician who is not readily available. The practice of psychology, for better or worse, functions on a more immediate, need now basis. In addition, if working part-time elsewhere, issues of multiple malpractice exposure and coverage are a concern.

2.30.2.7 Income Consideration

It should be clear to anyone that when you are earning an income based on a fee-for-service time frame, you are limited in your total gross income by the time you spend. If the overhead is fixed and the office is available, your income can increase if you have somebody else working the space for you. In a solo situation, your income would be vulnerable when you are on vacation or out marketing. If others are providing service and income is being generated, you are still free to go about your life, knowing that money is still being generated. (More will be spent on this later as we discuss contracts for outside service as a method to pay overhead.)

2.30.2.8 Values and Ethics

Do not compromise your values or your ethics for anything or anyone. Do not contract with providers or referral companies (i.e., employee assistance programs (EAPs) or managed care companies) who ask you to do things that you believe are questionable. One managed care company wanted us to submit signed HCFA 1500s to them with no other information other than the subscriber's name filled out. Do not give in to clients who attempt to manipulate you or "the system." I had an adult client's insurance salesman father inform me that his son's benefit plan was bad and that I needed to bill a second session per week instead of just the one he was seen—"who would ever know"—to allow for higher reimbursement and less out of pocket expense.

2.30.3 COMPONENTS OF A PRACTICE

In an effort to orient you to the practice about which you will be reading, let me describe the nature and form of the practice as it currently exists. The author is the sole owner and shareholder of a professional corporation, incorporated under the laws of the State of Ohio in 1978. I was trained as both an experimental and clinical psychologist, and began my career in 1968 as a committed behavioral-oriented therapist (social learning theory *a la* Bandura; cognitive behaviorist *a la* Meichenbaum, Mahoney, etc.). The practice began and remained "health psychology"-oriented (having been trained by William House, Tom Hyde, and Barbara Melamed in the early 1970s). At present, the practice consists of the following divisions or departments.

2.30.3.1 Division of Clinical Practice

This is the division of the practice that delivers clinical services in our outpatient offices. The work performed in these offices may be funded by traditional general indemnity insurance, capitated arrangements that we have had and will continue to have, managed care contracts (either in the form of PPO or EPO as we have with one large Ohio insurer for its HMO product), assessment services to national or local EAPs, etc. This division accounts for 20% of the total fee income generated. At the outset of the practice this unit produced 75% of fee income the remainder being earned by the Division of Mental Retardation and Development Disabilities.

2.30.3.2 Division of Employee Assistance Programs

Marketed as a separate contributing component of the group is a state-wide EAP division that has a number of capitated contracts to provide broad brush EAP services to company employees and eligible family members. Currently, there are seven contracted companies, covering 12 000 lives. This division accounts for 20% of the total fee income generated.

2.30.3.3 Division of Legal Consultation

This division provides psychological consultation to legal referral sources, ranging from public agencies (i.e., workers' compensation, Attorney General's office, Common Pleas or domestic relations courts, Social Security Administration [as a group we do not conduct Bureau of Disability Determination evaluations nor any longer conduct Bureau of Vocational Rehabilitation evaluations], Department of Labor) to attorneys, specializing in various aspects of the law. This division of the practice provides service throughout the state of Ohio and accounts for 20% of the total fee income generated by the group.

2.30.3.4 Division of Mental Retardation and Developmental Disabilities

Believe it or not, this was the core of the practice when it began in 1978 with a grant from the Ohio Office of Developmental Disabilities. To this day, we provide extensive consultative and direct service to facilities—public (state and county agencies) and private not-for-profit agencies—throughout the northern Ohio region. At the time of writing, facilities for the mentally retarded and developmentally disabled are under contract to the practice. When an opportunity became available to provide OBRA-PASARR evaluations, we marketed our services throughout the counties of northern Ohio, and provided coverage to 24 of the counties for this specialized need. This division in the past accounted for 50% of the fee income of the practice, but due to federal and state funding cuts has been reduced drastically to approximately 25%.

2.30.3.5 Division of Long-term Care

Staff are placed in facilities that provide long-term care services to the elderly or disabled. This division of the practice provides service throughout the entire state of Ohio, and has also been made available as a "turnkey" operation to psychologists in other states. One particular group in Utah is up and running with eight contracts since first marketing three months prior to this writing. Our practice provides the direction, forms, manuals, contracts, marketing materials, and more to get the program going, and helps to maintain the momentum in the following months through direct consultation. Within Ohio as a function of Medicare cutbacks I have "sold off" 32 of my contracts and maintained only those eight in the greater Cleveland area to allow for greater profitability resulting from lower overhead to service. This division will vary in its contribution to the total gross fee income dependent upon reimbursement rates and the local fiscal intermediary rate of pay. (On January 1, 1997 the reimbursement rate for Medicare service billed from a Cleveland provider dropped 27%.)

2.30.3.6 Division of Practice Development

The author provides consultation to psychologists nationwide on all facets of practice development and management. Consultation ranges from issues of managed care contracting, capitation, staffing, sale of practice, and long-term care development. This contributes approximately 5% of the gross revenues for the practice.

I believe that it would be helpful to set the stage for the present form of the practice by observing it developmentally. To do this, the reader will be asked to participate in an autobiographical sketch of the practice and owner. Having spent five years as a psychologist on the faculty of Case Western University School of Medicine, Departments of Pediatrics and Psychiatry, I precipitously left in December 1978. At that time, I had a state-funded grant to provide behavioral consultation and treatment to public agencies and citizens with developmental disabilities. The state invited me to take this grant into independent practice and so the doors of the "group" practice opened in December 1978 with money to subsidize the overhead. We began with four clinical staff, one support staff, and 850 square feet of office space divided into two consultation rooms where outpatient psychotherapy was conducted and one waiting room/secretarial area combined, a co-existence that lasted for many years. We began by consulting to six public and private not-for-profit facilities for the mentally retarded and developmentally disabled citizens of northeastern Ohio.

As I write today, the practice has six locations, totaling 7000 square feet where outpatient psychotherapy is conducted (three offices for which I pay the rent, three of which are satellite offices using other's space and staff), and over four dozen contracted facilities where my staff practices psychology, located throughout the entire state of Ohio. The current payroll consists of 35 individuals, 85% of whom are full-time, all of whom are considered "employees" by the Internal Revenue Service (IRS) and hence paid by a W-2. During this period of almost 18 years, the practice has had two statewide capitated contracts (inpatient and outpatient benefits for chemical dependency/substance abuse and mental health) of two years duration, servicing 25 000 members of an HMO, and one regional (northeastern Ohio) capitated contract (outpatient benefit of mental health and chemical dependency/substance abuse only), servicing 45 000 members of an HMO and lasting four years. Currently, the only form of capitation is within one specialized division of the practice, the EAP Division.

2.30.4 GENERAL PHILOSOPHY

You must have a strong, workable general philosophy for making decisions. For the first five years that I was in practice, I maintained the belief that I was a psychologist running a small business. At about that point, I began to recognize and enjoy the reality that I was a businessman with a doctorate in psychology, running a complex business. Very early I recognized that I could not do this on my own. One characteristic of the entrepreneurial process is seeing and seizing opportunities. A general business philosophy needs to be developed that drives and directs the practice so that the psychologist can remain proactive. To be reactive and attempt to play catch-up will typically leave you behind. Much of what has happened to the typical psychology practices (typically solo in days gone by) had been telegraphed ahead of time. Unfortunately psychologists allowed disbelief and denial to blind us while outside forces took control of us, and again we need to react from a defensive posture.

The general philosophy of your business needs to consist of a proactive, organized, goal-oriented cognitive statement(s) that you create. The general philosophy takes thought and planning, and is a vitally important activity for which you must make time as you plan and grow. Within this philosophy is the basis for decisions that need to be made and why they are made, such as deciding on the demographics of the populations to be served. This can be based on the location in which you wish to work (i.e., neighborhood, office-sharing with potential referral source that will help determine your practice, i.e., pediatric practice, internal medical practice, pedi-gi practice, making certain that this arrangement does not violate any Federal laws about the arrangement). For a period of time, we had an Orthodox Jewish social worker in the practice who maintained hours on Sunday. His being with us and his available hours attracted members of the large Orthodox Jewish community in Cleveland. Another decision may be to rent in a nonbarrier free building, but the consequences of that decision should be fully weighed. In many urban and suburban areas, there is the choice of locating on a mass transit/public transportation line or not. All of our managed care credentialing forms inquire about accessibility to public transportation. We always respond in the positive, yet nobody asks what percentage of the client pool utilizes mass transit. We have noted that on average less than one client per month uses a bus. If the choice is to not be on this route, then you are saying something to your potential public and are probably eliminating a segment from your clientele which can be a good decision as long as it is a well thought-out decision.

Most psychologists that I have known over the years have pieced together full-time positions for themselves. Typically, many have had positions on hospital psychiatric services or in psychiatric hospitals. In both settings, the psychologist has more often than not been used as an assessor. One can choose how and where to piece together work. Frequently, daytime school psychologists maintain evening independent practices, each position complimenting the other and rounding out the type of work and the settings for these professionals. I have also known a number of hospital-based, university-affiliated, and government-employed psychologists who maintain part-time independent practices. You can decide on the hours you wish to work and the quantity you wish to work. All of these are directly related to your conscious decision about the work philosophy, the amount of money you wish to earn, and the lifestyle you desire. In some cases, this may be more externally determined than internally. With all of the overhead required by managed care companies (i.e., credentialing, case management, time on the phone, etc.), this part-time private practice venture on the part of psychologists may no longer be viable. When our group first opened in the late 1970s, it was not difficult to fill daytime hours. Now, with two working parents and with managed care companies looking for evening and weekend hours for their subscribers, prizing that as a parameter by which they will choose you as a therapist; as a profession, psychologists tend to be working more evenings and weekends than we might wish. In addition, many of these piece-together clinical positions have allowed for psychologists to maintain nonemployee status (outside subcontractor) with the owner of the practice from whom they office share or sublet space and receive referrals. Many managed care companies will not allow for that arrangement in that they require all providers to be contracted with them individually and, if a member of a group, to be carried on the group malpractice. Therefore, the flexibility of these part-time jobs is being threatened.

As managed care began to remove the freedom of choice of providers for subscribers, much of this decision-making, self-control, and philosophy-setting goals needed to be reassessed; not necessarily thrown out, but realistically reevaluated. As noted more thoroughly below, in the past practitioners could determine their market niche and appeal directly to that specific target niche. With managed care as the predominant referral source for most of the practices that I have surveyed, there is now a new target market to which to appeal, that is to say, the case managers and other controllers from the "dark side" who can now control your client flow.

2.30.5 STRUCTURE/ADVISORS/STAFFING

It is imperative that the structure and staff of a psychology practice change and mold to the ideas, projects, needs, and circumstances that you create, control, and do not control over time. Market pressures create factors that we do not control; they, too, will affect your staffing needs. I believe that there are a number of givens that are the necessary basis of a viable group practice: advisors/consultants, office manager, support staff, and clinical staff with consideration to level of degree, certification in specialty areas, sex, age, and so on. Each of these will be briefly discussed below.

2.30.5.1 Advisors/Consultants

I strongly believe that to succeed in business you must surround yourself with competent advisors and use them often. Issues of accounting and taxes, corporate structure, employment and labor law, leases (office space, equipment, and automobiles), contracting, Board of Psychology requests (and how to respond appropriately and legally), fiscal structure, and creation of a sensible ledger system (more will become clear as we go on) are the province of other professionals. We can only expect ourselves to know our field (not just "therapy," but the business of psychology) and to know when to use consultants. The Division of Practice Development allows me to interact with psychologists across the USA. It gives me insight into how sparingly our colleagues call on what is available to them—and financially feasible—in the form of business advisors. I frequently see how unaware psychologists are about how and why to use these consultants. The two most frequently needed consultants are accountants and attorneys. Both need to be willing to familiarize themselves with mental health issues, for example, State Board of Psychology (Social Work, etc.) Code of Ethics, regulations, and guidelines. My personal experience has demonstrated that a solo practice or small firm attorney or accountant is perfect for the majority of our needs. My observations have led me to the sound conclusion that we are not good clients for the large firms in that we do not generate enough fees, the rates we are charged are typically too high for our budgets, and we can get the equivalent, if not better, attention elsewhere.

2.30.5.1.1 Accountants

My experience has been that a mid-sized firm of accountants (the one I have used for my entire practice career consists of three partners and 10 additional staff), working closely with my Office Manager, have been able to create general ledger and detailed accounting information to allow me to evaluate the productivity and profit of each division of the practice and each element within. The general ledger that is done in-house functions in such a way as to interface with the information in form and substance that the accountant needs to produce quarterly statements, payroll tax documents, and other financial information that we might wish to utilize. After being in practice for 15 years and having done payroll in-house, I finally switched to a payroll company that prepares the checks and handles the compilation of taxes. With a state-wide business, we need to be paying city tax in multiple locations, and it soon became too complicated and costly to continue to do so in-house. Although it took some time and effort to coordinate and prepare for the transition, the use of an outside service has been cost effective for both my office manager, but more importantly, for my accountant. I believe the importance of that is self-evident.

The availability of cost ledgers, allowing you to evaluate the profitability of a segment of your practice, is a vitally important tool. This is true even if you have only one segment. As you diversify and expand into various market places, you will want to assess the wisdom of continuing in a specific arena and you will need important data to do so. Let me cite three examples: (i) "Yellow Pages" advertising may have gone the way of Edsels and long-term therapy (with all due respect to M. Seligman). With the advent of managed care, the concept of "freedom of choice" has almost disappeared from the subscribers' repertoire of behaviors. There are fewer subscribers allowed to open a phone book and choose a therapist in this day and age. In 1993, over six locations (five different counties, including two different phone companies), I spent \$27 000 in Yellow Page advertising. In 1994, on a rolling basis as the contract year ended, I reduced my cost to \$13 000 and the next year to \$7000. This decision has been driven by the data we collect in the office pertaining to profitability of the money spent in a Yellow Page advertisement. It became evident that we were not paneled providers for many of the potential clients who called from the Yellow Pages and to do so would have involved the creation of another full-time support staff position in addition to continuing advertising all the while managed care companies were dropping their reimbursement rates; talk about unpaid overhead. It was also evident that in some areas callers using the Yellow Pages had the highest no show rate. Also, we began to notice that those individuals who called were those without case managers to guide them, without insurance coverage, and unable to pay our fees. To that extent, it began to feel that the Yellow Pages were no longer a viable source of referrals. With a strong emphasis placed on tracking the intake system within the office (i.e., source of referral, source of payment, number of sessions, total revenue), we could more easily evaluate the profitability of a particular cost center.

Using the same approach, you can assess the profitability of, for example, an office location, contract, or staff person. If you review your goal statement and realize that you want to make a certain predetermined profit or spend only a certain predetermined percentage of income in a specific area, then evaluating the data of cost and income will allow you to make decisions objectively. But only by being true to your philosophy, by keeping an eye to those predetermined goals and percentages will you allow yourself the opportunity to profitably continue in practice. (ii) The second example where financial ledgers were of great value in decision making was in assessing the profitability of multiple locations. As population shifts occur, as employers close or as referral sources change, the initial reason for opening in a certain location may be changing or may have changed. Instead of simply paying rent on an ongoing basis, you will want a system to evaluate the wisdom of staying in a location. One reason for maintaining a location in the absence of "specific site" profitability is to allow you to continue to serve a referral source that refers to your practice in all locations, even if one is not utilized as much. If you recall the general overriding philosophy of making it easy on your referral source, you may serve yourself better by remaining open in a location and not forcing the referral source to differentiate where to use you and where to use another practice. It remains a simple referral process if the referrer can think "psychologist" and think only of you. (iii) The third example is contract profitability. The practice has contracted with a large, for-profit nursing home chain. Of the 16 of their homes that we service two examples of profitability issues need pointing out: (a) one location is a facility smaller in size than our criteria calls for, and (b) three of the facilities are being served by a "less profitable provider" (by virtue of license, we collect less at these sites, but due to geography, we are forced to utilize this person). Both situations are counterbalanced by the desire to service their entire company and, hence, some of the sites may be loss lead item. On balance, the entire nursing home chain is profitable. This can be assessed in your financial feedback established by the accountant and you.

One last thought on the valuable use of an accountant. Frequently throughout your development you will be approached by life/disability insurance salespersons. I submit that most Ph.D. level readers of this paragraph have little ability to evaluate and integrate into your investment strategies the type and amount of insurance needed. It is at this time that your consultant accountant can also be helpful.

2.30.5.1.2 Attorneys

Another consultant of great value is the attorney. My legal needs have varied over the year, but the basic function of an attorney for most practices would include reviewing leases (office space, equipment, cars [do not forget to have first passed the idea of leasing a car by your accountant to see if it is a viable method based on your financial situation, then go to the next and probably more expensive level of advisor], etc); creating employment contracts; contracts for facilities and companies that contract with the practice; discussion with accountant regarding issues of corporate structure, benefits, pension and profit sharing plans, life insurance; and so on. I have had two primary lawyers for the life of the practice. The lawyer I have used for 13 years has functioned as the consultant attorney to the state psychological association's Committee on Ethical Practice: Needless to say, he is well-versed in issues of mental health practice law and, specifically, psychology. I have worked closely with the attorney to minimize the size and language in a legal document to cover all the issues he believes are necessary to protect my best interests and yet to not frighten the other party away. I have found that most psychologists find legal documents daunting and would rather avoid any affiliation rather then feel trapped by a document they feel they do not understand and, hence, avoid.

At one point in my career, 15 years ago, the State Board of Psychology requested a site visit to investigate our compliance with supervision guidelines. My attorney suggested that I request all questions in writing ahead of time and his presence at the site visit. The Board never made that visit. The other attorney I use practices business and tax law in another state, but is well-versed in Federal guidelines and regulations, for example, tax, retirement, insurance. I have, over the years, used him, in conjunction with my accountant, to review life insurance documents prior to purchase. The issues of life and disability insurance being held in or out of a professional corporation has been dealt with by consultation between this attorney and the accountant.

Surprisingly, I have had little use of attorneys as collection agents (I have used three professional collection agencies over the life of the practice). The only two occasions during which I used attorneys for collections have been when I was owed money from companies with whom I had contracted and they terminated the relationship, owing me funds. I used an attorney in both situations and won court orders for payment (clearly written and enforceable contracts prior to the problem helped prevail in court).

2.30.5.1.3 Insurance agent

Although many of the readers will have found the APAIT malpractice to be sufficient, I never thought that it met all of my group's needs. For years, I have had a true group with multiple locations and levels of professional staff, and I have found that I needed a local agent who could sell professional malpractice from a company that also would cover renters, fiduciary, personal property, and liability (nonprofessional) all in one policy. It made more sense to me to have one company cover everything so that in case of a claim, two or more companies did not simply point fingers and assert that it was the other's responsibility. Therefore, I have always had just one company write all of my insurance for the employees and office on one policy, and have a local agent (only two across a 19 year span) handle any of the details involved therein.

2.30.5.1.4 Support staff

It was once explained to me by a close friend that you cannot do multiple full-time functions such as being a clinician, administrator, or marketer, all within a 40–50 hour week. What has allowed me to find and initiate the projects within the practice has been an individual who has functioned as the "runner of the practice." The Office Manager implements and follows up on all aspects of the practice with both support staff and clinical staff reporting to her for all administrative issues. What has worked exceedingly well has been my "initiator spirit" and her "implementer behavior." The Office Manager has had a number of functions in the practice, all additive until they get delegated to a support staff member who reports directly to her. The Office Manager came from a background of public relations. In addition to the administrator function in overseeing the hiring of all support staff and purchasing all equipment and supplies for the practice, the Office Manager has been intimately involved in all financial aspects of the practice. She has overseen all payroll and payables function in consultation with the accountant. The creation of the general ledger and the type of computer-based accounting system was something that the two of them devised with the awareness of the type of information I believed that I needed to make decisions about the practice. The payroll function that has now been contracted out is still coordinated by the Office Manager who prepares all the data (gross wages) for the payroll company. The Office Manager works closely with the accountant to file all tax payments in a timely fashion.

Personnel issues are all handled through and by the Office Manager. Once clinical staff are hired, they deal directly with the Office Manager to request such things as vacation, sick leave, supplies, and continuing education seminars. These decisions are made and implemented either in consultation with the divisional director, myself, or by herself. In addition, the Office Manager fulfills a strong public relation function, handling all referrals and special needs from attorneys who refer to the practice. This has allowed for a personal relationship between the attorney's office and a single representative of the practice. As we developed a number of state-wide projects (managed care capitated project in 1984–86, our state-wide nursing home business currently ongoing, and our state-wide EAP), the Office Manager has functioned as the liaison with the outlying staff and the contract facilities. There is a Clinical Director for all projects who is always available to the clinical staff if that is the need, but short of that, most requests are for administrative and informational items. The Office Manager typically is not in the field and, hence, is available to those callers. There have been a number of practice consultations that I have been contracted to do that has included the Office Manager. We were hired by a local defense attorney to help in a malpractice suit against a psychologist and his Office Manager. Both of us advised the attorney after doing a site visit and interviewing the psychologist and his Office Manager. Through the Division of Practice Management, I have been to a number of states. Follow-up consultation is done by phone and frequently involves the clinician or staff consulting with the Office Manager around a variety of issues she knows more about than me, for example, computer programs, accounting functions, contract forms. To that end, her time is hired out by the contractor.

Obviously a critical support staff person is the receptionist/intake coordinator. This individual fulfills three major functions in the group. In addition to doing all phone intakes, she also schedules all appointments (save for depositions and court appearances that continue to go through the Office Manager) plus she checks on all incoming clients' benefit plans. Phone intakes in a group practice that bears an individual's name are very tricky. More often than not, the caller asks to speak to or schedule with the principal (typically, that is to whom the client is being referred). The receptionist/intake coordinator is aware of each clinician's area of specialty by training, area of interest, and so on, and so directs the caller to the appropriate clinician. These assignment duties are always done in conjunction with the Director of the clinical practice and never on her own. Scheduling for all clinicians and all locations is done in one master calendar book (not a computer) by the receptionist/intake coordinator. Although I resisted for years, my support staff convinced me that it would be better for the office if we were to check the insurance benefits prior to scheduling clients, informing them of what we were told by the insurance company. This focuses on deductibles and copays and, hence, what our financial expectation is from the client at the first and then each subsequent session. It also avoids the managed care client accessing the office, even having the first appointment, before it is apparent that the client had a particular panel to which they had to go, etc. The receptionist/intake coordinator gets all of the documentation completed for the first visit when it occurs at our main office; the clinicians are responsible for all intake documentation (e.g., Consent to Treat, Fee Agreement, Releases, etc.) when they are working in our satellite offices. The receptionist/intake coordinator is the first line in making available community resources to our EAP clients. She, along with the EAP manager, maintains an updated compilation of available community services, attorneys, financial support systems, and so on. In addition, the receptionist/intake coordinator doubles as the typist for all secretarial functions in the Division of Mental Retardation and Developmental Disabilities (producing hundreds of reports and behavioral programs per year).

We have over the years utilized an individual to function as a secretary/reception support and typist to the remainder of the practice. In addition, she has over the years shared phone duty with the receptionist/intake coordinator, preparing intakes and providing resource information to our EAP clients. She does not schedule new appointments, but does assist in such things as reschedules and cancellations.

Most critical perhaps is the billing department. Over the years, the billing function has evolved from one of the many tasks the Office Manager did to that of a single staff person's job. As the practice became more complicated and as we grew into doing a great deal of Medicare work, we developed the need for a billing department. This staff person needs exceptional computer skills and a curiosity to make a commercial program work maximally for the practice. This individual needs to understand the goals of the billing office to be more than just mailing statements (electronically or by post). The billing office is the arm of the group that has the data by which you, as the owner, will make decisions about future movement. The manner by which this information is maintained and massaged will make the decision-making function of the owner easier. Commercial billing packages can provide more information than simply meets the eye as long as the person working it is inquisitive and daring. In cooperation with the Office Manager, these two individuals can create data presentations that tell "all the news, nothing but the news" about the practice. As noted above, we have begun to spend 50% of our billing time dealing with the fiscal intermediary for Medicare in Ohio. The billing staff person must be capable of negotiating the darkest recesses of an insurance morass; and must be able to understand federal manuals plus carrier regulations, and have the personality to pursue unsatisfactory answers to attempt to guide the billing procedures of the practice.

As an adjunct to this division of the practice, the owner has established and incorporated a separate business that is a billing business for mental health professional practices. The knowledge, hard gained from working within the Medicare arena and over 20 years of insurance business, has presented an opportunity that has been seized and stroked. Remember—never tread water.

2.30.5.2 Clinical Staff

This is probably the most difficult realm to maintain, specifically with the type and style of personnel with whom you want to attract and work. If you believe that your practice has something special to offer employees (e.g., information about professional psychology, business information that you believe is proprietary, procedures and documentation that is special and yours, etc.), you will want to utilize a "covenant not to compete" in each employment contract. Most applicants for a position resist this concept and counter with the argument that they would never "steal" your referral sources, ideas, forms, or client list. The basis of an employment contract is to prohibit that from occurring. The employment status of clinical staff has always been a clear issue according to the IRS, yet seen as murky for convenience sake by our colleagues. Most "employees" (whether licensed or unlicensed) in psychology practices are by convenience treated as "outside subcontractors" ("independent contractors"), but by IRS definition they are indeed employees. This status demands that the employer take out and pay taxes on behalf of the staff. This is a practice that most psychologists have resisted. Malpractice considerations are also important at this point. You need to check with your carrier because frequently an independent contractor cannot be carried on a corporate or supervisor's malpractice. This begins to present complications for contracting and securing work for the group versus the tax and employment status of those performing and billing for the service. Medicare has made it abundantly clear that all providers who are employed by corporate structures for the purposes of contracting and billing and who are currently in the position of supervising another staff member must enjoy employee status.

2.30.5.3 Continued Reading

In addition to the professional journals that we must continue to read throughout our careers, those running "businesses" must be exposed to business publications. Some of these are specific to your locale and others are related generally to the business of running a practice. The former consists of publications such as *Crain's Cleveland Business Weekly* and *The Wall Street Journal*; the latter consists of publications such as *Psychotherapy Finances* and *Behavior Today*. Divisions within the American Psychological Association provide publications with practice information that is relevant to you both from a clinical but also business point of view. As you see new procedures developing, you are confronted by opportunities of new business niches. The information you can access in these publications exposes you to opportunities that you would probably otherwise not see. If your practice has an EAP component to it, you will read about businesses that are expanding or moving into your area. You can be aware of physician group and hospital mergers by reading these papers and see potential opportunities for your practice—all of which can be followed up on, but not all of which will come to fruition. Priortize and then watch your ledger sheets. Frequently, medical practices or hospital programs (e.g., pain and rehabilitation programs, mood disorder clinics, etc.) advertise or put out press releases in these papers, again providing you with an opportunity to expand your services. Many of these programs are in need of a psychology component for the appropriate accreditation. With the mergers and movement of for-profit hospital chains, managed care companies, and other entities entering your geographic region, there are opportunities for you to market your services, expand your scope of practice or simply meet new referral sources. These are opportunities you might not be aware of unless you are exposed to these types of publications. In some areas of the country downsizing or more euphamistically "right-sizing" has provided opportunities for psychologists to create new business ventures and seize the moment.

2.30.6 THE FUTURE

As you read this, there are changes and opportunities going on around you. For example, what began as an excellent and lucrative business venture into long-term care (i.e., nursing home work) has become a much less profitable and questionable venture for those with alternative opportunities. Watching your ledger and consulting with your accountant will tell you when to get out or at least allow you the choice to stick with the project and continue to evaluate the potential therein. The practice owner needs to be one step ahead of the next disaster to be prepared to maintain income flow and profitability. One stays aware of the field by reading and volunteering. Involvement with one's state, local, and national professional organizations is not only a responsibility that we all bear, but it is also an opportunity that is waiting to happen. While on committees or at meetings with other state or local agencies or bureaus, for example, one will see business ventures waiting to happen. While legislative chair of the state association, I was involved and instrumental in having psychologists become providers for the industrial commission of the state. From that time forth, my group has done a great deal of work for that agency. One of my colleagues is positioned on the state association's state and federal agency committee. In that position, she has seen an opportunity with the forthcoming Medicaid HMO and positioned her practice to be a major player in that arena. One gets back multifold from giving if you simply see what is in front of you as opportunities for the growth and development of your practice. After consulting to a large state-wide homehealth agency about issues of managed care, we evolved into a joint venture contractual relationship having nothing to do with managed care. Opportunities abound. The exciting aspect of change and expansion is the opportunity to learn new skills, to meet new people, and create new business opportunities. There is no question that due to outside forces our profession is dynamic and our practices must be prepared to change. To stick with what we used to do because we thought it was a good idea and it was the "right thing" may not pay the mortgage. To be flexible and attempt to work with those empowered to control the marketplace to help them understand appropriate psychological services, ethical, behavioral, and sound clinical practice works to the betterment of the client, clinician, and profession. One can do that alone, in groups, and in organizational ways. You stand a better chance of effecting change if you work through and with your national, state, and local organizations; we are a large and getting to be a loud and effective national profession with which to reckon.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.31

Technology in Practice

MARK R. McMINN
Wheaton College, IL, USA

2.31.1 INTRODUCTION 363
2.31.2 WELL-ESTABLISHED PRACTICE TECHNOLOGIES 364
2.31.2.1 Computerized Office Management 364
2.31.2.2 Facsimile 365
2.31.2.3 Telephone Answering and Voice Mail Systems 366
2.31.3 PARTIALLY ESTABLISHED PRACTICE TECHNOLOGIES 366
2.31.3.1 Computerized Assessment Tools 366
2.31.3.2 Computerized Interviewing and Diagnosis 368
2.31.3.3 Computerized Therapy Tools 369
2.31.3.4 Telephone Consultation 370
2.31.3.5 Electronic Mail 370
2.31.3.6 Clinical Databases 371
2.31.4 EMERGING PRACTICE TECHNOLOGIES 371
2.31.4.1 Teleconferencing 371
2.31.4.2 Global Networking Environments 371
2.31.4.3 Virtual Reality 373
2.31.5 TRAINING INNOVATIONS 373
2.31.6 CONCLUSION 373
2.31.7 REFERENCES 374

2.31.1 INTRODUCTION

The contemporary practice of psychology requires a prudent balance of traditional and emerging communication methods. Interpersonal interactions in the context of human relationship (e.g., speech, emotional expressions, and nonverbal gestures) have been a vital part of emotional healing throughout many centuries, and research findings in the 1990s underscore the importance of relational factors in effective psychological interventions (Whiston & Sexton, 1993). In addition to the time-honored interpersonal communication methods of professional psychology, rapid technological advances have propelled psychologists into another sphere of communication. Today's professional psychologist is increasingly expected to attain mastery in both of these communication methods—the very old and the very new.

It appears likely that technological changes will continue at a rapid or even an accelerating pace in the years ahead. This rapid rate of change suggests an ongoing need for psychologists to keep abreast of technological advances and the implications of these advances for professional practice. To this end, the American Psychological Association (APA) has published three editions of a software directory (Stoloff & Couch, 1992), the APA Practice Directorate has published *Organizing Your Practice Through Automation* (1996), national and international conferences have been established, and many journal articles and books on the topic have been published. Internet discussion groups regarding psychology and mental health practices abound, and the APA has charged the Media Psychology/New Technologies Committee of the Media Psychology Division to consider the implications of virtual reality, artificial intelligence, and other computer technologies. Psychologists are attempting to hold on to the value of established professional practices and communication methods while seeking the advantages available through technological advances.

Previous authors have distinguished between first- and second-wave innovations. First-wave technology allows for efficiency in maintaining records and office procedures but has little direct impact on clinical services. These technologies are now well-established and commonly used in the practice of psychology. They include office management computer software as well as facsimile, telephone answering, and voice mail services. Second-wave technologies have a greater direct impact on patient care and should be considered partially established in that many psychologists already use these technologies, though they will become increasingly prevalent in coming years. These include computerized assessment and interviewing tools, computerized therapeutic applications, telephone consultations, electronic mail, and clinical databases that directly impact patient care. Because of rapid technological changes, a third wave of technological advances must also be considered. This third-wave includes emerging innovations that have only started to affect the practice of psychology. Third-wave advances include teleconferencing, global networking environments (World Wide Web and Internet), virtual reality as a treatment tool, and innovative computer applications in training.

2.31.2 WELL-ESTABLISHED PRACTICE TECHNOLOGIES

A number of technological advances are now well-established as routine (first-wave) resources for professional psychologists. These include computerized office management systems, facsimile capabilities, and telephone answering and voice mail systems. The benefits and potential liabilities of each of these technologies are discussed below and summarized in Table 1.

2.31.2.1 Computerized Office Management

As early as 1987 a surprisingly high rate (90%) of psychologists surveyed in Canada were using computers for tasks indirectly related to clinical services including word processing, billing systems, and database management (Pollock & Maenpaa, 1990). It should be noted that the low response rate obtained in this survey research (23%) may have produced a spurious estimate of the number of psychologists using computers. Nonetheless, a decade has passed since the survey, and it is undoubtedly true that an increasing number of psychologists use computers in routine tasks associated with professional practice.

Word processing software allows a psychologist to reduce the time required for repetitive narrative. For example, a psychologist may use similar language in certain parts of clinical reports or treatment authorization requests. Rather than producing the narrative for each report, a template can be stored on computer disk and adapted for each individual report. Mailing merge options allow for personalized letters to be sent to various addresses for purposes of promotion or information. Also, word processing may improve the accuracy of written correspondence and reports. With tools such as spell checking, electronic thesaurus, and grammar checking, written documents can be scrutinized closely before leaving the psychologist's office. Word processing also allows for the evolving nature of forms and documents: For example, psychologists occasionally need to revise informed consent forms, release forms, or promotional statements. Word processing allows for such changes without requiring retyping.

Electronic billing software allows for the automation of the once tedious process of billing patients and health insurance carriers. Rather than typing or writing the patient's name, address, birthdate, and diagnosis on each bill, this identifying information can be stored electronically and automatically printed on a computer-generated billing form. Standard insurance forms are included in many software programs and printed onto plain paper with an inkjet or laser printer. In other cases the insurance forms are purchased separately and the printer is used to fill in the various fields in the standardized form. Most current billing programs and insurance companies also provide the option of electronic claims filing where the claim is filed via modem and no paper is transferred from the psychologist to the insurance company. Electronic billing promises to reduce much of the paper use traditionally associated with the billing process. Additional practice management software allows psychologists to keep electronic records of patient information, progress notes, financial records, accounts payable, payroll, and appointment scheduling (APA Practice Directorate, 1996).

Table 1 First-wave technologies: benefits and potential liabilities.

Technologies	*Benefits*	*Potential liabilities*
Computerized office management	Reduces time required for repetitive tasks Improves accuracy of written work Allows for evolving forms (e.g., informed consent) Billing efficiency, including reduced paper use Efficient management of patient records	Patient perceptions of increased bureaucracy Ease of duplication may compromise confidentiality and privacy rights In-session use of computers may inhibit patient–psychologist relationship
Facsimile	Immediate transfer of information (e.g., patient records)	Compromised patient confidentiality
Telephone answering and voice mail	Receive phone messages at any time without the expense of added secretarial support or answering services	Compromised patient privacy Technology failure may result in miscommunications

Although the use of first-wave computer software is commonplace in psychological practice, it remains important to consider the potentially negative effects of office computerization for patients. Because patients are generally less concerned about clerical efficiency than the nature of interpersonal interactions and quality of clinical services provided, some may experience highly visible computers as a bureaucratic hindrance or potential violation of privacy. Indeed, with the ease of duplicating electronic records there may be some legitimate basis for concern about violations of privacy and confidentiality. Also, some preliminary evidence from medical practice suggests that it may be important to avoid using computers for record keeping during the consultation itself. Physicians in the UK showed a diminished capacity to relate with patients when using a computer rather than pen and paper to issue prescriptions (Greatbatch, Heath, Campion, & Luff, 1995). Although this finding cannot be directly translated to psychological interventions, it is likely that using a laptop or notebook computer to take notes during a therapy session has an inhibiting effect on the human interactions that play an essential role in services provided by professional psychologists.

2.31.2.2 Facsimile

Facsimile (fax) machines are used to electronically transfer information to or from a remote site. The primary benefit of fax transmission is the immediacy with which information can be transfered. For example, a release form can be faxed to the office of another health care provider before holding a conversation about the patient, or patient records from a previous treatment can be instantly provided to the treating psychologist. Similarly, clinical reports can be faxed from a psychologist's office directly to the hospital floor or the referring professional's office. Contemporary computer software can make such transmissions paperless, with communications flowing from one computer to another.

The primary concern with fax transmission is the risk of compromising patient confidentiality. If a patient report is inadvertently sent to an incorrect fax station, for example, the psychologist has violated the patient's ethical and legal rights to confidentiality. This is especially likely with computer software that allows fax numbers to be stored in an electronic phonebook. Simply selecting the incorrect fax listing on the computer screen (with a single mouse click or keystroke) could result in sending a document to the wrong party and, therefore, in a violation of patient confidentiality. Even if the fax arrives at the correct office, the location and availability of the fax machine to office staff and others may compromise patient records. Because of these concerns, some jurisdictions have passed laws controlling fax transmissions of confidential materials. Thus, it is important for psychologists to know the legal guidelines of their regional or national jurisdictions. If faxing is legally permitted, it is incumbent on the psychologist to double-check the telefax number being used and to have a clear understanding of how the fax transmission will be handled on the receiving end.

2.31.2.3 Telephone Answering and Voice Mail Systems

Telephone answering systems allow callers to leave messages which can later be retrieved at the receiving end. The first generation telephone answering systems, usually depending on a cassette or microcassette recording of the incoming message, have become commonplace in many professional offices. This allows psychologists to receive incoming calls day or night without adding secretarial staff. In some cases psychologists use answering machines in lieu of an answering service. For emergency needs, an alternative phone number or electronic pager number can be left on the recorded greeting message.

The potential liabilities of using an answering machine for incoming messages include problems with patient privacy and technological failure. The patient may leave a message for the psychologist assuming the psychologist is the only one listening, especially if the greeting message is recorded by the psychologist. If a staff member or another professional in the office then listens to the recorded message, the patient may have disclosed information that was not intended for others in the office. Also, because low-cost answering machines have a relatively high failure or error rate, a patient or consulting professional may leave a message which the psychologist never receives. This breakdown in communication, caused by an uninformed faith in technology, can compromise patient care or even result in negligence in urgent situations.

Computerized voice mail systems have helped with both of these potential liabilities. Most voice mail systems have individual voice mail "boxes" which can be selected by the calling party. In addition, these boxes are usually password protected. These innovations reduce the inadvertent violations of patient privacy that frequently occur with low-cost answering machines. Computerized voice mail systems also reduce the likelihood of technological error because they depend on highly dependable digital technology rather than the mechanical and analog technologies used in free-standing answering machines.

2.31.3 PARTIALLY ESTABLISHED PRACTICE TECHNOLOGIES

A second group of technological advances are already used by many psychologists but will become more prominent in professional psychology. These second-wave technologies have more direct impact on clinical services than first-wave technologies, and thus also hold greater benefits and potential liabilities. Second-wave technologies, summarized in Table 2, include computerized assessment tools, computerized interviewing and diagnosis, computerized therapy tools, telephone consultations, electronic mail, and clinical databases.

2.31.3.1 Computerized Assessment Tools

In a survey of members of the Society for Personality Assessment and members of the Clinical Psychology Division of the APA, almost two-thirds of the respondents indicated that they use computers to assist with psychological testing (Ball, Archer, & Imhof, 1994). Over half used computers to help with test scoring and approximately one-third used computers to assist with test interpretation. Fewer than 10% used computers for test administration.

Many software programs are available for automated test scoring and interpretation. Among the most prominent test scoring and interpretation software is the Microtest Q system which generates reports for many tests, including the Minnesota Multiphasic Personality Inventories and the Millon Clinical Inventories. Two Rorschach software products are also widely used: the Rorschach Scoring Program (RSP) and the Rorschach Interpretative Assistance Program (RIAP).

One advantage of computerized test scoring for objective personality tests is improved accuracy. When Allard, Butler, Faust, and Shea (1995) had eight trained technicians and professionals hand score 43 protocols from the Personality Diagnostic Questionnaire-Revised (PDQ-R), they found 53% of the resulting profiles contained errors. Nineteen percent contained errors significant enough to affect clinical diagnosis—an alarming number of hand scoring errors. Hand scoring errors presumably affect other objective personality tests similarly, though it should be noted that the PDQ-R involves more complex hand scoring procedures than many other objective personality tests. In contrast to this high rate of human error, the error rate of computerized scoring is negligible (assuming correct data entry and software development).

Computers may also improve scoring accuracy of projective personality tests. Rorschach software such as the RSP and the RIAP provide algorithms to check for scoring inconsistencies. For example, a person may have coded both a whole blot response and a white space integra-

Table 2 Second-wave technologies: benefits and potential liabilities.

Technologies	*Benefits*	*Potential liabilities*
Computerized assessment tools	Greater accuracy in scoring	Unwarranted perception of interpretive accuracy
	Increases time efficiency	Use among unqualified users
	Access to interpretive expertise through expert system software	Changing professional identity
		Standardization assumptions may not hold with computer-administered tests
Computerized interviewing and diagnosis	More candid responses from patients	Greater difficulty in describing feelings and ideas
	Increases time efficiency	Possibility of excessive false positive diagnoses
	Increases screening of mental health needs in primary care	Must be quickly adapted to DSM changes
Computerized therapy tools	Used as an adjunctive tool for patient education, play therapy, cognitive rehabilitation, and behavior modification	Perceived as dehumanizing and reductive to those who view the human relationship as the essential ingredient of therapy
Telephone consultation	Immediate access to services	Lack of effectiveness data
	Decreases power differential between therapist and patient	Lack of nonverbal cues may inhibit emotional exploration or cause distractibility during session
	Gives privacy (or anonymity) to patient	May add to a sense of social alienation or dehumanization and thereby contribute to psychosocial problems
Electronic mail	As with telephone therapy	As with telephone therapy
		Unclear parameters for professional regulation
Clinical databases	Maintain personal outcome data to demonstrate effectiveness	Expense of equipment and on-line services
	Ready access to literature searches	Time required to input practice data

tion response but neglected to code the higher of the two z-scores. Rorschach software products detect scoring errors such as this and alert the psychologist or technician to the problem.

Computerization also reduces the time required for scoring psychological tests (Alexander & Davidoff, 1990). Whereas it might take 30–45 minutes to hand score a Millon Clinical Multiaxial Inventory, it can be computer scored in approximately five minutes. Similarly, the structural summary calculations for a Rorschach protocol may take almost an hour with a pocket calculator but only a few minutes with the appropriate software.

Test interpretation software provides psychologists with access to the interpretative expertise of test developers. Such software has been labeled "expert system" because it allows the computer to serve as an expert consultant to the psychologist. For example, Dr. John Exner has developed both the Comprehensive System for Rorschach scoring and the RIAP software that helps clinicians interpret Rorschach results. Similarly, Dr. Theodore Millon has been involved in the development of the various Millon tests (Millon Clinical Multiaxial Inventories, Millon Adolescent Clinical Inventory, and others) and in the development of the Microtest Q interpretive software. Automated interpretative systems such as these magnify the availability of experts for each test.

Despite these advantages of computerized test scoring and interpretation, legitimate concerns have been raised about these products and practices, especially regarding interpretative software. First, a computer printout may leave the psychologist with an unwarranted perception of accuracy. Although the scoring accuracy of well-reputed commercial products may generally be assumed, the interpretative accuracy may be suspect because the software developer has neither met the client being assessed nor had access to concurrent test results. Further, the interpretative narratives are often based on an expert clinician's personal experience with a test more than empirically validated results. Second, computerized test interpretation systems may unwittingly encourage the use of psychological tests among unqualified users. For example, a physician, marriage and family counselor, attorney, or social worker may attempt to bypass a psychologist's consulting services by gaining access to test interpretation software for the Minnesota Multiphasic Personality Inventory. To guard against misuse, test publishers closely monitor the qualifications of those who purchase test supplies and software. However, the publishers cannot monitor who actually uses the software once it is sold to the psychologist placing the order. Third, there is some risk that automated interpretative systems may move the identity of the professional psychologist toward that of a technician. At worst, a psychologist might administer tests, use computer-based interpretations, and write a report which includes much of the computer-generated narrative. This might involve little or no critical analysis and consideration of idiographic variables. The psychologist then would sign this report, thus taking legal and ethical responsibility for its contents, yet the computer-generated interpretative report on which the psychologist's report is based is not signed by a professional willing to share responsibility for the psychologist's conclusions and recommendations.

Lively interaction and debate will continue to play an important role in the development of professional standards for using technology in psychological assessment. Psychologists must collectively consider empirical evidence for computerized test interpretation and move forward with technological advances in a prudent manner.

Technological advances have shifted the debate from test interpretation software to automated administration of psychological tests. At the center of this debate is the question of test validity and standardization assumptions when administered by computer rather than pencil and paper.

It is unclear whether tests administered by computer produce comparable results to tests administered with pencil and paper. Sawyer, Sarris, and Baghurst (1991) reported only minor differences in Child Behavior Checklist scores when administered via computer, yet Davis and Cowles (1989) reported a greater tendency to "fake good" on a variety of personality tests administered by computer.

The burden to demonstrate that automated assessments do not significantly violate standardization assumptions rests on the profession of psychology and the individual test users. Under optimal circumstances each test would be restandardized and normative data collected for computer administration. A monumental task such as this would require many years and tremendous financial resources. But even this optimal solution may not completely address the problem because at least one study has found sizable differences among patients in the ways they respond to automated assessments (Spinhoven, Labbe, & Rombouts, 1993). Thus, renorming a test for computer administration may mask the significant individual differences in response styles. Spinhoven et al. (1993) recommend that a preassessment interview be used to determine if a patient is a good candidate for automated assessment.

The limited scope of this chapter does not permit a thorough analysis of other forms of computer-administered psychological testing. Promising areas of automated assessment include computer-assisted functional assessment (Smith, 1993), unobtrusive measurement (Meier & Wick, 1991), and behavioral assessment (Farrell, 1991).

2.31.3.2 Computerized Interviewing and Diagnosis

A variety of innovative diagnostic tools have been developed up to 1997. Some software is designed to be used as an adjunctive tool for diagnosis outside of the actual patient—psychologist interaction. Other software requires the computer to be used directly with the patient as a means for collecting information.

Software designed as an adjunctive tool for diagnosis is developed as an expert system in order to provide a structured algorithm for psychiatric diagnosis. First et al. (1993) report evidence of the effectiveness of the "DTREE: The *DSM* Expert," a computer-based expert system that guides clinicians through a series of decisions based on *Diagnostic and statistical manual of mental disorders* (*DSM*) criteria. Diagnoses obtained through a computerized system were compared with traditional diagnoses made during a case conference, resulting in a high rate of agreement, especially for patients with schizophrenia and major depression.

Software designed for direct patient use involves using the computer as interviewer in a screening interview. The self-administered Computerized Diagnostic Interview Schedule (C-DIS) is an adaptation of the original DIS, a standardized interview administered with pencil and paper. Preliminary research findings suggest that the computerized DIS produces data comparable to the original DIS (Bucholz et al., 1991; Erdman et al., 1992) and that adult

(Erdman et al.) and adolescent (Kight-Law, Mathisen, Calandra, Evans, & Salierno, 1989) patients respond favorably to computerized interviews. Additionally, there is evidence that patients disclose their concerns and symptoms more candidly in a computerized interview than in a traditional interview (Ferriter, 1993) and find a computerized interview less embarrassing (Erdman et al.).

Mini-SCID is a self-administered computerized adaptation of the Structured Clinical Interview for *DSM-III-R* (SCID) that can be used in screening for Axis I disorders. Unlike the C-DIS, Mini-SCID must be used in conjunction with the expert system DTREE (First, 1994).

A computerized version of the Diagnostic Interview for Children and Adolescents-Revised (DICA-R) has also been developed. Initial findings suggest that children find the computerized interview enjoyable and easy to use, and that reliability of results is respectable (Reich, Cottler, McCallum, Corwin, & VanEerdewegh, 1995). Comparability between the original and computerized versions of the DICA-R has not been reported up to 1997.

In addition to the obvious benefit of time efficiency and the benefit of increased patient candor discussed previously, a significant advantage of computer-based psychiatric diagnostic screening interviews is that they may ultimately help identify mental health needs earlier and more thoroughly. For example, if computerized psychiatric interviews were incorporated into primary medical care, it is likely that the mental health needs of patients would be more clearly and readily identified (Lewis, 1992).

One concern about computerized diagnostic interview tools is that patients report being better able to describe feelings and ideas to a human interviewer than to a computer (Erdman et al., 1992). In gaining efficiency with computerized interviews, there is some risk of losing the humanizing factors that have allowed professional psychology to prosper. A second concern is that the comparability of computerized interview data and human interview data has not been definitively established. Whereas two studies have shown comparability between the C-DIS and the DIS (Bucholz et al., 1991; Erdman et al.), a third study suggests that the C-DIS shows poor concurrence with the original (noncomputerized) SCID (Ross, Swinson, Larkin, & Doumani, 1994). Of particular concern is the high rate of false positive diagnoses made by the C-DIS. Finally, it should be noted that computerized interviews must be quickly adaptable to changes in *DSM* criteria. The computerized interviews described here were developed to fit *DSM-III-R* criteria, and will undoubtedly be adapted for *DSM-IV* criteria. Unfortunately, the process of changing the software could render obsolete much of the data supporting their usefulness.

2.31.3.3 Computerized Therapy Tools

Given the complexities of human interaction that are involved in effective psychotherapy, it is unimaginable to some psychologists that computers could ever be used effectively for psychotherapeutic interventions. Nonetheless, computer applications in psychotherapy have been steadily developing since the 1960s and are now used in some therapeutic contexts. The least controversial of these applications involves adjunctive educational programs to help prevent self-destructive behaviors and enhance desirable behaviors (Bloom, 1992). To illustrate, while waiting for an appointment with a psychologist a patient may find it helpful to participate in a educational computer simulation designed to enhance self-esteem, facilitate responsible sexual choices, or reduce alcohol consumption.

Others have found computers helpful to use adjunctively during the therapy session. For example, Kokish (1994) reports using a microcomputer as part of play therapy with children. Similarly, neuropsychologists and rehabilitation psychologists have found utility in using computer rehabilitation programs to restore executive functioning (Gianutsos, 1992; Giaquinto & Fiori, 1992).

Most controversial are the applications designed to replace or replicate part of the therapeutic intervention. Several programs have been designed to deliver behavioral and cognitive behavioral interventions (Bloom, 1992; Kenardy & Adams, 1993). These include the Dilemma Counseling System, designed to help patients choose between two unwanted alternatives, the Therapeutic Learning Program, used to reduce stressful interpersonal problems in the context of an educational group therapy setting, and a variety of specifically focused behavior modification programs used to treat depression, anxiety disorders, behavioral excesses, sexual dysfunction, and other problems. Some of these programs have been used in outcome studies that support their usefulness in providing therapeutic services.

Opinions about computer therapy tools will be partially determined by theoretical orientation. Those who see the human relationship as an essential part of the psychotherapeutic process are not likely to see great value in further developing therapeutic software. Moreover, they express concern that using computers as a substitute for human therapists is dehumanizing and harmful, robbing psychotherapy of the healing interpersonal process. Those who see the delivery of structured techniques as the active ingredient in psychotherapy are likely to be more encouraged by these recent software developments.

2.31.3.4 Telephone Consultation

Although telephones have been commonly available and an integral part of professional psychology for many years, interest in using telephones to deliver clinical services is a relatively recent phenomenon that is quickly gaining the attention of psychologists and professional organizations such as the APA. In a review of the literature, Haas, Benedict, and Kobos (1996) distinguish between telephone use as an adjunctive tool in therapy and telephone-only therapy. With regard to using the telephone as an adjunctive tool, telephones have been used for crisis intervention purposes, referral and screening purposes, and to provide services to established therapy clients when circumstances prevent face-to-face interactions. There is more controversy surrounding telephone-only therapy, which can involve either conventional therapeutic relationships conducted by telephone or pay-per-call services where the client is charged per minute of phone use.

One advantage is that telephone services can be obtained immediately. Rather than calling a psychologist's office to set an appointment several days or weeks in the future, a distressed patient can often make immediate contact with a psychologist offering phone services. Widespread availability of low cost cellular phone services further enhances the psychologist's availability to patients or potential patients. A second advantage is that the power differential between patient and therapist is reduced in a telephone exchange. No diplomas are visible, the patient has the power to hang up at any time, and the emotional intimacy is reduced by the absence of visual cues. Third, telephone therapy adds privacy (or anonymity, if desired) to the therapeutic encounter for those who might be concerned about being seen in the waiting room of a psychologist's office.

Each of these advantages has a corresponding disadvantage. Although there is little or no wait for telephone therapy, one might question the effectiveness of the treatment. Immediate ineffective treatment is no better than waiting for treatment. In the absence of compelling empirical effectiveness data, the question of whether or not telephone therapy works is largely determined by one's theoretical views of therapy. Second, though therapy may reduce the emotional intimacy and power differential of traditional therapy, it does so at the cost of eliminating virtually all nonverbal cues. Distractibility may be a greater problem without visual cues, and neither the therapist nor patient is able to experience the expressions or posture of the other. These factors may limit the emotional depth of exploration as well as understanding of transference and countertransference responses in the therapy process. Finally, although privacy and anonymity may be desirable to clients, one must contemplate the larger sociological dynamics that prompt this preference. Might the desire for anonymity (and the possibility that anonymous therapy can now be purchased) speak to a larger problem of social alienation and dehumanization that contributes to the same emotional problems that telephone psychologists purport to treat?

Several intriguing and promising applications for telephone therapy have been reported, including linking people with AIDS via conference calls (Rounds, Galinsky, & Stevens, 1991), providing support services for older adults (King, 1991), and even doing long-term intensive psychodynamic therapy by telephone (Spiro & Devenis, 1991). However, it is important that psychologists offering services by telephone observe all relevant ethical standards that pertain to providing psychological services. The long-term acceptability of telephone therapy will undoubtedly require persuasive empirical evidence of its efficacy.

2.31.3.5 Electronic Mail

The advent of electronic mail (e-mail), which consists of one person sending a message to another by way of a computer network, raises questions similar to those just discussed regarding telephone services. The Internet has made e-mail communication an easily accessible and global phenomenon. It is conceivable, for example, that a recognized expert in Rorschach interpretation could have a world wide consultation and supervision practice whereby psychologists from all over the world sent Rorschach protocols via e-mail and received interpretative assistance from the expert psychologist. Similarly, a psychologist might keep contact via e-mail with certain patients between therapy sessions, especially in rural settings where the traveling distances between patients and psychologists sometimes necessitate infrequent visits.

Most the advantages and disadvantages pertaining to telephone therapy also apply to e-mail services, except that even fewer cues can be retrieved from e-mail communications than from telephone (i.e., voice tone and inflection can be determined on the telephone). An additional problem, one that also applies to interstate or international telephone therapy, is the problem of professional regulation. When a psychologist in Switzerland offers supervision to a psychologist in the UK or direct services to a patient in the USA, in which country must the psychologist be licensed? To which national and state organizations is the psychologist accountable?

2.31.3.6 Clinical Databases

Although database computer software is commonly used by many psychologists to maintain patient and treatment information, additional clinical applications using database software are increasingly utilized in the practice of psychology. An excellent example is found among psychologists who maintain empirical outcome data for patients they treat (Clement, 1996). Although such databases do not lend themselves to controlled, double-blind outcome studies, they can be very useful in evaluating one's treatment efficacy with various disorders and types of clients. Moreover, as managed health care companies insist on assembling panels of competent therapists, the psychologist with positive outcome results will more likely be included on panels. Literature-based databases, such as PsychLit or various on-line services, are increasingly available for microcomputer use and can be very helpful in searching the literature for previous treatment approaches used for specific disorders. The variety and availability of other clinically relevant databases will expand exponentially as the Internet and World Wide Web grow in coming years. Potential concerns with clinical databases are limited to the expense of equipment plus on-line charges and the time required to input practice data.

2.31.4 EMERGING PRACTICE TECHNOLOGIES

In addition to the technologies discussed thus far, several other emerging technological advances may affect the practice of psychology in coming years. Third-wave technologies, summarized in Table 3, include teleconferencing, global networking environments (World Wide Web and Internet), virtual reality, and training innovations.

2.31.4.1 Teleconferencing

Teleconferencing, also known as videoconferencing, allows for visual and auditory communication through television monitors. Advances in electronics have made teleconferencing equipment increasingly affordable, resulting in experimental uses of teleconferencing among the health care professions. Whereas earlier prototypes required a satellite connection between television monitors, telecommunications can now be transmited much more economically through telephone lines.

Telecommunications are being used to offer medical care to some patients in rural areas of the USA. Technicians lead patients through diagnostic exercises while a physician observes and directs the examination from a remote location. Telemedicine saves travel time and money for many patients and makes medical care available to underserved areas. A similar concept has been developed at the University of Kansas (USA) to deliver neuropsychological services to rural areas. A master's level psychometrician leads a patient through diagnostic procedures while a neuropsychologist at the University observes and supervises the interaction. Patients have expressed satisfaction with the telecommunications services provided (Tröster, Paolo, Glatt, Hubble, & Koller, 1995). A similar strategy has also been used for remote cognitive testing of psychiatric patients in the UK (Ball, Scott, McLaren, & Watson, 1993). Teleconferencing also promises to provide an innovative means for continuing education for professional psychologists.

The primary benefit of telecommunication services is the availability of services and professional expertise to remote areas. Additionally, teleconferencing may enhance the availability of professional expertise around the world. For example, a patient in one part of the world may soon be able to consult with a leading expert in another part of the world about the symptoms he or she is experiencing. Similarly, psychologists throughout the world may participate in continuing education teleconferences with experts in various fields, thus raising the global level of professional expertise.

The main concern with teleconferencing relates to the setup costs. Although the telephone line costs of $10 to $35 per hour are not prohibitive, the site setup costs of $100 000 makes the widespread use of teleconferencing unlikely for most psychology practices, at least until the end of the twentieth century. Also, because most teleconferencing services are not yet covered by health insurance companies, there is little financial incentive for purchasing the necessary equipment.

2.31.4.2 Global Networking Environments

Computer networking allows one computer to communicate with another. The Internet is a

Table 3 Third-wave technologies: benefits and potential liabilities.

Technologies	*Benefits*	*Potential liabilities*
Teleconferencing	Can provide services to remote areas	Setup costs are high
	Will heighten the availability of professional expertise	Not yet covered by most health insurance companies
Global networking environments	May heighten awareness of mental health issues and resources May provide mechanism for supervision and consultation with other professionals	As with telephone therapy and e-mail (see Table 2)
Virtual reality	Provides the experience of *in vivo* exposure for phobias without having to leave the consulting office	Unknown generalizability of symptom relief
Training innovations	Promises to enhance learning and promote technological skills	Need to demonstrate didactic effectiveness

rapidly growing global network that allows communication among computers throughout the world. It is the electronic skeleton that allows for global e-mail communications and for the advent and growth of the World Wide Web. The World Wide Web is made possible by software innovations that allow information to be organized into "pages" which may contain text, graphics, video clips, sound, and connections with other web pages. Connections with other web pages, known as "links," usually appear as standard text printed in a distinct color. Thus, with a single mouse click the user can move from one web page to another. Because it is a global network, the user may (without knowing) be moving from a page in Australia to a page in Germany, then a page in the UK, then the USA and so on. This has led to the terms "web surfing" and "browsing." The software used to access the World Wide Web is called a web browser. Examples of web browsers include Netscape and Explorer.

With the growing use of the Internet and the user-friendly interface of the World Wide Web, many commercial services are now marketed on the Web, including psychological services. For example, Shrink-Link comprises a group of psychologists offering advice over the Internet for $20 per question.

In addition to direct services offered for profit, a number of no-cost resources are accessible through the Web. Newsgroups facilitate discussions among a variety of computers users throughout the world. Here one can find support groups for a variety of problems including depression, loneliness, addictions, and shyness. Professionals can subscribe to newsgroups or e-mail groups (called listserves) to discuss a variety of professional issues such as rural care, research funding, clinical training issues, and professional organizations. PsychNET, a service of the APA, provides free information on a variety of psychological disorders, lists and addresses of mental health organizations, and strategies for choosing a psychologist. Other organizations offer free self-help pages on a variety of topics. Some magazines and journals are available on the World Wide Web as well.

Because of the novelty and rapid growth of the World Wide Web, it is difficult to anticipate all the benefits and potential liabilities that will be forthcoming. What is clear is that the World Wide Web will continue to grow and will be a major technological influence on contemporary societies. Among the benefits for psychologists, it seems likely that the World Wide Web will make mental health information accessible to those who might not otherwise seek it. This heightened awareness of mental health needs may increase utilization of psychological services. Also, the Internet may become a mechanism for supervision and consultation among psychologists and other mental health professionals (Myrick & Sabella, 1995). The greatest liabilities of the Internet will probably be related to those who offer direct mental health services. All the concerns listed previously for telephone and e-mail consultations also apply for Internet consultations.

2.31.4.3 Virtual Reality

Virtual reality combines computer hardware and software to simulate real life experiences by presenting computer-generated graphics and sounds that are responsive to the movements and responses of the individual using the software. When using virtual reality, one wears a helmet that presents a visual field while simultaneously monitoring head movements. Moving one's head to the left, for example, would result in the visual field shifting accordingly. Many virtual reality systems also include gloves to monitor hand movements. In this way the user can pick up objects or move other objects aside while moving through the visual field.

Two reports of using virtual reality to treat acrophobia have recently been published (Rothbaum et al., 1995a, 1995b). By simulating graded exposure to heights, the researchers successfully provided many of the same benefits associated with *in vivo* graded exposure. Unlike *in vivo* interventions, virtual reality allows the treatment to occur in the office or laboratory while giving the psychologist greater control over the stimulus conditions. This appears to be a promising technology for psychology, with the only known liability being the uncertainty of generalizability to real life situations. Other applications of virtual reality in the practice of psychology are likely to be forthcoming.

2.31.5 TRAINING INNOVATIONS

Clinical training in psychology has already been influenced by the technological advances introduced by microcomputers, and it is inevitable that training advances will proliferate in the years ahead. Three types of training innovations are described here, though the list is not comprehensive.

First, one US university has developed a World Wide Web site for psychiatrists planning to take board certification examinations. The site provides sample questions and provides immediate notification of test results. A similar site for the Examination for Professional Practice in Psychology (EPPP) or other certification examinations would be helpful.

Second, a variety of simulations are available in which psychology students observe a situation on interactive videodisk and then are given options of how to best respond (see Engen, Finken, Luschei, & Kenney, 1994; Maple, 1994). Based on the option selected, they are then given more information and more choices. This type of learning environment is private enough to allow students to experiment with questionable options and observe the effects of their actions—something that would not be encouraged in real life practicum training. Many simulations are now available for training purposes, and many more are sure to come by the end of the 1990s. With the ease of newer object-oriented programming languages it is increasingly possible for psychology faculty to develop their own simulations without needing advanced programming expertise.

Third, expert system-based, computer-assisted training programs (ESCATs) utilize artificial intelligence technology to model the decision-making processes of experts in a particular subspecialty. Many aspects of professional training can be facilitated by ESCATs. For example, McMinn and Scanish (1996) developed The Rorschach Trainer, a feedback-rich tutorial program to help students master the scoring variables in Exner's Comprehensive System. There are varying levels of feedback that can be determined by the student or instructor. The Rorschach Trainer allows students to master much of the scoring system outside of the classroom, leaving more classroom time for teaching Rorschach interpretation. Todd (1996) describes the development and validation of an ESCAT to train clinicians to accurately diagnose eating disorders. As additional ESCATs are developed, specialized training may become much more interactive and practical than current lecture-oriented curriculum and continuing education programs. This could be especially important for psychologists practicing in remote areas where continuing education and in-person supervision are not readily obtained.

Appropriate and effective technological methods promise to enhance learning for students while simultaneously equipping them with technological skills to enter the changing mental health marketplace. With the rapid pace of technological advances, the biggest challenge for those developing training software will be to demonstrate the didactic effectiveness of the new training materials.

2.31.6 CONCLUSION

The tools of professional psychology are rapidly evolving in the midst of accelerating technological changes. Many new technologies can enhance the training and professional services offered by psychologists, especially for those who remain commited to the traditional tools, values, and standards of the profession.

2.31.7 REFERENCES

Alexander, J. E., & Davidoff, D. A. (1990). Psychological testing, computers and aging. *International Journal of Technology and Aging, 3,* 47–56.

Allard, G., Butler, J., Faust, D., & Shea, M. T. (1995). Errors in hand scoring objective personality tests: The case of the Personality Diagnostic Questionnaire-Revised (PDQ-R). *Professional Psychology: Research and Practice, 26,* 304–308.

American Psychological Association (1996). *Organizing your practice through automation: Managing information and data.* Washington, DC: Author.

Ball, C. J., Scott, N., McLaren, P. M., & Watson, J. P. (1993). Preliminary evaluation of a low-cost videoconferencing (LCVC) system for remote cognitive testing of adult psychiatric patients. *British Journal of Clinical Psychology, 32,* 303–307.

Ball, J. D., Archer, R. P., & Imhof, E. A. (1994). Time requirements of psychological testing: A survey of practitioners. *Journal of Personality Assessment, 63,* 239–249.

Bloom, B. L. (1992). Computer-assisted psychological intervention: A review and commentary. *Clinical Psychology Review, 12,* 169–197.

Bucholz, K. K., Robins, L. N., Shayka, J. J., Przybeck, T. R., Helzer, J. E., Goldring, E., Klein, M. H., Greist, J. H., Erdman, H. P., & Skare, S. S. (1991). Performance of two forms of a computer psychiatric screening interview: Version 1 of the DISSI. *Journal of Psychiatric Research, 25,* 117–129.

Clement, P. W. (1996). Evaluation in private practice. *Clinical Psychology: Science and Practice, 3,* 146–159.

Davis, C., & Cowles, M. (1989). Automated psychological testing: Method of administration, need for approval, and measures of anxiety. *Educational and Psychological Measurement, 49,* 311–320.

Engen, H. B., Finken, L. J., Luschei, N. S., & Kenney, D. (1994). Counseling simulations: An interactive videodisc approach. *Computers in Human Services, 11,* 283–298.

Erdman, H. P., Klein, M. H., Greist, J. H., Skare, S. S., Husted, J. J., Robins, J. N., Helzer, J. E., Goldring, E., Hamburger, M., & Miller, J. P. (1992). A comparison of two computer-administered versions of the NIMH diagnostic interview schedule. *Journal of Psychiatric Research, 26,* 85–95.

Farrell, A. D. (1991). Computers and behavioral assessment: Current applications, future possibilities, and obstacles to routine use. *Behavioral Assessment, 13,* 159–179.

Ferriter, M. (1993). Computer aided interviewing and the psychiatric social history. *Social Work & Social Sciences Review, 4,* 255–263.

First, M. B. (1994). Computer-assisted assessment of *DSM-III-R* diagnoses. *Psychiatric Annals, 24,* 25–29.

First, M. B., Opler, L. A., Hamilton, R. M., Linder, J., Linfield, L. S., Silver, J. M., Toshav, N. L., Kahn, D., Williams, J. B. W., & Spitzer, R. L. (1993). Evaluation in an inpatient setting of DTREE, a computer-assisted diagnostic assessment procedure. *Comprehensive Psychiatry, 34,* 171–175.

Gianutsos, R. (1992). The computer in cognitive rehabilitation: It's not just a tool anymore. *Journal of Head Trauma Rehabilitation, 7*(3), 26–35.

Giaquinto, S., & Fiori, M. (1992). THINKable, a computerized cognitive remediation: First results. *Acta Neurologica, 14,* 547–560.

Greatbatch, D., Heath, C., Campion, P., & Luff, P. (1995). How do desk-top computers affect the doctor–patient interaction? *Family Practice, 12,* 32–36.

Haas, L. J., Benedict, J. G., & Kobos, J. C. (1996). Psychotherapy by telephone: Risks and benefits for psychologists and consumers. *Professional Psychology: Research and Practice, 27,* 154–160.

Kenardy, J., & Adams, C. (1993). Computers in cognitive-behaviour therapy. *Australian Psychologist, 28,* 189–194.

Kight-Law, A., Mathisen, K. S., Calandra, F., Evans, F. J., & Salierno, C. A. (1989). Computerized collection of mental health information from emotionally disturbed adolescents. *Computers in Human Services, 5,* 171–181.

King, H. (1991). A telephone reassurance service: A natural support system for the elderly. *Journal of Gerontological Social Work, 16,* 159–177.

Kokish, R. (1994). Experiences using a PC in play therapy with children. *Computers in Human Services, 11,* 141–150.

Lewis, G. (1992). Computers in primary care. *International Review of Psychiatry, 4,* 307–310.

Maple, F. F. (1994). The development of goal-focused interactive videodiscs to enhance student learning in interpersonal practice methods classes. *Computers in Human Services, 11,* 333–346.

McMinn, M. R., & Scanish, J. D. (1996). *The Rorschach Trainer.* Lutz, FL: Psychological Assessment Resources.

Meier, S. T., & Wick, M. T. (1991). Computer-based unobtrusive measurement: Potential supplements to reactive self-reports. *Professional Psychology: Research and Practice, 22,* 410–412.

Myrick, R. D., & Sabella, R. A. (1995). Cyberspace: New place for counselor supervision. *Elementary School Guidance & Counseling, 30,* 35–44.

Pollock, N. L., & Maenpaa, M. (1990). How are psychologists using computers? *Canadian Psychology, 31,* 167–171.

Reich, W., Cottler, L., McCallum, K., Corwin, D., & VanEerdewegh, M. (1995). Computerized interviews as a method of assessing psychopathology in children. *Comprehensive Psychiatry, 36,* 40–45.

Ross, H. E., Swinson, R., Larkin, E. J., & Doumani, S. (1994). Diagnosing comorbidity in substance abusers: Computer assessment and clinical validation. *The Journal of Nervous and Mental Disease, 182,* 556–563.

Rothbaum, B. O., Hodges, L. F., Kooper, R., Opdyke, D., Williford, J. S., & North, M. (1995a). Virtual reality graded exposure in the treatment of acrophobia: A case report. *Behavior Therapy, 26,* 547–554.

Rothbaum, B. O., Hodges, L. F., Kooper, R., Opdyke, D., Williford, J. S., & North, M. (1995b). Effectiveness of computer-generated (virtual reality) graded exposure in the treatment of acrophobia. *American Journal of Psychiatry, 152,* 626–628.

Rounds, K. A., Galinsky, M. J., & Stevens, L. S. (1991). Linking people with AIDS in rural communities: The telephone group. *Social Work, 36,* 13–18.

Sawyer, M. G., Sarris, A., & Baghurst, P. (1991). The use of a computer-assisted interview to administer the Child Behavior Checklist in a Child Psychiatry Service. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 674–681.

Smith, R. O. (1993). Computer-assisted functional assessment and documentation. *The American Journal of Occupational Therapy, 47,* 988–992.

Spinhoven, P., Labbe, M. R., & Rombouts, R. (1993). Feasibility of computerized psychological testing with psychiatric outpatients. *Journal of Clinical Psychology, 49,* 440–447.

Spiro, R. H., & Devenis, L. (1991). Telephone therapy: Enhancement of the psychotherapeutic process. *Psychotherapy in Private Practice, 9,* 31–55.

Stoloff, M. L., & Couch, J. V. (Eds.) (1992). *Computer use in psychology: A directory of software* (3rd ed.). Washington, DC: American Psychological Association.

Todd, L. K. (1996). A computer-assisted expert system for clinical diagnosis of eating disorders: A potential learning tool for practitioners. *Professional Psychology: Research and Practice, 27,* 184–187.

Tröster, A. I., Paolo, A. M., Glatt, S. L., Hubble, J. P., & Koller, W. C. (1995). "Interactive video conferencing" in the provision of neuropsychological services to rural areas. *Journal of Community Psychology, 23,* 85–88.

Whiston, S. C., & Sexton, T. L. (1993). An overview of psychotherapy outcome research: Implications for practice. *Professional Psychology: Research and Practice, 24,* 43–51.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.32

The Evolution of Canada's Health-care System and its Implications for the Practice of Psychology as a Health Profession

PIERRE L.-J. RITCHIE and HENRY P. EDWARDS
University of Ottawa, ON, Canada

2.32.1 INTRODUCTION 378

2.32.2 THE ORGANIZATION OF PROFESSIONAL PSYCHOLOGY IN CANADA 378
- *2.32.2.1 Provincial/Territorial Organizations* 379
  - *2.32.2.1.1 Regulatory bodies* 379
  - *2.32.2.1.2 Societal bodies* 380
- *2.32.2.2 National Organizations* 380
  - *2.32.2.2.1 Canadian Psychological Association* 380
  - *2.32.2.2.2 Council of Provincial Associations of Psychologists* 380
  - *2.32.2.2.3 Canadian Register of Health Service Providers in Psychology* 380

2.32.3 HISTORY AND CURRENT STATUS OF CANADA'S HEALTH-CARE SYSTEM 381
- *2.32.3.1 The Broader Context* 381
- *2.32.3.2 Evolution of Canada's Health-care System* 381
- *2.32.3.3 Cooperative Federalism in Action* 382
- *2.32.3.4 A Public–Private Mixed Model of Health-care Financing* 383

2.32.4 HEALTH-CARE POLICY AND FINANCING IN RELATION TO PSYCHOLOGY 384
- *2.32.4.1 Psychology and Health-care Policy* 384
- *2.32.4.2 Psychology and Health-care Financing* 384

2.32.5 HEALTH-CARE REFORM AND ITS IMPLICATIONS FOR PSYCHOLOGY 385
- *2.32.5.1 Elements of a Renewed, Publicly Funded Health-care System* 385
- *2.32.5.2 Private Sector Initiatives* 386
- *2.32.5.3 Globalization* 386
- *2.32.5.4 Determinants of Health* 387
- *2.32.5.5 Scope of Health Care* 387

2.32.6 THE FUTURE OF PSYCHOLOGY AS A HEALTH-CARE PROFESSION IN CANADA 387
- *2.32.6.1 A Global Road Map: The Mississauga Conference* 387
- *2.32.6.2 Primary Care: An Illustration of Psychology's Future as a Health Service Profession* 388
  - *2.32.6.2.1 Primary health-care reform as a public policy issue* 388
  - *2.32.6.2.2 Principles for professional psychologists providing primary care* 389
  - *2.32.6.2.3 Competencies brought by psychologists to primary care* 389
  - *2.32.6.2.4 Funding for psychologists providing primary care* 389

2.32.7 CONCLUSIONS 390

2.32.8 REFERENCES 390

2.32.1 INTRODUCTION

This chapter discusses the evolution of psychology as an organized profession in Canada, the development and financing of Canada's health-care system, the relationship of health-care policy to the practice of psychology and, in the final section, future directions, with primary care reform as an illustration of trends that will help shape the future of psychology as a major health-care profession in Canada. Professional psychology is presented in historical perspective in relation to Canada's health care system.

There has long been a dichotomy between Canadian psychology as a profession and as an academic/scientific discipline. A cohesive national strategy for the profession has been lacking, due in part to the structure of Canadian society, as well as the nature and functions of national and provincial/territorial psychology associations.

By way of framework, Canada's version of constitutional federalism, which makes this country one of the most decentralized federations in the world, is reviewed in order to show how Canada's health-care system evolved within it. The Canada Health Act of 1984, the legislative cornerstone of Canada's publicly funded national health care system (Medicare), is assessed in this context. Brief comparative reference is made to other modern health-care systems.

Health policy is discussed next in relation to psychology as a health-care profession. For funding purposes, psychology has been associated traditionally with mental health services. Fee-for-service psychological services in private offices have not been funded by Medicare, although psychologists most often are paid from public funds for services provided in public facilities (e.g., hospitals). In general, Canadian health-care policy has not addressed psychological services explicitly. This has limited the public access, especially to primary care services provided by psychologists. Alternative methods of payment for psychological services are under consideration.

The concluding section addresses the future. For all the health professions, primary care reform is already a major issue throughout Canada. A strong emphasis on primary care is of considerable interest to Canadian psychologists, most of whose clinical services fall within definitions of primary care. In addition, the reduction of domestic and international trade barriers will likely prompt greater mobility and lead to greater emphasis on empirically supported approaches and competency-based credentialing.

2.32.2 THE ORGANIZATION OF PROFESSIONAL PSYCHOLOGY IN CANADA

A global review of psychology in Canada was recently provided by Adair, Paivio, and Ritchie 1996). It notes that the development of Canadian psychology as a profession has been distinct from its evolution as a discipline. This dichotomy, which has existed since at least the 1940s, as well as the tensions and debates surrounding it, have been captured in the reports of major discipline reviews (Appleby & Rickwood, 1968; Berlyne, Black, Berry, & Douglas, 1971; Bernhardt, 1961; Ferguson, 1977; Macleod, 1955; Ritchie, Hogan, & Hogan, 1988a; Webster, 1967). It was only in 1994 that the convocation of a national conference devoted exclusively to Canadian professional psychology (Dobson & King, 1995) heralded a new era in its maturation.

Lacking a cohesive national strategy in its organizations and in the realm of public policy, professional psychology has developed piecemeal even though Canadian scholarly contributions to both foundational and applied psychology have been strong and cohesive. Adair et al. (1996) cited the structure of Canadian society as offering a partial explanation. Canada is the world's largest jurisdictionally defined land mass, with a small population base, two official languages, multiple cultures, and a decentralized political system in which the provinces are constitutionally responsible for health services. These features make any national undertaking difficult indeed.

As noted by Adair et al. (1996), events internal to the discipline also contributed to the history and status of Canadian professional psychology. The Canadian Psychological Association (CPA), unlike its counterpart in the United States, largely distanced itself from provincially-based efforts to establish psychology as a human services profession. Only in the 1970s did CPA initiate a slow and difficult shift toward greater involvement in professional issues. Since then, substantial achievements have taken place (Dobson & Dobson, 1993), although tensions between professionals and traditional scientists remain (Craig, 1993; Ritchie & Sabourin, 1992). A promising trend acknowledges the presence of multiple psychologies in Canada, stresses the links of applied psychology to university-based empirical traditions, and advocates for the strengthening of common ground between the practice and science of psychology (e.g., Dobson, 1995; Wand, 1993). This perspective is well anchored in the value that Canadian society attaches to pluralism.

Contemporary professional psychology organizations in Canada are described by Craig (1993). Wright (1971) captured their early history. They may be divided into two types: regulatory organizations with a statutory mandate of public protection, and voluntary societal organizations whose primary mission is to advance the profession. All regulatory bodies are established by and accountable to provincial/territorial legislatures, in keeping with the Canadian constitution. Some have dual regulatory and societal functions. Nationally, there are membership-based societal organizations and consortia or federations of provincial/territorial and national bodies.

2.32.2.1 Provincial/Territorial Organizations

Professional psychology organizations have been established in the 10 provinces and in one of the two territories. In two provinces (New Brunswick, Quebec) and in the Northwest Territories, regulatory and societal functions are combined in the same body. In eight provinces (Alberta, British Columbia, Manitoba, Newfoundland, Nova Scotia, Ontario, Prince Edward Island, Saskatchewan), there are distinct regulatory and societal bodies.

2.32.2.1.1 Regulatory bodies

Approximately 13 000 psychologists and psychological associates are regulated in Canada for the practice of psychology. Statutory professional regulation is decentralized by province/territory. By statute, the regulatory bodies typically exercise five major functions: determination of eligibility criteria for registration (e.g., degree, curriculum, supervised experience, examinations); registration of persons who meet the eligibility criteria; maintenance of a register of regulated psychologists and, where applicable, psychological associates; investigation of complaints against registrants; and discipline for incompetence or misconduct. In addition, they may develop and promulgate guidelines and standards of professional conduct, carry out fitness to practice investigations, monitor individual practices by means of mandatory or voluntary peer reviews, and offer or sponsor continuing education activities. The earliest regulatory board was The Ontario Board of Examiners in Psychology, established in 1960 and renamed The College of Psychologists of Ontario. The most recent regulatory board is Prince Edward Island's, established in 1990.

The most important requirements for initial registration of psychologists are an acceptable graduate university degree, supervised experience, and successful completion of written and oral examinations. A degree in psychology is normally required, although some jurisdictions may accept a degree from a program whose content is primarily psychological, even if the degree title does not include the word "Psychology." Four provinces (British Columbia, Manitoba, Ontario, Saskatchewan) require the doctorate; four (New Brunswick, Newfoundland, Nova Scotia, Prince Edward Island) specify the doctorate but allow registration on the basis of a Masters degree subject to prescribed conditions or for a limited period of time; two (Alberta, Quebec) and one territory (Northwest Territories) require a Masters degree. Specific curriculum requirements are variable.

Supervised experience requirements also vary across jurisdictions. Doctoral entry jurisdictions typically require one or two years of supervised experience, one predoctoral and/or one postdoctoral. Among the Masters entry jurisdictions, Alberta requires 900 hours of supervised postdegree experience; New Brunswick four years; Newfoundland two years postdegree until the year 2001, when the doctoral entry standard becomes effective; the Northwest Territories one year; Nova Scotia six years; and Prince Edward Island two years postdegree. In these jurisdictions, the supervised experience requirements for applicants with doctorates are generally similar to those of doctorate-only jurisdictions. Quebec, however, the province with the largest number of Masters psychologists in Canada, does not specify a supervised postdegree experience requirement. The Examination for the Professional Practice of Psychology is a requirement in all jurisdictions except the Northwest Territories, Prince Edward Island, and Quebec. Oral examinations are also used in all jurisdictions except Newfoundland, the Northwest Territories, and Quebec. In addition to registering psychologists, three provinces (Manitoba, Ontario, Saskatchewan) also register Masters level "Psychological Associates." In Ontario, the psychological associates are autonomous practitioners, except that they may not communicate a diagnosis unless this has been formally and individually delegated. In Manitoba and Saskatchewan, they practice under supervision.

The above picture shows considerable diversity across Canada. Indeed, nowhere is the historic absence of a cohesive national strategy for the development of professional psychology more evident than in the variable education and training entry levels for the practice of psychology across provinces and territories. Despite its repeatedly expressed intent to resolve this problem, organized psychology has been unable to do so. Greater uniformity in entry level criteria for psychologist registration may be achieved, however, due largely to the Agreement on Internal Trade (AIT) and its implications for the regulated professions including psychology. This wide-ranging trade accord was concluded in 1994 by the federal, provincial, and territorial governments. It proposes to eliminate barriers to the free mobility of regulated professionals within Canada, and it outlines processes both for the mutual recognition of professional qualifications and for the development of uniform entry level criteria across the country.

2.32.2.1.2 Societal bodies

The provincial/territorial societal bodies have been vital in promoting the development and advancement of professional psychology in Canada (Adair et al., 1996). The British Columbia Psychological Association was the first to be established, in 1938. World War II precluded further developments, until the Ontario Psychological Association (OPA) was founded almost 10 years later in 1948 (Wright, 1971). Thereafter, voluntary membership-based associations were created throughout Canada, before the advent of regulatory bodies. These associations were most influential in the drive to secure statutory recognition for psychology.

As societal bodies, the associations engage in advocacy and lobbying, provide direct benefits to their members, organize meetings, and offer or organize continuing education and other competence enhancing activities. By mandate, they protect the profession and help to secure its future. Their membership ranges from approximately 40 in the Prince Edward Island Psychological Association to more than 1200 in OPA.

2.32.2.2 National Organizations

2.32.2.2.1 Canadian Psychological Association

CPA was created by 38 founding members in 1939. With 3000 members and 1000 student and other affiliates, it is the oldest national association. Adair et al. (1996) describe six functions provided by CPA: dissemination of original research and informed commentary by means of its three journals, quarterly newsletters, and annual meeting; accreditation of doctoral programs and internships in clinical psychology, counseling psychology, and neuropsychology; sponsorship of member benefits; development and promulgation of a code of ethics as well as several guidelines and standards of practice; advocacy of selected issues pertinent to practitioners and scientists, often in collaboration with other organizations; and attention to the subdisciplines of scientific/academic and applied psychology through formally organized sections and interest groups.

2.32.2.2.2 Council of Provincial Associations of Psychologists

The Council of Provincial Associations of Psychologists (CPAP) consists of both regulatory and societal bodies (Craig; 1993; Wand, 1990). CPAP is a federation that evolved from an advisory committee created within the CPA in 1968 (Berry, Davidson, & Gibson, 1974). It became an autonomous entity in 1983, with the CPA as an affiliate member, and was legally incorporated in 1995. Its primary purpose is to share information and resources among provincial/territorial bodies, and to facilitate national undertakings in the common interest of its member bodies. It fosters understanding, serves as a forum for debate on issues which are ultimately decided elsewhere, and ensures that provincial/territorial matters are duly considered by national organizations.

2.32.2.2.3 Canadian Register of Health Service Providers in Psychology

The establishment of the Canadian Register of Health Service Providers in Psychology (CRHSPP) in 1985 was prompted by two factors, client/patient access and the need for a national credential. In the previous decade, psychologists providing health services were increasingly frustrated in their efforts to secure enhanced access. Funding, particularly of services provided in independent practice settings, was at the source of the accessibility problem. Policymakers in both the public and private sectors increasingly stressed that they would only pursue the possibility of expanded funding once the profession had implemented a credible means of identifying and credentialing those psychologists who had the appropriate training and experience to offer health services. The legal right to practice, while a prerequisite, was an insufficient mechanism since psychologist registration across Canada is typically generic. The fact that the practice of psychology is regulated at the provincial/territorial level highlighted both the need for a national credentialing body for health service providers in psychology and the challenge of developing uniform criteria given varying regulatory standards. The historical development of CRHSPP is reviewed by Wand (1990) and Craig (1993), and the extent to which its purpose and structure were influenced by the National Register of Health Service Providers in

Psychology in the United States is described by Ritchie (1995a). Less than a decade after its establishment, Craig (1993) noted that CRHSPP is being increasingly recognized by governments, insurance companies, and quality assurance agencies.

Unlike its American counterpart, the Canadian Register's governance is determined by its corporate members, the provincial/territorial bodies of psychology, and CPA, each of which is entitled to nominate one person to its Board of Directors. CRHSPP lists more than 3000 psychologists and publishes an annual directory. Eligibility criteria for listing are based on training, experience, and statutory registration or licensure as psychologist in a province, territory, or non-Canadian jurisdiction.

CRHSPP also plays an important role in promoting the public's access to qualified health service psychologists, and it collaborates with its corporate members, other health professions in Canada, and international bodies in furthering the development of the profession. Research on issues related to the delivery of health services and continuing education complement its broader objectives. In this manner, its strategic priorities contribute to Canadian professional psychology as a major health profession (Hurley & Ritchie, 1995).

2.32.3 HISTORY AND CURRENT STATUS OF CANADA'S HEALTH-CARE SYSTEM

Canadian federalism has produced a publicly funded health-care system that is primarily under provincial/territorial control. Nonetheless, the federal government's funding role has enabled it to shape the main features of a national, publicly funded health-care system. This section traces its history, legislation, and the impact of federal funding cuts. It concludes by noting that the mix of private and public financing of the Canadian system often is not well understood.

2.32.3.1 The Broader Context

Canada's health-care system has been the object of much domestic and external scrutiny since the mid-1980s. In the early 1990s, the prospect of health-care reform in the USA generated substantial interest in the Canadian system (e.g., Adams, 1994; Haislmaier, 1992), including mental health issues (e.g., Freeman, 1994; Torrey, Bigelow, & Sladen-Dew, 1993) and factors pertinent to psychological practice (e.g., Wedding, Ritchie, Kitchen, & Baner, 1993). Some assessments have been more critical (e.g., Danzon, 1992; Enthoven, 1990), while others have been more positive (e.g., Barer & Evans, 1992). Similar to other industrialized nations which had developed a substantial publicly funded health-care system, Canada experienced strong internal pressures to embark on its own reforms. In the short-term, this was driven largely by immediate fiscal objectives, especially that of reducing the level of public debt. Longer term perspectives, however, included concern about the sustainability of a publicly funded universal health-care system, particularly in the context of an increasingly aged population.

From an international vantage point, the creation of the UK National Health Service in 1948 is often regarded as a major benchmark in publicly funded health care. Actually, the forerunner of the modern era dates back to 1883, when the Social Sickness Fund was established in Germany. Before the close of the nineteenth century, variations were implemented in Belgium (1894) and France (1898). Luxembourg followed in 1901. Greece, Ireland, Italy, the Netherlands, Portugal, and the UK all had a form of mandatory health insurance prior to World War II. The UK National Health Service, however, introduced universality, central financing, and public administration as features which would be adopted two decades later in Canada.

2.32.3.2 Evolution of Canada's Health-care System

To appreciate the status of the Canadian health-care system as well as to understand psychology's place within it requires some knowledge of how Canada's health-care system evolved. Canada is characterized by a largely confederal system in which substantial powers are held by the constituent provinces. Although executive federalism, exercised by the federal government, has increased greatly since World War II, Canada remains one of the most decentralized federations in the world. Canada's constitutional regime has been the object of continuing intense debate within Canada.

To understand the politics and policies which influence the nature of Canada's health-care system, it is essential to know that the administration of health care is a provincial responsibility under the Canadian constitution. The federal government, however, is unfettered in its authority (known as the federal spending power) to disburse money. Hence, the Canadian federal government can make funds available to the provinces for health care (or other areas within provincial responsibility such as welfare and various social assistance programs). In doing so, it can also set conditions for the use of these funds, notwithstanding that it has no authority to administer or deliver the actual services. The only exception to this constraint applies to persons in the direct employ of the federal government (e.g., members of the armed forces) or whom the constitution has placed under exclusive federal responsibility (e.g., aboriginal Canadians living on federal reservations).

In practice, nationwide programs and national standards have developed in Canada only when the federal government has made it attractive financially for the provinces to agree implicitly to relinquish some of their autonomy. Nonetheless, the provinces always retain regulatory control for areas within their constitutional jurisdiction. Hence they can, and sometimes do, prevent institutions (e.g., hospitals) under their authority from receiving federal funds. To avoid such stalemates, Canada has evolved a form of cooperative federalism whereby the federal and provincial governments typically establish formal agreements concerning the use of federal funds in areas of provincial jurisdiction. An important feature of such agreements is that the funds are typically transferred directly to the provincial governments. These funds may be conditional (i.e., subject to meeting established criteria) or unconditional. The federal government's financial participation in health care has led to the transfer of federal funds to the provinces/territories subject to more specific conditions than in any other area of federal/provincial transfer (e.g., education).

Deber (1992) summarized the early history of the federal government's involvement in health and the initial federal-provincial arrangements which preceded the current system. The federal parliament first investigated and reported on sickness insurance in 1928. In 1935, it made its first legislative attempt to mandate the collection of information on health insurance plans, as part of a broader employment and social insurance act which subsequently was found to be unconstitutional. A health insurance bill was drafted during World War II, and the federal government offered to contribute federal funds toward a comprehensive health insurance plan as part of its Reconstruction proposal to the provinces. A conference was convened to address the Reconstruction proposal, but it could not agree on the fiscal elements of the social security proposals of which health was a part.

The federal government resorted to its spending power in 1948 to fund general health grants. However, some provinces also chose to proceed on their own, especially in establishing hospital insurance plans. Health insurance returned actively to the agenda of the annual Conference of First Ministers (which brings together the federal Prime Minister and their provincial counterparts) in 1955. This culminated in the first major federal health legislation in the modern era, the Hospital Insurance and Diagnostic Services Act (HIDS), which was implemented in 1958. Consistent with the constitution, administration of services remained under the direct or regulatory control of the provinces but receipt of federal funds was made contingent on meeting specified conditions. Eligible services had to meet tests of comprehensiveness (linked to the notion of "medically necessary"), portability, public administration, and universality. Imprecise definitions (especially of "medically necessary," a problem for the interpretation and implementation of all subsequent federal-provincial health funding arrangements), were a feature of the HIDS. Nonetheless, it marked a turning point in moving toward a national health-care system.

2.32.3.3 Cooperative Federalism in Action

The zenith of cooperative federalism regarding a national health-care system (and arguably regarding any federal-provincial matter) was reached in 1965–66 with the passage of the National Medical Care Insurance Act (MCA) which commonly came to be known as "Medicare." It received the support of all provinces and was implemented in 1968. It retained the four criteria of the HIDS but substantially expanded the range of services covered to include virtually all services provided by physicians regardless of setting (inpatient or office based). Problems of interpretation and implementation remained.

Although provinces varied in the flexibility or rigidity of the options offered to physicians for participation in the Medicare system, the outcomes were similar across the country. Overwhelmingly, physicians joined the system. Soon, 98% of all physician services were reimbursed through the various provincial plans set up under the provisions envisioned by the federal legislation, a level which has remained steady since the 1960s, growing to 99% by 1994 (National Forum on Health, 1997). Hospital services also remained largely covered by the public system, varying between 90 and 94% of all costs from the 1970s (Federal, Provincial, and Territorial Advisory Committee on Population Health, 1996).

Other professional services were funded differently. In this respect, psychologists are situated similarly to most health providers other than physicians. Inpatient and outpatient services provided within or through public facilities are covered, although the facilities and the provincial/territorial governments that fund them have full discretion on which nonphysician services they will cover. Furthermore, unlike physicians, other health professionals including psychologists have not been reimbursed through the public purse for community services delivered in their private offices. These have been funded through a range of private mechanisms including extended health plans underwritten by insurance companies.

The last major piece of federal legislation in this area, the Canada Health Act (CHA), was passed in 1984. It was essentially similar to the MCA which it replaced. Its major purpose was to strengthen the enforcement provisions available to the federal government with respect to the five principles set out in the legislation. The four basic historic criteria were retained as well as the more precise definition of accessibility. The CHA is the current legislative cornerstone of the national health-care system.

In the two decades between the MCA and the CHA, important changes occurred in the nature of the federal government's financial participation in health care. It shifted from an open-ended 50–50 shared-cost formula linked to provincial spending to a complex system of cash and tax transfers called Established Programs Financing Arrangements (EPF) introduced in 1977. In principle, two-thirds of this federal-provincial transfer were flagged for health and one-third for postsecondary education.

In the mid-1980s, the federal government introduced increasingly stringent measures to address an escalating level of public debt. These arbitrary measures included a dramatic reduction in the transfers to the provinces under the provisions of the EPF. In concert with the federal government's earlier decision to constrain provincial discretion in the allocation of federal funds earmarked for health when it adopted the CHA, the progressive decrease in the transfer of federal funds has also yielded recurrent discord in federal-provincial relations. The impact is partly reflected in the decrease in gross domestic product (GDP) attributable to health expenditures. After rising from 7.4% in 1971 to more than 10% in the early 1990s, the proportion spent on health declined to 9.5% in 1995 and is projected to decrease a bit further. This proportion is comparable to that spent in several other Organization for Economic Co-operation and Development (OECD) countries (e.g., Austria, France, Germany), but significantly less than that spent in the USA, where it has risen from 7.6% to more than 14% of GDP in the same 25 years. Given that the provinces are the direct payers of facilities and professionals (largely physicians) reimbursed by the public system, it is noteworthy that health accounted for 28.7% of total provincial/territorial government expenditures in 1976 compared to 31.2% in 1996 (Federal, Provincial, and Territorial Advisory Committee on Population Health, 1996; National Forum on Health, 1997; Thomson, 1996).

In 1995, the federal government introduced a new program, the Canada Health and Social Transfer (CHST), which integrated several previously distinct federal-provincial transfer programs into a single system. The decline in federal cash contributions has continued. However, as a function of growing political pressure and the success of its general deficit reduction policies, the federal government has committed itself to maintaining a specified minimum level of support through the CHST for a five-year period. There is little doubt that confidence in the national health-care system has been shaken, but the public's support for the principles on which it was built remains high. Indeed, the political importance of Canada's health-care system was illustrated by the 1997 federal election. Health care was a priority for all five major political parties, reflecting the salience of this issue for Canadian voters. The incumbent government was re-elected in part on its expressed willingness to make a greater commitment to stable health-care funding, including a guaranteed higher minimum level of cash transfers to the provinces/territories.

2.32.3.4 A Public–Private Mixed Model of Health-care Financing

Outside the country, the public–private blend of Canada's health-care financing is often poorly understood. Notwithstanding the generally broad nature of the public component, the Canadian health-care system functions as a mixed economy. Even the publicly funded system operates with a certain duality, since most physicians are reimbursed on a fee-for-service basis for activities provided in their private offices. This led Stoddart (1987) to observe that the Canadian system is characterized by public financing with private management.

Given the decline in public funding described above, it is not surprising to find a trend toward a greater private share of the total health care dollar. The 76%:24% public–private ratio of 1975 shifted to 72%:28% by 1994. With physician services and hospital care largely covered by the public sector (99 and 90%, respectively), private funding is concentrated on prescription drugs and the noninstitutional care provided by most other regulated health professionals (68 and 86%, respectively). To put these numbers into perspective, total health expenditures in Canada were $72 billion in 1994, or $2010 per capita, third highest among OECD countries (National Forum on Health, 1997; OECD, 1996).

The Report on the Health of Canadians (Federal, Provincial, and Territorial Advisory Committee on Population Health, 1996) concluded that Canada generally benefits from a standard of health which is among the best in the world on most indicators (e.g., life expectancy, infant mortality). From the perspective of psychological factors, it is interesting to note that indicators such as quality of life, well-being, impact of health problems on everyday life, as well as measures of mental, social, and emotional health, have been added to traditional indicators of Canadians' health status. This shift underscores the importance increasingly accorded in Canada to psychological and social variables, not only in measuring health, but in setting health policy.

2.32.4 HEALTH-CARE POLICY AND FINANCING IN RELATION TO PSYCHOLOGY

Traditionally, the services of private practice psychologists have not been covered by Medicare, and Canadian health-care policy has been rarely explicit in addressing psychological services. However, the regular salaries of most providers of psychological services are paid from public funds. Alternative public and private methods for the payment of psychologists have been actively considered and promoted by organized psychology since the 1970s.

2.32.4.1 Psychology and Health-care Policy

Canadian health-care policy evolved primarily according to a medical model of health care. Within this model, from the origin of universal-access Medicare, the provision of psychological services was associated with mental health (psychiatric) services. Traditionally, this meant that the services of private practice psychologists were not covered by Medicare. Thus, the vast majority of psychological services providers in Canada have worked as publicly funded full-time salaried employees of provincial/territorial or federal agencies such as general hospitals, psychiatric hospitals, school boards, the military, correctional facilities, and mental health clinics. Private practitioners, working primarily in the large urban centers, have been paid directly by their clients/patients on a fee-for-service basis (Ritchie, Wilson et al., 1988).

Varied psychological services are offered in Canada (Martin, Ritchie, & Sabourin, 1988). Consistent with psychology's emergence as a major health profession, health services represent the largest concentration, in such areas as health psychology (Hearn & Evans, 1993), clinical psychology (Hunsley & Lefebvre, 1990), counseling psychology (Hiebert & Uhlemann, 1993), and neuropsychology (Fuerst & Rourke, 1995). These services are offered in a wide range of settings and characterized by considerable diversity. Focusing only on general hospitals, Hearn and Evans (1993) identified 65 areas of service provided by psychologists. The traditional distinctions among clinical, counseling, and health psychology are being increasingly incorporated into the broader designation of health service psychology (Adair et al., 1996). Cognitive-behavioral and behavioral approaches are the most prevalent among a comprehensive range of theoretical orientations (Hunsley & Lefebvre, 1990; Warner, 1991).

During the 1960s, limited third-party coverage for psychological services appeared in private health insurance plans, usually as an extended benefit. Starting in the 1970s, clinical psychology broadened its scope as a health service profession, a process which has continued to the 1990s, and, although psychological services often remain associated with mental health, psychologists are broadly defined healthcare providers in the Canadian health-care system. With this change in identity, psychologists desire increasingly to be funded in the same manner as other health professionals.

2.32.4.2 Psychology and Health-care Financing

A basic question about the public financing of psychological services is: Why were psychologists not included in Medicare on a basis comparable to physicians? Inclusion of psychologists in Medicare was considered during the 1965–68 development period, but the profession itself declined to pursue such inclusion. Few psychologists shared the view prevailing among physicians that government intervention, if at all, would be limited to the disabled and those on welfare. Professional psychology's provincially based leaders focused on the fact that there were relatively few professional psychologists in Canada at the time, and were concerned about the likelihood that initial inclusion might be limited to diagnostic services.

This failure to respond affirmatively to the opportunity for inclusion in Canada's first national healthcare system has bewildered and dismayed subsequent generations of psychologists. In retrospect, the decision is perhaps best understood in terms of three factors characterizing Canadian clinical psychology in the mid-1960s. First of these was the prevailing emphasis in Canadian clinical psychology on institutionally based public service as opposed to the private practice solo-office model espoused by most physicians. Second, and equally important, was the desire to see clinical psychologists legitimized as comprehensive practitioners. In the previous 20 years, psychologists had faced and increasingly overcome obstacles to move out of the "tester" role in order to assume broader clinical responsibilities, especially as psychotherapists. There was a fear that limiting Medicare inclusion to the diagnostic services provided by psychologists might make it more difficult to consolidate their role as treatment providers. Finally, at the national level, the issue was essentially irrelevant to CPA, given its decision two decades earlier not to support the emergence of psychology as a profession. If anything, Medicare inclusion was inimical to CPA's commitment to maintain psychology exclusively as a scientific discipline without professional aspirations. Hence, there was no national effort either to seek inclusion in Medicare as a goal in itself or within a framework that would secure coverage for both diagnostic and treatment services. Such a strategy would emerge only more than a decade later, by which time provincial governments were wary of adding new providers.

As a result, among the peculiarities of the Canadian system is the reimbursement of "psychotherapy" provided by general practice physicians (many of whom have virtually no training in this field) in their private offices, while similarly situated but well-trained private practice psychologists are ineligible for Medicare reimbursement. This is an excellent illustration of the implication of funding providers and facilities rather than services, something which is increasingly regarded as a flaw in Canada's publicly funded health-care system.

Canadian health-care policy has not been explicit in addressing psychological services, except under the rubrics of mental health and some specialized health services. The absence of broad inclusion in provincial/territorial government funding mechanisms has significantly limited the public's access to primary care services provided by psychologists. Since the late 1980s, downsizing in hospitals and other publicly funded institutions has resulted in a substantial reduction in the number of salaried psychologist positions, further limiting the public's access to psychologists, even for tertiary care. However, the public funding of all health-care services, including those provided by psychologists, is in transition. As discussed below, this situation presents opportunities as well as risks to health service psychology.

Given this context, it is important to articulate the key sources and methods of payment on which Canadian psychologists are most likely to rely in future years (Ritchie, Wilson et al., 1988). The sources of payment are either suppliers or users of psychological services. Suppliers are either private or public agencies, ranging from private clinics to government administered universal health-care plans. The users may be individuals, groups, agencies, corporations, or governments.

Several methods of payment commonly used by professionals are available to psychologists. Ritchie, Wilson, et al. (1988) identified 13 types of payment and 12 generic sources of payment actually or potentially available to Canadian psychologists. The vast majority of psychologists in Canada work in regular salaried positions. In large organizations, their salaries are typically within negotiated scales and include a fringe-benefits package. A significant proportion of salaried psychologists earn additional income from part-time private practice. Since the mid-1980s there have been increasing numbers of psychologists engaging in full-time private practice. Greater reliance on alternative methods of payment is required due to the negative influence of reduced federal transfer payments on the numbers and stability of salaried positions. In addition to government sponsored services, the actual or potential sources of funding are the users themselves, organized groups (e.g., employees, cooperatives), agencies (e.g, schools, clinics, hospitals), and businesses or corporations. Payment modalities include traditional fee-for-service, sessional rates, contracts, *per diem* consultations, retainers, and capitation.

2.32.5 HEALTH-CARE REFORM AND ITS IMPLICATIONS FOR PSYCHOLOGY

2.32.5.1 Elements of a Renewed, Publicly Funded Health-care System

The rapid pace of health-care reform will continue into the twenty-first century. The policy elements which will shape Canada's health-care system from 1997 to 2020 are increasingly clear. They are well captured in the final report of the National Forum on Health (1997), commissioned and chaired by the federal Prime Minister. Future health policy will be driven by a greater emphasis on a systems orientation and by fiscal sustainability. As a result, there will be more focus on defining the services to be funded rather than on simply funding providers. Given the traditional strong emphasis on funding physicians, this change potentially allows for the emergence of a larger number of health service providers, especially in primary care. Organized psychology in Canada is working vigorously to ensure the inclusion of psychologists in a reformed primary care system, discussed further below.

Although health services may become more broadly inclusive, funding will not be open-ended and is likely to be increasingly population-based. Variations on capitation as a basic funding mechanism are under renewed consideration for both public and private sector funding. In the public sector, Canada will almost certainly remain a "single payer" system. Evidence-based decision making is regarded as a key element in making the future health-care system appropriate to meeting health needs and in controlling costs.

The transformation of knowledge into action is essential to the functioning of such a system. Health service psychologists have an uncommon advantage relative to other health professionals. They have not only the capacity to be sophisticated consumers of advanced knowledge, they are also specifically trained to generate scientifically sound data. An implication for graduate education is that applied research skills will be especially important to the future marketability of health service psychologists.

The federal and most provincial/territorial governments have achieved their essential deficit reduction targets after introducing sufficient measures to reduce the public debt in the mid-1980s. Although there is no prospect of a return to large-scale public spending, developments indicate a restabilization of public sector health-care funding. The shift toward greater privatization of health care is unlikely to exceed a 70%:30% public–private ratio in the foreseeable future. Organized psychology will sustain its efforts for enhanced coverage of psychological services in both sectors. In the public sector, efforts will likely concentrate on primary care reform. In addition, focused opportunities such as that afforded by publicly administered "no-fault" auto insurance, with access to psychological services as a benefit, will continue to be pursued in the provinces with such insurance schemes.

2.32.5.2 Private Sector Initiatives

Organized psychology's private sector initiatives are concentrating on the traditional extended health and disability plans of the insurance industry, focusing on cost offsets associated with timely use of psychological care. American-style "managed care" has not been a prominent aspect of private sector thirdparty coverage in Canada, but elements of it have begun to emerge. Cognizant of the experience of American colleagues, there is considerable interest in enhancing the role which consumers and providers can play in determining the characteristics of such plans. OPA, with the active support of CRHSPP, is working on a provider-driven plan that would potentially encompass all nonphysician health professionals authorized by statute to practice autonomously. CPAP has expressed strong interest in the possibility of such a proposal being developed as a national plan.

2.32.5.3 Globalization

The reduction of domestic and international trade barriers brought about by the AIT and the North American Free Trade Agreement (NAFTA), respectively, is emerging as a previously unanticipated influence on the evolution of health care. Considerable concern was expressed when Canada and the United States established the Free Trade Agreement in the late 1980s, an agreement that was superseded by the NAFTA, of which Mexico is also a partner. In the short term, there has been no discernible impact on health-care delivery in any of the three countries. In part, this was due to the strong intervention of the health professions in Canada, and to a more limited degree in the United States, which resulted in each NAFTA country remaining unfettered to maintain its established mechanisms for regulating health-care and health professionals. However, both NAFTA and the AIT are likely to have an impact on the future mobility of health professionals. Mobility has emerged as a central topic for organized psychology in Canada, the United States, and Mexico.

The impact of a world economy characterized by the greater globalization of national and regional economies cannot be predicted with any certainty for psychology and other healthcare professions. Among the implications for health service psychology in Canada, none will present greater challenges than the impetus toward competency-based credentialing. In Canada, medicine and nursing have already devoted substantial resources to this issue, and several other health professions (e.g., physiotherapy) are already working actively in this area. In psychology, CPA, CPAP, and CRHSPP concluded an agreement in principle to mount a multiyear, collaborative effort to address this issue. Given the diversity of regulatory standards in psychology, the task will be difficult. On the other hand, external factors, especially government policies which could ultimately impose solutions insensitive to the interests of the public or the profession, provide the incentive for the profession to achieve appropriate solutions.

2.32.5.4 Determinants of Health

Canada has been a leader in generating public policy that attends to the broader determinants of health. Since the 1970s, two federal Ministers of Health made wellness and well-being, with a population-based focus, cornerstones of major federal government initiatives (Epp, 1986; Lalonde, 1974). The importance of attending to the broader determinants of health was reiterated by the Prime Minister's National Forum on Health (1997). Specific changes in health-care delivery and social policy have yet to match the policy rhetoric, but are likely to be incorporated in the reforms into the twenty-first century. For example, according to the Report on the Health of Canadians (1996), women did not score as well as men on measures of psychological well-being and depressions, and younger women appeared to be particularly vulnerable. The report also noted that much remains to be learned about the interpretation and application of such information. This is an opportunity virtually tailor-made for psychologists, whose knowledge and skills are well suited for the planning, implementation, and evaluation of action-oriented programs with targeted health outcomes.

2.32.5.5 Scope of Health Care

From a systems perspective, the largest and most overriding issue is the determination of what is or should be encompassed by health care. In Canada, this will be determined largely by decisions made in the public sector. Canadians expect that all essential elements of health care should be accessible within the publicly-funded system. Historically, this has been driven by a focus on physicians and hospitals as the fundamental elements of health-care delivery. Now, health-care reform is creating a shift toward greater overall emphasis on the need, appropriateness, and efficacy of health services. This is not occurring without considerable disruption and resistance.

Virtually no one believes that a return to the *status quo ante* is either achievable or desirable. In the future, a larger number of services will be provided in community settings that do not require the capital intensive investment of hospitals and large-scale facilities. However, consumers also want ready access to the range of services typically found in larger facilities. Health-policy planners want to control costs and prompt a shift toward greater emphasis on the maintenance of wellness relative to the treatment of disease and injury. Canadians remain strongly committed to the five core principles on which Medicare is based (accessibility, comprehensiveness, portability, public administration, universality) but are increasingly open-minded about the possibility of changing health service delivery mechanisms in view of economic and other realities, provided that the core principles are respected.

2.32.6 THE FUTURE OF PSYCHOLOGY AS A HEALTH-CARE PROFESSION IN CANADA

Two foci provide the context for assessing the future of psychology as a health service profession in Canada. The recommendations of the 1994 Mississauga Conference on Professional Psychology provide a broad vision. Primary care, as an area of specific application, links psychology to the broader structural and funding reforms which are prompting significant changes in Canada's health-care delivery system.

2.32.6.1 A Global Road Map: The Mississauga Conference

An important global portrait of professional psychology's preoccupations and the means by which it seeks to secure its future is found in the report of the 1994 National Conference on Professional Psychology (Dobson & King, 1995), known as the Mississauga Conference. The conference was initiated and organized by CPA in collaboration with CRHSPP, the National Health Research Development Program (an agency of the federal government), OPA, the Quebec College of Psychologists, and CPA's own sections on Clinical Neuropsychology, Clinical Psychology, and Counseling Psychology.

The conference developed some 50 guiding principles and an equivalent number of specific action plans for training, funding, and advocacy. The conference report stressed that psychologists must respond constructively to the needs and expectations of the Canadian public, especially at a time of fiscal restraint. It urged that the profession's capacity for creativity and flexibility, while remaining faithful to its scientific rigor and service-oriented values, be mediated by its ability to develop greater cohesion as an organized profession. Although encompassing professional psychology as a whole, the outcomes of the Mississauga Conference are especially germane to the future of health service psychology.

The training of future professional psychologists generated stimulating debate. Preparing them for the expectations of the marketplace was established as a legitimate training goal. In this context, Evans (1995) predicted that graduate education in professional psychology would become more focused on entrepreneurial skills and more market-oriented. The conference participants reaffirmed the doctoral degree for entry to practice as a psychologist, but agreed that training models could vary in their relative emphasis on practice and research. Furthermore, the diversity of applied psychology in Canada prompted a call for organized opportunities to support the development of proficiencies associated with specialties, as a complement to education and training for general practice.

The Mississauga Conference legitimized psychology's status as a business, complementary to its standing as a profession (Hearn, 1995). The conference also emphasized that advocacy is anchored in the public interest and must be responsive to public input (Ritchie, 1995b). This is consistent with the view that professional psychology serves a broader purpose than the narrow perpetuation of guild-driven goals. Viewed as a global road map, the approved action plans of the Mississauga Conference will help to shape professional psychology in Canada in the early decades of the twenty-first century.

2.32.6.2 Primary Care: An Illustration of Psychology's Future as a Health Service Profession

Primary health-care reform, being advocated in Canada and elsewhere, will likely have major implications for the practice of broadly defined health service psychology. Driven by broader public policy issues, primary care reform affords psychology a renewed opportunity to offer its unique competencies to the Canadian public.

2.32.6.2.1 Primary health-care reform as a public policy issue

Primary health-care reform is increasingly advocated in Western industrialized countries. Historically, Canada's primary care system, driven by physicians reimbursed on a fee-for-service basis, has encouraged an episodic, medically oriented approach. In Canada, a reformed primary health-care model is recommended, among other reasons, as the continuing entry point into the health-care system and because it can comprise comprehensive, continuing care (Manga & Campbell, 1994). A strong emphasis on such a model of primary care is of interest to psychologists because a comprehensive primary care system has the potential for more efficient primary consultation time (Hart, 1992), continuing rather than episodic care (Manga, 1993), and multidisciplinary teamwork (Fulop & Roemer, 1982).

The reform of primary health care is already a major public policy issue across Canada. The 1995 Conference of Ministers of Health received a report, commissioned from their senior officials who administer the respective provincial/territorial health-care systems, which called for a radical reorganization of primary care and for the introduction of population-based funding (capitation). Psychology was explicitly among the health professionals envisioned for inclusion in the proposed "Primary Care Organizations" (PCOs).

Provincially, Quebec has embarked on an ambitious effort to develop "Local Community Service Centres," blending some elements of primary health care with some social services. However, psychologists have been significantly under-represented in such settings. The inclusion of psychologists in a reformed primary care system has received the most systematic attention in the province of Ontario, where the Minister of Health has set primary care reform as a priority. The health professions, among them psychology, have been asked to contribute to the reform process by describing their role in primary care, their linkages to other professions, and their sources of funding. Several professions, again including psychology, were further invited to collaborate on the development of an interdisciplinary model for the provision of primary care services. Psychology participated under the auspices of OPA, with national support from the CRHSPP. Psychology's proposal established a clear vision and mandate for psychologists in a reformed primary care delivery system (OPA, 1997). The following considerations are based on this proposal.

2.32.6.2.2 Principles for professional psychologists providing primary care

Psychology committed to the establishment of the most appropriate roles for psychological services providers based on the principles of patient choice, rational utilization, and evidence-based decision making. Observance of these principles requires all professionals providing primary care as partners to have access to each other.

Because psychology has a broad scope of practice including the authorized act of diagnosis, its services span primary and tertiary care. Nonetheless, most clinical services provided by psychologists fall within contemporary definitions of primary care. However, some of the services properly construed as primary care have been historically provided in hospitals and other tertiary care settings, and may continue to be so provided in some communities although likely on a reduced scale. Psychologists will also continue to deliver services which are clearly part of tertiary care, such as sophisticated neuropsychological diagnosis of closed head injury patients, in tertiary care settings.

2.32.6.2.3 Competencies brought by psychologists to primary care

The particular competencies that health service psychologists bring to primary care include diagnosis, consultation, treatment, design and implementation of prevention and health promotion programs, as well as applied research.

Psychologists have particular expertise as diagnosticians. In community-based health-care settings, there is a high prevalence of patients with psychological problems or disorders associated with their physical complaints. Contrary to historic practice patterns, cost-effective patient care indicates that, in such cases, psychological diagnosis should be undertaken before other, considerably more expensive, diagnostic procedures. This will be feasible only when psychologists are as readily accessible to patients as physicians and nurses, and when psychologists' community-based services are publicly funded. It is expected that psychologists would both see patients for diagnostic consultation and provide partner health professionals with consultation around relevant diagnostic questions. Where the diagnostic question is primarily psychological, psychologists would be the most effective coordinators of the diagnostic process. This would be particularly important when the level of diagnostic complexity is moderate to high.

Treatment is a second major area of direct and indirect service. In future, the emphasis will likely be on short-term, empirically-supported approaches. Group therapy will be a preferred treatment modality, coupled with timely utilization of individual, couple, or family psychotherapies. Psychologists may provide treatment directly, supervise the treatment provided by other professionals, or provide consultation to other providers concerning the planning and implementation of appropriate treatments.

Psychologists will play a leading role in the design, adaptation and refinement of large-scale community based prevention and health promotion programs. Furthermore, since health-care settings will be expected to conduct in-house applied research, the fact that psychologists are educated as applied scientists will enable them to make a substantial contribution in this area, particularly in the areas of needs assessment, outcome evaluation, program evaluation, and clinical research.

Finally, the linkages of psychologists to other health professionals are expected to be strengthened in primary care agencies. In such settings, psychologists would probably have three types of linkages with other professions. Linkages would begin with day-to-day operations. All health-care providers would work in proximity to each other, and be accessible to their fellow providers and to each other's patients. An emphasis on patient need, rather than established patterns of service delivery, would benefit the patient, promote interdisciplinary collaboration, and enhance cost-effectiveness. The key would be access to the right provider for the most appropriate service at the right time.

Coordination and communication among various health practitioners, and across community agencies and services, are essential to the care of patients presenting with multiple physical and social needs. Primary care agencies would be well placed to play a central coordinating role, with participation from psychologists and other health professionals.

Psychology-specific linkages would likely derive from the health needs and referral patterns that are central to the work of psychologists. For example, psychologists would facilitate access to psychological services in tertiary facilities, and would make direct referrals to the specialized health services of psychiatrists, neurologists, pediatricians, and others.

2.32.6.2.4 Funding for psychologists providing primary care

The absence of province-wide public funding mechanisms across Canada has been the greatest impediment to the appropriate utilization of primary care services provided by psychologists. Traditionally, psychologists have provided some primary care diagnostic and treatment services in the outpatient departments of general hospitals, but downsizing has eliminated many positions and rendered psychological services less accessible.

The shift to a capitation funding system based on formulae driven by the demographics offers potential advantages to psychologists. In particular, funding will not be limited to physician services. However, there are utilization risks which will be borne by the providers rather than by the funder (i.e., provincial/territorial governments). Hence, careful attention to the derivation of capitation funding formulae will be required. Unfortunately, psychologists and other providers have little historic experience and data upon which to draw to bolster their negotiating strength in such negotiations. Governments will also have a stake in demonstrating good faith to ensure the success of primary care agencies to maintain the credibility of primary care reform with the public. When primary care agencies are set up, practitioners who are eligible for some private sector funding for other services may be able to offer an expanded range of services through a blend of public and private funding, provided that services reasonably construed as primary care remain publicly funded.

2.32.7 CONCLUSIONS

The long-standing dichotomy between psychology as a profession and psychology as an academic discipline had a retarding effect on the development of a cohesive national strategy for the profession in Canada. The emergence of stronger provincial/territorial associations, a reversal of CPA's previous antipathy to professional psychology, the creation of the CRHSPP, successful advocacy efforts which strengthened psychology's legitimacy through legislative recognition and the exemption of psychological health services provided by CRHSPP Listees from the federal Goods and Services Tax, have all enhanced Canadian psychology's standing significantly as a major health-care profession since 1980.

Although psychological services have not been covered directly by Medicare, the vast majority of psychologists in Canada have been reimbursed from public funds as regular salaried employees. Parallel services that emerged through private sector funding have improved public access to psychological services for some segments of the population. Full-time private practice remains the fastest growing segment of professional psychology in Canada. The development of innovative funding mechanisms is being pursued by organized psychology and entrepreneurially oriented psychologists across the country. Throughout Canada, health-care psychologists offer an increasingly wide range of high-quality services. Notwithstanding a decade of significant disruption and dislocation in publicly funded health services which has affected all health-care professions, there is in the late 1990s a reasonable likelihood of stable public funding, albeit at a reduced level when compared to historic baselines. The potential inclusion of psychologists in a reformed primary health-care system offers a promising opportunity to remedy limitations on the public's access to community-based psychological services engendered by anomalies in the Medicare program since the 1970s. On balance, after a period of sustained evolution since the 1940s, the realistic prospect for Canadian health service psychologists is bright for those who are prepared to adapt to emerging opportunities.

2.32.8 REFERENCES

Adair, J. G., Paivio, A., & Ritchie, P. (1996). Psychology in Canada. *Annual Review of Psychology, 47,* 341–370.

Adams, D. (1994). The future and long-range planning of health care services: A Saskatchewan perspective. *World Hospitals and Health Services, 30*(1), 25–31.

Appleby, M. H, & Rickwood, J. (1968). *Psychology in Canada: Special Study #3*. Ottawa, ON: Science Council of Canada.

Barer, M. L., & Evans, R. G. (1992). Interpreting Canada: Models, mind-sets, and myths. *Health Affairs, 11,* 44–61.

Berlyne, D. E., Black, A. H., Berry, R. G., & Douglas, V. I. (1971). *The Future of Canadian Psychology. Report to the Canadian Psychological Association and the Science Council of Canada.* Ottawa, ON: Science Council of Canada.

Bernhardt, K. S. (1961). *Training for Research in Psychology*. Toronto: University of Toronto Press.

Berry, R. G., Davidson, P. O., & Gibson, D. (1974). Advisory Council of Provincial Associations of Psychologists. *Canadian Psychology, 15,* 368–175.

Craig, K. D. (1993). The organization of professional psychology in Canada. In K. S. Dobson & D. J. G. Dobson (Eds.), *Professional psychology in Canada* (pp. 11–45). Toronto: Hogrefe & Huber.

Danzon, P. M. (1992). Hidden overhead costs: Is Canada's system really less expensive? *Health Affairs, 11,* 21–43.

Deber, R. (1992). Regulatory and administrative options for Canada's health care system. In *Exploring options for Canada's health care system*. Ottawa, ON: HEAL/ Canadian Nurses Association.

Dobson, K. S. (1995). Psychology in Canada: The future is not the past. *Canadian Psychology, 36*(1), 1–11.

Dobson, K. S., & Dobson, D. J. G. (Eds.) (1993). *Professional psychology in Canada.* Toronto: Hogrefe & Huber.

Dobson, K. S., & King, M. C. (Eds.) (1995). *The Mississauga Conference on Professional Psychology: Final Report*. Ottawa, ON: Canadian Psychological Association.

Enthoven, A. C. (1990). What can Europeans learn from Americans? In *Health care systems in transition. The search for efficiency*. OECD Social Policy Studies no. 7

(pp. 57–71). Paris: Organization for Economic Co-operation and Development.

Epp, J. (1986). *Achieving health for all: A framework for health promotion*. Ottawa, ON: Health and Welfare Canada.

Evans, D. (1995). Proceedings of the training group. In K. S. Dobson & M. C. King (Eds.), *The Mississauga Conference on Professional Psychology: Final Report* (pp. 8–12). Ottawa, ON: Canadian Psychological Association.

Federal, Provincial, and Territorial Advisory Committee on Population Health (1996). *Report on the health of Canadians*. Ottawa, ON: Health Canada.

Ferguson, G. A. (Ed.) (1977) *Report of the Vancouver Conference on the Organization and Representation of Psychology in Canada*. Ottawa, ON: Canadian Psychological Association.

Freeman, S. J. (1994). An overview of Canada's mental health system. *New Directions for Mental Health Services, 61,* 11–20.

Fuerst, K. B., & Rourke, B. P. (1995). Human neuropsychology in Canada: The 1980s. *Canadian Psychology, 36,* 12–45.

Fulop, T., & Roemer, M. I. (1982). *International Development of Health Manpower Policy*. Geneva, Switzerland: World Health Organization.

Haislmaier, E. F. (1992). Northern discomfort: The ills of the Canadian system. *Health Systems Review, 25*(4), 34–38.

Hart, J. T. (1992). Two paths for medical practice. *Lancet, 340,* 772–775.

Hearn, M. (1995). Proceedings of the funding group. In K. S. Dobson & M. C. King (Eds.), *The Mississauga Conference on Professional Psychology: Final Report* (pp. 13–17). Ottawa, ON: Canadian Psychological Association.

Hearn, M. T., & Evans, D. R. (1993). Applications of psychology to health care. In K. S. Dobson & D. J. G. Dobson (Eds.), *Professional psychology in Canada* (pp. 247–284). Toronto: Hogrefe & Huber.

Hiebert., B., & Uhlemann, M. R. (1993). Counselling psychology: development, identity, issues. In K. S. Dobson & D. J. G. Dobson (Eds.), *Professional psychology in Canada* (pp. 285–312). Toronto: Hogrefe & Huber.

Hunsley, J., & Lefebvre, M. (1990). A survey of the practices and activities of Canadian clinical psychologists. *Canadian Psychology, 31,* 350–358.

Hurley, G., & Ritchie, P. L.-J. (1995). Current and emerging strategic priorities of the Canadian Register of Health Service Providers in Psychology. *Register Report, 20*(1), 11–15.

Lalonde, M. (1974). *A New Perspective on the Health of Canadians*. Ottawa, ON: Health and Welfare Canada.

MacLeod, R. B. (1955). *Psychology in Canadian Universities and Colleges*. Ottawa, ON: Canadian Social Sciences Research Council.

Manga, P. (1993). Health economics and health care reform: Contributions and controversies. *Health and Canadian Society, 1*(1), 177–203.

Manga, P., & Campbell, T. (1994). *Health Human Resources Substitution: A Major Area of Reform Towards a More Cost-Effective Health System*. Working Paper. Ottawa, Canada: Queen's University of Ottawa Economic Projects.

Martin, R. M., Ritchie, P. L.-J., & Sabourin, M. (1988). The state of psychological practice in Canada: Current status and directions for the future. In P. L.-J. Ritchie, T. P. Hogan, & T. V. Hogan (Eds.), *Psychology in Canada: The state of the discipline in 1984* (pp. 73–119). Old Chelsea, PQ: Canadian Psychological Association.

National Forum on Health (1997). *Canada Health Action: Building on the Legacy-Final Report of the National Forum on Health*. Minister of Public Works and Government Services, Ottawa: National Forum on Health.

Organization for Economic Cooperation and Development (1996). *Health Data 1996*. Paris: Author.

Ontario Psychological Association (1997). *Psychological Services in Primary Care*. Toronto: Author.

Ritchie, P. L.-J. (1995a). Origins of the Canadian Register and its Relationship to the National Register. *Register Report, 20*(1), 11–15.

Ritchie, P. L.-J., (1995b). Proceedings of the advocacy group. In K. S. Dobson & M. C. King (Eds.), *The Mississauga Conference on Professional Psychology: Final Report* (pp. 18–28). Ottawa, ON: Canadian Psychological Association

Ritchie, P. L.-J., Hogan, T. P., & Hogan, T. V. (Eds.) (1988). *Psychology in Canada: The State of the Discipline, 1984*. Ottawa, ON. Canadian Psychological Association.

Ritchie, P. L.-J., & Sabourin, M. E. (1992). Sous un même toit: Canada's functional-structural approach to the unity of psychology. *International Journal of Psychology, 27,* 311–325.

Ritchie, P. L.-J., Wilson, R. F., Berry, R., Boulay, M., King, M., & Sabourin, M. (1988). The funding of psychological health services. *Canadian Psychology, 29*(2), 187–205.

Stoddart, G. L. (1987). Alternative modes for health care delivery. In *Aging With Limited Health Resources Proceedings of a Colloquium on Health Care* (pp. 68–71). Economic Council of Canada, Ottawa, ON: Supply and Services Canada.

Thomson, A. (1996). *Diminishing expectations: Implications of the CHST*. Ottawa, ON: HEAL/Canadian Nurses Association.

Torrey, E. F., Bigelow, D. A., & Sladen-Dew, N. (1993). Quality and cost of services for seriously mentally ill individuals in British Columbia and the United States. *Hospital and Community Psychiatry, 44*(10), 943–950.

Wand, B. (1990). *The Council of Provincial Associations of Psychologists: 1967–1990*. Ottawa, ON: Council of Provincial Associations of Psychologists.

Wand, B. (1993). The unity of the discipline: A challenge for the profession. *Canadian Psychology, 34*(2), 124–134.

Warner, R. E. (1991). A survey of theoretical orientations of Canadian clinical psychologists. *Canadian Psychology, 32,* 525–528.

Webster, E. C. (Ed.) (1967). *The Couchiching Conference on Professional Psychology in Canada*. Montréal: Canadian Psychological Association.

Wedding, D., Ritchie, P. L.-J., Kitchen, A., & Binner, P. (1993). Mental health services in a single payer system: Lessons from Canada and principles for an American plan. *Professional Psychology—Research and Practice, 24*(4), 387–393.

Wright, M. J. (1971). The psychological organizations of Canada. *Canadian Psychology, 15,* 112–113.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.33

Changing Health Care Environment in the United States: Steadily Evolving into the Twenty-first Century

PATRICK H. DeLEON and GARY R. VANDENBOS
American Psychological Association, Washington, DC, USA
MORGAN T. SAMMONS
National Naval Medical Center, Bethesda, MD, USA

and

ROBERT G. FRANK
University of Florida Health Sciences Center, Gainesville, FL, USA

2.33.1 INTRODUCTION 394
2.33.2 PROFESSIONAL PSYCHOLOGY: ONE OF THE "LEARNED PROFESSIONS" 394
2.33.3 MANAGED CARE 394
2.33.4 TELEHEALTH–TELEMEDICINE 395
  2.33.4.1 Federal Developments 395
  2.33.4.2 Professional Issues 397
  2.33.4.3 Federal Legislative Developments 398
2.33.5 PRESCRIPTION PRIVILEGES 399
  2.33.5.1 Fewer Prescribers, Greater Challenges—Rising Needs 400
  2.33.5.2 The Challenge of Psychiatric Residency Training 400
  2.33.5.3 Towards More Rational Prescribing Practices 401
  2.33.5.4 Aging Populations—Aging Programs 402
2.33.6 HEALTH CARE COSTS 402
  2.33.6.1 The National Health Care Debate 403
  2.33.6.2 The Town-gown Split 404
  2.33.6.3 Self-identity 404
  2.33.6.4 The Congressional Science Fellowship Program 405
  2.33.6.5 The Professional School Movement 406
2.33.7 OPPORTUNITIES FOR THE FUTURE 406
2.33.8 CONCLUSION 407
2.33.9 REFERENCES 407

2.33.1 INTRODUCTION

The US' health care system is undergoing unprecedented changes which will have a direct impact upon the future of professional psychology. These include: increasing acceptance of managed care, extraordinary technological advances within the computer and communications industries, drastically increasing numbers of psychologists, and the evolution of psychology into the generic health care arena. Advances in telehealth constitute major opportunities and challenges. Obtaining prescription privileges represents an entirely "new frontier." Psychology is one of the "learned professions" and should prosper if the practitioner and academic communities work collaboratively to address society's pressing needs. Is the next generation willing to become involved in the political and public policy efforts that are required?

2.33.2 PROFESSIONAL PSYCHOLOGY: ONE OF THE "LEARNED PROFESSIONS"

The USA is considered to be highly educated. Yet, the reality is that slightly less than 20% of adults possess a baccalaureate degree, and of those entering division one colleges, only half (56%) actually graduate within six years (DeLeon, Sammons, & Sexton, 1995).

Accordingly, it is contended that professional psychology, with its historical emphasis upon doctoral level training, has a truly unique and pressing societal responsibility to act as one of the "learned professions" (DeLeon & Wiggins, 1996). Society's pressing needs must be addressed affirmatively wherever psychological expertise could be deemed relevant. There is a societal obligation to work to ensure that all Americans have access to quality health care, including mental health care. There must be a collective responsiveness to the ever expanding opportunities for the psychology profession as it evolves within the nation's health care and educational arenas. Simply stated, psychologists have a responsibility to demonstrate effective leadership. As one of the "learned professions," psychologists must not "sit on our laurels" and be complacent.

Reflecting on the profession's past, psychology has an impressive history of successfully overcoming numerous obstacles to the establishment and expansion of professional practice. Psychologists are now authorized to autonomously provide psychotherapy in all states, are independently reimbursed for services rendered under nearly every federal and private health program, are routinely granted hospital privileges in the public and private sectors, and provide expert testimony in the courts (Resnick, DeLeon, & VandenBos, 1997). By 1996, 47 states had enacted psychology "freedom of choice" statutes.

The establishment of professional schools has resulted in numbers expanding beyond anyone's expectations. The current estimate being in excess of 70 000 licensed practitioners and an American Psychological Association (APA) total membership, across all categories of members, in excess of 150 000. Psychologist colleagues have been appointed to numerous positions of public trust. Approximately one third of US states have selected psychologists to direct their mental health system. Psychologists have served in numerous positions of high level federal responsibility; including in the US Congress, as Institute directors, head of the Federal Bureau of Prisons, and hospital (facility) commanders. The future should bring similar administrative and legislative recognition, not to mention broader clinical appreciation for the inherent value of the behavioral and psychological sciences throughout the delivery of quality health care (VandenBos, DeLeon, & Belar, 1991).

However, it must also be reported that some senior colleagues have begun raising serious questions about the profession's collective willingness to continue to engage in the administrative and legislative struggles that are required. It is as if the current generation of professional psychologists believes that they are entitled to a comfortable life, rather than being willing to work aggressively to obtain or maintain one. Stated slightly differently, recent graduates do not seem to appreciate the importance to their profession (and to society) of becoming personally involved in the public policy and political process. Interestingly, this phenomenon seems also to be the case within a number of the other nonphysician disciplines. It has further been observed that unlike many colleagues in the other disciplines, very few psychologists truly appreciate the long-term policy implications for practice of fostering viable clinical-training collaborative relationships (DeLeon, 1989).

2.33.3 MANAGED CARE

There can be little question that for most practitioners (and increasingly for those who are employed within academia) the 1990s can be summarized by the dreaded phrases "the advent of managed care" and the "industrialization of health care." Across the nation, many colleagues report that their incomes are being dramatically slashed. Practitioners, and particularly those in solo private practice, have become increasingly vocal in decrying the intrusion of managed care into their lives; allegations are made that the "quality of care" they are now able to provide may no longer even be within ethical boundaries; and that clinicians' complex judgments are simply not valued. In essence, many colleagues sincerely believe that efforts to curtail the costs of health care have supplanted providing necessary and quality health services. It is as if managed care companies are practicing health care, rather than administering a health care benefit (Miller, 1996).

Few would argue that there have not been unconscionable abuses under managed care. And, it is also the case that those who ultimately pay the bill have become increasingly involved in questioning what it is that they are actually receiving for their money. Many colleagues are fierce individualists; it would be an understatement to suggest that they passionately object to having anyone "overseeing" their clinical judgment. And yet, there is also a growing body of evidence which highlights that unnecessary variances exist in health care delivery, particularly in certain geographical areas (Wennberg, 1984). And, further, that credible indications continue to surface that many clinicians (regardless of discipline) always seem able to proffer "reasonable reasons" to provide services, as long as they will be reimbursed.

Notwithstanding the considerable emotional turmoil surrounding managed care expressed by numerous practitioners during the past years, the 1990s has also seen increasing numbers of psychologists working with, or being employed by, credible managed care organizations. Many report that they are very excited about their increasing ability to contribute significantly to improving their patients' overall health status, and that their practices are constantly expanding into entirely new and previously unforeseen arenas. A fundamental policy question that has been raised by the leadership of the Committee for the American Psychological Association of Graduate Students (APAGS) is: Have our training programs really kept up with the evolving changes in health care delivery? Or, is psychology instead training a generation of practitioners who will ultimately not possess the fundamental skills necessary to survive in the changing health care environment?

The practical, financial, and clinical reality is that increasing numbers of Americans are enrolling in managed care. Both the Clinton Administration and the Republican Congress continue to express great faith in the quality of care that is provided by managed care organizations. In 1996, nearly 57% of all Americans were covered by managed care, a doubling of the number from a decade ago. Nearly three-quarters of Americans covered by employer-based health insurance plans are enrolled in managed care plans, compared with only half in 1993. The Administration's Fiscal Year 1998 budget request for the Department of Health and Human Services (HHS) expressly noted that since 1993, the number of Medicare and Medicaid beneficiaries enrolled in managed care plans has experienced unprecedented growth. Focusing specifically upon the delivery of mental health care, it is estimated that 149 million people are insured through managed behavioral health firms (National Health Policy Forum, 1997). Previously enacted psychology "freedom of choice" laws will need to be amended to reflect the changing market place. The Balanced Budget Act of 1997 (P.L. 105–33) provided the states with greater flexibility to require individuals receiving Medicaid to enroll in managed care plans and specifically eliminated the previous statutory requirement that the states first needed to obtain a formal waiver from the federal government (i.e., HHS).

From a public policy perspective, it is evident that many simply do not appreciate that managed care is here, and that it is not a new conceptual entity. Managed care existed during the days of the great depression, and managed care was the cornerstone of President Nixon's health care agenda (DeLeon, Bulatao, & VandenBos, 1994; DeLeon, VandenBos, & Bulatao, 1991). The health programs of the various federal and state agencies (i.e., the Veterans Administration [VA], Indian Health Service, and Department of Defense) have always had to face the realities of providing comprehensive health care, for an identifiable patient population, within finite budgetary constraints (i.e., have been "managed"; Blanck & DeLeon, 1996). A related and equally significant fundamental health policy concept is that many psychologists are just beginning to appreciate that the delivery of psychological services extends far beyond the delivery of only mental health care *per se*; and that from a public policy perspective, psychological care is but a small subset of health care (DeLeon & VandenBos, 1991). Psychologists are health care professionals.

2.33.4 TELEHEALTH–TELEMEDICINE

2.33.4.1 Federal Developments

One of the most exciting and far reaching developments that has occurred within the health care arena since the mid-1980s has been the increased clinical applicability of computer and telecommunications technology. Historically, most Americans (and particularly practitioners) have viewed the process of receiving health care as a highly personal, one-on-one experience between the patient and practitioner (usually a physician or dentist). With the advent of the nurse practitioner and physician assistant movements during the mid-1960s (DeLeon, Kjervik, Kraut, & VandenBos, 1985), the notion and appropriateness of "physician extenders" providing care became increasingly acceptable, if not commonplace, that is, the fact that there might be someone (or something) in between the doctor–patient relationship began to seem normal. In the late 1990s, the unprecedented advances in technology which are occurring are taking this process a significant step further—the actual physical location (i.e., presence) of the practitioner is becoming almost irrelevant.

In 1994, a bipartisan, bicameral *ad hoc* Steering Committee on Telehealth and Health Care Informatics was established in the US Congress. Its objective was to allow interested members to informally explore how these technological advances might best be utilized to improve access to quality health care, particularly in rural America. It seems intuitive that the increasing sophistication and exponential growth of technology must have clinical implications for health care delivery—but exactly in what fashion has yet to be determined.

In many ways, the clinical application of telehealth–telemedicine is still in its developmental stages. Accordingly, it is pleasing to note the extent to which various governmental agencies and private organizations with health policy expertise have begun collaborating to establish a comprehensive policy framework for what might evolve. By the end of the first session of the 105th Congress (1997), impressive reports had already been issued by the Institute of Medicine (IOM, 1996), the former Office of Technology Assessment (US Congress, OTA, 1995), the General Accounting Office (GAO, 1997a), and the US Department of Commerce (1997). The federal government had established a Federal Interagency Joint Working Group on Telemedicine, which among other responsibilities has been tasked with establishing a world wide web-accessible inventory of federally-sponsored telemedicine projects. Note, as an aside, the consistent efforts by organized medicine and the Clinton Administration to utilize the policy phrase "telemedicine," notwithstanding the Congress' statutory preference for the more generic and inclusive term "telehealth."

Some of the highlights from these reports are: by the end of 1996, nearly 30% of rural hospitals were expected to use telehealth to deliver patient care; more than 40% of telehealth programs were in operation one year or less; and the most common clinical uses for telehealth were diagnostic consults, medical data transmissions, and management of chronic illnesses.

The Commerce Department report (1997) expressly addressed the potential applicability of telehealth for mental health services:

> Unlike some specialties such as cardiology, mental health telemedicine applications do not require a high degree of image specificity, thus the technology can be simpler and costs can be lower. Moreover, initial evidence indicates that telemedicine works well for group therapy as well as individual consults. (p. 19)

By the middle of the 105th Congress over 35 federal organizations, 10 state governments, and numerous private organizations had sponsored hundreds of initiatives in over 40 states. According to the Department of Commerce, over a three-year period, the Federal government alone invested more than $646 million, with different agencies developing their own unique priorities. The Library of Medicine of the National Institutes of Health was projected to be providing $42 million, targeting proposals that examined specific use of the technology in health care, whereas the HHS Office of Rural Health Policy was targeting multispecialty networks of care, stressing improving access to care in rural areas.

The Office of Rural Health Policy is the lead telehealth programmatic agency for HHS. In Fiscal Year 1996 it awarded $9.9 million in grants under one particular initiative, in which grantees were to belong to, or propose becoming members of, a network and thereby able to provide a minimum range of services. Each network was to provide a minimum of seven services, one of which was to be the stabilization of patients in emergency services. Although to date psychology has been only tangentially involved, the potential is definitely present since the office has demonstrated an explicit emphasis on encouraging multidisciplinary and community-focused care, with priority being given to projects that proposed providing care in such community-based settings as schools, assisted living facilities, individuals in the home, and so forth. That year, the Office awarded an additional $4.75 million to eight telemedicine projects under its rural health outreach grant program, as well as providing support to two major telemedicine demonstration projects in West Virginia and Iowa. Overall, between 1994 and 1997, the Office awarded approximately $36 million in telehealth funding.

2.33.4.2 Professional Issues

Some of the evolving clinical and public policy issues in telehealth that should be particularly relevant to professional psychology include: confidentiality of patient records, insurance reimbursement, interstate licensure decisions, malpractice liability, and the potential for "cookbook medicine" or requiring predetermined algorithms of care. Many of these policy issues will eventually be debated at the state level, both in the legislature and by administrative agencies. By the end of 1997, 15 states had decided to offer some telemedicine coverage under their Medicaid programs and the states of California and Louisiana had enacted legislation mandating coverage for telehealth consultation services under private insurance plans.

Health policy experts and clinicians can readily appreciate how this advancing technology might work for services that are primarily visual (or physical) in nature; that is, reading x-rays, performing dermatology consults, and transmitting medical records over long distances. For professional psychology the underlying question becomes: how can telehealth capabilities facilitate traditional psychological interventions which may not be as concrete (or physical) in nature? Computerized testing and diagnostic interviews and consultations are possible. What about providing psychotherapy? The more one gets involved, it also becomes clearer that there remains the critical issue of how to facilitate practitioner acceptance of this new technology.

Reflecting again upon the profession's training programs from a national frame of reference, the APA past-President Ron Fox has reported that there are approximately 26 identifiable telehealth programs within medical school environments, but none within psychology's training programs (personal communication, 1997). Fortunately, various elements of the APA governance (the Board of Educational Affairs, the Committee for the Advancement of Professional Practice, the Board of Professional Affairs, and the Committee on Rural Health) have begun addressing the policy and programmatic implications. The Department of Commerce report expressly referenced psychology's efforts. Telehealth is becoming a major factor in the delivery of mental health care. The *American Medical News* (1997, November 10) reported that mental health, constituting 20% of the 1996 consultations, has become telehealth's most common use.

For psychology, the fundamental issues of interstate licensure, privacy/confidentiality concerns, and patient outcome measures must be systematically addressed. Already a number of states (10 by 1997) have decided to require that telehealth practitioners possess multistate licensure. Professional nursing, however, has been considering urging their licensure boards to adopt a "driver's license" approach, under which the practitioner/consultant would only be required to be licensed in one state. Within medicine, the Federation of State Medical Boards is exploring what approach they should recommend as appropriate guidelines to their licensure boards.

The leadership of professional psychology, and in particular the leadership of its professional schools, must collectively become systematically involved in the ongoing HHS demonstrations and policy dialogues. For example, the Health Resources and Services Administration (HRSA) requested policy guidance on:

> the legal barriers to the cost-effective use of telemedicine and specific suggestions for overcoming these barriers. In particular, we seek suggestions for easing licensure barriers to physicians and other health professionals providing telemedicine services across state lines, and comments on specific alternatives, such as those recently proposed by the Federation of State Medical Boards and the Institute of Electrical and Electronics Engineers ... (R)espondents are encouraged to provide comments and suggestions on other legal issues associated with telemedicine such as liability/malpractice ... to identify the particular challenges in assuring privacy, confidentiality, and security in the conduct of telemedicine and provide suggestions for addressing those challenges. Comments will be reviewed and considered for incorporation into the final report to Congress. (*Federal Register*, August 9, 1996, p. 41 640)

We must ask ourselves: what elements of professional psychology, other than the professional schools, might become involved in the actual delivery of health care utilizing telehealth? And, thus, intimately affected by governmental decisions?

To again focus upon the federal government's role as a provider of health care and to give another perspective of the extent to which telehealth has already matured: in 1996, the health officials of the Department of Defense (DoD) testified before the Senate Appropriations Committee (US Department of Defense, 1996, S. Hrg. #104–756). The witnesses described their considerable enthusiasm for the clinical and teaching potential inherent in the technology revolution. Medical digital technology, telemedicine, advanced computer-based medical expert systems, information management—

for combat as well as for day-to-day business and clinical practice—were explicitly addressed. Allowing attending providers on board ships to transmit real-time ultrasound videos and full color images back to specialists at major treatment facilities hundreds of miles away was clearly "doable." It was noted that DoD's telehealth presence in Bosnia was extraordinary.

As we have indicated, the DoD is a major provider of health care. It administers care in 116 military hospitals and 513 clinics for a beneficiary population of over 8.3 million, with a medical budget of approximately $9.6 billion annually. In 1996/97 alone, DoD spent $48 million developing telemedicine infrastructure. What lessons can be collectively learned from military colleagues?

2.33.4.3 Federal Legislative Developments

The Kennedy–Kassebaum health insurance legislation, the Health Insurance Portability and Accountability Act of 1996 (P.L. #104–191), surfaced some very interesting policy implications for professional psychology, not the least of which was the extensive debate surrounding the possibility of enacting mental health parity legislation. The final enacted bill did not include the mental health parity provision but did recognize that advances in health care technology, and specifically telemedicine, will have a major impact upon how health care will be delivered in the future. The conference agreement

> directs the Health Care Financing Administration (HCFA) to complete its ongoing study of Medicare reimbursement of all telemedicine services and submit a report to Congress on reimbursement of telemedicine services by March 1, 1997. The report would . . . include an analysis of the cost of services provided via telemedicine, and include a proposal for Medicare reimbursement of telemedicine services. (H. Conf. Rpt. #104–736, p. 235)

The Balanced Budget Act of 1997 (P.L. 105–33) was subsequently enacted and took the next step by requiring the Secretary of HHS to continue the Department's three-year demonstration project under which Medicare will pay for telemedicine services at 57 Medicare-certified facilities and to make Part B payments under Medicare for telehealth "professional consultation" services available by no later than January 1, 1999. Interestingly, this same legislation also substantially modified the Graduate Medical Education (GME) program (which will be addressed in more depth later in the chapter) in ways which could have a major impact upon psychology as a health profession. Historically, the GME program has stressed hospital based services and has been physician dominated. The Budget Act expanded direct medical education payments to entities that were previously not eligible (and where psychologists are increasingly being employed), specifically to federally qualified health centers, rural health centers, and Medicare + Choice organizations. Most importantly for psychology, the conference report expressly "urges the Secretary to include physician assistants and psychologists under (the allied health) authority" (H. Conf. Rpt. #105–217, p. 822) for determining reimbursement.

During the conference deliberations on the Fiscal Year 1998 Appropriations bill for the Department of Defense (P.L.105–56, 1997), specific language was also included focusing upon telehealth:

> The conferees agree that pursuant to the GAO report "Telemedicine: Federal Strategy is Needed to Guide Investments," DoD Health Affairs should contract with an outside organization, which has previously demonstrated the ability to assess and plan strategic telehealth initiatives, to enumerate current telehealth efforts and recommend coordination activities among the tri-services to prevent duplication and enhance "dual use" telecommunications for health care delivery, education, and related public sector applications. This assessment should also include the worldwide computer-based patient record as a necessary component of telehealth. The resulting report and recommendations shall be provided to the congressional defense committees by April 1, 1998. (P. L. 105–56, 1997, H. Conf. Rpt. #105–265, p. 139)

Clearly major advances are evolving in telehealth within the federal sector. To cite another example, the Senate reports accompanying the DoD and VA appropriations bills for Fiscal Year 1998 also formally recognized this technology's potential for upgrading VA nursing staff professional skills.

> VA/DoD distance learning pilot project. The Committee encourages the DoD to continue the VA/DoD distance learning pilot project to transition clinical nurse specialists to the role of nurse practitioners. This collaborative program with the USUHS Graduate School of Nursing will increase the number of primary care providers. (Sen. Rpt. 105–45, p. 149)

And,

> The Committee encourages VA to continue the VA/DoD Distance Learning Pilot Program to transition clinical nurse specialists to the role of nurse practitioners. This collaborative program

> with the Uniformed Services University of the Health Sciences Graduate School of Nursing will add to the number of primary care providers required to care for the projected increase in the numbers of aging and female veterans. (FY 1998 Departments of Veterans Affairs, 1997, Sen. Rpt. #105–53, p. 16)

We would rhetorically ask: are VA psychologists involved in these activities?

2.33.5 PRESCRIPTION PRIVILEGES

The move towards prescriptive authority (DeLeon, Fox, & Graham, 1991; DeLeon, Sammons, & Sexton, 1995) takes place in the context of two major shifts in health care provision and training. The first of these is that organized medicine is gradually losing its monopoly on the provision of health care services. Other health care specialties—such as advanced practice nursing, podiatry, and optometry—are moving into fields which had previously been the sole purview of the physician. In the USA, though not the case globally, control of the prescription pad has long been the *sine qua non* of medicine. American medicine has found that control of prescriptive authority has proven to be an astonishingly potent political tool, and, since the beginning of the twentieth century, considerable effort has been expended by physicians to establish a relative monopoly on this function. Surrendering this authority, which is an emblematic step in the gradual dissolution of medicine's hegemony in health care provision, is the subtext which underlies the opposition by psychiatry and other branches of organized medicine to the acquisition of prescriptive authority by psychologists.

The data illustrating this shift are dramatic. An extraordinary illustration of this trend is revealed by the fact that the number of prescriptions written by nonphysician prescribers more than doubled between 1992 and 1996. In 1992, 0.22% of all prescriptions were written by nurse practitioners or physician assistants; in 1996, this figure was 0.57%. Though still small in comparison to the overall numbers of prescriptions written, this percentage represents over one million prescriptions per month (*American Medical News*, 1996, June 17). Since July, 1995, the Drug Enforcement Administration (DEA) has issued controlled substance prescription numbers to nurse practitioners and physician assistants.

These numbers have been driven by remarkable legislative successes by the various professional groups. In particular, those of advanced practice nurses (APNs) have demonstrated the changing scope of health care provision. In 1975, North Carolina became the first state to provide nursing with prescriptive authority. By 1996, APNs had some form of prescriptive authority in all states except Illinois. In 19 states, they could prescribe from formularies including controlled substances completely independent of physician involvement, a figure which had almost doubled from two years previously, when APNs could prescribe independently in only 11 states. In 1997 it was estimated that there are approximately 75 000 APNs throughout the nation. Optometrists and podiatrists possess prescriptive authority in all states (except the District of Columbia, for optometry). Physician assistants and pharmacists have acquired more limited forms of prescriptive authority, but at a rate which is quickly expanding. In the case of podiatry and optometry, these privileges are granted without physician oversight, and the degree of latitude afforded physician assistants and pharmacists has increased significantly. In the US Army there is a program for physical therapists, ongoing for more than two decades, under which they are granted prescriptive authority in the course of treating neuromuscular complaints (Benson, Schreck, Underwood, & Greathouse, 1995).

Although the medical press frequently implies that nonphysician prescribers imperil patients' well-being (i.e., are "public health hazards"), data have long been extant which demonstrate that nonphysician health care providers prescribe psychotropics at least as well as their medically trained counterparts. Studies in the 1970s demonstrated that nonmedically trained mental health providers, armed only with knowledge acquired during weekly seminars, did not differ from psychiatrists in their prescribing patterns (Platman, Dorgan, & Gerhardt, 1976). In the state of California, over a million patients were seen by nonphysician prescribers with no adverse consequences (State of California, 1982). Interestingly, less than 20% of all psychotropic medications are ordered by psychiatrists.

In addition to the results of demonstration projects and similar efforts, overall differences in safe practices between physicians and nonphysicians can be inferred by using the frequency of successful malpractice actions against both groups. Here, the extant data speaks clearly in favor of the argument that nonphysicians prescribe safely. In general, nonphysician groups are usually more successful than their physician counterparts at avoiding completed malpractice actions. For example, in a study of lawsuits related to failures to diagnose heart attacks, it was discovered that physician extenders were successfully sued least often among a varied group of medical specialties (including cardiologists and general practitioners), accounting for only four of 495 claims in the reporting period (Flannery, 1997).

Figures for prescribing psychologists are, of course, not available, but tangential evidence suggests parallel experiences for psychologists. A 13-year history of malpractice actions against psychologists in California indicated a total absence of successful claims for failure to diagnose medical conditions in patients seen by psychologists (Dorken, 1990). From these data one could reasonably infer that malpractice actions would not be a significant concern for those who prescribed. Further oblique support for this position comes from another interesting study which found that psychologists were as good as general physicians, and slightly (but not significantly) better than a small group of psychiatrists, in detecting Axis III conditions masquerading as Axis I problems (Sanchez & Kohn, 1991).

Patient acceptability of nonphysician prescribers has been good, particularly within organized systems of care. In England, nurses were granted prescriptive authority in 1994. A survey of patients after this change suggested that most patients were comfortable with this new arrangement. The principal benefits seen were that patients no longer had to visit two providers in order to get a prescription, that time was saved, and that patients could get medications from providers with whom they had previously established trusting relationships. No patients expressed opposition to the idea of nurse prescribing (Luker, Austin, Hogg, Ferguson, & Smith, 1997).

Among psychologists, the Psychopharmacology Demonstration Program (PDP) established under the auspices of the Department of Defense has provided a unique laboratory for experimentation in curriculum development for prescribing psychologists (Sammons & Brown, 1997). The program was established during Congressional deliberations on the Fiscal Year 1989 DoD Appropriations bill (P.L. 100–463, 1988) and has now graduated 10 Fellows who are privileged to prescribe at their military facilities. There have been several independent evaluations of the highly controversial program which have confirmed that the Fellows provide high-quality health care. There is disagreement, however, as to whether the program has been cost-effective (GAO, 1997b; Jones, Cohen, Munsat, Dorris, & Berson, 1996).

The second major shift is more specific to organized psychiatry. It is a subset of the first; that is, reflecting the gradual reduction in power of organized medicine. From our vantage point, it is clear that there is growing recognition that psychiatry as a profession has imperiled its existence by excessive reliance on biological models of psychopathology and biological interventions. And perhaps even more significant in the long run, psychiatry has not been able to attract a critical mass of American medical school graduates into its residencies.

A number of psychiatrists are now beginning to understand that the future of their profession as a discipline may depend to a great extent on its ability to combine psychological and pharmacological approaches to treatment. Unfortunately, a history of excessive reliance on ideologically grounded disciplines (first Freudian psychoanalysis, and later psychopharmacology), has limited the clinical evolution of psychiatry. In our judgement, psychology is a much more heterogeneously grounded discipline, and the variety of subdisciplines which make up the field have allowed us to sidestep the trap of falling into a single treatment modality. Nonetheless, it is important that psychologists recognize that, in order to prosper, they must be open to embracing ideologies and methodologies that had previously been regarded as outside their scope of practice. They must be careful not to emulate the mistakes of physician colleagues, who have found that reliance on unimodal treatment strategies quickly leads to obsolescence. Instead, they must develop and utilize the most appropriate investigatory techniques to determine which treatments, and in which combinations, will optimize the science and clinical practice of psychology.

2.33.5.1 Fewer Prescribers, Greater Challenges—Rising Needs

By the year 2020, psychological and behavioral disorders, such as major depression, bipolar disorder, panic disorder, schizophrenia, and alcoholism will have surpassed all but a very few other causes of worldwide morbidity and mortality. On a global level, unipolar major depression is projected to represent the second most significant contributor to disability adjusted life years, and neuropsychiatric conditions in general will contribute 10.5% to disability adjusted life years (Murray & Lopez, 1997). Again, a rhetorical question: what are the training programs doing to target these populations?

2.33.5.2 The Challenge of Psychiatric Residency Training

In 1994, only 3.2% of all US medical school graduates entered a psychiatric residency, the lowest number since 1929 (Sierles & Taylor, 1995). At the same time, the number of international medical graduates (IMGs) entering US residencies has escalated exponentially, and has increased 64.7% since 1990 (Peterson et al., 1997). The influx of IMGs who enter US residencies and then stay to practice medicine (as the majority of IMGs do) is not unique to psychiatry. Nationwide, 25.2% of all medical residency slots are filled by IMGs, but the problem is particularly acute in psychiatry where nearly half (41.8%–49.5%) of all US psychiatry residency programs are dependent on IMGs—that is, over 50% of their first year residents are IMGs (Whitcomb & Miller, 1995). It has been estimated that if Congress merely adopted the recommendations of organized medicine as to how to limit the percentage of GME payments currently going to international medical graduates, psychiatry would experience a 40% reduction of its residency slots the first year alone.

It has been noted that the oft-cited rationale for utilizing IMGs (i.e., that they would provide more services to indigent citizens) was not found to be true, with only a minority of IMGs being engaged in treating the poor and traditionally underserved. Could it be that the high percentages of IMGs in US psychiatry may be a factor in the astonishing utilization of psychotropic medications by their specialty? It has been reported that the percentage of psychiatrists' patients receiving medications rose from 27% in 1975 to 54.5% in 1988. Similar analysis of data from 1988 and 1989 indicated that 61.7% of all psychiatric patients received medication, and that IMGs were disproportionately represented in that group of psychiatrists defined as "high medication practice" (Olfson, Pincus, & Sabshin, 1994). As others have demonstrated, the percentage of psychiatric patients receiving medication continues to climb, with over 90% of patients receiving at least one drug in a 1996 survey of psychiatrists (West, Zarin, & Pincus, 1997).

During the DoD psychopharmacology training program, the Fellows became acutely aware of the relevance of their psychological training to the complex issues surrounding utilization of psychotropic medications. There are a number of prescribing areas in which psychology should be more actively involved; particularly, in asking relevant clinical utilization questions. Gender-specific training in clinical psychopharmacology is an open and challenging field. From a pharmacoepidemiological perspective, differential rates of prescribing for men and women are well recorded. Women are more likely than men to be prescribed psychotropic drugs, and this difference is not entirely explainable by generally higher rates of diagnosis of mental distress in women. In particular, elderly women are more likely than any other segment of the patient population to receive psychotropics. The factors which contribute to these discrepancies are less well understood. Is gender of the patient the sole determinant of differential rates of prescribing? This seems unlikely. It is more likely that an interaction exists between the gender of the prescriber and gender of the patient which can further explain such differences.

Another clinical issue which requires elucidation is the persistence of high risk prescribing patterns, despite the availability of data suggesting safer clinical strategies. It has been noted that this is an especially significant issue for elderly females, who receive disproportionate shares of psychotropic medication. On a more preclinical level, gender specific physiological functions, such as the effects of pregnancy, lactation, menopause, and estrogen levels on both the production of mental disorders and the pharmacokinetic and pharmacodynamic properties of psychotropic drugs, still await extensive investigation.

2.33.5.3 Towards More Rational Prescribing Practices

It is reasonably clear that a medical degree and completion of a psychiatric residency *per se* is not the prime determinant of best prescribing practice of psychotropics. Thus, a question that begs investigation is to clarify the essential components of the knowledge base which underlies skillful prescribing, and perhaps more importantly, what components of education (didactic and clinical) are necessary to allow providers to distinguish when a patient will be more responsive to medication or nonallopathic therapies.

Since the Second World War, the pharmacopoeia for mental disorders has expanded exponentially. The development of effective pharmacological treatments for many forms of mental distress has revolutionized mental health care and eased the suffering of untold numbers of patients. But in spite of advances in drug treatments and greater understanding of neurochemical processes in mental illness, collective insight into the mechanisms of action of almost all psychotropic drugs remains rudimentary. And, decades of clinical experience notwithstanding, there do not as yet exist universally defined, science-based criteria to guide the clinician in the selection and use of specific drugs for the treatment of any mental disorder.

These gaps in knowledge can partly be ascribed to the incomplete maturation of the mental health field. Additionally, a history of what must be considered fairly "shaky research" methods has led to questions regarding generalizability of clinical trials (Healy, 1990; Kraemer, 1995). Certain issues, such as the existence of a large placebo effect in psychotropic drug studies, continue to bedevil researchers in the field. The subject of generalizability has aroused considerable recent controversy, and the question of whether short-term clinical trials of medication can translate to the real-world, clinical environment is being answered with increasing skepticism.

Psychiatry's long tradition of clinical practice based on professional consensus rather than academic (or "hard") science has also served to retard the understanding of true clinical responses to psychotropic drugs. For example, it is common clinical practice to use serotonin reuptake inhibitors (SRIs) for the treatment of most forms of depression, in spite of evidence suggesting that tricyclic antidepressants (TCAs) may be more efficacious in treating severe depressions. In this instance, the negative effects of using consensus based, rather than science based, criteria are in all probability minimal. In other examples, however, consensus based psychiatric treatments have had significant negative repercussions on patient outcomes. The phenomenon of rapid neuroleptization, for instance, a common clinical practice in the past, is thought to have led to increased incidence of the potentially fatal neuroleptic malignant syndrome.

Although unappreciated by many psychologists, the gap between research and clinical psychopharmacology practice remains large. Some have attempted to rectify these disparities by constructing treatment algorithms that guide clinicians as to appropriate drug use. These efforts are still embryonic, and, unfortunately, treatment algorithms themselves may once again be developed on the basis of expert consensus rather than clinical science. Also, drug algorithms only address the pharmacologic aspects of treatment and rarely attempt to include psychosocial or other treatments (Rudorfer, 1995).

It is safe to assume that pharmacotherapy alone is not a sufficient condition for recovery from most forms of mental distress. Although it has been proposed that psychotherapy is an "occasionally indicated, delegatable" function (Klein, 1995, p. 81), examination of available evidence should temper any rush towards sole reliance on pharmacotherapy as an effective treatment strategy. For the most part, this is due to the reality that a substantial placebo effect operates in most psychotropic drug studies. The high rate of placebo responding to antidepressants is a particularly robust finding. Meta-analytic review of a large number of clinical trials of antidepressants indicated, for example, that while 60%–70% of depressed patients respond to drugs alone, approximately 35% displayed a therapeutic response to placebo. As a result, only around 25% of participants could be considered as true drug responders (Preskorn, 1996). Up to 70% of drug response in less severe depressions may be a placebo response. Indeed, the placebo response is so well demonstrated that some have suggested the clinical use of "pill placebo" as an effective intervention for depression (Brown, 1994). Sizable placebo responses have also been observed for other classes of psychoactive drugs as well, a fact which raises not only clinical concern but also methodological issues, as it is possible that many, if not most, patients in clinical trials can detect the differences between active drugs and placebos.

2.33.5.4 Aging Populations—Aging Programs

Although society continues to mature and age, as a profession, psychology continues not to prepare graduate students effectively for the challenges of caring for an older population (VandenBos & DeLeon, in press). Psychology training programs must be in the forefront of developing psychological based treatment programs targeted towards the needs of the elderly. Students in psychopharmacology will require far greater expertise than is currently available in most standard training programs if they are to provide adequately for the psychological and pharmacological needs of older Americans. The high rates of overprescription in elderly Americans has received wide comment. Programs educating physicians and other health care providers regarding rational use of psychopharmacologic agents have resulted in significant reductions in prescriptions written for elders. Given the importance of psychosocial factors in the care of the elderly, we would again rhetorically ask: are any of the professional schools demonstrating an interest in establishing psychology-based teaching programs in nursing homes?

2.33.6 HEALTH CARE COSTS

The growth in psychology reflects changes in health care in the USA. In 1950, the USA spent $13 billion or 4.4% of its gross national product (GNP), on health services. Twenty years later in 1970, the USA spent $75 billion or 7.5% of its GNP on health services. By 1980, this had increased to $250 billion or 9.4% of the GNP. In the 1990s, health spending increased to $740 billion or 13% of GNP, with estimates it would increase to more than 16% by the year 2000 and between 27% and 43% by the year 2030 (Bingaman, Frank, & Billy, 1993; Frank & Ross, 1995).

Increasing costs of this nature cannot be sustained by the US economy. Many believe that the increased cost of health care has been the driving force behind managed care. The federal government represents the largest health care payer in the USA. Federal health care programs (i.e., Medicare and Medicaid) have undergone substantial, uncontrolled growth. The federal government's share of health care spending rose from 28.1% in 1992 to 31.7% in 1993. In the six-year period between 1992 and 1998, Medicaid is projected to grow by 10.5%, and Medicare is expected to grow by 9.4%. This growth in spending stands in sharp contrast to other federal expenditure growth which is estimated, for the same period, to be 2.5% (Frank, Sullivan, & DeLeon, 1994).

2.33.6.1 The National Health Care Debate

Awareness of increasing costs of the health care system led to a national debate on overhauling the health care system in the early days of the Clinton Administration (Frank, 1993; Glueckauf, Frank, Bond, & McGrew, 1996). This debate resulted in no new immediate federal legislation. The functional outcome of the debate, however, was a transition to reform at the state level and reform through private markets. It has become clear that only incremental change is possible at the federal level (Frank & VandenBos, 1994).

Within private markets, fragmented care has often been viewed as a sizeable factor in the escalation of cost. Most patients have multiple providers, practicing in different settings, providing multiple work-ups and multiple prescriptions for treatment. This fragmentation of care leads to diminished accountability for outcomes. In addition, multiple providers create numerous opportunities for overdiagnosis, excessive treatment, and iatrogenic outcomes. Payers, aware of the cost of fragmentation, have focused on streamlining delivery systems that emphasize accountability by restricting access to patients.

Integrated service networks (ISNs) have been touted as an effective ingredient in the implementation of systems focusing on quality and accountability. In the 1993–1994 health care debates, several states considered proposals to legislate ISNs (Frank & Kivlahan, 1997). Although few states were successful in legislating this approach, the combined power of payers has functionally moved many delivery systems in this direction. ISNs are networks of organizations that provide, or arrange to provide, a coordinated continuity of services to a defined population for whom they are clinically and fiscally accountable for outcomes and the health status of the population served. Organized delivery systems may be created through the vertical integration of organizational services or through the virtual integration created through a series of contracts. The goals of organized delivery systems differ dramatically from the goals of single entities practicing in the fragmented health care system. While systems focus on singular patient encounters, ISNs or organized health care systems focus on the health of the community through prevention and health promotion. Integral to the development of these complex organizations is the use of "holographic organizations." In these organizations, constituent parts (i.e., key roles and services) are included in smaller delivery systems so that every aspect of the organization replicates the larger entity.

These evolving changes in organized delivery systems have critical implications for the practice of psychology. As organized delivery systems have become more dominant in the practice landscape, the solo practitioner has become an endangered species. Pressure to combine and to join specialty groups or multispecialty groups has been intense. Payers have used the "power of contracting" to exclude individuals who are not part of larger groups providing discounts and accountability.

It now appears that capitation will evolve differently in unique markets. While some markets will move towards massive levels of capitation as seen on the West Coast and some portions of the Midwest and Rocky Mountain areas, other areas such as the Southeast will be slower to become dominated by managed care. Indeed, there seem to be countervailing trends which suggest that even within high managed-care environments, the move towards capitation may be diminishing. Instead, there seems to be a trend towards point-of-service (POS) plans. In POS plans, participants are charged a slight surcharge in return for access to a wide range of providers. These plans have spread dramatically within California. Consumer demands for "choice" have increased options for broader provider panels. In many states legislation guaranteeing access to certain classes of providers, such as dermatology or OB/GYN, has become common.

There has been increased concern that capitated models place specialists at unreasonable risk. Primary care providers, by referring their most complex expensive cases to specialists, can remain under their cap and "dump" the most complex cases. In several highly penetrated managed-care markets, new trends emphasizing "episode of care" payment have become more apparent. In this model, care is "unbundled" so that the focus is on the individual case or the disease level. The Oxford Health Plan (Norwalk, CT) has committed its future growth to this model. Oxford is developing criteria for the start and completion of each episode of care. All care provided within a designated period is defined as part of this care episode. The practitioner receives a fixed, global price for services and all specialists bid on the provision of services within the disease episode. Importantly, in this model, there is emphasis upon cross-disciplinary teams.

Congruent with this approach is the growth of a number of specialized companies focusing on "disease management." For example, "for-profit" companies specializing in cancer, diabetes, asthma, and rehabilitation have now become common. Within ISNs, care of chronic disease conditions provided by several disciplines has become the method of delivery. Instead of vertical integration, these systems are horizontally integrated across the episode of care.

HealthSouth, a major rehabilitation company headquartered in Birmingham, AL, has often been touted as a classic example of this strategy. In 1982, HealthSouth was founded as an outpatient rehabilitation company. By 1997, it was fully operational in 30 large markets and HealthSouth facilities were located in all 50 states. Operating revenues grew from $25 million in 1991 to $2.4 billion in 1996, with a 28.2% operating margin in 1996.

Changes of this type have important implications for the practice of psychology. Increasingly, it has been recognized that psychology is a health specialty not merely a mental health specialty (DeLeon, Frank, & Wedding, 1995). Interestingly, a number of psychologists have taken another step and now consider themselves "primary care providers," rather than "specialists" of any type.

The slowness of this definitional recognition has come at high cost. During the 1993–1994 Congressional debate psychology mistakenly placed its primary emphasis upon mental health benefits. Arguing that each American was entitled to 52 sessions of psychotherapy per year, organized psychology emphasized an approach that had little likelihood of succeeding. Organized psychology pursued this strategy despite considerable evidence that this proposition would garner little external support. Late in the debate psychologists recognized that the narrow advocacy of mental health benefits, as opposed to suggesting changes in the broader health care arena, limited the ability of the profession to significantly impact the outcome of health reform conversations. This change in position led to recognition that psychology is a health profession. Interestingly, as many as 75% of all members of the APA identify themselves as "health care providers."

While psychologists initially contented themselves with narrow, self-defeating arguments, other groups staked out positions in the larger policy landscape. The American Medical Association, for example, led all other organizations in total contributions to political action committees during the national elections. In contrast, psychologists divided themselves into small, argumentative camps debating the veracity of each others' approach to the profession. The schism between academic psychology and the clinical practice community is marked, and in our judgment is extremely destructive.

2.33.6.2 The Town-gown Split

While American physicians have demonstrated loyalty to their training institutions, such leanings are rare in psychology. Many physicians have lamented the increasing pressure for clinical dollars within medical schools. This trend, however, demands medical school faculties engage in clinical practice and has assured that academic physicians are aware of trends within their profession and remain proficient clinicians. Psychology, in contrast, has followed a more "ivory tower science" approach. In many academic enterprises psychologists have little institutional incentive to be personally involved in clinical practice. Consequently, graduates of these programs who move into the practice community have little subsequent contact with their academic faculty. This creates schisms between the academic and clinical practitioners. Clinical practitioners come to believe that the academics understand little about their day-to-day lives. Such beliefs seem reasonable when few academic psychologists engage in actual clinical practice. Unfortunately, one result is that a true understanding of the drastic changes evolving in the health care delivery system is severely limited among those in many academic departments.

2.33.6.3 Self-identity

The failure of psychologists to identify with their academic training institutions has other implications. As the number of psychologists increase, they are moving into diverse settings. Increasingly, psychologists are moving into policy roles in addition to traditional clinical, academic, or science roles. There are numerous opportunities for psychologists in the public policy area. Behavioral science expertise is critical to health and educational programs. The demand for individuals with knowledge of these areas creates opportunities for psychologists. In addition, the broad purview of psychology training programs guarantees that graduates will continue to be sought for many diverse opportunities (Lorion, Iscoe, DeLeon, & VandenBos, 1996). Unfortunately, a tendency to be divorced from academic homes creates little common identification among psychologists. The tendency for different camps within psychology to devalue colleagues who work in different settings guarantees that a common identity is unlikely for the profession. As psychologists move into more diverse niches, this lack of common identity will continue to result in individual psychologists identifying more with their new opportunities, rather than with their profession in general.

It has been noted that physicians in policy settings identify strongly with their medical training. They assure that opportunities for physicians are developed through their policy positions and promulgated rules and regulations. In contrast, psychologists often seem to use a "scientific" approach to assessing the value of extending psychology into new settings. This approach requires empirical validation of the usefulness of psychologists for each and every niche. The tendency for psychologists within different camps to devalue each other exacerbates this problem.

2.33.6.4 The Congressional Science Fellowship Program

Several programs have created special opportunities for psychologists to learn more about health policy formulation on a "first hand" basis. Although psychology curriculae rarely emphasize either health policy or public health, the broad training psychologists receive is ideal for the evaluation of policy issues.

In 1974, the APA created the Congressional Science Fellowship program. This program has now been responsible for facilitating the involvement of 61 individual psychologists in policy-making and the political process. Through the APA Fellowship, psychology's expertise is systematically applied to society's priorities. The program allows psychologists to spend a year in Washington, DC working in the legislative or administrative branches. Through this "hands on" experience, psychologists have the opportunity to see (and participate in) national public policy being made on a first-hand basis. Upon completing their Capitol Hill experience, approximately half of the Fellows take policy-related positions within various agencies, while the other half return to their previous positions.

Another program that holds tremendous opportunity, but which unfortunately has been only rarely considered by psychologists, is the Robert Wood Johnson Health Policy Fellowship (RWJ). Like the APA Fellowship, this program was created in the mid-1970s to provide mid-career academic health center faculty the opportunity to learn about public policy through actual experiences. The RWJ Fellowship is administered by the Institute of Medicine (IOM) at the National Academies of Sciences. Each year, through a competitive process, a few are selected from academic health centers and invited to spend a year in Washington, DC working within the legislative or administrative branches. Again, most RWJ Fellows choose to work within the Legislative Branch. The program provides an extensive didactic training on health policy prior to the actual Fellowship experience. Fellows are exposed to leaders from virtually every sector of the American health care system. After the didactic sessions are completed, fellows spend 10–14 months working on federal health policy. To date, only two psychologists have taken advantage of the RWJ—Danny Wedding and Robert G. Frank.

Within health care, there are many critical issues which will affect the future of professional psychology. In the late 1990s most psychologists are concerned with those changes in health care delivery that impact their own economic status. Without question, such concerns are vital to the future of individual psychologists in practice. On the larger scale, however, psychology as a profession must be concerned about issues which open up new markets and expand the scope of practice, such as the acquisition of the ability to prescribe medication (DeLeon & Wiggins, 1996). In the 1980s concerns were voiced within the profession regarding whether it should seek authority to admit to hospitals. Each of these professional expansion decisions is critical to the future and vitality of the profession and each reflects the interface between psychology and organized medicine. For each issue, organized medicine has a history of standing between psychology and the ability to achieve the goal. As concerns regarding the number of providers and the level of utilization of health care services have increased, all types of health professionals have found themselves pitted against each other in a battle to win the health care dollar.

Since the 1890s organized medicine has demonstrated that it is the most effective health profession in the public relations arena. Starting at the turn of the twentieth century, organized medicine teamed with hospitals to create the ultimate public relations campaign. These organizations created the concept that health care must be provided by physicians in hospitals to assure the health and well-being of the American people. While there are clearly elements of truth within this premise, the most striking aspect of this partnership was the ability of these two groups to essentially vanquish other competitors such as osteopathic medicine.

2.33.6.5 The Professional School Movement

Organized medicine has been nothing short of remarkably successful at convincing the American public they should, using tax dollars, finance the education of American physicians. Starting with the passage of Medicare in 1965, taxpayers have subsidized the training of each American physician. These subsidies support both medical school education and, to a larger degree, training in residencies. Despite the fact that it now appears there is a clear excess of specialist physicians in the US, since the 1960s Americans have provided extraordinary subsidization for the training of American physicians. In 1997 alone, each American spent $48.91 on graduate medical education (GME). Annually, the federal government spends about $75 000 for each resident in training. The overall cost of the GME program exceeded $6 billion in 1997 (Medicare Payment Advisory Commission [MedPAC], 1997). Provisions included in the Balanced Budget Act of 1997 are expected to begin to curtail the program.

There is an important "reverse psychology" to the marketing support for graduate medical education. Simply put, if Americans spend this much to support the training of their physicians, physicians must be more valuable than any other health professionals. Indeed, if one examines the record of support for other health professions it is negligible in contrast to the monies provided to physician training.

Psychology, in particular, receives a very limited amount of support under Title VII (the Health Professions provisions) of the Public Health Service Act. Although medicine and the other nonphysician health care professions receive support under this initiative, support for psychology's internships or residencies is virtually nonexistent. It is the inclusion of physicians and professional nursing under GME and Title VII (or VIII for nursing) that creates the circular outcome that more money must be spent on physician training and advanced practice nursing and less on other health professions. Since psychologists are not formally included under the various federal government training programs, the limited value of the profession to the health of the nation is continuously demonstrated. Without inclusion, it is unlikely that psychology will become a critical health profession in the USA (Dunivin, 1994). And, since the vast majority of psychology's graduate students are now within professional schools, it is their leadership that must steer this legislative recognition.

2.33.7 OPPORTUNITIES FOR THE FUTURE

As psychologists have transitioned from mental health providers to health care providers, the opportunities have never been greater (Johnstone et al., 1995). Since the 1970s the field of psychology has changed substantially. One simple indication of the magnitude of this change is the increase in the number of psychologists. Between 1974 and 1990, it is estimated that the number of psychologists increased by 300%. This growth was accompanied by a 72% increase in the number of psychology training programs. If these estimates are correct, psychology's overall workforce increased from a ratio of 8.6 psychologists per 100 000 people in 1959 to nearly 22.4 psychologists per 100 000 people in 1989. There are definitely the numbers to be a major player in the legislative and health care arenas.

To cite one concrete example of an evolving opportunity: in the late 1990s, as many as 100 million Americans have chronic health conditions that limit one or more of their activities. Chronic health conditions engender significant cost. Individuals with activity limitations account for only 17% of the population but 40% of all medical expenditures. The per capita cost for individuals with chronic conditions are three times greater than individuals without these conditions.

Psychologists frequently care for individuals with chronic health conditions. Within settings such as rehabilitation hospitals, psychologists perform diverse activities ranging from provision of direct psychological care to the organization and management of all services to individuals with disabilities. In each of these examples psychologists are critical to the well-being of these individuals with chronic health conditions. The overall value of psychologist contributors in these types of health settings is intuitive but not well demonstrated empirically. Undoubtedly, as is reflected by the trend towards disease management models, more attention will be directed toward chronic health conditions in the future. Psychologists have the opportunity to demonstrate their critical role in treating those who are afflicted with chronic conditions.

It is a political and public policy decision as to which profession's training programs will be supported by the federal government. Psychology possesses the numbers, maturation, and clinical value to qualify. The underlying issue is: do psychologists collectively possess the necessary political will? There are many areas within the health care system in which psychological expertise can make a very real difference in the lives of the nation's citizens. The underlying question remains, however: will the next generation of psychologists take up the challenge?

2.33.8 CONCLUSION

Psychology is one of the "learned professions" and as such, possesses the responsibility of effectively addressing society's pressing needs. During the past decade, our nation's health care system has undergone tremendous structural changes which have had direct implications for the profession. These have not been easy times for the profession, especially for those in solo private practice. There has been an unprecedented acceptance of managed care, increasing requirements for professional accountability and the provision of cost-effective services, and a gradual awareness of the clinical potential inherent in technological advances within the communications and computer industries—that is, telehealth. Simultaneously, with the advent of the professional school movement there has been an exponential growth in the number of professional psychologists.

The period to 2010 will provide professional psychology with numerous opportunities and challenges. To survive and flourish, the profession must internally decide what are its strengths and where can it best excel—psychology is no longer merely one of the mental health professions. Psychologists must ensure that their academic and practice communities work together to address society's needs. Concerted efforts, especially by the leadership of the professional schools, must be made to ensure that psychology is expressly recognized as one of the health professions throughout the federal government's reimbursement and health training initiatives (i.e., GME and Title VII). The profession must further systematically seek to ensure that its "scope of practice" reflects advances in training, including obtaining prescription privileges and working collaboratively throughout the health care arena.

The behavioral sciences and specifically psychology have much to offer to the nation's health care struggles. It is unclear, however, whether the next generation of graduates is collectively willing to become personally involved in the necessary political and public policy endeavors.

ACKNOWLEDGMENT

The opinions expressed by Morgan T. Sammons represent his views as a private citizen and in no way represent the official views or positions of the US Navy or the Department of Defense.

2.33.9 REFERENCES

Balanced Budget Act of 1997, Pub. L. No. 105–33, 111 Stat. 251 (August 5, 1997). H. Conf. Rpt. #105–217.

Benson, C. J., Schreck, R. C., Underwood, F. B., & Greathouse, D. G. (1995). The role of army physical therapists as nonphysican health care providers who prescribe certain medications: observations and experiences. *Physical Therapy*, *75*(5), 380–386.

Bingaman, J., Frank, R. G., & Billy, C. L. (1993). Combining a global budget with a market driven delivery system: Can it be done? *American Psychologist*, *48*(3), 270–276.

Blanck, R. R., & DeLeon, P. H. (1996). News from Washington, DC—Managed care: Strongly conflicting views. *Professional Psychology: Research and Practice*, *27*(4), 323–324.

Brown, W. A. (1994). Placebo as a treatment for depression. *Neuropsychopharmacology*, *10*(4), 265–269.

DeLeon, P. H. (1989). New roles for "old" psychologists. *The Clinical Psychologist*, *42*(1), 8–11.

DeLeon, P. H., Bulatao, E. Q., & VandenBos, G. R. (1994). Federal government initiatives in managed health care. In S. A. Shueman, S. L. Mayhugh, & B. S. Gould (Eds.), *Managed behavioral health care: An industry perspective* (pp. 97–112). Springfield, IL: Charles C. Thomas.

DeLeon, P. H., Fox, R. E., & Graham, S. R. (1991). Prescription privileges: Psychology's next frontier? *American Psychologist*, *46*, 384–393.

DeLeon, P. H., Frank, R. G., & Wedding, D. (1995). Health psychology and public policy: The political process. *Health Psychology*, *14*(6), 493–499.

DeLeon, P. H., Kjervik, D. K., Kraut, A. G., & VandenBos, G. R. (1985). Psychology and nursing: A natural alliance. *American Psychologist*, *40*, 1153–1164.

DeLeon, P. H., Sammons, M. T., & Sexton, J. L. (1995). Focusing on society's real needs: Responsibility and prescription privileges? *American Psychologist*, *50*(12), 1022–1032.

DeLeon, P. H., & VandenBos, G. R. (1991). Psychotherapy in managed health care: Integrating federal policy with clinical practice. In C. S. Austad & W. H. Berman (Eds.), *Psychotherapy in managed health care: The optimal use of time & resources* (pp. 251–263). Washington, DC: American Psychological Association.

DeLeon, P. H., VandenBos, G. R., & Bulatao, E. Q. (1991). Managed mental health care: A history of the federal policy initiative. *Professional Psychology: Research and Practice*, *22*, 15–25.

DeLeon, P. H., & Wiggins, J. G. (1996). Prescription privileges for psychologists. *American Psychologist*, *51*(3), 225–229.

Dorken, H. (1990). Malpractice claims experience of psychologists: Policy issues, cost-comparisons with psychiatrists, and prescription privileges implications. *Professional Psychology: Research and Practice*, *21*, 150–152.

Dunivin, D. L. (1994). Health professions education: The shaping of a discipline. *American Psychologist, 49,* 868–878.

Fiscal Year 1989 Department of Defense Appropriations Act, Pub. L. No. 100–463, 102 Stat. 2270 (October 1, 1988).

Fiscal Year 1998 Department of Defense Appropriations Act, Pub. L. No. 105–56, 111 Stat. 1203 (October 8, 1997). S. Rpt. #105–45; H. Conf. Rpt. #105–265.

Fiscal Year 1998 Departments of Veterans Affairs and Housing and Urban Development, and Independent Agencies Appropriations Act, Pub. L. No. 105–65, 111 Stat. 1344 (October 27, 1997). Sen Rpt. #105–53.

Flannery, F. T. (1997). Liability in the diagnosis and treatment of myocardial infarction: The physician's insurers association of America study. *Legal Medicine Open File 1997,* 12–15.

Frank, R. G. (1993). Health care reform: An introduction. *American Psychologist, 48*(3), 258–260.

Frank, R. G., & Kivlahan, C. (1997). *The use of information and misinformation in a state health reform initiative*. Washington, DC: National Academy Press.

Frank, R. G., & Ross, M. J. (1995). The changing workforce. The role of health psychology. *Health Psychology, 14*(16), 519–525.

Frank, R. G., Sullivan, M. J., & DeLeon, P. H. (1994). Health care reform in the states. *American Psychologist, 49*(10), 855–867.

Frank, R. G., & VandenBos, G. R. (1994). Health care reform: The 1993–1994 evolution. *American Psychologist, 94*(10), 851–854.

Glueckauf, R. L., Frank, R. G., Bond, G. R., & McGrew, J. H. (Eds.) (1996). *Psychological practice in a changing health care system*. New York: Springer.

Health Insurance Portability and Accountability Act of 1966, Pub. L. No. 104–191, 110 Stat. 1936 (August 21, 1996). H. Conf. Rpt. #104–736.

Healy, D. T. (1990). The psychopharmacological era: Notes towards a history. *Journal of Psychopharmacology, 4*(3), 152–167.

Institute of Medicine. (1996). *Telemedicine: A guide to assessing telecommunications in health care*. Washington, DC: National Academy Press.

Johnstone, B., Frank, R. G., Belar, C., Berk, S., Bieliauskas, L. A., Bigler, E. D., Caplan, B., Elliott, T. R., Glueckauf, R. L., Kaplan, R. M., Kreutzer, J. S., Mateer, C. A., Patterson, D., Puente, A. E., Richards, J. S., Rosenthal, M., Sherer, M., Shewchuk, R., Siegel, L. J., & Sweet, J. J. (1995). Psychology in health care: Future directions. *Professional Psychology: Research and Practice, 26*(4), 341–365.

Jones, C. B., Cohen, S. A., Munsat, P. E., Dorris, J. F., & Berson, B. S. (1996). *Cost-effectiveness and feasibility of the DoD psychopharmacology demonstration project: Final report* (VRI-DSID3-1.08 WP96-1 (R.01)). Arlington, VA: Vector Research.

Klein, D. F. (1995). Psychopharmacological practice and health reform. *Psychiatric Annals, 25,* 79–83.

Kraemer, H. C. (1995). Methodological and statistical progress in psychiatric clinical research: A statistician's perspective. In F. E. Bloom & D. J. Kupfer (Eds.), *Psychopharmacology: The fourth generation of progress* (pp. 1849–1860). New York: Raven.

Lorion, R. P., Iscoe, I., DeLeon, P. H., & VandenBos, G. R. (Eds.) (1996). *Psychology and public policy: Balancing public service and professional need.* Washington, DC: American Psychological Association.

Luker, K., Austin, A., Hogg, C., Ferguson, B., & Smith, K. (1997). Patient's views of nurse prescribing. *Nursing Times, 93,* 51–54.

Medicare Payment Advisory Commission (MedPAC). (1997, October). *Selected medicare payment policies.* Washington, DC: Department of Health and Human Services.

Miller, I. J. (1996). Managed care is harmful to outpatient mental health services: A call for accountability. *Professional Psychology: Research and Practice, 27*(4), 349–363.

Murray, C. H., & Lopez, A. D. (1997). Global mortality, disability, and the contribution of risk factors: Global burden of disease study. *Lancet, 349,* 1436–1442.

National Health Policy Forum (1997). *Issue brief. Mental health parity: Unresolved issues affecting employers, consumers, and insurance companies.* (No. 709). Washington, DC: George Washington University.

Nonphysician prescribing doubles (1996, June 17). *American Medical News, 39*(23), 2.

Olfson, M., Pincus, H. A., & Sabshin, M. A. (1994). Pharmacotherapy in outpatient psychiatric practice. *American Journal of Psychiatry, 151*(4), 580–585.

Peterson, B. D., West, J., Pincus, H. A., Kohout, J., Pion, G. M., Wicherski, M. M., Vandivort-Warren, R. E., Palmiter, M. L., Merwin, E. I., Fox, J. C., Clawson, T. W., Rhodes, K. K., Stockton, R., Ambrose, J. P., Blankertz, L., Dwyer, K. P., Stanhope, V., Fleischer, M. S., Goldsmith, H. F., Witkin, M. A., Atay, J. E., & Manderscheid, R. W. (1997). An update on human resources in mental health. In R. W. Manderscheid & M. A. Sonnenschein (Eds.), *Mental health, United States, 1996* (pp. 168–204). Rockville, MD: US Department of Health and Human Services.

Platman, S. R., Dorgan, R., & Gerhardt, R. J. (1976). Some social and political ramifications of utilizing nonphysicians as chemotherapists. *International Journal of Social Psychiatry, 22,* 65–69.

Preskorn, S. H. (1996). A dangerous idea. *Journal of Practical Psychiatry and Behavioral Health, 1,* 231–234.

Request for comments on legal issues relating to telemedicine. (1996, August 9). *Federal Register, 61*(155), 41640.

Resnick, R. J., DeLeon, P. H., & VandenBos, G. R. (1997). Evolution of professional issues in psychology: Training standards, legislative recognition, and boundaries of practice. In J. R. Matthews & C. E. Walker (Eds.), *Basic skills and professional issues in clinical psychology* (pp. 281–303). Needham Heights, MA: Allyn and Bacon.

Rudorfer, M. V. (1995). Editor's note. *Psychopharmacology Bulletin, 31,* 453–455.

Sammons, M. T., & Brown, A. B. (1997). The Department of Defense psychopharmacology demonstration project: An evolving program for postdoctoral education in psychology. *Professional Psychology: Research and Practice, 28*(2), 107–112.

Sanchez, P. N., & Kohn, M. (1991). Differentiating medical from psychological disorders: How do medical and nonmedically trained clinicians compare? *Professional Psychology: Research and Training, 22,* 124–126.

Sierles, F. S., & Taylor, M. A. (1995). Decline of US medical student career choice of psychiatry and what to do about it. *American Journal of Psychiatry, 152,* 1416–1426.

State of California, Office of Statewide Health Planning and Development, Division of Health Professions Development. (1982, November). *Prescribing and dispensing pilot projects* (Final report to the legislature and to the healing arts licensing boards). Sacramento, CA: Author.

Telemedicine activity triples in one year, survey says (1997, November 10). *American Medical News, 40*(42), 17.

US Congress, Office of Technology Assessment (1995). *Telecommunications technology and Native Americans:*

Opportunities and challenges. (OTA-ITC-621). Washington, DC: US Government Printing Office.

US Department of Commerce (1997, January). *Telemedicine report to Congress.* (#1997–418–626/42023). Washington, DC: US Government Printing Office.

US Department of Defense (1996, June 5). Testimony on Fiscal Year 1997 Department of Defense Appropriations. S. Hrg #104–756. (ISBN 0–16–054069–0). Washington, DC: US Government Printing Office.

US General Accounting Office (1997a, February). *Telemedicine: Federal strategy is needed to guide investments.* (#GAO/NSIAD/HEHS-97–67). Washington, DC: US Government Printing Office.

US General Accounting Office (1997b, April). *Defense health care: Need for more prescribing psychologists is not adequately justified.* (GAO/HEHS-97–83). Washington, DC: US Government Printing Office.

VandenBos, G. R., & DeLeon, P. H. (in press). Clinical geropsychology and US federal policy. In M. Hersen & V. B. Van Hasselt (Eds.), *Handbook of Clinical Geropsychology*. New York: Plenum.

VandenBos, G. R., DeLeon, P. H., & Belar, C. D. (1991). How many psychologists are needed? It's too early to know! *Professional Psychology: Research and Practice, 22*(6), 441–448.

Wennberg, J. E. (1984, November). Testimony before the US Senate Appropriations Committee. Variations in Medical Practice. (S. Hrg. #98–1239). pp. 2–37.

West, J. C., Zarin, D. A., & Pincus, H. A. (1997). Clinical and psychopharmacologic practice patterns of psychiatrists in routine practice. *Psychopharmacology Bulletin, 33*(1), 79–86.

Whitcomb, M. E., & Miller, R. S. (1995). Participation of international medical graduates in graduate medical education and hospital care to the poor. *Journal of the American Medical Association, 274,* 696–699.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.34

The Evolution of Managed Behavioral Health Care: A Psychologist's Perspective on the Role of Psychology

CLARISSA C. MARQUES
Green Spring Health Services Inc., Columbia, MD, USA

2.34.1 CHANGE AND INSTABILITY: CONTEXT FOR MANAGED BEHAVIORAL HEALTH CARE 411

2.34.2 CONCEPTS OF MANAGED BEHAVIORAL HEALTH CARE 412
- *2.34.2.1 Types of Programs* 412
- *2.34.2.2 Populations Served* 414
  - *2.34.2.2.1 Need for service and utilization of resources* 414
  - *2.34.2.2.2 Demographic variables* 414
  - *2.34.2.2.3 Presence of comorbid medical/health conditions* 414
  - *2.34.2.2.4 Presence of previous mental health and/or substance abuse treatment* 414
- *2.34.2.3 Service Delivery Systems* 415
  - *2.34.2.3.1 Open access* 415
  - *2.34.2.3.2 Directed access* 415
  - *2.34.2.3.3 Guided access* 416
  - *2.34.2.3.4 Delegated access* 416
  - *2.34.2.3.5 Integrated delivery systems* 416
- *2.34.2.4 Benefits Managed* 417
- *2.34.2.5 Funding Arrangements Offered* 417

2.34.3 QUALITY AND ACCOUNTABILITY 418
- *2.34.3.1 Access to Health Care Data* 418
- *2.34.3.2 Management Information Systems to Support the Collection of Health Care Data* 419
- *2.34.3.3 Opportunities for Outcomes Measurement* 419
- *2.34.3.4 Infrastructure to Support Evaluation and Analysis* 419

2.34.4 IMPACT OF THE ADVOCACY AND CONSUMER MOVEMENTS 419

2.34.5 PROMISE AND CHALLENGE OF PSYCHOLOGY 420

2.34.6 REFERENCES 420

2.34.1 CHANGE AND INSTABILITY: CONTEXT FOR MANAGED BEHAVIORAL HEALTH CARE

The delivery of behavioral health care services can be characterized by a process of continuous, if not at times chaotic, changes in response to a rapidly evolving health care environment and instability in the role of the behavioral health care provider in providing services to consumers. It is easy, if not simplistic, to view the changes as precipitated by the advent of managed care itself, but the evolution of managed behavioral health care lies in the behavioral health service system itself and the economics of providing behavioral health care services (Feldman, 1993; Mechanic, Schlesinger, & McAlpine, 1995). The context for the development of managed behavioral health care was, at least in part, provided by the system for delivering behavioral health services. In particular, four aspects of the delivery system provided context for the development of managed behavioral health care.

(i) Diversity in the professional categories of providers delivering behavioral health care services and the training associated with each category of provider to provide behavioral health care services

(ii) Variety in the therapeutic orientation(s) of the behavioral health care providers even within any given professional category of provider.

(iii) Variance in the modalities of treatment available for individuals with mental disorders with limited information to compare the outcome of care by modality.

(iv) Fragmentation in the continuity of care for individuals across levels of care or between providers within the same level of care.

While these characteristics of the delivery system could have been valuable in generating multiple approaches to the challenges of providing services to individuals with mental disorders, instead these characteristics led to confusion for the consumer and added costs for the payor. There was little in the way of consensus-based decision-making for matching a consumer's behavioral health care needs with a type of provider (or providers) best able to deliver the recommended modality (or modalities) of treatment to achieve the optimal outcome for the individual for the dollar spent on service. Each individual was on his or her own to navigate the delivery system and each payer complained that there must be a better way to attain value for the health care dollar. Managed behavioral health care grew out of an effort to stem the raising costs of mental health and substance abuse services, while rationalizing a delivery system that appeared fragmented and confusing to consumers and payors (Berlant, Trabin, & Anderson, 1994).

2.34.2 CONCEPTS OF MANAGED BEHAVIORAL HEALTH CARE

Managed behavioral health care, as an industry, is frequently characterized as a single program without differentiation. However, the reality is that managed behavioral health care organizations vary by key dimensions in a matrix of the: type of program(s) offered, populations served, service delivery system(s) offered, benefit(s) managed, and funding arrangement(s) offered to purchasers. Each of these characteristics in combination can create a significantly different type of managed behavioral health care program with differing impacts on the consumers and providers of services.

2.34.2.1 Types of Programs

It is difficult to discuss the management of care separately from the delivery system that is providing the care. However, it is an important feature of managed behavioral health care to track where the roles of the manager and provider of care are the same and where these roles diverge. It is equally important to note the ability of the managed behavioral health care organization to direct consumers within the delivery system. The types of behavioral health care management can be seen in three broad categories of patient care management: (i) utilization management, (ii) staff practice management, or (iii) network management.

Utilization management is one of the earlier forms of managed behavioral health care derived from the peer review process. At its core, this form of managed behavioral health care consists of a review of the care, either inpatient or outpatient, conducted prospectively, concurrently, or retrospectively, to determine if the care being recommended or the care being provided meets the criteria of being medically necessary. A variant of utilization management is applying the utilization management principles for individuals with catastrophic or high-risk conditions only leading to intensive utilization management for a circumscribed group of individuals with high-risk conditions. For the consumer, utilization management does little to assist the individual in navigating among the categories of professional providers offering services, evaluating the appropriate treatment options, or monitoring the quality of care received. For the purchaser, utilization management contains the cost of care, but does little to assure that the health care dollar is being spent wisely. For the provider, utilization review entails constantly building a case for the medical necessity of the care being recommended and risking that a reviewer will not make the same determination as the provider putting payment for services at risk. One of the most appealing aspects of a utilization review program, from the consumer and provider perspectives, is that this form of managed care is essentially nondirectional in that the managed behavioral health care organization is not directing the consumer to any particular provider. Instead, the managed behavioral health care organization is making a medical necessity determination regarding the care that is being provided or recommended without directing the consumer to a particular provider. The provider continues to be the provider and manager of care with a peer reviewer overseeing the medical necessity of care for reimbursement purposes.

A staff model delivery of care has long been a mainstay of the publicly funded health care delivery system, but has become more common with staff model health maintenance organizations that delivered managed behavioral health care through a staff model environment (Feldman, 1994). At the core of staff model delivery systems is the delivery and management of most, if not all, of the behavioral health care services needed by a population of members through staff providers. All services required by members are provided through the staff model delivery system and any services needed outside of the staff model delivery system must be preauthorized by a representative of the staff model delivery system. Management of the care continues to rest with the staff model provider.

While this model worked well in areas with high geographic concentrations of members and where mental health care was integrated with other health care, it tended to not work as well for accounts with a wide geographic dispersion of members and where the health care was not well integrated with other health care options. National employers with multiple health care plans tended to not opt for staff models except in selected geographic areas with a high concentration of employees. For consumers their options were limited to receiving care through the staff model delivery system, with the attendant pressure on evening and weekend appointments. For providers there was no access to these consumers unless the provider was a member of the staff delivery system. For purchasers the options were clearer, tighter quality control was possible and costs contained, but the choice and access to providers was limited to the staff model delivery system.

Most staff model delivery systems have evolved into mixed staff and network model delivery systems, partly to respond to consumer need for greater choice and access to providers, but also to respond to the economic need to have providers in geographically distributed areas and to have a range of behavioral health care subspecialists, yet not having enough membership in areas to support a full staff model. Network providers augmented the existing staff model delivery system, providing greater geographic coverage, better access to appointments for consumers, and a wider array of subspecialty expertise. The development of mixed staff and network model delivery systems helped expand the range of health plans within geographic regions, but continued to have difficulty in meeting the needs of the national or regional employers who wanted consistency in the delivery of behavioral health care benefits. With the management of care within mixed staff and network model delivery systems, it became increasingly difficult to ensure that the treatment protocols and guidelines followed by the staff model delivery system were being applied equally and effectively across the panel of network providers. The larger the panel of network providers, the less likelihood of the provider feeling closely affiliated to the culture, the practice, and the economics of the health plan. The provider within the staff model delivery system maintained dual roles as the provider and internal manager of the care and an external manager of care provided within the network.

In response to national and regional employers, managed behavioral health organizations created national networks of participating providers all credentialed to meet standard requirements and included in practice management efforts to ensure some minimal consistency in the management of care. To guide the consumer to the manager of care was a representative of the managed behavioral health care organization, typically a behavioral health clinician, whose task was to assess the immediate needs of the consumer and match the consumer to the provider(s) within the network best able to initiate treatment. In the network model delivery system, the role of the provider and manager of care became truly separate. The manager of care became the consumer advocate coordinating aspects of care delivered through a provider delivery system and became responsible for overseeing the quality and outcome of care delivered by the provider. Purchasers of behavioral health care benefits could solve the dilemma of the diversity of providers, treatment modalities, outcome management methodologies and cost by holding a single managed behavioral health care organization accountable for all the complexity in the field, while retaining control of the quality of services and cost of care. For the consumer, the network model delivered wider access to providers with greater standardization in practice and a consumer advocate to coordinate care. For the provider, the network management model took away a key functional role in the management of care and left the provider struggling with the boundary between providing and managing care.

2.34.2.2 Populations Served

The population of patients to be served by managed behavioral health care organizations can also vary along key dimensions, such as the need for service and utilization of resources, demographic characteristics of the patient population, presence of comorbid medical/health conditions, and previous mental health and/or substance abuse treatment. An understanding of the population to be served provides the basis for the pricing and underwriting assumptions associated with the funding for the cost of care and for the design of the types of programs that will be offered.

2.34.2.2.1 Need for service and utilization of resources

A population-based perspective requires the examination of individuals in five broad clusters:

(i) those who *do not need* mental health and/or substance abuse services and *do not seek* those services;

(ii) those who *do need* mental health and/or substance abuse services and *do not seek* those services;

(iii) those who *do need* mental health and/or substance abuse service and *seek inappropriate* services;

(iv) those who *do need* mental health and/or substance abuse services and *do seek* those appropriately; and

(v) those who *do not need* mental health and/or substance abuse services, *yet seek* the services anyway.

Each cluster has a different impact on the managed behavioral health care organization, the broader health plan, the providers within the managed behavioral health care organization and the health plan organization, the consumer and his/her family, and the society in which the individual resides and the resources needed to sustain public health.

2.34.2.2.2 Demographic variables

Within these broad clusters, there are also the more traditional demographic characteristics of populations, including age, gender, marital status, and geographic location of residence. The demographic characteristics are typically available as part of a membership eligibility database determined at the time that the individual and his or her eligible dependents were enrolled within a health care program. These same membership eligibility databases tend to exclude information that would provide additional insight into the impacted populations, such as racial/ethnic affiliation, educational, occupational, and income level. These are variables that are sometimes inferred, but the absence of data on the broad population limits some of the programming that can be developed to address specific needs.

2.34.2.2.3 Presence of comorbid medical/health conditions

An intersecting population-based variable is the presence or absence of a comorbid medical/health condition. This variable is bidirectional. Individuals with one or more significant comorbid medical/health condition(s) may not have a specific mental health diagnosis, yet may experience an improved outcome with the presence of focused behavioral health intervention (Barlow, 1994a, 1994b). A good example is the improved mortality rate for individuals with breast cancer who receive a focused behavioral health intervention in conjunction with the prescribed treatment regimen as compared to those individuals receiving the treatment regimen without the behavioral health intervention. For individuals with an identifiable mental health diagnosis, the provision of behavioral health services may be complicated by the presence of one or more comorbid medical/health condition(s) (Pallak, Cummings, Dorken, & Henke, 1994). An example would be the difficulty in providing treatment for a reoccurring major depressive disorder with an individual who had recently undergone an organ transplant procedure. While the examples may seem obvious, the population-based variables related to the incidence and prevalence of medical/health conditions within a given population will influence the type of managed behavioral health care program and the funding to make the services available to individuals.

2.34.2.2.4 Presence of previous mental health and/or substance abuse treatment

If past behavior is the best predictor of future behavior, then previous mental health and/or substance abuse treatment is one of the indicators of the presence of serious and/or persistent mental illness. Approximately only one of five people suffering from a mental disorder seeks treatment, suggesting a significant unmet need in services. These individuals are likely to use the general medical delivery system, thereby increasing the overall health care costs (Pallak, et al., 1994).

2.34.2.3 Service Delivery Systems

Behavioral health care services have historically been offered in one of two types of delivery systems, either open access to any licensed provider or through staff model community based treatment programs. Managed behavioral health care organizations brought an array of other options for service delivery systems, including directed access, guided access, delegated access and integrated delivery systems.

2.34.2.3.1 Open access

Open access to any licensed provider had been the delivery system associated with indemnity insurance programs with the only limitation on service being an annual and/or lifetime benefit maximum. For consumers not covered under an indemnity insurance program, community mental health centers and other publicly funded community treatment resources were available for consumers seeking mental health and substance abuse specialty care. Additionally, mental health and substance abuse services were offered through staff model treatment programs associated with many staff model based health plans. The first limitation placed on an open-access delivery system was the establishment of a contracted panel or network of providers. However, most health insurance plans continued to market their traditional indemnity program as having in excess of 95% or more of all available physicians in a geographic area as an incentive for consumers to select their insurance product leading to an essentially open access to almost any licensed provider. These contracted networks, however, were built for a health insurance product, and, for the most part, determined that any licensed provider meant any licensed physician. Many of the contracted networks for indemnity insurance products excluded psychologists and rarely included any other mental health professional types. When psychologists were included, it was not with the intent to include broad participation, but tended to include psychologists in practice with a contracted network physician or psychologists with specific expertise that was not covered by the contracted network physician panel. Consumer access to nonphysician behavioral health care providers was limited to referral by a physician within the contracted network physician panel. Physicians made relatively few referrals, particularly for outpatient-based services, and referrals for inpatient-based services were made when a patient was in acute need of that level of care, which typically meant in need of a psychiatrist. As psychologists, other mental health professionals and consumers applied pressure to health insurance plans to include them as part of the contracted network of providers, health insurance plans turned to managed behavioral health care organizations to manage the diversity of behavioral health care providers and the challenges of credentialing and contracting with a network of specialists in the treatment of mental health and substance abuse disorders.

2.34.2.3.2 Directed access

Managed behavioral health care organizations brought a new twist to the development of contracted networks, the ability to contract with a wide array of behavioral health care providers by having credentialing standards specific to each professional category. Managed behavioral health care organizations also brought the subspecialty expertise needed to differentiate between the types of services offered by the behavioral health care provider and the levels of care available in the treatment continuum. Managed behavioral health care organizations created the network to handle all mental health and substance abuse service needs and created the basis for "carving out" the mental health and substance services from the health plan product. The consumer was offered the opportunity to access services through the managed behavioral health care organization directly rather than exclusively through the contracted network of physicians for the health plan. This development led to what has become known as a "gate-kept" model for service delivery. Because the consumer no longer needed to access services through a network physician for the health plan, access to the specialty mental health care network was directed by a representative, typically a clinician, of the managed behavioral health care organization to the provider (or providers) in the network able to address the consumers needs.

Directed access continues to exist, primarily in programs associated with health maintenance organizations, and where the services are provided within a more defined geographic area. The intent is to have a defined group of quality, multidisciplinary behavioral health care providers who can function as a part of an organized health care system and to discourage consumers from seeking services outside of the organized health care system both for cost and quality of care concerns. The development of directed access delivery systems also led to a significant decrease in the number of providers needed to deliver mental health and substance abuse services.

2.34.2.3.3 *Guided access*

Managed behavioral health care organizations also needed to address the requirements of purchasers, typically large employer groups, who desired access to a specialty mental health and substance abuse contracted network and wanted to permit consumers the ability to have broad access to providers. Consumer requirements drove the modifications to the directed access model of service delivery. In particular, consumers were concerned about the continuity of care and played an active role in nominating providers to be included in the network to minimize the disruption in care when a managed behavioral health care program was implemented. Consumers also found value in having a clinician at the managed behavioral health care organizations who was able to discuss the consumer's unique needs and able to assist the consumer by providing information about types of providers, types of treatment modalities, and expectations of treatment. The clinician at the managed behavioral health care organization became the consumer's advocate in jointly determining what services were needed, ensuring that the consumer had the referral(s) needed to address the situation, and creating options for flexible case management.

The guided access delivery system is inherently a more expensive program. However, from the consumer's perspective, it reduced the discontinuity of care created by more directed access programs and reduced the fragmentation that plagued the historic delivery system because the care was now more flexibly coordinated. Providers bemoaning the intrusion of managed behavioral health organizations into the treatment planning process would be surprised at the number of questions that consumers have regarding the provision of mental health and substance abuse services that consumers are unwilling or unable to address with a provider directly.

2.34.2.3.4 *Delegated access*

Perhaps one of the older forms of delegated access has been the community mental health center program. In delegated access delivery systems, a delivery system is designated as the entry point into the array of options offered by the delivery system. With community mental health centers, the designation as the point of entry was based on geographic catchment areas. To enhance access, many community mental health centers also created satellite sites to improve the availability of entry points. As the consumer entered the delivery system, the providers within the system were delegated the responsibility for managing the care within the limits of what was available using other systems, such as the state hospital system, when needed. Just as guided access to the delivery system was driven by consumer needs, delegated access to the delivery system was driven by provider needs, particularly organized group practices.

Many organized group practices have developed expertise in some of the key elements of the management of behavioral health care services and have positioned themselves as being capable of delivering those services with minimal oversight. Purchasers of managed behavioral health care services have also recognized the role that the organized group practices can play in minimizing administrative costs associated with managed behavioral, yet have remained concerned about handing the entirety of the process over to organized group practices who do not have the significant infrastructure found within managed behavioral health care organizations.

Managed behavioral health organizations have, for the most part, developed various programs for assessing the capacities of organized group practices to manage care, delegating key elements of the clinical management process, and overseeing the quality of care within the group practice. The implementation of a delegated access model can occur with or without any difference in the payment or funding methodology. In other words, organized group practices may have delegated access to outpatient services in which the group is paid on a fee-for-servicebasis, a discounted fee-for-service basis, a case rate basis, or a subcapitated basis. The payment methodology does not dictate the delegated arrangement, but may become a component of the arrangement. This is an important point to make in a time when many providers are seeking to take financial risk for care costs to gain control over the management of care. These are indeed separate issues and providers need to carefully evaluate their capacity to assume risk with the possibility of adverse selection across a smaller population base.

2.34.2.3.5 *Integrated delivery systems*

Integrated delivery system is a complex concept that means many different things depending upon one's perspective within the delivery system. One notion of an integrated delivery system was the concept of placing all the required levels of care within one integrated delivery system which will allow the purchaser and consumer of health care service to receive all the needed services within a seamless delivery system that would facilitate the needed access to the appropriate level of care at the appropriate time. The difficulty for integrated delivery systems within the behavioral health care arena is that relatively few individuals need the full range of integrated delivery system services. Integrated delivery systems do offer advantages under a case rate payment methodology for those individuals who will need the full continuum of services to manage an episode of care (i.e., inpatient hospitalization, partial hospitalization, intensive outpatient treatment, and ongoing outpatient treatment for maintenance and stabilization), but it is more difficult to have all individuals receive services from a single integrated delivery system because of the consumer issues raised under the directed access model.

Another notion of an integrated delivery system is the concept of providing managed behavioral health care services within a broader health care system. In some ways we have come full circle back to the original notion of delivering direct staff model services within a broader health plan delivery system such as a health maintenance program. Most managed behavioral health care organizations have created mechanisms to imbed the behavioral health services within a broader health plan delivery system by offering direct delivery of services (through staff or contracted clinicians) or by creating a subnetwork of specialty providers that function as the behavioral health service division within the broader health plan meeting the requirements of the health plan for integrated care.

The current state of the managed behavioral health care industry is that most managed behavioral health care organizations have some form of all of the options for delivery system access present within their networks. The terminology may vary from organization to organization, but the concept is a tiered network strategy built upon concentric circles. The widest element of the circle is the universe of all providers which is growing at an exponential rate, followed by the universe of contracted network providers, followed by yet smaller subsets of providers who have one or more specialty arrangements with a managed behavioral health care organization to address specific purchaser or consumer needs.

A provider's experience with a managed behavioral health care organization will be strongly shaped by the nature of the delivery system(s) within which he or she is providing services.

2.34.2.4 Benefits Managed

It goes without saying that health care benefits vary widely. But particularly in the area of mental health and/or substance abuse benefits, the variability is enormous. The variability in access to benefits has led to the extension of the Medicaid program to cover the uninsured and to the parity initiatives on the state and federal level.

Managed behavioral health care organizations can only manage an existing benefit. They can work to improve access and accountability for a limited benefit, but cannot change an existing benefit into a better benefit or create a benefit where one does not exist. Many consumers do not separately consider their mental health and/or substance abuse needs when selecting a health plan. Cost is still typically the primary consideration. Instead, the consumer typically becomes aware of a mental health and/or substance abuse benefit when a crisis emerges.

2.34.2.5 Funding Arrangements Offered

The funding arrangements for managed behavioral health care organizations have two components, the cost of care and the cost of the management of the care. There are three basic types of funding arrangements for managed behavioral health care organizations: (i) administrative services only (ASO), (ii) capitation, and (iii) performance-based risk corridor(s). Everything else is variation on a theme.

ASO is a funding arrangement in which the managed behavioral health care organization is paid for the services rendered in the management of the care. It does not include the cost of the care itself, although the services provided by the managed behavioral health care organization can be quite extensive. The services can include the costs associated with building and maintaining a broad national network, paying claims, handling all reporting requirements, providing 24 hour access to members, and coordinating all aspects of patient care.

Capitation is a funding arrangement in which the managed behavioral health care organization is provided the funds for the cost of care and for the services rendered in the management of care. The cost of care is determined by historical utilization of care when that data is available or by comparing the member characteristics and benefit structure with other comparable accounts and estimating utilization. The cost of care equation has two components, the utilization of services and the unit cost per service. Capitation has particularly impacted psychologists in clinical practice because of the cost per unit of service. The master's-prepared therapists have been willing to contract for a lower cost per unit of service. Physicians have been required inpatient and medication management services, which has enabled them to retain greater leverage on the cost per unit of service. Psychologists have been caught between the two.

Performance-based risk corridor funding has been a mechanism used by purchasers for ensuring accountability in managing the behavioral health care benefit. The managed behavioral health care organization places some or all of their administrative fee at risk for performance within a prescribed corridor along any number of dimensions. Corridors can be established around any number of performance indicators, such as utilization, member satisfaction, timeliness of reporting, and accuracy of claims processing. Corridors can also be established around the cost of care, allowing the managed behavioral health care organization to participate in the cost savings associated with containing the cost of care or to share in the risk associated with failing to meet the cost of care expectations.

Managed behavioral health care organizations are judged by the membership and the revenue associated with the membership. The revenue associated with ASO funding and performance-based risk corridor funding is much more limited than the funding associated with capitation because the cost of care is not included in the revenue equation. The risk associated with adverse selection is also embedded in the revenue associated with capitation, particularly in membership populations in which the historical utilization is not completely known or where the benefit has changed.

A survey conducted in 1995 suggested that of 185.7 million Americans with health insurance, over half (58.2% or 108 million) were enrolled in a managed behavioral health care organization. Managed behavioral health care revenues for 1994 were approximately $2.1 billion with approximately 20% of the membership enrolled in risk-based (capitation), network-based programs and 20% enrolled in nonrisk-based networks and the remainder in hybrid programs, including significant numbers of members involved in employee assistance programs (Oss, 1995).

2.34.3 QUALITY AND ACCOUNTABILITY

The value-added promise of managed behavioral health care organizations has been to improve the quality and accountability within the behavioral health care field by providing objective indicators of quality of care (Marques, Geraty, Harbin, Hoover, & Theis, 1994). The challenge lies more in the promise than in the reality. Managed behavioral health care organizations are poised to have all of the elements necessary to make a significant contribution to the quality of behavioral health care service. Poised because managed behavioral health care organizations have:

(i) access to health care data,
(ii) management information systems to support the collection of health-care data,
(iii) opportunities for outcomes measurement, and
(iv) infrastructures to support evaluation and analysis.

To date, the reality, for the most part, has been short of the opportunity.

2.34.3.1 Access to Health Care Data

All of the managed behavioral health care organizations are gathering enormous quantities of data. Expectations for the ways in which the data could be used have risen dramatically over the past five years. The major questions that impact managed behavioral health care organizations revolve around the internal issues related to the availability of the data, the currency of the data, the integrity of the data, the applicability of the data to questions pertaining to the outcome of care. Questions also revolve around the externally focused issues related to the ability to share the data within the concerns for privacy of health care data, the competitive nature of the organizations and the regulatory pressures impacting the industry. The good news is the increasing efforts at standardization of sets of data used by health care systems and managed behavioral health care organizations. The American Managed Behavioral Healthcare Association (AMBHA) established its own performance indicators for managed behavioral health care organizations in its Performance Measures for Managed Behavioral Health Care Programs (PERMS 1.0) initiative which prompted the members of AMBHA to begin capturing data and reporting comparative data in a common format structure (AMBHA, 1995). The bad news is the increasing expectations by purchasers and consumers about the timeliness, user-friendliness, and accessibility of the data. The corresponding expectations and costs to the managed behavioral health care organizations are shared with the providers, increasing the administrative workload on the providers to supply and audit needed data. There has also been another subtle, but significant, shift in the expectations regarding data availability. The shift is away from a retrospective framework to a more continuous, dynamic framework for the applicability of data. An example might help illustrate the shift. Managed behavioral health care organizations have been collecting patient or member satisfaction data for a number of years, yet that data reflects a retrospective evaluation of member satisfaction with a provider during a specified period of time that predates the survey process. The current purchaser and consumer environment places a premium on receiving that member feedback immediately following the receipt of service, incorporating the data into a dynamic provider profiling tool which then guides the next referral to that provider. The expectation in most cases exceeds the reality, however, the trend is definitely for data availability on a continuous or "real time" manner coupled with the expectation that the data will be used to manage providers and direct care (Bengen-Seltzer, 1995).

2.34.3.2 Management Information Systems to Support the Collection of Health Care Data

The search for the management information system optimized to meet the needs of managed care has been akin to the search for the Holy Grail. Most managed behavioral health care organizations are utilizing management information systems that were originally designed for managed health care, in particular, for utilization review programs. These management information systems do not provide the flexibility needed, in general, to address the current and changing needs of the managed behavioral health care environment.

Managed behavioral health care organizations that have grown through acquisition tend to have "legacy" systems that become a part of the system platform. Managed behavioral health care organizations that have grown through adding new membership in the rapid growth environment of the field have not been able to keep their management information systems updated due to the focus on rapid growth and expansion. Managed behavioral health care organizations that have not grown do not tend to have the resources to allocate to management information system infrastructure development.

On the positive side of the equation, managed behavioral health care organizations have placed a premium on flexibility in the design of the management information systems which has led to systems that are less rigid in architecture, more amenable to customization and change, more modular in their ability to connect to other applications, and better able to integrate with multiple claims payment systems. The fact that these management information systems are built from a flexible platform affords a better opportunity to electronically interface with providers than the structured systems used by health insurance plans.

2.34.3.3 Opportunities for Outcomes Measurement

A similar shift is occurring in the outcomes measurement activities. More traditional indicators of outcome had been evaluated against criteria of reliability and validity. Currently, indicators of outcome have the additional pressure of being confronted by business realities. Managed behavioral health care organizations are using outcomes measurement as one key aspect in the management of the network or providers. The demand is stemming from employers and purchasers of health care services to provide meaningful measures of patient improvement and treatment effectiveness and that the data is collected on an ongoing basis and used in decision-making regarding referral and continued care (Bengen-Seltzer, 1995).

2.34.3.4 Infrastructure to Support Evaluation and Analysis

This is the most vulnerable area within managed care organizations and has led to the development of strategic partnerships with academic and research institutions to provide the independence of evaluation and analysis. In the infancy of managed behavioral health care, much of the data was held as proprietary to the organization and not available for independent review. Given the highly competitive nature of the industry, the position was understandable. However, managed care organizations also had difficulty gaining acceptance of their research and evaluation efforts because of the taint of the information being collected for marketing purposes and had not being subjected to peer review processes. Currently the environment is opening up for collaborative approaches to addressing mental health and substance abuse policy issues with federally funded initiatives and through collaboration with research and training institutions.

2.34.4 IMPACT OF THE ADVOCACY AND CONSUMER MOVEMENTS

In evaluating managed behavioral health care organizations, one cannot underestimate the impact of the advocacy and consumer movements. Both forces have place mental health care at the forefront of a powerful political agenda to bring parity to the federal and state legislative bodies and worked to remove the stigma of mental illness. The parity movement itself would have had a more difficult uphill struggle without the real world experience of the managed behavioral health care organizations in managing care within even unlimited benefit environments. Parity, in the absence of methodologies for managing care, would have been a very difficult initiative to convince purchasers of health care services that the costs would not accelerate exponentially. The data from managed behavioral health care organizations currently providing services in states with parity initiatives implemented provided a basis for the cost estimates associated with the federal parity initiative. Increasingly, the coalitions to shape the mental health and substance abuse policies have forged partnerships between the managers of care and the individuals and advocates most interested in the behavioral services available for those most in need of those services.

2.34.5 PROMISE AND CHALLENGE OF PSYCHOLOGY

Training and education. There have been numerous discussions regarding the number of mental health providers who are entering the field each year and whether there is a surplus of providers. There is a perspective that has been shaped by years of working within managed behavioral health care organizations that can be shared. By all indicators, there is indeed a surplus of mental health providers. The more significant variable, though, is where are the providers electing to deliver service. In any major standard metropolitan area, there is a high likelihood that any given managed behavioral health care organization has a surplus of providers and has closed the network to new providers in that area or has restricted access into the network to providers with particular expertise (most frequently child and adolescent psychiatrists). However, the semirural and rural areas of the country continue to be seriously underserved by behavioral health care providers, as are the urban, inner-city areas, where some of the need for behavioral health care services is the greatest. If psychologists are to make a case for the early identification and timely treatment to address the unmet need and to support more equitable access to coverage for mental health and substance abuse disorders as afforded under various state and federal parity initiatives, the profession will need to actively address the distribution of resources to deliver the required services in underserved areas.

The second part of the discussion is focused on whether the training that psychologists receive has prepared them for the environment into which they will enter. Psychologists frequently report feeling professionally "squeezed" in managed behavioral health care networks between psychiatrists, who because of their medical training and licensure are uniquely situated to provide medication management and inpatient hospital based services, and other master's-prepared therapists, who may be more likely to accept a lower rate of reimbursement than psychologists. The concern seems to stem from the perceived lack of differentiation for psychologists relative to other mental health professionals, with the exception of psychological testing as a distinguishing skill. The unfortunate aspect of the discussion is the difficulty that psychologists have in seeing that the training provided to doctoral-level psychologists is unique among the mental health professionals, not just for the provision of behavioral health care services, but also for the management of behavioral health care systems of care. What most psychologists come to learn in professional practice is that the greatest strength is drawn from all of the subspecialty programs within psychology. Each subspecialty area contributes to a different knowledge and skill base, such as industrial and organizational psychology for the management of health care delivery systems, experimental psychology for the methodology needed to assess the outcome of care, statistical psychology for understanding the meaningfulness of data, social psychology for understanding the impact of social and family systems on the individual, and developmental psychology for understanding the staging process of human development. The list could continue, the profession has yet to prepare graduates to embrace the richness of the field and encourage psychologists to apply themselves to all aspects of the health care system.

There are many forces that will reshape managed behavioral health care. The opening of this chapter addressed the forces at work to create an ever-changing health care environment and those forces will continue to drive the managed behavioral health care organization to evolve to meet the needs of consumers and purchasers of behavioral health care services. The question remains whether psychologists, as a professional group representing multiple subspecialties, find a path for understanding and participation in the evolution of behavioral health care services.

2.34.6 REFERENCES

American Managed Behavioral Healthcare Association (1995). *Performance measures for managed behavioral healthcare programs*. Washington, DC: Author.

Barlow, D. H. (1994a). Psychological interventions in the era of managed competition. *Clinical Psychology: Science and Practice, 1,* 109–122.

Barlow, D. H. (1994b). Health care policy, psychotherapy research, and the future of psychotherapy. *American Psychologist, 51,* 1050–1058.

Bengen-Seltzer, B. (1995). *Fourth generation managed behavioral health: What does it look like?* Alexandria, VA: Manisses Communication Group.

Berlant, J., Trabin T., & Anderson, D. (1994). The value of mental health and chemical dependency benefits: more than meets the eye. In E. Sullivan (Ed.), *Driving down health care costs: Strategies and solutions.* Frederick, MD: Panel Publishers, Division of Aspen Publishers.

Feldman, S. (1993). Managed mental health services: Ideas and issues. In S. Feldman (Ed.), *Managed mental health* (pp. 3–26). Springfield, IL: Charles C. Thomas.

Feldman, S. (1994). Managed mental health: Community mental health revisited*? Managed Care Quarterly, 2,* 13–18.

Marques, C., Geraty, R., Harbin, H., Hoover, K., & Theis, J. (1994). Quality and access in the managed behavioral healthcare industry. *Behavioral Healthcare Tomorrow, 3,* 23–29.

Mechanic, D., Schlesinger, M., & McAlpine, D. (1995). Management of mental health and substance abuse services: State of the art and early results. *The Milbank Quarterly, 73,* 19–55.

Oss, M. (1995). *OPEN MINDS: The Behavioral Health Industry Analyst*, January.

Pallak, M., Cummings, N., Dorken. H., & Henke, C. (1994). Medical costs, Medicaid, and managed mental health treatment: The Hawaii study. *Managed Care Quarterly, 2,* 64–70.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.35

Demographic Characteristics of Psychologists and Supply/ Demand Issues: The Netherlands

PAUL M.G. EMMELKAMP
University of Amsterdam, The Netherlands
and
A. SCHOLING
University of Groningen, The Netherlands

2.35.1 INTRODUCTION 423
2.35.1.1 The Study of Psychology at the University 424
2.35.2 POSTGRADUATE EDUCATION 424
2.35.2.1 Psychotherapy Training Program 424
2.35.2.2 Clinical Psychology Training Program 425
2.35.3 HEALTH CARE PSYCHOLOGY TRAINING PROGRAM 425
2.35.3.1 Clinical Psychology Research Training Program 425
2.35.4 WORK SETTINGS OF CLINICAL PSYCHOLOGISTS AND PSYCHOTHERAPISTS 426
2.35.4.1 Psychologists Working in Mental Health Settings 426
2.35.4.2 Psychologists Working in General Hospitals 426
2.35.4.3 Psychotherapists 427
2.35.4.4 Relative Numbers of (Clinical) Psychologists, Psychotherapists, and Psychiatrists 427
2.35.5 JOB PROSPECTS 427
2.35.6 FUTURE DEVELOPMENTS 428
2.35.7 REFERENCES 428

2.35.1 INTRODUCTION

In 1996 there were approximately 20 000 psychologists in the Netherlands, a country which has nearly 15 million inhabitants. Less than one third of them was organized in the Dutch Psychological Association (Nederlands Instituut van Psychologen (NIP)). As far as clinical psychology is concerned, there are two formal professional registrations, that is, the license of "clinical psychologist" provided by the NIP and the license of "psychotherapist" provided by the government, but this situation will change, as will be discussed below. Since 1988, registration as a psychotherapist has been *conditio sine qua non* for registration as a clinical psychologist. Since 1986 formal course requirements and supervised practice have been required in order to be registered as a psychotherapist; since 1989 also for registration as clinical psychologist. The content of these programs are described in this chapter.

The first clinical psychologist in the Netherlands was appointed in 1942. His tasks were limited to diagnostics and he worked under the supervision of a psychiatrist. If there was interest in therapy in the 1940s and 1950s, this was primarily in psychodynamic therapy. With the rise of the nondirective and behavior therapies in the US and in Great Britain in the 1960s, an increasing number of clinical psychologists got involved in psychotherapy in the Netherlands.

There are some studies in the Netherlands which provide figures about the number and characteristics of psychologists who work in mental health settings (Beekman & Heike, 1991; Brunenberg, Neijmeijer, & Hutschemaekers, 1995; Hutschemaekers, Brunenberg, & Spek, 1993; Neijmeijer & Hutschemaekers, 1995; Trijsburg, Duivenvoorden, & Dijkstra, 1988; van Son & van der Staak, 1993). Estimates in this chapter are based on these studies.

2.35.1.1 The Study of Psychology at the University

In contrast to most other countries, in the Netherlands students follow primarily courses in psychology. After approximately one and a half years in which students have to follow courses in all fields of psychology and in methodology and statistics, students specialize in one of the main directions within psychology: developmental psychology, social psychology, psychonomy (general psychology), work and organizational psychology, personality psychology or clinical psychology, and health psychology. For another two and a half years students follow courses in their specialization, have to conduct research in this area, and undertake practice work under supervision.

Eight universities in the Netherlands offer university programs in clinical psychology or health psychology. The total number of graduates in psychology has stabilized since 1990 at about 1025. Clinical psychology has been one of the most popular specializations over the years; 25–30% of the students choose clinical psychology as their specialization.

2.35.2 POSTGRADUATE EDUCATION

After they have graduated from university (i.e., after the first phase of four years), students have several options. Most of them prefer further training in practice or research in the area of clinical psychology, but only few of them actually get the chance to enter one of the formal postdoctoral training programs. This training includes two training programs in clinical practice and one training program in clinical psychology research.

2.35.2.1 Psychotherapy Training Program

Students who want further education in clinical practice have several options. As noted above, until 1998 two formal registrations exist for psychologists working in the field of mental health care: psychotherapist and clinical psychologist. From 1998 this situation will change, as will be described below. First, some information about the contents and the numbers of trainees in the present training programs is given.

Both programs require several years of training after the academic study has been completed. The Psychotherapy Training program involves courses (440 hours) and supervised practice (at least 125 hours supervision over at least 500 hours psychotherapy). Further, trainees have to undergo some form of psychotherapy ("personal therapy," at least 50 hours). This training is accessible for graduates in (clinical) psychology and medicine, provided that they have completed the relevant courses in their university education. Introductory courses in the psychotherapy training program include psychopathology, general aspects of psychotherapy, skills for handling specific situations in clinical practice, systems theory, and introductions to the main psychotherapy orientations: (i) learning theory/cognitive-behavior therapy, (ii) client-centered therapy, and (iii) psychodynamic therapy.

After the introductory courses, students choose one of these three main orientations in psychotherapy. In this area they have to follow two of the following three specializations: children and adolescents, partner/family therapy, and group psychotherapy, both in theoretical work and in supervised practice. The supervised practice must extend over at least four and at most seven years; during this period the trainee must be working for at least two days a week in a mental health institute.

There is no requirement for the psychotherapy trainee to have a salaried appointment for the clinical practice part of the program; in fact, many trainees work as volunteers, while the costs of the courses total up to about $19 000, apart from the costs for personal therapy (about $4000) and the supervision (ranging from US$0 to 8000). Until 1998, no limitations were imposed on the numbers of psychotherapy trainees. In 1996, approximately 2500 people were registered as psychotherapists, while many were still in training. The estimate for 1997 is that 3016 persons will be registered as psychotherapists (Hutschemaekers, Brunenberg, & Spek, 1993).

2.35.2.2 Clinical Psychology Training Program

The Clinical Psychology Training program involves all the elements of the Psychotherapy Training program. Additionally, courses (*in toto* 740 hours) must be followed in clinical assessment, interventions not considered psychotherapy (crisis intervention and consultation, supportive interventions, pharmacotherapy), ethical aspects, and research in clinical practice. In addition to the requirements for psychotherapy training, an additional 140 hours of supervision are required, 70 about psychodiagnostic issues and another 70 about psychological interventions other than psychotherapy and about ethical aspects.

The training is spread over five years, in which students follow courses for one day a week while working in a mental health institute during the other four days. In contrast to the psychotherapy training program, applicants can follow this program only if they have paid employment in a mental health institute. The costs of this program total up to about $28 000, apart from the costs for personal therapy and supervision. In the period 1991–1997, approximately 350 people started the Clinical Psychology Training program.

2.35.3 HEALTH CARE PSYCHOLOGY TRAINING PROGRAM

Changes are taking place in the postgraduate training program. In 1998 the legal registration existed of two basic professions in the field of mental health care: "health care psychologist" and "psychotherapist." The new Psychotherapy Training program will be more or less comparable to the earlier one. However, in the new Psychotherapy Training program, systemic psychotherapy as an additional main orientation is included.

The Health Care Psychology Training program will spread over two years, in which trainees will have to follow courses (440 hours) during one day a week and acquire clinical experience during the other four days. The program is more or less comparable to the first two years of the earlier Clinical Psychology Training program. Some of the registered health care psychologists will be allowed, after completion of their training, to follow further specialization to be clinical psychologist, which will take another three years. For both the health care and the psychotherapy programs, only paid (partly subsidized) training posts will be acknowledged. As a consequence, the numbers of trainees will be drastically limited. In 1998 and 1999, only a total of 82 applicants each year will be allowed to start health care psychologist training and 42 psychotherapist training in the Netherlands.

Apart from these formal training programs, there is a possibility to obtain registration as a member of one of the various therapy associations, such as the associations of (cognitive) behavior therapy, client-centered therapy, psychodynamic therapy, group therapy, etc., but these registrations are not approved officially by the government. One of the "nonlegal" registrations that is becoming increasingly important is registration by the NIP as a "first-line psychologist." These psychologists offer low-threshold brief psychotherapy; a referral via a general practitioner or mental health institute is not required. Gradually more insurance companies pay a (limited) number of treatment sessions provided by these first-line psychologists.

2.35.3.1 Clinical Psychology Research Training Program

The main route to specialization in scientific research is via an appointment as an "assistant in training" ("Assistent in Opleiding" or AIO in Dutch). An AIO is appointed at a university, either for four years full-time, or for five years four days a week. The AIO appointment is a combination of training (including courses) and conducting a research project; in addition, it is customary that AIOs participate in teaching courses in the department where they are appointed.

The aim of the AIO appointment is to conduct, relatively independently, a scientific research project in the area of clinical psychology, resulting in a Ph.D. thesis. A typical AIO thesis consists of four to six (international) publications on a clinically relevant topic, for example, several experimental studies on cognitive processes, attentional bias (in patients and controls), or studies on the etiology, assessment, and treatment of specific anxiety or affective disorders. In the first year of their appointment, the ratio of research to training is supposed to be 55:45, and in the last year 85:15. In practice, most AIOs spend less time following courses than is formally prescribed.

There are no official data available about the numbers of new AIO vacancies; the estimate is that in the field of clinical psychology about 20 postgraduates start this route every year. The AIO appointment offers hardly any possibilities for obtaining formal clinical registrations, although most AIOs are involved in clinical work as part of their thesis.

2.35.4 WORK SETTINGS OF CLINICAL PSYCHOLOGISTS AND PSYCHOTHERAPISTS

In 1996 approximately 1200 people were licenced as clinical psychologists, but many more psychologists work in mental health settings. The chapter will include below figures for the numbers of clinical psychologists and psychotherapists and their respective fields.

2.35.4.1 Psychologists Working in Mental Health Settings

Beekman and Heike (1991) and Brunenberg, Neijmeijer, and Hutschemaeckers (1995) provide an overview of the kind of institutions in mental health care in which psychologists work. Based on these studies, the following estimates have been made of the number of full-time equivalents (FTEs) of psychologists working in mental health care settings (see Table 1).

Table 1 Number of psychologists (FTEs) working in specific settings.

Type of care	*Numbers*
Institutes	
Alcohol and drug centers	20
Psychiatric hospitals	545
Clinical child care: institution	15
Clinical child care: day center	60
Nursing homes	350
Institutes for mental retardates	380
Institutes for deaf and blind, etc.	60
Noninstitutional care	
School guidance	25
Community mental care	900
Centers for alcohol and drug abuse	55
Private practice	480
Other	85
Justice	
Penal institutes	75
Child custody	12

Source: Beekman and Heijke (1991); Brunenburg et al. (1995).

It is, therefore, estimated that 3082 FTE psychologists work in mental health care. These estimates exclude those psychologists working primarily in administrative functions, those who are primarily involved in teaching or research, and those who are employed as a psychotherapist or work in general hospitals (see below). Clearly, not all of the psychologists working in mental health settings are licenced clinical psychologists. The mean age of psychologists working in clinical settings is estimated to be 40 years; approximately 45% are below and 55% are over the mean age. Slightly more psychologists are female (54%) (Brunenberg et al., 1995). Based on the numbers of students enrolled in clinical psychology training programs, it is expected that the number of female clinical psychologists will increase in the future.

2.35.4.2 Psychologists Working in General Hospitals

Since the early 1950s, clinical psychologists have been employed in general hospitals. Although in the early days the emphasis was on assessment, more recently psychologists have been employed in treatment, teaching, and consultation with medical specialists. In 1988, the National Board of Health (Nationale Raad voor de Volksgezondheid, NRV) recommended that in every general hospital at least one psychologist should be employed. In 1995, there were 122 general hospitals in the Netherlands, with a mean of nearly 500 beds per hospital; 101 hospitals (82%) had a department of clinical psychology. The number of licenced clinical psychologists employed in general hospitals was estimated as 194 FTEs. An additional 70 nonlicensed FTE psychologists were also employed in general hospitals: half were trainees in a clinical psychology training program, a few others psychotherapists. In most hospitals clinical psychology is a separate department where clinical psychologists are considered to be "normal" members of the medical staff. Most patients are referred by the department of neurology, followed by the departments of child medicine, internal medicine, psychiatry, and cardiology (Neijmeijer & Hutschemaeckers, 1995).

The figure of the number of psychologists working in general hospitals in Neijmeijer and Hutschemaeckers' study (1995) is less than that in the study of Beekman and Heijke (1991). The higher number of psychologists in general hospitals in the latter study is presumably due to several factors. First, a number of smaller hospitals closed because of government regulations. Second, in the Beekman and Heijke study, psychologists who worked in administrative functions and psychotherapists were also counted as psychologists, which was not the case in Neijmeijer and Hutschemaeckers' study (1995).

2.35.4.3 Psychotherapists

In 1996, approximately 2500 psychotherapists were officially licenced by the Dutch government. Psychiatrists are not included in this figure, since they are not required to be licenced as a psychotherapists. Based on two studies (Hutschemaekers et al., 1993; Trijsburg et al., 1988), the mean age of psychotherapists is estimated as 48 years; nearly 10% are between the ages of 30 and 40 years; 55% are between the ages of 41 and 50 years and 35% are over age 50 years. Approximately half of the psychotherapists are female. Based on the numbers of students enrolled in postgraduate psychotherapy training programs, it is expected that female psychotherapists will outnumber male psychotherapists in the future. Approximately 70% of the psychotherapists have a university degree in psychology (Hutschemaeckers et al., 1993), but this figure will also change in the future. Nearly all students in psychotherapy training programs have a university degree in psychology.

As to the therapeutic schools, most psychotherapists are members of either the association of cognitive-behavior therapy (30%) or the association of family therapy (29%), followed by the associations of group psychotherapy, client-centered psychotherapy, and child psychotherapy. Only 11% are members of one of the three psychoanalytical psychotherapy associations which co-exist in the Netherlands.

Most psychotherapists (60%) work in community mental health centers (RIAGGS); 20% are employed in a psychiatric hospital and 10% in a general hospital. More than half of the psychotherapists also have a private practice (Hutschemaeckers et al., 1993).

2.35.4.4 Relative Numbers of (Clinical) Psychologists, Psychotherapists, and Psychiatrists

Estimates have been made of the relative number of (clinical) psychologists, psychotherapists, and psychiatrists in the Netherlands. For every 10 000 inhabitants there are two psychologists, one psychotherapist, and one psychiatrist available (Brunenberg et al., 1995). The mental health services are spread across the country, but there are relatively more psychotherapists in the western part of the Netherlands where most larger cities are.

2.35.5 JOB PROSPECTS

Until the early 1970s the demands from society equalled the numbers of psychologists available on the market. Nevertheless, at that time gloomy prognoses were given for the future. If one thing has become clear over the past three decades, it is that accurate forecasts are difficult to make. In the 1970s it was predicted that the numbers of psychology students would decrease in the ensuring years, whereas the demands from society were expected to increase. The contrary has proven true. As a consequence, official data from employment offices show increasing numbers of unemployed psychologists: 330 in 1976/77 growing to 1760 in 1985/86 (Van Drunen & Breijer, 1994). Data from employment offices, however, are unreliable. Some authors (e.g., Beekman & Heijke, 1991) have estimated that about half of the those registered as searching for a job are not searching any longer, partly because they had found a job and partly because they had decided to stop searching for work. In addition, these data do not give information about the type of work that is conducted by clinical psychologists, and more specifically, whether they do the kind of work they have been trained for.

Since 1980, several studies have been conducted on demand/supply issues among persons who recently graduated from university. The studies were published between 1981 and 1996, and were reviewed by Brenninkmeijer and Scholing (1996). Most of the studies concerned the graduates from one university only. The percentages of responders varied from 49.4 to 93%. The percentages of persons unemployed varied from 10.2 to 51.3%. The low percentage came from a study with a response rate of 55.1% (Dons, 1992), the highest from a study with a 93% response rate (Dijkstra & Bannenberg, 1986), which might indicate that, in general, nonresponders tend to be more often unemployed. Although the data in these studies lend no clear support for this hypothesis, such selection bias cannot be ruled out.

Beekman and Heijke (1991) studied the job prospects for clinical psychologists in the interval 1990–1995. They estimated the total number of clinical psychologists searching for a job (the recently graduated and the unemployed) at about 4900 for those five years.

Most studies are not designed to reveal trends over the years. Two studies that were conducted at the department of Clinical Psychology in Groningen, however, lend themselves best to a comparison of data over the years, because (i) they refer to students that followed training in clinical psychology at the same university, lived in the same geographical area, in the time intervals 1983–1986 and 1987–1993, respectively, and (ii) almost identical questionnaires were used. The first study of 120 people (Schaufeli & De Vries, 1988) had a response of 75%; of these 78% were employed. The second study, of 250 people (Scholing & Brenninkmeijer, 1995) had a response of 60%; 83.5% were employed. First, it must be noted that in both studies responders were classified as being employed when they had paid work for at least four hours per week. The percentages unemployed are more than twice as high as the mean unemployment rate in the Netherlands (7–8%), but more favorable than the mean unemployment rate of students from other Dutch programs in clinical psychology. The unemployed were over-represented in the recently graduated: from the graduates of 1992 and 1993, 31.5% were unemployed vs. 13.6%, 5.9%, and 11.1% in the years 1987–1989, respectively.

For nearly 36% of the subjects from the Brenninkmeijer and Scholing (1996) study it took at least six months before they got their first paid job; for almost 20% it took at least a year. This implies a slight improvement compared with the situation for the subjects from the Schaufeli and De Vries (1988) study. It seems that it takes some time and effort before new psychologists find a job, but that many of them succeed eventually. In the second study, 30% of the respondents had a full-time job, in contrast with 45% in the first study. This is in line with the general trend in the Netherlands towards more part-time jobs, usually between 17 and 32 hours a week.

The next question concerns the type of work the responders found. In Table 2 a comparison is made between both studies. It shows a shift to more typically clinical work at the expense of research work (mainly at the university).

Table 2 Type of work, in percentages of numbers of responders.

Type	*Study I*[a]	*Study II*[b]
Clinical psychologist	54	68
Research psychologist	16	13
Academic work, but not as clinical psychologist	12	6
Psychologist	8	6
Under academic level	5	5
Other	5	2

Source: Schaufeli and De Vries (1988);[a] Brenninkmeijer and Scholing (1996).[b]

2.35.6 FUTURE DEVELOPMENTS

There are a few areas that will require more psychologists in the future. Most importantly, it is expected that an increasing number of psychologists will be employed in the care of the elderly. Further, an increasing number of psychologists will be working in penal institutes and with juvenile delinquents. Given the relatively few people addicted to hard drugs in the Netherlands, it is not expected that there will be a major increase in the number of psychologists working with addicts. Further, in contrast to most other countries, relatively few people are HIV positive, so special measures to increase the number of psychologists working in this area are not expected.

There are two other areas in which it is expected that there might need to be an increase in the number of psychologists. First, there is a clear need for clinical psychologists to work in firms and factories, for example, providing stress management training. Another growth area is that of behavioral medicine. Given the emphasis in the Dutch health-care system on accountability, it is expected that an increasing number of psychologists will work with stress-related illnesses that were traditionally the domain of medical practitioners.

2.35.7 REFERENCES

Beekman, T. B. J., & Heijke, J. A. M. (1991). Prognoses voor 1990–1995. De arbeidsmarkt voor klinisch psychologen. [Prognoses for 1990–1995. Employment for clinical psychologists]. *De Psycholoog, 26*(4), 187–193.

Brenninkmeijer, J., & Scholing, A. (1996). Werk en werkloosheid onder klinisch psychologen [Work and unemployment among clinical psychologists]. *De Psycholoog, 31,* 31–35.

Brunenberg, W., Neijmeijer, L., & Hutschemaekers, G. (1995). *Beroep: Psycholoog/pedagoog.* [Profession: Psychologist/child psychologist]. Utrecht: Nederlands Centrum Geestelijke volksgezondheid.

Dons, W. J. (1992). *Waar zijn al die psychologen gebleven?* [Where are all those psychologists?]. Unpublished manuscript. Catholic Univerity of Brabant.

Drunen, P. van, & Breijer, M. (1994). Meer en meer, al maar meer. De ontwikkeling van het aantal psychologen in Nederland. [More and more: The development of the number of psychologists in the Netherlands] *De Psycholoog, 28,* 16–20.

Dijkstra, C., & Bannenberg, J. (1986). Er zijn te veel klinisch psychologen. [There are too many clinical psychologists]. *De Psycholoog, 21,* 178–182

Hutschemaekers, G., Brunenberg, W., & Spek, H. (1993). *Beroep: Psychotherapeut.* [Profession: Psychotherapist]. Utrecht: Nederlands Centrum Geestelijke volksgezondheid.

Neijmeijer, L., & Hutschemaekers, G. (1995). *GGZ in het algemeen ziekenhuis.* [Mental health care in the general hospital]. Utrecht: Nederlands Centrum Geestelijke volksgezondheid.

Schaufeli, W., & Vries, A. de (1988). Werk en werkloosheid onder klinisch psychologen. [Work and unemployment among clinical psychologists]. *De Psycholoog, 23,* 399–406.

Scholing, A., & Brenninkmeijer, J. (1995) *Werk en werkloosheid onder afgestudeerden in de richting klinische psychologie van de RUG*. [Work and unemployment among clinical psychologists graduated at the University of Groningen] Unpublished manuscript. University of Groningen.

Trijsburg, R. W., Duivenvoorden, H. J., & Dijkstra, S. (1988). Psychotherapie als beroep. [Psychotherapy as a profession]. *Nieuws voor Psychotherapeuten,* 1–16.

van Son, M. J. M., & van der Staak, C. P. F. (1993). Inleiding. [Introduction] In W. T. A. M. Everaerd et al. (Eds.), *Handboek Klinische Psychologie* (pp. 1–28). Bohn, Stafleu, van Loghum, Houten.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.36

Demographic Characteristics of Psychologists and Supply/Demand Issues: United States

JESSICA L. KOHOUT
American Psychological Association, Washington, DC, USA

2.36.1 INTRODUCTION 431
2.36.2 HOW MANY ARE THERE? 432
2.36.3 WHAT DO THEY LOOK LIKE? 433
2.36.4 WHERE ARE THEY FOUND? 434
2.36.5 ISSUES OF DEMAND: ARE THERE TOO MANY? 434
2.36.6 FUTURE DIRECTIONS 437
2.36.6.1 APA Recommendations 438
2.36.7 SUMMARY 438
2.36.8 REFERENCES 439

2.36.1 INTRODUCTION

Clinical psychology has flourished over the past two decades. The number of psychologists in the health service provider subfields has tripled during this time, from approximately 20 000 in 1975 to almost 70 000 by 1995 or about 2500 new doctorates in the health service provider subfields each year (National Science Foundation [NSF], 1996). Equally striking are changes in the "face" of professional psychology. In the 1970s, a newly minted doctorate in clinical psychology would have been white, male, and in his late twenties. More often than not he would have earned a Ph.D. and would be headed, after licensure, toward academic employment or into independent practice. In contrast, today the new doctorate is likely to be a woman and white, although the representation of persons of color has inched up over time. Graduates today are somewhat older than their colleagues two decades ago, with the median age of new Ph.D.s in 1995 at just over 33 years (Henderson, Clarke, & Reynolds, 1996). Increasingly, new graduates may have earned a Psy.D. degree. The new doctorate is less likely to find as many opportunities for academic employment and independent practice as was the case in the 1970s. In fact, clinical psychology may be faced with some tough decisions as it heads toward the twenty-first century, finding work in a system that is searching for the means to confront worsening social problems in light of shrinking financial resources and shifting national priorities.

This chapter begins by discussing the current supply of clinical psychologists and attempts to answer the question "How many are there?" For the most part, this chapter focuses on doctoral-level psychologists, and in some cases, due to limitations of available data, is forced to narrow its focus further on Ph.D. psychologists. However, no discussion of supply or demand can occur without some reference to the supply of master's-level psychological personnel or to the supply of professionals in mental health fields outside psychology.

The next section addresses the demographic characteristics of clinical psychologists, including the gender, racial/ethnic distributions, age, and types of degrees. One of the largest shifts, that of the changing gender composition or the gender "tip" and the implications for the field is given particular attention.

The third section of this chapter presents information about the employment and geographic distributions of clinical psychologists in the USA. What are clinical psychologists doing, in what types of settings and where?

Fourth, the issue of supply and demand is raised. The question of balance is not reducible simply to provider–client ratios but must consider the many recommendations that have urged psychologists to "shift and expand their focus" (Fox, 1994); look at "systems of delivery" (Kiesler & Morton, 1988; Schneider, 1990); explore a "universe of alternatives" (Sarason, 1981); and conduct a "cognitive reappraisal" (Humphreys, 1996). Discussion is worthwhile and, ultimately, encouraging.

The final section incorporates information on the social, economic, and legal milieus in which psychology is embedded and to which the field must respond, and suggests possible future directions for this major subfield of psychology as it moves into its second century. The chapter concludes with a short summary.

2.36.2 HOW MANY ARE THERE?

In order to answer this question as accurately as possible, there must be some agreement about what is meant by "clinical psychologist." As stated in Section 2.36.1, the data are limited to those with doctoral degrees. However, there still is ample room for variability. The answer varies depending upon the purpose driving this question. For example, one may choose to include those with degrees in all health service provider subfields, those with degrees in clinical only, those who are licensed or those with certain degree and licensure combinations. Training in a health service provider subfield does not ensure that the individual is working as a practitioner or provider of health services. Licensure alone does not guarantee preparedness to deliver services and in fact can be used by those with degrees in other psychological subfields such as applied experimental, social psychology or industrial/organizational psychology as an indication of additional "certification" and qualification. This has proven useful to psychologists who are employed in applied or business settings. It is the case that psychologists may be licensed in more than one state and thus may end up being counted more than once in any census activity.

The first count consists of the pool of clinically trained psychological personnel (all those eligible to provide services at the doctoral level). This pool includes:

(i) residents of the USA,

(ii) those with doctoral degrees in clinical psychology as well as other health service provider subfields,

(iii) licensed or not licensed, and

(iv) may or may not currently be providing direct services.

The second count comprises those who are actively providing direct services and consists of US residents with a doctoral degree in a health service provider subfield in psychology, licensed for the independent practice of psychology, and working in psychology as a health service provider. It is important to remember that the majority of doctoral degrees in the health service provider subfields go to clinical psychology and thus the numbers presented here should be highly representative of this specific subfield. Where possible, the numbers are presented separately for the subfield of clinical psychology.

The data for this chapter were derived by using the American Psychological Association (APA) membership numbers and state-by-state data on the number of licensed psychologists. The analyses originally were performed in preparation of a chapter that appears in a National Institute of Mental Health (NIMH) publication called *Mental health United States, 1996* (Peterson et al., 1996). In 1995, there were some 69 817 clinically trained psychologists in the USA. Approximately 76% (53 061) of this number was clinically active. Some 70% of new psychology Ph.D.s in health service provider subfields are granted in clinical psychology specifically (Henderson et al., 1996), which yields an estimate of almost 49 000 trained clinical psychologists and just over 37 000 active clinical psychologists in the USA. These estimates are likely to be conservative since they are based on survey returns rather than census counts and are likely to miss some of the Psy.D. graduates who come from programs outside the accreditation process.

Turning briefly to the master's level, there was a 600% increase in the number of master's-level psychological personnel between 1960 and 1991 (US Department of Education [USDOE], 1994, 1996). The first half of the 1990s witnessed at least 10 000 new master's graduates each year in psychology alone. There were at least 900 terminal master's programs in psychology in 1991 and other data indicate that the number of institutions awarding master's degrees has increased since the mid-1980s (Snyder & Hoffman, 1995; USDOE, 1994). Estimates placed the number of master's-level psychological personnel at 102 945 in 1994 an additional 20 000 by 1996). Data from the National Science Foundation indicate that about 32% (39 342) of the master's graduates were employed in psychology (NSF, 1988). Adding this number to a figure reported in the APA's 1983 Census of psychological personnel (Stapp, Tucker, & Vanden Bos, 1985) yielded approximately 55 551 master's-level psychological personnel working in psychology. At least two-thirds and probably closer to three-fourths of the master's graduates were in the health service provider subfields.

National data reveal two other fields that produced substantially more master's degrees from 1992 to 1993 than psychology. These were counselor education/guidance services in education, and social work (USDOE, 1994), and it is easy to see how the graduates in these fields could be competitors for jobs held by master's-level psychological personnel. Due to changes occurring in the health care delivery system, it also is clear that they could represent a substantial challenge to the full employment of doctoral-level psychologists.

The large numbers of master's degrees granted yearly in psychology, social work, and in counselor education in education; organized psychology's singular focus on the doctoral-level degree as the entry-level degree for the independent practice of psychology; efforts of master's-level providers to organize politically to influence state licensing laws, coupled with shifts in the health care delivery systems and an increasing emphasis on cost containment have pushed the question about the roles and functions of the master's-level psychologist in the provision of health services to the forefront. Increasingly, they are perceived as being in direct competition with doctoral-level providers. This has been called the "master's issue."

2.36.3 WHAT DO THEY LOOK LIKE?

The most obvious shift is in the gender composition of psychology (Pion et al., 1996). Enrolled students and degree recipients at all levels are predominantly female. For example, in 1995, women accounted for 67% of all full-time doctoral students and 74% of full-time master's students in psychology (Williams, 1996). From 1993/94, the most recent year for which data were available, women earned 73% of the bachelor's degrees and 72% of the master's degrees awarded in psychology (USDOE, 1996). Of all new psychology Ph.D.s in 1995, 64% were women, almost doubling the 32% of 20 years earlier (Henderson et al., 1996). The health service provider subfields in psychology have not been immune to this gender tip and it is not surprising that women comprised 44% of the clinically trained pool in 1995, up 5% from 1989 (Dial et al., 1990).

What are the implications of this momentous shift in gender composition, for the field and for services? A thorough discussion of this shift can be found in the Report of the Task Force on the Changing Gender Composition of Psychology available from the Women's Program Office of the APA. Case study findings also were presented in a 1996 *American Psychologist* article (Pion et al., 1996). A summary of the implications addressed in the case study includes the following. The pipeline data indicate that the face of the psychological workforce will continue to change. Rather than imply that these shifts are responsible for any decreases in earnings, or decline in attractiveness or autonomy, the report stresses the fact that the changes that have and are occurring in psychology are the result of external economic and legal forces and not the influx of women into the field.

The representation of persons of color in psychology remained small at about 6% of clinically trained psychologists (Peterson et al., 1996). The population of women was slightly more diverse than was the case for men. The presence of racial and ethnic minorities in psychology was smaller than their representation in the US adult population. Racial and ethnic minorities earned 14% of the Ph.D.s in psychology and 12% of those awarded in clinical psychology in 1995 (Henderson et al., 1996), but they represented 23% of full-time doctoral students in 1995; thus the pipeline is showing some diversification.

The median age of clinically trained psychologists has increased. In 1995 this number was just under 49 years, up somewhat from 1989 when it was just over 44 years. Median years since the doctorate also increased from 12 to 15 years between 1989 and 1995. Generally, women were younger than men and had earned their degrees more recently than men (Peterson et al., 1996).

2.36.4 WHERE ARE THEY FOUND?

The question is interpreted broadly and data are presented for the employment settings and activities as well as for the geographic locations of professional psychologists. In 1995 most of the clinically active psychologists were working full time and at least 50% held two or more positions. Part-time workers were more apt to be working at one position only. Again, most of these data are drawn from the publication by Peterson et al. (1996).

The data indicate that the largest single proportion of clinically active psychologists claimed independent practice as a primary employment setting (52.5%), mostly in individual (41.2%) as opposed to group (11.3%) arrangements. Academic settings were second at 14% and most of the clinical psychologists in these settings were located in higher education (e.g., universities and colleges (11%)), with the remaining 3% in secondary schools and other educational settings. Hospitals were claimed by just over 12%, 6.5% were working in clinics, while nursing homes employed less than 1%. Just over 14% of the practicing psychologists chose a generic "other human service setting" or "other" (see Figure 1).

For the 46% of active professional psychologists specifying a secondary employment setting, independent practice arrangements again were most frequent. Individual private practice was chosen most often (37.4%), while 9.4% specified group practices. Almost 18% were employed in universities and colleges. Hospitals were chosen by 8.9%, clinics by 5.1%, and nursing homes by less than 1%. Schools and other educational settings employed less than 3% and other human service or other settings claimed just over 18%.

A major point of interest in these data is the fact that despite the changes in health care delivery systems, the largest single proportion of practicing psychologists remains in independent practice These findings are supported by data from a 1995 Committee for the Advancement of Professional Practice (CAPP) survey of practicing members of the APA (APA Practice Directorate, 1996) .

Turning from setting to primary activity, just over three fourths of the clinically trained psychologists were involved in patient care and direct service, 9% were teaching, 8% were administrators and 4% were involved in research. Just under 3% indicated some other activity that was not specified (see Figure 2).

The question was asked about extent of involvement in each activity as opposed to primary or most important activity. Again, three-quarters of the clinically trained population was involved in patient care and direct service. This was not a surprise, given the overwhelming presence of professional psychologists in independent practice settings. However, it was interesting to discover that almost two thirds also taught, slightly more than a quarter was involved in administration, and just under one fifth conducted research. The other category climbed to 30% which is sizeable and warrants additional analysis to uncover these activities (see Figure 3).

Taking the setting and activity data together, it appears that clinical psychologists still do consider themselves to be independent practitioners, and that they continue to remain involved in a variety of activities, including teaching, administration and research. The other category is of great interest.

Geographically, where are psychologists located? Regionally, the East and West coasts of the USA held the single largest proportions of clinically trained psychologists, that is, 21% were in the Middle Atlantic region and 17% in the Pacific region. The center of the USA and the South and Mountain regions have markedly lower representation. At the state level, although insufficient to boost the Pacific region on its own to the most populous, California had the highest number of clinically trained psychologists, followed by New York, Pennsylvania, Massachusetts, and Ohio. The same patterns apply to the clinically active subset. A general rule of thumb with respect to understanding the geographic distribution of psychologists in the US is to look to the large metropolitan areas and the areas with numerous graduate programs in professional psychology. The large populations have supported the training programs while the training programs and later, their graduates, have used these concentrations of people for the internships and practices.

2.36.5 ISSUES OF DEMAND: ARE THERE TOO MANY?

When assessing the issue of whether there are too many doctoral psychologists, it depends on what you define as the appropriate scope of practice. It is becoming evident that there likely are a sufficient number of psychologists trained to provide mental health services to individual clients in independent practice settings. At the same time, it is equally evident that wrenching shortages continue to plague other parts of the systems of service delivery. Examples of shortages include rural and frontier areas to which traditionally clinically trained psychologists cannot be recruited and to which they are

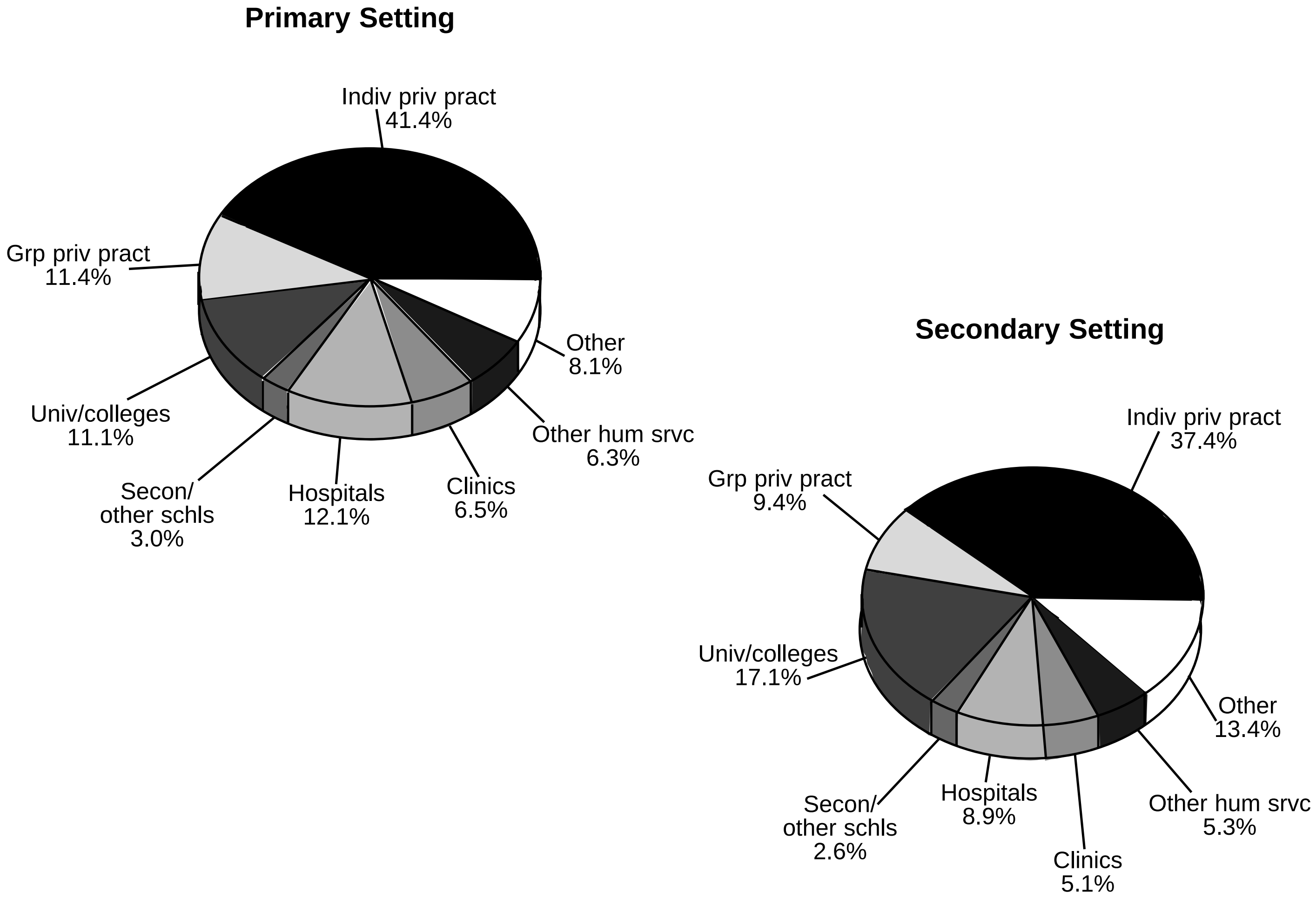

Figure 1 Primary and secondary employment settings of clinically active psychologists: 1995. Note: Percentages may differ slightly from the text due to rounding by the graphics package. Source: Peterson et al. (1996).

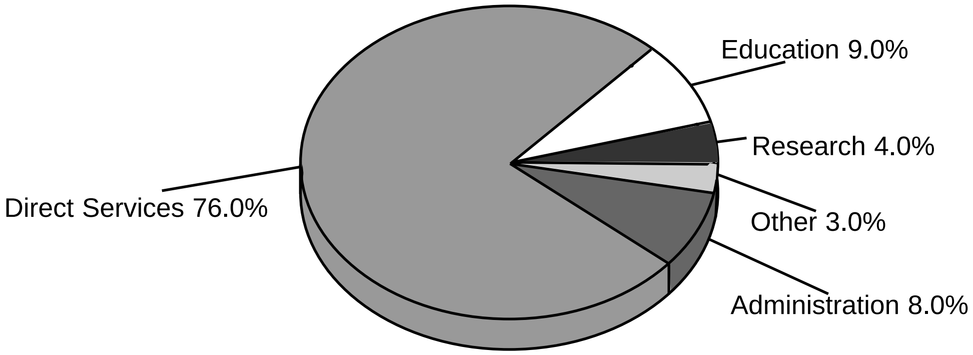

Figure 2 Primary activities of clinically trained psychologists: 1995. Source: Peterson et al. (1996).

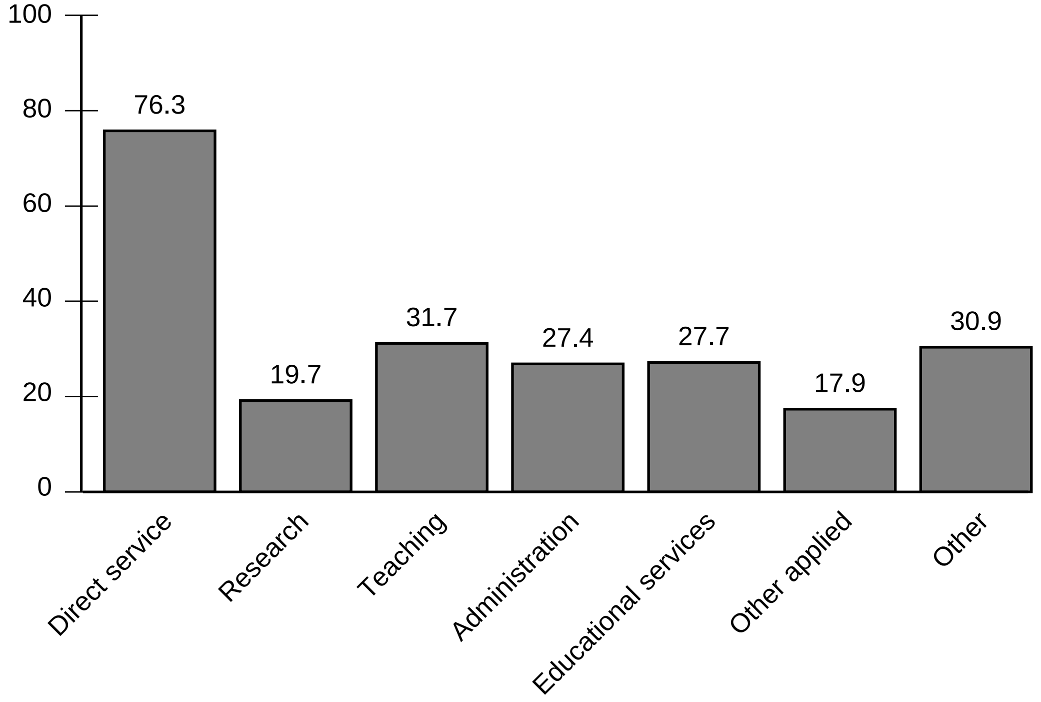

Figure 3 Any involvement in activities, clinically trained psychologists: 1995. Note: Percentages will total more than 100% as respondents could provide more than one response. Source: Peterson et al. (1996).

not attracted. Areas within urban settings that may appear dangerous or not able to support a solo practice also may be underserved. Other shortages can be defined with respect to the clients' needs (e.g., seriously mentally ill adults, seriously emotionally disturbed children, homeless persons, immigrant and minority populations, and the elderly) and organizational settings (clinics, hospitals, state agencies). A common characteristic of these groups of clients, settings, and areas is that they may not fit within the parameters/paradigm of a typical independent practice, and consequently will lack access to services.

Although a perceived oversupply of clinically-trained psychologists and threats posed by the health care revolution to the free practice of psychology may appear to be recent developments, this is not the case. In fact, a number of voices have been raised in common chorus for some time, urging psychology to expand its perspectives on what constitutes an appropriate scope of practice. Sarason (1975, 1981, 1983) has been urging an exploration of the "universe of alternatives" since the mid-1970s. Levy (1984) noted that training and practice in psychology could be more productive if they adopted a human services perspective rather than remaining limited to a focus on the treatment of individual mental health problems. Kovacs (1989) urged that we prepare psychologists to deal with challenges people face across the life span in a variety of settings. In the early 1980s and up to the present, Fox has written and talked about shifting and expanding the focus in professional psychology from an illness driven to a prevention driven model. Kiesler and Morton (1988), VandenBos, Belar, and DeLeon (1991), and Humphreys (1996) have joined the chorus, arguing for a shift to a larger proactive focus on designing systems/interventions/ways to promote healthy behavior and well-being at all levels. Yet, noted Schneider (1990), "it seems that the major practice preoccupations are geared toward a system that is disappearing rather than to the one replacing it."

For the most part the education and training structures in psychology have not changed, and students are not receiving the training they will need to personally succeed in, and effectively contribute to, the evolving health care environment. Despite the many calls for an expansion of the scope of practice and the obvious shifts in health care delivery systems, pessimism among established psychologists and a degree of panic among trainees and new graduates is already evident. At all levels of government and in many business settings, there is a general lack of understanding of what it is that psychologists do and can offer. The depth of understanding appears little better at the level of the individual citizen, where data from the evaluation of the Psychology Traveling Exhibit indicate fairly simplistic and uninformed views of what it is psychologists do. Positively, outcome interviews indicate that people are learning about the variety of research and the practical applications of psychology as a consequence of having attended the exhibit (APA Science Directorate, 1996).

If the educational system does not respond by providing training appropriate to the new health care systems and if clinically trained psychologists continue to establish traditional independent practices in which the client must come to the provider for individual or group treatment, a situation in which the psychologist actually is very passive, it is likely that the judgement will be that there are too many clinical psychologists and that they are irrelevant to the evolving systems and real world problems. This judgement will be held by psychologists, since they will have difficulty making a living, as well as by those responsible for establishing the new systems of health care and by those who are passing the laws and handing out funding for education, training, research, and outreach. The perception of oversupply can be as deadly as actual oversupply, dampening the interest of potential students in psychology as a career and lessening the support for the education and training of psychologists at all levels.

There is no oversupply. In fact, the persistent and worsening nature of many of society's ills and the fact that many of these have behavioral bases lend support to the argument that there are too few clinically trained psychologists. But psychology does appear to have hit a setback in that there is a mismatch between the training that students are receiving and the needs of the society. Suggestions about how to meet these needs is the subject of the next section.

2.36.6 FUTURE DIRECTIONS

Psychology should abandon efforts to model itself after medicine and move toward a proactive and self directing approach in order to define an appropriate scope of practice. This will require the combined efforts of all players within professional psychology. Examples include national, regional and state associations, the education and training structures within psychology, psychologists and students of psychology, and others.

> Professional psychology is that profession which is concerned with enhancing the effectiveness of human behavior. Therefore, a professional psychologist is one who has extensive expertise in the development and application of quality services to the public in a controlled, organized, ethical manner; based upon psychological knowledge, attitudes and skills, in order to enhance the effectiveness of human functioning. (Fox, 1982)

In the future, professional psychology should be prevention driven (Fox, 1994; Sarason, 1975, 1981, 1983) and it should help people to deal with the challenges of daily living across a variety of settings (Kovacs, 1989) and roles (e.g., family, work, parent, etc.). Emphasis is increasing on "designing and implementing interventions that help people pursue healthy behaviors and end or reduce risky behavior" (Humphreys, 1996). Psychology is moving away from a focus on individual-level problems toward the community, social problems, and human services. It will be part of a multidisciplinary effort in many instances in which the psychologist bands together with professionals from other fields to study social problems. The interventions that are designed as a consequence of this elaboration will be stronger by virtue of their multidisciplinary origins.

Defining professional psychology in this way immediately opens up many new avenues for productive and satisfying work. Some of the areas to which psychologists can contribute include human resources management, environmental design, training and supervision, and research on consumer behavior (VandenBos et al., 1991). VandenBos et al. also noted the variety of health related areas in which professional psychologists have made contributions, including: pain management, coping with stressful medical procedures, controlling drug side effects, physical self examination, health monitoring, recovery from illness or surgery, behavioral health, sticking to a health care regimen, and more.

How do we get to this stage? The recommendations that appear below can also be found in the Final Report and Recommendations of the APA Working Group on the Impact of Managed Care and Changes in the Health Services Delivery System on the Education and Training and Continuing Education of Professional Psychologists (American Psychological Association, Education Directorate and Research Office 1996). A copy of this report can be obtained from the APA's Research Office or from the Center For Mental Health Services, Washington, DC.

2.36.6.1 APA Recommendations

The educational system will have to change. Educational institutions will need to target clinical and research expertise to society's real needs (VandenBos et al., 1991). There needs to be greater emphasis on life-long learning and the content and the nature of the training will have to change. Some have called for a common core of general knowledge with specialization occurring postdoctorally. Practical experience or internships will need to be developed outside mental health settings to give trainees some experience with professionals from other fields in a variety of settings. Providing experiences with a diversity of populations and clients will be essential.

Technological literacy is critical to managing one's practice and evaluating and demonstrating effectiveness and efficiency. This involves computer literacy, ability to use information and delivery system technologies and research skills.

The ability to work with and communicate with professionals from other disciplines via interdisciplinary teamwork in multidisciplinary settings will be central to success in the evolving systems. Professional ethnocentrism is to be avoided.

Ultimately, the professional psychologist should be innovative, flexible and able to adapt to nontraditional settings, clients and modes of delivery by offering interventions that have been shown to work.

Practitioners will have to be able to demonstrate the cost effectiveness of psychological services and be able to predict and control costs. They will have to be able to fill public education and advocacy roles and they will need an understanding of the structure of evolving health care systems, including financing, organization, policy, and delivery principles. The ability to market one's services is necessary.

These recommendations and observations from the field indicate that health psychology will have a large role in the future (Cummings, 1995), by providing psychologists with the skills necessary to plan, research, and implement intervention programs for somaticizers and noncompliant chronically physically ill as well as programs for those with unhealthy behaviors (smoking, overeating, etc.) (e.g., Cummings, Dorken, Pallak, & Henke, 1993; Pallak, Cummings, Dorken, & Henke, 1993). Empirical outcomes research is also critical in that it will comprise the basis on which psychological interventions/therapies will rest, most of which are likely to be small group and psychoeducational in nature (Cummings, 1995).

2.36.7 SUMMARY

Clinical psychology has changed noticeably since the 1970s. Post-World War II shifts to a postindustrial society, in which the economy was based largely on the provision of services and an increased dependence on scientific and technical knowledge, brought with it an explosion of opportunities for professional psychology (Howard et al., 1986). The numbers of graduates have grown, tripling between 1975 and 1995 from 20 000 to almost 70 000. Recognizing opportunity, women flocked to psychology. Schools of professional psychology have been established that award a doctor of psychology (Psy.D.) rather than the Ph.D.

Primary employment shifted from organizational or institutional settings to private practice, nudged along by legal changes, increasing acceptance of psychological services and favorable reimbursement policies, and the psychologist was located most often in urban areas, near concentrations of population necessary to support a practice.

Questions about imbalances between supply and demand can be answered a number of ways. If one maintains a fairly limited view of what constitutes practice), then there appears to be an overabundance of practitioners and the supply is greater than the demand. This has negative consequences for the full employment of clinical psychologists. Indeed, data indicate that shifts in the health care delivery system in the USA have had negative impacts on health service providers in independent practice who must try to develop new niches.

Additionally, it is important that organized psychology takes some time to address the master's issue in a reasoned and grounded manner. Perhaps gathering data on the actual content of training at both levels would be a start. Revisiting the frequently mentioned rift between the science and practice of psychology is a needed exercise.

As many have argued, professional psychology must explore alternatives and expand its understanding of what constitutes an appropriate scope of practice. This shift from an "illness" model to a prevention model from a largely passive and reactive stance to one that is positive, proactive and in control—offers many new opportunities for professional psychology. Supply does not exceed demand and unemployment remains low. Above all, the benefits of such a shift extend beyond psychology, to society and individuals, by identifying healthy behaviors and the ways to practice those behaviors.

Although the goal of adopting a prevention approach has been discussed many times and is, from all appearances, a welcome outcome, a critical stumbling block remains. That is that the majority of the education and training systems in the USA have not begun to revise their training to meet the needs of a changing system. As a consequence, students in the health service provider subfields are not learning the knowledge and skills necessary to function effectively in the evolving health care system. The next challenge to reaching the goal of a prevention or human services model is to work with the education and training systems in psychology to design curriculum to meet this need.

2.36.8 REFERENCES

American Psychological Association, Education Directorate and Research Office (1996). *Final report and recommendations of the American Psychological Association working group on the impact of managed care and changes in the health services delivery system on the education and training and continuing education of professional psychologists*. Washington, DC: Author.

American Psychological Association, Practice Directorate (1996). *The Committee for the Advancement of Professional Practice (CAPP) practitioner survey results*. Washington, DC: Author.

American Psychological Association, Science Directorate (1996). Preliminary results of the evaluation of the museum exhibit, Psychology: Understanding ourselves, understanding each other. Unpublished analyses. Washington, DC: Author.

Cummings, N. A. (1995). Impact of managed care on employment and training: A primer for survival. *Professional Psychology: Research and Practice, 26,* 10–15.

Cummings, N. A., Dorken, H., Pallak, M. S., & Henke, C. J. (1993). The impact of psychological intervention on health care costs and utilization: The Hawaii Medicaid project. In *Medicaid: Managed behavioral health and implications for public policy: Volume 2. Healthcare and utilization cost series* (pp. 3–23). San Francisco: Foundation for Behavioral Health.

Dial, T. H., Tebbutt, R., Pion, G. M., Kohout, J., VandenBos, G., Johnson, M., Schervish, P. H., Whiting, L., Fox, J. C., & Merwin, E. I. (1990). Human resources in mental health. In R. W. Manderscheid, & M. A. Sonnenschein (Eds.), *Mental health, United States, 1990.* Center for Mental Health Services and National Institute of Mental Health, DHHS # (ADM)90–1708. Washington, DC: US Government Printing Office.

Fox, R. E. (1982). The need for a reorientation of clinical psychology. *American Psychologist, 37,* 1051–1057.

Fox, R. E. (1994). Training professional psychologists for the twenty-first century. *American Psychologist, 49,* 200–206.

Henderson, P. H., Clarke, J. E., & Reynolds, M. A. (1996). *Summary report 1995: Doctorate recipients from United States universities*. Washington, DC: National Academy Press.

Howard, A., Pion, G. M., Gottfredson, G. D., Flattau, P. E., Oskamp, S., Pfafflin, S. M., Bray, D. W., & Burstein, A. G. (1986). The changing face of American psychology: A report from the Committee on Employment and Human Resources. *American Psychologist, 41,* 1311–1327.

Humphreys, K. (1996). Clinical psychologists as psychotherapists: History, future and alternatives. *American Psychologist, 51,* 190–197.

Kiesler, C. A., & Morton, T. L. (1988). Psychology and public policy in the "health care revolution." *American Psychologist, 43,* 993–1003.

Kohout, J. L. (1993, August). *Employment opportunities in psychology: Current and future trends*. Paper presented at the annual meeting of the American Psychological Association, Toronto, Canada.

Kovacs, A. L. (1989, February). *The uncertain future of professional psychology*. Paper presented at the annual meeting of the Council of Graduate Departments. Tucson, Arizona.

Levy, L. H. (1984). The metamorphosis of clinical psychology: Toward a new charter as human services psychology. *American Psychologist, 39,* 486–494.

National Science Foundation (1988). *Profiles—psychology: Human resources and funding*. (NSF 88–325). Washington, DC: US Government Printing Office.

National Science Foundation (1996). *Characteristics of doctoral scientists and engineers in the United States: 1993*. (NSF 96–302). Arlington, VA: National Science Foundation.

Pallak, M. S, Cummings, N. A., Dorken, H., & Henke, C. J. (1993). Managed mental health, Medicaid, and medical cost offset. *New Directions for Mental Health Services, 59,* 27–40.

Peterson, B. D., West, J., Pincus, H. A., Kohout, J., Pion,

G. M., Wicherski, M. M., Vandivort-Warren, R. E., Palmiter, M. L., Merwin, E., Fox, J. C., Clawson, T. W., Rhodes, K. K., Stockton, R., Ambrose, J. P., Blankertz, L., Dwyer, K. P., Stanhope, V., Fleischer, M. S., Goldsmith, H. F., Witkin, M. J., Atay, J. E., & Manderscheid, R. W. (1996). An update on human resources in mental health. In R. W. Manderscheid & M. A. Sonnenschein (Eds.), *Mental health, United States, 1996*. Center for Mental health Services, DHHS# (SMA) 96-3098. Washington, DC: US Government Printing Office.

Pion, G. M., Mednick, M. T., Astin, H. S., Ijima Hall, C. C., Kenkel, M. B., Keita, G. P., Kohout, J. L., & Kelleher, J. C. (1996). The shifting gender composition of psychology: Trends and implications for the discipline. *American Psychologist, 51,* 509–528.

Sarason, S. B. (1975). Psychology of the Finland Station in the Heavenly City of the 18th century philosophers. *American Psychologist, 30,* 1072–1080.

Sarason, S. B. (1981). An asocial psychology and a misdirected clinical psychology. *American Psychologist, 36,* 827–837.

Sarason, S. B. (1983). Psychology and public policy: Missed opportunity. In R. D. Feiner, L. A. Jason, J. N. Moritsugu, & S. S. Ferber (Eds.), *Preventive psychology: Theory, research and practice* (pp. 245–250). New York: Pergamon.

Schneider, S. F. (1990). Psychology at a crossroads. *American Psychologist, 45,* 521–529.

Stapp, J., Tucker, A. M., & VandenBos, G. R. (1985). Census of psychological personnel. *American Psychologist, 40,* 1317–1351.

Snyder, T. D., & Hoffman, C. M. (1995). *Digest of education statistics.* (NCES 95–029). Washington, DC: US Department of Education.

US Department of Education, National Center for Education Statistics (1994). Earned degrees and completion surveys. Unpublished tabulations.

US Department of Education (1996). Degrees conferred by discipline, sex, and level of study. Unpublished raw data.

VandenBos, G. R., Belar, C. D., & DeLeon, P. H. (1991). How many psychological practitioners are needed? It's too early to know! *Professional Psychology: Research and Practice, 22,* 441–448.

Williams, S. (1996). Race/ethnicity and gender data for graduate students and faculty from the 1995–96 Graduate Study in Psychology survey. Unpublished analyses. Washington, DC: American Psychological Association.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.37

Psychologists' Work and the Public's Perceptions: A Dichotomy

RAYMOND FOWLER and RHEA K. FARBERMAN
American Psychological Association, Washington, DC, USA

2.37.1 THE PUBLIC'S PERCEPTIONS 441
    2.37.1.1 The Impression 441
    2.37.1.2 The Reality 441
2.37.2 THE HIDDEN CONTRIBUTIONS OF PSYCHOLOGY 442
2.37.3 THE WORK OF PSYCHOLOGISTS IN NUMEROUS AND VARIED SETTINGS 443
    2.37.3.1 Psychologists in Healthcare Settings 443
    2.37.3.2 Public Health Programs 444
    2.37.3.3 The Workplace 445
    2.37.3.4 Forensic Psychologists 445
    2.37.3.5 Human Factors Research 445
    2.37.3.6 Psychologists in Schools 446
    2.37.3.7 Military Psychologists 446
    2.37.3.8 Psychologists Study the Media 447
    2.37.3.9 Sport Psychology 447
2.37.4 FUTURE DIRECTIONS 447
2.37.5 REFERENCES 448

2.37.1 THE PUBLIC'S PERCEPTIONS

2.37.1.1 The Impression

Psychology is a broad discipline, but public understanding of its breadth and what it contributes to society is narrow. Psychologists work in so many settings that it is difficult to identify all of them, but the impression of psychologists many people have is limited to them as professors, researchers, or in the traditional counseling and clinical roles.

A national telephone survey conducted in 1995 by the American Psychological Association (APA) (Penn and Schoen Associates, 1995) found that 45% of all respondents believed that psychologists "generally help people with problems or offer counseling." Others (30%) said that psychologists "analyze people or study human behavior" (see Figure 1).

While it is true that many psychologists engage in clinical and counseling work, the world of psychology is much broader than the public often realizes because the work of psychologists in other settings occurs under other job titles and identities.

2.37.1.2 The Reality

Since the domain of psychology is the study of human and animal behavior, there are very few areas of inquiry which are not of interest to psychologists. In fact, psychological skills and

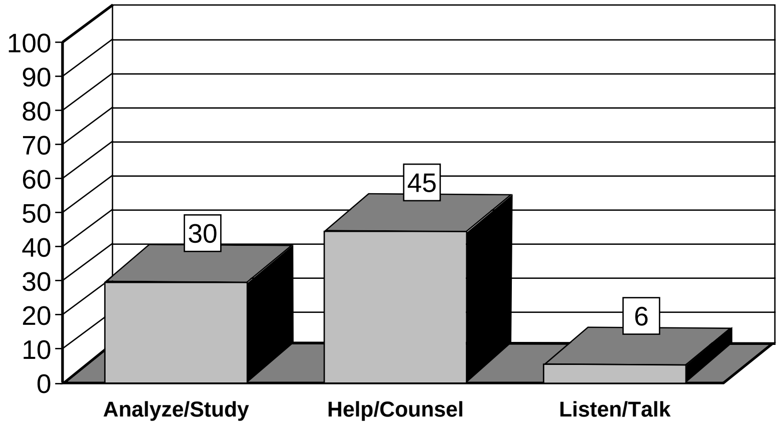

Figure 1 What Do Psychologists Do?
Source: American Psychological Association.

knowledge are relevant to a broad segment of social problems and public policy issues, for example, aging, AIDS, violence, depression, and drug abuse.

Today approximately 80 000 doctoral level psychologists (about half of the psychologists in the world) live and work in the USA. They work in a number of major subfields including clinical psychology, counseling psychology, developmental psychology, educational and school psychology, forensic psychology, health psychology, industrial/organizational and engineering psychology, neuropsychology, quantitative and measurement psychology, rehabilitation psychology, and social psychology. The first two subfields, clinical and counseling psychology, are somewhat familiar to the general public (although studies show that the public has very little understanding of how to differentiate among mental health providers). The remaining subfields, research shows, are unfamiliar to most members of the lay public.

Exact reports of how many professionals work in each of the subfields are difficult to ascertain because many in the field work in more than one role (see Figures 2 and 3). It is estimated however that approximately 20% of the discipline are in academic positions teaching and conducting research. Approximately 50–60% are thought to be primarily engaged in health service delivery (APA, 1996a).

2.37.2 THE HIDDEN CONTRIBUTIONS OF PSYCHOLOGY

There are numerous places where psychologists and psychological data collection and analysis are making contributions behind the scenes (e.g., social welfare, health and medicine, education, and corporate competitiveness) but are unknown to the general public. Some examples are:

(i) *Computer technology*. The computer and electronic communications revolution has benefited from, and continues to need, the input of psychologists. Computer learning and communications require that people learn to use the systems and function comfortably at them. Psychologists help in the design of teaching tools as well as operating systems and work stations that match human minds and bodies. Social psychologists also study the effect that communications via satellites and circuitry rather than face-to-face have on society and the way human beings interact.

(ii) *The new American demographics*. Towards the twenty-first century the world is a smaller place. America is more culturally, racially and ethnically diverse then it has ever been before. Psychology has a role in understanding the needs and ramifications of a more diverse population and helping institutions from governments to business to schools adjust to this new reality.

(iii) *Corporate downsizing*. Industrial/organizational psychologists and engineering psychologists teach corporations and business leaders to do more with less and deal with international economic challenges by finding work efficiencies, streamlining production processes, and promoting good management of human resources.

(iv) *Education reform*. Psychologists study the teaching and learning process and work in schools to make them more effective.

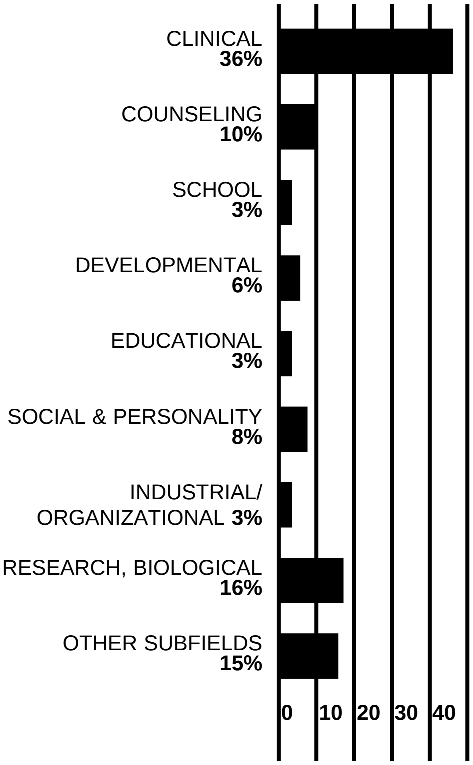

Figure 2 Subfields doctoral psychologists enter. Source: Adapted from 1993 *Survey of Doctorate Recipients*, National Research Council and National Science Foundation, Washington, DC.

2.37.3 THE WORK OF PSYCHOLOGISTS IN NUMEROUS AND VARIED SETTINGS

2.37.3.1 Psychologists in Healthcare Settings

It is estimated that between 50 000 and 60 000 psychologists work in a wide range of healthcare settings such as hospitals, clinics, in individual private practice, and as members of teams of primary care providers. Health psychologists study how biological, psychological, and social factors affect health and illness. In the last two decades clinical psychology has made large strides in evolving from a mental health profession to a health profession that teams with other healthcare providers—from primary care physicians to disease specialists—to provide comprehensive healthcare for patients. While most health disciplines specialize in the physical aspects of illness, psychologists are expert in the behavioral aspects of illness and the coping and adjustment skills needed to recover from illness. The connection between the mind and the body is well-established. Health psychologists understand how emotional illness can have physical ramifications and apply that knowledge to help people get well and stay well.

Psychologists begin to make important contributions to a patient's health at the initial stages of care giving. Psychologists evaluate patients by assessing personality traits and cognitive and motor abilities (memory or reaction time, for example). They work with other specialists to design and implement treatment plans to improve those skills.

Psychologists also help patients deal with the stress or discomfort which can be caused by the medical care, and play a critical role in keeping patients to their treatment regimes and recovery. Cancer patients, for example, often consult psychologists for help with medical decision making, treatment compliance, and the stress and depression that can accompany a cancer diagnosis.

Health psychologists also have an important role to play in preventing illness. They investigate health problems that affect large segments of the population such as smoking, sexually transmitted diseases, and poor diet. They can then design interventions to deal with these problems.

Research has shown that approximately half of all visits to primary care doctors are for symptoms for which psychological rather than physical causes are primary. The health psychologist in the USA therefore has a critical role to play in both the well-being of the American public and in reducing national healthcare expenditures.

Some psychologists have extensive training in psychopharmacology and work closely with physicians in determining appropriate medication for patients. Many psychologists believe that expanding their clinical skills to include prescribing drugs will allow them to better treat their patients. In a 1995 poll of its membership, the APA found that 62.7% of its members agreed with the statement, "appropriately trained psychologists should prescribe medication." Of the membership, 33% disagreed with the statement. No states currently license psychologists to prescribe drugs, but it is being considered by a few state legislatures.

Those who argue for prescription privileges believe that psychologists who prescribe will be better able to provide complete care for patients and will be able to appropriately coordinate medications with the psychotherapy needs of the patient. Opponents of prescribing, however, worry that the need to keep up with new advances in psychopharmacology will rob psychologists of the time necessary to continue to develop their expertise in psychotherapy.

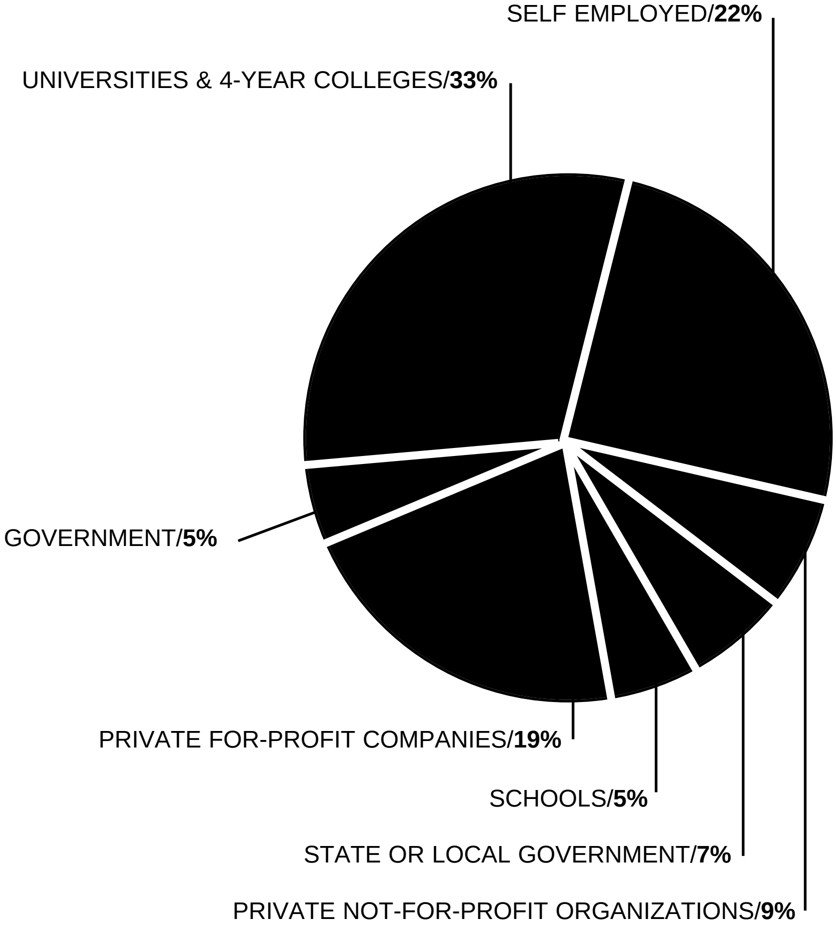

Figure 3 Where doctoral psychologists work.
Source: Data for psychologists working full time for others adapted from *Survey of Doctorate Recipients*, National Research Council and National Science Foundation, Washington, DC. Individual and group practice data are from Research office, APA Central Programs.

Up to 1997, four psychologists have graduated from a special military training program created to test the value of prescribing privileges for psychologists working with military personnel and their families and have been authorized to prescribe drugs. After some initial conflict with their military colleagues in the psychiatric ranks, all report good working relationships with other military doctors. Seven additional psychologists are in the training program in 1997.

2.37.3.2 Public Health Programs

Psychologists are tackling some of the late twentieth century's most troubling public health issues by both studying their root causes and designing interventions to solve them or help people cope with their ramifications.

Psychologists have been at the forefront of the battle to defeat AIDS since the epidemic was first identified. Psychologists are studying how to educate and motivate AIDS patients and those at high risk of contracting the virus to protect themselves against becoming infected. Needle exchange programs, for example, have been studied by many researchers. Psychologists also help patients and their families cope with the challenges of living with AIDS/HIV (Clay 1996; Sleek 1996).

Substances abuse is a problem that robs many people of their livelihoods, families, and even their lives. Psychologists, many specializing in the treatment of drug and alcohol abuse, treat patients with substance abuse. Research psychologists also study why some people become addicted to a substance while others do not, as well as what treatments programs have the greatest success rates.

Violence is a universal problem in the world. Violence takes places in urban centers, in rural communities, in the workplace, and at home. Psychologists have both studied the causes of violence and designed interventions to prevent it. In the workplace, psychologists theorize that violence often erupts as the results of unchecked stress and rage within an employee, but find it difficult to predict who may become violent in the work setting. Psychologists who study workplace violence have found that losing one's job is a better predictor of violence than age, gender, socioeconomic status, ethnicity, or diagnosable mental disorder. But, evidence does show that those who become violent often have a history of aggression, poor self-esteem, are socially isolated, and blame others for their problems (Martin, 1994).

Family violence is another area that psychologists have studied and work to prevent. Violence among family members is not restricted to any cultural or economic groups, but through research, psychologists have ascertained that family violence is circular—that it is passed from generation to generation within a family. Children who are abused and who witness adult members of the family hitting each other are more likely to be both the perpetrators and victims of violence as adults.

Psychologists intervene with families that are having trouble with domestic violence in numerous ways. Often helping people acquire better life skills, such as parenting skills and job skills, will help lessen the violence problem. A program in which psychologists make repeated home visits to families who are experiencing problems with violence has been found to be particularly effective in stemming that violence. The full involvement of many community resources in coping with the family violence problem has also proven to be very important (APA, 1996b).

2.37.3.3 The Workplace

Psychologists can be found in the workplace under a variety of job titles from human resources manager to quality performance director to wellness manager to consultant but most are there with the same basic mission; to make the workplace productive and safe.

Businesses today face numerous human resource challenges; getting the right employee in the right job, getting employees the training they need, making the workplace safe and keeping employees healthy, making the production process (whether the product is a widget or intellectual property) efficient, and keeping non work-related problems from interfering with employee performance.

Psychologists study, design, and implement intervention strategies to keep workers reporting to work healthy, ready to work, and to keep business moving ahead.

Clinical psychologists are also branching out into the fairly new subfield of executive coaching; working with corporate executives to strengthen their management, leadership, and interpersonal skills. Often both corporate leaders and line managers do not have a sufficient understanding of employee morale needs and how to build high functioning workplace teams and resolve conflict. Psychologists teach these skills where needed.

Workplace issues are likely to remain a strong area of work opportunities for psychologists. It is estimated that the direct and indirect costs of stress-related disorders to business in the USA now stands at $100 billion annually. Furthermore, about 600 000 workers a year are disabled by psychological disorders.

2.37.3.4 Forensic Psychologists

Forensic psychologists work within and for the justice system and attempt to bring psychological knowledge to bear in the search for justice and equitable outcomes in court cases. Psychologists who work in forensic psychology often interview defendants at the request of defense attorneys or the court to judge their competency and state of mind. Psychologists are typically involved in divorce and child custody cases usually offering an opinion about the developmental and emotional health of any children involved in the case.

Other psychologists study jury behavior and help attorneys select juries who will be most likely to side with their clients.

2.37.3.5 Human Factors Research

Designing machinery and manufacturing processes to fit comfortably with the labor force is the expertise of human factors and engineering psychologists.

Human factors psychologists also study human performance and match the limits of human performance to work demands. For example, psychologists working in the NASA Fatigue Countermeasures program are developing a means by which to measure pilot fatigue and its effect on performance. Concerns about pilot alertness especially during long transmeridian flights were first raised in the early 1980s. Today, psychologists working on or with the NASA team use basic behavioral research on sleep physiology and circadian rhythms to design instruments that measure alertnes and performance. They have also taken these instruments into the cockpit to test general theories and countermeasures where pilots actually work (Azur, 1996b).

This research has concluded that allowing pilots to rest during long flights can improve their alertness at the end of that flight. Findings from this research will have important cross over into many other fields that require around the clock schedules such as law enforcement, public safety, and medicine.

Another area in which psychologists are working behind the scenes to protect the public is in the design of medical devices. The US Food and Drug Administration (FDA) must approve all medical devices; FDA approval is based on product safety and effectiveness, although not whether the product is user-friendly. However, these devices are often used by staff in busy hospitals, who are working on severe time constraints, heavy workloads, and at times fatigued.

An added emphasis on good human factors engineering in the design of medical devices began in the early 1990s when the FDA started to ask why so many well-trained professionals were failing to use them properly. Psychologists who work in the area point out that as medical devices become more complex and as more and more healthcare takes place in the home or some other out-of-hospital settings making medical devices more user-friendly is more critical than ever (Azur, 1996a).

2.37.3.6 Psychologists in Schools

Schools mirror the communities in which they exist. Today they are asked to teach children in increasingly difficult circumstances. Psychologists, both by studying the teaching and learning process and by working in schools directly with administrators, teachers, and students, are helping schools fulfill their mission.

School psychologists bring health and social services into schools or expand those services which already exist. Such programs involve a holistic approach to problems which do not have simple solutions such as substance abuse, youth violence, or teenage pregnancy. For school-based health programs to be effective they must be appropriate to student ages and development levels. As schools move to meet the health needs of their students, psychologists are able to help with program design based on their knowledge of how children grow and develop. Psychologists, for example, have studied why teenagers join gangs or take health risks and can therefore develop strategies that will effectively deal with these circumstances.

Many school districts employ psychologists full-time. In other school districts, psychologists are consultants to administrators and teachers.

Educational psychologists concentrate their work on the study of how effective learning and teaching take place. They consider a variety of factors, such as human abilities, what motivates students, and the effect on the class of the diversity of races, ethnic, and cultural backgrounds that makes up most school populations.

The basic question which educational psychologists have sought to answer is what motivates children to learn and what is detrimental to that learning. Psychologists have found that much of the answer lies in showing students the personal stake they have in the learning process and in giving them choice and personalization in their work assignments.

2.37.3.7 Military Psychologists

The work of psychologists has long helped the military improve training programs and study the effects of military operations on soldiers. For example, psychologists completed a large scale study, involving over 10 000 Gulf War veterans, which examines the impact on soldiers and reservists of rapid mobilization, leaving home and family life, as well as exposure to combat and other war-related stresses. Results from this study convinced the US military to make mental health services more available for American troops when they were deployed into Bosnia in December of 1995. Thirty-seven percent of those soldiers interviewed in the Gulf War veterans study were diagnosed with a psychiatric condition, most commonly either depression (11%) and posttraumatic stress disorder (5%).

As it becomes more common for military troops to be called to international trouble spots to act as peacekeepers, psychologists are studying the psychological stresses involved in that assignment which often includes dealing with strict rules of engagement, a hostile local population, and scores of displaced persons.

2.37.3.8 Psychologists Study the Media

Psychologists were pioneers in studying the effects that television would have on its viewers, especially children. Psychological research done over the since the 1950s has shown that children who watch a lot of violent programming on television are:

(i) less sensitive to the pain and suffering of others,

(ii) more fearful of the world around them,

(iii) more likely to behave in aggressive or harmful ways toward other people (APA, 1991).

In the late 1990s, psychologists are at the forefront of the study and applications of the Internet and other forms of electronic communication and information exchange. Psychologists, for example, want to know what effects the Internet will have a person's ability and desire to deal with people on a face-to-face basis. Is teaching over the Internet an effective way for children or adults to learn? These are questions which psychologists who study the media and new media technologies are investigating.

Clinical application of new media technologies is an area of controversy within the profession. The Internet is an example of a new communications technology with numerous potential applications in the mental health arena.

Does psychotherapy via the Internet have value? Is offering therapy via the Internet an ethical thing for a provider to do? Some proponents of psychotherapy via the Internet say that for people who live in rural areas or for other reasons can not get to a mental health provider's office, help through the Internet is better than no help at all. Others say it is valuable as an adjunct to an occasional face-to-face sessions. However, some observers believe that the Internet is a poor replacement for traditional psychotherapy.

2.37.3.9 Sport Psychology

At the 1996 Summer Olympic Games in Atlanta few individual athletes and teams went into the competition without having first worked with a sports psychologist. Two high profile events, women's gymnastics and track and field, especially benefit from the attention of sports psychologists. Both events demand supreme concentration and put years of training and preparation into a few minutes of performance which will separate victory from defeat.

Psychologists who work with athletes help them refine their focus on competition goals and motivation. They also help athletes deal with the anxiety and fear of failure that often accompany competition.

Ground-breaking research by sports psychologists has shown that quality training and mental preparation for competition were the two consistent elements needed for Olympic caliber success. Quality training was characterized by intense focus during practice, clear daily goals, imagery training to perfect skills, athletes who imagine success, prepare for training, and motivate themselves. Mental preparation for competition included having a precompetition plan, having a competition focus plan, evaluating the lessons of competition, and controlling distractions.

2.37.4 FUTURE DIRECTIONS

As the twenty-first century approaches a few demographic realities are clear. The population of the industrialized nations is aging, the world population continues to grow, and the American population is more diverse than ever. Psychologists will have important roles to play in the study and application of knowledge to help individuals and society as a whole cope positively with these changes.

In 1992, the last year for which data is available, 3373 doctorates in psychology were awarded by universities in the USA. Of these graduates 46% said they would pursue work in clinical and counseling psychology and 16% planned to pursue work in psychological research; the rest varied widely. These newcomers to the discipline will comprise the psychology workforce of the twenty-first century. It may also be important to note that during 1992 63 513 students earned their bachelor's degree in psychology and another 10 215 earned a master's degree in the field (Wicherski & Kohout, 1995).

Psychology continues to be a very popular undergraduate major, ranking second as the most common major for undergraduates on many campuses. About 5% of these new bachelors degree holders go on to earn their doctorate. Others join such job fields as social services, sales and service personnel, business management, engineering and computer science (see Figure 4).

Since the 1890s, psychology has grown dramatically and made important contributions to society. The next 100 years holds the same potential for the discipline.

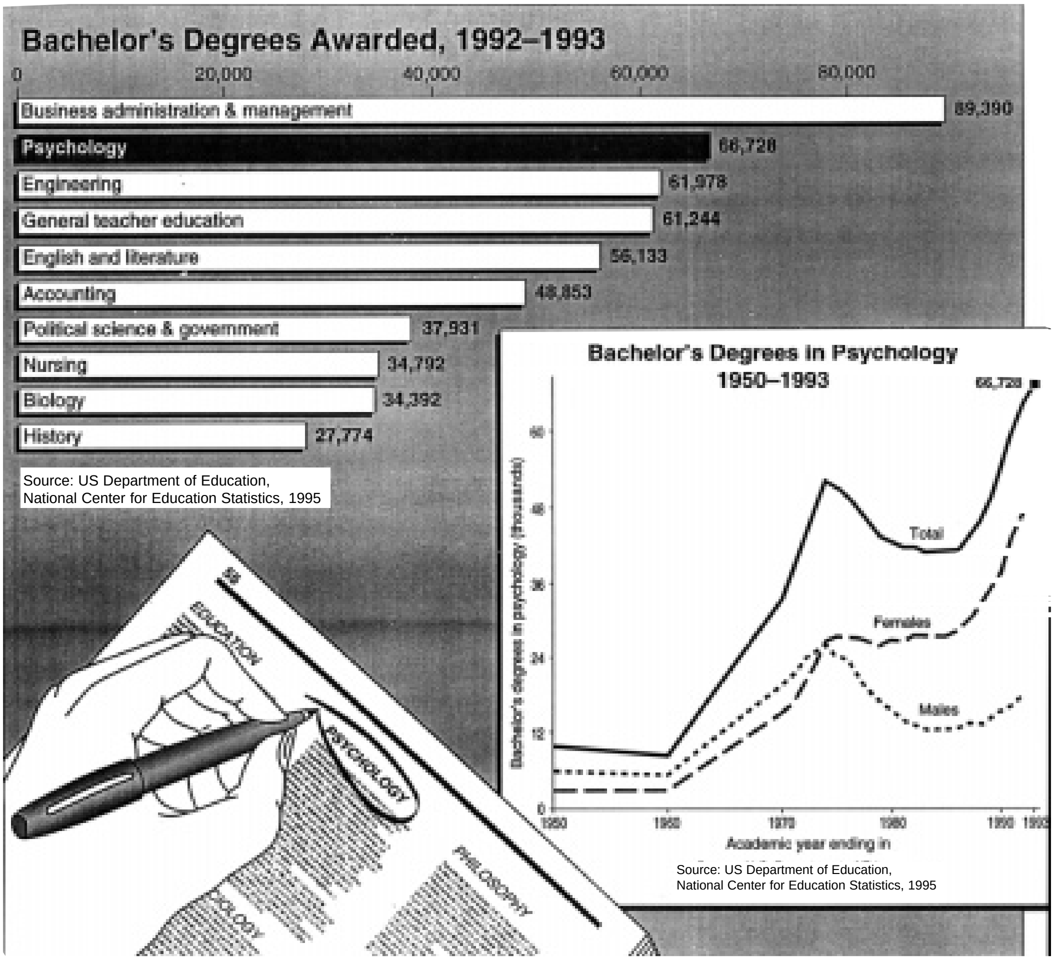

Source: US Department of Education, National Center for Education Statistics, 1995

Source: US Department of Education, National Center for Education Statistics, 1995

Figure 4 Most popular bachelor's degrees awarded, 1992–1993. Source: American Psychological Association.

2.37.5 REFERENCES

American Psychological Association (1991). *Violence on television: What do children learn, what can parents do?* Washington, DC: Author.

American Psychological Association (1996a). *Psychology: Careers for the 21st century*. Washington, DC: Author.

American Psychological Association (1996b). *Report of the APA presidential initiative on family violence*. Washington, DC: Author.

Azur, B. (1996a, May). FDA calls on psychologists to improve medical devices. *APA Monitor,* 23.

Azur, B. (1996b, May). Naps lead to safer piloting, study shows. *APA Monitor*, 22.

Clay, R. (1996, June). Targeted interventions curb the spread of HIV. *APA Monitor,* 29.

Martin, S. (1994, October). Workplace is no longer a haven from violence. *APA Monitor,* 29.

Penn and Schoen Associates (1995). *Public opinion on mental health issues*. Report to the American Psychological Association, Practice Directorate. Washington, DC: Author.

Sleek, S. (1996, June). AIDS therapy: Patchwork of pain, hope. *APA Monitor,* 1.

Wicherski, M., & Kohout, J. (1995). *1993 Doctorate employment survey*. Washington, DC: American Psychological Association, Research Office.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.38
The Image of Psychology in the Mass Media: The United States

RHEA K. FARBERMAN and RAYMOND D. FOWLER
American Psychological Association, Washington, DC, USA

2.38.1 INTRODUCTION 449
2.38.2 WHAT HAVE ALL THESE IMAGES ADDED UP TO CREATE? 450
2.38.3 WHAT IS LOST? 452
2.38.4 THE TELL-ALL AND TALK SHOW PHENOMENON 452
2.38.5 HOW SHOULD THE PROFESSION RESPOND? 456
2.38.5.1 Public Education Campaign 456
2.38.5.2 Psychology's Role in Reporting News Events 456
2.38.6 REFERENCES 457

2.38.1 INTRODUCTION

The reach and impact of television and films are unquestioned. Of all American homes with children, 99% own at least one television. Movies have been popular leisure time entertainment since the 1930s.

In their book, *Tuning in trouble: Talk TV's destructive impact on mental health*, psychologists Heaton and Wilson (1995) write that "television is at the heart of American culture; its power and influence are unparalleled."

About movies, Sklar (1994) writes in his book *Movie-made America*,

> ever since the peep shows first captivated urban audiences, film has had a revolutionary impact on American society transforming culture from the bottom up, radically revising attitudes toward pleasure and sexuality, and at the same time, cementing the myth of the American dream.

Many professions have been created in the public's mind through their portrayals in film and on television. For a generation of Americans, Marcus Welby set the image of what their family doctor should be. In the 1960s Perry Mason set the image of the laconic lawyer who fought for justice and always got to the truth. That image has been greatly changed by later television portrayals of lawyers on such hit shows as *LA Law* and *Murder One*.

People such as Bob Newhart, Woody Allen, and Nurse Ratched in *One Flew Over the Cuckoo's Nest*, have combined many characters to create the public image of the therapist and the patient; an image that in the movies and on television has revolved more around dysfunction than improved functioning.

Research has shown that the public's ability to differentiate between different types of mental health providers is poor (Woods, Jones, & Benjamin, 1986). What role did the media play in this lack of understanding? There are numerous examples of the media's stereotypical portrayal of mental health professionals. There are also numerous examples of the often poor job the media has done in portraying the therapeutic relationship.

During the 1930s numerous films featured a psychiatrist character. All had central European accents, were comical figures, and often appeared to be in need of treatment themselves. The actor Fritz Feld created the mold with his portrayal of a psychiatrist (complete with a facial tic) in *Bringing Up Baby* (1938).

In the 1947 classic, *A Miracle on 34th Street*, the character Kris Kringle is challenged by the Macy's store psychologist who pronounces Kris "dangerous and psychotic." Again however, the protagonist outwits the mental health provider, a small and nervous man who twists his eyebrows incessantly. This was "therapist as villain" movie-making. When an elderly man believes himself to be Santa Claus, eventually even the lawyers see his humanity and goodwill and begin to root for him. The villain of the film is the psychologist hired to screen new employees, he is a bundle of nervous habits and odd behavior—paranoid and downright mean.

In the 1958 film, *No Time for Sergeants*, Andy Griffith plays a Georgia farm boy who joins the air force and outwits the military psychiatrist. "Well I reckon I love my mother," Griffith tells the psychiatrist. "Don't you love your mother?" he then asks, evoking anger from the doctor.

The 1960s hit television sitcom *I Dream of Jeanie* featured a NASA psychiatrist, Dr Bellows. Poor Dr Bellows was always in a state of befuddlement as the other main characters, Jeanie, Roger, and Major Nelson, hide Jeanie's true identity (see Figure 1).

Probably the most far-reaching of all the television and film characterizations of psychologists and psychiatrists is the television sitcom of the early 1970s, *The Bob Newhart Show*. The Newhart show was extremely popular and held the Saturday evening time slot. The protagonist, Bob Newhart, a psychologist, was a caring but ineffectual therapist. A stream of patients came in and out of his office, each more strange and dysfunctional than the last. This one show, the only network show of the time featuring a psychologist as one of the main characters, left an indelible impression in the TV viewing public's mind—the psychologist as the middle-aged man in a now wildly out of style polyester suit, kind, but bumbling and unable to do much of anything to help his patients cope with their problems. His patients were an odd collection of neurotic souls—caricatures of people rather than real people. A virtual parade of dysfunction on both sides of the therapeutic relationship.

The cable television network, *Comedy Central*, has created a modern day psychotherapist as sitcom example, *Dr Katz: Professional Psychotherapist*. The show's producers describe it as an animated TV show about the funny, lovable psychotherapist who people love to take advantage of. The fact that this show is animated is somehow fitting with the way that the entertainment industry has treated and continues to portray the profession.

The best-selling novel by Pat Conroy and later a hit movie (1991), *The Prince of Tides* created a new portrayal of the psychiatrist—different from earlier characterization but also troubling. In the movie, Barbara Streisand plays a competent, compassionate, very human professional, but one with questionable ethics. The story line takes the Streisand character from psychiatrist to the main character's lover. The problem is the main character, Tom Wingo, who is also the brother of one of Streisand's current patients.

2.38.2 WHAT HAVE ALL THESE IMAGES ADDED UP TO CREATE?

The question is, what effect have these negative and inaccurate portrayals of psychiatrists and psychologists had on the viewers' psyche and health consumer behavior? The provider-character who is "crazy," strange, befuddled, and ill-tempered gets the most laughs and is most common. But is it educative and reassuring for the viewing public?

What has been the effect of the media portrayal of psychologists and other mental health providers? Many studies have shown that the public is not well educated when its comes to mental health issues nor types of mental health specialists. Americans are unsure as to how psychotherapy works and what psychologists and psychiatrists actually do.

In 1994, the American Psychological Association (APA) conducted both a national telephone survey and consumer focus groups to ascertain the public's attitudes towards and understanding of mental health issues and providers. Some important results were as follows:

(i) The public is nearly universally unable to differentiate between the different types of mental health providers.

(ii) Few people understand the educational and licensure requirements to practice psychology. Just 36% of those asked know that a doctoral degree is required to be a psychologist. Additionally, one in five believed that psychologists need a medical degree to practice; these data may reflect people confusing psychologists with psychiatrists.

(iii) Finally, there exists a great deal of confusion about the best way to evaluate and select a mental health provider.

A summary of the findings of the national focus group series reported the following,

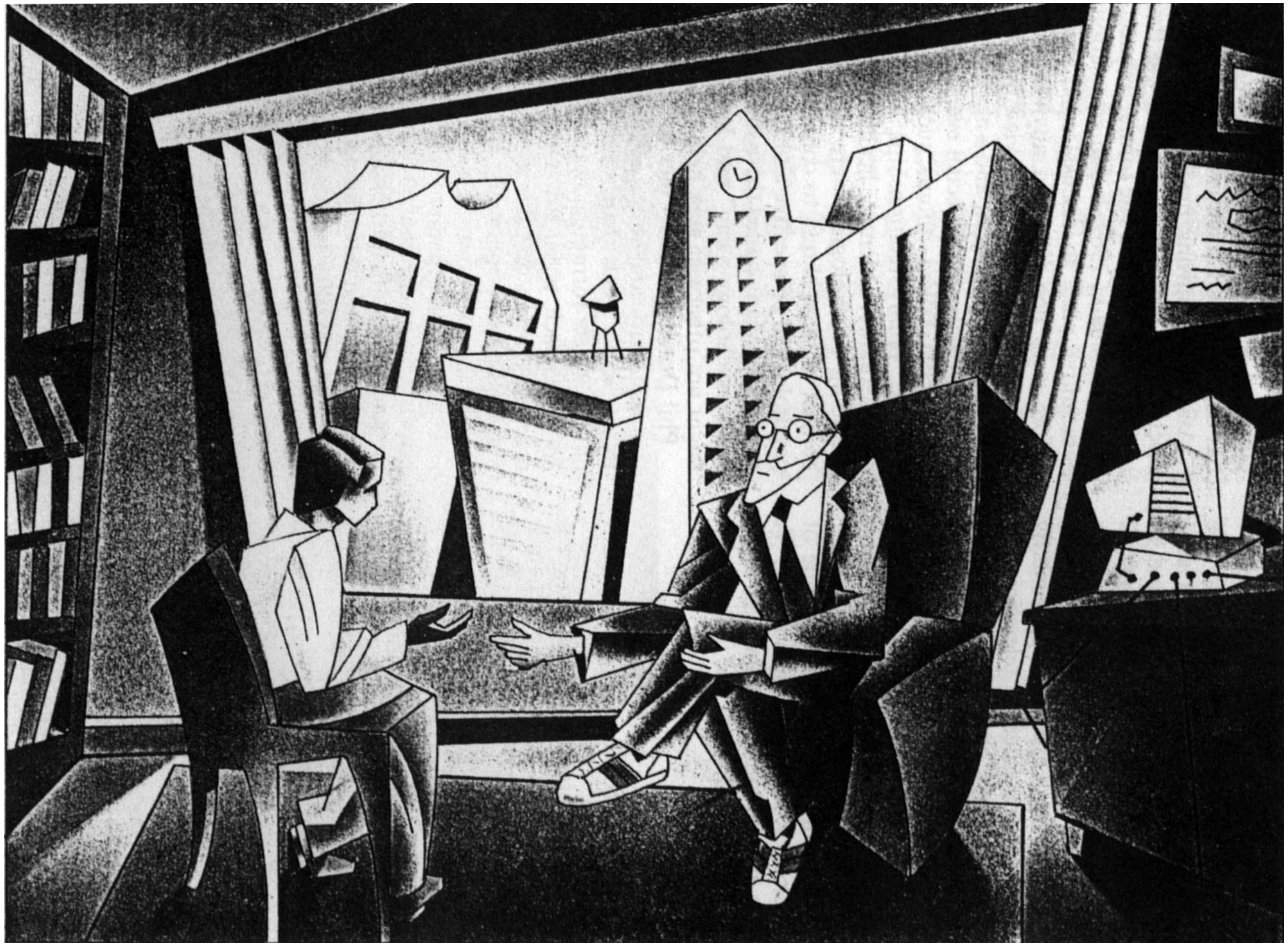

Figure 1 The media often portrays therapists as characters and humorous figures (Source: APA).

"across all the groups there was general confusion about the differences between the various types of mental health care providers. Most had limited experience in using or being exposed to mental health care services. Some made comments about the images of such professions being based on what they had seen in the media, particularly talk shows and programs like *Bob Newhart*."

Another focus group finding was that the term "mental health" was actually associated with "mental illness" by most people and to these people mental illness meant "being crazy."

This research also identified five primary barriers which prevent people from seeking psychological services. Three seem particularly related to the images created by television and movies—concerns about the treatment outcome (43%), the length of treatment (26%), and about how seeking treatment will affect their career (26%).

Where does the public get the impression that psychotherapy treatment would go on for indefinite lengths of time and the outcome of all of this time would be questionable? A Woody Allen movie comes to mind! "I don't think you ever get cured," said one focus group participant.

2.38.3 WHAT IS LOST?

Have the wrong impressions of the therapeutic process that the public has had an effect on its mental health? Data, again collected by APA, suggests that they have.

The APA found an information gap when it did public attitudes research. Most people knew that psychological health is important to ones' overall well-being, but did not feel that they knew when it was appropriate to seek mental health services or how to go about selecting a provider.

Of even greater concern than a lack of information is the negative impression that the media's handling of the mental health profession has given the public. Research has shown that the therapeutic set, the concept that if the client believes the therapist to be an expert in psychological techniques the client will improve, holds (Turkat, 1977). The question is, if the public generally believes that mental health professionals are incompetent how does that interfere with the therapeutic relationship?

2.38.4 THE TELL-ALL AND TALK SHOW PHENOMENON

While in the past, the mass media helped create a false image of mental health providers, in the 1990s, many media personalities are helping to decrease the stigma around mental health issues through their own public acknowledgments of mental health problems. In 1987, Patty Duke revealed her battle with manic depression in an autobiography and in numerous public appearances on behalf of mental health causes. Newsman Mike Wallace has talked openly about his battle with depression as has columnist Art Buchwald.

Another promising media direction for the discussion of mental health issues was the early television talk shows. The format, pioneered by Phil Donahue in 1967, began as a serious and substantive place to discuss social issues that before were not discussed anywhere else. At their inception, the talk/interview format shows were a breath of fresh air and a new opportunity to discuss issues that affected everyone but that few people talked about, such as domestic violence, family relationships, race relations, and sexuality. Mental health professionals including psychologists were often invited to be part of these shows to comment and add a psychological perspective and knowledge to the discussion. Many hailed these shows as a way to "give psychology away" and help people who might not otherwise seek help.

Phil Donahue's show met with great success and popularity, and not surprisingly was copied. By the early 1990s there were nearly a dozen talk/interview format shows on the airways in America, each trying to outdo the other for ratings by being that much more provocative, outrageous, and obnoxious.

With few exceptions, TV talk shows became little more than organized, electronic bar-room brawls. During the final 60 seconds of the show, the mental health professional was asked to appear on air and solve the guest's problem, be it marriage infidelity (my mother slept with my husband is not an atypical show topic), eating disorders, body image disorders, gambling addictions, or a host of other serious and intensely personal life problems.

Heaton and Wilson (1995) have studied the TV talk show phenomenon and have written about it in their book *Tuning in trouble: Talk TV's destructive impact on mental health*. They write, "TV talk shows make a mockery of the mental health profession by obscuring the fact that change and recovery are most often a long and painful process. Innocent viewers, moreover, succumb to a "copycat syndrome" being persuaded that they too not only have the same problem, but can solve it with similar quick-fix, sound-bite therapy. Television talk shows entertain their enormous viewing audience with a steady stream of wounded guests, self-serving gurus, and manipulative hosts who offer quick

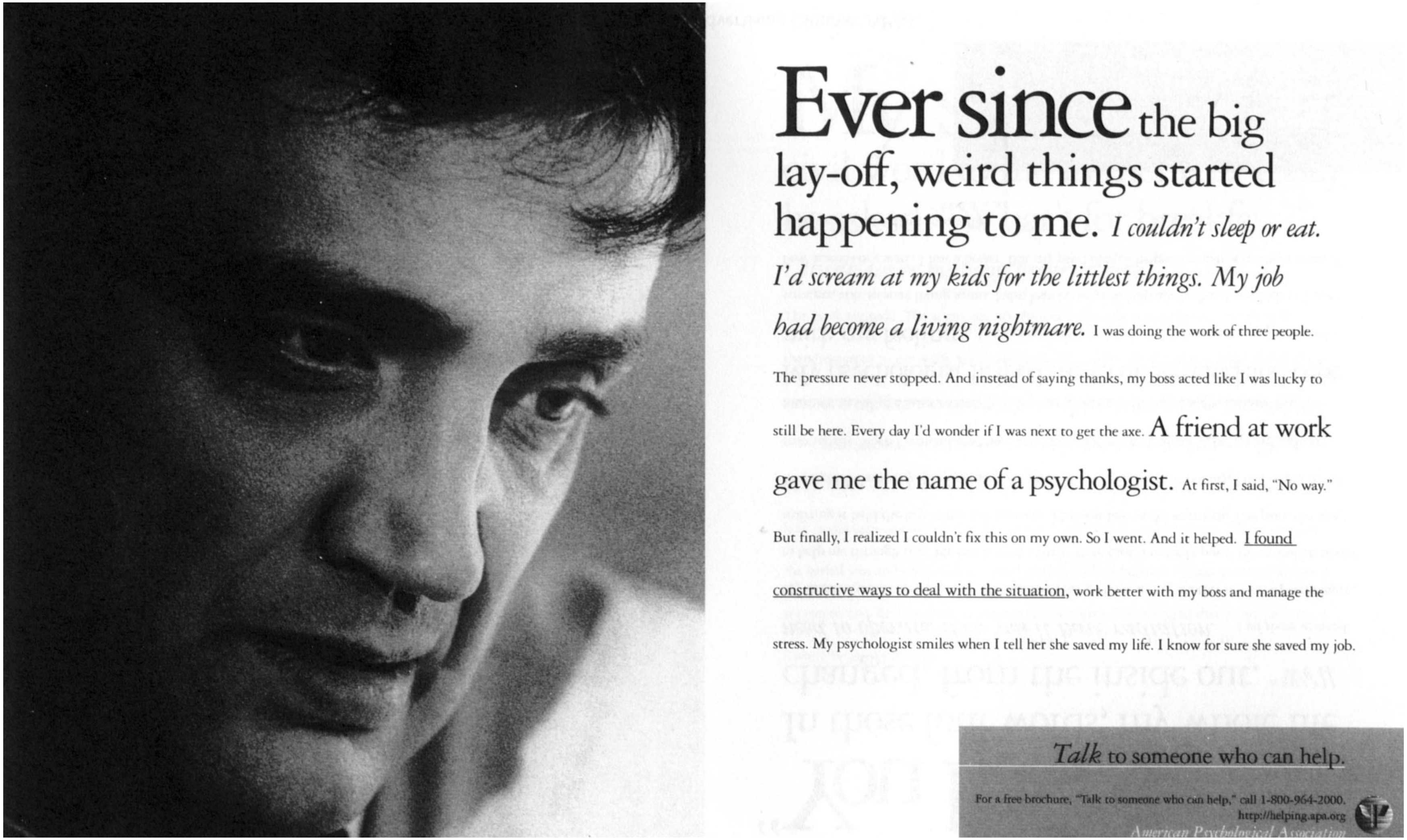

Ever since the big lay-off, weird things started happening to me. *I couldn't sleep or eat. I'd scream at my kids for the littlest things. My job had become a living nightmare.* I was doing the work of three people. The pressure never stopped. And instead of saying thanks, my boss acted like I was lucky to still be here. Every day I'd wonder if I was next to get the axe. A friend at work gave me the name of a psychologist. At first, I said, "No way." But finally, I realized I couldn't fix this on my own. So I went. And it helped. I found constructive ways to deal with the situation, work better with my boss and manage the stress. My psychologist smiles when I tell her she saved my life. I know for sure she saved my job.

Talk to someone who can help.

For a free brochure, "Talk to someone who can help," call 1-800-964-2000.
http://helping.apa.org
American Psychological Association

Figure 2 Public Education Advertising (Source: APA).

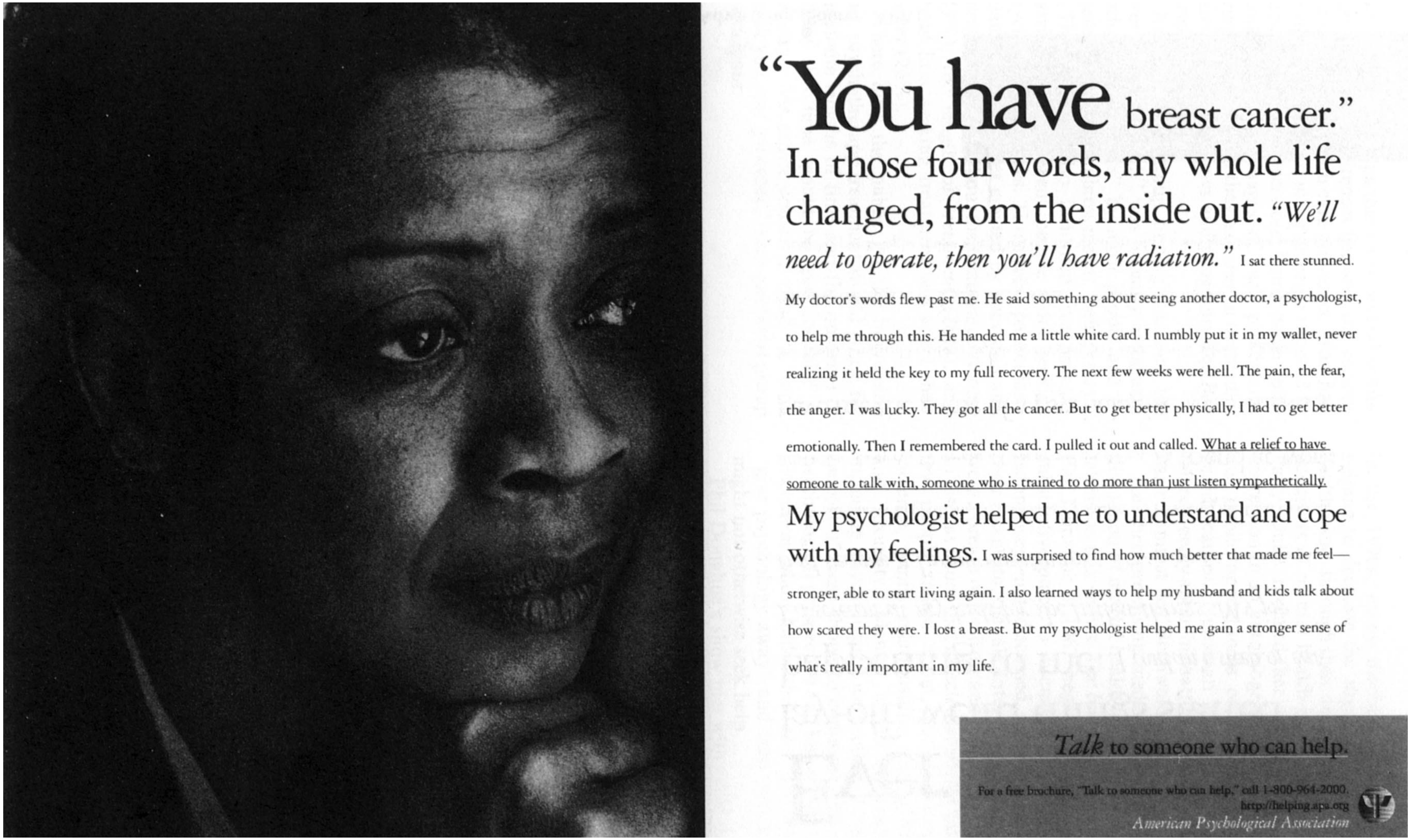

"You have breast cancer."

In those four words, my whole life changed, from the inside out. *"We'll need to operate, then you'll have radiation."* I sat there stunned. My doctor's words flew past me. He said something about seeing another doctor, a psychologist, to help me through this. He handed me a little white card. I numbly put it in my wallet, never realizing it held the key to my full recovery. The next few weeks were hell. The pain, the fear, the anger. I was lucky. They got all the cancer. But to get better physically, I had to get better emotionally. Then I remembered the card. I pulled it out and called. What a relief to have someone to talk with, someone who is trained to do more than just listen sympathetically.

My psychologist helped me to understand and cope with my feelings. I was surprised to find how much better that made me feel—stronger, able to start living again. I also learned ways to help my husband and kids talk about how scared they were. I lost a breast. But my psychologist helped me gain a stronger sense of what's really important in my life.

Talk to someone who can help.

For a free brochure, "Talk to someone who can help," call 1-800-964-2000.
http://helping.apa.org
American Psychological Association

Figure 3 Public Education Advertising (Source: APA).

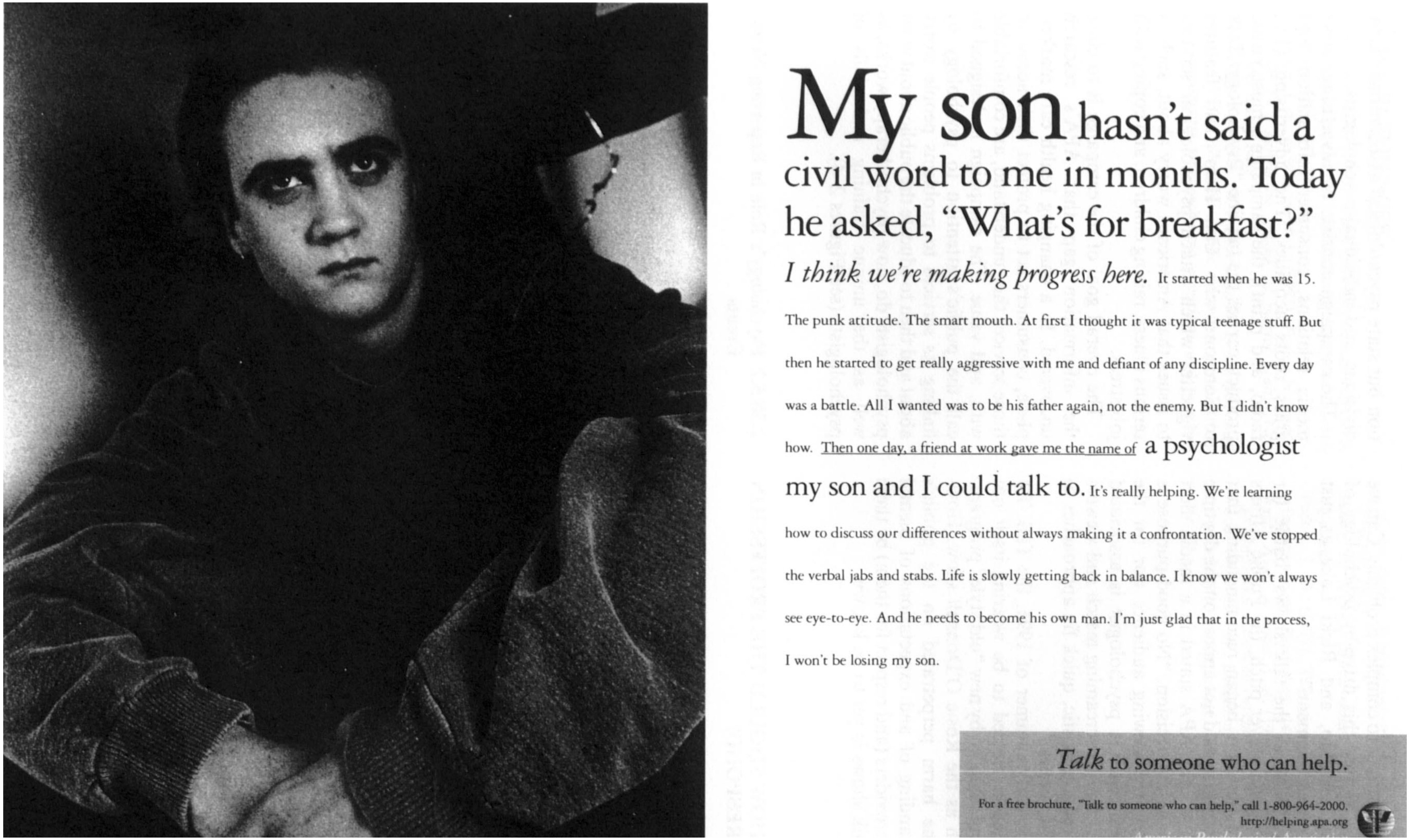

My son hasn't said a civil word to me in months. Today he asked, "What's for breakfast?"

I think we're making progress here. It started when he was 15. The punk attitude. The smart mouth. At first I thought it was typical teenage stuff. But then he started to get really aggressive with me and defiant of any discipline. Every day was a battle. All I wanted was to be his father again, not the enemy. But I didn't know how. Then one day, a friend at work gave me the name of a psychologist my son and I could talk to. It's really helping. We're learning how to discuss our differences without always making it a confrontation. We've stopped the verbal jabs and stabs. Life is slowly getting back in balance. I know we won't always see eye-to-eye. And he needs to become his own man. I'm just glad that in the process, I won't be losing my son.

Talk to someone who can help.

For a free brochure, "Talk to someone who can help," call 1-800-964-2000.
http://helping.apa.org

Figure 4 Public Education Advertising (Source: APA).

and easy solutions to complex problems. Can we afford to dismiss this drive-by psychology of Donahue, Geraldo, and Ricki Lake as just harmless entertainment?"

In late 1994, as the talk shows' battle for viewers came to fever pitch, the Public Affairs Office of the APA began recommending that psychologists should not appear on the daytime talk shows. The APA stated in a media alert announcing the decision, "No good purpose is served for the viewing audience nor for the profession when a psychologist is associated with a 27 minute screaming match and viewers are offered unrealistic, quick fix approaches to serious life issues."

By the early summer of 1996, the Talk TV phenomenon seemed to be wearing itself out and being replaced by new "old style" programing such as the Rosie O'Donnell show. However, the harm perpetrated on the public's understanding of and expectations of mental health providers (and respect for them) by these trash talk shows is yet to be known.

2.38.5 HOW SHOULD THE PROFESSION RESPOND?

Comical and inaccurate portrayals of mental health providers has created a stereotype image of the profession and has likely added to the stigma around mental health services. These issues challenge the profession to act to provide the public with more accurate information and a more accurate, more professional, more human image of its ranks. "It's critical that psychology communicate its value to many publics; consumers, the media, and policy-makers," says Dorothy Cantor, a past president of APA, "if psychology doesn't communicate about psychology, who will?" she asks.

2.38.5.1 Public Education Campaign

The public image of psychology has been the concern of "lone voices" in the profession for decades. Research studies of consumer information and attitudes have pointed toward the profession's need to better define and market itself to the public (Benjamin, 1986; Schindler et al., 1987; Webb & Speer, 1986).

In 1995, the APA Council of Representatives directed that the association undertake a national public education campaign designed to communicate the value of psychological services. The campaign was unveiled in February 1996 and is expected to be a long-term effort involving not only the national association but state psychological associations, APA divisions, and individual psychologists.

The campaign consists of advertising, news media relations, consumer information, and grass roots activities. The advertising (TV, radio, and print), the campaign's most visual product, carries the tag line "Psychology: Talk to Someone who Can Help" and features vignettes which attach psychological services to issues that Americans worry about, such as serious illness, raising children, and coping with job stress.

The overall goal of the campaign is to close the information gap that APA's research uncovered. In a changing health care marketplace, consumers act to protect their access to those services they understand, are comfortable with, and value. The campaign is designed to call the public's attention to psychology by linking its services to problems people worry about and then to educate the public about what psychologists do, how psychotherapy works, as well as the unique training and skills of psychologists (see Figures 2–4).

2.38.5.2 Psychology's Role in Reporting News Events

Gaining expertise in working with the news media is another important way in which psychology can reclaim its image and its image-making. The news media often call upon psychologists to help interpret or to comment on daily events. As major events unfold, the APA Public Affairs Office gets calls from print and electronic reporters looking for people to go on the air or on the record to help explain events. Examples of such events include the bombings, arrests, aircraft disasters, the controversy of recovered memories, post-traumatic stress disorder in earthquake survivors, and others. Tomorrow's news may be another opportunity for psychology to help redefine its image and display its knowledge of human behavior. However, in order to take advantage of this opportunity, individual psychologists must be prepared for the newsmaker role. The best way to get this preparation is to participate in media training programs (offered at APA's convention every year) and to practice those skills.

Without the resources to produce a new hit movie in which the hero is a competent, professional, sound, and humanistic psychologist, there are other ways in which individual psychologists can help redefine psychology's public image, for example, participation in grass roots public education activities and availability to the news media.

2.38.6 REFERENCES

Benjamin, L. T. (1986). Why Don't they understand Us? A history of psychology's public image. *American Psychologist, 41*(9), 941–946.

Heaton, J. A., & Wilson, N. L. (1995). *Tuning in trouble: Talk TV's destructive impact on mental health.* New York: Jossey-Bass.

Schindler, F., Berren, M. R., Hannah, M. T., Berigel, A., et al. (1987). How the public perceives psychiatrists, psychologists, nonpsychiatric physicians, and members of the clergy. *Professional Psychology: Research and Practice, 18*(4).

Sklar, R. (1994). *Movie-made America*. New York: Vintage.

Turkat, I. D. (1977). Television psychologists and the therapeutic set. *Psychology, 14*(2), 65–68.

Webb, A. R., & Speer, J. R. (1985). The public image of psychologists. *American Psychologist, 40*(9).

Wood, W., Jones, M., Benjamin, L. T. (1986). Surveying psychology's public image. *American Psychologist. 41*(9).

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.39

National Priority Needs: Implications for Clinical Psychologists in the United States

JOANNE E. CALLAN
California School of Professional Psychology, San Diego, CA, USA

2.39.1 INTRODUCTION 459
2.39.2 SELECTED NATIONAL PRIORITIES 460
2.39.2.1 Health Care 461
2.39.2.2 Education 463
2.39.2.3 Work World and Economics 465
2.39.2.4 Families and Communities: Relationships, Systems, and Communication 466
2.39.3 SUMMARY 467
2.39.4 REFERENCES 468

2.39.1 INTRODUCTION

Clinical psychology's very roots were planted firmly in its science base, yet also in its promise as a profession responsive to a wide range of human needs. Its essential focus was initially, and remains, on psychological and behavioral aspects of human functioning. Although its earliest foundations lay in general psychology emanating from Wilhelm Wundt's University of Leipzig laboratories in the 1870s, several traditions within psychological thinking are credited as contributing significantly to the development of clinical psychology as it is known today, and, presumably, to directions it will take in the future.

Korchin (1976), writing about major influences on clinical psychology, supported Watson's (1953) earlier identification of two traditions, other than the early European experimental influence, which have had marked impact: psychometrics and dynamic thinking. Rotter (1963), while noting the contributions of experimental and theoretical psychology to clinical psychology, pointed out the significance of clinical psychologists' gradual move from working predominantly in the 1900s–1940s with children and with assessment activities (in particular, intelligence testing) to functioning more as therapists and case managers since the 1950s. Arguably, a time of clinical psychology's greatest consolidation occurred during and just after World War II, when clinical psychologists joined other health care providers in responding to the many psychological needs of America's service men and women. Indeed, several psychologists considering the history of clinical psychology (e.g., Korchin, 1976; Shakow, 1969) relate the evolution of three major professional activities carried out by clinical psychologists—diagnosis, psychotherapy, and research—to World War II's special psychological needs. Korchin referred to these three activities as the "holy trinity of postwar clinical psychology" (1976, p. 45).

Later, fueled by political and social forces of the 1960s, the community mental health movement underscored psychology's importance to public health needs. By that time, clinical psychology was well established as one of the nation's major mental health professions. Just as the Veterans Administration (VA) had been singularly instrumental in advancing the education and training of psychologists to meet service needs following World War II (continuing its support to the present), the National Institute for Mental Health (NIMH) provided major support for the education and training of mental health professionals, including clinical psychologists in the 1960s, 1970s, and 1980s (with decreasing support in the 1980s and 1990s). Accordingly, clinical psychology has been responding to major public needs since the 1940s and, no doubt, will continue its commitment to national needs and priorities in the future.

Parallel developments in education and training also reflected clinical psychology's science and professional bases as well as its responsiveness to public needs. The Boulder Conference, led by American psychologists representing the American Psychological Association, the VA, and the NIMH, was held in 1949 for the purpose of developing standards for educating and training clinical psychologists (Raimy, 1950). One of the most enduring outcomes of the conference was the establishment of the scientist-practitioner training model, a model reaffirmed in several subsequent conferences (Korchin, 1976). In 1973, the Vail Conference addressed a lingering and also growing concern about how psychology could best educate and train clinicians to be responsive to changing and increasing demands for psychological services in both public and private sectors. A central question was how to assure clinical as well as research expertise, and conference deliberations on this matter led to the recognition of a new training model, one that emphasized professional training (Korchin). This recognition did not eliminate the scientist-practitioner model, which, in fact, continues to be observed by many graduate programs; rather, it broadened the ways in which graduate programs in clinical psychology would be designed and implemented.

Since the 1960s, as indicated previously, clinical psychology, both as a science and as a profession, has been recognized as a national player in public health, especially with regard to mental health. The NIMH, for example, has identified four disciplines as essential to addressing national mental health needs: clinical psychology; psychiatry; nursing; and social work. These four disciplines have been included in NIMH's applied and training programs, and as noted above, NIMH-supported clinical training programs provided major resources for several decades to assure an available pool of mental health service providers. Clinical training funding, divided among these four mental health professionals, was at the highest level of $93 million in FY 1967; regrettably, however, that level dropped to 10.9 million in FY 92 and to 2.9 million in FY 1993 (Callan, 1993).

Clearly, national leadership on mental health needs and services has been influenced considerably by NIMH-sponsored programs and directions. In the late 1980s, for example, NIMH in collaboration with others, including professional and advocacy groups as well as educators and trainers, identified four populations whose urgent mental health needs merited special attention: (i) seriously emotionally disturbed children; (ii) the seriously mentally ill; (iii) ethnic minorities; and (iv) the elderly. Clinical psychology was a central player in the series of conferences held between 1988 and 1992 on the education and training of health care providers for these particularly needy groups.

The mental health challenges of these groups, along with numerous other public problems, constitute a broad array of national needs to which clinical psychology has relevance. Arguably, they are apparent in all major human and social systems: in government, education, health, law and justice, transportation, environmental and community safety, communication, and in the worlds of work, art, and play. Individuals, families, and communities face myriad challenges in everyday living, ranging from meeting basic needs for food and shelter to higher order ones related more to quality of life. When special problems such as those of the underserved, under-represented, or the otherwise needy are also considered, the complexity and urgency are almost overwhelming. Identifying and addressing priorities among them, especially in the context of an increasingly complex and diverse society, are daunting tasks. Moreover, effectively responding to them on a national scale will certainly require the knowledge and skills as well as the caring of many. Not only must clinical psychology play a leadership role in this response, it must become an active, ongoing partner with other psychologists, other professions, and the public.

2.39.2 SELECTED NATIONAL PRIORITIES

In order to provide such leadership and build such partnerships, it is imperative that clinical psychologists become well informed on national needs. This chapter, building on the history of clinical psychology, focuses on selected national priorities, first identifying several key need categories and then elucidating the urgency associated with each and the implications for clinical psychology. Four need categories are presented: (i) health care (including mental health); (ii) education; (iii) work and economics; and (iv) families and communities: relationships, systems, and communication.

Although these categories take into account special populations, such as the four identified by NIMH, they provide a broader context, one identifying the larger human and societal needs that constitute national priorities for the USA as the twenty-first century nears. Reflecting overarching challenges, they highlight the broad scope as well as the complexity of national issues with which clinical psychology can engage in the next few decades and thereby make a real and dramatic difference for humankind. As clinical psychologists become both better informed and actively involved in addressing national priorities, the relevance and promise of clinical psychology as a science and as a practice will achieve heightened visibility and meaning.

2.39.2.1 Health Care

Rapidly escalating health costs since the 1970s as well as diminishing access to health care have mandated health care reform. According to Resnick and DeLeon (1995), cost estimates for health care in 1993 were around \$880 billion; moreover, access to health care is becoming nonexistent for many Americans with almost 40 million Americans lacking health care coverage and an estimated two million losing health care coverage each month. Mental health costs have been spiraling as well. Annual costs of depression exceed \$43 billion (Greenberg, Steglier, Finkelstein, & Berndt, 1993), while the annual overall expenses associated with mental health and substance abuse is \$273 billion (Rice, Kelman, Miller, & Dunmeyer, 1990). Inpatient substance abuse services grew 30% between 1986 and 1988, and adolescent inpatient treatment grew 65% ([see Frank & Salknero, 1990], as reported in Slosar & Lettieri, 1996).

Unquestionably, health care reform is a national priority, and only those disciplines actively engaged in the reform process will influence the directions it takes, an involvement that is critical in the face of such developments as expanding managed care. Indeed, in order to have a viable future, clinical psychology must exercise a leadership position in national health care planning and service delivery. Its leadership must be positioned from the broadest possible base, focusing on health care and not limited to mental health care, with which it has been associated in the past to a large extent. Although psychotherapy remains an established and effective modality (see below), it cannot stand as the *sine qua non* of clinical psychology. The developing trend since the 1980s in which clinical psychologists have been increasingly assuming roles more associated with general health care must persist.

Behavioral health research and service delivery constitute an area of general health care to which clinical psychology has natural relevance. In fact, clinical psychologists as well as health psychologists have become increasingly involved in professional activities related to behavioral health since the late 1970s. Among the many compelling behavioral health problems existing in the USA today requiring both further research and the development of more effective prevention and treatment strategies, one of the most pressing is drug abuse. Indeed, in its different forms and combinations, the abuse of drugs has reached such impressive levels as to have been identified as one of the nation's highest health and social priorities.

Alcohol dependence and abuse are among the most prevalent national disorders within America, for example, as reported in the *Diagnostic and statistical manual of mental disorders* (*DSM-IV*) (American Psychiatric Association [APA], 1994), and at a terrifically high price: in 1988, for example, costing $>$\$58 billion (National Institute on Drug Abuse, 1990). Of the 73% among the population reporting use of alcohol, $>$19 million are problem drinkers according to Winick (1992). Also according to Winick, another group, consisting of the approximately 36% of the population who are cigarette smokers, a figure representing a slight reduction from several years ago, actually constitute an especially alarming problem, given the upsurge of smoking among young people (Winick, 1992). A parallel problem involves the increasing number of adolescents and young adults who are using marijuana. Approximately one-third of the population reports the use of marijuana one or more times, making it among the most widely used drugs (APA, 1994). It has been described by Johnson and Muffler (1992) as the most frequently used illicit drug in the nation among those using drugs recreationally or intermittently and also among the most common secondary drugs used by abusers of other drugs, along with alcohol and cigarettes.

The same authors indicate that one of the more recent and threatening illicit drug developments has been the rapid invasion and increased usage of crack cocaine in the USA since 1985. Admissions for cocaine users rose from approximately 2000 in 1983 to >11 000 in 1989 (most for crack cocaine abuse rather than cocaine snorters), and treatment modalities and programs are insufficient to address the increased number of abusers. These authors caution that "much is unclear about how long the crack era will last and how long its effects will persist" and underscore the "crying need for a new treatment approach that is even modestly effective" (p. 127).

Referring to the "terrible tragedies" that arise out of drug abuse, Lowinger (1992) notes that it is among the nation's poorer citizens and also the infants of drug-addicted women that the greatest toll is exacted. His admonition is:

> Society has a stake at least in reducing, even if unable to eliminate completely, such tragedies. If it is to do so effectively, the true issues that must be dealt with are unemployment, poverty, racism, sexism, and feelings of powerlessness and alienation in our society. Every American has the right to hope for a future that can provide healthy sources of self-esteem, pleasure and affiliation. (p. 141)

Preventing and ameliorating costly and painful physical as well as psychological problems associated with such abusive behaviors are clearly critical health care goals. Psychologists are challenged to become professionally involved with interventions that effectively interrupt the addictive cycle, including behavioral as well as other approaches. Clinical psychologists may serve not only as consultants but also as program designers and managers for a variety of health programs and interventions, and, of course, they may continue to make critical contributions through their research expertise, as suggested by Barlow, Hayes, and Nelson (1984). As Nirenburg (1983) indicates, clinical psychologists work effectively in the treatment and management of addictive behaviors such as alcohol abuse, obesity, cigarette smoking, and drug abuse with a variety of treatment techniques including social skills training, problem-solving skills, vocational skills, chemical treatment, aversive conditioning, and relaxation.

Clinical psychologists provide valuable services for other health problems as well, and it is likely that such contributions will be in even greater demand in the future. They provide neuropsychological evaluations and treatment planning for brain-injured patients, for example, clinical services that can be expected to become even more sophisticated and accessible in parallel with anticipated technological advances. Clinical psychologists also contribute creatively by developing and managing prevention programs (Lorion, 1991), one of the most promising clinical efforts, from both cost-saving and disease prevention perspectives.

Even as clinical psychologists are assuming a broader role in general health care, however, they must continue to be skilled psychotherapists. Psychotherapy remains a viable and effective clinical tool, demonstrably a treatment of choice for a wide range of health problems. Even though further research will be helpful in demonstrating its value as a clinical treatment to a broader audience of health service providers and the general public, psychotherapy has already been shown to be an effective intervention for both physical and psychological problems. Reviews of recent research elucidate psychotherapy's efficacious and cost-saving contributions in alleviating or contributing to the cure of both physical and mental illnesses, a particularly critical finding in these times of managed care practices that delimit intensive and long-term psychotherapy. Lazar (1997), for example, identifies several mental health populations who benefit markedly from extended dynamic psychotherapy: patients with anxiety disorders, especially children; depressives, especially adolescents and also those with interpersonal and self-esteem symptoms; and post-traumatic stress disorder, borderline personality disorder, and multiple personality disorder patients as well as schizophrenics. Further, in describing positive outcomes from intensive psychotherapy with patients having certain physical illnesses, Lazar notes such dramatic results as increased survival rates for metastatic breast cancer patients in psychotherapy and points out as well the value of extended group psychotherapy in such health promotion programs as diet and exercise regimens and also refers to the greater stability of diabetic patients in three to four times a week psychotherapy.

It is ironic that a country with such a high degree of expertise in health care—supported by comprehensive research activities resulting in elucidating findings such as those on efficacious treatment noted above—has so many people to whom health care is either minimally provided or absent altogether. In the face of such inequity and, moreover, in the best interests of the general public, clinical psychology is challenged to bring to bear its knowledge and skills, as well as its essential regard for the human condition, as an energetic player in health care delivery. Not only must it address the wide range of health issues for which it has both relevance and promise, clinical psychology must assume energetic and assertive leadership for national health policy development and administration. Indeed, from both research and practice perspectives, clinical psychology has much to offer. Its expertise is key to the success of future health care delivery in this country.

2.39.2.2 Education

The USA's commitment to education is interwoven in its history as a nation. Both before and after the Revolution, leaders such as Thomas Jefferson and Benjamin Franklin were advocates for educating the masses, given what they saw as the "intimate relationship between educational opportunity and social, political, and economic education" (Gutek, 1991, p. 28). Jefferson strove diligently for many years to implement public education at all levels, from the lower grades through higher education; however, he achieved success only in the establishment of the University of Virginia. Although early resistances such as he encountered slowed national efforts to educate the masses, gradually throughout the nineteenth century progress was realized first with the establishment of common schools and then high schools. Cremen (1988), tracing America's commitment to education as reflected in the views of Thomas Jefferson, Horace Mann, and John Dewey, and also through conservative as well as progressive national administrations, pointed to the influential role of education in America:

> The notion of education as holding the key to the achievement of the nation's most fundamental political aspirations—of racial integration or social quality or economic productivity—remained constant, indeed no longer a characteristic progressive idea but increasingly a characteristic America idea, held with the same depth of commitment that had attached to religious belief during an earlier era in American history. As American education went, so would go the American people and the American nation. (p. 155)

Arguments supporting psychology's applications to education can be traced back to Herbart and, from a specifically more American perspective, to William James' *Principles of psychology* (1890) (James, 1952) and his subsequent writings on psychology and education. Indeed, among the many theoretical influences impacting American education in the late nineteenth and twentieth centuries, theories of psychology were among the most influential.

Even with continuing support for public as well as private education up to the present time, education has suffered strong criticism in the last couple of decades, in sharp contrast with the positive regard it enjoyed during the years of the education heyday that followed World War II. In the 1950s, for example, the promise of a high school education for all Americans and the expectation of an opportunity for higher education or training were fostered by the implementation of a number of federal programs, including the GI Bill. After the 1960s, however, not only have both dimmed considerably, but, as reflected in *Education: A nation at risk* (US National Commission on Excellence in Education, 1983), concerns about the overall lack of quality in America's educational programs have increased. As a result, the general public has joined educators and policymakers in demanding education reform. Holtzman (1992) writes that "major reform in public education, perhaps even radical changes that would scarcely have been voiced a generation ago, will be the driving force to ameliorate" many of the social problems such as "community renewal, family preservation, child development, and education" (p. xiv).

From a national perspective, specific priority areas of education reform have been identified and supported in important pieces of legislation, for example, those outlined in *Goals 200: The Educate America Act* (PL 103-227), passed in March, 1994, and *The Improving America's Schools Act of 1994* (IASA, PL 103-382). One section of IASA focuses on safe and drug-free schools and communities, acknowledging and responding to the nation's serious drug abuse and violence problems. Given the prevalence of these problems, especially among youth, and their invasion into the schools, clinical psychologists have a particular responsibility, and are in a special position, to address them. Findings from their research have already been invaluable to policymakers, for example, as in legislative decision-making around such matters as appropriate television programming for children (Wilcox & Kunkel, in press).

Another development which has become a major influencing factor in education reform is the view that early education programs as well as K-12 schools must address emotional and adjustment needs of students as well as content requirements. A related movement has advocated the provision of school-based services, including health care services. These developments are particularly inviting to clinical psychologists, whose research and applied expertise are so relevant to both: (i) the emotional development and adjustment of children and adolescents; and (ii) cognitive and learning strategies. Although this relevance is well established, it has been the case that for a good part of this century, psychologists, apart from school and educational psychologists, have not been consistently involved with schools. Many clinical psychologists, for example, may have worked with patients or clients on school-related problems (e.g., through professional activities with parents and students, including assessment and interventions), but not have spent many professional hours actually in school settings. As a result, they may not have been as informed on national education developments and needs in the past as will be required with anticipated greater state control, as, for example, in the context of school-based services.

A related development in education, the emphasis on early childhood education as well as parent–child development programs, has been of particular interest to psychologists for some years (e.g., Holtzman, 1992; Johnson, 1988; Johnson & Walker, 1987; Walker, Rodriguez, Johnson, & Cortez, 1995, Zigler, 1989; Zigler & Lag, 1991). A good deal of effort has been expended in implementing these programs in high-poverty areas, and both educators and psychologists have been encouraged by initial as well as subsequent reports indicating positive outcomes with respect to both cognitive and emotional development, for example, managing behavioral problems. More recent information, however, suggests that these promising outcomes are not maintained over time unless other kinds of interventions are included as well (D. L. Johnson, personal communication, July, 1996; Johnson & Blumenthal, in press). Specifically, positive gains seem more likely to persist when mothers of the children enrolled in the early education programs are supported through corollary education and training programs. On completion of the two-year Avance program in San Antonio, for example, more mothers went to work, and it is anticipated that this outcome will lead to better economic futures for the families which will, in turn, support a longer-term maintenance of the positive behavioral and academic gains realized by the children in the program (Walker et al., 1995). Results such as these suggest exciting possibilities for psychologists and educators interested in integrating and implementing early childhood education and parent–child development programs. Clinical psychology has much to offer in such efforts to advance early childhood education and development opportunities, and it can be a major player in assuring that these gains, which otherwise might remain only temporary, are maintained.

Other recent and mounting concerns, also related to the emotional needs of K-12 students, have been identified by Claire Ellen Weinstein, then President of the American Psychological Association's Division 15 (Educational Psychology), concerns supported by Gutek (1991) about persistent issues and challenges facing American education. While acknowledging the importance of *Goals 2000*, Weinstein extends the focus she believes educators and psychologists must pursue by identifying three key issues related to the current emphasis on the personal and social development of students: transitions, literacy, and self-system variables (C. E. Weinstein, personal communication, July, 1996). Enthusiastic in the hope that psychologists will more fully address these critical issues in the future, Weinstein stresses the necessity for collaboration between clinical psychologists and both educational and school psychologists. Such urging seems particularly valid, given: (i) the striking demographic and cultural changes taking place in our communities which are impacting our schools; and (ii) the resulting need for integrated responses and services.

Self-system variables are considered here briefly in respect of space limitations; they are selected from the three noted by Weinstein to emphasize the centrality of personal and social adjustment issues to the learning process. Moreover, not only is a focus on self-system variables particularly congruent with clinical psychology's domain, it is also especially relevant to understanding and addressing social and cultural challenges growing out of the increasing ethnic and cultural diversity in both urban and rural populations. Educators are experiencing great need for psychological understanding as they attempt to respond in school settings to the rapidly increasing, nationwide diversity. As children from diverse backgrounds and cultures enter local schools, speaking different languages and reflecting vastly different family as well as cultural values—for example, regarding such basic matters as parenting, religion, gender roles—challenges to educators are impressive. Administrators and teachers must have support in understanding, appreciating, and working in the context of self-system variables as they are reflected by students as well as their families and communities. Equally important, educators must have opportunities to consider how their own values relate to dramatically changing systems.

Student and family beliefs, expectations, and goals, for example, not only determine significantly how students perceive their school environments but also how well they will perform. How such views fit within specific community as well as administrative and teacher cultures can be a major determining factor regarding if and how students become and remain engaged in learning. As clinical psychologists work with families and children experiencing school-related problems, they are likely to find their understanding and interventions informed by determining the nature and fit of self-system variables that are obtained in a specific situation. From a school-based perspective, as psychologists gain greater appreciation for these kinds of interactions, they are likely to have opportunities for supporting an effective rapprochement among students, their families, and their teachers. An integrated and interactive approach on the parts of educators and psychologists seems congruent with Holtzman's view that psychologists can be "in the forefront of bringing about the necessary changes and scientific advancement" for education reform to succeed (Holtzman, 1992, p. xiv).

2.39.2.3 Work World and Economics

Radical changes in the world of work are impacting every family and community unit in our society in one way or another. Not only has the composition of professionals and workers changed significantly with respect to diversity (e.g., gender and ethnicity), the nature of work itself has changed remarkably. Technological advances alone, for example, have made for sea changes in the work setting. Computers and other digital products are used by 50 million workers every day (Cascio, 1995). A more gradual but nonetheless dramatic change in workforce composition is the increasing number of women working outside the home. Not only are more women working, more women are in the professional arena and more are working full-time. The increase of mothers in the workforce increased from 57% in 1980 to 68% in 1994 (Hartmann, Spalter-Roth, & Chu, 1996).

These workforce shifts contribute to major changes not only in the work place but also among families and communities. Included in the national-scale demands created by such changes are child and family care needs; health and safety issues in the work place, both physical and psychological; and job satisfaction and fulfillment, free of excessive stress. Changing standards and expectations on acceptable working conditions and settings as well as expected work place behaviors and relationships are leading to new and interesting developments. Harassment issues, for example, were almost unheard in the 1980s; yet, they are so much in the forefront today that alert human resources offices are conversant in the legal as well as everyday practical implications of such matters. Quite different from expectations a few decades ago, most workers today seek benefits for themselves and their families, and they expect protection from violence and injury in the work setting. Stress in the work place is another pervasive work problem, and clearly it is one to which clinical psychologists can bring considerable relevant research and applied expertise, as are the other critical work-related matters just noted.

Still another development, a key economic matter which has significant psychological ramifications, involves the changing nature and availability of jobs and raises questions about job security, thus, about family security. The fairly recent transition from an industrial society to a technologically driven one has been revolutionary with both positive and negative outcomes. Workers at all levels without computer skills are likely to find themselves unable to gain employment in many future work situations. On the one hand, business can be conducted more quickly and with greater ease as a result of technological advances, and more complex tasks can be completed and integrated than can be accomplished manually. On the other hand, companies and agencies, able to replace people with computers and other technological advances including mechanical operations, are downsizing, one result of which is that individuals who had expected lifetime job security are finding themselves among the unemployed. Tying downsizing to several other factors as well, including efforts to improve productivity, location moves, mergers, and new ways of organizing work, Cascio (1995) notes that seven million layoffs have occurred since 1987, and asserts that downward mobility is now the rule, not the exception.

What about the already unemployed? As expected welfare system changes are initiated in the next few years, communities will be faced with how the currently unemployed, including single mothers, will find employment providing them with a necessary income. Among the disadvantaged living in the USA, the largest group comprises single-mother families. They constitute 60% of all poor families with children, and those living in these families constitute 22% of all poor people (Hartman et al., 1996); or, put another way, in 1994: "44 percent of single-mother families were poor, compared with 8 percent of married-couple families with children and 12 percent of people over age 65" (p. 24). Clinical psychologists, along with others, such as the Institute for Women's Policy Research, surely are challenged to study factors contributing to this regrettable status and develop strategies for assuring upward change.

Based on studies investigating job availability in low-income areas, questions have been raised regarding the assumption that stopping state support will force welfare recipients to "join the labor market and support their children with honest jobs" (Newman, 1996, p. 20). Results indicate that the going may be tough, especially in geographical and target areas that will start from a low point—as in one inner-city neighborhood study, for example, where there was 18% unemployment with 40% below the poverty line and where new jobs were in short supply. One specific finding noted that "the ratio of jobseekers to successful hires in the fast-food restaurants studied in Harlem is approximately 14 to 1; moreover, among those rejected, 73% had not found work of any kind a year later" (p. 21). Those groups most frequently rejected were: younger applicants, especially 18 years and under; African-Americans, as compared with Latinos and others; residents of the neighborhoods, as compared with applicants commuting to work from other neighborhoods (a complicating factor for young mothers who need to work near their children); those without "connections," that is, those outside of friend and family networks; native-born applicants, as compared with legal immigrants; and those without prior work experience.

Such findings underscore the importance of developing and implementing programs to address job training and also those to create new jobs. The former need relates, interestingly enough, to reports on parent–child development programs (e.g., Walker et al., 1995). Although oftentimes in the past, such efforts have been viewed as largely the domain of private industry and business, or of government, it is clear that the development of new approaches informed by psychological understanding is a major economic priority for this country. Among some current views of welfare mothers are depictions of them as "welfare queens," as pathological, and as lazy, whereas others look to socio-cultural problems, including increasing numbers of teenage pregnancies, high divorce rate, lack of education and job training, as the major contributors (Newman, 1996). The resolution of this complex cluster of social and economic problems certainly could be aided by research as well as other expertise from clinical psychology. Moreover, in the context of broad collaboration with other scientists and professionals working closely with local communities as well as state and national policymakers, clinical psychologists could contribute significantly to new and better integrated approaches.

Such national work-related problems have vast implications for clinical psychologists, as they do for the general public. Clearly, considering such factors as emotional satisfaction, stress management, and the promotion of mastery and productivity, there are a number of national priorities which clinical psychology can appropriately address. There exist a number of opportunities involving collaboration with industrial and organizational psychologists, as supported by Cascio (1995): there is "an exciting agenda with large potential payoffs for individuals, organizations, and society as psychology moves into the 21st century" (1995, p. 938).

2.39.2.4 Families and Communities: Relationships, Systems, and Communication

That a central theme of both 1996 national party conventions, Republicans and Democrats, reflected the strong concern of Americans about the health, wealth, welfare, and happiness of the American nation's families and communities appears to be no accident or mere coincidence. Indeed, individuals across the USA openly express a yearning for greater individual opportunity but in the context of healthy and meaningful relationships, particularly those associated with family and community. They seek safe and supportive environments, ones that allow them to pursue their personal, family, and community goals. Clinical psychology is in an especially advantageous position for studying not only the kinds of changes that would respond to this yearning but also interventions which might best facilitate such changes. Outcomes from such programs as the Avance project, for example, are yielding provocative and exciting possibilities for bringing about changes in our communities (Walker et al., 1995). Clinical psychologists are challenged to work, not only with educators and collaboratively with other psychologists from various specialty areas, but also with local communities as well as state and national groups, to develop programs that better support children and their families. They can also serve as advisors to state and national legislators who are struggling to gain relevant, valid information about individuals, families, and communities in their various social, educational, and economic systems that would enable them to make better decisions in all public arenas on behalf of their constituents. Clinical psychologists must continue to be at the ready to provide information and counsel on matters of national priority to families and children, responding with critical leadership as they have in the past on such issues as teenage pregnancies and the negative influence of television violence on youth (Wilcox & Kunkel, in press).

Among the many family and community issues viewed as national priorities by leaders and policymakers as well as the general public, there are three that seem to be among the most pressing: violence, increasing diversity across the nation, and the social and economic problems associated with situations of disadvantage, the latter including poverty and the related high incidence of drug abuse. Again, given space considerations, violence is considered here, as an exemplar of the three, because of the particularly urgent need for relief through effective interventions, and also because, as a matter of national crisis, it has compelling implications for the future work of clinical psychologists.

Arguably, violence is one of the single greatest problems facing the USA and the world today. Possibly the USA's most devastating social and human problem, it strikes and is felt in every context of life: families, schools at all levels (preschool through 12 and on into higher education), the workplace, and in communities. Indeed, because the cost of violent events and actions is so great and the damage so extensive, the prevention of injury is an identified priority for federal and state public health offices. The Center for Disease Control and Prevention in Atlanta, part of the United States Public Health Service, has as one of its major priorities the development of further understanding and strategies for addressing the prevention of intentional and unintentional injuries. Recent violent and highly visible events, such as the April 1995 Oklahoma City bombing of the Alfred P. Murrah Federal Building, the 1995 Simpson trial, the June 1996 Saudi Arabian bombing of the American Marines' housing, and the July 1996 TWA plane crash in New York, have received wide media coverage, drawing attention to the pervasiveness of violence, or suspected violence, in American life. The visibility of these events demonstrates dramatically the destruction and pain that individual dissidents, those who are deeply psychologically disturbed, alienated groups, and others can inflict on families and communities.

Studies on the incidence and patterns of domestic violence indicate that at least half of the injuries sustained by women are related to partner abuse, and in some age groups, more than half of all reported rapes are committed by partners. Women are much more likely to be abused than men. Even though women may have reported injuries multiple times to their physicians and in emergency rooms, the violent causes of their repeated injuries may not be discerned. Recent litigation in several states involved violence to women in the workplace, from sexual harassment to alleged murder.

Every day in America: three children die from child abuse; 15 children die from gunshots, 135 000 children bring guns to school; and 1.2 million latchkey children come home to a house where there is a gun. Recent analyses show that in addition to the high price in psychological anguish paid by individuals, families, and communities as a result of violent behaviors and acts, physical injuries suffered by the victims of violence are increasing in number. The numbers are so staggering that the American Medical Association construes epidemic violence as a leading cause of death. Prevention of injury is an identified priority for federal and state public health offices.

Clinical psychologists are poised advantageously to address violence issues. Not only can they conduct requisite research to better understand the etiology and course of violent behaviors and acts in different contexts, they are already leaders actively involved in the development and implementation of effective interventions (e.g., Cohen & Mannarino 1996). Using their understanding about human motivations and conditions that lead to violence, they are poised to develop strategies and interventions to ameliorate, relieve, and dissipate individual and community violence. Clinical psychologists have a particular responsibility to work in interdisciplinary programs focusing on the development and implementation as well as the evaluation of such strategies and interventions, including interdisciplinary training programs (e.g., see Gallmeier & Bonner, 1992).

2.39.3 SUMMARY

Hubert Humphrey, Vice President of the USA from 1965 to 1969, once said:

> The moral test of government is how it treats those who are in the dawn of life—the children; those who are in the twilight of life—the aged; and those who are in the shadows of life—the sick, the needy, and the handicapped. (quoted in Edge & Groves, 1994, p. 2)

The subtitle of a 1996 publication titled *Psychology and public policy* (Lorion, Iscoe, DeLeon, & VandenBos, 1996), places this quote in an appropriate context with respect to clinical psychology's responsible and effective response to national priorities: "Balancing Public Service and Professional Need." In the foreword to that publication, Fowler says that "psychologists do not operate in a social vacuum" (p.ix), and he goes on to assert that psychologists' "theories must recognize social realities, their methods must allow for differing attitudes and behaviors, and their treatment must respond to socially defined psychological and behavioral problems" (Fowler, 1996, p. ix).

These authors identify a reciprocity between the importance of policy to the profession and that of the profession to policy formulation. As described, this reciprocity makes explicit the imperative that psychologists engage in national policymaking as well as in professional activities directly addressing national needs. Clinical psychologists, of course, are challenged accordingly, if not especially, to engage in clinical research relevant to national priorities and in applied contributions responsive to the country's most pressing social and human problems.

This chapter's focus serves that charge by presenting and elucidating four broad areas reflecting current national priorities and, then, by noting clinical psychology's relevance to, and encouraging its engagement with, them. With an impressive history in this regard and great promise for addressing human need, welfare, and happiness, clinical psychology should do no less.

2.39.4 REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Barlow, D. H., Hayes, S. C., & Nelson, R. O. (1984). *The scientist practitioner: Research and accountability in clinical and educational settings.* New York: Pergamon.

Callan, J. E. (1993). APA advocacy eyes scant federal dollars. *Monitor.* Washington, DC: American Psychological Association.

Cascio, W. F. (1995). Whither industrial and organizational psychology in a changing world of work? *American Psychologist, 50*(11), 928–939.

Cremen, L. A. (1988). *American education: The metropolitan experience—1876–1980.* New York: Harper & Row.

Edge, R., & Groves, J. (1994). *The ethics of health care: A guide for clinical practice.* Albany, NY: Delmar.

Cohen, J., & Mannarino, A. (1996). A treatment outcome study for sexually abused preschool children: Initial findings. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*(1), 42–50.

Fowler, R. D. (1996). Foreword. In R. Lorion, I. Iscoe, P. DeLeon, & G. VandenBos (Eds.), *Psychology and public policy* (pp. ix–xiv). Washington, DC: American Psychological Association.

Gallmeier, T., & Bonner, B. (1992). University-based interdisciplinary training in child abuse and neglect. *Child Abuse and Neglect, 16,* 513–521.

Greenberg, P., Steglier, L., Finkelstein, S., & Berndt, E. (1993). The economic burden of depression in 1990. *Journal of Clinical Psychology, 54,* 405–418.

Gutek, G. (1991). *An historical introduction to American education* (2nd ed.). Prospect Heights, IL: Waveland Press.

Hartmann, H., Spalter-Roth, R., & Chu, J. (1996). Poverty alleviation and single-mother families. *National Forum, 76*(13), 24–27.

Holtzman, W. (1992). *School of the future.* Washington, DC: American Psychological Association.

James, W. (1952). *The principles of psychology.* (1890). A *Great books of the Western World 53* publication. Chicago: Encyclopaedia Britannica.

Johnson, B. D., & Muffler, J. (1992). Sociocultural aspects of drug use and abuse in the 1990's. In J. H. Lowinson, P. Ruiz, R. B. Millman, & J. F. Langrod (Eds.), *Substance abuse: A comprehensive textbook* (2nd ed. pp. 118–137). Baltimore: Williams and Wilkins.

Johnson, D. L. (1988). Primary prevention of behavior problems in young children: The Houston parent–child development center. In R. Lorion & J. Ramos-McKay (Eds.), *14 ounces of prevention: A casebook for practitioners* (pp. 44–52). Washington, DC: American Psychological Association.

Johnson, D. L., & Blumenthal, J. (in press). Final report of the parent–child development center follow-up project (unpublished manuscript, received from author July 1996).

Johnson, D. L., & Walker, T. (1987). Primary prevention of behavior problems in Mexican–American children. *American Journal of Community Psycology, 15*(4), 375–385.

Korchin, S., Jr. (1976). *Modern clinical psychology.* New York: Basic Books.

Lazar, S. G. (1997). Epidemiology of mental illness in the United States: An overview of the cost effectiveness of psychotherapy for certain patient populations. In S. G. Lazar (Ed.), *1997 Psychoanalytic Inquiry Supplement (Extended dynamic psychotherapy: Making the case in an era of managed care).*

Lorion, R. (1991). Prevention and public health: Psychology's response to the nation's health care crisis. *American Psychologist, 46*(5), 516–519.

Lorion, R., Iscoe, I., DeLeon, P., & VandenBos, G. (Eds.) (1996). *Psychology and public policy.* Washington, DC: American Psychological Association.

Lowinger, P. (1992) Drug abuse: economic and political basis. In J. H. Lowinson, P. Ruiz, R. B. Millman, & J. G. Langrod (Eds.), *Substance abuse: A comprehensive textbook* (2nd ed., pp. 138–142). Baltimore: Williams and Wilkins.

National Institute on Drug Abuse (1990). *National household survey on drug abuse: Main findings 1988.* Rockville, MD: National Institute on Drug Abuse.

Newman, K. S. (1996). Job availability: Achilles heel of welfare reform. *National Forum, 76*(3), 20–23.

Nirenberg, T. D. (1983). Treatment of substance abuse. In C. E. Walker (Ed.), *The handbook of clinical psychology* (Vol. II, pp. 633–665). Homewood, IL: Dow Jones-Irwin (Dorsey Professional Issues).

Raimy, V. C. (Ed.) (1950). *Training in clinical psychology* New York: Prentice-Hall.

Resnick, R., & DeLeon, P. (1995). News from Washington, DC, the future of health care reform: Implications of 1994 elections (1995). *Professional psychology: Research and Practice, 26*(1), 3–4.

Rice, D., Kelman, S., Miller, L., & Dunmeyer, S. (1990). *Report on the economic costs of alcohol and drug abuse and mental illness: 1985.* US Department of Health and Human Services Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration.

Rotter, J. B. (1963). A historical and theoretical analysis of some broad trends in clinical psychology. In S. Koch (Ed.), *Psychology, a study of a science* (Vol. 5, pp. 780–830). New York: McGraw-Hill.

Shakow, D. (1969). *Clinical psychology as science and profession.* Chicago: Aldine.

Slosar, J. R., & Lettieri, R. (1996). Financing mechanisms in the delivery of mental health care services: Widgets or wisdom? *Psychologist Psychoanalyst, XVI*(1), 11–16.

US National Commission on Excellence in Education (1983). *A nation at risk.* Washington, DC: US National Commission on Excellence in Education.

Walker, T. B., Rodriguez, G. G., Johnson, D. L., & Cortez, C. P. (1995). Avance parent–child education program. In S. Smith (Ed.), *Two generation program for families in poverty: A new strategy* (pp. 67–90). Norwood, NJ: Ablex.

Watson, R. I. (1953). A brief history of clinical psychology. *The Psychological Bulletin, 50,* 321–346.

Wilcox, B. & Kunkel, D. (in press). Taking television seriously: Children and television policy. In E. Zigler, S. L. Kagan, & N. W. Hell (Eds.), *Children, families and government: Preparing for the 21st Century*. New York: Cambridge University Press.

Winick, C. (1992). Epidemiology of alcohol and drug use. In J. H. Lowinson, P. Ruiz, R. B. Millman, & J. G. Langrod (Eds.), *Substance abuse: A comprehensive textbook* (2nd ed., pp. 15–29). Baltimore: Williams and Wilkins.

Zigler, E. F. (1989). Addressing the nation's child care crisis: The school of the 21st century. *American Journal of Orthopsychiatry, 59,* 485–491.

Zigler, E. F., & Lang, M. (1991). *Child care choices: Balancing the needs of children, families, and society*. New York: Free Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.40

National Priority Needs: Implications for Clinical Psychologists in Japan

YOSHIKO SHIMONAKA and KATSUHARU NAKAZATO
Tokyo Metropolitan Institute of Gerontology, Japan

2.40.1 NEEDS OF OLDER ADULTS IN JAPANESE SOCIETY 471
  2.40.1.1 Clinical Psychological Services Sought in Japan 473
2.40.2 SPECIFIC PSYCHOLOGICAL PROBLEMS IN JAPANESE OLDER ADULTS 473
  2.40.2.1 Psychological Disturbance in Old Age 473
  2.40.2.2 Suicide Among the Elderly: Aspects of Suicide Unique to Japan 474
  2.40.2.3 Elderly Abuse: Abuse and Neglect 475
2.40.3 PSYCHOTHERAPEUTIC APPROACHES FOR THE ELDERLY IN CLINICAL GEROPSYCHOLOGY 475
2.40.4 REFERENCES 478

2.40.1 NEEDS OF OLDER ADULTS IN JAPANESE SOCIETY

By 1995, life expectancy in Japan reached 76.5 years for men and 82.98 years for women. Also by 1995, Japanese life expectancy had been the highest in the world for 10 consecutive years for men and 11 consecutive years for women. This trend will continue.

Japan joined the list of nations with great longevity in the 1970s. In the years before World War II, Japan was a developing country, lagging far behind Western countries in terms of life expectancy. In 1935–1936, life expectancies for Japanese men and women were 46.92 years and 49.63 years, respectively, while life expectancies for both men and women of other countries such as Holland and Sweden were over 60 years.

According to World Health Organization (WHO) statistics, a nation is categorized as having an aging society if the ratio of people aged 65 and older exceeds 7% of the entire population. Figure 1 (Japanese Ministry of Health and Welfare, 1994) shows changes in the ratio of people aged 65 and older in major countries between 1950 and 2025. In Japan, the ratio became 7% in 1970 and it took only 25 years for the figure to double, reaching 14% in 1995. Furthermore, it is evident that the aging of Japanese society is proceeding at a drastic speed when compared to an 85 year period for Sweden (7% in 1890 and 14% in 1975), a 70 year period for the USA (7% in 1945 and an estimated figure of 14% in 2015), a 50 year period for the UK (7% in 1930 and 14% in 1980), and a 45 year period for Germany (7% in 1930 and 14% in 1975).

It is also noteworthy that the population of those aged 65 and older in Japan will continue to increase at a rate that will not be observed in any other country, and will reach an estimated figure of 21.3%, the highest in the world, in 2010. An ultra-aging society is expected in Japan by 2020.

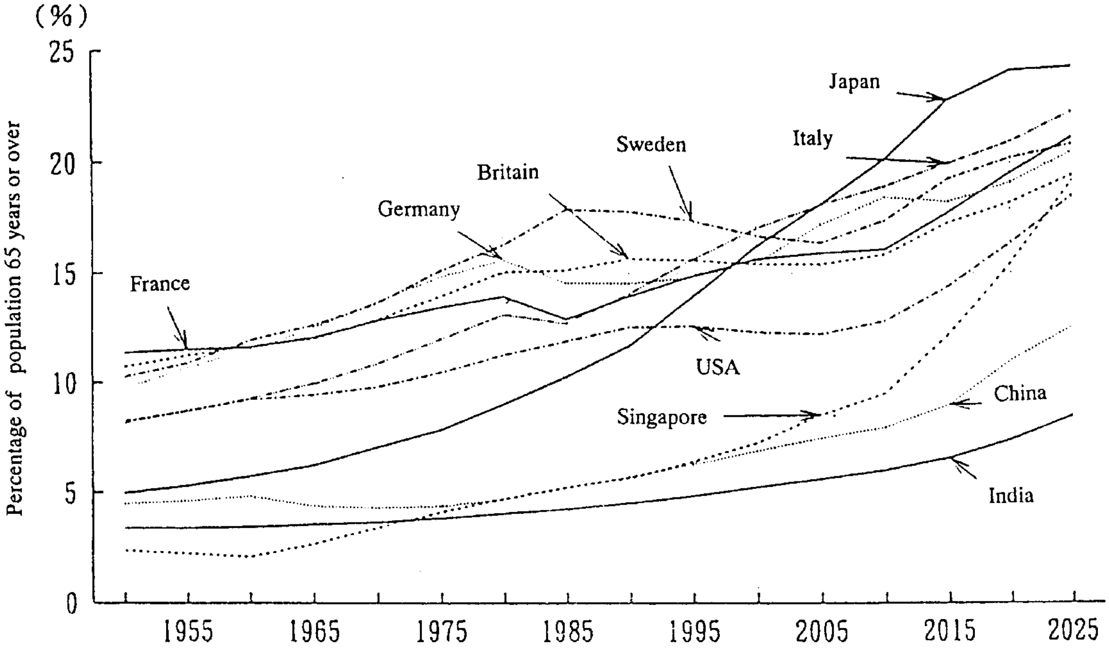

Figure 1 Changes in ratio of people 65 and over in Japan and other countries.

Now that Japan is facing an aging society, various issues concerning the aged are surfacing. One of the important issues is the change in family structure. The traditional Japanese family system (three generations in one household) started to break down after World War II, and "the nuclear family," which had long been common in Western countries, began to take over. The ratio of those aged 65 or older living with a child was 68% in 1975. The figure had dropped to 55.3% by 1994. This figure, however, is still high compared to such Western countries as the USA, the UK, and Sweden.

The ratio of those aged 60 and over who prefer to live with a child or grandchild is 53.6% in Japan, quite a high figure compared to that of Germany (15.4%), the USA (3.4%), and the UK (3.9%) (Japanese General Affairs Agency, 1992). Despite the fact there are more elderly in Japan who live with their family and who prefer to do so than in other countries, the ratio of those who felt lonely was over 50%, much higher than in the USA and the UK. This suggests that there is a wide gap between the expectations Japanese elderly people have of their family and the reality of their daily life. The same can be said about how they feel about work after retirement.

As to how Japanese elderly people feel about having a postretirement job, 35.9% replied that they would "work as long as they are healthy" and 22.7% replied that they would "work until they were about 70." These two figures together show that more than half the population of elderly people in Japan are willing to work (Japanese General Affairs Agency, 1995). In reality, however, the employment rate of those aged 65 years and older had dropped from 38.1% in 1962 to 24.3% in 1994.

The most common motivation for those who wanted to work was "the desire to make friends." The motivations given in other countries such as the USA and the UK were "earning money" or "enjoying work." This difference indicates that elderly people in Japan have little interaction with their community and reveals the fact that they are living with a feeling of loneliness.

Another difficult issue regarding the elderly in Japan in the 1990s is that of frail or demented elderly people. In 1990, there were approximately 1 million cases of senile dementia nationwide, and this figure is expected to climb to 1.5 million by the year 2000, and 2.47 million by 2015. The number of elderly people who are frail or bed-ridden was 810 000 in 1990 and it is estimated that this figure will triple to 2.29 million by 2025 (Japanese Ministry of Health and Welfare, 1992). At present, there are 15 caregivers (careworkers or family members) to one elderly person. By 2025, however, there will be only two caregivers for one elderly person, a very severe reduction.

According to a survey conducted by the Ministry of General Affairs in 1995, 41.5% of the elderly aged 65 and older wish to receive care at home, and 54.8% of those expect their family members to assume this task. But, since 1985, the number of people in the younger generation who consider looking after their aged parents an obligation has been decreasing drastically, and they tend to consider living with their parents and supporting their parents as separate matters (Japanese Ministry of Health and Welfare, 1996).

Despite the fact that this gap between the expectations of the elderly and the consciousness of the younger generation has started to become clear, in reality, one out of four people aged 65 and older who died had received care from their family members while bedridden for more than half a year. Female members of the family are more likely to assume the responsibility of taking care of elderly family members than men in the following order: the elderly person's wife (31.6%), the wife of the eldest son (27.6%), and the eldest daughter (15.5%). Only 5.0% of the elderly persons' husbands assume such responsibilities. It is reported that more than half of those who assumed these responsibilities had jobs and 20.6% of them had given up their work to take care of an elderly family member (Japanese Ministry of Health and Welfare, 1996).

The responsibility of caring for the elderly increases the fatigue and stress of the responsible family member. There have been reports in newspapers, albeit few, regarding elderly abuse and murder of old people. Abuse and murder of the elderly appear to be associated with their caregivers' added fatigue and stress.

2.40.1.1 Clinical Psychological Services Sought in Japan

Since the Japanese Association for Certification of Clinical Psychologists was established in 1989, the number of certified clinical psychologists has been increasing year by year. In 1997, more than 5000 clinical psychologists are working at various institutions or in private practice. The Association of Japanese Clinical Psychology has more than 7000 members, and the number of research papers presented at its annual congress has also been increasing. The Education and Training Committee of the Association of Japanese Clinical Psychology conducted a survey on the content of the clinical psychological services provided by its 2296 new members for a period of six years from 1986 to 1991 (Association of Japanese Clinical Psychology, 1993). Data were obtained by multiple choice responses. 53.6% of the clinical psychologists surveyed were full-time workers, their most common workplaces being hospitals (37.9%), followed by counseling centers in universities and schools (15.7%), and child guidance clinics (15.3%). Almost 70% worked at the aforementioned three workplaces. Other workplaces included welfare facilities for rehabilitation (6.3%), court (6.8%), private clinics (4.8%), research institutions (3.5%), mental health centers and health centers (3.3%), the Ministry of Justice (1.9%), and others (4.5%). 22.3% of the clinical psychologists surveyed were part-time workers.

The ratio of part-time workers working for hospitals was 24.2%, which was lower than that of full-time workers. However, the ratios of part-time workers working for child guidance clinics (29.0%), private clinics (12.4%), and mental health centers and health centers (11.8%) were higher than those of full-time workers.

Regarding the contents of the clinical psychologists' work, psychological therapy was the most common (70.6%), followed by psychological assessments (52.1%), interviews and consultations (50.1%), supervision and training (9.1%), and others (6.7%).

Clients for the services of clinical psychology can be categorized into the following age groups: children and youths (58.7%), middle aged people (55.0%), college students and working youths (41.7%), infants (39.6%), and the elderly (18.4%). Given the large number of the elderly people in Japan, 18.4% is a rather small figure for clinical psychology services, especially when we consider the fact that about 7% of those aged 65 and older are reported to have mental disorders.

Bullying in Japanese schools is a serious problem which has come to light. Due to the increasing number of cases of bullying, a school counseling system has been planned and discussed since 1995 for future implementation by the Ministry of Education. Appreciation for clinical psychology has started to permeate Japanese society owing to this problem. Clinical psychologists have also started tackling the pressing issue of HIV counseling. Furthermore, systems for crisis intervention such as the "Lifeline" system have gradually been put in order. It must, however, be pointed out that only a few elderly are taking advantage of these services. This is because they tend to think that revealing a personal psychological problem to a stranger is shameful. Also, they are not used to paying money for such intangible services as counseling.

2.40.2 SPECIFIC PSYCHOLOGICAL PROBLEMS IN JAPANESE OLDER ADULTS

2.40.2.1 Psychological Disturbance in Old Age

The Ministry of Health and Welfare conducted a survey on outpatients according to the *ICD-9* (*International statistical classification of* *diseases and related health problems-ninth revision*) in 1993 (Japanese Ministry of Health and Welfare, 1995).

The highest number of older adults sought treatment for senile/presenile organic psychosis (senile dementia) (43%), followed by neurosis (33%), and manic-depressive psychosis (26%). There is a tendency to consider depression as a more serious problem than neurosis among the aged, however it is evident from the above figures that the threat of neurosis should not be ignored (Japanese Ministry of Health and Welfare, 1995).

Figure 2 shows trends of outpatients by age group suffering from neurosis and manic-depressive psychosis. The curves for neurosis show a difference between the numbers of male and female outpatients receiving treatment. The peak for women is in the 70–74 age group, but not much difference among age groups exists in the men's curve. Little difference can be seen between the curves for men and women for manic-depressive psychosis, and the peak for both is in the 70–74 age group.

The results of this survey suggest that there is more chance of suffering from psychological disorders during the period from middle age to old age than at any other time in life. However, clinical services for these people are lagging behind the services available to younger people.

2.40.2.2 Suicide Among the Elderly: Aspects of Suicide Unique to Japan

Figure 3 (Association for Health and Welfare Statistics, 1994) shows trends of the suicide rate

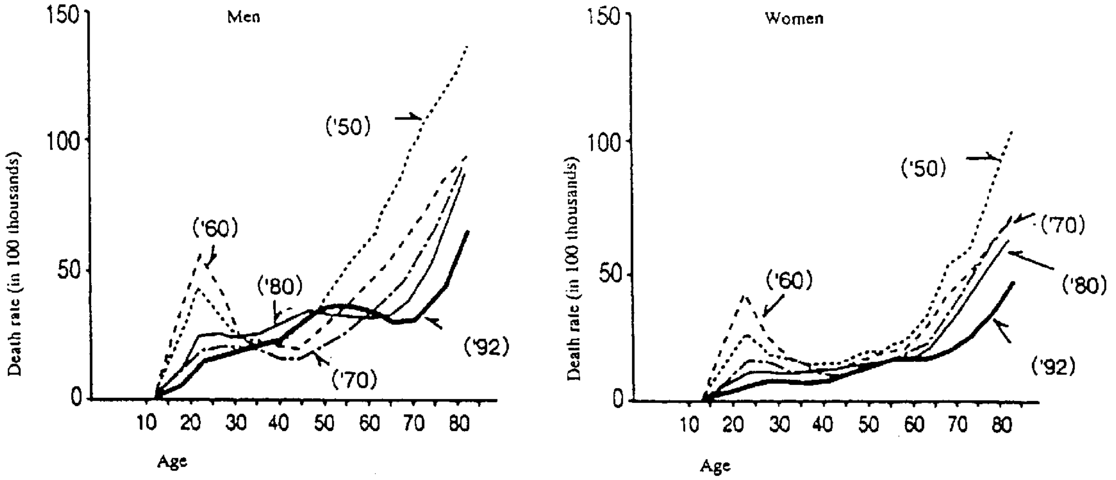

Changes in Suicide Rate in Japan

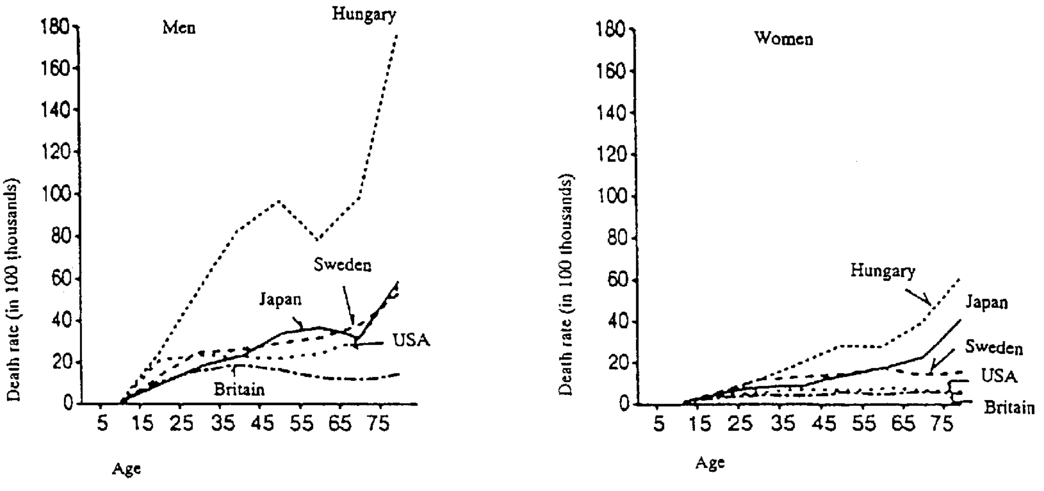

Cross-national Comparison of Suicide Rate

Figure 2 Suicide rate of elderly people: changes in Japan and cross-national comparison.

in different years by age group in Japan as compared to other countries. Although the suicide rate of people aged 65 years and older is declining year by year, it is still the highest among all age groups. The graphs comparing suicide rates of people aged 65 and older in different countries show that Hungary has the highest rate for both men and women, and that Japan has the second highest for both men and women. The leading cause of suicide among the elderly is illness (70%), followed by family problems, alcohol dependency, and psychological disturbance. Few suicides result from financial difficulties. This suggests that many older adults are suffering from illness or facing problems with their family while they are living in relative affluence. Older adults commit suicide because they "don't want to be a nuisance to the family." Committing suicide is not considered a sin in the religious views of most Japanese. The same can be said of double suicides of married couples which are discussed in this section.

What is considered a double suicide of a married couple is not the same in Japan as it is the West. Husband and wife mutually agree to commit suicide, or, in some cases, one spouse kills the other and then commits suicide. It must be understood here that Japanese people do not consider the murder/suicide of a married couple to be any different from a double suicide. Unlike the prevailing views in Western culture, this kind of murder/suicide is classified along with double suicide due to some profound cultural differences between Japan and the West. Older couples in Japan do not consider their spouses as individual persons, but rather as a part of themselves. Together, they are an inseparable whole, or "one flesh." This idea stems from the traditional Japanese concept of the extended family system. Older Japanese couples try to solve their problems by themselves, without outside help. When considering suicide, they tend to think, "It is better for my wife (or husband) to die with me, rather than for me to leave her (or him) alone to suffer." This attitude, while common in Japan, may be difficult for Westerners to understand.

Although cases have been few, murder of elderly people by a family member caregiver has become a social problem. In 1992, for example, there were 20 such cases, 17 of the victims being women. The murders were committed by husband or wife (9 cases), son or daughter (10 cases) and daughter in law (1 case) (Takeda, 1994). In most of the murder cases, the task of caring for the elderly had been done entirely within the family without the help of others or public service organizations, and thus, the caregiver's mental and physical fatigue had reached its limit.

2.40.2.3 Elderly Abuse: Abuse and Neglect

There is little substantiated data on elderly abuse in Japan due to the fact that people have been reluctant to discuss the issue. Tanaka (1994) surveyed 400 At-home Elderly Care Support Centers throughout Japan. These centers act as coordinators between families or older adults and public service organizations. The survey was conducted via mail and 220 replies were received (55%). Seventy-two of the respondants reported elderly abuse as one of their items for consultation.

What seems characteristic in Japan is that if elderly abuse is divided into abuse and neglect, there is more of the latter. The ratio of abuse, as opposed to neglect, was 75.1% in New York, 84.4% in London and 43.1% in Japan nationwide. This means that neglect made up more than half the cases of elderly abuse in Japan. In common with Western countries, more women are abused than men and the abuse rate increases with age. It is unique to Japan that most cases of abuse are committed by family members, and also that the abuse rate increases as the caregiver's age increases. The main cause of abuse is the fatigue associated with caregiving.

2.40.3 PSYCHOTHERAPEUTIC APPROACHES FOR THE ELDERLY IN CLINICAL GEROPSYCHOLOGY

Although the number of research papers presented at the Congress of the Association of Clinical Psychology has been increasing, much less research has been done with elderly people. This is due to the fact that gerontology has only recently been introduced and research into gerontological clinical psychology has not yet been established as a field in Japan.

The first and only gerontological research center in Japan, the "Tokyo Metropolitan Institute of Gerontology," was established in 1972, just before the population of people aged 65 years and older reached 7%. This is when research into the psychological characteristics of the elderly first began. Shimonaka and Nakazato (1980, 1986) compared personality characteristics of elderly people with those of young and middle-aged adults using a sentence completion test. With respect to family relationships, elderly people gave more positive responses to family images. As for self-perception, elderly people perceived their past and present selves as being more positive while the younger

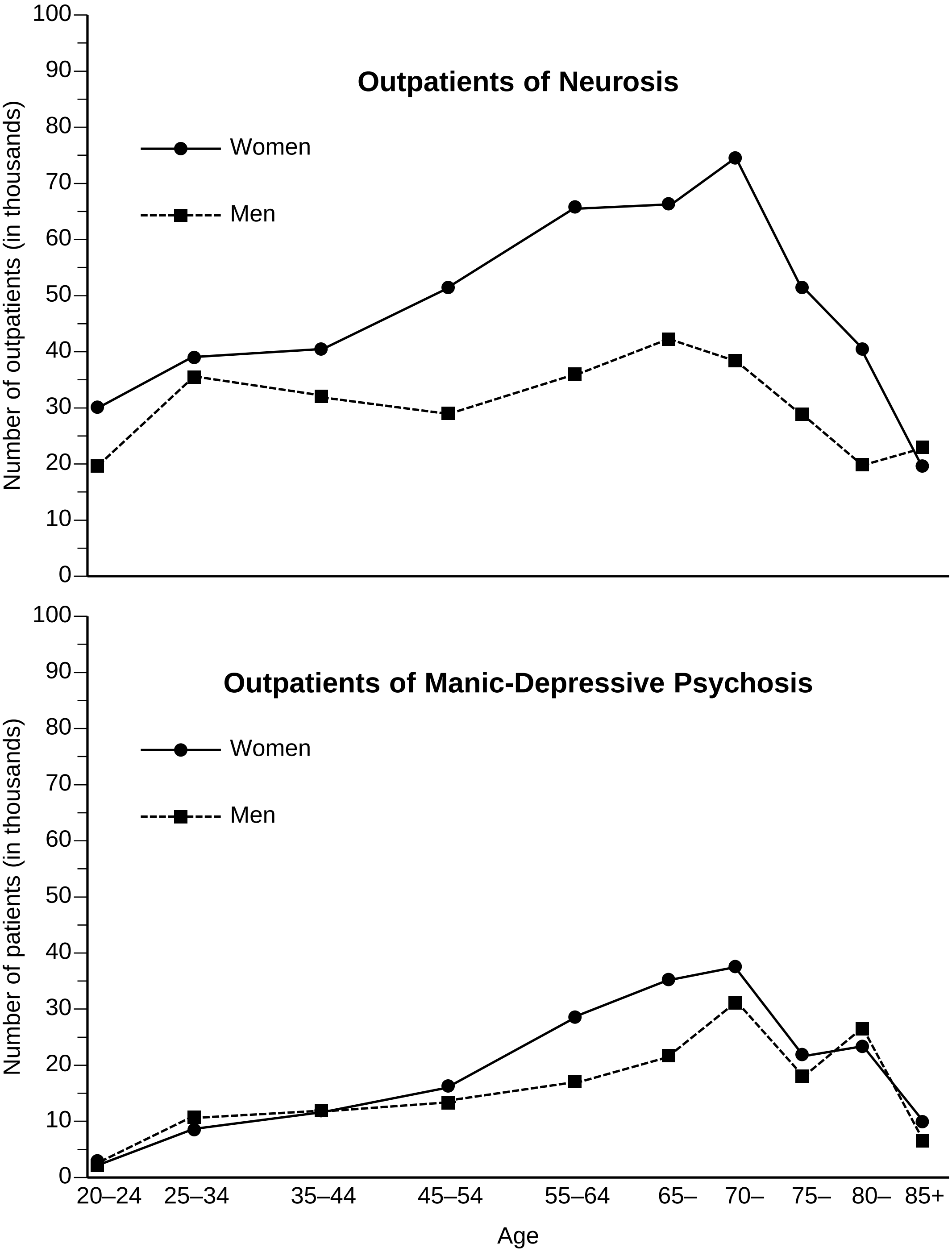

Figure 3 Number of outpatients of neurosis and manic-depressive psychosis according to age group and sex.

generations perceived their future selves as being more positive. However, future self-perception for elderly people was limited by the ultimate finality of death. More young adults viewed death with negative feelings such as fear, sorrow, or distress than did the older adults, who saw death in a neutral manner, as destiny, or a natural phenomenon.

Furthermore, in order to investigate stability or change in personality from young-old age to old-old age, a survey, which had been given to 70 year olds living in the community, was followed up 10 years later (1976–1986). It was revealed that positive family image increased with advancing age. With respect to self-perception, the past self changed from objective to positive. The present self maintained positive responses, while the future self shifted from objective to negative (Shimonaka & Nakazato, 1989).

The Rorschach Inkblot Test has often been used to analyze the personality of the elderly in Japan. Shimonaka and Nakazato (1991) used the Rorschach Inkblot Test six times over a 10 year period to study changes in personality related to aging and to determine whether significant changes in personality precede death.

Age-related changes were observed in the number of responses, responses regarding rejection, shading and popular, poor form level, and content range. Changes in these variables suggest that productivity, creative thinking, cooperativeness, and concern for and interest in society lessen with advancing age. Over the four to six years preceding death, popular responses and content range declined, with the decline accelerating near death. Tests given within two years of death showed less animal movement response, higher form (F(−)%), and higher human content. These results suggest that productivity, vital energy, and ego strength become reduced prior to death, while interest in and sensitivity to others is maintained to the end (Shimonaka & Nakazato, 1991).

People are inclined to have a negative image towards aging. This tendency was especially strong during the 1960s and 1970s. During this time, old age was considered to be a time of much negative change and anxiety (Cattell, 1965; Jarvik & Russell, 1979). However, from the results of the life-span developmental approach regarding anxiety (Costa et al., 1986), it has been suggested that feelings of anxiety decrease as age increases.

The anxiety of Japanese adults aged between 25 and 92 years was measured using a Japanese version of the State–Trait Inventory and found that the level of anxiety lowered linearly as age increased. These results suggest that anxiety about death is not, as expected, especially strong in old age (Nakazato & Shimonaka, 1989).

Psychotherapy for elderly people by clinical psychologists in Japan is quite uncommon. Since the 1980s, however, more effort has been made in this regard at institutions and day care centers for the elderly, as the population of the elderly has been increasing. This has been mainly group therapy, and there has been no reported research on behavior, cognitive, or family therapies yet. Moreover, many of the research papers available use results based on case studies and subjective evaluations by practitioners. Since most subjects of group therapy are those with senile dementia, reality orientation, reminiscence, and therapeutic recreation are the main activities provided. This is probably due to the fact that the drastic increase of people with senile dementia has become a social problem in Japan. Since these group therapy activities had not been used until 1990 in Japan, many research papers mainly examine their effectiveness.

The effectiveness of reality orientation has been examined using as subjects people with moderate senile dementia in a nursing home. Changes in the dementia scale (Hasegana, 1983; Kahn, Goldharb, Pollack, & Peck, 1960) were observed before, just after, and six months after the orientation. The participants showed improvement in the dementia scale scores, unlike the control group, which showed no change.

Furthermore, according to the evaluation of daily behavior in the ward by the charge nurse, improvement was seen in social communication. The improved dementia scale scores were even maintained six months after the orientation. These results suggest that reality orientation is very effective at the early stages of dementia (Shimonaka & Nakazato, 1989; Shimonaka, Nakazato, Habu, Haga, & Mizuno, 1988).

The effectiveness of reminiscence, which is as common as reality orientation in countries other than Japan, was examined with inpatients of a psychiatry ward as subjects, and evaluation was made using the Moses scale. Moses scale scores showed improvement after reminiscence. These results suggest that the psychological state, relationships with others, and the level of independence among people with senile dementia was improved by participating in reminiscence (Kurokawa, 1995).

There have been other reports on the effectiveness of therapeutic recreation incorporating play therapy, art therapy, sports therapy (Kawamuro, 1995; Koike, 1995; Sadaki, 1993), and music therapy (Kitamoto, 1996) which are carried out at geriatric hospitals, day care centers, and nursing homes.

Finally, there has been a unique approach, in which group therapy, provided during day care for outpatients with senile dementia who are receiving care at home from their family, is combined with group counseling provided for the outpatients' family members. This attempt aims to vitalize the residual mental functions of people suffering from senile dementia whose activity is restricted to their homes. At the same time, family members who are exhausted from care giving are provided with rest time and group counseling (Kato et al., 1987).

The various therapies above for people with senile dementia have just been introduced in Japan and they are still at a stage where their effectiveness is being researched. Storandt, Siegler, and Elias (1978) have pointed out in *The clinical psychology of aging* that there were too few clinical psychologists for people aged 65 and older. Despite the fact that 10% of those aged 65 and over were in need of some kind of mental health services in the US at that time, clinical psychology had long neglected the treatment of older adults. In the mid-1990s Japan is in the very same situation. Although there are few clinical psychologists in Japan, their efforts for proper application of therapies to the aged today will bear fruit and become a touchstone for tomorrow's clinical psychologists for the aged.

2.40.4 REFERENCES

Association for Health and Welfare Statistics (1994). *Trend of national health*. Tokyo: Association for Health and Welfare Statistics.

Cattell, R. B. (1965). *The scientific analysis of personality*. New York: Penguin.

Costa, P. T., Jr., McCrae, R. R., Zonderman, A. B., Barbano, H. E., Lebowitz, B. Z., & Larson, D. M. (1986). Cross-sectional study of personality in a national sample: 2. Stability in neuroticism, extravertion, and openness. *Psychology and aging*, 1, 144–149

Education and Training Committee Association of Japanese Clinical Psychology (1993). Questionnaire to the new members (1986–1991). *Journal of Japanese Clinical Psychology, 11,* 66–78.

Hasegawa, K. (1983). The clinical assessment of dementia in the aged: A dementia screening scale for psychogeriatric patients. In M. Bergener, U. Lehr, E. Lang, & R. Schmitz–Scherzer (eds.), *Aging in the eighties and beyond.* New York: Springer.

Institute of Population Problems, Japanese Ministry of Health and Welfare (1994). *Population trends*. Tokyo: Health and Welfare Statistics Association.

Japanese General Affairs Agency (1992). *Survey on the life and opinion in later years*. Tokyo: Chuohoki.

Japanese General Affairs Agency (1995). *Survey on attitudes among the elderly people*. Tokyo: Chuohoki.

Japanese Ministry of Health and Welfare (1992). *White paper on health and welfare*. Tokyo: Japanese Ministry of Health and Welfare.

Japanese Ministry of Health and Welfare (1996). *White paper on health and welfare*. Tokyo: Japanese Ministry of Health and Welfare.

Jarvik, L. & Russell, D. (1979). Anxiety, aging and the third emergency reaction. *Journal of Gerontology*, *34*, 196–200.

Kahn, R. L., Goldharb, A. I., Pollack, M., & Peck, A. (1960). Brief objective measures for determination of mental status in the aged. *American Journal of Psychiatry*, *117*, 326–328.

Kato, S., Ikeda, K., Hirata, N., Imai, Y., Ichihara, H., Homma, A., & Hasegawa, K. (1987). Attitudes toward the elderly with dementia among their family caregivers and their changes. *Japanese Journal of Psychogeriatrics*, *4*, 779–789.

Kawamuro, Y. (1995). Group psychotherapy. *Japanese Journal of Psychogeriatrics*, *6*, 1503–1511.

Kitamoto, F. (1996). An attempt at group music therapy for elderly persons. *Journal of Japanese Clinical Psychology*, *14*, 141–151.

Koike, K. (1995). Therapeutic recreation. *Japanese Journal of Psychogeriatrics*, *6*, 1497–1502.

Kurokawa, Y. (1995). Reminiscence group with Japanese dementia patients. *Journal of Japanese Clinical Psychology*, *13*, 169–179.

Nakazato, K., & Shimonaka, Y. (1989). The Japanese State–Trait Anxiety Inventory: Age and sex differences. *Perceptual and Motor Skills*, *69*, 611–617.

Sadaki, T. (1993). An example of a group therapy program for elderly people with dementia. *Journal of Japanese Clinical Psychology*, *11*, 53–58.

Shimonaka, Y., & Nakazato, K. (1980). Psychological characteristics of Japanese aged: A comparison of Sentence Completion Test responses of older and younger adults. *Journal of Gerontology*, *35*, 891–898.

Shimonaka, Y., & Nakazato, K. (1986). The development of personality characteristics of Japanese adults. *Journal of Genetic Psychology*, *147*, 37–46.

Shimonaka, Y., & Nakazato, K. (1989). Personality development among the Japanese elderly: A ten-year longitudinal study on a sentence completion test. *Journal of Cross-Cultural Gerontology*, *4*, 347–356.

Shimonaka, Y., & Nakazato, K. (1989). *Effectiveness of reality orientation applied to Japanese elderly with dementia.* Paper presented at the Tenth Biennial Meetings of ISSBD, Jyvaskyla, Finland.

Shimonaka, Y., & Nakazato, K. (1991). Aging and terminal change in Rorschach responses among the Japanese elderly. *Journal of Personality Assessment*, *57*, 10–18.

Shimonaka, Y., Nakazato, K., Habu, K., Haga, S., & Mizuno, H. (1988). Effects of group work for dementia in the nursing home residents. *Japanese Journal of Clinical Psychiatry*, *17*, 101–109.

Statistics and Information Department, Minister's Secretariat, Japanese Ministry of Health and Welfare (1995). *Patient survey* (Vol. 1). Tokyo: Japanese Ministry of Health and Welfare.

Statistics and Information Department, Minister's Secretariat, Japanese Ministry of Health and Welfare (1996). *Report on the socioeconomic survey of vital statistics.* Tokyo: Association for Health and Welfare Statistics.

Storandt, M., Siegler, I. C., & Elias, M. F. (Eds.) (1978). *The clinical psychology of aging*. New York: Plenum.

Takeda, K. (1994). *Why are elderly women killed by their family member?* Tokyo: Minerva.

Tanaka, S. (1994). *Preliminary study on elderly abuse in Japan*. Tokyo: The Study Group for Casework for the Elderly.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

2.41

National Advocacy for Psychology

JILL N. REICH and NINA G. LEVITT
American Psychological Association, Washington, DC, USA

2.41.1 INTRODUCTION 479
2.41.2 THE NATURE OF ADVOCACY IN THE POLITICAL PROCESS 480
2.41.2.1 What is Advocacy? 480
2.41.2.2 Why is Advocacy Necessary? 480
2.41.2.3 Who Does Advocacy? 481
2.41.2.4 What Makes Advocacy Effective? 481
2.41.3 ADVOCACY: IN ACTION 482
2.41.3.1 Advocacy: Goals and Plans 482
2.41.3.2 Advocacy: Making the Connection 483
2.41.3.3 Advocacy: Process or Outcome 484
2.41.3.4 Advocacy: Cost, Competition, and Champions 485
2.41.3.5 Advocacy: The Uncontrollable Element of Time 485
2.41.3.6 Summary 486
2.41.4 THE ADVOCATES 486
2.41.4.1 The Constituent 486
2.41.4.2 The Lobbyist 487
2.41.4.3 Third Parties: Organizations/Associations 487
2.41.5 FUTURE DIRECTIONS 487
2.41.6 REFERENCES 489

2.41.1 INTRODUCTION

The development of knowledge and its application occurs within a social construct. As such, a discipline is both influenced by the society in which it lives and, when that knowledge is disseminated and applied, it influences that society. This dynamic is especially true for a discipline like psychology, and very particularly clinical psychology, which has so much to contribute to issues of national and global concern. For example, the results of a massive new assessment of the world's state of health, and of the changes it is likely to undergo between 1997 and 2020, underscores the role of behavior in maintaining a healthy citizenry (Murray & Lopez, 1996). This study reports that the most overlooked cause of death and disability is injury, specifically using the word "injury" rather than "accident" in order to communicate that injurious events are avoidable. They note that the most important single variable expected to impact the world's health in 2020 is the use of tobacco. In a ranking of conditions purported to cause the greatest disease burden, severe depression moves from fourth highest in 1990 to second highest in 2020.

These data and their projections suggest that the discipline of psychology, especially through its subfield of clinical psychology, should be at the forefront of healthcare, and through these contributions valued highly by the society in which it lives. A prominent measure of the national status of the discipline is the support of that discipline in the corridors of national governance. For psychologists in the USA, this means looking at the perception, knowledge, and support for the discipline from members of the US Congress: members of the Senate (Senators) who represent their states each term for a period of six years and members of the House of Representatives (Representatives) who represent their districts (within each state) each term for a period of two years (despite much discussion about limiting the number of terms a member may serve, no action has ever been passed to do so). A second source of support and hence a measure of status lies in the Executive branch of government which through its agencies implements the legislation and policies set by Congress. It is through these governing bodies that the larger society identifies its interests, expresses its values, and develops the means to meet its national needs.

The role of advocacy in society's perception of psychology cannot be underestimated. This role is enlarged by the dynamic nature of advocacy which works to recognize psychology's potential to contribute to areas of national need and in so doing further reinforces the need for psychology's contributions to society. By definition, "advocacy" involves "pleading or defending the cause" (Webster; 1995), in this case, the cause of psychology. Thus, when advocating on behalf of psychology, one educates policymakers and implementers about the value of the discipline and hence the need or right of the discipline for support, resources, recognition, status, and the like. At the national level, most often advocacy involves working with policymakers and their staffs in Congress, and/or the implementors of these policies, the staffs in the various federal agencies.

In this chapter, the process of advocacy will be described, with its critical variables, and how these factors, together with existing structures for advocating, serve to enhance or hinder successful advocacy on behalf of psychology, in particular, utilizing examples from advocating for support for education in professional psychology. Moreover, as the advocacy process by its very nature is a changing one, emphasis should be placed on patterns rather than particulars as the specifics can be expected to change quickly and often.

2.41.2 THE NATURE OF ADVOCACY IN THE POLITICAL PROCESS

2.41.2.1 What is Advocacy?

The democratic process is one in which complementary and competing interests must vie for recognition and, through that recognition, for control and resources. Whose interests and which factors are recognized might be thought to be determined logically by the nature of the issue. But whose understanding of the issue and which priority among the many possible will prevail? How the policymaker understands the issue, its component parts, and what interests or factors are thought to be critical to the matter being deliberated necessarily is in the eye of the beholder. Thus, when advocating on behalf of psychology, that is, when pleading the case of the discipline, in effect the listener is being educated about the ways in which psychology can contribute to the issue at hand. To be effective, this must be done in ways that are meaningful, of interest, and value to the listener, and true to the discipline. In the political process, with its broad range of interests and decision making, this means that the job is to demonstrate to the listener, the policymaker, or legislator why and how psychology can serve the national issue being considered. Of course, the particular issue varies. Given the US national legislative agenda with its attention to healthcare, and the potential of clinical psychology to contribute to this agenda, the conversation is most likely to focus on one of the six major areas of health that have been demonstrated to be behaviorally based: violence, substance abuse, smoking, injury, diet, and AIDS. The goal then is to have the listener recognize that psychology has a significant contribution to make, so that when funding is being allocated or professional fields identified for inclusion in a particular practice or intervention, the listener will conclude that psychology should be a part of the policy and so named in the legislation. At the very least, the goal of advocacy is that the legislation should not preclude or deny psychology's participation in the bill being considered.

2.41.2.2 Why is Advocacy Necessary?

Ideally, advocacy is necessary for the same reasons that any education endeavor is necessary; so that the policymaker will have the benefit of the most pertinent information to the task, that is, so that the issues faced by legislators will have the benefit of the contributions of the discipline. In this scenario, the best the discipline has to offer is put forward to a particular cause and the results left to the good judgment of the legislators. But the political process is not always so ideal. There will be many ways to meet the demands being deliberated, many interests involved, and limited resources available for allocation. In these cases, advocacy also serves to bring attention and recognition to psychology. In the US political process, if there is no advocacy on behalf of the discipline, psychology is unlikely to be recognized. Herein lies a particular problem for clinical psychology in the complex and crowded healthcare arena in the 1990s.

This problem arises from the very nature of the discipline being a field with numerous subfields. It is not well understood by the general public. Among those who do know the discipline, some know psychologists as mental health practitioners; others are aware that psychologists are behavioral scientists who conduct research in areas like cognition and neuroscience; and still others remember their psychology professors from undergraduate courses in personality and perception. Yet often these separate parts are not a coherent whole, able to bring the full strength and richness of the discipline to bear on the range of policy matters before Congress. Moreover, in the process of advocating for a particular issue, and needing to keep the message simple, most often only a limited part of psychology will be addressed, for example, the needs of practice or the particular implications for science. Thus, despite the many roles that psychologists actually serve, there remains a limited public understanding of psychology as a whole and the many ways in which the discipline can be applied to real issues of public concern and national need: a fact that unfortunately is as true for policymakers as for the general public.

2.41.2.3 Who Does Advocacy?

There are many answers to this question. On the most general level, while this answer may come as a surprise to the reader, in fact, everyone in the discipline advocates. The problem of course being that psychologists are unaware that in the process of their work, they also serve to educate others about whether or not this work is valuable, deserves public recognition, support, and sometimes even protection. In spite of the lack of awareness, whether in classrooms or offices, how the discipline is presented, to whom, and with what level of respect and value will greatly influence how the listener will view the discipline now and in the future. Perhaps the most misunderstood arena for advocacy are classrooms where the citizens of tomorrow are most likely to be introduced to the discipline and its potential to contribute to areas of national need. If psychologists do not show respect for the discipline as a whole, others can hardly do so. If psychologists do not demonstrate how their knowledge base can contribute to national need, others cannot do so. And if the value and application of their work to areas of national interest are left unspoken, then these citizens, or their representatives, cannot be expected to allocate national resources to support psychologists' education, science, or practice.

On a more direct level, of course, advocacy follows a planned strategy, developed by professional lobbyists working with psychologists and focused on policymakers involved in issues of particular interest to psychologists, whether these be in the areas related to education, practice, science, or public interest. This approach, necessarily focused and targeted, remains dependent on the participation of psychologists at every level and activity as will be delineated as the process of advocacy for psychology is described.

2.41.2.4 What Makes Advocacy Effective?

Effectiveness is based on the same factors that make any human interaction effective: clarity, consistency, context, and credibility. Advocacy is not accomplished through magic nor are advocacy skills acquired through academic study. Advocacy, while a learned behavior, must be expanded through experience and perfected by practice. In other words, if advocacy is examined for its key elements, the most basic skills of human interaction will emerge. These skills then must be augmented by knowledge of the discipline, motivation to succeed, much persistence, and a little bit of good luck.

Some compare the components of successful advocacy to those of actual warfare, in this way placing key factors into the framework of a battle. Whether or not warfare is considered effective human interaction is debatable. However, the analogy serves to focus attention on the importance of reaching a conclusion in a competitive environment. At times, the analogy seems apt; the competition is real and the desire to win is strong. Zorack (1990) reports that the lobbying process parallels those factors developed by British Major-General J.F.C. Fuller during WWI as elements of a successful combat campaign:

(i) objective—clearly define an attainable goal,

(ii) offensive—exploit the initiative, be proactive, not reactive,

(iii) mass—concentrate efforts at optimum time and place,

(iv) economy of force—focus on one initiative at a time,

(v) maneuver—place opposition at disadvantage,

(vi) unity of command—maintain one leader,

(vii) security—know what information to share and what not to share,

(viii) surprise—employ unexpected legislative maneuvering when possible, and

(ix) simplicity—simplify and limit all requests.

The above-mentioned tenets for a successful campaign, regardless of the battlefield, are self-evident if advocacy is conceptualized as a campaign and the delivery of the discipline as the artillery. For this advocacy battlefield, the message is to recognize the need and value of psychology to society thereby establishing its right to support, recognition and involvement in legislation, and its implementation. In this view of the advocacy process, emphasis remains on the importance of communicating effectively, most often to individuals with little knowledge of the message and with great need to act efficiently and effectively for a broad range of very vocal constituents. Psychologists, especially, should be able to utilize their understanding of perception, memory, problem solving, social interaction, motivation, and the like to develop effective strategies for advocacy.

2.41.3 ADVOCACY: IN ACTION

2.41.3.1 Advocacy: Goals and Plans

Successful advocacy requires a plan of action consistent with the goal. As an example, the overarching goal for education is to increase federal funding for education and training programs in psychology. But there are many ways to go after this larger goal: through new legislation, by revising existing legislation to more effectively include psychology, or by changing regulatory language that controls how the legislation is implemented to name the most likely target areas. Careful analysis of the range of legislation that it is possible to effect, the likelihood of success, the ability of the results to benefit the discipline, as well as the more practical factors of time, cost, and resources will determine the particular target initiative. Because advocacy resources can rarely match the range of possible initiatives, priorities must be established. Within the American Psychological Association (APA), priorities are first established for Education by the Board of Educational Affairs with assistance from the Education Directorate and the Public Policy Office. Once the priorities are set for education, they are then tested with those of the other Directorates to achieve oversight, coordination, and priorities for the association as a whole. For example, in 1996, opportunities were available to impact the Higher Education Act and the Public Health Service Act as both were being reauthorized during the 105th Congress. Thus, it was decided to target, as the primary advocacy goals, expanded eligibility and increased participation for psychology in these bills.

Choice of a particular target initiative determines where the strategy will be focused since political combat may take place in either the legislative or executive branches of government. Although education advocacy is focused primarily on Congress and the legislative process, it sometimes targets the Executive Branch around the promulgation of regulations, thereby seeking to establish a positive and productive relationship with agency officials who administer the programs in which psychology participates. The two priority pieces of legislation for education, both in the reauthorization process, focus the strategy on Congress, particularly the members of those committees with jurisdiction over these bills who make final recommendations for action to the full Congress.

With the goals established and targets identified, education advocacy first develops a clear message. An information strategy was employed to educate the chairs and members of relevant committees and subcommittees in the Senate and House of Representatives about psychology's ability to meet national needs. If psychology is seen as important to the nation's well-being, then the education of psychologists should be supported by federal funds. In order to educate and thus garner support from these key legislators, two sources of information are needed. First, clear, databased, and well-presented information is necessary to establish psychology's contributions to national needs. In choosing from among the many possible topics for this information base, areas of interest and meaning to the audience should be sought and found. So, if the Chairperson of the Committee is very interested in health, say cancer treatment, then demonstrating psychology's role in meeting these particular needs is likely to be of great interest. Thus, topics are chosen and data and other relevant information about psychology's contributions to these topic areas are gathered from publications, the APA Research Office, federal agencies, and other sources to develop concise and clear informational materials. In making the case, psychology's ability to contribute in significant ways to the topic area must be demonstrated, but also the value and contribution of psychology must be established through benefit-analysis and cost-effectiveness data. Particularly, but not exclusively, in the 1990s healthcare arena, these two factors are dominant in considering the value of a discipline's ability to contribute.

Further, since psychology is not yet firmly established in the public's mind as a health profession, and certainly not as "primary care," efforts to justify and promote the contribution of psychology to healthcare are critical to successful advocacy. Awareness must be increased about clinical psychology's unique contributions and expanding knowledge, not only as mental health providers, but especially as behavioral health professionals. The reason for this emphasis is twofold. First, it is consistent with the work of many in the field; there is much for psychology to contribute as a behavioral health provider. Just as important is that behavioral health is consistent with the view policymakers have of healthcare in the 1990s and as they expect it to evolve in the twenty-first century. This approach points out that in making the case, it must be understood how listeners understand the issue and fit the discipline's work, worth, and contributions to that perspective.

Fortunately, this fit is often not difficult to make, but it must be remembered to take the time to do so. After all, the sign of a good educator is to begin at the student's level of knowledge and understanding and build from there. Knowing that the policymaker understands and values the role of behavior in healthcare sets the stage for placing clinical psychologists' unique contributions as educators, researchers, and practitioners into that context. In other words, the advocate starts with the perceptions of those making policy, understands how to fit the discipline into their national priorities, and elucidates what they need to know about psychology in order to make the relevant connections to garner their support.

Once the primary message of how the discipline can contribute to the area being considered and the value of psychology as a cost-effective player is established, then the advocate can move on to another essential aspect of the advocacy message which is to educate policymakers about concerns and interests of psychologists. Until this point, the advocate has been educating the policymakers about how the discipline meets their needs. Having achieved that part of the message, advocates move on to explain what the discipline needs in order to be able to continue to make these important contributions. This might be funding for education and training of a new generation of psychologists or new skills for the current cohort; support for continued research and new knowledge, or inclusion as providers of healthcare, eligible to access programs of federal support for healthcare.

2.41.3.2 Advocacy: Making the Connection

With the message developed and packaged, how to present it must be determined, coordinated, and implemented utilizing as many formats as possible, delivered by those most relevant to the audience: their constituents, the voter; their expert constituents, the psychologist voter; and finally national experts, the lobbyists and staff of national associations, in that order of importance and credibility. At the same time, it is the national experts whose full-time job is to advocate. The remaining participants are volunteers working within limitations on time, resources, and advocacy expertise and "the lobbyist who has the access has the best opportunity to wield the most influence" (Zorack, 1990). It is critical to establish rapport and a positive relationship with key legislators and their staff. A positive, ongoing relationship has implications for present and future legislative initiatives. The relationship must be based on honesty and trust, be mutually beneficial, provide expertise (directly or indirectly) and ensure constituent support.

Establishing a relationship with staff and/or legislator is accomplished primarily through visits to policymakers at the national, state, and local levels. Congressional briefings on a particular subject of interest to legislators and their staff, as well as site visits to places of interest in their districts/states, are also helpful in establishing relationships. Keeping in touch with staff by telephone is another important way to maintain ongoing communications. As members of Congress begin to use Internet, that will be another important and far easier method of communication for advocates. The use of fax machines and email, in fact, has revolutionized communication that needs to be timely and efficient. However, too much contact (e.g., phone calls, faxes, or email) can also be a serious nuisance to busy staff and legislators, and at times highly detrimental to one's efforts.

Constituents communicate best through visits, phone calls, and letters. In fact, constituent mail is still the most effective means of influencing how a member of Congress votes. The other is supporting members of Congress in their fundraising efforts to be re-elected, especially members of the House of Representatives who must run for office every two years. When thought of in terms of educating these members, consistency is a worthwhile investment.

Finally, successful connections result from a multifaceted strategy. Especially if the message is a new or unfamiliar one, it will benefit from several hearings, different perspectives, and several messengers. Thus, a successful strategy will depend not only on the ability to know what actions are needed, but ultimately on their implementation. Some initiatives will require working with one or more associations in a coalition, some will involve taking the message directly to the legislator, while others demand working with some or all of the committee staff, and some may need an accompanying public relations campaign while others will require a one time Congressional briefing. Of course, the quantity and quality of activities undertaken to meet an advocacy target initiative must be consistent with the availability of resources and time.

2.41.3.3 Advocacy: Process or Outcome

Following any policy issue in the popular press makes it obvious that legislation and policymaking are ongoing processes. So too is advocacy. Thus, it makes sense that establishing and maintaining relationships with legislators, staff, and agency officials must be a continuous process. This recognition makes the determination of the outcome of a particular advocacy initiative unclear, especially when working with the legislative branch. Therefore, it is unwise to measure the success or failure of a legislative initiative on whether or not a particular advocacy goal is achieved at a specific moment in time.

Indeed, federal legislative initiatives are unpredictable, especially when dealing with enacting or reauthorizing legislation. Even though budget and appropriations legislation occur according to a defined cycle, their underlying legislative processes are unpredictable such that along the way any anticipated outcome is far from certain; it is simply not done until it is done. In fact, as the entire legislative process becomes subject to greater political influences and tensions between the executive and legislative branches of government, it becomes ever less certain what is going to happen or when. For example, during the first session of the 104th Congress, the Republican-controlled House and Senate could not come to an agreement with the Democratic Administration on funding federal programs. As a result, 15 separate Continuing Resolutions were enacted to allow federal programs to continue during this fractious period. However, with each enactment of a Continuing Resolution, programs were targeted for a set percentage reduction in funding with no opportunity to change any other part of the legislation.

Certainly, enacting or reauthorizing legislation is the most uncertain process to effect, as thousands of bills are introduced each session but only a handful are enacted into law. In fact, as demonstrated in the above example, even those bills being reauthorized (i.e., those bills requiring review and extension) are not necessarily acted upon within the term of Congress specified for reauthorization. A good example is the effort to reauthorize the Public Health Service Act during the 103rd Congress. It passed both Houses of Congress, and was approved in conference by both houses. However, at the last minute, only the full House gave the bill its final *pro forma* approval; the Senate did not. Thus, the Public Health Service Act remained unchanged. The Senate Republicans looking to the November 1994 elections, when the Republicans made sweeping victories and took control of both Houses of Congress, did not want to give President Clinton even one small success regarding healthcare legislation. This result was most unfortunate for psychology because the proposed revised Public Health Service Act legislation had contained a hard-won provision to set aside funds for the education and training of minorities in psychology. As a result, this piece of legislation remains a high priority for education advocacy efforts, five years and three sessions of Congress after first beginning this work.

In summary, getting a bill drafted is the easy part, getting it passed is next to impossible. In fact, getting a bill through Congress takes an enormous amount of time and effort. Amending a piece of legislation depends on whether or not the statute is up for consideration (i.e., reauthorization), as well as what the proposed amendment(s) contain. It must be emphasized that the least amount of change, coupled with its degree of simplicity, positively effects its chances of being accepted.

Advocacy efforts regarding federal regulations are the easiest to control, which is not to suggest that they are easy. For instance, providing feedback to the agencies promulgating regulations happens according to a schedule and usually results from a consensus, that is, getting those who have a vested interest to agree with how a program is to be implemented. Working with agencies to direct more funds and attention to a specific field (e.g., psychology), or to place representatives of the field on advisory bodies, is much more challenging and far from certain. An example from education occurred, during the 103rd Congress, when the Bureau of Health Professions (a program authorized by the Public Health Service Act) chose to focus primarily on increasing the number of ethnic minority physicians. With decreasing federal funding, the Bureau of Health Professions was not interested in increasing the number of psychologists in the programs without first receiving an increase in funding. For this reason, education advocacy efforts to work with the agency to include psychology in their programs became a notable challenge. It is pleasing to report that this has met with continued success. This success has been gained in small steps over several years with still some distance to go before the goal will have been reached. Each step has required ongoing connections with the agency's staff, maintaining their attention to psychology's needs and psychology's ability to support their agency's mission and goals. Finally, along the way it is necessary to be sure that the agency's staff is cognizant of the links between Congressional decisions as evident in legislation, psychology's constituents, and agency implementation.

All of this serves to underscore the importance of having a clear and well-marked strategy, one that serves as a road map for a long and at times frustrating journey. Detours along the way must be expected and markers established against which to measure progress and ensure that the journey remains on track. Like any journey, separating process from outcome along the way can be distracting and miss the point. In the advocacy world, advancing the cause of psychology may occur in spite of missing any particular legislative target, and that advance will serve well in the next round of legislation and policymaking.

2.41.3.4 Advocacy: Cost, Competition, and Champions

As has been described, achieving advocacy goals may take years to accomplish during which time consideration of cost, competition, and champions are most relevant. Lobbying is of little consequence if the advocacy plan or activities lack organizational support that translates into real dollars. The more resources available, the more extensive and thorough the advocacy. Unfortunately, there is no consistent relationship between the amount of money spent and the likelihood of success. No matter, it is also always true that without adequate funding, the scope of advocacy goals and expectation for success must be limited from the start. Funding is needed for professional staff to create and prepare information materials, to contact and prepare psychologists' advocates, to organize and assist in the delivery of the advocacy message, to fund the necessary trips, and related support of the myriad of advocacy activities that are part of successful campaigns.

The cost of an advocacy initiative will be dependent not only on the expected activities of the initiative but also on the extent of competition or opposition encountered along the journey. For example, current beneficiaries of existing legislation often do not want to share the limited federal dollars that would result from adding psychology to the recipients of federal support. Also, it is not uncommon to find that there are those who want funds directed to their legislative priorities and away from others. Whatever the reason, the competition must be addressed and that requires more time, effort, and money. For example, opposition from others in the healthcare community during the Public Health Service Act debate resulted from their reluctance to share limited, and dwindling, federal funds. That particular obstacle was removed by forming a coalition with other professionals in the healthcare community to advocate jointly for issues of mutual interest. What was most effective in establishing this coalition was educating the healthcare community about what psychologists do, convincing them that together, better healthcare can be provided, each complementing and supplementing the other professional services. Nevertheless, in the competitive arena of healthcare, opposition from the medical community and others will remain a challenge for psychology to confront; each aspect will need to be addressed and argued one on one, for a long time to come.

Competition and opposition are constant companions of the advocacy process. As outlined, there are many stages in the legislative process. Each one presents an opportunity to block your competitors and likewise to be thwarted by them. Congressional champions are the single most significant factor in meeting these challenges. Especially when the opposition is stiff, a powerful legislator serving as your champion can be the deciding factor. Champions and supporters of the proposed advocacy initiative are critical to advancing the cause: it is extremely difficult, likely impossible, without them. These champions not only bring the message to the floor of Congress but also shepherd the votes of other policymakers in your favor.

2.41.3.5 Advocacy: The Uncontrollable Element of Time

Timing is everything. If one waits too long in the political process before presenting a case, there will not be sufficient time to carry forth the message through letters, visits, and testimony. As the voting nears, legislators, and their staff are frenzied, unwilling, and unable to process new information, especially information that will upset delicate balances built among existing constituents. But if a proposal is presented too early, it may be shelved by the legislators and/or their staff, forgotten in the midst of more pressing matters requiring immediate attention. Lobbyists must monitor their contacts with legislators and staff; no one likes someone who usurps their valuable time. Nevertheless, unless calls are made regularly to legislators and staff, it is easy to miss important opportunities to plead the cause. For example, a hearing on a particular bill provides an excellent opportunity to have a psychologist testify in Congress, in this way publicly making the case for the discipline to all Committee members.

The length of time to accomplish an advocacy initiative is the most uncontrollable of all aspects of advocacy. Indeed, complicated initiatives must be considered for the long haul, that is, more than one session of Congress, often several years (note: one session of Congress covers a two-year period). For example, the Public Health Service Act reauthorization entered the first session of the Congress, approximately five years after the initial efforts to reauthorize that legislation. Moreover, all along the way it is necessary to compromise, increasing the chances of getting less than what you at first wanted, and probably concluding with something quite different from what you first set out to accomplish. Nevertheless, as noted above, there are tangible benefits from the process during these long initiatives: educating policymakers about issue and discipline, establishing a positive relationship with key legislators and their staffs, and through this work garnering their support for psychology that will serve well for future endeavors.

During this time, however, the larger political context will change, a factor directly effecting the political process. Prevailing national priorities do not remain constant. Variables such as fiscal strength and social climate, whether real or perceived, national or international, impact what happens in Congress by influencing the policymakers' receptivity to your ideas, messages, and needs.

2.41.3.6 Summary

Successful advocacy is a complex, multidimensional process requiring discipline and flexibility, vigilance, and strategic action. It requires ongoing involvement by psychologists who often understand neither the reasons for their participation nor the process. Finally, vicissitudes of time act to introduce a level of uncertainty and lack of control particularly frustrating to the uninitiated. In all, for so important an activity, it is fraught with barriers to its success. Coordination, knowledge, experience, and motivation are key to meeting these challenges.

2.41.4 THE ADVOCATES

2.41.4.1 The Constituent

As Tip O'Neill, a legendary Speaker of the House of Representatives for over 15 years, was frequently heard to say: "Politics is people." Without question, the key player in any advocacy initiative is the constituent (individually or collectively). Individually, each constituent is important because they vote to elect the legislator, something that is especially significant to members of the House of Representatives who must stand for election every two years. Understanding this, many psychologists are politically active and play a role in the political process at all levels: local, state, and federal. Collectively, groups of constituents (e.g., state associations) have substantial clout, especially if there is an established network of communication and committed members who have financial and other means of support. Large national associations, such as the APA, are perhaps the "wealthiest" politically because of the number of members nationwide who can give their support.

APA members, like members of other associations, can be involved in advocacy through grassroots activities which are quantitative and qualitative. Quantitative grassroots refers to the broad-based approach aimed at generating as many letters, calls, and visits to legislators as possible. Qualitative grassroots emphasizes the quality of the relationship (or the connection) between individual constituent and their legislators. These relationships require thoughtful interaction about common interests; communication that exists over time, not limited to meeting self-serving causes for either side; and meaningful contributions and responses one to another, that is, contributions that each side acknowledges and values. For legislators, constituents need to know that their concerns have been heard and met in the legislative process. For constituents, legislators look to them as a critical factor in political fundraising. It is a reality of American politics that funds are required to run for office or to be re-elected. Once again, members of the House of Representatives are particularly vulnerable because of their election schedule. For this reason, members of Congress appreciate financial support, regardless of the amount. Further, when there are a number of issues to which they might respond, they will choose those that are popular with their supporters. A clear sign of interest can be communicated through financial support. Thus, in spite of discussions to the contrary, until there is a significant change in the political process, a fact of life remains that constituents and organizations who can contribute funds will very likely have the most influence in the legislative process.

2.41.4.2 The Lobbyist

The lobbyist's three mottos are: (i) tell the truth; (ii) burn no bridges; and (iii) if you don't ask, you don't get (Wolpe, 1990). Lobbyists are conveyors of information. Helpful information is appreciated; information supporting and justifying an issue is essential for accomplishing one's advocacy objectives. Indeed, the main job of the lobbyist is to get decision makers interested and involved in the lobbyist's issues. Communications must be continuous, occurring on a regular and timely basis, through the use of newsletters, newspaper articles, internet, broadcast faxing, phone banks, and 800 numbers. Lobbyists must also educate and train their advocate participants about the process of advocacy: how the legislative process works, how to develop relationships with legislators, how to communicate with legislators, and the like. At APA this is accomplished by special training weekends with a group of psychologists around a particular target goal and culminating in visits. Ideally, in the future, this training will take place throughout the country on a state or regional basis so as to expand the number of psychology constituents/voters effective and knowledgeable about the advocacy process.

Lobbyists also need information themselves. Lobbyists need to know the field they represent, and the issues of their employers in order to be prepared to describe, explain, and justify issues and needs with legislators and their staff. For example, in the area of psychology and healthcare, lobbyists need to know the many ways the discipline contributes to good health, they must be able to answer complex questions, and be able to make the necessary connections between these contributions and the discipline's needs, say in the area of education and training. During each visit to legislators, it is imperative for the lobbyist to know just what they want to accomplish. They need to know the interests and idiosyncracies of their listeners in order to make their case, often during a time of high tension and great demands. Lobbyists, like constituents, need to establish relationships over a period of time, and preferably nurture these relationship during the quiet times afforded by Congressional recesses.

In addition, lobbyists must anticipate the opposition and solidify support from allies. Lobbyists need to be prepared to deal with the opposition, directly and indirectly. Most of all, they need the support of allies. In the political process, there is strength in numbers and relationships built over time on foundations of trust.

2.41.4.3 Third Parties: Organizations/Associations

Numbers count in the political process as a measure of interest and because they represent votes. Thus, an organization or association has a more powerful voice than an individual and a coalition of many organizations has significantly more power than just one organization. It is especially helpful when organizations can work together toward shared goals of mutual benefit. For this reason, whenever possible in advocacy efforts, organizations are sought with which to coordinate activities, complement and supplement efforts of one another. In some instances, membership in coalitions requires a fee to support paid staff to work on behalf of all the coalition members (e.g., Council on Education Funding). However, working with coalitions is a luxury not always available because of competing interests among groups with similar needs.

2.41.5 FUTURE DIRECTIONS

While the nature of advocacy seems straightforward, its implementation remains complex for the discipline of psychology. To be effective, the discipline must be accepted as a credible player in the legislative arena. People who have participated in national advocacy efforts are likely to have been surprised by the naivety about the discipline by well-meaning professionals in a variety of congressional offices and federal agencies. For example, there is often confusion of psychology with other mental health fields; a common question posed to psychology advocates is to explain how psychology differs from psychiatry, social work, or counseling. If psychology is not recognized as a discipline with important and unique contributions to make, then there is little incentive for policymakers and implementors to support its needs. Moreover, even when the discipline is recognized for its unique qualities in one area, for example, practice, it may be unknown for its potential to contribute to another area, for example, science. Because advocacy for psychology most often occurs within specific and limited interests rather than speaking to the discipline as a whole, the full range of the discipline often remains unknown and its real potential fragmented. For example, if taken as a whole, the discipline could be presented around a particular topic area (e.g., smoking cessation) in terms of the implications for education, research, and practice. In this way, each interest would serve to reinforce the others, the needs could be reinforced optimally and over the long term the message would find a meaningful context to the listener on a variety of issues. Yet, such an approach demands considerable coordination, assumes consistency of interests among psychologists, and requires a common understanding of and involvement in the advocacy process. None of these requirements is being met in the late 1990s. As a result, at best, federal policymakers have an uneven understanding of the field, its potential to contribute to the national good, and its many needs.

In a world of limited resources and a crowded political environment, advocacy is needed more than ever to garner the support of federal policymakers in order to establish psychology as a field of national significance of benefit to all Americans. To do this, all psychologists are needed to participate in the advocacy process. Internally, all psychologists must be educated about why advocacy is important to them and how to participate. Externally, education is needed to bring the benefits of the discipline to the public, to garner their support and, through them, the support of their representatives in Congress. To manage these efforts requires a central organization drawing together the breadth of the discipline. APA, the largest scientific and professional organization representing psychology in the United States, is positioned uniquely to play this role for the discipline. As the world's largest association of psychologists with a membership of over 150 000, the APA includes researchers, educators, clinicians, consultants, and students representing 51 different subfields of psychology and maintaining affiliation with all State Psychological Associations and territorial associations. With a mission to advance psychology as a science, as a profession, and as a means of promoting human welfare, the APA fits well as the focal point for advocacy in psychology. Organized according to the major interests of the discipline, Education, Practice, Public Interest, and Science, these structures also serve to direct its advocacy foci. Such an organization ensures attention to each interest group and its advocacy priorities. At the same time, separating these functions leads to challenging demands for the coordination of activities and priorities.

Given this organization, it is not surprising that the development of advocacy for the discipline mirrors the development of its separate parts. Thus, the more established areas like practice and science have the clearest messages and the greatest influence on the political process, both internally and externally. The newer and broader area of education is just beginning to establish its message to the discipline as well as in Congress. In an environment of relative wealth, these distinctions can work. Each interest group can focus on policy areas of greatest relevance to them, and work with those policymakers involved in their issues. As resources dwindle and the government seeks ways to limit its scope, these distinctions become troublesome; most often confusion develops around priorities.

In this way the discipline of psychology in the advocacy world suffers from its riches. The kind of direct message so central to good advocacy is not easy to achieve. Because it is a field with many subfields, a singular theme is often elusive. Because it is a discipline with at least four major interests: education, practice, public interest, and research, priorities can be difficult to establish. Education seeks to increase funds to educate and train current and future psychologists; Science works primarily with agencies to ensure that support is available for psychological research; so too Practice works most particularly on issues of parity for the psychologist practitioner; and Public Interest focuses on AIDS, Children and Families, Diversity, and like issues of public service. When legislation and policy parallels these as separate interests, work flows smoothly albeit without forcing what could be the benefit of contributions from the other interest areas. When the policy environment overlaps, for example, when the same Committee handles legislation related to two or more of these interests, separate messages can be confusing and the need to meet with the same members of Congress about different initiatives and priorities can appear as an uncoordinated strategy. Such confusion can be further exacerbated by the fact that different organizations, often with similar names, can arrive at the doorstep of the decision maker carrying different messages of needs and requests.

The next step for psychology advocacy is to be coordinated and complementary. The discipline of psychology can and will benefit from advocacy regardless of its focus. Advocacy priorities for the field need to be assessed by each interest group represented by the directorates in APA as well as the association as a whole. This process has begun, but like all collaborations, it takes time as each part learns to work with and trust the others. Further, this path of coordinated advocacy is challenged by the different levels of maturity and resources of each interest group. Whereas Practice advocacy has a long history of work, grass roots development, and active member involvement, Education is just beginning, has the fewest resources, and a constituent group (i.e., educators) largely unaware of the need of their involvement.

There is much to be done; there is enough work, and enough glory, for everyone to shine. There are great successes on which to build. Successful advocacy must be done by psychologists for psychology.

2.41.6 REFERENCES

Merriam Webster (1995). *Webster's collegiate dictionary*. Springfield, MA: Author.

Murray, C. J. L., & Lopez, A. D. (Eds.) (1996). *The global burden of disease*. Boston: The Harvard School of Public Health, on behalf of the World Health Organization.

Wolpe, B. C. (1990). *Lobbying Congress: How the system works*. Washington, DC: Congressional Quarterly.

Zorak, J. L. (1990). *The lobbying handbook*. Washington, DC: Professional Lobbying and Consulting Center.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

3.01
Observational Methods

FRANK J. FLOYD, DONALD H. BAUCOM, JACOB J. GODFREY, and CARLETON PALMER
University of North Carolina, Chapel Hill, NC, USA

3.01.1 INTRODUCTION 1
3.01.2 HISTORICAL INFLUENCES ON BEHAVIORAL OBSERVATION 2
3.01.3 THE PROS AND CONS OF BEHAVIORAL OBSERVATION 3
3.01.4 OBTAINING OBSERVATIONS OF BEHAVIOR 4
  3.01.4.1 Selection of a Setting for Observing Behavior 4
  3.01.4.2 Selection of Live vs. Recorded Observation 5
  3.01.4.3 Selection of a "Task" or Stimulus Conditions for Observing Behavior 5
  3.01.4.4 Selection of a Time Period for Observing Behavior 6
3.01.5 ADOPTING OR DEVELOPING A BEHAVIORAL CODING SYSTEM 7
  3.01.5.1 Adopting an Existing Coding System 7
  3.01.5.2 Developing a New Coding System 9
    3.01.5.2.1 Cataloging relevant behaviors 9
    3.01.5.2.2 Selecting a unit of observation 9
    3.01.5.2.3 Creating code categories 10
    3.01.5.2.4 Content validation of codes 11
    3.01.5.2.5 Developing a codebook 11
3.01.6 TRAINING CODERS 11
3.01.7 RELIABILITY 12
  3.01.7.1 Enhancing Reliability 12
  3.01.7.2 Evaluating Reliability 13
3.01.8 ANALYZING DATA GATHERED THROUGH BEHAVIORAL OBSERVATION 14
  3.01.8.1 Data Reduction 14
    3.01.8.1.1 Measurement of base rates of behavior 15
    3.01.8.1.2 Measurement of sequential patterns of behavior 15
  3.01.8.2 Computer Technology for Recording and Coding Behavioral Observations 16
3.01.9 INTERPRETING DATA GATHERED THROUGH BEHAVIORAL OBSERVATION 17
3.01.10 CONCLUDING COMMENTS 19
3.01.11 REFERENCES 19

3.01.1 INTRODUCTION

Behavioral observation is a commonplace practice in our daily lives. As social creatures and "informal scientists," we rely upon observations of behavior to understand current social experiences and predict future social events. In fact, direct observation of behavior is one of the most important strategies we use to process our social world. Thus, it is not surprising that the field of psychology also is drawn to behavioral observation as a research method for understanding human behavior. The current chapter will focus upon behavioral observation as a formal research tool. In this context, behavioral observation involves the systematic observation of specific domains of behavior such that the resulting descriptions of behavior are replicable. In order to accomplish this task, the ongoing stream of behavior must be coded or broken down into recordable units, and the criteria for the assignment of labels or for making evaluations must be objectified. These practices of specifying units of behavior and objectifying coding criteria are the key steps in translating informal behavioral observations into formal, scientific observations. As will be seen below, the challenge of employing behavioral observation in research settings involves the myriad of decisions that an investigator must make in this translation process from informal to formal observation.

3.01.2 HISTORICAL INFLUENCES ON BEHAVIORAL OBSERVATION

The development of behavioral observation methodology is attributable to two major sources, the science of human behavior and the clinical practice of behaviorally oriented interventions. The science of human behavior is often traced to Watson and his colleagues (e.g., Watson & Raynor, 1924), who developed sophisticated methods to observe the behavior of children. Several important features distinguish this approach from early research that also used the observation of human actions as data. Most notably, unlike earlier trait-based researchers who measured behavior to make inferences about internal causes (e.g., Binet, Galton), the behavior itself is the focus of study. The approach also emphasizes direct observation in naturalistic settings as opposed to contrived responses that are elicited under artificial, controlled conditions. The further development of this approach was greatly stimulated by Skinner's (e.g., Skinner, 1938) theories that emphasized a focus on overt observable behavior, and by research manuals (e.g., Sidman, 1960) that further developed the rationale and methods for conducting and interpreting research on directly observed behaviors (Johnston & Pennypacker, 1993).

The second historical influence, behaviorally oriented clinical practice, also emphasizes direct observation under naturalistic circumstances. This approach defines pathology in terms of maladaptive behavior and then focuses on how environmental factors control maladaptive responding. Thus, behaviorally oriented clinicians conduct functional analyses to determine the environmental events that elicit and maintain maladaptive as opposed to adaptive behaviors, and they focus on observable behavior change as the criterion for treatment success. The need for precision both in measuring the behavior of interest and in identifying relevant environmental events has produced a body of scholarship on behavioral assessment as a clinical tool (e.g., Haynes, 1978; Hersen & Bellack, 1998), of which systematic behavioral observation is generally considered the hallmark strategy.

Ironically, Haynes (1998) notes that despite the theoretical importance of direct observation of behavior as a central feature of the behavioral approach, most research and clinical practice from a behaviorally oriented perspective relies on indirect measures such as self-report questionnaires. This increased reliance upon indirect assessment stems in part from the recognition that demonstrating improvement in subjective well-being is important, in addition to showing changes in overt behavior (e.g., Jacobson, 1985). Also, there has been an increasing emphasis on cognitions and other internal, nonobservable experiences (e.g., fear, dysphoria) as relevant focuses for behavioral intervention. However, Haynes suspects that most researchers and clinicians fail to conduct direct observations of relevant behaviors because they believe that the difficulty associated with conducting naturalistic observations outweighs the expected benefits in terms of incremental validity of the assessment.

Accordingly, recent advances in observational methodology and technology tend to come from fields of study and practice in which the veracity of self-reports is particularly suspect. For example, researchers and clinicians working with people who have mental retardation have produced important contributions in theory and methodology (e.g., Sackett, 1979b) as well as observational technology (Tapp & Walden, 1993). Similarly, research on infants and young children continues to emphasize direct observation over other types of measures (e.g., parent or teacher report).

Another more recent source for advances in behavioral observation is the growth of the marital and family systems perspective in clinical research and practice. This perspective emphasizes interpersonal exchanges as both focuses of interest in their own right and as contexts for individual functioning. The theoretical justification for this emphasis stems from family clinicians (e.g., Haley, 1971; Minuchin, 1974) who argue that couples, parent–child dyads, siblings, and larger family units are interacting subsystems that may cause and maintain pathological responding for individuals. In behavioral terms, marital and family interactions elicit and maintain maladaptive responding (e.g., Patterson, 1982). Because individuals are often biased reporters who have a limited perspective on the operation of the family system, observation of family interactions by outside observers is necessary for understanding these family processes. Thus, research and clinical work on marriage (e.g., Gottman, 1979), parenting (Patterson, 1982), and larger family units (e.g., Barton & Alexander, 1981) paved the way for many advances in observational technology and statistical approaches to analyzing observational data (e.g., Bakeman & Gottman, 1997).

3.01.3 THE PROS AND CONS OF BEHAVIORAL OBSERVATION

Behavioral observation is a fundamental form of measurement in clinical research and practice. The function of observation is the conversion of an ongoing, complex array of behavioral events into a complete and accurate set of data that can influence scientific or clinical decisions (Hawkins, 1979; Johnston & Pennypacker, 1993). As a measurement tool, behavioral observation is systematic; it is guided by a predetermined set of categories and criteria, and it is conducted by reliable observers (Bakeman & Gottman, 1997). In many instances, behavioral observation is the most objective form of measurement of relevant behaviors in relevant contexts because it is the most direct approach. Indeed, observations provide the yardstick against which other, more indirect measures, such as self-reports or rating scales, are evaluated.

In addition to providing direct, objective measurements, behavioral observation has several other strengths as an assessment tool. Hartmann and Wood (1990) note that (i) observation is flexible in providing various forms of data, such as counts of individual behaviors, or records of sequences of events; (ii) the measurements are relatively simple and noninferential, so that they are readily obtained by nonprofessional observers; and (iii) observations have a wide range of applicability across populations, behaviors, and settings. For clinical purposes, behavioral observation produces specificity in identifying problem behaviors that are targets for intervention, and in identifying causal and maintaining variables. Thus, it supports an idiographic, functional-analytic approach in which assessment leads to specific treatment goals. Furthermore, when observation is conducted in relevant settings, the data are criterion-referenced, so that treatment effects are likely to generalize to the natural environment.

Despite these strengths, behavioral observation also has several limitations. First, the "objectivity" of behavioral observation is far from absolute. Even when relatively simple forms of behavior are observed, the observational system imposes considerable structure on how behaviors are segmented and labeled which substantially affects the nature of the data that are obtained. Second, behavioral observation is expensive and labor-intensive as compared to self-reports. Third, observation cannot access inner experiences that are unobservable. Finally, observations provide only a limited snapshot of behavior under a specific set of circumstances, which often is not helpful for generalizing to other situations. Ironically, this latter limitation reveals how sensitivity to the effects of context on behavior is both a strength and limitation of this approach. Although behavioral observation is the method of choice for examining functional relationships that elicit and maintain actions in a particular context, such observations may have limited utility for predicting responses in different contexts or circumstances.

In designing a research study, an investigator has a variety of strategies available for gathering data, including self- and other-report measures, behavioral observation, and physiological indices, and must decide which and how many of these strategies are appropriate. Whereas behavioral observation may be more direct than self-report measures in many cases, observation is only a reasonable candidate if the investigator wishes to assess how an individual or group actually behaves in a given context. Understanding the different types of information that are obtained with each measurement strategy is critical when interpreting findings.

For example, in prevention studies to assist couples getting married, investigators often teach communication skills. Then, to assess whether the couples have learned these skills, they might employ behavioral observations of the couples' interactions, as well as obtaining self-report measures from the couples about their communication. In reviewing the findings across such investigations, an interesting pattern of results seems to have evolved. In most instances, investigators have been able to demonstrate that, based on behavioral observation, couples change the ways that they communicate with each other when asked to do so in a laboratory setting. However, on self-report measures, these same couples often report that their communication has not changed (Van Widenfelt, Baucom, & Gordon, 1997). How can we interpret these seemingly discrepant findings? Some investigators would argue that the behavioral observation findings are in some sense more meaningful because they reflect how the couple actually behaves. The self-report findings are more suspect because self-reports are subject to memory distortion, might reflect the couples' overall feelings about their marriages, and are impacted by response sets or test-taking attitudes. However, the results could be interpreted in other ways as well. Perhaps the couples' behavior in the laboratory does not reflect how they actually behave at home, and this discrepancy is demonstrated by the self-report data. Conversely, the behavioral observation in the laboratory might reflect ongoing behavior at home, but this behavior does not impact the couples' experience of their overall communication. That is, perhaps the intervention did not target, or the behavioral coding system did not capture, the critical elements of communication that impact the couples' subjective experience of how they communicate with each other.

When behavior is observed, the investigator must make a series of decisions that can significantly impact the results and interpretation of the findings. First, the investigator must decide what behavior to observe. This will include a consideration of (i) the setting in which the behavior will be observed, (ii) the length of time for observing the behavior on a given occasion, (iii) the number of occasions on which behavior will be observed, (iv) who will observe the behavior, and (v) the actual behavior that will be the focus of the investigation. After or while the data are gathered, it is coded according to some scheme or coding system. Thus, in evaluating a child's social behavior, the investigator must decide whether nonverbal facial cues will be coded, what verbal categories will be included in the coding system, whether the type of interaction such as play or classroom activities will be coded, and so forth. Once the behavior is coded, then the investigator must decide how to analyze the data. Often due to limited amounts of data or for conceptual reasons, specific codes are collapsed into larger codes; for example, many specific codes may be summarized as positive or negative behaviors. Finally, the investigator must decide how to analyze the data obtained from the observational coding scheme. Is the investigator interested in the frequency of various behaviors that were observed, or is the investigator focused upon the pattern of interaction and the interdependencies among various behaviors? These different questions of interest to the investigator will result in different data analytic strategies that will provide different information about the observed behavior.

3.01.4 OBTAINING OBSERVATIONS OF BEHAVIOR

The first challenge for an investigator who wishes to employ behavioral observation is deciding upon what behavior to observe, which actually involves a series of decisions. In almost all instances, the investigator is interested in drawing conclusions about a class of behaviors. However, the investigator can observe only a sample of that behavior while wishing to draw broader conclusions. For example, a family researcher might be interested in family interaction patterns and decide to observe a family while they interact around a dinner table; however, the investigator is interested in much more than dinner-time behavior. Similarly, a marital investigator might be interested in couples' communication fairly broadly, but observes their interaction for only 10 minutes in a laboratory setting. Or someone studying children's peer relationships might observe playground behavior at school but not in the neighborhood. In all of these instances, the investigator must be concerned with whether the sample of behavior is a representative or reasonable sample of the broader domain to which the investigator wishes to generalize.

3.01.4.1 Selection of a Setting for Observing Behavior

One major decision that the investigator must make is the setting for observing the behavior. A major distinction in this regard is whether the behavior is to be observed in a controlled or laboratory setting or in a more natural setting. Both strategies have their advantages and disadvantages. Laboratory settings have the asset of being more controlled, such that the behavior of various participants can be observed under more standard conditions. For example, a variety of parents and children can be brought into a laboratory setting and observed in the same setting with the same task. This standardized environment can be of help when interpreting the findings because it can help to exclude some alternative explanations of the findings based around differences among settings. For example, observing families in their home settings might be greatly impacted by whether there are interruptions from the outside, whether the home is excessively hot or cold, and so forth.

However, typically the investigator is interested in much more than laboratory behavior and wishes to draw conclusions about behavior in other settings in the natural environment. Not only are natural environments typically less controlled, but often the effort and expense involved in observing behavior in its natural environment is prohibitive. Therefore, in deciding which setting to employ in observing behavior, the investigator must address the question of the extent to which behavior observed in one setting is generalizable to other settings. Understandably, there is no general answer to this question, and it must be evaluated for the particular settings and participants of interest.

As an example of how this issue of generalizability of behavior across settings has been explored, Gottman (1979) evaluated couples' conversations both in a laboratory setting and at home. His findings indicated that although there is a tendency for couples to be more negative with each other at home than in a laboratory setting, along with more reciprocation of negative emotion at home, the couples generally demonstrated similar interaction patterns across settings. Even so, this finding applies only to the types of interaction that Gottman assessed with his particular sample, employing a given coding system. This issue of generalizability across settings from control/research settings to natural settings applies equally to generalizability within one of these domains, as is discussed later. For example, dinner-time behavior among family members in their own home might or might not reflect family interaction later in the evening when homework or bedtime activities are the focus of the family's discussion. Thus, even within a natural environment, behavior in one aspect of that setting might not reflect behavior in other aspects of the natural family environment. Consequently, the investigator should give a great deal of thought to the setting in which the behavior is observed in order to increase the likelihood that the resulting behavioral observations access the behavior of interest.

3.01.4.2 Selection of Live vs. Recorded Observation

The setting is altered by the experimenter when coders or recording devices are introduced into the environment. The decision to have coders present for the actual observation session, or instead to record the session for later coding, raises two major concerns. The first concern is the reactivity of live observation and recording equipment. Although research on this topic has a long history (e.g., Haynes & Horn, 1982; Kazdin, 1982), the findings are mixed and do not specifically address the relative effects on reactivity of having an observer present as opposed to using video or audio recording equipment. Much research shows that the presence of a live observer alters the behavior of subjects (Kazdin, 1982), and this reactivity may be greatest in intervention studies when the demands to display desired behaviors are relatively clear (e.g., Harris & Lahey, 1986). However, research on the reactivity of recording equipment is less certain. For example, studies using repeated sampling with recording equipment fail to detect habituation effects (e.g., Christensen & Hazzard, 1983; Pett, Wampold, Vaughn-Cole, & East, 1992), which suggests that the equipment does not evoke an initial orienting response. Further, studies that compare different recording methods show that relatively obtrusive as opposed to unobtrusive procedures produce few differences in the quality of most behaviors observed, although positivity may be increased somewhat (e.g., Carpenter & Merkel, 1988; Jacob, Tennenbaum, Seilhamer, Bargiel, & Sharon, 1994). Of course, it is possible that the mere presence of any type of recording equipment (or knowledge that it is present although unseen) may cause sustained changes in behavior similar to the effects of self-monitoring or participant observation (e.g., Jarrett & Nelson, 1984). Nevertheless, this set of studies suggests that using recording equipment with no observer present may be a relatively less reactive approach than live observation of behavior.

Another concern about live as opposed to recorded behavior is the accuracy of coded data. In general, we assume that video and audio recorded data help to improve coder accuracy because they provide the capacity to play back events repeatedly that are ambiguous or happen quickly. However, recorded data also may interfere with a coder's ability to attend selectively to salient behavior, particularly in a setting in which there is considerable background activity and noise. For example, Fagot and Hagen (1988) found that coders evaluating children in a preschool setting were less reliable and tended to miss relevant events when they coded from videotape as opposed to live observations. In part, the superiority of recorded observations depends on the ability to obtain excellent recordings. In many circumstances, audio recordings are particularly problematic because behavior is ambiguous without visual cues. When audio recordings are transcribed, some investigators also include live observers who make notes about nonverbal events that are added to the transcript.

3.01.4.3 Selection of a "Task" or Stimulus Conditions for Observing Behavior

Not only must investigators decide upon the physical setting for observing behavior, but the task or stimulus conditions within which the behavior is to be observed also must be decided. On many occasions, investigators ask the participants to engage in a specific task; on other occasions, the investigator merely decides to observe the participants' behavior in a given setting at a given time. If a particular task or interaction is structured by the investigator, then the effects of this particular task on the interaction must be considered. This is of less importance if the investigator is interested only in this particular type of interaction. For example, if an investigator is interested only in how a family would make weekend plans as a full family if asked to do so, then asking the full family to have such a discussion is relatively straightforward. However, the investigator might be interested in some more general question having to do with how the family makes decisions. If this is the issue of interest, then the investigator must carefully consider, and hopefully assess, the impact of this particular task involving planning a weekend. Deciding how to spend the weekend might or might not generalize to how the family makes decisions in other domains, such as how to divide household chores. Indeed, asking the family to sit and have a discussion resulting in weekend plans might not mirror how decisions are made in the family at all. Perhaps the parents make such decisions, or perhaps these decisions occur over the course of a number of informal interactions with different persons present at different times. More generally, when investigators structure particular types of interactions or ask the participants to engage in specific tasks, they must carefully consider whether the task or stimulus conditions that they have created mirror the ways in which the participants typically behave. Either the participants might behave differently when the task is different, or they might not typically engage in the task or situation that is constructed. The degree of concern raised by these issues is a function of the degree to which the investigator wishes to describe how the participants typically behave in their day-to-day lives.

How the task is selected also might impact the behavior. For example, Christensen and Heavey (1990) have described different interaction patterns among husbands and wives. This includes a well-known interaction pattern that they label "demand-withdraw," in which one partner presses for the discussion of a topic, and the other partner attempts to withdraw from the interaction. A number of investigations indicate that females are more likely to assume the "demand" role, and husbands are more likely to assume the "withdraw" role in marital interactions. However, Christensen and Heavey found that husbands were more likely to engage in the demand role during problem-solving interactions when the husbands selected the topic of conversation, compared to interactions in which the wife selected the topic to discuss. Thus, factors that influence an individual's interest in a task or motivation to participate in the task might significantly influence the resulting behavior that is observed.

3.01.4.4 Selection of a Time Period for Observing Behavior

In making decisions about the representativeness of the behavior that is observed, the investigator must also be concerned with the degree to which the observed behavior is generalizable across time. Classical test theory indicates that longer "tests" are more reliable in the sense that, keeping all other factors constant, they generally include a more representative sample of behavior and thus are more stable across time. In terms of behavioral observation, this raises two questions. First, how long should behavior be observed on a given occasion; second, on how many different occasions should behavior be observed? Whereas the answers to these questions should be based upon empirical findings, often pragmatic and logistic issues influence investigators' decisions along these lines. For example, sending observers into persons' homes or into classrooms can be troublesome and intrusive; sending observers to Africa to follow the social interactions among baboons can prove to be a logistical nightmare. Consequently, the difficulty, intrusiveness, and expense of behavioral observation often are a limiting factor in deciding how long to observe behavior. Similarly, the amount of time required to employ certain coding systems limits the length of behavioral observation. For example, some coding systems might require 20 hours to code one hour of recorded behavioral interaction. Therefore if the investigator intends to employ such a coding system, he or she might limit the amount of observed behavior to short time periods.

In deciding how long to observe behavior during a given observation period, several factors come into play. First and most generally, a long enough time period is needed such that the findings are relatively stable and replicable on other occasions. In part this is related to the frequency or base rate with which the behaviors of interest occur. If the behavior of interest is a relatively infrequent behavior, then longer observation periods are needed to obtain a stable assessment of its occurrence. However, if a behavior occurs with a high degree of frequency, then shorter observation periods can suffice. Second, the length of the observation period is influenced by the "complexity" of the phenomenon under consideration. For example, the investigator might be interested in whether there are different stages or phases in a couple's conversation as they attempt to reach resolution to a problem; in this instance, it would be essential to observe the entire problem-solving interaction. Or the investigator might be interested in whether there are different stages or phases in how a child responds to a new setting with the mother present. Attachment theory has explored this question and has provided indications of how securely and insecurely attached children initially respond in such settings, how they venture forth into a room after an initial exposure, and how their interaction with their mothers changes across time in this setting (Ainsworth, Blehar, Waters, & Wall, 1978). Therefore, if the investigator hypothesizes or wishes to explore whether there are different stages or phases in the interaction, this necessitates following the interaction for a sufficient time period to allow for an examination of the different phases of interest.

Second, the investigator must decide on how many occasions to observe the behavior. A given observation period might be significantly influenced by the occurrence of a major event, or the interaction of participants being observed might proceed in a given manner based upon what happens early in the interaction. For example, if a child is taunted on the playground early during a recess period, the child might withdraw from the group, which will significantly impact the child's behavior for the duration of the observation period on the playground. If the child is not taunted the next day, then his or her interaction pattern might proceed quite differently. Consequently, the number of occasions on which to observe interaction will differ according to the variability in the behavior being observed. If the behavior of interest occurs in a relatively consistent manner across different occasions, then fewer occasions are necessary for obtaining a stable sample of behavior.

Whereas there are far too few empirical investigations that have been conducted to determine the number of observational sessions needed to obtain stable results, some such investigations do exist. For example, Wieder and Weiss (1980) explored how many observational sessions were needed in order to obtain a stable assessment of couples' interaction when employing the Marital Interaction Coding System (MICS; Weiss, 1986). Based on generalizability theory, they concluded that observing couples on a single occasion could provide meaningful information about the couple, but observations across two separate evenings provide more stable interactions patterns when coded by the MICS. Similarly, Haynes, Follingstad, and Sullivan (1979) found that across three evenings, there was high stability on only 5 of 10 selected coding categories of communication between spouses. Interestingly, in spite of these findings, no marital therapy outcome investigations have observed couples' interactions across two or more consecutive evenings.

As can be seen based on the above discussion, there are a number of questions that the investigator must address in deciding what behaviors to observe. The decisions that are made will certainly impact the investigator's findings. Yet before these findings are obtained, there are many more decisions for the investigator to make that will influence the results.

3.01.5 ADOPTING OR DEVELOPING A BEHAVIORAL CODING SYSTEM

Foremost among the additional decisions to be made is the choice of coding system to employ to classify the data that have been observed. In fact, in order to address the questions raised above regarding what behavior to assess, the investigator needs to know ahead of time what coding system he or she will employ. At times the behavior is coded live during the interaction, so the coding system must be known. Even if the observed behavior is to be coded later based on video recordings of the behavior, it is important to decide ahead of time what coding system will be employed. For example, some coding systems might be appropriate only in certain settings or with certain tasks. Similarly, some coding systems might break behavior into a number of fine-grained categories that occur on a somewhat infrequent basis, thus necessitating longer observational periods. Therefore, deciding what coding system to employ is a significant factor in developing a thoughtful study based on behavioral observation.

An initial consideration in selecting a coding system is whether to adopt an existing coding system or to develop a new one. Each approach has some obvious benefits and limitations, as well as other subtle impacts that may not be apparent to investigators until they are well into the task of coding the behavioral observations.

3.01.5.1 Adopting an Existing Coding System

In many cases, it may be possible to adopt a coding system that has been used previously in the same research area, or one that can be imported from another area where similar constructs were assessed. Adoption of an existing system has the great advantage of saving the time and energy required to develop a reliable, valid, and practical coding scheme. It also links research across laboratories, samples, and locations, and thus provides for ready synthesis of research findings from various studies.

The selection of a coding system is guided by both theoretical and practical considerations. The primary theoretical issue is whether the coding system assesses behaviors that address the constructs of interest to the investigator. All coding systems organize data into some set of categories or units for coding, and these categories are based on issues of importance to the person who developed the coding system; however, they might or might not coincide with another investigator's concerns or theoretical model. Before beginning a search for a coding system, it is essential that the investigator first review theory and research to clarify the nature of the behavioral phenomena under investigation.

Behavioral phenomena related to a construct under one situation may take on different characteristics in a new situation, thus making an existing system inappropriate. For example, Ammerman, Van Hasselt, and Hersen (1991) coded problem-solving interactions between children with visual impairments and their parents using the Marital Interaction Coding System, III (MICS III; Weiss, 1986), a well-validated system for assessing problem-solving interactions within marital dyads. The study detected no differences between groups of families with and without children with disabilities, possibly because the coding system was inappropriate for the family context. Whereas the MICS III evaluates warm and hostile exchanges that indeed were similar across the groups, it does not assess factors such as behavior management, instruction, or other socialization practices that are important aspects of parent–child exchanges that are responsive to children's disabilities (Floyd & Costigan, 1997). Thus, a behavioral coding system may be sensitive to relevant variables only for certain types of people, relationships, or circumstances.

In addition to the substantive content of the system, various coding systems differ in the ways that they "chunk" or segment behavior into coding units. This "chunking" of behavior has both theoretical and practical implications for investigators. From a theoretical perspective, the nature of the phenomena being assessed should influence how the stream of ongoing behavior is segmented (Floyd, 1989). More specifically, relatively small, elemental units of analysis may be appropriate for evaluating discrete events that require little inference; however, larger units of behavior might be needed to capture more complex phenomena. Foster, Inderbitzen, and Nangle (1993) discuss a similar point regarding the selection of a coding system to evaluate the effectiveness of social skills training with children. They note that a frequent problem with interpreting the results of treatment outcome studies is that whereas the treatment teaches children specific social skills, such as offering to share a toy, the observational assessment evaluates only molar, or global codes, such as "positive interactions." From data such as these, it is impossible to know whether the behaviors that were trained actually were displayed during the assessment.

Alternatively, it also is important to question whether a complex phenomenon is accurately evaluated by merely summarizing elemental codes. For example, Jacob (1975) illustrates how power in interpersonal interactions may not be indicated by who speaks more frequently or wins more disagreements, but rather by the ability to prevail on the central, important conflicts. Evaluations such as these may require making judgments about relatively large units of behavior.

Choosing among systems with different units of observation also has practical implications. Microanalytic coding systems that parse behaviors into minute elements may be overly complex and labor-intensive for investigators who merely want to assess global levels of characteristics, such as positiveness, competence, anxiety, or irritability. In such cases, it may be more efficient to use a system that rates dimensions such as the intensity or quality of behavior exhibited over an extended period of time. On the other hand, small, elemental units of observation and analysis are useful for detecting situational influences on behavior and sequential patterns among minute events; larger, more integrative units are useful for understanding cross-situational consistency (Cairns & Green, 1979). Thus, findings based on larger units of observation may be more generalizable than microanalytic coding. Some investigators appear to assume that behavioral observation is synonymous with microanalytic coding; such assumptions can serve as a major impediment to the more widespread use of observational measures in research settings with limited resources. We encourage investigators to explore macroanalytical coding procedures as a practical and, in some cases, more informative alternative to microanalytic coding.

Every coding system incorporates an array of assumptions, biases, and procedural preferences that the originator used to guide coding decisions. These preferences are particularly relevant in decisions about how to handle ambiguous incidents. Many decision rules are not made explicit in published reports and coding manuals, so that it is difficult for investigators who did not participate in its development to apply an existing system accurately and in a way that is consistent with other research. Whenever possible, consultation with the originator is invaluable. Some originators of popular measures conduct periodic workshops to train new users (e.g., SASB, Humphrey & Benjamin 1989; SPAFF, Gottman, 1994). Most developers cannot be expected to provide ongoing consultation to support their system, but should be willing to share examples of coded data and advice about common problems.

3.01.5.2 Developing a New Coding System

New ideas focused on new constructs and employing new assumptions are the lifeblood of progress in the social sciences. Thus, there will always be a need to develop new coding schemes. Even when well-known constructs are studied, if observational procedures become highly standardized within a research area, the phenomenon of interest may become totally dependent on the existing measure. This situation can lead to replicated, well-established findings that are largely an artifact of a particular measurement procedure. The need to disentangle method variance from the phenomenon of interest is strong justification for the development of new coding systems.

Detailed instructions about developing coding systems are given in Bakeman & Gottman (1997) regarding interpersonal interaction, and by O'Neill, Horner, Albin, Storey, and Sprague (1990) regarding functional analysis of problem behaviors. The key steps in developing any coding system are summarized below.

3.01.5.2.1 Cataloging relevant behaviors

A useful initial step is to develop an exhaustive list of all relevant behaviors to be coded. In some cases, this may be accomplished by conducting initial observations and recording all relevant behaviors that occur. Animal researchers frequently use this procedure to develop an ethogram, which is a comprehensive list of all behaviors that are characteristic of a species. Several ethograms are published to describe the behavior repertoire of different animal species. However, because it is usually impossible to sample all possible behaviors for a species, investigators estimate the quality of sample coverage with the theta statistic, calculated as 1 – (number of different behaviors seen/total number of acts observed). As the value of theta approaches 1, the probability of encountering a new behavior approaches zero. That is, we assume that if new behaviors are not encountered with additional observations, the behavioral repertoire has been adequately sampled.

Of course, a strictly empirical approach such as this usually is not adequate for evaluating human behavior. As we noted at the beginning of the chapter, a stream of human behavior often can be classified according to an enormous variety of characteristics. In order to focus attention on a limited set of characteristics, the investigator should begin with a list of these characteristics and their manifestations as gleaned from previous research, experience, and theory. Pilot observations then can be directed toward refining this list by broadening some categories, tightening others to make finer distinctions between theoretically disparate behaviors, and adding new categories not suggested by previous research. For an excellent example of this process, see Jacob, Tennenbaum, Bargiel and Seilhamer's (1995) description of the development of their Home Interaction Scoring System.

One frequent concern while developing coding systems involves what to do with rare but theoretically important events. Rare events tend to decrease the reliability of coding systems; however, such rare events may be highly meaningful, and thus they cannot be excluded from a system without compromising validity and utility. It may be possible to collapse similar rare events into broader categories or to alter the observational situation in order to elicit these behaviors more consistently and reliably.

3.01.5.2.2 Selecting a unit of observation

An important component of identifying relevant behaviors is to determine the appropriate unit of observation. This involves the decision to evaluate behavioral states, events, or some combination of the two. In general, a state is any ongoing condition that persists over an extended period of time, whereas an event is a discrete action. States are usually measured with time-based indices such as duration or latency, whereas events are usually measured with frequency counts or sequential patterns. Both types of unit also can be rated for intensity. The distinction between states and events is blurred by the fact that the same behavior might be measured with both units, such as measuring both the duration of anxiety episodes or disruptive episodes in the classroom, as well as the frequency of the episodes. The type of unit is not always mandated by the content of the behavior and, once again, the decision about the appropriate unit for a particular purpose must be guided by theoretical, empirical, and practical considerations. At first glance, it may appear that recording onset and offset times for all behaviors is desirable so that information about duration and latency can always be retrieved. However, Bakeman and Gottman (1997) warn against "the tyranny of time" and propose that, even with sophisticated recording and analytical devices, the precise measurement of time adds substantial complexity to data recording and analysis, and can cause problems with reliability that outweigh the benefits of having these data.

The unit of observation also involves the sampling protocol. The two most common sampling protocols in clinical psychology research are event sampling, which involves noting each occurrence of events during the entire observation period, and time sampling or interval coding, which involves noting occurrences during selected time intervals. Most commonly, time sampling involves alternating periods of watching and recording, each lasting for several seconds. During the recording period, the coder usually records whether or not each possible code occurred during the preceding interval. This procedure is useful for live observations in which the coder must track a large number of behaviors, so that recording each event would interfere with accurate observation. The procedure assumes that recording periods are random, and will not distort the data. Several studies challenge this assumption and reveal that interval coding can distort the amount of behavior, sequential patterns, and observer agreement for both behavioral events (e.g., Mehm & Knutson, 1987) and behavioral states (e.g., Ary, 1984; Gardner & Griffin, 1989). However, other studies demonstrate that interval coding can accurately reflect actual behavior rates and durations (e.g., Klesges, Woolfrey, & Vollmer, 1985). A study by Mann, ten-Have, Plunkett, and Meisels (1991) on mother–infant interactions illustrates that the accuracy of data from interval sampling depends on how well the duration of the sampling interval matches with the duration of the behavior of interest. They found that the actual durations or frequencies of mother and infant behaviors, which tended to occur in relatively short episodes ("bouts"), were inaccurately assessed when the sampling interval was relatively long. Thus, before proceeding with a time-sampling/interval-coding approach, the investigator should test the time-sampled data against actual rates or durations, and adjust the length of the recording interval to produce the most accurate data possible. See Altmann (1974) for an extensive review of sampling protocols, and Bakeman and Quera (1995) for considerations about how to design sampling protocols and record the data for data analysis purposes.

3.01.5.2.3 Creating code categories

Typically, code categories will be mutually exclusive and exhaustive Mutual exclusivity means that coding categories are discrete and homogeneous, and that each event can be classified into one and only one category. Exhaustiveness means that all manifestations of a construct are included in the system. In most cases, exhaustiveness can be achieved by including a category, such as "other" or "neutral," to label all behaviors that do not fit well into any other category. For example, a measure of parent–adolescent interaction by Robin and Foster (1989) includes a "talk" code to cover all behaviors that are not instances of other categories. In some cases, the investigator may believe that it is necessary to violate these guidelines. For example, another measure of parent–adolescent interaction, the Constraining and Enabling Coding System (Leaper et al., 1989), allows behaviors by parents to receive both "constraining" and "enabling" codes because the authors believe that these "mixed messages" may be highly relevant to adolescent functioning. In other cases, investigators only label certain subsets of behaviors (i.e., as in scan sampling where, for example, only instances of target children's aggressive behavior and provocations by peers are recorded). Both situations create difficulties for the analysis of sequences of events, although Bakeman and Quera (1995) provide some solutions for managing these types of data.

Often it is useful to organize codes into groups and, if appropriate, to arrange the groups into a hierarchical classification. This arrangement makes the codes easier to recall; in addition, this hierarchical arrangement can help to fashion a set of decision steps to employ in the coding process, both of which may improve observer reliability. For example, children's social behaviors might first be classified as initiations versus reactions, and initiations could be classified as prosocial versus antagonistic, followed by more specific categories of events within each of these general categories. The coders can then use this hierarchical arrangement to organize their decision process, so that once they decide that a particular action is an initiation, and it is prosocial, there is a relatively small number of "prosocial initiation" codes to choose among.

An alternative approach to forming a hierarchical organization is exemplified in the Structural Analysis of Social Behavior (SASB) system (Humphrey & Benjamin, 1989). This system uses a rating scale format as an aid to categorical coding. In this system, all code categories are classified within a circumplex defined by two dimensions: Interdependence and Affiliation. Coders receive extensive instruction in the theory underlying each dimension, and they learn to rate behaviors in terms of their degree of interdependence, represented on a vertical axis, and their degree of affiliation, represented on a horizontal axis. The axes intersect at their midpoints. The location on the circumplex defined by the coordinates of these dimensional ratings is the code for that act. Thus, for example, a behavior rated as +5 for Interdependence (i.e., somewhat independent) and −4 for Affiliation (i.e., somewhat disaffiliative) is assigned to the category "walling off and distancing."

3.01.5.2.4 *Content validation of codes*

A useful, though often neglected, step in coding system development is content validation of the codes. This might be accomplished by having "experts" or the actors themselves classify codes into relevant categories, then comparing these categories to the expected categories. For example, in order to evaluate a family coding system that classified family behaviors as aversive and nonaversive, Snyder (1983) had family members rate the aversiveness of concrete examples of behaviors classified by the coding system.

3.01.5.2.5 *Developing a codebook*

After codes are labeled, defined, and classified, the next step is to produce a codebook. In general, most experts recommend that the more thorough, precise, and clearly written the codebook, the better the chances of training new coders to produce reliable, valid data (e.g., Hartmann & Wood, 1990; Herbert & Attridge, 1975); however, other studies demonstrate that naive coders can at times produce valid codes (e.g., Prinz & Kent, 1978). The codebook should include a list of all codes, a descriptive definition for each code, and examples of behaviors that represent each code, along with examples that do not match the code. In order to assist with reliable coding, it is particularly important to include examples of differential decisions in borderline or ambiguous cases. The APA guidelines for educational and psychological tests (APA, 1985) lists several other types of information that would also be useful to include in codebooks, including information about the theoretical underpinnings for the measure, and data on reliability and validity from research to date. Herbert and Attridge (1975) propose that providing this type of information to coders may help to facilitate training and improve coder reliability.

3.01.6 TRAINING CODERS

Once a coding system has been adopted or developed, the next step is to train coders in its use. A preliminary step in conducting efficient and effective coder training is the selection of coders who will perform well. Unfortunately, there is little research on personal characteristics or abilities that predict good performance as a behavioral coder. Research on interpersonal communication indicates that, in some circumstances, women tend to be more accurate than men at decoding the meaning of interpersonal behaviors (e.g., Noller, 1980). However, this effect is hardly large enough to warrant the exclusion of male coders in studies of interpersonal behavior. To the extent that coder characteristics such as gender, age, education, or ethnicity may be associated with biases that could influence coding decisions, it may be important to ensure that coders are diverse on these characteristics in order to randomize these effects and improve the validity of the data.

Ironically, one characteristic that may be important to select *against* is prior experience or personal investment in the research area under consideration. The use of naive, uninvolved observers controls for biases caused by factors such as familiarity with the research hypotheses and prior expectations. We believe that in many instances naive observers also tend to provide more reliable codes. Extensive research on clinical decision making demonstrates that highly experienced judges tend to employ idiosyncratic and inconsistent decision criteria that can reduce both intraobserver consistency and interobserver agreement as compared to naive observers (e.g., Dawes, 1979). Our experiences bear this out. When coders have extensive previous training or experience in the domain of interest, they tend to have difficulty adhering strictly to the coding criteria outlined in the coding manual, particularly if their experiences involved a different theoretical perspective or set of assumptions than those that undergird the coding system.

Wilson (1982) wisely proposed that observer training should address two equally important goals: teaching the skills needed to perform the coding task, and motivating the coders to perform well. Teaching the skills usually involves a combination of didactic and experiential training sessions. A helpful first step is to explain the rationale and theory that underlie the coding system. In cases where the coding involves considerable judgment in the application of a theory or model, the coders might benefit from readings and instruction about the model. The coders should read and demonstrate a thorough working knowledge of the coding manual and should be tested on this material before proceeding. Practice with the coding system should begin with the presentation of examples of behaviors, followed by an explanation for the coding decisions. Initial examples should be relatively unambiguous representations of the coding categories or examples of the extremes of the rating scales. The coders should be required to practice the coding with feedback about their accuracy and discussion of the rationale for coding decisions. Of course, the practice materials should match the actual coding situation as closely as possible. Training sessions should be relatively frequent and relatively brief to enhance alertness. Training continues until the coders reach acceptable levels of agreement with preestablished criterion codes. Because accuracy can be expected to decrease once actual coding begins (e.g., Taplin & Reid, 1973), most investigators set a training criterion that is higher than the minimal acceptable criterion during actual coding. Typically, this criterion is 80–90% agreement.

The maintenance of the coders' motivation to perform well probably involves the same types of factors that enhance performance in any work setting, including clarity of the task, investment in the outcome, personal responsibility, monitoring of performance, and a fair reward structure. One typical procedure used by investigators is to develop a contract that specifies the tasks to be completed, the expectations, and the reward structure. Coder investment in the project might be enhanced by providing them with the opportunity to participate as a member of the scientific or clinical team, to the extent that this does not bias coding, or by group interactions that build solidarity and cohesion among the coding team. As described below, reliability should be monitored on an ongoing basis. An unfortunate feature of reliability monitoring is that there is a built-in punishment schedule for inadequate performance, such as having to recode sessions or completing additional training, but the rewards for good performance are less tangible. Investigators should be sensitive to the need to instigate a reward system in whatever way possible, including providing raises for a good record, greater responsibility in the form of training and monitoring new coders, or public acknowledgment of good work.

3.01.7 RELIABILITY

In observational research, the reliability, or precision, of the measure is almost always evaluated with indices of interobserver agreement. Actually, according to psychometric theory, reliability concerns the extent to which coded data map onto "true scores," and thus, the reliability of coded data also relates to intraobserver consistency in applying the coding scheme (Bakeman & Gottman, 1997). However, because agreement between independent coders is a higher standard for judging precision, it is the focus of formal psychometric evaluations, and intraobserver consistency is addressed more informally through the implementation of training and monitoring procedures to prevent observer drift.

3.01.7.1 Enhancing Reliability

As noted throughout the previous sections, reliability can be enhanced in the way the codes are developed, in the procedures for training and monitoring the observers, and in the procedures for conducting the observations. Regarding code development, greater reliability can be expected when codes are clearly defined in operational terms, when categories are relatively elemental and overt as opposed to complex and inferential, and when coded behaviors occur at least moderately frequently. If some of these conditions cannot be met, coders likely will need relatively more training, practice, and experience in order to produce reliable data. Instructions that explicitly discourage or encourage specific expectancies about frequencies or patterns of codes either tend to reduce observer agreement or spuriously inflate it (Kazdin, 1977). Similar to guidelines for training sessions, frequent and relatively short observation sessions produce more reliable data than infrequent, longer observation sessions. In addition to booster sessions to reduce observer drift, periodically training new coders may help to reduce the biasing effects of prior experience on the data, and observations involving different subject groups or experimental conditions should be intermingled whenever possible to avoid systematic effects related to observer drift. Studies repeatedly demonstrate that coders are less reliable when they believe that their agreement is not being checked, although this effect may abate with increased experience (e.g., Serbin, Citron, &

Conner, 1978; Weinrott & Jones, 1984). Therefore, frequent, overt, and random assessments of reliability should help to maintain coder precision.

3.01.7.2 Evaluating Reliability

The appropriate procedure and calculations for evaluating reliability depend on the purposes of the evaluation and the nature of the inferences that will be drawn from the data. Thus, there is no one way to evaluate reliability, but rather an array of approaches that can be followed. Helpful reviews of various procedures and computational formulas are presented by Bakeman and Gottman (1997), Hartmann (1977, 1982), Jacob, Tennenbaum, and Krahn (1987), Stine (1989), and Suen, Ary, and Covalt (1990). Two important decisions include (i) whether to compute exact agreement for specific codes (point-by-point agreement) or aggregate agreement for larger units, and (ii) whether and how to correct for chance agreement. Below is a summary of some of the more common procedures, along with guidelines for selecting among them.

For the purpose of training, monitoring, and providing corrective feedback to coders, in most cases it is useful to assess the most precise form of point-by-piont agreement that is possible with the coding system. For example, with event recording, an observer's codes for each event are compared with a set of criterion codes or those of an independent coder. A contingency table that cross-lists the two sets of codes and tallies the frequencies of each pair of codes (i.e., a "confusion matrix") is a useful method of feedback for the coders. The total percent agreement for all codes (# agreements/total # events), and the percent agreement for specific codes (# agreements/(#agreements + #disagreements)) are summary scores that are easy to understand. During coder training, we compute these statistics beginning with the first practice assignments because the rapid improvement that usually occurs during the first few assignments can be highly rewarding and reinforcing for new coders. The contingency table also can display a pattern of disagreements that is instructive, such as when two coders consistently give a different code to the same type of behavior. Even when the actual observational procedure involves a larger or less precise unit of observation, such as in interval sampling, it may be helpful to employ an event-based tracking procedure during coder training in order to have more precise information about when disagreements occur. In our experience, the identification of consistent disagreements is helpful when follow-up instruction is used to clarify and resolve the confusion about the coding criteria. However, providing feedback about disagreements without instruction or resolution can make coders feel more confused and uncertain and can decrease reliability. Also, it usually is not helpful to coders to correct their scores for chance agreement, because corrected scores may be highly variable depending on the range of behaviors displayed in each session; thus they present a confusing picture of absolute agreement. Finally, if agreement statistics are used to identify unreliable coders in need of additional training, it is important to base this decision on multiple observation sessions, because occasional unreliable sessions can be expected due to the ambiguity in the target's behavior rather than because of coder error.

For the purpose of determining the precision of the measurement after coding has been completed, the method of computing reliability depends on the type of data that will be analyzed (Hartmann, 1977). That is, the investigator typically is interested in the reliability/precision of the scores that are actually computed from the data. Although it is common for investigators to report some form of point-by-point agreement for individual codes, most data analyses focus on aggregate indices such as the frequency, relative frequency, or rates of groups of behaviors. Thus, if several specific codes are combined or aggregated into a broader category of "positive behavior" for data-analytic purposes, then reliability estimates should be performed at the level of "positive behavior." We usually assume that observer agreement for specific codes ensures that aggregate scores are reliable; nonetheless, it is useful to determine the actual level of reliability for these aggregate codes. Furthermore, agreement at the specific code level is usually an overly stringent requirement that can unnecessarily prolong a study.

For the purpose of assessing the precision of the measure, percent agreement statistics are usually corrected for chance agreement between coders by using Cohen's kappa statistic. This statistic uses the information about agreements and disagreements from the confusion matrix, and evaluates the observed amount of agreement relative to the expected amount of agreement due to chance because of the base rates of the behaviors. The formula for kappa is: kappa = [p(Observed agreement) $-$ p(Expected agreement)]/[1 $-$ p(Expected agreement)], where p(Observed agreement) is the percent agreement for the two coders, and p(Expected agreement) is the percent agreement expected by chance. The probability of chance agreement is computed by calculating the relative frequencies for each code by each coder, obtaining the product of the two coders' relative frequency scores for each code, then summing these products. Usually, kappa is computed for each reliability session, and the mean and range across all reliability sessions are reported. One complication with the kappa statistic is that it can severely underrepresent observer agreement in sessions during which the subject displays a limited number of behaviors and/or some behaviors have exceptionally high base rates. This situation produces very large values for expected agreement, and thus can produce a very low score for kappa even when the proportion of observed agreement is very high. A potential solution in this situation is to use a sample-wise estimate of expected agreement derived from the base rates for the entire set of data (Bakeman & Gottman, 1997). Another complication is that point-by-point kappa may be overly stringent when, as frequently occurs, the actual scores used in the data analysis are aggregates (e.g., total frequency scores). Jacob et al. (1987) present a method for computing kappa for aggregate scores.

Often the measure of interest in the data analysis represents the data on an interval scale of measurement (e.g., relative frequency scores or average ratings). In this situation, the question of measurement precision concerns the relative positions of subjects on the interval scale rather than agreement between coders on individual coded behaviors. Coder agreement for interval data can be assessed with the correlation between the scores for pairs of coders; for example, the correlation between summary scores for a group of subjects, or the correlation between interval ratings within an observation session. Alternatively, the intraclass correlation is also appropriate for interval data. This approach is derived from the analysis of variance (Winer, 1971), and is the most commonly used application of generalizability theory for evaluating coder agreement. It assesses the proportion of variance in the scores that can be attributed to variation among subjects as opposed to variation among observers. An advantage of the approach is that when more than two coders are being evaluated, one score can reflect the level of agreement (i.e., proportion of variation attributed to differences among subjects) for the entire group of coders, rather than calculating individual correlations for all possible pairs of coders. Shrout and Fleiss (1979) outline the procedures for computing intraclass correlations under various assumptions about the data.

Alternative procedures for detecting agreement between coders may also be appropriate when the focus of study is the sequential patterns in the behavioral stream. Bakeman and Gottman (1997) vividly illustrate how point-by-point agreement between coders can be sharply deflated when one coder occasionally inserts extra codes, although the agreement about the sequential pattern of the codes is very high. Wampold and Holloway (1983) make a similar case that agreement about individual codes may be too stringent a criterion for sequential data. Gottman (1980) recommends an approach similar to intraclass correlation in which the investigator demonstrates that the measures of sequential dependency (e.g., lag z-scores, see below) are consistent across observers.

3.01.8 ANALYZING DATA GATHERED THROUGH BEHAVIORAL OBSERVATION

Apart from the usual decisions that guide the design of data analysis, observational data present some unique challenges for investigators. These include questions about data reduction, dimensions and scales of measurement, and the identification of sequential patterns of behavior.

3.01.8.1 Data Reduction

Most observational studies produce a large amount of data for each subject, particularly when the coding scheme exhaustively labels all behaviors emitted by one or more subjects during an observational session. Even sessions lasting only 5–10 minutes can produce several hundred codes when behavior units are small. The challenge is to reduce these data to make them interpretable, without sacrificing the rich descriptive information in the coded data.

A first step is to group individual codes into broader categories. For the purposes of data analysis, the primary problem with many observational coding schemes is that they include behavioral events that rarely or never occur for many subjects, and thus produce highly skewed distributions of scores for the sample, resulting in missing cells in contingency tables for sequential analyses. Whenever possible, it is important to incorporate rare events into categories of events that occur with sufficient frequency to be used in subsequent analyses. In most cases, categories are specified on an *a priori* basis by investigators, using theory and rational analysis to determine how codes are grouped.

Another approach is to factor analyze individual codes to identify clusters of codes that share common variance. This approach is probably most appropriate when the investigator is interested in examining behavior styles, because the factor analysis identifies groups of behaviors that covary across subjects. That is, behaviors are grouped based on their co-occurrence within subjects, irrespective of their meaning or functional impact. For example, if different children in classrooms tend to misbehave in different ways, with some children spending much time out of their seats, others talking out of turn, and others being withdrawn and inattentive, a factor analysis would likely identify these three types of behaviors as three separate clusters. Kuller and Linsten (1992) used this method and identified social behaviors and individual concentration as separate clusters of behaviors by children in classrooms. An alternative approach is to group behaviors according to their functional impact. If all three types of behaviors disrupt the activities of other children in the classroom, they can be grouped into one functional category of disruptive behavior. For example, Schaap (1982) used lag sequential analysis of marital interactions to identify a set of behaviors that were most likely to elicit negative responses from a spouse. These specific behaviors could then be grouped together as negative eliciting behaviors.

3.01.8.1.1 Measurement of base rates of behavior

Base rates of behavior can be expressed in various scales of measurement. For example, the frequency of events can be expressed as raw frequency, relative frequency, rate per minute, or ratio scores (i.e., ratio of positive to negative behaviors). The selection of a scale of measurement is usually determined by factors such as the focus of the investigation, the precedent in the field of study, and psychometric properties of various scales of measurement (e.g., distributions of the scores, reliability of the index). Tryon (1991) argues that scores should also be expressed in "natural" scientific units in order to make measurements as comparable as possible across time, situations, or different investigations. A compelling point in this argument is that relativistic measurements such as standardized *z*-scores, which provide information about the position of the subject relative to other subjects in the same study, are highly "unnatural" units that vary as a function of who participates in the study. Instead, scores should be expressed in relation to some objective criterion that remains invariant across samples, so that a body of knowledge about the phenomenon of interest can more readily develop. For example, because raw frequency counts will depend on the duration of the measurement period, relative frequency scores may be preferable because they are comparable across measurement situations. However, the comparability of relative frequency scores requires exhaustive coding of all behavior with the same set of codes. Thus, rate per minute scores may be preferable because they are less dependent on other characteristics of the coding scheme or measurement situation.

3.01.8.1.2 Measurement of sequential patterns of behavior

When events or states are coded exhaustively, it is possible to evaluate patterns of behavior for an individual or between individuals interacting with each other. The behaviors can occur concurrently (e.g., the co-occurrence of eyebrow arch and smile) or serially (e.g., reciprocal negative behaviors between spouses). Sequential patterns are usually derived from Bayesian statistics that relate the occurrence or nonoccurrence of an antecedent event or state with the occurrence or nonoccurrence of a consequent event or state. In lag-sequential analysis (e.g., Bakeman & Gottman, 1997; Sackett, 1979a), the investigator is concerned with the transitional probability of the antecedent/consequent sequence, which reveals the probability that the consequent occurs, given the antecedent event or state. The important question regarding sequential dependency is the extent to which this transitional probability exceeds (or is smaller than) the base rate for the consequent. If the consequent is more (or less) likely to occur in the presence of, or following, the antecedent than in other circumstances, then there is dependency between the antecedent and the consequent. If, for example, the probability that a mother smiles at her infant (consequent behavior) is greater after the infant smiles at her (antecedent behavior) than after other types of infant behaviors, then there is a sequential dependency in infant smile–mother smile exchanges.

Sequential dependency can be estimated with several statistics. One common metric is the lag sequential *z*-score developed by Sackett (1979a), and a modification developed by Allison and Liker (1982) that corrects for sampling error. This statistic compares the observed frequency of the antecedent/consequent sequence with the expected frequency of the sequence (based on the base rates for both events). More recently, Bakeman and colleagues (Bakeman & Gottman, 1997; Bakeman & Quera, 1995) recommended a slightly different formula derived from two-way contingency tables, which is the adjusted residual obtained in log-linear analysis. A convenient feature of this statistic is that because the scores are distributed as z, the statistical significance of sequential patterns is readily discerned (i.e., scores greater than $z = 1.96$). The statistic can be computed with standard statistical packages. Importantly, Bakeman and Quera (1995) also suggest statistical programs to compute these scores when codes in a sequence cannot repeat, such as when the coding system requires that new codes are assigned only when the behavior changes. Because a behavior can never follow itself, this situation produces "structural zeros" in the diagonal of the contingency table.

The z-statistic or adjusted residual is most useful for examining a sequential pattern in a single set of observations on an individual subject or a group of subjects. However, it is not recommended for use when separate sequential scores for each of multiple subjects are entered as data points into subsequent inferential analyses. The problem is that the size of adjusted residuals will vary (become larger) as a function of the number of observations that are made, even when the actual conditional probabilities remain constant. Thus, the z-scores are influenced by the base rates of behaviors, and can differ dramatically across subjects when response productivity differs. Wampold (1989) recommends a transformed kappa statistic as an alternative; however, this statistic requires selecting among three different computational formulas depending on the relative size of expected and actual probabilities. Other commonly used statistics are Yule's Q and phi. Bakeman and Casey (1995) provide computational formulas, discuss the conditions under which various statistics may be applicable, and suggest that investigators examine the distributions of scores for each statistic to determine which statistic provides the best distribution of scores for a particular set of data.

Investigators are commonly concerned about the amount of data needed to calculate sequential statistics. This issue is most relevant when the investigator is interested in studying sequences involving relatively rare events. Unfortunately, clinically relevant phenomena are often rare phenomena, so the problem is a common one in the field of clinical psychology. Bakeman and Gottman (1997) present a detailed analysis of this issue using guidelines employed in log-linear analysis of contingency tables. As a general rule of thumb, the investigator should obtain enough data so that each antecedent and consequent event in a contingency table occurs at least five times. In many cases, this criterion can be met by collapsing codes into more general categories. Another strategy is to conduct analyses using pooled observations on a group of subjects, rather than calculating sequential statistics for individual subjects. For example, Gottman, Markman, and Notarius (1977) examined sequences of effective and ineffective problem solving in groups of distressed and nondistressed married couples. Although incidents of positive behaviors such as "validate" and negative behaviors such as "put-down" were rare in some couples, within the two groups they occurred with sufficient frequency to identify patterns of supportive interactions that were more typical of happily married couples, and patterns of hostile exchange that were more typical of spouses in distressed marriages.

3.01.8.2 Computer Technology for Recording and Coding Behavioral Observations

Since the late 1970s, computer technology has become increasingly available for use in recording and coding observational data. Whereas the use of devices and software requires extra initial costs in time and money, they may ultimately increase the efficiency of data collection, coding, and data management. If developers find a market for these devices, we can expect that their availability will become more widespread. Because specific products are likely to undergo rapid changes and development, a list of currently available products would be immediately out of date, and thus would be of little use. However, it is possible to illustrate the range of current options and some of the factors to consider in selecting equipment and software.

For most situations in which behavioral observation is used, the major function of automated systems is to record the observations and the codes reported by the human observer. Although Tryon (1991) catalogs many automated devices that actually measure "actions," most of these devices track physiological responses or simple, gross motor activity. To date, no automated system can code the types of complex events that are the focus of this chapter. The first systems designed for event recording were "dedicated devices" because their sole function was to aid in data collection. Data could be stored in a temporary system and then uploaded to a computer for analysis. These devices have been rapidly replaced by software systems that run on standard computers. The newest automated systems combine data entry and management functions with computer control of video playback devices.

One consideration in selecting a system is the number of codes the system can handle. For example, one software package, the Observational Data Acquisition Program (ODAP; Hetrick, Isenhart, Taylor & Sandman, 1991), allows for recording frequency and duration of up to 10 behaviors. Duration is recorded by depressing a key during the occurrence of a behavior. Taylor et al. (1991) used this system to measure six types of self-injurious behaviors. Although this system is easy to use, the major limitation is that only a circumscribed range of behavior can be evaluated with simple, one dimensional codes. In contrast, other systems allow the investigator to configure a data entry file to handle as many as 1000 codes, and the codes can be multifaceted. For example, the Observer system (Noldus, 1991) has the capacity to hold 1000 different codes which can be classified as occurring only under certain circumstances. The system prompts the coder to correct errors when codes that violate the classification structure are entered (e.g., the same behavior entered twice in systems where codes cannot repeat, or when a teacher code is assigned to a student). The Multiple Option Observation System for Experimental Studies (MOOSES; Tapp, Wehby, & Ellis, 1995) offers similar flexibility. In a study by Shores et al. (1993), MOOSES was used to record classroom interactions of children with behavior disorders. Codes were devised to indicate the actor, the behavior exhibited, and the recipient of the behavior. In addition, conditional variables such as the presence of a teacher or the grouping of students who were present were also recorded.

A second consideration is whether the system can interface with video playback devices. Several systems are available that link the computer with professional quality videotape players that include a computer interface port. A machine-readable time code is stamped on the videotape, and the computer reads the time code to track onset and offset times for events, duration for states, or to access specific intervals. The advantage of these systems over simple real-time recording devices is that the videotape can be stopped, reversed, and started repeatedly without needing to reset a timer. Examples are the Observer system (Noldus, 1991), Observation Coding System Toolset (OCS Tools; Triangle Research Collaborative, Inc, 1994), and Procoder (Tapp & Walden, 1993). These systems differ greatly in terms of the number and complexity of codes they can handle and the ability and ease of controlling the video player from the computer keyboard. More recently, packages are becoming available that use CD-ROM instead of videotape input (e.g., vPrism; Digital LAVA Inc., 1996). These systems provide easier access to specific sections of an observation session because of the digital access features of CD-ROM, and easier replay as compared to rewinding a video player. They may also allow for greater precision in labeling exact time segments. However, they require converting videotaped observations to compact disks, which is expensive when completed by companies that provide this service, and time-consuming when equipment is available for converting data in-house. Nevertheless, the reduced storage requirements for CDs as opposed to videotapes is an advantage.

A third consideration is the statistical analysis that a particular program may offer. Whereas some systems include software for calculating coder agreement, linking multiple channels (e.g., behavioral codes with heart rate data), and conducting sequential analyses, others output the data so that it can be imported into other programs for these purposes. Graphic presentation of data is also an option with some packages. In some cases, statistical programs are included as part of the basic system; in other instances, the statistical software is optional.

3.01.9 INTERPRETING DATA GATHERED THROUGH BEHAVIORAL OBSERVATION

As can be seen, there are a number of steps before investigators obtain findings based on behavioral observation. The final step for the investigator is interpreting these findings. Whereas this interpretive process is psychological in nature and dependent upon the specifics of the particular investigation, our own research on couples' relationships, marital therapy, and family development illustrates how a pattern of findings that include both behavioral observation and self-report measures can help to elucidate family/marital functioning.

Baucom and his colleagues have conducted several treatment outcome studies with maritally distressed couples from a cognitive-behavioral perspective. Cognitive-behavioral conceptualizations of marital distress have placed a major emphasis on the centrality of communication as essential for effective marital functioning. However, the pattern of findings across our investigations indicates that communications training might be neither necessary nor sufficient for affecting changes in marital functioning. In these investigations, couples' communication was assessed by the MICS III (Weiss, 1986) after the couples attempted to solve problems on areas of difficulty in their marriage in a laboratory setting. In the first study, all couples received some set of behavioral interventions to assist their marriages (Baucom, 1982). The difference among the treatment conditions involved whether the couples received communications training.

The findings indicated that couples in all of the active treatment conditions improved equally on communication and marital adjustment, suggesting that communications training was not the critical element in the treatment. In a subsequent investigation, Baucom, Sayers, and Sher (1990) provided maritally distressed couples with a variety of cognitive-behavioral interventions, but all couples received communications training. Again, the findings indicated that couples in all active treatment conditions improved equally on communication and marital adjustment. However, additional analyses indicated that changes in communication were not correlated with changes in marital adjustment; thus, communication training could not be interpreted as the critical ingredient that altered marital adjustment (Iverson & Baucom, 1990). This is not to say that communication is unimportant in marriage; a number of investigations indicates that it is (e.g., Gottman & Krokoff, 1989). However, the results from these investigations and others have led cognitive-behavioral investigators to develop more complex and multifaceted models of marital distress that include, but are not limited to, a communications skills deficit explanation (e.g., Karney & Bradbury, 1995). Thus, in this instance, the findings from investigations involving behavioral observation of communication with couples have led to theoretical changes in the conceptualization of marital adjustment.

Research by Floyd and colleagues investigates associations among subsystems within the families of children who have disabilities, and examines how family relationships predict adaptive functioning for the child and the other family members. All observations and data collection are conducted in the families' homes in order to maximize the likelihood that observed behaviors are relevant to daily functioning in the family. One set of reports (Floyd, Gilliom, & Costigan, in press; Floyd & Zmich, 1991) focuses on the hypothesis that the quality of the parents' marital functioning and their parenting alliance influence the quality of parenting experiences. In order to test the hypothesis, observational measures were combined with self-reports of both marital functioning and parenting experiences.

Similar to procedures commonly employed in studies of marital relationships, and as illustrated in the studies by Baucom and colleagues, the parents' marital problem-solving skills were assessed by having them discuss and attempt to resolve a significant area of disagreement in their relationship. This procedure linked the investigation to the large body of previous observational research on marriage and marital therapy. Also, because the focus of the study was to understand how the parents functioned as a marital dyad, the marital discussions were conducted in a room separate from the children. Other studies interested in how children witness marital conflict have used procedures where children are present either as observers or participants in the discussions (e.g., Cummings, 1987). In order to reduce reactive effects associated with the presence of a live observer, the couples were left alone together for 10 minutes to complete the discussion, and the discussions were videotaped for later coding.

Parenting experiences were evaluated in a separate family interaction session. The procedures for this session were adopted from the work of Patterson, Reid, and colleagues (Patterson, 1982; Patterson, Reid, & Dishion, 1992) with aggressive children and their families. These procedures had been used to identify how behavior management and control are disrupted, and lead to the escalation of negative exchanges in the families of aggressive children. Because children with mental retardation also present behavior management challenges for parents, the possibility of negative escalation was potentially relevant to these families. In order to make the interaction as naturalistic as possible, families were videotaped in the home while completing a task of their own choosing. All family members were present. However, it was also necessary to structure the session somewhat in order to ensure that the family members interacted together and that the videotapes were sufficiently clear so that the behaviors could be reliably coded. Thus, families were instructed to complete an activity together (e.g., baking cookies, working on a crafts project), to refrain from watching television or making or taking telephone calls, and to remain together in one or two rooms within range of the video camera. The families were observed for a total of 50 minutes, which, for coding purposes, was divided into 10 minute segments. During each segment, one family member was identified as the focus of coding, and only behaviors that occurred by the focus person and anyone interacting with that person were coded by the observer. This procedure allowed the camera operator to focus on subsets of family members rather than trying to keep all family members in view at all times.

The findings support the value of these observational methods. Most notably, whereas self-report measures of marital quality, the parenting alliance, and parenting experiences failed to distinguish families of children with mental retardation from a comparison group with typically developing children, both the marital interactions and the family interactions demonstrated greater marital negativity and more parent–child behavior management struggles for the MR group (Floyd & Zmich, 1991). Furthermore, negative marital interactions were predictive of negative parent–child exchanges. A subsequent longitudinal evaluation demonstrated that marital quality, including the quality of marital problem-solving interactions, predicts changes in negative parent–child exchanges over a two-year period, with couples who are most negative together showing increases in their negative exchanges with children over time (Floyd, Gilliom, & Costigan, in press).

3.01.10 CONCLUDING COMMENTS

As can be seen in the above discussion, behavioral observation provides an extremely rich source of information for investigators as they attempt to understand the complexities of human behavior and interaction. This richness presents many opportunities for investigators and many challenges as well. These challenges are incorporated in the myriad of decisions that the investigator must make in employing behavioral observation, and obviously the route that the investigator chooses greatly impacts the findings. Thus, the investigator incurs responsibility for understanding the impact of these decisions on the questions under investigation. Often in reporting the results of investigations employing behavioral observation, the method section of the report involving the application of coding systems and data analytic strategies is presented in a few short paragraphs. Consequently, most readers will have only a general level of understanding of how the coding system was employed and how that impacts the findings. Therefore, the investigator must thoroughly understand the coding process so that discussions of the findings accurately represent what is indicated by the data. When this is done in a thoughtful manner, we have the opportunity to use one of our most natural strategies, observing people's behavior, as a major way to advance the science of clinical psychology.

3.01.11 REFERENCES

Ainsworth, M. D., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation*. Hillsdale, NJ: Erlbaum.

Allison, P. D., & Liker, J. K. (1982). Analyzing sequential categorical data on dyadic interaction: A comment on Gottman. *Psychological Bulletin, 91,* 393–403.

Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behaviour, 49,* 227–265.

American Psychological Association (1985). *Standards for educational and psychological testing*. Washington, DC: Author.

Ammerman, R. T., Van Hasselt, V. B., & Hersen, M. (1991). Parent–child problem-solving in families of visually impaired youth. *Journal of Pediatric Psychology, 16,* 87–101.

Ary, D. (1984). Mathematical explanation of error in duration recording using partial interval, whole interval, and momentary time sampling. *Behavioral Assessment, 6,* 221–228.

Bakeman, R., & Casey, R. L. (1995). Analyzing family interaction: Taking time into account. *Journal of Family Psychology, 9,* 131–143.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis* (2nd ed.). New York: Cambridge University Press.

Bakeman, R., & Quera, V. (1995). Log-linear approaches to lag-sequential analysis when consecutive codes may and cannot repeat. *Psychological Bulletin, 118,* 272–284.

Barton, C. & Alexander, J. F. (1981). Functional family therapy. In A. S. Gurman & D. P. Kniskern (Eds.), *Handbook of family therapy* (pp. 403–443). New York: Brunner/Mazel.

Baucom, D. H. (1982). A comparison of behavioral contracting and problem-solving/communications training in behavioral marital therapy. *Behavior Therapy, 13,* 162–174.

Baucom, D. H., Sayers, S. L., & Sher, T. G. (1990). Supplementing behavioral marital therapy with cognitive restructuring and emotional expressiveness training: An outcome investigation. *Journal of Consulting and Clinical Psychology, 58,* 636–645.

Cairns, R. B., & Green, J. A. (1979). How to assess personality and social patterns: Observations or ratings? In R. B. Cairns (Ed.), *The analysis of social interactions: Methods, issues, and illustrations* (pp. 209–226). Hillsdale, NJ: Erlbaum.

Carpenter, L. J., & Merkel, W. T. (1988). The effects of three methods of observation on couples in interactional research. *American Journal of Family Therapy, 16,* 144–157.

Christensen, A., & Hazzard, A. (1983). Reactive effects during naturalistic observation of families. *Behavioral Assessment, 5,* 349–362.

Christensen, A., & Heavey, C. L. (1990). Gender and social structure in the demand/withdraw pattern of marital conflict. *Journal of Personality and Social Psychology, 59,* 73–81.

Cummings, E. M. (1987). Coping with background anger in early childhood. *Child Development, 58,* 976–984.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist, 34,* 571–582.

Digital LAVA Inc. (1996). 10850 Wilshire Blvd., Suite 1260, LA, CA 90024.

Fagot, B., & Hagan, R. (1988). Is what we see what we get? Comparisons of taped and live observations. *Behavioral Assessment, 10,* 367–374.

Floyd, F. J. (1989). Segmenting interactions: Coding units for assessing marital and family behaviors. *Behavioral Assessment, 11,* 13–29.

Floyd, F. J., & Costigan, C. L. (1997). Family interactions and family adaptation. In N. W. Bray (Ed.), *International review of research in mental retardation* (Vol. 20, pp. 47–74). New York: Academic Press.

Floyd, F. J., Gilliom, L. A., & Costigan, C. L. (in press). Marriage and the parenting alliance: Longitudinal prediction of change in parenting perceptions and behaviors. *Child Development.*

Floyd, F. J., & Zmich, D. E. (1991). Marriage and the parenting partnership: Perceptions and interactions of parents with mentally retarded and typically developing children. *Child Development, 62,* 1434–1448.

Foster, S. L., Inderbitzen, H. M., & Nangle, D. W. (1993). Assessing acceptance and social skills with peers in childhood. *Behavior Modification, 17,* 255–286.

Gardner, W., & Griffin, W. A. (1989). Methods for the analysis of parallel streams of continuously recorded social behaviors. *Psychological Bulletin, 105,* 446–455.

Gottman, J. M. (1979). *Marital interaction: Experimental investigations*. New York: Academic Press.

Gottman, J. M. (1980). Analyzing for sequential connection and assessing interobserver reliability for the sequential analysis of observational data. *Behavioral Assessment, 2,* 361–368.

Gottman, J. M. (1994). *What predicts divorce?* Hillsdale, NJ: Erlbaum.

Gottman, J. M., & Krokoff, L. J. (1989). Marital interaction and satisfaction: A longitudinal view. *Journal of Consulting & Clinical Psychology, 57*(1), 47–52.

Gottman, J. M., Markman, H., & Notarius, C. (1977). The topography of marital conflict: A sequential analysis of verbal and nonverbal behavior. *Journal of Marriage and the Family, 39,* 461–477.

Haley, J. (Ed.) (1971). *Changing families*. New York: Grune & Stratton.

Harris, F. C. & Lahey, B. B. (1986). Condition-related reactivity: The interaction of observation and intervention in increasing peer praising in preschool children. *Education and Treatment of Children, 9,* 221–231.

Hartmann, D. P. (1977). Considerations in the choice of interobserver reliability estimates. *Journal of Applied Behavior Analysis, 10,* 103–116.

Hartmann, D. P. (Ed.) (1982). *Using observers to study behavior*. San Francisco: Jossey-Bass.

Hartmann, D. P., & Wood, D. D. (1990). Observational methods. In A. S. Bellack, M. Hersen, & A. E. Kazdin (Eds.), *International handbook of behavior modification and therapy* (2nd ed., pp. 109–138). New York: Plenum.

Hawkins, R. P. (1979). The functions of assessment: Implications for selection and development of devices for assessing repertoires in clinical, educational, and other settings. *Journal of Behavioral Assessment, 12,* 501–516.

Haynes, S. N. (1978). *Principles of behavioral assessment*. New York: Gardner Press.

Haynes, S. N. (1998). The changing nature of behavioral assessment. In M. Hersen & A. S. Bellack (Eds.), *Behavioral assessment: A practical handbook* (4th ed., pp. 1–21). Boston: Allyn & Bacon.

Haynes, S. N., Follingstad, D. R., & Sullivan, J. C. (1979). Assessment of marital satisfaction and interaction. *Journal of Consulting and Clinical Psychology, 47,* 789–791.

Haynes, S. N., & Horn, W. F. (1982). Reactivity in behavioral observation: A review. *Behavioral Assessment, 4,* 369–385.

Herbert, J., & Attridge, C. (1975). A guide for developers and users of observational systems and manuals. *American Educational Research Journal, 12,* 1–20.

Hersen, N., & Bellack, A. S. (1998). *Behavioral assessment: A practical handbook* (4th ed.). Boston: Allyn & Bacon.

Hetrick, W. P., Isenhart, R. C., Taylor, D. V., & Sandman, C. A. (1991). ODAP: A stand-alone program for observational data acquisition. *Behavior, Research Methods, Instruments, and Computers, 23,* 66–71.

Humphrey, L. L., & Benjamin, L. S. (1989). *An observational coding system for use with structural analysis of social behavior: The training manual.* Unpublished manuscript, Northwestern University Medical School, Chicago.

Iverson, A., & Baucom, D. H. (1990). Behavioral marital therapy outcomes: Alternate interpretations of the data. *Behavior Therapy, 21*(1), 129–138.

Jacob, T. (1975). Family interaction in disturbed and normal families: A methodological and substantive review. *Psychological Bulletin, 82,* 33–65.

Jacob, T., Tennenbaum, D., Bargiel, K., & Seilhamer, R. A. (1995). Family interaction in the home: Development of a new coding system. *Behavior Modification, 19,* 147–169.

Jacob, T., Tennenbaum, D. L., & Krahn, G. (1987). Factors influencing the reliability and validity of observation data. In T. Jacob (Ed.), *Family interaction and psychopathology: Theories, methods, and findings* (pp. 297–328). New York: Plenum.

Jacob, T., Tennenbaum, D., Seilhamer, R. A., Bargiel, K., & Sharon, T. (1994). Reactivity effects during naturalistic observation of distressed and nondistressed families. *Journal of Family Psychology, 8,* 354–363.

Jacobson, N. S. (1985). The role of observational measures in behavior therapy outcome research. *Behavioral Assessment, 7,* 297–308.

Jarrett, R. B., & Nelson, R. O. (1984). Reactivity and unreliability of husbands as participant observers. *Journal of Behavioral Assessment, 6,* 131–145.

Johnston, J. M., & Pennypacker, H. S. (1993). *Strategies and tactics of behavioral research* (2nd ed.). Hillsdale, NJ: Erlbaum.

Karney, B. R., & Bradbury, T. N. (1995). The longitudinal course of marital quality and stability: A review of theory, methods, and research. *Psychological Bulletin, 118*(1), 3–34

Kazdin, A. E. (1977). Artifact, bias and complexity of assessment: The ABC's of reliability. *Journal of Applied Behavior Analysis, 10,* 141–150.

Kazdin, A. E. (1982). Observer effects: Reactivity of direct observation. *New Directions for Methodology of Social and Behavioral Science, 14,* 5–19.

Klesges, R. C., Woolfrey, J., & Vollmer, J. (1985). An evaluation of the reliability of time sampling versus continuous observation data collection. *Journal of Behavior Therapy and Experimental Psychiatry, 16,* 303–307.

Kuller, R., & Linsten, C. (1992). Health and behavior of children in classrooms with and without windows. *Journal of Environmental Psychology, 12,* 305–317.

Leaper, C., Hauser, S., Kremen, A., Powers, S. I., Jacobson, A. M., Noam, G. G., Weiss-Perry, B., & Follansbee, D. (1989). Adolescent–parent interactions in relation to adolescents' gender and ego development pathway: A longitudinal study. *Journal of Early Adolescence, 9,* 335–361.

Mann, J., ten-Have, T., Plunkett, J. W., & Meisels, S. J. (1991). Time sampling: A methodological critique. *Child Development, 62,* 227–241.

Mehm, J. G., & Knutson, J. F. (1987). A comparison of event and interval strategies for observational data analysis and assessments of observer agreement. *Behavioral Assessment, 9,* 151–167.

Minuchin, S. (1974). *Families and family therapy*. Cambridge, MA: Harvard University Press.

Noldus, L. P. J. J. (1991). The Observer: A software system for collection and analysis of observational data. *Behavior Research Methods, Instruments, and Computers, 23,* 415–429.

Noller, P. (1980). Misunderstandings in marital communication: A study of couples' nonverbal communication. *Journal of Personality and Social Psychology, 39,* 1135–1148.

O'Neill, R. E., Horner, R. H., Albin, R. W., Storey, K., & Sprague, J. R. (1990). *Functional analysis of problem behavior: A practical assessment guide*. Sycamore, IL: Sycamore Publishing.

Patterson, G. R. (1982). *A social learning approach, Vol 3: Coercive family process*. Eugene, OR: Castalia Publishing Company.

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial Boys.* Eugene, OR: Castalia.

Pett, M. A., Wampold, B. E., Vaughn-Cole, B., & East, T. D. (1992). Consistency of behaviors within a naturalistic setting: An examination of the impact of context and repeated observations on mother–child interactions. *Behavioral Assessment, 14,* 367–385.

Prinz, R. J., & Kent, R. N. (1978). Recording parent–adolescent interactions without the use of frequency or interval-by-interval coding. *Behavior Therapy, 9,* 602–604.

Robin, A. L., & Foster, S. L. (1989). *Negotiating parent–adolescent conflict: A behavioral-family systems approach.* New York: Guilford.

Sackett, G. P. (1979a). The lag sequential analysis of contingency and cyclicity in behavioral interaction research. In J. D. Osofsky (Ed.), *Handbook of infant development* (pp. 623–649). New York: Wiley.

Sackett, G. P. (Ed) (1979b). *Observing behavior. Vol. 2: Data collection and analysis methods.* Baltimore: University Park Press.

Schaap, C. (1982). *Communication and adjustment in marriage.* Lisse, Holland: Swetts & Zeitlinger.

Serbin, L. A., Citron, C., & Connor, J. M. (1978). Covert assessment of observer agreement: An application and extension. *Journal of Genetic Psychology, 133,* 155–161.

Shores, R. E., Jack, S. L., Gunter, P. L., Ellis, D. N., Debreire, T. J., & Wehby, J. H. (1993). Classsroom interactions of children with behavior disorders. *Journal of Emotional and Behavioral Disorders, 1,* 27–39.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428.

Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology.* New York: Basic Books.

Skinner, B. F. (1938). *The behavior of organisms.* New York: Appleton-Century-Crofts.

Snyder, J. (1983). Aversive social stimuli in the Family Interaction Coding System: A validation study. *Behavioral Assessment, 5,* 315–331.

Stine, W. W. (1939). Interobserver relational agreement. *Psychological Bulletin, 106,* 341–347.

Suen, H. K., Ary, D., & Covalt, W. (1990). A decision tree approach to selecting an appropriate observation reliability index. *Journal of Psychopathology and Behavioral Assessment, 12,* 359–363.

Taplin, P. S. & Reid, J. B. (1973). Effects of instructional set and experimenter influence on observer reliability. *Child Development, 44,* 547–554.

Tapp, J., & Walden, T. (1993). PROCORDER: A professional tape control, coding, and analysis system for behavioral research using videotape. *Behavior Research Methods, Instruments, and Computers, 25,* 53–56.

Tapp, J., Wehby, J., & Ellis, D. (1995). A multiple option observation system for experimental studies: MOOSES. *Behavior Research Methods, Instuments, and Computers, 27,* 25–31.

Taylor, D. V., Hetrick, W. P., Neri, C. L., Touchette, P., Barron, J. L., & Sandman, C. A. (1991). Effect of naltrexone upon self-injurious behavior, learning, and activity: A case study. *Pharmacology, Biochemistry, and Behavior, 40,* 79–82.

Triangle Research Collaborative, Inc. (1994). P. O. Box 12167, 100 Park, Suite 115, Research Triangle Park, NC 27709.

Tryon, W. W. (1991). *Activity measurement in psychology and medicine.* New York: Plenum.

Van Widenfelt, B., Baucom, D. H., & Gordon, K. C. (1997). *The Prevention and Relationship Enhancement Program: An empirical analysis.* Manuscript submitted for publication.

Wampold, B. E. (1989). Kappa as a measure of pattern in sequential data. *Quality and Quantity, 23,* 171–187.

Wampold, B. E., & Holloway, E. L. (1983). A note on interobserver reliability for sequential data. *Journal of Behavioral Assessment, 5,* 217–225.

Watson, J. B., & Raynor, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3,* 1–12.

Weinrott, M. R. & Jones, R. R. (1984). Overt versus covert assessment of observer reliability. *Child Development, 5,* 1125–1137.

Weiss, R. L. (1986). *MICS-III manual.* Unpublished manuscript, Oregon Marital Studies Program, University of Oregon, Eugene.

Wieder, G. B., & Weiss, R. L. (1980). Generalizability theory and the coding of marital interactions. *Journal of Consulting and Clinical Psychology, 48,* 469–477.

Wilson, F. R. (1982). Systematic rater training model: An aid to counselors in collecting observational data. *Measurement and Evaluation in Guidance, 14,* 187–194.

Winer, B. J. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

3.02

Single Case Experimental Designs: Clinical Research and Practice

STEVEN C. HAYES and JOHN T. BLACKLEDGE
University of Nevada, Reno, NV, USA

3.02.1 INTRODUCTION 24
  3.02.1.1 Brief History of Single Case Experimental Designs 24
  3.02.1.2 Utility of Single-subject Designs 25
3.02.2 ESSENTIAL COMPONENTS OF SINGLE CASE EXPERIMENTAL DESIGNS 25
  3.02.2.1 Repeated Measurement 26
  3.02.2.2 Detailed Information 26
  3.02.2.3 Graphing of Data 26
  3.02.2.4 Creative Use of Design Elements 27
3.02.3 INTERPRETATION OF SINGLE-SUBJECT DATA 27
  3.02.3.1 Variability and Stability 27
  3.02.3.2 Level 27
  3.02.3.3 Trend 27
  3.02.3.4 Use of Statistics with Single-subject Data 27
3.02.4 VARIETIES OF SINGLE-SUBJECT DESIGNS 28
  3.02.4.1 Within-series Design Elements 28
    3.02.4.1.1 Simple phase change 29
    3.02.4.1.2 Complex phase change elements 29
    3.02.4.1.3 Parametric design 33
    3.02.4.1.4 Final suggestions for interpreting within-series data 34
  3.02.4.2 Between-series Designs 35
    3.02.4.2.1 Alternating treatment design 35
    3.02.4.2.2 Simultaneous treatment design 38
  3.02.4.3 Combined-series Elements 38
    3.02.4.3.1 Multiple baseline 38
    3.02.4.3.2 Crossover design 42
    3.02.4.3.3 Constant series control 42
3.02.5 FUTURE DIRECTIONS 43
  3.02.5.1 Potential for Research Production and Consumption by Practicing Clinicians 43
  3.02.5.2 Managed Health Care, Single-subject Design, and the Demonstration of Efficacy 43
3.02.6 CONCLUSION 43
3.02.7 REFERENCES 44

3.02.1 INTRODUCTION

The purpose of clinical research from the point of view of the consumer of research knowledge can be stated succinctly: "What treatment, by whom, is most effective for this individual with that specific problem under which set of circumstances, and how does it come about?" (Paul, 1969, p. 44). This question has always been of relevance to practicing clinical psychologists in the fee-for-service environment, but it is also of increasing relevance in the era of managed care. Mental health delivery systems cannot succeed, either in the world of public opinion or in the world of fiscal reality, without finding a way to deliver services that are both effective and efficient (Cummings, Cummings, & Johnson, 1997; Cummings, Pollack, & Cummings, 1996). In order to do that, Paul's clinical question above must be answered for the varieties of clients demanding and receiving services.

There is another way to say this. Clinical research must have external validity (must apply to the settings and clients of the research consumer), not merely internal validity (an unambiguous relationship between a dependent and independent variable). In group comparison research, external validity flows in principle from internal validity. In a classic group comparison design, researchers hope to randomly sample from a known population, randomly assign to a treatment and control group, and collect before and after measures on all. If these methodological requirements have been met, the results should apply to other random samples from that same population. In the practical world of clinical research, however, we do not have access to the entire population of interest (say, all panic disordered clients). We cannot randomly sample from this population because clients nonrandomly refuse to participate. We cannot randomly assign because clients nonrandomly drop out. And even if all this were not true, we never apply clinical research to other random samples from the same population—rather, a clinician treats the person who (nonrandomly) walked through the door.

External validity thus must be earned—whether in group comparison or single-case research—in a piecemeal inductive fashion by demonstrating that particular treatments work with particular clients with particular problems. In group comparison research, this is usually done by showing that treatments work in highly homogeneous groups of clients. By replicating effects across many such homogenous groups, external validity can be demonstrated in group research. Efforts to show broad external validity in a single group comparison experiment are usually extremely difficult. Blocking or stratifying samples on even a few factors in a single study can lead to huge designs that cannot be mounted without millions of dollars of research funds. One compromise is to use diagnostic categories as a way of establishing homogeneity; however, large and unexplained between-subject variation invariably results because the current diagnostic system is based on loose collections of signs and symptoms rather than functional processes.

An alternative approach is to build this knowledge base about treatment response from the ground up, person by person. In this approach, clinical replication across myriad clients provides the evidence that a treatment effect holds for a population and that it is moderated by specific subject, therapist, or setting variables. That is the approach of single case experimental designs (SCEDs) or what has also been termed "time-series designs." The former term is more popular but falsely suggests that the number of subjects is necessarily few in this approach. The latter correctly points to the source of the analysis but it is not very popular. Both terms will be used.

The purpose of this chapter is to show how single case designs are used in a research and practice environment, to provide an introduction to the various types of designs, and to explore the advantages of this research approach in the modern world of health care delivery.

3.02.1.1 Brief History of Single Case Experimental Designs

The origin of SCEDs in clinical psychology can be traced back to the beginnings of the scientist–practitioner model. Two years before at the first Boulder Conference, Thorne (1947) advocated the use of such designs as a practical way for clinicians to incorporate empirical science into their everyday interactions with clients, a goal concurrently explicated by Shakow et al. (1947). The type of experimental designs proposed by Thorne were a very significant improvement over the traditional case study because continual data collection over a series of phase changes were required, allowing more objective, data-driven decisions to be made. These designs were an adaptation of those previously used by experimental psychologists (Ferster & Skinner, 1957) working with animals.

Single case designs became considerably more popular with the popularization of behavioral approaches in the 1960s. For example, Baer, Wolf, and Risley (1968) described a number of these designs in an early issue of the *Journal of Applied Behavior Analysis*. Hersen and Barlow's groundbreaking text (1976) in turn brought these designs into the mainstream of behavior therapy.

Probably the biggest factors inhibiting the use of SCEDs by many clinicians has been a bias against ideographic research. Yet many methodological leaders of the field have been quite accepting of such an approach. For example, Cronbach (1975) advocated careful observation of individual clinical cases using SCED controls, maintaining that the use of such design tools allows a level of detail in observation and hypothesis testing not available in traditional group designs. Many others (Campbell, 1957; Koan & McGuire, 1973; Snow, 1974) have agreed, adding that tightly controlled, fixed condition, group experimentation is often not well-suited to a science of vastly differing humans with often vastly differing needs. Cone (1986) noted that ideographic, single-subject research is particularly well-suited for detecting point-to-point behavioral changes occurring in response to environmental variables, including psychological treatment.

3.02.1.2 Utility of Single-subject Designs

A clinician carefully conducting an appropriate single-subject design for a client could both circumvent the above difficulties and obtain scientifically useful results. Often no subject pool larger than one is needed to conduct a single-subject design (although it is necessary that the results of many such analysis be considered before any general conclusions can be made about a treatment's efficacy). Single-subject experiments are extremely flexible. For example, if a client is not responding as hoped to a treatment, components can be added or subtracted or an altogether new treatment can be implemented, without necessarily damaging the validity of conclusions drawn from the experiment. Essentially, the only limitation on the usefulness of single-subject designs in research and clinical practice is the flexibility of the researcher. If use of a planned design does not allow an adequate interpretation of emerging data, then the design can be altered at that point.

Perhaps more importantly, a properly used single-subject design can be extremely useful in facilitating responsible assessment and treatment. A clinician conducting a single-subject experiment is forced to pay close attention to repeated assessments of client behavior that provide invaluable information about target behaviors and treatment efficacy. With the continuous feedback that single-subject experiments provide, treatment components not working as planned can be altered, abandoned, or supplemented.

The comprehensive data recorded over several single-subject designs can also be used to provide linkage between client characteristics and treatment success or failure. As more detailed information is gathered in time-series designs than in a typical group experiment, events in individual client lives and various client characteristics that coincide with declines in treatment efficacy can be identified and taken into consideration for subsequent clients. The advantage at this level is that variability due to sources other than treatment can be identified at the level of the individual. This means that when many cases are collected and analyzed, correlations between subject characteristics and treatment responsiveness can be more refined. For example, it may become apparent after conducting several single-subject designs mapping the effects of a given intervention that subjects with certain characteristics respond in characteristically similar ways. Such a hypothesis can be followed up using additional single subjects, or by using a group design. Subjects that fail to respond positively to a given treatment may do so because of a detectable reason, and this reason may point to an important aspect of a relevant theory. Outcomes like this critically depend on the foresight of the researcher. Potentially important background data, as indicated by theory and common sense, should be collected in every single-subject experiment.

The use of time-series designs also compels a clinician to focus on the careful description of patient problems and treatment characteristics, and how this data relates to treatment outcome. In such a manner, variables that are functionally important for treatment can become evident. Over a series of SCEDs, generalizations concerning the active and efficacious components of treatment can be made from such data.

Finally, analysis of SCEDs concentrates on the magnitude of treatment effects rather than their statistical significance. If the treatment analyzed in a properly conducted SCED is clinically significant, it will be clearly so.

3.02.2 ESSENTIAL COMPONENTS OF SINGLE CASE EXPERIMENTAL DESIGNS

Although various types of individual time-series designs exist, and are discussed in some detail later, several procedures common to all time-series designs are first described.

3.02.2.1 Repeated Measurement

The bedrock of single case experimental designs is repeated measurements of client functioning taken in relevant domains throughout the course of treatment (Barlow & Hersen (1984), Hayes, Barlow, and Nelson (1997), and Kazdin (1980) give information on specific measurement strategies). Repeated measurements enable the estimation of variability within a case over time to provide the sources of information about treatment, measurement error, and extraneous factors in a time-series approach.

In the real world of clinical evaluation, the goal is to use measures that are both high quality and highly practical. It is advisable that such measurements begin as soon as possible, ideally during the first session. It is also advisable, in choosing which instruments to administer to the client or research subject, that the researcher err on the side of caution and administer as many instruments as may be even partially relevant to the current experiment's concerns. Theory and even common sense should be used to guide the choice as to what instruments are relevant. Practical constraints also play an important part. If an adequate but imperfect instrument already exists and no time is currently available to create a new one, the adequate instrument may be used. Client self-monitoring and self-report is entirely acceptable. In fact, any method of collecting data is acceptable with single-subject designs, so long as the researcher deems it appropriate. Flexibility, again, is the by-word. But, as with any instrument, pay attention to its established reliability and validity before using or interpreting data. As treatment progresses, instruments useless or irrelevant to the current case can be discarded, but data not initially gathered can never be collected. It is worth spending a little extra time administering an instrument on a hunch on the chance that it will yield valuable information. Measurements are then ideally administered as frequently as is practical and meaningful.

3.02.2.2 Detailed Information

Specification of the particular intervention made with the client, including when each component is delivered and any possibly significant deviations from the standardized treatment, allows meaningful inferences to be made from collected data. To be meaningful indicators of the effects of a treatment and its components, the clinician must be able to temporally link specific phases of their intervention with the ongoing flow of outcome data as described above. Details regarding the nature of the treatments, the environment that the therapy was delivered in, and characteristics of the therapist provide a level of detail conducive to proper replication, especially if the completed SCED is made available to others. Steps taken in group research to ensure treatment integrity can be taken here as well. For example, a colleague or student might be asked to assess a clinician's adherence to a designated treatment protocol, as well as the competence with which the treatment is delivered. Specification of the treatment requires that the researcher has a clear, theoretically-based idea of what they are attempting to accomplish and how to accomplish it. The type of treatment being used, its specific techniques, and the phases or subphases of each technique or strategy that are active should be noted. Enough detail should be added so that after the intervention is complete, an informed guess can be made as to what might have been responsible for observed effects.

Collection of detailed client information allows more meaningful inferences to be drawn from single-subject data. The recording of any information that might possibly affect the course and effectiveness of treatment may prove to be invaluable for data analysis. Seemingly relevant background information and significant events occurring in the client's life during the course of treatment qualify as important client information. If the client's spouse leaves her during the course of treatment, for example, notation of this event may serve as a possible explanation for a brief or sustained decline in the treatment's effectiveness. Of importance here is the chronicling of any information about the client or the client's life that might be expected to affect their response to treatment. Over a series of SCEDs, such information can be used to speculate as to why clients reacted differentially to treatment.

3.02.2.3 Graphing of Data

Analysis of an individual time-series design requires a visual representation of the data. A simple line graph, with time plotted on the x-axis and measurement score plotted on the y-axis, should be sufficient for most data because it presents an immediate picture of variability within a data series. Pragmatic considerations determine what unit of time to plot. Thus, if an instrument is measuring behaviors in frequent time intervals but the behavior itself is better thought of in temporally larger units, analysis may be facilitated if time intervals are collapsed (e.g., if daily measurements are summed and recorded as weekly measurements). Frequent and creative graphing of data, from various angles, using different units of time, composite scores from separate measures, etc., can allow insights into the effects of treatment.

3.02.2.4 Creative Use of Design Elements

Individual time-series design elements (discussed below) should be thought of as tools, not restrictive agents. If a complex phase change with an interaction element is initially planned, and the data indicates the client might benefit most from continuing with the current treatment, then continuation of treatment is clearly justified. Unexpected client or therapist vacations can be an opportunity for a return to baseline phase, allowing a chance to monitor the effectiveness, in contrast, of treatment. Clinical common sense and detailed understanding of the nature of the different design elements and when they are useful allows effective and serendipitous adaptation to events.

3.02.3 INTERPRETATION OF SINGLE-SUBJECT DATA

Time-series designs typically (though not always) consist of units of time called phases, each phase designating the continuous presence of a given condition (e.g., treatment, baseline, and so on). Data within each phase can be described as having various degrees of variability around a given level and trend. Use of statistical analyses is not necessary at the level of the individual, and indeed use of most inferential statistics with a single subject violates the assumptions of standard statistical tests. Clear graphing of data and a thorough understanding of interpreting such single-subject graphs are all that are needed to observe an effect. It is to the nature of variability, level, and trend, and to the opportunities they provide for the interpretation of the data, that we now turn.

3.02.3.1 Variability and Stability

Data within a phase is said to be stable to the extent that the effects of extraneous variables and measurement error, as reflected in variability within a subject across time, are sufficiently limited or identifiable that variability due to treatment can be ascertained. Determining stability requires that the clinician have some clarity about what treatment effects are large enough to be worth detection—the more variability due to extraneous factors or measurement error, the larger the treatment effect would have to be to be seen.

If the data are not stable (in the sense that important treatment effects might be obscured), the experimenter can (i) continue the phase until the data does become stable, (ii) examine the data using longer units of time if longer units are more sensible, or (iii) analyze the possible sources of variability. Each of these is defensible, but the last option is often the best because it can provide important information about the case.

3.02.3.2 Level

Level refers to the "height" on the y-axis at which the data tends to aggregate. For example, data taken during a treatment phase tending to stay at the upper points of the measurement scale would be notably distinguishable, in terms of level, from data congregating around the mid-point of a measurement scale.

3.02.3.3 Trend

Trend refers to the general linear direction in which data are moving across a given period of time. It takes a minimum of three data points to establish a trend and estimate variability around that trend.

Converging and diverging trends between phases can be analyzed to help differentiate variability due to treatment from variability due to extraneous sources. For example, if a client's data on a given measure shows a positive trend in baseline that continues when treatment is implemented and continues yet again when the baseline is reinstated, it is likely that something other than treatment is responsible for the improvement. Conversely, if a strong positive trend during a treatment phase levels out or especially reverses after a change to a baseline phase, the clinician can usually feel confident that treatment is responsible for the change (unless some potentially significant life change happened to co-occur with the phase change). Thus, trends are not only useful indicators as to the improvement or decline of clients on various measures, but are also useful indicators, coupled with phase changes, of the sources of those changes.

3.02.3.4 Use of Statistics with Single-subject Data

For the most part, inferential statistics were designed for use in between-group comparisons. The assumptions underlying the widely accepted classical model of statistics are usually violated when statistical tests based on the model are applied to single-subject data. To begin with, presentation of conditions is not generally random in single-subject designs, and randomization is a necessary prerequisite to statistical analysis. More importantly (and more constantly), the independence of data required in classical statistics is generally not achieved when statistical analyses are applied to time-series data from a single subject (Sharpley & Alavosius, 1988). Busk and Marascuilo (1988) found, in a review of 101 baselines and 125 intervention phases from various single-subject experiments, that autocorrelations between data, in most cases, were significantly greater than zero and detectable even in cases of low statistical power. Several researchers have suggested using analyses based on a randomization task to circumvent the autocorrelation problem (Edgington, 1980; Levin, Marascuilo, & Hubert, 1978; Wampold & Furlong, 1981). For example, data from an alternating treatment design or extended complex phase change design, where the presentation of each phase is randomly determined, could be statistically analyzed by a procedure based on a randomization task. Some controversy surrounds the issue (Huitema, 1988), but the consensus seems to be that classical statistical analyses are too risky to use in individual time-series data unless at least 35–40 data points per phase are gathered (Horne, Yang, & Ware, 1982). Very few researchers have the good fortune to collect so much data.

Time-series analyses where collected data is simply used to predict subsequent behavior (Gottman, 1981; Gottman & Glass, 1978) can also be used, and is useful when such predictions are desired. However, such an analysis is not suitable for series with less than 20 points, as serial dependence and other factors will contribute to an overinflated alpha in such cases (Greenwood & Matyas, 1990). In cases where statistical analysis indicates the data is not autocorrelated, basic inferential statistical procedures such as a *t*-test may be used. Finally, the Box–Jenkins procedure (Box & Jenkins, 1976) can technically be used to determine the presence of a main effect based on the departure of observed data from an established pattern. However, this procedure would require a minimum of about 50 data points per phase, and thus is impractical for all but a few single-subject analyses.

In addition, most statistical procedures are of unknown utility when used with single-subject data. As most statistical procedures and interpretations of respective statistical results were derived from between-group studies, use of these procedures in single-subject designs yields ambiguous results. The meaning of a statistically significant result with a lone subject does not mean the same thing as a statistically significant result with a group, and the assumptions and evidentiary base supporting classical statistics simply dose not tell us what a significant result with a single subject means. Beyond the technical incorrectness of using nomethetic statistical approaches to ideographic data, it is apparent that such use of these statistics is of extremely limited use in guiding further research and bolstering confidence about an intervention's efficacy with an individual subject. If, for example, a statistically significant result were to be obtained in the treatment of a given client, this would tell us nothing about that treatment's efficacy with other potential clients. Moreover, data indicating a clinically significant change in a single client would be readily observable in a well-conducted and properly graphed single-subject experiment. Statistics—so necessary in detecting an overall positive effect in a group of subjects where some improved, some worsened, and some remained unchanged—would not be necessary in the case of one subject exhibiting one trend at any given time.

3.02.4 VARIETIES OF SINGLE-SUBJECT DESIGNS

Time-series design elements can be classified either as within-series, between-series, or combined-series designs. Different types of time-series design elements within each group of designs are used for different purposes (Hayes et al., 1997). The nature of each type of element, as well as its purpose, will now be described.

3.02.4.1 Within-series Design Elements

A design element is classified as within-series if data points organized sequentially within a consistent condition are compared to other such sequential series that precede and follow. In such a design, the clinician is typically faced with a graphed data from a single measure or a homogenous group of measures, organized into a sequence of phases during which a consistent approach is applied. Phase changes occur, ideally, when data has stabilized. Otherwise, practical or ethical issues determine the time at which phases must change. For example, an extended baseline with a suicidal client would certainly not be possible, and time and financial constraints may determine phase changes in other cases. Aspects of specific design strategies that influence phase length are discussed at the appropriate points below.

Within-series design elements include the simple phase change, the complex phase change, the parametric design elements, and the changing criterion design. Each is described below. Note that specific phases are designated below by capital letters. Generally (but not always), the letter A refers to a baseline phase, and letters such as B and C refer to different interventions.

3.02.4.1.1 Simple phase change

If simply trying to determine a treatment's efficacy vs. no treatment at all for a given client, or comparing the relative efficacy of two validated treatments for a given client, then a simple phase change design is probably the right choice. Consider the case in which only a single treatment's efficacy is at issue. A similar approach can be taken to answer questions about the relative efficacy of two treatments.

In the standard simple phase change, baseline data on relevant client behaviors is typically taken for a period of time (enough time to estimate the level, trend, and variability around level and trend of the behavior of interest). This baseline phase can occur during initial sessions with the client while assessment (and no treatment *per se*) is taking place, while the client is on a waiting list, or through similar means. Treatment is then administered and while it is in place a second estimate is made of the level, trend, and variability around level and trend of the behavior of interest. If clear changes occur, treatment may have had an impact.

In order to control for extraneous events that might have co-occurred with the phase change from baseline to treatment, the phase change must be replicated. Usually this is done by changing from treatment back to baseline, but other options exist. Such a change, called a withdrawal, is useful in aiding the clinician in deciding whether it is indeed the treatment, and not some extraneous variable, that is responsible for any changes. However, certain ethical considerations must be made before a withdrawal is executed. If treatment definitely seems effective and is well-validated, it may not be necessary to withdraw the treatment to conclude that a treatment effect is likely. If the clinician is uncertain of the treatment's effectiveness, or if it is not well-validated, a withdrawal or some other means of replication is necessary in order to conclude that an effect has occurred. Even a short withdrawal can be sufficient if a clear trend is evident. Alternatively, a period of time where therapy is focused on a different, relatively unrelated problem can also be considered a withdrawal. Serendipitous opportunities for a return to a baseline phase can provide important information. Impromptu client or therapist vacations can be viewed as a chance to make more informed decisions regarding the treatment's efficacy.

More than one replication of the underlying phase change (such as an ABABAB design) may be necessary for the clinician to be confident of a treatment's effects for a particular client. Interpretation of the data is facilitated by referring to the specifics of stability, level and trends as discussed above.

Examples of well-conducted simple phase change designs include Gunter, Shores, Denny, and DePaepe (1994), who utilized the design to evaluate the effects of different instructional interactions on the disruptive behavior of a severely behaviorally disordered child. Gunter et al. (1994) used a simple phase change with reversal (in this case ABAB), with baseline or A phase consisting of a math assignment with between five and 15 difficult multiplication problems. The treatment or B phase consisted of equally difficult problems, but the experimenter would provide the correct answer and then present the same problem again whenever the subject incorrectly answered. Institution of the first treatment phase indicated a desirable effect, with the subject's rate of disruptive behavior falling from around 0.3 to around 0.1. Gunter et al. (1994) wisely decided to replicate the phase changes with the subject, allowing them to be more confident that extraneous variables such as time or (to some extent) order were not responsible for the changes.

Orsborn, Patrick, Dixon, and Moore (1995) provide another good, contemporary example of the simple phase change design (Figure 1). They used an ABAB design to evaluate the effect of reducing the frequency of teacher's questions and increasing the frequency of pauses on the frequency of student talk, using 21 first- and second-grade subjects. Both B phases showed a marked increase in student talk frequency relative to baseline phases. The strength of the design, however, could have been improved with the inclusion of more data points. Five data points were collected during baseline phases, and three in intervention phases. Three is an acceptable minimum, but more data points are advisable.

3.02.4.1.2 Complex phase change elements

A complex phase change combines a specific sequence of simple phase changes into a new logical whole.

(i) ABACA

When comparing the effectiveness of two (or more) treatments relative to each other and to

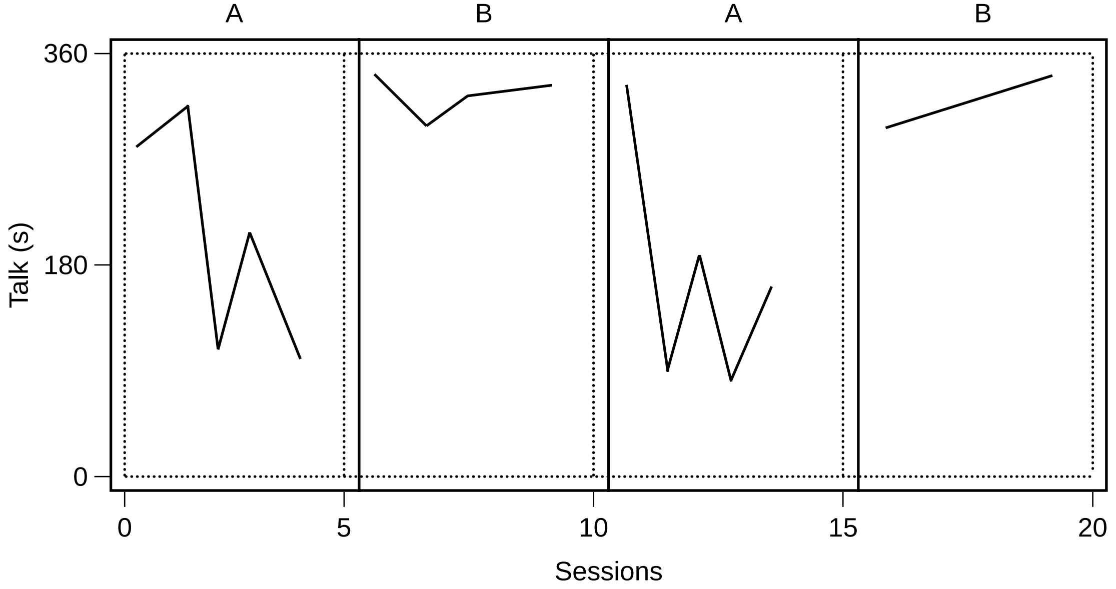

Figure 1 An adaptation of a simple phase change with reversal design. Data were relatively stable before phase changes (adapted from Orsborn et al., 1995).

no treatment, an ABACA complex phase change element might be used. Typically, the clinician would choose such a design when there is reason to believe that two treatments, neither yet well-validated, may be effective for a given client. When it is unclear whether either treatment will be effective, treatment phases can be interspersed with baseline phases. Phase changes conducted in such a manner will allow meaningful comparisons to be made between both treatments, as well as between each treatment and no treatment. After administering each treatment once, if one that results in clearly more desirable data level and trend, it may be reinstated. If each treatment's relative efficacy is still unclear, and the data gives no reason to believe that either may be iatrogenic, phase changes may be carried out as is practical. A sequence such as ABACACAB might not be unreasonable if the data and clinical situation warranted. Regardless of the sequence implemented, the clinician should remain aware that order effects in a complex phase change can be critical. The second treatment administered may be less effective simply because it is the second treatment. Counterbalancing of phase sequences in other cases can circumvent such ambiguity.

The clinician should stay alert to the possibility of introducing even a third treatment in such a design, if the original treatments do not appear to be having the desired effect. An example of such a situation is shown in a study by Cope, Moy, and Grossnickle (1988). In this study, McDonald's restaurants promoted the use of seat belts with an advertising campaign (phase B), and the distribution of instructive stickers for children asking them to "Make it Click" (phase C). These two interventions were tried repeatedly. Finally, an incentive program was implemented giving away soft drinks for drivers who arrived at McDonald's with their seat belt fastened. The design could be described as an ABCBCBDA. The spirit here, as always when using time-series approaches, should be one of flexible and data-driven decisions. With such a spirit, the clinician should be willing to abandon even an empirically validated treatment if it is clear, over a reasonable length of time, that there is no positive effect. No treatment is all things to all people.

Another well-conducted complex phase change design was reported by Peterson and Azrin (1992; Figure 2). Three treatments, including self-monitoring (phase B), relaxation training (phase C), and habit reversal (phase D), were compared with a baseline or no treatment (phase A). Six subjects were used, and the authors took advantage of the extra subjects by counterbalancing the presentation of phases. For example, while the first subject was presented with the phase sequence AAABCDCBDA, other subjects were presented with sequences such as AAADCBCD-BA, AAACDBDCBA, and AAACDBDCBA. A minimum of three data points were contained in each phase (generally four or more data points were used), and the authors more often than not waited until stability was achieved and clear trends were present before changing phases for subjects.

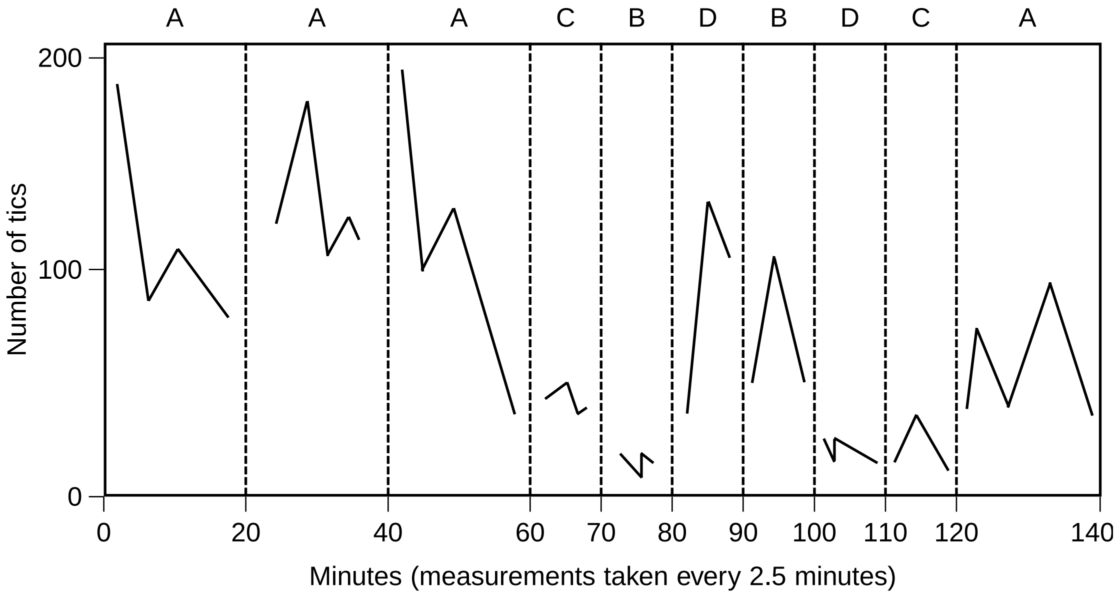

Figure 2 An example of a complex phase change. Waiting for the data in the first and third baseline (A) phase to stabilize (as well as the data in the first D phase) would have been preferable before initiating phase changes (adapted from Peterson and Azrin, 1992).

(ii) Interaction element

In an interaction element the separate and combined effects of intervention elements are examined. Use of this design element is appropriate both when a treatment is working and the clinician wishes to ascertain whether it will continue working without a particular (and costly in terms of time and/or money) component, or when a treatment is not maximally effective and the clinician believes adding or subtracting a specific component might enhance the treatment's efficacy.

White, Mathews, and Fawcett (1989) provide an example of the use of the interaction element. They examined the effect of contingencies for wheelchair push-ups designed to avoid the development of pressure sores in disabled children. Wheelchair push-ups were automatically recorded by a computer. After a baseline, two subjects were exposed to an alarm avoidance contingency (B), a beeper prompt (C), or a combination. An interaction design element was combined with a multiple baseline component. The design for one subject was an A/B + C/B/B + C/B/B + C/C/B + C, and for the other was A/B + C/C/B + C/C/B + C. Each component (B or C) was more effective than a baseline, but for both children the combined (B + C) condition was the most effective overall.

Shukla and Albin (1996) provided another good example when examining the effects of extinction vs. the effects of extinction plus functional communication training on problem behaviors of severely disabled subjects (Figure 3). Extinction had previously been found to be effective in reducing the frequency of target behaviors, but it had the unfortunate effect of sometimes causing more problematic behaviors to emerge. By alternating phases consisting of extinction alone and extinction plus communication training, the authors were able to show that the second condition resulted in uniform decreases in problematic behavior.

(iii) Changing criterion

The changing criterion design element consists of a series of shifts in an expected benchmark for behavior, such that the correspondence between these shifts and changes in behavior can be assessed. It is particularly useful in the case of behaviors that tend to change gradually, provided that some benchmark, goal, criterion, or contingency is a key component of treatment.

The establishment of increasingly strict limits on the number of times a client smokes per day provides a simple example. A changing criterion design could be implemented when it is unclear as to whether the criteria themselves, and no other variable, were responsible for observed changes in smoking. As another example, a changing criterion design could be implemented to assess the degree to which changing minimum numbers of social contacts per day affects actual social contacts in a socially withdrawn client.

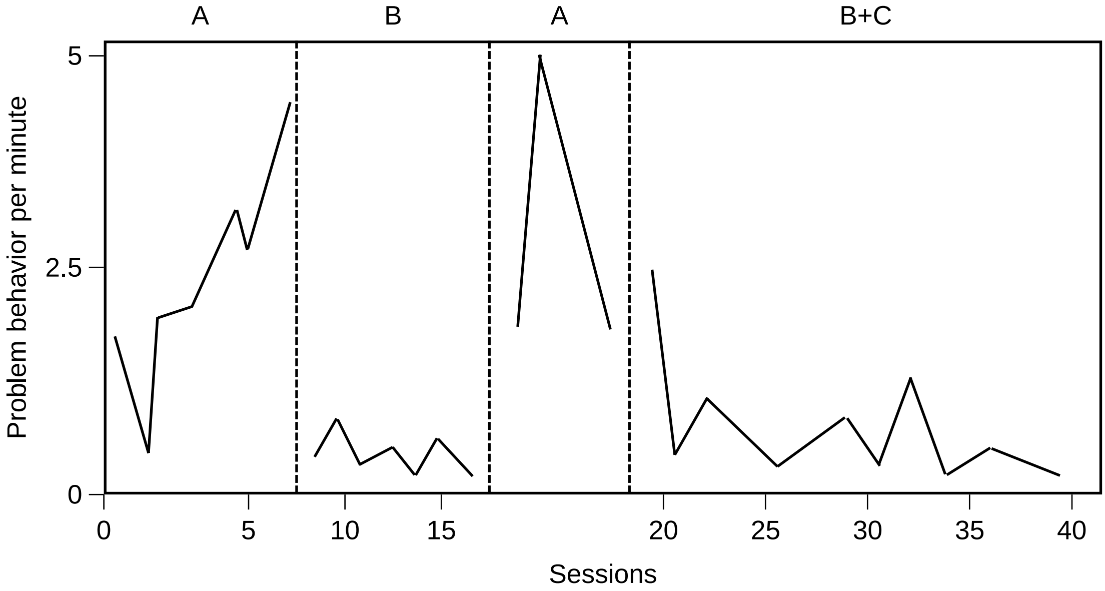

Figure 3 An interaction element design. The interaction (B + C) condition did not yield a better result than the B condition, but the demonstration of no further efficacy was still an important finding (adapted from Shukla & Albin, 1996).

In order to maximize the possibility that the data from a changing criterion design are correctly interpreted, five heuristics seem useful. First, the number of criterion shifts in the design should be relatively high. Each criterion shift is a replication of the effect of setting a criterion on subsequent client behavior. As with any type of experiment, the more frequently the results are replicated, the more confident the researcher can be of the effect and its causes. As a rule of thumb, four or more criterion shifts should occur when a changing criterion design is implemented.

Second, the length of the phase in which one level of the criterion is in effect should be long enough to allow the stability, level, and trend of the data to be interpreted relative to the criterion. Additionally, if a clear trend and level do not initially emerge, the criterion should remain in effect until a clear trend and level does emerge. Belles and Bradlyn (1987) provide a good example of properly timed criterion changes. The goal was to reduce the smoking rate of a long-time smoker who smoked several packs a day. The client recorded the number of cigarettes smoked each day (with reliability checks by the spouse). After a baseline period, goals were set by the therapist for the maximum number of cigarettes to be smoked. If the criterion was exceeded, the client sent a $25 check to a disliked charity. For each day the criterion was not exceeded, $3 went into a fund that could be used to purchase desirable items. Each criterion was left in place for at least three days, and the length of time, magnitude, and direction of criterion shifts varied. It should also be noted that, as in the Belles and Bradlyn (1987) study, criterion changes should occur at irregular intervals. As certain behaviors may change in a cyclical or gradual manner naturally, criteria should be shifted after differing lengths of time. Thus, if the length of one criterion's phase happens to correspond to changes occurring naturally, the phase lengths of other levels of the criterion will be unlikely to continue this trend.

Third, criterion changes should occur at irregular intervals. As certain behaviors may change in a cyclical or gradual manner naturally, criteria should be shifted after differing lengths of time. Thus, if the length of one criterion's phase happens to correspond to changes occurring naturally, the phase lengths of other levels of the criterion will be unlikely to continue this trend.

Fourth, the magnitude of criterion shifts should be altered. If the data can be shown to track criterion changes of differing magnitudes, the statement that the criterion itself is responsible for observed changes can be made with a greater level of assurance.

Finally, a periodic changing of the direction of criterion shifts can be useful in assisting interpretations of effects. Such a strategy is similar to the reversal common in simple and complex phase changes. If client behavior can be shown to systematically track increasing and decreasing criteria, the data can be more confidently interpreted to indicate a clear effect of the changing criteria on those behavioral changes.

DeLuca and Holborn (1992) used the changing criterion design in analyzing the effects of a variable-ratio schedule on weight loss in obese and nonobese subjects (Figure 4). Phase sequences consisted of a baseline (no treatment) phase, three different criterion phases, an additional baseline phase, and finally a return to the third criterion phase. The criterion involved the number of revolutions completed on a stationary exercise bike during the allotted time in each phase. Criterion phases were determined by a calculation of 15% over the criterion in place for the previous phase; when the criterion was met, a subject would receive reinforcement in the form of tokens exchangeable for established reinforcers. Each phase (except for the five-session final phase) lasted for eight 30-minute sessions, resulting in eight data points per phase. The increasing variable ratio schedules were shown, through use of this design, to exert control over increased frequencies of revolution. Although the design did not exhibit some typical features of the changing criterion design, such as staggered phase length, criterion direction reversals, and varying phase change depths, the observation of clear effects were facilitated by the return to baseline phase and subsequent replication of the third criterion phase. In addition, the use of the design in an exercise program was prudent, as exercise involves gradual performance improvements of a type easily detected by the changing criterion design.

3.02.4.1.3 Parametric design

When it is suspected that different levels or frequencies of a component of an intervention might have a differential effect on client behavior, a parametric design can be implemented. Such designs are frequently used to assess the effects of different psychotropic drug dosages, but design is relevant for many other purposes. Kornet, Goosen, and Van Ree (1991) demonstrated the use of the parametric design in investigating the effects of Naltrexone on alcohol consumption.

Ideally, a kind of reversal can be incorporated into a parametric design, where levels of the independent variable in question were systematically increased and then decreased. As with the changing criterion element, showing that an effect tracks a raised and lowered standard bolsters the confidence with which an interpretation is made. Baseline data should usually be taken before and after changes in the parameter of interest. If certain levels of the parameter provide interesting or unclear data, alternations between those levels, even if not originally planned, can aid in clarification (e.g., if the sequence A/B/B′/B″/B‴/B″/B′/B/A was originally planned and the data spurs increased interest in the B and B′ levels, a sequence such as B/B′/B/B′ could be inserted or added).

An example of acceptable use of parametric design is provided by Stavinoah, Zlomke, Adams, and Lytton (1996). The experimenters

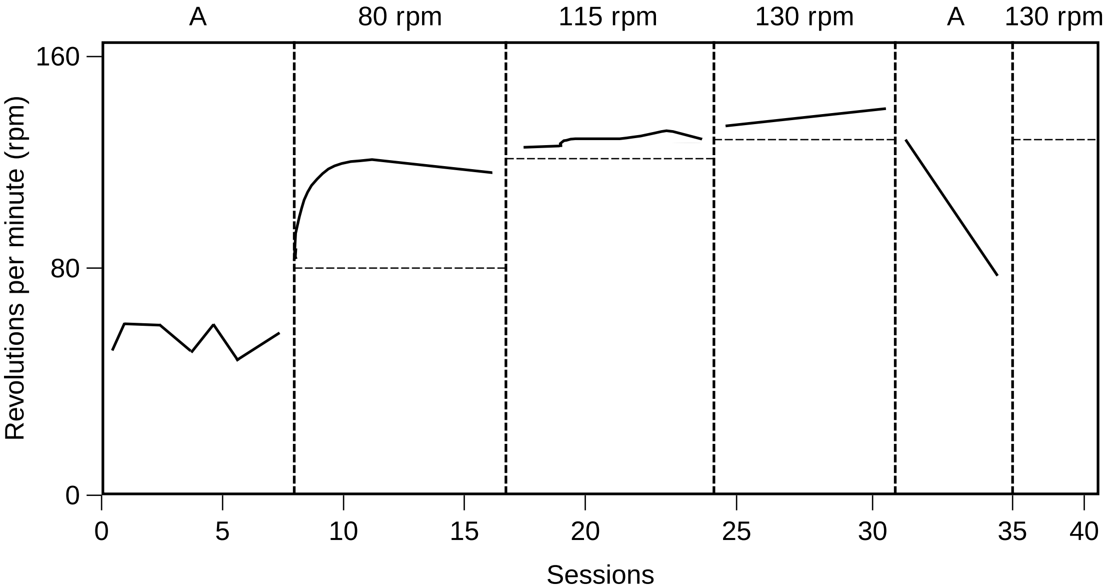

Figure 4 An example of a changing criterion design. Dotted horizontal lines indicate set criterion in each phase. The use of more than three criteria might have been preferable to further indicate experimental control over behavior, but the design seems adequate nonetheless (taken from DeLuca & Holborn, 1992).

systematically varied dosages of haloperidol and fluoxetine while measuring the frequency of the subject's impulsive aggressive behaviors (IABs). Dosages of haloperidol ranged from about 40 mg to 20 mg over the 40 weeks that it was administered. As a 40 mg dose initially resulted in an average of 10 IABs per week, dosage was reduced to 20 mg after 12 weeks. During the next 34 weeks, IABs increased to 13 per week, and the decision was made to increase the dosage back to 40 mg. This resulted in an average of 45 IABs per week for the four weeks this dosage was in place. The experimenters then administered a 20 mg dose of fluoxetine for the next 62 weeks, resulting in an average IAB frequency of near zero. A 40 mg dose of fluoxetine administered for 58 weeks yielded IAB frequencies of near zero. A subsequent reduction to 20 mg for five weeks resulted in an average IAB frequency of 12; the dosage was then returned to 40 mg, with a resulting IAB frequency of almost zero. Ideally, less time could have been spent at each dosage, and a greater variety of dosages could have been employed. But the experimenters did vary dosage and even drugs, and did so with a sufficient number of data points to determine the effect each drug and dosage had on behavior.

Lerman and Iwata (1996) provide a better example of use of the parametric design in treating the chronic hand mouthing of a profoundly retarded man (Figure 5). Sessions were conducted two or three times per day. Baseline frequencies (with no attempts to stop the hand mouthing behavior) were first calculated; baseline rates of three times per minute, on average, were recorded over several sessions. During the intervention phase, all subject attempts at hand mouthing were blocked by the experimenter putting his hand in front of the subject's mouth, resulting in less than one instance of hand mouthing per minute. Subsequently, attempts were blocked at a ratio of 1 block per 2 attempts, 1/4, 1/2, 2/3, and 3/4. The frequency of hand mouthing remained near zero over all levels. The experimenters properly used a descending/ascending order for levels, but also allowed subject behavior to determine what blocking schedule would be used. The experimenters thus remained responsive to the data, and their efforts yielded a less intensive intervention than one block per attempt.

3.02.4.1.4 Final suggestions for interpreting within-series data

Besides the suggestions listed above for interpreting various types of within-series data, additional general suggestions are offered here. First, as with any type of data-producing procedure, replicated effects are always more believable than nonreplicated effects. If an effect is consistently duplicated across several clients or across several different behaviors in the same client, the clinician can feel more confident in stating that the treatment in question is responsible for the effect. Each additional reinstatement of a treatment phase resulting in client improvement within a single series of data should also allow more confident statements about that treatment's efficacy with that client to be made. Second, effects are much more believable to the extent that they occur in a consistent manner. Third, changes of a greater magnitude (when parsing out changes apparently caused by extraneous variables) should generally be taken as more robust evidence of the treatment's effects. Fourth, effects occurring immediately after the onset of a treatment phase are logically stronger indicators that the treatment is responsible for the changes than are delayed effects, since fewer alternative explanations exist for the effects seen. Fifth, greater changes in the level and trend of the data are generally more indicative of a treatment's efficacy. Sixth, any effects not explainable by variables other than treatment should naturally be more convincing. Finally, all effects should be interpreted while considering the background variability of the data. The variability in the data around a given level and trends in a consistent condition provide an individual estimate of the impact of extraneous factors and measurement error against which any treatment effect is seen. If, for example, the level and/or trend of baseline data at times overlaps with the level and/or trend of treatment phase data, the clear possibility that factors other than treatment may be augmenting (or inhibiting) the treatment effect should be considered.

Brief mention of some commonly occurring conditions illustrate application of the guidelines discussed thus far. If a significant upward or downward data trend begins one to three data points before a phase change is planned, delay of the phase change until the data stabilizes is suggested. A trend is significant if it is of a greater degree than changes attributable to background variability, such as would be observed when a series of relatively stable data fluctuates around an average level. However, if an established and significant trend has emerged over a few (e.g., three or more) data points, a phase change might be acceptable if such results had been expected. Instability and unclear trends are obviously of less importance at the beginning of a phase than at the end; data at first uninterpretable often has a way of telling a clearer story after more data is collected. The

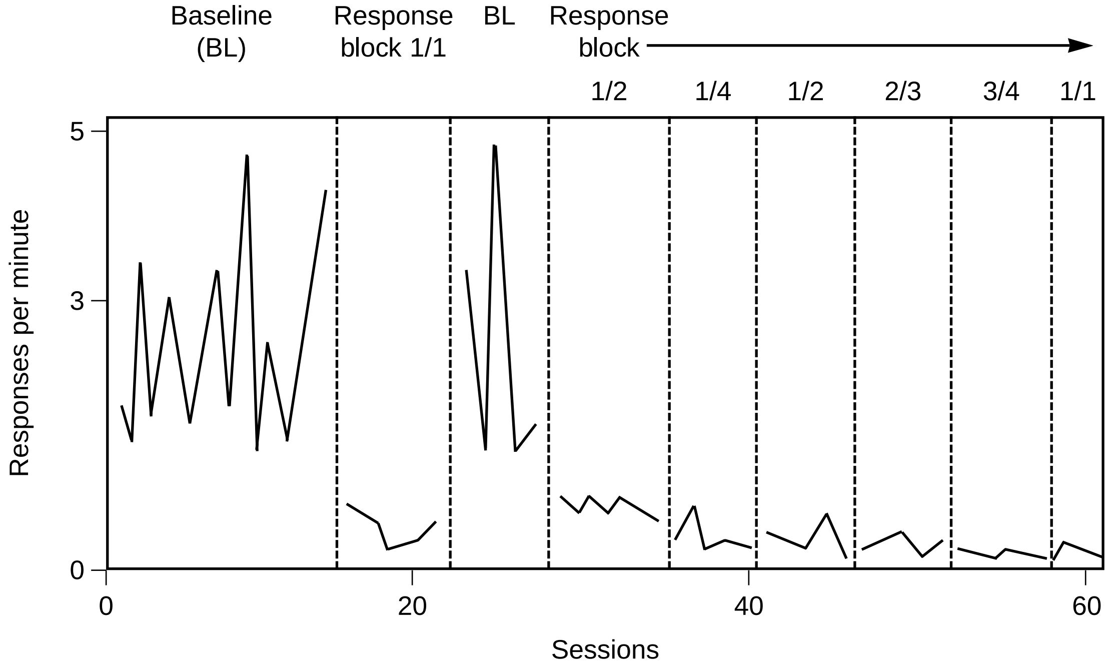

Figure 5 An example of a parametric design. Response block conditions refer to differing levels of intervention, that is, 14 translates to one response block per every four handmouthing attempts (adapted from Lerman & Iwata, 1996).

value of collecting as many data points as feasible (e.g., seven or more) becomes clear after only a few data points have been graphed.

3.02.4.2 Between-series Designs

Within-series designs involve comparisons of sequences of repeated measurements in a succession of consistent conditions. Between-series designs are based on a comparison of conditions that are concurrent or rapidly alternating, so that multiple data series are simultaneously created. Pure between-series designs consist of the alternating treatments design and the simultaneous treatment design.

3.02.4.2.1 Alternating treatment design

The alternating treatment design (ATD) consists of rapid and random or semirandom alteration of two or more conditions such that each has an approximately equal probability of being present during each measurement opportunity. As an example, it was observed during a clinical training case that a student therapist, during many sessions, would alternate between two conditions: leaning away from the client and becoming cold and predictable when he was uncomfortable, and leaning towards the client and becoming warm and open when feeling comfortable. The client would disclose less when the therapist leaned away, and more when he leaned forward. If it were assumed that the therapist had preplanned the within-session alternations, an ATD as shown in Figure 6 would be obtained. The condition present in the example at any given time of measurement is rapidly alternating. No phase exists; however, if the data in each respective treatment condition are examined separately, the relative level and trend of each condition can be compared between the two data series (hence the name between-series designs).

An example of the ATD is provided by Jordan, Singh, and Repp (1989). In this study, two methods of reducing stereotypical behavior (e.g., rocking, hand-flapping) in retarded subjects were examined: gentle reaching (the use of social bonding and gentle persuasion with the developmentally disabled) and visual screening (covering the client's eyes for a few seconds following stereotypic behavior, thus reducing visual stimulation including that provided by these movements). Each of the two conditions were randomly alternated with a baseline condition. After a baseline period, visual screening produced a dramatic reduction in stereotypy, whereas gentle teaching had only a transient effect.

Another proper use of the alternating treatments design is provided by Van Houten (1993; Figure 7). Four children were taught subtraction using two different procedures (one procedure involved the use of a general rule,

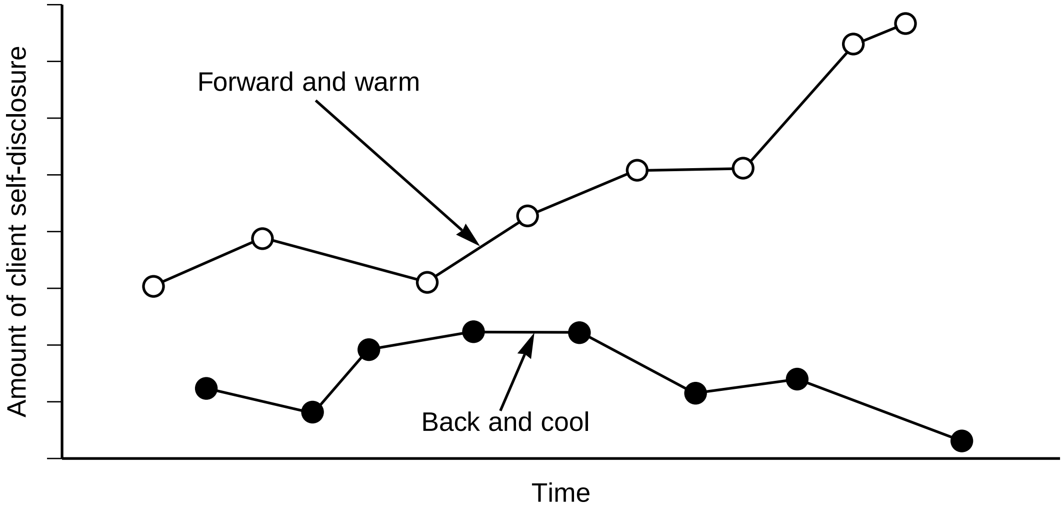

Figure 6 A hypothetical example of the use of an ATD to assess the effects of therapist behavior on client self-disclosure.

the other involved only rote learning). Use of the procedures was alternated randomly and every 15 minutes over the length of 15 or more sessions, and the subtraction problems used in each session were counterbalanced across the subjects so that effects could be attributed to the teaching methods and not the problem sets. The use of an ATD rather than a complex phase change was prudent, as the order the methods were presented in longer phases could probably have exerted a practice effect.

One of the benefits of the ATD is the simplicity with which it can be used to compare three or even more treatment conditions. Proper comparisons of three conditions in a within-series design can be difficult due to counterbalancing concerns, order effects, and the sheer number of phase changes that need to be executed over a relatively long period of time. With an ATD, three or even more conditions can be presented in a short time. The rapid and random alternations between conditions makes order effects less likely, but multiple treatment interference (the impact of one treatment is different due to the presence of another) is arguably likely. ATDs are ideally used with behaviors emitted at a relatively high frequency that correspondingly allows many instances of each alternate intervention to be applied. However, the design may be used with relatively infrequent behaviors if data is collected for a longer period of time. In addition, behaviors that tend not to have an impact for long after a discrete intervention is made and withdrawn make better targets for an ATD. If a change initiated by such a discrete intervention continues over a long period of time, effects of subsequent interventions are obscured and reliable data interpretation is often not possible.

ATDs hold several other advantages over standard within-series designs. First, treatment need not be withdrawn in an ATD—if treatment is periodically withdrawn, it can be for relatively short periods of time. Second, comparisons between components can be made more quickly. If a clear favorite emerges early in a well-conducted ATD, the clinician can be reasonably sure that its comparative efficacy will be maintained McCullough, Cornell, McDaniel, and Mueller (1974), for example, compared the relative efficacy of two treatments in four days using an ATD. ATDs can be used without collecting baseline data, or with baseline data through the creation of a concurrent baseline data series. Any background within-series trends (such as those due to maturation of the client or etiology of the disorder) are unlikely to obscure interpretation of the data because the source of data comparisons are purely between series, not within.

ATD requires a minimum of two alterations per data series. As both series can be combined to assist assessments of measurement error and extraneous factors, the number of data points required is less than with a within-series design. The collection of more than two data points per series is typical and highly useful, however. In a sense, each alternation is a replication and conclusions from all time-series designs can be stated with more confidence with each consistent replication.

When planning alternations, the clinician should be alert to the duration, after presentation, of a component's effect. An administered drug, for example, exerts an effect over a period of time, and presenting a new treatment component before that time has expired would confound interpretation.

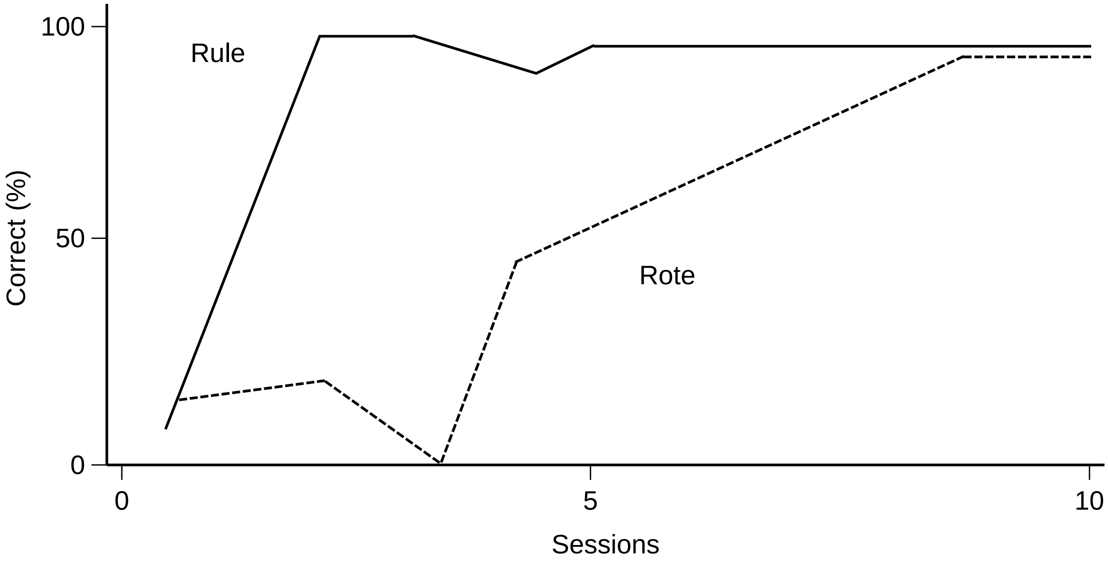

Figure 7 An ATD in which rule-learning trials are interspersed with rote-learning trials. A practice or generalization effect is apparent in the rote-learning condition beginning around session 8 (adapted from Van Houten, 1993).

One of the shortcomings of the ATD is that observed effects in the design can be due to the way in which conditions are presented and combined. Three areas of concern in this domain of multiple treatment interference are sequential confounding, carry-over effects, and alternation effects (Barlow & Hayes, 1979; Ullman & Sulzer-Aszaroff, 1975).

Sequential confounding occurs when there is a possibility that a treatment condition A yields different effects when presented before a treatment condition B than it does when presented after condition B. To control for sequential confounding, the clinician is encouraged to alternate treatment conditions randomly or at least semirandomly. With a randomly delivered sequence such as ABBBAABABBBABBAAAABBAAABAA, if consistent differences between each condition's effects continue to show up throughout the sequence, despite the fact that the order and frequency of each conditions' presence differs through the sequence, the clinician can be relatively certain that observed effects are not an artifact of order of condition presentation.

A carry-over effect occurs when the presentation of one condition somehow affects the impact of the subsequent condition, regardless of the presentation order of the conditions. Potentially this can occur in two ways. The effects of two conditions can change in opposite directions, or in the same direction. For example, when a strong reinforcer is delivered after a weak reinforcer, the weak reinforcer can subsequently cease to reinforce the desired behavior at all while the other reinforcer has a strong effect. At times, exposure to one condition results in a similar response to a somewhat similar second condition. Implementing each condition for a relatively short period of time can help reduce these problems (O'Brien, 1968), as might clear separations between each treatment condition (such as introducing only one treatment condition per session).

Several procedures exist to help detect multiple treatment interference (Sidman, 1960). A simple phase change where one treatment condition is preceded by a baseline phase, when compared to another AB design containing the other treatment, and finally compared to an ATD combining both conditions, could be used to parse out the separate and interactive effects of the treatment conditions. Alternatively, the intensity of one treatment condition could be increased, with any subsequent changes in the following conditions (as compared to changes already witnessed in an ATD containing both conditions) attributable to carry-over effects.

Some additional caveats regarding proper use of the ATD are of note. First, although a baseline phase is not necessary in an ATD, inclusion of baseline data can be useful for both gathering further information on client functioning and interpreting the magnitude of treatment effects. If periodic baseline points can be included within the ATD itself, valuable information regarding background level, trend, and variability can also be gleaned, over and above what would be interpretable if treatment conditions alone were present.

Second, it is important to realize that although ATDs can effectively be used with four or even more treatment conditions and corresponding data series, an upper limit exists on the number of data series that can be meaningfully interpreted. One useful heuristic (Barlow, Hayes, & Nelson, 1984) is to count the number of data points that will likely be collected for the ATD's purpose and then divide this number by the desired number of data series. If several data points will be collected for each series, the clinician should be able to proceed as planned.

Third, the clinician must consider the amount of data overlap between data series when interpreting ATD results. Overlap refers to the duplication of level between series. Several issues must be considered if considerable overlap exists. First, the percentage of data points, relative to all data points in an involved series, that indeed overlap can be calculated. If this percentage is low, the likelihood of a differential effect is higher. Second, the stability of the measured behavior should be considered. If the frequency of a given behavior is known to vary widely over time, then some overlap on measures of that behavior between ATD conditions would be expected. Third, the clinician must note if any overlapping trends occur in the two conditions. If data in two series are similar not only in level but also in trend, then it seems plausible that a background variable, rather than the treatment conditions, might affect the data.

One final example is shown in Figure 8. These are the data from an airplane-phobic client in the study on the effect of cognitive coping on progress in desensitization (Hayes, Hussian, Turner, Anderson, & Grubb, 1983). Notice that there is a clear convergence as the two series progress. The orderliness of the data suggested that the results from cognitive coping were generalizing to the untreated scenes. Alternating a reminder not to use the coping statements with the usual statements then tested this possibility. The data once again diverged. When the original conditions were then reinstated, the data converged once more. This showed that the convergence was a form of systematic generalization, rather than a lack of difference between the two conditions. This is also a good example of the combination of design elements to answer specific questions. This particular design does not have a name, but it is a logical extension of design tools discussed above.

3.02.4.2.2 Simultaneous treatment design

Simultaneous treatment design (STD) is similar to ATD in which the two treatments are continuously present but are accessed by the choice of the subject. What is plotted is not the impact of the treatment but the degree to which it is accessed. In other words, an STD measures preference. As an example, suppose a clinician wished to assess the motivation of a disabled child for different kinds of sensory stimulation. Several kinds of toys that produced different sensory consequences could be placed in a room with the child and the percentage of time played with each kind of toy could be recorded and graphed. This would be an STD.

3.02.4.3 Combined-series Elements

Combined-series designs contain elements from both within- and between-series designs, and combine them into a new logical whole. Although many examples in the literature contain elements of both between- and within-series designs, true combined-series designs involve more than merely piecing components together. What distinguishes combined series elements from any combination of elements is that the combination yields analytical logic.

3.02.4.3.1 Multiple baseline

One of the more often encountered SCEDs is the multiple baseline. An adaptation of the simple phase change, the multiple baseline design allows control over some variables that often confound interpretation of within-series phase change data. Upon a phase change from baseline to treatment in a simple phase change, for example, the data would ideally indicate a sudden and very strong treatment effect. It would be arranged in such a way that background variability could be easily detected against the effects of the treatment. Even when this ideal outcome occurs, however, there is also the possibility that some extraneous variable having a strong effect on measured outcomes might co-occur with the onset of a treatment phase. Such an occurrence would have dire implications for interpretation of the simple phase change result. The multiple baseline design allows considerable control over such threats to validity.

Essentially, the multiple baseline typically involves a sort of simple phase change across at least three data series, such that the phase change between baseline and treatment occurs at different times in all three data series (Figure 9). The logic of the design is elegantly simple. Staggering the implementation of each respective phase change allows the effects of extraneous variables to be more readily observed. It is generally unlikely that any given extraneous occurrence will have an equal effect on phase changes occurring at three different points in time.

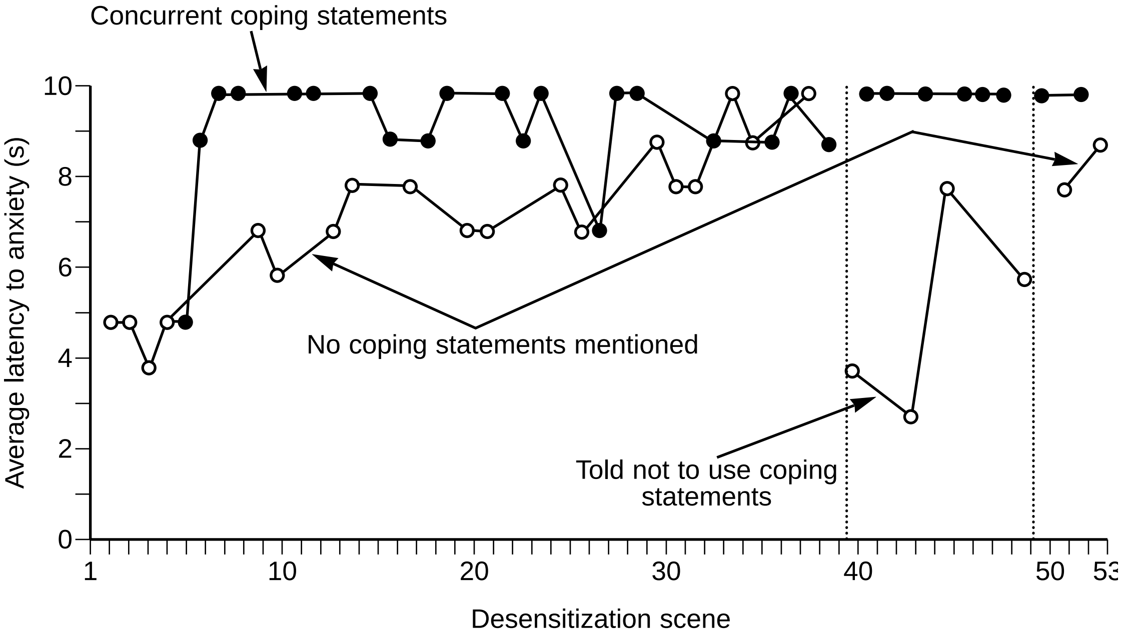

Figure 8 An example of series convergence in the ATD and its analysis by adding within-series components (redrawn from Hayes et al., 1983).

Implementation of a multiple baseline design greatly increases the potential number of comparisons that can be made between and within data series, ultimately strengthening the confidence with which conclusions are made from the data. Figure 3 details where such comparisons can be made. First, changes in level and trend within each data series (that is, between the baseline and treatment phase of each of the three data series) can be analyzed, just as with a simple AB design. Unlike a simple phase change, however, differences in level and trend between baseline and treatment can also be compared between three series of data. The design, in effect, contains replications of the phase change. If similar levels and trends are exhibited across all three series, the clinician can feel confident that the treatment is the most likely candidate for exerting the effect. Comparisons can also be made between the point of time at which the first phase change occurs, and the same points of time in the two remaining data series, where baseline data is still being collected. Such comparisons give the researcher information on whether some variable other than treatment might be responsible for observed changes. For example, a strong positive trend and marked change in level might be indicated by the data after the treatment phase is implemented. If similar changes occur in the other two data series at the same points in time, before treatment has even been implemented in those data series, it would seem clear that something besides treatment was exerting an influence on the data. A similar comparison can be made between the point at which the second phase change is implemented and the corresponding data points on the third series of data.

The type of data recorded in each of the three series must be similar enough so that comparisons can be made between the series, yet different enough that effects in one series are not expected from a phase change in another series. The context in which the data in each of the three series is collected can be of three varieties. The multiple baseline across behaviors requires that three relatively discrete and problematic behaviors, each of which might be expected to respond to a given treatment, be chosen. The treatment is then implemented in staggered, multiple baseline fashion for each behavior. The clinician would probably wish to choose behaviors that are unlikely to be subject to some generalization effect, so that a treatment implemented with behavior 1 results in concomitant change in one or both of the other behaviors (because, for example, the client begins to apply the principles learned immediately to those other behaviors). Such an effect would be easily observed in between-series comparisons (i.e., when a data trend in a condition where the intervention has not yet been initiated resembles the trend exhibited in an intervention condition, a generalization effect may very likely be present). However, the clinician could not be absolutely certain in such a case that the changes across behaviors were due to some generalization effect and not an extraneous variable. In

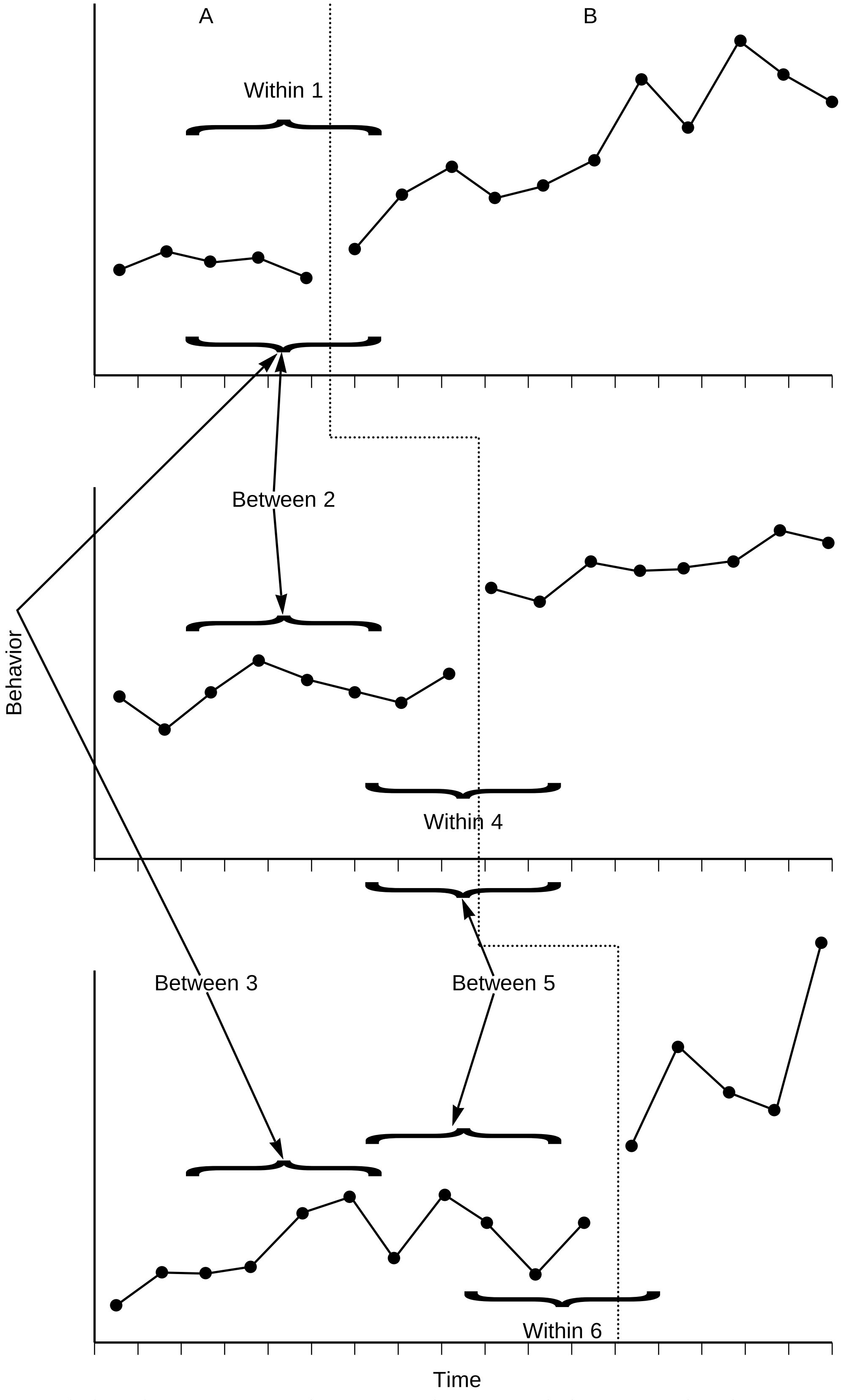

Figure 9 A hypothetical example of a multiple baseline, with between- and within-sources of data comparisons. Comparisons can be made between pre- and postphase change data within any of the three series, as well as between any two series at the points where an intervention is in place for one series and baseline conditions are still in effect for the second series.

general, if it seems theoretically likely that a generalization effect would occur under such circumstances, and no apparent confounding variable is present, then it is often safe to assume generalization occurred. Even if such an effect were observed, it might not be disastrous by any means. Implementing a treatment only to serendipitously find that it positively affects behaviors other than the targeted one could hardly be considered unfortunate.

An example of the multiple baseline across behaviors is provided by Trask-Tyler, Grossi, and Heward (1994). Developmentally disabled young adults were taught to use three cooking recipes of varying degrees of complexity (simple, trained, and complex), with the goal of the study being to determine whether or not specific skills would generalize across recipes. Simple tasks included preparing microwave popcorn while receiving specific instructions regarding its preparation. Trained tasks were analogs of previously taught simple tasks, where subjects used previously taught skills to prepare a different food (e.g., microwave french fries). Complex tasks consisted either of new tasks or novel combinations of previously trained tasks. Intervention phases were staggered across complexity levels, and occurred after a baseline phase where subjects completed recipes without instruction. Complexity level served as an appropriate analog for differing behaviors in this case, and the implementation of new phases were staggered with sufficient data points (i.e., three or over) occurring in each phase. The design use was also innovative in that it involved applied skills.

Alternately, the clinician may choose to implement a multiple baseline design across settings. If, for example, a client was socially withdrawn in a variety of circumstances (e.g., at work, at the gym, and on dates), social skills training might be implemented at different points in time across several of those circumstances. As another example, anger management training with an adolescent could be implemented in a staggered fashion at home, school, and at a part-time job site. As with the multiple baseline across behaviors, the clinician should be alert to the possibility of generalization effects.

An example of the multiple baseline across settings design is provided by Kennedy, Cushing, and Itkonen (1997; Figure 10). The study investigated the effects of an inclusion intervention on the frequency and quality of social contacts with nondisabled people, using two developmentally disabled children as subjects. The intervention included a number of components, such as placement in general school settings with nondisabled peers, and feedback. After a baseline where both subjects participated in special education classes only, the design took the shape of a multiple baseline across classes. One subject eventually participated in four classes, the other in two. Phase lengths were more than adequate (minimum of five data points), and phase changes were appropriately staggered (levels and trends stabilized before onset of the intervention in a new setting). The intervention was effective in increasing the frequency and quality of peer interactions.

Finally, a multiple baseline design can be implemented across persons. Such a manifestation of the design would, of course, require access to three clients with fairly similar presenting concerns subjected to the same treatment. As it is probably unlikely that anyone but a therapist at a university or health center would have simultaneous access to three such clients, it is acceptable to collect data on new clients as they present for services (Hayes, 1985). Objections exist to this approach (Harris and Jenson, 1985), but we feel strongly that the practicality of the approach and the control against extraneous factors that the design possesses greatly outweigh the potential risks.

The multiple baseline across persons design is exemplified by Kamps et al. (1992). The effects of social skills groups on the frequency of social skills interactions of high functioning autistic subjects with normal peers were analyzed across three subjects. Baseline consisted of frequency counts of social interactions before social skills training. Phase changes within each series did not occur until data stabilized, and phase changes across subjects were staggered approximately 10 sessions apart. A minimum of seven data points (usually more) were present in each phase. Social skills training had a positive effect on social interactions.

Ideally, phase changes should wait until the data indicates a clear level and trend, and is stable, before the phase change is executed. This is advisable when changing from baseline to treatment within a given series, and especially when executing staggered phase changes in the second and third data series. If no clear effect is observable in a series, then co-occurring data points between a series cannot be meaningfully interpreted.

An example of the multiple baseline is provided by Croce (1990). A multiple baseline across persons was used to evaluate the effects of an aerobic fitness program on three obese adults, while weekly body fat and cardiovascular fitness measurements were taken. Over three weeks of baseline data was collected in all subjects before treatment was delivered to the first subject. Treatment for the second subject was implemented three weeks after the first, and treatment for the third subject was delayed an

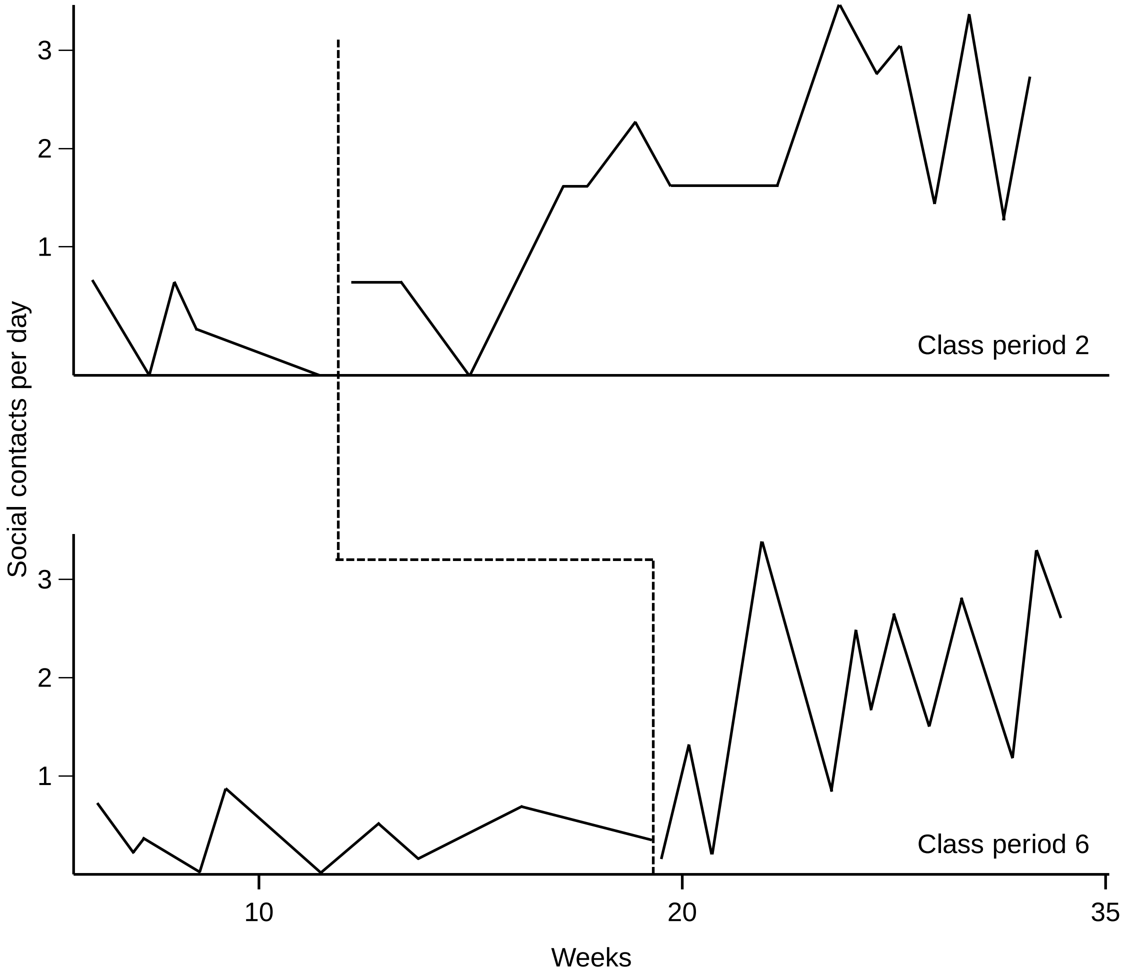

Figure 10 A multiple baseline across settings. Data was stable before phase changes in both settings, and comparisons between the settings indicate that the intervention was responsible for changes and that no generalization effect occurred (adapted from Kennedy et al., 1997).

additional three weeks. The data clearly indicated a desirable effect.

3.02.4.3.2 Crossover design

The crossover design essentially involves a modification of the multiple baseline allowing an additional degree of control over extraneous variables (Kazdin, 1980). It can be especially useful when only two (rather than three or even more) series of data can plausibly be gathered, as this added control tends to compensate for the control lost by the omission of a third series. This design has been widely used in pharmacological studies (Singh & Aman, 1981).

Execution of a crossover design simply involves simultaneous phase changes across two series of data such that, at any given point in time, opposite phases are in operation between each series. If series one's phase sequence was, for example, ABAB, a BABA sequence would be simultaneously delivered, with each phase change occurring at the same time in both series, and after treatment phases of precisely equal lengths. Such an arrangement can make even a relatively weak effect evident, as each treatment (B phase) always has a corresponding baseline (A phase) in the other data series to allow finer interpretations of stability, level, and trend.

3.02.4.3.3 Constant series control

One final combined series design warrants brief mention. A constant series control can be added to a series of data when baseline data alone is collected concurrently on some other person, problem, or situation of interest. In adding a constant series control to an in-school anger management intervention with a child, for example, relevant behavioral data might be collected at the child's home (where the treatment is not considered active) throughout the period where an ABAB phase sequence is being implemented at school. Treatment B effects from the school setting can then be compared to concurrently gathered baseline data from the home setting to assist in interpretation of the treatment's effects. Such a control is extremely useful when used in conjunction with a simple or complex phase change design.

A study by O'Reilly, Green, and Braunling-McMorrow (1990) provides an example of a baseline-only constant series control. O'Reilly et al. were attempting to change the accident-prone actions of brain-injured individuals. A written safety checklist that listed, but did not specifically prompt hazard remediation, was prepared for each of several areas of the home. If improvement was not maximal in a given area, individualized task analyses were prepared that prompted change in areas that still required mediation. The design used was a multiple baseline across settings (living room, kitchen, bedroom, and bathroom). Although phase changes from baseline to checklist prompts to task analysis occurred in staggered multiple baseline across the first three settings, a baseline condition remained in effect for the entire 29 weeks of the study. Results indicated very little evidence of generalization across responses, and the baseline-only constant series control provided additional evidence that training was responsible for the improvements that were seen.

3.02.5 FUTURE DIRECTIONS

Widespread use of SCEDs by practicing clinicians could provide a partial remedy to two currently omnipresent concerns in the mental health care delivery field: a virtual lack of use and production of psychotherapy outcome literature by clinicians, and the demand for demonstrated treatment efficacy by the managed care industry.

3.02.5.1 Potential for Research Production and Consumption by Practicing Clinicians

Line clinicians currently produce very little research. This is unfortunate since it is widely recognized that assessment of the field effectiveness of treatment technology is a critical and largely absent phase in the research enterprise (Strosahl, Hayes, Bergan, & Romano, in press).

Single-subject designs are well suited to answer many of the questions most important to a clinician. They are useful tools in keeping the clinician in touch with client progress and informing treatment decisions. Most of the requirements of these designs fit with the requirements of good clinical decision making and with the realities of the practice environment.

This has been true for some time, and the research production using these designs in the on-line practice environment is still limited. That may be about to change, however, for the reason discussed next.

3.02.5.2 Managed Health Care, Single-subject Design, and the Demonstration of Efficacy

The managed health care revolution currently underway in the USA represents a force that, in all likelihood, will soon encompass nearly all US mental health services delivery systems (Strosahl, 1994; Hayes et al., in press). The hallmark of managed care organizations (MCOs) is the provision of behavioral health services in a way that is directed, monitored, not merely compensated.

In generation I of the managed care revolution, cost savings accrued mostly to cost reduction strategies. That phase seems nearly complete. In generation II, cost savings are accruing to quality improvement strategies. Uppermost in this approach is the development of effective and efficient treatment approaches, and encouragement of their use through clinical practice guidelines.

Time-series designs are relevant to MCOs in three ways. First, they can allow much greater accountability. Single-subject designs provide an excellent opportunity for the clinician to document client progress and provide sufficient justification to MCOs, HMOs, and PPOs for implementing treatments or treatment components. Even a simple AB is a big step forward in that area. Second, when cases are complex or treatment resistant, these designs provide a way of evaluating clinical innovation that might be useful for improving the quality of treatment for other cases not now covered by empirically supported treatments. Finally, these designs can be used to evaluate the impact of existing treatment programs developed by MCOs.

3.02.6 CONCLUSION

This chapter provides a brief overview of current single-subject design methodologies and their use in the applied environment. The design elements are a set of clinical tools that are effective both in maximally informing treatment decisions and generating and evaluating research hypotheses. Single-subject designs fill a vital role in linking clinical practice to clinical science. With the evolution of managed care, this link is now of significant economic importance to a major sector of our economy.

3.02.7 REFERENCES

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1,* 91–97.

Barlow, D. H., & Hayes, S. C. (1979). Alternating treatments design: One strategy for comparing the effects of two treatments in a single subject. *Journal of Applied Behavior Analysis, 12,* 199–210.

Barlow, D. H., Hayes, S. C., & Nelson, R. O. (1984). *The scientist practitioner: Research and accountability in clinical and educational setting*. New York: Pergamon.

Barlow, D. H., & Hersen, M. (1984). *Single case experimental designs: Strategies for studying behavior change* (2nd ed.). New York: Pergamon.

Belles, D., & Bradlyn, A. S. (1987). The use of the changing criterion design in achieving controlled smoking in a heavy smoker: A controlled case study. *Journal of Behavior Therapy and Experimental Psychiatry, 18,* 77–82.

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. San Francisco: Holden-Day.

Busk, P. L., & Marascuilo, L. A. (1988). Autocorrelation in single-subject research: A counterargument to the myth of no autocorrelation: The autocorrelation debate. *Behavioral Assessment, 10*(3), 229–242.

Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin, 54,* 297–312.

Cone, J. D. (1986). Idiographic, nomothetic, and related perspectives in behavioral assessment. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 111–128). New York: Guilford Press.

Cope, J. G., Moy, S. S., & Grossnickle, W. F. (1988). The behavioral impact of an advertising campaign to promote safety belt use. *Journal of Applied Behavior Analysis, 21,* 277–280.

Croce, R. V. (1990). Effects of exercise and diet on body composition and cardiovascular fitness in adults with severe mental retardation. *Education and Training in Mental Retardation, 25*(2), 176–187.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist, 30,* 116–127.

Cummings, N. A., Cummings, J. L., & Johnson, J. N. (1997). *Behavioral health in primary care: A guide for clinical integration.* Madison, CT: Psychosocial Press.

Cummings, N. A., Pollack, M. S., & Cummings, J. L. (1996). *Surviving the demise of solo practice: Mental health practitioners prospering in the era of managed care*. Madison, CT: Psychosocial Press.

DeLuca, R. V., & Holborn, S. W. (1992). Effects of a variable ratio reinforcement schedule with changing criteria on exercise in obese and nonobese boys. *Journal of Applied Behavior Analysis, 25,* 671–679.

Edgington, E. S. (1980) Validity of randomization tests for one-subject experiments. *Journal of Educational Statistics, 5,* 235–251.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York: Appleton-Century-Crofts.

Gottman, J. M. (1981). *Time-series analysis: A comprehensive introduction for social scientists*. Cambridge, UK: Cambridge University Press.

Gottman, J. M., & Glass, G. V. (1978). Analysis of interrupted time-series experiments. In T. R. Kratochwill (Ed.), *Single subject research: Strategies for evaluating change* (pp. 197–235). New York: Academic Press.

Greenwood, K. M., & Matyas, T. A. (1990). Problems with the application of interrupted time series analysis for brief single subject data. *Behavioral Assessment, 12,* 355–370.

Gunter, P. L., Shores, R. E., Jac, K. S. L., Denny, R. K., & DePaepe, P. A. (1994). A case study of the effects of altering instructional interactions on the disruptive behavior of a child identified with severe behavior disorders. *Education and Treatment of Children, 17*(3), 435–444.

Harris, F. N., & Jenson, W. R. (1985). Comparisons of multiple-baseline across persons designs and AB designs with replication: Issues and confusions. *Behavioral Assessment, 7*(2), 121–127.

Hayes, S. C. (1985). Natural multiple baselines across persons: A reply to Harris and Jenson. *Behavioral Assessment, 7*(2), 129–132.

Hayes, S. C., Barlow, D. H., & Nelson, R. O. (1997). *The scientist practitioner: Research and accountability in the age of managed care*, (2nd ed.). Boston: Allyn & Bacon.

Hayes, S. C., Hussian, R. A., Turner, A. E., Anderson, N. B., & Grubb, T. D. (1983). The effect of coping statements on progress through a desensitization hierarchy. *Journal of Behavior Therapy and Experimental Psychiatry, 14,* 117–129.

Hersen, M., & Barlow, D. H. (1976). *Single case experimental designs: Strategies for studying behavior change*. New York: Pergamon.

Horne, G. P., Yang, M. C. K., & Ware, W. B. (1982). Time series analysis for single subject designs. *Psychological Bulletin, 91,* 178–189.

Huitema, B. E. (1988). Autocorrelation: 10 years of confusion. *Behavioral Assessment, 10,* 253–294.

Jordan, J., Singh, N. N., & Repp, A. (1989). An evaluation of gentle teaching and visual screening in the reduction of stereotypy. *Journal of Applied Behavior Analysis, 22,* 9–22.

Kamps, D. M., Leonard, B. R., Vernon, S., Dugan, E. P., Delquadri, J. C., Gershon, B., Wade, L., & Folk, L. (1992). Teaching social skills to students with autism to increase peer interactions in an integrated first-grade classroom. *Journal of Applied Behavior Analysis, 25,* 281–288.

Kazdin, A. E. (1980). *Research design in clinical psychology*. New York: Harper & Row.

Kennedy, C. H., Cushing, L. S., & Itkonen, T. (1997). General education participation improves the social contacts and friendship networks of students with severe disabilities. *Journal of Behavioral Education, 7,* 167–189.

Koan, S., & McGuire, W. J. (1973). The Yin and Yang of progress in social psychology. *Journal of Personality and Social Psychology, 28,* 446–456.

Kornet, M., Goosen, C., & Van Ree, J. M. (1991). Effect of naltrexone on alcohol consumption during chronic alcohol drinking and after a period of imposed abstinence in free-choice drinking rhesus monkeys. *Psychopharmacology, 104*(3), 367–376.

Lerman, D. C., & Iwata, B. A. (1996). A methodology for distinguishing between extinction and punishment effects associated with response blocking. *Journal of Applied Behavior Analysis, 29,* 231–233.

Levin, J. R., Marascuilo, L. A., & Hubert, L. J. (1978). N = nonparametric randomization tests. In T. R. Kratochwill (Ed.), *Single subject research: Strategies for evaluating change* (pp. 167–196). New York: Academic Press.

McCullough, J. P., Cornell, J. E., McDaniel, M. H., & Mueller, R. K. (1974). Utilization of the simultaneous treatment design to improve student behavior in a first-grade classroom. *Journal of Consulting and Clinical Psychology, 42,* 288–292.

O'Brien, F. (1968). Sequential contrast effects with human subjects. *Journal of the Experimental Analysis of Behavior, 11,* 537–542.

O'Reilly, M. F., Green, G., & Braunling-McMorrow, D. (1990). Self-administered written prompts to teach home accident prevention skills to adults with brain injuries. *Journal of Applied Behavior Analysis, 23,* 431–446.

Orsborn, E., Patrick, H., Dixon, R. S., & Moore, D. W. (1995). The effects of reducing teacher questions and increasing pauses on child talk during morning news. *Journal of Behavioral Education, 5*(3), 347–357.

Paul, G. L. (1969). Behavior modification research: Design and tactics. In C. M. Franks (Ed.), *Behavior therapy: Appraisal and status* (pp. 29–62). New York: McGraw-Hill.

Peterson, A. L., & Azrin, N. H. (1992). An evaluation of behavioral treatments for Tourette Syndrome. *Behavior Research and Therapy, 30*(2), 167–174.

Shakow, D., Hilgard, E. R., Kelly, E. L., Luckey, B., Sanford, R. N., & Shaffer, L. F. (1947). Recommended graduate training program in clinical psychology. *American Psychologist, 2,* 539–558.

Sharpley, C. F., & Alavosius, M. P. (1988). Autocorrelation in behavioral data: An alternative perspective. *Behavioral Assessment, 10,* 243–251.

Shukla, S., & Albin, R. W. (1996). Effects of extinction alone and extinction plus functional communication training on covariation of problem behaviors. *Journal of Applied Behavior Analysis, 29*(4), 565–568.

Singh, N. N., & Aman, M. G. (1981). Effects of thioridazine dosage on the behavior of severely mentally retarded persons. *American Journal of Mental Deficiency, 85,* 580–587.

Snow, R. E. (1974). Representative and quasi-representative designs for research in teaching. *Review of Educational Research, 44,* 265–291.

Stavinoah, P. L., Zlomke, L. C., Adams, S. F., & Lytton, G. J. (1996). Treatment of impulsive self and other directed aggression with fluoxetine in a man with mild mental retardation. *Journal of Developmental and Physical Disabilities, 8*(4), 367–373.

Strosahl, K. (1994). Entering the new frontier of managed mental health care: Gold mines and land mines. *Cognitive and Behavioral Practice, 1,* 5–23.

Strosahl, K., Hayes, S. C., Bergan, J., & Romano, P. (in press). Evaluating the field effectiveness of Acceptance and Commitment Therapy: An example of the manipulated training research method. *Behavior Therapy*.

Thorne, F. C. (1947). The clinical method in science. *American Psychologist, 2,* 159–166.

Trask-Tyler, S. A., Grossi, T. A., & Heward, W. L. (1994). Teaching young adults with developmental disabilities and visual impairments to use tape-recorded recipes: Acquisition, generalization, and maintenance of cooking skills. *Journal of Behavioral Education, 4,* 283–311.

Ulman, J. D., & Sulzer-Azaroff, B. (1975). Multielement baseline design in educational research. In E. Ramp & G. Semb (Eds.), *Behavior analysis: Areas of research and application* (pp. 377–391). Englewood Cliffs, NJ: Prentice-Hall.

Van Houten, R. (1993). Rote vs. rules: A comparison of two teaching and correction strategies for teaching basic subtraction facts. *Education and Treatment of Children, 16,* 147–159.

Wampold, B. E., & Furlong, M. J. (1981). Randomization tests in single subject designs Illustrative examples. *Journal of Behavioral Assessment, 3,* 329–341.

White, G. W., Mathews, R. M., & Fawcett, S. B. (1989). Reducing risks of pressure sores: Effects of watch prompts and alarm avoidance on wheelchair pushups. *Journal of Applied Behavior Analysis, 22,* 287–295.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

3.03
Group Comparisons: Randomized Designs

NINA R. SCHOOLER
Hillside Hospital, Glen Oaks, NY, USA

3.03.1 INTRODUCTION 47
3.03.2 COMPARING STRATEGIES FOR TREATMENT EVALUATION 48
  3.03.2.1 Case Reports and Summaries of Case Series 48
  3.03.2.2 Single Case Experimental Designs 48
  3.03.2.3 Quasi-experimental Designs 49
  3.03.2.4 Randomized Experimental Designs 49
3.03.3 HISTORICAL BACKGROUND 49
3.03.4 APPROPRIATE CONDITIONS FOR RANDOMIZED GROUP DESIGNS 49
  3.03.4.1 Ethical Requirements 50
  3.03.4.2 Stage of Treatment Development 51
  3.03.4.3 Treatment Specification: Definition of Independent Variables 52
  3.03.4.4 Client/Subject Characteristics 52
  3.03.4.5 Specification and Assessment of Outcome: Definition of Dependent Variables 53
  3.03.4.6 Study Duration 53
3.03.5 STUDY DESIGNS 54
  3.03.5.1 Simple Two-group Comparisons 54
  3.03.5.2 Multiple-group Comparisons 55
  3.03.5.3 Multiple-group Comparisons that Include Medication Conditions 56
  3.03.5.4 Factorial Designs 58
3.03.6 INDIVIDUAL DIFFERENCES: WHAT TREATMENT WORKS FOR WHOM? 59
  3.03.6.1 Individual Difference Variable 59
  3.03.6.2 Post hoc *Examination of Predictors of Treatment Response* 59
3.03.7 CONCLUSIONS 60
3.03.8 REFERENCES 60

3.03.1 INTRODUCTION

A major function of clinical psychology is to provide treatment or other interventions for clients who suffer from mental illness or seek relief from problems. This chapter addresses the role of experimental evaluation of treatments as a source of data in deciding about the utility of a treatment. In the medical literature, such studies are referred to as randomized clinical trials (RCTs) (Pocock, 1983) and many recent studies of psychological treatments also use this terminology (e.g., Bickel, Amass, Higgins, Badger, & Esch, 1997). This chapter will consider the strengths and weaknesses of a number of strategies for judging treatment utility; review some historical background of the experimental study of treatment; examine the experimental designs that have been used to study psychological treatments and the circumstances under which such experiments are appropriate. In each of the sections regarding specific experimental designs, exemplar studies will be discussed in some detail in order to highlight relevant issues for evaluating the literature and for the design of future studies.

3.03.2 COMPARING STRATEGIES FOR TREATMENT EVALUATION

In order to set the experimental study of treatment into context, this section will review a variety of strategies for evaluation of treatment efficacy attending to their strengths and weaknesses.

3.03.2.1 Case Reports and Summaries of Case Series

This strategy represents perhaps the method with the longest history. For example, Brill (1938) cites a case treated in 1880 by Josef Breuer (cf p. 7) of a young girl with symptoms of "paralyses with contractures, inhibitions and states of psychic confusion." Under hypnosis she was able to describe the connections between her symptoms and past experiences and "... by this simple method he freed her of her symptoms." Another case described by Brill concerns a four-year-old child who became nervous, refused to eat, and had frequent crying spells and tantrums. The symptoms began shortly after the mother was separated from the child and she was "cured" soon after the mother returned. The mechanism that was posited to account for the effect was a disturbance in libido.

Much more recently, *Consumer Reports* (1994, 1995) report the results of a survey of 4000 subscribers who reported their experiences with psychotherapy broadly defined. Respondents to the survey reported in general that they had benefitted from their psychotherapy and, of particular interest, that more and longer psychotherapy was associated with greater benefit. This survey has been reviewed and received careful critique by psychologists (Brock, Green, Reich, & Evans, 1996; Hunt, 1996; Kotkin, Daviet, & Gurin, 1996; Kriegman, 1996; Mintz, Drake, & Crits-Christoph, 1996; Seligman, 1995, 1996).

These examples differ in a number of ways. The most obvious is that they span over a century. The more relevant one for the present purpose is that the *Consumer Reports* survey carries the weight of numbers and implied authority as a result. However, both single case reports and reports that summarize large numbers of cases share limitations. Neither can address the very compelling question regarding the reported outcomes—"compared to what?" In the *Consumer Reports* survey, there are a number of obvious hypotheses that can be entertained regarding the absent comparison groups that call into question the validity of the reported finding. The finding that psychotherapy was helpful could be because those who did not feel they had benefitted from their psychotherapy did not respond to the questionnaire. The finding that longer psychotherapy was more helpful could be because those who experienced benefit continued in treatment longer or because those who discontinued early needed to justify their discontinuation. Statistical modeling of alternative causal interpretations cannot fully exclude these interpretations. The common characteristic of both the attributed interpretations of findings from case reports or case series and the criticism of them is that there is no formal procedure for judging the accuracy of the interpretation.

Such reports serve the important function of generating hypotheses regarding treatments and interventions that can lead to well-designed studies. But they are not designed to test hypotheses and therefore cannot provide definitive evidence for treatment efficacy.

3.03.2.2 Single Case Experimental Designs

Chapter 2, this volume, on single case designs, discusses the methodologies for assessing causal relationships in individuals. A striking advantage of single case designs is that they can be used to evaluate treatments for individuals with very specifically defined characteristics. These characteristics do not need to be shared with large numbers of other patients or clients for the experiments to be valid. The applicability of the specific interventions studied to broader populations of clients will, of course, depend on the similarity of such clients to the subjects studied. An important disadvantage of single case designs is that they are only applicable in conditions in which return of the target complaints can be reliably demonstrated once the treatment has been removed. If a treatment results (or is hypothesized to result) in a change in personality or in the orientation of an individual to himself or others, then the repeated administration and withdrawal of treatment that is the hallmark of single case designs will not provide useful information. A further disadvantage is that the comparisons are all internally generated and generalization is accomplished only through assessment of multiple cases with similar individual characteristics and treatment.

3.03.2.3 Quasi-experimental Designs

These methods, described in detail in Chapter 4, this volume, take advantage of naturalistic treatment assignment or self-selection and use statistical methods to model outcome and to provide controls for the major limitation of these methods—namely that assignment to treatment is not random. A further advantage is that evaluation of treatment outcome can be independent of treatment delivery (and even blind to the knowledge of treatment) so that the effects of expectancy on outcome are controlled. The major weakness of such designs is that the biases of treatment selection are allowed full rein.

3.03.2.4 Randomized Experimental Designs

Randomization is the hallmark of these studies. Random assignment to intervention provides the potential to control for biases that treatment self-selection and choice introduce into the evaluation of treatment outcome. The disadvantages that randomized treatment studies have include the logistical difficulties in implementation and restrictions in generalizability. Generalizability is restricted to clients who are willing to consider randomization. Individuals who come to treatment facilities with a strong preference or expectation may well reject randomization to treatment as an option and insofar as matching to treatment influences outcome, the absence of such subjects reduces generalization. Subject refusal to participate, particularly if it occurs after treatment assignment is known, can further compromise the advantage that random assignment introduces. Refusal after treatment assignment is known implies that subjects whose expectations are met continue and that those who did not get the treatment they preferred have chosen not to participate. If such refusal is frequent, the principle of randomization has been compromised and interpretation of findings will encounter limitations similar to those described for quasi-experimental designs. Attrition from all causes may introduce bias (Flick, 1988). Despite these disadvantages, randomization offers insurance that, within the population studied, bias—particularly from unknown sources, has been contained. Randomized experimental studies share with quasi-experimental designs the potential advantages afforded by the separation of evaluation of outcome from provision of the intervention.

3.03.3 HISTORICAL BACKGROUND

Early research in psychotherapy ran on two relatively independent tracks: the first emphasized the internal process of the psychotherapeutic encounter; the second was concerned with evaluating the effects of psychotherapy *per se*. In the latter studies the primary comparisons were between patients or clients who received psychotherapy and those who did not. Most controls centered around controls for treatment time such as deferred entry into treatment using waiting lists. In these experimental studies, attention to the content of the psychotherapy was descriptive. This position was strongly bolstered by the writings of theorists such as Frank (1971), who held that general constructs such as positive regard and the therapeutic alliance underlay the effects of psychotherapy.

Smith and Glass (1977) revolutionized the field with the publication of the first meta-analysis of the psychotherapy literature, demonstrating a moderate effect size for the comparison between psychotherapy and no treatment (see Chapter 15, this volume, for a discussion of current meta-analytic methods). At the same time, increased attention to measurement of change in psychotherapy (Waskow & Parloff, 1975) and the development of manualized treatments (see Chapter 9, this volume) signaled a change in the models that were available for the study of psychological treatments.

During the same period, attention to methodology in clinical psychology quickened (Campbell & Stanley, 1963, 1966; Cook & Campbell, 1979). The emergence of medications for the treatment of mental illness had inaugurated experimental studies of these treatments using rigorous designs drawn initially from clinical psychology that included randomization, double-blind administration of medication, placebo controls, diagnostic specificity, rating scales that paid attention to specific signs and symptoms of psychopathology (see Prien & Robinson, 1994, for a recent review of progress in methods for psychopharmacology). The demand for similar studies of psychological interventions became apparent. The modern era of experimental studies of psychotherapy and other psychological interventions was upon us.

3.03.4 APPROPRIATE CONDITIONS FOR RANDOMIZED GROUP DESIGNS

Experimentation is best suited to address comparative questions that would be biased if the groups being compared were self-selected. Random assignment of individuals to groups controls for biases introduced by self-selection. Some questions—for example, those regarding differences between young and old people or men and women—obviously will not be addressed experimentally, although later in the chapter the question of designs to study individual differences in relation to an experimental condition will be considered.

Some examples: does marital therapy prevent divorce?; is cognitive-behavior therapy helpful in depression?; does family psychoeducation potentiate the effects of medication in schizophrenia?; should children with attention deficit disorder receive medication or family support?; does respite care for Alzheimer's patients reduce physical illness in caregivers?; does medication reduce craving in alcohol dependent individuals? In all these examples, an intervention is either contrasted to others or in the cases where no comparison is stated, it is implied that the comparison is with a group that does not receive an alternate intervention.

The hallmark of all these questions is that they identify a group or condition and ask about an intervention. The implication is that the findings from the experiment will apply to members of the group who did not receive the intervention in the experimental study. This assumption represents a statistical assumption—it is required for the use of virtually all statistical tests applied in evaluating the questions under review. It also represents an important clinical concern. Experiments are not conducted solely for the benefit of the subjects who participate in the research. It is anticipated that the conclusions will be applicable to other individuals who share characteristics (ideally key characteristics) with those who were subjects in the research.

Perhaps the greatest threat to the validity of experimental studies is that there is something that distinguishes the members of the group (depressed patients, children with attention deficit disorder) who agree to randomization from those who will not participate in a study of randomized treatment. In the general medical clinical trials literature where the conditions to be treated are defined by readily measurable signs and symptoms or by physiological measurements (e.g., tumor size), that assumption is readily accepted and is therefore generally untested. In the psychological literature the question may be considered but there are very few studies that have actually drawn a sample of subjects from a population in order to know whether the subjects of a study are like the population from which they came. In other words, although we look to randomization as a control for bias in evaluating the effect of a treatment or intervention, there has been little attention paid to whether subjects who agree to randomization represent a bias that is negligible or substantial.

3.03.4.1 Ethical Requirements

Certain ethical conditions regarding the subjects and the treatments or interventions must be met (Department of Health and Human Services, 1994). Research participation requires informed consent from the participant or from an appropriately authorized surrogate, for example, a parent for a minor child. Even in cases where consent is obtained from a surrogate, the research participant (e.g., the child) must provide "assent." Assent is the term used for agreement to participate by individuals who are judged to lack the capacity to provide fully informed consent. Currently in the USA regulations are promulgated by the Department of Health and Human Services that define 12 elements of informed consent. They are:

(i) An explanation of the procedures to be followed, including specification of those that are experimental.

(ii) A description of the reasonably foreseeable attendant discomforts and risks and a statement of the uncertainty of the anticipated risks due to the inherent nature of the research process.

(iii) A description of the benefits that may be expected.

(iv) A disclosure of appropriate and available alternate procedures that might be advantageous for the subject.

(v) An offer to answer any inquiries concerning the procedures.

(vi) A statement that information may be withheld from the subject in certain cases when the investigator believes that full disclosure may be detrimental to the subject or fatal to the study design (provided, however, that the Institutional Review Board (IRB) has given proper approval to such withholding of information).

(vii) A disclosure of the probability that the subject may be given a placebo at some time during the course of the research study if placebo is to be utilized in the study.

(viii) An explanation in lay terms of the probability that the subject may be placed in one or another treatment group if randomization is a part of the study design.

(ix) An instruction that the subject may withdraw consent and may discontinue participation in the study at any time.

(x) An explanation that there is no penalty for not participating in or withdrawng from the study once the project has been initiated.

(xi) A statement that the investigator will inform the subject of any significant new information arising from the experiment or other ongoing experiments which may bear on the subject's choice to remain in the study.

(xii) A statement that the investigator will provide a review of the nature and results of the study to subjects who request such information.

All these essential elements must be included in an Informed Consent Form to insure adequate consent to participate in any research.

In studies that involve randomization to treatment, some elements are of particular importance because they represent concepts that are sometimes difficult for potential subjects to appreciate. The first, of course, is that of randomization itself—that treatment will be decided "as if tossing a coin." The second is that the clinician or treater will not choose "the best" treatment for the client. Other elements of informed consent that need to be emphasized for a truly informed process is that the subject is free to withdraw at any time and that those providing treatment may also discontinue the subject's participation if they believe it to be in the subject's best interest.

In most countries research involving human subjects must be reviewed by local Research Ethics Committees or IRBs that are mandated to insure that the research meets ethical standards, that all elements of informed consent are present, and that the interests of subjects are protected.

A number of issues regarding consent that go beyond the 12 critical items are currently a subject of debate. Participants in these deliberations include various regulatory bodies, independent commissions, and private individuals and groups that are self-declared protectors of those who may participate in experiments regarding treatments or interventions of interest to clinical psychology. Among these issues are the following: requiring that an independent person unaffiliated with a given research project oversee the consent process; expanding the definition of those who lack the capacity to provide informed consent to include populations such as all those diagnosed with schizophrenia and other illnesses; restricting the conduct of research that does not offer direct benefit to the participant subjects; elimination of the use of placebo treatment conditions in patient groups for whom there is any evidence of treatments that are effective. These potential changes may well alter the nature of experimental research regarding treatment and intervention.

In general, concerns about placebo treatment or no-treatment control conditions have been hotly debated in the arena of pharmacological treatment rather than psychological interventions. However, as evidence regarding the efficacy and effectiveness of psychological interventions becomes widely accepted, experimental strategies such as waiting list controls, "usual care," and even nonspecified psychotherapies may require re-evaluation. One strategy that may receive increasing attention and popularity is that of treatment dosage. In other words, a comparison group may be defined as one that receives the intervention of interest but receives less of it.

Interventions and control conditions being compared should be potentially effective; there should be evidence regarding benefit. If this assertion is accepted, how can a "no-treatment" comparison group be included in an experiment? First, there should be provision for the subject to receive the potentially more effective intervention following study participation. In other words, the "no-treatment" group is really a delayed treatment or the time hallowed waiting list condition. This may be relatively easy to implement with short-term treatment interventions but more difficult for long-term interventions. Alternate solutions include the use of minimal treatment conditions. Later in the chapter, both of these conditions will be discussed in more detail from the substantive perspective. In the present context, both represent plausible solutions to the ethical dilemma of providing appropriate treatment for client populations for whom there exist a corpus of information regarding treatment effects. Obviously, these concerns are moot in specific patient or client populations where there are no data regarding treatment effects. The ethical dilemma is heightened when the intervention is long term or there may be substantial risk if treatment is deferred. For example, studies of treatment for depression generally exclude subjects who are suicidal.

3.03.4.2 Stage of Treatment Development

The ideal sequence of treatment development includes the following stages: innovation; preliminary description in practice, usually by the innovators or developers of the treatment; comparative experimental studies to determine treatment efficacy and investigate client characteristics that may be linked to treatment response; dissemination into clinical practice; and evaluation of outcome under conditions of usual clinical care. Experimental, randomized designs may take several forms during the sequence outlined. The first is often a study that compares the new treatment with no treatment. As indicated earlier, that condition is most likely to be defined by a waiting list condition or deferred treatment. Other studies that provide comparisons to other established treatments may follow. At the present time, studies that examine psychological interventions in relationship to medication are often carried out. These studies may address direct comparative questions regarding medication and a psychological intervention or they may address relatively complex questions regarding the additive or interactive effects of medication and a psychological intervention.

3.03.4.3 Treatment Specification: Definition of Independent Variables

As indicated above, randomized designs are most valuable when the interventions can be well-specified or defined. Since the goal of such research is generalization, the advantage of well-specified treatment interventions is that their reproduction by other clinicians is possible. Manuals such as those discussed in Chapter 9, this volume, represent the ideal model of treatment specification, but other methods can be considered. The training of the treatment provider can provide a plausible statement regarding specification. For example, clinical psychologists with a Ph.D. who are board certified and who have completed an established training course may represent an appropriate statement regarding a treatment specification.

Specification should also include such elements as treatment duration, frequency, and number of contacts. Duration and number may appear to be relatively simple constructs but in examining psychological interventions the duration of treatment may not be rigidly defined. In some cases, treatment adherence by clients will affect duration, frequency, and number. In other circumstances, treatment may be continued until a specified outcome is reached. Under these conditions, duration of treatment until a criterion of improvement is reached may represent an outcome itself.

Finally, the most careful specification needs to be matched by measurement of adherence to the specified treatment. Did therapists do what they were supposed to do? This represents an interesting shift in research on interventions. Earlier in the history of psychotherapy research, when the emphasis was on the process of psychotherapy, the evaluation of what happened during the psychotherapeutic encounter was of interest because it was assumed that psychotherapeutic interventions altered outcomes. One could draw an analogy to recording the practices of a gifted cook in order to ascertain the recipe. As the emphasis has shifted to the evaluation of outcomes, the interest in the nature of the interpersonal interactions that comprise the intervention has come to be seen as assurance of adherence or "fidelity" (Hollon, Waskow, Evans, & Lowery, 1984). The question now is to determine whether a given cook has followed the recipe.

3.03.4.4 Client/Subject Characteristics

Characterization of the clients in the randomized trial provides the means for communicating attributes of the population of clients who may be appropriate for the treatment in routine clinical practice. Currently, the most common strategy for characterizing clients is the use of diagnostic criteria that provide decision rules for determining whether clients fit categories. The most widely used are those of the World Health Organization's *International classification of diseases* (World Health Organization, 1992) and the American Psychiatric Association's *Diagnostic and statistical manual of mental disorders* (American Psychiatric Association, 1987, 1994). The use of specified diagnostic criteria (and even standardized instruments for ascertaining diagnosis) provides some insurance that clients participating in a given study share characteristics with those who have been participants in other studies and with potential clients who may receive the treatment at a later time. However, under some circumstances it may be preferable to specify client inclusion based on other methods—such as the reason for seeking clinical assistance rather than a formal clinical diagnosis. An advantage of using inclusion criteria other than diagnosis is that problem-focused inclusion criteria may make translation of findings from a randomized trial to clinical practice easier.

A second issue regarding client characteristics in randomized studies is insuring that important client characteristics which may influence outcome are balanced across treatment groups. One strategy is to conduct relatively large studies. Randomization is designed to minimize imbalance but, as anyone who has seen heads come up 10 times in a row knows, in the short term randomization may well result in imbalance. One of the most common ways to achieve balanced groups is to randomize to treatment or condition within a prespecified group; for example, gender or age to insure against the possibility that by chance disproportionate numbers of one group are randomized to one treatment condition. A second strategy is to use an adaptive randomization algorithm (Pocock & Simon, 1975; White & Freedman, 1978). In this method, several client characteristics that are known or hypothesized to affect the outcomes of interest in the study are identified. The goal of the randomization algorithm is to insure that the groups are relatively well balanced considering all of the characteristics simultaneously. A particular algorithm may specify a number of characteristics but when the number exceeds three or four, the overall balance is unlikely to be affected. The characteristics can also be weighted so that some are more likely to influence treatment assignment than others. Adaptive randomization is a dynamic process in which subject characteristics are fed into a program that generates a treatment assignment such that the identified characteristics will be well balanced across all groups.

What is central to randomization within groups or adaptive randomization is the premise that the chosen characteristics are known or hypothesized to affect the outcomes of interest in the study. Characteristics that are not expected to influence outcome do not need to be considered in this way. Section 3.03.6 discusses the use of individual differences among study participants in attempting to understand outcome differences among treatments or interventions.

3.03.4.5 Specification and Assessment of Outcome: Definition of Dependent Variables

To state the obvious, outcome criteria should reflect the intended effect of the interventions studied and the reliability and validity of the assessment instruments used should be established. In general, there are advantages to using some measures that are established in the field. This provides the opportunity to compare findings across studies and aids in the cumulative development of knowledge (Waskow & Parloff, 1975). In addition, it is valuable to include study-specific measures that focus on the particular intervention(s) under investigation and hypotheses being tested. In discussing examples of specific studies in later sections of this chapter, the issue of whether differences were found on standard measures or on measures that were idiosyncratic to the study in question will be considered.

Also of relevance is who assesses outcome. In pharmacologic clinical trials, double-blind procedures represent standard operating procedure. Concerns are often expressed that side effect patterns and other "clues" may serve to break the blind, but blinding preserves a measure of uncertainty regarding treatment assignment (Cole, 1968). In studies of psychological intervention, double-blind conditions are not possible. The subject and the treating clinician know what treatment the subject is receiving and commitment and enthusiasm for treatment are assumed to be present. For this reason, a common strategy is to include independent assessors. It is not a perfect strategy. Such assessors have only limited opportunity to observe subjects and may not be sensitive to subtle but important cues because of their limited contact with the subjects. In this context, both initial training of assessors to insure reliability and ongoing monitoring of reliability are critical.

3.03.4.6 Study Duration

Study duration includes two components: duration of the intervention and duration of postintervention evaluation. Duration of the intervention should be theoretically driven, based on the nature of the clinical problem that is being addressed and the mechanism of action of the intervention. Short-term interventions are much easier to investigate, particularly short-term interventions for which manuals have been developed. It appears that manuals are easier to develop for short-term interventions. However, some questions will require longer-term intervention. Interventions whose duration are reckoned in months rather than weeks require additional care to avoid subject attrition and to insure that the treatment remains constant during the long period of time that it is being delivered.

Post-treatment evaluation or follow-up after treatment ends is common in studies of psychological interventions and addresses the important question of whether effects persist in the absence of continued treatment intervention. Such follow-up studies are subject to the criticism that uncontrolled interventions may have occurred during the post-treatment period and may bias the outcome. But, if an intervention is hypothesized to exert a long-lasting effect that may include change in personality or long-term functioning, such evaluation is required. Follow-up evaluations are also subject to increased problems of attrition and the risk that differential attrition may introduce bias. Comparisons of baseline and demographic characteristics are frequently made between those subjects who are ascertained at the end of follow-up and those who are lost. This provides some measure of comfort, but further important comparisons should include treatment group and measures of outcome at the completion of study treatment. An incorrect strategy that has fallen into disuse was to compare those not ascertained with the full cohort, including those lost to follow-up, so that those not ascertained were included in both groups.

In studies of pharmacologic treatment, reversal of effects on discontinuation of treatment has been taken as *prima facie* evidence of efficacy. In fact, a major research design strategy in psychopharmacology is to establish efficacy of treatment in a cohort and then to randomize subjects to continuation or discontinuation of medication. Differential levels of symptom severity after a fixed experimental period and time to symptom exacerbation are taken as evidence of treatment efficacy. In contrast, studies of psychological interventions have often considered persistence of effect following the end of treatment as evidence for the efficacy of the psychological intervention. The problem that these conflicting paradigms represent will be considered further in Section 3.03.5.3 which considers comparative effects of pharmacologic and psychological treatments.

3.03.5 STUDY DESIGNS

The choice of study design always involves compromise. The clever experimenter can always think of additional controls and variations. In the previous section some of the relevant issues have been highlighted. Ethical considerations may proscribe some scientifically appropriate conditions. For example, under some circumstances deferring treatment through the use of a "waiting list" control may be inappropriate. Financial resources may constrain designs. Aside from funding constraints, availability of appropriate subjects may limit the number of conditions that can be studied. This may be stating the obvious, but the more conditions and the larger the number of subjects in a study, the longer it will take to complete the study and the longer an important clinical question will remain unanswered. Each additional group in a study increases both the cost and the duration of a study. Finally, there is no single, ideal design. Ultimately, the design of a clinical experiment represents a decision that is driven by the hypotheses that are under investigation. What is critical is the recognition by the investigator (and by the reader/critic of the research) of the hypotheses that can be tested, and conversely, the questions that are simply not addressed by a given study.

3.03.5.1 Simple Two-group Comparisons

Designs involving only two groups are often used in the early stages of research with a given form of psychotherapy. The comparison group may be a no-treatment waiting list for the duration of the treatment, treatment "as usual" in the treatment setting, or a nonspecific psychological intervention. Nonspecific psychological interventions are sometimes referred to as "placebo" treatment in the literature. Comparisons of two specific psychological interventions also represent potential two-group comparisons. A final model may include a treatment and no-treatment comparison in the presence of a specified medication condition (see Section 3.03.5.3 for discussion of this design in the context of other designs that include medication conditions).

Such studies may evaluate the benefit of two specific psychological interventions, of a specific intervention vs. a nonspecific control (usual care) or of intervention vs. none. The hypotheses and specific clinical questions that drive a given investigation should, in principle, drive the particular nature of the experimental and control groups. However, sometimes logistic considerations such as the nature of the treatment setting and the treatments that are routinely provided in the setting may influence the design of the experiment. A further, important factor may be the clinical needs of the clients and the urgency of providing some intervention.

There are limitations of two-group designs. Whatever the outcome, the design will not allow testing of a range of alternate hypotheses. If the study does not include a no-treatment control group, there is no control for time, spontaneous remission, or improvement. If the study does not include an alternate intervention group there is no control for nonspecific factors in treatment or for the specific characteristics of the designated experimental treatment. Further, if the study includes two specified interventions, there is no control for either receipt of any treatment or for nonspecific factors in treatment.

Interpretation of the findings from two-group studies absent a no-treatment control is difficult if there are no differences in outcome between the groups. No difference can be interpreted as Lewis Carroll and the Red Queen would have us believe "that all have won and all shall have prizes" or that there is no effect of either. Interpretation of studies that do not include a no-treatment group may often depend on integrating findings from other prior studies that did include such controls—not necessarily a bad state of affairs. As the field of experimental studies of psychological interventions matures, it may become less appropriate to implement studies with no-treatment control groups.

In the following example of a two group design, Keane, Fairbank, Caddell, and Zimering (1989) compared implosive (flooding) therapy to a waiting list control in 24 subjects with post-traumatic stress disorder (PTSD). Interestingly, the study design had originally included a stress management group, but for unspecified reasons, this condition was not successfully implemented, and those subjects are not included in the report. Thus, the two-group comparison is between a specified treatment and a no-treatment control. Implosive therapy was manual driven, following a manual (Lyons and Keane, as cited in Keane et al., 1989), and included between 14 and 16 sessions. The experimental group received baseline, post-treatment (elapsed time from baseline not reported), and a six-month follow-up assessment. The wait-list control group was assessed twice: at baseline prior to randomization and after, on average, four months. Half of the implosive therapy group and the majority of subjects in the wait-list group received anxiolytic medication because of, as the authors note, "... concerns about placebo groups and no treatment controls ... we didn't attempt to withdraw these patients from the medications which were prescribed to them" (Keane et al., 1989, p. 249). The authors maintain that the comparison of implosive therapy to subjects who were receiving pharmacotherapy (even if it was not systematically administered) represented a more stringent test of implosive therapy—although it was not a part of the original design. Subjects were assessed using well-known standardized assessments scales for depression, trait and state anxiety, and instruments specifically designed by the investigators for the assessment of PTSD. Post-test assessments were completed by the therapist who treated the subject in the implosive therapy group and by one of the same therapists for the wait-list group. In addition, the subjects rated themselves on depression, anxiety, and satisfaction with social adjustment in several life areas.

Implosive therapy appeared to improve depression and anxiety according to self-report and specific features of PTSD as rated by the therapists. No changes in social adjustment were seen.

Strengths of the study include randomization, specification of diagnosis, apparent absence of dropouts in both conditions, the existence of a manual for the experimental treatment, and the use of both standard outcome measures and measures tailored to the specific study hypotheses. Although subjects were randomly assigned to treatment, the randomization was compromised by the fact that treatment in one of three groups to which subjects were randomized was, for unspecified reasons, not fully implemented. In other words, this study was a two-group design in execution rather than intent. Another weakness is that assessments were completed by the therapists who treated the subjects and who therefore knew the treatment assignment. As discussed in Section 3.03.4.5, when treatment assignment is known, there is a potential bias.

3.03.5.2 Multiple-group Comparisons

Multiple-group comparisons offer the opportunity to test hypotheses simultaneously regarding comparisons of interventions (specificity of effects) and of intervention to a no-intervention control. If the interventions being studied derive from theoretical models, their comparison may test specific hypotheses regarding the psychological mechanisms that underlie the condition being treated. Other more pragmatic questions that can be addressed in multiple-group comparisons include group vs. individual contact and treatment intensity. However, it should be noted that when intensity or dosage of the intervention becomes a condition, control for amount of contact is lost. Again, there is no single "best" design. The appropriate design depends on the questions being asked and the appropriate questions depend on the stage of research. For example, questions of treatment duration or frequency are more likely to be asked after a particular treatment has been shown to have an effect.

Brown and Lewinsohn (1984) compared three psychoeducational approaches in treating depression that all used the same course materials: class, individual tutoring, brief telephone contact, and a delayed treatment control. Sixty-three individuals who met diagnostic criteria for unipolar depression and paid a course fee were randomly assigned to one of the four groups. Half the course fee was refunded if subjects completed all evaluations. Subjects in the three immediate treatment groups were assessed pre- and post-treatment and at two later follow-up points. The delayed treatment control group was assessed at baseline and at eight weeks, the time equivalent of post-treatment after which they received the class condition. All treatments focused on specific behaviors believed to be dysfunctional in depression (although not necessarily in the study subjects) and followed a syllabus that included a text (Lewinsohn, Munoz, Youngren, & Zeiss, 1978) and a workbook (Brown & Lewinsohn, 1984). An independent interviewer completed a standardized diagnostic assessment at baseline and of symptoms at later assessment points. Subjects completed three standardized self-report measures. The primary outcome measure was a factor-analytically derived score that included all three self-report measures.

Improvement was seen in all three immediate treatment groups on this single composite measure compared to the delayed treatment group. There were no differences among the active treatment groups either in self-report or in rate of diagnosed depression at follow-up. Comparison of high and low responders to treatment found few differences; none were a function of specific treatment. No detailed assessment of psychopathology or self-report measures was made.

The strengths of the study include randomization, a design that tested specific effects of a treatment modality and the method of delivery, the absence of drop-outs from treatment or assessment, specification of treatment conditions, the use of standardized measures of outcome assessment, and the use of independent assessors. The major weakness of the study is not in its design but the limited use made of the assessment battery. Ratings made by the independent assessors were only used to assess the presence or absence of diagnosable depression at the six-month follow-up and the use of a single summary score for all self-report measures may well conceal more than it reveals. A second weakness is the absence of assessments of implementation—how well the instructors used the course materials in the three conditions and whether all three conditions were implemented equally well.

It is difficult to decide whether charging subjects for treatment is a strength or a weakness. On the one hand, payment for treatment is characteristic of clinical treatment settings. All subjects completed the study and it could be argued that motivation to receive a 50% rebate on the course fee was a factor in enhancing study completion. On the other hand, the authors report that one of the major reasons that screened potential participants did not enter the study was inability to pay. Thus, the study population may have been biased by the exclusion of these subjects.

This study exemplifies a model that deconstructs a psychotherapeutic approach—in this case a psychoeducational approach to the treatment of depression—in order to identify which treatment delivery strategy offers advantages in outcome given that the strategies differ in the amount of clinical time and resources required to deliver the treatment. All subjects, including the delayed treatment group, reported significant improvement. The implication of these findings is that the least costly method of treatment delivery, brief telephone contact, should be implemented. However, there was no report regarding the acceptability of the method of treatment delivery and as noted earlier, there was no assessment of post-treatment symptoms. In the absence of these data, perhaps a firm clinical recommendation is premature.

3.03.5.3 Multiple-group Comparisons that Include Medication Conditions

Although in principle, designs that compare medication to psychological interventions can be classified in terms of whether they are simple comparisons (Section 3.03.5.1), multiple comparisons (Section 3.03.5.2) or represent factorial designs (Section 3.03.5.4), they entail unique research challenges and therefore merit separate consideration. The first design challenge was introduced earlier, namely the difference in the model of assessing effects. Follow-up assessments after treatment ends are seen as important sources of information in studies of psychological treatments. In contrast, such assessments are rare in studies of pharmacotherapy. Relapse or re-emergence of symptoms after treatment discontinuation is taken as evidence of efficacy. Therefore, discontinuation designs to study efficacy of treatment are common in psychopharmacology. A second design challenge is whether psychological and pharmacologic treatments should be introduced at the same time or are appropriate for different stages of an illness. Third, pharmacologic studies use double-blind methods as the preferred strategy to control assessor bias and therefore studies often rely on treating clinicians to assess outcome. Because it is not possible to blind treating clinicians to the psychological treatment they are providing, studies of psychological interventions rely on independent assessors who, although they will be blind to treatment assignment, have reduced opportunity to observe, and have minimal opportunity to develop rapport with the subjects they are assessing. For these reasons, evaluations by independent assessors may be less sensitive than those made by treating clinicans. Fourth, psychological interventions are often hypothesized to affect different aspects of outcome than pharmacologic interventions, so that direct comparisons of effects may be difficult.

Historically, pharmacologic clinical trials have relied more heavily on medical diagnoses of mental illness than studies of psychological interventions. But current research in both pharmacologic and psychological treatment is largely wedded to specific diagnoses derived from the *DSM* of the American Psychiatric Association (1987, 1994) or the *ICD* of the World Health Organization (1992). Parenthetically, one can question whether the increased reliance on medical diagnostic schemes such as *DSM* and *ICD* represents an advance in studies of psychological intervention. It has been argued that diagnostic specificity enhances reliability and therefore enhances replication and communication. On the other hand, clients who seek treatment do not necessarily seek help for a specific diagnostic label.

A wide range of designs have been used to examine the relationship of medications and psychological treatments. In the simplest, a medication or placebo is added to a uniform psychological intervention, or the converse, a psychological intervention or control condition is added to an established medication condition. These designs examine the additive effect of the modality that is superimposed. Questions like does naltrexone enhance abstinence in alcoholic subjects who are in a therapeutic milieu or does social skills training enhance social functioning in schizophrenic subjects who are maintained on antipsychotic medication can be addressed in this manner. The major design challenges faced by such studies are: the inclusion of appropriate outcome measures to evaluate change on psychological interventions; insuring that the treatment modality which represents the background condition remains relatively constant during the course of the study; and insuring an adequate timeframe in which to evaluate effects of psychological treatments (in the design where psychological intervention is manipulated).

The second class of studies adds a medication condition (or conditions) to a multiple-group design such as those discussed in the previous section. These studies examine the comparative effects of medication and psychological treatments and face greater challenges. They have been called "horse races" and may thereby contribute to guild conflicts between psychologists and psychiatrists (Kendall & Lipman, 1991).

Appropriate inclusion criteria for medication and psychological treatments may differ and compromises may be necessary. Such studies will generally include double-blind conditions for medication and also benefit from inclusion of independent assessors. The studies require expertise in both clinical psychopharmacology and the relevant psychological interventions. Insuring this expertise for a study usually requires close collaboration and mutual respect by investigators drawn from psychology and psychiatry. Outcome criteria need to include measures of dimensions that are hypothesized to be sensitive to change in both modalities being studied and the reasons for inclusion of specific outcome measures should be explicit. A detailed review of these and other methodological considerations can be found in an article by Kendall and Lipman (1991) and in the article detailing the research plan for the National Institute of Mental Health Treatment of Depression Collaborative Research Program (Elkin, Parloff, Hadley, & Autry, 1985).

Finally, factorial designs provide explicit tests of both additive and interactive effects of medication and psychological treatment. These designs will be considered in Section 3.03.5.4.

An example of a study that compared medication and a specific psychotherapy is the study of prevention of recurrence of depression by Frank et al. (1990). Patients characterized by recurrent depression (at least two prior episodes) were treated with medication and manual-based interpersonal psychotherapy (IPT) (Klerman, Weissman, Rounsaville, & Chevron, 1984) until symptom remission was documented and maintained for 20 weeks. One hundred and twenty-eight subjects were then randomized to one of five treatment conditions: a maintenance form of interpersonal psychotherapy (IPT-M) that was characterized by less frequent visits; IPT-M and an antidepressant (imipramine) at the acute treatment dose; IPT-M and medication placebo; medication clinic visits and imipramine; medication clinic visits; and medication placebo. Subjects were treated for three years. Therapists were trained and certified by the developers of IPT. The major difference from the published manual is described as visit frequency. The major outcome examined was recurrence of depressive episodes during the three-year period. The two treatment arms that included imipramine (IPT-M and imipramine; imipramine and clinic visits) had the lowest risk of recurrence. Mean survival time was more than 83 weeks. The two groups that received IPT-M without medication (IPT-M alone; IPT-M and placebo) had a higher risk of recurrence; mean survival time was more than 60 weeks. The lowest survival time was seen in the group that received placebo and medication clinic visits; survival time was 38 weeks on average. The authors conclude that medication, at a relatively high maintenance dose, affords the greatest protection from recurrence in individuals who have experienced recurrent depression. However, psychotherapy, absent medication, represents a distinct advantage over clinic attendance coupled with a placebo.

The study has a number of strengths. Patients were randomized to treatment condition. There was an extremely low dropout rate (17%) over the three-year study duration. The design included control conditions for medication (placebo) and psychotherapy was administered under three conditions: alone, in combination with imipramine, and in combination with imipramine placebo. The psychotherapy followed an established manual and therapists were trained by the originators of the therapy. Survival analysis represents the most powerful statistical method for evaluating risk of recurrence.

Weaknesses include the limited examination of clinical outcomes. Although the authors mention the importance of judging interepisode social functioning, the article only examined recurrence risk. Second, the study population was, by design, limited to those who both had recurrent episodes and recovered from them. Of the 230 patients who met the first criterion, fully 42% did not enter the randomized trial. This represents a substantial limitation to generalization of the findings to depression broadly defined.

3.03.5.4 Factorial Designs

Designs that include at least two factors allow for detection of main effects in each factor as well as interactive effects. In the simplest of these, 2 × 2 designs, there is a control condition or at least two defined levels of each factor, as shown in Figure 1. If the factors represent independently defined treatment modalities, then such designs can detect the ability of one treatment to potentiate the other or even to inhibit the effect of the other. For this reason, such designs are particularly suited to examining the relationship of medication and psychological interventions. Uhlenhuth, Lipman, and Covi (1969), Schooler (1978), and Hollon and DeRubeis (1981) have considered factorial models for the study of the interaction of medication and psychological treatments. According to Uhlenhuth et al. (1969), there are four possible forms that the interaction can take. These are shown in Figure 1. An additive effect is defined as the simple sum of the effects of the two treatment modalities. The effect in cell 1 is equal to the combined effects seen in cells 2 and 3. The treatments may potentiate each other. In that case, the combined effect (cell 1) is greater than the additive sum of the two treatment effects (cells 2 and 3). The treatment modalities may inhibit one another. In that case the combined effect (cell 1) is less than the effects of each treatment independently (cells 2 and 3). Finally, there may be an effect that they call reciprocal, in which the combined effect (cell 1) is equal to the main effect of the more effective treatment (the effect seen in either cell 2 or cell 3). Detection of all these effects depends on the factorial design and the presence of cell 4 which represents the combined control condition or as it has sometimes been dubbed, the double placebo.

Factorial designs are potentially a valuable model for examination of other important variables in understanding treatment, for example, to control for the effect of setting in studies that are conducted in more than one clinical setting. However, the most important use, from the perspective of this chapter, is to study the relationship of medication and psychological treatment.

Marks and his colleagues (1993) studied the effects of medication and a psychological treatment for panic disorder. There was prior evidence of efficacy for both the medication, alprazolam, an antianxiety drug, and the experimental psychological treatment, live exposure. One hundred and fifty-four subjects were randomly assigned to four treatment conditions: alprazolam and exposure (AE cell 1, Figure 1); alprazolam and relaxation, the control for the psychological treatment (AR cell 2, Figure 1); placebo and live exposure (PE cell 3, Figure 1); and placebo and relaxation, so-called double-placebo (PR cell 4, Figure 1). Pharmacologic and psychological treatment lasted for eight weeks, medication was tapered during the following eight weeks and subjects were followed for an additional six months to assess relapse or maintenance of gains following discontinuation of treatment. Both exposure and relaxation followed published guides, not treatment manuals (Marks, 1978; Wolpe & Lazarus, 1966). The study was conducted at two centers; one in London, UK, and the other in Toronto, Canada.

The primary outcome measures were: ratings of avoidance by an assessor and the subject; the number of major panics; work and social disability; and the clinician's rating of improvement. After eight weeks of treatment, there were significant main effects of both alprazolam and exposure. The effects of alprazolam, as would be expected for a pharmacologic treatment, did not persist during follow-up, whereas the effects of exposure persisted after treatment discontinuation. Interpretation of findings is complicated by the fact that there was substantial improvement in total major panic in the "double-placebo" group that received medication placebo and relaxation.

In addition to reporting statistical significance of outcome measures, the article also reports the effect size of various outcome measures (Cohen, 1988, 1992). With large samples, differences that are clinically uninteresting may be statistically significant, that is, they are unlikely to be due to chance. The effect size is less influenced by sample size and may therefore be considered a better indicator of clinical relevance. During the first eight weeks, when both treatments were being delivered and were significantly better than their controls, the effect size is larger for exposure than for alprazolam. According to the definitions of Uhlenhuth et al. (1969), this would be a

		Psychological treatment	
		Experimental	Control
Pharmacologic treatment	Experimental	1 Alprazolam/exposure	2 Alprazolam/relaxation
	Control	3 Placebo/exposure	4 Placebo/relaxation

reciprocal effect. Effect sizes are moderate to large for alprazolam and very large for exposure according to criteria defined by Cohen (1992). The effect sizes diminish over time for exposure but are still in the large to very large range after six months. Those for alprazolam are absent to small during the follow-up period. Effect size for total major panics are absent to small throughout the study because of the effect in the control group. In general, the findings show a greater effect for exposure and no consistent pattern of interaction between psychological treatment and medication.

The strengths of the study include an appropriate population for study. All subjects met stringent criteria for panic disorder and none had been nonresponders to either of the treatment modalities prior to study entry. Further, assignment to treatment was random, assessment was double-blind to medication and single-blind to psychological treatment, assessment included both self- and assessor-completed measures, and analyses addressed the range of assessments included.

The design reveals one of the problems inherent in studies designed to assess interactions of medication and psychological treatments. The discontinuation design, that was appropriate for the psychological treatment in the present study, is not optimal for assessment of pharmacologic treatments and the fact that pharmacologic effects were not maintained after discontinuation should come as no surprise. Analysis was restricted to 134 subjects who completed at least six weeks of treatment (a 16% dropout rate). Although the rate is relatively low compared to other studies of alprazolam, and there were no differences in baseline characteristics between the 20 subjects who withdrew from the study and the 134 who did not, it is unclear how their inclusion in analyses might have altered the results. The general wisdom regarding randomized clinical trials is unequivocal in stating that all subjects randomized must be included in analysis but a review of a standard journal in the field, *Controlled Clinical Trials*, during the past four years did not reveal a single specific citation.

3.03.6 INDIVIDUAL DIFFERENCES: WHAT TREATMENT WORKS FOR WHOM?

From a clinical perspective, perhaps the most important question regarding treatment is the appropriate matching of patient or client to treatment. The idea that a given treatment approach is appropriate for all (even all who share certain characteristics) runs counter to clinical and research experience. Although randomized experiments may not represent the optimal method for evaluating this question, they can provide useful information. Two rather different methods can be used in the context of randomized trials. The first is the use of factorial designs and the other is the *post hoc* examination of individual characteristics as predictors of treatment response. This section will consider these methods.

3.03.6.1 Individual Difference Variable

The use of an individual difference variable as a factor in a factorial design allows a direct experimental test of a hypothesis regarding differential effects of a patient or client characteristic in relation to a psychological treatment. This strategy is difficult to implement—each client characteristic included doubles the number of subjects needed in a study. Perhaps the most common example of a client characteristic studied in this way is the use of more than one center in a study so that the generalization across centers can be tested. Examples of this strategy are relatively common (e.g., Elkin et al., 1989; Marks et al., 1993), but it is unusual for hypothesized differential effects to be proposed. A second variable that is sometimes investigated in a factorial design is gender.

3.03.6.2 *Post hoc* Examination of Predictors of Treatment Response

In general, *post hoc* analyses of individual client characteristics as differential predictors of treatment response have been disappointing. For example, several of the studies that are reviewed in this chapter have attempted to identify characteristics of clients who fare particularly well in the treatments studied and as has been noted in the descriptions, reliable differences have not been found. Review of the psychopharmacology literature yields a similar conclusion.

Perhaps the most reliable finding regarding prediction of treatment outcome has been severity of symptoms. In a number of studies where symptom severity has been studied, the evidence suggests that it usefully discriminated response (e.g., Elkin et al., 1989). In that study, discrimination of differences between medications and psychotherapies was clearer among subjects with greater symptom severity.

Difficulty in reliable identification of individual characteristics that predict differential response may stem from a number of causes. The nihilistic view is that there simply are no reliable predictors. Two alternatives seem more reasonable. The first is that although studies may have adequate power to detect treatment differences, they do not have adequate power to examine individual characteristics. If this is the case, meta-analytic strategies could allow the examination of data regarding characteristics that are frequently considered such as gender, age, symptom severity, duration of symptoms, referral source, comorbid psychological problems, or diagnoses. See Chapter 15, this volume, for a discussion of this problem and methods to deal with it. The final alternative is that client characteristics which are predictive of treatment outcome are elusive and are not captured adequately in experimental studies. Characteristics such as motivation and client–therapist rapport come immediately to mind.

3.03.7 CONCLUSIONS

Randomized, experimental designs represent a useful method in the evaluation of psychological treatments. They allow unambiguous conclusions regarding the treatments that are studied in the subjects who receive them. This chapter has reviewed in some detail several generally excellent individual studies in order to identify their strengths and weaknesses. The astute reader will have noted that all the studies had both strengths and weaknesses. The goal of drawing attention to these is twofold. The first is to provide readers of the research literature with a framework within which to evaluate other studies. The second is to hope that the observations in this chapter may contribute to the ongoing process of improving the quality of experimental studies of psychological treatments. In this way we can maximize the value of the contributions made by subjects who agree to participate in experiments and improve the treatment and care of clients and patients.

3.03.8 REFERENCES

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. Rev.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Bickel, W. K., Amass, L., Higgins, S. T., Badger, G. J., & Esch, R. A. (1997). Effects of adding behavioral treatment to opioid detoxification with buprenorphine. *Journal of Consulting and Clinical Psychology, 65,* 803–810.

Brill, A. A. (1938). *The basic writings of Sigmund Freud.* New York: Modern Library.

Brock, T. C., Green, M. C., Reich, D. A., & Evans, L. M. (1996). The *Consumer Reports* Study of Psychotherapy: Invalid is Invalid. *American Psychologist, 51,* 1083.

Brown, R. A., & Lewinsohn, P. M. (1984). A psychoeducational approach to the treatment of depression: Comparison of group, individual, and minimal contact procedures. *Journal of Consulting and Clinical Psychology, 52,* 774–783.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 171–246). Chicago: Rand McNally.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cole, J. O. (1968). Peeking through the double blind. In D. H. Efron (Ed.), *Psychopharmacology. A review of progress 1957–1967* (pp. 979–984). Washington, DC: US Government Printing Office.

Consumer Reports (1994). Annual questionnaire.

Consumer Reports (1995, November). Mental health: Does therapy help? *Consumer Reports*, 734–739.

Cook, T. D., & Campbell, D. T. (Eds.) (1979). *Quasi-experimentation: design and analysis issues for field settings.* Boston, MA: Houghton Mifflin.

Department of Health and Human Services, Office of the Secretary (1994). Protection of Human Subjects. Title 45 of the Code of Federal Regulations, Sub-part 46. *OPRR Reports*, Revised June 18, 1991, Reprinted March 15, 1994.

Elkin, I., Parloff, M. B., Hadley, S. W., & Autry, J. H. (1985). NIMH Treatment of Depression Collaborative Research Program. Background and Research Plan. *Archives of General Psychiatry, 42,* 305–316.

Elkin, I., Shea, T., Watkins, J. T., Imber, S. O., Sotsky, S. M., Collins, J. F., Glass, D. R., Pilkonis, P. A., Leber, W. R., Docherty, J. P., Fiester, S. J., & Parloff, M. B. (1989). National Institute of Mental Health treatment of depression collaborative research program. General effectiveness of treatments. *Archives of General Psychiatry, 46,* 971–982.

Flick, S. N. (1988). Managing attrition in clinical research. *Clinical Psychology Review, 8,* 499–515.

Frank, E., Kupfer, D. J., Perel, J. M., Cornes, C., Jarrett, D. B., Mallinger, A. G., Thase, M. E., McEachran, A. B., & Grochocinski, V. J. (1990). Three-year outcomes for maintenance therapies in recurrent depression. *Archives of General Psychiatry, 47,* 1093–1099.

Frank, J. D. (1971). Therapeutic factors in psychotherapy. *American Journal of Psychotherapy, 25,* 350–361.

Hollon, S. D., & DeRubeis, R. J. (1981). Placebo–psychotherapy combinations: Inappropriate representations of psychotherapy in drug-psychotherapy comparative trials. *Psychological Bulletin, 90,* 467–477.

Hollon, S. D., Waskow, I. E., Evans, M, & Lowery, H. A. (1984). System for rating therapies for depression. Read before the annual meeting of the American Psychiatric Association, Los Angeles, May 9, 1984. For copies of the Collaborative Study Psychotherapy Rating Scale and related materials prepared under NIMH contract 278-81-003 (ER), order "System for Rating Psychotherapy Audiotapes" from US Dept of Commerce, National Technical Information Service, Springfield, VA 22161.

Hunt, E. (1996). Errors in Seligman's "The Effectiveness of Psychotherapy: The *Consumer Reports* Study." *American Psychologist, 51*(10), 1082.

Keane, T. M., Fairbank, J. A., Caddell, J. M., & Zimering, R. T. (1989). Implosive (flooding) therapy reduces symptoms of PTSD in Vietnam combat veterans. *Behavior Therapy, 20,* 245–260.

Kendall, P. C., & Lipman, A. J. (1991). Psychological and pharmacological therapy: Methods and modes for comparative outcome research. *Journal of Consulting and Clinical Psychology, 59,* 78–87.

Klerman, G. L., Weissman, M. D., Rounsaville, B. J., and Chevron, E. S. (1984). *Interpersonal psychotherapy of depression.* New York: Basic Books.

Kotkin, M., Daviet, C., & Gurin, J. (1996). The *Consumer Reports* Mental Health Survey. *American Psychologist, 51,* 1080–1082.

Kriegman, D. (1996). The effectiveness of medication: The *Consumer Reports* study. *American Psychologist, 51,* 1086–1088.

Lewinsohn, P. M., Munoz, R. F., Youngren, M. A., & Zeiss, A. M. (1978). *Control your depression.* Englewood Cliffs, NJ: Prentice-Hall.

Marks, I. M. (1978). *Living with fear.* New York: McGraw-Hill.

Marks, I. M., Swinson, R. P., Basoglu, M., Kuch, K., Noshirvani, H., O'Sullivan, G., Lelliott, P. T., Kirby, M., McNamee, G., Sengun, S., & Wickwire, K. (1993). Alprazolam and exposure alone and combined in panic disorder with agoraphobia. A controlled study in London and Toronto. *British Journal of Psychiatry, 162,* 776–787.

Mintz, J., Drake, R. E., & Crits-Christoph, P. (1996). Efficacy and Effectiveness of Psychotherapy: Two Paradigms, One Science. *American Psychologist, 51,* 1084–1085.

Pocock, S. J. (1983). *Clinical trials: a practical approach.* New York: Wiley.

Pocock, S. J., & Simon, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics, 31,* 103–115.

Prien, R. F., & Robinson, D. S. (Eds.) (1994). *Clinical evaluation of psychotropic drugs—principles and guidelines.* New York: Raven Press.

Schooler, N. R. (1978). Antipsychotic drugs and psychological treatment in schizophrenia. In M. A. Lipton, A. DiMascio, & K. F. Killam (Eds.), *Psychopharmacology—a generation of progress* (pp. 1155–1168). New York: Raven Press.

Seligman, M. E. P. (1995). The effectiveness of psychotherapy: The *Consumer Reports* study. *American Psychologist, 50,* 965–974.

Seligman, M. E. P. (1996). Science as an ally of practice. *American Psychologist, 51,* 1072–1079.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32,* 752–760.

Uhlenhuth, E. H., Lipman, R. S., & Covi, L. (1969). Combined pharmacotherapy and psychotherapy: Controlled studies. *Journal of Nervous and Mental Diseases, 148*, 52–64.

Waskow, I. E., & Parloff, M. B. (1975). *Psychotherapy change measures.* Washington, DC: National Institute of Mental Health, US Government Printing Office.

White, S. J., & Freedman, L. S. (1978). Allocation of patients to treatment groups in a controlled clinical study. *British Journal of Cancer, 37,* 849–857.

Wolpe, J., & Lazarus, A. (1966). *Behaviour therapy techniques.* Oxford, UK: Pergamon.

World Health Organization (1992). *International classification of diseases (ICD-10)* (10th ed.). Geneva, Switzerland: Author.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

3.04

Multiple Group Comparisons: Quasi-experimental Designs

HANS-CHRISTIAN WALDMANN and FRANZ PETERMANN
Universität Bremen, Germany

3.04.1 INTRODUCTION 63
3.04.1.1 The Relevance of Experimentation in Clinical Psychology 63
3.04.1.2 Terms: "True," "Quasi-," and "Non-" Experimental Studies 64
3.04.1.3 The Concept of an "Effect" 64
3.04.2 THE LOGIC OF DESIGN IN THE COURSE OF EMPIRICAL RESEARCH 65
3.04.3 CRITERIA: SENSITIVITY, VALIDITY, AND CAUSALITY 67
3.04.3.1 Sensitivity 68
3.04.3.2 Validity 68
3.04.3.3 Causality, Practice, and Pragmatism 71
3.04.4 CONTROL IN QUASI-EXPERIMENTAL DESIGN 72
3.04.5 A PRIMER SYSTEM OF DESIGN 74
3.04.5.1 Selections from the General Scheme: Nonequivalent Group Designs 74
3.04.5.2 Using a Pretest 75
3.04.5.2.1 Using pretest as a reference for gain scores 75
3.04.5.2.2 Using pretests as approximations to initial group equivalence 76
3.04.5.2.3 When not to use a pretest 76
3.04.5.2.4 Outcome patterns 77
3.04.5.3 Using Multiple Observations 77
3.04.5.4 Using Different Forms of Treatment 78
3.04.5.5 Using Different Forms of Comparison Groups 79
3.04.5.6 Using Combinations of Different Designs 80
3.04.5.7 Regression Discontinuity 80
3.04.6 ANALYSIS 81
3.04.7 CONCLUSION 87
3.04.8 REFERENCES 88

3.04.1 INTRODUCTION

3.04.1.1 The Relevance of Experimentation in Clinical Psychology

Contributions of clinical psychology to health services have received widespread recognition and the benefits are undisputed. The parallel increase of costs in such services due to growing utilization of psychological aids, however, must be justified on grounds of the scientific method. There is a double need for valid and reliable demonstration of the value in technology derived from and offered by clinical psychology: efficacy in serving the customer and efficiency legitimizing it to the supplier must be shown. Quasi-experimental multiple group comparison designs are well-suited for this task: they account for the fact that randomization may often not be an adequate sampling strategy in real-life clinical research situations and still allow for assessment of absolute and relative efficacy of psychological interventions. They help evaluate the merits of different components in interventions and support decision making as regards implementation, monitoring, and optimization of respective programs. This practical, if not wholly pragmatic, research and reasoning contributes directly to either correction or refinement of the present state of scientific knowledge and development in the domain of clinical psychology. It is practice where results from clinical psychology are to be evaluated, and it will be seen that multiple group comparisons offer adequate tools to do so.

3.04.1.2 Terms: "True," "Quasi-," and "Non-" Experimental Studies

In this chapter, multiple group comparisons will be focused on by means of quasi-experimental designs. Delimiting characteristics to distinguish the three kinds of empirical research situations quoted above are subject assignment, modus of treatment, and control. It is agreed widely to speak of true experiments when both random subject assignment and creation of treatment or, in a broader sense, active manipulation of conditions occur. In a quasi-experimental setup the first condition is weakened, and in observational studies both conditions are completely released. Some authors refer to the term quasi-experimental as to the constraint that the researcher cannot deliberately assign subjects to conditions or that occurrence of a treatment is only a typical but not necessary feature (e.g., Spector, 1981). In this case, the differentiation to pure observational studies seems obsolete. It is argued that any experimental situation, be it "true" or "quasi," involves an arbitrary discernible kind of treatment, by creation or manipulation of conditions regardless of mode of subject assignment, and that quasi-experiments feature all elements of true experiments except randomization.

If there is no manipulation of subjects or conditions in order to evoke variation in dependent measures (and thus contrasts for testing), reference is made to "pure" observational studies (see Chapters 4 and 13, this volume). The main problem then lies with selecting a statistical model and further in devising an appropriate parametrization that reflects as closely as possible the structure of scientific hypotheses (in their original, verbal form). The objective of such inquiries may comprehend the identification of latent traits and their structural composition (factor analysis), the identification of the dimensional metrics according to which subjects determine the relative value of different objects (multidimensional scaling/conjoint measurement), or in the identification of a time series regression model in order to predict value and stability of a criterion. In neither case is a treatment strictly required.

3.04.1.3 The Concept of an "Effect"

Multiple groups comparison experiments are conducted in order to demonstrate relative effects of treatment. Effect hypotheses refer to presence, size, and variability of effects with respect to certain criterion measurements. The major objective of multiple group comparison designs lies with the generation of data in order to test such hypotheses. Subjects or groups methods representing different levels of independent variables or being subjected to different kinds of treatment, eventually measured at different points of time, are compared on data obtained in the same way in each group and at each occurrence. Then a classical concept of an "effect" is identified with the interaction of the subject or group factor and time/treatment with respect to the outcome measure. For example, the idea of a case–control study to determine therapy effectiveness might be pointed out as "if you apply X (treatment, independent variable) in one group, while not in another, and you observe a significant difference in Y (some functional status, dependent varable) between both groups, then you have reason to believe that there is a relationship between X and Y, given all else being equal." It will be seen that it is the *ceteris paribus* condition in the given clause, and how it is tentatively realized, that makes (quasi-)experimentation a delicate challenge to researchers' creativity and expertise.

A worked example of a quasi-experiment is chosen to introduce into the present chapter an illustration of some of these challenges and ways to counter them.

Worked example: Outline

In a very well designed study, Yung and Kepner (1996) compared absolute and relative efficacy of both muscular and cognitive relaxation procedures on blood pressure in borderline hypertensives employing a multiple-treatment–multiple-control extension of design Number 6 in Figure 2. It is reported that most clinical applications make use of an amalgam of cognitive techniques like suggestion, sensational focusing, and strict behavioral training of muscle stretching relaxation. The authors aim at partialing out the effects of these various components by clearly separating instructions that target muscles directly and such instructions that try to mediate muscle tension by mental efforts. As a consequence, special attention is given to operational validity and treatment integrity. In order to counter confounding of subject characteristics and procedure variables the authors devise a rigorous subject selection and assignment policy which makes their work an excellent example for the case here.

Before engaging in describing various designs and methods of data analysis, and in evaluating this particular experiment, it seems instructive to outline a general process of experimental research.

Generally, it will not be possible to observe all possible instances of variable relations (which would innecessitate inference), but evidence for the hypotheses will be sought by investigating subsets of the population(s) in which these relations are considered to hold true. Sampling variability and measurement error lead to the use of statistical models in order to determine the genuity of an observed effect. Effect sizes (ES) serves to model the above "difference" in terms of statistical distributions (which, in turn, provide both the mathematical background and the factual means for testing effect size). The significance of statistical significance and the concept of reason based hereupon are evaluated in Chapter 14, this volume, and the issue of interpreting such a relationship as causal is briefly discussed in Section 3.04.3. But how does one arrive at effect hypotheses and corresponding statistical models in the first place?

3.04.2 THE LOGIC OF DESIGN IN THE COURSE OF EMPIRICAL RESEARCH

It is suggested that experimentation as a procedure is a special realization of a general process of empirical research, as depicted in Figure 1. Whether experimentation is seen as an appropriate access to address the problems in question depends on preceding decisions; if so, choice of design and sampling structure is conceived as an embedded level. Furthermore, it is proposed that this procedural outline is paralleled by a conceptual process of hypotheses derivation. Figure 1 illustrates this correspondence with hypotheses levels according to the deductive four-step model presented by Hager (1992):

(i) Generally agreeing with the hypothetico-deductive method, scientific hypotheses are derived from proposition structures in theory. Sometimes, mere need for evaluation or striking observations from practice must serve as a starting point for the research process.

(ii) Relative to marginal conditions (resources, time, etc.) and constraints in construct operationalization, a subject matter model is derived in terms of variables and their interrelation. Temporal ordering and other prerequisites of model theory guaranteed, predictions regarding empirical manifestations are derivable ("should then be higher than"). Devise a design now in order to translate a hypothesized structure of variables into a measurable one and a sampling strategy in order to actually perform the measurement. Hypotheses stating that certain strategies of verbal intervention in psychotherapy are of use for depressives might be translated into a "better-off" proposition as quoted in the text above. It is clear that, at this level, an experiment rationale is suggested: comparison of groups. Also, global aspects of testing are implied: assessment strategy (will psychometric tests be applied or will therapists rate patients' depressiveness by virtue of their expertise?), or inference conditions (will an effect be required to instantiate simultaneously on all criteria and to be stable for at least a year, thus calling for multiple post-tests, or will the whole study be replicated using different subjects and other therapists?). In transition to the next level, a set of decisions to be made is encountered that are known to statistical consultants as "how many" questions in project and design planning (how many subjects are enough per group in order to balance sensibly factors of sensitivity?, how many time intervals make up the optimal retest interval?, how many outliers can be tolerated in and across groups?, etc.).

(iii) Multiple group comparison as a conceptual strategy implies suppositions about structures in the data. Here, conceptually identified effect sizes receive an accurate, quantitative, and thereby testable formulation. A "pre–postdifference" between groups may now be specified in terms of an expectation of a difference in central tendency of an outcome variable, a kind of second-order comprehensive measure that can be tested easily for significance at the next level. More generally, statistical predictions formally denote compound suppositions about the "behavior" of data, and a "good" design generates data behavior that by maximum degree either confirms or contradicts behavior that would be expected by hypothesis.

(iv) Statistical hypotheses are statements about specific values, variability, and ordinal or metric relations of parameters of distributions which follow from statistical predictions. A common prototype for such a measure is

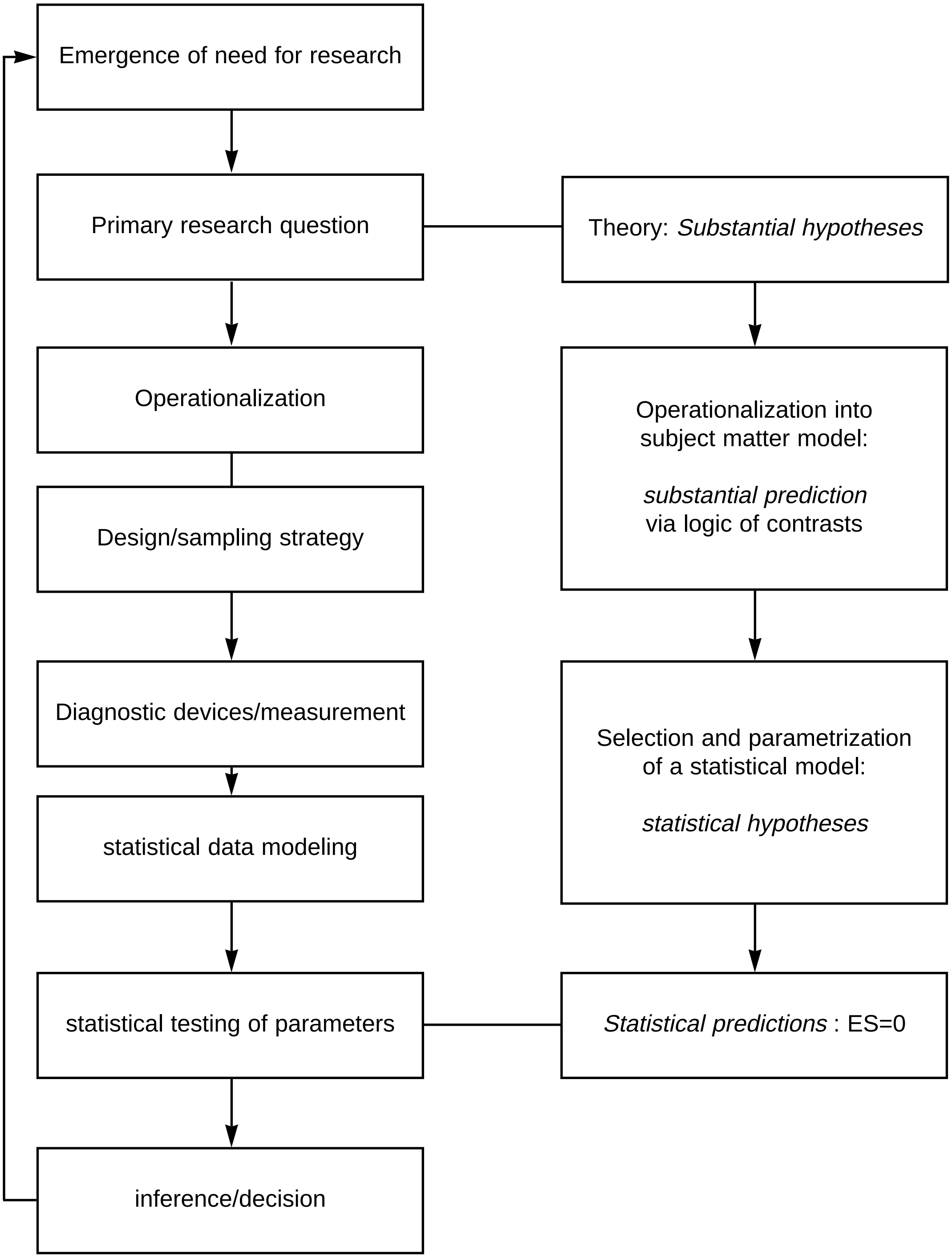
Emergence of need for research
Primary research question
Theory: *Substantial hypotheses*
Operationalization
Operationalization into subject matter model:
substantial prediction via logic of contrasts
Design/sampling strategy
Diagnostic devices/measurement
Selection and parametrization of a statistical model:
statistical hypotheses
statistical data modeling
statistical testing of parameters
Statistical predictions : ES=0
inference/decision

Figure 1 The process of empirical research.

given by $ES = (\mu_{\text{treat}} - \mu_{\text{control}})/\sigma_{\text{control}}$. Its problems not withstanding, this is considered a fairly reasonable and intuitive basic concept. Various other ES formulations can be translated to it (Cohen, 1988; Lipsey, 1990). For instance, in the example of therapy effectiveness, and effect is assumed to be present when statistical testing has determined that its ES differs from zero significantly, that is, beyond the expected limits of sampling error and with prespecified error probabilities of the α- and β-kind. In fact the probability is always estimated that there is no effect (therefore speaking of "null" hypotheses) or that, as it is frequently put in terms of sampling theory, comparison groups are samples that were drawn from the same population. In this case, observed differences merely reflect sampling error. Depending on the formulation of predictions, it is not always obvious at first sight whether these "null" probabilities are required to be high or low. Further complicating things, a statistical prediction must sometimes be further differentiated into a set of single-testable statistical hypotheses. It must then be decided whether an overall null (or alternative) hypothesis is conforming to the compound statistical prediction or what pattern of acceptance or rejection of individual hypotheses integrates to a favorable finding. Moreover, there is a need to decide whether integration of separate findings should follow a disjunctive (compensatory) or conjunctive rule (e.g., must depressives benefit on all scales applied in order to justifiably speak of an effect). As a consequence of this multiple testing, there is a need to control for accumulation of error probabilities. Note, however, that "difference" does not necessarily imply the comparison of means. It most cases, pretest-standardized differences in means are in fact used for this purpose because such measures have known statistical properties (e.g., distributions and standard errors) following normal theory and can thus be easily tested statistically. But a test might be made for a ratio to be greater than a prespecified number (as is done frequently when using differential model fit across groups as a criterion) or the inequality of regression intercept and slope parameters (as featured in regression continuity designs, see Section 3.04.5.7). The general rationale is to determine whether or not groups differ in any statistical parameter and to relate this difference to a reference sampling error estimate. (For details on procedures of significance testing see Cohen, 1988; Fisher, 1959; Giere, 1972; Mayo 1985.) Criteria like adequacy (on each level, hypotheses preserve to maximum degree classificatory, ordinal, or metric relations assumed on preceding level), sufficiency (a linearly predictable trend may not be tested for monotony only), decisiveness and unambiguous identification of theory-conforming empirical states (Hager, 1992) serve to estimate validity of hypothesis derivation. After statistical testing, however, there is a need to work backwards: what effect size is considered practically significant? Is there a failure in predicting things as they should occur in theory? Do the findings entitle inductive inferences to be drawn to higher levels in the hierarchy of hypotheses, leading to tentative confirmation of propositions in the theoretical frame of reference? Here have been introduced some of the most serious problems in philosophy of science, and there is a need to get still further into it before turning to techniques.

3.04.3 CRITERIA: SENSITIVITY, VALIDITY, AND CAUSALITY

Conventionally, the major objective of experimentation is seen in demonstrating a causal relationship between phenomena, or, as Cook and Campbell (1979) put it in more correct terms, in facilitating causal inference. The logic underlying experimental design is an implicit variation of the concepts of covariation, temporal precedence, and of ruling out alternatives. This means, in terms of experimentation, that the hypothesized "cause" should be capable of being omitted and control group esigns are preferable. One-shot case studies or pretest–post-test designs without at least one control or, in a broader sense, a comparison group, do by no degree allow for the kind of inference scientists are most interested in: causal ones. But, as was seen from Figure 1, experimentation ends up in statistics. The result of an experiment is that data are communicated in the form of statistical propositions. Whether statistical relationships should be taken as indicators only or be conceived as "causal," "functional," or "probabilistic" cannot be determined by simply rating how well the rules of experimentation were obeyed, but depends on affiliations in philosophy of science. Criteria for when to apply the predicate causal and be justified in doing so, are hardly available. But before drawing any such inferences it will be necessary to (i) ensure sensitivity in order to detect the effect and (ii) rely on it being a valid indicator for what is to be interpreted as causal. In the following section, a start is made on practice by introducing the key concepts whose consideration will hopefully establish a sensitive, valid, and meaningful design. It is clear, however, that an evaluation of meaning must rely on protoscientific argumentation.

3.04.3.1 Sensitivity

Sensitivity refers to the likelihood to detect an effect in sample-based investigations, given it is indeed present in the population from which samples were drawn. It is determined by sample size, total and direction of effect size, heterogenity of both subjects and treatment applications, and various undesired features of diagnostic devices (unreliability, etc.). While those factors affect experimental precision, others like prescribed error probabilities of both α and β type and the kind of statistical test used affect the associated concept on the level of statistical analysis: power. It must be borne in mind that sampling error stands for the variance of parameter estimates and is, thus, maximally dependent on sample size. As a consequence, an effect size however small will become significant in statistical terms with increasing sample size, while being wholly uninformative for practical considerations. For an ES of the kind presented above less than or equal to 0.3 to attain statistical significance at $\alpha = 0.05$, some $n = 300$ subjects would be in order; for an ES of 0.5 still a hundred or more. In a meta-analysis on psychotherapy efficacy Grawe (1992) found that 87% of 897 studies were using samples of $n = 30$ or less subjects in intervention groups. To attain statistical significance at the usual level an effect size of greater than 1.20 would then be required. This may be considered quite an unlikely prerequisite: in their meta-analysis on psychotherapy research Smith, Glass, and Miller (1980) reported an overall expectation of treatment effectiveness of no more than ES = 0.85. Tentative counterbalancing the determinants of effect sizes other than the "true" effect (population effect), namely sample size, error probability, and assumed ES, is best achieved using power tables as proposed in Cohen (1988). Since detecting an effect usually means observing an effect size that statistically differs from zero, the levels of design and data analysis in Figure 1 become mixed up when evaluating the whole procedure of quasi-experimental multiple group comparisons. It is the correspondence of design structure "giving" the data and statistical data model that establishes a "good" design. Sometimes data obtained from one design can be translated into different models (problem of parametrization) and effect sizes may be evaluated for their significance by various of statistical tests (problem of relative power).

3.04.3.2 Validity

Validity refers to the likelihood that the effect detected is in fact the effect of interest. Campbell and Stanley (1963), and Cook and Campbell (1979), originators of the concept, distinguish internal, external, and statistical validity as well as construct validity. The latter refers to operationalization problems (mono-operationalization bias, monomethod bias, evaluation apprehension, experimenter expectancies, etc.). Internal validity evaluates to which extent variation in measured dependent variables is attributable solely to variation of independent variables or treatment, respectively. Again, assuming an effect for a difference in the value of some statistical parameter calculated from data on dependent variables (e.g., the mean), internal validity assures that nothing but treatment "caused" this difference, and that, as Lipsey (1990) puts it, the observed difference parallels the fact that "the only difference between the treatment and the control conditions is the intervention of interest" (p. 11). This, of course, relies heavily on the experimenter's ability to control for other possible "sources of variation" or undesired confounding influences. External validity refers to the extent to which results obtained from a sample-based study may be generalized to a specified population from which samples were drawn in the first place and/or across slightly different (sub)populations. Such extrapolation must be reasonably justified as regards target population (i.e., persons), settings, and future time. Note that representativeness is a prerequisite to external validity that not only holds for subject selection but also for situative conditions of experimentation itself: ecological validity must be given careful consideration especially in clinical fields.

Campbell and Stanley (1963) have raised the argument that internal validity can be thought of as a necessary while not sufficient condition for external validity. Note, however, that some methods of control depicted in the next section show conflicting effects on internal and external validity, and to compensate for these imbalances with a reasonable trade-off certainly is an art of its own. Most flaws in experimental and quasi-experimental design boil down to violations of factors of validity which is why a short outline is presented (see Table 1) of their respective factors as designated by Cook and Campbell (1979). A far more comprehensive listing of possible threats to validity and various other shortcomings in planning experimental research is presented in Hoole (1978). Special attention to external validity in evaluation contexts is given in Bernstein (1976).

Admittedly, full consideration of all these caveats, if not mere apprehension of this rather lengthy list, seems to discourage any use of (quasi-)experimental designs at all. The contrary is true, though. There will be a benefit from

Table 1 Factors of internal and external validity.

Internal validity	
History	In the time interval between pretest and post-test measure, various influences or sources of variation that are not of interest or that may even distort “true” findings may be in effect besides the applied treatment. This clearly leads to confounding. Also, these factors may be differently effective for different subjects, thereby introducing additional error variance. Controlling for such intrusion by elimination or by holding it constant to some degree is an appropriate method in laboratory settings but is almost impossible to achieve in real-world clinical (field) settings (apart from identifying such factors as relevant in the first place). Moreover, many disorders are known to be subject to seasonal variation or others periodicities.
Maturation	In longitudinal studies with fairly large restest intervals, effects may be confounded with or entirely due to “natural” change in subjects’ characteristics that are correlated with the variables and treatments of interest. A prominent example in clinical research is spontaneous remission in psychotherapy studies. The hope is that such change takes effect equivalently in both groups of a case–control study (which again is rather a “large sample” argument but is usually held true for randomized samples).
Mortality	In repeated measures designs, attrition or loss of a certain proportion of subjects is to be expected on a regular basis. Differential loss, however, means that the observed drop-out pattern across groups seems not to be random but is dependent on variables that might correlate with treatment or outcome measure. This in turn introduces *post hoc* selection artifacts and is likely to erroneously contribute to effect size in that it may change value and variability of post-test statistics.
Statistical regression	A typical problem with measuring change lies with regression towards the mean. Subjects may shown gain or loss in scores from pretest to posttest solely due to the unreliability of the test. This may be compensated for to a certain degree by multivariate assessment of dependent variables. Moreover, differential effects are likely to fall into place depending on subjects” pretest or starting-off: “viewed more generally, statistical regression (a) operates to increase obtained pretest–post-test scores among low pretest scores, since this group’s pretest scores are more likely to have been depressed by error, (b) operates to increase obtained change scores among persons with high pretest scores since their pretest scores are likely to have been inflated by error, and (c) does not affect change scores among scorers at the center of the pretest distribution since the group is likely to contain as many units whose pretest scores are inflated by error as units whose pretest scores are deflated by it.”[a]
Testing, instrument reactivity, and sensitization	In repeated measures designs subjects are likely to become familiar with diagnostic devices or may carry over pretest scoring behavior. In addition, there are many ways that testing procedure itself may affect what is tested. Items that evoke strong emotional reactions may distort scoring in subsequent ability subtests by increasing subjects’ unspecific arousal. Moreover, pretests may sensitize subjects for treatment reception by enhancing attention for certain attributes or by facilitating self-communication and introspection. Sometimes it is possible to control for this effect in quasi-experimental settings by applying strategies similar to the Solomon four-group design (see Chapter 3, this volume).
Selection/ interactions with selection	Selection is a threat to internal validity when an effect size—a difference in some parameter across groups—reflects pre-experimental differences between cases and controls rather than differential impact of treatment. Note that inference from such experimentation relies heavily on the *ceteris paribus* condition (denoted in the given clause in the outline in Section 3.04.1.3), and that systematic differences across groups due to selection bias are most pervasive in quasi-experimentation where we have to do without randomization. This is especially true in case of clinical treatment: one would not want to allot inhouse patients or program participants to different treatments (or exclude them to serve as controls) by random numbers but according to their individual needs and aptitude. Patients assigned for cases on behalf of equal (that is, equivocal) diagnoses may yet still differ from other patients in their group (error variance) as well as from controls in a whole pattern, of other variables that are likely to be related to both reception of treatment and outcome measures (e.g., hospitalization, socioeconomic status, verbal fluency, etc.). Note that most factors of internal validity listed so far may interact with subject selection, producing differential effects and thereby increasing experimental error.

Table 1 (continued)

Instrumentation	In the first place it must be ensured, of course, that diagnostic devices meet criteria of reliability, objectivity, and validity in terms of measurement and testing theory. In repeated measures designs, discrimination and differentiation capabilities of tests may change. Thus a gain in scores from one occasion to another may partly be due to enhanced performance of observers who improved skills with practice, and less due to differential effects of applied treatment at these occasions. A related problem lies with so-called floor or ceiling effects. Subjects scoring at the upper end of a scale on one occasion, say 6 on a 1–7 Likert scale, may easily report a decrease by half but will find it difficult to state a major increase of similar proportion. Note that clinical samples are often defined by a shift in mean in certain indicators with respect to a "normal" population.
Control group behavior	Besides the problem of recruiting and maintaining a control group, "compensatory rivalry" of respondents receiving less desirable treatment and "resentful demoralization" of such subjects are further threats to internal validity. In the first case, being assigned to the nonreceiving group may motivate subjects to behave competetively as to reduce the "true" effect (due to treatment) in counter-hypothesized direction. In the second case, the deprived get more deprived and less motivated: a patient that suffers from his disease is assigned for a control and thus kept from a therapy that is thought to be efficient by the experimenter. The "lose-heart" effect would then artificially enlarge effects. Both variants are of great practical importance in therapy effectiveness research.[a]
External validity	
Interaction of pretest and treatment, reactivity	It is well-known that mere knowledge to participate in a study or the awareness of being observed does affect behavior. Furthermore, due to sensitization effects or conditioning effects (see above), results obtained from samples that have completed a pretest cannot be fully generalized to populations that have not done so.
Interaction of selection and treatment	People that deliberately demand for psychological care or for access to health care programs are likely to be more susceptible to treatment or motivated than are accidental samples of convenience (collective of inpatients, rooms, etc.). Often, selection processes that have led to admission of subjects are unknown. It is clear, however, that results obtained from such samples cannot be generalized unreservedly.
Interfering treatments	Parallel, overlapping or interacting treatments lead to confounding and thus constitute a major threat to internal validity. But there are undesirable consequences for external validity as well. Suppose patients are receiving medical treatment in ascending dosage. Due to idiosyncratic effects of cumulation, conclusions drawn with respect to an intermediate level of dosage that generalize across people may be erroneous.
Nonrepresentative samples	Generally, target populations must be clearly specified in advance. In quasi-experimental design, representativeness cannot be ensured by random sampling on a large scale but instead take advantage of "natural" (nonmanipulative) sample formation. Often characteristics of research context naturally differentiate subjects into groups (rooming policies of hospital, record analysis, etc.). Strictly speaking, since there is no real sampling there is no point in making inferences. But this aside, sampling should not refer to subjects only: generalizing also includes care givers (therapists, etc.), hospitals, and regions. Therefore, another strategy lies with deliberate sampling for heterogeneity with respect for certain attributes relevant to the research question. Inference would then be across a wide range of units holding subsets of these attributes, but still limited to the respective superset. When dealing with clinical samples, this method seems fairly reasonable.

Source: Cook & Campbell (1979).[a]

it in that the degree to which attention is given to these issues prior to experimentation parallels the degree to which results can be interpreted and inferences drawn after experimentation. Factors of validity thus serve as criteria for comparative evaluation of quasi-experimental studies, they are not meant as a checklist for building the perfect one (remember that several factors exclude mutually). Another precaution must be added. While the bearings of internal validity on technical issues in research planning are widely agreed upon and factors listed above have nearly become codified into a guide on design construction, most important issues of construct validity often receive less recognition. But re-examining the procedural structure of empirical research, as sketched in Figure 1, the view that emerges is that everything begins and ends with valid operationalization of constructs into measurable variables, and of hypothesized relations into exercisable methods for "visualizing" them. Greater appreciation of construct and derivation validity is advocated before turning too readily to tactics and techniques of experimentation. For instance, one should examine treatment integrity well before constructing sophisticated application schedules in an interrupted time series design (see Section 3.04.5.3). Nonetheless, within the entire validity framework, causal inference is predicated on maximum appreciation mostly of internal validity. The question is what is meant by causal.

3.04.3.3 Causality, Practice, and Pragmatism

However sensitive and valid a design might be by these terms, an effect or statistical relationship cannot be interpreted as truly causal in an objectivist ontological sense or in the traditional Humean readings (for an introduction, see Cook & Campbell, 1979, and Salmon, 1966). Therefore, and for the following reasons, a wholly alternative view is pledged on these issues in the tradition of James and other pragmatist reviewers of methodology. Scientific inquiry and its results (e.g., explanations) should always bear technological relevance. As one consequence, the importance of manipulation, is re-emphasized as a feature of treatment in quasi-experimentation. Consider the following argument raised by Collingwood (1940):

> Suppose one claimed to have discovered cause of cancer, but added that his discovery would be of no practical use because the cause he had discovered was not a thing that could be produced or prevented at will. His dicovery would be denounced a shame. (p. 300)

Collingwood's point is that causal explanations are often valued for the leads they give about factors that can be manipulated. If the essential cause of some effect does not imply controlling the effect (e.g., cancer) through manipulating some factor, then knowledge of this cause is not considered useful by many persons (Cook & Campbell, 1979). Scientific inquiry and quasi-experimental research designs are conceived as concretization devices, as tools to obtain technological knowledge. Thus, research designs need to take into account the feasibility of manipulation of conditions that produce effects in order to at least allow for transfer into practice. It is practice where results from research in clinical psychology should be evaluated, which is why a concept of causality and explanation is needed that is of relevance. Modern philosophers and methodologists have demonstrated successfully that pragmatist reasoning can be adapted successfully to methodology in social sciences, and have done so down to the level of hypotheses testing (Putnam, 1974). Of course there is no pragmatic explanation of its own right. But our attempts to explain phenomena on experimental or even statistical grounds can be justified pragmatically by introducing criteria beyond a "logic of discovery":

> The first and ineliminable criterion for the adequacy of an explanation must be based on what it does for a man who wants explanation. This depends on contextual factors that are not reflected in the forms of propositions or the structure of inferences. (Collins, 1966, p. 140)

> Knowledge of causal manipulanda, even the tentative, partial and probabilistic knowledge of which we are capable, can help improve social life. What policy makers and individual citizens seem to want from science are recipes that can be followed and that usually lead to the desired positive effects, even if understanding the micromediational is only partial and the positive effects are not *invariably* brought about. (Cook & Campbell, 1979, p. 28)

If efficacy of, say, a new drug has been demonstrated on a statistical basis by experimentation, it is not possible, strictly speaking, to conclude that deficiency in effective substance of the drug caused the disease or that the substance now reversed the pathogenetic process to regeneration. Having controlled for side effects, it is perfectly justifiable to maintain the drug as a successful therapy (and raise funds for further elaboration which might give new insights for science as well as for refinement of the drug). On a statistical level this "better-off" thinking in interpreting results from case–control studies translates into the concept of statistical relevance (Salmon, 1966) or into the notion of Granger causality. The latter derives from the theory of linear systems and defines causal relevance of independent variables ("causes") in terms of increased predictability of value and variability of dependent variables ("effects"). Prognosis should then be more precise with a relevant variable included in the model than with the variable removed, and good prognosis is definitely a very useful outcome of research (Granger, 1969). Good prognosis is, after all, a feature of operational models. This is what should be expected from experimentation in clinical psychology: identification of factors that influence etiology and the spread of disorders and derivation of operational models that in turn might aid to confront them.

3.04.4 CONTROL IN QUASI-EXPERIMENTAL DESIGN

Control aims at enhancing and balancing factors of internal and external validity following the max-con-min-rule formulated by Kerlinger (1973):

(i) *maxi*mize impact of independent variables on dependent ones,

(ii) while holding *con*stant factors of systematic intrusion, and

(iii) *mini*mizing the influence of unsystematic variables or error.

What is meant by factors, influences, and error? Control is always exerted over independent variables that contribute to effects. It is instructive to refine the term variable into four types relevant in experimentation. Systematic variance denotes differences between groups and, with reference to the definition of an effect size in Section 3.04.1.3, thus indicates the presence of "true" effects. In concept, it is solely attributed to explanatory variables, indicating treatment. Extraneous variables, however, are not a primary focus of a study but in cases of substantial correlation with dependent ones introduce error variance and do bias estimates of effects. Such variables may either be subjected to control or be left uncontrolled in one of two ways, depending on both sample size and subject assignment policy. In the latter case, randomized variables are let run freely in the hope that their effects will cancel out in the long run and thus equate groups in a probabilistic sense while confounded variables constitute a threat to validity that cannot be ruled out. It follows that controlled variables are extraneous but their impact on measures of effect is accounted for because error variance is reduced. If such erroneous influences are known to operate (by background knowledge introduced into the subject matter model, or simply by intuition), the techniques presented in Table 2 are used or at least an attempt is made to obtain accurate measures on them. Statistical techniques may, then, be used in order to separate *post hoc* the effect of treatment (in terms of ES) and selection differences.

In quasi-experimentation, control is exercised rather by means of subject selection or statistical analysis than by manipulating conditions. Because no randomization is used in quasi-experimental settings to wash out all initial differences in and across groups and because an effect has been defined in terms of post-test differences, credibility of findings will depend on the ability to demonstrate that groups have been alike as possible except for the difference in treatment they received. It will be noticed, however, that pretest equivalence does not lead, as might be implied by name, to equivalent groups. There may still be reliable and substantial differences on other factors relevant but left unconsidered (or even unmeasureable if known) that affect the phenomenon in question. The argument by Cook and Campbell (1979) is followed that only understanding the process of subject selection allows full understanding of subject differences that cannot be attributed to treatment, and that randomization must be considered the only model of such a process that can be completely understood.

The following is a worked example:—Control is by subject selection and assignment:

> Of 307 recruited individuals, only 40 were eligible for participation. Persons showing habituation effects were excluded, because their high pressure in multiple baseline readings was assumed to be a reaction to the measurement situation itself (sensitization). To control for treatment interference persons currently engaged in other pressure control techniques were also excluded. Persons on medication were asked to present medical records to assure constant dosage for the time of their participation (history). The authors cite evidence for a linear relationship between blood pressure level prior to behavioral intervention and treatment change. They further attribute contradictory results in recent research on relaxation efficacy to the common practice of matching experimental groups with respect to the average on certain blood pressure parameters, which leaves differential treatment susceptibility in higher degrees of hypertension uncontrolled. Subjects were therefore orthogonally matched on systolic blood pressure, age, and sex to assure pretest equivalence.

Experimental error also leads to bias in effect size estimates. While disturbing marginal influences (disruptions due to periodicities in hospitals" daily routines, noise, etc.) may be held approximately constant or be eliminated thereby having their influence equated in and across groups, other undeliberate variations of experimental conditions (e.g., treatment inflation or interferences) and measurement effects are to a far lesser extent under the control of the experimenter. But in many field settings there may be no desire to exert artificial control over such (in theory) irregular components of overall effect. It may be preferrable to preserve or enhance external or ecological validity when finally turning to generalization. Since practice is defined as primary, however, bias due to naturally varying concomittants is something to be faced. There is no need to put research back into the laboratory by excellent control when findings need to be transferred into applications for a noncompliant world.

The worked example is continued to demonstrate trial structure and procedural control:

Table 2 Controlling subject-induced extraneous variables.

Matching	The rationale of matching is to fully equate groups with respect to pretest scores, so that differential post-test scoring could legitimately be interpreted as differential change. The more variables are called in for matching the greater a shrinkage in sample size must be expected, sometimes deflating degrees of freedom to an extent where statistical analysis is no longer possible. Moreover, such samples may no longer be representative for target populations to which one finally intends to generalize findings. By the way, there is great appeal in the idea to match subjects with themselves across treatment levels as is the case, for example, in crossover designs.
Parallel groups or aggregate matching	This means of control is, in a way, a weaker version of matching by taking resort to group statistics rather than to individual values. The influence of confounds or, more generally speaking, nuisance factors is considered not relevant if it is distributed evenly across groups. If such a variable has equal distribution in each group (with respect to first- and second-order moments, e.g., mean and variance), groups are said to be parallel on this variable. When dealing with intact groups or when assignment is guided by need or merit, this strategy would entail exclusion of subjects in one group in order to approximate the other or vice versa. Again, sample size would be reduced and groups may finally differ on other variables by selective exclusion. Control always seems to trade off with external validity, notably generalizability. Parallel groups, however, are much easier to obtain than samples matched by score.
Blocking	Blocking further draws variance out of the error term when nuissance factor(s) are known and measurable on a categorical level. The rationale is to match subjects with equal score on this measure into as many groups as there are levels of the nuisance factor. If, for example, $n_j = 10$ subjects are to be assigned to each of $J = 3$ treatment conditions (factor A), and another factor B is known to contribute to variance, one would obtain $K = 10$ levels of factor B with $n_k = 3$ subjects each and then randomly assign $n_{jk} = 1$ subjects per B-level to each of the J levels in A. An increase in numbers of blocks will further decrease error variance since scores within a block of subjects matched this way are less likely to vary due to selection differences to the same extent than would have to be expected in blocks comprising more or all (no blocking) subjects. Note that in this model blocks represent a randomly sampled factor in terms of mixed-model ANOVA[a] and thus can enter statistical analysis to have their effect as well as possible interactions with treatment tested for. As a result, blocking could be viewed as transition of matching (above) into factorization (below). On the other hand, matching is but an extreme case of blocking where any pair of matched subjects can be considered a separate block.
Factorization	Confounding can be completely removed from analysis when confounding variables are leveled (measured by assignment to categories or by recording their continuous scale into discrete divisions) and introduced into design as a systematic source of variation. This generally results in factorial designs where effects of factors can be quantified and tested (main effect and interaction in fully crossed designs, main effect only in hierarchical or incomplete structures).[a,b,c]
Analysis of covariance	While factorization is a pre-experimental device of control (it gives, after all, a sampling structure), analysis of covariance serves to *post hoc* correct results for concommittant variables or pretest differences.[d]
Holding constant elimination	If hospitalization must be expected to substantially mediate results, a most intuitive remedy lies with selecting subjects with the same inhouse time. Strictly speaking, there is no way of generalizing to other levels of the hospitalization variable.

Source: Hays (1990).[a] Pedhazur & Pedhazur (1991).[b] Spector (1981).[c] Cook & Campbell (1979).[d]

All subjects received treatment, placebo, or pure assessment in equal time spacing in eight experimental sessions. They were instructed to replicate these sessions for 30 days after the laboratory trial and to keep a record of their practice. After this follow-up period, all subjects were finally assessed on their blood pressure. Though this procedure implicated a loss in standardization, it certainly enhanced ecological validity and increased the clinical relevance of the study's findings. Light and temperature were held constant in all measurement situations, all equipment was visually shielded to prevent feedback effects. Measurement equipment was calibrated before each trial. All participants spent equal time in the measurement setting and were blinded with respect to assessment results until completion of the measurement occasion. To control for experimenter effects subjects received live relaxation instructions in three sessions followed by five sessions with taped instructions. This procedure further improved standardization of treatment situations. A training time of 18–20 minutes was justified as optimal by recurring to recent findings that shorter training lacks any effects and that longer training often results in loss of motivation to practice, increased dropout rates, and reduced stability of effects over time.

3.04.5 A PRIMER SYSTEM OF DESIGN

As will be obvious from Figure 2, basic forms of designs mostly differ in whether or not they use randomization to assure initial group equivalence and control groups to determine effects, and whether or not they dispose of pretest, multiple observations, multiple forms of treatment (applied, removed, reversed, faded, etc.) and different controls. Following on from this, the regression discontinuity approach is presented as leading to a class of designs in its own right. In the case of more than one explanatory independent variable, additional decisions must be made regarding logical and temporal order of treatment application (factorization into fully crossed and balanced plans, incomplete and hierachical plans, balancing sequence effects, etc.). These topics are more fully covered in Chapter 4, this volume. Note that many important practical considerations and implications for choice of a statistical model for subsequent analysis are hidden completely in what is called the "data aquisition box" (or "O" in standard notation, see below). Here it must be decided on (i) number of dependent variables (uni- vs. multivariate analysis), (ii) various parameters of repeated measurement (number of occasions, lengths of retest intervals, etc.), (iii) order of dependent measures and error modeling (e.g., use of latent variable models), and (iv) scale characteristics of dependents (parametric vs. arbitrary distribution models).

Most authors, follow standard ROX terminology and notion introduced by Campbell and Stanley (1963): R stands for randomization, O stands for an observation or measurement occassion, X stands for treatment or intervention. Temporal sequence is expressed by left-to-right notation in a row, separate rows refer to different groups.

3.04.5.1 Selections from the General Scheme: Nonequivalent Group Designs

The designs shown in Figure 2, while not a focus of this chapter, will now be considered. Designs one and four are considered "pre-experimental" by Campbell and Stanley (1963):

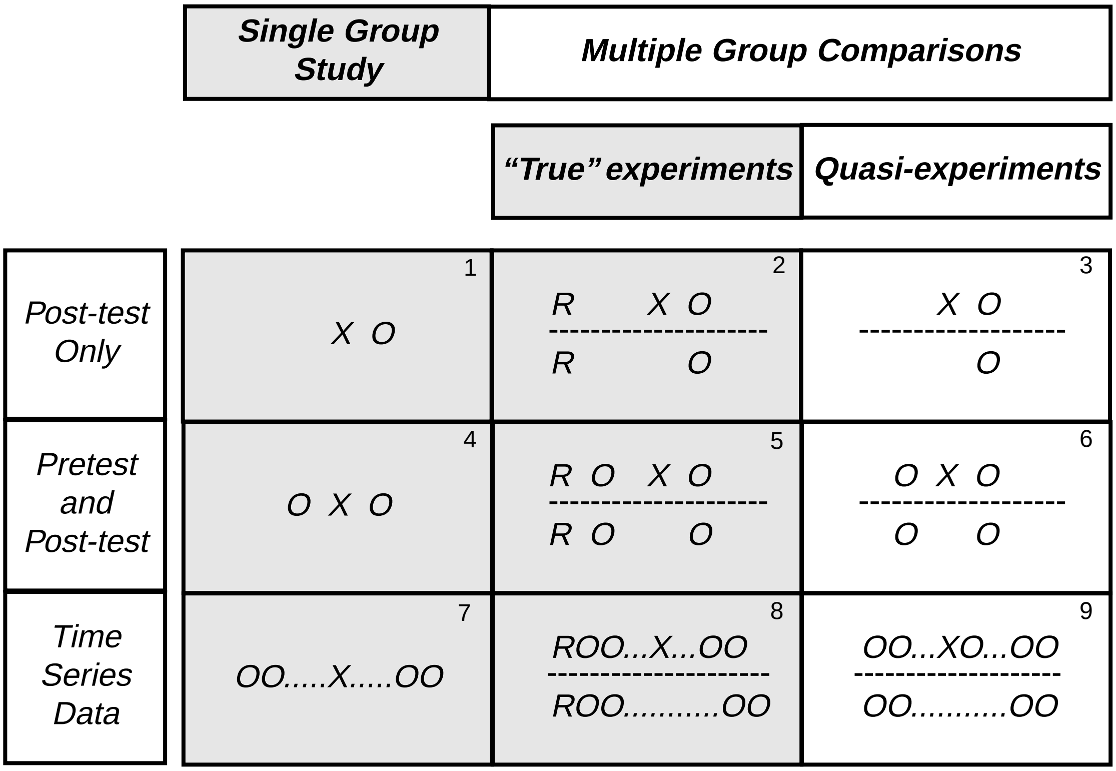

	Single Group Study	Multiple Group Comparisons	
		"True" experiments	Quasi-experiments
Post-test Only	1 X O	2 R X O -------------------- R O	3 X O -------------------- O
Pretest and Post-test	4 O X O	5 R O X O -------------------- R O O	6 O X O -------------------- O O
Time Series Data	7 OO.....X.....OO	8 ROO...X...OO -------------------- ROO..........OO	9 OO...XO...OO -------------------- OO..........OO

Figure 2 A basic system of designs.

findings from such studies are most ambiguous since almost any threat to internal validity listed above may be an effect and left uncontrolled. As a consequence, and because experimentation is meant to demonstrate effects by assessing comparative change due to treatment while ruling out alternative explanations, these designs are commonly held inconclusive of effects. They are not recommended for the purposes sketched in Sections 3.04.1 and 3.04.3. Design three does implement an independent control group, but still remains inconclusive as regards the genuity of effects due to selection differences. In design four, subjects serve as their own control. With two measures only, little can be inferred about true effects since maturation and other time-relevant threats to validity remain uncontrolled. Extending the principle to design seven, however, leads to a typical time series design for basic impact analysis that to some extent allows for more control. Suppose a trend line is fitted to both consecutive pretests and post-tests: a discernible shift in intercept and/or slope parallel to treatment application would indicate presence of an experimental effect (a conceptually different, but technically similar concept will be discussed with the regression discontinuity approach). Without determination of such a "natural" trend a higher post-test level might reflect what should be expected from extrapolation of regularities in pretest values only, disregarding treatment. In other words, dependent variables might show a change in value that is due to natural change over time (inflation, deflation, or periodicities of the phenomenon under study). Using multiple observations does not rule out this possibility as would using multiple observations and control groups, but still make it highly unlikely for such a trend to go undiscovered. Some authors argue that lack of a control group may be of less importance for this reason. Delay in effect instantiation or special forms of temporal calibration of effects can be modeled by univariate intervention (or interrupted time series) analyses (see McDowall, McCleary, Meidinger, & Hay, 1980). Note, however, that group designs are still being dealt with exclusively where $N > 1$. Single-subject designs, even when presented in the same notation, employ another rationale of contrasting and inference and are covered in Section 3.04.2. Multiple-group extensions to this design are covered in Section 3.04.5.3. Designs two, five, and eight have the most desirable properties: they implement randomization, repeated measurements, and control groups, with multiple observations additionally available in the last one. These designs are discussed in Chapter 4, this volume. From the preceding, it is clear that the rightmost column contains three nonequivalent group designs which do implement multiple group comparisons. Design three, however, is subject to the same reservations made for designs one and four. It lacks both randomization and pretesting as an approximation to the same end. Therefore designs six and nine, when used with appropriate care and caution, permit the testing of causal hypotheses, with causality being conceived as in Section 3.04.3.

The worked example is continued with aspects of design

> In quasi-experimental design, the treatments to be compared serve as independent variables and experimental groups are built by assigning subjects to treatment conditions by nonrandom procedures. In the worked example study by Yung and Kepner (1996), the treatment factor comprised stretch release (SR), tense release (TR), and cognitive relaxation (COG) as well as a placebo attention group and a test-only control group (TOC) that remained wholly untreated. The placebo group was given a medicament by a physician with the ruse that it would reduce blood pressure in order to control for positive treatment expectancy effects. The control group were given pure assessments of their blood pressure to control for effects of the measurement process itself. Systolic and diastolic blood pressure as well as heart rate were chosen as dependent variables, entailing a multivariate setup for subsequent statistical analysis. All subjects were measured repeatedly several occasions (multiple baseline readings, with assessments before and after trial, follow-ups), which gives a multiple-treatment–multiple-control extension of design six in Figure 2.

Note that the study outlined above does in fact realize most of those items that make up the following sectional structure: conduct of pretests, multiple observations, different forms of treatment, different forms of control groups, and, introduced *post hoc* by means of statistical analysis, even combination of different designs.

3.04.5.2 Using a Pretest

3.04.5.2.1 Using pretest as a reference for gain scores

The major step from design three to design six clearly lies with the introduction of pretesting, seemingly indicating analysis of change. A natural concept of an ES would derive by subtracting pretest averages from post-test means and normalize the result with respect to a control group for a comparison. There has been extensive work on the question whether or not to use difference scores as an estimate of ES. Some main issues addressed in this context comprise the fact that difference scores (i) carry the unreliability of both their constituents, (ii) depend on correlation with initial (pretest) scores, (iii) depend on scale characteristics, and (iv) are subject to regression towards the mean. Various alternative techniques have been proposed among which are (i) standardizing prior to calculating differences, (ii) establishing "true" difference scores (Lord, 1963), (iii) using regression residuals (Cronbach & Furby, 1970) and, (iv) most commonly used, regression adjustment by means of analysis of covariance. Further, in terms of mathematical models, specification of correlation structure for error components of total scores, decomposition of error itself into various time-depending and static components, as well as the formulation of model-fit criteria and appropriate statistical testing, with serial dependence of data contradicting the "usual" (IIND) assumptions on error terms, are of major concern when analyzing change. These items denote not mere problems of statistical theory, but reflect substantial problems about a concept of empirical change. There is, however, no concern about all the complications that arise when trying to measure change but with the logic of design. For an introduction and further reading on the above issues, see Eye (1990), Diggle, Liang, and Zeger (1994) and for special topics as indicated Kepner and Robinson (1988) and Thompson (1991) (ordinal measures), Hagenaars (1990) and Lindsey (1993) (categorical data), Christensen (1991) and Toutenburg (1994) (generalized linear models), Vonesh (1992) (nonlinear models), Farnum and Stanton (1989) and Box, Jenkins, and Reinsel (1994) (forecasting), Petermann (1996) and Plewis (1985) (measurement).

3.04.5.2.2 Using pretests as approximations to initial group equivalence

In Section 3.04.2 the concept was introduced of an ES as the observable and testable formulation of what is sought in experimentation: effects. As can be seen from Figure 2, repeated measures are not strictly required to demonstrate effects. As is the case in design three differences may be observed in dependent measures across independent samples that underwent two different treatments (or none for a control) and state that one group scored higher on the average (this being favorable). Drawing any inferences or preferring this treatment over the other (or none) as a result of such a post-test-only design would rest entirely on the assumption that groups were equivalent with respect to dependent variables prior to treatment application. In "true" experiments, random subject assignment and/or large samples may be relied upon to justify this assumption, but in quasi-experimental design performing a pretest is strictly required to obtain meaningful results. One still does not need to take gain scores or measures of different growing rates as a base for an effect size. In Section 3.04.2, they were not. But post-test differences can only be interpreted as monitoring effects when initial group equivalence is given as a common ground for comparison. As a consequence, subject assignment in quasi-experimental design is based upon techniques that require a pretest (see Table 2).

3.04.5.2.3 When not to use a pretest

Research situations can be imagined where conducting a pretest may not be recommended for even stronger reasons. Suppose that the phenomenon in question is likely to change qualitatively rather than quantitatively. A pretest would then presumably first measure something different than post-test. Second, a pretest may interact with treatment by differential sensitization (see factors of internal validity) or by producing such bias that results could not sensibly be attributed to treatment anymore (e.g., when using attitude tests). Third, retesting may be impossible because testing can only be set up once on the relevant variable because measurement must be assumed to distort or annul the phenomenon, or because no parallel form of a test is available. In such cases, the use of a proxy variable might be considered which, measuring something different but conceptually related to post-test criteria and, thus, (hopefully) being correlated to it in the statistical sense, serves as a substitute for evaluating initial group equivalence as a prerequisite for interpreting post-test differences (Rao & Miller, 1971). In extreme cases, when any kind of pretesting, be it on behalf of proxies or using the post-test instrument, seems impossible, using separate samples "from the same population" for pre- and post-test may be in order. From the preceding it is clear, however, that this variant of the basic design (six) provides a very weak basis for inference. Selection cohort designs as proposed in Cook and Campbell (1979) can be understood as a conceptually improved (while more demanding) extension of separate sample designs. When using self-selected samples it is obvious that subjects collected from cohorts that fall into place with regular periodicity because of characteristics of natural environment can be considered more alike than separate samples with obscure nonrandom selection processes. As this is most typically the case for institutions with cyclical (e.g., annual) turnover, like schools, such designs have been labeled "recurrent institutional cycle designs" (Campbell & Stanley, 1963).

3.04.5.2.4 Outcome patterns

It is trivial to state that informational complexity of outcome pattern depends on the factorial complexity of design. It should be borne in mind, however, that sophisticated group and treatment structures do not necessarily entail higher levels of information or yield better causal models of a post-test. For example, the higher an order of interaction is allowed in fully crossed factorial designs, the less readable will be the results obtained and interpretation will be difficult. Pedhazur and Pedhazur (1991) give examples of how adding predictors to regression equations alters coefficients of former included predictors and changes overall results. Though the foregoing statements about validity and control may suggest a kind of "the more the better" strategy when designating independent and dependent variables for a design, the ultimate purpose of all modeling, simplification, should be recalled. The ultimate prerequisite in modeling, however, is of equal importance to the same end: unambiguous specification. Design six, for instance, when including two pretest occasions and being carried out on near-perfectly matched groups of sufficient size with near-constant treatment integrity, clearly outperforms a time series extension of the same design (see below) with uncontrolled subject attrition and carry-over effects.

A naturally expected, and in most cases desired, outcome of basic design six would obtain when the average post-test response of the treatment group is found to differ significantly in desired (hypothesized) direction from the control group post-test scoring, given pretest equivalence and given zero or nonsignificant trend in the control group pretest–posttest mean scores. But there are several variations of that scheme, largely due to selection–maturation interaction, regression artifacts, and scale characteristics (see Table 1). Cook and Campbell (1979) discuss at length such different outcomes of design six in its basic form (parallel trend lines, differential growing rates for groups, nonconstant controls, controls outperforming cases, treatment leading to post-test equivalence, trendline crossover ["switching means"]). Testers should not, however, as might be suggested by these authors and the foregoing presentation of all those threats to validity, too readily attribute findings not in line with expectations to design misspecification, artifact operation, and limitations of diagnostic devices. It is true that there is always our obligation to evaluate and justify the whole procedure of research instead of its findings only. But having revised the process of hypotheses derivation that leads to design, and to data as a result, and having confirmed derivation validity to a maximum extent attainable within the limits offset by a concrete research context, there should be no complaints about reluctant and unruly reality.

3.04.5.3 Using Multiple Observations

It is obvious that design nine is but a generalization of standard design six, taking advantage of multiple observations pointed out above, and being subjected to the flaws of multiple repeated measurements. While the extension is straightforward as regards the logic of contrasts (the rationale for multiple group comparisons), the notion of an effect and according ES is more difficult to define. Time series analysis is usually carried out to detect and analyze patterns of change over time or to build models for forecasting. Multiple group comparisons enter when chosen to define an effect in terms of different time series evolution for cases and controls, conditional on a preintervention phase (baseline) and triggered by differential onset, intensity, and duration of treatment. Two major problems must be considered, then: first, how should treatment and observation phases be scheduled in the course of the whole study in order to set contrasts for obtaining effects while accounting for time-relevant threats to validity (see Section 3.04.5.4 using multiple forms of treatment); second, how to statistically analyze serially dependent data and interpret contrasts in the data in terms of within-subject (within-group) vs. between-subject (between groups) patterns (see Section 3.04.6).

The major benefit of using multiple observation has been presented in Section 3.04.5.1 (control for maturation and regression). Adding a control group further permits control of history and instrumentation for there is no *a priori* reason to suppose that untreated subjects experienced different environmental conditions during the study or that they should react differently to repeated measures. While nonspuriousness of treatment effects justifiably may be assumed by reference to significantly different pretreatment levels or increased/decreased slope of baseline, temporal persistence of effects can never be assured without comparison to an untreated control group.

As it is intended to dedicate the section on analysis mostly to handling design six and its nontime-series (i.e., simple pretest–post-test) extensions, a few words on statistical analysis of time series experiments are in order here. For more detail see Glass, Wilson, and Gottman (1975), Gottman (1981), and Lindsey (1993). Comparisons of intercepts (level), slopes (trend) and variance (stability, predictability) are mostly used for contrasting pre–post phases. Note that effects in time series designs need not instantiate immediately with treatment application and pertain. Impact of treatment may be delayed, with gradual or one-step onset, and various forms of nonconstancy show up over time (linear, nonlinear, or cyclic fading or gain, etc.). But such rather sophisticated parameters of time series data will only be used rarely for ES definition: statistical tests for directly comparing oscilation parameters or analytically derived characteristics like assymptotic behavior (e.g., speed of point convergence) are not readily available. While most one-sample time series parameters of interest to researchers concerned with intervention analyses can be identified, estimated, and tested for by transfer function models, between-sample comparisons often reduce to fitting an optimal model to one group and observing that fit indices decrease when obtained from application of the same model in another group. Cross-series statistics are, after all, hard to interpret in many cases (they require consideration of lead-lag relations, etc.). For an introduction to interrupted time series analysis for intervention designs see McDowall, et al., (1980) or Schmitz, (1989), and illustrated extensions to nonlinear and system models are presented in Gregson (1983).

3.04.5.4 Using Different Forms of Treatment

In this section concerns about implementing different forms of the same treatment instead of adding treatments like in a factorial design will be discussed. In basic pretest–post-test designs of type six, treatment is applied only once. Hence, variations concern application or removement only. Graduation in treatment application or withdrawal calls for multiple groups to undergo such graduation. This would be achieved, in concept, by stacking k–1 pretest–post-test designs without a control group over standard design six for k levels of treatment to be tested. The resulting model would conform to standard one-way analysis of variance for differences scores with a fixed or randomly sampled treatment factor (see Section 3.04 .6). In time series designs, however, some more challenging variations are at hand that further increase validity and clarity of inference, if properly implemented. Figure 3 gives an overview.

Design 10 replicates design four, with treatment removed in the second block in order for this block to serve as a "same-subjects" substitute for an otherwise nonavailable control group. Careful consideration must be given to the following questions:

(i) Are treatment effects of transient nature and be made to fade out (it is supposed, generally, that persistence of intervention effects is desired)? Ethical concerns, financial questions and potential attrition effects could be added to hesitations about mere feasibility of such designs.

(ii) Is it reasonable to expect that average scoring will return to a near-baseline level after removal of treatment?

(iii) Will it be expected that, after reintroduction of treatment as in design 12, effects will reinstall the same way (delay, size, stability, etc.) as in the first turn? (It is supposed that the reader is acquainted with general replication effects in experimenting like learning and resentment.) Then, and especially in case of high-amplitude treatment and obstrusive measurement, construct validity will be seriously affected.

If answers are favorable to these concerns then the designs that are particularly useful in evaluation of recurrent intervention components can be disposed of. Next should be considered the reversed treatment nonequivalent control design 11 with pretest and post-test. Cook and Campbell (1979) suggest that this design outperforms standard design six as regards construct validity since a treatment variable has to be specified rather rigorously in order to determine its logical reverse and an operational mode for it to be applied. Nevertheless the additional use of a wholly untreated control group is strongly advised in order to arrive at clear-cut results, in case trend lines do in fact depart from (assumed equivalent) pretest levels in opposite directions but in slopes that differ with respect to size and significance. Including a common reference for contradicting provides a very strong basis for inference. Finally design 13 (interrupted time series with switching replications) is considered. Due to delayed treatment application, one group can always serve as a control during treatment periods of the other ("reflexive controls," Rossi, Freeman, & Wright, 1979; "taking-turn controls," Fitz-Gibbon & Morris, 1987). The particular strength of this design lies with its capability to demonstrate simultaneously effects in two different settings and points of time with greatest parsimony. Still it is possible to overcome most threats to validity that are faced when using the above designs.

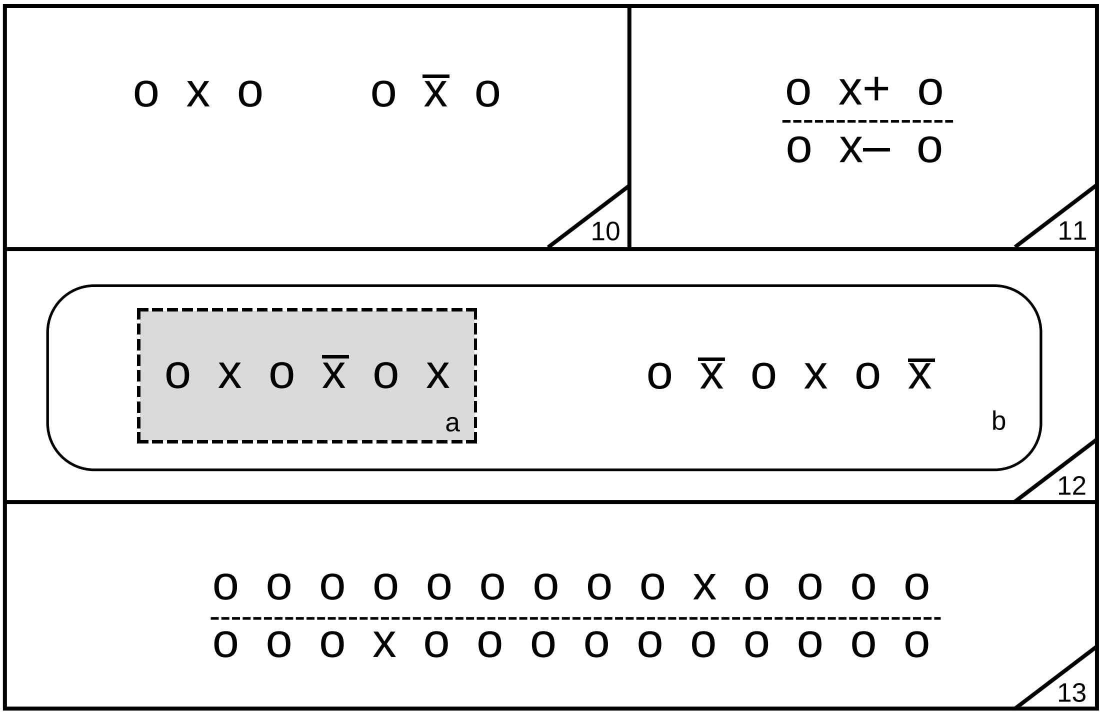

Figure 3 Designs using different forms of treatment.

3.04.5.5 Using Different Forms of Comparison Groups

It has been pointed out frequently that to use a control group is indispensible as a baseline for comparison in most cases of quasi-experimental research, and highly advocated in any other. Effects should be demonstrated by either relative change or post-test difference to rule out alternative explanations besides treatment impact. While this is widely agreed, there has been some controversy as regards various qualitative characteristics of controls. Are controls really untreated as assumed for inference? How can control groups be maintained (they do not receive, after all, any benefits of treatment)? Is there such a thing as placebo treatment? Do waiting groups (subjects scheduled for subsequent cycle of treatment, see design 13) differ in any substantial respect to wholly unaffiliated pools of control? In many cases the nature of a control group is determined simply by contextual factors like availability in terms of total number of subjects, financial resources, ethical considerations, or the time horizon of the study. While "real" control groups receive no treatment at all, placebo groups receive only irrelevant treatment. Conceptually, there are two forms of such irrelevance. In the first case, intervention neither in theory nor in statistical terms relates to the treatment variable of interest and is only meant to minimize motivation effects like resentful demoralization. A second version refers to treatments that in fact can be, or even are, expected to contribute to effects by operating through subjects' expectancies and external attributions that are hypothesized to go along with actual treatment. Perhaps the most prominent example is medication with a sugar pill. Note that such placebos work on psychological grounds and usually are implemented to separate such effects from "real," physical, or pharmacological treatment. As a consequence, it might be difficult to define a psychological placebo in psychotherapy research. For further detail on the issue of placebo groups, see a well-received article of Prioleau, Murdock, and Brody (1983), defending admission of such placebo treatment, and subsequent peer commentaries in *The Behavioral and Brain Sciences*. Waiting control groups differ from untreated groups in that they must be supposed to expect benefits from treatment to come. As a consequence, after assignment, such subjects may show more sensible behavior in any respects related to health and well-being or more medical compliance (thus "raising" baselines into trendlines). Waiting groups are, however, easiest to recruit and maintain, they naturally fall into place when treatment cannot be delivered unlimitedly, and there is no way of avoiding selection for admission. Still another concept of a control group emerges when using qualitatively different treatments in comparison groups, that is distributing treatment across groups. A way of treating a disease might, for instance, to compare against current or standard treatment, thus highlighting the relative merits of the new and saliently different features of the innovative treatment. Technically, within the ANOVA framework introduced below, treatment would represent a fixed factor with more than two levels (treat1, treat2, etc). It is clear that adding a further untreated control group is still preferable (see Section 3.04.5.6).

3.04.5.6 Using Combinations of Different Designs

There are three convincing arguments for combining designs. First, combining complete designs rather than extending one of their internal features (as discussed in above subsections) will generally enhance control. Second, broadly defining designs themselves as methods and using different methods and research plans in addressing the same question and phenomenon will always decrease monomethod bias and increase both construct validity and generalizability of results. Third, combining designs offers refined testing of hypotheses. For example, in Section 3.04.5.4 it was proposed, to multiply stack design four over design six. Consider now the application of different dosages of a drug as treatment variable. Given initially comparable groups (in the present, quasi-experimental sense), trend hypotheses about dosage could even be tested. Also, integral features of basic designs as depicted in Figure 2 may be tested for their mediating effects. For an example, if design six is stacked over design three a composite design is obtained that allows for separation of pretest and treatment effects and is named after its proposer Solomon (1949). Note that there are two fully crossed factors: whether treatment is applied or not (case–control factor), and whether pretesting is administered or not. This design may now be conceived as factorial (see Chapter 3, this volume) and an analysis of the main effect of treatment made (disregarding effects of pretest sensitization and the like), main effects of pretesting (effect obtained without regard to treatment), and interaction (treatment effects depend on pretest). If statistical testing has determined that such interaction is present, consideration of the main effects due to one sole factor (e.g., treatment) is sure to yield misleading results and inferences. But the latter is always true in cases where any design is used that both lacks randomization and pretest. As a consequence, this rather sophisticated design is mostly applied in research situations that allow for randomization. Trivial as it may seem, another way of combining designs lies with replicating the whole basic design using the same samples and measurement occasions by simply obtaining multivariate measures of posttest indicators and, thus, testing multiple predictions concerning the same outcome criterion.

3.04.5.7 Regression Discontinuity

As if to provoke further discussion in order to put an end to presumed underutilization of regression discontinuity (RD) designs, Trochim (1984) stated that RD "is inherently counterintuitive . . . not easy to implement . . . statistical analysis of the regression discontinuity design is not trivial . . . [there are] few good instance of the use" (pp. 46–47). So what is the point in these designs? In structure and objective, RD designs conform to the layout of design six. The major difference lies with assignment policy: while randomization is believed to guarantee fully equivalent groups (in the long run, that is) and pretest equivalence in nonequivalent group designs assures comparability with respect to the variables in questions only (leaving obscure presence and influence of other selection processes), it is maximum nonequivalence that serves as a rationale of regression discontinuity designs. This is achieved by ordering subjects according to their scores on the continuum of a pretest scale and subsequently defining a cut-off point for division into further subsets to be assigned to different conditions. Most commonly, subjects scoring below the cutting point would then be assigned to one group, subjects scoring below this point to another. If a "sharp" cut-off point is not desirable on theoretical grounds or because of known unreliability regions of scales, a midscale cut-off interval might be used instead, with interval width depending on estimates of measurement error similar standard deviation. Subjects within this range would then be assigned randomly to either treatment or control group condition. While it is true that the process of subject selection is still only imperfectly controlled because merely one factor (the dependent variable) is accounted for, it is perfectly known since being specified entirely by the researcher. It is clear that, in order to prevent cut-off point selection being rather arbitrary, strong theoretical grounds or consent on practical concerns (limited therapy resources, etc.) are called for.

Basically, analysis of RD designs means testing hypotheses of equality of regression equations across groups similar to analysis of covariance. The objective is to ask whether the trend or regression line obtained from analysis of the treatment group pretest and post-test scores is displaced significantly from the respective control group regression line. If displacement is due largely to greater (or lesser) intercept, while slopes are statistically equal (parallel line regression), a main efect has been successfully demonstrated. Stated differently, with no effect present, a single regression line would equally fit scores of both treatment and comparison group and, as a consequence, a trend in the treatment group could be predicted from the control group scores. Interaction effects are revealed by discontinuity in slope of regression lines at cut-off point. Note that, technically, regression need not necessarily be linear but can include parameters of any order, but that extrapolation of a trend line becomes extremely difficult then and often leads to false conclusions about effects (see Pedhazur & Pedhazur, 1991). These problems notwithstanding, more frequent use of these designs is recommended. In many intervention research contexts, subjects not only are in fact but also should be assigned to either treatment or control groups by need or similar, perhaps medical, criteria related to both treatment and outcome.

3.04.6 ANALYSIS

Having successfully derived and implemented a multiple group comparison design there is now a need to fit a statistical model to the data obtained in order to test the hypotheses on the "final" deductive level in the process model of empirical research (Figure 1). In particular, there is an interest in deriving ESs and whether they are different from zero. Note that from an epistemiological point of view, use of a different statistical model is made than observational methods (see Section 3.04.1.2). Here, models are considered as mere tools for hypotheses qualification and inference and are not of contentional value besides some implications of their structural assumptions (like additivity, linearity, etc.). Analysis of nonequivalent group designs is achieved in three major steps. First, the overall rationale of analysis is introduced by borrowing from the theory of generalized linear models. Second, its specification language is used to acquaint the reader with a system of models that indeed is a superset to nearly all models applicable in this context. Third, a closer look is taken at a subset of models that are most commonly employed in treating standard design six, namely analysis of variance with or without adjusting for a covariate and its respective conceptual analogies when ordinal or categorical data have been obtained. Since there are many textbooks on statistical data analysis that cover technique in detail, attention is drawn merely to the principle and major formulae and dispensed with (references are given, however).

It is important to understand that until now discussion has been located entirely in the realm of logic. Nothing has been implied as regards scale and complexity of measurement and data. The point is that anything can be stored in O (Figures 2 and 3) and that structural complexity of whatever resides in this data aquisition box is logically independent of complexity in design structure. Enlarging designs by stacking subelements simply adds data boxes and obtains lots of data and can sometimes lead to technical difficulties in statistics, but both of these features should not complicate the idea or inference behind the material. In fact, there is only one condition: statistical parameters obtained from the data must be comparable on formal (i.e., mathematical) grounds (distributions, standard error of estimate, etc.). Now dominance of comparison-of-means-style models is readily explained: differences in means of normally distributed variables again follow normal distribution, and departure from zero is most easily tested using direct probabilities. In essence, value and standard error of any parameter's estimate should be determined and related to each other (e.g., in a ratio, giving standard t-test in case of metric dependents) or to another pair of parameters (e.g., in a double ratio, giving standard F-test, respectively). But there are many other ways besides of comparing distributions, some less strict but more straightforward or creative.

Why not obtain entire time series for one single data box in a multiply repeated measurements design? A time series model may be fitted to data at each occassion resulting in microstructural analysis. Stated differently: time series information has been condensed into some quantities that now enter macrostructural analysis reflecting design logic. Here, one might compare such aggregate measures in a pre–postfashion analysis like in standard design six, thus evaluating treatment impact by relative change in parameters of local change. Other studies involve entire structural equations systems in a single data box, constructed from covariance structure in multivariate data obtained at one occassion (Bentler, 1995; Duncan, 1975; Möbus & Bäumer, 1986; Jöreskog & Sörbom, 1993).

Most textbooks on the statistical analysis of experiments focus on analysis of variance and associates. While preference for models assuming variables on an interval or rational scale is perfectly understandable from the researchers point of view, it is felt that there is a rationale of testing that allows for far more variety. Appreciation of this variety is badly needed in order to be able to choose an analysis model for the design present, thus preserving the rather deductive idea of the process model in Figure 1, instead of being forced to the contrary: adapting design and idea to some analysis model readily at hand from common textbooks. Variety need not imply incomprehensible diversity if there is a model rationale. Here, the rationale states: all analysis, excepting some tests for ordinal data working on special ranking methods, boils down to regression. This holds true even in cases where it might not be expected that a regression structure presents at first glance (joint space analysis in multidimensional scaling, etc.). Most important, it is also true for analysis of variance and companions since these models can be written as regression equations of metric dependent variables on a set of coded vectors. These dummy variables, then, represent treatment in all its aspects (intensity, frequency, schedule of application, etc.). Encoding logic of designs into what is now called a design or hypotheses matrix for statistical analysis certainly is an art of its own (for an introduction, see Kerlinger (1973), or Pedhazur & Pedhazur (1991)). But here, regression stands for a whole family of models obeying a certain relational form. Thus, the rationale translates to concrete statistics by proposing that all variable relations (and contrasts are just a special relation) be specified in the form $y = f(x, \beta, e)$, where y denotes a possibly vector-valued set of responses (dependents) and x denotes a set of explanatory variables (independents, respectively) that are used to approximate the response with determinable residual or stochastic error e. A set of coefficients β quantify the intensity of an element of x (which can be interpreted as effect information), these parameters are estimated conditional to the data by minimizing a quality function ("least squares," "$(-2)*$maximum likelihood"). Finally, $f(.)$ denotes a characteristic link function that relates the explanatory or predictive determistic part of the model and error component to the response. Mostly used for this purpose are identity link (giving standard multiple regression), logit link (for binary response), and logarithmic links (leading to regression models involving polytoneous categorical data). In this way of thinking, the data are thought of as determined by the model but jammed by additional random error. Note that "error" does not denote anything false but is meant to incorporate, by definition, all factors in effect that remained unspecified or did not explicitly enter the model (i.e., remain uncontrolled).

To sum up, ingredients of a model are variables, parameters, a link function to algebraically interconnect all these, and an estimation strategy to optimally fit the resulting model to the data. The point in mapping most different ideas and objectives into a common-models structure is variable designation: using intensities can still mean single-attribute utilities, attitude strength, blood pressure. But $y = f(x, \beta, e)$ is too much in the abstract to be of use for this purpose, and even giving an outline on mathematical formulation and derivation of some less hard-to-handle submodels or estimation techniques would be beyond the limits of this chapter. For an advanced reading and a complete build-up of generalized linear model theory and practice see Arminger, Clogg, and Sobel (1995), Christensen (1991), and Diggle et al. (1994). But model specification terms summarized in Figure 4 can be used to figure out the greater proportions of the generalized linear model. Note that not everything goes along with anything else when crossing items on the utmost right side and that some models use parametrization strategies that do not satisfy the condition of formal comparability of parameters across groups. What is required now is careful examination of the nature of the variables and the justifiability of assumed form of relations among them. Never follow the impression possibly implied that all it takes is simply picking up a model from the scheme and applying it to the data if the marginal parameters are met. While it is true that with the exceptions mentioned above almost any statistical model of interest for analysis of designs can be integrated in or extrapolated from the generalized linear models scheme, the fit of design and analysis model is never guaranteed from this alone.

Several submodels can be located from that scheme that are most commonly used in analyzing experiments of type three and six (in Figure 2) and that may now be arranged in a more readable tabular system. If groups or discrete time are assumed for a single independent variable of categorical nature, and an arbitrary scale for a single dependent manifests, variables measured in discrete time and everything else in Figure 4 is dispensed with, the following table results (Table 3). In view of practice, that is, software, some submodels present not as rigorously special cases of the Generalized Linear Model (though they all can be specified accordingly), but rather as "standalone" implementations of tests. They serve the same purpose and are integrated in all major statistical software packages.

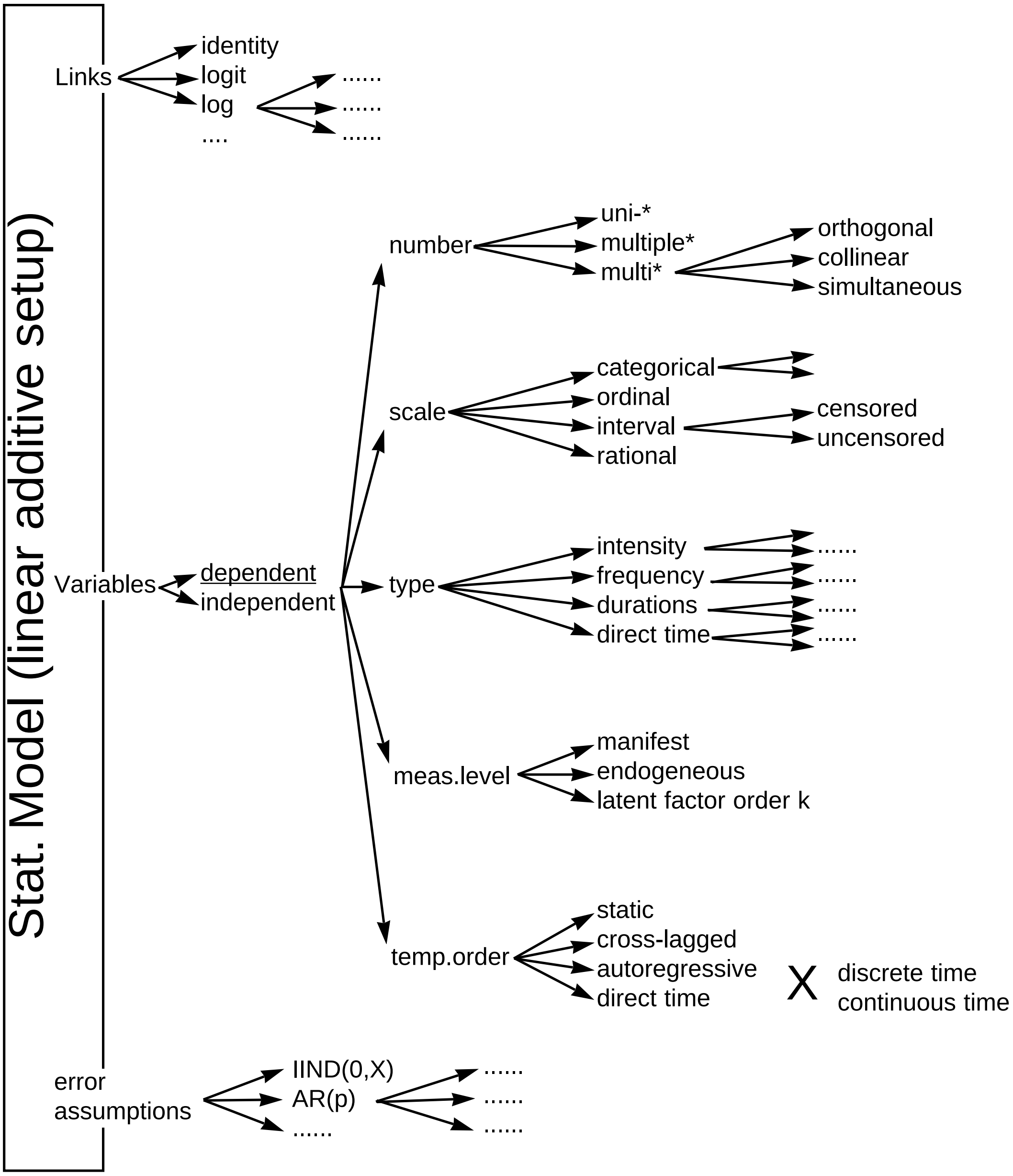
Stat. Model (linear additive setup)
Links
identity
logit
log
....
......
......
......
Variables
dependent
independent
number
uni-*
multiple*
multi*
orthogonal
collinear
simultaneous
scale
categorical
ordinal
interval
rational
censored
uncensored
type
intensity
frequency
durations
direct time
......
......
......
......
meas.level
manifest
endogeneous
latent factor order k
temp.order
static
cross-lagged
autoregressive
direct time
X
discrete time
continuous time
error
assumptions
IIND(0,X)
AR(p)
......
......
......
......

Figure 4 Building blocks in the Generalized Linear Model (GLM).

Table 3 Common models for analysis of designs three/six.

Idea of effect	*Independent (group-time)*	*Scale of single dependent*	*Analysis/test*
Simple change	One sample pre–post	categorical	McNemar test
		ordinal	Wilcoxon test
		metric	dep. t-test
	One sample k times	categorical	Cochran Q-test
		ordinal	Friedman test
		metric	One-way ANOVA rep. meas.
Post-test differences	Two samples	categorical	Simple χ^2
		ordinal	Mann–Whitney test
		metric	indep. t-test
m samples		categorical	Logit/Loglinear models
		ordinal	Kruskal–Wallis test
		metric	One-way ANOVA
Relative change (interaction)	m samples k times each	categorical	Logit/Loglinear models
		ordinal	Loglinear models
		metric	m-way ANOVA rep. meas.

Further insight can he gained by the introduction of the analyses of variance. These models are considered the standard for analyses of multiple group comparison designs, but there is no hope of covering its various features and the richness of its submodels in detail. For the time being, do not bother with the distinction of fixed-effect, random-effect, and mixed models but with the rationale of analysis (for details, see Hays, (1980). Especially in the case of repeated measurement models, there are, however, some ideas and assumptions that must be understood before actually using these techniques, most prominent and important among these are: heteroscedasticity, additivity, orthogonality, balance, degrees of freedom, compound symmetry sphericity. In addition, further precautions as regards covariance structure are in order when extending the model to multivariate analysis of several dependent variables. For advanced reading on these topics whose consideration is indispensible for enhanced statistical conclusion validity, see Toutenburg (1994) or Christensen (1991).

Consider that response data $\vec{\boldsymbol{y}} = [y_1 y_2 \ldots y_m \ldots y_M]$ on M dependent variables are obtained on interval-scale level for $k = 2 \ldots k \ldots K$ samples of equal size; for convenience, $n = 1 \ldots i \ldots N$ at $t = 1, 2 \ldots t \ldots T$ occasions. A categorical independent variable (or factor $f1$ denoting treatment(s) $f1 = (f1_1, \ldots, f1_{\bar{k}}, \ldots, f1^K)$ has been realized at least twice in one sample or independently across different samples at any occasion $1 < t < T$. Treatment $f1_{\bar{t}}$ denotes "no treatment" for implementing nonequivalent control group(s). If $\hat{\boldsymbol{y}}$ offers data on more than one dimension ($M > 1$), multivariate analysis (MANOVA) is called for, if it reduces to a scalar ($M = 1$) but extends in time ($T > 1$), univariate repeated measures ANOVA is applicable. For $M = 1$ and $T = 1$, effects are evaluated using a post-test-differences-only-given-pretest-equivalence rationale as presented in the first row in Table 3 (performed with a t-test for $k = 2$ both uni- and multivariate and any (M)ANOVA for $k > 2$). With repeated measures carrying effect information ($T > 1$) as it holds for both other rationales, data setup is a bit more sophisticated. Common software packages usually take repeated measures of one variable for separate variables (maybe with the same root name: p.ex. *out*1 *out*2 *out*3 … for criterion variable *out* at three occasions). It follows that a multivariate data setup of the input data matrix is required. Note that when more than one genuine variable besides *out* enters analysis, a doubly multivariate procedure is called for. Stated differently, M variables are to be analyzed whose respective T time-dependent realizations are coded as T different variables, giving $M \times T$ data columns in a typical data entry spreadsheet of minimal logical dimension $(N * K) \times (M * T)$ since data are obtained for K groups of N subjects each. This holds true, so far, for any one-way (M)ANOVA with a k-level treatment-factor, affecting one or more metric dependent variables. In most instances, however, it is likely that this root design will be extended by introducing:

(i) further experimental factors subject to manipulation (e.g., single vs. group therapy),

(ii) further cross-classifying factors that fall into place from basic sample analyses (e.g., age, sex, hospitality, etc.),

(iii) further metric time-invariant covariates (e.g., onset of disease, initial dosage of medication), and

(iv) further metric time-varying covariates (e.g., ongoing medication).

The general idea of analysis of variance is rather simple: if means between k groups and/or t occasions differ, and this is what is called an effect, roughly, then they are supposed to vary from the grand mean μ obtained from collapsing over all $M \times K$ data points disregarding groups and treatments or collapsing over occasions giving the grand mean as a "no-change" profile of t equal means. If they do vary significantly then this variation should be to a certain proportion greater than variation present within groups (i.e., with respect to the mean in a group). More specifically, a partitioning of total variance found in the data matrix into (at least) two components is attempted: variance of *groups'* means reflects the systematic ("causal model") part that is solely attributable to variation of treatment (be it in time or across groups), whereas variance of subjects' measurements within a group is generated by interindividual differences (i.e., violation pretest-equivalence or lack-in-treatment integrity) and is therefore taken as an "error term" against which the other part must be evaluated. Remember that the ratio of two variances or two sums of squares divided according to degrees of freedom constitutes on F-value when certain assumptions on data (Hays, 1980) are met: statistical significance is readily tested for. In this way, model variance due to effects, possibly containing interactions when more independent variables are present and error variance due to anything else, add up to give the total variance (assuming a fixed effect model). In the case of repeated measures analysis, there is also variance among t different measures of a single variable within one subject. To sum up: three main effect sources of variation are accounted for by standard ANOVA models:

(i) differences between individuals, to be further decomposed into differences between k groups (disregarding the individual by taking means and thereby constituting a "true value" outcome measure for this treatment) and differences between subjects within a group, obtained from subtracting the individual's measure from the group mean. In analyses with no repeated measures factor present, these differences contribute to error variance against which the first kind of variation is evaluated. In the case of f-way ANOVAs, interactions of independent variables (factors) must be taken into account as well.

(ii) Differences within repeated measures of a subject and according interactions with factors varied between groups in f-way analyses. This variation may be further decomposed into three sources: a main effect of retesting (supposed to be present in each subject's profile over time), an interaction of treatment and retesting, and differential reactions to retesting (specific to each of the subjects). In case of analyses with no between-factors present, the latter source of variation is taken to contribute to residual variance. In multiway repeated measures analyses of variance, determination of appropriate components for a F-ratio in order to test for statistical significance of a factor or an interaction of some factors, can be very complicated; excellent advice is given in Bortz (1993).

For design six, a differential change would be required for an effect by testing the interaction term of a repeated measures "within" factor and the case–control "between" factor for significance using the appropriate F ratio. Change takes place but in different directions (or none) depending on whether receiving treatment or not. Suppose one group of aggressive children received behavior modification training while another group of equal size is still waiting for admission and thus serves as a control group. Both groups are considered equivalent on the dependent variable of interest beforehand. Take it that $t = 3$ measures on each of $N_1 + N_2$ individuals in $k = 2$ groups: pretest, post-test after training, and a follow-up retest six months after intervention to estimate the stability of possible effects. Formally expressing the above sources of variation in a structural model for the individual measure, as is common for GLM submodels, and dispensing with explicit notation of associated dummies for discrete independent variables leads to:

$$y_{ikt} = \mu + a_k + \beta_t + (\alpha\beta)_{kt} + \gamma_{ik} + \epsilon_{ikt}$$

where α_k refers to effect of being in either treatment or control group, β_t refers to differences solely attributable to time or change, $(\alpha\beta)_{kt}$ stands for the interaction of both main effects and thus denotes differential change as sketched above and in greater detail in Section 3.04.1.3, γ_{ikt} refers to a random (= erroneous) effect of a single subject i in the kth kind of treatment (training vs. control, in the case present), maybe for individual specific reaction types to either form of treatment at any occassion, and ϵ_{ikt} accounts for anything not covered by the model (error term).

For full coverage of mathematical assumptions underlying this model, see Toutenburg (1994). As mentioned before, statistical hypotheses are formulated as null hypotheses conjecturing "no effect" of the according term in the model. In addition, for the present example, effects of several levels of an independent variable are required, coded by an according contrast in the hypothesis matrix (see above), which add up to zero. Hypotheses testable in this model include:

H_0: $\alpha_1 = \alpha_2$
There is no difference between treatment and control group with respect to the outcome variable (dependent), disregarding point of time of measurement. More generally speaking, such between hypotheses state homogeneity of measurement levels in groups. Note that no differences at pretest time are expected. Moreover, differences possibly present at post-test time (near termination of intervention) tend to attenuate up until the follow-up testing. It is not known whether group differences were at maximum by consideration of this main effect of group membership alone.

H_0: $\beta_{c1(t)} = \beta_{c2(t)}$
specific formulation depending on contrasts $c(t)$ for time effects.There is no overall change between occasions disregarding group membership. Again more generally speaking, the hypothesis states homogeneous influence of time effects in the course of experimentation. Considering this time main effect alone, it is known that whatever change occurred applied unspecifically to both groups and can therefore not be attributed to treatment. Obviously, this effect does not carry the information sought. When certain assumptions on covariance structure in time are met (sphericity condition, see above), a $t-1$ to orthonormal linear contrasts may be used to evaluate change at certain periods using univariate and averaged tests. These assumptions assure, roughly speaking, that present differences are due to inhomogeneity of means rather than of variances (for details, see Toutenberg, 1994). Major software packages offer correction methods to compensate for violations of these assumptions (e.g., by correcting components of the F-test) or have alternatives available.

H_0: $(\alpha\beta)_{ij} = 0$
This hypothesis states parallelism of progressive forms in groups or, equivalently, the absence of differential change. If this hypothesis can be rejected, there is an interaction of treatment and the repeated measurement factor in the sense that (at least) one group apparently took a different course from pretest to post-test and follow-up as regards the outcome criteria. This is precisely what was presented as the standard notion of an effect in Section 3.04.3. Such a change can be attributed to treatment because the overall effects like common linear trends and so on, have been accounted for by the time of main effect. In more complex cases, there may also be interest in differential growth rates or whether one group arrived sooner at a saturation level (suggesting earlier termination of treatment).

Actual testing of these hypotheses might be performed using "means models" available with the (M)ANOVA procedures or by parameterizing the above equation to fit the general linear model regression equation with identity link. In essence, this is done by building a parameter vector $\boldsymbol{\beta}$ of semipartial regression coefficients from all effects noted (μ, α_1, … α_{k-1}, $\beta_{c_1(t)}$ … β_{ct-1}), associated by the corresponding hypotheses matrix X (made up from coded vectors of contrasted independent variables) and an additive error term. This right-hand part of the equation is regressed on to the vector $\vec{\mathbf{y}}$ giving a standard Gauss–Markov–Aitken estimate of $\beta = (X'\Sigma^{-1}X)^{-1}\, X'\Sigma^{-1}y$ from least squares normal equations. A typical printout from respective software procedures includes parameter value, associated standard error, and t value, confidence limits for parameter estimate, η^2 effect size, F-noncentrality parameter, and power estimate. When interpreting results, remember that an effect is conceptualized as a difference of respective group (= treatment) mean from the grand mean, and that effects are usually normalized to add up to zero over all treatments (for details and exceptions, see Hays, 1980).

Analysis of the worked example study was carried out as follows:

> A one-way ANOVA with repeated measurements was used to test statistical significance of recorded difference in three dependent variables for five experimental conditions at pre, post, and follow-up occasions. There was no significant overall main group effect. Only systolic blood pressure was found to differ across groups at follow-up time with greatest reduction present in a comparison of TR to TOC. Group-time-interactions for testing differential change were carried out by an ANOVA using repeated contrasts (pre vs. post; post vs. follow-up). Systolic blood pressure decreased significantly over time in all relaxation groups and the placebo group, but not in the control group. TR muscle relaxation showed greatest reduction of diastolic pressure, whereas SR and COG significantly lowered heart rate. Pooled treatment effects were evaluated against nontreatment groups by special ANOVA contrasts, showing significant reduction in systolic pressure when compared to control group and significant reduction in diastolic pressure when compared to the placebo group. There were no effects on heart rate in comparison to either one of the nontreatment groups. The results indicate a strong placebo attention effect. It is suggested that independence of heart rate and significant blood pressure reduction is in line with the findings in the field.

With additional restrictions put on Σ, this rationale of testing easily extends to multivariate and doubly multivariate cases (Toutenburg, 1994). Also, there is much appeal in the idea of comparing groups on entire linear systems obtained from all measured variables by structural equations modeling (SEM; see Duncan, 1975). Here, measurement error in both independent and dependent variables is accounted for explicitely. For an introduction to multiple-group covariance structure analyses and to models additionally including structured means see Bentler (1995), Möbus and Bäumer (1986), and Sörbom (1979). MANOVA and SEM are compared on theoretical and practical grounds in Cole, Maxwell, Arvey, and Salas (1993). Note that the latest formulations of these models and associated software packages can now deal with nonnormal and categorical data.

When ordinal measures have been obtained, ANOVA on ranks of subjects" measures (with respect to the order in their respective group) rather than on absolute values may be applied instead of "genuine" ordinal methods as listed in Table 3. In the case of categorical measures, frequencies mostly have been obtained denoting the number of subjects with the characteristic feature of the category level in question present relative to the total number of subjects over the range of all possible levels of that categorical variable. Basically, analysis is carried out on an aggregate level (groups or subpopulations) rather than on the individual level, and link functions other than identity link are employed to regress independent variables (which have been of a categorical nature all the time: group membership, etc.) on to categorical outcome measures (now dependent variables have also become categorical). Regression coefficients must be interpreted differently, though, when dependent variables have been reconceptualized in this way. A positive β indicates that the probability of a specific category of a dependent variable is positively affected (risen) relative to another category, including the case of simply not being in that very category, when the characteristic feature of the category for an independent variable is present in a subject. In the case of a dichotomous dependent variable (logit models) interpretation is facilitated: β quantifies the effect of the independent variable on the odds, that is, the ratio of favorable to unfavorable responses (e.g., whether or not cancer is or will be present given the information available from independent variables). On a more formal level a nonzero regression coefficient in log-linear models indicates that the expected cell frequency in the category r of R categories or "levels" of the dependent variable departs from what would be expected in the "null" case of no effects, that is $n_r = n. = N/R$. As with the metric case, regression coefficients are considered effect parameters and tested for statistical significance by evaluating their standard error. Multiple group comparisons can be obtained in one of two ways: by introducing an independent variable denoting group membership, or by testing whether there is a significant change in the associated overall model fit indices from the more saturated model incorporating the effect in question (H_1) to the more restricted model dispensing with it (H_0).

3.04.7 CONCLUSION

In the last section are described the statistical models for multiple group comparisons. In most cases, however, analysis will be carried out using techniques presented in Table 3. It is reemphasized that it is not the complexity or elegance of the final data model that puts a good end to the process of quasi-experimental research: it is the fit of research questions or hypotheses and according methodological tools to tackle them. This is why a major part in this chapter is devoted to a presentation of an idealized procedural outline of this research process as sketched in Figure 1, and on discussing validity affairs. Generally speaking, the last step in this process lies with reversing it inductively and with introducing evaluative elements. It must be decided by which degree present patterns of statistical significance justify the induction that predictions derived from substantial hypotheses have instantiated as they should and whether validity issues have been attended to a degree that inference to the level of theory is admissible. Limitations in generalizability should be examined and problems encountered in the course of the study must be evaluated in order to point to future directions of research. With reference to Section 3.04.3.3, implications to clinical practice should be derived and be evaluated. In publications, these issues are usually covered in the "Results" and "Discussion" sections.

In the worked example study, the following discussion is offered:

> A result most relevant to clinical practice lies with the fact that cognitive and muscular relaxation both significantly reduce blood pressure and therapists may therefore choose from both methods with respect to their individual clients" special needs and aptitudes (differential indication). The critical prerequisite for this conclusion in terms of validity lies with clear separation of experimental instructions concerning either musculary or cognitively induced relaxation. It might be argued that any instruction relevant to improving irritating bodily sensations does feature cognitive elements, like expectancy, that contribute to intervention

effects. It might be better to differentiate between techniques involving direct muscular practice and a pool of techniques that dispense with the behavioral element. In the end, all relaxation techniques affect muscle tension. But persons might well vary on the degree to which they can benefit from direct or indirect techniques. It is known that more susceptible persons, or persons with greater ability to engage in and maintain vivid imagination and concentration, will benefit more from cognitive techniques, and that physical disabilities often rule out application of muscularly directed interventions. On the other hand, muscle relaxation instructions are considered easier to follow, and, after all, straight muscle relaxation techniques showed the greatest overall effect in the present study, thus pointing to a "safe-side" decision. These arguments refer to client selection in clinical practice as a parallel to the referred effects of subject selections in experimentation.

Accounting for the demand for both methodologically valid and practically relevant methods, the following guidelines for quasi-experimental research may be offered:

(i) preference for case–control designs,

(ii) preference for repeated measures designs, thus preference for diagnostic devices and statistical methods sensitive to (relative) change,

(iii) specification of hypotheses in terms of effect sizes,

(iv) consideration of multiple indicators per criterion, entailing a multivariate data setup, and

(v) special attention to balancing N (sample size), T (rep. measures level), ES (effect size), and other factors determining both statistical and practical significance of results presented in Section 3.04.3.1.

3.04.8 REFERENCES

Arminger, G., Clogg, C. C., & Sobel, M. E. (Eds.) (1995). *Handbook of statistical modeling for the social and behavioral sciences.* New York: Plenum.

Bentler, P. M. (1995). *EQS structural equations program manual.* Encino, CA: Multivariate Software.

Bernstein, I. N. (Ed.) (1976). *Validity issues in evaluative research.* Beverly Hills, CA: Sage.

Bortz, J. (1993). *Statistik.* Berlin: Springer.

Box, G. E., Jenkins, G. M., & Reinsel, G. C. (1994). *Time series analysis: forecasting and control.* Englewood Cliffs, NJ: Prentice-Hall.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental design for research on teaching.* In N. L. Gage (Ed.), *Handbook for research on teaching,* (pp. 171–246). Chicago: Rand McNally.

Christensen, R. (1991). *Linear models for multivariate, time series, and spatial data.* New York: Springer.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Cole, D. A., Maxwell, S. E., Arvey, R., & Salas, E. (1993). Multivariate group comparison of variable systems: MANOVA and structural equation modeling. *Psychological Bulletin, 114,* 174–184.

Collingwood, R. G. (1940). *An essay on metaphysics.* Oxford, UK: Oxford University Press.

Collins, A. W. (1966). The use of statistics in explanation. *British Journal for the Philosophy of Science, 17,* 127–140.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand McNally.

Cronbach, L. J., & Furby, L. (1970). How should we measure "change"—or should we? *Psychological Bulletin, 74,* 68–80.

Diggle, P. J., Liang, K. Y., & Zeger, S. L. (1994). *Analysis of longitudinal data.* Oxford, UK: Oxford University Press.

Duncan, O. D. (1975). *Introduction to structural equations models.* New York: Academic Press.

Eye, A. V. (Ed.) (1990). *Statistical methods for longitudinal research.* Boston: Academic Press.

Farnum, N. R., & Stanton, L. W. (1989). *Quantitative forecasting methods.* Boston: Kent Publishing.

Fisher, R. A. (1959). *Statistical method and statistical inference.* Edinburgh, UK: Oliver & Boyd.

Fitz-Gibbon, C. T., & Morris, L. L. (1991). *How to design a program evaluation. Program evaluation kit, 3.* Newbury Park, CA: Sage.

Giere, R. N. (1972). The significance test controversy. *British Journal for the Philosophy of Science, 23,* 170–181.

Glass, G. V., Willson, V. L., & Gottman, J. M. (1975). *Design and analysis of time series experiments.* Boulder, CO: Associated University Press.

Gottman, J. M. (1981). *Time series analysis.* Cambridge, UK: Cambridge University Press.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37,* 424–438.

Grawe, K. (1992). Psychotherapieforschung zu Beginn der neunziger Jahre. *Psychologische Rundschau, 43,* 132–168.

Gregson, R. (1983). *Times series in Psychology.* Hillsdale, NJ: Erlbaum.

Hager, W. (1992). *Jenseits von Experiment und Quasiexperiment: Zur Struktur psychologischer Versuche und zur Ableitung von Vorhersagen.* Göttingen, Germany: Hogrefe.

Hagenaars, J. A. (1990). *Categorical longitudinal data: log-linear panel, trend and cohort analysis.* Newbury Park, CA: Sage.

Hays, W. (1980). *Statistics for the social sciences.* New York: Holt, Rinehart, & Winston.

Hoole, F. W. (1978). *Evaluation research and development activities.* Beverly Hills, CA: Sage.

Jöreskog, K., & Sörbom, D. (1993). *LISREL 8 user's reference guide.* Chicago: Scientific Software International.

Kepner, J. L., & Robinson, D. H. (1988). Nonparametric methods for detecting treatment effects in repeated measures designs. *Journal of the American Statistical Association, 83,* 456–461.

Kerlinger, F. (1973). *Foundations of behavioral research* (2nd ed.). New York: Holt, Rinehart, & Winston.

Lindsey, J. K. (1993). *Models for repeated measurements. Oxford statistical science series, 10.* Oxford, UK: Claredon Press.

Lipsey, M. W. (1990). *Design sensitivity: statistical power for experimental research.* Newbury Park, CA: Sage.

Lord, F. M. (1963). Elementary models for measuring change. In C. W. Harris (Ed.), *Problems in measuring change.* Madison, WI: University of Wisconsin Press.

Mayo, D. (1985). Behavioristic, evidentialist and learning models of statistical testing. *Philosophy of Science, 52,* 493–516.

McDowall, D., & McCleary, R., Meidinger, E. E., & Hay, R. A. (1980). *Interrupted time series analysis.* Newbury Park, CA: Sage

Möbus, C., & Bäumer, H. P. (1986). *Strukturmodelle für*

Längsschnittdaten und Zeitreihen. Bern, Switzerland: Huber.

Pedhazur, E. J., & Pedhazur, L. (1991). *Measurement, design and analysis: an integrated approach*. Hillsdale, NJ: Erlbaum.

Petermann, F. (Ed.) (1996). *Einzelfallanalyse* (3rd ed.). München, Germany: Oldenbourg.

Prioleau, L., Murdock, M., & Brody, N. (1983). An analyis of psychotherapy versus placebo studies. *The Behavioural and Brain Sciences, 6*, 275–285.

Plewis, I. (1985). *Analysing change: measurement and explanation using longitudinal data*. New York: Wiley.

Putnam, H. (1991). The "Corroboration" of theories. In R. Boyd, P. Gasper, & J. D. Trout (Eds.), *The philosophy of science* (pp. 121–137). Cambridge, MA: MIT Press.

Rao, P., & Miller, R. L. (1971). *Applied econometrics*. Belmont, CA: Wadsworth.

Rossi, P. H., Freeman, H. E., & Wright, S. R. (1979). *Evaluation: a systematic approach*. Beverly Hills, CA: Sage.

Salmon, W. C. (1966). *The foundations of scientific inference*. Pittsburg, OH: University of Pittsburgh Press.

Schmitz, B. (1989). *Einführung in die Zeitreihenanalyse*. Bern, Switzerland: Huber.

Smith, M. L, Glass, V. G., & Miller, T. I. (1980). *The benefits of psychotherapy*. Baltimore: Johns Hopkins University Press.

Solomon, R. L. (1949). An extension of control group design. *Psychological Bulletin, 46*, 137–150.

Sörbom, D. (1979). A general model for studying differences in factor means and factor structure between groups. In K. G. Jöreskog & D. Sörbom (Eds.), *Advances in factor analysis and structural equation models* (pp. 207–217). Cambridge, MA: Abt Books.

Spector, P. E. (1981). *Research designs*. Beverly Hills, CA: Sage.

Thompson, G. L. (1991). A unified approach to rank tests for multivariate and repeated measures designs. *Journal of the American Statistical Association, 86*, 410–419.

Toutenburg, H. (1994). *Versuchsplanung und Modellwahl: statistische Planung und Auswertung von Experimenten mit stetigem oder kategorialem Response*. Heidelberg, Germany: Physica.

Trochim, W. M. (1984). *Research design for program evaluation: the regression-discontinuity approach. Contemporary evaluation research, 6*. Beverly Hills, CA: Sage.

Vonesh, E. F. (1992). Non-linear models for the analysis of longitudinal data. *Statistics in Medicine, 11*, 1929–1954.

Yung, P. M. B., & Kepner, A. A. (1996). A controlled comparison on the effect of muscle and cognitive relaxation procedures on blood pressure: implications for the behavioral treatment of borderline hypertensives. *Behavioral Research and Therapy, 34*, 821–826.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

3.05
Epidemiologic Methods

WILLIAM W. EATON
Johns Hopkins University, Baltimore, MD, USA

3.05.1 INTRODUCTION 92
3.05.1.1 Epidemiology and Clinical Psychology 92
3.05.1.2 Dichotomy and Dimension 92
3.05.1.3 Exemplar Study: Ecological Approach to Pellagra 93
3.05.2 RATES 94
3.05.2.1 Prevalence 94
3.05.2.2 Incidence 95
3.05.2.3 Incidence and Onset 96
3.05.3 MORBIDITY SURVEYS 99
3.05.3.1 Exemplar Study: The Epidemiologic Catchment Area (ECA) Program 99
3.05.3.2 Defining and Sampling the Population 100
3.05.3.3 Sample Size 101
3.05.3.4 Standardization of Data Collection 101
3.05.4 EXPOSURE AND DISEASE 103
3.05.4.1 Cohort Design 103
3.05.4.2 Exemplar Study: the British Perinatal Study 103
3.05.4.3 Case–Control Design 104
3.05.4.4 Exemplar Study: Depression in Women in London 105
3.05.5 USING AVAILABLE RECORDS 105
3.05.5.1 Government Surveys 105
3.05.5.2 Vital Statistics 106
3.05.5.3 Record-based Statistics and Record Linkage 106
3.05.5.4 Exemplar Study: the Danish Adoption Study 107
3.05.6 PREVENTIVE TRIALS 107
3.05.6.1 The Logic of Prevention 107
3.05.6.2 Attributable Risk 109
3.05.6.3 Developmental Epidemiology 109
3.05.6.4 Exemplar Study: the Johns Hopkins Prevention Trial 110
3.05.7 BIAS 112
3.05.7.1 Sampling Bias 112
3.05.7.2 Measurement Bias 113
3.05.8 CONCLUSION: THE FUTURE OF PSYCHOLOGY IN EPIDEMIOLOGY 113
3.05.9 REFERENCES 114

3.05.1 INTRODUCTION

3.05.1.1 Epidemiology and Clinical Psychology

Is psychological epidemiology an oxymoron? The primary goals of psychology differ from the primary goals of epidemiology. Psychology seeks to understand the general principles of psychological functioning of individual human beings. Epidemiology seeks to understand pathology in populations. These two differences—general principles vs. pathology, and individuals vs. populations—form the foundations for ideologic differences which have led to failures of linkage between the two disciplines, and many missed opportunities for scientific advances and human betterment. The gap between the two is widened by their location in different types of institutions: liberal arts universities vs. medical schools.

Various definitions of epidemiology reflect the notion of pathology in populations. Mausner and Kramer define epidemiology as "the study of the distribution and determinants of diseases and injuries in human populations" (1985, p. 1). Lilienfeld and Stolley use a very similar definition but include the notion that "epidemiologists are primarily interested in the occurrence of disease as categorized by time, place, and persons" (1994, p. 1). A less succinct, but nevertheless classic, statement is by Morris (1975), who listed seven "uses" of epidemiology:

(i) To study the hertory of health of populations, thus facilitating projections into the future.

(ii) To diagnose the health of the community, which facilitates prioritizing various health problems and identifying groups in special need.

(iii) To study the working of health services, with a view to their improvement.

(iv) To estimate individual risks, and chances of avoiding them, which can be communicated to individual patients.

(v) To identify syndromes, by describing the association of clinical phenomena in the population.

(vi) To complete the clinical picture of chronic diseases, especially in terms of natural hertory.

(vii) To search for causes of health and disease.

The seventh use is regarded by Morris as the most important, but, in reviewing the literature on psychiatric disorders for the third (1975) edition of his text, he was surprised at

> The sad dearth of hard fact on the causes of major as well as minor mental disorders, and so on how to prevent them ... [and] the dearth of epidemiological work in search of causes ... Lack of ideas? Methods? Money? It wants airing. Of course, the popular trend open-endedly to hand over to psychiatry (and to social work), in hope or resignation, the whole of the human condition is no help. (p. 220)

3.05.1.2 Dichotomy and Dimension

Pathology is a process, but diseases are regarded as present or absent. Psychologists have been interested historically in promoting health and in understanding the nature of each individual's phenomenology, leading to emphasis on dimensions of behavior; but epidemiologists have focused on disorders without much regard to individual differences within a given diagnostic category, leading to focus on dichotomies. For clinical psychologists adopting the epidemiologic perspective, the phrase "psychiatric epidemiology" could describe their work accurately (Tsuang, Tohen, & Zahner, (1995). Another area of work for psychologists in epidemiology focuses on psychological contributors to diseases in general, which can include psychiatric disorders but not be limited to them. This type of epidemiology might be termed "psychosocial epidemiology" (Anthony, Eaton, & Henderson, (1995), and receives less attention in this chapter.

Interest in normal functioning, or in deviations from it, has led psychologists to focus on statistical measures related to central tendency, such as the mean, and the statistical methods oriented toward that parameter, such as the comparison of means and analysis of variance. Epidemiology's base in medicine has led it to focus on the dichotomy of the presence or absence of disease. There is the old adage in epidemiology: "The world is divided into two types of persons: those who think in terms of dichotomies, and those who do not." The result for statistical analysis is recurring focus on forms involving dichotomies: rates and proportions (Fleiss, 1981), the two-by-two table (Bishop, Fienberg, & Holland, 1975), the life table (Lawless, 1982), and logistic regression (Kleinbaum, Kupper, & Morgenstern, 1982).

The epidemiologic approach is focused on populations. A population is a number of individuals having some characteristic in common, such as age, gender, occupation, nationality, and so forth. Population is a broader term than group, which has awareness of itself, or structured interactions: populations typically are studied by demographers, whereas groups typically are studied by sociologists. The general population is not limited to any particular class or characteristic, but includes all persons without regard to special features. In the context of epidemiologic research studies, the general population usually refers to the individuals who normally reside in the locality of the study. Many epidemiologic studies involve large populations, such as a nation, state, or a county of several hundred thousand inhabitants. In such large studies, the general population includes individuals with a wide range of social and biological characteristics, and this variation is helpful in generalizing the results. In the broadest sense, the general population refers to the human species. The fundamental parameters used in epidemiology often are conceptualized in terms of the population: for example, prevalence sometimes is used to estimate the population's need for health care; incidence estimates the force of morbidity in the population. Generally diseases have the characteristic of rarity in the population. As discussed below, rates of disease often are reported per 1000, 10 000, or 100 000 population.

Much of the logic that distinguishes the methods of epidemiology from those of psychology involves the epidemiologic focus on rare dichotomies. The case–control method maximizes efficiency in the face of this rarity by searching for and selecting cases of disease very intensively, for example, as in a hospital or a catchment area record system, and selecting controls at a fraction of their representation in the general population. Exposures can also be rare, and this possibility is involved in the logic of many cohort studies which search for and select exposed individuals very intensively, for example, as in an occupational setting with toxins present, and selecting nonexposed controls at a fraction of their representation in the general population. In both strategies the cases, exposed groups, and controls are selected with equal care, and in both strategies they are drawn, if possible, to represent the population at risk of disease onset: the efficiency comes in the comparison to a relatively manageable sample of controls.

3.05.1.3 Exemplar Study: Ecological Approach to Pellagra

The ecologic approach compares populations in geographic areas as to their rates of disease, and has been a part of epidemiology since its beginnings. A classic ecologic study in epidemiology is Goldberger's work on pellagra psychosis (Cooper & Morgan, 1973). The pellagra research also illustrates the concept of the causal chain and its implications for prevention. In the early part of the twentieth century, pellagra was most prevalent in the lower classes in rural villages in the southeastern US. As the situation may be in the 1990s for many mental disorders (Eaton & Muntaner, 1996), the relationship of pellagra prevalence to low social class was a consistent finding: a leading etiologic clue among many others with less consistent evidence. Many scientists felt that pellagra psychosis was infectious, and that lower class living situations promoted breeding of the infectious agent. But Goldberger noticed that there were striking failures of infection, for example, aides in large mental hospitals, where pellagra psychosis was prevalent, seemed immune. He also observed certain exceptions to the tendency of pellagra to congregate in the lower class, among upper-class individuals with unusual eating habits. Goldberger became convinced that the cause was a nutritional deficiency which was connected to low social class. His most powerful evidence was an ecologic comparison of high and low rate areas: two villages which differed as to the availability of agricultural produce. The comparison made a strong case for the nutritional theory using the ecological approach, even though he could not identify the nutrient. He went so far as to ingest bodily fluids of persons with pellagra to demonstrate that it was not infectious. Eventually a deficiency in vitamin B was identified as a necessary causal agent. Pellagra psychosis is now extremely rare in the US, in part because of standard supplementation of bread products with vitamin B.

Goldberger understood that low social class position increased risk of pellagra, and he also believed that nutritional deficiencies which resulted from lower class position were a necessary component in pellagra. With the advantage of hindsight, nutritional deficiency appears to have a stronger causal status, because there were many lower class persons who did not have pellagra, but few persons with pellagra who were not nutritionally deprived: vitamin B deprivation is a necessary cause, but lower social class position is a contributing cause. The concept of the causal chain points to both causes, and facilitates other judgements about the importance of the cause. For example, which cause produces the most cases of pellagra? (Since they operate in a single causal chain, the nutritional deficiency is more important from this point of view.) Which cause is antecedent in time? (Here the answer is lower social class position.) Which cause is easiest to change in order to prevent pellagra? The answer to this question is not so obvious, but, from the public health viewpoint, it is the strongest argument about which cause deserves attention. Which cause is consistent with an accepted framework of disease causation supportive of the ideology of a social group with power in society? Here the two causes diverge: funneling resources to the medical profession to cure pellagra is one approach to the problem consistent with the power arrangement in the social structure; another approach is redistributing income so that nutritional deficiency is less common. As it happens, neither approach would have been effective. The consensus to redistribute income, in effect, to diminish social class differences, is not easy to achieve in a democracy, since the majority are not affected by the disease. Bread supplementation for the entire population was a relatively cost-effective public health solution. The more general point is that Goldberger's work encompassed the complexity of the causal process.

Epidemiology accepts the medical framework as defining its dependent variable, but it is neutral as to discipline of etiology. Goldberger did not eliminate one or the other potential cause due to an orientation toward social vs. biological disciplines of study. He did not favor one or the other potential cause because it focused on the individual vs. the collective level of analysis. This eclecticism is an important strength in epidemiology because causes in different disciplines can interact, as they did in the case of pellagra, and epidemiology still provides a framework for study. The public health approach is to search for the most important causes, and to concentrate on those that are malleable, since they offer possibilities for prevention.

3.05.2 RATES

An early and well-known epidemiologist, Alexander Langmuir, used to say that "stripped to its basics, epidemiology is simply a process of obtaining the appropriate numerator and denominator, determining the rate, and interpreting that rate" (cited in Foege, 1996, p. S11). In the sense in which Langmuir was speaking, the term rates includes proportions such as the prevalence "rate" as well as the incidence rate, as explained below. Table 1 shows the minimum design requirements, and the definitions of numerators and denominators, for various types of the rates and proportions. The table is ordered from top to bottom by increasing difficulty of longitudinal follow-up.

3.05.2.1 Prevalence

Prevalence is the proportion of individuals ill in a population. Temporal criteria allow for several types of prevalence: point, lifetime, and period. Point prevalence is the proportion of individuals in a population at a given point in time. The most direct use of point prevalence is as an estimate of need for care or potential treatment load, and it is favored by health services researchers. It is also useful because it identifies groups at high risk of having a disorder, or greater chronicity of the disorder, or both. Finally, the point prevalence can be used to measure the impact of prevention programs in reducing the burden of disease on the community.

Lifetime prevalence is the proportion of individuals, who have ever been ill, alive on a given day in the population. As those who die are not included in the numerator or denominator of the proportion, the lifetime prevalence is sometimes called the proportion of survivors affected (PSA). It differs from the lifetime risk because the latter attempts to include the entire lifetime of a birth cohort, both past and future, and includes those deceased at the time of the survey. Lifetime risk is the quantity of most interest to geneticists. Lifetime prevalence also differs from the proportion of cohort affected (PCA), which includes members of a given cohort who have ever been ill by the study date, regardless of whether they are still alive at that time.

Lifetime prevalence has the advantage over lifetime risk and PCA in that it does not require ascertaining who is deceased, whether those deceased had the disorder of interest, or how likely those now alive without the disorder are to develop it before some given age. Thus, lifetime prevalence can be estimated from a cross-sectional survey. The other measures require either following a cohort over time or asking relatives to identify deceased family members and report symptoms suffered by them. Often these reports must be supplemented with medical records. The need for these other sources adds possible errors: relatives may forget to report persons who died many years before or may be uninformed about their psychiatric status; medical records may be impossible to locate or inaccurate; and the prediction of onset in young persons not yet affected requires the assumption that they will fall ill at the same rate as the older members of the sample and will die at the same ages if they do not fall ill. If risks of disorder or age-specific death rates change, these predictions will fail.

Lifetime prevalence requires that the diagnostic status of each respondent be assessed over his or her lifetime. Thus, accuracy of recall of symptoms after a possible long symptom-free period is a serious issue, since symptoms and disorders that are long past, mild, short-lived, and less stigmatizing are particularly likely to be forgotten. For example, data from several cross-sectional studies of depression seem to indicate a rise in the rate of depression in persons born after World War II (Cross-National Collaborative Group, 1992; Klerman & Weissman, 1989).

Table 1 Rates and proportions in epidemiology.

Rate	*Minimum design*	*Numerator*	*Denominator*
Lifetime prevalence	Cross-section	Ever ill	Alive
Point prevalence	Cross-section	Currently ill	Alive
Period prevalence (1)	Cross-section	Ill during period	Alive at survey
Period prevalence (2)	Two waves	Ill during period	Alive during period
First incidence	Two waves	Newly ill	Never been ill
Attack rate	Two waves	Newly ill	Not ill at baseline
Proportion of cohort affected	Birth to present	Ever ill	Born and still alive
Lifetime risk	Birth to death	Ever ill	Born

These persons are older, however, and it may be that they have forgotten episodes of depression which occurred many years prior to the interview (Simon & Von Korff, 1995). Data showing that lifetime prevalence of depressive disorder actually declines with age (Robins et al., 1984) is consistent with the recall explanation.

Period prevalence is the proportion of the population ill during a specified period of time. Customarily the numerator is estimated by adding the prevalent cases at the beginning of the defined period to the incident (first and recurrent) cases that develop between the beginning and the end of the period. This form is shown as period prevalence (2) in Table 1. In research based on records, all cases of a disorder found over a one-year period are counted. The denominator is the average population size during the interval. Thus, the customary definition of period prevalence requires at least two waves of data collection. Both Mausner and Kramer (1985) and Kleinbaum et al. (1982) have noted the advantages of period prevalence for the study of psychiatric disorders, where onset and termination of episodes is difficult to ascertain exactly (e.g., the failure to distinguish new from recurrent episodes is unimportant in the estimation of period prevalence). Furthermore, the number of episodes occurring during the follow-up is unimportant; it is important only to record whether there was one or more vs. none. The disadvantage of period prevalence is that it is not as useful in estimating need as point prevalence, nor as advantageous in studying etiology as incidence.

Another type of period prevalence, sometimes labeled by a prefix denoting the period, such as one-year prevalence, is a hybrid type of rate, conceptually mixing aspects of point and period prevalence, which has been found useful in the ECA Program (Eaton et al., 1985). This type of period prevalence is labeled period prevalence (1) in Table 1. It includes in the numerator all those surveyed individuals who have met the criteria for disorder in the past year, and as denominator all those interviewed. It is not a point prevalence rate because it covers a longer period of time, which can be defined as six months, two years, and so forth, as well as one year. But one-year prevalence is not a period prevalence rate because some individuals in the population who are ill at the beginning of the period are not successfully interviewed, because they either die or emigrate. As the time period covered in this rate becomes shorter, it approximates the point prevalence; as the time period becomes longer, the rate approaches the period prevalence. If there is large mortality, the one-year prevalence rate will diverge markedly from period prevalence.

3.05.2.2 Incidence

Incidence is the rate at which new cases develop in a population. It is a dynamic or time-dependent quantity and can be expressed as an instantaneous rate, although, usually, it is expressed with a unit of time attached, in the manner of an annual incidence rate. In order to avoid confusion, it is essential to distinguish first incidence from total incidence. The distinction itself commonly is assumed by epidemiologists but there does not appear to be consensus on the terminology. Most definitions of the incidence numerator include a concept such as "new cases" (Lilienfeld & Stolley, 1994, p. 109), or persons who "develop a disease" (Mausner & Kramer, 1985, p. 44). Morris (1975) defines incidence as equivalent to our "first incidence," and "attack rate" as equivalent to our "total incidence." First incidence corresponds to the most common use of the term "incidence," but since the usage is by no means universal, keeping the prefix is preferred.

The numerator of first incidence for a specified time period is composed of those individuals who have had an occurrence of the disorder for the first time in their lives and the denominator includes only persons who start the period with no prior history of the disorder. The numerator for attack rate (or total incidence) includes all individuals who have had an occurrence of the disorder during the time period under investigation, whether or not it is the initial episode of their lives or a recurrent episode. The denominator for total incidence includes all population members except those cases of the disorder which are active at the beginning of the follow-up period.

A baseline and follow-up generally are needed to estimate incidence. Cumulative incidence (not shown in Table 1) is the proportion of the sample or population who become a case for the first time between initial and followup interviews (Kleinbaum et al., 1982). But incidence generally is measured per unit of time, as a rate. When the follow-up period extends over many years, the exposure period of the entire population at risk is estimated by including all years prior to onset for a given individual in the denominator, and removing years from the denominator when an individual has onset or dies. In this manner, the incidence is expressed per unit time per population: for example, "three new cases per 1000 person years of exposure." This method of calculating facilitates comparison between studies with different lengths of follow-up and different mortality. When the mean interval between initial and follow-up interview is approximately 365 days, then the ratio of new cases to the number of followed up respondents, who were ascertained as being with a current or past history of the disorder at the time of the initial interview, can serve as a useful approximation of the disorder's annual first incidence.

First incidence also can be estimated by retrospection if the date of onset is obtained for each symptom or episode, so that the proportion or persons who first qualified in the year prior to the interview can be estimated. For this type of estimate (not shown in Table 1), only one wave of data collection is needed. This estimate of first incidence is subject to the effects of mortality, however, because those who have died will not be available for the interview.

The preference for first or total incidence in etiologic studies depends on hypotheses and assumptions about the way causes and outcomes important to the disease ebb and flow. If the disease is recurrent and the causal factors vary in strength over time, then it might be important to study risk factors not only for first but for subsequent episodes (total incidence): for example, the effects of changing levels of stress on the occurrence of episodes of neurosis (Tyrer, 1985) or schizophrenia (Brown, Birley, & Wing, 1972). For a disease with a presumed fixed progression from some starting point, such as dementia, the first occurrence might be the most important episode to focus on, and first incidence is the appropriate rate. In the field of psychiatric epidemiology, there are a range of disorders with both types of causal structures operating, which has led us to focus on this distinction in types of incidence.

The two types of incidence are related functionally to different measures of prevalence. Kramer et al. (1980) have shown that lifetime prevalence (i.e., the proportion of persons in a defined population who have ever had an attack of a disorder) is a function of first incidence and mortality in affected and unaffected populations. Point prevalence (i.e., the proportion of persons in a defined population on a given day who manifest the disorder) is linked to total incidence by the queuing formula $P = f(I \times D)$ (Kleinbaum et al., 1982; Kramer, 1957), that is, point prevalence is a function of the total number of cases occurring and their duration. In the search for risk factors that have etiologic significance for the disorder, comparisons based on point prevalence rates suffer the disadvantage that differences between groups as to the chronicity of the disorder: that is, the duration of episodes; the probability that episodes will recur; or the mortality associated with episodes; affect the comparisons (Kramer, (1957). For example, it appears that Blacks may have episodes of depression of longer duration than Whites (Eaton & Kessler, 1981). If so, the point prevalence of depression would be biased toward a higher rate for Blacks, based solely on their greater chronicity.

3.05.2.3 Incidence and Onset

Dating the onset of episodes is problematic for most mental disorders for many reasons. One is that the diagnostic criteria for the disorders themselves are not well agreed upon, and continual changes are being made in the definition of a case of disorder, such as the recent fourth revision of the *Diagnostic and statistical manual of mental disorders* (*DSM-III*; American Psychiatric Association, 1994). The *DSM-IV* has the advantage that criteria for mental disorders are more or less explicitly defined, but it is nevertheless true that specific mental disorders are often very difficult to distinguish from nonmorbid psychological states. Most disorders include symptoms that, taken by themselves, are part of everyone's normal experience: for example, feeling fearful, being short of breath or dizzy, and having sweaty palms are not uncommon experiences, but they are also symptoms of panic disorder. It is the clustering of symptoms, often with the requirement that they be brought together in one period of time or "spell," that generally forms the requirement for diagnosis. Although the clustering criteria are fairly explicit in the

DSM, it is not well established that they correspond to the characteristics generally associated with a disease, such as a predictable course, a response to treatment, an association with a biological aberration in the individual, or an associated disability. Thus, the lack of established validity of the criteria-based classification system exacerbates problems of dating the onset of disorder.

The absence of firm data on the validity of the classification system requires care about conceptualizing the process of disease onset. One criterion of onset used in the epidemiology of some diseases is entry into treatment, but this is unacceptable in psychiatry since people with mental disorders so often do not seek treatment for them. Another criterion of onset sometimes used is detectability, that is, when the symptoms first appear, but this is also unacceptable because experiences analogous to the symptoms of most psychiatric disorders are so widespread. It is preferable to conceptualize onset as a continuous line of development towards manifestation of a disease. There is a threshold at which the development becomes irreversible, so that at some minimal level of symptomatology it is certain that the full characteristics of the disease, however defined, will become apparent. This use of irreversibility is consistent with some epidemiological uses (Kleinbaum et al., 1982). Prior to this point, the symptoms are thought of as "subcriterial." At the state of knowledge in psychiatry in the 1990s, longitudinal studies in the general population, such as the ECA program and others mentioned above, are needed to determine those levels of symptomatology at which irreversibility is achieved.

There are at least two ways of thinking about development towards disease. The first way is the increase in severity or intensity of symptoms. An individual could have all the symptoms required for diagnosis but none of them be sufficiently intense or impairing. The underlying logic of such an assumption may well be the relatively high frequency of occurrence of the symptoms in milder form, making it difficult to distinguish normal and subcriterial complaints from manifestations of disease. For many chronic diseases, it may be inappropriate to regard the symptom as ever having been "absent," for example, personality traits giving rise to deviant behavior, categorized as personality disorders on Axis II of the *DSM* (APA, 1994). This type of development can be referred to as "symptom intensification," indicating that the symptoms are already present and have become more intense during a period of observation. This concept leads the researcher to consider whether there is a crucial level of severity of a given symptom or symptoms in which the rate of development towards a full-blown disease state is accelerated, or becomes irreversible.

A second way of thinking about progress towards disease is the occurrence of new symptoms which did not exist before. This involves the gradual acquisition of symptoms so that clusters are formed which increasingly approach the constellation required to meet specified definitions for diagnosis. A cluster of symptoms which occur more often together than would be expected by their individual prevalence in the population, that is, more often than expected by chance, is a syndrome. "Present" can be defined as occurrence either at a nonsevere or at a severe level, thus, decisions made about the symptom intensification process complicate the idea of acquisition. This idea leads the researcher to consider the order in which symptoms occur over the natural history of disease, and in particular, whether one symptom is more important than others in accelerating the process.

Figure 1 is an adaptation of a diagram used by Lilienfeld and Lilienfeld (1980, Figure 6.3) to visualize the concept of incidence as a time-oriented rate which expresses the force of morbidity in the population. In their original figure, (1(a)), the horizontal axis represents time and the presence of a line indicates disease. The adaptations (1(b) and 1(c)) give examples of the several distinct forms that onset can take when the disorder is defined by constellations of symptoms varying in intensity and co-occurrence, as is the case with mental disorders. In Figure 1(b), the topmost subject (1) is what might be considered the null hypothesis, and it corresponds to simple onset as portrayed in the original. Figure 1(b) shows how intensity, represented by the vertical width of the bars, might vary. The threshold of disease is set at four units of width, and in the null hypothesis subject 1 progresses from zero intensity to four units, becoming a case during the observation period. Subject 2 changes from nearly meeting the criteria (width of three units) to meeting it (four units) during the year. Both subjects 1 and 2 are new cases, even though the onset was more sudden in subject 1 than in subject 2, that is, the force of morbidity is stronger in subject 1 than subject 2. Subjects 3 and 4 are not new cases, even though their symptoms intensify during the year as much or more than those of subject 2.

Figure 1(c) adapts the same original diagram to conceptualize acquisition of symptoms and the development of syndromes. At one point in time there is no correlation between symptoms, but the correlation gradually develops, until there is a clear separation of the population into one group, with no association of symptoms,

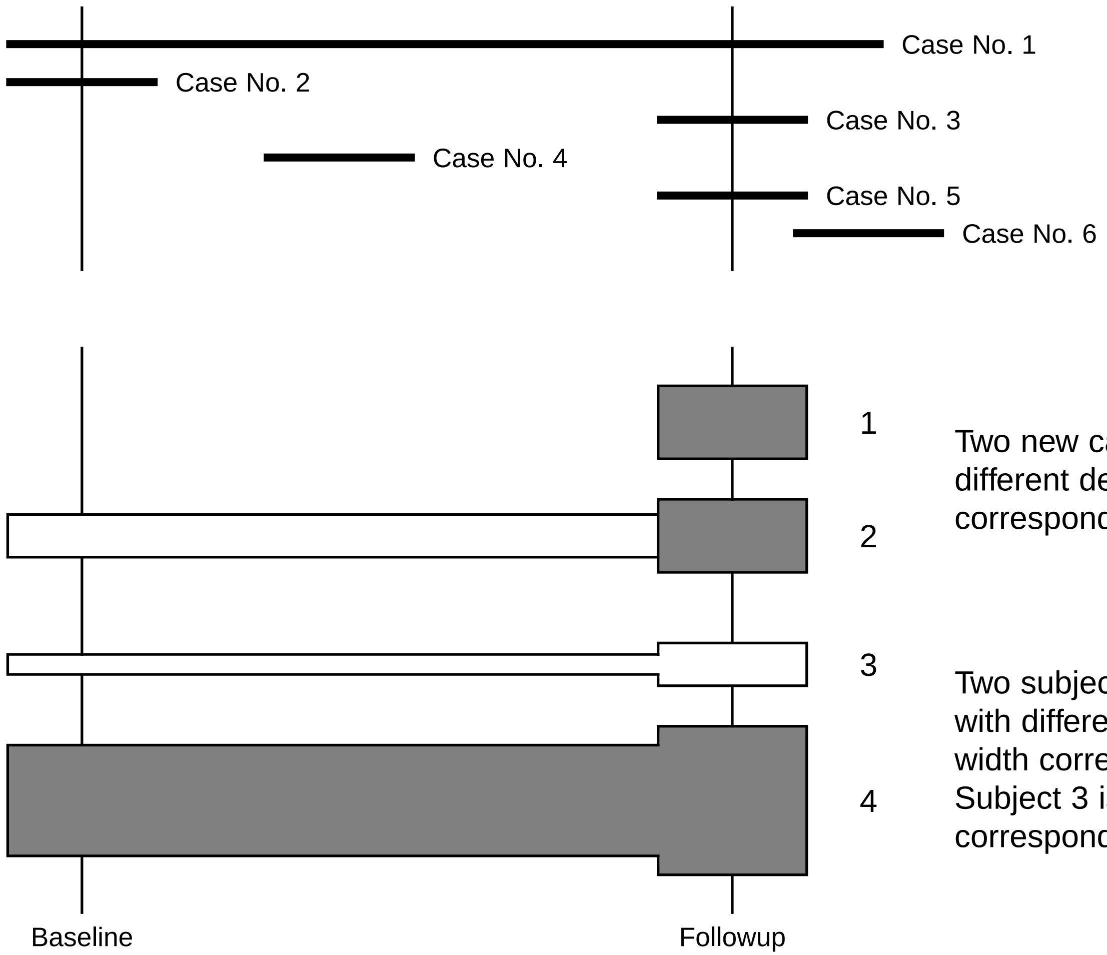
Case No. 1
Case No. 2
Case No. 3
Case No. 4
Case No. 5
Case No. 6
1
2
3
4
Baseline
Followup

Incidence and Intensification

Lilienfeld and Stolley, Figure 6.2; Length of line corresponds to duration of episode.

Two new cases, like No. 3 above, with different degrees of intensification; width corresponds to intensity of symptoms.

Two subjects not defined as new cases, with different degrees of intensification; width corresponds to intensity of symptoms. Subject 3 is never a case, and subject 4 corresponds to Case No. 1 above.

Figure 1(a) and (b) Incidence and intensification.

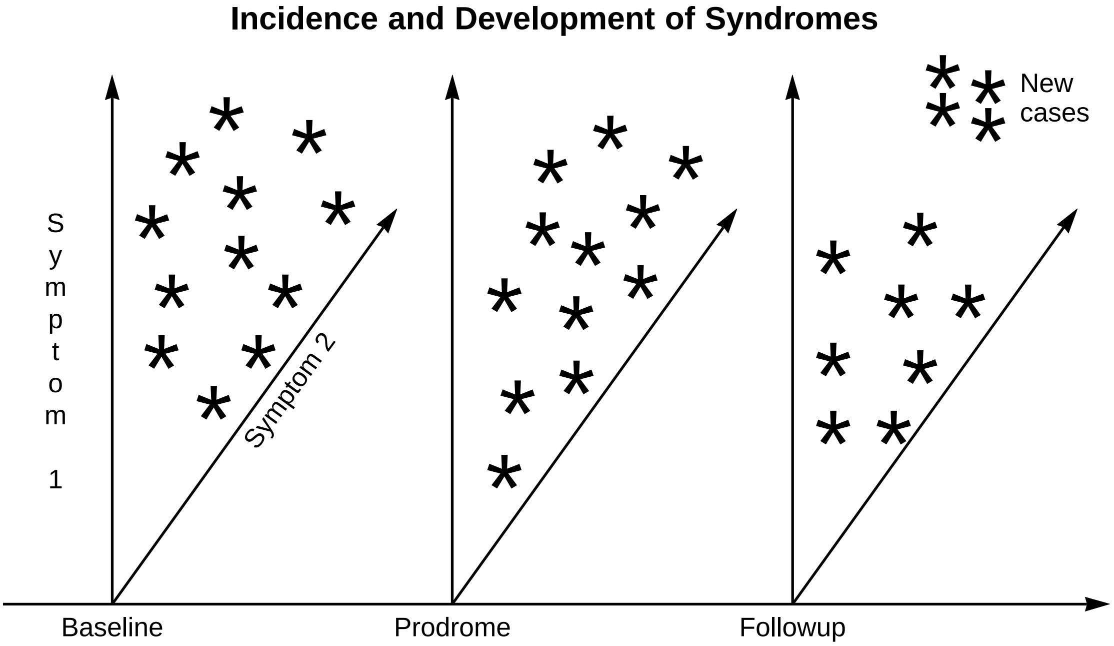

Figure 1(c) Incidence and the development of syndromes.

and another group where the two symptoms co-occur. This example could be expanded to syndromes involving more than two symptoms.

Acquisition and intensification are indicators of the force of morbidity in the population, as are more traditional forms of incidence rate. But they are not tied to any one definition of caseness. Rather, these concepts allow study of disease progression independently of case definition. Risk factors at different stages of the disease may be differentially related to disease progression only above or below the threshold set by the diagnosis. In this situation, we might reconsider the diagnostic threshold.

3.05.3 MORBIDITY SURVEYS

Morbidity surveys are cross-sectional surveys of the general population. They are used to estimate prevalence of disease in the population as well as to estimate need for care. Morbidity surveys address Morris's (1975, discussed above) second "use" of epidemiology, "diagnosing the health of the community, prioritizing health problems and identifying groups in need;" as well as the third use "studying the working of health services;" and the fifth use, "identifying syndromes." Morbidity surveys are sometimes called the descriptive aspect of epidemiology, but they can also be used to generate hypotheses about associations in the population, and to generate control samples for cohort and case–control studies.

3.05.3.1 Exemplar Study: The Epidemiologic Catchment Area (ECA) Program

An example of a morbidity survey is the Epidemiologic Catchment Area (ECA) Program, sponsored by the United States National Institute of Mental Health from 1978 through 1985. The broad aims of the ECA Program were "to estimate the incidence and prevalence of mental disorders, to search for etiological clues, and to aid in the planning of health care facilities" (Eaton, Rogier, Locke, & Taube, 1981). The Program involved sample surveys of populations living in the catchment areas of already designated Community Mental Health Centers. The broad goals of the ECA Program are described in Eaton et al. (1981), the methods are described in Eaton and Kessler (1985), and Eaton et al. (1984), the cross-sectional results are described in Robins and Regier (1991), and the incidence estimates are described in Eaton, Kramer, Anthony, Chee, and Shapiro (1989).

A principal advantage of the morbidity survey is that it includes several or many disorders, which helps in assessing their relative importance from the public health point of view. Another advantage is that estimates of prevalence and association are measured without regard to the treatment status of the sample (Shapiro et al., 1984). Figure 2 displays results from the ECA study, showing the relatively high prevalence of phobia and relatively low prevalence of schizophrenia. The figure also shows the proportion meeting criteria for a given disorder within the last six months, who had seen either a physician or mental health specialist during that period. These proportions are well under 50%, with the exception of schizophrenia and panic disorder. For depression, which is disabling and highly treatable, less than a third of those with the disorder have received treatment.

3.05.3.2 Defining and Sampling the Population

Defining the target population for the morbidity survey is the first step. The best way to define the target population is not always clear, and different definitions have implications for the ultimate value of the results, as well as the feasibility of the study. A good definition of a target population is an entire nation, such as in the National Comorbidity Survey, or, better yet, a birth cohort of an entire nation, such as the British Perinatal Study, which included all births in Britain during a single week in March of 1958 (Butler & Alberman, 1969). Other studies usually involve compromises of one form or another. The goal of the sampling procedure is that each respondent is selected into the sample with a known, and nonzero,

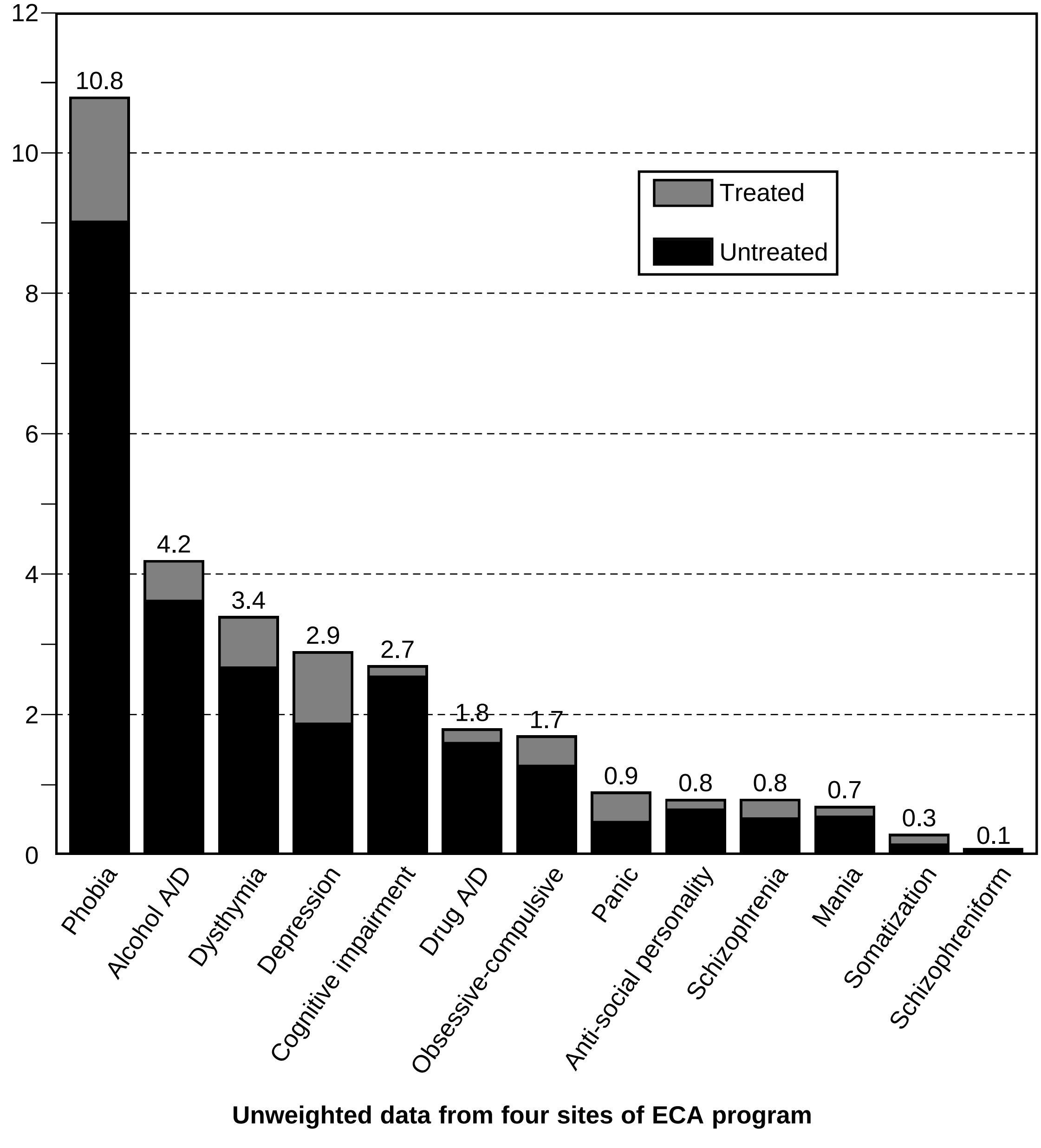

Figure 2 Prevalence of disorder in percent in six months prior to interview.

probability. Strict probabilistic sampling characterizes high-quality epidemiologic surveys, and is a requirement for generalization to the target population.

Most surveys are of the household residing, noninstitutionalized population, where the survey technology for sampling and interviewing individuals is strongest (e.g., Sudman, 1976). The ECA design defined the target population as "normal" residents of previously established catchment areas. Sampling was conducted in two strata, as shown in Figure 3. The household residing population was sampled via area probability methods or household listings provided by utility companies (e.g., as in Sudman, 1976). This stratum included short-stay group quarters such as jails, hospitals, and dormitories. After making a list of the residents in each household, the interviewer asked the person answering the door as to whether there were any other individuals who "normally" resided there but were temporarily absent. "Normally" was defined as the presence of a bed reserved for the individual at the time of the interview. Temporarily absent residents were added to the household roster before the single respondent was chosen randomly. If an individual was selected who was temporarily absent, the interviewer made an appointment for a time after their return, or conducted the interview at their temporary group quarters residence (i.e., in hospital, jail, dormitory, or other place). The ECA sampled the institutional populations separately, by listing all the institutions in the catchment area, as well as all those nearby institutions who admitted residents of the catchment area. Then the inhabitants of each institution were rostered and selected probabilistically. Sampling the institutional population required many more resources, per sampled individual, than the household sample, because each institution had to be approached individually. Inclusion of temporary and long-stay group quarters is important for health services research because many of the group quarters are involved in provision of health services, and because residents of group quarters may be high utilizers. The ultimate result of these procedures in the ECA was that each normal resident of the catchment area was sampled with a known probability.

It is not enough to crisply define a geographic area, because different areas involve different limitations on the generalizability of results. A nationally representative sample, such as the National Comorbidity Survey (NCS; Kessler et al., 1994), may seem to be the best. But how are the results of the national sample to be applied to a given local area, where decisions about services are made? In the ECA the decision was made to select five separate sites of research in order to provide the possibility of replication of results across sites, and to better understand the effects of local variation (Eaton et al., 1981). The ECA target population thus consisted, not of the nation, but rather of an awkward aggregation of catchment areas. Nevertheless, the ECA data were considered as benchmarks for a generation of research (Eaton, 1994) because there was sufficient variation in important sociodemographic variables to allow generalization to other large populations, that is, sufficiently large subgroups of young and old, men and women, married and unmarried, rich and poor, and so forth. Generalization to other target populations, such as Asian-Americans or Native Americans, or to rural areas, was not logical from the ECA. But note that generalization from a national random sample to all rural areas, or to small ethnic groups, would likewise not always be possible. The point is that the target population should be chosen with a view toward later generalization.

3.05.3.3 Sample Size

General population surveys are not efficient designs for rare disorders or unusual patterns of service use. Even for research on outcomes that are not rare, sample sizes are often? larger than one thousand. A common statistic to be estimated is the prevalence, which is a proportion. For a proportion, the precision is affected by the square root of the sample size (Blalock, 1979). If the distribution of the proportion is favorable, say, 30–70%, then a sample size of 1000 produces precision which may be good enough for common sample surveys such as for voter preference. For example, a proportion of 0.50 has a 95% confidence interval from 0.47 to 0.53 with a sample of 1000. For rarer disorders, the confidence interval grows relative to the size of the proportion (i.e., 0.82–0.118 for a proportion of 0.10; 0.004–0.160 for a proportion of 0.01). Often, there is interest in patterns broken down by subpopulations, thus challenging the utility of samples with as few as 1000 individuals. Many community surveys, such as the ECA, have baseline samples in excess of 3000. It is important to estimate the precision of the sample for parameters of interest, and the power of the sample to test hypotheses of interest, before beginning the data collection.

3.05.3.4 Standardization of Data Collection

Assessment in epidemiology should ideally be undertaken with standardized measurement instruments which have known and acceptable

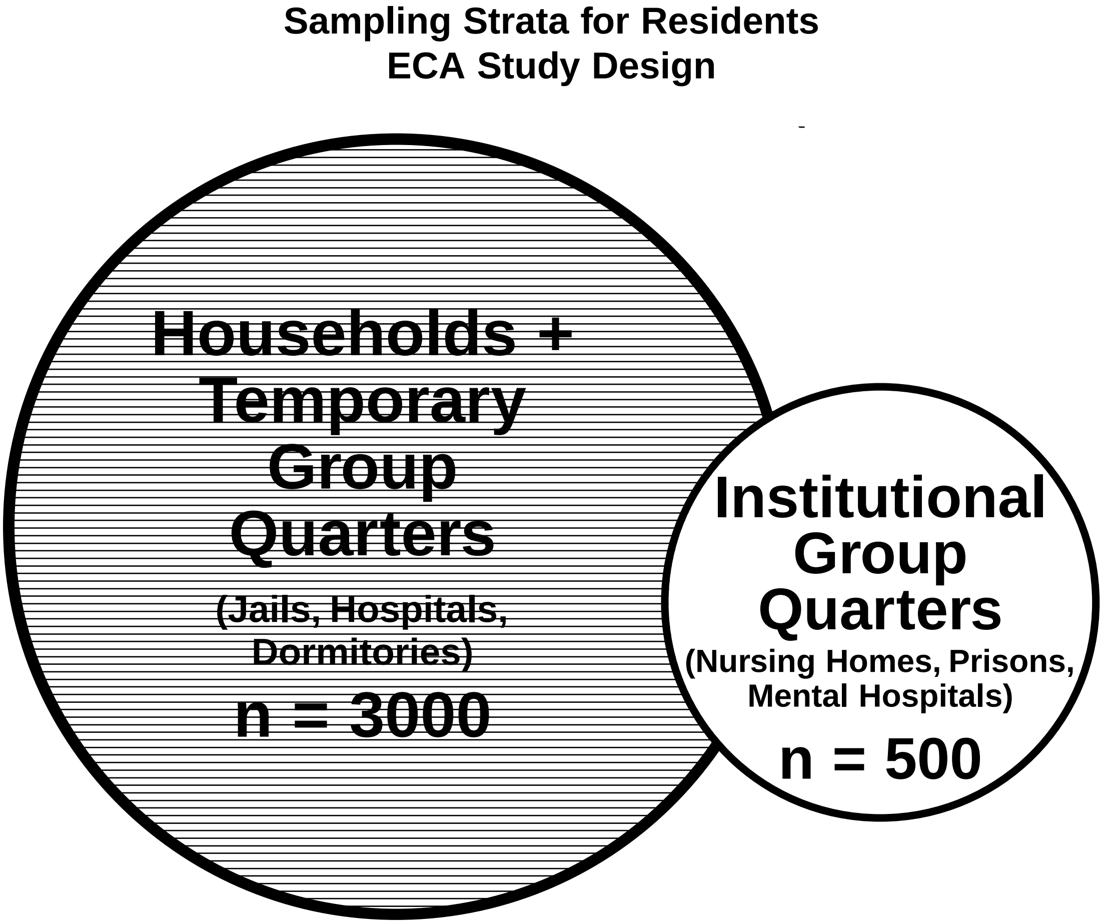
Sampling Strata for Residents
ECA Study Design
Households + Temporary Group Quarters
(Jails, Hospitals, Dormitories)
n = 3000
Institutional Group Quarters
(Nursing Homes, Prisons, Mental Hospitals)
n = 500

Figure 3 Sampling strata for residents: ECA study design.

reliability and validity. In community surveys, and automated record systems, reliable and valid measurement must take place efficiently and in "field conditions." The amount of training for interviewers in household studies depends on the nature of the study. Interviewers in the Longitudinal Study on Aging (LSOA; Kovar, 1992), a well-known cohort study in the United States, received about 1.5 days of training (Fitti & Kovar, 1987), while ECA interviewers received slightly more than a week (Munson et al., 1985). Telephone interviews, such as in the LSOA, can be monitored systematically by recording or listening in on a random basis (as long as the subject is made aware of this possibility), but it is difficult to monitor household interviews, since it cannot be predicted exactly when and where the interview will take place.

The ECA Program involved a somewhat innovative interview called the Diagnostic Interview Schedule, or DIS (Robins, Helzer, Croughan, & Ratcliff, 1981). The DIS was designed to resemble a typical psychiatric interview, in which questions asked are highly dependent on answers already given. For example, if an individual responds positively to a question about the occurrence of panic attacks, a series of questions about that particular response are asked, but if the response to the question on panic is negative, the interviewer skips to the next section. In effect, the interview adapts to the responses of the subject, so that more questions are asked where more information is needed. The high degree of structure in the DIS required more than one week of training, as well as attention to the visual properties of the interview booklet itself. The result was that the interviewer could follow instructions regarding the flow of the interview, and the recording of data, smoothly, so as not to offend or alienate the respondent. Household survey questionnaires are becoming increasingly adaptive and response dependent, because more information can be provided in a shorter amount of time with adaptive interviews. Inexpensive laptop computers will facilitate such adaptive interviewing. Self-administered, computerized admission procedures are becoming more widely disseminated in the health care system, expanding the database, and facilitating the retrieval and linkage of records.

The reliability and validity of measurement are usually assessed prior to beginning a field study. Usually, the assessment involves pilot tests on samples of convenience to determine the optimal order of the questions, the time required for each section, and whether any questions are unclear or offensive. Debriefing subjects in these pilot tests is often helpful. The next step is a test of reliability and validity. Many pretests select samples from populations according to their health or services use characteristics in order to generate enough variation on responses to adequately estimate reliability and validity. In order to economize, such pretests are often conducted in clinical settings. But such pretests do not predict reliability and validity under the "field" conditions of household interviews. The ECA Program design involved reliability and validity assessment in a hospital setting (Robins et al., 1981), and later under field conditions (Anthony et al., 1985). The reliability and validity of the DIS were lower in the household setting.

3.05.4 EXPOSURE AND DISEASE

Two basic research designs in epidemiology are the cohort and the case–control design. These designs are used mostly in addressing Morris' seventh "use" of epidemiology, the search for causes. This is sometimes called the analytic aspect of epidemiology. These designs differ in their temporal orientation to collection of data (prospective vs. retrospective), and in the criteria by which they sample (by exposure or by caseness), as shown in Table 2.

3.05.4.1 Cohort Design

In a cohort design, incidence of disorder is compared in two or more groups which differ in some exposure thought to be related to the disorder (Table 2). The cohort design differs from the morbidity survey in that it is prospective, involving longitudinal follow-up of an identified cohort of individuals. As well as the search for causes, the cohort design addresses Morris' fourth "use" of epidemiology, "estimating individual risks;" and the sixth "use," that is, "completing the clinical picture, especially the natural history." The entire cohort can be sampled, but when a specific exposure is hypothesized, the design can be made more efficient by sampling for intensive measurement on the basis of the putative exposure, for example, children living in an area of toxic exposures, or with parents who have been convicted of abuse, or who have had problems during delivery, and a control group from the general population. Follow-up allows comparison of incidence rates in both groups.

3.05.4.2 Exemplar Study: the British Perinatal Study

An example is the British Perinatal Study, a cohort study of 98% of all births in Great Britain during a single week in March, 1958 (Butler & Alberman, 1969; Done, Johnstone, & Frith, 1991; Sacker, Done, Crow, & Golding, 1995). The cohort had assessments at the ages of 7 and 11 years, and, later, mental hospital records were linked for those entering psychiatric hospitals between 1974 and 1986, by which time the cohort was 30 years old. Diagnoses were made from hospital case notes using a standard system. There was some incompleteness in the data, but not too large, for example, 12 of the 49 individuals diagnosed as "narrow" schizophrenia were excluded because they were immigrants, multiple births, or because they lacked data. It is difficult to say how many individuals in the cohort had episodes of disorder that did not result in hospitalization, but that does not necessarily threaten the results, if the attrition occurred equally for

Table 2 Case–control and cohort studies.

		Case–control study		
		Cases	Controls	
Cohort	Exposed	a	b	$a + b$
Study	Not exposed	c	d	$c + d$

$$\text{Relative risk} = \frac{a/(a+b)}{c/(c+d)}$$

$$\text{Relative odds} = \frac{a/b}{c/d} = \frac{ad}{bc}$$

different categories of exposure. Table 3 shows results for one of 37 different variables related to birth, that is, "exposures," that were available in the midwives' report: birth weight under 2500 g. With n given at the head of the columns in Table 3, the reader can fill in the four cells of a two-by-two table, as in Table 2, for each of the four disorders. For example, the cells, labeled as in Table 2, for narrow schizophrenia, are: a–5; b–706; c–30; d–16 106. The number with the disorder is divided by the number of births to generate the risk of developing the disorder by the time of follow-up: the cumulative incidence. The risks can be compared across those with and without the exposure. In Table 2, the incidence of those with exposure is $a/(a+b)$, and the incidence of those without exposure is $c/(c+d)$. The relative risk is a comparison of the two risks: $[a/(a+b)]/[c/(c+d)]$. For narrow schizophrenia the relative risk is (RR = [5/711]/[30/16 136]), or 3.78.

The relative risk is closely related to causality since it quantifies the association in the context of a prospective study, so the temporal ordering is clear. The relative risk is approximated closely by the relative odds or odds ratio, which does not include the cases in the denominator (i.e., OR = [5/706]/[30/16 106] = 3.80, not 3.78). The odds ratio has many statistical advantages for epidemiology. It is easy to calculate in the two-by-two table by the cross-products ratio (i.e., ad/bc). The odds ratio quantifies the association without being affected by the prevalence of the disorder or the exposure, which is important for the logic of cohort and case–control studies, where these prevalences may change from study to study, depending on the research design. This lack of dependence on the marginal distribution is not characteristic of many measures of association typically used in psychology, such as the correlation coefficient, the difference in proportions, or the kappa coefficient (Bishop et al., 1975). The odds ratio is a standard result of logistic regression, and can be adjusted by other factors such as gender, age, and so forth. The odds ratio for low birth weight and narrow schizophrenia in the British Perinatal Study, as it happens, was 3.9, after adjustment for social class and other demographic variables.

The logic of the cohort study includes assessing the natural history of the disorder, that is, the study of onset and chronicity, in a population context without specific intervention by the researcher. Study of natural history requires repeated follow-up observations. In the British Perinatal Study, there were assessments of the cohort at the ages of 7 and 11. Once the case groups had been defined, it was possible to study their reading and mathematics performance well before hospitalization. Those destined to become schizophrenic had lower reading and mathematics scores at both 7 and 11 years of age (Crow, Done, & Sacker, 1996). Males who would eventually be diagnosed schizophrenic had problems relating to conduct during childhood, while females were anxious, as compared to children who did not end up being hospitalized with a diagnosis of schizophrenia later. Later follow-ups in this study may economize by studying only those who have had onset, and a small random subsample of others, to estimate such factors as the length of episodes, the probability of recurrences, prognostic predictors, and long-term functioning.

3.05.4.3 Case–Control Design

In the case–control study, it is the disease or disorder which drives the logic of the design, and many factors can be studied efficiently as possible causes of the single disorder (Schlesselman, 1982). The case–control study may be the most important contribution of epidemiology to the advancement of public health, because it is so efficient in searching for causes when there is little knowledge. For example, adenocarcinoma of the vagina occurs so rarely, that, prior to 1966, not a single case under the age of 50 years

Table 3 Mental disorder and low birth weight: British Perinatal study.

	Cumulative incidence per 1000		
Disorder	*Low birth weight (n = 727)*	*Normal birth weight (n = 16 812)*	*Odds ratio*
Narrow schizophrenia (n = 35)	7.03	1.86	3.8
Broad schizophrenia (n = 57)	9.82	3.09	3.2
Affective psychosis (n = 32)	8.43	1.61	5.3
Neurosis (n = 76)	4.23	4.51	0.9

Source: Adapted from Sacker et al. (1995).

had been recorded at the Vincent Memorial Hospital in Boston. A time-space clustering of eight cases, all among young women born within the period 1946–1951, was studied (Herbst, Utfeder, & Poskanzer, 1971). There was almost no knowledge of etiology. A case–control study was conducted, matching four controls to each case. The study reported a highly significant ($p < 0.00001$) association between treatment of the mothers with estrogen diethylstilbestrol during pregnancy and the subsequent development of adenocarninoma of the vagina in the daughters. The results led to recommendations to avoid administering stilbestrol during pregnancy. Logistic regression, developed by epidemiologists and used in case–control studies, has distinct advantages over analysis of variance, ordinary least-squares regression, discriminant function analysis, and probit regression, developed by agronomists, psychologists, economists, and educational psychologists, especially when the dichotomous dependent variable has a very skew distribution, that is, is very rare.

3.05.4.4 Exemplar Study: Depression in Women in London

The most convincing demonstration of social factors in the etiology of any mental disorder is *The social origins of depression*, by Brown and Harris (1978). That study used the case–control method to demonstrate the importance of life-event stresses and chronic difficulties as causal agents provoking the onset of depression. The method involved indepth diagnostic and etiologic interviews with a sample of 114 patients and a sample of 458 household residents. The target population is that residing in the Camberwell area in south London.

The analysis presented by Brown and Harris is logical but does not always follow the standard epidemiologic style. Table 4 presents the crucial findings in the typical epidemiologic manner (following Table 2). Two case groups are defined: the 114 women presenting at clinics and hospitals serving the Camberwell area, diagnosed by psychiatrists using a standardized clinical assessment tool called the Present Examination, and 76 women in the community survey (17%) who were depressed at the time of the survey, as judged by a highly trained, nonmedical interviewer using a shortened version of the same instrument. Sixty-one percent of the patient cases (70/114) and 68% of the survey cases (52/76) experienced the "provoking agent" of a severe life event during the year prior to the onset of depression. Twenty percent (76/382) of the healthy controls experienced such an event in the year prior to the interview. Patient cases had 6.4 times the odds of having experienced a life event than the controls (i.e., [70/44]/[76/306]).

The case–control study is very efficient, especially when the disease is rare, because it approximates the statistical power of a huge cohort study with a relatively limited number of controls. For a disease like schizophrenia, which occurs in less than 1% of the population, 100 cases from hospitals or a psychiatric case register can be matched to 300 controls from the catchment area of the hospital. A cohort study of this number would involve measurements on 10 000 persons instead of 400! Since the disease is rare, it may be unnecessary to conduct diagnostic examinations on the group of controls, which would have only a small number of cases. Furthermore, a few cases distributed among the controls would weaken the comparison of cases to controls, generating a conservative bias. The statistical power of the case–control study is very close to that of the analogous cohort study, and, as shown above, the odds ratio is a close estimate of the relative risk. The case–control study can be used to test hypotheses about exposures, but it has the very great ability to make comparisons across a wide range of possible risk factors, and can be useful even when there are very few hypotheses available.

3.05.5 USING AVAILABLE RECORDS

3.05.5.1 Government Surveys

There are many sources of data for psychiatric epidemiologic research which are available to the public. These include statistics from

Table 4 Depression and life events and difficulties in London.

	Cases (patients)		*Cases (survey)*		*Controls (survey)*	
One or more severe event	70	61%	52	68%	76	20%
No severe events	44	39%	24	32%	306	80%
	114	100%	76	100%	382	100%

Source: Brown & Harris (1978).

treatment facilities and from large sample surveys of the populations conducted by a large organization such as national governments. Often the measures of interest to psychologists are only a small part of the survey, but the availability of a range of measures, drawn from other disciplines, can be a strong advantage. In the United States an important source of data is the National Center for Health Statistics (NCHS). For example, the Health and Nutrition Examination Survey (HANES) is a national sample survey conducted by the NCHS involving physical examinations of a sample of the United States population. Its first phase included the Center for Epidemiologic Studies Depression Scale as part of its battery of measures, which included anthropometric measures, nutritional assessments, blood chemistry, medical history, and medical examinations (Eaton & Kessler, 1981; Eaton & McLeod, (1984).

Later phases of the HANES included portions of the DIS. The Health Interview Survey conducts health interviews of the general population of the United States, and includes reports by the respondent of named psychiatric disorders such as schizophrenia, depressive disorder, and so forth. The National Medical Care Utilization and Expenditure Survey (MCUES) is conducted by the NCHS to help understand the health service system and its financing. The MCUES samples records of practitioners from across the nation, and there are several questions on treatment for psychological problems.

Some governments also sponsor national surveys which focus on psychological disorders. The ECA is one such example, although it was not, strictly speaking, a sample of the nation. The later National Comorbidity Survey includes 8058 respondents from a probability sample of the US, with its major measurement instrument being a version of the Composite International Diagnostic Interview (CIDI), a descendant of the DIS used in the ECA surveys (Kessler et al., 1994). The British government conducted a large survey of psychological disorders in a national sample, using a similar instrument to the DIS and CIDI (Meltzer, Gill, Pettirew, & Hindi, 1995a, 1995b). Anonymous, individual level data from the ECA, the British survey, and the NCS are available at nominal cost.

3.05.5.2 Vital Statistics

Governments generally assimilate data on births, deaths, and marriages from states, provinces, or localities, ensuring some minimal degree of uniformity of reporting and creating data files for public use. The mortality files usually list the cause of death, including suicide. These data files are often available from the government at nominal cost.

3.05.5.3 Record-based Statistics and Record Linkage

Statistics originating from treatment facilities can also be put to good use in psychiatric epidemiology. Many early epidemiologic studies used hospital statistics as the numerators in estimating rates, and census statistics in the denominators (Kramer, 1969). The utility of rates estimated in this manner is dependent on the degree to which the clinical disorder is associated with treatment, a stronger association for severe schizophrenia than for mild phobia, presumably. The value of these rates also depends on the relative scarcity or abundance of treatment facilities. In the United States, Medicaid and Medicare files, which include such data as diagnosis and treatment, are available for research use. Many Health Maintenance Organizations maintain data files which could be used in psychological epidemiological research.

Statistics from treatment facilities are enhanced considerably by linkage with other facilities in the same geographic area, serving the same population (Mortensen, 1995). Although the number has declined, there still exist several registers which attempt to record and link together all psychiatric treatment episodes for a given population (ten Horn, Giel, Gulbinar, & Henderson, 1986). Linkage across facilities allows multiple treatment episodes for the same individual (even the same episode of illness) to be combined into one record (so-called "unduplicating"). Linkage across time allows longitudinal study of the course of treatment. Linkage with general health facilities, and with birth and mortality records, provides special opportunities. The best known example of a comprehensive health registry is the Oxford Record Linkage Study (ORLS; Acheson, Baldwin, & Graham, 1987), which links together all hospital treatment episodes in its catchment area. The database of the ORLS consists of births, deaths, and hospital admissions, which is a strong limitation. However, due to the catchmenting aspect of the British National Health Service, the data are not limited to the household-residing population, as is the National Comorbidity Survey, for example, or the LSOA. In automated data collection systems such as the ORLS, the recordation is often done under the auspices of the medical record systems, with data recorded by the physician, such as the diagnosis. It cannot be presumed that physician's diagnosis is "standardized" and therefore more reliable than other interview or record data. In fact, there is significant variation in diagnosis among physicians. Research using measurements and diagnoses recorded by the physician, as in record linkage systems such as the ORLS, should ideally include studies of reliability and validity (e.g., Loffler et al., 1994).

3.05.5.4 Exemplar Study: the Danish Adoption Study

An example of the benefits of record linkage in psychological research is the adoption study of schizophrenia in Denmark (Kety, Rosenthal, Wender, Schulsinger, & Jacobsen, 1975). Familial and twin studies of schizophrenia suggested strongly that schizophrenia was inherited, but these studies were open to the interpretation that the inheritance was cultural, not genetic, because family members are raised together, and identical twins may be raised in social environments which are more similar than the environments of fraternal twins. The Danish Adoption Study ingeniously combined the strategy of file linkage with interviews of cases, controls, and their relatives. In Denmark, each individual receives a number at birth which is included in most registration systems. The registration systems are potentially linkable, after appropriate safeguards and clearances. In the adoption study, three registers were used. First, all 5483 individuals in the county and city of Copenhagen who had been adopted by persons or families other than their biological relatives, from 1924 through 1947, were identified from a register of adoptions (Figure 4). These were linked to the psychiatric case register, wherein it was determined that 507 adoptees had ever been admitted to a psychiatric hospital. From case notes in the hospitals, 34 adoptees who met criteria for schizophrenia (about 0.5% of of the total number of adoptees) were selected, and matched on the basis of age, sex, socioeconomic status of the adopting family, and time with biologic family, or in institution, prior to adoption. The relatives of these 68 cases and controls were identified by linkage with yet another register in Denmark which permits locating families, parents, and children. After allowing for mortality, refusal, and a total of three who were impossible to trace, a psychiatric interview was conducted on 341 individuals (including 12 on whom the interview was not conducted but sufficient information for diagnosis was obtained). Eleven of the 118 biologic relatives of the index adoptees were schizophrenic (9%), vs. one in the 35 relatives of adoptive families of the index adoptees (3%), and three of the 140 biologic relatives of the control adoptees (2%). The findings for schizophrenia spectrum disorders (including uncertain schizophrenia and schizoid personality) also show a pattern consistent only with genetic inheritance (26/118 in the biologic relatives of index cases, or 22%, vs. 16/140 biologic relatives of control adoptees, or 14%).

3.05.6 PREVENTIVE TRIALS

3.05.6.1 The Logic of Prevention

Some of the earliest epidemiologists used interventions in the community to stop an epidemic or to gather information about the causes of diseases in the population. This is sometimes called the experimental aspect of epidemiology. The best known experiment is that conducted by Snow in the cholera epidemic in London. Snow's work exemplifies epidemiologic principles in several ways (Cooper & Morgan, 1973). It was widely believed that cholera was spread through the air, in a miasma, leading many to seek safety by retreating to rural areas. Snow showed with ecologic data in London that areas serviced by a water company taking water from upstream in the Thames had lower rates of cholera than areas serviced by a company taking water from downstream. This ecologic comparison suggested that cholera was borne by water. In the context of a single cholera epidemic, Snow identified individual cases of cholera, showing that they tended to cluster around a single water pump at Broad Street. He further showed that many exceptional cases, that is, cases of cholera residing distant from the pump, had actually drawn or drank water from that pump (e.g., on their way home from work). His action to remove the handle of the pump, which roughly coincided with the termination of the epidemic, is regarded as an early instance of experimental epidemiology.

In epidemiology, as with medicine generally, intervention is valued if it is effective, regardless of whether the causation of the disease is understood or not. Table 5 shows examples of preventive measures which were implemented well prior to knowledge of the cause, discovered much later in many cases. This logic leads to experimentation even in the absence of causal information.

Various conceptual frameworks have been used to organize the area of prevention research. The Commission on Chronic Illness (1957) divided prevention into three types, dependent on the stage of the disease the intervention was designed to prevent or treat. Prior to onset of the disease, the prevention was primary, and its goal was to reduce incidence; after disease onset, the intervention was directed at promoting

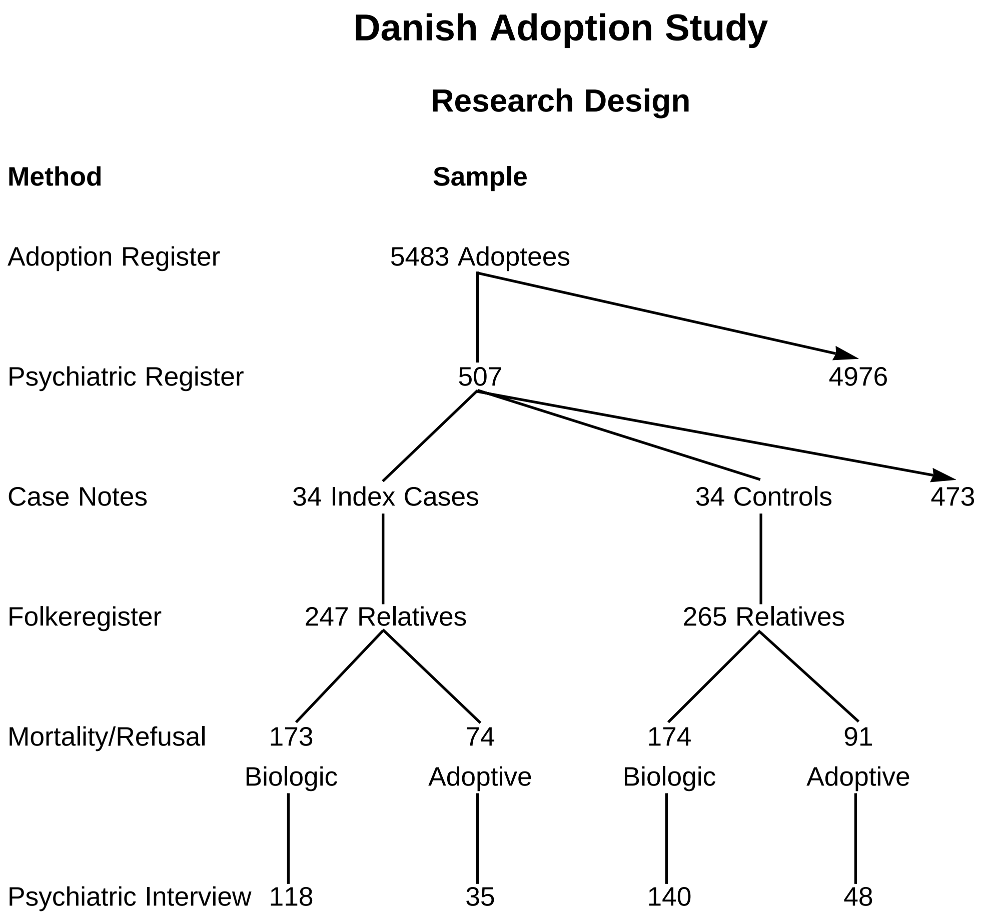
Danish Adoption Study
Research Design
Method
Sample
Adoption Register
5483 Adoptees
Psychiatric Register
507
4976
Case Notes
34 Index Cases
34 Controls
473
Folkeregister
247 Relatives
265 Relatives
Mortality/Refusal
173 Biologic
74 Adoptive
174 Biologic
91 Adoptive
Psychiatric Interview
118
35
140
48

Results

Schizophrenic:	11	1	3	2
Spectrum	26	3	16	5
Normal	81	31	121	41

Figure 4 Danish adoption study: research design and results.
Source: Kety, Rosenthal, Wender, Schulsinger, and Jacobsen (1975).

recovery, and its goal was to reduce prevalence, so-called secondary prevention. Tertiary prevention was designed to prevent impairment and handicap which might result from the disease. The Institute of Medicine (Mrazek & Haggerty, 1994) integrated this framework into that of Gordon (1983), who directed attention at the population to which the preventive intervention was directed: universal preventions at the entire general population; targeted interventions at subgroups at high risk of development of disorder; and indicated interventions, directed at individuals who have already manifest signs and symptoms of the disease. The results of integrating the two frameworks are shown in Figure 5. Curative medicine generally operates in the area of secondary prevention, with indicated interventions. Rehabilitative medicine is in the area of tertiary prevention. Primary, universal, and targeted interventions are the province of experimental epidemiology. Prominent examples of universal interventions in the area of epidemiology are various smoking cessation campaigns, and such studies as the Stanford Five-City Project, which was designed to reduce risk of heart disease by lowering levels of several associated risk factors (Farquhar et al., 1990).

Table 5 Knowledge of prevention and etiology.

	Prevention		*Etiology*		
Disease	*Discoverer*	*Year*	*Agent*	*Discoverer*	*Year*
Scurvy	Lind	1753	Ascorbic acid	Szent-Gyorgi	1928
Pellagra	Casal	1755	Niacin	Goldberger	1924
Scrotal cancer	Pott	1775	Benzopyrene	Cook	1933
Smallpox	Jenner	1798	Orthopoxvirus	Fenner	1958
Puerperal fever	Semmelwies	1847	Streptococcus	Pasteur	1879
Cholera	Snow	1849	Vibrio cholerae	Koch	1893
Bladder cancer	Rehn	1895	2-Napththylamine	Hueper	1938
Yellow fever	Reed	1901	Flavirus	Stokes	1928
Oral cancer	Abbe	1915	*N*-Nitrosonornicotine	Hoffman	1974

Source: Wynder (1994).

3.05.6.2 Attributable Risk

From among many possibilities, how should interventions be selected? Epidemiology provides a helpful tool in the form of the Population Attributable Risk, sometimes called the Attributable Fraction or the Etiologic Fraction (Lilienfeld & Stolley, 1994, p. 202). The attributable risk is the maximum estimate of the proportion of the incidence of disease that would be prevented if a given risk factor were eliminated. For a given disease, the attributable risk combines information from the relative risk for a given exposure with the prevalence of the exposure in the population. The formula for attributable risk is:

$$\text{Attributable Risk} = \frac{\text{P(RR} - 1)}{\text{P(RR} - 1) + 1}$$

where:
P = Prevalence of Exposure, and
RR = Relative Risk of Exposure to Disease

The relative risk can be estimated from a cohort study, as described above, and the prevalence of the exposure can be estimated from a separate survey. A simple case–control study also provides the information for attributable risk, under certain conditions. The relative risk is approximated by the relative odds, as discussed above. If the controls are selected from the general population, the level of exposure can be estimated from them. For example, case–control studies of smoking and lung cancer in the early 1950s showed that the odds of dying from lung cancer were about 15 times higher for smokers as for nonsmokers. About half the population smoked, leading to the attributable risk estimate of about 80%. In the United States, this meant that about 350 000 of the 400 000 annual deaths due to lung cancer would not occur if smoking were eliminated totally. In the situation of many possible risk factors, the attributable risk is a tool which helps prioritize them.

Epidemiologic cohort studies can provide information which may help to stage the intervention at the most appropriate time. The intervention should occur before most onsets in the population have taken place, but not so far before onset that the effects of the intervention wash out before the process of disease development has begun. The appropriate stage for prevention in many situations would appear to be that of precursors, in which there are subgroups which can be identified at high risk, but before the disease prodrome has started (Eaton et al., 1995).

3.05.6.3 Developmental Epidemiology

Epidemiologists are gradually becoming aware of the need for life course perspective (Kellam & Werthamer-Larson, 1986). In social and psychological research on human beings, cohort studies may suggest regularities in human development which can be considered etiologic clues. The clues consist of temporal sequences of various behaviors over several years, with only a probabilistic connection between each behavior. The behaviors may have a common antecedent, such as a genetic background, or they may depend on the sequence of environments that the individual experiences, and the interaction between environments, genetic traits, and habits formed by the history of prior behaviors. The epidemiologic notion of exposures, developed in the context of infectious agents and toxins, is too simple to be useful in these multicausal, developmental sequences. For example, aggressive behavior by boys in early elementary school years is followed by conduct problems in middle

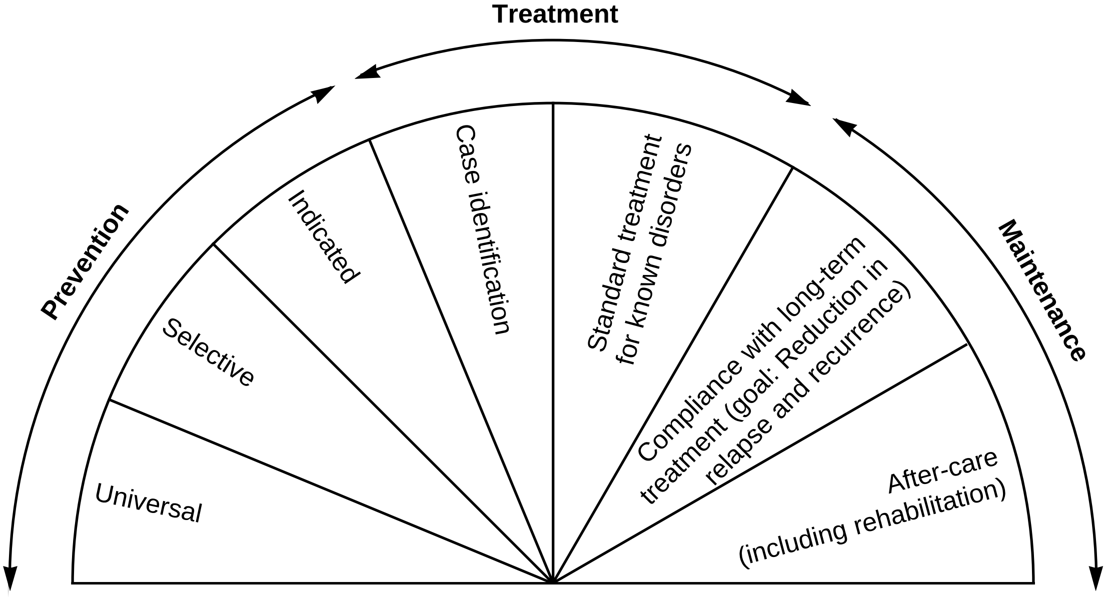
Treatment
Prevention
Maintenance
Universal
Selective
Indicated
Case identification
Standard treatment for known disorders
Compliance with long-term treatment (goal: Reduction in relapse and recurrence)
After-care (including rehabilitation)

Figure 5 The mental health intervention spectrum for mental disorders.

school; later behaviors such as smoking cigarettes, having sexual relations at an early age, and drinking alcohol, in high school, and, eventually, the diagnosis of antisocial personality disorder as an adult. Which of these are essential causes, and which simply are associated behaviors which have no influence on the causal chain to the outcome (in this example, the diagnosis of antisocial personality disorder)? Since the chain of causes is probabilistic and multicausal, the notion of attributable risk (discussed above) is too simplistic, and it is unlikely that any estimate of attributable risk would be even moderate in size. A preventive intervention trial with random or near-random assignment to conditions which manipulate a putative cause may be the only way to generate an understanding of the causal chain. The preventive trial serves the traditional purpose in epidemiology of testing the efficacy of the intervention, but it also serves as an examination of the developmental pathway leading to disorder, for building and testing theories.

3.05.6.4 Exemplar Study: the Johns Hopkins Prevention Trial

An example of a preventive intervention trial is the Johns Hopkins Prevention Trial (Kellam, Rebek, Ialongo, & Mayer, 1994). The trial was designed to examine developmental pathways generally occurring after two important characteristics which could be identified in first graders: success at learning (especially, learning to read) and aggressive behavior. Table 6 shows the design of the trials, using the notation of experimental and quasi-experimental design in psychology (Campbell & Stanley, 1971). The population for the trial was carefully defined in epidemiologic terms to include all first graders in public schools in an area of Baltimore. The Baltimore City Public School system was an active participant in the research. The 19 schools were divided into five levels, with three or four schools per level, based on the socioeconomic characteristics of the areas they served. At each of the five levels there was a control school and two experimental schools. For the study of aggressive behavior, 153 children were assigned at random to eight separate classrooms in which the teacher used a special classroom procedure called the "Good Behavior Game," which had been shown to reduce the level of aggressive behavior in classrooms. Nine classrooms with 163 children received an intervention to improve reading mastery, and there were 377 children who were in control classrooms. The control classrooms consisted of classrooms in the same school as the experimental, but who did not receive the intervention; and classrooms in schools in the same area of Baltimore, where there were no interventions at all. One outcome of the Good Behavior Game Trial was that the median level of aggressive behavior went down during the period of the trial for those in the experimental group. For controls, the level of aggressive behavior was relatively stable. The most impressive result was that aggressive behavior continued to decline in the experimental group, even after discontinuation of the intervention in third grade, at least through the spring of sixth grade.

Table 6 Intervention and assessment for Johns Hopkins Preventive Trials in 19 public elementary schools.

	Number of classrooms	*Number of students*	*Grade 1 1985–1986 F/S*	*Grade 2 1986–1987 F/S*	*Grade 3 1987–1988 F/S*	*Grade 4 1988–1989 F/S*	*Grade 5 1989–1990 F/S*	*Grade 6 1990–1991 F/S*
Good behavior game	8	153	RO/XO	XO/O	/O	/O	/O	/O
Mastery learning	9	163	RO/XO	XO/O	/O	/O	/O	/O
Controls	24	377	RO/O	O/O	/O	/O	/O	/O

Source: Kellam et al. (1994). F = fall; S = spring; O = observation; X = intervention; R = random assignment.

3.05.7 BIAS

Epidemiologic research entails bias, and it is beneficial to anticipate bias in designing research studies. Common biases in epidemiology take a slightly different form than in typical psychological research (e.g., as enumerated in Campbell & Stanley, 1971), but the principle of developing a language to conceptualize bias, and to attempts to anticipate, eliminate, or measure bias is common to both disciplines. In this section five basic types of bias are considered: sampling bias, divided into three subtypes of treatment bias, prevalence bias, and response bias; and measurement bias, divided into subtypes of diagnostic bias and exposure bias.

3.05.7.1 Sampling Bias

Sampling bias arises when the sample studied does not represent the population to which generalization is to be made. An important type of sampling bias is treatment bias. The case–control design often takes advantage of treatment facilities for finding cases of disorder to compare with controls. With severe diseases which invariably lead to treatment, there is less bias than for disorders which are less noticeable, distressing, or impairing to the individual. As shown in Figure 2, data from the ECA Program indicate that among the mental disorders, only for schizophrenia and panic disorder are as many as half the current cases under treatment. In 1946, Berkson showed that, where the probability of treatment for a given disorder is less than 1.0, cases in treatment overrepresented individuals with more than one disorder, that in, comorbid cases. In studying risk factors for disorder with the case–control design, where cases are found through clinics, the association revealed in the data may be an artifact of the comorbidity, that is, exposure x may appear to be associated with disease y, the focus disorder of the study, but actually related to disease z (perhaps not measured in the study). This type of bias is so important in epidemiologic research, especially case–control studies, it is termed “Berkson's bias,” or sometimes “treatment bias.” The existence of this bias is possibly one factor connected to little use of the case–control design in studies of psychiatric disorders.

Another type of sampling bias very common in case–control studies arises from another type of efficiency in finding cases of disorder, that in, the use of prevalent cases. This is sometimes termed prevalence bias, or the clinician's illusion (Cohen & Cohen, 1984). The ideal is to compare relative risks among exposed and nonexposed populations, which entails comparison of incidence rates, as discussed above. But the rate of incidence is often too low to generate sufficient cases for analysis, and necessitates extending the data collection over a period of months or years during which new cases are collected. Prevalent cases are easier to locate, either through a cross-sectional survey, or, easier yet, through records of individuals currently under treatment. The problem with study of current cases is that the prevalence is a function of the incidence and chronicity of the disorder, as discussed above. Association of an exposure with the presence or absence of disease mixes up influences on incidence with influences on chronicity. Influences on chronicity can include treatment. For example, comparing the brain structure of prevalent cases of schizophrenia to controls showed differences in the size of the ventricles, possibly an exciting clue to the etiology of schizophrenia. But a possible bias is that phenothiazine or other medication, used to treat schizophrenia, produced the changes in brain structure. Later studies of brain structure had to focus on schizophrenics who had not been treated in order to eliminate this possibility. The study of depression among women in London is possibly subject to prevalence bias since cases were selected on the basis of presentation to treatment, and controls via a cross-sectional (prevalence) survey. Thus, it may be that provoking agents such as life events contribute only to the recurrence of episodes of depression, not necessarily to their initial onset.

Response bias is a third general threat to the validity of findings from epidemiologic research. The epidemiologic design includes explicit statements of the population to which generalization is sought, as discussed above. The sampling procedure includes ways to designate individuals thought to be representative of that population. After designation as respondents, before measurements can be taken, some respondents become unavailable to the research, usually through one of three mechanisms: death, refusal, or change of residence. If these designated-but-not-included respondents are not a random sample, there will be bias in the results. The bias can result from differences in the distribution of cases vs. noncases as to response rate, or to differences in distribution of exposures, or both. In cross-sectional field surveys (Von Korff et al., 1985) and in follow-up surveys (Eaton, Anthony, Tepper, & Dryman, 1992), the response bias connected to psychopathology has not been extremely strong. Persons with psychiatric diagnoses are not more likely to refuse to participate, for example. Persons with cognitive impairment are more likely to die during a follow-up interval, and persons with antisocial personality disorder, or abuse of illegal drugs, are more likely to change address and be difficult to locate. Differential response bias is particularly threatening to studies with long follow-up periods, such as the British Perinatal Study and the Danish Adoption Study, since there was sizable attrition in both studies.

3.05.7.2 Measurement Bias

Bias in measurement is called invalidity in psychology, and this term is also used in epidemiology. But the study of validity in epidemiology has been more narrowly focused than in psychology. The concentration in epidemiology has been on dichotomous measures, as noted above. The medical ideology has ignored the notion that concepts for disease and pathology might actually be conventions of thought. Where the psychometrician takes as a basic assumption that the true state of nature is not observable, the epidemiologist tends to think of disease as a state whose presence is essentially knowable. As a result, discussion of face validity, content validity, or construct validity are rare in epidemiologic research. Instead the focus has been on sensitivity and specificity, which are two aspects of criterion validity (also a term not used much in epidemiologic research).

Sensitivity is the proportion of true cases that are identified as cases by the measure (Table 7). Specificity is the proportion of true noncases that are identified as noncases by the measure. In psychological research, validity is often estimated with a correlation coefficient, but this statistic is inappropriate because the construct of disease is dichotomous and differences in rarity of the disorder will constrain the value of the correlation coefficient. Furthermore, use of sensitivity and specificity has the advantage over correlation that it forces quantification of both types of error, that is, false-positive and false-negative error. These errors, and the calibration of the measurement to adapt to them, depend heavily on the prevalence of the disorder, on the importance of locating cases, and on the expense of dealing with those cases that are located. Choice of a threshold to define disorder as present or absent has important consequences, and is aided by detailed study of the effects of different thresholds (sometimes termed response operating characteristic (ROC) analysis, as in Murphy, 1987). There are simple procedures for correcting prevalence estimates according to the sensitivity and specificity of the measurement (Rogan & Gladen, 1978). Procedures for correcting measures of association, such as the odds ratio, are more complicated since they depend on the precise study design.

Bias also exists in the measurement of exposure. The epidemiologist's tendency to categorical measurement leads to the term "misclassification" for this type of bias. A well-known example is the study of Lilienfeld and Graham (1958: cited in Schlesselman, 1982, pp. 137–138), which compared male patients' declarations as to whether they were circumcised to the results of a doctor's examination. Of the 84 men judged to be circumcised by the doctor (the "gold standard" in this situation), only 37 reported being circumcised (44% sensitivity). Eighty-nine of the 108 men who were not circumcised in the view of the doctor reported not being circumcised (82% specificity). In psychological research, the exposures are often subjective psychological happenings, such as emotions, or objective events recalled by the subject, such as life events, instead of, for example, residence near a toxic waste dump, as might be the putative exposure in a study of cancer. The importance of psychological exposures, or subjective reporting of exposures, threatens the validity of the case–control design in psychiatric epidemiology, and may be one reason it has been used so little. The cases in a case–control study by definition have a disease or problem which the controls do not. In any kind of case–control study where the measure of exposure is based on recall by the subjects, the cases may be more likely than controls to recall exposures, because they wish to explain the occurrence of the disease. In the study of depression in London, depressed women may be more likely to recall a difficult life event that happened earlier in their lives, because of their current mood, than women who are not depressed; or depressed women may magnify the importance of a given event which actually did occur. These problems of recall raise the importance of strictly prospective designs in psychological epidemiological research.

3.05.8 CONCLUSION: THE FUTURE OF PSYCHOLOGY IN EPIDEMIOLOGY

Epidemiology is developing rapidly in ways that will promote effective collaboration with psychology. The traditional interest of psychologists in health as well as illness links up to epidemiologists' growing interest in the development of host resistance. The traditional interest of psychologists in phenomenology is beginning to link up with epidemiologists'

Table 7 Sensitivity and specificity.

		True disease status		
		Present	Absent	
Test results	Present	a (True-positives)	b (False-positives)	$a + b$
	Absent	c (False-negatives)	d (True-negatives)	$c + d$
		$a + c$	$b + d$	

Sensitivity = $a / (a + c)$
Specificity = $d / (b + d)$

growing interest in measurement and complex nosology. In statistics, techniques of multivariate analysis are being adapted increasingly well to the developmental perspective. Developments in the field of physiological and biological measurement, including nearly noninvasive assessments of DNA, hormones, and brain structure and functioning, have led to concepts like "molecular epidemiology," which will lead to vast increases in knowledge about how psychological functioning affects diseases across the spectrum of mental and physical illness. Many of these measures are efficient enough to be applied in the context of a large field survey. In a few decades, the concept of the "mind–body split" may be an anachronism. Finally, the increasing use of laptop computers for assistance in assessment of complex human behaviors and traits will increase the utility of questionnaires and interview data. For all these reasons psychologists should look forward to increasingly productive collaborations in the field of epidemiology.

ACKNOWLEDGMENTS

Production of this paper was supported by NIMH grants 47447, and by the Oregon Social Learning Center. The author is grateful to Nina Schooler and Ray Lorion for comments on early drafts.

3.05.9 REFERENCES

Acheson, E. D., Baldwin, J. A., & Graham, W. J. (1987). *Textbook of medical record linkage*. New York: Oxford University Press.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anthony, J. C., Eaton, W. W., & Henderson, A. S. (1995). Introduction: Psychiatric epidemiology (Special Issue of Epidemiologic Reviews). *Epidemiologic Reviews, 17*(1), 1–8.

Anthony, J. C., Folstein, M. F., Romanoski, A., Von Korff, M., Nestadt, G., Chahal, R., Merchant, A., Brown, C. H., Shapiro, S., Kramer, M., & Gruenberg, E. M. (1985). Comparison of the lay Diagnostic Interview Schedule and a standardized psychiatric diagnosis: Experience in Eastern Baltimore. *Archives of General Psychiatry, 42,* 667–675.

Berkson, J. (1946). Limitations of the application of fourfold table analysis to hospital data. *Biometrics, 2,* 47–53.

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis: Theory and practice*. Cambridge, MA: MIT Press.

Blalock, H. M. (1979). *Social statistics* (2nd ed.). New York: McGraw-Hill.

Brown, G. W., Birley, J. L. T., & Wing, J. K. (1972). Influence of family life on the course of schizophrenic disorders: a replication. *British Journal of Psychiatry, 121,* 241–258.

Brown, G. W., & Harris, T. (1978). *The social origins of depression: A study of psychiatric disorder in women*. London: Tavistock.

Butler, N. R., & Alberman, E. D. (1969). *Perinatal Problems: The Second Report of the 1958 British Perinatal Survey*. Edinburgh, Scotland: Livingstone.

Campbell, D. T., & Stanley, J. C. (1971). *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.

Cohen, P., & Cohen, J. (1984). The Clinician's Illusion. *Archives of General Psychiatry, 41,* 1178–1182.

Commission on Chronic Illness (1957). *Chronic illness in the United States*. Cambridge, MA: Harvard University Press.

Cooper, B., & Morgan, H. G. (1973). *Epidemiological psychiatry*. Springfield, IL: Charles C. Thomas.

Cross-National Collaborative Group (1992). The changing rate of major depression: cross-national comparisons. *Journal of the American Medical Association, 268,* 3098–3105.

Crow, T. J., Done, D. J., & Sacker, A. (1996). Birth cohort study of the antecedents of psychosis: Ontogeny as witness to phylogenetic origins. In H. Hafner & W. F. Gattaz (Eds.), *Search for the causes of schizophrenia* (Vol. 3, pp. 3–20). Heidelberg, Germany: Springer.

Done, D. J., Johnstone E. C., Frith, C. D., et al. (1991). Complications of pregnancy and delivery in relation to psychosis in adult life: data from the British perinatal mortality survey sample. *British Medical Journal, 302,* 1576–1580.

Eaton, W. W. (1994). The NIMH epidemiologic catchment area program: Implementation and major findings. *International Journal of Methods in Psychiatric Research, 4,* 103–112.

Eaton, W. W. (1995). Studying the natural history of psychopathology. In M. Tsuang, M. Tohen, & G. Zahner (Eds.), *Textbook in psychiatric epidemiology* (pp. 157–177). New York: Wiley Lisa, Inc.

Eaton, W. W., Anthony, J. C., Tepper, S., & Dryman, A. (1992). Psychopathology and Attrition in the Epidemiologic Catchment Area Surveys. *American Journal of Epidemiology, 135,* 1051–1059.

Eaton, W. W., Badawi, M., & Melton, B. (1995). Prodromes and precursors: Epidemiologic data for primary prevention of disorders with slow onset. *American Journal of Psychiatry, 152*(7), 967–972.

Eaton, W. W., Holzer, C. E., Von Korff, M., Anthony, J. C., Helzer, J. E., George, L., Burnam, A., Boyd, J. H., Kessler, L. G., & Locke, B. Z. (1984). The design of the Epidemiologic Catchment Area surveys: the control and measurement of error. *Archives of General Psychiatry, 41,* 942–948.

Eaton, W. W., & Kessler, L. G. (1981). Rates of symptoms of depression in a national sample. *American Journal of Epidemiology, 114,* 528–538.

Eaton, W. W., Kramer, M., Anthony, J. C., Chee, E. M. L., & Shapiro, S. (1989). Conceptual and methodological problems in estimation of the incidence of mental disorders from field survey data. In B. Cooper & T. Helgason (Eds.), *Epidemiology and the prevention of mental disorders* (pp. 108–127). London: Routledge.

Eaton, W. W., & McLeod, J. (1984). Consumption of coffee or tea and symptoms of anxiety. *American Journal of Public Health, 74,* 66–68.

Eaton, W. W., & Muntaner, C. (1996). Socioeconomic stratification and mental disorder. In A. V. Horwitz & T. L. Scheid (Eds.), *Sociology of mental health and illness*. New York: Cambridge University Press.

Eaton, W. W., Regier, D. A., Locke, B. Z., & Taube, C. A. (1981). The Epidemiologic Catchment Area Program of the National Institute of Mental Health. *Public Health Report 96,* 319–325.

Eaton, W. W., Weissman, M. M., Anthony, J. C., Robins, L. N., Blazer, D. G., & Karno, M. (1985). Problems in the definition and measurement of prevalence and incidence of psychiatric disorders. In W. W. Eaton & L. G. Kessler (Eds.), *Epidemiologic Field Methods in Psychiatry: The NIMH Epidemiologic Catchment Area Program* (pp. 311–326). Orlando, FL: Academic Press.

Farquhar, J. W., Fortmann, S. P., Flora, J. A., Taylor, C. B., Haskell, W. L., Williams, P. T., Maccoby, N., & Wood, P. D. (1990). Effects of communitywide education on cardiovascular disease risk factors: The Stanford Five-City Project. *Journal of the American Medical Association, 264*(3), 359–365.

Fitti, J. E., & Kovar, M. G. (1987). *The Supplement on Aging to the 1984 National Health Interview Survey. Vital and Health Statistics*. Washington, DC: Government Printing Office.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: Wiley.

Foege, W. H. (1996). Alexander D. Langmuir: His impact on public health. *American Journal of Epidemiology, 144*(8) (Suppl.), S11–S15.

Gordon, R. (1983). An operational classification of disease prevention. *Public Health Reports, 98,* 107–109.

Herbst, A. L., Ulfelder, H., & Poskanzer, D. C. (1971). Adenocarcinoma of the vagina. Association of maternal stilbestrol therapy with tumor appearance in young women. *The New England Journal of Medicine, 284*(16), 878–881.

Kellam, S. G., & Werthamer-Larsson, L. (1986). Developmental epidemiology: a basis for prevention. In M. Kessler & S. E. Goldston (Eds.), *A Decade of Progress in Primary Prevention* (pp. 154–180). Hanover, NH: University Press of New England.

Kellam, S. G., Rebok, G. W., Ialongo, N., & Mayer L.S. (1994). The course and malleability of aggressive behavior from early first grade into middle school: Results of a developmental epidemiologically-based preventive trail. *Journal of Child Psychology and Psychiatry, 35*(2), 259–281.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshelman, S., Wittchen, H., & Kendler, K. S. (1994). Lifetime and 12-month prevalence of *DSM-III-R* psychiatric disorders in the United States. *Archives of General Psychiatry, 51,* 8–19.

Kety, S. S., Rosenthal, D., Wender, P. H., Schulsinger, F., & Jacobsen, B. (1975). Mental illness in the biological and adoptive families of adopted individuals who have become schizophrenic: a preliminary report based on psychiatric interviews. In R. R. Fieve, D. Rosenthal, & H. Brill (Eds.), *Genetic Research in Psychiatry* (pp. 147–165). Baltimore, MD: Johns Hopkins University Press.

Kleinbaum, D. G., Kupper, L. L., & Morgenstern, H. (1982). *Epidemiologic Research. Principles and Quantitative Methods*. Belmont, CA: Lifetime Learning.

Klerman, G. L., & Weissman, M. M. (1989). Increasing rates of depression. *JAMA, 261,* 2229–2235.

Kovar, M. (1992). *The Longitudinal Study of Aging: 1984–90*. Hyattsville, MD: US Department of Health and Human Services.

Kramer, M. (1957). A discussion of the concepts of incidence and prevalence as related to epidemiologic studies of mental disorders. *American Journal of Public Health & Nation's Health, 47*(7), 826–840.

Kramer, M. (1969). *Applications of Mental Health Statistics: Uses in mental health programmes of statistics derived from psychiatric services and selected vital and morbidity records*. Geneva: World Health Organization.

Kramer, M., Von Korff, M., & Kessler, L. (1980). The lifetime prevalence of mental disorders: estimation, uses and limitations. *Psychological Medicine, 10,* 429–435.

Lawless, J. F. (1982). *Statistical models and methods for lifetime data* (Wiley Series in Probability and Mathematical Statistics). New York: Wiley.

Lilienfeld, A., & Lilienfeld, D. (1980). *Foundations of epidemiology*. New York: Oxford University Press.

Lilienfeld, D. E., & Stolley, P. D. (1994). *Foundations of epidemiology*. New York: Oxford University Press.

Loffler, W., Hafner, H., Fatkenheuer, B., Maurer, K., Riecher-Rossler, A., Lutzhoft, J., Skadhede, S., Munk-Jorgensen, P., & Stromgren, E. (1994). Validation of Danish case register diagnosis for schizophrenia. *Acta Psychiatrica Scandinavica, 90,* 196–203.

Mausner, J. S., & Kramer, S. (1985). *Mausner & Bahn epidemiology: An introductory text* (2nd ed.). Philadelphia: Saunders.

Meltzer, H., Gill, B., Petticrew, M., & Hinds, K. (1995a). *Physical complaints, service use and treatment of adults with psychiatric disorders*. London: Her Majesty's Stationery Office.

Meltzer, H., Gill, B., Petticrew, M., & Hinds, K. (1995b). *Prevalence of psychiatric morbidity among adults living in private households*. London: Her Majesty's Stationery Office.

Morris, J. N. (1975). *Use of epidemiology* (3rd ed.). Edinburgh: Churchill Livingstone.

Mortensen, P. B. (1995). The untapped potential of case registers and record-linkage studies in psychiatric epidemiology. *Epidemiologic Reviews, 17*(1), 205–209.

Mrazek, P. J., & Haggerty, R. J. (1994). *Reducing risks for mental disorders*. Washington, DC: National Academy Press.

Munson, M. L., Orvaschel, H., Skinner, E. A., Goldring, E., Pennybacker, M., & Timbers, D. M. (1985). Interviewers: Characteristics, training and field work. In W. W. Eaton & L. G. Kessler (Eds.), *Epidemiologic Field Methods in Psychiatry: The NIMH Epidemiologic Catchment Area Program* (pp. 69–83). Orlando, FL: Academic Press.

Murphy, J. M., Berwick, D. M., Weinstein, M. C., Borus, J. F., Budman, S. H., & Klerman, G. L. (1987). Performance of Screening and Diagnostic Tests: Application of receiver operating characteristic analysis. *Archives of General Psychiatry, 44,* 550–555.

Robins, L. N., Helzer, J. E., Croughan, J., & Ratcliff, K. S. (1981). National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics, and validity. *Archives of General Psychiatry, 38,* 381–389.

Robins, L. N., Helzer, J. E., Weissman, M. M., Orvaschel, H., Gruenberg, E. M., Burke, J. D., & Regier, D. A. (1984). Lifetime prevalence of specific psychiatric disorders in three sites. *Archives of General Psychiatry, 41,* 949–958.

Robins, L. N., & Regier, D. A. (Eds.) (1991). *Psychiatric disorders in America: The epidemiologic catchment area study*. New York: Free Press.

Rogan, W. J., & Gladen, B. (1978). Estimating Prevalence from the Results of a Screening Test. *American Journal of Epidemiology, 107,* 71–76.

Sacker, A., Done, J., Crow, T. J., & Golding, J. (1995). Antecedents of schizophrenia and affective illness obstetric complications. *British Journal of Psychiatry, 166,* 734–741.

Schlesselman, J. J. (1982). *Case–control studies: Design, conduct, analysis*. New York: Oxford University Press.

Shapiro, S., Skinner, E. A., Kessler, L. G., Von Korff, M., German, P. S., Tischler, G. L., Leaf, P. J., Benham, L., Cottler, L., & Regier, D. A. (1984). Utilization of health and mental health services, three epidemiologic catchment area sites. *Archives of General Psychiatry, 41,* 971–978.

Simon, G. E., & Vor Korff, M. (1995). Recall of Psychiatric History in Cross-sectional Surveys: Implications for Epidemiologic Research. *Epidemiologic Reviews, 17*(1), 221–227.

Sudman, S. (1976). *Applied sampling*. New York: Academic Press.

ten Horn, G. H., Giel, R., Gulbinat, W. H., & Henderson, J. H. (Eds.) (1986). *Psychiatric case registers in public health—A worldwide inventory 1960–1985*. Amsterdam: Elsevier Science.

Tsuang, M., Tohen, M., & Zahner, G. (1995). *Textbook in psychiatric epidemiology*. New York: Wiley-Liss.

Tyrer, P. (1985). Neurosis divisible? *Lancet, 1,* 685–688.

Von Korff, M., Cottler, L., George, L. K., Eaton, W. W., Leaf, P. J., & Burnam, A. (1985). Nonresponse and nonresponse bias in the ECA surveys. In W. W. Eaton & L. G. Kessler (Eds.), *Epidemiologic field methods in psychiatry: The NIMH Epidemiologic Catchment Area Program* (pp. 85–98). Orlando, FL: Academic Press.

Wynder, E. L. (1994). Studies in mechanism and prevention: Striking a proper balance. *American Journal of Epidemiology, 139*(6), 547–549.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

3.06

Qualitative and Discourse Analysis

JONATHAN A. POTTER
Loughborough University, UK

3.06.1 INTRODUCTION: QUALITATIVE RESEARCH IN CONTEXT 118
- *3.06.1.1 Historical Moments in Qualitative Research in Clinical Settings* 118
- *3.06.1.2 Background Issues* 120
  - *3.06.1.2.1 Philosophy, sociology, and changing conceptions of science* 120
  - *3.06.1.2.2 Investigatory procedures vs. justificatory rhetoric* 121
  - *3.06.1.2.3 Quality and quantity* 122
- *3.06.1.3 Qualitative Research and Theory* 122
- *3.06.1.4 Boundaries of Qualitative Research and Coverage of the Current Chapter* 122

3.06.2 GROUNDED THEORY 123
- *3.06.2.1 Questions* 124
- *3.06.2.2 Procedures* 125
  - *3.06.2.2.1 Materials* 125
  - *3.06.2.2.2 Coding* 125
  - *3.06.2.2.3 Method of constant comparison* 125
  - *3.06.2.2.4 Memo writing* 126
  - *3.06.2.2.5 Validity* 126
- *3.06.2.3 Example: Clegg, Standen, and Jones (1996)* 126
- *3.06.2.4 Virtues and Limitations* 127

3.06.3 ETHNOGRAPHY AND PARTICIPANT OBSERVATION 128
- *3.06.3.1 Questions* 128
- *3.06.3.2 Procedures* 129
  - *3.06.3.2.1 Access* 129
  - *3.06.3.2.2 Field relations* 129
  - *3.06.3.2.3 Fieldnotes* 129
  - *3.06.3.2.4 Interviews* 130
  - *3.06.3.2.5 Analysis* 130
  - *3.06.3.2.6 Validity* 130
- *3.06.3.3 Example: Gubrium (1992)* 131
- *3.06.3.4 Virtues and Limitations* 131

3.06.4 DISCOURSE ANALYSIS 132
- *3.06.4.1 Questions* 133
- *3.06.4.2 Procedures* 134
  - *3.06.4.2.1 Research materials* 134
  - *3.06.4.2.2 Access and ethics* 135
  - *3.06.4.2.3 Audio and video recording* 135
  - *3.06.4.2.4 Transcription* 135
  - *3.06.4.2.5 Coding* 136
  - *3.06.4.2.6 Analysis* 136
  - *3.06.4.2.7 Validity* 137
- *3.06.4.3 Example: Peräkylä (1995)* 137
- *3.06.4.4 Virtues and Limitations* 138

3.06.5 FUTURE DIRECTIONS 139
3.06.6 REFERENCES 140

3.06.1 INTRODUCTION: QUALITATIVE RESEARCH IN CONTEXT

For many researchers the words "qualitative method" spell out psychology's most notorious oxymoron. If there is one thing that qualitative methods are commonly thought to lack it is precisely an adequate methodic way of arriving at findings. Indeed, for much of the twentieth century quantification has been taken as the principle marker of the boundary between a mature scientific psychology and commonsense, intuitive approaches. Since the late 1980s, however, there has been a remarkable increase in interest in qualitative research in psychology. This partly reflects a broader turn to qualitative research across the social sciences, although qualitative research of one kind or another has long been an established feature of disciplines such as education, sociology, and, most prominently, anthropology.

In psychology there is a handbook of qualitative research (Richardson, 1996) as well as a range of volumes and special journal issues whose major focus is on developing qualitative approaches to psychological problems (Antaki, 1988; Bannister, Burman, Parker, Taylor, & Tyndall, 1994; Henwood & Parker, 1995; Henwood & Nicolson, 1995; Smith, Harré, & van Langenhove, 1995). Psychology methods books and collections are increasingly serving up qualitative methods to accompany the more usual diet of experiments, questionnaires, and surveys. At the same time, an increasing permeability of boundaries between the social sciences has provided the environment for a range of trans-disciplinary qualitative methods books including a useful doorstop-sized handbook (Denzin & Lincoln, 1994) and varied edited and authored works (Bogdan & Taylor, 1975; Bryman & Burgess, 1994; Coffey & Atkinson, 1996; Gilbert, 1993; Lofland & Lofland, 1984; Miles & Huberman, 1994; Miller & Dingwall, 1997; Silverman, 1993, 1997a). These general qualitative works are complemented by a mushrooming range of books and articles devoted to specific methods and approaches.

Why has there been this increase in interest in qualitative research? Three speculations are proferred. First, there is a widespread sense that traditional psychological methods have not proved successful in providing major advances in the understanding of human life. Despite regular promissory notes, psychology seems to offer no earth-moving equivalent of the transistor, of general relativity, or of molecular genetics. Second, as is discussed below, views of science have changed radically since the 1950s, making it much harder to paint qualitative researchers, as either antiscientific extremists or merely sloppy humanists. Third, psychology is no longer as insulated from other social sciences as it has been in the past. Of course, for much of its twentieth century existence psychology has been invigorated by infusions from other sciences such as physiology and linguistics. However, in recent years there has been increasing exchange with disciplines where qualitative methods have been more established such as sociology and anthropology. This is reflected in contemporary theoretical developments such as constructionism (Gergen, 1994) and poststructuralism (Henriques, Hollway, Irwin, Venn, & Walkerdine, 1984) that have swept right across the human sciences.

This is, then, an exciting time for qualitative researchers, with new work and new opportunities of all kinds. Yet it should also be emphasised that qualitative research in psychology is in a chaotic state, with a muddle of inconsistent questions and approaches being blended together. Much poor work has involved taking questions formulated in the metaphysics of experimental psychology and attempting to plug them into one or more qualitative methods. At its worst such research peddles unsystematic and untheorized speculations about the influences on some piece of behavior which are backed up with two or three quotes from an interview transcript. This expanding literature and variable quality creates some problems for the production of a useful overview. This chapter selects what are seen as the most coherent and successful qualitative approaches from a range of possibilities, as well as focusing on those approaches which are being used, and have prospects for success, in clinical settings. A range of references is provided for those who wish to follow up alternative methods.

3.06.1.1 Historical Moments in Qualitative Research in Clinical Settings

In one sense a large proportion of twentieth century clinical therapeutic practice was based on qualitative methods. The process of conducting some kind of therapy or counseling with clients and then writing them up as case histories, or using them as the basis for inferences about aspects of the psyche, behavior, or cognitive processes, has been a commonplace of clinical work. Freud's use of case histories in the development of psychoanalytic thinking is probably the most influential. Although it is an approach to clinical knowledge that overwhelmingly eschews quantification, it is hard to say much about its methodic basis. For good or bad, it is dependent on the unformulated skills and intuitions of the therapist/researcher. In the hands of someone as brilliant as Freud the result can be extraordinary; elsewhere the product has often been merely ordinary. The problem for the readers of such research is that they can do little except either take it on trust or disagree. The process through which certain claims are established is not open to scrutiny. However, Freud's study of the case of Little Hans is exceptional in this regard, and so it is briefly worth considering (see Billig, 1998).

Although Freud initially based his arguments for the existence of the Oedipus complex on the interpretation of what patients told him in the course of therapy sessions, he attempted, unusually, to support this part of psychoanalytic theory with more direct evidence. He asked some of his followers to collect observations from their own children. The music critic Max Graf was most helpful in this regard and presented Freud with copious notes on conversations between his son, Hans, and other family members, as well as descriptions of dreams he had recounted. The published case history (Freud, 1977 [1909]) contains more than 70 pages of reports about Hans which Freud describes as reproduced "just as I received them" without "conventional emendations" (1977, pp. 170). Here is an example:

> Another time he [Hans] was looking on intently while his mother undressed before going to bed. "What are you staring like that for?" she asked.
> *Hans*: "I was only looking to see if you'd got a widdler too."
> *Mother*: "Of course. Didn't you know that?"
> *Hans*: "No. I thought you were so big you'd have a widdler like a horse."
> (1977, p. 173)

Freud's fascinating materials and striking interpretations beg many of the questions that have been central to qualitative research ever since. For example, what is the role of Max Graf's expectations (he was already an advocate of Freud's theories) in his selection and rendering of conversations with Hans? How closely do the extracts capture the actual interactions (including the emphasis, nonvocal elements, and so on)? What procedure did Freud use to select the examples that were reproduced from the full corpus? And, most importantly, what is the basis of Freud's interpretations? His interpretations are strongly derived from his theory, as is shown by his willingness to straightforwardly rework the overt sense of the records. Take this example:

> Hans (aged four and a half) was again watching his little sister being given her bath, when he began laughing. On being asked why he was laughing, he replied. "I'm laughing at Hanna's widdler." "Why?" "Because her widdler's so lovely."
> Of course his answer was a disingenuous one. In reality her widdler had seemed to him funny. Moreover, this is the first time he has recognized in this way the distinction between male and female genitals instead of denying it. (1977, p. 184)

Note the way Graf here, and implicitly Freud in his text, treat the laughter as the real indicator of Hans understanding of events, and his overt claim to find his sister's "widdler" lovely as a form of dissembling. Hans is not delighted by the appearance of his sister's genitals but is amused, in line with psychoanalytic theory, by their difference from his own. Again, the issue of how to treat the sense of records of interaction, and what interpretations should be made from them to things going on elsewhere such as actions or cognitions, is a fundamental one in qualitative research (Silverman, 1993).

Some 40 years later another of clinical psychology's great figures, Carl Rogers, advocated the use of newly developed recording technology to study the use of language in psychotherapy itself, with the aim of understanding and improving therapeutic skills. For him such recordings offered "the first opportunity for an adequate study of counseling and therapeutic procedures, based on thoroughly objective data" (Rogers, 1942, p. 433). Rogers envisaged using such recordings in the development of a scale to differentiate the styles of different counselors and studies of the patterns of interaction; for example, "what type of counselor statement is most likely to be followed by statements of the client's feeling about himself?" (1942, p. 434).

Rogers' emphasis on the virtue of recordings was followed up in two major "microscopic" studies of psychotherapeutic discourse. The first by Pettinger, Hockett, and Danehy (1960) focused on the initial portion of an initial interview. A typical page of their study has just a few words of transcript coupled with an extended discussion of their sense. Much of the focus was on the prosodic cues—the intonation and stress—provided in the interview and their contextual significance. Prosody is, of course, a feature of interaction which is almost impossible to reliably capture in *post hoc* notes made by key informants and so highlights the virtue of the new technology. A second study by Labov and Fanshell (1977) also focused on the opening stages of a therapy session, in this case five episodes of interaction from the first 15 minutes of a psychoanalytic therapy session with Rhoda, a 19-year-old girl with a history of anorexia nervosa.

The classic example of ethnographic work in the history of clinical psychology is Goffman's study of the everyday life of a mental hospital published under the title *Asylums* (1961). It is worth noting that although Goffman was a sociologist, the various essays that make up *Asylums* were initially published in psychiatry journals. Rather than utilize tape recording technology to capture the minutiae of some social setting, Goffman used an ethnographic approach. He spent a year working ostensibly as an assistant to the athletic director of a large mental hospital, interacting with patients and attempting to build up an account of the institution as viewed by the patients. His justification for working in this way is instructive for how the strengths and weaknesses of qualitative work have been conceptualized:

> Desiring to obtain ethnographic detail regarding selected aspects of patient social life, I did not employ usual kinds of measurements and controls. I assumed that the role and time required to gather statistical evidence for a few statements would preclude my gathering data on the tissue and fabric of patient life. (1961, p. 8)

As an ethnographic observer, he developed an understanding of the local culture and customs of the hospital by taking part himself. He used the competence generated in this way as the basis for his writing about the life and social organization in a large mental hospital.

3.06.1.2 Background Issues

Before embarking on a detailed overview of some contemporary qualitative approaches to clinical topics there are some background issues that are worth commenting on, as they will help make sense of the aims and development of qualitative approaches. In some cases it is necessary to address issues that have been a long-standing source of confusion where psychologists have discussed the use of qualitative methods.

3.06.1.2.1 Philosophy, sociology, and changing conceptions of science

As noted above, the development of qualitative work in psychology has been facilitated by the more sophisticated understanding of the nature of science provided by philosophers and sociologists since the 1970s. The image of the lone scientist harvesting facts, whose truth is warranted through the cast-iron criterion of replication, has intermittently been wheeled on to defend supposedly scientific psychology against a range of apparently sloppier alternatives. However, this image now looks less than substantial (see Chalmers, 1992; Potter, 1996a; Woolgar, 1988).

The bottom-line status of scientific observation has been undermined by a combination of philosophical analyses and sociological case studies. Philosophers have highlighted the logical relationships between observation statements and theoretical notions (Hesse, 1974; Kuhn, 1962; Popper, 1959). Their argument is that even the simplest of scientific descriptions is dependent on a whole variety of theoretical assumptions. Sociologists have supplemented these insights with studies of the way notions of observations are used in different scientific fields. For example, Lynch (1994) notes the way the term observation is used in astronomy as a loose device for collecting together a range of actions such as setting up the telescope, attaching sensors to it, building up traces on an oscilloscope, converting these into a chart and canvassing the support of colleagues. Knorr Cetina (1997) documents the different notions of observation that appear in different scientific specialities, suggesting that high energy physicists and molecular biologists, for example, work with such strikingly different notions of what is empirical that they are best conceived of as members of entirely different epistemic cultures.

The idea that experimental replication can work as a hard criterion for the adequacy of any particular set of research findings has been shown to be too simple by a range of sociological studies of replication in different fields (Collins, 1981). For example, Collins (1985) has shown that the achievement of a successful replication is closely tied to the conception of what counts as a competent experiment in the first place—and this itself was often as much a focus of controversy as the phenomenon itself. In a study of gravity wave researchers, Collins found that those scientists who believed in gravity waves tended to treat replications that claimed to find them as competent and replications that failed to find them as incompetent. The reverse pattern was true of nonbelievers. What this meant was that replication did not stand outside the controversy as a neutral arbiter of the outcome, but was as much part of the controversy as everything else.

Philosophical and sociological analysis has also shown that the idea that a crucial experiment can be performed which will force the abandonment of one theory and demonstrate the correctness of another is largely mythical (Lakatos, 1970; Collins & Pinch, 1993). Indeed, historical studies suggest that so-called crucial experiments are not merely insufficient to effect the shift from one theory to another, they are often performed, or at least constructed as crucial, after the shift to provide illustration and legitimation (Kuhn, 1977).

Let us be clear at this point. This research does not show that careful observation, skilful replication, and theoretically sophisticated experiments are not important in science. Rather, the point is that none of these things are bottom-line guarantees of scientific progress. Moreover, these sociological studies have suggested that all these features of science are embedded in, and inextricable from, its communal practices. Their sense is developed and negotiated in particular contexts in accordance with *ad hoc* criteria and a wide range of craft skills which are extremely hard to formulate in an explicit manner (Knorr Cetina, 1995; Latour & Woolgar, 1986). The message taken from this now very large body of work (see Jasanoff, Markle, Pinch, & Petersen, 1995) is not that psychologists must adopt qualitative methods, or that qualitative methods will necessarily be any better than the quantitative methods that they may replace or supplement; it is that those psychologists who have argued against the adoption of such methods on the principle that they are unscientific are uninformed about the nature of science.

3.06.1.2.2 Investigatory procedures vs. justificatory rhetoric

There are a number of basic linguistic and metatheoretical difficulties in writing about qualitative methods for psychologists. Our terminology for methodological discussion—reliability, validity, sampling, factors, variance, hypothesis testing, and so on—has grown up with the development of quantitative research using experiments and surveys. The language has become so taken-for-granted that it is difficult to avoid treating it as obvious and natural. However, it is a language that is hard to disentangle from a range of metatheoretical assumptions about the nature of behavior and processes of interaction. Traditional psychology has become closely wedded to a picture of factors and outcomes which, in turn, cohabits with the multivariate statistics which are omnipresent where data is analyzed. For some forms of qualitative research, particularly most discourse and ethnographic work, such a picture is inappropriate. This does not mean that such research is incoherent or unscientific, merely that it should not be construed and evaluated using the family of concepts whose home is experimental journal articles. Likewise the psychological model of hypothesis testing is just one available across the natural and human sciences. Qualitative research that utilizes theoretically guided induction, or tries to give a systematic description of some social realm, should not be criticized on the grounds that it is unscientific, let alone illegitimate. Ultimately, the only consistent bottom line for the production of excellent qualitative work is excellent scholarship (Billig, 1988).

Another difference between traditional quantitative and qualitative work is that in the traditional work the justification of research findings is often taken to be equivalent to the complete and correct carrying out of a set of codified procedures. Indeed, methods books are often written as if they were compendia of recipes for achieving adequate knowledge. Sampling, operationalization of variables, statistical tests, and interpretation of significance levels are discussed with the aid of tree diagrams and flow charts intended to lead the apprentice researcher to the correct conclusion. In one sense, much qualitative work is very different to this, with the procedures for justifying the research claims being very different to the procedures for producing the work. Thus, the manner in which a researcher arrives at some claims about the various functions of "mm hm's" in psychiatric intake interviews, say, may be rather different from the manner in which they justify the adequacy of the analysis. Yet, in another sense the difference between qualitative and quantitative research is more apparent than real, for studies of the actual conduct of scientists following procedural rules of method show that such rules require a large amount of tacit knowledge to make them understandable and workable, and that they are often more of a rhetorical device used to persuade other scientists than an actual constraint on practice (Gilbert & Mulkay, 1984; Polyani, 1958). As Collins (1974) showed in an innovative ethnographic study, when a group of scientists wrote a paper offering the precise technical specification of how to make a new laser, the only people who were able to build a working laser of their own had actually seen one built; just reading the paper was not enough.

This presents something of a dilemma for anyone writing about qualitative methods. Should they write to help people conduct their research so as better to understand the world, or should they work to provide the sorts of formalized procedural rules that can be drawn on in the methods sections of articles to help persuade the psychologist reader? In practice, most writing does some of each. However, the difficulty that psychologists often report when attempting qualitative work is probably symptomatic of the failure to fully explicate the craft skills that underpin qualitative work.

3.06.1.2.3 Quality and quantity

There are different views on how absolute the quantity/quality divide is. Arguments at different levels of sophistication have been made for future integration of qualitative and quantitative research (Bryman, 1988; Silverman, 1993). It is undoubtedly the case that at times proponents of both quantitative and qualitative research have constructed black and white stereotypes with little attempt at dialog (although a rare debate about the relative virtues of quantitative and qualitative research on the specific topic of attitudes towards mental hospitalisation is revealing—Weinstein, 1979, 1980; Essex et al., 1980). It is suggested that quantification is perfectly appropriate in a range of situations, dependent on appropriate analytic and theoretical judgements.

In many other research situations the goal is not something that can be achieved through counting. For example, if the researcher is explicating the nature of "circular questioning" in Milan School Family Therapy, that goal is a prerequisite for a study which considers the statistical prevalence of such questioning. Moreover, there are arguments for being cautious about quantification when studying the sorts of discursive and interactional materials which have often been central to qualitative research because of distortions and information loss that can result (see Schegloff, 1993, and papers in Wieder, 1993). Some of the grounds for caution come from a range of qualitative studies of quantification in various clinical settings (Ashmore, Mulkay, & Pinch, 1989; Atkinson, 1978; Garfinkel, 1967a; Potter, Wetherell, & Chitty, 1991).

3.06.1.3 Qualitative Research and Theory

It is hard to overestimate how close the relationship is between the theories, methods, and questions used by psychologists. Theories specify different positions on cognition and behavior, different notions of science, different views of the role of action and discourse, different understandings of the role of social settings, and, most fundamentally, different objects for observation.

For both psychoanalytic theory and most of the mass of theories and perspectives that make up modern cognitivism the objects of observation are hypothetical mental entities (the Oedipus complex, attributional heuristics). Psychoanalytic researchers have generally preferred to engage in an interpretative exercise of reconstructing those entities from the talk of patients undergoing therapy. Cognitive psychologists have typically used some hypothetico-deductive procedure where predictions are checked in experiments which investigate, say, the attributional style of people classified as depressed. Note that in both of these cases they are using people's discourse—talk in the therapy session, written responses to a questionnaire—yet in neither case is the discourse as such of interest. In contrast, for researchers working with different perspectives such as social constructionism or discursive psychology, the talk or writing itself, and the practices of which it is part, is the central topic. For these researchers there is a need to use procedures which can make those practices available for study and allow their organization to be inspected and compared.

To take another example, behaviorist psychologists have done countless studies on the effects of particular regimes of reward and punishment on target behaviors such as compulsive hand washing. However, such studies are typically focused on outcomes and statistical associations, whereas a theoretical perspective such as symbolic interactionism or, to give a more psychological example, Vygotskyan activity theory, encourage a more ethnographic examination of the settings in which rewards are administered and of the sense that those behaviors have in their local context

Without trying to flesh out these examples in any detail, the important practical point they make is that it is a mistake to attempt simply to import a question which has been formulated in the problematics of one theoretical system, and attempt to answer it using a method developed for the problematics of another. The failure to properly conceptualize a research question that fits with the research method is a major source of confusion when psychologists start to use qualitative methods.

3.06.1.4 Boundaries of Qualitative Research and Coverage of the Current Chapter

What distinguishes qualitative research from quantitative? And what qualitative approaches are there? These questions are not as straightforward as they seem. In the past the qualitative/quantitative distinction has sometimes been treated as the equivalent of the distinction between research that produces objective and subjective knowledge—a distinction which makes little sense in the light of recent sociology and philosophy of science. Sometimes certain approaches using numbers have been treated as qualitative. For example, content analysis has occasionally been treated as a qualitative method because it is used to deal with "naturally occurring" objects such as diaries, novels, transcripts of meetings, and so on. Content analysis was meant to eliminate many of the potential "reactive" effects that bedevil social research and thereby avoid the problem in experimental and survey research, of how findings relate to what goes on in the real world; for these are records of (indeed, examples of) what is actually going on. However, in this chapter content analysis is treated as quantitative, and therefore outside the scope of this survey, on the grounds that it (i) transforms phenomena into numerical counts of one kind or another and (ii) typically attempts to statistically relate these counts to some broader factors or variables. For useful introductions to content analysis related to psychological topics see Holsti (1969) and Krippendorf (1980).

For similar reasons repertory grid analysis associated with personal construct theory, and the "Q" methodology developed by William Stephenson, have sometimes been treated as qualitative. The rationale for this was probably that they were often focused on understanding the reasoning or cognitive organization of single individuals rather than working exclusively from population statistics. However, as they involve quantitative manipulation of elicited responses from participants they will not be dealt with here. The ideographic/nomathetic distinction will be treated as orthogonal to the qualitative/quantitative one! For accessible introductions to these approaches, see Smith (1995) on repertory grid methods and Stainton Rogers (1995) on Q methodology.

In addition to these methods which are excluded as not properly qualitative, a wide range of methods have not been discussed which nevertheless satisfy the criterion of being at least minimally methodic and generally eschewing quantification. For simplicity, Table 1 lists nine methods or approaches which have been excluded, along with one or two references that would provide a good start point for any researcher who was interested in learning more about them. It would take some time to make explicit the reasons for excluding all of them. Generally, the problem is that they have not been, and are unlikely to be in the future, particularly useful for studying problems in the area of clinical psychology (e.g., focus groups—although, see Piercy & Nickerson, 1996). In some cases, the approaches are not coherent enough to warrant discussion. In others, their central problematics are better addressed by the approaches that are discussed.

The most controversial exclusion is probably humanistic methods given that humanistic psychology developed in settings which had a broad emphasis on therapy and psychological well-being. It is suggested that the romanticism of much humanistic psychology is attractive, but ultimately unconvincing. However, it is often, quite legitimately, more concerned with developing participants' skills and sensitivity than developing propositional claims and arguments; as such it is often offering a set of techniques for expanding human potential rather than developing methods for research. Feminist methods are excluded, despite an appreciation of the importance of feminist issues in clinical settings, because the arguments for the existence of specifically feminist methods (as opposed to theories or arguments) are not convincing. This is particularly true where such claims give a major epistemological role for experience or intuition (Ellis, Kiesinger, & Tillmann-Healy, 1997). These are topics for decomposition and analysis rather than bottom-lines for knowledge. For some arguments in both directions on this topic see Gelsthorpe (1992), Hammersley (1992), Ramazanoglu (1992), and Wilkinson (1986).

Finally, it should be stressed that qualitative work is not seen as having some overall coherence. Quite the contrary, it is fragmented and of highly variable quality. Nor is some overall coherence seen as a desirable goal. Those workers who talk of a qualitative paradigm (Guba & Lincoln, 1994, Reason & Rowan, 1981) unhelpfully blur over a range of theoretical and metatheoretical differences (see Henwood & Pidgeon, 1994).

3.06.2 GROUNDED THEORY

Grounded theory has the most clear-cut origin of any of the approaches discussed here. The term was coined by two sociologists in an influential book: *The discovery of grounded theory* (Glaser & Strauss, 1967). Some of its key features and concerns were a product of its birthplace within sociology, where it was developed to counter what the authors saw as a preoccupation, on the one hand, with abstract grand theories and, on the other, with testing those theories through large scale quantitative studies. Grounded theory was intended to link theoretical developments (conceived as plausible relations among concepts and sets of concepts—Strauss & Corbin, 1994) more closely to the particulars of settings, to ground

Table 1 Varieties of qualitative research not covered.

Qualitative research method	*Source*
Action research	Argyris, Putnam, & Smith, 1985, Whyte, 1991
Documentary studies	Hodder, 1994, Scott, 1990
Ethogenics	Harré, 1992
Feminist approaches	Olesen, 1994, Reinharz, 1992
Focus groups	Krueger, 1988, Morgan, 1997
Humanistic, participative research	Reason & Rowan, 1981, Reason & Heron, 1995
Life histories	Plummer, 1995, Smith, 1994
Role play	Yardley, 1995
Semiotics	Manning & Cullum-Swan, 1994

middle range theories in actual qualitative data rather than to start from preconceived hypotheses. It is important to stress that grounded theory is not a theory as such, rather it is an approach to theorising about data in any domain. Moreover, since its inception in the 1960s, work on the theory ladenness of data has made the idea of "grounding" theory increasingly problematic.

Much grounded theory research has been done outside of psychology; however, psychologists have become increasingly interested in the approach in general (Charmaz, 1995; Henwood and Pidgeon, 1995; Pidgeon, 1996; Pidgeon and Henwood, 1996; Rennie, Phillips & Quartaro, 1988), and have carried out specific studies in health (Charmaz, 1991, 1994) and clinical (Clegg, Standen, & Jones, 1996) settings.

3.06.2.1 Questions

Grounded theory is designed to be usable with a very wide range of research questions and in the context of a variety of metatheoretical approaches. Rather like the statistical analyses that psychologists are more familiar with, it deals with patterns and relationships. However, these are not relationships between numbers but between ideas or categories of things, and the relationships can take a range of different forms. In some respects the procedures in grounded theory are like the operation of a sophisticated filing system where entries are cross-referenced and categorized in a range of different ways. Indeed, this is one qualitative approach that can be effectively helped by the use of computer packages such as NUDIST, which was itself developed to address grounded theory notions.

Grounded theory has proved particularly appropriate for studying people's understandings of the world and how these are related to their social context. For example, Turner (1994; Turner & Pidgeon, 1997) has used grounded theory to attempt to explain the origins of manmade disasters like fires and industrial accidents; Charmaz (1991) has studied the various facets that make up people's experience of chronic illness; Clegg, Standen, and Jones (1996) focused on the staff members' understanding of their relationship with adults with profound learning disabilities. In each case, a major concern was to incorporate the perspectives of the actors as they construct their particular social worlds. Grounded theory methods can help explicate the relation of actions to settings (how does the behavior of key personnel in the evolution of a major fire follow from their individual understanding of events and physical positioning?); it can be used for developing typologies of relevant phenomena (in what different ways do sufferers of chronic illness conceptualize their problem?); and it can help identify patterns in complex systems (how does the information flowing between social actors help explain the development of a laboratory smallpox outbreak?).

Like most of the qualitative approaches discussed here, grounded theory is not well suited to the kinds of hypothesis testing and outcome evaluation that have traditionally been grist to the mill of clinical psychology, because of its open-ended and inductive nature. Although the researcher is likely to come to a topic with a range of more or less explicit ideas, questions, and theories, it is not necessary for any or all of these to be formally stated before research gets under way. The approach can start with a specific problem or it may be more directed at making sense of an experience or setting.

Grounded theory can be applied to a range of different textual materials such as documents, interview transcripts and records of interaction, and this makes it particularly suitable for certain kinds of questions. It can deal with records which exist prior to the research and it can deal with materials specifically collected. The processes of coding allow quite large amounts of material to be dealt with. For example, while Turner studied a single (lengthy) official report of a major fire in a holiday complex, Charmaz studied 180 interviews with 90 different people with chronic illnesses. The requirement is only that the material can be coded.

3.06.2.2 Procedures

The procedures for conducting grounded theory work are straightforward to describe, if less so to follow in practice. Pidgeon and Henwood (1996, p. 88) provide a useful diagram to explicate the process (Figure 1).

3.06.2.2.1 Materials

In line with the general emphasis on participants' perspectives and on understanding patterns of relationships, researchers often attempt to obtain rich materials such as documents and conversational, semistructured interviews. These may be supplemented by participant observation in the research domain, generating fieldnotes which can be added to other data sets or simply used to improve the researcher's understanding so they can better deal with the other materials.

After data is collected and stored the intensive process that is most characteristic of grounded theory is performed. This involves coding the data, refining the coding and identifying links between categories, and writing "memos" which start to capture theoretical concepts and relationships.

3.06.2.2.2 Coding

Different grounded theory researchers approach coding in different ways. For example, it can involve generating index cards or making annotations next to the relevant text. The researcher works through the text line by line, or paragraph by paragraph, labeling the key concepts that appear. Charmaz (1995, p. 38) suggests a series of specific questions that are useful for picking out the key concepts:

(i) What is going on?
(ii) What are the people doing?
(iii) What is the person saying?
(iv) What do these actions and statements take for granted?
(v) How do structure and context serve to support, maintain, impede, or change these actions and statements?

More broadly, Pidgeon and Henwood suggest that this phase of coding is answering the question: "what categories or labels do I need in order to account for what is of importance to me in this paragraph?" (1996, p. 92).

Such coding is intensive and time consuming. For example, Table 2 shows an example by Charmaz of line-by-line coding of just a brief fragment of one of her 180 interviews. Note the way that the interview fragment is coded under a number of different topics. There is no requirement in grounded theory that categories apply exclusively.

3.06.2.2.3 Method of constant comparison

Coding is not merely a matter of carefully reading and labeling the materials. As the coding continues the researcher will be starting to identify categories that are interesting or relevant to the research questions. They will refine their indexing system by focused coding which will pick out all the instances from the data coded as, for example "avoiding disclosure." When such a collection has been produced the researcher can focus on the differences in the use of this category according to the setting or the actors involved. This is what grounded theorists refer to as the "method of constant comparison." In the course of such comparisons the category system may be reworked; some categories will be merged together and others will be broken up, as the close reading of the data allows an increasingly refined understanding.

Table 2 Line-by-line coding.

Coding	*Interview*
Shifting symptoms, having inconsistent days	If you have lupus, I mean one day it's my liver; one day it's my joints; one day it's my head, and
Interpreting images of self given by others	it's like people really think you're a hypochondriac if you keep complaining about
Avoiding disclosure	different ailments. . . . It's like you don't want to say anything because people are going to start
Predicting rejection	thinking, you know, "God, don't go near her,
Keeping others unaware	all she is—is complaining about this."

Source: Charmaz (1995, p. 39)

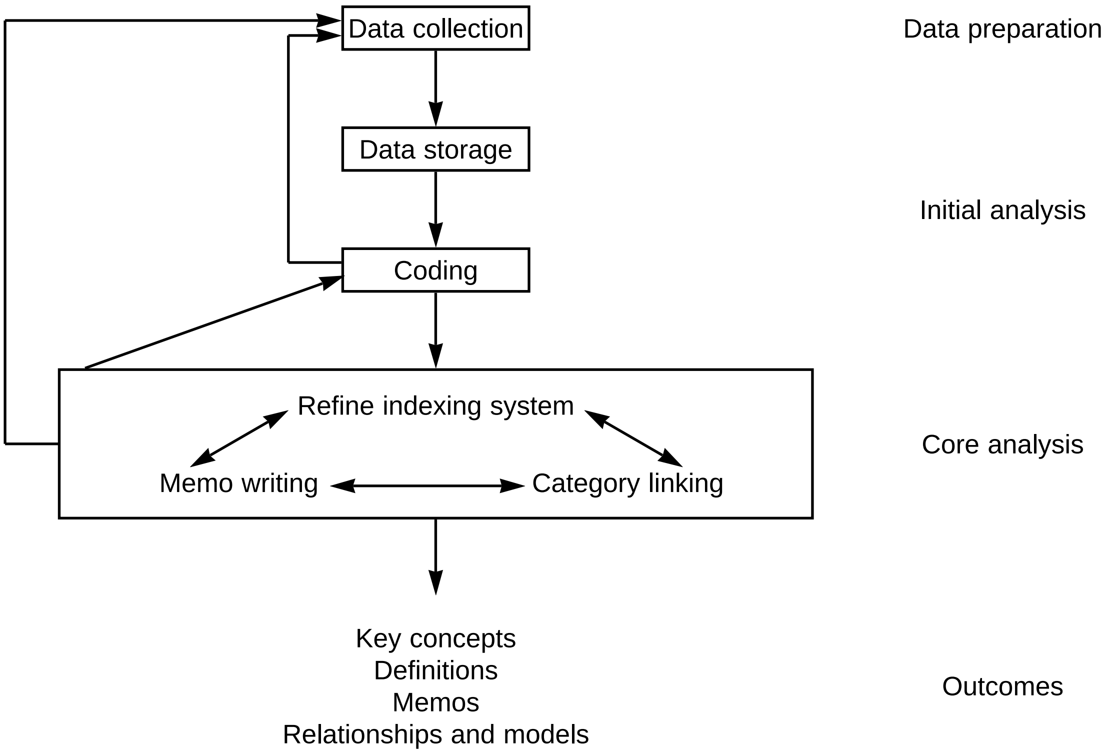

Figure 1 Procedures for conducting grounded theory research.

3.06.2.2.4 Memo writing

Throughout this stage of coding and comparison grounded theorists recommend what they call memo writing. That is, writing explicit notes on the assumptions that underly any particular coding. Memo writing is central to the process of building theoretical understandings from the categories, as it provides a bridge between the categorization of data and the writing up of the research. In addition to the process of refining categories, a further analytic task involves linking categories together. The goal here is to start to model relationships between categories. Indeed, one possibility may be the production of a diagram or flow chart which explicitly maps relationships.

As Figure 1 makes clear, these various elements in grounded theory analysis are not discrete stages. In the main phase of the analysis, refining the indexing system is intimately bound up with the linking of categories, and memo writing is likely to be an adjunct to both. Moreover, this analysis may lead the researcher back to the basic line-by-line coding, or even suggest the need to collect further material for analysis.

3.06.2.2.5 Validity

There is a sense in which the general methodological procedures of grounded theory are together designed to provide a level of validation as they force a thoroughgoing engagement with the research materials. Line-by-line coding, constant comparison, and memo writing are all intended to ensure that the theoretical claims made by the analyst are fully grounded in the data. That, after all, was the original key idea of grounded theory. However, some specific procedures of validation have been proposed.

Some grounded theorists have suggested that respondent validation could be used as a criterion. This involves the researcher taking back interpretations to the participants to see if they are accepted. The problem with such an approach is that participants may agree or disagree for a range of social reasons or they may not understand what is being asked of them. Other approaches to validation suggest that research should be generative, that is, facilitate further issues and questions; have rhetorical power, that is, prove effective in persuading others of its effectiveness; or that there could be an audit trail which would allow another researcher to check how the conclusions were reached (Henwood & Pidgeon, 1995; Lincoln & Guba, 1985).

3.06.2.3 Example: Clegg, Standen, and Jones (1996)

There is no example in clinical psychology of the sorts of full-scale grounded theory study that Charmaz (1991) conducted on the experience of illness or Glaser and Strauss (1968) carried out on hospital death. Nevertheless, a modest study by Clegg et al. (1996) of the relationships between staff members and adults with profound learning disabilities illustrates some of the potential of grounded theory for clinical work.

In line with the grounded theory emphasis on the value of comparison, 20 staff members from four different residential settings were recruited. Each member of staff was videotaped during eight sessions of interaction with a single client that they had known for over three months. Each staff member was subsequently interviewed about their work, their relationship with the client, a specific experience with the client, and their understanding of the client's view of the world. These conversational interviews were tape recorded and transcribed, and form the data for the major part of the study.

The study followed standard grounded theory procedures in performing a range of codings developing links between categories and building up theoretical notions from those codings. The range of different outcomes of the study is typical of grounded theory. In the first place, they identify four kinds of relationships that staff may have with clients: provider (where meeting of the client's needs is primary); meaning-maker (where making sense of the client's moods or gestures is primary); mutual (where shared experience and joy at the client's development is primary); companion (where merely being together is treated as satisfying). The authors go on to explore the way different settings and different client groups were characterized by different relationships. Finally, they propose that the analysis supports four propositions about staff–client relationships: some types of relationship are better than others (although this will vary with setting and client); staff see the balance of control in the relationship as central; families can facilitate relationships; professional networks create dilemmas.

3.06.2.4 Virtues and Limitations

Grounded theory has a range of virtues. It is flexible with respect to forms of data and can be applied to a wide range of questions in varied domains. Its procedures, when followed fully, force the researcher into a thorough engagement with the materials; it encourages a slow-motion reading of texts and transcripts that should avoid the common qualitative research trap of trawling a set of transcripts for quotes to illustrate preconceived ideas. It makes explicit some of the data management procedures that are commonly left inexplicit in other qualitative approaches. The method is at its best where there is an issue that is tractable from a relatively common sense actor's perspective. Whether studying disasters, illness, or staff relationships, the theoretical notions developed are close to the everyday notions of the participants. This makes the work particularly suitable for policy implementation, for the categories and understandings of the theory are easily accessible to practitioners and policy makers.

Some problems and questions remain, however. First, although there is a repeated emphasis on theory—after all, it is in the very name of the method—the notion of theory is a rather limited one strongly influenced by the empiricist philosophy of science of the 1950s. The approach works well if theory is conceived in a limited manner as a pattern of relationships between categories, but less well if theories are conceived of as, say, models of underlying generative mechanisms (Harré, 1986).

Second, one of the claimed benefits of grounded theory work is that it works with the perspective of participants through its emphasis on accounts and reports. However, one of the risks of processes such as line-by-line coding is that it leads to a continual pressure to assign pieces of talk or elements of texts to discrete categories rather than seeing them as inextricably bound up with broader sequences of talk or broader textual narratives. Ironically this can mean that instead of staying with the understandings of the participants their words are assigned to categories provided by the analyst.

Third, grounded theorists have paid little attention to the sorts of problems in using textual data that ethnomethodologists and discourse analysts have emphasised (Atkinson, 1978; Gilbert & Mulkay, 1984; Silverman, 1993; Widdicombe & Wooffitt, 1995). For example, how far is the grounding derived not from theorizing but from reproducing common sense theories as if they were analytic conclusions? How far are Clegg's et al. (1996) staff participants, say, giving an accurate picture of their relationships with clients, and how far are they drawing on a range of ideas and notions to deal with problems and work up identities in the interview itself?

Some practitioners are grappling with these problems in a sophisticated manner (Pidgeon & Henwood, 1996). As yet there is not a large enough body of work with clinical materials to allow a full evaluation of the potential of this method. For more detailed coverage of grounded theory the reader should start with the excellent introductions by Pidgeon (1996) and Pidgeon and Henwood (1996); Charmaz (1991) provides a full scale research illustration of the potential of grounded theory; Rafuls and Moon (1996) discuss grounded theory in the context of family therapy; and, despite its age, Glaser and Strauss (1967) is still an informative basis for understanding the approach.

3.06.3 ETHNOGRAPHY AND PARTICIPANT OBSERVATION

Ethnography is not a specific method so much as a general approach which can involve a number of specific research techniques such as interviewing and participant observation. Indeed, this has been a source of some confusion as rather different work is described as ethnography in anthropology, sociology, and other disciplines such as education or management science. The central thrust of ethnographic research is to study people's activities in their natural settings. The concern is to get inside the understanding and culture, to develop a subtle grasp of how members view the world, and why they act as they do. Typically, there is an emphasis on the researcher being involved in the everyday world of those who are being studied. Along with this goes a commitment to working with unstructured data (i.e., data which has not been coded at the point of data collection) and a tendency to perform intensive studies of small numbers of cases. Ethnography is not suited to the sorts of hypothetico-deductive procedures that are common in psychology.

Two important tendencies in current ethnographic work can be seen in their historical antecedents in anthropology and sociology. In the nineteenth century, anthropology often involved information collected by colonial administrators about members of indigenous peoples. Its use of key informants and its focus on revealing the details of exotic or hidden cultures continue in much work. Sociologists of the "Chicago School" saw ethnography as a way of revealing the lives and conditions of the US underclass. This social reformism was married to an emphasis on understanding their participants' lives from their own perspective.

3.06.3.1 Questions

Ethnography comes into its own where the researcher is trying to understand some particular sub-cultural group in a specific setting. It tends to be holistic, focusing on the entire experience participants have of a setting, or their entire cosmology, rather than focusing on discrete variables or phenomena. This means that ethnographic questions tend to be general: What happens in this setting? How do this group understand their world? In Goffman's (1961) *Asylums* he tried to reveal the different worlds lived by the staff and inmates, and to describe and explicate some of the ceremonies that were used to reinforce boundaries between the two groups. A large part of his work tracked what he called the "moral careers" of inmates from prepatient, through admission, and then as inpatients. Much of the force and influence of Goffman's work derived from its revelations about the grim "unofficial" life lived by patients in large state mental hospitals in the 1950s. In this respect it followed in the Chicago school tradition of exposé and critique. Rosenhan's (1973) classic study of hospital admission and the life of the patient also followed in this tradition. Famously it posed the question of what was required to be diagnosed as mentally ill and then incarcerated, and discovered that it was sufficient to report hearing voices saying "empty," "hollow," and "thud." This "pseudopatient" study was designed with a very specific question about diagnostic criteria in mind; however, after the pseudopatients were admitted they addressed themselves to more typically ethnographic concerns, such as writing detailed descriptions of their settings, monitoring patient contact with different kinds of staff, and documenting the experience of powerlessness and depersonalization.

Goffman's and Rosenhan's work picks up the ethnographic traditions of revealing hidden worlds and providing a basis for social reform. Jodelet (1991) illustrates another analytic possibility by performing an intensive study of one of the longest running community care schemes in the world, the French colony of Ainay-le-Château where mental patients live with ordinary families. Again, in line with the possibilities of ethnography, she attempted to explore the whole setting, including the lives of the patients and their hosts and their understandings of the world. Her work, however, is notable for showing how ethnography can explore the representational systems of participants and relate that system to the lives of the participants. To give just one small example, she shows the way the families' representation of a close link between madness and uncleanliness relates to practices such as taking meals separately from the lodgers.

Another topic for ethnographic work has been the practice of psychotherapy itself (Gubrium, 1992; Newfield, Kuehl, Joanning, & Quinn, 1990, 1991). These studies are interested in the experience of patients in therapy and their conceptions of what therapy is, as well as the practices and conceptions of the therapists. Such studies do not focus exclusively on the interaction in the therapy session itself, but on the whole setting.

Although ethnography is dependent on close and systematic description of practices it is not necessarily atheoretical. Ethnographic studies are often guided by one of a range of theoretical conceptions. For example, Jodelet's (1991) study was informed by Moscovici's (1984) theory of social representations and Gubrium's (1992) study of family therapy was guided by broader questions about the way the notion of family and family disorder are constructed in Western society. Ethnography has sometimes been treated as particularly appropriate for feminist work because of the possibility of combining concerns with experience and social context (e.g., Ronai, 1996).

3.06.3.2 Procedures

Ethnography typically involves a mix of different methods with interviews and participant observation being primary, but often combined with nonparticipant observation and the analysis of documents of one kind or another. Such a mixture raises a large number of separate issues which will only be touched on here. There are whole books on some elements of ethnographic work such as selecting informants (Johnson, 1990), living with informants (Rose, 1990), and interpreting ethnographic writings (Atkinson, 1990). The focus here is on research access, field relations, interviewing, observing, fieldnotes, and analysis (see Ellen, 1984; Fetterman, 1989; Fielding, 1993; Hammersley & Atkinson, 1995; Rachel, 1996; Toren, 1996; Werner & Schoepfle, 1987).

3.06.3.2.1 Access

Research access is often a crucial issue in ethnography, as a typical ethnographic study will require not only access to some potentially sensitive group or setting but may involve the researcher's bodily presence in delicate contexts. Sitting in on a family therapy session, for example, may involve obtaining consent from a range of people who have different concerns and has the potential for disrupting, or at least subtly changing, the interaction that would have taken place. There are only restricted possibilities for the ethnographer to enter settings with concealed identities, as Goffman did with his mental hospital study, and Rosenhan's pseudopatients did. Moreover, such practices raise a host of ethical and practical problems which increasingly lead ethnographers to avoid deception. Access can present another problem in sensitive settings if it turns out that it is precisely unusual examples where access is granted, perhaps because the participants view them as models of good practice.

3.06.3.2.2 Field relations

Field relations can pose a range of challenges. Many of these follow from the nature of the participation of the researcher in the setting. How far should researchers become full participants and how far should they stay uninvolved observers? The dilemma here is that much of the power of ethnography comes from the experience and knowledge provided by full participation, and yet such participation may make it harder to sustain a critical distance from the practices under study. The ethnographer should not be converted to the participants' cultural values; but neither should they stay entirely in the Martian role that will make it harder to understand the subtle senses through which the participants understand their own practices. Field relations also generate many of practical, prosaic, but nevertheless important problems which stem from the sheer difficulty of maintaining participant status in an unfamiliar and possibly difficult setting for a long period of time. At the same time there are a whole set of skills required to do with building productive and harmonious relationships with participants.

3.06.3.2.3 Fieldnotes

One of the central features of participant observation is the production of fieldnotes. Without notes to take away there is little point in conducting observation. In some settings it may be possible to take notes concurrently with the action but often the researcher will need to rely on their memory, writing up notes on events as soon as possible after they happened. A rule of thumb is that writing up fieldnotes will take just as much time as the original period of observation (Fielding, 1993). In some cases it may be possible to tape record interaction as it happens. However, ethnographers have traditionally placed less value on recording as they see the actual process of note taking as itself part of the process through which the researcher comes to understand connections between processes and underlying elements of interaction.

Ethnographers stress that fieldnotes should be based around concrete descriptions rather than consisting of abstract higher-order interpretations. The reason for this is that when observation is being done it may not yet be clear what questions are to be addressed. Notes that stay at low levels of inference are a resource that can be used to address a range of different questions. Fielding argues that fieldnotes are expected:

> to provide a running description of events, people and conversation. Consequently each new setting observed and each new member of the setting

> merits description. Similarly, changes in the human or other constituents of the setting should be recorded. (1993, p. 162)

It is also important to distinguish in notes between direct quotation and broad précis of what participants are saying. A final point emphasised by ethnographers is the value of keeping a record of personal impressions and feelings.

3.06.3.2.4 Interviews

Ethnographers make much use of interviews. However, in this tradition interviews are understood in a much looser manner than in much of psychology. Indeed, the term interview may be something of a misnomer with its image of the researcher running through a relatively planned set of questions with a single passive informant in a relatively formal setting. In ethnography what is involved is often a mix of casual conversations with a range of different participants. Some of these may be very brief, some extended, some allowing relatively formal questioning, others allowing no overt questioning. In the more formal cases the interview may be conducted with a planned schedule of questions and the interaction is recorded and transcribed.

3.06.3.2.5 Analysis

There is not a neat separation of the data collection and analytic phases of ethnographic research. The judgements about what to study, what to focus on, which elements of the local culture require detailed description, and which can be taken for granted, are already part of analysis. Moreover, it is likely that in the course of a long period of participant observation, or a series of interviews, the researcher will start to develop accounts for particular features of the setting, or begin to identify the set of representations shared by the participants. Such interpretations are refined, transformed, and sometimes abandoned when the fieldwork is completed and the focus moves on to notes and transcripts.

Fielding (1993, p. 192) suggests that the standard pattern of work with ethnographic data is straightforward. The researcher starts with the fieldnotes and transcripts and searches them for categories and patterns. These themes form a basis for the ethnographic account of the setting, and they also structure the more intensive analysis and probably the write-up of the research. The data will be marked or cut up (often on computer text files) to collect these themes together. In practice, the ethnographer is unlikely to attempt a complete account of a setting, but will concentrate on a small subset of themes which are most important or which relate to prior questions and concerns.

The analytic phase of ethnography is often described in only the sketchiest terms in ethnography guidebooks. It is clear that ethnographers often make up their own ways of managing the large amount of materials that they collect, and for using that material in convincing research accounts. At one extreme, ethnography can be considered an approach to develop the researcher's competence in the community being studied—they learn to be a member, to take part actually and symbolically, and they can use this competence to write authoritatively about the community (Collins, 1983). Here extracts from notes and interview transcripts become merely exemplars of the knowledge that the researcher has gained through participation. At the other extreme, ethnography blurs into grounded theorizing, with the notes and transcripts being dealt with through line-by-line coding, comparison of categories, and memo writing. Here the researcher's cultural competence will be important for interpreting the material, but the conclusions will ultimately be dependent on the quality of the fieldnotes and transcripts and what they can support.

3.06.3.2.6 Validity

One of the virtues of ethnography is its rich ecological validity. The researcher is learning directly about what goes on in a setting by observing it, by taking part, and/or by interviewing the members. This circumvents many of the inferences that are needed in extrapolating from more traditional psychological research tools—questionnaires, experimental simulations—to actual settings. However, the closeness of the researcher to the setting does not in itself ensure that the research that is produced will be of high quality.

The approach to validity most often stressed by ethnographers is triangulation. At the level of data this involves checking to see that different informants make the same sorts of claims about actions or events. At the level of method, it involves checking that conclusions are supported by different methods, for example, by both interviews and observation. However, triangulation is not without its problems. Discourse researchers have noted that in practice the sheer variability in and between accounts makes triangulation of only limited use (Potter & Wetherell, 1987) and others have identified conceptual problems in judging what a successful triangulation between methods would be (Silverman, 1993).

3.06.3.3 Example: Gubrium (1992)

There are only a small number of ethnographies done in clinical settings or on clinical topics. For instance, many of the clinical examples in Newfield, Sells, Smith, Newfield, & Newfield's (1996) chapter on ethnography in family therapy are unpublished dissertations. The body of ethnographic work is small but increasing rapidly. I have chosen Gubrium's (1992) study of family therapy in two institutions as an example because it is a book-length study, and it addresses the therapeutic process itself rather than concentrating solely on the patients' lives in hospital or community care schemes. However, it is important to stress that Gubrium's focus was as much on what the therapy revealed about the way notions of family and order are understood in American culture as in therapeutic techniques and effectiveness. He was concerned with the way behaviours such as truancy take their sense as part of a troubled family and the way service professionals redefine family order as they instigate programmes of rehabilitation.

Gubrium's choice of two contrasting institutions is a commonplace one in ethnography. The small number enables an intensive approach; having more than one setting allows an illuminating range of comparisons and contrasts. In this case one was inpatient, one outpatient; one more middle class than the other; they also varied in their standard approaches to treatment. The virtues of having two field sites shine through in the course of the write-up, although it is interesting to note that the selection was almost accidental as the researcher originally expected to be successful in gaining access to only one of the institutions.

The fieldwork followed a typical ethnographic style of spending a considerable amount of times at the two facilities, talking to counselors, watching them at work in therapy sessions, reviewing videos of counseling, and making fieldnotes. The study also drew on a range of documents including patients' case notes and educational materials. In some ways this was a technically straightforward setting for ethnographic observation as many of the participants were themselves university trained practitioners who made notes and videos as part of the general workings of the facilities.

One of the striking differences between this ethnographic study of therapy and a typical process or outcome study is that the therapy is treated as a part of its physical, institutional, and cultural contexts. For instance, time is spent documenting the organization of the reception areas of the two facilities and the way the counselors use the manner in which the families seat themselves in that area as evidence of family dynamics. Gubrium writes about the important role of tissue boxes in both signaling the potential for emotional display and providing practical support when such display occurs:

> I soon realized that tissues were about more than weeping and overall emotional composure during therapy. Tissues mundanely signaled the fundamental reality of the home as locally understood: a configuration of emotional bonds. For Benson [a counselor] their usage virtually put the domestic disorder of the home on display, locating the home's special order in the minutiae of emotional expression. (Gubrium, 1992, p. 26)

The ethnographic focus on events in context means that therapy is treated as a product of actual interactions full of contingency and locally managed understandings. It shows the way abstract notions such as family systems or tough love are managed in practice, and the way the various workers relate to each other as well as to the clients. It provides an insight into the world of family therapy quite different from most other styles of research.

3.06.3.4 Virtues and Limitations

Much of the power of ethnographic research comes from its emphasis on understanding people's actions and representations both in context and as part of the everyday practices that make up their lives, whether they are Yanomami Indians or family therapists. It can provide descriptions which pick out abstract organizations of life in a setting as well as allowing the reader rich access. Ethnography can be used in theory development and even theory testing (Hammersley & Atkinson, 1995). It is flexible; research can follow up themes and questions as they arise rather than necessarily keeping to preset goals.

Ethnographic research can be very time consuming and labor intensive. It can also be very intrusive. Although Gubrium was able to participate in some aspects of family therapy, this was helped by the sheer number of both staff and family members who were involved. It is less easy to imagine participant observation on individual therapy.

One of the most important difficulties with ethnographic work is that the reader often has to take on trust the conclusions because the evidence on which they are based is not available for assessment (Silverman, 1993). Where field notes of observations are reproduced in ethnographies—and this is relatively rare—such notes are nevertheless a ready-theorized version of events. Descriptions of actions and events are always bound up with a range of judgments (Potter, 1996a). Where analysis depends on the claims of key informants the problem is assessing how these claims relate to any putative activities that are described. Ethnographers deal with these problems with varying degrees of sophistication (for discussion see Nelson, 1994). However, some researchers treat them as inescapable and have turned to some form of discourse analysis instead.

For more detailed discussions of ethnography readers should start with Fielding's (1993) excellent brief introduction and then use Hammersley and Atkinson (1995) as an authoritative and up to date overview. Two of the most comprehensive, although not always most sophisticated, works are Werner and Schoepfle (1987) and Ellen (1984). Both were written by anthropologists, and this shows in their understanding of what is important. Newfield, Sells, Smith, Newfield, and Newfield (1996) provide a useful discussion of ethnography in family therapy research.

3.06.4 DISCOURSE ANALYSIS

Although both grounded theorizing and ethnographic work in clinical areas has increased, the most striking expansion has been in research in discourse analysis (Soyland, 1995). This work is of variable quality and often done by researchers isolated in different subdisciplines; moreover, it displays considerable terminological confusion. For simplicity, discourse analysis is taken as covering a range of work which includes conversation analysis and ethnomethodology (Heritage, 1984; Nofsinger, 1991), some specific traditions of discourse analysis and discursive psychology (Edwards & Potter, 1992a; Potter & Wetherell, 1987), some of the more analytically focused social constructionist work (McNamee & Gergen, 1992), and a range of work influenced by post-structuralism, Continental discourse analysis, and particularly the work of Foucault (Burman, 1995; Madigan, 1992; Miller & Silverman, 1995). In some research these different themes are woven together; elsewhere strong oppositions are marked out.

The impetus for discourse analysis in clinical settings comes from two directions. On the one hand, there are practitioner/researchers who have found ideas from social constructionism, literary theory, and narrative useful (e.g., Anderson & Goolishian, 1988; White & Epston, 1990). On the other, there are academic researchers who have extended analytic and theoretical developments in discourse studies to clinical settings (e.g., Aronsson & Cederborg, 1996; Bergmann, 1992; Buttny, 1996; Edwards, 1995; Lee, 1995). Collections such as Siegfried (1995), Burman, Mitchel, and Salmon (1996), and Morris and Chenail (1995) reflect both types of work, sometimes in rather uneasy combination (see Antaki, 1996).

This tension between an applied and academic focus is closely related to the stance taken on therapy. In much discourse analysis, therapy is the start point for research and the issue is how therapy gets done. For example, Gale's (1991; Gale & Newfield, 1992) intensive study of one of O'Hanlon's therapy sessions considered the various ways in which the goals of solution focused family therapy were realized in the talk between therapist and client. However, some conversation analysts and ethnomethodologists resist assuming that conversational interaction glossed by some parties as therapy (solution focused, Milan School, or whatever) must have special ingredient X—therapy—that is absent in, say, the everyday "troubles talk" done with a friend over the telephone (Jefferson, 1988; Jefferson & Lee, 1992; Schegloff, 1991; Watson, 1995).

This is a significant point for all researchers into therapy and counseling, so it is worth illustrating with an example. In Labov and Fanshel's classic study, the therapy session starts in the following manner:

Rhoda:	I don't (1.0) know, whether (1.5) I-I think I did- the right thing, jistalittle situation came up (4.5) an' I tried to uhm (3.0) well, try to (4.0) use what I- what I've learned here, see if it worked (0.3)
Therapist:	Mhm
Rhoda:	Now, I don't know if I did the right thing. Sunday (1.0) um- my mother went to my sister's again.
Therapist:	Mm-hm
Rhoda:	And she usu'lly goes for about a day or so, like if she leaves on Sunday, she'll come back Tuesday morning. So- it's nothing. But- she lef' Sunday, and she's still not home.
Therapist:	O- oh. (1977, p. 263)

Labov and Fanshel provide many pages of analysis of this sequence. They identify various direct and indirect speech acts and make much of what they call its therapeutic interview style, particularly the vague reference terms at the start: "right thing" and "jistalittle situation." This vagueness can easily be heard as the 19-year-old anorexia sufferer struggling to face up to her relationship with her difficult mother. However, in a reanalysis from a conversation analytic perspective, Levinson (1983) suggests that this sequence is characteristic of mundane news telling sequences in everyday conversation. These typically have four parts: the pre-announcement, the go ahead, the news telling, and the news receipt. For example:

D:	I forgot to to tell you the two best things that happen' to me today.	preannouncement
R:	Oh super = what were they.	go ahead turn
D:	I got a B+ on my math test . . . and I got an athletic award	news telling
R:	Oh excellent.	news receipt

(Levinson, 1983, p. 349—slightly modified)

A particular feature of preannouncements is their vagueness, for their job is to prefigure the story (and thereby check its newsworthiness), not to actually tell it. So, rather than following Labov and Fanshel (1977) in treating this vagueness as specific to a troubled soul dealing with a difficult topic in therapy, Levinson (1983) proposes that it should be understood as a commonplace characteristic of mundane interaction.

This example illustrates a number of features typical of a range of discursive approaches to therapy talk. First, the talk is understood as performing actions; it is not being used as a pathway to cognitive processes (repression, say) or as evidence of what Rhoda's life is like (her difficult mother). Second, the interaction is understood as sequentially organized so any part of the talk relates to what came immediately before and provides an environment for what will immediately follow. The realization of how far interaction gets its sense from its sequential context has critical implications for approaches such as content analysis and grounded theory which involve making categorizations and considering relations between them; for such categorizations tend to cut across precisely the sequential relations that are important for the sense of the turn of talk.

The third feature is that the talk is treated as ordered in its detail not merely in its broad particulars. For example, Levinson (1983) highlights a number of orderly elements in what we might easily mistake for clumsiness in Rhoda's first turn:

> R's first turn is . . . formulated to prefigure (i) the telling of something she did (*I think I did the right thing*), and (ii) the describing of the situation that led to the action (*jistalittle situation came up*). We are therefore warned to expect a story with two components; moreover the point of the story and its relevance to the here and now is also prefigured (*use what I've learned here, see if it worked*). (1983, p. 353)

Even the hesitations and glottal stops in Rhoda's first turn, which seem so redolent of a troubled young person are "typical markings of self-initiated self-repair, which is characteristic of the production of first topics" (Levinson, 1983, p. 353). This emphasis on the significance of detail has an important methodological consequence—if interaction is to be understood properly it must be represented in a way that captures this detail. Hence the use of a transcription scheme that attempts to represent a range of paralinguistic features of talk (stress, intonation) on the page as well as elements of the style of its delivery (pauses, cut-offs).

A fourth feature to note here is the comparative approach that has been taken. Rather than focus on therapy talk alone Levinson is able to support an alternative account of the interaction by drawing on materials, and analysis, taken from mundane conversations. Since the mid 1980s there has been a large amount of work in different institutional settings as well as everyday conversation, and it is now possible to start to show how a news interview, say, differs from the health visitor's talk with an expectant mother, and how that differs in turn from conversation between two friends over the telephone (Drew & Heritage, 1992a).

A fifth and final feature of this example is that it is an analysis of interaction. It is neither an attempt to reduce what is going on to cognitions of the various parties—Rhoda's denial, say, or the therapist's eliciting strategies—nor to transform it into countable behaviors such as verbal reinforcers. This style of discourse work develops a Wittgensteinian notion of cognitive words and phrases as elements in a set of language games for managing issues of blame, accountability, description, and so on (Coulter, 1990; Edwards, 1997; Harré & Gillett, 1994). Such a "discursive psychology" analyzes notions such as attribution and memory in terms of the situated practices through which responsibility is assigned and the business done by constructing particular versions of past events (Edwards & Potter, 1992b, 1993).

3.06.4.1 Questions

Discourse analysis is more directly associated with particular theoretical perspectives—

ethnomethodology, post-structuralism, discursive psychology—than either grounded theory or ethnography. The questions it addresses focus on the practices of interaction in their natural contexts and the sorts of discursive resources that are drawn on those contexts.

Some of the most basic questions concern the standardized sequences of interaction that take place in therapy and counseling (Buttny & Jensen, 1995; Lee, 1995; Peräkylä, 1995; Silverman, 1997b). This is closely related to a concern with the activities that take place. What is the role of the therapist's tokens such as "okay" or "mm hm" (Beach, 1995; Czyzewski, 1995)? How do different styles of questioning perform different tasks (Bergmann, 1992; Peräkylä, 1995; Silverman, 1997b)? What is the role of problem formulations by both therapists and clients, and how are they transformed and negotiated (Buttny, 1996; Buttny & Jensen, 1995; Madill & Barkham, 1997)? For example, in a classic feminist paper Davis (1986) charts the way a woman's struggles with her oppressive social circumstances are transformed into individual psychological problems suitable for individual therapy. While much discourse research is focused on the talk of therapy and counseling itself, studies in other areas show the value of standing back and considering clinical psychology as a set of work practices in themselves, including management of clients in person and as records, conducting assessments, delivering diagnoses, intake and release, stimulating people with profound learning difficulties, case conferences and supervisions, offering advice, and managing resistance (see Drew & Heritage, 1992a). Discourse researchers have also been moving beyond clinical settings to study how people with clinical problems or learning difficulties manage in everyday settings (Brewer & Yearley, 1989; Pollner & Wikler, 1985; Wootton, 1989).

Another set of questions are suggested by the perspective of discursive psychology. For example, Edwards (1997) has studied the rhetorical relationship between problem formulations, descriptions of activities, and issues of blame in counseling. Cheston (1996) has studied the way descriptions of the past in a therapy group of older adults can create a set of social identities for the members. Discursive psychology can provide a new take on emotions, examining how they are constructed and their role in specific practices (Edwards, 1997).

From a more Foucaultian inspired direction, studies may consider the role of particular discourses, or interpretative repertoires in constructing the sense of actions and experiences (Potter & Wetherell, 1987). For example, what are the discursive resources that young women draw on to construct notions of femininity, agency, and body image in the context of eating disorders (Malson & Ussher, 1996; Wetherell, 1996)? What discourses are used to construct different notions of the person, of the family, and of breakdown in therapy (Burman, 1995; Frosh, Burck, Strickland-Clark, & Morgan, 1996; Soal & Kotter, 1996)? This work is often critical of individualistic conceptions of therapy.

Finally, discourse researchers have stood back and taken the administration of psychological research instruments as their topic. The intensive focus of such work can show the way that the sort in idiosyncratic interaction that takes place when filling in questionnaires or producing records can lead to particular outcomes (Antaki & Rapley, 1996; Garfinkel, 1967b; Rapley & Antaki, 1996; Soyland, 1994).

Different styles of discourse work address rather different kinds of questions. However, the conversation analytic work is notable in commonly starting from a set of transcribed materials rather than preformulated research questions, on the grounds that such questions often embody expectations and assumptions which prevent the analyst seeing a range of intricacies in the interaction. Conversation analysis reveals an order to interaction that participants are often unable to formulate in abstract terms.

3.06.4.2 Procedures

The majority of discourse research in the clinical area has worked with records of natural interaction, although a small amount has used open-ended interviews. There is not space here to discuss the role of interviews in discourse analysis or qualitative research generally (see Kvale, 1996; Mischler, 1986; Potter & Mulkay, 1985; Widdicombe & Wooffitt, 1995). For simplicity discourse work will be discussed in terms of seven elements.

3.06.4.2.1 Research materials

Traditionally psychologists have been reluctant to deal with actual interaction, preferring to model it experimentally, or reconstruct it via scales and questionnaires. Part of the reason for this is the prevalent cognitivist assumptions which have directed attention away from interaction itself to focus on generative mechanisms within the person. In contrast, discourse researchers have emphasised the primacy of practices of interaction themselves. The most obvious practice setting for clinical work is the therapy session itself, and this has certainly received the most attention. After all, there is an elegance in studying the "talking cure" using methods designed to acquire an understanding of the nature of talk. However, there is a danger that such an exclusive emphasis underplays mundane aspects of clinical practices: giving advice, offering a diagnosis, the reception of new clients, casual talk to clients' relatives, writing up clinical records, case conferences, clinical training, and assessment.

Notions of sample size do not translate easily from traditional research as the discourse research focus is not so much on individuals as on interactional phenomena of various kinds. Various considerations can come to the fore here, including the type of generalizations that are to be made from the research, the time and resources available, and the nature of the topic being studied. For example, if the topic is the role of "mm hms" in therapy a small number of sessions may generate a large corpus; other phenomena may be much rarer and require large quantities of interaction to develop a useful corpus. For some questions, single cases may be sufficient to underpin a theoretical point or reveal a theoretically significant phenomena.

3.06.4.2.2 Access and ethics

One of the most difficult practical problems in conducting discourse research involves getting access to sometimes sensitive settings in ways which allow for informed consent from all the parties involved. Experience suggests that more often than not it is the health professionals rather than the clients who are resistant to their practices being studied, perhaps because they are sensitive to the difference between the idealized version of practices that was used in training and the apparently more messy procedures in which they actually engage. Sometimes reassurances about these differences can be productive.

Using records of interaction such as these raise particular issues for ensuring anonymity. This is relatively manageable with transcripts where names and places can be changed. It is harder for audio tape and harder still with video. However, technical advances in the use of digitized video allow for disguising of identity with relatively little loss of vocal information.

3.06.4.2.3 Audio and video recording

There is a range of practical concerns in recording natural interaction, some of them pulling in different directions. An immediate issue is whether to use audio or video recording. On the one hand, video provides helpful information about nonverbal activities that may be missed from the tape, and good quality equipment is now compact and cheap. On the other hand, video can be more intrusive, particularly where the recording is being done by one of the participants (a counselor, say), and may be hard to position so it captures gestures and expressions from all parties to an interaction. Video poses a range of practical and theoretical problems with respect to the transcription of nonvocal activity which can be both time consuming and create transcript that is difficult to understand. Moreover, there is now a large body of studies that shows high quality analysis can, in many cases, be performed with an audio tape alone. One manageable solution is to use video if doing so does not disrupt the interaction, and then to transcribe the audio and work with a combination of video tape and audio transcript. Whether audio or video is chosen, the quality (clear audibility and visibility) is probably the single most consequential feature of the recording for the later research.

Another difficulty is how far the recording of interaction affects its nature. This is a subtle issue. On the one hand, there are a range of ways of minimizing such influences including acclimatizing participants and giving clear descriptions of the role of the research. On the other, experience has shown that recording has little influence on many, perhaps most, of the activities in which the discourse researcher is interested. Indeed, in some clinical settings recordings may be made as a matter of course for purposes of therapy and training, and so no new disruption is involved.

3.06.4.2.4 Transcription

Producing a high-quality transcript is a crucial prerequisite for discourse research. A transcript is not a neutral, simple rendition of the words on a tape (Ochs, 1979). Different transcription systems emphasize different features of interaction. The best system for most work of this kind was developed by the conversation analyst Gail Jefferson using symbols that convey features of vocal delivery that have been shown to be interactionally important to participants (Jefferson, 1985). At the same time the system is designed to use characters and symbols easily available on wordprocessors making it reasonably easy to learn and interpret. The basic system is summarized in Table 3. For fuller descriptions of using the system see Button and Lee, (1987), Ochs, Schegloff, and Thompson (1996), and Psathas and Anderson (1990).

Producing high quality transcript is very demanding and time consuming. It is hard to give a standard figure for how long it takes

Table 3 Brief transcription notation.

Um::	colons represent lengthening of the preceding sound; the more colons, the greater the lengthening.
I've-	a hyphen represents the cut-off of the preceding sound, often by a stop.
↑Already	up and down arrows represent sharp upward and downward pitch shifts in the following sounds. Underlining represents stress, usually by volume; the more underlining the greater the stress.
settled in his = Mm = own mind.	the equals signs join talk that is continuous although separated across different lines of transcript.
hhh hh .hh P(h)ut	'h' represents aspiration, sometimes simply hearable breathing, sometimes laughter, etc.; when preceded by a superimposed dot, it marks in-breath; in parenthesis inside a word it represents laugh infiltration.
hhh[hh .hh] [I just]	left brackets represent point of overlap onset; right brackets represent point of overlap resolution.
.,?	punctuation marks intonation, not grammar; period, comma and 'question mark' indicate downward, 'continuative', and upward contours respectively.
( )	single parentheses mark problematic or uncertain hearings; two parentheses separated by an oblique represent alternative hearings.
(0.2)(.)	numbers in parentheses represent silence in tenths of a second; a dot in parentheses represents a micro-pause, less than two tenths of a second.
°mm hmm°	the degree signs enclose significantly lowered volume.

Source: Modified from Schegloff (1997, pp. 184–185).

because much depends on the quality of the recording (fuzzy, quiet tapes can quadruple the time needed) and the type of interaction (an individual therapy session presents much less of a challenge than a lively case conference with a lot of overlapping talk and extraneous noise); nevertheless, a ratio of one hour of tape to twenty hours of transcription time is not unreasonable. However, this time should not be thought of as dead time before the analysis proper. Often some of the most revealing analytic insights come during transcription because a profound engagement with the material is needed to produce good transcript—it is generally useful to make analytic notes in parallel to the actual transcription.

3.06.4.2.5 Coding

In discourse research the principle task of coding is to make the task of analysis more straightforward by sifting relevant materials from a large body of transcript. In this it differs from both grounded theory and traditional content analysis where coding is a more intrinsic part of the analysis. Typically coding will involve sifting through materials for instances of a phenomenon of interest and copying them into an archive. This coding will often be accompanied by preliminary notes as to their nature and interest. At this stage selection is inclusive—it is better to include material that can turn out to be irrelevant at a later stage than exclude it for ill-formulated reasons early on. Coding is a cyclical process. Full analysis of a corpus of materials can often take the researcher back to the originals as a better understanding of the phenomenon reveals new examples. Often initially disparate topics merge together in the course of analysis while topics which seemed unitary can be separated.

3.06.4.2.6 Analysis

There is no single recipe for analyzing discourse. Nevertheless, there are five considerations which are commonly important in analysis. First, the researcher can use variation in and between participants' discourse as an analytic lever. The significance of variation is that it can be used for identifying and explicating the activities that are being performed by talk and texts. This is because the discourse is constructed in the specific ways that it is precisely to perform these actions; a description of jealousy in couple counseling can be assembled very differently when excusing or criticizing certain actions (Edwards, 1995). The researcher will benefit from attending to variations in the discourse of single individuals, between different individuals, and between what is said and what might have been said.

Second, discourse researchers have found it highly productive to attend to the detail of discourse. Conversation analysts such as Sacks (1992) have shown that the details in discourse—the hesitations, lexical choice, repair, and so on—are commonly part of the performance of some act or are consequential in some way for the outcome of the interaction. Attending to the detail of interaction, particularly in transcript, is one of the most difficult things for psychologists who are used to reading through the apparently messy detail for the gist of what is going on. Developing analytic skills involves a discipline of close reading.

Third, analysis often benefits from attending to the rhetorical organization of discourse. This involves inspecting discourse both for the way it is organized to make argumentative cases and for the way it is designed to undermine alternative cases (Billig, 1996). A rhetorical orientation refocuses the analyst's attention away from questions about how a version—description of a psychological disorder, say—relates to some putative reality and focuses it on how it relates to competing alternatives.

Concern with rhetoric is closely linked to a fourth analytic concern with accountability. That is, displaying one's activities as rational, sensible, and justifiable. Ethnomethodologists have argued that accountability is an essential and pervasive character of the design and understanding of human conduct generally (Garfinkel, 1967c; Heritage, 1984). Again an attention to the way actions are made accountable is an aid for understanding precisely what those actions are.

A fifth and final analytic consideration is of a slightly different order. It is an emphasis on the virtue of building on prior analytic studies. In particular, researchers into interaction in an institutional setting such as a family therapy setting will almost certainly benefit from a familiarity with research on mundane talk as well as an understanding of how the patterning of turn taking and activities change in different institutional settings.

The best way to think of these five considerations is not as specific rules for research but as elements in an analytic mentality that the researcher will develop as they become more and more skilled. It does not matter that they are not spelled out in studies because they are separate from the procedures for validating discourse analytic claims.

3.06.4.2.7 Validity

Discourse researchers typically draw on some combination of four considerations to justify the validity of analytic claims. First, they make use of participants' own understandings as they are displayed in interaction. One of the features of a conversation is that any turn of talk is oriented to what came before and what comes next, and that orientation typically displays the sense that the participant makes of the prior turn. Thus, at its simplest, when someone provides an answer they thereby display the prior turn as a question and so on. Close attention to this turn-by-turn display of understanding provides one important check on analytic interpretations (see Heritage, 1988).

Second, researchers may find (seemingly) deviant cases most useful in assessing the adequacy of a claim. Deviant cases may generate problems for a claimed generalization, and lead the researcher to abandon it; but they may also display in their detailed organization precisely the reason why a standard pattern should take the form that it does.

Third, a study may be assessed in part by how far it is coherent with previous discourse studies. A study that builds coherently on past research is more plausible than one that is more anomalous.

Fourth, and most important, are readers' evaluations. One of the distinctive features of discourse research is its presentation of rich and extended materials in a way that allows the reader to make their own judgements about interpretations that are placed along side of them. This form of validation contrasts with much grounded theory and ethnography where interpretations often have to be taken on trust; it also contrasts with much traditional experimental and content analytic work where it is rare for "raw" data to be included or more than one or two illustrative codings to be reproduced.

Whether they appear singly or together in a discourse study none of these procedures guarantee the validity of an analysis. However, as the sociology of science work reviewed earlier shows, there are no such guarantees in science.

3.06.4.3 Example: Peräkylä (1995)

Given the wide variety of discourse studies with different questions and styles of analysis it is not easy to chose a single representative study. The one selected is Peräkylä's (1995) investigation of AIDS counseling because it is a major integrative study that addresses a related set of questions about interaction, counseling, and family therapy from a rigorous conversation analytic perspective and illustrates the potential of discourse work on clinical topics. It draws heavily on the perspective on institutional talk surveyed by Drew and Heritage (1992b) and is worth reading in conjunction with Silverman's (1997b) related study of HIV+ counseling which focuses more on advice giving.

Peräkylä focused on 32 counseling sessions conducted with HIV+ hemophilic mainly gay identified men and their partners at a major London hospital. The sessions were videotaped and transcribed using the Jeffersonian system. A wider archive of recordings (450 hours) was drawn on to provide further examples of phenomena of interest but not otherwise transcribed. The counselors characterized their practices in terms of Milan School Family Systems Theory and, although this is not the startpoint of Peräkylä's study, he was able to explicate some of the characteristics of such counseling.

Part of the study is concerned with identifying the standard turn-taking organization of the counseling. Stated baldly it is that (i) counselors ask questions; (ii) clients answer; (iii) counselors comment, advise, or ask further questions. When laid out in this manner the organization may not seem much of a discovery. However, the power of the study is showing how this organization is achieved in the interaction and how it can be used to address painful and delicate topics such as sexual behavior, illness, and death.

Peräkylä goes on to examine various practices that are characteristic of family systems theory such as "circular questioning," where the counselor initially questions the client's partner or a family member about the client's feelings, and "live open supervision," where a supervisor may offer questions to the counselor that are, in turn, addressed to the client. The study also identifies some of the strategies by which counselors can address "dreaded issues" in a manageable way. Take "circular questioning," for example. In mundane interaction providing your partial experience of some event or experience is a commonplace way of "fishing" for a more authoritative version (Pomerantz, 1980). For example:

A: Yer line's been busy.
B: Yeuh my fu (hh)- .hh my father's
wife called me

In a similar way, the use of questioning where a client's partner, say, offers their understanding of an experience "can create a situation where the clients, in an unacknowledged but most powerful way, elicit one another's descriptions of their inner experiences" (Peräkylä, 1995, p. 110). In the following extract the client is called Edward; his partner and the counselor are also present.

Counselor: What are some of things that you
think E:dward might have to
do. = He says he doesn't know
where to go from here maybe: and
awaiting results and things.
(0.6)
Counselor: What d'you think's worrying him.
(0.4)
Partner: Uh::m hhhhhh I think it's just fear
of the unknow:n.
Client: Mm[:
Counselor: [Oka:y.
Partner: [At- at the present ti:me. (0.2)
Uh:m (.) once: he's (0.5) got a better
understanding of (0.2) what could
happen
Counselor: Mm:
Partner: uh:m how .hh this will progre:ss
then: I think (.) things will be a
little more [settled in his =
Counselor: [Mm
Partner: = own mi:nd.
Counselor: Mm:
(.)
Client: Mm[:
Counselor: [Edward (.) from what you
know::
((sequence continues with Edward
responding to a direct question
with a long and detailed narrative
about his fears))
(Peräkylä, 1995, p. 110)

Peräkylä emphasizes the way that the client's talk about his fears is elicited, in part, through the counsellor asking the partner for his own view of those fears. The point is not that the client is forced to reveal his experiences, rather it is that the prior revelation of his partner's partial view produces an environment where such a revelation is expected and nonrevelation will itself be a delicate and accountable matter. In effect, what Peräkylä is documenting here are the conversational mechanisms which family therapists characterize as using circular questioning to overcome clients' resistance.

3.06.4.4 Virtues and Limitations

Given the variety of styles of work done under the rubric of discourse analysis it is difficult to give an overall summary of virtues and limitations. However, the virtue of a range of studies in the conversation and discourse analytic tradition is that they offer, arguably for the first time in psychology, a rigorous way of directly studying human social practices. For example, the Peräkylä study discussed above is notable in studying actual HIV + counseling in all its detail. It is not counseling as recollected by participants while filling in rating scales or questionnaires; it is not an experimental simulation of counseling; it does not depend on *post hoc* ethnographic reconstructions of events; nor are the activities immediately transformed into broad coding categories or used as a mere shortcut to underlying cognitions.

A corollary of this emphasis on working with tapes and transcripts of interaction is that these are used in research papers to allow readers to evaluate the adequacy of interpretations in a way that is rare in other styles of research.

Studies in this tradition have started to reveal an organization of interaction and its local management that has been largely missed from traditional psychological work from a range of perspectives. Such studies offer new conceptions of what is important in clinical practice and may be particularly valuable in clinical training which has often been conducted with idealized or at least cleaned up examples of practice.

Discourse research is demanding and requires a considerable investment of the researcher's time to produce satisfactory results. It does not fit neatly into routines that can be done by research assistants. Indeed, even transcription, which may seem to be the most mechanical element in the research, requires considerable skill and benefits from the involvement of the primary researchers. Analysis also requires considerable craft skills which can take time to learn.

With its focus on interaction, this would not necessarily be the perspective of choice for researchers with a more traditional cognitive or behavioral focus, although it has important implications for both of these. Some have claimed that it places too much emphasis on verbal interaction at the expense of nonverbal elements, and broader issues of embodiment. Others have claimed that it places too much emphasis on the importance of local contexts of interaction rather than on broader issues such as gender or social class. For some contrasting and strongly expressed claims about the role of discourse analysis in the cognitive psychology of memory, see papers in Conway (1992).

An accessible general introduction to various practical aspects of doing discourse analysis is provided in Potter and Wetherell (1987; see also Potter, 1996b, 1997). Potter and Wetherell (1995) discuss the analysis of broader content units such as interpretative repertoires, while Potter and Wetherell (1994) and Wooffitt (1993) discuss the analysis of how accounts are constructed. For work in the distinctive conversation analytic tradition Drew (1995) and Heritage (1995) provide clear overviews and Heath and Luff (1993) discuss analysis which incorporates video material; Gale (1996) explores the use of conversation analysis in family therapy research.

3.06.5 FUTURE DIRECTIONS

The pace of change in qualitative research in clinical settings is currently breathtaking, and it is not easy to make confident predictions. However, it is perhaps useful to speculate on how things might develop over the next few years. The first prediction is that the growth in the sheer quantity of qualitative research will continue for some time. There is so much new territory, and so many possibilities have been opened up by new theoretical and analytic developments, that they are bound to be explored.

The second prediction is that research on therapy and counseling talk will provide a particular initial focus because it is here that discourse analytic approaches can clearly provide new insights and possibly start to provide technical analytically grounded specifications of the interactional nature of different therapies in practice, as well as differences in interactional style between therapists. There may well be conflicts here between the ideological goals of constructionist therapists and the research goals of discourse analysts.

The third prediction is that the growth of qualitative work will encourage more researchers to attempt integrations of qualitative and quantitative research strategies. There will be attempts to supplement traditional outcomes research with studies of elements of treatment which are not easily amenable to quantification. Here the theoretical neutrality of grounded theory (ironically) is likely to make for easier integration than the more theoretically developed discourse perspectives. The sheer difficulty of blending qualitative and quantitative work should not be underestimated—research that has attempted this has often found severe problems (see Mason, 1994, for a discussion).

The final prediction is that there will be an increased focus on clinical work practices embodied within settings such as clinics and networks of professional and lay relationships. Here the richness of ethnographic work will be drawn on, but increasingly the conversation analytic discipline of working with video and transcript will replace field notes and recollections. Such work will have the effect of respecifying some of the basic problems of clinical research. Its broader significance, however, may depend on the course of wider debates in Psychology over the development and success of the cognitive paradigm and whether it will have a discursive and interaction based successor.

ACKNOWLEDGMENTS

I would like to thank Alan Bryman, Karen Henwood, Alexa Hepburn, Celia Kitzinger, and Denis Salter for making helpful comments on an earlier draft of this chapter.

3.06.6 REFERENCES

Anderson, H., & Goolishian, H. A. (1988). Human systems as linguistic systems: Preliminary and evolving ideas about the implications for clinical theory. *Family Process, 27,* 371–393.

Antaki, C. (Ed.) (1988). *Analysing everyday explanation.* London: Sage.

Antaki, C. (1996). Review of *The talk of the clinic. Journal of Language and Social Psychology, 15,* 176–81.

Antaki, C., & Rapley, M. (1996). "Quality of life" talk: The liberal paradox of psychological testing. *Discourse and Society, 7,* 293–316.

Argyris, C., Putnam, R., & Smith, D. M. (1985). *Action Science: Concepts, methods, and skills for research and intervention.* San Francisco: Jossey-Bass.

Aronsson, K., & Cederborg, A.-C. (1996). Coming of age in family therapy talk: Perspective setting in multiparty problem formulations. *Discourse Processes, 21,* 191–212.

Ashmore, M., Mulkay, M., & Pinch, T. (1989). *Health and efficiency: A sociological study of health economics.* Milton Keynes, UK. Open University Press.

Atkinson, J. M. (1978). *Discovering suicide: Studies in the social organization of sudden death.* London: Macmillan.

Atkinson, P. (1990). *The ethnographic imagination: The textual construction of reality.* London: Routledge.

Bannister, P., Burman, E., Parker, I., Taylor, M., & Tindall, C. (1994). *Qualitative methods in psychology: A research guide.* Buckingham, UK: Open University Press.

Beach, W. A. (1995). Preserving and constraining options: "Okays" and "official" priorities in medical interviews. In G. H. Morris & R. J. Chenail (Eds.), *The talk of the clinic: Explorations in the analysis of medical and therapeutic discourse* (pp. 259–289). Hillsdale, NJ: Erlbaum.

Bergmann, J. R. (1992). Veiled morality: Notes on discretion in psychiatry. In P. Drew & J. Heritage (Eds.), *Talk at work: Interaction in institutional settings* (pp. 137–162). Cambridge, UK: Cambridge University Press.

Billig, M. (1988). Methodology and scholarship in understanding ideological explanation. In C. Antaki (Ed.), *Analysing everyday explanation: A casebook of methods* (pp. 199–215). London: Sage.

Billig, M. (1996). *Arguing and thinking: A rhetorical approach to social psychology* (2nd ed.). Cambridge, UK: Cambridge University Press.

Billig, M. (1998). Dialogic repression and the Oedipus Complex: Reinterpreting the Little Hans case. *Culture and Psychology, 4,* 11–47.

Bogdan, R., & Taylor, S. J. (1975). *Introduction to qualitative research methods: A phenomenological approach to social sciences.* New York: Wiley.

Brewer, J. D., & Yearley, S. (1989). Stigma and conversational competence: A conversation-analytic study of the mentally handicapped. *Human Studies, 12,* 97–115

Bryman, A. (1988). *Quantity and quality in social research.* London: Unwin Hyman.

Bryman, A., & Burgess, R. G. (Eds.) (1994). *Analyzing qualitative data.* London: Routledge.

Burman, E. (1995). Identification, subjectivity, and power in feminist psychotherapy. In J. Siegfried (Ed.), *Therapeutic and everyday discourse as behaviour change: Towards a micro-analysis in psychotherapy process research* (pp. 469–490). Norwood, NJ: Ablex.

Burman, E., Mitchell, S., & Salmon, P. (Eds.) (1996). *Changes: An International Journal of Psychology and Psychotherapy (special issue on qualitative methods), 14,* 175–243.

Buttny, R. (1996). Clients' and therapists' joint construction of the clients' problems. *Research on Language and Social Interaction, 29,* 125–153.

Buttny, R., & Jensen, A. D. (1995) Telling problems in an initial family therapy session: The hierarchical organization of problem-talk. In G. H. Morris & R. J. Chenail (Eds.), *The talk of the clinic: Explorations in the analysis of medical and therapeutic discourse* (pp. 19–48). Hillsdale, NJ: Erlbaum.

Button, G., & Lee, J. R. E. (Eds.) (1987). *Talk and social organization.* Clevedon, UK: Multilingual Matters.

Chalmers, A. (1992). *What is this thing called science?: An assessment of the nature and status of science and its methods* (2nd ed.), Milton Keynes, UK: Open University Press.

Charmaz, K. (1991). *Good days, bad days: The self in chronic illness and time.* New Brunswick, NJ: Rutgers University Press.

Charmaz, K. (1994). Identity dilemmas of chronically ill men. *The Sociological Quarterly, 35,* 269–288.

Charmaz, K. (1995). Grounded theory. In J. A. Smith, R. Harré, & L. van Langenhove (Eds.), *Rethinking methods in pychology* (pp. 27–49). London: Sage.

Cheston, R. (1996). Stories and metaphors: Talking about the past in a psychotherapy group for people with dementia. *Ageing and Society, 16,* 579–602.

Clegg, J. A., Standen, P. J., & Jones, G. (1996). Striking the balance: A grounded theory analysis of staff perspectives. *British Journal of Clinical Psychology, 35,* 249–264.

Coffey, A., & Atkinson, P. (1996). *Making sense of qualitative data: Complementary research strategies.* London: Sage.

Collins, H. M. (1974). The TEA Set: Tacit knowledge and scientific networks. *Science Studies, 4,* 165–186.

Collins, H. M. (Ed.) (1981). Knowledge and controversy: Studies of modern natural science. *Social Studies of Science (special issue),* 11.

Collins, H. M. (1983). The meaning of lies: Accounts of action and participatory research. In G. N. Gilbert & P. Abell (Eds.), *Accounts and action* (pp. 69–76). Aldershot, UK: Gower.

Collins, H. M. (1985). *Changing order: Replication and induction in scientific practice.* London: Sage.

Collins, H. M., & Pinch, T. (1993) *The Golem: What everyone should know about science.* Cambridge, UK: Cambridge University Press.

Conway, M. (Ed.) (1992). Developments and debates in the study of human memory (special issue). *The Psychologist, 5,* 439–461.

Coulter, J. (1990). *Mind in action.* Cambridge, UK: Polity.

Czyzewski, M. (1995). "Mm hm" tokens as interactional devices in the psychotherapeutic intake interview. In P. ten Have & G. Psathas (Eds.), *Situated order: Studies in the social organization of talk and embodied activities* (pp. 73–89). Washington, DC: International Institute for

Ethnomethodology and Conversation Analysis & University Press of America.

Davis, K. (1986). The process of problem (re)formulation in psychotherapy. *Sociology of Health and Illness, 8,* 44–74.

Denzin, N. K., & Lincoln, Y. S. (Eds.) (1994) *Handbook of qualitative research.* London: Sage.

Drew, P. (1995). Conversation analysis. In J. Smith, R. Harré, & L. van Langenhove (Eds.), *Rethinking methods in psychology* (pp. 64–79). London: Sage.

Drew, P., & Heritage, J. (Eds.) (1992a). *Talk at work: Interaction in institutional settings.* Cambridge, UK: Cambridge University Press.

Drew, P., & Heritage, J. (1992b). Analyzing talk at work: An introduction. In P. Drew & J. Heritage (Eds.), *Talk at work: Interaction in institutional settings* (pp. 3–65). Cambridge, UK: Cambridge University Press.

Edwards, D. (1995). Two to tango: Script formulations, dispositions, and rhetorical symmetry in relationship troubles talk. *Research on Language and Social Interaction, 28,* 319–350.

Edwards, D. (1997). *Discourse and cognition.* London: Sage.

Edwards, D., & Potter, J. (1992a). *Discursive psychology.* London: Sage.

Edwards, D., & Potter, J. (1992b). The chancellor's memory: Rhetoric and truth in discursive remembering. *Applied Cognitive Psychology, 6,* 187–215.

Edwards, D., & Potter, J. (1993). Language and causation: A discursive action model of description and attribution. *Psychological Review, 100,* 23–41.

Ellen, R. F. (1984). *Ethnographic research: A guide to general conduct.* London: Academic Press.

Ellis, C., Kiesinger, C., & Tillmann-Healy, L. M. (1997) Interactive interviewing: Talking about emotional experience. In R. Hertz (Ed.), *Reflexivity and Voice.* Thousand Oaks, CA: Sage.

Essex, M., Estroff, S., Kane, S., McLanahan, J., Robbins, J., Dresser, R., & Diamond, R. (1980). On Weinstein's "Patient attitudes toward mental hospitalization: A review of quantitative research." *Journal of Health and Social Behaviour, 21,* 393–396.

Fetterman, D. M. (1989). *Ethnography: Step by step.* London: Sage.

Fielding, N. (1993). Ethnography. In N. Gilbert (Ed.), *Researching social life* (pp. 154–171). London: Sage.

Freud, S. (1977). *Case histories. I: "Dora" and "Little Hans."* London: Penguin.

Frosh, S., Burck, C., Strickland-Clark, L., & Morgan, K. (1996). Engaging with change: A process study of family therapy. *Journal of Family Therapy, 18,* 141–161.

Gale, J. E. (1991). *Conversation analysis of therapeutic discourse: The pursuit of a therapeutic agenda.* Norwood, NJ: Ablex.

Gale, J. E. (1996). Conversation analysis: Studying the construction of therapeutic realities. In D. H. Sprenkle & S. M. Moon (Eds.), *Research methods in family therapy* (pp. 107–124). New York: Guilford.

Gale, J. E., & Newfield, N. (1992). A conversation analysis of a solution-focused marital therapy session. *Journal of Marital and Family Therapy, 18,* 153–165.

Garfinkel, H. (1967a). "Good" organizational reasons for "bad" clinical records. In H. Garfinkel (Ed.), *Studies in ethnomethodology* (pp. 186–207). Englewood Cliffs, NJ: Prentice-Hall.

Garfinkel, H. (1967b). Methodological adequacy in the quantitative study of selection criteria and selection practices in psychiatric outpatient clinics. In H. Garfinkel (Ed.), *Studies in ethnomethodology* (pp. 208–261). Englewood Cliffs, NJ: Prentice-Hall.

Garfinkel, H. (1967c). *Studies in ethnomethodology.* Englewood Cliffs, NJ: Prentice-Hall.

Gelsthorpe, L. (1992). Response to Martyn Hammersley's paper "On feminist methodology." *Sociology, 26,* 213–218.

Gergen, K. J. (1994). *Realities and relationships: Soundings in social construction.* Cambridge, MA: Harvard University Press.

Gilbert, G. N. (Ed.) (1993). *Researching social life.* London: Sage.

Gilbert, G. N., & Mulkay, M. (1984). *Opening Pandora's box: A sociological analysis of scientists' discourse.* Cambridge, UK: Cambridge University Press.

Glaser, B., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Chicago: Aldine.

Glaser, B., & Strauss, A. L. (1968). *Time for dying.* Chicago: Aldine.

Goffman, E. (1961). *Asylums: Essays on the social situation of mental patients and other inmates.* London: Penguin.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105–117). London: Sage.

Gubrium, J. F. (1992). *Out of control: Family therapy and domestic disorder.* London: Sage.

Hammersley, M. (1992). On feminist methodology. *Sociology, 26,* 187–206.

Hammersley, M., & Atkinson, P. (1995). *Ethnography: Principles in practice* (2nd ed.). London: Routledge.

Harré, R. (1986). *Varieties of realism.* Oxford, UK: Blackwell.

Harré, R. (1992). *Social being: A theory for social psychology* (2nd ed.). Oxford, UK: Blackwell.

Harré, R., & Gillett, G. (1994). *The discursive mind.* London: Sage.

Heath, C., & Luff, P. (1993) Explicating face-to-face interaction. In N. Gilbert (Ed.), *Researching social life* (pp. 306–326) London: Sage.

Henriques, J., Hollway, W., Irwin, C., Venn, C., & Walkerdine, V. (1984). *Changing the subject: Psychology, social regulation and subjectivity.* London: Methuen.

Henwood, K., & Nicolson, P. (Eds.) (1995). Qualitative research methods (special issue). *The Psychologist, 8,* 109–129.

Henwood, K., & Parker, I. (1994). Qualitative social psychology (special issue). *Journal of Community and Applied Social Psychology, 4,* 219–223.

Henwood, K., & Pidgeon, N. (1994). Beyond the qualitative paradigm: A framework for introducing diversity within qualitative psychology. *Journal of Community and Applied Social Psychology, 4,* 225–238.

Henwood, K., & Pidgeon, N. (1995). Grounded theory and psychological research. *The Psychologist, 8,* 115–118.

Heritage, J. C. (1984). *Garfinkel and ethnomethodology.* Cambridge, UK: Polity.

Heritage, J. C. (1988). Explanations as accounts: A conversation analytic perspective. In C. Antaki (Ed.), *Analysing everyday explanation: A casebook of methods* (pp. 127–144). London: Sage.

Heritage, J. C. (1995). Conversation analysis: Methodological aspects. In U. Quasthoff (Ed.), *Aspects of oral communication.* (pp. 391–418). Berlin, Germany: De Gruyter.

Hesse, M. B. (1974). *The structure of scientific inference.* London: Macmillan.

Hodder, I. (1994). The interpretation of documents and material culture. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 395–402). London: Sage.

Holsti, O. R. (1969). *Content analysis for the social sciences and humanities.* Reading, MA: Addison-Wesley.

Jasanoff, S., Markle, G., Pinch T., & Petersen, J. (Eds.) (1995). *Handbook of science and technology studies.* London: Sage.

Jefferson, G. (1985). An exercise in the transcription and analysis of laughter. In T. van Dijk (Ed.), *Handbook of discourse analysis* (Vol. 3, pp. 25–34). London: Academic Press.

Jefferson, G. (1988). On the sequential organization of troubles-talk in ordinary conversation. *Social Problems, 35,* 418–441.

Jefferson, G., & Lee, J. R. E. (1992). The rejection of advice: Managing the problematic convergence of a "troubles-telling" and a "service encounter." In P. Drew & J. Heritage (Eds.), *Talk at work: Interaction in institutional settings* (pp. 521–548). Cambridge, UK: Cambridge University Press.

Jodelet, D. (1991). *Madness and social representations.* London: Harvester/Wheatsheaf.

Johnson, J. C. (1990). *Selecting ethnographic informants.* London: Sage.

Knorr Cetina, K. (1995). Laboratory studies: The cultural approach to the study of science. In S. Jasanoff, G. Markle, T. Pinch, & J. Petersen (Eds.), *Handbook of science and technology studies.* London: Sage.

Knorr Cetina, K. (1997). *Epistemic cultures: How scientists make sense.* Chicago: Indiana University Press.

Krippendorff, K. (1980). *Content analysis: An introduction to its methodology.* London: Sage.

Krueger, R. A. (1988). *Focus groups: A practical guide for applied research.* London: Sage.

Kuhn, T. S. (1977). *The essential tension: Selected studies in scientific tradition and change.* Chicago: University of Chicago Press.

Kvale, S. (1996). *InterViews: An introduction to qualitative research interviewing.* London: Sage.

Labov, W., & Fanshel, D. (1977). *Therapeutic discourse: Psychotherapy as conversation.* London: Academic Press.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91–195). Cambridge, UK: Cambridge University Press.

Latour, B., & Woolgar, S. (1986). *Laboratory life: The construction of scientific facts* (2nd ed.). Princeton, NJ: Princeton University Press.

Lee, J. R. E. (1995). The trouble is nobody listens. In J. Siegfried (Ed.), *Therapeutic and everyday discourse as behaviour change: Towards a micro-analysis in psychotherapy process research* (pp. 365–390). Norwood, NJ: Ablex.

Levinson, S. (1983). *Pragmatics.* Cambridge, UK: Cambridge University Press.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry.* London: Sage.

Lofland, J., & Lofland, L. H. (1984). *Analyzing social settings: A guide to qualitative observation and analysis.* Belmont, CA: Wadsworth.

Lynch, M. (1994). Representation is overrated: Some critical remarks about the use of the concept of representation in science studies. *Configurations: A Journal of Literature, Science and Technology, 2,* 137–149.

Madigan, S. P. (1992). The application of Michel Foucault's philosophy in the problem externalizing discourse of Michael White. *Journal of Family Therapy, 14,* 265–279.

Madill, A., & Barkham, M. (1997). Discourse analysis of a theme in one successful case of brief psychodynamic-interpersonal psychotherapy. *Journal of Counselling Psychology, 44,* 232–244.

Malson, H., & Ussher, J. M. (1996). Bloody women: A discourse analysis of amenorrhea as a symptom of anorexia nervosa. *Feminism and Psychology, 6,* 505–521.

Manning, P. K., & Cullum-Swan, B. (1994) Narrative, content and semiotic analysis. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 463–477). London: Sage.

Mason, J. (1994). Linking qualitative and quantitative data analysis. In A. Bryman & R. G. Burgess (Eds.), *Analyzing qualitative data.* London: Routledge.

McNamee, S., & Gergen, K. (Eds) (1992). *Therapy as social construction.* London: Sage

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd Ed.). London: Sage.

Miller, G., & Dingwall, R. (Ed.) (1997). *Context and method in qualitative research.* London: Sage

Miller, G., & Silverman, D. (1995). Troubles talk and counseling discourse: A comparative study. *The Sociological Quarterly, 36,* 725–747.

Mischler, E. G. (1986). *Research interviewing: Context and narrative.* Cambridge, MA: Harvard University Press.

Morgan, D. L. (1997). *Focus groups as qualitative research* (2nd ed.). London: Sage.

Morris, G. H., & Chenail, R. J. (Eds.) (1995). *The talk of the clinic: Explorations in the analysis of medical and therapeutic discourse.* Hillsdale, NJ: Erlbaum.

Moscovici, S. (1984). The phenomenon of social representations. In R. M. Farr & S. Moscovici (Eds.), *Social representations* (pp. 3–69). Cambridge, UK: Cambridge University Press.

Nelson, C. K. (1994). Ethnomethodological positions on the use of ethnographic data in conversation analytic research. *Journal of Contemporary Ethnography, 23,* 307–329.

Newfield, N. A., Kuehl, B. P., Joanning, H. P., & Quinn, W. H. (1990). A mini ethnography of the family therapy of adolescent drug abuse: The ambiguous experience. *Alcoholism Treatment Quarterly, 7,* 57–80.

Newfield, N., Sells, S. P., Smith, T. E., Newfield, S., & Newfield, F. (1996). Ethnographic research methods: Creating a clinical science of the humanities. In D. H. Sprenkle & S. M. Moon (Eds.), *Research Methods in Family Therapy.* New York: Guilford.

Newfield, N. A., Kuehl, B. P., Joanning, H. P., & Quinn, W. H. (1991). We can tell you about "Psychos" and "Shrinks": An ethnography of the family therapy of adolescent drug abuse. In T. C. Todd & M. D. Slekman (Eds.), *Family therapy approaches with adolescent substance Abusers* (pp. 277–310). London: Allyn & Bacon.

Nofsinger, R. (1991). *Everyday conversation.* London: Sage.

Ochs, E. (1979). Transcription as theory. In E. Ochs & B. Schieffelin (Eds.), *Developmental pragmatics* (pp. 43–47). New York: Academic Press.

Ochs, E., Schegloff, E., & Thompson, S. A. (Eds.) (1996). *Interaction and grammar.* Cambridge, UK: Cambridge University Press.

Olesen, V. (1994), Feminisms and models of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 158–174). London: Sage.

Peräkylä, A. (1995). *AIDS counseling: Institutional interaction and clinical practice.* Cambridge, UK: Cambridge University Press.

Pettinger, R. E., Hockett, C. F., & Danehy, J. J. (1960). *The first five minutes: A sample of microscopic interview analysis.* Ithica, NY: Paul Martineau.

Pidgeon, N. (1996). Grounded theory: Theoretical background. In J. T. E. Richardson (Ed.), *Handbook of qualitative research methods for psychology and the social sciences* (pp. 75–85). Leicester, UK: British Psychological Society.

Pidgeon, N., & Henwood, K. (1996). Grounded theory: Practical implementation. In J. T. E. Richardson (Ed.), *Handbook of qualitative research methods for psychology and the social sciences* (pp. 86–101). Leicester, UK: British Psychological Society.

Piercy F. P., & Nickerson, V. (1996). Focus groups in family therapy research. In D. H. Sprenkle & S. M.

Moon (Eds.), *Research methods in family therapy* (pp. 173–185). New York: Guilford.

Plummer, K. (1995). Life story research. In J. A. Smith, R. Harré, & L. van Langenhove (Eds.), *Rethinking methods in psychology* (pp. 50–63). London: Sage.

Polanyi, M. (1958). *Personal knowledge*. London: Routledge and Kegan Paul.

Pollner, M., & Wikler, L. M. (1985). The social construction of unreality: A case study of a family's attribution of competence to a severely retarded child. *Family Process, 24,* 241–254.

Pomerantz, A. M. (1980). Telling my side: "limited access" as a fishing device. *Sociological Inquiry, 50,* 186–198.

Popper, K. (1959). *The logic of scientific discovery*. London: Hutchinson.

Potter, J. (1996a). *Representing reality: Discourse, rhetoric and social construction*. London: Sage.

Potter, J. (1996b). Discourse analysis and constructionist approaches: Theoretical background. In J. T. E. Richardson (Ed.), *Handbook of qualitative research methods for psychology and the social sciences*. Leicester, UK: British Psychological Society.

Potter, J. (1997). Discourse analysis as a way of analysing naturally occurring talk. In D. Silverman (Ed.), *Qualitative research: Theory, method and practice* (pp. 144–160). London: Sage.

Potter, J., & Mulkay, M. (1985). Scientists' interview talk: Interviews as a technique for revealing participants' interpretative practices. In M. Brenner, J. Brown, & D. Canter (Eds.), *The research interview: Uses and approaches* (pp. 247–271). London: Academic Press.

Potter, J., & Wetherell, M. (1987). *Discourse and social psychology: Beyond attitudes and behaviour*. London: Sage.

Potter, J., & Wetherell, M. (1994) Analyzing discourse. In A. Bryman & B. Burgess (Eds.), *Analyzing qualitative data*. London: Routledge.

Potter, J., & Wetherell, M. (1995). Discourse analysis. In J. Smith, R. Harré, & L. van Langenhove (Eds.), *Rethinking methods in psychology* (pp. 80–92). London: Sage.

Potter, J., Wetherell, M., & Chitty, A. (1991). Quantification rhetoric—cancer on television. *Discourse and Society, 2,* 333–365.

Psathas, G., & Anderson, T. (1990). The "practices" of transcription in conversation analysis. *Semiotica, 78,* 75–99.

Rachel, J. (1996). Ethnography: Practical implementation. In J. T. E. Richardson (Ed.), *Handbook of qualitative research methods for psychology and the social sciences* (pp. 113–124). Leicester, UK: British Psychological Society.

Rafuls, S. E., & Moon, S. M. (1996). Grounded theory methodology in family therapy research. In D. H. Sprenkle & S. M. Moon (Eds.), *Research methods in family therapy* (pp. 64–80). New York: Guilford.

Ramazanoglu, C. (1992). On feminist methodology: Male reason versus female empowerment. *Sociology, 26,* 207–212.

Rapley, M., & Antaki, C. (1996). A conversation analysis of the "acquiescence" of people with learning disabilities. *Journal of Community and Applied Social Psychology, 6,* 207–227.

Reason, P., & Heron, J. (1995). Co-operative inquiry. In J. A. Smith, R. Harré, & L. van Langenhove (Eds.), *Rethinking methods in psychology* (pp. 122–142). London: Sage.

Reason, P., & Rowan, J. (Eds.) (1981). *Human inquiry: A sourcebook of new paradigm research*. Chichester, UK: Wiley.

Reinharz, S. (1992). *Feminist methods in social research*. New York: Oxford University Press.

Rennie, D., Phillips, J., & Quartaro, G. (1988). Grounded theory: A promising approach to conceptualisation in psychology? *Canadian Psychology, 29,* 139–150.

Richardson, J. E. (Ed.) (1996). *Handbook of qualitative research methods for psychology and the social sciences*. Leicester, UK: British Psychological Society.

Rogers, C. R. (1942). The use of electrically recorded interviews in improving psychotherapeutic techniques. *American Journal of Orthopsychiatry, 12,* 429–434.

Ronai, C. R. (1996). My mother is mentally retarded. In C. Ellis & A. P. Bochner (Eds.), *Composing ethnography: Alternative forms of qualitative writing*. Walnut Creek, CA: AltaMira Press.

Rose, D. (1990). *Living the ethnographic life*. London: Sage.

Rosenhan, D. L. (1973). On being sane in insane places. *Science, 179,* 250–258.

Sacks, H. (1992). *Lectures on conversation*. (Vols. I & II). Oxford, UK: Blackwell.

Schegloff, E. A. (1991). Reflections on talk and social structure. In D. Boden & D. H. Zimmerman (Eds.), *Talk and social structure* (pp. 44–70). Berkeley, CA: University of California Press.

Schegloff, E. A. (1993). Reflections on quantification in the study of conversation. *Research on Language and Social Interaction, 26*, 99–128.

Schegloff, E. A. (1997) Whose text? Whose Context? *Discourse and Society, 8*, 165–187.

Scott, J. (1990). *A matter of record: Documentary sources in social research*. Cambridge, UK: Polity.

Siegfried, J. (Ed.) (1995). *Therapeutic and everyday discourse as behaviour change: Towards a micro-analysis in psychotherapy process research*. Norwood, NJ: Ablex.

Silverman, D. (1993). *Interpreting qualitative data: Methods for analysing talk, text and interaction*. London: Sage.

Silverman, D. (Ed.) (1997a). *Qualitative research: Theory, method and practice*. London: Sage.

Silverman, D. (1997b). *Discourses of counselling: HIV counselling as social interaction*. London: Sage.

Smith, L. M. (1994). Biographical method. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 286–305) London: Sage.

Smith, J. A. (1995). Repertory grids: An interactive, case-study perspective. In J. A. Smith, R. Harré, & L. van Langehove (Eds.), *Rethinking methods in psychology* (pp. 162–177). London: Sage.

Smith, J. A., Harré, R., & van Langenhove, L. (Eds.) (1995). *Rethinking methods in psychology*. London: Sage.

Soal, J., & Kottler, A. (1996). Damaged, deficient or determined? Deconstructing narratives in family therapy. *South African Journal of Psychology, 26*, 123–134.

Soyland, A. J. (1994). Functions of a psychiatric case-summary. *Text, 14*, 113–140.

Soyland, A. J. (1995). Analyzing therapeutic and professional discourse. In J. Siegfried (Ed.), *Therapeutic and everyday discourse as behaviour change: Towards a micro-analysis in psychotherapy process research* (pp. 277–300). Norwood, NJ: Ablex.

Stainton Rogers, R. (1995). Q methodology. In J. A. Smith, R. Harré, & L. van Langenhove (Eds.), *Rethinking methods in psychology* (p. 178–192). London: Sage.

Strauss, A. L., & Corbin, J. (1994). Grounded theory methodology: An overview. In N. K. Denzin, & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 273–285). London: Sage.

Toren, C. (1996). Ethnography: Theoretical background. In J. T. E. Richardson (Ed.), *Handbook of qualitative research methods for psychology and the social sciences* (pp. 102–112). Leicester, UK: British Psychological Society.

Turner, B. A. (1994). Patterns of crisis behaviour: A qualitative inquiry. In A. Bryman & R. G. Burgess (Eds.), *Analyzing qualitative data* (pp. 195–215). London: Routledge.

Turner, B. A., & Pidgeon, N. (1997). *Manmade disasters* (2nd ed.). Oxford, UK: Butterworth-Heinemann.

Watson, R. (1995). Some potentialities and pitfalls in the analysis of process and personal change in counseling and therapeutic interaction. In J. Siegfried (Ed.), *Therapeutic and everyday discourse as behaviour change: Towards a micro-analysis in psychotherapy process research* (pp. 301–339). Norwood, NJ: Ablex.

Weinstein, R. M. (1979). Patient attitudes toward mental hospitalization: A review of quantitative research. *Journal of Health and Social Behavior, 20,* 237–258.

Weinstein, R. M. (1980). The favourableness of patients' attitudes toward mental hospitalization. *Journal of Health and Social Behavior, 21,* 397–401.

Werner, O., & Schoepfle, G. M. (1987). *Systematic fieldwork: Foundations of ethnography and interviewing* (Vol. 1). London: Sage.

Wetherell, M. (1996). Fear of fat: Interpretative repertoires and ideological dilemmas. In J. Maybin & N. Mercer (Eds.), *Using English: From conversation to canon* (pp. 36–41). London: Routledge.

White, M., & Epston, D. (1990). *Narrative means to therapeutic ends*. New York: Norton.

Whyte, W. F. (1991). *Participatory action research*. London: Sage.

Widdicombe, S., & Wooffitt, R. (1995). *The language of youth subcultures: Social identity in action*. Hemel Hempstead, UK: Harvester/Wheatsheaf.

Wieder, D. L. (Ed.) (1993). Colloquy: On issues of quantification in conversation analysis. *Research on Language and Social Interaction*, 26, 151–226.

Wilkinson, S. (1986). Introduction. In S. Wilkinson (Ed.), *Feminist social psychology* (pp. 1–6). Milton Keynes, UK: Open University Press.

Wooffitt, R. (1993). Analysing accounts. In N. Gilbert (Ed.), *Researching social life* (pp. 287–305). London: Sage.

Woolgar, S. (1988). *Science: The very idea*. London: Tavistock.

Wootton, A. (1989). Speech to and from a severely retarded young Down's syndrome child. In M. Beveridge, G. Conti-Ramsden, & I. Leudar (Eds.), *The language and communication of mentally handicapped people* (pp. 157–184). London: Chapman-Hall.

Yardley, K. (1995). Role play. In J. A. Smith, R. Harré & L. van Langenhove (Eds.), *Rethinking methods in psychology* (pp. 106–121). London: Sage.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

3.07

Personality Assessment

THOMAS A. WIDIGER and KIMBERLY I. SAYLOR
University of Kentucky, Lexington, KY, USA

3.07.1 INTRODUCTION 145
3.07.2 SELF-REPORT INVENTORIES 146
3.07.2.1 Item Analyses 146
3.07.2.2 Gender, Ethnic, and Cultural Differences 149
3.07.2.3 Individual Differences 150
3.07.2.4 Response Distortion 151
3.07.2.5 Automated Assessment and Base Rates 152
3.07.2.6 Illustrative Instruments 153
3.07.2.6.1 Minnesota Multiphasic Personality Inventory-2 153
3.07.2.6.2 Neo Personality Inventory-Revised 154
3.07.3 SEMISTRUCTURED INTERVIEWS 154
3.07.3.1 Personality, Mood States, and Mental Disorders 156
3.07.3.2 Dissimulation and Distortion 156
3.07.3.3 Intersite Reliability 157
3.07.3.4 Self and Informant Ratings 158
3.07.3.5 Convergent Validity Across Different Interviews 158
3.07.3.6 Illustrative Instruments 159
3.07.3.6.1 SIDP-IV 159
3.07.3.6.2 PDI-IV 160
3.07.4 PROJECTIVE TECHNIQUES 160
3.07.4.1 Rorschach 162
3.07.4.2 Thematic Apperception Test 163
3.07.5 CONCLUSIONS 163
3.07.6 REFERENCES 163

3.07.1 INTRODUCTION

Most domains of psychology have documented the importance of personality traits to adaptive and maladaptive functioning, including the fields of behavioral medicine (Adler & Matthews, 1994), psychopathology (Watson, Clark, & Harkness, 1994; Widiger & Costa, 1994), and industrial-organizational psychology (Hogan, Curphy, & Hogan, 1994). The assessment of personality is perhaps fundamental to virtually all fields of applied psychology. However, there are as many as 4500 different personality traits, or at least 4500 different trait terms within the English language (Goldberg, 1990) and there might be almost as many personality assessment instruments. There are instruments for the assessment of individual traits (e.g., tender-mindedness), for collections of traits (e.g., the domain of agreeableness, which includes tender-mindedness, trust, straightforwardness, altruism, compliance, and modesty), for constellations of traits (e.g., the personality syndrome of psychopathy, which includes such traits as arrogance, superficial charm, impulsivity, callousness, arrogance, deception, irresponsibility, and low empathy), and for traits identified by theorists for which there might not yet be a specific term within the language (e.g., extratensive experience balance; Exner, 1993).

There are also different methods for the assessment of personality traits. The primary methods used in personality research are self-report inventories, semistructured interviews, and projective techniques. Self-report inventories consist of written statements or questions, to which a person responds in terms of a specified set of options (e.g., true vs. false or agree vs. disagree along a five-point scale). Semistructured interviews consist of a specified set of verbally administered questions, accompanied by instructions (or at least guidelines) for follow-up questions and for the interpretation and scoring of responses. Projective techniques consist of relatively ambiguous stimuli or prompts, the possible responses to which are largely open-ended, accompanied by instructions (or at least guidelines) for their scoring or interpretation. The distinctions among these three methods, however, is somewhat fluid. A self-report inventory is equivalent to a fully structured, self-administered interview; semistructured interviews are essentially verbally administered self-report inventories that include at least some open-ended questions; and fully structured interviews are essentially verbally administered self-report inventories.

Each of the methods will be considered below, including a discussion of issues relevant to or problematic for that respective method. However, the issues discussed for one method of assessment will apply to another. Illustrative instruments for each method are also presented.

3.07.2 SELF-REPORT INVENTORIES

The single most popular method for the assessment of personality by researchers of normal personality functioning is a self-report inventory (SRI). The advantages of SRIs, relative to semistructured interviews and projective techniques, are perhaps self-evident. They are substantially less expensive and time-consuming to administer. Data on hundreds of persons can be obtained, scored, and analyzed at relatively little cost to the researcher. Their inexpensive nature allows collection of vast amounts of normative data that are unavailable for most semistructured interviews and projective techniques. These normative data in turn facilitate substantially their validation, as well as their interpretation by and utility to researchers and clinicians.

The high degree of structure of SRIs has also contributed to much better intersite reliability than is typically obtained by semistructured interviews or projective techniques. The findings of SRIs are very sensitive to the anxious, depressed, or angry mood states of respondents, contributing at times to poor test–retest reliability (discussed further below). However, the correlation between two SRIs is much more likely to be consistent across time and across research sites than the correlation between two semistructured interviews or two projective techniques. SRIs might be more susceptible to mood state distortions than semistructured interviews, but this susceptibility may itself be observed more reliably across different studies than the lack of susceptibility of semistructured interviews.

The specific and explicit nature of SRIs has also been very useful in researching and understanding the source and nature of subjects' responses. Much more is known about the effects of different item formats, length of scales, demographic variables, base rates, response sets, and other moderating variables from the results of SRIs than from semistructured interviews or projective techniques. Five issues to be discussed below are (i) item analyses; (ii) gender, ethnic, and cultural differences; (iii) individual differences; (iv) response distortion; and (v) automated assessments.

3.07.2.1 Item Analyses

There are a variety of methods for constructing, selecting, and evaluating the items to be included within an SRI (Clark & Watson, 1995). However, it is useful to highlight a few points within this chapter as such analyses are of direct importance to personality assessment research.

An obvious method for item construction, often termed the rational approach, is to construct items that describe explicitly the trait being assessed. For example, the most frequently used SRI in clinical research for the assessment of personality disorders is the Personality Diagnostic Questionnaire-4 (PDQ-4; Hyler, 1994). Its items were written to inquire explicitly with respect to each of the features of the 10 personality disorders included in the American Psychiatric Association's (APA) *Diagnostic and statistical manual of mental disorders* (*DSM-IV*; APA, 1994). For example, the dependent personality criterion, "has difficulty expressing disagreement with others because of fear of loss of support or approval" (APA, 1994, p. 668) is assessed by the PDQ-4 item, "I fear losing the support of others if I disagree with them" (Hyler, 1994, p. 4).

The content validity of an SRI is problematic to the extent that any facets of the trait being assessed are not included or are inadequately represented, items representing other traits are included, or the set provides a disproportionately greater representation of one facet of the trait relative to another (Haynes, Richard, & Kubany, 1995). "The exact phrasing of items can exert a profound influence on the construct that is actually measured" (Clark & Watson, 1995, p. 7) yet the content of items often receives little consideration in personality assessment research. Widiger, Williams, Spitzer, and Frances (1985) concluded from a systematic content analysis of the Millon Clinical Multiaxial Inventory (MCMI; Millon, 1983) that the MCMI failed to represent adequately many significant features of the antisocial personality disorder.

> Many of the [personality] traits are not sampled at all (e.g., inability to sustain consistent work behavior, lack of ability to function as a responsible parent, inability to maintain an enduring attachment to a sexual partner, failure to plan ahead or impulsivity, and disregard for the truth), including an essential requirement of the *DSM-III* antisocial criteria to exhibit significant delinquent behavior prior to the age of 15. (Widiger et al., 1985, p. 375)

Many MCMI-III items are written in part to represent an alternative theoretical formulation for the *DSM-IV* personality syndromes (Millon et al., 1996).

High content validity, however, will not ensure a valid assessment of a respective personality trait. Although many of the PDQ-4 dependency items do appear to represent adequately their respective diagnostic criteria, it is difficult to anticipate how an item will actually perform when administered to persons with varying degrees of dependency and to persons with varying degrees of other personality traits, syndromes, or demographic characteristics. Most authors of SRIs also consider the findings of internal consistency analyses (Smith & McCarthy, 1995). "Currently, the single most widely used method for item selection in scale development is some form of internal consistency analysis" (Clark & Watson, 1995, p. 313).

Presumably, an item should correlate more highly with other items from the same scale (e.g., dependent personality traits) than items from another scale (e.g., borderline personality traits), consistent with the basic principles of convergent and discriminant validity. However, these assumptions hold only for personality syndromes that are homogeneous in content and that are distinct from other syndromes, neither of which may be true for syndromes that involve overlapping constellations of personality traits (Pilkonis, 1997; Shea, 1995). Confining a scale to items that correlate highly with one another can result in an overly narrow assessment of a construct, and deleting items that correlate highly with other scales can result in false distinctions and a distorted representation of the trait being assessed (Clark & Watson, 1995; Smith & McCarthy, 1995). For example, the APA (1994) criteria set for the assessment of antisocial personality disorder does not include such items as lacks empathy and arrogant self-appraisal that are included in alternative criteria sets for this personality syndrome (Hare, 1991) in part because these items are already contained within the *DSM-IV* criteria set for the narcissistic personality disorder. Their inclusion within the criteria set for antisocial personality disorder would complicate the differentiation of the antisocial and narcissistic personality syndromes (Gunderson, 1992), but the failure to include these items may also provide an inadequate description and assessment of antisocial (psychopathic) personality traits (Hare, Hart, & Harpur, 1991).

Overlapping scales, on the other hand, have their own problems, particularly if they are to be used to test hypotheses regarding the relationships among or differences between the traits and syndromes being assessed (Helmes & Reddon, 1993). For example, the MCMI-III personality scales overlap substantially (Millon, Millon, & Davis, 1994). Scale overlap was in part a pragmatic necessity of assessing 14 personality disorders and 10 clinical syndromes with no more than 175 items. However, this overlap was also intentional to ensure that the scales be consistent with the overlap of the personality constructs being assessed. "Multiple keying or item overlapping for the MCMI inventories was designed to fit its theory's polythetic structure, and the scales constructed to represent it" (Millon, 1987, p. 130). However, the MCMI-III cannot then be used to assess the validity of this polythetic structure, or to test hypotheses concerning the covariation, differentiation, or relationship among the personality constructs, as the findings will be compelled by the scale overlap (Helmes & Reddon, 1993). For example, the MCMI-III provides little possibility for researchers to fail to confirm a close relationship between, or comparable findings for, sadistic and antisocial personality traits, given that eight of the 17 antisocial items (47%) are shared with the sadistic scale. The same point can be made for studies using the Minnesota Multiphasic Personality Inventory (MMPI-2) antisociality and cynicism scales, as they share approximately a third of their items (Helmes & Reddon, 1993). It was for this reason that Morey, Waugh, and Blashfield (1985) developed both overlapping and nonoverlapping MMPI personality disorder scales.

Scale homogeneity and interscale distinctiveness are usually emphasized in SRIs constructed through factor analyses (Clark & Watson, 1995). For example, in the construction of the NEO Personality Inventory-Revised (NEO-PI-R), Costa and McCrae (1992)

> adopted factor analysis as the basis for item selection because it identifies clusters of items that covary with each other but which are relatively independent of other item clusters—in other words, items that show convergent validity with respect to other items in the cluster and divergent validity with respect to other items [outside the cluster]. (p. 40)

Reliance upon factor analysis for scale construction has received some criticism (Block, 1995; Millon et al., 1996) and there are, indeed, instances in which factor analyses have been conducted with little appreciation of the limitations of the approach (Floyd & Widaman, 1995), but this is perhaps no different than for any other statistical technique. Factor analysis remains a powerful tool for identifying underlying, latent constructs, for data reduction, and for the validation of an hypothesized dimensional structure (Smith & McCarthy, 1995; Watson et al., 1994). In any case, the construction of the NEO-PI-R through factor analysis was consistent with the theoretical model for the personality traits being assessed (Clark & Watson, 1995; Costa & McCrae, 1992; Goldberg, 1990). Additional illustrations of theoretically driven factor analytic scale constructions are provided by the Dimensional Assessment of Personality Pathology—Basic Questionnaire (DAPP-BQ; Livesley & Jackson, in press), the Schedule for Nonadaptive and Adaptive Personality (SNAP; Clark, 1993), the Interpersonal Adjective Scales (IAS; Wiggins & Trobst, 1997), the Multidimensional Personality Questionnaire (MPQ; Tellegen & Waller, in press), and the Temperament and Character Inventory (TCI; Cloninger & Svrakic, 1994).

The most informative item analyses will be correlations with external validators of the personality traits, including the ability of items to identify persons with the trait in question, correlations with hypothesized indicators of the trait, and an absence of correlations with variables that are theoretically unrelated to the trait (Smith & McCarthy, 1995). These data typically are discussed under a general heading of construct validity, including data concerning items' concurrent, postdictive, predictive, convergent, and discriminant validity (Ozer & Reise, 1994). Illustrative studies with SRIs include Ben-Porath, McCully, and Almagor (1993), Clark, Livesley, Schroeder, and Irish (1996), Lilienfeld and Andrews (1996), Robins and John (1997), and Trull, Useda, Coasta, and McCrae (1995).

The MMPI-2 clinical scales exemplify the purely empirical approach to scale construction. "The original MMPI (Hathaway & McKinley, 1940) was launched with the view that the content of the items was relatively unimportant and what actually mattered was whether the item was endorsed by a particular clinical group" (Butcher, 1995, p. 302). This allowed the test to be "free from the restriction that the subject must be able to describe his own behavior accurately" (Meehl, 1945, p. 297). For example, one of the items on the MMPI hysteria scale was "I enjoy detective or mystery stories" (keyed false; Hathaway & McKinley, 1982).

Indicating that one does not enjoy detective or mystery stories does not appear to have any obvious (or perhaps even meaningful) relationship to the presence of the mental disorder of hysteria, but answering false did correlate significantly with the occurrence of this disorder. "The literal content of the stimulus (the item) is entirely unimportant, and even irrelevant and potentially misleading. Sophisticated psychometric use of test items dictates that the test interpreter ignore item content altogether lest he or she be misled" (Ben-Porath, 1994, p. 364). Limitations to this approach, however, are suggested below.

A sophisticated approach to item analysis is item response theory (IRT; Hambleton, Swaminathan, & Rogers, 1991). Its most unique datum is the probability of a response to a particular item given a particular level of the personality trait being assessed. IRT analyses demonstrate graphically how items can vary in their discriminative ability at different levels of the trait being assessed (assuming that the anchoring items are not themselves systematically biased). It is not the case that all items perform equally well, nor does any particular item perform equally well at all levels of the trait. It may be the case that none of the items on a scale provide any discrimination at particular levels of a trait, or that the scale is predominated by items that discriminate at high rather than moderate or low levels of the trait. For example, it is possible that the MMPI-2 items to assess the personality domain of neuroticism are weighted toward a discrimination of high levels of neuroticism, providing very little discrimination at low levels. IRT analyses might allow researchers to maximize the sensitivity of a test to a particular population (e.g., inpatient hospital vs. prison setting) by deleting items that work poorly within certain settings (Embretson, 1996). IRT analyses have been used widely with tests of abilities and skills, and are now being applied to measures of personality (Ozer & Reise, 1994). Illustrative applications include analyses of items from the MPQ (Tellegen & Waller, in press) by Reise and Waller (1993), items from a measure of alexithymia by Hendryx, Haviland, Gibbons, and Clark (1992) and items from a psychopathy semistructured interview by Cooke and Michie (1997).

3.07.2.2 Gender, Ethnic, and Cultural Differences

Differences among gender, ethnic, cultural, and other demographic populations with respect to personality traits are often socially sensitive, politically controversial, and difficult to explain (Eagly, 1995). There is substantial SRI research to indicate differences, on average, between males and females for a wide variety of personality traits, as assessed by research with SRIs (Sackett & Wilk, 1994). For example, males obtain higher scores on measures of assertiveness and self-esteem; whereas females obtain higher scores on measures of anxiousness, trust, gregariousness, and tender-mindedness (Feingold, 1994).

Separate norms therefore have been developed for the interpretation of most SRI personality scales (e.g., Costa & McCrae, 1992; Graham, 1993; Millon et al., 1994). However, the rationale for providing different norms for each sex (gender) is unclear. Sackett and Wilk (1994) "searched test manuals, handbooks, and the like for a discussion of the rationale for the practice. We typically found none" (p. 944). They indicated, with some surprise, that the provision of separate norms for males and females "does not seem to have been viewed as at all controversial" (Sackett & Wilk, 1994, p. 944). If the SRI is indicating an actual difference between males and females (Feingold, 1994) it is unclear why this difference showed then be eliminated or diminished in SRI assessments of males and females (Kehoe & Tenopyr, 1994; Sackett & Wilk, 1994).

For example, males and females with the same raw score on the psychopathic deviate scale of the MMPI-2 will be given different final scores due to the use of different norms. Different raw scores for males and females can then result in the same final score (Graham, 1993). The extent of psychopathy in females is then relative to other females; it is not a measure of the actual (absolute) extent to which they are psychopathic. A male who has more traits (or symptoms) of psychopathy could be described by the MMPI-2 as being less psychopathic than a female with the same number of traits. The separate norms provided for males and females on SRIs are never so substantial as to eliminate entirely the differences between the sexes, but the rationale for reducing or minimizing any differences is unclear.

The provision of separate norms for males and females by SRIs is inconsistent with other domains of assessment, including most semistructured interview and projective assessments of personality variables. For example, the same threshold is used for males and females by the most commonly used semistructured interview for the assessment of psychopathy, the (revised) Psychopathy Checklist (Hare, 1991). Separate norms are not provided for females and males in the assessment of intelligence, nor are they provided for different cultural and ethnic groups in the assessment of personality. Although statistically significant differences have also been obtained across ethnic and cultural groups for SRI personality measures, "none of the major personality measures ... offered norm scoring based on race or ethnicity either as a routine aspect of the scoring system or as a scoring option" (Sackett & Wilk, 1994, p. 947). It is unclear whether separate norms should be provided for ethnic groups, but it does appear to be inconsistent to provide separate norms for gender and not for ethnicity.

Separate norms would be appropriate if there is reason to believe that the SRI items are biased against a particular gender, ethnic, or cultural group. "Bias is the extent to which measured group differences are invalid ... Group differences are invalid to the extent that the constructs that distinguish between groups are different from the constructs the measures were intended to represent" (Kehoe & Tenopyr, 1994, p. 294). Consider, for example, the MMPI item cited earlier, "I enjoy detective or mystery stories" (Hathway & McKinley, 1982, p. 3). The MMPI reliance upon a blind empiricism for item selection can be problematic if the basis for the item endorsement is for reasons other than, or in addition to, the presence of the construct being assessed. Lindsay and Widiger (1995) suggested that this item might have been correlated with hysteria because hysteria was itself correlated with female gender. Significantly more females than males are diagnosed with hysteria and significantly more females than males respond negatively to an interest in detective or mystery stories; therefore, the item may have correlated with hysteria because it was identifying the presence of females rather than the presence of hysteria. In addition,

> because such an item concerns normal behavior that occurs more often in one sex than in the other, we would consider it to be sex biased because its errors in prediction (false positives) will occur more often in one sex than in the other. (Lindsay & Widiger, 1995, p. 2)

One would never consider including as one of the *DSM-IV* diagnostic criteria for histrionic personality disorder, "I have often enjoyed reading Cosmopolitan and Ms. Magazine" but many such items are included in SRIs to diagnose the presence of a personality disorder. For example, responding false to the item "in the past, I've gotten involved sexually with many people who didn't matter much to me," is used for the assessment of a dependent personality disorder on the MCMI-II (Millon, 1987, p. 190). Not getting involved sexually with many people who do not matter much is hardly an indication of a dependent personality disorder (or of any dysfunction), yet it is used to diagnose the presence of this personality disorder.

There are also data to suggest that the same items on an SRI may have a different meaning to persons of different ethnic, cultural, or gender groups (Okazaki & Sue, 1995), although the magnitude, consistency, and significance of these differences has been questioned (Ozer & Reise, 1994; Timbrook & Graham, 1994). Applications of IRT analyses to the assessment of bias would be useful (Kehoe & Tenopyr, 1994). For example, Santor, Ramsay, and Zuroff (1994) examined whether men and women at equal levels of depression respond differentially to individual items. They reported no significant differences for most items, but bias was suggested for a few, including a body-image dissatisfaction item. Females were more likely to endorse body-image dissatisfaction than males even when they were at the same level of depression as males, suggesting that the endorsement of the item by women (relative to men) reflected their gender independently of their level of depression.

3.07.2.3 Individual Differences

The provision of any set of norms may itself be questioned. Personality description with SRIs traditionally has been provided in reference to individual differences. "Raw scores on personality inventories are usually meaningless—responses take on meaning only when they are compared to the responses of others" (Costa & McCrae, 1992, p. 13). However, this individual differences approach can itself be problematic. "The individual differences research paradigm, which has thoroughly dominated empirical personality research throughout the present century, is fundamentally inadequate for the purposes of a science of personality" (Lamiell, 1981, p. 36). Lamiell's scathing critique of individual differences research raised many important and compelling concerns. For example, he noted how the test–retest reliability of a personality scale typically is interpreted as indicating the extent to which the expression of a trait is stable in persons across time, but most analyses in fact indicate the extent to which persons maintain their relative position on a scale over time. The test–retest reliability of relative position may itself correlate highly with the test–retest reliability of the magnitude within each person, but it need not. For example, the reliability of a measure of height, assessed across 100 persons using a product-moment correlation, would suggest substantial stability between the ages of eight and 15, despite the fact that each person's actual height would have changed substantially during this time. Height is not at all consistent or stable across childhood, yet a measure of stability in height would be very high if the relative height among the persons changed very little. This confusion, however, would be addressed by alternative reliability statistics (e.g., an intraclass correlation coefficient).

Interpreting personality test scores relative to a particular population is not necessarily problematic as long as the existence of these norms and their implications for test interpretation are understood adequately. For example, a score on an MMPI-2 social introversion scale does not indicate how introverted a person is, but how much more (or less) introverted the person is relative to a particular normative group. The MMPI-2 social introversion scale does not indicate the extent to which an Asian-American male is socially introverted, it indicates how more (or less) introverted he is relative to a sample of 1138 American males who provided the normative data, only 1% of whom were Asian-American (Graham, 1993).

Nevertheless, researchers and clinicians will at times refer to SRI results as if they are providing an absolute measure of a personality trait. For example, researchers and clinicians will describe a person as having high self-esteem, low self-monitoring, or high dependency, when in fact the person is simply higher or lower than a particular comparison group. It is common in personality research to construct groups of subjects on the basis of median splits on SRI scores for such traits as self-esteem, self-monitoring, dependency, autonomy, or some other personality construct. Baumeister, Tice, and Hutton (1989) searched the literature for all studies concerning self-esteem. "Most often ...

high and low self-esteem groups are created by performing a median split on the self-esteem scores across the sample" (p. 556). Persons above the median are identified as having "high self-esteem" whereas persons below the median are identified as having "low self-esteem." However, interpreting SRI data in this manner is inaccurate and potentially misleading. Persons below a median would indeed have less self-esteem than the persons above a median, but it is unknown whether they are in fact either high or low in self-esteem. All of the subjects might be rather high (or low). This method of assessing subjects is comparable to providing a measure of psychopathy to a sample of nuns to identify a group of psychopaths and nonpsychopaths, or a measure of altruism to convicts within a prison to identify a group of saints and sinners. Yet, many researchers will provide measures of self-esteem (self-monitoring, dependency, narcissism, or some other SRI measure) to a group of well-functioning college students to identify persons high and low in self-esteem. Baumeister et al. (1989) indeed found that "in all cases that we could determine, the sample midpoint was higher (in self-esteem) than the conceptual midpoint, and generally the discrepancy was substantial and significant" (p. 559).

3.07.2.4 Response Distortion

SRIs rely substantially on the ability of the respondent to provide a valid self-description. To the extent that a person does not understand the item, or is unwilling or impaired in his or her ability to provide an accurate response, the results will be inaccurate. "Detection of an attempt to provide misleading information is a vital and necessary component of the clinical interpretation of test results" (Ben-Porath & Waller, 1992, p. 24).

The presence of sufficiently accurate self-description is probably a reasonable assumption in most instances (Costa & McCrae, 1997). However, it may also be the case that no person will be entirely accurate in the description of this or her personality. Each person probably evidences some degree of distortion, either minimizing or exaggerating flaws or desirabilities. Such response distortion will be particularly evident in persons characterized by personality disorders (Westen, 1991). Antisocial persons will tend to be characteristically dishonest or deceptive in their self-descriptions, dependent persons may self-denigrate, paranoid persons will often be wary and suspicious, borderline persons will tend to idealize and devalue, narcissistic persons will often be arrogant or self-promotional, and histrionic persons will often be overemotional, exaggerated, or melodramatic in their self-descriptions (APA, 1994). Response distortion may also be common within some populations, such as forensic, disability, or psychiatric settings, due in part to the higher prevalence of personality disorder symptomatology within these populations but due as well to the pressures, rewards, and inducements within these settings to provide inaccurate self-descriptions (Berry, 1995).

A substantial amount of research has been conducted on the detection of response distortion (otherwise known as response sets or biases), particularly with the MMPI-2 (Butcher & Rouse, 1996). There are self-report scales to detect nonresponsiveness (e.g., random responding, yea-saying or nay-saying), overreporting of symptoms (e.g., malingering, faking bad, exaggeration, or self-denigration), and underreporting (e.g., faking good, denial, defensiveness, minimization, or self-aggrandizement). These scales often are referred to as validity scales, as their primary function has been to indicate the extent to which the scores on the personality (or clinical) scales are providing an accurate or valid self-description, and they do appear to be generally successful in identifying the respective response distortions (Berry, Wetter, & Baer, 1995).

However, it is not always clear whether the prevalence of a respective form of response distortion is frequent enough to warrant its detection within all settings (Costa & McCrae, 1997). Acquiescence, random responding, and nay-saying might be so infrequent within most settings that the costs of false positive identifications (i.e., identifying a personality description as invalid when it is in fact valid) will outweigh the costs of false negatives (identifying a description as valid when it was in fact invalid). In addition, much of the research on validity scales has been confined to analogue studies in which various response distortions are simulated by college students, psychiatric patients, or other persons. There is much less data to indicate that the inclusion of a validity scale as a moderator or suppressor variable actually improves the validity of the personality assessment. For example, it is unclear whether the correlation of a measure of the personality trait of neuroticism with another variable (e.g., drug usage) would increase when variance due to a response distortion (e.g., malingering or exaggeration) is partialled from the measure of neuroticism (Costa & McCrae, 1997).

In some contexts, such as a disability evaluation, response distortions are to be avoided, whereas in other contexts, such as the assessment of maladaptive personality traits, they may be what the clinician is seeking to identify (Ozer & Reise, 1994). Validity scales typically are interpreted as indicators of the presence of misleading or inaccurate information, but the response tendencies assessed by validity scales are central to some personality disorders. For example, an elevation on a malingering scale would indicate the presence of deception or dishonesty, suggesting perhaps that the information provided by the subject in response to the other SRI items was inaccurate or misleading. However, one might not want to partial the variance due to malingering from an SRI measure of antisocial or psychopathic personality. The validity of a measure of psychopathy would be reduced if variance due to deception or dishonesty was extracted, as dishonesty or deception is itself a facet of this personality syndrome (APA, 1994; Hare, 1991; Lilienfeld, 1994). It might be comparably misleading to extract variance due to symptom exaggeration from a measure of borderline personality traits, self-aggrandizement from a measure of narcissism, or self-denigration from a measure of dependency. Validity scales are not just providing a measure of response distortion that undermines the validity of personality description; they are also providing valid descriptions of highly relevant and fundamental personality traits.

For example, a response distortion that has been of considerable interest and concern is social desirability. Persons who are instructed to attribute falsely to themselves good, desirable qualities provide elevations on measures of social desirability, and many researchers have therefore extracted from a measure of personality the variance that is due to this apparent response bias. However, extracting this variance typically will reduce rather than increase the validity of personality measures because much of the socially desirable self-description does in fact constitute valid self-description (Borkenau & Ostendorf, 1992; McCrae & Costa, 1983). Persons who characteristically describe themselves in a socially desirable manner are not simply providing false and misleading information. Much of the variance within a measure of social desirability is due to persons either describing themselves accurately as having many desirable personality traits or, equally accurately, as having a personality disposition of self-aggrandizement, arrogance, or denial. This problem could be addressed by first extracting the variance due to valid individual differences from a measure of social desirability before it is used as a moderater variable, but then it might obviously fail to be useful as a moderater variable to the personality scales.

3.07.2.5 Automated Assessment and Base Rates

The structured nature of the responses to SRIs has also facilitated the development of automated (computerized) systems for scoring and interpretation. The researcher or clinician simply uses score sheets that can be submitted for computerized scanning, receiving in return a complete scoring and, in most cases, a narrative description of the subject's personality derived from the theoretical model of, and the data considered by, the author(s) of the computer system. Automated systems are available for most of the major personality SRIs, and their advantages are self-evident. "The computer can store and access a much larger fund of interpretive literature and base-rate data than any individual clinician can master, contributing to the accuracy, objectivity, reliability, and validity of computerized reports" (Keller, Butcher, & Slutske, 1990, p. 360).

Automated interpretive systems can also be seductive. They might provide the appearance of more objectivity and validity than is in fact the case (Butcher, 1995). Clinicians and researchers should always become closely familiar with the actual data and procedures used to develop an automated system, and the subsequent research assessing its validity, in order to evaluate objectively for themselves the nature and extent of the empirical support. For example,

> while any single clinician may do well to learn the norms and base rates of the client population he or she sees most often, a computer program can refer to a variety of population norms and, if programmed to do so, will always "remember" to tailor interpretive statements according to modifying demographic data such as education, marital status, and ethnicity. (Keller et al., 1990, p. 360)

This is an excellent sentiment, describing well the potential benefits of an automated report. However, simply providing a subject's education, marital status, and ethnicity to the automated computer system does not necessarily mean that this information will in fact be used. As noted above, very few of the SRIs consider ethnicity.

A purported advantage of the MCMI-III automated scoring system is that "actuarial base rate data, rather than normalized standard score transformations, were employed in calculating scale measures" (Millon et al., 1994, p. 4). The cutoff points used for the interpretation of the MCMI-III personality scales are based on the base rates of the personality syndromes within the population. "The BR [base rate] score was designed to anchor cut-off points to the prevalence of a particular attribute in the psychiatric population" (Millon et al., 1994, p. 26). This approach appears to be quite sophisticated, as it is seemingly responsive to the failure of clinicians and researchers to consider the effect of base rates on the validity of an SRI cutoff point (Finn & Kamphuis, 1995). "These data not only provide a basis for selecting optimal differential diagnostic cut-off scores but also ensure that the frequencies of MCMI-III diagnoses and profile patterns are comparable to representative clinical prevalence rates" (Millon et al., 1994 p. 4).

However, the MCMI-III automated scoring system does not in fact make adjustments to cutoff points depending upon the base rate of the syndrome within a local clinical setting. The MCMI-III uses the same cutoff point for all settings and, therefore, for all possible base rates. The advantage of setting a cutoff point according to the base rate of a syndrome is lost if the cutoff point remains fixed across different base rates (Finn & Kamphuis, 1995). "The use of the MCMI-III should be limited to populations that are not notably different in background from the sample employed to develop the instrument's base rate norms" (Millon et al., 1994, p. 35). It is for this reason that Millon et al. discourage the use of the MCMI-III within normal (community or college) populations (a limitation not shared by most other SRIs). In fact, the population sampled for the MCMI-III revision might have itself been notably different in background from the original sample used to develop the base rate scores. Retzlaff (1996) calculated the probability of having a personality disorder given the obtainment of a MCMI-III respective cutoff point, using the data provided in the MCMI-III test manual. Retzlaff concluded that "as determined by currently available research, the operating characteristics of the MCMI-III scales are poor" (p. 437). The probabilities varied from 0.08 to a high of only 0.32. For example, the probability of having an avoidant personality disorder if one surpassed the respective cutoff point was only 0.17, for the borderline scale it was only 0.18, and for the antisocial scale it was only 0.07. As Retzlaff indicated, "these hit rate statistics are well under one-half of the MCMI-II validities" (Retzlaff, 1996, p. 435).

Retzlaff (1996) suggested that there were two possible explanations for these results, either "the test is bad or that the validity study was bad" (p. 435) and he concluded that the fault lay with the MCMI-III cross-validation data. "At best, the test is probably valid, but there is no evidence and, as such, it cannot be trusted until better validity data are available" (Retzlaff, 1996, p. 435). However, it is unlikely that any new data will improve these statistics, unless the new study replicates more closely the population characteristics of the original derivation study. A third alternative is to have the cutoff points for the scales be adjusted for different base rates within local settings, but no automated scoring system currently provides this option.

3.07.2.6 Illustrative Instruments

There are many SRIs for the assessment of normal and maladaptive personality traits, including the SNAP (Clark, 1993), the DAPP-BQ (Livesley & Jackson, in press), the Personality Assessment Inventory (Morey, 1996), the MPQ (Tellegen & Waller, in press), the MCMI-III (Millon et al., 1994), the PDQ-4 (Hyler, 1994), the IAS (Wiggins & Trobst, 1997), the Wisconsin Personality Inventory (Klein et al., 1993), and the TCI (Cloninger & Svrakic, 1994). Space limitations prohibit a consideration of each of them. A brief discussion, however, will be provided for the two dominant SRIs within the field of clinical personality assessment, the MMPI-2 (Graham, 1993; Hathaway et al., 1989) and the NEO-PI-R (Costa & McCrae, 1992).

3.07.2.6.1 Minnesota Multiphasic Personality Inventory-2

The MMPI-2 is an SRI that consists of 566 true/false items (Graham, 1993; Hathaway et al., 1989). It is the most heavily researched and validated SRI, and the most commonly used in clinical practice (Butcher & Rouse, 1996). Its continued usage is due in part to familiarity and tradition, but its popularity also reflects the obtainment of a substantial amount of empirical support and normative data over the many years of its existence (Graham, 1993).

The MMPI-2 is described as "the primary self-report measure of abnormal personality" (Ben-Porath, 1994, p. 363) but its most common usage is for the assessment of anxiety, depressive, substance, psychotic, and other such (Axis I) mental disorders rather than for the assessment of personality traits (Ozer & Reise, 1994). "The MMPI-2 clinical scales are ... measures of various forms of psychopathology ... and not measures of general personality" (Greene, Gwin, & Staal, 1997, p. 21). This is somewhat ironic, given that it is titled as a personality inventory (Helmes & Reddon, 1993).

However, the MMPI-2 item pool is extensive and many additional scales have been developed beyond the basic 10 clinical and three validity scales (Graham, 1993). For example, some of the new MMPI-2 content scales (Ben-Porath, 1994; Butcher, 1995), modeled after the seminal research of Wiggins (1966), do assess personality traits (e.g., cynicism, social discomfort, Type A behavior, and low self-esteem). The Morey et al. (1985) personality disorder scales have also been updated and normed for the MMPI-2 (Colligan, Morey, & Offord, 1994) and an alternative set of MMPI-2, *DSM-IV* personality disorder scales are being developed by Somwaru and Ben-Porath (1995). Harkness and McNulty (1994) have developed scales (i.e., the PSY-5) to assess five broad domains of personality (i.e., neuroticism, extraversion, psychoticism, aggressiveness, and constraint) that are said to provide the MMPI-2 assessment of, or an alternative to, the five-factor model of personality (Ben-Porath, 1994; Butcher & Rouse, 1996). "Other measures of five-factor models of normal personality will need to demonstrate incremental validity in comparison to the full set of MMPI-2 scales (including the PSY-5) to justify their use in clinical practice" (Ben-Porath, 1994, p. 393). However, there are important distinctions between the PSY-5 and the five-factor model constructs, particularly for psychoticism and aggressiveness (Harkness & McNulty, 1994; Harkness, McNulty, & Ben-Porath, 1995; Widiger & Trull, 1997). In addition, the utility of the MMPI-2 for personality trait description and research may be limited by the predominance of items concerning clinical symptomatology and the inadequate representation of important domains of personality, such as conscientiousness (Costa, Zonderman, McCrae, & Williams, 1985) and constraint (DiLalla, Gottesman, Carey, & Vogler, 1993)

3.07.2.6.2 Neo Personality Inventory-Revised

The most comprehensive model of personality trait description is provided by the five-factor model (Saucier & Goldberg, 1996; Wiggins & Pincus, 1992). Even the most ardent critics of the five-factor model acknowledge its importance and impact (e.g., Block, 1995; Butcher & Rouse, 1996; Millon et al., 1996). And, the predominant measure of the five-factor model is the NEO-PI-R (Costa & McCrae, 1992, 1997; Ozer & Reise, 1994; Widiger & Trull, 1997).

The 240 item NEO-PI-R (Costa & McCrae, 1992) assesses five broad domains of personality: neuroticism, extraversion (vs. introversion), openness (vs. closedness), agreeableness (vs. antagonism), and conscientiousness. Each item is rated on a five-point scale, from strongly disagree to strongly agree. Each domain is differentiated into six underlying facets. For example, the six facets of agreeableness vs. antagonism are trust vs. mistrust, altruism vs. exploitiveness, compliance vs. oppositionalism or aggression, tender-mindedness vs. tough-mindedness (lack of empathy), modesty vs. arrogance, and straightforwardness vs. deception or manipulation. The domains and facets of the NEO-PI-R relate closely to other models of personality, such as the constructs of affiliation and power within the interpersonal circumplex model of personality (McCrae & Costa, 1989; Wiggins & Pincus, 1992). The NEO-PI-R assessment of the five-factor model also relates closely to the *DSM-IV* personality disorder nomenclature, despite its original development as a measure of normal personality functioning (Widiger & Costa, 1994).

The application of the NEO-PI-R within clinical settings, however, may be problematic, due to the absence of extensive validity scales to detect mood state and response-set distortion (Ben-Porath & Waller, 1992). A valid application of the NEO-PI-R requires that the respondent be capable of and motivated to provide a reasonably accurate self-description. This is perhaps a safe assumption for most cases (Costa & McCrae, 1997), but the NEO-PI-R might not be successful in identifying when this assumption is inappropriate. Potential validity scales for the NEO-PI-R, however, are being researched (Schinka, Kinder, & Kremer, 1997).

3.07.3 SEMISTRUCTURED INTERVIEWS

The single most popular method for clinical assessments of personality by psychologists working in either private practice, inpatient hospitals, or university psychology departments, is an unstructured interview (Watkins, Campbell, Nieberding, & Hallmark, 1995). Whereas most researchers of personality disorder rely upon semistructured interviews (SSI) (Rogers, 1995; Zimmerman, 1994), there are but a few SSIs for the assessment of normal personality functioning (e.g., Trull & Widiger, 1997) and none of the psychologists in a survey of practicing clinicians cited the use of a semistructured interview (Watkins et al., 1995).

Unstructured clinical interviews rely entirely upon the training, expertise, and conscientiousness of the interviewer to provide an accurate assessment of a person's personality. They are problematic for research as they are notoriously unreliable, idiosyncratic, and prone to false assumptions, attributional errors, and misleading expectations (Garb, 1997). For example, many clinical interviewers fail to provide a comprehensive assessment of a patient's maladaptive personality traits. Only one personality disorder diagnosis typically is provided to a patient, despite the fact that most patients will meet criteria for multiple diagnoses (Gunderson, 1992). Clinicians tend to diagnose personality traits and disorders hierarchically. Once a patient is identified as having a particular personality disorder (e.g., borderline), clinicians will fail to assess whether additional personality traits are present (Herkov & Blashfield, 1995). Alder, Drake, and Teague (1990) provided 46 clinicians with case histories of a patient that met the *DSM-III* criteria for four personality disorders (i.e., histrionic, narcissistic, borderline, and dependent). "Despite the directive to consider each category separately . . . most clinicians assigned just one [personality disorder] diagnosis" (Adler et al., 1990, p. 127). Sixty-five percent of the clinicians provided only one diagnosis, 28% provided two, and none provided all four.

Unstructured clinical assessments of personality also fail to be systematic. Morey and Ochua (1989) provided 291 clinicians with the 166 *DSM-III* personality disorder diagnostic criteria (presented in a randomized order) and asked them to indicate which personality disorder(s) were present in one of their patients and to indicate which of the 166 diagnostic criteria were present. Kappa for the agreement between their diagnoses and the diagnoses that would be given based upon the diagnostic criteria they indicated to be present, ranged from 0.11 (schizoid) to only 0.58 (borderline). In other words, their clinical diagnoses agreed poorly with their own assessments of the diagnostic criteria for each disorder. The results of this study were replicated by Blashfield and Herkov (1996). Agreement in this instance ranged from 0.28 (schizoid) to 0.63 (borderline). "It appears that the actual diagnoses of clinicians do not adhere closely to the diagnoses suggested by the [diagnostic] criteria" (Blashfield & Herkov, 1996, p. 226).

Clinicians often base their personality disorder assessments on the presence of just one or two features, failing to consider whether a sufficient number of the necessary features are present (Blashfield & Herkov, 1996). In addition, the one or two features that one clinician considers to be sufficient may not be consistent with the feature(s) emphasized by another clinician, contributing to poor interrater reliability (Widiger & Sanderson, 1995; Zimmerman, 1994) and to misleading expectations and false assumptions. For example, a number of studies have indicated that many clinicians tend to overdiagnose the histrionic personality disorder in females (Garb, 1997). Clinicians tend to perceive a female patient who has just one or two histrionic features as having a histrionic personality disorder, even when she may instead meet the diagnostic criteria for an alternative personality disorder (Blashfield & Herkov, 1996; Ford & Widiger, 1989; Morey & Ochoa, 1989). This overdiagnosis of histrionic personality disorder in females is diminished substantially when the clinician is compelled to assess systematically each one of the features. "Sex biases may best be diminished by an increased emphasis in training programs and clinical settings on the systematic use and adherence to the [diagnostic] criteria" (Ford & Widiger, 1989, p. 304).

The reluctance to use semistructured interviews within clinical practice, however, is understandable. Semistructured interviews that assess all of the *DSM-IV* personality disorder diagnostic criteria can require more than two hours for their complete administration (Widiger & Sanderson, 1995). This is unrealistic and impractical in routine clinical practice, particularly if the bulk of the time is spent in determining the absence of traits. However, the time can be diminished substantially by first administering an SRI to identify which domains of personality functioning should be emphasized and which could be safely ignored (Widiger & Sanderson, 1995).

Unstructured interviews are also preferred by clinicians because they find SSIs to be too constraining and superficial. Most clinicians prefer to follow leads that arise during an interview, adjusting the content and style to facilitate rapport and to respond to the particular needs of an individual patient. The questions provided by an SSI can appear, in comparison, to be inadequate and simplistic. However, SSIs are considered to be semistructured because they require (or allow) professional judgment and discretion in their administration and scoring. They are not simply mindless administrations of an SRI. The responsibility of an SSI interviewer is to assess for the presence of a respective trait, not to just record a subject's responses to a series of structured questions. Follow-up questions that must be sensitive and responsive to the mood state, defensiveness, and self-awareness of the person being interviewed are always required and are left to the expertise and discretion of the interviewer (Widiger, Frances, & Trull, 1989). There are only a few fully structured interviews for the assessment of personality traits and they may be inadequate precisely because of their excessive constraint and superficiality (Perry, 1992).

The questions provided by an SSI are useful in ensuring that each trait is assessed, and that a set of questions found to be useful in prior studies is being used in a consistent fashion. Systematic biases in clinical assessments are more easily identified, researched, and ultimately corrected with the explicit nature of SSIs and their replicated use across studies and research sites. A highly talented and skilled clinician can outperform an SSI, but it is risky to presume that one is indeed that talented clinician or that one is consistently skilled and insightful with every patient (Dawes, 1994). It would at least seem desirable for talented clinicians to be informed by a systematic and comprehensive assessment of a patient's personality traits.

3.07.3.1 Personality, Mood States, and Mental Disorders

Personality traits typically are understood to be stable behavior patterns present since young adulthood (Wiggins & Pincus, 1992). However, few SRIs emphasize this fundamental feature. For example, the instructions for the MMPI-2 make no reference to age of onset or duration (Hathaway et al., 1989). Responding true to the MMPI-2 borderline item, "I cry easily" (Hathaway et al., p. 6) could be for the purpose of describing the recent development of a depressive mood disorder rather than a characteristic manner of functioning. Most MMPI-2 items are in fact used to assess both recently developed mental disorders as well as long-term personality traits (Graham, 1993).

SRIs that required a duration since young adulthood for the item to be endorsed would be relatively insensitive to changes in personality during adulthood (Costa & McCrae, 1994), but SSIs are generally preferred over SRIs within clinical settings for the assessment of personality traits due to the susceptibility of SRIs to mood-state confusion and distortion (Widiger & Sanderson, 1995; Zimmerman, 1994). Persons who are depressed, anxious, angry, manic, hypomanic, or even just agitated are unlikely to provide accurate self-descriptions. Low self-esteem, hopelessness, and negativism are central features of depression and will naturally affect self-description. Persons who are depressed will describe themselves as being dependent, introverted self-conscious, vulnerable, and pessimistic (Widiger, 1993). Distortions will even continue after the remission of the more obvious, florid symptoms of depression (Hirschfeld et al., 1989).

Piersma (1989) reported significant decreases on the MCMI-II schizoid, avoidant, dependent, passive–aggressive, self-defeating, schizotypal, borderline, and paranoid personality scales across a brief inpatient treatment, even with the MCMI-II mood state correction scales. Significant increases were also obtained for the histrionic and narcissistic scales, which at first appeared nonsensical (suggesting that treatment had increased the presence of histrionic and narcissistic personality traits), until it was recognized that these scales include many items that involve self-confidence, assertion, and gregariousness. Piersma concluded that "the MCMI-II is not able to measure long-term personality characteristics ('trait' characteristics) independent of symptomatology ('state' characteristics)" (p. 91). Mood state distortion, however, might not be problematic within outpatient and normal community samples (Trull & Goodwin, 1993).

SSIs appear to be more successful than SRIs in distinguishing recently developed mental disorders from personality traits, particularly if the interviewers are instructed explicitly to make this distinction when they assess each item. Loranger et al. (1991) compared the assessments provided by the Personality Disorder Examination (PDE; Loranger, in press) at admission and one week to six months later. Reduction in scores on the PDE were not associated with depression or anxiety. Loranger et al. concluded that the "study provides generally encouraging results regarding the apparent ability of a particular semistructured interview, the PDE, to circumvent trait-state artifacts in diagnosing personality disorders in symptomatic patients" (p. 727).

However, it should not be presumed that SSIs are resilient to mood state distortions. O'Boyle and Self (1990) reported that "PDE dimensional scores were consistently higher (more symptomatic) when subjects were depressed" (p. 90) and Loranger et al. (1991) acknowledged as well that all but two of the PDE scales decreased significantly across the hospitalization. These changes are unlikely to reflect actual, fundamental changes in personality secondary to the brief psychiatric hospitalization.

The methods by which SRIs and SSIs address mood state and mental disorder confusion should be considered in future research (Zimmerman, 1994). For example, the *DSM-IV* requires that the personality disorder criteria be evident since late adolescence or young adulthood (APA, 1994). However, the PDE (Loranger, in press) requires only that one item be present since the age of 25, with the others present for only five years. A 45-year old adult with a mood disorder might then receive a dependent, borderline, or comparable personality disorder diagnosis by the PDE, if just one of the diagnostic criteria was evident since the age of 25.

3.07.3.2 Dissimulation and Distortion

Unstructured and semistructured interviewers base much of their assessments on the self-descriptions of the respondent. However, a substantial proportion of the assessment is also based on observations of a person's behavior and mannerisms (e.g., the schizotypal trait of odd or eccentric behavior; APA, 1994) the manner in which a person responds to questions (e.g., excessive suspiciousness in response to an innocuous question), and the consistency of the responses to questions across the interview. SSIs will also require that respondents provide examples of affirmative responses to ensure that the person understood the meaning or intention of a question. The provision of follow-up queries provides a significant advantage of SSIs relative to SRIs. For example, schizoid persons may respond affirmatively to an SRI question, "do you have any close friends," but further inquiry might indicate that they never invite these friends to their apartment, they rarely do anything with them socially, they are unaware of their friends' personal concerns, and they never confide in them. Widiger, Mangine, Corbitt, Ellis, and Thomas (1995) described a person who was very isolated and alone, yet indicated that she had seven very close friends with whom she confided all of her personal feelings and insecurities. When asked to describe one of these friends, it was revealed that she was referring to her seven cats.

In sum, none of the SSIs should or do accept a respondent's answers and self-descriptions simply at face value. Nevertheless, none of the personality disorder SSIs includes a formal method by which to assess defensiveness, dissimulation, exaggeration, or malingering. An assessment of exaggeration and defensiveness can be very difficult and complicated (Berry et al., 1995) yet it is left to the discretion and expertise of the interviewer in the assessment of each individual personality trait. Some interviewers may be very skilled at this assessment, whereas others may be inadequate in the effort or indifferent to its importance.

SSIs should perhaps include validity scales, comparable to those within self-report inventories. Alterman et al. (1996) demonstrated empirically that subjects exhibiting response sets of positive or negative impression management as assessed by the PAI (Morey, 1996) showed similar patterns of response distortion on two semistructured interviews, yet the interviewers appeared to be "essentially unaware of such behavior" (Alterman et al., 1996, p. 408). "The findings suggest that some individuals do exhibit response sets in the context of a structured interview and that this is not typically detected by the interviewer" (Alterman et al., 1996, p. 408). Much of the assessment of psychopathic personality traits by the revised Psychopathy Checklist (PCL-R; Hare, 1991) is based on a review of institutional file data rather than answers to interview questions, given the expectation that psychopathic persons will be characteristically deceptive and dishonest during an interview. It is unclear whether the PCL-R could provide a valid assessment of psychopathy in the absence of this additional, corrobatory information (Lilienfeld, 1994; Salekin, Rogers, & Sewell, 1996).

A notable exception to the absence of SSI validity scales is the Structured Interview of Reported Symptoms (SIRS; Rogers, Bagby, & Dickens, 1992) developed precisely for the assessment of subject distortion and dissimulation. The SIRS includes 172 items (sets of questions) organized into eight primary and five supplementary scales. Three of the primary scales assess for rare, improbable, or absurd symptoms, four assess for an unusual range and severity of symptoms, and the eighth assesses for inconsistencies in self-reported and observed symptoms. A substantial amount of supportive data has been obtained with the SIRS, particularly within forensic and neuropsychological settings (Rogers, 1995).

3.07.3.3 Intersite Reliability

The allowance for professional judgment in the selection of follow-up queries and in the interpretation of responses increases significantly the potential for inadequate interrater reliability. Good to excellent interrater reliability has been consistently obtained in the assessment of maladaptive personality traits with SSIs (Widiger & Sanderson, 1995; Zimmerman, 1994), but an SSI does not ensure the obtainment of adequate interrater reliability. An SSI only provides the means by which this reliability can be obtained. Most SSI studies report inadequate to poor interrater reliability for at least one of the domains of personality being assessed (see Widiger & Sanderson, 1995; Zimmerman, 1994). The obtainment of adequate interrater reliability should be assessed and documented in every study in which an SSI is being used, as it is quite possible that the personality disorder of particular interest is the one for which weak interrater reliability has been obtained.

The method by which interrater reliability has been assessed in SSI research is also potentially misleading. Interrater reliability has traditionally been assessed in SSI research with respect to the agreement between ratings provided by an interviewer and ratings provided by a person listening to an audiotaped (or videotaped) recording of this interview (Zimmerman, 1994). However, this methodology assesses only the agreement regarding the ratings of the respondents' statements. It is comparable to confining the assessment of the interrater reliability of practicing clinicians' personality assessments to the agreement in their ratings of a recording of an unstructured clinical interview. The poor interrater reliability that has been reported for practicing clinicians has perhaps been due largely to inconsistent, incomplete, and idiosyncratic interviewing (Widiger & Sanderson, 1995). It is unclear whether personality disorder SSIs have actually resolved this problem as few studies have in fact assessed interrater reliability using independent administrations of the same interview.

The misleading nature of currently reported SSI interrater reliability is most evident in studies that have used telephone interviews. Zimmerman and Coryell (1990) reported kappa agreement rates of 1.0 for the assessment of the schizotypal, histrionic, and dependent personality disorders using the Structured Interview for *DSM-III* Personality (SIDP; Pfohl, Blum, & Zimmerman, in press). However, 86% of the SIDP administrations were by telephone. Telephone administrations of an SSI will tend to be more structured than face-to-face interviews, and perhaps prone to brief and simplistic administrations. They may degenerate into a verbal administration of an SRI. Interrater agreement with respect to the ratings of subjects' affirmative or negative responses to MMPI-2 items (i.e., highly structured questions) is not particularly informative.

There is reason to believe that there might be poor agreement across research sites using the same SSI. For example, the prevalence rate of borderline personality disorder within psychiatric settings has been estimated at 15% (Pilkonis et al., 1995), 23% (Riso, Klein, Anderson, Ouimette, & Lizardi, 1994), 42% (Loranger et al., 1991), and 71% (Skodol, Oldham, Rosnick, Kellman, & Hyler, 1991), all using the same PDE (Loranger, in press). This substantial disagreement is due to many different variables (Gunderson, 1992; Pilkonis, 1997; Shea, 1995), but one distinct possibility is that there is an inconsistent administration and scoring of the PDE by Loranger et al. (1991), Pilkonis et al. (1995), Riso et al. (1994), and Skodol et al. (1991). Excellent interrater reliability of a generally unreliably administered SSI can be obtained within one particular research site through the development of local (idiosyncratic) rules and policies for the administration of follow-up questions and the scoring of respondents' answers that are inconsistent with the rules and policies developed for this same SSI at another research site.

3.07.3.4 Self and Informant Ratings

Most studies administer the SSI to the person being assessed. However, a method by which to address dissimulation and distortion is to administer the SSI to a spouse, relative, or friend who knows the person well. Many personality traits involve a person's manner of relating to others (McCrae & Costa, 1989; Wiggins & Trobst, 1997) and some theorists suggest that personality is essentially this manner of relating to others (Kiesler, 1996; Westen, 1991; Wiggins & Pincus, 1992).

A useful source for the description of these traits would then be persons with whom the subject has been interacting. These "informants" (as they are often identified) can be intimately familiar with the subject's characteristic manner of relating to them, and they would not (necessarily) share the subject's distortions in self-description. The use of peer, spousal, and other observer ratings of personality has a rich tradition in SRI research (Ozer & Reise, 1994).

An interview with a close friend, spouse, employer, or relative is rarely uninformative and typically results in the identification of additional maladaptive personality traits. However, it is unclear which source will provide the most valid information. Zimmerman, Pfohl, Coryell, Stangl, and Corenthal (1988) administered an SSI to both a patient and an informant, and obtained rather poor agreement, with correlations ranging from 0.17 (compulsive) to only 0.66 (antisocial). The informants identified significantly more dependent, avoidant, narcissistic, paranoid, and schizotypal traits, but Zimmerman et al. (1988) concluded that "patients were better able to distinguish between their normal personality and their illness" (p. 737). Zimmerman et al. felt that the informants were more likely to confuse patients' current depression with their longstanding and premorbid personality traits than the patients' themselves. Similar findings have been reported by Riso et al. (1994). Informants, like the patients themselves, are providing subjective opinions and impressions rather than objective descriptions of behavior patterns, and they may have their own axes to grind in their emotionally invested relationship with the identified patient. The fundamental attribution error is to overexplain behavior in terms of personality traits, and peers might be more susceptible to this error than the subjects themselves.

3.07.3.5 Convergent Validity Across Different Interviews

One of the most informative studies on the validity of personality disorder SSIs was provided by Skodol et al. (1991). They administered to 100 psychiatric inpatients two different SSIs on the same day (alternating in morning and afternoon administrations). Agreement was surprisingly poor, with kappa ranging from a low of 0.14 (schizoid) to a high of only 0.66 (dependent). Similar findings have since been reported by Pilkonis et al. (1995). If this is the best agreement one can obtain with the administration of different SSIs to the same persons by the same research team on the same day, imagine the disagreement that must be obtained by different research teams with different SSIs at different sites (however, both studies did note that significantly better agreement was obtained when they considered the dimensional rating of the extent to which each personality syndrome was present).

Skodol et al. (1991) concluded that the source of the disagreement they obtained between the Structured Clinical Interview for *DSM-III-R* Personality Disorders (SCID-II; First, Gibbon, Spitzer Williams, & Benjamin, in press) and the PDE (Loranger, in press) was the different questions (or items) used by each interview. "It is fair to say that, for a number of disorders (i.e., paranoid, schizoid, schizotypal, narcissistic, and passive–aggressive) the two [interviews] studied do not operationalize the diagnoses similarly and thus yield disparate results" (Skodol et al., p. 22).

There appear to be important differences among the SSIs available for the assessment of personality disorders (Clark,1992; Widiger & Sanderson, 1995; Zimmerman, 1994). Table 1 presents the number of structured questions (i.e., answerable by one word, such as "yes" or "frequently"), open-ended questions, and observational ratings provided in each of the five major personality disorder SSIs. It is evident from Table 1 that there is substantial variation across SSIs simply with respect to the number of questions provided, ranging from 193 to 373 with respect to structured questions and from 20 to 69 for open-ended questions. There is also variability in the reliance upon direct observations of the respondent. The PDE, Diagnostic Interview for *DSM-IV* Personality Disorders (DIPD-IV), and SIDP-IV might be more difficult to administer via telephone, given the number of observational ratings of behavior, appearance, and mannerisms that are required (particularly to assess the schizotypal and histrionic personality disorders).

There is also significant variability in the content of the questions used to assess the same personality trait (Clark, 1992; Widiger & Sanderson, 1995; Zimmerman, 1994). Table 2 presents the questions used by each of the five major SSIs to assess the borderline trait of identity disturbance, defined in *DSM-IV* as a markedly and persistently unstable sense of self-image or sense of self (APA, 1994). All five SSIs do ask about significant or dramatic changes in self-image across time. However, there are also notable differences. For example, the DIPD-IV and SIDP-IV refer specifically to feeling evil; the SIDP-IV highlights in particular a confusion regarding sexual orientation; the SCID-II appears to emphasize changes or fluctuations in self-image, whereas the DIPD-IV appears to emphasize an uncertainty or absence of self-image; and the Personality Disorder Interview-4 (PDI-IV) includes more open-ended self-descriptions.

It is possible that the variability in content of questions across different SSIs may not be as important as a comparable variability in questions across different SRIs, as the interviews may converge in their assessment through the unstructured follow-up questions, queries and clarifications. However, the findings of Pilkonis et al. (1995) and Skodol et al. (1991) suggest otherwise. In addition, as indicated in Table 1, most of the questions within the five SSIs are relatively structured, resulting perhaps in very little follow-up query.

3.07.3.6 Illustrative Instruments

The three most commonly used SSIs within clinical research for the assessment of the *DSM* personality disorders are the SIDP-IV (Pfohl et al., in press), the SCID-II (First et al., in press), and the PDE (Loranger, in press). The SIDP-IV has been used in the most number of studies. A distinctive feature of the PDE is the inclusion of the diagnostic criteria for the personality disorders of the World Health Organization's (WHO) *International Classification of Diseases* (*ICD-10*; 1992). However, using this international version of the PDE to compare the *DSM-IV* with the *ICD-10* is problematic, as the PDE does not in fact provide distinct questions for the *ICD-10* criteria. It simply indicates which of the existing PDE questions, developed for the assessment of the *DSM-IV* personality disorders, could be used to assess the *ICD-10* criteria. The *DSM-IV* and *ICD-10* assessments are not then independent. The DIPD (Zanarini, Frankenburg, Sickel, & Yong, 1996) is the youngest SSI, but it is currently being used in an extensive multi-site, longitudinal study of personality disorders.

3.07.3.6.1 SIDP-IV

The SIDP-IV (Pfohl et al., in press) is the oldest and most widely used SSI for the

Table 1 Amount of structure in personality disorder semistructured interviews.

Interview	*Number of questions*				*Observations*
	Structured	*Open-ended*	*Examples*	*Total*	
DIPD-IV	373	20	5	398	19
PDE	272	69	196	537	32
PDI-IV	290	35		325	3
SCID-II	193	35	75	303	7
SIDP-IV	244	58	35	337	16

Note. Examples = specified request for examples (PDI-IV instructs interviewers to always consider asking for examples, and therefore does not include a specified request for individual items); DIPD-IV = Diagnostic Interview for *DSM-IV* Personality Disorders (Zanarini et al., 1996); PDE = Personality Disorder Examination (Loranger, in press); PDI-IV = Personality Disorder Interview-4 (Widiger et al., 1995); SCID-II = Structured Clinical Interview for *DSM-IV* Axis II Personality Disorders (First, et al., in press); SIDP-IV = Structured Interview for *DSM-IV* Personality (Pfohl et al., in press).

assessment of the *DSM* personality disorders (Rogers, 1995; Widiger & Sanderson, 1995; Zimmerman, 1994). It includes 353 items (337 questions and 16 observational ratings) to assess the 94 diagnostic criteria for the 10 *DSM-IV* personality disorders. Additional items are provided to assess the proposed but not officially recognized negativistic (passive–aggressive), depressive, self-defeating, and sadistic personality disorders. It is available in two versions, one in which the items are organized with respect to the diagnostic criteria sets (thereby allowing the researcher to assess only a subset of the disorders) and the other in which the items are organized with respect to similar content (e.g., perceptions of others and social conformity) to reduce repetition and redundancy. Each diagnostic criterion is assessed on a four-point scale (0 = not present, 1 = subthreshold, 2 = present, 3 = strongly present). Administration time is about 90 minutes, depending in part on the experience of the interviewer and the verbosity of the subject. A computerized administration and scoring system is available. The accompanying instructions and manual are limited, but training videotapes and courses are available.

3.07.3.6.2 PDI-IV

The PDI-IV (Widiger et al., 1995) is the second oldest SSI for the assessment of the *DSM* personality disorders. It is comparable in content and style to the SIDP-IV, although it has fewer observational ratings and more open-ended inquiries. The PDI-IV has been used in substantially fewer studies than the SIDP-IV, PDE, or SCID-II, it lacks supportive training material, and only one of the published studies using the PDI-IV was conducted by independent investigators (Rogers, 1995). However, its accompanying manual is the most systematic and comprehensive, providing extensive information regarding the history, rationale, and common assessment issues for each of the *DSM-IV* personality disorder diagnostic criteria.

3.07.4 PROJECTIVE TECHNIQUES

Projective techniques are not used as often as SRIs in personality research, "and many academic psychologists have expressed the belief that knowledge of projective testing is not as important as it used to be and that use of projective tests will likely decline in the future" (Butcher & Rouse, 1996, p. 91). However, "of the top 10 assessment procedures [used by clinicians] ... 4 are projectives, and another (Bender-Gestalt) is sometimes used for projective purposes" (Watkins et al., 1995, p. 59). Butcher and Rouse (1996) document as well that the second most frequently researched clinical instrument continues to be the Rorschach. "Predictions about the technique's demise appear both unwarranted and unrealistic" (Butcher & Rouse, 1996, p. 91). "Whatever negative opinions some academics may hold about projectives, they clearly are here to stay, wishing will not make them go away ... and their place in clinical assessment practice now seems as strong as, if not stronger than, ever" (Watkins et al., 1995, p. 59).

The term "projective," however, may be somewhat misleading, as it suggests that these tests share an emphasis upon the interpretation of a projection of unconscious conflicts, impulses, needs, or wishes onto ambiguous stimuli. This is true for most projective tests but it is not in fact the case for the most commonly used scoring system (i.e., Exner, 1993) for the most commonly used projective test (i.e., the Rorschach). The Exner (1993) Comprehensive

Table 2 Semistructured interview questions for the assessment of identity disturbance.

DIPD-IV

1. During the past two years, have you often been unsure of who you are or what you're really like?
2. Have you frequently gone from feeling sort of OK about yourself to feeling that you're bad or even evil?
3. Have you often felt that you had no identity?
4. How about that you had no idea of who you are or what you believe in?
5. That you don't even exist?

PDE

1. Do you think one of your problems is that you're not sure what kind of person you are?
2. Do you behave as though you don't know what to expect of yourself?
3. Are you so different with different people or in different situations that you don't behave like the same person?
4. Have others told you that you're like that? Why do you think they've said that?
5. What would you like to accomplish during your life? Do your ideas about this change often?
6. Do you often wonder whether you've made the right choice of job or career? (If housewife, ask:) Do you often wonder whether you've made the right choice in becoming a housewife? (If student, ask:) Have you made up your mind about what kind of job or career you would like to have?
7. Do you have trouble deciding what's important in life?
8. Do you have trouble deciding what's morally right and wrong?
9. Do you have a lot of trouble deciding what type of friends you should have?
10. Does the kind of people you have as friends keep changing?
11. Have you ever been uncertain whether you prefer a sexual relationship with a man or a woman?

PDI-IV

1. How would you describe your personality?
2. What is distinct or unique about you?
3. Do you ever feel that you don't know who you are or what you believe in?
4. Has your sense of who you are, what you value, what you want from life, or what you want to be, been fairly consistent, or has this often changed significantly?

SCID-II

1. Have you all of a sudden changed your sense of who you are and where you are headed?
2. Does your sense of who you are often change dramatically?
3. Are you different with different people or in different situations so that you sometimes don't know who you really are?
4. Have there been lots of sudden changes in your goals, career plans, religious beliefs, and so on?

SIDP-IV

1. Does the way you think about yourself change so often that you don't know who you are anymore?
2. Do you ever feel like you're someone else, or that you're evil, or maybe that you don't even exist?
3. Some people think a lot about their sexual orientation, for instance, trying to decide whether or not they might be gay (or lesbian). Do you often worry about this?

Note. *DIPD-IV* = Diagnostic Interview for *DSM-IV* Personality Disorders (Zanarini et al., 1996, pp. 25–26); PDE = Personality Disorder Examination (Loranger, in press, pp. 58–60, 83, & 117); PDI-IV = Personality Disorder Interview-4 (Widiger et al., 1995, p. 92); SCID-II = Structured Clinical Interview for *DSM-IV* Axis II Personality Disorders (First et al., in press, p. 37); SIDP-IV = Structured Interview for *DSM-IV* Personality (Pfohl et al., in press, p. 19).

System does include a few scores that appear to concern a projection of personal needs, wishes, or preoccupations (e.g., morbid content) but much of the scoring concerns individual differences in the perceptual and cognitive processing of the form or structure of the shapes, textures, details, and colors of the ambiguous inkblots. Exner (1989) has himself stated that "unfortunately, the Rorschach has been erroneously mislabeled as a projective test for far too long" (p. 527).

The label "projective test" is also contrasted traditionally with the label "objective test" (e.g., Keller et al., 1990) but this is again misleading. The Exner (1993) Comprehensive System scoring for the Rorschach is as objective as the scoring of an MMPI-2 (although not as reliable). In addition, clinicians often can interpret MMPI-2 profiles in an equally subjective manner.

A more appropriate distinction might be a continuum of structure vs. ambiguity with respect to the stimuli provided to the subject and the range of responses that are allowed. Most of the techniques traditionally labeled as projective do provide relatively more ambiguous stimuli than either SRIs or SSIs (e.g., inkblots or drawings). However, many SRI items can be as ambiguous as an item from a projective test. Consider, for example, the MMPI-2 items "I like mechanics magazines," "I used to keep a diary," and "I would like to be a journalist" (Hathaway et al., 1989, pp. 5, 8, 9). These items are unambiguous in their content, but the trait or characteristic they assess is very ambiguous (Butcher, 1995). There is, perhaps, less ambiguity in the meaning of the stems provided in a sentence completion test, such as "My conscience bothered me most when," "I used to dream about" and "I felt inferior when" (Lah, 1989, p. 144) than in many of the MMPI-2 items.

SRIs, on the other hand, are much more structured (or constraining) in the responses that are allowed to these stimuli. The only responses can be "true" or "false" to an ambiguous MMPI-2 item, whereas anything can be said in response to a more obvious sentence completion stem. Projective tests are uniformly more open-ended than SRIs in the responses that are allowed, increasing substantially the potential for unreliability in scoring.

However, SSIs can be as open-ended as many projective tests in the responses that are allowed. For example, the PDI-IV SSI begins with the request of "having you tell me the major events, issues, or incidents that you have experienced since late childhood or adolescence" (Widiger et al., 1995, p. 245). This initial question is intentionally open-ended to ensure that the most important events or incidents are not neglected during the interview (Perry, 1992). However, such open-ended inquiries will contribute to problematic intersite reliability.

3.07.4.1 Rorschach

The Rorschach consists of 10 ambiguously shaped inkblots, some of which include various degrees of color and shading. The typical procedure is to ask a person what each inkblot might be, and to then follow at some point with inquiries that clarify the bases for the response (Aronow, Reznikoff, & Moreland, 1994; Exner, 1993). The Rorschach has been the most popular projective test in clinical practice since the Second World War, although the Thematic Apperception Test (TAT) and sentence completion tests are gaining in frequency of usage (Watkins et al., 1995).

There is empirical support for many Rorschach variables (Bornstein, 1996; Exner, 1993; Weiner, 1996), although the quality of some of this research has been questioned, including the support for a number of the fundamental Exner variables, such as the experience ratio (Kleiger, 1992; Wood, Nezworski, & Stejskal, 1996). The Rorschach can be administered and scored in a reliable manner, but the training that is necessary to learn how to score reliably the 168 variables of the Exner Comprehensive System is daunting, at best. SRIs and SSIs might provide a more cost-efficient method to obtain the same results (although profile interpretation of the MMPI-2 clinical scales can at times be equally complex; Helmes & Reddon, 1993). Exner (1996) has acknowledged, at least with respect to a depression index, that "there are other measures, such as the [MMPI-2], that might identify the presence of reported depression much more accurately than the Rorschach" (p. 12). The same point can perhaps be made for the assessment of personality traits, such as narcissism and dependency. Bornstein (1995), however, has argued that the Rorschach provides a less biased measure of sex differences in personality (dependency, in particular) because its scoring is less obvious to the subject. He suggested that the findings from SRIs and SSIs have provided inaccurate estimates of dependency in males because males are prone to deny the extent of their dependent personality traits in response to SRIs and SSIs.

An additional issue for the Rorschach is that its relevance to personality research and assessment is at times unclear. For example, the cognitive-perceptual mechanisms assessed by the Exner Comprehensive system do not appear to be of central importance to many theories of personality and personality disorder (Kleiger, 1992). "It is true that the Rorschach does not offer a precise measure for any single personality trait" (Exner, 1997, p. 41). How a person perceptually organizes an ambiguous inkblot may indeed relate to extratensive ideational activity, but constructs such as introversion, conscientiousness, need for affection, and empathy have a more direct, explicit relevance to current personality theory and research. More theoretically meaningful constructs are perhaps assessed by content (e.g., Aronow et al., 1994; Bornstein, 1996) or object-relational (e.g., Lerner, 1995) scoring systems. Content interpretations of the Rorschach are more consistent with the traditional understanding of the instrument as a projective stimulus, but this approach also lacks the empirical support of the cognitive-perceptual scoring systems and may only encourage a return to less reliable and subjective interpretations (Acklin, 1995).

3.07.4.2 Thematic Apperception Test

The TAT (Cramer, 1996) consists of 31 cards: one is blank, seven are for males, seven for females, one for boys or girls, one for men or women and one each for a boy, girl, man, and woman (the remaining 10 are for anyone). Thus, a complete set for any particular individual could consist of 20 stimulus drawings, although only 10 typically are used per person. Most of the drawings include person(s) in an ambiguous but emotionally provocative context. The instruction to the subject is to make up a dramatic story for each card, describing what is happening, what led up to it, what is the outcome, and what the persons are thinking and feeling. It is common to describe the task as a test of imaginative intelligence to encourage vivid, involved, and nondefensive stories.

The TAT is being used increasingly in personality research with an interpersonal or object-relational perspective (e.g., Westen, 1991). The TAT's provision of cues for a variety of interpersonal issues and relationships make it particularly well suited for such research, and the variables assessed are theoretically and clinically meaningful (e.g., malevolent vs. benevolent affect and the capacity for an emotional investment in relationships). The necessary training for reliable scoring is also less demanding than for the Rorschach, although a TAT administration remains time-consuming. There are many SRI measures of closely related interpersonal constructs that are less expensive and complex to administer and score (e.g., Kiesler, 1996; Wiggins & Trobst, 1997).

3.07.5 CONCLUSIONS

The assessment of personality is a vital component of clinical research and practice, particularly with the increasing recognition of the importance of personality traits to the development and treatment of psychopathology (Watson et al., 1994). The assessment of adaptive and maladaptive personality functioning is fundamental to virtually all fields of applied psychology.

It is then surprising and regrettable that clinical psychology training programs provide so little attention to the importance of and methods for obtaining comprehensive and systematic interviewing. The primary method for the assessment of personality in clinical practice is an unstructured interview that has been shown to be quite vulnerable to misleading expectations, inadequate coverage, and gender and ethnic biases. Training programs will devote a whole course, perhaps a whole year, to learning different projective techniques, but may never even inform students of the existence of any particular semistructured interview.

The preferred method of assessment in personality disorder research appears to be SSIs (Zimmerman, 1994), whereas the preferred method in normal personality research are SRIs (Butcher & Rouse, 1996). However, the optimal approach for both research and clinical practice would be a multimethod assessment, using methods whose errors of measurement are uncorrelated. No single approach will be without significant limitations. The convergence of findings across SRI, SSI, and projective methodologies would provide the most compelling results.

3.07.6 REFERENCES

Acklin, M. W. (1995). Integrative Rorschach interpretation. *Journal of Personality Assessment, 64,* 235–238.

Adler, D. A., Drake, R. E., & Teague, G. B. (1990). Clinicians' practices in personality assessment: Does gender influence the use of DSM-III Axis II? *Comprehensive Psychiatry, 31,* 125–133.

Adler, N., & Matthews, K. (1994) Health psychology: Why do some people get sick and some stay well? *Annual Review of Psychology, 45,* 229–259.

Alterman, A. I., Snider, E. C., Cacciola, J. S., Brown, L. S., Zaballero, A., & Siddiqui, N. (1996). Evidence for response set effects in structured research interviews. *Journal of Nervous and Mental Disease, 184,* 403–410.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders.* (4th ed.). Washington, DC: Author.

Aronow, E., Reznikoff, M., & Moreland, K. (1994). *The Rorschach technique: Perceptual basics, content, interpretation and applications.* Boston: Allyn and Bacon.

Baumeister, R. F., Tice, D. M., & Hutton, D. G. (1989). Self-presentational motivations and personality differences in self-esteem. *Journal of Personality, 57,* 547–579.

Ben-Porath, Y. S. (1994). The MMPI and MMPI-2: Fifty years of differentiating normal and abnormal personality. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (pp. 361–401). New York: Springer.

Ben-Porath, Y. S., McCully, E., & Almagor, M. (1993). Incremental validity of the MMPI-2 content scales in the assessment of personality and psychopathology by self-report. *Journal of Personality Assessment, 61,* 557–575.

Ben-Porath, Y. S., & Waller, N. G. (1992). "Normal" personality inventories in clinical assessment: General requirements and the potential for using the NEO Personality Inventory. *Psychological Assessment, 4,* 14–19.

Berry, D. T. R. (1995). Detecting distortion in forensic evaluations with the MMPI-2. In Y. S. Ben-Porath, J. R. Graham, G. C. N. Hall, R. D. Hirschman, & M. S. Zaragoza (Eds.), *Forensic applications of the MMPI-2* (pp. 82–102). Thousands Oaks, CA: Sage.

Berry, D. T. R., Wetter, M. W., & Baer, R. A. (1995). Assessment of malingering. In J. N. Butcher (Ed.), *Clinical personality assessment. Practical approaches* (pp. 236–248). New York: Oxford University Press.

Blashfield, R. K., & Herkov, M. J. (1996). Investigating clinician adherence to diagnosis by criteria: A replication of Morey and Ochoa (1989). *Journal of Personality Disorders, 10,* 219–228.

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin, 117,* 187–215.

Bornstein, R. F. (1995). Sex differences in objective and projective dependency tests: A meta-analytic review. *Assessment, 2,* 319–331.

Bornstein, R. F. (1996). Construct validity of the Rorschach oral dependency scale: 1967–1995. *Psychological Assessment, 8,* 200–205.

Borkenau, P., & Ostendorf, F. (1992). Social desirability scales as moderator and suppressor variables. *European Journal of Personality, 6,* 199–214.

Butcher, J. N. (1995). Item content in the interpretation of the MMPI-2. In J. N. Butcher (Ed.), *Clinical personality assessment. Practical approaches* (pp. 302–316). New York: Oxford University Press.

Butcher, J. N. (1995). How to use computer-based reports. In J. N. Butcher (Ed.), *Clinical personality assessment. Practical approaches* (pp. 78–94). New York: Oxford University Press.

Butcher, J. N., & Rouse, S. V. (1996). personality: Individual differences and clinical assessment. *Annual Review of Psychology, 47,* 87–111.

Clark, L. A. (1992). Resolving taxonomic issues in personality disorders. The value of large-scale analysis of symptom data. *Journal of Personality Disorders, 6,* 360–376.

Clark, L. A. (1993). *Manual for the schedule for nonadaptive and adaptive personality*. Minneapolis, MN: University of Minnesota Press.

Clark, L. A., Livesley, W. J., Schroeder, M. L., & Irish, S. L. (1996). Convergence of two systems for assessing specific traits of personality disorder. *Psychological Assessment, 8,* 294–303.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7,* 309–319.

Cloninger, C. R., & Svrakic, D. M. (1994). Differentiating normal and deviant personality by the seven factor personality model. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (pp. 40–64). New York: Springer.

Colligan, R. C., Morey, L. C., & Offord, K. P. (1994). The MMPI/MMPI-2 personality disorder scales. Contemporary norms for adults and adolescents. *Journal of Clinical Psychology, 50,* 168–200.

Cooke, D. J., & Michie, C. (1997). An item response theory analysis of the Hare Psychopathy Checklist-Revised. *Psychological Assessment, 9,* 3–14.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., & McCrae, R. R. (1994). Set like plaster? Evidence for the stability of adult personality. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change?* (pp. 21–40). Washington, DC: American Psychological Association.

Costa, P. T., & McCrae, R. R. (1997). Stability and change in personality assessment: The Revised NEO Personality Inventory in the year 2000. *Journal of Personality Assessment, 68,* 86–94.

Costa, P. T., Zonderman, A. B., McCrae, R. R., & Williams, R. B. (1985). Content and comprehensiveness in the MMPI: An item factor analysis in a normal adult sample. *Journal of Personality and Social Psychology, 48,* 925–933.

Cramer, P. (1996). *Storytelling, narrative, and the Thematic Apperception Test*. New York: Guilford.

Dawes, R. M. (1994). *House of cards. Psychology and psychotherapy built on myth*. New York: Free Press.

DiLalla, D. L., Gottesman, I. I., Carey, G., & Vogler, G. P. (1993). Joint factor structure of the Multidimensional Personality Questionnaire and the MMPI in a psychiatric and high-risk sample. *Psychological Assessment, 5,* 207–215.

Eagly, A. H. (1995). The science and politics of comparing women and men. *American Psychologist, 50,* 145–158.

Embretson, S. E. (1996). The new rules of measurement. *Psychological Assessment, 8,* 341–349.

Exner, J. E. (1989). Searching for projection in the Rorschach. *Journal of Personality Assessment, 53,* 520–536.

Exner, J. E. (1993). *The Rorschach: A comprehensive system* (Vol. 1). New York: Wiley.

Exner, J. E. (1996). A comment on "The Comprehensive System for the Rorschach: A critical examination." *Psychological Science, 7,* 11–13.

Exner, J. E. (1997). The future of Rorschach in personality assessment. *Journal of Personality Assessment, 68,* 37–46.

Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116,* 429–456.

Finn, S. E., & Kamphuis, J. H. (1995). What a clinician needs to know about base rates. In J. N. Butcher (Ed.), *Clinical personality assessment. Practical approaches* (pp. 224–235). New York: Oxford University Press.

First, M. B., Gibbon, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (in press). *User's guide for the Structured Clinical Interview for DSM-IV Axis II Personality Disorders*. Washington, DC: American Psychiatric Press.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7,* 286–299.

Ford, M. R., & Widiger, T. A. (1989). Sex bias in the diagnosis of histrionic and antisocial personality disorders. *Journal of Consulting and Clinical Psychology, 57,* 301–305.

Garb, H. N. (1997). Race bias, social class bias, and gender bias in clinical judgment. *Clinical Psychology: Science and Practice.*

Goldberg, L. R. (1990). An alternative "Description of personality": The Big Five factor structure. *Journal of Personality and Social Psychology, 59,* 1216–1229.

Graham, J. R. (1993). *MMPI-2. Assessing personality and psychopathology* (2nd ed.). New York: Oxford University Press.

Greene, R. L., Gwin, R., & Staal, M. (1997). Current status of MMPI-2 research: A methodologic overview. *Journal of Personality Assessment, 68,* 20–36.

Gunderson, J. G. (1992). Diagnostic controversies. In A. Tasman & M. B. Riba (Eds.), *Review of psychiatry* (Vol. 11, pp. 9–24). Washington, DC: American Psychiatric Press.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.

Hare, R. D. (1991). *Manual for the Revised Psychopathy Checklist* Toronto, Canada: Multi-Health Systems.

Hare, R. D., Hart, S. D., & Harpur, T. J. (1991). Psychopathy and the DSM-IV criteria for antisocial personality disorder. *Journal of Abnormal Psychology, 100,* 391–398.

Harkness, A. R., & McNulty, J. L. (1994). The Personality Psychopathology Five (PSY-5): Issue from the pages of a diagnostic manual instead of a dictionary. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (pp. 291–315). New York. Springer.

Harkness, A. R., McNulty, J. L., & Ben-Porath, Y. S. (1995). The Personality Psychopathology Five (PSY-5): Constructs and MMPI-2 scales. *Psychological Assessment, 7,* 104–114.

Hathaway, S. R., & McKinley, J. C. (1940). A multiphasic personality schedule (Minnesota): I. Construction of the schedule. *Journal of Psychology, 10,* 249–254.

Hathaway, S. R., & McKinley, J. C. (1982). *Minnesota Multiphasic Personality Inventory test booklet*. Minneapolis, MN: University of Minnesota.

Hathaway, S. R., McKinley, J. C., Butcher, J. N., Dahlstrom, W. G., Graham, J. R., & Tellegen, A. (1989). *Minnesota Multiphasic Personality Inventory test booklet*. Minneapolis, MN: Regents of the University of Minnesota.

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7,* 238–247.

Helmes, E., & Reddon, J. R. (1993). A perspective on developments in assessing psychopathology: A critical review of the MMPI and MMPI-2. *Psychological Bulletin, 113,* 453–471.

Hendryx, M. S., Haviland, M. G., Gibbons, R. D., & Clark, D. C. (1992). An application of item response theory to alexithymia assessment among abstinent alcoholics. *Journal of Personality Assessment, 58,* 506–515.

Herkov, M. J., & Blashfield, R. K. (1995). Clinician diagnoses of personality disorders: Evidence of a hierarchical structure. *Journal of Personality Assessment, 65,* 313–321.

Hirschfeld, R. M., Klerman, G. L., Lavori, P., Keller, M., Griffith, P., & Coryell, W. (1989). Premorbid personality assessments of first onset of major depression. *Archives of General Psychiatry, 46,* 345–350.

Hogan, R., Curphy, G. J., & Hogan, J. (1994). What we know about leadership. Effectiveness and personality. *American Psychologist, 49,* 493–504.

Hyler, S. E. (1994). *Personality Diagnostic Questionnaire-4 (PDQ-4)*. Unpublished test. New York: New York State Psychiatric Institute.

Kehoe, J. F., & Tenopyr, M. L. (1994). Adjustment in assessment scores and their usage: A taxonomy and evaluation of methods. *Psychological Assessment, 6,* 291–303.

Keller, L. S., Butcher, J. N., & Slutske, W. S. (1990). Objective personality assessment. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment* (2nd ed., pp. 345–386). New York: Pergamon.

Kiesler, D. J. (1996). *Contemporary interpersonal theory & research, personality, psychopathology, and psychotherapy*. New York: Wiley.

Kleiger, J. H. (1992). A conceptual critique of the EA;es comparison in the Comprehensive Rorschach System. *Psychological Assessment, 4,* 288–296.

Klein, M. H., Benjamin, L. S., Rosenfeld, R., Treece, C., Husted, J., & Greist, J. H. (1993). The Wisconsin Personality Disorders Inventory: development, reliability, and validity. *Journal of Personality Disorders, 7,* 285–303.

Lah, M. I. (1989). Sentence completion tests. In C. S. Newmark (Ed.), *Major psychological assessment instruments* (Vol. II, pp. 133–163). Boston: Allyn & Bacon.

Lamiell, J. T. (1981). Toward an idiothetic psychology of personality. *American Psychologist, 36,* 276–289.

Lerner, P. M. (1995). Assessing adaptive capacities by means of the Rorschach. In J. N. Butcher (Ed.), *Clinical personality assessment. Practical approaches* (pp. 317–325). New York: Oxford University Press.

Lilienfeld, S. O. (1994). Conceptual problems in the assessment of psychopathy. *Clinical Psychology Review, 14,* 17–38.

Lilienfeld, S. O., & Andrews, B. P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal populations. *Journal of Personality Assessment, 66,* 488–524.

Lindsay, K. A., & Widiger, T. A. (1995). Sex and gender bias in self-report personality disorder inventories: Item analyses of the MCMI-II, MMPI, and PDQ-R. *Journal of Personality Assessment, 65,* 1–20.

Livesley, W. J., & Jackson, D. (in press). *Manual for the Dimensional Assessment of Personality Pathology-Basic Questionnaire*. Port Huron, MI: Sigma.

Loranger, A. W. (in press). *Personality disorder examination*. Washington, DC: American Psychiatric Press.

Loranger, A. W., Lenzenweger, M. F., Gartner, A. F., Susman, V. L., Herzig, J., Zammit, G. K., Gartner, J. D., Abrams, R. C., & Young, R. C. (1991). Trait-state artifacts and the diagnosis of personality disorders. *Archives of General Psychiatry, 48,* 720–728.

McCrae, R. R., & Costa, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology, 51,* 882–888.

McCrae, R. R., & Costa, P. T. (1989). The structure of interpersonal traits: Wiggins' circumplex and the Five-Factor Model. *Journal of Personality and Social Psychology, 56,* 586–595.

Meehl, P. E. (1945). The dynamics of "structured" personality tests. *Journal of Clinical Psychology, 1,* 296–303.

Millon, T. (1983). *Millon Clinical Multiaxial Inventory manual* (3rd ed.). Minneapolis, MN: National Computer Systems.

Millon, T. (1987). *Manual for the MCMI-II* (2nd ed.). Minneapolis, MN: National Computer Systems.

Millon, T., Davis, R. D., Millon, C. M., Wenger, A. W., Van Zuilen, M. H., Fuchs, M., & Millon, R. B. (1996). *Disorders of personality. DSM-IV and beyond.* New York: Wiley.

Millon, T., Millon, C., & Davis, R. (1994). *MCMI-III manual.* Minneapolis, MN: National Computer Systems.

Morey, L. C. (1996). *An interpretive guide to the Personality Assessment Inventory (PAI)*. Odessa, FL: Psychological Assessment Resources.

Morey, L. C., & Ochoa, E. S. (1989). An investigation of adherence to diagnostic criteria: Clinical diagnosis of the DSM-III personality disorders. *Journal of Personality Disorders, 3,* 180–192.

Morey, L. C., Waugh, M. H., & Blashfield, R. K. (1985). MMPI scales for DSM-III personality disorders: Their derivation and correlates. *Journal of Personality Assessment, 49,* 245–251.

O'Boyle, M., & Self, D. (1990). A comparison of two interviews for DSM-III-R personality disorders. *Psychiatry Research, 32,* 85–92.

Okazaki, S., & Sue, S. (1995). Methodological issues in assessment research with ethnic minorities. *Psychological Assessment, 7,* 367–375.

Ozer, D. J., & Reise, S. P. (1994). Personality assessment. *Annual Review of Psychology, 45,* 357–388.

Perry, J. C. (1992). Problems and considerations in the valid assessment of personality disorders. *American Journal of Psychiatry, 149,* 1645–1653.

Pfohl B., Blum, N., & Zimmerman, M. (in press). *Structured Interview for DSM-IV Personality*. Washington, DC: American Psychiatric Press.

Piersma, H. L. (1989). The MCMI-II as a treatment outcome measure for psychiatric inpatients. *Journal of Clinical Psychology, 45,* 87–93

Pilkonis, P. A. (1997). Measurement issues relevant to personality disorders. In H. H. Strupp, M. J. Lambert, & L. M. Horowitz (Eds.), *Measuring patient change in mood, anxiety, and personality disorders: Toward a core battery* (pp. 371–388). Washington, DC: American Psychological Association.

Pilkonis, P. A., Heape, C. L., Proietti, J. M., Clark, S. W., McDavid, J. D., & Pitts, T. E. (1995). The reliability and validity of two structured diagnostic interviews for personality disorders. *Archives of General Psychiatry, 52,* 1025–1033.

Reise, S. P., & Waller, N. G. (1993). Traitedness and the assessment of response pattern scalability. *Journal of Personality and Social Psychology, 65,* 143–151.

Retzlaff, P. (1996). MCMI-III diagnostic validity: Bad test or bad validity study. *Journal of Personality Assessment, 66,* 431–437.

Riso, L. P., Klein, D. N., Anderson, R. L., Ouimette, P. C., & Lizardi, H. (1994). Concordance between patients and informants on the Personality Disorder Examination. *American Journal of Psychiatry, 151,* 568–573.

Robins, R. W., & John, O. P. (1997). Effects of visual perspective and narcissism on self-perception. Is seeing believing? *Psychological Science, 8,* 37–42.

Rogers, R. (1995). *Diagnostic and structured interviewing. A handbook for psychologists*. Odessa, FL: Psychological Assessment Resources.

Rogers, R., Bagby, R. M., & Dickens, S. E. (1992). *Structured Interview of Reported Symptoms (SIRS) professional manual.* Odessa, FL: Psychological Assessment Resources.

Sackett, P. R., & Wilk, S. L. (1994). Within-group norming and other forms of score adjustment in preemployment testing. *American Psychologist, 49,* 929–954.

Santor, D. A., Ramsay, J. O., & Zuroff, D. C. (1994). Nonparametric item analyses of the Beck Depression Inventory: Evaluating gender item bias and response option weights. *Psychological Assessment, 6,* 255–270.

Salekin, R. T., Rogers, R., & Sewell, K. W. (1996). A review and meta-analysis of the Psychopathy Checklist and Psychopathy Checklist-Revised: Predictive validity of dangerousness. *Clinical Psychology: Science and Practice, 3,* 203–215.

Saucier, G., & Goldberg, L. R. (1996). The language of personality: Lexical perspectives on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality. Theoretical perspectives* (pp. 21–50). New York: Guilford.

Schinka, J. A., Kinder, B. N., & Kremer, T. (1997). Research validity scales for the NEO-PI-R: Development and initial validation. *Journal of Personality Assessment, 68,* 127–138.

Shea M. T. (1995). Interrelationships among categories of personality disorders. In W. J. Livesley (Ed.), *The DSM-IV personality disorders* (pp. 397–406). New York: Guilford.

Skodol, A. E., Oldham, J. M., Rosnick, L., Kellman, H. D., & Hyler, S. E. (1991). Diagnosis of DSM-III-R personality disorders: A comparison of two structured interviews. *International Journal of Methods in Psychiatric Research, 1,* 13–26.

Smith, G. T., & McCarthy, D. M. (1995). Methodological considerations in the refinement of clinical assessment instruments. *Psychological Assessment, 7,* 300–308.

Somwaru, D. P., & Ben-Porath, Y. S. (1995). *Development and reliability of MMPI-2 based personality disorder scales.* Paper presented at the 30th Annual Workshop and Symposium on Recent Developments in Use of the MMPI-2 & MMPI-A. St. Petersburg Beach, FL.

Tellegen, A., & Waller, N. G. (in press). Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In S. R. Briggs & J. M. Cheek (Eds.), *Personality measures: Development and evaluation* (Vol. 1). Greenwich, CT: JAI Press.

Timbrook, R. E., & Graham, J. R. (1994). Ethnic differences on the MMPI-2? *Psychological Assessment, 6,* 212–217.

Trull, T. J., & Goodwin, A. H. (1993). Relationship between mood changes and the report of personality disorder symptoms. *Journal of Personality Assessment, 61,* 99–111.

Trull, T. J., Useda, J. D., Costa, P. T., & McCrae, R. R. (1995). Comparison of the MMPI-2 Personality Psychopathology Five (PSY-5), the NEO-PI, and the NEO-PI-R. *Psychological Assessment, 7,* 508–516.

Trull, T. J., & Widiger, T. A. (1997). *Structured Interview for the Five-Factor Model of Personality professional manual.* Odessa, FL: Psychological Assessment Resources.

Watkins, C. E., Campbell, V. L., Nieberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology: Research and Practice, 26,* 54–60.

Watson, D., Clark, L. A., & Harkness, A. R. (1994). Structure of personality and their relevance to psychopathology. *Journal of Abnormal Psychology, 103,* 18–31.

Weiner, I. B. (1996). Some observations on the validity of the Rorschach inkblot method. *Psychological Assessment, 8,* 206–213.

Westen, D. (1991). Social cognition and object relations. *Psychological Bulletin, 109,* 429–455.

Widiger, T. A. (1993). Personality and depression: Assessment issues. In M. H. Klein, D. J. Kupfer, & M. T. Shea (Eds.), *Personality and depression. A current view* (pp. 77–118). New York: Guilford.

Widiger, T. A., & Costa, P. T. (1994). Personality and personality disorders. *Journal of Abnormal Psychology, 103,* 78–91.

Widiger, T. A., Frances, A. J., & Trull, T. J. (1989). Personality disorders. In R. Craig (Ed.), *Clinical and diagnostic interviewing* (pp. 221–236). Northvale, NJ: Aronson.

Widiger, T. A., Mangine, S., Corbitt, E. M., Ellis, C. G., & Thomas, G. V. (1995). *Personality Disorder Interview-IV. A semistructured interview for the assessment of personality disorders.* Odessa, FL: Psychological Assessment Resources.

Widiger, T. A., & Sanderson, C. J. (1995). Assessing personality disorders. In J. N. Butcher (Ed.), *Clinical personality assessment. Practical approaches* (pp. 380–394). New York: Oxford University Press.

Widiger, T. A., & Trull, T. J. (1997). Assessment of the five factor model of personality. *Journal of Personality Assessment, 68,* 228–250.

Widiger, T. A., Williams, J. B. W., Spitzer, R. L., & Frances, A. J. (1985). The MCMI as a measure of DSM-III. *Journal of Personality Assessment, 49,* 366–378.

Wiggins, J. S. (1966). Substantive dimensions of self-report in the MMPI item pool. *Psychological Monographs, 80,* (22, Whole No. 630).

Wiggins, J. S., & Pincus, A. L. (1992). Personality: Structure and assessment. *Annual Review of Psychology, 43,* 473–504.

Wiggins, J. S., & Trobst, K. K. (1997). Prospects for the assessment of normal and abnormal interpersonal behavior. *Journal of Personality Assessment, 68,* 110–126.

Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1996). The Comprehensive System for the Rorschach: A critical examination. *Psychological Science, 7,* 3–10.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders. Clinical descriptions and diagnostic guidelines.* Geneva, Switzerland: Author.

Zanarini, M. C., Frankenburg, F. R., Sickel, A. E., & Yong, L. (1996). *Diagnostic Interview for DSM-IV Personality Disorders (DIPD-IV).* Boston: McLean Hospital.

Zimmerman, M. (1994). Diagnosing personality disorders. A review of issues and research methods. *Archives of General Psychiatry, 51,* 225–245.

Zimmerman, M., & Coryell, W. H. (1990). Diagnosing personality disorders in the community. A comparison of self-report and interview measures. *Archives of General Psychiatry, 47,* 527–531.

Zimmerman, M., Pfohl, B., Coryell, W., Stangl, D., & Corenthal, C. (1988). Diagnosing personality disorders in depressed patients. A comparison of patient and informant interviews. *Archives of General Psychiatry, 45,* 733–737.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

3.08

Assessment of Psychopathology: Nosology and Etiology

NADER AMIR
University of Georgia, Athens, GA, USA
and
CATHERINE A. FEUER
University of Missouri at St. Louis, MO, USA

3.08.1 INTRODUCTION 170

3.08.2 HISTORY OF CLASSIFICATION 170
  3.08.2.1 Medical Classification 170
  3.08.2.2 Early Nosology of Psychopathology 171

3.08.3 CLASSIFICATION OF MENTAL ILLNESS 171
  3.08.3.1 History of Classification of Mental Illness 171
  3.08.3.2 Current Classification Systems 172
  3.08.3.3 Tools for Clinical Assessment 173

3.08.4 CURRENT ISSUES IN THE CLASSIFICATION OF PSYCHOPATHOLOGY 174
  3.08.4.1 Definition of Psychopathology 174
  3.08.4.2 Comorbidity 175
    3.08.4.2.1 Actual co-occurrence of disorders 175
    3.08.4.2.2 Populations considered 176
    3.08.4.2.3 Range, severity, and base rates of disorders considered 176
    3.08.4.2.4 Assessment methods 176
    3.08.4.2.5 Structure of the classification system 177
  3.08.4.3 Clinical and Research Applications of Classification Systems 178
  3.08.4.4 Organizational Problems in the DSM 179

3.08.5 ALTERNATIVE APPROACHES TO CLASSIFICATION 180
  3.08.5.1 Types of Taxonomies 180
  3.08.5.2 Taxonomic Models 180
    *3.08.5.2.1 Neo-Kraepelinian (*DSM*) model* 180
    3.08.5.2.2 Prototype model 181
    3.08.5.2.3 Dimensional and categorical models 181
  3.08.5.3 The Use of Etiological Factors for Diagnosis 182

3.08.6 CONCLUSIONS 184

3.08.7 REFERENCES 184

3.08.1 INTRODUCTION

The history of clinical psychology, similar to the history of medicine, has been characterized by the quest for knowledge of pathological processes. This knowledge enables practitioners in both fields to treat and prevent disorders that threaten the quality of life. Health practitioners must be able to identify and conceptualize the problem, communicate research and clinical findings to other members of their field, and ideally, to reach a scientific understanding of the disorder. Diagnostic taxonomies have evolved in part to achieve these goals (Millon, 1991). Classification systems have a long history, first in the basic sciences and medicine, and later in psychopathology. Modern psychopathology taxonomies owe much to the work of earlier medical diagnosticians, and their development has often paralleled that of medical systems (Clementz & Iacono, 1990). Both medical and psychopathological classification systems have served the purposes of improving diagnostic reliability, guiding clinical conceptualization and treatment, and facilitating research and scientific advancement (Blashfield, 1991; Clark, Watson, & Reynolds, 1995). Despite the usefulness of classification systems in psychopathology, no system is accepted universally. Furthermore, some have questioned the utility of classification systems as a whole, favoring behavioral descriptions of individual presentations instead (Kanfer & Saslow, 1965; Ullman & Krasner, 1975). The purpose of this chapter is to review the issues related to the assessment, diagnosis, and classification of psychopathology. The chapter will begin with a brief outline of the history of medical and psychopathological classification. Next, it will review the systems in use in the late 1990s, and evaluate these systems from the standpoint of clinical utility, research facilitation, and scientific understanding. It will then discuss the alternative approaches and future directions in the classification of psychopathology. Finally, the role of assessment strategies in informing this debate will be examined.

3.08.2 HISTORY OF CLASSIFICATION

3.08.2.1 Medical Classification

As Clementz and Iacono (1990) point out, medical classification systems have relied historically on careful observation of patients and their symptoms. The importance of observation as an adjunct to theorizing was recognized as long ago as the seventeenth century by thinkers such as Thomas Sydenham (1624–1689). Observations were used to make inferences about the causes of disease. The practice of medical observation and classification gained momentum with the advent of medical instruments. The most notable of these was the stethoscope, developed in 1819 by Rene Theophile Hyacinthe-Laennec (1787–1826). This advance in instrumentation led to the use of objective signifiers of pathology in diagnosing disease and a diminished interest in less technologically sophisticated observations or subjective reports of symptoms. The continuous development of new technologies increased the ability of physicians to determine the function of organs and their modes of operation when healthy. This novel conceptualization of dysfunction as related to the functioning of organs led the French physician Broussais (1772–1838) to propose the radical idea that specific symptom clusters could not define a disease adequately because of the high overlap between different disorders. He suggested that in order to identify a disease one needs to study the patient's physiological processes and constitution. Mid-nineteenth century advances in the various basic sciences (such as anatomy, histology, and biology) led to the further decline of the practice of observation in favor of the experimental study of physiological disease processes. Although this emphasis on understanding function was a great benefit to the various basic sciences, it culminated in the near abandonment of the practice of clinical observation (Clementz & Iacono, 1990). The increasing popularity of laboratory research also resulted in an emphasis on specific symptoms, rather than on the phenomenology of disease. This topographic or symptom-based approach to the study of pathology lent itself to a descriptive system of classification. However, this approach had its critics, as far back as the 1800s. Some investigators (e.g., Trousseau, 1801–1867) believed that a more comprehensive approach to diagnosis, incorporating both clinical observation and laboratory findings, would better allow the recognition of disorders and their treatment.

The field of genetics provided another major development in the science of classification. Mendel, working in an isolated monastery, pioneered the science of genetics. His efforts were continued by others (e.g., Garrod, 1902) who applied his work to humans. Watson and Crick's (1953a, 1953b) detailed description of the structure of human genetic material supplied yet another powerful tool for identifying the presence and etiology of certain disorders. A classification system based on genetic and biological etiology may seem promising. However, researchers soon realized that the specificity of genetics is less than perfect. This lack of specificity seemed particularly evident in the psychiatric disorders. We now know several psychiatric disorders are at least partly genetic. This is revealed by the finding that individuals with identical genetic make-up have higher concordance rates for developing the same disease than those who do not. However, because the concordance rate between identical twins is less than 100% (e.g., schizophrenia; Dworkin; Lenzenwenger, Moldin & Cornblatt, 1987), not all the variance in psychiatric disorders can be explained by genetic inheritance. These findings have led to the view that genetic make-up provides a diathesis for a particular disorder, which, when interacting with the environmental factors, may produce a disease.

3.08.2.2 Early Nosology of Psychopathology

Emil Kraepelin (1856–1926) was the first to adapt a medical classification framework for psychiatric illnesses (Blashfield, 1991; Kraepelin, 1971). He considered psychopathology to be the result of underlying disease states and advocated the scientific investigation of illnesses. He coined the term "dementia praecox" (later known as schizophrenia) to describe what had previously been considered distinct pathologies. Kraepelin's thinking regarding psychopathology and classification was shaped by the rapid advances in German medical science in the nineteenth century and by his training in behaviorism from Wilhelm Wundt (Berrios & Hauser, 1988). German medicine during Kraepelin's lifetime emphasized the interpretation of mental disorders as diseases of the brain (Menninger, 1963).

Kraepelin's training with Wundt contributed to his use of behavioral descriptions for clusters of symptoms he believed were linked by a common etiology. Kraepelin's psychopathology classification became influential through the publication of his textbooks on psychiatry. His categories later became the basis for the first official classification adopted by the American Psychiatric Association (APA) in 1917 (Menninger, 1963), and revised in 1932 (APA, 1933).

A contemporary of Kraepelin's, Sigmund Freud (1856–1939), was also influential in forming our understanding of psychopathology. In contrast to Kraepelin, Freud placed little emphasis on the phenomenology of disorders, stressing the importance of diagnoses based on the underlying cause of patients' manifest symptoms. Kraepelin, however, opposed Freudian theory and psychoanalytic practice because of its nonempirical orientation (Kahn, 1959). Freud's theories were well articulated, however, and by the 1950s had come to dominate American psychiatry (Blashfield, 1991). Psychoanalytic thought did not lose its foothold until the community-based mental health movement of the 1960s, which pointed out that psychoanalysis was only accessible to the wealthy. Other factors that contributed to the decline of psychoanalytic thought include the development of alternative interventions such as psychotropic medications and behavioral intervention, as well as a general shift toward empiricism. Around this time, there was a resurgence of Kraepelinian thought, mainly by a group of psychiatrists at Washington University in St. Louis (Robins & Guze, 1970). They believed that psychiatry had drifted too far from its roots in medicine, and that it was necessary to identify the biological bases of psychopathology. They advocated an emphasis on classification as a tool for helping psychopathology to evolve as a field (Blashfield, 1991; Klerman, 1978).

3.08.3 CLASSIFICATION OF MENTAL ILLNESS

3.08.3.1 History of Classification of Mental Illness

After World War II, the armed forces developed a more comprehensive nomenclature to facilitate the treatment of World War II servicemen. During this same period, the World Health Organization (WHO) published the first version of the *International classification of diseases* (*ICD-6*; WHO, 1992) that included psychiatric disorders. The authors of the *ICD-6* relied heavily on the work of various branches of the armed forces in the development of their taxonomy of psychopathology. The first version of the *Diagnostic and statistical manual of mental disorders* (*DSM*), a variant on the *ICD-6*, was published in 1952 by the APA's Committee on Nomenclature and Statistics. This nomenclature was designed for use in the civilian population, and focused on clinical utility (APA, 1994). The *DSM* (APA, 1994) was originally created as a means of standardizing the collection of statistics in psychiatric hospitals in the early 1900s. By the release of the second edition of the *DSM* (*DSM-II*) (1975), the authors had moved toward the elimination of exclusion rules and provision of more explicit and descriptive diagnostic criteria. The inclusion of explicit diagnostic criteria in the *DSM-II* was inspired by the 1972 work of a group of Neo-Kraeplinian theorists, Feighner, Baran, Furman, and Shipman at the Washington University School of Medicine in St. Louis (Feighner et al., 1972). This group had originally created a six-category system as a guideline for researchers in need of homogenous subject groups.

In 1978 and 1985, Spitzer and his colleagues modified and expanded the Washington University system and labeled it the Research Diagnostic Criteria (RDC). The RDC established criteria for 25 major categories, with an emphasis on the differential diagnosis between schizophrenia and the affective disorders. The authors of the *DSM-III* (1980) followed the example of the Washington University and RDC groups, including diagnostic criteria for each disorder and exclusion criteria for 60% of all *DSM-III* disorders. Most of these exclusion rules described hierarchical relationships between disorders, in which a diagnosis is not given if its symptoms are part of a more pervasive disorder.

As a result of these changes, the *DSM-III* and *DSM-III-R* are based almost exclusively on descriptive criteria. These criteria are grouped into distinct categories representing different disorders. The *DSM-IV* continues this tradition, using a categorical model in which patients are assigned to classes that denote the existence of a set of related signs and symptoms. The constituent signs and symptoms are thought to reflect an underlying disease construct. Researchers who advocate a categorical model of disease (e.g., Kraepelin) propose that disorders differ qualitatively from each other and from the nondisordered state. Although later editions of the *DSM* have relied progressively more on research findings in their formulation, the stated goal of *DSM-IV* remains the facilitation of clinical practice and communication (Frances et al., 1991). The *DSM* system has evolved as a widely used manual in both clinical and research settings. However, the structure and the stated goals of the *DSM* have been at the core of many controversies regarding the assessment, diagnosis, and classification of psychopathology.

3.08.3.2 Current Classification Systems

Two systems of classification are currently in common use: the *DSM-IV* (APA, 1994) and the ICD-10 (WHO, 1992). A detailed description of these systems is beyond the scope of this chapter and the interested reader should consult recent reviews of these systems (Frances, 1998; Regier, et al., 1998; Spitzer, 1998). A brief overview of each system is provided below.

The *DSM-IV* is a multiaxial classification system. The individual is rated on five separate axes. Axis I includes all diagnoses except personality disorders and mental retardation, the later being rated on Axis II. The rationale for the separation of these axes is to ensure that the presence of long-term disturbances is not overlooked in favor of current pathology. Together, these axes constitute the classification of abnormal behavior. The remaining three axes are not needed to make a diagnosis but contribute to the recognition of factors, other than the individuals' symptoms, that should be considered in determining the person's diagnosis. General medical conditions are rated on Axis III, psychosocial and environmental problems are coded on axis IV, and the individual's current level of functioning is rated on axis V.

The 10th Revision of the *ICD* (*ICD-10*) is a multiaxial system. Axis I includes clinical diagnoses for both mental and physical disorders, Axis II outlines disabilities due to impairments produced by the disorder, and Axis III lists the environmental, circumstantial, and personal lifestyle factors that influence the presentation, course, or outcome disorders. *ICD-10* differs from other multiaxial classifications in that it: (i) records all medical conditions on the same axis (Axis I), (ii) assesses comorbid disorders (Axis II) in specific areas of functioning without ascribing a specific portion to each disorder, and (iii) allows expression of environmental (Axis III) factors determined by clinical practice and epidemiological evidence (Janca, Kastrup, Katschnig, Lopez-Ibor, 1996).

One purpose for the collaboration between *DSM-IV* and *ICD-10* authors was to foster the goal of the transcultural applicability of *ICD-10* (Uestuen, Bertelsen, Dilling, & van Drimmelen, 1996). A multiaxial coding format in both *DSM-IV* and *ICD-10* was adopted in order to provide unambiguous information with maximum clinical usefulness in the greatest number of cases (WHO, 1996). Although not explicitly stated by the authors of either manual, both the *DSM-IV* and *ICD-10* are examples of the "neo-Kraepelinian revolution" in psychiatric diagnostic classification (Compton & Guze, 1995).

The progress toward the shared goals of the two systems include two areas: general clinical use of the systems; and fostering international communication and cultural sensitivity. Regarding general clinical use, a multicenter field trial evaluating aspects of the *ICD-10* and *DSM-IV* involving 45 psychiatrists and psychologists in seven centers was conducted in Germany (Michels et al., 1996). Results revealed moderate inter-rater reliability for the *ICD-10* Axis I and Axis II. However, the number of relevant psychosocial circumstances coded on Axis III by the different raters varied greatly. The authors concluded that the multiaxial system was generally well accepted by participating clinicians, and that it is worth studying empirically and revising accordingly.

Clinicians specializing in several specific areas have not been as positive in their comments about the *ICD-10*. For instance, Jacoby (1996) has argued that neither the *DSM-IV* nor the *ICD-10* is adequate in its categorization of certain disorders associated with the elderly, such as certain dementias and psychotic disorders. Others have argued that specific *ICD-10* diagnoses show poor levels of stability of diagnosis across time. Specifically, a 1997 study found that neurotic, stress-related, adjustment, generalized anxiety, and panic disorders as well as some psychoses and personality disorders showed low rates of reliability across interviews at different time points (Daradkeh, El-Rufaie, Younis, & Ghubash, 1997).

Issues regarding the fostering of international communication and cultural sensitivity were addressed by a Swiss study published in 1995 that assessed the inter-rater reliability and confidence with which diagnoses could be made using the *ICD Diagnostic Criteria for Research* (*ICD-10 DCR*), as well as examining the concordance between *ICD-10 DCR*, *ICD-10 Clinical Descriptions and Diagnostic Guidelines*, and other classification systems including the *DSM-IV*. Field trials were carried out at 151 clinical centers in 32 countries by 942 clinician/researchers who conducted 11 491 individual patient assessments. The authors report that most of the clinician/researchers found the criteria to be explicit and easy to apply and the inter-rater agreement was high for the majority of diagnostic categories. In addition, their results suggested that the use of operational criteria across different systems increases levels of inter-rater agreement.

Regier, Kaelber, Roper, Rae, and Sartorius (1994) cited findings suggesting that while overall inter-rater reliability across 472 clinicians in the US and Canada was good, the clinician tended to make more use of multiple coding for certain disorders than clinicians from other countries. This suggests that certain aspects of the *DSM* system (e.g., its encouragement of multiple diagnoses) may make the transition and agreement between the two systems somewhat more difficult.

The *ICD's* efforts to place psychiatric disorders in the context of the world community's different religions, nationalities, and cultures has received praise in the literature (Haghighat, 1994), while the *DSM-IV* may be seen as somewhat less successful in achieving this goal. The *ICD's* increased cultural sensitivity was at the expense of increased length, but efforts to omit cultural-biased criteria were largely successful. However, other authors have argued that the attempts to improve cultural sensitivity to the degree of a "universal" understanding of mental illness may be misguided (Patel & Winston, 1994). Specifically, while mental illness as a phenomenon may be universal, specific categories of mental illness as outlined by the *DSM* and *ICD* may require identification and validation of particular diagnoses within specific cultures (Patel & Winston, 1994).

3.08.3.3 Tools for Clinical Assessment

A comprehensive review of assessment measures for psychopathology is beyond the scope of this chapter. The interested reader should consult comprehensive reviews of this topic (e.g., Baumann, 1995; Bellack & Hersen, 1998; Sartorius & Janca, 1996). One useful method of classifying assessment techniques is to consider two classes of measures: those that aim to aid in diagnosis and those that aim to assess the severity of symptoms relatively independent of diagnosis. Examples of the first category include The Structured Clinical Interview for *DSM-IV* Diagnoses (SCID; First, Spitzer, Gibbon, & Williams, 1995) and the Composite International Diagnostic Interview (CIDI; Robins et al., 1988). These measures are mapped closely on the diagnostic systems on which they are based (*DSM-IV* and *ICD-10*, respectively). The second class of assessment instruments use a dimensional approach and aim to assess severity of symptoms relatively independently of diagnosis. Examples of such instruments include the Hamilton Depression Inventory (HAM-D; Riskind, Beck, Brown, & Steer, 1987) and the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS; Goodman, Price, Rasmussen, & Mazure, 1989a). The advantages and disadvantages of these assessment techniques are tied closely to advantages and disadvantages of the categorical and dimensional approaches to classification and will be discussed in Section 3.08.5.2.3.

An important aspect of developing new knowledge in psychopathology is the improvement and standardized use of assessment tools across studies of psychopathology. The vast majority of studies of *DSM* disorders have used self-report or interview-based measures of symptoms. In some cases, behavioral checklists (either self- or other-reports) or psychological tests have been employed. In this section, issues such as the accuracy, reliability, and discriminate validity of such assessment tools and how they may influence findings in psychopathology studies will be examined.

Self-report measures may be the most time- and labor-efficient means of gathering data about psychological symptoms and behaviors. However, individuals often are inconsistent in their observations of their own behaviors, and the usefulness of these reports depends heavily on often limited powers of self-observation. Psychometric methods for enhancing the usefulness of self-report data generally focus on increasing the consistency (i.e., reliability and accuracy or validity) of the measures. A number of psychometrically sophisticated scales are available for several different types of disorders (e.g., anxiety; BAI; Beck, Epstein, Brown, & Steer, 1988; PSS; Foa, Riggs, Dancu, & Rothbaum, 1993; Y-BOCS; Goodman et al., 1989a, 1989b). Many of the traditional psychometric approaches classify disorders into state (transitory feelings) and trait (stable personality) attributes (e.g., state anxiety vs. trait anxiety, STAI; Spielberger, Gorsuch, & Lushene, 1970). The accuracy of these scales is most often evaluated by comparing results on the measure to results on other measures of the same emotion or disorder (e.g., interviews, physiological assessments, observations of behavior). By assessing various aspects of an emotion or disorder, the investigator tries to create a composite gauge of an emotion for which no single indicator offers a perfect yardstick.

One impediment to discovering which symptoms may be shared across disorders has been the structure of many clinical measures. Examples of widely-used measures of psychological states are the Beck Depression Inventory (BDI; Beck & Steer, 1987), the Beck Anxiety Inventory (BAI; Beck et al., 1988), the Minnesota Multiphasic Personality Inventory (MMPI-2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989), and structured interview scales such as the Diagnostic Inventory Scale (DIS; Helzer & Robins, 1988) and the Structured Clinical Interview for *DSM-IV* Diagnoses (SCID; First et al., 1995).

These self-report and clinician-rated scales usually assess "modal" symptoms, as they focus on core aspects of each syndrome rather than on all possible variants. Structured interviews, such as the SCID-IV or the DIS, often allow "skip outs" in which the interviewer need not necessarily assess all of the symptoms of all disorders. Many studies have examined the convergent and divergent validity patterns of self-report (e.g., modal) measures of various disorders (e.g., Foa et al., 1993). These measures tend to yield strongly convergent assessments of their respective syndromes, but there is little specificity in their measurement, especially in nonpatient samples. The data often suggest the presence of a large nonspecific component shared between syndromes such as anxiety and depression (Clark & Watson, 1991). Some scales appear to be more highly loaded with the nonspecific distress factor than others. Therefore, in order to efficiently discriminate between disorders, future research should emphasize identification of symptoms which optimally discriminate between diagnoses. Watson and Clark (1992) found that even when factor analytically derived mood measures such as the Profile of Moods (POMS; McNair, Lorr, & Droppleman, 1981) and the Positive and Negative Affectivity Scale (PANAS; Watson, Clark, & Tellegen, 1988) are used, certain basic affects (i.e., anxiety and depression) are only partially differentiable. Their data suggest that the overlap between basic affects represents a shared component inherent in each mood state, which must be measured in order to understand the overlap between different mood states and disorders.

3.08.4 CURRENT ISSUES IN THE CLASSIFICATION OF PSYCHOPATHOLOGY

As noted earlier, the classification of psychopathology has been a controversial topic since inception. Currently, the discussions revolve primarily around the *DSM* system, although many of the debates predate the *DSM*. The most common topics of controversy in the classification of psychopathology are: (i) the definition of psychopathology; (ii) the artificially high rates of comorbidity found between disorders when using the *DSM* system (Frances, Widiger, & Fyer, 1990; Robins, 1994; (iii) the balance between a focus on clinical utility and the facilitation of research and scientific progress (Follette, 1996); and (iv) organizational problems, both within and across disorders in the *DSM* (Brown & Barlow, 1992; Quay, Routh, & Shapiro, 1987).

3.08.4.1 Definition of Psychopathology

There is a lack of agreement about the definition of psychopathology. Early versions of both the *ICD* and the *DSM* attempted to guide the categorization of mental disorders without addressing the definition of psychopathology (Spitzer, Endicott, & Robins, 1978). Some contend that this lack of agreement about the definition of psychopathology remains the current state of affairs (e.g., Bergner, 1997). Others (e.g., Kendell, 1975, 1982) have noted that physicians frequently diagnose and treat disorders for which there is no specific definition, and which, in many cases, are not technically considered disorders (e.g., pregnancy). The controversy over the definition of mental disorder may be fueled by the fact that such illnesses are signified commonly by behaviors, and distinguishing between "normal" and "deviant" or "disordered" behaviors is viewed by some as having serious cultural and sociological implications (Mowrer, 1960; Szasz, 1960). According to Gorenstein (1992), attempts at defining mental illness have fallen historically into several categories. Some have relied on statistical definitions based on the relative frequency of certain characteristics in the general population. Others have employed a social definition in which behaviors which conflict with the current values of society are considered deviant. Still other approaches have focused on the subjective discomfort caused by the problem, as reported by the individual. Finally, some definitions of psychopathology have been based on psychological theories regarding states or behaviors thought to signify problems within the individual. The *DSM-IV* (APA, 1994) defines a mental disorder as a "clinically significant behavioral or psychological syndrome or pattern that occurs in an individual and ... is associated with present distress ... or disability ... or with a significant increased risk of suffering death, pain, disability or an important loss of freedom." The disorder should not be an "expectable and culturally sanctioned response to an event," and must be considered a manifestation of dysfunction in the individual. Wakefield has expanded and refined this idea in his concept of "harmful dysfunction" (Wakefield, 1992, 1997c). This concept is a carefully elaborated idea considered by some to be a workable solution to the problem of defining abnormality (e.g., Spitzer, 1997). In essence, the harmful dysfunction concept means that behaviors are abnormal to the extent that they imply something is not working properly as would be expected based on its evolutionary function. Specifically, Wakefield (1992) states that a mental disorder is present

> if and only if, a) the condition causes some harm or deprivation of benefit to the person as judged by the standards of the person's culture (the value criteria), and b) the condition results from the inability of some mental mechanism to perform its natural function, wherein a natural function is an effect that is part of the evolutionary explanation of the existence and structure of the mental mechanism (the explanatory criterion). (p. 385)

Critics of Wakefield's concept argue that it represents a misapplication of evolutionary theory (Follette & Houts, 1996). They contend that evolutionary selection was not meant to apply on the level of behavioral processes, that is, it is not possible to know the function of a part or process of the individual by examining evolutionary history because random variation precludes a straightforward interpretation of the "design" of various parts of organisms (Mayr, 1981; Tattersall, 1995). Proponents of the "harmful dysfunction" definition suggest that it may provide a useful starting point for research in psychopathology, although further research needs to identify the specific mechanisms that are not functioning properly (Bergner, 1997).

3.08.4.2 Comorbidity

The term comorbidity was first used in the medical epidemiology literature and has been defined as "the co-occurrence of different diseases in the same individual" (Blashfield, 1990; Lilienfeld, Waldman, & Israel, 1994). Many factors potentially may contribute to the comorbidity rates of psychiatric disorders. Reported comorbidity rates are influenced by the actual co-occurrence of disorders, the populations considered, the range, severity, and base rates of the disorders considered, the method of assessment, and the structure of the classification system used.

3.08.4.2.1 Actual co-occurrence of disorders

In the medical literature, comorbidity often refers to diagnoses which occur together in an individual, either over one's lifetime or simultaneously. This concept emphasizes the recognition that different diagnoses potentially are related in several ways. One disease can signal the presence of another, predispose the patient to the development of another, or be etiologically related to another disorder (Lilienfeld et al., 1994). Lilienfeld et al. suggest that the increased attention to comorbidity in psychopathology is due to acknowledgment of the extensive co-occurrence and covariation that exists between diagnostic categories (Lilienfeld et al. ; Widiger & Frances, 1985). For example, Kessler et al. (1994) reported the results of a study on the lifetime and 12-month prevalence of *DSM-III-R* disorders in the US, in a random sample of 8098 adults. Forty-eight percent of the subjects reported at least one lifetime disorder, with the vast majority of these individuals, 79%, reporting comorbid disorders. Other studies using large community samples report that more than 50% of the participants diagnosed with one *DSM* disorder also meet criteria for a second disorder (Brown & Barlow, 1992). It has been argued that the common etiological factors across different diagnoses are of greater importance than the etiological factors specific to one disorder (Andrews, 1991; Andrews, Stewart, Morris-Yates, Holt, & Henderson, 1990).

3.08.4.2.2 Populations considered

One factor that affects comorbidity rates is the population under study. Kessler et al. (1994) found that people with comorbid disorders were likely to report higher utilization of services than those without a comorbid disorder. Similarly, higher rates of comorbidity were found among urban populations than rural populations, despite the higher rates of single psychiatric disorders and financial problems among rural participants. The race of the population studied also seems to influence comorbidity rates. African-American participants in the Kessler et al. study reported lower comorbidity rates than Caucasian participants, after controlling for the effects of income and education. However, Caucasian participants reported lower rates of current comorbid disorders than Hispanics. Disparities in prevalence of reported comorbidity were also found between respondents of different age groups, with people aged between 25 and 34 years reporting the highest rates. The income and education of the participants were also associated with reported comorbidity in the Kessler et al. study.

3.08.4.2.3 Range, severity, and base rates of disorders considered

The rates of comorbidity are also influenced by the disorders studied. Specifically, the range, severity, and base rates of certain disorders increase the likelihood that these disorders will be comorbid with another disorder. Certain diagnostic categories, including childhood disorders (Abikoff & Klein, 1992; Biederman, Newcorn, & Sprich, 1991), anxiety disorders (Brown & Barlow, 1992; Goldenberg et al., 1996), and personality disorders (Oldham et al., 1992; Widiger & Rogers, 1989), appear to be comorbid with other disorders. For example, 50% of patients with a principal anxiety disorder reported at least one additional clinically significant anxiety or depressive disorder in a large-scale study by Moras, DiNardo, Brown, and Barlow (1991). Similarly, anxiety disorders are not only highly likely to be comorbid with each other but also with mood, substance use, and personality disorders (Brown & Barlow, 1992). These high rates of comorbidity may in part be due to the degree to which the different anxiety disorders include overlapping features (Brown & Barlow, 1992). Finally the degree of comorbidity is influenced directly by thresholds set to determine the presence or absence of various disorders. The choice of threshold appears to affect comorbidity rates differentially, depending on the disorder. For instance, certain disorders (e.g., social phobia) appear to be more likely to accompany other anxiety disorders when considered at subclinical levels (Rapee, Sanderson, & Barlow, 1988). Conversely, Frances et al. (1990) have suggested that the severity of mental health problems in a sample will influence comorbidity rates, in that a patient with a severe presentation of one disorder is more likely to report other comorbid disorders.

In addition to the range and severity thresholds of disorders, the base rates of a particular disorder have a strong influence on the apparent comorbidity rates. Conditions that are frequently present in a given sample will tend to be diagnosed together more often than those that are infrequent. This may help explain the comorbidity rates of certain highly prevalent disorders such as anxiety, depression, and substance abuse.

3.08.4.2.4 Assessment methods

The choice of assessment methods may influence comorbidity rates in much the same way as aspects of the disorders. Disagreement regarding comorbidity rates may be partly due to differences in the definition of comorbidity. For example, comorbidity has been defined as within-episode co-occurrence (or dual diagnosis) among disorders by some (August & Garfinkel, 1990; Fulop, Strain, Vita, Lyons, & Hammer, 1987), lifetime co-occurrence by others (Feinstein, 1970; Shea, Widiger, & Klein, 1992, p. 859), and covariation among diagnoses (i.e., across individuals) by still other researchers (Cole & Carpentieri, 1990; Lewinsohn, Rohde, Seeley, & Hops, 1991). Even when researchers agree on a definition, their estimates of comorbidity may differ based on the type of assessment tool they use. As Lilienfeld et al. (1994) point out, assessment techniques have error variance that by definition is not related to the construct(s) of interest or may artificially inflate the actual comorbidity rate.

Furthermore, individual raters may hold biases or differing endorsement thresholds for behaviors which are common to a variety of disorders. Likewise, raters may have specific beliefs about which disorders tend to covary. These types of biases may affect both self-report and interviewer-based data. Similarly, studies utilizing structured interviews may differ from studies in which severity thresholds are not described as concretely. Therefore, differing rates of comorbidity across studies may be an artifact of the types of measurement used, or the biases of raters involved (Zimmerman, Pfohl, Coryell, Corenthal, & Stangl, 1991).

3.08.4.2.5 Structure of the classification system

Frances et al. (1990) argue that the classification system used in the various versions of the *DSM* increases the chance of comorbidity in comparison to other, equally plausible systems. The early systems (e.g., *DSM-III*) attempted to address the comorbidity issue by proposing elaborate hierarchical exclusionary rules specifying that if a disorder was present in the course of another disorder that took precedence, the second disorder was not diagnosed (Boyd et al., 1984). Thus, disorders which may have been truly comorbid were overlooked. Examination of these issues resulted in the elimination of this exclusionary criterion in *DSM-III-R*. This new method of diagnosis, however, has been criticized for artificially inflating comorbidity between various disorders.

Frances et al. (1990) point out that new editions of the *DSM* expanded coverage by adding new diagnoses. This was done by "splitting" similar disorders into subtypes. They argue that the tendency to split diagnoses creates much more comorbidity than would "lumping" systems of classification. This is because disorders that share basic underlying features are viewed as separate.

Much of the "splitting" in the *DSM* has resulted from the increasing reliance on prototypical models of disorders (Blashfield, 1991). The creators of the *DSM* increasingly have relied on prototypes, defining a diagnostic category by its most essential features, regardless of whether these features are also present in other diagnoses. McReynolds (1989) argued that categories with a representative prototype and indistinct or "fuzzy" boundaries are the basis of the most utilitarian classification systems because they are representative of categories in nature. The use of this type of prototype-based classification has improved the sensitivity and clinical utility of the *DSM* system. However, these gains are achieved at the expense of increased comorbidity and decreased specificity of differential diagnosis due to definitional overlap. Thus, this system of diagnostic classification makes it unclear whether there is any true underlying affinity between disorders currently considered to have high rates of comorbidity.

The stated approach of the *DSM-IV* is a descriptive method relying on observable signs and symptoms rather than underlying mechanisms. The sets of signs and symptoms that constitute the different diagnoses are, by definition, categorical, that is, the presence of the required number and combination of symptoms indicates a diagnosis and the absence of the required number of symptoms precludes a diagnosis. The symptoms are not weighted, implying that they are of equal importance in defining the disorder. For many diagnoses, the structure of this system makes it possible for two individuals to meet the same diagnostic criteria without sharing many symptoms. Conversely, it is possible for one person to meet the criteria while another person who shares all but one feature does not meet the criteria. As a result, patients who actually form a fairly heterogeneous group may be "lumped" into one homogeneous diagnostic category.

Combined with unweighted symptoms and a lack of attention to severity of symptoms, this "lumping" can lead to what Wakefield (1997a) refers to as "overinclusiveness" of diagnostic criteria. Specifically, people who do not truly suffer from a mental disorder may nonetheless receive the diagnosis, thus lowering the conceptual validity of *DSM* diagnostic categories. Conversely, minute differences in reported symptoms may result in dichotomizing between disordered and nondisordered individuals, or between individuals with one disorder as opposed to another (e.g., avoidant personality disorder vs. generalized social phobia), creating heterogeneity where there may actually be none.

This point is further elaborated by Clark et al. (1995), who point out that within-category heterogeneity constitutes a serious challenge to the validity of the *DSM*. These authors view comorbidity and heterogeneity as related problems, that is, within-group heterogeneity of symptoms found across diagnostic categories leads to increased rates of comorbidity among the disorders. Homogenous categories, on the other hand, lead to patient groups that share defining symptoms and produce lower rates of both comorbidity and heterogeneity. This is in part because of polythetic systems that require some features of a disorder for diagnosis. In contrast, a monothetical system would require all features of a disorder for diagnosis. Because the monothetic approach to diagnosis produces very low base rates for any diagnosis (Morey, 1988), researchers generally have used the polythetic approach. However, this approach promotes within-category heterogeneity because individuals who do not share the same symptom profiles can receive the same diagnosis. This poses a problem for any categorical system, including the *DSM*. Because of the inherent heterogeneity in patient profiles, a categorical system must address this issue by either proposing artificial methods of limiting heterogeneity (e.g., *DSM-III*), using unrealistic homogeneous categories, or acknowledging the heterogeneity (e.g., *DSM-IV*; Clark et al., 1995). Schizophrenia is an example of a diagnosis that has proved to have within-group heterogeneity. For example, current nosology (e.g., *DSM-IV*) assumes that schizophrenia and affective illness are distinct disorders. However, shared genetic vulnerability has been proposed for schizophrenia and some affective disorders (Crow, 1994). Taylor (1992) reviewed the evidence for this continuum and suggested that the discrimination of these disorders by their signs and symptoms is inadequate (Taylor & Amir, 1994). The need to know whether psychoses vary along a continuum is obviously critical to our understanding of their pathogenesis and etiology.

To address within-group heterogeneity of disorders, researchers have attempted to create subtypes (e.g., positive vs. negative symptoms of schizophrenia; Andreasen, Flaum, Swayze, Tyrell, & Arndt, 1990). Consistent with this conceptualization, researchers have correlated negative symptoms of schizophrenia with cognitive deficits (Johnstone et al., 1978), poor premorbid adjustment (Pogue-Geile & Harrow, 1984), neuropathologic and neurologic abnormalities (Stevens, 1982), poor response to neuroleptics (Angrist, Rotrosen, & Gershon, 1980), and genetic factors (Dworkin & Lenzenweger, 1984). On the other hand, positive symptoms have been correlated with attention (Cornblatt, Lenzenweger, Dworkin, Erlenmeyer-Kimling, 1992) and a positive response to neuroleptics (Angrist et al., 1980). Despite these differences, some investigators have questioned the reliability of these findings (Carpenter, Heinrichs, & Alphs, 1985; Kay, 1990) and the clinical clarity of the subtypes (Andreasen et al., 1990). The majority of these studies included patients with schizophrenia but not affective disorders, and some assessed only limited symptoms (e.g., only negative symptoms; Buchanan & Carpenter, 1994; Kay, 1990). These approaches are potentially problematic because they assume implicitly that the psychoses are distinct, and that the specific types of psychopathology are easily discriminated.

In summary, the issue of comorbidity poses what is potentially the greatest challenge to diagnosis. Factors including the populations considered, the range, severity, and base rates of the disorders considered, the method of assessment, and the structure of the classification system contribute to reported comorbidity rates. These factors are often extremely difficult to distinguish from the true comorbidity of disorders. A full understanding of the common factors and shared etiologies of psychiatric disorders will most likely require an understanding of their true comorbidity. It is likely, then, that this will remain a highly controversial topic in psychopathology assessment.

3.08.4.3 Clinical and Research Applications of Classification Systems

Classification serves various purposes, depending on the setting in which it is used. In clinical settings, classification is used for treatment formulation, whereas in research settings it allows researchers to formulate and communicate ideas. The *DSM-IV* taskforce has stated that the various uses of classification are usually compatible (APA, 1994, p. xv). It is not clear whether this assertion has been tested empirically, or whether steps have been taken to resolve any divergence between the various uses of classification. The goals of clinical utility and outcome of research are not inherently inconsistent, but situations may arise in which the two diverge in their application. The majority of the modifications made to recent versions of the *DSM* were designed to improve clinical utility by simplifying or improving everyday diagnostic procedures. While not necessarily empirically driven, these changes do not appear to have adversely impacted the validity of the diagnoses they affect. Assigning diagnoses facilitates information storage and retrieval for both researchers and clinicians (Blashfield & Draguns; 1976; Mayr, 1981). However, problematic issues have been raised about the use and structure of diagnostic systems in both clinical and research settings.

The use of classification and the reliance on the *DSM* appears to be increasing (Follette, 1996). This is partly because of the trend toward the development and empirical examination of treatments geared toward specific diagnoses (Wilson, 1996). Although the use of diagnostic classification is beneficial in conceptualizing cases, formulating treatment plans, and communicating with other providers, some have argued that assigning a diagnosis may have a negative impact on patients (Szasz, 1960), that is, decisions about what constitutes pathology as opposed to normal reactions to stressors may be arbitrary (Frances, First, & Pincus, 1995), gender-based, and culturally or socioeconomically based. Furthermore, the choice of which behaviors are considered pathological (de Fundia, Draguns, & Phillips, 1971) may be influenced by the characteristics of the client (Broverman, Broverman, Clarkson, Rusenkrantz, & Vogel, 1970; Gove, 1980; Hamilton, Rothbart, & Dawes, 1986).

Diagnostic categories may also have adverse consequences on research. For example, although the diagnostic categories defined by the classifications systems are well studied and easily compared across investigations, patients who do not clearly fit a diagnosis are ignored. Furthermore, frequent and generally clinically-driven changes in the diagnostic criteria for a particular disorder make sustained investigation of the disorder difficult (Davidson & Foa, 1991). This situation is further exacerbated by the frequent addition of new diagnostic categories to the *DSM* (Blashfield, 1991). The *DSM-IV* taskforce addressed this issue by explicitly recommending that the diagnostic classifications be used as a "guideline informed by clinical judgment," as they "are not meant to be used in a cookbook fashion" (APA, 1994, p. xxiii). This suggestion, while useful in a clinical setting, may not be as applicable in research settings.

The proliferation of diagnostic criteria in recent versions of the *DSM* have successfully improved diagnostic reliability, but have done so at the expense of external validity (Clementz & Iacono, 1990). That may adversely affect research findings (Carey & Gottesman, 1978; Meehl, 1986). According to Clementz and Iacono (1990), the achievement of high reliability in a diagnostic classification is often taken to signify either validity or a necessary first step toward demonstrating validity. However, this is not the case in situations in which disorder criteria are designed significantly to increase reliability (i.e., by using a very restrictive definition). In such situations, validity may actually suffer greatly due to the increase in false categorizations of truly disordered people as unaffected. Conversely, the generation of specific behavioral criteria (e.g., Antisocial Personality Disorder in *DSM-III-R*) may increase reliability but lead to overinclusion (e.g., criminals in the antisocial category) while entirely missing other groups (e.g., "successful" psychopaths who have avoided legal repercussions of the behavior) (Lykken, 1984).

Clearly there are possible negative impacts of our current diagnostic system for research. More broadly, some have argued that nosological questions involve "value judgments" (Kendler, 1990, p. 971) and as such are nonempirical. But, as Follette and Houts (1996) point out, the classification of pathology, even in physical medicine, requires the identification of desirable endstates which are culturally rather than biologically defined. Furthermore, value judgments lie in the choices made between competing theories (Widiger & Trull, 1993) and the attempt by the authors of the *DSM* to remain theoretically "neutral" is inherently problematic. In its attempt to remain "neutral," the DSM actually adopts a model of shared phenomenology, which implicitly accepts a theory of the underlying structure of psychopathology (Follette & Houts, 1996). This implicit theory is problematic on several related fronts. First, the *DSM* professes to be theory-neutral while it may arguably be considered theory-bound. Second, because the authors of the *DSM* do not explicitly recognize any theory, specific theories cannot be empirically tested against competing theories of psychopathology.

3.08.4.4 Organizational Problems in the *DSM*

Several of the issues regarding assessment and classification revolve around the organization of the *DSM*. These problems can be classified into two broad categories: problems with the multiaxial system and problems concerning the placement and utilization of specific categories of disorders (Clark et al., 1995).

The multiaxial system of the *DSM* has no doubt accomplished its goal of increasing the attention given to various (e.g., personality) aspects of functioning. However, the distinction between certain personality (Axis II) disorders and some Axis I disorders are problematic. For example, there is now ample data suggesting that there is a high degree of overlap between avoidant personality disorder and generalized social phobia (Herbert, Hope, & Bellack, 1992; Widiger, 1992). Similarly, it is difficult to distinguish schizotypal personality disorder and schizophrenia on empirical grounds (Grove et al., 1991).

The second type of organizational issue relates to the placement and utilization of individual disorders. For example, because post-traumatic stress disorder (PTSD) shares many features (e.g., depersonalization, detachment) with the dissociative disorders, some (e.g., Davidson & Foa, 1991) have argued for its placement with the dissociative disorders instead of the anxiety disorders. Likewise, the overly exclusive diagnostic categories of the *DSM* have led to situations in which clearly disordered patients exhibiting symptoms of several disorders do not fit into any specific category. The authors of the *DSM* have attempted to remedy this by creating "not otherwise specified" (NOS) categories for several classes of disorders. While at times it is difficult to distinguish the meaning of NOS categories, they are often utilized in clinical practice. Clark et al (1995) view the high prevalence of subthreshold and atypical forms of the disorders commonly classified as NOS as contributing to the problem of heterogeneity. For example, various diagnoses including mood disorders (Mezzich, Fabrega, Coffman, & Haley, 1989), dissociative disorders (Spiegel & Cardena, 1991), and personality disorders (Morey, 1992) are characterized by high rates of NOS. One method of combating the high NOS rates would be to create separate categories to accommodate these individuals. For example, the rate of bipolar disorder NOS was reduced by the inclusion of the subcategories I and II in *DSM-IV*. Another method addressing the NOS problems is to include clear examples of individuals who would meet the NOS criteria. This second approach has been tested for some potential diagnostic groups, such as the mixed anxiety-depression category (Clark & Watson, 1991; Zinbarg & Barlow, 1991; Zinbarg et al., 1994).

3.08.5 ALTERNATIVE APPROACHES TO CLASSIFICATION

While many well-articulated criticisms of the classification system have been presented, the literature contains far fewer suggested alternatives (Follette, 1996). There is great disagreement among authors who have suggested other taxonomic models about not only the type of model, but also its content, structure, and methodology. The study of psychopathology has been pursued by researchers from different theoretical schools, including behaviorists, neurochemists, phenomenologists, and psychodynamicists (Millon, 1991). These approaches rely on unique methodologies and produce data regarding different aspects of psychopathology. No one conceptualization encompasses the complexity of any given disorder. These differing views are not reducible to a hierarchy, and cannot be compared in terms of their objective value (Millon, 1991). Biophysical, phenomenological, ecological, developmental, genetic, and behavioral observations have all been suggested as important contents to include in the categorization of psychopathology. How to structure or organize content has also been a topic of much debate. In this section the structural and methodological approaches to creating classification systems will be reviewed. Models of psychopathology will be discussed, including the Neo-Kraepelinian model, prototype models, and dimensional and categorical models. Finally, the use of etiological factors in the categorization of psychopathology (J. F. Kihlstrom, personal communication, September 2, 1997, SSCPnet) will be examined and other research and statistical methodologies will be outlined that may potentially lead to better systems of classification (J. F. Kihlstrom, personal communication, September 11, 1997, SSCPnet; D. Klein, personal communication, September 2, 1997, SSCPnet).

3.08.5.1 Types of Taxonomies

As knowledge about features best signifying various disorders increases, determinations must be made about the structure that best suits their classification. Before undertaking an investigation of the specific psychopathology classifications which have been proposed, the structural design of diagnostic taxonomies in general will be outlined briefly. The frameworks suggested fall into three categories: hierarchical, multiaxial, and circular. The hierarchical model organizes disorders into sets with higher-order diagnoses subsuming lower-order classifications. The multiaxial model assigns parallel roles for the different aspects of a disorder. Finally, the circular model assigns similar disorders to adjoining segments of a circle and dissimilar disorders to opposite sides of the circle (Millon, 1991). These three conceptualizations are not mutually exclusive, and many taxonomies share aspects of more than one structure. Within each of these three general frameworks, taxa may be considered categorical or dimensional. The current *DSM* combines both hierarchical and multiaxial approaches. Circular frameworks, which generally employ the dimensional approach, have been used in theories of personality. The model that is the basis for the *DSM* system, the Neo-Kraepelian model will be examined first, and then prototype and dimensional models will be considered Finally, suggested methodological and statistical approaches to improving the classification of psychopathology will be covered.

3.08.5.2 Taxonomic Models

3.08.5.2.1 Neo-Kraepelinian (DSM) *model*

The Neo-Kraepelinian model, inspired by the 1972 work of the Washington University group and embodied in the recent versions of the *DSM*, is the current standard of psychopathology classification. According to the Neo-Kraepelinian view, diagnostic categories represent medical diseases, and each diagnosis is considered to be a discrete entity. Each of the discrete diagnostic categories is viewed as having a describable etiology, course, and pattern of occurrence. Clearly specified operational criteria are used to define each category and foster objectivity. This type of classification is aided by the use of structured interviews in gathering relevant symptom information to assign diagnoses. Diagnostic algorithms specify the objective rules for combining symptoms and reaching a diagnosis (Blashfield, 1991). In this view, the establishment of the reliability of diagnostic categories is considered necessary before any type of validity can be established. Categories that refer to clearly described patterns of symptoms are considered to have good internal validity, while the utility of a diagnosis in predicting the course and treatment response of the disorder are seen as markers of good external validity. Despite its many shortcomings, the widespread adoption of the *DSM* system in both clinical work and research is a testament to the utility of the Neo-Kraepelinian model.

3.08.5.2.2 Prototype model

The prototype model has been suggested as a viable alternative to the current Neo-Kraepelinian approach (Cantor, Smith, French, & Mezzich, 1980; Clarkin, Widiger, Frances, Hurt, & Gilmore, 1983; Horowitz, Post, French, Wallis, & Siegelman, 1981; Livesley, 1985a, 1985b). In this system, patients' symptoms are evaluated in terms of their correlation with a standard or prototypical representation of specific disorders. A prototype consists of the most common features of members of a category, and is the standard against which patients are evaluated (Horowitz et al., 1981).

The prototype model differs from the Neo-Kraepelinian model in several ways. First, the prototype model is based on a philosophy of nominalism, in which diagnostic categories represent concepts used by mental health professionals (Blashfield, 1991). Diagnostic groups are not viewed as discrete, but individuals may warrant membership in a category to a greater or lesser degree. The categories are defined by exemplars, or prototypes, and the presentation of features or symptoms in an individual is neither necessary nor sufficient to determine membership in a category. Rather, the prototype model holds that membership in a category is correlated with the number of representative symptoms the patient has. The prototype model suggests that the degree of membership to a category is correlated with the number of features that a member has, so defining features are neither necessary nor sufficient.

Some authors have described the *DSM* system as a prototype model, primarily because it uses polythetic, as opposed to monothetic, definitions (Clarkin et al., 1983; Widiger & Frances, 1985). Although the *DSM* does use polythetic definitions, it does not constitute a prototypical model because specific subsets of symptoms are sufficient for making a diagnosis. Prototype and polythetic models allow variability among features within a category, however, they view category definition differently. Prototype models focus on definition by example, polythetic models focus on category membership as achieved by the presence of certain features that are sufficient. In a prototype model the level of membership is correlated with number of features present, and features are neither necessary nor sufficient since membership is not an absolute. Furthermore, categories in the prototype model have indistinct boundaries, and the membership decision relies largely on clinician judgment. It is likely that the adoption of this model would result in a decrease in reliability compared to the *DSM*. However, proponents argue that this model is more reflective of real-world categories in psychopathology (Chapman & Chapman, 1969).

3.08.5.2.3 Dimensional and categorical models

An alternative to the categorical classification system is the dimensional approach. In dimensional models of psychopathology, symptoms are assessed along several continua, according to their severity. Several dimensional models have been suggested (Eysenck, 1970; Russell, 1980; Tellegen, 1985). Dimensional models are proposed as a means of increasing the empirical parsimony of the diagnostic system. The personality disorders may be most readily adapted to this approach (McReynolds, 1989; Widiger, Trull, Hurt, Clarkin, & Frances, 1987; Wiggins, 1982) but this approach is by no means limited to personality disorders and has been suggested for use in disorders including schizophrenia (Andreasen and Carpenter, 1993), somatoform disorder (Katon et al., 1991), bipolar disorder (Blacker & Tsuang, 1992), childhood disorders (Quay et al., 1987), and obsessive-compulsive disorder (Hollander, 1993). Dimensional models are more agnostic (i.e., making fewer assumptions), more parsimonious (i.e., possibly reducing the approximately 300 diagnosis classifications in the *DSM* to a smaller subset of dimensions), more sensitive to differences in the severity of disorders across individuals, and less restrictive.

While a dimensional approach might simplify some aspects of the diagnostic process, it would undoubtedly create new problems. First, categorical models are resilient because of the psychological tendency to change dimensional concepts into categorical ones (Cantor & Genero, 1986; Rosch, 1978). Second, implementation of dimensional systems would require a major overhaul of current practice in the mental health field. Third, replacing the *DSM* categorical model with a dimensional model will undoubtedly meet with much resistance from proponents of clinical descriptiveness, who believe that each separate category provides a more richly textured picture of the disorder. Finally, there are currently no agreed upon dimensions to be included in such a classification model (Millon, 1991).

Thus, the task of advocates of the dimensional approach is twofold. First, they must determine the type and number of dimensions that are necessary to describe psychopathology. Second, they will need to demonstrate that it is possible to describe the entire range of psychopathology using a single set of dimensions. At this stage, the most useful starting point may be examination of the role of various dimensions in the description of psychopathology, as opposed to arguing the virtues and limitations of categorical and dimensional approaches to psychopathology.

3.08.5.3 The Use of Etiological Factors for Diagnosis

Another proposed remedy to the problems facing current classification systems is to examine the role of etiology. Andreasen and Carpenter (1993) point out the need to identify etiologic mechanisms in order to understand a disorder. In addition, understanding etiologic factors may help explain the heterogeneity and comorbidity of the disorders currently listed in the *DSM*. The authors of the *DSM* have generally avoided making statements regarding the etiology of disorders in keeping with the "theoretically neutral" stance. However, some authors have argued that this caveat is only loosely enforced in *DSM* as it is, as exemplified by life-stress based disorders such as PTSD and adjustment disorder (Brett, 1993).

Traditional models of etiology have focused on either the biological or environmental causes of psychopathology. Wakefield (1997b) has warned against the assumption that the etiology of psychopathological disorders will necessarily be found to be a physiological malfunction. He argued that the mind can begin to function abnormally without a corresponding brain disorder. More recent conceptualizations of the etiology of psychopathology acknowledge the role of both biological and environmental factors, and debate the ratio of the contribution from each. As would be expected, etiological studies of psychopathology tend to reflect the underlying theories of mental disorders accepted by the researchers who conduct them. For example, biological theorists attempt to identify biological markers of certain disorders (e.g., Clementz & Iacono, 1990; Klein, 1989). Environmental theories attempt to identify specific events that are necessary or sufficient to produce a disorder. These approaches have achieved varying degrees of success depending on which diagnostic groups were considered. One explanation for the limited success of attempts to identify etiological factors in psychopathology is the limited number of techniques available for the examination of potential factors. For example, the history of psychiatry and medicine is replete with examples of major findings due in part to advances in technology (e.g., computerized axial tomography [CAT] scans as a method of examining the function of internal organs). An alternative explanation for the limited success of etiological studies is that most researchers have relied on theoretical perspectives that assume distinct categories of illness. Specifically, the assumption of distinct diagnostic entities masks the possibility that multiple etiological factors may lead to the development of the same disorder, and that biological and environmental factors may ameliorate the effect of strong etiological factors. Even when a diagnosis may seem to have a clear etiology (e.g., PTSD), the picture is not clear. For example, although the diagnosis of PTSD requires a clear stressor it is known that not all individuals who experience that stressor develop PTSD and not all stressors are likely to create PTSD (Foa & Rothbaum, 1998). Furthermore, the presence of a stressor alone is not sufficient to warrant the diagnosis. In fact, research suggests that etiologic factors entirely outside the diagnostic criteria (e.g., IQ; Macklin et al., 1998; McNally & Shin, 1995) may ameliorate the effects of the identified etiologic factors on the development of the disorder.

Much of the controversy about assessment and classification in psychopathology stems from the conflict about the use of value judgments as opposed to data-driven theory testing in creating diagnostic categories. Some have suggested that a combination of the two perspectives, including the use of both theory and data, may be the most workable approach (Blashfield & Livesley, 1991; Morey, 1991). This approach would amount to a process of construct validation depending on both theory and evaluation of the theory by data analysis. As in other areas of theory development, testability and parsimony would play a crucial role in choosing between competing theories. In the following section, the need for the adoption of new research methodologies in the field of assessment and classification of psychopathology will be considered. Next some of the areas of research which illustrate promising methods, most of which focus on identifying etiologic factors in psychopathology, will be discussed.

As mentioned earlier, researchers have called for a move away from a system of diagnosis based on superficial features (symptoms) toward diagnosis based on scientifically-based theories of underlying etiology and disease processes. This focus parallels that of medical diagnosis of physical illness. Physicians do not diagnose based on symptoms. Instead, patient reports of symptoms are seen as indicators of potential illnesses. Diagnoses are not made until specific indicators of pathology (e.g., biopsies, masses, blood draws, etc.) have been examined. The interpretation of such laboratory results requires an understanding of the differences between normal and abnormal functioning of the cell or organ in question. In order for assessment of psychopathology to follow this route, researchers must continue to amass information on normal mental and behavioral functioning (J. F. Kihlstrom, personal communication, September 11, 1997, SSCPnet). This endeavor can be facilitated by technological advances in experimental paradigms and measurement techniques and devices. The issue of what constitutes a great enough deviation from "normal" functioning to warrant treatment has been and most likely will continue to be controversial. However, such decisions are a necessary hurdle if psychopathology is to evolve as a science.

It is likely that the identification of etiological factors in psychopathology will not rely entirely on biological factors. The validation of etiological constructs in psychopathology will undoubtedly include studies designed to identify potential contributing causes including environmental, personality, and physiological factors. Examples of research methods and paradigms which may prove useful in determining the etiology of psychiatric disorders are becoming increasingly evident in the literature. Possible methodologies include: psychopharmacological efficacy studies (Harrison et al., 1984; Hudson & Pope, 1990; Papp et al., 1993; Quitkin et al., 1993; Stein, Hollander, & Klein, 1994); family and DNA studies (Fyer et al., 1996); treatment response studies (Clark et al., 1995, Millon, 1991); *in vivo* monitoring of physiological processes (Martinez et al., 1996); and identification of abnormal physiological processes (Klein, 1993, 1994; Pine et al., 1994). These approaches may prove informative in designing future versions of psychopathology classification systems. Other researchers have chosen a more direct approach, as is evidenced in a series of articles in a special section of the *Journal of Consulting and Clinical Psychology* (*JCCP*) entitled "Development of theoretically coherent alternatives to the *DSM-IV*" (Follette, 1996). The authors of this special issue of *JCCP* pose radical alternatives to the *DSM*. The alternative classification systems are proposed from a clearly stated behavioral theoretical viewpoint, which differs considerably from many of the more biologically-based approaches described above.

A number of the alternatives and improvements suggested in the 1996 *JCCP* are based on functional analysis, the identification of the antecedents, and consequences of each behaviors (Hayes, Wilson, Gifford, Follette, & Strosahl, 1996; Scotti, Morris, McNeil, & Hawkins, 1996; Wulfert, Greenway, & Dougher, 1996). Wulfert et al. (1996) use the example of depression as a disorder which may be caused by a host of different factors (biological, cognitive, social skills deficits, or a lack of reinforcement). They argue that the fact that functionally similar behavior patterns may have very different structures may contribute to the heterogeneity found in the presumably homogenous *DSM* categories. These authors contend that functional analysis may constitute one means of identifying homogenous subgroups whose behavior share similar antecedents and consequences. This approach could be used to refine the existing *DSM* categories, and to inform treatment strategies. Hayes et al. (1996) describe a classification system based on functional analysis as a fundamentally different alternative to the current syndromal classification system. They proposed that a classification system could be based on dimensions derived from the combined results of multiple functional analyses tied to the same dimension. Each dimension would then be associated with specific assessment methods and therapy recommendations. The authors describe one such functional dimension, "experiential avoidance," and illustrate its utility across disorders such as substance abuse, obsessive-compulsive disorder, panic disorder, and borderline personality disorder (Hayes et al.). This model provides an alternative to the current *DSM* syndromal model, using the methodological approach of functional analysis. Scotti et al. (1996) proposed a similar system of categorization of childhood and adolescent disorders utilizing functional analysis. Hayes et al. (1996), Scotti et al. (1996), and Wulfert et al. (1996) have all successfully illustrated that alternatives or improvements to the current *DSM* system are possible.

However, functional analysis is a distinctly behavioral approach which assumes that a learned stimulus–response connection is an important element in the development or maintenance of psychopathology. Other authors, for example, proponents of genetic or biological explanations of psychopathology described above, might strongly oppose a classification system based purely on the methodology and tenets of a behavioral approach. Others disagree with the notion of any unified system of classification of psychopathology, arguing that no one diagnostic system will be equally useful for all of the classes of disorders now included in the *DSM* (e.g., what works for Axis I may not apply to the personality or childhood disorders). Several of these authors have taken a more radical stance, asserting the need for separate diagnostic systems for different classes of disorder (Kazdin & Kagan, 1994; Koerner, Kohlenberg, & Parker, 1996).

3.08.6 CONCLUSIONS

The assessment and classification of psychopathology is relatively new and controversies abound. However, the heated debates regarding issues such as comorbidity, types of taxonomies, and alternative approaches are indicators of the strong interest in this area. Comparisons between the classification of psychopathology and taxonomies in the basic sciences and medicine can be informative. However, the classification of psychopathology is a difficult task, and the methods used in other fields are not always applicable. It is likely that the systems of classification, informed by the continuing debates and research on the topic, will continue to evolve at a rapid pace. As Clark et al. (1995) remarked, the science of classification has inspired research in new directions and helped to guide future developments of psychopathology.

ACKNOWLEDGMENTS

We would like to thank Amy Przeworski and Melinda Freshman for their help in editing this chapter.

3.08.7 REFERENCES

Abikoff, H., & Klein, R. G. (1992). Attention-deficit hyperactivity and conduct disorder: Comorbidity and implications for treatment. *Journal of Consulting and Clinical Psychology, 60*(6), 881–892.

American Psychiatric Association (1933). Notes and comment: Revised classification of mental disorders. *American Journal of Psychiatry, 90,* 1369–1376.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Andreasen, N. C., & Carpenter, W. T. (1993). Diagnosis and classification of schizophrenia. *Schizophrenia Bulletin, 19*(2), 199–214.

Andreasen, N. C., Flaum, M., Swayze, V. W., Tyrrell, G., & Arndt, S. (1990). Positive and negative symptoms in schizophrenia: A critical reappraisal. *Archives of General Psychiatry, 47,* 615–621.

Angrist, B., Rotrosen, J., & Gershon, S. (1980). Differential effects of amphetamine and neuroleptics on negative vs. positive symptoms in schizophrenia. *Psychopharmacology, 72,* 17–19.

August, G. J., & Garfinkel, B. D. (1990). Comorbidity of ADHD and reading disability among clinic-referred children. *Journal of Abnormal Child Psychology, 18,* 29–45.

Baumann, U. (1995) Assessment and documentation of psychopathology, *Psychopathology, 28* (Suppl.1), 13–20.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56*(6), 893–897.

Beck, A. T., & Steer, R. A. (1987). *Beck depression inventory manual.* San Antonio, TX: The Psychological Corporation.

Bellack, A. S., & Hersen, M. (Eds.) (1998). *Behavioral assessment: A practical handbook.* Needham Heights, MA: Allyn & Bacon.

Bergner, R. M. (1997). What is psychopathology? And so what? *Clinical Psychology: Science and Practice, 4,* 235–248.

Berrios, G. E., & Hauser, R. (1988). The early development of Kraepelin's ideas on classification: A conceptual history. *Psychological Medicine, 18,* 813–821.

Biederman, J., Newcorn, J., & Sprich, S. (1991). Comorbidity of attention deficit hyperactivity disorder with conduct, depressive, anxiety, and other disorders. *American Journal of Psychiatry, 148*(5), 564–577.

Blacker, D., & Tsuang, M. T. (1992). Contested boundaries of bipolar disorder and the limits of categorical diagnosis in psychiatry. *American Journal of Psychiatry, 149*(11), 1473–1483.

Blashfield, R. K. (1990). Comorbidity and classification. In J. D. Maser & C. R. Cloninger (Eds.), *Comorbidity of mood and anxiety disorders* (pp. 61–82). Washington, DC: American Psychiatric Press.

Blashfield, R. K. (1991). Models of psychiatric classification. In M Hersen & S. M. Turner (Eds.), *Adult psychopathology and diagnosis* (pp. 3–22). New York: Wiley.

Blashfield, R. K., & Draguns, J. G. (1976). Toward a taxonomy of psychopathology: The purpose of psychiatric classification. *British Journal of Psychiatry, 129,* 574–583.

Blashfield, R. K., & Livesley, W. J. (1991). Metaphorical analysis of psychiatric classification as a psychological test. *Journal of Abnormal Psychology, 100,* 262–270.

Boyd, J. H., Burke, J. D., Gruenberg, E., Holzer, C. E., Rae, D. S., George, L. K., Karno, M., Stoltzman, R., McEvoy, L., & Nestadt, G. (1984). Exclusion criteria of *DSM-III. Archives of General Psychiatry, 41,* 983–989.

Brett, E. A. (1993). Classifications of posttraumatic stress disorder in *DSM-IV*: Anxiety disorder, dissociative disorder, or stress disorder? In J. R. T. Davidson & E. B. Foa (Eds.), *Posttraumatic stress disorder: DSM-IV and beyond* (pp. 191–204). Washington, DC: American Psychiatric Press.

Broverman, I. K., Broverman, D. M., Clarkson, F. E., Rosenkrantz, P. S., & Vogel, S. R. (1970). Sex-role stereotypes and clinical judgments of mental health. *Journal of Consulting and Clinical Psychology, 34*(1), 1–7.

Brown, T. A., & Barlow, D. H. (1992). Comorbidity among anxiety disorders: Implications for treatment and *DSM-IV. Journal of Consulting and Clinical Psychology, 60*(6), 835–844.

Buchanan, R. W., & Carpenter, W. T. (1994). Domains of psychopathology: an approach to the reduction of heterogeneity in schizophrenia. *The Journal of Nervous and Mental Disease, 182*(4), 193–204.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Minnesota multiphasic personality inventory (MMPI-2). Administration and scoring.* Minneapolis, MN: University of Minnesota Press.

Cantor, N., & Genero, N. (1986). Psychiatric diagnosis and natural categorization: A close analogy. In T. Millon & G. L. Klerman (Eds.) *Contemporary directions in psychopathology—toward the DSM-IV* (pp. 233–256). New York: Guildford Press.

Carey, G., & Gottesman, I. I. (1978). Reliability and validity in binary ratings: Areas of common misunderstanding in diagnosis and symptom ratings. *Archives of General Psychiatry, 35,* 1454–1459.

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100*(3), 316–336.

Clark, L. A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: Challenges to the current system and future directions. *Annual Review of Psychology, 46,* 121–153.

Clarkin, J. F., Widiger, T. A., Frances, A. J., Hurt, S. W., & Gilmore, M. (1983). Prototypic typology and the borderline personality disorder. *Journal of Abnormal Psychology, 92*(3), 263–275.

Clementz, B. A., & Iacono, W. G. (1990). Nosology and diagnosis. In L. Willerman & D. B. Cohen (Eds.), *Psychopathology,* New York: McGraw-Hill.

Cole, D. A., & Carpentieri, S. (1990). Social status and the comorbidity of child depression and conduct disorder. *Journal of Consulting and Clinical Psychology, 58,* 748–757.

Compton, W. M., & Guze, S. B., (1995). The neo-Kraepelinian revolution in psychiatric diagnosis. *European Archives of Psychiatry & Clinical Neuroscience, 245,* 196–201.

Cornblatt, B. A., Lenzenweger, M. F., Dworkin, R. H., & Erlenmeyer-Kimling, L. (1992). Childhood attentional dysfunctions predict social deficits in unaffected adults at risk for schizophrenia. *British Journal of Psychiatry, 161,* 59–64.

Crow, T. J. (1994). The demise of the Kraepelinian binary system as a prelude to genetic advance. In E. S. Gershon, & C. R. Cloninger (Eds.), *Genetic approaches to mental disorders* (pp. 163–192). Washington, DC: American Psychiatric Press.

Daradkeh, T. K., El-Rufaie, O. E. F., Younis, Y. O., & Ghubash, R. (1997). The diagnostic stability of ICD-10 psychiatric diagnoses in clinical practice. *European Psychiatry, 12,* 136–139.

Davidson, J. R. T., & Foa, E. B. (1991). Diagnostic issues in posttraumatic stress disorder: Considerations for the *DSM-IV*. *Journal of Abnormal Psychology, 100*(3), 346–355.

de Fundia, T. A., Draguns, J. G., & Phillips, L. (1971). Culture and psychiatric symptomatology: A comparison of Argentine and United States patients. *Social Psychiatry, 6*(1), 11–20.

Dworkin, R. H., & Lenzenweger, M. F. (1984). Symptoms and the genetics of schizophrenia: Implications for diagnosis. *American Journal of Psychiatry, 141*(12), 1541–1546.

Dworkin, R. H., Lenzenwenger, M. F., Moldin, S. O., & Cornblatt, B. A. (1987). Genetics and the phenomenology of schizophrenia. In P. D. Harvey & E. F. Walker (Eds.), *Positive and negative symptoms of psychosis: Description, research and future directives* (pp. 258–288). Hillsdale, NJ: Erlbaum.

Eysenck, H. J. (1970). A dimensional system of psychodiagnostics. In A. R. Mahrer (Ed.), *New approaches to personality classification* (pp. 169–207). New York: Columbia University Press.

Feighner, A. C., Baran, I. D., Furman, S., & Shipman, W. M. (1972). Private psychiatry in community mental health. *Hospital and Community Psychiatry, 23*(7), 212–214.

Feinstein, A. R. (1970). The pre-therapeutic classification of co-morbidity in chronic disease. *Journal of Chronic Diseases, 23,* 455–468.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, K. B. W. (1995). *Structured clinical interview for DSM-IV Axis I Disorders*. Washington, DC: American Psychiatric Press.

Foa, E. B., Riggs, D. S., Dancu, C. V., & Rothbaum, B. O. (1993). Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *Journal of Traumatic Stress, 6,* 459–473.

Foa, E. B., & Rothbaum, B. (1998). *Treating the trauma of rape*. New York: Guilford Press.

Follette, W. C. (1996). Introduction to the special section on the development theoretically coherent alternatives to the *DSM* system. *Journal of Consulting and Clinical Psychology, 64,* 1117–1119.

Follette, W. C., & Houts, A. C. (1996) Models of scientific progress and the role of theory in taxonomy development: A case study of the *DSM*. *Journal of Consulting and Clinical Psychology, 64,* 1120–1132.

Frances, A. J. (1998). Problems in defining clinical significance in epidemiological studies. *Archives of General Psychiatry, 55,* 119.

Frances, A. J., First, M. B., & Pincus, H. A. (1995). *DSM-IV guidebook*. Washington, DC: American Psychiatric Press.

Frances, A. J., First, M. B., Widiger, T. A., Miele, G. I., M., Tilly, S. M., Davis, W. W., & Pincus, H. A. (1991). An A to Z guide to *DSM-IV* conundrums. *Journal of Abnormal Psychology, 100*(3), 407–412.

Frances, A. J., Widiger, T. A., & Fyer, M. R. (1990). The influence of classification methods on comorbidity. In J. D. Maser & C. R. Cloninger (Eds.), *Comorbidity of mood and anxiety disorders* (pp. 41–59). Washington, DC: American Psychiatric Press.

Fulop, G., Strain, J., Vita, J., Lyons, J. S. & Hammer, J. S. (1987). Impact of psychiatric comorbidity on length of hospital stay for medical/surgical patients: A preliminary report *American Journal of Psychiatry, 144,* 878–882.

Fyer, A.J., Mannuzza, S., Chapman, T. F., Lipsitz, J., Martin, L. Y. & Klein, D. F. (1996). Panic disorder and social phobia: Effects of comorbidity on familial transmission. *Anxiety 2,* 173–178.

Goldenberg, I. M., White, K., Yonkers, K., Reich, J., Warshaw, M. G., Goisman, R. M., & Keller, M. B. (1996). The infrequency of "Pure Culture" diagnoses among the anxiety disorders. *Journal of Clinical Psychiatry, 57*(11), 528–533.

Goodman, W. K., Price, L. H., Rasmussen, S. A., & Mazure, C. (1989a). The Yale-Brown obsessive compulsive scale: I. Development, use and reliability. *Archives of General Psychiatry, 46,* 1006–1011.

Goodman, W. K., Price, L. H., Rasmussen, S. A., & Mazure, C. (1989b). The Yale-Brown obsessive compulsive scale: II. Validity. *Archives of General Psychiatry, 46,* 1012–1016.

Gorenstein, E. E. (1992). *The science of mental illness*. New York: Academic Press.

Gove, W. R. (1980). Mental illness and psychiatric treatment among women. *Psychology of Women Quarterly, 4,* 345–362.

Grove, W. M., Lebow, B. S., Clementz, B. A., Cerri, A., Medus, C., & Iacono, W. G. (1991). Familial prevalence and coaggregation of schizotypy indicators: A multitrait family study. *Journal of Abnormal Psychology, 100,* 115–121.

Haghighat, R. (1994). Cultural sensitivity: *ICD-10* versus *DSM-III–R*. *International Journal of Social Psychiatry, 40,* 189–193.

Hamilton, S., Rothbart, M., & Dawes, R. (1986). Sex bias, diagnosis and *DSM-III*. *Sex Roles, 15,* 269–274.

Harrison, W. M., Cooper, T. B., Stewart, J. W., Quitkin, F. M., McGrath, P. J., Liebowitz, M. R., Rabkin, J. R., Markowitz, J. S., & Klein, D. F. (1984). The tyramine challenge test as a marker for melancholia. *Archives of General Psychiatry, 41*(7), 681–685.

Hayes, S. C., Wilson, K. G., Gifford, E. V., Follette, V. M., & Strosahl, K. D. (1996). Experiential avoidance and behavioral disorders: A functional dimensional approach to diagnosis and treatment. *Journal of Consulting*

& Clinical Psychology, 64(6), 1152–1168.

Helzer, J. E., & Robins, L. N. (1988). The Diagnostic interview schedule: Its development, evolution, and use. *Social Psychology and Psychiatric Epidemiology, 23*(1), 6–16.

Herbert, J. D., Hope, D. A., & Bellack, A. S. (1992). Validity of the distinction between generalized social phobia and avoidant personality disorder. *Journal of Abnormal psychology, 101*(2), 332–339.

Hollander, E. (1993). Obsessive-compulsive spectrum disorders: An overview. *Psychiatric Annals, 23*(7), 355–358.

Horowitz, L. M., Post, D. L, French, R. S., Wallis, K. D., & Siegelman, E. Y. (1981). The prototype as a construct in abnormal psychology: II. Clarifying disagreement in psychiatric judgments, *Journal of Abnormal Psychology, 90*(6), 575–585.

Hudson, J. I., & Pope, H. G. (1990). Affective spectrum disorder: Does antidepressant response identify a family of disorders with a common pathophysiology? *American Journal of Psychiatry, 147* (5), 552–564.

Jacoby, R. (1996). Problems in the use of ICD-10 and DSM-IV in the psychiatry of old age. In N. S. Costas & H. Hanns (Eds.), *Neuropsychiatry in old age: An update. Psychiatry in progress series* (pp. 87–88). Gottingen, Germany: Hogrefe & Huber.

Janca, A., Kastrup, M. C., Katschnig, H., & Lopez-Ibor, J. J., Jr. (1996). The ICD-10 multiaxial system for use in adult psychiatry: Structure and applications. *Journal of Nervous and Mental Disease, 184,* 191–192.

Johnstone, E. C., Crow, T. J., Frith, C. D., Stevens, M., Kreel, L., & Husband, J. (1978). The dementia of dementia praecox. *Acta Psychiatrica Scandanavia, 57,* 305–324.

Kanfer, F. H., & Saslow, G. (1965). Behavioral analysis. *Archives of General Psychiatry, 12,* 529–538.

Kahn, E. (1959). The Emil Kraepelin memorial lecture. In D. Pasamanick (Ed.), *Epidemiology of mental disorders*. Washington, DC: American Association for the Advancement of Science. (As cited in M. Hersen & S. M. Turner (Eds.), *Adult psychopathology and diagnosis* (p. 20). New York: Wiley).

Katon, W., Lin, E., Von Korff, M., Russo, J., Lipscomb, P., & Bush, T. (1991). Somatization: A spectrum of severity. *American Journal of Psychiatry, 148*(1), 34–40.

Kay, S. R. (1990). Significance of the positive-negative distinction in schizophrenia. *Schizophrenia Bulletin, 16* (4), 635–652.

Kazdin, A. E., & Kagan, J. (1994). Models of dysfunction in developmental psychopathology. *Clinical Psychology-Science & Practice, 1*(1), 35–52.

Kendell, R. E. (1975). The concept of disease and its implications for psychiatry. *British Journal of Psychiatry, 127,* 305–315.

Kendell, R. E. (1982). The choice of diagnostic criteria for biological research. *Archives of General Psychiatry, 39,* 1334–1339.

Kendler, K. S. (1990). Toward a scientific psychiatric nosology. *Archives of General Psychiatry, 47,* 969–973.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshelman, S., Wittchen, H. U., & Kendler, K. S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. *Archives of General Psychiatry, 51,* 8–19.

Klein, D. F. (1989). The Pharmacological validation of psychiatric diagnosis. In L. N. Robins & J. E. Barrett (Eds.), *The validity of psychiatric diagnosis* (pp. 203–214). New York: Raven Press.

Klein, D. F. (1993). False suffocation alarms, spontaneous panics, and related conditions: An integrative hypothesis. *Archives of General Psychiatry, 50,* 306–317.

Klein, D. F. (1994). Testing the suffocation false alarm theory of panic disorder. *Anxiety, 1,* 1–7.

Klerman, G. L. (1978). The evolution of a scientific nosology. In J. C. Shershow (Ed.), *Schizophrenia: Science and practice* (pp. 99–121). Cambridge, MA: Harvard University Press.

Koerner, K., Kohlenberg, R. J., & Parker, C. R. (1996). Diagnosis of personality disorder: A radical behavioral alternative. *Journal of Consulting & Clinical Psychology, 64*(6), 1169–1176.

Kraeplin, E. (1971). *Dementia praecox and paraphrenia* (R. M. Barklay, Trans.). Huntington, NY: Krieger. (Original work published in 1919).

Lewinsohn, P. M., Rohde, P., Seeley, J. R., & Hops, H. (1991). Comorbidity of unipolar depression: I. Major depression with dysthymia. *Journal of Abnormal Psychology, 100*(2), 205–213.

Lilienfeld, S. O., Waldman, I. D., & Israel, A. C. (1994). A critical examination of the use of the term and concept of comorbidity in psychopathology research. Clinical Psychology-Science & Practice, 1(1), 71–83.

Livesley, W. J. (1985a). The classification of personality disorder: I. The choice of category concept. *Canadian Journal of Psychiatry, 30*(5), 353–358.

Livesley, W. J. (1985b). The classification of personality disorder: II. The problem of diagnostic criteria. *Canadian Journal of Psychiatry, 30*(5), 359–362.

Lykken, D. T. (1984). Psychopathic personality. In R. I. Corsini (Ed.), *Encyclopedia of psychology* (pp. 165–167). New York, Wiley. (As cited in B. A. Clementz & W. E. Iacono (1990). In L. Willerman & D. B. Cohen (Eds.), *Psychopathology*, New York: McGraw-Hill.)

Macklin, M. L., Metzger, L. J., Litz, B. T., McNally, R. J., Lasko, N. B., Orr, S. P., & Pitman, R. K. (1998). Lower precombat intelligence is a risk factor for posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology, 66,* 323–326

Martinez, J. M., Papp, L. A., Coplan, J. D., Anderson, D. E., Mueller, C. M., Klein, D. F., & Gorman, J. M. (1996). Ambulatory monitoring of respiration in anxiety. *Anxiety, 2,* 296–302.

Mayr, E. (1981). Biological classification: Toward a synthesis of opposing methodologies. *Science, 214*(30), 510–516.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1981). *POMS manual* (2nd ed.). San Diego: Educational and Industrial Testing Service.

McNally, R., & Shin, L. M. (1995). Association of intelligence with severity of posttraumatic stress disorder symptoms in Vietnam combat veterans. *American Journal of Psychiatry, 152*(6), 936–938.

McReynolds, P. (1989). Diagnosis and clinical assessment: Current status and major issues. *Annual Review of Psychology, 40,* 83–108.

Meehl, P. E. (1986). Diagnostic taxa as open concepts: Metatheoretical and statistical questions about reliability and construct validity in the grand strategy of nosological revision. In T. Millon & G. L. Klerman (Eds.), *Contemporary directions in psychopathology: Toward the DSM-IV* (pp. 215–231). New York: Guilford Press.

Menninger, K. (1963). *The vital balance: The life process in mental health and illness*. New York: Viking Press.

Mezzich, J. E., Fabrega, H., Coffman, G. A., & Haley, R. (1989). DSM-III disorders in a large sample of psychiatric patients: Frequency and Specificity of diagnoses. *American Journal of Psychiatry, 146*(2), 212–219.

Michels, R., Siebel, U., Freyberger, H. J., Stieglitz, R. D., Schaub, R. T., & Dilling, H. (1996). The multiaxial system of ICD-10: Evaluation of a preliminary draft in a multicentric field trial. *Psychopathology, 29,* 347–356.

Millon, T. (1991). Classification in psychopathology: Rationale, alternatives and standards. *Journal of Abnormal Psychology, 100,* 245–261.

Moras, K., DiNardo, P. A., Brown, T. A., & Barlow, D. H. (1991). Comorbidity and depression among the DSM-III-R anxiety disorders. Manuscript as cited in Brown, T.

A. & Barlow, D. H. (1992). Comorbidity among anxiety disorders: Implications for treatment and DSM-IV. *Journal of Consulting and Clinical Psychology, 60*(6), 835–844.

Morey, L. C. (1988). Personality disorders in DSM-III and DSM-III-R: Convergence, coverage, and internal consistency. *American Journal of Psychiatry, 145*(5), 573–577.

Morey, L. C. (1991). Classification of mental disorder As a collection of hypothetical constructs. *Journal of Abnormal Psychology, 100*(3), 289–293.

Morey, L. C. (1992). *Personality disorder NOS: Specifying patterns of the otherwise unspecified.* Paper presented at the 100th Annual Convention of the American Psychological Association. Washington, DC: USA.

Mowrer, O. H. (1960). "Sin," the lesser of two evils. *American Psychologist, 15,* 301–304.

Oldham, J. M., Skodal, A. E., Kellman, H. D., Hyler, S. E., Rosnick, L., & Davies, M. (1992). Diagnosis of DSM-II-R personality disorders by two structured interviews: Patterns of comorbidity. *American Journal of Psychiatry, 149,* 213–220.

Papp, L. A., Klein, D. F., Martinez, J., Schneier, F., Cole, R., Liebowitz, M. R., Hollander, E., Fyer, A. J., Jordan, F., & Gorman, J. M. (1993). Diagnostic and substance specificity of carbon dioxide-induced panic. *American Journal of Psychiatry, 150,* 250–257.

Patel, V., & Winston, M. (1994). "Universality of mental illness" revisited: Assumptions, artefacts and new directions. *British Journal of Psychiatry, 165,* 437–440.

Pine, D. S., Weese-Mayer, D. E., Silvestri, J. M., Davies, M., Whitaker, A., & Klein, D. F. (1994). Anxiety and congenital central hypoventilation syndrome. *American Journal of Psychiatry, 151,* 864–870.

Pogue-Geile, M. F., & Harrow, M. (1984). Negative and positive symptoms in schizophrenia and depression: A follow-up. *Schizophrenia Bulletin, 10*(3), 371–387.

Quay, H. C., Routh, D. K., & Shapiro, S. K. (1987). Psychopathology of childhood: From description to validation. *Annual Review of Psychology, 38,* 491–532.

Quitkin, F. M., Stewart, J. W., McGrath, P. J., Tricamo, E., Rabkin, J. G., Ocepek-Welikson, K., Nunes, E., Harrison, W., & Klein, D. F. (1993). Columbia atypical depression: A sub-group of depressives with better response to MAOI than to tricyclic antidepressants or placebo. *British Journal of Psychiatry, 163,* 30–34.

Rapee, R. M., Sanderson, W. C., & Barlow, D. H. (1988). Social phobia features across the DSM-III-R anxiety disorders. *Journal of Psychopathology and Behavioral Assessment, 10*(3), 287–299.

Regier, D. A., Kaelber, C. T., Rae, D. S., Farmer, M. E., Knauper, B., Kessler, R. C., & Norquist, G. S. (1998) Limitations of diagnostic criteria and assessment instruments for mental disorders. Implications for research and policy. *Archives of General Psychiatry, 55,* 109–115.

Regier, D. A., Kaelber, C. T., Roper, M. T., Rae, D. S. & Sartorius, N. (1994). The ICD-10 clinical field trial for mental and behavioral disorders: Results in Canada and the United States. *American Journal of Psychiatry, 151,* 1340–1350.

Riskind, J. H., Beck, A. T., Brown, G., & Steer, R. A. (1987). Taking the measure of anxiety and depression. Validity of the reconstructed Hamilton scale. *Journal of Nervous and Mental Disease, 175*(8), 474–479.

Robins, E., & Guze, S. B. (1970). Establishment of diagnostic validity in psychiatric illness: Its application to schizophrenia. *American Journal of Psychiatry, 126,* 983–987.

Robins, L. N. (1994). How recognizing "comorbidities" in psychopathology may lead to an improved research nosology. *Clinical Psychology: Science and Practice, 1,* 93–95.

Robins, L. N., Wing, J., Wittchen, H. J., Helzer, J. E., Barbor, T. F., Burke, J. D., Farmer, A., Jablenski, A., Pickens, R., Reiger, D. A., Sartorius, N., & Towle, L. H. (1988). The composite international Diagnostic Interview: An epidemiological instrument suitable for use in conjunction with different diagnostic systems and different cultures. *Archives of General Psychiary, 45,* 1069–1077.

Rosch, E. H. (1978). Principles of categorization. In E. H. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161–1178.

Sartorius, N., & Janca, A. (1996). Psychiatric assessment instruments developed by the World Health Organization. *Social Psychiatry and Psychiatric Epidemiological, 32,* 55–69.

Scotti, J. R., Morris, T. L., McNeil, C. B., & Hawkins, R. P. (1996). DSM-IV and disorders of childhood and adolescence: Can structural criteria be functional? *Journal of Consulting & Clinical Psychology, 64*(6), 1177–1191.

Shea, M. T., Widiger, T. A., & Klein, M. H. (1992). Comorbidity of personality disorders and depression: Implications for treatment. *Journal of Consulting and Clinical Psychology, 60,* 857–868.

Spiegel, D., & Cardena, E. (1991). Disintegrated experience: The dissociative disorders revisited. *Journal of Abnormal Psychology, 100*(3), 366–378.

Spielberger, C. D., Gorsuch, R. R., & Lushene, R. E. (1970). *State-trait anxiety Inventory: Test manual for form X.* Palo Alto, CA: Consulting Psychologists Press.

Spitzer, R. B. (1997). Brief comments for a psychiatric nosologist weary from his own attempts to define mental disorder: Why Ossorio's definition muddles and Wakefield's "Harmful Dysfunction" illuminates the issues. *Clinical Psychology: Science and Practice, 4,* 259–261.

Spitzer, R. B. (1998). Diagnosis and need for treatment are not the same. *Archieves of General Psychiatry, 55,* 120.

Spitzer, R. B., Endicott, J., & Robins, E. (1978). Research diagnostic criteria: Rationale and reliability. *Archives of General Psychiatry, 35*(6), 773–782.

Stein, D. J., Hollander, E., & Klein, D. F. (1994). Biological markers of depression and anxiety. *Medicographia, 16*(1), 18–21.

Stevens, J. R. (1982). Neuropathology of schizophrenia. *Archives of General Psychiatry, 39,* 1131–1139.

Szasz, T. (1960). The myth of mental illness. *American Psychologist, 15,* 113–118.

Tattersall, I. (1995). The fossil trail: How we know what we think we know about human evolution. New York: Oxford University Press. (As cited in Follette, W. C., & Houts, A. C. (1996). Models of scientific progress and the role of theory in taxonomy development: A case study of the DSM. *Journal of Consulting and Clinical Psychology, 64,* 1120–1132.)

Taylor, M. A. (1992). Are schizophrenia and affective disorder related? A selective literature review. *American Journal of Psychiatry, 149,* 22–32.

Taylor, M. A., & Amir, N. (1994). Are schizophrenia and affective disorder related?: The problem of schizoaffective disorder and the discrimination of the psychoses by signs and symptoms. *Comprehensive Psychiatry, 35*(6), 420–429.

Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 681–706). Hillsdale, NJ: Erlbaum.

Uestuen, T. B., Bertelsen, A., Dilling, H., & van Drimmelen, J. (1996). *ICD-10 casebook: The many faces of mental disorders—Adult case histories according to ICD-10.* Washington, DC: American Psychiatric Press.

Ullman, L. P., & Krasner, L. (1975). *A psychological* *approach to abnormal behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Wakefield, J. C. (1992). Disorder as harmful dysfunction: A conceptual critique of DSM-III-R's definition of mental disorder. *Psychological Review, 99,* 232–247.

Wakefield, J. C. (1997a). Diagnosing DSM-IV, Part I: DSM-IV and the concept of disorder. *Behavioral Research Therapy, 35*(7), 633–649.

Wakefield, J. C. (1997b). Diagnosing DSM-IV, Part II: Eysenck (1986) and the essential fallacy. *Behavioral Research Therapy, 35*(7), 651–665.

Wakefied, J. C. (1997c). Normal inability versus pathological disability: Why Ossorio's definition of mental disorder is not sufficient. *Clinical Psychology: Science and Practice, 4,* 249–258.

Watson, D., & Clark, L. A. (1992). Affects separable and inseparable: On the hierarchical arrangement of the negative affects. *Journal of Personality and Social Psychology, 62*(3), 489–505.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063–1070.

Watson, J. D., & Crick, F. H. C. (1953a). General implications of the structure of deoxyribose nucleic acid. *Nature, 171,* 964–967. (As cited in Clementz B. A., & Iacono, W. G. (1990). In L. Willerman & D. B. Cohen (Eds.), *Psychopathology*. New York: McGraw-Hill.)

Watson, J. D., & Crick, F. H. C. (1953b). A structure for deoxyribose nucleic acid. *Nature, 171,* 737–738. (As cited in Clementz, B. A., & lacono, W. G. (1990). In L. Willerman & D. B. Cohen (Eds.), *Psychopathology*. New York: McGraw-Hill.

Widiger, T. A. (1992). Categorical versus dimensional classification: Implications from and for research. *Journal of Personality Disorders, 6*(4), 287–300.

Widiger, T. A., & Frances, A. (1985). The DSM-III personality disorders: Perspectives form psychology. *Archives of General Psychiatry, 42,* 615–623.

Widiger, T. A., & Rogers, J. H. (1989). Prevalence and comorbidity of personality disorders. *Psychiatric Annals, 19*(3), 132–136.

Widiger, T. A., & Trull, T. J. (1993). The Scholarly development of DSM-IV. In J. A. Costa, E. Silva, & C. C. Nadelson (Eds.), *International Review of Psychiatry* (pp. 59–78). Washington DC, American Psychiatric Press.

Widiger, T. A., Trull, T. J., Hurt, S. W., Clarkin, J., & Frances, A. (1987). A multidimensional scaling of the DSM-III personality disorders. *Archives of General Psychiatry, 44,* 557–563.

Wiggins, J. S. (1982). Circumplex models of interpersonal behavior in clinical psychology. In P. C. Kendall & J. N. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 183–221). New York: Wiley.

Wilson, G. T. (1996). Empirically validated treatments: Realities and resistance. *Clinical Psychology-Science & Practice, 3*(3), 241–244.

World Health Organization (1992). *International statistical classification disease, 10th revision (ICD-10)*. Geneva, Switzerland: Author.

World Health Organization (1996). *Multiaxial classification of child and adolescent psychiatric disorders: The ICD-10 classification of mental and behavioral disorders in children and adolescents*. Cambridge, UK: Cambridge University Press.

Wulfert, E., Greenway, D. E., & Dougher, M. J. (1996). A logical functional analysis of reinforcement-based disorders: Alcoholism and pedophilia. *Journal of Consulting and Clinical Psychology, 64*(6), 1140–1151.

Zimmerman, M., Pfohl, B., Coryell, W. H., Corenthal, C., & Stangl, D. (1991). Major depression and personality disorder. *Journal of Affective Disorders, 22,* 199–210.

Zinbarg, R. E., & Barlow, D. D. (1991). Mixed anxiety-depression: A new diagnostic category? In R. M. Rapee & D. H. Barlow (Eds.), *Chronic anxiety: Generalized anxiety disorder and mixed anxiety-depression* (pp. 136–152). New York: Guilford Press.

Zinbarg, R. E., Barlow, D. H., Liebowitz, M., Street, L., Broadhead, E., Katon, W., Roy-Byrne, P., Lepine, J. P., Teherani, M., Richards, J., Brantley, P., & Kraemer, H. (1994). The DSM-IV field trial for mixed anxiety-depression. *American Journal of Psychiatry, 151*(8), 1153–1162.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

3.09
Intervention Research: Development and Manualization

JOHN F. CLARKIN
Cornell University Medical College, New York, NY, USA

3.09.1 INTRODUCTION 189
3.09.2 AIMS AND INGREDIENTS OF A TREATMENT MANUAL 190
3.09.3 DIMENSIONS ALONG WHICH THE MANUALS VARY 191
3.09.3.1 Process of Manual Generation and Elucidation 191
3.09.3.2 Patient Population 191
3.09.3.3 Knowledge Base of the Disorder 191
3.09.3.4 Treatment Strategies and Techniques 192
3.09.3.5 Treatment Duration 192
3.09.3.6 Treatment Format 192
3.09.3.7 Level of Abstraction 192
3.09.4 RANGE OF TREATMENT MANUALS 192
3.09.5 REPRESENTATIVE TREATMENT MANUALS 193
3.09.5.1 IPT: An Early Manual 193
3.09.5.2 Cognitive-behavioral Treatment of BPD 194
3.09.5.3 Psychosocial Treatment For Bipolar Disorder 195
3.09.5.4 Other Treatment Manuals 196
3.09.6 ADVANTAGES OF TREATMENT MANUALS 196
3.09.7 POTENTIAL DISADVANTAGES AND LIMITATIONS OF TREATMENT MANUALS 196
3.09.8 USE AND ROLE OF TREATMENT MANUALS 198
3.09.8.1 Efficacy to Effectiveness 198
3.09.9 SUMMARY 198
3.09.10 REFERENCES 199

3.09.1 INTRODUCTION

Psychotherapy research can progress only if the treatments that are investigated can be replicated by each of the therapists within the study, and the therapy can be described to the consumers of the research results. Standards of psychotherapy research have reached a point that one cannot communicate about results without a written manual describing the treatment and methods for demonstrating therapists' adherence and competence in the delivery of the treatment. In the recent history of psychotherapy research, the manual has played a vital role in the specification of clinical trials. In addition, if a science and not just an art of psychotherapy is to flourish, there must be methods to teach clinicians how to perform the various therapies with adherence to the specific treatment strategies and techniques, and with competence in this complex interpersonal process. It is in this context that the written description of psychotherapies for specific patient populations in the form of manuals has become an essential fixture in the psychotherapy world.

There has been much lament that clinical research has little if any impact on clinical practice (Talley, Strupp, & Butler, 1994). Researchers complain that their findings are ignored. Clinicians argue that the research on rarified samples yield findings that are detailed, tedious, and irrelevant to their heterogeneous patients. It may be that the psychotherapy treatment manuals will provide a bridge between clinical research and clinical practice. The treatment manual is an accessible yield of clinical research that the clinician may find helpful, if not necessary as health delivery systems change.

This chapter draws upon various sources to describe the contents of the psychotherapy manual, provides some examples of the therapy manual movement, and discusses the advantages and disadvantages of this technology both for research and the instruction of therapists. It is interesting that, although medication management is a social process between patient and physician, with many similarities to psychotherapy, there has been little development of manuals for medication management. The informative exception is the medication management manual used in the multisite treatment study of depression (Elkin et al., 1989). Despite the fact that medication adherence is often variable and generally poor, and that there is great variation in physician behavior in dispensing medications, this area has attracted little attention, apparently because researchers and clinicians assume that providing medication is a simple and standardized behavior. We suspect this assumption is false, and that more sophisticated medication and medication–psychotherapy studies in the future will have manuals with validity checks on both psychotherapy and medication management.

3.09.2 AIMS AND INGREDIENTS OF A TREATMENT MANUAL

The concept of a treatment manual has quickly evolved in the research and clinical communities. The treatment manual was acclaimed (Luborsky & DeRubeis, 1984) as a quantum leap forward which introduced a new era in clinical research. Since that time there has been the development of numerous treatment manuals. There are even special series geared around treatment manuals, and at times accompanying materials for patients as well. In addition, the intellectual community is moving forward in defining what should be included in a treatment manual to meet standards for treatment development.

An example of the development of standards for treatment manuals is the National Institute on Drug Abuse (NIDA, Moras, 1995). The experts on this panel suggested that the aims of a treatment manual are to:

(i) make the therapy reproducible by therapists other than the immediate research group;

(ii) include information on those issues the investigator thinks are important determinants of therapy response;

(iii) delineate the knowledge base and skills needed by a therapist to learn from the manual.

Given these aims, the group suggested the following content for the manual:

(i) theory and principles of the treatment (theory underpinning the therapy, decision rules and principles to guide the therapists, information on the therapy that therapists should have to help them adopt the requisite attitude about it, i.e., belief in the treatment);

(ii) elements of the treatment (therapists, primary interpersonal stance in relation to the patient, common and unique elements of therapy);

(iii) intervention strategies for handling problems commonly encountered in delivering the therapy (miscellaneous interventions, specification of interventions not to be used in the therapy);

(iv) companion videotapes for teaching the treatment; and

(v) criteria and method of assessing therapist competence in the treatment.

In essence, these are the criteria by which one can judge the adequacy of any treatment manual.

These criteria are used to assess existing treatment manuals, some of which are used as illustrations in this chapter. As noted previously (Kazdin, 1991), it seems clear that the manuals are often incomplete, as they do not reflect the complexity of treatment and the full scope of the exchanges between therapist and patient. However, the development of treatment manuals has and continues to help the field move forward.

It should be noted that there is a difference between a treatment manual and a book published and marketed that has many but not all of the attributes of a practical, ever-changing treatment manual. In an attempt to make a book of publishable size, things that might be included in the working manual, such as clinician work sheets or rating scales might be eliminated from the book. A notable exception is the work by Linehan which includes both a book (Linehan, 1993a) and an accompanying workbook (Linehan, 1993b) with therapist work sheets and other material most useful in practical delivery of the treatment. The treatment manual of the near future will probably be both a written description of the treatment in book form, and an accompanying CD-ROM that provides video demonstration of the treatment (using therapists and actor/patients).

3.09.3 DIMENSIONS ALONG WHICH THE MANUALS VARY

The existing treatment manuals can be compared not only to the content as specified by the NIDA experts, but also by differences in aspects of the treatment, including: the process by which the authors generated and elucidated their treatment; the patient population for which the treatment is intended; the knowledge base of the disorder being treated; treatment strategies and techniques, duration, and format; and level of abstraction in which the manual is written.

3.09.3.1 Process of Manual Generation and Elucidation

What was the process of manual generation by the authors? Was the manual generated from extensive experience with a patient population, followed by a codification of the techniques that seem to work on a clinical basis? Conversely, was the manual generated from a theoretical understanding of the patient population followed by treatment contact? Was it generated from a specific treatment orientation, for example, cognitive-behavioral, that is then applied to a specific patient population? Is the manual generated by clinical researchers or by clinicians? Was the manual generated during extensive clinical research, profiting from that research, or without research experience? Does the manual have accompanying aids, such as audiotapes and work sheets?

3.09.3.2 Patient Population

Therapy manuals differ in the extent to which they specify for which patients the manual is intended and for which patients the manual has not been attempted, or, indeed, for whom the treatment does not work. Since psychotherapy research funding is focused on *DSM*-diagnosed patient groups, many of the manuals describe treatments for patient groups defined by diagnosis. This fact has both positive and negative aspects. Some diagnoses are close to phenomenological descriptions of behavioral problems, so the diagnoses are immediate descriptions of problems to be solved. An example is the treatment of phobic anxiety, agoraphobia, and panic attacks. In contrast, the diagnosis of depression is more distant from specific behaviors, so the manuals describing treatments for depression have implicit theories about the nature of depression, such as the interpersonal treatment or the cognitive treatment. Linehan's manual for the treatment of parasuicidal individuals is an interesting transitional one, from a behavior (repetitive suicidal behavior) to a diagnosis that sometimes includes the behavior (borderline personality disorder). In contrast, there are manuals written not for patients with specific diagnoses but for those with certain problems such as experiencing emotions (Greenberg, Rice, & Elliott, 1993).

An additional way the manual being considered is specific to a population is the attention given to the details of patient assessment. Ideally, the assessment section would inform the reader of the patients included and excluded, and the subgroups of patients within the diagnosis/category treated by this manual, but treated with the differences in mind depending on such factors as comorbid conditions.

One difficulty that is becoming more clear is that, for those manuals that have been written for a specific disorder, how do they apply to patients with the disorder but with a common comorbid disorder? For example, the interpersonal psychotherapy manual was constructed with ambulatory, nonpsychotic patients with depression in mind, but gives little indication of its application when there are prominent Axis II disorders. Of course, this point also relates to the accumulation of knowledge that occurs all the time, so that a manual published a few years ago may need modification and updating as more data becomes available.

3.09.3.3 Knowledge Base of the Disorder

One could make a cogent argument that a treatment for a particular disorder cannot (should not) be developed until there is sufficient information on the natural course of the disorder at issue. Only with the background of the natural development of the disorder can one examine the impact of intervention. Furthermore, the age of the patients to whom the treatment is applied is most relevant in the context of the longitudinal pattern of the disorder in question. However, in light of the clinical reality of patients in pain needing intervention, the development of intervention strategies cannot wait for longitudinal investigations of the disorder. Both arguments are sound, but we can describe intervention manuals based on the natural history of the disorder versus those that do not have such a database upon which to build.

3.09.3.4 Treatment Strategies and Techniques

Obviously, treatment manuals differ in the strategies and techniques that are described. This would include manualization of treatments that use behavioral, cognitive, cognitive-behavioral, systems, supportive, and psychodynamic techniques. Less obvious is the fact that it may be easier to manualize some strategies/techniques more than others. In fact, some therapies may be more technique oriented and others relatively more interpersonally oriented, and this may have implications for the method of manualization and the results.

An issue related to strategies and techniques is the degree to which the treatment manual describes the stance of the therapist with respect to the patient, the induction of the patient into the treatment roles and responsibilities, and the development of the therapeutic relationship or alliance. Often, it is the ambivalent patient who questions the value of the treatment, is prone to drop out, or to attend therapy with little investment and involvement (e.g., does not carry out therapy homework) that provides a serious challenge to the therapist. The more complete manuals note the typical difficulties and challenges to full participation in the treatment that have been observed in experience with the patient population, and offer guidelines for the therapist in overcoming these challenges.

3.09.3.5 Treatment Duration

Most manuals describe a brief treatment of some 12–20 sessions. The methods for manualizing longer treatments may present a challenge to the field. A brief treatment can be described in a manual session by session. Each session can be anticipated in a sequence and described. This is especially true if the treatment is cognitive-behavioral in strategies and techniques, and can be anticipated as applied to the particular condition/disorder. In contrast, as the treatment becomes longer, and as the treatment deviates from a cognitive-behavioral orientation to one that depends more on the productivity and nature of the individual patient (e.g., interpersonal and psychodynamic treatment), the manual will of necessity become more principle based and reliant on what the patient brings to the situation.

3.09.3.6 Treatment Format

Psychotherapy is delivered in individual, group, marital, and family treatment formats. It would seem to be simpler, if not easier, to articulate a treatment manual for two participants (i.e., therapist and one patient), rather than a manual for therapist and many patients (i.e., family and group treatment). As the number of patients in the treatment room increases, the number of patient-supplied interactions increases, as does the potential for treatment difficulties and road blocks.

3.09.3.7 Level of Abstraction

The very term “manual” conjures up an image of the book that comes in the pocket of a new car. It tells how to operate all the gadgets in the car, and provides a guide as to when to get the car serviced. Diagrams and pictures are provided. It is a “how to” book on repair and maintenance of your new car. Thus, the term “treatment manual” promises to describe how to do the treatment in step-by-step detail.

More relevant to psychotherapy manuals are the manuals that inform the reader—often with graphs and pictures—how to sail a boat or play tennis. This is a more apt analogy because the manual attempts in written (and graphic) form to communicate how to achieve complex cognitive and motor skills. Opponents of treatment manuals will point to golfers who have read all the manuals, and hack their way around the course with less than sterling skill. (To us, this does not argue against the manual, but speaks of its limitations and how it should be used in the context of other teaching devises.)

Ignoring the critics for the moment, this discussion raises the issue of what level of concreteness or abstraction is the manual best formulated. Some manuals describe the treatment session by session, and within the individual session the flow of the session and details about its construction are indicated. Obviously, this is easier to do the shorter the treatment, and the more the treatment can be predicted from the beginning (i.e., the more the treatment is driven by the therapist and little influenced by patient response or spontaneous contribution). Probably the best manuals are those that constantly weave abstract principles and strategies of the treatment with specific examples in the form of clinical vignettes that provide illustrations of the application of the principles to the individual situation.

3.09.4 RANGE OF TREATMENT MANUALS

An exhaustive list of existing treatment manuals cannot be provided for many reasons, not the least of which is that the list would be out of date before this chapter goes to print. The American Psychological Association (APA), Division 12, Task Force on Psychological

Interventions lists the so-called efficacious treatments (Chambless et al., 1998) and has also listed treatment manuals that are relevant to the efficacious treatments (Woody & Sanderson, 1998). Some 33 treatment manuals are listed for 13 disorders and/or problem areas including bulimia, chronic headache, pain associated with rheumatic disease, stress, depression, discordant couples, enuresis, generalized anxiety disorder, obsessive-compulsive disorder, panic disorder, social phobia, specific phobia, and children with oppositional behavior. Not all of the manuals listed here have been published, and must be sent for to the originator. Further, this list is limited to only those treatments that have been judged by the Task Force of Division 12 to meet their criteria for efficaciousness.

3.09.5 REPRESENTATIVE TREATMENT MANUALS

In contrast, a listing of representative treatment manuals is provided here, using as the structure for such sampling the diagnoses in the *Diagnostic and statistical manual of mental disorders* (4th ed., *DSM-IV*).

This chapter uses three treatment manuals to illustrate the present state of the art in manualization of psychotherapy: *Interpersonal psychotherapy of depression* (IPT; Klerman, Weissman, Roundsaville, & Chevron, 1984), the dialectical behavioral treatment (DBT) for borderline personality disorder (BPD) and related self-destructive behaviors (Linehan, 1993a), and the family treatment of bipolar disorder (Miklowitz & Goldstein, 1997). One of these treatments and related treatment manuals (IPT) has achieved recognition in the APA Division 12 listing of efficacious treatments. One (DBT for BPD) is listed as a “probably efficacious treatment.” The third (family treatment for bipolar disorder) is not listed at this time. These manuals are also selected because they relate to existing *DSM* diagnostic categories which are central to research data and to insurance company procedures for reimbursement of treatment. These manuals for the specified disorders are also chosen because they represent three levels of severity of disorder that face the clinician. Ambulatory patients with depression of a nonpsychotic variety are treated with IPT. BPD is a serious disorder, often involving self-destructive and suicidal behavior. Finally, bipolar disorder is a psychiatric condition with biological and genetic causes that can be addressed by both the empirically efficacious medications and psychotherapy. In addition, these three treatment manuals provide some heterogeneity in terms of treatment format, strategies and level of treatment development. IPT is delivered in an individual treatment format (patient and individual therapist), the treatment format that is the simplest to manualize. DBT involves both an individual and a group treatment format. The family treatment of bipolar disorder involves the treatment of the individual with bipolar disorder and family or marital partner. There is less diversity among these three treatments in terms of treatment strategies and techniques. Two (DPT and family treatment of bipolar) are informed by cognitive-behavioral strategies and techniques, but the latter introduces the intriguing concepts of family dynamics and systems issues. The remaining (IPT) is interpersonal in focus and orientation. The three manuals vary in terms of treatment duration. IPT is brief, the family treatment of bipolar is intermediate, and DBT is the longest, presenting the most challenge to issues of manualization and therapist adherence across a longer period of time. A comparison of these three manuals gives a sense of development in this area, as the IPT manual was one of the first, and the latter two provide a view of recent manuals and their contents.

3.09.5.1 IPT: An Early Manual

One of the earliest and most influential manuals is *Interpersonal psychotherapy of depression* by Klerman et al. (1984). This manual was written as a time-limited, outpatient treatment for depressed individuals. The treatment focuses on the current rather than on past interpersonal situations and difficulties. While making no assumption about the origin of the symptoms, the authors connect the onset of depression with current grief, role disputes and/or transitions, and interpersonal deficits.

This brief intervention has three treatment phases which are described clearly in the manual. The first is an evaluation phase during which the patient and therapist review depressive symptoms, give the syndrome of depression a name, and induce the patient into a sick role. (Interestingly, the role of the therapist is not described explicitly.) The patient and therapist discuss and, hopefully, agree upon a treatment focus limited to four possibilities: grief, role disputes, role transitions, or interpersonal deficits. The middle phase of treatment involves work between therapist and patient on the defined area of focus. For example, with current role disputes the therapist explores with the patient the nature of the disputes and options for resolution. The final phase of treatment involves reviewing and consolidating therapeutic gains. A recent addition is the publication of a client workbook (Weissman, 1995) and client assessment forms.

IPT has been used in many clinical trials, the first of which was in 1974 (Klerman, DiMascio, Weissman, Prusoff, & Paykel, 1974). In addition to the IPT manual, a training videotape has been produced and an IPT training program is being developed (Weissman & Markowitz, 1994).

IPT provides a generic framework guiding patient and therapist to discuss current difficulties and this framework has been applied to symptom conditions other than depression. For example, the format has been applied to patients with bipolar disorder (Ehlers, Frank, & Kupfer, 1988), drug abuse (Carroll, Rounsaville, & Gawin, 1991; Rounsaville, Glazer, Wilber, Weissman, & Kleber, 1983), and bulimia (Fairburn et al., 1991). In addition to its initial use of treating depression in ambulatory adult patients, it has now been utilized as an acute treatment, as a continuation and maintenance treatment (Frank et al., 1990), and has been used for geriatric (Reynolds et al., 1992) and adolescent patients (Mufson, Moreau, Weissman, & Klerman, 1993) and in various settings such as primary care and hospitalized elderly. The success of IPT seems to be the articulation of a rather straightforward approach to discussion between patient and therapist, of current situations without the use of more complicated procedures such as transference interpretation.

The IPT manual was one of the first in the field, and its straightforward description of a common-sense approach to patients with depression is readily adopted by many clinicians. However, the process of treatment development and amplification is also relevant to this, one of the earliest and best treatment manuals. It is now clear that depression is often only partially responsive to brief treatments, such as IPT, and that many patients relapse. It would appear that maintenance treatment with IPT may be useful (Frank et al., 1991), and the IPT manual must therefore be amplified for this purpose. Furthermore, it has become clear that depressed individuals with personality disorders respond to treatment less thoroughly and more slowly than depressed individuals without personality disorders (Clarkin & Abrams, 1998). This would suggest that IPT may need modification for those with personality disorders, either in terms of how to manage the personality disorder during treatment, or to include treatment of relevant parts of the personality disorder to the depression.

In order to place IPT in perspective, one could compare it to the cognitive therapy of depression in its earliest articulation (Beck, Rush, Shaw, & Emery, 1979) and with more recent additions (Beck, 1995).

3.09.5.2 Cognitive-behavioral Treatment of BPD

Linehan's (1993a, 1993b) cognitive-behavioral treatment of the parasuicidal individual is an example of the application of a specific school of psychotherapy adapted to a specific patient population defined both by a personality disorder diagnosis (BPD) and repetitive self-destructive behavior.

The rationale and data upon which the patient pathology is understood as related to the treatment is well described and thorough. Linehan points out that theories of personality functioning/dysfunctioning are based upon world views and assumptions. The assumption behind DBT is that of dialectics. The world view of dialectics involves notions of inter-relatedness and wholeness, compatible with feminist views of psychopathology, rather than an emphasis on separation, individuation, and independence. A related principle is that of polarity, that is, all propositions contain within them their own oppositions. As related to the borderline pathology, it is assumed that within the borderline dysfunction there is also function and accuracy. Thus, in DBT, it is assumed that each individual, including the borderline clients, are capable of wisdom as related to their own life and capable of change. At the level of the relationship between borderline client and DBT therapist, dialectics refers to change by persuasion, a process involving truth not as an absolute but as an evolving, developing phenomenon.

Borderline pathology is conceptualized as a dialectical failure on the part of the client. The thinking of the BPD patient has a tendency to become paralyzed in either the thesis or antithesis, without movement toward a synthesis. There is a related dialectical theory of the development of borderline pathology. BPD is seen primarily as a dysfunction of the emotion regulation system, with contributions to this state of malfunction from both biological irregularities and interaction over time with a dysfunctional environment. In this point of view, the BPD client is prey to emotional vulnerability, that is, high sensitivity to emotional stimuli, emotional intensity, and a slow return to emotional baseline functioning. The invalidating environment, seen as crucial to the BPD development, is one in which interpersonal communication of personal and private experiences is met by others in the environment with trivialization and/or punishment. In such an environment, the developing individual does not learn to label private experiences, nor does the individual learn emotion regulation.

These assumptions and related data on the developmental histories and cross-sectional behaviors of those with BPD provide the rationale and shape of the treatment. The major tasks of the treatment are, therefore, to teach the client skills so that they can modulate emotional experiences and related mood-dependent behaviors, and to learn to trust and validate their own emotions, thoughts and activities. The relevant skills are described as four in type: skills that increase interpersonal effectiveness in conflict situations, strategies to increase self-regulation of unwanted affects, skills for tolerating emotional distress, and skills adapted from Zen meditation to enhance the ability to experience emotions and avoid emotional inhibition.

The manual provides extensive material on basic treatment strategies (e.g., dialectical strategies), core strategies of emotional validation (e.g., teaching emotion observation and labeling skills), behavioral validation (e.g., identifying, countering and accepting "shoulds"), and cognitive validation (e.g., discriminating facts from interpretations), and problem-solving strategies. Problem solving consists of analysis of behavior problems, generating alternate behavioral solutions, orientation to a solution behavior, and trial of the solution behavior. The core of the treatment is described as balancing problem-solving strategies with validation strategies.

This manual is exceptional, and provides a new and very high standard in the field for treatment manuals. There are a number of exemplary features. First, the patient population is defined by *DSM* criteria in addition to specific problematic behaviors, that is, parasuicidal behaviors. Second, the treatment manual was generated in the context of clinical research. The treatment was developed in the context of operationalization and discovery, that is, how to operationalize a treatment for research that fits the pathology of the patients who are selected and described by specific criteria. The skills training manual (Linehan, 1993b) notes that it has evolved over a 20 year period, and has been tested on over 100 clients. The treatment generated in this context has been used in diverse treatment settings to teach therapists of various levels of expertise and training to address BPD patients. This process of teaching the treatment to therapists of multiple levels of competence and training can foster the articulation of the treatment and enrich its adaptability to community settings. This treatment has a duration of one year or more because it is addressed to a very difficult and seriously disturbed group of individuals, and provides the challenge of extending treatment manuals beyond brief treatments. A most important and practical addition to the book is an accompanying workbook with work sheets for therapist and patients.

To put DBT in context one could compare it to other cognitive approaches to the personality disorders (Beck & Freeman, 1990), to an interpersonal treatment delivered in a group format (Marziali & Munroe-Blum, 1994), to a modified psychodynamic treatment (Clarkin, Yeomans, & Kernberg, in press), and to a supportive treatment (Rockland, 1992) for these patients.

3.09.5.3 Psychosocial Treatment For Bipolar Disorder

Capitalizing on their years of clinical research with patients with schizophrenia and bipolar disorder and their families, Miklowitz and Goldstein (1997) have articulated a treatment manual for patients with bipolar disorder and their spouses or family members.

The rationale for the treatment is threefold. An episode of bipolar disorder in a family member affects not only the patient but the entire family. Thus, with each episode of the disorder (and this is a recurring disorder) there is a crisis and disorganization of the entire family. Thus, the family treatment is an attempt to assist the patient and family to cooperatively deal with this chronic illness. Finally, both in the rationale for the treatment and in the psychoeducational component of the treatment, the authors espouse a vulnerability–stress model of the disorder, which would argue for a treatment that reduces patient and family stress.

The therapeutic stance of the clinician conducting this treatment is explicated well. The clinician is encouraged to be approachable, open and emotionally accessible. It is recommended that the clinician develop a "Socratic dialogue" with the family in providing information, and develop a give and take between all parties. Although the treatment has a major psychoeducational component, the clinician is encouraged to be ready to explore the emotional impact of the information, and not to be simply a classroom teacher. In dealing with both difficult news (e.g., you have a life-long illness) and tasks (e.g., problem solving helps reduce stress), the clinician is encouraged to be reasonably optimistic about what the patient and family can accomplish if they give the treatment and its methods a chance to succeed.

The treatment is described in the manual in terms of three major phases or modules that have a clear sequence to them. The first phase of psychoeducation is intended to provide the patient and family with information that gives them an understanding of bipolar disorder and its treatment. It is hoped that this information may have practical value in informing the patient and family as to things they should do (e.g., medication compliance) that assists in managing the illness. The second phase or module of treatment is a communication enhancement training phase. The aim of this phase is to improve and/or enhance the patient and family's ability to communicate with one another clearly and effectively, especially in the face of the stress-producing disorder. The third or problem-solving phase provides the patient and family with training in effective conflict resolution. Prior to extensive explanation of the three phases of the treatment is information on connecting and engaging with patient and family, and conducting initial assessment. The authors also examine carefully patient and family criteria for admission to this family intervention. Interestingly, they do not rule out patients who are not medication compliant at present, nor those who are currently abusing substances, both common situations in the history of many bipolar patients. This is an articulated treatment broad enough to encourage those patients who are on the cusp of what is actually treatable.

This manual raises directly the issue of patient medication compliance and the combination of psychosocial treatment with medication treatment. Bipolar patients are responsive to certain medications, and medication treatment is considered a necessity. Thus, this treatment manual provides psychoeducation around the need for medication, and encourages patient and family agreement about continuous medication compliance.

This treatment manual can be placed in context by comparing it to a treatment for bipolar disorder in the individual treatment format (Basco & Rush, 1996). This manual should also be seen in the context of quite similar treatments for patients with schizophrenia and their families (such as those articulated by Anderson, Reiss, & Hogarty, 1986; Bellack, Mueser, Gingerich, & Agresta, 1997; Falloon, Boyd, & McGill, 1984).

3.09.5.4 Other Treatment Manuals

Beyond the three manuals chosen for illustration, it is interesting to note the patient diagnoses and problem areas that are currently addressed by a manual describing a psychotherapy. As noted in Table 1, there are treatments described in manuals for major Axis I disorders such as schizophrenia, bipolar disorder, major depression, anxiety disorders, substance abuse and eating disorders. Treatments for Axis II disorders are less well developed and manualized, except for BPD. Problem areas such as marital discord, sexual dysfunction and problematic emotion expression are also addressed. Further development will probably come in addressing common comorbid conditions. In addition, there has been more funding and research for brief treatments with a cognitive-behavioral orientation. Longer treatments, maintenance treatment for chronic conditions, and the development of manuals for strategies and techniques other than cognitive-behavioral ones can be expected in the future. With the heavy incursion of managed care with its emphasis on cost-cutting, more development of treatments delivered in a group format may also be seen.

3.09.6 ADVANTAGES OF TREATMENT MANUALS

It is clear that the introduction of the treatment manual has had tremendous impact. The very process of writing a treatment manual forces the author to think through and articulate the details of the treatment that may not have been specified before. In this way, the era of the written treatment manual has fostered fuller articulation of treatment by the treatment originators, and furthered the ability of teachers of the treatment to explicate the treatment to trainees.

Manuals provide an operationalized statement of the treatment being delivered or researched so that all can examine it for its foci and procedures. It cannot be assumed that the treatment described in the written manual was delivered as described by the therapists in the study or in a treatment delivery system. This gap between the manual and the treatment as delivered highlights the need for rating scales to assess the faithful (i.e., adherent and competent) delivery of the treatment as described in the manual.

Manuals provide a training tool for clinical research and for clinical practice. It has been noted (Chambless, 1996; Moras, 1993) that students learn treatment approaches much more quickly from their systematic depiction in manuals than through supervision alone.

3.09.7 POTENTIAL DISADVANTAGES AND LIMITATIONS OF TREATMENT MANUALS

Dobson and Shaw (1988) have noted six disadvantages of treatment manuals: (i) the

Table 1 Representative treatment manuals.

Disorder/problem area	*Reference*
Panic	Barlow and Cerny (1988)
Obsessive-compulsive disorder	Steketee (1993) Turner and Beidel (1988)
PTSD	Foy (1992)
Depression	Beck, Rush, Shaw, and Emery (1979) Klerman, Weissman, Rounsaville, and Chevron (1984)
Bipolar	Miklowitz and Goldstein (1997) Basco and Rush (1996)
Schizophrenia	Bellack (1997) Anderson, Reiss, and Hogarty (1986) Falloon, Boyd, and McGill (1984)
Borderline	Linehan (1993a, 1993b) Clarkin, Yeomans, and Kernberg (in press) Marzialli and Munroe-Blum (1994) Rockland (1992)
Marital discord	Baucom and Epstein (1990)
Sexual dysfunction	Wincze and Carey (1991)
Alcohol abuse	Sobell and Sobell (1993)
Binge eating	Fairburn, Marcus, and Wilson (1993)
Problematic emotion schemas	Greenberg, Rice, and Elliott (1993)

inability to assess the effects of therapists' variables, (ii) the diminished ability to study the therapy process, (iii) a focus on treatment fidelity rather than on competence, (iv) the increased expense for research, (v) the over-researching of older and more codified therapy procedures; and (vi) the promotion of schools of psychotherapy. It is important to distinguish the limitations of the manuals as they have been developed up to now, from the limitations of manuals as they can be if the field continues to improve them. There is no reason why many of the concerns listed above cannot be incorporated into future manuals. Some are already being included, such as therapist variables, process, and competence.

Probably the most extensively debated issue around manuals is the issue of therapist flexibility in the execution of the treatment. Some would argue that a manual seriously curtails the flexibility and creativity of the therapist, thus potentially eliminating the therapeutic effectiveness of talented therapists. This is the type of issue that, unless infused with data, could be debated intensely for a very long time. Jacobson et al. (1989) compared two versions of behavioral marital therapy, one inflexible and the other in which therapists had flexibility. The outcome of the flexibility condition was no better than that of the inflexible condition. There was a trend for less relapse at follow-up in the flexibly treated couples.

Wilson (1996) notes that allowing therapists to pursue their own somewhat ideographic case formulations rather than following validated treatment in manuals on average might reduce effectiveness rather than enhance it (Schulte, Kuenzel, Pepping, & Schulte-Bahrenberg, 1992). When research therapy programs are transferred to a clinic setting, there tends to be an increase in effectiveness (Weisz, Donenberg, Han, & Weiss, 1995). A number of studies show that greater adherence to the psychotherapy protocol predicts better outcome (Frank et al., 1991; Luborsky, McLellan, Woody, O'Brien, & Auerbach, 1985).

However, it has been pointed out that adherence to manualized treatments may lead to certain disruptive therapists' behaviors (Henry, Schacht, Strupp, Butler, & Binder, 1993). There is the possibility that therapists delivering a manualized treatment, especially those that are just learning the treatment and adhering with concentration, will deliver it with strict adherence but without competence. At its extreme, it might be argued that the use of the treatment manual may de-skill the therapist and interfere with therapist competence. In fact, rigid adherence could lead to poor therapy, and a mockery of what the treatment is intended to be.

In a study reported by Castonguay, Goldfried, Wiser, Raue, and Hayes (1996), when therapists were engaged in an abstract cognitive intervention the outcome appeared worse. However, those interventions were related to a bad outcome only when practiced to the detriment of the therapeutic alliance.

Jacobson and Hollon (1996) point out that the measurement of therapist adherence to treatments as manualized has received much attention, but the measurement of competence is at a rudimentary stage. The manual should provide instruments that have been shown to reliably assess therapist adherence to the treatment as described in the manual, and competence in the delivery of that treatment. Many books that proport to be treatment manuals do not include such instrumentation. The instruments are helpful in further specifying necessary therapist behaviors, and may be useful to the clinical supervisor.

A major limitation of manual development is the number of patients who appear for treatment who are not addressed in the existing treatment manuals. Some of this is due to the research focus of manuals, that is, the need for clinical research to have a narrowly and carefully defined patient population for whom the treatment is intended. Unfortunately, many if not most patients who appear in clinical settings would not fit into a research protocol because of comorbidity, not quite meeting criteria for a specific diagnosis (e.g., personality disorder NOS, not otherwise specified). Some of this is simply due to the newness of manuals and the little time there has been in their development.

3.09.8 USE AND ROLE OF TREATMENT MANUALS

Treatment manuals have had a brief but exciting and productive history. There is an extreme position that something as complex, unique and creative as a psychotherapy between two individuals cannot be manualized. It is thought that experience indicates that this is an extreme view, and that manualization has been useful to the field. The issues are how best to utilize a manual and what is the role of the manual in clinical research and training?

Whether in the early stages of clinical research or in a clinical training program, the treatment manual can serve as a tool to be used by the expert therapist who is teaching the treatment. It has been noted that the manual enables the student to learn the treatment more quickly than with supervision alone (Chambless, 1996; Moras, 1993). It is our experience, however, that the manual has a place in the teaching toolbox, but is less important than supervision and watching experts doing the treatment on videotape. The manual provides a conceptual overview of the treatment. Videotapes provide a visual demonstration of the manual being applied to a particular patient, in the style of a particular therapist model. Supervision allows the expert to help the trainee apply the manual to a specific patient in a specific treatment, which will always produce some situations that are not exactly covered in the written manual.

It is a naive and false assumption that a clinician can simply read a treatment manual, and thereby be enabled to deliver the treatment with skill. It is interesting that the authors of a treatment manual (Turner & Beidel, 1988) felt the need to argue that just reading a manual is not enough to train an individual in the competent delivery of the treatment in question to the individual patient.

3.09.8.1 Efficacy to Effectiveness

There is currently much discussion about the need to extend clinical trials of empirically validated treatments, that is treatments that have been shown to be efficacious, to larger, community samples with average therapists (effectiveness research). It may be that treatment manuals can play an important role in this work. Indeed, the ultimate value of a treatment manual and the training system within which it operates will be the result that clinicians can perform the treatment with adherence and competence. Manuals that are most useful will contain scales developed to measure therapist adherence and competence, and most current manuals are lacking this feature.

3.09.9 SUMMARY

Although treatment manuals have been effective in specifying psychotherapy for clinical research, the step from clinical efficacy studies to demonstration of clinical effectiveness of the treatments that have been manualized is still lacking. This is an issue of knowledge transfer. That is, given the demonstration that a specific therapy (that has been manualized) has shown clinical efficacy in randomized clinical trials with a homogeneous patient population and with carefully selected therapists, can this treatment also show clinical benefits when transferred to a setting in which the patients are less homogeneous and the therapists are those who are working at the local level? The written treatment manual may play a role in this transfer of expertise from a small, clinical research group to a larger group of therapists. However, this step has yet to be demonstrated. Thus, the test of a manual, and the entire teaching package within which it resides, is to demonstrate that a wider group of therapists can do the treatment with adherence and competence.

It is interesting to speculate about the future of training in psychotherapy given the advances in the field, including the generation of treatment manuals. For sake of argument, we indicated that there are some 33 manuals for 13 disorders for which there are efficacious treatments in the field of clinical psychology. Should these manuals form the basis of the training in psychotherapy of future psychologists? Or can one generate principles of treatment out of these disparate treatments for various disorders and conditions, and teach these principles? There are obvious redundancies across the manuals, and one could imagine a supermanual with branching points for various disorders. For example, most manuals have an assessment phase, followed by the phase of making an alliance with the patient and describing the characteristics of the treatment to follow, with some indication of the roles of patient and therapist. These are obvious skills that a therapist must learn, with nuances depending upon the patient and the disorder in question. For those manuals that are cognitive behavioral in strategies and techniques, there seems to be great redundancy in terms of the selected finite number of techniques that are used.

What is missing is the assessment of the patient in which there is no indication of what the problem is, the diagnosis, or which manual or manuals to use for treatment. Each manual seems to presume that clinicians can properly identify patients for that manual. Unfortunately, one cannot train a psychologist to treat only the 13 disorders/problem areas in the list, as many patients suffer from other conditions not covered. To compound things even further, many patients (if not most, depending on the clinical setting) do not suffer from just one condition, but from several.

Our own approach to training is the one implied by Roth, Fonagy, Parry, Target, and Woods (1996), with emphasis on the initial assessment of the patient with specific clinical hypotheses about the situation, proceeding to the most relevant treatment that has been validated to various degrees. This is a less black-and-white world of the empirically supported treatment approach and more related to the complex condition we call clinical work. Treatment manuals will play some role in this process, but it might be less than the quantum leap suggested by Luborsky and DeRubeis (1984). In our experience, trainees read the treatment manual if they must, but they look forward to seeing experts do the treatment on videotape, and they see supervision from an expert in the treatment as a matter of course.

3.09.10 REFERENCES

Anderson, C. M., Reiss, D. J., & Hogarty, G. E. (1986). *Schizophrenia and the family*. New York: Guilford Press.

Barlow, D. H., & Cerny, J. A. (1988). *Psychological treatment of panic*. New York: Guilford Press.

Basco, M. R., & Rush, A. J. (1996). *Cognitive-behavioral therapy for bipolar disorder*. New York: Guilford Press.

Baucom, D. H., & Epstein, N. (1990). *Cognitive-behavioral marital therapy*. New York: Brunner/Mazel.

Beck, A. T., & Freeman, A. M. (1990). *Cognitive therapy of personality disorders*. New York: Guilford Press.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.

Beck, J. S. (1995). *Cognitive therapy: Basics and beyond*. New York: Guilford Press.

Bellack, A. S., Mueser, K. T., Gingerich, S., & Agresta, J. (1997). *Social skills training for schizophrenia: A step-by-step guide*. New York: Guilford Press.

Carroll, K. M., Rounsaville, B. J., & Gawin, F. H. (1991). A comparative trial of psychotherapies for ambulatory cocaine abusers: Relapse prevention and interpersonal psychotherapy. *American Journal of Drug and Alcohol Abuse, 17,* 229–247.

Castonguay, L. G., Goldfried, M. R., Wiser, S., Raue, P. J., & Hayes, A. M. (1996). Predicting the effect of cognitive therapy for depression: A study of unique and common factors. *Journal of Consulting and Clinical Psychology, 64,* 497–504.

Chambless, D. L. (1996). In defense of dissemination of empirically supported psychological interventions. *Clinical Psychology: Science and Practice, 3*(3), 230–235.

Chambless, D. L., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Crits-Christoph, P., Daiuto, A., DeRubeis, R., Detweiler, J., Haaga, D. A. F., Johnson, S. B., McCurry, S., Mueser, K. T., Pope, K. S., Sanderson, W. C., Shoham, V., Stickle, T., Williams, D. A., & Woody, S. R. (1998). Update on empirically validate therapies. II. (1998). *Clinical Psychologist, 51,* 3–13.

Clarkin, J. F., & Abrams, R. (1998). Management of personality disorders in the context of mood and anxiety disorders. In A. J. Rush (Ed.), *Mood and anxiety disorders* (pp. 224–235). Philadelphia: Current Science.

Clarkin, J. F., Yeomans, F., & Kernberg, O. F. (in press). *Psychodynamic psychotherapy of borderline personality organization: A treatment manual*. New York: Wiley.

Dobson, K. S., & Shaw, B. F. (1988). The use of treatment manuals in cognitive therapy: Experience and issues. *Journal of Consulting and Clinical Psychology, 56,* 673–680.

Ehlers, C. L., Frank E., & Kupfer, D. J. (1988). Social zeitgebers and biological rhythms: A unified approach to understanding the etiology of depression. *Archives of General Psychiatry, 45,* 948–952.

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., Glass, D. R., Pilkonis, P. A., Leber, W. R., Docherty, J. P., Fiester, S. J., & Parloff, M. B. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Archives of General Psychiatry, 46,* 971–982.

Fairburn, C. G., Jones, R., Peveler, R. C., Carr, S. J., Solomon, R. A., O'Connor, M. E., Burton, J., & Hope, R. A. (1991). Three psychological treatments for bulimia nervosa: A comparative trial. *Archives of General Psychiatry, 48,* 463–469.

Fairburn, C. G., Marcus, M. D. & Wilson, G. T. (1993). Cognitive-behavioral therapy for binge eating and bulimia nervosa: A comprehensive treatment manual. In C. G. Fairburn & G. T. Wilson (Eds.), *Binge eating: Nature, assessment and treatment* (pp. 361–404). New York: Guilford Press.

Falloon, I. R. H., Boyd, J. L., & McGill, C. W. (1984). *Family care of schizophrenia*. New York: Guilford Press.

Foy, D. W. (Ed.) (1992). *Treating PTSD: Cognitive-behavioral strategies. Treatment manuals for practitioners*. New York: Guilford Press.

Frank, E., Kupfer, D. J., Perel, J. M., Cornes, C., Jarrett, D. B., Mallinger, A. G., Thase, M. E., McEachran, A. B., & Grochociniski, V. J. (1990). Three-year outcomes for maintenance therapies in recurrent depression. *Archives of General Psychiatry, 47,* 1093–1099.

Frank, E., Kupfer, D. J., Wagner, E. F., McEachran, A. B., & Cornes, C. (1991). Efficacy of interpersonal psychotherapy as a maintenance treatment of recurrent depression. *Archives of General Psychiatry, 48,* 1053–1059.

Greenberg, L. S., Rice, L. N., & Elliott, R. K. (1993). *Facilitating emotional change: The moment-by-moment process*. New York: Guilford Press.

Henry, W. P., Schacht, T. E., Strupp, H. H., Butler, S. F., & Binder, J. L. (1993). Effects of training in time-limited dynamic psychotherapy: Mediators of therapists responses to training. *Journal of Consulting and Clinical Psychology, 61,* 441–447.

Jacobson, N. S., & Hollon, S. D. (1996). Prospects for future comparisons between drugs and psychotherapy: Lessons from the CBT-versus-pharmacotherapy exchange. *Journal of Consulting and Clinical Psychology, 64,* 104–108.

Jacobson, N. S., Schmaling, K. B., Holtzworth-Munroe, A., Katt, J. L., Wood, L. F., & Follette, V. M. (1989). Research-structured vs. clinically flexible versions of social learning-based marital therapy. *Behaviour Research and Therapy, 27,* 173–180.

Kazdin, A. E. (1991). Treatment research: The investigation and evaluation of psychotherapy. In M. Hersen, A. E. Kazdin, & A. S. Bellack (Eds.), *The clinical psychology handbook* (2nd ed., pp. 293–312). New York: Pergamon.

Klerman, G. L., DiMascio, A., Weissman, M., Prusoff, B., Paykel, E. S. (1974). Treatment of depression by drugs and psychotherapy. *American Journal of Psychiatry, 131,* 186–191.

Klerman, G. L., Weissman, M. M., Rounsaville, B. J., & Chevron, E. S. (1984). *Interpersonal psychotherapy of depression*. New York: Basic Books.

Linehan, M. M. (1993a). *Cognitive-behavioral treatment of borderline personality disorder*. New York: Guilford Press.

Linehan, M. M. (1993b). *Skills training manual for treating borderline personality disorder*. New York: Guilford Press.

Luborsky, L., & DeRubeis, R. J. (1984). The use of psychotherapy treatment manuals: A small revolution in psychotherapy research style. *Clinical Psychology Review, 4,* 5–14.

Luborsky, L., McLellan, A. T., Woody, G. E., O'Brien, C. P., & Auerbach, A. (1985). Therapists success and its determinants. *Archives of General Psychiatry, 42,* 602–611.

Marziali, E., & Munroe-Blum, H. (1994). *Interpersonal group psychotherapy for borderline personality disorder*. New York: Basic Books.

Miklowitz, D. J., & Goldstein, M. J. (1997). *Bipolar disorders: A family-focused treatment approach*. New York: Guilford Press.

Moras, K. (1993). The use of treatment manuals to train psychotherapists: Observations and recommendations. *Psychotherapy, 30,* 581–586.

Moras, K. (1995, January). Behavioral therapy development program workshop (Draft 2, 3/24/95). National Institute on Drug Abuse, Washington, DC.

Mufson, L., Moreau, D., Weissman, M. M., & Klerman, G. L. (Eds.) (1993). *Interpersonal psychotherapy for depressed adolescents*. New York: Guilford Press.

Reynolds, C. F., Frank, E., Perel, J. M., Imber, S. D., Cornes, C., Morycz. R. K., Mazumdar, S., Miller, M., Pollock, B. G., Rifai, A. H., Stack, J. A., George, C. J., Houck, P. R., & Kupfer, D. J. (1992). Combined pharmacotherapy and psychotherapy in the acute and continuation treatment of elderly patients with recurrent major depression: A preliminary report. *American Journal of Psychiatry, 149,* 1687–1692.

Rockland, L. H. (1992). *Supportive therapy for borderline patients: A psychodynamic approach*. New York: Guilford Press.

Roth, A., Fonagy, P., Parry, G., Target, M., & Woods, R. (1996). *What works for whom? A critical review of psychotherapy research*. New York: Guilford Press.

Rounsaville, B. J., Glazer, W., Wilber, C. H., Weissman, M. M., & Kleber, H. D. (1983). Short-term interpersonal psychotherapy in methadone-maintained opiate addicts. *Archives of General Psychiatry, 40,* 629–636.

Schulte, D., Kuenzel, R., Pepping, G., & Schulte-Bahrenberg, T. (1992). Tailor-made versus standardized therapy of phobic patients. *Advances in Behaviour Research and Therapy, 14,* 67–92.

Sobell, M. B., & Sobell, L. C. (1993). *Problem drinkers: Guided self-change treatment*. New York: Guilford Press.

Steketee, G. (1993). *Treatment of obsessive compulsive disorder*. New York: Guilford Press.

Talley, P. F., Strupp, H. H., & Butler, S. F. (Eds.) (1994). *Psychotherapy research and practice: Bridging the gap*. New York: Basic Books.

Turner, S. M., & Beidel, D. C. (1988). *Treating obsessive-compulsive disorder*. Oxford, UK: Pergamon.

Weissman, M. M. (1995). *Mastering Depression: A patient's guide to interpersonal psychotherapy*. San Antonio, TX: Psychological Corporation.

Weissman, M. M., & Markowitz, J. C. (1994). Interpersonal psychotherapy: Current status. *Archives of General Psychiatry, 51,* 599–606.

Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63,* 688–701.

Wilson, G. T. (1996). Empirically validated treatments: Realities and resistance. *Clinical Psychology, 3,* 241–244.

Wincz. J. P., & Carey, M. P. (1991). *Sexual dysfunction: A guide for assessment and treatment*. New York: Guilford Press.

Woody, S. R., & Sanderson, W. C. (1998). Manuals for empirically supported treatments: 1998 update. *Clinical Psychologist, 51,* 17–21.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

3.10

Internal and External Validity of Intervention Studies

KARLA MORAS

University of Pennsylvania, Philadelphia, PA, USA

3.10.1 INTRODUCTION 202
3.10.2 DEFINITIONS 202
  3.10.2.1 Overview of IV, EV, CV, and SCV 202
  3.10.2.2 Internal Validity 203
  3.10.2.3 External Validity 203
  3.10.2.4 Construct Validity 203
  3.10.2.5 Statistical Conclusion Validity 204
3.10.3 COMMON THREATS TO IV, EV, CV, AND SCV 204
3.10.4 EXPERIMENTAL DESIGNS AND METHODS THAT ENHANCE IV, EV, CV, AND SCV 205
  3.10.4.1 IV Methods and Designs 205
    3.10.4.1.1 Random assignment 205
  3.10.4.2 Untreated or Placebo-treated Control Group 208
  3.10.4.3 Specification of Interventions 209
  3.10.4.4 Process Research Methods 209
  3.10.4.5 EV Methods 210
    3.10.4.5.1 Naturalistic treatment settings 210
  3.10.4.6 Inclusive Subject Selection 211
    3.10.4.6.1 Staff therapists 212
  3.10.4.7 CV Methods 212
    3.10.4.7.1 Specification of interventions: treatment manuals 212
    3.10.4.7.2 Specification of interventions: therapist adherence measures 214
    3.10.4.7.3 Distinctiveness of interventions 215
    3.10.4.7.4 Adjunctive therapies 215
    3.10.4.7.5 Assessment procedures 215
  3.10.4.8 SCV Methods 216
    3.10.4.8.1 Statistical power analysis 216
    3.10.4.8.2 Controlling for the type 1 error rate 216
    3.10.4.8.3 Testing assumptions of statistical tests 218
3.10.5 FROM IV VS. EV TO IV + EV IN MENTAL HEALTH INTERVENTION RESEARCH 218
  3.10.5.1 Strategies for Efficacy + Effectiveness (IV + EV) Intervention Studies 218
    3.10.5.1.1 The conventional "phase" strategy 219
    3.10.5.1.2 A dimensional adaptation of the phase model 219
    3.10.5.1.3 Stakeholder's model to guide data selection 219
    3.10.5.1.4 Mediator's model 219
  3.10.5.2 Examples of Efficacy + Effectiveness Intervention Studies 220
    3.10.5.2.1 Schulberg et al. (1995) 220
    3.10.5.2.2 Drake et al. (1996) 220
3.10.6 A CONCLUDING OBSERVATION 221
3.10.7 REFERENCES 222

3.10.1 INTRODUCTION

The concepts of internal and external validity (IV and EV) were introduced by Campbell and Stanley in 1963. IV and EV concern the validity of inferences that can be drawn from an intervention study, given its design and methods. The concepts are used to assess the extent to which a study's outcome findings can be confidently: (i) interpreted as evidence for hypothesized causal relationships between interventions and outcomes (IV), and (ii) assumed to generalize beyond the study situation (EV). IV and EV are conceptual tools that guide deductive (IV) and inductive (EV) thinking about the impact of a study's design and methods on the validity of the conclusions that can be drawn from it. The concepts are not only of academic interest. Evaluation of a study's IV and EV is a logical, systematic way to judge if its outcome findings provide compelling evidence that an intervention merits implementation in public sector settings.

This chapter is written at a unique time in the history of IV and EV. The concepts have been at the forefront of a contemporary, often contentious debate in the USA about the public health value of alternative methods and designs for intervention research (e.g., Goldfried Wolfe, 1998; Hoagwood, Hibbs, Bren, & Jensen, 1995; Jacobson & Christensen, 1996; Lebowitz & Rudorfer, 1998; Mintz, Drake, & Crits-Christoph, 1996; Newman & Tejeda, 1996; Seligman, 1996; Wells & Sturm, 1996). The alternatives are referred to as "efficacy" vs. "effectiveness" studies. In current parlance, efficacy studies have high IV due to designs and methods that reflect a priority on drawing causal conclusions about the relationship between interventions and outcomes (e.g., Elkin et al., 1989). Effectiveness studies have high EV due to designs and methods that reflect the priority of obtaining findings that can be assumed to generalize to nonresearch intervention settings and clinic populations (e.g., Speer, 1994). Typically, effectiveness studies are done to examine the effects of interventions with demonstrated efficacy in high IV studies, when the interventions are used in standard community treatment settings. A related term, "clinical utility research," has started to appear (Beutler & Howard, 1998). The types of research questions connoted by clinical utility are broader than those connoted by effectiveness research (Howard, Moras, Brill, Martinovich, & Lutz, 1996; Kopta, Howard, Lowry, & Beutler, 1994; Lueger, 1998).

The efficacy vs. effectiveness debate is prominent in mental health intervention research. IV and EV often are viewed as competing rather than complimentary research aims (Kazdin, 1994; Roth & Fonagy, 1996). An alternative view is that EV questions only can be asked about a study's findings after the study's IV has been established (Flick, 1988; Hoagwood et al., 1995). A recent trend is to encourage investigators to design studies that can optimize both IV and EV (Clarke, 1995; Hoagwood et al., 1995). The topic is pursued later in the chapter.

This chapter is intended to provide a relatively concise, simplified explication of IV and EV that will be useful to neophyte intervention researchers and to consumers of intervention research. The main aims are to enhance the reader's sophistication as consumer of intervention research, and ability to use the concepts of IV and EV to design intervention studies. IV and EV are discussed and illustrated mainly from the perspective of research on interventions of a certain type: psychological therapies for mental health problems (Bergin & Garfield, 1971, 1994; Garfield & Bergin, 1978, 1986).The topics covered are: (i) definitions of IV and EV and of two newer, closely related concepts, construct validity (CV) and statistical conclusion validity (SCV); (ii) threats to IV, EV, CV, and SCV; (iii) designs and methods that are commonly used in mental health intervention research to enhance IV, EV, CV, and SCV; and (iv) suggested strategies from the efficacy vs. effectiveness debate to optimize the scientific validity, generalizability, and public health value of intervention studies. Finally, two intervention studies that were designed to meet both IV and EV aims are used to illustrate application of the concepts.

3.10.2 DEFINITIONS

3.10.2.1 Overview of IV, EV, CV, and SCV

Kazdin (1994) provided a concise overview of IV, EV, CV, and SCV, and of common threats to each type of validity (Table 1). The discussion of the four concepts in this chapter is based on Cook and Campbell's (1979) conceptualizations, as is Kazdin's table. Campbell and Stanley's (1963, 1966) original conceptualizations of IV and EV were extended and slightly revised by Cook and Campbell in 1979. The arguments and philosophy of science assumptions that underpin Cook and Campbell's (1979) perspective have been challenged (Cronbach, 1982). A key criticism is that they overemphasized the value of high IV studies in a way that was inconsistent with EV (generalizability) trade-offs often associated with such studies (Shadish, 1995). Cook and Campbell's views were adopted in this chapter because, challenges notwithstanding, they have had a major influence on mental health intervention research since the late 1960s.

Table 1 Types of experimental validity, questions they address, and their threats to drawing valid inferences.

Type of validity	*Questions addressed*	*Threats to validity*
Internal validity	To what extent can the intervention, rather than extraneous influences, be considered to account for the results, changes, or group differences?	Changes due to influences other than the experimental conditions, such as events (history) or processes (maturation) within the individual, repeated testing, statistical regression, and differential loss of subjects
External validity	To what extent can the results be generalized or extended to persons, settings, times, measures, and characteristics other than those in this particular experimental arrangement?	Possible limitations on the generality of the findings because of characteristics of the sample; therapists; or conditions, context, or setting of the study
Construct validity	Given that the intervention was responsible for change, what specific aspects of the intervention or arrangement were the causal agents; that is, what is the conceptual basis (construct) underlying the effect?	Alternative interpretations that could explain the effects of the intervention, that is, the conceptual basis of the findings, such as attention and contact with the subject, expectancies of subjects or experimenters, cues of the experiment
Statistical conclusion validity	To what extent is a relation shown, demonstrated, or evident, and how well can the investigation detect effects if they exist?	Any factor related to the quantitative evaluation that could affect interpretation of the findings, such as low statistical power, variability in the procedures, unreliability of the measurement, inappropriate statistical tests

Source: Kazdin (1994).

A basic premise of Cook and Campbell's definitions of IV and EV is that the term "validity" can only be used in an approximate sense. They would say, for example, that judgments of the validity of the conclusions that can be drawn from a study must always be understood as approximate because, from an epistemological perspective, we can never definitively know what is true, only that which has not been shown to be false.

3.10.2.2 Internal Validity

Simply stated, IV refers to the extent to which causal conclusions can be correctly drawn from a study about the relationship between an independent variable (e.g., a type of therapy) and a dependent variable (e.g., symptom change). The full definition reads: "Internal validity refers to the approximate validity with which we infer that a relationship between two variables is causal or that the absence of a relationship implies the absence of cause" (Cook & Campbell, 1979, p. 37).

3.10.2.3 External Validity

EV refers to the extent to which causal conclusions from a study about a relationship between interventions and outcomes can be assumed to generalize beyond the study's specific features (e.g., the patient sample, the therapists, the measures of outcome, the study setting). The full definition of EV is: " … the approximate validity with which we can infer that the presumed causal relationship can be generalized to and across alternative measures of the cause and effect and across different types of persons, settings, and times" (Cook & Campbell, 1979, p. 37).

3.10.2.4 Construct Validity

In their 1979 update of Campbell and Stanley (1963, 1966), Cook and Campbell highlighted two new concepts that are closely linked to IV and EV: CV and SCV. Both have been incorporated into contemporary thinking about experimental design and threats to IV and EV (e.g., Kazdin, 1994).

Cook and Campbell (1979) focused their discussion of CV on the "putative causes and effects" (i.e., interventions and outcomes) of intervention studies. CV reflects the addition of a concept developed earlier by Cronbach and Meehl (1955) to Cook and Campbell's (1979) model of experimental validity. Simply stated, CV is the goodness of fit between the methods used to operationalize constructs (interventions and outcome variables) and the referent constructs. In other words, CV is the extent to which the methods used to measure and operationalize interventions and outcomes are likely to reflect the constructs that the investigators say they studied (e.g., cognitive therapy and depression). A more precise definition of CV is: " … the possibility that the operations which are meant to represent a particular cause or effect construct can be construed in terms of more than one construct, each of which is stated at the same level of reduction" (p. 59). For example, the more possible it is to construe measures in terms of constructs other than those named by the investigators, the lower the CV.

CV also can be described as "what experimental psychologists are concerned with when they worry about 'confounding'" (Cook & Campbell, 1979, p. 59). An example of a threat to CV is the possibility that the effects of a medication for depression are due largely to the treating psychiatrist's concern for a patient and nonjudgmental reactions to his or her symptoms, rather than to neurochemical effects of the drug. Such alternative interpretations of the therapeutic cause of outcomes in medication intervention studies lead to: (i) the inclusion of pill placebo control conditions, (ii) randomized assignment to either active drug or pill placebo, and (iii) the use of double-blind procedures, that is, neither treater nor patient knows if the patient is receiving active drug or placebo.

Cook and Campbell (1979) link CV to EV. They say that generalizability is the essence of both. However, an integral aspect of CV is the adequacy with which the central variables of an intervention study (the interventions and outcomes) are operationalized. Thus, CV is also necessarily linked to IV: the CV of the methods used to measure and operationalize interventions and outcomes affects the validity of any causal conclusions drawn about a relationship between the designated interventions and outcomes.

3.10.2.5 Statistical Conclusion Validity

SCV is described by Cook and Campbell (1979) as a component of IV. SCV refers to the effects of the statistical methods used to analyze study data on the validity of the conclusions about a relationship (i.e., covariation) between interventions and outcomes. SCV concerns "particular reasons why we can draw false conclusions about covariates" (Cook & Campbell, 1979, p. 37). It "is concerned … with sources of random error and with the appropriate use of statistics and statistical tests" (p. 80). For example, a determinant of SCV is whether the assumptions of the statistical test used to analyze a set of outcome data were met by the data.

3.10.3 COMMON THREATS TO IV, EV, CV, AND SCV

A study's IV, EV, CV, and SCV are determined by its design and methods, by the psychometric adequacy of its measures and operationalizations of central constructs, and by the appropriateness of the statistical tests used to analyze the data. All must be assessed to evaluate a study's IV, EV, CV, and SCV. "Design" refers to elements of a study's construction (the situation into which subjects are placed) that determine the probability that causal hypotheses about a relationship between the independent and dependent variables can validly be tested. For example, a design feature is the inclusion of a no-treatment control group of some type (e.g., pill placebo) in addition to the treatment group. A no-therapy group provides outcome data to compare with the outcomes of the treated group. The comparison allows examination of the possibility that changes associated with the treatment also occur without it and, thus, cannot be causally attributed to it.

"Methods" refers to a wide variety of procedures that are used to implement study designs, such as random assignment of subjects to each intervention group included in a study. A study's methods and design together determine the degree to which any relationship found between interventions and outcomes can validly be attributed to the interventions rather than to something else. In other words, study methods and design determine whether alternative or rival interpretations of findings can be dismissed as improbable. Designs and methods that affect a study's IV, EV, CV, and SCV are discussed in Section 3.10.4.

The use of IV, EV, CV, and SCV to guide study design and critical appraisal of study findings is assisted by knowledge of common threats to each type of validity. Cook and Campbell (1979) discussed several threats to each type. They cautioned, however, that no list is perfect and that theirs was derived from their own research experience and from reading about potential sources of fallacious inferences.

The threats to each type of validity identified by Cook and Campbell are listed in Tables 2–5. The tables also present examples of designs and methods that can offset the various threats ("antidotes"). The interested reader is encouraged to review Cook and Campbell's discussion of threats to IV, EV, CV, and SCV; design and methodological antidotes to the threats; and limitations of common antidotes.

In the following sections, the four types of threats to validity are discussed in turn. In each section, examples of designs and methods are described that are used in contemporary psychological and pharmacological mental health treatment research to enhance that type of validity. Only a sampling of designs and methods relevant to each type of validity is described due to space limitations.

3.10.4 EXPERIMENTAL DESIGNS AND METHODS THAT ENHANCE IV, EV, CV, AND SCV

Each design and method discussed in the sections that follow is limited in the extent to which it can ensure that a particular type of experimental validity is achieved. Some limitations result mainly from the fact that patient-subjects are always free to deviate from research treatment protocols they enter, by dropping out of treatment, for example. Other limitations arise because research methods that can enhance IV can simultaneously reduce EV. For example, using psychotherapists who are experts in conducting a form of therapy can both increase IV and reduce EV.

One point merits emphasis. The experimental validity potential of alternative designs and methods is always contingent on the match between study hypotheses and a study's design and methods. For example, the IV potential of a particular design differs if the main aim of a study is to obtain data that can be interpreted as evidence for a therapy's hypothesized effects vs. to obtain data that can be interpreted as evidence for the comparative efficacy of alternative therapies for the same problem.

3.10.4.1 IV Methods and Designs

Table 2 lists and defines common IV threats. The text that follows elaborates on some of the information in Table 2.

3.10.4.1.1 Random assignment

Random assignment of subjects to all intervention groups in a study design is a *sine qua non* of IV. Broadly defined, random assignment means that a procedure is used to ensure that each new subject has an equal chance to be assigned to every intervention condition in a study. A simple random assignment procedure is the coin flip: heads the subject is assigned to one intervention in a two-intervention study; tails he or she receives the other one. Various procedures are used for random assignment including sophisticated techniques like urn randomization (Wei, 1978). Urn randomization simultaneously helps ensure that subjects are randomly assigned to interventions, and that the subjects in each one are matched on preidentified characteristics (e.g., co-present problems) that might moderate the effects of an intervention (Baron & Kenny, 1986).

Random assignment contributes to IV by helping to ensure that any differences found between interventions can be attributed to the interventions rather than to the subjects who received them. The rationale for randomization is that known and unknown characteristics of subjects that can affect outcomes will be equally distributed across all intervention conditions. Hence, subject features will not systematically affect (bias) the outcomes of any intervention.

Random assignment has limitations as a way to enhance IV. For example, it is not a completely reliable method to ensure that all outcome-relevant, preintervention features of subjects are equally distributed across intervention conditions. Even when randomization is used, by chance some potentially outcome-relevant subject characteristics can be more prevalent in one intervention than another (Collins & Elkin, 1985). This fact can be discovered *post hoc* when subjects in each intervention are found to differ on a characteristic (e.g., marital status) that is also found to relate to outcome. This happened, for example, in the US National Institute of Mental Health's Treatment of Depression Collaborative Research Program study (Elkin et al., 1989).

Attrition (e.g., of subjects, of subjects' outcome data) also limits the IV protection provided by random assignment (Flick, 1988; Howard, Cox, & Saunders, 1990; Howard, Krause, & Orlinsky, 1986). For example, subjects can always drop out of treatment or fail to provide data at all required assessment points. Both are examples of postinclusion attrition in Howard, Krause et al. (1986) terminology. The problems associated with, and types of, attrition have been carefully explicated (Flick, 1988; Howard, Krause et al. 1986). For example, the core IV threat associated with dropout is that the subjects who drop out of each study intervention might differ somehow (e.g., severity of symptoms). This would create differences in the subjects who complete each intervention and simultaneously

Table 2 Internal validity.

Threat	*Description*	*Example antidote*
Historical factors	Events that occur between pre- and postintervention measurement could be responsible for the intervention effects found (e.g., a public information campaign to reduce drug use in teenagers coincides with an adolescent drug treatment intervention study)	If historical the event does not overlap completely with the study period, the outcome data can be analyzed separately for subject cohorts that received the intervention before the event vs. during it, and results compared to estimate the intervention's effects independent of the event
Maturation	Changes in subjects that naturally occur with the passage of time could be responsible for pre- to postintervention changes in outcomes	Include a no or minimal treatment control condition in the study, and randomize subjects to treatment conditions
Testing	Pre- to postintervention changes on outcome measures could be due to repeated administration of the measures, rather than to change in the outcome variables assessed	Evaluate the impact of repeated testing by administering the outcome measures to a randomly selected subset of subjects only at postintervention. Compare the final scores with those of subjects to whom the measures were repeatedly administered. Use the Solomon Four-Group design (Campbell & Stanley, 1963; Rosenthal & Rosnow, 1991) to examine both the effects of repeated testing and any interaction of testing and interventions
Instrumentation	Pre- to postintervention differences in the way an outcome measure is administered could be responsible for intervention effects (e.g., different diagnosticians administer a structured diagnostic interview at pre- and at post-test to the same subject)	Keep instruments and administration procedures constant throughout a study (e.g., recalibrate laboratory instruments and human assessors during study)
Statistical regression	Change in pre- to postintervention scores include "normal" drift toward more moderate scores whenever an initial score is either higher or lower than the average (mean) tendency of the group	Use the statistical technique of covarying preintervention scores from termination scores in outcome analyses (e.g., analysis of covariance, multiple regression)
Selection	Differences between subjects assigned to a study's intervention conditions could account for outcome differences in the interventions	Randomly assign subjects to intervention conditions
Attrition or mortality	Differences between subjects who complete each study intervention (e.g., due to dropout from the interventions) could be responsible for differences in intervention outcomes	Collect outcome data at every planned assessment point from *all* subjects (including dropouts and those withdrawn due to deterioration) and conduct outcome analyses on the 'intention-to-treat' sample (see text)
Ambiguity about the direction of causal influence	The study design does not make it possible to determine if A causes B or B causes A, although the hypotheses assume undirectional causality (e.g., in a correlational study it cannot be determined whether improvement in the therapeutic alliance enhanced outcome or if a subject's improvement enhanced the therapeutic alliance)	Do not use cross-sectional designs and standard correlational statistics if causal conclusions are desired. When theory and empirical findings exist to support causal hypotheses, one alternative is to develop and test causal models on data that lack experimental controls needed to infer causality (e.g., Kenny, 1979)

Table 2 (continued)

Threat	*Description*	*Example antidote*
Diffusion or imitation of treatments	The putative active elements of a study intervention are somehow disseminated to subjects in other intervention and/or control conditions (e.g., subjects who receive a treatment for drug addiction give copies of their treatment workbooks to subjects in the control condition)	When informed consent is administered, emphasize the requirement to refrain from discussing one's treatment with other subjects. Guarantee all subjects the opportunity to receive a study treatment (e.g., at the end of the study)
Compensatory equalization of treatments	Personnel who have control over desirable interventions compromise the study design by somehow making them available to all subjects. For example, in a study in which therapists serve as their own controls (the same therapists administer the interventions and control treatment conditions), therapists have difficulty refraining from using elements of the study's "active" interventions in the therapy sessions of subjects who are in the control condition	Ensure that subjects who do not get the desirable intervention during the study will have access to them after the study. Let subjects know at the beginning of the study that this will happen. Antidotes for the therapist example are to audiotape all therapy sessions and monitor the tapes for therapists' adherence to the intervention conditions. Immediately give therapists feedback when they deviate from adherence requirements. Also, offer all subjects who are randomized to control conditions the opportunity to receive the (putative) active treatments at the end of the study; inform therapists at the study outset that control subjects will have the opportunity to receive an active treatment
Compensatory rivalry by respondents receiving less desirable treatments	Subjects are aware of all intervention conditions in the study, and those in a "standard" condition are motivated to outperform those in "special" intervention conditions. Thus, any differences or lack of differences found between intervention and/or control conditions cannot be attributed to the interventions	This threat applies only to some types of intervention studies. For example, it can occur when a study setting is an intact unit (such as a department) and subjects in the standard condition expect some negative personal effect (e.g., lose their jobs) of study findings. Investigators need to be aware of the potential relevance of this threat to a study and either redesign the project or take actions to reduce its potential influence (e.g., see above antidote for compensatory equalization)
Resentful demoralization of respondents receiving less desirable treatments	This threat is the opposite of the threat of compensatory rivalry. Subjects in a control or other intervention condition underperform relative to their capacity and thereby contribute to artificial differences between the intervention outcomes	Investigators need to be aware of the relevance of this threat to a particular study and either redesign the project or take actions to reduce its potential influence (promise subjects in the less desirable intervention condition the opportunity to receive it at the end of the study)

Statements in the first two columns and some in the third are abstracted from Cook and Campbell (1979).

in those who provide outcome data. This type of threat of attrition to IV has been called the "differential sieve" (Hollon, Shelton, & Loosen, 1991), an apt metaphor. If differential attrition by intervention occurs, any outcome differences could be due to the subjects from whom assessments were obtained rather than to the interventions examined.

One methodological antidote to the IV problems posed by attrition due to dropout is to conduct outcome analyses on the "intention-to-treat" sample (Gillings & Koch, 1991). The intention-to-treat sample can be defined in different ways. For example, one definition is all subjects who were randomized to study interventions. Another definition is all subjects who attended at least one intervention session. The latter type of definition is often used when investigators want to generalize outcome findings only to patients who at least are willing to try an intervention.

Intention-to-treat analyses ideally require outcome assessments of all dropouts (and of all others who discontinued treatment prematurely, such as by being withdrawn by the investigators due to treatment noncompliance or deterioration) at the time(s) when they would have been assessed if they had remained in treatment. The intent-to-treat analytic strategy also contributes to a study's EV by maximizing the probability that the findings can be assumed to generalize to all individuals who would meet the study entry criteria.

Unfortunately, intention-to-treat analyses have limitations with respect to protecting a study's IV. The reader is referred to Flick (1988) and to Howard, Krause et al. (1986) for comprehensive discussions of other methods to evaluate and reduce the effects of attrition on a study's IV and EV.

3.10.4.2 Untreated or Placebo-treated Control Group

The inclusion of a no-treatment group (e.g., wait list) or putative sham treatment (e.g., pill placebo) in a study design is required to validly answer causal questions like "Does an intervention have certain effects?" A no-treatment or placebo condition to which subjects are randomly assigned and which continues for the same duration as a putative active treatment allows several rival explanations for a relationship between a treatment and outcomes to be dismissed. For example, if the outcomes associated with an active treatment are statistically better than the outcomes associated with a control treatment, then the active treatment outcomes cannot be attributed solely to changes that occur naturally with the passage of time, maturation, or "spontaneous remission."

The inclusion of control treatment conditions of some type is logically very compelling as a way to increase an intervention study's IV. Unfortunately, the compelling logic is not mirrored in practice. Reams have been written on problems associated with designing placebo therapies that meet IV goals (e.g., O'Leary & Borkovec, 1978). For example, credibility is one issue (Parloff, 1986). How can a placebo psychotherapy be developed that has no theoretically active therapeutic ingredients but that seems equally credible to subjects as a putative active treatment?

Pill placebo conditions used in psychopharmacology research are often thought to be free of the IV limitations associated with psychotherapy placebos. However, pill placebo conditions are also limited in terms of IV. For example, active medications almost always have side effects. Thus, the placebo pill used should mimic the side effects of the active drug. It also is important that both the subject and the prescribing therapist be blind (uninformed) about who is getting placebo and who is getting active medication. Experienced psychopharmacologists often, probably typically, can tell which subjects are receiving active medication (Greenberg, Bornstein, Greenberg, & Fisher, 1992; Margraf et al., 1991; Ney, Collins, & Spensor, 1986; Rabkin et al., 1986). The failure of the blinding procedure is problematic because therapists' impressions could be associated with different behavior toward placebo and active medication subjects. For example, a therapist might be less encouraging about the possibility of improvement if he or she suspects that a subject is getting placebo rather than medication. Differential therapist behavior toward placebo and active medication subjects could affect outcomes and thereby reduce the IV of the study design.

Another common threat to the IV value of no-treatment and placebo treatment conditions is contamination or diffusion. Putative therapeutic elements of an intervention can somehow infiltrate the control condition. A well-known example of this is the "MRFIT" study conducted in the USA. It was designed to evaluate an educational intervention to reduce the incidence of heart disease in men. The intervention provided information about diet, smoking, and exercise as a way to lower cholesterol and other risks of heart disease (Gotto, 1997; Multiple Risk Factor Trial Interventions Group, 1977). The study design was a randomized, controlled comparison of men who received the intervention with men who received "usual care." During the years that the study was done, information used in the intervention was widely disseminated in the US media, thus threatening the IV value of the usual care control sample.

Another way contamination of a control treatment can occur is via contact between the control subjects and those receiving the active treatments. For example, subjects in active treatments might give control subjects intervention materials (e.g., copies of intervention self-help workbooks). To help prevent this threat to IV, subjects can be required, as part of the study informed consent, to refrain from sharing information about their respective treatments while a study is in progress.

3.10.4.3 Specification of Interventions

A relatively new set of IV methods focuses on detailed specification of the interventions examined in a study. These methods have been adopted widely in psychotherapy research since the 1980s. Their primary aims are to facilitate replication of studies and to ensure that interventions are properly implemented. The methods include the preparation of treatment manuals (e.g., Beck, 1995; Klerman, Weissman, Rounsaville, & Chevron, 1984), for example. Such manuals describe the treatment rationale, the therapeutic techniques to be used, and the circumstances that should prompt therapists to choose between alternative techniques (Lambert & Ogles, 1988). The methods also include the development of therapist adherence measures to be used by observer-judges of therapy session material (e.g., audiotapes) to assess the extent to which treatments were delivered by study therapists as specified in the manual (Waltz, Addis, Koerner, & Jacobson, 1993).

Treatment specification methods are described more fully in Section 3.10.4.7 on CV. They are cross-referenced here to emphasize the fundamental dependence of a study's IV on the CV of its intervention variables. The dependence is essentially the same as that between the psychometric adequacy of the measures and methods used to operationalize independent and dependent variables, and the possibility of conducting a valid test of hypotheses about the relationship between independent and dependent variables (Kraemer & Telch, 1992). The validity of conclusions about hypotheses is necessarily contingent on the adequacy of the operationalization of central study variables or constructs. The preceding point, albeit basic, merits emphasis because it is often ignored or not understood well in mental health intervention research (Kraemer & Telch, 1992).

3.10.4.4 Process Research Methods

"Process" research refers to a highly developed methodology that has been used in adult psychotherapy research since at the 1950s (Greenberg & Pinsof, 1986; Kiesler, 1973; Rice & Greenberg, 1984; Russell, 1987). Described generally, process research is the application of measurement methods to therapy session material, such as videotaped recordings or transcripts of therapy sessions. A typical strategy for process research consists of four basic steps. First, a measure is developed of a theoretically important therapy construct (e.g., therapeutic alliance). The measure is designed to be rated by trained judges, based on observation of some type of therapy session material (e.g., videotapes of sessions). Judges are selected and trained to apply the measure. Psychometric studies are done to evaluate and ensure that judges' ratings on the measure meet standards for inter-rater reliability. A plan for sampling material from therapy sessions for judges to rate is developed. The sampling plan is designed to provide a valid idea of the construct. Judges then rate the selected material. In addition to the foregoing basic steps, it is also important to conduct psychometric evaluations of measure (e.g., the internal consistency reliability of ratings).

Process methods are used to examine a host of research questions (for comprehensive reviews see Orlinsky & Howard, 1986; Orlinsky, Grawe, & Parks, 1994). One of the most valuable uses of process research methods is to test the theories of therapeutic change that are associated with different types of psychotherapy. A standard treatment outcome study is designed to ask questions like "Does a treatment work?" and "Does a therapy work better than an alternative one for the same problem?" Process methods allow the next logical questions to be asked, for example, "Does a therapy work for the reasons it is theoretically posited to?" and "What are the therapeutically active elements of a therapy?" Viewed in terms of Cook and Campbell's (1979) types of experimental validity, process research methods contribute mainly to IV and CV.

Process methods can contribute to a study's IV when they are used to test hypothesized relationships about theoretically causal elements of a therapy (such as specific therapeutic techniques) and targeted outcomes. The precision of the causal conclusions that can be drawn from intervention studies is limited when patients are randomly assigned to treatments, receive treatment, and then outcomes are measured. Even when intervention outcome studies yield positive findings, many fundamental and important questions remain unanswered. Are a therapy's effects mediated by the theoretically hypothesized activities of therapists? Alternatively, are the effects primarily due to nonspecific features of the therapy (e.g., the therapist's empathy and understanding) rather than to its specific techniques (e.g., cognitive restructuring)? (Ilardi & Craighead, 1994; Kazdin, 1979; Strupp & Hadley, 1979). Are all of the recommended therapeutic techniques necessary; are any necessary and sufficient?

Process methods can be used to examine questions of the foregoing type. When used in such ways, they contribute to a study's IV because they help elucidate causality, that is, the extent to which specific elements of a therapeutic intervention can validly be assumed to play a causal role in the outcomes obtained

(Gomes-Schwartz, 1978). Furthermore, when process research methods yield answers to mechanisms of action questions like those stated above, the findings can be used to refine and improve existing treatments, for example, to increase their efficiency and efficacy, and to modify the theories of therapeutic change on which treatments are founded.

Process methods can also contribute to CV. For example, process methods are used both to develop instruments to assess therapists' adherence to manualized treatments (e.g., Evans, Piasecki, Kriss, & Hollon, 1984), and to develop a rating strategy (e.g., plan for sampling from the therapy session material to be rated by judges) to assess therapist adherence for intervention studies (e.g., Hill, O'Grady, & Elkin, 1992). More detail is provided on process research methods that contribute to CV in Section 3.10.4.3.

3.10.4.5 EV Methods

Table 3 presents three types of threats to EV that were highlighted by Cook and Campbell (1979). Each threat is described as an interaction between a study feature and the study treatment(s). This is because Cook and Campell emphasized the difference between generalizing *across* subject populations, settings, etc., vs. generalizing *to* target subject populations, settings. Cook and Campbell chose to focus on the former because the latter requires sophisticated random sampling methods which are rarely used in intervention research.

As stated earlier, EV concerns the extent to which outcome findings validly can be assumed to generalize beyond a study's specific therapists, subjects, setting(s), measures, and point in time when a study was conducted. For about the past eight years in the USA, criticisms have been raised increasingly often about intervention studies with apparent low EV (e.g., Lebowitz & Rudorfer, 1998). The criticisms are part of the efficacy vs. effectiveness debate. For example, it is said that the outcomes of therapies done by study therapists who are trained in a therapy and then receive ongoing supervision during a study (to ensure that they continue to adhere to the therapy manual) cannot be assumed to generalize to community therapists who might provide the treatment. The generalizability point is a valid one. At minimum, it underscores the need for therapist training methods and materials to be made widely available for interventions found to be efficacious in high IV studies.

Contemporary controversies aside, methods that can contribute to EV are described next. Some of their limitations are also noted.

3.10.4.5.1 Naturalistic treatment settings

The ultimate aim of psychotherapy intervention studies is to identify therapies that can be provided effectively in settings where individuals who have the problems that are targeted by the therapy are treated. However, studies with the highest IV, for example, those that include random assignment and a placebo treatment, are often done in university-affiliated clinic settings. Such settings are where researchers tend to be located. The settings also tend to be most amenable to the implementation of research methods that are needed for high IV studies.

University-affiliated clinical settings differ in many ways from other settings in which mental health interventions are provided, such as urban and rural community mental health clinics and private practice offices. The differences between a study's setting(s) and the settings to which it is important to generalize its findings reduce EV only if the differences affect the outcomes of the interventions being tested. Thus, to accurately evaluate a study's EV with respect to nonstudy settings, one must know which setting features affect the outcomes of the interventions. Unfortunately, research generally is not done on the foregoing topic. In the absence of knowledge, the prevailing convention seems to be to adopt a conservative position: namely, the more similar to "typical" mental health treatment settings that an intervention study setting is, the greater the study's probable EV.

The foregoing convention results in recommendations like: a study protocol that has high IV should be implemented in a variety of public and private mental health treatment settings, both urban and rural, and in different sections of a country (i.e., in large countries where areas exist that have different subcultures). Unfortunately, the preceding strategy to enhance the EV of settings is rife with limitations (cf. Clarke, 1995). Cook and Campbell (1979) extensively discussed the problems, which they referred to as "major obstacles to conducting randomized experiments in field settings." They also discussed social forces and field settings that are conducive to randomized intervention studies.

A common obstacle is that all community clinics have standard operating procedures. The need to introduce the key IV-enhancing experimental method, random assignment, typically is an enormous challenge in such settings. At minimum, it requires that all regular clinic staff, or at least a representative sample of them, be willing to learn and to correctly implement a new intervention. Assuming that the staff are interested in being trained in a new intervention,

Table 3 External validity (generalizability).

Threat	*Description*	*Example antidote*
Interaction of selection and treatment	The study's subject inclusion and exclusion criteria and/or other requirements of study participation could result in a sample that responds differently to the intervention(s) than would other individuals of the type for whom the intervention(s) was designed.	Make entry into the study as convenient as possible to maximize the representativeness of the sample of the desired population. Attempt to replicate the study using modified inclusion criteria (e.g., drop the prerandomization study induction procedure that was used in the original study)
Interaction of setting and treatment	Features unique to the study setting(s) could contribute to the intervention effects (e.g., the setting was a well-known treatment and research clinic, famous for one of the interventions)	Attempt to replicate the study by simultaneously or consecutively carrying out the same design and procedure in two or more settings
Interaction of history and treatment	Features unique to the time when the study was done could contribute to the intervention effects (e.g., a study of therapies for depression was conducted during the same year that the US National Institute of Mental Health ran a major public information campaign on the efficacy of psychological therapies for depression)	Attempt to replicate the study

Statements in the first two columns and some in the third are abstracted from Cook and Campbell (1979). The threats are listed as interactions between variables and study treatments because the threats apply when the intention is to generalize *across* populations, settings, etc., rather than *to* target populations, settings, etc. Cook and Campbell provide an extended discussion of the difference.

training time must be made available to them. The need for training can pose a major hurdle because staff work days typically are fully booked with existing job responsibilities.

An alternative is to hire new clinic staff specifically to conduct treatments for a study. This strategy typically is associated with pragmatic limitations such as office space. Also, hiring new staff can compromise the setting EV goal, the main reason for wanting to conduct an intervention study in a field setting.

3.10.4.6 Inclusive Subject Selection

Mental health interventions are developed to be helpful to individuals who have particular types of problems. Hence, subject EV is a critical type of EV for an intervention study. Several threats to subject EV that are associated with procedures which enhance IV have been described (e.g., Howard, Cox, & Saunders, 1990; Howard, Krause et al., 1986). For example, IV requires that subjects accept random assignment to any intervention condition in a study. Some study applicants will refuse to participate due to this requirement. Such refusal is a threat to EV because the study sample will not be representative of all individuals for whom the study intervention is believed to be appropriate. Also, studies with high IV typically require subjects to participate in research assessments, often extensive ones. Some require subjects to agree to audio- or videorecording of therapy sessions. A portion of potential subjects are likely to refuse participation due to assessment requirements, thereby reducing the subject EV of study findings. All of the preceding examples are types of preinclusion attrition.

The fundamental threat to IV posed by subject dropout from treatment studies (post-inclusion attrition) was described previously. Dropout is also a serious, predictable, and often intractable EV threat. Researchers' careful selection, implementation, and description of subject inclusion and exclusion criteria all potentially enhance a study's EV by providing specific information on the types of individuals to whom the findings are most likely to generalize. The loss of outcome data that typically is associated with subject dropout compromises the confidence with which it can be concluded that a study's findings generalize to individuals like those selected for the study. The threat to EV of subject dropout is another reason why it is critically important to try to obtain data at all required assessment points from all subjects who enter a study, dropouts and remainers alike.

One of the most frequent subject EV criticisms of contemporary US studies with high IV is their psychodiagnostic inclusion and exclusion criteria (e.g., American Psychiatric Association, 1994). The criticism is that the use of specific diagnostic critieria for sample selection yields subjects who are unrepresentative of individuals who would receive the interventions in community settings. Study samples are said to be highly selected, consisting of "pure" types that are unrepresentatively homogeneous compared to treatment-seeking individuals who contact field clinical settings (e.g., Lebowitz & Rudorfer, 1998; Seligman, 1995).

The preceding subject of EV concern is a relatively new one. It relates directly to a mental health intervention research strategy that developed in the USA after 1980, when the third edition of the *Diagnostic and statistical manual of mental disorders* (*DSM*; American Psychiatric Association, 1980) was published. Shortly thereafter, sociopolitical forces converged to contribute to the requirement that applications for federal grant funding for mental health intervention efficacy studies be focused on specific diagnoses (e.g., major depressive disorder; social phobia) as defined in the *DSM*. Previously, the sample selection criteria for such studies were very broad, for example, treatment-seeking adults who complained of a variety of "problems in living" but who had no history of a psychotic disorder or current drug addictions.

The validity of the foregoing subject of EV criticism merits examination. It is true that the trend in intervention research in the USA since 1980 has been to develop and test interventions for relatively specific symptoms and syndromes (e.g., depression). However, study reports generally do not include comprehensive descriptions of co-present diagnoses (i.e., comorbidity patterns) of subjects. Thus, the extent to which most study samples are indeed more homogeneous than community setting patients for whom an intervention would be recommended typically cannot be determined. Also, some evidence from research clinics is beginning to appear that does not support the homogeneity criticism (e.g., Kendall & Brady, 1995).

3.10.4.6.1 Staff therapists

Another commonly mentioned threat to EV is the use of specially selected therapists. For example, sometimes investigators limit study therapists to a subset who meet a criterion of skillfulness. The rationale is that delivery of study interventions by therapists who are optimally skilled will provide the most accurate test of an intervention's efficacy. However, the strategy also poses a threat to EV. Can the outcomes of a study in which expert therapists delivered the interventions be assumed to generalize to less skilled therapists? The most conservative answer to this EV question is probably not. However, many contemporary intervention studies with high IV also specify (i) methods to identify therapists like those selected for a study, and (ii) therapist training programs that include ways to measure skill attained before therapists start to treat study patients. Thus, high IV studies that meet current methodological standards describe ways to identify and train nonstudy therapists who are like study therapists and, thus, who are likely to provide treatments that have outcomes similar to those obtained (cf. Kazdin, Kratochwill, & VandenBos, 1986).

3.10.4.6.2 CV Methods

Table 4 summarizes threats to CV and lists some antidotes. The text that follows elaborates on some of the information in Table 4.

In Cook and Campbell's 1979 discussion of CV, they opined that "most applied experimental research is much more oriented toward high construct validity of effects than of causes" (p. 63). Their observation was accurate for psychotherapy intervention research at the time but, notably, does not hold for therapy research published since about 1990 (e.g., Elkin et al., 1989). A methodological revolution in psychotherapy research enabled the CV of putative causes, that is, therapeutic interventions, to be enhanced. The relevant new methods are described in Sections 3.10.4.7.1–3.10.4.7.3.

3.10.4.6.3 Specification of interventions: treatment manuals

Cook and Campbell (1979) highlighted the dependence of CV on careful pre-operationalization explication of the essential features of study constructs. Prestudy preparation of treatment manuals, discussed in Section 3.10.4.1.3, is a research method that achieves pre-operationalization explication of the intervention construct, if done well. Criticisms of the treatment manual revolution in therapy research have been frequent and heated, despite the critical contribution of manuals to both the IV and potential EV of psychotherapy intervention studies

Table 4 Construct validity of putative causes and effects.

Threat	*Description*	*Example antidote*
Inadequate preoperational explication of construct	The researchers' elucidation of one or more of the main study constructs (i.e., the interventions and outcome) was not adequate to guide valid operationalization (measurement) of them	Devote necessary resources to identifying or developing valid operationalizations of interventions and outcomes because high construct validity of intervention and outcome variables is essential to conducting interpretable intervention research
Mono-operation bias	Reasonable variation in the implementation of an intervention does not occur (e.g., only male therapists are used) which reduces the validity of the operationalization of the intervention	Vary putatively irrelevant features of study interventions (e.g., use male and female therapists who represent a wide age range to deliver interventions)
Monomethod bias	Only one measurement method is used to assess central variables (e.g., all outcome measures use the paper and pencil self-report method) which confounds variance in scores due to method with variance that reflects the variable of interest	Include measures of central variables that are based on more than one assessment method (e.g., self-report and behavioral observation of symptoms)
Hypothesis-guessing within experimental conditions	Subjects try to intuit the study hypotheses and skew their responses based on what they think the hypotheses are	Try to make hypotheses hard to guess or intentionally tell subjects false hypotheses (and debrief them after the study is over). Tell subjects who are in a treatment outcome study that honest, candid responses are more useful to the research than trying to "help" the researcher by showing improvement on measures.
Evaluation apprehension	Subjects might present themselves more positively (or more negatively) than their honest self-evaluation when they are being assessed by experts in personality, psychopathology, or basic human skills	Use paper and pencil self-report measures of the same constructs that "expert" interviewers assess; include measures of social desirability response style, and use them as covariates if they are correlated with scores on outcome measures. Administer measures of socially undesirable variables at a prestudy baseline point and again immediately before study begins. Assume that the second administration is likely to be the more valid (candid) index
Experimenter expectancies	Investigators' biases influence the study findings	Investigators should not administer study interventions. Conduct the same study protocol in two or more settings, each of which is overseen by an investigator with a different preference (bias) for the interventions examined
Confounding constructs and levels of constructs	Some interventions might be differentially efficacious when provided at different levels or strengths (e.g., doses of psychotropic medications; a psychotherapy conducted with two sessions per week for the first 12 weeks vs. once weekly for the first 12 weeks)	Conduct a study to examine whether an intervention's effects differ by the strength of the "dose" administered

Table 4 (continued)

Threat	*Description*	*Example antidote*
Interaction of different treatments	Subjects concurrently receive nonstudy interventions and the study intervention (e.g., some subjects in a psychotherapy study start taking psychotropic medications; in a study of medication treatment for drug addiction, some subjects also receive medical and housing help at the study clinic)	Use procedures to reduce the probability that subjects will receive treatments other than the study treatments. For example, as a criterion for study participation, require subjects to refrain from taking psychotropic medications or having any other psychotherapeutic treatment while in the study. Check throughout a study for use of proscribed treatments and either withdraw subjects from study or exclude the data of subjects who use proscribed treatments from outcome analyses
Interaction of testing and treatment	The study assessment procedures affected the response of subjects to study interventions (e.g., a preintervention measure sensitized subjects in a way that enhanced their response to the study intervention)	Examine the presence of this threat experimentally by administering the measures to two subject samples. Administer tests to one sample at pre- and postintervention; administer to other only at postintervention. A complete Solomon four-group design (Campbell & Stanley, 1963) can also be used
Restricted generalizability across constructs	Interventions can affect many important outcome variables. However, effects on outcomes (constructs) of interest often are not highly correlated (generalizable)	Include measures that are closely related to the main outcomes sought with an intervention, in addition to the key outcomes (e.g., in a study of treatments for depression, include measures of severity of depression symptoms, work functioning [number of sick days], and marital conflict

Statements in the first two columns and some in the third are abstracted from Cook and Campbell (1979).

(Wilson, 1996). However, the CV of intervention constructs requires more than manuals. It also requires prestudy therapist training procedures to teach therapists how to conduct the interventions as specified in the manual. A final step is to evaluate and ensure the effectiveness of the training "manipulation" before therapists are permitted to conduct therapies with study subjects. Methods for this are discussed next.

3.10.4.6.4 Specification of interventions: therapist adherence measures

The previously mentioned process research methodology, observer-judge ratings of therapy session material, provides the most valid index available to date of the extent to which a therapist has acquired a requisite skill level with a study intervention. Some research findings support the preceding conclusion. For example, trainee therapists' scores on written tests about a therapy were not highly correlated with supervisors' judgments of trainees' skill in conducting a therapy. Likewise, supervisors' impressions of trainees' competence with a therapy, based on trainees' descriptions of therapy sessions, were not highly correlated with supervisors' ratings of trainees' competence based on videotapes of therapy sessions (Chevron & Rounsaville, 1983).

The use of process methods to assess the extent to which therapists conduct treatments as specified requires the development of adherence measures (e.g., Evans et al., 1984). Adherence measures typically consist of items that operationally define what the developers of an intervention believe to be its essential therapeutic elements. Judges review therapy session material and rate the extent to which a therapist enacts (verbally and/or behaviorally) each item as specified by the adherence measure. The development of a psychometrically sound therapist adherence measure for an intervention, including a system for sampling the therapy session material to which the measure should be applied, is a time-consuming and sophisticated research methodology.

The use of self-report therapist adherence measures recently has been proposed as a substitute for judge-rated measures (Carroll, Nich, & Rounsaville, in press). Self-report adherence measures have interesting uses in therapy research, perhaps particularly in therapy development research. However, the only psychometric examination to date of the validity of such methods as indices of therapists' adherence to a treatment as specified in a manual yielded negative results (Carroll et al., in press).

Therapist adherence measures are a way to assess the CV of study interventions. Adherence measures can be applied to systematically sampled therapy session material to yield an index of the extent to which study therapies were conducted as specified. The contribution of psychometrically sound therapist adherence measures to both CV and IV cannot be overemphasized: they allow us to evaluate whether or not the designated interventions (the independent variables) were delivered as intended in intervention studies.

The preceding discussion is a simplified explication of the issues involved. The advent of therapist adherence methodology was associated with recognition of the difference between two constructs, adherence and competence (e.g., Waltz et al., 1993). Basically, adherence denotes a therapist's *fidelity* to the therapeutic techniques that are described in a treatment manual; competence denotes *skill* in implementing the techniques (e.g., tailoring a technique to increase its acceptability to a particular client). Reseach to date is consistent with the conclusion that inter-rater reliability is more achievable with adherence measures than with competence measures. Recruiting experts in a particular form of therapy to judge competence does not solve the reliability problem (a "problem" which signals needed directions for psychotherapy outcome research at this time).

3.10.4.6.5 Distinctiveness of interventions

The main aim of some intervention studies is to compare the outcomes of alternate treatments for the same problem. A fundamental assumption of comparative treatment studies, albeit sometimes implicit, is that the interventions being compared are distinctive in terms of their putative therapeutically active elements. For example, a study that is designed to compare the outcomes of a form of cognitive therapy for depression with another type of psychotherapy for depression is based on the assumption that the therapies are substantively different somehow. The assumed distinction is what makes it worthwhile to compare therapies in a single study. In CV terms, a comparative intervention study design reflects the assumption that the intervention constructs are distinguishable. While the issue might seem moot, it is not. Therapies that sound distinct based on their theories of therapeutic change and prescribed techniques might not be distinguishable in practice, that is, when they are implemented (operationalized) in words and behaviors by therapists.

Therapist adherence measures for different interventions provide a systematic way to evaluate the distinctiveness of putatively different forms of psychotherapy and, hence, a way to evaluate the CV of comparative intervention studies. The general methodology is to apply the therapist adherence measure for each intervention to the intervention for which it was developed and to the other interventions in a study (e.g., cognitive therapy, psychodynamic therapy, and placebo therapy sessions each are rated on a cognitive therapy adherence scale and a psychodynamic therapy adherence scale). The scores obtained on the adherence measures for each intervention can then be compared to *a priori* criterion scores that are established to indicate the distinctiveness of the interventions. Hill et al. (1992) exemplifies the methodology.

3.10.4.6.6 Adjunctive therapies

When an intervention study is described as a test of a specific intervention, it is assumed that the intervention is the only one that the subjects received. However, sometimes study subjects are allowed to participate in adjunctive interventions (e.g., Alcoholics Anonymous) or are provided with additional interventions on an as-needed basis (e.g., aid from a case worker for housing and other needs; psychotropic medications). Participation in adjunctive interventions constitutes a threat to the CV of the study intervention, as well as to the study's IV. Outcome findings cannot be validly interpreted as due to the designated intervention(s) when adjunctive interventions are allowed but not controlled in some way.

3.10.4.6.7 Assessment procedures

Repeated testing on the same outcome measures is a CV threat that often is ignored. Cook and Campbell (1979) frame the issue as follows: can an intervention–outcome relationship be generalized to testing conditions other than those used in the study to assess outcomes?

For example, if an outcome measure was administered both pre- and postintervention, would the same outcomes have been obtained if the measure was administered only at postintervention? One way to examine the threat of repeated testing to the CV of outcome scores is to compare findings of two randomly assigned intervention groups: one that was both pre- and post-tested, and another that only was post-tested.

A related CV threat is the possibility that study outcome measures interact with interventions. For example, studies of cognitive-behavioral (CB) therapy for panic disorder have been criticized for the possibility that CB affects a primary outcome measure used in such studies, self-reported frequency of panic attacks, mainly by teaching subjects to redefine panic attack rather than by reducing the actual frequency of attacks. To the extent that the criticism is valid, differences in panic attack frequency found between a control intervention and CB have low IV because the outcome measure in the CB condition has low CV. The measure does not reflect change in frequency of panic attacks, it reflects change in subjects' definitions of panic attack. The Solomon Four-Group Design (Campbell & Stanley, 1963; Rosenthal & Rosnow, 1991) can be used to simultaneously evaluate main effects of repeated testing on outcome measures and interactions of interventions and measures.

A common way to reduce CV threats to an outcome measure like the one just described for panic is to include other measures of closely related constructs in a study. The other measures should not be vulnerable to the same CV threat. For example, a spouse could be asked to provide ratings of the frequency and intensity of a subject's panic attacks. (This solution is, of course, associated with other types of CV concerns.)

3.10.4.7 SCV Methods

Only a few SCV-relevant methods are discussed here due to space limitations and to the technical detail associated with many SCV issues. The reader is referred to Cook and Campbell (1979) for a more complete review and to Table 5 for a synopsis of their discussion of SCV.

3.10.4.7.1 Statistical power analysis

Conducting a statistical power analysis at the planning phase of an intervention study is critical to its SCV (Cohen, 1977, 1992; Kraemer & Thiemann, 1987; Rosenthal & Rosnow, 1991). A power analysis helps an investigator determine the minimum sample size needed to detect a statistically significant outcome difference of a magnitude that he or she regards as clinically meaningful (e.g., Jacobson & Truax, 1991), given his or her willingness to risk making a type II error (i.e., accept the null hypothesis of no difference when it should be rejected). The smaller the sample size, the greater the risk of type II error when standard statistical procedures like analysis of variance are used in intervention research. However, the larger the sample size, the more expensive, time-consuming, and less feasible a study typically becomes.

A prestudy power analysis is a computation that allows investigators to enhance a study's SCV by determining the minimum sample size needed to statistically detect an outcome difference of a specified magnitude, given expected variances in the outcome measures. Ideally, an investigator can draw on criteria of some type to determine the size of a difference (effect size) that will have applied significance, that is, be clinically meaningful (Jacobson & Truax, 1991). For example, a clinically meaningful difference is one that experienced therapists would regard as a compelling reason to recommend one treatment over another for a particular problem.

The rationale for a power analysis is as follows. The magnitude of the variances of outcome measures and the study sample size are major determinants of whether or not an outcome difference of a certain magnitude between interventions will be statistically significant (i.e., unlikely to be due to chance). The same absolute difference between mean outcome scores of two interventions could be statistically significant if the sample size in each group was 30, but not if it was 15. Another perspective on the situation is that the smaller a study's sample size, the higher the probability of a type II error if the conventional $p \leqslant 0.05$ is used to indicate statistical significance. Rosnow and Rosenthal (1989) present a compelling perspective on the fundamental importance of the foregoing issues.

3.10.4.7.2 Controlling for the type I error rate

Statistical tests commonly used in intervention research (e.g., the F-test in analysis of variance) are associated with probability values, for example, $p < 0.05$. A p value indicates the probability that a difference found between interventions is due to chance rather than a true difference. Thus, for example, a statistical test that meets the $p \leqslant 0.05$ significance level means that a difference of the same magnitude would

Table 5 Statistical conclusion validity.

Threat	*Description*	*Example antidote*
Low statistical power	Limited potential of a study to yield statistically significant intervention effects (i.e., low potential to reject the null hypothesis) due to sample size and to the probable variances of outcome measures	Conduct a power analysis (e.g., Cohen, 1977, 1992; Rosenthal & Rosnow, 1991) during the design phase of a study to determine the sample size needed to detect an intervention effect of a predetermined size
Violated assumptions of statistical tests	Failure to ensure that a study's outcome data meet the assumptions of the statistical tests used to analyze it	Know the crucial assumptions of all statistical tests used and conduct analyses needed to evaluate whether the study data meet the assumptions. Seek expert statistical consultation if needed.
Fishing and the error rate problem	Several unplanned (i.e., *a posteriori* rather than *a priori*) statistical tests of effects (fishing) are conducted but no adjustments are made to reduce the likelihood that statistically significant findings are due to chance. In other words, the probability of type II error is larger than indicated by the reported *p* values	Consult statistics textbooks for methods to handle the error rate problem (e.g., Bonferroni correction; Tukey or Scheffé multiple comparison tests)
Reliability of measures	Low test–retest reliability (stability) and/or low internal consistency reliability of outcome measures inflates the error variance in scores and thus reduces the probability of finding intervention effects	Rely on psychometric properties of measures to guide selection of study measures. Choose measures with strong psychometric properties over measures with poor or unknown psychometric properties that might have better "face validity" as measures of the construct of interest (Kraemer & Telch, 1992)
Reliability of treatment implementation	Variability in the implementation of an intervention with different subjects (lack of standardization across subjects) inflates error and, thus, reduces the probability of finding intervention effects	Train study therapists to a criterion level of adherence to interventions before they treat any study patients; monitor their adherence to the treatment manual throughout the study and intervene if drift occurs (e.g., Shaw, 1984)
Random irrelevancies in the experimental setting	Features of a study setting other than the interventions could affect outcome scores (e.g., two research assistants administer the outcome measures to subjects: one is warm and friendly and the other is abrupt and distant)	Code features of the study setting that are likely to affect outcome measures as variables and include them in the data analyses to reduce error (e.g., research assistant who administered outcome measures is coded as a variable and included as a factor in an analysis of variance that is used to test for treatment effects)
Random heterogeneity of subjects	Differences in subjects could affect scores on the outcome measures independent of any intervention effects	Measure subject characteristics that are likely to be correlated with scores on outcome measures; use the characteristics as covariates or as blocking variables

Statements in the first two columns and some in the third are abstracted from Cook and Campbell (1979).

be expected to be obtained on average five times or fewer if 100 of the statistical tests were done, if the null hypothesis of no difference between interventions were true. The use of *p* values in conjunction with statistical tests is the principal way investigators guard against type I error (rejecting the null hypothesis of no difference when it should be accepted) when interpreting study findings. Thus, conducting statistical tests in ways that maintain the interpretability of *p* values is crucial to a study's SCV.

The interpretation of a *p* value is compromised when multiple statistical tests are done but no precautions are taken to ensure that the degree of protection against a type I error remains constant. A variety of procedures exist to maintain a specified level of protection against type I error when several statistical tests are performed. Statistics textbooks describe procedures such as the Bonferroni correction (e.g., Rosenthal & Rosnow, 1991). The main point for present purposes is that conducting multiple statistical tests without adjusting the probability level of the tests compromises a study's SCV. Unfortunately, this is a common type of threat to SCV in intervention research (Dar, Serlin, & Omer, 1994).

3.10.4.7.3 Testing assumptions of statistical tests

One fundamental assumption of the statistical tests that are commonly used for outcome analyses is that the dependent (outcome) variable is normally distributed. If the data do not meet the normality assumption, then the validity of the statistical test is compromised, that is, the study SCV is threatened. A special class of statistical procedures, nonparametric statistics (Leach, 1979), exist for analyzing data that are not normally distributed. Alternatively, procedures exist to transform data (e.g., square root transformation) to make their distribution more symmetrical before applying statistical tests (Atkinson, 1985; Carroll & Rupert, 1988).

3.10.5 FROM IV VS. EV TO IV + EV IN MENTAL HEALTH INTERVENTION RESEARCH

One illustration of the centrality of IV and EV to intervention research is the aforementioned controversy on efficacy vs. effectiveness studies (e.g., Goldfried & Wolfe, 1998; Jacobson & Christensen, 1996; Lebowitz & Rudorfer, 1998; Mintz et al., 1996; Seligman, 1995, 1996). The controversy took center stage largely as a side effect of dramatic shifts in health care delivery in the USA since the mid-1980s. Two forces were particularly valent in drawing the attention of intervention researchers and consumers of their research to efficacy (IV) and effectiveness (EV) methods. The forces were (i) a move from indemnity insurance coverage for health care to managed care coverage to control costs, and (ii) the US Congress's attempt in 1993 to reform national health care. Both forces highlighted the need for valid information about interventions to guide decisions of government policymakers and managed care entrepreneurs. Both groups were motivated to make decisions rapidly that would impact the interventions that millions of Americans could receive in insured programs. The need for valid findings to inform the decisions was keenly felt by all interested parties, including researchers. Limitations of information from existing studies became widely and painfully evident. Lively debates later appeared about how intervention research in the USA, particularly federally funded research, should be designed (e.g., Clarke, 1995; Jacobson & Christensen, 1996; Lebowitz & Rudorfer, 1998; Newman & Tejeda, 1996; Seligman, 1995, 1996; Speer, 1994; Wells & Sturm, 1996).

A question at the core of the debate is how to design intervention studies to maximize three aims: (i) the value of findings in terms of definitive public health implications and enhancement, (ii) the speed with which valid and useful findings can be obtained, and (iii) the public health policy yield from relatively scarce federal funding for mental health intervention studies. Personal biases are evident in positions taken by many who have spoken on the issues (e.g., pragmatism vs. concern that policy decisions that affect large numbers of people will be made based on scientifically weak findings). However, a valuable product of the debate has been creative thinking about how IV and EV goals might be maximized, for example, via new study questions and methods, and the use of statistical techniques such as causal modeling (Kenny, 1979). Interestingly, Cook and Campbell (1979) were keenly aware of all three of the aforementioned aims in the late 1970s. The ideas developed in all editions of their book, particularly in the sections on quasiexperimental designs, were largely intended to address them.

The current debate illustrates the importance of being fully fluent with the concepts of IV and EV for both creators and consumers of intervention research. A few suggestions from the debate for mounting studies that have both high IV and EV are described next. Then, two illustrative IV + EV studies are reviewed in terms of IV and EV.

3.10.5.1 Strategies for Efficacy + Effectiveness (IV + EV) Intervention Studies

Most strategies for designing intervention studies that optimize both IV and EV that have emerged so far from the efficacy vs. effectiveness debate are intended to maximize the applied public health (utilitarian) information that is obtained from studies. However, to test treatments that hold promise as marked advances over currently available interventions, high IV designs and methods are likely to be the most informative research strategies, due to their suitability for theory-testing and for testing causal hypotheses (Jacobson & Christensen, 1996; Mintz et al., 1996). The IV + EV strategies described here were selected to illustrate the diversity of approaches that have been proposed. Each strategy is described briefly, in a way intended to highlight its essence.

3.10.5.1.1 The conventional "phase" strategy

Before the recent efficacy vs. effectiveness debate, federal agencies that funded mental health research in the USA endorsed a phase model to meet the goals of IV and EV inference. According to Hoagwood et al. (1995), the phase model was adapted from one used by the US National Cancer Institute. Described in a simplified way, the model is to conduct efficacy studies of interventions for which promising outcome data were obtained in prior, preliminary research; the next phase is to conduct effectiveness studies of interventions for which compelling efficacy data are obtained.

3.10.5.1.2 A dimensional adaptation of the phase model

Hoagwood et al. (1995) identified several reasons why the phase model, although elegant, is rarely fully realized. A major hurdle has been moving interventions with demonstrated efficacy to effectiveness studies. A typical stumbling block to the transfer is introducing research methods into community settings to conduct effectiveness studies. Hoagwood et al. presented an alternative strategy, the essence of which is to conceptualize efficacy and effectiveness goals "in terms of continuous dimensions, rather than discrete phases" (p. 685). The authors describe a conceptual framework, not a model *per se*. It is linked specifically to mental health interventions for children, adolescents, and families.

The framework emphasizes three bipolar dimensions of intervention research, each of which can be described as having an IV and an EV pole. The dimensions are types of validity, types of intervention modalities and parameters, and types of outcomes. For example, at the IV pole of the intervention dimension are short-term (e.g., three months), highly structured, single modality therapies that are operationalized in detailed manuals; at its EV pole are longer-term, less well-specified treatments that allow for the integration of more than one therapy modality.

3.10.5.1.3 Stakeholder's model to guide data selection

Newman and Tejada (1996) highlighted the need to collect data that are sought by all parties who have a vested interest in mental health interventions (the stakeholders). The authors endorsed the view that high IV studies have an essential role in intervention research because they provide scientifically well-founded, important information about interventions, for example, their potential effects and their safety. EV goals are accomplished in the stakeholders model mainly by *collecting certain kinds of outcome and other data* about interventions that can help inform the decisions of all stakeholders in mental health treatment (e.g., government policymakers and corporate officials of managed care companies; potential recipients of interventions; family members and others who are directly affected by those who receive mental health interventions).

Types of data emphasized by Newman and Tejada (1996) include costs of interventions and estimates of the amount of therapeutic effort required to obtain specified behavioral outcomes. Hoagwood et al. (1995) also included cost data in their framework. The need for cost data (e.g., cost-effectiveness, cost-offset, comparative cost) is a dominant theme in research strategies that have been proposed for efficacy + effectiveness studies. This emphasis is largely a response to increasing concerns about, and concomitant restrictions on, health care costs in the USA. Cost data have not been standard products of federally funded efficacy studies in the USA for the past 30 years. Yates (1997) has presented guidelines for psychotherapy researchers on how to collect cost data. Wolff, Helminiak, and Tebes (1997) illustrate the impact of different assumptions on cost estimates in mental health treatment research. Other relevant references are Knapp (1995) and Yates (1996). Research strategies to examine the amount of therapeutic effort required to obtain specific outcomes have been available for some time in the psychotherapy research literature (e.g., Howard, Koptka, Krause, & Orlinsky, 1986; Kopta et al., 1994).

3.10.5.1.4 Mediator's model

The crux of a model proposed by Clarke (1995) is to use IV experimental methods to examine crucial generalizability (EV) questions about an interventions' outcomes. For example, a study design could include two or more intervention conditions distinguished mainly by the degree to which an intervention is implemented as specified in its manual. The essence of Clarke's (1995) model is to use IV methods to evaluate the effects of several potential mediator variables on the outcomes that can be achieved with an intervention. Examples of relevant mediators are the degree to which therapists adhere to the structure of a therapy as specified in its manual, heterogeneity of the patients treated, and the adjunctive therapeutic context in which an intervention is provided (e.g., used as a stand-alone intervention vs. included in an intervention package).

3.10.5.2 Examples of Efficacy + Effectiveness Intervention Studies

This section illustrates the use of the concepts of IV and EV to evaluate intervention studies. Two highly regarded US studies that were designed to meet both IV and EV aims are reviewed: Drake, McHugo, Becker, Anthony, and Clark (1996) and Schulberg et al. (1995). The main intent of each review is to illustrate how the concepts that are the focus of this chapter can be applied to help determine the conclusions that can validly be drawn from an intervention study. The reviews obviously are not complete or definitive evaluations of either study.

3.10.5.2.1 Schulberg et al. (1995)

The investigators' interest in integrating IV and EV considerations in a single study is evident in the study's two main aims. One aim was to compare the outcomes of a type of psychotherapy and a medication when used to treat depression in patients who were being seen in general medical clinics, not psychiatric settings. Both of the interventions were tested previously in randomized controlled studies and found to be efficacious for depression in psychiatric outpatients. A second aim was to compare the effects on depression of each of the two tested interventions with the effects of a third intervention of "usual care," that is, treatments of the type typically provided to depressed patients by doctors in primary care settings.

Central IV features of the study were: random assignment of patients to all three treatment conditions, use of sample selection procedures to identify patients who met standard psychiatric diagnostic criteria for depression, and outcome analyses on an intention-to-treat sample. IV and EV were simultaneously enhanced by other methods. The same study protocol was implemented concurrently at four primary care clinic settings. Procedures were used to help ensure that the psychotherapy was conducted in ways consistent with its manual but that also were minimally intrusive into the conduct of the therapies. Central EV features of the study were: it was conducted in primary care settings (although all four settings had academic affiliations), outcome data were obtained from patients who were randomly assigned to usual care, and the natural course of the usual care condition was preserved as much as possible.

A noteworthy EV threat was substantial preinclusion attrition of eligible subjects due to refusal to participate and other reasons. An IV threat was that therapists were "nested" in a treatment condition. That is, the medical professionals who provided usual care were not the same professionals who provided the psychotherapy intervention or the pharmacological intervention. Thus, better outcomes of the previously tested (efficacious) interventions compared to usual care could not be attributed validly to the therapies *per se*. Findings of this type were, in fact, obtained. However, it can be argued that the confounding of therapists and treatments does not pose an important IV limitation to this particular study. Rather, the confound enhances the generalizability of findings to standard practice (EV) because in the public sector, primary medical care providers are not the same as those who provide specific forms of psychotherapy for depression or those who specialize in psychopharmacological treatments for depression.

3.10.5.2.2 Drake et al. (1996)

Drake et al. (1996) integrated IV and EV methods in a study designed to examine interventions of a very different type from those examined by Schulberg et al. (1995). Their main aim was to compare two vocational programs for individuals with severe mental disorders (SMD) such as schizophrenia. Drake et al. described the interventions as differing in two primary ways. One intervention was provided by a professional rehabilitation agency external to the mental health clinics where individuals with SMD were treated; the other was integrated with mental health services and provided at the clinics. A second difference was in the structure of the interventions. The external intervention ("group skills training," GST) started with three months of pre-employment skills training; the mental health clinic intervention ("individual placement and support," IPS) started immediately with job placement activities.

The central IV feature of the study was random assignment to each of the intervention conditions. A key EV enhancement design feature was that each intervention was concurrently tested at two sites. However, the full EV value of the multisite design was compromised because GST staff who had central training roles overlapped at the two GST sites. Also, the GST and IPS interventions each were conducted at different sites. Thus, as the authors point out, interventions and sites were confounded. This type of confound notably compromises IV, that is, the confidence with which outcome findings can be attributed primarily to interventions.

The study illustrates that sample selection criteria can simultaneously increase the homogeneity of a study sample and enhance EV. The inclusion and exclusion criteria constitute minimal, yet realistic, qualifications for SMD clients to receive vocational interventions (e.g., lack of noteworthy memory impairment; clinical stability as indicated by not being hospitalized for at least one month; unemployment for at least one month but interest in employment). The potential subject EV enhancement associated with the sample selection criteria was reduced by another procedure, a "prerandomization group experience." SMD clients who were interested in participating in the study were required to attend a minimum of four weekly research induction group sessions. One purpose of the prerandomization group attendance requirement was to ensure that potential participants were "motivated" for the project interventions.

Prerandomization and pretreatment "orientation" or "stabilization" phases are sometimes used in mental health intervention research, particularly with populations in which high treatment failure and/or dropout rates are expected, such as drug abusers. EV is one trade-off for the use of such procedures. On the other hand, the procedures can have an IV advantage. They can create samples of individuals who are most likely to respond to an intervention, thereby yielding findings on its efficacy with individuals hypothesized to be most responsive to it. One potential risk is that the sample will consist of individuals who are likely to benefit from many incidental life/environmental experiences and who, therefore, might not need the study interventions to achieve its targeted outcomes. The foregoing IV threat can be offset by the inclusion of a placebo control intervention of some type in the design.

Drake et al. (1996) also illustrates the critical impact on a study's CV, IV, and EV of the methods used to operationalize the interventions. The study report does not indicate if the interventions were specified in manuals or the type of training that counselors received before implementing them for the study. A related threat to the CV of the interventions and, thus, to the IV of study findings, is that the methods used to evaluate the success of the operationalization of the interventions (i.e., counselor adherence to GST and to IPS) are described too vaguely to evaluate. An additional IV threat is that it is not clear from the report if the IPS intervention was always provided by a counselor other than a client's SMD therapist. This is an IV consideration because the GST interventions were always provided by someone other than the client's SMD therapist.

One point illustrated by the preceding intervention-related aspects of the study is that the less clear and complete the description of methods used to implement interventions is, the less possible it is to evaluate their CV and, consequently, the IV of outcome findings. Thus, the public health significance of findings cannot be evaluated confidently. In turn, study findings can have less impact than they otherwise might.

An additional IV threat due to intervention procedures was site differences in the implementation of both interventions. The investigators intervened at one of the sites to correct deviations from the IPS model. However, the corrective attempts were not notably effective. The implementation of IPS is described as "weaker" at one site. Also, the GST counselors "concentrated their efforts" in one of their sites.

The study findings highlight the crucial importance to IV of intervention implementation procedures (intervention CV). In some analyses, the interventions differed on the major outcome variable (employment), favoring IPS; in others, the interventions did not differ. Interpretation of mixed findings about interventions' comparative effects is seriously compromised when the intervention CV is either poor or difficult to evaluate from information provided in a study report.

The foregoing review of Drake et al. (1996) from the perspective of IV and EV threats might seem inconsistent with the fact that it is widely regarded as one of the best IV + EV mental health intervention studies available. If it is assumed that the preceding IV and EV analysis itself is valid, what might be concluded? One conclusion is that difficulties maintaining the CV of interventions in studies that are done in natural settings to enhance EV can be costly in terms of IV. In turn, this analysis suggests that a cost-effective use of resources in high EV studies is to invest in methods that achieve and maintain the CV of interventions.

3.10.6 A CONCLUDING OBSERVATION

IV and EV considerations, and the related efficacy vs. effectiveness debate, highlight two vital applied aims of intervention research. The aims are: (i) to obtain findings that can be used directly and quickly in public mental health delivery systems, and (ii) to obtain findings that can markedly improve, ideally revolutionize, the effectiveness of treatments available in the public health sector. Both aims have clear public health significance. Aim (i) is most consistent with using designs and methods that are firmly tied to contemporary constraints on clinical practice in "real world" settings and obtaining findings that can affect practice in the short run. Aim (ii) is consistent with using scientific methods that are most likely to foster discovery and methods that can rule out rival interpretations of obtained covariation between interventions and outcomes. Immediate generalizability to current practice of findings from aim (ii) studies cannot be a central criterion of their public health significance.

The efficacy vs. effectiveness debate no doubt was partially fueled by efficacy researchers' frustration. Many of them observed that years of findings from high IV studies that support the value of certain types of psychotherapy (Task Force on the Promotion and Dissemination of Psychological Procedures, 1995) were eschewed by many who focused on the EV of those studies. Also, emphasizing EV limitations of well-done, high IV studies can impede implementation of tested therapies in public sector settings. By default, then, the use of untested, "treatment as usual" interventions is supported and maintained. This is the current situation in the USA.

A partial antidote to the preceding problem, a serious problem from the public health standpoint, was suggested by Kazdin et al. in 1986. Their suggestion is consistent with a point made earlier in this chapter. Mental health intervention research advances since the late 1980s include methods to systematically train therapists in therapies with demonstrated efficacy, and to evaluate the skillfulness that therapists attain from the training. Many therapist training and adherence assessment research methods generalize directly to public sector therapists. Thus, rather than look at findings from high IV studies and ask the therapist EV question: "Does the study provide evidence that therapists in the public health sector will be able to get the same results?," the question could be refocused: "Does the study provide methods that enable therapists to be identified and trained to get the same outcomes in public treatment settings?"

ACKNOWLEDGMENTS

Preparation of this chapter was supported by National Institute of Mental Health Grant K02 MH01443 to Karla Moras. Sincere thanks are extended to those who provided consultations, comments, or other contributions to the preparation of the manuscript: Phyllis Solomon, Elizabeth McCalmont, Christine Ratto, Heidi Grenke, J. R. Landis, and Jesse Chittams.

3.10.7 REFERENCES

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Atkinson, A. C. (1985). *Plots, transgressions and regression: An introduction to graphical methods of diagnostic regression analysis.* New York: Oxford University Press.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182.

Beck, J. S. (1995). *Cognitive therapy: Basics and beyond.* New York: Guilford Press.

Bergin, A. E., & Garfield, S. L. (Eds.) (1971). *Handbook of psychotherapy and behavior change* (1st ed.). New York: Wiley.

Bergin, A. E., & Garfield, S. L. (Eds.) (1994). *Handbook of psychotherapy and behavior change* (4th ed.). New York: Wiley.

Beutler, L. E., & Howard, K. I. (1998). Clinical utility research: An introduction. *Journal of Clinical Psychology, 54,* 297–301.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 171–246). Chicago: Rand McNally.

Campbell, D. T., & Stanley, J. C. (1996). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Carroll, K. M., Nich, C., & Rounsaville, B. J. (in press). Use of observer and therapist ratings to monitor delivery of coping skills treatment for cocaine abusers: Utility of therapist session checklists. *Psychotherapy Research.*

Carroll, R. J., & Ruppert, D. (1988). *Transformation and weighting in regression.* New York: Chapman and Hall.

Chevron, E. S., & Rounsaville, B. J. (1983). Evaluating the clinical skills of psychotherapists: A comparison of techniques. *Archives of General Psychiatry, 40,* 1129–1132.

Clarke, G. N. (1995). Improving the transition from basic efficacy research to effectiveness studies: Methodological issues and procedures. *Journal of Consulting and Clinical Psychology, 63,* 718–725.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.) New York: Academic Press.

Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin, 112,* 155–159.

Collins, J. F., & Elkin, I. (1985). Randomization in the NIMH treatment of depression collaborative research program. In R. F. Boruch & W. Wothke (Eds.), *Randomization and field experimentations: New directions for program evaluation.* San Francisco: Jossey-Bass.

Consumer Reports (1995, November). Mental health: Does therapy help? (pp. 734–739).

Cook, T. D., & Campbell, D. T. (Eds.) (1979). *Quasi-experimentation: design and analysis issues for field settings.* Boston: Houghton Mifflin.

Cronbach, L. J. (1982). *Designing evaluations of educational and social programs.* San Francisco: Jossey-Bass.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302.

Dar, R., Serlin, R. C., & Omer, H. (1994). Misuse of statistical tests in three decades of psychotherapy research. *Journal of Consulting and Clinical Psychology, 62,* 75–82.

Drake, R. E., McHugo, G. J., Becker, D. R., Anthony, W.

A., & Clark, R. E. (1996). The New Hampshire study of supported employment for people with severe mental illness. *Journal of Consulting and Clinical Psychology, 64,* 391–399.

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., Glass, D. R., Pilkonis, P. A., Leber, W. R., Docherty, J. P., Fiester, S. J., & Parloff, M. B. (1989). NIMH Treatment of Depression Collaborative Research Program: I. General effectiveness of treatments. *Archives of General Psychiatry, 46,* 971–982.

Evans, M. E., Piasecki, J. M., Kriss, M. R., & Hollon, S. D. (1984). *Raters' manual for the Collaborative Study Psychotherapy Rating Scale—form 6*. Minneapolis. University of Minnesota and the St. Paul-Ramsey medical Center. (Available from the US Department of Commerce, National Technical Information Service, Springfield, VA 22161.)

Flick, S. N. (1988). Managing attrition in clinical research. *Clinical Psychology Review, 8,* 499–515.

Garfield, S. L., & Bergin, A. E. (Eds.) (1978). *Handbook of psychotherapy and behavior change* (2nd ed.). New York: Wiley.

Garfield, S. L., & Bergin, A. E. (Eds.) (1986). *Handbook of psychotherapy and behavior change* (3rd ed.). New York: Wiley.

Gibbons, R. D., Hedeker, D., Elkin, I., et al. (1993). Some conceptual and statistical issues in analysis of longitudinal psychiatric data. *Archives of General Psychiatry, 50,* 739–750.

Gillings, D., & Koch, G. (1991). The application of the principle of intention-to-treat to the analysis of clinical trials. *Drug Information Journal, 25,* 411–424.

Goldfried, M. R., & Wolfe, B. E. (1998). Toward a more clinically valid approach to therapy research. *Journal of Consulting and Clinical Psychology, 66,* 143–150.

Gomes-Schwartz, B. (1978). Effective ingredients in psychotherapy: Prediction of outcome from process variables. *Journal of Consulting & Clinical Psychology, 46,* 1023–1035.

Gotto, A. M. (1997). The multiple risk factor intervention trial (MRFIT): A return to a landmark trial. *Journal of the American Medical Association, 277,* 595–597.

Greenberg, L., & Pinsof, W. M. (Eds.) (1986). *The psychotherapeutic process: A research handbook*. New York: Guilford Press.

Greenberg, R. P., Bornstein, R. F., Greenberg, M. D., & Fisher, S. (1992). A meta-analysis of anti-depressant outcome under "blinder" conditions. *Journal of Consulting and Clinical Psychology, 60,* 664–669.

Hill, C. E., O'Grady, K. E., & Elkin, I. (1992). Applying the Collaborative Study Psychotherapy Rating Scale to rate therapist adherence in cognitive-behavior therapy, interpersonal therapy, and clinical management. *Journal of Consulting and Clinical Psychology, 60,* 73–79.

Hoagwood, K., Hibbs, E., Bren, T. D., & Jensen, P. (1995). Introduction to the special section: Efficacy and effectiveness in studies of child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63,* 683–687.

Hollon, S. D., Shelton, R. C., & Loosen, P. T. (1991). Cognitive therapy and pharmacotherapy for depression. *Journal of Consulting and Clinical Psychology, 59,* 88–99.

Howard, K. I., Cox, W. M., & Saunders, S. M. (1990). Attrition in substance abuse comparative treatment research: the illusion of randomization. In L. S. Onken & J. D. Blaine (Eds.), *Psychotherapy and counseling in the treatment of drug abuse* (DHHS Publication No. (ADM)90–1722, pp. 66–79). Washington, DC: NIDA Research Monograph 104.

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–effect relationship in psychotherapy. *American Psychologist, 41,* 159–164.

Howard, K. I., Krause, M. S., & Orlinsky, D. E. (1986). The attrition dilemma: Toward a new strategy for psychotherapy research. *Journal of Consulting and Clinical Psychology, 54,* 106–107.

Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist, 51,* 1059–1064.

Ilardi, S. S., & Craighead, W. E. (1994). The role of nonspecific factors in cognitive-behavior therapy for depression. *Clinical Psychology: Science & Practice, 1,* 138–156.

Jacobson, N. S., & Christensen, A. (1996). Studying the effectiveness of psychotherapy: How well can clinical trials do the job? *American Psychologist, 51,* 1031–1039.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12–19.

Kazdin, A. E. (1979). Nonspecific treatment factors in psychotherapy outcome research. *Journal of Consulting & Clinical Psychology, 47,* 846–851.

Kazdin, A. E. (1994). Methodology, design, and evaluation in psychotherapy research. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 19–71). New York: Wiley.

Kazdin, A. E., Kratochwill, T. R., & VandenBos, G. R. (1986). Beyond clinical trials: Generalizing from research to practice. *Professional Psychology: Research and Practice, 17,* 391–398.

Kendall, P. C., & Brady, E. U. (1995). Comorbidity in the anxiety disorders of childhood. In K. D. Craig & K. S. Dobson (Eds.), *Anxiety and depression in adults and children*. Newbury Park, CA: Sage.

Kenny, D. A. (1979). *Correlation and causality*. New York: Wiley.

Kiesler, D. (1973). *The process of psychotherapy: Empirical foundations and systems of analysis*. Chicago: Aldine.

Klerman, G. L., Weissman, M. M., Rounsaville, B. J., & Chevron, E. S. (1984). *Interpersonal psychotherapy of depression*. New York: Basic Books.

Knapp, M. (Ed.) (1995). *The economic evaluation of mental health care*. London: Arena, Aldershot.

Kopta, S. M., Howard, K. I., Lowry, J. L., & Beutler, L. E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology, 62,* 1009–1016.

Kraemer, H. C., & Telch, C. F. (1992). Selection and utilization of outcome measures in psychiatric clinical trials: Report on the 1988 MacArthur Foundation Network I Methodology Institute. *Neuropsychopharmacology, 7,* 85–94.

Kraemer, H. C., & Thiemann, S. (1987). *How many subjects? Statistical power analysis in research*, Newbury Park, CA: Sage.

Lambert, M. J., & Ogles, B. M. (1988). Treatment manuals: Problems and promise. *Journal of Integrative and Eclectic Psychotherapy, 7,* 187–204.

Leach, C. (1979). *Introduction to statistics: A nonparametric approach for the social sciences*. New York: Wiley.

Lebowitz, B. D., & Rudorfer, M. V. (1998). Treatment research at the millennium: From efficacy to effectiveness. *Journal of Clinical Psychopharmacology, 18,* 1.

Levenson, H. (1995). *Time-limited dynamic psychotherapy: A guide to clinical practice*. New York: Basic Books.

Lueger, R. J. (1998). Using feedback on patient progress to predict outcome of psychotherapy. *Journal of Clinical Psychology, 54,* 383–393.

Margraf, J., Ehlers, A., Roth, W. T., Clark, D. B., Sheikh, J., Agras, W. S., & Taylor, C. B. (1991). How "blind" are double-blind studies? *Journal of Consulting and Clinical Psychology, 59,* 184–187.

McGrew, J. H., Bond, G. R., Dietzen, L., & Salyers, M. (1994). Measuring the fidelity of implementation of a mental health program model. *Journal of Consulting and Clinical Psychology, 62,* 670–678.

Mintz, J., Drake, R. E., & Crits-Christoph, P. (1996). Efficacy and effectiveness of psychotherapy: Two paradigms, one science. *American Psychologist, 51,* 1084–1085.

Multiple Risk Factor Intervention Trial Group. (1977). Statistical design considerations in the NHLI multiple risk factor intervention trial (MRFIT). *Journal of Chronic Disease, 30,* 261–275.

Ney, P. G., Collins, C., & Spensor, C. (1986). Double blind: Double talk or are there ways to do better research. *Medical Hypotheses, 21,* 119–126.

Newman, F. L., & Tejeda, M. J. (1996). The need for research that is designed to support decisions in the delivery of mental health services. *American Psychologist, 51,* 1040–1049.

O'Leary, K. D., & Borkovec, T. D. (1978). Conceptual, methodological, and ethical problems of placebo groups in psychotherapy research. *American Psychologist, 33,* 821–830.

Orlinsky, D. E., Grawe, K., & Parks, B. K. (1994). Process and outcome in psychotherapy—Noch einmal. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (pp. 270–376). New York: Wiley.

Orlinsky, D. E., & Howard, K. I. (1986). Process and outcome in psychotherapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed. pp. 311–381). New York: Wiley.

Parloff, M. B. (1986). Placebo controls in psychotherapy research: a sine qua non or a placebo for research problems? *Journal of Consulting and Clinical Psychology, 54,* 79–87.

Rabkin, J. G., Markowitz, J. S., Stewart, J., McGrath, P., Harrison, W., Quitkin, F. M., & Klein, D. F. (1986). How blind is blind? Assessment of patient and doctor medication guesses in a placebo-controlled trial of imipramine and phenelzine. *Psychiatry Research, 19,* 75–86.

Rice, L. N., & Greenberg, L. S. (1984). *Patterns of change: Intensive analysis of psychotherapy process.* New York: Guilford Press.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis.* New York: McGraw-Hill.

Rosnow, R. L., & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist, 44,* 1276–1284.

Roth, A., & Fonagy, P. (1996). *What works for whom? A critical review of psychotherapy research.* New York: Guilford Press.

Russell, R. L. (Ed.) (1987). *Language in psychotherapy: Strategies of discovery.* New York: Plenum.

Seligman, M. (1995). The effectiveness of psychotherapy. The *Consumer Report* study. *American Psychologist, 50,* 965–974.

Seligman, M. (1996). Science as an ally of practice. *American Psychologist, 51,* 1072–1079.

Shadish, W. R. (1995). The logic of generalization: Five priciples common to experiments and ethnographies. *American Journal of Community Psychology, 23,* 419–428.

Shadish, W. R., Matt, G. E., Navarro, A. N., Siegle, G., Crits-Christoph, P., Hazelrigg, M. D., Jorm, A. F., Lyons, L. C., Nietzel, M. T., Prout, H. T., Robinson, L., Smith, M. L., Svartberg, M., & Weiss, B. (1997). Evidence that therapy works in clinically representative conditions. *Journal of Consulting and Clinical Psychology, 65,* 355–365.

Shaw, B. F. (1984). Specification of the training and evaluation of cognitive therapists for outcome studies. In J. B. W. Williams & R. L. Spitzer (Eds.), *Psychotherapy research: Where are we and where should we go*? (pp. 173–189). New York: Guilford Press.

Schulberg, H. C., Block, M. R., Madonia, M. J., Scott, P., Rodriguez, E., Imber, S. D., Perel, J., Lave, J., Houck, P. R., & Coulehan, J. L. (1996). Treating major depression in primary care practice: Eight-month clinical outcomes. *Archives of General Psychiatry, 53,* 913–919.

Speer, D. C. (1994). Can treatment research inform decision makers? Nonexperimental method issues and examples among older outpatients. *Journal of Consulting and Clinical Psychology, 62,* 560–568.

Strupp, H. H., & Hadley, S. (1979) Specific vs. nonspecific factors in psychotherapy. *Archives of General Psychotherapy, 36,* 1125–1136.

Task Force on Promotion and Dissemination of Psychological Procedures (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist, 48,* 3–24.

Waltz, J., Addis, M. E., Koerner, K., & Jacobson, N. S. (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology, 61,* 620–630.

Wei, L. J. (1978). An application of an urn model to the design of sequential controlled trials. *Journal of the American Statistical Association, 73,* 559–563.

Wells, K. B., & Sturm, R. (1996). Informing the policy process: From efficacy to effectiveness data on pharmacotherapy. *Journal of Consulting and Clinical Psychology, 64,* 638–645.

Wilson, G. T. (1996). Manual-based treatments: The clinical application of research findings. *Behavior Research and Therapy, 34,* 295–314.

Wolff, N., Helminiak, T. W., & Tebes, J. K. (1997). Getting the cost right in cost-effectiveness analyses. *American Journal of Psychiatry, 154,* 736–743.

Woody, S. R., & Kihlstrom, L. C. (1997). Outcomes, quality, and cost: Integrating psychotherapy and mental health services research. *Psychotherapy Research, 7,* 365–381.

Yates, B. T. (1996). *Analyzing costs, procedures, process, and outcomes in human services.* Thousand Oaks, CA: Sage.

Yates, B. T. (1997). From psychotherapy research to cost–outcome research: What resources are necessary to implement which therapy procedures that change what processes to yield which outcomes? *Psychotherapy Research, 7,* 345–364.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

3.11
Mental Health Services Research

WILLIAM A. HARGREAVES
University of California, San Francisco, CA, USA
RALPH A. CATALANO, TEH-WEI HU
University of California, Berkeley, CA, USA
and
BRIAN CUFFEL
United Behavioral Health, San Francisco, CA, USA

3.11.1 INTRODUCTION 225
3.11.2 EFFICACY RESEARCH 226
  3.11.2.1 Examples of Efficacy Research 228
3.11.3 COST-OUTCOME RESEARCH 229
  3.11.3.1 Examples of Cost-outcome Studies 231
3.11.4 SERVICE SYSTEM RESEARCH 233
3.11.5 POLICY RESEARCH 238
3.11.6 CONCLUSION 239
3.11.7 REFERENCES 240

3.11.1 INTRODUCTION

The US National Institute of Mental Health blueprint for improving services (NIMH, 1991; Attkisson et al., 1992; Lalley et al., 1992; Mechanic et al., 1992; Steinwachs et al., 1992) broadened and clarified the scope of mental health services research. New and experienced investigators trying to understand mental health services research, however, may find it hard to comprehend the multiple perspectives on the field and how seemingly disparate research agendas contribute to a unified body of knowledge. Nevertheless, a unified view of the field can be gained by reflecting on key constructs and the causal paths that connect them.

Mental health services research is inherently multidisciplinary, drawing from clinical psychology, psychiatry, health economics, epidemiology, sociology, political science, social work, nursing, pharmacology, pharmacy, general medicine, and public health. This multidisciplinary character is reflected in the authorship of this article, including two clinical psychologists, an urban planner, and a health economist. Many clinical psychologists are contributing to mental health services research and more are needed.

In spite of great heterogeneity of methodology, all of services research can be viewed as examining some aspect of a nested system of causal processes that influence the mental health and well being of society. At the inner core are the effects of specific services on individual users of these services. These effects are contained within the broader effects of the way services are organized into local or other end-user service systems, how services are financed, and how these services systems impact relevant target populations. These effects of the organization and financing of services are in turn influenced by policy innovations of local, state, national, and international scope that influence the functioning of local service systems. These three domains of effects correspond to different types of services research: cost-outcome studies, service system studies, and policy studies.

The inner core of services research, cost-outcome studies, in turn grew out of an established tradition of clinical efficacy research utilizing the randomized clinical trial. The Agency for Health Care Policy Research of the US Government, in an attempt to foster services research, popularized a distinction between efficacy and effectiveness research in the evaluation of medical technologies (Salive, Mayfield, & Weissman, 1990). Efficacy studies evaluate the effect of an intervention under optimal conditions, when implemented by highly skilled clinicians using comprehensive treatment protocols under expert supervision, and where the intervention is usually delivered to patients with a single, well-defined disorder. Effectiveness studies complement the findings of efficacy studies by evaluating interventions in conditions, settings, and target populations that reflect the usual circumstances in the delivery of services, often giving attention to long-term effects of treatments on both costs and outcomes. The term "cost-outcome" is used to avoid confusion between "effectiveness" studies and "cost-effectiveness" studies, since the latter use a specific subset of cost-outcome methods.

Cost-outcome research remains conceptually and methodologically linked to efficacy research. The refinement of clinical trial methodology over several decades has established a useful state of the art that includes random assignment to treatment, appropriate blinding, complete follow-up with "intent-to-treat" analyses, and improved statistical handling of longitudinal data with missing or unevenly spaced measurement occasions, as discussed in Chapter 3, this volume. Cost-outcome studies involve the extension and improvement of clinical trials methodology necessary to address broader questions of generalizability, accessibility, acceptability, cost, and long-term outcome of interventions.

This chapter contrasts efficacy research and the three types of services research on several dimensions, and gives examples of each. These research models differ on a number of features: (i) their typical unit of analysis; (ii) the outcomes of primary interest; (iii) their typical design strategies; (iv) the independent variables usually experimentally controlled or sampled by the investigator; and (v) the hypothesized causal variables of greatest interest. Table 1 summarizes these features for each of the four models. As each model is discussed it will illustrate the causal pathways of usual interest, as portrayed in Figures 1 to 4.

3.11.2 EFFICACY RESEARCH

Efficacy studies examine the effects of a well-defined treatment or other service on the clinical and rehabilitation outcomes of individual recipients of care. Figure 1 is familiar to clinical trials investigators and shows the key groups of variables and causal paths of primary interest. Fifteen causal paths, numbered A1 to A15, encompass the scientific focus of efficacy research.

The ultimate outcomes of efficacy studies are reductions in illness symptoms, changes in behavior, or improvements in skills. Outcomes are usually measured over weeks or months rather than years. The unit of analysis is the person who consumes services. The typical study will compare an innovative treatment to either a placebo, an accepted "standard" treatment, or some other control for attention and expectation effects. Double-blind controls are employed to minimize bias that might be caused by expectation effects on outcomes and on judgments of outcomes. When the blind is likely to "leak" or is otherwise impossible to maintain, blinding of interviewers and raters is sometimes attempted as a partial effort to reduce bias.

The causal chain of primary interest is usually causal path A10, the effect of the implemented service on ultimate outcomes. Investigators may also be interested in the mediation of this effect through the use and "receipt" of service and through proximal outcomes. Thus, several aspects of the whole cascade of effects A1 to A8 may be examined. Even more broadly, efficacy studies may take account of the characteristics of the person at baseline such as diagnosis, severity of disability, history of treatment compliance, and stage of illness in order to understand their modulation of intervention effects (A11 and A13). Similarly, the effects of the person's preservice life context (e.g., homelessness, poverty, social support system, entitlements, insurance coverage) may be examined, along with life context during and after service (A14 to A15). In efficacy studies,

Table 1 Typical features of different research models.

	Research model			
Feature	*Efficacy*	*Cost-outcome*	*Service system*	*Policy*
Unit of analysis	Individual representing a defined disorder group	Individual representing a system target population	Individuals in target population Aggregate subgroup status, use, or change	Community service systems
Outcomes of primary interest	Symptom improvement Rehabilitative outcome Humanistic outcome Public safety	Efficacy outcomes Societal cost and payer mix Preference-weighted outcomes or utilities	Service demand, access, and use Service cost Prevalence of psychiatric impairment	Local service policy adoption, innovation Service cost Prevalence of psychiatric impairment
Typical design strategies	Randomized clinical trial	Randomized trial with epidemiological specification of target population Quasi-experiment	Policy experiment Population survey Org-network analysis MIS analysis Economic modeling and computer simulation	Policy experiment Population survey Org-network analysis MIS analysis Economic modeling and computer simulation
Variables experimentally controlled or set by investigator	Mental health system environment Clinical environment and service context Baseline characteristics of subjects Amount, type and delivery of services to study subjects	Mental health system environment Clinical environment and service context Characteristics of available services	Mental health system environment Policy innovation process	Population demographics Government structure
Hypothesized causal variables of interest	Service delivery and receipt	Service delivery and receipt	Local system organization, financing, management, and innovation.	Policy innovation above the local level

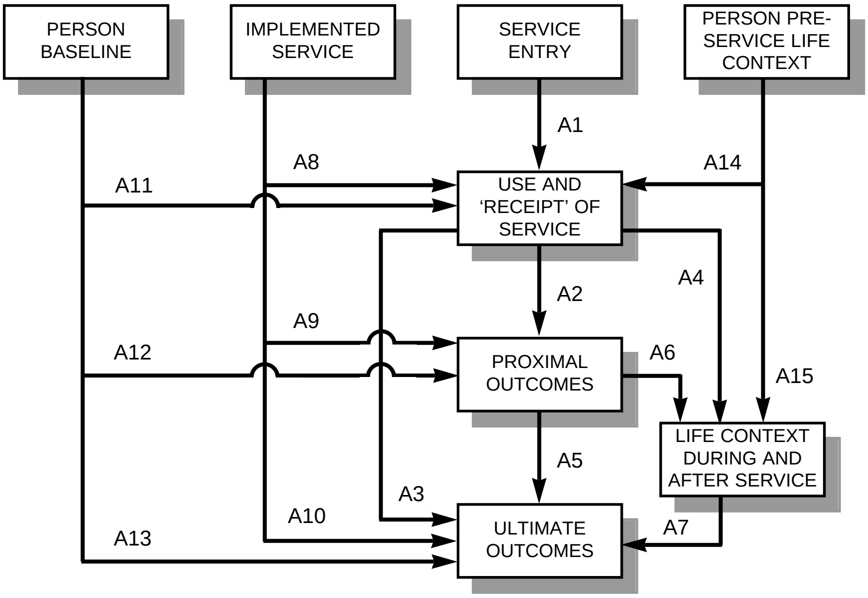

Figure 1 Efficacy studies.

however, the goal is usually focussed on causal path A10 while other causal paths are viewed as "nuisance" effects or "artifacts" that threaten a valid assessment of path A10.

The typical design strategy is the randomized trial. The design of efficacy studies typically emphasizes the "internal" validity of the estimate of the effect of a treatment, taking care to exclude alternative explanations for the observed effects, such as baseline or post-attrition differences in persons receiving different services.

Efficacy studies may focus on the mechanism of action of the intervention. Mechanism can often be inferred from analyses of "intermediate" variables that may account for the service effect, such as service use and "receipt" (e.g., medication or psychosocial treatment adherence, regularity of appointments, attitude toward the services), proximal outcomes (e.g., learning a communication skill, immediate symptomatic response to a medication), and changes in living setting or social support system resulting from the intervention.

The quest for internal validity and adequate statistical power often leads to considerable sacrifice of "external" validity, or generalizability. Service entry by study subjects is taken as a given, and usually little effort is made to relate the characteristics of the subject sample to those of the target population. Indeed, the usual efficacy study purposely employs narrow criteria to produce a homogeneous sample of subjects with few "distracting" sources of variance, such as comorbid illnesses, that may be quite prevalent in the population likely to receive the innovative intervention. Subjects are recruited from some convenient source that can yield enough subjects who qualify and will consent, making the assumption (usually unstated and unexamined) that the clients entering screening represent the target population on the dimensions that matter for generalization.

3.11.2.1 Examples of Efficacy Research

The National Institute of Mental Health "Treatment Strategies in Schizophrenia" cooperative study is an example of a late-stage, complex efficacy study. Schooler, Keith, Severe, & Matthews (1989) and Schooler et al. (1997) compared maintenance treatment strategies for people with a recent schizophrenia relapse whose family is in the picture. The study assessed the impact of dose reduction of antipsychotic medication and family treatment on relapse and rehospitalization during maintenance treatment (paths A1 to A15 in Figure 1). Standard dose medication was compared to continuous low-dose or targeted treatment, in which active medication was used only upon incipient signs of relapse. A widely studied and effective type of family intervention was compared to a much less intensive family intervention in the context of assertive medication management and ongoing supportive assistance to patients and families. It was hypothesized that the more intensive family intervention would make the conservative low-dose and targeted medication strategies more effective, perhaps as effective as standard-dose but with less burden from medication side-effects. The three-by-two design thus produced six treatment combinations, and the interaction between dosing strategy and family intervention level was of primary interest. The investigators found that both continuous low-dose and targeted medication increased the use of "rescue" medication and the risk of symptomatic relapse, but only targeted treatment increased hospitalization compared to standard-dose medication. The two family interventions showed essentially no differences in effectiveness, and family intervention did not interact with dose effects. There were two important policy implications of the findings: (i) targeted treatment, which had been advocated by some investigators, should not be adopted as a standard maintenance strategy, and (ii) the intensive family interventions that had been shown in previous research to be effective compared to no family intervention are more intensive than needed when provided in the context of good continuity of maintenance treatment and initial education and support of family members.

Like most efficacy studies, however, the primary focus was on mechanism of action rather than policy generalization. All sites were university-affiliated clinics, not sites picked to represent the "usual" circumstances of schizophrenia treatment. The design limited the ability to generalize to patients who did not reconstitute within six months of their most recent psychotic episode.

Nevertheless this study also had some features of cost-outcome research. "Ordinary" clinicians were recruited and trained as family clinicians, the study was carried out in five sites allowing a test of whether the findings generalized across site, and costs were assessed, although cost analysis was not part of the original design.

The methodology for efficacy research is well developed and relies heavily on the randomized clinical trial. The field reached a reasonable consensus on efficacy methodology in clinical psychopharmacology in the 1970s, and in psychotherapy efficacy research since the mid-1980s.

3.11.3 COST-OUTCOME RESEARCH

Cost-outcome studies of mental health services expand the scope of efficacy research to address key questions related to the effects of mental health services delivered in typical care settings. Figure 2 illustrates the expanded causal paths examined in cost-outcome studies. These include the same causal paths (A1 to A15) from efficacy studies (Figure 1), but include additional factors not addressed in efficacy trials (B1 to B21). Access to services is often examined in cost-outcome trials (Paths B1 to B9), and cost-outcome studies always measure the use and cost of services, including services other than the experimental interventions themselves (B10, B14, and B15).

The ultimate outcomes of cost-outcome studies are broad, including clinical outcomes (illness symptoms, side effects), rehabilitative outcomes (productivity), humanistic outcomes (life satisfaction, family psychological burden), and public safety outcomes (disruptive behavior, assaultiveness, suicidality, criminal behavior). Cost-outcome studies also may consider determinants of outcomes resulting from a wider range of causes including program case mix, consumer life context, and service cost/payer mix (B3, B7 to B9).

More importantly, perhaps, there is a fundamental shift in the perspective on outcomes in cost-outcome studies (Hargreaves, Shumway, Hu, & Cuffel, 1997). The ubiquitous fact of multiple, poorly correlated outcomes in mental disorders creates a problem that cannot be ignored in cost-outcome studies, since policy conclusions regarding the choice of alternative interventions depend on the relation of overall cost to overall outcome.

Different approaches to integrating outcomes are reflected in four different types of cost-outcome studies Cost-minimization studies examine only costs, in situations in which outcomes are known or assumed to be equal (B10, B14, B15). Cost-benefit studies focus on outcomes that can be monetized, such as labor force productivity, and analyze either cost-benefit ratios or net benefits (benefits minus costs) (B11, B16, B17). Cost-effectiveness studies leave outcomes in their natural units, and cost-effectiveness ratios are computed either on single outcomes (e.g., cost per suicide prevented) or on a global or integrated outcome score (e.g., cost per one unit of level-of-functioning score gained) (B12, B18, B19). Cost-utility studies utilize cost-utility ratios, where utility is a preference-weighted unitary outcome (e.g., cost per quality-adjusted life year gained) (B13, B20, B21). All of these studies are comparative—no intervention is "cost-effective" by itself. The cost-outcome performance of an intervention is always judged against a policy-relevant standard, such as treatment as usual. Cost-outcome ratios are not needed to compare interventions when the less costly intervention produces better outcomes. When an innovation is both more expensive and more

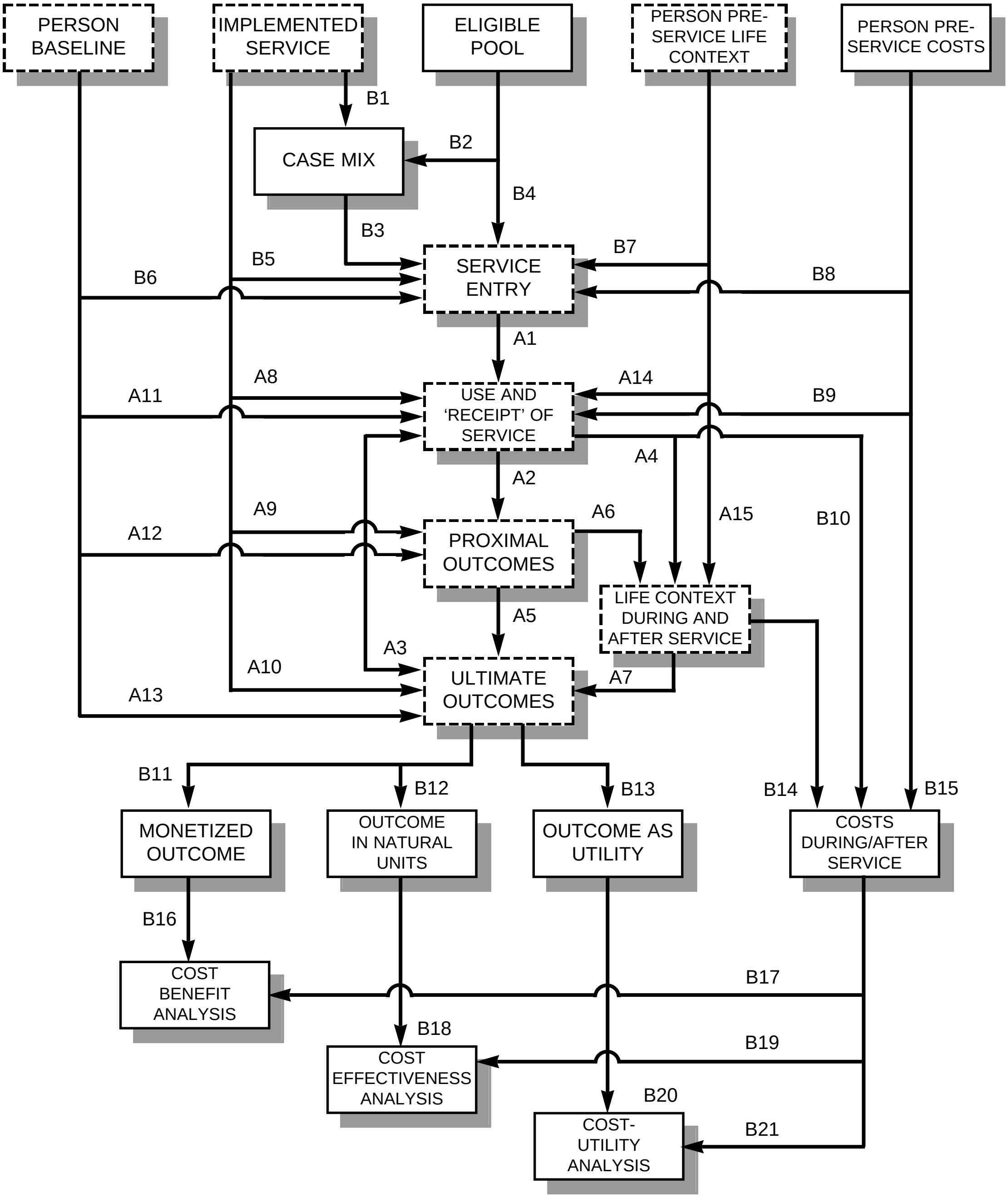
PERSON BASELINE
IMPLEMENTED SERVICE
ELIGIBLE POOL
PERSON PRE-SERVICE LIFE CONTEXT
PERSON PRE-SERVICE COSTS
B1
B2
CASE MIX
B4
B3
B7
B5
B6
SERVICE ENTRY
B8
A1
A8
A14
A11
USE AND 'RECEIPT' OF SERVICE
B9
A4
A2
A9
A6
A15
B10
A12
PROXIMAL OUTCOMES
LIFE CONTEXT DURING AND AFTER SERVICE
A5
A3
A10
A7
A13
ULTIMATE OUTCOMES
B11
B12
B13
B14
B15
MONETIZED OUTCOME
OUTCOME IN NATURAL UNITS
OUTCOME AS UTILITY
COSTS DURING/AFTER SERVICE
B16
B17
COST BENEFIT ANALYSIS
B18
B19
COST EFFECTIVENESS ANALYSIS
B20
B21
COST-UTILITY ANALYSIS

Figure 2 Cost-outcome studies.

effective than standard care, a common type of cost-outcome ratio compares the differential cost per differential outcome gained by the innovation compared to standard care.

Cost, as a consequence of an intervention, has its own measurement technology that has been carefully worked out in econometrics and health economics (Drummond, Brandt, Luce, & Rovira, 1993; Drummand, Stoddard, & Torrance, 1997; Sloan, 1995). Economic theory provides the conceptual underpinning of cost estimation methods. In a theoretically ideal competitive market, goods and services are exchanged at prices that reflect their true value. In the real world market quality varies, and in health care the operation of a free, competitive market is impaired because buyers and sellers have inadequate information, and because third-party payers and government regulation also introduce market distortions. The theory of welfare economics provides an alternative way to estimate the true value of goods and services when market pricing is not accurate. Under these circumstances the true value is estimated by considering the "opportunity cost" of resources. This value estimate, called a "shadow price," is the value of a resource were it to be used for its "next-best" purpose.

Economists also developed the concept of "societal" cost to describe a cost perspective of an entire economic system, without regard to who pays the cost. Societal cost is accepted widely as a standard reference method in cost-outcome research (Drummond et al., 1997). Most policy questions, however, also need to be informed by information about cost burdens to particular payers. For example, if a new intervention results in the same cost as usual care (from a societal perspective), but shifts the payment burden from the state and federal government to local government, or from government to the ill person's family, the payer information is important. Economists call shifts in payer burden a "distributional effect" of the innovative intervention, since it distributes the grief in a new way.

The difference in societal cost of two intervention strategies is the difference of the value of resources consumed or lost. Resources consumed (e.g., treatment activities) are called "direct" costs, while resources lost (e.g., time prevented from working, which produces resources) are called "indirect" costs. The amounts of resources are measured in natural units (hours of psychologist time, days of hospitalization), but values are assigned to these resource units by a separate process. In this way, questions about the accuracy of methods for measuring resource use can be examined separately from questions about the appropriateness of the values assigned to each type of resource.

Seven types of cost elements are often considered when planning mental health cost-outcome research: mental health treatment, physical health care, criminal justice procedures, social services, time and productivity of the patient and of family caregivers, other family costs, and sources and amounts of patient income. Different interventions and target populations vary greatly in the relative emphasis that needs to be given these seven types of cost. The most accurate measures of consumption and the most careful estimates of value are usually reserved for the most commonly used resources, especially the new activities introduced in the innovative intervention being studied.

Expert judgment is required to choose an appropriate balance of methods and to apply them skillfully. In most situations neither health economists nor clinical investigators can do it alone, and effective collaboration is required. Hargreaves et al. (1997) provide an introduction to cost-outcome methods for mental health research, designed to prepare noneconomist investigators for this type of collaboration.

The unit of analysis in cost-outcome studies is still the person, but attention is given to the person as a member of a target population, the paths from target population membership to service, the generalization of the findings to this target population, and even the impact of service provision or access policies on the whole target population.

The typical design strategy is the randomized cost-effectiveness trial focused on both internal and external validity, comparing a service innovation to standard services, and estimating the impact on the total target population as well as individual members of that target population. Ideally the size and character of the target population is estimated, and a random sample is invited to participate in the study. More often, a series of analyses tracks the biases in selection from target population, to potential subject pool, to recruited subject sample, to analyzable subject sample.

In these decision-oriented studies, certain "standard" features of efficacy trials, such as double-blind conditions, may be avoided in order to enhance external validity. For example, an open study does not require sham procedures such as unnecessary blood samples, or active placebos that mimic side effects, both of which may obscure the relative acceptability of services being compared. This may increase risks to internal validity, as when an open study allows rater bias and alters client expectation effects. These problems should not be ignored, although they are less troublesome when conclusions hinge on long-term social and vocation functioning improvement rather than short-term symptom improvement.

Despite the broader scope of cost-outcome studies, several types of variables are taken as given or set by the investigator, including the service model, the character of the target population, the consumer's baseline clinical characteristics and life context, and preservice utilization and costs. In service system research, by contrast, it will be seen that investigators may view these variables as outcomes.

3.11.3.1 Examples of Cost-outcome Studies

A study of assertive community treatment was an exemplary cost-benefit study (Stein & Test, 1980; Test & Stein, 1980; Weisbrod, 1983; Weisbrod, Test, & Stein, 1980). In this study the experimental treatment was applied for 14 months. Subjects in the experimental program had less time in the hospital, more time in independent living situations, better employment records, increased life satisfaction, decreased symptoms, and improved treatment adherence. Family and community burden did not appear to increase (A1, A3, A4). Societal cost was found to be slightly lower in the experimental treatment when wages earned were considered monetary benefits and welfare support was considered a cost (B10, B11, B14 to B17). Thus, the innovative program showed both an economic advantage and a psychosocial advantage.

Research on assertive community treatment (ACT) for persons with severe mental illness is the strongest and most coherent body of cost-effectiveness investigation in mental health. More than a dozen additional randomized trials of ACT were published from 1986 to 1996 and the findings have been critiqued and summarized in many reviews (Bond, McGrew, & Fekete, 1995; Burns & Santos, 1995; Chamberlain & Rapp, 1991; Hargreaves & Shumway, 1989; Olfson, 1990; Rubin, 1992; Santos, Henggeler, Burns, Arana, & Meisler, 1995; Scott & Dixon, 1995; Solomon, 1992; Test, 1992). Studies in the 1990s utilize stronger usual care conditions, reflecting the widespread adoption of linkage case management in the USA (Burns & Santos, 1995; Randolph, 1992). These studies are refining understanding of the conditions and target groups in which assertive community treatment is cost-efficient or has such superior outcome in spite of increased cost that it can be considered cost-effective relative to available service alternatives.

A randomized cost-effectiveness study of the atypical antipsychotic clozapine for persons with schizophrenia was carried out by Essock; (Essock, Hargreaves, Covell, & Goethe, 1996a; Essock et al., 1996b). It featured a complete enumeration of clozapine-eligible patients in an entire state hospital system, and a mix of random and purposive selection of members of the eligible pool to be offered study entry during a time when clozapine was rationed by the state hospital system and study participation was essentially the only route of access to clozapine treatment (B1 to B9). Study subjects were randomized to clozapine or usual antipsychotic medication in an open study (A1 to A15). An effectiveness and cost evaluation was carried out over two years of treatment (B10, B12, B14, B15, B18, B19). Clozapine showed fewer side effects with some advantage in reduced restrictiveness of care and reduced assaultive or disruptive behavior. In this long-stay state hospital patient group (mean length of their hospital episode was over eight years) no effect was seen on time from study entry to hospital discharge. In contrast, once discharged, clozapine patients were much less likely to be readmitted, resulting in saved hospital days. Full cost analyses have not been completed, but saved hospital costs may fully offset the higher price of the novel antipsychotic medication plus the increased cost of community care when patients are out of the hospital. This study illustrates how the results of an effectiveness trial comparing a new treatment to usual practice can be more conservative than an efficacy trial comparing a novel medication to a single alternative. Previous efficacy trials had shown larger effects on symptom improvement, typically in patients who were in an acute illness exacerbation, not the relatively stable but still symptomatic long-stay state hospital patients who showed lower baseline scores on symptom scales.

A study by Henggeler, Melton, and Smith (1992) illustrates how research can have an intermediate position between efficacy and cost-outcome research. Henggeler studied "multisystemic therapy" (MST) as an intervention for serious juvenile offenders and their multiproblem families. Yoked pairs of qualified subjects randomized to multisystemic therapy or usual services were referred by the South Carolina Department of Youth Services. Endpoints included family relations, peer relations, symptomatology, social competence, self-reported delinquency, number of arrests, and weeks of incarceration. Compared to youths who received usual services, youths who received MST had fewer arrests and self-reported offenses and spent an average of 10 fewer weeks incarcerated during a 59-week study period. Families assigned to MST reported increased family cohesion and decreased youth aggression in peer relations. Detailed cost analyses were not reported in this initial publication. This is one of the first controlled trials of interventions with this difficult target population to show significant beneficial outcomes and clinically important effects. The success of the intervention plausibly is related to its addressing each of the major factors found in multidimensional causal models of antisocial behavior in adolescents (Henggeler, 1991).

This study can be considered efficacy research in that it was delivered under close supervision of those who designed the intervention, and further evidence is needed that its effectiveness can be maintained when the intervention is taught and supervised by others. In addition, the statistical power of the initial study is limited, so while large and statistically significant effects were observed, effect size in the population still needs to be estimated more accurately before basic efficacy can be considered to be established. Features more like cost-outcome research include the construction of a subject pool that is plausibly representative of an important target population, and (except for service supervision) the intervention reflects much of the reality of providing such services in ordinary settings of care. The study can also provide preliminary information about the cost of the intervention and the cost savings that may result from its effective implementation.

The methodology for cost-outcome studies is a combination of cost-effectiveness methods from health economics, mental health epidemiology methods, and randomized controlled trial methods, applied to intervention processes that sometimes take place over several years. The merger seems to require ambitious and expensive multidisciplinary studies, but that conclusion is probably incorrect. The new investigator can work in this area by learning the methods and gathering appropriate consultants and collaborators. A workable state of the art is reasonably well advanced and accessible (Drummond et al., 1997; Hargreaves et al., 1997; Knapp, 1995; Sloan, 1995; Yates, 1995).

3.11.4 SERVICE SYSTEM RESEARCH

Research on local mental health service systems examines the effects of organization, financing, and service delivery mechanisms on utilization of services, costs, outcomes, cost-effectiveness, and community prevalence. The object of study, mental health service systems and community outcomes, are of broader scope than phenomena affecting the relative cost effectiveness of a circumscribed group of services for a specific problem or disorder. Therefore, cost-outcome research on mental health systems draws on economics, organizational sociology, epidemiology, and political science for explanatory theory and method. A mental health system serves an array of target populations whose service needs differ, yet these target populations also compete for many of the same services. Therefore it is reasonable to view system impact as its effect on a value-weighted sum of its impact on the prevalence of the array of mental disorders and their resulting impairments.

The ultimate outcomes in service system research are the prevalence and disability consequences of various mental disorders in the population. Intermediate outcomes include demand for mental health services, access and use of those services, and outcomes of those services. More broadly, the ultimate outcome is seen as the quality of the life of all members of the community, to the extent that it is affected by psychiatric disorders. Mental health systems and system managers work within a political, fiscal, and organizational context to focus service resources on high priority target populations so that the overall cost-effectiveness or cost-utility of the system is maximized.

Service system utility implies value judgments not only about trade-offs in outcomes for a particular target population, but trade-offs in the value of higher levels of functioning in one target population as against another. For example, after many years during which community programs were focused on improving the level of functioning of clients and families with serious life problems but who were not severely and persistently mentally ill, these values were called into question, and service programs shifted their emphasis to the latter group. While many factors influenced this change (Mechanic, 1987), including a rise in the relative cost of hospital care, the change also reflected an implicit value judgment that serving the severely and persistently mentally ill in the community was of overriding importance to the overall utility of the mental health service system. This implicit utility judgement could never be examined explicitly, because the complexity of the problem exceeded capacity to study it. An effort to identify the parameters of this problem may clarify the advances in theory and method needed to address such issues.

The system dynamics of prevalence, demand, utilization, and outcome can be seen as interacting with the service process, which in turn interacts with the system structure. Process and structure subsystems may experience innovation or innovation may be introduced experimentally. In either case, innovation can give the investigator an opportunity to learn something about causal phenomena in the system.

A simplistic view of system dynamics (Figure 3) is that different levels of prevalence cause different levels of service seeking behavior, that service seeking leads to utilization, that utilization influences outcomes for individuals, and these outcomes change prevalence and disability. Some variables of interest in each of these subsystems can illustrate the possibility for more realistic models.

Prevalence is used here to include not only the local prevalence of persons with psychiatric disorders and subgroups of such persons, but the distribution within these populations of living independence/freedom, productivity, homelessness, life satisfaction, public safety, and the societal cost of illness. These valued

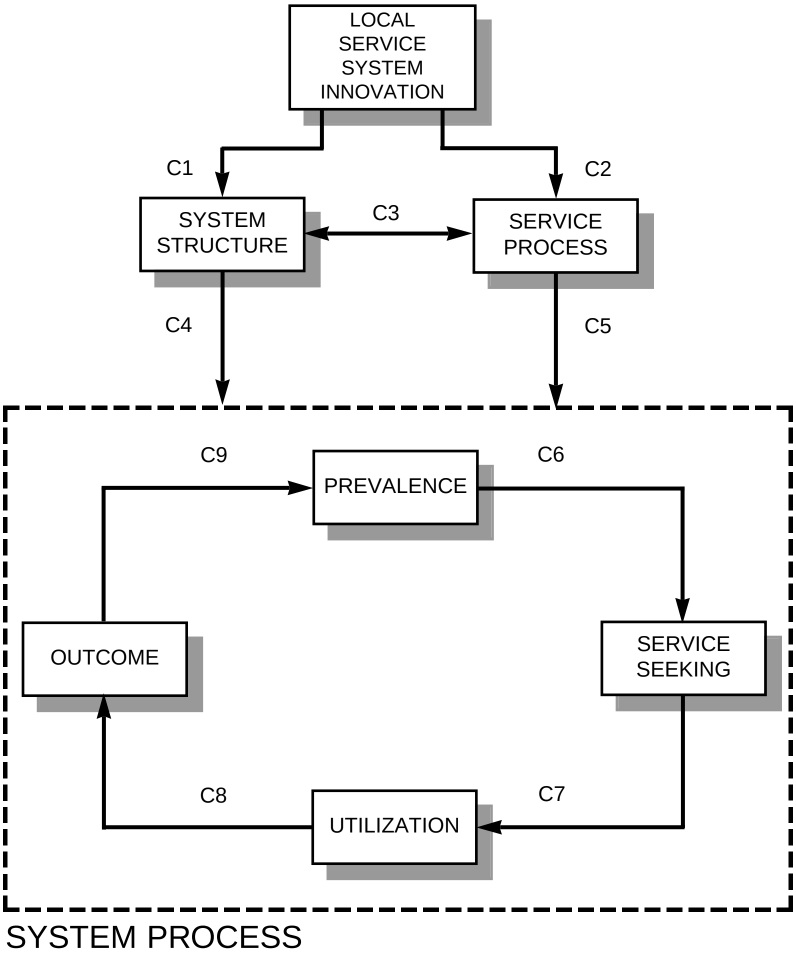

Figure 3 Service systems studies.

conditions and costs define the momentary utility of the human service system with regard to persons with psychiatric disorders. This utility can be influenced by effective human services. It can also be influenced by migration, system induced or otherwise. An innovation that incurs some marginal cost can be evaluated by comparing this cost to the change in utility that one is able to attribute to the innovation.

Priorities for various target populations and levels of disability are an unavoidable part of any definition of system utility. An egalitarian society aims to maximize cost utility without differential weighing of gender, age, socioeconomic status, ethnic group, residence location, or reason for disability. While real-world policy is formulated from narrower perspectives, it is useful for science to identify the larger impact of narrow policies. In the short term more modest research goals should be sought. Practical compromises may include restricting study samples to persons known by the service system to be in the target population, and considering only expenditures within the mental health system. This is a narrower perspective on target populations and cost than suggested for cost-outcome studies, in order to allow greater research investment to be focused on measuring a larger variety of intermediate effects and associated causes.

Service seeking results from experienced need and the lack of social supports (Alegria et al., 1991; Freeman et al., 1992), and is affected independently by economic conditions and job insecurity (Catalano, Rook, & Dooley, 1986). The need may be experienced by the ill person, by family members, or as a community response to illness-related disruptive behavior. Service seeking produces demand on each type of service provided in the system or seen as needed by each target population subgroup. There may be a pattern of over-capacity demand experienced by specific services, with demand-buffering patterns and consequent generation of secondary demand, that is, inappropriately high levels of care and lack of access to desired services. Cost-outcome and service system research should inform both providers and consumers about the relative value of alternative mental health services, enabling them to make informed choices about which services to supply or to consume.

Utilization is the actual amount of each type of service consumed by various subgroups of persons in the target population, and in some cases by other clients who compete for the same services. It is this consumption and its effects on outcome and cost compared to what would occur in the absence of service or with alternative services that are often the focus of research attention.

Outcome includes those changes in symptoms, functioning, life satisfaction, and public safety that result from the consumption of services. The strengthening of natural supports and the reduction of risk of relapse might be been as relevant outcomes. At the individual level these are the same variables as the prevalence conditions. Whether outcomes lead to changes in prevalence variables depends on the proportion of the target population that consumes appropriate and effective services, and the degree of offsetting changes that may occur in the rest of the target population.

The unit of analysis may be the individual for purposes of measurement although there is greater attention to the characteristics of the population from which samples are drawn. For example, service system research often attempts to make inferences about the effects of service system characteristics on mental health outcomes in the general population. In this regard, the unit of analysis can be the geographic area served by the local service system, particularly when the focus of study is aggregate statistics such as per capita rates of violence or suicide measured over time or across regions.

The causal chain of interest includes three broad aspects of local area service systems that are common sources of independent variables in service systems research. These are system structure, service process, and system innovation.

System structure includes a study of local system expenditures and expenditure decisions, the organization governance, networks, and environments, service system size, integration, and cohesiveness and local provider reimbursement policies, benefit eligibilities, risk-sharing incentives, and managed care strategies for monitoring and controlling the delivery of care.

Service process at the system level might be viewed as consisting of two types of variables: (i) service capacity mix (how many beds or slots of various types) and the unit costs of each type; and (ii) services integration methods such as linkage case management, assertive community treatment, entry and exit criteria, and treatment guidelines. For each service program or provider in the system one might examine management style, staff morale, and adherence to system-wide service delivery standards. Adequate measures of these aspects of the mental health delivery system need to be developed, especially such dynamic processes as eligibility criteria and service practices that shift in the demand-buffering process, or transient degradation of service capacity that may occur during rapid system change.

Local service system innovation can be viewed from a research perspective more as an event than a subsystem. Innovation, and indeed any change including financial disaster and dramatic downsizing, is of interest to the investigator because change offers the opportunity to study causal effects. This is particularly true when a similar change occurs or is introduced experimentally in a number of local mental health systems, allowing better strategies for distinguishing the effect of the change from other variations in system performance.

The typical design strategies that are employed in the study of local service systems draw from economics, organizational sociology, and epidemiology, in addition to traditional clinical research and cost-outcome research. Studies of the variability in local mental health policy often involve economic analysis of utilization and expenditure patterns across counties or other local geographical areas. Often system level changes resulting from imposed changes in conditions or local innovation facilitate naturalistic experiments, pre-post designs, or time series analyses that allow one to study the effects of these system changes on service seeking, use, and outcomes. States and the federal government in the USA are also seeking to learn from "pilot" innovation using experimental or quasi-experimental approaches, in collaboration with investigators.

The Substance Abuse and Mental Health Services Administration has sponsored a multistate experiment comparing two experimentally introduced innovations in local systems of care focused on homeless persons with dual disorders (severe mental illness and a substance use disorder). State mental health authorities were invited to propose pairs of sites to undertake a program to reduce homelessness among substance abusing persons with mental illness by increasing case management and housing supports. Both sites in each applicant state were to propose two specific innovations appropriate for their site: (i) methods for delivering augmented case finding and case management services, and (ii) methods to achieve increased system integration to enhance services for homeless persons with dual disorders. Site pairs were evaluated for their suitability for the demonstration and selected by reviewers. Selected pairs of sites received funding for initial engagement, case management, and follow-up of a comparable cohort of clients, while one site of each pair was randomly selected to receive additional "services integration" funding. Thus, this design addresses causal paths C1 to C5 and C6 to C8 in Figure 3, although it does not examine changes in the prevalence of this homeless subpopulation (C9). Preliminary results have been reported showing that the additional services integration funding augmented the effectiveness of services for the target population over and above the provision of augmented engagement and case management services (CMHS, 1996; Robert Rosenheck, personal communication, November 20, 1996).

Not all service system research requires a large budget. Some dramatic system changes occur with little warning but can be investigated retrospectively using time series methods. For example, Fenton, Catalano, and Hargreaves (1996) studied the impact of passage of Proposition 187, a ballot initiative in California mandating that undocumented immigrants be ineligible for most state-funded health services and requiring health providers to report persons suspected of being undocumented to the Immigration and Naturalization Service. Hispanic-Americans widely felt this as a racist initiative specifically directed at them. Though quickly blocked in court, publicity about the initiative and its level of support (passing in 50 of 58 counties) was hypothesized by the investigators to cause young Hispanics to avoid preventive psychiatric services (outpatient mental health treatment) putting them at risk for increased use of emergency services at some subsequent time. This hypothesis could be seen as involving causal path C6 in Figure 3, though (to anticipate) it also reaches up to causal paths D5 and D7 in Figure 4, discussed in the Policy Studies section.

The investigators analyzed data on weekly initiation of public outpatient mental health episodes and psychiatric emergency service contacts in the City and County of San Francisco for 67 weeks before and 23 weeks after the date of the election. Young Hispanics were compared to young non-Hispanic whites using Box-Jenkins time series methods. Initiation of outpatient services by young Hispanics dropped significantly the week of the election and remained low compared to use by non-Hispanic whites.

To refine their hypothesis of a delayed effect on emergency services use, the investigators examined the "natural" variation in the initiation of outpatient mental health services and its relation to emergency service episodes. Low rates of outpatient initiations were correlated significantly with high levels of emergency service episodes six weeks later in the client population as a whole. Looking at emergency use by young Hispanics (controlling for use by young non-Hispanic whites), the investigators found significantly higher use of emergency services six weeks after the election. Analyses of older populations showed neither of these effects of the election. Thus, the effect was specific to the typical age range of undocumented Hispanic residents. For a relatively non-technical introduction to time-series methods see Catalano and Serxner (1987), and for more detail see Box, Jenkins and Reinsel (1994).

Time-series analysis also can be used to test hypotheses about service system policies. In a provocative study, Catalano and McConnell (1996) tested the "quarantine" theory of involuntary commitment, which asserts that involuntary commitment of persons judged to be a danger to others reduces violent crime—a theory vigorously challenged by civil libertarians. This hypothesis could be seen as encompassing the entire causal loop in Figure 3 of paths C6 to C9.

No one has much enthusiasm for carrying out an experimental test of this theory. The investigators tested the specific hypothesis that the daily incidence of reported assaults and batteries by males is related negatively to the incidence of involuntary commitments of males as a danger to others during the preceding two days—the more men committed, the fewer the subsequent arrests. Involuntary commitments on the day of the reported assault were omitted because the direction of causality is ambiguous. Using data from the City and County of San Francisco for 303 days, the investigators carried out a time-series analysis that supported ($p < .05$) the conclusion that increased commitment of men judged a danger to others reduces the incidence of assault and battery offences. Score: paternalists, 1, civil libertarians, 0. The discussion continues.

These three examples illustrate that service system changes of interest may be introduced intentionally (a new intervention for homelessness), or arrive as unpredictable exogenous shocks (Proposition 187), or arise from endemic variations that all organizations exhibit by virtue of their inability to behave with perfect consistency over time (variation in the number of involuntary commitments). This third type of variation is of particular interest to managers of mental health systems. They manage organizations to minimize this inconsistency, but theories of rational management as well as common sense suggest that the effort to do so

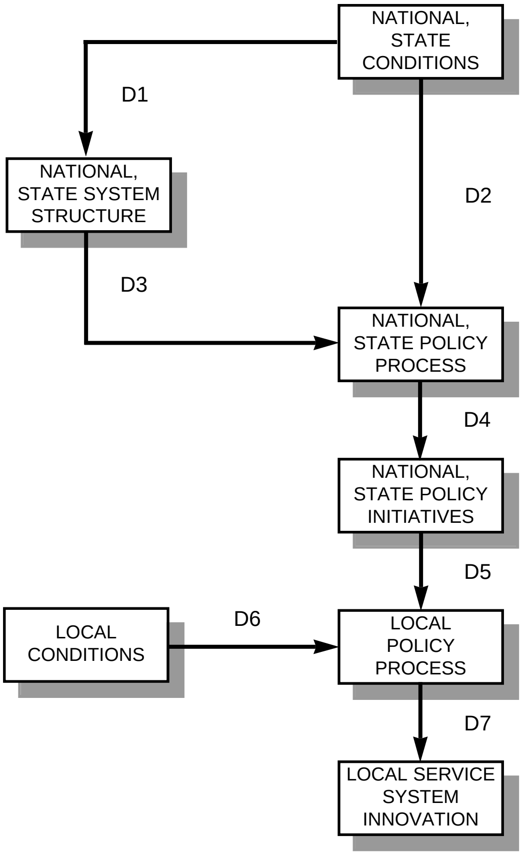
NATIONAL, STATE CONDITIONS
D1
NATIONAL, STATE SYSTEM STRUCTURE
D2
D3
NATIONAL, STATE POLICY PROCESS
D4
NATIONAL, STATE POLICY INITIATIVES
D5
LOCAL CONDITIONS
D6
LOCAL POLICY PROCESS
D7
LOCAL SERVICE SYSTEM INNOVATION

Figure 4 Policy studies.

should cost less than the benefit of greater standardization of output. Managed care is the health system's version of scientific management, and one of its consistency tools is the treatment guideline or algorithm.

Research on the impact of treatment guidelines and algorithms is an important and growing area of service systems research. Suppose an investigator wishes to test the hypothesis that new treatment guidelines improve outcomes while containing cost. The guidelines might be both restrictive (specifying when a procedure should not be provided even when sought by the client) and prescriptive (specifying when a procedure should be recommended proactively when the client fails to seek it). It is interesting to consider the design issues posed by this research goal.

To test the investigator's hypothesis for a particular disorder, the investigator needs to identify or articulate the relevant restrictive and prescriptive guidelines. Identifying potential health risks and benefits from these guidelines compared to usual practice will guide the choice of outcomes and the length of observation needed to detect each outcome.

In implementing the study the investigator needs to demonstrate that the implementation has induced optimal or near-optimal guideline adherence by health care providers, has induced it relatively rapidly, and has maintained adherence consistently over the long haul. This will usually require prompting and monitoring techniques, which are part of the cost of the intervention being tested. The investigator also will want to demonstrate that the usual-care control condition had not undergone similar practice changes. Another concern is that guidelines may alter the detection of the targeted health conditions. The investigator needs a way to determine whether detection of the target conditions is altered in either experimental or usual care. Given these possible diagnostic screening effects, the investigator has the problem of designing an unbiased way to select "tracer" patients to assay outcomes in both experimental and control conditions.

These competing threats to validity demand well-articulated hypotheses and carefully chosen design compromises. Rather than randomizing patients, an investigator may decide to randomize clinician practices or clinical sites to experimental and usual-care conditions. Cluster randomization often is used in educational research, where an innovative teaching method is randomized to teachers or classrooms (Murray et al., 1994). Thus, individual subjects are nested within the unit of randomization. This does not preclude examining individual baseline characteristics in the analysis, however.

Power is limited by the number of units randomized, but error variance is also reduced (and power increased) by increasing the number of subjects within each randomization unit. For each of the primary outcome measures, the power analysis informs the investigator about trade-offs between the number of randomization units and the number of subjects within each unit. The efficiency of the design will further depend on the cost of adding units versus the cost of adding subjects within units. Feasibility and cost constraints often limit the number of randomization units to a smaller than optimal number. Under these circumstances the investigator may be able to gain power if between-unit outcome variance can be reduced by using covariates (Raudenbush, 1997).

As these considerations illustrate, the important design issues in service systems research are the same as in efficacy and cost-outcome research. Intervention effects are most easily understood when interventions are introduced experimentally with some form of random assignment. Multistep causal processes can be studied most effectively with a clear theory about intermediate causal pathways and effects, and ways to demonstrate that expected intermediate effects are produced and potential confounding effects are avoided.

3.11.5 POLICY RESEARCH

National and state reform of the organization and finance of mental health systems is a major research issue in mental health economics. Financing processes and incentives are changing rapidly. Local government is being given increasing responsibility for providing mental health services for persons with severe mental illness, either through direct operation of the services or through contracting with the private sector. Mental health services research is examining the impact of these economic forces and changes at both the national and local levels. Figures 3 and 4 illustrate the linkage between national and state policy decisions and their effects on local service system innovation and adaptation. Research focuses on the impact of these changes on access to care, organizational adaptation, outcomes for the most vulnerable target populations, costs, and the changes in cost burden to various payers.

Figure 4 suggests that policy initiatives by national and state government agencies offer a useful research focus for the mental health services investigator. Examining the link between national/state initiatives and the cost-effectiveness of services at the client level involves "long-chain" causal paths in the extreme. Therefore, research designs need to be thoughtful and realistic. Some research questions may focus simply on the link from national/state policy initiatives to local service system innovation. This research strategy will be most defensible when a national policy aims to encourage specific types of local innovation and previous services research has demonstrated the value of those specific innovations in improving the cost effectiveness of services. Political scientists and organizational sociologists may also discover useful links among "national/state conditions," "national/state system structure," and the "national/state policy process" that generate policy initiatives. Such investigators may also fruitfully examine the comparable processes at the local level that influence local response to outside policy initiatives.

A third strategy is to take a national or state policy initiative, such as a state decision to enter into capitated contracts with local providers or managed care companies, and treat it as a case example in policy research, but to carry out an experimental or quasi-experimental study across local treatment systems of the impact on client access, client outcomes, and other endpoints. The evaluation of state adoption of managed care and capitation contracting in providing publicly funded mental health services in the USA is being evaluated in several states using this type of design.

Thus, the typical outcomes of national and state policy research are changes in local organization or financing of mental health services. It is also important, although more difficult, to determine the effect of national and state policy initiatives on the prevalence and consequences of mental disorder in the general population.

The causal chain of interest includes the effects of national and state conditions, system structure, and policy processes on those characteristics of local service systems thought to be related to their cost-effectiveness. The independent variables of causal interest to national and state policy research bear directly on local service systems as shown by the intersection of Figures 3 and 4 (C1 in Figure 3 and D5, D6, and D7 in Figure 4).

National/state conditions that influence the policy process include such variables as political stability, economic conditions, human resources mix, and cultural conditions such as family and community caretaking traditions. Corresponding factors presumably make up some of the relevant local conditions that influence local policy.

System structure at the national/state level includes such variables as centralization of control and funding, administrative standards, level of funding from sources above the local tax base, eligibility and benefit design, financing incentives for local innovation such as supply-side and demand-side payment incentives. In considering the economic concept of demand-side incentive as affecting local innovation behavior, copayment requirements imposed on persons with psychiatric illnesses and their families, might be included, and also copayment requirements that funding initiatives may impose on the local tax base.

Policy process at all levels includes such variables as leadership, research, and advocacy. Advocacy may come from research and service professionals, from industry, and from consumers and family members. Advocacy can be an important pathway from research findings to policy initiatives.

Mental health system reform in the USA, and specifically the introduction of capitated financing and competitive contracting for the provision of publicly funded mental health care for the severely and persistently mentally ill by many states, is of great interest (Brach, 1995). Several investigators are examining state policy initiatives such as carve-out capitation in mental health Medicaid, interventions related to causal paths D5 to D7 in Figure 4 (e.g., Christianson & Gray, 1994; Christianson & Osher, 1994; Dickey et al., 1996). Christianson, Gray, Kihlstrom, and Speckman (1995a) are studying Utah's prepaid mental health plan and have reported preliminary findings (Christianson et al., 1995b). There is also a study of Colorado's Medicaid carve-out capitated contracting for mental health services (Bloom et al., unpublished). Studies of state Medicaid innovations typically take advantage of a sequential conversion of different regions of a state from fee-for-service reimbursement to capitation contracts to study as a quasi-experiment.

In Colorado the state initially contracted with a for-profit contractor for coverage of a region of the state containing about one-third of the population, contracted with several existing community mental health centers to cover another third of the state's population, and delayed conversion in the rest of the state. Research assessments focus on organizational innovation and adaptation, service utilization, access to services, and individual client outcomes using a pre-post design in each of the three regions. Studies are in progress in several other states. Studies like this of statewide mental health system reform are of necessity relatively large projects, but state mental health authorities, services research investigators, and research funding agencies have joined together to mount major research efforts. By the end of the twentieth century the field will have a set of findings about cost control and outcomes that will set the policy and research agenda for the decade that follows.

As in the other areas of services research, small studies also make important contributions. Cuffel, Wait, and Head (1994) studied a 1990 innovation in the state of Arkansas in which the state transferred hospital funding and admission authority to local mental health agencies, essentially putting state hospitals into a competitive market. The investigators found, as expected, that state hospital use was reduced significantly after decentralization, although the reduction in urban areas was proportionally greater than in rural areas. Contrary to expectation, admissions were not limited to the most severely ill, disruptive, or substance-abusing patients, nor were discharged patients more likely to be readmitted. For patients treated in the state hospitals, communication between the community and the state hospital increased after decentralization.

Studies of national and state innovations such as the examples above make this aspect of mental health services research both feasible and increasingly important in guiding the evolution of mental health care.

3.11.6 CONCLUSION

Mental health services research encompasses a broad range of mental health services research projects. No one project can ordinarily study the whole system of variables at any of the three levels discussed in this Chapter. The scope of mental health services research, however, invites construction of integrated theory, especially in service system and policy domains. Such theory can lay the groundwork for a coordinated approach to empirical research. Furthermore, in spite of its great diversity, mental health services research shares a common objective to understand and improve the effectiveness and efficiency of mental health services for individual consumers and for entire target populations.

3.11.7 REFERENCES

Alegria, M., Robles, R., Freeman, D. H., Vera, M., Jimenez, A. L., Rios, C., & Rios R. (1991). Patterns of mental health utilization among island Puerto Rican poor. *American Journal of Public Health*, *81*, 875–879.

Attkisson, C., Cook, J., Karno, M., Lehman, A., McGlashan, T. H., Meltzer, H. Y., O'Connor, M., Richardson, D., Rosenblatt, A., Wells, K., Williams, J., & Hohmann, A. A. (1992). Clinical services research. *Schizophrenia Bulletin*, *18*, 561–626.

Bloom, J., Hu, T., Cuffel, B., Hausman, J., Wallace, N., & Scheffler, R. (Unpublished). Mental health costs under alternative capitation systems in Colorado.

Bond, G. R., McGrew, J. H., & Fekete, D. M. (1995). Assertive outreach for frequent users of psychiatric hospitals: A meta-analysis. *Journal of Mental Health Administration*, *22*, 4–16.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. (1994). *Time series analysis: Forecasting and control*, (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Brach, C. (1995). *Designing capitation projects for persons with severe mental illness: A policy guide for state and local officials*. Boston: Technical Assistance Collaborative.

Burns, B. J., & Santos, A. B. (1995). Assertive community treatment: An update of randomized trials. *Psychiatric Services*, *46*, 669–675.

Catalano, R. A., & McConnell, W. (1996). A time-series test of the quarantine theory of involuntary commitment. *Journal of Health and Social Behavior*, *37*, 381–387.

Catalano, R., Rook, K., & Dooley, D. (1986). Labor markets and help-seeking: A test of the employment security hypothesis. *Journal of Health and Social Behavior*, *27*, 277–287.

Catalano, R., & Serxner, S. (1987). Time series designs of potential interest to epidemiologists, *American Journal of Epidemiology*, *126*, 724–731.

Center for Mental Health Services. (1996). *Second year interim status report on the evaluation of the ACCESS demonstration program*. (Vol. I., Summary of Second year findings). Rockville, MD: Center for Mental Health Services.

Chamberlain, R., & Rapp, C. A. (1991). A decade of case management: A methodological review of outcome research. *Community Mental Health Journal*, *27*, 171–188.

Christianson, J. B., & Gray, D. Z. (1994). What CMHCs can learn from two states' efforts to capitate Medicaid benefits. *Hospital and Community Psychiatry*, *45*, 777–781.

Christianson, J. B., Gray, D. Z., Kihlstrom, L. C., & Speckman, Z. K. (1995a). Development of the Utah Prepaid Mental Health Plan. *Advances in Health Economics and Health Services Research*, *15*, 117–135.

Christianson, J. B., Manning, W., Lurie, N., Stoner, T. J., Gray, D. Z., Popkin, M., & Marriott, S. (1995b). Utah's prepaid mental health plan: The first year. *Health Affairs*, *14*, 160–172.

Christianson, J. B., & Osher, F. C. (1994). Health maintenance organizations, health care reform, and persons with serious mental illness. *Hospital and Community Psychiatry*, *45*, 898–905.

Cuffel, B. J., Wait, D., & Head, T. (1994). Shifting the responsibility for payment for state hospital services to community mental health agencies. *Hospital and Community Psychiatry*, *45*, 460–465.

Dickey, B., Normand, S. L., Norton, E. C., Azeni, H., Fisher, W., & Altaffer, F. (1996). Managing the care of schizophrenia: Lessons from a 4-year Massachusetts Medicaid study. *Archives of General Psychiatry*, *53*(10), 945–952.

Drummond, M., Brandt, A., Luce, B., Rovira, J. (1993). Standardizing methodologies for economic evaluation in health care. *International Journal of Technology Assessment in Health Care*, *9*, 26–36.

Drummond, M. F., Stoddard, G. L., & Torrance, G. W. (1997). *Methods for the economic evaluation of health care programmes* (2nd ed.). Oxford, UK: Oxford University Press.

Essock, S. M., Hargreaves, W. A., Covell, N. H., & Goethe, J. (1996a). Clozapine's effectiveness for patients in state hospitals: Results from a randomized trial. *Psychopharmacology Bulletin*, *32*, 683–697.

Essock, S. M., Hargreaves, W. A., Dohm, F-A., Goethe, J., Carver, L., & Hipshman, L. (1996b). Clozapine eligibility among state hospital patients. *Schizophrenia Bulletin*, *22*, 15–25.

Fenton, J., Catalano, R., & Hargreaves, W. A. (1996). Effect of Proposition 187 on mental health service use in California: A case study. *Health Affairs*, *15*, 182–190.

Freeman, D. H., Alegria, M., Vera, M., Munoz, C. A., Robles, R. R., Jimenez, A. L., Calderon, J. M., & Pena, M. (1992). A receiver operating characteristic (ROC) curve analysis of a model of mental health services use by Puerto Rican poor. *Medical Care*, *30*, 1142–1153.

Hargreaves, W. A., & Shumway, M. (1989). Effectiveness of mental health services for the severely mentally ill. In C. A. Taube, D. Mechanic, & A. A. Hohmann (Eds.), *The future of mental health services research*. Rockville, MD: National Institute of Mental Health.

Hargreaves, W. A., Shumway, M., Hu, T.-W., & Cuffel, B. (1997). *Cost-outcome methods for mental health*. San Diego, CA: Academic Press.

Henggeler, S. W. (1991). Multidimensional causal models of delinquent behavior. In R. Cohen & A. Siegel (Eds), *Context and development* (pp. 211–231). Hillsdale, NJ: Erlbaum.

Henggeler, S. W., Melton, G. B., & Smith, L. A. (1992). Family preservation using multisystem therapy: An effective alternative to incarcerating serious juvenile offenders. *Journal of Consulting and Clinical Psychology*, *60*, 953–961.

Knapp, M. R. J. (Ed.) (1995). *The economic evaluation of mental health care*. Aldershot, UK: Ashgate.

Lalley, T. L., Hohmann, A. A., Windle, C. D., Norquist, G. S., Keith, S. J., & Burke, J. D. (1992). Caring for people with severe mental disorders: A national plan to improve services. *Schizophrenia Bulletin*, *18*, 559–560.

Mechanic, D. (1987). Correcting misperceptions in mental health policy: Strategies for improved care of the seriously mentally ill. *Milbank Quarterly*, *65*, 203–230.

Mechanic, D., Bevilacqua, J., Goldman, H., Hargreaves, W., Howe, J., Knisley, M., Scherl, D. J., Stuart, G., Unhjem, M. B., & Lalley, T. L. (1992). Research resources. *Schizophrenia Bulletin*, *18*, 669–696.

Murray, D., McKinlay, S., Martin, D., Donner, A., Dwyer, J., Raudenbush, S., & Graubard, B. (1994). Design and analysis issues in community trials. *Evaluation Review*, *(August)*, 493–514.

National Institute of Mental Health. (1991). *Caring for people with severe mental disorders: A national plan of research to improve services*. DHHS Pub. No. (ADM)91–1762. Washington, DC: US Government Printing Office.

Olfson, M. (1990). Assertive community treatment: An evaluation of the experimental evidence. *Hospital and Community Psychiatry*, *41*, 634–641.

Randolph, F. L. (1992). NIMH funded research demonstration of (P/ACT) models. *Outlook*, *2*(2), 9–12.

Raudenbush, S. W. (1997) Statistical analysis and optimal design for cluster randomization trials. *Psychological Methods*, *2*, 173–185.

Rubin, A. (1992). Is case management effective for people with serious mental illness? A research review. *Health and Social Work*, *17*, 138–150.

Salive, M. E., Mayfield, J. A., & Weissman, N. W. (1990). Patient outcomes research teams and the Agency for Health Care Policy Research. *Health Services Research*, *25*, 697–708.

Santos, A. B., Henggeler, S. W., Burns, B. J., Arana, G. W., & Meisler, N. (1995). Research on field-based services: Models for reform in the delivery of mental health care to populations with complex clinical problems. *American Journal of Psychiatry*, *152*, 1111–1123.

Schooler, N. R., Keith, S. J., Severe, J. B., & Matthews, S. (1989). Acute treatment response and short term outcome in shizophrenia: First results of the NIMH Treatment Strategies in Schizophrenia study. *Psychopharmacology Bulletin*, *25*, 331–335.

Schooler, N. R., Keith, S. J., Severe, J. B., Matthews, S. M., Bellack, A. S., Glick, I. D., Hargreaves, W. A., Kane, J. M., Ninan, P. T., Frances, A., Jacobs, M., Lieberman, J. A., Mance. R., Simpson, G. M., & Woerner, M. (1997). Relapse and rehospitalization during maintenance treatment of schizophrenia: The effects of dose reduction and family treatment. *Archives of General Psychiatry*, *54*, 453–463.

Scott, J. E., & Dixon, L. B. (1995). Assertive community treatment and case management for schizophrenia. *Schizophrenia Bulletin*, *21*, 657–691.

Sloan, F. (Ed.) (1995). *Valuing health care: Costs, benefits, and effectiveness of pharmaceuticals and other medical technologies*. Cambridge, UK: Cambridge University Press.

Solomon, P. (1992). The efficacy of case management services for severely mentally disabled clients. *Community Mental Health Journal*, *28*, 163–180.

Stein, L., & Test, M. A. (1980). Alternative to mental hospital treatment: I. Conceptual model, treatment program, and clinical evaluation. *Archives of General Psychiatry*, *37*, 392–397.

Steinwachs, D. M., Cullum, H. M., Dorwart, R. A., Flynn, L., Frank, R., Friedman, M. B., Herz, M. I., Mulvey, E. P., Snowden, L., Test, M. A., Tremaine, L. S., & Windle, C. D. (1992). Service systems research. *Schizophrenia Bulletin*, *18*, 627–668.

Test, M. A. (1992). Training in community living. In R. P. Liberman (Ed.). *Handbook of Psychiatric Rehabilitation*. Boston: Allyn and Bacon.

Test, M. A., & Stein, L. (1980). Alternative to mental hospital treatment: III. Social cost. *Archives of General Psychiatry*, *37*, 409–412.

Weisbrod, B. A. (1983) A guide to benefit-cost analysis, as seen through a controlled experiment in treating the mentally ill. *Journal of Health Politics, Policy and Law*, *7*, 808–845.

Weisbrod, B., Test, M. A., & Stein, L. (1980). Alternative to mental hospital treatment: II. Economic benefit cost analysis. *Archives of General Psychiatry*, *37*, 400–405.

Yates, B. T. (1995). Cost-effectiveness analysis, cost-benefit analysis, and beyond: Evolving models for the scientist–manager–practitioner. *Clinical Psychology: Science and Practice*, *2*, 385–398.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

3.12
Descriptive and Inferential Statistics

ANDREW C. LEON
Cornell University Medical College, New York, NY, USA

3.12.1 INTRODUCTION 244
3.12.2 TABULAR AND GRAPHIC DISPLAYS 244
3.12.3 DESCRIPTIVE STATISTICS 245
  3.12.3.1 Measures of Central Tendency 245
  3.12.3.2 Measures of Variability 247
  3.12.3.3 Describing a Frequency Distribution 248
    3.12.3.3.1 Skewness 248
    3.12.3.3.2 Kurtosis 248
  3.12.3.4 Other Graphic Displays 248
  3.12.3.5 Statistical Software 249
3.12.4 INFERENTIAL STATISTICS 251
  3.12.4.1 One- vs. Two-tailed Tests 253
  3.12.4.2 Parametric vs. Nonparametric Tests 253
  3.12.4.3 Between-subject Designs: Dimensional and Ordinal Dependent Variables 254
    3.12.4.3.1 t-*test* 254
    3.12.4.3.2 Mann–Whitney test 255
    3.12.4.3.3 One-way analysis of variance 256
    3.12.4.3.4 ANOVA: Post hoc *tests* 257
    3.12.4.3.5 Kruskal–Wallis test 259
    3.12.4.3.6 Factorial analysis of variance 260
    3.12.4.3.7 Hotelling's T^2 264
    3.12.4.3.8 Multivariate analysis of variance 265
  3.12.4.4 Linear Relations 265
    3.12.4.4.1 Pearson correlation coefficient 265
    3.12.4.4.2 Test for a difference between two independent correlation coefficients 269
    3.12.4.4.3 Spearman rank correlation 270
    3.12.4.4.4 Simple linear regression analysis 272
    3.12.4.4.5 Multiple linear regression analysis 275
  3.12.4.5 Between-subject Designs: Categorical Dependent Variables 276
    3.12.4.5.1 Chi-square test 276
    3.12.4.5.2 Fisher's Exact test 277
    3.12.4.5.3 Logistic regression 278
    3.12.4.5.4 Survival analysis 279
  3.12.4.6 Within-subject Designs: Dimensional and Ordinal Dependent Variables 279
    3.12.4.6.1 Paired t-*test* 280
    3.12.4.6.2 Wilcoxon signed-rank test 280
    3.12.4.6.3 Repeated measures ANOVA 281
  3.12.4.7 Within-subject Designs: Categorical Dependent Variables 282
    3.12.4.7.1 McNemar test 282
3.12.5 SUMMARY 282

3.12.6 APPENDIX: SPSS COMMAND LINES 284
3.12.6.1 Descriptive Statistics 284
3.12.6.2 Inferential Statistics 284
3.12.6.2.1 Between-subject analyses 284
3.12.6.2.2 Within-group analyses 284
3.12.7 REFERENCES 284

3.12.1 INTRODUCTION

A fundamental motivation in science is to understand individual differences. It is *deviation* and *variability* that stimulate scientific curiosity. When do behavioral scientists study constants? One is not compelled to study differences in clinical psychology students who did and did not experience infancy. Instead a comparison of clinical psychology students who do and do not remember parental affect expression in infancy might be compelling. Similarly it is differences among scores on an achievement test that are exciting, not the proportion of psychology graduate students who took the Graduate Record Examination.

Psychological research is conducted to understand *individual differences*. In reporting research results, data are summarized with statistical analyses. In choosing among appropriate statistical procedures, several issues must be considered. First, are the results to be used strictly to describe one sample or will they also be used to draw inferences about the population from which the sample was drawn? *Descriptive statistics* are used for the former purpose, whereas *inferential statistics* are used for the latter. Here several of the more commonly used procedures are discussed. The choice among statistical procedures is also guided by the distribution, the sample size, and of course, the research question. This chapter gives an overview of those statistical procedures. It is not meant to provide a comprehensive explication of each technique, but instead provide a guide to the selection, implementation, and interpretation of appropriate procedures. References which provide comprehensive discussions of the procedures are cited in each section.

Descriptive statistics are used to summarize and describe data from a sample. Measures of *central tendency* and *variability* are fundamental. The former include quantities that represent the typical data from the sample, whereas the latter provide information on the *dispersion* of data within a sample. Data from a *hypothetical* clinical trial comparing the effects of two forms of psychotherapy for depression, interpersonal therapy (IPT) and cognitive-behavioral therapy (CBT), will be used repeatedly in this chapter. Table 1 presents only the data from the CBT group. Taking pedagogical license, the structure of the data will be changed to illustrate a variety of statistical procedures in this chapter. Throughout the chapter X_i will refer to an individual observation of the variable, X. For example, X_{17} represents the X value for subject number 17 (where i is each subject's arbitrarily assigned, unique identifying number, which ranges from 1 to N, in a study of N subjects).

3.12.2 TABULAR AND GRAPHIC DISPLAYS

There are a variety of tabular displays used to summarize data. Hamilton Rating Scale for Depression (HRSD) data will be used for illustration. Initially a *frequency distribution* is constructed by counting the number of occurrences, or frequency, of each value of the variable (X_i). In this case, X_i represents the HRSD rating for subject i, where i ranges from 1 to N. Five columns are included in the frequency distribution that is displayed in Table 2. The first column displays the values of X_i, in ascending order. The second column displays

Table 1 Hamilton Rating Scale for Depression (HRSD: X_i) for 20 patients (CBT group).

Subject number	X_i
1	12
2	12
3	13
4	14
5	15
6	17
7	18
8	19
9	21
10	21
11	21
12	22
13	24
14	25
15	26
16	27
17	27
18	28
19	29
20	29

Table 2 Frequency distribution of HRSD: CBT group.

X_i	f	$cum(n)$	%	cum (%)
12	2	2	10	10
13	1	3	5	15
14	1	4	5	20
15	1	5	5	25
17	1	6	5	30
18	1	7	5	35
19	1	8	5	40
21	3	11	15	55
22	1	12	5	60
24	1	13	5	65
25	1	14	5	70
26	1	15	5	75
27	2	17	10	85
28	1	18	5	90
29	2	20	10	100
Total	20		100	

the frequency for the value of X_i. The third column shows the number of subjects that have a value of X_i or lower, which is referred to as the *cumulative frequency*. In the fourth column the percentage of subjects with each value of X_i, or relative frequency, is displayed. In the fifth column, the cumulative percentage, or *cumulative relative frequency*, is shown, which is the percentage of subjects with a value less than or equal to X_i. Two column totals (sums, which are symbolized Σ) are presented. In the second column, the total of the frequencies is displayed, the value of which should be equal to the sample size (N). The total of the fourth column should equal 100%. Other cell entries worth noting include the final number in the column of cumulative frequencies which should equal N and the final entry in the cumulative relative frequencies column, which should be equal to 100%. Finally, the number of values of X_i provides an initial indication of its variability.

A graphic display of the contents of any of the last four columns can be informative. In these graphs, the x-axis of the graph represents the X_i values and the frequency data are represented on the y-axis. These graphs provide the reader with a sense of the distribution of the data at a glance. For instance, Figure 1 is a bar chart, or *histogram*, displaying the frequencies of each X value. (If there were no variability, only one bar would be displayed.) Figure 2 displays the *relative frequencies*. Figure 3 displays the *cumulative frequency* distribution. Finally, the *cumulative relative frequencies* of the HRSD ratings are displayed in Figure 4. Although Figures 1 and 3 display sample sizes, the proportions presented in Figures 2 and 4 can be readily compared across samples of diverse sizes.

3.12.3 DESCRIPTIVE STATISTICS

Descriptive statistics are summary measures that describe characteristics of data. The two general areas of descriptive statistics that will be discussed here are measures of *central tendency* and measures of variability. The data that were originally shown in Table 1 are now displayed in Table 3, with additional information that will be used for calculations throughout this section.

3.12.3.1 Measures of Central Tendency

The arithmetic mean ($\bar{X}$) of a sample, commonly referred to as the average, is equal to the sum of all observations in the sample, $\sum_{1}^{n} X_i$, divided by the number of observations (N) in the sample.

$$\bar{X} = \frac{\left(\sum_{i=1}^{N} X_i\right)}{N} = \frac{420}{20} = 21$$

The *median*, *Md*, is the middle observation in a sample of ordered data (i.e., ranked from lowest to highest value of X_i). In the case of ordered data, the subscript represents the ascending rank, ranging from 1 to n. With an

Table 3 HRSD: CBT group (N = 20).

X_i	X_i^2	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
12	144	−9	81
12	144	−9	81
13	169	−8	64
14	196	−7	49
15	225	−6	36
17	289	−4	16
18	324	−3	9
19	361	−2	4
21	441	0	0
21	441	0	0
21	441	0	0
22	484	1	1
24	576	3	9
25	625	4	16
26	676	5	25
27	729	6	36
27	729	6	36
28	784	7	49
29	841	8	64
29	841	8	64
Total 420	9460	0	640

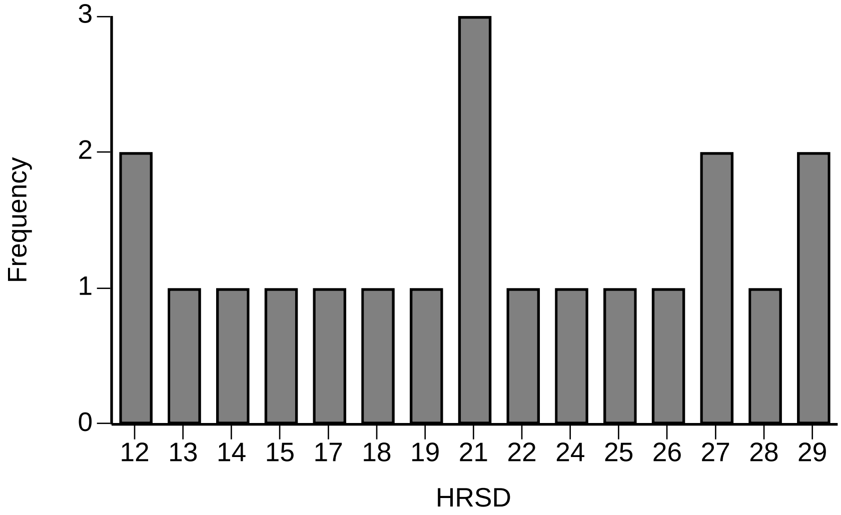

Figure 1 Plot of frequency distribution of HRSD: CBT group.

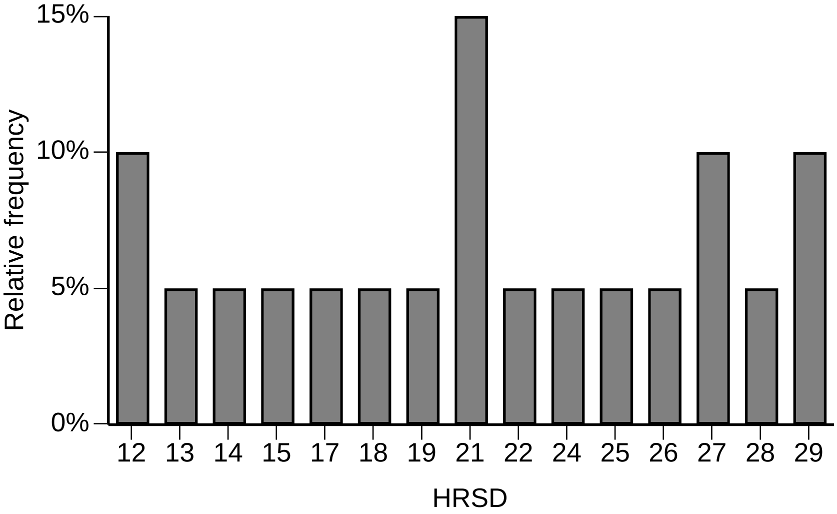

Figure 2 Plot of relative frequencies of HRSD: CBT group.

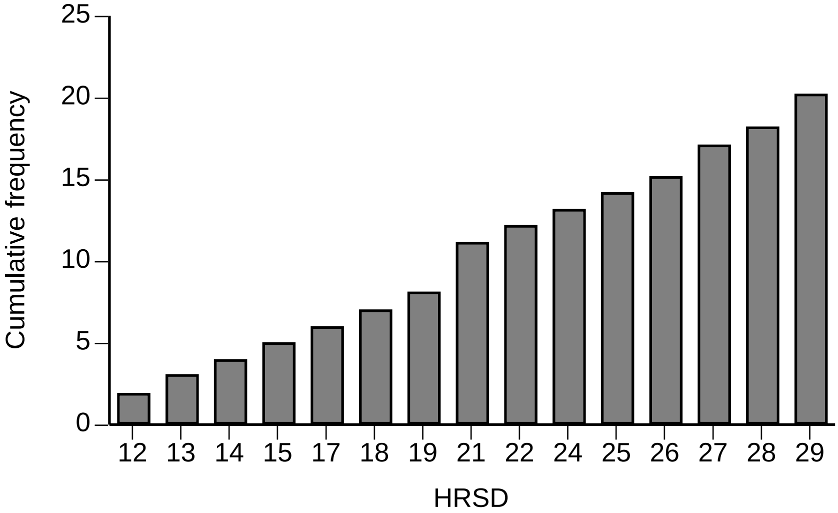

Figure 3 Plot of cumulative frequencies of HRSD: CBT group.

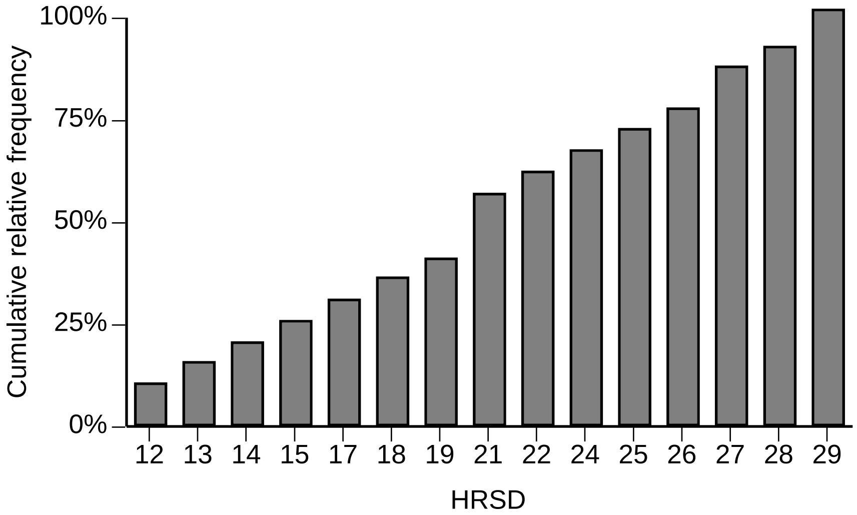

Figure 4 Plot of cumulative relative frequencies of HRSD: CBT group.

even number of observations, the median is the mean of the two middle observations. For instance, there are 20 observations in the sample presented in Table 1. Thus the median of the sample is the mean of the middle two values: 10 and 11:

$$Md = X_{(n+1)/2} = 21 \text{ (mean of 2 middle scores)}$$

The median represents the fiftieth percentile, the point below which half of the data fall. For example, in Table 1, a value of 21 is the median. When the data include a very small number of extreme values, the data set is referred to as skewed (discussed in greater detail below). With *skewed* data, the median is a more representative measure of central tendency than the mean because the mean is influenced by extreme values. For instance, the median is a more appropriate measure of central tendency of the US household income than the mean.

The mode, Mo, is the most frequently occurring value. The modal value can be determined by examining the frequency distribution which is presented in Table 1 and selecting the X_i with the largest value of f_i:

$$\text{Mo} = 21(f = 3)$$

3.12.3.2 Measures of Variability

This chapter began by stating that it is an understanding of individual differences that motivates a great deal of behavioral research. A wide range of statistical procedures are used to analyze differences. Each of the procedures that is discussed in this chapter has to do with differences: individual or group; temporal or conditional. Here the quantities used to assess differences, or *variability*, are discussed. The crudest measure of variability is the *range*, which is the difference between the maximum and minimum value in the data. For instance, in Table 1:

$$\begin{aligned}\text{Range} &= \text{Maximum} - \text{Minimum}\\ &= X_N - X_1 = 29 - 12 = 17\end{aligned}$$

There are several quantities that provide a more comprehensive description of the variability in data. Most are derived from the *Sum of Squares* (*SS*), which is the sum of squared deviations from the mean. The squared deviations are presented in the right-hand column of Table 3. Note that one of the properties of the mean is that the sum of the deviations from the mean is equal to zero:

$$\sum_{i=1}^{N}(X_i - \bar{X}) = 0$$

Notice that the sum of the deviations from the mean in Table 3 is equal to zero. However, the sum of squared deviations from the mean is a useful quantity:

$$SS = \sum_{i=1}^{N}(X_i - \bar{X}) = 640$$

The SS can be used for calculating the *variance* (s^2), which is approximately the mean of the squared deviations from the mean. The variance, a quantity that is commonly used to convey information regarding variability, is calculated as $\frac{SS}{n-1}$, and not $\frac{SS}{n}$, because it provides an unbiased estimator of the population variance. Further discussion of the reasons for this will not be discussed here. Note that, as the sample size increases the difference between $\frac{SS}{n-1}$ and $\frac{SS}{n}$ becomes trivial. The variance is calculated as follows:

$$s^2 = \frac{\sum_{i=1}^{N}(X_i - \bar{X})^2}{N-1} = \frac{640}{20-1} = 33.684$$

Or, if the SS is available:

$$s^2 = \frac{SS}{N-1} = \frac{640}{20-1} = 33.684$$

The *standard deviation* (s), a fundamental measure of variability, is equal to the square root of the variance. It is approximately equal to the average deviation from the mean. It is approximately equal because the deviations have been squared, summed, and divided by $N-1$, not by N, before the square root is computed. Note that by taking the square root, the original unit of measurement is restored.

$$s = \sqrt{s^2} = \sqrt{33.684} = 5.804$$

3.12.3.3 Describing a Frequency Distribution

Modality has to do with the number of modes contained within a frequency distribution. The HRSD data in Figure 1 has one mode and is thus referred to as unimodal. The most commonly occurring value in this sample is 21. There are three observations equal to 21 and there is no other value that is represented as often as 21. Because the y-axis represents the frequency, the mode is the highest point in the plot of the frequency distribution in Figure 1. In contrast to these data, some frequency distributions are bimodal, having two peaks, or multimodal, having more than two peaks.

3.12.3.3.1 Skewness

If one folds the distribution, vertically in the middle and one half can be superimposed on the other, the distribution is *symmetric*. The extent to which a frequency distribution diverges from symmetry is described as *skewness*. With help from Bill Gates and a few of his peers, incomes in the United States are positively skewed. The skewness, which is one aspect of the departure from symmetry, is described as either skewed to the right, positively skewed as shown in Figure 5, or skewed to the left, negatively skewed, as shown in Figure 6. For example, with a different set of HRSD data, a positively skewed distribution is one in which there are a very few severely depressed scores, whereas a distribution in which there are a very few number of euphoric scores might be negatively skewed.

3.12.3.3.2 Kurtosis

Kurtosis has to do with the extent to which a frequency distribution is peaked or flat. A normal bell-shaped distribution is referred to as a *mesokurtic* shape distribution. An example of this, a nicely rounded distribution, is shown in Figure 7. Many human traits are normally distributed including height and intelligence. A *platykurtic*, like a platypus, is a somewhat flat creature. For instance, as shown in Figure 8, this would describe data in which there are several modes, which cluster around, and include, the median. A *leptokurtic* distribution is more peaked. In Figure 9 hypothetical data show that such data would be characterized by a very high, unimodal peak, at the median.

3.12.3.4 Other Graphic Displays

Two other formats for graphic display are the *Stem and Leaf* and the *Box and Whisker* plots. These graphic displays, which are described in detail by Tukey (1977), each provide the researcher with a great deal of information about the distribution of the data. The importance of examining such plots cannot be overemphasized. One outlier (i.e., extreme score) in a data set can markedly alter the mean and exaggerate trivial group differences. The existence of outliers is most readily apparent in a good graph.

A *Stem and Leaf plot* is an efficient method of displaying a frequency distribution. In Figure 10, the vertical line, called the stem, displays the possible values, ranging from the minimum (at the top) to maximum value (at the bottom). Each leaf represents an individual with a particular value. A leaf could be an "x," an asterisk, or more commonly the value itself. A data value (on the stem) with several leaves is a frequently occurring value. Of course, the mode has the most leaves.

Box and Whisker plots are based on five points as shown in Figure 11. The first three of those have been discussed: the median, the minimum, and maximum. The final two points of the box and whisker plot are the values that fall midway between each extreme score and the median. These are the first and third *quartiles*. A long thin box is drawn with the latter two points serving as the *hinges*. A horizontal bar is drawn through the median. *Whiskers* are drawn at the two extreme values. The figure illustrates how easily outliers are detected in a Box and Whisker plot. In some situations, the plot is more informative when an identifier is used to label values such as an outlier.

When an outlier is identified, the researcher must determine if the value is real or if it is a keypunch error. The raw data forms will help

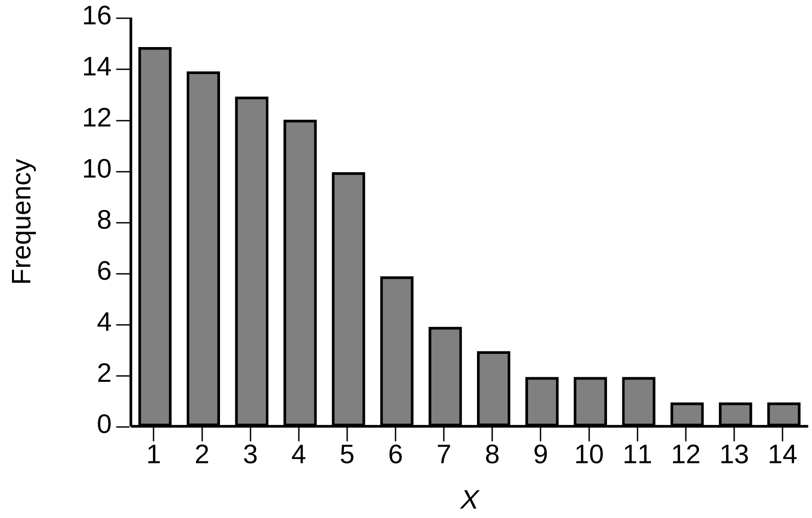

Figure 5 Example of positively skewed data.

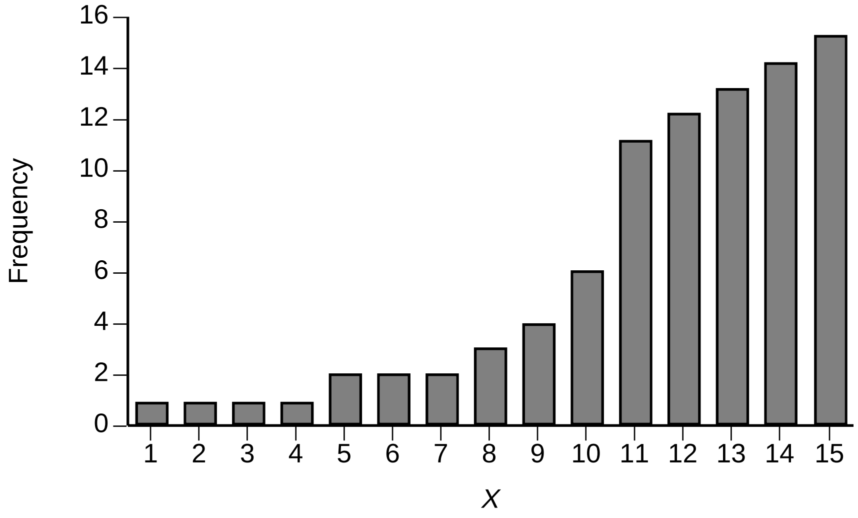

Figure 6 Example of negatively skewed data.

one correct keypunch errors. However, extreme scores are not necessarily incorrect values. In a study of depression, an HRSD score of 58 is possible, but extremely rare. In analyzing data in which valid extreme values have been identified, one has several strategies to choose from. First, a nonparametric procedure could be used in which the ranks of the raw data values are used. In such a case the minimum and maximum are equal to 1 and N, respectively. Second, a transformation of the data values using either the natural logarithm, where $y' = \ln(y)$, or square root of the original scale $y'' = \sqrt{y}$. Third, an extreme score might be *winsorized*, a process in which an outlier is simply recoded to the second highest value (see Tukey, 1977). Alternative strategies for outliers are discussed in detail by Wilcox (1996).

3.12.3.5 Statistical Software

Many of the statistical procedures that are described in this chapter involve tedious calculations. This leaves a great deal of room for computational errors. For all practical purposes, most statistical analyses in psychological research are conducted using a computer. There is a wide variety of statistical software that is commercially available for personal computers. Software that is commonly used in the social sciences includes SPSS, SAS, BMDP, and SYSTAT. Comprehensive reviews of statistical software are available (Lurie, 1995; Morgan, 1998). The choice among statistical software has to do with the statistical procedures that are included, the data management capabilities, and the ease of use. Nevertheless, a

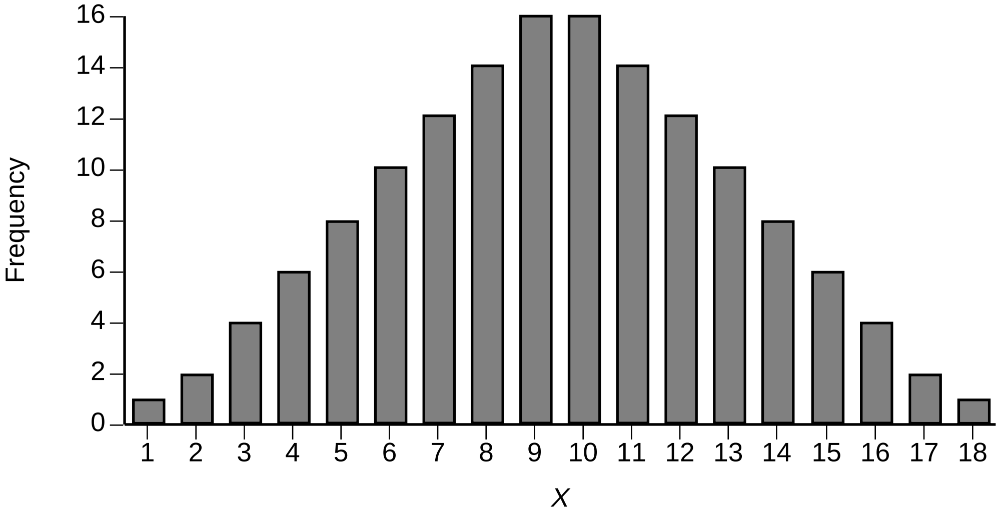

Figure 7 Example of normally distributed data.

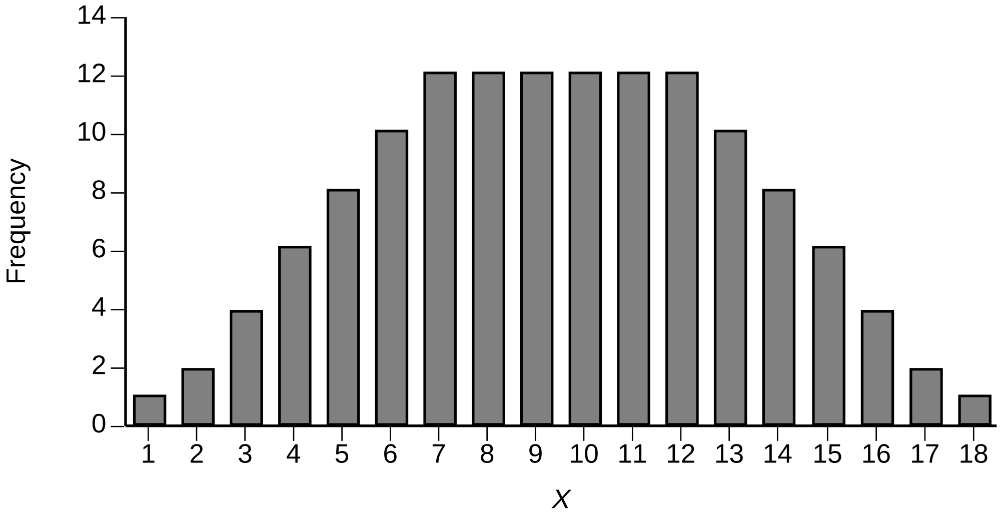

Figure 8 Example of data with a platykurtic distribution.

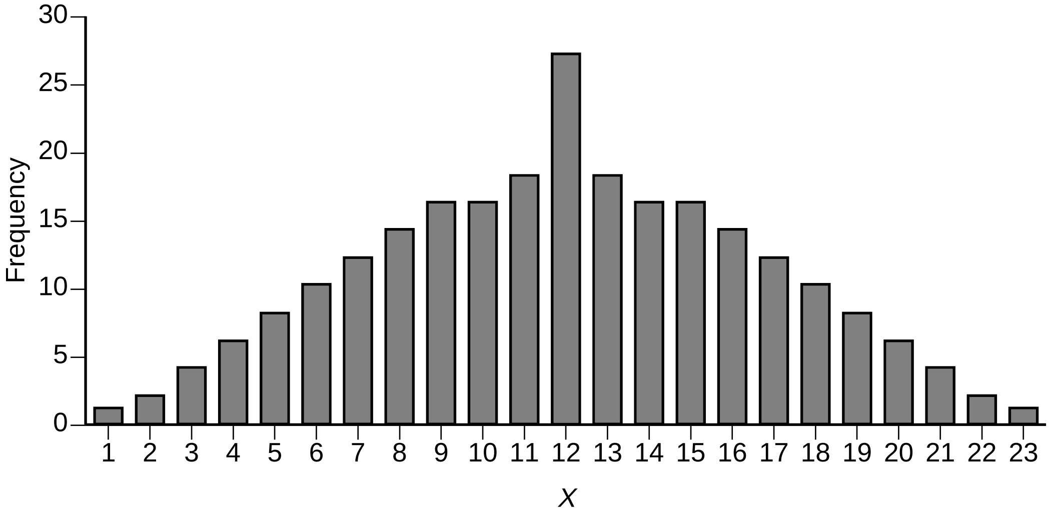

Figure 9 Example of data with a leptokurtic distribution.

```
   Frequency        Stem and Leaf

      4               1*2234
      4               1. 5789
      5               2*11124
      7               2. 5677899

   Stem width:          10.00
    Each leaf:        1 case(s)
```

Figure 10 HRSD: Stem and Leaf Plot.

statistical procedure should only be used in appropriate situations. A thorough understanding of the statistical technique, the assumptions, and a familiarity with the estimation procedure is necessary. Furthermore, the user must have the background necessary to interpret the results that are generated by the software. One of the goals of this chapter is to provide the reader with such a background for a variety of descriptive and inferential procedures. In addition, the appendix of this chapter includes SPSS command lines for each of the analyses that is discussed (SPSS, 1997). SPSS is used for illustration here because it is accessible with minimal programming skills, it covers a wide array of statistical procedures, and it has excellent data management capacity.

3.12.4 INFERENTIAL STATISTICS

Once again, an *inferential* statistical test is used to make an inference about a population based on data from a sample. In choosing among inferential procedures, one must consider the nature of both the independent and dependent variable, whether a between-subject design or within-subject design is employed, and the sample size. The procedures that are presented below are organized to reflect such considerations. Initially statistical procedures for between-subject designs are discussed. This is followed by a discussion of procedures for within-subject designs. In each of those sections both parametric and nonparametric procedures are considered. The distinction between parametric and nonparametric procedures is discussed in detail in Section 3.12.4.2.

First a digression into general features regarding inferential tests. A statistical test is used to draw an inference about a population(s) based on sample data. The test is used to examine the research question at hand. For instance, "Which of two forms of psychotherapy is more effective for the treatment of depression?" For purposes of hypothesis testing, the question is reframed into a null hypothesis (i.e., a hypothesis of *no difference*). For example, "There is no difference in the effects of CBT and IPT for the treatment of depression." In conducting such tests, the researcher might make a correct decision, but errors are certainly possible. Consider the correspondence between the *truth* and the research results (see Table 4). Assume that with omniscient (noninferential) ability, someone is able to describe the *truth* (e.g., population characteristics) without conducting a statistical test. Thus the omniscience. (Every department has at least one member with self-rated ability of that level.) For example, consider a clinical trial in which the post-treatment psychopathology of two groups (CBT and IPT) is being compared. Our colleague, Dr. Omniscient, could tell us whether or not the severity of psychopathology in the two groups differs. If there is no difference in the population, the correct decision would be to conclude that there is no difference (i.e., fail to reject the null hypothesis). In contrast, if there is no population difference, yet we find a difference, we would be in error. Specifically, rejecting a true null hypothesis is referred to as *Type I Error*. The probability of Type I Error is referred to as the alpha level (α).

On the other hand, if Dr. Omniscient tells us that there is, in fact, a treatment difference (i.e., population treatment differences), we would be correct if we detected the difference (rejected a false null hypothesis). The probability of such a rejection is referred to as *statistical power*. In contrast, we would be in error if we fail to detect the treatment difference (i.e., fail to reject a false null hypothesis). This is referred to as *Type II Error* and the $p(\text{Type II error}) = \beta$. In the absence of Dr. Omniscient's assistance, replication of results in subsequent studies provides convincing evidence that our inference was likely to be accurate.

One further general digression regarding inferential statistics. In general a test statistic

Table 4 Cross-classigication of the "truth" and research results.

Results of research	*Placebo = Active* (H_0 true)	*Placebo ≠ Active* (H_0 false)
No difference (CBT = IPT)	Correct decision	Type II error
Difference (CBT = IPT)	Type I error	Correct decision "Power"

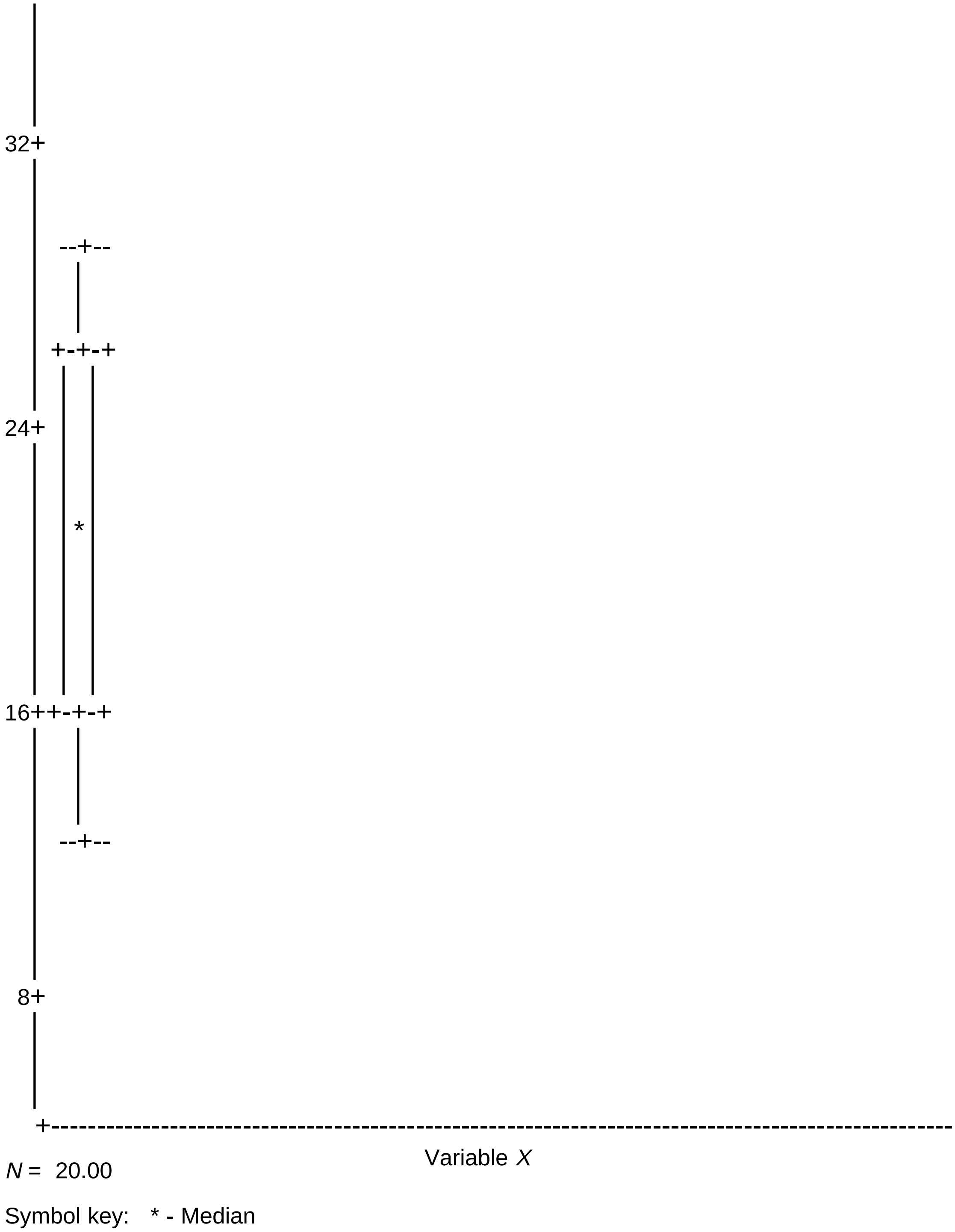

Figure 11 HRSD: Box and Whisker Plot.

is a ratio of a statistic (e.g., difference in means) to its *standard error*. Loosely speaking, the standard error is an estimate of the variability of the statistic in the population. Stated differently, and hypothetically, if the experiment were conducted repeatedly, say 10 000 times (with the same sample size each time), and the quantity of interest (e.g., difference between two means) were calculated in each of those experiments, the standard error would represent the standard deviation of that distribution (the sampling distribution) of that quantity (e.g., the standard deviation of the differences between two means).

As stated earlier, one component of each inferential statistical test is the null hypothesis,

Table 5 Hamilton Depression Ratings (X_i) for 20 patients (IPT group).

	X_i	X_i^2	$X_i-\bar{X}$	$(X_i-\bar{X})^2$
	6	36	-9	82
	7	49	-8	65
	8	64	-7	50
	9	81	-6	37
	12	144	-3	9
	12	144	-3	9
	12	144	-3	9
	13	169	-2	4
	17	289	2	4
	17	289	2	4
	18	324	3	9
	18	324	3	9
	19	361	4	16
	19	361	4	16
	19	361	4	16
	19	361	4	16
	19	361	4	16
	19	361	4	16
	19	361	4	16
	19	361	4	16
Total	301	4945	0	414.95

$$\bar{X} = \frac{\sum_{1}^{n} X_i}{n} = \frac{301}{20} = 15.05$$

$$\text{SS} = \sum (X_i - \bar{X})^2 = 414.95$$

$$s^2 = \frac{\text{SS}}{n-1} = \frac{414.95}{20-1} = 21.839$$

$$s = \sqrt{s^2} = \sqrt{21.839} = 4.673$$

H_0. In fact that is one of the eight components of the hypothesis testing procedure which are listed below:

(i) State research question.

(ii) State two hypotheses which are mutually exclusive and exhaustive:

 (a) null hypothesis (e.g., $H_0 : \mu_1 = \mu_2$)

 (b) alternative hypothesis (e.g., $H_A : \mu_1 \neq \mu_2$).

(iii) Choose an appropriate statistical test

(iv) Specify the critical value for the statistical test and the alpha level (both of which are discussed below).

(v) Collect data.

(vi) Perform calculations.

(vii) Make decision regarding H_0.

(viii) Conclusion. The results are interpreted as an inference about the population (which the sample represents) based on the sample data. The *generalizability* of the findings should be discussed. Inferences about the results cannot be generalized beyond the scope of the data that was examined. For instance, results from a clinical trial for the treatment of patients with depression cannot be used to draw inferences about the treatment of patients with panic disorder.

3.12.4.1 One- vs. Two-tailed Tests

In some settings, the researcher might have preconceived expectations regarding the outcome of the experiment. For instance, it might be expected that a novel treatment is superior to the standard. The researcher can then choose to specify the null hypothesis as that of *no treatment difference* or *superiority of the novel treatment*. A two-tailed null hypothesis is designed to detect any difference, positive or negative, whereas a one-tailed null hypothesis is designed to detect a difference in the prespecified direction (e.g., efficacy) but it would fail to detect the opposite (e.g., detrimental) effects. It is often as important, and ethically compelling, to detect detrimental treatments as it is to detect efficacious treatments. A two-tailed test is used for a nondirectional null hypothesis and a one-tailed test is used for a directional null hypothesis. Although there is debate regarding the exclusive use of two-tailed tests, there is a consensus that the rationale for a one-tailed test must be stated *a priori*. That is, if a one-tailed test is to be used, it must be so stated prior to study implementation and data collection. There has been an extensive discussion in the literature of the costs and benefits of one and two-tailed tests (e.g., Fleiss, 1981, pp. 27–29; Hopkins, Glass, & Hopkins, 1987, pp. 164–165, 205–206). All examples in this chapter employ two-tailed, nondirectional tests.

3.12.4.2 Parametric vs. Nonparametric Tests

Technically speaking, one must not use parametric tests such as a *t*-test and an ANOVA unless assumptions such as normality are fulfilled. By normality, it is assumed that the sample come from a population with normally distributed data. Briefly, in a normal distribution, which is bell-shaped, the mean, median, and mode are equal. The distribution is symmetrical around these measures of central tendency. The shape of a normal distribution is determined by the population means and variance. The nonparametric alternatives to the *t*-test and the ANOVA are the Mann–Whitney test and Kruskal–Wallis test. Neither of these makes the normality assumptions. Many of the nonparametric procedures require a simple rank transformation of the data (Conover, 1980; Sprent, 1989). This involves pooling the data from all subjects, regardless of treatment group, and ranking in ascending order based on the value of the dependent variable. The subject with the lowest value is assigned the rank of 1. The subject with the next lowest is assigned the value of 2, and so on, until the subject with the highest value is assigned a rank of N. If all assumptions have been reasonably fulfilled and all other aspects of the analysis are equal, parametric tests generally have more statistical power than nonparametric tests. However, if the assumptions are not met, parametric tests could provide misleading results about the population which the sample represents.

3.12.4.3 Between-subject Designs: Dimensional and Ordinal Dependent Variables

3.12.4.3.1 t*-test*

The t-test is used to compare the means of two groups. For instance, in a clinical trial, the severity of depression of two treatment groups (IPT and CBT) can be compared with a t-test. The HRSD is the dependent variable. The *null hypothesis* is that the population means do not differ: $H_0 : \mu_1 = \mu_2$ and the alternative hypothesis is that population means are not the same: $H_A : \mu_1 \neq \mu_2$.

The test is simply a ratio of the *between-group* differences relative to the *within-group* differences. It is used to ask the question, "Is the mean difference *between groups* large relative to the fluctuation that is observed *within* each group?" To illustrate this, the data from the CBT group (presented in Table 3) are compared with data from the other group of subjects in the hypothetical clinical trial that was described earlier, those assigned to IPT (see Table 5).

The algorithm for the t-test is a ratio of *between-group* variability to *within-group* variability:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}}$$

where the numerator represents the between-group variability (i.e., difference between the group means) and the denominator is derived from the pooled variance. The pooled variance is estimated as:

$$s_p^2 = \frac{(SS_1 + SS_2)}{(n_1 - 1) + (n_2 - 1)}$$

and

$$s_{(\bar{X}_1 - \bar{X}_2)} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$$

Applied to the CBT vs. IPT data set: the pooled variance estimate,

$$s_p^2 = \frac{(SS_1 + SS_2)}{(n_1 - 1) + (n_2 - 1)}$$

$$= \frac{640.0 + 414.95}{(20 - 1) + (20 - 1)} = 27.76$$

and

$$s_{(\bar{X}_1 - \bar{X}_2)} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} = \sqrt{\frac{27.76}{20} + \frac{27.76}{20}} = 1.67$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{(\bar{X}_1 - \bar{X}_2)}} = \frac{21.00 - 15.05}{1.67} = 3.56$$

This t-statistic is compared with a critical value from tabled values, which can be found in the appendix of any statistics book. To identify that critical value, one must determine the *degrees of freedom* (*df*). Although a discussion of the meaning of degrees of freedom is beyond the scope of this chapter, a value for *df* is often needed to find the critical value of a test statistic. In the case of a two-group t-test, *df* is equal to the total number of subjects minus the number of groups:

$$df = n_1 + n_2 = 20 + 20 - 2 = 38$$

Using this value for *df*, and a two-tailed alpha level of 0.05, the *critical* t-statistic is 2.024. The *observed* t of 3.56 exceeds this critical value. Thus the null hypothesis is rejected and one can conclude that the population means are not identical. IPT treatment is superior to CBT. Without the t-test it was apparent that the IPT mean indicated lower symptom severity as measured by the HRSD than the CBT mean. To get a handle on the magnitude of that difference, it is compared with the within-group standard deviations. The IPT mean was over 6 HRSD points lower, which exceeds one within-group standard deviation (i.e., $s_{\text{IPT}} = 4.67$ and $s_{\text{CBT}} = 5.80$). Using the t-test, one can conclude that the difference in group means is greater than would be expected by chance.

The *assumptions* of the t-test include:

(i) Independence among observations. For example, one subject cannot re-enroll in the clinical trial to try another treatment if the initial treatment does not result in symptom remission. If that were permitted, some subjects would contribute multiple observations to the sample and those observations would not be independent.

(ii) The data are normally distributed in the population. Technically, this assumption has to do with the distribution of the *difference in means*, $\bar{X}_1 - \bar{X}_2$ or the population sampling distribution of $\mu_1 - \mu_2$. This assumption is generally met if the sample distributions of the raw data (i.e., the actual measurements), as opposed to the mean differences, are normal, or if the sample sizes are *reasonably* large.

(iii) Homogeneity of variance. The population variances are presumed to be equal. There are several tests of homogeneity of variance including Levene's (1960) test of equality of variances, Bartlett's test for homogeneity of variances, or Hartley's F-Max test (1950). These tests will not be discussed here. In calculating the t-test statistic, one generally uses a pooled variance estimate. However, when the assumption cannot be met, separate variance estimates can be incorporated into the t-test, or another procedure altogether, such as the Mann–Whitney test (discussed below), can be used. In the case of *heteroscedasticity* (unequal group variances), a different algorithm for the t-test is used which incorporates separate variance estimates in the denominator. In addition, an adjustment in degrees of freedom is necessary. This will not be discussed here, but is presented in a variety of other references (Armitage & Berry, 1987; Fleiss, 1986; Zar, 1996).

In summary, the hypothesis testing steps for the t-test are as follows:

(i) State research question. Do CBT and IPT have similar effects for the treatment of patients with depression?

(ii) State two mutually exclusive and exhaustive hypotheses:

(a) $H_0: \mu_1 = \mu_2$

(b) $H_A: \mu_1 \neq \mu_2$

(iii) If the assumptions appear to be fulfilled, choose the t-test to evaluate the hypotheses.

(iv) The critical value for the t-test with $\alpha = 0.05$ and $df = 38$ is 2.024. If $|t_{\text{observed}}| > t_{\text{crititcal}}$, reject H_0.

(v) Data are displayed in Table 3 for the CBT group and Table 5 for the IPT group.

(vi) Calculations were described above.

(vii) Decision regarding H_0 Because $|t_{\text{observed}}| > t_{\text{crititcal}}$ (i.e., $3.56 > 2.024$), reject H_0.

(viii) Conclusion. IPT is a more effective treatment of depressive symptoms in these patients with major depressive disorder.

3.12.4.3.2 Mann–Whitney test

If the assumptions of the t-test cannot be met, yet the observations are independent, the Mann–Whitney test, (also called the Wilcoxon test), may be an appropriate method for two group comparisons, if the data are at least ordinal in nature. An ordinal measurement need not have equal intervals between units, but must be ordered numbers (e.g., 1st, 2nd, 3rd, . . .). The Mann–Whitney test can be used to compare two groups without making any distributional assumptions. It is used when there is reason to believe that the data are not normally distributed, when the sample sizes are small, or when the variances are heterogeneous. It can be used for dimensional data (e.g., HRSD), as well as ordered categorical data [e.g., illness severity categories rated on a Likert scale from 0 (none) to 10 (severe)].

Analyses are not conducted on the raw data, but instead on rank transformed data, as described above. For instance, using the data from the clinical trial comparing IPT and CBT, the ranked data are displayed in Table 6. Instead of analyzing the actual Hamilton scores (the dependent variable), each of the scores is ranked, with the lowest rank assigned a rank of 1 and the highest assigned a rank of N. For purposes of ranking the data are pooled across groups. In the case of ties, the midpoint is used. Subsequent calculations are conducted with the ranks.

The rationale behind the Mann–Whitney test is that if the treatments are equally effective, one would expect about half of each group to be below the median rank for the pooled sample. Another way to think of it is to consider all possible pairs of data across groups: subject 1 from group 1 with subject 1 from group 2; then subject 1 with subject 2 from group 2; . . . subject 20 from group 1 with subject 20 from group 2.

Table 6 Hamilton Rating Scale of Depression for patients in the IPT group ($N = 20$) and patients in the CBT group ($N = 20$): raw ratings (X_i) and rank-transformed ratings ($R(X_i)$).

	IPT		*CBT*	
	X_{i1}	$R(X_{i1})$	X_{i2}	$R(X_{i2})$
	12	7	13	10.5
	12	7	12	7
	17	15	14	12
	18	18	15	13
	19	24	17	15
	19	24	18	18
	19	24	21	30
	19	24	21	30
	19	24	21	30
	12	7	22	32
	13	10.5	12	7
	17	15	25	34
	18	18	24	33
	19	24	19	24
	19	24	28	38
	19	24	27	36.5
	6	1	26	35
	7	2	29	39.5
	8	3	29	39.5
	9	4	27	36.5
Total	301	299.5	420	520.5

Examine whether or not in more than half of the ranked pairs the group 1 member is lower than its counterpart from group 2. That is, we would count the number of times that a subject in group 1 has a higher rank than a subject in group 2. Computationally, the tedious process of all $n_1 \cdot n_2$ comparisons need not be made.

The computations for the Mann–Whitney U test that are now described are based on the formula for the case where both n_is are less than or equal to 20. The total number of subjects across groups is $N = n_1 + n_2$ and $U = n_1 \cdot n_2 + \frac{n_1(n_1+1)}{2} - R_1$, where R_1 is the sum of the ranks for group 1. Note that, in general, the sum of N ranks $= \frac{N(N+1)}{2}$. More specifically, in this case, $\frac{40(40+1)}{2} = 820$, which is equal to the sum of the ranks: $299.5 + 520.5 = 820$.

The observed U statistic (U_{obs}) is compared with a critical range which is tabled in most introductory statistics texts (see for example, Fleiss, 1986; Zar, 1996). The critical range is a function of the sample size of each group.

$$U_{obs} = n_1 \cdot n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$= 20 \cdot 20 + \frac{20 \cdot 21}{2} - 299.5 = 310.5$$

For the Mann–Whitney test, the hypothesis testing steps are as follows:

(i) State research question. Do CBT and IPT have similar effects for the treatment of patients with depression?

(ii) State two mutually exclusive and exhaustive hypotheses with respect to the population medians (M_k).

$H_0: M_1 = M_2$

$H_A: M_1 \neq M_2$

(iii) Choose the Mann–Whitney test to evaluate the hypotheses.

(iv) The critical values for the Mann–Whitney test, with $\alpha = 0.05$ and $n_1 = n_2 = 20$, are 127 and 273. Thus, if $U_{\text{obs}} \leqslant 127$ or $U_{\text{obs}} \geqslant 273$, reject H_0.

(v) The ranked data are displayed in Table 6 for the IPT and CBT groups.

(vi) Calculations were described above.

(vii) Make decision regarding H_0.

$U_{\text{obs}} \geqslant U_{\text{critical}}$ (i.e., $310.5 \geqslant 273$), thus reject H_0.

(viii) Conclusion: IPT is a more effective treatment of depressive symptoms in patients with major depressive disorder.

The assumptions of the Mann–Whitney test include:

(i) The two samples have been independently and randomly sampled from their respective populations.

(ii) The groups are independent.

(iii) The scale of measurement is at least ordinal.

For larger samples (i.e., when n_1 or n_2 is greater than 20), a Normal Approximation can be used for the Mann–Whitney test (see Conover, 1980).

3.12.4.3.3 One-way analysis of variance

When more than two groups are compared on a dimensional dependent variable, a fixed effects analysis of variance (ANOVA) is used. As with the *t*-test, the test statistic is a ratio of between-group differences to within-group differences. In this case the sum of squares is the quantity representing variability. The sums of squares was introduced earlier as the sum of squared deviations from the mean. For example, consider the clinical trial that has been described above, with the addition of a placebo group. Say that patients are randomly assigned to one of *three* treatments: CBT, IPT, or a psychotherapy placebo. The psychotherapy placebo patients see their clinicians for medical management (MM)—a brief nondirective/reflective talk therapy. The data from the MM group are presented in Table 7. The efficacy of the three treatments for depression in this study can be compared using an ANOVA, once again with the HRSD as the dependent variable.

Initially, three different types of sums of squared deviations are computed. The total sum of squared deviations (SS_{Total}), sum of squared deviations accounted for by treatment group ($\text{SS}_{\text{Treatment}}$), and the residual sum of squared deviations, also referred to as "sum of squares error" ($\text{SS}_{\text{Residual}}$), are calculated as described below. There are several aspects about these three quantities that are worth noting at this point. First, the SS_{Total} is the sum of squared deviations from the grand mean (i.e., for all subjects pooled, without regard to treatment group). Second, $\text{SS}_{\text{Treatment}}$ is a measure of the squared deviations of the treatment group means from the grand mean. It can be thought of as the sum of squares that is *explained* (by knowing what treatment the subjects received). Third, the sum of squared differences between each observation and its sample mean is the $\text{SS}_{\text{Residual}}$. It represents the sum of squares that is not explained by treatment, and in this context, is the difference between SS_{Total} and $\text{SS}_{\text{Treatment}}$. These quantities are used in the *F*-test of the fixed effects ANOVA that is presented in Table 8 (the calculation of each quantity is also described in Table 8).

The assumptions of fixed effects ANOVA are:

(i) The samples are independent.

(ii) The populations from which the sample are drawn are normally distributed.

(iii) Each of the populations has the same variance.

For a fixed effects ANOVA, the hypothesis testing steps are as follows:

(i) State research question. Do CBT, IPT and MM have similar effects for the treatment of patients with depression?

(ii) State two mutually exclusive and exhaustive hypotheses:

$H_0 : \mu_1 = \mu_2 = \mu_3$

$H_{A_1} : \mu_1 \neq \mu_2$ or $H_{A_2} : \mu_1 \neq \mu_3$ or $H_{A_3} : \mu_2 \neq \mu_3$

(iii) If assumptions are met, choose the fixed effects ANOVA to evaluate the hypotheses.

(iv) The critical value for the *F*-test with $\alpha = 0.05$ and $df = 2{,}57$ is 3.16 (*df* is defined in Table 8). If $F_{\text{observed}} > F_{\text{critical}}$, reject H_0.

(v) Data are displayed in Tables 3, 5, and 7 for the CBT, IPT, and MM groups, respectively.

(vi) Calculations are described above in text and in Table 8.

(vii) Make decision regarding H_0. Since $F_{\text{observed}} > F_{\text{critical}}$ (i.e., 10.966 > 3.16) reject H_0.

(viii) Conclusion: There is a significant difference in severity of depression among the treatments. And which is superior?

Table 7 Hamilton Depression Ratings (X_i) for 20 patients in the MM group.

	X_i	X_i^2	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
	16	256	−6.25	39.0625
	16	256	−6.25	39.0625
	17	289	−5.25	27.5625
	17	289	−5.25	27.5625
	19	361	−3.25	10.5625
	19	361	−3.25	10.5625
	19	361	−3.25	10.5625
	19	361	−3.25	10.5625
	20	400	−2.25	5.0625
	20	400	−2.25	5.0625
	21	441	−1.25	1.5625
	21	441	−1.25	1.5625
	22	484	−0.25	0.0625
	25	625	2.75	7.5625
	27	729	4.75	22.5625
	28	784	5.75	33.0625
	28	784	5.75	33.0625
	29	841	6.75	45.5625
	31	961	8.75	76.5625
	31	961	8.75	76.5625
Total	445	10385	0	483.75

$$\bar{X} = \frac{\sum_{1}^{n} X_i}{n} = \frac{445}{20} = 22.25$$

$$\text{SS} = \sum (X_i - \bar{X})^2 = 483.75$$

$$s^2 = \frac{\text{SS}}{n-1} = \frac{483.75}{20-1} = 25.46$$

$$s = \sqrt{s^2} = \sqrt{25.46} = 5.046$$

3.12.4.3.4 ANOVA: Post hoc *tests*

A significant omnibus *F*-test indicates that the null hypothesis is rejected: $H_0 : \mu_1 = \mu_2 = \mu_3$. Nonetheless, further analyses are necessary to determine where the difference lies, that is, which groups are significantly different from each other. A variety of *post hoc* tests can be used for this purpose. A *post hoc* test is designed to be used *only* after an omnibus *F*-test is statistically significant. Each test incorporates the mean square within (*MSW*), which comes from the calculations that are presented in the ANOVA table (Table 8). The *MSW* is the denominator of the *F*-test. A nonsignificant omnibus *F*-test, however, indicates that there is not a difference in means among the groups. Of course, when there is no difference, there is no need to look for differences. Consequently, the use of a *post hoc* test after a nonsignificant omnibus *F*-test is inappropriate.

Post hoc procedures provide a strategy for multiple comparisons while at the same time maintaining the probability of Type I error (i.e., the probability of detecting differences that do not actually exist). This is in stark contrast with the inappropriate strategy of following a significant omnibus *F*-test with several *t*-tests, each used to compare a pair of groups. That approach fails to account for the multiple tests, and as a consequence, increases the probability of Type I error. Three *Post hoc* tests are described below. The choice among these tests is based on the research question and the design of the study.

(i) Tukey's Honestly Significant Difference

Tukey's Honestly Significant Difference (HSD) test is used to compare all *pairs* of groups means. HSD can be applied when each of the pairwise comparisons is of interest to the researcher, or when the researcher does not have specific hypotheses about which group means will differ. Initially the groups means are calculated. In the hypothetical clinical trial example that has been discussed in this chapter, the groups are arranged in ascending order of their means: IPT ($\bar{X} = 15.05$), CBT ($\bar{X} = 21.00$), MM ($\bar{X} = 22.25$). Note that with k groups there are $\frac{k(k-1)}{2}$ pairwise comparisons. Thus, in this example, there are $\frac{k(k-1)}{2} = \frac{3(3-1)}{2} = 3$ comparisons: IPT vs. CBT, IPT vs. MM, and CBT vs. MM.

Table 8 ANOVA Table

Source	*SS*	*df*	*MS*	*F*
Treatment	592.033	2	296.017	10.966
Within	1538.697	57	26.995	
Total	2130.730	59		

$$
\begin{aligned}
\mathrm{SS}_{\mathrm{Total}} &= \sum_{i=1}^{n}(X_i - \bar{X}_{..})^2\\
&= (12 - 19.43)_2 + (12 - 19.43)^2\\
&\quad + \cdots + (31 - 19.43)^2(31 - 19.43)^2\\
&= 2130.733\\
\mathrm{SS}_{\mathrm{Treatment}} &= \sum_{j=1}^{p} n_j(X_j - \bar{X}_{..})^2\\
&= 20(21.00 - 19.43)^2\\
&\quad + 20(15.05 - 19.43)^2\\
&\quad + 20(22.25 - 19.43^2 = 592.033)\\
\mathrm{SS}_{\mathrm{Within}} &= \sum_{i=1}^{N}(\bar{X}_{ij} - \bar{X}_j)^2\\
&= (12 - 21)^2 + (12 - 21)^2\\
&\quad + \cdots + (31 - 21)^2\\
&\quad + (31 - 21)^2 = 1538.697\\
\mathrm{df}_{\mathrm{Treatment}} &= k - 1, \text{where } k = \text{number of groups}\\
\mathrm{df}_{\mathrm{Within}} &= N - k\\
\mathrm{df}_{\mathrm{Total}} &= N - 1\\
\mathrm{MS}_{\mathrm{Treatment}} &= \frac{\mathrm{SS}_{\mathrm{Treatment}}}{\mathrm{df}_{\mathrm{Treatment}}}\\
\mathrm{MS}_{\mathrm{Within}} &= \frac{\mathrm{SS}_{\mathrm{Within}}}{\mathrm{df}_{\mathrm{Within}}}\\
F &= \frac{\mathrm{MS}_{\mathrm{Treatment}}}{\mathrm{MS}_{\mathrm{Within}}}
\end{aligned}
$$

The three pairwise mean differences are then calculated. The HSD test statistic, q, a ratio of the mean difference over the standard error, is computed for each pairwise comparison. For each comparison, if the observed $q >$ critical q (a tabled value), the null hypothesis of no difference between the pair of means is rejected. When there are an equal number of subjects per group, the standard error (*se*) is calculated as:

$$se = \frac{\sqrt{MSE}}{n_k} = \frac{\sqrt{26.995}}{20} = 1.16$$

where MSE comes from the ANOVA table (Table 8) and n_k is the number of subjects per group. The algorithm for unequal sample sizes can be found in other texts (e.g., Zar, 1996). The results are presented in Table 9.

The hypothesis testing steps for Tukey's HSD are as follows:

(i) State research question. The ANOVA results indicated that the IPT, CBT, and MM do not have similar effects for the treatment of patients with depression. Which pairs of groups differ?

(ii) State two mutually exclusive and exhaustive hypotheses for each pairwise comparison:

(a) $H_0: \mu_{\mathrm{IPT}} = \mu_{\mathrm{CBT}}$
$H_A: \mu_{\mathrm{IPT}} \neq \mu_{\mathrm{CBT}}$
(b) $H_0: \mu_{\mathrm{IPT}} = \mu_{\mathrm{MM}}$
$H_A: \mu_{\mathrm{IPT}} \neq \mu_{\mathrm{MM}}$
(c) $H_0: \mu_{\mathrm{CBT}} = \mu_{\mathrm{MM}}$
$H_A: \mu_{\mathrm{CBT}} \neq \mu_{\mathrm{MM}}$

(iii) Choose Tukey's HSD as a *post hoc* test to evaluate the hypotheses.

(iv) The critical value for the q-test with $\alpha = 0.05$ and df = 3,56 is 3.408. (Note that $df = k, v$, where $k =$ number of groups and $v = n - k - 1$.) Reject H_0 if $q > 3.408$.

(v) Data are displayed in Table 9 for the three pairwise comparisons.

(vi) Calculations were described above.

(vii) Decisions regarding H_0.

IPT vs. CBT: $q_{\mathrm{observed}} > q_{\mathrm{critical}}$
(i.e., 5.13 > 3.408), thus reject H_0.
IPT vs. MM: $q_{\mathrm{observed}} > q_{\mathrm{critical}}$
(i.e., 6.21 > 3.408), thus reject H_0.
CBT vs. MM: $q_{\mathrm{observed}} < q_{\mathrm{critical}}$
(i.e., 1.07 < 3.408), thus do not reject H_0.

(viii) Conclusion: IPT was more effective than either CBT or MM for the treatment of depressive symptoms in patients with major depressive disorder. There was no difference between CBT and MM for the treatment of depressive symptoms.

(ii) The Dunnett test

The Dunnett test is to be used when a control group is to be compared with each of several other treatments, and when no other comparisons are of interest. For instance, it could be used to demonstrate the efficacy of each treatment relative to placebo, but not to examine differences between active treatments. In the example that has been discussed, the

Table 9 Tukey HSD pairwise comparisons

Comparison	*Mean difference*	*se*	*q*
IPT vs. CBT	5.95	1.16	5.13
IPT vs. MM	7.20	1.16	6.21
CBT vs. MM	1.20	1.16	1.07

Dunnett test could be used for two comparisons: (i) CBT vs. MM, and (ii) IPT vs. MM. It could be used to show that IPT and CBT each have *active ingredients*; but the Dunnett test could not be used by members of the IPT and CBT camps to settle the score of whose method is superior. The Dunnett test, which will not be illustrated here, is described in detail in Fleiss (1986).

(iii) Scheffe Post Hoc *Procedure*

The Scheffe test is a more general procedure than those described above. It is not restricted to pairwise comparisons, but can evaluate comparisons of any linear combination of means. For instance, in the above example, one might have two comparisons of interest: (i) active (CBT and IPT) vs. placebo (MM), and (ii) CBT vs. IPT. The Scheffe procedure would be appropriate for those comparisons. The Scheffe test is somewhat more conservative than others in that it has less statistical power. For that reason some criticize the test for being more likely to miss true differences. Others consider the conservative nature of the test appealing. The Scheffe test will not be applied here. It is discussed in detail in Fleiss (1986).

3.12.4.3.5 Kruskal–Wallis test

The Kruskal–Wallis test (1952) is a nonparametric approach to the one-way ANOVA. The procedure is used to compare three or more groups on a dependent variable that is measured on at least an ordinal level. Ordinal data extends beyond rating scores such as the HRSD, and can include ordered categorical variables such as Hollingshead and Redlich's (1958) four broad categories of socioeconomic status. As with the Mann–Whitney test, which is a special two-group case of the Kruskal–Wallis test, the data are pooled (across groups) and ranked from *1* for the lowest value of the dependent variable to *N* for the highest value. In the case of ties, the midpoint is used. For example, if two subjects had the third lowest value, they would each be given a rank of 3.5, the midpoint of 3 and 4. The data from the hypothetical clinical trial comparing CBT, IPT, and MM (originally displayed in Tables 3, 5, and 7, respectively) have been ranked and are presented in Table 10.

After the data have been ranked, the sum of the ranks for each group is calculated:

$$R_j = \sum_{i=1}^{n_j} r_i$$

The mean rank for each of the three groups is then calculated:

$$\bar{r}_j = \frac{\sum_{i=1}^{n_j} r_i}{n_j}$$

The respective rank sums and means for each group are also presented in Table 10. Note that the IPT (18.075) group has a substantially lower mean rank than either the CBT (34.975) or MM (38.45) groups. The Kruskal–Wallis test can be used to determine whether that difference is larger than might be expected by chance. Stated differently, consider an experiment in which 20 numbers are randomly selected from a box that contains 60 balls, numbered 1 though 60. Would it be unlikely to randomly select a group of 20 balls whose sum is 361.5 or lower? Or whose mean is 18.075 or lower? In essence, the Kruskal–Wallis test addresses this question.

Once the ranking has been conducted and the sum of ranks for each group has been determined, the calculations for the Kruskal–Wallis test statistic, *T*, can proceed as follows:

$$T = \frac{(N-1)(S_t^2 - C)}{S_r^2 - C}$$

where

$$S_t^2 = \sum_{i=1}^{k} \frac{R_i^2}{n_i}$$

$$S_r^2 = \sum_{i=1}^{N} r_{ij}^2$$

$$C = \frac{N(N+1)^2}{4}$$

Applying these algorithms to the data that is presented in Table 10:

$$S_t^2 = \sum_{i=1}^{k} \frac{R_i^2}{n_i} = \frac{699.5^2}{20} + \frac{361.5^2}{20} + \frac{769^2}{20}$$
$$= 60567.175$$

$$S_r^2 = \sum_{i=1}^{N} r_{ij}^2$$
$$= 1^2 + 2^2 + 3^2 + ... + 56^2 + 59^2 + 59^2$$
$$= 73587$$

$$C = \frac{N(N+1)^2}{4} = \frac{60(60+1)^2}{4} = 55815$$

$$T = \frac{(N-1)(S_t^2 - C)}{S_r^2 - C}$$
$$= \frac{(60-1)(60567.175 - 55815)}{73587 - 55815} = 15.776$$

Table 10 HRSD for 60 patients in the clinical trial of CBT, IPT, and MM: raw (X_{ij}) and ranked (R_{ij}) values.

	CBT		*IPT*		*MM*	
	X_{ij}	R_{ij}	X_{ij}	R_{ij}	X_{ij}	R_{ij}
	12	7	6	1	16	14.5
	12	7	7	2	16	14.5
	13	10.5	8	3	17	18
	14	12	9	4	17	18
	15	13	12	7	19	30
	17	18	12	7	19	30
	18	22	12	7	19	30
	19	30	13	10.5	19	30
	21	41	17	18	20	37.5
	21	41	17	18	20	37.5
	21	41	18	22	21	41
	22	44.5	18	22	21	41
	24	46	19	30	22	44.5
	25	47.5	19	30	25	47.5
	26	49	19	30	27	51
	27	51	19	30	28	54
	27	51	19	30	28	54
	28	54	19	30	29	57
	29	57	19	30	31	59.5
	29	57	19	30	31	59.5
Total	420	699.5	301	361.5	445	769
$\bar{R}$		34.975		18.075		38.45

Assumptions of the Kruskal–Wallis test include:

(i) The samples are independent random samples from their respective populations.

(ii) The scale of measurement (of the dependent variable) is at least ordinal.

The hypothesis testing steps for the Kruskal–Wallis test are as follows:

(i) State research question. Do CBT, IPT, and MM have similar effects for the treatment of patients with depression?

(ii) State two mutually exclusive and exhaustive hypotheses with regard to group medians:

(a) $H_0 : M_1 = M_2 = M_3$

(b) $H_A : M_1 \neq M_2 \neq M_3$

(iii) Use the Kruskal–Wallis test to evaluate the hypotheses.

(iv) The critical value for the Kruskal–Wallis test comparing k groups comes from an χ^2 distribution, with $k-1$ degrees of freedom and $\alpha = 0.05$. In this case there are three groups ($k = 3$) and df $= 3-1 = 2$. Therefore, the critical $\chi^2_{(2,.05)} = 5.99$. If $T_{\text{observed}} > 5.99$, reject H_0

(v) Data are displayed in Table 10 for the three groups.

(vi) Calculations are described above.

(vii) Decision regarding H_0. $T_{\text{observed}} > \chi^2_{\text{critical}}$ (i.e., $15.77 > 5.99$), thus reject H_0.

(viii) Conclusion. There is a significant difference in severity of depression among the treatments. (Which groups actually differ can only be determined using an appropriate *post hoc* test, as discussed below.)

This significant result in a Kruskal–Wallis test indicates that there are group differences, but does not indicate which groups differ. As with an ANOVA, a *post hoc* procedure that is analogous to the HSD for ANOVAs can be used to determine which groups are significantly different from each other. That procedure, which will not be discussed here, is described in detail by Conover (1980). Note that in this case, the Kruskal–Wallis results are similar to those of the ANOVA. This is not always the case. The choice between the two tests should be based on whether or not the assumptions of the test are fulfilled and *not* on a comparison of their results. When the results from an ANOVA and a Kruskal–Wallis test conflict, the degree of discrepancy between the two tests may, in part, reflect the extent to which the assumptions were fulfilled.

3.12.4.3.6 Factorial analysis of variance

The fixed effect ANOVA model that was just discussed can be extended to include more than one independent variable. Consider a clinical trial in which the two treatments (CBT and IPT) were compared among samples of two types of mood disorders (major depressive disorder and dysthymia). In such a design the two fixed *factors* (independent variables) are treatment and diagnosis, which each have two *levels*. The two-way ANOVA is used to test *main effects* of each independent variable and the *interaction* between the independent variables. The null hypotheses are:

Main effect of treatment	$H_0 : \mu_{1.} = \mu_{2.}$
Main effect of diagnosis	$H_0 : \mu_{.1} = \mu_{.2}$
Treatment by diagnosis interaction	$H_0 : \mu_{11} - \mu_{12} = \mu_{21} - \mu_{22}$

Note that the first subscript represents treatment level (where 1 = CBT and 2 = IPT), the second subscript represents diagnosis (where 1 = dysthymia and 2 = MDD). In the null hypotheses for the main effects, a dot (.) in the subscript indicates that the corresponding particular variable is ignored. These three null hypotheses are tested with data from one experiment. The *two* main effects in this two-way ANOVA represent an extension of the one main effect that is tested in a one-way ANOVA model described earlier. What is more novel about a *factorial* ANOVA design is the interaction between the factors, which is used to examine the moderating effect one factor has on the effect that the other factor has on the dependent variable. In the hypothetical study presented here, an interaction is used to compare treatment differences (CBT vs. IPT) across diagnostic groups. Stated differently, it is used to examine whether there is a larger difference between IPT and CBT for subjects with MDD or for subjects with dysthymia.

The data from the clinical trial shown earlier in Tables 3 and 5 have been stratified by diagnosis and are presented in Table 11. In a simple comparison of the means, it is clear that the IPT group has a substantially lower mean than the CBT group. Likewise, the dysthymia group has a lower mean than the MDD group. In both cases, the differences are of the order of one standard deviation. The factorial ANOVA is used to evaluate whether or not these differences are larger than might be expected by chance. Furthermore, the ANOVA will be used to compare the treatment effects across diagnostic groups.

The calculations for a two-way factorial ANOVA involve partitioning the total sums of squares into four components: SS_{Tx}, SS_{Dx}, SS_{TxbyDx}, SS_{within}. The algorithms are described in Table 12 and illustrated in Table 13. Note that the subscript g represents treatment ($g = 1,2$), the subscript h represents diagnosis ($h = 1,2$), and the subscript i represents individual subject ($i = 1$–N; $N = 40$ in this case). The formulae that are presented are applicable for designs with cells of equal size (n; $n = 10$ in this case) and are presented to illustrate the concept of the partitioning of sums of squares in a two-way ANOVA. Notice that each estimates a different form of sum of squared deviations (SS) from a mean that represents a component of the variability among the HRSD of all subjects in the clinical trial.

In order to conduct the F-tests of the ANOVA that correspond to the three null hypotheses, several additional calculations are required. The algorithms for these quantities are presented in Table 12. The first column of the table indicates the *source of variability*. The second column displays the SS (calculated above) that correspond to each of the four sources of variability. The third column of Table 12 presents the degrees of freedom that correspond to each source of variability. The calculation of degrees of freedom is described in Table 12. The fourth column in Table 12 presents the mean square (*MS*) for each source of variability which is equal to the sum of squares for that source divided by the corresponding degrees of freedom. In the final column of Table 12, the F-test for each effect is equal to the mean squares for that effect divided by the mean square within (MS_{within}). The results of the data from the example are shown in Table 13.

Assumptions of the factorial ANOVA are as follows:

(i) The observations in each of the cells are independent samples.

(ii) The population data are normally distributed.

(iii) The populations all have equal variances.

The hypothesis testing steps for the factorial ANOVA are as follows:

(i) State research questions.

(a) Do CBT and IPT have similar effects on depressive symptomatology for the treatment of patients with depression?

(b) Do patients with major depressive disorder and dysthymia have similar depressive symptomatology after the treatment of depression?

(c) Do CBT and IPT have similar effects on depressive symptomatology for patients with major depressive disorder and dysthymia?

(ii) State sets of mutually exclusive and exhaustive hypotheses:

(a) Main effect of treatment

$H_0 : \mu_{1\cdot} = \mu_{2\cdot}$

$H_A : \mu_{1\cdot} \neq \mu_{2\cdot}$

(b) Main effect of diagnosis

$H_0 : \mu_{\cdot 1} = \mu_{\cdot 2}$

$H_A : \mu_{\cdot 1} \neq \mu_{\cdot 2}$

(c) Interaction of treatment by diagnosis

$H_0: \mu_{11} - \mu_{21} = \mu_{12} - \mu_{22}$

$H_A: \mu_{11} - \mu_{21} = \mu_{12} - \mu_{22}$

(iii) If assumptions have been fulfilled, use the two-way factorial ANOVA to evaluate the hypotheses.

(iv) The critical F-values are selected from an F-table based on the degrees of freedom numerator, degrees of freedom denominator, and alpha-level. The critical value for the F-test with $\alpha = 0.05$ and $df = 1,36$ is 5.47. Reject H_0 if $F_{\text{observed}} > F_{\text{critical}}$.

(v) Data are displayed in Table 11.

(vi) Calculations are described above.

(vii) Decisions regarding the three H_0s.

(a) Main effect of treatment

The main effect of treatment is statistically significant:

$F_{\text{obs}} > F_{\text{critical}(1,36;.05)}$ (i.e., $17.114 > 5.47$).

(b) Main effect of diagnosis

The main effect of diagnosis is statistically significant:

$F_{\text{obs}} > F_{\text{critical}(1,36;.05)}$ (i.e., $14.359 > 5.47$).

(c) Interaction of treatment by diagnosis

The treatment by diagnosis interaction is not statistically significant:

$F_{\text{obs}} < F_{\text{critical}(1,36;.05)}$ (i.e., $0.639 < 5.47$).

(viii) Conclusion. IPT is a significantly more effective treatment than CBT for depressive symptoms. Patients with major depressive disorder have significantly more depressive symptomatology than those with dysthymia after treatment. The differential effects of IPT and CBT are similar for patients with major depressive disorder and patients with dysthymia.

In this case, *post hoc* tests such as those described earlier are not needed to interpret these results because there are only two levels of each significant factor and there is not a significant interaction. To determine which level of each factor had a better outcome in this clinical trial, simply compare the means of each level of the significant effects and determine which is lower (i.e., less depressed on the HRSD). For the treatment effect, the IPT group mean (sd) is substantially lower than that of the CBT group: 15.05 (4.67) vs. 21.00 (5.80). The two groups differ by more than one standard deviation, which is a very large difference. Similarly, the dysthymia group mean (sd) is substantially lower than that of the MDD group: 15.30 (5.57) vs. 20.75 (5.24). Here the groups differ by slightly less than one standard deviation, which is

Table 11 Hamilton Depression Ratings for 40 patients in the clinical trial of CBT and IPT stratified by diagnosis.

Treatment		CBT	CBT	IPT	IPT
Diagnosis		DYS	MDD	DYS	MDD
		12	14	6	12
		12	19	7	13
		13	26	8	17
		15	27	9	17
		17	22	12	18
		18	24	12	19
		21	25	18	19
		21	28	19	19
		21	29	19	19
		27	29	19	19
Four cells	sum	177	243	129	172
	mean	17.70	24.30	12.90	17.20
	sd	4.88	4.81	5.38	2.62
	SS	3347	6113	1925	3020
Diagnosis		DYS	MDD		
	sum	306	415		
	mean	15.30	20.75		
	sd	5.57	5.24		
	SS	5272	9133		
Treatment		CBT	IPT		
	sum	420.00	301.00		
	mean	21.00	15.05		
	sd	5.80	4.67		
	SS	9460	4945		

Table 12 Two-way ANOVA table.

Source	*SS*	*df*	*MS*	*F*
Treatment	SS_{Tx}	$p-1$	$\text{SS}_{\text{Tx}}/(p-1)$	$\text{MS}_{\text{Tx}}/\text{MS}_{\text{within}}$
Diagnosis	SS_{Dx}	$q-1$	$\text{SS}_{\text{Dx}}/(q-1)$	$\text{MS}_{\text{Dx}}/\text{MS}_{\text{within}}$
Tx by dx	$\text{SS}_{\text{Tx by Dx}}$	$(p-1)(q-1)$	$\text{SS}_{\text{Tx by dx}}/(p-1)(q-1)$	$\text{MS}_{\text{Tx by dx}}/\text{MS}_{\text{within}}$
Within	$\text{SS}_{\text{within}}$	$pq(n-1)$	$\text{SS}_{\text{within}}/pq(n-1)$	
Total	SS_{Total}	$n-1$		

$$\text{SS}_{\text{Total}} = \sum_{i=1}^{N}(X_{ghi} - \bar{X}...)^2$$

$$\text{SS}_{\text{T}_{\text{X}}} = nq\sum_{j=1}^{p}(\bar{X}_{g..} - \bar{X}...)^2$$

$$\text{SS}_{\text{D}_{\text{X}}} = np\sum_{k=1}^{q}(\bar{X}_{.h.} - \bar{X}_{...})^2$$

$$\text{SS}_{\text{TxbyDx}} = n\sum_{g=1}^{p}\sum_{h=1}^{q}\left[(\bar{X}_{gh.} - \bar{X}_{...}) - (\bar{X}_{g..} - \bar{X}_{...}) - (\bar{X}_{.h.} - \bar{X}_{...})\right]^2$$

$$\text{SS}_{\text{Within}} = \sum_{g=1}^{p}\sum_{h=1}^{q}\sum_{i=1}^{n}(\bar{X}_{ghi} - \bar{X}_{gh.})^2$$

again a substantial difference. In contrast, the interaction was not significant. Consider the differences in treatment means across diagnoses:

$$\bar{X}_{CBT/DYS} - \bar{X}_{IPT/DYS} = 17.70 - 12.90 = 4.80$$

and

$$\bar{X}_{CBT/MDD} - \bar{X}_{IPT/MDD} = 24.30 - 17.20 = 7.10$$

Relative to the variation within treatment (marginal standard deviations range from 4.67 to 5.80), the difference in mean differences (7.10 − 4.80 = 2.30), is not substantial enough to consider it more than might be expected by chance.

The *two-way factorial ANOVA* example that was just discussed included two factors (i.e., independent variables: treatment and diagnosis), each of which had two levels. The factorial ANOVA can incorporate factors which have more than two levels each. This was shown in the earlier example of the one-way ANOVA in which there were three levels of treatment (CBT, IPT, and MM). In addition, the factorial design can be extended to include more than two factors. For instance, the clinical trial could have been conducted to test the effect of gender as well as treatment and diagnosis. Such a design is called a *three-way ANOVA*, which includes tests of three main effects (treatment, diagnosis, and gender), three two-way interactions (treatment by diagnosis, treatment by gender, and diagnosis by gender), and one three-way interaction (treatment by diagnosis by gender). For an extensive discussion of ANOVA designs see Keppel (1991).

In this chapter, two multivariate statistical procedures will be introduced, but not discussed in depth. For a comprehensive discussion of these and other multivariate tests, see Bock (1985), Grimm and Yarnold (1995), Marcoulides and Hershberger (1997), or Stevens (1992). Here multivariate means that multiple dependent variables are being examined in one test. For instance, consider the comparison of affectively ill patients with schizophrenic patients on a battery of six personality scales. There are two schools of thought for use of a multivariate procedure such as this. One is that it should be used if there is a multivariate research question. For example, do the personalities of the two diagnostic groups differ as a whole? The researcher who asks this question, might be most interested in personality as a general construct, and less interested in the specific differences in ego resiliency, dependency and other personality characteristics. In contrast, others would argue that the multivariate procedures should be used to protect against Type I error whenever several dependent variables are examined.

Table 13 Hamilton depression ratings for clinical trial of CBT and IPT by diagnosis: ANOVA table.

Source	*SS*	*df*	*MS*	*F*
Treatment	354.025	1	354.025	17.114
Diagnosis	297.025	1	297.025	14.359
Tx by dx	13.225	1	13.225	0.639
Within	744.70	36	20.686	
Total	1408.975	39		

$$\begin{aligned}
\mathrm{SS}_{\mathrm{Total}} &= \sum_{i=1}^{n}(X_{ghi}-\bar{X}_{...})^2 \\
&= (12-18.025)^2+(12-18.025)^2 \\
&\quad + \cdots + (19-18.025)^2 \\
&\quad + (19-18.025)^2 = 1408.975
\end{aligned}$$

$$\begin{aligned}
\mathrm{SS}_{\mathrm{TX}} &= nq\sum_{j=1}^{q}(\bar{X}_{g..}-\bar{X}_{...})^2 \\
&= 10\cdot 2\cdot\Big[(21-18.025)^2 \\
&\quad + (15.05-18.025)^2\Big] = 354.025
\end{aligned}$$

$$\begin{aligned}
\mathrm{SS}_{\mathrm{DX}} &= np\sum_{k=1}^{q}(\bar{X}_{.h.}-\bar{X}_{...})^2 \\
&= 10\cdot 2\cdot\Big[(15.3-18.025)^2 \\
&\quad + (20.75-18.025)^2\Big] = 297.025
\end{aligned}$$

$$\begin{aligned}
\mathrm{SS}_{\mathrm{TxbyDx}} &= n\sum_{g=1}^{p}\sum_{h=1}^{q}\Big[(\bar{X}_{gh.}-\bar{X}_{...})-(\bar{X}_{g..}-\bar{X}_{...}) \\
&\quad -(\bar{X}_{.h.}-\bar{X}...)\Big]^2 \\
&= 10\cdot\Big[(17.7-18.025)+(21.0-18.025) \\
&\quad + (15.3-18.025)\Big]^2+\cdots \\
&\quad \cdots+\Big[(17.2-18.025)+(20.75-18.025) \\
&\quad + (15.05-18.025)\Big]^2 = 13.225
\end{aligned}$$

$$\begin{aligned}
\mathrm{SS}_{\mathrm{Within}} &= \sum_{g=1}^{p}\sum_{h=1}^{q}\sum_{i=1}^{n}(\bar{X}_{ghi}-\bar{X}_{gh})^2 \\
&= (12-17.7)^2+(12-17.7)^2 \\
&\quad + \cdots + (19-17.2)^2+(19-17.2)^2 \\
&= 744.70
\end{aligned}$$

3.12.4.3.7 Hotelling's T^2

Hotelling's T^2 is used to compare two groups on multiple continuous dependent variables, simultaneously. Consider that if a series of six univariate *t*-tests were performed on these scales, the probability of a Type I error would be substantially elevated from the conventional 0.05 level. In general, the experimentwise alpha level for *c* comparisons is calculated as follows: $\alpha_{\mathrm{EW}} = 1-(1-\alpha)^c$. Thus, in lieu of a multivariate procedure, the two groups might be compared using six univariate *t*-tests (one for each of the six dependent variables—the personality scales) and the resulting experimentwise alpha level would be $\alpha_{\mathrm{EW}} = 1-(1-0.05)^6 = 1-0.735 = 0.265$, which certainly exceeds the conventional Type I error level of 0.05.

Although the computational details of Hotelling's T^2 are beyond the scope of this chapter, some general issues regarding the procedure will be introduced. First, the null hypothesis is that the populations from which the two groups come do not differ on (a linear combination) of the scales (see below). Second, assumptions of Hotelling's T^2 include multivariate normality and independence of observations across subjects. Third, Hotelling's T^2 is an omnibus multivariate procedure. If, and only if, the multivariate null hypothesis is rejected, can subsequent *post hoc* univariate tests be used to identify the source of the difference (i.e., on which dependent variable(s) do the groups differ?). This is similar to the approach to the comparison of three or more groups in an ANOVA where, if the omnibus *F*-test is significant, a *post hoc* test must be used to determine which groups differ significantly. However, those *post hoc* tests cannot be used if the omnibus *F* of the univariate ANOVA is not significant. In the case of Hotelling's T^2, the *post hoc* tests would simply be a *t*-test for each dependent variable.

The hypothesis testing steps for Hotelling's T^2 are as follows:

(i) State research question. Are the personality profiles of affectively ill patients and schizophrenic patients similar (using a battery of six personality scales)?

(ii) State two mutually exclusive and exhaustive hypotheses:

$$H_0 : \begin{bmatrix}\mu_{11}\\ \mu_{21}\\ \mu_{31}\\ \mu_{41}\\ \mu_{51}\\ \mu_{61}\end{bmatrix} = \begin{bmatrix}\mu_{12}\\ \mu_{22}\\ \mu_{32}\\ \mu_{42}\\ \mu_{52}\\ \mu_{62}\end{bmatrix}$$

$$H_A : \begin{bmatrix}\mu_{11}\\ \mu_{21}\\ \mu_{31}\\ \mu_{41}\\ \mu_{51}\\ \mu_{61}\end{bmatrix} \neq \begin{bmatrix}\mu_{12}\\ \mu_{22}\\ \mu_{32}\\ \mu_{42}\\ \mu_{52}\\ \mu_{62}\end{bmatrix}$$

(iii) If the assumptions are met, use the Hotelling's T^2 evaluate the hypotheses.

(iv) The critical value, calculations, decision rules, and interpretation of Hotelling's T^2 are described elsewhere (e.g., Stevens (1992) or Marcoulides & Hershberger (1997)).

3.12.4.3.8 Multivariate analysis of variance

Multivariate analysis of variance (MANOVA) is an extension of the T^2 for the comparison of three or more groups. For example, three groups (e.g., mood disorders, schizophrenics, and no history of a mental disorder) can be compared on a battery of six personality scales using a MANOVA. Similar to the factorial ANOVA, MANOVA can also be extended to incorporate more than one factor. For instance, a two-way MANOVA could evaluate the main effects and interaction of gender and diagnosis (mood disorders vs. schizophrenics vs. no history of a mental disorder) in which the dependent variables are a battery of six personality scales. If any of those three omnibus multivariate tests (i.e, main effects of gender or diagnosis or the gender by diagnosis interaction) are statistically significant, subsequent *post hoc* tests must be conducted to determine the source of the group differences. The choice among multivariate test statistics (e.g., Pillais, Hotelling's, Wilks, or Roy) is considered by Olson (1976). *Post hoc* tests for MANOVA are described in detail by Stevens (1992).

Assumptions of MANOVA include:

(i) independence among observations (i.e., subjects)

(ii) a multivariate normal distribution among the dependent variables

(iii) equal population covariance matrices for the dependent variables.

The hypothesis testing steps for MANOVA are as follows:

(i) State research question. Are the personality profiles of affectively ill patients, schizophrenic patients, and those with no history of mental disorders similar (using a battery of six personality scales)?

(ii) State two mutually exclusive and exhaustive hypotheses:

$$H_0: \begin{bmatrix} \mu_{11} \\ \mu_{21} \\ \mu_{31} \\ \mu_{41} \\ \mu_{51} \\ \mu_{61} \end{bmatrix} = \begin{bmatrix} \mu_{12} \\ \mu_{22} \\ \mu_{32} \\ \mu_{42} \\ \mu_{52} \\ \mu_{62} \end{bmatrix} = \begin{bmatrix} \mu_{13} \\ \mu_{23} \\ \mu_{33} \\ \mu_{43} \\ \mu_{53} \\ \mu_{63} \end{bmatrix}$$

$$H_A: \begin{bmatrix} \mu_{11} \\ \mu_{21} \\ \mu_{31} \\ \mu_{41} \\ \mu_{51} \\ \mu_{61} \end{bmatrix} \neq \begin{bmatrix} \mu_{12} \\ \mu_{22} \\ \mu_{32} \\ \mu_{42} \\ \mu_{52} \\ \mu_{62} \end{bmatrix} \neq \begin{bmatrix} \mu_{13} \\ \mu_{23} \\ \mu_{33} \\ \mu_{43} \\ \mu_{53} \\ \mu_{63} \end{bmatrix}$$

(iii) If the assumptions are met, use a MANOVA to evaluate the hypotheses.

(iv) The critical value, calculations, decision rules, and interpretation of MANOVA are described elsewhere (e.g., Marcoulides & Hershberger, 1997; Pedhazur, 1982; Stevens, 1992).

3.12.4.4 Linear Relations

3.12.4.4.1 Pearson correlation coefficient

The strength of the bivariate linear relation between variables can be examined with a correlation coefficient. The Pearson product moment correlation coefficient, r, is most often used for this purpose. In correlational analyses, a distinction need not be made between the independent and dependent variables. Unlike the univariate analyses that were discussed earlier, each subject contributes data in the form of paired coordinates (x, y) to the analysis.

Consider the data in Figure 12 which represents the relation between depression (on the x-axis) and functional impairment (on the y-axis). (These same data are presented in the first two columns of Table 14.) If there were even a modest linear relation between depression and functional impairment, one would expect minimal impairment in the mildly depressed and greater impairment in the severely depressed. It is a correlation coefficient that quantifies the strength of that linear relationship. (Once again, as a point of clarification, unlike much of the prior discussion in this chapter, here we consider two variables from one group of subjects.)

A brief digression may help with this explanation. It is quite likely that the two variables in a correlational analysis are measured on very different scales. For instance, in this sample, the impairment and depression means (21.00 and 14.45, respectively) and standard deviations (5.80 and 4.93, respectively) differ somewhat. For that reason, consider a scale-free metric, a standardized score, referred to as a "z-score," which has the property of a mean of zero and standard deviation of one. A transformation from the original scale (X_i) to the standardized scale (Z_i) is quite simple. The original variable (X_i) is transformed by subtracting the mean and dividing by the standard deviation:

Table 14 Hamilton depression ratings and Sheehan Disability Scale for CBT group.

	X_i	Y_i	X_i^2	Y_i^2	X_iY_i	$X_i - \bar{X}$	$Y_i - \bar{Y}$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$	$x \cdot y$	z_x	z_y	$Z_x \cdot Z_y$
	12	11	144	121	132	−9.00	−3.45	81.00	11.90	31.05	−1.55	−0.70	1.09
	12	8	144	64	96	−9.00	−6.45	81.00	41.60	58.05	−1.55	−1.31	2.03
	13	6	169	36	78	−8.00	−8.45	64.00	71.40	67.60	−1.38	−1.72	2.36
	14	14	196	196	196	−7.00	−0.45	49.00	0.20	3.15	−1.21	−0.09	0.11
	15	11	225	121	165	−6.00	−3.45	36.00	11.90	20.70	−1.03	−0.70	0.72
	17	8	289	64	136	−4.00	−6.45	16.00	41.60	25.80	−0.69	−1.31	0.90
	18	16	324	256	288	−3.00	1.55	9.00	2.40	−4.65	−0.52	0.31	−0.16
	19	8	361	64	152	−2.00	−6.45	4.00	41.60	12.90	−0.34	−1.31	0.45
	21	15	441	225	315	0.00	0.55	0.00	0.30	0.00	0.00	0.11	0.00
	21	22	441	484	462	0.00	7.55	0.00	57.00	0.00	0.00	1.53	0.00
	21	12	441	144	252	0.00	−2.45	0.00	6.00	0.00	0.00	−0.50	0.00
	22	19	484	361	418	1.00	4.55	1.00	20.70	4.55	0.17	0.92	0.16
	24	15	576	225	360	3.00	0.55	9.00	0.30	1.65	0.52	0.11	0.06
	25	19	625	361	475	4.00	4.55	16.00	20.70	18.20	0.69	0.92	0.64
	26	16	676	256	416	5.00	1.55	25.00	2.40	7.75	0.86	0.31	0.27
	27	22	729	484	594	6.00	7.55	36.00	57.00	45.30	1.03	1.53	1.58
	27	18	729	324	486	6.00	3.55	36.00	12.60	21.30	1.03	0.72	0.75
	28	21	784	441	588	7.00	6.55	49.00	42.90	45.85	1.21	1.33	1.60
	29	11	841	121	319	8.00	−3.45	64.00	11.90	−27.60	1.38	−0.70	−0.97
	29	17	841	289	493	8.00	2.55	64.00	6.50	20.40	1.38	0.52	0.71
Total	420	289	9460	4637	6421.00	0.00	0.00	640.00	460.95	352.00	0.00	0.00	12.31
Mean	21.00	14.45	473.00	231.85	321.05	0.00	0.00	32.00	23.05	17.60	0.00	0.00	0.62
sd	5.804	4.925	239.790	142.538	165.468	5.803	4.925	28.038	22.404	23.171	1	1	0.8106

$x_i = X_i - \bar{X}$; $y_i = Y_i - \bar{Y}$; $Z_{x_i} = \frac{X_i - \bar{X}}{s_x}$; $Z_{y_i} = \frac{Y_i - \bar{Y}}{s_y}$.

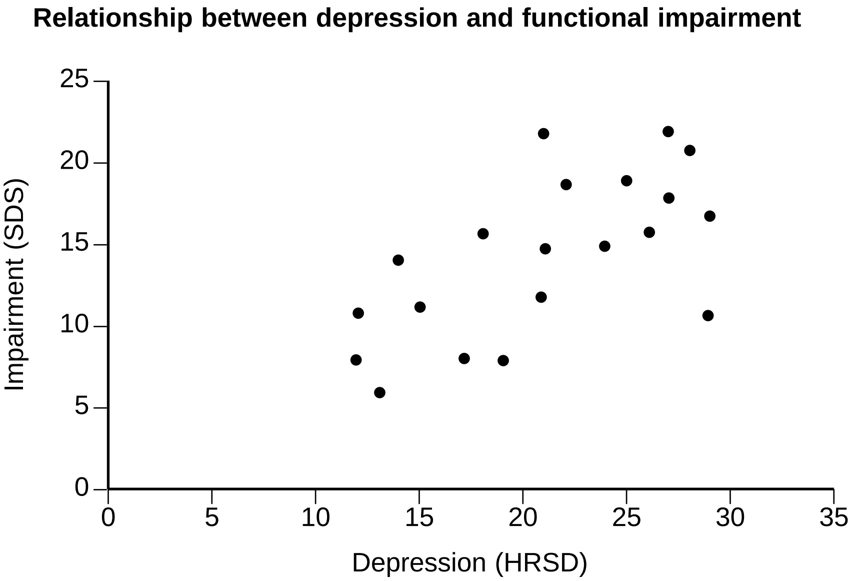

Figure 12 Plot of Hamilton Depression Ratings and Sheehan Disability Scale for CBT group.

$$Z_i = \frac{X_i - \bar{X}}{s_x}$$

The plot of standardized depression and functional impairment in Figure 13 looks identical to that of Figure 12, with the exception of the scale of the axes. (Note that in order to distinguish between the standardized X_is and standardized Y_is, Xs and Ys will be included in the subscript of the standardized scores: Z_{xi} and Z_{yi}.) As a consequence of the standardization, the Z_x-axis is plotted at $Y = 0$, which is the unstandardized y-mean, and the Z_y-axis is plotted at $X = 0$, the unstandardized x-mean.

As stated earlier, if there were a modest linear relation between depression and impairment, there would be minimal impairment in the mildly depressed and greater impairment in the severely depressed. Consider this positive linear relationship as presented in Figure 12. The points that are plotted tend to cluster either in quadrant I or in quadrant III. The latter fall in the upper-right quadrant, where subjects tend to have both relatively elevated depression and relatively elevated functional impairment. Stated differently, both the x-values and y-values are above the mean in Figure 12, and the standardized z_x and z_y scores are both positive in Figure 13. On the other hand, the subjects who have both relatively lower depression and relatively lower functional impairment are in the lower-left quadrant. That is, both the x-values and y-values are below the mean. In contrast, a negative linear relationship would characterize an association in which a high x-value tends to correspond with a low y-value, and a low x-value corresponds with a high y-value (e.g., depression and work productivity). Nunnally (1978) shows that the Pearson r can be calculated as approximately the mean of the sum of the product of z-scores:

$$r_{xy} = \frac{\sum_{i=1}^{N} Z_x Z_y}{N - 1}$$

With this algorithm it can be seen that if most subjects tend to have either high Z_xs and high Z_ys, or low Z_xs and low Z_ys, r will be larger and positive, approaching 1.0. In contrast, if most subjects tend to have either high Z_xs with low Z_ys, or low Z_xs with high Z_ys, r will have a large absolute value, but will be negative, approaching -1.0.

Using the data that are presented in Table 14:

$$r = \frac{\sum_{i=1}^{N} z_x z_y}{N - 1} = \frac{12.31}{20 - 1} = 0.648$$

The presentation of the algorithm for the Pearson correlation coefficient (r) usually comes from a seemingly different perspective which will now be discussed. The correlation coefficient is derived from the covariance. First, consider that the sum of cross-products ($\sum xy$) is a sum of the product of both the x-deviations (x) and y-deviations (y) from their respective means: $\sum xy = \sum (X_i - \bar{X})(Y_i - \bar{Y})$.

For instance, the first subject in Table 14 has an X-value (HRSD) of 12 and a Y-value (SDS) of 11. The respective deviation scores are:

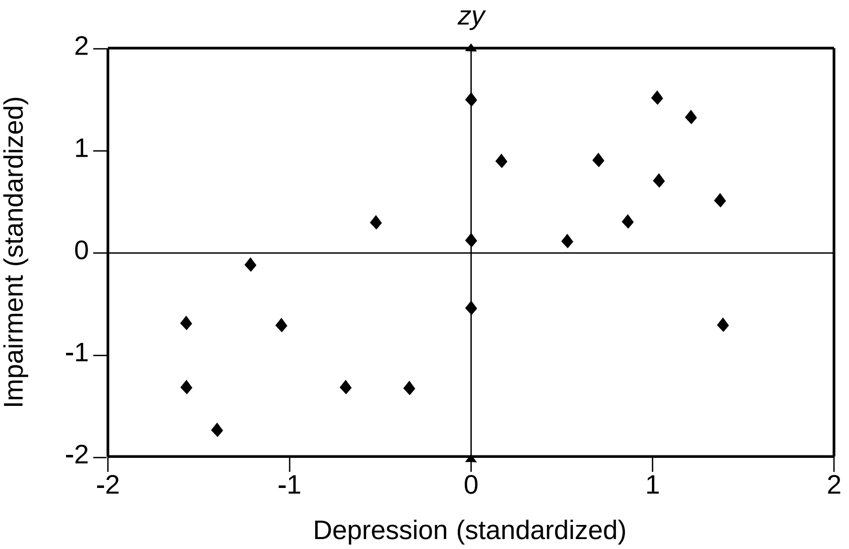

Figure 13 Plot of standardized Hamilton Depression Ratings and standardized Sheehan Disability Scale for subjects in the CBT group ($N = 20$).

$$x_1 = X_1 - \bar{X} = 12 - 21 = -9$$
$$y_1 = Y_1 - \bar{Y} = 11 - 14.45 = -3.45$$

The product of the deviation scores is called the cross-product. For the first subject, the cross-product is:

$$x_1 y_1 = -9 \cdot (-3.45) = 31.05$$

Once again, the sum of cross-products $(\sum_{i=1}^{N} x_i y_i)$ or $(\sum xy)$ is a sum of each subject's product of deviation scores. Notice that the sum of cross-products is a bivariate extension of the sum of squares that was described earlier, but the sum of cross-products incorporates two variables:

$$\sum x^2 = \sum (X_i - \bar{X})(X_i - \bar{X})$$

The covariance (s_{xy}), therefore, is a bivariate analogue of the variance. Consequently, the covariance is a function of the sum of cross-products:

$$s_{xy} = \frac{\sum xy}{n-1}$$

Using the equations above, the covariance (s_{xy}) can also be calculated:

$$s_{xy} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{n-1}$$

The covariance (s_{xy}) is a useful quantity that also measures the strength of the linear relationship between two variables. However, its interpretability is somewhat limited because it is not scale-free. As a consequence, unlike the correlation coefficient (r), which ranges from negative to positive (i.e., $-1.0 \leqslant r \leqslant 1.0$), the covariance does not have a fixed range.

Notice that the variance of a variable can be thought of as a special case of the covariance of that variable with itself: $S_{xx} = \sum (X_i - \bar{X})(X_i - \bar{X})$. A variance–covariance matrix is a convenient format to display the variances and covariances among several variables:

$$\begin{array}{cccc} & X & Y & Z \\ X & s_X^2 & s_{XY} & s_{XZ} \\ Y & s_{YX} & s_Y^2 & s_{YZ} \\ Z & s_{ZX} & s_{ZY} & s_Z^2 \end{array}$$

The variances are on the main diagonal and the covariances are the remaining elements of the matrix. The matrix is symmetrical about the main diagonal. That is, in general, $s_{ij} = s_{ji}$. Or in the case, $s_{XY} = s_{YX}$.

Using the data from Table 14, the covariance is:

$$S_{xy} = \frac{\sum xy}{n-1} = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{n-1}$$
$$= \frac{352}{20-1} = 18.526$$

Turning back to the correlation coefficient, r can be calculated using these quantities. It is equal to the covariance (s_{xy}) of X_i and Y_i, divided by the product of their standard deviations:

$$r_{XY} = \frac{S_{XY}}{S_X S_Y} = \frac{18.526}{5.803 \cdot 4.925} = 0.648$$

As stated earlier, the correlation coefficient (r) can range from -1 to $+1$. The greater the absolute value of r, the greater the strength of the bivariate linear relationship.

The statistical significance of the population correlation coefficient, ρ, can be tested using the sample data. The null hypothesis, H_0:$\rho = 0$, is tested with the following ratio:

$$t = \frac{r}{s_r}$$

where s_r is the standard error of r,

$$s_r = \sqrt{\frac{1 - r^2}{n - 2}}$$

This t-ratio is tested with $n-2$ dfs. Applying this to the data from Table 14, the standard error of r is:

$$s_r = \sqrt{\frac{1 - r^2}{n - 2}} = \sqrt{\frac{1 - 0.648^2}{20 - 2}} = 0.1795$$

The t-ratio is:

$$t = \frac{r}{s_r} = \frac{0.648}{0.1795} = 3.61$$

The hypothesis testing steps for the Pearson correlation coefficient are as follows:

(i) State research question. Is there a linear relationship between depression and impairment?

(ii) State two mutually exclusive and exhaustive hypotheses:

(a) $H_0: \rho_{XY} = 0$

(b) $H_A: \rho_{XY} \neq 0$

(iii) Choose the t-ratio to evaluate the hypotheses about the Pearson correlation coefficient.

(iv) The critical value for the t-ratio with $\alpha = 0.05$ and $\text{df} = 18$ is 2.101. If $t_{\text{observed}} > t_{\text{critical}}$, reject H_0.

(v) Data regarding depression and impairment are displayed in Table 14.

(vi) Calculations were described above.

(vii) Decision regarding H_0. Since $t_{\text{observed}} > t_{\text{critical}}$ (i.e., $3.61 > 2.101$), reject H_0.

(viii) Conclusion. Based on data from this sample, we conclude that there is a statistically significant linear relationship between depression and impairment.

Although the null hypothesis, H_0: $\rho = 0$, was rejected, be cautious not to overinterpret the results of such a test. The hypothesis simply examines the likelihood of observing a correlation of the *magnitude* seen in the sample, when the population correlation is zero (i.e., $\rho = 0$). Nevertheless, it ignores the meaningfulness of the magnitude of r. There is a disproportionate emphasis on the *statistical significance* of r while at the same time a tendency to ignore the *magnitude* of r in the research literature.

The magnitude of the correlation can be thought of as the proportion of variance in Y_i explained by X_i. In fact, because in correlational analyses a distinction between independent variable and dependent variable is unnecessary, the magnitude of the correlation can also be thought of as the proportion of variance in X_i explained by Y_i. In either case, the magnitude of the correlation coefficient r is calculated as the squared correlation coefficient, r^2, or, more commonly, R^2. Since the variability in human behavior and affect are not easily explained, meaningful R^2s in behavioral sciences can be quite small ($R^2 < 0.10$) and an $R^2 > 0.20$ may be considered substantial. Using the data from the *hypothetical* example in Table 14, 42% of the variance in functional impairment is explained by depression (i.e., $R^2 = r^2 = 0.648^2 = 0.420$).

3.12.4.4.2 Test for a difference between two independent correlation coefficients

The correlation coefficients of two groups can be compared by testing the following null hypothesis:

$$H_0\text{: } \rho_1 = \rho_2$$

The test is conducted in the following manner. First, each correlation coefficient (r_k) is transformed into a Z_k using a procedure called the r_k to Z_k transformation. Note that the subscript k represents group k. Unlike earlier notation, in this context the subscript is not used to differentiate among subjects.

$$Z_k = 0.5 \cdot \ln\left(\frac{1 + r_k}{1 - r_k}\right)$$

where $k = 1$ or 2 and ln represents the natural logarithm.

Next a standard error is calculated:

$$\rho_{z_1 - z_2} = \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}$$

Finally, the Z-statistic is calculated:

$$Z = \frac{Z_1 - Z_2}{\sigma_{z_1 - z_2}}$$

For example, this could be used to address the question, "Is the strength of the linear relation between depression and functional impairment different in the subjects in the CBT and IPT groups?" Although IPT impairment data are not presented in a table, for illustration assume that for the IPT group $r = 0.41$. Comparison of the correlation coefficients for the CBT and IPT groups proceeds as follows:

$$Z_1 = 0.5 \cdot \ln\left(\frac{1+\mathrm{r}_1}{1-\mathrm{r}_1}\right) = 0.5 \cdot \ln\left(\frac{1+.648}{1-.648}\right)$$
$$= .772$$
$$Z_2 = 0.5 \cdot \ln\left(\frac{1+\mathrm{r}_2}{1-\mathrm{r}_2}\right) = 0.5 \cdot \ln\left(\frac{1+.41}{1-.41}\right)$$
$$= .436$$
$$\sigma_{z_1-z_2} = \sqrt{\frac{1}{n_1-3}+\frac{1}{n_2-3}}$$
$$= \sqrt{\frac{1}{20-3}+\frac{1}{20-3}} = .343$$
$$Z = \frac{Z_1-Z_2}{\sigma_{z_1-z_2}} = \frac{.772-.436}{.343} = .980$$

The assumptions of the Z-test of correlation coefficients across two groups is that X and Y come from a bivariate normal distribution. What this means is that, if X comes from a normal distribution, and Y is normally distributed at every value of X.

The hypothesis testing steps for the Z-test to compare correlation coefficients across two groups are as follows:

(i) State research question. Is the strength of the linear relation between depression and functional impairment different for subjects in the CBT and IPT groups?

(ii) State two mutually exclusive and exhaustive hypotheses:

(a) $H_0: \rho_1 = \rho_2$

(b) $H_A: \rho_1 \neq \rho_2$

(iii) Choose the Z-test to evaluate the hypotheses.

(iv) The critical value for the Z-test with $\alpha = 0.05$ is 1.96. If $z_{\text{observed}} > z_{\text{critical}}$, reject H_0.

(v) Data are displayed in Table 14 for the CBT group. The correlation for the IPT group was specified in the text as $r = 0.41$.

(vi) Calculations were described above.

(vii) Decision regarding H_0. Since $z_{\text{observed}} < z_{\text{critical}}$ (i.e., $.980 < 1.96$), do not reject H_0.

(viii) Conclusion. There is not a statistically significant difference in the strength of the linear relationship between depression and impairment for those in the CBT and IPT groups in this study of treatment of depressive symptoms in patients with major depressive disorder.

There are several more points regarding the Pearson r that will be discussed briefly. First, the correlation of a *variable* with a *constant* is zero. One of the methods of calculating r that was described above includes the covariance (s_{xy}) in the numerator. Recall that the covariance is equal to the sum of cross-products over $N-1$. Also, recall that the sum of the cross-products: $\sum xy = \sum(X_i - \bar{X})(Y_i - \bar{Y})$. If Y_i is a constant, then for each subject $Y_i = \bar{Y}$ and thus the sum of cross-products is also equal to zero:

$$\sum xy = \sum(X_i - \bar{X})(Y_i - \bar{Y})$$
$$= \sum(X_i - \bar{X}) \cdot 0 = 0$$

If the covariance is equal to zero, $r = 0$.

There is another useful fact about how r is affected by a constant (c_i). If the correlation between depression (X_i) and impairment (Y_i) is r_{xy}, and a constant is added to either X_i or Y_i, the correlation is not affected. For instance, say that the 24-item HRSD ratings were to be collected for all acutely suicidal subjects in a study. However, in the interest of time the HRSD suicide item was not scored and not included in a modified HRSD 23-item total, call it Y_{M_i}. Technically, all of these subjects, being acutely suicidal would rate 2 on the HRSD suicide item. Thus a 24-item total, Y_i, could be calculated for each subject simply by adding 2 to the subject's 23-item total. The addition of a constant (c_i) to a variable (Y_i) increases the value of the y-mean by c_i (i.e., $\bar{Y}_M = \bar{Y} + c_i$) and it does not affect the standard deviation (i.e., $S_{Y_M} = S_Y$). As a consequence neither the covariance (s_{xy}) nor the sum of cross-products $\sum xy$ is changed, and thus the correlation coefficient is not affected: $r_{xy} = r_{xy_M}$.

Finally, correlation coefficients for pairs of several variables (x, y, and z) can be conveniently presented in the form of a correlation matrix:

$$\begin{array}{cccc} & X & Y & Z \\ X & 1.0 & r_{XY} & r_{XZ} \\ Y & r_{YX} & 1.0 & r_{YZ} \\ Z & r_{ZX} & r_{ZY} & 1.0 \end{array}$$

Notice that the elements of the correlation matrix are symmetric. That is, the correlation of X_i with Y_i (r_{XY}) is equal to the correlation of Y_i with X_i (r_{YX}). Furthermore, the correlation of a variable (e.g., X_i) with itself is equal to unity. All elements of the main diagonal of a correlation matrix are 1.0: $r_{XX} = r_{YY} = r_{ZZ} = 1.0$.

3.12.4.4.3 Spearman rank correlation

The Spearman rank correlation is a nonparametric correlation coefficient, which is computed using the ranks of the raw data. This is appropriate when the population data are believed to be non-normal. The Spearman rank correlation is especially useful when data contain an outlier (i.e., an extreme value),

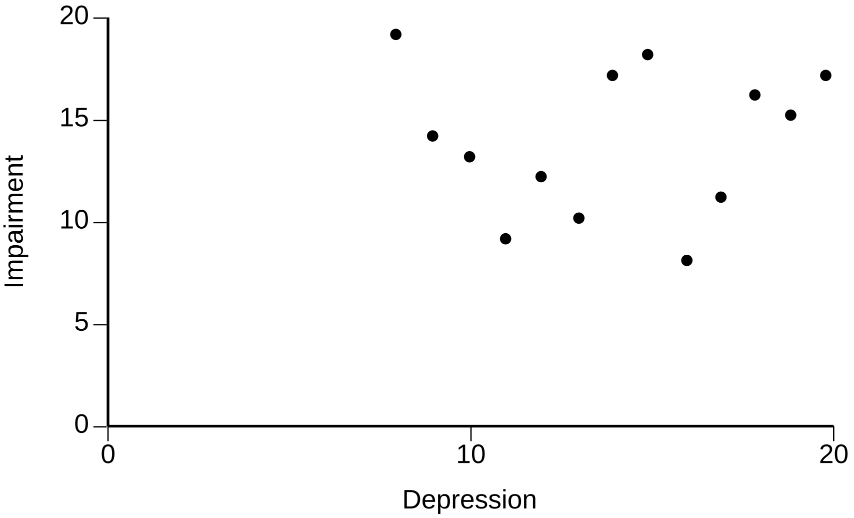

Figure 14 Plot of Hamilton Depression Ratings and Sheehan Disability Scores.

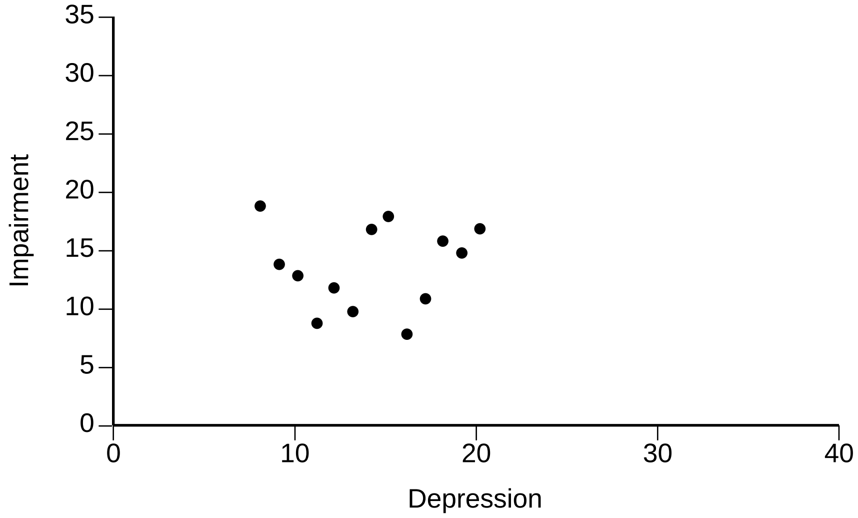

Figure 15 Plot of Hamilton Depression Ratings and Sheehan Disability Scores.

which can strongly, and inappropriately, increase the magnitude of the Pearson correlation coefficient. For instance, compare the plots presented in Figures 14 and 15. The only difference in these plots is that Figure 15 includes one additional case, which is an outlier. That is, the additional case has extreme values on both the x- and y-variables. In Figure 14 there is no linear relationship between depression (X) and impairment (Y). Stated differently, knowledge of the level of depression (X_i) will not help in "guessing" the level of impairment (Y_i). In fact, for the data in Figure 14, the Pearson correlation coefficient between depression and impairment is $r = 0.066$. In stark contrast, with the addition of one outlier, the Pearson correlation coefficient between depression and impairment is $r = 0.697$ for the data in Figure 15. In Figure 15, which includes one outlier, knowledge of the severity of a subject's depression would not help in guessing the corresponding level of impairment.

The discrepancy between these two Pearson correlation coefficients illustrates that the Pearson correlation coefficient can be quite sensitive to an outlier. The influence of outliers on the correlation coefficient can be minimized by calculating the correlation on the rank-transformed data to examine the linear relationship between the ranks of X_i and Y_i. The Spearman rank correlation is used for that purpose.

For example, the data displayed in Figure 15 are presented in Table 15. In addition to the actual depression (X_i) and impairment (Y_i) ratings, the ascending ranks of those ratings are presented in Table 15. The rank transformation procedure was described earlier in this chapter (see discussion of the Mann–Whitney test). Although this is a bivariate analysis, the rankings of the depression (R_{X_i}) and impairment (R_{Y_i}) ratings are performed independently. (The upper-case R is used to represent a ranked variable and is not to be confused with lower-case r, which is generally used to represent a Pearson correlation coefficient.) To calculate the Spearman rank correlation coefficient r_s, the difference (d_i) between the paired ranks for each subject is calculated ($d_i = R_{X_i} - R_{Y_i}$) and squared (d_i^2). The sum of the squared differences in ranks $\sum(d_i^2)$(is used in the calculations of r_s. In the case of no tied ranks, r_s is calculated:

$$r_s = 1 - \frac{6\sum_{i=1}^{N} d_i^2}{n^3 - n} = 1 - \frac{6 \cdot 345}{14^3 - 14} = 0.242$$

Notice that the value of d_i is minimized, and consequently the quantity, $6\sum_{i=1}^{N} d_i^2$ is minimized, as the pairs of ranks of X_i and Y_i for each subject correspond closely. Since the ratio in that equation is subtracted from unity, the value of r_s is closer to 1.0 when that ratio is small; and r_s is closer to zero when the ratio approaches 1.0. Thus, r_s is a quantity that reflects the strength of a linear relationship.

An r_s of .242 is certainly a more reasonable reflection of the strength of the linear relationship between X_i and Y_i in Table 15 than the Pearson r of .697.

The calculations for data involving tied data are more extensive (see, for example, Conover, 1980) and will not be discussed here. Incidentally, the data displayed in Figure 14 come from the first 13 of the 14 subjects that are presented in Table 15. The Pearson and Spearman rank correlation coefficients for the sample of 13 subjects are $r = 0.066$ and $r_s = 0.052$, respectively.

In using the Spearman rank correlation coefficient, it is assumed that the data are randomly sampled, that the subjects are independent, and that the scale of measurement is at least ordinal.

The hypothesis testing steps for the Spearman rank correlation coefficient are as follows:

(i) State research question. Is there a relationship between depression and impairment such that high values of depression tend to be associated with high values of impairment? Alternatively, one might hypothesize that there is a negative relationship between two variables X and Y (i.e., $r_s < 0$), such that high values of X tend to be associated with low values of Y.

(ii) State two mutually exclusive and exhaustive hypotheses:

(a) $H_0 : \rho_{S_{XY}} = 0$

(b) $H_A : \rho_{S_{XY}} \neq 0$

(iii) Use a table of critical values for the Spearman rank correlation coefficient (for example, see Conover, 1980, p. 456). The critical value with a two-tailed $\alpha = 0.05$ and $N = 14$ is 0.534. If $|r_{s_{\text{observed}}}| > r_{s_{\text{critical}}}$, reject H_0.

(iv) Data regarding depression and impairment are displayed in Table 15.

(v) Calculations were described above.

(vi) Decision regarding H_0. Since $|r_{s_{\text{observed}}}| < r_{s_{\text{critical}}}$ (i.e., $0.242 < 0.534$), do not reject H_0.

(viii) Conclusion. Based on data from this sample, we conclude that there is not a statistically significant correlation between depression and impairment.

3.12.4.4.4 Simple linear regression analysis

If a distinction between the independent and dependent variables is made, a correlational analysis can be extended to a *simple linear regression analysis*. In a regression analysis, the dependent variable (y) is regressed on the independent variable (x). Graphically, the x-axis represents the independent variable and the y-axis represents the dependent variable. As in a correlational analysis, each subject contributes data in the form of paired coordinates (x, y) to the regression analysis. The regression equation, $\hat{Y} = a + bX$, includes two variables, Y and X, and two constants, the y-intercept (a) and the slope (b). After the constants have been calculated, the equation can be used to estimate a conditional value of Y. That is, the equation is used to estimate a value of Y, *conditioned* on a value of X (i.e., $Y|X$). The algorithms which are used to estimate the regression constants, a and b, employ the *least-squares criterion*, which minimizes the sum of the squared distances from the observed data (Y_i) and those estimated with the regression equation ($\hat{Y}_i$). $\hat{Y}$ is called *predicted* Y or Y*-hat*. The estimated values of $\hat{Y}$ are points along the regression line, which is defined by the regression equation.

For example, consider the relation between depression (X) and functional impairment (Y), where the latter is the dependent variable. The data from Table 14 are plotted in Figure 16 where depression is measured using the HRSD and functional impairment is measured with the Sheehan Disability Scale. The slope (b) represents the expected change in functional impairment that is associated with a one-unit change in depression. Stated differently, consider two subjects with depression ratings of X_i and $X_i + 1$, respectively. If the former subject has an impairment rating of Y_i, the latter will have

Table 15 Depression (X_i) and impairment (Y_i) ratings of 14 patients with affective symptomatology.

X_i	$Rank(X_i)$	Y_i	$Rank(Y_i)$	d_i	d_i^2
8	1	19	13	−12	144
9	2	14	7	−5	25
10	3	13	6	−3	9
11	4	9	2	2	4
12	5	12	5	0	0
13	6	10	3	3	9
14	7	17	10	−3	9
15	8	18	12	−4	16
16	9	8	1	8	64
17	10	11	4	6	36
18	11	16	9	2	4
19	12	15	8	4	16
20	13	17	10	3	9
36	14	31	14	0	0
218	105	210	104	1	345

an expected impairment rating of $Y_i + b$. The intercept (a) represents the functional impairment that is expected with a *hypothetical* depression rating of 0. Of course, an HRSD of 0 is not realistic in a clinical trial for subjects meeting criteria for major depressive disorder, but it represents the point along the y-axis where an extrapolated regression line would cross.

The two regression constants are estimated as follows (using the summary data that are presented in Table 14):

$$b = \frac{\sum_{i=1}^{N} xy}{\sum_{i=1}^{N} x^2} = \frac{352}{640} = 0.55$$

$$a = \bar{Y} - b\bar{X} = 14.45 - 0.55 \cdot 21.0 = 2.90$$

Thus, the regression equation is:

$$\hat{Y} = a + bX = 2.90 + 0.55X$$

This indicates that a one-unit increase on the HRSD is associated with a 0.55 unit increase on the Sheehan Disability Scale. A regression line can be graphed by plotting two paired coordinates, $(0,a)$ and $(\bar{X}, \bar{Y})$, and drawing a line between the two points. Applying this to the data from Table 14, the regression line might be graphed by plotting: (0,2.9) and (21,14.45). However, it is inappropriate to extrapolate beyond the range of the sample data. This is because the relation between X and Y might look entirely different for values beyond the range of the data. In the example, the association between functional impairment and depression might be very different for those with minimal depressive symptomatology than for those meeting criteria for major depressive disorder.

In lieu of extending the regression line beyond the range of the sample data, consider the minimum and maximum HRSD (X_i) ratings. In this sample, the HRSD ratings range from 12 to 29. The corresponding predicted functional impairment (i.e., $\hat{Y}_i$) for the lowest and highest HRSD ratings in this sample are estimated by substituting the respective HRSD values in the regression equation:

$$Y = 2.90 + 0.55X = 2.90 + 0.55(12) = 9.50$$

$$Y = 2.90 + 0.55X = 2.90 + 0.55(29) = 18.85$$

Consequently, the regression line, which is presented in Figure 16, will go from (12,9.50) for the least severely ill to (29,18.85) for the most severely ill. Note that the regression line passes through the intersection of the X and Y means, $(\bar{X}, \bar{Y})$, which in this case is (21,14.45).

Recall that the proportion of variance in functional impairment that is explained by depression was presented above in the discussion of the correlation coefficient as: $R^2 = r^2 = 0.648^2 = 0.420$ These are the same data and thus the same value of R^2 applies. As with correlational analyses, R^2 is a measure of the magnitude of the association between Y_i and X_i in simple linear regression. One of the descriptive statistics that was discussed earlier in this chapter is a measure of variability called the sum of squared deviations, in this case, deviations from the Y-mean:

$$\sum y_i^2 = \sum_{i=1}^{N} \left(Y_i - \bar{Y}\right)^2$$

In the context of a regression analysis, this quantity can be partitioned into two components:

$$\sum y_i^2 = \sum \left(\hat{Y}_i - \bar{Y}\right) + \sum \left(Y_i - \hat{Y}_i\right)$$

or

$$\sum y_i^2 = SS_{reg} + SS_{res}$$

The former of the two components, the *sum of squares regression*, or SS_{reg}, represents the sum of squares in Y that is accounted for by X. If one were attempting to guess the level of impairment (Y_i), SS_{reg} can be thought of as the improvement over guessing the mean impairment ($\bar{Y}$) for each subject, and instead using the information known from the sample data about the $X_i - Y_i$ relation. SS_{reg} is estimated as follows (using the summary data from the total row in Table 14):

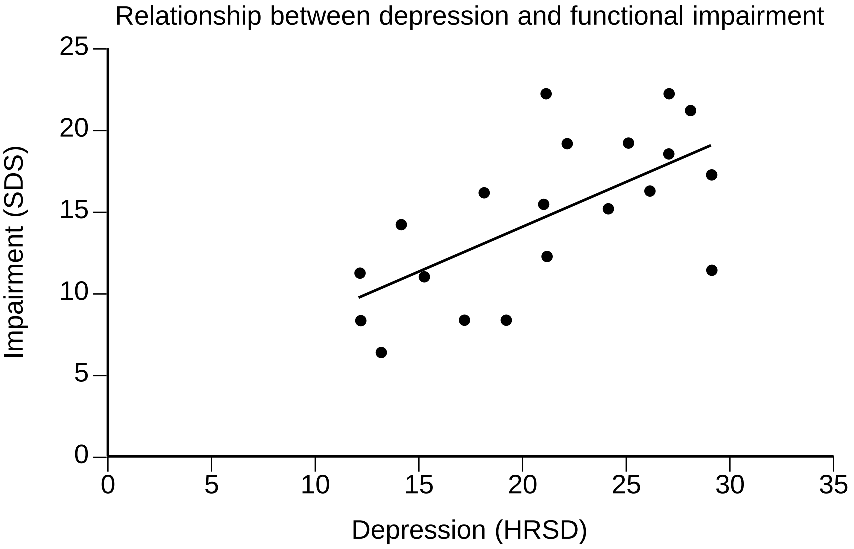

Figure 16 Relationship between depression and functional impairment.

$$SS_{reg} = \sum\left(\hat{Y} - \bar{Y}\right)^2 = \frac{\left(\sum xy\right)^2}{\sum x^2}$$
$$= \frac{352.00^2}{640.0} = 193.6$$

The sum of squares residual (SS_{res}) is the latter component of $\sum y^2$ and represents the sum of squared deviations of Y that is not accounted for by the $X_i - Y_i$ relation:

$$SS_{res} = \sum\left(Y_i - \hat{Y}\right)^2 = \sum y^2 - SS_{reg}$$
$$= 460.95 - 193.6 = 267.35$$

$R^2_{Y\cdot X}$ is the quantity that represents the magnitude of the association between Y and X. It represents the proportion of variance in Y that is accounted for by X:

$$R^2_{Y\cdot X} = \frac{SS_{reg}}{SS_{reg} + SS_{res}} = \frac{193.6}{193.6 + 267.35}$$
$$= 0.42$$

Being a proportion, $R^2_{Y\cdot X}$ can range from 0 to 1; $0 \leqslant R^2_{Y\cdot X} \leqslant 1$. $R^2_{Y\cdot X}$ can also be estimated as follows:

$$R^2_{Y\cdot X} = \frac{\left(\sum xy\right)^2}{\sum x^2} \cdot \frac{1}{\sum y^2} = \frac{352^2}{640} \cdot \frac{1}{460.95}$$
$$= 0.42$$

The first component is the numerator of the prior equation (SS_{reg}) and the second component is the denominator (SS_{res}) of that equation. Presented as one ratio, using the data from Table 14:

$$R^2_{Y\cdot X} = \frac{\left(\sum xy\right)^2}{\sum x^2 \sum y^2} = \frac{352^2}{640 \cdot 460.95} = 0.420$$

which is the square of the correlation between Y and X: $R^2_{Y\cdot X} = r^2_{YX} = 0.648^2 = 0.420$

Thus 42% of the variance in Y is accounted for by X. Although this seems to be a substantial proportion of variance, a statistical test is used to determine whether this is more than might be expected by chance. The null hypothesis, $H_0: \rho^2 = 0$, is tested with an F-test:

$$F = \frac{R^2/k}{(1 - R^2)/(N - k - 1)}$$

where k is equal to the number of independent variables, $k = 1$ in this case:

$$F = \frac{R^2/k}{(1 - R^2)/(N - k - 1)}$$
$$= \frac{.402/1}{(1 - .420)/(20 - 1 - 1)} = 13.034$$

The degrees of freedom for the F-test are k and $(N-k-1)$, or with these data, 1 and 18. The corresponding critical F, $F_{(1,18;\alpha = .05)} = 4.41$. The null hypothesis is rejected because 13.034 > 4.41. In other words, the amount of variance in functional impairment that is explained by depression is statistically significant.

The statistical significance of the regression coefficient b can also be tested using the null hypothesis as $H_0: B = 0$. Consider that if the null hypothesis is not rejected, the regression coefficient could not be considered to be different than zero. Then if $b = 0$, the X-term would fall out of the regression equation, and $\hat{Y}$

would not be a function of X. The statistical significance of the regression coefficient b is tested with a t-ratio, which is equal to the regression coefficient divided by its standard error:

$$t = \frac{b}{s_b}$$

where the standard error of the regression coefficient s_b is estimated as:

$$s_b = \frac{\sqrt{SS_{res}/N - k - 1}}{\sqrt{\sum_{i=1}^{N} x^2}}$$

Notice that since s_b is in the denominator of the t-ratio, a larger value of s_b reduces the value of t. Holding other aspects of the analysis constant, if SS_{res} is large, which can result from a small $X-Y$ relation, then s_b will be large, and t will be small. In contrast, with a larger sample size (N), s_b will be smaller and, as a consequence, t will be larger.

$$SS_{res} = \sum_{i=1}^{N} \left(Y_i - \hat{Y}\right) \sum y^2$$

This quantity incorporates the differences between the observed (Y_i) and predicted ($\hat{Y}_i$) values of Y. In lieu of calculating each of the predicted values, SS_{res} can be calculated using quantities that have been calculated above:

$$SS_{res} = \left(1 - R^2_{Y \cdot X}\right) \sum y^2$$
$$= (1 - .420) \cdot 460.95 = 267.351$$

The standard error of the regression coefficient is given as: s_b

$$s_b = \frac{\sqrt{SS_{res}/N - k - 1}}{\sqrt{\sum_{i=1}^{N} x^2}}$$
$$= \frac{\sqrt{267.351/(20 - 1 - 1)}}{\sqrt{640}} = .152$$

and the t-ratio as:

$$t = \frac{b}{s_b} = \frac{0.55}{0.152} = 3.61$$

In a simple linear regression analysis there is one independent variable. Thus the tests of significance of $R^2_{Y \cdot X}$ and b are synonymous. In fact, the square root of F is equal to t: $\sqrt{13.034} = 3.61$

The assumptions of simple linear regression analysis include:

(i) independence across observations (i.e., subjects);

(ii) the independent variable is fixed;

(iii) the error term (e) is not correlated with the independent variable (X);

(iv) the variance of the error terms is constant across all values of X. This is referred to as homoscedasticity.

The hypothesis testing steps for a simple linear regression analysis are as follows:

(i) State research question. Is there a linear relationship between depression and impairment?

(ii) State two mutually exclusive and exhaustive hypotheses:

(a) Hypotheses about the regression coefficient (β):

$H_0 : \rho_{XY} = 0$
$H_A : \rho_{XY} \neq 0$

(b) Hypotheses about the population R^2 (i.e., ρ^2_{XY}):

$H_0 : \rho^2_{XY} = 0$
$H_A : \rho^2_{XY} \neq 0$

(iii) Choose the t-ratio to evaluate the hypotheses about the regression coefficient (β). Use an F-test to test hypotheses about R^2.

(iv) The critical value for the t-ratio with two-tailed $\alpha = 0.05$ and df = 18 is 2.101. If $t_{\text{observed}} > t_{\text{critical}}$, reject $H_0 : \beta_{XY} = 0$. The critical value for the F-test with $\alpha = 0.05$ and df = 1,18 is 4.41. If $F_{\text{observed}} > F_{\text{critical}}$, reject $H_0 : \rho^2_{XY} = 0$.

(v) Data regarding depression and impairment are given in Table 14.

(vi) Calculations were described above.

(vii) Decision regarding H_0 about (β). Observed $t >$ critical t, thus reject H_0 (i.e., $3.61 > 2.101$). Decision regarding H_0 about R^2. Observed $F >$ critical F (i.e., $13.064 > 4.41$), thus reject H_0.

(vii) Conclusion. Based on data from this sample, we conclude that there is a statistically significant linear relationship between depression and impairment.

Linear regression is a useful statistical procedure for examining a bivariate relation. Unlike a correlational analysis in which $r_{yx} = r_{xy}$, linear regression is *not* a symmetric procedure. That is, the regression of Y_i on X_i will not result in the same regression constants (a and b) as the regression of Y_i on Y_i.

3.12.4.4.5 Multiple linear regression analysis

A *simple* linear regression analysis involves the regression of a dependent variable on one independent variable. *Multiple* linear regression analysis extends the statistical model such that one dependent variable is regressed on multiple independent variables. Multiple linear regression analysis is a general statistical model that can evaluate both dimensional and categorical independent variables. In fact it can test main effects and interactions of the ANOVA model and can be used to control for variables (i.e., covariates) if certain assumptions are fulfilled (see Pedhazur, 1982, Chap. 12).

Pedhazur (1982) distinguishes between the use of regression for prediction and explanation. The former has to do with estimating the dependent variable given a value of the independent variable (i.e., Y_i given X_i), whereas the latter focuses on the proportion of variance in the dependent variable that is explained by the independent variables (R^2). Although this chapter will not discuss multiple linear regression analysis in detail, several comprehensive examinations of multiple linear regression analysis are available (Aiken & West, 1991; Cohen & Cohen, 1983; Cook & Weisberg, 1982; Neter, Wasserman, & Kutner, 1989; Pedhazur, 1982; Pedhazur & Schmelkin, 1991).

3.12.4.5 Between-subject Designs: Categorical Dependent Variables

The statistical procedures that have been discussed until this point in the chapter are appropriate for continuous or ordinal dependent variables. Nevertheless, it is not unusual for the dependent variable to be categorical. There are a multitude of examples: pass or fail, sick or well, and live or die. There are many statistical procedures for categorical data and comprehensive texts devoted exclusively to categorical data analysis (see for example Agresti, 1996; Bishop, Feinberg, & Holland, 1975; Fleiss, 1981).

Consider the data from the hypothetical clinical trial that have been used throughout this chapter (originally presented in Tables 3 and 5). The HRSD is a continuous dependent variable that could be dichotomized such that the subjects with an HRSD rating of 14 or less are classified as "responders" and all other subjects are classified as "nonresponders." (For readers familiar with the HRSD, this is a rather crude and quite unconventional threshold, used only as an example for data analysis and not for the study of psychopathology.) Both the independent variable (treatment CBT vs. IPT) and the dependent variable (response: yes vs. no) are categorical variables. Although there is some literature discouraging categorization of continuous measures because of the loss of precision and statistical power, in the clinical setting decisions often must be dichotomous.

Table 16 Treatment response status of subjects in CBT and IPT groups

	Nonresponse	*Response*	*Total*
CBT	16 (a)	4 (b)	20
[row %]	[80%]	[20%]	[100%]
{column %}	{57.1%}	{33.3%}	{50%}
IPT	12 (c)	8 (d)	20
[row %]	[60%]	[40%]	[100%]
{column %}	{42.9%}	{66.7%}	{50%}
Total	28	12	40
[row %]	[70%]	[30%]	[100%]
{column %}	{100%}	{100%}	{100%}

3.12.4.5.1 *Chi-square test*

The chi-square (χ^2) test is generally used to compare rates or proportions across groups (Fleiss, 1981). In the example presented in Table 16, the χ^2 test is used to compare the response rates, and determine if rates as diverse as 20% and 40% can be expected by chance. This format of data display is referred to as a *contingency table*. Initially, notice that the IPT group had twice the response rate as the CBT group (40% vs. 20%). Also note the degree to which the rate for each group differs from that of the marginal (30%), which is presented in the margins of Table 16, and represents the response rate of the sample as a whole ($N = 40$).

The null hypothesis is that the population rates do not differ, $H_A: \pi_1 = \pi_2$. That hypothesis is tested by comparing the *observed* cell frequencies with those *expected* if treatment and response were independent. It is tested in the following manner:

$$\chi^2 = n_{..} \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{\left(\left|p_{ij} - p_i p_j\right| - 1/(2n_{..})\right)^2}{p_{i.} p_{.j}}$$

where p_{ij} is the proportion of the entire sample that is in row i and column j (i.e., $cell_{ij}$). Note that a dot (.) in the subscript represents a dimension (row or column) that is ignored. For example, the marginal value $p_{i.}$ represents the proportion of subjects in row i and $p_{.j}$ represents the proportion of subjects in column j. The product of those marginal proportions is the proportion of subjects that would be expected in the cell if the row and column variables were independent of each other. This product is referred to as the *expected value*. The χ^2 test essentially compares the *observed* and *expected* values. Applied to the data from Table 16:

$$\chi^2 = \frac{(|.10 - .50\cdot .70| - 1/(2\cdot 40))^2}{.50\cdot .70} + \frac{(|.40 - .50\cdot .30| - 1/(2\cdot 40))^2}{.50\cdot .30} + \frac{(|.3 - .50\cdot .70| - 1/(2\cdot 40))^2}{.50\cdot .70} + \frac{(|20 - .50\cdot .30| - 1/(2\cdot 40))^2}{.50\cdot .30} = 1.07$$

Another algorithm for a *fourfold table* (i.e., a two by two contingency table) is somewhat simpler, using cell and marginal frequencies, but no proportions:

$$\chi^2 = \frac{n..(|n_{11}n_{22} - n_{12}n_{21}| - \frac{1}{2}n_{..})^2}{n_{1.}n_{2.}n_{.1}n_{.2}}$$

With the data presented in Table 16:

$$\chi^2 = \frac{40(|16\cdot 8 - 4\cdot 12| - \frac{1}{2}40)^2}{20\cdot 20\cdot 28\cdot 12} = 1.07$$

Assumptions of the χ^2 test include:

(i) The samples are independent of each other.

(ii) The subjects within each group are independent of each other.

(iii) No subject can be classified in more than one category.

(iv) No cell has an expected value less than five.

The hypothesis testing steps for the χ^2 test are as follows:

(i) State research question. Do CBT and IPT have similar effects for the treatment of patients with depression?

(ii) State two mutually exclusive and exhaustive hypotheses:

(a) $H_0: \pi_1 = \pi_2$

(b) $H_A: \pi_1 \neq \pi_2$

(iii) If the assumptions are met, use the χ^2 test to evaluate the hypotheses.

(iv) The critical value for the χ^2 test with two-tailed $\alpha = 0.05$ and df = 1 is 3.84. (Note that df is equal to the product of # rows and # columns.) If $\chi^2_{\text{observed}} > \chi^2_{\text{critical}}$, reject H_0.

(v) Data are displayed in Table 16.

(vi) Calculations were described above.

(vii) Decision regarding H_0: $\chi^2_{\text{observed}} < \chi^2_{\text{critical}}$ (i.e., $1.07 < 3.84$), thus do not reject H_0.

(viii) Conclusion. In this sample of patients with major depressive disorder, those who were treated with IPT did not have different response rates than those who were treated with CBT.

The null hypothesis is not rejected because the difference in response rates across treatment groups (20% vs. 40%) is not more than is expected by chance. The chi square test can also be used for two-way contingency tables that have more than two rows and two columns.

3.12.4.5.2 Fisher's Exact test

When at least one expected frequency in a fourfold table is less than five, the Fisher's Exact test is more appropriate than the χ^2 test. The test assesses the probability of the cell frequencies, given the marginal frequencies. Stated differently, given the respective sample sizes and the response rates of the sample as a whole, the Fisher's Exact test evaluates the probability of the four cell frequencies. The algorithm, which is based on the hypergeometric probability distribution, is:

$$p = \frac{n_{1.}!n_{2.}!n_{.1}!n_{.2}!}{n_{..}!n_{11}!n_{12}!n_{21}!n_{22}!}$$

Where "*n*-factorial" ($n!$),

$$n! = n\cdot(n-1)\cdot(n-2)\cdot\ldots\cdot 1$$

and, by definition, $0! = 1$.

The algorithm for p calculates the probability of getting precisely the observed configuration of frequencies. There were no expected frequencies that were less than five in Table 16. For that reason, the data in Panel A of Table 17 will be used to illustrate the calculations involved in a Fisher's Exact test. These data, which will be referred to as the *observed* data, come from a *hypothetical pilot* study comparing electroconvulsive therapy (ECT) and psychodynamic psychotherapy (PP) for the treatment of dysthymia. The response rates were 93.33% for ECT and 85.0% for PP.

Applying the above algorithm:

$$p = \frac{15!\cdot 20!\cdot 4!\cdot 31!}{35!\cdot 1!\cdot 14!\cdot 3!\cdot 17} = 0.3266$$

That is, given the observed marginals, about one of every three *random* configurations will be that which is contained in Table 17. Consider that there are only four other configurations of cell frequencies that conform to the marginals of this contingency table, which are also presented in Table 17, Panels B–E. Fisher's Exact test involves calculating the probability of each configuration that conforms to the marginals. Applying the equation that was presented above, the respective probabilities of those tables are presented in the final column of Table 17. (Note that the sum of these four probabilities, and that of the *observed* table, 0.3266, is 1.00. That is, the probability of *any* of the possible outcomes is 1.00.)

For the Fisher's Exact test, one determines which of the configurations have a probability that is less than or equal to the probability

Table 17 Treatment by responder status.

	Configuration	*Treatment*	*Nonresponder*	*Responder*	*Total*	*p*
Panel A	#1	ECT	1	14	15	0.3266
		PP	3	17	20	
		Total	4	31	35	
Panel B	#2	ECT	0	15	15	0.0925
		PP	4	16	20	
		Total	4	31	35	
Panel C	#3	ECT	2	13	15	0.3810
		PP	2	18	20	
		Total	4	31	35	
Panel D	#4	ECT	3	12	15	0.1738
		PP	1	19	20	
		Total	4	31	35	
Panel E	#5	ECT	4	11	15	0.0261
		PP	0	20	20	
		Total	4	31	35	

associated with the *observed* contingency table, configuration #1. In this case, the probabilities of configurations #2 ($p = 0.0925$), #4 ($p = 0.1738$), and #5 ($p = 0.0261$) are all less than the observed configuration #1 ($p = 0.3266$). These probabilities, and that of the observed, are summed and compared to the alpha level:

$$p^* = 0.0925 + 0.1738 + 0.0261 + 0.3266 = 0.6190$$

Assumptions of the Fisher's Exact test are as follows:

(i) The data are in a contingency table with two rows and two columns.

(ii) The samples are independent of each other.

(iii) The subjects within each group are independent of each other.

(iv) No subject can be classified in more than one category.

(v) The row and column totals are given. They are not random.

The hypothesis testing steps for the Fisher's Exact test are as follows:

(i) State research question. Do ECT and PP have similar effects on *response*?

(ii) State two mutually exclusive and exhaustive hypotheses:

(a) $H_0: \pi_1 = \pi_2$

(b) $H_A: \pi_1 \neq \pi_2$

(iii) If assumptions are fulfilled, use Fisher's Exact test to evaluate the hypotheses.

(iv) The critical value for the Fisher's Exact test with two-tailed $\alpha = 0.05$ is 0.05. If $p^* \leqslant 0.05$, reject H_0.

(v) Data are displayed in Table 17.

(vi) Calculations were described above.

(vii) Decision regarding H_0: $p^*_{\text{observed}} > p_{\text{critical}}$ (i.e., $0.619 > 0.05$), thus do not reject H_0.

(viii) Conclusion. Dysthymic subjects treated with ECT and PP do not have differential response rates.

There is a growing body of literature on the use of *exact tests*, such as that just described, in other contexts (e.g., Agresti, 1992; Mehta & Patel, 1997). These procedures are based on minimal distributional assumptions. As a consequence they can be applied in situations when the use of other procedures is more tenuous, particularly with small data sets. Because exact procedures are often computationally extensive, their use was rare until software that implemented the exact algorithms were available. StatXact software (Mehta & Patel, 1997) is designed to conduct exact tests for a wide variety of types of data. It is available as a stand-alone package and also as an add-on module with statistical software such as SPSS and SAS.

3.12.4.5.3 Logistic regression

Logistic regression analysis is used to examine the association of (categorical or continuous) independent variable(s) with one *dichotomous* dependent variable. This is in contrast to *linear* regression analysis in which the dependent variable is a continuous variable. The discussion of logistic regression in this chapter is brief. Hosmer and Lemeshow (1989) provide a comprehensive introduction to logistic regression analysis.

Consider an example in which logistic regression could be used to examine the research question, “Is a history of suicide attempts associated with the risk of a subsequent (i.e., prospectively observed) attempt?” The logistic regression model compares the odds of a prospective attempt in those with and without prior attempts. The ratio of those odds is called the *odds ratio*. A logistic regression does not analyze the odds, but a natural logarithmic transformation of the odds, the log odds. Although the calculations are more complicated when there are multiple independent variables, computer programs can be used to perform the analyses. However, because of the logarithmic transformation of the odds ratio, the interpretation of results from the computer output is not necessarily straightforward. Interpretation requires a transformation back to the original scale by taking the inverse of the natural log of the regression coefficient, which is called *exponentiation*. The exponentiated regression coefficient represents the strength of the association of the independent variable with the outcome. More specifically, it represents the increase (or decrease) in risk of the outcome that is associated with the independent variable. The exponentiated regression coefficient represents the difference in risk of the outcome (e.g., suicide attempt) for two subjects who differ by one point on the independent variable. In this case, that is the difference between those with and without history of attempts (i.e., when history of attempts is coded: 0 = no and 1 = yes).

The logistic regression model can be extended to include several independent variables (i.e., hypothesized risk factors). For instance, are history of attempts, severity of depression, and employment status risk factors for suicidal behavior, controlling for diagnosis, age, and gender? Each odds ratio from such a model represents the change in risk of the outcome (i.e., a suicide attempt) that is associated with the independent variable, controlling for the other independent variables.

3.12.4.5.4 Survival analysis

In a study such as the example used above, the probability of the event is somewhat constrained by the follow-up time of the study. That is, the number of subjects with a suicide attempt will very likely be much smaller in a six-week study than in a six-year study. Survival analysis is a statistical technique that is used to examine the risk of an event *over time*. As with logistic regression analysis, this technique is appropriate for a dichotomous dependent variable. For example, survival analyses can be used to examine the risk of a suicide attempt over the course of a follow-up study. The survival model incorporates the differential risk period of the subjects. For that reason, it makes use of the available data in a follow-up study.

For example, in a 15-year follow-up study of subjects with affective disorders, some of the subjects will be followed for the entire study. Yet others will be followed for as short as a few weeks and others for several years, but not the duration of the entire 15-year study. The Kaplan–Meier product limit estimate (Kaplan & Meier, 1958) is used to estimate the proportion of subjects who have not had a postbaseline suicide attempt (i.e., the proportion *surviving* without a suicide attempt). Patients who do not attempt suicide by the end of follow-up, or who drop out of the follow-up study, are classified as *censored*. The Kaplan-Meier (1958) estimate is calculated as the probability of the event up to a given point in the follow-up. In this example, say that there are monthly assessments in which the investigators ask about psychopathology, suicide attempts, and treatment during the preceding month. The researcher could estimate the proportion of subjects who remained *suicide-free* after one month, two months, and so on to the end of follow-up.

There are several survival analytic procedures. For instance, the cumulative probability of an event (i.e., a suicide attempt) can be compared across groups using a logrank test (Peto & Peto, 1972). The survival model can be further extended to incorporate many independent variables that are hypothesized to be associated with risk of the event using a Cox proportional hazards regression model (Cox, 1972). There are several comprehensive texts on survival analytic techniques (e.g., Collett, 1994; Kalbfleisch & Prentice, 1980; Lawless, 1982; Lee, 1980; Parmar & Machin, 1995).

3.12.4.6 Within-subject Designs: Dimensional and Ordinal Dependent Variables

Subjects can serve as their own controls by being successively exposed to each treatment condition. The major advantage is that a design such as this reduces the between-subject variability that is inherent in a between-subjects design. As a consequence, a smaller number of subjects is needed in such a study. However, this design is inappropriate for many types of experimental intervention. For instance, the approach cannot be used in most studies of learning, unless it is feasible to get subjects to *unlearn* a task that has been mastered. The same holds true for acute treatment of psychopathology. Once a subject is well, the efficacy of subsequent treatments cannot be evaluated.

3.12.4.6.1 *Paired* t*-test*

The paired *t*-test is used for a within-subject comparison of dimensional dependent variables. For example, in a one-group design, the pretreatment HRSD scores could be compared with the post-treatment HRSD scores. (Note that there are many reasons that a multigroup design is preferable, but that is beyond the scope of this presentation. See Campbell & Stanley, 1963, for a comprehensive discussion of this.) The null hypothesis is $H_0: \mu_{\text{pre}} = \mu_{\text{post}}$, whereas the alternative hypothesis is $H_A: \mu_{\text{pre}} \neq \mu_{\text{post}}$. Stated differently, the data that are examined are the paired differences ($d_i = X_i - Y_i$) between preatment (X_i) and post-treatment (Y_i) for each subject. A mean difference ($\bar{d}$) and standard deviation of *differences* (s_d) are calculated.

For illustration, Table 18 presents HRSD scores for a group of 20 subjects before and after CBT treatment of depression. (At this point, focus only on the first three columns of Table 18.) Prior to treatment, the mean HRSD rating was 21.0 (sd = 5.8). After treatment, the mean HRSD rating was reduced to 15.95 (sd = 5.8). The paired *t*-test, which can generally be thought of as a test of whether the mean *difference* is equal to zero, is calculated as follows:

$$t = \bar{d}/s_{\bar{d}} \quad \text{where}$$

$$s_{\bar{d}} = \sqrt{\frac{s_d^2}{n}}$$

and s_d^2 is the variance of the *differences* (i.e., the square of the standard deviation of the *differences*):

$$s_d^2 = 5.46^2 = 29.84$$

$$s_{\bar{d}} = \sqrt{\frac{s_d^2}{n}} = \sqrt{\frac{29.84}{20}} = 1.22$$

$$t = \frac{\bar{d}}{s_{\bar{d}}} = \frac{5.10}{1.22} = 4.13$$

In a paired *t*-test, it is assumed that

(i) the data consist of correlated pairs of observations (i.e., paired data);

(ii) the subjects are randomly sampled;

(iii) the differences between each pair of observations are from a normally distributed population of differences.

The hypothesis testing steps for the paired *t*-test are as follows:

(i) State research question. Is there a difference in depression severity from pre- to post-CBT treatment?

(ii) State two mutually exclusive and exhaustive hypotheses:

(a) $H_0: \mu_d = 0$

(b) $H_A: \mu_d \neq 0$

(iii) If assumptions are met, use the paired *t*-test.

(vi) The critical value for the paired *t*-test with a two-tailed $\alpha = 0.05$ and $df = 19$, $t_{\text{critical}} = 2.093$ (Note that for a paired *t*-test, *df* is equal to one less than the number of pairs.) If $t_{\text{observed}} > t_{\text{critical}}$, reject H_0.

(v) Data are displayed in Table 18.

(vi) Calculations were described above.

(vii) Decision regarding H_0: $t_{\text{observed}} > t_{\text{critical}}$ (i.e., $4.13 > 2.093$), thus reject H_0.

(viii) Conclusion. There is a statistically significant reduction in severity of depression, as measured by the HRSD, from pre- to post-CBT treatment.

3.12.4.6.2 *Wilcoxon signed-rank test*

The Wilcoxon signed-rank test is used to test differences of paired data without the normal distribution assumption of the differences that is required for the paired *t*-test. The procedure is as follows. First, the difference between each pair of data is calculated. The differences are then ranked. Next, the ranks are assigned the sign ($\pm$) of the corresponding difference. Then two sums are computed: the sum of the positive signed ranks (T_+) and the sum of the absolute values of the negative signed ranks (T_-). (The differences of zero are ignored in these calculations.) Note that the sum of the totals should be: $T_+ + T_- = n(n+1)/2$, where *n* is the number of nonzero differences. If H_0 were true, one would expect about half of the differences to be positive and about half of the differences to be negative. If so, the sum of positive ranks would be approximately equal to the sum of the negative ranks: $T_+ \approx T_- \approx 1/2 \cdot n(n+1)/2$. The test statistic ($W_{\text{obs}}$) is used to examine this. W_{obs} is equal to the smaller of the two signed rank totals.

For example, see Table 18 which displays the pre- and post-HRSD scores and the difference scores. The median pretreatment rating (21.0) and the median post-treatment rating (15.0) appear to be quite different. The Wilcoxon Rank Sum test examines the magnitude of the difference. Separate columns display the ranks of the positive differences and the ranks of the absolute values of the

Table 18 HRSD scores: pre-(X) and post-(Y) CBT treatment ($N = 20$).

	X	*Y*	*Difference*	*Ranked |Difference|*	*Signed rank*	*Positive ranks*	*Negative ranks*
	12	11	1	5	5	5	
	12	4	8	15.5	15.5	15.5	
	13	15	−2	3	−3		3
	14	8	6	11	11	11	
	15	9	6	11	11	11	
	17	8	9	17	17	17	
	18	11	7	14	14	14	
	19	13	6	11	11	11	
	21	5	16	20	20	20	
	21	22	−1	4	−4		4
	21	28	−7	1	−1		1
	22	25	−3	2	−2		2
	24	21	3	6	6	6	
	25	15	10	18	18	18	
	26	22	4	7	7	7	
	27	22	5	8	8	8	
	27	14	13	19	19	19	
	28	22	6	11	11	11	
	29	23	6	11	11	11	
	29	21	8	15.5	15.5	15.5	
Total	420.00	319.00	101.00	210.00	190.00	200.00	10.00
Mean	21.00	15.95	5.05				
sd	5.80	7.16	5.46				

$d = X - Y$.

negative differences. The sums of those columns are:

$$T_+ = 200$$

$$T_+ = 10$$

$$W_{\text{observed}} = T_- = 10$$

Check the calculations the sum of the positive signed ranks and absolute values of the negative signed ranks:

$$T_+ + T_- = \frac{n(n+1)}{2}$$

$$T_+ + T_- = 200 + 10 = 210$$

$$\text{and} \quad \frac{\text{n(n+1)}}{2} = \frac{20(20+1)}{2} = 210$$

The assumptions of the Wilcoxon Signed Rank Sum test include:

(i) the data consist of correlated observations (i.e., paired data);

(ii) subjects with paired data are randomly and independently sampled;

(iii) The differences are measured on at least an ordinal scale.

The hypothesis testing steps for the Wilcoxon Signed Rank Sum test are as follows:

(i) State research question. Is there a difference in depression severity from pre- to post-CBT treatment?

(ii) State two mutually exclusive and exhaustive hypotheses with regard to location (e.g., median):

(a) $H_0 : \text{Median}_{\text{Time1}} = \text{Median}_{\text{Time2}}$

(b) $H_A : \text{Median}_{\text{Time1}} \neq \text{Median}_{\text{Time2}}$

(iii) If the assumptions are fulfilled, use the Wilcoxon Signed Rank Sum test to evaluate the hypotheses.

(iv) The critical value for the Wilcoxon Signed Rank Sum test, with two-tailed $\alpha = 0.05$ and $N = 20$, is $W_{\text{observed}} < W_{\text{critical}}$, reject H_0.

(v) Data are displayed in Table 18.

(vi) Calculations were described above.

(vii) Decision regarding $H_0 : W_{\text{observed}} < W_{\text{critical}}$ (i.e., $22.5 < 52$), thus reject H_0.

(viii) Conclusion. There is a significant reduction in depressive severity from pre- to post-CBT treatment.

3.12.4.6.3 Repeated measures ANOVA

The repeated measures ANOVA is a statistical procedure that can include both *within*-subject and *between*-subject factors. Initially the former will be considered. Most simply, it can be an extension of the paired t-test, used to compare subjects over time. Consider three HRSD assessments during the course of 12 weeks of CBT (e.g., weeks 0, 6, and 12). The null hypothesis is that there is no change in depression over the three assessments: $H_0 : \mu_1 = \mu_2 = \mu_3$. The alternative hypothesis is that the null hypothesis is false. The *F*-test of the repeated measures ANOVA tests whether that difference is more than can be expected by chance. If H_0 is rejected, the source of the difference must be examined more closely in a *post hoc* test. This is discussed in detail elsewhere (e.g., Keppel, 1991; Winer, 1971).

3.12.4.7 Within-subject Designs: Categorical Dependent Variables

The repeated measures design can be extended to include *between*-subject factors as well as *within*-subject factors. Such a design is referred to as a *split plot design*. There are two types of within- and between-subject factors: *fixed* and *random* effects. The specific levels of fixed variables are determined by the experimenter, whereas the levels of a random effect are randomly selected from a larger population of levels. Results of fixed effects cannot be generalized beyond the levels of the factors that are tested. In contrast, results of random effects can be generalized to the population from which the levels were selected. The within-subject factor that was discussed in the repeated measures model above, was a fixed effect. Random effects are discussed in detail elsewhere (e.g., Keppel, 1991; Winer, 1971). An extension of the previous example that includes two independent treatment groups (e.g., IPT vs. CBT) illustrates a design in which a repeated measures ANOVA with both within-subject and between-subject factors would be applied. Treatment would be the between-subjects factor, whereas time would be the within-subjects factor. Subjects are treated with only one modality, but have all levels of the within-subject factor (i.e., assessments at weeks 0, 6, and 12). This ANOVA includes three *F*-tests: the main effect of treatment, the main effect of time, and the interaction of treatment by time.

3.12.4.7.1 McNemar test

The McNemar test is used to examine paired dichotomous data. For example, one might compare the symptomatology pretreatment and post-treatment. Specifically, one might hypothesize that the sleep disturbance is neither developed nor overcome during the course of treatment with IPT for depression as presented in Table 19. The McNemar test is calculated as follows: $\chi^2 = (|b - c| - 1)^2/b + c$. (The algorithm, as presented with the term outside of the absolute value, incorporates a continuity correction in the numerator.) Note that the calculations focus on the *discordant* pairs and ignore the *concordant* pairs. More specifically, the test is a ratio of the squared difference in discordant frequencies relative to the total discordant frequencies. In the example above, the test would detect a disproportionate representation of an emergent sleep disturbance among those who had a change in sleep disturbance. This is illustrated with data in Table 19, where 11 of the 13 (84.5%) with changes in sleep disturbance developed the symptom during the course of treatment: $\chi^2 = (|11 - 2| - 1)^2/(11 + 2) = 4.92$. The McNemar χ^2 of 4.92 exceeds the critical χ_2 with 1 *df*, 3.84, and is thus statistically significant at the 0.05 level.

3.12.5 SUMMARY

There are several components of the analysis one must consider when choosing among statistical procedures. Is there a dependent variable? If so, is it dimensional, ordinal, or categorical? What are the comparisons that will answer the research question? Are there group comparisons? If so, how many groups are being compared? If not, is it bivariate linear relations that are of interest? In addition, the assumptions of the statistical procedures should be carefully considered before the procedure is applied. Figure 17 is a flow chart that is designed to facilitate selection of the appropriate inferential statistical procedure.

Finally, the value of descriptive statistical procedures must not be minimized. They are used not only to summarize data and describe a sample, but also to help a researcher become familiar with the data. The initial examination of descriptive statistics provides information that is useful when choosing among inferential statistical procedures.

In summary, this chapter began by stating that a fundamental motivation in science is to understand individual differences. A variety of

Table 19 Cross-classification of presence of symptoms before and after treatment.

	Post absent	*Post present*
Pre absent	76 (a)	11 (b)
Pre present	2 (c)	47 (d)

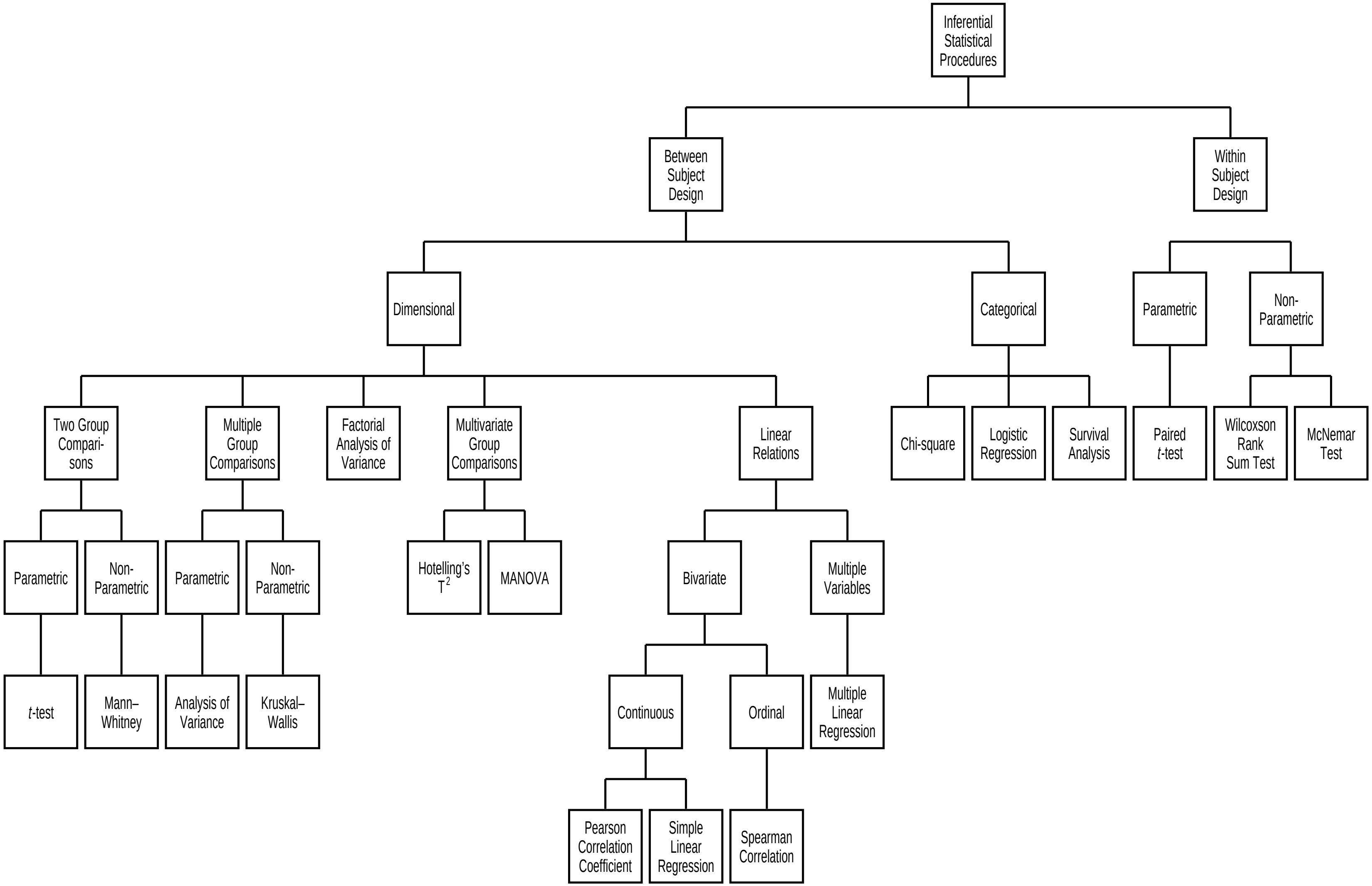
Inferential Statistical Procedures
Between Subject Design
Within Subject Design
Dimensional
Categorical
Parametric
Non-Parametric
Two Group Compari-sons
Multiple Group Comparisons
Factorial Analysis of Variance
Multivariate Group Comparisons
Linear Relations
Chi-square
Logistic Regression
Survival Analysis
Paired t-test
Wilcoxson Rank Sum Test
McNemar Test
Parametric
Non-Parametric
Parametric
Non-Parametric
Hotelling's T2
MANOVA
Bivariate
Multiple Variables
t-test
Mann–Whitney
Analysis of Variance
Kruskal–Wallis
Continuous
Ordinal
Multiple Linear Regression
Pearson Correlation Coefficient
Simple Linear Regression
Spearman Correlation

Figure 17 A guide for the selection of an appropriate inferential test.

statistical procedures that are used to examine *deviation* and *variability* have been outlined. The decision rules to choose among them has been discussed. Reference citations have been provided for a more comprehensive discussion of each technique. In closing, the reader is encouraged to consider not only the test statistic and corresponding *statistical significance* of research results, but also focus on the *clinical meaningfulness* of the findings. If a result cannot be replicated because it is trivial or if a finding has no practical value, it will serve no useful scientific or clinical purpose.

ACKNOWLEDGMENTS

The author is grateful to Laura Portera, M.S., who provided valuable feedback regarding the structure and content of earlier versions of this manuscript, and to Kira Lowell, M.A., who provided editorial assistance.

3.12.6 APPENDIX: SPSS COMMAND LINES

3.12.6.1 Descriptive Statistics

```
Descriptive variables= hrsd /statistics=all.
Frequency variables= hrsd
  /histogram=normal /statistics=all.
Examine variables= hrsd by tx /plot=boxplot
  stemleaf histogram /statistics=all.
```

3.12.6.2 Inferential Statistics

3.12.6.2.1 Between-subject analyses

```
Dimensional/Ordina Dependent Variables
  Two-group Comparisons
    t-test
      t-test group= tx (1,2) /variables=hrsd.
    Mann–Whitney
      Npar tests m–w=hrsd by tx (1,2)
  Multiple-group Comparisons
    Analysis of Variance (One-Way ANOVA)
      Oneway hrsd by tx (1,3)
      /posthoc=tukey, scheffe, dunnett
      /statistics=descriptives,homogeneity.
    Kruskal–Wallis
      Npar tests k–w=hrsd by tx (1,3)
  Factorial Analysis of Variance
    (Two-Way ANOVA)
      Anova variables= hrsd by tx (1,2)
        diagnosis (1,2)
      /maxorders=all.
  Bivariate Linear Relations
    Pearson Correlation Coefficient
      Correlations variables= depression with
        impairment.
    Spearman Rank
      Nonpar corr variables= depression with
        impairment
      /print=spearman.
    Simple Linear Regression
      Regression variables=impairment depression
      /descriptives=default
      /statistics=default
      /dependent=impairment
      /method=enter depression.
  Categorical Dependent Variables
    Chi-square
      Crosstab tx by response /cell=count row column /sta=chisq.
```

3.12.6.2.2 Within-group analyses

```
  Dimensional Ordinal Dependent Variables
    Paired t-test
      t-test pair=prehrsd with posthrsd.
    Wilcoxon
      Npar test wilcoxon= prehrsd with
          posthrsd
      /statistics=descriptives.
  Categorical Dependent Variables
    McNemar test
      Npar test mcnemar= presx with postsx.
```

3.12.7 REFERENCES

Agresti, A. (1992). A survey of exact inference for contingency tables. *Statistical Science, 7,* 131–177.

Agresti, A. (1996). *An introduction to categorical data analysis*. New York: Wiley.

Aiken L. S., & West S. G. (1991) *Multiple regression: testing and interpreting interactions*. Newbury Park, CA: Sage.

Armitage, P., & Berry, G. (1987). *Statistical methods in medical research* (2nd ed.). Oxford, UK: Blackwell Science.

Bishop, Y. M. M., Feinberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis: Theory and practice*. Cambridge, MA: MIT Press.

Bock, R. D. (1985). *Multivariate statistical procedures for behavioral research*. Chicago: Scientific Software Inc.

Campbell, D. T., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Collett, D. (1994). *Modelling survival data in medical research*. London: Chapman and Hall.

Conover, W. J. (1980). *Practical nonparametric statistics* (2nd. ed.). New York: Wiley.

Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression*. New York: Chapman and Hall.

Cox, D. R.(1972). Regression models and life tables (with discussion). *Journal of Royal Statistical Society, B34,* 187–220.

Fleiss, J. F. (1981). *Statistical methods for rates and proportions* (2nd. ed.). New York: Wiley.

Fleiss, J. F. (1986). *The design and analysis of clinical experiments*. New York: Wiley.

Grimm, L. G., & Yarnold, P. R. (Eds.) (1995). *Reading and understanding multivariate statistics*. Washington, DC: American Psychological Association.

Hartley, H. O. (1950). The maximum *F*-ratio as a short cut test for heterogeneity of variances. *Biometrika 37,* 308–312.

Hollingshead, A. B., & Redlich, F. C. (1958). *Social class and mental illness*. New York: Wiley.

Hopkins, K. D., Glass, G. V., & Hopkins, B. R. (1987). *Basic statistics for the behavioral sciences* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression*. New York: Wiley.

Kalbfleisch, J. D., & Prentice, R. L. (1980). *The statistical analysis of failure time data*. New York: Wiley.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association, 53,* 457–481.

Keppel, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Lawless, J. F. (1982). *Statistical models and methods for lifetime data*. New York: Wiley.

Lee, E. T. (1980). *Statistical methods for survival data methods*. Belmont, CA: Wadsworth Inc.

Levene, H. (1960). Robust tests for equality of variances. In I. Olkin (Ed.), *Contributions to probability and statistics*. Palo Alto, CA: Stanford University Press.

Lurie, P. M. (1995). A review of five statistical packages for Windows. *The American Statistician, 49,* 99–107.

Marcoulides, G. A., & Hershberger, S. L. (1997) *Multivariate statistical methods: A first course*. Hillsdale, NJ: Erlbaum.

Mehta, C., & Patel, N. (1997). *StatXact 3.1 for Windows: Statistical software for exact nonparametric inference: User manual*. Cambridge, MA: CYTEL Software Corporation.

Morgan, W. T. (1998). A review of eight statistics software packages for general use. *The American Statistician, 52,* 70–82.

Neter, J., Wasserman, W., & Kutner, M. H. (1989). *Applied linear regression models* (2nd ed.). Homewood, IL: Richard D. Irwin Inc.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Olson, C. L. (1976). On choosing a test statistic in multivariate analysis of variance. *Psychological Bulletin, 83,* 579–586.

Parmar, M. K. B., & Machin, D. (1995). *Survival analysis: A practical approach*. New York: Wiley.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research: Explanation and prediction*. New York: Holt, Rinehart and Winston.

Pedhazur, E. J., Schmelkin, L. P. (1991). *Measurement design and analysis: An integrated approach*. Hillsdale, NJ: Erlbaum.

Peto, R., & Peto, J. (1972). Asymptotically efficient rank invariant procedures (with discussion). *Journal of the Royal Statistical Society, Series A, 135,* 185–206

Sokal, R. R., & Rohlf, F. J. (1995) *Biometry* (3rd. ed.). New York: W. H. Freeman.

Sprent, P. (1989). *Applied nonparametric statistical methods*. London: Chapman and Hall.

SPSS (1997) *SPSS Base 7.5 Syntax reference guide*. Chicago: SPSS Inc

Stevens, J. (1992). *Applied multivariate statistics for the social sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Wilcox, R. R. (1996). *Statistics for the social sciences*. San Diego, CA: Academic Press.

Winer, B. J. (1971) *Statistical principles in experimental design* (2nd. ed.). New York: McGraw-Hill.

Zar, J. H. (1996). *Biostatistical analysis* (3rd. ed.). Upper Saddle River, NJ: Prentice-Hall.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

3.13

Latent Variables, Factor Analysis, and Causal Modeling

BRIAN S. EVERITT
Institute of Psychiatry, London, UK

3.13.1 INTRODUCTION 287
3.13.2 LATENT VARIABLES, OR HOW LONG IS A PIECE OF STRING? 288
3.13.3 FACTOR ANALYSIS 288
  3.13.3.1 Exploratory Factor Analysis 290
    3.13.3.1.1 Statements about pain 290
    3.13.3.1.2 Determining the number of factors 293
    3.13.3.1.3 Factor rotation 293
  3.13.3.2 Confirmatory Factor Analysis 295
    3.13.3.2.1 Confirmatory factor analysis models for length and intelligence 295
    3.13.3.2.2 Ability and aspiration 296
    3.13.3.2.3 Drug usage among American students 297
3.13.4 CASUAL MODELS AND STRUCTURAL EQUATION MODELING 300
  3.13.4.1 Examples of Causal Models 304
    3.13.4.1.1 Ability scores over time 304
    3.13.4.1.2 Stability of alienation 305
3.13.5 LATENT VARIABLES—MYTHS AND REALITIES 307
3.13.6 SUMMARY 308
3.13.7 APPENDIX 311
3.13.8 REFERENCES 311

3.13.1 INTRODUCTION

Many psychological investigations involve the collection of multivariate data, a term used when more than a single variable value is observed on each subject in a study. Table 1, for example, shows a set of data giving the examination scores obtained in four subjects by 15 students. For many such data sets the question of most interest is how the variables are related, and the first step in answering this question usually involves the calculation of the covariance or more commonly the correlation between each pair of variables (Table 2).

The techniques discussed in this chapter are aimed primarily at explaining and describing as concisely as possible the relationships between the variables in a set of multivariate data. Some of the methods are most applicable when the investigator has only limited expectation of what patterns to expect, and wishes to explore the covariance or correlation matrix in an effort to uncover hopefully some psychologically meaningful structure. Other methods are designed to test specific

Table 1 Examination scores for four subjects.

Student	*English*	*French*	*Algebra*	*Statistics*
1	77	82	67	67
2	63	78	80	70
3	75	73	71	66
4	55	72	63	70
5	63	63	65	70
6	53	61	72	64
7	51	67	65	65
8	59	70	68	62
9	62	60	58	62
10	64	72	60	62
11	52	64	60	63
12	55	67	59	62
13	50	50	64	55
14	65	63	58	56
15	31	55	60	57

theories or hypotheses as to why the variables have their observed relationships. Both approaches rely heavily on the concept of a latent variable, a topic discussed in the next section.

3.13.2 LATENT VARIABLES, OR, HOW LONG IS A PIECE OF STRING?

Table 3 shows the lengths of 15 pieces of string as measured by a ruler (R) and as estimated (guessed) by three different people (G, B, and D). Corresponding to each piece of string is a true but unknown length. The four observed measurements all give the true length plus some amount of measurement error, although R is likely to be considerably more accurate than the others. The four measurements are said to be fallible indicators of length.

Fascinating though the data in Table 3 may be (at least to the author), many readers of this article may quite reasonably ask what they have to do with psychology in general, and factor analysis in particular? To answer such questions consider again the data shown in Table 1, consisting of the examination scores of 15 individuals in four subjects. Analogous to the true string length of each of the 15 observations in Table 3, each individual in Table 1 might be assumed to have a particular, but unknown, level of cognitive ability (which some bolder psychologists might label "intelligence"), of which the four observed examinations scores are again fallible indicators.

And so in the data sets in both Tables 1 and 3, it can be assumed that the observed or manifest variables are fallible indicators of an underlying variable that cannot be measured directly—a so-called latent variable. Length is, of course, a far more straightforward latent variable than intelligence and one which most people would find little difficulty in accepting; its direct measurement is prevented simply by measurement error. Many of the latent variables postulated in psychology and related disciplines, however, are often more problematical and even in some cases controversial. Certainly there has been considerable debate about the nature of intelligence. The reason that such variables cannot be measured directly is often more complex than simple measurement error.

The question of how seriously latent variables should be taken in general is discussed in Section 3.13.5 but, as will be seen in later examples, invoking the concept of a latent variable is frequently extremely useful when exploring the relationships between manifest variables. In an exploratory analysis the aim will be to assess whether there is any evidence that the structure in an observed correlation matrix is due to a number of underlying latent variables and, if so, to describe and possibly label these variables.

When testing a specific hypothesis or model for the data, the relationships between the observed and latent variables will be specified *a priori,* and the investigator will be primarily interested in assessing the fit of the suggested model for the observed correlations or covariances. In addition to specifying how observed and latent variables are related, more complex models might also specify relationships between the latent variables themselves; examples are given later.

3.13.3 FACTOR ANALYSIS

Factor analysis is a generic term given to a class of multivariate statistical methods whose primary purpose is data reduction and summarization. In general terms factor analysis

Table 2 Covariances and correlations.

(i) A set of multivariate data is generally represented by a matrix **X** where

$$\mathbf{X}=\begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$$

where n is the number of observations in the data, p is the number of variables, and x_{ij} represents the value of variable j for observation i

(ii) The sample covariance s_{ij} of variables i and j is given by

$$s_{ij} = \frac{1}{n-1} \sum_{k=1}^{n} (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)$$

where $\bar{x}_i$ and $\bar{x}_j$ are the sample means of the two variables

(iii) A covariance can take any value from $-\infty$ to ∞. A value of zero indicates that there is no association (strictly, no linear association) between the variables. A positive value occurs when, as one of the variables increases in value then, on average, so does the other. A negative covariance indicates that as one variable increases, the other decreases, and vice versa

(iv) The sample covariance matrix is a matrix with the variances of each variable on the main diagonal and the covariances of each pair of variables as off-diagonal elements

(v) The sample correlation coefficient, r_{ij} between variables i and j is given by

$$r_{ij} = \frac{s_{ij}}{(s_{ii}s_{jj})^{\frac{1}{2}}}$$

where s_{ii} and s_{jj} are used to denote the variances of the two variables

(vi) A correlation takes values between -1 and 1, with the two extremes indicating that the two variables are perfectly linearly related

(vii) The sample correlation matrix is

$$\mathbf{R}=\begin{bmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ r_{p1} & r_{p2} & \dots & 1 \end{bmatrix}$$

techniques address themselves to the problem of analyzing the interrelationships amongst a possibly large set of observed variables and explaining these relationships in terms of a (hopefully) small number of underlying latent variables. By using factor analysis, the investigator should be able to identify the separate dimensions being measured by the manifest variables. When used in an exploratory mode, factor analysis methods place no constraints on the possible relationships between observed and latent variables. When used in a confirmatory manner, however, these relationships will be constrained in particular ways specified by the hypothesized model. The next section deals with exploratory factor analysis and Section 3.13.3.2 with confirmatory factor analysis.

Table 3 How long is a piece of string? Measurement by ruler (R) and guesses by three observers (G, B, and D) are given.

Piece	*R*	*G*	*B*	*D*
1	6.3	5.0	4.8	6.0
2	4.1	3.2	3.1	3.5
3	5.1	3.6	3.8	4.5
4	5.0	4.5	4.1	4.3
5	5.7	4.0	5.2	5.0
6	3.3	2.5	2.8	2.6
7	1.3	1.7	1.4	1.6
8	5.8	4.8	4.2	5.5
9	2.8	2.4	2.0	2.1
10	6.7	5.2	5.3	6.0
11	1.5	1.2	1.1	1.2
12	2.1	1.8	1.6	1.8
13	4.6	3.4	4.1	3.9
14	7.6	6.0	6.3	6.5
15	2.5	2.2	1.6	2.0

3.13.3.1 Exploratory Factor Analysis

The mathematical details of the factor analysis model are listed in Table 4, and the essential features of the technique are now demonstrated by considering an application.

3.13.3.1.1 Statements about pain

This illustration is based on a subset of the data reported in Skevington (1990). The study was concerned with beliefs about controlling pain and 123 individuals suffering from severe pain were presented with nine statements about pain. Each statement was scored on a scale from 1 to 6, ranging from disagreement to agreement. The nine statements and the observed correlations between them are shown in Table 5. One of the aims of the study was to ascertain whether the responses reflected the existence of subscales or groups of attitudes.

In an attempt to identify the latent variables that might account for the observed pattern of correlations between the pain statements, a particular form of factor analysis, maximum likelihood factor analysis was applied (described by Everitt & Dunn, 1991). The results from a factor analysis consist of the estimated regression coefficients of each observed variable on each latent variable (also known in this context as common factors). When the factor analysis has been carried out on the observed correlation matrix rather than the covariance matrix, the estimated regression coefficients are simply the correlations between each manifest variable and each latent variable. (In an exploratory factor analysis the choice of covariance or correlation matrix is not critical since there is a simple relationship between the solutions derived from each.) Table 6 shows these estimated correlations for both the two- and three-factor solutions. (Methods for deciding on the appropriate number of common factors are discussed later.)

How are the results given by a factor analysis interpreted? First, the estimated correlations (more commonly known as factor loadings) can be used to identify and perhaps name the underlying latent variables, although this is often more straightforward after the process of rotation, which is discussed in Section 3.13.3.1.3. But for now, examining the unrotated results in Table 6 it is seen that for both solutions the second factor is positively correlated, to a greater or lesser extent, with all nine statements. It would not require a great leap of imagination to suggest that this factor might be labeled "general pain level." The first factor is negatively correlated with statements taking personal responsibility for one's pain and positively correlated with statements in which the control of, and reasons for, pain are attributed elsewhere. Factors with a mixture of positive and negative loadings (often referred to as bipolar factors), usually become easier to understand after rotation and so further interpretation of the results is left until Section 3.13.3.1.3.

Apart from the factor loadings, a number of other quantities which need explanation are given in Table 6. First, the sum of squares of the factor loadings of a particular observed variable gives what is known as the communality of that variable, that is, the variance shared with the other manifest variables via their relationships with the common factors. Subtracting the communality of a variable

Table 4 The factor analysis model.

(i) In general terms factor analysis is concerned with whether the covariance, or correlations between a set of observed variables, $x_1, x_2, \dots, x_p$ can be explained in terms of a smaller number of unobservable latent variables (common factors), $f_1, f_2, \dots f_k$ where $k < p$ (hopefully k, the number of common factors, will be much less than the number of original variables p)

(ii) The factor analysis model is essentially a regression-type model in which the observed variables are regressed on the assumed common factors. In mathematical terms the factor analysis model can be written as

$$
\begin{aligned}
x_1 &= \lambda_{11}f_1 + \lambda_{12}f_2 + \dots + \lambda_{ik}f_k + e_1 \\
x_2 &= \lambda_{21}f_1 + \lambda_{22}f_2 + \dots + \lambda_{2k}f_k + e_2 \\
&\cdot \\
&\cdot \\
&\cdot \\
x_p &= \lambda_{p1}f_1 + \lambda_{p2}f_2 + \dots + \lambda_{pk}f_k + e_p
\end{aligned}
$$

The λs are factor loadings and the terms $e_1, e_2, \dots, e_p$ are known as specific variates—they represent that part of an observed variable not accounted for by the common factors

(iii) The following assumptions are made:

(a) The common factors are in standardized form and have variance one

(b) The common factors are uncorrelated

(c) The specific variates are uncorrelated

(d) The common factors are uncorrelated with the specific variates

(iv) With these assumptions the factor analysis model implies that the population variances (σ_i^2, $j = 1, 2, \dots, p$) and covariances (σ_{ij}) of the observed variables can be written as

$$\sigma_i^2 = \sum_{j=1}^{k} \lambda_{ij}^2 + \psi_i$$

$$\sigma_i^2 = \sum_{l=1}^{k} \lambda_{il}\lambda_{jl}$$

where ψ_i is the variance of specific variate e_i, i.e., the specific variance of variable x_i

(v) The model implies that the variance of an observed variable can be split into two parts $\Sigma_{j=1}^{k} \lambda_{ij}^2$ and ψ_i. The first of those is known as the communality of the variable x_i; it is the variance in the variable shared with the other observed variables via their relationships with the common factors

(vi) Note that the covariances of the observed variables are generated solely from their relationships with the common factors. The specific variates play no part in determining the covariances of the observed variables; they contribute only to the variances of those variables

(vii) There are a number of different methods for fitting the factor analysis model. The two most commonly used are principal factor analysis and maximum likelihood factor analysis—both are described in Everitt and Dunn (1991)

(viii) A method of factor analysis commonly used in practice is principal components analysis (Everitt & Dunn, 1991). Although this is an extremely useful technique for the summarization of multivariate data, it is not discussed in detail here because it is not a natural precursor to the confirmatory and causal models to be discussed later. It may, however, be worthwhile listing the main differences between the two approaches:

(a) Factor analysis (FA) and principal components analysis (PCA) each attempt to describe a set of multivariate data in a smaller number of dimensions than one starts with, but the procedures used to achieve this goal are essentially quite different in the two approaches

(b) FA, unlike PCA, begins with a hypothesis about the covariance (or correlational) structure of the variables, namely that there exists a set of k latent variables ($k < p$) and these are adequate to account for the interrelationships of the variables though not for their full variances

(c) PCA, however, is merely a transformation of the data and no assumptions are made about the form of the covariance matrix of the data. In particular PCA has no part corresponding to the specific variates of FA. Consequently, if the FA model holds and the specific variances are small, both forms of analysis would be expected to give similar results

(d) A clear advantage of FA over PCA is that there is a simple relationship between the solutions obtained from the covariance and correlation matrices

(e) It should be remembered that PCA and FA are both pointless if the observed variables are uncorrelated—FA because it has nothing to explain and PCA because it would lead to components which are essentially identical to the original variables

(ix) In many (perhaps most) examples the results from a principal components analysis and an exploratory factor analysis will be similar, with any differences not usually affecting the substantive interpretation

Table 5 Pain statements and their correlations.

Pain statements

(1) Whether or not I am in pain in the future depends on the skills of the doctors
(2) Whenever I am in pain, it is usually because of something I have done or not done
(3) Whether or not I am in pain depends on what the doctors do for me
(4) I cannot get any help for my pain unless I go to seek medical advice
(5) When I am in pain I know that it is because I have not been taking proper exercise or eating the right food
(6) People's pain results from their own carelessness
(7) I am directly responsible for my pain
(8) Relief from pain is chiefly controlled by the doctors
(9) People who are never in pain are just plain lucky

Correlations

	1	2	3	4	5	6	7	8	9
1	1.0000								
2	–0.0385	1.0000							
3	0.6066	–0.0693	1.0000						
4	0.4507	–0.1167	0.5916	1.000					
5	0.0320	0.4881	0.0317	–0.082	1.0000				
6	–0.2877	0.4271	–0.1336	–0.2073	0.4731	1.0000			
7	–0.2974	0.3045	–0.2404	–0.1850	0.4138	0.6346	1.0000		
8	0.4526	–0.3090	0.5886	0.6286	–0.1397	–0.1329	–0.2599	1.0000	
9	0.2952	–0.1704	0.3165	0.3680	–0.2367	0.1541	–0.2893	0.4047	1.000

Table 6 Maximum likelihood factor analysis solutions for pain statement correlations.

(i) Two-factor solution

Statement	*Factor 1*	*Factor 2*	*Communality*	*Specific variance*
1	0.643	0.211	0.458	0.542
2	–0.361	0.413	0.301	0.699
3	0.718	0.401	0.676	0.324
4	0.687	0.288	0.555	0.445
5	–0.311	0.569	0.420	0.580
6	–0.521	0.614	0.448	0.352
7	–0.563	0.477	0.544	0.456
8	0.709	0.254	0.567	0.433
9	0.482	0.004	0.233	0.767
Variance	2.950	1.451		

(ii) Three-factor solution

Statement	*Factor 1*	*Factor 2*	*Factor 3*	*Communality*	*Specific variance*
1	0.605	0.295	0.372	0.592	0.408
2	–0.455	0.291	0.431	0.477	0.523
3	0.613	0.498	0.192	0.641	0.339
4	0.621	0.399	0.000	0.545	0.455
5	–0.407	0.450	0.372	0.506	0.494
6	–0.671	0.594	–0.149	0.825	0.175
7	–0.625	0.342	–0.060	0.512	0.488
8	0.681	0.475	–0.273	0.253	0.237
9	0.449	0.162	–0.138	0.247	0.753
Variance	3.007	1.502	0.619		

from the value one gives the specific variance of a variable, that is, the variation in the variable *not* shared with the other variables. So, for example, in the two factor solution the communality of the statement "people who are never in pain are just plain lucky," is rather low at 0.23 and its specific variance consequently relatively high at 0.77. Variation in the response to this statement is largely unrelated to the two common factors.

The sum of squares of the loadings on a common factor gives the variation in the manifest variables accounted for by that factor. This is to be compared with the total variation in the observed variables, which since this example uses a correlation matrix and hence relates to variables standardized to have variance one, is simply equal to the number of variables, that is, nine. So, in the two-factor solution, the first factor has variance 2.95 and accounts for 33% of the variation in the observed variables. Both factors together in the two-factor solution account for 49% of the variance. The factors in the three-factor solution together account for 57% of the variance.

Note that factors are extracted in order of their variance, and are so constructed that they are uncorrelated, that is, independent—an alternative technical term that is sometimes encountered is "orthogonal."

3.13.3.1.2 Determining the number of factors

Before rotating and interpreting a factor solution the investigator needs to answer the important question, "How many factors?" A variety of informal and formal methods have been suggested. The former include, taking as many factors as account for an adequate amount of the variation in the observed variables (where "adequate" is usually interpreted as roughly around 60% or above), and plotting factor variances against factor number (a so-called scree plot) and identifying the point where the curve flattens out. When applying the maximum likelihood method of factor analysis a more formal significance testing procedure is available based on what is known as the likelihood function, which is essentially, a measure of how well the estimated factor solution fits the observed correlations. Everitt and Dunn (1991) give a specific definition. With this approach a sequential procedure is used to determine k, the number of common factors. Starting with some small value of k (usually one), the test for number of factors is applied and, if the test is nonsignificant, the current value of k is deemed acceptable; otherwise k is increased by one and the process repeated until an acceptable solution is found.

Each of the procedures described above can be applied to the pain statements data, and the results are shown in Table 7 and Figure 1. It is clear from these results that the three-factor solution is the one to choose and is consequently subjected to the process of rotation described in the next section.

3.13.3.1.3 Factor rotation

The initial factors extracted from a factor analysis are often difficult to interpret and

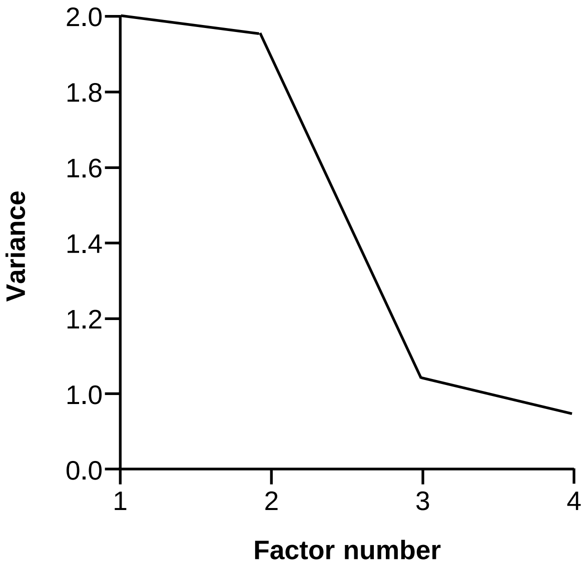

Figure 1 Scree plot.

Table 7 Determining number of factors for pain statement data.

(i) Percentage variance accounted for in a five-factor solution

Number of factors	*Percentage variance accounted for*
1	22
2	44
3	55
4	65
5	70

(ii) Significance test for number of factors

Model	*Test statistic*	*Degrees of freedom*	*P-value*
One-factor	172.34	27	< 0.001
Two-factor	58.95	19	< 0.001
Three-factor	18.19	12	0.11
Four-factor	10.01	6	0.124
Five-factor	0.85	1	0.356

name. A process which can frequently aid in these tasks is factor rotation whereby the initial solution is described in a different and, in many cases, a simpler fashion. A rotated and unrotated factor analysis solution are mathematically equivalent, but the former usually leads to a clearer picture of the nature of the underlying latent variables. The rotation methods usually employed are designed to lead to a factor solution with the properties that Thurstone (1947) referred to as a simple structure. In very general terms such a structure results when the common factors involve subsets of the original variables with as little overlap as possible, i.e., variables have high loadings on a particular factor and negligible loadings on the others. In this way the original variables are divided into groups relatively independent of each other. Methods of rotation operate by seeking, essentially, to make large loadings larger and small loadings smaller. This rather vague aim is translated into more specific mathematical terms by selecting a rotated solution so that the loadings optimize some suitable numerical criterion. For example, a well known method of rotation known as varimax attempts to maximize the within-factor variance of the squared loadings. Other methods (of which there are several) choose to optimize somewhat different criteria in their aim to achieve simple structure. In many examples the solutions given by the competing methods of rotation will be very similar.

To illustrate the application of rotation, Table 8 shows the varimax-rotated, three-factor solution for the pain statement data. Note that although the factor loadings have changed, the communalities of the variables are unaltered, as is the total variance accounted for by the solution. The variance attributable to each common factor, *has* however, changed.

A possible interpretation of the rotated three-factor solution is in terms of different aspects of the control of, and responsibility for, one's pain. The first factor attributes both to others, particularly doctors. The second factor, with high loadings on statements 6 and 7, involves complete personal responsibilty for one's pain and the third factor, having its highest loadings on statements 2 and 5, might be seen as attributing pain to deficiencies in one's lifestyle.

The possibility of rotating factor solutions arises because of the lack of uniqueness of the factor loadings in the basic factor analysis model described in Table 5 (Everitt, 1996). This property once caused many statisticians to view factor analysis with grave suspicion, since apparently it allows investigators licence to consider a large number of solutions (each corresponding to a different rotation of the factors) and to select the one closest to their *a priori* expectations (or prejudices) about the factor structure of the data. In general, however, such suspicion is misplaced and factor rotation can be a useful procedure for simplifying an exploratory factor analysis solution. Factor rotation merely allows the fitted factor analysis model to be described as simply as possible. Rotation does not alter the overall structure of a solution, but only how the

Table 8 Rotated three-factor solution for pain statement data.

Statement	*Factor 1*	*Factor 2*	*Factor 3*	*Communality*
1	0.654	–0.358	0.188	0.591
2	–0.121	0.200	0.650	0.478
3	0.793	–0.138	0.112	0.661
4	0.726	–0.10	–0.09	0.545
5	0.02	0.305	0.642	0.506
6	–0.09	0.827	0.365	0.825
7	–0.226	0.598	0.323	0.512
8	0.813	0.06	–0.312	0.763
9	0.435	–0.08	–0.229	0.247
Variances	2.509	1.340	1.279	

solution is described; rotation of factors is a process by which a solution is made more interpretable without changing its underlying mathematical properties.

It should be noted that there are two distinct types of rotation, orthogonal and oblique. With the former, the factors in the rotated solution remain independent of one another as they were in the initial solution but, with the latter, correlated factors are allowed. The consequence of allowing correlations between factors is that the sum of squares of a factor's loadings can no longer be used to determine the amount of variance attributable to a factor. Additionally, the sums of squares of factor loadings for each variable no longer give the communality of the variable. In practice, in an exploratory factor analysis, orthogonal rotation is far more commonly used than oblique rotation since the solutions are often satisfactory without introducing the complication of factor correlations.

3.13.3.2 Confirmatory Factor Analysis

In a confirmatory factor analysis, particular parameters of the model, for example, some of the factor loadings, are specified *a priori* to be fixed at some specified value, usually zero or one, whilst other parameters are free to be estimated from the data. The model considered may have arisen from theoretical considerations, or perhaps on the basis of the results of earlier exploratory factor analyses in the same area.

How is such a model fitted to an observed correlation or covariance matrix? The simple answer is by using the appropriate software as described in the Appendix. The essentials of the fitting process are, however, also given in Table 9. Once again, however, much can be learnt from examining a series of examples.

3.13.3.2.1 Confirmatory factor analysis models for length and intelligence

Consider once again the string data given in Table 3. The model suggested previously for these data is that the four observed measurements of string length are fallible estimates of the latent variable true string length. This model can be represented more specifically in mathematical terms as shown in Table 10, and in the form of a diagram as shown in Figure 2. The model is a very simple example of a confirmatory factor analysis model, and differs from the exploratory models considered in the previous section in two ways: (i) the number of factors, one, is specified *a priori*; and (ii) the loadings of the observed variables on the postulated latent variable are all fixed to be one.

The correlations between the four measurements are given in Table 11. Applying the procedure outlined in Table 9 to these correlations using the EQS package (Bentler, 1995), gives the results shown in Table 12.

Judged by the chi-squared statistic the model fits very well. As was to be expected, measurement by ruler is the most accurate; observer B appears to be the marginally less successful than G and D in guessing the string lengths. The residuals are all very small.

In the same way, the model detailed in Table 9 can be fitted to the examination scores data in Table 1. Details of the results are shown in Table 13. Here the model again fits reasonably well as judged simply by the chi-squared statistic, but rather less satisfactorily than for the string data, reflecting the less certain nature of intelligence as compared to length. In fact, the residuals in Table 13 indicate that the predicted and observed correlations differ quite considerably and that the model cannot really be considered adequate for these data despite the nonsignificant chi-squared value. The latter may be misleading here since it is not strictly

Table 9 Estimation and goodness of fit in confirmatory factor analysis models.

(i) Suppose that **S** is the observed covariance (or correlation) matrix with elements s_{ij}. Corresponding to **S** will be a matrix $\hat{\mathbf{S}}$ giving the correlations as predicted by the assumed model

(ii) The elements of the matrix $\hat{\mathbf{S}}$ will be functions of the parameters of the confirmatory factor analysis model, for example, factor loadings, specific variances, and factor correlations

(iii) Estimation of a confirmatory factor analysis (CFA) model involves finding values for the model's parameters (parameter estimates) so that the corresponding predicted and observed covariances and variances are as close as possible to one another

(iv) More specifically, a function of the differences between s_{ij} and $\hat{s}_{ij}$ is minimized with respect to the parameters

(v) One such function would be the sum of squares of the differences; this corresponds to least-squares estimation

(vi) Other estimation criteria are discussed in Everitt (1984)

(v) Having decided on a measure of closeness, some type of mathematical optimization algorithm (Everitt, 1987) is applied to find the parameter values that minimize the measure

(viii) The fitting procedure leads not only to parameter estimates but also to the standard errors of the estimates

(ix) A chi-squared statistic measuring the discrepancies between the observed variances and covariances and those predicted by the model is one of the methods that can be used to assess the fit of the model. The degrees of freedom are

$$\nu = p(p + 1)/2 - t$$

where p is the number of observed variables and t is the number of free parameters in the model

(x) Fit can (and should) also be assessed by examining the differences between the corresponding elements of the observed and predicted covariance or correlation matrix. Any differences (residuals) which are unacceptably large, or a pattern of moderately sized residuals sheds some doubt on the acceptability of the model

valid for samples as small as 15. Questions of the appropriate sample size necessary to achieve an adequate statistical power are as relevant when fitting confirmatory factor analysis as when applying a simple t-test (examples of how to determine sample size for the former are given in Dunn, Everitt, & Pickles, 1993).

A natural question that arises therefore is how the model might be amended to improve its fit. Some general points about how poorly fitting models might be amended are made later, but one relatively minor change that might be considered here is to allow the loading on each examination score to be a free parameter to be estimated rather than having a fixed value of one. Such a model is again easily fitted using the EQS software and the results are given in Table 14. Now the residuals are much more acceptable with only the correlation between algebra and statistics scores being poorly predicted. The estimated factor loadings show that the algebra scores are least related to the postulated latent variable.

3.13.3.2.2 Ability and aspiration

As a further example of fitting a confirmatory factor analysis model, the study of ability and aspiration described in Caslyn and Kenny (1977) is used. In this study the investigators observed the values of the following six variables for 556 white eighth-grade students: (1) self-concept of ability, (2) perceived parental evaluation, (3) perceived teacher evaluation, (4) perceived friend's evaluation, (5) educational aspiration, and (6) college plans.

The correlations between the six variables are given in Table 15. The model postulated to explain these correlations involves two latent variables, named by the authors as ability and aspiration. The first four observed variables are assumed to be indicators of ability and the last two observed variables are assumed to be indicators of aspiration. In addition the two latent variables are assumed to be correlated. Consequently the model is known as a correlated two-factor model. It may be represented in a diagram (a so-called path diagram (Everitt, 1996)) as shown in Figure 3. The mathematical form of the model appears in Table 16. Note again the differences between this model and the exploratory models of Section 3.13.3.1; here, each observed variable is constrained to have fixed zero loadings on one of the latent variables, and on the other, a free loading to be estimated from the data. The results of fitting the model to the observed correlations, using once again the EQS package, are shown in Table 17. Of particular note among these results is the estimate of the correlation between the two postulated latent variables. This estimate

Table 10 Mathematical specification of one-factor model for string data.

The factor analysis model for the string length data is

$$\begin{aligned} R &= f + e_1 \\ G &= f + e_2 \\ B &= f + e_3 \\ D &= f + e_4 \end{aligned}$$

where f represents the true length and e_1, e_2, e_3, and e_4 are the errors in the four observed length measurements. These errors are assumed to have zero mean and their estimated variances will reflect the accuracy of R, G, B, and D

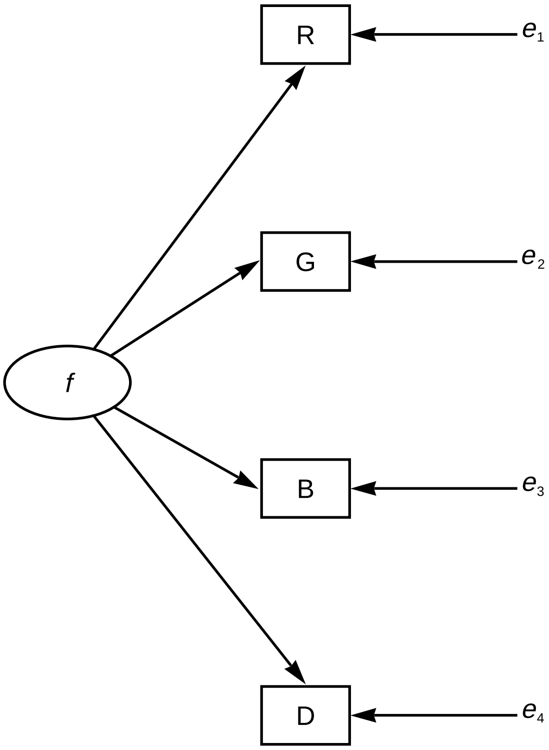

Figure 2 Single-factor model for string lengths.

(0.666 with a standard error of 0.03) is the disattenuated correlation, which represents the correlation between true ability and true aspiration, uncontaminated by measurement error in the observed indicators of these concepts. Note that the value of the disattenuated correlation is higher than any of the observed correlations.

3.13.3.2.3 Drug usage among American students

As a final example of fitting a confirmatory factor analysis model, an investigation of drug usage among American college students reported by Huba, Wingard, and Bentler (1981) is considered.

The majority of adult and adolescent Americans use psychoactive substances during an increasing proportion of their lifetime. Various forms of licit and illicit psychoactive substance use are prevalent, suggesting that patterns of psychoactive substance taking are major components of the individual's behavioral repertoire and have pervasive implications for the performance of other behaviors. In an investigation of these phenomena, Huba et al. (1981) collected data on drug usage rates for 1634 students in the seventh to ninth grades in 11 schools in the greater metropolitan area of Los Angeles. Each participant completed a questionnaire about the number of times he or she had ever used a particular substance.

Table 11 Correlation matrix for the four measurements of string length.

$$\mathbf{R} = \begin{array}{c} \\ R \\ G \\ B \\ D \end{array} \begin{array}{c} \begin{array}{cccc} \text{R} & \text{G} & \text{B} & \text{D} \end{array} \\ \left[\begin{array}{cccc} 1.0000 & & & \\ 0.9802 & 1.0000 & & \\ 0.9811 & 0.9553 & 1.0000 & \\ 0.9899 & 0.9807 & 0.9684 & 1.0000 \end{array}\right] \end{array}$$

Table 12 Results from fitting one-factor model with equal loadings for each variable to string data.

(i) Residuals (differences between observed and predicted correlations)

$$\mathbf{R} = \begin{array}{c} \\ R \\ G \\ B \\ D \end{array} \begin{array}{c} \begin{array}{cccc} \text{R} & \text{G} & \text{B} & \text{D} \end{array} \\ \left[\begin{array}{cccc} 0.007 & & & \\ -0.010 & -0.025 & & \\ -0.009 & -0.035 & -0.028 & \\ 0.000 & -0.009 & -0.022 & -0.007 \end{array}\right] \end{array}$$

(ii) Variances of error terms for each observed variable and their standard errors (SE)

R: variance = 0.003, SE = 0.005
G: variance = 0.035, SE = 0.014
B: variance = 0.038, SE = 0.016
D: variance = 0.017, SE = 0.008

(iii) Chi-squared goodness of fit statistic = 2.575, degrees of freedom = 5, *P*-value = 0.765

The substances for which data were collected were as follows: (1) cigarettes, (2) beer, (3) wine, (4) spirits, (5) cocaine, (6) tranquilizers, (7) drugstore medications used to get high, (8) heroin and other opiates. (9) marijuana, (10) hashish, (11) inhalents (glue, petrol, etc.), (12) hallucinogenics (LSD, mescalin, etc.), and (13) amphetamine stimulants.

Responses were recorded on a five-point scale which ranged from "never tried" to "used regularly." The correlations between the usage rates of the 13 substances are shown in Table 18. The model proposed by Huba et al. (1981) for these data arose from considering previously reported research in the area, and postulated the following three latent variables:

(i) Alcohol use (f_1) with nonzero loadings on beer, wine, spirits, and cigarettes.

(ii) Cannabis use (f_2) with nonzero loadings on marijuana, hashish, cigarettes, and wine. The cigarette variable is assumed to load on both the first and second latent variable because it sometimes occurs with both alcohol and marijuana and at other times does not. The nonzero loading on wine was allowed because of reports that wine is frequently used with marijuana and that consequently some of the use of wine may be an indicator of tendencies towards cannabis.

(iii) Hard drug use (f_3) with nonzero loadings on amphetamines, tranquilizers, hallucinogenics, hashish, cocaine, heroin, drugstore medication, inhalants, and spirits. The use of

Table 13 Results of fitting the one-factor model to examination scores, with loadings fixed at one.

(ii) Correlations

$$\begin{array}{l|cccc|} & \text{English} & \text{French} & \text{Algebra} & \text{Statistics} \\ \text{English} & 1.000 & & & \\ \text{French} & 0.683 & 1.000 & & \\ \text{Algebra} & 0.286 & 0.451 & 1.000 & \\ \text{Statistics} & 0.431 & 0.690 & 0.544 & 1.000 \end{array}$$

(ii) Residuals

$$\begin{array}{l|cccc|} & \text{English} & \text{French} & \text{Algebra} & \text{Statistics} \\ \text{English} & -0.113 & & & \\ \text{French} & 0.104 & 0.147 & & \\ \text{Algebra} & -0.293 & -0.128 & -0.265 & \\ \text{Statistics} & -0.148 & 0.111 & -0.035 & 0.025 \end{array}$$

(iii) Variances of error terms for each observed variable and their standard errors

English: variance = 0.534, SE = 0.244
French: variance = 0.274, SE = 0.156
Algebra: variance = 0.686, SE = 0.299
Statistics: variance = 0.396, SE = 0.196

(iv) Chi-squared goodness-of-fit statistic = 4.136, degrees of freedom = 5, *P*-value = 0.530.

each of these substances was considered to suggest a strong commitment to the notion of psychoactive drug use.

The path diagram of the proposed model is shown in Figure 4, and Table 19 shows the equivalent mathematical structure of the model. Note that each pair of the postulated variables is allowed to have a nonzero correlation. The results of fitting the model are detailed in Table 20. The chi-squared goodness-of-fit statistic takes a value 323.96 with 58 degrees of freedom and has a very small associated *P*-value. It appears that the proposed model does *not* provide an adequate explanation for the correlations between the recorded usage rates of the 13 substances, although the large sample size ($n = 1634$) may lead to even relatively trivial discrepancies between observed and predicted correlations being declared significant. In practice the chi-squared statistic is only one of the pieces of evidence that should be used when judging the fit of a model; other measures of fit are discussed in Dunn et al. (1993).

Amending a model that is considered not to provide an adequate fit will generally involve a mixture of the theoretical and empirical. Information from the data about possibly better fitting models is provided by two types of tests, the Lagrange multiplier test and Wald's test, both of which are described in detail in the EQS manual (Bentler, 1995). Essentially, however, the former evaluates whether, from a statistical viewpoint at least, the model could be improved by freeing a previously fixed parameter. The latter is designed to determine whether sets of parameters that were treated as free in the model could in fact be simultaneously set to zero without substantial loss in model fit. Examples of the use of the two test procedures are given in Dunn et al. (1993). In

Table 14 Results of fitting further single-factor model to examination scores; loadings now free parameters to be estimated.

(i) Residuals

$$\begin{array}{l c} & \begin{array}{cccc} \text{English} & \text{French} & \text{Algebra} & \text{Statistics} \end{array} \\ \begin{array}{l} \text{English} \\ \text{French} \\ \text{Algebra} \\ \text{Statistics} \end{array} & \left[\begin{array}{cccc} 0.000 & & & \\ 0.010 & 0.000 & & \\ -0.052 & -0.016 & 0.000 & \\ -0.070 & -0.002 & 0.197 & 0.000 \end{array} \right] \end{array}$$

(ii) Variances of error terms for each observed variable and their standard errors

English: variance = 0.513, SE = 0.226
French: variance = 0.069, SE = 0.217
Algebra: variance = 0.766, SE = 0.298
Statistics: variance = 0.485, SE = 0.221

(iii) Estimated factor loadings and their standard errors

English: loading = 0.698, SE = 0.247
French: loading = 0.965, SE = 0.225
Algebra: loading = 0.484, SE = 0.262
Statistics: loading = 0.718, SE = 0.246

(iv) Chi-squared goodness-of-fit statistic = 1.940, degrees of freedom = 2, P-value = 0.379

the original analysis of the drug usage data given in Huba et al. (1981) various amendments to the correlated three-factor model were tried in order to improve fit; for example, a number of error terms were allowed to be correlated. However, making such amendments to gain small improvements in fit is not always satisfactory and it should perhaps be pointed out that, in most cases, a theoretically justifiable model that provides a less adequate fit than a model adjusted on an *ad hoc* basis, is likely to be preferred.

Recent examples of the application of confirmatory factor analysis in clinical psychology are given in Osman et al. (1995) and Hittner (1995).

3.13.4 CAUSAL MODELS AND STRUCTURAL EQUATION MODELING

The confirmatory factor analysis models described in the previous section are a subset of the models for correlational data that can now be fitted routinely using software such as EQS or LISREL (see Appendix). More complex models that specify regression-type relationships between latent variables, as well as how the manifest and latent variables are linked, have now been used in many areas from models of the female orgasm (Newcomb & Bentler, 1983) to models to represent both psychological and physical processess in cognitive functioning (Hines, Chiu, McAdams, Bentler, & Lipcamon, 1992).

Such models are most generally (and most suitably) referred to as structural equation models. However, since they are often used as a means of describing some prespecified causal theory of the structure of a set of variables of interest, they are also often labeled causal models where "causal" implies that a change in one variable is assumed to result in the change of another variable. No matter how convincing, respectable, and reasonable a path diagram and its associated model may appear, it is important to recognize that seldom do structural equation models provide any direct test of the causal

Table 15 Observed correlations for ability and aspiration example.

$$\begin{array}{ccc} & & \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \mathbf{R} = & \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} & \left[\begin{array}{cccccc} 1.00 & & & & & \\ 0.73 & 1.00 & & & & \\ 0.70 & 0.68 & 1.00 & & & \\ 0.58 & 0.61 & 0.57 & 1.00 & & \\ 0.46 & 0.43 & 0.40 & 0.37 & 1.00 & \\ 0.56 & 0.52 & 0.48 & 0.41 & 0.72 & 1.00 \end{array} \right] \end{array}$$

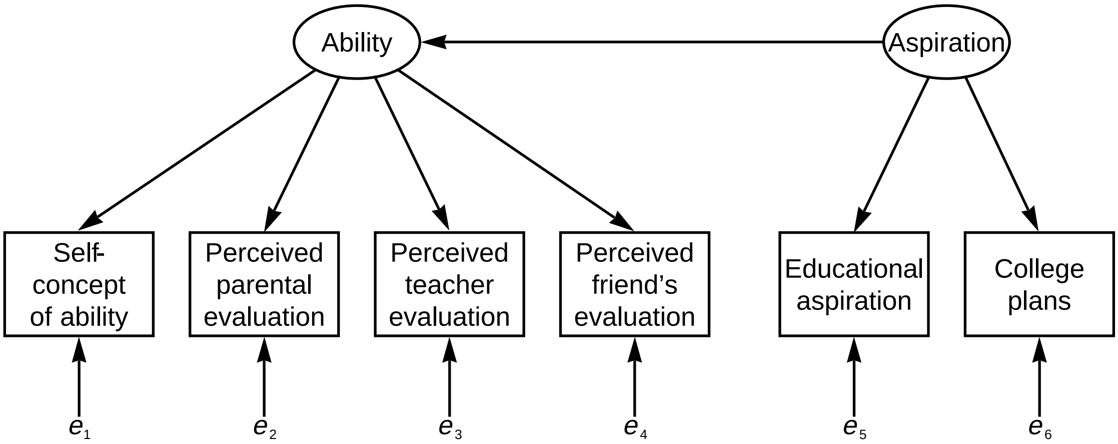

Figure 3 Path diagram for ability and aspiration.

assumptions on which these models are based. This is generally true simply because, in most cases, an investigator's belief about the causal structures which underlie a set of observed variables are based on little more than common sense and intuition; it is only rarely that there is strong evidence (such as that provided by controlled experimentation) about the causal structure involved, and the only satisfactory way to demonstrate causality is through the active control of variables. As pointed out by Cliff (1983) it is simply not possible, with correlational data, to isolate the empirical system sufficiently so that the nature of the relationships among variables can be unambiguously ascertained. It seems that the old aphorism much stressed in introductory statistics courses, "correlation does not imply causation" is still apposite even in the sophisticated world of structural equation modeling.

Many investigators proposing causal models might, of course, argue that they are using the term causal in a purely metaphorical fashion. As pointed out by deLeeuw (1985), however, such a cavalier attitude towards terminology becomes hard to defend if, for example, educational programs are based on the metaphors, such as "intelligence is largely genetically determined" or "allocation of resources to schools has only very minor impact on the careers of students."

Many of the problems of so called causal modeling mentioned above arise largely because of the relative lack of well specified causal theories in the social sciences. However, not all is doom and gloom for an investigator eager to try out such models. Although it may be the case

Table 16 Mathematical structure of correlated two-factor model for ability and aspiration.

(i) The two common factors or latent variables are ability (f_1) and aspiration (f_2) Both are assumed to have variance of one (this is necessary since they are unobserved and their scale needs to be set in some arbitrary way)

(ii) The proposed model postulates that the relationships between the observed variables and the latent variables are as follows:

$$\begin{aligned}
x_1 &= \lambda_1 f_1 + 0 f_2 + e_1\\
x_2 &= \lambda_2 f_1 + 0 f_2 + e_2\\
x_3 &= \lambda_3 f_1 + 0 f_2 + e_3\\
x_4 &= \lambda_4 f_1 + 0 f_2 + e_4\\
x_5 &= 0 f_1 + \lambda_5 f_2 + e_5\\
x_6 &= 0 f_1 + \lambda_6 f_2 + e_6
\end{aligned}$$

(iii) This may be rewritten as

$$\begin{aligned}
x_1 &= \lambda_1 f_1 + e_1\\
x_2 &= \lambda_2 f_1 + e_2\\
x_3 &= \lambda_3 f_1 + e_3\\
x_4 &= \lambda_4 f_1 + e_4\\
x_5 &= \lambda_5 f_2 + e_5\\
x_6 &= \lambda_6 f_2 + e_6
\end{aligned}$$

(iv) Note that, unlike an exploratory factor analysis, a number of loadings are fixed *a priori* at zero, that is, they play no part in the estimation process

(v) The model also allows for f_1 and f_2 to be correlated

(vi) The model has a total of 13 free parameters (six loadings, six error variances and one correlation). The observed correlation matrix has six variances and 15 correlations, a total of 21 terms. Consequently the postulated model has $21 - 13 = 8$ degrees of freedom

Table 17 Results from fitting the correlated two-factor model to correlations in Table 15.

Parameter	*Estimates*	*Standard error*	*Estimate/standard*
λ_1	0.863	0.035	24.558
λ_2	0.849	0.035	23.961
λ_3	0.805	0.035	22.115
λ_4	0.695	0.039	18.000
λ_5	0.775	0.040	19.206
λ_6	0.929	0.039	23.569
var(e_1)	0.255	0.023	19.911
var(e_2)	0.279	0.024	11.546
var(e_3)	0.352	0.027	13.070
var(e_4)	0.516	0.035	14.876
var(e_5)	0.399	0.038	10.450
var(e_6)	0.137	0.044	3.152
corr(f_1, f_2)	0.667	0.031	21.521

The chi-square test of the fit of the model takes the value 9.26 with 8 degrees of freedom. The associated *P*-value is 0.321. The model provides a very adequate fit for the data.

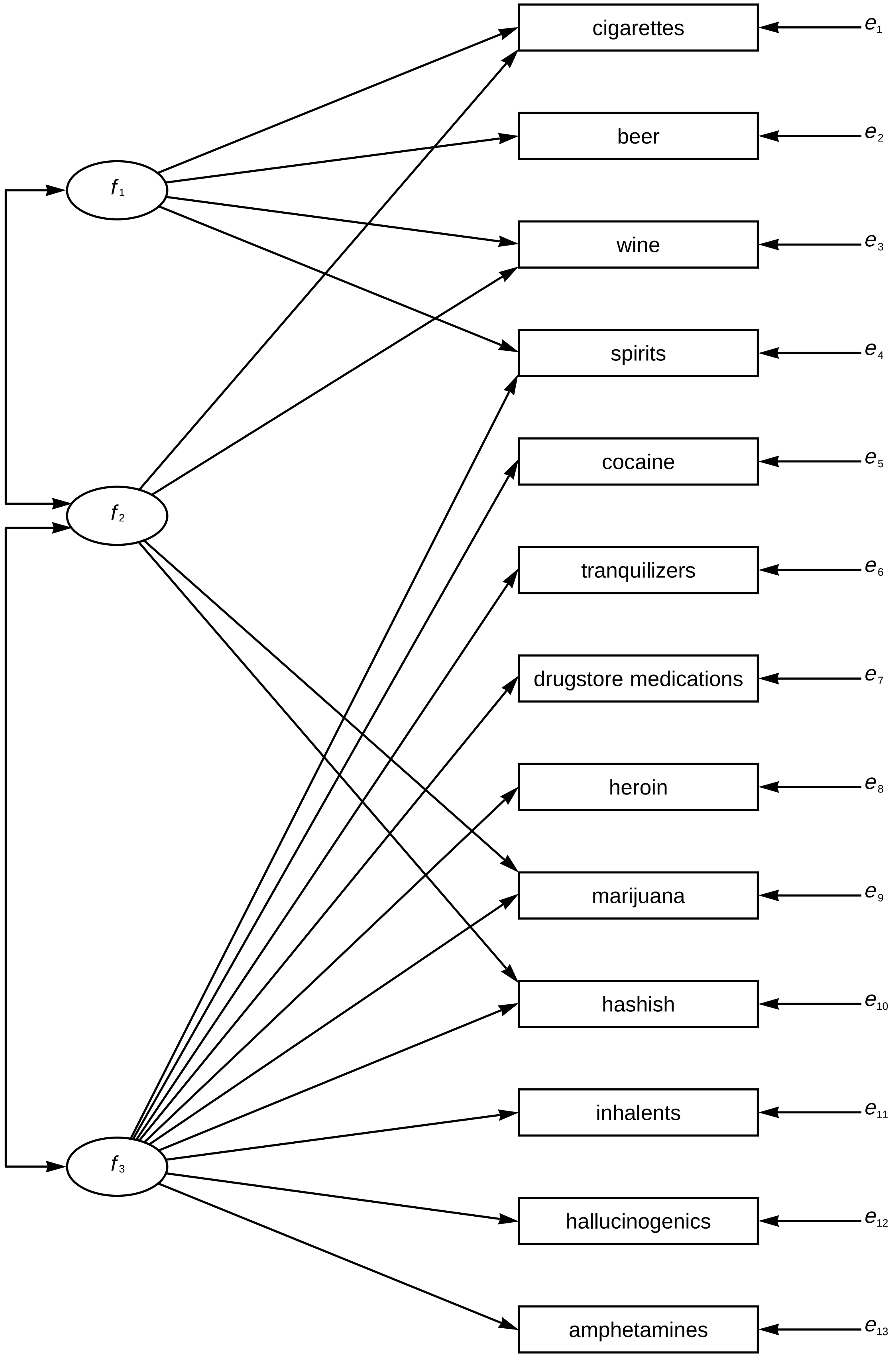
cigarettes
beer
wine
spirits
cocaine
tranquilizers
drugstore medications
heroin
marijuana
hashish
inhalents
hallucinogenics
amphetamines
f_1
f_2
f_3
e_1
e_2
e_3
e_4
e_5
e_6
e_7
e_8
e_9
e_{10}
e_{11}
e_{12}
e_{13}

Figure 4 Path diagram for three-factor model for drug usage example.

Table 18 Correlations between usage rates for 13 substances (key to substances is given in text).

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1.000												
2	0.447	1.000											
3	0.422	0.619	1.000										
4	0.436	0.604	0.585	1.000									
5	0.114	0.068	0.053	0.115	1.000								
6	0.203	0.146	0.139	0.258	0.349	1.000							
7	0.091	0.103	0.110	0.122	0.209	0.221	1.000						
8	0.082	0.063	0.066	0.097	0.321	0.355	0.201	1.000					
9	0.513	0.445	0.365	0.482	0.186	0.316	0.150	0.154	1.000				
10	0.304	0.318	0.240	0.368	0.303	0.377	0.163	0.219	0.534	1.000			
11	0.245	0.203	0.183	0.255	0.272	0.323	0.310	0.288	0.301	0.302	1.000		
12	0.100	0.088	0.074	0.139	0.279	0.367	0.232	0.320	0.204	0.368	0.340	1.000	
13	0.245	0.199	0.184	0.293	0.278	0.545	0.232	0.314	0.394	0.467	0.392	0.511	1.000

Table 19 Mathematical structure of the correlated three-factor model for the drug usage data.

(i) Three latent variables, alcohol use (f_1), cannabis use (f_2), and hard drug use (f_3) are postulated. All are assumed to have variance unity

(ii) The proposed model postulates the following relationship between the observed and latent variables:

$$\begin{aligned}
\text{cigarettes} &= \lambda_1 f_1 + \lambda_2 f_2 + e_1 \\
\text{beer} &= \lambda_3 f_1 + e_2 \\
\text{wine} &= \lambda_4 f_1 + \lambda_5 f_2 + e_3 \\
\text{spirits} &= \lambda_6 f_1 + \lambda_7 f_3 + e_4 \\
\text{cocaine} &= \lambda_8 f_3 + e_5 \\
\text{tranquilizers} &= \lambda_9 f_3 + e_6 \\
\text{drugstore} &= \lambda_{10} f_3 + e_7 \\
\text{heroin} &= \lambda_{11} f_3 + e_8 \\
\text{marjuanna} &= \lambda_{12} f_2 + e_9 \\
\text{hashish} &= \lambda_{13} f_2 + \lambda_{14} f_3 + e_{10} \\
\text{inhalants} &= \lambda_{15} f_3 + e_{11} \\
\text{hallucinogenics} &= \lambda_{16} f_3 + e_{12} \\
\text{amphetamines} &= \lambda_{17} f_3 + e_{13}
\end{aligned}$$

(iii) The proposed model also allows for nonzero correlations between each pair of latent variables. The proposed model has a total of 33 parameters to estimate (17 loadings, 13 error variances, and 3 between-factor correlations). Consequently, the model has $13 \times 14/2 - 33 = 58$ degrees of freedom

that correlation does not imply causation, it is equally true that a well specified causal model implies testable propositions about the structure of observed correlations, and so is amenable to falsification in the same way as is any other scientific theory. At the very least, such a model may provide a convenient and parsimonious description of a set of correlations and serve to rule out many alternative hypotheses about the structure of the data.

3.13.4.1 Examples of Causal Models

After the rather lengthy discussion of their merits or otherwise given above, it is now time to look at some examples of the application of structural equation/causal modeling.

3.13.4.1.1 Ability scores over time

Table 21 gives, for a group of children, the means, standard deviations, correlations and covariances for four measures of ability (percentage of test questions correct) made at the ages of 6, 7, 9, and 11 (Osbourne & Suddick, 1972). As one would expect, the means increase progressively over time, as do the standard deviations. The correlation matrix shows higher correlations between adjacent measures than between those further apart in time.

The simplest model that might be considered for these data is that specified by the path diagram shown in Figure 5; here, ability is assumed to be a fixed trait and that variation over time arises solely from measurement error. Fitting such a model to the covariance matrix in

Table 20 Results of fitting the correlated three-factor model to the drug usage data.

Parameter	*Estimates*	*Standard error*	*Estimate/standard error*
λ_1	0.358	0.035	10.371
λ_2	0.332	0.035	9.401
λ_3	0.792	0.023	35.021
λ_4	0.875	0.038	23.285
λ_5	–0.152	0.037	–4.158
λ_6	0.722	0.024	30.673
λ_7	0.123	0.023	5.439
λ_8	0.465	0.026	18.079
λ_9	0.676	0.024	28.182
λ_{10}	0.359	0.025	13.602
λ_{11}	0.476	0.026	18.571
λ_{12}	0.912	0.030	29.958
λ_{13}	0.396	0.030	13.379
λ_{14}	0.381	0.029	13.050
λ_{15}	0.543	0.025	21.602
λ_{16}	0.618	0.025	25.233
λ_{17}	0.763	0.023	32.980
var(e_1)	0.611	0.024	25.823
var(e_2)	0.374	0.020	18.743
var(e_3)	0.379	0.024	16.052
var(e_4)	0.408	0.019	21.337
var(e_5)	0.784	0.029	26.845
var(e_6)	0.544	0.023	23.222
var(e_7)	0.871	0.032	27.653
var(e_8)	0.773	0.029	26.735
var(e_9)	0.169	0.044	3.846
var(e_{10})	0.547	0.022	24.593
var(e_{11})	0.705	0.027	25.941
var(e_{12})	0.618	0.025	24.655
var(e_{13})	0.418	0.021	19.713
corr(f_1,f_2)	0.634	0.027	23.369
corr(f_1,f_3)	0.313	0.029	10.674
corr(f_2,f_3)	0.499	0.027	18.412

Table 21 gives the results shown in Table 22. The chi-squared goodness-of-fit statistic is significant at the 5% level, suggesting that the simple single factor model does not provide an adequate fit for the covariance matrix.

In this example, the observed covariance matrix has been used as the basis for fitting the model of interest. In general the question of whether the covariance or correlation matrix should be used as the basis for structural equation modeling is probably not of great importance, although there are some situations where the covariance matrix is definitely to be preferred (Cudeck, 1989).

A path diagram for a more plausible model for the ability data is shown in Figure 6; this model postulates causal effects between one latent variable and another and the presence of the disturbances terms on f_2, f_3, and f_4 (terms d_2, d_3, and d_4) means that latent ability is not regarded as an entirely fixed and stable trait, but may vary, increasing or decreasing relative to other children, from one time to the next. The results of fitting this model are shown in Table 23. The chi-squared test statistic now indicates that the model fits satisfactorily. The regression coefficients of each latent variable on the one preceding it in time are all highly significant (see Dunn et al., 1993, for a detailed accident).

3.13.4.1.2 Stability of alienation

As a further example of structural equation modeling a study reported by Wheaton, Muthen, Alwin, and Summers (1977) is used. The study was concerned with the stability over time of attitudes such as alienation and their relationships to background variables such as

Table 21 Ability scores over time.

Statistic	*Age 6*	*Age 7*	*Age 9*	*Age 11*
Mean	18.034	25.819	35.255	46.593
Standard deviation	6.374	7.319	7.796	10.386

(i) Correlation matrix

$$\begin{array}{lcccc} & \text{Age 6} & \text{Age 7} & \text{Age 9} & \text{Age 11} \\ \text{Age 6} & 1.000 & & & \\ \text{Age 7} & 0.809 & 1.000 & & \\ \text{Age 9} & 0.806 & 0.850 & 1.000 & \\ \text{Age 11} & 0.765 & 0.831 & 0.867 & 1.000 \end{array}$$

(ii) Covariance matrix

$$\begin{array}{lcccc} & \text{Age 6} & \text{Age 7} & \text{Age 9} & \text{Age 11} \\ \text{Age 6} & 40.628 & & & \\ \text{Age 7} & 37.741 & 53.568 & & \\ \text{Age 9} & 40.052 & 48.500 & 60.778 & \\ \text{Age 11} & 50.643 & 63.169 & 70.200 & 107.869 \end{array}$$

education and occupation. Data were collected on attitude scales from 932 people in two rural regions in Illinois at three points in time (1966, 1967, and 1971). Only that part of the data collected in 1967 and 1971 will be of concern here and Table 24 shows the covariances between six observed variables. The anomia and powerlessness subscales are taken to be indicators of a latent variable, alienation, and the two background variables, education (years of schooling completed) and Duncan's socioeconomic index (SEI) are assumed to relate to a respondent's socioeconomic status. The path diagram for the model postulated to explain the covariances between the observed variables is shown in Figure 7. The model involves a combination of a confirmatory factor analysis model with a regression model for the latent variables. One of the important questions here involves the size of the regression coefficient of alienation in 1971 on alienation in 1967, since this reflects the stability of the attitude over time. Note that the error terms of anomia and powerlessness are allowed to be correlated over time to account for possible memory or other retest effects. Some of the results of fitting the proposed model are shown in Table 25.

The chi-squared goodness-of-fit statistic takes a value of 4.73 with four degrees of freedom and suggests that the proposed model fits the observed covariances extremely well. The estimated regression coefficient of alienation on socioeconomic status in both 1967 and 1971 is negative, as might have been expected since higher socioeconomic status is likely to result in lower alienation and vice versa. The

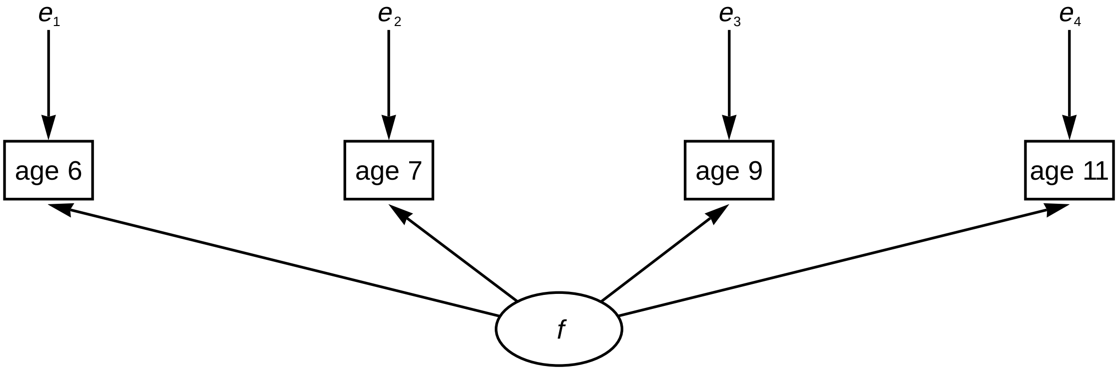

Figure 5 Single-factor model for ability data.

Table 22 Simple single-factor model—results for ability data.

Parameter	*Estimates*	*Standard error*	*Estimate/standard error*
λ_1	5.84	0.360	15.244
λ_2	6.684	0.397	16.842
λ_3	7.326	0.414	17.705
λ_4	9.476	0.564	16.813
var(e_1)	10.555	1.225	8.617
var(e_2)	8.89	1.197	7.427
var(e_3)	7.113	1.163	6.133
var(e_4)	18.084	2.424	7.460

(i) Chi-squared goodness-of-fit statistic 6.085, degrees of freedom = 2, *P*-value = 0.048.
(ii) Note that these estimates are obtained from the covariance matrix so that the factor loadings are no longer correlations between observed and latent variables. They now represent regression coefficients.
(iii) The statistical significance of each parameter can be judged by the *z*-statistics given in the final column. Values outside (–2,2) are roughly significant at the 5% level.

estimated regression coefficient for alienation in 1971 on alienation in 1967 is positive and highly significant. Clearly the attitude remains relatively stable over the time period.

An example of the application of structural equation modeling in clinical psychology is given in Taylor and Rachman (1994).

3.13.5 LATENT VARIABLES—MYTHS AND REALITIES

Having already commented in Section 3.13.4 that the "causal" in causal modeling is usually a misnomer, what can be said about the concept of a latent variable, central to the methods described in this chapter? In one sense, latent variables can never be anything more than is contained in the observed variables and never anything beyond what has been specified in the model. For example, in the statement that verbal ability is whatever certain tests have in common, the empirical meaning is nothing more than a shorthand for the observation of the correlations. It does not mean that verbal ability is a variable that is measurable in any manifest sense. In fact latent variables are essentially hypothetical constructs invented by a scientist for the purpose of understanding some research area of interest, and for which there exists no operational method for direct measurement. Consequently, a question that needs to be asked is "Can science advance by inferences based upon hypothetical constructs that cannot be measured or empirically tested?" According to Lenk (1986) the answer is a resounding "sometimes." For example, atoms in the eighteenth and nineteenth centuries were hypothetical constructs which allowed the foundation of thermodynamics; gravity and the electromagnetic field are further examples from physics. Clearly a science can advance using the concept of a latent variable, but their importance is not their reality or otherwise, but rather to what extent the models of which they are a part are able to describe and predict

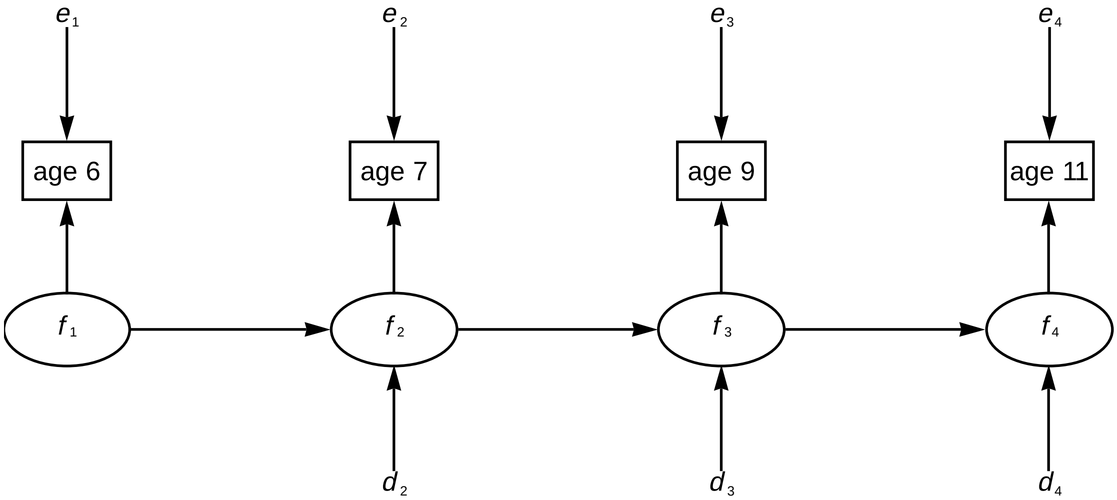

Figure 6 Causal model for ability data.

Table 23 Results from fitting causal model for ability data.

(i) Regression coefficients of latent variables on preceding latent variables

f_2 on f_1: 1.120, SE = 0.064
f_3 on f_2: 1.044, SE = 0.046
f_4 on f_3: 1.296, SE = 0.054

(ii) Chi-squared goodness-of-fit statistic = 1.433, degrees of freedom = 2, P-value = 0.489

phenomena (Lakatos, 1977). This point is nicely summarized by D. M. Fergusson and L. J. Horwood (personal communication, 1986):

> Scientific theories describe the properties of observed variables in terms of abstractions which summarize and make coherent the properties of observed variables. Latent variables, are, in fact, one of this class of abstract statements and the justification for the use of these variables lies not in an appeal to their "reality" or otherwise but rather to the fact that these variables serve to synthesize and summarize the properties of observed variables.

This point was also made by the participants in the Conference on Systems under Indirect Observation who concluded, after some debate (Bookstein, 1982) that latent variables are "as real as their predictive consequences are valid." Such a comment implies that the justification for postulating latent variables is their theoretical utility rather than their reality.

3.13.6 SUMMARY

The possibility of making causal inferences about latent variables has great appeal for the social and behavioural scientist, simply because many of the concepts in which they are most interested are not directly measurable. Many of the statistical and technical problems in applying the appropriate models to empirical data have largely been solved, and sophisticated software such as EQS means that researchers can investigate and fit extremely complex models routinely. Unfortunately, in their rush not to be left behind in the causal modeling stakes, many investigators appear to have abandoned completely their proper scientific skeptism, and accepted models as reasonable, simply because it has been possible to fit them to data. This would not be so important if it were not the case that much of the research involved is in areas where action, perhaps far-reaching action, taken on the basis of the findings of the research can have enormous implications, for example, in resources for education and

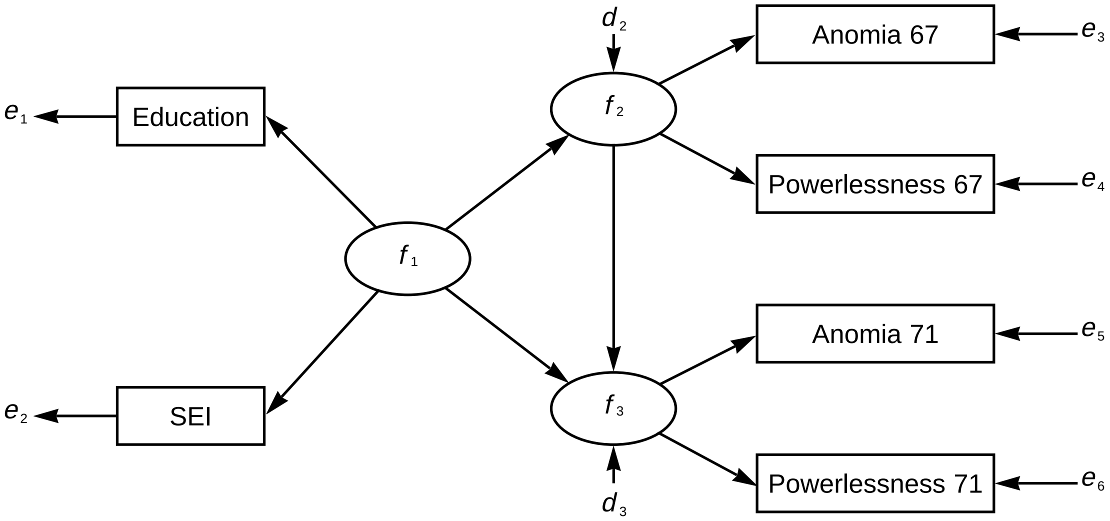

Figure 7 Causal model for stability of alienation.

Table 24 Covariance of manifest variables in the stability of alienation example.

	1	2	3	4	5	6
1	11.834					
2	6.947	9.364				
3	6.819	5.09	12.532			
4	4.783	5.028	7.495	9.986		
5	–3.839	–3.889	–3.841	–3.625	9.610	
6	–2.190	–1.883	–2.175	–1.878	3.552	4.503

Note. 1 = Anomia 67, 2 = Powerlessness 67, 3 = Anomia 71, 4 = Powerlessness 71, 5 = Education, 6 = Duncan's Socioeconomic Index.

Table 25 Regression coefficients for stability of the alienation model in Figure 5.

Alienation 67 on socioeconomic status: –1.500, SE = 0.124

Alienation 71 on alienation 67: 0.607, SE = 0.051

Alienation 71 on socioeconomic status: –0.592, SE = 0.131

legislation on racial inequality. Consequently, both producers of such research and audiences or consumers of it need to be particularly concerned that the conclusions reached are valid ones. With this in mind I would like to end with the caveat issued by Cliff (1983, p.125):

> beautiful computer programs do not really change anything fundamental. Correlational data are still correlational, and no computer program can take account of variables that are not in the analysis. Causal relations can only be established through patient, painstaking attention to all the relevant variables, and should involve active manipulation as a final confirmation

ACKNOWLEDGMENT

I would like to thank Dr. Nina Schooler for many helpful suggestions which substantially improved this chapter.

Table 26 Identification examples.

(i) Consider three variables, x_1, x_2, and x_3, with correlation matrix R given by

$$\mathbf{R} = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array} \begin{bmatrix} x_1 & x_2 & x_3 \\ 1.00 & & \\ 0.83 & 1.00 & \\ 0.78 & 0.67 & 1.00 \end{bmatrix}$$

(ii) Suppose we are interested in fitting a single-factor model, that is

$$\begin{aligned} x_1 &= \lambda_1 f + e_1 \\ x_2 &= \lambda_2 f + e_2 \\ x_3 &= \lambda_3 f + e_3 \end{aligned}$$

(iii) There are seven parameters to be estimated, namely

$$\lambda_1,\ \lambda_2,\ \lambda_3,\ \text{var}(f),\ \text{var}(e_1),\ \text{var}(e_2),\ \text{var}(e_3)$$

(iv) There are, however, only six statistics for use in parameter estimation:

$$\text{var}(x_1),\ \text{var}(x_2),\ \text{var}(x_3),\ \text{corr}(x_1, x_2),\ \text{corr}(x_1, x_3),\ \text{corr}\ (x_2, x_3)$$

(v) Consequently, the model is underidentified

(vi) If var(f) is set equal to one then the model is just identified—there are exactly the same number of parameters to estimate as there are informative sample statistics

(vii) Equating observed and expected variances and correlations will give the required estimates:

$$\begin{aligned} \hat{\lambda}_1\hat{\lambda}_2 &= 0.83 \\ \hat{\lambda}_1\hat{\lambda}_3 &= 0.78 \\ \hat{\lambda}_2\hat{\lambda}_3 &= 0.67 \\ \widehat{\text{var}}(e_1) &= 1.0 - \hat{\lambda}_1^2 \\ \widehat{\text{var}}(e_2) &= 1.0 - \hat{\lambda}_2^2 \\ \widehat{\text{var}}(e_3) &= 1.0 - \hat{\lambda}_3^2 \end{aligned}$$

(where "hats" indicate estimates)

(ix) Solving these equations leads to the estimates

$$\hat{\lambda}_1 = 0.99,\ \hat{\lambda}_2 = 0.84,\ \hat{\lambda}_3 = 0.79,\ \text{var}(e_1) = 0.02,\ \text{var}(e_2) = 0.30,\ \text{var}(e_3) = 0.38$$

(x) Now consider an analogous measurement model with four observed variables and again set var(f) = 1 (this model is now the one used on the string data and the examination scores data)

(xi) Equating observed and expected variances and correlations in this case will lead to more than a single unique estimate for some of the parameters. The model is overidentified and represents a genuinely more parsimonious description of the structure of the data. Here a better strategy for estimation is clearly needed (see Table 9)

3.13.7 APPENDIX

Explanatory factor analysis methods such as principal factor analysis and maximum likelihood factor analysis are available in all major packages, for example, SPSS, SAS, MINITAB, and SYSSTAT. Options that need to be selected include (i) method of factor analysis, (ii) method of estimating initial communalities, and (iii) method of rotation.

For many examples the solutions given by a different combination of the available options will be very similar. In most cases the available software makes the derivation of the results of a factor analysis very simple; interpretation of these results is often, however, a different matter.

Confirmatory factor analysis and structural equation modeling are generally undertaken using either the LISREL or EQS package. The former is described in Jöreskog and Sörborm (1993) and the latter in Dunn et al. (1993). Using either piece of software requires some degree of understanding of how to translate proposed models into equations that specify the model explicitly (see Tables 10 and 19). One of the problems of structural equation modeling conveniently ignored in the text is that of model identification which refers to the degree to which there is a sufficient number of equations to solve for each of the parameters to be estimated. Models can be (i) underidentified (too few equations), (ii) just identified (no degrees of freedom remain for testing the fit of the model—the model will fit perfectly but will not give a more parsimonious description of the data than is provided by the observed correlations or covariances), or (iii) overidentified (more equations than parameters—the fit of model can be tested). Table 26 illustrates these different situations with a number of simple models, and a full discussion of identification is given in Dunn et al. (1993). In many cases the identification status of a complex model (and, at times, relatively simple models) is very difficult to ascertain *a priori.* Nonidentified models will cause problems for both EQS and LISREL and users need to be very wary of results which are accompanied by messages that point out that a parameter is, for example, at its lower or upper bound (i.e., a variance is zero or a correlation is one) or that some parameters are linearly dependent on others.

3.13.8 REFERENCES

Bentler, P. M. (1995). *EQS Structural Equations Program Manual.* Encino, CA: Multivariate Software.

Bookstein, F. L. (1982). Panel discussion—modeling method. In K. Jöreskog & H. Wold (Eds.), *Systems under indirect observation: Causality, structure and prediction,* (pp. 317–322). Amsterdam: North-Holland.

Caslyn, R. J. & Kenny, D. A. (1977). Self-concept of ability and perceived evaluation of others. Cause or effect of academic achievement? *Journal of Educational Psychology, 69,* 136–145.

Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research, 18,* 115–126.

Cudeck, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin, 105,* 317–327.

deLeeuw, J. (1985). Book review. *Psychometrika, 50,* 371–375.

Dunn, G., Everitt, B. S., & Pickles, A. (1993). *Modelling covariances and latent variables using EQS.* London: Chapman and Hall.

Everitt, B. S. (1984). *An introduction to latent variable models.* London: Chapman and Hall.

Everitt, B. S. (1987). *An introduction to optimization methods and their application in statistics.* London: Chapman and Hall.

Everitt, B. S. (1996). *Making sense of statistics in psychology.* Oxford: Oxford University Press.

Everitt, B. S. & Dunn, G. (1991). *Applied multivariate data analysis.* London: Arnold.

Hines, M., Chiu, L., McAdams, L. A., Bentler, P. A., & Lipcamon, J. (1992). Cognition and the corpus callosum: Verbal fluency, visuospatial ability and language lateralization related to midsagittal surface areas of callosal subregions. *Behavioral Neuroscience, 106,* 3–14.

Huba, G. J., Wingard, J. A., & Bentler, P. M. (1981). A comparison of two latent variable causal models for adolescent drug use. *Journal of Personality and Social Psychology, 40,* 180–193.

Hittner, J. B. (1995). Factorial validity and equivalence of the alcohol expectancy questionnaire tension-reduction subscale across gender and drinking frequency. *Journal of Clinical Psychology, 51,* 563–576.

Jöreskog, K. & Sörbom, D. (1993). *Lisrel 8 structural equation modeling with the simplis command language.* Hillsdale, NJ: Erlbaum.

Lakatos, I. (1977). *The methodology of scientific research programmes.* Cambridge, UK: Cambridge University Press.

Lenk, P. J. (1986). Book review. *Journal of the American Statistical Association, 81,* 1123–1124.

Newcomb, M. D. & Bentler, P. M. (1983). Dimensions of subjective female orgasmic responsiveness. *Journal of Personality and Social Psychology, 44,* 862–873.

Osborne, R. T. & Suddick, D. E. (1972). A longitudinal investigation of the intellectual differentiation hypothesis. *Journal of Genetic Psychology, 121,* 83–89.

Osman, A., Barrios, F. X., Kopper, B., Osman, J. R., Grittman, L., Troutman, J. A., & Panak, W. J. (1995). The pain behavior check list (PBCL); psychometric properties in a college sample. *Journal of Clinical Psychology, 51,* 775–782.

Skevington, S. M. (1990). A standardised scale to measure beliefs about controlling pain: A preliminary study. *Psychology and Health, 4,* 221–232.

Taylor, S. & Rachman, S. J. (1994). Stimulus estimation and the overprediction of fear. *British Journal of Clinical Psychology, 33,* 173–181.

Thurstone, L. L. (1947). *Multiple factor analysis.* Chicago: University of Chicago Press.

Wheaton, B., Muthen, B., Alwin, D., & Summers, G. (1977). Assessing reliability and stability in panel models. In D. R. Heise (Ed.), *Sociological methodology* (pp. 84–136). San Francisco: Jossey-Bass.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

3.14

The Shift from Significance Testing to Effect Size Estimation

MICHAEL BORENSTEIN
Hillside Hospital, Glen Oaks, NY, USA

3.14.1 INTRODUCTION 314

3.14.2 THE SIGNIFICANCE TEST 314
  3.14.2.1 The Logic of the Significance Test 314
  3.14.2.2 Errors of Inference 315
    3.14.2.2.1 Type I error 315
    3.14.2.2.2 Type II error 315

3.14.3 POWER ANALYSIS 316
  3.14.3.1 Role of Effect Size in Power Analysis 316
    3.14.3.1.1 The effect size used in power analysis is not necessarily the population effect size 317
    3.14.3.1.2 Conventions for effect size 317
  3.14.3.2 Role of Alpha in Power Analysis 318
  3.14.3.3 Role of Tails in Power Analysis 319
  3.14.3.4 Role of Sample Size in Power Analysis 319
  3.14.3.5 Computing Power in Power Analysis 320
  3.14.3.6 The Null Hypothesis vs. the Nil Hypothesis 320
  3.14.3.7 Retrospective Power Analysis 320
  3.14.3.8 Ethical Issues in Power Analysis 321
  3.14.3.9 Power Analysis in Perspective 321
  3.14.3.10 Some Real World Issues in Power Analysis 322
  3.14.3.11 Computer Programs for Power Analysis 322

3.14.4 PROBLEMS WITH THE SIGNIFICANCE TEST 324
  3.14.4.1 Significance Tests Address the Wrong Question 325
  3.14.4.2 The Null Hypothesis is not a Viable Hypothesis 325
  3.14.4.3 The Gratuitous Use of Significance Tests 326
  3.14.4.4 p-*Values: Responding with a Misdirected Answer* 327
    3.14.4.4.1 Misinterpreting the significant p-*value* 327
    3.14.4.4.2 Misinterpreting the nonsignificant p-*value* 328
  3.14.4.5 Summary 331

3.14.5 FOCUSED SIGNIFICANCE TESTS—A DIGRESSION 331

3.14.6 EFFECT SIZE AND CONFIDENCE INTERVALS 332
  3.14.6.1 The Key Advantage of this Approach 332
  3.14.6.2 Effect Size and Confidence Intervals can be Computed for Any Type of Study 333
  3.14.6.3 Factors Affecting the Confidence Interval Width 333
  3.14.6.4 Information Conveyed by the Effect Size and Confidence Interval 334

3.14.7 THE RELATIONSHIP BETWEEN SIGNIFICANCE TESTING AND EFFECT SIZE ESTIMATION 334
  3.14.7.1 Confidence Intervals as a Surrogate for the Significance Test 335
    3.14.7.1.1 Confidence intervals should serve only as a surrogate for the significance test 336
    3.14.7.1.2 Confidence intervals can optionally be used as a surrogate for the significance test 336
    3.14.7.1.3 Confidence intervals should never be used as a surrogate for the significance test 337

3.14.7.2 Should the Significance Test be Banned? 337
3.14.7.3 Should the Analysis of the Single Study be Banned? 339
3.14.7.4 Looking Ahead 340
3.14.8 USE OF CONFIDENCE INTERVALS IN STUDIES WHOSE PURPOSE IS TO PROVE THE NULL 340
3.14.9 COMPUTATIONAL ISSUES IN CONFIDENCE INTERVALS 341
3.14.10 STUDY PLANNING: PRECISION ANALYSIS VS. POWER ANALYSIS 341
3.14.10.1 Planning for Precision 342
3.14.11 META-ANALYSIS 344
3.14.12 CONCLUSIONS 345
3.14.13 REFERENCES 346

3.14.1 INTRODUCTION

A statistical wit once remarked that researchers often pose the wrong question and then proceed to answer that question incorrectly. Statistical analyses in psychological research traditionally have taken the form of significance tests. The near-universal use of significance tests has made them the *de facto* standard of proof in research and the logic that underlies significance tests has played an important role in shaping the development of psychological theory. In fact, though, significance tests are generally not appropriate for the kinds of questions that are addressed in psychological research. First, significance tests address the wrong question. Researchers are concerned with clinical significance (i.e., the magnitude of the effect) but significance tests address only statistical significance (whether or not the effect is zero). Second, the significance test, by focusing attention on the p-value, lends itself to mistakes of interpretation. Significant p-values are assumed to reflect clinically important effect sizes, while in fact a significant p-value may reflect a large sample size rather than a large effect size. Similarly, nonsignificant p-values are taken as evidence that the treatment has no impact, though this conclusion is almost always incorrect.

To address these issues researchers have been moving away from significance tests and toward effect size estimation. Rather than report that "$p=0.01$," the researcher would report that "the treatment improved the response rate by 20 percentage points (95% confidence interval of 0.15–0.25)." While there is widespread consensus that the shift toward effect size estimation is appropriate, there is considerable disagreement over the nature of the proposed shift. Some feel that the logic of the significance test should be retained, with the shift to the use of confidence intervals being one of format rather than substance. From this perspective, the confidence intervals serve as a surrogate for the significance test. Others argue that we need to rethink the role of the single study as an element in the research process, and that any vestige of the significance test should be excluded. From this perspective, confidence intervals serve as an index of precision only.

We also consider the implications of this paradigm for study planning. The discussion of significance tests is followed by a section on power analysis, which is used in study planning to ensure that the study will yield a statistically significant result. The discussion of effect size estimation is followed by a discussion of precision analysis, which may be used in study planning to ensure that the study will yield a precise estimate of the treatment effect.

3.14.2 THE SIGNIFICANCE TEST

3.14.2.1 The Logic of the Significance Test

Research trials are always carried out on a sample of finite size. The results obtained in the sample are assumed to be representative of the larger population but because of random sampling error will rarely, if ever, mirror that population exactly.

For example, consider a study whose goal is to determine whether or not a new drug being tested with acute schizophrenics yields a higher response rate than the current treatment. Assume that the drug really is effective, and (in the population) increases the response rate by 20 percentage points. The mean effect over an infinite number of samples would be 20 points, but the effect observed in any single sample would almost invariably fall somewhat below or above 20 points. Similarly, if the drug really has zero effect in the population then the mean effect over an infinite number of samples would be zero points, but the effect observed in any single sample would likely fall somewhat below zero (the response rate being lower for new drug), or above zero.

Therefore, if we run a study and observe a difference in response rates for the two groups, we need to determine whether or not the sample may have been drawn from the second population above (where the treatment effect is zero). To this end we pose the null hypothesis that the two treatments are equally effective. If the null hypothesis is true, then the effect observed in our sample, if not zero, should fall within a "reasonable" distance of zero. If our sample were to yield a rate difference that was "compelling," in the sense that it was not "reasonably close to zero," we would conclude that the assumption of equal response rates is not viable.

Our ability to conclude that the observed effect is "compelling" will depend on three factors, each of them intuitive. The first is the size of the effect. An observed difference of 40 points in response rates is more compelling than an observed difference of 20 points. The second is the size of the sample. An observed difference of 20 points may be discounted if based on a sample of 10 cases per group since a single aberrantly responding patient could pull the response rate by 10 percentage points in either direction. The same size effect would be more convincing in a sample of 40 patients per group, and would probably be seen as definitive in a study that employed 100 patients per group. The third element is the stringency of the criteria we require as evidence that a difference exists. If the treatment carried little risk of harm and was inexpensive, we might be willing to declare it "effective" even if the evidence was relatively weak. On the other hand, if it carried a nontrivial risk of harm or a high cost we might require that it meet a more stringent criterion.

This logic (how large is the effect, how large is the sample, how certain do I want to be) is formalized in the test of significance that is routinely applied to studies. Significance tests vary from one statistical procedure to the next but are all variations on the theme

$$\text{Test statistic} = \frac{\text{Observed difference}}{\text{Dispersion of the difference}}$$

The numerator in this equation is the observed difference, for example, a rate difference of 10 points or 20 points between groups. The denominator reflects the sample-to-sample dispersion expected in the numerator, and incorporates information about the sample size (when the sample is large, the expected dispersion is small). The test statistic is computed as the ratio of these two values. As such, it will be high when the observed effect is large and/or the sample size is large. In a sense, the test statistic serves as a kind of signal-to-noise ratio: if the population effect is zero we expect the sample effect to vary within some range of zero. The numerator in the equation reflects the observed effect and the denominator defines the expected range in which sample effects will normally fall. The test statistic tells us whether or not the sample effect falls outside the normal range.

The test statistic is compared with a table of values showing the expected range of the test statistic under the assumption that the null hypothesis is true. If our test statistic falls within this range, we cannot reject the null. If our test statistic falls outside this range, however, we would conclude that the null hypothesis is most likely false, and the treatment does have an effect. We need to decide what level of evidence we will require prior to rejecting the null, and this threshold is known as alpha. For example, if alpha is set at 0.05, then the study result will be deemed significant if, on the assumption that the null is true, the study effect (or larger) would be observed in only 5% of cases.

3.14.2.2 Errors of Inference

The significance test carries the potential for two types of error: We may reject as false a null hypothesis that in fact is true (type I error), or we may accept as true a null hypothesis that in fact is false (type II error).

3.14.2.2.1 Type I error

As noted earlier, study results would be declared "significant" if the sample effect was "compelling." The type I error rate is denoted by alpha. With alpha set at 0.05, over an infinite number of studies where the null is true, our sample will nevertheless meet the criterion and be declared significant in 5% of the studies. This situation—where the null is true but we decide in error that the treatment is effective—is referred to as a type I error.

This point is repeated for emphasis. There is a general perception that 5% of studies will result in a type I error, but this is not correct. When the null hypothesis is true 5% of studies will result in a type I error. However, if the treatment really is effective, then by definition a type I error is not possible and the type I error rate is zero.

3.14.2.2.2 Type II error

To meet the criterion for statistical significance a study must yield an effect size that is large enough to establish significance, given the sample size and the criterion alpha. Even if the treatment is effective, and even if the size of the effect is substantial (in the population and/or in the sample), there is no guarantee that the study will yield a statistically significant result. The effect observed in the sample may be smaller than the population effect (in half of all unbiased studies, it will be), which may yield a nonsignificant result. Indeed, even if the sample effect is as large (or larger) than the population effect, unless the sample size is adequate the results may not meet the criterion for statistical significance. The type II error rate is denoted by beta. If the treatment really is effective and beta is 0.20, then 20% of all studies will yield nonsignificant results.

3.14.3 POWER ANALYSIS

The likelihood that a study will yield a statistically significant effect is the study's power and the process of designing a study to ensure an appropriate level of power is referred to as a power analysis. If the purpose of the study is to test the null hypothesis, then a power analysis is a critical element in study planning, to ensure that the study will be able to meet this goal.

The statistical significance computed subsequent to the study is a function of three elements—effect size, sample size, and alpha. It follows that power, which reflects the likelihood that the study will yield statistical significance, is determined by the same three elements. Specifically, the larger the effect size used in the power analysis, the larger the sample size, and/or the more liberal the criterion required for significance (alpha), the higher the expectation that the study will yield a statistically significant effect. These three factors, together with power, form a closed system—once any three are established, the fourth is completely determined. The goal of a power analysis is to find an appropriate balance among these factors by taking into account the substantive goals of the study and the resources available to the researcher.

3.14.3.1 Role of Effect Size in Power Analysis

Power analysis gives power for a specific effect size. For example, the researcher might report "If the treatment increases the response rate by 20 percentage points the study will have power of 80% to yield a significant effect." For the same sample size and alpha, if the treatment effect is less than 20 points then power will be less than 80%. If the treatment effect exceeds 20 points, then power will exceed 80%.

Since our computation of power ensures, to some likelihood, that the study will succeed in rejecting the null given a specific effect size, it follows that the effect size used in the power analysis should represent the smallest effect that would be of clinical or substantive significance. In clinical trials, for example, the selection of an effect size might take account of the following factors:

(i) The severity of the illness being treated. If we are testing a treatment that will reduce the likelihood of relapse among schizophrenics, we might decide that a reduction of as little as 10% is clinically important since the implications of a relapse can be relatively severe. By contrast, if we are testing a treatment that was expected to reduce the likelihood of an anxiety attack, a 10% reduction might be of little clinical import since the attacks are transitory. In this case, we might decide that only a reduction of 20% or more is clinically important.

(ii) The availability of alternate treatments. If no treatments exist for a particular condition, then a treatment that was effective in even a small proportion of patients might be clinically important. By contrast, if treatments already exist for the condition, then a new treatment might need to surpass these to be considered clinically important. This might mean that the new treatment was effective in a larger proportion of affected persons, or had a larger impact for individual patients.

(iii) Treatment cost and side effects. In deciding what constitutes a clinically important effect we might want to take account of such issues as treatment costs and side effects. A drug that carried a risk of serious side effects or that carried a high price might be considered clinically useful only if it was effective in a substantially higher proportion of cases than available treatments, or was being used to treat a severe condition.

The factors that are taken into account for identifying a clinically important effect will of course vary from one study to the next and the three items mentioned above are intended only as examples. The general point is that the selection of an effect size should be made on the basis of substantive issues and that these will vary from one study to the next.

A study that has adequate power to detect a relatively large effect will not have adequate power to detect a small effect. By contrast, a study that has adequate power to detect a relatively small effect will, of course, have more than enough power to detect moderate or large effects. While one might therefore be tempted to set the "clinically important effect" at a small value to ensure high power for even a small effect, this determination cannot be made in isolation. Small effects will require a larger investment of resources than large effects. The selection of an effect size reflects the need for balance between the size of the effect that we can detect, and the resources available for the study. Figure 1 shows power as a function of sample size for three levels of effect size (assuming alpha, two-tailed, is set at 0.05). For the smallest effect (30% vs. 40%), we would need a sample of 376 per group to yield power of 80%. For the intermediate effect (30% vs. 50%), we would need a sample of 103 per group to yield this level of power. For the highest effect size (30% vs. 60%), we would need a sample of 49 per group to yield power of 80%. In this example we may decide that it would make sense to enroll 103 patients per group to detect the intermediate effect but inappropriate to enroll 376 patients per group to detect the smallest effect.

3.14.3.1.1 The effect size used in power analysis is not necessarily the population effect size

Researchers often assume that the effect size used in a power analysis is the "true" (population) effect size. In fact, though, the "true" effect size is not known. While the effect size in the power analysis is assumed to reflect the population effect size for the purpose of calculations, the power analysis is more appropriately expressed as "If the true effect is 20 percentage points power would be ... " rather than "The true effect is 20 percentage points, and therefore power is ... "

This distinction is an important one. Researchers sometimes assume that a power analysis cannot be performed in the absence of pilot data. In fact, it is usually possible to perform a power analysis based entirely on a logical assessment of what constitutes a clinically (or theoretically) important effect. Indeed, while the effect observed in prior studies might help to provide an estimate of the true effect it is not likely to be the true effect in the population—if we knew that the effect size in these studies was accurate, there would be no need to run the new study.

Since the effect size used in power analysis is not the "true" population value, the researcher may elect to present a range of power estimates. For example (assuming $N = 103$ per group and alpha = 0.05, two-tailed), "The study will have power of 80% to detect a treatment effect of 20 points (30% vs. 50%), and power of 99% to detect a treatment effect of 30 points (30% vs. 60%)."

The nature of the effect size will vary from one statistical procedure to the next. In all cases, however, it serves the same function of providing a pure index of the effect, that is, an index that focuses exclusively on the effect, independent of the sample size, and that is not affected by the metric in which the effect is measured.

3.14.3.1.2 Conventions for effect size

Cohen (1988, 1992) has suggested conventional values for "small," "medium," and "large" effects in the social sciences. For tests of a mean difference between two groups the effect size index is *d*, the standardized mean difference (i.e., mean difference divided by the

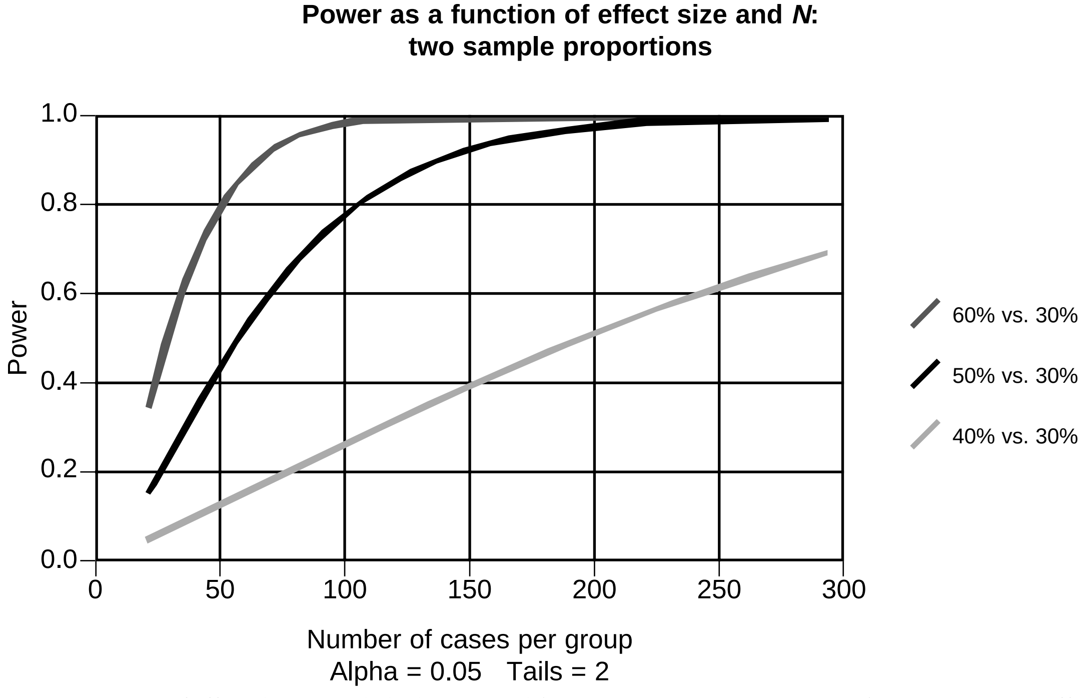

Figure 1 Impact of effect size on statistical power. With 100 cases per group, power to detect a treatment effect of 10 percentage points (response rates of 40% vs. 30%) is about 0.25, power to detect a treatment effect of 20 points is about 0.80, and power to detect a treatment effect of 30 points is about 0.99.

common within-group standard deviation). A "small" effect is given as $d = 0.20$ (i.e., 20% of the common within-group standard deviation); a medium effect as $d = 0.50$; and a large effect as $d = 0.80$. For tests of a difference between proportions a small effect is given as proportions of 40% vs. 50%; a medium effect corresponds to proportions of 40% vs. 65%; and a large effect corresponds to proportions of 40% vs. 78%. Note that the effect size for proportions cannot be specified simply by giving the difference between the two proportions, but requires that the absolute proportions be specified. For example, the effect size of 10% vs. 20% is larger than the effect size of 40% vs. 50%, despite the fact that the difference is 10 points in either case. For tests of a single correlation, the effect size is the correlation itself. A small effect is given as a correlation of 0.10, a medium effect is a correlation of 0.30, and a large effect is a correlation of 0.50. Cohen also provides conventions for other tests including analysis of variance and multiple regression. These conventions are also included in some computer programs (Borenstein & Cohen, 1988; Borenstein, Cohen, Rothstein, Pollack, & Kane, 1990, 1992; Borenstein, Rothstein, & Cohen, 1997; Rothstein, Borenstein, Cohen, & Pollack, 1990).

Cohen himself cautions that these conventions should not be used routinely, since it is preferable to select an effect size based on the substantive issues involved in the specific study (as above). Nevertheless, these conventions do serve two functions. In all cases the researcher may want to use these values as a kind of reality-check, to ensure that the values specified make sense relative to these anchors. Additionally, in cases where the researcher has difficulty deriving any kind of effect size index, they may elect to fall back on these conventions (see also Kraemer & Thiemann, 1987).

3.14.3.2 Role of Alpha in Power Analysis

The significance test yields a *p*-value that reflects the likelihood of obtaining an effect as large (or larger) than the observed effect, under the assumption that the null hypothesis is true. For example, a *p*-value of 0.02 means that, assuming that the treatment has no effect, and given the sample size, an effect as large as the observed effect would be seen in only 2% of studies. The *p*-value obtained in the study is evaluated against the criterion alpha. If alpha is set at 0.05, then a *p*-value of 0.05 or lower is required to reject the null hypothesis and establish statistical significance.

If our only concern in study design were to prevent a type I error, it would make sense to set alpha as conservatively as possible (e.g., at 0.001). However, alpha does not operate in isolation. If we select a more stringent criterion for alpha, then our ability to meet this criterion

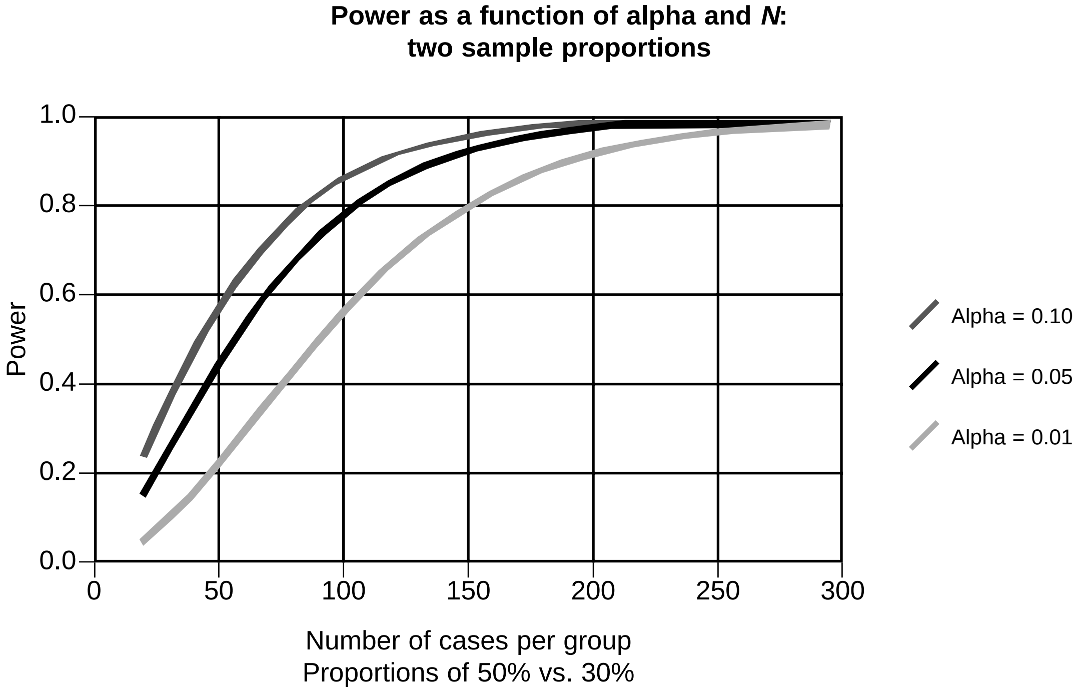

Figure 2 Impact of alpha on statistical power. With 100 cases per group and alpha set at 0.01, power is less than 0.60; with alpha at 0.05, power is about 0.80; and with alpha at 0.10, power is about 0.85. These computations are based on response rates of 50% vs. 30%, corresponding to the center line in Figure 1.

is reduced. By moving alpha from (say) 0.10 toward 0.01, we reduce the likelihood of a type I error but increase the likelihood of a type II error. (More accurately, and to emphasize a point made earlier: We reduce the likelihood of a type I error if the null is true, but increase the likelihood of a type II error if the null is false.)

Figure 2 shows power as a function of sample size for three levels of alpha (assuming an effect size of 30% vs. 50%, which is the intermediate effect size in the previous figure). With a sample size of 100 per group, with alpha set at 0.10 power exceeds 85%; with alpha set at 0.05 power is about 80%; and with alpha set at 0.01, power is under 60%.

Traditionally, researchers in some fields have accepted the notion that alpha should be set at 0.05 and power at 80% (corresponding to a beta of 0.20). This notion is implicitly based on the assumption that a type I error is four times as harmful as a type II error (the ratio of alpha to beta is 0.05–0.20), which notion has no basis in fact. Rather, it should fall to the researcher to strike a balance between alpha and beta as befits the issues at hand. For example, if the study will be used to screen a new drug for further testing we might want to set alpha at 0.20 and power at 95%, to ensure that a potentially useful drug is not overlooked. On the other hand, if we were working with a drug that carried the risk of side effects and the study goal was to obtain Federal Drug Administration (FDA) approval for use, we might want to set alpha at 0.05 while keeping power at 95%.

3.14.3.3 Role of Tails in Power Analysis

The significance test is always defined as either one- or two-tailed. A two-tailed test is a test that will be interpreted if the effect meets the criterion for significance and falls in either direction. As such, it is appropriate for the vast majority of research studies. A one-tailed test is a test that will be interpreted only if the effect meets the criterion for significance and falls in the expected direction (i.e., the treatment improves the cure rate).

A one-tailed test is appropriate only if an effect in the unexpected direction would be functionally equivalent to no effect. For example, assume that the treatment we are using for depression is relatively inexpensive and carries a minimal risk of side effects. We will be testing a new treatment which is more expensive but carries the potential for a greater effect. The possible conclusions are that (i) the old treatment is better, (ii) there is no difference, or (iii) the new treatment is better. For our purposes, however, conclusions (i) and (ii) are functionally equivalent since either would lead us to retain the standard treatment. In this case, a one-tailed test, whose only goal is to test whether or not conclusion (iii) is true, might be an appropriate choice.

Note that a one-tailed test should be used only in a study in which, as in this example, an effect in the reverse direction is, for all intents and purposes, identical to "no effect." It is not appropriate to use a one-tailed test merely because one is able to specify the expected direction of the effect prior to running the study. In psychological research, for example, we typically expect that the new procedure will increase, rather than decrease, the cure rate. Nevertheless, a finding that it decreases the cure rate would be important, since it would demonstrate a possible flaw in the underlying theory. Even in the example cited, one would want to be certain that a profound effect in the reverse direction could safely be ignored—under a one-tailed test, it cannot be interpreted. In behavioral research, the use of a one-tailed test can be justified only rarely.

For a given effect size, sample size, and alpha, a one-tailed test is more powerful than a two-tailed test (a one-tailed test with alpha set at 0.05 has the same power as a two-tailed test with alpha set at 0.10). However, the number of tails should be set based on the substantive issue ("Will an effect in the reverse direction be meaningful?"). In general, it would not be appropriate to run a test as one-tailed rather than two-tailed as a means of increasing power. (Note also that power is higher for the one-tailed test only under the assumption that the observed effect falls in the expected direction. When the test is one-tailed, the power for an effect in the reverse direction is zero by definition).

3.14.3.4 Role of Sample Size in Power Analysis

For any given effect size and alpha, increasing the sample size will increase the power. As is true of effect size and alpha, sample size cannot be viewed in isolation but rather as one element in a complex balancing act. In some studies it might be important to detect even a small effect while maintaining high power. In this case it might be appropriate to enroll many thousands of patients (as was done in the "Physicians" study that found a relationship between aspirin use and cardiovascular events).

Typically, though, the number of available cases is limited. The researcher might need to find the largest *N* that can be enrolled, and work backwards from there to find an appropriate balance between alpha and beta. They may need to forgo the possibility of finding a small effect, and acknowledge that power will be adequate for a large effect only.

For studies that involve two groups, power is generally maximized when the subjects are divided evenly between the two groups. When the number of cases in the two groups is uneven the "effective *N*" for computing power falls much closer to the smaller sample size than the larger one.

There is one exception to the rule that an increase in sample size will always yield an increase in power. When we are working with exact formulas for discrete distributions, such as the binomial test for a single proportion or the Fisher exact test for a crosstabulation, it is possible that a modest increase in sample size will serve to lower power (Borenstein et al., 1997).

3.14.3.5 Computing Power in Power Analysis

Power is the fourth element in this closed system—for given effect size, alpha, and sample size, power is completely determined. A convention exists that power should be set at 80% but this convention has no clear *a priori* basis. The appropriate level of power should be decided on a case-by-case basis, taking into account the potential harm attendant on a type I error, the determination of a clinically important effect, the available sample size, as well as the importance of identifying a small, medium, or large effect.

While power is a function of three elements (effect size, sample size, and alpha), as a practical matter the effect size tends to dominate the calculation of power. Sample size plays a secondary role, and the impact of alpha is relatively modest. For example, with response rates of 30% vs. 50%, a sample size of 100 per group, and alpha (two-tailed) set at 0.05, power would be 83%. Consider, then, the impact of changes to these factors. If we elect to work with an effect size of 25 percentage points rather than 20 percentage points (i.e., 55% vs. 30% rather than 50% vs. 30%) power would be increased from 83% to 95%. To achieve the same 12-point increase in power by manipulating the sample size we would have to increase the sample size from the initial value of 100 per group to 160 per group. To achieve the same increase in power by selecting a more liberal value for alpha we would need to move alpha from 0.05 to 0.20.

Because the effect size plays such an important role in the computation of power, it is imperative for the researcher to use an appropriate effect size in the computations. If we use an effect size that is too large, then the study will not have adequate power to detect a more modest effect. If we use an effect size that is too low, then we might increase the sample size in an effort to reach acceptable levels of power, thus inappropriately putting patients at risk. Rather, the effect size used in the power analysis should be selected carefully, based on the kinds of substantive issues outlined earlier.

3.14.3.6 The Null Hypothesis vs. the Nil Hypothesis

Power analysis focuses on the study's potential for rejecting the null hypothesis. In most cases the null hypothesis is the null hypothesis of no effect, or the "nil" hypothesis. In some studies, however, the researcher might want to test another null hypothesis. For example, rather than testing the null hypothesis that a correlation coefficient is zero, a researcher might want to test the null hypothesis than the correlation is 0.80 (i.e., to demonstrate that the correlation exceeds this value). Intuitively, it is easier to show that an observed correlation exceeds zero than to show that it exceeds 0.80. In the latter case the effect size is smaller and a substantially larger sample would be required to ensure adequate power.

3.14.3.7 Retrospective Power Analysis

One occasionally sees a retrospective power analysis of the following form. A researcher completes a study, and reports that the effect size in the study was, for example, a standardized mean difference of $d=0.2$. The researcher then goes on to report that, based on this effect size, the study actually had power of only, for example, 40%. However well intentioned, this type of retrospective power analysis is generally misguided for several reasons. First, as noted earlier, the role of effect size in a power analysis is not "The effect size is $d=0.4$ and therefore power is 80%" but rather "If the effect size is $d=0.4$, then power would be 80%." The effect size $d=0.4$ is supposed to represent the smallest effect that would be important to detect. As such, it is not affected by evidence that the actual size of the effect might be larger or smaller than this value. Second, even if one wanted to base power on the "true" effect, the fact remains that the value observed in any single study is not likely to be the true effect. Indeed, if we had confidence that observed effect size was the true effect we would have no need to run another study to pinpoint the effect size, much less to test the hypothesis that it was zero.

While this application of a retrospective power analysis is inappropriate, there are two related applications where it may be useful. One is the use of study data to identify a base rate.

Assume, for example, that our goal is to reduce by 20% the proportion of patients having a panic attack within two days of treatment. Initially, we assume that the base rate for these attacks is 30%, compute power for a difference of 30% vs. 24% (corresponding to a 20% decrease), and set the sample size on this basis. It emerges in the study that the base rate is 20%, which means that we should have computed power for a difference of 20% vs. 16% (corresponding also to a 20% decrease). In this case, one could argue that the original power calculations were based on incorrect assumptions, and that for the *a priori* effect size the study was underpowered. Similarly, if the effect size was computed as a mean difference in raw units (such as SAT [Scholastic Achievement Test] scores) but we underestimated the standard deviation of the scores, one could argue that our initial computations had been in error and that power for the desired effect was too low. A second logical application along these lines is to use the study data to obtain a more informed estimate of the likely effect size by computing the confidence intervals based on the sample. This information could prove useful in the design of a subsequent study.

Retrospective power analyses are usually reported subsequent to a study that failed to yield a significant effect. Occasionally, though, researchers who have obtained a significant result may discover that their study had been under-powered and then wonder if the study results are still valid. Of course, this is a nonissue. Power tells us the likelihood is that the study will yield a significant effect and is used for purposes of planning. If the study is completed, then we know the outcome—either the effect was significant, or it was not. To worry in retrospect that the chances of success were small is akin to the parent whose teenager has taken the car for a ride during a bad rain but returned home safely. There might have been reason for worry while the teenager was still out (and the parent might want to be more careful the next time) but it makes no sense to worry about the outcome of this particular trip after the fact.

3.14.3.8 Ethical Issues in Power Analysis

Some studies involve putting patients at risk. At one extreme, the risk might involve a loss of time spent completing a questionnaire. At the other extreme, the risk might involve the use of an ineffective treatment for a potentially fatal disease. While an extensive discussion of research ethics is outside the scope of this chapter, it must be emphasized that ethical issues should play a central role in power analysis. If a study to test a new drug will have adequate power with a sample of 100 patients, then it would be inappropriate to use a sample of 200 patients since the second 100 are being put at risk unnecessarily. At the same time, if the study requires 200 patients to yield adequate power, it would be inappropriate to use only 100. These 100 patients consent to take part in the study on the assumption that the study is likely to yield useful information. If the study is underpowered, then the 100 patients will have been put at risk for no reason.

Of course, the actual decision-making process is complex. One can argue about whether "adequate" power for the study is 80%, or 90%, or 99%. One can argue about whether power should be set based on an improvement of 10 points, or 20 points, or 30 points. One can argue about the appropriate balance between alpha and beta. The point being made here is that these kinds of issues need to be addressed explicitly as part of the decision making process.

3.14.3.9 Power Analysis in Perspective

To this point we have focused on the details of a power analysis, but it is important also that the entire procedure be seen in an appropriate context. The use of a power analysis in study planning is appropriate only when the study goal is to test the null hypothesis. By contrast, if the study goal is to estimate the magnitude of a treatment effect, then statistical power has no bearing on sample size. Later in this chapter it is argued that most studies should focus on effect size estimation rather than power analysis, and so this becomes a critical point.

More generally, before performing a power analysis or any other procedure to set sample size, it behooves the researcher to ensure that the study should be conducted at all. Chalmers and co-workers have shown that many studies have put patients at risk in attempts to test treatments when, in fact, the data required to answer the question already existed from prior studies. In one famous example (Lau et al., 1992), a series of 33 randomized trials were carried out between 1959 and 1988 to assess the impact of intravenous (iv) streptokinase in preventing mortality subsequent to a myocardial infarction. In fact, the life-saving power of these drugs could have been proven as early as 1973, if the data available at that time had been subjected to a meta-analysis. At that time, a total of eight studies had been completed using a total of 2432 patients. The p-value for a test of the nil was < 0.01 and the effect size (odds ratio) 0.74 with a 95% confidence interval of

0.59–0.92. Subsequent to 1973, researchers ran an additional 25 studies which employed a total of 34 542 additional patients, with approximately half of these assigned to placebo and denied the potential benfits of the treatment. In fairness, it must be noted that meta-analysis was not widely accepted in 1973, and even if the researchers had access to meta-analytic methods, results based on a these procedures might not have been accepted by the medical community. Our intention is not to argue that this specific set of studies was not needed, but to make the point that the study must hold the potential to provide useful information and that power analysis is only one element in this process.

3.14.3.10 Some Real World Issues in Power Analysis

In some ways, the funding process and other constraints ensure that studies will continue to be run with low power. Feinstein (1975) may have identified one part of the problem with his straightforward account of how studies are designed in real life. The statistician (i) computes the maximum number of patients for which funding can be obtained, (ii) finds the effect size required to yield power of 80%, given the sample size, and (iii) develops a rationale to present to the finding agency to show that the effect size is the smallest effect size that would be clinically important to detect. The process identified by Feinstein leads to grant applicants asserting that a clinically important effect would be one that improves the response rate by precisely "22 percentage points."

In fact, the thrust of this approach is not necessarily inappropriate. The selection of an effect size should take into account any number of factors, and it might be appropriate to allocate resources to detect a large effect but not a small one. However, the grant application would be more credible if the applicant acknowledged that the effect size was set in this way, rather than arguing that the smallest effect of clinical importance would be a rate difference of precisely 22 percentage points.

3.14.3.11 Computer Programs for Power Analysis

Power is defined as the proportion of samples that will yield a statistically significant effect, given a specific set of assumptions. The computation of this value is a two-step process. First, we compute the effect size required to yield a significant effect, for the given alpha and sample size. Second, we compute the proportion of samples that will yield an effect this large (or larger), given the effect size in the population.

As a practical matter, computation of power is performed by means of computer programs designed for this purpose. The example that follows is taken from *Power and Precision*, a program developed by Borenstein et al. (1997). The example is for a study in which patients will be assigned at random to either a new treatment or the standard treatment, and we will compare the proportion responding in the two groups.

Computation of power proceeds as follows (Figure 3):

(i) Optionally, enter names for the two groups: "New treatment" and "Standard treatment."

(ii) Enter the proportion responding in either group: 40% for the standard group and 60% for the new treatment.

(iii) Click "Find N." The program shows that a sample of 107 per group will yield power of 80%.

(iv) At this point the researcher may elect to vary the study parameters, for example, to modify the effect size or sample size and see the impact on power. This may be done interactively. Alternatively, the program will automatically generate a table and graph that allows the researcher to simultaneously take account of several factors.

One possible graph is shown in Figure 4. The three lines in the graph represent treatment effects of 25 points, 20 points, and 15 points (specifically, response rates of 65% vs. 40%, 60% vs. 40%, and 55% vs. 40% for the new treatment vs. the standard treatments). The sample size required to yield power of 80% is 70 patients per group for the largest effect, 107 per group for the second effect, and 186 per group for the third.

In this example, taking account of substantive issues and available resources, we decide that we will base the study on the center line. In other words, we decide that a 20-point improvement in response rates is important enough to justify a trial with 107 patients per group. By following the trajectory of the center line we see that 139 patients per group would yield power of 90% for this effect, but decide that we cannot commit the additional resources. By comparing the position of the three lines at "107 per group" we note that if the treatment effect is actually 25 points (65% vs. 40%) rather than 20 points, then the study's power is actually 95% rather than 80%. If the true effect is actually 15 points (55% vs. 40%) then power is actually 54%.

The program will also generate a text report of the computation which serves as an educational tool and may also be copied into

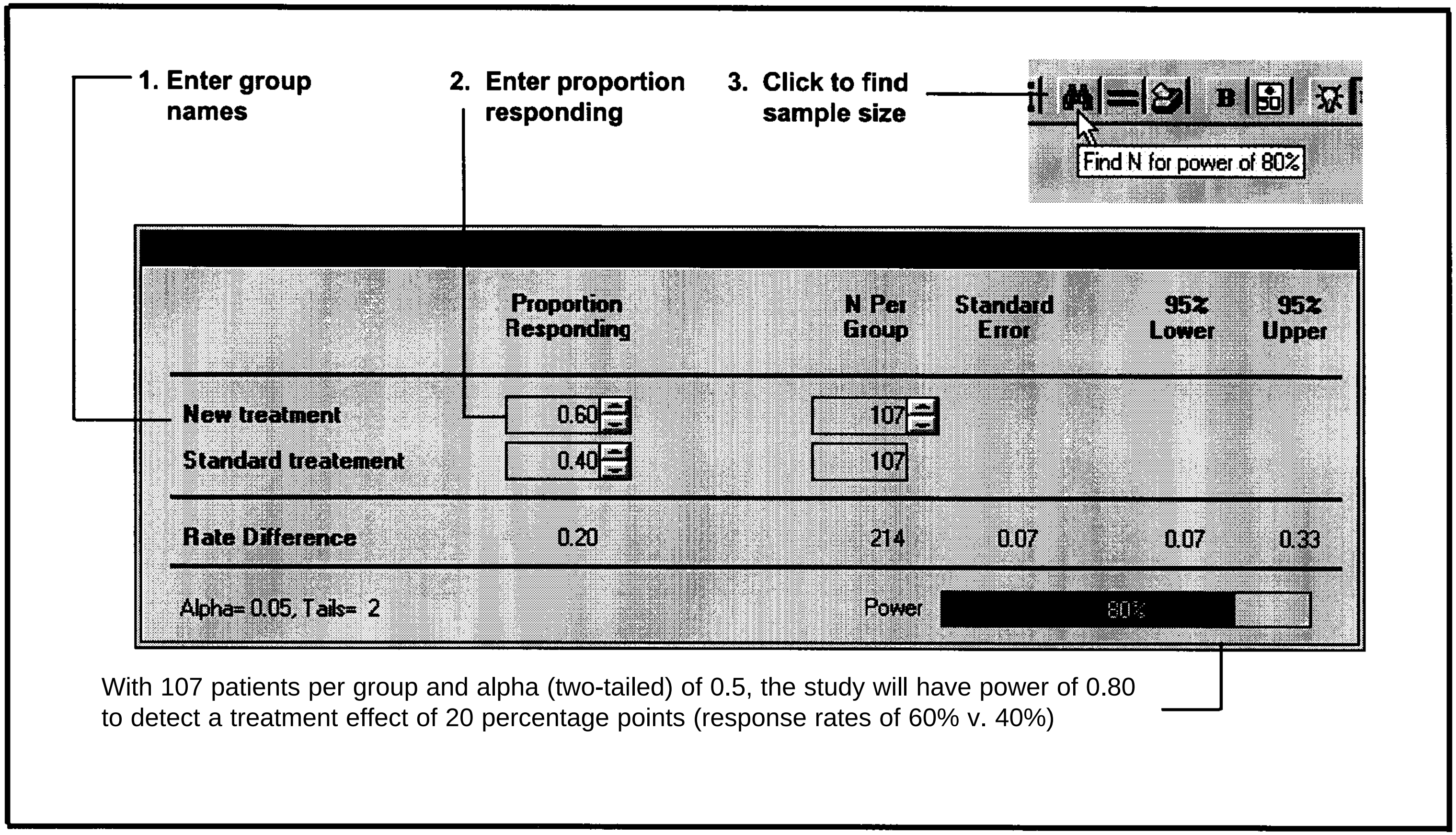

1. Enter group names

2. Enter proportion responding

3. Click to find sample size

	Proportion Responding	N Per Group	Standard Error	95% Lower	95% Upper
New treatment	0.60	107			
Standard treatement	0.40	107			
Rate Difference	0.20	214	0.07	0.07	0.33

Alpha= 0.05, Tails= 2

Power 80%

With 107 patients per group and alpha (two-tailed) of 0.5, the study will have power of 0.80 to detect a treatment effect of 20 percentage points (response rates of 60% v. 40%)

Figure 3 Screen from *Power and Precision* showing computation of power. The researcher enters an effect size (response rates of 60% vs. 40%), alpha (0.05, two-tailed) and clicks "Find N" on the toolbar. The program shows that a sample size of 107 per group will yield a power of 80%. The researcher may modify any of the study parameters, and the program will update power automatically in response.

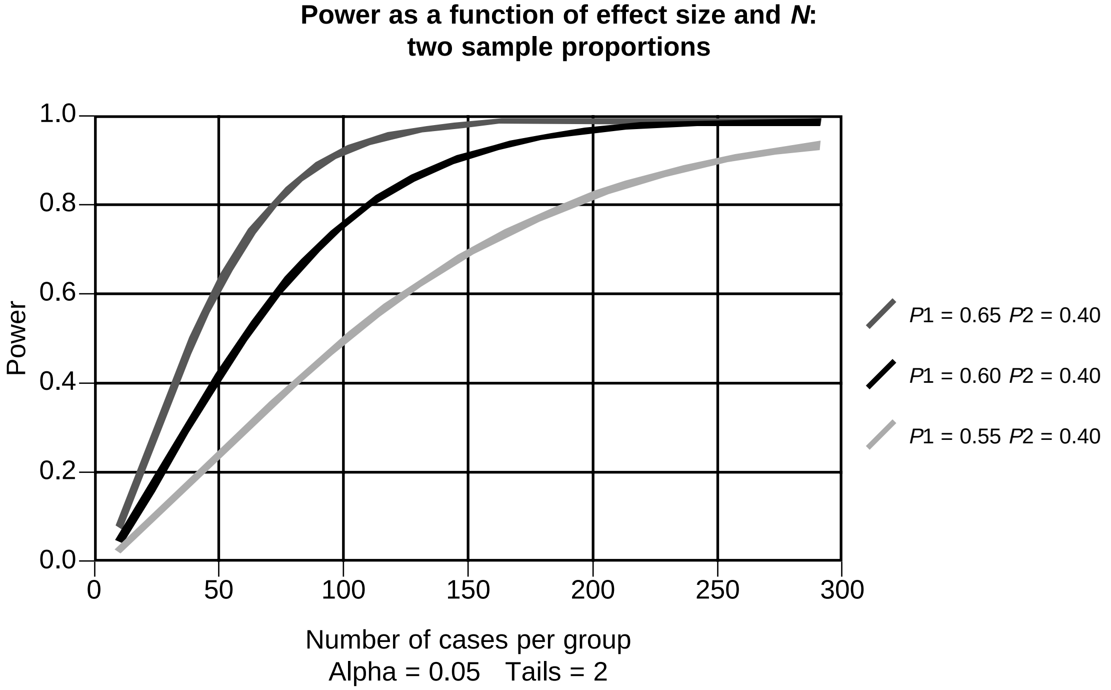

Figure 4 Power as a function of effect size and sample size. This graph, which is an extension of Figure 3, allows the researcher to take account of multiple factors simultaneously while planning the study. We can design the study to detect the largest effect (top line) which would require only about 70 cases per group but would result in low power to detect the more modest effects. We can design the study to detect the middle effect, which would require 107 cases per group but would result in low power to detect the smallest effect. To ensure power for the smallest effect (with the least potential benefit) we would need to enroll some 180 cases per group.

a word-processing program. In this example the text of the report reads (in part) as follows: "One goal of the proposed study is to test the null hypothesis that the proportion responding is identical in the two populations. The criterion for significance (alpha) has been set at 0.05. The test is two-tailed, which means that an effect in either direction will be interpreted. With the proposed sample size of 107 and 107 for the two groups, the study will have power of 80.1% to yield a statistically significant result. This computation assumes that the difference in proportions is 0.20 (specifically, 0.60 vs. 0.40). This effect was selected as the smallest effect that would be important to detect, in the sense that any smaller effect would not be of clinical or substantive significance. It is also assumed that this effect size is reasonable, in the sense that an effect of this magnitude could be anticipated in this field of research."

The software available for power analysis is constantly changing but up-to-date information is available on the Internet. Information on the program used in this example is available at http://www.PowerAndPrecision.com. Thomas and Krebbs (1997) maintain a site with links to any number of programs for power analysis at http://www.Interchg.ubc.ca/cacb/power.

3.14.4 PROBLEMS WITH THE SIGNIFICANCE TEST

To this point we have highlighted the key features of significance testing and power analysis. The vast majority of statistical analyses in psychological research take the form of significance tests and an understanding of the process is essential to an understanding of the research literature.

However, this was intended primarily as an introduction. The thesis of this chapter is that while the significance test is ubiquitous in the literature of psychological and medical research, its use in these fields is inappropriate for a number of reasons. In this section we will make the following points:

(i) The significance test addresses the wrong question. Researchers are concerned with clinical significance (i.e., the magnitude of the treatment effect), but significance tests address statistical significance (whether or not the effect is zero).

(ii) The question that is addressed by significance tests (whether or not the null hypothesis is true) is inappropriate not only because it is tangential to clinical significance but also because it is almost always false in psychological research.

Additional problems derive from the fact that significance tests lend themselves to mistakes of interpretation. In particular:

(i) A significant p-value is assumed to reflect a clinically important effect, and the level of significance is assumed to reflect the magnitude of the effect. In fact, though, the p-value incorporates information about the size of the sample as well as the size of the effect. Therefore, the use of the p-value for this purpose results in confusion and errors of interpretation.

(ii) Historically, a majority of studies have been run with low power, which means that even moderate or large treatment effects may go undetected. Additionally, the absence of significance is often interpreted, inappropriately, as evidence that the treatment effect is zero. The combination of these two types of error has seriously impeded research in psychology.

3.14.4.1 Significance Tests Address the Wrong Question

The key problem with significance testing is that it addresses the wrong question. The question that researchers want to ask is "How large is the treatment effect?" By contrast, the question actually addressed by the significance test is "Can we conclude that the treatment effect is not zero?"

Researchers typically assume that reports of "significance" or "nonsignificance" refer to clinical significance. This assumption follows from the fact that clinical significance is the question of interest, and also from the fact that "significance" refers to substantive (or clinical) significance in common parlance. Nevertheless, this assumption is incorrect. A significant p-value tells us only that the treatment (probably) increases the response rate by some amount. It does not provide information about the size of the effect.

The clinician, of course, needs information about the magnitude of the treatment effect—Does the treatment increase the likelihood of response by five percentage points, or 10 points, or 50 points?—since this is the kind of information required to balance the treatment's potential benefits against the costs and potential side effects. It is also the kind of information required to allow for an informed choice between the treatment in question and other possible treatments. The only information that the significance test can provide—that the treatment effect exceeds zero—is at best of tangential interest.

The same point being made here about implications for clinical practice applies also to the development of psychological theory. For example, Tukey (1969) points out that if the nature of elasticity had been defined as "when you pull it, it gets longer" the science of elasticity would have progressed very slowly. In psychological research as well, we need to address the size of the effect in order for theory to develop, and this is not addressed by the significance test.

3.14.4.2 The Null Hypothesis is not a Viable Hypothesis

There is a second problem with the use of the null hypothesis in psychological research. In addition to being irrelevant, it is also not a viable hypothesis, in the sense that it is known to be false in the overwhelming majority of psychological research.

For example, consider studies that compare the impact of two treatments. Hunter (1997) points to a survey conducted by Lipsey and Wilson (1993) which reviewed 302 meta-analyses of treatment studies in psychology and education. The average number of studies in each meta-analysis is 60, and the total number of studies included is 18 120. The treatment effect was zero (the null hypothesis was true) in less than 1% of the domains studied. Similarly, studies that look at the relationship between behaviors or traits are designed to test the null hypothesis that the correlation between traits is zero. In fact, though, it would be difficult to name two human characteristics whose correlation with each other is 0.0000000. In this kind of study the null hypothesis is always false.

The typical retort to the argument that the null is always false is that we do not intend to test whether or not the response rates are identical—rather, we intend to test whether or not they differ by some important amount. But this is exactly the problem. What we care about is whether or not the difference is clinically important, but what we are testing is whether or not the difference is exactly zero (Cohen, 1990).

This fact, followed to its logical conclusion, leads to an interesting paradox. Given that the effect is not zero, and since the p-value is a function of both effect size and sample size, it follows that with a large enough sample size, any effect will meet the criterion for significance. One can extend this argument in several ways:

(i) With a large enough sample size we know the study will yield a significant effect, so there is no need to run the study (Cohen, 1994).

(ii) With a large enough sample size we know the study will yield a significant effect, so if our result is not significant, then by definition we are committing a type II error (Schmidt, 1992).

(iii) With a large enough sample size the study will yield a significant effect, so what we are really testing is not whether the effect is large, but whether or not we have run enough subjects. Thompson (1992) suggests that instead of testing for significance we can assess how tired the researchers are. If they are tired, then we probably ran a lot of subjects and have a "significant" effect. If they are very tired, then we probably ran even more subjects and have a "very significant" effect.

There are exceptions to this point, which are discussed later, but these are rarely encountered in psychological research.

3.14.4.3 The Gratuitous Use of Significance Tests

Even those who argue that significance tests can serve an important role in psychological research recognize that the tests are almost invariably abused. Abelson (1997a) wrote a chapter in which he argues that the null hypothesis should not be banned, but includes a section entitled "The Null Hypothesis: Merely Misused, or Really Idiotic?" He uses the term "gratuitous" to describe the ritualistic application of significance tests to all kinds of data where the test could not possibly provide any useful information.

Indeed, significance tests are sometimes applied in ways that defy common sense. Abelson (1997a) cites the example of a researcher who divided a cohort of subjects at the group median into "low" and "high" groups and then performed a significance test to find out whether or not the groups were, in fact, different. In doing so the researcher missed the point that we use significance tests to make inferences about whether or not the groups differ in some systematic way, but in this case we know that they differ systematically because we created them that way. Simply put, one is high and one is low. In fact, this is the archetypical case of the situation cited earlier, where the groups clearly differ, and the only thing being tested is whether or not we have a large enough sample. What if the test had failed to yield a significant effect? Would the researcher have concluded that he had assigned persons to the two groups at random?

Cohen (1994) cites the case of a researcher who wanted to test the hypothesis that a particular ailment would never exist in a specific population. He tested 30 patients and found the ailment in only one. He wanted to know whether or not this (one case in 30) was "significant." In doing so, he missed the point that the theory required that there be no cases in the population, and if he found any then the theory was clearly false.

Papers are often seen in which persons report significance tests for reliability coefficients, in effect making the point that the scale's reliability is not zero. To show that the test has reliability better than zero is hardly reassuring to potential users of the test. Surely, the relevant question is whether or not the reliability exceeds 0.70 or 0.80. Abelson (1997b) suggests that "... declaring a reliability coefficient to be nonzero constitutes the ultimate in stupefyingly vacuous information" (p. 13).

Cohen (1994) traces the history of harangues against null hypothesis significance testing (NHST) and some of his quotes are reproduced here. Rozeboom (1960) wrote "The statistical folkways of a more primitive past continue to dominate the local scene" (p. 417). Bakan (1966), writing that "a great deal of mischief has been associated [with the significance test]" noted that the idea was hardly original with him and that "to say it 'out loud' is ... to assume the role of the child who pointed out that the emperor was really outfitted in his underwear" (p. 423). Meehl (1967) likened NHST to "a potent but sterile intellectual rake who leaves in his merry path a long train of ravished maidens but no viable scientific offspring." Rothstein (personal communication) noted that "... researchers [using NHST] often pose the wrong question and then proceed to answer that question incorrectly."

Cohen (1994) goes on to say that " ... we, as teachers, consultants, authors, and otherwise perpetrators of quantitative methods, are responsible for the ritualization of null hypothesis significance testing ... to the point of meaninglessness and beyond ... NHST has not only failed to support the advance of psychology as a science but also has seriously impeded it" (p. 997) (see also Cohen, 1990).

Rothman (1986b) writes "Testing for statistical significance today continues not on its own merits as a methodological tool but on the momentum of tradition. Rather than serving as a thinker's tool, it has become for some a clumsy substitute for thought, subverting what should be a contemplative exercise into an algorithm prone to error" (p. 447).

Additional citations on this topic include Abelson (1995, 1997a, 1997b); Borenstein (1994b); Carver (1978); Dar (1987); Estes (1997); Fisher (1959); Gonzalez (1994); Hagen (1997); Harlow, Mulaik, and Steiger (1997); Harris (1997a, 1997b); Hunter (1997); Loftus (1991); Lykken (1968); Meehl (1967, 1978, 1990); Murphy (1990); Neyman and Pearson (1932a, 1932b); Oakes (1986); Peto and Doll (1977); Scarr (1997); Schmidt (1992, 1996); Schmidt and Hunter (1996); Shaver (1993), and Shrout (1997).

3.14.4.4 *p*-Values: Responding with a Misdirected Answer

To this point we have shown that significance tests address a question that is at best tangential to the question of interest and at often entirely irrelevant. The use of significance tests carries with it an additional problem: Significance tests produce *p*-values that lend themselves to mistakes of interpretation.

The "*p*-value" highlighted by the significance test is a function of two elements—the size of the effect, and the precision of the estimate. When we consider the logic of the significance test, that is, to determine whether or not the population effect is zero, we can appreciate the simple elegance of the *p*-value. Either a large effect, or a large sample (yielding a precise estimate of the effect), or some appropriate combination of the two, provides evidence that the effect is not zero. By the same logic, however, it would not be appropriate to press the *p*-value into service as an indicator of effect size. Since the *p*-value incorporates information about both the size of the effect and the size of the sample, it does not allow us to distinguish between the two.

3.14.4.4.1 Misinterpreting the significant p*-value*

Statistical significance is often assumed to reflect substantive significance. Almost invariably, the first question asked by the reader, and the first point made by the researcher, is that the results were "significant." This is the point highlighted at meetings, in abstracts, and in the results section of publications. Often, the discussion of effect does not proceed beyond the question of significance at all. Even when it does, the issue of significance, that is, statistical significance, is emphasized over clinical significance or effect size. In fact, though, the only information imparted by a statistically significant *p*-value is that the true effect is (probably) not nil. A significant *p*-value could reflect a clinically meaningful effect. It could also reflect a clinically trivial effect that had been found in a large sample (because of the large sample the effect size is reported precisely, and though small is known to be non-nil).

Cohen (1965) writes "Again and again, the results section of an article describing an effect as significant or highly significant is followed by a discussion section which (usually implicitly) proceeds to treat the effect as if it had been found to be large or very large" (p. 102). The same point has been documented repeatedly in the field of psychology and medicine (Feinstein, 1975, 1976, 1977; Friedman & Phillips, 1981; Mainland, 1982; Nelson, Rosenthal, & Rosnow, 1986; Rothman, 1986b; Tversky & Kahneman, 1971; Wonnacott, 1985).

A review of a manuscript is recalled that purported to show a substantial advantage for a novel antipsychotic drug. The basis for this claim was a difference between treatment groups on a critical variable, with $p < 0.001$. As it happens, the sample size was of the order of 2000 patients, and a *p*-value <0.001 could reflect a difference that was so small as to be trivial. In fact, some of the baseline differences between the groups (which had, appropriately, been dismissed as negligible in size) were larger (and more significant!) than the post-treatment differences being submitted as evidence of the drug effect.

It gets worse. If the confusion of statistical significance with clinical significance is a problem in the interpretation of single studies, the situation is even worse when researchers use *p*-values to compare the results in different studies. This type of comparison is common when we want to know if the treatment is more effective in men than it is for women, or if one treatment is more effective than another.

Since the *p*-value incorporates information about both the sample size and effect size, a *p*-value of 0.05 could represent response rates in two groups of 50% vs. 70% (a 20-point effect) with a sample size of 50 cases per group. It could also represent response rates of 50% vs. 90% (a 40-point effect) with a sample size of 10 cases per group. In the second case the effect size is substantially larger than it is in the first case, but this fact is lost in the *p*-values, which are identical. Similarly, a *p*-value of 0.01 could represent response rates of 40% vs. 65% (25 points) with 50 cases per group, or 40% vs. 95% (55 points) with 10 cases per group. Again, the difference in effect sizes is not evident from the *p*-value.

In fact, Tversky and Kahneman (1971) found that students presented with information about *p*-values and sample sizes tend to make exactly the wrong conclusions about the effect size. Students were presented with two studies where the *p*-value was 0.05, and told that the sample size was 10 per group in one and 50 per group in the other. Invariably, students assumed that the effect size in the second case was more impressive, while exactly the reverse is true (see also Berkson, 1938, 1942; Friedman & Phillips, 1981; Rosenthal & Gaito, 1963, 1964; Rozeboom, 1960).

The possibilities for mistakes expand when we consider the possibility of comparing results between studies when both the *p*-value and the sample size differ. If one study yielded a *p*-value of 0.05 and another yielded a *p*-value of 0.01, then in the absence of any additional information a reader might assume that the effect size was stronger in the latter case. In fact, though, if the first study ($p = 0.05$) used a sample of 10 per group and the second ($p = 0.01$) used 50 per group, then the effect size would have been substantially larger in the study with the modest p-value (a 40-point effect as compared with a 25-point effect).

3.14.4.4.2 Misinterpreting the nonsignificant p*-value*

The complement to misinterpreting the significant p-value is to misinterpret the nonsignificant p-value. The only information imparted by a nonsignificant p-value is that we have failed to reject the null. Assume, for example, that refractory schizophrenic patients currently in relapse are assigned to be treated either with clozapine or with haldol. In this hypothetical example we find that the proportion meeting a remission criterion within six weeks is not significantly different in the two groups. The nonsignificant p-value could reflect a finding in a large sample that the proportion remitting was virtually identical in the two groups (the large sample ensuring to a high degree of certainty that the same would hold true in the population). Or, it could reflect a finding in a small sample that the remission rate was twice as high (or more) for the patients treated with clozapine (the finding failing to prove significant in part because of the low sample size). The first scenario would justify a conclusion that clozapine does not substantively increase the likelihood of remission and the second scenario would not justify this conclusion. However, researchers and readers almost invariably fail to make this distinction, and routinely interpret the absence of significance to mean that an effect does not exist. A related problem is the fact that many studies in psychological research suffer from low statistical power. In this section we will make four points.

(i) Power for research in psychology is abysmally low;

(ii) Rule (i) appears to be impervious to change;

(iii) The absence of significance should be interpreted as "more information is required" but is interpreted in error as meaning "no effect exists"; and

(iv) Rule (iii) appears to be impervious to change.

While Cohen has often made the point that he discovered, rather than invented, power analysis, the fact is that Cohen's papers over a series of decades have played a key part in making researchers aware of the power issue, and providing them with the mechanisms for computing power. At present, Cohen's papers and texts (Cohen, 1962, 1965, 1969, 1977, 1988, 1990, 1994) also serve as a kind of historical map that traces the role of power analysis since the 1960s.

Cohen (1962) surveyed papers published in the *Journal of Abnormal and Social Psychology* in 1960. Mean power to detect a small, medium, or large effect, respectively, was 0.18, 0.48, and 0.83. (Definitions of small, medium, and large were developed by Cohen with reference to the effect sizes typically found in social science research; the definitions were modified slightly subsequent to the 1962 paper.) The implications of this were spelled out clearly in the paper. At a minimum, a great deal of time and effort is being wasted on studies that have little chance of meeting their goals. Worse, when the studies with "negative" results are published, readers tend to interpret the absence of statistical significance as evidence that the treatment has been proven ineffective.

What was the response to this paper? Woody Allen has spoken of the time that he was kidnapped as a child. It took his father a while to catch on ("He had bad reading habits—he started reading the ransom note, but in the middle he fell asleep") but eventually he "sprang into action and rented out my room." In that vein, it took the scientific community a while to appreciate the implications of Cohen's paper but eventually it sprang into action also. First one, then another researcher found that what Cohen had done for the field of psychology could be done equally as well for any field of research, and in the years that followed a kind of cottage industry developed of papers that documented the fact of low power in any number of journals in the area of behavioral research. Many of these are cited in Sedlmeier and Gigerenzer (1989) and Rossi (1990). Similar papers were published to document the same problem in the field of medicine (Borenstein, 1994b; Hartung, Cottrell & Giffen, 1983; Phillips, Scott, & Blasczcynski, 1983; Reed & Slaichert, 1981; Reynolds, 1980) and psychiatry (Kane & Borenstein, 1985).

One oft-cited paper by Frelman, Chalmers, Smith and Kuebler (1978) surveyed reports of controlled clinical trials that had been published in a number of medical journals (primarily the *Lancet*, the *New England Journal of Medicine*, and the *Journal of the American Medical Association* during the period 1960–1977), and selected 71 that had reported negative results. The authors found that if the true drug effect had been on the order of 50% (e.g., a mortality rate of 30% for placebo vs. 15% for drug), median power would have been 60%. In other words, even if the drug cut the mortality rate in half there was still a 40% probability that the study would have failed to obtain a significant result.

How does the story end? In Woody Allen's tale the FBI surrounded the kidnappers. They wanted to "toss in the tear gas," but they had no tear gas so instead they put on the death scene from Camille. The kidnappers were overcome. The FBI and the kidnappers bargained with each other and eventually reached an agreement that the kidnappers would "throw out their guns and keep the kid."

The power analysis story has a similar ending. In keeping with the general move away from isolated studies and toward meta-analysis, papers that document low power in specific journals are now being replaced by meta-analysis of these papers. Now, researchers have been able to document the fact that low power exists not only in specific journals, but also in every field of research.

Sedlmeier and Gigerenzer (1989) published a paper entitled "Do studies of statistical power have an effect on the power of statistical studies?" They found that in the 25 years since Cohen's initial survey power had not changed in any substantive way. Specifically, they reported that in the 1984 volume of the *Journal of Abnormal Psychology*, mean power to detect small, medium, and large effects was 0.21, 0.50, and 0.84, which are essentially similar to the values reported by Cohen for the 1960 volume (in 1960, for the *Journal of Abnormal and Social Psychology*). Similarly, Rossi (1990) reviewed papers published in 1982 in the *Journals of Abnormal Psychology, Consulting and Clinical Psychology*, and *Personality and Social Psychology*. Mean power to detect small, medium, and large effects, respectively, was 0.17, 0.57, and 0.83. In fact, since papers with higher power are more likely to yield significant results and be published, the mean power reported by these surveys is almost certainly higher than the mean power for all studies in the field.

Given that the fact of low power is well documented, one would hope that researchers would adapt their thinking to accommodate to this state of affairs. In fact, though, the mistake of interpreting nonsignificant results as implying the absence of an effect runs through the psychological and medical literature. One could say that it practically gallops.

Cohen (1990) relates the story of a doctoral candidate who completed a study with 20 cases per group, and power of 0.33 to detect a medium sized effect. As Cohen recalls "He ended up with non-significant results—with which he proceeded to demolish an important branch of psychoanalytic theory" (p. 1304).

Earlier, the paper by Freiman et al. (1978) was cited, showing that power for papers published in some of the most prestigious medical journals was clearly at unacceptable levels. The authors went on to make a second point: Despite the fact that power was terribly low, in most cases the absence of significance was interpreted as meaning that the drug was not effective. The authors wrote: "The conclusion is inescapable that many of the therapies discarded as ineffective after inconclusive 'negative' trials may still have a clinically meaningful effect" (p. 694). In fact, it is possible (or likely) that some of the therapies discarded on this basis might well have had very substantial therapeutic effects.

In fact, Sedlmeier and Gigerenzer reported that power had actually dropped slightly over the 25 years subsequent to Cohen's initial survey. The magnitude of the drop is of little consequence, but the reason for the drop is interesting—researchers in 1990 were more likely to be making adjustments to alpha when running multiple tests (e.g., using alpha of 0.025 rather than 0.05) to ensure that the type I error rate was kept at 0.05. Schmidt (1996) points to the irony that researchers were protecting themselves against an error that could not occur (since the null hypothesis was false) while oblivious to the type II error rate, which exceeds 50%.

As was the case for misinterpretation of significant results, the misinterpretation of nonsignificant results becomes even more complicated when researchers try to compare results across studies. One would hope that the error of interpretation would be obvious when there are a series of studies, with some yielding a significant effect and others failing to do so. In some such series the effect size is remarkably consistent from one study to the next—the studies with a large sample size met the criteria for significance and those with a smaller sample size did not. Even in these cases, however, there is usually a perception that "the studies yield conflicting results."

Even in 1998 a cardiologist remarked that thrombiolytics are incredibly effective in reducing mortality for patients suffering a myocardial infarction, and that he sees clearly that they save lives, but that the studies of effect "flip back and forth." In fact, thrombiolytics represent the textbook case (literally) cited earlier (Lau et al., 1992), where the treatment effect is consistent from one study to the next, and the presence or absence of statistical significance varies with sample size.

Even worse, when the significant study was performed in one type of sample and the nonsignificant study was performed in another type of sample, researchers sometimes try to interpret this difference as meaning that the effect exists in one population only (e.g., males but not females). Abelson (1997a) notes that if a treatment effect yields a p-value of 0.07 for wombats and 0.05 for dingbats we are likely to see a discussion explaining why the treatment is effective only in the latter group—completely missing the point that the treatment effect may have been virtually identical in the two. The treatment effect may have even been larger in the nonsignificant group, if the sample size was smaller.

A more serious example of this is cited by Poole (1987a). Selikoff, Hammond, and Churg (1968) found a statistically significant association between asbestos exposure and lung cancer among cigarette smokers (the risk of death from lung cancer was 90 times as high as that for nonsmokers in the general population), but failed to find this same association among nonsmokers (no deaths were reported in this group during the observation period). The authors, taking into account the low statistical power for detecting this effect, drew no conclusions about the nonsmoking group. Others, however, were less cautious and proceeded to (mis)interpret this negative finding. Hoffman and Wynder (1976) wrote: "We conclude from [this study] that asbestos induces mesothelioma of the pleura and peritoneum, but not by itself [cancer] of the bronchus." Similarly, Cole and Goldman (1975) interpreted the study to mean that "[A]pparently, asbestos will produce lung cancer only in smokers." Poole reports that this finding was eventually summarized in a pamphlet distributed to workers in the asbestos industry: "Studies show that if you don't smoke cigarettes, asbestos does not increase your risk of lung cancer." Poole used the data reported in the original study to compute confidence intervals for the risk of cancer associated with asbestos exposure for the nonsmokers, and found the odds ratio to fall in the range of zero to 46.1.

We have shown that power for many studies is low, which leads to nonsignificant results even when treatments are actually effective. And, we have shown that the nonsignificant results are often taken as evidence that the treatment is not effective. Either of these two issues would be problematic by itself, and together these items seriously impede research in psychology.

Schmidt (1996) outlines the impact of this practice on research and policy: an idea is proposed (school integration will result in better test scores for African-American children). A number of studies are run. All of the studies show that scores are increased and the magnitude of the increase is consistent from one study to the next, but the sample sizes are small and, in some of the studies, not significant. Researchers report that the evidence is "conflicting." A series of studies are funded to determine why the integration had a positive effect in some studies but not others (is it the teacher's attitude? Is it the students' socioeconomic status; is it the students' age?), entirely missing the point that the effect was actually consistent from one study to the next. No pattern can be found (since none exists). Eventually, researchers decide that the issue cannot be understood. A promising idea is lost, and a perception builds that research is not to be trusted. A similar point is made by Meehl (1978, 1990).

Rossi (1997) gives an example of this same phenomenon from the field of memory research. The issue of whether or not researchers could demonstrate the spontaneous recovery of previously extinguished associations had a bearing on a number of important learning theories, and some 40 studies on the topic were published between 1948 and 1969. Evidence of the effect (i.e., significant findings) was obtained in only about half the studies which led most texts and reviews to conclude that the effect was ephemeral and "the issue was not so much resolved as it was abandoned" (page 179). Recently, Rossi returned to these studies and found that the average effect size (d) for the studies was 0.39. If we assume that this is the population effect size then, given the sample size in the various studies, the mean power for these studies would have been slightly under 50%. On this basis we would expect about half the studies to yield a significant effect, which is exactly what happened.

It seems unlikely that these errors of interpretation will change as long as researchers continue to work under the current system of significance testing. These types of problems have continued despite educational efforts such as those of Cohen (1962, 1990). They have continued despite the development of simple techniques for computing power, including rule-of-thumb tables (Cohen, 1988; Kraemer & Thiemann, 1987) and computer programs (Borenstein & Cohen, 1988; Borenstein et al., 1990, 1992, 1997; Goldstein, 1989; Rothstein et al., 1990). They have continued despite research that has addressed the kinds of thought processes on the part of researchers that appears to drive these kinds of errors (Tversky & Kahneman, 1971). There are many researchers who are very cognizant of power issues and insist that their studies be properly powered. Most large-scale and multicenter studies are also designed to ensure adequate power for important tests. In this sense, there has been some very substantial progress since the late

1960s. Still, it is clear from Sedlmeier and Gigerenzer, and Rossi, that low power continues to be a problem.

3.14.4.5 Summary

Significance tests are not the appropriate vehicle for the vast majority of psychological research, both because they are inherently inappropriate for this role and additionally because they foster mistakes of interpretation. Specifically:

(i) Significance tests focus attention on statistical significance (whether or not we can reject the nil) rather than clinical significance (which is the question of interest). The test can tell us about the direction of the effect, but in virtually all cases we want also to address the size of the effect.

(ii) There is a general perception that significance tests allow us to distinguish between effects that are "real" and those that are not. In fact, though, virtually all effects we test are nonzero, so the only question resolved by the significance test is whether or not we used a large enough sample.

(iii) The significance test yields *p*-values that lend themselves to misinterpretation. A significant *p*-value is assumed to reflect a clinically important effect, but it may not. A nonsignificant *p*-value is taken to reflect a zero effect, but it does not. These problems are compounded when a series of studies are seen as yielding "conflicting results" for one population, or are seen as evidence that the effect exists in one population but not another.

3.14.5 FOCUSED SIGNIFICANCE TESTS—A DIGRESSION

To this point we have used the term significance test to refer to the null hypothesis of no effect. Overwhelmingly, this is the manner in which statistical tests are applied in the literature. For example, the null hypothesis being tested is that the mean in two groups is identical (i.e., the intervention had no effect), or the response rates in two groups are identical, or the correlation between two variables is zero.

In fact, though, we can adopt a more focused approach by formulating and testing a hypothesis that predicts the magnitude of an effect. This would be appropriate if a theory had developed to the point that questions of effect size would bear on the validity of the theory. It would also be appropriate if we were working with a prospective treatment and the question of interest was whether or not the treatment was clinically useful.

For example, assume that we are planning to test a new drug as a maintenance medication for schizophrenic patients. The drug is known to increase the level of negative symptoms slightly, but would nevertheless be considered useful if it increased the response rate (defined by control of positive symptoms) by 20 percentage points. The hypothesis to be tested is not "Does it increase the response rate at all" but rather "Does it increase the response rate by 20 points *or more*."

The same logic can be extended to a test of means. Assume, for example, that we want to evaluate a new teaching strategy that is expected to yield an improvement in SAT scores. The strategy will be expensive to implement, and a decision is made that it would be adopted on a wide scale only if it increased the mean score by at least 50 points. The hypothesis to be tested is obvious.

The idea of using significance testing in this way is not new. In the psychological literature the null hypothesis tested is almost invariably the null hypothesis of no effect and it is commonly assumed that the word "null" in "null hypothesis" refers to the size of the effect. In fact, though, it is a reference to the hypothesis to be tested, or "nullified" (Cohen, 1997). While the hypothesis to be nullified is usually the hypothesis that two response rates are identical it could also be the hypothesis that, for example, the response rates differ by 20 percentage points. To avoid confusion, Cohen suggests that the term "null hypothesis" be used to refer to any hypothesis to be tested, while the term "nil hypothesis" be used to refer to the null hypothesis of no effect.

Similarly, in the clinical literature the term "significance test" is almost invariably used to refer to a test of the nil hypothesis, but it can refer to any null hypothesis. Cohen suggests that the term "significance test" be used to refer to a test of the nil, while the term "hypothesis test" be used to refer to other tests (such as the 20 point difference). These conventions will be followed here to avoid confusion. The reader should note, however, that these terms are used interchangeably elsewhere.

The use of hypothesis testing rather than significance testing would provide some obvious benefits. For one thing, it would force researchers to think about the meaning of the null hypothesis. Earlier, the example of researchers testing the null hypothesis that a reliability coefficient is zero was cited. With a focused test the researcher could test the hypothesis that the reliability was significantly better than, for example, 0.80. If these kinds of tests were in common use, researchers would have an incentive to formulate the hypotheses more clearly. Additionally, in those cases where the null hypothesis was rejected, one could assert that the effect size fell in a clinically important range ("the relapse rate is reduced by at least 20%").

However, many of the problems described for the use of significance tests with the nil hypothesis would apply also to the use of hypothesis tests for other null hypotheses.

First, the problems in logic that were identified for significance tests apply also, albeit in modified form, to hypothesis tests. It was argued earlier that the significance test addresses the question of whether or not the effect is precisely nil, which is of little import. Similarly, the focused hypothesis test might address the question of whether or not the effect is precisely 20 points. A 20-point difference might be of more interest than a nil effect but is still an arbitrary value. In practice, what we would like to do is to get a feel for the size of the effect as being trivial, or moderate, or substantial, rather than to rule out any specific value. Similarly, it was argued earlier that the nil hypothesis was rarely true, and that with a large enough sample it would almost always be disproved. Similarly, the null hypothesis that the rate difference is *precisely* 20 points is virtually certain to be false and with a large enough sample will almost always be rejected. Under the system of hypothesis testing, with a large enough sample, an observed rate difference of 21 points, or even 20.1 points, will require us to reject the null of 20 points. Of course, this is not the intention in running the test, but it serves as a stark example of why the test is not the appropriate vehicle for this kind of application.

Second, the mistakes of interpretation that were identified for significance tests apply, again in modified form, to focused hypothesis tests. One problem with the significance test is that researchers use *p*-values (improperly) as surrogates for effect size indices. A *p*-value of 0.05 was assumed to reflect a modest effect and a *p*-value of 0.01 was assumed to reflect a strong effect. While the focused hypothesis test does introduce the concept of effect size (a 20-point effect in this example) the key index continues to be the *p*-value. A *p*-value of 0.05 is likely to be seen as reflecting a modest departure from 20 points, and a *p*-value of 0.01 as reflecting a substantial departure from 20 points. Similarly, the absence of significance has traditionally been mistaken as evidence that the effect is zero. With the focused hypothesis test it would be mistaken as evidence that the effect is precisely 20 points.

Additionally, if a move was made to focused hypothesis tests, low power would become even more of a problem than it is now. For example, with a sample of 70 per group using a 50-point difference as the effect size (and $SD = 100$), power to show that one educational method is superior to another is 84%. With the same assumptions, power to show that it is superior by at least 20 points is only 42%. Logically, in fact, we might want to use 50 points as the effect size and also test the hypothesis that the effect is 50 points. When the effect size and the effect being tested are identical to each other, then no sample size will yield power greater than alpha (e.g., 5%).

In sum, the adoption of hypothesis tests would be appropriate for some studies but would not address the key problems outlined above for significance tests, and would additionally introduce some new problems. In any event, the possibility of working with hypothesis tests rather than significance tests is something of an academic point, since the field has bypassed this approach in favor of a more comprehensive approach—the report of effect size with confidence intervals, and it is to this approach that we turn now.

3.14.6 EFFECT SIZE AND CONFIDENCE INTERVALS

For the reasons outlined above researchers and journals are now moving away from significance tests and toward a focus on effect size estimation. Rather than report that an effect is (or is not) significantly different from zero, we report the size of the effect, and specify additionally a confidence interval, the range of treatment effects in the population that are likely to have given rise to the observed data.

In the example cited earlier, assume that we had completed the study and obtained response rates of 30% on drug A vs. 50% for drug B, with a sample size of 100 patients per group. We would report that drug B increased the response rate by 20 percentage points, with a 95% confidence interval of 7–33. The effect size observed in the sample (a rate difference of 20 points) is our best guess about the size of the effect in the population. The confidence interval (seven points to 33 points) reflects the likely upper and lower bounds of the effect in the population (Altman, Gore, Gardner, & Pocock, 1983; Borenstein, 1994a, 1994b; Gardner & Altman, 1986, 1989a, 1989b; Morgan, 1989; Simon, 1986, 1987).

3.14.6.1 The Key Advantage of this Approach

If the key problem with significance tests is that they address the wrong question, then the key advantage of effect size estimation is that it addresses the right question. Working with significance tests we defined an effect and asked whether or not that effect was probably true. It was problematic that the effect was typically of little interest (the nil effect) or selected arbitrarily (the focused hypothesis test). Working with effect size and confidence intervals, rather than report on what the effect is not we report on what the effect is. By doing so we focus on the question of interest. Additionally, we bypass the need to specify an arbitrary null hypothesis.

The effect size with its corresponding confidence interval provides three key pieces of information:

(i) An estimate of effect size that is separate and distinct from the precision of this estimate. Because this is a pure measure of effect, it focuses our attention on the substantive impact. For example, the treatment increases the response rate by 25 percentage points.

(ii) A value for the lower bound of this value in the population (for a given confidence level). With the traditional test we know only that the response rate is increased by some amount exceeding zero. By contrast, with the confidence interval we may be able to report that it is increased by at least 20 percentage points.

(iii) A value for the upper bound of this value in the population. The traditional test provides no estimate at all of the upper bound. By contrast, with the confidence interval we might report that the upper bound for the treatment's impact is 10 percentage points (in which case we know that the impact is of limited substantive value) or that the upper bound is 50 percentage points.

3.14.6.2 Effect Size and Confidence Intervals can be Computed for Any Type of Study

Any study that may be used to compute a *p*-value may also be used to generate an effect size with confidence intervals. This follows from the fact that the *p*-value includes both components (the effect size and the precision). The *p*-value combines the two to yield a single value, but each value can be presented separately. As was the case with significance tests, the exact form of the report will vary with the nature of the data, but the differences are minor.

For studies that focus on the mean difference between groups a natural index of effect is the mean difference itself (either in the original metric, such as SAT points, or standardized to a common metric with a standard deviation of 1.0). For studies that focus on the proportion of cases responding to a treatment, one index of effect is the rate difference, that is, the difference in the proportion responding in either group. Other indices, common in medicine and epidemiology, are the relative risk and the odds ratio, either of which is often presented in log units. For studies that deal with correlations, the correlation itself serves as an index of effect size. For studies that focus on time to an event, or survival, the effect size reported is often the hazard ratio, that is, the relative risk per unit time in one group vs. the other. Computational details for any of these computations are readily available (see, for example, Harris & Albert, 1991; Kleinbaum, Kupper, & Morgenstein, 1982; Lawless, 1982; Rosner, 1990; Selvin, 1991). The goal here is to point out that these exist, and that the effect size can be presented for any type of study.

The effect size is reported with a corresponding confidence interval. The computation of confidence intervals is always some variation on the theme

Lower Limit = Observed effect minus
(Standard Error* 1.96)
Upper Limit = Observed effect plus
(Standard Error* 1.96)

At the core of either equation is the observed effect, about which the confidence interval is constructed. The lower and upper limits represent the lower and upper bounds for the effect size in the population from which the sample was drawn. Computation of confidence intervals for proportions, correlations, or survival times is more complicated than suggested by this summary, and is discussed in more detail below. These computational details aside, the formula shown here offers a straightforward conceptual overview.

3.14.6.3 Factors Affecting the Confidence Interval Width

The width of the confidence interval is based on two factors. The first of these is the sample-to-sample dispersion of the effect size: The higher the dispersion within a sample, and/or the lower the sample size, the more the effect size will vary from one sample to the next. When working with means the dispersion within a sample is indexed by the standard deviation of the scores. Intuitively, if working with a select group of students and the SAT scores cluster in the range of 600–750, sample-to-sample dispersion will be modest. By contrast, if we are working with a broader population, in which the scores cover the full range of 200–800, sample-to-sample dispersion will be wider. The impact of sample size on the standard error of the effect size is also intuitive. If the sample size is small the group mean might be affected substantially by a few outliers. By contrast, if the sample size is large the impact of outliers will be dissipated. It should be noted that when working with proportions or correlations, the first of these elements, the within-sample dispersion, is a function of the index itself. In the case of proportions, when the population proportion is close to 50% the standard deviation is relatively high, and it declines as the proportions approach either zero or unity. Intuitively, one can imagine that samples drawn from a population where the proportion of responders is 99% will yield values close to 99%—it would be difficult to draw many nonresponders since they do exist. By contrast, if the proportion of responders is 50% it would not be hard to draw one sample with a 40% response rate and another with a 60% response rate. Similarly, for correlations, the standard deviation is highest when the correlation is zero, and it declines as the correlation approaches minus one or plus one.

The second factor controlling the width of the confidence interval is the level of confidence desired. For example, one could choose to present the effect size plus/minus one standard error. In approximately 68% of all studies where the confidence intervals are reported in this manner, the interval will include the population effect size. One could also choose to present the effect size ± 1.96 standard errors (assume for this example that the variance is known). On the normal curve, 95% of all observations will fall inside the interval of -1.96 to $+1.96$. Therefore, if the effect size is presented as ± 1.96 standard error units, over an infinite number of studies 95% of the intervals will include the population effect.

The researcher may elect to compute two-tailed or one-tailed bounds for the confidence interval. A two-tailed confidence interval extends from some finite value below the observed effect to another finite value above the observed effect. A one-tailed confidence interval extends from minus infinity to some value above the observed effect, or from some value below the observed effect to plus infinity (the term "interval" is a misnomer in the one-tailed case but is used to maintain readability).

3.14.6.4 Information Conveyed by the Effect Size and Confidence Interval

Figure 5 shows the results of six fictional studies. For each study, the effect (the difference in response rates between placebo and drug) is depicted as a vertical line, bounded on either side by the 95% confidence interval. A solid line from the top to the bottom of the graph marks the null effect of no difference. An effect to the left of this line indicates that the drug group was inferior to placebo, while an effect to the right indicates that the drug group was superior. The studies with narrow confidence intervals are based on a sample of 700 patients per group. Those with wide confidence intervals are based on a sample of between 10 and 15 patients per group. All examples assume that the response rate on placebo is 30%. We can visualize a hierarchy of possible study results, some of which are depicted on this schematic.

Study A: Zero difference (-30 to $+30$). We cannot rule out the possibility that the effect is nil. Nor, however, can we rule out the possibility that the effect is large enough to be clinically useful or clinically harmful.

Study B: Zero difference (-4 to $+4$). We cannot rule out the possibility that the effect is nil. More to the point, however, it is clear that the effect is trivial, at best.

Study C: Twenty-five point difference (-10 to $+54$). The effect size in the sample (and our best estimate of the population effect) is 25 points. We cannot rule out the possibility that the effect is nil (nor can we rule out the possibility that it is quite potent).

Study D: Five-point difference (1–9). The effect is probably not nil. More to the point, however, we can assert with a high level of certainty that the effect is not clinically important (at best, a nine-point advantage in favor of the drug).

Study E: 40-point difference (10–63). The effect is probably not nil. The possible magnitude of the effect ranges from small to extremely potent (the interval is intentionally asymmetric).

Study F: 40-point difference (36–44). The drug is quite potent. The likely range of effects falls entirely within the "very potent" range.

3.14.7 THE RELATIONSHIP BETWEEN SIGNIFICANCE TESTING AND EFFECT SIZE ESTIMATION

The formulas for significance tests and for confidence intervals are mathematically congruent. The standard error of the effect size in confidence intervals is identical to the standard error used in significance testing. The confidence level used for confidence intervals can be complementary to the significance level used in significance tests (with the 95% confidence level corresponding to alpha of 5%).

Given the correspondence between the two methods, any significance test of the form "the difference is zero" and tested with alpha of 0.05 may also be tested by computing the 95% confidence interval and determining whether or

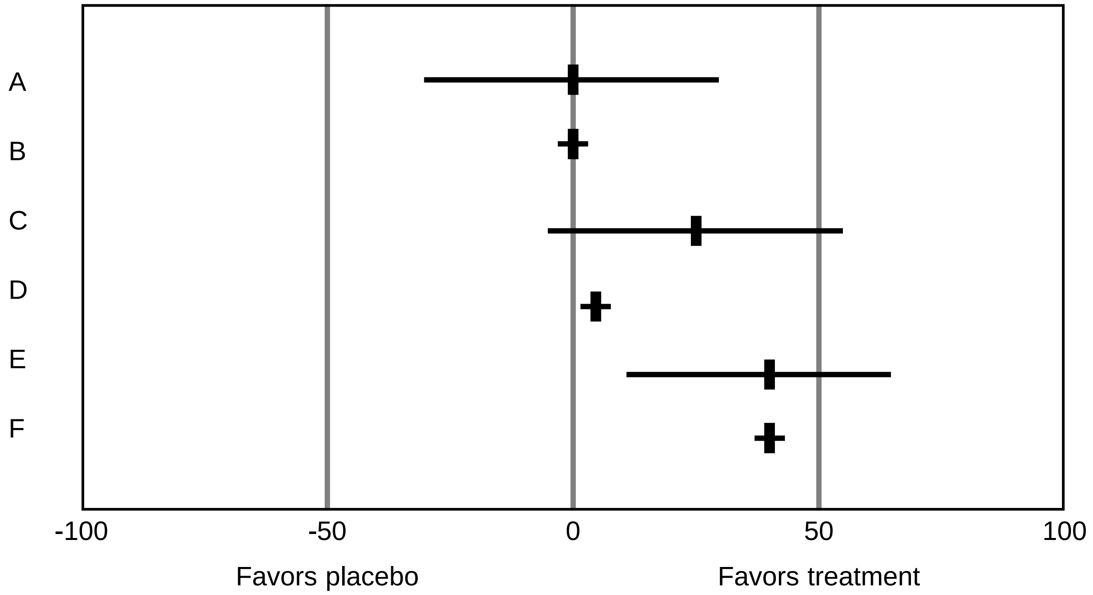

Figure 5 Data for six fictional studies. The effect size for each study is depicted by a vertical line bounded by the corresponding 95% confidence interval. The effect size highlights the magnitude of the effect while the interval width provides information about the precision with which the effect is estimated. This schematic also provides information about the significance test—if the 95% confidence interval excludes zero, the study is significant at the 0.05 level.

not this interval includes zero. More generally, any hypothesis test of the form "the difference is 20 points" and tested with alpha of 0.05 may also be tested by computing the 95% confidence interval and determining whether or not the difference includes 20 points. If the confidence interval includes the null, then we do not reject the possibility that the true effect is the null. If the confidence interval does not include the null, then we conclude that the effect is (probably) not the null. The same holds true, of course, for a significance test with alpha set at 0.01, and a report of the 99% confidence interval.

When we report a confidence interval and then focus on the question of whether or not the interval includes zero, the distinction between confidence intervals and significance tests is one of format rather than substance. However, as suggested by Figure 6, this shift in format is critically important. Working within the traditional framework of significance testing (shown by the columns in this figure) the key distinction is between studies A, B, C (nonsignificant) on the one hand, as opposed to D, E, F (significant) on the other. By contrast, working within the framework of effect size estimation (shown by the rows) we would group studies B and D (not clinically important), then A, C, E (possibly clinically important), while study F (clinically important) would be classified by itself. (A 20-point difference in response rates would be clinically important in the context of this study.)

3.14.7.1 Confidence Intervals as a Surrogate for the Significance Test

Mathematically, then, confidence intervals can serve as a surrogate for the significance test. Should they be used in this way? Researchers have staked out three positions:

(i) Confidence intervals should be used only as a surrogate for the significance test. Otherwise the results are entirely subjective, and we have no basis to make a decision about whether or not the treatment is effective.

(ii) Confidence intervals can serve as a surrogate for the significance test but they can also stand on their own. Some studies will serve best if used to provide an estimate of effect size, accompanied by an estimate of precision. Others would additionally take note of whether or not the confidence interval included specific values.

(iii) Confidence intervals should never serve as a surrogate for the significance test. The significance test has no role in psychological research, and any vestige of the significance test should be removed.

There is a substantial literature developing on these points of view. In general, the debate in the medical literature is between the first two options while the debate in the psychological literature is between the first and the third. The rationale for each of the positions is outlined here.

		Statistically significant	
		No	Yes
Clinically important	No	B	D
	Possibly	A,C	E
	Yes		F

Figure 6 Effect size estimation vs. significance testing. The columns correspond to significance tests—Studies A, B, C are not significant while studies D, E, F are significant. The rows correspond to effect size estimation—Studies B and D show that the treatment is not clinically important, studies A, C, E show that the treatment may be clinically important, and study F shows that the treatment is clinically important.

3.14.7.1.1 Confidence intervals should serve only as a surrogate for the significance test

Some researchers have taken the position that the basic idea of significance testing is a good one, and that the problems outlined above stem from the misuse of these tests rather than from any flaw inherent in the method. They argue that it is critical to retain the logic of the significance test since the researcher and the reader need to make a decision at the end of the study about whether or not the treatment effect was probably "real." By using confidence intervals as a surrogate for the significance test these decision rules are maintained while gaining the following critical advantages over the traditional significance test.

First, attention is focused on the magnitude of the treatment effect. By doing so the tendency is avoided to press the *p*-value into service as an indicator of effect size. A rate difference of five percentage points is seen as being a small effect (in most contexts), even if the sample is so large that the confidence interval does not include zero (and the test is statistically significant). By the same logic an effect of 50 percentage points is large, even if the sample size is so small that the confidence interval does include zero (and the test is not statistically significant).

Second, unambiguous information is provided about the precision of the effect. An effect reported with a precision of ±5 percentage points is definitive—the effect is known to be useless if small, or is known to be useful if large. An effect reported with a precision of ±30 percentage points is likely to leave open the question of clinical utility, but this point will be obvious to the researcher and the reader. Mistakes of interpretation will be less likely than they are with significance tests.

At the same time, this approach preserves the protective mechanisms that are incorporated in the significance test. Specifically, if the confidence interval excludes zero then it is assumed that the effect is not nil; if it includes zero then it is acknowledged that the true effect may in fact be zero (see, for example, Chow, 1988, 1989; Walker, 1986a, 1986b).

3.14.7.1.2 Confidence intervals can optionally be used as a surrogate for the significance test

Others have taken the position that confidence intervals can be used as a surrogate for the significance test, but that it is not necessary to focus on significance when interpreting the results. In the running example one could articulate a series of goals, including the following:

(i) We would like to be certain (where "certain" corresponds to a given level of certainty) that the treatment is not actually harmful.

(ii) We would like to be certain that the treatment effect exceeds zero.

(iii) We would like to be certain that the treatment effect is large enough to be clinically important.

As suggested by Figure 5, the 95% confidence interval might or might not exclude effects that are clinically harmful, that are zero, or that are clinically trivial. More generally, the effect size and confidence intervals might allow us to draw a conclusion about some of these points but not others. Or, they might allow us to answer some questions with one level of certainty, and others at another level of certainty.

In fact, rather than focus on the issue of whether or not the interval includes any single value we can consider the full range of population parameters that are consistent with the study results, recognizing the fact that those falling closer to the observed effect are more probable than those closer to the extremes. This approach was suggested by Birnbaum, (1961). More recently, it has been popularized by Miettenan and by Poole (1987a) who suggests that this is displayed most effectively as a graph of confidence interval by confidence level (see also Foster & Sullivan, 1987; Poole, 1987b, 1987c; Sheehe, 1993; Smith & Bates, 1993; Sullivan & Foster, 1990).

In the running example, assume that 20 patients were followed under each treatment condition and the proportion responding was 25% in one treatment group as compared with 50% in the other. The continuous confidence interval for this data is shown in Figure 7. This graph highlights the fact that the most likely impact of the new treatment mode is to increase the response rate by about 25 percentage points. The 60% confidence interval includes a benefit of between 12 and 37 percentage points for the new treatment; the 80% confidence interval extends from five percentage points to 42 percentage points; the 90% confidence interval extends from zero to 46 percentage points; and the 95% confidence interval extends from minus five (a five-point advantage for the standard treatment) to plus 49 (a 49 percentage point advantage for the new treatment).

While it might seem at first glance that this approach introduces an element of subjectivity into the analysis, the fact is that any test of significance involves an element of subjectivity since the decision to set alpha at a particular level is a subjective one, based (at least in theory) on the researcher's need to balance type-I error against type-II error. Additionally, the decision to focus on whether or not the interval includes the nil effect, rather than (say) an effect of five points or 10 points, is also arbitrary. Therefore, the use of confidence intervals in the manner being discussed here, rather than introducing an element of subjectivity into the analysis, merely shifts the subjectivity that is already there from the researcher and on to the reader (Fleiss, 1986a, 1986b; Thompson, 1987a, 1987b, 1987c, 1987d. See also Cohen, 1962; Derouen, 1987; Lachenbruch et al., 1987; Rosnow & Rosenthal, 1989, Rothman, 1986b; Savitz, 1987).

Perhaps more to the point, the reader may not feel a need to frame the answer as a dichotomy, and to conclude that the treatment does (or does not) have any impact on recurrence rates. Rather, the take-home message might be the more comprehensive picture, that is, (in Figure 7 and introducing subjective judgments) "the 40% confidence interval falls entirely in the range of substantial effects; the 80% interval falls in the range of substantial or moderate effects; and the 95% interval includes effects that are trivial" with the entire gamut of possibilities being factored into a decision about the utility of this treatment (see also Burnand, Kernan & Feinstein, 1990; Poole, 1987a, 1987b, 1987c; Rothman, 1986a, p. 119; Walker, 1986a, 1986b; Walter, 1991).

3.14.7.1.3 Confidence intervals should never be used as a surrogate for the significance test

Others have taken the position that even if the abuse associated with the significance test could be avoided, there is a fatal flaw inherent in this approach. The argument cited above in favor of the first position is also the key argument cited for this position: "If we use confidence intervals as surrogates for the significance test, then we retain the basic protective mechanisms of the significance test. Specifically, if the confidence interval excludes zero then we assume that the effect is not nil; if it includes zero then we explicitly acknowledge that the true effect may in fact be zero." Proponents of the third position would add "And therein lies the problem."

It was argued earlier that the null hypothesis is of little interest. If this is true, then when reporting an effect size we should care about the precision of the estimate, but not about whether or not the confidence interval (CI) includes zero. Put another way, an effect of 30 points with CI of 20–40 is informative; with CI of 10–50 is less informative; with CI of 0–60 is less informative, still. But there is no reason that we should be more concerned about "ruling out zero" than about "ruling out" a five-point effect or a 10-point effect. Knowing that the effect is probably not zero does not tell us that the effect is probably clinically important.

Similarly, it was argued that significance tests are misleading since they allow us to confuse effect size with precision. Studies meet the criterion for significance in part because the sample size is large; studies fail to meet the criterion for significance in part because the sample size is not large enough. By exactly the same logic, if the confidence interval includes zero it is in part because the sample size is too small. If the key issue being addressed by the significance test is whether or not there was a large enough sample, then it makes no sense to perform a significance test, in any format. To do so is to propagate and codify exactly the same error that attempts are being made to avoid (see Schmidt, 1992; Schmidt & Hunter, 1996, 1997).

3.14.7.2 Should the Significance Test be Banned?

In 1998 the significance test continues to serve as the *de facto* standard of proof in most research

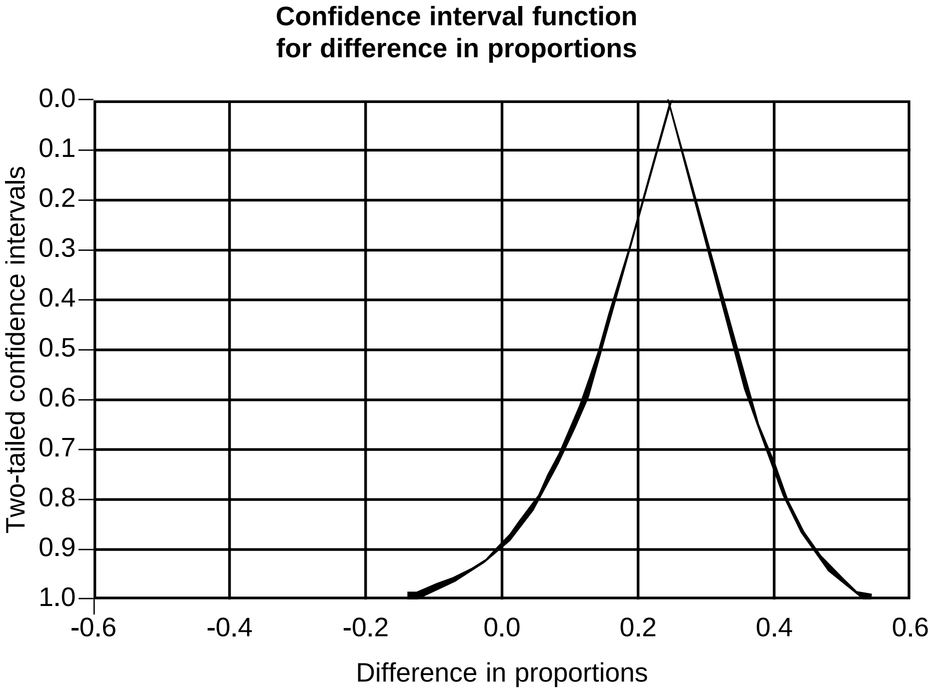

Figure 7 Continuous confidence intervals. Rather than focus on whether or not we can exclude a specific effect size with an arbitrary level of confidence, we can consider the full range of population parameters that are consistent with the study results, recognizing that those falling closer to the observed effect are more probable than those falling closer to the extremes.

studies. As a practical matter, the grant applicant who proposes to skip the significance test is not likely to be taken seriously by the funding agency; drug manufacturers are not likely to submit studies for FDA approval without data on statistical significance; and papers submitted for publication without tests of significance will face a serious obstacle in most journals. In the context of the three positions outlined above, the current state is somewhere prior to the first position. Nevertheless, recent debate has focused on the proposition that the significance test should be banned, which falls at or beyond the third position.

The most vocal proponents in favor of a ban are Schmidt and Hunter (1997) who attack the null hypothesis test on three fronts. First, they note that the null hypothesis is almost always false. Rather than thinking of studies as "significant" or "not significant" we would do well to think of them as "correct" and "incorrect." As long as researchers are permitted to report significance tests, type-II error will continue to run rampant in psychological research. Second, they argue that the role of the single study should be to accumulate data which can later be incorporated in a meta-analysis, and so there is no need to perform a significance test for the single study. As long as significance tests are permitted, the significant studies will be accepted for publication at a higher rate than non-significant studies, which will introduce a bias into the research literature and impair ability to conduct a meta-analysis at a later point. Finally, they contend that the significance test serves no useful function at all. In support of this latter point they proposed a ban on significance tests and then systematically collected objections to this proposal. They summarize these objections by reporting that "Each of these objections is intuitively appealing and plausible but is easily shown to be logically and intellectually bankrupt" (p. 37). Some of the objections to a ban follow, together with comments.

Some have taken issue with the contention that the nil hypothesis is always false. For example, if using a randomized trial to test an entirely new class of drug that actually has zero effect, then the null hypothesis would be true. This is a valid point but largely irrelevant, since it applies to only a trivial number of studies.

Some have argued that significance tests assure objectivity. In fact, one could argue that the hypothesis test is not objective since the selection of a criterion alpha, and also the selection of a specific null to be tested, are made subjectively. Even if one accepts the idea that confidence intervals are less objective, the fact remains that this objectivity leads to ubiquitous errors of interpretation. To paraphrase Tukey, "Be approximately right rather than exactly wrong."

Others have taken issue with Schmidt and Hunter's contention that a single study cannot provide definitive evidence to support or refute a theory. Abelson, for example, cites the classic Festinger experiment, which raised questions about the prevailing theory and set the stage for the studies that followed. (In this study one theory predicted that one group would score higher and the competing theory predicted that the second group would score higher. Ratings were conducted on an arbitrary scale, and the only issue to be resolved was the direction of the effect). This is a single study that had a substantial impact on the direction of the field, inasmuch as it set the direction for numerous confirmatory studies that followed. Again, these arguments are valid but they apply in only a small fraction of studies.

Perhaps the key point of Schmidt and Hunter is not that any single objection to the of banning significance tests is or is not valid, but the fact that researchers cling so hard to their beliefs about significance tests. For example, researchers keep coming back to the idea that we need the tests to distinguish between "real" effects and "chance effects," thereby missing the key point: All effects are real. The tests distinguish between studies that showed this and those that (in error) did not. Schmidt and Hunter write "We attribute this psychologically interesting fact to the virtual brainwashing in significance testing that all of us have undergone, beginning with our first undergraduate course in psychological statistics."

Should significance tests be banned, then? Abelson (1997a) wrote that the idea of such a ban is appealing "In the words of a famous empathizer 'I feel your pain'" (p. 118). The abuse of significance tests is so pervasive that it is likely to require some kind of absolute ban to prevent the continued abuse of this tool. If significance tests are allowed to continue, they will continue to be abused, and if the confidence interval is allowed to serve as a surrogate for the significance test, then the door to abuse will be wide open. He notes that abuse would be not only possible, but inevitable: "Indeed, under the Law of Diffusion of Idiocy, every foolish application of significance testing will beget a corresponding foolish practice for confidence limits." Despite these concerns, Abelson concludes that the tests should not be banned. He writes, "Create a list of things that people misuse—for example, oboes, ice skates, band saws, skis, and college educations. Would you be inclined to ban them because people make errors with them? Will we want to ban effect sizes, too, when their misuse escalates?" (p. 13).

While Abelson is correct that the idea of imposing a ban is unappealing, it is also clear that some kind of decisive action is required. Without some concrete and clear action on the part of educators, funding agencies, and journal editors, the shift away from significance tests will drag out over an indeterminate period of time. During this period a great deal of effort will be spent in repeatedly making the case for the shift, harm will be caused, and a great deal of potentially important work will go undone.

It is submit that the appropriate course of action would be not to ban the significance test but rather to require that researchers report effect size indices and confidence intervals. Scientifically, this is more defensible. Pragmatically, this can take effect immediately without further debate. In fact, this requirement is implicit in guidelines now, in the sense that analyses are expected to address the research question and conclusions are expected to follow logically from the analyses. Only rarely do we care about excluding a specific value, and even more rarely do we care about excluding the nil. In virtually all studies, the requirement that analyses address the research question will require that the report highlight the effect size and the precision with which this is reported.

While researchers would be free to report *p*-values in addition to the effect size and confidence interval, it is likely that this practice would disappear because it would become evident that the significance test is of tangential interest, at best. The experience of Rothman is informative in this regard. As an editor at the *American Journal of Public Health*, he took a firm stance against the inappropriate use of significance tests, and in doing so both educated and empowered researchers to focus on effect size in their analyses. More recently, as the editor of *Epidemiology*, Rothman has established a policy that prohibits references to significance testing, but reports that "Interestingly, few readers seem to be aware of this policy, which tells me that they don't miss the absence of claims regarding statistical significance. As you urge, we emphasize the measurement of effect size, and this gets a clear message across to readers."

3.14.7.3 Should the Analysis of the Single Study be Banned?

Similarly, while one can empathize with the desire to ban the analysis of single studies, Schmidt and Hunter overstate their case when they argue that no single study can provide definitive information on its own. Abelson writes, "Indeed, one might say that Hunter and Schmidt use the 'Brooklyn' method of argumentation: The Brooklynite asks you a question, tells you his answer, and then says that you are wrong." The fact is that some studies are well designed, with adequate power and appropriate controls. In fact, some single studies contain more subjects and better designed controls than some meta-analyses. To balance Hunter and Schmidt's point that no substantive issue in psychology has ever been resolved by a single study and that all testing should be deferred to the meta-analysis, we cite a recent editorial in the *New England Journal of Medicine*. Bailar (1997, 1998) noted that no substantive issue in medicine has ever been resolved by a meta-analysis, and that the randomized clinical trial is still considered to be the gold standard. (Bailar acknowledges that meta-analysis hold great potential but objects to the improper use of this procedure. Recall Abelson's prediction, cited earlier.) Additionally, even when the single study cannot yield adequate precision there is a need to provide advice about treatment options based on the evidence at hand, or to use this evidence to help design subsequent research. In such cases the analysis of a single study may be required.

Schmidt and Hunter highlight an important problem when they note that the practice of allowing significance testing for individual studies contributes to the problem of publication bias, but this problem exists, in large part, because researchers and editors focus on the *p*-value. Abelson wrote that in deciding whether or not a study should be published he takes into account five components, including such issues as the role of the study in the context of other research and theory. The *p*-value is a subset of the fifth component, and the least interesting one, at that. If the recommendations outlined above for focusing on effect size are followed, the illogic of basing a publication decision on the significance test should become clear, and the publication bias minimized.

3.14.7.4 Looking Ahead

It seems likely that the shift from significance tests to effect size estimation with confidence intervals is likely to take place over the next several years. Earlier, the history of researchers railing against the null hypothesis tests was outlined, beginning in the 1950s and proceeding to the present. The impact of these writings has been modest. Nevertheless, whether or not significance tests are banned, it is likely that the next few years will be marked by a shift away from significance tests and toward effect size estimation with confidence intervals. This prediction is based on the convergence of several factors (Schmidt, 1996).

The first is a shift in educational resources. Some statistics texts now place more of an emphasis on confidence intervals than they had in the past. Computer programs used for analysis are more likely now than in the past to present confidence intervals as well as tests of significance. Some programs for study planning now incorporate the ability to plan for confidence intervals as well as statistical significance (see, for example, Borenstein et al., 1997 or Elashoff, 1997).

The second is the acceptance of effect size estimation by important segments of the mainstream research community. Journal editors in some fields now encourage or require the use of confidence intervals in research reports (Altman, 1991; Altman et al., 1983; Rothman, 1978, 1986b; Rothman & Yankauer, 1986). The use of confidence intervals is also encouraged in the *Uniform Requirements for Manuscripts Submitted to Biomedical Journals* (International Committee of Medical Journal Editors (International Committee of Medical Journal Editors, 1991) [p. 425]). There is now a task force working on behalf of the American Psychological Association, charged with the goal of studying the issues raised in this paper and making recommendations. Harlow et al. (1997) edited a book entitled "*What if there were no significance tests*" and discuss eight points that are endorsed by some or all of the chapter contributors. There is unanimous rejection of the traditional use of significance tests—specifically the exclusive focus on the nil hypothesis and *p*-values. There is unanimous agreement that researchers should present effect size estimates with confidence intervals, should be cognizant of power issues, and should evaluate the results in the context of a theory rather than in isolation. There is strong support, with a few dissenters, for the use of tests that are designed to confirm or refute specific, hypothesis-based theories other than the nil hypothesis.

The third, and most important, factor in this move away from significance tests is the increasingly common application of meta-analyses. Meta-analyses serve both to highlight the flaws in significance tests and to provide a realistic mechanism for the application of effect size estimation. Meta-analysis is the next logical step in the research process and we return to this point at the conclusion of this chapter.

3.14.8 USE OF CONFIDENCE INTERVALS IN STUDIES WHOSE PURPOSE IS TO PROVE THE NULL

Some studies are conducted for the express purpose of showing that the null is true—for example, that two drugs are bioequivalent or that two treatments are equally effective. These studies are framed along the following lines: A study compares the mortality rates in patients treated with either of two drugs. The study has power of 95% to detect a difference of, say, five percentage points in mortality rates and so if the study fails to yield a significant effect we can conclude (with 95% certainty) that the groups do not differ by five percentage points.

The same logic is sometimes applied after the fact to a study which was initially run in hopes of finding a significant group effect. It was shown earlier that the absence of an effect is often interpreted without justification as meaning that the groups are identical. While a conclusion that the treatments are equivalent cannot be justified, one can make the argument that the study had power to detect a difference of some magnitude, and the absence of significance indicates that the groups do not differ by this much. While this approach is technically correct it suffers from two basic flaws.

First, it involves the loss of useful information. Assume, for example, that a study is planned in which schizophrenic patients currently in remission are assigned to treatment with either a low dose or a standard dose of neuroleptic treatment ($N = 40$ per group) and followed for a year. The study has power of 95% to detect a difference of 38 percentage points in recurrence rates (specifically, rates of 0.80 vs. 0.42). Therefore, if the study failed to yield significance it would be fair to conclude that the impact of treatment on recurrence rates, if any, is less than 38 percentage points.

However, the study that fails to yield significance could yield a rate difference of 18 percentage points (0.80 vs. 0.62) with $p = 0.08$ and a 95% confidence interval of -0.02 to 0.35. Or, it could yield a rate difference of two percentage points (0.80 vs. 0.78) with $p = 0.83$ and a confidence interval of -0.15 to 0.18. Once the study is completed it would be inappropriate to fall back on the original computations and report that the rate difference in the population is less than 38 percentage points. Rather, we could report (in the first example) that the rate difference could be as much as 35 points and (in the second) that it is at most 18 percentage points.

Second, the logic of using a power analysis to argue that two groups are comparable is muddled and often misunderstood. We are arguing that if the effect had been of a certain size we would have found it, and we did not, so we can assert that the effect is not that large. Consider, by contrast, the elegant simplicity by which the same information is conveyed in the context of confidence intervals. In the examples cited above we would report that the rate difference falls in the range of -0.02 to 0.35, or the rate difference falls in the range of -0.15 to 0.18. In other words, confidence intervals allow us to report what the effect is rather than what the effect is not (Blackwelder, 1982; Detsky & Sackett, 1985, 1986; Makuch & Johnson, 1986).

The response which had been framed in terms of significance tests and p-values is confusing because the sole function of the significance test is to evaluate the null hypothesis. If we try to reframe the null to mean almost null because we are trying to estimate the magnitude of an effect, then we are using the technique in a manner for which it was not intended and the results are awkward.

3.14.9 COMPUTATIONAL ISSUES IN CONFIDENCE INTERVALS

Earlier, it was noted that a confidence interval could be approximated as the observed effect plus/minus a specified distance. This is a shortcut that yields approximate results, and two caveats are in order. First, even for this approximation the computation is somewhat more complex than suggested here. If the index is based on a mean difference, then to compute the 95% interval we would multiply not by 1.96 but by a factor, based on the t-distribution, that takes sample size into account. Additionally, for most indices, including those based on proportions or correlations, we would transform the data prior to computing the confidence interval and then convert back to the original metric prior to reporting the results. Formulas for these transformations are widely available (Fleiss, 1981; Kleinbaum et al., 1982; Rothman, 1986a; Selvin, 1991) and are applied as a matter of course by computer programs. Second, even when these transformation are applied, for most indices the result is still approximate. While the approximation yields an interval that is accurate enough for most purposes, it is not identical to the interval obtained by more sophisticated procedures (see, for example, Cornfield, 1956; Fleiss, 1979, 1981). The assertion that the 95% confidence interval will exclude the null if and only if the significance test is significant at 0.05 assumes that exact formulas (or the identical approximation) are used for computing the p-value and the confidence interval.

3.14.10 STUDY PLANNING: PRECISION ANALYSIS VS. POWER ANALYSIS

If the goal of a study is to yield a test of significance, it follows that the study should have adequate power to yield a significant effect. By the same logic, if the goal of a study is to provide an estimate of effect size, it follows that the study should be able to provide this estimate with an acceptable degree of precision. While the research community is clearly moving toward effect size estimation for data analysis, relatively little has been said about the need for a corresponding shift in the mechanisms used for study planning. In fact, it is not unusual to see a paper that focuses on effect size in the analysis but whose sample size was set on the basis of a power analysis. There is neither a mathematical nor a logical basis for this mixture of apples and estimates. The two goals, significance tests and effect size estimation, have in common that they are both things we can do with data. Beyond this, however, the two criteria are quite different.

Power addresses the likelihood that a study will yield a significant effect. As such, it is controlled by effect size, sample size, and alpha—the same elements that control statistical significance. Precision addresses the width of the confidence interval surrounding the effect size estimate. As such, it is controlled by the confidence level and the sample size—the same elements that control the confidence interval width. Of particular note, effect size is the dominant factor in computation of power but plays little (if any) role in the computation of precision. The sample size required to yield adequate power and the sample size required to yield adequate precision will differ from each other in almost all cases, and often by very substantial amounts.

3.14.10.1 Planning for Precision

To determine the study's precision during the process of study planning, we may "plug in" the sample size and confidence level, and compute the confidence interval width that will result. As was the case for power analysis, this process is facilitated by computer programs designed for this purpose.

Figure 8 is an example of this process using the program cited earlier for power analysis. On this panel the researcher has entered response rates of 60% and 40% for the new treatment and the standard treatment, respectively, and a sample size of 190 patients per group. The program displays the effect size as a rate difference of 20 points. This effect will be reported with a 95% confidence interval of plus/minus 10 points (0.10–0.30).

The researcher may modify the sample size—increases to sample size will yield a more precise estimate. The researcher may modify the confidence level—setting the level at 90% rather than 95% will yield a narrower interval. If the study index is a mean difference, the researcher may modify the standard deviation (SD) within groups—a smaller SD will yield a narrower interval. The researcher may also modify the effect size but, as noted earlier, the confidence interval width will be affected only slightly, if at all.

It should be noted that Figure 8 is an extension of the example cited earlier for power (Figure 3). In that case the researcher had determined that a sample of 107 patients per group would yield power of 80%, and the confidence interval is shown as 0.07–0.33. When planning for precision the researcher has increased the sample size to 190 to yield a more precise estimate of the effect size, and power has increased as well, from 80% to 97%.

As it did for power, the program will generate a table or graph of precision that allows us to quickly take account of the larger picture (Figure 9). With 50 patients per group we will be able to report the rate difference with a precision (95% confidence interval) ± 20 percentage points. With 100 patients per group we will report the rate difference with a precision of ±13 percentage points. With 200 patients per group we will report the rate difference with a precision of ±10 percentage points.

The program will also generate a summary of the analysis. This report reads (in part) "A second goal of this study is to estimate the difference between the two populations ... The study will enable us to report the difference in proportions with a precision (95.0% confidence level) of approximately ± 0.098 points. Specifically, an observed difference of 0.200 would be reported with a 95.0% confidence interval of 0.100–0.296. The precision estimated here is the approximate expected precision. Precision will vary as a function of the observed proportions (as well as sample size), and in any single study will be narrower or wider than this estimate."

In this example the power analysis led to a sample size requirement of 107 per group while the precision analysis led to a sample size of 190 per group. In other studies the relationship between the two sample sizes will be different, and could even be reversed (with the sample for power being larger than the sample for precision). The key point is that the two goals of significance testing and effect size estimation are distinct. Precision analysis focuses on the precision of the estimate rather than the power to test the null hypothesis, and as such is the appropriate method to use when the study goal is to estimate the magnitude of the treatment effect (Bristol, 1989; Burnand et al., 1990; Cobb, 1985; Feinstein, 1990; Gordon, 1987;

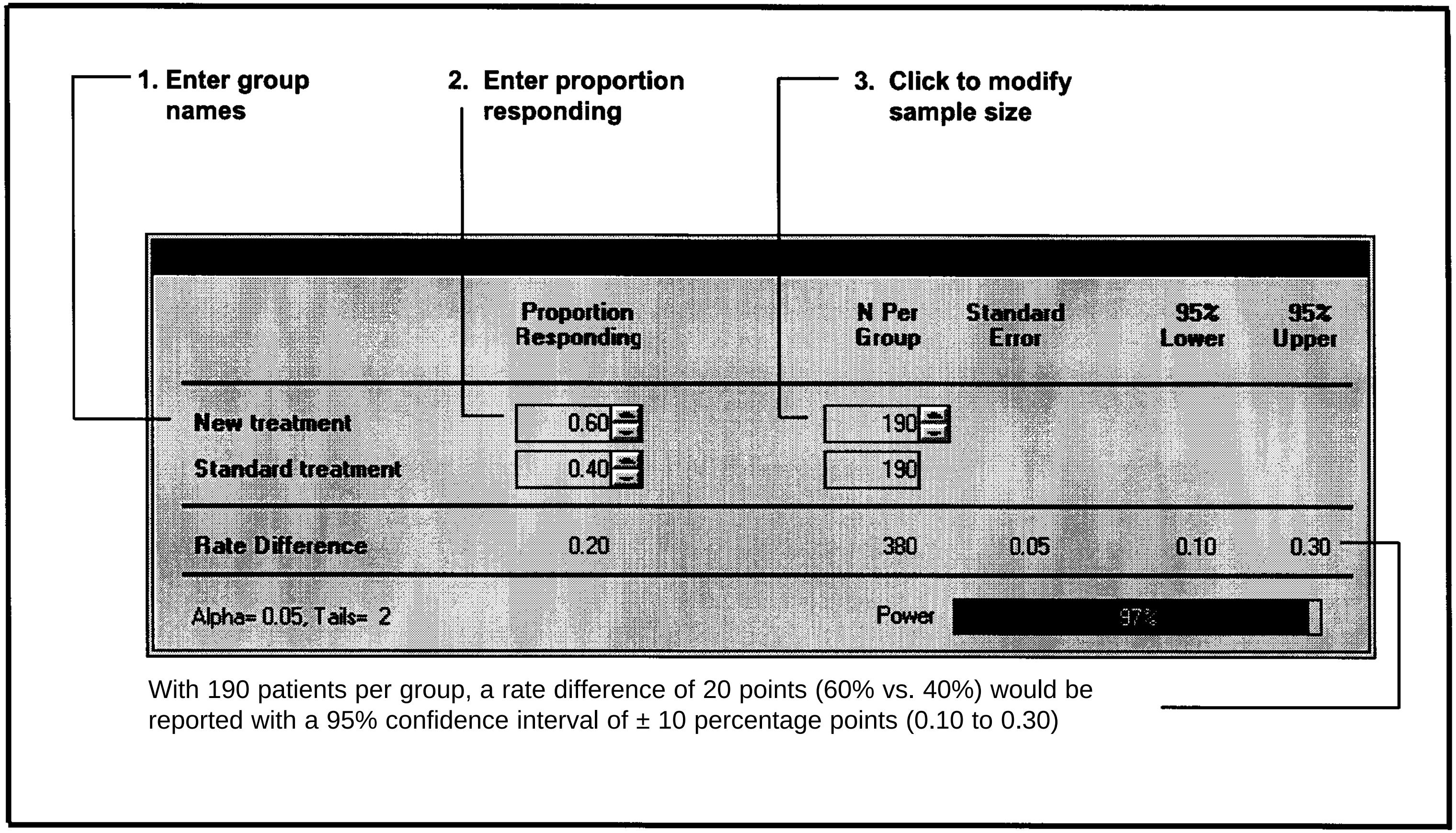

1. Enter group names

2. Enter proportion responding

3. Click to modify sample size

	Proportion Responding	N Per Group	Standard Error	95% Lower	95% Upper
New treatment	0.60	190			
Standard treatment	0.40	190			
Rate Difference	0.20	380	0.05	0.10	0.30

Alpha= 0.05, Tails= 2

Power 97%

With 190 patients per group, a rate difference of 20 points (60% vs. 40%) would be reported with a 95% confidence interval of ± 10 percentage points (0.10 to 0.30)

Figure 8 Screen from power and precision showing computation of precision. The effect size is shown as a rate difference of 20 points with a 95% confidence interval of 0.10–0.30. The researcher may modify any of the study parameters such as sample size and confidence level and immediately see the impact on the confidence interval width.

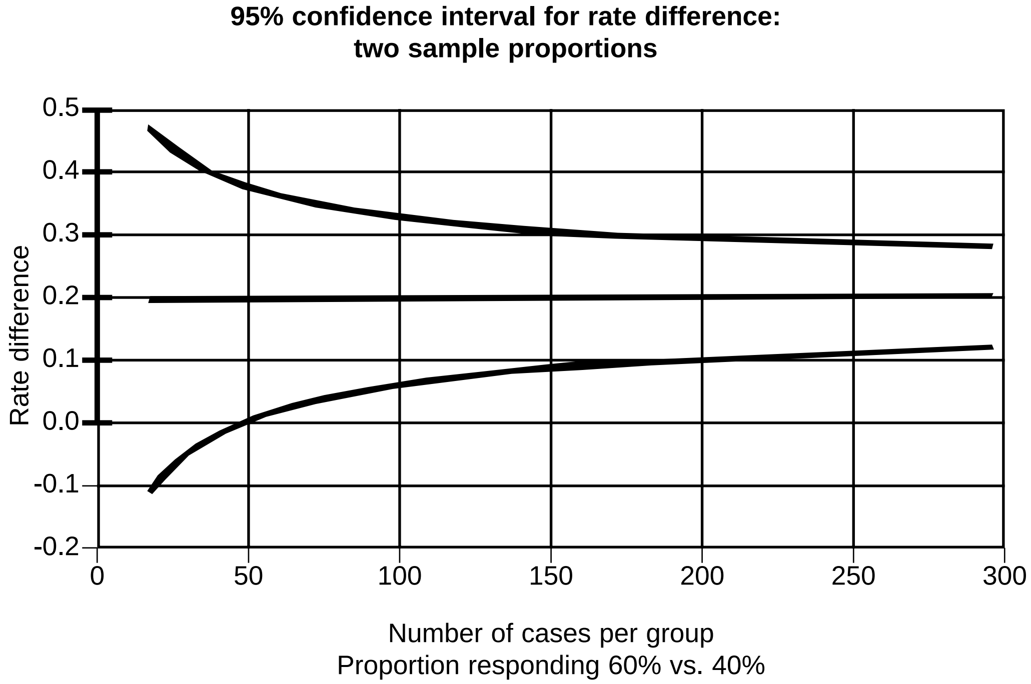

Figure 9 Planning for precision. With a sample of 50 cases per group the treatment effect will be reported with a 95% confidence interval some 40 percentage points wide. With 100 cases per group the interval will narrow to 30 points, and with 200 per group the interval will narrow to 20 points.

Greenland, 1988; Oakes, 1986; Smith & Bates, 1992; Walter, 1991).

Of course, even a fairly narrow interval may include some effects that are clinically meaningful and others that are not, but as the interval narrows in width it becomes increasingly likely that the study results will provide the information required for an informed clinical decision. In many cases, the sample size required for a precise estimate will exceed the available resources. For example, if only 50 cases per group are available, then the 95% interval will be nearly 40 points wide. In this case the study could still provide data that would be pooled with data from other studies, in a meta-analysis.

The method presented here for planning for precision shows provides an estimate of the confidence interval width, but the actual confidence interval in any given study will almost always be narrower or wider than the width estimated in this way. When working with means, for example, the confidence intervals are based in part on the sample standard deviation, which will vary from one sample to the next. When working with proportions, correlations, or survival data the confidence intervals are based in part on the absolute effect size obtained which will vary from one sample to the next. The procedure outlined here is based on the assumption that the sample effect will mirror the population effect precisely. As such, it yields the median confidence interval width for some procedures and the expected (mean) width for others. Computer programs use more complex formulas that take account of the sampling distribution of the confidence intervals. As such they may allow us to report, for example, that "With a sample of 50 cases per group and a standard deviation of 10 rating points within groups, the median width of the confidence interval will be 7.9 points. In 80% of studies, the confidence interval will be no wider than 8.4 points." If the researcher needs to ensure that the confidence interval width will fall within a specific range, then these formulas should be employed (Borenstein et al., 1997; Elashoff, 1997; Hahn & Meeker, 1991).

3.14.11 META-ANALYSIS

Meta-analysis is the process of developing a comprehensive picture of a field of research by performing analyses on summary data from individual studies, and as such is the next logical step in the research process. The thesis of this chapter has been that the goal of research should be to identify the size of a treatment effect. The single study, because it includes a limited number of subjects, typically yields an imprecise estimate. By combining data from multiple studies we effectively increase the sample size and yield a more precise estimate of the effect. Additionally, the single study typically includes one treatment and one treatment population. The meta-analysis, by contrast, may synthesize data from multiple treatments or multiple populations, and as such may be used to identify the type of treatment that will work best for a specific type of patient. Meta-analyses are being used to provide treatment recommendations based on the results of completed studies. They are also being used to identify areas where the questions of interest have not yet been answered, and to help shape the development of new studies.

The relatively widespread acceptance of meta-analysis since the late 1980s has important implications for the topic of this chapter in a more direct sense as well. The thesis of this chapter is that in psychological research the null hypothesis is rarely true, and what appear to be "conflicting results" (one study is statistically significant while another is not) often reflects normal sample-to-sample variation in effect size coupled with variations in sample size. In past decades this argument had been made in the abstract. With the advent of meta-analysis it is possible to actually "see" the bigger picture. Almost invariably, it becomes clear that the null hypothesis is false, and the study-to-study variations can be seen in context. This has been a critical factor in the shift toward effect size estimation (Cook et al., 1992; Cooper, 1989; Cooper & Hedges, 1994; Eagly & Wood, 1994; Hedges, 1990; Hedges & Olkin, 1985; Schmidt & Hunter 1997; Shadish, 1992).

3.14.12 CONCLUSIONS

Researchers typically want to know whether or not a treatment has any impact, and also what the size of this impact might be. Researchers working exclusively with significance tests are able to address only the first of these issues, and leave the second unanswered. Worse, because the issue of effect magnitude is such a compelling one, the significance tests are often pressed into service to address this issue as well, which often leads to study results—whether statistically significant or not—being misinterpreted. Statistical significance is equated with clinical significance, though this interpretation may not be warranted. Absence of statistical significance is taken as evidence that a treatment is not effective, though this interpretation is rarely justified.

By contrast, estimates of effect size bounded by confidence intervals provide the kind of information that researchers and clinicians require. The report of effect size focuses attention on the magnitude of the effect, and the confidence intervals tell us how much confidence we can have that the population effect lies within a given range. The data may allow us to conclude that the treatment is clearly not effective, or that it clearly is effective, or that additional information is required. Critically, though, this type of report focuses attention on the relevant issues. If and when we elect to make a dichotomous decision, the format of the presentation helps to ensure that the meaning of this decision is clearly understood.

There is an ongoing debate about the role of the single study as part of the research process. Some take the position that any single study should allow researchers to conclude that a specific effect size is or is not tenable. Others have argued that the single study cannot provide this kind of information reliably, and that hypothesis testing should be deferred to the time that the data from many studies can be synthesized in a meta-analysis.

This debate has carried over to the role of confidence intervals. There is complete agreement that researchers should present the effect size bounded by confidence intervals since this will focus attention of the issue of interest and will help to prevent mistakes of interpretation. However, there is sharp disagreement over how the confidence interval should be interpreted. Some take the position that confidence intervals should serve as a surrogate for the significance test, and the focus should be in the question of whether or not the interval includes a specific value. When used in this way the shift from significance tests to confidence intervals is one of format rather than substance—it helps to focus attention on the proper issue and to avoid mistakes of interpretation. Others feel that this approach will merely codify the kinds of errors inherent in significance tests. They argue that the key element in a shift to confidence intervals should be the abandonment of significance testing, and that one should refrain from any discussion of whether or not specific points fall within the interval.

The issues addressed in study planning should correspond to the issues that will be addressed in the data analyses. If the goal of a study is to yield a test of significance, the study should have adequate power to yield a significant effect. Similarly, if the goal of a study is to provide an estimate of effect size that can stand on its own, the study should be able to provide this estimate with an acceptable degree of precision. Prior to ensuring that the power analysis is performed properly it behooves the researcher to ensure that it makes sense to test the null, and indeed that the study should be carried out at all.

The argument that effect size estimation and confidence intervals should replace significance testing in behavioral research is not new but the research community is beginning to take concrete steps in this direction. This shift has been driven by the advent of meta-analysis. Meta-analyses have served to highlight the kinds of errors that are fostered by the use of significance tests, and at the same time provide a vehicle for estimating treatment effects precisely. Additionally, they provide a vehicle for moving beyond the single study and for synthesizing data from multiple sources, which is the next logical step in the research process.

ACKNOWLEDGMENTS

This work was supported in part by the following grants: NIMH/SBIR MH-43083, NIMH/SBIR MH-52969, NIMH/SBIR MH 558483, and NIMH 41960. I am grateful for their comments to Hannah Rothstein, Nina Schooler, Kenneth Rothman, and the late Jacob Cohen, to whom this chapter is dedicated.

3.14.13 REFERENCES

Abelson, R. P. (1995). *Statistics as principled argument.* Mahwah, NJ: Erlbaum.

Abelson, R. P. (1997a). A retrospective on the significance test ban of 1999 (if there were no significance tests, they would be invented). In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 117–144). Mahwah, NJ: Erlbaum.

Abelson, R. P. (1997b). On the surprising longevity of flogged horses: Why there is a case for the significance test. *Psychological Science, 8*(1), 12–15.

Altman, D. G. (1991). Statistics in medical journals: Developments in the 1980s. *Statistics in Medicine, 10,* 1897–1913.

Altman, D. G., Gore, S. M., Gardner, M. J., & Pocock, S. J. (1983). Statistical guidelines for contributors to medical journals. *British Medical Journal, 286,* 1489–1493.

Bailar, J. C. (1997). The promise and problems of meta analysis. *New England Journal of Medicine, 337,* 559–561.

Bailar, J. C. (1998). Letter. *New England Journal of Medicine, 338,* 62.

Bakan, D. (1966). The test of significance in psychological research. *Psychological Bulletin, 66,* 423–437.

Berkson, J. (1938). Some difficulties of interpretation encountered in the application of the chi-square test. *Journal of the American Statistical Association, 33,* 526–542.

Berkson, J. (1942). Tests of significance considered as evidence. *Journal of the American Statistical Association, 37,* 325–335.

Birnbaum, A. (1961). Confidence curves: An omnibus technique for estimation and testing statistical hypotheses. *Journal of the American Statistical Association, 56,* 246–249.

Blackwelder, W. C. (1982). "Proving the null hypothesis" in clinical trials. *Controlled Clinical Trials, 3,* 345–353.

Borenstein, M. (1994a). Planning for precision in survival studies. *Journal of Clinical Epidemiology, 47*(11), 1277–1285.

Borenstein, M. (1994b). The case for confidence intervals in controlled clinical trials. *Controlled Clinical Trials, 15,* 411–428.

Borenstein, M., & Cohen, J. (1988). *Statistical power analysis: A computer program.* Hillsdale, NJ: Erlbaum.

Borenstein, M., Cohen, J., Rothstein, H., Pollack, S., & Kane, J. (1990). Statistical power analysis for one-way analysis of variance: A computer program. *Behavior Research Methods, Instruments, and Computers, 22,* 271–282.

Borenstein, M., Cohen, J., Rothstein, H., Pollack, S., & Kane, J. (1992). A visual approach to statistical power analysis on the microcomputer. *Behavior Research Methods, Instruments & Computers, 24,* 565–572.

Borenstein, M., Rothstein, H., & Cohen, J. (1997). *Power and precision.* Teaneck, NJ: Biostat. http://www.PowerAndPrecision.com

Bristol, D. R. (1989). Sample sizes for constructing confidence intervals and testing hypotheses. *Statistics in Medicine, 8,* 803–811.

Burnand, B., Kernan, W. N., & Feinstein, A. R. (1990). Indexes and boundaries for "quantitative significance" in statistical decisions. *Journal of Clinical Epidemiology, 43,* 1273–1284.

Carver, R. P. (1978). The case against statistical significance testing. *Harvard Educational Review, 48,* 378–399.

Chow, S. L. (1988). Significance test or effect size? *Psychology Bulletin, 103,* 105–110.

Chow, S. L. (1989). Significance tests and deduction: Reply to Folger (1989). *Psychological Bulletin, 106,* 161–165.

Cobb, E. B. (1985). Planning research studies: An alternative to power analysis. *Nursing Research, 34,* 386–388.

Cohen, J. (1962). The statistical power of abnormal–social psychological research: A review. *Journal of Abnormal and Social Psychology, 65*(3), 145–153.

Cohen, J. (1965). Some statistical issues in psychological research. In B. B. Wolman (Ed.), *Handbook of clinical psychology* (pp. 95–121). New York: McGraw-Hill.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences.* New York: Academic Press.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J. (1990). Things I have learned (so far). *American Psychologist, 45,* 1304–1312.

Cohen, J. (1992). A power primer. *Psychology Bulletin, 112,* 155–159.

Cohen, J. (1994). The Earth is round ($p < 0.05$). *American Psychologist, 49,* 997–1003.

Cohen, J. (1997). The earth is round ($p < .05$). In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 21–36). Mahwah, NH: Erlbaum.

Cole, P., & Goldman, M. B. (1975). Occupation. In J. F. J. Fraumeni (Ed.), *Persons at high risk of cancer: An approach to cancer etiology and control* (pp. 167–183). New York: Academic Press.

Cook, T. D., Cooper, H., Cordray, D. S., Hartman, H., Hedges, L. V., Light, L. V., Louis, T. A., & Mosteller, F. (1992). *Meta-analysis for explanation.* New York: Russell Sage Foundation.

Cooper, H. (1989). *Integrating research: A guide for literature reviews* (2nd ed.). Newbury Park, CA: Sage.

Cooper, H., & Hedges, L. V. (1994). Research synthesis as a scientific enterprise. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 3–14). New York: Russell Sage Foundation.

Cornfield, J. (1956). A statistical problem arising from retrospective studies. In J. Neyman (Ed.), *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* (pp. 135–148). Berkeley, CA: University of California Press.

Dar, R. (1987). Another look at Meehl, Lakatos, and the scientific practices of psychologists. *American Psychologist, 42,* 145–151.

Derouen, T. A. (1987). Letter. *American Journal of Public Health, 77,* 237.

Detsky, A. S., & Sackett, D. L. (1985). When was a "negative" clinical trial big enough? How many patients you need depends on what you found. *Archives of International Medicine, 145,* 709–712.

Detsky, A. S., & Sackett, D. L. (1986). Establishing therapeutic equivalency: What is clinically significant difference? *Archives of International Medicine, 146,* 861–862.

Eagly, A. H., & Wood, W. (1994). Using research synthesis to plan future research. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 485–500). New York: Russell Sage Foundation.

Elashoff, J. (1997). *nQuery.* Cork, Ireland: Statistical Solutions.

Estes, W. K. (1997). Significance testing in psychological research: Some persisting issues. *Psychological Science, 8*(1), 18–20.

Feinstein, A. R. (1975). Clinical biostatistics XXXIV: The other side of "statistical significance": Alpha, beta, delta, and the calculation of sample size. *Clinical Pharmacology and Therapeutics, 18,* 491–505.

Feinstein, A. R. (1976). Clinical biostatistics XXXVII. Demeaned errors, confidence games, nonplussed minuses, inefficient coefficients, and other statistical disruptions of scientific communication. *Clinical Pharmacology and Therapeutics, 20,* 617–631.

Feinstein, A. R. (1977). Clinical biostatistics XL: Stochastic significance, consistency, apposite data, and some other remedies for the intellectual pollutants of statistical vocabulary. *Clinical Pharmacology and Therapeutics, 22,* 113–123.

Feinstein, A. R. (1990). The unit fragility index: An additional appraisal of "statistical significance" for a contrast of two proportions. *Journal Clinical Epidemiology*, 43, 201–209.

Fisher, R. A. (1959). *Statistical methods and scientific inference* (2nd ed.). Edinburgh, UK: Oliver & Boyd.

Fleiss, J. (1979). Confidence intervals for the odds ratio in case-control studies: The state of the art. *Journal of Chronic Diseases, 32,* 69–77.

Fleiss, J. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: Wiley.

Fleiss, J. L. (1986a). Confidence intervals vs. significance tests: Quantitative interpretation (letter). *American Journal Public Health, 76,* 587.

Fleiss, J. L. (1986b). Significance tests have a role in epidemiologic research: Reactions to A. M. Walker. *American Journal Public Health, 76,* 559–560.

Foster, D. A., & Sullivan, K. M. (1987). Computer program produces *p*-value graphics (letter). *American Journal Public Health, 77,* 880–881.

Freiman, J. A., Chalmers, T. C., Smith, H., Jr., & Kuebler, R. R. (1978). The importance of beta, the type-II error, and sample size in the design and interpretation of the randomized control trial. Survey of 71 "negative" trials. *New England Journal Medicine, 299,* 690–694.

Friedman, S. B., & Phillips, S. (1981). What's the difference? Pediatric residents and their inaccurate concepts regarding statistics. *Pediatrics, 68,* 644–646.

Gardner, M. J., & Altman, D. G. (1986). Confidence intervals rather than *P* values: Estimation rather than hypothesis testing. *British Medical Journal, 292,* 746–750.

Gardner, M. J., & Altman, D. G. (1989a). *Confidence intervals analysis.* London: British Medical Journal.

Gardner, M. J., & Altman, D. G. (1989b). *Statistics with confidence—confidence intervals and statistical guidelines.* London: British Medical Journal.

Goldstein, R. (1989). Power and sample size via MS/PC–DOS computers. *American Statistician, 43,* 253–260.

Gonzalez, R. (1994). The statistics ritual in psychological research. *Psychological Science, 5,* 321–328.

Gordon, I. (1987). Sample size estimation in occupational mortality studies with use of confidence interval theory (see comments). *American Journal of Epidemiology, 125,* 158–162.

Greenland, S. (1988). On sample-size and power calculations for studies using confidence intervals. *American Journal of Epidemiology, 128,* 231–237.

Hagen, R. L. (1997). In praise of the null hypothesis significance test. *American Psychologist, 52,* 15–24.

Hahn, G. J., & Meeker, W. Q. (1991). *Statistical intervals.* New York: Wiley.

Harlow, L. L., Mulaik, S. A., & Steiger, J. H. (1997). *What if there were no significance tests?* Mahwah, NJ: Erlbaum.

Harris, E. K., & Albert, A. (1991). *Survivorship analysis for clinical studies.* New York: Marcel Dekker.

Harris, R. J. (1997a). Reforming significance testing via three-valued logic. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 145–174). Mahwah, NJ: Erlbaum.

Harris, R. J. (1997b). Significance tests have their place. *Psychological Science, 8*(1), 8–11.

Hartung, J., Cottrell, J. E., & Giffen, J. P. (1983). Absence of evidence is not evidence of absence. *Anesthesiology, 58,* 298–300.

Hedges, L. V. (1990). Directions for future methodology. In K. W. Wachter & M. L. Straf (Eds.), *The future of meta-analysis* (pp. 11–26). New York: Russell Sage Foundation.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Boston: Academic Press.

Hoffman, D., & Wynder, E. L. (1976). Smoking and occupational cancers. *Preventative Medicine, 5,* 245–261.

Hunter, J. E. (1997). Needed: A ban on the significance test. *Psychological Science, 8*(1), 3–7.

International Committee of Medical Journal Editors (1991). Uniform requirements for manuscripts submitted to biomedical journals. *New England Journal of Medicine, 324,* 424–428.

Kane, J. M., & Borenstein, M. (1985). Compliance in the long-term treatment of schizophrenia. *Psychopharmacology Bulletin, 21,* 23–27.

Kleinbaum, D. G., Kupper, L. L., & Morgenstern, H. (1982). *Epidemiologic research: Principles and quantitative methods.* Belmont, CA: Lifetime Learning Publications.

Kraemer, H. C., & Thiemann, S. (1987). *How many subjects? Statistical power analysis in research.* Newbury Park, CA: Sage.

Lachenbruch, P. A., Clark, V. A., Cumberland, W. G., Chang, P. C., Afifi, A. A., Flack, V. F., & Elashoff, R. M. (1987). Letter. *American Journal of Public Health, 77,* 237.

Lau, J., Antman, E. M., Jimenez Silva, J., Kupelnick, B., Mosteller, F., & Chalmers, T. C. (1992). Cumulative meta-analysis of therapeutic trials for myocardial infarction. *New England Journal of Medicine, 327,* 248–254.

Lawless, J. F. (1982). *Statistical models and methods for lifetime data.* New York: Wiley.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment. Confirmation from meta-analysis. *American Psychologist, 48,* 1181–1209.

Loftus, G. R. (1991). On the tyranny of hypothesis testing in the social sciences. *Contemporary Psychology, 36,* 102–105.

Lykken, D. T. (1968). Statistical significance in psychological research. *Psychology Bulletin, 70,* 151–159.

Mainland, D. (1982). Medical statistics—thinking vs. arithmetic. *Journal of Chronic Diseases, 35,* 413–417.

Makuch, R. W., & Johnson, M. F. (1986). Some issues in the design and interpretation of "negative" clinical studies. *Archives of Internal Medicine, 146,* 986–989.

Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science, 34,* 103–115.

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress in soft psychology. *Journal of Consulting and Clinical Psychology, 46,* 806–834.

Meehl, P. E. (1990). Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports, 66,* 195–244.

Morgan, P. P. (1989). Confidence intervals: From statistical significance to clinical significance. *Canadian Medical Association Journal, 141,* 881–883.

Murphy, K. R. (1990). If the null hypothesis is impossible, why test it? *American Psychology, 45,* 403–404.

Nelson, N., Rosenthal, R., & Rosnow, R. L. (1986). Interpretation of significance levels and effect sizes by psychological researchers. *American Psychologist, 41,* 1299–1301.

Neyman, J., & Pearson, E. S. (1932a). On the use and interpretation of certain test criteria for purposes of statistical inference: Part II. *Biometrika, 20A,* 263–294.

Neyman, J., & Pearson, E. S. (1932b). On the use and interpretation of certain test criteria for purposes of statistical inference: Part 1. *Biometrika, 20A,* 175–240.

Oakes, M. (1986). *Statistical inference: A commentary for the social and behavioral sciences.* New York: Wiley.

Peto, R., & Doll, R. (1977). When is significant not significant. *British Medical Journal, 2,* 259.

Phillips, W. C., Scott, J. A., & Blasczcynski, G. (1983). The significance of "No significance": What a negative statistical test really means. *American Journal of Roentgenology, 41,* 203–206.

Poole, C. (1987a). Beyond the confidence interval. *American Journal of Public Health, 77,* 195–199.

Poole, C. (1987b). Confidence intervals exclude nothing. *American Journal of Public Health, 77,* 492–493.

Poole, C. (1987c). Mr. Poole's response (letter). *American Journal of Public Health, 77,* 880.

Reed, J. F., & Slaichert, W. (1981). Statistical proof in inconclusive "Negative" trials *Archives of Internal Medicine, 141,* 1307–1310.

Reynolds, T. B. (1980). Type II error in clinical trials (Editor's reply to letter). *Gastroenterology, 79,* 180.

Rosenthal, R., & Gaito, J. (1963). The interpretation of levels of significance by psychological researchers. *Journal of Psychology, 55,* 33–38.

Rosenthal, R., & Gaito, J. (1964). Further evidence for the cliff effect in the interpretation of levels of significance. *Psychological Reports, 15,* 570.

Rosner, B. (1990). *Fundamentals of biostatistics* (3rd ed.). Boston: PWS-Kent Publishing.

Rosnow, R. L., & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychology, 44,* 1276–1284.

Rossi, J. (1990). Statistical power of psychological research: What have we gained in 20 years? *Journal of Consulting and Clinical Psychology, 58,* 646–656.

Rossi, J. (1997). A case study in the failure of psychology as a cumulative science: The spontaneous recovery of verbal learning. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 175–198). Mahwah, NJ: Erlbaum.

Rothman, K. J. (1978). A show of confidence (letter). *New England Journal of Medicine, 299,* 1362–1363.

Rothman, K. J. (1986a). *Modern epidemiology.* Boston: Little, Brown and Company.

Rothman, K. J. (1986b). Significance questing (editorial). *Annals of Internal Medicine, 105,* 445–447.

Rothman, K. J., & Yankauer, A. (1986). Editors' note. *American Journal of Public Health, 76,* 587–588.

Rothstein, H., Borenstein, M., Cohen, J., & Pollack, S. (1990). Statistical power analysis for multiple regression/correlation: A computer program. *Educational and Psychological Measurement, 50,* 819–830.

Rozeboom, W. W. (1960). The fallacy of the null hypothesis significance test. *Psychological Bulletin, 57,* 416–428.

Savitz, D. (1987). Letter. *American Journal of Public Health, 77,* 237–238.

Scarr, S. (1997). Rules of evidence: A larger context for the statistical debate. *Psychological Science, 8*(1), 16–17.

Schmidt, F. L. (1992). What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist, 47,* 1173–1181.

Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. *Psychological Methods, 1,* 115–129.

Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods, 1,* 199–223.

Schmidt, F. L., & Hunter, J. E. (1997). Eight common but false objections to the discontinuation of significance testing in the analysis of research data. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 37–64). Mahwah, NJ: Erlbaum.

Sedlmeier, P., & Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies? *Psychological Bulletin, 105,* 309–316.

Selikoff, I. J., Hammond, E. C., & Churg, J. (1968). Asbestos exposure, smoking and neoplasia. *Journal of the American Medical Association, 204,* 106–112.

Selvin, S. (1991). *Statistical analysis of epidemiologic data.* New York: Oxford University Press.

Shadish, W. R. (1992). Do family and marital psychotherapies change what people do? A meta-analysis of behavioral outcomes. In T. D. Cook & H. Cooper (Eds.), *Meta-analysis for explanation* (pp. 129–208). New York: Russell Sage.

Shaver, J. P. (1993). What statistical significance testing is, and what it is not. *Journal of Experimental Education, 61,* 293–316.

Sheehe, P. R. (1993). A variation on a confidence interval theme (letter). *Epidemiology, 4,* 185–186.

Shrout, P. E. (1997). Should significance tests be banned? Introduction to a special section exploring the pros and cons. *Psychological Science, 8,* 1–2.

Simon, R. (1986). Confidence intervals for reporting results of clinical trials. *Annals of Internal Medicine, 105,* 429–435.

Simon, R. (1987). The role of overviews in cancer therapeutics. *Statistics in Medicine, 6,* 389–396.

Smith, A. H., & Bates, M. N. (1992). Confidence limit analyses should replace power calculations in the interpretation of epidemiologic studies. *Epidemiology, 3,* 449–452.

Smith, A. H., & Bates, M. N. (1993). A variation on a confidence interval theme (Reply to letter). *Epidemiology, 4,* 186–187.

Sullivan, K. M., & Foster, D. A. (1990). Use of the confidence interval function. *Epidemiology, 1,* 39–42.

Thomas, L., & Krebs, C. (1997). A comprehensive list of power analysis software for microcomputers. http://www.Interchg.ubc.ca/cacb/power.

Thompson, B. (1992). Two and one-half decades of leadership in measurement and evaluation. *Journal of Counseling and Development, 70,* 434–438.

Thompson, W. D. (1987a). Exclusion and uncertainty (letter). *American Journal of Public Health, 77,* 879–880.

Thompson, W. D. (1987b). Letter. *American Journal of Public Health, 77,* 238.

Thompson, W. D. (1987c). On the comparison of effects. *American Journal of Public Health, 77,* 491–492.

Thompson, W. D. (1987d). Statistical criteria in the interpretation of epidemiologic data (published erratum appears in *American Journal of Public Health,* 1987(Apr), 77(4), 515). *American Journal of Public Health, 77,* 191–194.

Tukey, J. W. (1969). Analyzing data: Sanctification or detective work? *American Psychologist, 24,* 83–91.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin, 76,* 105–110.

Walker, A. M. (1986a). Reporting the results of epidemiologic studies. *American Journal of Public Health, 76,* 556–558.

Walker, A. M. (1986b). Significance tests represent consensus and standard practice (letter). *American Journal of Public Health, 76,* 1033.

Walter, S. D. (1991). Statistical significance and fragility criteria for assessing a difference of two proportions. *Journal of Clinical Epidemiology, 44,* 1373–1378.

Wonnacott, T. (1985). “Statistically significant.” *Canadian Medical Association Journal, 133,* 843.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

3.15

Meta-analytic Research Synthesis

SHARON H. KRAMER and ROBERT ROSENTHAL
Harvard University, Cambridge, MA, USA

3.15.1 INTRODUCTION 352

3.15.1.1 A Brief History 352

3.15.2 WHY DO A META-ANALYSIS? 352

3.15.2.1 The Problem of Poor Cumulation 352
3.15.2.2 Primary Purposes of Meta-analyses 353
3.15.2.3 Pooling Pilots 353

3.15.3 CONDUCTING A META-ANALYSIS 353

3.15.3.1 Formulating the Question 353
3.15.3.2 Defining Criteria for Inclusion 354
3.15.3.3 Searching the Literature 354
3.15.3.4 Recording Study Characteristics and Identifying Moderators 355
3.15.3.5 Coding 355
3.15.3.6 Descriptive Data Displays 356
3.15.3.6.1 Stem and leaf 356
3.15.3.6.2 Box plot and summary table 356

3.15.4 QUANTITATIVELY ANALYZING THE STUDIES 356

3.15.4.1 Comparing Two Studies 356
3.15.4.1.1 Significance testing 356
3.15.4.1.2 Effect size estimation 358
3.15.4.2 Combining Two Studies 360
3.15.4.2.1 Significance testing 360
3.15.4.2.2 Effect size estimation 360
3.15.4.3 Comparing Three or More Studies: Diffuse Test 360
3.15.4.3.1 Significance testing 361
3.15.4.3.2 Effect size estimation 361
3.15.4.3.3 Comparing three or more studies: focused tests 361
3.15.4.3.4 Significance levels 361
3.15.4.3.5 Effect size estimation 362
3.15.4.4 Combining Three or More Studies 362
3.15.4.4.1 Significance testing 362
3.15.4.4.2 Effect size estimation 363
3.15.4.5 Weighting Studies 363

3.15.5 SPECIAL CASES 364

3.15.5.1 Imprecisely Reported p*'s* 364
3.15.5.2 Repeated Measures Designs 364
3.15.5.3 Random vs. Fixed Effects 364

3.15.6 CRITICISMS OF META-ANALYSIS 365

3.15.6.1 Sampling Bias and the File Drawer Problem 365
3.15.6.2 Combining Apples and Oranges: Problems of Heterogeneity 365
3.15.6.3 Problems of Independence 366
3.15.6.3.1 Nonindependence of subjects 366
3.15.6.3.2 Nonindependence of experimenters 366
3.15.6.4 Exaggerated Significance Levels 366
3.15.6.5 Too Small Effects 366

3.15.6.6 Abandonment of Primary Studies 367
3.15.7 REFERENCES 367

3.15.1 INTRODUCTION

3.15.1.1 A Brief History

A little over 20 years ago, educational researcher, Gene V. Glass was interested in the question of the whether psychotherapy works. In other words, does it help patients any more than would a placebo treatment? Furthermore, if in fact psychotherapy overall is effective, can it be determined whether one form is more effective than another? Glass was not the first researcher interested in this question. Hundreds of studies had previously been conducted in this area. Yet the results from these studies, and even from reviews which summarized the studies, had reached contradictory conclusions. What was unique about Glass' attempt to answer this question was that he was the first to conduct a quantitative synthesis of the research, a procedure which he called a "meta-analysis" (Glass, 1976; Smith, Glass, & Miller, 1980). In conducting the meta-analysis, Glass first located studies which asked this question of whether psychotherapy works and then he quantitatively characterized the outcomes and features of the studies and statistically described his findings. Glass ultimately found that psychotherapy shows positive results and that there is little difference between the different types of therapy.

While Glass is responsible for coining the term and for carrying out the first large-scale meta-analysis, meta-analytic enterprises had previously been undertaken. For example, in 1904, Karl Pearson, the inventor of the Pearson product moment correlation coefficient, had been investigating the degree to which inoculation against smallpox saved lives. He collected the following six correlation coefficients (r's) which measured the relationship of inoculation to survival: 0.58, 0.58, 0.60, 0.63, 0.66, and 0.77, and averaged them, finding a mean correlation of 0.6. Thus, his "meta-analysis" or synthesis of these six studies found a rather large effect of inoculation on survival.

In the 1930s, Ronald Fisher was experimenting with combining significance levels of independent studies (Fisher, 1938) as were Frederick Mosteller and Robert Bush in the 1950s (Mosteller & Bush, 1954). Mosteller and Bush also showed the usefulness of combining effect sizes as well as significance levels. Even in George Snedecor's 1946 statistics textbook is included an example drawn from agriculture of combining correlation coefficients measuring the relationship between the initial weight of steers and their subsequent gain in weight. Finally, back to the field of psychology, in early work on experimenter expectancy effects in the 1960s, there is an example of combining three probability levels to give an overall test of significance of three experiments (Rosenthal, 1966).

Since Glass' use of the term, meta-analyses have become widely used for research in the behavioral and physical sciences and in applied situations, such as for policy making. By one estimate, nearly 2000 meta-analytic reviews have been conducted in the social and health sciences between 1980 and 1994 (Bausell, Li, Gau, & Soeken, 1995). In addition to the growing numbers of meta-analyses that have been conducted, a number of different procedures for conducting meta-analyses have also been described. Indeed, whole books, not to mention numerous individual papers and chapters, have addressed the issue of conducting meta-analyses (e.g., Cooper, 1989; Cooper & Hedges, 1994; Glass, McGaw, & Smith, 1981; Hedges & Olkin, 1985; Hunter & Schmidt, 1990; Light & Pillemer, 1984; Rosenthal, 1991). This chapter will illustrate the basic quantitative procedures for conducting a meta-analysis. This chapter draws on material presented in Rosenthal, (1984, 1991, 1993, 1995a, 1995b, 1998) and in Rosenthal and Rosnow (1991). For more detailed information about the procedures discussed here, the reader may wish to consult these sources. As the chapter will demonstrate, the level of quantitative skill and training required to employ basic meta-analytic procedures is so modest that any researchers capable of analyzing the results of their own research will be capable of learning the small number of calculations required to conduct a first-rate meta-analysis.

3.15.2 WHY DO A META-ANALYSIS?

3.15.2.1 The Problem of Poor Cumulation

A meta-analysis is literally an analysis of analyses or a procedure for quantitatively cumulating research results addressing specific research questions. Unlike a typical literature review, a meta-analysis addresses the problem of "poor cumulation" in the social sciences. It is quite common to find a literature review in which dozens of studies are discussed but which ends with a call for more research. When these studies are discussed within the review, only the statistical significance or nonsignificance of the results, without any information about the size of the effects, are presented. Until somewhat recently, when there existed only a small number of studies in any given research area, it was easier to form conclusions about a body of research, but now, as hundreds of studies have been conducted in many fields in the social sciences, the problem of poor cumulation continues to grow.

Unlike the physical sciences in which newer work builds directly upon the older work, in the social sciences each succeeding volume of our scientific journals almost seems to be starting anew, posing the same research questions time and again. The common explanation for this problem is that there is more variability or less agreement in our field which prevents the kind of building seen in other fields. Yet, this commonly held belief has not been supported. A recent comparison of 13 research areas in particle physics with 13 areas in psychology did not actually find more conflicting results in our field (Hedges, 1987). For the researcher trying to develop theory, for the practitioner trying to decide on treatment, for the policy maker trying to change policy, and for the funding agency trying to allocate research money, meta-analysis holds great promise.

3.15.2.2 Primary Purposes of Meta-analyses

The first primary purpose of a meta-analysis is to summarize for a set of studies the overall relationship between the two variables that had been investigated in each study. It can be used, as with the Glass psychotherapy meta-analysis, for pulling together hundreds of studies, or it can also be used simply by an investigator asking whether a study and its replication produce similar or different results and what the net results of these two studies might be. By quantitatively summarizing the existing research, a meta-analysis is able to discover what has so far been learned and help to discover what has not yet been learned.

Another equally important purpose of meta-analytic work is to learn from comparisons between studies. By comparing the studies in a meta-analysis, the researcher hopes to determine factors that are associated with variations in the magnitudes of the relationships between two variables. These factors are known as moderator variables because they moderate or alter the magnitude of a relationship. For example, a meta-analysis might suggest that the effectiveness of a given type of therapy is moderated by the age of the patient, such that it is more effective with adolescents but less so with adults. By comparing studies, the meta-analysis uncovers new relationships or provides support for suspected relationships, thus paving the way for future research.

3.15.2.3 Pooling Pilots

While meta-analysis is potentially valuable in all fields, it can be particularly helpful for clinical research. Often in clinical research, the number of patients available is so small that all the researcher is able to do is to conduct pilot studies, none of which will ever reach a significance level of $p < 0.05$. However, the researcher is able to combine these individual studies meta-analytically. A recent example of this appeared in an issue of *Science* (Cohen, 1993) which described two pilot studies in which experimental monkeys were vaccinated using SIV (akin to HIV), while control monkeys were not. In the first pilot, six monkeys were available, three vaccinated and three controls. The results showed two of the three vaccinated monkeys and none of the three controls in better health ($p = 0.20$, one-tailed). In the second pilot, 11 animals were available and two of five of the experimentals and none of six of the controls wound up in better health ($p = 0.18$, one-tailed). While the results of neither one of these tiny pilot studies was significant, if treated meta-analytically, they show dramatic benefits of vaccination.

3.15.3 CONDUCTING A META-ANALYSIS

3.15.3.1 Formulating the Question

The first step in conducting a meta-analysis, as with any research enterprise, is a clear formulation of the question being examined. The only constraint on the question is that primary research on the topic must have already been conducted. The specificity and clarity of the question is especially important in a meta-analysis in that it will help the researcher define the criteria for inclusion (see below). The primary question the meta-analysis seeks to answer is: what is the relationship between any variable X and any variable Y? In order to answer this question, we need to have

(i) an estimate of the level of significance of the difference between the obtained effect size and the effect size expected under the null hypothesis (usually an effect size of 0). Estimates of levels of significance are expressed by p values associated with the given significance test, such as t, F, chi-square or Z;

(ii) an estimate of the magnitude of the relationship between X and Y (an effect size). An effect size commonly used by Glass and his colleagues is the number of standard deviation units that separate outcome scores of experimental and control group, expressed as d.

According to Cohen (1977), d's around 0.2 are considered small, around 0.5 are considered medium, and around 0.8 are large. A concrete example of the meaning of an effect size which is drawn from educational research is that a d of 0.2 would raise a student's performance from the 50th percentile to the 58th percentile, a 0.5 would raise it to the 69th, and a 0.8 would raise it to the 79th percentile (Cohen, 1977). Another common effect size estimate is given by r. The relationship between r and d is $r = d \,/\, [(d^2 + 4)^{12}]$.

The relationship between an effect size and a significance test is given by:

$$\text{test of significance} = \text{size of effect} \times \text{size of study} \qquad (1)$$

Table 1 provides several examples of this relationship.

As this basic equation illustrates, as a study increases its sample size, it will obtain more significant results (except in the unusual case where an effect size is truly zero in which case a larger study will not produce a result any more significant than a smaller study). An effect size, however, would not be affected by the size of the sample used in the study.

3.15.3.2 Defining Criteria for Inclusion

Once the research question has been formulated, the next step in the meta-analysis is a careful consideration of what criteria must be met in order for a study to be included in the synthesis. For example, a study comparing individual to group therapy would need to define what constitutes group therapy. Would couples therapy meet these criteria; would therapy which combines group meetings with individual sessions meet these criteria? Deciding on the criteria should be theory or hypothesis driven. For example, if previous research had found that couples therapy is fundamentally different from group therapy, then studies with couples should be excluded from the meta-analysis.

Regardless of the specific question being examined, a basic criterion for inclusion in the meta-analysis is that the studies have reported, or have provided the necessary information for us to obtain, an estimate of an effect size and a significance level for each study. Studies which do not provide this information or a way for the reader to derive this information are not ordinarily included in the analysis.

3.15.3.3 Searching the Literature

There are four basic types of documents from which studies may be retrieved: (i) published books or book chapters; (ii) published journal articles, magazines, newsletters, and newspapers; (iii) bachelor's, master's, and doctoral theses; (iv) unpublished works such as technical reports, convention papers, or raw data. There are a number of computer databases available to help with the retrieval process, such as PsycLit or PsycInfo (see Dickersin, 1994; Reed & Baxter, 1994; Rosenthal, 1994; White, 1994 for specific details on helpful retrieval techniques). Because successful retrieval from these databases is dependent on identifying the right keyword or key subject, the searcher may occasionally miss some references. Therefore, retrieval should also rely on the "ancestry approach," also known as "footnote chasing" (White, 1994), which involves searching reference lists of relevant articles and books. Past issues of relevant

Table 1 Examples of the relationship between tests of significance and effect size: chi-square, Z, t, and F.

Test of significance	=	*Size of effect*	*x*	*Size of study*
$\chi^2_{(1)}$	=	ϕ^2	x	N
Z	=	ϕ	x	$\sqrt{N}$
t	=	$\frac{r}{\sqrt{1-r^2}}$	x	$\sqrt{df}$
t	=	d	x	$\frac{\sqrt{df}}{2}$
F	=	$\frac{r^2}{\sqrt{1-r^2}}$	x	df
t^{a}	=	d	x	$\sqrt{df}$

[a] Correlated observtions.

journals should also be hand searched. Finally, if primary researchers in the field can be contacted, they should be consulted for their knowledge of any unpublished work in the field by themselves or by their colleagues. This last method has been referred to as using the "invisible college" grapevine (Crane, 1972).

A comprehensive search for all the research on the given topic is necessary in order to avoid conducting a biased retrieval of only the major journals which may selectively publish results characterized by lower *p* values. If the field we are searching is too vast for us to analyze every study, it is better to sample randomly from this complete selection than to analyze only the more readily retrievable results.

3.15.3.4 Recording Study Characteristics and Identifying Moderators

In addition to an estimate of effect size and a level of significance, there are always other features of the retrieved studies that will be of interest both for purely descriptive purposes and for analysis as potential moderating variables. For example, the meta-analyst should always record and report the following study characteristics: (i) descriptions of the participant population, such as number, age, sex, education, and volunteer status; (ii) descriptions of how the study was conducted, such as whether it was conducted in a laboratory or in the field, whether it was observational or randomized; and (iii) year and form of publication of the study. In addition to listing each study and its relevant characteristics, it is also good meta-analytic practice to provide the range and median of these features, such as the range and median of dates of the studies and ages of the participants.

All of the study characteristics recorded can also be employed as moderator variables, that is, variables correlated with the magnitude of the effect size obtained for the different studies. For example, a meta-analysis conducted by social psychologist Alice Eagly on sex differences in influenceability found that the time when a study was published was strongly related to the study's results. Studies published before 1970 were somewhat more likely to show greater influenceability among females, while studies published during the period of the women's movement in the 1970s were less likely to find differences between sexes (Eagly & Carli, 1981).

In addition to these general characteristics, each area of research will also have its own specific moderator variables which are of particular interest or meaning. For example, a meta-analysis in the field of the effects of teachers' expectations on pupils' IQ gains specifically looked at the possible moderator variable of how long teachers had known their pupils before the expectations were assigned (Raudenbush, 1994). This analysis found that the longer teachers had known their pupils before the experiment began, the smaller were the effects of experimentally-induced teacher expectations. One method of organizing the various relevant variables that has often proven helpful is to break down the variables into substantive and methodological variables. Substantive variables are those relating to the phenomenon under investigation while methodological variables are those that relate to the procedural aspect of the study (Lipsey, 1994). Some examples of substantive features might be type of treatment administered or characteristics of the subject population. Some methodological variables of interest might be variations in the research design or procedure or internal validity of the study (e.g., random assignment of participants to conditions or use of a control group).

An example of a methodological variable from the study of female influenceability mentioned above that was found to play a large role was the sex of the researcher (Eagly & Carli, 1981). In this meta-analysis, male researchers were more likely to find female influenceability in their studies than female researchers. An explanation for this finding may be that if a difference was found between the sexes, it was equally as likely to be reported by male and female researchers. However, if no difference was found, a male researcher would tend to report the results less often than a female researcher who would be more likely to see the absence of an effect as important in that it disputes the stereotype of females being more conforming and persuasible. After forming a preliminary list of moderators which have been identified, this form should be discussed with colleagues, advisors, and other workers in the field for suggestions of other moderators that should be coded.

3.15.3.5 Coding

Once relevant features have been identified, they need to be coded and entered into a database type of system. Some variables, such as subject sex or number of subjects in the study, can be easily coded (e.g., a 1 for female subjects or a 0 for males). Others, such as an assessment of the quality of the study, will require more sophisticated coding. For these types of codings, each study should be coded by several knowledgeable methodologists who are not necessarily specialists in the field being examined. The agreement or reliability of the judges can then be assessed (for details on how to compute reliability, see Rosenthal, 1991 or Rosenthal & Rosnow, 1991) and when agreement is satisfactory the feature is coded. Table 2 provides an example of a partial database of relevant variables which were coded from a recent meta-analysis examining individual vs. group problem-solving performance (Kramer & Rosenthal, 1997).

3.15.3.6 Descriptive Data Displays

Once all the studies have been collected and the relevant information extracted and recorded from each, the next step is to display what has been found. One of the most important parts of a meta-analysis is a graphic display of the effect sizes and the summary of their distribution and central tendency. We will describe two visual displays that may be used, but there are a great many more to choose from (e.g., Cooper, 1989; Glass et al., 1981; Greenhouse & Iyengar, 1994; Hedges & Olkin, 1985; Light & Pillemer, 1984; Light, Singer, & Willett, 1994; Rosenthal & Rosnow, 1991; Tukey, 1977).

3.15.3.6.1 Stem and leaf

Table 3 is a stem and leaf display (Tukey, 1977) from a recent meta-analysis of the effects of gender on judgments of sexual harassment (Blumenthal, 1997). Each of the 83 effect sizes is recorded with the first digit found in the column labeled "stem" and the second digit found in the column labeled "leaf." The top two entries of Table 3, therefore, are read as two *r*'s of 0.65, and 0.43. The stem and leaf display provides a clear overall impression of all of the results; in this particular case, the display quickly and easily demonstrates that most of the effects of gender differences found in these studies tend to be small. It can also point out when there is an unusual distribution of effect sizes, for example, a bimodal distribution may suggest that a certain innovation works well for one group but not for another.

3.15.3.6.2 Box plot and summary table

The box plot or box graph, originally called a box-and-whisker plot (Tukey, 1977), provides a pictorial summary of five descriptive statistics from the distribution of effect sizes: the maximum, the minimum, the median, the 75th percentile, and the 25th percentile. Figure 1 displays a box plot of the data in Table 3. The top and bottom dots represent the maximum and minimum scores, the top and bottom of the rectangle represent the 75th and 25th percentile and the line dividing the rectangle represents the median. The box plot is especially useful when there are data to display from several subsamples.

In addition to the graphic display, it is also instructive to provide a summary table of these measures of central tendency and variability. Several indices of central tendency should be reported: the unweighted mean effect size, the weighted mean effect size, (see below for more details on weighting), and the median. The number of independent effect sizes on which these indices are based should be reported. The standard deviation, the maximum and minimum effect size, and the effect sizes found at the 75th percentile (Q3) and the 25th percentile (Q1) should also be given. The following may also be reported: the proportion of studies showing effect sizes in the predicted direction; the total number of subjects on which the weighted mean is based; and the median number of subjects per obtained effect size. Table 4 provides a summary of the data displayed in the box plot in Figure 1.

3.15.4 QUANTITATIVELY ANALYZING THE STUDIES

In addition to the descriptive techniques just discussed, there are also procedures for inferential analysis of the retrieved data. As Table 5 shows, there are two major ways to evaluate the results of research studies: (i) terms of their statistical significance (e.g., *p* levels); and (ii) in terms of their effect sizes (e.g., *r*). When evaluating, there are two major analytic processes that can be applied to the set of studies: comparing and combining. Because there are some especially convenient computational procedures for the two-study situation, Table 5 separates the procedures applicable to the case of combining and comparing two studies from three or more studies. In cases of three or more studies being represented, we are able to compare studies on two levels: with diffuse tests and with focused tests. When diffuse tests are used to show whether significance levels or effect sizes differ, the researcher may learn that they differ but not how they differ. By using focused tests, or contrasts, we are able to expand on the knowledge of the diffuse tests by learning whether the studies differ in a theoretically predictable or meaningful way.

3.15.4.1 Comparing Two Studies

3.15.4.1.1 Significance testing

Although we are generally more interested in comparing the results of the effect sizes of the

Table 2 An example of a partial coding database of four studies from a meta-analysis of individual vs. group problem-solving performance.

Author	*Year*	*Type of publication*	*N*	*Subject occupation*	*Size of group*	*Type of task*	*Signfi-cance level*	*Effect size estimate*
Barnlund	1959	Journal article	143	Students	5	Logic problem	$Z = 6.3$	$r = 0.88$
Campbell	1968	Journal article	48	Managers	4	Human relations problem	$Z = -1.4$	$r = -0.43$
Kanekar	1982	Journal article	36	Students	2	Anagrams	$Z = 6.9$	$r = 0.80$
Knight	1921	Master's thesis	35	Students	6	Judgment task	$Z = 5.2$	$r = 0.74$

Table 3 A stem and leaf display of 83 effect sizes of the effect of gender on judgments of sexual harassment.

Stem	*Leaf*
0.6	5
0.6	
0.5	
0.5	
0.4	
0.4	3,3
0.3	7,7,9,9
0.3	0,0,2,3
0.2	5,5,5,5,5,6,7
0.2	2,2,3,3,4,4
0.1	5,5,5,5,5,6,6,6,7,7,7,8,8,8,8,9,9,9
0.1	0,1,1,1,2,2,2,3,3,3,3,4,4,4
0.0	5,6,6,6,7,7,8,8,8,8,8,8,9,9,9,9,9
0.0	0,0,0,0,0,0,0,0,0,3

two studies, sometimes this information is not available and all we are able to compare is p values. The following is the procedure for comparing the significance levels of two studies (Rosenthal & Rubin, 1979); all p levels are one-tailed. First, for each of the studies, we obtain as exact a p level as possible. For example, if we obtain $t(150) = 2.32$, our $p = 0.011$ not <0.05. Extended tables of the t distribution are helpful here (e.g., Federighi, 1959; Rosenthal & Rosnow, 1991) or calculators and software packages with built-in distributions of Z, t, F, and chi-square that will provide exact p values for the various statistical tests. Second, for each p, we find Z, the standard normal deviate corresponding to the p value (using a Z table or calculator or statistical software that converts from p to Z). Since only one-tailed p's were used, if both Z results are in the same direction, they will both have positive signs; if the results are in the opposite direction, one of the two Z's will be negatively signed. The two Z's are then compared by the following formula:

$$Z = \frac{Z_1 - Z_2}{\sqrt{2}} \qquad (2)$$

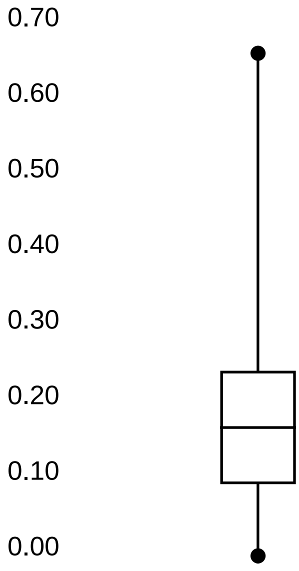

Figure 1 A box plot of the data displayed in Table 3.

Example 1

Studies A and B yield results in the same direction; one is significant at the 0.05 level, the other is not. The p level of Study A is 0.04, the p level of Study B is 0.30. The Z's corresponding to these p's are found in a table of the normal curve to be 1.75 and 0.52 respectively. Using Equation (2), we find that:

$$\frac{Z_1 - Z_2}{\sqrt{2}} = \frac{1.75 - 0.52}{\sqrt{2}} = 0.87$$

The p value associated with this Z is 0.19 one-tailed. Thus, when comparing these two studies, one of which reaches "significance," while the other does not, we find that the difference between them does not come close to the conventional levels of significance. The examples in the chapter are hypothetical in order to keep the computational examples small and manageable. For illustrations of various meta-analytic procedures with real-life examples, see the final chapter in Rosenthal (1991).

Example 2

Studies A and B yield results in different directions and both are significant. One p is 0.001, the other is 0.03. The Z's corresponding to these p's are 3.09 and -1.88 (note the opposite signs to indicate results in opposite directions). From Equation (2), we have:

$$\frac{(3.09) - (-1.88)}{\sqrt{2}} = 3.51$$

which has a corresponding p of 0.0002, which indicates that the two p values differ significantly.

3.15.4.1.2 Effect size estimation

When comparing studies, we are usually more interested in looking at the consistency or heterogeneity of the effect sizes than of the significance levels since a difference in Z's might simply be due to the difference in the size of the studies but not truly reflective of whether the two studies differ in effect size.

Researchers do not routinely report an effect size estimate together with their test of significance. Yet, in most cases, by rearranging the formulas in Table 1 into the following formulas, the meta-analyst can compute the

Table 4 A summary table of the measures of central tendency and variability of the distribution of effect sizes in Table 3.

Central tendency		*Variability*	
Unweighted mean	0.17	Maximum	0.65
Weighted mean	n.a.	Q3	0.23
Median	0.15	Median	0.15
Proportion of studies greater than 0.00	0.89	Q1	0.09
		Minimum	0.00
Median number of subjects per effect size	n.a.	Q3–Q1	0.14
		Standard deviation	0.13
Total number of independent effect sizes	83		
Total number of subjects	34 350		

n.a. = information not available

necessary effect size from the information that has been reported. There are several ways of computing effect sizes (*r*) from significance test results (*t*, *Z*, chi-square, and *F*) as shown below.

$$r = \sqrt{\frac{t^2}{t^2 + df}}$$

$$r = \sqrt{\frac{x^2(1)}{N}}$$

$$r = \frac{Z}{N}$$

$$r = \sqrt{\frac{F(1,\_)}{F(1,\_) + df_{error}}}$$

The effect size emphasized in this chapter will be *r* but analogous procedures are available for comparing other effect size indicators such as Hedge's *g* or *d′* (Hedges, 1982; Hsu, 1980; Rosenthal, 1991; Rosenthal & Rubin, 1982a; Rosenthal, Rosnow, & Rubin, 1997).

As the population value of *r* gets further and further from zero, the distribution of *r*'s sampled from that population becomes more and more skewed. To adjust for this skew, it is recommended that all *r*'s first be transformed to Fisher z_rs (Fisher, 1928), which are distributed nearly normally, before further computations are carried out. The relationship between *r* and *z* is given by $1/2 \log_e [(1 + r)/(1 - r)]$. There are handy tables available which convert *r* to z_r and z_r to *r* (e.g., Rosenthal & Rosnow, 1991) and most statistical software packages will easily perform this transformation as well. After converting obtained *r*'s to z_r's, the following formula may be used:

$$\frac{z_{r_1} - z_{r_2}}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}} \qquad (3)$$

which is distributed as *Z* (Snedecor & Cochran, 1989).

Example 3

Studies A and B yield results in the same direction with effect sizes of $r = 0.70$ ($n = 25$) and $r = 0.25$ ($n = 100$), respectively. The Fisher z_r's corresponding to these *r*'s are 0.87 and 0.26, respectively. From Equation (3), we have

$$Z = \frac{z_{r_1} - z_{r_2}}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}} = \frac{0.87 - 0.26}{\sqrt{\frac{1}{22} + \frac{1}{97}}} = 2.58$$

Table 5 An overview of 10 meta-analytic procedures.

	Significance testing	*Effect size estimation*
Comparing		
Two studies		
Three or more studies: diffuse tests		
Three or more studies: focused tests		
Combining:		
Two studies		
Three or more studies		

which has an associated p of 0.005, one-tailed. Thus, these two studies agree on a significant positive relationship between variables X and Y but disagree significantly in their estimates of the size of that relationship.

3.15.4.2 Combining Two Studies

3.15.4.2.1 Significance testing

While comparisons of two studies allow us to ask whether our studies differ in some way, it is also of interest to combine the two studies. By combining the significance levels of two studies, we are able to obtain an overall estimate of the probability that the two p levels might have been obtained if there truly were no relationship between the variables X and Y. We will present here the simplest and most versatile procedure combining probability levels, the method of adding Z's called the Stouffer method (Mosteller & Bush, 1954); other methods are summarized elsewhere (Rosenthal, 1978, 1991).

Like the method for comparing p values, the first step is to obtain accurate p levels for each of the two studies and then to find the Z's corresponding to those p's. The following equation is then applied:

$$\frac{Z_1 + Z_2}{\sqrt{2}} = Z \qquad (4)$$

Example 4

Studies A and B yield results in the same direction but neither are significant. One p value is 0.10, the other is 0.15, with corresponding Z's of 1.28 and 1.04. From Equation (4), we have

$$\frac{Z_1 + Z_2}{\sqrt{2}} = \frac{1.28 + 1.04}{\sqrt{2}} = 1.65$$

the p associated with this Z is 0.05. Thus, we have an example of two studies, neither of which yielded significant results when considered individually but which were significant when meta-analytically combined.

Example 5

The results of Studies A and B are in opposite directions and are both significant. One p is 0.00001, one-tailed, and the other is 0.03, one-tailed, but in the opposite direction. The Z's corresponding to these p's are 4.26 and -1.88 (note the opposite signs indicate results in opposite directions). Using Equation (4), we have

$$\frac{(4.26) + (-1.88)}{\sqrt{2}} = 1.68$$

with an associated p of 0.046. Thus, the combined p supports the finding of the more significant of the two results. However, since the two p's were so significantly different, we should be cautious in our interpretation and should search for that factors may have led Studies A and B to have obtained such different results.

3.15.4.2.2 Effect size estimation

Once we have computed the associated Fisher z_r's for each r, the combined effect size estimate is simply

$$\frac{z_{r_1} - z_{r_2}}{2} = \bar{z}_{r_1} \qquad (5)$$

or the Fisher z_r corresponding to our mean r. We can then look up the r associated with this mean z_r (using an r to z_r or z_r to r table or statistical software).

Example 6

Studies A and B yield results in the same direction, one $r = 0.95$, the other $r = 0.25$. The Fisher z_r's corresponding to these r's are 1.83 and 0.26, respectively. Applying Equation 5, we have

$$\frac{z_{r_1} + z_{r_2}}{2} = \frac{1.83 + 0.26}{2} = 1.045$$

which is our mean Fisher z_r. This value can then be converted back to an r of 0.78 which can be interpreted as showing a rather large combined mean effect (Cohen, 1977, 1988) from these two studies.

Example 7

Study A and B yield effect sizes of $r = 0.30$ and 0.00 respectively. The Fisher z_r's corresponding to these two effect sizes are 0.31 and 0.00. Again applying Equation (5) we have

$$\frac{0.31 + 0.00}{2} = 0.155$$

as our mean z_r which is associated with a mean r of 0.15.

These last two examples also illustrate that the use of Fisher's z_r gives heavier weight to r's that are further from zero in either direction but makes little difference when working with very small r's. If we had averaged the two r's from Example 6 without first transforming them to Fisher z_r's, we would have found a mean r of $(0.95 + 0.25)/2 = 0.60$, substantially smaller than the 0.78 which we found using the Fisher z_r's. However, if we recompute Example 7 using r's instead of Fisher z_r's, we would obtain the same result: $(0.30 + 0.00)/2 = 0.15$.

3.15.4.3 Comparing Three or More Studies: Diffuse Test

The following procedures are generalizations of the procedures just described which can be used to compare more than two studies. As mentioned earlier, when dealing with three or more studies, we are able to conduct both diffuse comparisons as well as focused comparisons. We will begin with the procedures for diffuse tests.

3.15.4.3.1 Significance testing

As with two studies, the first step is to find the Z associated with each p level. The studies can then be compared using the following equation:

$$\sum(Z_j - \bar{Z})^2 = \chi^2 \tag{6}$$

which is distributed as chi-square with $K - 1$ *df*. In this equation, Z_j is the Z for any one study, $\bar{Z}$ is the mean of all the Z's obtained, and K is the number of studies being compared. A significant chi-square tells us that the p's associated with our studies differ significantly among themselves.

Example 8

Studies A, B, C, and D yield one-tailed p values of 0.05, 0.001, 0.17, and 0.01, respectively. However, Study C shows results in the opposite direction from Studies A, B, and D. We find the Z's corresponding to these p's to be 1.64, 3.09, -0.95, and 2.33 (note the negative sign for Study C whose results were in the opposite direction). Using Equation (6) we find a chi-square value of

$$\sum(Z_j - \bar{Z})^2 = [(1.64) - (1.53)]^2 + [(3.09) - (1.53)]^2 + [(-0.95) - (1.53)]^2 + [(2.33) - (1.53)]^2 = 9.24$$

which for $K - 1 = 3$ *df* is significant at $p = 0.03$. The p values from these four studies, then, are significantly heterogeneous.

3.15.4.3.2 Effect size estimation

The statistical significance of the heterogeneity of the effect sizes of three or more studies is also obtained from a chi-square using the following equation

$$\sum(n_j - 3)(z_{r_1} - \bar{z}_r)^2 = \chi^2 \tag{7}$$

where n is the number of sampling units on which each r is based, z_{rj} is the Fisher z_r corresponding to each r, and $\bar{z}_r$ is the weighted mean z_r, that is

$$\bar{z}_r = \frac{\sum(n_j - 3)z_{r_j}}{\sum(n_j - 3)} \tag{8}$$

Example 9

Studies A, B, C, and D yield effect sizes of $r = 0.70$ ($n = 30$), $r = 0.45$ ($n = 45$), $r = 0.10$ ($n = 20$), and $r = -0.15$ ($n = 25$), respectively. The Fisher z_r's corresponding to these r's are 0.87, 0.48, 0.10, -0.15. First, using Equation (8), the weighted mean, z_r, is

$$\frac{[27(0.87) + 42(0.48) + 17(0.10) + 27(-0.15)]}{27 + 42 + 17 + 22} = 0.39$$

Then, applying Equation (7), we find that

$$27(0.87 - 0.39)_2 + 42(0.48 - 0.39)_2 + 17(0.10 - 0.39)_2 + 22(-0.15 - 0.39)^2 = 14.4$$

which for $K - 1 = 3$ *df* is significant at $p = 0.0024$. These four effect sizes, then, are significantly heterogeneous.

3.15.4.3.3 Comparing three or more studies: focused tests

While the two examples above showed us that the set of four studies were found to be significantly heterogeneous both in terms of their significance levels and effect sizes, they merely told us that overall the four studies differed. It is generally more useful to have more focused information about how the four studies differed. For example, did the prescribed treatment differ when given to women vs. men? Or, for example, did the studies in which the subjects lived in nursing homes differ from those studies in which the subjects lived with their families? In order to answer this question, it is necessary to conduct focused tests, also known as contrasts.

3.15.4.3.4 Significance levels

As with the diffuse comparisons, we first find the Z's that correspond to the p level of each study. We then compute a Z from the following equation

$$\frac{\sum \lambda_j Z_j}{\sqrt{\sum \lambda_j^2}} = Z \tag{9}$$

In this equation λ_j is the theoretically derived prediction or contrast weight for any one study, chosen such that the sum of the λ_j's will be zero. Z_j is the Z for any one study.

Example 10

Studies A, B, C, and D were conducted to examine the efficacy of a new type of behavioral therapy. They yielded the following one-tailed p values: 0.0001, 0.03, 0.15, and 0.10, all in the same direction. We calculate the Z's corresponding to these p's to be 3.72, 1.88, 1.04, and 1.28, respectively. We also know that the amount of therapy given in each of the studies differed such that Studies A, B, C, and D involved 8, 6, 4, and 2 h of therapy per month, respectively. A focused question that would be of interest to us in this example might be whether there was a significant linear relationship between the number of hours of therapy and the statistical significance of the result. The weights of a linear contrast involving four studies are 3, 1, -1, -3, as obtained from a table of orthogonal polynomials (e.g., Rosenthal & Rosnow, 1991). Therefore, from Equation (9) we have

$$\frac{\sum \lambda_j Z_j}{\sqrt{\sum \lambda_j^2}} = \frac{(3)3.72 + (1)1.88 + (-1)1.04 + (-3)1.28}{\sqrt{(3)^2 + (1)^2 + (-1)^2 + (-3)^2}} = 1.82$$

as our Z value which is significant at $p = 0.03$. Thus, the four p values tend to grow more significant as more hours of therapy are received per month.

3.15.4.3.5 Effect size estimation

It is almost always quite valuable to the researcher to compute a focused comparison on a series of effect sizes. For example, given a set of effect sizes for studies of a new behavioral therapy, we might want to know whether these effects are increasing or decreasing linearly with the number of hours of therapy received per month. It is from the focused comparison of effects that we are able to test moderator variables and often form new theories.

As was the case for diffuse tests, we begin by computing the effect size r, its associated Fisher z_r and $n-3$ where n is the number of sampling units on which each r is based. The statistical significance of the focused test or contrast then is obtained from a Z computed as follows (Rosenthal & Rubin, 1982a):

$$\frac{\sum \lambda_j z_{r_j}}{\sqrt{\frac{\sum \lambda_j^2}{W_j}}} = Z \qquad (10)$$

Once again, λ_j is the theoretically derived prediction or contrast weight for any one study, chosen such that the sum of the λ_j's will be zero. The z_{rj} is the Fisher z_r for any one study and w_j is the inverse of the variance of the effect size for each study. When using Fisher z_r transformations of the effect size r, the variance is $1/(n_j - 3)$, so for the present example, $w_j = n_j - 3$.
Example 11

Studies A, B, C, and D yield effect sizes of 0.89, 0.76, 0.23, and 0.59, respectively, all with $n = 12$. The Fisher z_r's corresponding to these r's are found to be 1.42, 1.00, 0.23, and 0.68, respectively. We also know that Studies A, B, C, and D involved 8, 6, 4, and 2 h of therapy per month, respectively. We are interested in whether there is a relationship between number of hours of therapy received and size of observed effect favoring the new therapy. As in Example 10, the appropriate λ's are 3, 1, -1, and -3. Therefore, applying Equation (10) we find

$$\frac{\sum \lambda_j z_{r_j}}{\sqrt{\frac{\sum \lambda_j^2}{W_j}}} = \frac{(3)1.42 + (1)1.00 + (-1)0.23 + (-3)0.68}{\sqrt{\frac{(3)^2}{9} + \frac{(1)^2}{9} + \frac{(-1)^2}{9} + \frac{(-3)^2}{9}}} = 2.01$$

as our Z value which is significant at $p = 0.22$. Thus, the four effect sizes tend to grow larger as the number of hours of therapy increases.

(i) Interpretation of moderator variables

Examples 10 and 11 seem to show that number of hours of therapy is linked linearly to effectiveness of the therapy. Yet, in these hypothetical examples, subjects were not assigned at random to the four studies, each of which studied the effects of treatment employing a different number of hours per week of therapy. Therefore, our interpretation of this finding must be cautious. We cannot be sure that differences between the four studies are due to the different number of therapy hours and not to some other variable that is also correlated with this moderator. The finding should not be taken as evidence for a causal relationship, rather as being suggestive of the possibility of a causal relationship, a possibility which can then be studied experimental in future research (Hall, Tickle-Degnen, Rosenthal, & Mosteller, 1994).

3.15.4.4 Combining Three or More Studies

3.15.4.4.1 Significance testing

By combining the p levels of a set of studies, we are able to obtain an overall estimate of the probability that the set of p levels might have been obtained if the null hypothesis of no relationship between X and Y were true. The method presented here is the generalized version of the method presented earlier as applying to the results of two studies; other methods are described in detail elsewhere (e.g., Rosenthal, 1991).

As before, we first obtain the Z's corresponding to our p levels (with Z's disagreeing in direction given negative signs). The sum of the Z's is then divided by the square root of the number of studies (K) yielding a new statistic which is distributed as Z, as follows:

$$\frac{\sum Z_j}{\sqrt{k}} = Z \tag{11}$$

Example 12

Studies A, B, C, and D yield one-tailed p values of 0.05, 0.001, 0.17, and 0.01, respectively. However, Study C shows results in the opposite direction from Studies A, B, and D. We find the Z's corresponding to these p's to be 1.64, 3.09, -0.95, and 2.33 (note the negative sign for Study C whose results were in the opposite direction). Applying Equation (11) we then find:

$$\frac{\sum Z_j}{\sqrt{k}} = \frac{(1.64) + (3.09) + (-0.95) + (2.33)}{\sqrt{4}}$$
$$= 3.06$$

as our new Z value which is associated with a p of 0.001, one-tailed. (We would normally use a one-tailed p if we had correctly predicted the bulk of the findings but would use a two-tailed p value if we had not.) Thus, we find a combined p level for the four studies that would be highly unlikely if the null hypothesis were true. Yet, we should be very cautious about drawing any simple overall conclusion from this combined significance level because of the heterogeneity of these four p levels. Example 8 employed the same p values and found that they were significantly different at $p = 0.03$. This heterogeneity, however, could be due to heterogeneity of effect sizes, of sample sizes, or both. We should always look carefully at the effect sizes and sample sizes of the studies before drawing any conclusions based on the significance levels above.

3.15.4.4.2 Effect size estimation

When combining the results of three or more studies, we are at least as interested in the combined effect size estimate as in the combined probability discussed above. In computing this value, we once again convert our r's to the associated Fisher z_r's and simply find the mean Fisher z_r as follows:

$$\frac{\sum z_r}{k} = \bar{z}_r \tag{12}$$

where K refers to the number of studies being combined. We then find the r associated with this mean z_r.

Example 13

Studies A, B, and C yield effect sizes of $r = 0.70$, 0.45, and 0.10, respectively, all in the same direction. The Fisher z_r values associated with these r's are 0.87, 0.48, and 0.10, respectively. Applying Equation (12), we find

$$\frac{\sum z_r}{k} = \frac{0.87 + 0.48 + 0.10}{3} = 0.48$$

as our mean Fisher z_r, which corresponds to an r of 0.45, a rather strong combined effect of the three studies. In this example, all the effect sizes were in the same direction; however, if we had a case of three studies where effect sizes seemed to be substantially heterogeneous, we would be cautious in our interpretation of the combined effect size, just as we were in our treatment of the combined probabilities in Example (12).

3.15.4.5 Weighting Studies

When combining two or more studies, the meta-analyst may choose to weight each study by size of study (*df*), estimated quality of study, or any other desired weights (Mosteller & Bush, 1954; Rosenthal, 1978, 1980, 1984). The general procedure for weighting Z's is to (i) first assign to each study a given weight (symbolized by w and which is assigned before inspection of the data); (ii) multiply each Z by the desired weight, w; (iii) add the weighted Z's and; (iv) divide this sum by the square root of the sum of the squared weights, as follows:

$$\frac{\sum w_j Z_j}{\sqrt{\sum w_j^2}} = Z_w \tag{13}$$

Example 14

We are interested in finding the combined Z from the four studies discussed in Example 12 and now we want to weight them each by their degrees of freedom which are 24, 75, 4, and 50, respectively. From Example 12, we found Z's for the four studies to be 1.64, 3.09, -0.95, and 2.33. Applying Equation (13) we now find:

$$\frac{\sum w_j z_j}{\sqrt{\sum w_j^2}}$$
$$= \frac{24(1.64) + 75(3.09) + 4(-0.95) + 50(2.33)}{\sqrt{(24)^2 + (75)^2 + (4)^2 + (50)^2}}$$
$$= 4.11$$

as our combined Z which is associated with a p of 0.00002. We should keep in mind that when weighting Z's by *df*, the size of the study is actually playing a doubly large role since it has already been entered into the determination of each Z initially.

We can similarly weight effect sizes by *df* or any other desired weight using the following equation:

$$\bar{z}_{r_{weighted}} = \frac{\sum w_j z_{r_j}}{\sqrt{\sum w_j}} \tag{14}$$

Example 15

From Example 13, we found that Studies A, B, and C yield Fisher z_r values of 0.87, 0.48, and 0.10, respectively. The *df*'s for each study are 56, 120, and 24, respectively. Applying Equation (14), we find:

$$\begin{aligned}\bar{z}_{r_{weighted}} &= \frac{\sum w_j z_{r_j}}{\sqrt{\sum w_j}} \\ &= \frac{56(0.87) + 120(0.48) + 24(0.10)}{56 + 120 + 24} \\ &= 0.54\end{aligned}$$

as our mean weighted Fisher z_r which corresponds to an r of 0.49, slightly larger than the unweighted mean effect size r of 0.45 which we found above.

3.15.5 SPECIAL CASES

3.15.5.1 Imprecisely Reported *p*'s

The meta-analytic procedures discussed above for combining and comparing significance levels require the meta-analyst to find the Z associated with the given p from each study. Unfortunately, it is quite common for p's to be reported imprecisely as $p < 0.05$ or < 0.01 so that p might be 0.001 or 0.0001 or even 0.00001. When this is the case, the researcher needs to go back to the original test statistic employed, for example, t, F, Z, or chi-square, which many journals require their contributors to report. The *df* for the t and for the denominator of the F test tell us about the size of the study. The *df* for chi-square is analogous to the *df* for the numerator of the F test and so tells us about the number of conditions, not the number of sampling units. Fortunately, the 1983 edition of the Publication Manual of the American Psychological Association has added the requirement that the total N be reported together with the *df* for conditions for chi-square tests. Using the reported test statistic and the information about the size of the study, we should be able to obtain the exact p value for the given test statistic and the Z associated with that p.

Occasionally, we may have information about an effect size estimator (such as r, d, or g) but not a test statistic. Using the equations in section 3.15.4.1.2 we can often derive the associated test statistic, its exact p value, and the associated Z.

If a result is simply reported as "nonsignificant," and if no further information is available, we have little choice but to treat the result as a p of 0.50, which corresponds to a Z of 0.00. Although this procedure is ordinarily conservative leading to effect size estimates that are too small, the alternative of not using those studies is likely to lead to effect size estimates that are too large and almost surely to p values that are too significant. The recommended approach is to conduct the analysis both ways—with the studies and without the studies and to see how much of a difference it will make to the overall view of the data.

3.15.5.2 Repeated Measures Designs

A recent meta-analysis (Kramer & Rosenthal, 1997) introduced the concept of the repeated measures meta-analysis. In this situation, the question of interest concerns the difference between two or more effect sizes for each study. The meta-analysis compared the improvement in problem-solving performance when individuals were formed into noninteracting groups vs. their performance when they actually interacted as a group. A similar situation where one may be interested in a repeated measures question might be where patients' improvement is first measured one month after receiving a new treatment and this improvement is then compared to their improvement a year after receiving the treatment. One useful way to approach this type of meta-analysis is to treat it as if it were two meta-analyses whose results are then compared. In the first example above, a meta-analysis was conducted of the effect sizes associated with improvement from the individual to the noninteracting group condition. The results from this meta-analysis were then compared with the meta-analytic results of the size of the effect associated with improvement due to interaction.

3.15.5.3 Random vs. Fixed Effects

A common area of confusion is the distinction between random and fixed effects models of analysis. The fixed effects model is in the same spirit as a regression model or a fixed analysis of variance in which the treatment levels are considered fixed and the random variation is from the sample of individuals within treatment levels. With a fixed effects analysis, we are able to generalize to participants of the type found in the studies we included in the meta-analysis but we are not able to generalize to other studies.

This is because the source of sampling error in the fixed model is the variation among the people in the studies. Although our generalizability is somewhat restricted by using a fixed effects model, we gain in statistical power because our *df* are drawn from the total *N* of all the studies.

With a random effects model, our generalizability increases to all studies of the type from which our sample of studies was drawn. However, by employing a random effects model, we lose in statistical power compared to the fixed effects analysis. A simple one-sample *t*-test on the mean effect size (Mosteller & Bush, 1954) is often an appropriate random effects analysis. It will typically be more conservative than the results of the Stouffer method, but would allow for more generalizability. It is highly recommended to conduct an analysis both with a random and a fixed effects model and to compare the two methods. (For more detailed discussion of these issues, see Hedges, 1994; Raudenbush, 1994; Rosenthal, 1995a; Shadish & Haddock, 1994).

3.15.6 CRITICISMS OF META-ANALYSIS

3.15.6.1 Sampling Bias and the File Drawer Problem

One of the most frequent criticisms leveled against meta-analyses is that they are inherently biased because there is a greater likelihood that published studies will have significant results and thus might not be representative of the population of studies which have been conducted. This criticism, while well taken, could equally well be applied to narrative reviews of the literature. However, with a meta-analysis, there are now certain computational procedures that can be employed to address this problem. One visual way of examining the issue of publication bias is through the use of a Funnel Plot (Light & Pillemer, 1984), a scatterplot of sample size vs. estimated effect sizes. Another method is to compute a "file drawer analysis" (Rosenthal, 1991; Rosenthal & Rubin, 1988) as follows.

To find the number (*X*) of new, filed, or unretrieved studies averaging null results required to bring the new overall *p* to any desired level of significance (e.g., $p = 0.05$, $Z = 1.645$), the following equation may be used:

$$X = \frac{(\sum Z)^2}{2.706} - k \qquad (15)$$

where *K* is the number of studies combined and ΣZ is the sum of the *Z*'s obtained for the *K* studies.

Example 16

57 experiments have been conducted examining the effect of a new treatment for alcoholism. The ΣZ of these 57 studies was 98.6. How many new, filed, or unretrieved studies (*X*) would be required to bring this *Z* down to a barely significant level ($Z = 1.645$)? From Equation (15) we have

$$X = \frac{(98.6)^2}{2.706} - 57 = 3535$$

Thus, there would need to be over 3000 studies averaging null results tucked away in file drawers before one could conclude that the overall results of the 57 experiments were due to sampling bias.

3.15.6.2 Combining Apples and Oranges: Problems of Heterogeneity

Meta-analysis has been criticized for combining apples and oranges, in other words, for combining studies with different operationalizations of independent and dependent variables and different sampling units. This is in fact true, and a good meta-analysis will examine these differing factors as potential moderating variables as was discussed above.

Critics of meta-analysis have also pointed out the dangers of combining studies of different quality and treating them all as equal. This is an issue which can be dealt with by weighting studies according to their quality using the same procedures illustrated in Examples 14 and 15 where studies were weighted by study size. Consider the following example.

Example 17

Studies A, B, and C yield effect sizes of Fisher $z_r = 0.70$, 0.23, and 0.37. A group of judges rated these three studies for their internal validity and found mean ratings of 1.5, 3.4, and 4.0. Using these mean ratings as weights, the weighted combined effect size can be calculated using Equation (14):

$$\frac{\sum w_j z_{r_j}}{\sum w_j} = \frac{1.5(0.70) + 3.4(0.23) + 4.0(3.7)}{1.5 + 3.4 + 4.0}$$
$$= 0.37$$

We can conclude that weighting by quality of research in this case led to a somewhat different result than not weighting in that it lowered the perhaps somewhat elevated effect size of the poorer quality study.

Coding of studies for their quality involves having raters check each study for the presence of certain aspects, such as random assignment, experimenter blind to hypothesis, presence of demand characteristics, etc. and then adding the number of desirable features present for each study. Ratings of studies involves having raters make a more global, overall assessment of the methodological quality of a study. When establishing weightings of quality, the coding or rating should be done, as in hypothetical Example 17, by a group of disinterested judges, rather than by the researchers conducting the meta-analysis. This will safeguard against the understandable tendency to think of our own studies, of those of our students and our friends, and of those who successfully replicate our work as good studies while thinking of those studies conducted by our enemies and by those who failed to replicate our work as bad studies. Reliability of coding or rating should also be reported.

3.15.6.3 Problems of Independence

3.15.6.3.1 Nonindependence of subjects

It would not be unusual for the meta-analyst to discover multiple studies in which the same subjects have participated or that within one study subjects have contributed to more than one dependent variable. For example, a study may present the assessments of a group of subjects on a number of different measures of sensitivity to nonverbal cues. In this situation, the results generated by each dependent variable (each measure of nonverbal sensitivity) cannot be considered independent of each other since the same subjects contributed toward each one. Although it would be an error to treat multiple results from one study as though they were independent for significance testing, there is nothing wrong in doing so for the purposes of effect size estimation. By doing so, the researcher weights each study in proportion to the number of different effect sizes it generates, a procedure which meta-analysts might not choose to do but one which would not be unjustifiably wrong. Our recommendation is to use any of the procedures mentioned in order to produce only one effect size estimate and significance level for each study for the overall analysis. One option for combining the various significance levels and effect sizes would simply be to take the mean or the median Z or z_r of all the dependent variables. (The mean Z could be computed by averaging the Z's or by first averaging the z_r's and then getting the Z associated with the mean z_r; the mean z_r could also be computed by averaging the p levels and then computing the effect size that corresponds to that mean p level. These different procedures can yield different results, none of which is intrinsically more correct, but one should be chosen beforehand and used consistently through out the meta-analysis.)

However, unless the dependent variables are perfectly correlated, using the mean or the median estimates is likely to be somewhat conservative. If the meta-analyst has available the degrees of freedom and typical intercorrelation among the dependent variables, another approach, described in more detail elsewhere (Rosenthal & Rubin, 1986), can provide a more accurate combination of the effect sizes which can then be entered into the meta-analysis. This approach also allows the meta-analyst to compute contrasts among the effect sizes for the multiple dependent variables.

3.15.6.3.2 Nonindependence of experimenters

Given that researchers within the same laboratory or who once worked within the same laboratory group may have similar interests, it quite frequently occurs that several of the studies in the meta-analysis have been conducted by the same experimenter or experimental team. This may be problematic in that the results of researcher A may be correlated with those of researcher B despite the fact that they used different participants. It would be sound practice to analyze the studies by laboratory as well as by study. For example, a recent meta-analysis of the predictive power of short samples of nonverbal behavior (Ambady & Rosenthal, 1992) compared studies authored by either of the two meta-analysts to studies conducted by other researchers. The comparison showed all the effect sizes to be quite homogenous with a mean effect size of studies by the two researchers of $r = 0.38$ compared to a mean r of 0.39 by the others.

3.15.6.4 Exaggerated Significance Levels

The "criticism" that meta-analyses yield more significant results is to a great extent true. As we saw with Equation(1) (Significance Test = Size of Effect × Size of Study), as more subjects are added, either to a single study or to a meta-analysis of many studies, the results will become more significant. In fact, this is one of the benefits of conducting meta-analyses in that it allows for the cumulation of knowledge that was discussed above so that we no longer dismiss studies which have not found p's < 0.05 but rather cumulate them so that the body of work may show its overall significance level and effect size.

3.15.6.5 Too Small Effects

In contrast to the argument above, another criticism of meta-analysis is that the results of

Table 6 A binomial effect size display (BESD) of the relationship between aspirin usage and heart attacks.

	Heart attack	*No heart attack*	*Total*
Aspirin	48.3	51.7	100
Placebo	51.7	48.3	100
Total	100	100	200

some meta-analyses really only show "small effects" because the obtained *r*'s and subsequent *r*'s are small. While this may be true, it is also true that small effects can have great importance. A classic example of this is the Physician's Aspirin Study (see Rosenthal, 1995b). In 1988, a study of the effects of aspirin on reducing heart attacks was prematurely terminated. It had become clear to the physicians involved that it would be unethical to continue to give half of the subjects a placebo, since aspirin clearly prevented heart attacks. What was the *r* of this study? The *r* was 0.034, with an r^2 of 0.0011.

Table 6 presents the results of the aspirin study as a Binomial Effect Size Display (BESD), a display which allows us to see the practical importance of the effect (Rosenthal & Rubin, 1982b). The correlation is shown as the difference in outcome rates between the experimental and the control group. As the BESD shows, aspririn lead to a 4% decrease in heart attacks, a small but rather important, effect.

3.15.6.6 Abandonment of Primary Studies

Finally, critics have claimed that the growing use of meta-analysis has destroyed the motivation (or even the need) to conduct primary research studies. Certainly not: first of all, the primary studies are needed in order to do a meta-analysis. Second, the research suggested by meta-analyses can only be addressed by newly designed primary studies employing randomization. A great benefit of meta-analysis is that it will prevent unnecessary duplication, replication, and the wasting of scientifically and socially valuable resources. Conducting more meta-analyses may very well reduce the number of primary studies conducted, but it should increase the potential value of those that are conducted.

3.15.7 REFERENCES

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111*, 256–274.

Bausell, R. B., Li, Y., Gau, M., & Soeken, K. L. (1995). The growth of meta-analytic literature from 1980 to 1993. *Evaluation and the Health Professions, 18*, 238–251.

Blumenthal, J. A. (1998). The reasonable woman standard: A meta-analytic review of gender differences in perceptions of sexual harassment. *Law and Human Behavior, 22,* 33–57.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press.

Cohen, Jon (1993). A new goal: Preventing disease, not infection. *Science, 262*, 1820–1821.

Cooper, H. M. (1989). *Integrating research: A guide for literature reviews* (2nd ed.). Newbury Park, CA: Sage.

Cooper, H. M., & Hedges, L. V. (Eds.) (1994). *The handbook of research synthesis.* New York: Sage.

Crane, D. (1972). *Invisible colleges: Diffusion of knowledge in scientific communities.* Chicago: University of Chicago Press.

Dickersin, K. (1994). Research registers. In H. M. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 71–83). New York: Sage.

Eagly, A. H., & Carli, L. L. (1981). Sex of researchers and sex-typed communications as determinants of sex differences in influenceability: A meta-analysis of social influence studies. *Psychological Bulletin, 90*, 1–20.

Federighi, E. T. (1959). Extended tables of the percentage points of Student's t-distribution. *Journal of the American Statistical Association, 54*, 683–688.

Fisher, R. A. (1928). *Statistical methods for research workers* (4th ed.). London: Oliver & Boyd.

Fisher, R. A. (1938). *Statistical methods for research workers* (7th ed.). London: Oliver & Boyd.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*(10), 3–8.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Greenhouse, J. B., & Iyengar, S. (1994). Sensitivity analysis and diagnostics. In H. M. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 383–398). New York: Sage.

Hall, J. A., Tickle-Degnen, L., Rosenthal, R., & Mosteller, F. (1994). Hypotheses and Problems in Research Synthesis. In H. M. Copper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 17–28). New York: Sage.

Hedges, L. V. (1982). Fitting categorical models to effect sizes from a series of experiments. *Journal of Educational Statistics, 7,* 119–137.

Hedges, L. V. (1987). How hard is hard science, how soft is soft science? *American Psychologist*, *42*, 443–455.

Hedges, L. V. (1994). Fixed effects models. In H. M. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 285–299). New York: Sage.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Hsu, L. M. (1980). Tests of differences in p levels as tests of differences in effect sizes. *Psychological Bulletin, 88,* 705–708.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

Kramer, S. H., & Rosenthal, R. *Why are two heads better than one: A meta-analytic comparison of individual, statistical, and interacting group problem solving.* Manuscript submitted for publication.

Light, R. J., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research.* Cambridge, MA: Harvard University Press.

Light, R. J., Singer, J. D., & Willett, J. B. (1994). The visual presentation and interpretation of meta-analyses. In H. M. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 439–453). New York: Sage.

Lipsey, M. W. (1994). Identifying potentially interesting variables and analysis opportunities. In H. M. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 111–124). New York: Sage.

Mosteller, F. M., & Bush, R. R. (1954). Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology: Vol. 1. Theory and method* (pp. 289–334). Cambridge, MA: Addison-Wesley.

Raudenbush, S. W. (1994). Random effects models. In H. M. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 301–322). New York: Sage.

Reed, J. G., & Baxter, P. M. (1994). Using reference databases. In H. M. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 57–70). New York: Sage.

Rosenthal, R. (1966). *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts.

Rosenthal, R. (1978). Combining results of independent studies. *Psychological Bulletin, 85,* 185–193.

Rosenthal, R. (Ed.) (1980). *New directions for methodology of social and behavioral science: Quantitative assessment of research domains* (No. 5). San Francisco: Jossey-Bass.

Rosenthal, R. (1984). *Meta-analytic procedures for social research*. Beverly Hills, CA: Sage.

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.

Rosenthal, R. (1993). Cumulating evidence. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 519–559). Hillsdale, NJ: Erlbaum.

Rosenthal, M. C. (1994). The fugitive literature. In H. M. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 85–95). New York: Sage.

Rosenthal, R. (1995a). Writing meta-analytic reviews. *Psychological Bulletin, 118*(2), 183–192.

Rosenthal, R. (1995b). Progress in clinical psychology: Is there any? *Clinical Psychology: Science and Practice, 2,* 133–50.

Rosenthal, R. (1998). Meta-analysis: Concepts, corollaries and controversies. In J. G. Adair D. Bellanger, & K. L. Dion (Eds.), *Advances in psychological science. Vol. 1: Social personal and cultural aspects* (pp. 371–384). Hove, UK: Psychology Press.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York: McGraw-Hill.

Rosenthal, R., Rosnow, R. L., & Rubin, D. B. *Contrasts and effect sizes in behavioral research: A correlational approach*. Unpublished Volume.

Rosenthal, R., & Rubin, D. B. (1979). Comparing significance levels of independent studies. *Psychological Bulletin, 86,* 1165–1168.

Rosenthal, R., & Rubin, D. B. (1982a). Comparing effect sizes of independent studies. *Psychological Bulletin, 92,* 500–504.

Rosenthal, R., & Rubin, D. B. (1982b). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74,* 166–169.

Rosenthal, R., & Rubin, D. B. (1986). Meta-analytic procedures for combining studies with multiple effect sizes. *Psychological Bulletin, 99*(3), 400–406.

Rosenthal, R., & Rubin, D. B. (1988). Comment: Assumptions and procedures in the file drawer problem. *Statistical Science, 3,* 120–125.

Shadish, W. R., & Haddock, C. K. (1994). Combining estimates of effect size. In H. M. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 261–281). New York: Sage.

Smith, M. L., Glass, G. V., & Miller, T. I. (1980). *The benefits of psychotherapy*. Baltimore, MD: Johns Hopkins University Press.

Snedecor, G. W. (1946). *Statistical methods applied to experiments in agriculture* (4th ed.). Ames, IA: Iowa State University Press.

Snedecor, G. W., & Cochran, W. G. (1967). *Statistical methods* (6th ed.). Ames, IA: Iowa State University Press.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

White, H. D. (1994). Scientific communication and literature retrieval. In H. M. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 41–55). New York: Sage.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.01

The Role of Assessment in Clinical Psychology

LEE SECHREST, TIMOTHY R. STICKLE, and MICHELLE STEWART
University of Arizona, Tucson, AZ, USA

4.01.1 INTRODUCTION 2
4.01.1.1 Useful Clinical Assessment is Difficult but not Impossible 2
4.01.2 WHY ARE ASSESSMENTS DONE? 4
4.01.2.1 Bounded vs. Unbounded Inference and Prediction 5
4.01.2.2 Prevalence and Incidence of Assessment 5
4.01.2.3 Proliferation of Assessment Devices 7
4.01.2.4 Over-reliance on Self-report 9
4.01.3 PSYCHOMETRIC ISSUES WITH RESPECT TO CURRENT MEASURES 10
4.01.3.1 Reliability 10
4.01.3.2 Validity 10
4.01.3.3 Item Response Theory 11
4.01.3.4 Scores on Tests 11
4.01.3.5 Calibration of Measures 12
4.01.4 WHY HAVE WE MADE SO LITTLE PROGRESS? 12
4.01.4.1 The Absence of the Autopsy 14
4.01.5 FATEFUL EVENTS CONTRIBUTING TO THE HISTORY OF CLINICAL ASSESSMENT 14
4.01.5.1 The Invention of the Significance Test 14
4.01.5.2 Ignoring Decision Making 14
4.01.5.3 Seizing on Construct Validity 16
4.01.5.4 Adoption of the Projective Hypothesis 16
4.01.5.5 The Invention of the Objective Test 17
4.01.5.6 Disinterest in Basic Psychological Processes 17
4.01.6 MISSED SIGNALS 19
4.01.6.1 The Scientist–Practitioner Model 19
4.01.6.2 Construct Validity 19
4.01.6.3 Assumptions Underlying Assessment Procedures 20
4.01.6.4 Antecedent Probabilities 20
4.01.6.5 Need for Integration of Information 20
4.01.6.6 Method Variance 21
4.01.6.7 Multiple Measures 21
4.01.7 THE ORIGINS OF CLINICAL ASSESSMENT 22
4.01.7.1 The Tradition of Assessment in Psychology 22
4.01.7.1.1 Witmer 23
4.01.7.1.2 Army Alpha 23
4.01.8 THE RORSCHACH INKBLOT TECHNIQUE AND CLINICAL PSYCHOLOGY 23
4.01.8.1 The Social and Philosophical Context for the Appearance of the Rorschach 23
4.01.8.2 The Birth of the Rorschach 24
4.01.8.3 Clinical vs. Statistical Prediction 25
4.01.8.4 Old Tests Never Die, They Just Fade Away 26

4.01.9 OTHER MEASURES USED IN CLINICAL PSYCHOLOGY 27
4.01.9.1 The Thematic Apperception Test 27
4.01.9.2 Sentence Completion Tests 28
4.01.9.3 Objective Testing 28
4.01.9.4 The Clinician as a Clinical Instrument 28
4.01.9.5 Structured Interviews 29
4.01.10 CONCLUSIONS 29
4.01.11 REFERENCES 29

4.01.1 INTRODUCTION

In this chapter we will describe the current state of affairs with respect to assessment in clinical psychology and then we will attempt to show how clinical psychology got to that state, both in terms of positive influences on the directions that efforts in assessment have taken and in terms of missed opportunities for alternative developments that might have been more productive psychology. For one thing, we really do not think the history is particularly interesting in its own right. The account and views that we will give here are our own; we are not taking a neutral—and innocuous—position. Readers will not find a great deal of equivocation, not much in the way of "a glass half-empty is, after all, half-full" type of placation. By assessment in this chapter, we refer to formal assessment procedures, activities that can be named, described, delimited, and so on. We assume that all clinical psychologists are more or less continuously engaged in informal assessment of clients with whom they work. Informal assessment, however, does not follow any particular pattern, involves no rules for its conduct, and is not set off in any way from other clinical activities. We have in mind assessment procedures that would be readily defined as such, that can be studied systematically, and whose value can be quantified. We will not be taking account of neuropsychological assessment nor of behavioral assessment, both of which are covered in other chapters in this volume. It will help, we think, if we begin by noting the limits within which our critique of clinical assessment is meant to apply. We, ourselves, are regularly engaged in assessment activities, including developmemt of new measures, and we are clinicians, too.

4.01.1.1 Useful Clinical Assessment is Difficult but not Impossible

Many of the comments about clinical assessment that follow may seem to some readers to be pessimistic and at odds with the experiences of professional clinicians. We think our views are quite in accord with both research and the theoretical underpinnings for assessment activities, but in at least some respects we are not so negative in our outlook as we may seem. Let us explain. In general, tests and related instruments are devised to measure constructs, for example, intelligence, ego strength, anxiety, antisocial tendencies. In that context, it is reasonable to focus on the construct validity of the test at hand: how well does the test measure the construct it is intended to measure? Generally speaking, evaluations of tests for construct validity do not produce single quantitated indexes. Rather, evidence for construct validity consists of a "web of evidence" that fits together at least reasonably well and that persuades a test user that the test does, in fact, measure the construct at least passably well. The clinician examiner especially if he or she is acquainted in other ways with the examinee, may form impressions, perhaps compelling, of the validity of test results. The situation may be something like the following:

Test <—construct

That is, the clinician uses a test that is a measure of a construct. The path coefficient relating the test to the construct (in the convention of structural equations modeling, the construct causes the test performance) may well be substantial. A more concrete example is provided by the following diagram:

IQ Test <—0.80—intelligence

This diagram indicates that the construct of intelligence causes performance on an IQ test. We believe that IQ tests may actually be quite good measures of the construct of "intelligence." Probably clinicians who give intelligence tests believe that in most instances the test gives them a pretty good estimate of what we mean by intelligence, for example, 0.80 in this example. To use a term that will be invoked later, the clinician is "enlightened" by the results from the test.

As long as the clinical use of tests is confined to enlightenment about constructs, many tests may have reasonably good, maybe even very good "validity." The tests are good measures of the constructs. In many, if not most, clinical uses of tests, however, the tests are used in order to make decisions. Tests are used, for example to decide whether a parent should have custody of a child, to decide whether a patient is likely to benefit from some form of therapy, to decide whether a child "should" be placed in a social classroom, or to decide whether a patient should be put on some particular medication. Using our IQ test example, we get a diagram of the following sort:

IQ Test<—0.80—intelligence—0.50—> School grades

This diagram, which represents prediction rather than simply enlightenment, has two paths, and the second path is almost certain to have a far lower validity coefficient than the first one. Intelligence has a stronger relationship to performance on an IQ test than to performance in school. If an IQ test had construct validity of 0.80, and if intelligence as a construct were correlated 0.50 with school grades, which means that intelligence would account for 25% of the total variance in school grades, then the correlation between the IQ test and school grades would be only 0.80 x 0.50 = 0.40 (which is about what is generallly found to be the case).

IQ Test<—0.40—>School grades

A very good measure of ego strength may not be a terribly good predictor of resistance to stress in some particular set of circumstances. Epstein (1983) pointed out some time ago that tests cannot be expected to be related especially well to specific behaviors, but it is in relation to specific behaviors that tests are likely to be used in clinical settings.

It could be argued and has been, (e.g., Meyer & Handler 1997), that even modest validities like 0.40 are important. Measures with a validity of 0.40, for example, can improve ones prediction from that 50% of a group of persons will succeed at some task to the prediction that 70% will succeed. If the provider of a service cannot serve all eligible or needy persons, that improvement in prediction may be quite useful.

In clinical settings, however, decisions are made about individuals, not groups. To recommend that one person should not receive a service because the chances of benefit from the service are only 30% instead of the 50% that would be predicted without a test, could be regarded as a rather bold decision for a clinician to make about a person in need of help. Hunter and Schmidt (1990) have developed very useful approaches to validity generalization that usually result in estimates of test validity well above the correlations reported in actual use, but their estimates apply at the level of theory, construct validity, rather than at the level of specific application as in clinical settings.

A recommendation to improve the clinical uses of tests can actually be made: test for more things. Think of the determinants of performance in school, say college, as an example. College grades depend on motivation, persistence, physical health, mental health, study habits, and so on. If clinical psychologists are serious about predicting performance in college, then they probably will need to measure several quite different constructs and then combine all those measures into a prediction equation. The measurement task may seem onerous, but it is worth remembering Cronbach's (1960) band width vs. fidelity argument: it is often better to measure more things less well than to measure one thing extraordinarily well. A lot of measurement could be squeezed into the times usually allotted to low bandwidth tests. The genius of the profession will come in the determination of what to measure and how to measure it. The combination of all the information, however, is likely best to be done by a statistical algorithm for reasons that we will show later.

We are not negative toward psychological testing, but we think it is a lot more difficult and complicated than it is generally taken to be in practice. An illustrative case is provided by the differential diagnosis of attention deficit hyperactivity disorder (ADHD). There might be an ADHD scale somewhere but a more responsible clinical study would recognize that the diagnosis can be difficult, and that the validity and certainty of the diagnosis of ADHD is greatly improved by using multiple measures and multiple reporting agents across multiple contexts. For example, one authority recommended beginning with an initial screening interview, in which the possibility of an ADHD diagnosis is ruled in, followed by an extensive assessment battery addressing multiple domains and usually including (depending upon age): a Wechsler Intelligence Scale for Children (WISC-III; McCraken & McCallum, 1993), a behavior checklist (e.g., Youth Self-Report (YSR); Achenbach & Edelbrock, 1987), an academic achievement battery (e.g., Kaufmann Assessment Battery for Children; Kaufmann & Kaufmann, 1985), a personality inventory (e.g., Millon Adolescent Personality Inventory (MAPI); Millon & Davis, 1993), a computerized sustained attention and distractibility test (Gordon Diagnostic System [GDS]; McClure & Gordon, 1984), and a semistructured or a stuctured clinical interview (e.g., Diagnostic Interview Schedule for Children [DISC]; Costello, Edelbrock, Kalas, Kessler, & Klaric, 1982).

The results from the diagnostic assessment may be used to further rule in or rule out ADHD as a diagnosis, in conjunction with child behavior checklists (e.g., CBCL, Achenbach & Edelbrock, 1983; Teacher Rating Scales, Goyette, Conners, & Ulrich, 1978), completed by the parent(s) and teacher, and additonal school performance information. The parent and teacher complete both a historical list and then a daily behavior checklist for a period of two weeks in order to adequately sample behaviors. The information from home and school domains may be collected concurrently with evaluation of the diagnostic assessement battery, or the battery may be used initially to continue to rule in the diagnosis as a possibility, and then proceed with collateral data collection. We are impressed with the recommended ADHD diagnostic process, but we do recognize that it would involve a very extensive clinical process that would probably not be reimbursable under most health insurance plans. We would also note, however, that the overall diagnostic approach is not based on any decision-theoretic approach that might guide the choice of instruments corresponding to a process of decision making. Or alternatively, the process is not guided by any algorithm for combining information so as to produce a decision. Our belief is that assessment in clinical psychology needs the same sort of attention and systematic study as is occurring in medical areas through such organizations as the Society for Medical Decision Making.

In summary, we think the above scenario, or similar procedures using similar instruments (e.g., Atkins, Pelham, & White, 1990; Hoza, Vollano, & Pelham, 1995), represent an exemplar of assessment practice. It should be noted, however, that the development of such multimodal batteries is an iterative process. One will soon reach the point of diminishing returns in the development of such batteries, and the incremental validity (Sechrest, 1963) of instruments should be assessed. ADHD is an example in which the important domains of functioning are understood, and thus can be assessed. We know of no examples other that ADHD of such systematic approaches to assessment for decision making. Although approaches such as described here and by Pelham and his colleagues appear to be far from standard practice in the diagnosis of ADHD, we think they ought to be. The outlined procedure is modeled after a procedure developed by Gerald Peterson, Ph.D., Institute for Motivational Development, Bellevue, WA.

4.01.2 WHY ARE ASSESSMENTS DONE?

Why do we "test" in the first place? It is worth thinking about all the instances in which we do not test. For example, we usually do not test our own children—nor our spouses. That is because we have ample opportunities to observe the "performances" in which we are interested. That may be one reason that psychotherapists are disinclined to test their own clients: they have many opportunities to observe the behaviors in which they are interested, that is, if not the actual behaviors than reasonably good indicators of them. As we see it, testing is done primarily for one or more of three reasons: efficiency of observation, revealing cryptic conditions, and quantitative tagging.

Testing may provide for more efficient observation than most alternatives. For example, "tailing" a person, that method so dear to detective story writers, would prove definitive for many dispositions, but it would be expensive and often impractical or even unethical (Webb, Campbell, Schwartz, Sechrest, & Grove, 1981). Testing may provide for more efficient observation than most alternatives. It seems unlikely that any teacher would not have quite a good idea of the intelligence and personality of any of her pupils after at most a few weeks of a school year, but appropriate tests might provide useful information from the very first day. Probably clinicians involved in treating patients do not anticipate much gain in useful information after having held a few sessions with a patient. In fact, they may not anticipate much gain under most circumstances, which could account for the apparent infrequent use of assessment procedures in connection with psychological treatment.

Testing is also done in order to uncover "cryptic" conditions, that is, characteristics that are hidden from view or otherwise difficult to discern. In medicine, for example, a great many conditions are cryptic, blood pressure being one example. It can be made visible only by some device. Cryptic conditions have always been of great interest in clinical psychology, although their importance may have been exaggerated considerably. The Rorschach, a prime example of a putative decrypter, was hailed upon its introduction as "providing a window on the mind," and it was widely assumed that in skillful hands the Rorschach would make visible a wide range of hidden dispositions, even those unknown to the respondent (i.e., in "the unconscious"). Similarly, the Thematic Apperception Test was said to "expose underlying inhibited tendencies" of which the subject is unaware and to permit the subject to leave the test "happily unaware that he has presented the psychologist with what amounts to an X-ray picture of his inner self" (Murray, 1943, p. 1).

Finally, testing may be done, is often done, in order to provide a quantitative "tag" for some dispositions or other characteristic. In foot races, to take a mundane example, no necessity exists to time the races; it is sufficient to determine simply the order of the finish.

Nonetheless, races are timed so that each one may be quantitatively tagged for sorting and other uses, for example, making comparisons between races. Similarly, there is scarcely ever any need for more than a crude indicator of a child's intelligence, for example, "well above average," such as a teacher might provide. Nonetheless, the urge to seemingly precise quantification is strong, even if the precision is specious, and tests are used regularly to provide such estimates as "at the 78th percentile in aggression" or "IQ = 118." Although quantitative tags are used, and may be necessary, for some decision-making, for example, the awarding of scholarships based on SAT scores, it is to be doubted that such tags are ever of much use in clinical settings.

4.01.2.1 Bounded vs. Unbounded Inference and Prediction

Bounded prediction is the use of a test or measure to make some limited inference or prediction about an individual, couple, or family, a prediction that might be limited in time, situation, or range of behavior (Levy, 1963; Sechrest, 1968). Some familiar examples of bounded prediction are that of a college student's grade point average based on their SAT score, assessing the likely response of an individual to psychotherapy for depression based on MMPI scores and a SCID interview, or prognosticating outcome for a couple in marital therapy given their history. These predictions are bounded because they are using particular measures to predict a specified outcome in a given context. Limits to bounded predictions are primarily based on knowledge of two areas. First, the reliability of the information, that is, interview or test, for the population from which the individual is drawn. Second, and most important, these predictions are based on the relationship between the predictor and the outcome. That is to say, they are limited by the validity of the predictor for the particular context in question.

Unbounded inference or prediction, which is common in clinical practice, is the practice of making general assessment of an individual's tendencies, dispositions, and behavior, and inferring prognosis for situations that may not have been specified at the time of assessment. These are general statements made about individuals, couples, and families based on interviews, diagnostic tests, response to projective stimuli, and so forth that indicate how these people are likely to behave across situations. Some unbounded predictions are simply descriptive statements, for example, with respect to personality, from which at some future time the clinician or another person might make an inference about a behavior not even imagined at the time of the original assessment. A clinician might be asked to apply previously obtained assessment information to an individual's ability to work, ability as a parent, likelihood of behaving violently, or even the probability that an individual might have behaved in some way in the past (e.g., abused a spouse or child). Thus, they are unbounded in context. Since reliability and validity require context, that is, a measure is reliable in particular circumstances, one cannot readily estimate the reliability and validity of a measure for unspecified circumstances.

To the extent that the same measures are used repeatedly to make the same type of prediction or judgment about individuals, the more the prediction becomes of a bounded nature. Thus, an initially unbounded prediction becomes bounded by the consistency of circumstances of repeated use. Under these circumstances, reliability, utility, and validity can be assessed in a standard manner (Sechrest, 1968). Without empirical data, unbounded predictions rest solely upon the judgment of the clinician, which has proven problematic (see Dawes, Faust, & Meehl, 1989; Grove & Meehl, 1996; Meehl, 1954). Again, the contrast with medical testing is instructive. In medicine, tests are generally associated with gathering additional information about specific problems or systems. Although one might have a "wellness" visit to detect level of functioning and signs of potential problems, it would be scandalous to have a battery of medical tests to "see how your health might be" under an unspecified set of circumstances. Medical tests are bounded. They are for specific purposes at specific times.

4.01.2.2 Prevalence and Incidence of Assessment

It is interesting to speculate about how much assessment is actually done in clinical psychology today. It is equally interesting to realize how little is known about how much assessment is done in clinical psychology today. What little is known has to do with "incidence" of assessment, and that only from the standpoint of the clinician and only in summary form. Clinical psychologists report that a modest amount of their time is taken up by assessment activities. The American Psychological Association's (APA's) Committee for the Advancement of Professional Practice (1996) conducted a survey in 1995 of licensed APA members. With a response rate of 33.8%, the survey suggested that psychologists spend about 14% of their time conducting assessments—roughly six or seven hours per week. The low response rate, which ought to be considered disgraceful in a profession that claims to survive by science, is indicative of the difficulties involved in getting useful information about the practice of psychology in almost any area. The response rate was described as "excellent" in the report of the survey. Other estimates converge on about the same proportion of time devoted to assessment (Wade & Baker, 1977; Watkins, 1991; Watkins, Campbell, Nieberding, & Hallmark, 1995). Using data across a sizable number of surveys over a considerable period of time, Watkins (1991) concludes that about 50–75% of clinical psychologists provide at least some assessment services. We will say more later about the relative frequency of use of specific assessment procedures, but Watkins et al. (1995) did not find much difference in relative use across seven diverse work settings.

Think about what appears not to be known: the number of psychologists who do assessments in any period of time; the number of assessments that psychologists who do them actually do; the number or proportion of assessments that use particular assessment devices; the proportion of patients who are subjected to assessments; the problems for which assessments are done. And that does not exhaust the possible questions that might be asked. If, however, we take seriously the estimate that psychologists spend six or seven hours per week on assessment, then it is unlikely that those psychologists who do assessments could manage more than one or two per week; hence, only a very small minority of patients being seen by psychologists could be undergoing assessment. Wade and Baker (1977) found that psychologists claimed to be doing an average of about six objective tests and three projective tests per week, and that about a third of their clients were given at least one or the other of the tests, some maybe both. Those estimates do not make much sense in light of the overall estimate of only 15% of time (6–8 hours) spent in testing.

It is almost certain that those assessment activities in which psychologists do engage are carried out on persons who are referred by some other professional person or agency specifically for assessment. What evidence exists indicates that very little assessment is carried out by clinical psychologists on their own clients, either for diagnosis or for planning of treatment. Nor is there any likelihood that clinical psychologists refer their own clients to some other clinician for assessment. Some years ago, one of us (L. S.) began a study, never completed, of referrals made by clinical psychologists to other mental health professionals. The study was never completed in part because referrals were, apparently, very infrequent, mostly having to do with troublesome patients. A total of about 40 clinicians were queried, and in no instance did any of those clinical psychologists refer any client for psychological assessment.

Thus, we conclude that only a small minority of clients or patients of psychologists are subjected to any formal assessment procedures, a conclusion supported by Wade and Baker (1977) who found that relatively few clinicians appear to use standard methods of administration and scoring. Despite Wade and Baker's findings, it also seems likely that clinical psychologists do very little assessment on their own clients. Most assessments are almost certainly on referral. Now contrast that state of affairs with the practice of medicine: assessment is at the heart of medical practice. Scarcely a medical patient ever gets any substantial treatment without at least some assessment. Merely walking into a medical clinic virtually guarantees that body temperature and blood pressure will be measured. Any indication of a problem that is not completely obvious will result in further medical tests, including referral of patients from the primary care physician to other specialists.

The available evidence also suggests that psychologists do very little in the way of formal assessment of clients prior to therapy or other forms of intervention. For example, books on psychological assessment even in clinical psychology may not even mention psychotherapy or other interventions (e.g., see Maloney & Ward, 1976), and the venerated and authoritative *Handbook of psychotherapy and behavior change* (Bergen & Garfield, 1994) does not deal with assessment except in relation to diagnosis and the prediction of response to therapy and to determining the outcomes of therapy, that is, there is no mention of assessment for planning therapy at any stage in the process. That is, we think, anomalous, especially when one contemplates the assessment activities of other professions. It is almost impossible even to get to speak to a physician without at least having one's temperature and blood pressure measured, and once in the hands of a physician, almost all patients are likely to undergo further explicit assessment procedures, for example, auscultation of the lungs, heart, and carotid arteries. Unless the problem is completely obvious, patients are likely to undergo blood or other body-fluid tests, imaging procedures, assessments of functioning, and so on. The same contrast could be made for chiropractors, speech and hearing specialists, optometrists, and, probably, nearly all other clinical specialists. Clinical psychology appears to have no standard procedures, not much interest in them, and no instruments for carrying them out in any case. Why is that?

One reason, we suspect, is that clinical psychology has never shown much interest in normal functioning and, consequently, does not have very good capacity to identify normal responses or functioning. A competent specialist in internal medicine can usefully palpate a patient's liver, an organ he or she cannot see, because that specialist has been taught what a normal liver should feel like and what its dimensions should (approximately) be. A physician knows what normal respiratory sounds are. An optometrist certainly knows what constitutes normal vision and a normal eye. Presumably, a chiropractor knows a normal spine when he or she sees one.

Clinical psychology has no measures equivalent to body temperature and blood pressure, that is, quick, inexpensive screeners (vital signs) that can yield "normal" as a conclusion just as well as "abnormal." Moreover, clinical psychologists appear to have a substantial bias toward detection of psychopathology. The consequence is that clinical psychological assessment is not likely to provide a basis for a conclusion that a given person is "normal," and that no intervention is required. Obviously, the case is different for "intelligence," for which the conclusion of "average" or some such is quite common.

By their nature, psychological tests are not likely to offer many surprises. A medical test may reveal a completely unexpected condition of considerable clinical importance, for example, even in a person merely being subjected to a routine examination. Most persons who come to the attention of psychologists and other mental health professionals are there because their behavior has already betrayed important anomalies, either to themselves or to others. A clinical psychologist would be quite unlikely to administer an intelligence test to a successful business man and discover, completely unexpectedly, that the man was really "stupid." Tests are likely to be used only for further exploration or verification of problems already evident. If they are already evident, then the clinician managing the case may not see any particular need for further assessment.

A related reason that clinical psychologists appear to show so little inclination to do assessment of their own patients probably has to do with the countering inclination of clinical psychologists, and other similarly placed clinicians, to arrive at early judgments of patients based on initial impressions. Meehl (1960) noted that phenomenon many years ago, and it likely has not changed. Under those circumstances, testing of clients would have very little incremental value (Sechrest, 1963) and would seem unnecessary. At this point, it may be worth repeating that apparently no information is available on the specific questions for which psychologists make assessments when they do so.

Finally, we do believe that current limitations on practice imposed by managed care organizations are likely to limit even further the use of assessment procedures by psychologists. Pressures are toward very brief interventions, and that probably means even briefer assessments.

4.01.2.3 Proliferation of Assessment Devices

Clinical psychology has experienced an enormous proliferation of tests since the 1960s. We are referring here to commercially published tests, available for sale and for use in relation to clinical problems. For example, inspection of four current test catalogs indicates that there are at least a dozen different tests (scales, inventories, checklists, etc.) related to attention deficit disorder (ADD) alone, including forms of ADD that may not even exist, for example, adult ADD. One of the test catalogs is 100 pages, two are 176 pages, and the fourth is an enormous 276 pages. Even allowing for the fact that some catalog pages are taken up with advertisements for books and other such, the amount of test material available is astonishing. These are only four of perhaps a dozen or so catalogs we have in our files.

In the mid-1930s Buros published the first listings of psychological tests to help guide users in a variety of fields in choosing an appropriate assessment instrument. These early uncritical listings of tests developed into the *Mental measurements yearbook* and by 1937 the listings had expanded to include published test reviews. The *Yearbook*, which includes tests and reviews of new and revised tests published for commercial use, has continued to grow and is now in its 12th edition (1995). The most recent edition reviewed 418 tests available for use in education, psychology, business, and psychiatry. *Buros Mental Measurements Yearbook* is a valuable resource for testers, but it also charts the growth of assessment instruments. In addition to instruments published for commercial use, there are scores of other tests developed yearly for noncommercial use that are never reviewed by Buros. Currently, there are thousands of assessment instruments available for researchers and practitioners to choose from.

The burgeoning growth in the number of tests has been accompanied by increasing commercialization as well. The monthly *Monitor* published by the APA is replete with ads for test instruments for a wide spectrum of purposes. Likewise, APA conference attendees are inundated with preconference mailings advertising tests and detailing the location of the test publisher's booth at the conference site. Once at the conference, attendees are often struck by the slick presentation of the booths and hawking of the tests. Catalogs put out by test publishers are now also slick, in more ways than one. They are printed in color on coated paper and include a lot of messages about how convenient and useful the tests are with almost no information at all about reliability and validity beyond assurances that one can count on them.

The proliferation of assessment instruments and commercial development are not inherently detrimental to the field of clinical psychology. They simply make it more difficult to choose an appropriate test that is psychometrically sound, as glib ads can be used as a substitute for the presentation of sound psychometric properties and critical reviews. This is further complicated by the availability of computer scoring and software that can generate assessment reports. The ease of computer-based applications such as these can lead to their uncritical application by clinicians. Intense marketing of tests may contribute to their misuse, for example, by persuading clinical psychologists that the tests are remarkably simple and by convincing those same psychologists that they know more than they actually do about tests and their appropriate uses.

Multiple tests, even several tests for every construct, might not necessarily be a bad idea in and of itself, but we believe that the resources in psychology are simply not sufficient to support the proper development of so many tests. Few of the many tests available can possibly be used on more than a very few thousand cases per year, and perhaps not even that. The consequence is that profit margins are not sufficient to support really adequate test development programs. Tests are put on the market and remain there with small normative samples, with limited evidence for validity, which is much more expensive to produce than evidence for reliability, and with almost no prospect for systematic exploration of the other psychometric properties of the items, for example, discrimination functions or tests of their calibration (Sechrest, McKnight, & McKnight, 1996).

One of us (L. S.) happens to have been a close spectator of the development of the SF-36, a now firmly established and highly valued measure of health and functional status (Ware & Sherbourne, 1992). The SF-36 took 15–20 years for its development, having begun as an item pool of more than 300 items. Over the years literally millions of dollars were invested in the development of the test, and it was subjected, often repeatedly, to the most sophisticated psychometric analyses and to detailed scrutiny of every individual item. The SF-36 has now been translated into at least 37 languages and is being used in an extraordinarily wide variety of research projects. More important, however, the SF-36 is also being employed routinely in evaluating outcomes of clinical medical care. Plans are well advanced for use of the SF-36 that will result in its administration to 300 000 patients in managed care every year. It is possible that over the years the Wechsler intelligence tests might have a comparable history of development, and the Minnesota Multiphasic Inventory (MMPI) has been the focus of a great many investigations, as has the Rorschach. Neither of the latter, however, has been the object of systematic development efforts funded centrally, and scarcely any of the many other tests now available are likely to be subjected to anything like the same level of development effort (e.g., consider that in its more than 70-year history, the Rorschach has never been subjected to any sort of revision of its original items).

Several factors undoubtedly contribute to the proliferation of psychological tests (not the least, we suspect, being their eponymous designation and the resultant claim to fame), but surely one of the most important would be the fragmentation of psychological theory, or what passes for theory. In 1995 a taskforce was assembled under the auspices of the APA to try to devise a uniform test (core) battery that would be used in all psychotherapy research studies (Strupp, Horowitz, & Lambert, 1997). The effort failed, in large part because of the many points of view that seemingly had to be represented and the inability of the conferees to agree even on any outcomes that should be common to all therapies. Again, the contrast with medicine and the nearly uniform acceptance of the SF-36 is stark.

Another reason for the proliferation of tests in psychology is, unquestionably, the seeming ease with which they may be "constructed." Almost anyone with a reasonable "construct" can write eight or 10 self-report items to "measure" it, and most likely the new little scale will have "acceptable" reliability. A correlation or two with some other measure will establish its "construct validity," and the rest will eventually be history. All that is required to establish a new projective test, it seems, is to find a set of stimuli that have not, according to the published literature, been used before and then show that responses to the stimuli are suitably strange, perhaps stranger for some folks than others. For example, Sharkey and Ritzler (1985) noted a new Picture Projective Test that was created by using photographs from a photo essay. The pictures were apparently selected based on the authors' opinions about their ability to elicit "meaningful projective material," meaning responses with affective content and activity themes. No information was given pertaining to comparison of various pictures and their responses nor relationships to other measures of the target constructs; no comparisons were made to pictures that were deemed inappropriate. The "validation" procedure simply compared diagnoses to those in charts and results of the TAT. Although rater agreement was assessed, there was no formal measurement of reliability.

New tests are cheap, it seems. One concern is that so many new tests appear also to imply new constructs, and one wonders whether clinical psychology can support anywhere near as many constructs as are implied by the existence of so many measures of them. Craik (1986) made the eminently sensible suggestion that every "new" or infrequently used measure used in a research project should be accompanied by at least one well-known and widely used measure from the same or a closely related domain. New measures should be admitted only if it is clear that they measure something of interest and are not redundant, that is, have discriminant validity. That recommendation would likely have the effect of reducing the array of measures in clinical psychology by remarkable degrees if it were followed.

The number of tests that are taught in graduate school for clinical psychology is far lower than the number available for use. The standard stock-in-trade are IQ tests such as the Wechsler Adult Intelligence Scale (WAIS), personality profiles such as the MMPI, diagnostic instruments (Structured Clinical Interview for *DSM-III-R* [SCID]), and at some schools, the Rorschach as a projective test. This list is rounded out by a smattering of other tests like the Beck Depression Inventory and Millon. Recent standard application forms for clinical internships developed by the Association of Psychology Postdoctoral and Internship Centers (APPIC) asked applicants to report on their experience with 47 different tests and procedures used for adult assessment and 78 additional tests used with children! It is very doubtful that training programs actually provide training in more than a handful of the possible devices.

Training in testing (assessment) is not at all the same as training in measurement and psychometrics. Understanding how to administer a test is useful but cannot substitute for evaluating the psychometric soundness of tests. Without grounding in such principles, it is easy to fall prey to glib ads and ease of computer administration without questioning the quality of the test. Psychology programs appear, unfortunately, to be abandoning training in basic measurement and its theory (Aiken, West, Sechrest, & Reno, 1990).

4.01.2.4 Over-reliance on Self-report

"Where does it hurt?" is a question often heard in physicians' offices. The physician is asking the patient to self-report on the subjective experience of pain. Depending on the answer, the physician may prescribe some remedy, or may order tests to examine the pain more thoroughly and obtain objective evidence about the nature of the affliction before pursuing a course of treatment. The analog heard in psychologists' offices is "How do you feel?" Again, the inquiry calls forth self-report on a subjective experience and like the physician, the psychologist may determine that tests are in order to better understand what is happening with the client.

When the medical patient goes for testing, she or he is likely to be poked, prodded, or pricked so that blood samples and X-rays can be taken. The therapy client, in contrast, will most likely be responding to a series of questions in an interview or answering a pencil-and-paper questionnaire. The basic difference between these is that the client in clinical psychology will continue to use self-report in providing a sample, whereas the medical patient will provide objective evidence.

Despite the proliferation of tests in recent years, few rely on evidence other than the client's self-report for assessing behavior, symptoms, or mood state. Often assessment reports remark that the information gleaned from testing was corroborated by interview data, or vice versa, without recognizing that both rely on self-report alone. The problems with self-report are well documented: poor recall of past events, motivational differences in responding, social desirability bias, and malingering, for example. Over-reliance on self-report is a major criticism of psychological assessment as it is currently conducted and was the topic of a recent conference sponsored by the National Institute of Mental Health.

What alternatives are there to self-report? Methods of obtaining data on a client's behavior that do not rely on self-report do exist. Behavioral observation with rating by judges can permit the assessment of behavior, often without the client's awareness or outside the confines of an office setting. Use of other informants such as family members or co-workers to provide data can yield valuable information about a client. Yet, all too often these alternatives are not pursued because they involve time or resources—in short, they are demanding approaches. Compared with asking a client about his or her mood state over the last week, organizing field work or contacting informants involves a great deal more work and time.

Instruments are available to facilitate collection of data not relying so strongly on self-report and for collection of data outside the office setting, for example, the Child Behavior Checklist (CBCL; Achenbach & Edelbrock, 1983). The CBCL is meant to assist in diagnosing a range of psychological and behavior problems in children, and it relies on parent, teacher, and self-reports of behavior. Likewise, neuropsychological tests utilize functional performance measures much more than self-report. However, as Craik (1986) noted with respect to personality research, methods such as field studies are not widely used as alternatives to self-report. This problem of over-reliance on self-report is not new (see Webb, Campbell, Schwartz, & Sechrest, 1966).

4.01.3 PSYCHOMETRIC ISSUES WITH RESPECT TO CURRENT MEASURES

Consideration of the history and current status of clinical assessment must deal with some fundamental psychometric issues and practices. Although psychometric is usually taken to refer to reliability and validity of measures, matters are much more complicated than that, particularly in light of developments in psychometric theory and method since the 1960s, which seem scarcely to have penetrated clinical assessment as an area. Specifically, generalizability theory and Item Response Theory (IRT) offer powerful tools with which to explore and develop clinical assessment procedures, but they have seen scant use in that respect.

4.01.3.1 Reliability

The need for "reliable" measures is by now well accepted in all of psychology, including clinical assessment. What is not so widespread is the necessary understanding of what constitutes reliability and the various uses of that term. In their now classic presentation of generalizability theory, Cronbach and his associates (Cronbach, Gleser, Nanda, & Rajaratnam, 1972) used the term "dependability" in a way that is close to what is meant by reliability, but they made especially clear, as classical test theory had not, that measures are dependable (generalizable) in very specific ways, that is, that they are dependable across some particular conditions of use (facets), and assessments of dependability are not at all interchangeable. For example, a given assessment may be highly dependable across particular items but not necessarily across time. An example might be a measure of mood, which ought to have high internal consistency (i.e., across items) but that might not, in fact, should not, have high dependability over time, else the measure would be better seen as a trait rather than as a mood measure.

An assessment procedure might be highly dependable in terms of internal consistency and across time but not satisfactorily dependable across users, for example, being susceptible to a variety of biases characteristic of individual clinicians. Or an assessment procedure might not be adequately dependable across conditions of its use, as might be the case when a measure is taken from a research to a clinical setting. Or an assessment procedure might not be dependable across populations, for example, a projective instrument useful with mental patients might be misleading if used with imaginative and playful college students.

Issues of dependability are starkly critical when one notes the regrettably common practice of justifying the use of a measure on the ground that it is "reliable," often without even minimal specification of the facet(s) across which that reliability was established. The practice is even more regrettable when, as is often the case, only a single value for reliability is given when many are available and when one suspects that the figure reported was not chosen randomly from those available. Moreover, it is all too frequently the case that the reliability estimate reported is not directly relevant to the decisions to be made. Internal consistency, for example, may not be as important as generalizability over time when one is using a screening instrument. That is, if one is screening in a population for psychopathology, it may not be of great interest that two persons with the same scores are different in terms of their manifestations of pathology, but it is of great interest whether if one retested them a day or so later, the scores would be roughly consistent.

In short, clinical assessment in psychology is unfortunately casual in its use of reliability estimates, and it is shamefully behind the curve in its attention to the advantages provided by generalizability theory, originally proposed in 1963 (Cronbach, Rajaratnam, & Gleser, 1963).

4.01.3.2 Validity

It is customary to treat validity of measures as a topic separate from reliability, but we think that is not only unnecessary but undesirable. In our view, the validity of measures is simply an extension of generalizability theory to the question of what other performances aside from those involved in the test is the score generalizable. A test score that is generalizable to another very similar performance, say on the same set of test items or over a short period of time, is said to be reliable. A test score that is generalizable to a score on another similar test is sometimes said to be "valid," but we think that a little reflection will show that unless the tests demand very different kinds of performances, generalizability from one test to another is not much beyond the issues usually regarded as having to do with reliability. When, however, a test produces a score that is informative about another very different kind of performance, we gradually move over into the realm termed validity, such as when a paper-and-pencil test of "readiness for change" (Prochaska, DiClemente, & Norcross, 1992) predicts whether a client will benefit from treatment or even just stay in treatment.

We will say more later about construct validity, but a test or other assessment procedure may be said to have construct validity if it produces generalizable information and if that information relates to performances that are conceptually similar to those implied by the name or label given to the test. Essentially, however, any measure that does not produce scores by some random process is by that definition generalizable to some other performance and, hence, to that extent may be said to be valid. What a given measure is valid for, that is, generalizable to, however, is a matter of discovery as much as of plan. All instruments used in clinical assessment should be subjected to comprehensive and continuing investigation in order to determine the sources of variance in scores. An instrument that has good generalizability over time and across raters may turn out to be, among other things, a very good measure of some response style or other bias. The MMPI includes a number of "validity" scales designed to assess various biases in performance on it, and it has been subjected to many investigations of bias. The same cannot be said of some other widely used clinical assessment instruments and procedures. To take the most notable example, of the more than 1000 articles on the Rorschach that are in the current PsychInfo database, only a handful, about 1%, appear to deal with issues of response bias, and virtually all of those are on malingering and most of them are unpublished dissertations.

4.01.3.3 Item Response Theory

Although Item Response Theory (IRT) is a potentially powerful tool for the development and study of measures of many kinds, its use to date has not been extensive beyond the area of ability testing. The origins of IRT go back at least to the early 1950s and the publication of Lord's (1952) monograph, *A theory of test scores,* but it has had little impact on measurement outside the arena of ability testing (Meier, 1994). Certainly it has had almost no impact on clinical assessment. The current PsychInfo database includes only two references to IRT in relation to the MMPI and only one to the Rorschach, and the latter one, now 10 years old, is an entirely speculative mention of a potential application of IRT (Samejima, 1988).

IRT, perhaps to some extent narrowly imagined to be relevant only to test construction, can be of great value in exploring the nature of measures and improving their interpretation. For example, IRT can be useful in understanding just when scores may be interpreted as unidimensional and then in determining the size of gaps in underlying traits represented by adjacent scores. An example could be the interpretation of Whole responses on the Rorschach. Is the W score a unidimensional score, and, if so, is each increment in that score to be interpreted as an equal increment? Some cards are almost certainly more difficult stimuli to which to produce a W response, and IRT could calibrate that aspect of the cards. IRT would be even more easily used for standard paper-and-pencil inventory measures, but the total number of applications to date is small, and one can only conclude that clinical assessment is being short-changed in its development.

4.01.3.4 Scores on Tests

Lord's (1952) monograph was aimed at tests with identifiable underlying dimensions such as ability. Clinical assessment appears never to have had any theory of scores on instruments included under that rubric. That is, there seems never to have been proposed or adapted any unifying theory about how test scores on clinical instruments come about. Rather there seems to have been a passive, but not at all systematic, adoption of general test theory, that is, the idea that test scores are in some manner generated by responses representing some underlying trait. That casual approach cannot forward the development of the field.

Fiske (1971) has come about as close as anyone to formulating a theory of test scores for clinical assessment, although his ideas pertain more to how such tests are scored than to how they come about, and his presentation was directed toward personality measurement rather than clinical assessment. He suggested several models for scoring test, or otherwise observed, responses. The simplest model is what we may call the cumulative frequency model, which simply increments the score by 1 for every observed response. This is the model that underlies many Rorschach indices. It assumes that every response is equivalent to every other one, and it ignores the total number of opportunities for observation. Thus, each Rorschach W response counts as 1 for that index, and the index is not adjusted to take account of the total number of responses. A second model is the relative frequency model, which forms an index by dividing the number of observed critical responses by some indicator of opportunities to form a rate of responding, for example, as would be accomplished by counting W responses and dividing by the total number of responses or by counting W responses only for the first response to each card. Most paper-and-pencil inventories are scored implicitly in that way, that is, they count the number of critical responses in relation to the total number possible.

A long story must be made short here, but Fiske describes other models, and still more are possible. One may weight responses according to the inverse of their frequency in a population on the grounds that common responses should count for less than rare responses. Or one may weight responses according to the judgments of experts. One can assign the average weight across a set of responses, a common practice, but one can also assign as the score the weight of the most extreme response, for example, as runners are often rated on the basis of their fastest time for any given distance. Pathology is often scored in that way, for example, a pathognomic response may outweigh many mundane, ordinary responses.

The point is that clinical assessment instruments and procedures only infrequently have any explicit basis in a theory of responses. For the most part, scores appear to be derived in some standard way without much thought having been given to the process. It is not clear how much improvement in measures might be achieved by more attention to the development of a theory of scores, but it surely could not hurt to do so.

4.01.3.5 Calibration of Measures

A critical limitation on the utility of psychological measures of any kind, but certainly in their clinical application, is the fact that the measures do not produce scores in any directly interpretable metric. We refer to this as the *calibration* problem (Sechrest, McKnight, & McKnight, 1996). The fact is that we have only a very general knowledge of how test scores may be related to any behavior of real interest. We may know in general that a score of 70, let us say, on an MMPI scale is "high," but we do not know very well what might be expected in the behavior of a person with such a score. We would know even less about what difference it might make if the score were reduced to 60 or increased to 80 except that in one case we might expect some diminution in problems and in the other some increase. In part the lack of calibration of measures in clinical psychology stems from lack of any specific interest and diligence in accomplishing the task. Clinical psychology has been satisfied with "loose calibration," and that stems in part, as we will assert later, from adoption of the uninformative model of significance testing as a standard for validation of measures.

4.01.4 WHY HAVE WE MADE SO LITTLE PROGRESS?

It is difficult to be persuaded that progress in assessment in clinical psychology has been substantial in the past 75 years, that is, since the introduction of the Rorschach. Several arguments may be adduced in support of that statement, even though we recognize that it will be met with protests. We will summarize what we think are telling arguments in terms of theory, formats, and validities of tests.

First, we do not discern any particular improvements in theories of clinical testing and assessments over the past 75 years. The Rorschach, and the subsequent formulation of the projective hypothesis, may be regarded as having been to some extent innovations; they are virtually the last ones in the modern history of assessment. As noted, clinical assessment lags well behind the field in terms of any theory of either the stimuli or responses with which it deals, let alone the connections between them. No theory of assessment exists that would guide selection of stimuli to be presented to subjects, and certainly none pertains to the specific format of the stimuli nor to the nature of the responses required. Just to point to two simple examples of the deficiency in understanding of response options, we note that there is no theory to suggest whether in the case of a projective test responses should be followed by any sort of inquiry about their origins, and there is no theory to suggest in the case of self-report inventories whether items should be formulated so as to produce endorsements of the "this is true of me" nature or so as to produce descriptions such as "this is what I do."

Given the lack of any gains in theory about the assessment enterprise, it is not surprising that there have also not been any changes in test formats since the introduction of the Rorschach.

Projective tests based on the same simple (and inadequate) hypothesis are still being devised, but not one has proven itself in any way better than anything that has come before. Item writers may be a bit more sophisticated than those in the days of the Bernreuter, but items are still constructed in the same way, and response formats are the same as ever, "agree–disagree," "true–false," and so on.

Even worse, however, is the fact that absolutely no evidence exists to suggest that there have been any mean gains in the validities of tests over the past 75 years. Even for tests of intellectual functioning, typical correlations with any external criterion appear to average around 0.40, and for clinical and personality tests the typical correlations are still in the range of 0.30, the so-called "personality coefficient." This latter point, that validities have remained constant, may, of course, be related to the lack of development of theory and to the fact that the same test formats are still in place.

Perhaps some psychologists may take exception to the foregoing and cite considerable advances. Such claims are made for the Exner (1986) improvements on the Rorschach, known as the "comprehensive system," and for the MMPI-2, but although both claims are superficially true, there is absolutely no evidence for either claim from the standpoint of validity of either test. The Exner comprehensive system seems to have "cleaned up" some aspects of Rorschach scoring, but the improvements are marginal, for example, it is not as if inter-rater reliability increased from 0.0 to 0.8, and no improvements in validity have been established. Even the improvements in scoring have been demonstrated for only a portion of the many indexes. The MMPI-2 was only a cosmetic improvement over the original, for example, getting rid of some politically incorrect items, and no increase in the validity of any score or index seems to have been demonstrated, nor is any likely.

An additional element in the lack of evident "progress" in the validity of test scores may be lack of reliability (and validity!) of people being predicted. (One wise observer suggested that we would not really like it at all if behavior were 90% predictable! Especially our own.) We may just have reached the limits of our ability to predict what is going to happen with and to people, especially with our simple-minded and limited assessment efforts. As long as we limit our assessment efforts to the dispositions of the individuals who are clients and ignore their social milieus, their real environmental circumstances, their genetic possibilities, and so on, we may not be able to get beyond correlations of 0.3 or 0.4.

The main "advance" in assessment over the past 75 years is not that we do anything really better but that we do it much more widely. We have many more scales than existed in the past, and we can at least assess more things than ever before, even if we can do that assessment only, at best, passably well.

Woodworth (1937/1992) wrote in his article on the future of clinical psychology that, "There can be no doubt that it will advance, and in its advance throw into the discard much guesswork and half-knowledge that now finds baleful application in the treatment of children, adolescents and adults" (p. 16). It appears to us that the opposite has occurred. Not only have we failed to discard guesswork and half-knowledge, that is, tests and treatments with years of research indicating little effect or utility, we have continued to generate procedures based on the same flawed assumptions with the misguided notion that if we just make a bit of a change here and there, we will finally get it right. Projective assessments that tell us, for example, that a patient is psychotic are of little value. Psychologists have more reliable and less expensive ways of determining this. More direct methods have higher validity in the majority of cases. The widespread use of these procedures at high actual and opportunity cost is not justified by the occasional addition of information. It is not possible to know ahead of time which individuals might give more information via an indirect method, and most of the time it is not even possible to know afterwards whether indirectly obtained "information" is correct unless the information has also been obtained in some other way, that is, asking the person, asking a relative, or doing a structured interview. It is unlikely that projective test responses will alter clinical intervention in most cases, nor should it.

Is it fair to say that clinical psychology has no standards (see Sechrest, 1992)? Clinical psychology gives the appearance of standards with accreditation of programs, internships, licensure, ethical standards, and so forth. It is our observation, however, that there is little to no monitoring of the purported standards. For example, in reviewing recent literature as background to this chapter, we found articles published in peer-reviewed journals using projective tests as outcome measures for treatment. The APA ethical code of conduct states that psychologists "... use psychological assessment ... for purposes that are appropriate in light of the research on or evidence of the... proper application of the techniques." The APA document, *Standards for educational and psychological testing*, states:

> ... Validity however, is a unitary concept. Although evidence may be accumulated in may ways, validity always refers to the degree to which that evidence supports the inferences that are made from the scores. The inferences regarding specific uses of a test are validated, not the test itself. (APA, 1985, p. 9)

Further, the section titled, *Professional standards for test use* (APA, 1985, p. 42, Standard 6.3) states:

> When a test is to be used for a purpose for which it has not been previously validated, or for which there is no supported claim for validity, the user is responsible for providing evidence of validity.

No body of research exists to support the validity of any projective instrument as the sole outcome measure for treatment—or as the sole measure of anything. So not only do questionable practices go unchecked, they can result in publication.

4.01.4.1 The Absence of the Autopsy

Medicine has always been disciplined by the regular occurrence of the autopsy. A physician makes a diagnosis and treats a patient, and if the patient dies, an autopsy will be done, and the physician will receive feedback on the correctness of his or her diagnosis. If the diagnosis were wrong, the physician would to some extent be called to account for that error; at least the error would be known, and the physician could not simply shrug it off. We know that the foregoing is idealized, that autopsies are not done in more than a fraction of cases, but the model makes our point. Physicians make predictions, and they get feedback, often quickly, on the correctness of those predictions. Surgeons send tissue to be biopsied by pathologists who are disinterested; internists make diagnoses based on various signs and symptoms and then order laboratory procedures that will inform them about the correctness of their diagnosis; family practitioners make diagnoses and prescribe treatment, which, if it does not work, they are virtually certain to hear about.

Clinical psychology has no counterpart to the autopsy, no systematic provision for checking on the correctness of a conclusion and then providing feedback to the clinician. Without some form of systematic checking and feedback, it is difficult to see how improvement in either instruments or clinicians' use of them could be regularly and incrementally improved. Psychologist clinicians have been allowed the slack involved in making unbounded predictions and then not getting any sort of feedback on the potential accuracy of even those loose predictions. We are not sure how much improvement in clinical assessment might be possible even with exact and fairly immediate feedback, but we are reasonably sure that very little improvement can occur without it.

4.01.5 FATEFUL EVENTS CONTRIBUTING TO THE HISTORY OF CLINICAL ASSESSMENT

The history of assessment in clinical psychology is somewhat like the story of the evolution of an organism in that at critical junctures, when the development of assessment might well have gone one way, it went another. We want to review here several points that we consider to be critical in the way clinical assessment developed within the broader field of psychology.

4.01.5.1 The Invention of the Significance Test

The advent of hypothesis testing in psychology had fateful consequences for the development of clinical assessment, as well as for the rest of psychology (Gigerenzer, 1993). Hypothesis testing encouraged a focus on the question whether any predictions or other consequences of assessment were "better than chance," a distinctly loose and undemanding criterion of "validity" of assessment. The typical validity study for a clinical instrument would identify two groups that would be expected to differ in some "score" derived from the instrument and then ask the question whether the two groups did in fact (i.e., to a statistically significant degree) differ in that score. It scarcely mattered by how much they differed or in what specific way, for example, an overall mean difference vs. a difference in proportions of individuals scoring beyond some extreme or otherwise critical value. The existence of any "significant" difference was enough to justify triumphant claims of validity.

4.01.5.2 Ignoring Decision Making

One juncture had to do with bifurcation of the development of clinical psychology from other streams of assessment development. Specifically, intellectual assessment and assessment of various capacities and propensities relevant to performance in work settings veered in the direction of assessment for decision-making (although not terribly sharply nor completely), while assessment in clinical psychology went in the direction of assessment for enlightenment. What eventually happened is that clinical psychology failed to adopt any rigorous criterion of correctness of decisions made on the basis of assessed performance, but adopted instead a conception of assessments as generally informative or "correct."

Simply to make the alternative clear, the examples provided by medical assessment are instructive. The model followed in psychology would have resulted in medical research of some such nature as showing that two groups that "should" have differed in blood pressure, for example, persons having just engaged in vigorous exercise vs. persons having just experienced a rest period, differed significantly in blood pressure readings obtained by a sphygmomanometer. Never mind by how much they differed or what the overlap between the groups. The very existence of a "significant" difference would have been taken as evidence for the "validity" of the sphygmomanometer.

Instead, however, medicine focused more sharply on the accuracy of decisions made on the basis of assessment procedures. The aspect of biomedical assessment that most clearly distinguishes it from clinical psychological assessment is its concern for sensitivity and specificity of measures (instruments) (Kraemer, 1992). Kraemer's book, *Evaluating medical tests: Objective and quantitative guidelines,* has not even a close counterpart in psychology, which is, itself, revealing. These two characteristics of measures are radically different from the concepts of validity used in psychology, although "criterion validity" (now largely abandoned) would seem to require such concepts.

Sensitivity refers to the proportion of cases having a critical characteristic that are identified by the test. For example, if a test were devised to select persons likely to benefit from some form of therapy, sensitivity would refer to the proportion of cases that would actually benefit which would be identified correctly by the test. These cases would be referred to as "true positives." Any cases that would benefit from the treatment but that could not be identified by the test would be "false-negatives" in this example. Conversely, a good test should have high specificity, which would be avoiding "false-positives," or incorrectly identifying as good candidates for therapy persons who would not actually benefit. The "true negative" group would be those persons who would not benefit from treatment, and a good test should correctly identify a large proportion of them.

As Kraemer (1992) points out, sensitivity and specificity as test requirements are nearly always in opposition to each other, and are reciprocal. Maximizing one requirement reduces the other. Perfect sensitivity can be attained by, in our example, a test that identifies every case as suitable for therapy; no amenable cases are missed. Unfortunately, that maneuver would also maximize the number of false-positives, that is, many cases would be identified as suitable for therapy who, in fact, were not. Obviously, the specificity of the test could be maximized by declaring all cases as unsuitable for therapy, thus ensuring that the number of false-positives would be zero—while at the same time ensuring that the number of false-negatives would be maximal, and no one would be treated.

We go into these issues in some detail in order to make clear how very different such thinking is from usual practices in clinical psychological assessment. The requirements for Receiver Operating Curves (ROC), which is the way issues of sensitivity and specificity of measures are often labeled and portrayed, are stringent. They are not satisfied by simple demonstrations that measures, for example, suitability for treatment, are "significantly related to" other measures of interest, for example, response to treatment. The development of ROC statistics almost always occurs in the context of the use of tests for decision-making: treat–not treat, hire–not hire, do further tests–no further tests. Those kinds of uses of tests in clinical psychological assessment appear to be rare.

Issues of sensitivity-specificity require the existence of some reasonably well-defined criterion, for example, the definition of what is meant by favorable response to treatment and a way of measuring it. In biomedical research, ROC statistics are often developed in the context of a "gold standard," a definitive criterion. For example, an X ray might serve as a gold standard for a clinical judgment about the existence of a fracture, or a pathologist's report on a cytological analysis might serve as a gold standard for a screening test designed to detect cancer. Clinical psychology has never had anything like a gold standard against which its various tests might have been validated. Psychiatric diagnosis has sometimes been of interest as a criterion, and tests of different types have been examined to determine the extent to which they produce a conclusion in agreement with diagnosis (e.g., Somoza, Steer, Beck, & Clark, 1994), but in that case the gold standard is suspect, and it is by no means clear that disagreement means that the test is wrong.

The result is that for virtually no psychological instrument is it possible to produce a useful quantitative estimate of its accuracy. Tests and other assessment devices in clinical psychology have been used for the most part to produce general enlightenment about a target of interest rather than to make a specific prediction of some outcome. People who have been tested are described as "high in anxiety," "clinically depressed," or "of average intelligence." Statements of that sort, which we have referred to previously as unbounded predictions, are possibly enlightening about the nature of a person's functioning or about the general range within which problems fall, but they are not specific predictions, and are difficult to refute.

4.01.5.3 Seizing on Construct Validity

In 1955, Cronbach and Meehl published what is arguably the most influential article in the field of measurement: *Construct validity in psychological tests* (Cronbach & Meehl, 1955). This is the same year as the publication of *Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores* (Meehl & Rosen, 1955). It is safe to say that no two more important articles about measurement were ever published in the same year. The propositions set forth by Cronbach and Meehl about the validity of tests were provocative and rich with implications and opportunities. In particular, the idea of construct validity required that measures be incorporated into elaborated theoretical structure, which was labeled the "nomological net." Unfortunately, the fairly daunting requirements for embedding measures in theory were mostly ignored in clinical assessment (the same could probably be said about most other areas of psychology, but it is not our place here to say so), and the idea of construct validity was trivialized.

The trivialization of construct validity reflects in part the fact that no standards for construct validity exist (and probably none can be written) and the general failure to distinguish between necessary and sufficient conditions for the inference of construct validity. In their presentation of construct validity, Cronbach and Meehl did not specify any particular criteria for sufficiency of evidence, and it would be difficult to do so. Construct validity exists when everything fits together, but trying to specify the number and nature of the specific pieces of evidence would be difficult and, perhaps, antithetical to the idea itself. It is also not possible to quantify level or degree of construct validity other than in a very rough way and such quantifications are, in our experience, rare. It is difficult to think of an instance of a measure described as having "moderate or "low" construct validity, although "high" construct validity is often implied.

It is possible to imagine what some of the necessary conditions for construct validity might be, one notable requirement being convergent validity (Campbell & Fiske, 1959). In some manner that we have not tried to trace, conditions necessary for construct validity came to be viewed as sufficient. Thus, for example, construct validity usually requires that one measure of a construct correlates with another. Such a correlation is not, however, a sufficient condition for construct validity, but, nonetheless, a simple zero-order correlation between two tests is often cited as "evidence" for the construct validity of one measure or the other. Even worse, under the pernicious influence of the significance testing paradigm, any statistically significant correlation may be taken as evidence of "good construct validity." Or, for another example, construct validity usually requires a particular factor structure for a measure, but the verification of the required factor structure is not sufficient evidence for construct validity of the measure involved. The fact that a construct is conceived as unidimensional does not mean that a measure alleged to represent the construct does so simply because it appears to form a single factor.

The net result of the dependence on significance testing and the poor implementation of the ideas represented by construct validity has been that the standards of evidence for the validity of psychological measures has been distressingly low.

4.01.5.4 Adoption of the Projective Hypothesis

The projective hypothesis (Frank, 1939) is a general proposition stating that whatever an individual does when exposed to an ambiguous stimulus will reveal important aspects of his or her personality. Further, the projective hypothesis suggests that indirect responses, that is, those to ambiguous stimuli, are more valid than direct responses, that is, those to interviews or questionnaires. There is little doubt that indirect responses reveal something about people, although whether that which is revealed is, in fact, important is more doubtful. Moreover, what one eats, wears, listens to, reads, and so on are rightly considered to reveal something about that individual. While the general proposition about responses to ambiguous stimuli appears quite reasonable, the use of such stimuli in the form of projective tests has proven problematic and of limited utility.

The course of development of clinical assessment might have been different and more useful had it been realized that projection was the wrong term for the link between ambiguous stimuli and personality. A better term would have been the "expressive hypothesis," the notion that an individual's personality may be manifest (expressed) in response to a wide range of stimuli, including ambiguous stimuli. Personality style might have come to be of greater concern, and unconscious determinants of behavior, implied by projection, might have received less emphasis.

In any case, when clinical psychology adopted the projective hypothesis and bought wholesale into the idea of unconscious determinants of behavior, that set the field on a course that has been minimally productive but that still affects an extraordinarily wide range of clinical activities. Observable behaviors have been downplayed and objective measures treated with disdain or dismissed altogether. The idea of peering into the unconscious appealed both to psychological voyeurs and to those bent on achieving the glamour attributed to the psychoanalyst.

Research on projective stimuli indicates that highly structured stimuli which limit the dispositions tapped increase the reliability of such tests (e.g., Kagan, 1959). In achieving acceptable reliability, the nature of the test is altered in such a way that the stimulus is less ambiguous and the likelihood of an individual "projecting" some aspect of their personality in an unusual way becomes reduced. Thus, the dependability of responses to projective techniques probably depends to an important degree on sacrificing their projective nature. In part, projective tests seem to have failed to add to assessment information because most of the variance in responses to projective stimuli is accounted for by the stimuli themselves. For example, "popular" responses on the Rorschach are popular because the stimulus is the strongest determinant of the response (Murstein, 1963).

Thorndike (Thorndike & Hagen, 1955, p. 418), in describing the state of affairs with projective tests some 40 years ago, stated:

> A great many of the procedures have received very little by way of rigorous and critical test and are supported only by the faith and enthusiasm of their backers. In those few cases, most notable that of the Rorschach, where a good deal of critical work has been done, results are varied and there is much inconsistency in the research picture. Modest reliability is usually found, but consistent evidence of validity is harder to come by.

The picture has not changed substantially in the ensuing 40 years and we doubt that it is likely to change much in the next 40. As Adcock (1965, cited in Anastasi, 1988) noted, "There are still enthusiastic clinicians and doubting statisticians." As noted previously (Sechrest, 1963, 1968), these expensive and time-consuming projective procedures add little if anything to the information gained by other methods and their abandonment by clinical psychology would not be a great loss. Despite lack of incremental validity after decades of research, not only do tests such as the Rorschach and TAT continue to be used, but new projective tests continue to be developed. That could be considered a pseudoscientific enterprise that, at best, yields procedures telling clinical psychologists what they at least should already know or have obtained in some other manner, and that, at worst, wastes time and money and further damages the credibility of clinical psychology.

4.01.5.5 The Invention of the Objective Test

At one time we had rather supposed without thinking about it too much that objective tests had always been around in some form or other. Samelson (1987), however, has shown that at least the multiple-choice test was invented in the early part of the twentieth century, and it seems likely that the true–false test had been devised not too long before then. The objective test revolutionized education in ways that Samelson makes clear, and it was not long before that form of testing infiltrated into psychology. Bernreuter (1933) is given credit for devising the first multiphasic (multidimensional) personality inventory—only 10 years after the introduction of the Rorschach into psychology.

Since 1933, objective tests have flourished. In fact, they are now much more widely used than projective tests and are addressed toward almost every imaginable problem and aspect of human behavior. The Minnesota Multiphasic Personality Inventory (1945) was the truly landmark event in the course of development of paper-and-pencil instruments for assessing clinical aspects of psychological functioning. "Paper-and-pencil" is often used synonymously with "objective" in relation to personality. From that time on, other measures flourished, of recent in great profusion.

Paper-and-pencil tests freed clinicians from the drudgery of test administration, and in that way they also made testing relatively inexpensive as a clinical enterprise. They also made tests readily available to psychologists not specifically trained on them, including psychologists at subdoctoral levels. Paper-and-pencil measures also seemed so easy to administer, score, and interpret. As we have noted previously, the ease of creation of new measures had very substantial effects on the field, including clinical assessment.

4.01.5.6 Disinterest in Basic Psychological Processes

Somewhere along the way in its development, clinical assessment became detached from the mainstream of psychology and, therefore, from the many developments in basic psychological theory and knowledge. The Rorschach was conceived not as a test of personality *per se* but in part as an instrument for studying perception and Rorschach referred to it as his "experiment" (Hunt, 1956). Unfortunately, the connections of the Rorschach to perception and related mental processes were lost, and clinical psychology became preoccupied not with explaining how Rorschach responses come to be made but in explaining how Rorschach responses reflect back on a narrow range of potential determinants: the personality characteristics of respondents, and primarily their pathological characteristics at that.

It is testimony to the stasis of clinical assessment that three-quarters of a century after the introduction of the Rorschach, a period of time marked by stunning (relatively) advances in understanding of such basic psychological processes as perception, cognition, learning, and motivation and by equivalent or even greater advances in understanding of the biological structures and processes that underlie human behavior, the Rorschach continues, virtually unchanged, to be the favorite instrument for clinical assessment. The Exner System, although a revision of the scoring system, in no way reflects any basic changes in our advancement of understanding of the psychological knowledge base in which the Rorschach is, or should be, embedded. Take, just for one instance, the great increase of interest in and understanding of "priming" effects in cognition; those effects would clearly be relevant to the understanding of Rorschach responses, but there is no indication at all of any awareness on the part of those who write about the Rorschach that any such effect even exists. It was known a good many years ago that Rorschach responses could be affected by the context of their administration (Sechrest, 1968), but without any notable effect on their use in assessment.

Nor do any other psychological instruments show any particular evidence of any relationship to the rest of the field of psychology. Clinical assessment could have benefited greatly from a close and sensitive connection to basic research in psychology. Such a connection might have fostered interest in clinical assessment in the development of instruments for the assessment of basic psychological processes.

Clinical psychology has—is afflicted with, we might say—an extraordinary number of different tests, instruments, procedures, and so on. It is instructive to consider the nature of all these tests; they are quite diverse. (We use the term "test" in a somewhat generic way to refer to the wide range of mechanisms by which psychologists carry out assessments.) Whether the great diversity is a curse or a blessing depends on one's point of view. We think that a useful perspective is provided by contrasting psychological measures with those typically used in medicine, although, obviously, a great many differences exist between the two enterprises. Succinctly, however, we can say that most medical tests are very narrow in their intent, and they are devised to tap basic states or processes. A screening test for tuberculosis, for example, involves subcutaneous injection of tuberculin which, in an infected person, causes an inflammation at the point of injection. The occurrence of the inflammation then leads to further narrowly focused tests. The inflammation is not tuberculosis but a sign of its potential existence. A creatinine clearance test is a test of renal function based on the rate of clearance of ingested creatinine from the blood. A creatinine clearance test can indicate abnormal renal functioning, but it is a measure of a fundamental physiological process, not a state, a problem, a disease, or anything of that sort. A physician who is faced with the task of diagnosing some disease process involving renal malfunction will use a variety of tests, not necessarily specified by a protocol (battery) to build an information base that will ultimately lead to a diagnosis.

By contrast, psychological assessment is, by and large, not based on measurement of basic psychological processes, with few exceptions. Memory is one function that is of interest to neuropsychologists, and occasionally to others, and instruments to measure memory functions do exist. Memory can be measured independently of any other functions and without regard to any specific causes of deficiencies. Reaction time is another basic psychological process. It is currently used by cognitive psychologists as a proxy for mental processing time, and since the 1970s, interest in reaction time as a marker for intelligence has grown and become an active research area.

For the most part, however, clinical assessment has not been based on tests of basic psychological functions, although the Wechsler intelligence scales might be regarded as an exception to that assertion. A very large number of psychological instruments and procedures are aimed at assessing syndromes or diagnostic conditions, whole complexes of problems. Scales for assessing attention deficit disorder (ADD), suicide probability, or premenstrual syndrome (PMS) are instances. Those instruments are the equivalent of a medical "Test for Diabetes," which does not exist. The Conners' Rating Scales (teachers) for ADD, for example, has subscales for Conduct Problem, Hyperactivity, Emotional Overindulgent, Asocial, Anxious-Passive, and Daydream-Attendance. Several of the very same problems might well be represented on other instruments for entirely different disorders. But if they were, they would involve a different set of items, perhaps with a slightly different twist, to be integrated in a different way. Psychology has no standard ways of assessing even such fundamental dispositions as "asocial."

One advantage of the medical way of doing things is that tests like creatinine clearance have been used on millions of persons, are highly standardized, have extremely well-established norms, and so on. Another set of ADD scales, the Brown, assesses "ability to activate and organize work tasks." That sounds like an important characteristic of children, so important that one might think it would be widely used and useful. Probably, however, it appears only on the Brown ADD Scales, and it is probably little understood otherwise.

Clinical assessment has also not had the benefit of careful study from the standpoint of basic psychological processes that affect the clinician and his or her use and interpretation of psychological tests. Achenbach (1985), to cite a useful perspective, discusses clinical assessment in relation to the common sources of error in human judgment. Achenbach refers to such problems as illusory correlation, inability to assess covariation, and the representativeness and availability heuristics and confirmatory bias described by Kahneman, Slovic, and Tversky (1982). Consideration of these sources of *human*, that is, general, error in judgment would be more likely if clinical assessment were more attuned to and integrated into the mainstream developments of psychology.

We do not suppose that clinical assessment should be limited to basic psychological processes; there may well be a need for syndrome-oriented or condition-oriented instruments. Without any doubt, however, clinical assessment would be on a much firmer footing if from the beginning psychologists had tried to define and measure well a set of fundamental psychological processes that could be tapped by clinicians faced with diagnostic or planning problems.

Unfortunately, measurement has never been taken seriously in psychology, and it is still lightly regarded. One powerful indicator of the casual way in which measurement problems are met in clinical assessment is the emphasis placed on brevity of measures. "... entire exam can be completed... in just 20 to 30 minutes" (for head injury), "completed in just 15–20 minutes" (childhood depression), "39 items" (to measure six factors involved in ADD) are just a few of the notations concerning tests that are brought to the attention of clinician-assessors by advertisers. It would be astonishing to think of a medical test advertised as "diagnoses brain tumors in only 15 minutes," or "complete diabetes workup in only 30 minutes." An MRI examination for a patient may take up to several hours from start to finish, and no one suggests a "short form" of one. Is it imaginable that one could get more than the crudest notion of childhood depression in 15–20 minutes?

4.01.6 MISSED SIGNALS

At various times in the development of clinical psychology, opportunities existed to guide, or even redirect, assessment activities in one way or another. Clinical psychology might very well have taken quite a different direction than it has (Sechrest, 1992). Unfortunately, in our view, a substantial number of critical "signals to the field were missed, and entailed in missing them was failure to redirect the field in what would have been highly constructive ways.

4.01.6.1 The Scientist–Practitioner Model

We do not have the space to go into the intricacies of the scientist–practitioner model of training and practice, but it appears to be an idea whose time has come and gone. Suffice it to say here that full adoption of the model would not have required every clinical practitioner to be a researcher, but it would have fostered the idea that to some extent every practitioner is responsible for the scientific integrity of his or her own practice, including the validity of assessment procedures. The scientist–practitioner model might have helped clinical psychologists to be involved in research, even if only as contributors rather than as independent investigators.

That involvement could have been of vital importance to the field. The development of psychological procedures will never be supported commercially to any appreciable extent, and if they are to be adequately developed, it will have to be with the voluntary—and enthusiastic—participation of large numbers of practitioners who will have to contribute data, be involved in the identification of problems, and so on. That participation would have been far more likely had clinical psychology stuck to its original views of itself (Sechrest, 1992).

4.01.6.2 Construct Validity

We have already discussed construct validity at some length, and we have explained our view that the idea has been trivialized, in essence abandoned. That is another lost opportunity, because the power of the original formulation by Cronbach and Meehl (1955) was great. Had their work been better understood and honestly adopted, clinical psychology would by this time almost certainly have had a set of well-understood and dependable measures and procedures. The number and variety of such measures would have been far less than exists now, and the dependability of them would have been circumscribed, but surely it would have been better to have good than simply many measures.

4.01.6.3 Assumptions Underlying Assessment Procedures

In 1952, Lindzey published a systematic analysis of assumptions underlying the use of projective techniques (Lindzey, 1952). His paper was a remarkable achievement, or would have been had anyone paid any attention to it. The Lindzey paper could have served as a model and stimulus for further formulations leading to a theory, comprehensive and integrated, of performance on clinical instruments. A brief listing of several of the assumptions must suffice to illustrate what he was up to:

> IV. The particular response alternatives emitted are determined not only by characteristic response tendencies (enduring dispositions) but also by intervening defenses and his cognitive style.
>
> XI. The subject's characteristic response tendencies are sometimes reflected indirectly or symbolically in the response alternatives selected or created in the test situation.
>
> XIII. Those responses that are elicited or produced under a variety of different stimulus conditions are particularly likely to mirror important aspects of the subject.
>
> XV. Responses that deviate from those typically made by other subjects to this situation are more likely to reveal important characteristics of the subject than modal responses which are more like those made by most other subjects.

These and other assumptions listed by Lindzey could have provided a template for systematic development of both theory and programs of research aimed at supporting the empirical base for projective—and other—testing. Assumption XI, for example, would lead rather naturally to the development of explicit theory, buttressed by empirical data, which would indicate just when responses probably should and should not be interpreted as symbolic.

Unfortunately, Lindzey's paper appears to have been only infrequently cited and to have been substantially ignored by those who were engaged in turning out all those projective tests, inventories, scales, and so on. At this point we know virtually nothing more about the performance of persons on clinical instruments than was known by Lindzey in 1952. Perhaps even less.

4.01.6.4 Antecedent Probabilities

In 1955 Meehl and Rosen published an exceptional article on antecedent probabilities and the problem of base rates. The article was, perhaps, a bit mathematical for clinical psychology, but it was not really difficult to understand, and its implications were clear. Whenever one is trying to predict (or diagnose) a characteristic that is quite unevenly distributed in a population, the difficulty in beating the accuracy of the simple base rates is formidable, sometimes awesomely so. For example, even in a population considered at high risk for suicide, only a very few persons will actually commit suicide. Therefore, unless a predictive measure is extremely precise, the attempt to identify those persons who will commit suicide will identify as suicidal a relatively large number of "false-positives," that is, if one wishes to be sure not to miss any truly suicidal people, one will include in the "predicted suicide" group a substantial number of people not so destined. That problem is a serious to severe limitation when the cost of missing a true-positive is high, but so, relatively, is the cost of having to deal with a false-positive.

More attention to the difficulties described by Meehl and Rosen (1955) would have moved psychological assessment in the direction taken by medicine, that is, the use of ROCs. Although ROCs do not make the problem go away, they keep it in the forefront of attention and require that those involved, whether researchers or clinicians, deal with it. That signal was missed in clinical psychology, and it is scarcely mentioned in the field today. Many indications exist that a large proportion of clinical psychologists are quite unaware that the problem even exists, let alone that they have an understanding of it.

4.01.6.5 Need for Integration of Information

Many trends over the years converge on the conclusion that psychology will make substantial progress only to the extent that it is able to integrate its theories and knowledge base with those developing in other fields. We can address this issue only on the basis of personal experience; we can find no evidence for our view. Our belief is that clinical assessment in psychology rarely results in a report in which information related to a subject's genetic disposition, family structure, social environment, and so on are integrated in a systematic and effective way.

For example, we have seen many reports on patients evaluated for alcoholism without any attention, let alone systematic attention, to a potential genetic basis for their difficulty. At most a report might include a note to the effect that the patient has one or more relatives with similar problems. Never was any attempt made to construct a genealogy that would include other conditions likely to exist in the families of alcoholics. The same may be said for depressed patients. It might be objected that the responsibilities of the psychologist do not extend into such realms as genetics and family and social structure, but surely that is not true if the psychologist aspires to be more than a sheer technician, for example, serving the same function as a laboratory technician who provides a number for the creatinine clearance rate and leaves it to someone else, "the doctor," to put it all together.

That integration of psychological and other information is of great importance has been implicitly known for a very long time. That knowledge has simply never penetrated training programs and clinical practice. That missed opportunity is to the detriment of the field.

4.01.6.6 Method Variance

The explicit formulation of the concept of method variance was an important development in the history of assessment, but one whose import was missed or largely ignored. The concept is quite simple: to some extent, the value obtained for the measurement of any variable depends in part on the characteristics of the method used to obtain the estimate. (A key idea is the understanding that any specific value is, in fact, an estimate.) The first explicit formulation of the idea of method variance was the seminal Campbell and Fiske paper on the "multitrait-multimethod matrix" (Campbell & Fiske, 1959). (That paper also introduced the very important concepts of "convergent" and "discriminant" validity, now widely employed but, unfortunately, not always very well understood.) There had been precursors of the idea of method variance. In fact, much of the interest in projective techniques stemmed from the idea that they would reveal aspects of personality that would not be discernible from, for example, self-report measures. The MMPI, first published in 1943 (Hathaway & McKinley), included "validity" scales that were meant to detect, and, in the case of the K-scale, even correct for, methods effects such as lying, random responding, faking, and so on. By 1960 or so, Jackson and Messick had begun to publish their work on response styles in objective tests, including the MMPI (e.g., Jackson & Messick, 1962). At about the same time, Berg (1961) was describing the "deviant response tendency," which was the hypothesis that systematic variance in test scores could be attributed to general tendencies on the part of some respondents to respond in deviant ways. Nonetheless, it was the Campbell and Fiske (1959) paper that brought the idea of method variance to the attention of the field.

Unfortunately, the cautions expressed by Campbell and Fiske, as well as by others working on response styles and other method effects, appear to have had little effect on developments in clinical assessment. For the most part, the problems raised by methods effects and response styles appear to have been pretty much ignored in the literature on clinical assessment. A search of a current electronic database in psychology turned up, for example, only one article over the past 30 years or so linking the Rorschach to any discussion of method effects (Meyer, 1996). When one considers the hundreds of articles having to do with the Rorschach that were published during that period of time, the conclusion that method effects have not got through to the attention of the clinical assessment community is unavoidable. The consequence almost surely is that clinical assessments are not being corrected, at least not in any systematic way, for method effects and response biases.

4.01.6.7 Multiple Measures

At least a partial response to the problem of method effects in assessment is the use of multiple measures, particularly measures that do not appear to share sources of probable error or bias. That recommendation was explicit in Campbell and Fiske (1959), and it was echoed and elaborated upon in 1966 (Webb et al., 1966), and again in 1981 (Webb et al., 1981). Moreover, Webb and his colleagues warned specifically against the very heavy reliance on self-report measures in psychology (and other social sciences). That warning, too, appears to have made very little difference in practice. Examination of catalogs of instruments meant to be used in clinical assessment will show that a very large proportion of them depend upon self-reports of individual subjects about their own dispositions, and measures that do not rely directly on self-reports nonetheless do nearly all rely solely on the verbal responses of subjects. Aside from rating scales to be used with parents, teachers, or other observers of behavior, characteristics of interest such as personality and psychopathology almost never require anything of a subject other than a verbal report. By contrast, ability tests almost always require subjects to do something, solve a problem, complete a task, or whatever. Wallace (1966) suggested that it might be useful to think of traits as abilities, and following that lead might very well have expanded the views of those interested in furthering clinical assessment.

4.01.7 THE ORIGINS OF CLINICAL ASSESSMENT

The earliest interest in clinical assessment was probably that used for the classification of the "insane" and mentally retarded in the early 1800s. Because there was growing interest in understanding and implementing the humane treatment of these individuals, it was first necessary to distinguish between the two types of problems. Esquirol (1838), a French physician, published a two-volume document outlining a continuum of retardation based primarily upon language (Anastasi, 1988).

Assessment in one form or another has been part of clinical psychology from its beginnings. The establishment of Wundt's psychological laboratory at Leipzig in 1879 is considered by many to represent the birth of psychology. Wundt and the early experimental psychologists were interested in uniformity rather than assessment of the individual. In the Leipzig lab, experiments investigated psychological processes affected by perception, in which Wundt considered individual differences to be error. Accordingly, he believed that since sensitivity to stimuli differs, using a standard stimulus would compensate and thus eliminate individual differences (Wundt, Creighton, & Titchener, 1894/1896).

4.01.7.1 The Tradition of Assessment in Psychology

Sir Francis Galton's efforts in intelligence and heritability pioneered both the formal testing movement and field testing of ideas. Through his Anthropometric Laboratory at the International Exposition in 1884, and later at the South Kensington Museum in London, Galton gathered a large database on individual differences in vision, hearing, reaction time, other sensorimotor functions, and physical characteristics. It is interesting to note that Galton's proposition that sensory discrimination is indicative of intelligence continues to be promoted and investigated (e.g., Jensen, 1992). Galton also used questionnaire, rating scale, and free association techniques to gather data.

James McKeen Cattell, the first American student of Wundt, is credited with initiating the individual differences movement. Cattell, an important figure in American psychology, (Fourth president of the American Psychological Association and the first psychologist elected to the National Academy of Science) became interested in whether individual differences in reaction time might shed light on consciousness and, despite Wundt's opposition, completed his dissertation on the topic. He wondered if, for example, some individuals might be observed to have fast reaction time across situations and supposed that the differences may have been lost in the averaging techniques used by Wundt and other experimental psychologists (Wiggins, 1973). Cattell later became interested in the work of Galton and extended his work by applying reaction time and other physiological processes as measures of intelligence. Cattell is credited with the first published reference to a mental test in the psychological literature (Cattell, 1890).

Cattell remained influenced by Wundt in his emphasis on psychophysical processes. Although physiological functions could be easily and accurately measured, attempts to relate them to other criteria, however, such as teacher ratings of intelligence and grades, yielded poor results (Anastasi, 1988).

Alfred Binet conducted extensive and varied research on the measurement of intelligence. His many approaches included measurements of cranial, facial, and hand form, handwriting analysis, and inkblot tests. Binet is best known for his work in the development of intelligence scales for children. The earliest form of the scale, the Binet–Simon, was developed following Binet's appointment to a governmental commission to study the education of retarded children (Binet & Simon, 1905). The scale assessed a range of abilities with emphasis on comprehension, reasoning, and judgment. Sensorimotor and perceptual abilities were relatively less prominent, as Binet considered the broader process, for example, comprehension, to be central to intelligence. The Binet–Simon scale consisted of 30 problems arranged in order of difficulty. These problems were normed using 50 3–11-year-old normal children and a few retarded children and adults.

A second iteration, the 1908 scale, was developed. The 1908 scale was somewhat longer and normed on approximately 300 3–13-year-old normal children. Performance was grouped by age according to the level at which 80–90% of the normal children passed, giving rise to the term "mental age."

The Binet–Simon has been revised, translated, and adapted in numerous languages. Perhaps the most well-known revision was directed by Lewis Terman (1916) at Stanford University and this test is what is known as the Stanford–Binet. The Stanford–Binet was the origin of the intelligence quotient (IQ), the ratio between chronological and mental ages.

4.01.7.1.1 Witmer

Lightner Witmer, who studied with both Cattell and Wundt, established the first American psychological clinic at the University of Pennsylvania in 1896. This event is considered by many as the beginning of clinical psychology (Garfield, 1965; McReynolds, 1987, 1996). Witmer's approach to assessment was focused on determining the causes of children's problems and then to make recommendations for treatment. Diagnoses, *per se*, were not considered important, however, Witmer did make use of the Stanford–Binet and other formal assessment tools. McReynolds (1996) noted that Witmer strongly emphasized both direct observation and extensive background data as especially important for assessment.

Although Witmer characterized his work as practical, he remained committed to a scientific basis for psychology (McReynolds, 1996). It seems reasonable to conclude that Witmer was interested in assessment for bounded inference and prediction. That is, he wanted information as it might relate to specific problems for the express purpose of treating those problems (Witmer, 1996/1907).

4.01.7.1.2 Army Alpha

Robert M. Yerkes initiated and administered a program to test 1.75 million army recruits during World War I. This program, which Yerkes developed in conjunction with Terman and H. H. Goddard, administered the Army Alpha written mental test to recruits. Illiterate recruits and those failing the Alpha were given a picture-based test called the Army Beta.

Yerkes hoped that the army could be "engineered" by classifying the intelligence and capabilities of all recruits. To that end, recruits were graded from A through E and Yerkes recommended that they be assigned rank and tasks according to their tested ability. Although the army did not use the results uniformly, in many instances recruits for officer training were required to have an A or B grade on the Alpha. The tests results were later used in controversial ways by both Yerkes and E. G. Boring to assess average American intelligence levels (see Yerkes, 1921, 1941). Despite whatever controversy may have arisen over the years, the army continues to use testing to assess aptitudes (Jensen, 1985).

4.01.8 THE RORSCHACH INKBLOT TECHNIQUE AND CLINICAL PSYCHOLOGY

The history of the Rorschach Inkblot Technique is in many ways a reflection of the history of clinical psychology in America. Clinical psychology continues to struggle with competing world views focusing on the nature of reality, the mind, and human behavior. In clinical psychology the debate about how to view the mind and behavior is usually expressed, broadly speaking, as poles of a dimension anchored by only observable behavior at one end, the influences of conscious mental processes (i.e., cognition) more in the center, and unconscious mental processes anchoring the other end. The relative importance of observable behavior and unconscious mental processes alternate with the intellectual fashions of the times.

The role of the clinical psychologist as scientist, diagnostician, and therapist continue to change, with a growing fracture between the scientifically and the clinically oriented. A central focus of debate has to do with molar vs. molecular views of personality and the ways in which personality is assessed. Conflict over the use of the Rorschach is characteristic of the debate and perturbing in light of long-standing doubts about the psychometric adequacy and the clinical usefulness of the instrument. An additional factor in the ongoing conflict in psychology seems to be that in psychology, alas, like old soldiers, theories never die. Even if refuted, they are not replaced, they only very gradually fade away (Meehl, cited by Lykken, 1991).

4.01.8.1 The Social and Philosophical Context for the Appearance of the Rorschach

Although the Rorschach was first introduced in the United States in 1925, it was during the 1940s and 1950s that the Rorschach rose to prominence in clinical psychology. The prevailing theoretical views in American academic psychology during the early years of the Rorschach were Gestalt and behaviorism. In many ways the interest and devotion of Rorschach proponents to the technique seems to have been a reaction against what they saw as reductionist and positivistic approaches to personality assessment on the part of behaviorists and often atheoretical psychometricians. Additionally, behaviorists focused on environmental determinants of behavior at the same time that psychoanalytic theory, in spite of its rejection in much of academia, was beginning to flourish in clinical psychology. Moreover, by the late 1940s, many psychologists were interested in reviving the notion of the self, which had been rejected by behaviorism and psychoanalysis (Reisman, 1991).

Proponents of the Rorschach believed that underlying dimensions of "true" personality could be elicited only by indirect, projective methods; defense mechanisms, repression, and perhaps other unconscious processes prevented an individual from having access to critical information about him- or herself. Direct assessment of personality was narrow and incomplete, but the ambiguity of the inkblot stimulus material would elicit true responses. Because during the 1940s and 1950s testing was virtually the only applied professional activity performed by clinical psychologists (Millon, 1984), it is not surprising that the Rorschach would generate a great deal of interest and activity. What is surprising is that a test criticized even then and continuously until now as being too subjective in administration, scoring, and interpretation, of questionable reliability, and of dubious validity, would be continually used for 70 years.

Rorschach proponents did claim to view the technique as scientific, and there were attempts to establish norms and to approach the Rorschach scientifically, but we view the Rorschach ultimately as what Richard Feynman (1986) refers to as "Cargo Cult Science:"

> In the South Seas there is a cargo cult of people. During the war, they saw airplanes land with lots of good materials, and they want the same thing to happen now. So they've arranged to make things like runways, to put fires along the sides of the runways, to make a wooden hut for a man to sit on, with two wooden pieces on his head like headphones and bars of bamboo sticking out like antennas—he's the controller—and they wait for the airplanes to land. They're doing everything right. The form is perfect. It looks just the way it looked before. But it doesn't work. No airplanes land. So I call these things cargo cult science, because they follow all the apparent percepts and forms of scientific investigation, but they're missing something essential, because the planes don't land.

The Rorschach technique is missing something essential. Although as we stated earlier, people almost certainly "project" aspects of their personality on to ambiguous stimuli, use of the Rorschach has failed to demonstrate convincing evidence of validity in decades of attempts to find it. The planes still don't land.

4.01.8.2 The Birth of the Rorschach

Whether and how to use the Rorschach has been a source of controversy since its introduction. Perhaps much of the controversy and dissent about scoring and interpretation of responses to the inkblots among advocates of the technique were a result of its founder's death a few months after the publication of his initial monograph detailing 10 years of studies with inkblots, leaving a nascent method open to various interpretations. The original notions of using the technique tentatively and experimentally began fading with its founder's death, being replaced by an overriding concern for clinical uses.

Herman Rorschach, the son of a Swiss art teacher, began experimenting with inkblots in various psychopathic hospitals in 1911, the year after completing his medical training (Klopfer & Kelley, 1942). The Rorschach method was introduced in the United States in 1925 by David Levy, a psychologist and psychiatrist (Hertz, 1986; Klopfer & Kelley, 1942), who had been a student of Emil Oberholzer, Rorschach's closest medical colleague and who continued Rorschach's work after his death. Levy taught the technique to Samuel Beck, who wrote his dissertation on the technique and published the first manual on the Rorschach in 1937 (Exner, 1969; Hertz, 1986; Klopfer & Kelley, 1942). Beck and Bruno Klopfer were probably the most influential individuals in terms of widening the use of the technique, as well as in fomenting debate about how to score and interpret Rorschach responses. Beck was more behavioral and experimental in his approach and strongly advocated establishing norms and testing the validity of responses. Klopfer, a German who had studied with Jung in Switzerland after fleeing Hitler and before coming to the United States, was much more inferential in his interpretation and scoring. Rorschach himself was considerably more tentative about his findings than subsequent proponents of the technique were, or than they seem to be to this day.

It is likely that dissemination of the Rorschach was actually helped by the controversy and dissent within the ranks of Rorschach adherents, as well as by the fight against perceived rigid standards of psychometrics and nomothetic personality theories. The internal debate among adherents of various systems of scoring and interpretation seemed to foster beliefs that the findings finally proving them right were just around the corner. This belief of imminent justification seems to characterize even present day Rorschach proponents.

Another faction of Rorschach adherents with more interest in applying the Rorschach to clinical cases took the view that assessment and prediction based on clinical judgment and acumen are inherently superior to psychometric and statistical assessment and prediction. During the 1950s and 1960s, the emphasis shifted from scores and scoring systems to the utilization of clinical acumen and sensitivity, and attempts to understand subtle aspects of the entire testing situation (Sarason, 1954). As the role of the clinical psychologist expanded into more applied clinical activity, practitioners' attention to the experimental scientific roots of the discipline began fading. With this movement further from a scientific basis for theories and techniques, the theories promoted by academic psychologists were considered mechanistic by most practitioners. As a result, the academics' criticisms about projectives such as the Rorschach were increasingly viewed as invalid (or, perhaps worse, as irrelevant). In our reading of the literature, it appears that even those Rorschach supporters who believe science is important cling to the "Cargo Cult Science" of ratios and scoring systems lacking in empirical support but with the expectation of redemption almost momentarily expected.

This shift in the 1950s and 1960s to a focus on clinical skills was in the context of the emergence of psychotherapy as a primary professional activity for psychologists. Erikson's theory of psychosocial development was embraced, psychodynamic theory in various forms (Adler, Rank, Horney, Sullivan) was popular with clinicians, and Rogerian humanistic psychology emerged along with behavior modification and systematic desensitization (Reisman, 1991). In psychiatry there were rapid advances in psychotropic medications. These changes in the field seemed to steel the resolve of clinicians who believed that human psychology could not be reduced to biology, classification, and statistical formulas. Despite the lack of any demonstrated validity of the Rorschach from research studies, clinicians focused on the feedback they received, or thought they received, from clients, and believed the Rorschach helped them to better understand their clients. At about the same time as these developments, Paul Meehl (1954) published an analysis of the general problem of clinical vs. statistical prediction.

4.01.8.3 Clinical vs. Statistical Prediction

The central issue in relation to comparisons of clinical and statistical (actuarial) prediction is simple: when there is a database of cases with known outcome, can a skilled clinician use his or her judgment to combine the relevant information about a client into the correct formulations (predictions) as well as or better than a statistical formula that uses the same information? The answer, based on numerous studies in which clinicians had as much or more information as was entered into the statistical prediction, is *no*. Clinicians occasionally equal but never exceed statistical predictions of behavior, diagnoses, psychotherapy outcome, and like events of interest. The preponderance of evidence favors statistical prediction.

Even when statistical models are based upon the information used by clinicians, the models outperform the clinicians on whom they are based (Dawes et al., 1989). Exceptions do occur in circumstances of events that reverse the actuarial formula or of judgments mediated by theories that are, therefore, difficult or even impossible to duplicate statistically (Dawes et al., 1989). When such information is available to clinicians, and those circumstances may be infrequent, they are likely to outperform statistical models. Meehl (1954) referred to these rare events as the broken leg phenomenon. That name was derived from an illustration in which a statistical formula is highly successful in predicting an individual's weekly attendance at a movie, but should be discarded upon discovering that the subject is in a cast with a fractured femur.

One reason for the superiority of statistical prediction is that clinicians tend to think that too many cases are exceptions to ordinary rules and, even in the case of rare events, they ultimately perform better when they rely strictly on statistical conclusions (Goldberg, 1968). The human mind is a poor computer and does not do a good job at quantifying and weighting observations, the very things that regression equations were invented for (Goldberg, 1991). We do not mean to suggest that statistical formulas can be used to perform psychotherapy, or that the predictions could be made without first gathering the relevant observations from clinicians. We also bear in mind that a great many clinical decisions are made in circumstances in which there is no known outcome.

The debate about clinical vs. statistical prediction has been characterized by *ad hominem* attacks, and Meehl (1954) started his book, *Clinical versus statistical prediction,* with lists of invective from both sides of the argument. Briefly, opponents of statistical prediction have suggested that the approach is atomistic, inhuman, arbitrary, and oversimplified, while its proponents suggest that it is objective, reliable, rigorous, and scientific. Conversely, negative appraisals of clinical prediction suggest the method is sloppy, muddleheaded, unscientific, and vague, while its proponents suggest that the method is dynamic, sensitive, meaningful, and holistic (Meehl, 1954).

The case for the use of psychodiagnostic tests such as the Rorschach and the validity of clinical observation of relationships between thoughts, behavior, and personality characteristics becomes questionable considering the findings about the questionable validity of clinical judgments. Further, it has been known for a long while that statements from clinical assessments and psychological reports are often of universal applicability (Forer, 1949). When previously prepared statements representative of those in psychological evaluations are presented to a variety of individuals, the individuals enthusiastically agree that the statements uniquely apply to them. Therefore, it seems that the very evidence often used by clinicians, that their clients believe assessments to be accurate and that they are helped by assessment and treatment, affords no reassurance. Much information provided by typical psychodiagnostic feedback is general and applies to almost anyone. The associations with various personality characteristics, signs, and indicators may be more related to what any astute observer has learned to associate with them through observation, folklore, and literature, that is, "illusory correlations" (Chapman & Chapman, 1967, 1969; Reisman, 1991). It is likely that such illusory correlations are involved in accounts of individuals known as "Rorschach Savants," who are purported anecdotally to see phenomenal amounts of information in Rorschach responses.

It is astonishing that the Rorschach continues to be not only very popular, but in many states is required as part of forensic psychological assessment in child custody disputes (Robyn Dawes, personal communication). Reisman (1991) suggests that the failure of clinical psychologists to modify their behavior no matter how much aversive stimulation is applied is less a refutation of Skinner's theory than evidence of a great capacity to distort information. Many clinicians and even some researchers continue to believe in the validity of the Rorschach (and other projective tests) in spite of overwhelming evidence to the contrary and almost universal agreement among the scientific community that the central assumption on which the Rorschach is based is faulty.

> The entire Rorschach is based on a fallacious assumption, namely that indirect (projective) methods are more valid than direct (self-rating or questionnaire) methods because people are so repressed that they cannot describe their real emotions and impulses. A large body of literature indicates the fallacy of this assumption. Even within self-report items, more content obvious items prove to be more valid than subtle ones. Why give an hour test, with another hour to score, to get a crude estimate of anxiety or depression which is usually less reliable and valid than a short true–false scale which takes a few minutes and where there is no unreliability of scoring? I have compared direct and indirect (Rorschach and TAT) measures of dependency, anxiety, depression, and hostility using peer ratings as criteria. The most indirect methods have zero validity, the most direct methods have low to moderate validity, and methods which are intermediate in directness (e.g., sentence completion) are intermediate in validity. A great deal of effort was expended in scoring content from TAT and Rorschach and consensus agreement was obtained where disagreement in scoring occurred. All this was to no avail because the two projectives did not correlate with each other let alone with the criteria or any of the direct methods. (Marvin Zuckerman, SSCPNET, April 22, 1996)

Although yet another scoring system for the Rorschach has been used and researched for the past 20 years (Exner, 1974, 1993) with a greater emphasis on standardization of scoring and interpretation, it has yielded no apparent improvement on the predictive or incremental validity of the technique. Criticisms of the research are nearly identical to those expressed in the 1940s and 1950s. Disturbingly, in spite of overwhelming evidence of their invalidity, clinicians tend to continue to rely on their impressions and interpretations of the content of Rorschach responses (Reisman, 1991). It is not precisely fair to say that the Rorschach is unrelated to anything, but its validity is so limited as to leave virtually no real utility for its use. Most problematic, it is inferior to and more time-consuming than instruments with better reliability and validity and the Rorschach appears to have zero incremental validity (Sechrest, 1963).

4.01.8.4 Old Tests Never Die, They Just Fade Away

The continued drift of psychology away from its scientific roots does not appear to be slowing. This drift seems additionally fueled by economic and employment concerns and continued training of too many practitioners. The current conflict is unlikely to slow as managed health care and cutbacks in federal funding lessen job opportunities, and the future of psychology is uncertain. Clinical psychology, even in the halcyon days of the scientist–practitioner model, was never resolute in its commitment to science. For example, students coming into the field were generally not required to have any particular prior training in science, or its principal handmaiden, mathematics, and they needed only to declare a personal fealty to the idea of research. That situation has almost certainly become much worse over the past two or three decades of drift toward practitioner–scientist, then practitioner–scholar, and then frankly practitioner programs. The net result is that clinical psychology has a huge number of practitioners who are not only ill-equipped to handle the demands of evaluating the scientific basis for practice, but they are ill-disposed even to doing so. Economic pressures and their own incapacities make scientific evidence, which is at best likely to be disappointing, a threat. Anecdotes, "clinical experience," and so on are far more reassuring and, hence, attractive. Better to believe in an unproven instrument or procedure than to be deprived of any basis for pride and survival.

Lykken (1991) noted that present knowledge in psychology is very broad but very shallow. Most recently trained clinical psychologists probably have little acquaintance with the philosophy of science and not much knowledge of the clinical vs. statistical prediction literature; certainly they have inadequate training in measurement, statistics, and probability. This ignorance of the roots of psychological theory and scientific psychology contributes to the continued use of a completely unjustifiable procedure such as the Rorschach. It is difficult to refute disproven techniques and theories when a class of the profession basis its identity and livelihood on them. The problem of theories fading away and reviving as suggested by Meehl's "old soldiers" simile is not restricted to clinical psychology; psychology as a whole operates in this way.

In other sciences, each generation builds on the foundations of the discipline's previous scientists. Psychology seems to view its predecessors as "intrepid explorers who came back empty-handed" (Lykken, 1991). To be fair, establishing a psychological science is extremely difficult because it is difficult to operationalize psychological constructs and because there is notable measurement error. The profession and practice of clinical psychology would be helped immensely, however, if we could better educate graduate students in philosophy of science, measurement, and statistics, in addition to psychological theory.

The Rorschach did not come into prominence originally because of evidence for its superiority over existing measures, for example, questionnaires and checklists. It was adopted eagerly, we think, more because of discontent with the obvious inadequacies of existing alternatives. We suspect that whatever its own inadequacies, the Rorschach will not die but will only fade away when some alternative instrument or procedure becomes available and seems potentially to be a better one.

4.01.9 OTHER MEASURES USED IN CLINICAL PSYCHOLOGY

The list of measures that have been used in clinical psychology is very long, and many appear simply to have faded away. For example, two projective tests that once had a spate of popularity are the Blacky Test and the Make-a-Picture Story Test (MAPS) (Shneidman, 1986). The Blacky Test seems to have disappeared altogether, and the MAPS is rarely encountered in the literature. Neither was ever demonstrated to be less reliable or less valid than other tests; each simply appears to have faded away, the Blacky probably because its version of psychoanalytic theory has also faded somewhat and the MAPS because it was cumbersome and slow to administer. There is not much point in recounting the histories of the many now deservedly (even if not uniquely deserved) forgotten tests.

4.01.9.1 The Thematic Apperception Test

Morgan and Murray (1935) introduced the Thematic Apperception Test (TAT) based on what they termed the "well-recognized fact" that when presented with ambiguous stimuli people reveal their own personality. The TAT consists of a series of pictures of ambiguous social situations in which the examinee describes the social situation as they see it. The TAT was originally designed to be interpreted in light of psychoanalytic theory, the theory driving its design. There were subsequently a variety of scoring systems from different perspectives, although none has improved on the recurrent problem of inconsistency in use from clinician to clinician.

The TAT, as one might imagine, can be scored more or less reliably, depending on the nature of the variable involved and the adequacy of its definition. The major problem is what the scores may be related to and how they may be interpreted. Over the many years of its existence, TAT scores have been related to many different phenomena, sometimes with moderate success. The literature would show that achievement has been extensively studied by way of the TAT (see Keiser & Prather, 1990) as have other needs or motives. Although the research is reasonably consistent in showing some evidence for validity of some TAT scores and the instrument has proven to be of some value in research, the evidence was never strong enough to justify use of the TAT for individual decision-making in clinical settings. The TAT, like most other clinical measures, can at best be considered enlightening.

4.01.9.2 Sentence Completion Tests

Another variety of quasiprojective instruments is the sentence completion test, which consists of a stem, for example, "When I was a child," that the respondent is supposed to make into a complete sentence by writing down his or her own thoughts. The sentence completion test, of which the Rotter Incomplete Sentences Blank (Rotter & Rafferty, 1950) is the best known version, probably evolved from word association tests, which go back to Galton, Cattell, and Kraepelin in the latter part of the nineteenth century (Anastasi, 1988). The Rotter ISB was considered to be a measure of psychological conflict and, therefore, adjustment, and like so many other measures, under the right circumstances, it could be scored in a reasonably dependable way and could result in "significant" validity coefficients. That is to say, the ISB could be shown variously and not invariably to be correlated around 0.30 with criteria thought by someone to be of interest. Those correlations might be useful for some research purposes, but they were not grounds for much confidence in clinical settings. They may, however, in the minds of many clinicians have inspired more confidence and, therefore, more use than was warranted.

4.01.9.3 Objective Testing

The term "objective test" usually refers to a self-report measure that presents a stimulus item to a respondent and that requires a constrained response such as "True/False," "Agree/Disagree," and so forth. There are many, many objective tests, but the dominant one is, and virtually always has been, the MMPI (Hathaway & McKinley, 1943). We have already discussed various aspects of the MMPI under other topics, but it is worth noting here that the durability of the MMPI has been impressive. Its clinical utility has not. It yields profiles that seem impressive, and it certainly can, in general, serve as a screening instrument for psychopathology: people who get really high scores on one or more of the MMPI scales probably have *something* awry in their lives. No relationships have ever been consistently demonstrated between the MMPI and functional capacities or incapacities that would justify clinical decisions other than to seek further information about the client or patient.

The MMPI more than other available instruments has been automated, to the extent of producing computer-based interpretations of test profiles. An unfortunate limitation of computer-based interpretations is that, because of their proprietary nature, the algorithms underlying them are not available. Consequently, one cannot know which interpretations are based on empirical evidence and which, perhaps, on clinical lore, let alone how good the evidence might be. Such interpretations must be accepted on faith. When the MMPI is used in a fully automatic mode, it is questionable whether it even should be considered a clinical assessment.

4.01.9.4 The Clinician as a Clinical Instrument

Clinical psychology has never been completely clear about whether it wishes to distinguish between the test—a tool—and the test-in-the-hands-of-a-user. The perspective of standardized testing implies that the test is a tool that, in the hands of any properly trained user, should produce the same results for any given examinee. Many clinical instruments, however, cannot be considered to be so tightly standardized, and it is to be expected that results might differ, perhaps even substantially, from one examiner to another, even for the same examinee. Within reason, at least, an examinee's performance on a vocabulary test or a trail-making test should be little affected by the characteristics of the examiner, nor should the scoring and interpretation of the performance. By contrast, an examinee's responses might be affected to a considerable degree by the characteristics of an examiner administering a Rorschach or a TAT, let alone the interpretation of those responses.

The field of clinical psychology abounds in tales of diagnostic acumen of marvelous proportions manifested by legendary clinicians able to use the Rorschach, an MMPI profile, or some other instrument as a stimulus. Unfortunately, no such tales have advanced beyond the bounds of anecdote, and none of these legendary clinicians appears to have been able to pass along his or her acumen to a group of students—let alone passing it along across several generations. Consequently, if clinicians are to be part of the clinical assessment equation, then it seems inevitable that individual clinicians will have to be validated individually, that is, individual clinicians will have to be shown to be reliable and valid instruments. That will not further progress in the field.

4.01.9.5 Structured Interviews

A fairly recent development in clinical assessment is the structured interview schedule. These schedules are intended to produce a diagnostic judgment related to the *DSM* (American Psychiatric Association, 1994), a narrow, bounded purpose. There are several such interview schedules currently available, but we will discuss the Structured Clinical Interview for *DSM-IV* (SCID) as an example and because it is probably the one most widely used.

As noted earlier, most psychological assessment appears to be done for purposes of enlightenment rather than for decision-making. Nevertheless, diagnoses are often required for reimbursement, medication referrals, custody evaluations, and forensic assessments. The SCID (Spitzer, Gibbon, & Williams, 1997) appears to be used quite infrequently in other than research settings, for example, it is not mentioned on any list of instruments used by clinicians. That neglect is interesting in view of the attention that was paid to the development of the SCID and its established dependability.

Use of the SCID in clinical practice would probably contribute to improved assessment (and presumably to more appropriate treatment), whether for specific *DSM* diagnostic purposes or simply for gathering pertinent information. The SCID was designed to capitalize on clinical skills and to be more "clinician-friendly" than other structured interviews (Spitzer, Williams, Gibbon, & First, 1992). The SCID is meant to be used by precisely those people who can already conduct an interview, and although the SCID is somewhat time-consuming, but probably less so than, say, the Rorschach, psychologists interview all patients, and for most clinicians to do so in a structured manner would not be a significant departure. That is, the time would be spent interviewing the patient and the SCID would not add much if anything in terms of time or cost to standard practice. The SCID demonstrates good reliability (test–retest and inter-rater) for most disorders, with kappa coefficients averaging 0.60–0.80 or greater (Segal, Hersen, & Van Hasselt, 1995; Williams, Gibbon, First, & Spitzer, 1992). Agreement between diagnoses obtained by SCID and by traditional clinical interviews is poor to moderate with average kappa coefficients of 0.25 (Steiner, Tebes, Sledge, & Walker, 1995), suggesting strongly that reliance on unstructured clinical interviews is unwise.

Perhaps the SCID is not used because it takes some training and practice to become proficient in its use. That requirement is certainly different from the typical assessment instruments advertised in psychological publications, which boast their quick and easy use and say nothing about their reliability and validity. It may also be that beliefs about the superiority of clinical judgment over other more structured practices, for example, the use of projective tests, contributes strongly as well. Whatever the reasons for lack of clinical use of the SCID, and we suspect that it is both training time and beliefs about clinical skill, it is an unfortunate omission from assessment practice.

4.01.10 CONCLUSIONS

Progress in psychological assessment, at least for clinical applications, has been disappointing over the century since the field started. Conceptual and theoretical developments have been minimal, although we might except some observational methods used primarily in behavioral work and some research settings. The field continues to move away from its scientific roots in psychology, and clinical assessment has no other base on which to build any conceptual structure. Moreover, clinical assessment has never been more than minimally guided by psychometric theory and analysis, for example, scarcely beyond superficial concern with "reliability" of measures, and graduate education and training in research methods and measurement is at an ebb and is maybe still decreasing. Overall, clinical assessment as an enterprise seems to be cut adrift from any important sources of rigor, and almost anything goes. Perhaps it is fortunate, then, that despite the frequent insistence on assessment as a cornerstone of the practice of clinical psychology, there is much less evidence for its importance and prevalence than would be expected.

4.01.11 REFERENCES

Achenbach, T.M. (1985). *Assessment and taxonomy of child and adolescent psychopathology*. Beverly Hills, CA: Sage.

Achenbach, T.M., & Edelbrock, C. S. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, VT: Department of Psychiatry, University of Vermont.

Achenbach, T. M., & Edelbrock, C. S. (1986). *Manual for the Teachers Report Form and Teacher Version of the Child Behavior Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M., & Edelbrock, C. S. (1987). *Manual for the Youth Self-Report Form and Youth Version of the Child Behavior Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Aiken, L., West, S. G., Sechrest, L., & Reno, R. (1990). Graduate training in statistics, methodology, and measurement in psychology: A survey of Ph.D. programs in North America. *American Psychologist, 45,* 721–734.

American Psychiatric Association (1994). *Diagnostic and statistical manual for mental disorders* (4th ed.). Washington, DC: Author.

American Psychological Association (1985). *Standards for educational and psychological testing*. Washington, DC: Author.

APA Practice Directorate (1996). Practitioner survey results offer comprehensive view of psychological practice. *Practitioner Update, 4*(2).

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.

Atkins, M. S., Pelham, W. E., & White, K. J. (1990). Hyperactivity and attention deficit disorders. In M. Hersen & V. B. Van Hasselt (Eds.), *Psychological aspects of developmental and physical disabilities: A casebook*. Newbury Park, CA: Sage.

Berg, I. A. (1961). *Measuring deviant behavior by means of deviant response sets*. New York: Harpers.

Bergen, A. E., & Garfield, S. L. (1994). *Handbook of psychotherapy and behavior change*. New York: Wiley.

Bernreuter, R. G. (1933). Validity of the personality inventory. *Personality Journal, 11,* 383–386.

Binet, A., & Simon, T. H. (1905). Methodes nouvelles pour le diagnostic du niveau intellectuel des anormaux. *Annee Psychologique, 11,* 191–244.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by multitrait–multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Cattell, J. M. (1890). Mental tests and measurements. *Mind, 15,* 373–380.

Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology, 72,* 193–204.

Chapman, L. J., & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology, 74,* 271–280.

Costello, A., Edelbrock, C. S., Kalas, R., Kessler, M., & Klaric, S. A. (1982). *Diagnostic Interview Schedule for Children (DISC)*. Bethesda, MD: National Institute for Mental Health.

Craik, K. H. (1986). Personality research methods: An historical perspective. *Journal of Personality, 54*(1), 18–51.

Cronbach, L. J. (1960). *Essentials of psychological testing* (2nd ed.). New York: Harper and Row.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements*. New York: Wiley.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302.

Cronbach, L. J., Rajaratnam, N., & Gleser, G. C. (1963). Theory of generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology, 16,* 137–163.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243,* 1668–1674.

Epstein, S. (1983). Aggregation and beyond: Some basic issues in the prediction of behavior. *Journal of Personality, 51,* 360–392.

Esquirol, J. E. D. (1838). *Des maladies mentales considerees sous les rapports medical, hygienique, et medico-legal* (2 Vols.). Paris: Bailliere.

Exner, J. E. (1969). *The Rorschach systems*. New York: Grune & Stratton.

Exner, J. E. (1974). *The Rorschach systems*. New York: Grune & Stratton.

Exner, J. E. (1986). *The Rorschach: A comprehensive system*. New York: Wiley.

Exner, J. E. (1993). *The Rorschach: A comprehensive system*. New York: Wiley.

Feynman, R. (1986). *Surely you're joking, Mr. Feynman!* New York: Bantam Books.

Fiske, D. W. (1971). *Measuring the concepts of personality*. Chicago: Aldine Press.

Forer, B. (1949). The fallacy of personal validation: A classroom demonstration of gullibility. *Journal of Abnormal and Social Psychology, 44,* 118–123.

Frank, L. K. (1939). Projective methods for the study of personality. *Journal of Psychology, 8,* 389–413.

Gigerenzer, G. (1993). The superego, the ego, and the id in statistical reasoning. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 311–338). Hillsdale, NJ: Erlbaum.

Goldberg, L. R. (1968). Simple models or simple processes? Some research on clinical judgments. *American Psychologist, 23,* 483–496.

Goldberg, L. R. (1991). Human mind versus regression equation: Five contrasts. In D. Cicchetti & W. M. Grove (Eds.), *Thinking clearly about psychology: Vol. 1. Matters of public interest: essays in honor of Paul E. Meehl* (pp. 173–184). Minneapolis, MN: University of Minnesota Press.

Goyette, C. H., Conners, C. K., & Ulrich, R. E. (1978). Normative data on the Conner's parent and teacher rating scales. *Journal of Abnormal Child Psychology, 6*(2), 221–236.

Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, & Law, 2*(2), 293–323.

Hathaway, S. R., & McKinley, M. N. (1943). *The Minnesota Multiphasic Personality Inventory* (Rev. ed.). Minneapolis, MN: University of Minnesota Press.

Hertz, M. R. (1986). Rorschachbound: A 50-year memoir. *Journal of Personality Assessment, 50*(3), 396–416.

Hoza, B., Vallano, G., & Pelham, W. E. (1995). Attention-deficit/hyperactivity disorder. In R. T. Ammerman & M. Hersen (Eds.), *Handbook of child behavior therapy in psychiatric setting*. New York: Wiley.

Hunt, W. C. (1956). *The clinical psychologist*. Springfield, IL: Thomas.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.

Jackson, D. N., & Messick, S. (1962). Response styles on the MMPI: Comparison of clinical and normal samples. *Journal of Abnormal and Social Psychology, 65,* 285–299.

Jensen, A. R. (1985). Description & utility of Armed Services Vocational Aptitude Battery-14. *Measurement & Evaluation in Counseling & Development, 18*(1), 32–37.

Jensen, A. R. (1992). The importance of intraindividual variation in reaction time. *Personality & Individual Differences, 13*(8), 869–881.

Kaufman, A. S., & Kaufman, N. L. (1985). *Kaufman Test of Educational Achievement (K-TEA)*. Circle Pines, MN: American Guidance Service.

Kagan, J. (1959). The stability of TAT fantasy and stimulus ambiguity. *Journal of Consulting Psychology, 23,* 266–271.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.) (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.

Keiser, R. E., & Prather, E. N. (1990). What is the TAT? A review of ten years of research. *Journal of Personality Assessment, 55* (3–4), 800–803.

Klopfer, B., & Kelley, D. M. (1942). *The Rorschach technique*. Yonkers-on-Hudson, NY: World Book Company.

Kraemer, H. C. (1992). *Evaluating medical tests: Objective and quantitative guidelines*. Newbury Park, CA: Sage.

Levy, L. H. (1963). *Psychological interpretation*. New York: Holt, Rinehart, and Winston.

Lindzey, G. (1952). Thematic Apperception Test: Interpretive assumptions and related empirical evidence. *Psychological Bulletin*.

Lord, F. M. (1952). A theory of test scores. *Psychometric Monographs,* No. 7.

Lykken, D. T. (1991). What's wrong with psychology anyway? In D. Cicchetti & W. M. Grove (Eds.), *Thinking clearly about psychology* (pp. 3–39). Minneapolis, MN: University of Minnesota Press.

Maloney, M. P., & Ward, M. P. (1976). *Psychological assessment: A conceptual approach*. New York: Oxford University Press.

McClure, D. G., & Gordon, M. (1984). Performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. *Journal of Abnormal Child Psychology, 12*(4), 561–571.

McCraken, B. A., & McCallum, S. R. (1993). *Wechsler Intelligence Scale for Children* (3rd ed.). Brandon, VT: Clinical Psychology Publishing.

McReynolds, P. (1987). Lightner Witmer: Little known founder of clinical psychology. *American Psychologist, 42,* 849–858.

McReynolds, P. (1996). Lightner Witmer: A centennial tribute. *American Psychologist, 51*(3), 237–240.

Meehl, P. E. (1954). *Clinical versus statistical prediction*. Minneapolis, MN: University of Minnesota Press.

Meehl, P. E. (1960). The cognitive activity of the clinician. *The American Psychologist, 15,* 19–27.

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin, 52,* 194–216.

Meier, S. L. (1994). *The chronic crisis in psychological measurement and assessment: A historical survey*. New York: Academic Press.

Meyer, G. J. (1996). The Rorschach and MMPI: Toward a more scientific differential understanding of cross-method assessment. *Journal of Personality Assessment, 67,* 558–578.

Meyer, G. J., & Handler, L. (1997). The ability of the Rorschach to predict subsequent outcome: a meta-analysis of the Rorschach Prognostic Rating Scale. *Journal of Personality Assessment, 69,* 1–38.

Millon, T. (1984). On the renaissance of personality assessment and personality theory. *Journal of Personality Assessment, 48*(5), 450–466.

Millon, T., & Davis, R. D. (1993). The Millon Adolescent Personality Inventory and the Millon Adolescent Clinical Inventory. *Journal of Counseling and Development*.

Mitchell, J. V., Jr. (Ed.) (1985). *The mental measurements yearbook*. Lincoln, NE: Buros Institute of Mental Measurements, University of Nebraska.

Morgan, C. D., & Murray, H. A. (1935). A method for investigating fantasies. *Archives of Neurological Psychiatry, 35,* 289–306.

Murray, H. A. (1943). *Manual for the Thematic Apperception Test*. Cambridge, MA: Harvard University Press.

Murstein, B. I. (1963). *Theory and research in projective techniques*. New York: Wiley.

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist, 47*(9), 1102–1114.

Rotter, J. B., & Rafferty, J. E. (1950). *Manual: The Rotter Incomplete Sentences Blank*. San Antonio, TX: Psychological Corporation.

Reisman, J. M. (1991). *A history of clinical psychology* (2nd ed.). New York: Hemisphere.

Samejima, F. (1988). Comprehensive latent trait theory. *Behaviormetrika, 24,* 1–24.

Samelson, F. (1987). Was early mental testing (a) racist inspired, (b) objective science, (c) a technology for democracy, (d) the origin of multiple-choice exams, (e) none of the above? (Mark the RIGHT answer). In M. M. Sokal (Ed.) *Psychological testing and American society 1890–1930* (pp. 113–127). New Brunswick, NJ: Rutgers University Press.

Sarason, S. B. (1954). *The clinical interaction, with special reference to the Rorschach*. New York: Harper.

Sechrest, L. (1963). Incremental validity: A recommendation. *Educational and Psychological Measurement, 33*(1), 153–158.

Sechrest, L. (1968). Testing, measuring, and assessing people. In W. W. Lambert & E. G. Borgatta (Eds.), *Handbook of personality theory and research*. Chicago: Rand McNally.

Sechrest, L. (1992). The past future of clinical psychology: A reflection on Woodworth (1937). *Journal of Consulting and Clinical Psychology, 60*(1), 18–23.

Sechrest, L., McKnight, P. E., & McKnight, K. M. (1996). Calibration of measures for psychotherapy outcome studies. *American Psychologist, 51,* 1065–1071.

Segal, D. L, Hersen, M., & Van Hasselt, V. B. (1994). Reliability of the structured clinical interview for DSM-III-R: An evaluative review. *Comprehensive Psychiatry, 35*(4), 316–327.

Sharkey, K. J., & Ritzler, B. A. (1985). Comparing the diagnostic validity of the TAT and a New Picture Projective Test. *Journal of Personality Assessment, 49,* 406–412.

Shneidman, E. S. (1986). MAPS of the Harvard Yard. *Journal of Personality Assessment, 50*(3), 436–447.

Somoza, E., Steer, R. A., Beck, A. T., & Clark, D. A. (1994). Differentiating major depression and panic disorders by self-report and clinical rating scales: ROC analysis and information theory. *Behaviour Research and Therapy, 32,* 771–782.

Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1997). *Structured Clinical Interview for DSM-IV Disorders (SCID-I)-Clinician Version*. Washington, DC: American Psychiatric Press.

Spitzer, R. L, Williams, J. B. W., Gibbon, M., & First, M. B. (1992). The Structured Clinical Interview for DSM-III-R (SCID): I. History, rationale, and description. *Archives of General Psychiatry, 49*(8), 624–629.

Steiner, J. L., Tebes, J. K., Sledge, W. H., & Walker, M. L. (1995). A comparison of the Structured Clinical Interview for *DSM-III-R* and clinical diagnoses. *Journal of Nervous & Mental Disease, 183*(6), 365–369.

Strupp, H. H., Horowitz, L. M., & Lambert, M. J. (1997). *Measuring patient changes in mood, anxiety, and personality disorders: Toward a core battery*. Washington, DC: American Psychological Association.

Terman, L. M. (1916). *The measurement of intelligence*. Boston: Houghton Mifflin.

Thorndike, R., & Hagen, E. (1955). *Measurement and evaluation in psychology and education*. New York: Wiley.

Wade, T. C., & Baker, T. B. (1977). Opinions and use of psychological tests: A survey of clinical psychologists. *American Psychologist, 32,* 874–882.

Wallace, J. (1966). An abilities conception of personality: Some implications for personality measurement. *American Psychologist, 21*(2), 132–138.

Ware, J. E., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): 1. Conceptual framework and item selection. *Medical Care, 30*(6), 473–483.

Watkins, C. E. (1991). What have surveys taught us about the teaching and practice of psychological assessment? *Journal of Personality Assessment, 56,* 426–437.

Watkins, C. E., Campbell, V. L., Nieberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology: Research and Practice, 26,* 54–60.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences*. Chicago: Rand McNally.

Webb, E. J., Campbell, D. T., Schwartz, R. D., Sechrest, L., & Grove, J. B. (1981). *Nonreactive measures in the social sciences*. Boston: Houghton Mifflin.

Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment*. Reading, MA: Addison-Wesley.

Williams, J. B. W., Gibbon, M., First, M. B., & Spitzer, R. L (1992). The Structured Clinical Interview for DSM-III-R (SCID): II. Multisite test–retest reliability. *Archives of General Psychiatry, 49*(8), 630–636.

Witmer, L. (1996). Clinical Psychology. *American Psychologist, 51*(3), 248–251. (Original work published 1907.)

Woodworth, R. S. (1992). The future of clinical psychology. *Journal of Consulting and Clinical Psychology, 60,* 16–17. (Original work published 1937.)

Wundt, W., Creighton, J. E., & Titchener, E. B. (1894/1896). *Lectures on human and animal psychology*. London: Swan Sonnenschein.

Yerkes, R. M. (Ed.) (1921). Psychological examining in the United States army. *Memoirs of the National Academy of Sciences, 15*.

Yerkes, R. M. (1941). Man power and military effectiveness: The case for human engineering. *Journal of Consulting Psychology, 5,* 205–209.

Zuckerman, M. (1996, April 22). *Society for a Science of Clinical Psychology Network* (SSCPNET; electonic network).

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.02
Fundamentals of Measurement and Assessment in Psychology

CECIL R. REYNOLDS
Texas A&M University, College Station, TX, USA

4.02.1 INTRODUCTION 33

4.02.2 NORMS AND SCALES OF MEASUREMENT 34
- *4.02.2.1 Scales of Measurement* 34
  - *4.02.2.1.1 Nominal scales* 34
  - *4.02.2.1.2 Ordinal scales* 34
  - *4.02.2.1.3 Interval scales* 35
  - *4.02.2.1.4 Ratio scales* 35
- *4.02.2.2 Norms and Reference Groups* 35

4.02.3 UNITS OF MEASUREMENT 37

4.02.4 ACCURACY OF TEST SCORES 41
- *4.02.4.1 True Score Theory* 41
- *4.02.4.2 Generalizability Theory* 43

4.02.5 VALIDITY 43

4.02.6 THE ASSESSMENT PROCESS 45

4.02.7 MODELS AND METHODS OF ASSESSMENT 46
- *4.02.7.1 Traditional Norm-referenced Assessment* 46
  - *4.02.7.1.1 Intelligence, achievement, and special abilities* 46
- *4.02.7.2 Norm-referenced, Objective Personality Measures* 48
- *4.02.7.3 Projective Assessment* 49
- *4.02.7.4 Behavioral Assessment* 50
- *4.02.7.5 Neuropsychological Assessment* 51

4.02.8 CLINICAL VS. STATISTICAL PREDICTION 52

4.02.9 ACCESSING CRITICAL COMMENTARY ON STANDARDIZED PSYCHOLOGICAL TESTS 53

4.02.10 CONCLUDING REMARKS 53

4.02.11 REFERENCES 54

4.02.1 INTRODUCTION

Measurement is a set of rules for assigning numbers to objects or entities. A psychological measuring device (typically a test), then, is a set of rules (the test questions, directions for administration, scoring criteria, etc.) for assigning numbers to an individual that are believed to represent a level of some particular psychological trait, attribute, or behavior of the individual. These characteristics may be observable directly or may be inferred or observed indirectly through changes in behavior or responses to a set or a variable stimulus. Assessment is a more comprehensive process of deriving meaning from test scores and clinical information in order to describe the individual both broadly and in depth. Psychological tests are the nonexclusive tools of assessment. A proper assessment must also consider the background and current cultural milieu of the individual and actual observed behavior. This chapter does not attempt to deal with all aspects of the assessment process. An introduction to basic measurement technology and theory will be provided along with material concerning different methods of measurement intended to enhance understanding of other chapters in this work.

There are many problems and controversial issues in psychological and educational assessment and, obviously, all cannot be treated in this work. As one example, assessment and the testing that accompanies it occur within a particular situation or context. The results that are obtained may thus be strongly influenced by situational factors in the case of some individuals but less so or not at all for others. The question of the generalizability of test results obtained under a specified set of conditions takes on major importance in interpreting test scores. Not all variables that influence generalizability are known and few that are have been well researched. Test anxiety is one factor thought to influence strongly the generalizability of results across settings and has been researched extensively, yet the complete articulation of the relationship among test anxiety, test performance, and the validity of test-score interpretations across settings is far from complete. The assessment of children, in particular, poses special problems because of the rapid growth and development as well as their susceptibility to external environmental factors. Many of these factors are treated at length in Anastasi (1981), Cronbach (1983), Kaufman (1994), Reynolds (1985), and Reynolds and Kamphaus (1990a, 1990b), and the interested reader is referred to these sources for further reading on the problems, issues, and limitations of educational and psychological testing, as well as to the other chapters in this volume and to Volume 10.

4.02.2 NORMS AND SCALES OF MEASUREMENT

Many pieces of information are necessary before one can attach the proper meaning to a test score. Among the basic are knowledge of what scale of measurement has been employed and with what sort of reference group the individual is being compared, if any. Different scales have different properties and convey different levels and types of information just as they do in other arenas; for example, four inches of water conveys a very different meaning than a reference to four gallons of water. The four basic scales of measurement are nominal, ordinal, interval, and ratio scales. As one moves from nominal scales toward ratio scales, increasingly sophisticated levels of measurement are possible.

4.02.2.1 Scales of Measurement

4.02.2.1.1 Nominal scales

A nominal scale is a qualitative system of categorizing people (or objects, traits, or other variables) or individual observations regarding people typically into mutually exclusive classes or sets. Sex is an example of a nominal scale; one is either male or female. Diagnostic categories such as hyperactivity, learning disabled, aphasia, severely emotionally disturbed, or major depressive disorder represent nominal scaling categories that are not mutually exclusive. Nominal scales provide so little quantitative information about members of categories that some writers prefer to exclude nominal scales from the general rubric of measurement. As Hays (1973) points out, the term measurement typically is reserved for a situation where each individual is assigned a relational number. Because the quantitative relationship among nominal categories is unknown, many common statistical tests cannot be employed with nominal scale data. However, since nominal scales do allow for the classification of an event into a discrete category, many writers (e.g., Nunnally, 1978) do include them as one type of measurement.

4.02.2.1.2 Ordinal scales

Ordinal scales provide considerably more quantitative information regarding an observation than nominal scales. Ordinal scales allow one to rank objects or people according to the amount of a particular attribute displayed. Ordering usually takes the form of the "most" to the "least" amount of the attribute in question. If children in a classroom were weighed and then ranked from heaviest to lightest with the heaviest child assigned the rank of 1, the next heaviest a 2, and so on, until all children had been assigned a number, the resulting measurement would be on an ordinal scale. Although an ordinal scale provides certain quantitative information about each individual, it does not tell how far apart each observation is from the next one. Between adjacent pairs of ranks there may be a different degree of difference. The difference in weight between child 1 and child 2 may be 10 pounds, but the difference between child 2 and child 3 may be one pound or even less. Ordinal scales thus designate relative positions among individuals, an advance over nominal scaling, but are still crude with regard both to describing individuals and to the possible statistical treatments that can be meaningfully applied. Means and standard deviations are usually without meaning when applied to ordinal scales, although the median and mode can be determined and used meaningfully. Age and grade equivalents are examples of common ordinal scales.

4.02.2.1.3 Interval scales

Interval scales afford far more information about observations and can be mathematically manipulated with far greater confidence and precision than nominal or ordinal scales. To have an interval scale of measurement, one must have an ordinal scale on which the difference between any two adjacent points on the scale is equal. Most of the measurement scales and tests used in psychology and education assume an interval scale. Intelligence tests are one good example of an interval scale and can also illustrate the distinction between interval and the highest level of measurement, ratio scales. Although nearly all statistical methods can be applied to measurements on an interval scale, the interval scale has no true zero point, where zero designates total absence of an attribute. If one were to earn an IQ of zero on an intelligence test (by failing to answer a single question correctly), this would not indicate the absence of intelligence, for without intelligence no human could remain alive (it is not possible on most tests of intelligence to earn an IQ of zero even if all test questions are answered incorrectly).

4.02.2.1.4 Ratio scales

Ratio scales possess the attributes of ordinal and interval scales but also have a true zero point—a score of zero indicates the complete absence of the attribute under consideration. Length and width are ratio scales. There are few instances of ratio scales in psychology outside of measurement of simple sensory and motor functions. Ratio scales have useful quantitative features, in particular, as indicated by the name: ratios are meaningful—six feet is twice three feet. Ratios are not meaningful with interval scales. A person with an IQ of 100 cannot be said to be twice as intelligent as a person with an IQ of 50. Fortunately, it is not necessary to have ratio scales to attack the vast majority of problems in psychological assessment.

This discussion of scales of measurement has necessarily been limited to the most basic elements and distinctions among scales. The reader who desires to explore this topic from a technical perspective will find an excellent and extensive mathematical presentation of scales of measurement in Hays (1973).

4.02.2.2 Norms and Reference Groups

To understand the individual's performance as represented by a score on a psychological measurement device, it is necessary, except with certain very specific tests, to evaluate the individual's performance relative to the performance of some preselected group. To know simply that an individual answers 60 out of 100 questions correctly on a history test, and 75 out of 100 questions correctly on a biology test, conveys very little information. On which test did this individual earn the better score? Without knowledge of how a comparable or other relevant group of persons would perform on these tests, the question of which score is better cannot be answered.

Raw scores on a test, such as the number or percentage of correct responses, take on meaning only when evaluated against the performance of a normative or reference group of individuals. For convenience, raw scores are typically converted to a standard or scaled score and then compared against a set of norms. The reference group from which the norms are derived is defined prior to the standardization of the test. Once the appropriate reference population has been defined, a random sample is tested, with each individual tested under as nearly identical procedures as possible. Many factors must be considered when developing norms for test interpretation. Ebel (1972), Angoff (1971), and Petersen, Kolen, and Hoover (1989) have provided especially good discussions of the necessary conditions for appropriate development and use of normative reference data. The following points are taken principally from these three sources, with some elaboration by the present author. Some of these conditions place requirements on the test being normed, some on the psychological trait being measured, and others on the test user.

(i) The psychological trait being assessed must allow the ranking of individuals along a continuum from high to low, that is, it must be amenable to at least ordinal scaling. If a nominal scale was employed, only the presence or absence of the trait would be of interest and relative amounts of the trait could not be determined; norms, under this unusual condition, would be superfluous if not distracting or misleading.

(ii) The content of the test must provide an adequate operational definition of the psychological trait under consideration. With a proper operational definition, other tests can be constructed to measure the same trait and should yield comparable scores for individuals taking both tests.

(iii) The test should assess the same psychological construct throughout the entire range of performance.

(iv) The normative reference group should consist of a large random sample representative of the population on whom the test is to be administered later.

(v) The normative sample of examinees from the population should "have been tested under standard conditions, and ... take the test as seriously, but no more so, than other(s) to be tested later for whom the norms are needed" (Ebel, 1972, p. 488).

(vi) The population sampled to provide normative data must be appropriate to the test and to the purpose for which the test is to be employed. The latter point is often misinterpreted, especially with regard to evaluation of exceptional children. Many adequately normed psychological tests are inappropriately maligned for failure to include significant numbers of handicapped children in their normative sample. The major intelligence scales designed for use with children (i.e., the various Wechsler scales and the McCarthy Scales of Children's Abilities) have been normed on stratified random samples of children representative of children in the United States. Some authors (e.g., Salvia & Ysseldyke, 1981) criticize tests such as the Wechsler scales as inappropriate for measuring the intellectual level of various categories of children with disabilities because large numbers of these children were not included in the test's standardization sample. Whether this is a valid criticism depends on the purpose to which the test is applied. If knowledge of an emotionally disturbed child's level of intellectual functioning relative to age mates in the United States is desired, comparing the child's performance to that of other similarly emotionally disturbed children, then a reference group of emotionally disturbed children would be appropriate. The latter information is not sought frequently nor has it been shown to be more useful in the diagnosis or development of appropriate intervention strategies. Salvia and Ysseldyke (1981) contend that it would be inappropriate to base predictions of future intellectual or academic performance on test scores for an exceptional child that have been derived through comparison with the larger, normal population's performance. To make predictions, they would first require that the reference group from which scores are derived be a group of similar sociocultural background, experience, and handicapping condition. Although this may be an appropriate, if not noble, hypothesis for research, implementation must await empirical verification, especially since it runs counter to traditional psychological practice. Indeed, all interpretations of test scores should be guided principally by empirical evidence. Once norms have been established for a specific reference group, the generalizability of the norms becomes a matter of actuarial research; just as norms based on one group may be inappropriate for use with another group, the norms may also be appropriate and *a priori* acceptance of either hypothesis would be incorrect (Reynolds & Brown, 1984). A large, cumulative body of evidence demonstrates clearly that test scores predict most accurately (and equally well for a variety of subgroups) when based on a large, representative random sample of the population, rather than open highly specific subgroups within a population (e.g., Hunter, Schmidt, & Rauschenberger, 1984; Jensen, 1980; Reynolds, 1982, 1995, in press-a, in press-b).

(vii) Normative data should be provided for as many different groups as it may be useful for an individual to be compared against. Although this may at first glance seem contradictory to the foregoing conclusions, there are instances when it is useful to know how a patient compares to members of other specific subgroups. The more good reference groups available for evaluating a patient's performance on a test, the potentially more useful the test may become.

The normative or reference group most often used to derive scores is the standardization sample, a sample of the target population drawn using a set plan. The best tests, and most publishers and developers of tests, aspire to a standardization sample that is drawn using population proportionate stratified random sampling. This means that samples of people are selected based on subgroups of a larger group to ensure that the population as a whole is represented. In the USA, for example, tests are typically standardized via a sampling plan that stratifies the sample by gender, age, ethnicity, socioeconomic background, region of residence, and community size based on population statistics provided by the US Bureau of the Census. If the Census Bureau data were to indicate, for example, that 1% of the US population consisted of African-American males in the middle range of socioeconomic status residing in urban centers of the south region, then 1% of the standardization sample of the test would be drawn to meet this same set of characteristics.

Once the normative reference group has been obtained and tested, tables of standardized or scaled scores are developed. These tables are based on the responses of the standardization sample and are frequently referred to as norms tables. There are many types of scaled scores or other units of measurement that may be reported in the "norms tables" and just which unit of measurement has been chosen may greatly influence score interpretation.

4.02.3 UNITS OF MEASUREMENT

Raw scores such as number correct are tedious to work with and to interpret properly. Raw scores are thus typically transformed to another unit of measurement. Scaled scores are preferred, but other units such as age and grade equivalents are common. Making raw scores into scaled scores involves creating a set of scores with a predetermined mean and standard deviation that remain constant across some preselected variable such as age.

The mean is simply the sum of the scores obtained by individuals in the standardization sample divided by the number of people in the sample ($\Sigma X_i/N$). In a normal distribution of scores (to be described in the next paragraph), the mean breaks performance on the test into two equal parts, with half of those taking the test scoring above the mean and half scoring below the mean, though the *median* is formally defined as the score point which breaks a distribution into two equal parts; in a normal distribution, the mean and median are the same score.

The standard deviation (SD) is an extremely useful statistic in describing and interpreting a test score. The SD is a measure of the dispersion of scores about the mean. If a test has a mean of 100 and an individual earns a score of 110 on the test, we still have very little information except that the individual performed above average. Once the SD is known, one can determine how far from the mean the score of 110 falls. A score of 110 takes on far different meaning depending on whether the SD of the scores is 5, 15, or 30. The SD is relatively easy to calculate once the mean is known; it is determined by first subtracting each score from the mean, squaring the result, and summing across individuals. This sum of squared deviations from the mean is then divided by the number of persons in the standardization sample. The result is the variance of the test scores; the square root of the variance is the SD.

Once the mean and SD of test scores are known, an individual's standing relative to others on the attribute in question can be determined. The normal distribution or normal curve is most helpful in making these interpretations. Figure 1 shows the normal curve and its relationship to various standard score systems. A person whose score falls 1 SD above the mean performs at a level exceeding about 84% of the population of test-takers. Two SDs will be above 98% of the group. The relationship is the same in the inverse below the mean. A score of 1 SD below the mean indicates that the individual exceeds only about 16% of the population on the attribute in question. Approximately two-thirds (68%) of the population will score within 1 SD of the mean on any psychological test.

Standard scores such as those shown in Figure 1 (*z* scores, *T* scores, etc.) are developed for ease of interpretation. Though standard scores are typically linear transformations of raw scores to a desired scale with a predetermined mean and SD, normalized scaled scores can also be developed. In a linear transformation of test scores to a predetermined mean and SD, equation (1) must be applied to each score:

$$\text{scaled score} = X_{ss} + \mathrm{SD}_{ss}\,\frac{(X_i - X)}{\mathrm{SD}_x} \qquad (1)$$

where X_i = raw score of any individual i, X = mean of the raw scores, SD_x = standard deviation of the raw scores, SD_{ss} = standard deviation scaled scores are to have, and X_{ss} = mean scaled scores are to have.

Virtually all tests designed for use with children along with most adult tests standardize scores and then normalize them within age groups so that a scaled score at one age has the same meaning and percentile rank at all other ages. Thus a person age 10 who earns a scaled score of 105 on the test has the same percentile rank within his or her age group as a 12-year-old with the same score has in his or her age group. That is, the score of 105 will fall at the same point on the normal curve in each case.

Not all scores have this property. Grade and age equivalents are very popular types of scores that are much abused because they are assumed to have scaled score properties when in fact they represent only an ordinal scale. Grade equivalents ignore the dispersion of scores about the mean although the dispersion changes from age to age and grade to grade. Under no circumstances do such equivalent scores qualify as standard scores. Consider the calculation of a grade equivalent. When a test is administered to a group of children, the mean raw score is calculated at each grade level and this mean raw score then is called the "grade equivalent" score for a raw score of that magnitude. If the mean raw score for beginning fourth graders (grade 4.0) on a reading test is 37, then any person

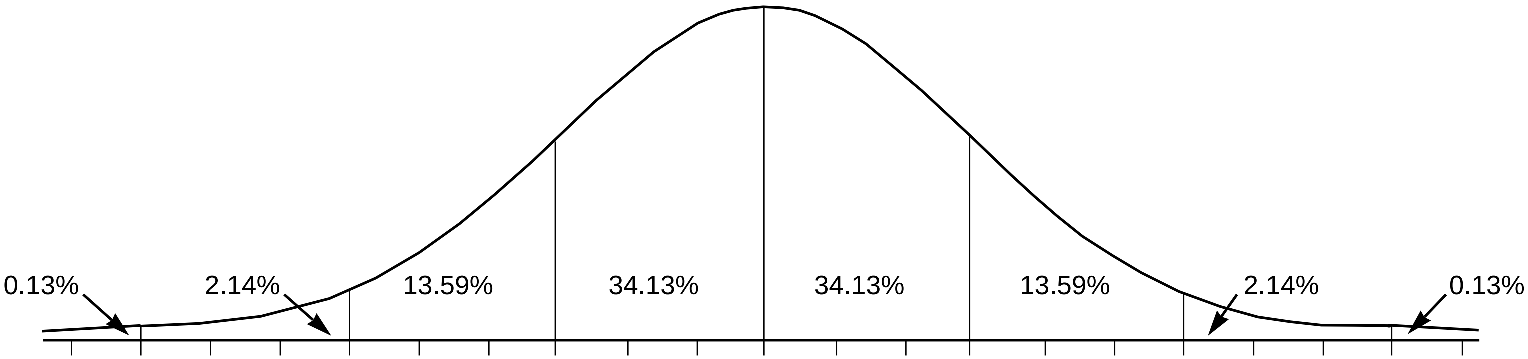

z scores	-3.33	-3	-2.67	-2.33	-2	-1.67	-1.33	-1	-0.67	-0.33	0	0.33	0.67	1	1.33	1.67	2	2.33	2.67	3	3.33
T scores	17	20	23	27	30	33	37	40	43	47	50	53	57	60	63	67	70	73	77	80	83
Wechsler IQ (and others)	50	55	60	65	70	75	80	85	90	95	100	105	110	115	120	125	130	135	140	145	150
Wechsler scale		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
Binet IQ	47	52	57	63	68	73	79	84	89	95	100	105	111	116	121	127	132	137	143	148	153
Binet scale	23	26	29	31	34	37	39	42	45	47	50	53	55	58	61	63	66	69	71	74	77
SAT/GRE scores		200	233	267	300	333	367	400	433	467	500	533	567	600	633	667	700	733	767	800	
Percentile ranks	0.04	0.13	0.38	1	2	5	9	16	25	37	50	63	75	84	91	95	98	99	99.62	99.87	99.96

Stanines	1	2	3	4	5	6	7	8	9
	4%	7%	12%	17%	20%	17%	12%	7%	4%

Figure 1 Relationships among the normal curve, relative standing expressed in percentiles, and various systems of derived scores.

earning a score of 37 on the test is assigned a grade equivalent score of 4.0 regardless of the person's age. If the mean raw score of fifth graders (grade 5.0) is 38, then a score of 38 would receive a grade equivalent of 5.0. A raw score of 37 could represent a grade equivalent of 4.0, 38 could be 5.0, 39 could be 5.1, 40 be 5.3, and 41, 6.0. Thus, differences of one raw score point can cause dramatic differences in the grade equivalents received, and the differences will be inconsistent across grades with regard to the magnitude of the difference in grade equivalents produced by constant changes in raw scores.

Table 1 illustrates the problems of using grade equivalents to evaluate a patient's academic standing relative to his or her peers. Frequently in both research and clinical practice, children of normal intellectual capacity are diagnosed as learning disabled through the use of grade equivalents such as "two years below grade level for age" on a test of academic attainment. The use of this criterion for diagnosing learning disabilities or other academic disorders is clearly inappropriate (Reynolds, 1981a, 1985). As seen in Table 1, a child with a grade equivalent score in reading two years below the appropriate grade placement for age may or may not have a reading problem. At some ages this is within the average range, whereas at others a severe reading problem may be indicated.

Grade equivalents tend to become standards of performance as well, which they clearly are not. Contrary to popular belief, grade equivalent scores on a test do not indicate what level of reading text a child should be using. Grade equivalent scores on tests simply do not have a one-to-one correspondence with reading series placement or the various formulas for determining readability levels.

Grade equivalents are also inappropriate for use in any sort of discrepancy analysis of an individual's test performance, diagnosis of a learning disability or developmental disorder, or for use in many statistical procedures for the following reasons (Reynolds, 1981a).

(i) The growth curve between age and achievement in basic academic subjects flattens at upper grade levels. This can be seen in Table 1 where there is very little change in standard score values corresponding to two years below grade level for age after about grade 7 or 8. In fact, grade equivalents have almost no meaning at this level since reading instruction typically stops by high school. Consider the following analogy with height as an age equivalent. Height can be expressed in age equivalents just as reading can be expressed as grade equivalents. It might be helpful to describe a tall first grader as having the height of an 8½ year old, but what happens to the 5 feet, 10 inch tall 14-year-old female since at no age does the mean height of females equal 5 feet, 10 inches? Since the average reading level in the population changes very little after junior high school, grade equivalents at these ages become virtually nonsensical with large fluctuations resulting from a raw score difference of two or three points on a 100-item test.

(ii) Grade equivalents assume that the rate of learning is constant throughout the school year and that there is no gain or loss during summer vacation.

(iii) Grade equivalents involve an excess of extrapolation, especially at the upper and lower ends of the scale. However, since tests are not administered during every month of the school year, scores between the testing intervals (often a full year) must be interpolated on the assumption of constant growth rates. Interpolations between sometimes extrapolated values on an assumption of constant growth rates is a somewhat ludicrous activity.

(iv) Different academic subjects are acquired at different rates and the variation in performance varies across content areas so that "two years below grade level for age" may be a much more serious deficiency in math than in reading comprehension.

(v) Grade equivalents exaggerate small differences in performance between individuals and for a single individual across tests. Some test authors even provide a caution on record forms that standard scores only, and not grade equivalents, should be used for comparisons.

Age equivalents have many of the same problems. The standard deviation of age equivalents varies substantially across tests, subsets, abilities, or skills assessed, and exist on an ordinal, not interval scale. It is inappropriate to add, subtract, multiply, or divide age or grade equivalents or any other form of ordinal score. Nevertheless, the use of such equivalent scores in *ipsative* analysis of test performance remains a common mistake in clinical, educational, and neuropsychological assessment.

The principal advantage of standardized or scaled scores lies in the comparability of score interpretation across age. By standard scores of course, I refer to scores scaled to a constant mean and SD such as the Wechsler Deviation IQ and not to ratio IQ types of scales employed by the early Binet and original Slosson Intelligence Test, which give the false appearance of being scaled scores. Ratio IQs or other types of quotients have many of the same problems as grade equivalents and should be avoided for many of these same reasons. Standard scores of the deviation IQ type have the same percentile rank across age since they

Table 1 Standard scores and percentile ranks corresponding to performance "two years below grade level for age" on three reading tests.

		Wide range achievement test		*Woodcock Reading Mastery Test*[a]		*Stanford Diagnostic Reading Test*[a]	
Grade placement	*Two years below placement*	*SS*[b]	*%R*[c]	*SS*	*%R*	*SS*	*%R*
2.5	K.5	67	1				
3.5	1.5	69	2	64	1	64	1
4.5	2.5	73	4	77	6	64	1
5.5	3.5	84	14	85	16	77	6
6.5	4.5	88	21	91	27	91	27
7.5	5.5	86	18	94	34	92	30
8.5	6.5	87	19	94	34	93	32
9.5	7.5	90	25	96	39	95	34
10.5	8.5	85	16	95	37	95	37
11.5	9.5	85	16	95	37	92	30

[a] Total test. [b] All standard scores in this table have been converted for ease of comparison to a common scale having a mean of 100 and an SD of 15. [c] Percentile rank.
Source: Adapted from Reynolds (1981a).

are based not only on the mean but the variability in scores about the mean at each age level. For example, a score that falls two-thirds of a SD below the mean has a percentile rank of 25 at every age. A score falling two-thirds of a grade level below the average grade level or an age equivalent six months below chronological age have different percentile ranks at every age.

Standard scores are more accurate and precise. When constructing tables for the conversion of raw scores into standard scores, interpolation of scores to arrive at an exact score point is typically not necessary, whereas the opposite is true of age and grade equivalents. Extrapolation is also typically not necessary for scores within 3 SDs of the mean, which accounts for more than 99% of all scores encountered.

Scaled scores can be set to any desired mean and standard deviation, with the fancy of the test author frequently the sole determining factor. Fortunately, a few scales can account for the vast majority of standardized tests in psychology and education. Table 2 illustrates the relationship between various scaled score systems. If reference groups are comparable, Table 2 can also be used to equate scores across tests to aid in the comparison of a patient's performance on tests of different attributes, provided normalized scores are provided.

What has been said thus far about scaled scores and their equivalency applies primarily to scores that have been forced to take the shape of the Gaussian or bell curve. When test-score distributions derived from a standardization sample are examined, the scores frequently deviate significantly from normal. Often, test developers will then transform scores, using one of a variety of statistical methods (e.g., see Lord & Novick, 1968, for a mathematical review and explication), to take a normal distribution. Despite what is often taught in early courses in psychological statistics and measurement, this is not always appropriate. It is commonplace to read that psychological variables, like most others, are normally distributed within the population; many are. Variables such as intelligence, memory skill, and academic achievement will closely approximate the normal distribution when well measured. However, many psychological variables, especially behavioral ones such as aggression, attention, and hyperactivity, deviate substantially from the normal curve within the population of humans.

When a score distribution then deviates from normality, the test developer is faced with the decision of whether to create normalized scores via some transformation or to allow the distribution to retain its shape with perhaps some smoothing to remove irregularities due to sampling error. In the later case, a linear transformation of scores is most likely to be chosen.

To make this determination, the test developer must ascertain whether the underlying construct measured by the test is normally distributed or not and whether the extant sample is adequate to estimate the distribution, whatever its shape. For applied, clinical devices, the purpose of score transformations that result in normalization of the distribution is to correct for sampling error and presumes that the underlying construct is, in fact, normally or near normally distributed. Normalization of the score distribution then produces a more

Table 2 Conversion of standard scores based on several scales to a commonly expressed metric.[a]

Scales									
X = 0 *SD = 1*	*X = 10* *SD = 3*	*X = 36* *SD = 6*	*X = 50* *SD = 10*	*X = 50* *SD = 15*	*X = 100* *SD = 15*	*X = 100* *SD = 16*	*X = 100* *SD = 20*	*X = 500* *SD = 100*	*Percentile rank*
2.6	18	52	76	89	139	142	152	760	>99
2.4	17	51	74	86	136	138	148	740	99
2.2	17	49	72	83	133	135	144	720	99
2.0	16	48	70	80	130	132	140	700	98
1.8	15	47	68	77	127	129	136	680	96
1.6	15	46	66	74	124	126	132	660	95
1.4	14	44	64	71	121	122	128	640	92
1.2	14	43	62	68	118	119	124	620	88
1.0	13	42	60	65	115	116	120	600	84
0.8	12	41	58	62	112	113	116	580	79
0.6	12	40	56	59	109	110	112	560	73
0.4	11	38	54	56	106	106	108	540	66
0.2	11	37	52	53	103	103	104	520	58
0.0	10	36	50	50	100	100	100	500	50
−0.2	9	35	48	47	97	97	96	480	42
−0.4	9	34	46	44	94	94	92	460	34
−0.6	8	33	44	41	91	90	88	440	27
−0.8	8	31	42	38	88	87	84	420	21
−1.0	7	30	40	35	85	84	80	400	16
−1.2	6	29	38	32	82	81	76	380	12
−1.4	6	28	36	29	79	78	72	360	8
−1.6	5	26	34	26	76	74	68	340	5
−1.8	5	25	32	23	73	71	64	320	4
−2.0	4	24	30	20	70	68	60	300	2
−2.2	3	23	28	17	67	65	56	280	1
−2.4	3	21	26	14	64	62	52	260	1
−2.6	2	20	24	13	61	58	48	240	1

[a] *X* = mean; SD = standard deviation.

accurate rendition of the population distribution and improves the utility of the standardized scaled scores provided. If the population distribution of the construct in question is not normal, for example, aggressive behavior (see Reynolds & Kamphaus, 1992), then a different form of transformation, typically linear, is required to be accurate. This decision affects how clinicians best interpret the ultimately scaled scores.

If score distributions have been normalized for a battery of tests or subtests of a common test, for example, the Wechsler scales, the same scaled score on any part-test will have the same percentile rank. On the Wechsler Intelligence Scale for Children-III (WISC-III; Wechsler, 1992), for example, a subtest scaled score of 13 is 1 SD above the mean and, for all 13 subtests of the WISC-III, will have a percentile rank of approximately 86. If the scores had not been transformed through the nonlinear methods necessary to approximate a normal distribution, this would not be true. For a linear transformation, a scaled score of 13 could still be 1 SD above the mean on all of the subtests but the percentile rank could vary, and could vary substantially the more the underlying distribution deviates from that of the normal curve. It is thus important for clinicians to review test manuals and ascertain the methods of scaling that have been applied to the raw score distributions. This becomes increasingly important as scores are to be compared across different tests or batteries of tests. This effect is magnified as the distance from the mean increases.

4.02.4 ACCURACY OF TEST SCORES

4.02.4.1 True Score Theory

When evaluating test scores, it is also necessary to know just how accurately the score reflects the individual's true score on the test. Tests typically do not ask every possible question that could be asked or evaluate every possible relevant behavior. Rather a domain of possible questions or test items is defined and a sample taken to form the test. Whenever less than the total number of possible behaviors within a domain is sampled, sampling error occurs. Psychological and educational tests are thus destined to be less than perfect in their accuracy. Certainly, psychological tests contain errors produced from a variety of other sources, most of which are situational. Error resulting from domain sampling is the largest contributor to the degree of error in a test score, however (Feldt & Brennan, 1989; Nunnally, 1978), and is the type of error about which measurement theory has the greatest concern. Fortunately, this type of error is also the easiest and most accurately estimated.

Error caused by domain sampling is determined from an analysis of the degree of homogeneity of the items in the test, that is, how well the various items correlate with one another and with an individual's true standing on the trait being assessed. The relative accuracy of a test is represented by a reliability coefficient symbolized as r_{xx}. Since it is based on the homogeneity or consistency of the individual items of a test and no outside criteria or information are necessary for its calculation, r_{xx} is frequently referred to as internal consistency reliability or as an estimate of item homogeneity. Error caused by domain sampling is also sometimes estimated by determining the correlation between two parallel forms of a test (forms of a test that are designed to measure the same variable with items sampled from the same item domain and believed to be equivalent). The correlation between the two equivalent or alternate forms is then taken as the reliability estimate and is usually symbolized as r_{xx}, r_{ab}, or r_{xy} (although r_{xy} is generally used to represent a validity coefficient).

Split-half reliability estimates can also be determined on any specific test as a measure of internal consistency. Split-half reliability is typically determined by correlating each person's score on the one-half of the items with his or her score on the other half of the test with a correction for the original length of the test, since length will affect reliability. Predetermined or planned split-half comparisons such as correlating scores on odd numbered items with scores on the even numbered items may take advantage of chance or other factors resulting in spuriously high estimates of reliability. A reliability coefficient called alpha is a better method for estimating reliability since it is the mean of all possible split-half comparisons, thus expunging any sampling error resulting from the method of dividing the test for the purposes of calculating a correlation between each half.

As noted earlier, a number of techniques exist for estimating reliability. Throughout this chapter, reliability has been referred to as estimated. This is because the absolute or "true" reliability of a psychological test can never be determined. Alpha and all other methods of determining reliability are, however, considered to be lower bound estimates of the true reliability of the test. One can be certain that the reliability of a test is at least as high as the calculated estimate and possibly even higher.

Once the reliability of a test has been estimated, it is possible to calculate a sometimes more useful statistic known as the standard error of measurement. Since there is always some error involved in the score a person obtains on a psychological test, the obtained score (X_i) does not truly represent the individual's standing with regard to the trait in question. Obtained scores estimate an individual's true score on the test (the score that would be obtained if there was no error involved in the measurement). Since this is not possible, the true score (X_∞) is defined as the mean score of an individual if administered an infinite number of equivalent forms of a test and there were no practice effects or other intervening factors. The standard error of measurement (S_{em}) is the SD of the individual's distribution of scores about his or her true score. To determine the S_{em} it is necessary to know only the SD and the reliability (preferably an internal consistency estimate) of the test in question. The calculation of X_∞ and S_{em} are only estimates, however, since the conditions for determining a true score never actually exist.

Since the distribution of obtained scores about the true score is considered to be normal, one can establish a degree of confidence in test results by banding the estimated true score by a specified number of S_{em}s. A table of values associated with the normal curve (pictured in Figure 1) quickly tells us how many S_{em}s are necessary for a given level of confidence. In a normal distribution, about 68% of all scores fall within 1 SD of the mean, and about 95% of all scores fall within 2 SDs of the mean. Therefore, if one wanted to be 68% certain that a range of scores contained a person's true score, X_∞ would be banded by $\pm 1\ S_{em}$. To be 95% certain that a range of scores contained the true score, a range of $X_\infty \pm 2\ S_{em}$s would be necessary.

When evaluating a test or performance on a test, it is important to ascertain just what type of reliability estimate is being reported. S_{em}s should typically be calculated from an internal consistency estimate. Comparisons of reliability estimates across tests should be based on the same type of estimate. For example, one should not compare the reliability of two tests based on alternate form correlations for one test and estimation of the alpha coefficient for the other.

Test–retest correlations, also frequently referred to as reliability coefficients, should not be confused with measures of the accuracy or precision of a test at a given point in time.

Test–retest "reliability" is one of the most often confused concepts of psychometric theory. Even Anastasi (1976), in introducing reliability, refers to reliability as a measure of the degree to which a person would obtain the same score if tested again at a later time. In the earlier stages of development of psychology when traits were considered unchanging, test–retest reliability was properly considered to be a characteristic of the test and indeed was believed to be an indication of the degree to which a person would obtain the same score if tested again. Test–retest reliability speaks principally of the stability of the trait being measured and has little to do with the accuracy or precision of measurement unless the psychological construct in question is considered to be totally unchangeable. Given that traits such as anxiety and even intelligence do in fact change over time and that testing from one time to the next is positively correlated, it is still possible to use the test–retest correlation to determine estimates of what score a person would obtain upon retesting. Internal consistency estimates, however, should not be interpreted in such a manner. When psychological constructs are not highly labile and believed to change only over long periods of time, test–retest correlations may be considered to reflect the accuracy of a test if the two testings occur at close points in time during which the trait under consideration is believed to be stable.

4.02.4.2 Generalizability Theory

Generalizability theory is an extension of true score theory (also known as classical test theory) that is achieved principally through use of analysis of variance (ANOVA) procedures. Often, more than one type of error is acting on a reliability coefficient. For example, in true score theory, errors due to domain sampling (e.g., not asking about every possible symptom of depression), errors due to faulty administration, scoring errors by the examiner, and errors associated with time sampling may all act to lower the average interitem correlation, which will reduce the internal consistency reliability of the test score. Under true score theory, it is impossible to partial the relative contributions, that is, to determine how much error is contributed by each subset of error to the total amount of unreliability. Even test–retest or stability coefficients are confounded by internal consisting errors. The maximum r_{12} is equal to the square root of r_{xx} or max $r^{12} = (r_{xx})^{1/2}$.

Generalizability theory takes advantage of the capabilities of ANOVA in partitioning variance components to develop a model of unreliability (as opposed to concentrating on statistical significance). Through ANOVA, generalizability theory is able to partition the error variance of a set of scores into the components listed above, such as domain sampling error and the like, along with some additional components not considered in true score theory. Generalizability theory is no more difficult mathematically than true score theory. Generalizability theory is surprisingly absent from the measurement repertoire of most clinicians but is becoming increasingly popular among measurement scientists. However, the understanding and application of generalizability theory does require an understanding of methods and designs for partitioning variance components in ANOVA, a skill that is perhaps on the decline in clinical training programs in favor of statistical methods more aligned with structural equation modeling.

The basic foundations of generalizability theory can be found in Cronbach, Rajaratnam, and Gleser (1963). A current, detailed explanation appears in Feldt and Brennan (1989) along with the necessary mathematical models necessary to apply generalizability theory to the concept of error in test scores of groups and individuals.

4.02.5 VALIDITY

Reliability refers to the precision or accuracy of test scores. Validity refers to the appropriateness of the interpretations of test scores and not to the test or the score itself. "Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and the appropriateness of inferences and actions based on test scores or other modes of assessment" (Messick, 1989, p. 13). As is reliability, validity is a matter of degree and not an all or none concept. Reliability will, however, enter into evaluation of the validity of an inference drawn from a test score. Reliability is a necessary but insufficient condition for validity. As reliability approaches zero, the amount of random error in a test score increases. The greater the relative proportion of random error present, the less confidence one can have in any interpretation of a score since, by definition, random error is unrelated to anything meaningful. Validation is not static but is an ongoing process, not just of the corroboration of a particular meaning, but for the development of sounder and better interpretations of observations that are expressed as scores on a psychological test.

Although it is often done as a matter of convenience or as simple short hand, it should be obvious by now that it is not correct technically to refer to the validity of a test. Validity is a characteristic of the interpretation given to performance on a test. It makes no sense, for example, to ask a question such as "Is this Wechsler scale a valid test?" Rather, one might pose the superior question "Is the interpretation of performance on this Wechsler scale as reflecting intelligence or intellectual level valid?" This is more than a game of semantics as such subtle differences in language affect the way we think about our methods and our devices. The difference in language and its implications are considered powerful enough that *Educational and Psychological Measurement*, one of the oldest and most respected journals in psychometrics, founded originally by Frederic Kuder, no longer allows authors in its pages to refer to the validity of a test or the reliability of a test. Reviewers for this journal are asked routinely to screen manuscripts for improper or imprecise use of such terminology.

Just as reliability may take on a number of variations, so may validity. Quite a bit of divergent nomenclature has been applied to validity. Messick (1980) identified 17 "different" types of validity that are referred to in the technical literature! Traditionally, validity has been broken into three major categories: content, construct, and predictive or criterion-related validity. These are the three types of validity distinguished and discussed in the joint *Standards for Educational and Psychological Tests* (American Psychological Association, 1985). Construct validity cuts across all categories, and criterion-related validity is definitely a question of the relationship of test performance to other methods of evaluating behavior.

Content validity is determined by how well the test items and their specific content sample the set of behaviors or subject matter area about which inferences are to be drawn on the basis of the test scores. Criterion-related or predictive validity refers to either comparisons of test scores with performance on accepted criteria of the construct in question taken in close temporal relationship to the test or the level of prediction of performance at some future time. Criterion-related validity is determined by the degree of correspondence between the test score and the individual's performance on the criterion. If the correlation between these two variables is high, no further evidence may be considered necessary (Nunnally, 1978). Here, reliability has a direct, and known, limiting effect on validity. A correlation between a predictor (x) and a criterion (y), a validity coefficient, typically expressed as r_{xy}, is restricted in magnitude. Its maximum true value is equal to the square root of the product of the internal consistency reliability coefficients of the scores being compared: $r_{xy}\ \max = (r_{xx}\, r_{yy})^{1/2}$.

Construct validity of the interpretations given to psychological tests is one of the most complex issues facing the psychometrician and permeates all aspects of test development and test use. Psychology for the most part deals with intangible constructs. Intelligence is one of the most intensely studied constructs in the field of psychology, yet it cannot be directly observed or evaluated. Intelligence can only be inferred from the observation and quantification of what has been agreed upon as "intelligent" behavior. Personality variables such as dependence, anxiety, need achievement, mania, and on through the seemingly endless list of personality traits that psychologists have "identified" also cannot be observed directly. Their existence is only inferred from the observation of behavior. Construct validity thus involves considerable inference on the part of the test developer and the researcher; construct validity is evaluated by investigating just what psychological properties a test measures.

Prior to being used for other than research purposes, interpretations given to a test must be shown clearly to demonstrate an acceptable level of validity. For use with various categories of psychopathology, validation with normally functioning individuals should be considered insufficient. The validity of an interpretation needs to be demonstrated for each group with whom it is used. This can be a long and laborious process but is nevertheless a necessary one. There are many subtle characteristics of various classes of exceptional children, for example, that may cause an otherwise appropriate interpretation of a test to lack validity with special groups (e.g., see Newland, 1980).

As has been noted by Cronbach (1971) and others, the term "test validation" can cause some confusion. In thinking about and evaluating validity, we must always keep in mind that one does not ever actually validate a test but only the interpretation that is given to the score on the test. Any test may have many applications and a test with originally a singular purpose may prove promising for other applications. Each application of a test or interpretation of a test score must undergo validation. Whenever hearing or reading that a test has been validated, we need to know for what purpose it has been validated, and what interpretations of scores from the instrument in question have been shown empirically to be justifiable and accurate.

4.02.6 THE ASSESSMENT PROCESS

As noted at the opening of this chapter, assessment is an involved, comprehensive process of deriving meaning from test scores to achieve a broad but detailed description and understanding of the individual. The description here of assessment as a process is important. Assessment, properly carried out, is not a static collection of information, but an ongoing dynamic synthesis and evaluation of data, reliably obtained, from multiple sources relevant to the current, and possibly future, status of the individual. Assessment is open ended—new information can occur daily that can properly alter one's perception of the ecological validity of prior impressions and recommendations.

Crucial to the assessment process, and far too frequently neglected or overlooked, is follow-up evaluation that should occur after more formal diagnostic assessments have been made and habilitative recommendations implemented. There are no absolutes in psychological and educational testing; no profile of assessment information is inexorably linked with a single method of treatment, remediation, or intervention that will always be successful. Currently, the opposite is the case; the search for the aptitude × treatment interaction is nearly as elusive as that for the neural engram. The follow-up component of the assessment process is crucial to the fine-tuning of existing intervention procedures and in some cases more massive overhauling of an intervention.

Psychological and educational testing and assessment are far from exact, just as are the clinical assessment procedures of medicine and related specialties. When used in diagnosis, assessment allows one simply to narrow the number of disorders under serious consideration. Similarly, when used in the search for an appropriate method of habilitation for a handicapped youngster, the assessment process allows the psychologist to narrow the number of strategies (i.e., hypotheses) from which to choose one that is believed to be most effective. There are no guarantees that the first strategy adopted will be the most effective program of treatment (or be effective at all for that matter). Kaufman (1994) described the proper attitude of the psychologist involved in assessment to be that of a "detective" who evaluates, synthesizes, and integrates data gleaned from the assessment process with knowledge of psychological theories of development and the psychology of individual differences (also see Reynolds, 1981b; Reynolds & Clark, 1982). As described here, the assessment process is a major component in psychological problem-solving. Individuals are not randomly selected for an expensive, time-consuming psychological evaluation. They are referred to a psychologist for some more or less specific reason; a problem of some kind exists. The assessment process then plays a major role in accurately identifying and describing the problem, suggesting solutions, and properly carried through, provides ideas for modifying the initially proposed interventions.

It is necessary in the assessment process to entertain and evaluate information from a variety of sources if the assessment is to be ecologically valid. Each situation will dictate the relevance and appropriate weighting of each piece of information. Age and physical condition are two obvious factors that influence the gathering of information regarding child and adult patients. Palmer (1980), Newland (1980), Salvia and Ysseldyke (1981), and Sattler (1988) have discussed factors to be included in the assessment process when evaluating exceptional children in the schools. The following are generally accepted to be important aspects of assessment: medical condition, sensory and motor skills, school performance and behavior (e.g., group achievement tests, grades, teacher checklists), individual intelligence test scores, special aptitude and achievement test performance, affective characteristics (e.g., personality tests), teacher reports on behavior and peer interaction, the child–school interaction, characteristics of the classroom, parent reports on behavior, the social and cultural milieu of the home, and the child's developmental history. Each of these factors takes on more or less importance for individual patients. With adult patients, many of the same types of information will be relevant with a conceptual shift toward adulthood (Anastasi & Urbina, 1997). The patient's vocational functioning and relationships including parents, spouse, and children will all need to be considered when designing the assessment and later when interpreting the results. More specialized types of knowledge may be required for any given case. For example, in certain genetically-based disorders, a complete family history may be necessary to achieve a good understanding of the nature of the patient's difficulty.

Numerous methods of psychological testing can be used in the assessment process. Each will have its own strengths and weaknesses. There are frequent debates in the psychological literature over the relative merits of one category of assessment over another, with some respondents carrying on with nearly religious fervor. However, these arguments can be resolved quickly by recalling that tests are tools of assessment and most certainly not an end in themselves. Different methods and techniques of testing are best seen and used as complementary in assessment, which is a problem-solving process requiring much information. With these admonitions in mind, it is time to turn to a discussion of various methods of testing and their role in the assessment process.

4.02.7 MODELS AND METHODS OF ASSESSMENT

A variety of assessment methods are available for evaluating adults and exceptional children. Some of these methods grew directly from specific schools of psychological thought, such as the psychoanalytic view of Freud (projective assessment techniques) or the behavioral schools of Watson, Skinner, and Bandura (applied behavior analysis). Other methods have grown out of controversies in and between existing academic disciplines such as personality theory and social psychology. New and refined methods have come about with new developments in medicine and related fields, whereas other new testing methods stem from advances in the theory and technology of the science of psychological measurement. Unfortunately, still other new techniques stem from psychological and educational faddism with little basis in psychological theory and little if any empirical basis. Any attempt to group tests by characteristics such as norm-referenced vs. criterion-referenced, traditional vs. behavioral, maximum vs. typical performance, and so on, is doomed to criticism. As will be seen in the pages that follow, the demarcations between assessment methods and models are not so clear as many would contend. In many cases, the greatest distinctions lie in the philosophical orientation and intent of the user. As one prominent example, many "behavioral" assessment techniques are as bound by norms and other traditional psychometric concepts as are traditional intelligence tests (Cone, 1977). Even trait measures of personality end up being labeled by some as behavioral assessment devices (e.g., Barrios, Hartmann, & Shigetomi, 1981). The division of models and methods of assessment to follow is based in some part on convenience and clarity of discussions but also with an eye toward maintaining the most important conceptual distinctions among these assessment methods.

4.02.7.1 Traditional Norm-referenced Assessment

4.02.7.1.1 Intelligence, achievement, and special abilities

These assessment techniques have been grouped together primarily because of their similarity of content and, in some cases, their similarity of purpose. There are, however, some basic distinctions among these measures. Intelligence tests tend to be broad in terms of content; items sample a variety of behaviors that are considered to intellectual in nature. Intelligence tests are used both to evaluate the current intellectual status of the individual and to predict future behavior on intellectually demanding tasks and to help achieve a better understanding of past behavior and performance in an intellectual setting. Achievement tests measure relatively narrowly defined *content*, sampled from a specific subject matter domain that typically has been the focus of purposeful study and learning by the population for whom the test is intended. Intelligence tests by contrast are oriented more toward testing intellectual processes and use items that are more related to incidental learning and not as likely to have been specifically studied as are achievement test items. Tests of special abilities, such as memory, mechanical aptitude, and auditory perception, are narrow in scope as are achievement tests but focus on process rather than content. The same test question may appear on an intelligence, achievement, or special ability test, however, and closely related questions frequently do. Tests of intelligence and special abilities also focus more on the application of previously acquired knowledge, whereas achievement tests focus on testing just what knowledge has been acquired. One should not focus on single items; it is the collection of items and the use and evaluation of the individual's score on the test that are the differentiating factors.

(i) Intelligence tests

Intelligence tests are among the oldest devices in the psychometric arsenal of the psychologist and are likely the most frequently used category of tests in the evaluation of exceptional children, especially in the cases of mental retardation, learning disabilities, and intellectual giftedness. Intelligence and aptitude tests are used frequently in adult assessment as well and are essential diagnostic tools when examining for the various dementias. They are used with adults in predicting a variety of other cognitive disorders and in the vocational arena. Since the translation and modification of Alfred Binet's intelligence test for French schoolchildren was introduced in the United States by Lewis Terman (of Stanford University, hence the Stanford–Binet Intelligence Scale), a substantial proliferation of such tests has occurred. Many of these tests measure very limited aspects of intelligence (e.g., Peabody Picture Vocabulary Test, Columbia Mental Maturity Scale, Ammons and Ammons Quick Test), whereas others give a much broader view of a person's intellectual skills, measuring general intelligence as well as more specific cognitive skills (e.g., the various Wechsler scales). Unfortunately, while intelligence is a hypothetical psychological construct, most intelligence tests were developed from a primarily empirical basis, with little if any attention given to theories of the human intellect. Empiricism is of major importance in all aspects of psychology, especially psychological testing, but is insufficient in itself. It is important to have a good theory underlying the assessment of any theoretical construct such as intelligence.

Intelligence tests in use today are for the most part individually administered (i.e., a psychologist administers the test to an individual in a closed setting with no other individuals present). For a long time, group intelligence tests were used throughout the public schools and in the military. Group tests of intelligence are used more sparingly today because of their many abuses in the past and the limited amount of information they offer about the individual. There is little of utility to the schools, for example, that can be gleaned from a group intelligence test that cannot be obtained better from group achievement tests. Individual intelligence tests are far more expensive to use but offer considerably more and better information. Much of the additional information, however, comes from having a highly trained observer (the psychologist) interacting with the person for more than an hour in a quite structured setting, with a variety of tasks of varying levels of difficulty. The most widely used individually administered intelligence scales today are the Wechsler scales, the Kaufman scales, and the Stanford–Binet Intelligence Scale (Fourth Edition). Though the oldest and best known of intelligence tests, the Binet has lost much of its popularity and is now a distant third.

Intelligence testing, which can be very useful in clinical and vocational settings, is also a controversial activity, especially with regard to the diagnosis of mild mental retardation among minority cultures in the United States. Used with care and compassion, as a tool toward understanding, such tests can prove invaluable. Used recklessly and with rigidity, they can cause irreparable harm. Extensive technical training is required to master properly the administration of an individual intelligence test (or any individual test for that matter). Even greater sensitivity and training are required to interpret these powerful and controversial devices. Extensive knowledge of statistics, measurement theory, and the existing research literature concerning testing is a prerequisite to using intelligence tests. To use them well requires mastery of the broader field of psychology, especially differential psychology, the psychological science that focuses on the psychological study and analysis of human individual differences and theories of cognitive development. Extensive discussions of the clinical evaluation of intelligence can be found in Kaufman (1990, 1994) and Kaufman and Reynolds (1983).

(ii) Achievement tests

Various types of achievement tests are used throughout the public schools with regular classroom and exceptional children. Most achievement tests are group tests administered with some regularity to all students in a school or system. Some of the more prominent group tests include the Iowa Test of Basic Skills, the Metropolitan Achievement Test, the Stanford Achievement Test, and the California Achievement Test. These batteries of achievement tests typically do not report an overall index of achievement but rather report separately on achievement in such academic areas as English grammar and punctuation, spelling, map reading, mathematical calculations, reading comprehension, social studies, and general science. The tests change every few grade levels to accommodate changes in curriculum emphasis. Group achievement tests provide schools with information concerning how their children are achieving in these various subject areas relative to other school systems throughout the country and relative to other schools in the same district. They also provide information about the progress of individual children and can serve as good screening measures in attempting to identify children at the upper and lower ends of the achievement continuum. Group administered achievement tests help in achieving a good understanding of the academic performance of these individuals but do not provide sufficiently detailed or sensitive information on which to base major decisions. When decision-making is called for or an in-depth understanding of a child's academic needs is required, individual testing is needed.

Psychologists use achievement measures with adult clients as well. With the elderly, acquired academic skills tend to be well preserved in the early stages of most dementias and provide a good baseline of promorbid skills. Academic skills can also be important in recommending job placements, as a component of child custody evaluations, in rehabilitation planning, and in the diagnosis of adult learning disorders and adult forms of attention deficit hyperactivity disorder.

(iii) Tests of special abilities

These are specialized methods for assessing thin slices of the spectrum of abilities for any single individual. These measures can be helpful in further narrowing the field of hypotheses about an individual's learning or behavior difficulties when used in conjunction with intelligence, achievement, and personality measures. The number of special abilities that can be assessed is quite large. Some examples of these abilities include visual–motor integration skills, auditory perception, visual closure, figure–ground distinction, oral expression, tactile form recognition, and psychomotor speed. While these measures can be useful, depending on the questions to be answered, one must be particularly careful in choosing an appropriate, valid, and reliable measure of a special ability. The use and demand for these tests are significantly less than that for the most popular individual intelligence tests and widely used achievement tests. This in turn places some economic constraints on development and standardization procedures, which are very costly enterprises when properly conducted. One should always be wary of the "quick and dirty" entry into the ability testing market. There are some very good tests of special abilities available, although special caution is needed. For example, simply because an ability is named in the test title is no guarantee that the test measures that particular ability. As with all tests, what is actually measured by any collection of test items is a matter for empirical investigation and is subject to the process of validation.

To summarize, norm-referenced tests of intelligence, achievement, and special abilities provide potentially important information in the assessment process. Yet each supplies only a piece of the required data. Equally important are observations of how the patient behaves during testing and in other settings, and performance on other measures.

4.02.7.2 Norm-referenced, Objective Personality Measures

Whereas tests of aptitude and achievement can be described as maximum performance measures, tests of personality can be described as typical performance measures. When taking a personality test, one is normally asked to respond according to one's typical actions and attitudes and not in a manner that would present the "best" possible performance (i.e., most socially desirable). The "faking" or deliberate distortion of responses is certainly possible, to a greater extent on some scales than others (e.g., Jean & Reynolds, 1982; Reynolds, 1998b), and is a more significant problem with personality scales than cognitive scales. Papers have even been published providing details on how to distort responses on personality tests in the desired direction (e.g. Whyte, 1967). Although there is no direct solution to this problem, many personality measures have built-in "lie" scales or social desirability scales to help detect deliberate faking to make one look as good as possible and F or infrequency scales to detect the faking of the presence of psychopathology.

The use and interpretation of scores from objective personality scales also has implications for this problem. Properly assessed and evaluated from an empirical basis, response to the personality scale is treated as the behavior of immediate interest and the actual content conveyed by the item becomes nearly irrelevant. As one example, there is an item on the Revised-Children's Manifest Anxiety Scale (RCMAS; Reynold's & Richmond, 1978, 1985), a test designed to measure chronic anxiety levels in children, that states "My hands feel sweaty." Whether the child's hands actually do feel sweaty is irrelevant. The salient question is whether children who respond "true" to this question are in reality more anxious than children who respond "false" to such a query. Children who respond more often in the keyed direction on the RCMAS display greater general anxiety and exhibit more observed behavior problems associated with anxiety than children who respond in the opposite manner. Although face validity of a personality or other test is a desirable quality, it is not always a necessary one. It is the actuarial implications of the behavioral response of choosing to respond in a certain manner that holds the greatest interest for the practitioner. Scales developed using such an approach are empirical scales.

Another approach is to devise content scales. As the name implies, the item content of such scales is considered more salient than its purely empirical relationship to the construct. Individuals with depression, especially men, may be edgy and irritable at times. Thus, the item "I sometimes lash out at others for no good reason" might show up on an empirically derived scale assessing depression, but is unlikely to find its way onto a content scale. "I am most often sad" would be a content-scale item assessing depression. Content scales are typically derived via expert judgments, but from an item pool that has passed muster at some empirical level already.

The emphasis on inner psychological constructs typical of personality scales poses special problems for their development and validation. A reasonable treatment of these issues can be found in most basic psychological measurement texts (e.g., Anastasi & Urbina, 1997; Cronbach, 1983).

Objective personality scales are probably the most commonly used of all types of tests by clinical psychologists. They provide key evidence for the differentiation of various forms of psychopathology including clinical disorders and especially personality disorders. Omnibus scales such as the MMPI-2 and the Millon Clinical Multiaxial Inventory-3 are common with adult populations and also have adolescent versions. Omnibus scales directed at children and adolescents specifically however may be more appropriate for these younger age ranges. Among the many available, the Personality Inventory for Children and the Self-report of Personality from the Behavior Assessment System for Children (BASC: Reynolds & Kamphaus, 1992) are the most commonly used. Omnibus scales that are multidimensional in their construction are important to differential diagnosis in the early stages of contact with a patient. As diagnosis is established and treatment is in place, narrower scales that coincide with the clinical diagnosis become more useful for follow-up assessment and monitoring of treatment. In this regard, scales such as the Beck Depression Inventory or the State-Trait Anxiety Inventory are common examples. The tremendous cultural diversity of the world and how culture influences perceptions of items about the self and about one's family places special demands of cultural competence and cultural sensitivity on psychologists interpreting personality scales outside of their own cultural or ethnic heritage (e.g., see Dana, 1996; Hambleton, 1994; Moreland, 1995).

4.02.7.3 Projective Assessment

Projective assessment of personality has a long, rich, but very controversial history in the evaluation of clinical disorders and the description of normal personality. This controversy stems largely from the subjective nature of the tests used and the lack of good evidence of predictive validity coupled with sometimes fierce testimonial and anecdotal evidence of their utility in individual cases by devoted clinicians.

The subjectiveness of projective testing necessarily results in disagreement concerning the scoring and interpretation of responses to the test materials. For any given response by any given individual, competent professionals would each be likely to interpret differently the meaning and significance of the response. It is primarily the agreement on scoring that differentiates objective from subjective tests. If trained examiners agree on how a particular answer is scored, tests are considered objective; if not, they are considered subjective. Projective is not synonymous with subjective in this context but most projective tests are closer to the subjective than objective end of the continuum of agreement on scoring of responses. Projective tests are sets of ambiguous stimuli, such as ink blots or incomplete sentences, and the individual responds with the first thought or series of thoughts that come to mind or tells a story about each stimulus. Typically no restrictions are placed on individuals' response options. They may choose to respond with anything desired; in contrast, on an objective scale, individuals must choose between a set of answers provided by the test or at least listed out for the examiner in a scoring manual. The major hypothesis underlying projective testing is taken from Freud (Exner, 1976). When responding to an ambiguous stimulus, individuals are influenced by their needs, interests, and psychological organization and tend to respond in ways that reveal, to the trained observer, their motivations and true emotions, with little interference from the conscious control of the ego. Various psychodynamic theories are applied to evaluating test responses, however, and herein too lie problems of subjectivity. Depending on the theoretical orientation of the psychologist administering the test, very different interpretations may be given. Despite the controversy surrounding these tests, they remain very popular.

Projective methods can be divided roughly into three categories according to the type of stimulus presented and the method of response called for by the examiner. The first category calls for the interpretation of ambiguous visual stimuli by the patient with an oral response. Tests in this category include such well-known techniques as the Rorschach and the Thematic Apperception Test (TAT). The second category includes completion methods, whereby the patient is asked to finish a sentence when given an ambiguous stem or to complete a story begun by the examiner. This includes the Despert Fables and a number of sentence completion tests. The third category includes projective art, primarily drawing techniques, although sculpture and related art forms have been used. In these tasks, the child is provided with materials to complete an artwork (or simple drawing) and given instructions for a topic, some more specific than others. Techniques such as the Kinetic-Family-Drawing, the Draw-A-Person, and the Bender–Gestalt Test fall in this category.

Criterion-related and predictive validity have proven especially tricky for advocates of projective testing. Although techniques such as the TAT are not as amenable to study and validation through the application of traditional statistical and psychometric methods as objective tests may be, many clinical researchers have made such attempts with less than heartening results. None of the so-called objective scoring systems for projective devices has proved to be very valuable in predicting behavior, nor has the use of normative standards been especially fruitful. This should not be considered so surprising; however, it is indeed the nearly complete idiographic nature of projective techniques that can make them useful in the evaluation of a specific patient. It allows for any possible response to occur, without restriction, and can reveal many of a patient's current reasons for behaving in a specific manner. When used as part of a complete assessment, as defined in this chapter, projective techniques can be quite valuable. When applied rigidly and without proper knowledge and consideration of the patient's ecology, they can, as with other tests, be detrimental to our understanding of the patient. For a more extensive review of the debates over projective testing, the reader is referred to Exner (1976), Jackson and Messick (1967, Part 6), O'Leary and Johnson (1979), and Prevatt (in press), as well as to other chapters in this volume, especially Chapter 15.

4.02.7.4 Behavioral Assessment

The rapid growth of behavior therapy and applied behavior analysis has led to the need for tests that are consistent with the theoretical and practical requirements of these approaches to the modification of human behavior. Thus, the field of behavioral assessment has developed and grown at an intense pace. Book length treatments of the topic became commonplace in the 1970s and 1980s (e.g., Haynes & Wilson, 1979; Hersen & Bellack, 1976; Mash & Terdal, 1981) and entire journals are now devoted to research regarding behavioral assessment (e.g., *Behavioral Assessment*). The general term "behavioral assessment" has come to be used to describe a broad set of methods including some traditional objective personality scales, certain methods of interviewing, physiological response measures, naturalistic observation, norm-referenced behavior checklists, frequency counts of behavior, and a host of informal techniques requiring the observation of a behavior with recording of specific responses. Behavioral assessment will be discussed here in its more restricted sense to include the rating (by self or others) of observable behavioral events, primarily taking the form of behavior checklists and rating forms that may or may not be normed. Although I would certainly include psychophysiological assessment within this category, the scope of the work simply will not allow us to address this aspect of behavioral assessment except to say that it is indeed a most useful one in the treatment of a variety of clinical disorders.

The impetus for behavioral assessment comes not only from the field of behavior therapy but also from a general revolt against the high level of inference involved in such methods of assessing behavior as the Rorschach and the TAT. The greatest distinguishing characteristic between the behavioral assessment of psychopathological disorders and most other techniques is the level of inference involved in moving from the data provided by the assessment instrument to an accurate description of the patient and the development of an appropriate intervention strategy. This is a most useful strength for behavioral assessment strategies but is their greatest weakness when it is necessary to understand what underlies the observed behaviors.

Many of the early conceptual and methodological issues have been resolved in this area of assessment, for example, the importance of norms and other traditional psychometric concepts such as reliability and validity (Cone, 1977; Nelson, Hay, & Hay, 1977). Problems of interobserver reliability and observer drift remain but are well on their way to being resolved. Unquestionably, behavioral assessment is an exciting and valuable part of the assessment process. Behavioral assessment grew from a need to quantify observations of a patient's current behavior and its immediate antecedents and consequences, and this is the context within it that remains most useful today.

There are a number of formal behavior rating scales or behavior checklists now available. These instruments typically list behaviors of interest in clearly specified terms and have a trained observer or an informant indicate the frequency of occurrence of these behaviors. Interpretation can then take on a normative or a criterion-reference nature depending on the purpose of the assessment and the availability of norms. Clusters of behaviors may be of interest that define certain clinical syndromes such as attention deficit hyperactivity disorder. On the other hand, individual behaviors may be the focus (e.g., hitting other children). More frequently, behavioral assessment occurs as an "informal" method of collecting data on specific behaviors being exhibited by a patient and is dictated by the existing situation into which the psychologist is invited. An informal nature is dictated by the nature of behavioral assessment in many instances. Part of the low level of inference in behavioral assessment lies in not generalizing observations of behavior across settings without first collecting data in multiple settings. In this regard, behavioral assessment may for the most part be said to be psychosituational. Behavior is observed and evaluated under existing circumstances, and no attempt is made to infer that the observed behaviors will occur under other circumstances. Comprehensive systems that are multimethod, multidimensional, and that assess behavior in more than one setting have been devised and provide a good model for the future (Reynolds & Kamphaus, 1992).

Another area of assessment that stems from behavioral psychology and is considered by many to be a subset of behavioral assessment is task analysis. Whereas behavioral assessment typically finds its greatest applicability in dealing with emotional and behavioral difficulties, task analysis is most useful in evaluating and correcting learning problems. In task analysis, the task to be learned (e.g., brushing one's teeth or multiplying two-digit numbers) is broken down into its most basic component parts. The learner's skill at each component is observed and those skills specifically lacking are targeted for teaching to the child. In some cases, hierarchies of subskills can be developed, but these have not held up well under cross-validation. Task analysis can thus be a powerful tool in specifying a learner's existing (and needed) skills for a given learning problem. Task analysis could, for example, form an integral part of any behavioral intervention for a child with specific learning problems. The proper use of these procedures requires a creative and well-trained individual conversant with both assessment technology and behavioral theories of learning, since there are no standardized task analysis procedures. Those involved in task analysis need to be sensitive to the reliability and validity of their methods. As with other behavioral assessment techniques, some contend that behavioral assessment techniques need only demonstrate that multiple observers can agree on a description of the behavior and when it has been observed. Though not having to demonstrate a relationship with a hypothetical construct, behavioral techniques must demonstrate that the behavior observed is consistent and relevant to the learning problems. For behavior checklists and more formal behavioral assessment techniques, most traditional psychometric concepts apply and must be evaluated with regard to the behavioral scale in question.

4.02.7.5 Neuropsychological Assessment

Along with behavioral assessment, perhaps the most rapidly growing area in patient evaluation is neuropsychological assessment. Many view neuropsychological assessment as the application of a specific set of tests or battery of tests. Far from being a set of techniques, the major contribution of neuropsychology to the assessment process is the provision of a strong paradigm from which to view assessment data (Reynolds, 1981b, 1981c, 1997). Without a strong theoretical guide to test score interpretation, one quickly comes to rely upon past experience and illusory relationships and trial and error procedures when encountering a patient with unique test performance. As with most areas of psychology, there are competing neuropsychological models of cognitive functioning, any one of which may be most appropriate for a given patient. Thus considerable knowledge of neuropsychological theory is required to evaluate properly the results of neuropsychological testing.

Since the 1950s, clinical testing in neuropsychology has been dominated by the Halstead–Reitan Neuropsychological Test Battery (HRNTB), although the Luria–Nebraska Neuropsychological Battery and the Boston process approach have made significant inroads. The prevalence of use of the HRNTB is partly responsible for perceptions of clinical neuropsychology as primarily a set of testing techniques. However, a brief examination of the HRNTB should quickly dispel such ideas. The HRNTB consists of a large battery of tests taking a full day to administer. There is little that can be said to be psychologically or psychometrically unique about any of these tests. They are all more or less similar to tests that psychologists have been using for the past 50 years. The HRNTB also typically includes a traditional intelligence test such as one of the Wechsler scales. The HRNTB is unique in the particular collection of tests involved and the method of evaluating and interpreting performance. While supported by actuarial studies, HRNTB performance is evaluated by the clinician in light of existing neuropsychological theories of cognitive function, giving the battery considerable explanatory and predictive power.

Neuropsychological approaches to clinical assessment are rapidly growing and can be most helpful in defining areas of cognitive-neuropsychological integrity and not just in evaluating deficits in neurological function. Neuropsychological techniques can also make an important contribution by ruling out specific neurological problems and pointing toward environmental determinants of behavior. The well-trained neuropsychologist is aware that the brain does not operate in a vacuum but is an integral part of the ecosystem of the patient. As with other methods of assessment, neuropsychological assessment has much to offer the assessment process when used wisely; poorly or carelessly implemented, it can create seriously false impressions, lessen expectations, and precipitate a disastrous state of affairs for the patient it is designed to serve. Clinicians who use neuropsychological approaches to their work or make neuropsychological interpretations of tests or behaviors are in high demand but require specialized training that takes years to acquire.

4.02.8 CLINICAL VS. STATISTICAL PREDICTION

Given a set of test results and/or additional interview or historical data on a patient, there are two fundamental approaches a clinician can apply to diagnostic decision-making or to the prediction of future behavior. The first, and likely most common, is the simple human judgment of the clinician who reviews the data in conjunction with prior experiences and knowledge of diagnostic criteria, psychopathology, and other psychological literature. As a result of the application of this *clinical method*, a diagnosis or prediction is made. Alternatively, the clinician may apply a formal algorithm or set of mathematical equations to predict membership in a class (a diagnosis) or the probability of some future behavior (e.g., whether a defendant will or will not molest children again if placed on probation). The use of such mechanistic approaches constitutes the *statistical method.*

The ability of clinicians to use test data in optimal ways and to combine multiple sources of data optimally to arrive at correct diagnoses and to make predictions about future behavior has been the subject of much scrutiny. Meehl (1954) addressed this problem early on and concluded that formula-based techniques, derived by mathematical models that search for optimal combinations of data, are superior to clinicians in decision-making. This has been difficult for clinicians to accept even as more and more actuarial systems for test interpretation find their way into our office computers. In 70+ years of research on this topic, actuarial modeling continues to be superior (Faust & Ackley, 1998; Grove & Meehl, 1996), yet I listened to a clinical psychologist testify in February of 1998 that clinical judgment was always better and that actuarial predictions were used only when you had nothing else to use. In 136 studies since 1928, over a wide range of predictions, actuarial modeling is invariably equal to or superior to clinical methods (Grove & Meehl, 1996). Grove and Meehl (1996) have addressed this reluctance (or perhaps ignorance) about actuarial modeling in clinical-decision-making and clinicians' objections to acceptance of more than 70 years of consistent research findings, as has Kleinmuntz (1990), who did seminal research on developing expert systems for MMPI interpretation in the 1960s.

Grove and Meehl (1996) state the most common objection of clinicians to statistical modeling is that they (the clinicians) use a combination of clinical and statistical methods, obviating the issue since they are used in a complementary model. This is a spurious argument because, as research shows, the clinical method and the statistical method often disagree and there is no known way to combine the two methods to improve predictions; we simply do not know under what conditions to conclude the statistical model will be wrong (also see Faust & Ackley, 1998, and Reynolds, 1998a). Grove and Meehl (1996) illustrate this quandary by examining the actions of an admissions committee. Suppose the applicant's test scores, grade point average, and rank in class predict successful college performance but the admissions commission believes, perhaps based on an interview and letters of recommendation, that the applicant will not be successful. The two methods cannot be combined when they specify different outcomes. One is right and one is wrong. Consider a crucial prediction in a forensic case. A psychologist treating an offender on parole for aggravated sexual assault of a child is asked whether the offender might molest a child of the same age and gender as a prior victim if the child is placed in the offender's home (he recently married the child's mother). Actuarial tables with rearrest rates for offenders with many similar salient characteristics and in treatment indicate that 10–11% of those individuals will be arrested again for the same offense. The psychologist concludes that the offender is virtually a zero percent risk, however, because "he has worked so hard in treatment." The two methods again cannot be resolved; yet, as clinicians we persist in believing we can do better than the statistical models even in the face of a substantial body of contradictory evidence. Grove and Meehl (1996) review some 16 other objections clinicians make to statistical methods in diagnosis and other clinical predictions. Most resolve to questions of validity, reliability, or cost. Cost and inconveniences are rapidly becoming a nonissue as computerized models are widely available and some of these cost less than one dollar per application for a patient (e.g., Stanton, Reynolds, & Kamphaus, 1993).

Statistical models work better, consistently, and for known reasons. Clinicians cannot assign optimal weights to variables they use to make decisions, do not apply their own rules consistently, and are influenced by relatively unreliable data (e.g., see Dawes, 1988; Faust, 1984; Grove & Meehl, 1996; Meehl, 1954). As the reliability of a predictor goes down, the relative accuracy of any prediction will be reduced and, consequently, the probability of being wrong increases. The decisions being made are not trivial. Every day, thousands of diagnostic decisions are being made that influence treatment choices, freedom for parolees and probationers, large monetary awards in personal injury cases, custody of children, and others. What clinical psychologists do is important and the increased use of statistical models based on sound principles of data collection that includes data from well-standardized, objective psychological tests seems imperative from an ethical standpoint and from the standpoint of the survival of the profession as accountability models are increasingly applied to reviews of the need for and effectiveness of our services.

4.02.9 ACCESSING CRITICAL COMMENTARY ON STANDARDIZED PSYCHOLOGICAL TESTS

Not every practitioner can nor should be an expert on the technical adequacy and measurement science underlying each and every standardized test that might be useful to their clinical practice. With a clear understanding of the fundamentals of measurement, however, clinicians can make intelligent choices about test selection and test use based on the test manuals and accompanying materials in most cases. However, when additional expertise or commentary is required, critical reviews of nearly every published test can be accessed with relative ease.

Many journals in psychology routinely publish reviews of new or newly revised tests including such publications as the *Archives of Clinical Neuropsychology, Journal of Psychoeducational Assessment, Journal of Learning Disabilities, Journal of School Psychology, Child Assessment News,* and the *Journal of Personality Assessment*. However, the premier source of critical information on published tests are the publications of the Buros Institute of Mental Measurement.

In the late 1920s, Oscar Krisen Buros began to publish a series of monographs reviewing statistics texts. He noted the rapid proliferation of psychological and educational tests beginning to occur at the same time and rapidly turned his attention to obtaining high-quality reviews of these tests and publishing them in bound volumes known as the *Mental Measurements Yearbook* (MMY). The first MMY was published by Buros in 1938 and the series continues today. Buros died in 1978, during the final stages of production of the Eighth MMY (though "Yearbooks," they are not published annually), and his spouse, art director, and assistant Luella Buros saw the Eighth MMY to completion. Subsequently, she opened a competition for proposals to adopt the Institute and continue the work of her late husband. A proposal written by this author (then on the faculty of the University of Nebraska-Lincoln) was chosen and the Buros Institute of Mental Measurement was established in 1979 at the University of Nebraska-Lincoln, where it remains permanently due to a generous endowment from Luella Buros.

The Institute continues to seek out competent reviewers to evaluate and provide critical commentary on all educational and psychological tests published in the English language. These reviews are collected in an ongoing process, as tests are published or revised, under a strict set of rules designed to ensure fair reviews and avoid conflicts of interest. The collected reviews are published on an unscheduled basis approximately every five to eight years. However, as reviews are written and accepted for publication, they are added quickly to the Buros Institute database which may be searched on-line by subscription to the master database or through most major university libraries. Information on how to access current reviews in the Buros database can be obtained through nearly any reference librarian or through a visit to the Buros Institute website. The Institute established a sterling reputation as the "consumer reports" of the psychological testing industry under the 50 year leadership of Oscar Buros and this reputation and service have been continued at the University of Nebraska-Lincoln. Consumers of tests are encouraged to read the Buros reviews of the tests they choose to use.

4.02.10 CONCLUDING REMARKS

Knowledge of measurement science and psychometrics is critical to good practice in clinical assessment. Hopefully, this review, targeted at practice, has provided a better foundation for reading about psychological testing and for developing better skills in the application of methods of testing and assessment. Old tests continue to be revised and updated, and many new tests are published yearly. There is a large, rapidly growing body of literature on test interpretation that is too often ignored in practice (e.g., Reynolds & Whitaker, in press) but must be accessed in practice. It is difficult but necessary to do so. Measurement science also progresses and practitioners are encouraged to revisit their mathematical backgrounds every few years as part of what has become a nearly universal requirement for continuing education to continue in practice. New paradigms will always emerge as they have in the past.

It is from basic psychological research into the nature of human information processing and behavior that advances in psychological assessment must come. While some of these advances will be technological, the more fruitful area for movement is in the development of new paradigms of test interpretations. With each advance, with each "new" test that appears, we must proceed with caution and guard against jumping on an insufficiently researched bandwagon. Fruitful techniques may be lost if implemented too soon to be fully understood and appreciated; patients may also be harmed by the careless or impulsive use of assessment materials that are poorly designed (but attractively packaged) or without the necessary theoretical and empirical grounding. When evaluating new psychological assessment methods, surely *caveat emptor* must serve as the guard over our enthusiasm and our eagerness to provide helpful information about patients in the design of successful intervention programs.

4.02.11 REFERENCES

American Psychological Association (1985). *Standards for educational and psychological tests*. Washington, DC: Author.

Anastasi, A. (1976). *Psychological testing* (4th ed.). New York: Macmillan.

Anastasi, A. (1981). *Psychological testing* (5th ed.). New York: Macmillan.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed.). Washington, DC: American Council on Education.

Barrios, B. A., Hartmann, D. P., & Shigetomi, C. (1981). Fears and anxieties in children. In E. J. Marsh & L. G. Terdal (Eds.), *Behavioral assessment of childhood disorders*. New York: Guilford.

Cone, J. D. (1977). The relevance of reliability and validity for behavioral assessment. *Behavior Therapy, 8,* 411–426.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed.). Washington, DC: American Council on Education.

Cronbach, L. J. (1983). *Essentials of psychological testing* (4th ed.). New York: Harper & Row.

Cronbach, L. J., Rajaratnam, N., & Gleser, G. C. (1963). Theory of generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology, 16,* 137–163.

Dana, R. H. (1996). Culturally competent assessment practices in the United States. *Journal of Personality Assessment, 66,* 472–487.

Dawes, R. M. (1988). *Rational choice in an uncertain world*. Chicago, IL: Harcourt Brace Jovanovich.

Ebel, R. L. (1972). *Essentials of educational measurement*. Englewood Cliffs, NJ: Prentice-Hall.

Exner, J. E. (1976). Projective techniques. In I. B. Weiner (Ed.), *Clinical methods in psychology*. New York: Wiley.

Faust, D. (1984). *The limits of scientific reasoning*. Minneapolis, MN: University of Minnesota Press.

Faust, D., & Ackley, M. A. (1998). Did you think it was going to be easy? Some methodological suggestions for the investigation and development of malingering detection techniques. In C. R. Reynolds (Ed.), *Detection of malingering during head injury litigation* (pp. 1–54). New York: Plenum.

Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. Linn (Ed.), *Educational measurement*, (3rd ed., pp. 105–146). New York: Macmillan.

Grove, W. M., & Meehl, P. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychology, Public Policy, and Law, 2,* 293–323.

Hambleton, R. K. (1994). Guidance for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment, 10,* 229–244.

Haynes, S. N., & Wilson, C. C. (1979). *Behavioral assessment*. San Francisco: Jossey-Bass.

Hays, W. L. (1973) *Statistics for the social sciences*. New York: Holt, Rinehart & Winston.

Hersen, M., & Bellack, A. S. (1976). *Behavioral assessment: A practical handbook*. New York: Pergamon.

Hunter, J. E., Schmidt, F. L., & Rauschenberger, J. (1984). Methodological and statistical issues in the study of bias in mental testing. In C. R. Reynolds & R. T. Brown (Eds.), *Perspectives on bias in mental testing* (pp. 41–99) New York: Plenum.

Jackson, D. N., & Messick, S. (Eds.) (1967). *Problems in human assessment*. New York: McGraw-Hill.

Jean, P. J., & Reynolds, C. R. (1982). *Sex and attitude distortions: The faking of liberal and traditional attitudes about changing sex roles*. Paper presented to the annual meeting of the American Educational Research Association, New York, March.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Kaufman, A. S. (1990). *Assessment of adolescent and adult intelligence*. Boston: Allyn & Bacon.

Kaufman, A. S. (1994) *Intelligent testing with the WISC-III*. New York: Wiley.

Kaufman, A. S., & Reynolds, C. R. (1983). Clinical evaluation of intellectual function. In I. Weiner (Ed.), *Clinical methods in psychology* (2nd ed.). New York: Wiley.

Kleinmuntz, B. (1990). Why we still use our heads instead of the formulas: Toward an integrative approach. *Psychological Bulletin, 107,* 296–310.

Lord, F. M., & Novick, M. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Mash, E. J., & Terdal, L. G. (1981). *Behavioral assessment of childhood disorders*. New York: Guilford.

Meehl, P. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis, MN: University of Minnesota Press.

Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist, 35,* 1012–1027.

Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–104). New York: Macmillan.

Moreland, K. L. (1995). Persistent issues in multicultural assessment of social and emotional functioning. In L. A. Suzuki, P. J. Meller, & J. G. Ponterrotto (Eds.), *Handbook of multicultural Assessment: Clinical, psycho-*

logical, and educational applications. San Francisco: Jossey-Bass.

Nelson, R. O., Hay, L. R., & Hay, W. M. (1977). Comments on Cone's "The relevance of reliability and validity for behavioral assessment." *Behavior Therapy, 8,* 427–430.

Newland, T. E. (1980). Psychological assessment of exceptional children and youth. In W. M. Cruickshank (Ed.), *Psychology of exceptional children and youth* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Nunnally, J. (1978). *Psychometric theory*. New York: McGraw-Hill.

O'Leary, K. D., & Johnson, S. B. (1979). Psychological assessment. In H. C. Quay & J. S. Werry (Eds.), *Psychopathological disorders of childhood.* New York: Wiley.

Palmer, D. J. (1980). Factors to be considered in placing handicapped children in regular classes. *Journal of School Psychology, 18,* 163–171.

Petersen, N. S., Kolen, M., & Hoover, H. D. (1989). Scaling, norming, and equating. In R. Linn (Ed.) *Educational measurement* (3rd ed., pp. 221–262). New York: Macmillan.

Prevatt, F. (in press). Personality assessment in the schools. In C. R. Reynolds & T. B. Gutkin (Eds.), *The handbook of school psychology* (3rd ed.). New York: Wiley.

Reynolds, C. R. (1981a). The fallacy of "two years below grade level for age" as a diagnostic criterion for reading disorders. *Journal of School Psychology, 19,* 350–358.

Reynolds, C. R. (1981b). Neuropsychological assessment and the habilitation of learning: Considerations in the search for the aptitude X treatment interaction. *School Psychology Review, 10,* 343–349.

Reynolds, C. R. (1981c). The neuropsychological basis of intelligence. In G. Hynd & J. Obrzut (Eds.), *Neuropsychological assessment of the school-aged child.* New York: Grune and Stratton.

Reynolds, C. R. (1982). The problem of bias in psychological assessment. In C. R. Reynolds & T. B. Gutkin (Eds.), *The handbook of school psychology*. New York: Wiley.

Reynolds, C. R. (1985). Critical measurement issues in learning disabilities. *Journal of Special Education, 18,* 451–476.

Reynolds, C. R. (1997). Measurement and statistical problems in neuropsychological assessment of children. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *The handbook of clinical child neuropsychology* (2nd ed., pp. 182–203). New York: Plenum.

Reynolds, C. R. (1998a). Common sense, clinicians, and actuaralism in the detection of malingering during head injury litigation. In C. R. Reynolds (Ed.), *Detection of malingering during head injury litigation* (pp. 261–282). New York: Plenum.

Reynolds, C. R. (Ed.) (1998b). *Detection of malingering during head injury litigation.* New York: Plenum.

Reynolds, C. R. (in press-a). Need we measure anxiety separately for males and females? *Journal of Personality Assessment.*

Reynolds, C. R. (in press-b). Why is psychometric research on bias in mental testing so often ignored? *Psychology, Public Policy and Law.*

Reynolds, C. R. & Brown, R. T. (1984). Bias in mental testing: An introduction to the issues. In C. R. Reynolds & R. T. Brown (Eds.), *Perspectives on bias in mental testing*. New York: Plenum.

Reynolds, C. R., & Clark, J. (1982). Cognitive assessment of the preschool child. In K. D. Paget & B. Bracken (Eds.), *Psychoeducational assessment of the preschool and primary aged child.* New York: Grune & Stratton.

Reynolds, C. R., & Kamphaus, R. W. (Eds.) (1990a). *Handbook of psychological and educational assessment of children: Vol I. Intelligence and achievement*. New York: Guilford.

Reynolds, C. R., & Kamphaus, R. W. (Eds.) (1990b). *Handbook of psychological and educational assessment of children: Vol II. Personality, behavior, and context.* New York: Guilford.

Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior assessment system for children.* Circle Pines, MN: American Guidance Service.

Reynolds, C. R., & Richmond, B. O. (1978). What I think and feel: A revised measure of children's manifest anxiety. *Journal of Abnormal Child Psychology, 6,* 271–280.

Reynolds, C. R., & Richmond, B. O. (1985). *Revised children's manifest anxiety scale*. Los Angeles: Western Psychological Services,

Reynolds, C. R., & Whitaker, J. S. (in press). Bias in mental testing since Jensen's "Bias in mental testing." *School Psychology Quarterly.*

Salvia, J., & Ysseldyke, J. E. (1981). *Assessment in special and remedial education* (2nd ed.). Boston: Houghton Mifflin.

Sattler, J. (1988). *Assessment of children's intelligence and special aptitudes*. San Diego, CA: Author.

Stanton, H., Reynolds, C. R., & Kamphaus, R. W. (1993). *BASC plus software scoring program for the Behavior Assessment System for Children.* Circle Pines, MN: American Guidance Service.

Weschsler, D. (1992). *Weschsler Intelligence Scale for Children-III.* San Antonio, TX: Psychological Corporation.

Whyte, W. H. (1967). How to cheat on personality tests. In D. Jackson & S. Messick (Eds.), *Problems in human assessment*. New York: McGraw-Hill.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.03
Diagnostic Models and Systems

ROGER K. BLASHFIELD
Auburn University, AL, USA

4.03.1 INTRODUCTION 57
4.03.2 PURPOSES OF CLASSIFICATION 58
4.03.3 DEVELOPMENT OF CLASSIFICATION SYSTEMS IN THE USA 59
4.03.4 KRAEPLIN 62
4.03.5 EARLY *DSMS* AND *ICDS* 62
4.03.6 NEO-KRAEPELINIANS 63
4.03.7 *DSM-III, DSM-III-R,* AND *DSM-IV* 64
4.03.8 *ICD-9, ICD-9-CM,* AND *ICD-10* 67
4.03.9 CONTROVERSIES AND ISSUES 69
  4.03.9.1 Organizational Models of Classification 69
  4.03.9.2 Concept of Disease 72
  4.03.9.3 Two Views of a Hierarchical System of Diagnostic Concepts 75
  4.03.9.4 Problem of Diagnostic Overlap/Comorbidity 76
4.03.10 CONCLUDING COMMENTS 78
4.03.11 REFERENCES 78

4.03.1 INTRODUCTION

> To all the People and Inhabitants of the United States and all the outlying Countries, Greetings: I, John Michler, King of Tuskaroras, and of all the Islands of the Sea, and of the Mountains and Valleys and Deserts; Emperor of the Diamond Caverns, and Lord High General of the Armies thereof; First Archduke of the Beautiful Isles of the Emerald Sea; Lord High Priest of the Grand Lama, etc., etc., etc.: Do issue this my proclamation. Stand by and hear, for the Lord High Shepherd speaks. No sheep have I to lead me around, no man have I to till me the ground, but the sweet little cottage is all of my store, and my neat little cottage has ground for the floor. No children have I to play me around, no dog have I to bark me around, but the three-legged stool is the chief of my store, and my neat little cottage has ground for the floor. Yea, verily, I am the Mighty King, Lord Archduke, Pope, and Grand Sanhedrim, John Michler. None with me compare, none fit to comb my hair, but the three-legged stool is the chief of my store, and my neat little cottage has ground for the floor. John Michler is my name. Selah!
> I am the Great Hell-Bending Rip-Roaring Chief of the Aborigines! Hear me and obey! My breath overthrows mountains; my mighty arms crush the everlasting forests into kindling wood; I am the owner of the Ebony Plantations; I am the owner of all the mahogany groves and of all the satin-wood; I am the owner of all the granite; I am the owner of all the marble; I am the owner of all the owners of Everything. Hear me and obey! I, John Michler, stand forth in the presence of the Sun and of all the Lord Suns and Lord Planets of the Universe, and I say, Hear me and obey! I, John Michler, on this eighteenth day of August, 1881, do say, Hear me and obey! for with me none can equal, no, not one, for the three-legged stool is the chief of my store,

> and my neat little cottage has ground for the floor. Hear me and obey! Hear me and obey! John Michler is my name.
> John Michler, First Consul and Dictator of the World, Emperor, Pope, King and Lord High Admiral, Grand Liconthropon forever! (Hammond, 1883, pp. 603–604)

A physician in private practice in New York City reported that a man brought his brother, John Michler, to see him. John was acting strangely, and his brother wanted to know what to do. The brother gave the physician a proclamation that John Michler had written (see above). Clearly, to most observers, there would be no question that John Michler was "crazy." However, what is the diagnosis of John Michler?

When this proclamation was shown to mental health professionals, the most common diagnostic possibilities that are mentioned are schizophrenia and bipolar disorder (manic episode). What do these various diagnoses mean? Why did clinicians not assign a diagnosis of narcissistic personality disorder to this patient? Certainly this man would fit the vernacular meaning of self-centered and self-aggrandizing that is often associated with a narcissistic personality. How is a manic episode differentiated from schizophrenia? What does it mean to say that Michler appears to be psychotic? Does that diagnosis mean that he has some type of disease that has affected part of his brain, or does it suggest that his childhood was so unusual and abnormal that he has developed a strange and unusual way of coping with life?

4.03.2 PURPOSES OF CLASSIFICATION

The vernacular word that was used to describe John Michler was "crazy." This word is frequently used in descriptions of persons who have mental disorders. The reason for the applicability of this word is that one common feature of most psychiatric patients is that their behaviors are statistically abnormal. That is, psychiatric patients behave in ways that are deviant. Their interpersonal actions are not expected within the cultural context of our society.

Classification is a fundamental human activity that people use to understand their world. For instance, a classification of animals is helpful when understanding the variations among diverse forms of living organisms. In forming a classification, the classifier is attempting to use observed similarities and differences among the things being classified, so as to find some order among those things. In psychopathology, the general goal of classification is an attempt to use similarities and differences among people who behave in deviant and socially abnormal ways in order to understand their behaviors.

More specifically, there are five purposes to the classification of mental disorders: (i) forming a nomenclature so that mental health professionals have a common language; (ii) serving as a basis of information retrieval; (iii) providing a short-hand description of the clinical picture of the patient; (iv) stimulating useful predictions about what treatment approach will be best; and (v) serving as a concept formation system for a theory (or theories) of psychopathology.

The first reason to have a classification system, providing a nomenclature, is the most fundamental (World Health Organization, 1957). At a minimum, a classification system provides a set of nouns for clinicians to use to discuss their objects of interest—people. Thus, a nomenclature is a set of terms that refer to groups of people that mental health professionals see in their various professional roles.

The second reason, information retrieval, has a pragmatic focus on how well a classification organizes a research literature, so that clinicians and scientists can search for information that they need (Feinstein, 1967). In biology, there is a dictum: "the name of a plant is the key to its literature" (Sneath & Sokal, 1973). The same is true in the area of mental disorders. If a student clinician is assigned a patient who attempts to control weight by inducing vomiting, the name bulimia becomes important for helping the student locate the literature about this disorder in books, journal articles, and even on the internet.

The third reason for having a classification is description. There are many ways of creating classifications that could satisfy the first two purposes. For instance, clinicians could decide to classify all of their patients on eye color. Using eye color would allow clinicians to have a language to discuss patients ("I have seen 17 brown eyed, eight blue eyed, and four mixed eye color patients in the last month."). Also, eye color categories could be used as names to store information about patients. However, using eye color to classify patients would not be a satisfactory solution to either researchers or clinicians because patients with similar eye colors are unlikely to have similar symptom patterns. To meet the purpose of description, patients are grouped on the basis of similarity. Patients who have the same diagnosis are expected to be relatively similar in terms of symptoms (Lorr, 1966). In addition, these patients should be dissimilar when compared to patients with different diagnoses (Everitt, 1974). In the case of John Michler, if he is having a manic episode of a bipolar disorder, then we would expect that Michler's brother would report that John had been spending large sums of money that he did not have, that his speech was extremely rapid, and that his sleep pattern was markedly disturbed. In contrast, his brother is unlikely to report that Michler usually sat around the house with an emotionless, cold, detached interpersonal style and that he would tell his brother about voices in his head that were in discussion about Michler's behaviors. The latter symptoms typically occur in individuals with schizophrenia. Thus, diagnoses become descriptive short-hand names for clusters of co-occurring symptoms.

The fourth purpose is prediction. This purpose typically involves two types of inferences: (i) predicting the course of the patient's condition if there is no treatment or intervention ("Diagnosis is prognosis" as stated by Woodruff, Goodwin, & Guze, 1974); and (ii) predicting which treatment approach will be most effective with the patient (Roth & Fonagy, 1996). In the field of psychopathology, prediction has proved to be an elusive goal. Recently, there was an important multisite study that was performed in the USA in which three different treatment approaches to alcoholism were compared. An attempt was made to see whether particular groups of patients improved with specific treatments. The initial results have been negative. Except for differences related to the severity of patient symptomatology, other patient characteristics did not predict which treatments were most effective (Project Match Research Group, 1997).

The final goal of a classification is concept formation (Hempel, 1965). This goal is probably the most distant. In biological classification, Linneaus and his contemporaries made noteable gains in the classification of living organisms by creating classifications that served to describe most of the known information about these organisms. Almost a century later, Darwin formulated a theory of evolution which explained many of the organizational phenomena that Linneaus had used (Hull, 1988). In particular, Darwin's evolutionary theory provided a basis for understanding why there is a hierarchical arrangement about the categories in biological classification. The field of biological classification has continued to change. During the twentieth century, three different, competing approaches to biological classification have appeared (Hull, 1988). In oversimplified terms, one of these approaches emphasized the nomenclature and information retrieval purposes, the second focused on description, and the third was based on a theoretical view. The third view, the one based on theory, has become the dominant approach to biological classification (Nelson & Platnick, 1981).

4.03.3 DEVELOPMENT OF CLASSIFICATION SYSTEMS IN THE USA

The classification of mental disorders has an extensive history that can be traced back to some of the earliest writings known to man. A nineteenth century BC Egyptian writer discussed a disorder in women in which they would report vague and inconsistent physical symptoms that appeared to shift in body location over time (Veith, 1965). Psalm 102 provides an excellent clinical description of depression.

However, like many others areas of modern science, the first major commentaries on mental disorders were found in the writings of the Greeks. The Greek medical writers introduced four terms, all of which are still used today: "melancholia," "hysteria," "mania," and "paranoia." Melancholia was a Greek term that referred to a condition that now would be described by the word depression. Hippocrates believed that the sadness and the slowed bodily movements associated with this disorder were caused by an abundance of black bile, which he considered to be one of the four main ingredients in the human body. Hence, he named this disorder *melan* (black) + *cholia* (bile). The second term, hysteria, was the Greek word for the condition originally described by the Egyptians in which women had multiple, inconsistent and changing somatic complaints. The Hippocratic writers used the name of hysteria, which means pertaining to the uterus, because they believed that this disorder was caused by a dislodged, floating uterus. The last two terms were mania and paranoia. Mania, to the Greeks, referred to persons who were delusional. During the nineteenth century, individuals who were delusional but had few other symptoms were diagnosed with monomania (Berrios, 1993). Mania became an umbrella term for almost any type of psychotic state (Spitzka, 1883). The meaning of mania, however, changed again in the twentieth century to its contemporary denotation of grandiosity, excitement, expansiveness, and elation. The final Greek term, paranoia, has undergone a similar transformation. Paranoia initially meant craziness (*para* = abnormal + *nous* = mind). Now the term refers to people who are suspicious and often have delusions in which others are plotting against them.

After the Greeks, psychopathology did not attract much scientific interest until the nineteenth century. During the Middle Ages, mental disorders were associated with evilness. Thus, mental disorders were under the domain of religious authorities, not physicians or scientists. This approach to mental disorders began to change in the late 1700s as exemplified by the fact that King George III of England, who was psychotic for most of the last decade of his reign, received medical care rather than religious counseling.

The first major American physician to be interested in mental disorders was Benjamin Rush who, as a signer of the Declaration of Independence, was one of the prominent American physicians of the late eighteenth century. He was very interested in the forms of insanity. Rush also published a book on the topic which he titled *Medical inquires and observations upon the diseases of the mind* (Alexander & Selesnick, 1966). He believed in a theory of neurosis. According to this theory, mental disorders were caused by overstimulation of the nervous system. Thus, environmental phenomena such as the pace of urban living, overuse of alcohol, excessive sexual behavior, masturbation, and smoking were all seen as casual factors in the development of mental disorders. As a result, asylums were the appropriate way to treat insane patients. Asylums could provide the quiet and tranquility that was necessary to allow the nervous system to heal and to repair itself.

About the same time that Rush was writing about psychopathology in the USA, there was an important discovery in France that was to markedly influence thinking about mental disorders. In 1822, a physician named Bayle (Quetel, 1990) performed autopsies on a number of patients who presented with gradiose delusions and dementia (i.e., who had lost their mind from the French *de-* (not) + *ment* (mind)). Bayle discovered that all of the patients in his study had marked changes in their brains. In addition to their dementia, all of these patients developed motor paralysis before they died. The brains of these patients had shrunk to almost half the weight of a normal brain, the skin of the brain (i.e., the meninges and the arachnoid) was thickened, and the color of the brain was strikingly different from that of normal brains. Bayle's name for this disorder was chronic arachnitis since he believed that this disorder was caused by a chronic infection of arachnoid tissue (Quetel, 1990). Later, the common name for this disorder was changed to dementia paralytica, a descriptive name that referred to the combined presence of a psychosis together with the progressive paralysis of the patient's limbs.

The discovery of dementia paralytica was the first instance in which a mental disorder had been shown to be associated with demonstrable changes in the brain. A number of autopsy studies appeared in the medical journals. These studies attempted to identify the exact neuropathology of this disorder as well as to understand its cause. Dementia paralytica was also a clinically serious disorder because it accounted for about one-sixth of all admissions to insane asylums during the nineteenth century. The prognosis of the disorder was very poor because death typically would occur within three years of the diagnosis (Austin, 1859/1976). For most of the nineteenth century, many different etiologies were proposed as potential causes of this disorder. Austin (1859/1976), for instance, listed the following as moral causes of dementia paralytica: death of a son, sudden loss of two good situations, wife's insanity, worry, and commercial ruin.

With the increasing interest in psychopathology during the nineteenth century, a number of classifications for mental disorders appeared. One example of these classifications was published by William A. Hammond (Blustein, 1991). Hammond, like Freud, was a nineteenth century neurologist. As a young physician, Hammond had published a set of interesting experimental studies on human physiology. At the age of 34, he became Surgeon General for the US Army during the Civil War and was credited with a number of important innovations at the time including hospital design, the development of an ambulance corps, and the removal of a mercury compound from the medical formulary. His political clashes with the Secretary of War, however, led to his court martial and dismissal. After the Civil War, he moved to New York City and developed a lucrative private practice as a neurologist—a remarkable accomplishment during a time when most physicians were generalists. His interests extended to psychiatry and to writing novels as well as to physiology, studies of sleep, hypnosis, and the use of animal hormonal extracts. Hammond wrote extensively in scientific journals. He was one of the founders of the *Journal of Nervous and Mental Disease* which is still published. In addition, he authored important textbooks of neurology and psychiatry.

In Hammond's textbook of mental disorders, he argued that there were six possible principles that could be used to organize a classification system:

(i) anatomical (organized by the part of the brain that is affected);

(ii) physiological (organized by the physiological system in the brain);

(iii) etiological (supposed causes);

(iv) psychological (based upon a functional view of the mind);

(v) pathological (observable, morbid alterations in the brain); and

(vi) clinical (descriptive, based upon clusters of symptoms).

Of these six principles, Hammond said that the anatomical, the physiological, and the pathological are the best, but he could not use them because the science of his time was insufficient. Hammond also rejected the etiological organization of categories, because he felt that an etiological classification, given nineteenth century knowledge, would be too speculative. Thus, the main choice was between the clinical (descriptive) approach and the psychological (mentalistic) approach. Hammond preferred the latter because he thought that a classification which did not have a strong theoretical basis would fail.

Hammond adopted a functional view of psychology that was common in his day. He believed that mental functioning could be organized into four areas: perception, cognition, affect, and will. Hence, he organized his classification of mental disorders into six major headings:

(i) perceptual insanities (e.g., hallucinations);

(ii) intellectual insanities (e.g., delusional disorders);

(iii) emotional insanities (e.g., melancholia);

(iv) volitional insanities (e.g., abulomania);

(v) compound insanities (i.e., disorders affecting more than one area of the mind); and

(vi) constitutional insanities (i.e., disorders with specific causes such as choreic insanity).

There were a total of 47 specific categories of mental disorders that were organized under these six major headings. Most names of these specific categories would not be recognized by modern clinicians. The descriptions of these disorders, together with case histories that he included in this textbook do, however, suggest that many of the disorders he was discussing would have modern counterparts. For instance, under the heading "intellectual insanities," Hammond classified four disorders whose names seem odd by modern standards: intellectual monomania with exaltation, chronic intellectual monomania, reasoning mania, and intellectual subjective morbid impulses. In modern terms, these disorders probably would be called biopolar I disorder (manic episode), schizophrenia (continuous), narcissistic personality disorder, and obsessive compulsive disorder.

In Hammond's textbook, the lengthiest discussion was devoted to general paralysis, for which Hammond's name was dementia paralytica. As part of this discussion, Hammond included the proclamation by John Michler (quoted at the beginning of this chapter). In his discussion of general paralysis, Hammond emphasized the many medical symptoms associated with this disorder.

By the end of the nineteenth century, there were three broad theories about the etiology of this disorder. First, one school of thought believed that it was caused by alcoholism because the disorder primarily affected men, the age of onset was typically during the 30s and 40s (which is the same time of onset for severe alcoholism), and most men with dementia paralytica had substantial drinking histories. Second, was the theory that the disorder was caused by syphilis. Epidemiological surveys had found that over 80% of men with dementia paralytica had had syphilis. However, since no survey had documented 100% with a history of syphilis, many investigators suggested that syphilis was an important precondition to the development of dementia paralytica but was not the single cause. Hammond, for instance, was clear that syphilis was not the etiology because syphilis was associated with other forms of insanity. The third theory was more psychological in that the disorder was believed to be caused by moral depravity because persons who drank, who frequented prostitutes, and who were in the military were more likely to have the disorder. As additional evidence, dementia paralytica was known to be rare among priests and Quakers.

Research, attempting to provide evidence for or against these theories, was performed. For instance, a famous German psychiatrist named Kraft-Ebbing performed a study in which he injected serum from patients with syphilis into the blood streams of patients with dementia paralytica. Since it was known at the time that a person could only develop a syphilitic infection once, if any of the patients with dementia paralytica developed syphilis, it would prove that they had not had syphilis previously. Hence syphilis could not be the cause of the disorder. None of the 32 patients with dementia paralytica developed syphilis. Kraft-Ebbing concluded that syphilis was the cause of this disorder.

The conclusive evidence regarding the etiology of dementia paralytica occurred in the early twentieth century. In 1906, the bacillus that causes syphilis was isolated, and the Waserman test for syphilis was developed. Plaut (1911) demonstrated that patients with dementia paralytica had positive Waserman tests from blood samples and also from samples of cerebrospinal fluid. In 1913, two Americans, Noguchi and Moore, were able to identify the presence of the syphillitic baccilus in the brains of patients who had died from dementia paralytica (Quetel, 1990). The name of this disorder was changed again to reflect the new understanding of this disorder. It was called paresis or general paralysis associated with tertiary syphilis of the central nervous system. However, even after the discovery of the cause of dementia paralytica, the development of antibiotics to treat the disorder was another 30 years in the future. Thus, many patients were treated by inoculating them with malaria (Braslow, 1995).

4.03.4 KRAEPELIN

Another important development at the turn of the century was the international focus on the classificatory ideas of a German psychiatrist named Emil Kraepelin (Berrios & Hauser, 1988). Kraepelin was a researcher who initially learned his approach to research in a laboratory organized by Wundt, one of the founders of modern experimental psychology. After completing his medical degree, Kraepelin became the medical director for an insane asylum in east Prussia. While in that setting, Kraepelin published a number of experimental psychology studies on persons with mental disorders. He also began to write textbooks about psychopathology for German medical students. Like most textbook authors, Kraepelin organized his chapters around the major categories of mental disorders that he recognized. The sixth edition of Kraepelin's textbook (Kraepelin, 1902/1896) attracted major international attention because of two chapters in these texts.

One chapter was about a disorder that Kraepelin described as dementia praecox (praecox = early) which was a form of psychosis that had a typical age of onset in adolescence. Kraepelin recognized three subtypes to this disorder: hebephrenic, catatonic, and paranoid. Kraepelin's chapter on dementia praecox paralleled the immediately preceding chapter on dementia paralytica. Just as dementia praecox had three descriptive subtypes, dementia paralytica also had three descriptive subtypes: a depressed form, a grandiose form, and a paranoid form.

The second chapter to attract attention was what Kraepelin named manic-depressive insanity. Kraepelin's observations of patients in asylums had led him to believe that the mania and depression (= melancholia) had the same type of course when these patients were observed over time. Both were episodic disorders. Moreover, nineteenth century clinicians had recognized that there were some patients who went from episodes of mania to episodes of depression and vice versa. These observations led Kraepelin to hypothesize that mania and depression were essentially flip sides of the same coin. Hence, he combined, what had been recognized since the times of the ancient Greeks as two disorders, into one mental disorder.

In 1917, the newly formed American Psychiatric Association (APA) adopted a classification system that was quite similar to the classification contained in Kraepelin's sixth edition of his textbook. This early twentieth century American classification included the concepts of dementia praecox and manic-depressive disorder. The classification also adopted the fundamental Kraepelinian distinction between the organic disorders, the functional psychoses, and the neurotic/character disorders.

In 1932, the APA officially adopted a new classification system as part of the Standard Classified Nomenclature of Diseases (APA, 1933). This new classification, however, did not attract much attention (Menninger, 1963).

4.03.5 EARLY *DSMS* AND *ICDS*

World War II led to a renewed emphasis on classification. During the war, nearly 10% of all military discharges were for psychiatric reasons. By the time the war ended, there were four major competing classification systems at use in the USA: (i) the standard system adopted by the APA in 1932; (ii) the US Army classification; (iii) the US Navy classification; and (iv) the Veterans Administration system (Raines, 1952). In response to this disorganization, the APA formed a task force to create a system that would become standard in the USA. The result was the *Diagnostic and statistical manual of mental disorders* (*DSM*; APA, 1952). This classification is usually known by its abbreviated name: *DSM-I*. The *DSM-I* contained 108 different diagnostic categories.

The *DSM-I* was important for a number of reasons. First, the major rationale behind the *DSM-I* was to create a classification system that represented a consensus of contemporary thinking. Care was taken to include all diagnostic concepts that were popular in American psychiatry during the 1940s and 1950s. Thus, the *DSM-I* emphasized communication among professionals as the major purpose of classification and emphasized the need for psychiatric classification to be an accepted nomenclature that members of a profession can use to discuss their clinical cases. Consistent with this emphasis on communication, early versions of the *DSM-I* were revised, based on comments elicited in a questionnaire sent to 10% of the members of the APA. The *DSM-I* was finally adopted by a vote of the membership of the APA (Raines, 1952).

The emphasis on communication in the *DSM-I* led to a similar organizing movement at an international level. The international psychiatric community had adopted a classification of mental disorders that was part of the *International statistical classification of disease, injuries and causes of death* (6th ed.) (*ICD-6* World Health Organization (WHO), 1948). The first *ICD* had been created in 1900 and was a medical classification of causes of death. The *ICD-6* was the first edition to include all diseases, whether they led to death or not. The classification of mental disorders in the *ICD-6* did not gain broad acceptance. A committee, chaired by the British psychiatrist Stengel, was formed to review the classification systems used by various countries and to make any necessary recommendation for changes to the WHO. What Stengel (1959) found was a hodgepodge of diagnostic systems between, and sometimes within, different countries. Stengel despaired over the confused state of international classification and said that the *ICD-6* did not serve as a useful nomenclature. A positive note in his review, however, concerned the *DSM-I*, which Stengel considered an advance over the other national classifications because of its emphasis on representing a well-organized nomenclature for a country.

As a result of Stengel's review, there was an international movement to create a consensual system that would be adopted by the WHO. The final product was the mental disorders section of the *ICD-8*. In the USA, the APA revised its *DSM* classification to correspond with the *ICD-8*. The US version of the *ICD-8* was known as the *DSM-II* (APA, 1968). The *DSM-II* had 185 categories. These categories were subdivided by a hierarchical organizational system. First, there was a distinction between psychotic and nonpsychotic disorders. The psychotic disorders were further subdivided into organic and nonorganic disorders. The classification of the organic disorders was in terms of etiology (e.g., tumors, infections, heredity, etc.). The nonorganic psychotic disorders primarily contained the Kraepelinian categories of schizophrenia and manic-depressive insanity. The nonpsychotic disorders were subdivided into eight subheadings including the neuroses (now called anxiety disorders), personality disorders, mental retardation, etc.

4.03.6 NEO-KRAEPELINIANS

After the publication of the *DSM-II*, psychiatric classification became a very unpopular topic. There were three general lines of criticism that were aimed at classification. First, the diagnosis of mental disorders was unreliable as shown by empirical research. Second, a number of critics attacked the implicit medical model approach to psychopathology that was associated with the *DSM-I* and *DSM-II*. Third, sociologists became interested in a theory of labeling that suggested the process of classification was a process that stigmatized human beings who adopted unusual patterns of behavior and the act of diagnosis could lead to self-fulfilling prophecies.

The first of these criticisms was summarized in three different review articles by Kreitman (1961), Zubin (1967), and Spitzer and Fleiss (1974). All discussed various problems associated with the classification of mental disorders and why poor reliability was a significant issue. Zubin (1967) made an excellent case that the lack of uniform statistical procedures for estimating reliability was a serious methodological problem. Spitzer and Fleiss (1974) suggested a solution to this problem and showed how this solution could be applied retrospectively to earlier data. Kreitman (1961) probably had the most far-reaching analysis of the issue because he said that the unreliability problem had been overemphasized and that the more serious issue was the unexplored validity of diagnostic concepts.

The second issue criticism of the early *DSM*s and *ICD*s was the implicit acceptance of a medical model. Despite the dramatic etiological solution to dementia paralytica, most of the twentieth century research has been disappointing to those who believed that mental disorders are caused by underlying biological processes. Large amounts of research have attempted to discover the etiology of disorders such as schizophrenia, yet, despite interesting advances, a clear understanding of the cause of this disorder is not available. A psychiatrist named Thomas Szasz published a book titled *The myth of mental illness* (Szasz, 1961). He argued that mental disorders are not diseases, but instead are better conceptualized as "problems in living." He argued that psychiatrists had placed themselves into the role of moral policeman to control individuals with deviant behavior patterns. Szasz is now considered one of a group of critics of classification known by the title "antipsychiatrists." Others in this group are the British psychiatrist Laing (1967), the French psychoanalyst Lacan (1977), and recent authors such as Sarbin (1997), and Kirk and Kutchins (1992).

The third criticism of classification that became popular in the 1960s and 1970s was the labeling criticism. Sociologists such as Matza (1969) and Goffman (1961) suggested that the act of psychiatric diagnosis could lead to self-fulfilling prophecies in which the behavior of deviant individuals was constrained to become even more deviant. A dramatic demonstration of the labeling criticism was contained in a controversial paper published by Rosenhan (1973). In this paper, Rosenhan and his colleagues gained admission to mental hospitals even though they reported everything about themselves factually except their names and one auditory hallucination. All but one of these pseudopatients were admitted with a diagnosis of schizophrenia and all were released with a diagnosis of schizophrenia in remission. The pseudopatients commented that most of the patients were aware that they did not belong there, even though the hospital staff never figured that out. In addition, the experiences of the pseudopatients supported the labeling concern. For instance, one pseudopatient reported being bored while being on a ward and walking around. A nurse noticed him pacing and asked if he was feeling anxious. Following the publication of the Rosenhan paper, an issue of the *Journal of Abnormal Psychology* in 1975 was devoted to commentaries on this controversial paper.

Partially in reaction to these criticism of classification, a new school of thought was formed in psychiatry called the neo-Kraepelinians (Klerman, 1978). This group of psychiatrists, initially an active collection of psychiatric researchers at Washington University in St. Louis, believed that psychiatry, with its psychoanalytic emphasis, had drifted too far from its roots in medicine. The neo-Kraepelinians emphasized that psychiatry should be concerned with medical diseases, that extensive research was needed on the biological bases of psychopathology, and that much more emphasis needed to be placed upon classification if knowledge about psychopathology was to grow. Klerman (1978) summarized the perspective implicit in the neo-Kraepelinian approach to psychiatry by listing the following tenets:

(i) Psychiatry is a branch of medicine.

(ii) Psychiatry should utilize modern scientific methodologies and base its practice on scientific knowledge.

(iii) Psychiatry treats people who are sick and who require treatment.

(iv) There is a boundary between the normal and the sick.

(v) There are discrete mental illnesses. Mental illnesses are not myths. There is not one, but many mental illnesses. It is the task of scientific psychiatry, as of other medical specialties, to investigate the causes, diagnosis, and treatment of these mental illnesses.

(vi) The focus of psychiatric physicians should be particularly on the biological aspects of mental illness.

(vii) There should be an explicit and intentional concern with diagnosis and classification.

(viii) Diagnostic criteria should be codified, and a legitimate and valued area of research should be to validate such criteria by various techniques. Further, departments of psychiatry in medical schools should teach these criteria and not depreciate them, as has been the case for many years.

(ix) In research efforts directed at improving the reliability and validity of diagnosis and classification, statistical techniques should be utilized.

In 1972, a group of the psychiatric researchers at Washington University published a paper entitled "Diagnostic criteria for use in psychiatric research" (Feighner, Robins, Guze, Woodruff, Winokur, & Munoz, 1972). This paper listed 15 mental disorders that they believed had sufficient empirical evidence to establish their validity, and listed a set of diagnostic criteria for defining these disorders. The authors argued that a major problem in research about these disorders had stemmed from the lack of uniform definitions by different researchers of the disorders. They suggested that future research on any of these disorders should utilize the diagnostic criteria proposed in their paper.

The paper by Feighner et al. had a dramatic impact on American psychiatry. It was a heavily cited paper, probably the most frequently referenced journal article of the 1970s in psychiatry. The diagnostic criteria were immediately adopted and became the standard for psychiatric research. Moreover, the 15 categories described by Feighner et al. were expanded into a much larger set of categories, focusing primarily on the schizophrenic and affective disorders (Spitzer, Endicott, & Robins, 1975). This new classification was called the Research Diagnostic Criteria (RDC) and had an associated structured interview known as the SADS. Since the lead author of the RDC, Robert Spitzer, had been appointed as the new chairperson responsible for organizing the *DSM-III*, the RDC became the initial foundation from which the *DSM-III* developed.

4.03.7 *DSM-III*, *DSM-III-R*, AND *DSM-IV*

The *DSM-III* (APA, 1980) was a revolutionary classification. First, unlike the *DSM-I* and *DSM-II*, which had emphasized using consensus as the major organizing principle, the *DSM-III* attempted to be a classification based on scientific evidence rather than clinical consensus. For instance, the classification of depression was very different from the view of depression in the *DSM-I* and *DSM-II*, largely because of family history data gathered in research studies performed by the neo-Kraepelinians. In the earlier

*DSM*s, the primary separation of affective disorders was in terms of a psychotic vs. neurotic distinction. The *DSM-III* dropped this differentiation and, instead, emphasized a separation of bipolar vs. unipolar mood disorders. Second, the *DSM-III* discontinued the use of prose definitions of the mental disorders. The neo-Kraepelinians were impressed by the research data suggesting that the reliability of psychiatric classification, as represented in the *DSM-I* and *DSM-II*, was less than optimal (Spitzer & Fleiss, 1974). To try to help improve diagnostic reliability, virtually all mental disorders in the *DSM-III* were defined using diagnostic criteria stimulated by the innovative system used in the Feighner et al. paper. Third, the *DSM-III* was a multiaxial classification. Because the *DSM-I* and *DSM-II* were "committee products," the subsections of these classifications had different implicit organizing principles. For instance, in the *DSM-I/DSM-II*, the organic brain syndromes were organized by etiology, the psychotic disorders were organized by syndromes, and the neurotic disorders were organized according to ideas from psychoanalytic theory. In order to avoid the confusion inherent in the use of multiple organizing principles, the *DSM-III* adopted a multiaxial system that permitted the separate description of the patient's psychopathology syndrome (Axis I), personality style (Axis II), medical etiology (Axis III), environmental factors (Axis IV), and role disturbances (Axis V). The *DSM-III* was published in 1980 and contained 265 mental disorders. Moreover, the size of the manuscript for the *DSM-III* was 482 pages, a huge increase over the 92 pages of the *DSM-II*.

The revolutionary impact of the *DSM-III* led to changes in many areas of the mental health professions. One area of impact was in terms of research. As soon as versions of the *DSM-III* began to be disseminated to researchers interested in mental disorders, new studies began to appear that explored the adequacy of the diagnostic criteria used in this classification. The *DSM-III* was a marked stimulus for descriptive research in psychiatry and in the other mental health professions. Another area of impact was political. There was a major controversy that erupted in the late 1970s over the issue of whether the term "neurosis" should appear in the *DSM-III*. Spitzer and the neo-Kraepelinians had exorcized this term from the classification because of its psychoanalytic associations. The psychoanalysts lobbied intensely within the APA to have the term reintroduced. Although a compromise was achieved (Bayer, 1981), the psychoanalysts lost ground in this struggle, and their influence in organized psychiatry has continued to wane since that time. A third area of impact for the *DSM-III* was economic. The *DSM-III* became very popular, sold well in the USA, and even became a surprisingly large seller to the international community as translations appeared. The sizeable revenues that accrued to the APA led to the formation of the American Psychiatric Press, which published subsequent versions of the *DSM* as well as many other books of clinical interest.

Despite its innovations and generally positive acceptance by mental health professionals, a number of criticisms were leveled at the *DSM-III*. One focus of criticism concerned the diagnostic criteria. Despite the intention to make decisions about the classificatory categories using scientific evidence, most diagnostic criteria were based on the intuitions of experts in the field. In addition, even though the goal when formulating diagnostic criteria was to make them as behavioral and explicit as possible, not all criteria met this goal. Consider part of the *DSM-III* criteria for histrionic personality disorder, for instance.

> Characteristic disturbances in interpersonal relationships as indicated by at least two of the following:
>
> 1) perceived by others as shallow and lacking genuineness, even if superficially warm and charming
> 2) egocentric, self-indulgent and inconsiderate of others
> 3) vain and demanding
> 4) dependent, helpless, constantly seeking reassurance
> 5) prone to manipulative suicidal threats, gestures or attempts

Note that common language terms such as "shallow" were highly subjective. In addition, a criterion such as 3) requires an inference about motivations and reasons for behavior, rather than direct observation of behaviors. Finally, subsequent research showed that criterion 5) above actually was observed more frequently in borderline rather than histrionic patients (Stangl, Pfohl, Zimmerman, Bowers, & Corenthal, 1985).

A second major criticism of the *DSM-III* concerned the multiaxial system. First, diagnosing multiple axes required increased time and effort by clinicians, an exercise they were unlikely to do unless they were certain that the gain in information was significant. Second, the relative emphasis on these five axes in the *DSM-III* was sizeably different. In the *DSM-III* manual, almost 300 pages were devoted to defining the Axis I disorders, another 39 pages were spent on Axis II disorders, whereas only two pages each were devoted to Axes IV and V.

Moreover, Axes I and II were assessed using diagnostic categories, whereas Axes IV and V were measured using relatively crude, ordinal rating scales. Third, the choice of particular axes was also criticized. For instance Rutter, Shaffer, and Shepard (1975) had advocated the use of one axis for the clinical syndrome in childhood disorders with a second focusing on intellectual level. Instead, both clinical syndromes and mental retardation were Axis I disorders in the *DSM-III*. A group of psychoanalysts argued that defense mechanisms should have been included as a sixth axis. Psychiatric nurses advocated an additional axis for a nursing diagnosis, relevant to the level of care required by a patient.

Only seven years later, a revision to the DSM-III was published. This version, known as the *DSM-III-R* (APA, 1987), was intended primarily to update diagnostic criteria using the research that had been stimulated by the *DSM-III*. It was called a revision because the goal was to keep the changes small. However, the differences between the *DSM-III* and the *DSM-III-R* were substantial. Changes included renaming some categories (e.g., paranoid disorder was renamed delusional disorder), changes in specific criteria for disorders (e.g., the criteria for schizoaffective disorder), and reorganization of categories (e.g., panic disorder and agoraphobia were explicitly linked). In addition, six diagnostic categories originally in the *DSM-III* were deleted (e.g., egodystonic homosexuality and attention deficit disorder without hyperactivity) while a number of new specific disorders were added (e.g., body dysmorphic disorder and trichotillomania). As a result, the *DSM-III-R* contained 297 categories compared to the 264 categories in the *DSM-III*.

Associated with the *DSM-III-R* was the development of a major controversy that had political overtones. Among the changes proposed for the *DSM-III-R* was the addition of three new disorders: premenstrual syndrome (PMS), masochistic personality disorder and paraphilic rapism. These additions raised the ire of a number of groups, especially feminists. Concerning PMS, feminists argued that the inclusion of this disorder into the *DSM* would be the implicit assumption that the emotional state of women can be blamed on their biology. If it were to be a disorder, the feminists argued that PMS should be classified as a gynecological disorder rather than a psychiatric disorder. Masochistic personality disorder had been suggested for inclusion by psychoanalysts who pointed to the extensive literature on this category. Feminists, however, believed that this diagnosis would be assigned to women who had been physically or sexually abused. Thus, this diagnosis would have the unfortunate consequence of blaming these women for their roles as victims. Finally, the proposal to include paraphilic rapism was also attacked. The critics argued that this diagnosis would allow chronic rapists to escape punishment for their crimes because their behaviors could be attributed to a mental disorder. Thus, these men would not be held responsible for their behaviors.

Because of the ensuing controversy, a compromise somewhat similar to the earlier compromises about homosexuality and neurosis was attempted. The authors of the *DSM-III-R* revised the names of the first two disorders (PMS and masochistic personality disorder) to periluteal phase dysphoric disorder and self-defeating personality disorder. They also deleted the proposal to add paraphilic rapism. In addition, another disorder, sadistic personality disorder, was added presumably to blame abuser as well as victims, thereby balancing the potential antifeminine connotations of self-defeating/masochistic personality disorder. This compromise was not successful. As a result, the executive committee for the American Psychiatric Classification decided not to include these categories in the body of the *DSM-III-R*. Instead, they were placed in an appendix as disorders needing more research (Walker, 1987).

The *DSM-IV* was published in 1994, contained 354 categories, and was 886 pages in length, a 60% increase over the *DSM-III* and almost seven times longer than the *DSM-II* (APA, 1994). There are 17 major categories in the *DSM-IV*:

disorders usually first diagnosed in childhood
cognitive disorders
mental disorders due to a general medical condition
substance-related disorders
schizophrenia and other psychotic disorders
mood disorders
anxiety disorders
somatoform disorders
factitious disorders
dissociative disorders
sexual disorders
eating disorders
sleep disorders
impulse control disorders
adjustment disorders
personality disorders
other conditions that may be a focus of clinical attention

The *DSM-IV* retained a multiaxial system and recognized five axes (dimensions) along which patient conditions should be coded:

Axis I clinical disorders
Axis II personality disorders/mental retardation
Axis III general medical conditions
Axis IV psychosocial and environmental problems
Axis V global assessment of functioning

As with the *DSM-III-R*, a major focus in the *DSM-IV* revision concerned diagnostic criteria. A total of 201 specific diagnoses in the *DSM-IV* were defined using diagnostic criteria. The average number of criteria per diagnosis was almost eight. Using this estimate, the *DSM-IV* contains slightly over 1500 diagnostic criteria for the 201 diagnoses.

To give the reader a glimpse of how the diagnostic criteria have changed from the *DSM-III* to the *DSM-IV*, the criteria for histrionic personality disorder are listed below:

> A pervasive pattern of excessive emotionality and attention seeking, beginning by early adulthood and present in a variety of contexts, as indicated by five (or more) of the following:
>
> (1) is uncomfortable in situations in which he or she is not the center of attention
> (2) interaction with others is often characterized by inappropriate sexually seductive or provocative behavior
> (3) displays rapidly shifting and shallow expression of emotions
> (4) consistently uses physical appearance to draw attention to oneself
> (5) has a style of speech that is excessively impressionistic and lacking in detail
> (6) shows self-dramatization, theatricality, and exaggerated expression of emotion
> (7) is suggestible, i.e., easily influenced by others or circumstances
> (8) considers relationships to be more intimate than they actually are

In addition to presenting diagnostic criteria, the *DSM-IV* contains supplementary information about the mental disorders in its system. For instance, there are three pages of information about histrionic personality disorder including

diagnostic features (prose description of symptoms)
associated features and disorders (mental disorders that are likely to co-occur)
specific culture, age and gender features
prevalence
differential diagnosis (how to differentiate the disorder from others with which it is likely to be confused)

In order to help ensure that the *DSM-IV* would be the best possible classification of mental disorders, the steering committee for this classification contained 27 members, including four psychologists. Reporting to this committee were 13 work groups composed of 5–16 members. Each work group had a large number of advisors (typically over 20 per work group).

There were three major steps associated with the activities of each work group. First, all work groups performed extensive literature reviews of the disorders under their responsibility. Many of these literature reviews were published in the journal literature. Second, the work groups solicited contributions of descriptive data from researchers in the field. Using these data, the work groups reanalyzed the data to decide which diagnostic criteria needed revision. Third, a series of field trials was performed on specific topics. For instance, the personality disorders work group performed a multicenter study on antisocial personality disorder which led to a significant alteration in the diagnostic criteria for that disorder.

The *DSM-IV* was not without controversy. For instance, the issues that had been raised in the *DSM-III-R* regarding premenstrual syndrome, masochistic personality disorder, and sadistic personality disorder continued in the *DSM-IV*. Interestingly, none of these three disorders were included in the *DSM-IV*. In fact, two (masochistic and sadistic personality disorders) completely disappeared from the classification. PMS remained in an appendix as a disorder "for further study." Interestingly, 17 other disorders joined PMS in this appendix as did three possible new axes for the multiaxial system (defense mechanisms, interpersonal functioning, and occupational functioning).

Earlier editions of the *DSM* had few, if any, references to document the sources for any factual claims in these classifications. The *DSM-IV* attempt to overcome this problem was by publishing a five-volume companion set of sourcebooks. These sourcebooks are edited papers by members of the work groups. The intent of the sourcebooks is to provide a scholarly basis for understanding the specific decisions that were made by the work groups.

4.03.8 *ICD-9, ICD-9-CM,* AND *ICD-10*

Earlier in this chapter, the point was made that the *DSM-II* and *ICD-8* were virtually identical because the American psychiatric community had joined an international movement to create a consensual classification. With the revolutionary *DSM-III*, American psychiatry reversed itself and created a radically new classification based upon the purpose of description, rather than emphasizing a system that would be acceptable world-wide.

The editions of the *ICD*s were intended to be revised every 10 years. The *ICD-8* was published in 1966; the *ICD-9* came out in 1977. The mental disorders section of the *ICD-9* was very similar to the *ICD-8/DSM-II* (WHO, 1978). The psychotic/nonpsychotic distinction was the primary hierarchical distinction among categories. The psychotic disorders were further subdivided into organic and functional psychoses. There were 215 categories in this system, and the *ICD-9* was published in a monograph that was 95 pages in length.

The USA has signed an international treaty that obliges it to use the *ICD* as the official medical classification. Thus, when the *DSM-III* was created, an odd numeric coding scheme was incorporated so that the *DSM-III* categories could be incorporated with the *ICD-9* framework. To understand this, below is an overview of the specific diagnostic categories under the general heading of anxiety disorders in the *DSM-III*:

Phobic disorders
300.21 Agoraphobia with panic attacks
300.22 Agoraphobia without panic attacks
300.23 Social phobia
300.29 Simple phobia
Anxiety states
300.01 Panic disorder
300.02 Generalized anxiety disorder
300.30 Obsessive compulsive disorder
Post-traumatic stress disorder
308.30 Acute
309.81 Chronic or delayed
300.00 Atypical anxiety disorder

Notice that the coding scheme for the anxiety disorders in the *DSM-III* is not what one might expect. Most of the anxiety disorders are coded with 300.xx numbers. However, the two forms of post-traumatic stress disorder are coded 308.30 and 309.81. Notice also that the first number after the decimal point is somewhat irregular. The phobic disorders, listed first in the *DSM-III*, are given 300.2x codes whereas the anxiety states are coded 300.0x or 300.30.

To understand why the *DSM-III* codes appear this way, below is a listing of the specific neurotic disorders in the *ICD-9*:

Mental disorders (290–319)
Nonpsychotic mental disorders (300–316)
Neurotic disorders (300)
300.0 Anxiety states
300.1 Hysteria
300.2 Phobic states
300.3 Obsessive-compulsive disorders
300.4 Neurotic depression
300.5 Neurasthenia
300.6 Depersonalization syndrome
300.7 Hypochondriasis
300.8 Other neurotic disorders
300.9 Unspecified

In the *ICD-9* system, all diagnoses have four-digit codes. The codes for all mental disorders range from 290 to 319. The 29x disorders are the psychotic disorders; 300–315 are reserved for nonpsychotic disorders; and 316–318 are codes for classifying mental retardation. The first subheading under the nonpsychotic disorder is the neurotic disorders. Notice that this subheading includes what the *DSM-III* recognizes as anxiety disorders but it also includes categories that the *DSM-III* placed under other headings (e.g., neurotic depression = dysthymic disorder and depersonalization syndrome). Because the *DSM-III* anxiety disorders were mostly found under the *ICD-9* neurotic disorders, these categories were given 300.xx codes. However, post-traumatic stress disorder (chronic; PTSD) was given a code number of 309.89 because it was included under the general *ICD-9* heading of adjustment reactions. Notice also that PTSD has an xxx.8 coding. In the *ICD-9* coding system, all diagnoses with an xxx.8 code represent country-specific categories that are generally recognized by the international psychiatric community. Thus, PTSD (309.89) is a US category that has no clear equivalent in the international diagnostic system. Another *DSM-III* category with a coding that reflects a similar status is borderline personality disorder (301.83).

In order to blend the *DSM-III* with the *ICD-9* so that a consistent coding system would be used, a US version of the *ICD-9* was created. This new version was the *ICD-9-CM* (where CM stands for "clinical manual"). The *ICD-9-CM* is the official coding system that all physicians and mental health professionals must use when assigning diagnostic codes. However, American clinicians do not need to refer to the *ICD-9-CM* because the applicable codes are listed in the printed versions of the *DSM-III*, *DSM-III-R* and *DSM-IV*.

As noted earlier, the *DSM-III* and its successors (*DSM-III-R* and *DSM-IV*) were resounding successes. Not only did these systems become dominant in the USA, but they also achieved substantial popularity among European mental health professionals (Mezzich, Fabrega, Mezzich, & Coffman, 1985). The proponents of the *ICD* were somewhat resentful (Kendell, 1991). Thus, when the *ICD-10* was created, substantial changes were made that utilized innovations from the *DSM-III*.

First, like the *DSM-III*, the *ICD-10* went through extensive field testing (Regier, Klaeber, Roper, Rae, & Sartorius, 1994). There were two major goals in the field testing: (i) to ensure that the *ICD-10* could be used in a reliable way across clinicians, and (ii) to examine the acceptability of the mental disorder categories contained in this system. The data reported regarding both of these goals have given a favorable view of the *ICD-10*.

The second important innovation of the *ICD-10* is that the mental disorders section is published in two forms. One form, the blue manual, is subtitled *Clinical descriptions and diagnostic guidelines* (WHO, 1992). The blue manual contains prose definitions of categories and is primarily intended for clinical use. The other, the green manual, is like the *DSM-III* in that the categories are defined using explicit diagnostic criteria with rules regarding how many criteria must be met in order for a diagnosis to be made (WHO, 1993). The green manual is intended for research use.

The complete *ICD-10* is organized into a series of 21 chapters, one of which is Chapter V (labeled with the prefix F) about "Mental and behavioural disorders." Other chapters in the *ICD-10* are:

Chapter I A–B Infections and parasitic diseases

Chapter II C–D Neoplasms

Chapter X J Diseases of the respiratory system

Chapter XXI Z Factors influencing health status and contact with health services

In terms of classificatory size, the mental disorders section of the *ICD-10* and the *DSM-IV* are reasonably similar. The *DSM-IV* contains 354 categories organized under 17 major headings. The *ICD-10* has 389 categories that are structured under 10 major headings. One ironic feature of the *ICD-10* is that it did not adopt a multiaxial system of classification. This decision is ironic because the idea originated in Europe, most prominently by a Swedish psychiatrist, Essen-Moller (1971).

4.03.9 CONTROVERSIES AND ISSUES

Although the *DSM-III* and its successors are usually viewed as substantial improvements in the classification of mental disorders, a number of controversies and issues still remain concerning the classification of mental disorders. The remainder of this chapter attempts to provide an overview of some of these issues.

4.03.9.1 Organizational Models of Classification

There are four organizational models of classification that have been frequently discussed in the literature. Often these models are seen as competing and incompatible. For instance, most discussions of the first two models, the categorical and the dimensional, are presented as if the mental health professions must choose between them. However, hybrid models are possible (Skinner, 1986) and perhaps even the most probable.

Mental disorders are usually discussed as if they are categories. Thus, a categorical model is often implicitly assumed to be the structural model for psychopathology (Kendell, 1975). The tenets of this model are listed below:

> 1.1 The basic objects of a psychiatric classification are patients
> 1.2 Categories should be discrete, in the sense that the conditions for membership should be able to be clearly ascertained
> 1.3 Patients either belong or do not belong to specified classes (categories)
> 1.4 The members of a category should be relatively homogeneous
> 1.5 Categories may or may not overlap
> 1.6 In the borderline areas where categories may overlap, the number of overlapping patients should be relatively small
> 1.7 Cluster analytic methods can be used to initially identify categories
> Discriminant analysis is used to validate categories. (Blashfield, 1991, p. 14)

The *DSM*s, particularly the *DSM-III* and its successors, are seen as fitting a categorical model. According to the categorical model, the unit of analysis is the patient. Diagnoses refer to classes of patients. Patients either are or are not the members of the categories. The categorical model assumes that some type of definitional rule exists by which the membership in a category can be determined. Moreover, membership in a category is considered to be a discrete, all-or-nothing event. An animal is either a cat or not a cat. A patient is either a schizophrenic or not a schizophrenic.

An important assumption of the categorical model is that members of a category are relatively homogeneous. All animals that belong to the class of "birds" should be reasonably similar morphologically. This is not to say that all birds must be alike. Certainly a robin and sparrow have a number of obvious differences. Yet they are more like each other than either is to a lynx, for instance. In the same way, two schizophrenic patients should be relatively similar. Both may have different symptom pictures, but their symptom pictures should be more similar to each other than either is to an antisocial patient (Lorr, 1966).

Classes in a categorical model may or may not overlap. Most uses of the categorical model typically assume that nonoverlapping categories are an ideal condition, but recognize that this condition may not always happen. Thus, overlap among categories is often treated like measurement error—a condition to be tolerated, but minimized. However, there are categorical models that have been developed in which the categories are assumed to overlap (Jardine & Sibson, 1971; Zadeh, 1965). In these models, overlap is not error. Categories are fuzzy sets whose boundaries of membership do not need to result in mutually exclusive groupings.

According to the assumption of relative homogeneity, the number of patients who clearly belong to one and only one category should be relatively frequent, whereas patients who fall in the overlapping, borderline areas between categories should be relatively infrequent. A sparrow–lynx should not occur if categories are to have the necessary homogeneity that allows them to be separable constructs (Grinker, Werble, & Drye, 1968).

Finally, the methods that have been developed to find the boundaries among categories are called cluster analytic methods (Everitt, 1974). Generally, these methods analyze a large matrix of descriptive data on patients, and attempt to form clusters (categories) of relatively homogeneous patients in terms of the descriptive variables that were gathered on the patients. Cluster analysis was used in the 1960s and 1970s to create new descriptive classifications. In the last decade, most researchers have abandoned the use of these methods because of pragmatic difficulties in their application and because of unsolved statistical issues. Meehl, however, has developed a related method for isolating categories that he believes has promise (Meehl, 1995).

Although the categorical model, as presented above, seems to be a common sense model of psychiatric classification, the recent *DSM*s clearly do not adhere to this model. First, as noted above, a categorical model assumes that the unit of analysis is the patient and that groups of patients are members of similar sets called mental disorders. The authors of the *DSM-III*, *DSM-III-R and DSM-IV* explicitly reject this approach. They state that these classifications are not intended to classify individual patients (Spitzer & Williams, 1987). Instead, these recent *DSM*s state that they are classifying disorders (rather than patients).

A second structural model is the dimensional model. The tenets for this model are:

> 2.1 The basic unit of the dimensional model is a descriptive variable (e.g., a symptom, a scale from a self-report test, a laboratory value, etc.)
> 2.2 Dimensions refer to higher-order, abstract variables
> 2.3 A dimension refers to a set of correlated descriptive variables
> 2.4 There are a relatively small number of dimensions compared to the number of descriptive variables, yet the dimensions account for almost as much reliable variance as do the larger number of descriptive variables
> 2.5 Dimensions themselves may be correlated or independent
> 2.6 The methods used to identify dimensions are exploratory factor analysis and multidimensional scaling. Confirmatory factor analysis can be used to test a specific dimensional model. (Blashfield, 1991, p. 15)

For the dimensional model, the basic units of analysis are the descriptive variables. Thus, the dimensional model focuses on symptoms, behaviors, diagnostic criteria, scales from self-report tests, and the like. The dimensional model summarizes these descriptive variables by forming higher-order abstract variables that can serve to represent the original measurement variables. Each of these higher-order abstract variables constitutes a dimension through its conceptualization as occurring on a continua. Patients can have scores anywhere along these dimensions.

A major test of a dimensional model is parsimony. The specific dimensions in such a system should account for most of the systematic, reliable variance that exists within the original set of descriptive variables. If the dimensions do not account for the reliable variance in the original descriptive variables, then using the dimensions will sacrifice a great deal of information and the original variables should be used rather than the smaller set of dimensions.

A third structural model that is often discussed regarding the classification of mental disorders is a disease model. The basic assumption in this model is that all diagnostic categories refer to medical diseases (Wing, 1978). In effect, this model is a modern extension of Griesinger's famous nineteenth-century dictum that all mental disorders are diseases of the brain (Stromgren, 1991). The tenets of the disease model are:

> 3.1 The fundamental units are biological diseases of individual patients (essentialism)
> 3.2 Each diagnosis refers to a discrete disease
> 3.3 Diagnostic algorithms specify objective rules for combining symptoms to reach a diagnosis
> 3.4 Adequate reliability is necessary before any type of validity can be established

3.5 Good internal validity will show that the category refers clearly described patterns of symptoms
3.6 Good external validity will mean that the diagnosis can be used to predict prognosis, course and treatment response. (Blashfield, 1991, p. 8)

Some authors have assumed that a categorical model and a disease model are the same. These models are not identical. A categorical is neutral about the existential status of the categories in its system. A disease model adopts a stronger view. Diseases do have an existential status. Diseases are real. The goal of medical research is to identify, describe, understand and eventually treat these diseases. This belief in the reality of diseases is associated with a broader view about the status of scientific concepts known as essentialism.

Notice also that diseases are not necessarily categorical, at least as this model was described above. For instance, more than one disease can occur in the same patient. In fact, some diseases are very likely to co-occur (e.g., certain sarcomas have high frequency in patients with AIDS). Thus, diseases do not refer to mutually exclusive sets. In addition, there are diseases that are conceptualized as dimensional constructs. Hypertension is the most common example. Patients with hypertension vary along a continuum. A categorical scaling of this continuum is possible, but imposing a categorical separation on this continuum is arbitrary.

The fourth model of classification is the prototype model. Cantor, Smith, French, and Mezzich (1980) have suggested that this model is superior to the implicit categorical model of psychiatric classification. For those readers who do not know what the prototype model is, the easiest way to conceptualize this model is through an example.

According to the prototype model, if a mother wanted to teach a child what "anger" means, she would not say "Steven, you need to understand that anger is an emotion that many of us feel when we are frustrated or upset." Instead, at some point when little Steven is upset because he has to go to bed and he tries to hit his mother, she would say "I know that you are angry, Steven, but it is time to go to bed." And on another day, when Steven is upset because another child took one of his toys, his mother might say, "You are feeling angry, Steven. Being angry is natural, but you should not hit the other child. Maybe if you ask Carol she will return your toy."

In effect, a child learns a concept by being presented with instances of the concept. Once the child is able to associate the instances with a verbal label (i.e., the word "angry"), then the child will have a grasp of the concept. Later, the child will learn to abstract the essential features of the concept. This occurs by making observations about similarities that occur among instances of the concept (e.g., internal feelings, interpersonal context, etc.). Russell and Fehr (1994) provide an interesting and more complete discussion of the concept of anger from a prototype perspective.

Another important aspect of the prototype model is the idea that not all instances of a concept are equally good representatives of the concept. A robin, for instance, is a good exemplar of a bird. Robins are about the same size as most birds; they have feathers; they can fly; etc. Penguins, however, are not a good exemplar. Penguins are larger than most birds; they cannot fly; they do have feathers, although, to a child, their covering probably seems more like fur than feathers; etc.

The above presentation of the prototype model is easy to understand and seems like a common-sense view of classification. However, advocates of the prototype model argue that this model is radically different than a categorical model (Barasalou, 1992; Russell, 1991). According to the categorical model, classificatory concepts are defined by listing the features that are sufficient for making a diagnosis. If a given instance has a sufficient number of these features, then that instance is a member of the classificatory concept. Moreover, all members of a concept are equal. A square is a square. One square is not squarer than another square. In contrast, the prototype model does stipulate that some instances of a concept are better exemplars than others. The Glenn Close character in *Fatal attraction* is a better representation of borderline personality disorder than the Jessica Walter character in *Play Misty for me*.

The basic tenets of the prototype model are presented below:

4.1 Diagnoses are concepts that mental health professionals use (nominalism)
4.2 Categories are not discrete
4.3 There is a graded membership among different instances of a concept
4.4 Categories are defined by exemplars
4.5 Features (symptoms) are neither necessary nor sufficient to determine membership
4.6 Membership in a category is correlated with number of features (symptoms) that a patient has. (Blashfield, 1991, p. 11)

The major difference between the disease and the prototype model is that the latter is associated with nominalism. Nominalism is the position that the names of diseases are just convenient fictions that clinicians use to organize information. Diagnostic terms do not have some underlying reality of their own. Diagnostic concepts are names. These concepts exist simply to meet some functional goal.

The preceding discussion of organizational models of the classification of psychopathology is overly simplistic. Each of these models, when examined more closely, can become quite complex, and the apparent distinctions between the models blur. Two instances of this complexity will be described. First, although the categorical and dimensional are usually presented as if they are competing and antagonistic, Skinner (1986) has suggested that these models are actually complementary. He suggested that the measurement model associated with a dimensional perspective is the more fundamental of the two models. The dimensional model only assumes that, in order to assess a patient, a clinician should gather information on specific descriptive variables that are correlated and that can be summarized by higher-order variables (dimensions). The categorical model also assumes that descriptive variables can be sorted into dimensions. But, in addition, the categorical model asserts that the patients themselves "cluster" into groups that are relatively homogeneous on these dimensions. Thus, from Skinner's hybrid perspective, a pure categorical model makes stronger assumptions about descriptive data than does a dimensional model. However, psychological models of human social cognition suggest that categorical models are more basic than are dimensional models (Wyer & Srull, 1989).

A second example of the complexity of these models is associated with Barsalou's distinction between "prototype" and "exemplar" models. There are two types of approach that can be used to define a concept: intensional definitions and extensional definitions. An intensional definition lists the features that can be used to separate the concept from related concepts (e.g., a square is a four-sided closed figure whose sides are equally long and occur at right angles to each other). In contrast, an extensional definition is a definition by listing the members of the category (e.g., the 1957 New York Yankees included Roger Marris, Yogi Berra, Micky Mantle, etc.). Barsalou says that a prototype model uses an intensional definition for categories in which the prototype represents the average (centroid) of the concept. Using the example of a child learning about birds, Barasalou suggests that the reason that a robin is a prototype for bird, whereas a penguin is not, is that robins have the average features of most birds (small size, bright coloring, migration, food choices, etc.). Penguins, in contrast, are statistically unusual on these dimensions. An exemplar model, according to Barsalou, uses a particular type of extensional definition in which a concept is defined by an outstanding or exemplary instance of the concept. Thus, Micky Mantle might be a good exemplar of the 1957 New York Yankees even though Mantle's batting prowess was hardly average. In the same way, Abraham Lincoln might be seen as exemplar of American presidents, even though he was not average for this set of individuals.

4.03.9.2 Concept of Disease

The discussion of the disease model vs. the prototype model led to a brief introduction regarding the dualism of essentialism vs. nominalism. This dualism is associated with a complicated problem that has bothered writers about classification throughout the last two centuries. What is a disease? Do the two concepts of "mental disease" and "mental disorder" have the same meaning? To discuss the issues associated with the meaning of "disease," the writings of a British internist named Scadding will be discussed. At the end of this section, other approaches to the concepts of "disease" and "disorder" are briefly introduced.

Scadding's (1959) first attempt to discuss the meaning of disease occurred in a short essay. This essay offered his first general definition of disease which read:

> The term "a disease" refers to those abnormal phenomena which are common to a group of living organisms with disturbed structure or function, the group being defined in the same way.

In effect, Scadding was saying that a disease was associated with a cluster of signs and symptoms (i.e., abnormal phenomena) that are associated with some functional or structural disturbance in the human body. Scadding went on to argue that a disease had (i) defining characteristics and (ii) diagnostic criteria. The defining characteristics refer to the indications that prove the presence of the disease (e.g., locating syphilitic bacilli in the brains of individuals with paresis). In contrast, the diagnostic criteria are signs and symptoms of the disease that may or may not be present (e.g., motor paralysis, grandiose delusions, and sluggish pupillary response to light would be possible diagnostic criteria for paresis).

Ten years later, Scadding (1969) revised his definition of disease to read as follows:

> A disease is the sum of the abnormal phenomena displayed by a group of living organisms in association with a specified common characteristic or set of characteristics by which they differ from the norm for their species in such a way as to place them at biological disadvantage.

There are four important points to note about this second definition of disease. First, the emphasis is on abnormal phenomena. Scadding wanted to be quite clear that the name of the disease does not refer to the etiologic agent causing the disease. That is, tuberculosis is not simply defined by the presence of a particular bacterium, *Mycobacterium tuberculosis*. To have tuberculosis a patient must manifest the symptoms of the disease as well as the anatomical changes (i.e., the formation of characteristic lesions called tubercles in the lung) associated with this disease. This distinction is important because there are other bacilli, besides *Mycobacterium tuberculosis*, which can cause these lesions and the symptom pattern of tuberculosis.

Second, the definition contains the rather vague phrase "common characteristic." Scadding argued that there are three general ways of characterizing any individual disease: (i) a clinical-descriptive approach, (ii) a morbid anatomical approach, and (iii) an etiological approach. (Note that these approaches were presented almost a century earlier by Hammond (1883), as previously discussed.) The clinical-descriptive approach is simply the description of the "syndrome." That is, the clinical-descriptive approach outlines a loose cluster of signs and symptoms that are correlated in their appearance in patients. For instance, the clinical-descriptive approach to defining diabetes focuses on frequent urination, an intense thirst, and rapid loss of weight as indications of this disorder. The clinical-descriptive approach dominated when the *DSM-III* and its successors were created. The second approach concerns morbid anatomy. This refers to the anatomical changes in the body's structure associated with the disease. For diabetes mellitus, a morbid anatomy view might define this disease in terms of the destruction of β-insulin-producing cells in the pancreas. Finally, the etiological approach would be to define a disease in terms of the syndrome caused by a known and specifiable etiological process. For Type I diabetes mellitus, this might be an autoimmune process whose exact details have yet to be specified. For paresis, the etiological agent is the effect of the syphilitic bacillus on the central nervous system of the affected individual.

Scadding commented that, historically, knowledge about diseases typically proceeds from clinical-description to morbid anatomy to etiology. Certainly his observation seems to be correct when applied to the history of paresis. He argued that any of these approaches to characterizing a disease are appropriate. That is, a disease can be defined in terms of a clinical syndrome; or it can be defined by some associated morbid anatomy; or it can be defined through a recognition of its etiological cause.

The third point to note about Scadding's definition is its emphasis on norms, in that disease refers to an abnormality. To be a disease, the condition must refer to phenomena that are statistically deviant. For instance, most of us have various types of bacteria that normally reside in our intestines and which are important in the digestive process. The presence of these bacteria do not define a disease. The effects of these bacteria are normative. In fact, individuals without these bacteria have abnormal digestive processes.

Finally, the definition ends with the term "biological disadvantage." Scadding introduced this term because he recognized that not all nonnormative aspects of human structure and functioning should be called diseases. For instance, some individuals produce an abnormal amount of ear wax. However, this should not define a disease unless there is some biological disadvantage associated with this condition. Although the term biological disadvantage is not more precisely specified, its general meaning seems clear: syphilis and diabetes place an individual at biological disadvantage since both can lead to death if untreated.

In 1979, Scadding and two Canadian authors (Campbell, Scadding and Roberts, 1979) extended their ideas about disease by studying what physicians and nonphysicians meant by the concept of disease. They published a survey that they had conducted regarding the meaning of disease. To conduct their survey, these authors read a list of possible diseases to four groups of individuals: (i) a group of medical faculty, (ii) a group of nonmedical faculty, (iii) a sample of general practice physicians, and (iv) a sample of youth in British and Canadian schools. The subjects in this study were asked to note whether the terms being read aloud referred to diseases or not. In addition, the subjects were asked to assign degree of confidence ratings to their decisions.

At the top of the list of conditions that were viewed as diseases are infections (malaria, tuberculosis, syphilis, measles). Virtually everyone in the four groups, whether physicians or nonphysicians, agreed that these terms referred to diseases. Syphilis, for instance, was considered a disease by over 90% of the subjects in all groups.

At the bottom of the list were concepts that were not considered diseases by these subjects. Two terms that were seen as referring to diseases by less than 30% of all four groups were drowning and starvation. Many of the terms at the bottom of Scadding's list might be described as injuries, i.e., traumas that affect bodily functioning and that were caused by identifiable external events such as a car accident (e.g., skull fracture) or ingestion of a substance (e.g. barbituate overdose, poisoning).

The psychiatric concepts in the list (schizophrenia, depression, and alcoholism) were ranked in the middle. There was considerable variance among the four groups regarding whether these concepts referred to diseases. For instance, faculty of medical schools rated these three concepts in the following order: schizophrenia (78%), depression (65%), and alcoholism (62%). Children in secondary schools had quite different impressions of what is considered to be a disease: alcoholism (76%), schizophrenia (51%) and depression (23%).

One factor that had a large influence on whether a term referred to a disease concerned the role of a physician in the diagnosis or treatment of the disorder. Malaria and syphilis require a doctor to diagnose and treat. In contrast, starvation can be identified and treated by nonmedical individuals. The latter is also true of acne and hemorrhoids, although the intervention of physicians can prove useful for both. Consistent with this view, acne and hemorrhoids were ranked in the middle of the list. The potential role of nonphysicians in the treatment of mental disorders may also account for the occurrence of schizophrenia, depression and alcoholism in the middle of the same list. Scadding et al (1979, p. 760) concluded their paper with the following interesting comment:

> Most people without medical training seem to think of a disease as an agent causing illness. The common concept of "disease" is essentialist: diseases exist, each causing a particular sort of illness. Doctors tend to adopt a more nominalist position, but they obviously retain remnants of belief in the real existence of diseases.

When viewed from this dualism of an essentialist vs. a nominalist perspective, Scadding had started his definitional attempts from an essentialist perspective but, by the time of his last writings on the topic, he was suggesting that a nominalist view was preferred. Interestingly, the writings of a prominent British psychiatrist, Kendell, who has also tried to solve this issue, have followed the same progression. His ideas shifted from a paper trying to settle on an essentialist meaning of disease (Kendell, 1976) to a skeptical discussion of how the disease model fails to explain alcoholism (Kendell, 1979) to a nominalist view (Kendell, 1986).

Because this nominalism vs. essentialism dualism is so important, the approach of Wulff, Pedersen, and Rosenberg (1986) is discussed briefly. To understand the dualism, Wulff et al. discussed possible ways to classify defective grandfather clocks as an example.

Suppose that one examines how people who work in a repair shop might classify clocks. The receptionist, who knows very little about the workings of the clocks, might classify them descriptively. Thus, some clocks would be placed together because they do not work after being wound; others have broken faces, arms or other parts; and still others do not keep time accurately. Another person who might classify the grandfather clocks would be the bookkeeper of the shop. This individual might classify the clocks according to the manufacturer and cost of the clock. A third person who might classify the clocks is the repairman. He might organize clocks anatomically into those with accumulated dirt impeding their normal functioning, those needing replacement parts, and those with weighting mechanisms that have become unbalanced. Finally, the owner of the repair shop, when reporting back to various manufacturers about the causes of clock malfunctions, might classify the clocks etiologically. That is, she might report about clocks that have had little care, clocks that become worn over various time intervals of ownership, and clocks that developed problems after being moved or damaged.

Which of these classifications—descriptive, cost oriented, anatomical, or etiological—is the true or best classification of defective grandfather clocks? From the nominalist perspective, none of these classifications is inherently the best. Each of these classifications is imposed by the needs of the particular individual using the classification. Each classification serves a function. For any particular function, one classificatory system may be preferable. But none of these is the true classification of defective clocks.

Notice that this apocryphal classification of defective clocks is analogous to the approaches to defining disease suggested by Scadding: clinical-descriptive, morbid anatomical and etiological. The cost oriented classification was simply added as an analogy to how medical classifications are used by the insurance industry in the USA.

Wulff's defective clock analogy was borrowed from the British philosopher, John Locke. Locke had argued that classificatory systems are inherently nominalist, even though the ultimate goal is often essentialist:

> Therefore we in vain pretend to rank things in sorts, and dispose them into certain classes, under names, by their *real essences*, that are so far from our discovery or comprehension. (Wulff et al., 1986, p. 75)

In this regard, it is interesting to contrast Locke's view of classification to those of his physician friend, Thomas Sydenham. Believing in an essentialist view of disease, Sydenham made the following statement which has been quoted repeatedly since then:

> Nature, in the production of disease, is uniform and consistent ... The selfsame phenomena that you observe in the sickness of a Socrates you would observe in the sickness of a simpleton.

In other words, diseases do exist. They do have an essence. It is the business of medical research to discover what these essences are. It was the belief in this essentialist perspective that led nineteenth century researchers to solve the etiological issues associated with dementia paralytica.

Scadding and Wulff et al. warned about the dangers of essentialist thinking when applied to disease. For instance, Scadding noted that often we mistake the disease for the cause of the disease. Noguchi and Moore (Quetel, 1990) discovered the syphilitic bacilli in the brains of individuals with paresis. Hence, we might say that paresis occurs when syphilis invades the central nervous system. But the last sentence is misleading. Syphilis is not an organism. The bacteria, *Treponema pallidum*, is an organism and it could be said to invade the central nervous system. But even if this bacteria were present in the brain of an individual, that presence does not mean that the individual has paresis. To have paresis the individual must manifest the characteristic symptoms and anatomical changes associated with paresis.

Wulff et al. end their discussion of nominalism vs. essentialism with the following statement:

> The philosophical problem which underlies the discussion in this chapter is the age-old dispute about universals, and we have tried to navigate between the Scylla of essentialism (or Platonism) and the Charybdis of extreme nominalism. Essentialism underlines correctly that any classification of natural phenomena must reflect the realities of nature, but it ignores the fact that classifications also depend on our choice of criteria and that this choice reflects our practical interests and the extent of our knowledge. Nominalism, on the other hand, stresses correctly the human factor, but the extreme nominalist overlooks that classifications are not arbitrary but must be molded on reality as it is. (Wulff et al., 1986, pp. 87–88)

As mentioned earlier, defining the concepts of "disease" and/or "disorder" raise complicated issues, and the preceding discussion does not adequately cover the literature. One American author, in particular, has attracted substantial attention in the 1990s for his writings on this issue. Wakefield (1992, 1993) initially addressed this issue by providing a detailed critique of the definition of mental disorder that appeared in the *DSM-III*. Following this seminal paper, other theoretical papers (Wakefield, 1997a, 1997b) proposed a "harmful dysfunction" view of how to define mental disorders. A special section of the *Journal of Abnormal Psychology* has been devoted to a discussion of Wakefield's ideas .

Besides Wakefield's writings, there are other important discussions of this definitional issue including an overview by Reznek (1987), a book by Wing (1978), and an edited book on philosophical issues associated with classification (Sadler, Wiggins & Schwartz, 1994).

4.03.9.3 Two Views of a Hierarchical System of Diagnostic Concepts

Categories in the classification of mental disorders are organized hierarchically. This structural arrangement is commonly recognized but, since the publication of the *DSM-III*, two different views about this hierarchical structure have been discussed. Since these two views are often confused, the next section briefly discussed them.

The first approach to the meaning of hierarchy is nested set approach. Consider, for instance, the *DSM-II* classification of mental disorders. In this system, there are two broad categories of disorders: (I) psychotic disorders and (II) nonpsychotic disorders. The psychotic disorders are further subdivided into (I.A) the organic disorders and (I.B) the nonorganic disorders. The nonorganic psychotic disorders in the *DSM-II* were subdivided into three categories: (I.B.1) schizophrenic disorders, (I.B.2) major affective disorders, and (I.B.3) paranoid disorders. Then the schizophrenic disorders were subdivided into various subtypes: (I.B.1.a) simple type, (I.B.1.b) hebephrenic type, (I.B.1.c) catatonic type, and so on. Notice that this organization of mental disorders has a similar outline to the organization of categories in the biological classification. Any patient, for instance, who would be diagnosed as being hebephrenic, would also be considered as having schizophrenia disorder (the next highest, inclusive category) as well as having a psychotic disorder (an even higher level, inclusive category). Thus, this approach to hierarchy is called a nested set approach because the categories low in the system refer to sets of patients that are included (nested) in higher order categories. This approach parallels the classification of biological organisms in which any animal who is a member of the species *Felis catus* (housecat) is also member of the genus *Felis* (cats) and a member of an even higher order category of *Mammalia* (mammal).

The other approach to hierarchy is called a pecking order view. This view can be best understood by making analogy to the hierarchical organization of rank among military officers. A colonel is higher in rank than a major who in turn is higher in rank than a lieutenant. In this pecking order structure, a colonel can give orders to a lieutenant, but a lieutenant cannot issue orders to a colonel. Thus, the pecking order in military rank concerns lines of authority. A colonel has authority over a lieutenant. Notice, however, that there is no membership nesting in these categories. If a particular individual is a lieutenant, then that individual is not a colonel even though a colonel is higher in the hierarchy than a lieutenant.

To understand how this analogy to the hierarchical arrangement of military rank can be applied to psychiatric classification, consider the following order of general mental disorders:

organic mental disorders
schizophrenic disorders
affective disorders
anxiety disorders
personality disorders.

In terms of the pecking order meaning of hierarchy, this order means that disorders higher in this order should be diagnosed over disorders lower in the hierarchy. Thus, in terms of standard diagnostic practice, the presence of organic mental disorders should be ruled out first, then the schizophrenic disorders, then affective disorders, etc. This principle of diagnostic precedence is analogous to the authority relationship among different levels of rank in the military.

Notice that the pecking order relationship among these five general mental disorders also carries another implication. If a patient has an organic mental disorder, the patient can (and often does) have the symptoms associated with disorders that are lower in the hierarchy. Thus, a patient with Alzheimer's disease can develop hallucinations like a schizophrenic, can have marked sleep disturbance like someone who is depressed, can develop a fear of leaving the house like someone with anxiety disorder, and show the rigidity and need to be controlled like someone with an obsessive-compulsive personality disorder. However, a patient with anxiety disorder such as agoraphobia should not show the disturbed memory patterns and disorientation of a patient with an organic mental disorder. The important point to note is that disorders placed higher in this pecking order view of hierarchy can explain any of the symptoms of disorders lower in the hierarchy; however, the reverse should not occur. There should be symptoms that will be manifest in patients with schizophrenia that will not occur in patients with personality disorders.

The pecking order approach to the hierarchical arrangement of mental disorder categories was popularized by Foulds and Bedford (1975). The specifics of their approach to the classification of mental disorders differs from that presented above, but the general outline is the same. An important corollary to this pecking order view of the hierarchical arrangement among mental disorder categories is that this view suggests that there will be a strong severity dimension in any descriptive approach to the classification of these disorders. Mental disorders higher in this system will be associated with many more symptoms than are mental disorders lower in this system. Descriptive studies of psychopathology have repeatedly found a strong severity dimension that the pecking order view would predict.

4.03.9.4 Problem of Diagnostic Overlap/Comorbidity

When discussing the categorical model of classification, one of the tenets that was attributed to the model stated: "In the borderline areas where categories may overlap, the number of overlapping patients should be relatively small." Diagnostic overlap refers to the relative percentage of patients with one diagnosis who also meet the criteria for another diagnosis. As the tenet above states, some diagnostic overlap is expected. But the relative amount of overlap should be small.

One terminological note should be made before proceeding. The literature on this issue is grouped under the general heading of comorbidity. This term is from the medical literature, because it is well recognized that some medical disorders tend to go together. For instance, individuals who develop AIDS are relatively likely to develop yeast infections, sarcomas, and other disorders because of their compromised autoimmune system. The term comorbidity refers to the pattern of co-occurrences of these medical disorders. However, because the concept of comorbidity implies the acceptance of a disease model, the preferred term in this chapter will be "diagnostic overlap."

One of the earliest studies that focused on diagnostic overlap was by Barlow, DiNardo, Vermilyea, Vermilyea, and Blanchard (1986).

These investigators reported on 126 patients who were referred for the treatment of anxiety. These patients were administered structured interviews. Of the 126 patients interviewed, 108 were assigned one of seven diagnoses that fit within the anxiety/affective disorder spectrum. Of these 108 patients 65% were given at least one additional diagnosis. This is a high level of diagostic overlap and apparently was much higher than these researchers had expected *a priori*.

A large number of other empirical studies have confirmed the high levels of diagnostic overlap using the *DSM-III* and subsequent classifications. Many of these studies are discussed in an excellent review by Clark, Watson, and Reynolds (1995). Examples of the results found in their review are listed below. For example, these reviewers noted one study of personality disorder diagnoses in a state hospital population that found that these patients met the diagnostic criteria for an average of 3.75 Axis II disorders. In addition to the personality disorders, the depressive disorders also have striking overlap with many other disorders. For instance, over half of the patients with major depressive disorder as well as patients with dysthymic disorder were found to have at least one co-occuring mental disorder. Depression shows significant overlap even with disorders that one might not expect overlap. For instance, a sample of antisocial personality disorder patients showed that one-third of these individuals also had a depressive diagnosis (Dinwiddie & Reich, 1993). These antisocial patients, less surprisingly, also had high rates of alcoholism (76%) and other substance use (63%). Even broad epidemiological studies on normal community samples show high rates of diagnostic overlap. Clark et al. reported that, in two national studies, over half of the individuals who had one mental disorder diagnosis had at least one more diagnosis. In some samples, the rate of diagnostic overlap is even higher. For instance, a study of suicidal patients showed that these individuals averaged about four mental disorders (Rudd, Dahm, & Rajab, 1993).

Together, these and many other studies suggest that the number of overlapping patients among mental disorder diagnoses is not small at all. Instead, diagnostic overlap is a standardly occurring phenomenon. Blashfield, McElroy, Pfohl, and Blum (1994) studied 151 patients who had been administered a structured interview to assess personality disorders. In this sample, only 24% of the patients met the criteria for one and only one personality disorder. Exactly the same percentage met the diagnostic criteria for at least four personality disorders, with 6% meeting the criteria for six of the eleven *DSM-III-R* personality disorders! When Blashfield et al. attempted to identify prototypic patients (i.e., individuals who met at least eight of the diagnostic criteria for a specific disorder), they found that only 15% of the patients would qualify as prototypic. However, most of these individuals also satisfied the diagnostic criteria for other disorders. When a prototypic patient was defined as an individual with eight or more criteria for a personality disorder and the lack of an additional personality diagnosis, only 1% of the patients were prototypes. In effect, patients with mixed symptom patterns are much more typical than are patients who represent relatively pure forms of a disorder.

Clark, Watson and Reynolds (1995) suggested that there are three related issues associated with the problem of diagnostic overlap. The first issue concerns the hierarchical organization of categories. As discussed earlier, one view of hierarchy is a pecking order approach. This view was implicitly adopted by the *DSM-III* because a number of exclusion rules were included in the diagnostic criteria for different disorders, so that a lower diagnosis would not be made if a higher order diagnosis was possible. However, research on the exclusionary criteria suggested that these criteria were arbitrary. Exclusionary criteria were mostly deleted in the *DSM-III-R* and the *DSM-IV*.

The second issue associated with the diagnostic overlap issue concerns the heterogeneity with diagnostic categories. In effect, the extensive diagnostic overlap suggests that the definitions of the various mental disorders are too broad and inclusive. Evidence of excessive heterogeneity comes from other sources, according to Clark et al. For instance, direct studies of the variability in symptom patterns have shown high levels of variability within disorders such as schizophrenia, depression, eating disorders and anxiety disorders. Another line of evidence of heterogeneity is the frequency with which mixed or atypical diagnoses are used. For instance, Mezzich, Fabrega, Coffman, and Haley (1989) found that the majority of patients with a dissociative disorder fit the criteria for an atypical dissociative disorder.

The third issue associated with the comorbidity finding is the increasing support associated with replacing the categorical approach to classification with a dimensional view. In the discussion of these two models, it was noted that the dimensional model is the simpler of the two. Unless there is clear evidence of the existence of discrete categories, a dimensional approach is the more parsimonious.

A number of researchers, when confronted with high rates of diagnostic overlap, have suggested that a dimensional model should be used. For instance, in the area of the personality disorders where some of the highest rates of diagnostic overlap have been found, interest in a "Big Five" dimensional approach to the personality disorders has been attracting increasing support. Another example in which a dimensional model is gaining popularity is in the subclassification of schizophrenia. The classic Kraepelinian subtypes do not generate sufficiently homogeneous categories. Various dimensional schemes (process–reactive, paranoid–nonparanoid, positive vs. negative symptoms) have more empirical support than the use of categorical subtypes.

4.03.10 CONCLUDING COMMENTS

This chapter provides a rather simplified overview of the issues associated with the classification of psychopathology. It attempts to help the reader gain a better understanding of classification by starting with a reasonably detailed history of classificatory systems, and thereby give some idea how many of the features of contemporary classificatory systems have evolved. The text also presents a succinct and readable presentation of four issues that currently attract a reasonable degree of attention in the literature. However, justice has not been done to many of the other complex issues that face both clinicians and scientists interested in classification, such as whether to use semistructured interviews as the "gold standard" for measurement (Mezzich, 1984), the role of values in the diagnostic practice of mental health professionals, whether or not certain mental disorders are sexually or racially biased (Busfield, 1996; Nuckolls, 1992; Widom, 1984), the relevance of life-span measures of psychopathology (Roff & Ricks, 1970), and the problem of focusing on the individual patient as the basic unit of psychopathology (as opposed to families or systems or interpersonal relationship patterns) (Clarkin & Miklowitz, 1997; Francis, Clarkin & Ross, 1997; Williams, 1997)

Like any general topic, the classification of psychopathology becomes a very complex topic when analyzed in detail.

> Perhaps I believe that the world can get forward most by clearer and clearer definitions of fundamentals. Accordingly I propose to stick to the tasks of nomenclature and terminology, unpopular and ridicule-provoking though they may be. (Southard, as quoted by Menninger, 1963, p. 3)

4.03.11 REFERENCES

American Psychiatric Association (1933). Notes and comments: Revised classified nomenclature of mental disorders. *American Journal of Psychiatry, 90,* 1369–1376.

American Psychiatric Association (1952). *Diagnostic and statistical manual of mental disorders* (1st ed.). Washington, DC: Author.

American Psychiatric Association (1968). *Diagnostic and statistical manual of mental disorders* (2nd ed.). Washington, DC: Author.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. Rev.). Washington, DC: American Psychiatric Press.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Press.

Alexander, F. G., & Selesnick, S. T. (1966). *The history of psychiatry*. New York: Harper and Row.

Austin, T. J. (1859/1976). *A practical account of general paralysis*. New York: Arno Press.

Barlow, D. H., DiNardo, P. A., Vermilyea, B. B., Vermilyea, J., & Blanchard, E. B. (1986). Co-morbidity and depression among the anxiety disorders: Issues in diagnosis and classification. *Journal of Nervous and Mental Disease, 174*, 63–72.

Barsalou, L. W. (1992). *Cognitive psychology: An overview for cognitive scientists*. Hillsdale, NJ: Erlbaum.

Bayer, R. (1981). *Homosexuality and American psychiatry*. New York: Basic Books.

Berrios, G. E. (1993). Personality disorders: A conceptual history. In P. Tyrer & G. Stein (Eds.), *Personality disorder reviewed* (pp. 17–41). London: Gaskell.

Berrios, G. E., & Hauser, R. (1988). The early development of Kraepelin's ideas on classification: A conceptual history. *Psychological Medicine, 18*, 813–821.

Blashfield, R. K. (1991). Models of psychiatric classification. In M. Hersen & S. M. Turner (Eds.), *Adult psychopathology and diagnosis* (2nd ed. pp. 3–22). New York: Wiley.

Blashfield, R. K., McElroy, R. A., Pfohl, B., & Blum, N. (1994). Comorbidity and the prototype model. *Clinical Psychology: Science and Practice, 1*, 96–99.

Blustein, B. E. (1991). *Preserve your love for science: Life of William A. Hammond, American neurologist*. Cambridge, UK: Cambridge University Press.

Braslow, J. T. (1995). Effect of therapeutic innovation on perception of disease and the doctor–patient relationship: A history of general paralysis of the insane and malaria fever therapy, 1910–1950. *American Journal of Psychiatry, 152,* 660–665.

Busfield, J. (1996). *Men, women and madness: Understanding gender and mental disorder*. New York: New York University Press.

Campbell, E. J. M., Scadding, J. G., & Roberts, R. J. (1979). The concept of disease. *British Medical Journal, 2,* 757–762.

Cantor, N., Smith, E. E., French, R. D., & Mezzich, J. (1980). Psychiatric diagnosis as a prototype categorization. *Journal of Abnormal Psychology, 89,* 181–193.

Clark, L. A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: Challenges to the current system and future directions. *Annual Review of Psychology, 46,* 121–153.

Clarkin, J. F., & Mikowitz, D. J. (1997). Marital and family communication difficulties. In T. A. Widiger, A. J. Francis, H. A. Pincus, R. Ross, M. B. First, & W. Davis (Eds.), *DSM-IV sourcebook* (Vol. 3, pp. 631–672). Washington, DC: American Psychiatric Press.

Dinwiddie, S. H., & Reich, T. (1993). Attribution of antisocial symptoms in coexistent antisocial personality disorder and substance abuse. *Comprehensive Psychiatry, 34,* 235–242.

Essen-Moller, E. (1971). Suggestions for further improvement of the international classification of mental disorders. *Psychological Medicine, 1,* 308–311.

Everitt, B. S. (1974). *Cluster analysis.* New York: Halstead Press.

Feighner, J. P., Robins, E., Guze, S., Woodruff, R. A., Winokur, G., & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry, 143,* 57–63.

Feinstein, A. R. (1967). *Clinical judgment.* Huntington, VA: Krieger.

Foulds, G. A., & Bedford A. (1975). Hierarchy of classes of personal illness. *Psychological Medicine, 5,* 181–192.

Francis, A. J., Clarkin, J. F., & Ross, R. (1997). Family/relational problems. In T. A. Widiger, A. J. Francis, H. A. Pincus, R. Ross, M. B. First, & W. Davis (Eds.), *DSM-IV sourcebook* (Vol. 3, pp. 521–530). Washington, DC: American Psychiatric Press.

Goffman, E. (1961). *Asylums.* London: Penguin.

Grinker, R. R., Werble, B., & Drye, R. C. (1968). *The borderline syndrome.* New York: Basic Books.

Hammond, W. A. (1883). *A treatise on insanity in its medical relations.* New York: Appleton.

Hempel, C. G. (1965). *Aspects of scientific explanation.* New York: Free Press.

Hull, D. L. (1988). *Science as a process.* Chicago: University of Chicago Press.

Jardine, N., & Sibson, R. (1971). *Mathematical taxonomy.* New York: Wiley.

Kendell, R. E. (1975). *The role of diagnosis in psychiatry.* Oxford, UK: Blackwell.

Kendell, R. E. (1976). The concept of disease. *British Journal of Psychiatry, 128,* 508–509.

Kendell, R. E. (1979). Alcoholism: A medical or a political problem. *British Medical Journal, 1,* 367–381.

Kendell, R. E. (1986). What are mental disorders? In A. M. Freedman, R. Brotman, I. Silverman, & D. Huston (Eds.), *Issues in psychiatric classification* (pp. 23–45). New York: Human Sciences Press.

Kendell, R. E. (1991). Relationship between the DSM-IV and ICD-10. *Journal of Abnormal Psychology, 100,* 297–301.

Kirk, S. A., & Kutchins, H. (1992). *The selling of DSM: The rhetoric of science in psychiatry.* Hawthorne, NY: Walter deGruyter.

Klerman, G. L. (1978). The evolution of a scientific nosology. In J. C. Shershow (Ed.), *Schizophrenia: Science and practice* (pp. 99–121). Cambridge, MA: Harvard University Press.

Kraepelin, E. (1902/1896). *Clinical psychiatry: A text-book for students and physicians* (6th ed., translated by A. R. Diefendorf). London: Macmillan.

Kreitman, N. (1961). The reliability of psychiatric diagnosis. *Journal of Mental Science, 107,* 878–886.

Lacan, J. (1977). *Ecruits: A selection.* New York: Norton.

Laing, R. D. (1967). *The politics of experience.* London: Penguin.

Lorr, M. (1966). *Explorations in typing psychotics.* New York: Pergamon.

Matza, D. (1969). *Becoming deviant.* Englewood Cliffs, NJ: Prentice-Hall.

Meehl, P. E. (1995). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist, 50,* 266–275.

Menninger, K. (1963). *The vital balance.* New York: Viking.

Mezzich, J. E. (1984). Diagnosis and classification. In S. M. Turner & M. Hersen (Eds.), *Adult psychopathology and diagnosis* (pp. 3–36). New York: Wiley.

Mezzich, J. E., Fabrega, H., Coffman, G. A., & Haley, R. (1989). DSM-III disorders in a large sample of psychiatric patients: Frequency and specificity of diagnosis. *American Journal of Psychiatry, 146,* 212–219.

Mezzich, J. E., Fabrega, H., Mezzich, A. C., & Coffman, G. A. (1985). International experience with DSM-III. *Journal of Nervous and Mental Disease, 173,* 738–741

Nelson, G., & Platnick, N. (1981). *Systematics and biogeography: Cladistics and vicariance.* New York: Columbia University Press.

Nuckolls, C. W. (1992). Toward a cultural history of the personality disorders. *Social Science and Medicine, 35,* 37–47.

Plaut, F. (1911). *The Wasserman sero-diagnosis of syphilis in its application to psychiatry.* New York: Journal of Nervous and Mental Disease Publishing Company.

Project Match Research Group (1997). Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol, 58,* 7–29.

Quetel, C. (1990). *History of syphilis.* Baltimore: Johns Hopkins University Press.

Raines, G. N. (1952). Forward. In American Psychiatric Association, *Diagnostic and statistical manual of mental disorders* (1st ed., pp. v–xi). Washington, DC. American Psychiatric Assocation.

Regier, D. A., Kaelber, C. T., Roper, M. T., Rae, D. S., & Sartorius, N. (1994). The ICD-10 clinical field trial for mental and behavioral disorders: Results in Canada and the United States. *American Journal of Psychiatry, 151,* 1340–1350.

Reznek, L. (1987). *The nature of disease.* London: Routledge & Kegan Paul.

Roff, M., & Ricks, D. F. (Eds.) (1970). *Life history research in psychopathology.* Minneapolis, MN: University of Minnesota Press.

Rosenhan, D. L. (1973). On being sane in insane places. *Science, 179,* 250–258.

Roth, A., & Fonagy, P. (1996). *What works for whom: A critical review of psychotherapy research.* New York: Guilford.

Rudd, M. D., Dahm, P. F., & Rajab, M. H. (1993). Diagnostic comorbidity in persons with suicidal ideation and behavior. *American Journal of Psychiatry, 147,* 1025–1028.

Russell, J. A. (1991). In defense of a prototype approach to emotion concepts. *Journal of Personality and Social Psychology, 60,* 37–47.

Russell, J. A., & Fehr, B. (1994). Fuzzy concepts in a fuzzy hierarchy: Varieties of anger. *Journal of Personality and Social Psychology, 67,* 186–205.

Rutter, M., Shaffer, D., & Shepard, M. (1975). *A multiaxial classification of child psychiatry disorders.* Geneva, Switzerland: World Health Organization.

Sadler, J. Z., Wiggins, O. P., & Schwartz, M. A. (1994). *Philosophical perspectives on psychiatric diagnostic classification.* Baltimore: Johns Hopkins University Press.

Sarbin, T. R. (1997). On the futility of psychiatric diagnostic manuals (DSMs) and the return of personal agency. *Applied and Preventive Psychology, 6,* 233–243.

Scadding, J. G. (1959). Principles of definition in medicine with special reference to chronic bronchitis and emphysema. *Lancet, 1,* 323–325.

Scadding, J. G. (1969). Diagnosis: The clinician and the computer. *Lancet, 2,* 877–882.

Skinner, H. A. (1986). Construct validation approach to psychiatric classification. In T. Millon & G. L. Klerman (Eds.), *Contemporary directions in psychopathology: Toward the DSM-IV* (pp. 307–331). New York: Guilford Press.

Sneath, P. H. A., & Sokal, R. R. (1973). *Numerical taxonomy.* San Francisco: Freeman.

Spitzer, R. L., Endicott, J., & Robins, E. (1975). Research diagnostic criteria. *Archives of General Psychiatry, 35*, 773–782.

Spitzer, R. L., & Fleiss, J. L. (1974). A re-analysis of the reliability of psychiatric diagnosis. *British Journal of Psychiatry, 125*, 341–347.

Spitzer, R. L., & Williams, J. B. W. (1987). Revising DSM-III: The process and major issues. In G. L. Tischler (Ed.), *Diagnosis and classification in psychiatry* (pp. 425–434). New York: Cambridge University Press.

Spitzka, E. C. (1883). *Insanity: Its classification, diagnosis and treatment*. New York: Bermingham.

Stangl, D., Pfohl, B., Zimmerman, M., Bowers, W., & Corenthal, C. (1985). Structured interview for the DSM-III personality disorders. *Archives of General Psychiatry, 42,* 519–596.

Stengel, E. (1959). Classification of mental disorders. *Bulletin of the World Health Organization, 21,* 601–663.

Stromgren, E. (1991). A European perspective on the conceptual approaches to psychopathology. In A. Kerr & H. McClelland (Eds.), *Concepts of mental disorders: A continuing debate* (pp. 84–90). London: Gaskell.

Szasz, T. (1961). *The myth of mental illness*. New York: Hoeber-Harper.

Veith, I. (1965). *Hysteria: The history of a disease*. Chicago: University of Chicago Press.

Wakefield, J. C. (1992). The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist, 47,* 373–388.

Wakefield, J. C. (1993). Limits of operationalization: A critique of Spitzer and Endicott's (1978) proposed operational criteria for mental disorder. *Journal of Abnormal Psychology, 102,* 160–172.

Wakefield, J. C. (1997a). Diagnosing DSM-IV—Part I: DSM-IV and the concept of disorder. *Behavioral Research and Therapy, 35,* 633–649.

Wakefield, J. C. (1997b). Diagnosing DSM-IV—Part II: Eysenck (1986) and the essentialist fallacy. *Behavioral Research and Therapy, 35,* 651–665.

Walker, L. (1987). Inadequacies of the masochistic personality disorder diagnosis for women. *Journal of Personality Disorders, 1,* 183–189.

Widom, C. (Ed.) (1984). *Sex, roles and psychopathology*. New York: Plenum.

Williams, J. B. W. (1997). The DSM-IV multiaxial system. In T. A. Widiger, A. J. Francis, H. A. Pincus, R. Ross, M. B. First, & W. Davis (Eds.), *DSM-IV sourcebook* (Vol. 3). Washington, DC: American Psychiatric Press.

Wing, J. K. (1978). *Reasoning about madness*. Oxford, UK: Oxford University Press.

Woodruff, R. A., Goodwin, D. W., & Guze, S. B. (1974). *Psychiatric diagnosis*. New York: Oxford University Press.

World Health Organization (1948). *Manual of the international statistical classification of diseases, injuries, and causes of death*. Geneva, Switzerland: Author.

World Health Organization (1957). Introduction to *Manual of the international statistical classification of diseases, injuries, and causes of death* (7th ed.). Geneva, Switzerland: Author.

World Health Organization (1978). *Mental disorders: Glossary and guide to their classification in accordance with the ninth revision to the International Classification of Diseases*. Geneva, Switzerland: Author.

World Health Organization (1992). *The ICD-10 classification of mental and behavioral disorders: Clinical descriptions and diagnostic guidelines*. Geneva, Switzerland: Author.

World Health Organization (1993). *The ICD-10 classification of mental and behavioural disorders: Diagnostic criteria for research*. Geneva, Switzerland: Author.

Wulff, H. R., Pedersen, S. A., & Rosenberg, R. (1986). *Philosophy of medicine*. Boston: Blackwell Scientific.

Wyer, R. S., & Srull, T. K. (1989). *Memory and cognition in a social context*. Hillsdale, NJ: Erlbaum.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*, 338–353.

Zubin, J. (1967). Classification of behavior disorders. *Annual Review of Psychology*, *28*, 373–406.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.04
Clinical Interviewing

EDWARD L. COYLE
Oklahoma State Department of Health, Oklahoma City, OK, USA
and
DIANE J. WILLIS, WILLIAM R. LEBER, and JAN L. CULBERTSON
University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

4.04.1 PURPOSE OF THE CLINICAL INTERVIEW 82
- *4.04.1.1 Gathering Information for Assessment and Treatment* 82
- *4.04.1.2 Establishing Rapport for Assessment and Treatment* 82
- *4.04.1.3 Interpersonal Style/Skills of the Interviewer* 82
- *4.04.1.4 Structuring the Interview* 83
  - *4.04.1.4.1 Setting variables* 83
  - *4.04.1.4.2 Preparing for the patient* 84
- *4.04.1.5 Introductory Remarks* 84
- *4.04.1.6 How to Open the Interview* 85
- *4.04.1.7 The Central Portion of the Interview* 85
- *4.04.1.8 Closing the Interview* 86
- *4.04.1.9 The Collateral Interview* 86

4.04.2 DEVELOPMENTAL CONSIDERATIONS IN INTERVIEWING 87
- *4.04.2.1 Interviewing Children (Preschool Age through Older Elementary)* 87
- *4.04.2.2 Interviewing Parents* 87
- *4.04.2.3 Social Context* 88
- *4.04.2.4 Developmental Context* 89
- *4.04.2.5 Direct Interview of Children* 89
- *4.04.2.6 Adolescents* 90
  - *4.04.2.6.1 Separation–individuation* 91
  - *4.04.2.6.2 Resolving conflict with authority figures* 91
  - *4.04.2.6.3 Peer group identification* 91
  - *4.04.2.6.4 Realistic appraisal and evaluation of self-qualities* 92
- *4.04.2.7 Interviewing Young Adults (18–40 Years)* 92
- *4.04.2.8 Interviewing Adults in Middle Adulthood (40–60 Years)* 93
- *4.04.2.9 Interviewing Older Adults (60–70 Years)* 93
- *4.04.2.10 Interviewing In Late Adulthood (70 Years–End of Life)* 93

4.04.3 INTERVIEWING SPECIAL POPULATIONS OR SPECIFIC DISORDERS 94
- *4.04.3.1 Interviewing Depressed Patients* 94
- *4.04.3.2 Interviewing Anxious Patients* 94

4.04.4 SUMMARY 95

4.04.5 REFERENCES 96

4.04.1 PURPOSE OF THE CLINICAL INTERVIEW

The clinical interview is extremely important as a diagnostic tool in the assessment and treatment of patients. Clinicians who do thorough and competent interviews have a much better understanding of the developmental course of symptoms presented by the patient. Indeed, before there were any personality inventories, before the Rorschach and one-way mirror behavioral observation, there was the clinical interview. The purpose of the clinical interview is to gain sufficient information from the informant or informants to formulate a diagnosis, assess the individual's strengths and liabilities, assess the developmental and contextual factors that influence the presenting concerns, and to allow planning for any interventions to follow. The interview is in many instances the ultimate clinical tool, and effective interviewing must be an integral part of any clinician's professional abilities. Although the clinical interview is used primarily to gather information for clinical evaluation or psychotherapeutic treatment, it can also serve the purpose of preparing the patient for therapy, and less frequently the interview process itself provides some relief from psychological distress. The interview may be performed in many different settings: outpatient private practice, community mental health center, psychiatric hospital, prison, emergency or medical hospital room, school, and others.

While the amount of time devoted to the interview, the setting, and the purposes may vary, the features of an effective clinical interview remain the same. When completed, the interviewer has created a relatively comprehensive portrait of the patient that can be communicated readily to others and will provide the basis for making important judgments about the subject of the interview. The relative importance of various symptoms or concerns should be established, along with an estimate of the individual's overall functioning. The relative importance of various symptoms or concerns should be established, and some estimate of the individual's responses in a variety of settings can be made with an acceptable degree of validity. These features can be said to be a part of any clinical interview. Some specific purposes of the interview are described next, along with suggestions about different approaches and emphases the clinician may be required to take.

4.04.1.1 Gathering Information for Assessment and Treatment

The most common purposes of the clinical interview are to gather information to establish a diagnosis, evaluate mental status and historical data that impact upon the individual, and provide a full understanding of the important personality, biological, and environmental variables that have brought the patient to this point. All treatment planning begins with some type of formal or informal evaluation. The clinical interview is the most effective way to gain an understanding of the current functioning and the difficulties faced by the patient, and is a necessary adjunct to the data gathered from other assessment approaches. In the clinical interview, the clinician inquires directly and in a focused manner about the patient's development, adaptation, and current difficulties. When the interview is part of a comprehensive evaluation, some features may be emphasized as a result of the specific referral reason that would not be as prominent in the interview conducted for psychotherapy treatment planning. As an example, the interview conducted for an initial psychoeducational evaluation of an elementary grade school child to determine reasons for school failure is likely to entail considerable emphasis upon academic and learning history and the involvement of one or more of the child's teachers as collateral informants. If the same child were to present later for psychotherapeutic interventions to address the depression and oppositional behaviour problems identified by previous evaluation to be the cause of his academic failure, the interviewer would likely spend more time and effort in determining family interactions, parenting skills, and social supports.

4.04.1.2 Establishing Rapport for Assessment and Treatment

One function of the clinical interview is to prepare the patient for the clinical interventions that follow, including additional formal assessment procedures. In order to obtain valid psychometric data, the patient must be adequately cooperative and invested in the testing process. The interview can help the clinician achieve this end by providing a sense of professional intimacy and a feeling of compassion and interest in the patient's well-being. Thus prepared, the respondent is more willing to give themself over to the process, and to perceive it as being something that will provide them with some beneficial outcome.

4.04.1.3 Interpersonal Style/Skills of the Interviewer

While the basic purpose of gathering relatively concrete information may be accomplished by individuals with a minimum of training and sensitivity, there are a number of personal qualities that tend to improve the quality of information gained and to result in a more helpful and pleasant experience on the part of the informant. Chief among these is the quality of empathy, which must be readily recognized by the informant through subtle and overt communications from the interviewer.

Empathy means identifying with and understanding someone else's feelings, motives, or word view. It means entering the private perceptual world of another person and being at home in it—to be sensitive to what they feel (Egan, 1994; Luborsky, 1996). An intellectual understanding of empathy, however, does not provide one with the interpersonal skills and experience that result in the ability to truly resonate to the informant's experience and to respond in ways that will ease the flow of information of a personal and often sensitive nature. The skill and art of attuning oneself not only to the overt communications of the patient, but also to the underlying feelings and meanings, must become a continuing focus of attention for the interviewing clinician. While much of this process is not fully accessible to conscious awareness, there are some components that lend themselves readily to examination and modification. For example, the interviewer's responses that communicate to the informant negative value judgments are perhaps more easily modified.

Although the mental health fields and their practitioners have often been vilified for their purported moral relativism, no reasonable clinician would believe himself or herself to be free of individual prejudices and deeply-held convictions regarding right and wrong. These values are a part of each person, and to truly expunge them would result in an insipid and ineffective shell of a human being. The relevance to a discussion of clinical interviewing is this: the effective interviewer takes care to be aware of his or her own expectations and biases regarding human behaviour and strives to avoid making explicit negative judgments of the informant in order to provide a comfortable and supportive environment during the interview. This skill can be and is developed and improved through careful attention to the process, internal changes within the interviewer during the interaction, and by effective supervision and review of actual interviews with other clinicians.

Often such judgments can be communicated to the respondent with no more than a change in facial expression or in a shift in questioning. Specific wording and follow-up questions sometimes can have the effect of casting a chill upon the interview process. For example, the interviewer who learns the informant is homosexual and then avoids asking questions about sexual functioning and current relationships readily communicates their discomfort with an important aspect of the informant's personality. The effective interviewer does not perfect an unexpressive mask, but does develop the ability to decrease the immediate translation of visceral responses into explicit behaviours during the interview. Introspection about areas that increase the clinician's anxiety and honest confrontation about discriminatory beliefs are necessary if one is to perform clinical tasks competently and ethically.

4.04.1.4 Structuring the Interview

As with any therapeutic or evaluative intervention, the setting and structure of the interview have a significant effect on the outcome of the interaction. Because the actual face-to-face time spent with patients must be as productive and positive as possible, the clinician should take care to prepare for the clinical interview prior to contact. While the goals of the interview may vary somewhat as discussed previously, many factors common to all clinical interviews should at least be in one's mind prior to and during the interview.

4.04.1.4.1 Setting variables

Some basic attention should be given to simple environmental factors when preparing for the interview. Although many fruitful interviews have been conducted with patients, families, and other sources under conditions that might be charitably described as less-than-optimal, doing so in a comfortable and soothing environment will often add to an informant's ease in discussing delicate and/or emotionally charged matters. Seating accommodations should be given consideration, as hard, uncomfortable, or rickety, precarious seats can add a tinge of anxiety or discomfort to the gestalt of the interview process and thus to the evaluation and/or treatment that follows the interview. It should go without saying that the space being used for the clinical interview should be held relatively inviolate from intrusions, including external noises and conversation to the degree possible. While most people are able to tolerate minor interruptions such as a ringing telephone, having another clinician open the door while your patient is tearfully recounting a past trauma is likely to be somewhat harmful to the tentative alliance you are developing. Therefore, if you work in a setting with multiple users you will do well to take precautions to avoid such disruptions. A white-sound generator can help decrease the penetration of external sounds and add somewhat to the intimacy of the interaction. Throughout the interview the clinician will be carefully observing the behaviours of the subject, noting congruencies and incongruencies, attending to shifts in voice and posture. One sometimes overlooked source of information that may add to the interview process is that of behavioral observations made while the patient and collaterals are in the waiting area of the clinic or office. Often it is possible to observe interactions and general demeanor while you organize paperwork or make other preparations before formally introducing yourself. It may then be helpful to comment during interview on some salient interaction or response. Of course, as has been particularly noted by various custody and forensic evaluators (Bricklin, 1990), the behavior in the waiting room must not be taken as necessarily representative of the person's usual response outside of the clinic. However, one often can observe telling interactional patterns, particularly between parents and children, and this may provide opportunity for addressing problematic areas during the subsequent interview.

4.04.1.4.2 Preparing for the patient

It is common practice now to present the patient with relatively extensive questionnaires or personal history forms prior to the first clinic visit. With this information in hand, the clinician may be able to focus in quickly on the salient symptomatology and current concerns. When available, this information should be used to tailor the interview, allotting time during the face-to-face interview in the most effective manner. If the clinician does choose to utilize such instruments, he or she would be well served to take the time necessary to review the data prior to entering the room with the informant. Watching the professional read and murmur over the forms for minutes while the informant is ignored or asked disconnected questions can be expected to result in a sense of devaluation for the informant. It also gives the impression of disorganization and lack of preparation on the part of the clinician. Neither of these will be helpful in the ensuing interview process.

4.04.1.5 Introductory Remarks

It is helpful to develop a standard approach to the clinical interview, including the introduction and beginning of the interview. One should introduce oneself, and give the informant information about the purpose of the interview and the expected duration. The interviewer's role and title should be clarified, and any supervisory or other training relationship must be disclosed prior to beginning the interview. It is essential that issues of confidentiality be fully addressed, and the informant be given opportunity and encouragement to ask questions about disclosure of information. If any of the data obtained will be shared with other individuals, this must be explained clearly. This is of particular importance in forensic or custody evaluations. When interviewing children and parents, keep in mind the fact that in many jurisdictions the noncustodial parent may retain full rights to examine medical records, including data from the clinical interview. Even if the informant has signed a general disclosure or consent for treatment, it is the clinician's ethical responsibility to review duties to warn and the possible limits on confidentiality. The legal definition of informed consent in many jurisdictions is not necessarily satisfied by the presence of a signature on a form, but rather is established by questioning the informant about their understanding at the time the information was given. The best practice is for the clinician to do their best to make certain that the person with whom they are communicating for professional purposes is fully informed of such issues. In illustration, imagine for a moment being 30 minutes into an interview with a man who informs you very clearly that he intends to use the pistol in his car to shoot his wife when he returns home. If you have fully informed him of the limits of confidentiality, you are in a very distressing situation. If you have not done so, your position is much worse. The growth of managed care and its attendant prospective treatment review process may complicate the ethical duties involved in the clinical interview. As Corcoran and Vandiver (1996) point out, "There can be no doubt that managed care has restricted clients autonomy and interferes with the confidential relationship" (p. 198). During the initial interview and prospective utilization review of a patient whose care is managed by a third (or possibly fourth) party, the clinician may find him or herself in the uncomfortable position of being more the agent of the managed-care organization than the advocate of the patient. In such relationships, it is imperative that the patient be made fully aware at the outset of the interview of the additional limits of confidentiality imposed by the managed-care entity. This may include multiple reviews of the data gained during interview and any subsequent treatment sessions. An additional ethical concern arises in the clinical interview with regard to the establishment of a professional relationship and responsibility for the clinical care of the patient. Does performing the clinical interview and prospective review obligate the clinician to provide service even if the managed-care entity denies authorization? Again, the only way to avoid difficulties, misunderstandings, and possible litigation or board complaints is to be absolutely clear with interviewees and any involved third-party payer about these issues prior to the professional contact. If it is possible that the interviewing clinician will not receive reimbursement from the managed-care company for services, any alternative financial arrangements should also be discussed with the prospective patient before any formal clinical contact. If there are inherent limitations to the number of sessions or type of interventions that are covered by the third-party payer, the potential client should also be made aware of these before ending the interview. Of course, it is possible that no treatment will be necessary; thus it seems sensible to leave discussing the mechanics of paying for it until it is determined to be needed.

4.04.1.6 How to Open the Interview

The best way to open the interview is with a very general, open-ended question about the circumstances that have brought the patient to the interview. Morrison (1995) recommends taking approximately eight to 10 minutes in the typical one-hour interview to allow the respondent to explain in their own words their needs and history. Morrison points out that, among other things, this provides the clinician an opportunity to obtain a true flavor for the respondent's personality and communication style, and to make general observations of behavior, affect, and thought process relatively free from the clinician's direction. An example of an opening question might be "Please tell me about the things that are concerning you most right now" or "I would like for you to tell me what you need some assistance with now" or even "Please give me an idea of how you came to be here today." The amount of information gathered during this portion of the interview will be to some degree dependent upon the respondent's intellectual ability and verbal facility. Many people are characterologically unwilling to self-disclose, even within the confines of the clinical interview, and may require additional urging. The clinician should generally respond to hesitations with supportive restatement of the opening question, or with gentle encouragement and reflection of any apprehension that is detected. If hesitation or lack of content appear to be due to cognitive limitations, disorientation, or distractibility, it may be helpful to ask more direct and close-ended questions based upon previously obtained information or the patient's brief verbalizations. It is generally not desirable to lead the patient any more than necessary, as the more you query the less likely you will be able to distinguish between accurate responses and those that are colored by the demands experienced by the patient. However, in some cases the clinician must take a more directive approach to complete the interview successfully.

The topics to be included in every interview are:

(i) Introduction, purpose of interview, credentials/role of interviewer;
(ii) Confidentiality and exceptions;
(iii) Presenting problems (preferably phrased in general, open-ended manner);
(iv) Mood/Anxiety symptoms;
(v) Impulse control and direct inquiry of suicidal ideation/history;
(vi) Current social, academic, and vocational functioning;
(vii) System of social support;
(viii) Environmental factors, including current basic needs/shelter;
(ix) Developmental factors (especially for children) that may influence symptom presentation;
(x) Medical history, including family health history and previous treatment/hospitalization;
(xi) Substance use;
(xii) Legal involvement and history; and
(xiii) Vegetative symptoms.

4.04.1.7 The Central Portion of the Interview

After the initial introduction, housekeeping, and rapport-building, it is time to focus upon the most salient features of the person being evaluated, and the circumstances that maintain the current dysfunction. Once the presenting problems have been identified and an adequate alliance with the respondent established, the clinician must utilize their knowledge of psychopathology and diagnostic criteria to fully understand and classify the presenting problems, as well as to identify the primary strengths and resources that will be drawn upon by the patient and the professionals involved in subsequent interventions. The central portion of the interview is dedicated to adding to the framework established by queries about the presenting problem. One mistake made by novice (as well as by some more seasoned but overly concrete) interviewers is to rigidly adhere to an interviewing framework, disregarding the natural flow of conversation.

If one is unable to recognize the more subtle verbal and nonverbal messages that should be probed and instead forces one's way forward, the clinician will end up with less information than they should. Thus, it is essential to attend carefully to shifts in mood during the interview, both within the patient and the interviewer.

Luborsky (1996) details the utilization of momentary shifts in mood during a therapy session to focus upon vital underlying thoughts that are salient to the therapeutic issues. The interviewing clinician can also benefit by noticing changes in voice tone, volume, and content of speech. During the central portion of the interview the clinician continues to focus on the problems and possible explanations for present distress. When possible, avoid becoming involved in digressive topics, as some respondents may prefer to spend most of the available time presenting problems that are not central to the services being sought. By the same token, it is the clinician's responsibility to follow any significant leads in the interview, and to be aware of any tendencies on their own part to avoid distressing topics. Experience shows that clinicians tend to be vulnerable to this type of error in particular with regard to sexual functioning, substance use, and racial/ethnic discrimination issues. It may be helpful to keep in mind that while the interview shares many commonalties with social conversation, it is by definition not a run-of-the-mill social interaction. Thus, inhibitions that prevent the interviewer from querying these admittedly uncomfortable topics must be dealt with. Because many clinicians may find themselves having completed much of their formal training without ever overcoming the discomfort experienced when such topics are broached, it may be necessary to practice on colleagues in role-play activity designed to help the clinician become adept at obtaining the necessary information despite initial resistance from within as well as from the respondent. As one of the primary purposes of the clinical interview is accurate diagnosis according to current syndromal criteria from the *Diagnostic and statistical manual of mental disorders* (4th ed., *DSM-IV*), the clinician must have a solid working knowledge of the criteria for major disorders. Many of the diagnostic categories require precise time qualifiers, so any reports of significant symptoms should be followed by the clinician's efforts to establish their time of onset, duration, and severity. The respondent should be encouraged by the clinician to employ descriptive terms, and to indicate in some way the intensity of the symptoms with a numerical scaling or comparative descriptors.

4.04.1.8 Closing the Interview

As the time for the interview draws to a close, the clinician should consolidate the information gained. It is helpful to review one's notes so that any lingering questions may be answered, and to clarify any dates, names, or other details that may have become confused during the course of the interview. An additional responsibility of the clinician conducting the interview is to assist the informant in achieving closure. Many times the clinical interview results in emotional dilation and some level of cognitive disorganization as distressing events are recalled and exposed to another person. The skilled clinician will continue to structure the interview with reminders about the amount of time remaining, summarizing the information provided, and giving appropriate feedback to the informant regarding what to expect next. Avoid rushing the informant out of the room, but be prepared to set limits about the closing of the interview.

When possible, it is beneficial to give the informant a good idea of your diagnostic formulation, and to outline possible intervention strategies. If this is not possible or appropriate at the close of interview, convey to the informant what steps will be taken to complete the evaluation, or provide the informant with an idea of how the information provided will be utilized. If possible the informant should leave the interview with the feeling that they have been heard, understood, and will be benefiting in some way from having participated.

4.04.1.9 The Collateral Interview

Collateral interviewing refers to any direct interviewing done with persons other than the identified patient. Common collateral individuals who are interviewed in the clinical setting include parents, spouses, siblings, and other close relatives. In the case of children and adolescents, school teachers, administrators, and counselors are also often interviewed directly about the behavior and adaptive functioning of the patient. The same skills used in interviewing the identified patient will be employed in these interviews. Empathy, a lack of criticism, and an appropriate use of humor are just as indispensable in talking with a spouse or school principal as they are with the individual presenting for assessment and/or treatment. In many cases, the collateral interview is conducted because the patient is unable to provide the needed information on their own because of disorganizing pathology or other limiting factors, making a collateral information source even more important. In conducting the collateral interview, one must also determine, to the extent possible, the degree of reliability or weight to place upon the information thus gathered.

The clinician should consider the amount of time the informant has known the patient, and the circumstances under which the patient has been observed by the informant. In the case of the school teacher, beginning with questions regarding the amount of time spent with the patient and the subjects taught provides an opportunity to gather useful information about the school setting. In addition, this allows the clinician to evaluate to some extent the affective responses of the informant toward the patient, for example, excessive anger or frustration on the part of the teacher may point to possible distortions in reporting. It is helpful to probe gently for the teacher's experience level, to avoid being unduly influenced by the observations of one who has relatively little comparative knowledge of normative classroom behavior. Begin by asking how long they have been in this particular school, or teaching this particular grade. If the teacher is special education certified, ask how long they have been certified and in what areas. Usually these queries will be sufficient to obtain a good estimate of the experience base of the teacher, and most will actually respond to the general probes with more than enough information to make appropriate judgments.

In the case of the parent interview, take care to establish current custody arrangements and responsibilities as clearly as possible. Depending upon the jurisdiction in which the clinician works, noncustodial parents may not have the right to seek mental health services for the child. It is the clinician's responsibility to be aware of all the legal constraints on service, as well as the ethical duties peculiar to working with children or others who are unable to consent to treatment legally. Be cautious about taking the word of one parent involved in a visitation or custody dispute who reports that the other parent has no interest in the child, or that the other parent would be completely uninterested in assisting in assessment or treatment for the child. Experience indicates that while this may be true in some cases, this attempt to shut the other parent out of clinical work may result in significant distortion of the presenting facts, and can hamper effective work with the child. Thus, if the parent bringing the child for services indicates that their (ex)spouse will not participate in the interview, go ahead and obtain consent from the present parent to contact the reportedly uninvolved parent. This action would only be contraindicated if the clinician has convincing evidence that contacting the other parent would present a significant danger to the patient.

4.04.2 DEVELOPMENTAL CONSIDERATIONS IN INTERVIEWING

4.04.2.1 Interviewing Children (Preschool Age through Older Elementary)

Because children are usually brought into the clinic setting by their parents, clinicians typically schedule an interview with the parents to obtain information about current concerns and past history. Parents are in a unique position to provide a chronology of significant events in the child's life, leading up to the present concerns and reasons for referral. Often collateral interviews will be scheduled with others who play a significant role in the child's life, such as grandparents, teachers, day care providers, etc. Indeed, some diagnoses (such as attention deficit hyperactivity disorder [ADHD]) require that symptoms be documented across at least two settings, and it is helpful to have informants from settings such as school to add to the history provided by parents. One should not limit interviewing to only the adults who are significant in the child's life, however. To do so would create a risk of overlooking important information that could be obtained from the child directly about the child's perceived fears, anxieties, mood, and critical events in the child's world. The child's perspective is often overlooked in situations where the child is not articulate about feelings or is immature in language development. It is necessary for the clinician to develop skill in obtaining interview information from children even in these circumstances. An excellent resource for interviewing or observing children, including infants, can be found in Sattler (1998, pp. 96–132).

4.04.2.2 Interviewing Parents

The purpose of the interview with parents is similar to that discussed earlier in the chapter, in that the clinician attempts to clarify the reasons for concern, identify strengths and weaknesses that moderate the presenting problems in the child, and obtain information that could assist with treatment planning. However, there are important ecological variables that are salient for children and should be addressed in the interview. These include placing the child's current and past problems into a social and developmental context, assessing possible risk and resilience factors that may relate to the child's problems, and assessing the consequences or developmental impact of the child's problems on their future development.

4.04.2.3 Social Context

Schroeder and Gordon (1993) outlined several steps in assessing the problems of young children, including clarifying the referral questions and determining the social context of the problem. Parents often present to clinicians feeling anxiety and/or frustration about their child's problems. This may lead to emotionally-laden, imprecise descriptions of behavior (e.g., "He never minds!" or "She is always disrespectful to her parents!"). The first task in the interview is to help parents define the specific behaviors that cause concern, and to obtain information about the frequency, intensity, and nature of the problem. For instance, a three-year-old child who displays temper tantrums once per week may be of mild concern, but, one who has tantrums three to five times per day would be of much greater concern. The intensity of the child's problems might be gauged by the degree of distress caused to the child or the disruption to typical family activities. For instance, tantrums that occur occasionally at home may cause less distress than if they occur with regularity at church, school, or in other public places. Finally, the nature of the child's problems will be an indicator of severity. Children who engage in cruelty to animals or other people, who are destructive, or who engage in a pattern of fire-setting behavior with the intent to destroy property are of more concern than those who have less serious oppositional and defiant symptoms. As clinicians interview parents about the specific behaviors of concern, important information about the frequency, nature, and severity of the problems can be assessed.

The social context is best assessed by asking simple questions such as, "Who is concerned about the child?," "Why is this person concerned?," and "Why is this person concerned now vs. some other time?" (Schroeder & Gordon, 1993). Although parents or teachers may refer children for assessment or treatment, this does not mean that the child necessarily has a problem that needs treatment. A teacher who refers several active children from a first grade class may be feeling overwhelmed by the sheer number of active children in the class at one time, although a given child's behavior may not be severe enough to warrant a diagnosis of ADHD. Rutter and Schroeder (1981) provided a case of example of a mother who presented with concerns about her daughter occasionally masturbating while watching television. In an attempt to determine why this mother was concerned and the best approach to intervention, the clinician asked about the mother's perception of what this behavior means. The mother responded by saying that she knew her daughter was at a developmental age when exploring her body was normal, and she knew that nothing bad would happen (e.g., such as growing hair on the palms of her hands) as a result of masturbation. The additional question ('Why is the mother concerned now vs. any other time?") yielded the most salient information about the mother's concerns. The mother revealed that her mother-in-law was coming for a visit the next week, and she was concerned that this relative would have a negative reaction to seeing her granddaughter masturbate. The intervention was simplified by understanding the true reason for the mother's concern. The clinician recommended that the mother provide rules about when and where it was acceptable to masturbate (e.g., when her daughter was alone, in her bedroom, or in the bathroom) and institute a behavioral reward system for remembering not to masturbate while watching television.

Other social contextual information can be obtained about family status (who is living in the home), recent transitions (moves, job changes, births, recent deaths, or illnesses of significant family members), and other family stresses (marital problems, financial stresses, etc.). The presence of persons who are supportive to the child, or who may provide a buffer in the face of other stresses, is important. The literature on resilience is replete with examples of children who have lived with adversity but have developed and functioned normally due to protective factors in their social history (Routh, 1985). The interview with parents also provides an opportunity for assessing possible psychopathology in the parents, such as significant depressive or anxiety symptoms; problems with anger management and self-control, as is often seen in abusive parents; substance abuse problems that may lead to parental instability or neglect; or problems with reality testing, as in schizophrenia. One mother, for example, described her 14-year-old son as being afraid of the dark and reporting seeing ghosts at night. This was viewed by the clinician as an example of a fear that was developmentally inappropriate for a 14-year-old; it also raised questions about possible hallucinations. The context became more clear when the mother revealed that she saw nothing inappropriate about this behavior. The mother reported that she, too, needed to sleep with a light on due to her fear of the dark, and that she also imagined seeing ghosts in her bedroom. This mother reported that she and her son had many discussions about their mutual fears. The context of the son's fears was changed by the mother's revelation, and the clinician decided to include a more thorough interview regarding the mother's mental status in this case. Even when no concerns about parental psychopathology exist, parental stress levels and affect must be considered when interpreting their reports about child behavior. A parent who is calm and rational in providing a history of their child's behavior may be viewed as more objective than a parent who is extremely upset, tearful, or angry and uses exaggerated descriptors of the child's behavior.

4.04.2.4 Developmental Context

Developmental context provides an essential lens from which to view children's behavior, allowing the clinician to evaluate the child's behavior relative to that of other children of the same chronological and/or mental age. For instance, enuresis may not be unusual in a four-year-old, but would be of concern in a 14-year-old. Likewise, enuresis may not be unusual in a six-year-old youngster with a moderate degree of mental retardation. Some behavioral problems of young children are transient, reflecting their responses to normative developmental challenges (e.g., a five-year-old girl who displays a regression to thumb-sucking and infantile speech patterns following birth of a new sibling). Other problems are more serious and persistent, and suggest risk for later maladjustment. Familiarity with developmental theory and the rich empirical literature in clinical child psychology and developmental psychopathology can provide the clinician with guidance in making these discriminations.

Knowledge of the sequence and transitions in social/emotional development are helpful to the clinician in judging the appropriateness of children's behavior at various ages. For instance, a toddler who has never displayed a strong attachment to a primary caregiver (usually a parent or parents) and who seems to form attachments indiscriminately with others would raise concerns about possible attachment relational problems. A seven-year-old child who cannot delay gratification or consider the feelings of others would be of concern for not having learned appropriate self-control and capacity for emotional empathy that would be expected at that age. Critical developmental tasks for the preschool age child include establishing effective peer relations (e.g., learning to share material resources and adult attention with peers, establishing reciprocal play relationships) and developing flexible self-regulatory skills (e.g., adjusting to the authority of preschool or daycare teachers and classroom routines). In contrast, children in middle to late elementary years (seven to 12 years of age) encounter developmental tasks related to mastery of knowledge and intellectual skills, leading to feelings of productivity and competence. Children with learning disorders or other developmental problems that interfere with academic progress may be at risk for secondary behavioral or emotional problems related to their primary problems with learning during this developmental period. The clinician must tailor the interview to exploration of the child's strengths and weaknesses in the context of appropriate developmental expectations for particular ages.

The newly emerging field of developmental psychopathology has provided a theoretical and empirical base for better understanding the developmental precursors of psychopathology in children, and the impact of this psychopathology on subsequent functioning (cf, Cicchetti & Cohen, 1995a, 1995b). There is a growing body of research addressing risk factors for the onset and continuity of various childhood disorders. For example, Loeber and colleagues have made important contributions to understanding the developmental pathways to childhood disruptive behavior disorders, in which different constellations of risk factors lead to different outcomes. In their longitudinal study of inner city boys at ages seven, 10 and 13, they found that initiation into antisocial behavior was predicted by some factors (e.g., poor parent–child relations, symptoms of physical aggression) that were present across all three ages, while others (e.g., shyness at age seven, depression at age 10) were age specific (Loeber, Stouthamer-Loeber, Van Kammen, & Farrington, 1991). Further, the environments of children who remained antisocial differed from those whose antisocial behavior dropped out; good supervision was more important in helping older children (age 13 at intake) while attitude toward school was more important for the younger children. Studies such as these illustrate the importance of understanding the contextual variables related to parenting style and parent–child relational issues, as well as specific child behaviors, in determining the significance of presenting problems and their possible trajectory over time.

4.04.2.5 Direct Interview of Children

Perhaps the best and most comprehensive resource guide for interviewing children and adolescents who present with a variety of problems is Sattler's (1998) book on clinical and forensic interviewing of children. Basically, the goals of the initial interview of the child depends upon the referral questions as well as the age and verbal ability of the child (Sattler, 1998). When interviewing children and their families the information sought often includes the following:

(i) to obtain informed consent to conduct the interview (for older children) or agreement to be at the interview (for younger children);

(ii) to evaluate the children's understanding of why they are at the interview and how they feel about being at the interview;

(iii) to gather information about the children's perception of the situation;

(iv) to identify antecedent and consequent events related to the children's problems;

(v) to estimate the frequency, magnitude, duration, intensity, and pervasiveness of the children's problems;

(vi) to identify the circumstances in which the problems are most or least likely to occur;

(vii) to identify potentially reinforcing events related to the problems;

(viii) to identify factors associated with the parents, school, and environment that may contribute to the problems;

(ix) to gather information about the children's perceptions of their parents, teachers, peers, and other significant individuals in their lives;

(x) to assess the children's strengths, motivations, and resources for change;

(xi) to evaluate the children's ability and willingness to participate in formal testing;

(xii) to estimate what the children's level of functioning was before an injury; and

(xiii) to discuss the assessment procedures and possible follow-up procedures. (Sattler, p. 98).

A part of the interview process with children includes observation of parent and child and obtaining collateral information from the schools or others if the presenting problem relates to learning or behavior problems outside the home. Recognizing the developmental tasks that children must master at varying ages helps the clinician understand the child's behavior. Thus, a comprehensive, detailed developmental history of the child and family milieu is an integral part in establishing an appropriate treatment.

Clinicians must also consider interviewing the child at some stage during the evaluation process. Very young children may be observed using a free-play setting and using observational guides during the play. The clinician can learn a great deal about the child's energy level, physical appearance, spontaneity, organization, behavior, affect, and attitude through their play and through a diagnostic play interview. School-aged children are able to share thoughts and feelings with the clinician unless they are unusually shy or oppositional (Sattler, 1998). Obviously establishing rapport and maintaining the child's cooperation during the interview is crucial. Kanfer, Eyberg, and Krahn (1992) identified five basic communication techniques that can aid the clinician in attaining rapport and cooperation. First, the clinician can use descriptive statements to describe the clients ongoing behavior, for example, "You're stacking the toys so nice." Second, using reflective statements to mirror the childs statements can be nonthreatening. For example, if the child says she wants to play with blocks the clinician merely reflects "you want to play with the blocks." Third, labeled praise helps the child feel good and feel that the clinician approves of them. Fourth, the clinician must avoid critical statements that suggest disapproval or make the child feel as though they are bad. Finally, open-ended questions avoid yes or no answers and provide opportunities for children to elaborate on their responses (Kanfer et al., 1995).

4.04.2.6 Adolescents

Interpersonal style may play a greater role in good interviewing with this age group than with any other. Adolescents tend to be intensely attuned to any communications that concern their personal appearance, skills, or competence, and the interviewer must avoid at all costs even the hint of condescension. As numerous authors have pointed out, older clinicians tend to identify readily with the parents of adolescents, while younger ones may easily align themselves with the youth. The clinician who remains unaware of their tendencies in this regard runs the risk of making insensitive or intrusive statements that will inhibit rapport rather than increase it. In the first case, the clinician who approaches the adolescent with a parental attitude may unconsciously interact in a way that increases the informant's anxiety, guilt, and hostility. Questions that presuppose information the adolescent has not provided may mirror intrusive interactions with other adults, resulting in defensive efforts and guardedness. Similarly, clinicians who identify easily with the adolescent may also appear "hokey" and insincere when they misuse popular language, or try too hard to relate their own somewhat misty adolescent experiences to those of the youth they are interviewing. These errors result from incautious use of the same techniques that will be necessary for successful adolescent interviewing. That is, to obtain good information and develop adequate rapport, the adolescent must perceive that the clinician is clearly on their side within the boundaries of the relationship. Judicious use of self-disclosure can help the adolescent believe that the interviewer is not attempting to take away from the interaction without reciprocating. Earnest discussion of the limits of confidentiality and the purposes of the interview will help allay some of the suspicions the informant may have about the clinician's role, and will serve to make a distinction between the clinician–informant relationship and those the adolescent has with parents, teachers, parole officers, and other adults.

The adolescent patient presents a number of challenges to the interviewer that are often less present or significant in interactions with both older and younger people. Because of the unique developmental pressures and challenges of adolescence, special care must be taken in the interview to ensure adequate cooperation as well as to make the interview process a helpful one to the patient. It is essential that the interviewing clinician possess a basic knowledge of the common demands and urges present in the adolescent and their family to effectively assess the patient's functioning. Listed next are those tasks commonly believed to be operating in the adolescent period of life according to various developmental theorists (Erikson, 1963; Rae, 1983).

4.04.2.6.1 Separation–individuation

Separation–individuation refers to the need of the adolescent to identify those qualities in themselves that set them apart from their family. Many of the issues bringing adolescents to treatment involve conflicts that are direct results of this process. The adolescent during this time begins testing family boundaries and experimenting with beliefs and behaviors that differ from those held by their caretakers. This process often produces considerable anxiety for all family members, and the adolescent's interpersonal relations may become quite variable. Often, the adolescent moves between the poles of autonomy from, and dependence upon, the family. An important portion of the adolescent interview is that of identifying the severity of the stressors resulting from this natural process.

4.04.2.6.2 Resolving conflict with authority figures

Related to the individuation task is the frequent occurrence of conflict with authority figures outside of the family as well as within. For younger adolescents this involves primarily their teachers and other school personnel, and for later adolescents this includes work supervisors as well. Conflicts with authority figures outside the home often have their roots in greater-than-average difficulties in resolving the family relationship struggles. Thus, when interviewing the adolescent, it is helpful to identify both positive and negative relationships with other adults in their life.

Often classroom performance for the adolescent presenting for services is related strongly to the quality of the relationship with the teacher, so discussion of academic performance (usually a relatively nonthreatening issue in the context of the clinical interview) can elicit useful information about this area of functioning as well. Adolescents, as well as younger children, may readily express relational difficulties in response to the question "Is he/she a good teacher?" This often elicits the adolescent's opinion regarding the desirable qualities in an important adult, and allows the interviewer to follow up with questions regarding the adolescent's ability to recognize their own role in any positive or negative interactions.

4.04.2.6.3 Peer group identification

As adolescence is inarguably a time of shifting focus from family relations to peer relations, it is vital to gather information regarding the patient's friendships and any identification with a social subgroup. Some effective ways of eliciting this information include discussion of music topics, such as taste and dress, that will provide clues to the adolescent's social presentation and degree of inclusion or exclusion from social groups.

To effectively interview adolescents regarding social issues, it is necessary for the clinician to maintain a moderate degree of understanding of popular culture. Thus, one would be well served by making an effort to watch television programming, read magazines, and spend time taking in the various electronic media that are aimed at people in this age group. The interviewer should not attempt to present as an authority on the adolescent's culture, but will benefit from being able to recognize specific music groups, current movies, video games and Internet activities, and other elements that are part of the adolescent's milieu. It is often helpful to enlist adolescents' aid in delineating the social groups present in their school, then ask them to identify the group to which they feel they most belong. This question can usually be asked rather directly, and many teens are pleased by the opportunity to display their understanding of the social complexities in their school.

Follow-up inquiry should establish with whom the adolescent spends most time and how they see themself as fitting into the groups at school. Many youth social strata include a group delineated primarily by drug/alcohol use as well as different groups for aggressive or delinquent behavior that may be gang-affiliated or gang-emulating. Thus, the social categories to which the adolescent assigns themself may also point the interviewer toward necessary inquiries into these possible problem areas as well as providing information about the degree of social integration in the adolescent's life.

4.04.2.6.4 Realistic appraisal and evaluation of self-qualities

As the focus of evaluation or treatment is likely to include assessing and modifying self-image, it is necessary to include questions regarding the ways in which the adolescent views themself. Adolescents generally display both overly optimistic and excessively pessimistic appraisals of personal qualities. One purpose of the interview is to assist in determining when these perceptions area faculty and result in impaired functioning. It is often helpful to present questions about self-image in terms of strengths and liabilities, and to follow up on both. Questions about the adolescent's physical capacities as well as social and emotional abilities are necessary components of the interview. This portion of the interview can be directed toward uncovering problems with perception of body image and behaviors related to physical health. The interviewer should attend carefully to clues that might indicate the need for more focused exploration of possible eating disorders, and to somatic complaints indicative of anxiety or depression.

4.04.2.7 Interviewing Young Adults (18–40 Years)

The psychological distinction between adolescence and young adulthood is frequently blurred, and many of the same traits and problems may be observed in individuals both over and under the chronological age of majority. However, since the age of majority is generally 18 years, a higher proportion of patients over age 18 will be self-referred and hence will present in a more open and cooperative manner than some adolescents. Additionally, young adults are more likely to present with some subjective description of their distress and their situation. Therefore, the client may be more likely to identify a problem area spontaneously. Despite the fact that more patients in this age group may independently seek services, many of the adolescent issues related to establishing an autonomous role-identity may surface in the interactions with the interviewer, especially with the "youngest" adults. Therefore the interviewer may frequently call upon the skills used in interviewing adolescents.

Erikson (1963) identified the primary developmental conflict for the various stages of adulthood, and these stages suggest important interview topics (see Table 1). The primary conflict of young adulthood is intimacy vs. isolation. Consequently, many of the psychological problem areas frequently encountered will revolve around commitment to interpersonal relationships and establishing trust. Establishment of a working relationship with the patient is also affected by these issues.

A relatively greater amount of the interview might be devoted to exploration of existing relationships or those the patient wishes existed. One type of relationship to consider is that with parents and family of origin. Establishing the degree of desired independence continues to be an issue with some young adults. Issues relevant to these ties might be financial (e.g., parents may be paying college expenses), or they may be more interpersonal in nature (e.g., parents controlling social relationships or defining goals for the patient).

Intimate relationships with individuals of the same or opposite sex may also be a source of psychological discomfort and play a part in the development of anxiety disorders or depression. Inquiry about social functioning should include peer relationships, such as partners in love relationships, friends, and acquaintances.

Individuals in the young adult age group generally will have established some degree of independence, and the relative importance of work and employment will be much greater than at younger ages. The interview should therefore include specific inquiry into current job status, job satisfaction, goals, and relationships with co-workers. The further one progresses into this stage, the greater is the importance of establishment of a stable intimate relationship and mutual trust, and the higher the probability that the issue of procreation will arise. Therefore inquiry should include questions about intentions and concerns associated with having children and child rearing and any differences with one's partner about children.

Finally, the initial episodes of many severe psychiatric disorders are most likely to occur within the young adult period. Initial episodes of depression, and post-partum depression, are likely to occur in those affected before they pass through this period (Kaelber, Moul, & Farmer, 1995). Therefore screening for affective disorders should be included in the interview. A

Table 1 Interview topics for each developmental stage.

Young adult	Independence from family, relationships with peers, stable intimate relationships, trust in relationships, establishment of a family, issues related to having and rearing children, education, and career goals.
Middle adult	Achievement of work and family goals, career or family role changes, responsibility for aging parents, death of grandparents and parents, reducing responsibility for children, changes in physical appearance and characteristics, and anticipating retirement.
Older adult	Accepting status of family and career, developing identity as grandparent or "elder advisor," coping with reduced physical capability and/or health changes, specific plans for retirement, loss of siblings, spouse, and friends. Increased reliance on children or caretakers.
Late adult	Coping with deteriorating health, decreased mobility, dependence on caretakers, and anticipation of death.

later section of this chapter deals with interviewing depressed and anxious patients. Additionally, first episodes of schizophrenia or bipolar disorder generally take place in adolescence or young adulthood and the interviewer should be sensitive to symptoms of these disorders.

4.04.2.8 Interviewing Adults in Middle Adulthood (40–60 Years)

Interview techniques need not differ with this age group, but the relevant topics from a developmental perspective are somewhat different (see Table 1). This period encompasses much of the creative and productive portion of the life span in western culture. The emphasis is not on starting, but on completing tasks begun in young adulthood. The focus of individuals at this stage of life is much less on goal setting than on goal attainment. The growth and nurturing of an established family, the attainment of successive career goals, and nurturing of one's parents and grandparents occur in this time span. One's children come into adulthood and begin to establish their identities and families. Inquiry into the relationships with the former and succeeding generations should be made. Towards the middle of this period, individuals are able to anticipate the likelihood of reaching family and career goals, and become aware of the fact that certain goals for themselves and their children may not be met. Biological changes associated with mid-life, which are well-defined for women, but also may be present for men, should be queried since they may be associated with depression or anxiety. Possible mid-life existential crises related to loss should also be assessed. The losses may result from death of parents or grandparents, or changes in roles as parent, spouse, or worker.

4.04.2.9 Interviewing Older Adults (60–70 Years)

For many adults in this age range, the predominant life circumstance deals with additional impending changes in the area of life roles. Retirement usually occurs within this time frame, and inquiries might reveal difficulties in psychological adjustment to one's own retirement or the retirement of a significant other. The frequency of death in the patient's social circle gradually increases, and may include a spouse, close friends, or even an adult child.

Due to the possibility of some early decline in cognitive capacity in this age group, the response to inquiry may be defensiveness and denial. The patient with some early impairment may deny the need for the evaluation, object to questions, and become resentful if the interview serves to demonstrate difficulties with memory. Therefore, it becomes more important to interview a collateral person or include a collateral person in the patient interview. In addition to a spouse or family member, a collateral person to be considered with older adults is an adult caretaker, who may or may not be related to the patient. This may give rise to some special issues of confidentiality.

Attention to the collateral person's nonverbal behavior may sometimes suggest that they are uncomfortable reporting the patient's difficulties, especially in the patient's presence. In such circumstances a separate collateral interview is desirable.

4.04.2.10 Interviewing In Late Adulthood (70 Years–End of Life)

Adults in the latest stages of life have their own unique set of circumstances of which the interviewer must be aware. The losses that may have begun earlier may become more frequent.

Physical changes, often represented by medical problems, may interfere with some life activities, and there may be a need to accept reduced independence. At some point anticipation of the end of life is common. The combination of these forces often lead the elderly to have a perspective on life and the situation giving rise to the interview that differs considerably from younger adults, in that they may be unconcerned and see no need for the evaluation. Often the reasons for the interview are more important to someone else than to the patient. As with children and adolescents, it is more likely that someone other than the client identified the need for and arranged for the mental health contact.

It is also common for the oldest adults to answer questions more slowly, either because of difficulty accessing information or because a more tangential and elaborate route is taken to reach a point in conversation. Patience on the part of the examiner in these situations is important, both for maintaining rapport and to show the proper respect due the patient.

It has been estimated that the incidence of cognitive decline in people over age 65 is 10–20% (Brody, 1982). Estimates are as high as 25% of those 80 years and older (Hooper, 1992). Thus, the likelihood of cognitive impairment is even greater in this age group than those discussed previously. For those with cognitive dysfunction, cooperation may be minimal and denial, and even belligerence, may be present. Again, the availability of a collateral person for interview may be very important, as the patient may not cooperate or may be impaired in their ability to provide information.

4.04.3 INTERVIEWING SPECIAL POPULATIONS OR SPECIFIC DISORDERS

4.04.3.1 Interviewing Depressed Patients

Interviewing depressed adults may require some adjustment in the tempo and the goals of the interview. Due to low energy and psychomotor retardation, it may not be possible to gather all the desired information within the time available. Hence, some prioritization of information is necessary, so that issues such as suicidality, need for hospitalization, and need for referral for medication may be addressed. Beck (1967) and later, Katz, Shaw, Vallis, and Kaiser (1995) pointed out that the interpersonal interaction with the depressed patient may be frustrating for the interviewer, not only due to the slowness mentioned above, but also because of the negative affect and negative tone of information provided.

It is also particularly important with depressed patients, who are prone to hopelessness, to provide encouragement and attempt to impart hope to the patient during the interview. This may be done by recognizing areas of strength, either in terms of personal qualities or successful areas of functioning.

Specific inquiry is necessary to diagnose depression appropriately, and a variety of sources are available to guide this inquiry. Diagnostic criteria for depression are clearly delineated in the *DSM-IV* (American Psychiatric Association [APA], 1994). A number of structured interviews have been developed that may serve as guides for inquiry or provide sample questions. Formal training is required for the reliable use of these interviews for diagnostic purposes. The Schedule for Affective Disorders and Schizophrenia (SADS; Endicott & Spitzer, 1978) is a relatively early forerunner of current interviews that slightly preceded the *DSM-III* (APA, 1980), and includes probe questions for depressive symptoms as well as other disorders.

The Structured Clinical Interview for *DSM-III-R* (SCID; Spitzer, Williams, Gibbon, & First, 1992) is a more current instrument with a modular format so that sections for each disorder may be used independently. Table 2 also lists sample questions that might be used to probe for the presence of various depressive symptoms.

4.04.3.2 Interviewing Anxious Patients

The anxious patient may also present some special difficulties during the interview. If the patient is acutely distressed at the time of the interview, as might be true of someone with a generalized anxiety disorder, they may provide a rush of disorganized information so that it may be difficult to obtain a coherent history. Anxiety interferes with attention and concentration, so that repetition may be necessary. Experience has shown that in such a situation, some initial intervention using brief relaxation techniques, is helpful before proceeding with the interview. Anxious patients also frequently seek reassurance that treatment will be effective in reducing their anxiety. It is appropriate to indicate that treatment techniques have been helpful to other anxious patients, and that these techniques will be available to them.

The diagnostic symptoms of various anxiety disorders are identified in *DSM-IV*, and the structured interviews mentioned earlier also provide some guidance for the inquiry for specific anxiety symptoms. In addition to the diagnostic information, it is important to

Table 2 Sample questions for depressive symptoms.

Mood (depressed)	How would you describe your mood? Have you been feeling down or sad much of the time? How much of the time do you feel down or sad?
Mood (irritable)	Have you been more short-tempered than usual for you? Do others say you are more irritable or lose your temper more easily than usual?
Interest and pleasure	Are you as interested as ever in things like your work, hobbies, or sex? Do you continue to enjoy the things you usually like to do, like hobbies, doing things with friends, or your work? Has your interest declined in things which used to be really interesting for you?
Energy/fatigue	Do you have enough energy to do the things you want to do or need to do? Do you have the energy to do the things you find interesting? Do you tire out more easily than usual for you?
Weight loss/gain	Have you gained or lost weight since . . . (specify a time period)? If the patient does not know, you may inquire about whether clothes fit properly, or what others may have said about weight. Insomnia/hypersomnia
Insomnia/hypersomnia	How well are you sleeping? Do you have difficulty getting to sleep? (initial insomnia). Do you awaken frequently during the night and have trouble getting back to sleep? (middle insomnia) Do you awaken too early in the morning? (terminal insomnia)
Psychomotor agitation/retardation	Have other people commented on your being too active or being very slowed down? Are there times when you just can't sit still, when you have to be active, like pacing the floor or something similar? Are there times when you are very slowed down, and can't move as quickly as usual?
Worthlessness/guilt	How do you feel about yourself? Do you think of yourself as worthwhile? Do you often feel guilty or have thoughts of being guilty for something? Is guilt a problem for you?
Concentration/decisiveness	Is it difficult for you to keep your attention on things you are doing? Do you lose track of things, like conversations or things you are working on? Is there a problem with making decisions? Does it seem that your thoughts are slowed down, so it takes a long time to make a decision?
Thoughts of death/suicide	Do you frequently have thoughts of death? Do you think a lot about friends or loved ones who have died? (Inquire if someone close to the patient has recently died or is near death.) Do you sometimes think it would be better if you were dead? Have you thought abut hurting yourself or killing yourself? Have you planned a particular way in which you would kill yourself? What would keep you from killing yourself?

inquire about ways the patient has attempted to cope with the anxiety, and to provide some reinforcement for such efforts.

4.04.4 SUMMARY

The clinical interview provides rich diagnostic information that can aid in the assessment and treatment of patients. Interpersonal style of the clinician interview, structuring the interview, the setting in which the interview takes place, preparing the patient, and the beginning, middle, and ending phases of the interview are discussed. Developmental considerations and suggestions are offered in interviewing children, adolescents, and adults. Sample questions are primarily for interviewing depressed patients.

4.04.5 REFERENCES

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Beck, A. T. (1967). *Depression: Clinical, experimental and therapeutic aspects*. New York: Harper and Row.

Bricklin, B. (1990). *The custody evaluation handbook: Research-based solutions and applications*. New York: Brunner-Mazel.

Brody, J. A. (1982). An epidemiologist views senile dementia: Facts and fragments. *American Journal of Epidemiology, 115,* 155–160.

Cicchetti, D., & Cohen, D. J. (Eds.) (1995a). *Developmental psychopathology. Vol. 1: Theory and methods*. New York: Wiley.

Cicchetti, D., & Cohen, D. J. (Eds.) (1995b). *Developmental psychopathology. Vol. 2: Risk, disorder, and adaption.* New York: Wiley.

Corcoran, K., & Vandiver, V. (1996). *Maneuvering the maze of managed care: Skills for mental health professionals*. New York: Simon & Schuster.

Egan, G. (1994). *The skilled helper: A problem management approach to helping*. Pacific Grove, CA: Brooks/Cole Publishing.

Endicott, J., & Spitzer, R. (1978). A diagnostic interview: The Schedule for Affective Disorders and Schizophrenia. *Archives of General Psychiatry, 35,* 837–844.

Erikson, E. H. (1963). *Childhood and society* (2nd ed.). New York: Norton.

Hooper, C. (1992). Encircling a mechanism in Alzheimer's disease. *Journal of National Institutes of Health Research, 4,* 48–54.

Kaelber, C. T., Moul, D. E., & Farmer, M. E. (1995). Epidemiology of depression. In E. E. Beckham & W. R. Leber (Eds.), *Handbook of depression* (2nd ed., pp. 3–35). New York: Guilford Press.

Kanfer, R., Eyberg, S., & Krahn, G. L. (1992). Interviewing strategies in child assessment. In M. Roberts & C. E. Walker (Eds.), *Handbook of clinical child psychology* (pp. 49–62). New York: Wiley.

Katz, R., Shaw, B., Vallis, M., & Kaiser, A. (1995). The assessment of the severity and symptom patterns in depression. In E. E. Beckham & W. R. Leber (Eds.), *Handbook of depression* (2nd ed., pp. 61–85). New York: Guilford Press.

Loeber, R., Stouthamer-Loeber, M., Van Kammen, W., & Farrington, D. P. (1991). Initiation, escalation and desistance in juvenile offending and their correlates. *The Journal of Criminal Law and Criminology, 82,* 36–82.

Luborsky, L. (1996). *The symptom–context method.* Washington, DC: APA Publications.

Morrison, J. (1995). *The first interview*. New York: Guilford Press.

Rae, W. A. (1992). Teen–parent problems. In M. C. Roberts & C. E. Walker (Eds.), *Handbook of clinical child psychology* (pp. 555–564). New York: Wiley.

Routh, M. (1985). Masturbation and other sexual behaviors. In S. Gabel (Ed.), *Behavioral problems in childhood* (pp. 387–392). New York: Grune & Stratton.

Rutter, D. K., & Schroeder, C. S. (1981). Resilience in the face of adversity: Protective factors and resistance to psychiatric disorder. *British Journal of Psychiatry, 147,* 598–611.

Sattler, J. (1998). *Clinical and forensic interviewing of children and families* (pp. 96–132) San Diego, CA: J. M. Sattler.

Schoeder, C. S., & Gordon, B. N. (1993). Assessment of behavior problems in young children. In J. L. Culbertson & D. J. Willis (Eds.), *Testing young children* (pp. 101–127). Austin, TX: ProEd.

Spitzer, R., Williams, J. B. W., Gibbon, M., & First, M. (1992). The Structured Clinical Interview for *DSM-III-R* (SCID): I. History, rationale and description. *Archives of General Psychiatry, 49,* 624–636.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.05
Structured Diagnostic Interview Schedules

JACK J. BLANCHARD and SETH B. BROWN
University of New Mexico, Albuquerque, NM, USA

4.05.1 INTRODUCTION 98
4.05.1.1 Evaluating Structured Interviews 99
4.05.1.1.1 Reliability 99
4.05.1.1.2 Validity 101
4.05.1.2 Overview 101
4.05.2 ADULT DISORDERS 101
4.05.2.1 Schedule for Affective Disorders and Schizophrenia 101
4.05.2.2 Reliability 102
4.05.2.2.1 Summary 104
4.05.2.3 Present State Examination 104
4.05.2.3.1 Reliability 105
4.05.2.3.2 Supplements to the PSE 106
4.05.2.3.3 Summary 107
4.05.2.4 Structured Clinical Interview for DSM-IV*/Axis I Disorders* 107
4.05.2.4.1 Reliability 108
4.05.2.4.2 Summary 109
4.05.2.5 Comprehensive Assessment of Symptoms and History 109
4.05.2.5.1 Reliability 110
4.05.2.5.2 Summary 110
4.05.2.6 Diagnostic Interview for Genetic Studies 111
4.05.2.6.1 Reliability 111
4.05.2.6.2 Summary 112
4.05.2.7 Diagnostic Interview Schedule 112
4.05.2.7.1 Reliability 112
4.05.2.7.2 Summary 113
4.05.2.8 Composite International Diagnostic Interview 113
4.05.2.8.1 Reliability 114
4.05.2.8.2 Summary 114
4.05.3 PERSONALITY DISORDERS 114
4.05.3.1 Structured Interview for DSM-IV *Personality Disorders* 114
4.05.3.1.1 Reliability 115
4.05.3.1.2 Summary 115
4.05.3.2 International Personality Disorder Examination 116
4.05.3.2.1 Reliability 117
4.05.3.2.2 Summary 117
4.05.3.3 Structured Clinical Interview for DSM-IV *Personality Disorders* 118
4.05.3.3.1 Reliability 119
4.05.3.3.2 Summary 119
4.05.3.4 Personality Disorder Interview-IV 120
4.05.3.4.1 Reliability 120
4.05.3.4.2 Summary 120
4.05.4 CHILD AND ADOLESCENT DISORDERS 121

4.05.4.1 Schedule for Affective Disorders and Schizophrenia for School Age Children 121
4.05.4.1.1 Reliability 121
4.05.4.1.2 Summary 121
4.05.4.2 Child Assessment Schedule 122
4.05.4.2.1 Reliability 122
4.05.4.2.2 Summary 123
4.05.4.3 Child and Adolescent Psychiatric Assessment 123
4.05.4.3.1 Reliability 124
4.05.4.3.2 Summary 124
4.05.4.4 Diagnostic Interview Schedule for Children 124
4.05.4.4.1 Reliability 124
4.05.4.4.2 Summary 125
4.05.4.5 Diagnostic Interview for Children and Adolescents 125
4.05.4.5.1 Reliability 126
4.05.4.5.2 Summary 126

4.05.5 SUMMARY 126

4.05.6 REFERENCES 126

4.05.1 INTRODUCTION

As early as the 1930s, and certainly by the early 1960s, it was apparent that clinical diagnostic interviews were fallible assessments. Evidence suggested that clinicians frequently arrived at different diagnoses, often agreeing at no more than chance levels (e.g., Beck, Ward, Mendelson, Meck, & Erbaugh, 1962; Matarazzo, 1983; Spitzer & Fleiss, 1974). These findings gave raise to the study of the causes of diagnostic unreliability as well as the development of methods to improve psychiatric diagnosis.

In the first study systematically to examine reasons for diagnostic disagreement, Ward Beck, Mendelson, and Erbaugh (1962) summarized three major sources of disagreement. These were inconsistency on the part of the patient (5% of the disagreements), inconsistency on the part of the diagnostician (32.5%), and inadequacies of the nosology (62.5%). Thus, nearly one-third of the diagnostic disagreements arose from the diagnosticians. Factors associated with differences between diagnosticians included interviewing techniques that led to differences in information obtained, weighing symptoms differently, and differences in how the symptomatology was interpreted (Ward et al., 1962). It is interesting that these problems arose despite methods of study which included the use of experienced psychiatrists, preliminary meetings to review and clarify diagnostic categories, the elaboration of diagnostic descriptions to minimize differences, and the compilation of a list of instructions to guide diagnosis (Beck et al., 1962).

The study of Ward et al. (1962) identified two major sources of disagreement that have been termed "criterion variance" and "information variance" (Endicott & Spitzer, 1978). Criterion variance refers to the errors in diagnostic assignment attributable to how clinicians summarize patient information into existing definitions of psychiatric diagnoses. Inadequacies of early diagnostic systems (e.g., the *Diagnostic and statistical manual of mental disorders* [*DSM-I* and *DSM-II*]) generally arose from the lack of explicit diagnostic criteria. The development of newer diagnostic schemes such as the Research Diagnostic Criteria (RDC; Spitzer, Endicott, & Robins, 1978), and subsequently the *DSM-III* (American Psychiatric Association, 1980), provided inclusion and exclusion criteria and specific criteria relating to symptoms, signs, duration and course, and severity of impairment.

In addressing errors that arise from inadequate nosology, clinicians and researchers are still faced with information variance, that is, errors originating from differences in what information is obtained and how that information is used by the interviewers. As reviewed above, Ward et al. (1962) found that nearly a third of diagnostic disagreements were related to the interviewers. Structured interviews were developed in order to address this source of error variance, and the history of structured interviews goes back to the 1950s (Spitzer, 1983). All structured interviews seek to minimize information variance by ensuring that clinicians systematically cover all relevant areas of psychopathology. Although specific methods vary across instruments, common techniques that characterize structured interviews are the specification of questions to be asked to assess domains of psychopathology and the provision of anchors and definitions in order to determine the ratings of symptoms (e.g., do the descriptions obtained within the interview achieve diagnostic threshold or not?). Despite some shared qualities it is also clear that available structured interviews differ markedly on a number of dimensions. The reliability of diagnoses based on refined diagnostic criteria paired with structured interviews was found to be greatly improved (e.g., Endicott & Spitzer, 1978; Spitzer, Endicott, & Robins, 1978).

4.05.1.1 Evaluating Structured Interviews

The selection of a structured interview will be driven by a number of other concerns. Some of the potential considerations are summarized in Table 1 and are derived from the reviews of Page (1991) and Zimmerman (1994). Questions that should be asked in selecting an instrument will include the diagnoses covered by the interview, the nosological criteria adhered to in generating these diagnoses (e.g., *DSM*, the *International classification of diseases* [*ICD*], or other criteria such at the RDC), and the population studied (e.g., adult or child). Additionally, the context in which the interview is conducted may also be relevant. That is, who will be administering the questionnaire and under what circumstances? Some interviews were developed to be used by lay interviews as in community epidemiological studies while other instruments require extensive clinical experience and are to be administered only by mental health professionals. Other concerns will relate to the guidelines and support available for an instrument. Some measures have extensive book-length user's manuals with available training videotapes and workshops; however, other measures have only sparse unpublished manuals with no training materials available. Finally, major concerns arise regarding reliability.

4.05.1.1.1 Reliability

The reliability of a diagnostic interview refers to the replicability of diagnostic distinctions obtained with the interview. Methods for evaluating agreement between two or more raters can take a variety of forms (Grove, Andreasen, McDonald-Scott, Keller, & Shapiro, 1981). Importantly, these differing methods may have an impact on indices of reliability. The most stringent evaluation involves raters conducting separate interviews on different occasions with diagnoses based only on interview data (i.e., no access to medical records or collateral informants such as medical staff or family members). Reliability assessment based on this method is referred to as "test–retest reliability," given the two occasions of interviewing. This methodology ensures a rigorous evaluation of the ability of the interview to limit information variance and to yield adequate information for the determination of diagnoses. However, in addition to interviewing style and methods of diagnosticians, two other factors contribute to rater disagreements in test–retest designs. First, the information provided by the patient may be different during the two interviews. Even with structured interviews this can continue to be a relevant source of variance. In a review of test–retest data using the Schedule for Affective Disorders and Schizophrenia, Lifetime Anxiety, Fyer et al. (1989) found that more than 60% of diagnostic disagreements were due to variation in the information provided by subjects. Second, there may be a true change in the clinical status of the individual who is interviewed. As the test–retest period increases, the potential contribution of changing clinical status will increase.

Other methods sometimes utilize a single interview that is either observed or videotaped and rated by independent (nonparticipating) raters, yielding inter-rater agreement. This method may yield inflated estimates of reliability as information variance is absent; that is, both raters are basing diagnostic decisions on the same subject responses. Also, interviewer behavior may disclose diagnostic decisions to the second rater. For example, during a structured interview an interviewer may determine that a module of the interview is not required based on subject responses and the interviewer's interpretation of rule-out criteria. The observing rater is aware of this diagnostic decision as the interviewer skips questions and moves to another module. Given the importance of methods used in assessing reliability, throughout this chapter we will attempt to indicate clearly the techniques by which reliability was assessed for each instrument.

In addition to considering study designs used in evaluating diagnostic reliability, it is also important to examine the statistics used to compute reliability. One method that has been used as an index of reliability is percent agreement. As outlined by Shrout, Spitzer, and Fleiss (1987), percent agreement is inflated by chance agreement and the base rates with which disorders are diagnosed in a sample. In their example, if two clinicians were randomly to assign *DSM* diagnoses to 6% of a community sample of 100 persons, chance agreement would produce an overall rate of agreement of about 88.8% (Shrout et al., 1987).

The index of agreement, kappa (K), was developed to address this statistical limitation (Cohen, 1960). Kappa reflects the proportion of agreement, corrected for chance; it varies from negative values reflecting agreement below chance, zero for chance agreement, and positive values reflecting agreement above chance to 1.0 for perfect agreement (Spitzer, Cohen Fleiss, & Endicott, 1967). The statistic weighted K was

Table 1 Relevant questions for selecting a diagnostic interview.

Content

Does the interview cover the relevant diagnostic system (e.g., RDC, *DSM-IV, ICD-10*)?
As an alternative to adhering to a single diagnostic system, does the interview provide polydiagnostic assessment (i.e., diagnoses for multiple diagnostic systems can be generated)?
Does the interview cover the relevant disorders?
Can irrelevant disorders be omitted?
Does the interview provide a sufficiently detailed assessment? That is, aside from diagnostic nosology is adequate information in other domains assessed (e.g., course of illness, family environment, social functioning)?
How are signs and symptoms rated (e.g., dichotomous ratings of presence vs. absence or continuous ratings of severity)?

Population

Is the interview applicable to the target population?
Relevant population considerations include adult vs. child, patient or nonpatient, general population, or family members of psychiatrically ill probands and whether the instrument will be used cross-culturally (is it available in languages for the cultures to be studied?).
Aside from general population considerations are there other exclusionary conditions to be aware of (e.g., age, education, cognitive functioning, or exclusionary clinical conditions)?

Time period

Does the interview cover the relevant time period (e.g., lifetime occurrence)?
Can the interview be used in longitudinal assessments to measure change?

Logistics of interview

How long does the interview take?
Does interview require or suggest use of informant (e.g., with child interviews)?

Administration/interviewer requirements

Who can administer the interview (e.g., lay interviewers, mental health professionals)?
How much training or experience is required to administer interview?
Does interview provide screening questionnaire to assist in expediting the interview (e.g., in personality disorder assessments)?
Are computer programs required and compatible with available equipment?

Guidelines and support

What guidelines for administration and scoring are available (e.g., user's manual)?
What training materials are available (videotapes, workshops)?
Is consultation available for training or clarification of questions regarding administration or scoring?

Reliability and validity

Is the interview sufficiently reliable?
Are reliability data available for the diagnoses and populations to be studied?
Are validity data available for the interview (e.g., concordance with other structured interviews, expert-obtained longitudinal data, and other noninterview measures)?

developed for distinguishing degrees of disagreement, providing partial credit when disagreement is not complete (Spitzer et al., 1967).

Standards for interpreting kappa suggest that values greater that 0.75 indicate good reliability, values between 0.50 and 0.75 indicate fair reliability, and values below 0.50 reflect poor reliability (Spitzer, Fleiss, & Endicott, 1978). Although the present review will adhere to the recommendations of Spitzer, Fleiss & Endicott (1978) that kappas below 0.50 indicate poor agreement or unacceptable reliability, it should be noted that other authors have proposed what appear to be more lenient criteria for evaluating kappa. For example, Landis and Koch (1977) suggest that kappas as low as 0.21–0.40 be considered as indicating "fair" agreement.

It is important to understand that reliability is not a quality inherent within an instrument that is transportable to different investigators or populations. All the measures described herein require interviewer training, and some instruments require an extensive amount of prior clinical experience and professional training.

Ultimately, the reliability of any structured interview will be dependent on the user of that interview. Although the present review will invite comparisons of reliability across studies it should be noted that reliability statistics such as kappa are influenced by a number of factors that constrain such comparisons. Differences in population heterogeneity, population base rates, and study methods will all influence reliability and should be considered in evaluating the literature.

4.05.1.1.2 Validity

In addition to reliability, the validity of a diagnostic assessment can also be evaluated. In the absence of an infallible or ultimate criterion of validity for psychiatric diagnosis, Spitzer (1983) proposed the LEAD standard: longitudinal, expert, and all data. Longitudinal refers to the use of symptoms or information that may emerge following an evaluation in order to determine a diagnosis. Additionally, expert clinicians making independent diagnoses based on all sources of information make a consensus diagnosis that will serve as a criterion measure. These expert clinicians should use all sources of data that have been collected over the longitudinal assessment including direct evaluation of the subject, interviewing of informants, and information from other professionals such as ward nurses and other personnel having contact with the subject. Typically, validity data such as that suggested by Spitzer are rarely available for structured interviews.

4.05.1.2 Overview

Within this chapter we provide an overview of the major structured interviews available for use with adult and child populations. The interviews included in this chapter are listed in Table 2. Due to space limitations we have focused on the review of broad diagnostic instruments and have not reviewed more narrow or specialized instruments that may address a single diagnosis or category of diagnoses. Each instrument will be reviewed with regard to its history and development and description of the instrument including its format and the diagnoses covered. Reliability data available will be presented and reviewed. Finally, a summary will be provided that intends to highlight the advantages and disadvantages inherent in each instrument. Interviews reviewed will address adult disorders, including personality disorders, and interviews for children and adolescents.

4.05.2 ADULT DISORDERS

4.05.2.1 Schedule for Affective Disorders and Schizophrenia

In order to address sources of diagnostic error arising from criterion variance, Spitzer, Endicott, and Robins (1978) developed the RDC. The RDC contains specific inclusion and exclusion criteria for 25 major diagnostic categories as well as subtypes within some categories. Disorders covered include schizophrenia spectrum disorders, mood disorders (depression and bipolar disorders), anxiety disorders (panic, obsessive-compulsive disorder, phobic disorder, and generalized anxiety disorder), alcohol and drug use disorders, some personality disorders (cyclothymic, labile, antisocial), and two broad categories of unspecified functional psychosis, and other psychiatric disorder. The major source of data for determining RDC diagnoses is the use of the Schedule for Affective Disorders and Schizophrenia (SADS; Endicott & Spitzer, 1978).

As originally developed there were three versions of the SADS: the regular version (SADS), the lifetime version (SADS-L), and a change version (SADS-C). A lifetime anxiety version of the SADS (SADS-LA; Fyer et al, 1989; Mannuzza et al., 1989) has been developed to assess RDC, *DSM-III*, and *DSM-III-R* criteria for almost all anxiety disorder diagnoses, in addition to all the diagnoses covered in the original SADS. The SADS has two parts: Part I provides a detailed description of current condition and functioning during the one week preceding the interview, Part 2 assesses past psychiatric disturbance. The SADS-L is similar to Part 2 of the SADS but the SADS-L focuses on both past and current disturbance. It is appropriate for use in populations where there is likely no current episode or when detailed information regarding the current condition is not required. Endicott and Spitzer (1978) estimate that the SADS can be completed in one and one-half to two hours depending on the disturbance of the individual being interviewed.

The SADS provides probe questions for each symptom rating. However, in addition to using the interview guide the rater is instructed to use all sources of information and to use as many supplemental questions as is required to make ratings. Part 1 of the SADS rates severity of symptoms when they were at their most extreme. Many items are rated for severity during the week prior to the interview and for severity during their most extreme during the current episode. Ratings are made on a seven-point scale from zero (no information) to six (e.g., extreme). The SADS provides defined

Table 2 Structured interviews included in review.

Adult

Schedule for Affective Disorders and Schizophrenia
Present State Examination
Structured Clinical Interview for *DSM-IV*/Axis I Disorders
Comprehensive Assessment of Symptoms and History
Diagnostic Interview for Genetic Studies
Diagnostic Interview Schedule
Composite International Diagnostic Interview

Personality disorders

Structured Interview for *DSM-IV* Personality
International Personality Disorder Examination
Structured Clinical Interview for *DSM-IV* Personality Disorders
Personality Disorder Interview-IV

Child and adolescent

Schedule for Affective Disorders and Schizophrenia for School Age Children Child Assessment Schedule
Child and Adolescent Psychiatric Assessment
Diagnostic Interview Schedule for Children
Diagnostic Interview for Children and Adolescents

levels of severity for each item. For example, in the screening items for manic criteria the item "less need for sleep than usual to feel rested" ratings are as follow: 1 = no change or more sleep needed; 2 = up to 1 hour less than usual; 3 = up to 2 hours less than usual; 4 = up to 3 hours less than usual; 5 = up to 4 hours less than usual; 6 = 4 or more hours less than usual.

In addition to the item ratings and the assignment of RDC diagnoses, the SADS can be used to provide eight summary scales: Depressive Mood and Ideation, Endogenous Features, Depressive-Associated Features, Suicidal Ideation and Behavior, Anxiety, Manic Syndrome, Delusions-Hallucinations, Formal Thought Disorder. These scales were determined by factor-analytic work using similar content scales, and an evaluation of clinical distinctions that are made in research of affective and schizophrenic disorders. The intent of the scales is to provide a meaningful summary of SADS information.

The SADS is intended for use by individuals with experience in clinical interviewing and diagnostic evaluation. Since clinical judgments are required in determining the need for supplemental questioning and in determining ratings, Endicott and Spitzer (1978) suggest that administration of the interview be limited to psychiatrists, clinical psychologists, or psychiatric social workers. However, these authors do note that interviewers with different backgrounds may be used but will require additional training. In one study using highly trained raters (Andreasen et al., 1981), reliability in rating videotaped SADS interviews was not affected by level of education (from medical degrees and doctorates in psychology to masters and other degrees) or years of prior clinical experience (from less than four years to more than 10 years).

Training videotapes and training seminars are available from the developers of the SADS at the New York State Psychiatric Institute, Columbia University. One example of training in the SADS is provided by Mannuzza et al. (1989) in the use of the SADS-LA. Training was conducted in three phases consisting of 50–60 hours over three to four months. In Phase 1 raters spent 20 hours attending lectures covering diagnosis, systems of classification, interviewing technique and the SADS-LA rater manual, reviewed RDC and *DSM-III* vignettes, and rated videotapes of interviews. In Phase 2 each rater administered the SADS-LA with one patient with an expert trainer and other raters observed. Interviews were subsequently reviewed and discussed. In the final phase, raters independently interviewed three or four patients who had already received the SADS-LA from an expert rater. This final phase allowed for test–retest reliability to be examined and provided an opportunity to discuss discrepancies in ratings during a consensus meeting.

4.05.2.2 Reliability

Initial reliability data for the SADS were reported by Spitzer, Endicott, and Robins

(1978) and Endicott and Spitzer (1978) for both joint interviews ($N = 150$) and independent test–retest interviews, separated by no more than one week ($N = 60$). For joint interviews, present and lifetime RDC diagnoses obtained kappas greater than 0.80 (median $K = 0.91$), with the exception of minor depressive disorder ($K = 0.68$). For test–retest interviews, reliability was somewhat attenuated but remained high with kappas greater than 0.55 for all disorders (median $K = 0.73$) with the exception of bipolar I (0.40). Endicott and Spitzer (1978) also reported reliability of the SADS items and eight summary scales using these same samples. For the 120 items of the current section of the SADS, reliability was high for both joint (90% of items interclass correlation coefficients [ICCs] equal to or greater than 0.60) and test–retest interviews (82% of items ICCs greater than or equal to 0.60). Summary scales also yielded high reliability for joint (ICC range = 0.82–0.99, median = 0.96) and test–retest interviews (ICC range = 0.49–0.91, median = 0.83). Spitzer et al. (1978) also examined the reliability of the SADS-L with first-degree relatives of patient probands ($N = 49$). All kappas were 0.62 or higher with the exception of other psychiatric disorder (0.46), median kappa = 0.86.

Two subsequent studies examined test–retest reliability (separate interviews conducted on the same day) of the SADS (Andreasen et al., 1981; Keller et al., 1981). In a study of 50 patients using the SADS-L, Andreasen et al. (1981) found ICCs equal to or greater than 0.62 for the major RDC diagnoses of bipolar I and II, major depressive disorder, alcoholism, and never mentally ill. The RDC subtypes of major depression also achieved ICCs equal to or greater than 0.60 with the exception of the subtype of incapacitating. Keller et al. (1981), using the SADS-L in a sample of 25 patients, obtained kappas equal to or greater than 0.60 for the RDC diagnoses of schizophrenia, schizoaffective-depressed, manic, major depressive disorder, and alcoholic. The major diagnoses of schizoaffective-manic and hypomanic had low reliability with kappas of 0.47 and 0.26, respectively. Keller et al. (1981) also found high reliability for background information items on social and educational background, and history of hospitalization (kappas greater than 0.73). Finally, individual items from manic, major depressive disorder, psychosis, alcohol and drug abuse, suicidal behavior, and social functional all achieved kappas above 0.56.

McDonald-Scott and Endicott (1984) evaluated the influence of rater familiarity on diagnostic agreement using the SADS. In this study modified SADS-C ratings were compared for two raters: one with extensive familiarity about subject's psychiatric history and course of illness and prior history and one who was blind to this history and had no prior contact with the subject. Quasi-joint interviews were conducted with the two raters. The nonblind rater was allowed to ask additional questions following the joint SADS interview, in the absence of the blind rater. Of four SADS-C summary scale scores all achieved ICCs of 0.79 or greater. At the item level, 92% of the 52 items had ICCs of 0.60 or greater. Rater differences in scoring suggested that the blind rater may have been somewhat more sensitive to items relating to dysphoria while the nonblind rater was more likely to identify some symptoms that may have been minimized or missed in the blind rater's interview (e.g., mania). However, these discrepancies were subtle and suggest that the SADS can achieve accurate assessment of current cross-sectional functioning whether or not raters have familiarity with the patient.

The inter-rater reliability of the SADS-derived *DSM-III* diagnoses in adolescents has been examined by Strober, Green, and Carlson (1981). Joint interviews were conducted with 95 inpatient adolescents and a family member. Raters independently reviewed all available collateral information prior to the SADS interview (e.g., medical and psychiatric records, school records, current nurses' observations). All diagnoses achieved kappas of 0.63 or greater with the exception of anxiety disorders of childhood (0.47) and undiagnosed illness (0.47). Although encouraging, these data should be viewed in the context of the use of joint interviews and the extensive use of collateral information to supplement the SADS.

Mannuzza et al. (1989) examined the reliability of the SADS-LA in a sample of 104 patients with anxiety disorders. Independent interviews were conducted with test–retest periods ranging from the same day to 60 days. Collapsing across RDC, *DSM-III*, and *DSM-III-R* anxiety disorder diagnoses, agreement for lifetime disorders achieved kappas of 0.60 or greater, with the exception of simple phobia. Examining lifetime anxiety diagnoses separately for each diagnostic system again suggested adequate reliability for most disorders (K range = 0.55–0.91), with the exception of RDC and *DSM-III-R* diagnoses of simple phobia and generalized anxiety disorder (Ks less than 0.49). Using this same sample, Fyer et al. (1989) assessed item reliability and factors contributing to disagreements. In general, symptoms were reliably rated with the exception of stimulus-bound panic (typical of simple phobia), near panic attacks, persistent generalized anxiety, six social and nine nonsocial irrational fears. Review of narratives and consensus meeting forms by Fyer et al. (1989) suggested that the largest source of disagreement was variation in information provided by subject (more than 60% of disagreements). Differences in rater interpretation of criteria resulted in 10–20% of the disagreements and rater error accounted for 10% of the disagreements.

The prior studies have examined test–retest reliability of lifetime diagnoses over brief periods. Two studies have examined the long-term test–retest reliability of SADS-L diagnoses. Bromet, Dunn, Connell, Dew, and Schulberg (1986) examined the 18-month test–retest reliability of the SADS-L in diagnosing lifetime major depression in a community sample of 391 women. Whenever possible, interviewers conducted assessments with the same subject at both interviews. Overall, reliability of lifetime diagnoses of RDC episodes of major depression was quite low. Of those women reporting an episode of major depression at either interview for the period preceding the first assessment, only 38% consistently reported these episodes at both interviews (62% reported a lifetime episode on one occasion but not another). For those women meeting lifetime criteria for a depressive episode at the first interview, fully 52% failed to meet criteria at the time of the second interview.

In a large-scale study of 2226 first degree relatives of probands participating in the National Institute of Mental Health (NIMH) Collaborative Program on the Psychobiology of Depression study, Rice, Rochberg, Endicott, Lavori, and Miller (1992) examined the stability of SADS-L-derived RDC diagnoses over a six-year period. The rater at the second interview was blind to the initial SADS-L. A large degree of variability in reliability was obtained for RDC diagnoses, with kappas ranging from 0.16 to 0.70. Diagnoses with kappas greater than 0.50 included major depression, mania, schizoaffective-mania, alcoholism drug use disorder, and schizophrenia. Diagnoses with low reliability as reflected by kappas below 0.50 were hypomania, schizoaffective-depressed, cyclothymia, panic disorder, generalized anxiety disorder, phobic disorder, antisocial personality, and obsessive-compulsive disorder. Rice et al. (1992) suggested that diagnostic reliability increases with symptom severity. In the studies of Bromet et al. (1986) and Rice et al. (1992), results indicated that there may be substantial error in the temporal stability of some SADS-L-derived lifetime diagnoses. This error may be particularly problematic in nonclinical community samples as studied in these investigations.

4.05.2.2.1 Summary

The development of the SADS represented significant progress in clinical assessment. The SADS has been used extensively in a number of research studies and a wealth of reliability data are available. The SADS provides a broad assessment of symptoms as well as severity ratings for many symptoms. However, the range of disorders covered is somewhat narrow (with an emphasis on schizophrenia, mood disorders, and anxiety disorders in the SADS-LA). Additionally, diagnostic criteria are based on the RDC, with the exception of anxiety disorders covered in the SADS-LA which provide *DSM-III-R* diagnoses.

4.05.2.3 Present State Examination

The Present State Examination (PSE) grew out of research projects in the UK requiring the standardization of clinical assessment. The PSE was not developed as a diagnostic instrument, as with the SCID and SADS. Rather, the PSE was intended to be descriptive and facilitate investigation of diagnostic rules and practices. At the time of the first publication of this instrument (Wing, Birley, Cooper, Graham, & Isaacs, 1967), the PSE was in its fifth edition. Currently, the ninth edition is widely used (Wing, Cooper, & Sartorius, 1974) and the tenth edition of the PSE is available (Wing et al., 1990). The PSE has been translated into over 40 languages and has been used in two large-scale international studies: the US–UK Diagnostic Project (Cooper et al., 1972) and the International Pilot Study of Schizophrenia (IPSS; World Health Organization, 1973).

The standardization of the PSE is achieved through the provision of a glossary of definitions for the phenomena covered by the interview. Additionally, specific series of questions with optional probes and cut-off points are also provided. Detailed instructions for rating the presence and severity of symptoms is also available. Despite this standardization, the developers have emphasized that it remains a clinical interview. The examiner determines the rating provided, evaluates the need for additional probe questions, and uses a process of cross-examination to address inadequate or inconsistent responses. As the name implies, the PSE was developed to ascertain present symptomatology and focuses on functioning in the month prior to the interview.

The eighth edition of the PSE was comprised of 500 items which were then reduced to 140 symptom scores. The ninth edition of the PSE reduced the number of items by having the 140 symptoms rated directly (the presence or absence of a symptom can be determined without asking as many questions, although additional probe questions are maintained in the ninth edition). Items receive one of three ratings. A zero indicates that a symptom is not present. If present, a symptom is rated as either one (moderate) or two (severe). Items are grouped into symptom scores based on item content and infrequency (Wing et al., 1974). The eighth edition takes approximately one hour to complete while the ninth edition takes approximately 45 minutes (Wing et al., 1974).

Symptoms can be further reduced to 38 syndrome scores by grouping together symptoms of similar content. For example, in the ninth edition the symptoms of worrying, tiredness, nervous tension, neglect through brooding, and delayed sleep are combined into the syndrome score of "Worrying, etc." These syndrome scores were intended to aid in the process of diagnosis by reducing the information to be considered, provide descriptive profiles, and provide a brief method of summarizing clinical information from other, non-interview, sources such as medical records by using a syndrome checklist.

Following the rating of items, a computer program (CATEGO) can be used to summarize PSE ratings further. For the ninth edition, the CATEGO program provides syndrome scores along with summary data for each syndrome (e.g., scores on constituent items). In the next stage, the program further summarizes the syndrome scores into six descriptive categories. The certainty of each descriptive category is also indicated (three levels of certainty are provided). Finally, a single CATEGO class (of 50) is assigned. Importantly, Wing (1983) has emphasized that the PSE and CATEGO program were not developed as diagnostic instruments *per se*. The CATEGO category or class assignments should not be considered diagnoses in the sense of DSM or ICD nosology. Rather, these summaries are provided for descriptive purposes. However, data from the US–UK Diagnostic Project and the IPSS have indicated reasonable convergence between CATEGO classes and clinical project diagnoses, especially when clinical history information is added to the PSE (reviewed in Wing et al., 1974).

Although short training courses lasting one week are available at various centers including the Institute of Psychiatry in London, Wing (1983) suggests that more extensive training is necessary. Wing (1983) recommends that at least 20 interviews be conducted under supervision in order to determine competency in administration of the PSE. Luria and Berry (1980) describe the stages of training used by these authors to achieve reliable PSE administration. In this study a general introduction and experience with unstructured symptom assessment was followed by reading and discussion of the PSE manual, the rating and discussion of 13 videotaped PSE interviews, and finally, participation in and observation of 12 live student-conducted PSE interviews followed by discussion.

4.05.2.3.1 Reliability

Early evaluations of the reliability of the PSE indicated promising agreement between raters. In the first reliability study conducted on early versions of the PSE (up to PSE-5), rater agreement was evaluated with both independent interviews and observation of interviews, or listening to audiotapes (Wing et al., 1967). Assignment to main categories suggested reasonable agreement, using percent agreement, of 83.7%. Examining five nonpsychotic symptoms, agreement also seemed satisfactory (range across studies $r = 0.53–0.97$). Reliability for nine psychotic symptoms, calculated for single interviews (tape recorded or observed) was also adequate (range of $r = 0.62–0.97$). Kendell, Everitt, Cooper, Sartorius, and David (1968) found a mean kappa for all items to be 0.77.

Luria conducted two reliability studies using the PSE-8 (Luria & McHugh, 1974; Luria & Berry, 1979). Luria and McHugh (1974) examined agreement using six videotaped PSE interviews. The authors examined agreement for 19 profiles of their own design. Patients were ranked on each category based on ratings of examiners. Reliability for these categories was generally adequate with Kendall's W coefficients greater than 0.73 except for behavioral items such as psychomotor retardation (0.66); excitement, agitation (0.47), catatonic, bizarre behavior (0.44); blunted, inappropriate, incongruous affect (0.60). In a subsequent study, Luria and Berry (1979) examined agreement on 20 symptoms deemed of diagnostic importance, 19 psychopathology profiles, and eight syndromes. Thirteen interviews were rated for reliability on the basis of videotapes; 12 were rated based on joint observational interviews. Reliability for videotape and live symptom ratings were adequate with median ICCs of 0.84 and 0.86, respectively (however, agitation or retardation and bizarre behaviors were judged to have poor reliability). Of the 19 profile ratings, the 13 had adequate reliability for videotaped (0.97) and live interviews (0.95). The six behavioral profiles were somewhat lower at 0.72 and 0.66, respectively. Syndrome agreement was high with generalized kappas above 0.91.

Three studies have examined inter-rater agreement for abbreviated versions of the PSE-8 and PSE-9 when used by nonpsychiatric raters (Cooper, Copeland, Brown, Harris, & Gourlay, 1977; Wing, Nixon, Mann, & Leff, 1977; Rodgers & Mann, 1986). Cooper et al. (1977) examined the agreement between ratings of a psychiatrist or psychologist and those obtained by a sociologist or sociology graduate student. Agreement was evaluated for both joint interviews and test–retest over one week. For joint interviews, with the exception of situational anxiety ($r = 0.34$), inter-rater agreement for the remaining 13 section scores was good, with correlations ranging from 0.65 to 0.96 (mean $r = 0.77$). Test–rest reliability was lower with five section scores having correlations below 0.40 and the mean for the 14 sections decreasing to 0.49. The correlation between total scores was 0.88 for inter-rater agreement and 0.71 for test–retest. Finally, presence vs. absence decisions for the 150 rated items indicated good reliability with a mean inter-rater kappa of 0.74, and a mean test–retest kappa of 0.54.

Wing et al. (1977) conducted two studies of a brief version of the PSE-9. In the first, 95 patients were interviewed independently (5–84 days between interviews) by a nonmedical interviewer and a psychiatrist. Agreement was examined for 13 syndromes and was unacceptably low with a mean kappa of 0.34 (range 0–0.49). The authors examined agreement on five symptoms relating to anxiety and depression. Poor agreement was found for these symptoms with kappas below 0.32. In the second study, 28 interviews were conducted by a nonmedical interviewer. Audiotapes of these interviews were rated by a psychiatrist. The mean kappa for syndrome scores was 0.52 (range = 0.25–0.85). Ratings of the five symptoms yielded kappas above 0.62 with the exception of free-floating anxiety ($K = 0.34$).

In a large population study, Rodgers and Mann (1986) assessed inter-rater agreement between nurses and a psychiatrist's rating of audiotapes. Audiotapes of 526 abbreviated PSE-9 interviews were evaluated. A statistic of index of association was used, although the authors report that this measure was highly correlated with kappa. Of 44 symptoms rated, six were considered too infrequent to evaluate. Of the remaining 38 symptoms the median index of association was 0.73 (range 0–0.96); seven items (18%) were unacceptably low in level of agreement (index of association less than 0.45): Expansive Mood, Ideomotor Pressure, Obsessional Checking/Repeating, Obsessional Ideas/Rumination, Hypochondriasis, Suicidal Plans or Acts, and Ideas of Reference. Thirteen syndrome scores derived from symptom ratings ranged from 0.29–0.94 (median = 0.76). Two syndrome scores were unacceptably low in agreement, Ideas of Reference (0.44) and Hypochondriasis (0.29).

In a recent study Okasha, Sadek, Al-Haddad and Abdel-Mawgoud (1993) examined rater agreement for assigning diagnosis based on *ICD-9*, *ICD-10*, and *DSM-III-R* criteria. The Arabic version of the PSE-9 was modified to collect extra data needed to make *ICD* and *DSM-III-R* diagnoses. One hundred adult inpatients and outpatients were interviewed by a single rater. An abstract form with PSE scores and other demographic and clinical information was then rated and diagnoses assigned. Overall kappa for nine broad diagnostic categories was acceptable (*ICD-9*, $K = 0.79$; *ICD-10*, $K = 0.82$; *DSM-III-R*, $K = 0.64$). Overall kappa values for more specific 18 diagnoses diminished somewhat but remained adequate (*ICD-9*, 0.62; *ICD-10*, 0.80, *DSM-III-R*, 0.63). Although this study indicates that PSE-9-derived information can be used to assign *ICD* and *DSM* diagnoses reliably it does not address the reliability of PSE-9 interviews themselves as diagnostic ratings were made from a single PSE abstract.

4.05.2.3.2 Supplements to the PSE

Two supplements to the PSE have been developed to address limitations in this instrument. These supplements address the assessment of lifetime psychopathology (McGuffin, Katz, & Aldrich, 1986) and change ratings (Tress, Bellenis, Brownlow, Livingston, & Leff, 1987). Because of the PSE's focus on the last month, its use in epidemiological studies is somewhat limited as these population-based investigations generally require the assessment of lifetime psychopathology. This concern led McGuffin et al. (1986) to modify the PSE. A Past History Schedule was developed to determine the dates of onset of worst episode of psychopathology, first psychiatric contact, and severest disturbance and recovery. Based on information obtained with the Past History Schedule, the PSE is then administered in three time formats: focusing on the last month, the most serious past episode, and modifying each PSE obligatory question with "have you ever experienced this?" Reliability assessment of this modified PSE using audiotaped interviews (McGuffin et al., 1986) has suggested adequate inter-rater agreement for the PSE CATEGO classes for past month (kappa range = 0.48–0.74), first episode (kappa range = 0.87–1), and ever (kappa range = 0.88–0.92). Rater agreement for dating past episodes was also found to be satisfactory (rank-order correlation coefficients, median = 0.83, range 0.54–0.99).

Tress et al. (1987) modified the PSE for purposes of obtaining change ratings. The authors suggest that the advantage of the PSE over other instruments available for ratings of clinical change are that the PSE gives data for clinical classification, provides clear definitions of items, and uses a structured interview format. The PSE Change Rating is administered following a standard PSE assessment. Items not rated positively on the initial assessment are discarded (as well as items that were not rated). Subsequent ratings are only made on these remaining items. These items are subsequently rated on an eight-point scale from zero (Completely Remitted) to seven (Markedly Worsened). Inter-rater agreement based on observed interviews was high for grouped symptom ratings (ICC range = 0.75–0.99) and selected individual symptoms (ICC range = 0.70–1).

4.05.2.3.3 Summary

As the first semistructured clinical interview the PSE has an extensive history with application in a number of studies. Additionally, the PSE has been translated into over 40 languages and has been employed in cross-cultural studies. A potential advantage of the PSE is that it is not constrained by a particular diagnostic system; however, the PSE-10 was designed to yield *ICD-10* and *DSM-III-R* diagnoses (Wing et al., 1990). The reliability data for the PSE are constrained in that assessments have included a variety of versions and modifications of the PSE using raters with a variety of training with different populations. Caution should be exercised in applying these data to an investigator's own intended use. Furthermore, reliability data for the PSE-10, which has undergone substantial revision, are not yet available, although a multisite investigation has been conducted (Wing et al., 1990). Additionally, examination of diagnostic reliability achieved with the PSE, while encouraging, has been limited to a few diagnoses and are not available for *DSM-IV*.

4.05.2.4 Structured Clinical Interview for *DSM-IV*/Axis I Disorders

The Structured Clinical Interview for *DSM-IV* (SCID-I) is a semistructured interview designed to assist in determining *DSM-IV* Axis I diagnoses (First, Gibbon, Spitzer, & Williams, 1996). Construction of the interview began in 1983 following the introduction of the *DSM-III*, which introduced operationalized, specific behavioral criteria. At this time existing clinical structured diagnostic interviews became limited in that they did not conform to the *DSM-III* criteria (e.g., the SADS and PSE). Although the Diagnostic Interview Schedule (DIS) was developed to yield *DSM-III* diagnoses, the DIS was designed to be used by lay interviewers in epidemiological studies. It was argued by Spitzer (1983) that the most valid diagnostic assessment required the skills of a clinician so that the interviewer could rephrase questions, ask further questions for clarification, challenge inconsistencies, and use clinical judgment in ultimately assigning a diagnosis. Thus, the SCID was initially developed as a structured, yet flexible, clinical interview for *DSM-III*, and subsequently *DSM-III-R*, diagnoses (Spitzer, Williams, Gibbon, & First, 1992).

The SCID-I has been revised several times due to criteria changes and field trials. The interview was primarily developed for use with adults, but may be used with adolescents. It is contraindicated for those with less than an eighth grade education, severe cognitive impairments, and experiencing severe psychotic symptoms (First et al., 1996). The SCID-I is available in Spanish, German, Portuguese, Dutch, and Hebrew, as well as English.

Separate versions of the SCID-I have been developed for research and clinical applications. The clinical version, the SCID-I-CV, is briefer than the research version and focuses primarily on key diagnostic information (excluding the supplementary coverage provided in the research version) and on the most commonly occurring diagnoses (First et al., 1996). Within the research version, three variations of the interview provide differing comprehensive coverage of the disorders, subtypes, severity, course specifiers, and history. The research versions have been used historically for inclusion, exclusion, and data collection of study participants (in over 100 studies), and are distributed in loose page format to allow the investigator to customize the SCID-I to meet the needs of their research. The SCID-P (patient edition) was designed to address psychiatric patients and provides thorough coverage of psychotic disorders and past psychiatric history. The SCID-NP (nonpatient edition) was developed to focus on nonpsychiatric patients, and subsequently screens for psychotic disorders and provides less comprehensive coverage of psychiatric history. The SCID-P with Psychotic Screen was developed for patients where a psychotic disorder is not expected (and therefore only screens for psychotic disorders), but has thorough coverage of psychiatric history.

The SCID-I can usually be administered in 60 to 90 minutes, contingent on the quantity of symptoms and disorders, and the ability of the interviewee to describe problems succinctly. It begins with an introductory overview followed by nine diagnostic modules. The overview provides open and closed questions that not only gather background information, but allows the interviewer to establish rapport with the interviewee before more detailed (and potentially uncomfortable) diagnostic questions are asked. The overview gathers information on demographics, work history, medical and psychiatric history, current stressors, substance use, and the interviewee's account of current and past problems (First et al., 1996).

There are nine diagnostic modules focusing on both current (usually defined as the past month) and lifetime assessment of diagnostic criteria: Mood Episodes, Psychotic Symptoms, Psychotic Disorders, Mood Disorders, Substance Use Disorders, Anxiety Disorders, Somataform Disorders, Eating Disorders, and Adjustment Disorders. An optional module covers Acute Stress Disorder, Minor Depressive Disorder, Mixed Anxiety Depressive Disorder, and symptomatic details of past Major Depressive/Manic episodes. Each page of the modules contains questions, reprinted *DSM-IV* criteria, ratings, and instructions for continuation. Initial questions are closed-ended and followed up with open-ended elaboration questions. If further clarification is needed, the interviewer is encouraged to ask supplementary (their own) questions, give examples, present hypothetical situations, and challenge inconsistencies. In essence, the interviewer is testing diagnostic hypotheses. The ratings are based not on the question response, but on fulfillment of *DSM-IV* criteria which are provided alongside the questions. The interviewer is encouraged to use alternate sources of information to assist in rating the criteria, such as observed behavior, medical records, and informants. Each criteria is rated as one of the following: ? = inadequate information, 1 = symptom clearly absent or criteria not met, 2 = subthreshold condition that almost meets criteria, and, 3 = threshold for criteria met.

Unlike other diagnostic interviews such as the SADS, PSE, or DIS, where diagnostic algorithms are applied following the interview, the SCID incorporates diagnostic criteria and decision making within the interview. The use of a "decision-tree approach" allows the interviewer to test hypotheses and concentrate on more problematic areas (Spitzer, Williams, Gibbons & First, 1992). In addition, this approach makes the interview more time efficient, allowing the interviewer to "pass over" areas of no concern. Following the interview, the interviewer is provided with concise summary scoring sheets to indicate the lifetime absence or threshold, and current presence of each disorder.

As a prerequisite for the SCID-I, the interviewer must possess adequate clinical experience and knowledge of psychopathology and diagnostic issues. The test developers recommend the following training: reading the administration sections of the *User's guide for the SCID-I*, reading the entire test booklet, reading the questions orally, practicing the SCID-I on a colleague/friend, watching a six-hour didactic training videotape titled *SCID-I 201*, role playing sample cases in the *User's guide for the SCID-I*, administering on an actual subject, conducting joint interviews (with independent ratings) followed by discussion sections, and examining inter-rater and test–retest reliability among interviewers (First et al., 1996). The following training materials and services are available: *User's guide for the SCID-I*, *SCID-I 201* video tape, videotape samples of interviews, on-site training, off-site monitoring, and SCID-I certification (under development). Following training, interviewers would benefit from ongoing supervision and feedback from an experienced SCID-I interviewer.

4.05.2.4.1 Reliability

Inter-rater agreement for the *DSM-III-R* version of the SCID-I was examined for 592 patients in five inpatient sites (one in Germany) and two nonpatient sites (Williams et al., 1992). At each site, two clinicians independently interviewed and diagnosed patients at least 24 hours but less than two weeks apart. In order to limit access to other information (e.g., chart review), interviewers were provided with only a brief summary of the hospital admission evaluation (circumstances of admission, number of prior hospitalizations, presenting problems). Diagnostic terms were excluded from the summary. For patients, overall weighted kappa was 0.61 for 18 current and 0.68 for 15 lifetime *DSM-III-R* diagnoses common to these sites. Disorders with poor agreement (i.e., *K*s below 0.50) were current diagnoses of dysthymia, agoraphobia without panic disorder, and social phobia, and the lifetime diagnosis of agoraphobia without panic disorder. Agreement for specific substance dependence diagnoses at a drug and alcohol treatment facility was high with all diagnoses having kappas above 0.61 except cannabis dependence and polydrug dependence (both kappas below 0.35). For nonpatients, overall weighted kappa was 0.37 for five current diagnoses and 0.51 for seven lifetime diagnoses common to these sites. The only diagnoses in nonpatients with a kappa of 0.50 or greater were current panic disorder, and lifetime diagnoses of alcohol dependence/abuse, other drug dependence/abuse, and panic disorder. Due to low occurrences, data were inconclusive for infrequent diagnoses.

Although generally satisfactory, these findings do indicate low agreement for some diagnoses. Williams et al. (1992) suggest several possible causes for low rater agreement in this study including the restriction of noninterview information, the focus on a broad range of diagnoses, and the flexible nature of the SCID in using clinical judgment. With regard to this last point, a review of a sample of audiotapes indicated that diagnostic disagreements were largely due to one interviewer's acceptance of a yes response without requesting elaboration while the other interviewer asked follow-up questions that ultimately determined that an initial yes response did not meet diagnostic criterion. As concluded by Williams et al. (1992) maximizing reliability on the SCID requires extensive training in the diagnostic criteria and an emphasis on not taking shortcuts but requiring that descriptions of behavior are elicited to justify each criterion.

Several other studies offer data on the reliability of the SCID-I, but the findings are confounded by small number of participants, changing *DSM* criteria and SCID-I revisions during the study, low base rates of disorders, and limited range of disorders (Segal, Hersen, & Van Hasselt, 1994). However, higher inter-rater agreement was observed in these studies ($K = 0.70–1$) compared to that obtained by Williams et al. (1992). The differences may have been due to the use of joint interviews (which controls for subject report variance) rather than independent interviews, access to noninterview information such as medical records and reports from other clinical staff, and the focus on a narrower range of diagnoses assessed.

4.05.2.4.2 Summary

The SCID is a well-established structured interview for determining *DSM-III-R* and *DSM-IV* diagnoses. Users may find the inclusion of diagnostic algorithms within the SCID and the use of skip-outs to result in a time-efficient interview. Reliability data from multiple sites indicate that the SCID can provide reliable *DSM-IV* diagnoses. Additionally, the SCID has some of the most extensive training materials and support available for any structured interview. The interview, user's guide, and all training materials have been completely updated for *DSM-IV*.

There are, however, a few disadvantages of the SCID-I. The interview does not cover a number of disorders, including infant, childhood, adolescence, cognitive, factitious, sexual, sleep, and impulse control disorders. Also, for those individuals interested in other diagnostic nosologies or needing to obtain broader clinical assessments, the restriction of the SCID to *DSM-IV* might be limiting. As with other structured interviews there is as yet no information currently available on the reliability of the SCID-I for the *DSM-IV* criteria. However, minor changes in the diagnostic criteria should not adversely affect reliability obtained with the *DSM-III-R* version.

4.05.2.5 Comprehensive Assessment of Symptoms and History

The Comprehensive Assessment of Symptoms and History (CASH; Andreasen, 1987) was developed without adherence to existing diagnostic systems (such as the *DSM* or *ICD*). The CASH adopted this approach based on observations that diagnostic criteria change over time and that methods of collecting information that conform to these criteria may be quickly outdated (Andreasen, Flaum, & Arndt, 1992). The CASH was designed for the study of psychosis and affective syndromes and is intended to provide a standardized assessment of psychopathology that will, ideally, yield diagnoses based on multiple criteria (both existing and future).

The CASH consists of nearly 1000 items divided into three sections: present state, past history, and lifetime history. The present state section consists of sociodemographic information intended to establish rapport and, subsequently, items pertaining to present illness. This section includes symptoms relating to the psychotic syndrome, manic syndrome, major depressive syndrome, treatment, cognitive assessment (laterality and a modified Mini-Mental Status Examination), a Global Assessment scale, and a summary of diagnoses for current episode.

Past history includes history of onset and hospitalization, past symptoms of psychosis, characterization of course, and past symptoms of affective disorder. To provide a detailed evaluation of phenomenology over time, for each symptom or sign, interviewers determine whether it was present during the first two years of illness, and whether it has been present for much of the time since onset. Finally, the lifetime history section includes history of somatic therapy, alcohol and drug use, premorbid adjustment, personality (schizotypal and affective characteristics), functioning in the past five years, Global Assessment scale and diagnoses for lifetime. Most items are given detailed definitions with suggested interview probes. Items are typically rated on a six-point Likert-type scale.

A number of measures are embedded within the CASH. Scales within the CASH include the Scale for Assessment of Negative Symptoms (Andreasen, 1983), Scale for Assessment of Positive Symptoms (Andreasen, 1984), most items for the Hamilton depression scale (Hamilton, 1960), and the Brief Psychiatric Rating Scale (Overall & Gorham, 1962), the Mini-Mental Status Exam (Folstein, Folstein, & McHugh, 1975), and the Global Assessment Scale (Endicott, Spitzer, Fleiss, & Cohen, 1976). These measures make the CASH useful for repeat assessments.

The CASH was intended for use by individuals with experience and training in working with psychiatric patients (e.g., psychologists, psychiatrists, nurses, or social workers). A training program has been developed for its use, which includes training videotapes conducted with patients presenting a range of psychopathology. Narratives and calibrated ratings for the CASH items are available from the authors

4.05.2.5.1 Reliability

A small reliability study conducted with 30 patients has been reported (Andreasen et al., 1992). Two forms of rater agreement were evaluated. First, patients were interviewed by a primary rater with a second rater observing and asking clarifying questions when necessary. Second, test–retest reliability was evaluated with a third rater interviewing the patient within 24 hours of the initial interview. All raters had access to medical records. Agreement between the two initial raters was generally good for the spectrum diagnoses (Schizophrenia Spectrum, $K = 0.86$; Affective Spectrum, $K = 1$). For specific *DSM-III-R* diagnoses (focusing on diagnoses with more than one case), the results were positive for schizophrenia ($K = 0.61$), bipolar affective disorder ($K = 1$), and major depression ($K = 0.65$). However, reliability for schizoaffective disorder was somewhat low ($K = 0.45$). Test–rest reliability was similarly positive with kappas above 0.74 for spectrum diagnoses and above 0.64 for specific *DSM-III-R* diagnoses with the exception of schizoaffective disorder ($K = 0.52$).

Because of the intent of the CASH to provide a reliable assessment of symptoms and functioning independent of diagnostic classification, it is important to examine the reliability of the individual items. Given the number of items, Andreasen et al. (1992) provide summaries of the intraclass correlation coefficients for the inter-rater and test–retest administrations. For inter-rater agreement, ICCs were generally high with three-quarters of the items having ICCs greater than or equal to 0.65. For the test–retest design, reliability was somewhat lower with approximately one-half of the items demonstrating ICC greater than or equal to 0.65.

Reliability data have been published for some more critical items or content areas (Andreasen et al., 1992). For history of illness, ICC values for both inter-rater and test–retest designs were quite adequate with values generally above 0.60 (median ICCs above 0.70). Reliability for items relating to manic and depressive syndromes was acceptable (median ICCs = 0.68 and 0.58, respectively). For positive and negative symptoms inter-rater and test–retest reliability was generally acceptable for "current" and "much of time since onset" time frames (global symptom scores ICCs greater than 0.65). However, test–retest reliability for negative symptoms rated for the "first two years of illness" and "worst ever" were unacceptably low (ICCs = 0 and 0.48, respectively). Test–retest data on premorbid and prodromal symptoms were very low (median ICCs = 0.37 and 0.25, respectively), while residual symptom ratings were somewhat better (median ICC = 0.60).

4.05.2.5.2 Summary

The CASH presents several advantages including its lack of adherence to any diagnostic system. This may afford the opportunity to collect a rich body of information on individuals. The comprehensiveness of the items is intended to allow for diagnoses for *DSM* and *ICD* to be generated while not narrowing the collection of information to these systems. Available reliability data are encouraging, with some exceptions as noted above. The availability of training materials including videotapes and consensus ratings is also attractive. The CASH also has companion instruments that are useful in the context of longitudinal assessments, providing baseline and follow-up assessment of psychosocial functioning and symptomatology.

The CASH is limited in several respects. First, because it seeks a full assessment of symptoms and history without regard to diagnostic criteria the entire CASH must be administered (however, some syndromes can be skipped if the interviewer already knows that a syndrome is not applicable). With nearly 1000 items this ensures lengthy assessment. Second, the CASH is limited to schizophrenia and affective syndromes and alcohol and drug use. Thus, it may not provide the breadth of symptom and diagnostic evaluation that some settings may require. Finally, although intended to be capable of assigning diagnoses based on extant nosological systems, the CASH may not always be capable of achieving this goal. Diagnostic criteria may require symptom information that does not conform to the information obtained with the CASH. Interested users should carefully evaluate the content of the CASH to ensure that relevant diagnoses can be made.

4.05.2.6 Diagnostic Interview for Genetic Studies

The Diagnostic Interview for Genetic Studies (DIGS; Nurnberger et al., 1994) was developed by participants in the NIMH Genetics Initiative. The need for the DIGS arose from perceptions that inconsistent findings in the genetics of psychiatric illnesses might be, in part, the result of differences in phenotypic assessment. Problems in assessment are highlighted in genetics research where family members are more likely to evince spectrum disorders and subclinical symptomatology.

The DIGS adopts a polydiagnostic approach that collects clinical information in sufficient detail to allow a variety of diagnostic definitions to be applied including *DSM-III-R* (a new version for *DSM-IV* is now available), modified RDC, RDC, Feighner criteria, *ICD-10*, and the European Operational Criteria (OPCRIT) Checklist (McGuffin & Farmer, 1990). As with the CASH, the advantage of this feature includes the collection of a broad data set for diagnostic entities whose definitions are sometimes ambiguous and evolving. However, the DIGS (unlike the CASH) explicitly collects information that conforms to several diagnostic systems including *DSM-IV*. Items from other interviews have been incorporated into the DIGS including the SADS, CASH, and DIS.

Like the SADS the DIGS provides standard probe questions and criterion-based definitions. Additionally, the DIGS requires clinical judgment for item ratings and in determining the need for further probe questions. Sections of the DIGS begin with one or two screening questions that, if denied, allow the interviewer to skip out of the remainder of the section. Questions are integrated so as to cover the various diagnostic criteria covered while maintaining an efficient flow of questioning. The interview can take 30 minutes to four hours depending on the complexity of the symptomatology (median time for symptomatic individuals is two and one-half hours).

The DIGS begins with a modified Mini-Mental Status examination in order to determine if the interview should be terminated as a result of cognitive impairment. The Introduction continues with demographics and extensive medical history screening questions. Somatization follows to enhance flow from medical history. An overview section assesses psychiatric history and course of illness and this information is summarized graphically in a time line to provide chronology of symptoms and episodes of illness. Mood disorders include major depression, mania/hypomania, and cyclothymic personality disorder. The DIGS also provides a detailed assessment of substance use history. Psychotic symptoms are assessed in great detail and psychotic syndromes are distinguished. Additionally, schizotypy is also assessed. A unique feature of the DIGS is an assessment for comorbid substance use. The aim of this section is to determine the temporal relationship between affective disorder, psychosis, and substance use. Suicidal behaviors, major anxiety disorders (except generalized anxiety disorder), eating disorders, and sociopathy are also evaluated. Finally, at the conclusion of the DIGS, the interviewer completes a Global Assessment scale (Endicott et al., 1976), the Scale for the Assessment of Negative Symptoms (Andreasen, 1983), and the Scale for the Assessment of Postive Symptoms (Andreasen, 1984).

Appropriate personnel to administer the DIGS are mental health professionals with clinical experience and familiarity with multiple diagnostic systems. In the study of Nurnberger et al. (1994), all but one interviewer had prior experience administering semistructured clinical interviews. Training as outlined in the reliability studies (Nurnberg et al., 1994) consisted of demonstration interviews by senior clinicians, the administration of a DIGS under supervision, and supplemental training involving three videotaped patient interviews.

4.05.2.6.1 Reliability

Test–retest reliability for the DIGS has been evaluated for major depression, bipolar disorder, schizophrenia, schizoaffective disorder and an "other" category (Faraone et al., 1996; Nurnberger et al., 1994). Test–retest reliability was evaluated within participating research sites as well as across sites. For the intrasite study, participants were independently interviewed with the DIGS over a period of no more than three weeks. For the intersite study, interviewers traveled to other research centers so that interviewers from different sites could assess the same subjects. These pairs of interviews were conducted within a 10-day period. With the exception of *DSM-III-R* schizoaffective disorder (K less than 0.50), *DSM-III-R* and RDC target diagnoses showed excellent reliability with kappas above 0.72 across the two studies.

4.05.2.6.2 Summary

The DIGS appears to be an excellent instrument for the study of the genetics of schizophrenia and affective disorders and other comorbid conditions. It provides an exhaustive assessment of symptomatology that allows for the comparison of findings across a number of diagnostic systems. Furthermore, it targets spectrum disorders and other comorbid conditions that may be relevant in family studies of schizophrenia and affective disorders. Although reliability has been shown to be high across different sites, these data are limited to schizophrenia and the affective disorders (but are not available for bipolar II and schizotypal personality); data are not available for other disorders such as the anxiety and eating disorders. As emphasized by the developers, the DIGS is not designed for routine clinical use. It is designed to be used by highly trained clinical interviewers for use in research settings.

4.05.2.7 Diagnostic Interview Schedule

The DIS (Robins, Helzer, Croughan, & Ratcliff, 1981) is a highly structured interview developed for use in large-scale epidemiological studies by the NIMH (the Epidemiological Catchment Area, ECA, projects). Because of the logistical constraints in general population studies, the use of traditional structured interviews administered by clinicians is prohibitive. The DIS was developed for use by lay interviewers with relatively brief training (one to two weeks). Thus, unlike structured interviews such as the SADS or SCID, the DIS minimizes the amount of discretion that an interviewer exercises in either the wording of questions or in determining the use of probe questions. Additionally, diagnoses are not made by the interviewer; rather, the data are scored, and diagnoses assigned, by a computer program.

The DIS was designed to provide information that would allow diagnoses to be made according to three diagnostic systems: *DSM-III* (APA, 1980), the Feighner criteria (Feighner et al. 1972), and the RDC. The interview covers each item or criterion in the form of one or more close-ended questions. Questions assess the presence of symptoms, whether they meet criteria for frequency and clustering in time, and whether they meet the age-at-onset criterion. The use of a Probe Flow Chart provides probes needed to determine severity and address alternative explanations. Rules concerning when and what probe questions to use are explicit in the interview. Nearly all questions can be read by lay interviewers as written. The DIS can take between 45 and 75 minutes to complete.

4.05.2.7.1 Reliability

vRobins et al. (1981) addressed the question of whether lay interviewers could obtain psychiatric diagnoses comparable to those obtained by a psychiatrist. In a test–retest design, subjects (mostly current or former psychiatric patients) were separately interviewed by a lay interviewer and a psychiatrist, both using the DIS. With the exception of *DSM-III* panic disorder ($K = 0.40$), kappas for all lifetime diagnoses across each diagnostic system were 0.50 or greater. Mean kappas for each diagnostic system were quite adequate: *DSM-III*, $K = 0.69$; Feighner, $K = 0.70$; RDC, $K = 0.62$. Further analysis of these data (Robins, Helzer, Ratcliff, & Seyfried, 1982) suggested that current disorders and severe disorders are more reliably diagnosed with the DIS than disorders in remission or borderline conditions.

Although the findings of Robins et al. (1981 1982) suggested acceptable concordance between lay interviewers and psychiatrists using the DIS, these data do not address whether the DIS would yield similar diagnoses as obtained by psychiatrists with broader clinical assessment than that allowed by the DIS alone. Additionally this first study was conducted with a largely psychiatric sample and may not be generalizable to the nonclinical populations the DIS was designed for. Anthony et al. (1985), using data from the ECA obtained in Eastern Baltimore, compared DIS-obtained diagnoses with psychiatrist-conducted clinical reappraisal examinations ($N = 810$). These clinical reappraisals were based on an augmented PSE consisting of 450 signs and symptoms and included all items of the PSE-9. Additionally, psychiatrists reviewed all available records. The two assessments were independent and the majority were completed within four weeks of each other. Diagnostic comparisons were for conditions present at the time of the interview or within one month prior to the interview. Results indicated substantial differences between DIS and psychiatrists' diagnoses. Except for schizophrenia and manic episode, one month prevalence rates for *DSM-III* diagnostic categories were significantly different for the two methods. Additionally, there was very low concordance between DIS-based diagnoses and those obtained by psychiatrists with kappas for selected diagnoses all below 0.36.

In a second major study based on ECA data collected in St. Louis, Helzer et al. (1985) compared lay-interview diagnoses with those obtained by a psychiatrist ($N = 370$). The psychiatrist re-examined subjects with the

DIS and were also allowed to ask whatever supplemental questions they deemed necessary to resolve diagnostic uncertainties following the DIS. The majority of psychiatrist interviews were conducted within three months of the lay interview. Diagnostic comparisons were made for lifetime diagnoses. Helzer et al.'s summary of the data was somewhat optimistic, indicating that corrected concordance was 0.60 or better for eight of the 11 lifetime diagnoses. However, when kappa statistics are examined, the results are less promising. Only one of eleven diagnoses obtained a kappa greater than or equal to 0.60, and eight diagnoses had kappas below 0.50. As summarized by Shrout et al. (1987) "for most diagnoses studied, the agreement in community samples between the DIS and clinical diagnoses is poor" (p. 177).

A number of other studies have been conducted comparing lay interviewer-administered DIS diagnoses with clinical diagnoses (e.g., Erdman et al., 1987; Escobar, Randolph, Asamen, & Karno, 1986; Ford, Hillard, Giesler, Lassen, & Thomas, 1989; Spengler & Wittchen, 1988; Wittchen, Semler, & von Zerssen, 1985; also see recent review by Wittchen, 1994). These studies are difficult to summarize and their interpretability is sometimes limited due to the use of a variety of assessment methodologies, diagnostic systems, and populations. Although some diagnoses achieve acceptable concordance levels between the DIS and clinical diagnoses, in total the results of these studies suggest limitations of the DIS consistent with the observations of Shrout et al. (1987). Wittchen (1994) has summarized particular problems apparent in the panic, somatoform, and psychotic sections of the DIS.

4.05.2.7.2 Summary

The DIS marked a significant development in the epidemiological study of psychopathology. The ECA findings based on the DIS have yielded important information about the epidemiology of a variety of disorders. However, studies examining the concordance between the DIS and clinical interviews conducted by psychiatrists suggest that there may be appreciable diagnostic inaccuracy in DIS-based diagnoses. Although the use of the DIS in epidemiological studies may continue to be warranted given the logistical constraints of such studies and the important data the DIS does obtain, the concern with diagnostic reliability should preclude the use of the DIS in settings where other structured diagnostic interviews can be used (e.g., the SADS or SCID).

4.05.2.8 Composite International Diagnostic Interview

The Composite International Diagnostic Interview (CIDI; Robins et al., 1988; World Health Organization [WHO], 1990) was developed at the request of WHO and the United States Alcohol, Drug Abuse, and Mental Health Administration. The CIDI was designed to serve as a diagnostic instrument in cross-cultural epidemiological and comparative studies of psychopathology. The initial version of the CIDI was based on the DIS, to cover DSM-III diagnoses, and initially incorporated aspects of the PSE since the PSE has been used in cross-cultural studies and reflects European diagnostic traditions. Some items from the DIS were altered either to provide further information for the PSE or to address language and content that would allow cross-cultural use. Additional questions were added to provide adequate coverage of the PSE-9. These PSE items were rewritten into the closed-ended format of the DIS.

Initial versions of the CIDI provided *DSM-III* diagnoses and updated versions, used in Phase II WHO field trials, now provide *DSM-III-R* and *ICD-10* diagnoses (WHO, 1990). The latest version of the CIDI has also eliminated all questions that are not needed for *DSM-III-R* (deletion of Feighner and RDC criteria) and has added items to assess *ICD-10* criteria. Furthermore, the PSE approach was abandoned and only same questions from the PSE were retained because they were relevant to *ICD-10* diagnoses. Revisions of the CIDI to meet *DSM-IV* criteria are in progress. A Substance Abuse Module was developed for the CIDI to be used alone or substituted for the less detailed coverage of drug abuse in the CIDI proper (Cottler, 1991). Other modules that have been developed or are under development include post-traumatic stress disorders, antisocial disorder, conduct disorder, pathological gambling, psychosexual dysfunctions, neurasthenia, and persistent pain disorder (Wittchen, 1994).

Like the DIS, the CIDI was intended to be used by lay interviewers with modest training (one week), and to be capable of rapid administration. In a multicenter study conducted in 15 countries the CIDI was found to be judged appropriate by the majority of interviewers (Wittchen et al., 1991). However, 31% of interviewers rated parts of the CIDI as inappropriate, in particular sections for schizophrenia and depression. The CIDI also takes a long time to administer: one-third of the interviews took one to two hours and another third lasted two to three hours (Wittchen et al., 1991). In this study the duration of the interviews may have been extended because of the assessment of predominantly clinical populations. One might expect briefer administration times with general population samples.

As with the DIS, training for the CIDI can be conducted in five days. No professional experience is necessary as the CIDI is intended to be used by lay interviewers. Training sites participating in the WHO field trials may be available to conduct courses. However, there is a CIDI user manual, a standardized training manual with item-by-item specifications, and computer scoring program available from WHO (1990).

4.05.2.8.1 Reliability

In an evaluation of the "prefinal" version of the CIDI (*DSM-III* and PSE) across 18 clinical centers in 15 countries, Wittchen et al. (1991) found high inter-rater agreement. Kappas for diagnoses were all 0.80 or greater with the exception of somatization (0.67). Wittchen (1994) has summarized test–retest reliability of the CIDI across three studies involving independent interviews conducted within a period of three days. Kappa coefficients for *DSM-III-R* diagnoses were all above 0.60 with the exception of bipolar II (0.59), dysthymia (0.52), and simple phobia (0.59).

Two studies have also examined the concordance between the CIDI and clinical ratings. Farmer, Katz, McGuffin, and Bebbington (1987) evaluated the test–retest concordance between CIDI PSE-equivalent items obtained by a lay interviewer and PSE interviews conducted by a psychiatrist. Interviews were conducted no more than one week apart. Concordance at the item level was unacceptably low. Of 45 PSE items, 37 (82%) achieved kappas below 0.50. Agreement was somewhat higher at the syndrome level but remained low for the syndromes of behavior, speech, and other syndromes (Spearman $r = 0.44$) and specific neurotic syndromes (Spearman $r = 0.35$).

Janca, Robins, Cottler, and Early (1992) examined diagnostic concordance between a clinical interviewer using the CIDI and a psychiatrist in a small sample of patients and nonclinical subjects ($N = 32$). Psychiatrists asked free-form questions and completed an *ICD-10* checklist following either the observation of lay-administered CIDI interview, or following their own administration of the CIDI. Overall diagnostic agreement appeared adequate with an overall kappa of 0.77. High concordance was also found for the *ICD-10* categories of anxiety/phobic disorders ($K = 0.73$), depressive disorders ($K = 0.78$), and psychoactive substance use disorders ($K = 0.83$).

4.05.2.8.2 Summary

The CIDI may be considered the next step beyond the DIS. This instrument incorporated lessons learned from the development of the DIS and has been subject to repeated revisions to enhance its reliability and cross-cultural application. The CIDI appears to have achieved somewhat better reliability than the DIS and it covers the latest diagnostic standards of *ICD-10* and *DSM-III-R* (soon to cover *DSM-IV*). However, the concordance between CIDI-obtained diagnoses and diagnoses obtained by other structured interviews administered by clinicians (e.g., SADS, SCID) remains unclear.

4.05.3 PERSONALITY DISORDERS

4.05.3.1 Structured Interview for *DSM-IV* Personality Disorders

Introduced in 1983, the Structured Interview for *DSM-III* Personality Disorders (SIDP) was the first structured interview to diagnose the full range of personality disorders in *DSM-III* (Stangl, Pfohl, Zimmerman, Bowers, & Corenthal, 1985). Subsequent versions have addressed Axis II criteria for the *DSM-III-R* (SIDP-R; Pfohl, Blum, Zimmerman, & Stangl, 1989) and the *DSM-IV* (SIDP-IV; Pfohl, Blum, & Zimmerman, 1995).

SIDP-IV is organized into 10 topical sections: Interests and Activities, Work Style, Close Relationships, Social Relationships, Emotions, Observational Criteria, Self-perception, Perception of Others, Stress and Anger, and Social Conformity. This format is intended to provide a more conversational flow and is thought to enhance the collection of information from related questions and facilitate the subsequent scoring of related criteria.

The SIDP-IV can be administered to a patient and to an informant, requiring 60–90 and 20–30 minutes, respectively (Pfohl et al., 1995). In addition, the interview takes 20–30 minutes to score. Each page of the interview provides questions, prompts, diagnostic criteria, and scoring anchors. The informant interview is composed of a subset of questions from the patient interview. Two alternate versions of the SIDP-IV are available. The SIDP-IV Modular Edition is organized by personality disorders rather than by topical sections. This modular form permits the interviewer to focus on disorders of interest and to omit others. The Super SIDP is an expanded version that includes all questions and criteria necessary to assess *DSM-III-R*, *DSM-IV*, and *ICD-10* personality disorders.

Several instructions for administering the SIDP-IV are noteworthy. First, the SIDP uses a "five year rule" to operationalize criteria involving an enduring pattern that represents personality. Thus, "behavior, cognititions, and feelings that have predominated for most of the last five years are considered to be representative of the individual's long-term personality functioning" (Pfohl et al., 1995, p. ii). The SIDP-IV is intended to follow assessment of episodic (Axis I) disorders in order to assist the interviewer in ruling out the influence of temporary states of behavior described by the patient. Second, the patient's responses are not given a final rating until following the interview. This is intended to allow for all sources of information to be reviewed before rating. However, unlike previous versions of the SIDP, the SIDP-IV now provides a rater with the opportunity to rate or refer to specific *DSM-IV* criteria associated with each set of questions. Third, use of an informant is optional and Pfohl et al. (1995) note that while the frequency of personality diagnoses may increase with the use of informants there appears to be little effect on predictive validity.

Each diagnostic criterion is scored as one of the following: 0 = not present or limited to rare isolated examples, 1 = subthreshold: some evidence of the trait but it is not sufficiently pervasive or severe to consider the criterion present, 2 = present: criterion is present for most of the last five years (i.e., present at least 50% of the time during the last five years), and 3 = strongly present: criterion is associated with subjective distress or some impairment in social or occupational functioning or intimate relationships. Unlike other personality interviews (e.g., the IPDE and SCID-II), scores of both 2 and 3 count towards meeting diagnostic criteria (Pilkonis et al., 1995).

The SIDP-IV is an interview requiring clinical skill in determining the need for additional probe questions and to discriminate between personality (Axis II) disorders and episodic (Axis I) disorders. The developers of the SIDP-IV recommend one month of intensive training to administer the interview properly (Pfohl et al., 1995; Standage, 1989). Pfohl et al. (1995) have reported success with interviewers having at least an undergraduate degree in the social sciences and six months of previous experience with psychiatric interviewing. Training videotapes and workshops are available from the developers of the SIDP-IV.

4.05.3.1.1 Reliability

Several investigations of inter-rater reliability reveal poor to good agreement. Using the SIDP-R, Pilkonis et al. (1995) found that inter-rater agreement for continuous scores on either the total SIDP-R score or scores from Clusters A, B, and C, was satisfactory (ICCs ranging from 0.82 to 0.90). Inter-rater reliability for presence or absence of any personality disorder with the SIDP-R was moderate with a kappa of 0.53. Due to infrequent diagnoses, mixed diagnoses, and the number of subthreshold protocols, in this study kappas for individual diagnoses were not provided.

Stangl et al. (1985) conducted SIDP interviews on 63 patients (43 interviews were conducted jointly, and 20 interviews were separated by up to one week). The kappa for presence or absence of any personality disorder was 0.66. Only five personality disorders occurred with enough frequency to allow kappa to be calculated: dependent (0.90), borderline (0.85), histrionic (0.75), schizotypal (0.62), and avoidant (0.45). Using the SIDP among a small sample of inpatients, Jackson, Gazis, Rudd, & Edwards (1991) found inter-rater agreement to be adequate to poor for the five specific personality disorders assessed: borderline ($K = 0.77$), histrionic (0.70), schizotypal (0.67), paranoid (0.61), and dependent (0.42).

The impact of informant interviews on the diagnosis of personality disorders and inter-rater agreement for the SIDP was assessed by Zimmerman, Pfohl, Stangl, and Corenthal (1986). Inter-rater agreement (kappa) for the presence or absence of any personality disorder was 0.74 before the informant interview and 0.72 after the informant interview. Kappas for individual personality disorders were all 0.50 or above. Reliability did not appear to benefit or be compromised by the use of informants. However, the informant generally provided additional information on pathology and, following the informant interview, diagnoses that had been established with the subject only were changed in 20% of the cases (Zimmerman et al., 1986).

In an examination of the long-term test–retest reliability of the SIDP, Pfohl, Black, Noyes Coryell, and Barrash (1990) administered the SIDP to a small sample of depressed inpatients during hospitalization and again 6–12 months later. Information from informants was used in addition to patient interviews. Of the six disorders diagnosed three had unacceptably low kappas (below 0.50): passive-aggressive (0.16), schizotypal (0.22), and histrionic (0.46). Adequate test–retest reliability was obtained for borderline (0.58), paranoid (0.64), and antisocial (0.84).

4.05.3.1.2 Summary

The SIDP-IV represents the latest version of the first interview to assess the spectrum of

DSM personality disorders. Although originally developed to be administered in a topical format to assess *DSM* personality disorders the SIDP-IV now provides alternative formats for assessing specific disorders without administering the entire SIDP-IV and for assessing *ICD* diagnoses. Reliability data are encouraging for some disorders. However, these data are limited to selected disorders using the SIDP and reliability data have not been presented for specific disorders using the SIDP-R (Pilkonis et al., 1995). No reliability data are available for the SIDP-IV. Little data are available concerning the long-term test–retest reliability of the SIDP. The SIDP-IV does not come with a screening questionnaire to assist in identifying personality disorders that might be a focus of the interview.

4.05.3.2 International Personality Disorder Examination

The International Personality Disorder Examination (IPDE; Loranger et al., 1995), a modified version of the Personality Disorder Examination (PDE), is a semistructured interview designed to assess personality disorders in both the *DSM-IV* and *ICD-10* classification systems. The PDE was initially developed in the early 1980s to assist in the diagnosis of personality disorders. At that time, only structured interviews existed that focused on Axis I mental disorders. A highly structured lay-administered interview for personality disorders was thought to be inappropriate due to the complexity of diagnostic criteria and level of inference required (Loranger et al., 1994). The first version of the PDE was completed in 1985. Beginning in 1985, international members of the psychiatric community attended several workshops to formulate an international version of the PDE, the IPDE, which was developed under the WHO, and the US Alcohol, Drug Abuse, and Mental Health Administration System. The purpose of the IPDE was to assess personality disorders in different languages and cultures.

The IPDE interview surveys behavior and life experiences relevant to 157 criteria and can be used to determine *DSM-IV* and *ICD-10* categorical diagnoses and dimensional scores of personality disorders. The IPDE is not recommended for use on individuals who are under the age of 18, agitated, psychotic, severely depressed, below normal intelligence, or severely cognitively impaired. The interview is available in the following languages: English, Spanish, French, German, Italian, Dutch, Danish, Norwegian, Estonian, Greek, Russian, Japanese, Hindi, Kannada, Swahili, and Tamil (Loranger et al., 1995).

The IPDE contains materials for determining both *DSM-IV* and *ICD-10* diagnoses. However, due to the long length of the interviews noted in the field trials (mean interview length was 2 hours, 20 minutes), the interview is distributed in two modules for each classification system. Furthermore, clinicians and researchers can easily administer specific personality modules to suit their purpose (Loranger et al., 1994).

A self-administered IPDE screening questionnaire may be administered prior to the interview in order to eliminate subjects who are unlikely to have any or a particular personality disorder. Similar to the SCID-II (described below), a corresponding low diagnostic threshold (for endorsement) is set for each question. If three or more items are endorsed for a specific personality disorder, then the interview is administered for that personality disorder.

In the attempt to establish reliable diagnoses, the IPDE interview utilizes two distinct guidelines. First, the behavior or trait must be present for at least five years to be considered an expression of personality with some manifestations (based on the disorder) occurring within the past 12 months (Loranger et al., 1995). This strict criterion is adopted to ensure the enduring nature of behavior and life experiences, and rule out transient or situational behaviors. A second guideline for the IPDE is that one criterion must be met before the age of 25. However, if an individual develops a personality disorder later in life (with no criterion exhibited prior to age 25) the IPDE provides an optional "late onset" diagnosis (Loranger et al, 1995).

The developers constructed the IPDE interview not only to be clearly organized, but to "flow" naturally. As a result, the diagnostic criteria are not ordered by cluster or disorder, but by six sections that assess major life domains: Work, Self, Interpersonal Relationships, Affects, Reality Testing, and Impulse Control. Each section begins with open-ended questions that provide a transition between sections and allow the interviewer to gather general background information. Closed-ended and elaboration questions follow for each criterion (Loranger et al., 1995). Each individual page of the IPDE is designed to optimally assist the interviewer in correctly determining if the diagnostic criteria is met. Each page contains: personality disorder and criterion number, structured questions, reprinted *DSM-IV* or *ICD-10* criteria, notes on determining criteria, descriptions of scoring criteria, and scoring areas for both interview and informants.

The scoring of the IPDE interview is similar to other semistructured interviews. Prior to the interview, the developers recommend that collecting information or conducting interviews on Axis I disorders be completed to assist in scoring the criteria. Each trait or behavior (i.e., criterion) is scored as one of the following: absent or normal (0), exaggerated or accentuated (1), criterion level or pathological (2), and interviewee is unable to furnish adequate information (?). Some items may not be applicable to all interviewees and scored not applicable. The IPDE also allows the interviewer to rate each criterion based on informants (Loranger et al., 1995). The IPDE manual provides information on the scope and meaning of each criterion, and provides guidelines and anchor points for scoring. The manual does not recommend challenging the interviewee on inconsistencies with informants during the interview, due to the potential threat to rapport. However, challenging discrepancies occurring within the interview is encouraged. The IPDE may be hand scored or computer scored (program provided by publishers). Hand-scored algorithms and summary sheets are provided to assist in determining categorical diagnoses and dimensional scores.

The IPDE developers recommend that only those with the clinical experience and training to make psychiatric diagnoses independently use the IPDE (Loranger et al., 1994). The IPDE manual strongly discourages the use of IPDE by clinicians early in their training, research assistants, nurses, medical students, and graduate students. In addition, the interviewer should have familiarity with the *DSM-IV* and *ICD-10* classification systems (Loranger et al., 1995). The test manual recommends the following training: read the interview and manual thoroughly, practice on several participants to get familiar with the interview, administer with an IPDE-experienced interviewer, and discuss any problems in administration or scoring. Before administering the IPDE, the interviewer should have thorough knowledge of the scope and meaning of each criterion and scoring guidelines. IPDE training courses are offered at the WHO training centers.

4.05.3.2.1 Reliability

The IPDE field trial conducted in 11 countries and 14 centers evaluated inter-rater reliability in joint interviews as well as test–retest reliability with an average test–retest interval of six months (the test–retest interviews were conducted by the same interviewer). Results indicated overall weighted kappa for individual definite personality disorders to be 0.57 for the *DSM-III-R* and 0.65 for the *ICD-10* (Loranger et al., 1994). Median kappas for definite or probable personality diagnoses were 0.73 for *DSM-III-R* and 0.77 for *ICD-10*. Using broader definite or probable criteria, kappa for any personality disorder increased to 0.70 for *DSM-III-R* and 0.71 for *ICD-10*. For temporal stability, kappas for the presence or absence of a personality disorder were 0.62 for *DSM-III-R* and 0.59 for *ICD-10*. Inter-rater reliability was higher for dimensional scores with ICCs ranging from 0.79 to 0.94 for the *DSM-III-R* and 0.86 to 0.93 for the *ICD-10*. Temporal stability for dimensional scores was also high with ICCs ranging from 0.68 to 0.92 for *DSM-III-R* and from 0.65 to 0.86 for *ICD-10*.

Pilkonis et al. (1995) also evaluated the reliability of the third version of the PDE. Intraclass correlations for total scores or cluster scores ranged from 0.85 to 0.92. Inter-rater agreement (kappa) for the presence or absence of any personality disorder was 0.55.

Loranger et al. (1991) examined inter-rater agreement and test–retest reliability of the PDE in a sample of psychiatric inpatients. Second administrations of the PDE were conducted one week to six months later by a separate interviewer blind to the initial assessment. Inter-rater agreement between two raters was assessed at both the initial and repeated interview. At the first interview kappas for inter-rater reliability ranged from 0.81 to 0.92 (median = 0.87). At the repeat interview kappas for inter-rater reliability ranged from 0.70 to 0.95 (median = 0.88). At follow up there was a significant reduction in the number of criteria met on all disorders except schizoid and antisocial. Stability of the presence or absence of any personality disorder was moderated with a kappa of 0.55.

O'Boyle and Self (1990) interviewed 20 patients with a depressive disorder for a personality disorder. Eighteen patients were re-interviewed across a mean of 63 days for the presence or absence of personality disorder. Intraclass correlations were 0.89 to 1 and inter-rater agreement (kappa) was 0.63. Depressive disorders did not consistently affect categorical diagnoses, but dimensional scores were higher during depressed periods.

4.05.3.2.2 Summary

The IPDE has a number of strengths to recommend its use. First, the IPDE (and PDE) has demonstrated medium to high inter-rater agreement and temporal reliability for both categorical diagnoses and dimensional scores. In addition, preliminary investigations into the influence of Axis I disorders (e.g. depression) on the assessment of personality disorders indicate no significant influence on PDE categorical diagnoses. Second, a detailed training manual accompanies the interview, which provides thorough instructions and scoring algorithms. Third, a unique feature of the IPDE is dual coverage of the *DSM-IV* and *ICD-10* criteria. Fourth, in addition to providing categorical diagnoses, the IPDE measures dimensional scores which provide information about accentuated normal traits below the threshold required for a personality disorder. Finally, the IPDE is available in several languages and has been studied in 11 countries.

The IPDE, while quite comprehensive, is flexible enough to permit more economical administration. The *DSM-IV* and *ICD-10* modules can be administered separately. Furthermore, rather than administer in thematic organization, the IPDE can be limited to diagnostic modules of interest. A self-administered screening questionnaire is available to assist in identifying personality disorders that might be of focus in the interview.

Inter-rater agreement between the SCID-II and IPDE have led some researchers to conclude that the IPDE (and PDE) presents more stringent guidelines to fulfill personality disorder criteria (Hyler, Skodol, Kellman, Oldhan, & Rosnik, 1990; Hyler, Skodol, Oldham, Kellham, & Doldge, 1992; O'Boyle & Self, 1990). This stringent determination is probably due to the consistent five-year time period requirement for personality traits. In conclusion, the specificity of the instrument is increased (fewer false positives) but this may be at the cost of decreased sensitivity (more false negatives).

4.05.3.3 Structured Clinical Interview for *DSM-IV* Personality Disorders

The Structured Clinical Interview for *DSM-IV* Personality Disorders (SCID-II) is a structured interview that attempts to provide an assessment of the 11 *DSM-III-R* personality disorders, including the diagnosis of self-defeating personality disorder, which is included in Appendix A of *DSM-III-R* (First, Spitzer, Gibbon, & Williams, 1995). The SCID-II interview can be used to make categorical or dimensional personality assessments (based on the number of items judged present). The SCID-II was developed as a supplementary module to the SCID-I, but was redesigned in 1985 to be a separate and autonomous instrument due to different assessment procedures and length of interview (First et al., 1995).

In conducting the SCID-II, it is extremely important to evaluate the interviewee's behavior out of the context of an Axis I disorder (Spitzer et al., 1990). The test developers recommend an evaluation of Axis I disorders prior to the SCID-II with either a SCID-I or some other Axis I evaluation.

A self-report screening questionnaire is provided to improve time efficiency. Each of the 113 items on the questionnaire corresponds to a diagnostic criterion and is purposefully constructed to have a low threshold for a positive response, and is therefore for screening purposes only. Interviewers should probe all items coded "yes" in the questionnaire. Under most circumstances, the interviewer does not need to interview for the negatively endorsed criteria, due to the low probability of psychopathology. Negative questionnaire responses should be followed up when either the interviewer suspects that the criterion or personality disorder is actually present or when the number of items endorsed positively in the interview is only one item below that required for making a diagnosis (in which case all questions for that diagnosis should be probed).

Utilizing the screening questionnaire, the SCID-II interview can usually be administered in 30–45 minutes (First et al., 1995). First, Spitzer, Gibbons, Williams, Davies, et al. (1995) interviewed 103 psychiatric patients and 181 nonpatients, and calculated a mean interview time of 36 minutes. A unique feature of the SCID-II is that the interview begins with a brief overview which gathers information on behavior and relationships, and provides information about the interviewee's capacity for self-reflection. This allows the interviewer not only to establish rapport, but also allows interviewees to provide a description and consequences of their behavior in their own words (Spitzer et al., 1990). Following the overview, the interview progresses through each relevant disorder.

The format and structure of the SCID-II is very similar to that of the SCID for Axis I disorders. Each page of the interview contains questions, reprinted *DSM-IV* criteria, guidelines for scoring, and space for scores (Spitzer et al., 1990). Initial questions are open-ended and followed up with prompts for elaboration and examples. If further clarification is needed, the interviewer is encouraged to ask supplementary (their own) questions, give examples, present hypothetical situations, and challenge inconsistencies (Spitzer et al., 1990). There are usually two to three interview questions for each personality disorder criterion. In essence, the interviewer is testing diagnostic hypotheses. The ratings are based not on the question response, but an fulfillment of *DSM-IV* criteria. The interviewer is encouraged to use alternate sources of information to assist in rating the criteria, such as observed behavior, medical records, and informants. With slight modifications, the SCID-II can be administered to an informant (First et al., 1995). Each criterion is rated as one of the following: ? = inadequate information, 1 = symptom clearly absent or criteria not met, 2 = subthreshold condition that almost meets criteria, and 3 = threshold for criteria met.

A rating of "3" is scored only when the interviewee provides convincing, elaborative, and/or exemplary information. A rating of "3" is reserved only for criteria that fulfill the following three guidelines: pathological (outside normal variation), pervasive (occurs in a variety of places and/or with a variety of people), and persistent (occurs with sufficient frequency over at least a five-year period). Specific guidelines for a "3" rating are provided for each criterion in the body of the interview.

Due to the similarity in interview procedures, SCID-II training procedures are almost identical to SCID-I training. As with the SCID-I, clinical judgment is required in the administration and scoring of the SCID-II and thus requires interviewers to have a full understanding of *DSM* nosology and experience in diagnostic interviewing. A user's manual is available for the SCID-II and demonstration videotapes are available. Training workshops can also be arranged with the developers.

4.05.3.3.1 Reliability

The test–retest reliability of the SCID-II was examined within an investigation of the reliability of the Axis I SCID (Williams et al., 1992). In this study, First, Spitzer, Gibbons, Williams, Davies, et al. (1995) administered the SCID-II to 103 psychiatric patients and 181 nonpatients. Two raters independently administered the SCID-II between 1 and 14 days apart. Each SCID-II was preceded by an Axis I SCID evaluation. The SCID-II Personality Questionnaire was given only on the occasion of the first assessment (both SCID-II interviews used the same questionnaire results). Overall weighted kappas were 0.53 and 0.38 for patients and nonpatients, respectively. For the patient sample, kappas were above 0.5 for avoidant, antisocial, paranoid, histrionic, and passive-aggressive personality disorders, and below 0.5 for dependent, self-defeating, narcissistic, borderline, and obsessive-compulsive personality disorders. For the nonpatient sample, kappas were above 0.5 for dependent, histrionic, and borderline personality disorders, and below 0.5 for avoidant, obsessive-compulsive, passive-aggressive, self-defeating, paranoid, and narcissistic personality disorders.

Using the Dutch version of the SCID-II, Arntz et al. (1992) randomly selected 70 mental health center outpatients and conducted SCID-II interviews. Inter-rater reliability was determined by comparing criteria scores between a primary interviewer and an observer. With the exception of a few criteria, inter-rater reliability for each criterion was good to excellent. Eighty-four of 116 *DSM-III-R* criteria had ICCs higher than 0.75, and 14 had reliability ranging from 0.60 to 0.75. Inter-rater reliability was not able to be rated for 12 criteria due to lack of variance. Inter-rater agreement for specific personality disorders was good with kappas ranging from 0.65 to 1.

Several other studies report good to excellent inter-rater reliability and agreement using the SCID-II (Brooks, Baltazar, McDowell, Munjack, & Bruns, 1991; Fogelson, Nuechterlein, Asarnow, Subotnik, & Talovic, 1991; Malow, West, Williams, & Sutker, 1989; Renneberg, Chambless, & Gracely, 1992). However, these studies contained two or more of the following limitations: restricted number of personality disorders, a homogeneous population and a small number of participants.

4.05.3.3.2 Summary

The SCID-II differs from other personality interviews in several respects. Although other interviews have a disorder-based format available, only the SCID-II has this format as its primary (and only) format of administration. First et al. (1995) maintain that the grouping of questions based on disorder may more closely approximate clinical diagnostic practice and that this grouping forces interviewers to consider criteria in the context of the overarching theme of the disorder. One disadvantage is that the lack of a thematically organized format limits an interviewer's choices, and some have raised concerns that disorder-based organization results in redundancy and repetition with similar items across different diagnoses. Also, the organization of items by disorder may create "halo" effects where a positive criterion rating may bias an interviewer's rating of similar items.

Although the SCID-II screening questionnaire is unusual the IPDE now has a screening questionnaire as well (neither the SIDP-IV or the PDI-IV use screening questionnaires). The SCID-II has shown reliability comparable to other interviews and has been used in a number of studies. The shared format between the SCID-II and the SCID for Axis I disorders should facilitate training on the two measures and may ease the typical requirement that Axis I disorders are assessed and taken into consideration when conducting personality disorder examinations.

4.05.3.4 Personality Disorder Interview-IV

The Personality Disorder Interview-IV (PDI-IV; Widiger, Mangine, Corbitt, Ellis, & Thomas, 1995) is a semistructured interview developed to assess 10 *DSM-IV* personality disorders as well as the two *DSM-IV* personality criteria sets provided for further study (depressive personality disorder, passive-aggressive personality disorder). The PDI-IV is the fourth edition of the Personality Interview Questionnaire (PIQ). The name change was based, in part, on the intent to provide a more descriptive title as the PDI focuses on the assessment of disordered personality.

The PDI-IV provides questions for the assessment of the 94 diagnostic criteria that relate to the 12 *DSM-IV* personality disorders. Criterion ratings are made on a three-point scale (0 = not present, 1 = present according to *DSM-IV* criteria, 2 = present to more severe or substantial degree). Questions from the PDI-IV were selected as useful in determining criterion ratings and additional questions are provided for further elaboration if time allows. However, given the questionnaire's semistructured format, the interviewer may deviate from questions to obtain further information or to address inconsistencies. It is suggested that all positive responses be followed by a request for examples.

The PDI-IV can be administered in a manner either organized by thematic content (as with the IPDE and SIDP) or by *DSM-IV* diagnostic category (as with the SCID-II). Separate interview booklets are provided for these two forms of administration. For occasions when all personality disorders will be assessed, it is recommended that the thematic administration be used. Content areas in the thematic format include Attitudes Towards Self, Attitudes Toward Others, Severity or Comfort with Others, Friendships and Relationships, Conflicts and Disagreements, Work and Leisure, Social Norms, Mood, and Appearance and Perception. The diagnostic format may be preferable when only particular disorders must be assessed. Ratings can be used either to derive categorical or dimensional ratings for *DSM-IV* personality disorders.

The PDI-IV comes with an extensive manual that discusses general issues regarding administration but also provides a thorough discussion of each personality disorder in separate chapters. Each chapter provides an overview of the history and rationale for the personality disorder including discussion of the development of the criterion in *DSM* as well as *ICD* and other criterion sets. Each criterion item is discussed with regard to revisions and rationale for each item, ratings and questions, and issues and problems relevant to assessing that criterion.

The PDI-IV does not include the use of a self-report questionnaire. However, Widiger et al. (1995) do recommend that a stand-alone self-report inventory be used to assess personality. Scores from such a questionnaire may then be used to inform selection of the most relevant personality disorders to assess on the PDI-IV. Widiger et al. (1995) suggest that the use of such self-report measures will serve the same purpose as screening questionnaires but also will provide more information than measures simply designed for screening purposes.

The PDI-IV manual indicates that, although lay interviewers can administer the PDI-IV, extensive training and supervision is required. Even then, it is recommended that final item scoring be done by an experienced clinician. Ideally, the PDI-IV should be administered and scored by a professional clinician with training and experience in diagnosis and assessment. The PDI-IV manual outlines suggested training beginning with study of the *DSM-IV*, articles on the diagnosis or assessment of personality disorders, and the PDI-IV manual and interview booklets. Following discussion of this literature it is recommended that trainees conduct pilot interviews with nonclinical subjects. Tapes of these initial interviews should be reviewed and feedback provided. It is then suggested that 5–10 patient interviews be conducted and taped for review. Continued taping and systematic review of interviews is recommended to avoid interviewer drift.

4.05.3.4.1 Reliability

Inter-rater agreement for presence vs. absence of personality disorders ranges from 0.37 (histrionic) to 0.81 (antisocial), with a median kappa of 0.65. Agreement for the number of personality disorder criteria met ranges from 0.74 (histrionic, narcissistic, and schizotypal) to 0.90 (obsessive-compulsive and passive-aggressive) and 0.91 (sadistic). Median reliability for the number of PD criteria met was 0.84. Although these data are generally encouraging there are some concerns. Unfortunately, the population on which these reliability data were obtained is not specified nor are the methods for determining rater agreement described. More detailed information may be available from the unpublished dissertation from which these data are derived (Corbitt, 1994).

4.05.3.4.2 Summary

The PDI-IV is built upon the extensive history and experience derived from prior editions of this interview. The PDI-IV manual is one of the more extensive, thorough, and informative manuals available for the assessment of personality disorders. The flexibility afforded by the choice of either thematic content format or diagnostic category format is also attractive. Despite the accumulation of research on prior versions of the PDI-IV, there is limited reliability data for the PDI-IV. However, the PDI-IV is the only personality interview that has reliability data available for the *DSM-IV* personality disorders.

4.05.4 CHILD AND ADOLESCENT DISORDERS

4.05.4.1 Schedule for Affective Disorders and Schizophrenia for School Age Children

The Schedule for Affective Disorders and Schizophrenia for School Age Children (K-SADS; Puig-Antich & Chambers, 1978) is a semistructured interview designed for research or clinical assessment by a trained clinician. The K-SADS was developed as a child and adolescent version of SADS resulting from research in childhood depression. The K-SADS covers a wide range of childhood disorders but has a strong emphasis on major affective disorders (Roberts, Vargo, & Ferguson, 1989). The interview is intended to assess both past and current episodes of psychopathology in children aged 6–17 years old.

The K-SADS-III-R is compatible with *DSM-III-R* criteria. This version of the SADS provides 31 diagnoses within affective disorders (including depression, bipolar disorder, dysthymia, and cyclothymia), eating disorders, anxiety disorders, behavioral disorders (e.g., conduct disorder, substance abuse/dependence), psychoses, and personality disorders (i.e., schizotypal and paranoid)

The K-SADS is composed of three parts. It begins with an unstructured interview that aims to put the patient at ease and gather information regarding present problems, onset and duration of problems, and treatment history. Following this general interview, the interviewer asks questions relevant to specific symptoms and diagnostic criteria. Sample questions are provided only as a guideline, and modification is encouraged. If initial probe questions are negative, follow-up questions are skipped over. At the conclusion of the interview, observational items are rated (Roberts et al., 1989).

The parent interview should be conducted first, followed by the child interview. The child and parent interview require approximately 90 minutes each. The K-SADS focuses on the last week and most intense symptoms over the last 12 months. Each time period is rated independently and a summary score is made. Diagnostic criteria are rated as present or absent, and then rated on severity (Ambrosini, Metz, Prabucki, & Lee, J., 1989). Ultimately, diagnoses are given based on clinical judgment (Hodges, McKnew, Burbach, & Roebuck, 1987).

As with the SADS, the K-SADS requires extensive training and experience in psychiatric interviewing but has an added burden of conducting interviews with adults (parent/guardian) and children. Full familiarity with *DSM-III-R* is required. Training typically requires viewing videotapes and the conduct of practice interviews with ongoing supervision.

4.05.4.1.1 Reliability

Chambers et al. (1985) examined test–retest reliability of the K-SADS administered to children and parents. Test–retest reliability of major diagnoses was generally adequate with kappas ranging from 0.54 to 0.74, with the exception of anxiety disorder ($K = 0.24$). Individual symptoms and symptom scales generally showed adequate test–retest reliability with anxiety-related symptoms showing the lowest reliability. Agreement between the parent and child interviews varied greatly, ranging from poor to excellent. This later finding suggests the nonredundant aspect of these two interviews.

Inter-rater reliability was examined in videotaped K-SADS-III-R interviews by Ambrosini, et al. (1989). Inter-rater agreement among child, parent, combined interview ratings, and across time frames (present episode and last week) ranged from acceptable ($K = 0.53$) to excellent ($K = 1$) for major depression, minor depression, overanxious disorder, simple phobia, separation anxiety, oppositional, and attention deficit. Of the 36 kappa values, 30 were 0.75 or higher.

Apter, Orvaschel, Laseg, Moses, and Tyano (1989) examined inter-rater and test–retest agreement in a sample of adolescent inpatients (aged 11 to 18 years). Overall inter-rater and test–retest agreement was high with kappas of 0.78. Reliability of symptom scales was also adequate with ICCs of 0.63–0.97 for inter-rater agreement and ICCs of 0.55–0.96 for test–retest agreement. Diagnostic agreement between parent and child interviews (conducted by different clinicians for each informant) was generally low with an overall kappa of 0.42. Parent–child agreement for symptom scales was particularly low for anxiety symptoms.

4.05.4.1.2 Summary

The K-SADS extensively covers the major affective disorders and has adequate coverage of other childhood disorders. It has been one of the main diagnostic interviews available for use with children and adolescents. Reliability data are very positive for a number of disorders. However, reliability data are largely for *DSM-III* diagnoses and limited data are available for *DSM-III-R* diagnoses (no data are available for *DSM-IV*).

4.05.4.2 Child Assessment Schedule

The Child Assessment Schedule (CAS; Hodges, Kline, Stern, Cytryn, & McKnew, 1982) is a structured interview that is unique in that it is modeled after traditional clinical interviews with children. The interview is organized around thematic topics (e.g., school, friends) with diagnostic items inserted within these topics. Structured questions are provided in a format that is intended to develop rapport. Hodges (1993) has noted that about half of the CAS material related to clinical content does not reflect directly on diagnostic criteria.

The CAS is organized into three parts. In the first part 75 questions are organized into 11 content areas: school, friends, activities and hobbies, family, fears, worries, self-image, mood, somatic concerns, expression of anger, and thought disorder symptoms. Items are rated true (presence of symptom), false (absence of symptom), ambiguous, no response, or not applicable. In the second part the onset and duration of symptoms is assessed. In the third part of the CAS, following completion of the interview, 56 items are rated based on observations during the interview. These items include the following areas: insight, grooming, motor coordination, activity level, other spontaneous physical behaviors, estimate of cognitive ability, quality of verbal communications, quality of emotional expression, and impressions about the quality of interpersonal interactions. A parallel form of the CAS is available for administration to parents (P-CAS). The same inquiries are made, with parents being asked about the child.

Quantitative scales can be obtained for a total score, scores for content areas and for symptom complexes. The internal consistency of the scale scores has been examined and are generally adequate with a few exceptions. Symptom scales have been found to be internally consistent (Hodges, Saunders, Kashani, Hamlett, & Thompson, 1990), especially in a psychiatric sample with some attenuation in medically ill and community samples (particularly for anxiety symptoms). Hodges and Saunders (1989) examined the internal consistency of content scales for both the CAS and P-CAS. For the CAS, content scales generally had alphas greater than 0.70 but low internal consistency was found for Activities and Reality Functioning. For the P-CAS, content scales with alphas below 0.60 were Activities, Reality-testing Symptoms, Self-image, and Fears. Diagnoses for *DSM-III-R* can be derived in addition to these scale scores. The CAS takes approximately 45 minutes to one hour to complete.

It is recommended that the CAS be administered by trained clinicians (although lay interviewers have been used; Hodges et al., 1987). Guidelines for administering, scoring, and interpreting the CAS are contained in the CAS manual (Hodges, 1985) and in guidelines established for achieving rater reliability (Hodges, 1983).

4.05.4.2.1 Reliability

In the initial rater reliability study, Hodges, McKnew, Cytryn, Stern and Kline (1982) examined inter-rater agreement using videotaped interviews. For symptom scores, mean kappas were generally close to or exceeded 0.60. Mean correlations across raters for content areas was 0.63 or above with the exception of worries (0.59). For symptom complexes mean correlations were 0.68 or above except for attention deficit without hyperactivity (0.58), separation anxiety (0.56), and socialized conduct (0.44). Hodges et al. (1982) also report inter-rater reliability on a small sample ($N = 10$). Correlations for items, content scores, and symptom complex scores were all above 0.85. Verhust, Althaus, and Berden (1987) also have reported inter-rater reliability for contents scores using a small number ($N = 10$) of videotaped interviews. Correlations for content areas ranged from 0.70 to 0.97.

In the only test–retest reliability study of the CAS, Hodges, Cools, and McKnew (1989) examined agreement over a mean of five days with an inpatient sample. Intraclass correlations indicated good reliability for the total CAS score and scale scores. Kappas for *DSM-III* diagnoses of conduct disorder, depression, and anxiety were above 0.70. However, the kappa for attention deficit disorder was only 0.43.

The concordance between the CAS and the K-SADS was examined by Hodges et al. (1987). Lay interviewers were used and agreement was examined for both child and parent interviews for four major diagnostic categories (attention deficit disorder, conduct disorders, anxiety disorders, and affective disorders). Only present episodes were evaluated. Child only diagnostic concordance between the CAS and K-SADS was poor for attention deficit disorder and anxiety disorders (kappas less than 0.40). Better agreement was obtained for parent only interviews or in combinations of child and parent interviews. Anxiety disorders generally had low concordance across informant conditions.

The concordance between child and parent interviews has also been examined with the CAS. Verhulst et al. (1987) found low to moderate correlations between parent- and child-derived content areas, somatic concerns, and observational judgments. Of 22 correlations, only four exceeded 0.40. The total score correlation was 0.58, indicating approximately 34% shared variance. Hodges, Gordon, and Lennon (1990) also found low to moderate correlations between parent and child interview ratings. For diagnostic areas, the lowest correlations (those below 0.30) were obtained for overanxious disorder, oppositional disorder, and separation anxiety. Low correlations (again below 0.30) were also found in the content areas of fears, worries and anxieties, and physical complaints. These data indicate reasonable parent–child agreement for conduct/behavioral problems, moderate agreement for affective symptoms, and low agreement for anxiety symptoms. As with other child assessment measures, the greatest parent–child agreement appears to be for observable behavior and the lowest for subjective experiences (Hodges et al., 1990; Hodges, 1993).

4.05.4.2.2 Summary

The CAS appears to provide a reliable assessment of a range of symptoms and shows reasonable convergence with noninterview measures. It does not cover a number of disorders including sleep disorders, eating disorders, alcohol or drug use disorders, or mania. Although it provides a broad clinical assessment, some users may find that the presence of many CAS items that do not reflect directly on diagnostic criteria is inefficient. Inter-rater agreement for diagnoses studied appears adequate. However, only one small-scale study has examined test–retest reliability for a subset of diagnoses. No reliability data are available for *DSM-IV* diagnoses.

4.05.4.3 Child and Adolescent Psychiatric Assessment

The Child and Adolescent Psychiatric Assessment (CAPA; Angold et al., 1995) was developed in order to assess a wide range of diagnostic classifications including *DSM-III*, *DSM-III-R*, *ICD-9*, and *ICD-10*. Additionally, other symptoms of clinical interest are evaluated. As with other interviews, the CAPA can be administered to children and parents. The CAPA has four sections, three of which pertain to the interview proper. The time period addressed is the three months preceding the interview. In the Introduction, the interview is conducted in a conversational manner in order to establish rapport. Questions within the Introduction address three areas: home and family life, school life, peer groups and spare time activities.

The second section is the Symptom Review which has a disorder-based organization. A wide range of disorders are covered including anxiety disorders, obsessive-compulsive disorders, depressive disorders, manic disorders, somatization disorders, food-related disorders, sleep disorders, elimination disorders, tic disorders and trichotillomania, disruptive behavior disorders, tobacco use, alcohol, psychotic disorders, life events and post-traumatic stress disorder (PTSD), and drugs. Due to problems in child report with some disorders, only the parent interview assesses sleep terror disorder, sleepwalking disorder, and attention deficit hyperactivity disorder. Alternatively, because parents may be a poor source of information for children's substance use, delusions, hallucinations, and thought disorder these items are abbreviated in the parent interview with more extensive coverage in the child interview.

The third section of the interview assesses incapacity. At this point the interviewer reviews symptom information and questions about the effects of symptoms in 17 areas of psychosocial impairment. Impairment is evaluated in the three domains of home and family life, school life, peer groups and spare time activities. Finally, following the interview, observations of interview behavior are rated for 67 items. These items cover level of activity, child's mood state, quality of child's social interaction during interview, and psychotic behavior.

Detailed questions are provided for each interview item in the CAPA. There are three levels of questions. Screening questions allow a skip-out of a section. If the screen question is positive two levels of probes are provided. Emphasized probes are required and should be asked for all subjects. Discretionary probes are provided if further information is required. A glossary is provided to be used in conjunction with the standardized questions. The glossary provides operational definitions of symptom items. These definitions were based on a review of several of the existing clinical child interviews. The glossary also provides explicit rating principles including a formal definition of the item, ratings of intensity (from 0, absent, to 3, present at a higher intensity level), duration, length of time symptom is occurring, and psychosocial impairment related to the symptom.

A wealth of information is obtained with the CAPA. Fortunately, a computer program is available in order to summarize these data with a series of diagnostic algorithms (the CAPA Originated Diagnostic Algorithms; CODA). The CODA can generate diagnoses according to *DSM-III*, *DSM-III-R*, *DSM-IV*, and *ICD-10* systems as well as symptom scores for particular diagnostic areas.

Angold et al. (1995) report that the CAPA has been used with a variety of populations (both clinical and general population) in both the UK and the USA. Training requires four to five weeks with emphasis on practice administering the CAPA and group ratings of tapes. Based on its use in multiple clinical centers explicit training criteria have been developed and details about the CAPA and its training requirements can be obtained from Angold.

4.05.4.3.1 Reliability

Angold and Costello (1995) examined test–retest reliability in a clinical sample. Interviews were conducted with children only and were completed within 11 days of each other. Kappas for specific *DSM-III-R* diagnoses were all above 0.73, with the exception of conduct disorder ($K = 0.55$). Agreement on symptom scales for these disorders was also high with ICCs above 0.60 except for conduct disorder (ICC = 0.50). No reliability data were available for a number of disorders covered in the CAPA including obsessive-compulsive disorders, manic disorders, food-related disorders, sleep disorders, elimination disorders, tic disorders, psychotic disorders, or life events and PTSD.

4.05.4.3.2 Summary

The CAPA appears to offer a thorough clinical evaluation that incorporates several diagnostic criteria. It provides a broader assessment with more contemporary diagnostic nosology than some other instruments. However, this breadth of assessment does come at a cost. The CAPA administered to the child alone can take one hour and coding can take another 45 minutes. The CAPA is not recommended for use with children under the age of eight. Additionally, the CAPA is limited to the three months preceding the interview. Although reliability data are encouraging, these data are limited to child only interviews, are not available for a number of disorders covered by the CAPA, and are not available for *DSM-IV* diagnoses. It will be important to determine the reliability of other disorders as well as that of parent interviews, and whether diagnostic agreement is improved with both child and parent administration.

4.05.4.4 Diagnostic Interview Schedule for Children

The Diagnostic Interview Schedule for Children (DISC) is a highly structured interview designed to assess most child and adolescent psychiatric disorders (Jensen et al., 1995) The interview was introduced in 1982 as a child version of the Diagnostic Interview Schedule (DIS). The DISC was intended to be administered by lay interviewers and used for epidemiological research (Shaffer et al., 1993). The version current in the late 1990s, DISC-2.1 covers 35 diagnostic criteria for the *DSM-III-R*, and contains a child (DISC-C) and parent (DISC-P) interview. The DISC was designed for children and adolescents ranging from 6 to 18 years old.

DISC interviewers are encouraged not to deviate from the order, wording, and scoring procedures. The child and parent interviews of the DISC-2.1 require approximately 60–75 minutes each (Jensen et al., 1995). Questions, organized into six separate diagnostic modules, inquire about current and past symptoms, behaviors, and emotions of most child and adolescent diagnoses. Diagnostic criteria are initially assessed with a broad "stem question" (with a low diagnostic threshold) and, if endorsed, followed with "contingent questions" to determine criteria requirements, duration, frequency, impairment, and other modifiers (Fisher et al., 1993). The DISC-2.1 focuses on the last six months and a graphic timeline is used to assist in recall (Fisher et al., 1993; Jensen et al., 1995). At the end of each module, supplementary questions are provided to assess onset, current impairment, treatment history, and precipitating stressors. Questions are rated as: "no," "yes," or "sometimes" or "somewhat," and a computer algorithm generates diagnoses.

The DISC was specifically developed for use by lay interviewers in epidemiological research. Interviewer training typically takes one to two weeks. No differences in performance have been found between clinicians and lay interviewers using the DISC-1 (Shaffer et al., 1993). A user's manual for the DISC is available.

4.05.4.4.1 Reliability

Jensen et al. (1995) examined test–retest reliability in both a clinical and community sample across three sites. In the clinic sample, for major diagnostic categories, test–retest agreement was adequate for parents (*K* range = 0.58–0.70) and was generally higher than that obtained for child interviews (*K* range = 0.39–0.86). Using a combined diagnostic algorithm, test–retest agreement was adequate (*K* range = 0.50–0.71). Inter-rater agreement was lower for the community sample with test–retest agreement lower for parents (*K* range = 0.66) and children (*K* range = 0.23–0.60). The combined diagnostic algorithm for the community sample continued to provide low agreement (*K* range = 0.26–0.64). Instances of diagnostic disagreement in the clinic sample appeared to be related to an absolute decrease in the number of symptoms at the time of the second interview. Low reliability in the community sample was attributed to decreased symptom severity, the presence of threshold cases, and other unknown factors.

Other studies have generally found adequate test–retest reliability for the DISC. One general pattern that has emerged is greater agreement when examining parent interviews. Schwab-Stone et al. (1993) interviewed 41 adolescents (aged 11–17 years) and 39 parents twice (1–3 weeks apart) with the DISC-R. Inter-rater agreement ranged from poor (*K* = 0.16) to good (*K* = 0.77) for the child interviews, and ranged from fair (*K* = 0.55) to excellent (*K* = 0.88) for the parent interviews. Schwab-Stone, Fallon, Briggs, and Crowther (1994) interviewed 109 preadolescents (aged 6–11 years) and their parents twice (7–18 days apart) with the DISC-R. Inter-rater agreement ranged from poor (*K* = 0) to fair (*K* = 0.56) for the child interviews, and from poor (*K* = 0.43) to excellent (*K* = 0.81) for the parent interviews. Based on the lower inter-rater agreement for preadolescents, Schwab-Stone et al. (1994) concluded that highly structured interviews were not appropriate for directly assessing young children due to lower endorsement of symptoms and unreliable reporting within the interview.

Most of the reliability studies on the DISC have examined only the most common childhood diagnoses and little information is available on uncommon diagnoses. From a clinical sample of relatively uncommon diagnoses, Fisher et al. (1993) interviewed 93 children (aged 8–19 years) and 75 parents with the DISC-2.1. Using the clinic diagnosis as a standard, the DISC-2.1 had good (0.73) to excellent (1) sensitivity in identifying eating disorders, major depressive disorders, obsessive-compulsive disorder, psychosis, substance use disorders, and tic disorders. The DISC-2.1 was noted to ;be less sensitive for major depressive disorder than other interviews (K-SADS, DICA, CAS).

4.05.4.4.2 Summary

The DISC's design for epidemiological research constitutes several advantages. First, the highly structured interview may be administered by nonclinicians. Second, the DISC contains the full range of disorders. Finally, the DISC has been thought to contain a lower threshold for disorders than other interviews, which makes it ideal for screening and use in the general population (Roberts et al., 1989). However, the design of the DISC has several disadvantages. It may be too restrictive, at times not allowing the interviewer to probe further and adapt the interview to accommodate special situations. The DISC has shown to be unreliable among young children and is fairly long requiring 60–75 minutes each for the two interviews. Research on the concordance between the DISC and clinical structured interviews such as the K-SADS has not been examined.

4.05.4.5 Diagnostic Interview for Children and Adolescents

The Diagnostic Interview for Children and Adolescents (DICA) is a highly structured interview designed to be used by lay interviewers for clinical and epidemiological research. The interview assesses the present episode of a wide range of psychopathology among children aged 6–17 years (Roberts et al., 1989). The interview initially appeared in 1969, and was revised in 1981 to emulate the organization of the DIS and based on *DSM-III* criteria (Welner, Reich, Herjanic, Jung, & Amado, 1987). The DICA was subsequently revised to conform to *DSM-III-R* diagnoses (DICA-R; Kaplan & Reich, 1991). In addition to coverage of *DSM-III-R*, the DICA-R was also revised so that questions were presented in a more conversational style.

The DICA-R is organized into 15 sections. Two sections cover general and sociodemographic information, 11 sections relate to disorders and symptomatology, and remaining sections address menstruation, psychosocial stressors, and clinical observations. The DICA consists of a separate parent (DICA-P) and child (DICA-C) interview. The child interview requires 30–40 minutes to administer, while the parent interview takes longer due to additional questions on developmental history, medical history, socioeconomic status, and family history (Roberts et al., 1989). For each diagnostic criterion, one or more questions elicit information. Follow-up questions are skipped if primary questions are responded negatively. Responses on the DICA-R are coded on a four-point scale: "no," "rarely," "sometimes" or "somewhat," and "yes." Following each diagnostic section, specific *DSM* criteria are listed to assist in deriving diagnoses (Welner et al., 1987).

4.05.4.5.1 Reliability

Limited data on the inter-rater agreement of the DICA is available. Only one study has provided data pertaining to individual diagnoses with an adequate description of sample and methods. Welner et al. (1987) administered two independent interviews (1–7 days apart) to 27 psychiatric inpatients (aged 7–17 years). Using lay interviewers, inter-rater agreement was excellent across diagnostic categories (K range = 0.76–1, median = 0.86).

Similar to other interviews, diagnoses derived from the parent and child interviews vary. Welner et al. (1987) examined concordance between child and parent interviews among 84 outpatients (ages 7–17 years). Fair to excellent concordance was noted (K range = 0.49–0.80, median = 0.63). However, other studies have found more modest concordance between parent and child interviews with median kappas below 0.30 (Earls, Reich, Jung, & Cloninger, 1988; Sylvester, Hyde, & Reichler, 1987).

4.05.4.5.2 Summary

The DICA-R appears to be a well-developed instrument that has taken special care in the writing and sequencing of questions. Although the DICA has been used in a number of studies, only a limited amount of reliability information is available. No reliability information is available for the DICA-R. Other child and adolescent interviews may be more attractive because of the relatively greater inter-rater reliability information.

4.05.5 SUMMARY

There has been an enormous amount of research conducted on the development and use of structured clinical interviews since the late 1960s. This research has yielded diagnostic interviews that address an array of clinical diagnoses in both adult and child populations. The use of structured interviews can not only provide reliable diagnostic evaluations but can also serve to ensure a broad and thorough clinical assessment. Although most readily embraced in research settings, it is anticipated (and hoped) that structured diagnostic interviews will become more commonplace in clinical applied settings as part of the standard assessment tools that clinicians use regularly.

4.05.6 REFERENCES

Ambrosini, P. J., Metz, C., Prabucki, K., & Lee, J. (1989). Videotape reliability of the third revised edition of the K-SADS. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 723–728.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

Andreasen, N. C. (1983). *The Scale for the Assessment of Negative Symptoms (SANS).* Iowa City, IA: The University of Iowa.

Andreasen, N. C. (1984). *The scale for the assessment of positive symptoms (SAPS).* Iowa City, IA: The University of Iowa.

Andreasen, N. C. (1987). *Comprehensive assessment of symptoms and history.* Iowa City, IA: The University of Iowa.

Andreasen, N. C., Flaum, M., & Arndt, S. (1992). The comprehensive assessment of symptoms and history (CASH): An instrument for assessing diagnosis and psychopathology. *Archives of General Psychiatry, 49,* 615–623.

Andreasen, N. C., Grove, W. M., Shapiro, R. W., Keller, M. B., Hirschfeld, R. M. A., & McDonald-Scott, P. (1981). Reliability of lifetime diagnoses: A multicenter collaborative perspective. *Archives of General Psychiatry, 38,* 400–405.

Angold, A., & Costello, E. J., (1995). A test–retest study of child-reported psychiatric symptoms and diagnoses using the Child and Adolescent Psychiatric Assessment (CAPA-C). *Psychological Medicine, 25,* 755–762.

Angold, A., Prendergast, M., Cox, A., Harrington, R., Simonoff, E., & Rutter, M. (1995). The Child and Adolescent Psychiatric Assessment (CAPA). *Psychological Medicine, 25,* 739–753.

Anthony, J. C., Folstein, M., Romanoski, A. J., Von Korff, M. R., Nestadt, G. R., Chahal, R., Merchant, A., Brown, H., Shapiro, S., Kramer, M., & Gruenberg, E. M. (1985). Comparison of the lay Diagnostic Interview Schedule and a standardized psychiatric diagnosis: Experience in eastern Baltimore. *Archives of General Psychiatry, 42,* 667–675.

Apter, A., Orvaschel, H., Laseg, M., Moses, T., & Tyano, S. (1989). Psychometric properties of the K-SADS-P in an Israeli adolescent inpatient population. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 61–65.

Arntz, A., van Beijsterveldt, B., Hoekstra, R., Hofman, A., Eussen, M., & Sallaerts, S. (1992). The inter-rater reliability of a Dutch version of the Structured Clinical Interview for DSM-III-R personality disorders. *Acta Psychiatrica Scandinavica, 85,* 394–400.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. E., & Erbaugh, J. K. (1962). Reliability of psychiatric diagnoses: 2. A study of consistency of clinical judgments and ratings. *American Journal of Psychiatry, 119,* 351–357.

Bromet, E. J., Dunn, L. O., Connell, M. M., Dew, M. A., & Schulberg, H. C. (1986). Long-term reliability of diagnosing lifetime major depression in a community sample. *Archives of General Psychiatry, 43,* 435–440.

Brooks, R. B., Baltazar, P. L., McDowell, D. E., Munjack, D. J., & Bruns, J. R. (1991). Personality disorders co-occurring with panic disorder with agoraphobia. *Journal of Personality Disorders, 5,* 328–336.

Chambers, W. J., Puig-Antich, J., Hirsch, M., Paez, P., Ambrosini, P. J., Tabrizi, M. A., & Davies, M. (1985). The assessment of affective disorders in children and adolescents by semistructured interview. *Archives of General Psychiatry, 42,* 696–702.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37–46.

Cooper, J. E., Copeland, J. R. M., Brown, G. W., Harris, T., & Gourlay, A. J. (1977). Further studies on interviewer training and inter-rater reliability of the Present State Exam (PSE). *Psychological Medicine, 7,* 517–523.

Cooper, J. E., Kendell, R. E., Gurland, B. J., Sharpe, L., Copeland, J. R. M., & Simon, R. (1972). *Psychiatric diagnosis in New York and London. Maudsley monographs.* London: Oxford University Press.

Corbitt, E. M. (1994). *Sex bias and the personality disorders: A reinterpretation from the five-factor model.* Unpublished doctoral dissertation, University of Kentucky, Lexington.

Cottler, L. B. (1991). The CIDI and CIDI-Substance Abuse Module (SAM): Cross-cultural instruments for assessing DSM-III, DSM-III-R and ICD-10 criteria. *Research Monographs, 105,* 220–226.

Earls, R., Reich, W., Jung, K. G., & Cloninger, C. R. (1988). Psychopathology in children of alcoholic and antisocial parents. *Alcoholism: Clinical and Experimental Research, 12,* 481–487.

Endicott, J., & Spitzer, R. L. (1978). A diagnostic interview: The Schedule for Affective Disorders and Schizophrenia. *Archives of General Psychiatry, 35,* 837–844.

Endicott, J., Spitzer, R. L., Fleiss, J. L., & Cohen, J. (1976). The Global Assessment Scale: A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry, 33,* 766–771.

Erdman, H. P., Klein, M. H., Greist, J. H., Bass, S. M., Bires, J. K., & Machtinger, P. E. (1987). A comparison of the Diagnostic Interview Schedule and clinical diagnosis. *American Journal of Psychiatry, 144,* 1477–1480.

Escobar, J. I., Randolph, E. T., Asamen, J., & Karno, M. (1986). The NIMH-DIS in the assessment of DSM-III schizophrenic disorder. *Schizophrenia Bulletin, 12,* 187–194.

Faraone, S. V., Blehar, M., Pepple, J., Moldin, S. O., Norton, J., Nurnberger, J. I., Malaspina, D., Kaufman, C. A., Reich, T., Cloning, C. R., DePaulo, J. R., Berg, K., Gershon, E. S., Kirch, D. G., & Tsuang, M. T. (1996). Diagnostic accuracy and confusability analyses: An application to the Diagnostic Interview for Genetic Studies. *Psychological Medicine, 26,* 401–410.

Farmer, A. E., Katz, R., McGuffin, P., & Bebbington, P. (1987). A comparison between the Present State Examination and the Composite International Interview. *Archives of General Psychiatry, 44,* 1064–1068.

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G., & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry, 26,* 57–63.

First, M. B., Gibbon, M., Spitzer, R. L., & Williams, J. B. W. (1996). *User's guide for the Structured Clinical Interview for DSM-IV Axis I Disorders-Research Version (SCID-I, version 2.0, February 1996 Final version).* New York: Biometrics Research Department, New York State Psychiatric Institute.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1995). The Structured Clinical Interview for DSM-III-R Personality Disorders (SCID-II). Part I: Description. *Journal of Personality Disorders, 9,* 83–91.

First, M. B., Spitzer, R. L., Gibbon, M., Williams, J. B. W., Davies, M., Borus, J., Howes, M. J., Kane, J., Pope, H. G., & Rounsaville, B. (1995). The Structured Clinical Interview for DSM-III-R Personality Disorders (SCID-II). Part II: Multi-site test–retest reliability study. *Journal of Personality Disorders, 9,* 92–104.

Fisher, P. W., Shaffer, D., Piacentini, J. C., Lapkin, J., Kafantaris, V., Leonard, H., & Herzog, D. B. (1993). Sensitivity of the Diagnostic Interview Schedule for Children, 2nd Edition (DISC-2.1) for specific diagnoses of children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 666–673.

Fogelson, D. L., Nuechertlein, K. H., Asarnow, R. F., Subotnik, K. L., & Talovic, S. A. (1991). Inter-rater reliability of the Structured Clinical Interview for DSM-III-R, Axis II: schizophrenia spectrum and affective spectrum disorders. *Psychiatry Research, 39,* 55–63.

Folstein, M. F., Folstein, S. E., & McHugh, P. (1975). "Mini Mental State": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Ford, J., Hillard, J. R., Giesler, L. J., Lassen, K. L., & Thomas, H. (1989). Substance abuse/mental illness: Diagnostic issues. *American Journal of Drug and Alcohol Abuse, 15,* 297–307.

Fyer, A. J., Mannuzza, S., Martin, L. Y., Gallops, M. S., Endicott, J., Schleyer, B., Gorman, J. M., Liebowitz, M. R., & Klein, D. F. (1989). Reliability of anxiety assessment, II: Symptom assessment. *Archives of General Psychiatry, 46,* 1102–1110.

Grove, W. M., Andreasen, N. C., McDonald-Scott, P., Keller, M. B., & Shapiro, R. W. (1981). Reliability studies of psychiatric diagnosis: Theory and practice. *Archives of General Psychiatry, 38,* 408–413.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry, 23,* 56–62.

Helzer, J. E., Robins, L. N., McEvoy, L. T., Spitznagel, E. L., Stolzman, R. K., Farmer, A., & Brockington, I. F. (1985). A comparison of clinical and diagnostic interview schedule diagnoses: Physician reexamination of lay-interviewed cases in the general population. *Archives of General Psychiatry, 42,* 657–666.

Hodges, K. (1983). *Guidelines to aid in establishing inter-rater reliability with the Child Assessment Schedule.* Unpublished manuscript.

Hodges, K. (1985). *Manual for the Child Assessment Schedule.* Unpublished manuscript.

Hodges, K. (1993). Structured interviews for assessing children. *Journal of Child Psychology and Psychiatry, 34,* 49–68.

Hodges, K., Cools, J., & McKnew, D. (1989). Test–retest reliability of a clinical research interview for children: The Child Assessment Schedule (CAS). *Psychological Assessment: Journal of Consulting and Clinical Psychology, 1,* 317–322.

Hodges, K., Gordon, Y., & Lennon, M. P. (1990). Parent–child agreement on symptoms assessed via a clinical research interview for children: The Child Assessment Schedule (CAS). *Journal of Child Psychology and Psychiatry, 31,* 427–436.

Hodges, K., Kline, J., Stern, L., Cytryn, L., & McKnew, D. (1982). The development of a child assessment interview for research and clinical use. *Journal of Abnormal Child Psychology, 10,* 173–189.

Hodges, K., McKnew, D., Burbach, D. J., & Roebuck, L. (1987). Diagnostic concordance between the Child Assessment Schedule (CAS) and the Schedule for Affective Disorders and Schizophrenia for School-age Children (K-SADS) in an outpatient sample using lay interviewers. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 654–661.

Hodges, K., McKnew, D., Cytryn, L., Stern, L., & Kline, J. (1982). The Child Assessment Schedule (CAS) diagnostic interview: A report on reliability and validity. *Journal of the American Academy of Child Psychiatry, 21,* 468–473.

Hodges, K., & Saunders, W. (1989). Internal consistency of a diagnostic interview for children: The Child Assessment Schedule. *Journal of Abnormal Child Psychology, 17,* 691–701.

Hodges, K., Saunders, W. B., Kashani, J., Hamlett, K., & Thompson, R. J. (1990). *Journal of the American*

Academy of Child and Adolescent Psychiatry, 29, 635–641.

Hyler, S. E., Skodol, A. E., Kellman, H. D., Oldham, J. M., & Rosnik, L. (1990). Validity of the Personality Diagnostic Questionnaire-Revised: Comparison with two structured interviews. *American Journal of Psychiatry, 147,* 1043–1048.

Hyler, S. E., Skodol, A. E., Oldham, J. M., Kellman, D. H., & Doldge, N. (1992). Validity of the Personality Diagnostic Questionnaire-Revised: A replication in an outpatient sample. *Comprehensive Psychiatry, 33,* 73–77.

Jackson, H. J., Gazis, J., Rudd, R. P., & Edwards, J. (1991). Concordance between two personality disorder instruments with psychiatric inpatients. *Comprehensive Psychiatry, 32,* 252–260.

Janca, A., Robins, L. N., Cottler, L. B., & Early, T. S. (1992). Clinical observation of assessment using the Composite International Diagnostic Interview (CIDI): An analysis of the CIDI field trials—Wave II at the St Louis Site. *British Journal of Psychiatry, 160,* 815–818.

Jensen, P., Roper, M., Fisher, P., Piacentini, J., Canino, G., Richters, J., Rubio-Stipec, M., Dulcan, M., Goodman, S., Davies, M., Rae, D., Shaffer, D., Bird, H., Lahey, B., & Schwab-Stone, M. (1995). Test–retest reliability of the Diagnostic Interview Schedule for Children (DISC 2.1). *Archives of General Psychiatry, 52,* 61–71.

Kaplan, L. M., & Reich, W. (1991). *Manual for Diagnostic Interview for Children and Adolescents-Revised (DICA-R).* St Louis, MO: Washington University.

Keller, M. B., Lavori, P. W., McDonald-Scott, P., Scheftner, W. A., Andreasen, N. C., Shapiro, R. W., & Croughan, J. (1981). Reliability of lifetime diagnoses and symptoms in patients with a current psychiatric disorder. *Journal of Psychiatric Research, 16,* 229–240.

Kendell, R. E., Everitt, B., Cooper, J. E., Sartorius, N., & David, M. E. (1968). Reliability of the Present State Examination. *Social Psychiatry, 3,* 123–129.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159–174.

Loranger, A. W., Andreoli, A., Berger, P., Buchheim, P., Channabasavanna, S. M., Coid, B., Dahl, A., Diekstra, R. F. W., Ferguson, B., Jacobsberg, L. B., Janca, A., Mombour, W., Pull, C., Ono, Y., Regier, D. A., Sartorius, N., & Sumba R. O. (1995). *The International Personality Disorder Examination (IPDE) manual.* New York: World Health Organization.

Loranger, A. W., Lenzenweger, M. F., Gartner, A. F., Susman, V. L., Herzig, J., Zammit, G. K., Gartner, J. D., Abrams, R. C., & Young, R. C. (1991). Trait-state artifacts and the diagnosis of personality disorders. *Archives of General Psychiatry, 48,* 720–728.

Loranger, A. W., Sartorius, N., Andreoli, A., Berger, P., Buchheim, P., Channabasavanna, S. M., Coid, B., Dahl, A., Diekstra, R. F. W., Ferguson, B., Jacobsberg, L. B., Mombour, W., Pull, C., Ono, Y., & Regier, D. A. (1994). The international personality disorder examination. *Archives of General Psychiatry, 51,* 215–224.

Luria, R. E., & Berry, R. (1979). Reliability and descriptive validity of PSE syndromes. *Archives of General Psychiatry, 36,* 1187–1195.

Luria, R. E., & Berry, R. (1980). Teaching the Present State Examination in American. *American Journal of Psychiatry, 137,* 26–31.

Luria, R. E., & McHugh, P. R. (1974). Reliability and clinical utility of the "Wing" Present State Examination. *Archives of General Psychiatry, 30,* 866–971.

Malow, R. M., West, J. A., Williams, J. L., Sutker P. B. (1989). Personality disorders classification and symptoms in cocaine and opioid addicts. *Journal of Consulting and Clinical Psychology, 57,* 765–767.

Mannuzza, S., Fyer, A. J., Martin, L. Y., Gallops, M. S., Endicott, J., Gorman, J., Liebowitz, M. R., & Klein, D. F. (1989). Reliability of anxiety assessment, I: Diagnostic agreement. *Archives of General Psychiatry, 46,* 1093–1101.

Matarazzo, J. D. (1983). The reliability of psychiatric and psychological diagnosis. *Clinical Psychology Review, 3,* 103–145.

McDonald-Scott, P., & Endicott, J. (1984). Informed versus blind: The reliability of cross-sectional ratings of psychopathology. *Psychiatry Research, 12,* 207–217.

McGuffin, P., & Farmer, A. E. (1990). *Operational Criteria (OPCRIT) Checklist. Version 3.0.* Cardiff, UK: University of Wales.

McGuffin, P., Katz, R., & Aldrich, J. (1986). Past and Present State Examination: The assessment of "lifetime ever" psychopathology. *Psychological Medicine, 16,* 461–465.

Nurnberger, J. I., Blehar, M. C., Kaufman, C. A., York-Cooler, C., Simpson, S. G., Harkavy-Friedman, J., Severe, J. B., Malaspina, D., Reich, T., & collaborators from the NIMH Genetics Initiative (1994). Diagnostic Interview for Genetic Studies: Rationale, unique features, and training. *Archives of General Psychiatry, 51,* 849–859.

O'Boyle, M, & Self, D. (1990). A comparison of two interviews for DSM-III-R personality disorders. *Psychiatry Research, 32,* 85–92.

Okasha, A., Sadek, A., Al-Haddad, M. K., & Abdel-Mawgoud, M. (1993). Diagnostic agreement in psychiatry: A comparative study between ICD-9, ICD-10, and DSM-III-R. *British Journal of Psychiatry, 162,* 621–626.

Overall, J., & Gorham, D. (1962). Brief Psychiatric Rating Scale. *Psychological Reports, 10,* 799–812.

Page, A. C. (1991). An assessment of structured diagnostic interviews for adult anxiety disorders. *International Review of Psychiatry, 3,* 265–278.

Pfohl, B., Black, D. W., Noyes, R., Coryell, W. H., & Barrash, J. (1990). Axis I/Axis II comorbidity findings: Implications for validity. In J. Oldham (Ed.), *Axis II: New perspectives on validity* (pp. 147–161). Washington, DC: American Psychiatric Association.

Pfohl, B., Blum, N., & Zimmerman, M. (1995). *The Structured Interview for DSM-IV Personality Disorders (SIDP-IV).* Iowa City, IA: University of Iowa College of Medicine.

Pfohl, B., Blum, N., Zimmerman, M., & Stangl, D. (1989). *Structured Interview for DSM-III-R Personality Disorders (SIDP-R).* Iowa City, IA: University of Iowa College of Medicine.

Pilkonis, P. A., Heape, C. L., Proietti, J. M., Clark, S. W., McDavid, J. D., & Pitts, T. E. (1995). The reliability and validity of two structured diagnostic interviews for personality disorders. *Archives of General Psychiatry, 52,* 1025–1033.

Puig-Antich, J., & Chambers, W. J. (1978). *Schedule for Affective Disorders and Schizophrenia for School-age Children: Kiddie SADS (K-SADS).* New York: Department of Child and Adolescent Psychiatry, New York State Psychiatric Institute.

Renneberg, B., Chambless, D. L., & Gracely, E. J. (1992). Prevalence of SCID-diagnosed personality disorders in agoraphobic outpatients. *Journal of Anxiety Disorders, 6,* 111–118.

Rice, J. P., Rochberg, N., Endicott, J., Lavori, P. W., & Miller, C. (1992). Stability of psychiatric diagnoses: An application to the affective disorders. *Archives of General Psychiatry, 49,* 824–830.

Roberts, N., Vargo, B., & Ferguson, H. B. (1989). Measurement of anxiety and depression in children and adolescents. *Psychiatric Clinics of North America, 12,* 837–860.

Robins, L. N., Helzer, J. E., Croughan, J., Ratcliff, K. S. (1981). National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics, and validity. *Archives of General Psychiatry, 38,* 381–389.

Robins, L. N., Helzer, J. E., Ratcliff, K. S., & Seyfried, W. (1982). Validity of the diagnostic interview schedule, version II: DSM-III diagnoses. *Psychological Medicine, 12,* 855–870.

Robins, L. N., Wing, J., Wittchen, H.-U., Helzer, J. E., Babor, T. F., Burke, J., Farmer, A., Jablenski, A., Pickens, R., Regier, D. A., Sartorius, N., & Towle, L. H. (1988). The Composite International Diagnostic Interview: An epidemiological instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry, 45,* 1069–1077.

Rodgers, B., & Mann, S. (1986). The reliability and validity of PSE assessments by lay interviewers: A national population survey. *Psychological Medicine, 16,* 689–700.

Schwab-Stone, M., Fallon, T., Briggs, M., & Crowther, B. (1994). Reliability of diagnostic reporting for children aged 6–11 years: A test–retest study of the Diagnostic Interview Schedule for Children-Revised. *American Journal of Psychiatry, 151,* 1048–1054.

Schwab-Stone, M., Fisher, P., Piacentini, J., Shaffer, D., Davies, M., & Briggs, M. (1993). The Diagnostic Interview Schedule for Children-Revised version (DISC-R): II. Test–retest reliability. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 651–657.

Segal, D. L., Hersen, M., & Van Hasselt, V. B. (1994). Reliability of the structured clinical interview for DSM-III-R: An evaluative review. *Comprehensive Psychiatry, 35,* 316–327.

Shaffer, D., Schwab-Stone, M., Fisher, P., Cohen, P., Piacentini, J., Davies, M., Connors, C. K., & Regier, D. (1993). The Diagnostic Interview Schedule for Children-Revised version (DISC-R): I. Preparation, field testing, inter-rater reliability, and acceptability. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 643–650.

Shrout, P. E., Spitzer, R. L., & Fleiss, J. L. (1987). Quantification of agreement in psychiatric diagnosis revisited. *Archives of General Psychiatry, 44,* 172–177.

Spengler, P. A., & Wittchen, H. -U. (1988). Procedural validity of standardized symptom questions for the assessment of psychotic symptoms: A comparison of the DIS with two clinical methods. *Comprehensive Psychiatry, 29,* 309–322.

Spitzer, R. L. (1983). Psychiatric diagnosis: Are clinicians still necessary? *Comprehensive Psychiatry, 24,* 399–411.

Spitzer, R. L., Cohen, J., Fleiss, J. L., & Endicott, J. (1967). Quantification of agreement in psychiatric diagnosis: A new approach. *Archives of General Psychiatry, 17,* 83–87.

Spitzer, R. L., Endicott, J., & Robins, E. (1978). Research Diagnostic Criteria: Rationale and reliability. *Archives of General Psychiatry, 35,* 773–782.

Spitzer, R. L., & Fleiss, J. L. (1974). A re-analysis of the reliability of psychiatric diagnosis. *British Journal of Psychiatry, 125,* 341–347.

Spitzer, R. L., Fleiss, J. L., & Endicott, J. (1978). Problems of classification: Reliability and validity. In M. A. Lipton, A. DiMascio, & K. F. Killam (Eds.), *Psychopharmacology: A generation of progress* (pp. 857–869). New York: Raven Press.

Spitzer, R. L., Williams, J. B. W., Gibbon, M., & First, M. B. (1990). *User's guide for the Structured Clinical Interview for DSM-III-R (SCID).* Washington, DC: American Psychiatric Press.

Spitzer, R. L., Williams, J. B. W., Gibbon, M., & First, M. B. (1992). The Structured Clinical Interview for DSM-III-R (SCID). I: History, rationale, and description. *Archives of General Psychiatry, 49,* 624–629.

Standage, K. (1989). Structured interviews and the diagnosis of personality disorders. *Canadian Journal of Psychiatry, 34,* 906–912.

Stangl, D., Pfohl, B., Zimmerman, M., Bowers, W., & Corenthal, C. (1985). A structured interview for the DSM-III personality disorders. *Archives of General Psychiatry, 42,* 591–596.

Strober, M., Green, J., & Carlson, G. (1981). Reliability of psychiatric diagnosis in hospitalized adolescents: inter-rater agreement using DSM-III. *Archives of General Psychiatry, 38,* 141–145.

Sylvester, C., Hyde, T., & Reichler, R. (1987). The Diagnostic Interview for Children and Personality Interview for Children in studies of children at risk for anxiety disorders or depression. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 668–675.

Tress, K. H., Bellenis, C., Brownlow, J. M., Livingston, G., & Leff, J. P. (1987). The Present State Examination change rating scale. *British Journal of Psychiatry, 150,* 201–207.

Verhulst, F. C., Althaus, M., & Berden, G. F. M. G. (1987). The Child Assessment Schedule: Parent–child agreement and validity measures. *Journal of Child Psychology and Psychiatry, 28,* 455–466.

Ward, C. H., Beck, A. T., Mendelson, M., Mock, J. E., & Erbaugh, J. K. (1962). The psychiatric nomenclature: Reasons for diagnostic disagreement. *Archives of General Psychiatry, 7,* 198–205.

Welner, Z., Reich, W., Herjanic, B., Jung, K. G., & Amado, H. (1987). Reliability, validity, and parent–child agreement studies of the Diagnostic Interview for Children and Adolescents (DICA). *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 649–653.

Widiger, T. A., Mangine, S., Corbitt, E. M., Ellis, C. G., & Thomas, G. V. (1995). *Personality Disorder Interview-IV: A semistructured interview for the assessment of personality disorders.* Odessa, FL: Psychological Assessment Resources.

Williams, J. B. W., Gibbon, M., First, M. B., Spitzer, R. L., Davies, M., Borus, J., Howes, M. J., Kane, J., Pope, Jr., H. G., Rounsaville, B., & Wittchen, H.-U. (1992). The Structured Clinical Interview for DSM-III-R (SCID). II: Multisite test–retest reliability. *Archives of General Psychiatry, 49,* 630–636.

Wing, J. K. (1983). Use and misuse of the PSE. *British Journal of Psychiatry, 143,* 111–117.

Wing, J. K., Babor, T., Brugha, T., Burke, J., Cooper, J. E., Giel, R., Jablenski, A., Regier, D., & Sartorius, N. (1990). SCAN: Schedules for Clinical Assessment in Neuropsychiatry. *Archives of General Psychiatry, 47,* 589–593.

Wing, J. K., Birley, J. L. T., Cooper, J. E., Graham, P., & Isaacs, A. (1967). Reliability of a procedure for measuring and classifying present psychiatric state. *British Journal of Psychiatry, 113,* 499–515.

Wing, J. K., Cooper, J. E., & Sartorius, N. (1974). *The measurement and classification of psychiatric symptoms.* London: Cambridge University Press.

Wing, J. K., Nixon, J. M., Mann, S. A., & Leff, J. P. (1977). Reliability of the PSE (ninth edition) used in a population study. *Psychological Medicine, 7,* 505–516.

Wittchen, H.-U. (1994). Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): A critical review. *Journal of Psychiatry Research, 28,* 57–84.

Wittchen, H.-U., Robins, L. N., Cottler, L. B., Sartorius, N., Burke, J. D., Regier, D., & participants in the multicentre WHO/ADAMHA field trials (1991). Cross-cultural feasibility, reliability and sources of variance of the Composite International Diagnostic Interview (CIDI). *British Journal of Psychiatry, 159,* 645–653.

Wittchen, H. -U., Semler, G., & von Zerssen, D. (1985). A comparison of two diagnostic methods: Clinical ICD diagnoses versus DSM-III and Research Diagnostic Criteria using the Diagnostic Interview Schedule (Version 2). *Archives of General Psychiatry, 42,* 677–684.

World Health Organization (1973). *The international pilot study of schizophrenia, Vol. 1.*: Geneva: Author.

World Health Organization (1990). *Composite International Diagnostic Interview (CIDI): a) CIDI-interview (version 1.0), b) CIDI-user manual, c) CIDI-training manual, d) CIDI-computer programs.* Geneva: Author.

Zimmerman, M. (1994). Diagnosing personality disorders: A review of issues and research methods. *Archives of General Psychiatry, 51,* 225–245.

Zimmerman, M., Pfohl, B., Stangl, D., & Corenthal, C. (1986). Assessment of DSM-III personality disorders: The importance of interviewing an informant. *Journal of Clinical Psychiatry, 47,* 261–263.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.06
Principles and Practices of Behavioral Assessment with Children

THOMAS H. OLLENDICK
Virginia Tech, Blacksburg, VA, USA
and
ROSS W. GREENE
Harvard Medical School, Boston, MA, USA

4.06.1 INTRODUCTION 132
4.06.2 HISTORY AND DEVELOPMENT 133
4.06.3 THEORETICAL UNDERPINNINGS 134
4.06.4 DESCRIPTION OF ASSESSMENT PROCEDURES 136
  4.06.4.1 Behavioral Interviews 136
  4.06.4.2 Ratings and Checklists 138
  4.06.4.3 Self-report Instruments 140
  4.06.4.4 Self-monitoring 142
  4.06.4.5 Behavioral Observation 142
4.06.5 RESEARCH FINDINGS 144
  4.06.5.1 Behavioral Interviews 145
  4.06.5.2 Ratings and Checklists 145
  4.06.5.3 Self-report Instruments 146
  4.06.5.4 Self-monitoring 146
  4.06.5.5 Behavioral Observation 147
4.06.6 FUTURE DIRECTIONS 148
  4.06.6.1 Developmental Factors 148
  4.06.6.2 The Utility of the Multimethod Approach at Different Age Levels 149
  4.06.6.3 Cultural Sensitivity 149
  4.06.6.4 Measures of Cognitive and Affective Processes 150
  4.06.6.5 The Role of the Child 150
  4.06.6.6 Ethical Guidelines 151
4.06.7 SUMMARY 151
4.06.8 REFERENCES 151

4.06.1 INTRODUCTION

While treatment strategies derived from behavioral principles have a long and rich tradition in clinical child psychology (e.g., Holmes, 1936; Jones, 1924; Watson & Rayner, 1920), assessment practices based on these same principles have lagged, especially in the area of child behavioral assessment. In fact, many child behavioral assessment practices have been adopted, sometimes indiscriminately, from those used with adults. This practice is of dubious merit and, as we have argued elsewhere (Ollendick & Greene, 1990), it has frequently led to imprecise findings and questionable conclusions. As a result, greater attention has been focused on the development of behavioral assessment practices for children in recent years (e.g., Mash & Terdal, 1981, 1989; Ollendick & Hersen, 1984, 1993; Prinz, 1986).

As first suggested by Mash and Terdal (1981) and elaborated by Ollendick and Hersen (1984, 1993), child behavioral assessment can be defined as an ongoing, exploratory, hypothesis-testing process in which a range of specific procedures is used in order to understand a given child, group, or social ecology, and to formulate and evaluate specific intervention techniques. As such, child behavioral assessment is a dynamic, self-correcting process. It seeks to obtain information from a variety of sources in order that we might understand diverse child behavioral problems in their rich and varied contexts, and plan and evaluate behavioral interventions based upon the information obtained. Thus, assessment from this perspective is fluid (i.e., responsive to feedback and open to change(s) based on information obtained throughout the process), and it is linked intimately with treatment (i.e., assessment serves treatment). Moreover, child behavioral assessment entails more than the identification of discrete target behaviors and their controlling variables. While the importance of direct observation of target behaviors in simulated and natural settings should not be underestimated, recent advances in child behavioral assessment have incorporated a range of assessment procedures, including behavioral interviews, self-reports, ratings by significant others, and self-monitoring in addition to behavioral observations. An approach combining these procedures can best be described as a multimethod one in which an attempt is made to obtain a complete picture of the child and his or her presenting problems. Such a picture is intended to be useful in the understanding and modification of specific child behavior problems (Ollendick & Cerny, 1981; Ollendick & Hersen, 1984, 1993).

Two other primary features characterize child behavioral assessment procedures (Ollendick & Hersen, 1984, 1993). First, they must be sensitive to development, and second, they must be validated empirically. As noted by Lerner (1986, p. 41), the concept of development implies "systematic and successive changes over time in an organism." Descriptors such as "systematic" and "successive" suggest that these changes are, for the most part, orderly and that changes seen at one point in time will be influenced, at least in part, by changes that occurred at an earlier point in time. Thus development is not random nor, for that matter, discontinuous. Changes that occur at an early age (whether due to learning, an unfolding of basically predetermined structures, or some complex, interactive process) have a direct impact on subsequent development. Changes associated with development, however, create problems in selecting appropriate methods of assessment, as well as in identifying specific target behaviors for change (Ollendick & King, 1991). Behavioral interviews, self-reports, other-reports, self-monitoring, and behavioral observation may all be affected by these rapidly changing developmental processes. Further, due to "systematic and successive" change, some of these procedures may be more useful at one age than another. For example, interviews may be more difficult to conduct and self-reports less reliable with younger children, whereas self-monitoring and behavioral observations may be more reactive at older ages (Ollendick & Hersen, 1984). Age-related constraints are numerous and must be taken into consideration when selecting specific methods of assessment.

Just as child behavioral assessment procedures must be developmentally sensitive, they must also be validated empirically. All too frequently, professionals working with children have used assessment methods of convenience without sufficient regard for their psychometric properties, including their reliability, validity, and clinical utility (i.e., the degree to which assessment strategies contribute to beneficial treatment outcomes; see Hayes, Nelson, & Jarrett, 1987, for a discussion of treatment utility). Although child behavior assessors have fared somewhat better in this regard, they too have tended to design and use idiosyncratic, "convenient" tools for assessment. As we have suggested elsewhere (Ollendick & Hersen, 1984), comparison across studies is made difficult, if not impossible, and the advancement of an assessment science and technology, let alone an understanding of child behavior disorders and their effective treatments, is compromised with such an idiosyncratic approach.

While a multimethod approach that is based on developmentally sensitive and empirically validated procedures is recommended, it should be clear that a "test battery" approach is not being espoused. The specific devices to be selected depend on a host of factors, including the child's age, the nature of the referral question, the contexts in which the problematic behavior occurs, and the personnel, time, and resources available (Ollendick & Cerny, 1981). Nonetheless, given inherent limitations in each of the various procedures, as well as the desirability of obtaining as complete a picture of the child as possible, we recommend multimethod assessment whenever possible. Any one procedure, including direct behavioral observation, is not sufficient to provide a composite view of the child. The multimethod approach, if implemented, is not only helpful in assessing specific target behaviors and in determining response to behavior change, but also in understanding child behavior disorders and advancing assessment as a scientific endeavor. Based on these considerations, we offer the following tentative conclusions regarding child behavioral assessment:

(i) Children are a special and unique population. The automatic extension of adult behavioral assessment methods to children is not warranted and is often inappropriate. Further, not all "children" are alike. Clearly, a 16-year-old adolescent differs from a 12-year-old preadolescent who in turn differs from an 8-year-old middle-age child and a young 4-year-old child. Age-related variables affect the choice of methods as well as the procedures employed.

(ii) Given rapid developmental change observed in children as they grow, normative comparisons are required to ensure that appropriate target behaviors are selected and that change in behavior is related to treatment effects, and not normal developmental processes. Such comparisons require identification of suitable reference groups and information about the "natural course" of diverse child behavior problems (Ollendick & King, 1994).

(iii) Thorough child behavioral assessment involves multiple targets of change, including overt behavior, affective states, and cognitive processes. Further, such assessment entails determining the context (e.g., familial, social, cultural) in which the child's behavior occurs and the function(s) the target behaviors serve.

(iv) Given the wide range of targets for change and the imprecision of extant measures, multimethod assessment is desirable. Multimethod assessment should not be viewed simply as a test battery approach; rather, methods should be selected on the basis of their appropriateness to the referral question. Regardless of the measures used, they should be developmentally sensitive and empirically validated.

4.06.2 HISTORY AND DEVELOPMENT

As indicated above, assessment of children's behavior problems requires a multimethod approach in which data are gathered from clinical interviews and self- and other-report sources as well as from direct behavioral observations. In this manner, important information from the cognitive and affective modalities can be obtained and integrated with behavioral data to provide a more complete picture of the child. In addition, the multimethod approach provides the clinician with necessary detail regarding perceptions and reactions of significant others in the child's life (e.g., parents, teachers, peers). It should be noted, however, that this comprehensive and inclusive assessment approach is of relatively recent origin.

In its earliest stages, behavioral assessment of children relied almost exclusively on identification and specification of discrete and highly observable target behaviors (cf. Ullmann & Krasner, 1965). As such, assessment was limited to gathering information solely from the motoric (i.e., behavioral) response modality. This early assessment approach followed logically from theoretical assumptions of the operant school of thought which was in vogue at the time. Early on, behaviorally oriented clinicians posited that the only appropriate behavioral domain for empirical study was that which was directly observable (Skinner, 1953). Contending that objective demonstration of behavior change following intervention was of utmost importance, behaviorists relied upon data that could be measured objectively. Subjectivity, and the inferential process associated with it, were eschewed. Hence the frequency, intensity, and duration of problematic behaviors (i.e., "hard core" measures) were pursued. Although existence of cognitions and affective states was not denied, they were not deemed appropriate subject matter for experimental analysis.

As behavioral treatment approaches with children were broadened to include cognitive and self-control techniques in the 1970s (e.g., Bandura, 1977; Kanfer & Phillips, 1970; Kendall & Hollon, 1980; Meichenbaum, 1977), it became apparent that assessment strategies would have to expand into the cognitive and affective domains as well. Furthermore, even though operant techniques were shown to be efficacious in producing behavior change under controlled conditions, the clinical significance of these changes was less evident. The issue of clinical significance of behavior change is especially important in child behavioral assessment because children are invariably referred for treatment by others (e.g., parents, teachers). Once treatment goals have been identified, the ultimate index of treatment efficacy lies in the referral source's perceptions of change. Hence, other-report measures become as important as direct observational ones.

Furthermore, the scope of behavioral assessment has been expanded to include the impact of large-scale social systems (e.g., schools, neighborhoods) on the child's behavior (e.g., Patterson, 1976; Wahler, 1976). Although inclusion of these additional factors serves to complicate the assessment process, they are an indispensable part of modern-day child behavioral assessment. The ideologies and expectations of seemingly distal social systems often have immediate and profound effects on individual behavior (see Winett, Riley, King, & Altman, 1989, for discussion of these issues).

In sum, child behavioral assessment has progressed from sole reliance on measurement of target behaviors to a broader approach that takes into account cognitive and affective processes of the child that serve to mediate behavior change. Further, the social contexts (i.e., families, schools, communities) in which the problematic behaviors occur have been targeted for change. The assessment techniques that accompany this approach include behavioral interviews and self- and other-report instruments. These measures are utilized in addition to direct behavioral observation which remains the cornerstone of behavioral assessment (Mash & Terdal, 1981, 1989; Ollendick & Hersen, 1984, 1993).

4.06.3 THEORETICAL UNDERPINNINGS

Although behaviorism has had an historical development of its own, it is safe to state that the increased popularity of the behavioral approach has flourished, at least in part, due to dissatisfaction with the psychodynamic approach. A reflection of this dissatisfaction is that virtually all discussions of behavioral assessment are carried out through comparison and contrast with traditional assessment approaches (e.g., Bornstein, Bornstein, & Dawson, 1984; Cone & Hawkins, 1977; Goldfried & Kent, 1972; Hayes, Nelson, & Jarrett, 1986; Mash & Terdal, 1981, 1989; Mischel, 1968; Ollendick & Hersen, 1984, 1993). Though such comparisons often result in oversimplification of both approaches, they serve to elucidate theoretical underpinnings of the behavioral approach and its unique contributions. In this section, we will contrast the theoretical assumptions that guide behavioral and traditional assessment and discuss the practical implications of these assumptions for child behavioral assessment.

The most fundamental difference between traditional and behavioral assessment lies in the conception of "personality" and behavior (we place the construct "personality" in quotations because early behaviorists would have objected to use of this term, given its subjectivity and imprecise meaning). In the traditional assessment approach, personality is viewed as a reflection of underlying and enduring traits, and behavior is assumed to be caused by these internal personality characteristics ("personalism"). Aggressive behavior, for example, is assumed to reside "in" the child and to be caused by an underlying dynamic process attributed, perhaps, to hostility or anger and resulting from deep-seated intrapsychic conflict. "Aggression," it is said, is caused by the underlying hostility/anger. In contrast, behavioral approaches have avoided references to underlying personality constructs, focusing instead on what the child does under specific conditions. From the behavioral perspective, "personality" refers to patterns rather than causes of behavior (Staats, 1975, 1986). Furthermore, behavior is viewed as a result of current environmental factors ("situationalism") and of current environmental factors interacting with organismic variables ("interactionism"). Thus the role of the current environment is stressed more in behavioral assessment than in traditional assessment. The focus of assessment is on what the child does in that situation rather than on what the child has or "is" (Mischel, 1968). As a result, a lower level of inference is required in behavioral assessment than in traditional assessment.

It is important not to oversimplify the behavioral view of the causes of behavior, however. It has often been erroneously asserted that the behavioral approach focuses on external determinants of behavior to the exclusion of organismic states or internal cognitions and affects. To be sure, behavioral views of childhood disorders have emphasized the significant role of environmental factors in the manifestation and maintenance of behavior. However, organismic variables that influence behavior are not ignored or discounted. Among the organismic variables—dubbed cognitive social learning person variables (CSLPVs) by Mischel (1973)—that have been found to be important are competencies (skills which children possess such as social skills, problem-solving skills), encoding strategies (the manner in which children perceive or encode information about their environment), expectancies (expectancies about performance, including self-efficacy and outcome expectancies), subjective values (children's likes or dislikes, preferences or aversions), and self-regulatory systems and plans (children's capacity for and manner of self-imposing goals and standards and self-administering consequences for their behavior). A wide array of self-report instruments tapping CSLPVs and related cognitive and affective modalities for use in child behavioral assessment have been reviewed recently by us (Greene & Ollendick, in press). A thorough behavioral assessment should attempt to identify controlling variables, whether environmental or organismic in nature. As Mash and Terdal (1981) point out, "the relative importance of organismic and environmental variables and their interaction . . . should follow from a careful analysis of the problem" (p. 23).

The traditional conception of personality as made up of stable and enduring traits implies that behavior will be relatively persistent over time and consistent across situations. The behavioral view, in contrast, has been one of situational specificity; that is, because behavior is in large part a function of situational determinants and CSPLVs that are enacted only under specified conditions, a child's behavior will change as these situational factors are altered or the person variables are engaged. Similarly, consistency of behavior across the temporal dimension is not necessarily expected. Hence, as noted above, an aggressive act such as a child hitting another child would be seen from the traditional viewpoint as a reflection of underlying hostility which, in turn, would be hypothesized to be related to early life experiences or intrapsychic conflict. Little or no attention would be given to specific situational factors or the environmental context in which the aggressive act occurred. From the behavioral perspective, an attempt is made to identify those variables that elicit and maintain the aggressive act in that particular situation. That the child may aggress in a variety of situations is explained in terms of his or her learning history in which reinforcing consequences have been obtained for past aggressive acts (which help shape CSLPVs), and not in terms of an underlying personality trait of hostility. From this analysis, it is clear that actual behavior is of utmost importance to behaviorists, because it represents a sample of the child's behavioral repertoire in a specific situation. From the traditional viewpoint, the behavior assumes importance only insofar as it is a sign of some underlying trait.

These differing assumptions have implications for the assessment process. In behavioral assessment, the emphasis on situational specificity necessitates an assessment approach that samples behavior across a number of settings. Hence assessment of the child's behavior at home, in school, and on the playground is important in addition to information obtained in the clinic setting. Furthermore, it is not assumed that information obtained from these various settings will be consistent. The child may behave aggressively in school and on the playground with peers but not at home with siblings or parents. Or conversely, the child might behave aggressively at home but not at school or when with his or her peers. This lack of consistency in behavior would be problematic for the traditional approach, but not for the behavioral approach. Similarly, the notion of temporal instability requires the child's behavior be assessed at several points in time from a behavioral perspective, whereas such measurements across time would be less critical for the traditional approach.

At one point in time, it was relatively easy to differentiate behavioral from traditional assessment on the basis of the methods employed. Direct behavioral observation was the defining characteristic and often the sole assessment technique of the behavioral approach, whereas clinical interviews, self-report measures, and projective techniques characterized traditional assessment. However, as behavioral assessment was expanded to include a wider repertoire of assessment methods, differentiating behavioral and traditional assessments simply on the basis of assessment methods used has become more difficult. It is not uncommon for behaviorists to utilize information from clinical interviews and self-report instruments, and to pursue perceptions and expectancies of significant others in the child's environment. Thus there is considerable overlap in actual assessment practices, with one notable exception. Rarely, if ever, would projective techniques be utilized by the child behavioral assessor. The primary difference between traditional and behavioral assessment lies then not in the methods employed, but rather in the manner in which data from assessment sources are utilized. Traditional approaches interpret assessment data as signs of underlying personality functioning. These data are used to diagnose and classify the child and to make prognostic statements. From the behavioral perspective, assessment data are used to identify target behaviors and their controlling conditions (again, be they overt or covert). Information obtained from assessment serves as a sample of the child's behavior under specific circumstances. This information guides the selection of appropriate treatment procedures. Because behavioral assessment is ongoing, such information serves as an index by which to evaluate critically the effects of treatment and to make appropriate revisions in treatment. Further, because assessment data are viewed as samples of behavior, the level of inference is low, whereas a high level of inference is required when one attempts to make statements about personality functioning from responses to interview questions or test items.

In addition to these differences, Cone (1986) has highlighted the nomothetic and ideographic distinction between traditional and behavioral assessment. Stated briefly, the nomothetic approach is concerned with the discovery of general laws as they are applied to large numbers of children. Usually, these laws provide heuristic guidelines as to how certain variables are related to one another. Such an approach can be said to be variable-centered because it deals with particular characteristics (traits) such as intelligence, achievement, assertion, aggression, and so on. In contrast, the ideographic approach is concerned more with the uniqueness of a given child and is said to be child-centered rather than variable-centered. Unlike the nomothetic approach, the ideographic perspective emphasizes discovery of relationships among variables uniquely patterned in each child. The ideographic approach is most akin to the behavioral perspective, whereas the nomothetic approach is closely related to the traditional approach. As Mischel (1968) observed, "Behavioral assessment involves an exploration of the unique or idiosyncratic aspects of the single case, perhaps to a greater extent than any other approach" (p. 190). Cone (1986) illustrates how the ideographic/nomothetic distinction relates to the general activities of behavioral assessors by exploring five basic questions: What is the purpose of assessment? What is its specific subject matter? What general scientific approach guides this effort? How are differences accounted for? And, to what extent are currently operative environmental variables considered? Although further discussion of these important issues is beyond the scope of the present chapter, Cone's schema helps us recognize the pluralistic nature of behavioral assessment and calls our attention to meaningful differences in the practices contained therein. As Cone (1986) concludes, "There is not one behavioral assessment, there are many" (p. 126). We agree.

In sum, traditional and behavioral assessment approaches operate under different assumptions regarding the child's behavior. These assumptions, in turn, have implications for the assessment process. Of paramount importance for child behavior assessors is the necessity of tailoring the assessment approach to the specific difficulties of the child in order to identify the problem accurately, specify treatment, and evaluate treatment outcome. Such tailoring requires ongoing assessment from a number of sources under appropriately diverse stimulus conditions.

4.06.4 DESCRIPTION OF ASSESSMENT PROCEDURES

Multimethod behavioral assessment of children entails use of a wide variety of specific procedures. As behavioral approaches with children evolved from sole reliance on operant procedures to those involving cognitive and self-control procedures, the methods of assessment have changed accordingly. Identification of discrete target behaviors has been expanded to include assessment of cognitions and affects, as well as large-scale social systems that affect the child (e.g., families, schools, communities).

Information regarding these additional areas can be obtained most efficiently through behavioral interviews, self-reports, and other-reports. Cone (1978) has described these assessment methods as indirect ones; that is, while they may be used to measure behaviors of clinical relevance, they are obtained at a time and place different from when the behaviors actually occurred. In both behavioral interviews and self-report questionnaires, a verbal representation of the behaviors of interest is obtained. Other-reports, or ratings by others such as parents or teachers, are also included in the indirect category because they involve retrospective descriptions of behavior. Generally, a significant person in the child's environment (e.g., at home or school) is asked to rate the child based on previous observations in that setting (recollections).

As noted by Cone (1978), ratings such as these should not be confused with direct observation methods, which assess behaviors of interest at the time and place of their occurrence. Of course, information regarding cognition and affects, as well as the situations or settings in which they occur, can also be obtained through direct behavioral observations, either by self-monitoring or through trained observers. In the sections that follow, both indirect and direct methods are reviewed.

4.06.4.1 Behavioral Interviews

The first method of indirect assessment to be considered is the behavioral interview. Of the many procedures employed by behavioral clinicians, the interview is the most widely used (Swann & MacDonald, 1978) and is generally considered an indispensable part of assessment (Gross, 1984; Linehan, 1977). Behavioral interviews are frequently structured to obtain information about the target behaviors and their controlling variables and to begin the formulation of specific treatment plans. While the primary purpose of the behavioral interview is to obtain information, we have found that traditional "helping" skills including reflections, clarifications, and summary statements help put children and their families at ease and greatly facilitate collection of this information (Ollendick & Cerny, 1981). As with traditional therapies, it is important to establish rapport with the child and family and to develop a therapeutic alliance (i.e., agreement on the goals and procedures of therapy) in the assessment phase of treatment (Ollendick & Ollendick, 1997).

Undoubtedly, the popularity of the behavioral interview is derived in part from practical considerations associated with its use. While direct observation of target behaviors remains the hallmark of behavioral assessment, such observations are not always practical or feasible. At times, especially in outpatient therapy in clinical settings, the clinician might have to rely on children's self-report as well as that of their parents to obtain critical detail about problem behaviors and their controlling variables. Further, the interview affords the clinician the opportunity to obtain information regarding overall functioning in a number of global areas (e.g., home, school, neighborhood), in addition to specific information about particular problem areas. The flexibility inherent in the interview also allows the clinician to build a relationship with the child and the family and to obtain information that might otherwise not be revealed. As noted early on by Linehan (1977), some family members may be more likely to divulge information verbally in the context of a professional relationship than to write it down on a form to be entered into a permanent file. In our experience, this is not an uncommon occurrence. That is, certain family members report little or no difficulties on intake reports or on self-report measures, yet they divulge a number of problem areas during the structured behavioral interview.

In addition, the interview allows the clinician the opportunity to observe the family as a whole and to obtain information about the familial context in which the problem behaviors occur. Several interrelated issues may arise when child behavioral assessment is expanded to include the family unit (Evans & Nelson, 1977; Ollendick & Cerny, 1981). First, children rarely refer themselves for treatment; invariably, they are referred by adults whose perceptions of problems may not coincide with the child's view. This is especially true when problems are centered around externalizing behaviors such as oppositional or disruptive behaviors, less so with internalizing behaviors (e.g., anxiety or depression). Moreover, it is not uncommon for the perception of one adult to differ from that of another (i.e., the mother and father disagree, or the teacher and the parents disagree; cf, Achenbach, McConaughy, & Howell, 1987). A second issue, related to the first, is the determination of when child behaviors are problematic and when they are not. Normative developmental comparisons are useful in this regard (Lease & Ollendick, 1993; Ollendick & King, 1991). It is not uncommon for parents to refer 3-year-olds who wet the bed, 5-year-olds who reverse letters, 10-year-olds who express interest in sex, and 13-year-olds who are concerned about their physical appearance. Frequently, these referrals are based on parental uneasiness or unrealistic expectations rather than genuine problems (see Campbell, 1989, for further discussion of these issues). Finally, problematic family interactions (especially parent–child interactions) are frequently observed in families in which a particular child has been identified and referred for treatment (cf. Dadds, Rapee, & Barrett, 1994; Patterson, 1976, 1982). These interactions may not be a part of the parents' original perception of the "problem." Furthermore, assessment of such interactions allows the clinician an opportunity to observe the verbal and nonverbal behaviors of the family unit in response to a variety of topics, and of family members in response to each other. Structured interviews assessing parent–child interactions have been developed for a number of behavior problems (e.g., Barkley, 1987; Dadds et al., 1994).

Ideally, evaluation of parental perceptions and parent–child interactions will enable the clinician to conceptualize the problematic behaviors and formulate treatment plans from a more comprehensive, integrated perspective. However, the above discussion is not meant to imply that the behavioral interview should be limited to the family; in many instances, the practices described above should be extended to adults outside the family unit, such as teachers, principals, and physicians, and to environments beyond the home, including schools and daycare centers. For example, if a problem behavior is reported to occur primarily at school, assessing the perceptions and behavioral goals of a teacher and/or principal will be necessary (Greene, 1995, 1996), and evaluating teacher–

student interactions may prove more productive than observing parent–child interactions during the clinical interview. Finally, the clinician should approach the behavioral interview with caution and avoid blind acceptance of the premise that a "problem" exists "in" the child. Information obtained in a comprehensive assessment may reveal the behavior of the identified client is only a component of a more complex clinical picture involving parents, siblings, other adults, and/or social systems.

In sum, an attempt is made during the behavioral interview to obtain as much information as possible about the child, his or her family, and other important individuals and environments. While the interview is focused around specific target behaviors, adult–child interactions and adult perceptions of the problem may also be assessed. These perceptions should be considered tentative, however, and used primarily to formulate hypotheses about target behaviors and their controlling variables and to select additional assessment methods to explore target behaviors in more depth (e.g., rating scales, self-reports, self-monitoring, and behavioral observations). The behavioral interview is only the first step in the assessment process.

Brief mention should also be made here of structured diagnostic interviews and their role in child behavioral assessment. In some instances, most notably when a diagnosis is required, it may be desirable for the clinician to conduct a structured diagnostic interview. In general, diagnostic interviews are oriented toward obtaining specific information to determine if a child "meets" diagnostic criteria for one or more specific diagnoses included in the *Diagnostic and statistical manual of mental disorders* (4th ed., *DSM-IV*) (American Psychiatric Association, 1994) or the *International classification of diseases* (10th ed., *ICD-10*; World Health Organization, 1991). Such interviews facilitate collection of data relative to a broad range of "symptoms" (i.e., behaviors) and psychiatric diagnoses. Several "omnibus" diagnostic interviews are available, including the Diagnostic Interview Schedule for Children-Version 2.3 (Shaffer, 1992), which was recently revised to reflect *DSM-IV* criteria. Other diagnostic interviews are oriented toward a specific domain such as anxiety (e.g., the Anxiety Disorders Interview Schedule for Children; Silverman & Nelles, 1988). It, too, has recently been revised to incorporate *DSM-IV* criteria (Silverman & Albano, 1996). Both child and parent forms of these interviews are available. Although these structured diagnostic interviews provide a wealth of information, they are limited by an overemphasis on diagnostic categories (to the exclusion of important details regarding specific target behaviors and their controlling variables), weak or untested reliability for children under age 11, low correspondence between responses of children and their parents, and categorical vs. dimensional scoring criteria (McConaughy, 1996). Further, structured diagnostic interviews often do not yield specific information about contextual factors associated with the child's problematic behavior; thus, when a diagnostic interview is used, it needs to be supplemented with a problem-focused interview. In our opinion, diagnostic interviews should not be considered as replacements for problem-focused interviews; rather they should be viewed as complementary.

4.06.4.2 Ratings and Checklists

Following the initial behavioral interview and, if necessary, the diagnostic interview, significant others in the child's environment may be requested to complete rating forms or checklists. In general, these forms are useful in providing an overall description of the child's behavior, in specifying dimensions or response clusters that characterize the child's behavior, and in serving as outcome measures for the evaluation of treatment efficacy. Many of these forms contain items related to broad areas of functioning such as school achievement, peer relationships, activity level, and self-control. As such, they provide a cost-effective picture of children and their overall level of functioning. Further, the forms are useful in eliciting information that may have been missed in the behavioral interview (Novick, Rosenfeld, Bloch, & Dawson 1966). Finally, the forms might prove useful in the search for the best match between various treatments (e.g., systematic desensitization, cognitive restructuring, and self-control) and subtypes of children as revealed on these forms (Ciminero & Drabman, 1977).

The popularity of omnibus rating forms and checklists is evident in the number of forms currently available (McMahon, 1984). Three of the more frequently used forms are described here: the Behavior Problem Checklist (Quay & Peterson, 1967, 1975) and its revision (Quay & Peterson, 1983); the Child Behavior Checklist (Achenbach, 1991a, 1991b); and the recently developed Behavior Assessment System for Children (Reynolds & Kamphaus, 1932). Based on Peterson's (1961) early efforts to sample diverse child behavior problems, the Revised Behavior Problem Checklist consists of 89 items, each rated on a three-point severity scale. While some of the items are general and require considerable inference (e.g., lacks self-confidence, jealous), others are more specific (e.g., cries, sucks thumb). Six primary dimensions or response clusters of child behavior have been identified on this scale: conduct problems, socialized aggression, attention problems, anxiety-withdrawal, psychotic behavior, and motor excess. Interestingly, the two primary problem clusters found on this checklist are similar to those found in numerous factor-analytic studies of other rating forms and checklists. These two factors or response clusters represent consistent dimensions of child behavior problems, reflecting externalizing (e.g., acting out) and internalizing (e.g., anxiety, withdrawal) dimensions of behavior (Achenbach, 1966).

While the Behavior Problem Checklist has a rather lengthy history and is one of the most researched scales, it does not include the rating of positive behaviors and, hence, does not provide a basis on which to evaluate more appropriate, adaptive behaviors. A scale that does assess appropriate behaviors, as well as inappropriate ones, is the Child Behavior Checklist (CBCL; Achenbach, 1991a, 1991b; Achenbach & Edelbrock, 1989). The scale, designed for both parents and teachers, contains both social competency and behavior problem items. The parent-completed CBCL is available in two formats depending on the age of the child being evaluated (i.e., 2–3 years and 4–18 years). The CBCL 4–18, for example, consists of 112 items rated on a three-point scale. Scored items can be clustered into three factor-analyzed profiles: social competence, adaptive functioning, and syndrome scales. The latter includes eight syndrome scales: withdrawn, somatic complaints, anxious/depressed, social problems, thought problems, attention problems, aggressive behavior, and delinquent behavior. Social competency items examine the child's participation in various activities (e.g., sports, chores, hobbies) and social organizations (e.g., clubs, groups), as well as performance in the school setting (e.g., grades, placements, promotions). The teacher-completed CBCL (TRF; Teacher Report Form) also consists of 112 items which are fairly similar, but not completely identical to, those found on the CBCL completed by parents. The scored items from the teacher form cluster into the same three factor-analyzed profiles; further, the eight syndrome scales are the same for the two measures, allowing for cross-informant comparisons. As with Quay and Peterson's Behavior Problem Checklist, some of the items are general and require considerable inference (e.g., feels worthless, acts too young, fears own impulses), while others are more specific and easily scored (e.g., wets bed, sets fires, destroys own things). Broadband grouping of the factors reflects the aforementioned internalizing and externalizing behavioral dimensions.

Although the Behavior Problem Checklist and Child Behavior Checklist have enjoyed considerable success, the recently developed Behavior Assessment System for Children (BASC; Reynolds & Kamphaus, 1992) represents a challenge to both of these well-established rating scales. Like these other instruments, the BASC is an omnibus checklist composed of parent, teacher, and child versions. It also contains a developmental history form and a classroom observation form. Most similar to Achenbach's Child Behavior Checklist (Achenbach, 1991a), Teacher's Report Form (Achenbach, 1991b), and Youth Self-Report (Achenbach, 1991c), the parent, teacher, and self-report forms of the BASC contain items that tap multiple emotional and behavioral domains and produce scale scores that represent pathological and adaptive characteristics of the child. Unlike the empirically derived scales of Achenbach's checklists, however, the scales of the BASC were created conceptually to represent content areas relevant to assessment and classification in clinical and educational settings.

For example, the BASC Parent Rating Scale (BASC-PRS) yields *T* scores in broad externalizing and internalizing domains as well as in specific content areas, including aggression, hyperactivity, conduct problems, attention problems, depression, anxiety, withdrawal, somatization, and social skills.

In addition, it provides *T* scores in areas of social competency such as leadership and adaptability. Preschool (ages 4–5), child (ages 6–11), and adolescent (ages 12–18) forms are available. Recent findings suggest the utility of this instrument with both clinical and educational populations and in identifying youth at risk for maladaptive outcomes (cf, Doyle, Ostrander, Skare, Crosby, & August, 1997). Although initial findings associated with its use appear promising, much more research is needed before its routine acceptance can be endorsed.

In addition to these more general rating forms, rating forms specific to select problem areas are also available for use in child behavioral assessment. Three such forms have been chosen for purposes of illustration: one used in the assessment of an internalizing dimension (fears/anxiety), another in the assessment of an externalizing dimension (defiance/noncompliance), and the final one in measuring a specific area of social competency.

The Louisville Fear Survey Schedule for Children (Miller, Barrett, Hampe, & Noble,

1972) contains 81 items that address an extensive array of fears and anxieties found in children and adolescents. Each item is rated on a three-point scale by the child's parents. Responses to specific fear items can be used to subtype fearful children. For example, Miller et al. (1972) were able to differentiate among various subtypes of school-phobic children on the basis of this instrument.

The Home Situations Questionnaire (HSQ; Barkley, 1981; Barkley & Edelbrock, 1987) contains 16 items representing home situations in which noncompliant behavior may occur (e.g., while playing with other children, when asked to do chores, and when asked to do homework). For each situation, parents indicate whether noncompliant behavior is a problem and then rate each of the 16 problematic situations on a nine-point scale (mild to severe); thus the scale assesses both the number of problem situations and the severity of noncompliant behavior. The HSQ has been shown to be sensitive to stimulant-drug effects (Barkley, Karlsson, Strzelecki, & Murphy, 1984), to discriminate children with behavior problems from normal children (Barkley, 1981), and to be sensitive to the effects of parent training programs (Pollard, Ward, & Barkley, 1983). The HSQ was selected for inclusion in this chapter because it may also be used in conjunction with a companion scale, the School Situations Questionnaire (SSQ; Barkley, 1981; Barkley & Edelbrock, 1987), which is completed by teachers. This scale includes 12 school situations most likely to be problematic for clinic-referred, noncompliant children, including "during lectures to the class," "at lunch," and "on the bus." Teachers rate the occurrence and severity of noncompliant behavior on a scale identical to that of the HSQ. In earlier sections, we emphasized the importance of assessing child behavior in multiple environments; the HSQ and SSQ are representative of recent efforts to develop measures for this purpose, thus providing us important contextual information about specific problem behaviors.

In some instances, it may be useful to obtain more information about a positive, adaptive domain of behavior—such as social skills or self-regulation—than that provided for by omnibus rating scales such as the Revised Behavior Problem Checklist (Quay & Peterson, 1983) or the Child Behavior Checklist (Achenbach, 1991a, 1991b). For example, the Social Skills Rating System (Gresham & Elliot, 1990), a 55-item questionnaire, provides specific information about a child's behavior in three domains (social skills, problem behaviors, and academic competence). Parent, teacher, and self-rating forms are available. In general, this instrument provides important and detailed information about academic and social competence that can be used to supplement information obtained from the more generic rating scales.

In sum, a variety of other-report instruments are available. As noted earlier, these forms must be considered indirect methods of assessment because they rely on retrospective descriptions of the child's behavior. For all of these scales, an informant is asked to rate the child based on past observations of that child's behavior. Global scales such as the Revised Behavior Problem Checklist, Child Behavior Checklist, and Behavior Assessment Scale for Children comprehensively sample the range of potential behavior problems, while more specific scales such as the Louisville Fear Survey Schedule for Children, the Home Situations Questionnaire, and the Social Skills Rating System provide detailed information about particular maladaptive or adaptive behaviors. Both types of other-report instruments provide useful, albeit different, information in the formulation and evaluation of treatment programs.

4.06.4.3 Self-report Instruments

Concurrent with the collection of other-reports regarding the child's behavior from significant others, self-reports of attitudes, feelings, and behaviors may be obtained directly from the child. As noted earlier, behaviorists initially eschewed such data, maintaining that the only acceptable datum was observable behavior. To a large extent, this negative bias against self-report was an outgrowth of early findings indicating that reports of subjective states did not always coincide with observable behaviors (Finch & Rogers, 1984). While congruence in responding is, in fact, not always observed, contemporary researchers have cogently argued that children's perceptions of their own behavior and its consequences may be as important for behavior change as the behavior itself (Finch, Nelson, & Moss, 1993; Ollendick & Hersen, 1984, 1993). Furthermore, as noted earlier, although different assessment procedures may yield slightly different information, data from these sources should be compared and contrasted in order to produce the best picture of the child and to derive treatment goals and procedures. Although self-report instruments have specific limitations, they can provide valuable information about children and their presenting problem; furthermore, they can be used as an index of change following treatment.

A wide array of self-report instruments have been developed for children. Some self-report instruments focus on a broad range of behavioral, cognitive, and affective functioning, as in the case of the Youth Self-report (Achenbach, 1991c). Other self-report instruments tap more specific areas of interest, such as anger (Nelson & Finch, 1978), anxiety (Reynolds & Richmond, 1985; Spielberger, 1973), assertion (Deluty, 1979; Ollendick, 1983a), depression (Kovacs, 1985), and fear (Scherer & Nakamura, 1968). Each of these instruments has been carefully developed and empirically validated. Three of the more frequently used instruments we have found to be useful in our clinical practice will be described briefly.

Spielberger's State–Trait Anxiety Inventory for Children (1973) consists of 20 items that measure state anxiety and 20 items that tap trait anxiety. The state form is used to assess transient aspects of anxiety, while the trait form is used to measure more global, generalized aspects of anxiety. Combined, the two scales can provide both process and outcome indices of change in self-reported anxiety. That is, the state form can be used to determine session-by-session changes in anxiety, while the trait form can be used as a pretreatment, posttreatment, and follow-up measure of generalized anxiety. A clear advantage of this instrument is that the state scale is designed so responses to relatively specific anxiety-producing situations can be determined. For example, children can be instructed to indicate how they feel "at this moment" about standing up in front of class, leaving home for summer camp, or being ridiculed by peers. Further, cognitive, motoric, and physiologic indicators of anxiety can be endorsed by the child (e.g., feeling upset, scared, mixed up, jittery, or nervous). Responses to items are scored on a three-point scale (e.g., "I feel very scared/scared/not scared"). Finally, the pervasiveness of the anxiety response can be measured by the trait form. The Spielberger scales are most useful for children aged 9–12, but have been used with both younger children and adolescents as well.

A second instrument that has been used frequently in child behavioral assessment is the Fear Survey Schedule for Children (Scherer & Nakamura, 1968) and its revision (Ollendick, 1983b). In the revised scale, designed to be used with younger and middle-age (9–12) children, children are instructed to rate their level of fear to each of 80 items on a three-point scale. They are asked to indicate whether a specific fear item (e.g., having to go to school, snakes, dark places, riding in a car) frightens them "not at all," "some," or "a lot." Factor analysis of the scale has revealed five primary factors: fear of failure or criticism, fear of the unknown, fear of injury and small animals, fear of danger and death, and medical fears. This pattern of fear has been shown to be relatively invariant across several nationalities, including American (Ollendick, Matson & Hetsel, 1985), Australian (Ollendick, King, & Frary, 1989), British (Ollendick, Yule, & Ollier, 1991), Chinese (Dong, Yang, & Ollendick, 1994), and Nigerian youth (Ollendick, Yang, King, Dong, & Akande, 1996). Further, it has been shown that girls report more fear than boys in these various countries, that specific fears change developmentally, and that the most prevalent fears of boys and girls have remained unchanged over the past 30 years (although some differences have been noted across nationalities). Such information is highly useful when determining whether a child of a specific age and gender is excessively fearful. Further, the instrument has been used to differentiate subtypes of phobic youngsters whose fear of school is related to separation anxiety (e.g., death, having parents argue, being alone) from those whose fear is due to specific aspects of the school situation itself (e.g., taking a test, making a mistake, being sent to the principal). When information from this instrument is combined with that from parents on the Louisville Fear Survey Schedule for Children (Miller et al., 1972), a relatively complete picture of the child's characteristic fear pattern can be obtained.

The final self-report instrument to be reviewed is Kovac's (1985) Children's Depression Inventory (CDI). Since the mid-1980s, no other area in clinical child psychology has received more attention than depression in children. A multitude of issues regarding its existence, nature, assessment, and treatment have been examined (Cantwell, 1983; Rutter, 1986). One of the major obstacles to systematic investigations in this area has been the absence of an acceptable self-report instrument, and the CDI appears to meet this need. The instrument is a 27-item severity measure of depression based on the well-known Beck Depression Inventory. Each of the 27 items consists of three response choices designed to range from mild depression to fairly severe and clinically significant depression. Kovacs reports that the instrument is suitable for middle-age children and adolescents (8–17 years of age). We have found the instrument to be useful with younger children as well, especially when items are read aloud and response choices are depicted on a bar graph. Smucker, Craighead, Craighead, and Green (1986) have provided additional psychometric data on the CDI. Overall, they conclude it is a reliable, valid, and clinically useful instrument for children and adolescents.

In sum, a variety of self-report instruments are available. As with other-report forms, self-reports should be used with appropriate caution and due regard for their specific limitations. Because they generally involve the child's retrospective rating of attitudes, feelings, and behaviors, they too must be considered indirect methods of assessment (Cone, 1978). Nevertheless, they can provide valuable information regarding children's own perception of their behavior.

4.06.4.4 Self-monitoring

Self-monitoring differs from self-report in that it constitutes an observation of clinically relevant target behaviors (e.g., thoughts, feelings, actions) at the time of their occurrence (Cone, 1978). As such, it is a direct method of assessment. Self-monitoring requires children to observe their own behavior and then to record its occurrence systematically. Typically, the child is asked to keep a diary, place marks on a card, or push the plunger on a counter as the behavior occurs or shortly thereafter. Although self-monitoring procedures have been used with both children and adults, at least three considerations must be attended to when such procedures are used with younger children (Shapiro, 1984): behaviors should be clearly defined, prompts to use the procedures should be readily available, and rewards for their use should be provided. Some younger children will be less aware of when the target behavior is occurring and will require coaching and assistance prior to establishing a monitoring system. Other young children may have difficulty remembering exactly what behaviors to monitor and how those behaviors are defined. For these reasons, it is generally considered advisable to provide the child with a brief description of the target behavior or, better yet, a picture of it, and to have the child record only one or two behaviors at a time. In an exceptionally sensitive application of these guidelines, Kunzelman (1970) recommended the use of COUNTOONS, simple stick figure drawings that depict specific behaviors to be self-monitored. Children are instructed to place a tally mark next to the picture when the behavior occurs. For example, a girl monitoring hitting her younger brother may be given an index card with a drawing of a girl hitting a younger boy and instructed to mark each time she does what the girl in the picture is doing. Of course, in a well-designed program, the girl might also be provided with a picture of a girl and a younger boy sharing toys and asked as well to mark each time she emits the appropriate behavior. Such pictorial cues serve as visual prompts for self-monitoring. Finally, children should be reinforced profusely following successful use of self-monitoring.

In general, methods of self-monitoring are highly variable and depend on the specific behavior being monitored and its place of occurrence. For example, Shapiro, McGonigle, and Ollendick (1980) had mentally retarded and emotionally disturbed children self-monitor on-task behavior in a school setting by placing gummed stars on assignment sheets; Ollendick (1981) had children with tic disorders place tally marks upon the occurrence of tics on a colored index card carried in the child's pocket; and Ollendick (1995) had adolescents diagnosed with panic disorder and agoraphobia indicate the extent of their agoraphobic avoidance on a 1–5 scale each time they encountered the feared situation. He also had the adolescents indicate their confidence (i.e., self-efficacy) in coping with their fear on a similar 1–5 scale. In our clinical work, we have also used wrist counters with children whose targeted behaviors occur while they are "on the move." Such a device is not only easy to use, but serves as a visual prompt to self-record. The key to successful self-monitoring in children is the use of recording procedures that are uncomplicated. They must be highly portable, simple, time-efficient, and relatively unobtrusive (Greene & Ollendick, in press).

In sum, self-monitoring procedures represent a direct means of obtaining information about the target behaviors as well as their antecedents and consequences. While specific monitoring methods may vary, any procedure that allows the child to monitor and record presence of the targeted behaviors can be used. When appropriate procedures are used, self-monitoring represents a direct and elegant method of assessment (Ollendick & Greene, 1990; Ollendick & Hersen, 1993).

4.06.4.5 Behavioral Observation

Direct observation of the child's behavior in the natural environment is the hallmark of child behavioral assessment. As described by Johnson and Bolstad (1973), the development of naturalistic observation procedures represents one of the major, if not the major, contributions of the behavioral approach to assessment and treatment of children. A direct sample of the child's behavior at the time and place of its occurrence is obtained with this approach. As such, it is the least inferential of the assessment methods described heretofore. However, behavioral observations in the naturalistic environment should not be viewed as better than other methods of assessment. Rather, direct observations should be viewed as complementary to the other methods, with each providing different and valuable information.

In behavioral observation systems, a single behavior or set of behaviors that have been identified as problematic (generally through the aforementioned procedures) are operationally defined, observed, and recorded in a systematic fashion. In addition, events that precede and follow behaviors of interest are recorded and subsequently used in development of specific treatment programs. Although Jones, Reid, and Patterson (1975) have recommended use of "trained impartial observer-coders" for collection of these data, this is rarely possible in the practice of child behavioral assessment in the clinical setting. Frequently, time constraints, lack of trained personnel, and insufficient resources mitigate against the use of highly trained and impartial observers. In some cases, we have used significant others in the child's environment (e.g., parents, teachers, siblings) or the children themselves as observers of their own behavior. Although not impartial, these observers can be trained adequately to record behaviors in the natural environment. In other cases, behavioral clinicians have resorted to laboratory or analogue settings that are similar to, but not identical to, the natural environment. In these simulated settings, children may be asked to behave as if they are angry with their parents, to role play assertive responding, or to approach a highly feared object. Behaviors can be directly observed or videotaped (or audiotaped) and reviewed retrospectively. The distinguishing characteristic of behavioral observations, whether made in the naturalistic environment or in simulated settings, is that a direct sample of the child's behavior is obtained.

A wide variety of target behaviors have been examined using behavioral observation procedures. These behaviors have varied from relatively discrete behaviors like enuresis and tics, that require relatively simple and straightforward recording procedures, to complex social interactions that necessitate extensive behavioral coding systems (e.g., Dadds et al., 1994; O'Leary, Romanczyk, Kass, Dietz, & Santogrossi, 1971; Patterson, Ray, Shaw, & Cobb, 1969; Wahler, House, & Stambaugh, 1976).

The utility of behavioral observations in naturalistic and simulated settings is well illustrated in Aylion, Smith, and Rogers' (1970) behavioral assessment of a young school-phobic girl. In this case study, impartial observers in the child's home monitored the stream of events occurring on school days in order to identify the actual school-phobic behaviors and to determine the antecedent and consequent events associated with them. In this single-parent family, it was noted that the mother routinely left for work about one hour after the targeted girl (Valerie) and her siblings were to leave for school. Although the siblings left for school without incident, Valerie was observed clinging to her mother and refusing to leave the house and go to school. As described by Ayllon et al. (1970), "Valerie typically followed her mother around the house, from room to room, spending approximately 80 percent of her time within 10 feet of her mother. During these times there was little or no conversation" (p. 128). Given her refusal to go to school, the mother took Valerie to a neighbor's apartment for the day. However, when the mother attempted to leave for work, Valerie frequently followed her at a 10-foot distance. As a result, the mother had to return to the neighbor's apartment with Valerie in hand. This daily pattern was observed to end with the mother "literally running to get out of sight of Valerie" so she would not follow her to work. During the remainder of the day, it was observed that Valerie was allowed to do whatever she pleased: "Her day was one which would be considered ideal by many grade-school children—she could be outdoors and play as she chose all day long. No demands of any type were placed on her" (p. 129). Based on these observations, it appeared that Valerie's separation anxiety and refusal to attend school were related to her mother's attention and to the reinforcing environment of the neighbor's apartment where she could play all day.

However, because Valerie was also reported to be afraid of school itself, Ayllon et al. (1970) designed a simulated school setting in the home to determine the extent of anxiety or fear toward specific school-related tasks. (Obviously, observation in the school itself would have been desirable but was impossible because she refused to attend school.) Unexpected, little or no fear was evinced in the simulated setting; in fact, Valerie performed well and appeared to enjoy the school-related setting and homework tasks. In this case, these detailed behavioral observations were useful in ruling upon differential hypotheses related to school refusal. They led directly to a specific and efficacious treatment program based on shaping and differential reinforcement principles. The utility of behavioral observations for accurate assessment and treatment programming has been noted in numerous other case studies as well (e.g., Ollendick, 1995; Ollendick & Gruen, 1972; Smith & Sharpe, 1970).

A major disadvantage of behavioral observations in the natural environment is that the target behavior may not occur during the designated observation periods. In such instances, simulated settings that occasion the target behaviors can be used. Simulated observations are especially helpful when the target behavior is of low frequency, when the target behavior is not observed in the naturalistic setting due to reactivity effects associated with being observed, or when the target behavior is difficult to observe in the natural environment due to practical constraints. Ayllon et al.'s (1970) use of a simulated school setting illustrated this approach under the latter conditions. A study by Matson and Ollendick (1976) illustrates this approach for low-frequency behaviors. In this study, parents reported that their children bit either the parent or siblings when they "were unable to get their way or were frustrated." Direct behavioral observations in the home confirmed the parental report, but it was necessary to observe the children for several hours prior to observing an occurrence of the behavior. Further, parents reported that their children were being "good" while the observers were present and that frequency of biting was much lower than its usual, "normal" rate. Accordingly, parents were trained in observation procedures and instructed to engage their children in play for four structured play sessions per day. During these sessions, parents were instructed to prompt biting behavior by deliberately removing a preferred toy. As expected, removal of favored toys in the structured situations resulted in increases in target behaviors, which were then eliminated through behavioral procedures. The structured, simulated play settings maximized the probability that biting would occur and that it could be observed and treated under controlled conditions.

It is often intimated that behavioral observation systems may not be suitable for more complex behavior problems, such as parent–child interactions. Sophisticated systems developed by Dumas (1989) and Dadds et al. (1994) to capture family interactions and processes suggest otherwise. For example, Dadds et al. (1994) developed the Family Anxiety Coding Schedule in order to measure anxious behavior in both child and parent, and the antecedents and consequences each provided the other to occasion anxiety in the other. This schedule was developed following the observation that children learned to process information about threat cues through interactions with their parents. More specifically, they observed that anxious children tended to view "neutral" situations as more threatening after discussing the situations with their parents than they did in the absence of such interactions. To learn more about how this happened, Dadds and colleagues observed the moment-to-moment process whereby parents of anxious children influenced their children to change from a nonthreatened stance to an avoidant, threatened stance. To examine the interdependency of the parents and the child, they coded each family member's utterances in real time sequence so that conditional probabilities could be computed between different family members' behaviors. Using this system, they were able to show the process by which, and through which, the anxiety response was activated and maintained in the child. Thus a very complicated process of parent–child interactions was broken down into its constituent parts, recorded with a sophisticated observation system, and analyzed sequentially over time. Moreover, the observations suggested that, in this sample of overanxious children, anxiety did not exist solely "in" the child; rather, it existed in a context that was highly dependent upon parental influences. Such a demonstration illustrates the importance of contextual influences in understanding, assessing, and treating diverse child behavior disorders.

In sum, direct behavioral observation—either in the natural or simulated environment—provides valuable information for child behavioral assessment. When combined with information gathered through behavioral interviews, self- and other-reports, and self-monitoring, a comprehensive picture of children and their behaviors, as well as their controlling variables, is obtained. As with other assessment procedures, however, direct behavioral observation alone is not sufficient to meet the various behavioral assessment functions required for a thorough analysis of a child's problem behavior.

4.06.5 RESEARCH FINDINGS

As noted earlier, use of assessment instruments and procedures that have been empirically validated is one of the primary characteristics of child behavioral assessment. However, the role of conventional psychometric standards in evaluating child behavioral assessment procedures is a controversial one (e.g., Barrios & Hartman, 1986; Cone, 1981, 1986; Cone & Hawkins, 1977; Mash & Terdal, 1981). Given the theoretical underpinnings of child behavioral assessment and the basic assumptions regarding situational specificity and temporal instability of behavior, traditional psychometric standards would appear to be of little or no value. After all, how can behaviors thought to be under the control of specific antecedent and consequent events be expected to be similar in different settings and at different times? Yet, if there is no consistency in behavior across settings and time, prediction of behavior is impossible and the generalizability of findings obtained from any one method of assessment would be meaningless. Such an extreme ideographic stance precludes meaningful assessment, except of highly discrete behaviors in very specific settings and at very specific points in time (Ollendick & Hersen, 1984).

Research findings suggest that it is not necessary totally to dismiss notions of cross-situational and cross-temporal consistency of behavior (e.g., Bem & Allen, 1974). Although a high degree of behavioral consistency cannot be expected, a moderate degree of behavioral consistency can be expected across situations that involve similar stimulus and response characteristics and are temporally related. When multimethod assessment procedures are used under these conditions, a modest relationship among the measures and a fair degree of predictability and generalizability can be expected. Under such circumstances, application of conventional psychometric standards to evaluation of child behavioral assessment procedures is less problematic and potentially useful (Cone, 1977; Ollendick & Hersen, 1984, 1993). The value of psychometric principles has already been demonstrated for certain classes of behavior when obtained through methods such as behavioral observation (e.g., Olweus, 1979), self-report (e.g., Ollendick, 1981), and other-report ratings (e.g., Cowen, Pederson, Barbigian, Izzo, & Trost, 1973). Further, when multiple methods of behavioral assessment have been used in the same studies, a modest degree of concurrent and predictive validity has been reported (e.g., Gresham, 1982).

It is beyond the scope of the present chapter to review specific research findings related to the reliability, validity, and clinical utility of the various procedures espoused in the multimethod approach. Nonetheless, brief mention will be made of specific directions of research and ways of enhancing the psychometric qualities of each procedure.

4.06.5.1 Behavioral Interviews

As noted by Evans and Nelson (1977), data based on retrospective reports obtained during the interview may possess both low reliability (agreement among individuals interviewed may differ and responses may vary over time) and low validity (reported information may not correspond to the "facts"). Such inaccurate or distorted recollections may result not only in delayed clarification of the presenting complaints, but also in faulty hypotheses about causal agents and maintaining factors. For example, Chess, Thomas, and Birch (1966) reported that parents inaccurately reported certain behavior problems developed at times predicted by popular psychological theories. For example, problems with siblings were recalled to have begun with the birth of a younger sibling, and problems with dependency were reported to have begun when the mother became employed outside the home. In actuality, these behaviors were present prior to these events; nonetheless, they were "conveniently" recalled to have begun coincident with commonly accepted "life" points. In a similar vein, Schopler (1974) noted that many parents of autistic children inaccurately blame themselves for their child's problematic behaviors and that many therapists inadvertently "buy into" this notion that parents are to blame. Such scapegoating accomplishes little in the understanding, assessment, and treatment of the child's problematic behavior (Ollendick & Cerny, 1981).

While the reliability and validity of general information about parenting attitudes and practices are suspect, findings suggest parents and children can be reliable and valid reporters of current, specific information about problematic behaviors (e.g., Graham & Rutter, 1968; Gross, 1984; Herjanic, Herjanic, Brown, & Wheatt, 1973). The reliability and validity of the information are directly related to recency of behaviors being discussed and specificity of information requested. Thus, careful specification of precise behaviors and conditions under which they are occurring is more reliable and valid than vague descriptions of current behaviors or general recollections of early childhood events (Ciminero & Drabman, 1977). When the interview is conducted along such guidelines, it is useful in specifying behaviors of clinical interest and in determining appropriate therapeutic interventions. As we have noted, however, it is only the first step in the ongoing, hypothesis-generating process that is characteristic of child behavioral assessment.

4.06.5.2 Ratings and Checklists

As with behavioral interviews, issues related to reliability and validity are also relevant to ratings and checklists. Cronbach (1960) has noted that the psychometric quality of rating scales is directly related to the number and specificity of the items rated. Further, O'Leary and Johnson (1986) have identified four factors associated with item-response characteristics and raters that enhance reliability and validity of such scales: (i) the necessity of using clearly defined reference points on the scale (i.e., estimates of frequency, duration, or intensity), (ii) the inclusion of more than two reference points on the scale (i.e., reference points that quantify the behavior being rated), (iii) a rater who has had extensive opportunities for observing the child being rated, and (iv) more than one rater who has equal familiarity with the child.

The rating forms and checklists described earlier (e.g., Revised Behavior Problem Checklist, Child Behavior Checklist, Behavior Assessment System for Children, the Louisville Fear Survey Schedule for Children, and the Home Situations Questionnaire) incorporate these item and response characteristics and are generally accepted as reliable and valid instruments. For example, the interrater reliability of the Revised Behavior Problem Checklist is quite high when raters are equally familiar with the children being rated and when ratings are provided by raters within the same setting (Quay, 1977; Quay & Peterson, 1983). Further, stability of these ratings has been reported over two-week and one-year intervals. These findings have been reported for teachers in the school setting and parents in the home setting. However, when ratings of teachers are compared to those of parents, interrater reliabilities are considerably lower. While teachers seem to agree with other teachers, and one parent tends to agree with the other parent, there is less agreement between parents and teachers. Such differences may be due to differential perceptions of behavior by parents and teachers or to the situational specificity of behavior, as discussed earlier (also see Achenbach et al., 1987). These findings support the desirability of obtaining information about the child from as many informants and from as many settings as possible.

The validity of the Revised Behavior Problem Checklist has also been demonstrated in numerous ways. It has been shown to distinguish clinic-referred children from nonreferred children, and to be related to psychiatric diagnosis, other measures of behavioral deviance, prognosis, and differential effectiveness of specific treatment strategies (see Ollendick & Cerny, 1981, for a discussion of these findings).

Findings similar to these have been reported for the Child Behavior Checklist, Behavior Assessment System for Children, Louisville Fear Survey Schedule, and the Home Situations Questionnaire. These rating forms and checklists, as well as others, have been shown to possess sound psychometric qualities and to be clinically useful. They not only provide meaningful data about the child's adaptive and problem behaviors but are also useful in orienting parents, teachers, and significant others to specific problem or asset areas and in alerting them to observe and record specific behaviors accurately and validly.

4.06.5.3 Self-report Instruments

Of the various methods used in child behavioral assessment, the self-report method has received the least empirical support, although this picture is rapidly changing. As noted earlier, child behavioral assessors initially eschewed use of self-report instruments, largely on the basis of their suspected low reliability and validity. As we have noted, however, data from self-report instruments can be meaningfully used to understand and describe the child, plan treatment, and evaluate treatment outcome.

As with interview and checklist or rating data, self-report of specific behaviors (including cognitions and affects) and events is more reliable and valid than more general, global reports of life experiences. Such self-reports of specific states can be used to identify discrete components of more general constructs (e.g., determining the exact fears of a phobic child and the exact situations that are associated with withdrawn behavior in an unassertive child). Illustratively, Scherer and Nakamura's (1968) Fear Survey Schedule for Children and its revision. (Ollendick, 1983b) can be used to pinpoint specific fears and classes of fear. Further, this instrument has been shown to be reliable over time, to possess high internal consistency and a meaningful and replicable factor structure, to distinguish between phobic and nonphobic children, and to discriminate among subtypes of phobic youngsters within a particular phobic group (Ollendick & Mayer, 1984; Ollendick, King, & Yule, 1994).

Clearly, more research is needed in this area before routine use of self-report instruments can be endorsed. Nonetheless, instruments that measure specific aspects of behavior such as anxiety or depression rather than global traits hold considerable promise for child behavioral assessment.

4.06.5.4 Self-monitoring

In self-monitoring, children observe their own behavior and then systematically records its occurrence. As with other measures, concerns related to the reliability and validity of this method exist. What is the extent of interobserver agreement between children who are instructed to monitor their own behavior and objective observers? How accurate are children in recording occurrences of behavior? How reactive is the process of self-monitoring?

The literature in this area is voluminous. Even though all necessary studies have not been conducted, the findings are in general agreement. First, children as young as seven years of age can be trained to be reliable and accurate recorders of their own behavior. However, the specific behaviors should be clearly defined, prompts to self-record should be available, and reinforcement for self-monitoring should be provided. Under such conditions, children's recordings closely approximate those obtained from observing adults. For example, in a study examining the effects of self-monitoring and self-administered overcorrection in the treatment of nervous tics in children, Ollendick (1981) showed that 8–10-year-old children who were provided clear prompts to self-record highly discrete behaviors were able to do so reliably. Estimates of occurrence closely paralleled those reported by parents and teachers, even though children were unaware that these adults were recording their nervous tics. In another study, Ackerman and Shapiro (1985) demonstrated the accuracy of self-monitoring by comparing self-recorded data with a permanent product measure (the number of units produced in a work setting). Again, accuracy of self-monitoring was confirmed.

Second, self-monitoring may result in behavior change due to the self-observation process and result in altered estimates of target behaviors. This effect is known as reactivity. Numerous factors have been shown to influence the occurrence of reactivity: specific instructions, motivation, goal-setting, nature of the self-recording device, and the valence of the target behavior (e.g., Nelson, 1977, 1981). Among the more important findings are that desirable behaviors (e.g., study habits, social skills) may increase while undesirable behaviors (e.g., nervous tics, hitting) tend to decrease following self-monitoring, and that the more obtrusive the self-recording device, the greater the behavior change. For example, Nelson, Lipinski, and Boykin (1978) found that hand-held counters produced greater reactivity than belt-worn counters. Holding a counter in one's hand was viewed as more obtrusive, contributing to increased reactivity. Reactivity is a concern in the assessment process because it affects the actual occurrences of behavior. However, if one is aware of the variables that contribute to reactive effects, self-monitoring can be used as a simple and efficient method for data collection (Shapiro, 1984).

In short, self-monitoring has been found to be useful in the assessment of a wide range of child behavior problems across a wide variety of settings. When issues related to the reliability, accuracy, and reactivity of measurement are addressed, self-monitoring represents another clinically useful strategy that is highly efficient and clinically useful.

4.06.5.5 Behavioral Observation

As with other assessment strategies, behavioral observation procedures must possess adequate psychometric qualities and be empirically validated before their routine use can be endorsed. Although early behaviorists accepted the accuracy of behavioral observations based on their face validity, subsequent investigators enumerated a variety of problems associated with their reliability, validity, and clinical utility (e.g., Johnson & Bolstad, 1973; Kazdin, 1977). These problems include the complexity of the observation code, the exact recording procedures to be used (e.g., frequency counts, time sampling, etc.), observer bias, observer drift, and the reactive nature of the observation process itself (see Barton & Ascione, 1984, for further discussion of these issues). Our experience suggests that the greatest threat to the utility of observational data comes from the reactive nature of the observational process itself, especially when the observer is present in the natural setting. It is well known that the presence of an observer affects behavior, usually in socially desirable directions. We have found two strategies to be useful in reducing such reactive effects: recruiting and training observer-coders already present in the natural setting (e.g., a teacher or parent), and if this is not possible, planning extended observations so children can habituate to the observers and so that the effects of reactivity will diminish. However, in the latter instance, it should be noted that several sessions of observations may be required, since reactive effects have been observed to be prolonged (Johnson & Lobitz, 1974). Reactive effects, combined with the aforementioned practical issues of personnel, time, and resources, have led us to place greater emphasis on recruiting observer-coders already in the children's natural environment or training children themselves to record their own behavior.

In brief, behavioral observations are the most direct and least inferential method of assessment. Even though a variety of problems related to their reliability and validity have been commented upon, behavioral observations are highly useful strategies and represent the hallmark of child behavioral assessment. Whenever possible, behavioral observations in the natural setting should be obtained.

4.06.6 FUTURE DIRECTIONS

A number of directions for future research and development in child behavioral assessment may be evident to the reader. What follows is our attempt to highlight those areas that appear most promising and in need of greater articulation.

4.06.6.1 Developmental Factors

First, it seems to us that greater attention must be given to developmental factors as they affect the selection of child behavioral assessment procedures. Although we have argued that these procedures should be developmentally sensitive, child behavioral assessors have frequently not attended to, or have ignored, this recommendation. As we noted earlier, the most distinguishing characteristic of children is developmental change. Such change encompasses basic biological growth and maturity as well as affective, behavioral, and cognitive fluctuations that may characterize children at different age levels. While the importance of accounting for developmental level when assessing behavior may be obvious, ways of integrating developmental concepts and principles into child behavioral assessment are less evident. Edelbrock (1984) has noted three areas for the synthesis of developmental and behavioral principles: (i) use of developmental fluctuations in behavior to establish normative baselines of behavior, (ii) determination of age and gender differences in the expression and covariation of behavioral patterns, and (iii) study of stability and change in behavior over time. Clearly, these areas of synthesis and integration are in their infancy and in need of greater articulation (e.g., Harris & Ferrari, 1983; Ollendick & Hersen, 1983; Rutter & Garmezy, 1983; Sroufe & Rutter, 1984).

Recently, Ollendick and King (1991) addressed this developmental–behavioral synthesis in some detail. In reference to normative data, they suggested that such information could be used to determine which behavior problems represent clinically significant areas of concern, examine appropriateness of referral, and evaluate efficacy of interventions. Essentially, this normative-developmental perspective emphasizes the central importance of change over time and the need for relevant norms against which children can be compared, both at the time of assessment and following treatment. Another way in which developmental principles can be integrated into ongoing child behavioral assessment is to identify age differences in the relations or patterns among behaviors (Edelbrock, 1984). Ollendick and King (1991) have shown such patterning of behavior across development for a number of measures, including diagnostic interviews, self- and other-report instruments, and behavioral coding systems.

Finally, developmental principles can be useful in child behavioral assessment in our attempts to examine and understand continuity and discontinuity of certain behavioral patterns. Basically, this issue can be addressed from two vantage points, a descriptive one and an explanatory one. From a descriptive standpoint, we are interested in determining whether a behavior or set of behaviors seen at one point in time can be described in the same way at another point in time. If it can be described in the same way, descriptive continuity is said to exist; if it cannot, descriptive discontinuity is said to obtain (Lerner, 1986). We are simply asking, does the behavior look the same or different? Does it take the same form over time? For example, if 4-year-old, 8-year-old, and 12-year-old children all emitted the same behaviors to gain entry in a social group, we would conclude that descriptive continuity exists for social entry behavior.

For the most part, it has been shown that the expression and patterning of a large number of behaviors change across development and that descriptive discontinuity is more likely the case (Ollendick & King, 1991). Changes in behavior observed with development can, of course, occur for many different reasons. If the same explanations are used to account for behavior over time, then that behavior is viewed as involving unchanging laws or rules and explanatory continuity is said to exist. However, if different explanations are used to account for changes in behavior over time, explanatory discontinuity prevails (Lerner, 1986). For the most part, behaviorally oriented theorists and clinicians maintain changes over time are due to a set of learning principles that are largely the same across the child's life span. No new principles or laws are needed as the child grows. Developmental theorists, on the other hand, maintain a progressive differentiation of the organism which suggests a different set of principles be invoked across different stages of development. Unfortunately, the evidence on explanatory continuity versus discontinuity is scarce; the jury is out on these issues. Much work remains to be done in this area; however, as Ollendick and King (1991) note, the emergence of "developmental–behavioral assessment" is on the horizon.

4.06.6.2 The Utility of the Multimethod Approach at Different Age Levels

Second, and somewhat related to the first area, greater attention must be focused on the incremental validity of the multimethod approach when used for children of varying ages. Throughout this chapter, we have espoused a multimethod approach consisting of interviews, self- and other-reports, self-monitoring, and behavioral observations. Some of these procedures may be more appropriate at some age levels than others. Further, the psychometric properties of these procedures may vary with age. For example, self-monitoring requires the ability to compare one's own behavior against a standard and accurately to judge occurrence or nonoccurrence of targeted events and behaviors. Most children below six years of age lack the requisite ability to self-monitor and may not profit from such procedures. In fact, the limited research available suggests self-monitoring may be counter-productive when used with young children, resulting in confusion and impaired performance (e.g., Higa, Thorp, & Calkins, 1978). These findings suggest that self-monitoring procedures are better suited for children who possess sufficient cognitive abilities to benefit from their use (Shapiro, 1984). In a similar vein, age-related variables place constraints on use of certain self-report and sociometric measures with young children. It has often been noted that sociometric devices must be simplified and presented in pictorial form to children under six years of age (Hops & Lewin, 1984). The picture-form sociometric device provides young children with a set of visual cues regarding children to be rated and, of course, does not require them to read names of children being rated. The roster-and-rating method, used so frequently with older children, is simply not appropriate with younger children. Ollendick and Hersen (1993) review additional age-related findings for other procedures and suggest caution in using these procedures without due regard for their developmental appropriateness and related psychometric properties.

If certain procedures are found to be less reliable or valid at different age levels, their indiscriminate use with children can not be endorsed. Inasmuch as these strategies are found to be inadequate, the combination of them in a multimethod approach would serve only to compound their inherent limitations (Mash & Terdal, 1981). The sine qua non of child behavioral assessment is that the procedures be empirically validated. In addition, the different procedures might vary in terms of their treatment utility across different ages. Treatment utility refers to the degree to which assessment strategies are shown to contribute to beneficial treatment outcomes (Hayes et al., 1987). More specifically, treatment utility addresses issues related to the selection of specific target behaviors and to the choice of specific assessment strategies. For example, we might wish to examine the treatment utility of using self-report questionnaires to guide treatment planning, above and beyond that provided by direct behavioral observation of children who are phobic of social encounters. All children could complete a fear schedule and be observed in a social situation, but the self-report data for only half of the children would be made available for treatment planning. If the children for whom self-reports were made available improved more than those whose treatment plans were based solely on behavioral observations, then the treatment utility of using self-report data would be established (for this problem with this age child). In a similar fashion, the treatment utility of interviews, role plays, and other devices could be evaluated (Hayes et al., 1987). Of course, it would be important to examine treatment utility from a developmental perspective as well. Certain procedures might be shown to possess incremental validity at one age but not another. Although the concept of treatment utility is relatively new, it shows considerable promise as a strategy to evaluate the incremental validity of our multimethod assessment approach. We should not necessarily assume that "more" assessment is "better" assessment.

4.06.6.3 Cultural Sensitivity

Considerable energy must be directed to the development of child behavioral assessment methods that are culturally sensitive. Numerous observers have called attention to the internationalization of the world and the "browning of America" (e.g., Malgady, Rogler, & Constantino, 1987; Vasquez Nuttall, DeLeon, & Del Valle, 1990). In reference to this chapter, these developments suggest that the assessment process is increasingly being applied to non-Caucasian children for whom English is not the primary language in America, and that many procedures developed in America and other Western countries are being applied, sometimes indiscriminately, in other countries as well. Development of assessment procedures that are culture-fair (and language-fair) is of utmost importance. Of course, many cultural issues need to be considered in the assessment process. Cultural differences may be expressed in child-rearing practices, family values, parental expectations, communication styles, nonverbal communication patterns, and family structure and dynamics (Vasquez et al., 1990). As an example, behaviors characteristic of ethnic minority children may be seen as emotionally or behaviorally maladaptive by persons who have little or no appreciation for cultural norms (e.g., Prewitt-Diaz, 1989). Thus, cultural differences (biases?) are likely to occur early in the assessment process. Fortunately, Vasquez-Nuttall, Sanchez, Borras Osorio, Nuttall, & Varvogil (1996) have suggested several steps that can be taken to minimize cultural biases in the assessment process.

Vasquez et al. (1996) have offered the following suggestions: (i) include extended family members in the information-gathering process; (ii) use interpreters, if necessary, in interviewing the child and family members; (iii) familiarize oneself with the culture of specific groups; and (iv) use instruments that have been translated into the native language of the children and for which norms are available for specific ethnic groups. With regard to this latter recommendation, significantly greater progress has been witnessed for the translation component than the establishment of well-standardized normative information. For example, while the Conners' Parent Rating Scales (Conners, 1985) and Conners' Teacher Rating Scale (Conners, 1985) have been translated into Spanish and other languages, group norms are lacking and the reliability and validity of the translations have not been examined systematically. Similarly, the Fear Survey Schedule for Children-Revised (Ollendick, 1983b) has been translated into over 10 languages, yet normative data are lacking and the psychometric properties of the instrument have not been fully explored or established. In sum, a clear challenge before us in the years ahead is to attend to important cultural factors that impinge on our assessment armamentarium, and to develop and promulgate culturally sensitive methods that are developmentally appropriate and empirically validated.

4.06.6.4 Measures of Cognitive and Affective Processes

More effort must be directed toward the development of culturally relevant, developmentally sensitive, and empirically validated procedures for assessment of cognitive and affective processes in children. In recent years, child behavioral assessors have become increasingly interested in the relation of children's cognitive and affective processes to observed behaviors. The need for assessment in this area is further evidenced by the continued increase of cognitive-behavioral treatment procedures with children, a trend first observed in the late 1970s and early 1980s (e.g., Kendall, Pellegrini, & Urbain, 1981; Meador & Ollendick, 1984). As noted by Kendall et al. (1981), there is a particularly pressing need to develop procedures that can examine the very cognitions and processes that are targeted for change in these intervention efforts. For example, the reliable and valid assessment of self-statements made by children in specific situations would facilitate the empirical evaluation of cognitive-behavioral procedures such as self-instructional training and cognitive restructuring (cf. Zatz & Chassin, 1983; Stefanek, Ollendick, Baldock, Francis, & Yaeger, 1987).

4.06.6.5 The Role of the Child

We must concentrate additional effort on the role of the child in child behavioral assessment. All too frequently, "tests are administered to children, ratings are obtained on children, and behaviors are observed in children" (Ollendick & Hersen, 1984, p. ix). This process views the child as a passive responder, someone who is largely incapable of actively shaping and determining behaviors of clinical relevance. Although examination of these organismic variables is only beginning, it would appear that concerted and systematic effort must be directed to their description and articulation. For example, children's conceptions of their own behavior is a critical area of further study. To what causes do children attribute aggressive or withdrawn behavior in themselves or in their peers? Are there aggregated trends in these attributions? Do they differ by culture? Do causal attributions (as well as self-efficacy and outcome expectancies) mediate treatment outcomes? Again, are there age-related effects for these effects or culturally relevant effects? The answers to these questions are of both theoretical interest and applied clinical significance.

The process described above also implies that child behavior (problematic or otherwise) occurs in a vacuum, and that the perceptions and behaviors of referral sources (parents, teachers) and characteristics of the environments in which behavior occurs are somehow less critical to assess. Recent efforts to develop reliable methods for assessing parent–child interactions are indicative of an increased awareness of the need to broaden the scope of assessment to include specific individuals with whom, and environments in which, child behavior problems commonly occur (cf. Dadds et al., 1994; Dumas, 1989; Greene, 1995, 1996; Ollendick, 1996). However, much additional work remains to be done in this area.

4.06.6.6 Ethical Guidelines

Finally, we must continue to focus our attention on ethical issues in child behavioral assessment. A number of ethical issues regarding children's rights, proper and legal consent, professional judgment, and social values are raised in the routine practice of child behavioral assessment (Rekers, 1984). Are children capable of granting full and proper consent to a behavioral assessment procedure? At what age and in what cultures are children competent to give such consent? Is informed consent necessary? Or might not informed consent be impossible, impractical, or countertherapeutic in some situations? What ethical guidelines surround the assessment procedures to be used? Current professional guidelines suggest our procedures should be reliable, valid, and clinically useful. Do the procedures suggested in this chapter meet these professional guidelines? What are the rights of parents and of society? It should be evident from these questions that a variety of ethical issues persists. Striking a balance between the rights of parents, society, and children is no easy matter but is one that takes on added importance in the increasingly litigious society of the USA.

In short, future directions of child behavioral assessment are numerous and varied. Even though a technology for child behavioral assessment has evolved and is in force, we need to begin to explore the issues raised before we can conclude the procedures are maximally productive and in the best interests of children throughout the world.

4.06.7 SUMMARY

Child behavioral assessment strategies have been slow to evolve. Only recently has the chasm between child behavior therapy and child behavioral assessment been narrowed. Increased awareness of the importance of developing assessment procedures that provide an adequate representation of child behavior disorders has spurred research into assessment procedures and spawned a plethora of child behavioral assessment techniques. The growing sophistication of child behavior assessment is witnessed by the appearance of self- and other-report strategies that are beginning to take into account developmental, social, and cultural influences as well as cognitive and affective mediators of overt behavior. At the same time, attention to psychometric properties of assessment procedures has continued.

Certain theoretical assumptions guide child behavioral assessment. Foremost among these is the premise that behavior is a function of situational determinants and not a sign of underlying personality traits. To assess adequately the situational determinants and to obtain as complete a picture of the child as is possible, a multimethod assessment approach is recommended, utilizing both direct and indirect methods of assessment. Direct methods include self-monitoring as well as behavioral observation by trained observers in naturalistic or simulated analogue settings. Indirect measures include behavioral interviewing and self- and other-report measures. These sources of information are considered indirect ones because they involve retrospective reports of previous behavior.

Even though direct behavioral observation remains the hallmark of child behavioral assessment, information from these other sources is considered not only valuable but integral in the understanding and subsequent treatment of child behavior disorders. Hence, whereas identification and specification of discrete target behaviors were once considered sufficient, current child behavioral assessment involves serious consideration and systematic assessment of cognitive and affective aspects of the child's behavior and of developmental, social, and cultural factors that influence the child, as well as direct observation of the problematic behavior *in situ*.

Several areas of future research remain. These include clearer specification of developmental variables, a closer examination of the utility of the multimethod approach at different age levels, the influence of culture and the need for models of assessment that take cultural forces into consideration, development of specific measures to examine cognitive and affective processes in children, articulation of the role of the child in child behavioral assessment, and continued development of ethical guidelines. While the basis for a technology of child behavioral assessment exists, considerable fine-tuning remains to be done. Child behavioral assessment is at a critical crossroad in its own development; which path it takes will determine its long-term future.

4.06.8 REFERENCES

Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: A factor-analytic study. *Psychological Monographs, 80,* 1–37.

Achenbach, T. M. (1991a). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, VT: University of Vermont Department of Psychiatry.

Achenbach, T. M. (1991b). *Manual for the Teacher Report Form and 1991 Profile*. Burlington, VT: University of Vermont Department of Psychiatry.

Achenbach, T. M. (1991c). *Manual for the Youth Self-Report and 1991 Profile*. Burlington, VT: University of Vermont Department of Psychiatry.

Achenbach, T. M., & Edelbrock, C. S. (1989). Diagnostic, taxonomic, and assessment issues. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of child psychopathology* (2nd ed., pp. 53–69). New York: Plenum.

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101,* 213–232.

Ackerman, A. M., & Shapiro. E. S. (1985). Self-monitoring and work productivity with mentally retarded adults. *Journal of Applied Behavior Analysis, 17,* 403–407.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Ayllon, T., Smith, D., & Rogers, M. (1970). Behavioral management of school phobia. *Journal of Behavior Therapy and Experimental Psychiatry, 1,* 125–138.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84,* 191–215.

Barkley, R. A. (1981). *Hyperactive children: A handbook for diagnosis and treatment*. New York: Guilford Press.

Barkley, R. A. (1987). *Defiant children: A clinician's manual for parent training*. New York: Guilford Press.

Barkley, R. A., & Edelbrock, C. S. (1987). Assessing situational variation in children's behavior problems: The home and school situations questionnaires. In R. Prinz (Ed.), *Advances in behavioral assessment of children and families* (Vol. 3, pp. 157–176). Greenwich, CT: JAI Press.

Barkley, R. A., Karlsson, I., Strzelecki, E., & Murphy, J. (1984). Effects of age and Ritalin dosage on the mother–child interactions of hyperactive children. *Journal of Consulting and Clinical Psychology, 52,* 750–758.

Barrios, B., & Hartmann, D. P. (1986). The contributions of traditional assessment: Concepts, issues, and methodologies. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 81–110). New York: Guilford Press.

Barton, E. J., & Ascione, F. R. (1984). Direct observations. In T. H. Ollendick & M. Hersen (Eds.), *Child behavioral assessment: Principles and procedures* (pp. 166–194). New York: Pergamon.

Bem, D. I., & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review, 81,* 506–520.

Bornstein, P. H., Bornstein, M. T., & Dawson, B. (1984). Integrated assessment and treatment. In T. H. Ollendick & M. Hersen (Eds.), *Child behavioral assessment: Principles and procedures* (pp. 223–243). New York: Pergamon.

Campbell, S. B. (1989). Developmental perspectives in child psychopathology. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of child psychopathology* (2nd ed., pp. 5–28). New York: Plenum.

Cantwell, D. P. (1983). Childhood depression: A review of current research. In B. B. Lahey & A. E. Kazdin (Eds.), *Advances in clinical child psychology* (Vol. 5, pp. 39–93). New York: Plenum.

Chess, S., Thomas, A., & Birch, H. G. (1966). Distortions in developmental reporting made by parents of behaviorally disturbed children. *Journal of the American Academy of Child Psychiatry, 5*, 226–231.

Ciminero, A. R., & Drabman, R. S. (1977). Current developments in the behavioral assessment of children. In B. B. Lahey & A. E. Kazdin (Eds.), *Advances in clinical child psychology* (Vol. I, pp. 47–82). New York: Plenum.

Cone, J. D. (1977). The relevance of reliability and validity for behavioral assessment. *Behavior Therapy, 8*, 411–426.

Cone, J. D. (1978). The behavioral assessment grid (BAG): A conceptual framework and taxonomy. *Behavior Therapy, 9*, 882–888.

Cone, J. D. (1981). Psychometric considerations. In M. Hersen & A. S. Bellack (Eds.), *Behavioral assessment: A practical handbook* (2nd ed., pp. 38–68). Elmsford, NY: Pergamon.

Cone, J. D. (1986). Ideographic, nomothetic, and related perspectives in behavioral assessment. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 111–128). New York: Guilford Press.

Cone, J. D., & Hawkins, R. P. (Eds.) (1977). *Behavioral assessment: New directions in clinical psychology*. New York: Brunner/Mazel.

Conners, C. K. (1985). *The Conners rating scales: Instruments for the assessment of childhood psychopathology*. Unpublished manuscript, Children's Hospital National Medical Center, Washington, DC.

Cowen, E. L., Pederson, A., Barbigian, H., Izzo, L. D., & Trost, M. A. (1973). Long-term follow-up of early-detected vulnerable children. *Journal of Consulting and Clinical Psychology, 41*, 438–445.

Cronbach, L. J. (1960). *Essentials of psychological testing*. New York: Harper & Row.

Dadds, M. R., Rapee, R. M., & Barrett, P. M. (1994). Behavioral observation. In T. H. Ollendick, N. J. King, & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in children and adolescents* (pp. 349–364). New York: Plenum.

Deluty, R. H. (1979). Children's Action Tendency Scale: A self-report measure of aggressiveness, assertiveness, and submissiveness in children. *Journal of Consulting and Clinical Psychology, 41*, 1061–1071.

Dong, Q., Yang, B., & Ollendick, T. H. (1994). Fears in Chinese children and adolescents and their relations to anxiety and depression. *Journal of Child Psychology and Psychiatry, 35*, 351–363.

Doyle, A., Ostrander, R., Skare, S., Crosby, R. D., & August, G. J. (1997). Convergent and criterion-related validity of the Behavior Assessment System for Children–Parent Rating Scale. *Journal of Clinical Child Psychology, 26*, 276–284.

Dumas, J. E. (1989). Interact: A computer-based coding and data management system to assess family interactions. In R. J. Prinz (Ed.), *Advances in behavioral assessment of children and families* (Vol. 3, pp. 177–202). Greenwich, CT: JAI Press.

Edelbrock, C. S. (1984). Developmental considerations. In T. H. Ollendick & M. Hersen (Eds.), *Child behavioral assessment: Principles and procedures* (pp. 20–37). Elmsford, NY: Pergamon.

Evans, I. M., & Nelson, R. O. (1977). Assessment of child behavior problems. In A. R. Ciminero, K. S. Calhoun, & H. E. Adams (Eds.), *Handbook of behavioral assessment* (pp. 603–681). New York: Wiley-Interscience.

Finch, A. J., Nelson, W. M., III, & Moss, J. H. (1983). Stress innoculation for anger control in aggressive children. In A. J. Finch, W. M. Nelson, & E. S. Ott (Eds.), Cognitive-behavioral procedures with children: A practical guide (pp. 148–205). Newton, MA: Allyn & Bacon.

Finch, A. J., & Rogers, T. R. (1984). Self-report instruments. In T. H. Ollendick & M. Hersen (Eds.), *Child behavioral assessment: Principles and procedures* (pp. 106–123). Elmsford, NY: Pergamon.

Goldfried, M. R., & Kent, R. N. (1972). Traditional versus behavioral personality assessment: A comparison of methodological and theoretical assumptions. *Psychological Bulletin, 77,* 409–420.

Graham, P., & Rutter, M. (1968). The reliability and validity of the psychiatric assessment of the child—II. Interview with the parents. *British Journal of Psychiatry, 114,* 581–592.

Greene, R. W. (1995). Students with ADHD in school classrooms: Teacher factors related to compatibility, assessment, and intervention. *School Psychology Review, 24,* 81–93.

Greene, R. W. (1996). Students with ADHD and their teachers: Implications of a goodness-of-fit perspective. In T. H. Ollendick & R. J. Prinz (Eds.), *Advances in clinical child psychology* (Vol. 18, pp. 205–230). New York: Plenum.

Greene, R. W., & Ollendick, T. H. (in press). Behavioral assessment of children. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment* (3rd ed.). Boston: Allyn & Bacon.

Gresham, F. M. (1982). Social interactions as predictors of children's likability and friendship patterns: A multiple regression analysis. *Journal of Behavioral Assessment, 4,* 39–54.

Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating system manual.* Circle Pines, MN: American Guidance Service.

Gross, A. M. (1984). Behavioral interviewing. In T. H. Ollendick & M. Hersen (Eds.), *Child behavioral assessment: Principles and procedures* (pp. 61–79). Elmsford, NY: Pergamon.

Harris, S. L., & Ferrari, M. (1983). Developmental factors in child behavior therapy. *Behavior Therapy, 14,* 54–72 .

Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1986). Evaluating the quality of behavioral assessment. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 463–503). New York: Guilford.

Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1987) The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist, 42,* 963–974.

Herjanic, B., Herjanic, M., Brown, F., & Wheatt, T. (1973). Are children reliable reporters? *Journal of Abnormal Child Psychology, 3,* 41–48.

Higa, W. R., Tharp, R. G., & Calkins, R. P. (1978). Developmental verbal control of behavior: Implications for self-instructional testing. *Journal of Experimental Child Psychology, 26,* 489–497.

Holmes, F. B. (1936). An experimental investigation of a method of overcoming children's fears. *Child Development, 1,* 6–30.

Hops, H., & Lewin, L. (1984). Peer sociometric forms. In T. H. Ollendick & M. Hersen (Eds.), *Child behavioral assessment: Principles and procedures* (pp. 124–147). New York: Pergamon.

Johnson, S. M., & Bolstad, O. D. (1973). Methodological issues in naturalistic observations: Some problems and solutions for field research. In L. A. Hammerlynck, L. C. Handyx, & E. J. Mash (Eds.), *Behavior change: Methodology, concepts, and practice* (pp. 7–67). Champaign, IL: Research Press.

Johnson, S. M., & Lobitz, G. K. (1974). Parental manipulation of child behavior in home observations. *Journal of Applied Behavior Analysis, 1,* 23–31.

Jones, M. C. (1924). The elimination of children's fears. *Journal of Experimental Psychology, 7,* 382–390.

Jones, R. R., Reid, J. B., & Patterson, G. R. (1975). Naturalistic observation in clinical assessment. In P. McReynolds (Ed.), *Advances in psychological assessment* (Vol. 3, pp. 42–95). San Francisco: Jossey-Bass.

Kanfer, F. H., & Phillips, J. S. (1970). *Learning foundations of behavior therapy.* New York: Wiley.

Kazdin, A. E. (1977). Artifact, bias, and complexity of assessment: The ABCs of reliability. *Journal of Applied Behavior Analysis, 4,* 7–14.

Kendall, P. C., & Hollon, S. D. (Eds.) (1980). *Cognitive-behavioral intervention: Assessment methods.* New York: Academic Press.

Kendall, P. C., Pellegrini, D. S., & Urbain, E. S. (1981). Approaches to assessment for cognitive-behavioral interventions with children. In P. C. Kendall & S. D. Hollon (Eds.), *Assessment strategies for cognitive-behavioral interventions* (pp. 227–286). New York: Academic Press.

Kovacs, M. (1985). Children's Depression Inventory (CDI). *Psychopharmacology Bulletin, 21,* 995–998.

Kunzelman, H. D. (Ed.) (1970). *Precision teaching.* Seattle, WA: Special Child Publications.

Lease, C. A., & Ollendick, T. H. (1993). Development and psychopathology. In A. S. Bellack & M. Hersen (Eds.), *Psychopathology in adulthood* (pp. 89–102). Boston: Allyn & Bacon.

Lerner, R. M. (1986). *Concepts and theories of human development* (2nd ed.). New York: Random House.

Linehan, M. (1977). Issues in behavioral interviewing. In J. D. Cone & R. P. Hawkins (Eds.), *Behavioral assessment: New directions in clinical psychology* (pp. 30–51). New York: Brunner/Mazel.

Malgady, R., Rogler, L., & Constantino, G. (1987). Ethnocultural and linguistic bias in mental health evaluation of Hispanics. *American Psychologist, 42,* 228–234.

Mash, E. J., & Terdal, L. G. (1981). Behavioral assessment of childhood disturbance. In E. J. Mash & L. G. Terdal (Eds.), *Behavioral assessment of childhood disorders* (pp. 3–76). New York: Guilford Press.

Mash, E. J., & Terdal, L. G. (Eds.) (1989). *Behavioral assessment of childhood disorders* (2nd ed.). New York: Guilford Press.

Mash, E. J., & Terdal, L. G. (1989). Behavioral assessment of childhood disturbance. In E. J. Mash & L. G. Terdal (Eds.), *Behavioral assessment of childhood disorders* (2nd ed., pp. 3–65). New York: Guilford Press.

Matson, J. L., & Ollendick, T. H. (1976). Elimination of low-frequency biting. *Behavior Therapy, 7,* 410–412.

McConaughy, S. H. (1996). The interview process. In M. J. Breen & C. R. Fiedler (Eds.), *Behavioral approach to assessment of youth with emotional/behavioral disorders: A handbook for school-based practitioners* (pp. 181–224). Austin, TX: ProEd.

McMahon, R. J. (1984). Behavioral checklists and rating forms. In T. H. Ollendick & M. Hersen (Eds.), *Child behavioral assessment: Principles and procedures* (pp. 80–105). Elmsford, NY: Pergamon.

Meador, A. E., & Ollendick, T. H. (1984). Cognitive behavior therapy with children: An evaluation of its efficacy and clinical utility. *Child and Family Behavior Therapy, 6,* 25–44.

Meichenbaum, D. H. (1977). *Cognitive-behavior modification.* New York: Plenum.

Miller, L. C., Barrett, C. L., Hampe, E., & Noble, H. (1972). Comparison of reciprocal inhibition, psychotherapy, and waiting list control for phobic children. *Journal of Abnormal Psychology, 79,* 269–279.

Mischel, W. (1968). *Personality and assessment.* New York: Wiley.

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review, 80,* 252–283.

Nelson, R. O. (1977). Methodological issues in assessment via self-monitoring. In J. D. Cone & R. P. Hawkins (Eds.), *Behavioral assessment: New directions in clinical psychology* (pp. 217–240). New York: Brunner/Mazel.

Nelson, R. O. (1981). Theoretical explanations for self-monitoring. *Behavior Modification, 5,* 3–14.

Nelson, R. O., Lipinski. D. P., & Boykin, R. A. (1978). The effects of self-recorder training and the obtrusiveness of the self-recording device on the accuracy and reactivity of self-monitoring. *Behavior Therapy, 9,* 200–208.

Nelson, W. M., III, & Finch, A. J., Jr. (1978). *The new children's inventory of anger*. Unpublished manuscript, Xavier University, OH.

Novick, J., Rosenfeld, E., Bloch, D. A., & Dawson, D. (1966). Ascertaining deviant behavior in children. *Journal of Consulting and Clinical Psychology, 30,* 230–238.

O'Leary, K. D., & Johnson, S. B. (1986). Assessment and assessment of change. In H. C. Quay & J. S. Werry (Eds.), *Psychopathological disorders of children* (3rd ed., pp. 423–454). New York: Wiley.

O'Leary, K. D., Romanczyk, R. G., Kass, R. E., Dietz, A., & Santogrossi, D. (1971). *Procedures for classroom observations of teachers and parents*. Unpublished manuscript, State University of New York at Stony Brook.

Ollendick, T. H. (1981). Self-monitoring and self-administered overcorrection: The modification of nervous tics in children. *Behavior Modification, 5,* 75–84.

Ollendick, T. H. (1983a). Development and validation of the Children's Assertiveness Inventory. *Child and Family Behavior Therapy, 5,* 1–15.

Ollendick, T. H. (1983b). Reliability and validity of the Revised-Fear Survey Schedule for Children (FSSC-R). *Behaviour Research and Therapy, 21,* 685–692.

Ollendick, T. H. (1995). Cognitive-behavioral treatment of panic disorder with agoraphobia in adolescents: A multiple baseline design analysis. *Behavior Therapy, 26,* 517–531.

Ollendick, T. H. (1996). Violence in society: Where do we go from here? (Presidential address). *Behavior Therapy, 27,* 485–514.

Ollendick, T. H., & Cerny, J. A. (1981). *Clinical behavior therapy with children.* New York: Plenum.

Ollendick, T. H., & Greene, R. W. (1990). Behavioral assessment of children. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment* (2nd ed., pp. 403–422). Elmsford, NY: Pergamon.

Ollendick, T. H., & Gruen, G. E. (1972). Treatment of a bodily injury phobia with implosive therapy. *Journal of Consulting and Clinical Psychology, 38,* 389–393.

Ollendick, T. H., & Hersen, M. (Eds.) (1983). *Handbook of child psychopathology*. New York: Plenum.

Ollendick, T. H., & Hersen, M. (Eds.) (1984). *Child behavioral assessment: Principles and procedures.* New York: Pergamon.

Ollendick, T. H., & Hersen, M. (1993). Child and adolescent behavioral assessment. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of child and adolescent behavioral assessment* (pp. 3–14). New York: Pergamon.

Ollendick, T. H., & King, N. J. (1991). Developmental factors in child behavioral assessment. In P. R. Martin (Ed.), *Handbook of behavior therapy and psychological science: An integrative approach* (pp. 57–72). New York: Pergamon.

Ollendick, T. H., & King, N. J. (1994). Assessment and treatment of internalizing problems: The role of longitudinal data. *Journal of Consulting and Clinical Psychology, 62,* 918–927.

Ollendick, T. H., King, N. J., & Frary, R. B. (1989). Fears in children and adolescents in Australia and the United States. *Behaviour Research and Therapy, 27,* 19–26.

Ollendick, T. H., King, N. J., & Yule, W. (Eds.) (1994). *International handbook of phobic and anxiety disorders in children*. Boston: Allyn & Bacon.

Ollendick, T. H., Matson, J. L., & Hetsel, W. J. (1985). Fears in children and adolescents: Normative data. *Behaviour Research and Therapy, 23,* 465–467.

Ollendick, T. H., & Mayer, J. (1984). School phobia. In S. M. Turner (Ed.), *Behavioral treatment of anxiety disorders* (pp. 367–411). New York: Plenum.

Ollendick, T. H., & Ollendick, D. G. (1997). General worry and anxiety in children. *In Session: Psychotherapy in Practice, 3,* 89–102.

Ollendick, T. H., Yang, B., King, N. J., Dong, Q., & Akande, A. (1996). Fears in American, Australian, Chinese, and Nigerian children and adolescents: A cross-cultural study. *Journal of Child Psychology and Psychiatry, 37,* 213–220.

Ollendick, T. H., Yule, W., & Ollier, K. (1991). Fears in British children and their relation to manifest anxiety and depression. *Journal of Child Psychology and Psychiatry, 32,* 321–331.

Olweus, D. (1979). Stability of aggressive reaction patterns in males: A review. *Psychological Bulletin, 86,* 852–875.

Patterson, G. R. (1976). The aggressive child: Victim and architect of a coercive system. In E. J. Mash, L. A. Hammerlynck, & L. C. Hardy (Eds.), *Behavior modification and families* (pp. 267–316). New York: Brunner/Mazel.

Patterson, G. R. (1982). *Coercive family process.* Eugene, OR: Castalia.

Patterson, G. R., Ray, R. S., Shaw, D. A., & Cobb, J. A. (1969). *Manual for coding family interaction* (6th ed.). Unpublished manuscript, University of Oregon.

Peterson, D. R. (1961). Behavior problems of middle childhood. *Journal of Clinical and Consulting Psychology, 25,* 205–209.

Pollard, S., Ward, E., & Barkley, R. A. (1983). The effects of parent training and Ritalin on the parent–child interactions of hyperactive boys. *Child and Family Behavior Therapy, 5,* 51–69.

Prewitt-Diaz, J. (1989). *The process and procedures for identifying exceptional language minority children.* State College, PA: Pennsylvania State University.

Prinz, R. (Ed.) (1986). *Advances in behavioral assessment of children and families.* Greenwich, CT: JAI Press.

Quay, H. C. (1977). Measuring dimensions of deviant behavior: The Behavior Problem Checklist. *Journal of Abnormal Child Psychology, 5,* 277–287.

Quay, H. C., & Peterson, D. R. (1967). *Manual for the Behavior Problem Checklist.* Champaign, IL: University of Illinois.

Quay, H. C., & Peterson, D. R. (1975). *Manual for the Behavior Problem Checklist.* Unpublished manuscript .

Quay, H. C., & Peterson, D. R. (1983). *Interim manual for the Revised Behavior Problem Checklist.* Unpublished manuscript, University of Miami.

Rekers, G. A. (1984). Ethical issues in child behavioral assessment. In T. H. Ollendick & M. Hersen (Eds.), *Child behavioral assessment: Principles and procedures* (pp. 244–262). Elmsford, NY: Pergamon.

Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior assessment system for children.* Circle Pines, MN: American Guidance Service.

Reynolds, C. R. & Richmond, B. O. (1985). *Revised children's manifest anxiety scale manual.* Los Angeles: Western Psychological Services.

Rutter, M. (1986). The developmental psychopathology of depression: Issues and perspectives. In M. Rutter, C. E. Izard, & P. B. Read (Eds.), *Depression in young people: Clinical and developmental perspectives* (pp. 3–30). New York: Guilford Press.

Rutter, M., & Garmezy, N. (1983). Developmental psychopathology. In E. M. Hetherington (Ed.), *Socialization, personality, and social development: Vol 14. Mussen's Handbook of child psychology* (pp. 775–911). New York: Wiley.

Scherer, M. W., & Nakamura, C. Y. (1968). A fear survey schedule for children (FSS-FC): A factor-analytic comparison with manifest anxiety (CMAS). *Behaviour Research and Therapy, 6,* 173–182.

Schopler, E. (1974). Changes of direction with psychiatric children. In A. Davids (Ed.), *Child personality and* *psychopathology: Current topics* (Vol. I, pp. 205–236). New York: Wiley.

Shaffer, D. (1992). *NIMH diagnostic interview schedule for children, Version 2.3.* New York: Columbia University Division of Child & Adolescent Psychiatry.

Shapiro, E. S. (1984). Self-monitoring. In T. H. Ollendick & M. Hersen (Eds.), *Child behavioral assessment: Principles and procedures* (pp. 148–165). Elmsford, NY: Pergamon.

Shapiro, E. S., McGonigle, J. J., & Ollendick, T. H. (1980). An analysis of self-assessment and self-reinforcement in a self-managed token economy with mentally retarded children. *Journal of Applied Research in Mental Retardation, 1,* 227–240.

Silverman, W. K., & Albano, A. M. (1996). *Anxiety Disorders Interview Schedule for DSM-IV.* San Antonio, TX: The Psychological Corporation.

Silverman, W. K., & Nelles, W. B. (1988). The anxiety disorders interview schedule for children. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 772–778.

Skinner, B. F. (1953). *Science and human behavior.* New York: Macmillan.

Smith, R. E., & Sharpe, T. M. (1970). Treatment of a school phobia with implosive therapy. *Journal of Consulting and Clinical Psychology, 35,* 239–243.

Smucker, M. R., Craighead, W. E., Craighead, L. W., & Green, B. J. (1986). Normative and reliability data for the Children's Depression Inventory. *Journal of Abnormal Child Psychology, 14,* 25–39.

Spielberger, C. D. (1973). *Preliminary manual for the State–Trait Anxiety Inventory for Children ("how I feel questionnaire").* Palo Alto, CA: Consulting Psychologist Press.

Sroufe, L. A., & Rutter, M. (1984). The domain of developmental psychopathology. *Child Development, 55,* 17–29.

Staats, A. W. (1975). *Social behaviorism.* Homewood, IL: Dorsey Press.

Staats, A. W. (1986). Behaviorism with a personality: The paradigmatic behavioral assessment approach. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 242–296). New York: Guilford Press.

Stefanek, M. E., Ollendick, T. H., Baldock, W. P., Francis, G., & Yaeger, N. J. (1987). Self-statements in aggressive, withdrawn, and popular children. *Cognitive Therapy and Research, 11,* 229–239.

Swann, G. E., & MacDonald, M. L. (1978). Behavior therapy in practice: A rational survey of behavior therapists. *Behavior Therapy, 9,* 799–807.

Ullmann, L. P., & Krasner, L. (Eds.) (1965). *Case studies in behavior modification.* New York: Holt, Rinehart, & Winston.

Vasquez Nuttall, E., DeLeon, B., & Del Valle, M. (1990). Best practice in considering cultural factors. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology II,* (pp. 219–233). Washington, DC: National Association of School Psychologists.

Vasquez Nuttall, E., Sanchez, W., Borras Osorio, L., Nuttall, R. L., & Varvogil, L. (1996). Assessing the culturally and linguistically different child with emotional and behavioral problems. In M. J. Breen & C. R. Fiedler (Eds.), *Behavioral approach to assessment of youth with emotional/behavioral disorders: A handbook for school-based practitioners* (pp. 451–502). Austin, TX: ProEd.

Wahler, R. G. (1976). Deviant child behavior in the family: Developmental speculations and behavior change strategies. In H. Leitenberg (Ed.), *Handbook of behavior modification and behavior therapy* (pp. 516–543). Englewood Cliffs, NJ: Prentice-Hall.

Wahler, R. G., House, A. E., & Stambaugh, E. E. (1976). *Ecological assessment of child problem behavior: A clinical package for home, school, and institutional settings.* Elmsford, NY: Pergamon.

Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3,* 1–14.

Winett, R. A., Riley, A. W., King, A. C., & Altman, D. G. (1989). Preventive strategies with children and families. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of child psychopathology* (2nd ed., pp. 499–521). New York: Plenum.

World Health Organization (1991). *International classification of mental and behavioral disorders: Clinical descriptions and diagnostic guidelines* (10th ed.). Geneva, Switzerland: Author.

Zatz, S., & Chassin, L. (1983). Cognitions of test-anxious children. *Journal of Consulting and Clinical Psychology, 51,* 526–534.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.07
Principles and Practices of Behavioral Assessment with Adults

STEPHEN N. HAYNES
University of Hawaii at Manoa, Honolulu, HI, USA

4.07.1 INTRODUCTION 158

4.07.2 CLINICAL JUDGMENTS AND FUNCTIONAL ANALYSIS IN BEHAVIORAL ASSESSMENT 158

4.07.3 CONCEPTUAL FOUNDATIONS OF BEHAVIORAL ASSESSMENT 162
- *4.07.3.1 Assumptions About the Causes of Behavior Disorders* 162
  - *4.07.3.1.1 Multiple causality* 163
  - *4.07.3.1.2 Multiple causal paths* 163
  - *4.07.3.1.3 Individual differences in causal variables and paths* 163
  - *4.07.3.1.4 Environmental causality and reciprocal determinism* 163
  - *4.07.3.1.5 Contemporaneous causal variables* 164
  - *4.07.3.1.6 Interactive and additive causality* 165
  - *4.07.3.1.7 Situations, setting events, and systems factors as causal variables* 165
  - *4.07.3.1.8 Dynamic causal relationships* 166
  - *4.07.3.1.9 Additional assessment implications of causal assumptions* 166
- *4.07.3.2 Assumptions About the Characteristics of Behavior Problems* 166
  - *4.07.3.2.1 Behavior problems can have multiple response modes* 166
  - *4.07.3.2.2 Behavior problems have multiple parameters* 167
  - *4.07.3.2.3 Client can have multiple behavior problems* 168
  - *4.07.3.2.4 Behavior problems are conditional* 168
  - *4.07.3.2.5 The dynamic nature of behavior problems* 168

4.07.4 METHODOLOGICAL FOUNDATIONS OF BEHAVIORAL ASSESSMENT 169
- *4.07.4.1 An Empirically Based Hypothesis Testing Approach to Assessment* 169
- *4.07.4.2 An Individualized Approach to Assessment* 170
- *4.07.4.3 Time-series Assessment Strategies* 170
- *4.07.4.4 Quantitative and Qualitative Approaches to Behavioral Assessment* 171

4.07.5 BEHAVIORAL ASSESSMENT METHODS 171
- *4.07.5.1 Behavioral Observation* 172
  - *4.07.5.1.1 Behavioral observation in the natural environment* 173
  - *4.07.5.1.2 Analogue observation* 175
- *4.07.5.2 Self-monitoring* 176
- *4.07.5.3 Psychophysiological Assessment* 177
- *4.07.5.4 Self-report Methods in Behavioral Assessment* 178
- *4.07.5.5 Psychometric Foundations of Behavioral Assessment* 179

4.07.6 BEHAVIORAL AND PERSONALITY ASSESSMENT 180

4.07.7 SUMMARY 181

4.07.8 REFERENCES 182

4.07.1 INTRODUCTION

Psychological assessment is the systematic evaluation of a person's behavior. The components of psychological assessment include the variables selected for measurement (e.g., beliefs, social behaviors), the measurement methods used (e.g., interviews, observation), the reduction and synthesis of derived data (e.g., whether summary scores are calculated for a questionnaire), and the inferences drawn from the data (e.g., inferences about treatment effectiveness). Psychological assessment affects the evolution of all social, cognitive, and behavioral science disciplines. The accuracy with which variables can be measured affects the degree to which relationships among behavioral, cognitive, environmental, and physiological events can be identified and explained. For example, our understanding of the impact of traumatic life stressors on immune system functioning, the relationship between depressed mood and self-efficacy beliefs, the effect of social response contingencies on self-injurious behavior, and the degree to which presleep worry moderates the impact of chronic pain on sleep quality depends on the strategies we use to measure these constructs.

Psychological assessment also affects clinical judgments. In adult treatment settings, clinical psychologists make judgments about a client's risk of suicide, whether treatment is warranted for a client, whether a client should be hospitalized, the variables that affect a client's behavior problems, and the best strategy for treating a client. Psychological assessment also helps the clinician select intervention goals and evaluate intervention effects.

There are many paradigms in psychological assessment. A psychological assessment paradigm is composed of a coherent set of assessment principles, values, assumptions, and methods of assessment. It includes assumptions about the relative importance of different types of behavior problems, the variables that most likely cause behavior problems, the most likely mechanisms through which causal variables operate, the importance and role of assessment in treatment design, and the best methods and strategies of assessment. A psychological assessment paradigm also includes guidelines for problem-solving, decision-making strategies, and data interpretation.

One powerful and evolving psychological assessment paradigm is behavioral assessment (Haynes & O'Brien, in press). Influenced by behavior-analytic, social-learning, and cognitive-behavioral therapy construct systems, the paradigm incorporates diverse methods of assessment but emphasizes naturalistic and analogue observation, self-monitoring, and electrophysiological measurement. The behavioral assessment paradigm also has many methodological elements, including an emphasis on the use of minimally inferential constructs, time-series measurement, hypotheses testing, and an idiographic (i.e., focus on an individual client) approach to assessment.

This chapter focuses on clinical applications of behavioral assessment with adults. The chapter will outline the conceptual and methodological elements of behavioral assessment and indicate how they aid clinical judgment and decision-making with adult clients. To illustrate the underlying assumptions, methods, and strategies of behavioral assessment, the first section presents a functional analytic causal model of a client—a vector diagram of a behavioral clinical case conceptualization. Following a discussion of the principles and methods of behavioral assessment, subsequent sections briefly discuss the history of behavioral assessment and psychometric considerations. Developments in behavioral assessment and differences between behavioral and nonbehavioral assessment paradigms are also discussed.

4.07.2 CLINICAL JUDGMENTS AND FUNCTIONAL ANALYSIS IN BEHAVIORAL ASSESSMENT

One of the most important and complex clinical judgments is the clinical case conceptualization. The clinical case conceptualization is a metajudgment—it is composed of many lower-level judgments regarding a client's behavior problems and the factors that affect them. It is a synthesis of assessment- and research-based hypotheses about a client and its primary application is for designing the most effective treatment.

In behavioral assessment, clinical case conceptualization is often termed "functional analysis" (Haynes & O'Brien, 1990). (Terms with similar meanings include "clinical pathogenesis map" [Nezu & Nezu, 1989] and "behavioral case formulation" [Persons, 1989]. The term "functional analysis" is often used in applied behavior analysis to refer to the systematic manipulation of variables, in a controlled assessment setting, to deteremine their effect on designated dependent variables.) Functional analysis is a central component in the design of behavior therapy programs because of individual differences between clients—two clients can manifest the same primary behavior problem for different reasons and, consequently, receive different treatments. Behavioral interventions are often designed to modify variables that are hypothesized to affect (i.e., account for variance in, trigger, maintain, moderate) problem behaviors and goals (Haynes, Spain, & Oliveira, 1993). Many permutations of causal variables can result in identical behavior problems, thereby resulting in different functional analyses and warranting different interventions.

Behavioral interventions are designed on the basis of many judgments about a patient reflected in a functional analysis. Clinical judgments with important implications for treatment decisions include the importance (e.g., severity, degree of risk associated with) of a client's multiple behavior problems; the relationships (e.g., strength, correlational vs. causal) among a client's multiple behavior problems, and the effects of behavior problems. Judgments about causal variables that affect the behavior problem (their importance, functional form, modifiability) are particularly important elements of the functional analysis.

There can be many errors in the clinical judgments that compose a clinical case conceptualization. Books by Eels (1997), Nezu and Nezu (1989), and Turk and Salovey (1988) discuss many of these errors. In brief, a clinician's judgments about a client can be affected by the clinician's preconceived beliefs, recent or particularly salient clinical experiences, selective attention to data that confirms the clinician's expectations, training-related biases, premature diagnoses, decisions based on initial impressions, and insufficient integrative abilities. These errors can reduce the validity of the case conceptualization and reduce the chance that the most effective treatment strategy will be selected for a patient.

The supraordinate goal of behavioral assessment is to reduce error and increase the validity in clinical judgments. The behavioral assessment paradigm suggests that clinical judgment error can be reduced to the degree that the assessor uses multiple sources of information, validated assessment instruments, time-series measurement strategies, focuses on multiple response modes, minimizes the inferential characteristics of variables, addresses behavior–environment interactions, and is guided by data from previously published studies (e.g, Persons & Fresco, 1998).

Haynes (1994), Haynes et al. (1993), Nezu and Nezu (1989), and Nezu et al. (1996) have outlined two methods to help the clinician systematically integrate the complex information contained in a functional analysis. These methods are the clinical pathogenesis map and functional analytic causal models (FACMs). Both involve systematic construction of vector diagrams of the component clinical judgments in a functional analysis and are designed to promote less intuitive intervention decisions. The clinical pathogenesis map and FACMs graphically model the clinician's hypotheses about a patient's behavior problems and goals and their relative importance, interrelationships, sequela and the strength, modifiability, and direction of action of causal variables. The FACM allows the clinician to estimate, quantitatively or qualitatively, the relative magnitude of effect of a particular treatment focus, given the clinician's hypotheses about the patient.

The FACM of a client presented in Figure 1 will be used to illustrate several underlying assumptions and methods of the behavioral assessment paradigm. The graphics in Figure 1 are explained in Figure 2. The client was a 35-year-old pregnant, married, unemployed woman (Mrs. A) who came to an outpatient mental health center complaining of constant headaches and sleeping problems. (This FACM was modified from one developed by Akiko Lau, University of Hawaii and discussed in Haynes, Leisen, & Blaine, 1997.) She was referred by her neurologist and had undergone extensive neurological and endocrinological examinations with negative results.

Mrs. A was first interviewed in an unstructured, open-ended manner, with the goal of encouraging her to talk about her behavior problems, goals, and the factors affecting them (Haynes, 1978). Structured interviews were then conducted to acquire more detailed information on specific behavior problems mentioned in the unstructured interview, such as her anxiety symptoms (Brown, DiNardo, & Barlow, 1994), marital relationship concerns (O'Leary, 1987), headache (James, Thorn, & Williams, 1993), sleep disturbance (Lichstein & Riedel, 1994), and other factors depicted in Figure 1. Validated questionnaires on marital satisfaction, child behavior problems, anxiety, and life stressors were also administered to provide additional information on issues raised in the interview process.

Because headaches and sleep difficulties were important problems, Mrs. A began daily self-monitoring after the second interview and continued to self-monitor throughout the assessment-treatment process. She recorded headache intensity and symptoms four times per day and each morning she recorded sleep-onset and awakenings for the previous night.

Marital conflict was a major concern for Mrs. A and one possible cause of her psychophysiological problems. Consequently, a one and a half hour assessment session (the third session) was conducted with her and her husband. During the session, the couple underwent a

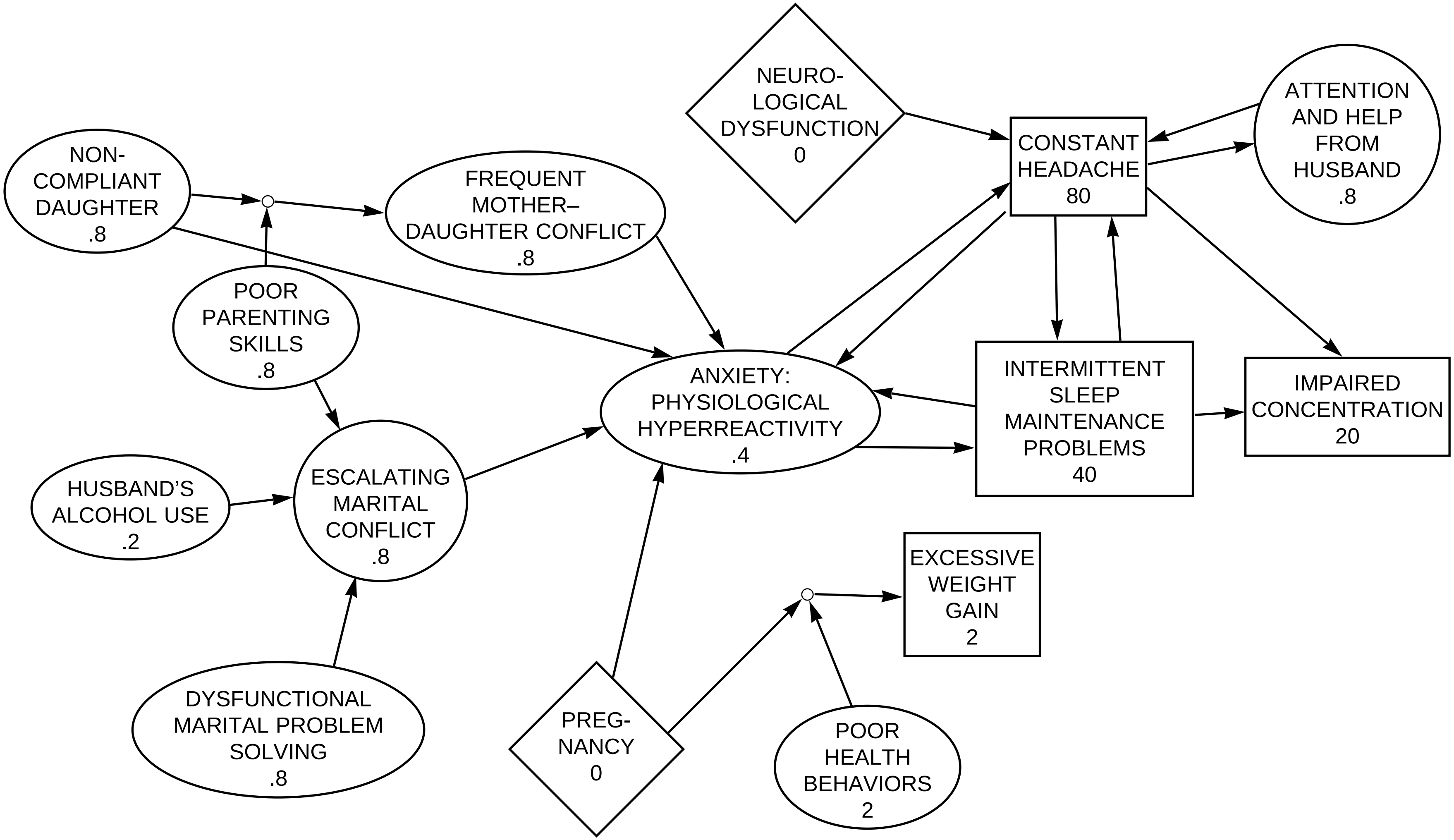
NEURO-LOGICAL DYSFUNCTION 0
ATTENTION AND HELP FROM HUSBAND .8
NON-COMPLIANT DAUGHTER .8
FREQUENT MOTHER–DAUGHTER CONFLICT .8
CONSTANT HEADACHE 80
POOR PARENTING SKILLS .8
ANXIETY: PHYSIOLOGICAL HYPERREACTIVITY .4
INTERMITTENT SLEEP MAINTENANCE PROBLEMS 40
IMPAIRED CONCENTRATION 20
HUSBAND'S ALCOHOL USE .2
ESCALATING MARITAL CONFLICT .8
EXCESSIVE WEIGHT GAIN 2
DYSFUNCTIONAL MARITAL PROBLEM SOLVING .8
PREG-NANCY 0
POOR HEALTH BEHAVIORS 2

Figures 1 and 2 An FACM of an outpatient woman with headaches and sleep disorders. The figures illustrate the relative importance of behavior problems, interrelationships among behavior problems, behavior problem sequalea, casual relationships, and the modifiability of casual variables.

ILLUSTRATING A FUNCTIONAL ANALYSIS WITH A FUNCTIONAL ANALYTIC CAUSAL MODEL

IMPORTANCE/MODIFIABILITY OF VARIABLES
(using width of variable boundary and coefficients)

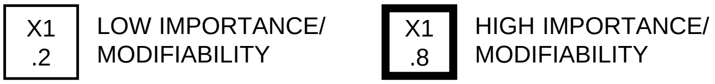
X1
.2
LOW IMPORTANCE/
MODIFIABILITY
X1
.8
HIGH IMPORTANCE/
MODIFIABILITY

TYPE AND DIRECTION OF RELATIONSHIP BETWEEN VARIABLES

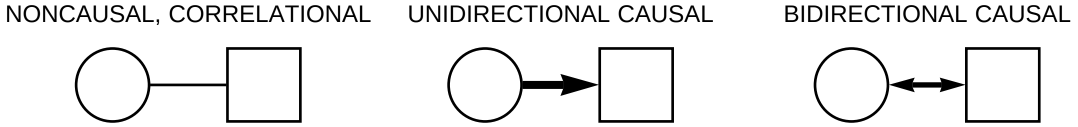
NONCAUSAL, CORRELATIONAL
UNIDIRECTIONAL CAUSAL
BIDIRECTIONAL CAUSAL

SYMBOLS

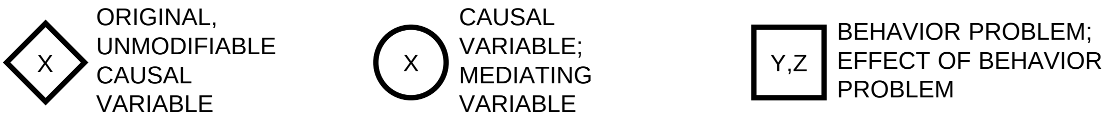
X
ORIGINAL, UNMODIFIABLE CAUSAL VARIABLE
X
CAUSAL VARIABLE; MEDIATING VARIABLE
Y,Z
BEHAVIOR PROBLEM; EFFECT OF BEHAVIOR PROBLEM

STRENGTH OF RELATIONSHIP BETWEEN VARIABLES

INDICATED BY ARROW THICKNESS; MORE PRECISELY BY COEFFICIENTS

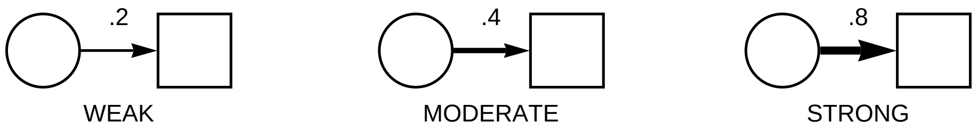
.2
.4
.8
WEAK
MODERATE
STRONG

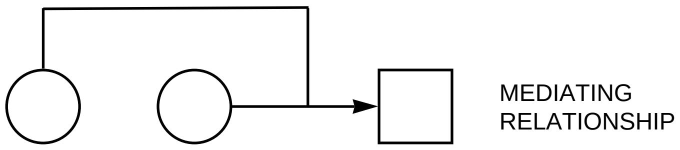
MEDIATING
RELATIONSHIP

conjoint structured interview about their marital relationship (e.g., perceived conflicts, strengths, spousal excesses and deficits, marital goals), and Mr. A also completed a marital satisfaction questionnaire. The couple participated in an analogue communication assessment, in which they discussed for 10 minutes their conflicts regarding disciplining their 12-year-old daughter. The conversation was recorded and later coded by the assessor.

Based on interview and questionnaire data, conflicts with her daughter were another source of distress for Mrs. A. A joint assessment session (the fourth session) was conducted in which the daughter was interviewed about her perception of family issues. Also, the mother and daughter were observed in two structured assessment settings: while trying to resolve one of their frequent sources of conflict (the daughter's refusal to do her school work) and while Mrs. A attempted to help her daughter with some school work.

The functional analytic causal model of Mrs. A emphasizes many elements of a clinical case conceptualization that are important for behavioral treatment decisions. (Many other factors—e.g., treatment history, client cognitive resources, cost-efficiency of treatments, responses of persons in the client's environment—affect treatment design in addition to those included in a FACM.) Important and controllable functional relationships are highlighted in the FACM because of their clinical utility. The FACM for Mrs. A recognizes some unmodifiable variables but these have no clinical utility. Unidirectional and bidirectional causal relationships are shown because they can significantly affect decisions about what variables should be targeted in treatment. Treatment decisions are also affected by the strength of causal relationships and the degree of modifiability of causal variables, depicted in Figure 1.

Before considering the specific assumptions of the behavioral assessment paradigm that influenced the assessment strategy outlined above and the clinical judgments summarized in Figure 1, several additional attributes of the functional analysis should be briefly noted. First, the functional analysis (and the FACM) reflects the clinician's current judgments about a client. It is a subjectively derived (although data-influenced), hypothesized model. It is also unstable, in that it can change with changes in naturally occurring causal variables, with the acquisition of additional data, and as a result of treatment. For example, a change in the variables that affected Mr. A's drinking could lead to a significant change in the FACM for Mrs. A. A FACM for a client may be conditional. For example, some variables affecting Mrs. A's behavior problems may change after the birth of her child.

A final note on the limitations of the functional analysis. Despite its central role in behavior therapy, the functional analysis is limited in several ways: (i) the best assessment methods for developing a functional analysis have not been identified, (ii) the best methods for formulating a functional analysis from assessment data have not been determined, and (iii) for many behavior problems, the incremental utility and cost-effectiveness of the functional analysis have yet to be established.

4.07.3 CONCEPTUAL FOUNDATIONS OF BEHAVIORAL ASSESSMENT

Many methodological elements of the behavioral assessment paradigm, such as the preferred methods of assessment and the variables targeted in assessment, are influenced by its underlying assumptions. The following sections review two sets of assumptions: (i) those concerning the causal factors associated with behavior problems and goals, and (ii) those concerning the characteristics of behavior problems. This section also discusses implications of these assumptions for behavioral assessment strategies with adults. More extensive discussions of underlying assumptions in behavioral assessment can be found in Bandura (1969), Barrios (1988), Bellack and Hersen, (1988), Bornstein, Bornstein, and Dawson (1984), Ciminero, Calhoun, and Adams (1986), Cone (1988), Eysenck, (1986), Haynes (1978), Hersen and Bellack (1998), Johnston and Pennypacker, (1993), Kratochwill and Shapiro (1988), Mash and Terdal (1988), Nelson and Hayes (1986), O'Donohue and Krasner, (1995), Ollendick and Hersen (1984, 1993), Strosahl and Linehan (1986), and Tryon (1985).

4.07.3.1 Assumptions About the Causes of Behavior Disorders

Psychological assessment paradigms differ in the assumptions they make regarding the causes of behavior disorders. Although causal assumptions and the identification of causal variables in pretreatment assessment are less important for treatment paradigms with limited treatment options (e.g., Gestalt, transactional therapies), the identification of potential causal variables is a primary objective in pretreatment behavioral assessment. This is because hypothesized controlling variables are targeted for modification in behavior therapy and it is presumed that causal variables may vary across patients with the same behavior problems. The variables presumed to cause Mrs. A's sleep problems may not operate for other patients with identical sleep problems. Consequently, other patients with the same sleep disorder would be treated differently.

Behavioral assessment strategies are guided by several empirically based and interrelated assumptions about the causes of behavior problems. These assumptions include: multiple causality, multiple causal paths, individual differences in causal variables and paths, environmental causality and reciprocal determinism, contemporaneous causal variables, the dynamic nature of causal relationships, the operation of moderating and mediating variables, interactive and additive causality; and situations, setting events, and systems factors as causal variables and dynamical causal relationships.

4.07.3.1.1 Multiple causality

Behavior problems are often the result of multiple causal variables acting additively or interactively (Haynes, 1992; Kazdin & Kagan, 1994). This is illustrated in Figure 1 by the multiple factors influencing Mrs. A's headaches. Although some behavior problems and the behavior problems of some individuals (e.g., asthma episodes that are mostly triggered by exposure to specific allergens) may primarily be the result of single causal variables, multivariate causal models have been proposed for most adult behavior disorders, including schizophrenia, chronic pain, sleep disorders, paranoia, personality disorders, child abuse, and many other behavior disorders (see reviews in Gatchel & Blanchard, 1993; Sutker & Adams, 1993).

4.07.3.1.2 Multiple causal paths

A causal variable may also affect a behavior problem through multiple paths. Note that for Mrs. A, physiological hyperreactivity can directly influence sleep but hyperreactivity can also influence sleep because it produces headaches. Similarly, there may be many paths through which social isolation increases the risk of depression (e.g., by restricting the potential sources of social reinforcement, by increasing dependency on reinforcement from a few persons) and many paths through which immune system functioning can be impaired by chronic life stressors (e.g., dietary changes, reduction of lymphocyte levels).

4.07.3.1.3 Individual differences in causal variables and paths

Models of causality for behavior problems are further complicated because the permutations of causal variables and causal mechanisms can differ across clients with the same behavior problem. For example, there can be important differences in the causal variables and causal paths among persons reporting chronic pain (Waddell & Turk, 1992), exhibiting self-injurious behaviors (Iwata et al., 1994), or who complain of difficulties in initiating and maintaining sleep (Lichstein & Riedel, 1994). Some differences in causality may covary with dimensions of individual differences. For example, the causes of depression, marital distress, and anxiety may vary as a function of ethnicity, age, gender, religion, sexual orientation, and economic status (see discussions in Marsella & Kameoka, 1989).

4.07.3.1.4 Environmental causality and reciprocal determinism

The behavioral assessment paradigm also stresses the importance of environmental causality and behavior–environment interactions (McFall & McDonel, 1986). Many studies have shown that it is possible to account for variance in many behavior problems by examining variance in response contingencies (e.g., how others respond to self-injurious behaviors, depressive statements, or asthma episodes can affect the parameters of those behaviors—a "parameter" of a behavior refers to a quantitative dimension, such as rate, duration, magnitude, and cyclicity), situational and antecedent stimulus factors (e.g., alcohol use may vary reliably across social settings, anxiety episodes may be more likely in more interpersonally stressful environments), and other learning principles (e.g., modeling, stimulus pairings; see discussions in Eysenck & Martin, 1987; O'Donohue & Krasner, 1995).

An important element of the principle of environmental causality is reciprocal determinism, (i.e., bidirectional causality, reciprocal causation; Bandura, 1981)—the idea that two variables can affect each other. In a clinical context, reciprocal determinism refers to the assumption that clients can behave in ways that affect their environment which, in turn, affects their behavior. For example, a client's depressive behaviors (e.g., reduced social initiations and positive talk) may result in the withdrawal of the client's friends, increasing the client's loss of social reinforcement and increasing the client's depressive mood and behaviors. A hospitalized paranoid patient may behave suspiciously with staff and other patients. These behaviors may cause others to avoid the patient, talk about him/her and behave in many other ways that confirm and strengthen the patient's paranoid thoughts. With Mrs. A, we presume that there are some ways that Mrs. A is behaving that might contribute to her marital distress and difficulties. An emphasis on bidirectional causation does not negate the possibility of unidirectional environmental causal factors. In some distressed marriages, for example, a spouse may independently be contributing to marital distress by being verbally abusive or unsupportive. However, pure unidirectional causation may be rare.

Viewing clients within a reciprocal determinism framework effects the focus of assessment and treatment. Clients are considered active participants in their own lives, as active contributors to their goal attainment and to their behavior problems. Consequently, clients are encouraged to participate actively in the assessment–treatment process. Assessment goals include identifying the ways that clients may be contributing to their behavior problems and ways they can contribute to the attainment of treatment goals.

One consequence of reciprocal determinism is that labels such as "behavior problem (dependent variable)" or "causal variable (independent variable)" become less distinct. Often, either variable in a bidirectional relationship can be described as a behavior problem and a causal variable—each variable can be either or both. Which variables are described as problem vs. cause depends more on convention or the intent of the assessor and client than on the characteristic of the functional relationships. As indicated in the functional analytic causal models (e.g., Haynes, 1994), treatment decisions are dictated more by estimates of the strength of causal relationships than by the label of the variable.

The concept of reciprocal determinism also promotes a behavioral skills focus in assessment and treatment. A client's behavior problems are presumed to be a partial function of their behavioral repertoire. Their behavioral excesses, deficits, and strengths are presumed to affect whether they will experience problems in some situations, the type and magnitude of behavior problem experienced and how long the problem persists. For example, a behavior skills assessment with a socially anxious client might focus on specific deficits that prevent the client from forming more frequent and satisfying friendships. Similar to a task analysis, the necessary skills for attaining a treatment goal (e.g., establishing positive friendships) are broken down into molecular components and the client's abilities on these components are evaluated. With Mrs. A, it would be important to determine what additional parenting and marital communication skills might help Mrs. A develop a more positive relationship with her daughter and husband. An example would include the ability to clearly and positively talk about her ideas and concerns.

Cognitive skills are often targeted by behavioral assessors. The clients' beliefs, expectancies, deductions, and other thoughts regarding their capabilities in specific situations (e.g., Linscott & DiGiuseppe, 1998) are often considered essential elements for effective functioning. A molar-level skill is adaptive flexibility—an overarching goal of behavior therapy is to help the client to develop behavior repertoires that facilitate adaptability to various, novel and changing environments.

4.07.3.1.5 Contemporaneous causal variables

The behavioral assessment paradigm emphasizes the relative importance and utility of contemporaneous rather than historical, causal factors. It is presumed that a more clinically useful, and sometimes more important, source of variance in a client's behavior problems can be identified by examining the client's current, rather than historical, learning experiences, social contingencies, and thoughts. For example, suspicious thoughts and behaviors can undoubtedly be learned as a child from parental models (e.g., parents who teach a child to be mistrustful of others' intentions; Haynes, 1986). However, early parent–child learning experiences are difficult to identify and "treat" in therapy. Assessment efforts might more profitably be focused on contemporaneous causal variables for paranoid thoughts—such as restricted social network that precludes corrective feedback about misperceptions, social skills deficits, hypersensitivity to negative stimuli or negative scanning, or failure to consider alternative explanations for ambiguous events. These can also be important causal variables for a client's paranoid behaviors and are more amenable than historical events to intervention.

The emphasis on contemporaneous, reciprocal, behavior–environment interactions dictates an emphasis on particular methods of assessment. For example, naturalistic observation, analogue observation, and self-monitoring are better suited than retrospective questionnaires to measuring contemporaneous, reciprocal dyadic interactions. Additionally, in behaviorally oriented interviews and questionnaires clients are more often asked about current than about past behavior–environment interactions (Jensen & Haynes, 1986; Sarwer & Sayers, 1998).

An emphasis on contemporaneous reciprocal determinism is compatible with the causal role of genetic and physiological factors, and early learning experiences. Evidence from many sources suggests that genetic factors, neurophysiological mechanisms, medical disorders, and early learning (e.g., early traumatic experiences) can serve as important causal variables for behavior problems (see reviews in Asteria, 1985; Haynes, 1992; Sutker & Adams, 1993). Sometimes, physiological, behavioral, and cognitive variables are different modal expressions of the same phenomena.

The emphasis on contemporaneous behavior and environmental causality is evident in the contemporaneous focus of many behavioral assessment interviews. However, behavioral assessors differ in their emphasis on historical data. Joseph Wolpe, for example, emphasized the importance of gathering a complete clinical history for patients before therapy (Wolpe & Turkat, 1985). Often historical information can aid in the development of hypotheses regarding the time-course and causes of behavior problems. For example, a careful interview about past behaviors, events and treatment experiences can help determine if Mrs. A may be experiencing neurological deficits (e.g., she had a minor head injury two years prior to this assessment) and may help estimate the degree to which her health-related behaviors (e.g., poor diet and exercise habits) are modifiable.

4.07.3.1.6 Interactive and additive causality

In the section on multiple causality, I noted that a behavior problem often results from multiple causal factors acting concurrently—this is an additive model of causality. Causal variables can also interact—this is a multiplicative model of causality. Interactive causality occurs when the causal effects of one variable vary as a function of the values of another causal variable (see discussion in Haynes, 1992). Furthermore, the effects of the variables in combination often cannot be predicted by simply summing their independent effects. A longitudinal study by Schlundt, Johnson, and Jarrell (1986) demonstrated interactive causal effects with bulimic clients. The probability of postmeal purging was significantly related to a history of recent purges (i.e., purging tended to occur in cycles). However, the social context within which eating occurred affected the strength of the relationship between those two variables. The chance of purging was higher when the person had recently purged, but especially higher if the person ate alone. The effect of each causal variable (purging history, social context) depended on the value of the other causal variable.

Diathesis-stress models of psychopathology are common exemplars of interactive causality (e.g., Barnett & Gotlib, 1988). Diathesis-stress models suggest that environmental stressors and physiological or genetic vulnerability (or genetic and later physiological challenges) interact to affect the probability that a particular behavior disorder will occur.

4.07.3.1.7 Situations, setting events, and systems factors as causal variables

One assumption of the behavioral assessment paradigm is that the probability (or another parameter such as magnitude) of behavior problems varies across situations, settings, and antecedent stimuli (e.g., discrete and compound antecedent stimuli, contexts, discriminative stimuli for differential reinforcement contingencies); behavior problems are conditional. The conditional nature of behavior problems has important causal implications because it marks the differential operation of causal factors. Mrs. A was more likely to experience anxiety symptoms in the presence than in the absence of her daughter. The presence of the daughter marked the operation of a causal relationship and suggested to the assessor that the mother–daughter interactions should be explored in greater detail.

A situational model of behavior problems contrasts with traditional personality trait models, which emphasize a higher degree of cross-situational consistency of behavior and some enduring trait of the person as the primary causal variable for behavior problems (see subsequent discussion of personality assessment). However, situational and trait models of behavior problems can be compatible. Knowledge of the robust behaviors of a client (e.g., those behaviors that do not vary to an important degree across conditions) and knowledge of the situational factors that influence the client's behavior can both contribute to a functional analysis of the client. This "interactional" perspective is a welcomed refinement of exclusively trait models (see discussions by Mischel, 1968; McFall & McDonel, 1986).

Cross-situational consistency can vary across different behaviors, individuals, and situations; relative cross-situational consistency in behavior can occur, but may not. Because the assessor does not have prior knowledge of the degree of cross-situational consistency of a client's behavior problems, the assessor must evaluate their conditional nature. Unfortunately, strategies and classification schema for situations have not been developed.

Although the behavioral assessment paradigm emphasizes contemporaneous causal factors (e.g., a SORC [stimulus, organism, response, contingency] model; Goldfried, 1982), extended social systems can play an important causal role. Mrs. A's marital satisfaction may be affected by her relationships with her friends and family. Her daughter's behavior problems in the home may be affected by the social and academic environment of her school. Assessment efforts cannot be confined to individual elements extracted from a complex array of interacting variables. Chaos theory and dynamic modeling also suggest that it may be difficult to develop powerful predictive or explanatory models unless we measure behavior within the complex dynamical systems in which the behavior is imbedded (Vallacher & Nowack, 1994)

4.07.3.1.8 Dynamic causal relationships

All elements of a functional analytic causal model—the causal variables that affect a client's behavior problems, the strengths of causal relationships, moderating variables, for example, are nonstationary (Haynes, Blaine, & Meyer, 1995). Causal relationships for a client can be expected to change across time in several ways. First, new causal variables may appear: Mr. or Mrs. A may develop new health problems; Mr. A may lose his job. Second, a causal variable may disappear: Mr. A may stop drinking; Mrs. A may give birth to her baby. Third, the strength and form of a causal relationship are likely to change over time. There may be a decrease in sleep disruption originally triggered by a traumatic event; marital distress that originally caused depressive symptoms may be exacerbated by continued depressive reactions. Fourth, moderating variables may change: Clients may change their expectancies about the beneficial effects of alcohol (Smith, 1994). In causal models of behavior disorders, a moderating variable is one that changes the relationship between two other variables. For example, "social support" would be a moderating variable if it affected the probability that an environmental disaster would be associated with symptoms of posttraumatic sress disorder (PTSD).

4.07.3.1.9 Additional assessment implications of causal assumptions

Emphases on multivariate, idiosyncratic, interactive, reciprocal deterministic, and dynamic causal models have several implications for behavioral assessment strategies that were briefly noted in the previous sections and will be discussed in greater detail later in this chapter. The assessment implications include: (i) pretreatment assessment should be broadly focused on multiple variables; (ii) classification will usually be insufficient to identify the causal variables operating for a particular client and insufficiently precise to identify the client's behavior problems; (iii) assessors should avoid "premature" or "biased" presumptions of causal relationships for a client and draw data-based inferences whenever possible; (iv) a valid functional analysis is most likely to result from assessment that focuses on multiple response modes, uses multiple methods; and gathers data from multiple sources; (v) it is important to identify the mechanisms underlying causal relationships (see also, discussion in Shadish, 1996); and (vi) a time-series assessment strategy can be an effective method of identifying and tracking behavior problems and potential causal factors.

4.07.3.2 Assumptions About the Characteristics of Behavior Problems

Behavioral assessment strategies and the resulting clinical case conceptualizations are strongly affected by assumptions of the behavioral assessment paradigm about the characteristics of behavior problems. Several of these assumptions were mentioned earlier in this chapter. They include the multimodal and multiparameter characteristics of behavior problems, differences among clients in the importance of behavior problem modes and parameters, the complex interrelationships among a client's multiple behavior problems, and the conditional and dynamic natures of behavior problems.

4.07.3.2.1 Behavior problems can have multiple response modes

Adult behavior problems can involve more than one response mode. For example, PTSD can involve physiological hyperreactivity, subjective distress, avoidance of trauma-related situations and thoughts, and distressing recollections and dreams of the traumatic event (American Psychiatric Association, 1994; Kubany, 1994). The degree of covariation among modes of a behavior problems can vary across persons (note that the *Diagnostic and statistical manual of mental disorders* [4th ed.] requires that a client manifest only one of five major [category B] symptoms for a diagnosis of PTSD). For example, some PTSD clients show only slight evidence of distressing recollections of the event but show significant physiological reactivity to event-related cues while others show the opposite pattern.

Low levels of covariation among the multiple modes of behavior problems have been found in both group nomothetic and single-subject time-series research (see discussion in Gannon & Haynes, 1987). Acknowledging that the apparent magnitude of covariation reflects the ways in which the modes are measured, these findings suggest that behavior problem modes can be discordant across persons and for one person across time. (See discussions in Cone (1979) and Eifert and Wilson (1991). Different response modes are often measured with different methods. Different response modes can also have different response latencies which can reduce apparent magnitudes of covariation if all modes are sampled simultaneously.)

Discordance among response modes for some clients has many clinical implications. Different response modes of a behavior problem may have different latencies in response to treatment. Furthermore, some treatments may have a stronger effect on some modes than on others. Consequently, the effects of treatment may be judged differently depending on which response mode is measured by which method. More important for a functional analysis of a patient, different response modes can be affected by different causal factors and respond differently to the same causal factor. Therefore, the selection of the most important mode for a client's behavior problem can affect the functional analysis and the intervention program designed for that client.

Assessment strategies should be congruent with the multimodal nature of behavior problems. First, because causal inferences are an important component of the functional analysis and guide treatment decisions, the primary mode of a client's behavior problem should be identified. Second, as the prior example of PTSD illustrated, diagnosis may be helpful but is usually an insufficient basis for the identification of the most important response modes for a client. Third, inferences regarding treatment effects for one mode may not be generalizable to other modes. In sum, behavioral assessment should have a multimodal focus.

4.07.3.2.2 Behavior problems have multiple parameters

As previously mentioned, each behavior problem mode can have multiple parameters. Parameters are quantitative dimensions, such as duration, magnitude, and rate, that can be used to describe a behavior. As with response modes, there are important between-client differences in the relative importance of different behavior problem parameters. For example, Mrs. A reported mildly intense but constant headaches and intermittent but severe sleep disruption. Other clients could report the same problems with different parameters (e.g., infrequent but debilitatingly intense headaches). Similarly, some clients report frequent but severe episodes of depression that last for only a few days; others report episodes of mild to moderate depression that can last months.

Multiple parameters of behavior disorders have important implications for the functional analysis because different parameters may be affected by different causal variables. For example, Barnett and Gotlib (1988) reviewed the literature on depression and suggested that learned helplessness beliefs seem to affect the duration and magnitude of depressive behaviors. However, learned helplessness beliefs could not account for the onset of depressive behaviors. Consequently, the functional analysis and treatment of a client with frequent depressive episodes might be different from the functional analysis and treatment of a client with infrequent but persistent depressive episodes.

One assessment implication that permeates many assumptions underlying behavioral assessment is that aggregated measures are insufficiently precise for a functional analysis. Between-person differences in the importance of behavior problem modes and parameters mandate careful specification and measurement of multiple modes and parameters. Measures of behavior problems that aggregate across modes, parameters, situations, for example, a single measure of "depression" or "anxiety," will often be insufficiently precise for the development of a functional analysis and for the design of intervention programs. Aggregated indices may be helpful as a general index of the magnitude or generality (or unconditional probability) of a behavior problem, but are insufficient for the precise inferences that are necessary for the evolution of assessment paradigms, functional analyses, and treatment strategies.

One helpful strategy for gathering more precise data on behavior problem parameters is the construction of a time-course for behavior problem—a time line reflecting occurrence, magnitude, and duration of behavior problems. An example of this method is the "timeline followback" by Sobell, Toneatto, and Sobell (1994) to establish a time-course of substance use.

4.07.3.2.3 Client can have multiple behavior problems

Many clients have multiple behavior problems. For example, Beck and Zebb (1994) reported that 65–88% of panic disorder patients have a coexisting behavior disorder, Figley (1979) reported a high incidence of comorbidity for PTSD, Regier et al., (1990) noted that drug users often have other concurrent behavior problems. Similar observations of comorbidity have been reported for panic disorders (Craske & Waikar, 1994) and depression (Persons & Fresco, 1996).

A client with multiple behavior problems challenges the clinical judgment capabilities of the assessor and complicates the functional analysis because the mode and parameter of each behavior problem can be affected by multiple causal variables: functional analytic causal models were developed partly as a method of organizing these clinical judgments. Additionally, multiple behavior problems can have complex causal and noncausal relationships. Note the relationships between sleep and headache problems for Mrs. A illustrated in Figure 1. Beach, Sandeen, and O'Leary (1990) observed a reciprocal causal relationship between marital distress depression for some patients (with many variables moderating that relationship). Hatch (1993) observed that depression can affect pain perception of headache patients and that headaches may contribute to the occurrence of depressive episodes for some patients.

The assumption that clients may have more than one behavior problem has several implications for behavioral assessment strategies. First, initial assessment, such as the intake interview (Kerns, 1994; Sarwer & Sayers, 1998) must be broadly focused to identify a client's multiple behavior problems. Premature narrowing of the assessment focus may preclude the identification of important behavior problems. Following a broad survey, subsequent assessment efforts can be focused on problem specification and the functional relationships relevant to each behavior problem. The functional analysis and intervention decisions will also be affected by estimates of the form and strength of relationship among, and relative importance of, a client's behavior problems. Multiple behavior problems also mandate a multivariate focus in treatment outcome evaluation.

Sometimes, the identification of functional response groups will aid in treatment decisions (Sprague & Horner, 1992). A functional response group is a set of behaviors, which may differ in form, that are under the control of the same contingencies (a set of behaviors that has the same function, (Haynes, 1996). Adaptive elements of the response class can sometimes be strengthened to weaken maladaptive elements of that class. Relaxation skills may be taught as a substitute for dysfunctional ways of reducing physiological hyperarousal. Effective communication skills may reduce the frequency of self-injurious behavior for some developmentally disabled individuals (Durand & Carr, 1991).

4.07.3.2.4 Behavior problems are conditional

As noted earlier in this chapter, behavior problems seldom occur randomly or unconditionally. Although it is difficult to predict the occurrence of many behavior problems, the probability of occurrence often varies as a function of settings, antecedent stimuli, environmental contexts, physiological states, and other discriminative stimuli (Gatchel, 1993; Glass, 1993; Ollendick & Hersen, 1993). It was previously noted that identifying sources of variance in behavior problems can help the assessor to identify causal variables and mechanisms. For example, behavioral assessors attempt to identify the conditions that trigger the startle responses of a client with PTSD (Foa, et al., 1989), that triggers a client's asthma episodes (Creer & Bender, 1993), and the conditions associated with marital violence (O'Leary, Vivian, & Malone, 1992) to develop a functional analysis and design the most appropriate intervention program.

The conditional nature of behavior problems further diminishes the clinical utility of aggregated measures of a behavior problem—assessment instruments that provide a "score" without providing more precise information of the conditional nature of the behavior problem. Assessment instruments should help the assessor examine the conditional probabilities of the behavior problem or the magnitude of shared variance between the behavior problem and multiple situational factors. For Mrs. A, the assessor would try to determine the situations that provoked conflict with her daughter, and to determine the events that increased or decreased the intensity of her headaches. Behavioral questionnaires and interviews, self-monitoring, and naturalistic observation can aid in identifying the conditional aspects of behavior problems.

4.07.3.2.5 The dynamic nature of behavior problems

The parameters and qualitative aspects (e.g., topography, form, characteristics) of behavior problems can change over time. The frequency, intensity, and content of arguments between Mr. and Mrs. A will probably change within a few weeks and months. The magnitude, frequency, duration, and form of a clients' PTSD symptoms, paranoid delusions, nightmares, excessive dieting, and pain can change in important ways across time. Also, new behavior problems may occur and some behavior problems may become less important.

Dynamic behavior problems and other variables can only be sensitively tracked by measuring them frequently, using time-series assessment strategies. The frequency with which dynamic variables should be sampled depends on their rate of change. Collins and Horn (1991), Heatherton and Weinberger (1994), Johnston and Pennypacker (1993), Kazdin (1997), and Kratochwill and Levin (1992) discuss instruments, strategies, and issues in the measurement of dynamic variables. Frequent measurement of behavior problems can also help the assessor identify important causal relationships. For example, recent changes in the magnitude of a client's depressive symptoms may provide cues about environmental or cognitive causal factors. Changes in Mrs. A's sleep patterns could trigger inquiries about events preceding the change. Self-monitoring, brief structured interviews and short questionnaires are particularly amenable to time-series assessment.

4.07.4 METHODOLOGICAL FOUNDATIONS OF BEHAVIORAL ASSESSMENT

The methodological elements of the behavioral assessment paradigm, the preferred strategies of behavioral assessment, are dictated by the assumptions about behavior and its causes outlined in the previous sections. Many of these methodological elements were introduced earlier in this chapter and are outlined in Table 1.

This section will discuss three of the elements of the behavioral assessment paradigm delineated in Table 1: (i) the emphasis on empirical hypothesis-testing, (ii) the idiographic emphasis, and (ii) the use of time-series assessment strategies. These and other methodological elements from Table 1 have been presented in Cone and Hawkins (1977), Haynes (1978), Hersen and Bellack (1996), Johnston and Pennypacker, (1993), Mash and Terdal (1988), Nelson and Hayes (1986), Ollendick and Hersen (1993), and Shapiro and Kratochwill (1988).

Most of the methodological elements of the behavioral assessment paradigm, particularly the assessment strategies, serve the priority placed on a scholarly, empirically based approach to clinical assessment. It is assumed that clinically useful knowledge about behavior problems is best acquired through the frequent administration of carefully constructed assessment instruments that are precisely focused of multiple, minimally inferential variables. Behavioral assessors are likely to eschew assessment instruments that are poorly validated, provide indices of vaguely defined and highly inferential constructs, and have unsubstantiated clinical utility.

4.07.4.1 An Empirically Based Hypothesis Testing Approach to Assessment

A hypothesis testing and refinement climate guides the behavioral assessment of adults. The assessor makes many tentative judgments about the client beginning early in the preintervention assessment phase. The clinician estimates the relative importance of the client's behavior problems and goals, the variables that influence problems and goals, and other elements of the functional analysis. Based on these early clinical judgments, the assessor also begins to estimate the most effective methods of intervention and selects additional assessment instruments, (see discussions in Eels, 1996; Haynes et al., 1997; Nezu et al., 1996; O'Brien & Haynes, 1995; Persons & Bertagnolli, 1994; Turk & Salovey, 1988). These hypotheses are tested and refined as assessment continues. With Mrs. A, for example, initial judgments that deficits in marital communications skills were functionally related to their marital conflicts would be evaluated subsequently through analogue communication assessment and by teaching the couple more positive discussion strategies. If their communication skills increased but their marital conflicts did not decrease invalid hypotheses may have been initially drawn about the causes of this couple's marital conflicts. There are other possible explanations for a lack of observed covariation in a causal relationship. For example, if another causal factor or moderating variable became operational while communication skills were strengthened, there could appear to be no causal relationship between communication skills and conflict (Shadish, 1996).

A scholarly hypotheses-testing approach to psychological assessment requires that the results of assessment and contingent clinical inferences be viewed skeptically. Clinical judgments are always based on imperfect measurements of nonstationary data and are intrinsically subjective. The validity and utility

Table 1 Methodological emphases of the behavioral assessment paradigm.

Assessment strategies

1. Idiographic assessment; a focus on the client's specific goals, individual behavior problems; individually tailored assessment; a de-emphasis on normatively based assessment (e.g., trait-focused questionnaires).
2. A hypothesis-testing approach to assessment and treatment (including the use of interrupted time-series designs).
3. Time-series assessment strategies (as opposed to single-point or pre–post measurement strategies); frequent measures of important variables.
4. Multimethod and multi-informant measurement of variables across multiple situations.
5. Quantification of variables (measurement of rates, magnitudes, durations).
6. The use of validated assessment instruments in conditions in which sources of measurement error are minimized.

The focus of assessment

7. Precisely specified, lower-level, and less inferential constructs and variables.
8. Observable behavior (as opposed to hypothesized intrapsychic events).
9. Client–environment interactions, sequences, chains, and reciprocal interactions.
10. Behavior in the natural environment.
11. Events that are temporally contiguous to behavior problems.
12. Multiple client and environmental variables in pretreatment assessment (multiple behavior problems, causal variables, moderating and mediating variables).
13. Multiple targets in treatment outcome evaluation (e.g., main treatment effects, side effects, setting and response generalization).
14. Multiple modes and parameters of behavior and other events.
15. Extended systems (e.g., family, work, economic variables).

Source: Haynes (1996b).
These are relative emphases whose importance varies across behavioral assessment subparadigms (e.g., behavior anlaysis, cognitive-behavioral). Many are compatible with other psychological assessment paradigms (e.g., neuropsychological, educational achievement).

of clinical judgments will covary with the degree to which they are guided by the assessment principles outlined in Table 1. The assessor can also reduce some of the biases in clinical judgments by basing them on assessment data and being receptive to clinical assessment data that is inconsistent with those judgments.

4.07.4.2 An Individualized Approach to Assessment

Partially due to between person differences in behavior problems, goals, and functional relationships, the behavioral assessment paradigm emphasizes individualized assessment—an idiographic approach to assessment (Cone, 1986). An individualized approach to assessment is manifested in several ways: (i) self-monitoring targets and sampling procedures are often tailored to the individual client; (ii) role play scenarios and other assessment instruments are often individually tailored (e.g., Chadwick, Lowe, Horne, & Higson, 1994); (iii) client-referenced and criterion-referenced assessment instruments, in contrast to norm-referenced assessment instruments, are often used; (iv) treatment goals and strategies are often individually tailored (e.g., de Beurs, Van Dyck, van Balkom, Lange, & Koele, 1994), and (v) within-subject, interrupted time-series and multivariate time-series designs are often used in research.

4.07.4.3 Time-series Assessment Strategies

As indicated in previous sections, time-series measurement of independent and dependent variables across time (e.g., the frequent [e.g., >40] daily samples of a client's behavior problems and causal variables) is an important strategy in behavioral assessment. It is a powerful strategy in clinical research and has many advantages in clinical assessment. First, it can help estimate causal and noncausal relationships for a participant's behavior problems. By subjecting the data to sequential analyses or cross-lagged correlations time-series assessment can provide information on conditional probabilities and suggest possible causal mechanisms for behavior problems (Bakeman & Gottman, 1986; Moran, Dumas, & Symons, 1992; Tryon, 1998).

A major advantage is that time-series assessment allows the researcher and clinician to control for and examine the temporal precedence of causal variables. As O'Leary, Malone, and Tyree (1994) noted in their discussion of marital violence, it is difficult to draw inferences about causal factors unless they are measured well ahead of the targeted behavior problem. A concurrent measurement strategy (in which hypothesized independent and dependent variables are measured concurrently) cannot be sufficient for causal inferences because the form (e.g., correlational, bidirectional causal) of the relationships cannot be established.

Statistical analysis of time-series data can be cumbersome in many clinical assessment contexts (however, see subsequent discussion of computer aids). However, time-series assessment is also the best strategy for tracking the time-course of behavior problems and goal attainment during therapy. Frequent monitoring increases the chance that failing treatment will be detected early (Mash & Hunsley, 1993) and that the clinician can make needed changes in the functional analysis and intervention programs.

Time-series assessment is congruent with an emphasis on professional accountability and empirically based judgments in clinical treatment. Time-course plots can be a useful means of documenting treatment effects and of providing feedback to clients about behavior change and possible causal relationships for their behavior problems. It can help the clinician identify naturally occurring causal mechanisms for a client's mood fluctuations, panic episodes, and other behavior problems. Time-series data was acquired with Mrs. A when she self-monitored daily her headaches and sleep problems.

Finally, time-series measurement is an essential element in interrupted time-series designs, such as the A–B–A–B, multiple baseline, or changing-criterion designs (Kazdin, 1997; Shapiro & Kratochwill, 1988). These designs can strengthen the internal validity of inferences about treatment effects and mechanisms.

4.07.4.4 Quantitative and Qualitative Approaches to Behavioral Assessment

The empirical elements of the behavioral assessment paradigm are partially responsible for the current emphasis on treatment outcome evaluation through frequently applied, minimally inferential, validated assessment instruments. This quantitative approach to clinical inference reflects and accentuates the growing importance of systematic evaluation by professionals delivering clinical services. The behavioral assessment paradigm provides a useful structure for the evaluation of clinical service delivery and intervention outcome. However, it is possible for clinicians and clinical researchers to overemphasize excessively molecular measures and excessive quantification. Quantification is an indispensable component of clinical inference but an exclusive reliance on quantification can promote a focus on variables and effects with minimal practical importance. The emphasis on clinical significance (e.g., Jacobson & Truax, 1991) of treatment effects and functional relationships reflects the practical and clinical importance of effects, as well as their statistical significance.

Qualitative analyses (Haberman, 1978) can compliment quantitative analyses. Behavioral assessors can generate clinically useful hypothesis by supplementing quantitative with qualitative analyses of clinical phenomena. Using time-sampling measurement strategies to code specific dyadic interaction sequences in a communication analogue between Mrs. and Mr. A provided data that helped identify dysfunctional verbal exchanges and provided the data base for judging the effects of communication training. However, it was also beneficial for the clinician to "passively" listen to the couple discuss their marital problems. Qualitative observation can be a rich source of ideas about the beliefs, attitudes, and behaviors that may contribute to marital distress.

Qualitative analyses can also promote the development of the behavioral assessment paradigm. We have only an elementary understanding of the causes of behavior disorders and of the best methods of treating them. An exclusive reliance on quantification can impair the evolution of any nascent assessment paradigm. Consequently, we must adopt a Steinbeckian attitude—generate and consider new ideas about functional relationships, presume that a new idea may be true, and then rigorously examine it.

Although I am advocating that qualitative methods can contribute to behavioral assessment, scientific methods remain the core element in the behavioral assessment paradigm. The stagnant nature of many psychological construct systems can be attributed to their focus on a rigidly invoked core set of assumptions about the nature of behavior and treatment, rather than on a core set of beliefs about the best way to learn about behavior and treatment. The behavioral assessment paradigm will continue to evolve to the degree that it emphasizes scientific methods for studying behavior, rather than a set of *a priori* beliefs about the nature of behavior disorders and their treatment.

4.07.5 BEHAVIORAL ASSESSMENT METHODS

There are many methods of behavioral assessment. Some, such as self-monitoring and behavioral observation, are congruent with and influenced by the conceptual and methodological elements of the behavioral assessment paradigm outlined earlier in this chapter. Others, such as trait-focused self-report questionnaires, are less congruent with the behavioral assessment paradigm. Surveys of journal publications and of the assessment methods used by practicing behavior therapists show that it is difficult to reliably categorize an assessment instrument as "behavioral" or "nonbehavioral"—categories of behavioral and nonbehavioral assessment methods are becoming increasingly indistinct. For example, many cognitive assessment instruments used by behavior therapists are aggregated and trait-based. They provide an unconditional "score" of some cognitive construct such as "self-efficacy beliefs" or "locus of control" (see discussion of cognitive assessment in Linscott & DiGiuseppe, 1998). Other assessment instruments and methods used by behavior therapists include neuropsychological assessment, sociometric status evaluation, the Minnesota Multiphasic Personality Inventory and other trait-based personality tests, aggregated mood scales, historically focused interviews, and tests of academic achievement (see Hersen & Bellack, 1998).

There are several bases for the decreasing distinction between behavioral and nonbehavioral assessment methods. First, many variables currently targeted in behavioral assessment (e.g., subjective distress, beliefs, mood) were not the focus of behavior therapy several decades ago. As the response modes in causal models of behavior disorders and the targets of behavioral treatments expanded beyond motor and verbal behavior, the array and focus of assessment instruments used in behavioral assessment expanded correspondingly.

Second, behavioral assessors are less predisposed to immediately reject all traditional self-report assessment instruments. Some well-validated "personality" assessment instruments can provide useful information. However, care should be taken to avoid unwarranted inferences from personality assessment instruments and to insure that, if used, they are part of an assessment program that includes more precisely focused assessment methods (see discussions of behavioral and personality assessment in *Behavior Modification*, 1993, No. 1, and general problems in personality assessment in Heatherton & Weinberger, 1994).

Third, in the 1960s and 1970s many behavior analysts denounced trait concepts and emphasized situational sources of behavior variance. The current person × situation interactional model of behavior variance, discussed earlier, suggests that clinical judgments can sometimes be aided by some trait measures, when used with situationally sensitive assessment instruments.

Fourth, the power of behavioral assessment methods often surpasses their clinical utility. Some behavioral assessment methods, such as behavioral observation, are costly and time-consuming, which decreases their usefulness in clinical assessment (however, as discussed in subsequent sections, many technological advances have enhanced the clinical utility of some behavioral assessment methods).

Finally, behavioral assessors are sometimes insufficiently educated in psychometric principles and in the degree to which assessment strategies match the conceptual and methodological aspects of the behavioral assessment paradigm. For example, behavioral assessors sometimes use an aggregated score from an assessment instrument that has multiple uncorrelated factors. At other times an aggregated score is used when there is important between-situation variance in the measured construct. Also, norm-referenced assessment instruments are sometimes applied when they are not psychometrically appropriate or useful with a particular client (norms may not be available for the client's ethnic group, age, or economic status; Cone, 1996; Haynes & Wai'alae, 1995; Silva, 1993).

The following sections briefly present four categories of behavioral assessment methods: (i) behavioral observation, in the natural environment and in analogue situations; (ii) self-monitoring; (iii) self-report interviews and questionnaires; and (iv) psychophysiological assessment. The specific strategies, conceptual foundations, utility, psychometric properties, disadvantages, technical advancements, and contribution to clinical judgment of each category will be discussed. Coverage of these methods is necessarily limited and more extensive presentations of behavioral assessment methods and instruments can be found in books by Hersen and Bellack (1998), Mash and Terdal (1988), Ollendick and Hersen (1993), and Shapiro and Kratochwil (1988).

4.07.5.1 Behavioral Observation

Behavioral observation involves the time-series observation of molecular, precisely defined behaviors, and environmental events. Usually, an observation session is divided into smaller time samples and the occurrence of discrete events within each time sample is recorded by external observers (Foster & Cone, 1986; Foster, Bell-Dolan, & Burge, 1988; Mash & Hunsley, 1990; Tryon, 1998).

Two observation strategies, and variants of each, are discussed below: behavioral observation in the natural environment and behavioral observation in analogue environments.

4.07.5.1.1 Behavioral observation in the natural environment

Systematic behavioral observation of clients is congruent with most of the underlying assumptions of the behavioral assessment paradigm. Quantitative, minimally inferential data are derived on clients in their natural environment using external observers (nonparticipant observers). Behavior observation systems can be constructed for individual patients, some sources of measurement error can be examined through interrater reliability coefficients, and the acquired data can provide valuable information for the functional analysis and for treatment outcome evaluation. Observation in the natural environments has been used in the assessment of self-injurious, delusional, and hallucinatory behaviors in institutions; eating and drinking in restaurants, bars, and in the home; marital and family interactions in the home; student, teacher, and peer interactions in schools; pain and other health-related behaviors at home and in medical centers; community-related behaviors (e.g., driving, littering); and many other behaviors.

Typically, the client is observed for brief periods (e.g., one hour) in his or her natural environment (e.g., at home) several times in "sensitive" or "high-risk" situations—situations or times with an elevated probability that important behaviors or interactions will occur (e.g., at dinnertime when observing problematic family interactions; at mealtime when observing the social interactions of a psychiatric inpatient). Trained observers record the occurrence/nonoccurrence of specified and carefully defined behaviors within *time-sampling* periods (e.g., 15-second periods). Sequences of events (e.g., sequential interactions between a depressed client and family members) and the duration of events can also be recorded. Observers can also use momentary time sampling and record the behaviors that are occurring at predetermined time sampling points. An example of this latter sampling strategy would be a nurse recording the social interactions of a psychiatric inpatient at the beginning of every hour.

Observation in unrestricted environments (e.g., in a client's home) can be problematic because clients are sometimes out of sight of the observers and sometimes engage in behaviors that are incompatible with the purposes of the observation (e.g., talking on the phone; watching TV). Consequently, the behavior of the individual to be observed is often constrained. For example, a marital couple being observed at home might be requested to remain within sight of the observer, to avoid long phone conversations, and keep the TV off. Such constraints compromise the generalizability of the obtained data but increase the efficiency of the observation process.

The temporal parameters (the length, frequency, spacing, and sampling intervals) of an observation session are influenced by the characteristics of the targeted behaviors. For example, higher rate behaviors require shorter time sampling intervals. Highly variable behaviors required more observation sessions. Suen and Ary (1989) discuss temporal parameters of behavioral observation in more detail.

Behavioral observation in the natural environment has many applications. It is a powerful method of treatment outcome evaluation because it minimizes many sources of measurement error associated with self-report and is sensitive to behavior change. It has also been used as a basic research tool in the behavioral and social sciences and to gather data for functional analyses.

The clinical utility of behavioral observation in the natural environment is limited in several ways. It is not cost-efficient for the assessment of very low frequency behaviors, such as stealing, seizures, and some aggressive behaviors. Also, "internalized," covert behavior problems such as mood disorders and anxiety are less amenable to external observation (although some verbal and motoric components of these disorders can be observed). Socially sensitive or high-valence behaviors, such as sexual dysfunctions, paraphiliac behaviors, substance use, and marital violence may not be emitted in the presence of observers. In most outpatient clinical settings, observation in the natural environment with external observers is prohibitively expensive and time-consuming.

Behavioral observation is an expensive assessment method but technological advances have enhanced its clinical utility. Audio and video tape recorders and other instrumentation can facilitate the acquisition of observation data on clients in the natural environment without having to send observers (Tryon, 1991). Observers can also use hand-held computers to record and analyze observation data in real-time (Tryon, 1998).

Because many behaviors can be observed, behavior sampling is an indispensable component of behavioral observation. Typically, observers use a standardized behavior coding system (e.g., Marital Interaction Coding System; Weiss & Summers, 1983) that contains preselected behavior codes. Behaviors selected for inclusion in behavioral observation include: (i) client problem behaviors (e.g., social initiation behaviors by a depressed psychiatric inpatient), (ii) causal variables for the client's behavior problems and goals (e.g., compliments and insults emitted during distressed marital interaction), (iii) behaviors correlated with client problem behaviors (e.g., members of a functional response class such as verbal and physical aggressive behaviors), (iv) salient, important, high-risk behaviors (e.g., suicide talk), (v) client goals and positive alternatives to undesirable behaviors (e.g., positive social interactions by a delusional or socially isolated client), (vi) anticipated immediate, intermediate, and final outcomes of treatment, and (vii) possible positive and negative side- and generalized effects of treatment.

Although observers sometimes focus on only one individual, it is more common in behavioral observation to monitor interpersonal interactions. To this end, observers can record sequences of interactions between two or more individuals (see discussions in Bakeman & Gottman, 1986; Moran et al., 1992). When the goal of observation is to draw inferences about a group of persons, subject sampling can be used. A few persons may be selected for observation from a larger group. For example, several patients may be selected for observation if the goal is to document the effects of a new token program on patients in a psychiatric unit.

Behavioral observation is often considered the "gold standard" for assessment. However, there are several sources of error which can detract from the accuracy and validity of obtained data and of the inferences drawn from them. Sources of error in behavioral observation include: (i) the degree to which observers are trained, (ii) the composition and rotation of observer teams, (iii) observer bias and drift, (iv) the behaviors selected for observation, (v) the specificity of code definitions, (vi) the methods of evaluating interobserver agreement, (vii) the match between time samples and the dynamic characteristics of the observed behaviors, (viii) variance in the situations or time of day in which observation occurs (Alessi, 1988; Hartmann, 1982; Suen & Ary, 1989; Tryon, 1996).

Several types of data can be derived from behavioral observation. Often, assessors are interested in the rate of targeted events. This is usually calculated as the percent of sampling intervals in which a behavior occurs. Sequential analyses and conditional probabilities are often more important for developing functional analyses. For example, observation of family interaction in the home can provide data on negative reciprocity—the relative probability that one family member will respond negatively following a negative (in comparison to a nonnegative) response by another family member. Particularly with computer-aided observation, data on the duration of behaviors can also be acquired. Some observation systems use ratings by observers, rather than event recordings, although these are more frequently conducted outside formal observation session (Segal & Fal, 1998; Spector, 1992). According to McReynolds (1986), the first rating scale was published by Thomasius in 1692. Four basic characterological dimensions were rated—sensuousness, acquisiteveness, social ambition, and rational love.

"Reactivity" refers to the degree to which asessment modifies the targets of assessment. Reactivity is a potential source of inferential error in all assessment procedures, but particularly in behavioral observation (Foster et al., 1988; Haynes & Horn, 1982). The behavior of clients, psychiatric staff members, spouses, and parents may change as a function of whether observers are present or absent. Therefore, reactivity threatens the external validity or situational and temporal generalizability of the acquired data. In the cases of some highly socially sensitive behaviors (e.g., sexual or antisocial behaviors), the magnitude of reactivity may be sufficient to preclude observation in the natural environment.

Participant observation is an alternative to the use of nonparticipant observers. Participant observation is behavioral observation, as described above, using observers who are normally part of the client's natural environment. (In ethnography and cultural anthropology, the term "participant observation" usually refers to qualitative observation by external observers who immerse themselves in the social system they are observing.) Examples of participant observation include: (i) parents observing the play behavior of their children, (ii) nurses observing the delusional speech of psychiatric inpatients, and (iv) a spouse observing a client's depressive or seizure behaviors. Participant observers often use time and behavior sampling procedures similar to those used by nonparticipant observers. However, participant observers are usually less well trained and apply less complex observation systems (e.g., fewer codes, use of momentary time sampling). For example, a staff member on a psychiatric unit might monitor the frequency of social initiations by a client only during short mealtime periods or only briefly throughout the day.

The primary advantages of participant observation are its cost-efficiency and applicability. Participant observation can be an inexpensive method of gathering data on clients in their natural environment. It may be particularly advantageous for gathering data on low frequency behaviors and on behaviors that are highly reactive to the presence of external observers.

There are several potential sources of error in participant observation. First, it is susceptible to most of the sources of error mentioned for nonparticipant observation (e.g., behavior definitions, time sampling errors). Additionally, participant observation may be particularly sensitive to observer biases, selective attention by observers, and recent interactions with the target of observation. The observer is likely to be less well trained and often is not a neutral figure in the social context of the client.

Participant observation may be associated with reactive effects. Sometimes, the reactive effects would be expected to be less for participant than for nonparticipant observation. One determining factor in reactivity is the degree to which the assessment process modifies the natural environment of the client. Because participant observation involves less change in the natural environment of the client, it may be less reactive. However, the reactive effects of participant observation may be strengthened, depending on the method of recording, the behaviors recorded, and the relationship between the observer and client. In some situations participant observation might be expected to alter the monitored behavior to an important degree or to adversely affect the relationship between the observer and target (e.g., an individual monitoring the sexual or eating behavior of a spouse).

Critical event sampling is another infrequently used but clinically useful and cost-efficient method of acquiring natural environment observation data. Critical event sampling involves video or audio tape recording of important interactions in the client's natural environment (Jacob, Tennenbaurm, Bargiel, & Seilhamer, 1995; Tryon, 1998). The critical interactions are later qualitatively or quantitatively analyzed. For example, a tape recorder could be self-actuated by a distressed family during mealtime; a marital couple could record their verbal alterations at home; a socially anxious individual could record conversations while on a date.

4.07.5.1.2 Analogue observation

Analogue observation involves the systematic behavioral observation of clients in carefully structured environments. The assessment environment is arranged to increase the probability that clinically important behaviors and functional relationships can be observed. It is a powerful, clinically useful, and idiographic behavioral assessment method. Many elements of the analogue assessment setting are similar to those of the client's natural environment. However, the participants, social and physical stimuli, and instructions to the client may differ from those of the client's natural environment.

The behavior of clients in analogue assessment is presumed to correlate with their behavior in the natural environment. For example, a distressed marital couple might discuss a problem in their relationship while in a clinic and being observed from a one-way mirror. It is presumed that the problem-solving strategies they use will be similar to those they use outside the clinic.

Analogue observation has many applications. The role play is often used in the behavioral assessment of social skills. A psychiatric patient or socially anxious adult might be observed in a clinic waiting room while attempting to initiate and maintain a conversation with a confederate. A client might be placed in a simulated social situation (e.g., simulated restaurant) and asked to respond to social stimuli provided by a confederate. The Behaviour Avoidance Test (BAT; e.g., Beck & Zebb, 1994) is another form of analogue observation. In a BAT, a client is prompted to approach an object that is feared or avoided.

Analogue methods have been used in the assessment of many clinical phenomena, such as pain (Edens & Gil, 1995), articulated thoughts (Davison, Navarre, & Vogel, 1995), and social anxiety and phobia (Newman, Hofmann, Trabert, Roth, & Taylor, 1994). Other applications include the study of self-injurious behaviors, dental anxiety, stuttering, heterosexual anxiety, alcohol ingestion, panic episodes, cigarette refusal skills, parent–child interaction, speech anxiety, animal phobias, test anxiety, and eating patterns.

Data can be acquired on multiple response modes in analogue observation. For example, during exposure to an anxiety provoking social exchange, clients can report their level of anxiety and discomfort, electrophysiological measures can be taken, observers can record the behavior of the client, and clients can report their thoughts.

Analogue observation is a cost-efficient and multimodal method of assessment and can be a useful supplement to retrospective interview and questionnaire methods. It provides a means of directly observing the client in sensitive situations and of obtaining *in vivo* client reports to supplement the client's retrospective report of thoughts, emotions, and behavior. In comparison to observation in the natural environment, it is particularly useful for observing some important but low-rate events (e.g., marital arguments).

When used in conjunction with systematic manipulation of hypothesized controlling variables analogue observation can be exceptionally useful for identifying causal relationships and for constructing a functional analysis of behavior problems. For example, social attention, tangible rewards, and task demands can be systematically presented and withdrawn before and after the self-injurious behavior of developmentally disabled individuals (e.g., Iwata et al., 1994). Systematic manipulation of hypothesized controlling factors during analogue observation has also been used to identify the cognitive sequelae of hyperventilation during anxiety episodes (Schmidt & Telch, 1994), the most effective reinforcers to use in therapy (Timberlake & Farmer-Dougan, 1991), and the factors associated with food refusal (Munk & Repp, 1994).

Analogue observation is associated with several unique sources of variance, measurement error, and inferential error (e.g., Hughes & Haynes, 1978; Kern, 1991; Torgrud & Holborn, 1992). First, because the physical environment and social stimuli are more carefully controlled in analogue observation but may differ from those in the client's naturalistic environment, external validity may be reduced concomitantly with an increase in behavioral stability. The behavior of clients and the data acquired during analogue observation can covary with: (i) the physical characteristics of the assessment environment; (ii) the instructions to participants; (iii) observer and sampling errors, such as those outlined in naturalistic observation, time and behavior sampling; and (iv) the content validity of the assessment environment (i.e., the degree to which the stimuli are relevant to the construct being measured by the analogue assessment situation).

The primary disadvantage to analogue observation is its representational nature: the data acquired in analogue assessment are only presumed to correlate with data that would be acquired in the natural situations the analogue assessment is designed to represent. It is an indirect measure of the individual's behavior in the natural environment. The results of many studies have supported the discriminant and criterion-related validity of analogue observation; the results of other validation studies have suggested more cautious conclusions. Given the presumption that many behaviors are sensitive to situational sources of variance, clients can be expected to behave differently in analogue and natural environments. That is, the behavior of clients in analogue settings may not accurately reflect or match their behavior in the natural environment. Nevertheless, analogue assessment should be expected to be valid in other ways. For example, socially anxious and nonanxious clients should behave differently during analogue observation even if their behaviors do not match their behaviors in the natural environment. The validity and clinical utility of analogue observation should be considered dependent variables. They are likely to vary across the purposes of the assessment, subjects, target behaviors, settings, and observation methods.

4.07.5.2 Self-monitoring

In self-monitoring, clients systematically record their behavior and sometimes relevant environmental events (Bornstein, Hamilton, & Bornstein, 1986; Gardner & Cole, 1988; Shapiro, 1984). The events to be recorded by the client are first identified and specified by the client and assessor during an interview. A recording form is developed or selected and the client monitors the designated events, sometimes at designated times or in specified situations. To reduce errors associated with retrospective reports, recording usually occurs immediately before or after the monitored event.

One particularly innovative development is self-monitoring via hand-held computer and computer-assisted data acquisition and analysis (Agras, Taylor, Feldman, Losch, & Burnett, 1990; Shiffman, 1993). Hand-held computers allow the collection of real-time data and simplify the analysis and presentation of self-monitoring data. Computerization should significantly increase the clinical utility of self-monitoring.

Time-sampling is sometimes used with self-monitoring, depending on the characteristics of the behavior. Clients can easily record every occurrence of very low-rate behaviors such as seizures or migraine headaches. However, with high-rate or continuous behaviors, such negative thoughts, blood pressure, and mood, clients recordings may be restricted to specified time periods or situations.

Many clinically important behaviors have been the targets of self-monitoring. These include ingestive behaviors (e.g., eating, caffeine intake, alcohol and drug intake, smoking), specific thoughts (e.g., self-criticisms); physiological phenomena and medical problems (e.g., bruxism, blood pressure, nausea associated with chemotherapy, Raynaud's symptoms, arthritic and other chronic pain, heart rate, seizures); and a variety of other phenomena such as self-injurious behaviors, electricity use, startle responses, sexual behavior, self-care behaviors, exercise, panic and asthma episodes, social anxiety, mood, marital interactions, study time, sleeping patterns, and nightmares.

Many response modes are amenable to measurement with self-monitoring. Clients can monitor overt motor behavior, verbal behavior, subjective distress and mood, emotional responses, occurrence of environmental events associated with their behavior, physiological responses, thoughts, and the qualitative characteristics of their behavior problems (e.g., location of headaches, specific negative thoughts, multimodal aspects of panic episodes). Self-monitoring can also be used to track multiple response parameters: response durations, magnitudes, and frequencies. Of particular relevance for the functional analysis, the client can concurrently monitor behaviors, antecedent events, and consequent events to help identify functional relationships. Johnson, Schlundt, Barclay, Carr-Nangle, and Engler (1995), for example, had eating disordered clients monitor the occurrence of binge eating episodes and the social situations in which they occurred in order to calculate conditional probabilities for binge eating.

Self-monitoring is a clinically useful assessment method. Self-monitoring can be tailored for individual clients and used with a range of behavior problems. It is an efficient and inexpensive assessment method for gathering data on functional relationships in the natural environment and is another important supplement to retrospective self-report. It is suitable for time-series assessment and for the derivation of quantitative indices of multiple response modes. Self-monitoring is applicable with many populations—adult outpatients, children, inpatients, parents and teachers, and developmental disabled individuals. Events that, because of their frequency or reactive effects, are not amenable to observation by participant and nonparticipant observers may be more amenable to assessment with self-monitoring.

Although many validation studies on self-monitoring have been supportive (see reviews in Bornstein et al., 1986; Gardner & Cole, 1988; Shapiro, 1984) there are several threats to the validity of this assessment method. Two important sources of error in self-monitoring are clients' recording errors and biases. The resultant data can reflect the client's abilities to track and record behaviors, client expectancies, selective attention, missed recording periods, the social valence and importance of the target behaviors, fabrication, and the contingencies associated with the acquired data. Data can also be affected by how well the client was trained in self-monitoring procedures, the demands of the self-monitoring task, the degree to which target events have been clearly specified, reactions from family and friends to the client's monitoring, and the frequency and duration of the targeted behaviors. Clinical researchers have also reported difficulty securing cooperation from clients to self-monitor for extended periods of time. One particularly powerful source of inferential error is reactivity (Bornstein et al., 1986). The reactive effects of self-monitoring are frequently so great that self-monitoring is sometimes used as a method of treatment.

4.07.5.3 Psychophysiological Assessment

Psychophysiological measurement is an increasingly important method in behavioral assessment (Haynes, Falkin, & Sexton-Radek, 1989). The increased emphasis on psychophysiological assessment is due, in part, to a growing recognition of the importance of physiological components of behavior problems, such as depression, anxiety, and many psychotic behavior problems. Also, behavior therapists are increasingly involved in the assessment and treatment of disorders that have traditionally been the focus of medical interventions—cancer, chronic pain, diabetes, cardiovascular disorders. A third reason for the importance of psychophysiological assessment is that many behavioral intervention procedures, such as relaxation training and desensitization, focus partly on the modification of physiological processes. Advances in ambulatory monitoring, computerization, and other technologies have increased the clinical utility of psychophysiological measurement. Finally, psychophysiological measurement can easily be combined with other behavioral assessment methods, such as self-monitoring and analogue observation.

The recognition of the importance of the physiological response mode in behavior problems mandates the inclusion of electrophysiological and other psychophysiological measurement methods. Electromyographic, electrocardiovascular, electroencephalographic, and electrodermal measures are particularly applicable to behavioral assessment with adults. A range of behavior problems (e.g., panic disorders, PTSD, schizophrenic behaviors, obsessive-compulsive behaviors, worry, depression, substance abuse, disorders of initiating and maintaining sleep) have important physiological components. The low magnitude of covariance between physiological and other response modes, noted earlier in this chapter, suggests that they may be a function of different causal variables and respond differently to the same treatment.

Psychophysiological measurement is a complex, powerful, and clinically useful assessment method in many assessment contexts and for many clients. It is amenable to idiographic assessment, can be applied in a time-series format, and generates quantitative indices. The validity of the obtained measures can be affected by electrode placement, site resistance, movement, instructional variables, time-sampling parameters, data reduction and analysis, equipment intrusiveness, and equipment failures. Books by Andreassi (1995) and Cacioppo and Tassinary (1990) cover instrumentation, measurement methods, technological innovations, clinical applications, and sources of measurement error.

4.07.5.4 Self-report Methods in Behavioral Assessment

Many interview formats and hundreds of self-report questionnaires have been adopted by behavioral assessors from other assessment paradigms. A comprehensive presentation of these methods is not possible within the confines of this chapter. Here, I will emphasize how behavioral and traditional self-report methods differ in format and content. The differences reflect the contrasting assumptions of behavioral and nonbehavioral assessment paradigms.

More extensive discussions of self-report questionnaire and interview methods, and applicable psychometric principles are provided by Anastasi (1988), Jensen and Haynes (1986), Nunnally and Bernstein (1994), Sarwer and Sayers (1998), and Turkat (1986)

Behavioral assessors, particularly those affiliated with an applied behavior analysis paradigm, have traditionally viewed self-report questionnaires and interviews with skepticism. Objections have focused on the content and misuses of these methods. Many questionnaires solicit retrospective reports, stress situationally insensitive aggregated indices of traits, focus on molar level constructs that lack consensual validity, and are unsuited for idiographic assessment. Biased recall, demand factors, item interpretation errors, and memory lapses, further challenged the utility of self-report questionnaires. Data from interviews have been subject to the same sources of error with additional error variance associated with the behavior and characteristics of the interviewer. Despite these constraints, interviews and questionnaires are the most frequently used methods used by behavior therapists (e.g., Piotrowski & Zalewski, 1993).

The interview is an indispensable part of behavioral assessment and treatment and undoubtedly is the most frequently used assessment method. All behavioral interventions require prior verbal interaction with the client or significant individuals (e.g., staff) and the structure and content of that interview can have an important impact on subsequent assessment and treatment activities.

As illustrated with Mrs. A, an assessment interview can be used for multiple purposes. First, it can help identify and rank order the client's behavior problems and goals. It can also be a source of information on the client's reciprocal interactions with other people, and, consequently, provides important data for the functional analysis. Interviews are the main vehicles for informed consent for assessment and therapy and can help establish a positive relationship between the behavior assessor and client. Additionally, interviews are used to select clients for therapy, to determine overall assessment strategies, to gather historical information, and to develop preliminary hypotheses about functional relationships relevant to the client's behavior problems and goals.

The behavioral assessment interview differs from nonbehavioral interviews in content and format. First, the behavioral interview is often more quantitatively oriented and structured (although most behavioral interviews involve unstructured, nondirective, and client-centered phases). The focus of the behavioral interview reflects assumptions of the behavioral assessment paradigm about behavior problems and causal variables and emphasizes current rather than historical behaviors and determinants. Behavioral interviewers are more likely to query about situational sources of behavioral variance and to seek specification of molecular behaviors and events.

A systems perspective also guides the behavioral assessment interview. The behavioral interviewer queries about the client's extended social network and the social and work environment of care-givers (e.g., the incentives at a psychiatric institution that encourage cooperation by staff members). The interviewer also evaluates the effects that treatment may have on the client's social system—will treatment effect family or work interactions?

Some of the concerns with the interview as a source of assessment information reside with its traditionally unstructured applications. Under unstructured conditions, data derived from the interview may covary significantly with the behavior and biases of the interviewer. However, structured interviews and technological advances in interview methods promise to reduce such sources of error (Hersen & Turner, 1994; Sarwer & Sayers, 1998). Computerization, to guide the interviewer and as an interactive system with clients, promises to reduce some sources of error in the interview process. Computerization can also increase the efficiency of the interview and assist in the summarization and integration of interview-derived data.

Other structured interview aids, such as the Timeline Followback (Sobell, Toneatto, & Sobell, 1994) may also increase the accuracy of the data derived in interviews. In Timeline Followback, memory aids are used to enhance accuracy of retrospective recall of substance use. A calendar is used as a visual aid, with the client noting key dates, long periods in which they abstained or were continuously drunk, and other discreet events associated with substance use.

Some interviews are oriented to the information required for a functional analysis. For example, the Motivation Assessment Scale is used with care-givers to ascertain the factors that may be maintaining or triggering self-injurious behavior in developmentally disabled persons (Durand & Crimmins, 1988).

Questionnaires, including rating scales, self-report questionnaires, and problem inventories, are also frequently used in behavioral assessment; they have probably been frequently used with all adult behavior disorders. Many questionnaires used by behavioral assessors are identical to those used in traditional nonbehavioral psychological assessment. As noted earlier, questionnaires are often adopted by behavioral assessors without sufficient thought to their underlying assumptions about behavior and the causes of behavior problems, content validity, psychometric properties, and incremental clinical utility. They are often trait-focused, insensitive to the conditional nature of the targeted behavior, and provide aggregated indices of a multifaceted behavioral construct (Haynes & Uchigakiuchi, 1993). Questionnaires are sometimes helpful for initial screening or as a nonspecific index of program outcome but are not useful for a functional analysis or for precise evaluation of treatment effects. The integration of personality and assessment is addressed further in a subsequent section of this chapter.

Some questionnaires are more congruent with the behavioral assessment paradigm. These usually target a narrower range of adult behavior problems or events, such as panic and anxiety symptoms, outcome expectancies for alcohol, recent life stressors, and tactics for resolving interpersonal conflicts. Most behaviorally oriented questionnaires focus on specific and lower-level behaviors and events and query about situational factors. However, the developers of behaviorally oriented questionnaires sometimes rely on the face validity of questionnaires and do not follow standard psychometric principles of questionnaire development (see special issue on "Methodological issues in psychological assessment research" in *Psychological Assessment*, 1995, Vol. 7). Deficiencies in the development and validation of any assessment instrument reduce confidence in the inferences that can be drawn from resulting scores.

Questionnaires, given appropriate construction and validation, can be an efficient and useful source of behavioral assessment data. Most are inexpensive, quick to administer and score, and are well received by clients. Computer administration and scoring can increase their efficiency and remove several sources of error (Honaker & Fowler, 1990). They can be designed to yield data on functional relationships of variables at a clinically useful level of specificity.

4.07.5.5 Psychometric Foundations of Behavioral Assessment

The application of psychometric principles to behavioral assessment instruments has been discussed in many books and articles (e.g., Cone, 1988; 1996; Foster & Cone, 1995; Haynes & Wai'alae, 1995; Silva, 1993; see also "Methodological issues in psychological assessment research," *Psychological Assessment*, September, 1995). Psychometric principles were originally applied to tests of academic achievement, intelligence, and abilities. Because many of the principles were based on estimating measurement error with presumably stable and molar-level phenomena, the relevance of psychometric principles to behavioral assessment has been questioned. However, "psychometry" is best viewed as a general validation process that is applicable to any method or instrument of psychological assessment.

The ultimate interest of psychometry is the construct validity of an assessment instrument or, more precisely, the construct validity of the data and inferences derived from an assessment instrument. Construct validity is comprised of the multiple lines of evidence and rationales supporting the trustworthiness of assessment instrument data interpretation (Messick, 1993). Indices of construct validity are also conditional—an index of validity does not reside unconditionally with the instrument (Silverman & Kurtines, 1998, discuss contextual issues in assessment). Elements of construct validation are differentially applicable, depending on the method, target, and purpose of assessment.

The validity of data derived from an assessment instrument establishes the upper limit of confidence in the clinical judgments to which the instrument contributes. Consequently, the validity of every element of the functional analysis is contingent on the validity of the assessment instruments used to collect contributing data. The validity of other clinical judgments (e.g., risk factors for relapse, and the degree of treatment effectiveness) similarly depends on the validity of assessment data.

The applicability of psychometric principles (e.g., internal consistency, temporal stability, content validity, criterion-related validity) to behavioral assessment instruments varies with their methods, targets, and applications. The data obtained in behavioral assessment differ in the degree to which they are presumed to measure lower-level less-inferential variables (e.g., number of interruptions in a conversation, hitting) or higher-level more inferential variables (e.g., positive communication strategies, aggression). With lower-level variables, psychometric indices such as internal consistency and factor structure are not useful indications of validity of the obtained data. Interobserver agreement and content validity may be more useful indices.

The validity of data from an assessment instrument depends on how it will be used—on the clinical judgments that it affects. For example, accurate data may be obtained from analogue observation of clients social interactions. That is, there may be perfect agreement among multiple observers about the client's rate of eye contact, questions, and reflections. However, those rates may demonstrate low levels of covariance (i.e., low criterion-referenced validity) with the same behaviors measured in natural settings. The relative importance of accuracy and other forms of validity varies with the purpose of the assessment (see Cone, 1998). If the analogue data is used to evaluate the effectiveness of a social skills training program, accuracy is an important consideration. If the data is to be used to evaluate generalization of treatment effects, accuracy is necessary but not sufficient.

The interpretation of temporal and situational stability coefficients is complicated in behavioral assessment by the conditional and unstable nature of some of the targeted phenomena (e.g., appetitive disorders, social behaviors, mood, expectancies). Indices of instability (e.g., low test–retest correlations) can reflect variability due to true change across time in the variable (e.g., change in the social behavior of an observed client) as well as measurement error (e.g., poorly defined behavior codes, observer error). Consequently, temporal stability coefficients, by themselves are weak indices of validity. A multimethod/multi-instrument assessment strategy, by providing indices of covariance among measures of the same targeted phenomena, however, can help separate true from error variance. Additionally, a low magnitude of temporal stability in a time-series measurement strategy has implications for the number of samples necessary to estimate or capture the time course of the measured phenomena—unstable phenomena require more sampling periods than do stable phenomena.

Behavioral assessment often involves multiple methods of assessment, focused on multiple modes and parameters. As noted earlier in this chapter, sources of measurement error and determinants can vary across methods, modes, and parameters. A multimethod approach to assessment can strengthen confidence in subsequent clinical judgments. However, estimates of covariance are often used as indices of validity and can be attenuated in comparison to monomethod or monomode assessment strategies (see discussion of psychometric indices of multiple methods in Rychtarik & McGillicuddy, 1996).

The individualized nature of behavioral assessment enhances the importance of some construct validity elements. For example, accuracy, content validity, and interobserver agreement are important considerations in behavioral observation coding systems. Idiographic assessment reduces the importance of construct validity elements such as nomothetically based discriminant and convergent validity.

4.07.6 BEHAVIORAL AND PERSONALITY ASSESSMENT

As noted earlier in this chapter, behavioral assessors often use traditional personality questionnaires and several possible reasons for this integration were given. The positive cost-efficiency of personality trait measures is one factor. One of the more empirically based rationales for integration is a person × situation interactional model for assessment: if we want to predict a person's behavior, it helps to know something about the relatively stable aspects of the person and something about the situations that promote instability, at least sometimes (McFall & McDonel, 1986). Personality questionnaires are often used in initial screening, followed by more specifically focused, molecular, and less inferential assessment instruments. Noted in this section are several additional issues concerning the integration of personality and behavioral assessment. These issues were discussed in Haynes and Uchigakiuchi (1993) and in other articles in a special section of *Behavior Modification*, 1993, *17*(1).

There are several complications associated with adopting the situation × person interaction model and with the use of personality assessment instruments. Given that there are hundreds of traits measurable by extant instruments, it is difficult to determine which traits to measure, and how best to measure them. Also, despite a growing literature on situational factors in behavior disorders, we still do not know which aspects of situations can be most important in controlling behavioral variance for a particular client (e.g., Kazdin, 1979). Nor do we know under which conditions a person–situation interaction model, as opposed to a situational or trait model, will assume the greatest predictive efficacy.

Several additional issues regarding the trait × situation interactional model of behavior and the utility of personality assessment strategies for behavioral assessment have already been discussed in this chapter and in many previously published articles. First, personality traits vary in theoretical connotations and the theoretical connotations of a trait measure influence its utility for behavioral assessment. Many constructs measures by personality assessment instruments have psychodynamic and intrinsically causal connotations. Some, such as "emotional instability," "hardiness," and "passive–aggressive," refer to an internal state that is presumed to control observed behavior—these traits invoke causal models that are inconsistent with aspects of the behavioral assessment paradigm. In these cases "psychological processes" are inferred from cross-situational consistencies in behavior. In a circular fashion, the processes become explanations for the behaviors that are their indicators. The processes cannot be independently validated, are difficult to measure, and the inferential process can inhibit a scientific investigation of these behaviors.

Personality questionnaires invariable invoke molar-level traits whose interpretation require normative comparison. Consequently, trait measures are less amenable to idiographic assessment of lower-level variables. Clinical inferences about a person on a trait dimension are derived by comparing the person's aggregated trait score to the trait scores of a large sample of persons. Such comparative inferences can be helpful but can also be in error if there are important differences between the person and the comparison group, such as on dimensions of gender, ethnicity, and age.

Most behavioral assessors would acknowledge that molar self-report measures can contribute to clinical inferences, when used within a multimethod assessment program and care is taken to address many sources of measurement and inferential error noted above. However, there are several other complications associated with personality assessment: (i) many traits measured by personality assessment instruments are poorly defined and faddish; (ii) molar variables are less likely than molecular variables to reflect the dynamic nature of behavior—they are momentary snap-shots of unstable phenomena; (iii) personality trait measures may be more useful for initial screening than for the construction of a detailed functional analysis and treatment planning; (iv) personality traits can also be conditional: their probability and magnitude can vary across situations; (v) inferences about a client's status on a trait dimension varies across assessment instruments; and (vi) because of their aggregated nature, many response permutations can contribute to a particular score on a trait dimension.

In sum, the integration of person–situation interactional models and personality assessment in the behavioral assessment paradigm can benefit clinical judgments. However, this integration has sometimes occurred too readily, without the thoughtful and scholarly reflection characteristics of the behavioral assessment paradigm.

4.07.7 SUMMARY

Behavioral assessment is a dynamic and powerful assessment paradigm designed to enhance the validity of clinical judgments. One of the most important and complex clinical judgments in behavioral assessment functional analysis—a synthesis of the clinicians hypotheses about the functional relationships relevant to a clients behavior problems. The functional analysis is a central component in the design of behavior therapy programs. The behavioral assessment paradigm suggests that errors in clinical judgments can be reduced to the degree that the judgments are based on multiple assessment methods and sources of information, validated assessment instruments, time-series measurement strategies, data on multiple response modes and parameters, minimally inferential variables, and the assessment of behavior–environment interactions. The Clinical Pathogenesis Map and Functional Analytic Causal Model were introduced as ways of graphically depicting and systematizing the functional analysis.

The methods of behavioral assessment and clinical case conceptualizations are influenced by several interrelated assumptions about the causes of behavior problems. The behavioral assessment paradigm emphasizes multiple causality; multiple causal paths; individual differences in causal variables and paths; environmental causality and reciprocal determinism; contemporaneous causal variables; the dynamic nature of causal relationships; the operation of moderating and mediating variables; interactive and additive causality; situations, setting events, and systems factors as causal variables; and dynamical causal relationships.

The methods of behavioral assessment and clinical case conceptualizations are also affected by assumptions about the characteristics of behavior problems. These include an emphasis on the multimodal and multiparameter characteristics of behavior problems, differences among clients in the importance of behavior problem modes and parameters, the complex interrelationships among a client's multiple behavior problems, and the conditional and dynamic natures of behavior problems.

Three of many methodological foundations of behavioral assessment were discussed: the emphasis on empirical hypothesis-testing, the idiographic emphasis, and the use of time-series assessment strategies.

The decreasing distinctiveness of behavioral and nonbehavioral assessment, and reasons for this change, were discussed. Four caegories of behavioral assessment methods were presented (i) behavioral observation, (ii) self-monitoring, (iii) self-report methods, and (iv) psychophysiological assessment. The specific strategies, conceptual foundations, clinical utility, psychometric properties, disadvantages, technical advancements, and contribution to clinical judgment of each category were presented.

The application of psychometric principles to behavioral assessment was discussed. The applicability of specific principles varies across methods, targets, and applications.

Several issues relating to the integration of behavioral and personality assessment were presented. These included poor definitions for some traits, the molar nature of personality assessment variables, insensitivity to dynamic aspects of behavior, reduced utility for functional analysis and treatment planning, the conditional nature of personality traits, differences among personality assessment instruments, and the aggregated nature of trait measures.

4.07.8 REFERENCES

Agras, W. S., Taylor, C. B., Feldman, D. E., Losch, M., & Burnett, K. F. (1990). Developing computer-assisted therapy for the treatment of obesity. *Behavior Therapy, 21,* 99–109.

Alessi, G. (1988). Direct observation methods for emotional/behavior problems. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Behavioral assessment in schools: Conceptual foundations and practical applications* (pp. 14–75). New York: Guilford Press.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.

Andreassi, J. L. (1995). *Psychophysiology: Human behavior and physiological response* (3rd ed.). Hillsdale, NJ: Erlbaum.

Asterita, M. F. (1985). *The physiology of stress.* New York: Human Sciences Press.

Bakeman, R., & Gottman, J. M. (1986). *Observing interaction: An introduction to sequential analysis.* New York: Cambridge University Press.

Bandura, A. (1969). *Principles of behavior modification.* New York: Holt, Rinehart and Winston.

Bandura, A. (1981). In search of pure unidirectional determinants. *Behavior Therapy, 12,* 315–328.

Barlow, D. H., & Cerny, J. A. (1988). *Psychological treatment of panic.* New York: Guilford Press.

Barnett, P. A., & Gotlib, I. H. (1988). Psychosocial functioning and depression: Distinguishing among antecedents, concomitants, and consequences. *Psychological Bulletin, 104,* 97–126.

Barrios, B. A. (1988). On the changing nature of behavioral assessment. In A. S. Bellack & M. Hersen (Eds.), *Behavioral assessment: A practical handbook* (pp. 3–41). New York: Pergamon.

Beach, S., Sandeen, E., & O'Leary, K. D. (1990). *Depression in marriage.* New York: Guilford Press.

Beck, J. G., & Zebb, B. J. (1994). Behavioral assessment and treatment of panic disorder: Current status, future directions. *Behavior Therapy, 25,* 581–612.

Bellack, A. S., & Hersen, M. (1988). *Behavioral assessment: A practical handbook.* New York: Pergamon.

Bornstein, P. H., Bornstein, M. T., & Dawson, D. (1984). Integrated assessment and treatment. In T. H. Ollendick & M. Hersen (Eds.), *Child behavioral assessment: Principles and procedures* (pp. 223–243). New York: Pergamon.

Bornstein, P. H., Hamilton, S. B., & Bornstein, M. T. (1986). Self-monitoring procedures. In A. R. Ciminero, C. S. Calhoun, & H. E. Adams (Eds.), *Handbook of behavioral assessment* (pp. 176–222). New York: Wiley.

Brown, T. A., DiNardo, P. A., & Barlow, D. H. (1994). *Anxiety disorders interview schedule for DSM-IV (ADIS-IV).* Albany, NY: Graywind Publications.

Cacioppo, J. T., & Tassinary, L. G. (1990). *Principles and psychophysiology: Physical, social, and inferential elements.* New York: Cambridge University Press.

Chadwick, P. D. J., Lowe, C. F., Horne, P. J., & Higson, P. J. (1994). Modifying delusions: The role of empirical testing. *Behavior Therapy, 25,* 35–49.

Ciminero, A. R., Calhoun, K. S., & Adams, H. E. (1986). *Handbook of behavioral assessment.* New York: Wiley.

Collins, L. M., & Horn, J. L. (Eds.) (1991). *Best methods for the analysis of change.* Washington, DC: American Psychological Association.

Cone, J. D. (1979). Confounded comparisons in triple response mode assessment research. *Behavioral Assessment, 11,* 85–95.

Cone, J. D. (1986). Idiographic, nomothetic and related perspectives in behavioral assessment. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment.* (pp. 111–128). New York: Guilford Press.

Cone, J. D. (1988). Psychometric considerations and the multiple models of behavioral assessment. In A. S. Bellack & M. Hersen (Eds.), *Behavioral assessment: A practical handbook* (pp. 42–66). New York: Pergamon.

Cone, J. D. (1998). Psychometric considerations: Concepts, contents and methods. In M. Hersen & A. S. Bellack (Eds.), *Behavioral assessment: A practical handbook* (4th ed.). Boston: Allyn & Bacon.

Cone, J. D., & Hawkins, R. P. (Eds.) (1977). *Behavioral assessment: New directions in clinical psychology*. New York: Brunner/Mazel.

Craske, M. G., & Waikar, S. V. (1994). Panic disorder. In M. Hersen & R. T. Ammerman (Eds), *Handbook of prescriptive treatments for adults*. (pp. 135–155). New York: Plenum.

Creer, T. L., & Bender, B. G. (1993). Asthma. In R. J. Gatchel & E. B. Blanchard (Eds.), *Psychophysiological disorders, research and clinical applications* (pp. 151–204) Washington, DC: American Psychological Association.

Davison, G. C., Navarre, S., & Vogel, R. (1995). The articulated thoughts in simulated situations paradigm: A think-aloud approach to cognitive assessment. *Current Directions in Psychological Science, 4,* 29–33.

de Beurs, E., Van Dyck, R., van Balkom, A. J. L. M., Lange, A., & Koele, P. (1994). Assessing the clinical significance of outcome in agoraphobia research: A comparison of two approaches. *Behavior Therapy, 25,* 147–158.

Durand, V. M., & Carr, E. G. (1991). Functional communication training to reduce challenging behavior: Maintenance and application in new settings. *Journal of Applied Behavior Analyses, 24,* 251–264.

Durand, V. M., & Crimmins, D. M. (1988). Identifying the variables maintaining self-injurious behaviors. *Journal of Autism and Developmental Disorders, 18,* 99–117.

Eels, T. (1997). *Handbook of psychotherapy case formulation*. New York: Guilford Press.

Eifert, G. H., & Wilson, P. H. (1991). The triple response approach to assessment: A conceptual and methodological reappraisal. *Behaviour Research and Therapy, 29,* 283–292.

Edens, J. L., & Gil, K. M. (1995). Experimental induction of pain: Utility in the study of clinical pain. *Behavior Therapy, 26,* 197–216.

Evans, I. (1993). Constructional perspectives in clinical assessment. *Psychological Assessment, 5,* 264–272.

Eysenck, H. J. (1986). A critique of contemporary classification and diagnosis. In T. Millon & G. L. Klerman (Eds.), *Contemporary directions in psychopathology: Toward the DSM-IV* (pp. 73–98). New York: Guilford Press.

Eysenck, H. J., & Martin, I. (1987). *Theoretical foundations of behavior therapy*. New York: Plenum.

Figley, C. R. (Ed.) (1979), *Trauma and its wake: Volume 1: The study of post-traumatic stress disorder*. New York: Brunner/Mazel.

Foster, S. L., Bell-Dolan, D. J. & Burge, D. A. (1988). Behavioral observation. In A. S. Bellack & M. Hersen (Eds.), *Behavioral assessment: A practical handbook* (pp. 119–160). New York: Pergamon.

Foster, S. L., & Cone, J. D. (1986). Design and use of direct observation systems. In A. R. Ciminero, C. S. Calhoun, & H. E. Adams (Eds.), *Handbook of behavioral assessment* (pp. 253–324). New York: Wiley.

Gannon, L. R., & Haynes, S. N. (1987). Cognitive-physiological discordance as an etiological factor in psychophysiologic disorders. *Advances in Behavior Research and Therapy, 8,* 223–236.

Gardner, W. I., & Cole, C. L. (1988). Self-monitoring procedures. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Behavioral assessment in schools: Conceptual foundations and practical applications* (pp. 206–246). New York: Guilford Press.

Gatchel, R. J. (1993). Psychophysiological disorders: Past and present perspectives. In R. J. Gatchel & E. B. Blanchard (Eds.), *Psychophysiological disorders, research and clinical applications* (pp. 1–22). Washington, DC: American Psychological Association.

Gatchel, R. J., & Blanchard, E. B. (1993). *Psychophysiological disorders, research and clinical applications*. Washington, DC: American Psychological Association.

Glass, C. (1993). A little more about cognitive assessment. *Journal of Counseling and Development, 71,* 546–548.

Goldfried, M. R. (1982). Behavioral Assessment: An overview. In A. S. Bellack, M. Hersen, & A. E. Kazdin (Eds.), *International handbook of behavior modification and therapy* (pp. 81–107). New York: Plenum.

Haberman, S. J. (1978). *Analysis of qualitative data* (Vol. 1). New York: Academic Press.

Hartmann, D. P. (Ed.) (1982). *Using observers to study behavior*. San Francisco: Jossey-Bass.

Hatch, J. P. (1993). Headache. In: R. J. Gatchel & E. B. Blanchard (Eds.), *Psychophysiological disorders, research and clinical applications* (pp. 111–150) Washington, DC: American Psychological Association.

Haynes, S. N. (1978). *Principles of behavioral assessment*. New York: Gardner Press.

Haynes, S. N. (1986). The design of intervention programs. In R. O. Nelson & S. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 386–429). New York: Guilford Press.

Haynes, S. N. (1992). *Models of causality in psychopathology: Toward synthetic, dynamic and nonlinear models of causality in psychopathology*. Des Moines, IA: Allyn & Bacon.

Haynes, S. N. (1994). Clinical judgment and the design of behavioral intervention programs: Estimating the magnitudes of intervention effects. *Psichologia Conductual, 2,* 165–184.

Haynes, S. N. (1996a). Behavioral assessment of adults. In M. Goldstein and M. Hersen (Eds.), *Handbook of psychological assessment*.

Haynes, S. N. (1996b). The changing nature of behavioral assessment. In: M. Hersen & A. Bellack (Eds.), *Behavioral assessment: A practical guide* (4th ed.).

Haynes, S. N. (1996c). The assessment–treatment relationship in behavior therapy: The role of the functional analysis. *The European Journal of Psychological Assessment*. (in press).

Haynes, S. N., Blaine, D., & Meyer, K. (1995). Dynamical models for psychological assessment: Phase–space functions. *Psychological Assessment, 7,* 17–24.

Haynes, S. N., Falkin, S., & Sexton-Radek, K. (1989). Psychophysiological measurement in behavior therapy. In G. Turpin (Ed.), *Handbook of clinical psychophysiology* (pp. 263–291). London: Wiley.

Haynes, S. N., & Horn, W. F. (1982). Reactive effects of behavioral observation. *Behavioral Assessment, 4,* 369–385.

Haynes, S. N., Leisen, M. B., & Blaine, D. D. (1997). Design of individualized behavioral treatment programs using functional analytic clinical case models. *Psychological Assessment, 9,* 334–348.

Haynes, S. N., & O'Brien, W. O. (1990). The functional analysis in behavior therapy. *Clinical Psychology Review, 10,* 649–668.

Haynes, S. N., & O'Brien, W. O. (1998). *Behavioral assessment. A functional approach to psychological assessment*. New York: Plenum.

Haynes, S. N., & O'Brien, W. O. (in press). *Behavioral assessment*. New York: Plenum.

Haynes, S. N., Spain, H., & Oliviera, J. (1993). Identifying causal relationships in clinical assessment. *Psychological Assessment, 5,* 281–291.

Haynes, S. N., & Uchigakiuchi, P. (1993). Incorporating personality trait measures in behavioral assessment:

Nuts in a fruitcake or raisins in a mai tai? *Behavior Modification, 17,* 72–92.

Haynes, S. N., Uchigakiuchi, P., Meyer, K., Orimoto, Blaine, D., & O'Brien, W. O. (1993). Functional analytic causal models and the design of treatment programs: Concepts and clinical applications with childhood behavior problems. *European Journal of Psychological Assessment, 9,* 189–205.

Haynes, S., N., & Wai'alae, K. (1995). Psychometric foundations of behavioral assessment. In: R. Fernández-Ballestros (Ed.), *Evaluacion conductual hoy: (Behavioral assessment today)*(pp. 326–356). Madrid, Spain: Ediciones Piramide.

Haynes, S. N., & Wu-Holt, P. (1995). Methods of assessment in health psychology. In M. E. Simon (Ed.), *Handbook of health psychology* (pp. 420–444). Madrid, Spain: Sigma

Heatherton, T. F., & Weinberger, J. L. (Eds.) (1994). *Can personality change*. Washington, DC: American Psychological Association.

Hersen, M., & Bellack, A. S. (Eds.) (1998). *Behavioral assessment: A practical handbook* (4th ed.). Boston: Allyn & Bacon.

Hersen, M., & Turner, S. M. (Eds.) (1994). *Diagnostic interviewing* (2nd ed.). New York: Plenum.

Hughes, H. M., & Haynes, S. N. (1978). Structured laboratory observation in the behavioral assessment of parent–child interactions: A methodological critique. *Behavior Therapy, 9,* 428–447.

Honaker, L. M., & Fowler, R. D. (1990). Computer-assisted psychological assessment. In: G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment*. (pp. 521–546.) New York: Pergamon.

Iwata, B. A. (and 14 other authors). (1994). The functions of self-injurious behavior: An experimental–epidemiological analysis. *Journal of Applied Behavior Analysis, 27,* 215–240.

Jacob, T. Tennenbaurm, D., Bargiel, K., & Seilhamer, R. A. (1995). Family interaction in the home: Development of a new coding system. *Behavior Modification, 12,* 249–251.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12–19.

James, L. D., Thorn, B. E., & Williams, D. A. (1993). Goal specification in cognitive-behavioral therapy for chronic headache pain. *Behavior Therapy, 24,* 305–320.

Jensen, B. J., & Haynes, S. N. (1986). Self-report questionnaires. In A. R. Ciminero, C. S. Calhoun, & H. E. Adams (Eds.), *Handbook of behavioral assessment* (pp 150–175). New York: Wiley.

Johnson, W. G., Schlundt, D. G., Barclay, D. R., Carr-Nangle, R. E., & Engler, L. B. (1995). A naturalistic functional analysis of binge eating. *Behavior Therapy, 26,* 101–118.

Johnston, J. M., & Pennypacker, H. S. (1993). *Strategies and tactics of behavioral research* (2nd ed.). Hillsdale, NJ: Erlbaum.

Kanfer, F. H. (1985). Target selection for clinical change programs. *Behavioral Assessment, 7,* 7–20.

Kazdin, A. E. (1979). Situational specificity: The two edged sword of behavioral assessment. *Behavioral Assessment, 1,* 57–75.

Kazdin, A. (1997). *Research design in clinical psychology* (2nd ed.). New York: Allyn & Bacon.

Kazdin, A. E., & Kagan, J. (1994). Models of dysfunction in developmental psychopathology. *Clinical Psychology: Science and Practice, 1,* 35–52.

Kern, J. M. (1991). An evaluation of a novel role-play methodology: The standardized idiographic approach. *Behavior Therapy, 22,* 13–29.

Kerns, R. D. (1994). Pain management. In M. Hersen & R. T. Ammerman (Eds), *Handbook of prescriptive treatments for adults* (pp 443–461). New York: Plenum.

Kratochwill, T. R., & Levin, J. R. (1992). *Single-case research design and analysis: New directions for psychology and education*. Hillsdale, NJ: Erlbaum.

Kratochwill, T. R., & Shapiro, E. S. (1988). Introduction: Conceptual foundations of behavioral assessment. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Behavioral assessment in schools: Conceptual foundations and practical applications* (pp. 1–13). New York: Guilford Press.

Kubany, E. S. (1994). A cognitive model of guilt typology in combat-related PTSD. *Journal of Traumatic Stress, 7,* 3–19.

Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist, 50,* 519–525.

Lichstein, K. L., & Riedel, B. W. (1994). Behavioral assessment and treatment of insomnia: A review with an emphasis on clinical application. *Behavior Therapy, 25,* 659–688.

Linscott, J., & DiGiuseppe, R. (1998). Cognitive assessment. In M. Hersen & A. S. Bellack (Eds.), *Behavioral assessment: A practical handbook* (4th ed.). Boston: Allyn & Bacon.

Malec, J. F., & Lemsky, C. (1996). Behavioral assessment in medical rehabilitation: Traditional and consensual approaches. In L. Cushman & M. Scherer (Eds.) *Psychological assessment in medical rehabilitation* (pp. 199–236). Washington, DC: American Psychological Association.

Marsella, A. J., & Kameoka, V. (1989). Ethnocultural issues in the assessment of psychopathology. In S. Wetzler (Ed.), *Measuring mental illness: Psychometric assessment for clinicians* (pp. 157–181). Washington, DC: American Psychiatric Association.

Mash, E. J., & Hunsley, J. (1990). Behavioral assessment: A contemporary approach. In A. S. Bellack, M. Hersen, & A. E. Kazdin (Eds.), *International handbook of behavior modification and therapy* (2nd ed., pp. 87–106). New York: Plenum.

Mash, E. J., & Hunsley, J. (1993). Assessment considerations in the identification of failing psychotherapy: Bringing the negatives out of the darkroom. *Psychological Assessment, 5,* 292–301.

Mash, E. J. & Terdal, L. G. (1988). *Behavioral assessment of childhood disorders*. New York: Guilford Press.

McConaghy, N. (1998). Assessment of sexual dysfunction and deviation. In M. Hersen & A. S. Bellack (Eds.), *Behavioral assessment: A practical handbook* (4th ed). Boston: Allyn & Bacon.

McFall, R. M. & McDonel, E. (1986). The continuing search for units of analysis in psychology: Beyond persons, situations and their interactions. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 201–241). New York: Guilford Press.

McReynolds, P. (1986). History of assessment in clinical and educational settings. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 42–80). New York: Guilford Press.

Mischel, W. (1968). *Personality and assessment*. New York: Wiley.

Moran, G., Dumas, J., & Symons, D. K. (1992). Approaches to sequential analysis and the description of contingency tables in behavioral interaction. *Behavioral Assessment, 14,* 65–92.

Munk, D. D., & Repp, A. C. (1994). Behavioral assessment of feeding problems of individuals with severe disabilities. *Journal of Applied Behavior Analysis, 27,* 241–250.

Nelson, R. O., & Hayes, S. C. (1986). *Conceptual foundations of behavioral assessment*. New York: Guilford Press.

Newman, M. G., Hofmann, S. G., Trabert, W., Roth, W. T., & Taylor, C. B. (1994). Does behavioral treatment of social phobia lead to cognitive changes? *Behavior Therapy, 25,* 503–517.

Nezu, A. M., & Nezu, C. M. (1989). *Clinical decision making in behavior therapy: A problem-solving perspective*. Champaign, IL: Research Press.

Nezu, A., Nezu, C., Friedman, & Haynes, S. N. (1996). Case formulation in behavior therapy. In T. D. Eells (Ed.), *Handbook of psychotherapy case formulation*. New York: Guilford Press.

Nunnally, J. C., & Burnstein, I. H. (1994). *Psychometric theory* (3rd ed.) New York: McGraw-Hill.

O'Brien, W. H., & Haynes, S. N. (1995). A functional analytic approach to the conceptualization, assessment and treatment of a child with frequent migraine headaches. *In Session, 1,* 65–80.

O'Donohue, W., & Krasner, L. (1995). *Theories of behavior therapy*. Washington, DC: American Psychological Association.

O'Leary, D. K. (Ed.) (1987). *Assessment of marital discord: An integration for research and clinical practice*. Hillsdale, NJ: Erlbaum.

O'Leary, K. D., Malone, J., & Tyree, A. (1994). Physical aggression in early marriage: Prerelationship and relationship effects. *Journal of Consulting and Clinical Psychology, 62,* 594–602.

O'Leary, K. D., Vivian, D., & Malone, J. (1992). Assessment of physical aggression against women in marriage: The need for multimodal assessment. *Behavioral Assessment, 14,* 5–14.

Ollendick, T. H., & Hersen, M. (1984). *Child behavioral assessment, principles and procedures*. New York: Pergamon.

Ollendick, T. H., & Hersen, M. (1993). *Handbook of child and adolescent assessment*. Boston: Allyn & Bacon.

Persons, J. B. (1989). *Cognitive therapy in practice: A case formulation approach*. New York: Norton.

Persons, J. B., & Bertagnolli, A. (1994). Cognitive-behavioural treatment of multiple-problem patients: Application to personality disorders. *Clinical Psychology and Psychotherapy, 1,* 279–285.

Persons, J. B., & Fresco, D. M. (1998). Assessment of depression. In M. Hersen & A. S. Bellack (Eds.), *Behavioral assessment: A practical handbook* (4th ed.). Boston: Allyn & Bacon.

Piotrowski, C., & Zalewski, C. (1993). Training in psychodiagnostic testing in APA-approved PsyD and PhD clinical psychology programs. *Journal of Personality Assessment, 61,* 394–405.

Regier, D. A., Farmer, M. E., Rae, D. S., Locke, B. Z., Keith, S. J., Judd, L. L., & Goodwin, F. K. (1990). Comorbidity of mental disorders with alcohol and other drug abuse. *Journal of the American Medical Association, 264,* 2511–2518.

Rychtarik, R. G., & McGillicuddy, N. B. (1998). Assessment of appetitive disorders: Status and empirical methods in alcohol, tobacco, and other drug use. In M. Hersen & A. S. Bellack (Eds.), *Behavioral Assessment: A practical handbook (4th ed.)*. Boston: Allyn & Bacon.

Sarwer, D., & Sayers, S. L. (1998). Behavioral interviewing. In M. Hersen, & A. S. Bellack (Eds.), *Behavioral assessment: A practical handbook* (4th ed.). Boston: Allyn & Bacon.

Schmidt, N. B., & Telch, M. J. (1994). Role of fear and safety information in moderating the effects of voluntary hyperventilation. *Behavior Therapy, 25,* 197–208.

Schlundt, D. G., Johnson, W. G., & Jarrell, M. P. (1986). A sequential analysis of environmental, behavioral, and affective variables predictive of vomiting in bulimia nervosa. *Behavioral Assessment, 8,* 253–269.

Segal, D. L., & Fal, S. B. (1998). Structured diagnostic interviews and rating scales. In M. Hersen & A. S. Bellack (Eds.), *Behavioral assessment: A practical handbook* (4th ed.). Boston: Allyn & Bacon.

Shadish, W. R. (1996). Meta-analysis and the exploration of causal mediating processes: A primer of examples, methods, and issues. *Psychological Methods. 1,* 47–65.

Shapiro, E. S. (1984). Self-monitoring. In T. H. Ollendick & M. Hersen (Eds.), *Child behavioral assessment: Principles and procedures* (pp. 350–373). Elmsford, NY: Pergamon.

Shapiro, E. W., & Kratochwill, T. R. (Eds.) (1988). *Behavioral assessment in schools, Conceptual foundations and practical applications*. New York: Guilford Press.

Shiffman, S. (1993). Assessing smoking patterns and motives. *Journal of Consulting and Clinical Psychology, 61*, 732–742.

Silva, F. (1993). *Psychometric foundations and behavioral assessment*. Newbury Park, CA: Sage

Silverman, W. K., & Kurtines, W. M. (1998). Anxiety and phobic disorders: A pragmatic perspective. In M. Hersen, & A. S. Bellack (Eds.), *Behavioral assessment: A practical handbook* (4th ed.). Boston: Allyn & Bacon.

Smith, G. T. (1994). Psychological expectancy as mediator of vulnerability to alcoholism. *Annals of the New York Academy of Sciences, 708*, 165–171.

Smith, G. T., & McCarthy, D. M. (1995). Methodological considerations in the refinement of clinical assessment instruments. *Psychological Assessment, 7*, 300–308.

Sobell, L. C., Toneatto, T., & Sobell, M. B. (1994). Behavioral assessment and treatment planning for alcohol, tobacco, and other drug problems: Current status with an emphasis on clinical applications. *Behavior Therapy, 25*, 523–532.

Spector, P. E. (1992). *Summated rating scale construction: An introduction*. Beverly Hills, CA: Sage.

Sprague, J. R., & Horner, R. H., (1992). Covariation within functional response classes: Implications for treatment of severe problem behavior. *Journal of Applied Behavior Analysis, 25*, 735–745.

Strosahl, K. D., & Linehan, M. M. (1986). Basic issues in behavioral assessment. In A. Ciminero, K. S. Calhoun, & H. E. Adams (Eds.), *Handbook of behavioral assessment* (pp. 12–46). New York: Wiley.

Suen, H. K., & Ary, D. (1989). *Analyzing quantitative observation data*. Hillsdale, NJ: Erlbaum.

Sutker, P. B., & Adams, H. E. (Eds.) (1993). *Comprehensive handbook of psychopathology*. New York: Plenum.

Taylor, C. B., & Agras, S. (1981). Assessment of phobia. In D. H. Barlow (Ed.), *Behavioral assessment of adult disorders* (pp. 280–309). New York: Guilford Press.

Timberlake, W., & Farmer-Dougan, V. A. (1991). Reinforcement in applied settings: Figuring out ahead of time what will work. *Psychological Bulletin, 110*, 379–391.

Torgrud, L. J., & Holborn, S. W. (1992). Developing externally valid role-play for assessment of social skills: A behavior analytic perspective. *Behavioral Assessment, 14*, 245–277.

Turk, D. C., & Salovey, P. (Eds.) (1988). *Reasoning, inference, and judgment in clinical psychology*. New York: Free Press.

Turkat, I. (1986). The behavioral interview. In A. Ciminero, K. S. Calhoun, & H. E. Adams (Eds.), *Handbook of behavioral assessment* (pp. 109–149). New York: Wiley.

Tryon, W. W. (1985). Behavioral assessment in behavioral medicine. New York: Springer.

Tryon, W. W. (1991). *Activity measurement in psychology and medicine*. New York: Plenum.

Tryon, W. (1996). Observing contingencies: Taxonomy and methods. *Clinical Psychology Review* (in press).

Tryon, W. W. (1998). Behavioral observation. In

M. Hersen & A. S. Bellack (Eds.). *Behavioral assessment: A practical handbook* (4th ed.). Boston: Allyn & Bacon. (in press)

Waddell, G., & Turk, D. C. (1992). Clinical assessment of low back pain. In D. C. Turk & R. Melzack (Eds.), *Handbook of pain assessment* (pp. 15–36). New York: Guilford Press.

Weiss, R. L., & Summers, K. J., (1983). Marital interaction coding system—III. In E. E. Filsinger (Ed.), *Marriage and family assessment: A sourcebook for family therapy* (pp. 85–115). Beverly Hills, CA: Sage.

Wolpe, J., & Turkat, I. D. (1985). Behavioral formulation of clinical cases. In I. Turkat (Ed.), *Behavioral cases formulation* (pp. 213–244). New York: Plenum.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.08

Intellectual Assessment

ALAN S. KAUFMAN
Yale University School of Medicine, New Haven, CT, USA
and
ELIZABETH O. LICHTENBERGER
The Salk Institute, La Jolla, CA, USA

4.08.1 INTRODUCTION 188
  4.08.1.1 Brief History of Intelligence Testing 188
  4.08.1.2 Controversy Over Intelligence Testing 190
  4.08.1.3 Principles of the Intelligent Testing Philosophy 192
4.08.2 MEASURES OF INTELLIGENCE 193
  4.08.2.1 Wechsler's Scales 193
    4.08.2.1.1 Wechsler Primary and Preschool Intelligence Scale-Revised (WPPSI-R) 193
    4.08.2.1.2 Wechsler Intelligence Scale for Children-3rd Edition (WISC-III) 195
    4.08.2.1.3 WISC-III Short Form 198
    4.08.2.1.4 Wechsler Adult Intelligence Scale-Revised (WAIS-R) 199
    4.08.2.1.5 Wechsler Adult Intelligence Scale-Third Edition (WAIS-III) 203
    4.08.2.1.6 Kaufman Assessment Battery for Children (K-ABC) 208
    4.08.2.1.7 Kaufman Adolescent and Adult Intelligence Test 211
    4.08.2.1.8 Overview 214
    4.08.2.1.9 The Stanford–Binet: Fourth edition 216
    4.08.2.1.10 Woodcock–Johnson Psycho-Educational Battery-Revised: tests of cognitive ability (WJ-R) 217
    4.08.2.1.11 Detroit Tests of Learning Aptitude (DTLA-3) 220
    4.08.2.1.12 Differential Abilities Scales (DAS) 221
    4.08.2.1.13 Cognitive Assessment System (CAS) 223
4.08.3 INSTRUMENT INTEGRATION 224
  4.08.3.1 K-ABC Integration with Wechsler Scales 224
  4.08.3.2 Integration of KAIT with Wechsler Scales 224
  4.08.3.3 Integration of Binet IV with Wechsler Scales 225
  4.08.3.4 Integration of WJ-R with Wechsler Scales 225
  4.08.3.5 DTLA-3 Integration with Wechsler Scales 226
  4.08.3.6 Integration of DAS with Wechsler Scales 226
  4.08.3.7 Integration of CAS with Wechsler Scales 227
4.08.4 FUTURE DIRECTIONS 227
4.08.5 SUMMARY 228
4.08.6 ILLUSTRATIVE CASE REPORT 229
  4.08.6.1 Referral and Background Information 229
  4.08.6.2 Appearance and Behavioral Characteristics 231
  4.08.6.3 Tests Administered 231
  4.08.6.4 Test Results and Interpretation 232
  4.08.6.5 Summary and Diagnostic Impressions 233
  4.08.6.6 Recommendations 234
4.08.7 REFERENCES 235

4.08.1 INTRODUCTION

The assessment of intellectual ability has grown and continued to develop and flourish since the nineteenth century. This chapter gives a foundation for understanding the progression of intellectual assessment through a brief historical review of IQ testing. Some of the controversy surrounding intelligence testing will be introduced, and an "intelligent" approach to testing (Kaufman, 1979, 1994b) will be discussed in response to the critics of testing. There are multiple available tests for assessing child, adolescent, and adult intelligence, and this chapter will address a select group of measures. A description and brief overview is provided on the following intelligence tests: Wechsler Primary and Preschool Intelligence Scale-Revised (WPPSI-R; Wechsler, 1989), Wechsler Intelligence Scale for Children-Third Edition (WISC-III; Wechsler, 1974), Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981), Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983), Kaufman Adolescent and Adult Intelligence Test (KAIT; Kaufman & Kaufman, 1993), Stanford–Binet: Fourth Edition (Binet IV; Thorndike, Hagen, & Sattler, 1986b), Woodcock-Johnson Psycho-Educational Battery-Revised: Tests of Cognitive Ability (WJ-R; Woodcock & Johnson, 1989), Detroit Tests of Learning Aptitude (DTLA-3; Hammill, 1991), Differential Abilities Scale (DAS; Elliott, 1990), and Das–Naglieri Cognitive Assessment System (CAS; Naglieri & Das, 1996).

A thorough cognitive assessment contains supplemental measures in addition to the main instrument used to measure IQ, as will be discussed later in this chapter. Accordingly, following the description and overview of these multiple instruments is a section which integrates each of them with the Wechsler scales, focusing on how they may be used with Wechsler's tests to further define cognitive functioning. The final part of this chapter provides an illustrative case report that combines a number of different measures in the assessment of a 13-year-old female with academic difficulties.

4.08.1.1 Brief History of Intelligence Testing

Exactly when intelligence testing began is difficult to pinpoint. IQ tests as they are known in the 1990s stem from nineteenth century Europe. Study of the two extremes of intelligence, giftedness and retardation, led to breakthroughs in intellectual assessment. Tracing the roots of assessment, the early pioneers were Frenchmen, who worked with the retarded. In the 1830s, Jean Esquirol distinguished between mental retardation and mental illness, unlumping idiocy from madness (Kaufman, 1983). He focused on language and speech patterns, and even on physical measurements such as the shape on the skull, in testing "feebleminded" and "demented" people.

Another contribution of Esquirol was a system for labeling the retarded. He formed a hierarchy of mental retardation, with "moron" at the top. Those less mentally adept were classified as "imbecile" and those at the bottom rung of intelligence were "idiots." The 1990s classification systems, which use terms like profoundly, severely, or moderately retarded, appear to most as less offensive than Esquirol's labels.

In the mid-1800s, another innovator joined Esquirol in testing retarded individuals. As opposed to Esquirol's use of verbal tests, Edouard Seguin tested these individuals using nonverbal methods, oriented toward sensation and motor activity (Kaufman, 1983). A link between Seguin's work and the twentieth century can be seen, as many of the procedures he developed were adopted or modified by later developers of performance and nonverbal tasks. Intelligence testing and education became intertwined during this time when Seguin convinced authorities of the desirability of educating the "idiots" and "imbeciles." Seguin was also the inspiration for Maria Montessori. Many of his methods and materials are present in the Montessori approach to education.

In an approach similar to Seguin's, stressing discrimination and motor control, Galton studied individual differences in the ordinary man, not just the tail ends of the normal curve. He was committed to the notion that intelligence is displayed through the use of the senses (sensory discrimination and sensory motor coordination), and believed that those with the highest IQ should also have the best discriminating abilities. Therefore, he developed mental tests that were a series of objective measurements of sensory abilities like keenness of sight, color discrimination, and pitch discrimination; sensory motor-abilities like reaction time and steadiness of hand; and motor abilities like strength of squeeze and strength of pull (Cohen, Montague, Nathanson & Swerdlik, 1988).

Galton's theory of intelligence was simplistic: people take in information through their senses, so those with better developed senses ought to be more intelligent. Although his theory of intelligence was quite different than what is considered as intelligence today, he is credited with establishing the first comprehensive individual intelligence test. He also influenced two basic notions of intelligence: the idea that intelligence is a unitary construct, and that individual differences in intelligence are largely genetically determined (possibly influenced by the theory of his cousin, Charles Darwin) (Das, Kirby, & Jarman, 1979).

Galton's concepts were brought to the USA by James McKeen Cattell, an assistant in Galton's laboratory (Roback, 1961). In 1890, Cattell established a Galton-like mental test laboratory at the University of Pennsylvania, and he moved his laboratory to Columbia University in New York the next year. He shared Galton's philosophy that intelligence is best measured by sensory tasks, but expanded his use of "mental tasks" to include standardized administration procedures. He urged for the establishment of norms, and thereby took the assessment of mental ability out of the arena of abstract philosophy and demonstrated that mental ability could be studied experimentally and practically. Studies conducted around the turn of the century at Cattell's Columbia laboratory showed that American versions of Galton's sensory-motor test correlated close to zero with meaningful criteria of intelligence, such as grade-point average in college.

Following Esquirol's lead by focusing on language abilities, Alfred Binet began to develop mental tasks with his colleagues Victor Henri and Theodore Simon (Binet & Henri, 1895; Binet & Simon, 1905, 1908). His tests were complex, measuring memory, judgment, reasoning, and social comprehension, and these tasks survive to the 1990s in most tests of intelligence for children and adults.

The Minister of Public Instruction in Paris appointed Binet to study the education of retarded children in 1904. The Minister wanted Binet to separate retarded from normal children in the public schools. Thus, with 15 years worth of task development behind him, the Binet–Simon scale quickly was assembled (Sattler; 1988).

Binet used a new approach in his tests; he ordered tasks from easy to hard within the scale. In 1908 and 1911 he revised his test to group tasks by age level, to add levels geared for adults, to introduce the concept of mental age, and to give more objective scoring rules (Sattler, 1988). If someone passed the nine-year-old level tasks, but failed the ones at the 10-year level, then that person had the intelligence of the typical nine-year old, whether the person was 6, 9, or 30. The measurement adult's intelligence, except mentally retarded individuals, was almost an after thought. Binet's untimely death in 1911 prevented him from actualizing the many applications of his tests in child development, education, medicine, and research (Kaufman, 1990).

Every IQ test in existence has been impacted greatly by Binet's work, and incorporates many of the same kinds of concepts and test questions that he developed. Lewis Terman was one of several Americans who translated the Binet–Simon for use in the USA. Terman published a "tentative" revision in 1912. Terman then took four years to carefully adapt, expand, and standardize the Binet–Simon. After much painstaking work, in 1916 the Stanford–Binet was born. This test used the concept of mental quotient and introduced the intelligence quotient. The Stanford–Binet left its competitors in the dust, and became the leading IQ test in America.

Like Binet, Terman viewed intelligence tests as useful for identifying "feeblemindedness," or weeding out the unfit. Terman also saw the potential for using intelligence tests with adults for determining ability to perform well in certain occupations. He believed that minimum intelligence quotients were necessary for success in specific occupations.

With the onset of World War I, the field of adult assessment grew quickly due to practical recruitment issues. The USA needed a way to evaluate the mental abilities of thousands of recruits and potential officers in an expedient manner. Due to the large volume of individuals tested, a group version of Binet's test was created by Arthur Otis, a student of Terman. This group-administered Stanford–Binet was labeled the Army Alpha. The Army Beta was also created during World War I to assess anyone who could not speak English or who was suspected of malingering. This was a nonverbal problem solving test, which was a forerunner of today's nonverbal ("Performance") subtests. The Army Alpha and Army Beta tests, published by Yerkes in 1917, were validated on huge samples (nearly two million). The tests were scores "A" to "D-" with the percent scoring "A" supporting their validity: 7% of recruits, 16% of corporals, 24% of sergeants, and 64% of majors. The best evidence of validity, though, was the Peter Principle in action. Second lieutenants (59% "A") outperformed their direct superiors, first lieutenants (53%) and captains (53%), while those with ranks higher than major did not do as well as majors (Kaufman, 1990).

The subtests developed by Binet and World War I psychologists were borrowed by David Wechsler in the mid-1930s to develop the Wechsler-Bellevue Intelligence Scale. His innovation was not in the selection of tasks, but in his idea that IQ was only in part a verbal intelligence. He also assembled a Performance Scale from the nonverbal, visual-motor subtests that were developed during the war to evaluate people who could not speak English very well or whose motivation to succeed was in doubt. Wechsler paired the verbally laden Army Alpha and the Stanford–Binet to create the verbal scale, and the Army Group Examination Beta and the Army Individual Performance scale to create the Performance scale. These two scales together were thought to contribute equally to the overall intelligence scale. The Full Scale IQ, for Wechsler, is an index of general mental ability (*g*).

To Wechsler, these tests were dynamic clinical instruments, more than tools to subdivide retarded individuals (Kaufman, 1990). However, the professional public was leery. They wondered how tests developed for the low end of the ability spectrum could be used to test normal people's intelligence. The professionals and publishers had a difficult time accepting that nonverbal tests could be used as measures for all individuals, not just foreigners. The postwar psychological community held the belief that IQ tests were primarily useful for predicting children's success in school, but were critical of Wechsler for developing a test primarily for adolescents and adults.

He persisted with his idea that people with poor verbal intelligence may be exceptional in their nonverbal ability, and vice versa. He met with resistance and frustration, and could not find a publisher willing to subsidize his new test. Thus, with a group of psychologist friends, Weschler tested nearly 2000 children, adolescents, and adults in Brooklyn, New York. Although it was a very urban sample, he managed to obtain a well stratified sample.

Once it had been standardized, Wechsler had no problem finding a publisher in The Psychological Corporation. The original Wechsler-Bellevue (Wechsler, 1939) has grandchildren, including the Wechsler Intelligence Scale for Children-Revised (WISC-R), and the Wechsler Adult Intelligence Scale-Revised (WAIS-R); more recently in 1991 a great-grandchild was born, the WISC-III.

Loyalty to the Stanford–Binet prevented Wechsler's test from obtaining instant success. However, gradually, Wechsler overtook the Binet during the 1960s as the learning disabilities movement gained popularity. The Verbal IQ and Performance IQ provided by Wechsler's tests helped to identify bright children who had language difficulties or visual-perceptual problems. The Stanford–Binet offered just one IQ, and the test was so verbally oriented that people with exceptional nonverbal intelligence were penalized.

Terman's Stanford–Binet lost favor when revisions of the battery after his death in 1956 proved to be expedient and shortsighted. In the 1990s, Wechsler's scales have proven themselves by withstanding challenges by other test developers, including the Kaufman Assessment Battery for Children (K-ABC) (Kaufman & Kaufman, 1983), the Kaufman Adolescent and Adult Intelligence Test (KAIT) (Kaufman & Kaufman, 1993), the Differential Abilities Scale (DAS) (Elliott, 1990), and Woodcock–Johnson (Woodcock & Johnson, 1989). These many other tests are used widely, but generally remain as alternatives or supplements to the Wechsler scales.

4.08.1.2 Controversy Over Intelligence Testing

The measurement of intelligence has long been the center of debate. In the past, critics have spoken of IQ tests as "biased," "unfair," and "discriminatory." The critics' arguments in the 1990s center more around what the IQ tests truly measure, as well as how or if they should be interpreted, their relevance to intervention, and their scope. Despite the controversy, there is great interest and need for measurement of intelligence, especially in the educational context, in order to help children and adolescents. Amidst the criticisms and limitations of IQ testing, these instruments remain a most technologically advanced and sophisticated tool of the profession for providing essential and unique information to psychologists so they may best serve the needs of children and adults. When used in consideration of the American Psychological Association's Ethical Principles of Psychologists (American Psychological Association, 1990) Principle 2-Competence, which encourages clinicians to recognize differences among people (age, gender, socioeconomic, and ethnic backgrounds) and to understand test research regarding the validity and the limitations of their assessment tools, these tests can be beneficial despite the controversy.

Three controversial themes associated with IQ testing were noted by Kaufman (1994). The first involves opposition to the common practice of subtest interpretation advocated by Wechsler (1958) and Kaufman (1979, 1994b). The second includes those who would abandon the practice altogether. Finally, the third group suggests that the concept of intelligence testing is sound, but more contemporary instrumentation could improve the effectiveness of the approach.

The first group of psychologists has encouraged practitioners to "just say 'no' to subtest analysis" (McDermott, Fantuzzo, & Glutting, (1990) (p. 299; also see Glutting, McDermott, Prifitera, & McGrath, (1994), and Watkins & Kush, (1994)). McDermott and his colleagues argue that interpreting a subtest profile is in violation of the principles of valid test interpretation because the ipsative method fails to improve prediction (McDermott, Fantuzzo, Glutting, Watkins, & Baggaley, 1992) and therefore does not augment the utility of the test. It is agreed that the results of studies conducted by McDermott et al. (1992), do suggest that using the WISC-III in isolation has limitations, but using the ipsative approach in conjunction with other relevant information such as achievement test results and pertinent background information may be beneficial. Kaufman (1994) further suggests that by shifting to the child's midpoint score a more equally balanced set of hypotheses can be developed which can be integrated with other findings to either strengthen or disconfirm hypotheses. When the ipsative assessment approach is used to create a base from which to search for additional information to evaluate hypothesized strengths and weaknesses in the child's subtest profile its validity is extended beyond that which can be obtained using the Wechsler subtests alone. If support is found for the hypotheses, then such a strength or weakness can be viewed as reliable, because of its cross-validation (Kaufman, 1994). When considering this position and that represented by McDermott et al., as well as Glutting et al. (1994) and Watkins and Kush (1994), it is important to recognize that these authors are against subtest profile analysis not the use of IQ tests in general. This is in contrast to others who hold a more extreme negative view of IQ testing.

One extremist group that opposes IQ testing includes those who advocate throwing away Verbal and Performance IQs, along with the subtest profile interpretation, and finally the Full Scale IQ because they insist that all that Wechsler scales measure is *g* (MacMann & Barnett, 1994). They argue that differences between the Verbal and Performance Scales on Wechsler tests hold no meaning, that conventional intelligence tests only measure *g* (and a measure of *g* is not enough to warrant the administration of such an instrument) and that Wechsler scale data do not have instructional value. These authors fail to recognize a wealth of data that illustrates that differences between the Verbal and Performance Scales can be very important. Any clinician using intelligence tests cannot ignore the numerous studies that are available that point to significant Verbal Performance differences in patients with right-hemisphere damage (Kaufman, 1990, Chapter 9), in Hispanic and Navajo children (McShane & Cook, 1985; McShane & Plas, 1984; Naglieri, 1984), and in normal adults (Kaufman, 1990, Chapter 7). If only the Full Scale IQ is interpreted, following MacMann and Barnett's (1994) advisement that the Verbal and Performance scales are meaningless, then it prevents the fair use of these tests with those groups who have inconsistent V–P discrepancies. Moreover, contrary to what MacMann and Barnett (1994) suggest, it is clear that when a child earns very poor Verbal and average Performance scores there are obvious implications for instruction and a high probability that such results will be reflected in poor verbal achievement (Naglieri, 1984).

Another extremist group opposed to IQ testing is Witt and Gresham (1985) who state, "The WISC-R lacks treatment validity in that its use does not enhance remedial interventions for children who show specific academic skill deficiencies" (p. 1717). It is their belief that the Wechsler test should be replaced with assessment procedures that have more treatment validity. However, as Kaufman (1994) points out, Witt and Gresham (1985) do not provide evidence for their statements. Another pair of researchers (Rechsly & Tilly, 1993) agree with the Witt and Gresham (1985) statements about the lack of treatment validity of the WISC-R, but only provide references that are not specific to the Wechsler scales. Thus, the Wechsler scales appear to have been rejected by these researchers without ample relevant data.

Witt and Gresham (1985) also complain that the WISC-R (as well as the WISC-III) only yields a score, and does not provide school psychologists with direct strategies of what to do with and for children, which are what teachers are requesting. As Kaufman (1994) points out, however, it is not the instrument's responsibility to provide direct treatment information; rather, he states, "It is the examiner's responsibility ... to provide recommendations for intervention" (p. 35). The examiner should not be just taking the bottom-line IQ scores or standard scores, but should provide statements about a child's strengths and weaknesses that have been cross-validated through the observations of behavior, background information, and the standardized intelligence and achievement tests.

Finally, there is a group of professionals who have suggested that the Wechsler has limits that should be recognized, but these limits could be addressed by alternative methods rather than abandoning the practice of intelligence testing altogether. Some have argued for a move toward alternative conceptualizations of intelligence and methods to measure new constructs that are based on factor analytic research (e.g., Woodcock, 1990) while others have used neuropsychology and cognitive psychology as a starting point (e.g., Naglieri & Das, 1996). The results of these efforts have been tests such as the Das–Naglieri Cognitive Assessment System

(Naglieri & Das, 1996), Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983), Kaufman Adult Intelligence Test (Kaufman & Kaufman, 1993), and the Woodcock–Johnson Tests of Cognitive Ability (Woodcock & Johnson, 1989). This chapter will show how these tests, and others, can be utilized in conjunction with the Wechsler to gain a more complete view of the child.

The main goal of this chapter is to use Kaufman's (1994) philosophy of "intelligent" testing to address some of the concerns about Wechsler interpretation through a careful analysis of the results and integration with other measures. Much of this discussion is based on the principles of IQ testing as outlined by Kaufman, which focus on the view that "WISC-III assessment is of the individual, by the individual, and for the individual" (Kaufman, 1994, p. 14). Through research knowledge, theoretical sophistication, and clinical ability examiners must generate hypotheses about an individual's assets and deficits and then confirm or deny these hypotheses by exploring multiple sources of evidence. Well-validated hypotheses must then be translated into meaningful, practical recommendations. A brief description of those five principles of intelligent testing follows.

Clinician-scientists must come well equipped with state of the art instrumentation, good judgment, knowledge of psychology, and clinical training to move beyond the obtained IQs (Kaufman, 1994). Integration of information from many sources and different tests is very important if the child referred for evaluation is to remain the focus of assessment because it is impossible to describe fully a person by just presenting a few numbers from the Wechsler protocol or those obtained from a computer program. Each adult and child who comes for an assessment has unique characteristics, a particular way of approaching test items, and may be affected differently by the testing situation than the next individual. Through the use of an integrated interpretation approach the various dimensions that influence a child can become apparent.

4.08.1.3 Principles of the Intelligent Testing Philosophy

The first principle of intelligent testing is that "the WISC-III subtests measure what the individual has learned" (Kaufman, 1994, p. 6). The WISC-III is like an achievement test, in that it is a measure of past accomplishments and is predictive of success in traditional school subjects. Research indicates that intelligence tests consistently prove to be good predictors of conventional school achievement. The WISC-III Manual (Wechsler, 1991, pp. 206–209) gives many such correlations between the WISC-III IQs or Factor Indexes and achievement measures. Although this connection between the WISC-III and achievement in school is well documented empirically, it should not be accepted ultimately as a statement of fate, that if a child scores poorly on the WISC-III that they will do poorly in school (Kaufman, 1994). Instead, constructive interpretation of a test battery can lead to recommendations which may alter helpfully a child's development.

The second principle is that WISC-III subtests are samples of behavior and are not exhaustive. Because the subtests only offer a brief glimpse into a child's overall level of functioning, examiners must be cautious in generalizing the results to performance and behaviors in other circumstances. The Full Scale "should not be interpreted as an estimate of a child's global or total intellectual functioning; and the WISC-III should be administered along with other measures, and the IQs interpreted in the context of other test scores" (Kaufman, 1994, p. 7). It is important that the actual scores are not emphasized as the bottom-line; rather, it is more beneficial to elaborate on what children can do well, relative to their own level of ability. Such information can be used to create an individualized education program which will tap a child's areas of strength and help improve areas of deficit.

Principle three states, "The WISC-III assesses mental functioning under fixed experimental conditions" (Kaufman, 1994, p. 8). Rigid adherence to the standardized procedures for administration and scoring, outlined in the WISC-III manual (Wechsler, 1991), helps to ensure that all children are measured in an objective manner. However, parts of the standardized procedure make the testing situation very different than a natural setting. For example, it is not very often in a children's every day life that someone is transcribing virtually every word they say or timing them with a stopwatch. The standardization procedures are important to follow, but must be taken into account as limitations when interpreting the scores obtained in the artificial testing situation. The value of the intelligence test is enhanced when the examiner can meaningfully relate observations of the child's behaviors in the testing situation to the profile of scores.

The fourth principle is that "The WISC-III is optimally useful when it is interpreted from an information-processing model" (Kaufman, 1994, p. 10). This is especially beneficial for helping to hypothesize functional areas of strength and dysfunction. This model suggests examining how information enters the brain from the sense organs (input), how information is interpreted and processed (integration), how information is stored for later retrieval (storage), and how information is expressed linguistically or motorically (output). Through this model, examiners can organize the test data, including fluctuations in subtest scores, into meaningful underlying areas of asset and deficit.

The fifth and very important principle of intelligent testing is that, "Hypotheses generated from WISC-III profiles should be supported with data from multiple sources" (Kaufman, 1994, p. 13). Although good hypotheses can be raised from the initial WISC-III test scores, such hypotheses do not hold water unless verified by diverse pieces of data. Such supporting evidence may come from careful observation of a child's behavior during test administration; from the pattern of responses across various subtests; from background information obtained from parents, teachers, or other referral sources; from previous test data; and from the administration of supplemental subtests. The integration of data from all these different sources is critical in obtaining the best and most meaningful clinical interpretation of a test battery.

4.08.2 MEASURES OF INTELLIGENCE

Intelligence tests are administered for a variety of reasons including identification (of mental retardation, learning disabilities, other cognitive disorders, giftedness), placement (gifted and other specialized programs), and as a cognitive adjunct to a clinical evaluation. The following comprehensive intelligence tests are discussed in the next sections: WPPSI-R, WISC-III, WAIS-R, K-ABC, KAIT, Binet-IV, WJ-R Tests of Cognitive Ability, DTLA-3, DAS, and CAS.

4.08.2.1 Wechsler's Scales

As discussed in the brief history of IQ tests, Wechsler's scales reign as leaders of measures of child, adolescent, and adult intelligence. The WISC-III is a standard part of a battery administered to children by school psychologists and private psychologists to assess level cognitive functioning, learning styles, learning disabilities, or giftedness. The WAIS-R is administered invariably as a part of a battery to assess intellectual ability for a clinical, neurological, or vocational evaluation of adolescents and adults. The WPPSI-R may be used to measure intellectual ability from ages three to seven years, three months; intellectual assessment may be done from age six up to age 16 with the WISC-III, while the WAIS-R may be used from ages 16–74.

The different Wechsler scales overlap at ages 6–7 and 16. Kaufman (1994) recommends that the WISC-III be used at both these overlapping age periods rather than the WPPSI-R or the WAIS-R. One of the reasons cited for these recommendations is that the WISC-III has a much better "top" than the WPPSI-R for children who are ages six or seven. On the WPPSI-R a child can earn a maximum score of 16 or 17 (rather than 19) on six of the 10 subtests when age seven. The recommendation to use the WISC-III rather than the WAIS-R at age 16 is made because the WAIS-R norms are outdated relative to the WISC-III norms. Kaufman (1990) recommends that the WAIS-R norms for ages 16–19 be used cautiously, and states that the "even apart WISC-III norms, the Performance scale is more reliable for the WISC-III (0.92) than the WAIS-R (0.88) at age 16" (p. 40).

Wechsler (1974) puts forth the definition that "intelligence is the overall capacity of an individual to understand and cope with the world around him [or her]" (p. 5). His tests, however, were not predicated on this definition. Tasks developed were not designed from well-researched concepts exemplifying his definition. In fact, as previously noted, virtually all of his tasks were adapted from other existing tests.

Like the Binet, Wechsler's definition of intelligence also ascribes to the conception of intelligence as an overall global entity. He believed that intelligence cannot be tested directly, but can only be inferred from how an individual thinks, talks, moves, and reacts to different stimuli. Therefore, Wechsler did not give credence to one task above another, but believed that this global entity called intelligence could be ferreted out by probing a person with as many different kinds of mental tasks as one can conjure up. Wechsler did not believe in a cognitive hierarchy for his tasks, and he did not believe that each task was equally effective. He felt that each task was necessary for the fuller appraisal of intelligence.

4.08.2.1.1 Wechsler Primary and Preschool Intelligence Scale-Revised (WPPSI-R)

(i) Standardization and properties of the scale

The WPPSI-R is an intelligence test for children aged three years, through seven years, three months. The original version of the WPPSI was developed in 1967 for ages four to six and a half years, and the WPPSI-R was revised in 1989. Several changes were made to the revised version of the WPPSI-R. The norms were updated, the appeal of the content to young children was improved, and the age range was expanded.

The WPPSI-R is based on the same Wechsler–Bellevue theory of intelligence, emphasizing intelligence as a global capacity but having Verbal and Performance scales as two methods of assessing this global capacity (Kamphaus, 1993). The Verbal scale subtests include: Information Comprehension, Arithmetic, Vocabulary, Similarities, and Sentences (optional subtest). The Performance scale subtests include: Object Assembly, Block Design, Mazes, Picture Completion, and Animal Pegs (optional subtest).

Like the K-ABC and the Differential Abilities Scales (DAS), the WPPSI-R allows the examiner to "help" or "teach" the client on early items on the subtests to assure that the child understands what is expected. Providing this extra help is essential when working with reticent preschoolers (Kamphaus, 1993).

Subtest scores have a mean of 10 and a standard deviation of three. The overall Verbal, Performance, and Full Scale IQs have a mean of 100 and a standard deviation of 15. The examiner manual provides interpretive tables that allow the examiner to determine individual strengths and weaknesses as well as the statistical significance and clinical rarity of Verbal and Performance score differences.

The WPPSI-R was standardized on 1700 children from age three through seven years, 3 months. The standardization procedures followed the 1986 US Census Bureau estimates. Stratification variables included gender, race, geographic region, parental occupation, and parental education.

The WPPSI-R appears to be a highly reliable measure. The internal consistency coefficients across age groups, for the Verbal, Performance, and Full Scale IQs are 0.95, 0.92, and 0.96, respectively. For the seven-year-old age group, the reliability coefficients are somewhat lower. The internal consistency coefficients for the individual Performance subtests vary from 0.63 for Object Assembly to 0.85 for Block Design, with a median coefficient of 0.79. The internal consistency coefficients for the individual Verbal subtests vary from 0.80 for Arithmetic to 0.86 for Similarities, with a median coefficient of 0.84. The test–retest coefficient for the Full Scale IQ is 0.91.

The WPPSI-R manual provides some information on validity; however, it provides no information on the predictive validity of the test. Various studies have shown that concurrent validity between the WPPSI-R and other tests is adequate. The correlation between the WPPSI and the WPPSI-R Full Scale IQs was reported at 0.87, and the correlations between WPPSI-R and WISC-III Performance, Verbal, and Full Scale IQs for a sample of 188 children were 0.73, 0.85, and 0.85, respectively. The correlations between the WPPSI-R and other well known cognitive measures are, on average, much lower. The WPPSI-R Full Scale IQ correlated 0.55 with the K-ABC Mental Processing Composite (Kamphaus, 1993) and 0.77 with the Binet IV Test Composite (McCrowell & Nagle, 1994). In general, the validity coefficients provide strong evidence for the construct validity of the WPPSI-R (Kamphaus, 1993).

(ii) Overview

The WPPSI-R is a thorough revision of the 1967 WPPSI, with an expanded age range, new colorful materials, new item types for very young children, a new icebreaker subtests (Object Assembly), and a comprehensive manual (Kaufman, 1990). The revision of the test has resulted in an instrument that is more attractive, more engaging, and materials that are easier to use (Buckhalt, 1991; Delugach, 1991).

The normative sample is large, provides recent norms and is representative of the 1986 US Census data (Delugach, 1991; Kaufman, 1990). The split-half reliability of the IQs and most subtests are exceptional, the factor analytic results for all age groups are excellent, and the concurrent validity of the battery is well supported by several excellent correlational studies (Delugach, 1991; Kaufman, 1990). The manual provides a number of validity studies, factor analytic results, research overviews, and state-of-the-art interpretive tables, which provide the examiner with a wealth of information. Kaufman (1990) noted that the WPPSI-R has a solid psychometric foundation.

In spite of its reported strengths, the WPPSI-R has flaws. In publishing the WPPSI-R, great effort was made to ensure that all subtests had an adequate "top" and "bottom" (Kaufman, 1992). However, the WPPSI-R has an insufficient floor at the lowest age levels, which limits the test's ability to diagnose intellectual deficiency in young preschoolers (Delugach, 1991). For example, a child at the lowest age level (2–11–16 to 3–2–15) who earns only one point of credit on all subtests will obtain a Verbal IQ of 75, a Performance IQ of 68, and a Full Scale IQ of 68, making it impossible to adequately assess the child's degree of intellectual deficiency. The WPPSI-R subtests are able to distinguish between gifted and nongifted children at all ages, but the top of some subtests is not adequate to discriminate among gifted children. Kaufman (1992) indicates that at the youngest ages (3–4.5 years), all subtests are excellent. However, at age five, Geometric Design begins to falter, and at ages 6.5 and above, it only allows a maximum scaled score of 16. Other subtests, such as Object Assembly and Arithmetic also have problems with the ceiling. Although the ceilings on the subtests described are not ideal, the IQ scales do allow maximum IQs of 150 for all ages and IQs of 160 for children ages 3–6.25.

Another major problem with the WPPSI-R is the role played by speed of responding. From both early developmental perspectives and common-sense perspectives, giving bonus points for speed is silly (Kaufman, 1992). Young children may respond slowly for a variety of reasons that have little to do with intellect. A three- or four-year-old child might respond slowly or deliberately because of lack immaturity, lack of experience in test taking, underdeveloped motor coordination, or a reflective cognitive style. The WPPSI-R Object Assembly and Block Design place an overemphasis on speed. For example, if a six and a half or seven year old child solves every Object Assembly item perfectly, but does not work quickly enough to earn bonus points, they would only receive a scaled score of 6 (ninth percentile). Because of the age-inappropriate stress on solving problems in with great speed, a child's IQ may suffer on two of the 10 subtests (Kaufman, 1992). In addition, the directions on some of the Performance subtests are not suitable for young children because they are not developmentally appropriate (Kaufman, 1990). However, Delugach (1991) reports that if the directions are too difficult, the test provides procedures to ensure that the child understands the demands of the task.

The WPPSI-R is a useful assessment tool, but, like all others, it possesses certain weaknesses that limit its usefulness (Delugach, 1991). Examiners should be aware of the WPPSI-R's inherent strengths and weaknesses and keep them in mind during administration, scoring, and interpretation. The WPPSI-R may provide the examiner with useful information; however, "it does little to advance our basic understanding of the development and differentiation of intelligence or our understanding of the nature of individual differences in intelligence" (Buckhalt, 1991).

4.08.2.1.2 Wechsler Intelligence Scale for Children-3rd Edition (WISC-III)

(i) Standardization and properties of the scale

The WISC-III was standardized on 2200 children ranging in age from six through 16 years. The children were divided into 11 age groups, one group for each year from six through 16 years of age. The median age for the each age group was the sixth month (e.g., 7 years, 6 months). The standardization procedures followed the 1980 US Census data and the manual provides information by age, gender, race/ethnicity, geographic region, and parent education. "Overall, the standardization of the WISC-III is immaculate ... a better-standardized intelligence test does not exist" (Kaufman, 1993, p. 351).

The WISC-III yields three IQs, a Verbal Scale IQ, a Performance Scale IQ, and a Full Scale IQ. All three are standard scores (mean of 100 and standard deviation of 15) obtained by comparing an individual's score with those earned by the representative sample of age peers. Within the WISC-III, there are 10 mandatory and three supplementary subtests all of which span the age range of six through 16 years. The Verbal scale's five mandatory subtests include: Information, Similarities, Arithmetic, Vocabulary, and Comprehension. The supplementary subtest on the Verbal Scale is Digit Span. Digit Span is not calculated into the Verbal IQ unless it has been substituted for another Verbal subtest because one of those subtests has been spoiled (Kamphaus, 1993; Wechsler, 1991).

The five mandatory Performance scale's subtests include Picture Completion, Picture Arrangement, Block Design, Object Assembly, and Coding. The two supplementary subtest on the Performance scale are Mazes and Symbol Search. The Mazes subtest may be substituted for any Performance scale subtest; however, Symbol Search may only be substituted for the Coding subtest (Kamphaus, 1993; Wechsler, 1991).

"Symbol Search is an excellent task that should have been included among the five regular Performance subtests instead of Coding. Mazes is an awful task that should have been dropped completely from the WISC-III" (Kaufman, 1994, p. 58). He goes further to say that "there's no rational reason for the publisher to have rigidly clung to Coding as a regular part of the WISC-III when the new Symbol Search task is clearly a better choice for psychometric reasons" (p. 59). Therefore, for all general purposes, Kaufman (1994) strongly recommends that Symbol Search be substituted routinely for coding as part of the regular battery, and to use Symbol Search to compute the Performance IQ and Full Scale IQ. The manual does not say to do this, but neither does it prohibit it.

Reliability of each subtest except Coding and Symbol Search was estimated by the split-half method. Stability coefficients were used as reliability estimates for the Coding and Symbol Search subtests because of their speeded nature.

Across the age groups, the reliability coefficients range from 0.69 to 0.87 for the individual subtests. The average reliability, across the age groups, for the IQs and Indexes are: 0.95 for the Verbal IQ, 0.91 for the Performance IQ, 0.96 for the Full Scale IQ, 0.94 for the Verbal Comprehension Index, 0.90 for the Perceptual Organization Index, 0.87 for the Freedom from Distractibility Index and 0.85 for the Processing Speed Index (Wechsler, 1991).

Factor analytic studies of the WISC-III standardization data were performed for four age group subsamples: ages 6–7 (n = 400), ages 8–10 (n = 600), ages 11–13 (n = 600), and ages 14–16 (n = 600) (Wechsler, 1991). Compiling the results of the analysis, a four-factor solution was found for the WISC-III. Like the WISC-R, Verbal Comprehension and Perceptual Organization remain the first two factors. Verbal Comprehension involves verbal knowledge and the expression of this knowledge. Perceptual Organization, a nonverbal dimension, involves the ability to interpret and organize visually presented material. The third factor consists of the Arithmetic and Digit Span subtests. Factor III has been described as Freedom from Distractibility since common among tasks is the ability to focus, to concentrate, and to remain attentive. Other interpretations of this factor have included facility with numbers, short-term memory, and sequencing because the three tasks which comprise the factor all involve a linear process whereby numbers are manipulated. Success is either facilitated by or wholly dependent on memory (Kaufman, 1979). The fourth factor consists of Coding and Symbol Search, and is referred to as the Processing Speed factor. Taken together, the Verbal Comprehension and Perceptual Organization factors offer strong support for the construct validity of the Verbal and Performance IQs; substantial loadings on the large, unrotated first factor (g) supports the construct underlying Wechsler's Full Scale IQ.

(ii) Analyzing the WISC-III data

To obtain the most information from the WISC-III, the psychologist should be more than familiar with each of the subtests individually as well as with the potential information that those subtests can provide when integrated or combined. The WISC-III is maximally useful when tasks are grouped and regrouped to uncover a child's strong and weak areas of functioning, so long as these hypothesized assets and deficits are verified by multiple sources of information.

As indicated previously, the WISC-III provides examiners with a set of four Factor Indexes in addition to the set of three IQs. The front page of the WISC-III record form lists the seven standard scores in a box on the top right. The record form is quite uniform and laid out nicely; however, it is difficult to know just what to do with all of those scores. Kaufman (1994) has developed seven steps to interpretation which offer a systematic method of WISC-III interpretation that allows the clinician to organize and integrate the test results in a step-wise fashion. The seven steps (see Table 1) provide an empirical framework for profile attack while organizing the profile information into hierarchies.

(iii) Overview

Professionals in the field of intelligence testing have described the third edition of the Wechsler Intelligence Scale for Children in a number of different ways. Some critics feel that the WISC-III reports continuity, the status quo, but makes little progress in the evolution of the assessment of intelligence. Such critics note that despite more than 50 years of advancement in theories of intelligence, the Wechsler philosophy of intelligence (not actually a formal theory), written in 1939, remains the guiding principle of the WISC-III (Schaw, Swerdilik, & Laurent, 1993). One of the principal goals for developing the WISC-III stated in the manual was merely to update the norms, which is "hardly a revision at all" (Sternberg, 1993). Sternberg (1993) suggests that the WISC-III is being used to look for a test of new constructs in intelligence, or merely a new test, the examiner should look elsewhere.

In contrast to these fairly negative evaluations, Kaufman (1993) reports that the WISC-III is a substantial revision of the WISC-R and that the changes that have been made are considerable and well done. "The normative sample is exemplary, and the entire psychometric approach to test development, validation, and interpretation reflects sophisticated, state-of-the-art knowledge and competence" (Kaufman, 1993). For Kaufman, the WISC-III is not without its flaws but his overall review of the test is quite positive. One of Kaufman's (1993) main criticisms is that the Verbal tasks are highly culturally-saturated and school-related, which tend to penalize bilingual, minority, and learning-disabled children. He suggests that perhaps a special scale could have been developed to provide a fairer evaluation of the intelligence of children who are from the non-dominant culture or who have academic difficulties. Another criticism raised by Kaufman is that too much emphasis is (placed on a child's speed of responding on the WISC-III. It is difficult to do well on the WISC-III if you do not solve problems very quickly. This need for

Table 1 Summary of seven steps for interpreting WISC-III profiles.

Step	
1	*Interpret the full scale IQ* Convert it to an ability level and percentile rank and band it with error, preferable a 90% confidence interval (about ± 5 points)
2	*Determine if the verbal-performance IQ discrepancy is statistically significant* Overall values for V–P discrepancies are *11 points* at the 0.05 level and *15 points* at the 0.01 level. For most testing purposes, the 0.05 level is adequate
3	Determine if the V–P IQ discrepancy is interpretable, or if the VC and PO factor indexes should be interpreted instead Ask four questions about the Verbal and Performance Scales *Verbal Scale* (i) Is there a significant difference ($p < 0.05$) between the child's standard scores in VC vs. FD? size needed for significant (VC–FD) = 13+ points (ii) Is there abnormal scatter (highest minus lowest scaled score) among the five Verbal subtests used to compute V-IQ? Size needed for abnormal verbal scatter = 7+ points *Performance Scale* (iii) Is there a significant difference ($p < 0.05$) between the child's standard scores on PO vs. PS? Size needed for significant (PO–PS) = 15+ points (iv) Is there abnormal scatter (highest minus lowest scaled score) among the five Performance subtests used to compute P-IQ? Size needed for abnormal performance scatter = 9+ points If all answers are no, the V–P IQ discrepancy is interpretable. If the answer to one or more questions is yes, the V–P IQ discrepancy may not be interpretable. Examine the VC–PO discrepancy. Overall values for VC–PO discrepancies are *12 points* at the 0.05 level and *16 points* at the 0.01 level Determine if the VC and PO indexes are unitary dimensions: 1. Is there abnormal scatter among the four VC subtests? Size needed for abnormal VC scatter = 7+ points 2. Is there abnormal scatter among the four PO subtests? Size needed for abnormal PO scatter = 8+ points If the answer to either question is yes, then you probably shouldn't interpret the VC–PO Index discrepancy – unless the discrepancy is to big to ignore (see Step 4). If both answers are no, interpret the VC–PO differences as meaningful
4	*Determine if the V–P IQ discrepancy (or VC–PO discrepancy) Is abnormally large* Differences of at least *19 points* are unusually large for both the V–P and VC-PO discrepancies. Enter the table with the IQs or Indexes, whichever was identified by the questions and answers in Step 3 If neither set of scores was found to be interpretable in Step 3, they may be interpreted anyway if the magnitude of the discrepancy is unusually large (19+ points)
5	*Interpret the meaning of the global verbal and nonverbal dimensions and the meaning of the small factors* Study the information and procedures presented in Chapter 4 (verbal/nonverbal) and Chapter 5 (FD and PS factors). Chapter 5 provides the following rules regarding when the FD and PS factors have too much scatter to permit meaningful interpretation of their respective Indexes (both Chapters 4 and 5 are on *Intelligent Testing with the WISC-III*): (i) Do not interpret the FD Index if the Arithmetic and Digit Span scaled scores differ by *four or more points* (ii) Do not interpret the PO Index if the Symbol Search and Coding scaled scores differ by *four or more points*
6	*Interpret significant strengths and weaknesses in the WISC-III subtest profile* If the V–P IQ discrepancy is less than 19 points, use the child's mean of all WISC-III subtests administered as the child's midpoint If the V–P IQ discrepancy is 19 or more points, use the child's mean of all Verbal subtests as the midpoint for determining strengths and weaknesses on Verbal subtests, and use the Performance mean for determining significant deviations on Performance subtests Use either the specific values in Table 3.3 of *Intelligent Testing with the WISC-III*, rounded to the nearest whole number, or the following summary information for determining significant deviations: ±3 points: Information, similarities, arithmetic, vocabulary ±4 points: Comprehension, digit span, picture completion, picture arrangement, block design, object assembly, symbol search ±5 points: Coding
7	*Generate hypotheses about the fluctuations in the WISC-III subtest profile* Consult Chapter 6 in *Intelligent Testing with the WISC-III*, as it deals with the systematic reorganization of subtest profiles to generate hypotheses about strengths and weaknesses

Source: Kaufman (1994b). Reprinted with permission.

speed penalizes children who are more reflective in their cognitive style or who have coordination difficulties. The speed factor may prevent a gifted child from earning a high enough score to enter into an enrichment class or may lower a learning disabled child's overall IQ score to a below average level, just because they do not work quickly enough. Although the WISC-III clearly has had mixed reviews, it is one of the most frequently used tests in the field of children's intelligence testing.

4.08.2.1.3 WISC-III Short Form

Short forms of the Wechsler scales were developed shortly after the original tests were developed (Kaufman, Kaufman, Balgopal, & McLean, 1996). Short forms are designed to have sound psychometric qualities and clinical justification, but should also be practical to use. Clinicians and researchers utilize short form when wanting to perform a screen of intellectual ability or when doing research which does not permit the needed time to complete an entire Wechsler scale. In a study using three different WISC-III short forms, the clinical, psychometric, and practical qualities for each form were examined (Kaufman, Kaufman, Balgopal et al. 1996). A psychometric and clinically strong short form was examined, and included the following subtests: Similarities, Vocabulary, Picture Arrangement, and Block Design. A practical short form, based on its brevity and ease of scoring, included the following subtests: Information, Arithmetic, Picture Completion, and Symbol Search. A short form which combines psychometric, clinical, and practical qualities was also examined: Similarities, Arithmetic, Picture Completion, and Block Design.

The results of this study using the WISC-III standardization sample of 2200 children, 6–16 years old, revealed important information about the utility of these three different short forms (Kaufman, Kaufman, Balgopal et al., 1996). The split-half reliability coefficients, standard error of measurement (SEM), validity coefficients, and standard errors of estimate for the three selected tetrads are presented in Table 2. The form which had both psychometric and clinical properties (S-V-PA-BD) was compared to the form which had the quality of practicality in addition to psychometric and clinical properties (S-A-PC-BD). The results indicated that they were equally valid and about equally reliable (see Table 2). Each of the three short form tetrads had reliability coefficients of above 0.90 for the total sample. The brief tetrad (I-A-PC-SS) had a lower correlation with the Full Scale of 0.89, compared to the other two forms which each correlated 0.93 with the Full Scale. Although they were each about equally reliable, the S-A-PC-BD subtests are quicker to administer and only Similarities requires some subjectivity to score. It is quicker to score than the S-V-PA-BD form because it uses Arithmetic instead of Vocabulary.

The authors recommend that the extra 25–30% savings in time in using the S-A-PC-BD form, in addition to the added validity in comparison to the practical tetrad (I-A-PC-SS), makes the S-A-PC-BD short form an excellent choice. The very brief practical form was not recommended for clinical use or screening purposes because of its lower validity. Kaufman, Kaufman, Balgopal, et al. (1996) present an equation for converting a person's sum of scaled scores of the four subtests to estimated FSIQs. The magnitude of the intercorrelations among the component subtests provide that data from which the exact equation is derived. The intercorrelations vary to some extent as a function of age, which leads to slightly different equations at different ages. However, the authors state that the equations for the total sample for each tetrad represent a good overview for all children ages 6–16. The following conversion equation is for the total sample for the recommended S-A-PC-BD short form:

Estimated FSIQ = $1.6c + 36$, (for other specific equations for varying ages, see Kaufman, Kaufman, Balgopal et al., 1996, p. 103).

To use this conversion equation, the child's scaled scores on the four subtests (S-A-PC-BD) must first be summed. The sum (Xc) must then be entered into the equation. For example, if examiners give the recommended psychometric/clinical/practical form to an eight-year-old, the child's scores on the four subtests would need to be summed. Suppose that the child's sum is 50. The above equation would show:

$$\begin{aligned}\text{Estimated FSIQ} &= 1.6(50) + 36\\ &= 80 + 36 = 116\end{aligned}$$

It is important to note that examiners should not take the good psychometric qualities of the short form to mean that the short form can be regularly substituted for the complete battery. There is a wealth of information, both clinical and psychometric, that the examiner benefits from when administering the complete battery. It is important not to favor the short forms just because they are shorter, on account of all the important information that is derived from a complete administration. Kaufman (1990) suggests that the following are a few instances in which the savings in administration time may justify use of the short form: (i) when only a

Table 2 Reliability, standard error of measurement, and validity of the three selected short forms by age.

Age	*Split-half reliability*[a]			*Standard error of measurement*			*Validity: Correlation with full scale*[a]			*Standard error of estimate*		
(Years)	*SF1*	*SF2*	*SF3*	*SF1*	*SF2*	*SF3*	*SF1*	*SF2*	*SF3*	*SF1*	*SF2*	*SF3*
6	92	92	90	4.2	4.2	4.7	92	94	89	6.0	5.2	7.0
7	91	90	89	4.5	4.7	5.0	91	91	89	6.4	6.4	7.0
8	93	93	92	4.0	4.0	5.5	92	94	90	6.0	5.2	6.7
9	91	91	89	4.5	4.5	5.0	93	93	89	5.6	5.6	7.0
10	93	92	90	4.0	4.2	4.7	91	89	87	6.4	4.2	7.6
11	91	92	91	4.5	4.2	4.5	94	91	89	5.2	6.4	7.0
12	94	92	90	3.7	4.2	4.7	94	93	89	5.2	5.6	7.0
13	93	92	90	4.0	4.2	4.7	94	92	86	5.2	6.0	7.9
14	94	92	90	3.7	4.2	4.7	93	93	90	5.6	5.6	6.7
15	94	94	93	3.7	3.7	4.0	96	94	91	4.2	5.2	6.4
16	94	93	92	3.7	4.0	4.2	94	95	90	5.2	4.7	6.7
Total	93	92	91	4.0	4.2	4.5	93	93	89	5.6	5.6	7.0

Source: Kaufman et al. (1996).

Notes: SFI = Short Form 1 (Psychometric/Clinical; Similarities-Vocabulary-Picture Arrangement-Block Design), SF2 = Short Form 2 (Psychometric/Clinical/Practical; Similarities-Arithmetic-Picture Completion-Block Design), SF3 = Short Form 3 (Practical; Information-Arithmetic-Picture Completion-Symbol Search).

[a] Decimal points have been omitted.

global assessment of IQ is needed in the context of a complete personality evaluation; (ii) when a thorough evaluation has been completed recently and just a brief check of present intellectual functioning is needed; and (iii) when an individual does not require categorization of their intellectual ability for placement or for diagnosis of a cognitive disorder.

4.08.2.1.4 Wechsler Adult Intelligence Scale-Revised (WAIS-R)

The Wechsler Adult Intelligence Scale-Third Edition (WAIS-III; The Psychological Corporation, 1997) came out in August 1997 and soon will be replacing the WAIS-R. Based on our experience with previous versions of the WAIS, it is likely that there will be a transition period of 3–4 years during which time clinicians will be gradually moving to use primarily the newer instrument. Because of this predicted transition time, we are including information about both the WAIS-R and the WAIS-III. Additionally, much of the research on the WAIS-R will be directly relevant and applicable to the WAIS-III and is therefore included here.

(i) Standardization and properties of the scale

Similar to the other Wechsler scales discussed, three IQ scores are derived from the WAIS-R subtests. Each of these scores are standard scores with a mean of 100 and a standard deviation of 15, which are created by comparing an individual's score to scores earned by the normative group of the same age. The Verbal IQ is comprised of six verbal subtest scores (Information, Digit Span, Vocabulary, Arithmetic, Comprehension, and Similarities). The Performance IQ is comprised of five nonverbal subtests (Picture Completion, Picture Arrangement, Block Design, Object Assembly, and Digit Symbol). The Full Scale represents the average of the Verbal and Performance IQs.

The WAIS-R was standardized by administering the full scale to 1880 adult subjects, selected according to current US Census data tested between 1976 and 1980. Subjects were stratified according to age, sex, race (white–nonwhite), geographic region, occupation, education, and urban–rural residence. Subjects were divided into nine age groups, corresponding to categories often used by the US Census Bureau. The number in each age group ranged from 160–300, and the age groups spanned from ages 16–74. A problem with the standardization sample noted by Kaufman (1985) was that there was apparent systematic bias in the selection of 16- to 19-year-olds, leading to very questionable teenage norms. Also Hispanics were not included systematically in the total sample. There was no mention of how Hispanics were categorized, if any were tested (Kaufman, 1985).

Reliability coefficients for the 11 tests and the Verbal, Performance, and Full Scale IQs were computed using the split-half method (except for Digit Span and Digit Symbol). Average reliability, across the nine age groups, are as follows: 0.97 for the Verbal IQ; 0.93 for the Performance IQ; and 0.97 for the Full Scale IQ.

Stability coefficients for the three IQs are: 0.97, 0.90, and 0.96 for Verbal, Performance, and Full Scale, respectively.

Many factor analytic studies are available which examine the underlying structure of the WAIS (i.e., *WAIS-R manual,* 1981). Three basic factors have been reported: a "verbal comprehension" factor, a "perceptual organization" factor, and a "memory" factor, which has also been assigned labels like Freedom from Distractibility, Sequential Ability, and Number Ability. These findings are noted to confirm the appropriateness of separating the tests of the WAIS into the Verbal and Performance Scales. Researchers have disagreed about how many factors do underlie the WAIS-R. Some researchers regard the WAIS-R as a one-factor test, stating that the common ability factors account for only a small measure of intellectual ability (O'Grady, 1983). Some have interpreted as many as four or five meaningful WAIS factors for various normal and clinical samples (Cohen, 1957). However, Kaufman (1990) states that there is not any justification for interpreting five WAIS-R factors.

When only two factors are rotated for the WAIS-R, the results do not quite correspond to Wechsler's division of subtests into the Verbal and Performance Scales, although the fit is adequate. In a comparison of six cross-validation samples and the total normative sample using two-factor solutions, all of the Verbal subtests loaded more highly on the Verbal Comprehension than the Perceptual Organization factor (Kaufman, 1990). The loadings from the standardization sample ranged from 0.47 to 0.84 for the Verbal Conceptualization factor, and ranged from 0.45 to 0.72 for the Perceptual Organization factor. Two Verbal tests (Digit Span and Arithmetic) did, however, show strong secondary loadings on the Performance factor. Digit Span and Arithmetic's loadings on the Verbal dimension are also not as consistently strong as the other four Verbal Subtests. Each of the five Performance subtests also consistently loaded more highly on the Perceptual Organization than the Verbal Comprehension factor for the total standardization sample and for the various supplementary groups. Picture Arrangement was the exception, with equal loadings on both factors of the normative group (Kaufman, 1990).

The three-factor solutions for the normal WAIS-R standardization sample demonstrated factors that were fairly well anticipated. Kaufman (1990) discusses the three factor solutions, and presents a table summarizing the data from six samples plus the normative sample (p. 244). The Verbal Comprehension factor was defined by loadings ranging from 0.67 to 0.81, including: Information, Vocabulary, Comprehension, and Similarities. The triad of Picture Completion, Block Design, and Object Assembly comprised the Perceptual Organization factor with loadings in the 0.56–0.73 range. The third factor comprises Digit Span and Arithmetic, with factors of 0.64 and 0.55, respectively. Picture Arrangement and Digit Symbol are more or less unaccounted for in the three-factor solution. Picture Arrangement loads equally on the verbal and nonverbal dimensions. Digit Symbol achieves loadings of only 0.32, 0.38, and 0.36 for each of the factors, loading only marginally on each, but not definitively on any. Depending on the profile obtained by any given individual, examiners may choose to interpret either two or three factors (Kaufman, 1990). The decision to interpret two or three factors should be based on whether the small third dimension is interpretable for a given person.

Studies on gender differences on the WAIS-R have shown that males' earned IQs were higher (although not significantly so) than females' earned IQs (Kaufman, 1990). In a sample of 940 males and 940 females, males scored about two points higher on the VIQ, 1.5 points higher on the PIQ, and two points higher on the FSIQ (Renolds, Chastain, Kaufman, & McLean, 1987). When the gender differences are examined within different age groups, there are larger differences for ages 20–54 than at the extreme age groups of 16–19 and 55–74. For the 20–54 year age range, males scored higher by about 2.5 points on VIQ and PIQ, and by about three points on the FSIQ (Kaufman, 1990).

In examining gender differences on the individual subtests (Kaufman, McLean, & Reynolds, 1988), males and females were found to perform differently on some of the 11 subtests. On Information, Arithmetic, and Block Design males significantly and consistently outperformed females. However, females were far superior on Digit Symbol. On a less consistent basis, males showed superiority on Comprehension, Picture Completion, Picture Arrangement, and Object Assembly. No gender differences for any age group were found for Digit Span, Vocabulary, and Similarities.

Research on WAIS-R profiles has also focused on the area of neuropsychology. In this area it has been hypothesized that lesions in the left cerebral hemisphere are associated with diminished language and verbal abilities, whereas lesions in the right cerebral hemisphere are accompanied by visual–spatial deficits (Reitan, 1955). The hypothesis that has grown from these expected differences is that individuals with left brain lesions will have WAIS-R profiles with $P > V$, and individuals with right hemisphere lesions will have a profile with V > P (Kaufman, 1990). On the basis of numerous WAIS and WAIS-R studies of patients with brain lesions, two general conclusions may be drawn. A summary of empirical data (Kaufman, 1990) leads to a few main conclusions, as follows (see Table 3). First, patients with right hemisphere damage (unilateral or some bilateral damage as well) will most likely demonstrate a V > P profile. Second, patients with left hemisphere, unilateral damage may show a slight P > V profile, but not large enough in size or consistently enough that it is beneficial diagnostically.

A further area of study in subjects with unilateral brain lesions and cognitive ability is with gender differences. Males and females are believed to differ in various aspects of brain functioning. Kaufman (1990) presents data from eight studies that included males and females with brain lesions. The accumulated data are reported to support the alleged gender-related interaction between side of brain lesion and direction of Verbal IQ–Performance IQ difference. Damage to the right hemisphere for both males and females lead to more striking V–P differences than damage to the left hemisphere. However, the V > P of 12 points for males with right lesions is nearly twice the value of 6.5 points for females. For males, the six-point P > V difference for patients with left damage supports the hypothesis of depressed Verbal IQ for left hemisphere lesions. However, the P > V discrepancy of only 1.6 points for females with left hemisphere lesions does not support the reversed hypothesis (Kaufman, 1990). This difference across genders for adults with brain lesions may indicate that women have different cerebral organization than men. However, data supporting the reason for the interaction with gender is not definitive (Kaufman, 1990).

Turkheimer and Farace (1992) performed a meta-analysis of 12 different studies which used Wechsler IQ data to examine both male and female patients with right or left hemisphere damage, including a variety of etiologies. The researchers noted a problem in the previous literature to be the use of the difference between the PIQ and VIQ in measuring the effects of lesions. The V–P differences are determined by potentially separate effects of each hemisphere on the IQs. Thus, in this meta-analysis, separate VIQ and PIQ means were reported for men and women with either right or left hemisphere lesions (Turkheimer & Farace, 1992). The results of the repeated-measures analysis revealed that left hemisphere lesions produce substantial and roughly equal VIQ deficits in male and female patients, but lower mean PIQ scores in female than in male patients. Right hemisphere lesions produce PIQ deficits in both genders, but lower mean VIQ scores in female patients. Mean scores from Turkheimer and Farace's (1992) data are presented in Table 4.

The main effect indicated by the data presented in Table 4 is that "female patients are more sensitive to lesions in the hemisphere

Table 3 Effects of unilateral brain damage on WAIS/WAIS-R VIQ–PIQ discrepancies.

Group	*Sample size*	*Mean VIQ minus mean PIQ*	
		Left damage	*Right damage*
Stroke			
Men	124	−10.1	+16.8
Women	81	+0.1	+9.5
Total	248	−6.4	+13.5
Tumor (generalized or posterior)	200	+0.2	+8.4
Frontal lobe	104	−2.2	+2.6
Temporal lobe epilepsy			
Preoperative	101	−3.1	+2.4
Postoperative	101	−6.4	+6.0
Acute lesion	109	−2.4	+14.2
Chronic lesion	131	−2.5	+5.5
Age 20–34	664	−5.0	+6.7
Age 35–54	1245	−3.9	+9.5
Age 55+	168	−2.9	+14.9
Whites	50	−5.2	+15.1
Blacks	50	+5.7	+10.4

Source: Kaufman (1990).

Table 4 Gender differences and brain damage on WAIS/WAIS-R.

	Men	*Women*
Left damage		
VIQ	91	91
PIQ	95	91
V–P	−4	0
Right damage		
VIQ	104	99
PIQ	90	91
V–P	+14	+8

Source: Turkheimer and Farace (1992).
Note: V–P = Verbal IQ minus Performance IQ.
Total sample size = 983.

opposite to that thought to be 'odominant' for a function" (Turkheimer & Farace, 1992, p. 499). Although these results are consistent with previously reported greater V–P differences in males, the analysis show that there is also no difference between male and female patients in the effects of left hemisphere lesions on VIQ, or right hemisphere lesions on PIQ. The females demonstrated a pattern of lower mean scores following lesions to hemisphere opposite to the "dominant" hemisphere for each function. This pattern is supportive of a model which asserts that there is a greater degree of bilateral processing in women (Turkheimer & Farace, 1992). This gender difference could be the result of many things including: degree of hemispheric lateralization, differences in problem-solving strategy, or callosal function.

According to Turkheimer, Farace, Yfo, and Bigler (1993), two major findings have been suggested by earlier studies. Individuals with lesions in the left hemisphere have smaller Verbal IQ–Performance IQ differences than subjects with lesions in the right hemisphere, and this difference is greater for males than females. Theories of why gender differences exist can be evaluated through the lesion data. The degree of lateralization in males and females cannot account for gender differences in PIQ and VIQ, because a "statistical model in which the genders have the same degree of lateralization fits the data as well as a model in which the genders are allowed to differ" (Turkheimer et al., 1993, p. 471). There was also not support for the hypothesis that the gender difference results from differences in the within-hemisphere organization of verbal skills.

In a study examining 64 patients through archival data, Turkheimer et al. (1993) did find Inglis and Lawson's (1982) hypothesis that females use verbal strategies in solving PIQ items to be supported by their data. In females, a single model of lesion effects could account for deficits in VIQ and PIQ, but this was not found for males. The most striking observation made was that females with left-hemisphere lesions had substantial deficits in PIQ related to lesion parameters, but males with left-hemisphere lesions did not (Turkheimer, 1993). Notably, this difference could be present because in the left-hemisphere females may have more nonverbal abilities relevant to PIQ, or females may use more verbal strategies in solving PIQ items. Further research examining problem-solving strategy is necessary to clarify the reason for these gender differences.

(ii) Overview

The WAIS-R has proven itself as a leader in the field of adult assessment. Kaufman (1985) stated, "The WAIS-R is *the* criterion of adult intelligence, and no other instrument is even close." Matarazzo (1985) had an equally favorable review of the WAIS-R, applauding its strong psychometric qualities and clinical usefulness. It has strong reliability and validity for Verbal, Performance, and Full Scale IQs, as did its predecessor, the WAIS. The separate subtests, however, have split-half reliability coefficients that are below 0.75 for six of the 11 tasks at ages 16–17, and for tasks across the age range (Picture Arrangement and Object Assembly) (Kaufman, 1985). The sample selection included apparent systematic bias in the selection of individuals ages 16–19, leading to very questionable teenage norms (Kaufman, 1985). However, the rest of the sample selection was done with precision, leading to an overall well-stratified sample.

Administration is not difficult with the clear and easy to read WAIS-R manual, which provides good scoring directions (Spruill, 1984). The administration and scoring rules of the WAIS-R were made more uniform with the WISC-R rules, which facilitates transfer (Kaufman, 1985). In addition, for the Verbal items with subjective scoring systems, the scoring criteria has been expanded to reduce ambiguity; and to facilitate administration, all words spoken by the examiner are placed on separate lines of the administration manual (Kaufman, 1985).

The WAIS-R does have its limitations; some of which are the nonuniformity of the scaled scores, and the limited floor and ceiling (Spruill, 1984). Individuals who are extremely gifted or severely retarded cannot be assessed adequately with the WAIS-R because the range of possible Full Scale IQ scores is only 45–150. Several subtests have insufficient floors for adequate assessment of retarded individuals: Picture Arrangement, Arithmetic, Similarities, Picture Completion, and Block Design (Kaufman, 1985). If evaluating an individual who falls at the extreme low end, this is a distinct disadvantage. In addition, even if a subject receives a raw score of zero on a subtest, they can receive one scaled-score point on that subtest.

The WAIS's method of using a reference group (ages 20–34) to determine everyone's scaled scores was retained in the development of the WAIS-R. Kaufman (1985) stated that this method is "indefensible," because use of this single reference group impairs profile interpretation below age 20 and above 34. Profile interpretation is further impaired for individuals aged 16–17 because of low subtest reliability. The WAIS and WAIS-R studies at ages 35–44 cannot be generalized to individuals aged 16–19 because of poor teenage norms, again negatively impacting a clinician's ability to interpret the profile. In the WAIS-R appendix, clinicians must utilize separate scaled-score tables which are arranged by age group. These separate tables invite clerical errors and confusion in case reports (Kaufman, 1985). The WAIS-R manual itself fails to provide appropriate empirical guidelines for profile interpretation, showing a limited awareness of clinicians' practical needs (Kaufman, 1985). However, despite these limitations, the WAIS-R is still one of the most readily chosen instruments in the assessment of intelligence.

4.08.2.1.5 *Wechsler Adult Intelligence Scale-Third Edition (WAIS-III)*

(i) Description

The newest member of the Wechsler family of tests is the WAIS-III (The Psychological Corporation, 1997). The WAIS-III is an instrument for assessing the intellectual ability of individuals aged 16–89. Like the other Wechsler scales discussed, three IQ scores (Verbal, Performance, and Full Scale) and four factor indices (Verbal Comprehension, Perceptual Organization, Working Memory, and Processing Speed) are derived from the WAIS-III subtests. Each of these scores are standard scores with a mean of 100 and a standard deviation of 15, which are created by comparing an individual's score to scores earned by the normative group of the same age. The Verbal IQ is comprised of six verbal subtest scores (Vocabulary, Similarities, Arithmetic, Digit Span, Information, and Comprehension), plus a new supplementary test to substitute for Digit Span if necessary (Letter–Number Sequencing). The Performance IQ is comprised of five nonverbal subtests (Picture Completion, Picture Arrangement, Block Design, Matrix Reasoning, and the renamed Digit Symbol-Coding). In addition, two supplemental subtests are provided on the Performance scale: Symbol Search (which may be used to replace Digit Symbol-Coding) and Object Assembly (which is an optional subtest that may be used to replace any performance subtest for individuals younger than 75). In addition to its new name, Digit Symbol-Coding also has two new optional procedures not used in IQ computation, which may be used to help the examiner rule out potential problems. These new procedures include Digit Symbol-Incidental Learning and Digit Symbol-Copy. The Full Scale represents the average of the Verbal and Performance IQs.

New to the WAIS-III are additional factor indices, which can be helpful in further breaking down and understanding an individual's performance. Like the WISC-III, there are four factor indices: Verbal Comprehension, Perceptual Organization, Working Memory, and Processing Speed. The two new subtests on the WAIS-III, Letter–Number Sequencing and Symbol Search, are used in calculation of the Working Memory and Processing Speed Indices, respective. Table 5 shows which tests comprise each of the IQs and factor indices.

(ii) Standardization properties of the scale

The WAIS-III was standardized by administering the full scale to 2450 adult subjects, selected according to 1995 US Census data. Subjects were stratified according to age, sex, race/ethnicity, geographic region, and education level. Subjects were divided into 13 age groups, which is an improvement over the nine age groups tested in the WAIS-R standardization sample. The number in each WAIS-III standardization age group ranged from 100 to 200, and the age groups spanned from age 16 to 89. Due to the fact that US citizens are living longer, the WAIS-III developers extended norms beyond the highest age group (74) provided in the WAIS-R. In the collection of normative data, an additional 200 African American and Hispanic individuals were also administered the WAIS-III without discontinue rules. "This over sampling provided a sufficient number of item scores across all items for item bias analyses" (The Psychological Corporation, 1997, p. 20).

Reliability coefficients for the 14 subtests and the Verbal, Performance, and Full Scale IQs were computed using the split-half method (except for Digit Symbol-Coding and Symbol Search). Average reliability, across the 13 age

Table 5 Subtests comprising WAIS-III IQs and Index Scores.

Subtest	*IQ scale*		*Factor index*			
Vocabulary	VIQ		VCI			
Similarities	VIQ		VCI			
Information	VIQ		VCI			
Comprehension	VIQ					
Arithmetic	VIQ				WMI	
Digit Span	VIQ				WMI	
Letter–number Sequencing[a]					WMI	
Picture Arrangement		PIQ				
Picture Completion		PIQ		POI		
Block Design		PIQ		POI		
Matrix Reasoning		PIQ		POI		
Digit Symbol-coding		PIQ				PSI
Symbol Search[a]						PSI
Object Assembly[a]						

Note. Verbal IQ (VIQ); Performance IQ (PIQ); Verbal Comprehension Index (VCI); Perceptual Organization Index (POI); Working Memory Index (WMI); Processing Speed Index (PSI).

[a] The Letter–Number Sequencing, Symbol Search, and Object Assembly subtests can substitute for other subtests under certain circumstances (see The Psychological Corporation, 1997).

groups, is as follows: 0.97 for the Verbal IQ; 0.94 for the Performance IQ; and 0.98 for the Full Scale IQ. The average individual subtests' reliabilities ranged from 0.93 (Vocabulary) to 0.70 (Object Assembly), with a median coefficient of 0.85. Stability coefficients for the three IQs are: 0.96, 0.91, and 0.96 for Verbal, Performance, and Full Scale, respectively. The stability coefficients for individual subtests ranged from an average of 0.94 (Information) to 0.69 (Picture Arrangement) with a median coefficient of 81.5.

The WAIS-III manual (The Psychological Corporation, 1997) reports that numerous factor analytic studies (exploratory and confirmatory) examined the underlying structure of the WAIS-III. There were four basic factors predicted to be underlying the WAIS-III: Verbal Comprehension, Perceptual Organization, Working Memory, and Processing Speed. Overall, results of exploratory and confirmatory factor analysis support the appropriateness of separating the tests of the WAIS into the four factors. The manual reports that the four factor model is a "clearly superior solution to a one-, two-, or three-factor solution and more parsimonious than a five-factor one" (p. 110). Except for the oldest age group, the findings across all ages are similar. However, in the 75–89 year age range, many more subtests loaded on the Processing Speed factor than the Perceptual Organization Factor (i.e., Block Design, Picture Completion, and Picture Arrangement all load on Processing Speed). Only Matrix Analogies had a factor loading above 0.40 on the Perceptual Organization factor for the oldest age group. From the data presented with the standardization sample, it appears that the WAIS-III is best represented by the four factors that were originally predicted to underlie it.

Across all ages, the Verbal Comprehension factor was defined by loadings ranging from 0.76 to 0.89, including: Information, Vocabulary, Comprehension, and Similarities. Picture Completion, Block Design, Matrix Reasoning, and Picture Arrangement comprised the Perceptual Organization factor with loadings in the 0.47–0.71 range. The third factor is comprised of Digit Span, Arithmetic, and Letter–Number Sequencing with factor loadings of 0.71, 0.51, and 0.62, respectively. Symbol Search and Digit Symbol-Coding are assumed in the Processing Speed factor, with loadings of 0.63 and 0.68, respectively. The Symbol Search subtest requires the examinee to determine whether a pair of target symbols are present in a larger group of shapes within a specified time limit. The addition of the new subtests seems to have strengthened the factor structure, as in the previous version of the WAIS, some of the subtests did not load strongly on any of the factors or loaded similarly across the factors (i.e., Picture Arrangement and Digit Symbol).

(iii) Preliminary research with the WAIS-III

The WAIS-R and the WAIS-III were compared to see how well they were related (The Psychological Corporation, 1997). A sample of 192 individuals with a mean age of 43.5 years (ranging from 16 to 74) were administered the two tests. The median time between administrations was 4.7 weeks. As would be predicted by work done by Flynn (1984), subjects scored 2.9 points lower on the WAIS-III FSIQ than on the WAIS-R FSIQ. The WAIS-III VIQ and PIQ were 1.2 points and 4.8 points lower than the respective WAIS-R scales. The overall correlations between the WAIS-III and WAIS-R global scales were high. The correlation coefficients for the VIQ, PIQ, and FSIQ were 0.94, 0.86, and 0.93, respectively.

The WAIS-III and WISC-III were also administered to a sample of adolescents to determine how well the two tests correlated (The Psychological Corporation, 1997). The sample consisted of 184 16-year olds who were administered the two tests from 2 to 12 weeks apart (median time = 4.6 weeks). The correlations between the global scales of the two tests were very high, indicating that the two instruments appear to be measuring very similar constructs. The VIQ, PIQ, and FSIQ correlation coefficients were 0.88, 0.78, and 0.88, respectively. The Index scores from the WAIS-III and WISC-III were also compared. The Indices' correlations were 0.87, 0.74, 0.80, and 0.79 for the VCI, POI, WMI, and PSI, respectively. The differences between the mean WISC-III and WAIS-III IQs were all less than one point. The differences between the two tests' mean VCI and POI were also each less than one point. The difference between the WAIS-III and WISC-III mean WMI was 1.7 standard score points. On the PSI, the difference between the means on the two tests was 2.7 points. Thus, overall, the IQ and Indices of the two tests correspond quite highly.

The WAIS-III Technical Manual (The Psychological Corporation, 1997) also presents some studies collected from clinical groups with neurological, psychiatric, and developmental disorders. Reviewed here are a select group of these studies, including those from a sample of patients with mild Alzheimer's disease, a sample of individuals who are mentally retarded, and one from individuals with attention-deficit hyperactivity disorder (ADHD).

Individuals with probable Alzheimer's disease ($N = 35$) were administered both the WAIS-III and the Wechsler Memory Scales-Third Edition (WMS-III). Decrements in cognitive ability and memory were predicted. This sample was reported to have a significantly higher level of education than the normal population, with 48.6% of the sample having completed at least four years of college. The results of this study (The Psychological Corporation, 1997) show that the individuals with probable Alzheimer's disease had lower scores on all IQ scales than the general population. The mean VIQ, PIQ, and FSIQ scores were 92.2, 81.7, and 86.6, respectively. The PIQ scores tend to be more sensitive to the effects of this neurological condition; thus, the mean VIQ score is predictably higher than the PIQ score. The mean factor index scores demonstrate more differentiation in their cognitive profile. This sample had a mean factor indices of 79.6 (PSI), 84.8 (POI), 87.2 (WMI), and 93.0 (VCI).

A total of 108 individuals diagnosed with mental retardation were administered the WAIS-III. Six of these, 46 were categorized as having mild mental retardation, while the other 62 had moderate mental retardation (The Psychological Corporation, 1997). The results demonstrated deficits across all areas of cognitive functioning, as expected. In the mildly mentally retarded group, mean IQ scores were as follows: 60.1 (VIQ), 64.0 (PIQ), and 58.3 (FSIQ). The subjects with moderate mental retardation exhibited lower scores, earning mean VIQ, PIQ, and FSIQ scores of 54.7, 55.3, and 50.9, respectively. The variability in scores of each of these clinical groups is much smaller than found in the general population. The standard deviations are more than 50% smaller than those found in the general population.

ADHD sufferers was another group studied and reported in the WAIS-III technical manual (The Psychological Corporation, 1997). Traditionally, IQ scores have not be useful in discriminating ADHD from non-ADHD individuals. However, examining subtests patterns on tests of cognitive ability has been more fruitful in discriminating those with ADHD from those without.

The WAIS-III was administered to 30 individuals diagnosed with ADHD (mean age 19.8 years). The mean level of intellectual functioning was found to be in the Average range for this group (mean FSIQ = 103.00). In addition, there was no significant difference found between Verbal and Performance IQs for the group (104.2 and 100.9, respectively). The WAIS-III factor indices were also examined, and the pattern of performance on the indices was found to differ in comparison to the general population. The ADHD sample scored on average 8.3 points lower on the WMI than the VCI. About 30% of the ADHD sample had WMI scores at least 1 SD lower than their VCI scores, whereas 13% of the WAIS-III standardization sample had obtained discrepancies of this magnitude. On average, the ADHD sample scored 7.5 points lower on the PSI than the POI. In the ADHD sample, 26% of the group had PSI scores at least 1 SD lower than their POI scores, but only 14% of the WAIS-III standardization sample had such discrepancies.

(iv) Analyzing the WAIS-III data

The WAIS-III manual provides a general description of how a clinician may begin to interpret the plethora of data obtained in its 14 subtests. However, to obtain the maximum amount of information from the profile, one should utilize an approach to profile interpretation that will group and regroup subtests (Kaufman & Lichtenberger, in press). Similar to the WISC-III interpretation, the WAIS-III may be examined from the global level (IQs) to the individual profile (subtest) level. An organized, systematic approach will be advantageous in obtaining the most accurate picture of the individual.

The WAIS-III record form provides a nice beginning to profile interpretation. However, a structure is needed to work through the large amount of data in a systematic fashion. Kaufman and Lichtenberger, (in press) have developed a series of 10 steps to aid the clinician in interpreting and integrating all of the data obtained from the WAIS-III's three IQs and four factor indices, while not becoming overwhelmed with the multiple scores and difference scores. The 10 steps, a step-by-step approach to WAIS-III profile interpretation, is presented in summary form below. The 10 steps are presented in Table 6. Using these 10 steps can help to organize the information to generate meaningful hypotheses about personal cognitive strengths and weaknesses in preparation for clear presentation in the form of a written report.

(v) Overview

WAIS-III is likely to follow in its footsteps of the WAIS-R, which has proven itself as a leader in the field of adult assessment. The new norms and psychometric improvements of the WAIS-III are much welcomed by the assessment community. The WAIS-III has strong reliability and validity for Verbal, Performance, and Full Scale IQs, as did its predecessor, the WAIS-R. The subtest with the lowest split half reliability has been removed from the computation of the IQs (Object Assembly). However, Picture Arrangement still exhibits split half reliability coefficients below 0.75 at several ages, and it remains part of the Performance IQ. The WAIS-III sample selection was done with precision, leading to an overall well-stratified sample.

Many visual and practical improvements were made in the development of the WAIS-III. Administration is not difficult with the clear and easy to read WAIS-III manual, in addition to the record form with ample space and visual icons (Kaufman & Lichtenberger, in press). The administration and scoring rules of the WAIS-III were made more uniform throughout the entire test, which reduces chances of examiner error. In addition, the scoring rules are listed right in the administration manual for the Verbal items with subjective scoring systems. This change from the WAIS-R has eased the work of clinicians (no more flipping back and forth during administration).

The WAIS-III attempted to improve its floor and ceiling in comparison to the earlier version. Several step-down items have been added on each subtest for lower functioning individuals. However, like the WAIS-R, individuals who are extremely gifted or severely retarded cannot be adequately assessed with the WAIS-III because the range of possible Full Scale IQ scores is only 45–150. Studies on individuals in the lower extreme range of the WAIS-III are yet to determine whether evaluating an individual who falls at the extreme low end is a distinct disadvantage. As on the WAIS-R, even if a subject receives a raw score of zero on a subtest, they can receive one to five scaled-score points on that subtest. Uniformity is not found across the range of scaled scores for each subtest. On certain subtests subjects may reach a ceiling more quickly than on others. At certain ages, ordinarily the highest scaled score that can be obtained on the subtest is 19; however, on the Arithmetic or Picture Arrangement subtest, 17 is the maximum score (Kaufman & Lichtenberger, in press). Profile analysis is made difficult because of this nonuniformity across subtests, especially for the extremely gifted subjects.

The method of the WAIS-R of using a reference group (ages 20–34) to determine everyone's scaled scores was not retained in the development of the WAIS-III. Kaufman (1985) stated that this WAIS-R method is "indefensible," because use of this single reference group impairs profile interpretation below age 20 and above 34; thus, the WAIS-III change in determining scaled scores is a significant improvement. The process of profile interpretation has been made much less confusing by the removal of the reference group scores. (However, if one wants to calculate the scores by using the 20–34 reference group, this is still possible.) Fewer clerical errors and less confusion in case reports will be present because of these changes (Kaufman & Lichtenberger, in press). The WAIS-III manual and record form themselves provide the beginning to interpretation, with clearly laid out tables to calculate score discrepancies, and so forth. However, to meet clinicians' practical needs for more specific empirical guidelines for profile interpretation in a systematic and step-by-step fashion, other

Table 6 Ten tips for WAIS-III interpretation.

Step 1. Interpret the Full Scale IQ.
(i) Report the FSIQ confidence interval (The Psychological Corporation, 1997; Table A.5 p. 197).
(ii) Report the FSIQ percentile rank (The Psychological Corporation, 1997; Table A.5 p. 197).
(iii) Report the FSIQ ability level (The Psychological Corporation, 1997, Table 2.3 p. 25).
(iv) If in STEPS 2–6, it is determined that there is a significant difference between the component parts of the FSIQ (i.e., VIQ & PIQ or VCI & POI), the FSIQ should not be interpreted as a meaningful representation of the individual's overall performance.

Step 2. Determine if the Verbal–Performance IQ discrepancy is statistically significant.
(i) For all ages VIQ–PIQ difference of *6 points* is significant at *0.15 level.*
(ii) For all ages VIQ–PIQ difference of *9 points* is significant at *0.05 level.*
(iii) Values for specific ages are presented in Table B.1 (The Psychological Corporation, 1997; p. 205).

Step 3. Determine if the VIQ and PIQ are interpretable. Four questions to consider about the Verbal and Performance Scales.

Verbal Scale
(i) Is the difference between VCI and WMI statistically significant ($p < 0.05$)?
Size needed for difference = *10+ points.*
(ii) Is there abnormal scatter among the VIQ subtests?
Highest of 6 VIQ subtest scaled scores minus lowest = *8+ points.*

Performance Scale
(iii) Is the difference between POI and PSI statistically significant ($p < 0.05$)?
Size needed for difference = *13+ points.*
(iv) Is there abnormal scatter among the PIQ subtests?
Highest of 5 PIQ subtest scaled scores minus lowest = *8+ points.*

Step 4. Determine if VIQ–PIQ discrepancy is interpretable or if the VCI and POI should be interpreted instead.
(i) If all answers in Step 3 are *no*, then the VIQ–PIQ discrepancy is interpretable. Skip Step 5 and go directly to Step 6.
(ii) If the answer to one or more of the questions in Step 3 is *yes*, then the VIQ–PIQ discrepancy may not be interpretable.
(iii) If VIQ–PIQ is not interpretable, then examine the VCI–POI discrepancy in Step 5.

Step 5. Determine whether VCI and POI are interpretable and significantly different from one another.
(i) Is there abnormal scatter among the VCI subtests?
Highest of 3 VCI subtest scaled scores minus lowest = *5+ points.*
(ii) Is there abnormal scatter among the POI subtests?
Highest of 3 POI subtest scaled scores minus lowest = *6+ points.*
(iii) If the answer to either (i) or (ii) is yes, then VCI–POI difference may not be interpretable. Otherwise, if both answers are no, examine the interpretable VCI–POI difference:
    (a) Is the difference between VCI and POI statistically significant ($p < 0.05$)?
    (b) Size needed for difference = *10+ points.*

Step 6. Determine if the VIQ–PIQ discrepancy (or the VCI–POI discrepancy) is abnormally large.
(i) *17 Point* difference is abnormally large for the VIQ–PIQ.
(ii) *19 Point* difference is abnormally large for the VCI–POI.
(iii) Exact point values according to ability level are available in Appendix D (The Psychological Corporation, 1997, pp. 300–309)
(iv) If the discrepancies are abnormally large, this indicates that they are too big to ignore (see Steps 4 & 5), and they may be interpreted anyway.

Step 7. Determine whether the Working Memory and Processing Speed indices are interpretable.
(i) Do not interpret WMI if scatter among the 3 subtests is = *6+ points.*
(ii) Do not interpret PSI if difference among 2 subtests is = *4+ points.*

Step 8. Interpret the Global Verbal and Nonverbal Dimensions, as well as the small factors, if they were found to be interpretable.
Study the information and procedures presented in (Kaufman & Lichtenberger, in press).

Step 9. Interpret significant strengths and weaknesses in the WAIS-III subtest profile.
(i) If the VIQ–PIQ discrepancy is less than 17 points, use the individual's mean of all WAIS-III subtests as the person's midpoint.
(ii) I the VIQ–PIQ discrepancy is 17 or more points, use 2 separate means:
    (a) Use the individual's mean of all the Verbal subtests as the midpoint for determining strengths and weaknesses on Verbal subtests;

Table 6 (continued)

(b) Use the individual's mean of all the Performance subtests as the midpoint for determining strengths and weaknesses on Performance subtests.

(iii) Subtract the individual's mean from each of the subtest scaled scores to determine strengths and weaknesses. Round to the nearest whole number.

(iv) Values are presented in Table B.3 (The Psychological Corporation, 1997, p. 208) for determining if a subtest significantly deviates from the individual's own mean. The following summary information may also be used to determine significance:

±2 points: Vocabulary
±3 points: Similarities, Arithmetic, Digit Span, Information, Comprehension, Coding, Block Design, Matrix Reasoning,
±4 points: Letter-Number Sequencing, Picture Completion, Picture Arrangement, Symbol Search
±5 points: Object Assembly.

Step 10. Generate hypotheses about the fluctuations in the WAIS-III subtest profile.

Review the information presented in Kaufman and Lichtenberger (in press) which detail how to reorganize subtest profiles to systematically generate hypotheses about strengths and weaknesses.

---

sources are available (Kaufman & Lichtenberger, in press). Undoubtedly, like its predecessors, the WAIS-III is likely to become one of the most readily chosen instruments in the assessment of intelligence.

4.08.2.1.6 Kaufman Assessment Battery for Children (K-ABC)

The K-ABC is a battery of tests measuring intelligence and achievement of normal and exceptional children aged 2.5–12.5. It yields four scales: Sequential Processing, Simultaneous Processing, Mental Processing Composite (Sequential and Simultaneous), and Achievement.

The K-ABC is becoming a frequently used test in intelligence and achievement assessment that is used by both clinical and school psychologists (Kamphaus, Beres, Kaufman, & Kaufman, 1995). In a nationwide survey of school psychologists conducted in 1987 by Obringer (1988), respondents were asked to rank the following instruments in order of their usage: Wechsler's scales, the K-ABC, and both the old and new Stanford–Binets. The Wechsler scales earned a mean rank of 2.69, followed closely by the K-ABC with a mean of 2.55, the L-M version of the Binet (1.98), and the Stanford–Binet Fourth Edition (1.26). Bracken (1985) also found similar results of the K-ABC's popularity. Bracken surveyed school psychologists and found that for ages 5–11 years the WISC-R was endorsed by 82%, the K-ABC by 57%, and the Binet IV by 39% of the practitioners. These results suggest that clinicians working with children should have some familiarity with the K-ABC (Kamphaus et al., 1995).

The K-ABC has been the subject of great controversy from the outset, as evident in the strongly pro and con articles written for a special issue of the *Journal of Special Education* devoted to the K-ABC (Miller & Reynolds, 1984). Many of the controversies, especially those regarding the validity of the K-ABC theory, will likely endure unresolved for some time (Kamphaus et al., 1995). Fortunately, the apparent controversy linked to the K-ABC has resulted in numerous research studies and papers that provide more insight into the K-ABC and its strengths and weaknesses.

(i) Theory

The K-ABC intelligence scales are based on a theoretical framework of Sequential and Simultaneous information processing, which relates to how children solve problems rather than what type of problems they must solve (e.g., verbal or nonverbal). In stark contrast is Wechsler's theoretical framework of the assessment of "g," a conception of intelligence as an overall global entity. As a result, Wechsler used the Verbal and Performance scales as a means to an end. That end is the assessment of general intelligence. In comparison, the Kaufmans emphasize the individual importance of the Sequential and Simultaneous scales in interpretation, rather than the overall Mental Processing Composite (MPC) score (Kamphaus et al., 1995).

The Sequential and Simultaneous framework for the K-ABC stems from an updated version of a variety of theories (Kamphaus et al., 1995). The foundation lies in a wealth of research in clinical and experimental neuropsychology and cognitive psychology. The Sequential and Simultaneous theory was primarily developed from two lines of theory: the information processing approach of Luria (e.g., Luria, 1966), and the cerebral specialization work of Sperry (1968, 1974), Bogen (1969), Kinsbourne (1975), and Wada, Clarke, and Hamm (1975).

The neuropsychological processing model, which originated with the neurophysiological observations of Luria (1966, 1973, 1980) and Sperry (1968), the psychoeducational research of Das (1973; Das et al., 1975; Das, Kirby, & Jarman, 1979; Naglieri & Das, 1988, 1990), and the psychometric research of Kaufman and Kaufman (1983), possesses several strengths relative to previous models in that it (i) provides a unified framework for interpreting a wide range of important individual difference variables; (ii) rests on a well-researched theoretical base in clinical neuropsychology and psychobiology; (iii) presents a processing, rather than a product-oriented, explanation for behavior; and (iv) lends itself readily to remedial strategies based on relatively uncomplicated assessment procedures (Kaufman & Kaufman, 1983; McCallum & Merritt, 1983; Perlman, 1986).

This neuropsychological processing model describes two very distinct types of processes which individuals use to organize and process information received in order to solve problems successfully: successive or sequential, analytic-linear processing vs. holistic/simultaneous processing (Levy & Trevarthen, 1976; Luria, 1966). These processes have been identified by numerous researchers in diverse areas of neuropsychology and cognitive psychology (Perlman, 1986). From Sperry's cerebral specialization perspective, these processes represent the problem-solving strategies of the left hemisphere (analytic/sequential) and the right hemisphere (Gestalt/holistic). From Luria's theoretical approach, successive and simultaneous processes reflect the "coding" processes that characterize "Block 2" functions.

Regardless of the theoretical model, successive processing refers to the processing of information in a sequential, serial order. The essential nature of this mode of processing is that the system is not totally surveyable at any point in time. Simultaneous processing refers to the synthesis of separate elements into groups. The essential nature of this mode of processing is that any portion of the result is, at once, surveyable without dependence on its position in the whole. The model assumes that the two modes of processing information are available to the individual. The selection of either or both modes of processing depends on two conditions: (i) the individual's habitual mode of processing information as determined by social–cultural and genetic factors, and (ii) the demands of the task (Das et al., 1975).

In reference to the K-ABC, Simultaneous processing refers to the mental ability to integrate information all at once to solve a problem correctly. Simultaneous processing frequently involves spatial, analogic, or organizational abilities (Kaufman & Kaufman, 1983; Kamphaus & Reynolds, 1987). There is often a visual aspect to the problem and visual imagery used to solve it. A prototypical example of a Simultaneous subtest is the Triangles subtest on the K-ABC, which is similar to Wechsler's Block Design. To solve both of these subtests, children must be able to see the whole picture in their mind and then integrate the individual pieces to create the whole.

In comparison, Sequential processing emphasizes the ability to place or arrange stimuli in sequential or serial order. The stimuli are all linearly or temporally related to one another, creating a form of serial interdependence within the stimulus (Kaufman & Kaufman, 1983). The K-ABC subtests assess the child's Sequential processing abilities in a variety of modes. For example, Hand Movements involves visual input and a motor response, Number Recall involves auditory input with a vocal response, and Word Order involves auditory input and visual response. These different modes of input and output allow the examiner to assess the child's sequential abilities in a variety of ways. The Sequential subtests also provide information on the child's short-term memory and attentional abilities.

According to Kamphaus et al. (1995), one of the controversial aspects of the K-ABC was the fact that it took the equivalent of Wechsler's Verbal Scale and redefined it as "achievement." The Kaufmans' analogs of tests such as Information (Faces & Places), Vocabulary (Riddles and Expressive Vocabulary), and Arithmetic (Arithmetic) are included on the K-ABC as achievement tests and viewed as tasks that are united by the demands they place on children to extract and assimilate information from their cultural and school environment. The K-ABC is predicated on the distinction between problem solving and knowledge of facts. The former set of skills are interpreted as intelligence; the latter is defined as achievement. This definition presents a break from other intelligence tests, where a person's acquired factual information and applied skills influence greatly the obtained IQ (Kaufman & Kaufman, 1983).

(ii) Standardization and properties of the scale

Stratification of the K-ABC standardization sample closely matched the 1980 US Census data on the variables of age, gender, geographic region, community size, socioeconomic status (SES), race or ethnic group, and parental occupation and education. Additionally, unlike most other intelligence measures for children, stratification variables also included educational placement of the child (see Table 7).

Table 7 Representation of the K-ABC standardization sample by educational placement.[a]

Educational placement	*N*	*%*	*%*
Regular classroom	1862	93.1	91.1
Speech impaired	28	1.4	2.0
Learning disabled	23	1.2	2.3
Mentally retarded	37	1.8	1.7
Emotionally disturbed	5	0.2	0.3
Other[b]	15	0.8	0.7
Gifted and talented	30	1.5	1.9[c]
Total K-ABC sample	2000	100.0	100.0

[a] Data from US Department of Education, National Center for Education Statistics. 1980. Table 2.7, The Condition of Education, Washington, DC, US Government Printing Office. [b] Includes other health impaired, orthopedically handicapped, and hard of hearing. [c] Data from US Office for Civil Rights, 1980, *State, Regional, and National Summaries of Data from the 1978 Child Rights Survey of Elementary and Secondary Schools* (p. 5). Alexandria, VA: Killalea Associates.

Reliability and validity data provide considerable support for the psychometric aspects of the test. A test–retest reliability study was conducted with 246 children after a two- to four-week interval (mean interval = 17 days). The coefficients for the Mental Processing Composite were 0.83 for age two years, six months through four years, eleven months; 0.88 for ages five years through eight years, eleven months; and 0.93 for ages nine years to 12 years, five months. Test–retest reliabilities for the Achievement scale composite for the same age groups were 0.95, 0.95, and 0.97, respectively (Kamphaus et al., 1995). The test–retest reliability research reveals that there is a clear developmental trend, with coefficients for the preschool ages being smaller than those for the school-age range. This trend is consistent with the known variability over time that characterizes preschool children's standardization test performance in general (Kamphaus & Reynolds, 1987). Split-half reliability coefficients for the K-ABC global scales range from 0.86 to 0.93 (mean = 0.90) for preschool children, and from 0.89 to 0.97 (mean = 0.93) for children aged 5–12.5 (Kamphaus et al., 1995).

There has been a considerable amount of research done on the validity of the K-ABC. The *K-ABC interpretive manual* (Kaufman & Kaufman, 1983) includes the results of 43 such studies. Construct validity was established by looking at five separate topics: developmental changes, internal consistency, factor analysis (principal factor, principal components, and confirmatory), convergent and discriminant analysis, and correlations with other tests. Factor analysis of the Mental Processing Scales offered clear empirical support for the existence of two, and only two, factors at each age level, and for the placement of each preschool and school-age subtest on its respective scale. Analyses of the combined processing and achievement subtests also offered good construct validation of the K-ABC's three-scale structure (Kaufman & Kamphaus, 1984).

Although the K-ABC and the WISC-III differ from one another in a number of ways, there is strong evidence that the two measures correlate substantially (Kamphaus & Reynolds, 1987). In a study of 182 children enrolled in regular classrooms, the Mental Processing Composite (MPC) correlated 0.70 with WISC-R Full Scale IQ (FSIQ), thus, sharing a 49% overlap in variance (Kamphaus et al., 1995; Kaufman & Kaufman, 1983). There have also been numerous correlational studies conducted with handicapped and exceptional populations that may be found in the *Interpretative manual*. The overall correlation between the K-ABC and the WISC-R range from 0.57 to 0.74, indicating that the two tests overlap a good deal, yet also show some independence (Kamphaus et al.).

(iii) Overview

Although the K-ABC has been the subject of past controversy, it appears that it has held its own and is used often by professionals. The K-ABC is well designed with easy to use easels and manuals. The information in the manuals is presented in a straightforward, clear fashion, making use and interpretation of the K-ABC relatively easy (Merz, 1985). There has been a considerable amount of research done on the validity of the K-ABC and the authors have done a thorough job of presenting much of that information in the manual. The reporting of the reliability and validity data in the manual is complete and understandable. However, there is not enough information presented on the content validity of the test. The various tasks on the subtests on the K-ABC are based on clinical, neuropsychological, and/or other research-based validity; however, a much clearer explication of the rationale behind some of the novel subtests would have been quite helpful (Merz, 1985).

The K-ABC measures intelligence from a strong theoretical and research basis, evident in the quality of investigation in the amount of research data presented in the manual (Merz, 1985). The K-ABC was designed to measure the intelligence and achievement of children aged 2.5–12.5 and the research done to date suggests that in fact the test does just that. The Nonverbal Scale significantly contributes to the effort to addressing the diverse needs of minority groups and language handicapped children. Overall, it appears that the authors of the K-ABC have met the goals listed in the interpretative manual and that this battery is a valuable assessment tool (Merz, 1985).

Keith and Dunbar (1984) present an alternate means of interpreting the K-ABC, based on exploratory and confirmatory factor analytic data. The two K-ABC Reading subtests are eliminated in their alternate analysis, and factors labeled Verbal Memory, Nonverbal Reasoning, and Verbal Reasoning are presented. For school-aged children whose Achievement Scale splits in half, this model may help interpret their profile. A problem with the Keith and Dunbar labels is that they do not offer evidence to support their Verbal Memory and Nonverbal Reasoning labels. Keith and Dunbar conclude that considerable caution be used when interpreting K-ABC results. In the K-ABC Interpretive Manual (Kaufman & Kaufman, 1983), it is also stressed that a child's profile may need to be approached from an alternative model, if the author's model does not create a good interpretation of the profile.

4.08.2.1.7 Kaufman Adolescent and Adult Intelligence Test

The Kaufman Adolescent and Adult Intelligence Test (KAIT) (Kaufman & Kaufman, 1993) is an individually administered intelligence test for individuals between the ages of 11 and more than 85 years. It provides Fluid, Crystallized, and Composite IQs, each a standard score with a mean of 100 and a standard deviation of 15.

(i) Theory

The Horn–Cattell theory forms the foundation of the KAIT and defines the constructs believed to be measured by the separate IQs; however, other theories guided the test development process, specifically the construction of the subtests. Tasks were developed from the models of Piaget's formal operations (Inhelder & Piaget, 1958; Piaget, 1972) and Luria's (1973, 1980) planning ability in an attempt to include high-level, decision-making, more developmentally advanced tasks. Luria's notion of planning ability involves decision-making, evaluation of hypotheses, and flexibility, and "represents the highest levels of development of the mammalian brain" (Golden, 1981, p. 285).

Cattell and Horn (Cattell, 1963; Horn & Cattell, 1966, 1967) postulated a structural model that separates fluid from crystallized intelligence. Fluid intelligence traditionally involves relatively culture-fair novel tasks and taps problem solving skills and the ability to learn. Crystallized intelligence refers to acquired skill, knowledge, and judgments which have been taught systematically or learned via acculturation. The latter type of intelligence is influenced highly by formal and informal education and often reflects cultural assimilation. Tasks measuring fluid ability often involve more concentration and problem solving than crystallized tasks which tend to measure retrieval and application of general knowledge.

Piaget's formal operations depicts a hypothetical-deductive abstract reasoning system that has as its featured capabilities the generation and evaluation of hypotheses and the testing of propositions. The prefrontal areas of the brain associated with planning ability mature at about ages 11–12 years (Golden, 1981), the same ages that characterize the onset of formal operational thought (Piaget, 1972). The convergence of the Luria and Piaget theories regarding the ability to deal with abstractions is striking; this convergence provided the rationale for having age 11 as the lower bound of the KAIT, and for attempting to measure decision making and abstract thinking with virtually every task on the KAIT (Kaufman & Kaufman, 1993).

Within the KAIT framework (Kaufman & Kaufman, 1993), crystallized intelligence "measures the acquisition of facts an problem solving ability using stimuli that are dependent on formal schooling, cultural experiences, and verbal conceptual development" (p. 7). Fluid intelligence "measures a person's adaptability and flexibility when faced with new problems, using both verbal and nonverbal stimuli" (Kaufman & Kaufman, 1993, p. 7). It is important to note that this crystallized-fluid construct split is not the same as Wechsler's 1974, 1981, 1991) verbal–nonverbal split. This was documented in the results of a factor analysis done with the WISC-R and the KAIT that showed the KAIT crystallized subtests loaded highly on the Crystallized/Verbal factor (0.47–0.78), Fluid subtests loaded 0.51–0.88 on the Fluid factor, and Memory for Block Designs also loads 0.41 on the Perceptual Organization factor (Kaufman & Kaufman, 1993; Kaufman, Ishikuma, & Kaufman, 1994). The KAIT Fluid subtests stress reasoning rather than visual–spatial ability, include verbal comprehension or expression as key aspects of some tasks, and minimize the role played by visual-motor speed for correct responding. In addition, the KAIT scales measure what Horn (1989) refers to as broad fluid and broad crystallized abilities, rather than the purer and more specific skill areas that have emerged in Horn's expansion and elaboration of the original Horn–Cattell Gf-Gc theory.

The Core Battery of the KAIT is composed of three Crystallized and three Fluid subtests, and these six subtests are used to compute the IQs. The Expanded Battery also includes two supplementary subtests and two measures of delayed recall that evaluate the individual's ability to retain information that was learned previously in the evaluation during two of the Core subtests. The Core Battery of the KAIT consists of subtests one through six, and subtests one through 10 comprise the Expanded Battery. Each subtest except the supplementary Mental Status task yields age-based scaled scores with a mean of 10 and a standard deviation of three. Sample and teaching items are included for most subtests to ensure that examinees understand what is expected of them for each subtest.

The delayed recall subtests are administered, without prior warning, about 25 and 45 minutes after the administration of the original, related subtests. The two delayed recall subtests provide good measure of an ability that Horn (1985, 1989) calls TSR (Long-Term Storage and Retrieval). TSR "involves the storage of information and the fluency of retrieving it later through association" (Woodcock, 1990, p. 234).

The Mental Status subtest is comprised of 10 simple questions that assess attention and orientation to the world. Most normal adolescents and adults pass at least nine of the 10 items, but the task has special use with retarded and neurologically impaired populations. The Mental Status subtest may be used as a screener to determine if the KAIT can be validly administered to an individual.

(ii) Standardization and properties of the scale

The KAIT normative sample, composed of 2000 adolescents and adults between the ages of 11–94 years was stratified on the variables of gender, racial/ethnic group, geographic region, and SES (Kaufman & Kaufman, 1993).

Mean split-half reliability coefficients for the total normative sample were 0.95 for Crystallized IQ, 0.95 for Fluid IQ, and 0.97 for Composite IQ (Kaufman & Kaufman, 1993). Mean test–retest reliability coefficients, based on 153 identified normal individuals in three age groups (11–19, 20–54, 55–85 +), retested after a one-month interval, were 0.94 for Crystallized IQ, 0.87 for Fluid IQ, and 0.94 for Composite IQ (Kaufman & Kaufman, 1993). Mean split-half reliabilities of the four Crystallized subtests ranged from 0.89 to 0.92 (median = 0.90). Mean values for the four Fluid subtests ranged from 0.79 to 0.93 (median = 0.88) (Kaufman & Kaufman, 1993). Median test–retest reliabilities for the eight subtests, based on the 153 people indicated previously, ranged from 0.72 to 0.95 (median = 0.78). Rebus Delayed Recall had an average split-half reliability of 0.91 and Auditory Delayed Recall had an average value of 0.71; their respective stability coefficients were 0.80 and 0.63 (Kaufman & Kaufman, 1993).

Factor analysis, both exploratory and confirmatory, gave strong construct validity support for the Fluid and Crystallized Scales, and for the placement of each subtest on its designated scale. Crystallized IQs correlated 0.72 with Fluid IQs for the total standardization sample of 2000 (Kaufman & Kaufman, 1993).

Table 8 summarizes the results of correlational studies involving the KAIT and other well known intelligence tests. The values shown in Table 8 support the construct and criterion-related validity of the three KAIT IQs.

Two separate exploratory joint factor analyses were conducted to analyze the KAIT with both the WISC-R and the WAIS-R. Data were obtained from 118 adolescents in the WISC-R sample and 338 adults in the WAIS-R sample. Two-, three-, four-, five-, and six-factor solutions were examined for each analysis using both varimax and oblimin rotations. Three robust factors came out in the three-factor solutions for both the WISC-R and the WAIS-R. Loadings from the first unrotated principal factor (*g* loadings) along with the three-factor solution for the joint analysis of the KAIT and the WISC-R are presented in Table 9. The analogous data for the KAIT and WAIS-R are presented in Table 10.

The results of these joint analyses indicate that the Wechsler subtests and KAIT subtests are about equal as measures of general intelligence. The KAIT subtests have a mean *g* loading of 0.71 and the Wechsler subtests also have a mean *g* loading of 0.71. The most important finding from these analyses is that the

Table 8 Correlations of the three KAIT IQs with standard scores and IQs yielded by other major intelligence tests

Scale		*Age*	*Sample size*	*KAIT crystallized*	*KAIT fluid*	*KAIT composite*
WAIS-R IQ	Verbal	16–19	71	0.85	0.74	0.86
		20–34	90	0.78	0.66	0.78
		35–49	108	0.79	0.74	0.85
		50–83	74	0.85	0.70	0.86
	Performance	16–19	71	0.64	0.70	0.72
		20–34	90	0.60	0.74	0.73
		35–49	108	0.57	0.73	0.73
		50–83	74	0.74	0.66	0.77
	Full scale	16–19	71	0.84	0.79	0.88
		20–34	90	0.77	0.76	0.83
		35–49	108	0.74	0.78	0.85
		50–83	74	0.84	0.70	0.85
WISC-R IQ	Verbal	11–16	118	0.79	0.74	0.83
	Performance	11–16	118	0.67	0.67	0.72
	Full scale	11–16	118	0.78	0.75	0.82
K-ABC	Sequential	11–12	124	0.46	0.44	0.50
	Simultaneous	11–12	124	0.53	0.62	0.63
	Mental processing composite	11–12	124	0.57	0.62	0.66
	Achievement	11–12	124	0.81	0.64	0.82
Stanford–Binet-4	Composite intelligence	11–42	79	0.81	0.84	0.87

Source: Kaufman and Kaufman (1993).

KAIT Fluid subtests and Wechsler Performance subtests seem to measure markedly different constructs. The Fluid and Perceptual Organization factors correlate about as highly with each other as they do the Crystallized/Verbal Factor. The differences between Fluid IQ and Perceptual Organization abilities have been studied and discussed by Woodcock (1990). Woodcock presented evidence that Wechsler's Performance IQ primarily measures Horn's Gv or broad visualization, and not fluid intelligence.

The KAIT benefits from an integration of theories that unite developmental (Piaget), neuropsychological (Luria), and experimental-cognitive (Horn–Cattell) models of intellectual functioning. The theories work well together and do not compete with one another. Together, the theories give the KAIT a solid theoretical foundation that facilitates test interpretation across the broad 11–94-year age range on which the battery was normed.

The changes in crystallized and fluid abilities on the KAIT were examined in a study of 1500 adults aged 17–94 (Kaufman & Horn, 1996). The results of this study indicated that fluid reasoning (Gf) declined steadily across adulthood, and this decline accelerated during the period beginning at about age 55. This finding that fluid ability reaches a peak in development in young adulthood and declines thereafter, at first quite gradually, but more rapidly as age progresses, is in agreement with results from previous research, although a more steep decline in ability was found in the research directed by Horn (1985). Through the 20s, the measure of crystallized knowledge was found to increase, but then showed no increase or decrease until about age 60. Similar findings have been reported in other cross-sectional investigations of fluid and crystallized measures (Kaufman, Kaufman, Chen, & Kaufman, 1996). After age 60, crystallized knowledge was found to decrease as well. Individual differences in education, gender, and ethnicity were found not to alter the basic findings (Kaufman & Horn, 1996).

The KAIT has also been examined with respect to its relationship to adult interests, as demonstrated on the Strong Interest Inventory (SII; Hansen & Campbell, 1985). Kaufman and McLean (1992, November) examined 936 individuals who were administered both the KAIT and the SII. The SII includes six General Occupational Themes (GOTs), including Realistic, Investigative, Artistic, Social,

Table 9 Exploratory joint principal-factor analysis of the KAIT and the WISC-R ($N = 118$).

		Oblimin factor pattern		
Subtest	*First unrotated factor (g)*	*Crystallized/verbal I*	*Perceptual organization II*	*Fluid III*
KAIT				
Crystallized				
Definitions	82	**47**		**45**
Auditory comprehension	69	**51**		37
Double meanings	79	**47**		**48**
Famous faces	65	**78**		
Fluid				
Rebus learning	69			**64**
Logical steps	66			**60**
Mystery codes	78			**88**
Memory for block designs	74		**41**	**51**
WISC-R				
Verbal				
Information	86	**62**		
Vocabulary	84	**75**		
Arithmetic	73			**59**
Comprehension	67	**46**		
Similarities	79	**50**	32	
Performance				
Picture completion	67		**43**	
Picture arrangement	53		36	
Block design	78		**79**	
Object assembly	73		**80**	
Coding	55			32

Source: Kaufman and Kaufman (1993).
Note: Decimal points are omitted. Rotated loadings <0.25 are omitted; those $\geqslant 0.4$ are in bold print. Correlations among factors are as follows: Factors I and II (0.59); Factors I and III (0.69); Factors II and III (0.65).

Enterprising, and Conventional. The findings indicated that there were two General Occupational Themes that produced substantial mean differences between IQ levels. Individuals with higher IQ were more Investigative and more Artistic (Artistic mean score = 49) than those with average IQ (Artistic mean score = 45) or low IQs (Artistic mean score = 42). The authors explain that in light of the Investigative person's interest in science and in solving abstract problems, the relationship of the Investigative theme to IQ level makes sense. The relationship between the Artistic theme and IQ was not hypothesized by the researchers, but was nonetheless significant, even with the effect of Investigative partialed out.

An examination of the KAIT with the Myers–Briggs Type Indicator has been conducted to understand more clearly the commonly accepted relationship between personality style and cognition (Kaufman, McLean, & Lincoln, 1996). The researchers had hypothesized that subjects who favored Intuition and Thinking on the Myers–Briggs would be more intelligent and would also favor fluid over-crystallized intelligence, compared to those subjects who favored Sensing and Feeling on the Myers–Briggs. Just as hypothesized, individuals classified as Intuitive earned higher KAIT Composite IQs than those classified as Sensing. However, the Fluid IQ was not found to be favored over the Crystallized IQ, as had been predicted (Kaufman et al., 1996). Thus, a modest association is present between personality dimensions and intellectual ability (as evidenced on the Myers–Briggs and KAIT).

4.08.2.1.8 Overview

The KAIT represents a reconceptualization of the measurement of intelligence that is more consistent with current theories of intellectual development (Brown, 1994). The fluid-crystallized dichotomy, the theory underlying the KAIT, is based on the original Horn–Cattell theory of intelligence, thus offering a firm and well researched theoretical framework (Flanagan, Alfonso, & Flanagan, 1994). The fluid-crystallized dichotomy enhances the richness of the clinical interpretations that can be drawn from this instrument (Brown, 1994). The test materials are well constructed and attractive,

Table 10 Exploratory joint principal-factor analysis of the KAIT and the WAIS-R (N = 338).

		Oblimin factor pattern		
Subtest	*First unrotated factor (g)*	*Crystallized/verbal I*	*Perceptual organization II*	*Fluid III*
KAIT				
Crystallized				
Definitions	78	**65**		34
Auditory comprehension	73	**69**		
Double meanings	79	**64**		**47**
Famous faces	68	**69**		
Fluid				
Rebus learning	65			**60**
Logical steps	66			**57**
Mystery codes	64		32	**56**
Memory for block designs	54			**62**
WAIS-R				
Verbal				
Information	79	**89**		
Vocabulary	81	**91**		
Arithmetic	77	**47**		
Comprehension	79	**78**		
Similarities	77	**59**	32	
Digit Span	60	**25**		29
Performance				
Picture completion	66		49	
Picture arrangement	64	28	57	
Block design	70		76	
Object assembly	64		80	
Coding	50		60	

Source: Kaufman and Kaufman (1993)
Note: Decimal points are omitted. Rotated loadings <0.25 are omitted; those ⩾0.4 are in bold print. Correlations among factors are as follows: Factors I and II (0.53); Factors II and III (0.53).

and the manual is well organized and helpful (Dumont & Hagberg, 1994; Flanagan et al., 1994). Furthermore, the test materials are easy to use and stimulating to examinees (Flanagan et al.).

"The KAIT has been standardized by state-of-the-art measurement techniques" (Brown, 1994). The psychometric properties of the KAIT regarding standardization and reliability are excellent and the construct validity evidence that is reported in the manual provides a good foundation for its theoretical underpinnings (Flanagan et al., 1994).

The theoretical assumption that formal operations is reached by early adolescence limits that application of the KAIT with certain adolescent and adult populations (Brown, 1994). If an individual has not achieved formal operations, many of the subtests will be too difficult for them and perhaps frustrating and overwhelming. Examiners should be aware of this limitation when working with such individuals in order to maintain rapport. The KAIT can be a useful assessment tool when working with high-functioning, intelligent individuals; however, it can be difficult to use with borderline individuals and some elderly clients. Elderly clients' scores on some of the subtests may be negatively impacted by poor reading, poor hearing, and poor memory (Dumont & Hagberg, 1994).

Flanagan et al. (1994) report that the inclusion of only three subtests per scale may limit or interfere with the calculation of IQs if a subtest is spoiled. The usefulness of the Expanded Battery and Mental Status subtest of clinical populations is questionable given the reliability and validity data presented in the manual, suggesting that interpretations be made with caution (Flanagan et al.).

Although there clearly are some limitations in the use of the KAIT with some populations, overall, the test appears to be well thought out and validated (Dumont & Hagberg, 1994). The KAIT represents an advancement in the field of intellectual assessment with its ability to measure fluid and crystallized intelligence from a theoretical perspective and, at the same time, maintain a solid psychometric quality (Flanagan et al., 1994).

4.08.2.1.9 The Stanford–Binet: Fourth edition

(i) Theory

Like its predecessors, the fourth edition is based on the principal of a general ability factor, *g*, rather than on a connection of separate functions. The fourth edition has maintained, albeit to a much lesser degree, its adaptive testing format. No examinee takes all the items on the scale, nor do all examinees of the same chronological age respond to the same tasks. Like its predecessors, the scale provides a continuous appraisal of cognitive development from ages two through young adult.

One of the criticisms in the previous versions is that they tended to underestimate the intelligence of examinees whose strongest abilities did not lie in verbal skills (or overestimate the intelligences of those whose verbal skills excelled). Therefore, consideration when developing the Binet IV was to give equal credence to several areas of cognitive functioning. The authors set out to appraise verbal reasoning, quantitative reasoning, abstract/visual reasoning, and short-term memory (in addition to a composite score representing *g*).

This model is based on a three-level hierarchical model of the structure of cognitive abilities. A general reasoning factor is at the top level (*g*). The next level consists of three broad factors: crystallized abilities, fluid analytic abilities, and short-term memory. The Horn–Cattell theory forms a foundation for the test, with measures of Gc being Verbal and Quantitative, and Abstract-Visual being a Gf scale. The third level consists of more specific factors, similar to some of Thurstone's eight primary mental factors: verbal reasoning, quantitative reasoning, and also abstract/visual reasoning. The selection of these four areas of cognitive abilities came from the authors' research and clinical experience of the kinds of cognitive abilities that correlate with school progress.

The Binet IV contains previous tasks, combining old with new items, and some completely new tasks. In general, test items were accepted if (i) they proved to be acceptable measurements of the construct, (ii) they could be reliably administered and scored, (iii) they were relatively free of ethnic and/or gender bias, and (iv) they functioned adequately over a wide range of age groups.

(ii) Standardization and properties of the scale

Standardization procedures followed 1980 US Census Data. There appears to be an accurate sample representation from geographic region, size of community, race/ethnic group, and gender. The standardization falls short, however, in terms of age, parental occupation, and parental education. The total sample size was large (5013), with age representation extending from two years to 23 years, 11 months. The concentration of the sample is on children aged 4–9 years old (41%). Not only were adults 24 years and older not represented, but also representation beyond age 17 years, 11 months was negligible (4%).

In order to assess characteistics of SES, information regarding parental occupation and parental education was obtained. A review of Table 11 demonstrates that children whose parents came from managerial/professional occupations and or who were college graduates and beyond were grossly over-represented in the sample. In other words, the norms are based on a large percentage of individuals from upper-socioeconomic classes. In order to adjust for this discrepancy, a weighting procedure was applied, which makes the norming sample suspect. Unquestionably, SES has been shown time and again to be the single most important stratification variable regarding its relationship to IQ (Kaufman, 1990, Chapter 6; Kaufman & Doppelt, 1976).

Internal consistency estimates for the Stanford–Binet IV Composite Scale are excellent, ranging from 0.95 to 0.99 across the age groups (median = 0.97) (Sattler, 1988). The internal reliabilities are also high for the Verbal Reasoning, Abstract/Visual Reasoning, Quantitative Reasoning, and Short-term Memory Area scores (typically in the upper 0.80s–0.90s). Subtest reliabilities are also good, with the exeption of Memory for Objects which had a median of 0.73 (Thorndike et al., 1986b). Test–retest reliability estimates are also good for preschool (Composite coefficient = 0.91) and elementary school aged (Composite coefficeint = 0.90) samples (Thorndike, Hagen, & Sattler, 1986a). From an internal reliability perspective, this measure is generally good.

Construct validity for *g* and for the four factors was studied using a variant of confirmatory factor analysis. The subtests had impressive high to substantial loadings on *g* (0.51–0.79). Unfortunately, the four factors were given weak support by the confirmatory procedure. Additionally, exploratory factor analysis gave even less justification for the four Binet Scales; only one or two factors were identified by Reynolds, Kamphaus, and Rosenthal (1988) for 16 of the 17 age groups studied. Clearly, the factor analytic structure does not conform to the theoretical framework used to construct the test. Therefore, once again the composite score is left as the only clearly valid representation of a child's cognitive abilities.

Table 11 Representation of the Stanford–Binet fourth edition.

	Sample percent	*US population percent*
By parental occupation		
Managerial/professional	45.9	21.8
Technical sales	26.2	29.7
Service occupations	9.7	13.1
Farming/forestry	3.2	2.9
Precision production	6.7	13.0
Operators, fabricators, other	8.3	9.5
Total	100.0	100.0
By parental education		
College graduate or beyond	43.7	19.0
1–3 years of college	18.2	15.3
High school graduate	27.5	36.5
Less than high school	10.6	29.2
Total	100.0	100.0

Correlational studies, using nonexceptional children, between the Binet IV and the Stanford–Binet (Form L-M), WISC-R, WAIS-R, WPPSI, K-ABC have ranged from 0.80 to 0.91 (comparing full-scale composites). Correlational studies using exceptional children (gifted, learning impaired, mentally retarded) produced generally lower correlations, probably because of restricted variability in the test scores. For example, for gifted students the mean composite score on the Binet IV correlated 0.69 with the WISC-R Full Scale IQ. These data and data from similar validity investigations are presented more extensively in the *Technical manual* for the Binet IV (Thorndike et al., 1986a). Despite the presentation of ample evidence of concurrent validity, the substantial problems with construct validity, the data collection method, and other difficulties with the Binet IV have led at least one reviewer to recommend that the battery be laid to rest (Reynolds, 1987): "To the S-B IV, *Requiescat in pace*" (p.141).

(iii) Overview

The Binet IV was developed in an attempt to increase the popularity of the test as well as address some of the negative reviews that had plagued the previous edition. The test authors attempted to make the fourth edition significantly different from the previous L-M edition; however, it appears that this goal has achieved only limited success. Canter (1990) describes the "rebirth" of the Binet as giving way to "confusion and even dismay as the primary consumers of intelligence tests learned that the new edition offered a more complicated route to the same destination." Another reviewer describes the Binet IV as "in most respects, a completely new version of a very old test" (Spruill, 1987). This author also questioned whether or not the weighting procedure that was used to correct for sample bias was adequate (Spruill, 1987) that was not outweighed by the large size of the standardization sample. Finally, it is not clear why a test described as for individuals aged two to adult does not include persons over the age of 23 in the standardization sample.

Although there appears to be a number of difficulties with the Binet IV, the test is still used and it is not without its strengths. The administration of some of the subtests allow the examiner flexibility, and young children seem to find the items challenging and fun. The scale has excellent internal reliability and provides a flexible administration format. Despite its shortcomings, Binet IV continues to be a very good assessment of cognitive skills related to academic progress (Spruill, 1987). It also includes several excellent, well-constructed tasks that offer valuable information when they are administered in addition to the Wechsler scales (Kaufman, 1990, 1994b).

4.08.2.1.10 Woodcock–Johnson Psycho-Educational Battery-Revised: tests of cognitive ability (WJ-R)

The WJ-R is one of the most comprehensive test batteries available for the clinical assessment of children and adolescents (Kamphaus, 1993). The WJ-R is a battery of tests for individuals from age 2 to 90 + , and is composed of two sections, Cognitive and Achievement. The focus of this discussion is the Cognitive portion of the WJ-R battery.

(i) Theory

The WJ-R Cognitive battery is based on Horn's (1985, 1989) expansion of the Fluid/Crystallized model of intelligence (Kamphaus, 1993; Kaufman, 1990). The standard and supplemental subtests of the WJ-R are aligned with eight of the cognitive abilities isolated by Horn (1985, 1989) (Kamphaus, 1993; Kaufman 1990). The cognitive battery measures seven Horn abilities: Long-term Retrieval, Short-term Memory, Processing Speed, Auditory Processing, Visual Processing, Comprehension-knowledge, and Fluid Reasoning. An eighth ability, Quantitative Ability, is measured by several Achievement subtests on the WJ-R.

The four subtests that measure Long-term Retrieval (Memory for Names, Visual–Auditory Learning, Delayed Recall/Memory for Names, Delayed Recall/Visual–Auditory Learning), require the subject to retrieve information stored minutes or a couple of days earlier. In contrast, the subtests that measure Short-term Memory (Memory for Sentences, Memory for Words, Numbers Reversed) require the subject to store information and retrieve it immediately or within a few seconds. The two Processing Speed subtests (Visual Matching, Cross Out) assess the subject's ability to work quickly, particularly under pressure to maintain focused attention.

Within the Auditory Processing domain, three subtests (Incomplete Words, Sound Blending, Sound Patterns) assess the subject's ability to perceive fluently patterns among auditory stimuli. The three Visual Processing subtests (Visual Closure, Picture Recognition, Spatial Relations) assess the subject's ability to manipulate fluently stimuli that are within the visual domain.

Picture Vocabulary, Oral Vocabulary, Listening Comprehension, and Verbal Analogies are the four subtests that are linked to the Comprehension-knowledge factor, also known as crystallized intelligence within Horn's theoretical model. These subtests require the subject to demonstrate the breadth and depth of their knowledge of a culture. Analysis-synthesis, Concept Formation, Spatial Relations, and Verbal Analogies (which also loads on the Comprehension-knowledge factor) assess the subject's Fluid Reasoning, or "new" problem-solving ability. Finally, from the Achievement portion of the WJ-R, both the Calculation and Applied Problems subtests assess the individual's Quantitative Ability.

The cognitive battery consists of 21 subtests, seven of which comprise the standard battery; the remaining 14 are part of the supplemental battery (one per ability as described by Horn). There are two composite scores, Broad Cognitive Ability and Early Development (for preschoolers), which are both comparable to an overall IQ. The individual subtest scores as well as the composite scores have a mean of 100 and a standard deviation of 15.

Computer software is available for scoring the WJ-R and is essential if all of the information is to be obtained that the WJ-R is capable of providing. The WJ-R provides the examiner with percentile ranks, grade-based scores, age-based scores, and the Relative Mastery Index (RMI). The RMI is a unique kind of ratio, with the second part of the ratio set at a value of 90. The denominator of the ratio means that children in the norm sample can perform the intellectual task with 90% accuracy. The numerator of the ratio refers to that child or adolescent's proficiency on that subtest (Kamphaus, 1993). For example, if a child obtains an RMI of 60/90, it would mean that the child's proficiency on the subtest is at a 60% level whereas the typical child of that age (or grade) mastered the material at a 90% level of accuracy.

The entire battery is quite lengthy and therefore can be timely to administer. The seven-subtest Standard Battery takes approximately 40 minutes to administer; however, all the clinician will obtain from it is, essentially, a measure of *g*. In order to obtain all of the information that the WJ-R is capable of providing, a clinician should administer most of the subtests in both the Cognitive and Achievement batteries. Administration of a thorough cognitive and achievement assessment using the WJ-R would take approximately 3.5–5 hours depending on the subject's age, abilities, and speed. However, individual subtests may be administered to test specific hypothesis without administering the entire battery. The WJ-R tests also provides measures of differential scholastic aptitudes including reading, mathematics, written language, and knowledge. An aptitude-achievement comparison may be made if the WJ-R Tests of Achievement are given in addition. Such a discrepancy reflects the amount of disparity between certain intellectual capabilities of an individual and their actual academic performance.

Evidence has been presented that supports the use of the WJ-R, standard cognitive and achievement tests in the identification and classification of school-aged children as gifted, learning-disabled, and mentally retarded (Evans, Carlsen, & McGrew, 1993). Significant group differences were found on mean scores of all WJ-R standard cognitive and achievement clusters. Together, the WJ-R cognitive and achievement demonstrated the ability to predict group membership (gifted, L.D., or M.R.). This was shown in an overall classification agreement of 93.5% for ages 8–10 and 84.3% for ages 16–18. These levels of classification agreement support the use of the WJ-R batteries in the identification of exceptional students (Evans et al., 1993). It should be noted, however, that clinicians must supplement the process of assessment and diagnosis by including other factors beyond statistical classification, such as social-emotional considerations, medical conditions, vision and hearing measures, and other environmental considerations, to make the determination of classification of the aforementioned groups.

(ii) Standardization and properties of the scale

The WJ-R was normed on a reasonably representative sample of 6359 individuals selected to provide a cross-section of the US population aged 2–90+ (Woodcock & Mather, 1989). The sample included 705 preschool children, 3245 students in grades K-12, 916 college/university students, and 1493 individuals aged 14–90+ who were not enrolled in school. Stratification variables included gender, geographic region, community size, and race. However, Kaufman (1990) reports that, although representation on important background variables was adequate, it was necessary to use a weighting procedure to adjust the data that was collected so it would match US population statistics.

The internal consistency estimates for the standard and supplemental battery subtests are good, with median scores from ages 2 to 79 ranging from 0.69 to 0.93. The Broad Cognitive Ability composite score for the seven standard battery subtest yields a median internal consistency coefficient of 0.94 and the Broad Cognitive Ability Early Development Scale yields a coefficient of 0.96 at ages two and four (Kamphaus, 1993).

The validity of the WJ-R has been called into question when used with a learning disabilities population (Hoy et al., 1993). This is because the mean broad cognitive ability scores from the WJ-R Tests of Cognitive Ability have been found to be one standard deviation lower than mean full scale scores on the WISC-R. However, in an adult sample, significant and consistently high correlations were found between the WAIS-R and the WJ scores (Siehen, 1985). Hoy et al. studied 27 male and 20 university students with a previously diagnosed learning disability, as well as 47 learning disabled individuals from a rehabilitation clinic. The results indicated a low positive correlation between the WJ-R broad cognitive score and the WAIS-R Verbal Scale scores ($r = 0.44$) for the rehabilitation subjects, but a more moderate positive correlation was found for the university subjects ($r = 0.73$). The WAIS-R Full Scale IQ also correlated moderately with the Broad Cognitive in university subjects ($r = 0.72$), and had a lower correlation with rehabilitation subjects ($r = 0.58$). The correlation between WJ-R Broad Cognitive Scores and the Performance Scale IQ was low for both the university and rehabilitation subjects ($r = 0.40$). The low correlation between the PIQ and Broad Cognitive suggests that the two instruments are providing different information, and are therefore both potentially useful.

(iii) Overview

The WJ-R Cognitive battery was developed based on Horn's expansion of the Cattell–Horn Fluid–Crystallized model of intelligence. This theoretical rationale allows for further empirical analysis of both the WJ-R and the theory (Webster, 1994). The standardization of the battery appears to be sound and the various age groups are represented adequately.

The Cognitive battery is quite thorough and, when administered in its entirety, can provide the examiner with a wealth of information about an individual's intellectual functioning and abilities. The test materials and manuals are easy to use and well designed. The administration is fairly simple; however, scoring the test, especially when the Achievement battery is administered as well, can be quite a lengthy and, initially, a difficult process. The scoring can be done by hand but is done more efficiently with the computer scoring program. The computer scoring program is easy to use and provides the examiner with the individual's raw scores, standard scores, percentile ranks, and age and grade equivalents for each subtest (Webster, 1994).

The WJ-R Cognitive battery is a well standardized test developed on a theory of intelligence. However, the test is not without shortcomings. Webster (1994) raises issues with the specific psychometric procedures used in developing test items. Data are lacking that show the efficacy of the WJ-R to predict, from a time based perspective, actual functional levels of academic achievement and to identify children at risk for failure early in the educational process (Webster, 1994). Kaufman (1990) points to another shortcoming with the small number of tests that comprise each scale. The Standard Scale measures each of the seven scales with one subtest apiece.

The *Woodcock–Johnson Psycho-Educational Battery: Revised examiner's manual* reports that

"Items included in the various tests were selected using item validity studies as well as expert opinion" (Woodcock & Mather, 1989, p. 7). Kamphaus (1993) states that the manual should have included more information on the results of the experts' judgments or some information on the methods and results of the studies that were used to assess validity.

It is clear that the WJ-R Cognitive battery is quite comprehensive, providing the clinician with a wealth of information. The standardization sample is large, the factor loadings reveal generally strong factor analytic support for the construct validity for the battery for adolescents and adults, and the reliability coefficients are excellent (Kaufman, 1990).

4.08.2.1.11 Detroit Tests of Learning Aptitude (DTLA-3)

DTLA-3 was developed by Hammill (1991) and was designed to measure different, but interrelated, mental abilities for individuals aged six through 17 years, 11 months. It is a battery of 11 subtests and yields 16 composites that measure both general intelligence and discrete ability areas. Hammill and Bryant (1991) report that the DTLA-3 was influenced greatly by Spearman's two-factor theory (1927). This theory of "aptitude" consisted of a general factor *g* that is present in all intellectual pursuits, and specific factors that vary from task to task (McGhee, 1993).

The 11 subtests are used to form the 16 composite scores. The subtests are grouped into different combinations according to various hypothetical constructs that exist in theories of intelligence and information processing. In general, the composite scores estimate general mental ability; however, they all do so in a somewhat different manner. The General Mental Ability Composite is formed by combining the standard scores of all 11 subtests, and, thus, has been referred to as the best estimate of *g*. The Optimal Level Composite is composed of the four largest standard scores that the individual earns. This individualized score is often referred to as the best estimate of a person's overall "potential." The Domain Composites may be divided into three areas: Linguistic, Attentional, and Motoric. Furthermore, there is a Verbal and Nonverbal Composite in the Linguistic domain, an Attention-enhanced and Attention-reduced Composite in the Attentional Domain, and a Motor Enhanced and a Motor-reduced composite in the Motoric Domain. Finally, there are the Theoretical Composites of the DTLA-3 on which the battery's subtests are constructed. The major theories that the subtests were developed from include Horn and Cattell's (1966) fluid and crystallized intelligences, Das' (1973) simultaneous and successive processes, Jensen's (1980) associative and cognitive levels, and Wechsler's (1974, 1981, 1989) verbal and performance scales.

The DTLA-3 yields five types of scores: raw scores, subtest standard scores, composite quotients, percentiles, and age equivalents. Standard scores for the individual subtests have a mean of 10 and a standard deviation of three and the Composite Quotients have a mean of 100 and a standard deviation of 15.

The individual subtest reliabilities range from 0.77 to 0.94, with a median of 0.87, and the averaged alphas for the composites ranged from 0.89 to 0.96, with a median of 0.94. To assess the DTLA-3's stability over time, the test–retest method was used with a sample of 34 children residing in Austin, Texas. The children, aged six through 16, were tested twice, with a two-week period between testings (Hammill, 1991). The results of this test–retest analysis indicate that individual subtest reliabilities range from 0.75 to 0.96, with a median of 0.84, and Composite reliabilities range from 0.81 to 0.96, with a median of 0.90.

(i) Overview

The DTLA-3 was designed to measure both general intelligence and discrete abilities for children aged six through 17 years, 11 months. The DTLA-3 is not grounded in one specific theory but rather can be linked to a number of different theorists and their views on intelligence and achievement. This "eclectic" theorizing has resulted in the DTLA-3's numerous subtests, composites and various combinations of the two that yield potentially important information about an individual's abilities.

Reliability and validity studies are encouraging but are based on specific and limited samples (VanLeirsburg, 1994). Additional research in this area would be beneficial. Furthermore, test–retest reliability data were collapsed across age levels, which makes it impossible to determine the stability of scores of the various age levels (Schmidt, 1994). The standardization sample was representative of the US population but more information on socioeconomic level is needed (Schmidt, 1994). Also, there is no normative data reported for subjects with handicapping conditions and sample stratification for age was not equalized (VanLeirsburg, 1994).

The testing manual suggests that individual testing time may vary but that on average it takes 50 minutes to two hours to administer. Scoring and interpretation of the results is easy, yet it can be quite time consuming without the aid of the computer program (VanLeirsburg, 1994). Despite apparent shortcomings, the DTLA-3 may be useful for eligibility or placement purposes as well as a useful research tool (Schmidt, 1994).

4.08.2.1.12 Differential Abilities Scales (DAS)

The DAS was developed by Elliott (1990) and is an individually administered battery of 17 cognitive and achievement tests for use with individuals aged 2.5 through 17 years. The DAS Cognitive Battery has a preschool level and a school-age level. The school-age level includes reading, mathematics, and spelling achievement tests that are referred to as "screeners." The same sample of subjects was used to develop the norms for the Cognitive and Achievement Batteries; therefore, intra- and intercomparisons of the two domains are possible.

The DAS is not based on a specific theory of intelligence. Instead, the test's structure is based on tradition and statistical analysis. Nonetheless, the test is not theory free, and, in fact, is based in part on *g* and the view of intelligence as hierarchical in nature (McGhee, 1993). Elliott (1990) described his approach to the development of the DAS as "eclectic" and cited researchers such as Cattell, Horn, Das, Jensen, Thurstone, Vernon, and Spearman. Indeed, there are some clear-cut relationships between several DAS scales and theoretical constructs. For example, Horn's (1985, 1989) concepts of fluid and crystallized intelligence are measured quite well by the Nonverbal Reasoning and Verbal Ability scales, respectively. Elliott endorses Thurstone's ideas that the emphasis on intellectual assessment should be on the assessment and interpretation of distinct abilities (Kamphaus, 1993). He also stresses that in the assessment of children with learning and developmental disabilities, clinicians need more fine detail than is provided by a global IQ score (Elliott, 1997). Therefore, subtests were constructed to emphasize their unique variance which should translate into unique abilities. Although it was expected that meaningful composites would be derived from the subtests, the primary focus in test development was at the subtest level. One of the main distinctions between the DAS and other batteries is its emphasis on the subtest level.

Elliott (1997) noted how in psychology there has long been a link between cognitive abilities and neurological structure. The DAS uses the link between the factor structure of abilities and the neurological evidence of the nature of the structures. For example, the DAS has two major ability clusters which reflect two major systems for receiving, perceiving, remembering, and processing information in the visual and auditory modalities. The systems are represented by Verbal and Visualization/Spatial factors. There is strong neuropsychological evidence for the existence of these systems (Elliott, 1997), which tend to be specific in the left and right cerebral hemispheres, respectively. The DAS specifically measures each of these factors by the Verbal cluster and the Spatial Cluster.

Normally, the auditory and visual systems do not operate completely independently. There is interaction of the systems. The integrative system is represented factorally by the fluid reasoning factor in the Cattell–Horn theory. Analysis of both verbal and visual information is usually required in measures of fluid reasoning. The neuropsychological function is an integrative system of the frontal lobe, which is critical to complex mental functioning (Luria, 1973). The DAS measures fluid ability by the Nonverbal Reasoning cluster, which requires integrated analysis and transformation of both visual and verbal information.

Elliott (1997) notes that there is much evidence from cognitive psychology that indicates that verbal and visual short-term memory systems are quite distinct. However, some cognitive tests (i.e., Stanford–Binet IV) represent memory with a single factor. The DAS, on the other hand, represents visual and auditory short-term memory with distinct measures, rather than with one unitary task. The DAS also provides a measure of intermediate-term memory (in the Horn–Catell model), which is usually measured by tasks that have both verbal and visual components. The DAS has a measure in which pictures are presented, but they have to be recalled verbally (Recall of Objects). This visual–verbal short-term memory measures another distinct information processing system (Elliott, 1997).

The cognitive portion of the DAS consists of "core" and "diagnostic" subtests designed to assess intelligence at the preschool level and the school-age level. The core subtests measure complex processing and conceptual ability, which is strongly *g*-related. The diagnostic subtests measure less cognitively complex functions, such as short-term memory and processing speed, thereby having less of a *g*-saturation (Elliott, 1997). The achievement portion measures skills in the areas of word reading, spelling, and basic number skills. The core subtests are averaged to obtain the General Conceptual Ability (GCA) score and, depending on the age of the individual, additional composite scores are calculated which are referred to as Cluster scores.

The individual Cognitive subtests have a mean of 50 and a standard deviation of 10. The GCA scores, Cluster scores, and Achievement scores, have a mean of 100 and a standard deviation of 15. Percentile ranks, age equivalents, and score comparisons are also available in the examiner's manual. Score comparisons provide a profile analysis and allow the examiner to ascertain information regarding aptitude-achievement discrepancies.

Interpretation of the DAS subtests and composites is facilitated by the framework provided in the *Handbook* (Elliott, 1990). Elliott (1990) notes one positive aspect of interpreting the DAS is that the "design of scoring procedures on the Record Form enables statistically significant high and low scores to be identified immediately" (p. 37). Significant discrepancies between subtests, cluster scores, and ability and achievement can be obtained immediately. Like Kaufman's (1994b) approach to interpreting the WISC-III, an ipsative approach is used to examine differences between subtests, requiring the examiner to compare the child's mean score on core subtests to his or her individual subtest scores. Discrepancy between ability and achievement is analyzed by examining the GCA (or Special Nonverbal Composite) and each of the achievement tests.

(i) Standardization and properties of the scale

Elliott (1997) noted that exceptionally careful and effective standardization and data-analytic procedures were used in the development of the DAS. The DAS was standardized on 3475 children tested between 1987 and 1989. The normative sample includes 200 cases for each age level between the ages of five and 17. The younger part of the sample consisted of 350 children between the ages of 2.5 year, and four years, 11 months. Exceptional children were also included in the standardization sample. Gender, race, geographic region, community size, and enrollment (for ages 2–5 through 5–11) in an educational program were controlled. SES was estimated using the average education level of the parent or parents living with the child (Kamphaus, 1993).

Over and above the requirements of the norm sample, an additional 600 cases of Black and Hispanic children were collected. The reason that this oversample was collected was to perform statistical analysis for item bias and prediction bias. The test developers wanted to ensure that the rules for scoring would be sensitive to minority children's responses (Elliott, 1997). Only a small number of items were deleted due to item bias, and there was "no evidence" that the DAS is biased against either Blacks or Hispanics, compared with Whites (Elliott, 1990).

The DAS has a median reliability estimate of 0.95 for the GCA. Internal consistency reliability estimates for the cluster scores range from 0.83 for Nonverbal Reasoning at age five to 0.94 for Spatial at several ages (Kamphaus, 1993). The test–retest reliability coefficients for the preschool composite scores are 0.84 for Verbal Ability and 0.79 for Nonverbal Ability. The individual subtests' reliabilities vary with an average coefficient of 0.78.

Correlational research has shown good evidence of concurrent validity for the DAS (Kamphaus, 1993). With a sample of 27 children aged 7–14, the WISC-III Full Scale IQ correlated very highly with the DAS GCA score (0.92), and the WISC-III Verbal IQ score correlated highly with the DAS Verbal Ability score (0.87). The WISC-III Performance IQ correlated 0.78 with Nonverbal Reasoning and 0.82 with Spatial Ability. Additionally, the DAS Speed of Information Processing subtest score correlated 0.67 with the WISC-III Processing Speed Index score. The Binet IV Composite IQ correlated 0.88 with the DAS GCA for nine- and 10-year-olds and 0.85 with the DAS GCA for a sample of gifted children. The K-ABC Mental Processing Composite correlated 0.75 with the DAS GCA for 5–7-year-olds (Kamphaus, 1993).

Elliott (1997) presents three validity studies not published in the DAS manual. One of these studies included a confirmatory factor analysis of the DAS by Keith (1990), which concluded that "the constructs measured by the DAS are remarkably consistent across overlapping age levels of the test" (Elliott, 1997, p. 20). Elliott also discusses a joint factor analysis of the DAS and WISC-R. In a reanalysis of data, Elliott (in press) reported five factors emerging: crystallized intelligence (including DAS verbal and WISC-R verbal subtests), spatial or broad visualization (including DAS spatial and four of the five major WISC-R Performance subtests), nonverbal reasoning or fluid intelligence (defined only by DAS Nonverbal Reasoning subtests), auditory short-term memory, and speed of processing.

(ii) Overview

In general, the professional reviews of the DAS seem to be quite positive. Sandoval (1992) reports that the DAS is one of the least biased tests available. The test appears to be a relatively culture fair measure; however, its use with linguistically different children needs to be explored further (Sandoval, 1992). In examining different cultural group's performance on the

DAS, the group differences found typically to be present on traditional IQ tests are also found on the DAS. For example, African-American and Hispanic children score between half and two-thirds of a standard deviation below White children, and Asian children score above White children on all but verbal areas of the test. Caution is necessary when assessing Hispanic children because the DAS overpredicts achievement for this group based upon group achievement results (Bain, 1991). The author of the DAS suggests that children who are not proficient in English be given the nonverbal tests on the Special Nonverbal scale in the primary language. However, this can be problematic as the test developers did not provide directions in other common languages, such as Spanish. In addition, the utility of the English norms for assessing a child who is administered the test in Spanish or another non-English language has not been explored.

The DAS manual has recommendations for administering the test to deaf or limited-English-proficient children, however, these recommendations are lacking in a couple of areas. Braden (1992) notes, "The recommendations that age equivalents be used to represent the performance of retarded person is common, but it is potentially misleading" (p. 93). Another problem with the recommendations is that no mention is made of the use of interpreters for hearing-impaired children and nonverbal children, which may have a detrimental effect on deaf children's test scores.

According to Braden (1992), the Technical Manual includes extensive research data which suggest that the DAS is a psychometric improvement over existing techniques for measuring intellectual abilities and for determining intracognitive and aptitude-achievement discrepancies. The GCA of the DAS is largely independent of tasks known to be difficult for learning disabled children and is able to assist in the identification of learning disabilities or processing deficits.

The DAS can be a useful tool in assessing intelligence and achievement in both children and adolescents. However, there are a few characteristics of the DAS which do not promote ease of administration, especially for novices (Braden, 1992). These difficulties include having to apply two rules for subtest discontinuation, and having to convert raw scores to ability scores prior to obtaining subtest scaled scores. The DAS examiner's manual provides interpretive information and a framework for interpretation for the composite scores and subtests. The level and/or depth of information that the interpretative portion of the manual provides is quite thorough and is easy to use, making interpretation of the profiles, and individual and composite scores much easier.

4.08.2.1.13 Cognitive Assessment System (CAS)

(i) Theory

The Das–Naglieri Cognitive Assessment System (Naglieri & Das, 1996) was developed according to the Planning, Attention, Simultaneous, and Successive (PASS) theory of intelligence. The subtests are organized into four scales designed to provide an effective measure of the PASS cognitive processes. Planning subtests require the child to devise, select, and use efficient plans of action to solve the test problems, regulate the effectiveness of the plans, and self-correct when necessary. Attention tests require the child to selectively attend to a particular stimulus and inhibit attending to distracting stimuli. Simultaneous processing tests require the child to integrate stimuli into groups to form an interrelated whole and Successive processing tests require the child to integrate stimuli in their specific serial order or appreciate the linearity of stimuli with little opportunity for interrelating the parts.

The CAS yields Planning, Attention, Simultaneous, Successive, and Full Scale normalized standard scores (mean of 100 and standard deviation of 15). The Planning scale's subtests include Matching Numbers, Planned Codes, Planned Connections, and Planned Search; the Attention scale subtests include Number Detection, Receptive Attention, and Expressive Attention; the Simultaneous Scale subtests are Nonverbal Matrices, Verbal–Spatial Relations, and Figure Memory; and the Successive Scale subtests are Word Series, Sentence Repetition, Sentence Questions, and Successive Speech Rate. All subtests are set at a normative mean of 10 and SD of three.

(ii) Standardization and properties of the scale

The CAS was standardized on 2200 children ranging in age from five through 17 years. The sample was stratified by age, gender, race, ethnicity, geographic region, and parent education according to US. Census reports and closely matches the US population characteristics on the variables used. In addition to administration of the CAS, most of the standardization sample was also administered several achievement tests from the Woodcock–Johnson Tests of Achievement. (Woodcock & Johnson, 1989). This provided for both validity evidence and analysis of the relationships between PASS and achievement. No further data were available on the CAS when this chapter went to press.

4.08.3 INSTRUMENT INTEGRATION

It is to the advantage of clinical and school psychologists that there are so many instruments available to assess a child's, adolescent's, or adult's intellectual functioning. Often when one instrument is administered, such as the WISC-III or WAIS-R, and then analyzed, the examiner will find that questions and hypotheses are raised regarding the individual's functioning in specific areas. Creativity and a bit of detective work are required to uncover exactly where a person's true deficits and strengths lie in their cognitive abilities. Part of the detective work in this process involves the integration of information from various instruments to support or clarify hypotheses raised as initial results are examined. Thus, examiners ultimately have to be able to integrate data from multiple instruments. As suggested by Kaufman (1994) "crucial educational decisions are sometimes made on the basis of a psychological evaluation, and these decisions should be supported by ample evidence" (p. 326) so that initial hypotheses are verified. This section describes and discusses several cognitive tests in terms of their value when integrated with results from the Wechsler Scales.

4.08.3.1 K-ABC Integration with Wechsler Scales

The K-ABC measures some of the same abilities as the WISC-III and WPPSI-R, but also measures ability in ways that are different from the Wechsler scales, thereby contributing unique information about a child's cognitive functioning. The K-ABC Simultaneous Processing Scale is believed by some researchers to involve the same cognitive requirements as Wechsler's Performance Scale (Das, Naglieri, & Kirby, 1994) and by others to be a measure of Visual Processing (Gv) (Horn, 1991). However, two Simultaneous subtests (Matrix Analogies and Photo Series), involve more reasoning (and load on two of Woodcock's (1990) factors: Fluid Reasoning (Gf) and Visual Processing (Gv)) than Wechsler's Performance subtests: The Sequential Processing Scale of the K-ABC is an excellent addition to the Wechsler because it measures sequential processing more efficiently than any Wechsler subtest. That is, because the only Wechsler test that can be viewed as measuring sequential processing is Digit Span Forward (Das et al., 1994) but the Digit Span subtest score includes Backwards span which involves more than sequential processing (Schofield & Ashman, 1986), there is no efficient measure of sequential processing on the Wechsler. In addition, the K-ABC Achievement Scales are highly related to the Wechsler Verbal IQ and crystallized abilities (Kaufman & Kaufman, 1983; Naglieri & Jensen, 1987).

Given the above characteristics of the K-ABC, examiners may note that the entire Simultaneous Processing Scale serves as a good measure for children with motor and/or speed problems who earn low Wechsler Performance IQs, because it minimizes both of these variables. From the Horn view, the K-ABC offers good supplemental subtests to measure Gc including Faces and Places and Riddles, in the Achievement Scale. These tasks measure range of general knowledge by identifying visual stimuli (Faces and Places), and require the child to use verbal reasoning to demonstrate word knowledge (Riddles). This is unlike many tests, such as WISC-III Vocabulary, and similar Binet IV and DAS tasks, which measure word knowledge by requiring a child to retrieve word definitions from long-term storage. Like the WJ-R crystallized subtests, Riddles requires a one-word response; it is, therefore, a good Wechsler supplement to help discern whether a low V-IQ is due more to conceptual problems or expressive difficulties. The K-ABC offers some alternative modalities of receiving input and expression of response, to supplement Wechsler subtests that mainly use the auditory-vocal and visual-motor channels of communication. The K-ABC offers three subtests which call for use of the visual and vocal modalities (Magic Window, Faces and Places, and Gestalt Closure), and has one subtest that uses the auditory-motor channel (Word Order). In addition, the K-ABC taps the visual and semantic-motor channels with Reading Understanding, which requires a child to read a stimulus and do what it says (i.e., "Stand up").

4.08.3.2 Integration of KAIT with Wechsler Scales

The KAIT was developed from the Horn–Cattell theory and yields both a Crystallized IQ and Fluid IQ. The three subtests comprising the KAIT Fluid Scale are very good supplements to the Wechsler scales. As noted previously, there is controversy regarding how well the Wechsler scales measure fluid abilities; thus, it is wise to administer supplemental tests to tap an individual's fluid reasoning ability and learning ability. Assessment of planning ability, formal operational thought, and learning ability may be obtained through KAIT Fluid Subtests: Mystery Codes and Logical Steps. Problem-solving through verbal reasoning and verbal comprehension is required in Logical Steps, and Rebus Learning demands vocal responding; therefore, the KAIT Fluid Scale measures an ability that is quite different from Wechsler's P-IQ. If questions about an individual's planning speed arise from the primary battery administered, examiners may administer Mystery Codes to further assess planning speed.

To supplement Wechsler's Verbal Scale, the KAIT Crystallized subtests may be used. For assessing an individual's base of general factual knowledge, Famous Faces may be administered to supplement WISC-III or WAIS-R Information. Famous Faces uses pictorial stimuli integrated with verbal clues about famous people; whereas Information is a purely auditory-vocal task. Formal operational thought within the Crystallized domain can be assessed through Double Meanings. Double Meanings challenges examinees to unify apparently disparate semantic stimuli. KAIT subtests, Double Meanings and Definitions, can also provide follow-up to questionable performance on tasks of word knowledge and verbal concept formation, such as Wechsler's Vocabulary or Similarities. Auditory Comprehension can be used for questions regarding an individual's memory and comprehension ability. This subtest mimics a real-life situation, in requiring an individual to listen to a mock news broadcast and answer questions about it. The two delayed recall (TSR) KAIT subtests are also very good WISC-III supplements when hypotheses are raised regarding an individual's memory.

4.08.3.3 Integration of Binet IV with Wechsler Scales

As noted in the earlier discussion of the Binet IV, there is controversial and weak factor-analytic support for the four Binet IV area scores (Verbal, Abstract-Visual, Quantitative, and Short-term Memory). The relationship between Wechsler's Verbal and Performance IQs and the Binet IV Area scores is not clear-cut. In a correlational analysis with the WISC-R and Binet IV (Thorndike et al., 1986), the WISC-R and Binet IV Verbal scales were found to relate substantially to each other. However, Kaufman (1994b) notes that the Absurdities subtest probably lowered the relationship with the Verbal IQ and increased the correlation with the Performance IQ because it uses visual stimuli. In Woodcock's (1990) factor analysis, the Binet IV Quantitative subtests loaded on a separate Quantitative Factor (Gq), which also included Wechsler Arithmetic and WJ-R math achievement subtests. The Binet IV Abstract-Visual subtests, except Matrices, were on the Gv factor, along with most Wechsler Perceptual Organization subtests. Matrices, however, had a substantial loading on the Gf factor; it is, therefore, an excellent addition to Wechsler's Performance Scale.

The Binet IV can be integrated with Wechsler results, and can be especially helpful in assessing young children and mentally retarded individuals because of the extension of its norms down to age two. Response time is relatively unimportant on the Binet IV; therefore, it provides several subtests to further evaluate hypotheses regarding a low score on the WISC-III P-IQ or PO Index. If poor fluid intelligence is suspected, Pattern Analysis, Paper Folding and Cutting, Matrices, and Number Series can be administered. To further assess the comprehension knowledge ability measured by the WISC-III, without requiring verbal comprehension, Absurdities is especially good to administer because the stimulus is visual and minimally verbal. For assessing whether a child's fluid reasoning ability generalizes to number manipulation activities, the two Binet IV Quantitative subtests are useful. The tasks not included on the Wechsler scales such as Matrices, Equation Building, Number Series, and Verbal Relations, can be used to further explore the reasoning abilities of an individual.

4.08.3.4 Integration of WJ-R with Wechsler Scales

Wechsler's Verbal IQ primarily can be viewed as a measure of crystallized intelligence and short-term memory; Performance IQ as a blend of fluid reasoning, visual processing, and processing speed (Kaufman, 1994b). However, some researchers view Wechsler's Perceptual Organization to be a measure of visual processing (McGrew & Flanagan, 1996; Woodcock, 1990). Long-term retrieval is not specifically measured by the Wechsler scales nor is auditory processing. And, if Woodcock and others are correct, then fluid reasoning also is not measured very well by the Wechsler scales. Therefore, WJ-R Cognitive subtests provide excellent tasks for extending assessment from Wechsler subtests using the WJ-R tests which were developed to reflect Horn's pure factors. The WJ-R provides subtests which are controlled learning tasks, allowing the assessment of a person's learning ability. Conventional intelligence tests, including Wechsler's, do not typically measure this ability. The controlled learning subtests include the following: Memory for Names and Visual–Auditory Learning (both Long-term Retrieval tasks), and Analysis–Synthesis and Concept Formation (both Fluid Reasoning tasks). Whereas the Wechsler Performance subtests emphasize visual-motor coordination and speed of response, Analysis–Synthesis and Concept Formation involve no motor coordination at all, and speed of response is not a major variable in determining a person's performance level.

The WJ-R provides several subtests from which to choose, so when questions arise regarding Wechsler's Perceptual Organization construct (including visual processing and fluid abilities, as noted above), the WJ-R is an excellent tool to investigate these hypotheses. The different aspects of the information-processing model are measured by four WJ-R factors including Gv (input), Gf (integration), Glr (storage), and Gs (output). A high or low score on Wechsler's Performance scale should be explored further to determine what aspects of an individual's information-processing may have affected this asset or deficit. If an individual is suspected of having a deficit or strength in their nonverbal visual-spatial ability, requiring further testing for clarification, then WJ-R Spatial Relations is a useful tool to make that determination. The other Fluid Reasoning subtests, previously mentioned, have a heavy verbal component and do not assess visual-spatial skills, although they do use figural material. (One precaution to note is that cognitive tests in the WJ-R battery are heavily entrenched in the tradition of measuring intelligence through predominantly verbal means (Kaufman, 1990)).

For assessing strengths or weaknesses within the auditory-vocal channel, the following WJ-R factors may be used: Ga (input), Gc (integration), and Gsm (storage). These factors can be helpful in clarifying questions raised in the Verbal scale of the Wechsler test. On the WJ-R, two Gc tasks require one-word responses, which make them good when you do not know if a low Verbal score reflects poor concepts or poor expression. Auditory-perceptual tasks on the Ga scale assess whether a child can perceive words in isolation (through filling in the gaps or by blending sounds). However, for assessing a processing deficit of longer auditory input, additional subtests may be needed (such as the Cognitive Assessment System subtests Verbal-Spatial Relations, Sentence Repetition, and Sentence Questions).

Wechsler scales do not assess long-term memory over the period of a few minutes, although they do measure short-term memory with Digit Span and remote memory with Picture Completion and Information. The Long-term Retrieval subtests of the WJ-R provide a good assessment of the long-term memory function; therefore, these tasks complement the Wechsler subtests for supplementary analysis. In addition, the WJ-R tests provide a measure Auditory Processing and Visual Processing, which have strong perceptual components. These perceptual processes are not typically evaluated in most tests of intelligence, but need to be assessed in cases with possible neuropsychological difficulties.

4.08.3.5 DTLA-3 Integration with Wechsler Scales

The DYLA-3 has several theoretical underpinnings, including models such as fluid and crystallized intelligence, simultaneous and successive processes, and verbal and performance abilities. DTLA-3 subtests may be used to augment the Wechsler scales in several instances. To further assess perceptual organization ability, fluid ability, and simultaneous processing, Design Reproduction or Symbolic Relations may be administered. For hypotheses regarding similar fluid abilities, but tapping sequential processing, examiners may administer Design Sequences. DTLA-3 Design Reproduction is also a good supplement Wechsler Performance subtests if there is a question regarding a person's ability being hampered by response speed tests. This test does require visual-motor coordination but places minimal demands on speeded performance. Like the K-ABC, the DTLA-3 offers some alternative modalities of receiving input and expression of response to supplement Wechsler subtests that mainly use the auditory-vocal and visual-motor channels of communication. The DTLA-3 offers two subtests which call for use of the visual and vocal modalities (Story Construction and Picture Fragments), and has one subtest that uses the auditory-motor channel (Reversed Letters).

4.08.3.6 Integration of DAS with Wechsler Scales

The six Core subtests of the DAS create three separate scales for children, namely: Verbal, Spatial, and Nonverbal Reasoning. The WISC-III Verbal Comprehension subtests (specifically Vocabulary and Similarities) have been noted to be quite similar to the DAS Verbal Scale (Kaufman, 1994). The DAS Verbal, Spatial, and Nonverbal Reasoning scales have been shown to correspond to the Woodcock–Johnson Revised factors of Gc, Gv and Gf, respectively (McGhee, 1993).

The two subtests comprising the DAS Nonverbal Reasoning Scale provide an excellent addition to the WISC-III because they are quite different from WISC-III subtests. The Nonverbal Reasoning subtests (Matrices, and Sequential and Quantitative Reasoning) measure nonverbal reasoning without time limits but they do require visual-motor coordination, and minimize visualization. Thus, they can provide good measures of an individual's pure fluid ability. DAS subtests requiring visual-motor coordination, but placing minimal demands on speeded performance include Recall of Designs and Pattern Construction (when the latter test is administered via special procedures). Thus, these subtests can be useful in following up hypotheses generated from Wechsler's Performance subtests, which reward quick performance.

4.08.3.7 Integration of CAS with Wechsler Scales

Like the other tests included in this chapter, the CAS has some overlap with the WISC-III, but because its conceptualization is based on the PASS theory unique information about a child's cognitive functioning can be obtained. The CAS Planning and Attention Scales require processes that can not be effectively assessed by the Wechsler Scales (Das et al., 1994). In order to measure planning adequately, tests that evaluate the child's ability to decide how to solve problems, and determine their effectiveness are required. This means that the child must be given the opportunity to compete tasks in planful ways, unencumbered by rules imposed by the test. Additionally, items that are influenced by the child's plan rather than other factors (e.g., spatial or verbal skills) are needed. Tests of this type are not found on the WISC-III. Attention tests should demand the focus of cognitive activity and selective attending to particular information while avoiding distraction. Carefully constructed measures of attentional processes are not included on the Wechsler, yet this, as well as planning processes, are especially important when evaluating children, especially those with attention deficits and learning disabilities, for example. The measurement of Planning and Attention offer important cognitive functions that extend beyond the WISC-III and therefore offer additional information for diagnosis as well as intervention (Das et al., 1994).

The CAS, like the K-ABC, provides a measure of simultaneous processing that is similar to the demands of Wechsler's Performance Scale (Das et al., 1994) but there are important distinctions. The CAS offers a verbal test of simultaneous processing (Verbal-Spatial Relations), one that involves memory (Figure Memory), and one with complex demands (Nonverbal Matrices). The addition of the Verbal-Spatial Relations subtest is important because it integrates both nonverbal and verbal stimuli for the comprehension of logical grammatical sentences. Similarly, the Successive processing Scale of the CAS provides tests that demand immediate recall of information (Word Series), and also measures that demand comprehension of syntax (Sentence Repetition and Setence Questions) and where the involvement of immediate memory is markedly reduced (Successive Speech Rate).

The CAS offers a view of ability that reduces the influence of language and achievement, and therefore, provides additional information from the WISC-III or the WPPSI-R to evaluate the performance of children who are bilingual or whose educational history is problematic. Because the CAS does not have achievement or language based tests like the Wechsler (e.g., Arithmetic or Vocabulary) the reduction in the involvement of acquired knowledge provides an opportunity to evaluate children whose poor school history or language difference may have lowered their Wechsler scores. In such a situation the CAS scores can assist the psychologist in determining the extent to which low Wechsler scores may reflect language/achievement issues rather than low intellectual ability.

4.08.4 FUTURE DIRECTIONS

Intellectual assessment has changed a great deal in the twentieth century. It has moved from assessments based on language and speech patterns to sensory discrimination, with most early assessments being for the mentally deficient. Gradually, the assessment instruments developed into the precursors of the standardized instruments used in the 1990s, which measure more complex cognitive tasks for all levels of cognitive ability. The most commonly used tests, the Wechsler scales, were not developed out of theory, but were guided rather by clinical experience. In the progression of test development, the relative alternates to the Wechsler have been more theory-based. The direction in test development in the 1990s seems to be continuing to lead to theory being at the base of intelligence tests, instead of being just purely clinically driven. Psychometric theories and neurological theories are growing as the basis for new instruments. However, despite this proliferation of new theory-driven tests, there is a definite conservatism that holds on to the past, reluctant to let Wechsler tests be truly rivaled. Part of this hold on the past is research-based and part is clinically-based. Because the Wechsler tests have been reigning supreme for so long, there has been a mountain of research studies using the Wechsler scales. Thus, clinicians have a good empirical basis to form their understanding of what a specific Wechsler profile may be indicating. Clinically, psychologists are also quite comfortable and familiar with the Wechsler scales. A good clinician who has done many assessments may be familiar enough with every nook and cranny of the WAIS-R to barely need the manual to administer it. Thus, the field so far has changed relatively slowly. Computer based technology is likely going to ultimately shape the field of assessment by 2020. Computer scoring programs and computer assisted reports are already in use, and the future is likely to include a much greater progression of technologically advanced instruments for assessing intelligence.

4.08.5 SUMMARY

This chapter first introduces the assessment of intellectual ability through a brief overview of its development throughout history. Some important faces in the history of assessment include those such as Jean Esquirol, Edouard Seguin, Sir Francis Galton, James McKeen Cattell, Alfred Binet, Lewis Termin, and David Wechsler. The early pioneers in assessment focused mainly on tests for the mentally deficient, but more recently in history tests were developed for assessing all levels of intellectual functioning. The progression of test development to the standardized instruments we know today is reviewed in the beginning of this chapter.

The debate over intelligence testing is also discussed in this chapter. Three groups of critics are presented. One controversy is raised by those opposed to subtest interpretation advocated by Wechsler; another is those that insist that all the Wechsler scales measure is *g*, rendering the different scales meaningless; and a final controversy is raised by a group who complains that the instruments do not enhance remedial interventions. Kaufman's (1979, 1994b) intelligent testing approach is presented in response to the criticism. The aim of the basic principles of intelligent testing presented in this chapter is to encourage clinicians to approach the task of profile interpretation using their knowledge of research, theory, and clinical skills, rather than being overly dependent on specific scores.

With this important philosophy of intelligent testing in mind, 10 measures of child, adolescent, and adult intelligence are discussed. The Wechsler scales are generally used by examiners as the primary instrument in an assessment battery. However, there are multiple other excellent instruments available for assessing cognitive ability. The following instruments are discussed: WPPSI-R, WISC-III, WAIS-R, K-ABC, KAIT, Binet-IV, WJ-R Tests of Cognitive Ability, DTLA-3, DAS, and CAS. For each, the theory or theories underlying the instrument is presented, followed by the standardization and properties of the scale, including research using the scale, and each is completed with an overview of the instrument.

One of the main principles of the intelligent testing philosophy discussed is that hypotheses generated from the profile of the main assessment instrument should be supported with data from multiple sources. Accordingly, this chapter discusses how to supplement the Wechsler scales with additional cognitive tasks to support or clarify hypotheses raised from the initial cognitive battery. It is important for examiners to be knowledgeable and insightful in choosing supplemental measures to uncover exactly what an individual's cognitive strengths and weaknesses are, in order to form the basis for good recommendations. Specific examples are given for how examiners may supplement the Wechsler scales with each cognitive instrument presented. For example, hypotheses raised in the Wechsler profile regarding fluid reasoning may be further assessed by WJ-R subtests or KAIT subtests. If questions arise regarding the effect of response speed, the Binet IV or K-ABC Simultaneous subtests may be useful supplements. The DAS provides good supplemental information about nonverbal reasoning ability. Alternative modalities of receiving input and expression of response may be assessed by the DTLA-3 or K-ABC subtests. The K-ABC Achievement subtests may also provide additional information about an individual's verbal ability or crystallized abilities. Unique information about a child's planning ability may be found by supplementing a battery with the CAS. The multiple pieces of evidence provided by the supplemental tests, can be carefully integrated to confirm or deny hypotheses raised in an individual's cognitive profile. An illustrative case report is presented at the conclusion of this chapter, which provides an example of how many different sources of evidence are combined to provide a clear description of a 13-year-old female's cognitive and academic functioning.

The future direction of intellectual assessment appears to be in theory-driven instruments. The psychological community has held tightly onto the clinically-based Wechsler scales because of their large research base, clinical familiarity, and just tradition. Thus, psychometric and neurological theories are becoming an important base for the newer instruments of assessment, although their use by the majority of clinicians has only slowly been occurring. Technological advances are also eminently going to be further impacting the field of assessment. The use of computer science and neurological measurement will likely begin to change the face of cognitive assessment within the next few decades.

4.08.6 ILLUSTRATIVE CASE REPORT

The following case report is of Laura S., a 13-year-old seventh grader, experiencing difficulty on standardized tests in the areas of vocabulary, reading comprehension, and writing mechanics. This report illustrates the methods and procedures for test integration and interpretation described earlier in this chapter. Table 12 provides the specific scores earned by Laura on each instrument administered.

4.08.6.1 Referral and Background Information

Laura was brought in for an evaluation by her parents, Linda and Rob (Mr. & Mrs. S) who were referred for an assessment by her current school, due to concern about the inconsistency between Laura's high grades at school and her lower standardized test scores in some areas. From Laura's scores on the Educational Record Bureau's (ERB) Comprehensive Testing Program (CTPII), the main area of concern to Laura's parents was in her Verbal Ability, specifically, Vocabulary, Reading Comprehension, and Writing Mechanics. Mr. and Mrs. S. would like to gain a better understanding of Laura's difficulty with her mathematics courses, reading comprehension, and memory retention. Laura's parent's reported that she does not always seem to understand written instructions and asks a lot of questions. They would specifically like recommendations to help Laura improve her learning ability, enhance her memory, and develop a greater aptitude for understanding directions.

Laura is a 13.5-year-old adolescent girl who has lived at home with her mother and step-father since age three. At age 2.5., her mother and biological father separated, and, other than for one month of her life, she has lived with her mother. Laura also has a 20-year-old step-sister, Kelly, and a 23-year-old step-brother, Rick, who do not live with her at home. Mr. and Mrs. S. both work full-time as professionals outside the home. Laura's biological father lives out of state and Laura visits him once or twice a year. Laura has also had other adults serve as caretakers for her. From ages 1–3, her mother indicated that two other people helped care for her, and from ages 4–13, five housekeepers also provided care for her.

Mrs. S. reported that she had a normal, full-term pregnancy with Laura. There were no problems during the birth, and Laura was born weighing six pounds, nine ounces after a short labor of only 15 minutes. Laura's medical history is relatively unremarkable, with her parents noting only that she experienced a bilateral hernia at nine months and had several earaches and colds as a younger child. Mr. and Mrs. S. stated that their daughter has never been hospitalized, has had no major injuries, or diseases. All of Laura's developmental milestones were "on time" or "early" according to her mother.

A medical question of Laura's parents had been her hearing ability, due to her noted difficulty sometimes distinguishing particular sounds in school and occasional single-word substitutions. After discussing this at the current evaluation's intake interview, Mr. and Mrs. S. took Laura to a physician at a local university's Medicine Ear Institute to rule out any potential hearing loss. According to a letter sent by the physician, he performed a physical exam and audiogram for Laura. He stated that her audiogram "revealed normal thresholds bilaterally with excellent speech discrimination," and "she appears to have normal hearing."

Laura's educational history began when she entered preschool at age three. According to her parents, she had no difficulties with beginning preschool or transitioning to her next school. All of Laura's report cards from first through seventh grades indicate consistent, excellent achievement. Most of her teachers commented on her conscientious and serious approach to learning, and express their great pleasure having her as a student.

A classroom observation of Laura was performed for this evaluation. Laura was observed in her Honors math class. Laura was very friendly, chatting with her friends and brightly greeting the teacher right before class. However, as soon as the class began, she got right to work. When the instructor asked questions of the class, Laura raised her hand to every question. She was focused on the work all throughout the class, and seemed motivated to do well in math. She asked the teacher for assistance several times during the observation. In an interview, the teacher stated that Laura frequently asks for extra help, but it is her teacher's belief that this is mainly because "she

Table 12 Laura: Psychometric summary.

WISC-III

Scale	*IQ*	*Percentile rank*	*Factor*	*Index*	*Percentile rank*
Verbal scale	114±5	82	Verbal comprehension	120±5	91
Performance scale	95±5	37	Perceptual organization	94±6	34
Full scale	106±4	66	Freedom from distractibility	101±8	53
			Processing speed	111±7	77

Subtest	*Scaled score*	*Percentile rank*	*Subtest*	*Scaled score*	*Percentile rank*
Information	11	63	Picture Completion	8	25
Similarities	14	91	Coding	14	91
Arithmetic	8 -W	25	Picture Arrangement	8	25
Vocabulary	12	75	Block Design	11	63
Comprehension	17 -S	99	Object Assembly	9	37
Digit span	12	75	Symbol Search	10	50

KAIT

	IQ	*Percentile rank (age)*
Fluid scale	120±5	91

Subtest	*Standard score*	*Percentile rank (age)*
Rebus Learning	16	98
Logical Steps	12	75
Mystery Codes	13	84

WJ-R	*Standard score*	*Percentile rank*
Broad reading	118	89
Basic reading skills	125	95
Reading comprehension	113	81
Letter word identification	125	95
Passage comprehension	106	65
Word attack	117	87
Reading vocabulary	118	88
Broad mathematics	111	77
Basic math skills	115	85
Calculation	117	86
Applied problems	102	56
Quantitative concepts	108	70
Broad written language	105	64
Basic writing skills	97	41
Written expression	130	98
Dictation	95	37
Writing samples	123	94
Proofing	99	48
Writing fluency	134	99
Punctuation and capitalization	97	41
Spelling	100	50
Usage	98	44
Broad knowledge	99	48
Science	97	42
Social studies	96	40
Humanities	110	74

Table 12 (continued)

WJ-R (intra-achievement discrepancies)	*Actual standard score*	*Predicted standard score*	*Standard deviation difference*	*Percentile rank*
Broad reading	118	105	+1.48	93
Broad mathematics	111	106	+0.52	70
Broad written language	105	109	−0.45	32
Broad knowledge	99	113	−1.20	12

does not trust her gut." The teacher said that Laura always begins the class with a happy, cheerful mood, but she tends to "stress out" when problems become difficult.

In discussing family history related to Laura's difficulties, Mrs. S. indicated that she feels she also has "a poor memory and retrieving skills." Mr. and Mrs. S. noted that Laura has several strengths, including her strong intuitive abilities, her confidence, and her persistence. Laura is very popular with her peers and is viewed as a leader by many of her teachers.

In an interview with Laura, she stated that she does not enjoy reading but likes using her creativity in writing. She said that she takes school more seriously than most of her peers, and has set very high standards for herself. She explained that she is harder on herself than her parents are, especially when it comes to academics. In her free time Laura works diligently on her homework and enjoys playing tennis and soccer, and socializing with her friends.

4.08.6.2 Appearance and Behavioral Characteristics

Laura is a mature, pretty, 13-year-old seventh grader. She wore her thick brown hair parted stylishly down the middle. For each of the evaluation sessions she dressed casually and neatly in jeans and a sweatshirt. She talked comfortably when asked questions by the examiner and spoke in soft voice with an air of politeness and good manners. She tended to respond with short phrases rather than in complete sentences when answering questions posed by the examiner. However, this did not detract from her ability to articulate her thoughts clearly.

She seemed eager to please the examiner, often asking for clarification about what was the right thing to do during a certain subtest. For example, during a task which required her to look at a picture and tell what important part was missing, she asked, "Can I tell you what's wrong with it?" Her frequent questions to the examiner indicated not only her anxiousness to please, but also her uncomfortableness with ambiguity in a situation. She appeared much more relaxed when the situation was structured and she knew clearly everything that was expected of her.

Laura demonstrated a strong ability to concentrate and focus. She had great stamina throughout the mentally challenging evaluation. She was persistent in always refusing breaks offered to her and worked straight through during each session, displaying extreme self-control. She was cooperative and friendly, often helping the examiner put away stimulus materials and always following directions. Encouragement from the examiner was welcomed warmly with a smile by Laura. She gradually became less cautious in her casual conversation with the examiner as the testing progressed. She began to share bits and pieces of her life and showed a well-rounded self.

Laura always tried her best. Even after having attempted problems that were difficult for her, she did not lose her motivation to keep trying. In solving problems she worked quickly, but was reflective. She would continue to check her work, even after she was done, always being careful in her responses. Laura expressed anxiety about having to solve mathematical problems mentally, saying, "I can't do things in my head, without pencils and things." She was more tentative in answering such questions. On more difficult nonverbal tasks, Laura tended to analyze the situation first and then proceed with the task. Again, evidence of her trying carefully to do her best and not make mistakes. On the basis of her behavior during the evaluation these results should be considered a valid representation of her ability.

4.08.6.3 Tests Administered

(i) Wechsler Intelligence Scale for Children-3rd edition (WISC-III)

(ii) Woodcock–Johnson-Revised (WJ-R): Tests of Achievement

(iii) Kaufman Adolescent and Adult Intelligence Test (KAIT): Fluid Subtests

(iv) Rotter Incomplete Sentences.

4.08.6.4 Test Results and Interpretation

Laura was administered a series of cognitive tests to assess her information processing abilities. According to the WISC-III, Laura is functioning currently within the Average to High Average range of intelligence. She obtained a Verbal IQ score of 114 ± 5 (82nd percentile), Performance IQ score of 95 ± 5 (37th percentile), and Full Scale IQ score of 106 ± 4 (66th percentile). In addition, she also obtained a Verbal Comprehension Index of 120 ± 5, which was significantly higher than her Freedom from Distractibility Index of 101 ± 8. However, this difference within her verbal scale was due mainly to her difficulty computing arithmetic problems mentally. Laura's Processing Speed Index of 111 ± 7 was significantly higher than her Perceptual Organization Index of 94 ± 6, indicating that she performed better on tests of visual processing speed than on tests of nonverbal reasoning.

It is important to note that Laura's Verbal subtest scores exhibit a significant amount of scatter, suggesting that her Verbal Comprehension and Perceptual Organization Indices provide a more meaningful picture of her abilities than the overall Full Scale IQ, Verbal IQ, or Performance IQ. The variance in her Verbal Scale indicates that some of her abilities are more well developed than others. The difference between her verbal and nonverbal abilities, as reflected by her Factor Indices, is unusually large and statistically significant. Because there is a 26-point difference in favor of her Verbal Comprehension Index over her Perceptual Organization Index (occurring in less than 5% of normal children), her Full Scale IQ should not be used as an indication of her overall ability. A fuller and clearer picture of Laura's abilities can be found by looking at her performance in individual areas rather than considering the statistical average of these various abilities.

In the verbal area, Laura demonstrated a significant weakness in Arithmetic, earning a score in the 25th percentile. The WISC-III arithmetic does not allow use of paper and pencil to do computation, so Laura was required to manipulate the numbers mentally to solve auditorally and visually presented problems. She expressed several times during this test that she needed to see the numbers and have them down concretely in front of her to figure the problems out. Her weakness on WISC-III Arithmetic was in contrast to her higher scores in the mathematics area on Woodcock–Johnson-Revised Tests of Achievement (WJ-R). All of the mathematics on the WJ-R allowed the use of paper and pencil in solving the problems, which Laura was more comfortable with. Her performance on problems of calculation, such as addition, subtraction, multiplication, and division with multiple digit numbers, decimals, and fractions, was significantly better than her performance on applied problems requiring her to use mathematics to solve problems involving scenarios with money, distance, and weight. On the Woodcock–Johnson-Revised Mathematics Subtests, she scored at the 86th percentile on calculation and 56th percentile on problems that were applied. Thus, her weakness in Arithmetic on the WISC-III is not due to poor computation ability, but rather due to the fact that she needs the concrete visual stimulus of written numbers in order to utilize the mathematical knowledge that she does have.

Laura's variability in her performance on the Verbal Scale of the WISC-III was accentuated by her extremely high score (99th percentile) on a test of social judgment, verbal reasoning, and practical knowledge. Her strong verbal ability was apparent during this subtest, as well as on a subtest requiring her to use abstract reasoning to describe how two things are similar. This strength was also paralleled by her excellent performance on WJ-R tests of written expression, on which she earned an overall score at the 98th percentile. She can come up easily with vocabulary to express her ideas and is able to formulate alternative ways to get her point across if it is unclear. However, the mechanical details of written expression are more difficult for Laura. For example, she earned a lower score on a test of dictation (37th percentile), which examined her spelling, punctuation, and word usage. Laura's performance in spelling, punctuation, and word usage was at a lower level than her overall Written Expression ability (41st, 50th, and 44th percentiles, in contrast to 98th percentile). Thus, she is able to get her verbal ideas across, but is lacking skill in the grammatical rules and details of written expression.

Laura's verbal reasoning ability also appeared stronger than her knowledge of general factual knowledge. She scored at the 63rd percentile on a WISC-III task requiring her to answer questions about information acquired from to formal schooling. However, her performance on WJ-R tests of Broad Knowledge was lower than expected, given her high level of academic achievement at school, as well as her performance on other WJ-R tests of achievement. On subtests covering areas such as science, social studies, and humanities, Laura scored in the Average level at the 48th percentile. A person with Laura's total achievement performance would have been expected to earn a slightly higher score on these tests of Broad Knowledge. Her actual standard score was 1.2 standard deviations lower than what was predicted in this area, indicating that she is not achieving at a level consistent with what would be expected given her level of achievement. This is also reflected in her lower ERB group standardized test scores that are discrepant from her higher school grades. The ERB scores reflect strictly facts derived from multiple choice exams. However, her interaction with teachers, participation and performance at school allow the teachers to understand Laura as a more complete person, which positively influences her grades.

Laura's performance on WJ-R Broad Reading (89th percentile) is also commensurate with her strong verbal reasoning skills evidenced. However, her score on a task measuring comprehension of a written passage was somewhat lower than expected (65th percentile) given her strong verbal abilities. Nonetheless, it is in the average range and not low enough to be of concern. Overall, Laura's strong performance on Broad Reading ability was 1.48 standard deviations above what would be predicted for an individual with her total achievement performance, so she is demonstrating strong ability with her reading skills. Only 7% of students Laura's age, who had the same expected standard score as she, scored as high or higher than Laura on Broad Reading Subtests of the WJ-R.

Laura's significantly higher score on the Verbal Comprehension Index than the Perceptual Organization Index of the WISC-III indicates that her verbal abilities are more well developed than her nonverbal and perceptual abilities. However, even on the nonverbal performance tests, all of her scores were at or above the Average level (25th to 91st percentiles). For example, Laura performed well on a task requiring her to use short-term memory in copying a series of different symbols from a visually presented key. She scored at the 91st percentile on this task and scored at the 63rd percentile on a another nonverbal task requiring her to reproduce a model using blocks.

Her cognitive skills were further assessed by the Fluid Scale of the Kaufman Adolescent and Adult Intelligence Test (KAIT). The KAIT Fluid subtests measure one's ability to solve novel problems using reasoning, memory, paired-associate learning, verbal comprehension, and perceptual organization. Laura scored well above average (standard score 120 ± 5) on these tasks involving her ability to solve novel problems. She scored at the 98th percentile on a task that essentially required her to learn a new language through paired-associate learning. On this task she used her good verbal concentration, expression ability and memory to succeed. She also performed quite well on tasks that required her to use abstract reasoning with novel stimuli and planning ability.

4.08.6.5 Summary and Diagnostic Impressions

Laura is a mature, pretty 13-year-old girl who was brought in for a psychoeducational evaluation by her parents due to their concern with inconsistency between her standardized tests scores and high grades at La Jolla Country Day School. Her parents' main areas of concern involve her abilities in vocabulary, reading comprehension, writing mechanics, and mathematics. Mr. and Mrs. S. wanted a better understanding of these discrepancies in her performance, as well as her difficulty with memory retention and understanding written directions.

Laura demonstrated strong motivation to please the examiner, persistence, and stamina during her evaluation. She was reflective in her problem solving and careful in her responses, trying hard to perform to the best of her ability. She earned scores in Average to High Average range of intelligence on the WISC-III. Her factor indices and specific strengths and weaknesses gave the most meaningful picture of Laura's abilities, due to the scatter in her verbal subscales. She performed significantly better on tests of verbal reasoning and word knowledge than on tests of nonverbal ability and visual-perceptual skills. She demonstrated a weakness on a task requiring her to mentally solve auditorally and visually presented arithmetic problems. However, this was due to her difficulty performing mathematical calculations in her head without the concrete visual help of pencil and paper, as evidenced by her higher scores on WJ-R tests of Mathematics that allow use of pencil and paper for problem-solving.

Her strengths were in the area of verbal reasoning. This was congruent with her strong ability to express herself in a written format on the WJ-R and also with her overall reading ability demonstrated on the WJ-R. Although her ability to express herself verbally and in writing was good, her skills in the details of writing, such as spelling, punctuation, and usage were not as strong. In addition, her performance on a test of passage comprehension was not as high as expected given her overall verbal abilities. In the area of general knowledge, including science, social studies, and humanities, Laura did not perform as well as one would predict given her other achievement scores.

On an additional cognitive test measuring Laura's ability to solve novel problems, she demonstrated well above average abilities. On the KAIT Fluid subtests, Laura's performance indicated that she has strong ability to learn new material through paired-associate learning. She also evidenced above average ability to use abstract reasoning and planning ability with novel stimuli.

Laura has learned very well to compensate for her areas of weakness, which is evident in her high grades at school. Her ability to ask questions when uncertain, and create a structured environment for herself so she is most comfortable are examples of such compensatory strategies. Additionally, this strength is reflected in her higher grades at school compared to her ERB scores. Her ability to express herself well in writing and vocally, such as on essay tests or class discussions, helps her grades at school but is not able to aid her in the cut-and-dry responses needed for the ERB tests. Her strong ability to solve novel problems will be quite useful to Laura in continuing to creatively create other strategies when new difficulties appear in her life.

4.08.6.6 Recommendations

The following recommendations have been made to assist Laura and her parents in enhancing her learning ability.

(i) Laura is a highly self-motivated student who has created very high expectations for herself, and has been working very hard to meet those high standards. At times this causes anxiety for her, and when this anxiety reaches an unmanageable level it may cause difficulty and decreased performance. It will therefore be useful for Laura and her parents to discuss that it is okay not to be perfect in every area of her life, every minute of the day, and to allow Laura to experience those instances of nonperfection. Tolerance and appreciation for her own continuum of strengths to weaknesses needs to be developed to keep Laura feeling good about herself.

(ii) As Laura works so hard at not making any mistakes, she is sometimes hesitant to proceed without asking many questions to prevent making any sort of error. She demonstrated strong abilities in many areas and therefore should be encouraged to go with her gut feeling, and try to attempt problems that normally she may immediately ask for help on. This will further foster her sense of independence and confidence in her own abilities.

(iii) Laura expressed being uninterested in some subjects that have not stressed problem-solving or reasoning as learning methods, such as Social Studies. If her school classes present material in a factual, "memorize this" style, then Laura needs to study in more creative way to enhance her learning and reduce boredom. For example, she can make up a story about a character that may have lived in a period of history that she is learning about. As another example, Laura can draw and illustrate a timeline of important historical events to remember.

(iv) To incorporate a different method of studying subjects that are especially tedious to Laura, she could study with another diligent student. Similarly, she may want to get together with a group of peers and create a quiz show to reinforce and study facts that may seem boring to her when she is studying alone.

(v) To encourage her to become comfortable with and reinforce information that may be presented in a test format such as the ERB tests, Laura may create a competition for herself. For example, on a weekly basis she could give herself a test on a content area (such as those in books created for SAT preparation), and then she could chart her own progress from week to week.

(vi) It is important to reassure Laura that it is fine to use the compensatory strategies that work for her, such as writing down arithmetic problems. Her ability to figure out such strategies is a strength that was evident in her ability to solve novel problems and can be used to help her in areas that are more difficult for her.

(vii) To help Laura increase her overall academic abilities, Laura will benefit from broadening her base of what she considers "studying." Studying includes not only completing assignments given at school and preparing for exams, but also includes an awareness of one's environment outside of the school context. In broadening her conceptualization of studying, for example, Laura may incorporate more pleasure reading of nonschool books or magazines into her weekly routine, which will benefit her grammar and vocabulary. She may watch a movie or television program and relate it to some topic she is studying in school. This is important as people who improve their general learning, tend to do better overall on standardized types of testing.

(viii) Laura demonstrated strong written expression ability, but weaker ability to incorporate correctly details of grammar such as spelling and punctuation. These details may be less interesting to Laura, thus, again she may want to use her problem-solving ability to create more interesting ways to learn such details. There are many excellent computer programs that help teach grammatical details, in an interesting manner. As Laura tends to strive for excellence, she may set goals for herself according to a computer program she is working with to help her.

(ix) Laura is very conscientious and focused in her school work. She may benefit and find it rewarding to become a peer tutor for a lower grade child (such as a fourth- or fifth-grade student). Tutoring a child who is having trouble with grammar will reinforce rules in Laura's mind as she gains self respect by being appointed to teach someone else.

ACKNOWLEDGMENTS

The authors would like to thank Drs. Kristee A. Beres, Randy Kamphaus, Nadeen L. Kaufman, Jack A. Naglieri, and Mitch Perlman for their contributions to this chapter.

4.08.7 REFERENCES

American Psychological Association (1990). *Standards for educational and psychological tests and manuals*. Washington, DC: Author.

Bain, S. K. (1991). Test reviews: Differential ability scales. *Journal of Psychoeducational Assessment, 9,* 372–378.

Binet, A., & Henri, V. (1895). La psychologie individuelle. *L'Année Psycholgique, 2,* 411–465.

Binet, A., & Simon, T. (1905). Méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux. *L'Année Psychologique, 11,* 191–244.

Binet, A., & Simon, T. (1908). Le développement de l'intelligence chez les enfants. *L'Année Psychologique, 14,* 1–94.

Bogen, J. E. (1975). Some educational aspects of hemispheric specialization. *UCLA Educator, 17,* 24–32.

Bracken, B. A. (1985). A critical review of the Kaufman Assessment Battery for Children (K-ABC). *School Psychology Review, 14,* 21–36.

Braden, J. P. (1992). Test reviews: The differential ability scales and special education. *Journal of Psychoeducational Assessment, 10,* 92–98.

Brown, D. T. (1994). Review of the Kaufman Adolescent and Adult Intelligence Test (KAIT). *Journal of School Psychology, 32,* 85–99.

Buckhalt, J. A. (1991). A critical review of the Wechsler Preschool and Primary Scale of Intelligence Revised (WPPSI-R). *Journal of Psychoeducational Assessment, 9,* 271–279.

Canter, A. (1990). A new Binet, an old premise: A mismatch between technology and evolving practice. *Journal of Psychoeducational Assessment, 8,* 443–450.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54,* 1–22.

Cohen, J. (1957). A factor-analytically based rationale for the Wechsler-Adult Intelligence Scale. *Journal of Consulting Psychology, 6,* 451–457.

Cohen, R. J., Montague, P., Nathanson, L. S., & Swerdlik, M. E. (1988). *Psychological testing*. Mountain View, CA: Mayfield.

Das, J. P. (1973). Structure of cognitive abilities: Evidence for simultaneous and successive processing. *Journal of Educational Psychology, 65,* 103–108.

Das, J. P., Kirby, J. R., & Jarman, R. F. (1975). Simultaneous and successive synthesis: An alternative model for cognitive abilities. *Psychological Bulletin, 82,* 87–103.

Das, J. P., Kirby, J., & Jarman, R. F. (1979). *Simultaneous and successive cognitive processes*. New York: Academic Press.

Das, J. P., Naglieri, J. A., & Kirby, J. (1994). *Assessment of cognitive processes*. Boston: Allyn & Bacon.

Delugach, R. (1991). Test review: Wechsler Preschool and Primary Scale of Intelligence-Revised. *Journal of Psychoeducational Assessment, 9,* 280–290.

Dumont, R., & Hagberg, C. (1994). Test reviews: Kaufman Adolescent and Adult Intelligence Test (KAIT). *Journal of Psychoeducational Assessment, 12,* 190–196.

Elliott, C. D. (1990). *Differential Ability Scales (DAS) administration and scoring manual*. San Antonio, TX: Psychological Corporation.

Elliott, C. D. (1997). The Differential Ability Scales (DAS). In D. P. Flanagan, J. L. Gensaft, & P. L. Harrison (Eds.), *Beyond traditional intellectual assessments: Contemporary and emerging theories, tests, and issues* (pp. 183–208). New York: Guilford Press.

Evans, J. H., Carlsen, R. N., & McGrew, K. S. (1993). Classification of exceptional students with the Woodcock–Johnson Psycho-Educational Battery-Revised. In R. S. MacCallum (Ed.), *Journal of Psychoeducational Assessment monograph series. Advances in psychoeducational assessment: Woodcock–Johnson Psycho-educational Battery-Revised* (pp. 6–19). Germantown, TN: Psychoeducational Corporation.

Flanagan, D. P., Alfonso, V. C., & Flanagan, R. (1994). A review of the Kaufman Adolescent and Adult Intelligence Test: An advancement in cognitive assessment? *School Psychology Review, 23,* 512–525.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95,* 29–51.

Glutting, J. J., McDermott, P. A., Prifitera, A., & McGrath, E. A. (1994). Core profile types for the WISC-III and WIAT: Their development and application in identifying multivariate IQ-achievement discrepancies. *School Psychology Review, 23,* 619–639.

Goldern, C. J. (1981). The Luria-Nebraska Children's Battery: Theory and formulation. In G. W. Hynd & J. E. Obrzut (Eds.), *Neuropsychological assessment of the school-age child issues and procedures* (pp. 277–302). New York: Grune and Stratton.

Hammill, D. D. (1991). *Interpretive Manual for Detroit Tests of Learning Aptitude:* (3rd ed.). Austin, TX: PRO-ED.

Hammill, D. D., & Bryant, B. R. (1991). *Interpretive Manual for Detroit Tests of Learning Aptitude-Primary: Second Edition*. Austin, TX: PRO-ED.

Hansen, J. C., & Campbell, D. P. (1985). *Manual for the SVIB-SCII* (4th ed.). Stanford, CA: Stanford University Press (Distributed by Consulting Psychologists Press).

Horn, J. L. (1985). Remodeling old model in intelligence. In B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurements, and applications* (pp. 267–300). New York: Wiley.

Horn, J. L. (1989). Cognitive diversity: A framework of learning. In P. L. Ackerman, R. J. Sternberg, & R. Glaser (Eds.), *Learning and individual differences* (pp. 61–116). New York: Freeman.

Horn, J. L. (1991). Measurement of intellectual capabilities: A review of theory. In K. S. McGrew, J. K. Werder, & R. W. Woodcock (Eds.), *Woodcock–Johnson Technical manual: A reference on theory: and current research* (pp. 197–246). Allen, TX: DLM Teaching Resources.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology, 57,* 253–270.

Horn, J. L., & Cattell, R. B. (1967). Age difference in fluid and crystallized intelligence. *Acta Psychologica, 26,* 107–129.

Hoy, C., Gregg, N., Jagota, M., King, M., Moreland, C., & Manglitz, E. (1993). Relationship between the Wechsler

Adult Intelligence Scale-Revised and the Woodcock–Johnson Test of Cognitive Ability-Revised among adults with learning disabilities in university and rehabilitation settings. In R. S. MacCallum (Ed.), *Journal of Psychoeducational Assessment monograph series. Advances in psychoeducational assessment: Woodcock–Johnson Psycho-educational Battery-Rivised* (pp. 54–63). Germantown, TN: Psychoeducational Corporation.

Inglis, J., & Lawson, J. (1982). A meta-analysis of sex differences in the effects of unilateral brain damage on intelligence test results. *Canadian Journal of Psychology, 36*, 670–683.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Kamphaus, R. W. (1993). *Clinical assessment of children's intelligence*. Boston: Allyn & Bacon.

Kamphaus, R. W., Beres, K. A., Kaufman, A. S., & Kaufman, N. L. (1995). The Kaufman Assessment Battery for Children (K-ABC). In C. S. Newmark (Ed.), *Major psychological assessment instruments* (2nd ed., pp. 348–399). Boston: Allyn & Bacon.

Kamphaus, R. W., & Reynolds, C. R. (1987). *Clinical and research applications of the K-ABC*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S. (1979). *Intelligent testing with the WISC-R*. New York: Wiley.

Kaufman, A. S. (1983). Intelligence: Old concepts—new perspectives. In G. Hynd (Ed.), *The school psychologist* (pp. 95–117). Syracuse, NY: Syracuse University Press.

Kaufman, A. S. (1985). Review of Wechsler Adult Intelligence Scale-Revised. In J. V. Mitchell (Ed.), *The ninth mental measurement yearbook* (pp. 1699–1765). Lincoln, NE: University of Nebraska Press.

Kaufman, A. S. (1990). *Assessing adolescent and adult intelligence*. Boston: Allyn & Bacon.

Kaufman, A. S. (1992). Evaluation of the WISC-III and WPPSI-R for gifted children. *Roeper Review, 14*, 154–158.

Kaufman, A. S. (1993). King WISC the third assumes the throne. *Journal of School Psychology, 31*, 345–354.

Kaufman, A. S. (1994a). A reply to MacMann and Barnett: Lessons form the blind men and the elephant. *School Psychology Quarterly, 9*, 199–207.

Kaufman, A. S. (1994b). *Intelligent testing with the WISC-III*. New York: Wiley.

Kaufman, A. S., & Doppelt, J. E. (1976). Analysis of WISC-R standardization data in terms of the stratification variables. *Child Development, 47*, 165–171.

Kaufman, A. S., & Horn, J. L. (1996). Age changes on test of fluid and crystallized ability for women and men on the Kaufman Adolescent and Adult Intelligence Test (KAIT) at ages 17–94 years. *Archives of Clinical Neuropsychology, 11*, 97–121.

Kaufman, A. S., & Kamphaus, R. W. (1984). Factor analysis of the Kaufman Assessment Battery for Children (K-ABC) for ages $2^1{}_2$ through $12^1{}_2$ years. *Journal of Educational Psychology, 76*, 623–637.

Kaufman, A. S., & Kaufman, N. L. (1983). *Interpretive manual for the Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S., & Kaufman, N. L. (1993). *Interpretive Manual for Kaufman Adolescent & Adult Intelligence Test*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S., Ishikuma, T., & Kaufman, N. L. (1994). A Horn analysis of the factors measured by the WAIS-R, Kaufman Adolescent and Adult Intelligence Test (KAIT), and two new brief cognitive measures for normal adolescents and adults. *Assessment, 1*, 353–366.

Kaufman, A. S., Kaufman, J. C., Balgopal, R., & McLean, J. E. (1996). Comparison of three WISC-III short forms: Weighing psychometric, clinical, and practical factors. *Journal of Clinical Child Psychology, 25*, 97–105.

Kaufman, A. S., Kaufman, J. C., Chen, T., Kaufman, N. L. (1996). Differences on six Horn abilities for 14 age groups between 15–16 and 75–94 years. *Psychological Assessment, 8*, 1–11.

Kaufman, A. S., & Lichtenberger, E. O. (in press). *WAIS-III assessment made simple*. New York: Wiley.

Kaufman, A. S., & McLean, J. E. (1992, November). *An investigation into the relationship between interests and intelligence*. Paper presented at Annual meeting of the Mid-South Educational Research Association, Knoxville, TN.

Kaufman, A. S., McLean, J. E., & Lincoln, A. (1996). The relationship of the Myers–Briggs Type Indicator to IQ level and fluid-crystallized discrepancy on the Kaufman Adolescent and Adult Intelligence Test (KAIT). *Assessment, 3*, 225–239.

Kaufman, A. S., McLean, J. E., & Reynolds, C. R. (1988). Sex, race, residence, region, and education differences on the 11 WAIS-R subtests. *Journal of Clinical Psychology, 44*, 213–248.

Keith, T. Z. (1990). Confirmatory and hierarchical confirmatory analysis of the Differential Ability Scales. *Journal of Psychoeducational Assessment, 8*, 391–405.

Keith, T. Z., & Dunbar, S. B. (1984). Hierarchical factor analysis of the K-ABC: Testing alternate models. *Journal of Special Education, 18*, 367–375.

Kinsbourne, M. (Ed.) (1978). *Asymmetrical function of the brain*. Cambridge, MA: Cambridge University Press.

Levy, J., & Trevarthen, C. (1976). Metacontrol of hemispheric function in human split-brain patients. *Journal of Experimental Psychology: Human Perception and Performance, 2*, 299–312.

Luria, A. R. (1966). *Higher cortical functions in man*. New York: Basic Books.

Luria, A. R. (1973). *The working brain: An introduction to neuro-psychology*. London: Penguin.

Luria, A. R. (1980). *Higher cortical functions in man* (2nd ed.). New York: Basic Books.

MacMann, G. M., & Barnett, D. W. (1994). Structural analysis of correlated factors: Lessons form the Verbal-Performance dichotomy of the Wechsler scales. *School Psychology Quarterly, 9*, 161–197.

Matarazzo, J. D. (1985). Review of Wechsler Adult Intelligence Scale-Revised. In J. V. Mitchell (Ed.), *The ninth mental measurement yearbook* (pp. 1703–1705). Lincoln, NE: University of Nebraska Press.

McCallum, R. S., & Merritt, F. M. (1983). Simultaneous-successive processing among college students. *Journal of Psychoeducational Assessment, 1*, 85–93.

McDermott, P. A., Fantuzzo, J. W., & Glutting, J. J. (1990). Just say no to subtest analysis: A critique on Wechsler theory and practice. *Journal of Psychoeducational Assessment, 8*, 290–302.

McDermott, P. A., Fantuzzo, J. W., & Glutting, J. J., Watkins, M. W., & Baggaley, A. R., (1992). Illusions of meaning in the ipsative assessment of children's ability. *Journal of Special Education, 25*, 504–526.

McGhee, R. (1993). Fluid and crystallized intelligence: Confirmatory factor analysis of the Differential Ability Scales, Detroit Tests of Learning Aptitude-3, and Woodcock–Johnson Psycho-Educational Battery-Revised. In B. A. Bracken & R. S. McCallum (Eds.), *Journal of Psychoeducational/Assessment monograph series, advances in psychoeducational assessment: Woodcock–Johnson Psycho-Educational Battery-Revised* (pp. 39–53). Germantown, TN: Psychoeducational Corporation.

McGrew, K. S., & Flanagan, D. P. (1996). The Wechsler Performance scale debate: Fluid intelligence (Gf) or visual processing. *NASP Communique, 24*, 15–17.

McShane, D., & Cook, V. (1985). Transcultural intellectual assessment: Performance by Hispanics on the Wechsler scales. In B. B. Wolman (Ed.), *Handbook of intelligence* (pp. 385–426). New York: Wiley.

McShane D. A., & Plas, J. M. (1984). The cognitive functioning of American Indian children: Moving from the WISC to the WISC-R. *School Psychology Review, 13,* 61–73.

Merz, W. R. (1985). Test review of Kaufman Assessment Battery for Children. In D. J. Keyser & R. C. Sweetland (Eds.), *Test overviews* (pp. 393–405). Kansas City, MO: Test Corporation of America.

Miller, T. L., & Reynolds, C. R. (1984). Special issue ... The K-ABC. *Journal of Special Education, 8,* 207–448.

Naglieri, J. A. (1984). Concurrent and predictive validity of the Kaufman Assessment Battery for Children with a Navajo sample. *Journal of School Psychology, 22,* 373–380.

Naglieri, J. A., & Das, J. P. (1988). Planning–Arousal–Simultaneous–Successive (PASS): A model for assessment. *Journal of School Psychology, 26,* 35–48.

Naglieri, J. A., & Das, J. P. (1990). Planning, Attention, Simultaneous, and Successive (PASS) cognitive processes as a model for intelligence. *Journal of Psychoeducational Assessment, 8,* 303–337.

Naglieri, J. A., & Das, J. P. (1996). *Das–Naglieri Cognitive Assessment System*. Chicago: Riverside.

Naglieri, J. A., & Jensen, A. R. (1987). Comparison of black-white differences on the WISC-R and the K-ABC: Spearmen's hypothesis. *Intelligence, 11,* 21–43.

Obringer, S. J. (1988, November). *A survey of perceptions by school psychologists of the Stanford–Binet IV*. Paper presented at the meeting of the Mid-South Educational Research Association, Louisville, KY.

O'Grady, K. E. (1983). A confirmatory maximum likelihood factor analysis of the WAIS-R. *Journal of Consulting and Clinical Psychology, 51,* 826–831.

Perlman, M. D. (1986). *Toward an integration of a cognitive-dynamic view of personality: The relationship between defense mechanisms, cognitive style, attentional focus, and neuropsychological processing*. Unpublished doctoral dissertation, California School of Professional Psychology, San Diego.

Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. *Human Development, 15,* 1–12.

The Psychological Corporation (1997). *Wechsler Adult Intelligence Scale-Third edition (WAIS-III)*. San Antonio, TX: Author.

Reitan, R. M. (1955). Certain differential effects of left and right cerebral lesions in human adults. *Journal of Comparative and Physiological Psychology, 48,* 474–477.

Reschly, D. J., & Tilly, W. D. (1993, September). The WHY of system reform. *Communique*, pp. 1, 4–6.

Reynolds, C. R. (1987). Playing IQ roulette with the Stanford–Binet, 4th edition. *Measurement and Evaluation in Counseling and Development, 20,* 139–141.

Reynolds, C. R., Chastain, R. L., Kaufman, A. S., & McLean, J. E. (1987). Demographic characteristics and IQ among adults: Analysis of the WAIS-R standardization sample as a function of the stratification variables. *Journal of School Psychology, 25,* 323–342.

Reynolds, C. R., Kamphaus, R. W., & Rosenthal, B. L. (1988). Factor analysis of the Stanford–Binet Fourth Edition for ages 2 years through 23 years. *Measurement and Evaluation in Counseling and Development, 2,* 52–63.

Roback, A. A. (1961) *History of psychology and psychiatry*. New York: Philosophical Library.

Sandoval, J. (1992). Test Reviews: Using the DAS with multicultural populations: Issues of test bias. *Journal of Psychoeducational Assessment, 10,* 88–91.

Sattler, J. M. (1988). *Assessment of children* (3rd ed.). San Diego, CA: Sattler.

Schaw, S. R., Swerdlik, M. E., & Laurent, J. (1993). Review of the WISC-III. In B. A. Bracken & R. S. McCallum (Eds.), *Journal of Psychoeducational Assessment monograph series, advances in psychoeducational assessment: Wechsler Intelligence Scale for Children-Third edition* (pp. 151–160). Germantown, TN: Psychoeducational Corporation.

Schmidt, K. L. (1994). Review of Detroit Tests of Learning Aptitude-Third Edition. *Journal of Psychoeducational Assessment, 12,* 87–91.

Schofield, N. J., & Ashman, A. F. (1986). The relationship between Digit Span and cognitive processing across ability groups. *Intelligence, 10,* 59–73.

Siehen, F. A. (1985). *Correlational study of Woodcock–Johnson deviation IQ scores and WAIS-R with adult population*. Unpublished manuscript, Arizona State University.

Sperry, R. W. (1968). Hemisphere deconnection and unity in conscious awareness. *American Psychologist, 23,* 723–733.

Sperry, R. W. (1974). Lateral specialization in the surgically separated hemispheres. In F. O. Schmitt & F. G. Worden (Eds.), *The neurosciences: Third study program*. Cambridge, MA: MIT Press.

Spruill, J. (1984). Wechsler Intelligence Scale-Revised. In D. J. Keyser & R. C. Sweetland (Eds.), *Test overviews* (pp. 728–739). Kansas City, MO: Test Corporation of America.

Spruill, J. (1987). Review of Stanford–Binet Intelligence Scale, Fourth edition. In D. J. Keyser & R. C. Sweetland (Eds.), *Test overviews* (pp. 544–559). Kansas City, MO: Test Corporation of America.

Sternberg, R. J. (1993). Rocky's back again: A review of the WISC-III. In B. A. Bracken & R. S. McCallum (Eds.), *Journal of Psychoeducational Assessment monograph series, advances in psychoeducational assessment: Wechsler Intelligence Scale for Children-Third Edition* (pp. 161–164). Germantown, TN: Psychoeducational Corporation.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986a). *Technical manual for the Stanford–Binet Intelligence Scale-Fourth Edition*. Chicago: Riverside.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986b). *Stanford–Binet Intelligence Scale: Fourth Edition*. Chicago: Riverside.

Turkheimer, E., & Farace, E. (1992). A reanalysis of gender differences in IQ scores following unilateral brain lesions. *Psychological Assessment*, *4,* 498–501.

Turkheimer, E., Farace, E., Yfo, R. A., & Bigler, E. D. (1993). Quantitative analysis of gender differences in the effects of lateralized lesions on verbal and performance IQ. *Intelligence*, *17,* 461–474.

VanLeirsburg, P. (1994). Review of Detroit Tests of Learning Aptitude-3. In D. J. Keyser & R. C. Sweetland (Eds.), *Test overviews* (pp. 219–225). Kansas City, MO: Test Corporation of America.

Wada, J., Clarke, R., & Hamm, A. (1975). Cerebral hemisphere asymmetry in humans. *Archives of Neurology*, *37,* 234–246.

Watkins, M. W., & Kush, J. C. (1994). Wechsler subtest analysis: The right way, the wrong way, or no way? *School Psychology Review*, *23,* 640–651.

Webster, R. E. (1994). Review of Woodcock–Johnson Psycho-educational Battery-Revised. In D. J. Keyser & R. C. Sweetland (Eds.), *Test overviews* (pp. 804–815). Kansas City, MO: Test Corporation of America.

Wechsler, D. (1939). *Measurement of adult intelligence*. Baltimore: Williams & Wilkins.

Wechsler D. (1958). *Measurement and appraisal of adult intelligence* (4th ed.). Baltimore: Willilams & Wilkens.

Wechsler D. (1974). *Manual for the Wechsler Intelligence Scale for Children-Revised*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale-Revised (WAIS-R)*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1989). *Manual for the Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R)*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children-Third Edition, (WISC-III)*. San Antonio, TX: Psychological Corporation.

Witt, J. C., & Gresham, F. M. (1985). Review of the Wechsler Intelligence Scale for Children-Revised. In J. V. Mitchell (Ed.), *Ninth mental measurements yearbook* (pp. 1716–1719). Lincoln, NE: University of Nebraska Press.

Woodcock, R. W. (1990). Theoretical foundations of the WJ-R measures of cognitive ability. *Journal of Psychoeducational Assessment*, *8,* 231–258.

Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock–Johnson Tests of Cognitive Ability: Standard and supplemental batteries*. Chicago: Riverside.

Woodcock, R. W., & Mather, N. (1989). WJ-R Tests of Cognitive Ability-Standard and Supplemental Batteries: Examiner's Manual. In R. W. Woodcock & M. B. Johnson (Eds.) *Woodcock–Johnson psycho-educational battery-revised*. Allen, TX: DLM Teaching Resources.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.09
Assessment of Memory, Learning, and Special Aptitudes

ROBYN S. HESS
University of Nebraska at Kearney, NE, USA
and
RIK CARL D'AMATO
University of Northern Colorado, Greeley, CO, USA

4.09.1 INTRODUCTION 239
  4.09.1.1 Assessment Approaches 241
  4.09.1.2 Evaluation of Domain Areas 243
  4.09.1.3 Intervention Approaches 244
4.09.2 ASSESSMENT OF MEMORY 245
  4.09.2.1 Attention 246
  4.09.2.2 Short-term and Long-term Memory 247
  4.09.2.3 Memory: Implications for Intervention 249
4.09.3 ASSESSMENT OF LEARNING 249
  4.09.3.1 Models of Learning 249
  4.09.3.2 Learning Processes: Input and Integration 252
  4.09.3.3 Academic Achievement: Output 253
  4.09.3.4 Learning: Implications for Intervention 254
4.09.4 ASSESSMENT OF SPECIAL APTITUDES 254
  4.09.4.1 Sensory Perception 255
    4.09.4.1.1 Sensory perception: implications for intervention 256
  4.09.4.2 Motor: Fine and Gross 256
  4.09.4.3 Sensory-motor Integration 257
    4.09.4.3.1 Motor: implications for intervention 257
  4.09.4.4 Communication/Language 257
    4.09.4.4.1 Communication/language: implications for intervention 259
4.09.5 FUTURE DIRECTIONS 259
4.09.6 SUMMARY 260
4.09.7 REFERENCES 261

4.09.1 INTRODUCTION

The inner workings of the human mind and the way in which people process information has intrigued researchers for centuries. The traditional categories of cognition, such as attention, memory, language, and learning are terms frequently used and considered all important to our daily adaptive functioning and ability to learn new information. Yet all have defied simple explanation and manipulation (e.g., Gaddes & Edgell, 1994; Lezak, 1995). That is, as professionals we are able to identify when a child or adult is having difficulty processing information, but the exact interrelations between an individual's different capacities in areas such as attention, learning style, and sensory integration still eludes educational and clinical specialists. More puzzling still is finding effective rehabilitation strategies to address deficits in memory, learning, and other psychological processes. These areas are frequently addressed in the growing volumes of neuropsychological research, but are relevant to the practices of many traditionally trained psychologists as well (Hamsher, 1984). Not surprisingly, all of the information needed for an adequate understanding and interpretation of cognitive processes cannot be provided in a single chapter or obtained in an individual university course. Thus, the purpose of this chapter is to provide a brief description of our higher cognitive processes and introduce a variety of strategies and measures for evaluating these functions.

Problems in attention, memory, and learning are not isolated to the very young or the very old. Adult learning problems often become apparent in employment settings and after injuries resulting from strokes, accidents, or diseases. Epidemiological studies suggest that in the decade following the late 1990s there will be a dramatic increase in the number of individuals suffering from organically related disorders resulting from the abuse of alcohol and other types of toxic substances (Touyz, Byrne, & Gilandas, 1994). The recent documented increase in the number of head injuries caused by motor vehicle accidents has become an intrinsic fact in today's society. High speed transportation and the growing prevalence of violent street crimes have further increased the incidence of head injuries (Touyz et al., 1994). So too, the growing popularity of certain contact sports (e.g., hockey, boxing), noncontact sports (e.g., rock climbing, mountaineering, bicycling), and recreational activities (e.g., skateboarding, roller blading) has contributed significantly to the number of individuals suffering from traumatic brain injury (Drew & Templer, 1992; Templer & Drew, 1992). Internal processes such as eating disorders, depression, diseases, epilepsy, and tumors can result in impaired executive functioning as well (Black & Strub, 1994). These realities of today's society make it necessary for clinicians to be able to accurately evaluate a client's strengths and weaknesses in everyday functioning and find the key elements to fostering effective behavioral change through rehabilitation or educational improvement.

Psychology has made many new inroads into understanding the learning process and the subsequent development of corrective or adaptive programs for children and adults with learning disorders and traumatic brain injuries. As assessment specialists, psychologists must quickly and accurately wade through the cumulative data available about the individual in order to select the most viable of alternative hypotheses to explain the findings and offer appropriate interventions (D'Amato & Dean, 1989a; D'Amato, Rothlisberg, & Leu, in press; Gutkin & Reynolds, 1990). Although administering a test may be a routine activity, conducting a thorough, valid assessment is an extremely complex process. The clinician is required to make decisions regarding which skills to evaluate and the best instrument to use with a particular client, and to generate accurate interpretations of the results in order to create the most effective intervention plan. Adding to the immensity of this task are the wide range of client variables that can impact the assessment process, including motivation, environment, culture, age, developmental level, language, training, educational quality, personal experience, and attitude to name just a few (Golden, Sawicki, & Franzen, 1984; Hynd & Semrud-Clikeman, 1990).

The clinician must recognize that the context of the client may influence or even define the outcome of the assessment (Dana, 1993; Figueroa & Garcia, 1994). For example, many of the psychological, educational, and personality instruments available to practitioners have been criticized as culturally biased, as traditionally not including individuals from diverse ethnic backgrounds in the norming sample, and as measuring acquired knowledge rather than an individual's responsiveness to instruction or the learning process (Cole & Siegel, 1990; Dana, 1993; Figueroa & Garcia, 1994; Sattler, 1992). Because of these problems, several researchers believe that standardized assessment may have questionable validity for those clients who represent diverse cultural groups (Cole & Siegel, 1990; Dana, 1993; Figueroa & Garcia, 1994; Martinez, 1985). From the beginning, one robust assumption of standardized testing was that all individuals who take the tests would have had equal or comparable exposure to the contents of the assessment materials prior to the assessment (e.g., Colvin, 1921; Dearborn, 1921; Woodrow, 1921). In direct contrast to this supposition, current statistics indicate that the US immigrant population is not only growing rapidly but is also quickly expanding in diversity (Figueroa & Garcia, 1994). These authors conclude that tests, although given high status in US society, are actually quite fragile because of the founding assumption regarding homogeneity and general shortcomings in technical properties. Nevertheless, standardized tests can be useful in evaluating current functioning especially when multiple sources of information and multidimensional functions are evaluated to measure individual processes (Sattler, 1992). The responsible clinician must recognize both the assets and the limitations when using standardized measures with ethnic minority clients.

The strategy the practitioner uses to accomplish an effective assessment must, of necessity, be based upon well-grounded, empirically validated theories of cognition and behavior. Only through the use of a theoretical framework are specific predictions regarding performance under a given set of ecological circumstances made possible (Dean, 1985a, 1986; Rothlisberg, 1992). Unfortunately, no single, diagnostic paradigm or theory has proven sufficient to explain fully the vagaries of behavior (D'Amato & Rothlisberg, 1992). Psychoanalytically, behaviorally, and biologically based approaches, as well as other theoretical positions, have been continually challenged not only to describe behavior, but also to provide effective interventions for the populations whom they serve (D'Amato & Dean, 1989b; Gutkin & Reynolds, 1990). Prepackaged programs dealing with psycholinguistic or visual-motor training, and sensory integration training have attempted this, but typically failed to meet the demands of this challenge. Gradually, the field has acknowledged that the effective use of assessment procedures, including educational and psychological tests, is reliant upon a theoretical foundation, which allows the incorporation of information from multiple data sources and environments in such a manner as to increase the amount of effective and appropriate interventions generated.

A framework that is particularly useful is one reliant on an ecological approach. From this perspective, it is critical to evaluate several different aspects of clients' lives in order to develop a better understanding of their functioning within a variety of contexts. The purpose of this chapter is to examine the particular areas of attention and memory, learning processes (input and output), and the special aptitudes of sensory perception, sensory-motor integration, and language/communication to facilitate making informed assessment decisions. One must possess knowledge of the cognitive processes that these tests purport to measure to make judgments about the usefulness of any given instrument with the client's presenting issue. Furthermore, the clinician is provided with an introduction to the types of instruments used to measure memory, learning, and special aptitudes as well as a brief description of those instruments that have strong empirical support for use with children and adult populations. Any one of these areas, on its own, represents a very narrow picture of the overall functioning of an individual. However, when used in conjunction with a more thorough assessment, these areas can provide the missing pieces to the puzzle.

Client difficulties may be attributed not only to intra-individual characteristics, but also to the *domain* of functioning (e.g., social, vocational, educational); the *context* or environment in which the client is expected to function (e.g., job site, classroom, independent living); the requirements of particular assignments, jobs, or responsibilities, *task*; and the strategy used to teach or remediate a difficulty, *intervention* (Geil & D'Amato, 1996). Although the focus of this chapter is memory, learning, and special aptitudes, it may be helpful for the clinician to view these processes as key components within the conceptual framework presented in Figure 1. These areas represent extremely important aspects of psychological functioning and can help to complete the diagnostic picture of an individual by providing critical information to assist the clinician in accurate diagnosis and intervention planning.

4.09.1.1 Assessment Approaches

Before addressing the particular areas of concern, a brief discussion of the assessment process is warranted. Both quantitative and qualitative assessment procedures help to provide a breadth of information concerning individual functioning. A quantitative or product-oriented approach uses standard performance data to assess individuals within and across all the functional domains to be measured by comparing the findings to a normative group (D'Amato, Rothlisberg, & Rhodes, 1997; Dean, 1985a, 1985b; Lezak, 1995). This process detects whether the client's skills show a discrepancy when they are compared to other individuals performing within a normal range. Patterns of performance can also be carefully analyzed to determine the individual client's strengths and weaknesses. Data is usually considered in several ways: level of performance or current functioning (compared to normative standards); pattern of performance (uniqueness of strengths and weaknesses); right–left differences (comparing tests that evaluate both hemispheres including both sides of the body); pathognomic signs (indications of abnormal signs or brain damage); qualitative analysis (behavioral observations of problem solving); intervention

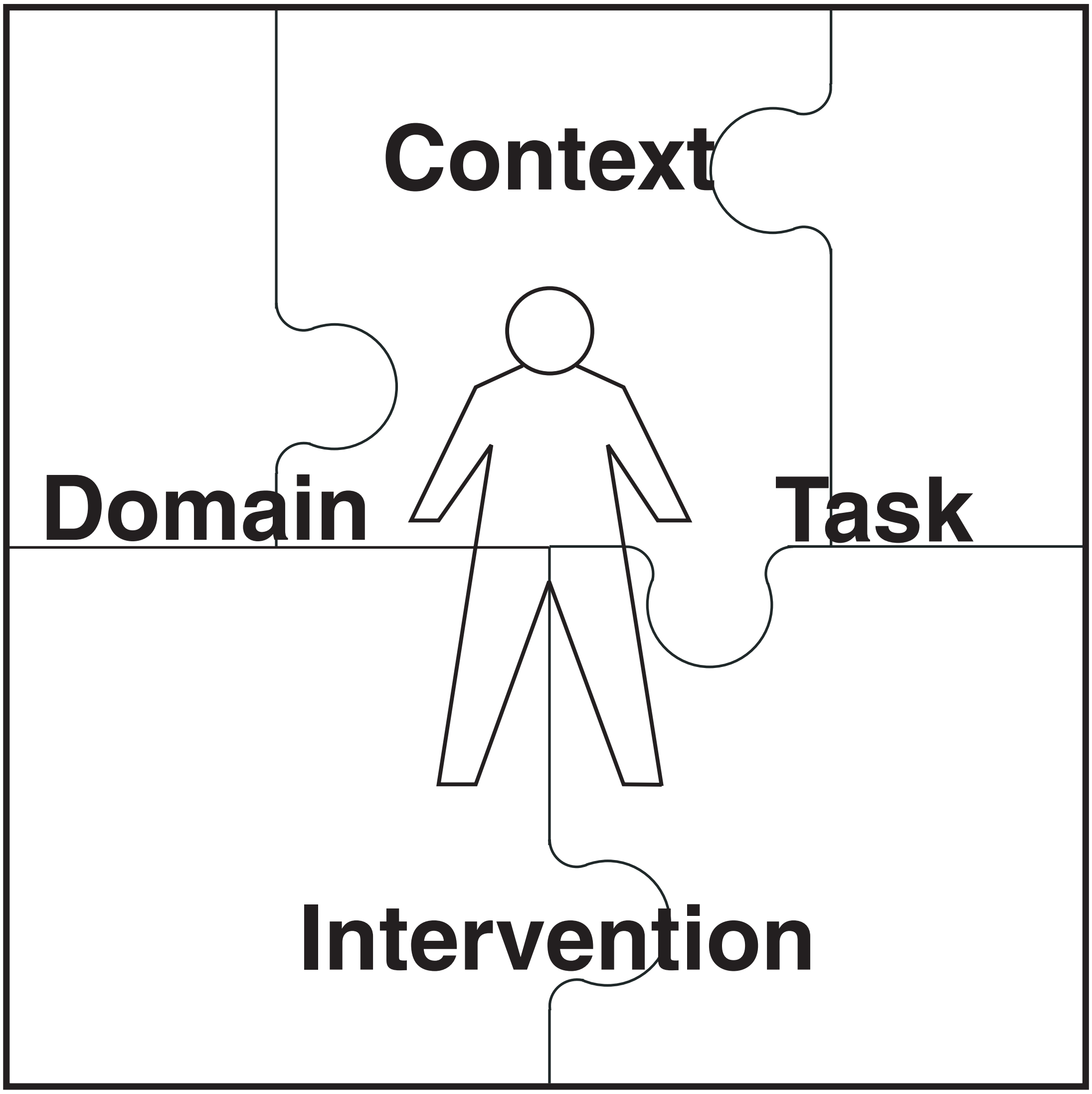

- Domain (social, vocational, educational)
- Context (job site, classroom, independent living)
- Task (assignment, job, responsibility)
- Intervention (remediation, counseling, training)

Figure 1 Conceptual framework of client functioning.

planning (recommendations for appropriate rehabilitation) (Hynd & Semrud-Clikeman, 1990; Jarvis & Barth, 1994; Reitan & Wolfson, 1985, 1993; Sattler, 1992; Selz, 1981).

Most proponents of a quantitative approach recommend a standard or fixed battery of tests. A fixed battery, such as a typical psychoeducational battery (i.e., Wechsler Adult Intelligence Scale-Revised [WAIS-R] or Wechsler Intelligence Scale for Children-3rd Edition [WISC-III], Minnesota Multiphasic Personality Inventory-2nd Edition, Bender Visual-Motor Gestalt Test, Woodcock–Johnson Psychoeducational Battery-Revised) or neuropsychological battery (e.g., *Halstead–Reitan Neuropsychological Test Battery*; Reitan & Wolfson, 1993) involve the same set of instruments for each individual tested (Hynd & Semrud-Clikeman, 1990; Hynd & Willis, 1988). A standard battery format insures that a broad array of appropriate tools is used to cover all significant domains and therefore provide documented results that may be interpreted with ease. In fact, a standard battery approach may be the best choice when and if potential litigation is an issue because this method offers a normative data base to which client profiles can be compared and contrasted (Guilmette & Giuliano, 1991; Reitan & Wolfson, 1995). Despite the apparent strengths of a standardized approach, it has been argued that when developing treatment options, the use of qualitative methods that explore the process of learning or behavior may be better suited than a purely quantitative or product-oriented approach (D'Amato, Rothlisberg, & Leu, in press). For example, if verbal instruction with verbal response is a strength for the client, a preference for left hemisphere processing might be entertained and interventions utilizing a verbal component could be tailored with that hypothesis in mind. Likewise, if a client demonstrated a strength in simultaneous processing of information, a global concept or visual chart could be introduced before presenting the individual skills necessary to accomplish the particular task.

A second strategy, the qualitative approach or process-oriented approach, uses informal procedures such as direct observation of particular skills to analyze the specific patterns and processes in order to understand better the intricacies of the client's psychological processes (Lezak, 1995). Practitioners utilize a client's individual pattern of responses or results to guide the assessment process. That is, if a client was observed to have difficulty with memory tasks, that particular area would be investigated in more detail through the use of additional measures of memory. A decision-making process (i.e., whether to explore an area further or move on to another area of functioning) occurs after each item and is based on clinical judgment. By employing this strategy, it is argued that a clinician is better able to understand the complexities of an individual's performance and focus on the impaired functional system (D'Amato, Rothlisberg, & Leu, in press; Golden, 1981; Luria, 1980). Unique and individualized sets of procedures, questions, or tasks shape the evaluation process and might include an individual case study approach consisting of a mental status exam, observation, and symptom checklists. From an educational perspective, a psychologist using a qualitative approach might gather information using work samples, classroom observations, or dynamic assessment strategies (e.g., Campione & Brown, 1987; Feuerstein, Rand, & Hoffman, 1979). While the major areas traditionally covered in a qualitative evaluation seem comprehensive (e.g., investigations of motor functions, expressive speech, writing, reading; see Hynd & Semrud-Clikeman, 1990), this view does not rely on standardized batteries or clear comparisons to normative populations. Instead, the selection of strategies utilized follows significant clinical patient–practitioner interactions. Although the flexibility and individualization apparent in this method is appealing, it requires a great deal of clinical experience to make accurate interpretations of behaviors, and problems with reliability and validity are ever-present (Lezak, 1995).

A third approach, and one which is likely used by the majority of clinicians, is the use of integrated data. Indeed, any time examiners note an examinee's reaction to a task, the response time involved, or any problem-solving strategies employed (e.g., rehearsal, verbal cuing), they are inferring the underlying processes being used (D'Amato, Rothlisberg, & Leu, in press; Taylor, 1988; Taylor & Fletcher, 1990). Because all individuals show a distinctive pattern of learning and behavioral characteristics, it is improbable that any given test, or even battery of tests, in isolation, can capture the range of skills exhibited by that individual. Furthermore, test scores that are interpreted without consideration to the context of the examination may be objective but are meaningless in their individual application (Lezak, 1995). Likewise, clinical observations unsupported by standardized and quantifiable testing may provide a rich picture of the client's current functioning but lack the comparability necessary for many diagnostic and planning decisions. Thus it is expected that most practitioners are integrated in their assessment practices, relying on both norm-referenced comparisons and qualitative procedures and observations to enrich their views of their clients. In fact, Lezak (1995) suggests that either method is incomplete without the other.

4.09.1.2 Evaluation of Domain Areas

Regardless of the position of the examiner along the quantitative–qualitative continuum, it is helpful to conceptualize an evaluation of domain areas, rather than simply focusing on tests or specific problem behaviors. The following domains are offered because of their importance to daily functioning and usefulness to intervention development in educational and vocational settings (Begali, 1994; D'Amato & Rothlisberg, 1992; D'Amato, Rothlisberg, & Leu, in press; Gaddes & Edgell, 1994). These domains include:

(i) intelligence/cognitive abilities,
(ii) personality/behavior/family information,
(iii) memory and attention,
(iv) learning processes,
(v) academic achievement,
(vi) sensory/perceptual systems,
(vii) motor functions, and
(viii) communication/language skills.

The areas of intelligence and personality assessment are covered in more depth in Chapters 8 and 12 of this volume. The remaining areas are divided into the subareas presented in Table 1 to provide a better understanding of the types of skills encompassed in each of these domains. Given the complexity of these areas, all should be considered both formally and informally. Direct observations and interviews with the client and family members are vital components in evaluating any individual's performance. The selection of tests utilized to evaluate these abilities will vary greatly depending on the unique needs of the individual, considered in tandem with the reason for referral.

Data on the functioning of these domains provides useful information for the clinical psychologist. As demonstrated in Table 2, a variety of testing instruments are appropriate in each of these areas. It should be noted that different authors have suggested various subsets of domains for analysis as well as recommending literally hundreds of other measures as appropriate for children and adults (Batchelor, 1996a; Begali, 1992; Dean & Gray, 1990; Gaddes & Edgell, 1994; Hynd & Willis, 1988; Lezak, 1995). Thus the instruments categorized in Table 2 represent only a sampling of available measures. The practitioner must take responsibility for carefully matching the individual with potential assessment options, after considering the distinct features of the instruments and the unique needs of the client.

Table 1 Subdomains of attention and memory, learning process, and special aptitudes.

Attention and memory
Attention
Concentration or vigilance
Visual memory
Verbal memory
Recall
Recognition
Short-term memory
Long-term memory
Learning processes (input and output)
Visual processing
Motoric processing
Auditory processing
Linguistic/verbal processing
Simultaneous processing
Sequential processing
Academic achievement
Sensory/perceptual
Visual
Auditory
Tactile-kinesthetic
Integrated
Motor functions
Strength
Speed
Coordination
Lateral preference
Sensory-motor integration
Communication/language skills
Receptive vocabulary
Expressive vocabulary
Speech/language
Written language

Source: Adapted from D'Amato & Rothlisberg (in press) and D'Amato, Rothlisberg & Rhodes (1997).

4.09.1.3 Intervention Approaches

The referral question for any client is rarely "how is this individual functioning today?"; instead the referral source is most often interested in the extent of decline following an injury or illness, the expected future performance in school or work settings, or how to maximize a client's potential given certain difficulties (e.g., head injury, learning disability; Long, 1996). Several decisions must be made in relation to the intervention strategy and will be reliant on the quality of the information provided by the assessment. Intervention may be conceptualized using one of three approaches: remediation (retraining a previously learned skill), compensation (learning to use other strengths to offset a lost skill), or a combination of both (D'Amato & Rothlisberg, 1996). In particular, it is critical to determine the level of intervention on which to focus one's efforts and the ideal combination of strategies that will work best with an individual. Rehabilitative efforts emphasize enabling clients to reach their goals in educational, vocational, social, and recreational settings despite difficulties related to their deficits. To reach this end, intervention strategies may focus on:

(i) remediating or retraining impaired cognitive processes (if there is a reason to believe that the process can be improved with practice),

(ii) helping the client to develop new skills to compensate for residual deficits,

(iii) creating classroom or workplace adaptations and other environmental compensations that permit effective performance despite residual deficits,

(iv) choosing instructional or therapeutic procedures that best fit the client's profile of strengths and weaknesses, and

(v) promoting improved metacognitive awareness of strengths and needs so that the client can become an active participant in selecting goals and interventions strategies (Ylvisaker, Szekeres, & Hartwick, 1994).

Table 2 Common instruments and procedures used to evaluate attention and memory, learning processes and special aptitudes.

Attention and memory

Test of Variables of Attention (TOVA™)
Visual Search and Attention Test
Tests of Memory and Learning
Wechsler Memory Scale-Revised
Wide range assessment of memory and learning

Learning processes: input and integration

Detroit Tests of Learning Aptitude-3
Swanson's Cognitive Processing Test
Children's Auditory Verbal Learning Test
Tactile Performance Test (Halstead–Reitan Battery)
Speech–Sounds Perception Test (Halstead–Reitan Battery)
Wisconsin Card Sort

Academic achievement: output

Woodcock–Johnson Psycho-educational Battery-Revised: Achievement
Peabody Individual Achievement Test-Revised
Wechsler Individual Achievement Test
Kaufman Test of Educational Achievement
Keymath-revised
Woodcock Reading Mastery Test-Revised
Test of Reading Comprehension-3
Test of Written Language-3

Sensory perception

Observations
Developmental history
Mental status examination
Motor-free Visual Perception Test
Vision and hearing screening

Motor (fine and gross)

Bender Visual-Motor Gestalt Test
Detroit Test of Learning Aptitude-3 (Motoric Composite)
Developmental Test of Visual-Motor Integration
Finger Oscillation Test
Grip Strength Test
K-ABC Nonverbal Scale (e.g., Hand Movements subtest)
McCarthy Scales of Children's Abilities (Motor Scale)
WISC-III and WAIS-R (Block Design, Object Assembly, Coding subtests)
Bruininks–Oseretsky Test of Motor Proficiency

Communication/language skills

Revised Token Test
Peabody Picture Vocabulary Test-Revised
Test of Adolescent Language
Test of Language Development-2 (Primary and Intermediate)
Test of Language Competence

Source: Adapted from D'Amato, Rothlisberg, & Rhodes (1997).

4.09.2 ASSESSMENT OF MEMORY

Memory is one of the most important cognitive functions to be assessed. It is a highly complex cognitive function that encompasses several relatively discrete stages: reception and registration of sensory stimuli, temporary short-term storage of information, storage of the information in a more permanent form (long-term memory), and recall and retrieval of previously stored information (Shiffrin & Atkinson, 1969; Taylor, Fletcher, & Satz, 1984). Functioning at each stage depends upon the integrity of the previous steps, with any interruption in the hierarchy having the potential to interfere with memory storage or retrieval. For example, difficulties with attention, which most closely relates to the first stage of memory, would obviously lead to problems in short- and long-term storage as well as later retrieval of the information. A further source of complexity in understanding and measuring this skill is the variety of theoretical approaches from which memory can be conceptualized, including information processing, neuropsychological, and behavioral perspectives.

Memory testing is very useful for assessing the possibility of organic disease, in helping to differentiate between organic and psychiatric disorders, and in determining the functional significance of a memory problem (Black & Strub, 1994). Most of the major neurobehavioral disorders such as dementia, confusional states, amnesia, material-specific memory/learning defects, and attentional dysfunction are those in which disturbances of memory and attention are the prominent clinical features (Hamsher, 1984). However, individuals with depression, post-traumatic stress disorder, dissociative disorders (e.g., dissociative amnesia, dissociative identity disorders) might also demonstrate attention and memory deficits (American Psychiatric Association, 1994). Memory skills represent a difficult area to address because of the variety of levels (e.g., working, short-term, long-term) and the potential implications of a deficit. So too, memory to some degree is modality specific; that is, for example, some individuals may have impaired verbal memory but intact visual memory. Thus, it is important to look at various components of memory rather than obtaining a simple global memory score. Tests that provide a single memory score offer a myopic and problematic view of the multifaceted quality of memory.

4.09.2.1 Attention

One of the key components to memory and learning is the ability to attend selectively to relevant information that we are presented with during the course of daily functioning. Attention is an extremely important basic function which refers to the client's ability to maintain awareness and to focus on a specified environmental stimulus, while screening out other stimuli that are potentially distracting (Black & Strub, 1994). Being able to attend has three major benefits for an individual: accuracy, speed, and maintenance of mental processing (LaBerge, 1995). Attention deficits appear as distractibility or impaired ability for focused behavior, regardless of the individual's intention (Lezak, 1995). Intact attention is a necessary condition of concentration which requires an individual to sustain attention over an extended period of time. Concentration problems may be due to a simple attentional disturbance, or to inability to maintain a purposeful attentional focus or, as is often the case, to both problems. This skill is important for adequate performance on any cognitive task, and can be impaired as a result of either an organic or emotional disorder (D'Amato, 1990; Dean, 1985a). Several psychological difficulties have been associated with attentional problems such as impulsivity, distractibility, and poor social judgment. Tests that require mental effort and persistence can measure an individual's ability to select, sustain, and shift attention (Slomka & Tarter, 1993). By comparing performance on various types of tasks, the practitioner is able to distinguish a global attention deficit from the more discrete, task-specific problems of concentration and tracking.

It is important to clarify the nature of an attention problem by observing people's general behavior as well as their performance on tests involving concentration. An interview with family members can provide important information about attentiveness and susceptibility to distraction. So too, formal or quantitative measures of attention and short-term memory can be derived from the Digit Span and Coding or Digit Symbol subtests of the WISC-III or WAIS-R tests. Some of the more specific, informal measures of attention and concentration might include observation, a digit span task, and a vigilance task as outlined by Strub and Black (1993). The individual is given orally a series of random letters with the letter "A" occurring with greater frequency than the other letters. The individual is instructed to signal whenever the targeted letter (i.e., A) is heard. The individual's performance is scored for errors, which are rarely made by those without attention or vigilance difficulties (Black & Strub, 1994). More formal measures such as the *Visual Search and Attention Test* (Trenerry, Crosson, DeBoe, & Leber, 1990) can be used for adults. This test purports to measure sustained attention and visual scanning. The test consists of four 60-second trials and is made up of four tasks which become increasingly complex. The respondent is required to cross out letters or symbols that match a target. Normative tables are provided and arranged in four 10-year age bands, an 18–19 year age band, and a 60+ age band, and the statistical properties appear to be adequate (Hooper, 1995). Based on how the stimuli are presented, these tasks can provide a measure of either visual or auditory vigilance.

Vigilance tests are often referred to as continuous performance tests (CPTs), which are automated tasks, now computer-administered, that purport to measure sustained attention (Greenberg & Waldman, 1993; Lassiter, D'Amato, Raggio, Whitten, & Bardos, 1994; Rosvold, Mirsky, Sarason, Bransome, & Beck, 1956). CPTs have become a popular tool for clinicians to measure attentional performance, response inhibition, and medication monitoring in both children and adults (Eliason & Richman, 1987; Lassiter et al., 1994). Many versions of CPT have been developed since the original but the basic methodology of these tasks remains fairly constant. Clients are presented with a variety of stimuli that are displayed for a short period of time, and are instructed to respond to a predefined "target" stimulus. A number of different indices can be recorded with these tasks including omissions errors (i.e., failing to detect target stimulus), commissions (i.e., responding to nontarget stimulus), and response times for correct detections (Greenberg & Waldman, 1993). Commission errors are considered to be indicative of impulsivity and omission errors are thought to denote inattention (Eliason & Richman, 1987; Lassiter et al., 1994). Examples of these types of tests include the CPT-2 (Lindgren & Lyon, 1983), the *Raggio Evaluation of Attention Deficit Disorder* (Raggio, 1991), and the *Test of Variables of Attention* (*TOVA*$^{\mathrm{TM}}$; Greenberg, 1993).

The *TOVA*$^{\mathrm{TM}}$ is a nonlanguage-based, visual continuous performance test. This test runs for 23 minutes on a fixed-interval schedule and presents two easily discriminated visual stimuli for 100 milliseconds every two seconds. It was designed for use in the diagnosis and monitoring of pharmacotherapy of children and adults with attention deficit disorders and can be used with individuals age five to adulthood. The test does not require right-left discrimination and has negligible practice effects. Recently, the authors of the *TOVA*$^{\mathrm{TM}}$ have created developmental norms for children aged 6 to 16 which are available for few other CPT versions (Greenberg & Waldman, 1993). This type of tool may also be useful in assisting the clinician in the differential diagnosis of children and adolescents experiencing externalizing problems (e.g., attention deficit disorder, oppositional defiant disorder, conduct disorder, and aggression) and/or learning disabilities (Eliason & Richman, 1987; Greenberg & Waldman, 1993).

Despite the technological advances and newly defined norms, CPTs present a quandary to practitioners because of the variety of attributes that the tests reportedly measure. Some see these tests as measuring attention and impulsivity (Klee & Garfinkel, 1983), educational achievement (Campbell, D'Amato, Raggio, & Stephens, 1991), behavior (Lassiter et al. 1994), general neuropsychological functioning (Halperin, Sharma, Greenblatt, & Schwartz, 1991), and information processing (Swanson, 1981). Given this variance, a practitioner is left with the question of how to interpret the test results of a particular client. While research supports many of these claims, different versions of the CPTs have been used in these studies, with different samples of children and adults. So too, the validity of CPTs have been related to material collected from teachers, parents, and peers, and from standardized intelligence, achievement, and personality tests. While it is obvious that CPTs measure issues critical to learning and memory, the specificity of these instruments remains unclear. In conclusion, Morris (1996) noted that many of the measures that purported specifically to measure sustained attention often measured other variables, and thus many of these tests have poor construct validation and may be more appropriately viewed as multidimensional in nature. As an alternative, Barkley (1996) advocates the use of more natural tasks to study attention in an individual. He concludes that CPT-type tasks are unrelated to our daily functioning and thus, an individual's performance on such tasks is irrelevant. In response to this concern, several investigators have reportedly used television viewing, performance on classroom tasks, video games, and driving performance as a means of studying attention and its deficits in various groups of children and young adults (Barkley, 1996).

4.09.2.2 Short-term and Long-term Memory

As already noted, each component of the memory process is reliant upon the previous steps. If information in sensory storage undergoes additional processing, it becomes a more lasting memory (short-term or long-term). Furthermore, these different memory functions must be systematically reviewed through visual and aural modalities using both recall and recognition tasks. Lezak (1995) suggests that at a minimum, the memory examination should include: *immediate retention* tasks, including short-term memory with interference; *learning* in terms of extent of recent memory, learning capacity, and how well newly learned material is retained; and efficiency of *retrieval* of both recently learned and long-stored information (i.e., remote memory).

Informal methods of assessment include tests of immediate recall such as digit repetition and/or sentence repetition, interviewing for information from remote memory (e.g., "where were you born?"), and new learning ability (e.g., immediate recall for a verbal story, asking the individual to remember four unrelated words for a span of 5, 10, and 30 minutes). During this last task, the examiner can provide recognition cues if the individual is having difficulty remembering the words. It is expected that those without difficulties will remember all words, while those with brain damage might be expected to remember one (Black & Strub, 1994). For aphasic clients or those with other speech or language problems, an informal measure of visual memory can be completed by hiding five objects around the interview room as the client names each item as it is hidden. After 10 minutes, the client is asked for name and location of each item. Reportedly, both normal and lower IQ clients should be able to find all five objects, with slightly lower performance for older patients (approximately four objects) (Black & Strub, 1994; Simpson, Black, & Strub, 1986). These memory tasks should be supplemented with observations and interviews with family members. So too, if an ability measure such as the WISC-III or WAIS-R is administered, performance on Digit Span can provide information on immediate verbal retention and the information subtest can be an indicator of the extent of remote memory in an individual.

To complete an assessment of the major dimensions of memory, Lezak (1995) has suggested including:

(i) a test of configural recall and attention such as the visual reproduction subtest on the *Wechsler Memory Scale* (Wechsler, 1987) or the *Benton Visual Retention Test* (Benton-Sivan, 1992);

(ii) a paragraph for recall to examine learning and retention of meaningful verbal material; and

(iii) a test of learning ability that gives a learning curve and includes a recognition trial, such as Rey's Auditory-Verbal Learning Test (for review see Lezak, 1995).

These techniques should be integrated into the general clinical interview to create a varied testing format, to enable the practitioner to use nonmemory tasks as interference activities, and to reduce stress in those clients who have memory impairments and are concerned about their abilities (Black & Strub, 1994).

There are numerous formal instruments available which measure different dimensions of memory. For children and adolescents, the *Wide Range Assessment of Memory and Learning* (Sheslow & Adams, 1990), and the *Test of Memory and Learning* (*TOMAL*; Reynolds & Bigler, 1994) can be used to evaluate individual strengths and weaknesses in the areas of memory and attention. In particular, the *TOMAL* represents a reliable, empirically sound measure for children and adolescents. The *TOMAL* consists of four core indexes comprising Verbal Memory, Nonverbal Memory, Composite Memory, and Delayed Recall. Supplementary indexes for Learning, Attention and Concentration, Sequential Memory, Free Recall, and Associative Recall are also provided. Subtests include Memory for Stories, Facial Memory, Word Selective Reminding, Visual Selective Reminding, Object Recall, Abstract Visual Memory, Digits Forward, Visual Sequential Memory, Paired Recall, Memory-for-Location, Manual Imitation, Letters Forward, Digits Backward, and Letters Backward. The *TOMAL* was standardized for children aged 5 to 19.

The *TOMAL* boasts many unique features, including a great variety of memory indexes (Reynolds & Bigler, 1994). While some of the subtests appear similar to other memory measures, some unique features of this test include a learning index where teaching is permissible, a sequential memory index, and an attention and concentration index. Delayed recall subtests are also available and are offered as an evaluation of forgetting or memory decay. It is possible to compare the examinee's own personal learning curve with a standardized learning curve. The test is easy to administer and generally user-friendly. Its psychometric properties appear to be well-developed. In the *TOMAL* subtests, 63% of the reliability coefficients are at or exceed 0.9, 31% are between 0.8 and 0.89, and only 6% fall below 0.8. Test–retest coefficients range from 0.71 to 0.91. Support for the validity of this instrument was determined through indices of content validity, construct validity (e.g., factor analytic studies), and criterion-related validity.

Assessment of memory dysfunction in adults is easier than in children because their period of rapid intellectual, academic, and physical (including neurological) development has ended (Reynolds & Bigler, 1994). In adults, memory dysfunction is associated with a variety of well-defined disorders, and in many individuals is one of the earliest and key symptoms such as in Korsakoff's disease and various other dementias including Alzheimer's disease. Because of the key role of evaluating memory in the clinical setting, there are a number of instruments designed for memory assessment in older populations including the *Doors and People: A Test of Visual and Verbal Recall and Recognition* (Baddeley, Emslie, & Nimmo-Smith, 1994), the *Memory Assessment Scales* (*MAS*; Williams, 1991), and the *Wechsler Memory Scale-Revised* (*WMS-R*; Wechsler, 1987).

The *WMS-R* (Wechsler, 1987) provides an extensive measure of several dimensions of memory. It consists of eight short-term memory tests, four delayed-recall subtests, and a brief screening measure of mental status (i.e., information and orientation questions). The eight short-term memory tests yield four composite scores: Verbal Memory, Visual Memory, Total General Memory, and Attention/Concentration. The delayed-recall measures can be combined to derive a fifth composite score, Delayed Recall. The test is intended for use for individuals ranging in age from 16 to 74 and requires approximately 50 minutes to administer. The psychometric properties of the *WMS-R* are questionable in terms of low reliability coefficients for the composite scores (average $r = 0.74$), but provides stronger support for the General Memory and Attention/Concentration (average $r = 0.81$) scores. Although the *WMS-R* demonstrated satisfactory discrimination power between various clinical groups, factor analyses supported a two-factor rather than the hypothesized five-factor model. Huebner (1992) concluded that this instrument must be used cautiously in making clinical decisions about individuals and interpretation should be restricted to General Memory and Attention/Concentration ability.

For adolescents and adults, the *MAS* (Williams, 1991) also provides a valid, reliable, and comprehensive measure of memory functioning. The *MAS* was standardized for use with adults aged 18 to 90. The major functions measured by the *MAS* include: verbal and nonverbal learning and immediate memory; verbal and nonverbal attention, concentration, and short-term memory; and memory for verbal and nonverbal material following delay. In addition, measures of recognition, intrusions during verbal learning recall, and retrieval strategies are also available. The test consists of 12 subtests based on seven memory tasks.

Five of the subtests assess the retention of information learned in a subtest administered earlier in the sequence. Total testing time is approximately one hour. Test-retest reliability for the *MAS* was estimated using generalizability coefficients and these correlations averaged 0.85 for the subtests, 0.9 for the summary scales (i.e., Short-Term Memory, Verbal Memory, and Visual Memory), and 0.95 for the global memory scale. The validity of the *MAS* was established using three types of studies: convergent and discriminant validity, factorial validity, and group differentiation. Despite these strengths, Berk (1995) concluded that clinicians should use caution in interpreting the scores until some technical problems (e.g., inadequate samples, lack of evidence for content validity) can be corrected.

4.09.2.3 Memory: Implications for Intervention

Because memory is multifaceted, interventions in the memory domain must also be multidimensional. Interventions may be divided several ways; those involving language, those that are nonverbal, those requiring long, short, or intermediate memory, and those that use a combined approach to aid in retention (Gaddes & Edgell, 1994; Lezak, 1995). Strategy selection depends on accessing the strengths of the clients or, in the case of injury or disease, accessing those parts of the brain which have been least impacted. For example, learners who have difficulty with nonverbal memory tasks but have retained verbal skills may benefit from memory interventions that use language. Mnemonic devices may be used to assist in recall of information if a series of problem-solving steps is required (Mastropieri & Scruggs, 1989). For those with more difficulty remembering, some simple techniques such as writing all meetings in an appointment book or using grocery shopping lists and daily "to do" lists are practical.

4.09.3 ASSESSMENT OF LEARNING

Memory is a ubiquitous component of daily life and is fundamental to the process of learning. One must be able to remember in order to demonstrate learning. The classic definition of learning describes it as changes in behavior as a result of experience. Some have even considered this definition of learning as also defining memory (e.g., see Kolb & Whishaw, 1990). Despite this relatively simple definition, the learning process itself defies easy explanation. Learning can be approached from a neuropsychological (e.g., planning, attention, simultaneous, successive [PASS] model; Das, Naglieri, & Kirby, 1994), behavioral (e.g., classical and operant conditioning; Baldwin & Baldwin, 1986), cognitive (e.g., information processing; Pressley & Levin, 1983), or social (e.g., social learning and modeling; Bandura, 1977) perspectives, to name but a few. Furthermore, the distinction drawn between measures of memory and measures of learning is tenuous at best since all instruments that evaluate an individual's learning process will automatically include aspects of memory functioning. Because learning takes center stage as one of our functions of daily living, it represents a critical area to evaluate. By fully evaluating several components of learning, clinicians can determine where the process is breaking down and provide recommendations for rehabilitation. An understanding of how a client best learns also has important implications for the types of therapeutic strategies that will most likely be successful. For example, if an adolescent has difficulty processing and remembering auditory information, talk therapy may not be the best approach. Supplementing discussion with role play, videos, and other visual cues may be necessary to facilitate the client's acquisition of new knowledge.

Difficulties in learning can be attributed to a number of disorders, including the general category of learning disorders (e.g., dyslexia, dyscalculia), traumatic brain injuries, drug and alcohol abuse, and medical disorders (e.g., strokes, Alzheimer's disease). Indeed, all of the variables that can affect attention and memory will also impact learning. Furthermore, certain chronic medical conditions can play a role in learning difficulties. For example, childhood diabetes is associated with subtle problems with respect to visuospatial and visuomotor processing (e.g., Rovet, Ehrlich, & Hoppe, 1988), verbal abilities (Kovacs, Goldston, & Ivengar, 1992), and memory and attention problems, which translate into increased risk for difficulties in academic achievement, (Kovacs et al., 1992; Rovet, Ehrlich, Czuchta & Akler, 1993). So too, sickle cell anemia, an inherited disorder in people of African descent, often produces some subtle cognitive impairments that can affect school achievement negatively (Brown, Armstrong, & Eckman, 1993). Among individuals with traumatic brain injuries, learning process problems may be reflected as an uncertainty as to whether a concept has been learned or not (Cohen, 1991).

4.09.3.1 Models of Learning

Information-processing theories have proved extremely useful in conceptualizing learning because this model can be applied to any given cognitive task and allows the practitioner to specify where the learning process is breaking down. Silver (1993) proposed an information-processing model based on four steps: input (how information from the sense organs enters the brain), integration (interpreting and processing the information), storage (storing the information for later retrieval), and output (expressing information via language or muscle activity). Learning is reliant upon each of the first three steps and is observed or inferred from the fourth step. Other models of information processing highlight the importance of the working memory in skill acquisition and learning (Baddeley, 1986; Just & Carpenter, 1992; Swanson, 1995). Working memory has traditionally been defined as a system of limited capacity for the temporary maintenance and manipulation of information (e.g., Baddeley, 1986; Just & Carpenter, 1992) and most closely corresponds to the integration step in Silver's model. Tasks that measure working memory are those that require the client to remember a small amount of material for a short time while simultaneously carrying out further operations. In daily life, these tasks might include remembering a person's address while listening to instructions about how to reach a specific destination (Swanson, 1995). When viewed from this perspective, working memory differs from the related concept of short-term memory which is typically described as remembering small amounts of material and reproducing it without integrating or transforming the information in any way (e.g., repeating back a series of numbers) (Cantor, Engle, & Hamilton, 1991; Just & Carpenter, 1992). Working memory appears to be extremely important to an individual's ability to learn, and in adult samples has correlations of 0.55–0.92 with reading and intelligence measures (e.g., Daneman & Carpenter, 1980; Kyllonen & Christal, 1990).

In an effort to promote the notion that input and integration of stimuli can impact subsequent learning, Cronbach and Snow (1977) have advanced a theory suggesting that some types of individuals might benefit from one form of treatment, whereas others might benefit from another type of treatment: an aptitude by treatment interaction (ATI). Many researchers and educators alike believe that matching learner characteristics with treatment approaches can enhance learning (e.g., Cronbach & Snow, 1977; Resnick, 1976; Reynolds, 1981b). However, subsequent studies have demonstrated little support for this theory (e.g., Arter & Jenkins, 1977; Tarver & Dawson, 1978). Initially, theories of input examined learner modalities (e.g., visual, auditory, kinesthetic), which were later deemed to be too simplistic (Arter & Jenkins, 1977; Kaufman, 1994; Tarver & Dawson, 1978). More recently, neuropsychological models have been applied to ATIs and offer promise for identifying aptitudes and prescribing treatments (D'Amato, 1990; Hartlage & Telzrow, 1983). One of the major techniques that Cronbach and Snow (1977) suggested for matching treatment approaches with learner aptitudes was "capitalization of strengths." Our increasing knowledge of how the brain functions allows clinicians to obtain a more detailed understanding of how a client learns new information. For example, although the cerebral hemispheres act in concert, the right hemisphere seems to be specialized for holistic, spatial, and/or nonverbal reasoning whereas the left shows a preference for verbal, serial, and/or analytic type tasks (Gaddes & Edgell, 1994; Lezak, 1995; Reynolds, 1981a; Walsh, 1978). Similarly, models of cognitive processing have been proposed that agree with the specialization of how scientists think the brain processes information; some have called this preferential processing styles (D'Amato, 1990). For example, simultaneous processing ability has been affiliated with the right hemisphere because of its holistic nature; it deals with the synthesis of parts into wholes and is often implicitly spatial (Das, Kirby, & Jarman, 1979). In contrast, the left hemisphere processes information using a more successive/sequential method, considering serial or temporal order of input (Dean, 1984, 1986). Models of brain organization have also been proposed that attempt to explain the diversity and complexity of behavior.

An expansion of the hemispheric specialization approach is offered in the planning, attention, simultaneous, successive (PASS) cognitive processing model (Das et al., 1994) which proposes four processing components. This model is based on the neuropsychological model of Luria (1970, 1973, 1980; Reynolds, 1981a) and presents a comprehensive theoretical model by which cognitive processes can be examined. On the basis of his clinical investigations with brain-injured patients, Luria (1973) suggested that there are three functional units that provide three classes of cognitive processes (i.e., memory, conceptual, and perceptual) responsible for all mental activity. Figure 2 provides a graphic presentation of the PASS model of cognitive processing. The functional units work in concert to produce behavior and provide arousal and attentional (first unit), simultaneous-successive (second unit), and planning (third unit) cognitive processes. The PASS model separates the second unit into two individual processes (i.e., simultaneous and sequential). Instruments can be used to measure individual strengths in these different styles of processing.

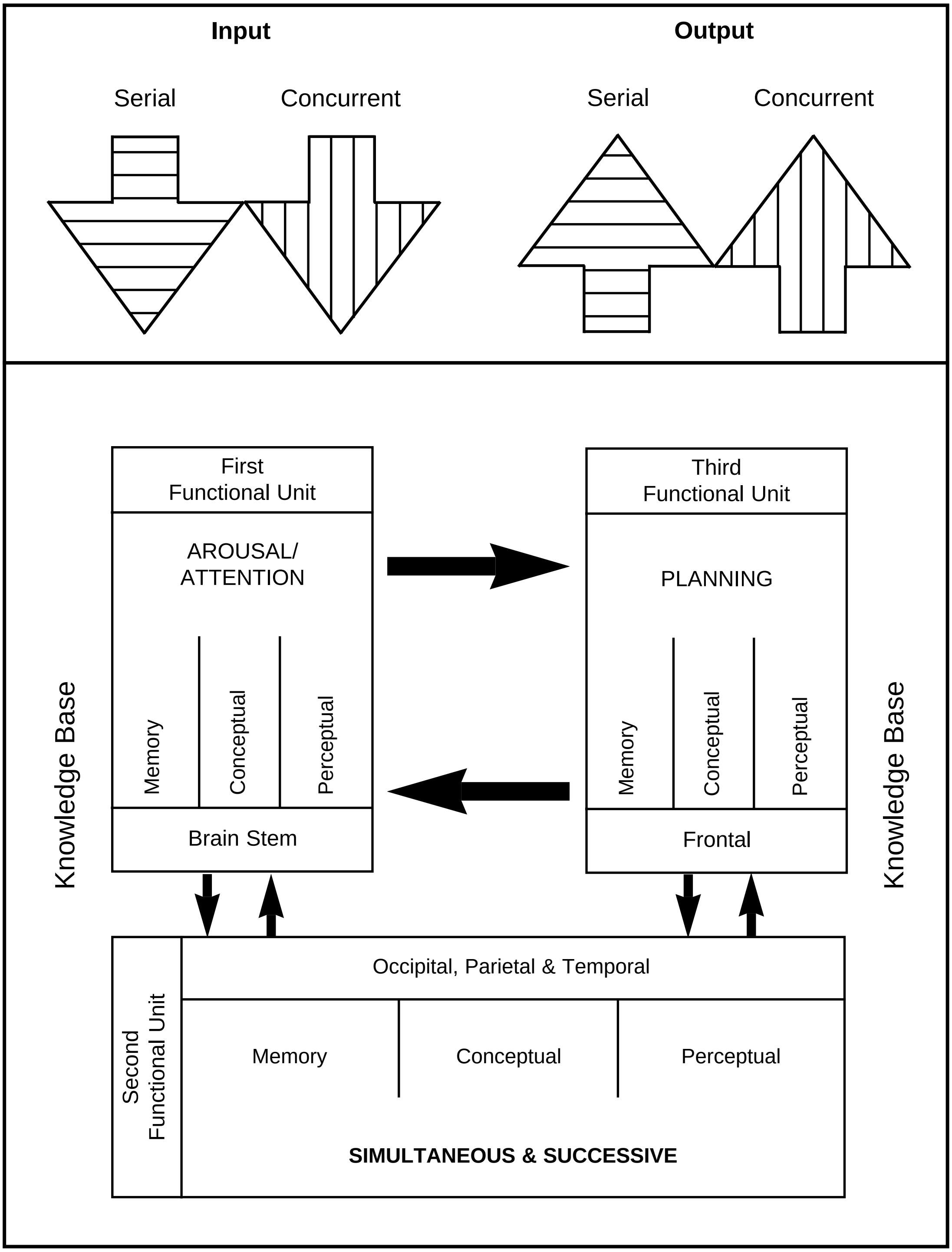
Input
Output
Serial
Concurrent
Serial
Concurrent
Knowledge Base
First
Functional Unit
AROUSAL/
ATTENTION
Memory
Conceptual
Perceptual
Brain Stem
Third
Functional Unit
PLANNING
Memory
Conceptual
Perceptual
Frontal
Knowledge Base
Second
Functional Unit
Occipital, Parietal & Temporal
Memory
Conceptual
Perceptual
SIMULTANEOUS & SUCCESSIVE

Figure 2 PASS model of cognitive processing. (*Assessment of Cognitive Processes: The Pass Theory of Intelligence* (p. 21), by J. P. Das, J. A. Naglieri, and J. R. Kirby, 1994, New York: Allyn & Bacon. Copyright 1994, by Allyn & Bacon. Reprinted with permission.)

Knowledge of the brain and theories governing information processing can determine the types of data collected during the assessment phase. For example, instead of simply observing whether the individual was successful at a task or set of measures, the practitioner looks beyond the product to determine the influence of related factors. These factors can include the nature of the stimuli used (visual, verbal, tactile), the method of presentation (visual, verbal, concrete, social), the type of response desired (verbal, motor, constructional), and the response time allowed (timed, untimed; Cooley & Morris, 1990). Other researchers have advocated a move to an even more intense examination of processing through the use of dynamic assessment strategies (Campione & Brown, 1987; Feuerstein et al., 1979; Palincsar, Brown, & Campione, 1991). Theoretically, this strategy allows the examiner to obtain information about the client's responsiveness to hints or probes, and thus elicits processing potential (Swanson, 1995). When an examinee is having difficulty, the examiner attempts to move the individual from failure to success by modifying the format, providing more trials, providing information on successful strategies, or offering increasingly more direct cues, hints, or prompts (Swanson, 1995). This approach allows the examiner an opportunity to evaluate performance change in the examinee with and without assistance. However, there is little if any standardized information available on this technique and it has been criticized for its clinical nature and poor reliability (e.g., Palincsar et al., 1991).

4.09.3.2 Learning Processes: Input and Integration

Generally, three approaches have been utilized when evaluating how individuals preferentially process information. The first approach, seen as the traditional approach, employs established measures (such as the WISC-III) with the practitioner seeking to understand information processing through an analysis of common test results such as reviewing global scores, subtests, and clusters of subtests (Kaufman, 1990, 1994). For example, a pattern of strengths on the Picture Completion, Block Design, Object Assembly and corresponding weaknesses in Picture Arrangement and Coding might suggest meaningful differences in a client's mental processing style (i.e., right hemispheric functioning vs. left hemispheric functioning from cerebral specialization theory, or simultaneous vs. successive coding from Luria theory). The second view of information processing, considered the informal approach, considers observations, checklists, and learning style inventories to understand how individuals learn. From this view, individuals who seem to profit most from visual clues may be seen as visual learners, and might be taught utilizing overheads, visual diagrams, and worksheets. The final approach to understanding processing stems from the administration and analysis of the many unique measures that have been offered as learning style or processing tests. This approach is seen as a nontraditional test approach (D'Amato, Rothlisberg, & Rhodes, 1997). These specialized measures of performance in learning do not fall neatly within the traditional domains of intelligence, achievement, or neuropsychological processing. These tests, including the *Detroit Tests of Learning Aptitude-3* (*DTLA-3*; Hammill, 1991), the *Swanson Cognitive Processing Test* (*S-CPT*; Swanson, 1996), the *Children's Auditory Verbal Learning Test-2* (Talley, 1993), and others can offer valuable information concerning how individuals attend to and deal with new information. While practitioners have used these instruments to document clients' strengths and weaknesses, diagnose problems, and chart the course of disorders, these instruments offer more practical information concerning rehabilitation or program planning than for diagnostic activities.

Neuropsychological tests have also been seen by some to evaluate variables related to learning processes. In fact, the *Halstead Reitan Neuropsychological Test Battery* (Reitan & Wolfson, 1985, 1993) reportedly measures problem solving, tactual discrimination, sensory recognition, spatial memory, verbal-auditory discrimination, attention, nonverbal auditory discrimination, psychomotor speed, and manual dexterity as well as several other skills (D'Amato, 1990; Dean, 1985a, 1985b, 1986; Lezak, 1995). Individuals interested in a neuropsychological approach to processing should consult some of the recommended references and Chapters 10 and 11, this volume.

The area of learning processing is more difficult to evaluate and is often subsumed within the intelligence or achievement domains. One measure that has a long history in the evaluation of processing styles is the *DTLA-3*. This instrument was designed for use with individuals aged 6 to 17. More recently, two other versions of this test, the *Detroit Tests of Learning Aptitude-Adult* (Hammill & Bryant, 1991a) and *Detroit Test of Learning Aptitude-Primary* (2nd ed.; Hammill & Bryant, 1991b) have expanded the usefulness of this instrument to include individuals from age 2 to 79. The *DTLA-3* consists of 11 subtests comprising: Word Opposites, Design Sequences, Sentence Imitations, Reversed Letters, Story Construction, Design Reproduction, Basic Information, Symbolic Relations, Word Sequences, Story Sequences, Picture Fragments; and 16 composite scores: General Mental Ability Composite, Optimal Level Composite, Domain Composites (Verbal, Nonverbal, Attention-Enhanced, Attention-Reduced, Motor-Enhanced, Motor-Reduced), Theoretical Composites (Fluid Intelligence, Crystallized Intelligence, Associative Level, Cognitive Level, Simultaneous Processing, Successive Processing, Verbal Scale, and

Performance Scale). The testing time is estimated to vary from 50 minutes to two hours. The internal consistency reliabilities for the subtests are sufficiently high; however, the data on stability are limited. It was also noted that the factor analysis does not support the construct validity of the different composites. In fact, only four factors (one being a residual or difficult to interpret category) were identified in the manual. Despite these concerns, Poteat (1995) notes that the *DTLA-3* can be recommended as an adjunct to some of the better developed measures of intelligence and it provides some potentially valuable information about diverse abilities. The *DTLA* has been especially helpful when evaluating children who suffer from learning disabilities or traumatic brain injuries. The *Detroit Tests of Learning Aptitude-Adult* (Hammill & Bryant, 1991a) comprises 12 subtests and 16 composites and measures areas similar to the *DTLA-3*. Internal consistency reliability of all scores approximates 0.9 for all ages. This instrument represents a useful tool in practice because of the type of information it can provide regarding client cognitive functioning in relation to learning new information.

A very recent contribution to the field of cognitive processing is the *S-CPT* (Swanson, 1996) which purports to measure different aspects of intellectual functioning and information processing potential. The battery, designed for use with persons age five to adulthood, draws from the work on information processing theory and dynamic assessment. The subtests in this measure are as follows: Rhyming Words, Visual Matrix, Auditory Digit Sequencing, Mapping and Directions, Story Retelling, Picture Sequencing, Phrase Recall, Spatial Organization, Semantic Association, Semantic Categorization, and Nonverbal Sequencing. This standardized test battery can be administered in an abbreviated form (five subtests) or in a complete form under traditional or interactive testing conditions. Normative data for the *S-CPT* were gathered on 1611 children and adults. The author reports high levels of internal reliability and high construct and criterion-related validity (Swanson, 1995). This instrument may offer a promising alternative to product-oriented evaluation strategies while still allowing for normative comparison.

4.09.3.3 Academic Achievement: Output

It is likely that those individuals who experience difficulties in processing or learning will display academic difficulties as well. Although some might hold that there is little difference in the measure of ability and the measure of achievement (e.g., Anastasi, 1988; Dean, 1977, 1983), it would seem that the operationalization of the two areas allows for a comparison of more generic problem-solving and verbal tasks to those directly involved in scholastic performance. Thus, a measure of ability may be conceived of as attempting to address the concept of underlying skills or capacities, whereas the measure of achievement is tied to the notion of the individual's proficiency in applying that ability in a functional way to real world skills (e.g., academics). A measure of academic achievement can help provide information as to the degree of impairment experienced by individuals, especially among children and adolescents. New learning, however, is not isolated to the school years. Adults required to learn new skills as part of job training, vocational rehabilitation, or after brain injuries are all placed in very real learning situations. It is critical to have an understanding of the client's basic skills in order to facilitate vocational, academic, and intervention decision making.

Assessment of academic achievement can occur through a blend of formal and informal measures as well. For example, in a school setting, reviewing student clients' work samples, interviewing the students and teacher about their learning and the classroom, and classroom observations can provide essential information. So too, curriculum-based measurement, where informal reading, writing, and math probes (Shinn, 1989) are obtained to determine the clients' current level of functioning and progress during intervention phases, are particularly useful for monitoring the effectiveness of treatment approaches to learning difficulties (Fuchs, 1994). There are also several types of norm-referenced instruments that are available, which, because of the availability of a standard normative base, permit comparison across a wide variety of curricular contexts (D'Amato, Rothlisberg, & Rhodes, 1997). Some of these instruments measure a particular area of achievement such as math or reading (e.g., *Keymath Revised*, Connolly, 1988; *Test of Reading Comprehension-3*, Brown, Hammill, & Wiederholt, 1995; *Test of Written Language-3*, Hammill & Larsen, 1996) while others provide a broad-based screening of a number of academic areas (e.g., *Peabody Individual Achievement Test-Revised* [*PIAT-R*], Markwardt, 1989; *Woodcock–Johnson Psychoeducational Battery-Revised* [*WJPB-R*], Woodcock & Johnson, 1989). These broad-based tests all have a similar organizational structure. For example, measures in a particular area, such as reading, are typically divided into basic skill areas (e.g., reading decoding) and some form of applied skill area (e.g., reading comprehension) so that variations in the aspects of the academic tasks can be noted. The difference between measures often lies in the method by which they obtain their information (e.g., whether visual-motor or oral responses are required); that is, whether they require the examinee to indicate the response through nonverbal (e.g., pointing) or verbal output.

The achievement test that is designed for the broadest range of individuals is the *WJPB-R* with norms ranging from 2 to 95 years of age. The *WJPB-R* consists of both a cognitive and an achievement component. The tests of achievement are divided into a standard battery consisting of four broad areas: Reading (Letter-Word Identification, Passage Comprehension), Mathematics (Calculations, Applied Problems), Written Language (Dictation, Writing Samples), and Broad Knowledge (Science, Social Studies, Humanities). A supplemental battery is also available to expand the standard battery coverage. It includes Word Attack, Reading Vocabulary, Quantitative Concepts, Proofing, and Writing Fluency. Employing one or more of the supplemental subtests gives the examiner the option of computing additional areas of achievement such as Basic Reading Skills and Reading Comprehension which is consistent with the language of the Individuals with Disabilities Educational Act of 1990 and some state legislative guidelines for identifying specific areas of learning disability. This instrument is statistically sound and ample amounts of research have been conducted and support the use of this test.

Another measure of achievement, the *PIAT-R* (Markwardt, 1989) has also been supported as a well-developed and psychometrically sound instrument (Williams & Vincent, 1991). It consists of five subtest scores (General Information, Reading Recognition, Reading Comprehension, Mathematics, Spelling) that are provided in addition to the Total Reading and Total Test scores. A Written Expression and optional Written Language score are also available. The *PIAT-R* was normed for individuals aged 5–18 years. It is different from other tests in that it includes a larger pictorial component in its item types, letting children avoid the need for verbal reply, and instead expecting them to point at the correct answer (out of four) for reading, spelling, and mathematics items. Since the task demands for recognition of information do not appear to be the same as for recall, this response format may aid children with retrieval difficulties or those that have developed some background knowledge of the area in question. It should be noted, though, that this response-type advantage may not give a good indication of the expectations for student performance in the classroom where recall and more integrated answers are the norm (D'Amato, Rothlisberg, & Rhodes, 1997).

4.09.3.4 Learning: Implications for Intervention

A number of authors have related how knowledge of the way individuals process information can contribute to the development of treatment based on neuropsychological processes (D'Amato, 1990; Reynolds, 1981b, 1986; Telzrow, 1985). For example, when learning how to read, individuals who display a simultaneous/visual spatial strength in processing might benefit from being taught using a whole word approach whereas individuals who display a strength in sequential/auditory processing can be taught using a phonetic approach (Whitten, D'Amato, & Chittooran, 1992). For both children and adults, cognitive rehabilitation is an emerging discipline which includes the retraining or use of compensatory strategies in thinking and problem-solving skills (Wedding, Horton, & Webster, 1986). Cognitive retraining can include assistance in strategy development for attention and concentration, memory, language, perceptual and cognitive deficits, and social behavior. Thus, the term cognitive retraining encompasses all areas of functioning that may have been negatively impacted by neuropsychological disorders or traumatic brain injury (D'Amato, Rothlisberg, & Leu, in press; Gray & Dean, 1989). Assisting learners with cognitive remediation or compensation often includes the use of metacognitive strategies. Metacognition includes analyzing the processes an individual uses to generate an idea or thought. By receiving assistance in breaking down problems and understanding the processes needed to solve problems, clients may learn how to generalize the process to many problem types and improve overall learning and functioning. Although cognitive retraining is time consuming, the generalizability of the strategies has been seen as appropriate to many settings (Gray & Dean, 1989; Kavale, Forness, & Bender, 1988).

4.09.4 ASSESSMENT OF SPECIAL APTITUDES

In determining the basis for a client's difficulty, it is critical to explore the building blocks of memory and learning to obtain an understanding of how the individual processes information (sensory input). For instance, sensory and perceptual skills are essential to receiving stimuli from the environment and making sense of what is received. So too, a clinician must examine the output or production that the client demonstrates in response to stimuli via action (e.g., motor skills) or communication (e.g., spoken language, writing). That is, clients may understand a task, but, because of integration difficulties or language impairments, be unable to demonstrate their knowledge. For example, the reproduction of a visual stimulus in response to a request involves both perceptual discrimination and fine motor development, as well as the ability to integrate visual, tactile, and auditory skills. Therefore, inadequate performance in copying geometric designs developed to assess these skills may stem from: a misperception, or faulty interpretation of the input information; problems in executing the fine motor response, or output; and/or difficulties integrating the input and output, otherwise known as integrative or central processing difficulties. By evaluating the domains of sensory perception, sensory-motor integration, and communication/language, the practitioner is in a better position to understand the client's ability to receive information adequately, integrate these basic skills, and demonstrate the products of memory and learning processes.

4.09.4.1 Sensory Perception

Perception of stimuli is a complex process involving many different aspects of brain functioning (Lezak, 1983). Typically, perception includes recognizing features and relationships among features. It is affected by context (figure–ground) and intensity, duration, significance, and familiarity of the stimuli (Ylvisaker, Szekeres, & Hartwick, 1994). Sensory perception skills are vital to an individual's understanding and response to the environment because they form the basis of each individual's interaction with the world (D'Amato, Rothlisberg, & Rhodes, 1997; Lezak, 1995). Difficulties may manifest themselves in the individual's ability to use information gained through the senses. For example, a client may be able to hear sounds well, but have trouble understanding what is heard (auditory processing). Likewise, a client may be able to see words clearly but have problems reproducing them when writing (visual-motor difficulties). Sensory perception tasks often form the foundation for the later performance of higher order cognitive skills. Without the ability to accurately sense and perceive cues from the environment, the learner is placed in the position of trying to decode a message when the code is scrambled and often changing.

In the assessment of sensory perception, it is important to evaluate visual, auditory, and haptic (tactile) functions. For children and older adults it is especially important that actual sensory deficits have been ruled out through the administration of a thorough vision and hearing test. If these senses appear to be intact, an in-depth evaluation of functioning is warranted. Sensory perception can be evaluated informally through clinical observations, formally through standardized tests, or via other methods of data collection. However, at times these strategies may prove inconclusive regarding the etiology of performance difficulties and more formal assessment is necessary to evaluate a client's functioning.

Several instruments are available to measure a client's functioning within this domain. Most are inexpensive and relatively quick to administer. Within the visual modality, the *Motor-Free Visual Perception Test* (*MFVPT*; Colarusso & Hammill, 1996) allows the clinician to evaluate visual perception without motor involvement in children. This 36-item measure assesses five facets of visual perception: spatial relations, visual discrimination, figure–ground, visual closure, and visual memory. The *MFVPT* is intended for children four to eight years of age. The *MFVPT* can offer information essential for the differential diagnosis of motor vs. visual processing problems. However, when used in isolation from other measures or techniques, the *MFVPT* offers information regarding visual processing difficulties but is unable to rule out motor concerns. For adults, the *Benton Revised Visual Retention Test* (Benton-Sivan, 1992) is a widely used measure of visuoperceptual ability, constructional skills, and immediate visual memory (Youngjohn, Larrabee, & Crook, 1993). Clients are required to reproduce abstract geometric designs from memory.

Some clients have difficulty discriminating sounds even when thresholds for sound perceptions are intact (Lezak, 1995). Auditory discrimination can be tested by having the client repeat words and phrases spoken by the clinician, or by asking the client whether two spoken words are the same or different. On this task, the clinician will want to use word pairs that sound alike such as "cat" and "cap" along with identical word pairs (Lezak, 1995). This technique has been formalized through the development of *Wepman's Auditory Discrimination Test* (Wepman & Reynolds, 1987) which allows the clinician to determine whether the client is able to discriminate similar sounding words adequately. Although the test was originally devised to identify auditory discrimination problems in young school children, and the present norms were developed on samples of four to eight year olds, norms for the oldest age group (8–0 to 8–11) are adequate for adults since auditory discrimination is generally fully developed by this age (Lezak, 1995).

The perception of tactile stimuli is regularly measured as a component of a thorough neuropsychological examination, but less often in nonspecialized clinical settings. Informal strategies for evaluating this area include asking clients to indicate whether they feel the sharp or the dull end of a pin, pressure from one or two points (applied simultaneously and close together), or pressure from a graded set of plastic hairs, the "Von Frey hairs" (Lezak, 1995). The eyes should be closed or the hand kept out of sight when tactile sensory functions are tested. More formal measures include the *Tactile Form Perception Test* (Benton, Hamsher, Varney, & Spreen, 1983) and the *Tactual Performance Test* (Reitan & Davison, 1974). Deficits in tactile senses are often associated with damage to the right hemisphere of the brain and may have important implications for a client's vocational functioning (Lezak, 1995).

4.09.4.1.1 Sensory perception: implications for intervention

If the client is having difficulty in one or more areas of sensory perception, this information is critical for intervention planning. That is, the client's unique pattern of receiving information from the environment can be used to create effective education, rehabilitation, or therapeutic intervention. If a client is weak in auditory processing but strong in visual processing, for example, visual cues such as drawings, videos, or demonstrations may be the most effective means for training them in new skills.

4.09.4.2 Motor: Fine and Gross

The motor domain involves a range of both fine and gross motor movement. Fine motor skill is commonly thought of as movement which does not involve the entire body. Writing, opening a letter, or tying a shoe are all examples of fine motor movements. Gross motor movement involves large extremities and often the entire body. Activities such as walking or sitting down involve gross motor capacities. Intentional movement, using fine and gross motor skills, involves a series of brain-based systems and is learned with repetition. With repeated action, the movement becomes rote or, as Luria (1973) described it, a "kinesthetic melody." Movement also can consist of both discrete and continuous patterns. Movements that are discrete might involve something as simple as lifting a finger, while continuous movements include an integrated set of skills like skipping. Movements may be disrupted if damage exists in the premotor cortex where the "kinesthetic melody" is believed to be formed. If this occurs, the individual may not be able to perform serial-continuous movements but may be able to demonstrate the specific discrete movements. Because of the complexity of motor patterns, the individual's posture, movement in isolation, and movement in serial order should be assessed for possible intervention. This can be accomplished by observing individuals completing tasks such as writing their name (uses one hand), tying their shoes (uses both hands), and also performing novel tasks such as repeated tapping or clapping patterns. It should be noted if there is difficulty integrating the use of both hands. Both fine and gross motor skills should always be evaluated.

Although informal methods will yield a great deal of information regarding a client's fine and gross motor functioning, several standardized instruments are also available which measure various specific or broad components of motoric functioning. For example, the *Bruininks–Oseretsky Test of Motor Proficiency* (Bruininks, 1978) provides a comprehensive picture of an individual's motor development. The instrument was designed for children aged 4–5 to 14–15 and can be administered in 15–60 minutes, depending on whether the complete or short form of the battery is used. Three composite scores are provided in the areas of: Gross Motor Development (Running speed and agility, Balance, Bilateral coordination, Strength), Gross and Fine Motor Development (Upper-limb coordination), and Fine Motor Development (Response speed, Visual-motor control, Upper-limb speed and dexterity). Specific fine motor abilities can also be measured by using the Finger Tapping Test and the Grip Strength Test which are both a part of the Halstead-Reitan Battery (Reitan & Wolfson 1993).

To measure a client's lateral preference, the *Lateral Preference Schedule* (Dean, 1988) can be administered to obtain a better understanding of clients' lateral preference in the use of their eyes, ears, arms, hands, and feet (Rothlisberg, 1991). Atypical patterns of lateral preference have been hypothesized to indicate potential predictors of reading difficulty (Bemporad & Kinsbourne, 1983; Dean, Schwartz, & Smith, 1981). Determining lateral preference can be useful in interpreting assessment findings and in creating a plan for rehabilitation (Lezak, 1995).

4.09.4.3 Sensory-motor Integration

An additional component of our motoric functioning is the ability to integrate what is received by the senses with what is produced through action. For example, an individual may be able to perceive letters correctly and have adequate fine motor control, but still have difficulty correctly copying material presented in visual form. Numerous paper-and-pencil tests have been developed to assess motor function as it relates to visual-motor integration. Two of the most popular measures for this purpose are the *Bender Visual-Motor Gestalt* (Bender, 1938) and the *Developmental Test of Visual-Motor Integration* (*VMI*; Beery, 1989).

The *Visual-Motor Gestalt Test* (Bender, 1938) is an individually administered test containing nine geometric figures which the client copies on to a blank sheet of paper. While historically this test was seen as a general measure of organicity, it is more appropriately used as a measure of visual-motor skills. Standard scores are provided in the developmental scoring system for children ages 5–0 to 11–11, although it is frequently used with adults as well. Most commonly known as the "Bender," this measure is perhaps the best known and most widely used visual-motor assessment procedure available today (Bender, 1938; Reynolds & Kamphaus, 1990). As a component of a comprehensive assessment battery, performance on the Bender has long been thought to reveal visual-motor difficulties that may be associated with cerebral impairment (Sattler, 1992). Traditionally used to assess an individual's constructional praxic skills, the Bender provides an evaluation of motor integration employed in the execution of complex learned movements (Hartlage & Golden, 1990). The information generated through this process may then be compared with levels of performance across other measures of functioning.

Alternate uses of the Bender include its administration as a memory test as well as a copying test. This dual administration process can be employed to assess different mental functions (short-term visual memory and visual perception) which utilize the same modalities in perception and task execution (Sattler, 1992). An additional technique available when interpreting the Bender performance is to have individuals compare the figure which they produced with the corresponding stimulus design. If the client is unable to recognize obvious differences between the two designs, a perceptual deficit may be involved. Likewise, if the client is able to detect a difference between the two figures, but is unable to make them alike, motor involvement may be influencing performance (Hartlage & Golden, 1990). In the personality area, performance on the Bender may also be used to develop hypotheses regarding impaired performance due to poor planning, impulsivity, or compulsivity. Extremely large or small figures, heavily reinforced lines, and second attempts are examples of the item reproduction difficulties which are thought to indicate emotional concerns on the part of the individual.

The *VMI* is an individual or group administered test that involves copying a sequence of 24, increasingly complex, geometric figures. The test requires a relatively short administration time and is designed primarily for ages 4 to 13. The *VMI* offers several advantages as a tool for assessment and is widely used in psychological evaluation and research. The most common use of the *VMI*, now in its third edition, seems to be in assisting with the diagnosis of children who are suspected of having learning problems due to visual-motor difficulties. The *VMI* is also frequently employed when investigating the reliability and validity of other tests of visual-motor integration, such as the Bender, self drawing tasks, progressive matrices, and neuropsychological tests (Goldstein, Smith, & Waldrep, 1986; Palisano & Dichter, 1989). Because the *VMI* does not require a verbal response, it has also been used to assess visual-motor processes among non-English-speaking children (Brand, 1991; Frey & Pinelli, 1991).

4.09.4.3.1 Motor: implications for intervention

Motor problems and sensory-motor integration difficulties can impair a client's ability to write or to learn new skills requiring motor coordination, and generally can have a negative impact on daily functioning. In the classroom setting, possible suggestions for accommodating these difficulties might include modifying instructions to compensate for motor difficulties (e.g., allowing pointing to the correct choice rather than writing, allowing students to tape record notes or copy them from others). In a rehabilitation setting, the client may need to learn alternative methods for writing such as using word processing programs on a computer or, if serious difficulties exist, using voice-activated programs. Consultation with an occupational therapist or a physical therapist will be extremely helpful in treatment planning when motor difficulties are evident.

4.09.4.4 Communication/Language

Language is the basic tool of human communication and hence essential to evaluate when working with any client (Black & Strub, 1994). It should be viewed as a key skill because it serves as a primary means of conveying information from the individual to others and from others to the individual. Thus, communication difficulties have the power to influence all areas of life (D'Amato, Rothlisberg, & Rhodes, 1997). Difficulties or dysfunction found on tests of higher level functioning (e.g., learning processes) may well be secondary to a language disorder. Accordingly, language should be evaluated early in the course of an assessment to rule out problems in this area of functioning (Black & Strub, 1994; Lezak, 1995). Another obvious reason to evaluate language and communication skills is that language disorders occur as the result of a wide range of neurologic diseases and can manifest in a variety of forms of aphasia (e.g., Wernicke's, anomia, global, alexia, agraphia; Kolb & Whishaw, 1990). To aid in the interpretation of test findings, it is important for the clinician to be familiar with the various classic clinical aphasia presentations (see Gaddes & Edgell, 1994; Kolb & Whishaw, 1990; Lezak, 1995).

The language evaluation should be systematic and include an assessment of a range of relatively specific language functions. Assessment must evaluate both receptive and expressive verbal and nonverbal abilities to determine if adaptations are needed to enhance the individual in academic, vocational, and social situations. As part of an informal evaluation of expressive language, the clinician will want to evaluate spontaneous speech and verbal fluency by asking the client open-ended questions (Black & Strub, 1994). While the client is responding, the clinician can listen carefully for abnormal articulation, dysarthria (incoordination of the speech apparatus), verbal apraxia (difficulty carrying out purposeful speech), dysfluency, loss of prosody (melodic intonation), and disturbances of syntax or paraphasic errors (production of unintended syllables, words, or phrases) (Black & Strub, 1994; Kolb & Whishaw, 1990). Another important component of language, pragmatics, can also be evaluated. Pragmatics refers to the knowledge and activities of socially appropriate communication, which takes in much of the nonverbal aspects of communication, such as gestures, loudness of speech, as well as verbal appropriateness (Sohlberg & Mateer, 1990).

Evaluation of verbal fluency can be accomplished by counting the number of words the client is able to produce without repetition within a restricted category (e.g., animals or words beginning with a particular letter) and time (e.g., 60 seconds). The average adult should produce approximately 20 animal names and a total of 40–60 words with performance depending to some degree on the client's intelligence, education, and social/linguistic background (Black & Strub, 1994). Additional methods of examining expressive language include having the client repeat back meaningful verbal phrases or sentences of increasing length and semantic complexity. Word finding and naming difficulties can be detected by the client's responses to the open-ended questions or by having the client describe a picture containing a series of objects or actions (Black & Strub, 1994).

Another major area of language functioning is receptive language or an individual's ability to understand what has been said. A comprehensive assessment should include an evaluation of the individual's ability to analyze and integrate information presented in a verbal format, since a common difficulty among those experiencing traumatic head injury is a decreased capacity to coordinate the social aspects of language (Ylvisaker, Szekeres, Haarbauer-Krupa, Urbanczyk, & Feeney, 1994). It is not sufficient to evaluate language comprehension based on open-ended questions because this method relies on expressive skills and does not examine comprehension in isolation (Black & Strub, 1994). Language comprehension can be evaluated informally by asking the client to point to common objects in the room or by asking a series of increasingly complex questions that require only a "yes" or "no" response (e.g. "Do dogs have four legs?") (Black & Strub, 1994). An evaluation of a client's reading and writing skills could also be included in an evaluation of language and communication (Black & Strub, 1994). The client can be asked to read sentences of increasing difficulty, spell words to dictation, and compose a paragraph in response to a prompt (e.g., "Tell me how to change a tire.") (Black & Strub, 1994).

In addition to the informal methods, several instruments are available that can prove useful for the clinical evaluation of language. There are a number of aphasia tests and batteries (e.g., *Boston Diagnostic Aphasia Examination*, Goodglass & Kaplan, 1983; *Multilingual Aphasia Examination*, Benton & Hamsher, 1989) which involve lengthy, well-controlled procedures and are best left to speech pathologists who have more extensive training in the specialized techniques of aphasia examinations (Lezak, 1995). As an alternative, aphasia screening tests can be used to indicate the presence of an aphasic disorder and may even highlight its specific characteristics, but do not provide the fine discriminations of the complete aphasia test batteries. Furthermore, these screening tests do not require technical knowledge of speech pathology for adequate administration or determination of whether a significant aphasic disorder is present (Lezak, 1995). One of the most comprehensive aphasic screening tests available is the *Revised Token Test* (McNeil & Prescott, 1978). This expanded version of the original *Token Test* (De Renzi & Vignolo, 1962) contains 10 10-item subtests. McNeil and Prescott (1978) sought to ameliorate psychometric weaknesses of the original with this revision as well as seeking to develop an evaluative system for describing the nature and quantifying the degree of language deficit in order to facilitate treatment planning. Using tokens of various shapes and sizes, the clinician gives the client a series of increasingly complex instructions to follow. Though simple to administer, this instrument is reportedly very sensitive to disrupted linguistic processes that are central to the aphasic disability (Lezak, 1995).

Clinicians wanting a basic measure of different aspects of communication and language may wish to consider using the *Peabody Picture Vocabulary Test-Revised* (*PPVT-R*; Dunn & Dunn, 1981), the *Test of Language Development* (*TOLD-2*; Hammill & Newcomer, 1988), and the *Clinical Evaluation of Language Fundamentals-Revised* (Semel, Wiig, & Secord, 1987) or the *Test of Language Competence* (*TLC*; Wiig & Secord, 1989). Unfortunately, most of these tests are normed exclusively on children and adolescents and, therefore, have limited application to adults. One of the most frequently used tests, and one which has adult norms, is the *PPVT-R*. This test measures receptive vocabulary only and was normed for individuals aged two and a half to adulthood. The *PPVT-R* is untimed and requires the examinee to select from each plate of four pictures the one that best represents the target word. The test requires no reading ability, nor is the ability to point or provide an oral response essential (Shea, 1989). The *PPVT-R* can help to establish the level of verbal understanding a client has when expressive language is not required. Comparing such receptive skills with those expressive skills needed for other tests may help in developing hypotheses about the qualitative nature of verbal performance and in framing potential treatment (D'Amato, Rothlisberg, & Rhodes, 1997).

The *TOLD-2* is available in a form designed for primary ages (4–0 to 8–11) and intermediate ages (8–6 to 12–11). This test purports to measure receptive and expressive language proficiency. The results for the *TOLD-2* provide quotients for an Overall Spoken Language score, and for the composites of Listening, Speaking, Semantics, Syntax, and Phonology (on the primary form only). Overall, the *TOLD-2* instruments are reliable and valid as language screening tools for younger clients (Wochnick Fodness, McNeilly, & Bradley-Johnson, 1991; Westby, 1988). Some tests have been developed to measure more complex language usage, such as the TLC (Wiig & Secord, 1989) which purports to measure metalinguistic abilities. The four subtests involve producing multiple meanings for ambiguous sentences, recognizing inferences on the basis of incomplete information, creating sentences given three words and a context, and recognizing the meaning of figurative language. This type of test may be useful for identifying subtle problems in language usage (Crosson, 1996).

4.09.4.4.1 Communication/language: implications for intervention

For clients who are experiencing difficulty with either or both receptive and expressive language, the clinician can modify verbal interaction by shortening the length of information presented or presenting information in steps. Additional ideas for the school-age client might include recommending that the teacher repeat directions and have the student also repeat and explain the directions back to the teacher, pairing verbal instructions with nonverbal cues, and using nonverbal cues. Many of these strategies could be adapted to adults in rehabilitation and other types of therapeutic settings as well. The clinician must be careful to check frequently with clients to ensure understanding and to assist the clients and their families to adjust to these communication or language deficits.

4.09.5 FUTURE DIRECTIONS

Although knowledge about how individuals process information has grown exponentially since the late 1970s, researchers and practitioners alike are left with many questions regarding how individuals remember and learn. How do age, gender, and ethnicity impact a client's functioning on these specific instruments? Do individuals with brain damage process information differently than individuals with "normal" brain functioning? How do the results of an assessment translate into effective treatment strategies that will help individuals function better on the job? Despite these questions and more, as a field we do know that current measures of processing can allow practitioners to make predictions with a reasonable level of confidence. However, we must also recognize that future research investigating the prediction accuracy of various tests is needed to expand the range and precision of clinical prediction (Long, 1996). Indeed, limited research exists that evaluates the effect of various treatment approaches on success in clinical pediatric or adult populations (Batchelor, 1996b; Ris & Noll, 1994). As our knowledge of the nervous system is expanded, and we begin to understand the intricacies of the brain's organization, we can begin to see how perceptions are formed, information stored and integrated, and action taken. Until that time, the explanation for behavior and certain learning difficulties can only be inferred (D'Amato, Rothlisberg, & Leu, in press).

Another component complicating our enhanced understanding of information processing is the need for common language and goals between neurologists, psychologists, educational researchers, and vocational rehabilitation specialists. Bigler (1996) notes that it is

> essential that physicians and psychologists work toward some common understanding of normal and abnormal behavioral manifestations of brain functioning, particularly aspects of complex attention and integration of sensory experiences, memory, motivation, organization of verbal and nonverbal cognition, abstract thinking, problem solving, executive functions, and self-monitoring of behavior. Without agreement on a detailed and relatively comprehensive model of neurobehavioral development, psychologists will be limited in their ability to develop appropriate assessment and intervention strategies. (p. 50)

Unfortunately, outcomes in rehabilitation research have also suffered from a lack of consensus with importance placed on different variables depending on the orientation of the author and audience (Batchelor, 1996b). For example, clinical researchers have examined pre- and postperformance measures of cognitive, motivational, and behavioral functions (Ris & Noll, 1994), while service providers have focused on outcome constructs such as employment and independent living (e.g., Adunsky, Hershkowitz, Rabbi, Asher-Sivron, & Ohry, 1992). Concurrently, third-party payers are interested in length of stay, cost, and effectiveness in allocation of resources (Fratalli, 1993), while families and consumers are interested in quality of care. Batchelor (1996b) concluded that the majority of research has emphasized short-term outcomes and the meaningful questions generated by service providers, consumers, and third-party payers have been overlooked and present an important direction for the field to pursue.

In terms of the practical aspects of assessment, the practitioner's task is not an easy one. In order to generate an accurate diagnosis or provide the most sound recommendations, a complete understanding of the client is necessary. A practitioner could spend hours evaluating each of the areas outlined with careful consideration of all subdomains using a multidimensional approach. Although this approach would yield a bounty of information, it may not be practical given time limitations and insurance policy guidelines. The key for the clinician is to find a balance between finding out the most important information about client functioning through the use of instruments with the best predictive ability and spending a limited amount of time on assessment. By creating an efficient and effective assessment approach, more time is available to implement treatment. The task of generating recommendations for interventions that are likely to enhance client functioning is of central importance to the issue of assessment. To this end, continued information is needed on the relationship between assessed cognitive processing and predicted future performance in real-world settings (Sbordone & Long, 1996). Furthermore, we must gain knowledge about the most effective intervention strategies for all types of individuals. Future research will enable us to understand how the brain processes information and what treatments are effective with what types of clients. Indeed, this additional knowledge may allow us to match client subtypes with specific treatments which will increase the effectiveness and efficacy of psychological services.

4.09.6 SUMMARY

The assessment of children, adolescents, and adults encompasses a wide range of domains from which a clinician may view a client's functioning. Consideration must be given to the many layers of the client context (e.g., family support, socioeconomic status, domain of functioning), client characteristics (e.g., motivation, education level, ethnicity, age), as well as the specific cognitive processes under question (e.g., memory, sensory perception). Although often overlooked or subsumed within the broader arenas of intelligence or achievement, the specific areas of memory, learning, and special aptitudes are critical to our daily functioning. That is, one will have difficulty demonstrating intelligence or learning new tasks, if there is a severe memory deficit or difficulty in accurately perceiving stimuli. The descriptions of the domains presented in this chapter have offered insight into the field's current understanding of these systems as well as the breadth of evaluation strategies available to probe the diverse nature of cognitive processes. Once the practitioner identifies the assessment needs of the individual client, and a decision is made as to the components most relevant for exploring the referral question, the process can begin. By generating quality data, our ability to predict outcomes and provide effective potential intervention strategies is increased. Indeed, the goal of any assessment is to respond correctly to the question presented by the referral source, provide accurate predictions of future outcomes, and generate effective strategies for improving the client's adaptation or functioning.

4.09.7 REFERENCES

Adunsky, A., Hershkowitz, M., Rabbi, R., Asher-Sivron, L., & Ohry, A. (1992). Functional recovery in young stroke patients. *Archives of Physical Medicine and Rehabilitation*, *73*, 859–862.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington DC: Author.

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.

Arter, J. A., & Jenkins, J. R. (1977). Examining the benefits and prevalence of modality considerations in special education. *The Journal of Special Education*, *11*, 291–298.

Baddeley, A. (1986). *Working memory*. Oxford, UK: Oxford University Press.

Baddeley, A., Emslie, H., & Nimmo-Smith, I. (1994). *Doors and People: A Test of Visual and Verbal Recall and Recognition*. Suffolk, UK: Thames Valley Test Co.

Baldwin, J. D., & Baldwin, J. I. (1986). *Behavior principles in everyday life* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Barkley, R. A. (1996). Critical issues in research on attention. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function* (pp. 45–56). Baltimore: Brookes.

Batchelor, E. S. (1996a). Neuropsychological assessment of children. In E. S. Bachelor & R. S. Dean (Eds.), *Pediatric neuropsychology: Interfacing assessment and treatment for rehabilitation* (pp. 9–26). Boston: Allyn & Bacon.

Batchelor, E. S. (1996b). Future considerations for rehabilitation research and outcome studies. In E. S. Bachelor & R. S. Dean (Eds.), *Pediatric neuropsychology: Interfacing assessment and treatment for rehabilitation* (pp. 347–352). Boston: Allyn & Bacon.

Beery, K. E. (1989). *Developmental Test of Visual-Motor Integration*. Odessa, FL: Psychological Assessment Resources.

Begali, V. (1992). *Head injury in children and adolescents: A resource and review for school and allied professionals* (2nd ed.). Brandon, VT: Clinical Psychology Publishing Company.

Begali, V. (1994). The role of the school psychologist. In R. C. Savage & G. F. Wolcott (Eds.), *Educational dimensions of acquired brain injury* (pp. 453–473). Austin, TX: PRO-ED.

Bemporad, B., & Kinsbourne, M. (1983). Sinistrality and dylexia: A possible relationship between subtypes. *Topics in Learning and Learning Disabilities*, *3*(1), 48–65.

Bender, L. (1938). A visual motor gestalt test and its clinical use. *American Orthopsychiatric Association Research Monograph*, No. 3. New York: American Orthopsychiatric Association.

Benton, A. L., & Hamsher, K. deS. (1989). *Multilingual Aphasia Examination*. Iowa City, IA: AJA Associates.

Benton, A. L., Hamsher, K. deS., Varney, N. R., & Spreen, O. (1983). *Contributions to neuropsychological assessment*. New York: Oxford University Press.

Benton-Sivan, A. (1992). *The Revised Visual Retention Test* (5th ed.). New York: The Psychological Corporation.

Berk, R. A. (1995). Review of the Memory Assessment Scale. In J. C. Conoley & J. C. Impara (Eds). *The twelfth mental measurement yearbook* (pp. 593–594). Lincoln, NE: Buros.

Bigler, E. D. (1996). Bridging the gap between psychology and neurology: Future trends in pediatric neuropsychology. In E. S. Bachelor & R. S. Dean (Eds.), *Pediatric neuropsychology: Interfacing assessment and treatment for rehabilitation* (pp. 27–54). Boston: Allyn & Bacon.

Black, F. W., & Strub, R. L. (1994). The bedside and office mental status examination. In S. Touyz, D. Byrne, & A. Gilandas (Eds.), *Neuropsychology in clinical practice* (pp. 38–60) Boston: Academic Press.

Brand, H. J. (1991). Correlation for scores on revised tests of visual-motor integration and copying test in a South African sample. *Perceptual and Motor Skills*, *73*, 225–226.

Brown, R. T., Armstrong, F. D., & Eckman, J. R. (1993). Neurocognitive aspects of pediatric sickle cell disease. *Journal of Learning Disabilities*, *26*, 33–45.

Brown, V. L., Hammill, D. D., & Wiederholt, J. L. (1995). *Test of Reading Comprehension-3 (TORC-3)*. Austin, TX: PRO-ED.

Bruininks, R. H. (1978). *Bruininks–Oseretsky Test of Motor Proficiency*. Circle Pines, MN: American Guidance Service.

Campbell, J. W., D'Amato, R. C., Raggio, D. J., & Stephens, K. D. (1991). Construct validity of the computerized Continuous Performance Test with measures of intelligence, achievement, and behavior. *Journal of School Psychology*, *29*, 143–150.

Campione, J. C., & Brown, A. L. (1987). Linking dynamic assessment with school achievement. In C. S. Lidz (Ed.), *Dynamic assessment: Foundations and fundamentals* (pp. 82–115). New York: Guilford.

Cantor, J., Engle, R. W., & Hamilton, G. (1991). Short-term memory, working memory, and verbal abilities: How do they relate? *Intelligence*, *15*, 229–246.

Cohen, S. B. (1991). Adapting educational programs for students with head injuries. *Journal of Head Trauma Rehabilitation*, *1*, 56–63.

Colarusso, R. P., & Hammill, D. D. (1996). *Motor-Free Visual Perception Test-Revised (MFPT-R)*. Novato, CA: Academic Therapy.

Cole E., & Siegel, J. A. (1990). School psychology in a multicultural community: Responding to childrens' needs. In E. Cole & J. A. Siegel (Eds.), *Effective consultation in school psychology* (pp. 141–169). Toronto, ON: Hogrefe & Huber.

Colvin, S. S. (1921). Intelligence and its measurement: A symposium (IV). *Journal of Educational Psychology*, *12*, 136–139.

Connolly, A. J. (1988). *Keymath-revised: A diagnostic inventory of essential mathematics*. Circle Pines, MN: American Guidance Service.

Cooley, E. L., & Morris, R. D. (1990). Attention in children: A neuropsychology based model of assessment. *Developmental Neuropsychology*, *6*, 239–274.

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods. A handbook for research on interactions*. New York: Irvington.

Crosson, B. (1996). Assessment of subtle language deficits in neuropsychological batteries: Strategies and implications. In R. J. Sbordone & C. J. Long (Eds.), *Ecological* *validity of neuropsychological testing* (pp. 243–259). Delray Beach, FL: GR Press/St Lucie Press.

D'Amato, R. C. (1990). A neuropsychological approach to school psychology. *School Psychology Quarterly*, *5*, 141–160.

D'Amato, R. C., & Dean, R. S. (Eds.) (1989a). *The school psychologist in nontraditional settings: Integrating clients, services, and settings*. Hillsdale, NJ: Erlbaum.

D'Amato, R. C., & Dean, R. S. (1989b). The past, present, and future of school psychology in nontraditional settings. In R. C. D'Amato & R. S. Dean (Eds.), *The school psychologist in nontraditional settings: Integrating clients, services, and settings* (pp. 185–209). Hillsdale, NJ: Erlbaum.

D'Amato, R. C., & Rothlisberg, B. A. (1992). *Psychological perspectives on intervention: A case study approach to prescriptions for change*. New York: Longman.

D'Amato, R. C., & Rothlisberg, B. A. (1996). How education should respond to students with traumatic brain injuries. *Journal of Learning Disabilities*, *29*, 670–683.

D'Amato, R. C., Rothlisberg, B. A., & Leu, P. W. (in press). Neuropsychological assessment for intervention. In C. R. Reynolds & T. B. Gutkin (Eds.), *The handbook of school psychology* (3rd ed.). New York: Wiley.

D'Amato, R. C., Rothlisberg, B. A., & Rhodes, R. L. (1997). Utilizing a neuropsychological paradigm for understanding common educational and psychological tests. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (2nd ed.). New York: Plenum.

Dana, R. H. (1993). *Multicultural assessment perspectives for professional psychology*. Boston: Allyn & Bacon.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.

Das, J. P., Kirby, J., & Jarman, R. F. (1979). *Simultaneous and successive cognitive processes*. New York: Academic Press.

Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes. The PASS theory of intelligence* New York: Allyn & Bacon.

Dean, R. S. (1977). Canonical analysis of a jangle fallacy. *Multivariate Experimental Clinical Research*, *3*, 17–20.

Dean, R. S. (1983). Intelligence-achievement discrepancies in diagnosing pediatric learning disabilities. *Clinical Neuropsychology*, *3*, 58–62.

Dean, R. S. (1984). Functional lateralization of the brain. *Journal of Special Education*, *18*, 239–256.

Dean, R. S. (1985a). Neuropsychological assessment. In R. Michels, J. O. Cavenar, H. K. H. Brodie, A. M. Cooper, S. B. Guze, L. L. Judd, G. L. Klerman, & A. J. Solnit (Eds.), *Psychiatry* (pp. 1–16). Philadelphia: Lippincott.

Dean, R. S. (1985b). Foundation and rationale for neuropsychological bases of individual differences. In L. C. Hartlage & C. F. Telzrow (Eds.), *The neuropsychology of individual differences: A developmental perspective* (pp. 7–39). New York: Plenum.

Dean, R. S. (1986). Perspectives on the future of neuropsychological assessment. In B. S. Plake & J. C. Witt (Eds.), *Buros-Nebraska series on measurement and testing: Future of testing and measurement* (pp. 203–241). Hillsdale, NJ: Erlbaum.

Dean, R. S. (1988). *Lateral Preference Schedule*. Odessa, FL: Psychological Assessment Resources.

Dean, R. S., & Gray, J. W. (1990). Traditional approaches to neuropsychological assessment. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence and achievement* (pp. 371–388). New York: Guilford Press.

Dean, R. S., Schwartz, N. H., & Smith, L. S. (1981). Lateral preference patterns as a discriminator of learning difficulties. *Journal of Consulting and Clinical Psychology*, *49*, 227–235.

Dearborn, W. F. (1921). Intelligence and its measurement: A symposium (XII). *Journal of Educational Psychology*, *12*, 210–212.

De Renzi, E., & Vignolo, L. A. (1962). The Token Test: A sensitive test to detect disturbances in aphasics. *Brain*, *85*, 665–678.

Drew, R. H., & Templer, D. I. (1992). Contact sports. In D. I. Templer, L. C. Hartlage, & W. G. Cannon (Eds.), *Preventable brain damage: Brain vulnerability and health* (pp. 15–29). New York: Springer.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN: American Guidance Service.

Eliason, M. J., & Richman, L. C. (1987). The Continuous Performance Test in learning disabled and nondisabled children. *Journal of Learning Disabilities*, *20*, 614–619.

Feuerstein, R., Rand, Y., & Hoffman, M. (1979). *The dynamic assessment of retarded performers: The Learning Potential Assessment Device: Theory, instruments, and techniques*. Baltimore: University Park.

Figueroa, R. A., & Garcia, E. (1994). Issues in testing students from culturally and linguistically diverse backgrounds. *Multicultural Education*, *2*, 10–19.

Frattali, C. M. (1993). Perspectives on functional assessment: Its use for policy making. *Disability and Rehabilitation*, *15*, 1–9.

Frey, P. D., & Pinelli, B. (1991). Visual discrimination and visuomotor integration among two classes of Brazilian children. *Perceptual and Motor Skills*, *72*, 847–850.

Fuchs, L. S. (1994). Integrating curriculum-based measurement with instructional planning for students with learning disabilities. In N. C. Jordan & J. Goldsmith-Phillips (Eds.), *Learning disabilities: New directions for assessment and intervention* (pp. 177–195). Boston: Allyn & Bacon.

Gaddes, W. H., & Edgell, D. (1994). *Learning disabilities and brain function: A neuropsychological approach* (3rd ed.). New York: Springer-Verlag.

Geil, M., & D'Amato, R. C. (1996). *Contemporary ecological neuropsychology: An alternative to the medical model for conceptualizing learning disabilities*. Manuscript submitted for publication.

Golden, C. J. (1981). The Luria–Nebraska Children's Battery: Theory and formulation. In G. W. Hynd & J. E. Obrzut (Eds.), *Neuropsychological assessment and the school-aged child: Issues and procedures* (pp. 277–302). New York: Grune & Stratton.

Golden, C. J., Sawicki, R. F., & Franzen, M. D. (1984). Test construction. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment* (pp. 19–37). New York: Pergamon.

Goldstein, D. J., Smith, K. B., & Waldrep, E. E. (1986). Factor analytic study of the Kaufman Assessment Battery for Children. *Journal of Clinical Psychology*, *42*, 890–894.

Goodglass, H., & Kaplan, E. (1983). *Boston Diagnostic Aphasia Examination (BDAE)*. Philadelphia: Lea and Febiger. Distributed by Psychological Assessment Resources, Odessa, FL.

Gray, J. W., & Dean, R. S. (1989). Approaches to the cognitive rehabilitation of children with neuropsychological impairment. In C. R. Reynolds & F. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (pp. 397–408). New York: Plenum.

Greenberg, L. (1993). *Test of variables of attention* (*T.O.V.A.*$^{\mathrm{TM}}$). Wood Dale, IL: Stoetling.

Greenberg, L. M., & Waldman, I. D. (1993). Developmental normative data on the test of variables of attention (*T.O.V.A.*$^{\mathrm{TM}}$). *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *34*, 1019–1030.

Guilmette, T. J., & Giuliano, A. J. (1991). Taking the stand: Issues and strategies in forensic neuropsychology. *The Clinical Neuropsychologist*, *5*, 197–219.

Gutkin, T. B., & Reynolds, C. R. (Eds.) (1990). *The handbook of school psychology* (2nd ed.). New York: Wiley.

Halperin, J. M., Sharma, V., Greenblatt, E., & Schwartz, S. (1991). Assessment of the Continuous Performance Test: Reliability and validity in a nonreferred sample. *Psychological Assessment*, *3*, 603–608.

Hammill, D. D. (1991). *Detroit Tests of Learning Aptitude (DTLA-3)* (3rd ed.). Austin, TX: PRO-ED.

Hammill, D. D., & Bryant, B. R. (1991a). *Detroit Tests of Learning Aptitude-Adult (DTLA-A)*. Austin, TX: PRO-ED.

Hammill, D. D., & Bryant, B. R. (1991b). *Detroit Tests of Learning Aptitude-Primary (DTLA-P:2)* (2nd ed.). Austin, TX: PRO-ED.

Hammill, D. D., & Larsen, S. C. (1996). *Test of Written Language-3 (TOWL-3)*. Austin, TX: PRO-ED.

Hammill, D. D., & Newcomer, P. L. (1988). *Test of Language Development Intermediate (TOLD-2)* (2nd ed.). Austin, TX: PRO-ED.

Hamsher, K. de S. (1984). Specialized neuropsychological assessment methods. In G. Goldstein & M. Hersen (Eds.). *Handbook of psychological assessment* (pp. 235–256). New York: Pergamon.

Hartlage, L. C., & Golden, C. J. (1990). Neuropsychological assessment techniques. In T. B. Gutkin & C. R. Reynolds (Eds.), *The handbook of school psychology* (2nd ed., pp. 431–457). New York: Wiley.

Hartlage, L. C., & Telzrow, C. F. (1983). The neuropsychological basis of educational intervention. *Journal of Learning Disabilities*, *16*, 521–528.

Hooper, S. R. (1995). Review of the Visual Search and Attention Test. In J. C. Conoley & J. C. Impara (Eds.), *The twelfth mental measurements yearbook* (pp. 1081–1082). Lincoln, NE: Buros.

Huebner, E. S. (1992). Review of the Wechsler Memory Scale-Revised. In J. J. Kramer & J. C. Conoley (Eds.), *The eleventh mental measurement yearbook* (pp. 1023–1024). Lincoln, NE: Buros.

Hynd, G. W., & Semrud-Clikeman, M. (1990). Neuropsychological assessment. In A. S. Kaufman (Ed.), *Assessing adolescent and adult intelligence* (pp. 638–695). Boston: Allyn & Bacon.

Hynd, G. W., & Willis, W. G. (1988). *Pediatric neuropsychology*, Boston: Allyn & Bacon.

Jarvis, P. E., & Barth, J. T. (1994). *The Halstead-Reitan Neuropsychological Battery: A guide to interpretation and clinical applications*. Odessa, FL: Psychological Assessment Resources.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149.

Kaufman, A. S. (1990). *Assessing adolescent and adult intelligence*. Boston: Allyn & Bacon.

Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. New York: Wiley.

Kavale, K. A., Forness, R. F., & Bender, M. (1988). *Handbook of learning disabilities: Volume II: Methods and interventions*. Boston: College-Hill.

Klee, S. H., & Garfinkel, B. D. (1983). The computerized Continuous Performance Task: A new measure of inattention. *Journal of Abnormal Child Psychology*, *11*, 489–495.

Kolb, B., & Whishaw, I. Q. (1990). *Fundamentals of human neuropsychology* (3rd ed.). New York: Freeman.

Kovacs, M., Goldston, D., & Ivengar, S. (1992). Intellectual development and academic performance of children with insulin-dependent diabetes mellitus: A longitudinal study. *Developmental Psychology*, *28*, 676–684.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*, 389–433.

LaBerge, D. (1995). *Attentional processing: The brain's art of mindfulness*. Cambridge, MA: Harvard University Press.

Lassiter, K. S., D'Amato, R. C., Raggio, D. J., Whitten, J. C. M., & Bardos, A. N. (1994). The construct specificity of the Continuous Performance Test: Does inattention relate to behavior and achievement? *Developmental Neuropsychology*, *10*, 179–188.

Lezak, M. D. (1983). *Neuropsychological assessment* (2nd ed.). New York: Oxford University Press.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Lindgren, S. D., & Lyon, D. (1983). *PACE: Pediatric assessment of cognitive efficiency*. Iowa City, IA: University of Iowa, Department of Pediatrics.

Long, C. J. (1996). Neuropsychological tests: A look at our past and the impact that ecological issues may have on our future. In R. J. Sbordone & C. J. Long (Eds.), *Ecological validity of neuropsychological testing* (pp. 1–14). Delray Beach, FL: GR Press/St Lucie Press.

Luria, A. R. (1970). The functional organization of the brain. *Scientific American*, *222*(3), 66–78.

Luria, A. R. (1973). *The working brain: An introduction to neuropsychology*. New York: Basic Books.

Luria, A. R. (1980). *Higher cortical functions in man* (2nd ed.). New York: Basic Books.

Markwardt, (1989). *Peabody Individual Achievement Test-Revised (PIAT-R)*. Circle Pines, MN: American Guidance Service.

Martinez, M. A. (1985). Toward a bilingual school psychology model. *Educational Psychology*, *20*, 143–152.

Mastropieri, M. A., & Scruggs, T. E. (1989). Constructing more meaningful relationships: Mnemonic instruction for special populations. *Educational Psychology Review*, *1*, 83–111.

McNeil, M. M., & Prescott, T. E. (1978). *Revised Token Test*. Austin, TX: PRO-ED.

Morris, R. D. (1996). Relationships and distinctions among the concepts of attention, memory, and executive function: A developmental perspective. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function* (pp. 11–16). Baltimore: Brookes.

Palincsar, A., Brown, A. L., & Campione, J. C. (1991). Dynamic assessment. In H. L. Swanson (Ed.), *Handbook on the assessment of learning disabilities: Theory, research, and practice* (pp. 75–95). Austin, TX: PRO-ED.

Palisano, R. J., & Dichter, C. G. (1989). Comparison of two tests of visual-motor development used to assess children with learning disabilities. *Perceptual and Motor Skills*, *68*, 1099–1103.

Poteat, G. M. (1995). Review of the Detroit Tests of Learning Aptitude, Third Edition. In J. C. Conoley & J. C. Impara (Eds.), *The twelfth mental measurement yearbook* (pp. 277–278). Lincoln, NE: Buros.

Pressley, M., & Levin, J. R. (Eds.) (1983). *Cognitive strategy research: Psychological foundations*. New York: Springer-Verlag.

Raggio, D. (1991). *Raggio Evaluation of Attention Deficit Disorder (Computerized test)*. Jackson, MS: University of Mississippi Medical Center, Infant and Child Development Clinic.

Reitan, R. M., & Davison, L. A. (1974). *Clinical neuropsychology: Current status and applications*. New York: Winston/Wiley.

Reitan, R. M., & Wolfson, D. (1985). *The Halstead–Reitan Neuropsychological Test Battery: Theory and clinical interpretation*. Tucson, AZ: Neuropsychology Press.

Reitan, R. M., & Wolfson, D. (1993). *The Halstead–Reitan Neuropsychological Test Battery: Theory and clinical interpretation* (2nd ed.). Tucson, AZ: Neuropsychology Press.

Reitan, R. M., & Wolfson, D. (1995, October). *Cognitive and emotional consequences of mild head injury*. Paper presented at the fall conference of the Colorado Neuropsychological Society, Colorado Springs, CO.

Resnick, L. B. (Ed.) (1976). *The nature of intelligence*. Hillsdale, NJ: Erlbaum.

Reynolds, C. R. (1981a). The neuropsychological basis of intelligence. In G. W. Hynd & J. E. Obrzut (Eds.), *Neuropsychological assessment and the school-aged child: Issues and procedures* (pp. 87–124). New York: Grune & Stratton.

Reynolds, C. R. (1981b). Neuropsychological assessment and the habilitation of learning: Considerations in the search for the aptitude x treatment interaction. *School Psychology Review*, *10*, 343–349.

Reynolds, C. R. (1986). Transactional models of intellectual development, yes. Deficit models of process remediation, no. *School Psychology Review*, *15*, 256–260.

Reynolds, C. R., & Bigler, E. D. (1994). *Test of memory and learning*. Austin, TX: PRO-ED.

Reynolds, C. R., & Kamphaus, R. W. (1990). *Handbook of psychological and educational assessment of children: Intelligence and achievement*. New York: Guilford Press.

Ris, D., & Noll, R. B. (1994). Long-term neurobehavioral outcome in pediatric brain-tumor patients: Review and methodological critique. *Journal of Clinical and Experimental Neuropsychology*, *16*(1), 21–42.

Rosvold, H., Mirsky, A., Sarason, I., Bransome, L., & Beck, L. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology*, *20*, 343–350.

Rothlisberg, B. A. (1991). Factor stability of the Lateral Preference Schedule. *International Journal of Neuroscience*, *61*, 83–85.

Rothlisberg, B. A. (1992). Integrating psychological approaches to intervention. In R. C. D'Amato & B. A. Rothlisberg (Eds.), *Psychological perspectives on intervention: A case study approach to prescriptions for change* (pp. 190–198). New York: Longman.

Rovet, J. F., Ehrlich, R. M., Czuchta, D., & Akler, M. (1993). Psychoeducational characteristics of children and adolescents with insulin-dependent diabetes mellitus. *Journal of Learning Disabilities*, *26*, 7–22.

Rovet, J. F., Ehrlich, R. M., & Hoppe, M. (1988). Specific intellectual deficits in children with early onset diabetes mellitus, *Child Development*, *59*, 226–234.

Sattler, J. M. (1992). *Assessment of children* (3rd ed., rev.). San Diego, CA: Sattler.

Sbordone, R. J., & Long, C. J. (Eds.) (1996). *Ecological validity of neuropsychological testing*. Delray Beach, FL: GR Press/St Lucie Press.

Selz, M. (1981). Halstead-Reitan neuropsychological test batteries for children. In G. W. Hynd & J. E. Obrzut (Eds.), *Neuropsychological assessment and the school-aged child: Issues and procedures* (pp. 195–235). New York: Grune & Stratton.

Semel, E., Wiig, E. H., & Secord, W. (1987). *Clinical Evaluation of Language Fundamentals-Revised (CELF-R)*. San Antonio, TX: Psychological Corp.

Shea, V. (1989). Peabody Picture Vocabulary Test-Revised. In C. S. Newmark (Ed.), *Major psychological assessment instruments* (Vol. II, pp. 271–283). Boston: Allyn & Bacon.

Sheslow, D., & Adams, W. (1990). *Wide Range Assessment of Memory and Learning* (*WRAML*). Wilmington, DE: Jastak.

Shiffrin, R. M., & Atkinson, R. C. (1969). Storage and retrieval processes in long-term memory. *Psychological Review*, *76*, 179–193.

Shinn, M. R. (Ed.) (1989). *Curriculum-based measurement: Assessing special children*. New York: Guilford.

Silver, L. B. (1993). Introduction and overview to the clinical concepts of learning disabilities. *Child and Adolescent Psychiatric Clinics of North America: Learning Disabilities*, *2*, 181–192.

Simpson, N., Black, F. W., & Strub, R. L. (1986). Memory assessment using the Strub-Black mental status examination and the Wechsler Memory Scale. *Journal of Clinical Psychology*, *42*, 147–155.

Slomka, G. T., & Tarter, R. E. (1993). Neuropsychological assessment. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of child and adolescent assessment* (pp. 208–223). Boston: Allyn and Bacon.

Sohlberg, M. M., & Mateer, C. A. (1990). Evaluation and treatment of communicative skills. In J. S. Kreutzer & P. Wehman (Eds.), *Community integration following traumatic brain injury*. Baltimore: Paul H. Brookes.

Strub, R. L., & Black, F. W. (1993). *The mental status examination in neurology* (3rd ed.). Philadelphia: F. A. Davis.

Swanson, H. L. (1981). Vigilance deficits in learning disabled children: A signal detection analysis. *Journal of Psychology and Psychiatry*, *2*, 339–398.

Swanson, H. L. (1995). Using the Cognitive Processing Test to assess ability: Development of a dynamic assessment measure. *School Psychology Review*, *24*, 672–693.

Swanson, H. L. (1996). *Swanson Cognitive Processing Test (S-CPT)*. Austin, TX: PRO-ED.

Talley, J. L. (1993). *Children's Auditory Verbal Learning Test-2 (CAVLT-2)*. Odessa, FL: Psychological Assessment Resources.

Tarver, S. G., & Dawson, M. M. (1978). Modality preference and the teaching of reading: A review. *Journal of Learning Disabilities*, *11*, 5–17.

Taylor, H. G. (1988). Learning disabilities. In E. J. Mash & L. G. Terdal (Eds.), *Behavioral assessment of childhood disorders* (2nd ed., pp. 402–450). New York: Guilford Press.

Taylor, H. G., & Fletcher, J. M. (1990). Neuropsychological assessment of children. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment* (2nd ed., pp. 228–255). New York: Pergamon.

Taylor, H. G., Fletcher, J. M., & Satz, P. (1984). Neuropsychological assessment in children. In G. Goldstein & M. Hersen (Eds.). *Handbook of psychological assessment* (pp. 211–234). New York: Pergamon.

Telzrow, C. F. (1985). The science and speculation of rehabilitation in developmental neuropsychological disorders. In L. C. Hartlage & C. F. Telzrow (Eds.), *The neuropsychology of individual differences: A developmental perspective* (pp. 271–307). New York: Plenum.

Templer, D. I., & Drew, R. H. (1992). Noncontact sports. In D. I. Templer, L. C. Hartlage, & W. G. Cannon (Eds.), *Preventable brain damage: Brain vulnerability and health* (pp. 30–40). New York: Springer.

Touyz, S., Byrne, D., & Gilandas, A. (1994). *Neuropsychology in clinical practice*. Boston: Academic Press.

Trenerry, M. R., Crosson, B., DeBoe, J., & Leber, W. R. (1990). *Visual search and attention test*. Odessa, FL: Psychological Assessment Resources.

Walsh, K. W. (1978). *Neuropsychology: A clinical approach*. New York: Churchill Livingstone.

Wechsler, D. (1987). *Wechsler Memory Scale-Revised manual*. San Antonio, TX: The Psychological Corporation.

Wedding, D., Horton, A. M., & Webster, J. S. (1986). *The neuropsychology handbook: Behavioral and clinical perspectives*. New York: Springer.

Wepman, J. M., & Reynolds, W. M. (1987). *Wepman's Auditory Discrimination Test* (2nd ed.) Los Angeles: Western Psychological Services.

Westby, C. (1988). Test review: Test of Language Development-2 Primary, Test of Language Development-2 Intermediate. *The Reading Teacher*, *42*, 236–237.

Whitten, J. C., D'Amato, R. C., & Chittooran, M. M.

(1992). A neuropsychological approach to intervention. In R. C. D'Amato & B. A. Rothlisberg (Eds.), *Psychological perspectives on intervention: A case study approach to prescriptions for change* (pp. 112–136). White Plains, NY: Longman.

Wiig, E. H., & Second, W. (1989). *Test of Language Competence-Expanded Edition (TLC)*. San Antonio, TX: The Psychological Corporation.

Williams, J. M. (1991). *Memory Assessment Scales (MAS)*. Odessa, FL: Psychological Assessment Resources.

Williams, R. E., & Vincent, K. R. (1991). Review of the Peabody Individual Achievement Test-Revised. In D. J. Keyser & R. C. Sweetland (Eds.), *Test critiques* (Vol. 8, pp. 557–562). Kansas City, MO: Test Corporation of America.

Wochnick Fodness, R., McNeilly, J., & Bradley-Johnson, S. (1991). Test–retest reliability of the Test of Language Development-2: Primary and Test of Language Development-2: Intermediate. *Journal of School Psychology*, *29*, 161–165.

Woodcock, R., & Johnson, M. B. (1989). *Woodcock–Johnson Psychoeducational Battery-Revised (WJPB-R)*. Chicago: Riverside.

Woodrow, H. (1921). Intelligence and its measurement: A symposium (XI). *Journal of Educational Psychology*, *12*, 207–210.

Ylvisaker, M., Szekeres, S. F., Haarbauer-Krupa, J., Urbanczyk, B., & Feeney, T. J. (1994). Speech and language intervention. In R. C. Savage & G. F. Wolcott (Eds.), *Educational dimensions of acquired brain injury* (pp. 185–235). Austin, TX: PRO-ED.

Ylvisaker, M., Szekeres, S. F., & Hartwick, P. (1994). A framework for cognitive intervention. In R. C. Savage & G. F. Wolcott (Eds.), *Educational dimensions of acquired brain injury* (pp. 35–67). Austin, TX: PRO-ED.

Youngjohn, J. R., Larrabee, G. J., & Crook, T. H. (1993). New adult age- and education-correction norms for the Benton Visual Retention Test. *The Clinical Neuropsychologist*, *7*, 155–160.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.10

Neuropsychological Assessment of Children

CYNTHIA A. RICCIO and CECIL R. REYNOLDS
Texas A&M University, College Station, TX, USA

4.10.1 INTRODUCTION 267
  4.10.1.1 Assessment Process 269
4.10.2 MEASURES USED IN THE ASSESSMENT OF CHILDREN 271
  4.10.2.1 Neuropsychological Interpretation of Children's Measures 272
  4.10.2.2 Development of New Measures for Children 273
  4.10.2.3 Current and Future Trends 274
    4.10.2.3.1 Memory 274
    4.10.2.3.2 Attention 275
    4.10.2.3.3 Computer-administered assessment 276
    4.10.2.3.4 Integration of neuroimaging and electrophysiology 276
    4.10.2.3.5 Integration of cognitive and developmental psychology 277
  4.10.2.4 Measurement Issues 277
  4.10.2.5 Approaches to Test Selection with Children 279
    4.10.2.5.1 Nomothetic approaches 280
    4.10.2.5.2 Idiographic approaches 283
    4.10.2.5.3 Combined approaches 285
  4.10.2.6 General Organization of the Neuropsychological Assessment of the Child 287
  4.10.2.7 Interpretation Issues 289
    4.10.2.7.1 Performance level 290
    4.10.2.7.2 Profile patterns 290
    4.10.2.7.3 Functional asymmetry 290
    4.10.2.7.4 Pathognomonic signs 291
    4.10.2.7.5 Combination approaches 291
4.10.3 CONCLUSIONS 291
4.10.4 REFERENCES 293

4.10.1 INTRODUCTION

The area of clinical neuropsychology has only recently been established as a viable specialty area (Woody, 1997). By definition, neuropsychology is the study of brain–behavior relationships that uses the theory and methodologies of both neurology and psychology. Historically, neuropsychology has been used for the diagnostic assessment of adults with known brain damage or injury; the clinical research has been derived from the study of adults with identified insult to the brain. The major premise of neuropsychological assessment is that different behaviors, including higher order cognitive skills, involve differing neurological structures or functional systems (Luria, 1980). As such, the neuropsychological approach to assessment involves assessment of various behavioral domains believed to be related to functional systems and making inferences about brain integrity based on the individual's performance across these domains. Neuropsychological assessment samples behaviors known to depend on the integrity of the central nervous system (CNS) using measures that correlate with cognitive, sensorimotor, and emotional functioning based on clinical research (Dean & Gray, 1990).

As more became known about brain–behavior relationships, clinical findings and theories were applied to the understanding of learning and behavior problems of adults where brain damage/injury was not identified. Thus, neuropsychological assessment as it is practiced today grew out of the need to clarify pathophysiological conditions where brain damage was not indicated by neurological, neuroradiological, or electrophysiological methods, in order to make differential diagnoses and provide information that would be useful in treatment planning and follow up (Dean & Gray, 1990). Based on the neuropsychological study of adults, this progressed to the application of neuropsychological methods and perspectives to the understanding of learning and other problems in children (L. C. Hartlage & Long, 1997).

Luria's (1970, 1980) theory, while based on adults, can be applied to children and adolescents. Neuropsychological techniques have been incorporated into the assessment of children for special education for some time (e.g., Haak, 1989; Hynd, 1981) with increasing interest by neuropsychologists in educational problems such as learning disability and attention deficit hyperactivity disorder (ADHD). The influence of theories specific to child psychology, school psychology, and education, are evident in the composition of neuropsychological assessment batteries, procedures, and measures used with children (Batchelor, 1996a). Increased interest and emphasis in the application of neuropsychology to educational issues and children may be due to a variety of factors: the emergence of clinical neuropsychology as a specialty area; advances in neuroscience and clinical evidence, specific to brain–behavior relationships, based on localized brain damage in childhood and youth; advances in technology (e.g., functional imaging) that are adding to the knowledge base regarding brain development and function; and continued research efforts specific to problems encountered by children and their neuropsychological functioning.

Several positive outcomes of the application of neuropsychology to children and adolescents have been identified. These include extending the range of diagnostic techniques available, providing for better integration of behavioral data (Dean, 1986; Gray & Dean, 1990; Obrzut & Hynd, 1983), increasing the ability to deal adequately with the multidimensionality of observed behavior, creating a unified or holistic picture of a student's functioning (Rothlisberg & D'Amato, 1988), and providing documentation of changes in behavior and development (Hynd & Willis, 1988). Clinical child neuropsychology provides a theoretical framework for understanding identified patterns of strengths and weaknesses, the relationships between strengths and weaknesses, and the extent to which these patterns remain stable or are subject to change over the course of development (Fletcher & Taylor, 1984; Temple, 1997). Increased understanding of the child's strengths and weaknesses can potentially be used to identify areas that may provide difficulty for the child in the future, as well as compensatory strategies or methods to circumvent these difficulties. It has further been argued that neuropsychological assessment of children can provide a better understanding of the ways in which neurological conditions impact on behavior and the translation of this knowledge into educationally relevant information (Allen, 1989). Although not all psychologists agree with the application of neuropsychological principles to children (see Riccio, Hynd, & Cohen, 1993), the growing significance of clinical child neuropsychology is evident in the increasing number of child clinical and school psychology graduate programs that offer coursework in neuropsychology (e.g., D'Amato, Hammons, Terminie, & Dean, 1992).

As a result of this growing interest in clinical child neuropsychology, the extent of knowledge available regarding the developing brain has increased dramatically since the 1980s. This includes advances in the understanding of typical development of neuropsychological functions (e.g., Ardila & Roselli, 1994; Halperin, McKay, Matier, & Sharma, 1994; Miller & Vernon, 1996; Molfese, 1995). Research has also begun to explore physiological processes and subcortical motivational systems that, together with environmental influences, are believed to impact on how the relevancy of information is determined and, ultimately, on the formation of cognitive representations in typically developing children (Derryberry & Reed, 1996). Additional advances in educational arenas have been made in the understanding of learning disabilities (e.g., Feagans, Short, & Meltzer, 1991; Geary, 1993; Obrzut & Hynd, 1983; Riccio, Gonzalez, & Hynd, 1994; Riccio & Hynd, 1995, 1996) as well as in the understanding of the short- and long-term problems associated with traumatic brain injury (e.g., Bigler, 1990; Snow & Hooper, 1994); the sequelae of neurological impairment of known causes such as lead poisoning, meningitis, and so on (e.g., Bellinger, 1995; Taylor, Barry, & Schatschneider, 1993); the impact of cancer treatment on CNS function (e.g., Copeland et al., 1988); and the short- and long-term sequelae in children identified as at-risk for learning problems due to perinatal or prenatal difficulties (e.g., Breslau, Chilcoat, DelDotto, & Andreski, 1996; Cohen, Beckwith, Parmalee, & Sigman, 1996; Gatten, Arceneaux, Dean, & Anderson, 1994; Saigal, 1995; Waber & McCormick, 1995).

While more research has focused on educational problems, increased risk for psychiatric disorder and long-term adjustment problems have been found to be associated with brain injury, both in adults and children (e.g., Breslau & Marshall, 1985; Rutter, Graham, & Yule, 1970; Seidel, Chadwick, & Rutter, 1975). Research consistently demonstrates that adjustment and behavioral problems are associated with children who have neurodevelopmental deficits (e.g., Hooper & Tramontana, 1997; Tramontana & Hooper, 1997). Children with neurological impairment have been found to be six times more likely to develop emotional, behavioral, or motivational problems secondary to, if not as a direct result of, the neurological impairment (Dean, 1986). At one time it was believed that specific relationships between brain dysfunction and child psychopathology would be found; it is now posited that the relationships between brain integrity and psychopathology are nonspecific and impacted by secondary influences including failure, frustration, social stigma, family reaction, and so on (Tramontana & Hooper, 1997). Many advances have been made in the area of psychopathology, including the development of models specific to the underlying neurological basis of ADHD (see Riccio, Hynd, & Cohen, 1996), autism (e.g., Damasio & Maurer, 1978; Hooper, Boyd, Hynd, & Rubin, 1993; Hurd, 1996; Maurer & Damasio, 1982; Shields, Varley, Broks, & Simpson, 1996), schizophrenia in childhood and adolescence (e.g., Asarnow, Asamen, Granholm, & Sherman, 1994; Asarnow, Brown, & Strandburg, 1995; Hendren, Hodde-Vargas, Yeo, & Vargas, 1995), conduct disorder (e.g., Moffitt, 1993), and anxiety (e.g., Gray, 1982). Various models (e.g., Gray, 1982; Kinsbourne, 1989; Nussbaum et al., 1988; Rourke, 1989; Tucker, 1989) have been proposed to explain the interface between brain function and behaviors associated with childhood psychopathology.

This chapter will provide an overview of the neuropsychological assessment process for children, both historically and in the context of current practices and future trends. Continuing concerns with regard to the translation to children and adolescents of what is known about adult functioning and neuropsychological assessment, as well as concerns with measurement and methodology, will be discussed. Finally, future directions and issues that need to be addressed in the neuropsychological assessment of children, if clinical child neuropsychology is to continue to add to the understanding of underlying processes in children's learning and behavior, as well as the application of that understanding to intervention programs, will be addressed.

4.10.1.1 Assessment Process

Neuropsychological assessment generally includes assessment of a number of functional domains that are, based on clinical evidence, associated with functional systems of the brain. This is considered important for the development of hypotheses and potential interventions (L. C. Hartlage & Telzrow, 1986; Whitten, D'Amato, & Chitooran, 1992). Areas evaluated generally include cognition, achievement, and behavior/personality/emotionality, as would be assessed as part of a general psychological evaluation. A neuropsychological evaluation provides for consideration of a wider array of functions, however, than is addressed in a typical psychological or psychoeducational evaluation (Dean, 1985, 1986; Obrzut, 1981). In general, the neuropsychological evaluation is more thorough and also includes the assessment of perceptual, motor, and sensory areas, and of attention, executive function (planning, organization), and learning/memory (e.g., Dean & Gray, 1990; Obrzut, 1981; Shurtleff, Fay, Abbot, & Berninger, 1988).

Given that the neurodevelopment of Luria's functional systems and the experiences of the child interact in a reciprocal manner (Spreen, Risser, & Edgell, 1995), as well as the potential for adjustment/behavioral difficulties, the use has been advocated of a transactional model that takes into consideration the reciprocal interactions of the child, home and family members, classroom (teacher and peers), and other social environments in which the child functions (Batchelor, 1996b; D'Amato & Rothlisberg, 1996; D'Amato, Rothlisberg, & Leu, in press; Teeter, 1997; Teeter & Semrud-Clikeman, 1997). This should incorporate information from a variety of sources (e.g., parents, teachers, physicians, medical records, school records, and so on) in order to enable cross-comparison (Batchelor, 1996b). In addition, it has been suggested that motivational factors (Batchelor, 1996b) and the child's ability to cope with the injury/impairment need to be determined (Dean, 1986). Thus, the neuropsychological assessment process not only incorporates a more complete review of information regarding the child but attempts to integrate this information with an understanding of brain–behavior relations and environmental factors (Taylor & Fletcher, 1990).

This means that a neuropsychological assessment involves a wide range of tasks focused on the child, as well as measures/observations of the various contexts in which the child functions and the associated expectations. Some critics of neuropsychological assessment have argued that so extensive an evaluation is not time- or cost-effective (e.g., Little & Stavrou, 1993). For example, neuropsychological assessment of a learning disability goes beyond identifying the academic deficit(s) to the identification of the child's processing strengths and deficits as well as the child's ability to function in a variety of contexts (Morris, 1994). The assessment of a wider range of higher cortical functions is supported by research findings that neurological disorders are seldom expressed as a single dysfunction (Dean & Gray, 1990), and it has been shown to improve differential diagnosis of learning problems (D'Amato, Rothlisberg, & Rhodes, 1997; Morris, 1994; Rourke, 1994). It is further argued that the process of deriving hypotheses for intervention planning is a complex process that requires a comprehensive assessment battery coupled with neuropsychological foundations and familiarity with the contexts and task demands of the child (Gaddes, 1983; Rourke, 1994). The cumulative performances of the child on neuropsychological measures are seen as behavioral indicators of brain function (Fennell & Bauer 1997). Based on all of the data generated in the evaluation process, hypotheses are generated which are specific to how and why a child processes information (D'Amato, 1990; Dean 1986; Leu & D'Amato, 1994; Whitten et al., 1992). Inferences are then made based on the child's performance on a variety of measures and the theoretical perspective of the clinician.

Much of the skepticism regarding the application of neuropsychology to problems of learning and behavior in children has centered on the assessment–intervention interface (Little & Stavrou, 1993; Samuels, 1979; Sandoval & Halperin, 1981). It is argued by some that neuropsychological perspectives do not add to the ability to develop remedial and treatment programs and may even lead to a sense of hopelessness (Little & Stavrou, 1993). Others have argued that the information obtained from neuropsychological assessment can be used as the basis for developing appropriate intervention programs (Gaddes & Edgell, 1994; Reynolds, 1981b; Rourke, 1991, 1994; Teeter & Semrud-Clikeman, 1997). Data about the additional areas of functioning included in the neuropsychological assessment are needed to support inferences about the integrity of various functional systems of the brain (Shurtleff et al., 1988). The neuropsychological perspective leads to better understanding of underlying causes of learning and behavior problems; this in turn results in an increased ability to develop appropriate interventions or circumvent future problems (D'Amato et al., 1997).

Ultimately, data generated from the neuropsychological assessment process are used to develop recommendations regarding whether the individual would profit from compensatory strategies, remedial instruction, or a combination of approaches (Gaddes & Edgell, 1994). Through the use of information about how various skills correlate in the developmental process, neuropsychological assessment allows one to make inferences not only about those skills measured, but also about skills that have not been evaluated. Further, by understanding the neurological correlates of these skills and of instructional methods, neuropsychological assessment can assist in the formulation of hypotheses regarding potential instructional methods/materials for a particular child (Reynolds, Kamphaus, Rosenthal, & Hiemenz, 1997). For example, based on the neuropsychological evaluation of two children with autism, differing nonverbal teaching strategies were identified for each child in order to improve their individual outcomes (Hurd, 1996).

While "treatment" within the framework of education is generally considered to consist of eligibility, placement decisions, and the development of an educational plan, "treatment" resulting from a neuropsychological assessment frequently includes assistance with specific medical management, vocationally related goals, speech/language areas, and physical issues (Cohen, Branch, Willis, Weyandt, & Hynd, 1992; Dean & Gray, 1990). Effective interventions need to take into consideration the myriad psychosocial contexts in which the child functions and adjustment and motivational issues, and to identify those environmental modifications that can ameliorate or reduce the behavioral effects of brain dysfunction (Batchelor, 1996b; Teeter & Semrud-Clikeman, 1997). It has been suggested that the interventions developed must also be multidimensional and incorporate not only academic, behavioral, and psychosocial techniques, but also include motivational, metacognitive, medical, and classroom management techniques (Batchelor, 1996b; Teeter & Semrud-Clikeman, 1997). However, correct diagnosis and early implementation of treatment strategies that work have been shown to be cost-effective in dollars and in the quality of a child's life (Reynolds, Wilen, & Stone, 1997).

4.10.2 MEASURES USED IN THE ASSESSMENT OF CHILDREN

As previously noted, the application of neuropsychological theory and assessment with children was derived from applications with adults. In the development of clinical child neuropsychology, historically, one of the basic avenues used in determining the assessment measures and processes to be used consisted of modifying, for use with children, existing neuropsychological batteries and other measures already used for adults (L. C. Hartlage & Long, 1997). In some cases, this involved modifying some tasks in the battery or adding tasks. An alternative strategy involved collecting some normative data on children for existing tasks. Both of these strategies were based on the clinical efficacy of the measures with adults, not with children, and on the assumption that tasks for adults measure the same thing when used with children. Similarly, in the assessment and hypothesis generation process, it is tempting to assume that neuropsychological findings from adults will be useful with children; however, this has not been shown to be a valid assumption.

When applied to children and adolescents, the premise that behavior can be used to make inferences about brain function and integrity has to be expanded to include consideration of neurodevelopmental differences that exist as a function of the age of the child. To directly apply adult inferences/hypotheses to children ignores what is known about changes in the functional organization of the brain as children grow (Cohen et al. 1992; Fletcher & Taylor, 1984). Research has provided evidence of age-based differences in children for verbal memory (Kail, 1984; Miller & Vernon, 1996), language (Segalowitz, 1983) and right hemisphere functions (Bakker, 1984; Wittelson, 1977). Recent research, for example, found that the relationship of memory, general intelligence, and speed of processing in children was not consistent with adult models (Miller & Vernon, 1996). Research has also suggested that typologies generated from the Halstead–Reitan Neuropsychological Battery (HRNB) used with children differed from typologies generated with adults, in that the child groups were more homogeneous, but provided less coverage (28–42%) as compared to adult typologies (Livingston et al., 1997). Because of neurodevelopmental changes, it is also not possible to view brain dysfunction on a continuum based on behavioral deficits as these may change over time (Fletcher & Taylor, 1984).

Further, there is often an over-reliance on signs of dysfunction in adults as reflecting pathology in children when these may be developmental. For example, on clock-face drawing in adults, identification of unilateral spatial neglect has been of particular interest (e.g., Heilman, Watson, & Valenstein, 1985; Mesulam, 1985). Developmental study of clock-face drawing with children found that this hemispatial neglect was developmental and not infrequent through the age of seven years (Edmonds, Cohen, Riccio, Bacon, & Hynd, 1993). It was concluded that this developmental pattern was consistent with the development of the frontal lobes and planning ability in children.

Neurodevelopment follows an ontogenetic course with primary cortical zones generally mature by birth (Luria, 1980). Secondary and tertiary areas continue to develop postnatally. These include the integrative systems involved in the higher order functions of learning, memory, attention, emotion, cognition, and language as well as the association areas. The association areas are the last of these areas to develop and myelinate (Goldman & Lewis, 1978; Goldman-Rakic, 1987). Vygotsky (1980) suggested that not only is there continued development of secondary and tertiary areas, but that the interaction of primary, secondary, and tertiary areas is likely to change with chronological age (Merola & Leiderman, 1985; Rutter, 1981). Although the developmental sequence for the formation of neural pathways and the myelination of specific locations corresponding to specific behaviors have been identified, these do not correspond directly to models of cognitive development (Spreen et al., 1995).

Knowledge of typical neurodevelopmental progress has increased since the 1980s; however, most of what is practiced today, as well as the theoretical bases in neuropsychology, is grounded on observations and informal assessment of individuals with identified brain damage (Reynolds, 1997b). Extensive research regarding typical neurodevelopment, particularly in relation to higher order cognitive skills, is limited, and the changing organization over time of brain function in children is only beginning to be understood (Hynd & Willis, 1988). Thus, there are still many unanswered questions regarding the developmental progression of many functional systems, particularly at the associative and integrative levels, and concerning how the neurodevelopmental progression maps onto the cognitive functioning observed.

It is often assumed, for example, based on earlier theory, that children reach adult levels of performance at 8–10 years of age. For example, Luria (1966) suggested that the frontal lobes become functional between the ages of four and seven years. This in turn led to the assumption that executive functioning would approach adult levels by age 8–10 years. It has been suggested that the greatest period of frontal lobe development occurs at the six- and eight-year-old levels, which is consistent with Luria's initial hypothesis (Passler, Isaac, & Hynd, 1985). Subsequent research, however, has demonstrated that the development of frontal lobe functioning continues at least through age 12 and possibly through age 16 (e.g., Becker, Isaac, & Hynd, 1987; Chelune & Baer, 1986; Levin et al., 1991; Welsh, Pennington, & Grossier, 1991). Further, while cognitive ability does not appear to be a factor for particular measures of frontal lobe functioning after age 12, it has been suggested that cognition can impact performance on frontal lobe measures in younger children (Chelune & Thompson, 1987; Riccio, Hall, et al., 1994). Thus, it is important to first have a strong foundation of understanding of the normal neurodevelopmental course before it is possible to interpret accurately and differentiate behaviors that represent an alteration or deviance from expected neurodevelopment.

Not only do neurodevelopmental courses need to be considered, there are complex differences between children and adults in the mechanisms of brain pathology that lead to neuropsychological and behavioral/affective problems and these do not necessarily follow a similar progression in children as for adults (Fennell & Bauer, 1997; Fletcher & Taylor, 1984). The developing brain of the child needs to be considered in that the impact of neurological insult is influenced by age as well as location and nature of injury, gender, socioeconomic status, level of emotional adjustment and coping, and the individual's own adaptive skills (Bolter & Long, 1985). With the development of the child occurring on a continuous basis and at a rapid rate, it is often difficult to obtain sufficient consistency from the premorbid history (Batchelor, 1996a). Accurate estimation of premorbid ability levels is best obtained from previous individualized standardized cognitive or achievement assessment, or if this is unavailable, from results of group-administered standardized data from school records with some consideration for potential regression effects (Reynolds, 1997c). For young children, this information is not generally available. Prenatal and perinatal, as well as postnatal developmental histories may be inaccurate, incomplete, or unknown, particularly in very young children (Batchelor, Gray, Dean, & Lowery, 1988; Gray, Dean, & Rattan, 1987). Even in school-aged children, teacher reports, grades, and so on may result in inaccurate estimations of premorbid ability (Reynolds, 1997c).

Given the different mechanisms and progression involved in the pathology, it is clear that the inferences drawn from and the interpretations of neuropsychological performance need to be different for adults and children. For children, the nature and persistence of learning problems is dependent on the status of development of various brain structures, the effects of the injury/insult, and the interactions between functional and dysfunctional neurological systems, as well as genetic and environmental influences (Teeter & Semrud-Clikeman, 1997).

Neuropsychological assessment of children and adolescents requires not only tests/measures that are age-appropriate and have sufficient empirical support for the inferences being made between neurological substrates and the behavioral performance of the child, but the generation of inferences also needs to take into consideration these developmental issues (Cohen et al., 1992). Further, it is important to document the sensitivity of the measures to neurobehavioral and neurodevelopmental functioning in children (Fletcher & Taylor, 1984). Although the measures are derived predominantly from neuropsychological study and clinical evidence regarding adults with known brain injury, a developmental perspective needs to be maintained in the application of neuropsychology to children (Hooper & Tramontana, 1997). Unfortunately, many of the measures used with adults do not have the sensitivity necessary to reflect developmental issues and, as a result, the utility of procedures used with adults in the neuropsychological assessment of children has multiple pitfalls and has been questioned (e.g., Cohen et al., 1992; Fletcher & Taylor, 1984).

4.10.2.1 Neuropsychological Interpretation of Children's Measures

Another approach to applying neuropsychological principles in the assessment of children took measures already in use for children (e.g., standardized intelligence tests) and interpreted these measures from a neuropsychological perspective; where existing child measures did not exist, these measures were then developed. L. C. Hartlage and Long (1997) indicated that most practitioners preferred this method (interpreting child-based measures from a neuropsychological perspective) as opposed to using adult measures with child norms. As with adults, this has occurred most frequently with the Wechsler scales. General summary scores of Wechsler scales have been found to be reliable indicators of brain integrity (Black, 1976; Hynd & Willis, 1988). Various subtests of the WISC-R also have been found to correlate with neuropsychological measures (see Batchelor, Sowles, Dean, & Fischer, 1991) and have been used to formulate hypotheses (Kaplan, 1988). Multiple efforts have been made with regard to recategorizing or clustering various subtests to provide for neuropsychological interpretation of the WISC-R. L. C. Hartlage (1982), for example, suggested that the functional integrity of the right and left hemispheres could be estimated by comparing the Similarities and Picture Arrangement subtests (temporal lobe) and the Arithmetic and Block Design subtests (parietal lobe). Bannatyne (1974) proposed four categories of neuropsychological function that could be assessed and interpreted based on combinations of subtests on the WISC-R: verbal comprehension, sequencing, spatial, and acquired knowledge. Kaufman (1979) recategorized the subtests into successive and simultaneous tests, based on Luria's theory.

Concerns with this practice have been evidenced in the literature. Interpretations based on isolated measures of a child's behavior (e.g., a single subtest) have limited reliability and validity (Kamphaus, 1993; Lezak, 1995) and this is often what occurs in this process. Recategorizations of multiple subtests (e.g., Bannatyne, 1974; Kaufman, 1979), appear to have greater reliability, but the validity of these recategorizations continues to be questionable (see Kamphaus, 1993). Further, in many cases there is no attempt to translate the inferences made, using these methods, into effective interventions.

4.10.2.2 Development of New Measures for Children

As opposed to trying to "make do" with existing children's measures, additional measures have been developed with an underlying neuropsychological basis. For example, the Luria–Das model of successive/simultaneous processing (Das, Kirby, & Jarman, 1979) in conjunction with the cerebral lateralization research by Sperry (1968, 1974), Kinsbourne (1975), and others, served as the basis for the development of the Kaufman Assessment Battery for Children (KABC; Kaufman & Kaufman, 1983a). As such, the design of the KABC is compatible with current neuropsychological models of higher order cognitive function (Reynolds & Kamphaus, 1997). Unlike the Wechsler scales, where mode of presentation determines the scale with which a task is associated, on the KABC the cognitive processing demands of the task (e.g., simultaneous or sequential) determine the scale with which it is associated (Kaufman & Kaufman, 1983b). Further, rather than conceptualizing lateralization based on content or method of presentation, the lateralization component of the KABC is based on the way in which the information is processed or manipulated. Within each scale, there is a variation of mode of presentation and response that allows for further evaluation of complex functional systems (Reynolds & Kamphaus, 1997). KABC interpretation is intended to identify cognitive neuropsychological strengths of the child, and the related instructional methods and learning activities that will exploit these strengths and circumvent deficit areas. Research on the effectiveness of this model for intervention is, however, limited.

Evaluation of the KABC with regard to its relevance to Luria's approach and to child neuropsychology has been positive (e.g., Donders, 1992; Majovski, 1984; Snyder, Leark, Golden, Grove, & Allison, 1983). It has been suggested that the KABC is a good complement to other neuropsychological tests. Specifically with regard to the use of the KABC as a component of a neuropsychological battery, it has been shown to provide useful information in the differential diagnosis of learning disability subtypes (e.g., Hooper & Hynd, 1985; Telzrow, Century, Harris, & Redmond, 1985) and right hemisphere dysfunction, which is consistent with physical evidence (Morris & Bigler, 1985; Shapiro & Dotan, 1985). Similar positive results were found in the comparison of dichotic listening performance and KABC results (Dietzen, 1986). Thus, the KABC has been shown to be sensitive to traumatic brain injury to specific cortical regions (Donders, 1992). Research also indicated that the pathognomonic and intellect scales of the Luria Nebraska Neuropsychological Battery-Children's Revision were closely related to performance on the global scales of the KABC (Leark, Snyder, Grove, & Golden, 1983). Research results overall tend to support the use of the KABC in neuropsychological assessment, and subtests of the KABC are frequently used in eclectic batteries (e.g., Nussbaum et al., 1988; Branch, Cohen, & Hynd, 1995).

The KABC may well be the test of choice for children under age five (Reynolds et al., 1997); the use of sample and teaching items adds to the likelihood that a neurological substrate or functional system is being assessed as opposed to language, experience, or culture (Reynolds & Kamphaus, 1997). The KABC has strong validity and reliability (Kamphaus, 1993), is sensitive to developmental changes in information processing/functional organization (Reynolds & Kamphaus, 1997), and is considered an appropriate instrument for use with US ethnic minorities (e.g., Fan, Willson, & Reynolds, 1995; Kamphaus & Reynolds, 1987). While further research with the KABC in conjunction with neuropsychological assessment is needed, available research supports the potential for the KABC to be a useful tool for child neuropsychologists with results providing implications for the habilitation of learning problems (Reynolds & Kamphaus, 1997).

4.10.2.3 Current and Future Trends

The development of new measures, specifically designed and normed for children may not only reflect current interest areas in children's learning and behavior, but may in many ways dictate the future directions of neuropsychological assessment of children. In particular, since the late 1980s a number of measures have been developed which are specific to memory and attention. At the same time, there is also an increase in the use of technology, with or without the inclusion of electrophysiological or imaging methods, which is evident in the research literature and clinical practice.

4.10.2.3.1 Memory

Nearly every disorder that involves the CNS and higher cognitive functions includes some form of memory complaint; memory is incorporated in almost all daily activities (Reynolds & Bigler, 1997a). Research across neurological disorders points to the importance of memory in evaluating brain integrity (Reynolds & Bigler, 1997a); 80% of a sample of clinicians who performed testing noted memory as important (Snyderman & Rothman, 1987). Standard psychoeducational batteries used with children tend to focus solely on cognitive ability as defined by IQ, achievement, and behavioral status. In the area of learning disabilities, there has been recent interest in examining the underlying psychological processes, and particularly learning and memory (Zurcher, 1995). Research in the area of memory and the development of new measures to assess memory functions may lead to further interest in the learning process itself (Reynolds, 1992). It has been argued, additionally, that the assessment of learning and memory would provide useful information for instructional planning (Wasserman, 1995).

Historically, assessment of memory in children relied on the use of subtests from various tests including the KABC, the WISC-III and its earlier versions, and so on (e.g., Nussbaum et al., 1988; Branch et al., 1995). All too frequently, inferences regarding verbal memory in particular relied on the Digit Span subtest of the Wechsler scales. Multiple concerns about relying on Digit Span can be found in the research literature (e.g., Reynolds, 1997a; Talley, 1986). Recent research in the area of memory has suggested that the traditional combining of forward and backward digits may be inappropriate and that these tasks represent quite different cognitive demands (Ramsey & Reynolds, 1995; Reynolds, 1997a) with distinct neuropsychological substrates. Initial findings suggest, for example, that forward memory span may be more directly impacted by attention while backward memory span may be more a reflection of general intelligence. Additional investigation into the distinction between forward and backward memory span, as well as into other areas of memory continues to be needed. Due to the increased interest in this area, children's norms for measures used in the assessment of memory in adults have been developed (e.g., Delis, Kramer, Kaplan, & Ober, 1994). In addition, three comprehensive measures for the assessment of memory/learning have been developed specifically for children and adolescents since the mid-1980s. The development of these measures has in many ways been due to the perceived inappropriateness of adult measures of memory for use with children and the inability to relate results from adult measures to the contexts (e.g., school) in which children function.

The first of the measures developed for the assessment of memory in children, the Wide Range Assessment of Memory and Learning (WRAML; Sheslow & Adams, 1990), consists of 12 subtests which yield verbal memory and visual memory scores, with normative data for children ages 5–17 years. Delayed recall trials can be given for four of the subtests. Initial factor analysis of the WRAML corroborated the two-factor structure (Haut, Haut, Callahan, & Franzen, 1992); however, with at-risk children and a clinical population, three factors were extracted (Aylward, Gioia, Verhulst, & Bell, 1995; Phelps, 1995). Some concern has been voiced with regard to the multiple items/tasks that may tap attention as opposed to memory and the absence of consideration of attention/concentration (Haut et al., 1992). Further, evaluation of the WRAML for children with, compared with those without, ADHD or learning disabilities indicated that the WRAML provided little additional information for discriminating between clinical groups (Phelps, 1996).

The Test of Memory and Learning (TOMAL; Reynolds & Bigler, 1994) consists of 10 core subtests (five verbal and five nonverbal) yielding separate verbal memory and nonverbal memory scale scores as well as a composite memory score. A delayed recall procedure can be implemented to provide a delayed recall index. Additional supplemental indices (e.g., sequential recall, free recall, attention/concentration, and learning) can also be computed. Using a variety of factor analytic methods, Reynolds and Bigler (1996) examined the latent structure of the TOMAL. Factor analytic study of the TOMAL indicated that the factor solutions obtained were highly stable across all age groups. Notably, none of the solutions obtained matched the verbal–nonverbal dichotomy usually considered and represented by the two scales of the TOMAL. Instead, what emerged were components representing various levels of complexity in memory tasks and processing demands that cut across modalities. Alternative methods of interpretation based on the factor analytic results are available (see Reynolds & Bigler, 1996). The TOMAL does provide separate scores for forward and backward recall, in contrast to many scales that combine these inappropriately. Unlike most neuropsychological measures (Reynolds, 1997b), the TOMAL included studies of ethnic and gender bias during standardization; items showing cultural biases were eliminated.

Most recently, the Children's Memory Scale (CMS; Cohen, 1997) was developed with linkages to the WISC-III built in to the standardization process. The composition of the CMS was based on extensive clinical practice with initial tasks and items, field trials of the measures, and feedback from clinicians involved in the field trials. The CMS consists of six core subtests representing verbal memory, attention/concentration, and visual/nonverbal memory as well as three supplemental subtests. The CMS provides for evaluation of immediate recall as well as delayed recall of the verbal and nonverbal memory areas. For scoring purposes, seven index scores can be calculated to examine differences between immediate/delayed verbal/visual memory, learning, recognition, and attention/concentration. Factor analytic studies of the standardization sample were conducted and four models evaluated to determine the "best fit." Results indicated that the three-factor solution (attention/concentration, verbal memory, visual memory) was the most consistent (Cohen, 1997).

4.10.2.3.2 Attention

It has been argued that the most frequent symptoms associated with childhood neuropsychological disorder include attention/concentration, self-regulation and emotional/behavioral problems (Nussbaum & Bigler, 1990). Further, it is the neural traces left by attention that are likely the root of memory. It is not surprising that there is increased interest in the measurement of attentional processes or that these are seen as an important component of the assessment process. The assessment of attention, more so than of other domains, has moved to computerized approaches. The most comprehensive battery of computerized measures is the Gordon Diagnostic System (GDS; Gordon, 1983). This is a microcomputer-based assessment that includes 11 tasks specific to attention and self-regulation. Since the development of the GDS, a number of other computer-based measures of attention and impulsivity have been developed and marketed. These programs tend to vary with regard to the actual paradigm used; there are variations in the modality employed, the type of stimuli, and the nature of the task (Halperin, 1991).

Continuous performance tests (CPTs), for example, may require a response only when a specified target stimulus is presented (if X) or only when the target stimulus follows another specified stimulus (if AX) and so on. A further variation of this is a similar task where the required "response" to the presentation of the target stimuli is, however, to inhibit responding (Conners, 1995). The stimuli may be presented in a visual or auditory format, or in a combination format requiring a modality shift. Also, depending on the program used, the scores may be limited to correct responses, commission errors, and omission errors, or may include reaction time information.

Through the use of computerized measures of attention, knowledge specific to the developmental nature of attentional processes has been gleaned (Mitchell, Chavez, Baker, Guzman, & Azen, 1990). Research has demonstrated the usefulness of computerized measures of attention and self-regulation for monitoring the effects of medical management (e.g., Barkley, DuPaul, & McMurray, 1991; Barkley, Fischer, Newby, & Breen, 1988; Hall & Kataria, 1992). It was anticipated that computerized assessment of attention would provide more objective data in the assessment process for ADHD as well as providing information specific to attentional deficits associated with traumatic brain injury or other neurological disorders (Timmermans & Christensen, 1991). The results of studies with various paradigms for CPTs are equivocal with regard to discriminant validity specific to ADHD (e.g., Barkley et al., 1991; Wherry et al., 1993) as well as concerning the extent to which results are consistent with teacher perceptions (Barkley, 1991, 1994). Interpretation of these measures is limited by the availability of comprehensive research with any one software program. The extent to which cultural differences, gender differences, cognitive ability, order effects, and so on impact on CPT performance is unknown. Further, the extent to which the particular paradigms used provide predictive information that may be helpful in intervention planning has not been studied.

4.10.2.3.3 Computer-administered assessment

Measures of attention are not the only computer-based assessment tools. A computerized neuropsychological test battery for adults has been developed (Powell et al., 1993), computer-administered interviews and self-reports are available, and specific neuropsychological tests or their analogs can be administered via computer (e.g., Burin, Prieto, & Delgado, 1995; Heaton, 1981). Computerized assessment of children's reading skills has been investigated with indications of high coefficients of equivalence with traditional assessment (Evans, Tannehill, & Martin, 1995). With advances in microcomputers, the use of computerized assessment will likely increase in the near future.

The use of computers and technology in assessment has a number of advantages and clearly allows for the development of an increasing variety of tasks without excessive and cumbersome testing materials; computerized assessment may be less time-consuming and, as such, cost- and time-effective. Further, the speed or measure of time to task completion is considered one of the most sensitive indices in neuropsychological assessment and computer programs can provide increased accuracy in the measurement of speed of processing (Kane & Kay, 1997). Kane and Kay (1997) point out a number of additional advantages to the use of computers in the assessment process, including presentation of items at a fixed rate (computer-paced) as well as providing for accurate measure of time to completion (child-paced). Computers can also be used to generate multiple forms of a test, thus providing baseline data as well as a means of monitoring change over time. With computerized assessment, standard/uniform administration is ensured and results are free of potential bias. Computers further facilitate the production of relevant test statistics (Kane & Kay, 1997).

There are however, multiple concerns and disadvantages with "diagnosis by computer." Predominant among these is the loss of information from not being able to observe the process and strategy used by the individual in reaching the solution (Powell, 1997). First (1994) concluded that computerized assessment processes were advantageous, but cautioned that clinicians must continue to be a strong component in the diagnostic process in order to provide for diagnostic validity. Computers cannot replace the information gained from interaction and clinical observation of process nor can a computer draw conclusions regarding level of attention, motivation, fatigue, and so on that may be cues to discontinue testing for a brief period. Computers also cannot provide the child with prompts and encouragement as needed to maintain performance over time (Kane & Kay, 1997). At the extreme, there is the potential for computers to be used as a substitute for a complete evaluation and this is of concern (Kane & Kay, 1997). First (1994) asserted that clinicians needed to be well advised of the limitations as well as the strengths of computerized assessment procedures. As the number of computer-driven assessments increases, there will need to be an analogous increase in the research field comparing the various programs and their psychometric properties with each other and with more traditional methods of assessments. At the time of writing, in the late 1990s, many computerized assessment methods fail to meet established testing standards (Kane & Kay, 1997).

4.10.2.3.4 Integration of neuroimaging and electrophysiology

With advances in neuroscience, clinical and research protocols may more frequently include neuroradiological methods in conjunction with neuropsychological techniques in order to enhance understanding of childhood disorders. This type of "partnership" is already occurring in a number of research areas (e.g., Bigler, 1991; Denckla, LeMay, & Chapman, 1985; Duffy, Denckla, McAnulty, & Holmes, 1988; Hynd, Marshall, & Semrud-Clikeman, 1991). The integration of information from neuroradiology with neuropsychological assessment has already established relationships for specific lesions and associated behaviors and is beginning to establish a better understanding of the relationship between myelination differences and white/gray matter ratios (e.g., Harbord et al., 1990; Jernigan & Tallal, 1990; Turkheimer, Yeo, Jones, & Bigler, 1990). The availability of imaging using ultrasound has added to the knowledge of relationships between gross abnormalities evident *in vitro* and later negative outcomes (e.g., Beverley, Smith, Beesley, Jones, & Rhodes, 1990; Iivaneihan, Launes, Pihko, Nikkinen, & Lindroth, 1990). Measurement issues in imaging, such as differences in resolution from one magnetic resonance image to another, continue to be problems in this area, but will hopefully be resolved in the future. While in the past routine EEG of children did not offer much utility in the evaluation of learning or behavior problems, the development of computer-assisted analysis has improved the interpretability of electrophysiological measures (Duffy & McAnulty, 1990).

Computerized measures have been developed to examine more closely the speed of information processing, through reaction time paradigms that have included linguistic (e.g., Lovrich, Cheng, & Velting, 1996) as well as visual stimuli (Novak, Solanto, & Abikoff, 1995), in conjunction with electrophysiological measures.

This integration of methods across neuroscience and neuropsychology is providing further evidence concerning brain–behavior relationships and adding to the knowledge base related to neurodevelopmental processes in children and adolescents. Functional imaging and other imaging quantification methods hold promise for furthering the future understanding of neuropsychological performance (Bigler, 1996). Similarly, it has also been argued that a comprehensive and integrative assessment process, that involves both the neurologist and neuropsychologist with the tools and expertise of both disciplines, may enhance the value and role of neuropsychological assessment (Batchelor, 1996a, 1996b).

4.10.2.3.5 Integration of cognitive and developmental psychology

Neuropsychological assessment of children is being influenced more and more by developmental and cognitive psychology. This is most apparent in the areas of language, attentional processes, and executive functions (Williams & Boll, 1997). Integration across fields has been suggested specifically with regard to metacognition (from cognitive psychology) and executive function (Torgesen, 1994). The domain of "executive function" may incorporate a variety of constructs (e.g., attention, self-regulation, working memory) but the "executive" processes generally focus more on effortful and flexible organization, strategic planning, and proactive reasoning (Denckla, 1994). Denckla further asserted that executive function cannot be dealt with as a "composite" of scores on various measures, but must be fractionated. The measurement of executive function in children is exceptionally difficult due to the ongoing development and maturation of the frontal lobes through adolescence. Factor analytic study of executive function tasks (Welsh et al., 1991) yielded factors that appeared to be divided according to developmentally related constructs as opposed to theoretical ones (Denckla, 1994). The majority of measures for executive function which are used with children are downward extensions of adult measures and many lack sufficient normative data and psychometric study. In addition, the emphasis on the use of novel tasks in the assessment of executive function places significant limitations on the evaluation of interventions in the area of executive processes. Torgesen (1994) argued that current measures of executive function, with the presumed assumption for a need for novelty, evidence a lack of cross-theoretical integration between neuropsychology and the information-processing paradigms. He further stated that there is a need to include assessment of tasks that are ecologically based and require executive function, in order to enhance the evaluation of treatment programs designed to remediate executive processes. Certainly, the production of child-centered, developmentally sensitive measures of executive processing, that are more directly linked to real-life activities thus facilitating the development of interventions, and that have sufficient flexibility to allow for pre- and postevaluation, is needed. Overall, integration of neuropsychological assessment and models of cognitive development may lead not only to a better understanding of deficit processes but also to better remediation/habilitation programs (Williams & Boll, 1997).

4.10.2.4 Measurement Issues

Although research methods and statistical tools have greatly improved since the early 1970s, clinical child neuropsychology has been criticized for its failure to attend to principles of research and to incorporate psychometric advances (Cicchetti, 1994; Parsons, & Prigatano, 1978; Reschly & Gresham, 1989; Ris & Noll, 1994; Sandoval, 1981; Willson & Reynolds, 1982). Problems with statistical methods and design in clinical neuropsychology have been frequently noted (e.g., Adams, 1985; Dean, 1985; Reynolds, 1986a, 1986b, 1997b).

One major concern relates to the extent and nature of normative data for many measures used in the neuropsychological assessment of children. Although clinical insight may be gained by observation of test performance, sound normative data provides a backdrop against which to evaluate that insight (Reynolds, 1997b). The systematic development and presentation of normative data across the lifespan for many tools used in neuropsychological assessment have received far too little attention to date (Reynolds, 1986b) and greater attention in this area is needed. Good normative data require extensive systematic and stratified sampling of a population in order to obtain a reliable standard against which to judge the performance of others. The provision of adequate normative data has multiple benefits for the field of neuropsychology, including improved communication among clinicians and researchers, increased accuracy in diagnosis, and facilitation of training for new members to the discipline. In addition, good normative data provide the opportunity to deflate and expose a variety of clinical myths (Reynolds, 1986b).

Most of what is known about the measures used is specific to the performance of those with identified brain injury/insult as opposed to typically developing individuals. Normative data that are available in the literature are often based on small samples, may have been collected in a single geographical region, and do not reflect the ethnic diversity, socioeconomic levels, or gender composition of the general population. In many cases, the sample is predominantly male and Caucasian, yet research suggests that gender and cultural differences may also contribute to variations in brain organization (e.g., McGlone & Davidson, 1973).

The lack of sufficiently large, stratified samples in the development and standardization of neuropsychological assessment inhibits the understanding of demographic influences (Reynolds, 1997b) and thus complicates test interpretation (Dean, 1985). That cultural differences exist on standardized measures is well documented. Mostly, the use of neuropsychological measures with Hispanic populations has been studied (e.g., Ardila & Roselli, 1994; Ardila, Roselli, & Putente, 1994; Arnold, Montgomery, Castenada, & Langoria, 1994), but overall research on the effects of cultural differences (e.g., differences in the value of speed of responding) is sparse. Differences between ethnic groups have also been examined with respect to specific measures of memory (e.g., Mayfield & Reynolds, 1997). However, for most neuropsychological measures there has been no study of ethnic and gender differences; cultural/ethnic differences are infrequently accounted for in the collection of normative data and therefore cannot be used in the interpretation process.

All too frequently, neuropsychologists rely on the "clinical" nature of the test and overlook the psychometric concepts of reliability and validity. The need for the establishment of reliability of neuropsychological measures has been cited in the literature (e.g., Parsons & Prigatano, 1978, Reynolds, 1982); reliability information on neuropsychological measures is not routinely reported in research studies and is frequently not included in the test manuals (Reynolds, 1997b). Reliability of test scores is important as it relates to the amount of variance that is "real," systematic, and related to true differences between individuals. Therefore, it is important to determine the reliability of neuropsychological measures for purposes of individual diagnosis as well as for research, in that reliability influences the likelihood that any experimental or treatment effects will be detected (Reynolds, 1986b). Reliability is also the foundation on which validity is built.

Related to issues of reliability and validity, the method of scaling/measurement used with any test or measure is "crucial" (Reynolds, 1997b, p. 189). Scaling across neuropsychological measures, however, is inconsistent. Frequently what are obtained are raw scores for number correct, time for completion, or number of errors. This results in the need to use score transformations, based on insufficient normative data, in order to make any kind of meaningful comparison. Alternatively, clinicians may use inappropriate scales such as age or grade equivalents in an attempt to give meaning to raw data (Reynolds, 1997b). Grade equivalents, in particular, are inappropriate due to the extent of extrapolation that is used in their derivation as well as faulty assumptions that are made with regard to learning and growth over time (e.g., from lower to upper grades, across subject areas, and throughout the calendar year). Further, grade equivalents exaggerate small differences in performance between individuals and for a single individual across tests (Reynolds, 1986b). It is, therefore, imperative that standard score conversions, based on adequate normative data, be provided for measures used in neuropsychological assessment.

It has been asserted that by age 10 or 12 children perform at adult levels in some areas, and for many older neuropsychological measures most of the normative sampling, in addition to using small numbers, often stopped at age 12. This is despite the fact that many researchers have suggested that neurodevelopment continues through at least age 14 (Boll, 1974) and possibly through age 16 (Golden, 1981). This further limitation in the provision of normative data impedes the interpretation process for adolescents, bolsters the assumption that adolescents should function as adults, and promotes the use of downward extensions of adult measures that often are not appropriate. Neuropsychological function is developmental, and distinct age-related norms are required. It has also been recommended that item response theory (IRT) be used to ensure that neuropsychological measures include an adequate range of difficulty levels, thus ensuring coverage of developmental levels (Morris, 1994). Most existing neuropsychological measures, however, have not been subjected to this type of analysis.

The standardization of administration procedures is also an area of concern. Reynolds (1986b) commented on the availability of at least four versions of Halstead's category test, three of which were somewhat similar and the fourth with significant differences in terms of administration. Despite these differences, however, the same normative data are used. Similarly, administration of the Wisconsin Card Sorting Test (Heaton, 1981) can be done traditionally or via computer, yet there is a single normative data set to be used for scoring and interpretation. Differences in administration impact on the validity and reliability of the measure and normative data, including validity studies, for each variation of administration (unless controlled in the standardization process) are necessary.

Sensitivity, specificity, and diagnostic accuracy need to be further researched as well. Sensitivity is the extent to which a given test accurately predicts brain impairment and is often gauged by statistical power (Pedhazur, 1973). Sensitivity is dependent on validity. No single neuropsychological measure demonstrates high sensitivity (Boll, 1978); combined scores from a given battery may be more successful (e.g., Selz & Reitan, 1979a). In contrast, specificity is dependent on the nature of the behavioral, cognitive, and emotional functions of the task (Batchelor, 1996b). The extent of specificity can only be determined by comparing clinical groups to each other as opposed to focusing on differences between a specific clinical group and the normal population. Cross-clinical group comparisons are frequently not done however. When research is based on comparisons across clinical groups, the results are generally inconclusive (Koriath, Gualtieri, van Bourgondien, Quade, & Werry, 1985).

In comparing clinical groups, it is important to control for comorbidity and family history (Seidman et al., 1995) as well as to differentiate between subtypes of a given disorder when these have been identified (e.g., Halperin, 1991). For many disorders, subtypes have been validated, yet frequently research studies with clinical groups rely on the more global rubric. With regard to learning disabilities, Rourke (1994) asserted that this "lumping" together may lead to gross misunderstanding, if not to conflicting results across studies. For example, the need to develop and incorporate typologies/subtypes for homogeneous grouping of children with dyslexia, for the purposes of developing appropriate intervention as well as for research purposes, has been recognized for some time, yet much of the educational and psychological literature and practices relating to dyslexia continue to address heterogeneous groups of children without regard for subtype (Reynolds, 1986b). Differing typologies and comorbidity have rarely been considered in the extant literature on many neuropsychological measures and likely contribute to the conflicting results of differing studies. Further, the constructs being measured by given tasks may be influenced by gender, premorbid status, the task itself, neuropsychological functions, and so on, making a high level of specificity difficult to attain (Batchelor, 1996b). Batchelor (1996b) suggested that many neuropsychologists compromise between sensitivity and specificity through the selection, administration, and interpretations of neuropsychological measures that are needed to effect such a balance.

Often, in an attempt to provide accurate differential diagnosis, a large set of behaviors is typically assessed. Researchers then use multivariate classification procedures for determination of group information or to determine the effectiveness of specific measures in the diagnostic process. The sample sizes in many of the studies, however, are too small for multivariate analysis, given the large number of variables involved. As a result, in the absence of cross-validation, many diagnoses or classifications may be due to random relationships (Willson & Reynolds, 1982). Problems with the lack of consistency in the diagnosis/classification of disorders also impede the research process and ultimately, clinical practice (Hooper & Tramontana, 1997).

4.10.2.5 Approaches to Test Selection with Children

In addition to selecting tests based on psychometric properties, it has been suggested that child neuropsychologists should select measures that vary along a continuum of difficulty, include both rote and novel tasks, and vary the tasks with regard to processing and response requirements within modalities (Rourke, 1994). Many neuropsychologists continue to include observation and informal assessment; others have adopted more actuarial approaches; many use a combination of observation, informal assessment, and actuarial approaches (Reynolds, 1997b), with a focus on direct appraisal of functions and abilities in order to obtain detailed information on the behavioral effects of brain impairment (Tramontana, 1983). It has been argued that the use of actuarial methods that rely on standardized measures to obtain information may not, however, be useful in intervention planning (D'Amato, 1990).

In addition to quantitative measures, neuropsychological assessment may incorporate not only Luria's theory but also his qualitative assessment model (Luria, 1966, 1970). Luria described assessment that was flexible and varied from individual to individual depending on the functional system that was of concern (Teeter, 1986). Although more dependent on clinician interpretation, qualitative methods can add to information related to the process of learning and may be better suited to the development of intervention/treatment plans (D'Amato et al., in press). In the incorporation of qualitative tasks, child neuropsychologists make use of work samples, informal tasks, criterion-referenced measures and clinical observations of interactions throughout the assessment process (D'Amato et al., 1997). Qualitative procedures can also be used to complete a task analysis and determine specifically which components of a more complex task are problematic for the child (Taylor, 1988). Others may use standardized measures but administer them in other than standardized fashion (Kaplan, 1988). Modifications of tasks presented (e.g., provision of cues, adjustment of rate, changing modality of presentation or response, adjustment of task complexity) can provide insight into processing differences (Clark & Hostettler, 1995; Harrington, 1990; Ylvisaker et al., 1990) and have been recommended for use in the evaluation of children who are culturally or linguistically diverse (Gonzalez, Brusca-Vega, & Yawkey, 1997). Unfortunately, tests administered with these types of modifications are no longer consistent with standardization procedures, and clinicians need to exercise caution in the interpretation of brain–behavior relations based on qualitative data (D'Amato et al., in press). A strictly qualitative approach using experimental/ad hoc measures and nonquantitative/ nonstandardized interpretation of standardized measures may provide additional information, but does not allow for verification of diagnostic accuracy, is not easily replicated, and does not allow for formal evaluation of treatment methods (Rourke, 1994). In practice, most clinicians prefer a combination of quantitative and qualitative measures (Rourke, 1994).

Test selection in the neuropsychological assessment of children and young people varies considerably from clinician to clinician due to differences in philosophy and theoretical foundations. Evaluation may take the form of various published battery approaches (e.g., Golden, 1997; Reitan, 1974; Selz, 1981) or may use a more eclectic approach (e.g., Benton, Hamsher, Varney, & Spreen, 1983; Gaddes, 1980; Hynd & Cohen, 1983; Knights & Norwood, 1979; Obrzut, 1981; Obrzut & Hynd, 1986; Rourke, Bakker, Fisk, & Strang, 1983; Rutter, 1983; Spreen & Gaddes, 1979; Teeter, 1986; Tramontana & Hooper, 1987). Generally, however, the approaches can be categorized as nomothetic, idiographic, or a combination of these two approaches. Additional variation within categories, however, is evident in the extent to which clinicians rely on quantitative, qualitative, or both types of information in the process.

4.10.2.5.1 Nomothetic approaches

The fixed/standardized battery or nomothetic approach uses the same assessment protocol for all children being assessed. An example of the nomothetic approach would be the administration of a published neuropsychological battery, usually in conjunction with IQ and achievement tests. It may also be a predetermined set grouping of selected tests that remains constant across children evaluated, regardless of the referral problem (Sweet, Moberg, & Westergaard, 1996). These tend to be more actuarial in nature (Lezak, 1995) and often rely on cut-off scores, pathognomonic signs, or a combination, for determination of the presence of brain damage. The choice of a fixed battery approach is generally related to an orientation and preference consistent with standardized procedures, objective methods, and psychometric development. It may also reflect a preference for "blind" assessment such that the referral problem does not dictate the measures used (Goldstein, 1997). This approach has the advantage of covering a breadth/depth of functions, provides for extensive databases, and facilitates the collection of data for clinical interpretation of large numbers of clinical groups. Standardized/nomothetic batteries, however, often do not take into consideration education, age, and experiential variables, and may or may not specifically address the referral question. Further, diagnosis with a nomothetic approach may be driven by the base rates of the clinical problems in a particular setting, due to sampling bias, and therefore may not be useful for detecting disorders in other population samples (Tramontana & Hooper, 1987). Use of a standardized battery/nomothetic approach appears to be declining in the general area of clinical neuropsychology (Sweet et al., 1996); however it may be the preferred method if litigation is a potential issue (Reitan & Wolfson, 1985).

The published neuropsychological batteries most frequently used with school-aged children are the Luria Nebraska Neuropsychological Battery-Children's Revision (LNNB-CR; Golden, 1984), the Halstead–Reitan Neuropsychology Battery (HRNB; Reitan & Davison, 1974) and the Reitan Indiana Neuropsychological Test Battery for Children (RINB; Reitan, 1969). All of these batteries require extensive training for appropriate administration and interpretation of results. The neuropsychological battery is often supplemented with a traditional test of cognitive ability as well as achievement testing.

The HRNB and RINB are considered to be the most widely used in clinical practice (Howieson & Lezak, 1992; Nussbaum & Bigler, 1997). Both of these use a multiple inferential approach to interpretation, including level of performance, pathognomonic signs, patterns of performance, and right–left differences (Reitan, 1986, 1987). The batteries contain numerous measures that are considered necessary for understanding brain–behavior relationships in children and adolescents. Descriptions of these measures are provided in Tables 1 and 2. Both the HRNB and the RINB can be used in clinical practice for the assessment of a child with identified brain damage as well as with those children where specific brain damage has not been documented through neuroradiological methods (Nussbaum & Bigler, 1997).

Strong correlations have been found between the Wechsler Intelligence Scale for Children-Revised (WISC-R; Wechsler, 1974) and the RINB and HRNB (Klesges, 1983) suggesting the ability of the latter tests to predict neuropsychological dysfunction. A number of factor analytic studies have been completed comparing results from the WISC-R and the Reitan batteries (e.g., Batchelor et al., 1991; D'Amato, Gray, & Dean, 1988; Snow & Hynd, 1985a); these consistently suggest that most of the variance is due to factors of language, academic achievement, and visual spatial skills. With the addition of other measures, up to eight factors were found (D'Amato et al., 1988; Batchelor et al., 1991). Although factor analytic research has demonstrated a good deal of common information when the WISCR and HRNB were both given, it has also been determined that the HRNB offers unique information (Klonoff & Low, 1974) and the addition of the HRNB to the typical psychoeducational battery has been found to increase the extent to which variability in school achievement can be accounted for (Strom, Gray, Dean, & Fischer, 1987). Research regarding the efficacy of the HRNB in the differential diagnosis of children with learning problems is equivocal (Arffa, Fitzhugh-Bell, & Black, 1989; Batchelor, Kixmiller, & Dean, 1990; Selz & Reitan, 1979a, 1979b). Factor analytic research with the RINB has been less conclusive (Crockett, Klonoff, & Bjerring, 1969; Foxcroft, 1989; Teeter, 1986). The RINB has been found to be sensitive to mild levels of traumatic brain injury within four months of injury in the absence of obvious lags in academic achievement (Gulbrandson, 1984). The HRNB and RINB are both downward extensions of the adult version with some modifications for children (Teeter, 1986) and do not fully reflect the developmental continuum of childhood and youth (Cohen et al., 1992).

The LNNB-CR was developed on the basis of the neurodevelopmental stages of the child (Golden, 1981, 1997) and was revised four times in the process (Plaisted, Gustavson, Wilkening, & Golden, 1983). It is designed for children ages 8–12 years and in addition to IQ and achievement provides information specific to motor, rhythm, tactile, visual, receptive speech, expressive language, and memory functions. A description of the LNNB-CR is provided in Table 3. The development of the LNNB-3 represents an extensive revision and major expansion of the LNNB-CR and the adult version. It includes tasks from the previous two measures, but also includes additional tasks, with a total of 27 domains being evaluated. With this major revision, both lower level and more complex tasks and items have been added. The LNNB-3 is intended for use with individuals from age five through adulthood. Interpretation of the LNNB-CR and LNNB-3 focuses predominantly on scale patterns and intrascale (intraindividual) differences, as opposed to levels of performance or pathognomonic signs.

Due to its recent development, there is little research available on the LNNB-3 and most is specific to adults (e.g., Crum, Bradley, Teichner, & Golden, 1997; Crum, Golden, Bradley, & Teichner, 1997). Extensive research has, however been completed with the LNNB-CR. Factor analytic studies (e.g., Karras, Newton, Franzen, & Golden, 1987; Pfeiffer, Naglieri, & Tingstrom, 1987; Sweet, Carr, Rossini, & Kasper, 1986) have resulted in varying factor structures. It has been determined consistently that the LNNB-CR offers unique information not otherwise obtained in psychoeducational assessment, with particular sensitivity to deficits in language, writing, reading, and rhythm (Geary & Gilger, 1984). The pathognomonic scale of the LNNB-CR has been found to account for increased variance, independently of the WISC-R, and to be a better predictor of academic achievement in spelling and reading (McBurnett, Hynd, Lahey, & Town, 1988). It was also found that the LNNB-CR had greater shared variance than the WISC-R with measures of achievement (Hale & Foltz, 1982). The LNNB-CR has been found to be more sensitive to improvement in functioning following medical intervention (e.g., shunt placement) than either cognitive or achievement measures (Torkelson, Liebrook, Gustavson, & Sundell, 1985), as well as supporting differential diagnosis (Carr, Sweet & Rossini, 1986) and the understanding of academic deficits in children with emotional/behavioral problems (Tramontana, Hooper, Curley, & Nardolillo, 1990). The utility of the LNNB-CR in the differentiation of learning disability as opposed to other forms of brain damage, however, has been questioned (e.g., Morgan & Brown, 1988; Oehler-Stinnett,

Table 1 Halstead–Reitan Neuropsychological Battery for Children (ages 9–14 years).

Subtest	*Description*	*Function(s) assessed*
Category test	Requires individual to select colors or numbers corresponding to some abstract problem-solving criteria. Immediate feedback is provided for both correct and incorrect responses.	This task assesses general abstraction and concept formation as well as general neuropsychological functioning (Reitan & Wolfson, 1985, 1988).
Tactual performance test	The individual is blindfolded and required to place blocks in slots on a form board using the dominant hand, the nondominant hand, and both hands together. The individual is then asked to draw a diagram of the board with the blocks in their proper spaces.	This task measures tactual discrimination, sensory recognition, and spatial memory. The drawing component is a measure of incidental learning/memory (Reitan & Wolfson, 1988; Selz, 1981).
Speech sounds perception	A taperecorded voice presents a sequence of 60 spoken nonsense words from which the individual must select the correct word each time from three written choices.	This task measures alertness, attention/concentration, and verbal ability (Reitan & Wolfson, 1988)
Seashore rhythm test	The individual is required to differentiate between 30 pairs of rhythmic patterns which are sometimes the same and sometimes different.	This test is thought to be an indicator of generalized cerebral function as well as a measure of alertness and attention/concentration. (Reitan & Wolfson, 1988)
Trail making test	This test uses two tracking tasks, one with numbers (A) and one with letters and numbers (B). First, the individual must connect numbered circles in order; then, the individual must connect circles in sequence, alternating numbers and letters.	The test is believed to measure conceptual flexibility, symbolic recognition, and visual tracking under time constraints (Selz, 1981). It is also used as a measure of overall functioning (Reitan & Wolfson, 1985, 1988).
Finger oscillation test	This test requires the individual to depress a lever as quickly as possible with the index finger of each hand.	This measures motor speed and manual dexterity (Selz, 1981) and lateral dominance (Reitan & Wolfson, 1988.)
Aphasia screening test	This test includes enunciation of spoken language (repeating), naming, reading, writing, spelling, and arithmetic. It also includes copying of a square, circle, and Greek cross.	It is a measure of verbal ability. The drawings are indicative of the verbal-to-motor process (Reitan & Wolfson, 1988).
Sensory perceptual examination		All measure receptive sensory function (Reitan & Wolfson, 1985, 1988).
Tactile perception	The individual is asked to report whether right hand, left hand, right side of face, or left side of face is touched; touches are done unilaterally and bilaterally.	
Auditory perception	Examiner lightly rubs fingers together at the individual's right, left or both ears and the individual is asked to localize the sound produced.	
Visual perception	The individual is asked to report peripheral, unilateral and bilateral single movements produced by the examiner, to assess all four quadrants of the visual field.	

Table 1 (continued)

Subtest	*Description*	*Function(s) assessed*
Tactile form recognition	The individual must identify a cross, triangle, square or circle when put in the dominant hand behind a board (unseen) and point to that same object with the nondominant hand; the same process is then carried out with the object in the nondominant hand.	This test is believed to measure tactile perception as well as attention (Nussbaum & Bigler, 1997).
Fingertip number writing	This requires the individual to identify numerals written on their fingertips (both hands).	This is a measure of sensory perceptual functioning (Reitan & Wolfson, 1988).
Grip strength test	Using a hand dynamometer, the strength of grip for the dominant and nondominant hand is determined.	This measures motor functioning and lateral dominance (Reitan & Wolfson, 1988)

Stinnett, Wesley, & Anderson, 1988; Snow & Hynd, 1985b; Snow, Hynd, & Hartlage, 1984). Hynd (1992) also questioned the appropriateness of the standardization sample.

More recently, the Neuropsychological Investigation for Children (NEPSY; Korkman, Kirk, & Kemp, 1997) has been developed for young children. Based on Luria's model (1970), the NEPSY consists of 27 subtests that are summarized in test profiles of strengths and weaknesses. Initially developed in Finnish, in its English version the test includes subtests specific to attention and executive functions, and language, sensorimotor, visuospatial and memory/learning functions (Korkman et al., 1997). It is intended for use with children age 3–12 years. Early research on the NEPSY appears positive (e.g., Korkman, 1988; Korkman, & Hakkinen-Rihu, 1994; Korkman, Liikanen, & Fellman, 1996); however, additional research with this measure, particularly in comparison to the KABC for younger children, will be needed following publication.

In addition, a new battery, the Dean–Woodcock Neuropsychological Assessment System (DWNAS: Dean & Woodcock, in press) is in the process of development. Based on the work of Catell and Horn (Horn, 1988, 1991), this battery combines the cognitive battery of the Woodcock–Johnson Psychoeducational Battery-Revised with a newly developed battery of sensorimotor tests, a structured interview, and a mental status exam. The sensory and motor portion is projected to include eight tests of sensory function and nine tests of motor function. Although some of these tests are similar to those on other neuropsychological batteries, standardized administration, objective scoring criteria, and normative data will be provided for all tasks. The structured interview and mental status exam are intended to provide information specific to emotional state, motivation, temperament, and prior medical conditions as well as to premorbid history, age at onset and emotional reaction (coping) that may influence neuropsychological performance (Dean & Woodcock, in press). It is projected that this battery will be available in both English and Spanish, with general as well as focused norms that account for age and education using regression methods. As with the NEPSY, clinical research with the DWNAS will be needed once all components of the system are available.

Nomothetic approaches may also be eclectic and use selected measures to sample behaviors from the differing functional systems. Several examples of eclectic batteries can be found in the research literature (e.g., Nussbaum et al., 1988; Rourke, 1994). Since the combinations of measures vary in an eclectic battery from clinical setting to clinical setting, research regarding the efficacy of any given combination of tasks as compared to other combinations or to the published batteries is not feasible, and factor analytic studies of eclectic batteries are not routinely found in the research literature.

4.10.2.5.2 Idiographic approaches

At the other end of the continuum, an idiographic approach tailors the assessment battery to the referral question and the child's individual performance on initial measures administered (Christensen, 1975; Luria, 1973). This type of approach is intended to isolate neurobehavioral mechanisms that underlie the problem of a particular individual rather than

Table 2 Reitan Indiana Neuropsychological Battery (ages 5–8).

Subtest	*Description*	*Function(s) assessed*
Category Test	Requires the individual to select colors or numbers corresponding to some abstract problem-solving criteria. Immediate feedback is provided. This version has fewer items and only five categories.	This task assesses general abstraction and concept formation as well as general neuropsychological functioning. (Reitan & Wolfson, 1985).
Matching pictures test	The child matches a single picture to the same picture or to a picture from a more general category.	This task assesses abstraction and concept formation (Reitan & Wolfson, 1988).
Color form test	The child must alternately touch shapes and colors.	This task assesses abstraction and concept formation (Reitan & Wolfson, 1988).
Progressive figures test	The child must use the small shape within a large shape to select the outer shape of the next figure in sequence in a timed condition.	This task assesses concept formation (Reitan & Wolfson, 1988). Also involves cognitive flexibility and attention (Nussbaum & Bigler, 1997)
Tactual performance test	The individual is blindfolded and required to place blocks in slots on a form board using the dominant hand, the nondominant hand, and both hands together. The individual is then asked to draw a diagram of the board with the blocks in their proper spaces.	This task measures tactual discrimination, sensory recognition, and spatial memory. The drawing component is a measure of incidental learning (Reitan & Wolfson, 1988; Selz, 1981).
Finger oscillation test	This test requires the individual to depress a lever as quickly as possible with the index finger of each hand.	This measures motor speed and manual dexterity (Selz, 1981) and lateral dominance (Reitan & Wolfson, 1988.)
Fingertip symbol writing	This requires the individual to identify Xs and Os written on their fingertips (both hands).	This is a measure of sensory perceptual functioning and attention (Reitan & Wolfson, 1988).
Marching test	The child is required to touch a sequence of circles as quickly as possible.	This measures motor functioning (Reitan & Wolfson, 1988).
Sensory perceptual examination		All measures are sensitive to receptive sensory function (Reitan & Wolfson, 1985, 1988).
Tactile perception	The individual is asked to report whether right hand, left hand, right side of face, or left side of face is touched; touches are done unilaterally and bilaterally.	
Auditory perception	Examiner lightly rubs fingers together at the individual's right, left or both ears and the individual is asked to localize the sound produced.	
Visual perception	The individual is asked to report peripheral, unilateral and bilateral single movements produced by the examiner, to assess all four quadrants of the visual field.	
Grip strength test	Using a hand dynamometer, the strength of grip for the dominant and nondominant hand is determined.	This measures motor functioning and lateral dominance of the upper body (Reitan & Wolfson, 1988).

Table 2 (continued)

Subtest	*Description*	*Function(s) assessed*
Tactile form recognition	The individual must identify a cross, triangle, square, or circle when put in the dominant hand behind a board (unseen) and point to that same object with the nondominant hand; the same process is then carried out with the object in the nondominant hand.	This test is believed to measure tactile perception as well as attention (Nussbaum & Bigler, 1997).
Aphasia screening test	This test includes enunciation of spoken language, naming, reading, writing, and arithmetic; naming, identifying body parts, left/right, numerals and letters. It also includes drawing a square circle, and Greek cross. It is an abbreviated version of the screening test for older children and adults.	It is a measure of verbal ability. The drawings are indicative of the verbal-to-motor process (Reitan & Wolfson, 1988).
Individual performance		These measure visual perceptual and spatial abilities (Reitan & Wolfson, 1988).
Matching Vs, figures	Child must match figures or Vs.	
Star, concentric squares	Child must copy figures of varying difficulty.	

providing a detailed evaluation of all areas of functioning. With no predetermined uniformity across evaluations and the dependence on the individual's presenting problems, this approach requires substantial clinical knowledge to determine the components of the battery in order to meet this goal. Due to the limited data on neuropsychological functions and organization of behavior in children, this approach is less frequently used (Fennell, 1994). However, it may be more cost-effective because of the small number of domains which are assessed (Goldstein, 1997). A major drawback to the idiographic approach is the limited research base which is generated and the inability to verify or study the efficacy of this approach as compared to other approaches.

4.10.2.5.3 Combined approaches

The most frequently used approach represents a combination of the nomothetic and idiographic approaches and has been referred to as the flexible battery approach (Sweet et al., 1996). A core set of the same tests is administered to all children, as in the nomothetic approach, and serves as the basis for initial hypothesis generation; this may constitute an initial screening battery. To this core set, further tests are added that are specific to the referral question or that are believed, based on initial observations and performance, to enhance the information provided (Bauer, 1994; L. C. Hartlage & Telzrow, 1986; Rourke, Fisk, & Strang, 1986). The components of the flexible battery itself generally reflect the theoretical position taken by the neuropsychologist with regard to the manner in which behavioral performance reflects brain pathology and the reasons for referral for neuropsychological evaluation for a given individual (Bauer, 1994). According to surveys completed in the 10 years since 1987, the flexible battery approach is generally that preferred by neuropsychologists working with populations of varying ages (Sweet & Moberg, 1990; Sweet et al., 1996). This approach is believed to more accurately identify specific deficits (Batchelor, 1996b).

The Boston Process Approach (Kaplan, 1988; Lezak, 1995) is one example that incorporates a flexible battery. Specific measures with low specificity are used to assess multiple constructs in a variety of neuropsychological domains (Batchelor, 1996b). Hypotheses are then made based on the initial measures, and additional measures with higher levels of specificity are then selected and used to differentiate within and between various functions. Hypotheses initially generated from the screening battery are thus either confirmed or nullified. Inferences are then made regarding brain function based on the specific deficits identified. The flexible battery used in the Boston Process Approach is not limited to quantitative data but also includes qualitative information that is believed to be important in

Table 3 Luria Nebraska Neuropsychological Battery-Children's Revision.

Scale	*Description*	*Function(s) assessed*
C1 (motor)	Items cover a variety of motor skills (bilateral and unilateral) including simple hand movements, drawing, and constructional skills.	These tests measure motor domains but are sensitive to many types of motor problems. (Golden, 1997).
C2 (rhythm)	Items include a variety of tasks in which the child is required to report whether one of two groups of tones is higher or lower, reproduce tones and rhythmic patterns, and identify the number of beeps in groups of sounds.	These items are considered to be most sensitive to attention and concentration (Golden, 1997).
C3 (tactile)	Items include tasks in which the child is asked to report where they are touched, how hard they are touched, as well as to name and identify objects through touch.	These items measure the extent of cutaneous sensation and stereognostic perception (Golden, 1997).
C4 (visual)	Items include tasks in which the child is required to identify an object or picture, overlapping pictures, pictures that are difficult to perceive, and mirror image versions; items also include progressive matrices, and spatial rotation.	These items measure visual–spatial organization and perception as well as right hemisphere function (Golden, 1997).
C5 (receptive speech)	The child is required to repeat phonemes, repeat phonemes at various levels of pitch, name objects, point to objects, identify and define words, and respond to sentences.	These items measure receptive language and auditory skills as well as left hemisphere function (Golden, 1997).
C6 (expressive speech)	The child is required to repeat phonemes, words, and sentences as well as to generate speech forms including naming objects, counting forward and backward, spontaneous discourse in response to a picture, story, or discussion topic.	These items measure expressive language as well as left hemisphere function. Results may be impacted by reading ability (Golden, 1997).
C7 (writing)	Tasks include copying of letters and words, writing first and last name, writing sounds, words, and phrases from dictation.	These items measure visual motor and auditory motor skills and are believed to measure functioning of the temporal–parietal–occipital area (Golden, 1997).
C8 (reading)	The child is asked to generate sounds from letters, name letters, read simple words, sentences, and paragraphs.	These items measure reading as well as left hemisphere function (Golden, 1997).
C9 (arithmetic)	Child is asked to write arabic and roman numerals from dictation, compare numbers, complete simple computation problems, and generate serial threes.	These tasks measure arithmetic skills, but are considered the most sensitive to educational deficits as well to all/any dysfunction (Golden, 1997).
C10 (memory)	Tasks required include having the child memorize words as well as predicting their own performance on various memory tasks.	These items measure short-term memory functions and are most sensitive to verbal dysfunction (Golden, 1997).
C11 (intellectual)	The child is asked to complete a variety tasks including interpretation of pictures, arranging pictures in order, identification of what is comical/absurd, interpretation of story, determination of similarities, simple arithmetic problems, identification of logical relations and so on.	These are considered to be reflective of general neuropsychological function, concept formation, and reasoning (Golden, 1997).

understanding the child's problems and in developing effective intervention programs (Batchelor, 1996b; Milberg, Hebben, & Kaplan, 1996). There is less of a focus on the results of standardized test performance with greater attention paid to developmental history, presentation of symptoms, strategy use in task completion, and error analysis. As such, the "process" approach uses both standardized measures and experimental measures as well as "testing of limits" that may involve procedural modifications in order to gain insight into brain–behavior relationships (Kaplan, 1988; Milberg, et al., 1996). Concern has, however, been expressed regarding the reliability of scores obtained on standardized measures when the standardization procedures have been compromised (e.g., Rourke et al., 1986). Further, most of the research and clinical study, with the Boston Process Approach in particular, has been with adult populations as opposed to children, and it is not recommended for other than research applications.

4.10.2.6 General Organization of the Neuropsychological Assessment of the Child

When the neurologist examines a child, the physical examination looks principally for structural defects in the CNS, trauma to the CNS, or specific disease entities or toxins. An assessment of history is an integral component of both the neurological and the neuropsychological assessment of children and includes, for the neurologist, assessment of the gestational period, delivery, postnatal history, and the family medical history through at least two generations. The physical examination that lay people view as the neurological examination proper is based largely on observations of the neurologist and is conducted in the context of a brief interview and physical manipulation to assess tone, muscle strength, deep tendon reflexes, sensation, and brain stem and spinal reflexes. Electrophysiologic, serologic and/or imaging studies may then be ordered as may be suggested by such results. Neuropsychological testing may also be ordered when there are suspicions of intellectual delay or functional sequelae are suspected, related to trauma, disease, or toxins. As recently as the 1970s and into the 1980s, neuropsychological testing was used to evaluate lesion site and size and to assist in the differential diagnosis of a variety of neurologic diseases, but this function has been largely supplanted by advances in neuroimaging, clinical serology, and the linking of a variety of cancers to mental symptoms.

The neuropsychological examination of children is focused more directly on an analysis of the functional concomitants and sequelae than on the identification of strictly neurologic disorders, but it is also useful in the diagnosis and identification of more subtle conditions (e.g., learning disabilities, ADHD) or other neurologic disorders, especially in early stages (such as childhood onset of Huntington's disease) that are more resistant to diagnosis via neurologic examination. (In adulthood, differential diagnosis, such as depression versus dementia or differentiation of malingering or among various dementias takes on greater importance.) At all ages, the neuropsychological examination is also focused on rehabilitation.

A thorough history is important to a proper neuropsychological assessment. The length of time since trauma or disease onset, premorbid levels of functions, family history of related problems, and problems related to gestation, delivery, and the postnatal period are all relevant to accurate interpretation of the results of neuropsychological testing. If a school-aged or college-aged individual is involved, it is important to review educational history with specific performance data including standardized tests scores and grades along with any special education history. With children who are, developmentally, a moving target, it is important to be always cognizant of the educational implications of the reason for referral. Following a review of history and obtainable records, there are nine key points to consider in the organization of the neuropsychological assessment.

(i) *All or at least a significant majority of the child's educationally relevant cognitive skills or higher order information process skills should be assessed.* This will often involve an assessment of general intellectual level via a comprehensive IQ test such as a Wechsler scale or KABC. Evaluation of the efficiency of mental processing as assessed by strong measures of g, is essential to provide a baseline for interpreting all other aspects of the assessment process. Assessment of basic academic skills including reading, writing, spelling, and math will be necessary, along with tests of memory and learning such as the TOMAL which also have the advantage of including performance-based measures of attention and concentration. Problems with memory, attention and concentration, and new learning are the most common of all complaints following CNS compromise and are frequently associated with more chronic neurodevelopmental disorders (e.g., learning disability, ADHD).

(ii) *Testing should sample the relative efficiency of the right and of the left hemispheres of the brain.* Asymmetries of performance are of interest in their own right, but different brain systems are involved in each hemisphere that have different implications for treatment as well. Even in a diffuse injury such as anoxia, it is possible to find greater impairment in one portion of an individual's brain than in another. Specific neuropsychological tests like those of Halstead and Reitan or the LNNB-CR are useful here.

(iii) *Sample anterior and posterior regions of cortical function*. The anterior portion of the brain is generative and regulatory while the posterior region is principally receptive. Deficits and their nature in these systems will have great impact on treatment choices. Many common tests such as receptive (posterior) and expressive (anterior) vocabulary tests may be applied here along with a systematic and thorough sensory perceptual examination. In conjunction with key point (ii), this allows for evaluation of the integrity of the four major quadrants of the neocortex.

(iv) *Determine the presence of specific deficits*. Any specific functional problems a child is experiencing must be determined and assessed. In addition to those being of importance in the assessment of children with neurodevelopmental disorders, traumatic brain injury (TBI), stroke, even some toxins can produce very specific changes in neocortical function that are addressed best by the neuropsychological assessment. Similarly, research with children with leukemia suggests the presence of subtle neuropsychological deficits following chemotherapy that may not be detected by more traditional psychological measures. Neuropsychological tests tend to be less *g*-loaded as a group and to have greater specificity of measurement than many common psychological tests. Noting areas of specific deficit is important in both diagnosis and treatment planning.

(v) *Determine the acuteness versus the chronicity of any problems or weaknesses found*. The "age" of a problem is important to diagnosis and to treatment planning. Combining a thorough history with the pattern of test results obtained, it is possible, with reasonable accuracy, to distinguish chronic neurodevelopmental disorders such as dyslexia or ADHD from new acute problems resulting from trauma, stroke, or disease. Care must be taken especially in developing a thorough, documented history when such a determination is made. When designing intervention/treatment strategies, rehabilitation and habilitation approaches take differing routes depending upon the age of the child involved and the acuteness or chronicity of the problems evidenced.

(vi) *Locate intact complex functional systems*. It is imperative in the assessment process to locate strengths of the child and intact systems that can be used to overcome the problems the child is experiencing. Treatment following CNS compromise involves habilitation and rehabilitation with the understanding that some organic deficits will represent permanently impaired systems. As the brain is a complex interdependent systemic network of complex organizations that produce behavior, the ability to identify intact systems is crucial to enhancing the probability of designing successful treatment. Identification of intact systems also suggests the potential for a positive outcome to parents and teachers, as opposed to fostering low expectations and fatalistic tendencies on identification of brain damage or dysfunction.

(vii) *Assess affect, personality, and behavior*. Neuropsychologists sometimes ignore their roots in psychology and focus on assessing the neural substrates of a problem. However, CNS compromise will result in changes in affect, personality, and behavior. Some of these changes will be transient, some will be permanent, and due to the developmental nature of children, some will be dynamic. Some of these changes will be direct (i.e., a result of the CNS compromise at the cellular and systemic levels) and others will be indirect (i.e., reactive to loss or changes in function, or to how others respond to and interact with the individual). A thorough history, including onset of problem behaviors, can assist in determination of direct versus indirect effects. Comprehensive approaches such as the Behavior Assessment System for Children (BASC; Reynolds & Kamphaus, 1992) which contain behavior rating scales, omnibus personality inventories, and direct observation scales seem particularly useful. Such behavioral changes will also require intervention and the latter may vary depending on whether the changes noted are direct or indirect effects or whether there were behavior problems evident on a premorbid basis.

(viii) *Test results should be presented in ways that are useful in school settings, not just in acute care or intensive rehabilitation facilities*. Schools are a major context in which children with chronic neurodevelopmental disorders must function. Children who have sustained insult to the CNS (i.e., TBI, stroke) will eventually return to a school or similar educational setting. Schools are where the greatest long-term impact on a child's outcome after CNS compromise is seen and felt. Results should speak to academic and behavioral concerns, reflecting what a child needs to be taught next in school, how to teach to the child's strengths through the engagement of intact complex functional systems, and how to motivate and manage positive behavioral outcomes. For children with TBI, additional information regarding potential for recovery and the tenuousness of evaluation results immediately post-injury needs to be communicated as does the need for reassessment of both the child and the intervention program at regular intervals.

(ix) *If consulting directly to a school, be certain the testing and examination procedures are efficient*. School systems, which is where one finds children, do not often have the resources for funding the type of diagnostic workups neuropsychologists prefer. Therefore, when consulting to the school, it is necessary to be succinct and efficient in planning the neuropsychological evaluation. If the school can provide the results of a very recent intellectual and academic assessment as well as the behavioral assessment information, this can be then integrated into the neuropsychological assessment by the neuropsychologist. If a recent intellectual and academic assessment has not been completed, it may be cost-efficient for qualified school district personnel to complete this portion of the assessment for later integration with other data obtained and interpretated by the neuropsychologist. For children in intensive rehabilitation facilities or medical settings, it may be appropriate for school personnel to participate in the evaluation prior to discharge (i.e., for children with TBI being released and returned to the schools). This collaborative involvement can facilitate program planning with the receiving school district and is preferable to eliminating needed components of the neuropsychological evaluation.

When considering rehabilitation of the child with a focal injury or TBI, several additional considerations are evident. It is important to determine what type of functional system is impaired. Impaired systems may, for example, be modality-specific or process-specific. The nature or characteristics of the impairments must be elucidated before an intelligent remedial plan can be devised.

The number of systems impaired should be determined and prioritized. Children may not be able to work out everything at once and a system of priorities should be devised so that the most important of the impairments to impact overall recovery is the first and most intensely addressed area of impairment. The degree of impairment, a normative question, is also an important consideration in this regard. At times, this will require the neuropsychologist to reflect also on the indirect effects of a TBI, as an impaired or dysfunctional system may adversely affect other systems that are without true direct organic compromise.

The quality of neuropsychological strengths that exist will also be important and tends to be more of an ipsative as opposed to a normative determination. Furthermore, certain strengths are more useful than others. Preserved language and speech are of great importance for example, while an intact sense of smell (an ability often impaired in TBI) is of less importance in designing treatment plans and outcome research. Even more important to long-term recovery are intact planning and concept formation skills. The executive functioning skills of the frontal lobes take on greater and greater importance with age, and strengths in those areas are crucial to long-term planning (as is the detection of weaknesses). These will change, however, with age as the frontal lobes become increasingly prominent in behavioral control after age nine years, again through puberty, and continuing into the 20s.

There are of course times when the scope of the neuropsychological assessment of a child is less broad. On occasion, referrals may be very specific (e.g., "Does Susan have memory or attention problems?). Even when such seemingly succinct questions are asked, it is commonly a good practice to inquire of the referral source as to whether other questions may be anticipated (e.g., Is memory an issue because of poor school achievement? Possible learning disability?).

This section draws in part upon the writings, teachings, and workshops of Lawrence C. Hartlage and Byron Rourke.

4.10.2.7 Interpretation Issues

Neuropsychological assessment of children yields not only an accumulation of test data and impressions, but also a variety of paradigms for understanding and interpreting that data. There are a number of competing paradigms and theories (e.g., Ayers, 1974; Das et al.,and 1979; Luria, 1966; Reynolds, 1981b), and as a result, not only is there considerable variability in the quality and choice of measures used in the neuropsychological assessment of children, there are considerable differences in the ways in which the data obtained are used for making inferences and eventually interpreted (Batchelor, 1996a; Nussbaum & Bigler, 1997). Interpretation of the accumulated data is dependent to a great extent on the neuropsychologist's clinical skills and acumen (D'Amato et al., 1997). Interpretation may be based on overall performance level (e.g., Reitan, 1986, 1987), performance patterns (e.g., Mattarazzo, 1972; Reitan, 1986, 1987), asymmetry of function (e.g., L. C. Hartlage, 1982), the presence of "organic" signs (Kaplan, 1988; Lezak, 1995), or on some combination of features. It is not necessarily the case that only one paradigm is appropriate; which paradigm is most suitable may depend on the child being evaluated. Most importantly, the model used for interpretation should allow the neuropsychologist to make predictions about the child's ability to perform in a variety of contexts and about the efficacy of treatment/intervention plans (Reynolds et al., 1997).

4.10.2.7.1 Performance level

With the use of this indicator, the child's overall level of performance is compared to normative data and conclusions are reached based on deviations from the norm. The extent of variability among typically developing children on some measures at given ages (e.g., when the standard deviation approximates the mean score) may preclude interpretation of results using this approach. In addition, this approach can be misleading, particularly in those individuals with higher cognitive ability (Jarvis & Barth, 1984; Reitan & Wolfson, 1985). Further, there is a tendency for this method to yield a large number of false positives due to the potential for other factors (e.g., motivation, fatigue) to impact on a child's performance (Nussbaum & Bigler, 1997).

4.10.2.7.2 Profile patterns

Application of the neuropsychological model to learning problems has been criticized as being too aligned with a medical model and an emphasis on pathology (Gaddes & Edgell, 1994). As asserted by Little and Stavrou (1993), merely identifying that brain integrity has in some way been compromised is not in and of itself particularly helpful to the child or to those who need to develop interventions to help the child. Neuropsychologists look beyond diagnosis or categorization to an understanding of brain–behavior relations. In order to accomplish this, neuropsychological assessment involves consideration of associations and dissociations of performance across measures (Fletcher, 1988; Rutter, 1981). Performance patterns or intraindividual differences provide a means of conceptualizing functional vs. dysfunctional organizational systems. Strengths and weaknesses are then identified based on the discrepancies between the domains studied. This method has, however, been used frequently for the identification or classification of subtypes of learning disabilities (Branch et al., 1995; Nussbaum & Bigler, 1986; Rourke, 1984).

In interpreting data obtained using this type of evaluation, clinicians differ with regard to emphasis on child strengths, child weaknesses or a combination of strengths and weaknesses (L. C. Hartlage & Telzrow, 1983; Reynolds, 1981b, 1986a; Teeter, 1997). It has been argued that a strength model is more efficacious, with habilitation based on those complex functional systems that are sufficiently intact, and therefore potentially capable of taking over and moderating the acquisition of the skills needed (Reynolds et al., 1997). Emphasis on weaknesses, generally referred to as the deficit model, is not supported by research (e.g., Adams & Victor, 1977; L. C. Hartlage, 1975; P. L. Hartlage & Givens, 1982; P. L. Hartlage & Hartlage, 1978), and deficit approaches to intervention (e.g., remediation of the deficit process) have not been found to be effective and may even be harmful (L. C. Hartlage & Reynolds, 1981).

There are, however, some problems with this method of interpretation regardless of whether the focus is on strengths, weaknesses, or a combination of these. This approach may be misleading as other variables may account for these intra-individual differences (Jarvis & Barth, 1984). Additionally, some such intra-individual differences (e.g., verbal IQ–performance IQ differences) have been found to occur with frequency in the general population (e.g., Kaufman, 1976b) and seemingly abnormal levels of subtest scatter (WISC-R) have been found to be relatively common (Gutkin & Reynolds, 1980; Kaufman, 1976a, 1976b; Reynolds, 1979). Base rates in the general population of specific intra-individual differences for various other combinations of measures have not been studied, and what appears to be a "difference" may not be unusual or unique at a given age level. Further, the stability of these profile patterns over at least very short periods of time needs to be investigated (Reynolds, 1997b).

4.10.2.7.3 Functional asymmetry

Examination of asymmetries in performance across measures is another method of intra-individual consideration. Replicable asymmetries in performance are generally considered signs of CNS dysfunction (Batchelor, 1996b). Most frequently, the comparison is made between those functions that are believed to be right hemisphere-dominated as opposed to left hemisphere-dominated. These differences, however, may be difficult to interpret, particularly for younger children (Reynolds et al., 1997). Further, understanding of the lateralization of cortical functions is frequently based on evidence from adults as opposed to children and assumes that the lateralization is stable over time, despite differing rates of brain maturity (Spreen et al., 1995). Reliance on left–right differences and measures based on lateralization of function have also been criticized as ignoring the role of hemispheric interaction on behavior (e.g., Efron, 1990; Hiscock & Kinsbourne, 1987).

As with the patterns of performance method, right–left differences have been used in the characterization of children with right hemisphere dysfunction as suggestive of learning disability or ADHD (Rourke, 1989). The results of studies are, however, equivocal (e.g., Branch et al., 1995; Gross-Tsur, Salev, Manor, & Amir, 1995; Voeller, 1995). Research, in general, regarding lateralization of function and hemispheric specialization is fraught with conflicting results (e.g., Bever, 1975; Das et al., 1979; Dean, 1984; Reynolds, 1981b), and it has been suggested that the traditional verbal–nonverbal distinction between hemispheres is an oversimplification of a complex system (Dean, 1984; Reynolds, 1981a). Based on Luria's theories, asymmetries of function are not content- or modality-specific but rather are "process"-specific (Reynolds, 1981a, 1981b). Bever (1975) posited two fundamental lateralized processing types, the analytic and holistic; these were translated into sequential and simultaneous in the KABC based on Das et al. (1979). In the research literature, however, there is often a preponderance of emphasis placed on content and modality, as opposed to process, in the interpretation of functional asymmetries; it is believed that this may account for the conflicting results across studies (Reynolds et al., 1997).

Nussbaum et al., (1988) proposed an alternate method of examining asymmetry. In the model of Nussbaum and colleagues, the neuropsychological protocol and interpretation reconceptualized neurobehavioral functioning along the anterior–posterior gradient as opposed to left–right differences. The recommended protocol includes tasks from the HRNB as well as from other batteries. Nussbaum and colleagues asserted that this model may provide additional information in the investigation of asymmetries in children with learning and behavioral problems. Initial research in this area suggested that weaknesses on anterior measures were associated with psychological/behavioral problems (Teeter & Semrud-Clikeman, 1997). Two later studies, however, failed to support the anterior–posterior gradient theory (Matazow & Hynd, 1992)

4.10.2.7.4 Pathognomonic signs

The pathognomonic signs approach involves the identification of specific deficits or performance errors that are not frequently found in typically developing individuals (Nussbaum & Bigler, 1997). Identification of "organic" signs is generally completed through qualitative analysis of errors (Kaplan, 1988; Lezak, 1995). The presence of specific types of errors is then seen as an indication of a compromise to brain integrity. This method has been used reliably with adult populations; however, the utility of this approach in the neuropsychological assessment of children has not been demonstrated (Batchelor, 1996b). The range of variability associated with the developmental process in children would seem to make it more difficult to interpret specific errors as signs of organic impairment (Nussbaum & Bigler, 1997). Unlike the performance levels method, the use of pathognomonic signs has been found to result in a large number of false negatives (Boll, 1974). This may be related to the potential for reorganization/recovery of function in children (Nussbaum & Bigler, 1997).

4.10.2.7.5 Combination approaches

Boll (1981) proposed utilizing performance levels, patterns of performance, pathognomonic signs, and asymmetry of function in concert, in order to account for the potential limitations to the use of any single approach in the interpretation of neuropsychological assessment data. This multiple inferential levels approach is used in the HRNB and is supported by others as well (e.g., Rourke, 1994). The "rules approach" (Selz & Reitan, 1979b) also combines approaches, but in a different manner. Using the "rules approach," each of 37 aspects of neuropsychological performance is rated on a four-point scale in order to provide an objective system for measuring the extent of impairment.

More recently, Taylor and Fletcher (1990) proposed that the child's performance on neuropsychological measures be used to identify and clarify the functional aspects of the child's problems, with the understanding that the biological or neurological substrates of the learning or behavior problem serve to set limits on the child's performance. Levine (1993) has posited still another model for interpretation of neuropsychological data. The "observable phenomenon" model places the emphasis on observable behaviors, that may impact on classroom performance and the changing demands placed on the child over time, as opposed to test results.

4.10.3 CONCLUSIONS

Neuropsychological assessment and the field of clinical child neuropsychology in general have much to offer in the way of understanding the functional systems of the brain and the mechanisms involved in the learning and self-regulation process. Not only is this important in understanding and designing treatment programs for children with problems, but increased understanding of brain functions and their relation to behavior can also improve the overall outcomes for all children (Gaddes, 1983). Historically, neuropsychological assessment of children has taken its lead from research and practice with adults. Issues relating to neurodevelopment, task appropriateness, varying contexts for children, progression following brain injury with children, and so on, render the continuing use of this approach inappropriate. A variety of theoretical models exist; however, many of these are adult-based and used without consideration of developmental issues. Only by developing its own theories and clinical assessment procedures, that are sensitive to developmental features and responsive to educational issues, can the field of clinical child neuropsychology continue to advance and make meaningful contributions to the understanding of learning and behavior problems in children. Development and incorporation of typologies within the childhood disorders, based on clinical experience with children and neuropsychological theory that addresses habilitation, programming, and research needs and that has intuitive appeal to psychologists, educators, and neurologists, would be viewed as a major conceptual contribution to the field of child neuropsychology (Reynolds, 1986b).

Continued methodological and measurement problems in the research that serves as a foundation for the interpretation of neuropsychological data impede progress in the field of clinical child neuropsychology and impact on the accuracy of diagnosis and on the appropriateness of treatment planning. Lack of attention to standard psychometric methods within the field of clinical neuropsychology is all too rampant and poses serious limitations in research in clinical arenas; intuitive appeal, clinical acumen, and perceived utility are not sufficient, but must be combined with sound empirical research (Reynolds, 1986b). While it is anticipated that new measures will have sufficient normative samples for evaluation of associations with demographic variables, and assessment of validity, reliability, sensitivity, and specificity issues, many existing measures continue to have insufficient normative data. All too often, sensitivity is a focus; specificity is also necessary if results are to be useful in treatment/intervention planning. This requires further investigation of contrasting clinical groups. In research with clinical groups, there is a need to consider comorbidity and be as specific as possible in descriptions of clinical subgroups (e.g., Fletcher, Shaywitz, & Shaywitz, 1994). Regardless of the perspective used in interpretation, the value of that interpretation is only as good as the measures used in the assessment process and their sensitivity and specificity (Batchelor, 1996b) in combination with the skills and knowledge of the user (Golden, 1997). Failure to resolve these measurement and methodology issues has impeded and will continue to impede progress in the field of neuropsychology (Reynolds, 1997b).

The field of clinical child neuropsychology is in part driven by the development and application of standardized diagnostic procedures that are sensitive to higher cognitive process as related to brain function (Reynolds, 1997b). While the development of new measures of memory, attention, information processing, and so on provide alternatives for clinical child neuropsychologists (e.g., Reynolds, & Bigler, 1994), further research is needed with these as well as with other new and existing measures in order to determine their utility as part of a comprehensive neuropsychological battery. The incorporation of computer-based assessment is likely to increase in the next decades with the potential for incorporation of computer simulation, interactive types of tasks, virtual reality and so on, as means of measuring neuropsychological function. Computerized testing may facilitate the interface with electrophysiological and neuroradiological methods and, ultimately, bring about significant advances in the understanding of learning and behavior problems. Technological advances in assessment, however, will require the same types of research regarding psychometric properties, confounding factors, cultural/gender differences, and so on. Even with those existing tests that currently include an option for computerized assessment, there are some indications of differences in results following computer administration as opposed to the more traditional administration. If this is the case, then it may be appropriate for separate normative data to be obtained for each mode of administration. Furthermore, children with substantial CNS compromise will have difficulty manipulating computerized test materials, and careful validity research will be required at a time when publishers and others are looking for ways to reduce costs associated with health care products.

A major concern with regard to the increased emphasis on reducing health care costs is that neuropsychologists will shorten tests and attempt to streamline batteries, and in the process lessen both the quantity of time required and the quality of the assessment provided (Woody, 1997). This not only impacts on clinical practice, but also on the knowledge generated through future research. Conflict or dissonance among clinicians will not be well tolerated and the future of clinical child neuropsychology will need the support of public policy (Woody, 1997). This means that reasonable agreement on theoretical foundations, training, and procedures will need to be reached (Woody, 1997). Provision of neuropsychological services needs to be predicated on academics, research, and training consistent with both a clinical psychology orientation and specialized training in brain–behavior relationships, in order to ensure sound foundations (Woody, 1997). With regard to neuropsychological assessment of children, there is a definite need on the part of child neuropsychologists to be well grounded in the developmental process, from a cognitive as well as neurological perspective. Children are a moving target and need even more sophisticated assessment devices than do adults. Considerable work remains in this domain alone.

Neuropsychological assessment of children with learning or behavioral/emotional problems is not necessarily essential. It does, however, provide for a comprehensive evaluation of cognitive skills and emotional factors as well as of environmental influences (L. C. Hartlage & Long, 1997). Neuropsychological assessment may be most appropriate for those children who exhibit characteristics of a disorder that includes cognitive deficits (e.g., learning disability) or significant behavioral problems (e.g., behavior disorder) in the absence of neurophysiological evidence of brain damage, for those with a known neurological syndrome or disease as evidenced by neurophysiological methods, and for those children who, because of genetic predisposition or prenatal/perinatal complications, are believed to be at high risk for neurological disorders (Allen, 1989).

4.10.4 REFERENCES

Adams, R. L. (1985). Review of the Luria Nebraska Neuropsychological Battery. In J. V. Mitchell (Ed.), *The ninth mental measurements yearbook*. Lincoln, NE: University of Nebraska.

Adams, R. D., & Victor, M. (1977). *Principles of neurology*. New York: McGraw-Hill.

Allen, C. (1989). Why use neuropsychology in the schools? *The Neuro-Transmitter, 1,* 1–2.

Ardila, A., & Roselli, M. (1994). Development of language, memory, and visuospatial abilities in 5- to 12-year old children using a neuropsychological battery. *Developmental Neuropsychology, 10,* 97–120.

Ardila, A., Roselli, M., & Puente, T. (1994). *Neuropsychological evaluation of the Spanish speaker*. New York: Plenum.

Arffa, S., Fitzhugh-Bell, K., & Black, F. W. (1989). Neuropsychological profiles of children with learning disabilities and children with documented brain damage. *Journal of Learning Disabilities, 22,* 635–640.

Arnold, B. R., Montgomery, G. T., Castenada, I., & Langoria, R. (1994). Acculturation of performance of Hispanics on selected Halstead–Reitan neuropsychological tests. *Assessment, 1,* 239–248.

Asarnow, R. F., Asamen, J., Granholm, E., & Sherman, T. (1994). Cognitive/neuropsychological studies of children with a schizophrenic disorder. *Schizophrenia Bulletin, 20,* 647–669.

Asarnow, R. F., Brown, W., & Strandburg, R. (1995). Children with a schizophrenic disorder: Neurobehavioral studies. *European Archives of Psychiatry and Clinical Neuroscience, 245*(2), 70–79.

Ayers, A. J. (1974). *Sensory integration and learning disorders*. Los Angeles: Western Psychological Services.

Aylward, G. P., Gioia, G., Verhulst, S. J., & Bell, S. (1995). Factor structure of the Wide Range Assessment of Memory and Learning in a clinical population. *Journal of Psychoeducational Assessment, 13,* 132–142.

Bakker, D. J. (1984). The brain as a dependent variable. *Journal of Clinical Neuropsychology, 6,* 1–16.

Bannatyne, A. (1974). Diagnosis: A note on recategorizations of the WISC scale scores. *Journal of Learning Disabilities, 7,* 272–274.

Barkley, R. A. (1991). The ecological validity of laboratory and analogue assessments of ADHD symptoms. *Journal of Abnormal Child Psychology, 19,* 149–178.

Barkley, R. A. (1994). The assessment of attention in children. In G. R. Lyon (Ed.), *Frames of reference of the assessment of learning disabilities: New views on measurement issues* (pp. 69–102). Baltimore: Brookes.

Barkley, R. A., DuPaul, G. J., & McMurray, M. B. (1991). Attention deficit disorder with and without hyperactivity: Clinical response to three dose levels of methylphenidate. *Pediatrics, 87,* 519–531.

Barkley, R. A., Fischer, M., Newby, R., & Breen, M. (1988). Development of multimethod clinical protocol for assessing stimulant drug responses in ADHD children. *Journal of Clinical Child Psychology, 20,* 163–188.

Batchelor, E. S. (1996a). Introduction. In E. S. Batchelor, Jr. & R. S. Dean (Eds.), *Pediatric neuropsychology: Interfacing assessment and treatment for rehabilitation* (pp. 1–8). Boston: Allyn & Bacon.

Batchelor, E. S. (1996b). Neuropsychological assessment of children. In E. S. Batchelor, Jr. & R. S. Dean (Eds.), *Pediatric neuropsychology: Interfacing assessment and treatment for rehabilitation* (pp. 9–26). Boston: Allyn & Bacon.

Batchelor, E. S., Gray, J. W., Dean, R. S., & Lowery, R. (1988). Interactive effects of socioeconomic factors and perinatal complications. *NASA Program Abstracts, F-169,* 115–116.

Batchelor, E. S., Kixmiller, J. S., & Dean, R. S. (1990). Neuropsychological aspects of reading and spelling performance in children with learning disabilities. *Developmental Neuropsychology, 6,* 183–192.

Batchelor, E. S., Sowles, G., Dean, R. S., & Fischer, W. (1991). Construct validity of the Halstead Reitan Neuropsychological Battery for children with learning disabilities. *Journal of Psychoeducational Assessment, 9,* 16–31.

Bauer, R. M. (1994). The flexible battery approach to neuropsychological assessment. In R. D. Vanderploeg (Ed.), *Clinician's guide to neuropsychological assessment* (pp. 259–290). Hillsdale, NJ: Erlbaum.

Becker, M. G., Isaac, W., & Hynd, G. W. (1987). Neuropsychological development of nonverbal behaviors attributed to "frontal lobe" functioning. *Developmental Neuropsychology, 3,* 275–298.

Bellinger, D. (1995). Lead and neuropsychological function in children: Progress and problems in establishing brain–behavior relationships. In M. G. Tramontana & S. R. Hooper (Eds.), *Advances in Child Neuropsychology*

(Vol. 3, pp. 12–47). New York: Springer-Verlag.

Benton, A. L., Hamsher, K., Varney, N. R., & Spreen, O. (1983). *Contributions to neuropsychological assessment*. New York: Oxford University Press.

Bever, T. G. (1975). Cerebral asymmetries in humans are due to the differentiation of two incompatible processes: Holistic and analytic. In D. Aaronson & R. Reiber (Eds.), *Developmental neurolinguistics and communication disorders*. New York: New York Academy of Sciences.

Beverley, D. W., Smith, I. S., Beesley, P., Jones, J., & Rhodes, N. (1990). Relationship of cranial ultrasonography, visual and auditory evoked responses with neurodevelopmental outcome. *Developmental Medicine and Child Neurology, 32,* 210–222.

Bigler, E. R. (1990). *Traumatic brain injury: Mechanisms of damage, assessment, intervention and outcome*. Austin, TX: Pro-Ed.

Bigler, E. R. (1991). Neuropsychological assessment, neuroimaging, and clinical neuropsychology. *Archives of Clinical Neuropsychology, 6,* 113–132.

Bigler, E. R. (1996). Bridging the gap between psychology and neurology: Future trends in pediatric neuropsychology. In E. S. Batchelor, Jr. & R. S. Dean (Eds.), *Pediatric Neuropsychology* (pp. 27–54). Boston: Allyn & Bacon.

Black, F. W. (1976). Cognitive, academic, and behavioral findings in children with suspected and documented neurological dysfunction. *Journal of Learning Disabilities, 9,* 182–187.

Boll, T. J. (1974). Behavioral correlates of cerebral damage in children age 9–14. In R. M. Reitan & L. A. Davison (Eds.), *Clinical neuropsychology: Current status and applications* (pp. 91–120). Washington, DC: Winston.

Boll, T. J. (1978). Diagnosing brain impairment. In B. B. Wolman (Ed.), *Clinical diagnosis of mental disorders*. New York: Plenum .

Boll, T. J. (1981). The Halstead Reitan Neuropsychological Battery. In S. Filskov & T. J. Boll (Eds.), *Handbook of clinical neuropsychology* (pp. 577–607). New York: Wiley-Interscience.

Bolter, J. F., & Long, C. J. (1985). Methodological issues in research in developmental neuropsychology. In L. C. Hartlage & C. F. Telzrow (Eds.), *Neuropsychology of individual differences: A developmental perspective* (pp. 41–59). New York: Plenum.

Branch, W. B., Cohen, M. J., & Hynd, G. W. (1995). Academic achievement and attention-deficit/hyperactivity disorder in children with left- or right-hemisphere dysfunction. *Journal of Learning Disabilities, 28,* 35–43.

Breslau, N., Chilcoat, H., DelDotto, J., & Andreski, P. (1996). Low birth weight and neurocognitive states at six years of age. *Biological Psychiatry, 40,* 389–397.

Breslau, N., & Marshall, I. A. (1985). Psychological disturbance in children with physical disabilities: Continuity and change in a 5-year follow-up. *Journal of Abnormal Child Psychology, 13,* 199–216.

Burin, D. I., Prieto, G., & Delgado, A. (1995). Solution strategies and spatial visualization strategies: Design of a computerized test for their assessment. *Interdisciplinaria, 12*(2), 123–137.

Carr, M. A., Sweet, J. J., & Rossini, E. (1986). Diagnostic validity of the Luria Nebraska Neuropsychological Battery-Children's Revision. *Jounal of Consulting and Clinical Psychology, 54,* 354–358.

Chelune, G. J., & Baer, R. A. (1986). Developmental norms for the Wisconsin Card Sorting Test. *Journal of Clinical and Experimental Neuropsychology, 8,* 219–228.

Chelune, G. J., & Thompson, L. L. (1987). Evaluation of the general sensitivity of the Wisconsin Card Sorting Test among younger and older children. *Developmental Neuropsychology, 3,* 81–89.

Christensen, A. L. (1975). *Luria's neuropsychological investigation*. New York: Spectrum.

Cicchetti, D. V. (1994). Multiple comparison methods: Establishing guidelines for their valid application in neuropsychological research. *Journal of Clinical and Experimental Neuropsychology, 16,* 155–161.

Clark, E., & Hostettler, C. (1995). *Traumatic brain injury: Training manual for school personnel*. Longmont, CO: Sopris West.

Cohen, M. J. (1997). *The Children's Memory Scale*. San Antonio, TX: Psychological Corporation.

Cohen, M. J., Branch, W. B., Willis, W. G., Weyandt, L. L., & Hynd, G. W. (1992). Childhood. In A. E. Puente & R. J. McCaffrey (Eds.), *Handbook of neuropsychological assessment* (pp. 49–79). New York: Plenum.

Cohen, S. E., Beckwith, L., Parmalee, A. H., & Sigman, M. (1996). Prediction of low and normal school achievement in early adolescents born preterm. *Journal of Early Adolescence, 16,* 46–70.

Conners, C. (1995). *Continuous performance test computer program 3.0: User's manual*. Toronto, London: Multi-Health Systems.

Copeland, D. R., Dowell, R. E., Jr., Fletcher, J. M., Bordeaux, J. D., Sullivan, M. P., Jaffe, N., Frankel, L. S., Ried, H. L., & Cangir, A. (1988). Neuropsychological effects of childhood cancer treatment. *Journal of Child Neurology, 3,* 53–62.

Crockett, D., Klonoff, H., & Bjerring, J. (1969). Factor analysis of neuropsychological tests. *Perceptual and Motor Skills, 29,* 791–802.

Crum, T. A., Bradley, J. D., Teichner, G., & Golden, C. J. (1997, November). *Analysis of the general intelligence subtest of the Luria Nebraska neuropsychological battery III*. Paper presented at the 17th Annual Conference of the National Academy of Neuropsychologists, Las Vegas, NV.

Crum, T. A., Golden, C. J., Bradley, J. D., & Teichner, G. (1997, November). *Analyzing the concurrent validity of the memory scales of the Luria Nebraska neuropsychological battery-Third edition*. Paper presented at the 17th Annual Conference of the National Academy of Neuropsychologists, Las Vegas, NV.

Damasio, A. R., & Maurer, R. G. (1978). A neurological model for childhood autism. *Archives of Neurology, 37,* 504–510.

D'Amato, R. C. (1990). A neuropsychological approach to school psychology. *School Psychology Quarterly, 5,* 141–160.

D'Amato, R. C., Gray, J. W., & Dean, R. S. (1988). A comparison between intelligence and neuropsychological functioning. *Journal of School Psychology, 26,* 283–292.

D'Amato, R. C., Hammons, P. F., Terminie, T. J., & Dean, R. S. (1992). Neuropsychological training in APA-accredited and nonaccredited school psychology programs. *Journal of School Psychology, 30,* 175–183.

D'Amato, R. C., & Rothlisberg, B. A. (1996). How education should respond to students with traumatic brain injuries. *Journal of Learning Disabilities, 29,* 670–683.

D'Amato, R. C., Rothlisberg, B. A., & Leu, P. W. (in press). Neuropsychological assessment for intervention. In C. R. Reynolds & T. B. Gutkin (Eds.), *The handbook of school psychology* (3rd. ed.). New York: Wiley.

D'Amato, R. C., Rothlisberg, B. A., & Rhodes, R. L. (1997). Utilizing neuropsychological paradigms for understanding common educational and psychological tests. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (2nd ed., pp. 270–295). New York: Plenum.

Das, J. P., Kirby, J. R., & Jarman, R. F. (1979). *Simultaneous and successive cognitive processes*. New York: Academic Press.

Dean, R. S. (1984). Functional lateralization of the brain. *Journal of Special Education, 8,* 239–256.

Dean, R. S. (1985). Foundation and rationale for neuropsychological bases of individual differences. In

L. D. Hartlage & C. F. Telzrow (Eds.), *The neuropsychology of individual differences: A developmental perspective* (pp. 203–244). New York: Plenum.

Dean, R. S. (1986). Lateralization of cerebral functions. In D. Wedding, A. M. Horton, & J. S. Webster (Eds.), *The neuropsychology handbook: Behavioral and clinical perspectives* (pp. 80–102). Berlin, Germany: Springer-Verlag.

Dean, R. S., & Gray, J. W. (1990). Traditional approaches to neuropsychological assessment. In C. R. Reynolds, & R. W. Kamphaus (Eds.). *Handbook of psychological and educational assessment of children* (pp. 317–388). New York: Guilford.

Dean, R. S., & Woodcock, R. W. (in press). *Dean Woodcock neuropsychological assessment system professional manual.* Manuscript in preparation, Ball State University.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1994). *CVLT-C Children's California Verbal Learning Test: Manual.* San Antonio, TX: Psychological Corporation.

Denckla, M. B. (1994). Measurement of executive function. In G. R. Lyon, *Frames of reference of the assessment of learning disabilities: New views on measurement issues* (pp. 117–142). Baltimore: Brookes.

Denckla, M. B., LeMay, M., & Chapman, C. A. (1985). Few CT scan abnormalities found even in neurologically impaired learning disabled children. *Journal of Learning Disabilities, 18,* 132–135.

Derryberry, D., & Reed, M. A. (1996). Regulatory processes and the development of cognitive representations. *Development and Psychopathology, 8,* 215–234.

Dietzen, S. R. (1986). *Hemispheric specialization for verbal sequential and nonverbal simultaneous information processing styles of low-income 3 to 5 year olds.* Unpublished doctoral dissertation, Washington State University.

Donders, J. (1992). Validity of the Kaufman Assessment Battery for Children when employed with children with traumatic brain injury. *Journal of Clinical Psychology, 48,* 225–229.

Duffy, F. H., Denckla, M. B., McAnulty, G. B., & Holmes, J. A. (1988). Neurophysiological studies in dyslexia. In F. Plum (Ed.), *Language, communication, and the brain* (pp. 105–122). New York: Raven.

Duffy, F. H., & McAnulty, G. (1990). Neurophysiological heterogeneity and the definition of dyslexia: Preliminary evidence for plasticity. *Neuropsychologia, 28,* 555–571.

Edmonds, J. E., Cohen, M. J., Riccio, C. A., Bacon, K. L., & Hynd, G. W. (1993, October). *The development of clock face drawing in normal children.* Paper presented at the annual meeting of the National Academy of Neuropsychology, Phoenix, AZ.

Efron, R. (1990). *The decline and fall of hemipsheric specialization.* Hillsdale, NJ: Erlbaum.

Evans, L. P., Tannehill, R., & Martin, S. (1995). Children's reading skills: A comparison of traditional and computerized assessment. *Behavior Research Methods, Instruments, and Computers, 27*(2), 162–165.

Fan, X., Willson, V. L., & Reynolds, C. R. (1995). Assessing the similarity of the construct structure of the KABC for black and white children from 7 to 12½ years in age. *Journal of Psychoeducational Assessment, 13,* 120–131

Feagans, L. V., Short, E. J., & Meltzer, L. J. (1991). *Subtypes of learning disabilities: Theoretical perspectives and research.* Hillsdale, NJ: Erlbaum.

Fennell, E. B. (1994). Issues in child neuropsychological assessment. In R. Venderploeg (Ed.), *Clinician's guide to neuropsychological assessment* (pp. 165–184). Hillsdale, NJ: Erlbaum.

Fennell, E. B., & Bauer, R. M. (1997). Models of inference in evaluating brain–behavior relationships in children. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (2nd ed., pp. 204–215). New York: Plenum.

First, M. B. (1994). Computer-assisted assessment of DSM III-R diagnoses. *Psychiatric Annals, 24,* 25–29.

Fletcher, J. M. (1988). Brain-injured children. In E. J. Mash & L. G. Terdal (Eds.), *Behavioral assessment of childhood disorders* (Vol. 2, pp. 451–589). New York: Guilford.

Fletcher, J. M., Shaywitz, B. A., & Shaywitz, S. E (1994). Attention as a process and as a disorder. In G. R. Lyon (Ed.). *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 103–116). Baltimore: Brookes.

Fletcher, J. M., & Taylor, H. G. (1984). Neuropsychological approaches to children: Toward a developmental neuropsychology. *Journal of Clinical Neuropsychology, 6,* 139–156.

Foxcroft, C. D. (1989). Factor analysis of the Reitan-Indiana Neuropsychological Test Battery. *Perceptual and Motor Skills, 69,* 1303–1313.

Gaddes, W. H. (1980). *Learning disabilities and brain function.* Berlin, Germany: Springer-Verlag.

Gaddes, W. H. (1983). Applied educational neuropsychology: Theories and problems. *Journal of Learning Disabilities, 16,* 511–514.

Gaddes, W. H., & Edgell, D. (1994). *Learning disabilities and brain function: A neurodevelopmental approach.* New York: Springer-Verlag.

Gatten, S. L., Arceneaux, J. M., Dean, R. S., & Anderson, J. L. (1994). Perinatal risk factors as predictors of developmental functioning. *International Journal of Neuroscience, 75,* 167–174.

Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114,* 345–362.

Geary, D. C., & Gilger, J. W. (1984). The Luria-Nebraska Neuropsychological Battery-Children's Revision: Comparison of learning disabled and normal children matched on full scale IQ. *Perceptual and Motor Skills, 58,* 115–118.

Golden, C. J. (1981). The Luria-Nebraska children's battery: Theory and formulation. In G. W. Hynd & J. E. Obrzut (Eds.), *Neuropsychological assessment and the school-age child: Issues and procedures* (pp. 277–302). New York: Grune & Stratton.

Golden, C. J. (1984). *Luria-Nebraska neuropsychological battery: Children's revision.* Los Angeles: Western Psychological Services.

Golden, C. J. (1997). The Nebraska neuropsychological children's battery. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (2nd ed., pp. 237–251). New York: Plenum.

Goldman, P. S., & Lewis, M. E. (1978). Developmental biology of brain damage and experience. In C. W. Cotman (Ed.), *Neuronal plasticity.* New York: Raven.

Goldman-Rakic, P. S. (1987). Development of cortical circuitry and cognitive function. *Child Development, 58,* 601.

Goldstein, G. (1997). The clinical utility of standardized or flexible battery approaches to neuropsychological assessment. In G. Goldstein & T. M. Incagnoli (Eds.), *Contemporary approaches to neuropsychological assessment* (pp. 67–92). New York: Plenum.

Gonzalez, V., Brusca-Vega, R., & Yawkey, T. (1997). *Assessment and instruction of culturally and linguistically diverse students with or at-risk of learning problems.* Boston: Allyn & Bacon.

Gordon, M. (1983). *The Gordon Diagnostic System. DeWitt, NY: Gordon System.*

Gray, J. A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system.* Oxford, UK: Oxford University Press.

Gray, J. W., & Dean, R. S. (1990). Implications of neuropsychological research for school psychology. In T. B. Gutkin & C. R. Reynolds (Eds.), *The handbook of school psychology* (pp. 269–288). New York: Wiley.

Gray, J. W., Dean, R. S., & Rattan, G. (1987). Assessment of perinatal risk factors. *Psychology in the Schools, 24,* 15–21.

Gross-Tsur, V., Salev, R. S., Manor, O., & Amir, N. (1995). Developmental right hemisphere syndrome: Clinical spectrum of the nonverbal learning disability. *Journal of Learning Disabilities, 28,* 80–86.

Gulbrandson, G. B. (1984). Neuropsychological sequelae of light head injuries in older children 6 months after trauma. *Journal of Clinical Neuropsychology, 6,* 257–268.

Gutkin, T. J., & Reynolds, C. R. (1980, September). *Normative data for interpreting Reitan's index of Wechsler subtest scatter*. Paper presented at the annual meeting of the American Psychological Association, Montreal, Canada.

Haak, R. (1989). Establishing neuropsychology in a school setting: Organization, problems, and benefits. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (pp. 489–502). New York: Plenum.

Hale, R. L., & Foltz, S. G. (1982). Prediction of academic achievement in handicapped adolescents using a modified form of the Luria Nebraska Pathognomonic Scale and WISC-R Full Scale IQ. *Clinical Neuropsychology, 4,* 99–102.

Hall, C. W., & Kataria, S. (1992). Effects of two treatment techniques on delay and vigilance tasks with attention deficit hyperactive disorder (ADHD) children. *Journal of Psychology, 126,* 17–25.

Halperin, J. M. (1991). The clinical assessment of attention. *International Journal of Neuroscience, 58,* 171–182.

Halperin, J. M., McKay, K. E., Matier, K., & Sharma, V. (1994). Attention, response inhibition, and activity level in children: Developmental neuropsychological perspectives. In M. G. Tramontana & S. R. Hooper (Eds.), *Advances in child neuropsychology* (Vol. 2., pp. 1–54). New York: Springer-Verlag.

Harbord, M. G., Finn, J. P., Hall-Craggs, M. A., Robb, S. A., Kendall, B. E., & Boyd, S. G. (1990). Myelination patterns on magnetic resonance of children with developmental delay. *Developmental Medicine and Child Neurology*, *32,* 295–303.

Harrington, D. E. (1990). Educational strategies. In M. Rosenthal, E. R. Griffith, M. R. Bond, & J. D. Miller (Eds.), *Rehabilitation of the adult and child with traumatic brain injury* (2nd ed., pp. 476–492). Philadelphia: Davis.

Hartlage, L. C. (1975). Neuropsychological approaches to predicting outcome of remedial education strategies for learning disabled children. *Pediatric Psychology*, *3,* 23–28.

Hartlage, L. C. (1982). Neuropsychological assessment techniques. In C. R. Reynolds & T. B. Gutkin (Eds.), *The handbook of school psychology* (pp. 296–313). New York: Wiley.

Hartlage, L. C., & Long, C. J. (1997). Development of neuropsychology as a professional specialty: History, training, and credentialing. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (2nd ed., pp. 3–16). New York: Plenum.

Hartlage, L. C. & Reynolds, C. R. (1981). Neuropsychological assessment and the individualization of instruction. In G. W. Hynd & J. E. Obrzut (Eds.), *Neuropsychological assessment of the school-aged child* (pp. 355–378). Boston: Allyn & Bacon.

Hartlage, L. C., & Telzrow, C. F. (1983). The neuropsychological basis of educational intervention. *Journal of Learning Disabilities*, *16,* 521–528.

Hartlage, L. C., & Telzrow, C. F. (1986). *Neuropsychological assessment and intervention with children and adolescents.* Sarasota, FL: Professional Resource Exchange.

Hartlage, P. L., & Givens, T. S. (1982). Common neurological problems of school age children. In C. R. Reynolds & T. B. Gutkin (Eds.), *The handbook of school psychology* (pp. 1009–1222). New York: Wiley.

Hartlage, P. L., & Hartlage, L. C. (1978). Clinical consultation to pediatric neurology and developmental pediatrics. *Journal of Clinical Child Psychology*, *12,* 52–53.

Haut, J. S., Haut, M. W., Callahan, T. S., & Franzen, M. D. (1992, November). *Factor analysis of the Wide Range Assessment of Memory and Learning (WRAML) scores in a clinical sample*. Paper presented at the 12th Annual Meeting of the National Academy of Neuropsychology, Pittsburgh, PA.

Heaton, R. K. (1981). *A manual for the Wisconsin Card Sorting Test*. Odessa, FL: Psychological Assessment Resources.

Heilman, K. M., Watson, R. T., & Valenstein, E. (1985). Neglect and related disorders. In K. M. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology* (2nd ed., pp. 243–293). New York: Oxford University Press.

Hendren, R. L., Hodde-Vargas, J., Yeo, R. A., & Vargas, L. A. (1995). Neuropsychophysiological study of children at risk for schizophrenia: A preliminary report. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 1284–1291.

Hiscock, M., & Kinsbourne, M. (1987). Specialization of the cerebral hemispheres: Implications for learning. *Journal of Learning Disabilities, 20,* 130.

Hooper, S. R., Boyd, T. A., Hynd, G. W., & Rubin, J. (1993). Definitional issues and neurobiological foundations of selected severe neurodevelopmental disorders. *Archives of Clinical Neuropsychology, 8,* 297–307.

Hooper, S. R., & Hynd, G. W. (1985). Differential diagnosis of subtypes of developmental dyslexia with the Kaufman Assessment Battery for Children (K-ABC). *Journal of Clinical Child Psychology, 14,* 145–152.

Hooper, S. R., & Tramontana, M. G. (1997). Advances in neuropsychological bases of child and adolescent psychopathology: Proposed models, findings, and on-going issues. *Advances in Clinical Child Psychology, 19,* 133–175.

Horn, J. L. (1988). Thinking about human abilities. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate psychology* (2nd ed., pp. 645–685). New York: Academic.

Horn, J. L. (1991). Measurement of intellectual capabilities: A review of theory. In K. S. McGrew, J. K. Werder, & R. W. Woodcock (Eds.), *WJ-R technical manual.* Chicago: Riverside.

Howieson, D. B., & Lezak, M. D. (1992). The neuropsychological evaluation. In S. C. Yudofsky & R. E. Hales, (Eds.), *The American Psychiatric Press textbook of neuropsychiatry* (2nd ed., pp. 127–150). Washington, DC: American Psychiatric Press.

Hurd, A. (1996). A developmental cognitive neuropsychological approach to the assessment of information processing in autism. *Child Language Teaching and Therapy, 12,* 288–299.

Hynd, G. W. (1981). Neuropsychology in schools. *School Psychology Review, 10,* 480–486.

Hynd, G. W. (1992). *Neuropsychological assessment in clinical child psychology*. Newbury Park, CA: Sage.

Hynd, G. W., & Cohen, M. J. (1983). *Dyslexia: Neuropsychological theory, research, and clinical differentiation.* New York: Grune & Stratton.

Hynd, G. W., Marshall, R. M., & Semrud-Clikeman, M. (1991). Developmental dyslexia, neurolinguistic theory and deviations in brain morphology. *Reading and Writing: An Interdisciplinary Journal, 3,* 345–362.

Hynd, G. W., & Willis, W. G. (1988). *Pediatric neuropsychology*. New York: Grune & Stratton.

Iivaneihan, M., Launes, J., Pihko, H., Nikkinen, P., &

Lindroth, L. (1990). Single photon emission computed tomography of brain perfusion: Analysis of 60 pediatric cases. *Developmental Medicine and Child Neurology, 32,* 63–68.

Jarvis, P. E., & Barth, J. T. (1984). *Halstead–Reitan Test Battery: An interpretive guide*. Odessa, FL: PAR.

Jernigan, T. L., & Tallal, P. (1990). Late childhood changes in brain morphology observable with MRI. *Developmental Medicine and Child Neurology, 32,* 379–385.

Kail, R. (1984). *The development of memory in children*. San Francisco: Freeman.

Kamphaus, R. W. (1993). *Clinical assessment of children's intelligence*. Boston: Allyn & Bacon.

Kamphaus, R. W., & Reynolds, C. R. (1987). *Clinical and research applications of the K-ABC*. Circle Pines, MN: American Guidance Service.

Kane, R. L., & Kay, G. G. (1997). Computer applications in neuropsychological assessment. In G. Goldstein & T. M. Incagnoli (Eds.), *Contemporary approaches to neuropsychological assessment* (pp. 359–392). New York: Plenum.

Kaplan, E. (1988). A process approach to neuropsychological assessment. In T. Boll & B. K. Bryant (Eds.). *Clinical neuropsychology and brain function* (pp. 125–167). Washington, DC: American Psychological Association.

Karras, D., Newton, D. B., Franzen, M. D., & Golden, C. J. (1987). Development of factor scales for Luria-Nebraska Neuropsychological Battery: Children's revision. *Journal of Clinical Child Psychology, 16,* 19–28.

Kaufman, A. S. (1976a). A new approach to the interpretation of test scatter on the WISC-R. *Journal of Learning Disabilities, 9,* 160–167.

Kaufman, A. S. (1976b). Verbal-performance IQ discrepancies on the WISC-R. *Journal of Learning Disabilities, 9,* 739–744.

Kaufman, A. S. (1979). Cerebral specialization and intelligence testing. *Journal of Research and Development in Education, 12,* 96–107.

Kaufman, A. S., & Kaufman, N. L. (1983a). *Kaufman Assessment Battery for Children (K-ABC) administration and scoring manual*. Circle Pines, MN: American Guidance Services.

Kaufman, A. S., & Kaufman, N. L. (1983b). *Kaufman Assessment Battery for Children (K-ABC) interpretative manual*. Circle Pines, MN: American Guidance Services.

Kinsbourne, M. (1975). Cerebral dominance, learning, and cognition. In H. R. Myklebust (Ed.), *Progress in learning disabilities*. New York: Grune & Stratton.

Kinsbourne, M. (1989). A model of adaptive behavior related to cerebral participation in emotional control. In G. Gainotti & C. Caltagirone (Eds.), *Emotions and the dual brain* (pp. 248–260). New York: Springer-Verlag

Klesges, R. C. (1983). The relationship between neuropsychological, cognitive, and behavioral assessments of brain functioning in children. *Clinical Neuropsychology, 5,* 28–32.

Klonoff, H., & Low, M. (1974). Disordered brain function in young children and early adolescents: Neuropsychological and electroencephalographic correlates. In R. M. Reitan & L. A. Davison (Eds.), *Clinical neuropsychology: Current status and application* (pp. 76–94). Washington, DC: Winston.

Knights, R. M., & Norwood, J. W. (1979). *A neuropsychological test battery for children: Examiner's manual.* Ottawa, Canada: Knights Psychological Consultants.

Koriath, U., Gualtieri, C. T., van Bourgondien, M. E., Quade, D., & Werry, J. S. (1985). Construct validity of clinical diagnosis in pediatric psychiatry: Relationship among measures. *Journal of the American Academy of Child Psychiatry, 24,* 429–436.

Korkman, M. (1988). NEPSY: An adaptation of Luria's investigation for young children. *Clinical Neuropsychologist, 2,* 375–392.

Korkman, M., & Hakkinen-Rihu, P. (1994). A new classification of developmental language disorders. *Brain & Language, 47*(1), 96–116.

Korkman, M., Kirk, U., & Kemp, S. (1997). *The neuropsychological investigation for children*. San Antonio, TX: Psychological Corporation.

Korkman, M., Liikanen, A., & Fellman, V. (1996). Neuropsychological consequences of very low birth weight and asphyxia at term: Follow-up until school age. *Journal of Clinical Neuropsychology, 18,* 220–233.

Leark, R. A., Snyder, T., Grove, T., & Golden, C. J. (1983, August). *Comparison of the KABC and standardized neuropsychological batteries: Preliminary results.* Paper presented at the annual meeting of the American Psychological Association, Anaheim, CA.

Leu, P. W., & D'Amato, R. C. (1994, April). *Right children, wrong teachers? Using an ecological assessment for placement decisions*. Paper presented at the 26th Annual Convention of the National Association of School Psychologists, Seattle, WA.

Levin, H. S., Culhane, K. A., Hartmann, J., Evankovich, K., Mattson, A. J., Harward, H., Ringholz, G., Ewing-Cobbs, L., & Fletcher, J. M. (1991). Developmental changes in performance on tests of purported frontal lobe functioning. *Developmental Neuropsychology, 7,* 377–395.

Levine, M. D. (1993). *Developmental variation and learning disorders*. Cambridge, MA: Education Publishers Service.

Lezak, M. D. (1995). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.

Little, S. G., & Stavrou, E. (1993). The utility of neuropsychological approaches with children. *The Behavior Therapist, 16,* 104–106.

Livingston, R. B., Pritchard, D. A., Moses, J. A., Haak, R. A., Marshall, R., & Gray, R. (1997). Modal profiles for the Halstead–Reitan neuropsychological battery for children. *Archives of Clinical Neuropsychology, 12,* 450–476.

Lovrich, D., Cheng, J. C., & Velting, D. M. (1996). Late cognitive brain potentials, phonological and semantic classification of spoken words, and reading ability in children. *Journal of Clinical Neuropsychology, 18,* 161–177.

Luria, A. R. (1966). *Higher cortical functions in man*. New York: Basic Books.

Luria, A. R. (1970). Functional organization of the brain. *Scientific American, 222,* 66–78.

Luria, A. R., (1973). *The working brain*. New York: Basic Books.

Luria, A. R. (1980). *Higher cortical functions in man* (2nd ed.). New York: Basic Books.

Majovski, L. V. (1984). The K-ABC: Theory and applications for child neuropsychological assessment and research. *Journal of Special Education, 18,* 266–268.

Matazow, G., & Hynd, G. W. (1992, February). *Analysis of the anterior–posterior gradient hypothesis as applied to attention deficit disordered children.* Paper presented at the annual meeting of the International Neuropsychological Society, San Diego, CA.

Mattarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence*. Baltimore: Williams & Wilkins.

Maurer, R. G., & Damasio, A. R. (1982). Childhood autism from the point of view of behavioral neurology. *Journal of Autism and Developmental Disorders, 12,* 195–205.

Mayfield, J. W., & Reynolds, C. R. (1997). Black–white differences in memory test performance among children and adolescents. *Archives of Clinical Neuropsychology, 12,* 111–122.

McBurnett, K., Hynd, G. W., Lahey, B. B., & Town, P. A. (1988). Do neuropsychological measures contribute to the prediction of academic achievement? The predictive validity of the LNNB-CR pathognomonic scale. *Journal of Psychoeducational Assessment, 6,* 162–167.

McGlone, J., & Davidson, W. (1973). The relation between spatial ability with special reference to sex and hand preference. *Neuropsychologia, 11,* 105–113.

Merola, J. L., & Leiderman, J. (1985). The effect of task difficulty upon the extent to which performance benefits from between hemisphere division of inputs. *International Journal of Neuroscience, 51,* 35–44.

Mesulam, M. M. (1985). *Principles of behavioral neurology*. Philadelphia: F. A. Davis.

Milberg, W. B., Hebben, N., & Kaplan, E. (1986). The Boston process approach to neuropsychological assessment. In I. Grant & K. M. Adams (Eds.), *Neuropsychological assessment and neuropsychiatric disorders* (2nd ed., pp. 58–80). New York: Oxford University Press.

Miller, L. T., & Vernon, P. A. (1996). Intelligence, reaction time, and working memory in 4- to 6-year-old children. *Intelligence, 22,* 155–190.

Mitchell, W. G., Chavez, J. M., Baker, S. A., Guzman, B. L., & Azen, S. P. (1990). Reaction time, impulsivity, and attention in hyperactive children and controls: A video game technique. *Journal of Child Neurology, 5,* 195–204.

Moffitt, T. E. (1993). The neuropsychology of conduct disorder. *Developmental Psychopathology, 5,* 135–151.

Molfese, D. L. (1995). Electrophysiological responses obtained during infancy and their relation to later language development: Further findings. In M. G. Tramontana & S. R. Hooper, (Eds.), *Advances in Child Neuropsychology* (Vol. 3, pp. 1–11). New York: Springer-Verlag.

Morgan, S. B., & Brown, T. L. (1988). Luria-Nebraska Neuropsychological Battery-Children's Revision: Concurrent validity with three learning disability subtypes. *Journal of Consulting and Clinical Psychology, 56,* 463–466.

Morris, J. M., & Bigler, E. (1985, January). *An investigation of the Kaufman Assessment Battery for Children (KABC) with neurologically impaired children.* Paper presented at the annual meeting of the International Neuropsychological Society, San Diego, CA.

Morris, R. (1994). Multidimensional neuropsychological assessment models. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement* (pp. 515–522). Baltimore: Brookes

Novak, G. P., Solanto, M., & Abikoff, H. (1995). Spatial orienting and focused attention in attention deficit hyperactivity disorder. *Psychophysiology, 32,* 546–559.

Nussbaum, N. L., & Bigler, E. D. (1986). Neuropsychological and behavioral profiles of empirically derived subgroups of learning disabled children. *International Journal of Clinical Neuropsychology, 8,* 82–89.

Nussbaum, N. L., & Bigler, E. D. (1990). *Identification and treatment of attention deficit disorder*. Austin, TX: Pro-Ed.

Nussbaum, N. L., & Bigler, E. D. (1997). Halstead–Reitan neuropsychological test batteries for children. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (2nd ed., pp. 219–236). New York: Plenum.

Nussbaum, N. L., Bigler, E. D., Koch, W. R., Ingram, J. W., Rosa, L., & Massman, P. (1988). Personality/behavioral characteristics in children: Differential effects of putative anterior versus posterior cerebral asymmetry. *Archives of Clinical Neuropsychology, 3,* 127–135.

Obrzut, J. E. (1981). Neuropsychological procedures with school-age children. In G. W. Hynd & J. E. Obrzut (Eds.), *Neuropsychological assessment and the school-age child: Issues and procedures* (pp. 237–275). New York: Grune & Stratton.

Obrzut, J. E., & Hynd, G. W. (1983). The neurobiological and neuropsychological foundations of learning disabilities. *Journal of Learning Disabilities, 16,* 515–520.

Obrzut, J. E., & Hynd, G. W. (1986). Child neuropsychology: An introduction to theory and research. In G. W. Hynd & J. E. Obrzut (Eds.), *Child neuropsychology* (Vol. 1, pp. 1–12). New York: Academic Press.

Oehler-Stinnett, J., Stinnett, T. A., Wesley, A. L., & Anderson, H. N. (1988). The Luria Nebraska Neuropsychological Battery-Children's Revision: Discrimination between learning disabled and slow learner children. *Journal of Psychoeducational Assessment, 6,* 24–34.

Parsons, O. A., & Prigatano, G. P. (1978). Methodological considerations in clinical neuropsychological research. *Journal of Consulting and Clinical Psychology, 46,* 608–619.

Passler, M., Isaac, W., & Hynd, G. W. (1985). Neuropsychological development of behavior attributed to frontal lobe functioning in children. *Developmental Neuropsychology, 1,* 349–370.

Pedhazur, E. J. (1973). *Multiple regression in behavioral research: Explanation and prediction.* New York: CBS College Publishing.

Pfeiffer, S. I., Naglieri, J. A., & Tingstrom, D. H. (1987). Comparison of the Luria Nebraska Neuropsychological Battery-Children's Revision and the WISC-R with learning disabled children. *Perceptual and Motor Skills, 65,* 911–916.

Phelps, L. (1995). Exploratory factor analysis of the WRAML with academically at-risk students. *Journal of Psychoeducational Assessment, 13,* 384–390.

Phelps, L. (1996). Discriminative validity of the WRAML with ADHD and LD children. *Psychology in the Schools, 33,* 5–12.

Plaisted, J. R., Gustavson, J. C., Wilkening G. N., & Golden, C. J. (1983). The Luria Nebraska Neuropsychological Battery-Children's Revision: Theory and current research findings. *Journal of Clinical Child Psychology, 12,* 13–21.

Powell, H. (1997). Comment on computerized assessment of arithmetic computation skills with MicroCog. *Journal of the International Neuropsychological Society, 3,* 200.

Powell, D. H., Kamplan, E. F., Thitla, D., Weintraub, S., Catlin, R., & Funkenstein, H. H. (1993). *MicroCog assessment of cognitive functioning manual.* San Antonio, TX: Psychological Corporation.

Ramsey, M. C., & Reynolds, C. R. (1995). Separate digits tests: A brief history, a literature review, and reexamination of the factor structure of the Test of Memory and Learning (TOMAL). *Neuropsychology Review, 5,* 151–171.

Reitan, R. M. (1969). *Manual for the administration of neuropsychological test batteries for adults and children.* Indianapolis, IN: Author.

Reitan, R. M. (1974). *Clinical neuropsychology: Current status and applications.* New York: Winston.

Reitan, R. M. (1986). *Theoretical and methodological bases of the Halstead–Reitan Neuropsychological Test Battery.* Tucson, AZ: Neuropsychological Press.

Reitan, R. M. (1987). *Neuropsychological evaluation of children.* Tucson, AZ: Neuropsychological Press.

Reitan, R. M., & Davison, L. A. (1974). *Clinical neuropsychology: Current status and applications.* Washington, DC: Winston.

Reitan, R. M., & Wolfson, D. (1985). *The Halstead Reitan neuropsychological battery: Theory and clinical interpretation.* Tucson, AZ: Neuropsychological Press.

Reitan, R. M., & Wolfson, D. (1988). The Halstead Reitan Neuropsychological Test Battery and REHABIT: A model for integrating evaluation and remediation of cognitive impairment. *Cognitive Rehabilitation*, May-June, 10–17.

Reschly, D., & Gresham, F. M. (1989). Current neuropsychological diagnosis of learning problems: A leap of faith. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (pp. 503–520). New York: Plenum.

Reynolds, C. R. (1979). Interpreting the index of abnormality when the distribution of score differences is known: Comment on Piotrowski. *Journal of Consulting and Clinical Psychology, 47,* 401–402.

Reynolds, C. R. (1981a). The neuropsychological basis of intelligence. In G. W. Hynd & J. E. Obrzut (Eds.), *Neuropsychological assessment and the school-aged child: Issues and procedures* (pp. 87–124). New York: Grune & Stratton.

Reynolds, C. R. (1981b). Neuropsychological assessment and the habilitation of learning: Considerations in the search for the aptitude × treatment interaction. *School Psychology Review, 10,* 343–349.

Reynolds, C. R. (1982). The importance of norms and other traditional psychometric concepts to assessment in clinical neuropsychology. In R. N. Malathesha & L. C. Hartlage (Eds.), *Neuropsychology and cognition* (Vol. 3, pp. 55–76). The Hague, The Netherlands: Nijhoff.

Reynolds, C. R. (1986a). Transactional models of intellectual development, yes. Deficit models of process remediation, no. *School Psychology Review, 15,* 256–260.

Reynolds, C. R. (1986b). Clinical acumen but psychometric naivete in neuropsychological assessment of educational disorders. *Archives of Clinical Neuropsychology, 1*(2), 121–137.

Reynolds, C. R. (1992). Two key concepts in the diagnosis of learning disabilities and the habilitation of learning. *Learning Disability Quarterly, 15*(1), 2–12.

Reynolds, C. R. (1997a). Forward and backward memory span should not be combined for clinical analysis. *Archives of Clinical Neuropsychology, 12,* 29–40.

Reynolds, C. R. (1997b). Measurement and statistical problems in neuropsychological assessment of children. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (2nd ed., pp. 180–203). New York: Plenum.

Reynolds, C. R. (1997c). Postscripts on premorbid ability estimation: Conceptual addenda and a few words on alternative and conditional approaches. *Archives of Clinical Neurpsychology, 12,* 769–778.

Reynolds, C. R., & Bigler, E. D. (1994). *Manual for the Test of Memory and Learning.* Austin, TX: PRO-ED.

Reynolds, C. R., & Bigler, E. D. (1996). Factor structure, factor indexes, and other useful statistics for interpretation of the Test of Memory and Learning (TOMAL). *Archives of Clinical Neuropsychology, 11,* 29–43.

Reynolds, C. R., & Bigler, E. D. (1997a). Clinical neuropsychological assessment of child and adolescent memory with the Test of Memory and Learning. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (2nd ed., pp. 296–319). New York: Plenum.

Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior assessment system for children*. Circle Pines, MN: American Guidance Services.

Reynolds, C. R., & Kamphaus, R. W. (1997). The Kauffman assessment battery for children: Development, structure and applications in neuropsychology. In A. M. Horton, D. Wedding, & J. Webster, (Eds.), *The neuropsychology handbook* (Vol. 1, pp. 291–330). New York: Springer.

Reynolds, C. R., Kamphaus, R. W., Rosenthal, B. L., & Hiemenz, J. R. (1997). Application of the Kaufman assessment battery for children (K-ABC) in neuropsychological assessment. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (2nd ed., pp. 252–269). New York: Plenum.

Reynolds, C. R., Wilen, S., & Stone, B. (1997, November). *The economy of neuropsychological evaluations.* Paper presented at the annual meeting of the National Academy of Neuropsychology, Las Vegas, NV.

Riccio, C. A., Gonzalez, J. J., & Hynd, G. W. (1994). Attention-deficit hyperactivity disorder (ADHD) and learning disabilities. *Learning Disability Quarterly*, *17,* 311–322.

Riccio, C. A., Hall, J., Morgan, A., Hynd, G. W., Gonzalez, J. J., & Marshall, R. M. (1994). Executive function and the Wisconsin card sorting test: Relationship with behavioral ratings and cognitive ability. *Developmental Neuropsychology*, *10,* 215–229.

Riccio, C. A., & Hynd, G. W. (1995). Contributions of neuropsychology to our understanding of developmental reading problems. *School Psychology Review*, *24,* 415–425.

Riccio, C. A., & Hynd, G. W. (1996). Neuroanatomical and neurophysiological aspects of dyslexia. *Topics in Language Disorders*, *16*(2), 1–13.

Riccio, C. A., & Hynd, G. W., & Cohen, M. J. (1993). Neuropsychology in the schools: Does it belong? *School Psychology International*, *14,* 291–315.

Riccio, C. A., Hynd, G. W., & Cohen, M. J. (1996). Etiology and neurobiology of Attention-Deficit Hyperactivity Disorder. In W. Bender (Ed.), *Understanding ADHD: A practical guide for teachers and parents* (pp. 23–44). New York: Merrill.

Ris, M. D., & Noll, R. B. (1994). Long-term neurobehavioral outcome in pediatric brain tumor patients: Review and methodological critique. *Journal of Clinical and Experimental Neuropsychology*, *16,* 21.

Rothlisberg, B. A., & D'Amato, R. C. (1988). Increased neuropsychological understanding seen as important for school psychologists. *Communique*, *17*(2), 4–5.

Rourke, B. P. (1984). *Subtype analysis of learning disabilities.* New York: Guilford.

Rourke, B. P. (1989). *Nonverbal learning disabilities: The syndrome and the model.* New York: Guilford.

Rourke, B. P. (1991). *Neuropsychological validation of learning disability subtypes.* New York: Guilford.

Rourke, B. P. (1994). Neuropsychological assessment of children with learning disabilities: Measurement issues. In G. R Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 475–514). Baltimore: Brookes.

Rourke, B. P., Bakker, D. J., Fisk, J. L., & Strang, J. D. (1983). *Child neuropsychology: An introduction to theory, research, and practice.* New York: Guilford.

Rourke, B. P., Fisk, J. L., & Strang, J. D. (1986). *Neuropsychological assessment of children: A treatment oriented approach.* New York: Guilford.

Rutter, M. (1981). Psychological sequelae of brain damage in children. *American Journal of Clinical Neuropsychology, 138,* 1533–1544.

Rutter, M. (1983). *Developmental neuropsychiatry.* New York: Guilford.

Rutter, M., Graham, P., & Yule, W. (1970). *A neuropsychiatric study in childhood.* London: Lavenham Press.

Saigal, S. (1995). Long term outcome of very low-birthweight infants: Kindergarten and beyond. *Developmental Brain Dysfunction, 8,* 109–118.

Samuels, S. J. (1979). An outside view of neuropsychological testing. *Journal of Special Education, 13,* 57–60.

Sandoval, J. (1981, August). *Can neuropsychology contribute to rehabilitation in educational settings? No.* Paper presented at the annual meeting of the American Psychological Association, Los Angeles, CA.

Sandoval, J., & Halperin, R. M. (1981). A critical commentary on neuropsychology in the schools: Are we ready? *School Psychology Review, 10,* 381–388.

Segalowitz, S. (1983). *Language functions and brain organization.* New York: Academic Press.

Seidel, U. P., Chadwick, O., & Rutter, M. (1975). Psychological disorders in crippled children: A comparative study of children with and without brain damage. *Developmental Medicine and Child Neurology, 17,* 563.

Seidman, L. J., Biederman, J., Faraone, S. V., Milberger, S., Norman, D., Seiverd, K., Benedict, K., Guite, J., Mick, E., & Kiely, K. (1995). Effects of family history and comorbidity on the neuropsychological performance of children with ADHD: Preliminary findings. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 1015–1024.

Selz, M. (1981). Halstead–Reitan neuropsychological test batteries for children. In G. W. Hynd & J. E. Obrzut (Eds.), *Neuropsychological assessment of the school-aged child: Issues and procedures*, (pp. 195–235). New York: Grune & Stratton.

Selz, M. & Reitan, R. M. (1979a). Rules for neuropsychological diagnosis: Classification of brain function in older children. *Journal of Consulting and Clinical Psychology, 47,* 258–264.

Selz, M. & Reitan, R. M. (1979b). Neuropsychological test performance of normal, learning disabled, and brain damaged older children. *Journal of Nervous and Mental Disease, 167,* 298–302.

Shapiro, E. G., & Dotan, N. (1985, October). *Neurological findings and the Kaufman Assessment Battery for Children*. Paper presented at the annual meeting of the National Association of Neuropsychologists, Philadelphia, PA.

Sheslow, D., & Adams, W. (1990). *Wide Range Assessment of Memory and Learning*. Wilmington, DE: Jastak Associates.

Shields, J., Varley, R., Broks, P., & Simpson, A. (1996). Hemisphere function in developmental language disorders and high level autism. *Developmental Medicine & Child Neurology, 38,* 473–486.

Shurtleff, H. A., Fay, G. E., Abbott, R. D., & Berninger, V. W. (1988). Cognitive and neuropsychological correlates of academic achievement: A levels of analysis assessment model. *Journal of Psychoeducational Assessment, 6,* 298–308.

Snow, J. H., & Hooper, S. R. (1994). *Pediatric traumatic brain injury*. Thousand Oaks, CA: Sage.

Snow, J. H., & Hynd, G. W. (1985a). Factor structure of the Luria-Nebraska Neuropsychological Battery-Children's Revision. *Journal of School Psychology, 23,* 271–276.

Snow, J. H., & Hynd, G. W. (1985b). A multivariate investigation of the Luria-Nebraska Neuropsychological Battery-Children's Revision with learning disabled children. *Journal of Psychoeducational Assessment, 2,* 23–28.

Snow, J. H., Hynd, G. W., & Hartlage, L. H. (1984). Differences between mildly and more severely learning disabled children on the Luria Nebraska Neuropsychological Battery-Children's Revision. *Journal of Psychoeducational Assessment, 2,* 23–28.

Snyder, T. J., Leark, R. A., Golden, C. J., Grove, T., & Allison, R. (1983, March). *Correlations of the K-ABC, WISC-R, and Luria Nebraska Children's Battery for exceptional children*. Paper presented at the annual meeting of the National Association of School Psychologists, Detroit, MI.

Snyderman, M., & Rothman, S. (1987). Survey of expert opinion on intelligence and aptitude testing. *American Psychologist, 42,* 137–144.

Sperry, R. W. (1968). Hemisphere deconnection and unity in conscious awareness. *American Psychologist, 23,* 723–733.

Sperry, R. W. (1974). Lateral specialization in the surgically separated hemispheres. In F. O. Schmitt & F. G. Worden (Eds.), *The neurosciences: Third study program*. Cambridge, MA: MIT Press.

Spreen, O., & Gaddes, W. H. (1979). Developmental norms for 15 neuropsychological tests age 6–15. *Cortex, 5,* 813–818.

Spreen, O., Risser, A. H., & Edgell, D. (1995). *Developmental neuropsychology*. London: Oxford University Press.

Strom, D. A., Gray, J. W., Dean, R. S., & Fischer, W. E. (1987). Incremental validity of the Halstead–Reitan neuropsychological battery in predicting achievement for learning disabled children. *Journal of Psychoeducational Assessment, 5,* 157–165.

Sweet, J. J., Carr, M. A., Rossini, E., & Kasper, C. (1986). Relationship between the Luria Nebraska Neuropsychological Battery-Children's Revision and the WISC-R: Further examination using Kaufman's factors. *International Journal of Clinical Neuropsychology, 8,* 177–180.

Sweet, J. J., & Moberg, P. (1990). A survey of practices and beliefs among ABPP and non-ABPP clinical neuropsychologists. *The Clinical Neuropsychologist, 4,* 101–120.

Sweet, J. J., Moberg, P., & Westergaard, C. K. (1996). Five year follow-up survey of practices and beliefs of clinical neuropsychologists. *The Clinical Neuropsychologist, 10,* 202–221.

Talley, J. L. (1986). Memory in learning disabled children: Digit span and the Rey Auditory Verbal Learning Test. *Archives of Clinical Neuropsychology, 1,* 315–322.

Taylor, H. G. (1988). Neuropsychological testing: Relevance for assessing children's learning disabilities. *Journal of Consulting and Clinical Psychology, 56,* 795–800.

Taylor, H. G., Barry, C. T., & Schatschneider, C. W. (1993). School-age consequences of haemophilus influenzae type b meningitis. *Journal of Clinical Child Psychology, 22,* 196–206.

Taylor, H. G., & Fletcher, J. M. (1990). Neuropsychological assessment of children. In G. Goldstein & M. Hersen (Eds.), *Handbook of neuropsychological assessment* (pp. 228–255). New York: Plenum.

Teeter, P. A. (1986). Standard neuropsychological batteries for children. In G. W. Hynd & J. E. Obrzut (Eds.), *Child neuropsychology* (Vol. 2, pp. 187–227). New York: Academic Press.

Teeter, P. A. (1997). Neurocognitive interventions for childhood and adolescent disorders: A transactional model. In C. R. Reynolds & E. Fletcher-Janzen, *Handbook of clinical child neuropsychology* (2nd ed., pp. 387–417). New York: Plenum.

Teeter, P. A., & Semrud-Clikemen, M. (1997). *Child neuropsychology: Assessment and interventions for neurodevelopmental disorders*. Boston: Allyn & Bacon.

Telzrow, C. F., Century, E., Harris, B., & Redmond, C. (1985, April). *Relationship between neuropsychological processing models and dyslexia subtypes*. Paper presented at the annual meeting of the National Association of School Psychologists, Las Vegas, NV.

Temple, C. M. (1997). Cognitive neuropsychology and its application to children. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 38,* 27–52.

Timmermans, S. R., & Christensen, B. (1991). The measurement of attention deficit in TBI children and adolescents. *Cognitive Rehabilitation, 9,* 26.

Torgesen, J. K. (1994). Issues in the assessment of executive function: An information processing perspective. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 143–162). Baltimore: Brookes.

Torkelson, R. D., Leibrook, L. G., Gustavson, J. L., & Sundell, R. R. (1985). Neurological and neuropsychological effects of cerebral spinal fluid shunting in children with assumed arrested ("normal pressure") hydrocephalus. *Journal of Neurology, Neurosurgery, and Psychiatry, 48,* 799–806.

Tramontana, M. G. (1983). Neuropsychological evaluation in child/adolescent psychiatric disorders: Current status. *Psychiatric Hospital, 14,* 158–162.

Tramontana, M., & Hooper, S. (Eds.) (1987). *Neuropsychological assessment with children*. New York: Plenum.

Tramontana, M., & Hooper, S. (1997). Neuropsychology of child psychopathology. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child* *neuropsychology* (2nd ed., pp. 120–139). New York: Plenum.

Tramontana, M. G., Hooper, S. R., Curley, A. S., & Nardolillo, E. M. (1990). Determinants of academic achievement in children with psychiatric disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 265–268.

Tucker, D. M. (1989). Neural substrates of thought and affective disorders. In G. Gainotti & C. Caltagirone (Eds.), *Emotions and the dual brain* (pp. 225–234). New York: Springer-Verlag.

Turkheimer, E., Yeo, R. A., Jones, C., & Bigler, E. D. (1990). Quantitative assessment of covariation between neuropsychological function and location of naturally occurring lesions in humans. *Journal of Clinical and Experimental Neuropsychology, 12,* 549–565.

Voeller, K. K. S. (1995). Clinical neurologic aspects of the right hemisphere deficit syndrome. *Journal of Child Neurology, 10,* 516–522.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological process*. Cambridge, MA: Harvard University Press.

Waber, D. P., & McCormick, M. C. (1995). Late neuropsychological outcomes in preterm infants of normal IQ: Selective vulnerability of the visual system. *Journal of Pediatric Psychology, 20,* 721–735.

Wasserman, J. (1995, February). *Assessment and remediation of memory deficits in children*. Paper presented at the meeting of the Supervisors of School Psychologists for the New York Board of Education. New York.

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children-Revised*. New York: Psychological Corporation.

Welsh, M. C., Pennington, B. F., & Grossier, D. B. (1991). A normative developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology, 7,* 131–139.

Wherry, J. N., Paal, N., Jolly, J. B., Adam, B., Holloway, C., Everett, B., & Vaught, L. (1993). Concurrent and discriminant validity of the Gordon Diagnostic System: A preliminary study. *Psychology in the Schools, 30,* 29–36.

Whitten, C. J., D'Amato, R. C., & Chitooran, M. M. (1992). The neuropsychological approach to interventions. In R. C. D'Amato & B. A. Rothlisberg (Eds.), *Psychological perspectives on intervention: A case study approach to prescriptions for change* (pp. 112–136). New York: Longman.

Williams, M. A., & Boll, T. J. (1997). Recent advances in neuropsychological assessment of children. In G. Goldstein & T. M. Incagnoli (Eds.), *Contemporary approaches to neuropsychological assessment* (pp. 231–267). New York: Plenum.

Willson, V. L., & Reynolds, C. R. (1982). Methodological and statistical problems in determining membership in clinical populations. *Clinical Neuropsychology, 4,* 134–138.

Wittelson, S. F. (1977). Early hemisphere specialization and interhemispheric plasticity: An empirical and theoretical review. In S. Segalowitz & F. A. Gruber (Eds.), *Language development and neurological theory* (pp. 213–287). New York: Academic Press.

Woody, R. H. (1997). Psycholegal issues for clinical child neuropsychology. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (2nd ed., pp. 712–725). New York: Plenum.

Ylvisaker, M., Chorazy, A. J. L., Cohen, S. B., Mastrilli, J. P., Molitor, C. B., Nelson, J., Szekeres, S. F., Valko, A. S., & Jaffe, K. M. (1990). Rehabilitative assessment following head injury in children. In M. Rosenthal, E. R. Griffith, M. R. Bond, & J. D. Miller (Eds.), *Rehabilitation of the adult and child with traumatic brain injury* (2nd ed., pp. 521–538). Philadelphia: Davis.

Zurcher, R. (1995). Memory and learning assessment: Missing from the learning disabilities identification process for too long. *LD Forum, 21,* 27–30.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.11
Neuropsychological Assessment of Adults

C. MUNRO CULLUM
University of Texas Southwestern Medical Center at Dallas, TX, USA

4.11.1 INTRODUCTION 304
4.11.2 APPROACHES TO ASSESSMENT IN CLINICAL NEUROPSYCHOLOGY 305
  4.11.2.1 The Standard Battery 305
  4.11.2.2 Test Batteries for Specific Populations 308
  4.11.2.3 The Hypothesis-driven Approach 308
  4.11.2.4 Quantitative and Qualitative Examinations of Neurobehavioral Competence 309
  4.11.2.5 Other Issues in Test Interpretation 309
    4.11.2.5.1 Test cutoff scores 310
    4.11.2.5.2 T-scores 310
  4.11.2.6 Computer Interpretation of Neuropsychological Assessment Results 310
  4.11.2.7 Cognitive Screening 311
4.11.3 METHODS OF NEUROPSYCHOLOGY 311
  4.11.3.1 Clinical Interview and Background Information 312
  4.11.3.2 Neuropsychological Measurement of Brain Function 312
4.11.4 PRINCIPAL COGNITIVE DOMAINS FOR ASSESSMENT 315
  4.11.4.1 Global Cognitive/Intellectual Functioning 315
  4.11.4.2 Academic Achievement 322
  4.11.4.3 Executive Functioning, Problem-solving, and Reasoning 323
    4.11.4.3.1 Tests of reasoning and problem-solving 323
  4.11.4.4 Arousal and Orientation 324
  4.11.4.5 Attention/Concentration 325
    4.11.4.5.1 Assessment of attention 325
  4.11.4.6 Language 326
    4.11.4.6.1 Global language assessment 327
  4.11.4.7 Visuospatial Skills 328
    4.11.4.7.1 Visuospatial tasks 329
  4.11.4.8 Memory 332
    4.11.4.8.1 Clinical assessment of learning and memory 332
    4.11.4.8.2 Clinical memory tests 335
    4.11.4.8.3 Memory batteries 335
    4.11.4.8.4 Verbal memory tests 336
    4.11.4.8.5 Nonverbal memory tests 337
  4.11.4.9 Motor and Sensory Function 338
    4.11.4.9.1 Psychometric measures of motor and sensory function 338
  4.11.4.10 Assessment of Motivation 339
  4.11.4.11 Personality and Emotional Functioning 340
  4.11.4.12 Test Selection Issues 341
  4.11.4.13 Relationships Between Neuropsychometry and the Behavioral Geography of the Brain 341

4.11.4.14 Training in Neuropsychology 342
4.11.4.15 Challenges for Neuropsychological Assessment 343
4.11.5 REFERENCES 343

4.11.1 INTRODUCTION

Clinical neuropsychology represents one of the most rapidly growing areas within the field of psychology and was one of the first to gain specialty status by the American Psychological Association in 1996. At a basic level, the neuropsychological examination represents a combination of the traditional behavioral neurologic examination with a psychometric approach to the evaluation of brain–behavior relationships. From the broader field of psychology, neuropsychology derived its emphasis in the application of psychometric procedures to quantify behavior. From its other parent discipline of neurology came the interest in evaluating brain function. The term "neuropsychology" originally evolved in the 1930s and 1940s and its popularization is often attributed to Hans-Lukas Teuber. Many early neuropsychological procedures were developed during war time to assess cognitive status and suitability of individuals for special military service. Subsequently, penetrating missile wounds to the brain became the focus of localization studies. Accordingly, many tests were created with the goal of being sensitive to focal brain insults. Other measures were developed to assess for "organicity," a now archaic term that had been used to grossly refer to brain damage or neurological deficit. Prior to the advent of modern neuroimaging procedures, neuropsychological techniques emerged as a front-line assessment procedure for the identification and localization of acquired cerebral damage. While neuropsychological techniques continue to provide this aspect of neurodiagnostic assessment to some extent, more commonly the procedures are used to describe and quantify behavior. The results of these evaluations can be used to infer cerebral integrity vs. dysfunction, to delineate cognitive strengths and weaknesses, and to assist in differential diagnosis. Assisting other professionals in differential diagnostic situations and documenting level of impairment can be a primary role for these evaluations. Other goals of neuropsychological assessment include making treatment and rehabilitation recommendations, assisting in placement issues, and in evaluating treatment response (e.g., in cases of neurosurgical or pharmacological intervention).

The neuropsychological examination can also provide useful information to patients regarding their functioning, and feedback may be used therapeutically in terms of adjustment issues. In many cases, informing the patient that their cognitive difficulties can be documented and explained (or at least are not unexpected) can be most reassuring. Discussing the nature of a disorder and the various cognitive and emotional sequelae it may have, as well as the typical course of recovery or changes that can be expected, is also helpful in terms of setting realistic expectations and goals. Furthermore, presenting information and potential compensatory strategies or interventions that are helpful given a patient's particular situation can be rewarding. In some medical settings in particular, the neuropsychologist may play a central role in helping patients understand the nature of their difficulties by providing this information in understandable terms within an emotionally supportive context.

Among existing neurodiagnostic procedures, the neuropsychological evaluation remains the most sensitive means of assessing human brain function. Knowing the presence and location of a given lesion, for example, provides only limited information about an individual's functioning. For example, the patient in Figure 1 sustained an infarction of the right posterior cerebral artery.

How such a patient might be functioning in his or her daily life, however, remains a question that simple structural neuroimaging cannot address. For example, the individual depicted in Figure 1 might be expected to show a contralateral visual field cut (hemianopsia) based on the location of the lesion, although the impact of this as well as any other associated cognitive processing difficulties would be unknown without clinical examination.

Another example is provided in the case of the 65-year-old patient whose magnetic resonance imaging (MRI) scan is presented in Figure 2. This individual demonstrated grossly normal brain structure, with no major evidence of atrophy or particular neuropathology. Despite the normal appearance of her brain, a severe level of dementia was observed as reflected by her score of 10/30 on the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975). Although dementia is often associated with cortical atrophy, clinicians should keep in mind that dementia is a clinical diagnosis and the degree of atrophy on neuroimaging shows only a modest association

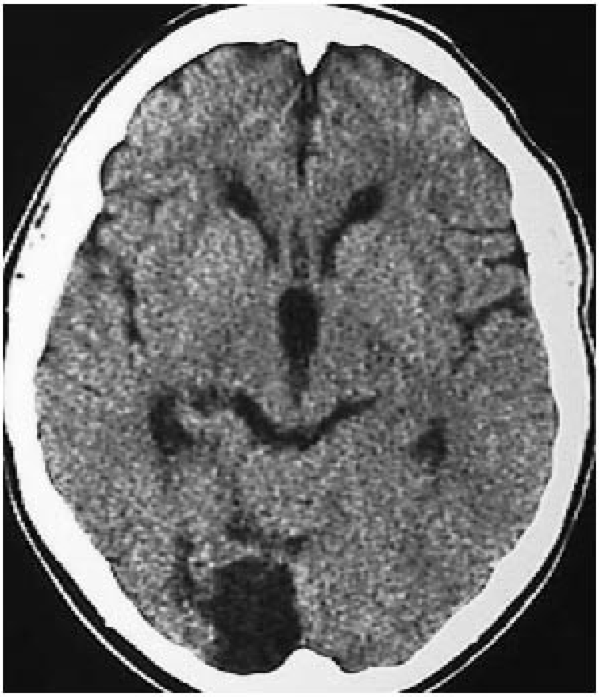

Figure 1 CT scan depicting posterior right hemisphere lesion (note that right is depicted on the left).

with level of cognitive impairment (Naugle, Cullum, Bigler, & Massman, 1986).

Given the utility of neuropsychological techniques in terms of documenting and understanding the neurobehavioral sequelae of cerebral dysfunction, a discussion of some of the major approaches and common measures used in clinical neuropsychology is in order.

4.11.2 APPROACHES TO ASSESSMENT IN CLINICAL NEUROPSYCHOLOGY

Contemporary clinical neuropsychology owes a debt to many of its early founders, and because of the youth of the field, many of these individuals remain heavily involved in the field. A listing of the major contributors to the field of clinical neuropsychology is beyond the scope of this chapter and involves a wide variety of individuals from different theoretical backgrounds and disciplines. In terms of major contributors to the development of neuropsychological assessment, such a list would have to include the following: Benton, Butters, Goldstein, Goodglass, Halstead, Kaplan, Meier, Parsons, Reitan, and Spreen. Citations and representative examples of their work can be found in various textbooks and merit review by those interested in obtaining more information about clinical neuropsychology (e.g., see Grant & Adams, 1996; Lezak, 1995; Naugle, Cullum, & Bigler, 1998; Spreen & Strauss, 1998).

The concept of "approaches" to neuropsychological assessment can be considered from several perspectives. Historically, the "fixed" vs. "flexible" test battery approach was a topic of discussion and even heated debate for many years. Since many contemporary neuropsychologists tend to utilize a "core" battery of favorite measures for many of the patients they examine, but modify the assessment by adding or deleting specific tests depending upon the clinical population and/or referral issues in question, this has become an essentially moot issue. Furthermore, the modification of even the most "fixed" of test batteries to address individual patient needs and referral questions is now commonplace. Thus, the distinction between fixed and flexible approaches to neuropsychological assessment is somewhat arbitrary, although some discussion of the central issues involved in test selection and the composition of "standard" or core test batteries is in order.

4.11.2.1 The Standard Battery

In its strictest sense, a fixed standard battery approach involves the administration of the same set of tests to all patients, regardless of

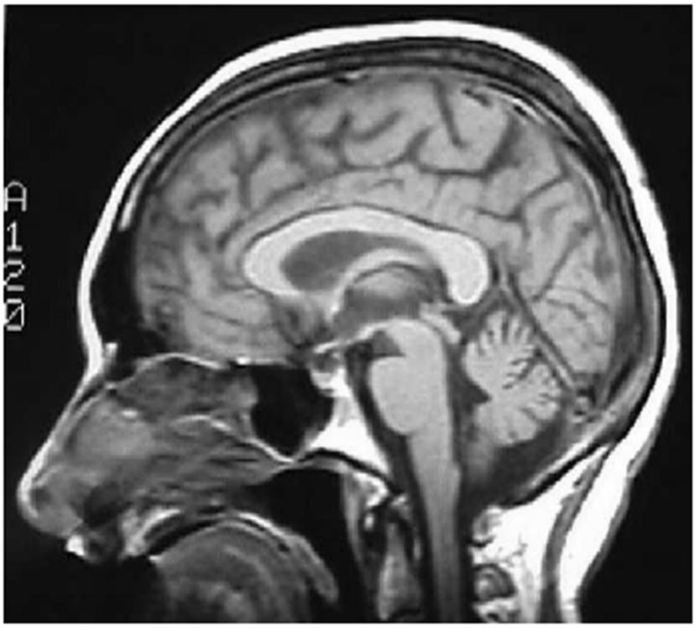

Figure 2 MRI depicting normal gross brain structure in a patient with severe dementia.

diagnostic or referral question, level of impairment, patient complaints, or clinical presentation. Applied too rigidly, such an approach may not yield appropriately detailed information regarding a specific area of deficiency (e.g., memory) unless that ability area is adequately represented in the standard or core battery. The application of an extensive omnibus test battery to all individual cases may also prove less efficient in some instances by oversampling behavior. For example, in the case of a rigid standard battery, a lack of impairment on a higher level ability measure may be followed by the routine administration of a similar task that tends to be less sensitive to deficits in that same domain. However, it has also been argued that a standard battery approach may allow for a broader assessment of neurobehavioral abilities in some settings, and hence, may detect cognitive impairment that a less comprehensive battery might miss. For example, a patient referred for memory assessment might be administered only a test of global intelligence (e.g., Wechsler Adult Intelligence Scale-Revised; WAIS-R) and selected measures of memory (e.g., the Logical Memory and Visual Reproduction subtests from one of the versions of the Wechsler Memory Scale are commonly used in various settings). Such an evaluation may not be sensitive to deficits in executive functioning or higher-order cognitive integration that might be manifest upon more detailed assessment and that can impact memory test performance. It should be noted, however, that the limited comprehensiveness of the battery would be a more important issue than whether this represents a fixed or more flexible assessment approach.

Perhaps the best example of a standard neuropsychological battery is the group of tests that was originally created in the 1940s by Ward Halstead and further developed by one of his most prominent students and colleagues, Ralph Reitan. Halstead had strong interests in the assessment of biological intelligence and in the quantification of brain–behavior relationships, particularly at a time when localization of cerebral lesions was often a focus of neuropsychological evaluations. The work of Halstead and Reitan contributed monumentally to the development of American neuropsychology, and the Halstead–Reitan Neuropsychological Battery (HRB) continues to be used in many settings, whether used *in toto* or by its component tests.

The core HRB consists of the following tests and the primary corresponding cognitive functions (also see Dodrill, 1997): Category Test (abstract reasoning, logical analysis), Tactual Performance Test (complex psychomotor problem solving); Trail Making Test (psychomotor speed and cognitive sequencing); Aphasia Screening Test (brief assessment of language and graphomotor construction); Speech Sounds Perception Test (verbal/auditory attention); Seashore Rhythm Test (nonverbal auditory attention); Finger Tapping Test (fine motor speed); Grip Strength (gross motor strength); Sensory-Perceptual Examination (basic sensory and perceptual abilities); and Tactile Form Recognition (sensory perception, agnosia). Composite summary indices from the core

HRB can also be calculated in order to provide an overall index of cognitive impairment. These include the General Neuropsychological Deficit Scale (GNDS), which is derived from 42 scores from the HRB (0 = no impairment to a maximum possible score of 168; see Reitan & Wolfson, 1993), the Average Impairment Rating (Russell, Neuringer, & Goldstein, 1970) which ranges from 0.00 (no impairment) to 5.00, and the original Halstead Impairment Index (HII; Halstead, 1947), which utilizes a proportion of HRB test results based upon cutoff scores from selected tests (0.0 = no impairment to 1.0, reflecting impairment on all of the key variables). Like IQ scores, such summary indices suffer from a number of limitations (e.g., see Lezak, 1995, pp. 23–24), although they can arguably provide some useful reference information regarding an individual's overall functioning if not overinterpreted or used too rigidly.

Reitan and Wolfson (1996) provide a detailed overview of the HRB in terms of its application and interpretation, and many of the assessment and interpretation issues they discuss have relevance beyond the HRB *per se*. For example, the authors include a discussion of four of the principal ways in which test data can be interpreted: (i) level of performance; (ii) pathognomonic signs (i.e., specific test findings/deficits that strongly suggest brain dysfunction and are uncommon in nonbrain injured subjects; (iii) patterns of test results (e.g., those typifying different disorders); and (iv) comparison of test results using each side of the body (e.g., right vs. left hand) to examine hemispheric lateralization issues.

Clinicians vary with respect to their relative reliance upon each of these interpretive strategies, and depending upon the case, some findings may need to be given greater weight than others. For example, the pathognomonic sign of focal motor weakness in the absence of a peripheral motor injury clearly merits careful attention to the possibility of contralateral cerebral dysfunction. Similarly, a visual field cut (e.g., hemianopsia) should alert the clinician to focal cerebral involvement in the appropriate corresponding neuroanatomical area.

In other cases, the pattern of test results can be more important than the level of performance. Test findings that fall technically within normal limits may show a pattern that is suggestive of cerebral dysfunction. To illustrate this point using a more familiar measure, a patient may demonstrate overall scores within normal limits on the WAIS-III or WAIS-R, yet show relative weaknesses on those subtests most sensitive to cerebral dysfunction which may reflect neurologically-based deficits. Similarly, a normal summary score on other neuropsychological tests can occur even when subscores and/or qualitative performance characteristics suggest clear abnormality or demonstrate a pattern consistent with a particular disorder.

In practice, many clinicians who utilize the HRB administer a modified version of the battery that selectively includes individual HRB tests. For example, many settings that do not routinely use the core HRB nevertheless administer the Category Test as a measure of higher-order problem-solving and because of its high sensitivity to cerebral dysfunction. Alternatively, clinical evaluations relying upon the complete HRB often include additional measures to provide a more detailed assessment of areas of cerebral functioning not well represented in the core battery. For example, one area that often is supplemented is memory, since the core HRB does not provide much in the way of memory assessment, even though this represents a common complaint among neurological populations.

Luria's Neuropsychological investigation (Christensen, 1984) represents a battery of measures that involves the systematic application of the assessment techniques originally described by Luria (1966). The battery is organized into various sections assessing an array of abilities, including basic and higher-level motor functions, receptive and expressive language, memory, and higher-order thinking. Even though it is presented here as a "standard" battery, it involves a sequential hypothesis-testing approach that includes a variety of individual tasks and relies upon qualitative judgments. Many clinicians utilize some of the component tasks selectively in their examinations. Some of the more commonly used tasks are the "go-no-go" paradigm (e.g., "When I hold up one finger, you hold up two," and vice-versa) and what has come to be called the "Luria 3-step" motor sequence by some (i.e., rapidly alternating the hand from flat to fist to edge) in order to assist in the examination of aspects of executive functioning and cognitive initiation/inhibition.

A very different type of standard test battery involves computer administration. Although several such test batteries exist, one example is the Microcog: Assessment of Cognitive Functioning (Powell et al., 1993). This is a computer-administered and scored battery of tests designed to screen for gross cognitive impairment in adults. It is available in short and long forms consisting of 12 and 18 subtests that assess the following general areas: attention/mental control, memory, reasoning/calculation, spatial processing, and reaction time. The test software calculates standard scores and percentile comparisons using age-referenced norms that range from 18 to 89. While Microcog may be useful in some screening situations, the limitations of the brief and selective nature of the tasks, the lack of alternate forms, and the limitations inherent in computer-based assessment must be carefully considered in light of the specific clinical questions and patient populations in need of evaluation.

4.11.2.2 Test Batteries for Specific Populations

Various standard test batteries have been developed and/or assembled for specific purposes and populations, and many of these are reviewed in the excellent neuropsychological test compendiums by Spreen and Strauss (1998) and Lezak (1995). A few of the more commonly used batteries include the following:

(i) The consortium to establish a registry for Alzheimer's disease (CERAD) neuropsychological assessment battery (Morris, Heyman, Mohs, et al., 1989) was developed for use in patients with known or suspected dementia. This is a 30–45 minute examination of cognitive skills that includes measures of orientation, verbal fluency, naming, verbal learning and memory, and graphomotor constructional skills. Good reliability and validity have been demonstrated, and normative data for older adult populations are available (Welsh et al., 1994).

(ii) The NIMH core neuropsychological battery (Butters et al., 1990) was developed to evaluate cognitive changes associated with HIV infection. Both a brief (1–2 hour) and an extended (7–9 hours) battery comprised of standard clinical neuropsychological tests and measures used in cognitive psychology were assembled in order to assess the following cognitive ability areas: premorbid IQ, attention, speed of information processing, learning and memory, abstract reasoning, language, visuoperceptual and graphomotor constructional abilities, and psychomotor skills. It includes tasks that provide information regarding patterns of cognitive deficits as commonly seen in degenerative conditions that affect cortical and particularly subcortical cerebral functions.

(iii) The Pittsburgh Occupational Exposures Test battery (Ryan, Morrow, Bromet, & Parkinson, 1989) was developed to evaluate those cognitive functions most commonly affected by exposure to environmental toxins. It comprises a series of 15 standard and experimental neuropsychological measures that assess a variety of abilities in a reasonably brief amount of time (i.e., approximately 90 minutes). Factor analysis of the battery revealed the following five domains: general intelligence, attention, learning and memory, visuospatial, and psychomotor speed and manual dexterity.

As noted, many neuropsychologists develop their own groups of tests for selected populations, such that a core battery for dementia might have some overlap with a battery for epilepsy (e.g., perhaps in terms of intellectual assessment). Such batteries would likely have different components and stepdown procedures, however. For example, additional or more comprehensive memory measures might be selected for patients with epilepsy when lateralized temporal lobe dysfunction is known or suspected. Furthermore, an assessment of academic achievement skills would typically be more important in the young adult with epilepsy than in the older individual with Alzheimer's disease. Thus, the issue of "core" test batteries depends largely upon neuropsychological training and experience, and many clinicians will use somewhat different sets of tests for different populations. In any case, it is important for the neuropsychologist to be familiar with a wide array of neuropsychological measures, not only in order to stay current with new test developments and research, but also to have a large test repertoire from which to select measures for individual or unique cases, as well as to be best prepared to review test results from other centers.

4.11.2.3 The Hypothesis-driven Approach

The hypothesis-driven or more flexible approach to neuropsychological assessment espouses the selection of tests based on referral questions, known or suspected diagnoses/pathology, and individual patient complaints and symptom presentation. Such examinations tend to be more "customized" or individually tailored, even though a relatively standard "core" battery of tests may be included as part of the evaluation. Tests may also be added or removed from the planned evaluation as the testing process progresses, depending upon performance results and patterns of strengths and weaknesses observed on various measures. As noted, many neuropsychologists administer a core group of tests to many or most of the patients they examine, subsequently tailoring aspects of the evaluation to follow up on areas of particular clinical interest. For example, a patient referred for evaluation of memory complaints might undergo a standard battery of measures that provides an evaluation of multiple cognitive domains, although the focus of the examination might be a more extensive assessment of various aspects of memory designed to thoroughly examine the patient's complaints or known/suspected pathology. Care must obviously be used in test selection, lest an examination become too cursory or overly focused, or, at the other extreme, too lengthy and redundant. For example, an examination that consisted only of memory measures to evaluate a patient's memory complaints would neglect other areas of functioning (e.g., attention/concentration, executive functioning, language disturbance) that might play a prominent role in an individual's complaints of memory problems. Another issue is that once an area of deficiency is identified, the question arises as to how much exploration/assessment of the deficit is necessary. Certainly addressing the referral question is of great importance, and to the extent that this has been addressed, the clinician may decide to stop the evaluation at that point. In other cases, however, such as in the differential diagnosis of Parkinson's disease-related cognitive decline vs. depression vs. early Alzheimer's disease, more detailed exploration of multifaceted aspects of memory and other neurocognitive functioning would be required beyond the question, "Is there evidence of cognitive impairment?" that is often posed in certain settings.

4.11.2.4 Quantitative and Qualitative Examinations of Neurobehavioral Competence

Quantitative assessment refers to the use of test scores for interpretation, and qualitative examination often refers more to observations that pertain to the process by which patients perform tasks. To illustrate this latter point, even though a particular score on a test reflects a certain level of proficiency, that level of achievement may be attained via different cognitive strategies. An inefficient trial-and-error approach to a problem-solving task, for example, may yield the same final result as a more systematized strategy, even though these approaches reflect different underlying processes. Furthermore, the reliance upon either of these processes may have very different implications for patients in terms of their effectiveness in everyday functioning and in their ability to cope with novel situations. This approach to neuropsychological assessment is perhaps nowhere better illustrated than in the Boston Process Approach to clinical neuropsychology (see Milberg, Hebben, & Kaplan, 1996, for an overview) which incorporates an examination of the process or manner in which tasks are performed rather than relying upon final or summary scores alone. In some approaches to neuropsychological assessment, there is an emphasis on specific test scores, with less attention paid to how scores are achieved, the cognitive functions underlying those scores, the relative preservation of component functions that may be indicated, and the manner in which those scores may relate to the patient's functioning in everyday life situations.

As noted, many clinical neuropsychologists implement aspects of flexible and standard, quantitative and qualitative approaches to assessment in practice, and these should not be viewed as mutually exclusive. Perhaps more important is the approach to test interpretation and performance analysis, which represents a multidimensional process. As noted earlier, the dimensions of level and pattern of performance on neuropsychological tests are very important to consider. In some cases, the level of performance may suggest normal functioning, while the pattern of test results suggests the presence of an abnormal process. Along these lines, it is important to keep in mind that scores in the "normal" or "average" range may reflect impairment in some individuals compared to their premorbid or baseline level of functioning. For example, the average IQ and memory scores in the previously high-functioning college professor who sustained a traumatic brain injury (TBI) may represent a significant decline in functioning, even though the obtained scores remain within the "normal" range (e.g., see Naugle, Cullum, & Bigler, 1990). Furthermore, if the patient achieves a good score on a particular test, but does so in an abnormal manner, this, too, may reflect altered mental function. Careful consideration of these factors, along with the pattern of test results and the degree to which the findings are consistent with deficits commonly seen in TBI, are essential in arriving at a correct diagnosis. Furthermore, the quantitative and qualitative information gained through a comprehensive neuropsychological examination may be very important in providing appropriate feedback and making realistic recommendations to the patient. Thus, as with the combination of standard and flexible test batteries, utilization of quantitative as well as qualitative data from the neuropsychological evaluation can be of extreme importance.

4.11.2.5 Other Issues in Test Interpretation

Regardless of a clinician's primary approach to neuropsychological assessment, the determination of what tests are to be given to each patient requires forethought and planning, and may vary depending upon the patient's age, level of education, socioeconomic status and ethnic background, and clinical presentation. For example, the neuropsychological evaluation of a retired physician following a mild brain injury would require the administration of a different set of measures than an 87-year-old laborer referred for diagnostic confirmation of

Alzheimer's disease who is known to have at least a 10-year history of progressive decline. As a more extreme example, the workup of an elderly individual with three years of education and limited English speaking abilities would require careful test selection to avoid measures with high cultural and educational biases.

4.11.2.5.1 Test cutoff scores

The use of "cutoff" or "cut" scores is another issue that merits comment, because this comes up frequently in clinical practice and often serves as a topic of debate. The cutoff score is based on the notion that most normal individuals tend to perform above a certain level on a given test, while most brain-injured individuals score below that level. From a statistical perspective, the notion of cutoff scores holds some merit, since many test results are normally distributed, and depending upon sensitivity and specificity qualities desired, cut points in a distribution can be set in order to maximize one or the other or both. This approach will produce definable true-positive and false-positive hit rates for groups, although statistical inferences may differ when applied to any given individual case. As a result, definitions of what constitutes "impairment" vs. "normalcy" may vary depending upon a variety of factors such as age and education, to take two of the more extensively studied demographic variables. Different cutoff scores for impairment would be needed to adjust for the effects of age on a given test, since age has a significant effect on so many cognitive measures. Using the original cutoff score of 90 seconds for the Trail Making Test—Part B, for example, would result in many healthy elderly individuals being misclassified as "impaired," when in fact their performances on this test may fall well within normal limits or even above average for their age. Similar considerations must be made with respect to different educational levels and estimated premorbid intellectual functioning, which can have a profound impact on the interpretation of certain neuropsychological test scores.

4.11.2.5.2 T-scores

Heaton, Grant, and Matthews (1991) provide normative reference values (*t* scores) for an extensive battery of tests that includes the HRB. This work, and the accompanying supplement for the WAIS-R (Heaton, 1992), represents a monumental contribution to the field by providing age-, education-, and where appropriate, gender-adjusted standard scores for an array of neuropsychological tests. The use of demographically-corrected *t* scores (derived from the same normative population) places all test results on the same metric (mean = 50, SD = 10, with lower scores reflecting poorer performances), thereby allowing for ready comparisons across areas of cognitive functioning. For example, questions such as, "Is memory more impaired than expected given an individual's background and overall level of functioning?" can be addressed through the use of such standardized scores. Also, the effects of age and education on specific test results can be more readily appreciated, even when such factors may be thought to have little influence on a given task. For example, if it takes a 30 year old female with a high school education 90 seconds to complete Trails B, the corresponding *t* score is 38, which reflects a mild impairment, just over one standard deviation below average. If that individual were 75 years old, however, that same level of performance would fall in the above-average range (*t* score = 56).

The use of *t* scores allows for an individual's performance on specific tests to be compared with results from healthy groups of subjects of similar gender, age, and educational backgrounds rather than relying upon strict cutoff scores in order to help determine the degree of normality–abnormality of findings. If enough scores fall below expectation and/or form a consistent pattern of deficits, the likelihood of cerebral dysfunction is increased.

Given the error variance of any behavioral measure, *t* scores, like IQ or other standard scores, should not be used as rigid neurobehavioral markers that are absolute or "true" scores, but rather as interpretive guidelines to assist in evaluating levels and patterns of performance in a particular case. Furthermore, caution must be used in the strict application of interpretive scores and guidelines derived from normal populations to cases of brain-injured individuals (e.g., see Reitan & Wolfson, 1996). Another potential risk of over-reliance upon standardized scores in neuropsychological interpretation is that it can potentially give the (particularly inexperienced) clinician a false sense of security, when in fact, it is the neuropsychologist's training, knowledge base, experience, and skill in interpreting various aspects of test performance and behavior that result in valid neuropsychological conclusions regarding cerebral integrity.

4.11.2.6 Computer Interpretation of Neuropsychological Assessment Results

Computer-derived interpretive programs have been developed for several standard test batteries (e.g., HRB, Luria–Nebraska Neuropsychological Battery). While such programs attempt to make general statements regarding the likelihood and even pattern of cerebral dysfunction based upon normative and neuropathological data and "typical" profiles, these programs may lend a false sense of interpretive security to those with more limited training and experience in clinical neuropsychology. While certain features of some such programs are arguably useful in certain situations (e.g., rapid scoring and normative referencing of results), they must be used with caution. Because brain damage affects individuals in different ways (i.e., a lesion in a specific location may produce different symptoms across individual patients depending upon a host of interindividual neuroanatomic, neuropathologic, genetic, experiential, and personality factors), any blanket statements about the nature of neurobehavioral disturbance in an individual must be examined carefully within the context of the patient in question. Particular care must be used when computer interpretations yield purported localization indices, since this process requires knowledge of underlying neuroanatomical systems, neuropsychiatric disorders, and individual neurobehavioral variations in order to optimize diagnostic accuracy. Attempts to distill numerically the complex process of clinically weighing various test scores, combinations of results, and qualitative performance features into "focal" interpretive patterns is a most challenging prospect, indeed, and erroneous interpretations can be rendered through over-reliance upon some of the available test interpretation programs.

4.11.2.7 Cognitive Screening

In keeping with the current managed-care *Zeitgeist*, neuropsychological evaluations should be appropriately detailed and comprehensive, while at the same time designed to provide the maximum amount of information in a time- and cost-efficient manner. It should be kept in mind, however, that brief cognitive assessments should not be carried out at the expense of thoroughness. In some cases, a brief examination may be adequate to address referral questions, as in the case of documenting cognitive impairment in a patient with severe dementia. A thorough examination of such a patient might be accomplished readily through the use of cognitive screening tools, and in fact, the administration of an eight- or 10-hour battery of measures in such a case might not yield any more diagnostic or clinically useful information than a brief assessment. However, evaluations that are overly brief may be insensitive to mild or even more significant cognitive deficits. Take, for example, the ever-popular MMSE, which is arguably the most widely used cognitive screening tool and provides a very brief assessment of orientation, simple language, and recent memory skills. Whereas such instruments have utility in quantifying gross level of impairment when more severe brain dysfunction exists, scores that fall in the "normal" range on this test do not rule out cognitive abnormality. That is, simple screening measures such as the MMSE tend to have a high false-negative rate (i.e., identifying patients as being intact, when in fact, they are not). For example, diagnostic error rates using traditional cutoff scores on the MMSE (i.e., <24/30) can be around 15% even among patients with a diagnosis of possible or probable Alzheimer's disease (Cullum & Rosenberg, 1998); in less impaired populations, the false-negative rate likely will be higher. Although the use of cognitive screening tools like the MMSE is preferable to a more idiosyncratic and non-standardized clinical mental status evaluation, the limitations of any brief cognitive screening tool must be kept carefully in mind. Along these lines, as with other measures of cognitive function, the use of cutoff or specific reference scores on such tests must be carefully considered in light of the individual patient's background. Demographic factors often influence cognitive tests, and appropriate norms should be used even in the case of brief screening tools (e.g., see Crum, Anthony, Bassett, & Folstein, 1993 for age- and education-corrected MMSE scores).

4.11.3 METHODS OF NEUROPSYCHOLOGY

Neuropsychological assessment provides the most sensitive and comprehensive means by which brain function and cognition can be assessed given our current state of technology. The techniques are used widely in various settings across many patient populations, and are recognized by the American Academy of Neurology as accepted diagnostic procedures (Report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology, 1996). Neuropsychologists often serve as consultants to neurologists, neurosurgeons, psychiatrists, clinical psychologists, and other medical and health professionals.

Neuropsychological tests are the means by which neurobehavioral samples are elicited, and specific test scores represent summary statements about observed behaviors and cognitive capabilities. Obtaining valid behavioral samples is critical, as is a thorough knowledge of the range of responses made by patients with and without various neurobehavioral disorders. As noted in neuropsychology training programs, neuropsychological tests are no better than the clinician interpreting them. Thus, the best tests in poor or inadequately trained hands may fail to yield the correct neurobehavioral conclusions, and it is the clinician that is the instrument of evaluation—not the tests alone. Appropriate specialized training in neuropsychology is essential to conduct detailed neuropsychological evaluations beyond a cognitive screening level, and this is the topic of a section later in this chapter.

4.11.3.1 Clinical Interview and Background Information

Prior to a discussion of individual cognitive domains and specific representative measures, an overview of some of the other information that is essential to the neuropsychological examination is in order. In addition to a formal examination of mental functions, information pertaining to the patient's history, presenting complaint, and overall functional status is essential. In order to interpret results from any neurobehavioral task, an understanding of the behavioral geography of the brain and the underlying cerebrocortical systems is critical. It should be kept in mind that many cognitive processes represent inter-related phenomena and that individual tasks seldom assess highly specific skills or involve circumscribed cerebral areas in isolation. Second, background variables of the patient, including their age, level of education, SES, and cultural background, must be kept in mind while asking them to perform various tasks. Third, it should be kept in mind that the clinician interpreting the test data is the key to a good neuropsychological evaluation. Even the most sensitive tests in the hands of the inadequately trained clinician may yield erroneous conclusions. This is particularly true in cases of subtle behavioral abnormalities that may only be elicited through the use of specific cognitive tasks or detected via the careful and experienced observation of characteristic symptoms or behavior patterns. Table 1 provides an overview of much of the information that should be collected as part of the clinical interview within the context of the typical neuropsychological evaluation. Such data form a basis for the selection of assessment procedures as well as interpretation and integration of findings.

After obtaining this information, the selection of neuropsychological tests can be done, with particular areas of interest being stressed depending upon the referral question, nature of the disorder(s) under consideration, and patient complaints and presentation. In many settings, neuropsychological tests are administered by highly trained psychometrists or technicians. These individuals are responsible for standardized test administration and scoring under the supervision of the clinical neuropsychologist. This is a recognized standard practice in the field that is endorsed and/or utilized by a majority of neuropsychologists (Sweet & Moberg, 1990).

Table 1 Cognitive domains typically assessed by the comprehensive neuropsychological evaluation.

1	Global cognitive/intellectual functioning
2	Academic achievement
3	Executive functioning/problem-solving
4	Language
5	Visuospatial processing
6	Attention/concentration
7	Learning and memory
8	Psychomotor functions
9	Sensory perceptual abilities
10	Psychological functioning

4.11.3.2 Neuropsychological Measurement of Brain Function

The comprehensive neuropsychological evaluation includes the administration of tests that comprise multiple cognitive domains. Although the specific tests selected by individual neuropsychologists may vary, the domains listed in Table 2 are commonly represented.

Excellent descriptions and reviews of the most frequently used cognitive measures (as well as normative data for some tests) can be found in Lezak (1995) and Spreen and Strauss (1998), and the reader is encouraged to consult these comprehensive reference sources. Table 3 presents some of the more commonly used tests to assess these cognitive domains. It should also be noted that there may be significant overlap between some of the listed tests, and across domains. The tests are listed for illustrative purposes and are not intended to provide an exhaustive list or to reflect a prototypical clinical neuropsychological test battery.

The question sometimes arises as to why a neuropsychological evaluation cannot be extremely focused, that is, "Why not administer a single test of memory to a patient if that is the only presenting complaint?" The answer obviously is that by using only one test or assessing only one cognitive domain, the clinician might completely overlook the actual cause of the patient's reported difficulties. For example, an individual presenting for memory disturbance may actually have primary difficulties in attention and concentration which preclude their ability to efficiently encode new material. When that individual attempts to recall information at

Table 2 Common neuropsychological measures.

Representative measures	*Functions assessed*
Intellectual	
Wechsler Adult Intelligence Scale-3 (WAIS-III) and Wechsler Adult Intelligence Scale-Revised (WAIS-R)	Global cognitive capacity/ Intelligence
National Adult Reading Test-Revised (NART-R)	Premorbid IQ estimation
Global cognitive	
Dementia Rating Scale (DRS)	Presence and level of dementia
Mini-Mental State Examination (MMSE)	Gross cognitive screening
Academic achievement	
Wide Range Achievement Test-3 (WRAT-3)	Reading, spelling, math
Pearowstart Individual Achievement Test-Revised (PIAT-R)	Reading, spelling, math
Wechsler Individual Achievement Test (WIAT)	Basic academic skills
Executive function and problem-solving	
Wisconsin Card Sorting Test (WCST)	Cognitive flexibility/problem-solving
Category Test	Abstraction and reasoning
California Sorting Test (CST)	Cognitive flexibility/idea generation
Similarities (WAIS-III, WAIS-R)	Abstraction
Trail Making Test (Part B)	Mental sequencing/flexibility
Raven's Matrices (standard, colored, advanced)	Nonverbal reasoning
Language	
Vocabulary (WAIS-R)	Word knowledge
Boston Naming Test (BNT)	Confrontation naming
Word fluency; Controlled Oral Word Association	Verbal fluency
Token Test	Comprehension
Boston Diagnostic Aphasia Examination (BDAE)	Overall language function
Aphasia Screening Test	Language screening
Visuospatial	
Block Design (WAIS-R)	Visuoconstructional ability
Clock drawings (command and copy)	Graphomotor construction
Cross drawings (command and copy)	Graphomotor construction
Rey–Osterrieth Complex Figure (copy)	Graphomotor construction/planning
Hooper Visual Organization Test (HVOT)	Visuoperceptual integration
Object Assembly (WAIS-R)	Visuospatial integration
Line bisection, Visual cancellation tests	Hemispatial inattention
Ruff Figural Fluency Test	Nonverbal fluency
Attention/Concentration	
Digit Span forward (WAIS)	Simple attention
Digit Vigilance (number cancellation)	Sustained concentration and attention
Paced Auditory Serial Addition Test (PASAT)	Auditory concent/processing speed
Digit Symbol (WAIS) or Symbol Digit Modalities Test	Visual attention/psychomotor speed
Trail Making Test (Part A)	Visual tracking/psychomotor speed
Continuous Performance Test (various versions)	Sustained vigilance and reaction time
Visual Sustained Attention Test (VSAT)	Sustained concentration and attention
Learning and memory	
Wechsler Memory Scale-3 (WMS-III) and Wechsler Memory Scale-Revised (WMS-R)	Global memory function
California Verbal Learning Test (CVLT)	Verbal learning and memory
Logical Memory (WMS-III, WMS-R)	Verbal memory
Visual Reproduction (WMS-III, WMS-R)	Nonverbal memory
Rey–Osterrieth Complex Figure (recall)	Complex nonverbal memory
Hopkins Verbal Learning Test	Brief measure of verbal memory
Benton Visual Retention Test	Nonverbal memory
Warrington Recognition Memory Test	Verbal/nonverbal recog. memory
Rey Auditory Verbal Learning Test	Verbal learning and memory
Buschke Selective Reminding Test	Verbal learning and memory
Motor and sensory	
Finger Tapping Test	Fine motor speed
Hand Dynamometer	Grip strength
Grooved Pegboard	Fine motor dexterity
Luria three-step hand sequence	Coordinated/integrated sequencing
Sensory-Perceptual Examination (Reitan–Klove)	Sensory and perceptual skills

Table 3 Information to be collected from the history and behavioral observations.

History of symptoms
- Description of present illness and specific symptoms (e.g., Why is patient being evaluated?)
- Nature of onset (e.g., insidious vs. acute; are symptoms new or a recurrence?)
- Duration
- Course of symptoms
- Associated behavioral or personality/emotional changes

Past medical history
- Previous neurological disease
- Previous psychiatric disease (psychiatric treatment, known diagnoses, etc.)
- Significant head trauma (details: presence and length of LOC/coma; PTA—assess)
- Seizures (circumstances of onset, frequency, characterization of seizures)
- Other medical illnesses requiring hospitalization
- Toxic Exposure (may relate to work history)

Birth and developmental history
- Birth complications
- Developmental delays (motor, language, intellectual)

Social history
- Abnormalities in behavior past and present
- History of and current family and peer interactions

Educational history
- Highest grade attained/ degrees awarded
- Specific problems learning
- Problems in school (e.g., truancy, being expelled, etc.)
- Typical grades in school (e.g., below average, average, above average)
- Best and worst subjects in school
- Reasons for termination of education (if relevant; also, if not HS, did get GED?)

Vocational history
- Present occupational status (if retired, what was last position)
- Length of current job
- Nature of responsibilities
- Types of previous jobs
- Frequency of job changes
- Recent difficulties with work (self-report, work appraisals, raises, etc.)

Substance use history
- Duration, frequency, and amounts of typical alcohol intake
- Use of other drugs
- Use of prescription medications (list current medications)

Family history
- Neurologic disease in family (especially parents, siblings)
- Psychiatric disease in family
- Familial predilection for diseases (e.g., learning disabilities, dementia, epilepsy, etc.)
- Cause and age of death of parents and siblings

Psychiatric symptoms
- Unusual or bizarre behavior
- Attention/concentration problems—distractibility
- Problems with social judgment
- Depression
- Anxiety
- Paranoid ideation/suspiciousness
- Hallucinations
- Delusions
- Tangentiality, looseness of associations

General appearance
- Appropriateness of appearance for age
- Posture (examine for asymmetries)
- Facial expression
- Eye contact
- Personal cleanliness, grooming, dress

Mood
- Range and appropriateness of affect, responsiveness to conversation/interview
- Depression
- Apathy
- Lability

Table 3 (continued)

- Mania or hypomanic symptoms

Behavioral
- Cooperativeness with examiner
- Anxiety
- Suspiciousness
- Anger
- Insight
- Response to interview

Motor
- General activity level
- Abnormal posturing, movements, tremors, etc.
- Gait (broad-based, shuffling, unsteady, symmetry of movement)
- Handedness (and familial history thereof, particularly if left-handed)

Speech
- Rate
- Volume
- Modulation
- Expression
- Dysarthria

a later time, they fail, thereby producing what subjectively seems like a deficit in memory. In actuality, their problem may not be at the retrieval or recall stage, but rather, at earlier stages of information processing. Similarly, a patient with aphasia may present with word-finding difficulty that might also mimic a memory impairment if the individual were not thoroughly evaluated. Some patients with frontal lobe damage demonstrate inefficient encoding secondary to organizational deficits, when in fact their ability to recall learned material may be normal.

The following section discusses the major cognitive domains typically assessed by the comprehensive neuropsychological evaluation, including an overview of some common representative measures from each domain. Each section begins with a clinical overview of the area to be examined and includes a brief discussion of relevant, less structured or "bedside" clinical examination/observation procedures prior to a presentation of specific psychometric assessment tools.

4.11.4 PRINCIPAL COGNITIVE DOMAINS FOR ASSESSMENT

4.11.4.1 Global Cognitive/Intellectual Functioning

The concept of intelligence has a long history and has been the topic of much discussion and controversy since the early 1900s. Various definitions of intelligence have been posited (Sternberg, 1997), and it is clear that intelligence is a multidimensional construct that is made up of many types of abilities that are demonstrated by each of us to varying degrees. Despite ongoing debate regarding the use of IQ measures, the concept of global intellectual capacity has utility within the context of assessing cognitive functioning and adaptive ability as it relates to brain function.

From a clinical bedside perspective, the clinician must rely upon the patient's educational and vocational history, in addition to their overall presentation style, use of language (when possible), and performance during various cognitive tasks to estimate a gross level of intellectual functioning. Next, one must begin to evaluate whether there seems to be evidence of gross intellectual decline based on their current presentation. Ideally, premorbid IQ and other cognitive test scores would be available for direct comparison, but this is rarely the case. Thus, detailed examination of this issue requires formal psychometric assessment, and based on specific cognitive test results and patterns of test scores, indices of current intellectual functioning as well as estimates of premorbid intelligence can be derived (Franzen, Burgess, & Smith-Seemiller, 1998).

The most commonly used intellectual assessment tools for adults are the Wechsler Intelligence Scales. In addition to the information provided by the various component subtests of the Wechsler Scales, some assessment of an individual's global level of cognitive capacity is important in establishing an interpretive background for other specific test results. For example, knowing that a patient has a Full-Scale IQ (FSIQ) score in the borderline range can be very useful in accurately interpreting their low average to borderline performance on memory tests. Although some variability across cognitive skills is commonly seen within individuals, higher intellectual levels tend to be associated with more advanced skills in various domains.

Thus, the 75-year-old retiree with eight years of formal education who obtains an FSIQ of 120 would be expected to perform at a higher level on memory or attention/concentration measures than the elderly individual with a similar educational background, but an IQ in the low average range.

One of the limitations of intellectual assessment that is worth noting is the use of global IQ scores in depicting an individual's cognitive capabilities. An FSIQ of 100 can be arrived at in various ways, for example, by achieving an very high Verbal IQ and low Performance IQ score, or vice-versa. On the WAIS-R, note that each of the following sets of VIQ/PIQ scores result in an FSIQ of 100 for a 30-year-old: (i) VIQ = 121, PIQ = 79; (ii) VIQ = 80, PIQ = 138; (iii) VIQ = 100, PIQ = 101. Also, some subtest scaled scores may be very high and offset lower ones, resulting in an "average" overall IQ score. These examples represent significantly different cognitive processing styles and differential ability patterns, although these distinctions would be lost if there is an over-reliance upon global IQ scores.

From a neuropsychological perspective, intellectual tests such as the WAIS-R and WAIS-III can be quite useful in describing cognitive strengths and weaknesses and eliciting relevant neurobehavioral samples, even though these tests were not originally designed for assessment of cognitive dysfunction. However, many of the changes in the WAIS-R reflected in the WAIS-III, in fact, were influenced by neuropsychological principles, and the WAIS-III includes several new subtests that attempt to provide a more detailed exploration of cognitive abilities beyond "IQ" or "g" alone. Whereas intellectual tests tend to be less sensitive to cerebral dysfunction than many specialized neuropsychological measures of specific abilities, the relative sensitivity of some subtests to cerebral dysfunction have proven useful in detecting gross cognitive decline in some cases. For example, the Vocabulary subtest of the WAIS-R historically has been known as one of the tests that is least sensitive to acquired cerebral dysfunction. This subtest relies upon highly crystallized and overlearned knowledge, and the ability to define familiar words tends to be a function that is relatively less affected by acquired brain damage than many other tasks that require novel learning or reasoning. In contrast, the Digit Symbol subtest of the WAIS-R, which requires rapid psychomotor speed, new learning, and attention, is one of the more sensitive subtests to cerebral dysfunction. Knowledge regarding which subtests are more or less sensitive to brain injury, educational background, age, gender, and so on, is essential in making inferences about cerebral dysfunction (see Heaton, Ryan, Grant, & Matthews, 1996). Pattern analysis of the Wechsler scales can thus be quite useful in the assessment of brain damage, as the subtests are differentially susceptible to the effects of acquired brain injury and can be further impacted depending upon the nature of the disorder and/or location of primary cerebrocortical involvement. These tests are described now in detail.

(i) *Wechsler scales*. (a) Wechsler Adult Intelligence Scale-Revised. The Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) is the most popular measure of intellectual functioning in adults, and has enjoyed widespread use since its introduction in 1981. A plethora of studies using the WAIS-R or its components has been published in the neuropsychological literature, and a great deal has been learned about its utility as a neuropsychological tool. This is underscored by the development and publication of the WAIS-R as a neuropsychological instrument (WAIS-R-NI; Kaplan, Fein, Morris, & Delis, 1991). This modification of the WAIS-R reflects the application of specific neuropsychological principles and procedures to the administration of the WAIS-R. For example, in individuals with difficulty responding to open-ended questions or those having initiation difficulties, the Information subtest of the WAIS-R-NI includes a provision for multiple choice responding. Some of the timed subtests include extended time limits for those with motoric slowing to permit a more valid assessment of visuoconstructional skills apart from the effects of motor dysfunction.

In settings where an overall estimate of intellectual functioning is desired, various short forms of the WAIS-R have been developed. One very brief example is the combination of Vocabulary and Block Design in the estimation of FSIQ (Brooker & Cyr, 1986; Silverstein, 1982). Such brief forms and lengthier procedures (e.g., Adams, Smigielski, & Jenkins, 1984) have shown reasonable correlations with the complete WAIS-R, and these estimates may be adequate in some situations. For example, in a case of head injury wherein the focus is upon more subtle difficulties with attention/concentration and memory, less of an emphasis may be placed on global intellectual status, particularly when there are no questions of a decline in this area. In selecting the shortened version of any test such as the WAIS-R, clinicians must rely upon experience with a large number of cases and with appropriately supportive literature. Among the individual subtests that show the strongest correlations with overall FSIQ scores is the Vocabulary subtest ($r = 0.81$), followed by Information ($r = 0.81$). At the other end of the spectrum lies Digit Symbol, which, along with Object Assembly, shows a correlation of 0.57 with FSIQ. In general, the five Performance subtests, which rely more upon novel thinking and "fluid" cognitive skills, tend to be more sensitive to cerebral dysfunction than the more crystallized abilities tapped by the six Verbal subtests. Because the various subtests of the WAIS-R assess a variety of cognitive abilities, some of these will be discussed in more detail under specific cognitive domains.

(b) Wechsler Adult Intelligence Scale-III. The WAIS-III is a newly revised version of the Wechsler Adult Intelligence Scale (WAIS-III; Wechsler, 1997a). It contains a variety of changes and additional subtests, and some previously familiar measures have been dropped or are now optional. The addition of a matrix reasoning task and measures of sustained concentration and working memory promise to provide a more well-rounded evaluation of global cognitive functioning and may enhance the test's utility in the context of the neuropsychological evaluation. Most of the familiar subtests remain, although these have been updated for the 1990s, with improved norms. Many subtests that include scenes for visual material now are presented in color, with larger figures and a more representative depiction of our multicultural society. As a quick review of the test evaluators and consultants will indicate, there is a decidedly more neuropsychological flair to the WAIS-III. As we have learned more about brain–behavior relationships and the effects of acquired cerebral dysfunction on intellectual functioning, test developers felt it important to incorporate aspects of this knowledge into the current test design. Table 4 lists the subtests for the WAIS-III.

Table 4 WAIS-III subtests.

Vocabulary	Picture Completion
Similarities	Digit Symbol–Coding
Arithmetic	Block Design
Digit Span	Matrix Reasoning
Information	Picture Arrangement
Comprehension	Symbol Search
Letter–Number Sequencing	Object Assembly

As reported in the WAIS-III technical manual (Wechsler, 1997a), correlations involving VIQ, PIQ, and FSIQ scores between the WAIS-III and WAIS-R are high (0.94, 0.86, and 0.93, respectively). WAIS-III VIQ, PIQ, and FSIQ scores were found to be an average of 1.2, 4.8, and 2.9 points lower than the corresponding values on the WAIS-R. Obviously, given the recent release of this test, many studies of its neuropsychological utility and psychometric relationships to other neuropsychological measures will be forthcoming.

(c) The Wechsler scales and patterns in neurological populations: VIQ–PIQ Differences. Despite the utility of instruments like the Wechsler scales in neuropsychological evaluations, appropriate interpretation of scores and patterns of scores in cases of brain dysfunction requires extensive training in neuropsychology and familiarity with a broad range of neuropathological, neuropsychiatric, and normal conditions. For example, deficits on PIQ subtests are often discussed in association with right hemisphere dysfunction. Whereas this may hold true in many cases, decreased nonverbal test scores may occur for a variety of reasons (e.g., psychomotor slowing on timed tasks, greater sensitivity to acquired brain injury than most of the VIQ tasks). The following cases serve to illustrate this point.

First it should be noted that all three cases depict at least a 22-point difference between VIQ and PIQ scores, although there is obviously some variability across the profiles in terms of specific scaled scores. One initial impression might be that all three cases likely represent some sort of acquired brain injury, as 20-point differences between IQ indices are rare in the general population. A second level of interpretation, based in part on the clinical lore associated with the test, might include an inference about lateralization, that is, all three cases might be suggestive of greater impairment of right hemisphere functions, given the large VIQ–PIQ differences and relative deficits on the "nonverbal" subtests. As alluded to earlier, knowledge that a variety of factors may result in decreased PIQ scores can provide the clinician with information that is essential in neurodiagnostic situations.

Case 1 is an engineer in his 40s who sustained a right hemisphere stroke. CT findings (Figure 3) depict a hemorrhage deep in the right hemisphere that included intraventricular bleeding.

His WAIS-R profile of decreased nonverbal abilities and a 22-point VIQ–PIQ discrepancy is what is often associated with a "classic" pattern of right hemisphere dysfunction (see Figure 4). Whereas this holds true for this particular individual, Cases 2 and 3 illustrate that simplistic interpretations or overgeneralizations can lead to erroneous conclusions.

Case 2 is a 50-year-old nurse with a history of stroke four months prior to evaluation. MRI results (Figure 5) revealed a chronic egg-sized hemorrhage involving the left internal capsule and basal ganglia.

Her primary presenting symptoms included mild right-sided hemiparesis and mild difficulties with verbal fluency and word finding. Much of the decrease in PIQ subtest scores (Figure 6)

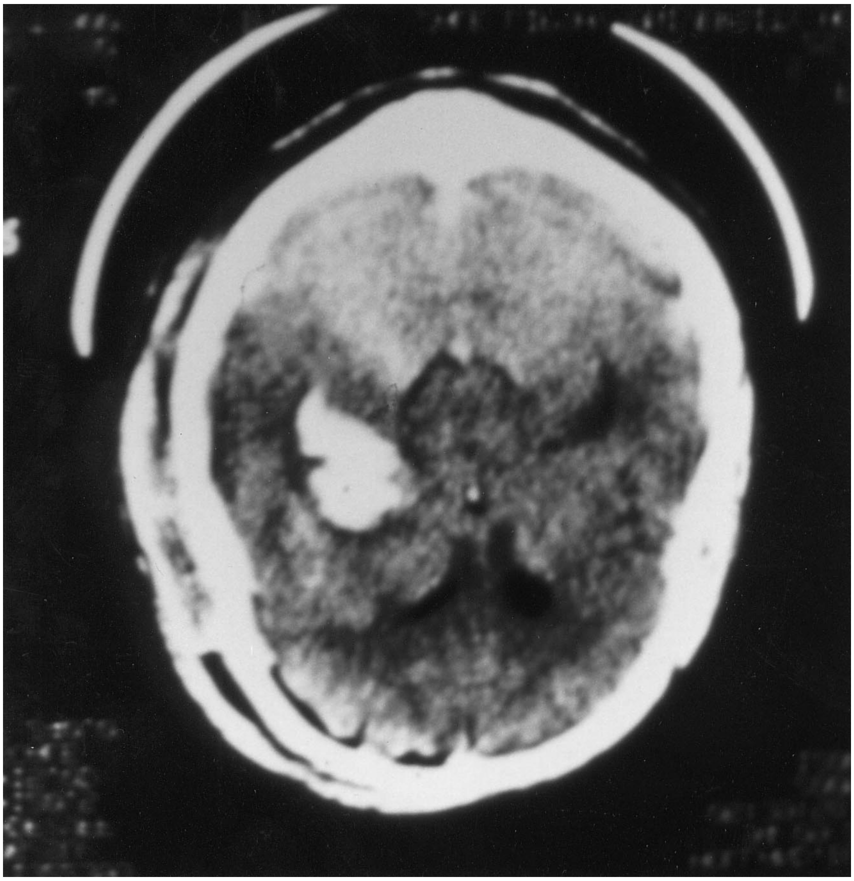

Figure 3 CT from Case #1 showing right hemisphere infarct.

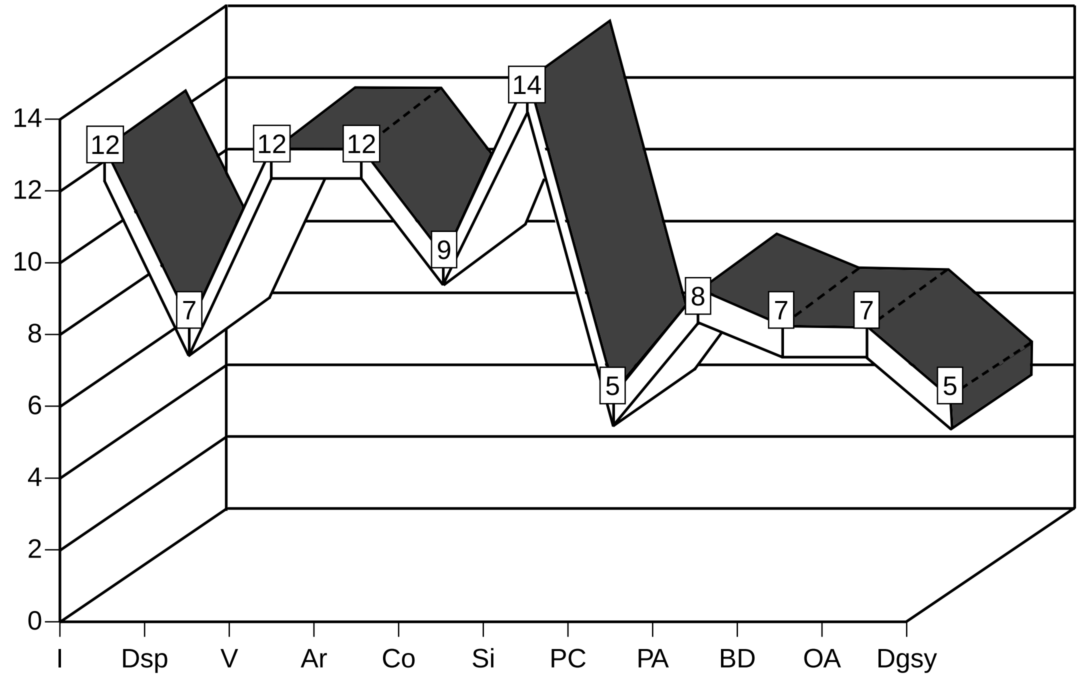

Figure 4 WAIS-R scaled scores from case #1. VIQ = 108, PIQ = 86, FSIQ = 98.

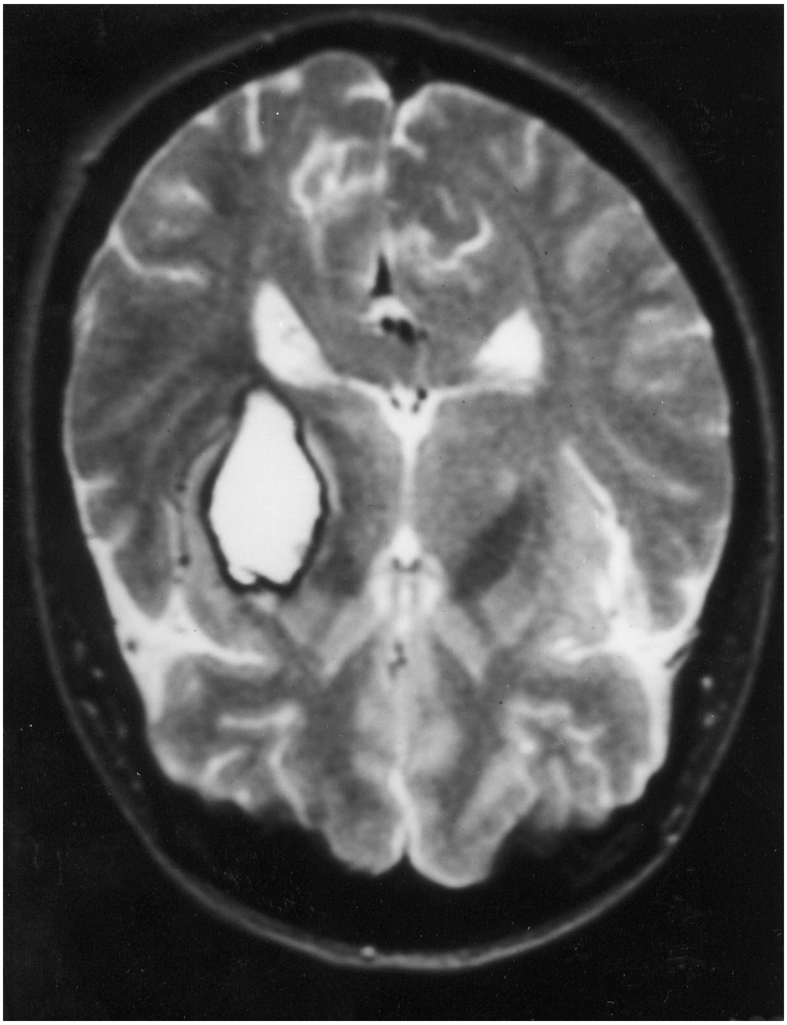

Figure 5 MRI showing left basal ganglia/internal capsule hemorrhage in case #2 (note that left is depicted on the right).

was felt to be related to the patient's hemiparesis and having to perform some tasks with her nondominant left hand. However, the relative sensitivity of nonverbal vs. verbal WAIS-R subtests to any type of acquired brain damage should be kept in mind, and in fact, additional time allowances on some of the PIQ subtests (e.g., Block Design) still resulted in evidence of cognitive impairment; thus, psychomotor factors alone could not completely explain the patient's difficulties in performing these tasks.

Case 3 results are from a 40-year-old long-term alcoholic who showed evidence of alcohol-related dementia. Axial MRI results (Figure 7) depict generalized sulcal widening and ventricular dilation.

The patient demonstrated generalized neuropsychological deficits, with particular difficulties in visuospatial abilities (see Figure 8). Results such as this are not uncommon in cases of extensive and long-term (i.e., 10+ years) alcohol abuse, and this pattern of findings on the WAIS-R was amply supported by the patient's impaired performances on other visuoconstructional tasks.

(d) WAIS subtest "scatter." The issue of "scatter" among WAIS-R subtest scaled scores represents another area of clinical lore that merits comment within the context of assessing patients with known or suspected cerebral dysfunction. It has been observed by some that, in cases of brain damage, there tends to be a greater than average range of scores from highest to lowest across subtests. This represents a complex issue, since most individuals with or without brain damage have relative cognitive strengths and weaknesses, and as noted, the subtests differ with respect to their individual sensitivity to acquired cerebral dysfunction.

Finally, what constitutes clinically "significant" scatter across Wechsler subtest scores also

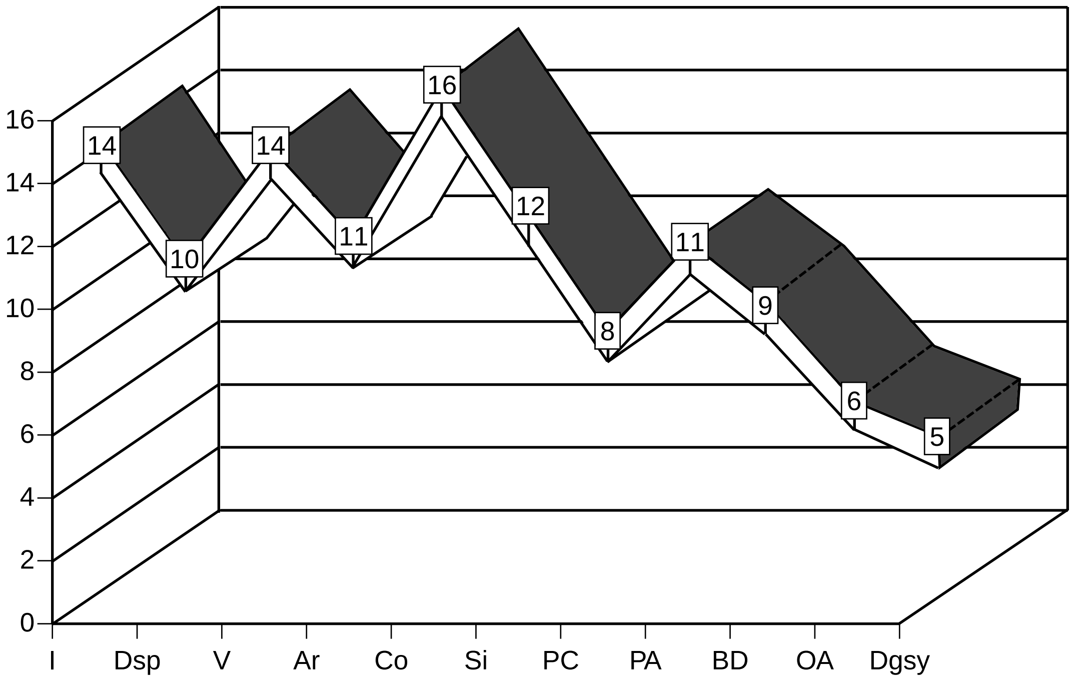

Figure 6 WAIS-R scaled scores from case #2. VIQ = 121, PQ = 101, FSIQ = 112.

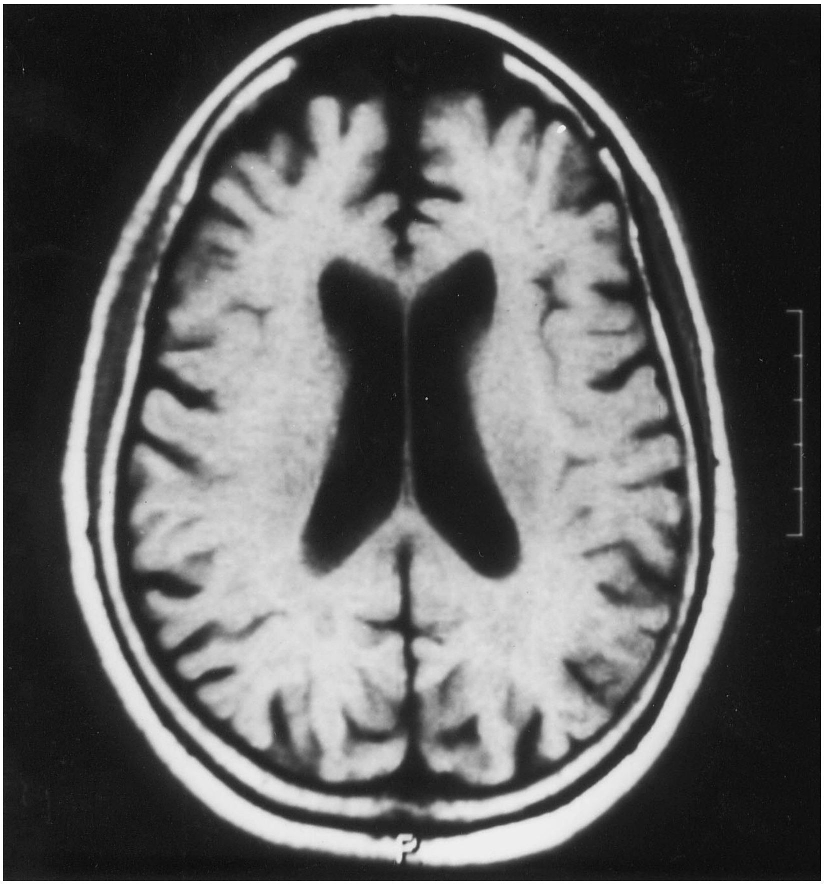

Figure 7 MRI results from case #3 showing generalized atrophy.

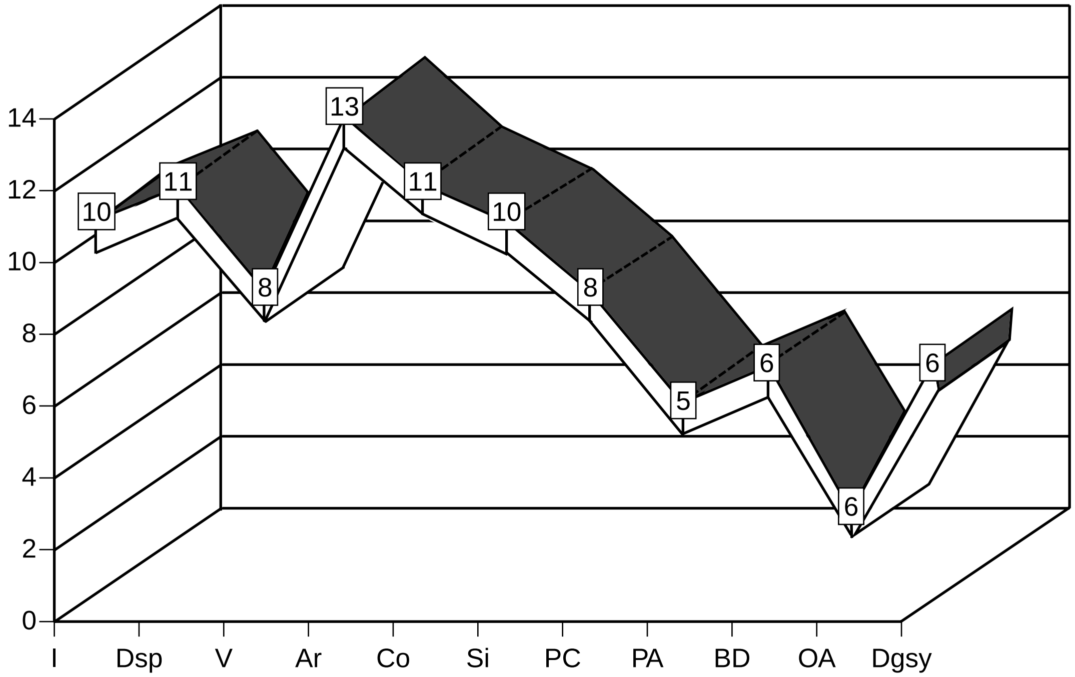

Figure 8 WAIS-R scaled scores from case #3. VIQ = 103, PQ = 78, FSIQ = 90.

merits careful review. In his APA Presidential address, Matarazzo (1990) provides data that clearly calls to question the notion of the relevance of "scatter." For example, a table is presented (p. 1010) that was derived from previous work (Matarazzo & Prifitera, 1989) which includes the percentage of cases from the WAIS-R standardization sample that demonstrate various levels of scatter (the difference between highest and lowest of the 11 subtest scaled scores). While extreme score differences of greater than 10 points were rare (i.e., seen in less than 5% of the 1880 standardization sample subjects), discrepancies of 10 points were seen in approximately 9% of the sample, while differences of seven points occurred in almost half of the sample, and roughly 86% showed a difference of five points. Thus, it should be kept in mind that "scatter" among WAIS subtest scaled scores is commonly seen in non-neurological populations. With these numbers in mind, the issue of the neurological significance of scatter among Wechsler scaled scores should be considered cautiously.

(ii) *Ravens Progressive Matrices*. This series of tests includes the Colored (Raven, 1995), Standard (Raven, 1996), and Advanced (Raven, 1994) Progressive Matrices, providing different levels of assessment of nonverbal reasoning. In the Colored Matrices, subjects are shown a stimulus array at the top of a booklet, with one small piece cut away. Below this, six small pieces corresponding to the outline of the missing piece are presented; however, each typically has a different pattern depicted. The task is to select the piece that best completes the stimulus array. The Colored matrices are particularly well suited for children and impaired elderly, and the Advanced matrices are most appropriate for high functioning individuals when an assessment of nonverbal reasoning is desired. IQ equivalent scores can be derived from the matrices, and their nonverbal nature (particularly the Colored version) allows for those with language impairments or limited fluency in English to apprehend the nature of the task quite readily in many cases.

(iii) *National Adult Reading Test-Revised (NART-R)*. This is a measure of estimated premorbid intellectual functioning (Blair & Spreen, 1989) as derived from the original NART (Nelson, 1982; Nelson & O'Connell, 1978). It comprises 60 printed words that are presented to subjects to read aloud. Thus, rather than relying upon any sort of integrative or associative abilities, the NART-R and its various derivations rely only upon sight-word reading. Most of the words included in the NART- R are phonetically irregular, and thus, correct pronunciation generally depends upon prior familiarity with the words. The stimuli are presented in order of increasing difficulty, and range from words like "asterisk" to "ennui." As with any performance-based cognitive task, the NART-R does show some sensitivity to acquired brain dysfunction; nevertheless, it tends to remain relatively stable, even in the face of progressive dementia (Matt-Maddrey, Cullum, Weiner, & Filley, 1996). The NART-R does have limitations in terms of its applicability to individuals with very high and very low intellectual status, but correlates well with Verbal and Full-Scale IQ in many cases when there is little or no intellectual decline.

(iv) *Dementia Rating Scale (DRS)*. This is a popular instrument for the assessment of global cognitive functioning in patients with known or suspected dementia (Mattis, 1988). It has been widely used clinically as well as for research purposes, and has proven to be a reliable and valid screening tool that typically requires only 20–30 minutes (up to 45 minutes in some demented individuals). The total number of possible points on the DRS is 144, and Table 5 provides a general description of level of impairment on this measure. It should be noted, obviously, that any such listing provides only a global guideline for interpretation, as results on the DRS do vary according to age, education, and other background factors.

The DRS is made up of the following subtests: attention, initiation/perseveration, conceptualization, construction, and memory. Because of the screening nature of the test, higher functioning or minimally impaired individuals may achieve normal scores yet demonstrate evidence of impairment on more sensitive tests. For example, the memory component of the DRS assesses general orientation and includes some simplistic recognition memory tasks, while only two tasks focus on the assessment of delayed recall.

Although designed as a tool to screen for the presence or absence of dementia, patterns on the DRS have also been shown to be useful in helping to distinguish some types of dementias, although it should be noted that once cognitive impairment becomes severe, such differentiation on the basis of cognitive test results can be difficult or impossible. Despite is relative brevity, total scores from the DRS are quite useful in documenting cognitive decline over time and have shown to be predictive of nursing home placement (Smith et al., 1994).

(v) *Mini-Mental State Examination*. This is a very popular screening measure of global cognitive status. It is particularly useful in cases of dementia and may be used even in severely impaired individuals (Folstein et al., 1975). It is very brief (i.e., 5–10 minutes) and comprises 30 total points. Test items include an assessment of orientation to time and place, attention, language, constructional skill, and three word repetition and recall. It is recommended that standardized recall items (e.g., apple, table, penny or rose, ball, key) are used, as opposed to the original instructions of "name three objects" for patients to recall. Total MMSE scores correlate best with orientation and recall items, and significant variability can be seen among healthy subjects on three word recall (Cullum, Smernoff, & Thompson, 1993). Despite the limitations inherent in any brief measure of overall cognition, the MMSE has proven to be a valid and reliable screening tool, with utility in a number of populations and settings. Although the traditional cutoff for impairment is $<24/30$, the MMSE is sensitive to age and education effects, and appropriate norms should be used (Crum et al., 1993; Tombaugh, McDowell, Krisjansson, & Hubley, 1996).

Table 5 Dementia Rating Scale (DRS) and level of dementia.

115–130	Mild
100–114	Moderate
90–99	Moderate to severe
<90	Severe

4.11.4.2 Academic Achievement

These tests sample a range of abilities related to academic or scholastic levels of attainment. A number of omnibus achievement batteries exist and each has its relative strengths and limitations depending upon the purpose of assessment. Within the context of the adult neuropsychological evaluation, various components of different achievement tests are often administered. The more common measures sample a variety of abilities related to academic achievement (e.g., spelling, sight reading, reading comprehension, math) which can be compared with other neuropsychological skills and performances on related tests. Assistance in the diagnosis of learning disabilities can also be provided by these measures to the extent that specific deficiencies are identified.

(i) *Wide-Range Achievement Test-3rd Edition (WRAT-3)*. The WRAT-3, like its predecessors (WRAT, WRAT-R), provides for an assessment of written spelling, written mathematical, and sight-word reading skills (Wilkinson, 1993). It enjoys widespread use across the lifespan (age 5–75) and is relatively brief (i.e., approximately 20–30 minutes).

(ii) *Woodcock–Johnson Psychoeducational Battery-Revised (WJ-R)*. This is one of the most comprehensive batteries when used in its entirety for the individual assessment of global cognitive skill, general information, academic achievement, and interests (Woodcock & Mather, 1989). It has been used with children to adults, ranging in age from two to 90 years. The WJ-R Achievement Supplemental Battery provides a more focused assessment of academic achievement through the use of subtests that assess reading (both nonsense words and for meaning), math concepts, proof-reading, writing to dictation, and writing/sentence composition. Various cluster scores pertaining to different aspects of achievement and level of competence across domains can be examined in order to establish a profile of abilities.

(iii) *Peabody Individual Achievement Test-Revised (PIAT-R)*. The PIAT-R is used primarily for children, but offers norms from kindergarten through high school (up to age 19). It comprises the following subtests: general information, reading recognition, reading comprhension, mathematics, spelling, and written expression (Markwardt, 1989). Various subtests may be administered in order to screen for specific areas of difficulty, and because it is infrequently assessed by many other tests of cognition, the reading comprehension subtest is perhaps one of the most commonly used subtests of the PIAT-R as part of the adult neuropsychological evaluation.

4.11.4.3 Executive Functioning, Problem-solving, and Reasoning

This group of abilities is referred to by various names, and encompasses a variety of skills and functions. Judgment, reasoning, problem-solving, and abstraction are some of the more common skills within this domain that involve some of the most complex and integrative functions of the human brain. While such abilities are not specifically localized and rely upon the integration of various cortical and subcortical regions and pathways, the frontal lobes tend to play a relatively prominent role in many of these skills.

Formal neuropsychological evaluation represents the most efficient means of assessing these functions, although more informal administration of similarities, proverb interpretation, inhibition, and reasoning and judgment tasks also can be useful. Fluency tasks have also been shown to be useful in assessing executive or anterior cerebral functions—letter or phonemic fluency ("tell me as many words as you can think of that start with "R' within a minute") to assess left anterior hemispheric functions, and design fluency for anterior right hemispheric function (e.g., having patients draw as many unique nonreplicated designs as possible within a minute). Such tasks may be modified to further tax executive systems, for example, by alternating between different procedures ("name an animal, then a city"), although normative data on such procedures are lacking.

Judgment or reasoning tasks include questions such as "What would you do if you were stranded at the airport with only one dollar,?" or "Why should people visit the doctor when they are not ill?," or "The tiger was eaten by the Lion—who did the eating?" As with some other tests of higher cognitive function, proverb interpretation involves a variety of skills that cannot be readily distilled into a specific set of abilities. Administration of familiar (e.g., "A rolling stone gathers no moss") as well as unfamiliar proverbs (e.g., "The used key is always brightest") is recommended, however, since an individual's ability to interpret novel material may rely upon different skills compared to their ability to retrieve the common response to a more well-rehearsed task. Information about patients' planning abilities is also useful, whether assessed by a formal cognitive task or simply by inquiring about their plans for the future given their current situation. Malloy and Richardson (1998) outline a number of bedside and formalized procedures for the assessment of frontal/executive functioning. In evaluating results from various tests of executive abilities or frontal lobe function, it must be kept firmly in mind that such measures are not uniformly or invariably affected by damage to frontal systems (Bigler, 1988; Stuss & Benson, 1984), and that normal test results can be seen in some patients with clear evidence of frontal damage, just as impairment on the same tests can occur in patients with pathology outside the frontal lobes.

4.11.4.3.1 Tests of reasoning and problem-solving

(i) Wisconsin Card Sorting Test (WCST)

This is a popular measure of nonverbal reasoning and cognitive flexibility (Heaton, 1981; Heaton, Chelune, Talley, Kay, & Curtis, 1993). It requires subjects to sort a series of colored geometric designs into piles according to specific principles. Feedback is given as to whether each individual response is correct or incorrect, although subjects are not instructed as to what strategies they should be using. The WCST is sometimes touted as a measure of "frontal" functioning. Whereas the test has been shown to be relatively more sensitive to anterior brain damage in some studies (Heaton, 1981), it is important to keep in mind that the WCST is not exclusively a measure of frontal functioning, and can be sensitive to damage in other areas of the brain. Thus, impairment on the WCST in and of itself cannot necessarily be interpreted to suggest disruption of frontal brain systems. Nevertheless, patients who are disinhibited and have difficulty monitoring their responses will tend to be perseverative in their responses to this task. Furthermore, given the appropriate supporting pattern of anterior dysfunction based on behavior and results from other measures, results and observations from the WCST may assist in determining whether or not there is evidence of frontal brain system involvement. The most sensitive score from the WCST has proven to be the number of perseverative responses (Heaton, 1981), as many brain-injured individuals will tend to repeat the same strategy, despite corrective feedback. The ability to shift response strategies and perform such trial-and-error learning may also be a good predictor of various aspects of everyday functioning.

(ii) Category Test

The Category Test (Reitan & Wolfson, 1993) is one of the more sensitive measures of cerebral dysfunction, probably because of its complexity. This task requires subjects to generate and implement various problem-solving strategies across its 208 items. It comprises seven subtests, most of which require different principles to arrive at correct solutions. The subtests are generally arranged in increasing levels of difficulty, although two of the subtests actually require the same solution strategy, and the final subtest relies more upon memory, insofar as subjects must recall which response was correct when they saw that item in a previous subtest. The Category Test tends to be sensitive to dysfunction regardless of primary site of damage, although executive function impairment tends to have a particularly disruptive effect on results. Several versions of the Category Test are available (see Spreen & Strauss, 1998), including the standard machine version, a booklet version, a pamphlet version, and several computerized adaptations of the original test. Various short forms and child versions are also available.

(iii) Delis–Kaplan Executive Function Scale (DKEFS)

This is a battery of executive function tasks based on the Boston process approach that is being published by The Psychological Corporation (Delis, Kaplan, & Kramer, in press). Among other measures, it includes the California Sorting Test (CST), a measure designed to delineate more qualitative and multidimensional aspects of problem-solving, concept generation, and mental flexibility than other sorting tasks. The test consists of a series of cards that differ in shape and color and have different words printed on them. The cards can be sorted according to any of eight different principles. Initially, subjects are asked to sort the cards into two equal groups, without any further instruction. This procedure is repeated until no new groupings are performed. After each sort, subjects are asked to indicate what rule or principle they used. Subsequent trials include stating the sorting principle used by the examiner, and sorting in different ways according to cues. The test provides for an assessment of verbal and nonverbal concept formation and perseveration and allows for information to be derived regarding which component process may be principally affected. Preliminary studies of the CST have been promising in terms of the test's clinical utility (e.g., Beatty, 1993; Delis, Squire, Bihrle, & Massman, 1992; Greve, Farrwell, Besson, & Crouch, 1995). Also, because of the flexibility of the procedure, it is appropriate for a variety of ages, and normative data on individuals aged eight to 89 will be available.

4.11.4.4 Arousal and Orientation

Evaluation of these basic functions is a necessary prelude to the examination of higher level cognitive skills. These abilities represent some of the basic building blocks upon which higher abilities rely. For example, memory cannot usually be adequately assessed in a patient who is unable to attend to stimuli. Arousal refers to basic level of consciousness as a function of the reticular activating system. Arousal is often described as ranging from alert, drowsy, lethargic, stuporous, to comatose. Alternatively, hyperarousal can also be seen in terms of restlessness, agitation, and delirium. The delirious patient is confused and transiently disoriented. Patients may seem to come "in and out of it" from day to day or hour to hour, and mental status examination results may fluctuate accordingly. Assessing level of arousal is a clinical judgment, and often, if there are difficulties in this area, a behavioral description of what it takes to arouse a patient can be most important (particularly since a more detailed examination of cognitive abilities may not be possible).

Orientation is typically thought of in terms of the three general spheres of time, place, and person. Orientation to time is the most readily disrupted of these and can be assessed by asking patients questions about the current time, day, year, month, and date. Assessment of place is obvious (i.e. "Where are we now?"), although questioning about situational variables ("Why are you here?") is also useful along these lines. Orientation to person is the most resistant to disrupted cerebral functioning relative to time and place, and refers to a patient's ability to provide personal identifying information as well as being able to identify others who are familiar to him or her.

4.11.4.5 Attention/Concentration

The concept of attention is complex (e.g., see Mirsky, Anthony, Duncan, Ahearn, & Kellam, 1991). As some of the more basic cognitive abilities, attention and concentration are required for adequate performance on essentially all other cognitive tasks. Thus, it should be kept in mind that if a severe attentional impairment is present, associated difficulties should be manifest on tests of other cognitive abilities as well. General or undirected attention is related to overall arousal level and one's awareness of the environment. Additional divisions of attention into immediate vs. sustained attention also has clinical merit, as some patients can perform simple tasks of immediate attention, yet may have difficulty when asked to maintain their concentration longer. Measures of sustained attention, often referred to as "concentration" tasks, include measures designed to focus an individual's attentional skills for a more extended period of time, for example, over the course of a minute or longer. Informal assessment procedures may include having the examiner say a long series of random letters or numbers and having the patient indicate whenever they hear a designated target number or letter (e.g., the "A" test of attention; Strub & Black, 1988, p. 58).

Although it reflects a "global" brain activity, attention has some focal and lateralized correlates. Subcortical systems such as the reticular activating system in the core of the brainstem and midbrain, as well as diencephalic structures (e.g., thalamus) play a central role in the brain's ability to maintain arousal and focus attention. In terms of laterality issues in the visual domain, the left hemisphere attends to right hemispace and the right hemisphere attends to left hemispace. Hemi-attentional disorders, less severe than hemispatial neglect, may require double simultaneous stimulation to elicit, although evidence of hemi-inattention can also be detected in many cases by using line bisection tasks, number or symbol cancellation tasks, or even simple drawings to command or copy. As an extreme example, the patient with hemineglect may ignore one-half of their visual world, denying that their arm or leg belongs to them upon questioning, or eating only the food on one side of their plate. Hemi-inattention and neglect phenomena most often occur on the left side of hemispace, and the right hemisphere has been implicated as perhaps playing a predominant role in some aspects of attention. One simple test of neglect involves line bisection, wherein the examiner draws a series of horizontal lines of different length randomly on a page. The task for the patient is to bisect each of these with a small mark. Patients with neglect will show a tendency to skew their marks to one side (i.e., away from the neglected side) of each stimulus. A wide array of standardized clinical and experimental assessment techniques is available to measure aspects of attention and concentration, and only a few of these will be discussed.

4.11.4.5.1 Assessment of attention

(i) Digit Span (WAIS-III, WMS-III, WAIS-R, WMS-R)

The familiar Digit Span task requires subjects to repeat series of digits of increasing length. Digit span forward is a good measure of simple attention, and most healthy individuals perform within the seven plus/minus two span of apprehension range. While some consider the digit span test to be representative of a memory task, the demands on memory *per se* are minimal, and digit span forward is best considered as a measure of attention. To illustrate this principle, amnestic patients with Alzheimer's disease and Korsakoff's syndrome often demonstrate a normal digit span forward despite severe anterograde amnesia. Digit Span backwards represents a qualitatively different type of task that relies more upon working memory skills that should be considered separately from digits forward (Reynolds, 1997).

(ii) Digit Symbol (WAIS-III, WAIS-R)

This is a complex measure of attention, working memory, visual scanning, and psychomotor speed. Subjects are required to learn to associate numbers with symbols and must fill in the symbol that is paired with each of nine numbers under time constraints (90 seconds). As noted, it tends to be the most sensitive of the WAIS-R subtests to brain dysfunction, and is often used separately from the Wechsler scales.

(iii) Symbol Digit Modalities Test

This is very similar to the Wechsler Digit Symbol test, in that it requires the written substitution of numbers and symbols. In this test, however, subjects must fill in the specified number that goes with the indicated symbol (Smith, 1982). An oral version is also available wherein the patient responds verbally with the number associated with each symbol they come to. Ninety seconds are allowed in this test as well, and it has been used in a large number of neurobehavioral studies since its inception. In some research settings to assess medication effects, subjects complete the test repeatedly prior to the beginning of the study in order to decrease subsequent practice effects.

(iv) Digit Vigilance Test

This represents a popular and well-normed (Heaton, Grant, & Matthews, 1991) version of the traditional number cancellation task. In this test, subjects must cross out all of a specified number they come to on pages that are filled with rows and columns of single digits (Lewis & Kupke, 1977; Lewis & Rennick, 1979). The task requires visual scanning, psychomotor speed, and sustained concentration, insofar as it typically takes several minutes to complete the two-page version of this task. Additional information regarding lateralized inattention can also be derived from this and similar tasks.

(v) Visual Search and Attention Test (VSAT)

This provides similar information to Digit Vigilance regarding sustained visual concentration, but includes verbal and nonverbal stimuli (Trenerry, Crosson, DeBose, & Leber, 1990). The VSAT consists of four tasks (two for practice) that require subjects to cross out target items from 10 rows of 40 stimuli (letters or numbers) on each version. Scoring for left- vs. right-sided errors in addition to an overall attention index is included.

(vi) Bells test

This is a symbol cancellation task similar to Digit Vigilance, although the target in this case is a small picture of a bell that is repeated within a quasirandom array of 315 small visual figures on a page (Gauthier, Dehaut, & Joanette, 1989). One of the advantages of this task can be that patients who utilize more systematized visual search strategies (e.g., scanning left to right and from row to row) on more symmetrically organized tasks such as Digit Vigilance may show impairment more readily on the more randomly arranged bells.

(vii) Continuous Performance Test (CPT)

Several versions of this computerized test of sustained concentration have been developed, and variations on the basic task have been used in research and clinical settings since the 1950s. In one common format, the CPT presents several series of single visual stimuli (usually letters) projected at brief intervals. Subjects must respond when a letter was preceded by a designated target letter. In Connor's CPT (Connors, 1995), subjects are asked to press the appropriate key for any letter presented, except for the letter S. Multiple trials of different interstimulus intervals are presented, and scoring includes reaction times, number correct, and errors of omission and comission.

4.11.4.6 Language

Language is a central part of the human experience, and a careful evaluation of the core components of language function are an important part of the neurobehavioral examination. Because of its complexity and multifaceted nature, an assessment of "language" requires multiple tasks. Note that some components of language (e.g., basic comprehension and expression) are amenable to observation during even casual interactions with patients, as well as during most of the traditional mental status examination. Recall that the left hemisphere is dominant for language functions in most people (particularly in right-handers, wherein over 95% are left-hemisphere dominant for language), although a minority of individuals (particularly left-handers) have bilateral speech representation or may even have their primary language center in the right hemisphere.

In screening for language disturbance, several primary areas should be assessed: verbal expression, naming, repetition, comprehension (oral and written), reading, and writing. Depending upon the function disrupted, performances on tasks involving these skills may provide useful information in terms of localization of dysfunction. Anterior dominant hemisphere systems are involved in expressive language (e.g., Broca's area), while more posterior regions are involved in receptive language (e.g., Wernicke's area). Aphasic disorders can occur as a result of a disruption of either of these areas, interconnecting cortical and subcortical pathways, or in cortical association areas nearby. The major aphasic syndromes are listed in Table 6.

Several comprehensive batteries for assessing aphasia have been developed, including the Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1987), the Multilingual Aphasia Examination (MAE; Benton & Hamsher, 1989; Benton, Hamsher, Rey, & Sivan, 1994), and the Western Aphasia Battery (Kertesz, 1982). Routine neuropsychological evaluations vary widely in terms of the depth with which language is examined. At a basic level, to begin to assess language function, the following areas and procedures should be considered.

Table 6 Common aphasic syndromes and abilities affected.

Aphasia Type	*Fluency*	*Naming*	*Repetition*	*Comp.*	*Rdg. comp.*
Broca's	Poor	Poor	Poor	Good	Good
Wernicke's	Good	Poor	Poor	Poor	Poor
Conduction	Good	Poor	Poor	Good	Good
Global	Poor	Poor	Poor	Poor	Poor
Transcortical motor	Poor	Poor	Good	Good	Good
Transcortical sensory	Fluent	Poor	Good	Poor	Poor
Mixed transcortical	Poor	Poor	Good	Poor	Poor
Anomic	Good	Poor	Good	Good	Poor

(i) Expressive language/fluency. Spontaneous speech can be examined during clinical interview and history-taking. This includes an assessment of rate of speech, prosody, grammatical complexity, and use of vocabulary.

(ii) Confrontation naming refers to the ability to name objects or representations of objects. This can be assessed in a rudimentary way by having the patient name things in the room or things you are wearing or holding. A much more sensitive means of assessing naming ability is to show the patient pictures of objects and request that they say what it is. Clinical experience suggests that the further the representation is from real life (e.g., a line drawing vs. a photo or real object), the more sensitive the task will be to dysnomia. An important note is that some degree of dysnomia is seen in most types of aphasia.

(iii) Repetition is examined by having the patient repeat single words (e.g., "paper," "ball," "comb") and phrases of increasing length and complexity (e.g., "The boys played baseball," "The vat leaks," "He is the one who did it"). Careful notation of the specific errors made (e.g., omissions, paraphasias) should be made, as these may merit further exploration using more standardized tasks.

(iv) Comprehension is also assessed in interview and during the performance of tasks upon request by the examiner (i.e., does the patient seem to understand what is being asked, or do tasks require additional explanation?). Examples of comprehension tasks include: "Point to the ceiling, then to the floor" "Do two pounds of flour weigh more than one?" "Place the quarter to the right of the penny and pick up the dime."

(v) Reading is examined readily by having the patient read single words, something they have written, and/or read sentences written by the examiner. Reading comprehension can also be tested by having the patient read a passage from a newspaper or magazine and then asking them questions about what they read (a number of standard brief tasks are available).

(vi) Writing is tested by having the patient write their name (a highly over learned task), specific words or a sentence to dictation, or perhaps a sentence or paragraph on a topic of interest (which also allows for an assessment of sentence structure, punctuation, and grammar use, etc).

The assessment of general expressive and receptive language skills encompasses a large array of tests and procedures, and is well beyond the scope of this section. Nevertheless, various language skills should be sampled in the comprehensive neuropsychological evaluation, at least at a cursory level. This may entail the administration of specific tests (e.g., the Token Test to evaluate comprehension), or may reflect more of a basic examination of an individual's ability to repeat, speak, write, and comprehend. Various types of linguistic errors can be seen during speech and these various tasks and should be noted when they occur. Paraphasias can be semantic or verbal (i.e., the substitution of an incorrect word), phonemic or literal (i.e., the incorrect substitution of phonemes within a word or a similar sounding word), or neologistic (i.e., made-up, nonsense words).

4.11.4.6.1 Global language assessment

(i) The Aphasia Screening Test

This test is a popular and extremely brief measure that employs single- or several-item sampling of a variety of language skills (e.g., reading, writing, naming, comprehension) (Reitan & Wolfson, 1986). For tests such as this, it must be kept in mind that a single error on a particular item, while potentially significant, may not reflect an underlying deficit. Instead, this may reflect the brief nature of the task and its sensitivity to response variability during a neuropsychological examination. Hence, individual errors on such tasks should be followed up with additional tests, lest an overinterpretation of results occur.

(ii) The Wechsler Vocabulary Subtest

This is a subtest of the Wechsler scales (WAIS-III, WAIS-R) and provides a good assessment of vocabulary and basic word knowledge. As previously noted, the Vocabulary subtest tends to be relatively insensitive to acquired cerebral damage in most cases, and may provide an index from the Wechsler scales regarding an individual's likely level of premorbid functioning. Unlike the NART, however, the Vocabulary subtest, with its reliance upon subject-generated word definitions, tends to be somewhat more susceptible to the effects of brain injury, and may also be impacted by conditions that affect initiation of responding. The WAIS-R-NI version includes provisions for multiple-choice responding for those patients who may have difficulty with verbal expression. The correlation between WAIS-R Vocabulary and FSIQ is 0.81 (0.85 with VIQ), and the WAIS-III Vocabulary correlation with FSIQ is 0.80 (0.83 with VIQ), also makes the Vocabulary subtest appealing when a brief estimate of current verbal intellectual functioning is desired.

(iii) Boston Naming Test (BNT)

This is a popular quantitative measure of confrontation naming ability (Kaplan, Goodglass, & Weintraub, 1983). Because of the prominence of dysnomia in the various aphasic syndromes, the assessment of word-finding has become a component of many neuropsychological test batteries. The standard BNT consists of 60 line drawings of objects of decreasing familiarity (e.g., ranging from "comb" to "abacus"). Several short forms of the BNT are also available (Franzen, Haut, Ranking, & Keefover, 1995; Lansing, Ivnik, Cullum, & Randolph, in press) for use when a more brief assessment of naming is desired. As noted, clinical experience indicates that the ability to name drawings of items is a much more sensitive means of assessing dysnomia than having patients name articles that are physically present in the room.

(iv) Verbal fluency (letter and category)

Clinical tests of verbal fluency come in two primary varieties: letter or phonemic fluency and category or semantic fluency (Benton & Hamsher, 1967; Borkowski, Benton, & Spreen, 1967). For letter fluency, subjects are asked to generate as many words as possible in 60 seconds that begin with a specified letter. Norms and alternate forms are available (Kozora & Cullum, 1994; Spreen & Strauss, 1998), although the most common sets of letters in use today include F, A, S, and C, F, L, with P, R, W serving as an alternate form. Category fluency is most commonly assessed by having subjects generate as many animals as they can think of in 60 seconds, with other variations of this task including fruits, vegetables, first names of people, supermarket items, cities, and US states.

Performance on letter (phonemic) and category (semantic) fluency tasks has been analyzed in number of populations, and it has been shown that the abilities involved in each task are dissociable. For example, in patients with Alzheimer's disease, it has been commonly reported that patients show a greater difficulty on category fluency, and often do somewhat better on letter fluency. This is thought to be related to the breakdown in semantic knowledge as seen in patients with Alzheimer's disease, unlike patients with subcortical dementias (e.g., see Chan et al., 1995). Along these lines, Moscovitch (1994) provides evidence to suggest that letter fluency is more of a frontally-mediated task, while category fluency relies more upon temporal lobe structures.

4.11.4.7 Visuospatial Skills

This cognitive domain refers to the ability to process visual or nonverbal information. In the most common visuospatial task, patients are asked to draw a specific shape (e.g., intersecting pentagons). More difficult tasks include drawings of three-dimensional figures (cube, house). Drawings to command can be compared with patients' copies of drawings to see if the presence of a model facilitates performance. Three-dimensional constructional tasks also exist, such as replicating designs using blocks. These and related tasks require various levels of visuoperceptual and visual-integrative skills, in addition to a motor output (e.g., graphomotor or manipulospatial).

Various standard tasks and stimuli are available for clinical use. Prominent visuospatial errors can be readily detected on such tasks, most often in association with right hemisphere damage (primarily parietal dysfunction). Lateralized errors in drawings, embellishment or paucity of creations, perseveration, rotations, intrusions, and loss of *gestalt* or detail can all be observed and may provide useful diagnostic and descriptive information. For example, patients with right hemisphere damage will tend to have more difficulty maintaining the overall *gestalt* of designs, whereas left hemisphere dysfunction more often results in problems recreating details of figures. Because many of these tasks require visual perception and a motor response, the clinician must use caution in interpreting visuoconstructional impairments when primary visual, perceptual, or motor disturbances are present.

4.11.4.7.1 Visuospatial tasks

A variety of drawings are commonly used in neuropsychological settings to assess graphomotor visuoconstructional ability. Some drawings have gained widespread popularity such as clocks, drawings of crosses, and three-dimensional cubes. A dissociation can be seen in some patients between their ability to spontaneously draw a specified design or item to command and their ability to reproduce it from a model presented to them. Obviously, the production of any drawing involves a variety of skills, and an interpretation of impaired visuoconstructional ability should only be done once a thorough assessment of the various component processes has been undertaken. For example, patients with basic motor or movement disorders will often produce impaired drawings, which should not be interpreted as a visuoconstructional deficit if their level of motor impairment can just as easily explain their poor productions.

The term "constructional dyspraxia" is often used to describe impaired visuoconstructional abilities. Unfortunately, the term is often misused and overused in clinical practice. Technically, the term refers to an impairment in the ability to construct visuospatial figures or drawings, whether two- or three-dimensional. Thus, a patient who demonstrates an impaired copy of a key from the aphasia screening test should not be considered to have constructional dyspraxia if they show normal performances on other visuoconstructional tasks such as WAIS-R Block Design, for example. Furthermore, it is important to obtain more than one drawing sample from a patient, as the degree of complexity of one stimulus may fail to elicit an impairment, whereas another one of greater complexity or less familiarity may reveal evidence of difficulty in this domain. The clinician's familiarity with developmental and general norms on a given drawing procedure are critical. This is particularly important on drawing tasks that do not lend themselves to highly detailed scoring procedures. Even when such scoring procedures are available, clinical interpretation based on a careful review of the component processes and the actual drawings produced can be critical. For example, the Rey–Osterrieth complex figure has seen the development of a variety of scoring systems over the years (see Lezak, 1995; Spreen & Strauss, 1998). Although more recent attempts such as the Boston qualitative scoring system (Stern et al., 1994) have attempted to capture more of the essential details and qualitative aspects of the design, quantitative scoring alone may prove inadequate in the absence of careful inspection by a seasoned clinician.

(i) Clock drawing test

This is a staple for many clinicians across disciplines involved in mental status examinations and neurobehavioral assessment, and actually can provide a wealth of information regarding a variety of cognitive skills. First, the reproduction of the face of a clock from memory requires some familiarity with such a stimulus. In some cultures, this represents a highly overlearned icon, and thus, should be reproduced even by children without significant difficulty, providing they are able to tell time.

Typical instructions (Goodglass & Kaplan, 1987) for the clock drawing test are to "Draw the face of a clock showing the numbers and the hands set to ten after eleven." The additional instruction to include all of the numbers on the clock can be helpful in some cases, particularly since some clocks and watches do not include a complete set of alphanumeric symbols. The instruction to set the hands at a specified time is also critical, and even when gross visuoconstructional abilities and the overall *gestalt* of the clock may be correct, time-setting errors may reflect problems with executive functions and/or memory. It is also important that the instructions include setting the hands to a time that is not too readily obvious or overlearned, so as to reduce the likelihood of an individual getting the time correct by chance. Various scoring procedures for clock drawings have been established (e.g., see Freedman, Leach, & Kaplan, 1994; Kozora & Cullum, 1994; Royall, Cordes, & Polk, 1998), although careful visual inspection and qualitative analysis, in addition to the application of a particular preferred scoring system, is important from a clinical interpretive perspective.

Along these lines, some of the more common errors that can be seen on the clock drawing test include confusion, gross distortion, and constructional dyspraxia (Figure 9(a)), aspects of unilateral neglect (Figure 9(b)), stimulus-boundedness and perseveration (Figure 9(c)), and errors in hand-setting (Figure 9(d)). Figure 9(d) was drawn by a patient with mild Alzheimer's disease and demonstrates a common conceptual stimulus-bound response by their setting of the hands at the 10 and 11 in order to depict "ten after eleven." Comparison of clock drawings done to command and under copy conditions have also demonstrated utility in some differential diagnostic situations (e.g., Libon, Malamut, Swenson, Sands, & Cloud, 1996).

(ii) Aphasia Screening Test

Although designed primarily as a brief screening measure of gross language disturbance, the aphasia screening test contains several drawings,

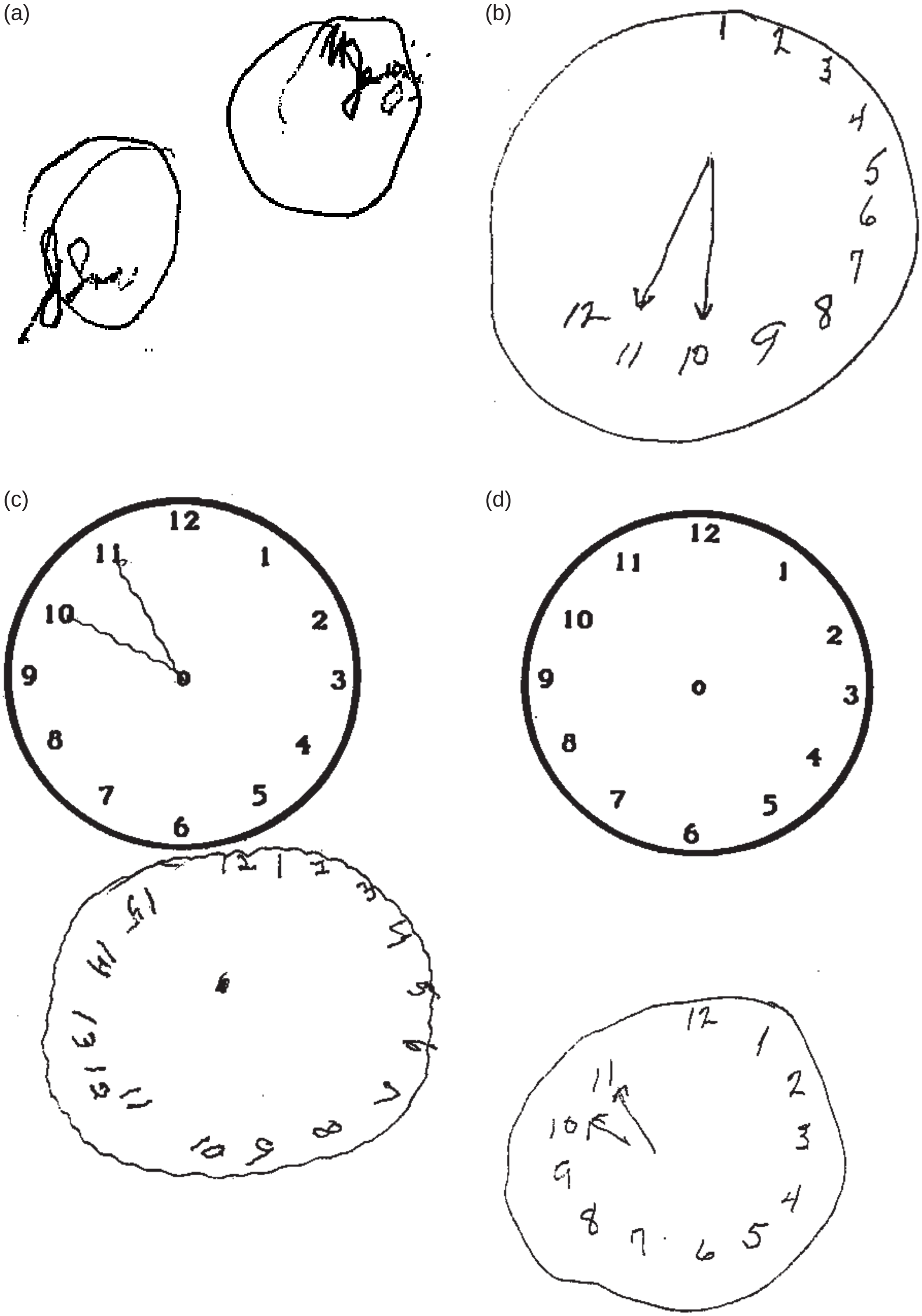

Figure 9 Clock drawings from patients showing various error types to command (9a, 9b) and copy (9c, 9d).

including a square, triangle, Greek cross, and a skeleton key (Reitan, 1984). Traditional qualitative scoring procedures for these drawings exist, although as with other similar tasks, visual inspection by a trained clinician is extremely important, and one needs to be familiar with the wide range of variability seen in the normal population as well. The copy of the Greek cross or similar crosses is a very common graphomotor constructional task, and despite its relative simplicity, nevertheless does require a variety of skills, including visuoperception, planning, graphomotor control, and sequencing.

(iii) Block Design

The Block Design subtest from the WAIS-III or WAIS-R provides one of the most common means of assessing visuoconstructional ability. This familiar task requires the replication of red and white designs using three-dimensional colored blocks. As with other visuoconstructional tasks, careful observation of the patient's performance during the procedure can yield very important information, as can the qualitative nature of the designs that are produced. This serves as some of the rationale for including spaces for noting the designs produced by patients in the WAIS-III summary form, as the same score can be achieved for a variety of reasons. To illustrate, Figure 10 depicts the designs from two brain injured patients.

Technically speaking, both productions would receive a WAIS-III score of zero, yet they obviously reflect very different underlying aspects of dysfunction. As has been discussed by Kaplan and co-workers (Kaplan, 1990; Milberg, Hebben, & Kaplan, 1996), many patients with right hemisphere dysfunction tend to break the overall square configuration of the design, as they have difficulties in processing the *gestalt* of figures. In contrast, patients with damage to the left hemisphere tend to more frequently make internal rotation errors. In Figure 10 , the design on the left was, in fact, produced by a patient with a large subcortical infarct in the left-hemisphere, and the design on the right was produced by a patient with a large middle cerebral artery stroke.

(iv) Rey–Osterrieth complex figure

Although the Rey–Osterrieth figure is perhaps most well known as a measure of nonverbal memory, the copy of the figure provides an excellent sampling of visuoconstructional skills (Corwin & Bylsma, 1993; Osterrieth, 1944; Rey, 1941). In addition to providing data pertaining to an individual's organizational approach to the task (assuming the sequential replication of

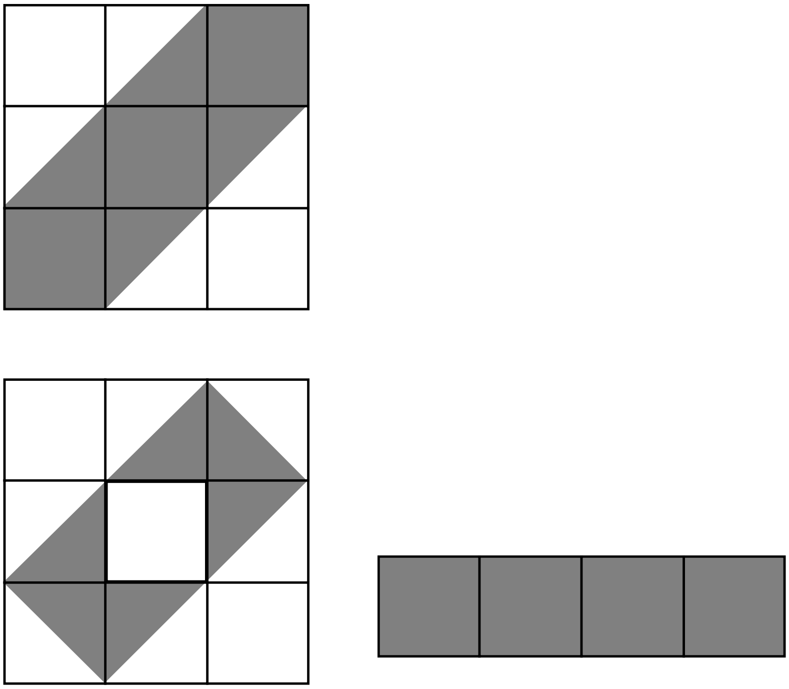

Figure 10 Block Design reproductions (top = model) of a patient with left hemisphere damage (left) and right hemisphere damage (right) showing characteristic internal detail and configural errors, respectively, Note that technically, both reproductions would receive a score of zero, although they reflect very different types of errors.

the figure is somehow carefully tracked and recorded), visuoperceptual as well as visuospatial and graphomotor constructional skills are necessary components to this task. Information pertaining to lateralized dysfunction can also be obtained in some cases (Rapport, Dutra, Webster, Charter, & Morrill, 1995). This requires not only careful attention to the *gestalt* vs. internal details of the figure, but also to left–right and quadrant by quadrant analysis, in addition to an assessment of item and figural rotations and perseverations. Gross omission or distortion of part of the figure might also implicate lateralized dysfunction. For example, Figure 11(a) depicts the Rey–Osterrieth from a patient following the rupture of a right middle cerebral artery aneurysm. The Rey–Osterrieth stimulus figure is presented in Figure 11(b).

4.11.4.8 Memory

Memory is one of the most common clinical complaints of patients referred for neuropsychological evaluation. This represents a complex set of abilities that are all too often subsumed under a generic and unitary rubric of "memory." While multiple memory systems and skills have been identified and discussed over the years (Schacter & Tulving, 1994), some of the more useful concepts from a clinical standpoint include immediate, recent, and remote memory. Because of the multifactorial nature of memory and the use of various labels and terms to describe this set of abilities, some basic definitions of memory concepts are in order before proceeding to a discussion of common clinical memory assessment tools.

Immediate or working memory refers to a very brief memory store where information is retained for a matter of seconds before it decays, in the absence of rehearsal. A prime example is going to the phone book to look up a new number, walking across the room to the phone, and dialing the number. Unless the number is of particular relevance or is rehearsed, it will probably be forgotten immediately or shortly after the call. Evidence suggests a particularly prominent role of the prefrontal cortex in working memory.

Recent memory is the ability to encode new information and remember things learned in the recent past. Temporal lobe structures (particularly the hippocampal formation) are particularly important in this stage of information acquisition. Examples include remembering a few words or an address presented several minutes ago, or remembering what you had for breakfast. This type of memory is what is routinely assessed clinically and is what allows us to learn new information each day.

Remote memory pertains to the ability to recall information from the distant past, for example, from many years ago. Remote memory involves the recollection of previously stored data (over the course of months or years) that has been rehearsed or retrieved multiple times and in various contexts (e.g., where you went to high school, aspects of your job, stories you have learned, etc.).

In most cases of acquired cerebral dysfunction, recent memory is the most vulnerable to damage (i.e., anterograde amnesia), while remote memory tends to remain relatively more intact. However, retrograde amnesia does occur, and often extends back into time (sometimes decades, as in Korsakoff's syndrome) in a temporal gradient fashion, with a greater disruption of more recent memories. The clinical assessment of memory involves multiple steps and procedures. At a basic level, obtaining a history from a patient involves their remote memory abilities. Questions pertaining to familiar information historical events can also be useful (e.g., as in the WAIS-III Information subtest).

Recent memory is best informally assessed by presenting the patient with a few words, name and address, or a paragraph to recall. Here it is essential to stress that delayed recall of such information must be examined. For example, amnestic patients can often repeat back even a brief story immediately after it is presented, and it is only after additional time has elapsed (e.g., 10–15 minutes) that their full amnestic deficit may be manifest. The use of multiple memory procedures is also recommended, since significant variability even among healthy normal subjects can be seen on singular brief tests of memory. Furthermore, the use of standard or favorite memory stimuli is also encouraged (i.e., rather than inventing new stimuli for every patient), since the specific content of items can have a profound effect on their ability to be recalled (e.g. "rose, ball, key" or "apple, table, penny" are easier for most people to remember than "brown, honesty, tulip," and these sets are no doubt easier to recall than "abstemious, benign, pusillanimous"). Regardless of what brief memory tasks are employed at the bedside, it must be kept in mind that more lengthy psychometric procedures may be required to elicit deficits in many cases.

4.11.4.8.1 Clinical assessment of learning and memory

Memory can fail for a wide variety of reasons. Problems with attention/concentration which result in an inefficiency in the ability to attend to and learn new information may be perceived as

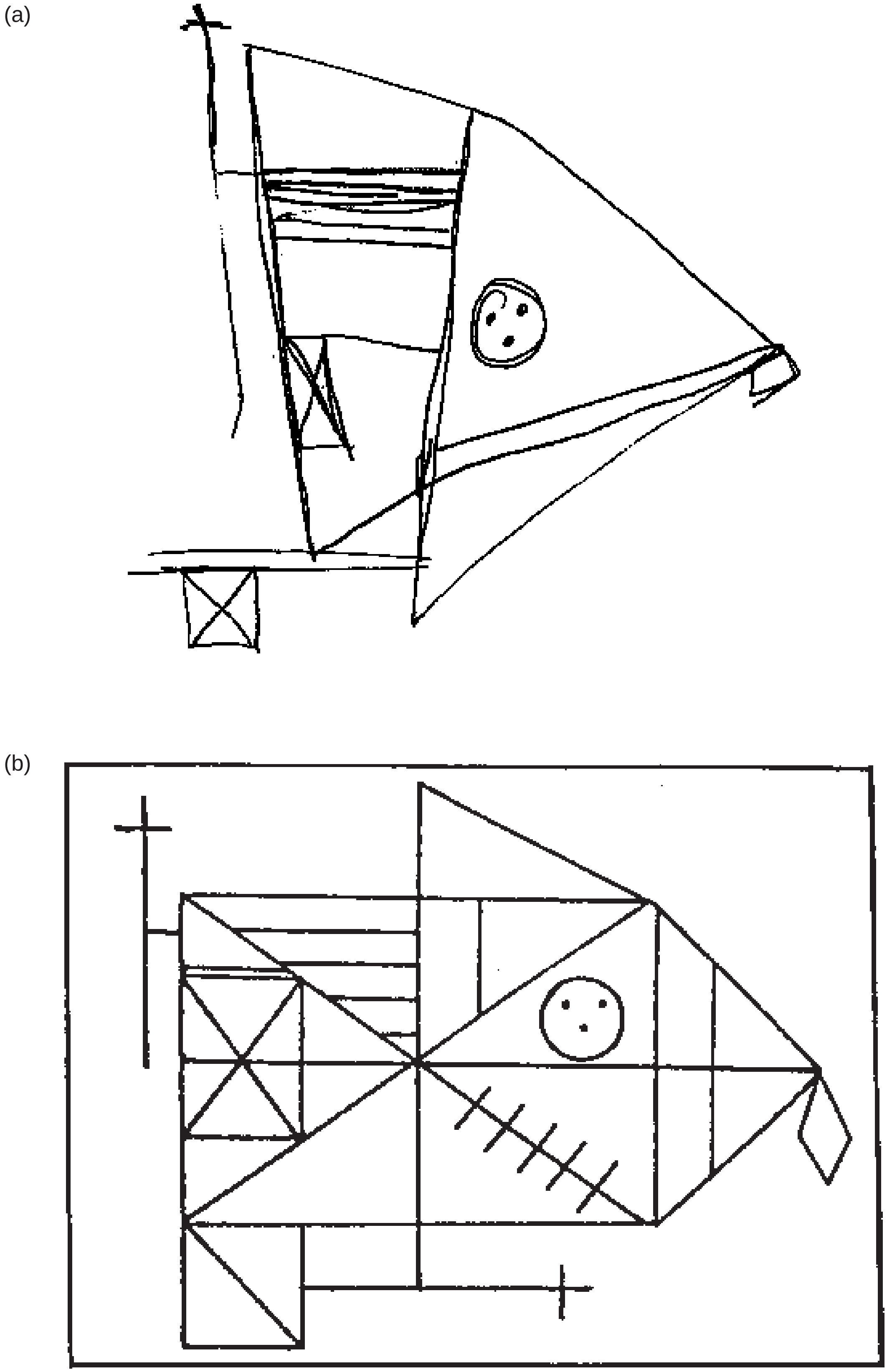

Figure 11 Rey–Osterrieth copy by a patient with right hemisphere damage (11a). Note the distortion of the left side of the figure in particular. The original Rey–Osterrieth figure is (11b).

a problem with memory, for example. If information is not efficiently encoded, it obviously cannot be stored for later recall. Subjectively, this can be perceived as a problem with remembering. Other component abilities that can interfere with efficient memory functioning include language disturbance and problems in visuospatial processing, depending upon the nature of the material that is to be remembered.

Some of the different learning and memory types that have been proposed include episodic, semantic, procedural, implicit, explicit, and working memory. A discussion of these is beyond the purpose of this chapter, although detailed elaborations can be found in sources such as Schacter and Tulving (1994) and Squire (1987). At a basic level, most of the clinical assessment approaches to memory involve an examination of episodic memory, insofar as patients are required to recall material that was presented to them earlier. Typically this material is presented in an explicit memory paradigm wherein patients are instructed to learn and remember information, although in some procedures, implicit recall is required. One example of the latter would be the typical administration of the Rey–Osterrieth complex figure, wherein subjects are asked to copy the complex design, and once that is done, they are asked to recall it from memory, without ever having been instructed to "remember" the figure.

(i) Verbal and nonverbal memory

Verbal learning and memory tasks (e.g., word lists, verbal paired associates, and story recall) have been shown to be strongly associated with left temporal lobe functioning (Chelune, 1995; Hermann, Wyler, Richey, & Rea, 1987). More debate exists about so-called "nonverbal" measures of memory and their respective sensitivity to right temporal lobe dysfunction (Barr et al., 1997). Some of this discussion centers around the nature of nonverbal stimuli that are employed in many popular memory tests. For example, many of the figures that are used are verbally encodable or lend themselves quite readily to verbal labels, thereby perhaps involving left temporal lobe structures to a significant degree. Other debate exists with regard to the specificity of functioning of the right temporal lobe, as some argue whether there is a "nonverbal" memory system at all. This is perhaps nowhere better illustrated than in cases of right temporal lobectomy in the treatment of intractable epilepsy, wherein scores on some nonverbal memory tests may show no significant change following surgery (Naugle, Chelune, Cheek, Luders, & Awad, 1993).

(ii) Forgetting rates

An important concept in memory assessment includes the examination of delayed recall for newly learned information. The simple assessment of the ability to immediately recall newly presented information, for example, may not be significantly impaired in individuals with even gross memory disorders. This can be illustrated by many patients with Alzheimer's disease, as their ability to recall a brief paragraph that was just presented may be relatively intact, even though they show a rapid rate of forgetting over time. The seminal work by Butters and colleagues (e.g., Butters, 1985) clearly demonstrated the importance of including delayed recall techniques in the assessment of memory disorders. Along these lines, the calculation of forgetting rates or percent loss over time has proven most useful in clinical memory assessment.

(iii) Patterns of memory dysfunction

In addition to examining immediate vs. delayed recall and looking at forgetting rates, it is important to carefully analyze patterns of memory performance across individual patients. Different disorders are now known to demonstrate characteristic memory profiles in many cases, and the degree to which an individual's profile matches what is known in the literature may be of great assistance in differential diagnostic situations. As one example, patients with Alzheimer's disease typically show impoverished learning, rapid forgetting, faulty encoding and storage of information, and a tendency to rely upon more passive, less efficient learning strategies (Bondi, Salmon, & Butters, 1994). Their recall of stimuli that has been recently presented to them is often marked with intrusion errors (i.e., inserting features of related or unrelated stimuli during their recall of recently presented material). Like the amnestic Korsakoff's disease patient, individuals with Alzheimer's disease may tend to confabulate during recall in an apparent attempt to "fill in" for the material which they are no longer able to effectively organize or recall.

Another example of a memory pattern that has been shown useful in clinical assessment is the distinction between recall and recognition. Free recall involves the recollection of previously presented material. Cued recall additionally involves the presentation of a cue to help "jog" an individual's memory. For example, if a recently presented story was about a woman who was robbed, a recall cue might be, "it was about a woman." The next step in assisted memory is the recognition paradigm, wherein subjects are presented with several possible stimulus items that were presented to them earlier, and the task is simply to identify which item was previously presented. It has been argued that such recognition paradigms represent "easier" tasks, insofar as the load on memory processing is diminished. An everyday analogy would be driving to a new location, being unable to remember the name of the street, but then quickly recognizing the name of the street once it is seen. Some patients with various forms of amnesia demonstrate the ability to learn, even in the absence of awareness (e.g., Schacter & Tulving, 1994), and such mechanisms may be involved in recognition memory performance.

4.11.4.8.2 Clinical memory tests

An enormous array of published measures exist for the clinical examination of memory skills, and an even larger number of tasks have been developed in the experimental and cognitive psychology literatures. A few of the tests that are commonly used in clinical practice are discussed below, and an attempt has been made to avoid too many overlapping measures. Before presenting a brief overview of individual memory assessment tools in common clinical use, one popular omnibus memory battery will be discussed. It should be noted, however, that many of the subtests which comprise the more global memory batteries are often used individually; hence, brief mention of some of the more popular subtests will be made in the following subsections.

4.11.4.8.3 Memory batteries

(i) Wechsler memory scales

This series of memory assessment batteries has a long tradition and represents some of the most popular clinical memory tests in use in the late 1990s. Because each version of the test contains various significant revisions, they will be discussed separately, although there is some overlap of core concepts and general approaches to memory testing.

(a) Wechsler Memory Scale (WMS). This was the original version (Wechsler, 1945) of the popular Wechsler Memory Scales. The WMS was designed to provide a brief assessment of overall memory function by combining several tests of attention and memory. The subtests include the following: Personal and Current Information, Mental Control (e.g., counting down from 20, reciting the alphabet, counting by threes), Digit Span, Logical Memory (paragraph recall), Visual Reproduction (immediate recall of a series of geometric figures), and Associate Learning (learning verbal associations such as "metal–iron," "crush–dark." A total Memory Quotient (MQ) score could be derived to permit a more direct comparison with traditional IQ scores. Thus, a significant discrepancy between MQ and IQ scores would indicate a particular impairment of recent memory. In its original version, the WMS assessed only immediate recall of verbal and nonverbal material, in addition to orientation/mental control and verbal paired associate learning. Importantly from a memory assessment standpoint, Russell (1975) added delayed recall procedures to the Logical Memory and Visual Reproduction subtests, resulting in what became know as the Russell revision of the WMS.

(b) Wechsler Memory Scale-Revised (WMS-R). Administration and scoring procedures were revised and made more specific in this version of the WMS, and the floor effects of the previous test were addressed (Wechsler, 1987). Importantly, delayed recall measures were built into several of the subtests so as to assess forgetting over time. A number of new subtests to assess various aspects of memory were added, which allows for a more comprehensive assessment of different aspects of learning and memory, but also adds substantially to the time. Individual subtests include Information and Orientation, Mental Control, Digit Span, Visual Memory Span, Logical Memory I (immediate recall) and II (delayed recall), Verbal Paired Associates I and II, Figural Memory, Visual Paired Associates I and II, and Visual Reproduction I and II. The following memory index scores can be derived: Attention/Concentration, Verbal Memory, Visual Memory, General Memory, and Delayed Memory. The WMS-R was a welcome revision to the WMS, addressing some of the concerns about the earlier scale and providing much-needed improvements in standardization and norming procedures. While it has proven to be useful in differentiating various memory-disordered populations (Tröster, Jacobs, Butters, Cullum, & Salmon, 1989), the utility of the Verbal and Nonverbal Memory indices in distinguishing lateralized cerebral dysfunction is questionable (Loring, Lee, Martin, & Meador, 1989). In addition to the norms available in the test manual (age 16–74), normative data for the elderly (up to age 94) are available (Ivnik et al., 1992).

(c) Wechsler Memory Scale-III (WMS-III). This represents the most recent revision of the WMS (Wechsler, 1997). The WMS-III was normed along with the WAIS-III, thereby providing for IQ and memory scores derived from the same standardization population. This allows for more direct comparisons between IQ and memory scores which should facilitate

inferences regarding the likelihood of acquired memory impairment. The standardization sample was selected to be more representative of the US population, and included individuals up to age 89. New subtests were added, and some are optional in terms of the derivation of memory index scores. Scoring procedures were further refined, updated, and expanded, and the available computer-generated scoring greatly facilitates the comparison of WMS-III and WAIS-III scores alike. Separate scores comparing working memory and immediate memory, auditory (previously referred to as "verbal") and visual (previously "nonverbal") memory, immediate and delayed recall, single-trial vs. multitrial learning, and percent retention composites in each domain are now provided. Recognition memory procedures were also added to several of the WMS-III subtests. The subtests of the WMS-III include: Information and Orientation, Logical Memory, Verbal Paired Associates, Word Lists, Faces, Family Pictures, Visual Reproduction, Letter–Number Sequencing, Spatial Span, Digit Span, and Mental Control. Whereas this is a lengthy list of measures, five of the subtests (Information and Orientation, Word Lists, Visual Reproduction, Mental Control, and Digit Span) are supplemental or optional and do not contribute to the primary index scores. From various combinations of the core subtests, the primary index scores can be derived: Auditory Immediate, Visual Immediate, Immediate (combined), Auditory Delayed, Visual Delayed, Auditory Recognition Delayed, General Memory, and Working Memory.

The WMS-III is obviously comprehensive in its scope, now including aspects of a variety of memory assessment stimuli and procedures. It is lengthy and overlaps with some other standard memory tests in common use, but contains some interesting new procedures based on research and clinical experience in neuropsychology. Because of its recent release, studies regarding its ultimate clinical utility in various settings and populations will be needed, in addition to the development of procedures for deriving short forms and estimating memory index scores. Perhaps even more so than with the WMS-R, it is anticipated that when administration of the entire test is not feasible, many clinicians will routinely utilize selected subtests of the WMS-III to fit their particular memory assessment needs.

4.11.4.8.4 Verbal memory tests

(i) California Verbal Learning Test (CVLT)

This is a 16-item word list-learning task that contains items from four semantic categories (Delis, Kramer, Kaplan, & Ober, 1987). The list is presented five times to subjects, in the same order, which allows for not only an assessment of learning, but of primacy-recency recall effects. The CVLT was developed with principles of cognitive neuroscience and clinical memory disorders in mind, and provides a quantitative assessment of a wide variety of qualitative features of memory performance. It has been shown to be reliable (Paolo, Troster, & Ryan, 1997) and clinically useful in a variety of clinical populations, particularly in the differential diagnostic assessment of memory disorders (Cullum, Filley, & Kozora, 1995; Peavy et al., 1994).

After the fifth presentation of the word list, an interference list is presented to permit an examination of the effects of interference on recall. Next, free recall of the initial word list is examined, followed by cued recall that involves the presentation of semantic categories for the words (e.g., "Tell me all the items that were fruits"). Long delay free recall is examined following a 20-minute delay. This is immediately followed by another series of cued recall trials, and finally, a recognition trial wherein a longer list of items containing target items as well as distractors is presented, and subjects must simply indicate "yes" or "no" whether or not these items appeared in the originally presented list. As noted, the CVLT provides for a quantified assessment of a variety of qualitative aspects of verbal learning and memory, and scoring software is available to provide normative reference scores for a host of verbal learning and memory indices. An alternate form of the CVLT has been developed that has shown good correlations with the original version (Delis et al., 1991), and a revised CVLT is in preparation in 1998. A nine-item dementia version of the CVLT has also been derived for use in more impaired populations (Libon, Mattson, et al., 1996).

(ii) Hopkins Verbal Learning Test (HVLT)

This test was developed to provide an abbreviated assessment of word list-learning and memory (Brandt, 1991). It is particularly well suited for more impaired patients who might not be able to complete measures such as the CVLT. A list of 12 items from three semantic categories is presented across three trials, followed by recognition testing. A modification of the original HVLT provides for an assessment of delayed recall prior to recognition in order to assess forgetting over time. A unique aspect of the HVLT is that it includes six alternate forms, thereby reducing practice effects and making it a good choice in serial assessment situations.

(iii) Logical Memory (WMS, WMS-R, WMS-III)

This is the classic story learning and memory test that is one of the most widely used clinical indices of memory function. It is a measure of the ability to learn and retain new structured verbal information. The original WMS included two stories that were read to patients, followed by an assessment of immediate recall. The WMS-R updated the stories, norms, and scoring procedures, and importantly, added a delayed recall trial. Each of the WMS-R stories contains 25 bits of information, and a standard 30-minute delayed recall procedure provides for a ready assessment of forgetting rates. The WMS-III version includes an initial story that is administered in standard fashion (i.e., one trial, followed by immediate and then delayed recall), and a second story that is repeated in order to assess the effects of learning on storage and retrieval of material.

Patients' responses should be written down verbatim in order to glean additional qualitative information, and to allow for cross-checking of scoring. Scoring for Logical Memory is done largely based upon identical or almost-identical recall of the story material, although several scoring procedures allow partial credits to be assigned for close approximations, and the WMS-III version also incorporates scores for thematic units or gist. Logical Memory has been shown to be highly useful in the assessment of a wide array of memory disorders and senstive to left hippocampal damage in particular (Sass et al., 1992).

4.11.4.8.5 Nonverbal memory tests

(i) Rey–Osterrieth complex figure

This is a very popular measure of nonverbal memory and is described and presented in several sources (e.g., Corwin & Bylsma, 1993; Lezak, 1995; Rey, 1941; Spreen & Strauss, 1998). It consists of a complex figure made up of various subcomponents. Administration procedures vary somewhat (see Meyers & Meyers, 1995), although in one of the more common administrations the task requires patients to first copy the figure, and then immediately upon its removal (and without instructions to remember the figure), reproduce it from memory. Fifteen minutes later, delayed recall is assessed. Recognition procedures have also been developed and used in some settings. Several scoring systems for the test have been developed (e.g., see Lezak, 1995; Loring, Martin, Meador, & Lee, 1990), although qualitative interpretation of patients' reproductions often provide much more detailed information than quantitative scores alone. The Boston scoring approach (Stern et al., 1994) to the Rey–Osterrieth attempts to provide a quantitative assessment of many of the qualitative features of the figure.

(ii) Visual Reproduction (WMS, WMS-R, WMS-III)

This is one of the most commonly used measures of visual or nonverbal memory, and is often used in conjunction with its verbal counterpart, Logical Memory, even when the WMS are not used in their entirety. Although each version of the WMS contains somewhat different stimulus figures (only two of the original WMS figures have been included in both the WMS-R and WMS-III, and Visual Reproduction is optional in the latest rendition), the administration procedures are similar. It consists of the presentation of several cards with geometric designs on each. Each card is presented to subjects for 10 seconds, and after each is removed from view, immediate recall is assessed by having the subject draw the design from memory. Following a delay of 20–30 minutes (25–35 minutes in the WMS-III), subjects are asked to reproduce the designs again from memory.

As noted, Visual Reproduction has long been used as a measure of nonverbal memory. While some literature indicates that the test does not readily lend itself to localization of cerebral dysfunction in large groups and may be insensitive to nondominant temporal lobe resection (Barr et al., 1997; Naugle et al., 1993), it can be useful in some individual cases as there may be a tendency for patients with right temporal lobe dysfunction to do more poorly (Chelune & Bornstein, 1988). Furthermore, it can be of assistance in providing a more detailed examination of memory function beyond the assessment of learning and memory for words and stories.

(iii) Warrington Recognition Memory Test (RMT)

This is a test of word and face recognition that allows for a comparison between basic verbal and nonverbal abilities (Warrington, 1984). First, a series of 50 single words is presented in a booklet, followed by yes/no recognition testing. The same procedure is performed for a series of 50 black and white pictures of unfamiliar male faces. While some studies of patients with lateralized brain damage have demonstrated the expected associations between left hemisphere/word recognition and right hemisphere/face recognition deficits (Warrington, 1984), this relationship does not always hold true (Naugle, Chelune, Schuster, Luders, & Comair, 1994), and the test is perhaps best used as a supplemental test of memory in conjunction with more primary measures. In clinical practice, many patients seem to have more difficulty with the facial recognition portion of the RMT, and the association between words and faces with left vs. right hemisphere dysfunction is variable.

4.11.4.9 Motor and Sensory Function

A variety of clinical procedures have been developed in neuropsychology and behavioral neurology to examine simple and complex motor functions. These range in complexity, degree of standardization, and level of knowledge required for appropriate interpretation/analysis. At a most basic observational level, how a patient walks and moves during an evaluation can provide useful information regarding aspects of their motor functioning. Evidence of unilateral motor weakness, tremor, or other unusual movements may reflect a neurological condition or even suggest a focal cerebral deficit. These observations can also spur the astute observer to inquire more carefully about specific symptoms and may help guide the neuropsychological evaluation in terms of test selection. To take an obvious example, the patient presenting with unilateral left-sided motor weakness (e.g., foot drag and/or decreased limb mobility) should undergo careful inquiry about their symptoms, and depending upon the situation, the clinician may wish to follow up on this observation by including appropriate sensory, motor, and cognitive tasks in his/her evaluation in order to assess abilities associated with right hemisphere function.

The primary motor and sensory areas of the cortex, as well as their multiple ascending and descending fiber tracts, are well known and charted. It must be kept in mind, however, that sensory and motor regions have significant overlap in terms of cortical surface representation, and deficits are often multimodal in nature, with more complex integrative functions having a broader distribution. In addition to the assessment of two-point discrimination and simple touch, for example, more subtle cortical deficits may be detectable by the use of double simultaneous stimulation techniques. These are typically conducted in the visual, auditory, and haptic modalities, and a variety of simple quantifiable procedures can be used. Double simultaneous stimulation in the tactile modality involves repeated trials of unilateral light touch to the back of either hand, interspersed with bilateral touch. Subjects are asked to indicate which hand was touched. Tactile extinctions or suppressions occur when one side is not perceived, and if consistent, this may suggest contralateral parietal lobe dysfunction.

Basic motor function can be quantified using various techniques and instruments, depending upon the degree of detail that is desired about a given function. Simple finger tapping speed can be assessed crudely by having the patient tap their first finger to their thumb as quickly as possible, first with the dominant, then non-dominant hands. Grip strength can be assessed by having the patient squeeze the examiner's fingers, but this technique tends to be extremely crude and is better left to a dynamometer when accurate information is desired. Fine motor dexterity can be assessed using a variety of procedures, including having the patient touch each finger consecutively to their thumb in rapid, repetitive fashion. Stereognosis can be readily examined by having the patient name objects placed in their hand without the aid of vision.

Whatever motor or sensory assessment task used, is important to note that noncortical or noncentral nervous system (non-CNS) (i.e., subcortical, spinal, peripheral) damage can produce deficits in motor and/or sensory function. Furthermore, other factors such as inattention or impairments in concentration or memory may result in poor scores or errors on sensorimotor tasks. In the absence of neuropathology, gross errors or variable responses should alert the clinician to the possibility of poor effort. In practice, motor and sensory deficits tend to be consistent, such that repeated trials later in the examination should yield similar results. Interpretation of results from these tests should always begin with these considerations in mind.

4.11.4.9.1 Psychometric measures of motor and sensory function

(i) Finger Tapping Test

This test is a simple assessment of index finger tapping speed using a mechanical tapping key device (Reitan & Wolfson, 1993). Patients are instructed to tap using their dominant hand index finger as fast as possible, and the number of taps performed within a series of 10-second intervals is recorded and averaged. These results are compared with similar scores from the nondominant hand. Most right-handed individuals tend to show approximately a 10% dominant hand superiority. Gross aberrations from this may reflect contralateral anterior cerebral dysfunction, although it is important to note that arthritis or peripheral injuries to the brachial plexus, arm, hand, or fingers can have profound effects on this test and must be carefully ruled out before cerebral implications are made.

(ii) Grooved Pegboard

This is a test of fine motor dexterity that requires subjects to pick up small (one inch) metal pegs that are rounded on one side and squared on the other (Kløve, 1965). The task is to pick them up one at a time and place them in similarly shaped holes on the test platform as quickly as possible. First, the dominant hand is tested, followed by the nondominant hand. The task requires fine motor coordination and speed.

(iii) Hand dynamometer

Various manufacturers provide grip strength meters to assess basic hand strength. This is a device that requires patients to simply squeeze a handle as hard as they can, and the amount of force they exert is indicated in pounds or kilograms on the instrument. Two or three trials are usually administered with the dominant and nondominant hands. Dominant hand grip strength usually tends to be approximately 10% stronger than nondominant in most right-handed individuals.

(iv) Sensory–Perceptual Examination (SPE)

This test is included in the HRB, but is often used apart from this battery in its entirety or in portions (Reitan & Wolfson, 1985). The basic functions assessed by the SPE include finger gnosis, graphesthesia, stereognosis, and basic sensation and extinctions to simple stimulation and double simultaneous stimulation in the visual, auditory, and tactile modalities. It essentially represents a quantified approach to some of the sensory testing often performed in the traditional neurological examination. In the absence of peripheral damage factors, arthritis, severe cerebral dysfunction, or prominent attentional difficulties (which can make the interpretation of patient responses more difficult), deficits associated with these various tasks may suggest some degree of impairment of the sensory cortex and/or associated pathways.

4.11.4.10 Assessment of Motivation

The sensitivity and clinical utility of neuropsychological measures is largely dependent upon the cooperation and effort of the patient being evaluated. Poor effort can result in artificially low or otherwise abnormal neuropsychological test scores, and the clinician must always be alert to this potential confound during test administration and interpretation. Careful analysis of levels and patterns of neuropsychological performance in relation to the disorder in question is critical (i.e., "Do the results make sense given what has happened to the patient?). For example, the patient with an uncomplicated mild traumatic brain injury should not demonstrate severe, widespread neuropsychological deficits, and this finding in such a case should raise questions regarding levels of motivation and effort. Furthermore, as serial testing frequently occurs within the context of forensic neuropsychological evaluations, the examination of consistency of scores within and across test sessions can be a very good source of information regarding the likelihood of invalidity or unreliability of results (Cullum, Heaton, & Grant, 1991; Reitan & Wolfson, 1997).

The issue of malingering in relation to neuropsychological assessment has been reviewed in several excellent sources (e.g., Franzen, Iverson, & McCracken, 1990; Nies & Sweet, 1994). Suspected malingering or willful distortion/exaggeration of deficits continues to pose challenges in various clinical settings. Along these lines, a number of specialized tests and procedures have been developed to help detect exaggerated or feigned deficits. Some of these involve the administration of simple tasks that are presented as being difficult (e.g., Rey's 15-item test; Lezak, 1995). Others that show promise are based upon the time to complete shorter or less difficult vs. longer tasks (e.g., the Dot Counting Test; Lezak, 1995; Rey, 1941). Another technique involves the comparison of unusual patterns of performance with results known to be associated with a particular deficit or disorder (e.g., Brandt, 1988). Perhaps the most popular approach has been to utilize forced-choice recognition testing (Guilmette, Hart, & Giuliano, 1993; Hiscock & Hiscock, 1989), which relies upon probability theory by identifying performances that fall below a chance level (e.g., the Portland Digit Recognition Test; Binder, 1993).

The detection of malingering poses a complex challenge to clinicians, particularly as a variety of psychological and situational factors (e.g., gain issues) may further complicate the diagnostic picture. While it is difficult to definitively prove malingering in a particular case, the existing research suggests that it is often possible to detect evidence of malingering, although this requires skillful and deliberate inquiry by the clinician.

As with diagnosis, the clinical history can prove invaluable in establishing the likelihood of malingering. Consistency of reporting is important, and information from various sources can be compared, including clinical interview data and reports from the patient to other professionals. The consistency or "fit" between neuropsychological test results and reports of everyday functioning can also be highly useful. For example, the patient who performs in the severely impaired range on formal memory testing should demonstrate difficulties with memory in their everyday functioning. Likewise, neuropsychological test data obtained from the patient who has gross difficulty with coordination, sequencing, memory, and reaction times who rides skillfully away on his/her motorcycle following the evaluation should be carefully scrutinized.

As indicated by Nies and Sweet in their review of the literature, "Note that, at present, multiple measures and methodologies are needed, in that no single measure or methodology has proven sufficient to date" (1994, p. 544). They recommend some of the following strategies that may assist in the clinical detection of malingering: (i) use specific tests of malingering as outlined above, (ii) evaluate patterns on clinical neuropsychological tests where below-chance performances can be assessed, (iii) examine unusual or aberrant test results, (iv) examine consistency of test scores within and across measures and test sessions, (v) obtain independent information regarding a patient's reports of their everyday functioning, and (vi) obtain a detailed clinical history and carefully examine the information provided, both in terms of accuracy and consistency.

4.11.4.11 Personality and Emotional Functioning

Changes in the functional integrity of brain systems can result in alterations in emotional reactivity (e.g., lability) and personality. In some cases, subtle changes in behavior and personality may represent the primary or earliest harbingers of brain dysfunction. The evaluation of emotional or personality change associated with neurological disease or damage poses a complex set of challenges. Damage to the frontal and prefrontal cortex can be particularly associated with personality change. This may present variably as decreased responsivity (e.g., apathy or amotivation) or increased responsivity (e.g., hypomania, increased or indiscrimant drives), depending upon the primary site of damage. In cases of lateralized brain dysfunction, left hemisphere lesions (particularly more anterior) have been shown to be associated with depressive symptomatology, while lesions of the right hemisphere tend to be more frequently associated with varying degrees of unawareness or denial (anosognosia) of deficit (e.g., see Lacritz & Cullum, 1998).

Right hemisphere damage can also result in aprosodia, which refers to the impaired ability to communicate meaning using intonations and inflections during speech (emotional language). Some data suggest that the aprosodic syndromes secondary to right hemisphere dysfunction parallel the aphasias (Ross, 1985), with anterior right hemisphere lesions being more associated with expressive dysprosody, and more posterior lesions associated with receptive dysprosody. Expressive prosody testing can be readily accomplished by asking the patient pretend to be an actor and say a neutral sentence such as "I am going to the movie" using different affective intonations (happy, sad, angry, indifferent). Receptive prosody can be assessed by the examiner (while outside of the patient's view) saying a sentence in different tones of voice and having the patient indicate if the examiner sounded happy, sad, angry, indifferent, etc. The patient with expressive dysprosody will generally sound monotone and be unable to alter the affective tone of their voice across test conditions aside from perhaps becoming louder or softer at times. The patient with receptive dysprosody will be able to say the sentence with variable affective intonations, but will have difficulty guessing the affective valence of the examiner's sentences.

The issue of changes in emotional functioning following brain damage represents a complex biopsychosocial interaction that merits careful investigation, since multiple factors may be involved. Obviously, an insult to the brain can have a direct neurobiological effect upon the way in which information and affect is processed. However, individual psychological reactions to changes in cognitive and/or physical functioning are also commonly seen. To help establish whether neurobiologically-induced changes have occurred, the clinician must be careful to obtain information regarding patients' premorbid functioning from family members or others who know the patient, as well as patients themselves. Defensiveness or unawareness on the part of the patient and/or family may exist, and it must be kept in mind that the clinician may be hearing a biased picture of the situation. The issue of secondary gain and malingering as discussed above must also be considered, particularly in litigation situations. Finally, the potential effects of pre-existing psychopathology should be examined, and psychiatric and/or more detailed psychological evaluation may be useful along these lines. Direct questioning about symptoms and behavioral changes obviously may be productive, although in some instances, the line of inquiry must be more subtle, requiring indirect probing of evidence for personality/emotional changes. Rephrasing questions and asking about related symptoms in a variety of contexts may also be a useful means of not only verifying information, but in some cases, uncovering important data.

Formal psychological assessment in patients with neurological disorders requires some adjustments to standard interpretations, insofar as the endorsement of certain symptoms may simply relate to the neurological condition rather than implying psychopathology. For example, patients with neurologic disease often demonstrate clinical elevations on the Minnesota Multiphasic Personality Inventory-2 (MMPI-2) that should not be necessarily attributed to psychological disturbance. Scale 8 in particular contains a number of items dealing with unusual symptoms, and scales 1 and 3 relate to aspects of physical functioning which may be endorsed by neurological patients due to their physical status, in the absence of psychopathology. Table 7 presents some of the more commonly used measures of psychological functioning within the context of the neuropsychological evaluation.

4.11.4.12 Test Selection Issues

As noted, specialized core batteries of tests are often assembled by neuropsychologists for specific populations. Test selection is based on training and experience in combination with information obtained from the relevant literature. Table 8 presents an example of a test battery designed to assess dementia (both in terms of early detection as well as assistance with differential diagnosis and staging of disease). Note that individual measures can readily be added or dropped (i.e., a "stepdown battery") from the "core" set of tests, depending upon specific patient needs and clinical questions.

An example of a modification to this battery in the case of a high functioning business executive referred for evaluation secondary to suspected incipient Alzheimer's disease might include the following: administration of a complete WAIS-III to obtain a more precise assessment of current intellectual functioning and allow for analysis of subtest patterns; possible omission of the DRS since it may not be sufficiently sensitive to subtle deficits in such an individual in the early stages of dementia; addition of the category test to provide a more challenging evaluation of higher-order cognitive abilities than the WCST; substitution of the Ravens Advanced Matrices for the coloured version, as the latter would likely be too simplistic to be of much value in such a case; administration of a complete WMS-III to provide a more thorough examination of memory function.

Alternatively, the test battery might be dramatically altered in the case of a patient with known dementia who is being followed to track their rate of cognitive decline (or perhaps to assess for any positive effects of cognitive enhancing medications). As a more extreme example, if an individual presents with clear evidence of severe dementia during clinical interview (or if the patient had a previous MMSE score of 10), it may suffice to administer the DRS and obtain only a brief screening assessment of cognitive domains (e.g., by administering all or parts of the CERAD neuropsychological test battery). In such a case, a brief neuropsychological examination may nevertheless be considered "comprehensive" insofar as a range of abilities is assessed, and the addition of multiple complex measures may actually yield little additional neurobehavioral data, depending on the individual patient and referral questions.

4.11.4.13 Relationships Between Neuropsychometry and the Behavioral Geography of the Brain

Neuropsychological procedures, like the neurological examination, are designed to provide information about the integrity of human brain function. While these measures are sensitive to impaired cognition and may provide information regarding lateralized or focal dysfunction, it

Table 7 Common measures of psychological and emotional functioning.

Minnesota Multiphasic Personality Inventory-2 (MMPI-2)	Psychopathology/emotional status
Personality Assessment Inventory (PAI)	Psychopathology/emotional function
Millon Clinical Multiaxial Inventory (MCMI)	Personality styles, psychopathology
Beck Depression Inventory (BDI)	Depression
Inventory of Depressive Symptomatology (IDS)	Depression
Profile of Mood States (POMS)	Brief review of psychiatric symptoms
Symptom Checklist-90 (SCL-90)	Brief review of physical and psychiatric symptoms

Table 8 Sample test battery for the evaluation of dementia.

Global cognitive functioning
  - Dementia Rating Scale
  - WAIS-III/WAIS-R IQ or IQ estimate (e.g., based on Vocabulary and Block Design)
  - NART-R estimated premorbid IQ

Reasoning/Cognitive flexibility
  - Wisconsin Card Sorting Test
  - Similarities and proverbs testing (WAIS-III)
  - Trail Making Test, Part B

Language
  - WAIS-III Vocabulary
  - Boston Naming Test
  - Verbal fluency (letter and category, e.g., FAS, animals)
  - Token Test and/or Comprehension subtest from BDAE

Visuospatial skills
  - WAIS-III Block Design
  - Clock drawing
  - Rey–Osterrieth Complex Figure (copy)

Attention/concentration
  - WAIS-III Digit Span (forward)
  - Trail Making Test, Part A

Learning and Memory
  - California Verbal Learning Test
  - Wechsler Memory Scale-3 (WMS-III) or Wechsler Memory Scale-Revised (WMS-R) or portions thereof (e.g., Logical Memory and Visual Reproduction)
  - Rey–Osterrieth Complex Figure (immediate and delayed recall)

Motor functioning
  - Finger tapping test
  - Rapid Alternating Movements/Luria three-step

should be kept in mind that the neuroanatomical specificity of most of the tests is limited (Dodrill, 1998). Furthermore, while the frontal, temporal, parietal, and occipital cortices subserve specific functions to some degree, the amount of overlapping functional representations and inter-relationships cannot be underscored. It is worth recalling that even "primary" motor and sensory regions of the cortex have overlapping cellular representations and functions to some degree. Thus, the extent to which various cognitive functions are said to be "localized" is somewhat relative, although a number of functional neuroanatomic correlations exist with various measures. Higher cognitive functions in particular tend to defy localization, although frontal systems (particularly prefrontal cortices) tend to be heavily involved in the many abilities associated with executive functioning (e.g., planning, reasoning, judgment, insight) and personality. Limbic structures of the temporal lobes play a predominant role in learning and memory for new information, with verbal memory skills being strongly associated with left hemisphere (specifically temporohippocampal) functions in most individuals. Although less well supported, association is also often seen between nonverbal memory abilities and nondominant temporal lobe functioning. The limbic system is also intimately associated with aspects of emotional processing, and links between affective memory and limbic functioning have been shown. While functional neuroanatomic distinctions exist (by lobe as well as in the anterior–posterior and lateral planes), the myriad interconnections that exist between these areas and other cortical and subcortical systems indicate that various brain functions may be disrupted when specific circuits or systems are damaged, even when the cortical region known to be largely responsible for a given function may appear to be intact. Thus, a careful probing of multiple aspects of brain–behavior relationships is required to provide a thorough assessment of function.

4.11.4.14 Training in Neuropsychology

Because of the complexities inherent in the neurobehavioral assessment of human brain function, specialized training is needed for the clinical neuropsychologist. Neuropsychologists must have a firm understanding of basic brain organization and function, and need to be familiar with neurological disorders and neuropathology as well as their myriad behavioral sequelae. As such, specific knowledge bases must be mastered that go beyond the traditional training of the doctoral degree in psychology.

Several documents relevant to training and practice in the specialty of clinical neuropsychology have been published over the years (e.g., Adams & Rourke, 1992; Eubanks, 1997; Reports of the INS-Division 40 Task Force on Education, Accreditation, and Credentialing, 1987), although the most comprehensive and far-reaching proviso to date was published in 1997. The Houston Conference on Specialty Training and Education in Neuropsychology was held in September 1997 and outlines recommended training and education guidelines for the future in clinical neuropsychology (Hannay et al., 1998). For this conference, neuropsychologists representing various backgrounds, employment settings, and organizations were selected from applications across the USA and Canada and invited to participate. One of the goals of the Houston Conference was to develop the guidelines for an integrated model of education and training in clinical neuropsychology, and this includes the following generic definition of a clinical neuropsychologist:

> A clinical neuropsychologist is a professional psychologist trained in the science of brain–behavior relationships. The clinical neuropsychologist specializes in the application of assessment and intervention principles based on the scientific study of human behavior across the lifespan as it relates to normal and abnormal functioning of the central nervous system. (p. 161)

The recommendations for coursework, training, and skills are provided in some detail, including a model of integrated education across the doctoral, internship, and postdoctoral or residency levels. In addition to core courses in clinical psychology (e.g., psychopathology, assessment, ethics), coursework in areas pertaining specifically to brain–behavior relationships are recommended, including neuroanatomy, psychopharmacology, and various aspects of neuropsychology *per se*. It is noted in the Policy Statement of the Houston Conference that the proposed guidelines were not intended to be applied retroactively or to those currently in training at the time of the conference and proceedings, but rather, to serve as recommendations for future education and training in clinical neuropsychology. Whereas the specific functions of neuropsychologists vary depending upon setting, the following seven professional activities were put forth as being important for the clinical neuropsychologist: assessment, intervention, consultation, supervision, research and inquiry, consumer protection, and professional development. A postdoctoral residency of the equivalent of two years full time experience is included in these recommendations. Some of the exit criteria for the postdoctoral residency include demonstration of advanced skills in neuropsychology, eligibility for state or provincial licensure, and eligibility for board certification in clinical neuropsychology.

4.11.4.15 Challenges for Neuropsychological Assessment

The neuropsychological evaluation represents the most sensitive existing means by which to assess the functional integrity of the human brain. Modern structural imaging provides us with wonderfully detailed images of intracranial contents, but even knowing that an abnormality exists in a given region cannot provide much insight into how that patient is able to function and process information in day-to-day life. Functional imaging has seen major advancements in recent years and is able to indicate areas of relative dysfunction—but this, too, cannot indicate what specific cognitive abilities are impaired, or to what extent. Such techniques also offer little or no insight into how to help brain-injured patients adapt and better compensate for their altered cognitive processing, as this is the purview of clinical neuropsychology.

Neuropsychological assessment is faced with the ongoing challenge of improving diagnostic and prognostic procedures and integrating findings from clinical neuropsychology, cognitive psychology, and neuroscience into assessment tools that can provide increasingly detailed examinations of some of the intricacies of human brain function. Many of the neuropsychological tests in common use in the 1990s were developed years or decades ago, and even though some procedures are "tried and true" and continue to provide important neurodiagnostic and neurodescriptive information, the field must strive to keep up with the rapid advances being made in neuroimaging and neuroscience and be open to incorporating this new information into assessment procedures.

ACKNOWLEDGMENTS

Thanks to Kathy Saine, Ph.D., and Eric Smernoff, M.A., for their assistance with this manuscript.

4.11.5 REFERENCES

Adams, K. M., & Rourke, B. P. (1992). *The TCN guide to professional practice in clinical neuropsychology*. Amsterdam/Lisse: Swets & Zeitlinger.

Adams, R. L., Smigielski, J., & Jenkins, R. L. (1984). Development of a Satz-Mogel short form of the WAIS-R. *Journal of Consulting and Clinical Psychology, 52*, 908.

Baddeley, A. (1992). Working memory. *Science, 255,* 556–559.

Barr, W. B., Chelune, G. J., Hermann, B. P., Loring, D. W., Perrine, K., Strauss, E., Trenerry, M. R., & Westerveld, M. (1997). The use of figural reproduction tests as measures of nonverbal memory in epilepsy surgery candidates. *Journal of the International Neuropsychological Society, 3,* 435–443.

Beatty, W. W. (1993). Age differences on the California Card Sorting Test: Implications for the assessment of problems solving by the elderly. *Bulletin of the Psychonomic Society, 31,* 511–514.

Benton, A. L., & Hamsher, K.de-S. (1989). *Multilingual Aphasia Examination.* Iowa City, IA: AJA Associates.

Benton, A. L., Hamsher, K.de-S., & Sivan, A. B. (1994). *Multilingual Aphasia Examination* (3rd ed.). Iowa City, IA: AJA Associates.

Bigler, E. D. (1988). Frontal lobe damage and neuropsychological assessment. *Archives of Clinical Neuropsychology, 3,* 279–297.

Binder, L. M. (1993). An abbreviated form of the Portland digit recognition test. *The Clinical Neuropsychologist, 7,* 104–107.

Blair, J. R., & Spreen, O. (1989). Predicting premorbid IQ: A revision of the National Adult Reading Test. *The Clinical Neuropsychologist, 3,* 129–136.

Bondi, M. W., Salmon, D. P., & Butters, N. (1994). Neuropsychological features of memory disorders in Alzheimer's disease. In R. D. Terry, R. Katzman, & K. L. Bick (Eds.), *Alzheimer's disease* (pp. 41–63). New York: Raven Press.

Borkowski, J. G., Benton, A. L., & Spreen, O. (1967). Word fluency and brain damage. *Neuropsychologia, 5,* 135–140.

Brandt, J. (1988). Malingered amnesia. In R. Rogers (Ed.), *Clinical assessment of malingering and deception.* New York: Guilford Press.

Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new verbal memory test with six equivalent forms. *The Clinical Neuropsychologist, 5,* 125–142.

Brooker, B. H., & Cyr, J. J. (1986). Tables for clinicians to use to convert WAIS-R short forms. *Journal of Clinical Psychology, 42,* 982–985.

Butters, N. (1985). Alcoholic Korsakoff's syndrome: An update. *Seminars in Neurology, 4,* 226–244.

Butters, N., Grant, I., Haxby, J., Judd, L. L., Martin, A., McClelland, J., Pequegnat, W., Schacter, D., & Stover, E. (1990). Assessment of Aids-related cognitive changes: Recommendations of the NIMH Workshop on neuropsychological assessment approaches. *Journal of Clinical and Experimental Neuropsychology, 12,* 963–978.

Cermak, L. S. (1994). *Neuropsychological explorations of memory and cognition.* New York: Plenum.

Chan, A. S., Butters, N., Salmon, D. P., Johnson, S. A., Paulsen, J. S., & Swenson, M. R. (1995). Comparison of the semantic networks in patients with dementia and amnesia. *Neuropsychology, 9,* 177–186.

Chelune, G. J. (1995). Hippocampal adequacy versus functional reserve: Predicting memory functions following temporal lobectomy. *Archives of Clinical Neuropsychology, 10,* 413–432.

Chelune, G. J., & Bornstein, R. A. (1988). WMS-R patterns among patients with unilateral brain lesions. *The Clinical Neuropsychologist, 2,* 121–132.

Christensen, A.-L. (1984). The Luria method of examination of the brain-impaired patient. In P. E. Logue & J. M. Schear (Eds.), *Clinical neuropsychology: A multidisciplinary approach.* Springfield, IL: C.C. Thomas.

Conners, C. K. (1995). *Conners' continuous performance test.* New York: Multi-Health Systems.

Corwin, J., & Bylsma, F. W. (1993). Translations of excerpts from Andre Rey's *Psychological examination of traumatic encephalopathy* and P.A. Osterrieth's *The Complex Figure Copy Test. The Clinical Neuropsychologist, 7,* 3–21.

Crum, R. M., Anthony, J. C., Bassett, S. S., & Folstein, M. F. (1993). Population-based norms for the Mini-Mental State Examination by age and educational level. *Journal of the American Medical Association, 269,* 2386–2391.

Cullum, C. M., Filley, C. M., & Kozora, E. (1995). Episodic memory function in advanced aging and early Alzheimer's Disease. *Journal of The International Neuropsychological Society, 1,* 100–103.

Cullum, C. M., Heaton, R. K., & Grant, I. (1991). Psychogenic factors influencing neuropsychological perforance: Somatoform disorders, factitious disorders, and malingering. In H. O. Doerr & A. Carlin (Eds.), *Forensic neuropsychology* (pp. 141–171). New York: Guilford Press.

Cullum, C. M., & Rosenberg, R. N. (1998). Memory loss—when is it Alzheimer disease? *Journal of the American Medical Association, 279,* 1689–1690.

Cullum, C. M., Thompson, L. L., & Smernoff, E. N. (1993). Three word recall as a measure of memory. *Journal of Clinical and Experimental Neuropsychology, 15,* 321–329.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California verbal learning test. Research edition.* San Antonio, TX: The Psychological Corporation.

Delis, D. C., Massman, P. J., Kaplan, E., McKee, R., Kramer, J. H., & Gettman, D. (1991). Alternate form of the California Verbal Learning Test: Development and reliability. *The Clinical Neuropsychologist, 5,* 154–162.

Delis, D. C., Squire, L. R., Bihrle, A., & Massman, P. (1992). Componential analysis of problem-solving ability: Performance of patients with frontal lobe damage and amnesic patients on a new sorting test. *Neuropsychologia, 30,* 683–697.

Dodrill, C. B. (1997). Myths of neuropsychology. *The Clinical Neuropsychologist, 11,* 1–17.

Eubanks, J. D. (1997). Clinical neuropsychology summary information prepared by Division 40, Clinical Neuropsychology, American Psychological Association. *The Clinical Neuropsychologist, 11,* 77–80.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state." *Journal of Psychiatric Research, 12,* 189–198.

Franzen, M. D., Burgess, E. J., & Smith-Seemiller, L. (1998). Methods of estimating premorbid functioning. *Archives of Clinical Neuropsychology, 12,* 711–738.

Franzen, M. D., Iverson, G. L., & McCracken, L. M. (1990). The detection of malingering in neuropsychological assessment. *Neuropsychology Review, 1,* 247–279.

Freedman, M., Leach, L., & Kaplan, E. (1994). *Clock drawing: A neuropsychological analysis.* New York: Oxford University Press.

Gauthier, L., Dehaut, F., & Joanette, Y. (1989). The Bells Test: A quantitative and qualitative test for visual neglect. *International Journal of Clinical Neuropsychology, 11,* 49–54.

Goodglass, H., & Kaplan, E. (1987). *The assessment of aphasia and related disorders* (2nd ed.). Philadelphia: Lea & Febiger.

Grant, I., & Adams, K. M. (Eds.) (1996). *Neuropsychological Assessment of neuropsychiatric disorders* (2nd ed.). New York: Oxford University Press.

Greve, K. W., Farrwell, J. F., Besson, P. S., & Crouch, J. A. (1995). A psychometric analysis of the California Card Sorting Test. *Archives of Clinical Neuropsychology, 10,* 265–278.

Guilmette, T., Hart, K., & Giuliano, A. (1993). Malingering detection: The use of a forced choice method in identifying organic versus simulated memory impairment. *The Clinical Neuropsychologist, 7,* 59–69.

Halstead, W. C. (1947). *Brain and intelligence.* Chicago: University of Chicago Press.

Hannay, H. J., Bieliauskas, L. A., Crosson, B. A., Hammeke, T. A., Hamsher, K. De-S., & Koffler, S. P. (1998). Proceedings of the Houston Conference on Specialty Education and Training in Clinical Neuropsychology. *Archives of Clinical Neuropsychology, 13,* 157–250.

Heaton, R. K. (1981). *Wisconsin Card Sorting Test (WCST)*. Odessa, FL: Psychological Assessment Resources.

Heaton, R. K. (1992). *Comprehensive norms for an expanded Halstead–Reitan battery: A supplement for the Wechsler Memory Scale-Revised*. Odessa, FL: Psychological Assessment Resources.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtis, G. (1993). *Wisconsin Card Sorting Test (WCST) Manual Revised and expanded*. Odessa, FL: Psychological Assessment Resources.

Heaton, R. K., Grant, I., & Matthews, C. G. (1991). *Comprehensive norms for an expanded Halstead–Reitan battery: Demographic corrections, research findings, and clinical applications*. Odessa, FL: Psychological Assessment Resources.

Heaton, R. K., Ryan, L., Grant, I., & Matthews, C. G. (1996). Demographic influences on neuropsychological test performance. In K. M. Adams & I. Grant (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (pp. 141–163). New York: Oxford University Press.

Hermann, B. P., Wyler, A. R., Richey, E. T., & Rea, J. M. (1987). Memory function and verbal learning ability in patients with complex partical seizures of temporal lobe origin. *Epilepsia, 28,* 547–554.

Hiscock, M., & Hiscock, C. K. (1989). Refining the forced-choice method for the detection of malingering. *Journal of Clinical and Experimental Neuropsychology, 11,* 967–974.

Ivnik, R. J., Malec, J. F., Smith, G. E., Tangalos, E. G., Petersen, R. C., Kokman, E., & Kurland, L.T. (1992). Mayo's older Americans normative studies: WMS-R norms for ages 56 to 94. *The Clinical Neuropsychologist, 6,* 49–82.

Kaplan, E. (1990). The process approach to neuropsychological assessment of psychiatric patients. *Journal of Neuropsychiatry, 2,* 72–87.

Kaplan, E., Fein, D., Morris, R., & Delis, D. C. (1991). *WAIS-R as a neuropsychological instrument*. San Antonio, TX: The Psychological Corporation.

Kaplan, E. F., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test* (2nd ed.). Philadelphia: Lea & Febiger.

Kertesz, A. (1982). *Western Aphasia Battery*. San Antonio, TX: The Psychological Corporation.

Kløve, H. (1963). *Grooved pegboard*. Lafayette, IN: Lafayette Instruments.

Kozora, E., & Cullum, C.M. (1994). Qualitative features of clock drawings in normal aging and Alzheimer's disease. *Assessment, 1,* 179–187.

Kozora, E., & Cullum, C. M. (1995). Generative naming in normal aging: Total output and qualitative changes using phonemic and semantic constraints. *The Clinical Neuropsychologist, 9,* 313–320.

Lacritz, L. H., & Cullum, C. M. (1998). Stroke and depression. In A. J. Rush (Ed.), *Mood & anxiety disorders* (pp. 339–362). Philadelphia: Williams and Wilkins.

Lansing, A. E., Ivnik, R. J., Cullum, C. M., & Randolph, C. (in press). An empirically derived short form of the Boston Naming Test. *Archives of Clinical Neuropsychology*.

Lewis, R., & Kupke, T. (1977). *The Lafayette Clinic Repeatable Neuropsychological Test Battery: Its development and research applications*. Paper presented at the annual meeting of the Southeastern Psychological Association, Hollywood, Florida.

Lewis, R. F., & Rennick, P. M. (1979). *Manual for the Repeatable Cognitive–Perceptual–Motor Battery*. Clinton Township, MI: Ronald F. Lewis.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Libon, D. J., Malamut, B. L., Swenson, R., Sands, L. P., & Cloud, B. S. (1996). Further analyses of clock drawings among demented and nondemented older subjects. *Archives of Clinical Neuropsychology, 11,* 193–205.

Libon, D. J., Mattson, R. E., Glosser, G., Kaplan, E., Malamut, B. L., Sands, L. P., Swenson, R., & Cloud, B. S. (1996). A nine-word dementia version of the California Verbal Learning Test. *The Clinical Neuropsychologist, 10,* 237–244.

Loring, D.W., Lee, G.P., Martin, R.C., & Meador, K.J. (1989). Verbal and visual memory index discrepancies from the Wechsler Memory Scale-Revised: Cautions in interpretation. *Psychological Assessment, 1,* 198–202.

Loring, D. W., Martin, R. C., Meador, K. J., & Lee, G. P. (1990). Psychometric construction of the Rey–Osterrieth complex figure: Methodological considerations and interrater reliability. *Archives of Clinical Neuropsychology, 5,* 1–14.

Luria, A. R. (1966). *The working brain: An introduction to neuropsychology*. New York: Basic Books.

Malloy, P. F., & Richardson, E. D. (1998). Assessment of frontal lobe functions. *Journal of Neuropsychiatry, 6,* 399–410.

Markwardt, F. C., Jr. (1989). *The Peabody Individual Achievement Test-Revised.* Circle Pines, MN: American Guidance Service.

Matarazzo, J. D. (1990). Psychological assessment versus psychological testing: Validation from Binet to the school, clinic, and courtroom. *American Psychologist, 45,* 999–1017.

Matarazzo, J. D., & Prifitera, A. (1989). Subtest scatter and premorbid intelligence: Lessons from the WAIS-R standardization sample. *Psychological Assessment, 1,* 186–191.

Mattis, S. (1988). *Dementia Rating Scale (DRS)*. Odessa, FL: Psychological Assessment Resources.

Matt-Maddrey, A., Cullum, C. M., Weiner, M. F., & Filley, C. M. (1996). Premorbid intelligence estimation and level of dementia in Alzheimer's disease. *Journal of The International Neuropsychological Society, 2,* 1–5.

Meyers, J. E., & Meyers, K. R. (1995). Rey complex figure test under four different administration procedures. *The Clinical Neuropschologist, 9,* 63–67.

Milberg, W. P., Hebben, N., & Kaplan, E. (1996). The Boston process approach to neuropsychological assessment. In I. Grant & K. M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (2nd ed., pp. 58–80). New York: Oxford University Press.

Mirsky, A. F., Anthony, B. J., Duncan, C. C., Ahearn, M. B., & Kellam, S. G. (1991). Analysis of the elements of attention: A neuropsychological approach. *Neuropsychology Review, 2,* 109–145.

Morris, J. C., Heyman, A., Mohs, R. C., et al. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology, 39,* 1159–1165.

Moscovitch, M. (1994). Cognitive resources and dual-task interference effects at retrieval in normal people: The role of the frontal lobes and medial temporal cortex. *Neuropsychology, 8,* 524–534.

Naugle, R. I., Chelune, G. J., Cheek, R., Luders, H., & Awad, I. A. (1993). Detection of changes in material-specific memory following temporal lobectomy using the Wechsler Memory Scale-Revised. *Archives of Clinical Neuropsychology, 8,* 381–395.

Naugle, R. I., Chelune, G. J., Schuster, J., Luders, H. O., & Comair, Y. (1994). Recognition memory for words and faces before and after temporal lobectomy. *Assessment, 1,* 373–381.

Naugle, R. I., Cullum, C. M., & Bigler, E. D. (1998). *Introduction to clinical neuropsychology: A casebook*. Austin, TX: ProEd.

Naugle, R. I., Cullum, C. M., Bigler, E. D., & Massman, P. J. (1986). Neuropsychological and computerized axial tomography volume characteristics in senile and presenile dementia. *Archives of Clinical Neuropsyhology, 1,* 219–230.

Nelson, H. E. (1982). *The National Adult Reading Test (NART): Test Manual.* Windsor, UK: NFER-Nelson.

Nelson, H. E., & O'Connell, A. (1978). Dementia: The estimation of premorbid intelligence levels using the National Adult Reading Test. *Cortex, 14,* 234–244.

Nies, K. J., & Sweet, J. J. (1994). Neuropsychological assessment and malingering: A critical review of past and present strategies. *Archives of Clinical Neuropsychology, 9,* 501–552.

Osterrieth, P. A. (1944). Le test de copie d'une figure complexe. *Archives de Psychologie, 30,* 206–356; translated by J. Corwin & F.W. Bylsma (1993), *The Clinical Neuropsychologist, 7,* 9–15.

Paolo, A. M., Tröster, A. I., & Ryan, J. J. (1997). Test-retest stability of the California Verbal Learning Test in older persons. *Neuropsychology, 11,* 613–616.

Peavy, G., Jacobs, D., Salmon, D., Butters, N., Delis, D., Taylor, M., Massman, P, Stout, J., Heindel, W., Kirson, E., Atkinson, J., Chandler, J., Grant, I., and the HIV Neurobehavioral Research Center Group (1994). Verbal memory performance of patients with human immunodeficiency infection: Evidence of subcortical dysfunction. *Journal of Clninical and Exeperimental Neuropsychology, 16,* 508–523.

Powell, D. H., Kaplan, E. F., Whitla, D., Weintraub, S., Catlin, R., & Funkenstein, H. H. (1993). *Manual for MicroCog: Assessment of cognitive functioning*. San Antonio, TX: The Psychological Corporation.

Rapport, L. J., Dutra, R. L., Webster, J. S., Charter, R., & Morrill, B. (1995). Hemispatial deficits on the Rey–Osterrieth complex figure drawing. *The Clinical Neuropsychologist, 9,* 169–179.

Raven, J. C. (1994). *Advanced progressive matrices sets I & II.* Oxford, UK: Oxford Psychologists Press.

Raven, J. C. (1995). *Colored progressive matrices sets A, Ab, B.* Oxford, UK: Oxford Psychologists Press.

Raven, J. C. (1996). *Progressive Matrices: A perceptual test of intelligence. Individual Form.* Oxford, UK: Oxford Psychologists Press.

Reitan, R. M., & Wolfson, D. (1993). *The Halstead–Reitan Neuropsychological Test Battery: Theory and clinical interpretation.* Tucson, AZ: Neuropsychology Press.

Reitan, R. M., & Wolfson, D. (1996). Relationships of age and education to Wechsler Adult Intelligence Scale IQ values in brain-damaged and non-brain-damaged groups. *The Clinical Neuropsychologist, 10,* 293–304.

Reports of the INS-Division 40 Task Force on Education, Accreditation, and Credentialing (1987). *The Clinical Neuropsychologist, 1,* 29–34.

Reports of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology (1996). Assessment: Neuropsychological testing of adults. *Neurology, 47,* 592–599.

Rey, A. (1941). L'examen psychologique dans les cas d'encéphalopathie traumatique. *Archives de Psychologie, 28,* 286–340.

Reynolds, C. R. (1997). Forward and backward memory span should not be combined for clinical analysis. *Archives of Clinical Neuropsychology, 12,* 29–40.

Ross, E. D. (1985). Modulation of affect and nonverbal communication by the right hemisphere. In M. M. Mesulam (Ed.), *Principles of behavioral neurology* (pp. 239–257). Philadelphia: F.A. Davis.

Royall, D. R., Cordes, J. A., & Polk, M. (1998). CLOX: An executive clock drawing task. *Journal of Neurology, Neurosurgery, and Psychiatry, 64,* 588–594.

Russell, E. W. (1975). A multiple-scoring method for the assessment of complex memory functions. *Journal of Consulting and Clinical Psychology, 43,* 800–809.

Russell, E.W., Neuringer, C., & Goldstein, G. (1970). *Assessment of brain damage: A neuropsychological key approach.* New York: Wiley.

Ryan, C. M., Morrow, L. A., Bromet, E., & Parkinson, D. K. (1987). Assessment of neuropsychological dysfunction in the workplace: Normative data from the Pittsburgh occupational exposures test battery. *Journal of Clinical and Experimental Neuropsychology, 9,* 665–679.

Sass, K. J., Sass, A., Westerveld, M., Lencz, T., Novelly, R. A., Kim, J. H., & Spencer, D. D. (1992). Specificity in the correlation of verbal memory and hippocampal neuron loss: Dissociation of memory, language, and verbal intellectual ability. *Journal of Clinical and Experimental Neuropsychology, 14,* 662–672.

Schacter, D. L., & Tulving, E. (1994). What are the memory systems of 1994? In D. L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 1–38). Cambridge, MA: MIT Press.

Silverstein, A. B. (1982). Two-and four-subtest short forms of the Wechsler Adult Intelligence Scale-Revised. *Journal of Consulting and Clinical Psychology, 50,* 415–418.

Smith, G. E., Ivnik, R. J., Malec, J. F., Kokmen, Tangalos, E., & Petersen, R. C. (1994). Psychometric properties of the Mattis Dementia Rating Scale. *Assessment, 1,* 123–131.

Spreen, O., & Strauss, E. (1998). *Compendium of neuropsychological tests* (2nd ed.). New York: Oxford University Press.

Squire, L. R. (1987). *Memory and brain.* New York: Oxford University Press.

Stern, R. A., Singer, E. A., Duke, L. M., Singer, N. G., Morey, C. E., Daughtrey, E. W., & Kaplan, E. (1994). The Boston qualitative scoring system for the Rey–Osterrieth Complex Figure: Description and interrater reliability. *The Clinical Neuropsychologist, 8,* 309–322.

Sternberg, R. J. (Ed.) (1997). Intelligence and lifelong learning [Special issue]. *American Psychologist, 52.*

Strub, R. L., & Black, F. W. (1988). *Neurobehavioral disorders: A clinical approach.* Philadelphia: F.A. Davis.

Stuss, D. T., & Benson, D. F. (1984). Neuropsychological studies of the frontal lobes. *Psychological Bulletin, 95,* 3–28.

Sweet, J. J., & Moberg, P. J. (1990). A survey of practices and beliefs among ABPP and non-ABPP clinical neuropsychologists. *The Clinical Neuropsychologist, 4,* 101–120.

Tombaugh, T. N., McDowell, I., Krisjansson, B., & Hubley, A. M. (1996). Mini-Mental State Examination (MMSE) and the Modified MMSE (3MS): A psychometric comparison and normative data. *Psychological Assessment, 8,* 48–59.

Trenerry, M. R., Crosson, B., DeBoe, J., & Leber, W. R. (1990). *Visual Search and Attention Test.* Odessa, FL: Psychological Assessment Resources.

Tröster, A. I., Jacobs, D., Butters, N., Cullum, C. M., & Salmon, D. P. (1989). Differentiating Alzheimer's disease from huntington's disease with the Wechsler Memory Scale-Revised. *Clinics in Geriatric Medicine, 5,* 611–632.

Warrington, E. K. (1984). *Recognition Memory Test.* Windsor, UK: NFER-Nelson.

Wechsler, D. (1945). A standardized memory scale for clinical use. *The Journal of Psychology, 19,* 87–95.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1987). *Wechsler Memory Scale-Revised.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale-Third Edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale-Third Edition*. San Antonio, TX: The Psychological Corporation.

Welsh, K. A., Butters, N., Mohs, R. C., Beekly, D., Edland, S., Fillenbaum, G., & Heyman, A. (1994). The consortium to establish a registry for Alzheimer's disease (CERAD). Part V. A normative study of the neuropsychological battery. *Neurology, 44,* 609–614.

Wilkinson, G. S. (1993). *The Wide Range Achievement Test-3 manual*. Wilmington, DE: Wide Range, Inc.

Woodcock, R. W., & Mather, N. (1989). *Woodcock–Johnson Tests of Achievement*. Allen, TX: DLM Teaching Resources.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.12

Principles of Personality Assessment

JERRY S. WIGGINS and KRISTA K. TROBST
University of British Columbia, Vancouver, BC, Canada

4.12.1 INTRODUCTION 350

4.12.2 PSYCHODYNAMIC TRADITION 350

- *4.12.2.1 Historical Background* 350
- *4.12.2.2 Conceptual Framework* 351
- *4.12.2.3 Assessment Instruments* 351
- *4.12.2.4 Interpretive Principles* 351
  - *4.12.2.4.1 Projective hypothesis* 352
  - *4.12.2.4.2 Levels of functioning* 352
  - *4.12.2.4.3 Psychological adjustment* 352
- *4.12.2.5 Applications and Current Status* 352

4.12.3 PERSONOLOGICAL TRADITION 353

- *4.12.3.1 Historical Background* 353
- *4.12.3.2 Conceptual Framework* 353
- *4.12.3.3 Assessment Instruments* 354
- *4.12.3.4 Interpretive Principles* 354
- *4.12.3.5 Applications and Current Status* 355

4.12.4 MULTIVARIATE TRADITION 355

- *4.12.4.1 Historical Background* 355
- *4.12.4.2 Conceptual Framework* 356
- *4.12.4.3 Assessment Instruments* 356
- *4.12.4.4 Interpretive Principles* 356
- *4.12.4.5 Applications and Current Status* 357

4.12.5 EMPIRICAL TRADITION 357

- *4.12.5.1 Historical Background* 357
- *4.12.5.2 Conceptual Framework* 358
- *4.12.5.3 Assessment Instruments* 359
- *4.12.5.4 Interpretive Principles* 359
  - *4.12.5.4.1 Empirical strategy* 359
  - *4.12.5.4.2 Constructive strategy* 359
  - *4.12.5.4.3 Interpersonal strategy* 359
- *4.12.5.5 Applications and Current Status* 360

4.12.6 INTERPERSONAL PARADIGM 360

- *4.12.6.1 Historical Background* 360
- *4.12.6.2 Conceptual Framework* 361
- *4.12.6.3 Assessment Instruments* 361
  - *4.12.6.3.1 Interpersonal Adjective Scales* 361
  - *4.12.6.3.2 Inventory of Interpersonal Problems* 361
  - *4.12.6.3.3 Impact Message Inventory* 362
- *4.12.6.4 Interpretive Principles* 362
- *4.12.6.5 Applications and Current Status* 363

4.12.7 CONCLUSION 364
4.12.8 REFERENCE 365

4.12.1 INTRODUCTION

Consider a hypothetical clinician who is both well trained and well informed about available personality tests. In her daily practice, she employs a range of assessment instruments (singly or in combination) to address the range of diagnostic issues posed by her clients. When "thought disorder" is suspected, she might occasionally administer a Rorschach and from most clients she will obtain a case history. Her clients themselves may be asked to fill out the NEO-Personality Inventory to clarify dimensions of normal personality functioning or the Minnesota Multiphasic Personality Inventory-Revised (MMPI-2) to assess gross abnormality. She might also routinely administer the Interpersonal Adjective Scales to assess the character and quality of clients' interpersonal relationships.

Like many of her colleagues, our clinician may describe her theoretical orientation as "eclectic" and indeed that is what she appears to be. But she is also, perhaps unknowingly, operating within larger historical contexts of assumptions and beliefs that may include traditions of thought as diverse as psychoanalysis, psychobiography, psychometrics, empirical realism, and symbolic interactionism (respectively for the tests just mentioned). We will refer to such contexts as different "traditions" of personality assessment, by which we mean the background of a set of unquestioned beliefs or orienting attitudes within and against which personality tests are administered and interpreted.

The traditions that will be considered in this chapter are referred to as psychodynamic, personological, multivariate, empirical, and interpersonal. Such traditions are related, but *not* equivalent, to certain theories of personality (Hall & Lindzey, 1978) and psychopathology (Millon, 1969). Similarly, these traditions are related, but *not* equivalent, to certain kinds of personality tests. Thus, the MMPI has been associated with a distinctive philosophy of science (empirical realism) rather than with a particular theory of psychopathology. The Rorschach test and the Wechsler scales have figured prominently within the ego psychological perspective of the psychodynamic tradition, although both tests were constructed for quite different purposes.

We believe that these traditions reflect different and potentially useful ways of viewing the entire psychodiagnostic enterprise. We also believe that these traditions have contributed to more or less invisible barriers to communication being erected among clinicians who work, knowingly or not, within them. This state of affairs is unfortunate because "membership" in a tradition, as it were, is mainly determined by where one went to graduate school, with whom one studied, and the particular clinical setting in which one finds oneself operating. Further, we are convinced that, at the present time, there are some remarkably similar ideas being advocated about the nature of personality and its assessment among those working within several different traditions, and that recognition of these commonalties might advance the field as a whole (see Wiggins, in press).

We would venture to suggest that most readers will have familiarity with a few of these traditions, but not with some of the others. For example, those intimately familiar with the psychodynamic tradition are less likely to be familiar with the multivariate and empirical traditions, and vice versa. Those who are fond of obtaining life history data from interviews may not be fully aware of the rich history of the personological tradition nor of the exciting recent developments that have occurred within that tradition. And many may not be aware that the interpersonal tradition has fostered a strong "underground movement" within clinical psychology for many decades.

4.12.2 PSYCHODYNAMIC TRADITION

4.12.2.1 Historical Background

David Rapaport (1911–1960) was a major systematist of psychoanalytic theory, the originator of a now standard psychodiagnostic test battery, and a highly influential mentor of the principal architects of the psychodynamic tradition in personality assessment. After receiving his PhD from the Royal Hungarian University, he emigrated to the USA in 1938 and shortly thereafter joined the staff of the Menninger Clinic where he eventually became chief psychologist and head of the research department. During and shortly after World War II, the results of an extensive program of collaborative research were summarized in a two-volume *Manual of diagnostic psychological testing* (Rapaport, 1944–46) that eventually "revolutionized clinical psychology and influenced clinical psychologists the world over" (Gill & Klein, 1967, p. 18).

Rapaport moved to the Austin Riggs Center in 1948 and continued his extensive collaborations with colleagues at institutions such as the Menninger Foundation, Yale University Department of Psychiatry, and the Research Center for Mental Health at New York University. Speaking collectively for workers at these and other institutions, Roy Schafer (1967) observed, "All of us are working within the psychoanalytic psychodiagnostic tradition crystallized by David Rapaport" (p. 2).

4.12.2.2 Conceptual Framework

It is difficult to make a coherent set of distinctions among the many "psychodynamic" theories that have been proposed as alternatives to Freud's classical psychoanalytic theory (Westen, 1990). It would seem to be generally true however that since, and even before, Freud's death, there has been an increased emphasis on the autonomous and adaptive functions of the ego and on the internal representation of self and others within the ego domain. The aspect of classical psychoanalysis that has been de-emphasized, or even abandoned by some, is the proposition that personality and its development are largely determined by unconscious libidinal and aggressive drives.

Rapaport (1960) must be counted among the most rigorous systematists of classical psychoanalytic theory and he was also a major contributor to the alternative perspective known as "ego psychology." In both his theoretical and assessment work, his principal interest was in thought processes and their course of development from drive-oriented to reality-oriented processes. In this context, Hartmann's (1939) concept of the independence (from drive) of the ego and its role in fostering active processes of adaptation were appealing to him, as was Erikson's (1950) emphasis on the cultural and historical contexts in which the ego develops. Rapaport (1960) developed and expanded the metaconcept of structure which he felt was the "prime requisite for progress toward dimensional quantification" (p. 98) and which he applied broadly to many aspects of psychological functioning and cognitive organization (see Gill & Klein, 1967).

The generally increased emphasis upon ego functions in postclassical psychoanalysis was also reflected in the currently important psychodynamic alternative of object relations theory. Proponents of this view challenged the classical idea that cathexes of external "objects" (including persons) served mainly as vehicles through which instinctual energies were discharged. Instead, it was postulated that early interactions with significant others ("objects") led to internalized representations of both others (Jacobson, 1954) and self (Kohut, 1971) that serve as "internal working models" (Bowlby, 1973) for later interpersonal relationships. From this perspective, the person is, from birth (Main, Kaplan, & Cassidy, 1985), an object seeker (Fairbain, 1952) who establishes a "gratifying involvement" with other persons (Behrends & Blatt, 1985).

The shift in emphasis within psychoanalytic theory from drives to internalized representations has been accompanied by a corresponding shift in rationales for personality assessment, as exemplified by the ego-psychological approach of Allison, Blatt, and Zimet (1968) and the object-relations approach of Blatt and Lerner (1983).

4.12.2.3 Assessment Instruments

Prior to Rapaport's influential contributions in the 1940s, assessment psychologists were primarily technicians who administered IQ tests. Since that time, they have become clinicians who administer batteries of both structured and projective tests of personality and cognition. The multitest battery was advocated in view of the apparent complexity of personality and cognition and their interrelated functions, as well as for the purpose of gathering normative data that would shed light on those complexities (Rapaport, Gill, & Schafer, 1946). At Menninger, the composition of the battery reflected judgments regarding the potential of each instrument to yield measures that might be interpreted within the ego psychological framework of Rapaport and his associates.

The projective component of the original Menninger battery included the Rorschach Test (Rorschach, 1921), the Thematic Apperception Test (TAT; Morgan & Murray, 1935), and a locally constructed Word Association Test. The nonprojective component included the Bellevue Scale (Wechsler, 1941), the Babcock Test of mental efficiency (Babcock, 1933), the Sorting Test of concept formation (Goldstein & Scheerer, 1941), and the Hanfmann–Kasanin test of concept formation (Hanfmann & Kasanin, 1937). As this test battery evolved over a 20-year period, the Rorschach, TAT, and WAIS (Wechsler, 1958) became the more or less standard core of the psychodynamic test battery (e.g., Allison et al., 1968).

4.12.2.4 Interpretive Principles

The vast literature on interpretive principles associated with the psychodynamic paradigm defies easy summarization. Rapaport's theoretical work (e.g., Gill, 1967) and the writings of those whom he influenced (e.g., Holt, 1967) are, understandably, highly technical in nature. Of greater help to the clinician are the case summaries (Allison et al., 1968; Prelinger & Zimet, 1964; Rapaport et al., 1946; Schafer, 1948) and the individual treatments of the Rorschach (Schafer, 1954), TAT (Holt, 1951), and WAIS (Blatt & Allison, 1981). The following statements regarding the projective hypothesis, levels of functioning, and psychological adjustment are meant only to convey some of the flavor of three, among many, important interpretive principles.

4.12.2.4.1 Projective hypothesis

The projective hypothesis states that "All behavior manifestations of the human being, including the least and the most significant, are revealing and expressive of his personality, by which we mean that individual principle of which he is the carrier" (Rapaport, 1942, p. 92). Thus, people's possessions—clothes, automobiles, furniture—are expressive of their personalities and reflect single acts of choice; in their totality, they reflect the organization of such choices. Responses to the ambiguous stimuli of projective tests may also be thought of in terms of "choice," although such choices are much less conscious or volitional in nature. Thus, responses to a Rorschach inkblot may be thought of as reflecting "choices" between forms, colors, shadings, and so forth, to which a subject imparts meaning through organization. Responses to a TAT card also involve both choice (e.g., which figure to identify with) and organization (e.g., sequence of events). Responses to intelligence and concept formation tests involving choice and organizational processes may be used as "nonprojective tests of personality," given an adequate theory of "functions underlying the reactions and achievements on these tests" (Rapaport, 1946, p. 228).

4.12.2.4.2 Levels of functioning

The psychoanalytic model of primary (pleasure principle) and secondary (reality principle) modes of thought is meant to account for both a developmental sequence and characteristics of the mature adult (Rapaport, 1951). Consequently, there is a continuum of adult psychological functioning that may be assessed with an appropriate battery of tests.

> This continuum ranges from functioning in situations which put a premium on highly logical, reality-oriented secondary modes of thought (WAIS) to those which allow for more personal, less conventionally constrained thinking (TAT) and finally those which allow for considerably novel, personalized, and regressive modes of thinking (Rorschach). (Allison et al., 1968, p. vii)

Assessment of level of functioning has been greatly facilitated by Holt's innovative procedures for assessing primary and secondary process in the Rorschach (Holt & Havel, 1960).

4.12.2.4.3 Psychological adjustment

The Rorschach, TAT, and WAIS may be employed to assess both adaptive capacities and impairments in psychological functioning and to identify the functions impaired in different psychiatric diagnostic groups. Adjustment and maladjustment may be assessed with reference to the following postulated sequence:

> Certain patterns of defense mechanisms are adopted and these determine specific strengths and weaknesses in psychological functioning which then become characteristic of the adjustment of the personality; with the onset of maladjustment, an exaggeration or breakdown in these strengths and weaknesses characteristic for that maladjustment occurs which can be measured; this leads to a diagnostic differentiation. (Rapaport, Menninger, & Schafer 1947, p. 249)

4.12.2.5 Applications and Current Status

Since the late 1960s, theory and method within the psychodynamic paradigm of personality assessment have evolved into an object relations perspective that is highly compatible with contemporary formulations of social cognition, information processing, attachment research, and ego development (see Westen, 1990). Sidney Blatt and his associates at Yale have been primarily responsible for this paradigm shift. From the early 1950s until his untimely death in 1960, Rapaport contributed to the development of ego psychology and its application to personality assessment; since the early 1970s, Blatt has contributed to the development of object relations theory and its application to current personality assessment methods.

"The study of the representation of the human form on the Rorschach is an ideal data base for assessing an individual's representational world—his conception of people, including himself, and their actual and potential interactions. The representation of people, that is, object representations, have both structure and content" (Blatt & Lerner, 1983, p. 8). The structural aspects of object representations are emphasized in the Rorschach scoring system developed by Blatt, Brenneis, Schimek, and Glick (1976) that provides a developmental analysis of object representations in terms of such categories as differentiation, articulation, and integration of object and action. The content and affective themes of object representations are emphasized in the Rorschach scoring system developed by Mayman (1967) that emphasizes phenomenological dimensions such as affect states, ego states, experience of self, and sense of identity.

An object relations scoring system for the TAT has also been developed by Westen (1991) in which stories are rated for complexity of representations of people, affective tone of relationship paradigms, capacity for emotional investment, and understanding of social causality. A similar scoring system has been developed for the Picture Arrangement subtest of the WAIS (Westen, 1991, p. 72). In conclusion, the three major instruments of the original Menninger battery continue to show promise under a revised conceptual orientation that is most compatible with current thinking in personality, social, clinical, and developmental psychology.

4.12.3 PERSONOLOGICAL TRADITION

4.12.3.1 Historical Background

The "storied" nature of human conduct (Sarbin, 1986) has been recognized for many centuries within the disciplines of history, literature, and most of the social sciences. Within the much more recent personological tradition in personality assessment, the person is taken to be the basic unit of observation and the person's life story is considered to be the preferred way in which the person is to be understood (McAdams, 1994). The origins of the personological tradition, and in fact the origins of academic personality psychology itself, can be traced most directly to Harvard University in the 1930s. At that institution, Gordon Allport (1937) wrote the first textbook on personality, which defined the person as the basic unit of observation. Subsequently, he emphasized the use of personal documents in psychobiography (Allport, 1942, 1965) and spent a lifetime pondering the question of "How shall a psychological life history be written?" (Allport, 1967). At the Harvard Psychological Clinic, Murray (1938) introduced a multiform organismic method for studying "lives in progress" (White, 1952), in which an interdisciplinary team of investigators applied a variety of assessment procedures to the study of individuals over the course of their lives. During World War II, a variant of this procedure was employed in the selection of intelligence agents and saboteurs (Office of Strategic Services, 1948), and subsequently this approach spawned a variety of successful peacetime assessment programs (e.g., MacKinnon, 1975; Stern, Stein, & Bloom, 1956).

Murray's enduring contributions to the personological paradigm go well beyond those just mentioned. During his 60 years at Harvard, Murray was a source of personal inspiration for an extraordinarily talented and diverse group of students, colleagues, and visitors who, over the years, have applied and expanded his "personology" up until the present time. Early associates at the clinic included Samuel Beck, Erik Erikson, Jerome Frank, Daniel Levenson, Donald MacKinnon, Silvan Tomkins, and Robert White. Dan McAdams and William McKinley Runyan are more recent products of this Harvard tradition.

4.12.3.2 Conceptual Framework

Murray's (1938) personology provided a set of orienting attitudes that guided the multiform (many assessors, many tests), organismic (holistic) approach to the study of lives over time (e.g., White, 1966, 1975). Murray's (1959) conceptual framework was meant to serve as "the scaffold of a comprehensive system," rather than as a completed formal theory. Nevertheless, Murray (1938) introduced a number of concepts that have subsequently proven useful in guiding the study of lives. By taking the life cycle of the individual as the largest unit of study, Murray was committed to studying personality "the long way" (White, 1981), and he introduced units of time varying in duration and complexity from the proceeding (single episode) to the unity-thema (central life motif). A thema may be understood in terms of the interaction of needs (for which Murray developed his famous taxonomy) and press (environmental facilitation or obstruction of a need). The TAT (Murray, 1943) was developed as a projective measure of need–press interaction and has been utilized in large-scale studies of needs for achievement (McClelland, 1961), power (Winter, 1973), and intimacy (McAdams, 1989).

Erik Erikson's (1950) psychodynamic theory of personality development provided a more explicit account of the stages of psychosocial development over the entire life span (from the original psychosocial crisis of "trust versus mistrust" to the final crisis of "ego-integrity versus despair"). His theory informed his own classic psychobiographies, such as those of

Martin Luther (Erikson, 1958) and Gandhi (Erikson, 1969), as well as contributing to more recent psychobiographical efforts (e.g., Stewart, Franz, & Layton, 1988). More recently, in a book entitled *The stories we live by: Personal myths and the making of the self*, McAdams (1993) presented a neo-Eriksonian theory of identity development that is enriched by more contemporary concepts such as image, prototype, and script, and that is guided by the metatheoretical concepts of agency and communion (Bakan, 1966). This work has been hailed as "the most original and important new book on personality theory since George Kelly's *Psychology of personal constructs*" (Hogan, 1994, p. 356). The heuristic potential of McAdams' conceptual framework is suggested by the quality and quantity of critical response it has elicited from personality psychologists representing diverse theoretical perspectives (see McAdams, 1996 and peer commentary therein). Although McAdams' framework is more formal, explicit, and testable than earlier efforts, its roots in the formulations of Allport and Murray are easily traceable (McAdams, 1994).

4.12.3.3 Assessment Instruments

In what is now widely recognized as the definitive general treatise on life histories and psychobiography, Runyan (1982) observed:

> It should be clear that there is no single life history method, any more than there is a single personality research method, and that life histories may be studied through phenomenological self-reports, archival research, prospective longitudinal research, and experimental research. (p. 6)

Personality assessment psychologists from Murray to McAdams have recognized the need for methodological pluralism (Craik, 1986) in the study of lives, and for that reason the "boundaries" of the personological paradigm have historically been somewhat more permeable than those of the other paradigms considered in this chapter. Nevertheless, it would be inaccurate to characterize this paradigm as "eclectic," because the diversity of methods employed reflects the different levels of analysis encompassed by the term "life history," rather than the selective use of theories and methods to "explain" a given life. Following Kluckhohn and Murray (1953), Runyan (1982) distinguished three levels of generality in the social sciences: the universal (general laws of human behavior); the group (differences in sex, social class, culture); and the individual (distinctive, distinguishing characteristics of a particular person) and surveyed attempts to establish acceptable criteria for interpretations at each level. Although personologists in general, and clinicians in particular, are primarily concerned with the individual, an awareness of the larger enterprise that Runyan portrays serves to clarify the nature and goals of individual personality assessment. And for those few who might seriously be considering becoming psychobiographers, Alan Elms' (1994) book is essential reading.

The 25 different assessment procedures employed for the study of individuals at the Harvard Clinic (Murray, 1938) were indeed "multiform" in terms of both methods and the assessors who employed them; for example, autobiography (Murray), hypnosis (White), level of aspiration (Frank), dramatic productions (Erikson), and Rorschach (Beck). Perhaps of greatest interest to the contemporary clinician is the case study method in which some form of interview is employed with all the attendant advantages and limitations of retrospective, introspective, and qualitative methods (see Runyan, 1982, chap. 8). The procedures employed by McAdams (1993, chap. 10) provide an excellent example of a theory-driven approach to the interview.

4.12.3.4 Interpretive Principles

The variety of personality research methods for studying an individual's life history is reflected in a similar variety of interpretive principles associated with different methods. In this context, the interpretive principles elucidated by Alexander (1990) would appear to be "mainstream" in the sense that they evolved directly from the framework of Murray and his co-workers (especially Tomkins 1947, 1979) at the Harvard Psychological Clinic.

In analyzing data from directed interviews or autobiographical essays, Alexander (1990) looks for recurring dynamic sequences (scripts, themas, guiding messages) that may be revealed by "letting the data speak" or by "asking the data a question." In the former method, the typically large data set may be reduced by identifying "significant" material according to the following nine criteria of salience: (i) primacy (that which occurs first), (ii) frequency (that which occurs often), (iii) uniqueness (that which is unusual or odd), (iv) negation (that which is denied or disavowed), (v) emphasis (that which is either overemphasized or underemphasized), (vi) omission (that which is missing by normative standards or by implication), (vii) error or distortion (factual errors and Freudian slips), (viii) isolation (that which does not "fit" and non sequiturs), and (ix) incompletion (that which is not finished or lacks closure).

The method of "asking the data a question" is similar to that employed in the analysis of TAT protocols for predetermined categories such as "power motivation" (Winter, 1973). The database of interview material is reduced by selecting every sequence or incident related to the questions posed by the investigator. When applied to interview material, this method retains the "projective" advantage of addressing issues that the storyteller did not consciously intend to describe or deal with, while circumventing problems such as the identification of the "hero" of the narrative (Alexander, 1990).

4.12.3.5 Applications and Current Status

In recent years, there appears to have been a "back to basics" movement in personality assessment research (Wiggins & Pincus, 1992) that has rekindled interest in the fundamental assumptions of several paradigms, including the personological: "Once, again, it is okay to study the 'whole person.' Better, contemporary personologists insist, as did pioneers like Gordon Allport, that such an endeavor is the personologist's *raison d'être*" (McAdams, 1988, p. 1).

In terms of clinical applications, it should be borne in mind that the personological tradition has tended to focus on relatively normal, nonclinical samples and that within this tradition, life histories have been viewed as relatively veridical accounts of what has happened in a person's life (Runyan, 1982) or as imaginative reconstructions of the past (McAdams, 1993). It is also important to recognize that personality assessment within the personological paradigm occurs on a different level than traditional psychodiagnostic assessment (Alexander, 1990). Whereas personological assessment attempts to understand current personality functioning in terms of an individual's own life history, traditional psychodiagnostic assessment is concerned with determining the individual's membership in a group or class of individuals that differs from other groups in terms of psychopathology.

> A major difference between the two approaches is that the results of the former [personological] can easily be directed toward answering the questions intended by the latter [psychodiagnostic]. The reverse is, unfortunately, a low probability event in anything other than a global sense. To designate someone as obsessive or hysteric, or depressive, or schizophrenic will place the focus on particular salient aspects of functioning but say little about the dynamics in that individual leading to that particular form of functioning. (Alexander, 1990, p. 7)

4.12.4 MULTIVARIATE TRADITION

4.12.4.1 Historical Background

Sir Francis Galton's (1888) method for analyzing the "co-relations" between twin pairs was refined by Pearson (1896) and utilized effectively by Spearman (1904) in identifying a "general factor" of intelligence (*g*) that appeared to "underlie" various tests of mental abilities. Factors are statistical abstractions that summarize the relations among test scores in terms of underlying or "latent" variables which may be thought of as operating at several levels. For example, the intelligence quotient (IQ) could be considered to be a general higher-order factor that summarizes relations between verbal and performance major group factors, which in turn, summarize relations among multiple minor group factors (Vernon, 1950).

Raymond B. Cattell, the founding father of the multivariate tradition of personality assessment, was a student of Spearman and was among the first to apply factor-analytic methodology to the study of temperament (e.g., Cattell, 1933). Using cluster-analytic procedures, Cattell (1943) reduced Allport and Odbert's (1936) exhaustive compilation of trait-descriptive terms from Webster's unabridged dictionary to 35 "surface traits" (see John, 1990). He administered the 35 surface-trait variables (e.g., talkative vs. silent) to respondents in a series of peer-rating studies and eventually concluded that no less than 16 "source traits" could account for the interrelations among the 35 surface clusters (Cattell, 1949).

The Guilfords (1936, 1939) were also pioneers in the application of factor analysis to personality data, and they assembled evidence for the four interpretable factors of "shyness," "rhathymia," "depression," and "liking for thinking" which would now be called " extraversion," "conscientiousness," "neuroticism," and "intellect," and which (together with "agreeableness") constitute the currently important five-factor model (FFM) of personality (Digman, 1996). A decade later, Eysenck (1947) factored the intercorrelations among presenting symptoms of psychiatric patients and found two factors (extraversion and neuroticism) that had both theoretical (Jung, 1971) and empirical (MacKinnon, 1944) precedents (see Eysenck, 1990, pp. 95–98). By the 1950s, the systems of Eysenck (1953), Cattell (1957), and Guilford (1959) had become extensive alternative theories of personality structure.

The history of the multivariate tradition is a curious one that has recently been reconstructed by Digman (1996). Briefly, Digman argues that as early as the 1930s, there was evidence for the generality and replicability of a FFM of personality that consisted of surgency/extraversion (E), agreeableness (A), conscientiousness (C), neuroticism (N), and intellect/openness (O). Over a period of almost 50 years, a series of cumulative studies may be traced (e.g., Fiske, 1949; Norman, 1963; Tupes & Christal, 1961) that led to an emerging consensus on the utility and generalizability of this five-factor representation (e.g., Digman, 1979; Goldberg, 1977; Wiggins, 1973b). By the 1980s, there were three distinctive interpretations of the nature of these five factors (Costa & McCrae, 1985; Goldberg, 1981; Hogan, 1986), that were followed in the 1990s by still more interpretive perspectives (e.g., Buss, 1996; Wiggins & Trapnell, 1996). It should be noted, however, that although there are a number of workers within the multivariate paradigm who agree on the number, if not the nature, of the factors in the FFM, there are many who strongly contest both the number of factors and their significance (e.g., Block, 1995; Eysenck, 1992; Hough, 1992; Waller & Ben-Porath, 1987).

4.12.4.2 Conceptual Framework

Cattell's (1957) master plan for constructing a theory of personality presupposed a rigorous and representative sampling of the language personality sphere that, in effect, defined the universe of content of behaviors to be explained (Wiggins, 1984). To make certain that this universe of content was properly defined, Norman (1967) replicated the earlier lexical research of Allport and Odbert (1936), using a more recent unabridged dictionary, and subsequently employing more rigorous clustering procedures than those used by Cattell (see John, 1990). Norman's taxonomy of personality attributes was subsequently employed by Goldberg (1981, 1990, 1993) in a program of methodologically elegant research on the natural language of personality that firmly established what he termed the "Big Five" dimensions identified by earlier investigators.

Goldberg's taxonomic research program revitalized interest in the five-factor model and within a relatively brief period of time a variety of new instruments were developed (e.g., Botwin & Buss, 1989; Costa & McCrae, 1985; Goldberg, 1992; Hogan, 1986; Trapnell & Wiggins, 1990) that measured these five dimensions within an equal variety of conceptual frameworks, for example, evolutionary theory (Buss, 1996), trait theory (McCrae & Costa, 1996), lexical theory (Saucier & Goldberg, 1996), socioanalytic theory (Hogan, 1996), and interpersonal theory (Wiggins & Trapnell, 1996). Moreover, a series of investigations by Costa and McCrae revealed that instruments employed in the other four traditions of personality assessment—psychodynamic (Myers & McCaulley, 1985), personological (Jackson, 1984), empirical (Hathaway & McKinley, 1983), and interpersonal (Wiggins, Trapnell, & Phillips, 1988)—showed meaningful convergences with some or all of the dimensions of the FFM (Wiggins & Trapnell, 1997).

4.12.4.3 Assessment Instruments

The NEO Personality Inventory (NEO PI-R; Costa & McCrae, 1992b) is the instrument of choice for measuring the dimensions of the FFM of personality. This test provides global measures (domain scores) for each of the five factors and six more specific measures (facet scores) within each of the five domains: (i) neuroticism (anxiety, angry hostility, depression, self-consciousness, impulsiveness, vulnerability); (ii) extraversion (warmth, gregariousness, assertiveness, activity, excitement-seeking, positive emotions); (iii) openness (fantasy, aesthetics, feelings, actions, ideas, values); (iv) agreeableness (trust, straightforwardness, altruism, compliance, modesty, tender-mindedness); and (v) conscientiousness (competence, order, dutifulness, achievement striving, self-discipline). The "astonishingly fruitful research collaboration" (Block, 1995) of Costa and McCrae has produced scores of publications that document the substantive, structural, and empirical validity of this instrument.

4.12.4.4 Interpretive Principles

There are notable differences in interpretive principles suggested by proponents of different multivariate instruments and these would appear to reflect different claims as to the extensiveness and importance of the "universe of content" measured by each instrument, and the theoretical basis on which that content is interpreted for each instrument. With respect to the former, the differences among alternative theories of personality structure within the multivariate tradition have often reflected preferences for different methods of factoring correlations among the phenotypic variables of personality study. Thus, Cattell, who argued for 16 primary source traits, as well as the originator of the multiple-factor method himself (Thurstone, 1934), have both been accused of "overfactoring" (Digman, 1996); while Eysenck (1992), who disputes anything beyond three factors, has been accused of "underfactoring" (Costa & McCrae, 1992a). That almost half a century of essentially methodological disagreements may be approaching resolution can be seen in the emerging consensus that the FFM of personality is an appropriate "working model" of the universe of content of personality structure that may (and should) be expanded, as necessary, whenever the incremental validity of additional dimensions are demonstrated (Wiggins, 1996).

The theoretical bases on which a five-factor profile may be interpreted are of greater interest to the practitioner. And, as is true of a number of instruments that originated in other personality assessment traditions, it is possible to interpret the FFM from a variety of theoretical perspectives. The NEO PI-R (Costa & McCrae, 1992b) provides explicit principles for interpreting normal and abnormal behavior from the standpoint of modern trait theory (McCrae & Costa, 1996). The Hogan Personality Inventory (Hogan & Hogan, 1992) provides explicit interpretive principles for personnel selection in the workplace from the standpoint of socioanalytic theory (Hogan, 1996). A new edition of the venerable Sixteen Personality Factor Questionnaire (Cattell, Cattell, & Cattell, 1993) provides interpretive principles for five "global factor scales" (Cattell, 1994) from the standpoint of a long-standing "mainstream" personality theory (Wiggins, 1984); and the possibilities for interpreting other well-established instruments, such as the Personality Research Form (Jackson, 1984), from an FFM perspective, are under active investigation (e.g., Jackson, Paunonen, Fraboni, & Goffin, 1996).

4.12.4.5 Applications and Current Status

Historically, there has been little consensus regarding the relation between the normal dimensions of personality, measured by inventories developed within the multivariate tradition, and the dimensions and/or categories of psychopathology. Recently, however, there has been an increased conceptual and empirical focus on this issue, stimulated in part by the definition of personality disorders in terms of personality traits by the American Psychiatric Association (APA; 1980) and by the availability of new instruments and techniques for addressing this problem (Strack & Lorr, 1994).

It would appear that the global dispositions measured by the NEO PI-R provide a comprehensive assessment of emotional (N), interpersonal (E & A), motivational (C), and cognitive (O) styles that are not measured by traditional clinical instruments. As Costa (1991b) noted, "This portrait can be used in psychotherapy: to formulate a tentative diagnosis; to aid the clinician's development of empathy; to help select appropriate treatments; to identify the client's strengths; and to anticipate the outcome, duration, and course of therapy" (p. 395). A detailed and perceptive account of how this might be done has been provided by Miller (1991). The notable longitudinal stability of NEO PI-R dimensions in the adult personality (McCrae & Costa, 1990) has led to a revised view of the therapeutic enterprise:

> we expect that clients will bring these dispositions with them to psychotherapy, and that the therapist should (a) take them into account when trying to understand the individual and his or her problems, and (b) tailor the therapy to fit the needs and styles of the client. (McCrae, 1991, p. 406)

The most extensive psychodiagnostic application of the FFM in general, and of the NEO PI-R in particular, has been to the diagnosis and understanding of personality disorders (Costa & Widiger, 1994). The definition of these disorders in terms of dysfunctional personality traits, by a committee of the APA (1980), stimulated an unprecedented collaborative effort among psychiatrists, psychologists, and psychometricians to find common grounds for characterizing these disorders. It was found, for example, that both the APA diagnostic criteria and the clinical literature on personality disorders were compatible with descriptions of each disorder generated from the facet and domain scales of the NEO PI-R (see Widiger, Trull, Clarkin, Sanderson, & Costa, 1994, Table 1, p. 42).

Subsequent advances in relating the FFM to the personality disorders are too numerous to summarize, but may be found in current journals, particularly the *Journal of Personality Disorders*. The current literature may be viewed as reflecting yet another milestone in the long-standing relation between the FFM and the multivariate paradigm, which Digman (1994) characterized as follows:

> Now, after many years of lying on the closet shelf of personality theory, the model has been dusted off, "as good as new," and appears to be for many researchers ... a very meaningful theoretical structure for organizing the myriad specifics implied by the term *personality*. (p. 13)

4.12.5 EMPIRICAL TRADITION

4.12.5.1 Historical Background

The MMPI (Hathaway & McKinley, 1943) is, and has been for many years, the most widely used inventory of personality and psychopathology. To understand the 50-year history of this instrument is to understand the history of objective personality assessment for the same time period. As Craik (1986) put it, in his historical survey of personality research methods, "the MMPI came to serve as the centerpiece of this period's [post-WWII] predominant or mainstream agenda" (p. 21). Within that agenda, one can discern the operation of two separate dialectical processes over time: ongoing disputes between the developers of the MMPI and their critics and bipolar shifts in conceptualization that have occurred within the evolving conceptual and interpretive frameworks of the developers of the MMPI. The ongoing disputes reflect the fact that, because of its prominence, the MMPI came to be associated with such contentious issues as clinical versus statistical prediction (Meehl, 1954) and response styles (Wiggins, 1962). The shifts in conceptualization reflect the virtuosity and intellectual flexibility of the developers of the MMPI who were able to shift from typological categories to continuous trait dimensions; from discriminant validity of differential diagnoses to the construct validity of scales and profiles; and from denigration of self-reports to the canonization of item content (Ben-Porath, 1994).

The dialectical processes just described were enacted by a remarkably diverse group of scientists/practitioners whose program of clinical research had its roots in the unique combination of behavioral, biological and psychometric thinking that prevailed at the University of Minnesota in the 1930s and 1940s (Meehl, 1989). Stark R. Hathaway was an originator of the MMPI, a mentor for successive generations of Minnesota students and a merciless critic of his own instrument (e.g., Hathaway, 1972). Paul E. Meehl—learning theorist, philosopher of science, psychoanalyst, and taxonometrician—is a legend in his own time, the principal theorist of the MMPI, and arguably the major figure in the field of personality assessment since the late 1950s. W. Grant Dahlstrom is the Talmudic scholar of the MMPI, whose meticulous organizations of that vast literature have informed and influenced generations of clinicians and researchers (e.g., Dahlstrom & Dahlstrom, 1980; Dahlstrom, Welsh, & Dahlstrom, 1972–1975; Welsh & Dahlstrom, 1956). James N. Butcher is a distinguished editor, scholar, and researcher who is principal author of the revised MMPI (MMPI-2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989), the new MMPI content scales (Butcher, Graham, Williams, & Ben-Porath, 1990), the adolescent version of the MMPI (MMPI-A; Butcher et al., 1992), and the MMPI automated clinical report (Butcher, 1993).

4.12.5.2 Conceptual Framework

The rationales underlying four distinguishable strategies of objective (as opposed to projective) test construction are based on differing views regarding the meaning of subjects' responses to objective test items (Wiggins, 1973b)

(i) *The rational (correspondence) strategy*. This assumes a one-to-one correspondence between a subject's "report" ("I am anxious") and a palpable internal state of the subject (the experience of anxiety) (Buchwald, 1961). This assumption was the cornerstone of introspectionism in early experimental psychology and it was the principal rationale underlying the earliest personality inventories (e.g., Woodworth, 1917).

(ii) *The empirical (instrumental) strategy*. This views a subject's "utterance" as "an intrinsically interesting and significant bit of verbal behavior, *the nontest correlates of which must be discovered by empirical means*" (Meehl, 1945, p. 297, emphasis added). If it were established empirically that a false response to the item "I am anxious" was more characteristic of hospitalized hysterics than of a normal control group, that particular bit of verbal behavior would become meaningful, and would illustrate the instrumental value of verbal behavior as a tool in psychodiagnostics. This philosophy was evident in the early behavioristic critique of introspectionism in experimental psychology and in later critiques of existing rational personality inventories of the 1930s (Humm & Wadsworth, 1935; Landis & Katz, 1934; Landis, Zubin, & Katz, 1935). The earliest example of a major inventory constructed under an empirical strategy was the Vocational Interest Blank (Strong, 1927), which served as the prototype for the construction of the MMPI.

(iii) *The constructive (substantive) strategy*. This views a subject's response as a manifestation of an underlying personality construct that is embedded in an interlocking system of laws that relate constructs to one another and to observable properties of the environment (Cronbach & Meehl, 1955). The meaning of a construct is given by the empirical laws into which it enters and the significance of a construct is given by the number of such lawful relations discovered. The construct validity of a personality scale is estimated from "the proportion of test score variance that is attributable to the construct variable" (p. 289).

(iv) *The interpersonal (self-presentational) strategy*. This views a subject's response as an interpersonal communication between the subject and the tester (or the institution that the tester represents) (Carson, 1969b; Leary, 1957). The meaningfulness of such a communication is dependent upon both the subject's and the tester's views of the meaning of test responses in a particular testing situation (e.g., hospital admission, vocational guidance). Thus, the subject may mean to communicate a complaint ("I am anxious") and the tester (or the scoring system employed by the institution) may encode this as "neuroticism." The subject may view the testing situation as an opportunity for self-disclosure (Jourard, 1964) and the examiner may view it as an opportunity for impression management (Goffman, 1959). Elsewhere it has been argued that the interpersonal view of test responses has much to commend it, not the least of which is the likelihood that this is the frame of reference which subjects themselves adopt (Wiggins, 1966).

4.12.5.3 Assessment Instruments

Construction of the MMPI began with the generation of a large pool of items that were considered representative of "behaviors of significance to the psychiatrist" (Hathaway & McKinley, 1940). These items were presented in true–false format to several groups of diagnosed psychiatric inpatients (e.g., depressives) and to a large group of normal control subjects. The proportion of patients in a given diagnostic group who responded "true" ("endorsement") to each item (e.g., 80%) was compared with the proportion of normal subjects who had responded "true" to that item (e.g., 10%), in order to identify discriminating items that might be keyed (in this instance, "true") on a clinical scale (e.g., depression scale; Hathaway & McKinley, 1942). Innovative procedures were then applied in the development of three validity scales that served to identify sources of invalidity or "faking" in test responding: a "lie scale" (*L*) with items of high desirability and low endorsement, an "infrequency scale" (*F*) with items of low desirability and low endorsement, and a subtle "correction scale" (*K*) that measures defensiveness when high and self-criticalness when low. The resultant three validity scales, eight clinical scales (hypochondriasis, depression, hysteria, psychopathic deviate, paranoia, psychasthenia, schizophrenia, hypomania), a masculinity/femininity scale, and a social introversion scale constitute the MMPI profile that was normed and standardized on the original sample of normal control subjects.

4.12.5.4 Interpretive Principles

4.12.5.4.1 Empirical strategy

The original purpose of the MMPI was to aid psychiatrists in their assignment of patients to the traditional psychiatric diagnostic categories of that time. The conservative rationale of the empirical strategy permitted such statements as "This subject's pattern of verbal behavior resembles more closely that of a group of hospitalized schizophrenics than it does that of a normal control group." This fact, together with evidence from other sources, was suggestive of the working diagnosis of "schizophrenia." Unfortunately, the individual clinical scales were not successful in subsequent attempts to discriminate diagnostic groups from normals (Hathaway, 1960), due most likely to lack of power in the original statistical comparisons and to unreliability of both predictor and criterion measures.

4.12.5.4.2 Constructive strategy

Ben-Porath (1994) provides an excellent historical summary of the "reinvention of the MMPI" following the failure of the empirical strategy and of its subsequent "evolution into an omnibus measure of personality" within a construct-oriented perspective. The highlights of this shift over a 50-year period of unprecedented research and clinical work include emphases upon: (i) MMPI profile configurations, rather than on single scales (e.g., Gough 1946); (ii) normal personality correlates of profile types (e.g., Black, 1956); (iii) additional empirically and substantively derived scales that broadened the nomological network of MMPI investigation (e.g., Morey, Waugh, & Blashfield, 1985), and (iv) actuarial prediction systems based on configural profile types (e.g., Marks & Seeman, 1963).

4.12.5.4.3 Interpersonal strategy

Loevinger's (1957) formulation of the substantive component of construct validity provided an even greater contrast to the original empirical perspective of the MMPI. Nevertheless, workers within the MMPI tradition for the most part continued to pursue the other components of construct validity just described. Interest in the "content" of patients' communication of complaints via the MMPI remained minimal until the appearance of Wiggins' (1966) MMPI Content Scales. In terms of Loevinger's distinctions, these content scales appeared to provide: (i) representative coverage of the universe of content of the MMPI item pool (see Johnson, Butcher, Null, & Johnson, 1984), (ii) psychometrically sound measures of 13 dimensions of self-report that were interpretable with reference to previous studies of the factorial structure of the MMPI (Welsh, 1956), and (iii) empirical evidence of convergent and discriminant validity with reference to psychiatric diagnostic categories (Wiggins, 1966; Payne & Wiggins, 1972). Perhaps of greater interest to the MMPI community was the potential usefulness of these content scales in individual psychodiagnostic appraisal (e.g., Nichols, 1987). Variants of the original content scales are now an official supplement to the clinical scales of MMPI-2 (Butcher et al., 1990) and their incremental contribution to the differential diagnosis of psychopathology appears highly promising (e.g., Ben-Porath, Butcher, & Graham, 1991).

4.12.5.5 Applications and Current Status

The MMPI developed in the 1940s was, by contemporary standards, based on inadequate and outdated norms, replete with offensive and anachronistic items, and patently unsuited to the diagnostic task for which it was devised. The clinical scales were derived by less than optimal contrasted group procedures (Jackson, 1971) that, from the beginning, appeared unreplicable (Benton, 1949). In the absence of a coherent theoretical rationale, the profiles, code-types, and actuarial systems based on these scales have an ambiguous conceptual status (Helmes & Reddon, 1993). Pressures to restandardize, revise, or replace this relic have existed almost from its inception, culminating in a historically important summit meeting (Butcher, 1972) in which the third option was seriously considered (Wiggis, 1973a). Once again, however, the resilient and perdurable Minneapolis Group (Butcher et al., 1989) rose to the challenges from within and without to produce what was optimistically dubbed "MMPI-2" (leaving open the possibility of MMPI-3 in the year 2035).

Although the original MMPI is a disgruntled psychometric critic's dream come true, it is difficult to ignore the past and present status of the MMPI as our most widely used objective personality test. Given what appeared to be the lesser of two evils, the personnel of the MMPI Restandardization Project opted to emphasize continuity with the vast interpretive and empirical database accumulated for the original clinical and validity scales while simultaneously removing inappropriate items, correcting psychometric deficiencies where possible, and generating new content scales. Shorn of its dated and politically incorrect items, the lamb-like content of the MMPI-2 is now clothed in the lupine garments of computer technology and state-of-the-art methodology (e.g., uniform *T*-scores, adaptive testing, measurement of variable response inconsistency). MMPI-2 appears to be replacing the original in most clinical settings (e.g., Webb, Levitt, & Rojdev, 1993), although there continue to be criticisms from both without and within the MMPI community (e.g., Caldwell, 1997).

In 1972, Hathaway concluded: "So, in summary, we are stuck, I think, with the MMPI ... for a dreary while longer, although a prophet may even now be wandering in the psychometric wilderness" (1972 p. 40). There is little doubt that we will be "stuck" with MMPI-2 for many years to come, although the paths of the prophets are now more discernible (e.g., Costa, 1991a; Morey, 1991).

4.12.6 INTERPERSONAL PARADIGM

4.12.6.1 Historical Background

The interpersonal circumplex model had its origins in the writings of Harry Stack Sullivan (1953) who introduced a number of original and radical ideas into the American psychiatry of the 1940s. Although not entirely liberated from Freud's spatial-hydraulic metapsychology, Sullivan was "behavioral" in the sense that his primary emphasis was upon the things that persons do to one another in interpersonal transactions, rather than upon internal processes that persons might "have." His oft-cited definition of personality as "the relatively enduring pattern of recurrent interpersonal situations which characterize a human life" represents an even greater departure from psychoanalytic thought, as well as a unique perspective within personality theory itself.

Sullivan was influenced by Lewin's (1938) field-theoretical conceptualization of contemporaneous, bidirectional influences in "psychological fields," which led him to define the interpersonal situation as the basic unit of observation in psychiatry. This idea does not connote a separate "person" or "persons" who may be considered independently from the complex field (situation) of bidirectional causalities in which they are embedded; nor does it hold out the hope of capturing such complexities as "interaction" terms in analyses of variance.

The Kaiser Foundation Research Project, in which Timothy Leary (1957) was the most preeminent investigator, attempted to operationalize Sullivanian concepts in terms of concrete measurement procedures. An ordinary language analysis of clinicians' observations of the things that patients did to each other (and to themselves) in group psychotherapy led to a taxonomy of interpersonal behaviors that appeared to be empirically well-captured by a circular arrangement of interpersonal variables organized around the coordinates of "dominance" and "affiliation." LaForge, Leary, Naboisek, Coffey, and Freedman (1954) derived the basic trigonometry of what would later be called a "circumplex," without knowledge of Guttman's (1954) related work on that topic.

Notable contributions during the 1960s included the conceptualization of both maternal and child behavior within an interpersonal circumplex framework (Schaefer, 1961), a psychometrically sophisticated replication of the circumplex within a clinical population (Lorr & McNair, 1963), and a highly influential integration of the circumplex with the clinical, social, and experimental psychology of that time (Carson, 1969a). Alternative conceptual formulations of the interpersonal model were presented in the 1970s (e.g., Benjamin, 1974; Kiesler, 1979; Wiggins, 1979), and during the 1980s the model was applied to psychotherapy (Anchin & Kiesler, 1982), complementarity (Kiesler, 1983), and interpersonal problems (Horowitz, Rosenberg, Baer, Ureno, & Villasenor, 1988). Among the contributions of the 1990s have been an updated and psychometrically-sound version of the original interpersonal checklist (Wiggins, 1995), a comprehensive exposition of contemporary interpersonal theory and research (Kiesler, 1996), and a presentation of the impressive variety of contexts in which the interpersonal circumplex model has been applied (Plutchik & Conte, 1997).

4.12.6.2 Conceptual Framework

The radical interpersonalism of Sullivan (1953) may be difficult to comprehend on first exposure because of our ingrained "individualist language" (p. 50) for describing personality as reflecting attributes of a discrete individual who is "separate" from others and from a shared social environment. Theorists within the interpersonal paradigm have attempted to operationalize Sullivan's conceptualization of personality in ways that avoid this individualistic bias, by defining personality as "nothing more (or less) than the patterned regularities that may be observed in an individual's relations with other persons, who may be real in the sense of actually being present, real but absent and hence 'personified' or 'illusory'" (Carson, 1969a, p. 26).

On a metatheoretical level, it is helpful to think of the dominance and affiliation axes of the interpersonal circumplex in terms of Bakan's (1966) concepts of "agency" (mastery, self-assertion, self-expansion) and "communion" (union with a larger social or spiritual entity), respectively (Wiggins, 1991). These concepts are compatible with Sullivan's emphasis on the communal context of interpersonal transactions and with his belief in the importance of the social sciences for understanding interpersonal situations. When conjoined with Foa and Foa's (1974) theory of the exchange of love (communion) and status (agency) in interpersonal transactions, these concepts provide a rich conceptual framework for the measurement and understanding of interpersonal behavior (Wiggins & Trapnell, 1996).

4.12.6.3 Assessment Instruments

4.12.6.3.1 Interpersonal Adjective Scales

The Interpersonal Ajective Scales (IAS; Wiggins, 1995) evolved from a psychological taxonomy of trait-descriptive terms that was developed within the framework of a larger program of collaborative research on language and personality (Goldberg, 1977). Within a representative pool of trait-descriptive adjectives selected from an unabridged dictionary, an "interpersonal domain" was distinguished from other domains, such as characterological, temperamental, and cognitive domains (Wiggins, 1979). Approximately 800 interpersonal adjectives were assigned to the original categories of the interpersonal circumplex (Leary, 1957) on both conceptual and empirical grounds. Using computer-based multivariate procedures it was found that scales based on the original categories failed to meet certain circumplex criteria that could be better met with scales based on the revised categories that now constitute the IAS (Wiggins, 1995). The IAS consists of 64 adjectives that respondents rate for self-descriptive accuracy on an eight-place Likert scale ranging from "extremely inaccurate" to "extremely accurate." Eight scales, or "octants," of eight items each, assess the interpersonal dispositions listed in the first column of Table 1. The PA octant, for example, includes items such as "dominant," "forceful," and "assertive."

4.12.6.3.2 Inventory of Interpersonal Problems

Horowitz (1979) transcribed statements of interpersonal problems expressed by psychiatric outpatients in the course of videotaped intake interviews and employed these statements as items in the construction of the Inventory of Interpersonal Problems (IIP), in which respondents are required to indicate the extent to which each of 127 statements is problematic on a five-place Likert scale ranging from "not at

Table 1 Octant scales from three interpersonal assessment instruments.

	IAS Interpersonal Adjective Scales	*IAS Inventory of Interpersonal Problems*	*IAS Impact Message Inventory*
PA	Assured–dominant	Domineering	Dominant
BC	Arrogant–calculating	Vindictive	Hostile–dominant
DE	Cold-hearted	Cold	Hostile
FG	Aloof-introverted	Socially avoidant	Hostile–submissive
HI	Unassured–submissive	Nonassertive	Submissive
JK	Unassuming–ingenuous	Exploitable	Friendly–submissive
LM	Warm–agreeable	Overly nurturant	Friendly
NO	Gregarious–extraverted	Intrusive	Friendly–dominant

all" to "extremely" (Horowitz et al., 1988). Subsequently, Alden, Wiggins and Pincus (1990) developed a circumplex version of this inventory (IIP-C) that consists of eight scales, of eight items each, that assess the interpersonal problems listed in the second column of Table 1. The PA octant, for example, includes items such as "I try to control other people too much."

4.12.6.3.3 Impact Message Inventory

The Impact Message Inventory (IMI) is a highly original and promising method of assessment that is based on Kiesler's (1988) theory of interpersonal communication in psychotherapy. The theory postulates that disordered individuals are unaware of the unintended, inappropriate, and ambiguous messages they repetitively "send" to others and that they are therefore confused and distressed by the pattern of negative responses they consistently evoke or "pull" from others. The IMI attempts to identify the location within the interpersonal circumplex of these patterns of negative response evoked in others, as a means of gaining insight into a client's maladaptive transactional behavior.

Items were generated from a content analysis of free responses to 15 interpersonal vignettes that described characters enacting 15 different interpersonal styles, similar to those found in the first column of Table 1. Respondents were asked to imagine themselves in the company of each of these characters and to record their covert reactions using the stem, "He makes me feel..." Content analysis of responses suggested three categories of covert reaction: direct feelings, action tendencies, and perceived evoking messages. The Octant version of the IMI consists of six items for each of the target stimuli listed in the third column of Table 1. For each octant, there are two items for each of the three categories of covert reaction to the target stimulus. Thus, for example, the PA scale includes two items each for direct feelings (e.g., "bossed around"), action tendencies (e.g., "I want to tell him to give someone else a chance to make a decision"), and perceived evoking messages (e.g., "He thinks he's always in control of things"). Respondents (e.g., a psychotherapist) are asked to imagine themselves in the company of a particular person (e.g., a psychotherapy patient) and to indicate the extent to which they experience the covert reactions on a four-place scale. Covert reactions to a target person (e.g., feeling "bossed around") are thus revealing of the personality and the interpersonal impact of that target person (e.g., rigidly dominant, PA).

4.12.6.4 Interpretive Principles

One of the most valuable features of the circumplex model is the opportunity it provides for interpreting interpersonal variables with reference to the geometric principles of a circle (LaForge et al., 1954). The variables of traditional multiscale inventories are typically displayed as "factor lists" (Hogan, 1983) in which the ordering and interrelations among variables lack both conceptual and interpretive significance. In contrast, the trigonometric procedures that can be applied to the variables of an interpersonal circumplex permit descriptions and diagnostic inferences that cannot be generated from traditional scales (Wiggins, Phillips, & Trapnell, 1989).

Some general principles of circumplex interpretation may be illustrated with reference to Figure 1, which presents the IAS profile of a 44-year-old woman who was employed as a senior bank manager. At the bottom of the figure are *T*-scores on the octant variables, expressed with reference to an appropriate normative group. These scores have been plotted on shaded sectors of the circle and they are interpreted as representing eight vectors in two-dimensional space. The mean or average directionality of these eight vectors is of critical diagnostic significance because it determines the "typological" category to which this woman will be assigned. By trigonometric procedures, this average directionality was determined to be 90°, which falls exactly at the midpoint of the assured-dominant category (PA). Such an angular location is considered to be prototypical of individuals classified as pure "PA types." Had the location been at 110°, a more hostile manifestation of dominance (BC) would have been suggested; had it been at 65°, a more affiliative expression of dominance (NO) would have been suggested. But the protypicality of the present classification increases our confidence in asserting that the principal interpersonal style of this woman is one that emphasizes the exercise of power over others in a social context, by such activities as taking charge, making decisions, and winning arguments (Wiggins, 1995, p. 22).

The general shape of the profile in Figure 1 closely resembles the characteristic configuration of IAS profiles that is found in the average profiles of subjects in all IAS typological groups (Wiggins et al., 1989). The characteristic configuration of an IAS profiles is one in which the principal elevation occurs on the defining octant (in this case, PA), with secondary and approximately equal elevations occurring on adjacent octants (BC and NO), and diminishing and approximately equal elevations occurring on subsequent pairs of octants (DE and LM), (FG and JK), down to a highly truncated "opposite" (to the principal) octant (HI). Thus, across interpersonal situations, we would expect this woman frequently to behave in a forceful, assertive, dominant, and self-confident manner (PA); to somewhat less frequently behave in aggressive (BC) and gregarious (NO) ways; to seldom, if ever, behave in a submissive (HI) fashion, and so forth.

The vector length of an interpersonal profile is a measure of "deviance" in both a statistical and a psychiatric sense. In the former sense, vector length is the standard deviation of the eight interpersonal variables that indicates the "intensity" with which a pattern of interpersonal behavior is expressed. In the latter sense, high-variance profiles are associated with interpersonal "rigidity" and are found most often in psychiatric groups. The length of the arrow indicating angular location in Figure 1 is the vector length of the profile, which is approximately 1.8 standard deviations above that of the normative group. This would suggest that the pattern of interpersonal behaviors described above would be expressed vigorously, and quite possibly rigidly or intemperately.

Having established that this individual's profile is representative of those obtained by assured dominant "types," and that this pattern of behavior is expressed in a clearly differentiated fashion, interpretation would proceed with reference to the empirical literature of the interpersonal tradition that has examined both the dimension of assured-dominant (PA) and the characteristics of dominant "types" in both normal and psychiatric populations (Kiesler, 1996). This would include the literature of other instruments that have studied this location (PA) in contexts such as interpersonal problems (IIP-C), impact messages (IMI), and other contexts that have been studied by interpersonal researchers.

4.12.6.5 Applications and Current Status

The definition of the personality disorders in terms of personality traits by the APA (1980), and the availability of new instruments and techniques for assessing personality disorder, has resulted in an increased understanding of this diagnostic axis. For example, many of the *Diagnostic and statistical manual of mental disorders* (*DSM*) personality disorder categories have been shown to be well captured by the two-dimensional structures of the IAS (e.g., Wiggins & Pincus, 1989, 1994) and the IIP-C (e.g., Pincus & Wiggins, 1990; Soldz, Budman, Demby, & Merry, 1993). Similarly, clinicians' ratings of *DSM* personality disorder categories have been found to be well captured by an interpersonal circumplex model (e.g., Blashfield, Sprock, Pinkston, & Hodgin, 1985; Plutchik & Conte, 1986). And more recently, Benjamin (1995) has provided a detailed and perceptive description of procedures for the diagnosis of personality disorders within the framework of her variant of the interpersonal circumplex model.

The interpersonal tradition originated in the context of psychotherapy and, perhaps more than any other tradition, has contributed to an understanding of the therapeutic process itself (e.g., Anchin & Kiesler, 1982; Benjamin, 1995; Kiesler, 1988; Safran & Segal, 1990; Sullivan, 1953). The most notable contribution of this tradition to psychotherapy has been the circumplex structural model that provides a framework for representing dyadic interactions in the therapeutic relationship. A number of different assessment devices based on the interpersonal circumplex, and variants of that model (e.g., Benjamin, 1974), have proven useful in studies of psychotherapy process and outcome (Kiesler, 1996).

Gurtman (1996) has emphasized the construct validity of the circumplex version of Horowitz's Inventory of Interpersonal Problems (IIP-C; Alden et al., 1990) within the psychotherapy context. Henry (1996) has made

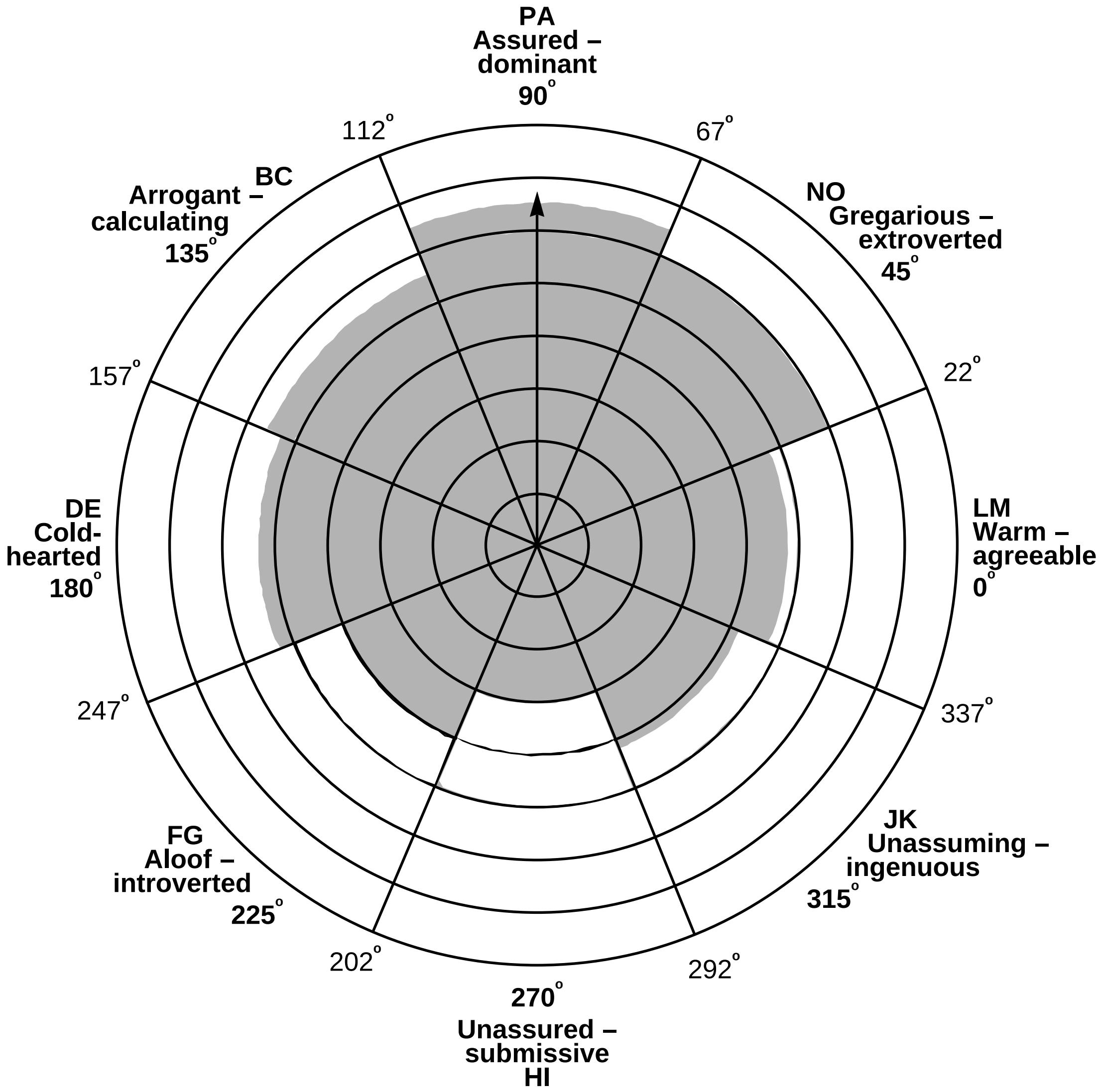

	PA	BC	DE	FG	HI	JK	LM	NO
T-Score	65	56	52	40	30	41	49	60

Angular location = 90°

Vector length T-score = 68°

Figure 1 IAS profile of a 44-year-old bank manager.

a similar case for the Structural Analysis of Social Behavior (Benjamin, 1974) as a common metric for programmatic psychotherapy research. Kiesler, Schmidt, and Wagner (1997) have stressed the conceptual advantages of the IMI in analyzing the psychotherapeutic relationship. And Kiesler (1996) has also emphasized the potential of the revised Check List of Psychotherapy Transactions (Kiesler, Goldston, & Schmidt, 1991) for measuring the interpersonal behavior of interactants in the therapy relationship. Taken together, the studies summarized in the aforementioned review articles constitute an impressive empirical literature that attests to the heuristic potential of the interpersonal circumplex model in psychotherapy research. The future of the interpersonal circumplex tradition for research in both psychodiagnostics and psychotherapy appears to be a bright one (Wiggins & Trobst, 1997).

4.12.7 CONCLUSION

In the introduction to this chapter, we considered the diagnostic work of a hypothetical clinician who was both well trained and well informed about personality tests, but who was perhaps unaware of the conceptual and interpretive richness, and particularly of the distinctiveness, of the five personality assessment traditions just considered. Assuming that she persevered in reading through this highly condensed overview, there are a few "take home" messages that might now be conveyed to her.

First, in Kelly's (1955) terms, the five traditions arose and were developed within different "foci of convenience" and therefore are often most usefully applied within their specific areas of focus. The psychodynamic tradition of Freud arose from the revolutionary idea that the causes and reasons for behavior are frequently not apparent and that we must look "within" the individual, as it were, to discover the sources (often historical) of contemporary problems of living. The personological tradition arose in response to the question of "How shall a psychological life history be written?" The multivariate tradition arose as an extension of the mulifactor assessment of intelligence to the personality domain, and was subsequently concerned with the determination of the basic descriptive dimensions of personality. The empirical tradition arose in the context of descriptive psychiatry, and was originally most concerned with the differential diagnosis of psychopathology. The interpersonal tradition was originally focused on the character and quality of interpersonal relationships and on what persons did to one another within such relationships.

Second, instruments developed or employed within the five traditions have been found to have different "ranges of convenience" (Kelly, 1955), that is, to be generally or specifically applicable to a particular spectrum of assessment problems. Here it must be reiterated that there is not a one-to-one relation between instruments and traditions in this respect. A few examples must suffice: The Rorschach test may be useful in assessing cognitive and perceptual deficits; the WAIS may be useful in detecting covert motivational conflicts; the MMPI item pool has proven a fruitful source for the construction of literally hundreds of scales serving different assessment purposes; the TAT, and its associated need-press framework, has proven useful within many different traditions, as has the NEO PI-R, and the IAS.

Finally, our hypothetical clinician should be aware of the armamentarium of conceptual orientations and instruments that may be brought to bear on a specific referral or assessment task. It is extremely unlikely that there is one test, or one orientation, that will be optimal for all assessment tasks. Practical limitations, especially managed care in the USA, often preclude the once-favored idea of a "standard battery" of tests. It would also not be fair to expect our hypothetical clinician to master all available orientations and tests. But perhaps she could become familiar with a few more. Attending a workshop or two (outside of her present area of expertise) might enhance the professional life of our hypothetical clinician.

ACKNOWLEDGMENTS

We would like to express our indebtedness to Yossef S. Ben-Porath, Sidney J. Blatt, Dan P. McAdams, David S. Nichols, and William McKinley Runyan for their helpful suggestions concerning earlier drafts of sections within their own areas of expertise. However, we accept full responsibility for any errors of fact or interpretation that remain in the present version.

4.12.8 REFERENCES

Alden, L. E., Wiggins, J. S., & Pincus, A. L. (1990). Construction of circumplex scales for the Inventory of Interpersonal Problems. *Journal of Personality Assessment, 53,* 521–536.

Alexander, I. E. (1990). *Personology: Method and content in personality assessment and psychobiography.* Durham, NC: Duke University Press.

Allison, J., Blatt, S. J., & Zimet, C. N. (1968). *The interpretation of psychological tests.* New York: Harper and Row (Reprinted in 1988 by Taylor & Francis).

Allport, G. W. (1937). *Personality: A psychological interpretation.* New York: Holt, Rinehart, & Winston.

Allport, G. W. (1942). *The use of personal documents in psychological science.* New York: Social Science Research Council.

Allport, G. W. (1965). *Letters from Jenny.* New York: Harcourt, Brace, & World.

Allport, G. W. (1967). Autobiography. In E. G. Boring & G. Lindzey (Eds.), *A history of psychology in autobiography* (Vol. 5, pp. 1–25). New York: Appleton-Century-Crofts.

Allport, G. W., & Odbert, H. S. (1936). Trait names: A psycho-lexical study. *Psychological Monographs, 47* (Whole Issue No. 211).

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

Anchin, J. C., & Kiesler, D. J. (Eds.) (1982). *Handbook of interpersonal psychotherapy.* Elmsford, NY: Pergamon.

Babcock, H. (1933). *A short form of the Babcock examination for the measurement of mental deterioration.* Chicago: Stoelting.

Bakan, D. (1966). *The duality of human existence: Isolation and communion in Western man.* Boston: Beacon Press.

Behrends, R. S., & Blatt, S. J. (1985). Internalization and psychological development through the life cycle. *Psychoanalytic Study of the Child, 40,* 11–39.

Benjamin, L. S. (1974). Structural analysis of social behavior. *Psychological Review, 81,* 392–425.

Benjamin, L. S. (1995). *Interpersonal diagnosis and treatment of personality disorders* (2nd ed.). New York: Guilford Press.

Ben-Porath, Y. S. (1994). The MMPI and MMPI-2: Fifty years of differentiating normal and abnormal personality. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (pp. 361–401). New York: Springer.

Ben-Porath, Y. S., Butcher, J. N., & Graham, J. R. (1991). Contribution of the MMPI-2 content scales to the differential diagnosis of psychopathology. *Psychological Assessment, 3,* 634–640.

Benton, A. L. (1949). Review of Minnesota Multiphasic Personality Inventory. In O. K. Buros (Ed.), *The third mental measurements yearbook* (pp. 104–107). Highland Park, NJ: Gryphon Press.

Black, J. D. (1956). Adjectives associated with various MMPI codes. In G. S. Welsh & W. G. Dahlstrom (Eds.), *Basic readings on the MMPI in psychology and medicine* (pp. 151–172). Minneapolis, MN: University of Minnesota Press.

Blashfield, R., Sprock, J., Pinkston, K., & Hodgin, J. (1985). Exemplar prototypes of personality disorder diagnoses. *Comprehensive Psychiatry, 26,* 11–21.

Blatt, S. J., & Allison, J. (1981). The intelligence test in personality assessment. In A. I. Rabin (Ed.), *Assessment with projective techniques* (pp. 137–231). New York: Springer.

Blatt, S. J., Brenneis, C. B. Schimek, J. G., & Glick, M. (1976). The normal development and psychopathological impairment of the concept of the object on Rorschach. *Journal of Abnormal Psychology, 85,* 364–373.

Blatt, S. J., & Lerner, H. D. (1983). The psychological assessment of object representation. *Journal of Personality Assessment, 47,* 7–28.

Block, J. (1995). A contrarian view of the five-factor approach to personality *Psychological Bulletin, 117,* 187–215.

Botwin, M. D., & Buss, D. M. (1989). Structure of act-report data: Is the five-factor model of personality recaptured? *Journal of Personality and Social Psychology, 56,* 988–1001.

Bowlby, J. (1973). *Attachment and loss, Vol. 2: Separation.* New York: Basic Books.

Buchwald, A. M. (1961). Verbal utterances as data. In H. Feigl & G. Maxwell (Eds.), *Current issues in the philosophy of science* (pp. 461–468). New York: Holt.

Buss, D. M. (1996). Social adaptation and five major factors of personality. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 180–207). New York: Guilford Press.

Butcher, J. N. (Ed.) (1972). *Objective personality assessment: Changing perspectives.* New York: Academic Press.

Butcher, J. N. (1993). *User's guide for the Minnesota Clinical Report.* Minneapolis, MN: National Computer Systems.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *The Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring.* Minneapolis, MN: University of Minnesota Press.

Butcher, J. N., Graham, J. R., Williams, C. L., & Ben-Porath, Y. S. (1990). *Development and use of the MMPI-2 content scales.* Minneapolis, MN: University of Minnesota Press.

Butcher, J. N., Williams, C. L., Graham, J. R., Archer, R. P., Tellegen, A., Ben-Porath, Y. S., & Kaemmer, B. (1992). *Minnesota Multiphasic Personality Inventory (MMPI-A): Manual for administration, scoring, and interpretation.* Minneapolis, MN: University of Minnesota Press.

Caldwell, A. B. (1997). Whither goest our redoubtable mentor, the MMPI/MMPI-2? *Journal of Personality Assessment, 68,* 47–68.

Carson, R. C. (1969a). *Interaction concepts of personality.* Chicago: Aldine.

Carson, R. C. (1969b). Interpretive manual to the MMPI. In J. N. Butcher (Ed.), *MMPI: Research developments and clinical applications* (pp. 41–53). New York: McGraw-Hill.

Cattell, H. E. P. (1994). Development of the 16 PF fifth edition. In S. R. Conn & M. L. Rieke (Eds.), *The 16 PF fifth edition technical manual* (pp. 3–20). Champaign, IL: Institute of Personality and Ability Testing.

Cattell, R. B. (1933). Temperament tests. II: Tests. *British Journal of Psychology, 23,* 308–329.

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology, 38,* 476–506.

Cattell, R. B. (1949). *The Sixteen Personality Factor Questionnaire* (1st ed.). Champaign, IL: Institute of Personality and Ability Testing.

Cattell, R. B. (1957). *Personality and motivation structure and measurement.* Yonkers-on-Hudson, NY: World Book.

Cattell, R. B., Cattell, A., & Cattell, H. E. P. (1993). *Sixteen Personality Factor Questionnaire* (5th ed.). Champaign, IL: Institute of Personality and Ability Testing.

Costa, P. T., Jr. (Ed.) (1991a). Clinical use of the five-factor model of personality [Special series]. *Journal of Personality Assessment, 57*(3), 393–464

Costa, P. T., Jr. (1991b). Clinical use of the five-factor model: An introduction. *Journal of Personality Assessment, 57,* 393–398.

Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory manual.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., & McCrae, R. R. (1992a). Four ways five factors are basic. *Personality and Individual Differences, 13,* 653–665.

Costa, P. T., Jr., & McCrae, R. R. (1992b). *NEO PI-R professional manual.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., Jr., & Widiger, T. A. (Eds.) (1994). *Personality disorders and the five-factor model of personality.* Washington, DC: American Psychological Association.

Craik, K. H. (1986). Personality research methods: An historical perspective. *Journal of Personality, 54,* 18–51.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302.

Dahlstrom, W. G., & Dahlstrom, L. E. (Eds.) (1980). *Basic readings on the MMPI: A new selection on personality measurement.* Minneapolis, MN: University of Minnesota Press.

Dahlstrom, W. G., Welsh, G. S., & Dahlstrom, L. E. (1972–1975). *An MMPI handbook.* (Vols. 1 & 2). Minneapolis, MN: University of Minnesota Press.

Digman, J. M. (1979, November). *The five major domains of personality variables: Analysis of questionnaire data in light of the five robust factors emerging from studies of rated characteristics.* Paper presented at the annual meeting of the Society of Multivariate Experimental Psychology, Los Angeles, CA.

Digman, J. M. (1994). Historical antecedents of the five-factor model. In P. T. Costa, Jr. & T. A. Widiger (Eds.), *Personality disorders and the five-factor model of personality* (pp. 13–18). Washington, DC: American Psychological Association.

Digman, J. M. (1996). The curious history of the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 1–20). New York: Guilford Press.

Elms, A. C. (1994). *Uncovering lives: The uneasy alliance of biography and psychology.* New York: Oxford University Press.

Erikson, E. H. (1950). *Childhood and society.* New York: Norton.

Erikson, E. H. (1958). *Young man Luther: A study in psychoanalysis and history*. New York: Norton.

Erikson, E. H. (1969). *Ghandi's truth: On the origins of militant nonviolence*. New York: Norton.

Eysenck, H. J. (1947). *Dimensions of personality*. London: Routledge & Kegan Paul.

Eysenck, H. J. (1953). *The structure of human personality*. New York: Wiley.

Eysenck, H. J. (1990). *Rebel with a cause*. London: Allen.

Eysenck, H. J. (1992). Four ways five-factors are *not* basic. *Personality and Individual Differences, 6,* 667–673.

Fairbain, W. R. D. (1952). *Psychoanalytic studies of the personality: The object relation theory of personality*. London: Routledge & Kegan Paul.

Fiske, D. W. (1949). Consistency of the factorial structure of personality ratings from different sources. *Journal of Abnormal and Social Psychology, 44,* 329–344.

Foa, U. G., & Foa, E. B. (1974). *Societal structures of the mind*. Springfield, IL: Charles C. Thomas.

Galton, F. (1888). Co-relations and their measurement. *Proceedings of the Royal Society, 45,* 135–145.

Gill, M. M. (Ed.) (1967). *The collected papers of David Rapaport*. New York: Basic Books.

Gill, M. M., & Klein, G. S. (1967). The structuring of drive and reality: David Rapaport's contributions to psychoanalysis and psychology. In M. M. Gill (Ed.), *The collected papers of David Rapaport* (pp. 8–34). New York: Basic Books.

Goffman, E. (1959). *The presentation of self in everyday life*. Garden City, NY: Doubleday Anchor.

Goldberg, L. R. (1977, August). *Language and personality: Developing a taxonomy of trait-descriptive terms.* Invited address to the Division of Evaluation and Measurement at the 86th annual meeting of the American Psychological Association, San Francisco.

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of personality and social psychology* (Vol. 2, pp. 141–165). Beverly Hills, CA: Sage.

Goldberg, L. R. (1990). An alternative "Description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology, 59,* 1216–1229.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4,* 26–34.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48,* 26–34.

Goldstein, K., & Scheerer, M. (1941). Abstract and concrete behavior: An experimental study with special tests. *Psychological Monographs, 53*(2) (Whole Issue No. 239).

Gough, H. G. (1946). Diagnostic patterns on the MMPI. *Journal of Clinical Psychology, 2,* 23–37.

Guilford, J. P. (1959). *Personality*. New York: McGraw-Hill.

Guilford, J. P., & Guilford, R. B. (1936). Personality factors S, E, and M, and their measurement. *Journal of Personality, 34,* 21–36.

Guilford, J. P., & Guilford, R. B. (1939). Personality factors D, R, T, and A. *Journal of Abnormal and Social Psychology, 34,* 21–36.

Gurtman, M. B. (1996). Interpersonal problems and the psychotherapy context: The construct validity of the Inventory of Interpersonal Problems. *Psychological Assessment, 8,* 241–255.

Guttman, L. (1954). A new approach to factor analysis: The radex. In P. R. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences* (pp. 258–348). Glencoe, IL: Free Press.

Hall, C. S., & Lindzey, G. (1978). *Theories of personality* (3rd ed.). New York: Wiley.

Hanfmann, E., & Kasanin, J. (1937). A method for the study of concept formation. *Journal of Psychology, 3,* 521–540.

Hartmann, H. (1939). *Ego psychology and the problem of adaptation.* New York: International Universities Press.

Hathaway, S. R. (1960). Forward. In W. G. Dahlstrom, G. S. Welsh, & L. E. Dahlstrom (Eds.), *An MMPI handbook: A guide to use in practice and research* (pp. vii–xi). Minneapolis, MN: University of Minnesota Press.

Hathaway, S. R. (1972). Where have we gone wrong? The mystery of the missing progress. In J. N. Butcher (Ed.), *Objective personality assessment: Changing perspectives* (pp. 21–43). New York: Academic Press.

Hathaway, S. R., & McKinley, J. C. (1940). A multiphasic personality schedule (Minnesota): I Construction of the schedule. *Journal of Psychology, 10,* 249–254.

Hathaway, S. R., & McKinley, J. C. (1942). A multiphasic personality schedule (Minnesota): III. The measurement of symptomatic depression. *Journal of Psychology, 14,* 73–84.

Hathaway, S. R., & McKinley, J. C. (1943). *The Minnesota Multiphasic Personality Inventory*. Minneapolis, MN: University of Minnesota Press.

Hathaway, S. R., & McKinley, J. C. (1983). *The Minnesota Multiphasic Personality Inventory manual.* New York: Psychological Corporation.

Helmes, E., & Reddon, J. R. (1993). A perspective on developments in assessing psychopathology: A critical review of the MMPI and MMPI-2. *Psychological Bulletin, 113,* 453–471.

Henry, W. P. (1996). The structural analysis of social behavior as a common metric for programmatic psychotherapy research. *Journal of Consulting and Clinical Psychology, 64,* 1263–1275.

Hogan, R. (1983). A socioanalytic theory of personality. In M. M. Page (Ed.), *1982 Nebraska symposium on motivation: Personality—current theory and research* (pp. 55–89). Lincoln, NE: University of Nebraska Press.

Hogan, R. (1986). *Hogan Personality Inventory manual.* Minneapolis, MN: National Computer Systems.

Hogan, R. (1994). Reinventing ourselves (Review of D. P. McAdams, *The stories we live by*). *Contemporary Psychology, 39,* 355–356.

Hogan, R. (1996). A socioanalytic perspective on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 163–179). New York: Guilford Press.

Hogan, R., & Hogan, J. (1992). *Hogan Personality Inventory manual.* Tulsa, OK: Hogan Assessment Systems.

Holt, R. R. (1951). The Thematic Apperception Test. In H. A. Anderson & G. L. Anderson (Eds.), *An introduction to projective techniques* (pp. 181–229). New York: Prentice-Hall.

Holt, R. R. (Ed.) (1967). Motives and thought: Psychoanalytic essays in honor of David Rapaport. *Psychological Issues, 5,* Monograph 18/19.

Holt, R. R., and Havel, J. (1960). A method for assessing primary and secondary process in Rorschach responses. In M. A. Rickers-Ovsiankina (Ed.), *Rorschach psychology* (pp. 263–315). New York: Wiley.

Horowitz, L. M. (1979). On the cognitive structure of interpersonal problems treated in psychotherapy. *Journal of Consulting and Clinical Psychology, 47,* 5–15.

Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureno, G., & Villasenor, V. S. (1988). The Inventory of Interpersonal Problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology, 56,* 885–892.

Hough, L. (1992). The "Big Five" personality variables—construct confusion: Description versus prediction. *Human Performance, 5,* 139–155.

Humm, D. G., & Wadsworth, G. W. (1935). The Humm-

Wadsworth temperament scale. *Journal of Applied Psychology, 92,* 163–200.

Jackson, D. N. (1971). The dynamics of structured personality tests: 1971. *Psychological Review, 78,* 229–248.

Jackson, D. N. (1984). *Personality Research Form manual* (3rd ed.). Fort Huron, MI: Research Psychologists Press.

Jackson, D. N., Paunonen, S. V., Fraboni, M., & Goffin, R. D. (1996). A five-factor versus six-factor model of personality structure. *Personality and Individual Differences, 20,* 33–45.

Jacobson, E. (1954). *The self and the object world.* New York: International Universities Press.

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66–100). New York: Guilford Press.

Johnson, J. H., Butcher, J. N., Null, C., & Johnson, K. N. (1984). Replicated item level factor analysis of the full MMPI. *Journal of Personality and Social Psychology, 47,* 105–114.

Jourard, S. M. (1964). *The transparent self.* Princeton, NJ: Van Nostrand.

Jung, C. G. (1971). *Psychological types* (H. G. Baynes, Trans.; revised by R. F. C. Hull). Princeton, NJ: Princeton University Press (Original work published 1923).

Kelly, G. A. (1955). *The psychology of personal constructs* (2 vols.). New York: Norton.

Kiesler, D. J. (1979). An interpersonal communication analysis of relationship in psychotherapy. *Psychiatry, 42,* 299–311.

Kiesler, D. J. (1983). The 1982 Interpersonal Circle: A taxonomy for complementarity in human transactions. *Psychological Review, 90,* 185–214.

Kiesler, D. J. (1988). *Therapeutic metacommunication: Therapist impact disclosure as feedback in psychotherapy.* Palo Alto, CA: Consulting Psychologists Press.

Kiesler, D. J. (1996). *Contemporary interpersonal theory and research: Personality, psychopathology, and psychotherapy.* New York: Wiley.

Kiesler, D. J., Goldston, C. S., & Schmidt, J. A. (1991). *Manual for the Check List of Interpersonal Transactions-Revised (CLOIT-R) and the Check List of Psychotherapy transactions-revised (CLOPT-R).* Richmond, VA: Virginia Commonwealth University.

Kiesler, D. J., & Schmidt, J. A. (1993). *The Impact Message Inventory: Form IIA Octant Scale version.* Palo Alto, CA: Mind Garden.

Kiesler, D. J., Schmidt, J. A., & Wagner, C. C. (1997). A circumplex inventory of impact messages: An operational bridge between emotion and interpersonal behavior. In R. Plutchik & H. R. Conte (Eds.), *Circumplex models of personality and emotions* (pp. 221–244). Washington, DC: American Psychological Association.

Kluckhohn, C., & Murray, H. A. (1953). Personality formation: The determinants. In C. Kluckhohn, H. A. Murray, & D. M. Schneider (Eds.), *Personality in nature, society, and culture* (pp. 53–67). New York: Knopf.

Kohut, H. (1971). *The analysis of the self: A systematic psychoanalytic approach to the treatment of narcissistic personality disorders.* New York: International Universities Press.

LaForge, R., Leary, T. F., Naboisek, H., Coffey, H. S., & Freedman, M. B. (1954). The interpersonal dimension of personality: II. An objective study of repression. *Journal of Personality, 23,* 129–153.

Landis, C., & Katz, S. E. (1934). The validity of certain questions which purport to measure neurotic tendencies. *Journal of Applied Psychology, 18,* 343–356.

Landis, C., Zubin, J., & Katz, S. E. (1935). Empirical validation of three personality adjustment inventories. *Journal of Educational Psychology, 26,* 321–330.

Leary, T. (1957). *Interpersonal diagnosis of personality.* New York: Ronald.

Lewin, K. (1938). *The conceptual representation and measurement of psychological forces.* Durham, NC: Duke University Press

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, 3,* 635–694.

Lorr, M., & McNair, D. M. (1963). An interpersonal behavior circle. *Journal of Abnormal and Social Psychology, 67,* 68–75.

MacKinnon, D. W. (1944). The structure of personality. In J. McV. Hunt (Ed.), *Personality and the behavior disorders* (Vol. 1, pp. 3–48). New York: Ronald.

MacKinnon, D. W. (1975). IPAR's contribution to the conceptualization and study of creativity. In I. A. Taylor & J. W. Getzels (Eds.), *Perspectives in creativity* (pp. 60–89). Chicago: Aldine.

Main, M., Kaplan, N., & Cassidy, J. (1985). Security in infancy, childhood, and adulthood: A move to the level of representation. *Monographs of the Society for Research in Child Development, 50,* 66–104.

Marks, P. A., & Seeman, W. (1963). *The actuarial description of abnormal personality: An atlas for use with the MMPI.* Baltimore, MD: Williams & Wilkins.

Mayman, M. (1967). Object representations and object relationships in Rorschach responses. *Journal of Projective Techniques, 31,* 17–25.

McAdams, D. P. (1988). Biography, narrative, and lives: An introduction. *Journal of Personality, 56,* 1–18.

McAdams, D. P. (1989). *Intimacy: The need to be close.* New York: Doubleday.

McAdams, D. P. (1993). *The stories we live by: Personal myths and the making of the self.* New York: William Morrow.

McAdams, D. P. (1994). *The person: An introduction to personality psychology* (2nd ed.). Fort Worth, TX: Harcourt Brace.

McAdams, D. P. (1996). Personality, modernity, and the storied self: A contemporary framework for studying persons (Target article). *Psychological Inquiry, 7,* 295–321.

McClelland, D. C. (1961). *The achieving society.* New York: Van Nostrand.

McCrae, R. R. (1991). The five-factor model and its assessment in clinical settings. *Journal of Personality Assessment, 57,* 399–414.

McCrae, R. R., & Costa, P. T., Jr. (1990). *Personality in adulthood: Emerging lives, enduring dispositions.* New York: Guilford Press.

McCrae, R. R., & Costa, P. T., Jr. (1996). Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives.* (pp. 51–87). New York: Guilford Press.

Meehl, P. E. (1945). The dynamics of "structured" personality tests. *Journal of Clinical Psychology, 1,* 296–303.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and review of the evidence.* Minneapolis, MN: University of Minnesota Press.

Meehl, P. E. (1989). Autobiography. In G. Lindzey (Ed.), *A history of psychology in autobiography* (Vol. 3, pp. 337–389). Stanford, CA: Stanford University Press.

Miller, T. R. (1991). The psychotherapeutic utility of the five-factor model of personality: A clinician's experience. *Journal of Personality Assessment, 57,* 415–433.

Millon, T. (1969). *Modern psychopathology.* Philadelphia: Saunders.

Morey, L. C. (1991). *Personality Assessment Inventory manual.* Odessa, FL: Psychological Assessment Resources.

Morey, L. C., Waugh, M. H., & Blashfield, R. K. (1985).

MMPI scales for *DSM-III* personality disorders: Their derivation and correlates. *Journal of Personality Assessment, 49,* 245–251.

Morgan, C. D., & Murray, H. A. (1935). A method for investigating fantasies: Thematic Apperception Test. *Archives of Neurology and Psychiatry, 34,* 209–306.

Murray, H. A. (1938). *Explorations in personality.* New York: Oxford University Press.

Murray, H. A. (1943). *The Thematic Apperception Test: Manual.* Cambridge, MA: Harvard University Press.

Murray, H. A. (1959). Preparations for the scaffold of a comprehensive system. In S. Koch (Ed.), *Psychology: A study of science* (Vol. 3, pp. 7–54). New York: McGraw-Hill.

Myers, I. B., & McCaulley, M. H. (1985). *Manual: A guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press.

Nichols, D. S. (1987). Interpreting the Wiggins MMPI content scales. In K. L. Moreland & J. N. Butcher (Eds.), *Clinical notes on the MMPI* (No. 10, pp. 3–26). Minneapolis: National Computer Systems.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology, 66,* 574–583.

Norman, W. T. (1967). *2800 personality trait descriptors: Normative operating characteristics in a university population.* Ann Arbor, MI: University of Michigan, Department of Psychology.

Office of Strategic Services Assessment Staff (1948). *Assessment of men.* New York: Rinehart.

Payne, F. D., & Wiggins, J. S. (1972). MMPI profile types and the self-report of psychiatric patients. *Journal of Abnormal Psychology, 79,* 1–8.

Pearson, K. (1896). Mathematical contributions to the theory of evolution: Regression, heredity, and panmixia. *Philosophical Transactions, 187a,* 253–318.

Pincus, A. L., & Wiggins, J. S. (1990). Interpersonal problems and conceptions of personality disorders. *Journal of Personality Disorders, 4,* 342–352.

Plutchik, R., & Conte, H. R. (Eds.) (1997). *Circumplex models of personality and emotions.* Washington, DC: American Psychological Association.

Plutchik, R., & Conte, H. R. (1986). Quantitative assessment of personality disorders. In R. Michels & J. O. Cavenar, Jr. (Eds.), *Psychiatry* (Vol. 1, pp. 1–13). Philadelphia: Lippincott.

Prelinger, E., & Zimet, C. N. (1964). *An ego-psychological approach to character assessment.* Glencoe, IL: The Free Press.

Rapaport, D. (1942). Principles underlying projective techniques. *Character and Personality, 10,* 213–219 (reprinted in Gill, 1967).

Rapaport, D. (1944–1946). *Manual of diagnostic psychological testing* (2 vols.). New York: Josiah Macy, Jr. Foundation.

Rapaport, D. (1946). Principles underlying nonprojective tests of personality. *Annals of the New York Academy of Sciences, 46,* 643–652 (reprinted in Gill, 1967).

Rapaport, D. (1951). The conceptual model of psychoanalysis. *Journal of Personality, 20,* 56–81 (reprinted in Gill, 1967).

Rapaport, D. (1960). The structure of psychoanalytic theory: A systematizing attempt. *Psychological Issues, 2,* Monograph 6.

Rapaport, D., Gill, M., & Schafer, R. (1946). *Diagnostic psychological testing* (2 vols.). Chicago: Year Book.

Rapaport, D., Menninger, K. A., & Schafer, R. (1947). The new role of psychological testing in psychiatry. *American Journal of Psychiatry, 103,* 473–476 (reprinted in Gill, 1967).

Rorschach, H. (1921). *Psychodiagnostik*. Bern, Switzerland: Bircher. (Trans. Hans Huber Verlag, 1942).

Runyan, W. McK. (1982). *Life histories and psychobiography: Explorations in theory and method.* New York: Oxford University Press.

Safran, J. D., & Segal, Z. V. (1990). *Interpersonal process in cognitive therapy*. New York: Basic Books.

Sarbin, T. R. (Ed.) (1986). *Narrative psychology: The storied nature of human conduct.* New York: Oxford University Press.

Saucier, G., & Goldberg, L. R. (1996). The language of personality: Lexical perspectives on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 21–50). New York: Guilford Press.

Schaefer, E. S. (1961). Converging conceptual models for maternal behavior and for child behavior. In J. G. Glidewell (Ed.), *Parental attitudes and child behavior* (pp. 124–146). Springfield, IL: Charles C. Thomas.

Schafer, R. (1948). *The clinical application of psychological tests.* New York: International Universities Press.

Schafer, R. (1954). *Psychoanalytic interpretation in Rorschach testing: Theory and application.* New York: Grune & Stratton.

Schafer, R. (1967). *Projective testing and psychoanalysis.* New York: International Universities Press.

Soldz, S., Budman, S., Demby, A., & Merry, J. (1993). Representation of personality disorders in circumplex and five-factor space: Explorations with a clinical sample. *Psychological Assessment, 5,* 53–63.

Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology, 15,* 201–293.

Stern, G. G., Stein, M. I., & Bloom, B. S. (1956). *Methods in personality assessment.* Glencoe, IL: Free Press.

Stewart, A. J., Franz, C., & Layton, L. (1988). The changing self: Using personal documents to study lives. *Journal of Personality, 56,* 41–74.

Strack, S. & Lorr, M. (1994). Introduction. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (pp. xiii–xviii). New York: Springer.

Strong, E. K., Jr. (1927). *Vocational Interest Blank (Form A).* Stanford, CA: Stanford University Press.

Sullivan, H. S. (1953). *The interpersonal theory of psychiatry.* New York: Norton.

Thurstone, L. L. (1934). The vectors of mind. *Psychological Review, 41,* 1–32.

Tomkins, S. S. (1947). *The Thematic Apperception Test.* New York: Grune & Stratton.

Tomkins, S. S. (1979). Script theory: Differential magnification of affects. In H. E. Howe, Jr. & R. A. Dienstbier (Eds.), *Nebraska symposium on motivation* (Vol. 26, pp. 201–236). Lincoln, NE: University of Nebraska Press.

Trapnell, P. D., & Wiggins, J. S. (1990). Extension of the Interpersonal Adjective Scales to include the Big Five dimensions of personality. *Journal of Personality and Social Psychology, 59,* 781–790.

Tupes, E. C., & Christal, R. E. (1961). *Recurrent personality factors based on trait ratings* (USAF ASD Tech. Rep. No. 61–97). Lackland Air Force Base, TX: US Air Force.

Vernon, P. E. (1950). *The structure of human abilities.* London: Methuen.

Waller, N. G., & Ben-Porath, Y. S. (1987). Is it time for clinical psychology to embrace the five-factor model of personality? *American Psychologist, 42,* 887–889.

Webb, J. T., Levitt, E. E., & Rojdev, R. (1993, March). *After three years: A comparison of the clinical use of the MMPI and MMPI-2.* Paper presented at the 53rd Annual Meeting of the Society for Personality Assessment, San Francisco, CA.

Wechsler, D. (1941). *The measurement of adult intelligence.* Baltimore: Williams & Wilkins.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (4th ed.). Baltimore: Williams & Wilkins.

Welsh, G. S. (1956). Factor dimensions A and R. In G. S. Welsh & W. G. Dahlstrom (Eds.), *Basic readings on the MMPI in psychology and medicine* (pp. 264–281). Minneapolis, MN: University of Minnesota Press.

Welsh, G. S., & Dahlstrom, W. G. (1956). *Basic readings on the MMPI in psychology and medicine*. Minneapolis, MN: University of Minnesota Press.

Westen, D. (1990). Psychoanalytic approaches to personality. In L. Pervin (Ed.), *Handbook of personality theory and research* (pp. 21–65). New York: Guilford Press.

Westen, D. (1991). Clinical assessment of object relations using the TAT. *Journal of Personality Assessment, 56,* 56–74.

White, R. W. (1952). *Lives in progress* (1st ed.). New York: Holt, Rinehart, & Winston.

White, R. W. (1966). *Lives in progress* (2nd ed.). New York: Holt, Rinehart, & Winston.

White, R. W. (1975). *Lives in progress* (3rd ed.). New York: Holt, Rinehart, & Winston.

White, R. W. (1981). Exploring personality the long way: The study of lives. In A. I. Rabin, J. Arnoff, A. M. Barclay, & R. A. Zucker (Eds.), *Further explorations in personality* (pp. 3–19). New York: Wiley.

Widiger, T. A., Trull, T. J., Clarkin, J. F., Sanderson, C., & Costa, P. T., Jr. (1994). A description of the *DSM-III-R* and *DSM-IV* personality disorders with the five-factor model of personality. In P. T. Costa, Jr. & T. A. Widiger (Eds.), *Personality disorders and the five-factor model of personality* (pp. 41–71). Washington, DC: American Psychological Association.

Wiggins, J. S. (1962). Strategic, method, and stylistic variance in the MMPI. *Psychological Bulletin, 59,* 224–242.

Wiggins, J. S. (1966). Substantive dimensions of self-report in the MMPI item pool. *Psychological Monographs, 80,* (22, Whole No. 630).

Wiggins, J. S. (1973a). Despair and optimism in Minneapolis (A review of J. N. Butcher (Ed.), Objective personality assessment: Changing perspectives). *Contemporary Psychology, 18,* 605–606.

Wiggins, J. S. (1973b). *Personality and prediction: Principles of personality assessment*. Reading, MA: Addison-Wesley.

Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology, 37,* 395–412.

Wiggins, J. S. (1984). Cattell's system from the perspective of mainstream personality theory. *Multivariate Behavioral Research, 19,* 176–190.

Wiggins, J. S. (1991). Agency and communion as conceptual coordinates for the understanding and measurement of interpersonal behavior. In W. Grove & D. Cicchetti (Eds.), *Thinking clearly about psychology: Essays in honor of Paul E. Meehl* (Vol. 2, pp. 89–113). Minneapolis, MN: University of Minnesota Press.

Wiggins, J. S. (1995). *Interpersonal Adjective Scales: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Wiggins, J. S. (1996). Preface. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. vii–xi). New York: Guilford.

Wiggins, J. S. (in press). *Paradigms of personality assessment*. New York: Guilford Press.

Wiggins, J. S., Phillips, N., & Trapnell, P. (1989). Circular reasoning about interpersonal behavior: Evidence concerning some untested assumptions underlying diagnostic classification. *Journal of Personality and Social Psychology, 56,* 296–305.

Wiggins, J. S., & Pincus, A. L. (1989). Conceptions of personality disorders and dimensions of personality. *Psychological Assessment, I,* 303–316.

Wiggins, J. S., & Pincus, A. L. (1992). Personality: Structure and assessment. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual review of psychology* (Vol. 43, pp. 473–504). Palo Alto, CA: Annual Reviews.

Wiggins, J. S., & Pincus, A. L. (1994). Personality structure and the structure of personality disorders. In P. T. Costa, Jr. & T. A. Widiger (Eds.), *Personality disorders and the five-factor model of personality* (pp. 73–93). Washington, DC: American Psychological Association.

Wiggins, J. S., & Trapnell, P. D. (1996). A dyadic-interactional perspective on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 88–162). New York: Guilford.

Wiggins, J. S., & Trapnell, P. D. (1997). Personality structure: The return of the Big Five. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 737–764). San Diego, CA: Academic Press.

Wiggins, J. S., & Trapnell, P. D., & Phillips, N. (1988). Psychometric and geometric characteristics of the revised Interpersonal Adjective Scales (IAS-R). *Multivariate Behavioral Research, 23,* 517–530.

Wiggins, J. S., & Trobst, K. K. (1997). Prospects for the assessment of normal and abnormal interpersonal behavior. *Journal of Personality Assessment, 68,* 110–126.

Winter, D. G. (1973). *The power motive*. New York: Free Press.

Woodworth, R. S. (1917). *Personal data sheet*. Chicago: Stoelting.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.13

Observations of Parents, Teachers, and Children: Contributions to the Objective Multidimensional Assessment of Youth

DAVID LACHAR
University of Texas-Houston Medical School, Houston, TX, USA

4.13.1 ISSUES IN OBJECTIVE MULTIDIMENSIONAL MULTISOURCE ASSESSMENT 372
- *4.13.1.1 Overview* 372
- *4.13.1.2 Multidimensional Measurement* 372
- *4.13.1.3 Multisource Assessment* 373
- *4.13.1.4 Psychometric and Technical Considerations* 374
  - *4.13.1.4.1 Scale stability and homogeneity* 374
  - *4.13.1.4.2 Scale construction methods and scale character* 375
  - *4.13.1.4.3 Evidence of scale and questionnaire validity* 376

4.13.2 MULTIDIMENSIONAL MULTISOURCE RATING SCALES 377
- *4.13.2.1 Overview* 377
- *4.13.2.2 Conners' Rating Scales-Revised* 378
  - *4.13.2.2.1 Conners' Parent Rating Scale-Revised* 378
  - *4.13.2.2.2 Conners' Teacher Rating Scale-Revised* 379
  - *4.13.2.2.3 Conners–Wells' Adolescent Self-Report Scale* 379
  - *4.13.2.2.4 Commentary* 380
- *4.13.2.3 Child Behavior Checklist/4–18 and 1991 Profile (CBCL); Teacher's Report Form and 1991 Profile (TRF); Youth Self-report and 1991 Profile (YSR)* 380
  - *4.13.2.3.1 Child Behavior Checklist/4–18 (CBCL) and 1991 Profile* 382
  - *4.13.2.3.2 Teacher's Report Form (TRF) and 1991 Profile* 383
  - *4.13.2.3.3 Youth Self-report (YSR) and 1991 Profile* 383
  - *4.13.2.3.4 Commentary* 384
- *4.13.2.4 Behavior Assessment System for Children* 385
  - *4.13.2.4.1 BASC Parent Rating Scales* 386
  - *4.13.2.4.2 BASC Teacher Rating Scales* 386
  - *4.13.2.4.3 BASC Self-Report of Personality (SRP)* 386
  - *4.13.2.4.4 BASC Monitor for ADHD* 387
  - *4.13.2.4.5 Commentary* 387
- *4.13.2.5 Personality Inventory for Children, 2nd Edition (PIC-2); Personality Inventory for Youth (PIY); Student Behavior Survey (SBS)* 388
  - *4.13.2.5.1 Personality Inventory for Children, 2nd Edition (PIC-2); Personality Inventory for Youth (PIY)* 389
  - *4.13.2.5.2 Student Behavior Survey* 397
  - *4.13.2.5.3 Commentary* 398

4.13.3 CONCLUSIONS 398
4.13.4 REFERENCES 399

4.13.1 ISSUES IN OBJECTIVE MULTIDIMENSIONAL MULTISOURCE ASSESSMENT

4.13.1.1 Overview

The use of objective questionnaires in the clinical assessment of child and adolescent adjustment continues to grow in popularity. The five years that followed my last review of objective assessment of children and adolescents (Lachar, 1993), have brought additional psychometric advances in instrument development and application. The importance of objective measures in current practice has also been demonstrated by the relative coverage provided in monographs and textbooks available for graduate education and continuing education (Kamphaus & Frick (1996) is an excellent example that supports this observation).

It is also important to consider the effect recent changes in the provision of health care may have upon the use of psychological testing in the diagnostic process. Managed care may influence the use of objective assessment in two quite different ways. First, in the pursuit of cost containment, the use of time-intensive individually administered psychological tests is less likely to be authorized in the routine evaluation of youth, quite independently of the documented value of such procedures. Relatively inexpensive diagnostic questionnaires, in contrast, are more likely to be accepted as a routine component of initial evaluations because they improve upon the efficiency of the intake data collection process. Second, these families of multidimensional measures can contribute substantial value to the program evaluation aspects of managed care. Objective questionnaires have the potential to independently document both the need for treatment as well as to quantify treatment efficacy through repeated application.

This chapter will expand upon many of the themes presented in Lachar (1993) and will describe in some detail four "families" of assessment instruments that meet current standards for intake assessment and treatment evaluation. These materials: (i) assess multiple dimensions of problem behavior, (ii) collect observations from parents, teachers, and youth, and (iii) provide standard scores based upon contemporary national samples. In 1993 only one set of measures met these criteria (Achenbach, 1992); at the conclusion of 1997 four sets of such scales were either commercially available or in press.

4.13.1.2 Multidimensional Measurement

Multidimensional instruments require standardized responses to similarly formatted descriptive statements which generate estimates of a variety of dimensions of child behavior or psychopathology. These estimates are usually visually displayed in the form of a profile of standard scores. These standard scores represent a common metric that has been derived from one normative reference sample or overlapping or similar samples, and are therefore directly comparable (vs. the case of scores from several single-dimension measures that have been standardized on normative samples collected in very different places and times, as well as under different conditions).

Advances in psychopathology and measurement technology demonstrate the clinical advantage of multidimensional measurement. It is the case, rather than the exception, in clinical assessment that several scaled dimensions of adjustment, such as depression and anxiety or inattention and noncompliance, are concurrently elevated in the clinical range. Not only are these dimensions—often classified into externalizing (behavioral excess) or internalizing (behavioral deficit) dimensions—problematic concurrently, the externalizing and internalizing diagnoses are often found to be comorbid. For example, adolescents hospitalized on a psychiatric unit are often described by more than one diagnosis or by a diagnosis that reflects different dimensions of psychopathology. It is not unusual for such adolescents to demonstrate significant levels of depression or anxiety and at the same time demonstrate pathologies of behavioral excess (oppositional defiant disorder, conduct disorder, attention deficit hyperactivity disorder (ADHD); see, for example, O'Connor, McGuire, Reiss, Hetherington, & Plomin, 1998). The most useful assessment procedure will provide a quantified estimate of the pattern of problems in need of remediation and answer relevant questions such as the role that family conflict may play in symptom development or maintenance, and the context of such symptoms.

Multidimensional assessment is uniquely designed to describe patterns of adjustment characterized by dimensions of both poor adjustment and relatively symptom-free adjustment. In many cases the documented absence of certain problems is as important as is the documented presence of others. The following common clinical challenge documents the importance of multidimensional assessment. Children and adolescents are often referred for evaluation by parents, educators, physicians, and mental health professionals because they have been observed to be hyperactive and/or inattentive. Exclusive application of homogeneous measures of ADHD or hyperactivity in such cases can only prove to be problematic. These referred children are often found to be "inattentive" for reasons other than ADHD (depression, anxiety, situational adjustment, learning disability, acquired cognitive disability, etc.). When significant elevation is not obtained on focused hyperactivity scales, guidance in the search for likely alternative diagnoses is not provided. On the other hand, youth found to meet all of the criteria of ADHD often present with coexisting conditions (see, for example, August, Realmuto, MacDonald, Nugent, & Crosby, 1996; Vaughn, Riccio, Hynd, & Hall, 1997). When such comorbid conditions remain unrecognized, their presence cannot be considered in treatment planning. In such cases, the resulting treatment may be compromised by the absence of such information.

4.13.1.3 Multisource Assessment

Although the evaluation of adjustment or psychopathology in adults usually relies on self-report, such self-description is rarely adequate in the evaluation of children and adolescents. Indeed, the context of assessment is fundamentally different for children and adolescents who are unlikely to refer themselves for evaluation or treatment and may not possess the academic, cognitive, or motivational status to complete a comprehensive self-report instrument. The first consideration in this regard is that young elementary school children, perhaps students in K through 3, may be unable to describe themselves adequately through response to questionnaire statements. These children are unlikely to have mastered the range of vocabulary necessary to adequately describe dimensions of adjustment; such language competence is not usually attained before the fourth or fifth grade.

Another consideration is the reality that youth are most often referred for evaluation because they are either noncompliant with the requests of the significant adults in their lives or exhibit problems in academic achievement, often presenting with inadequate reading skills. It is therefore not unusual that completion of a self-report inventory of several hundred items by a high school student (e.g., the Minnesota Multiphasic Personality Inventory-Adolescent consisting of 478 statements) could present quite an assessment challenge.

Parents and teachers not only refer youth for assessment, they are also the primary sources of useful systematic observation. Certainly adults should be the ideal informants to report noncompliance of a child to their requests. Parents are the only consistently available source for report of early childhood development and description of child behavior in the home. Teachers offer the most accurate observations of the age-appropriateness of a child's adjustment in the classroom and academic achievement, as well as of the attentional, motivational, and social phenomena unique to the classroom and to the school. It is likely, however, that such observational accuracy decreases after the elementary school grades, as teachers have very little continuous observation of students who appear for only 45 minutes a day in a classroom of 30 or more other students. (An exception perhaps would be the ratings of counselors, resource room teachers who work with small groups of students, and teachers in self-contained special education classrooms.) Youth self-description, regardless of problems that have been documented for this source of information (Greenbaum, Dedrick, Prange, & Friedman, 1994; Jensen et al., 1996), still must be seen as the most direct and accurate expression of personal thoughts and feelings. (Note that Michael & Merrell (1998) have demonstrated adequate short-term temporal stability for the self-report of third to fifth graders.)

The availability of a coordinated "family of instruments" in which uniquely designed questionnaires are completed by two or three different informant sources, each providing independent sets of child descriptions, offers a natural opportunity for critical comparison of the value of these sources of information. Achenbach, McConaughy, and Howell (1987) conducted a comprehensive literature review and found, although relatively greater between-source agreement was obtained for scales representing externalizing behaviors, very limited concordance in general between the report of parent, teacher, and youth. A review of similar studies that evaluated the responses to objective interviews from parent and child concluded that greater agreement was obtained for visible behaviors and for child–parent pairs with increasing child age (Lachar & Gruber, 1993).

Although one reasonable approach to the interpretation of differences between parent, teacher, and youth is to assign such differences to situation-specific variation (e.g., the child is actually oppositional at home, not in the classroom), other explanations are equally plausable. One alternative explanation of cross-informant variance requires a close examination of the measures being applied. It is a frequent occurrence to find that scales with similar names contain significantly different content. In this situation an apparent lack of agreement between informants on similarly named scales (such as "depression") may more accurately reflect the differences in statement manifest content. At one extreme a clinician might consider accepting any evidence of symptom presence from any informant source, at the other extreme it would also be possible to focus exclusively on problems demonstrated by at least two or even all three informant sources.

On the other hand, the development and application of valid identical across-source measures may not provide optimal assessment data. Such an approach may restrict the diagnostic potential of each informant source by excluding the measurement of attributes that may be uniquely obtained from only one informant source (e.g., a parent-completed measure of developmental delay, a teacher-completed measure of academic skills, or a youth-completed measure of achievement motivation).

Another explanation for poor cross-informant agreement for data gathered in a clinical context acknowledges the potential effect of response sets on the accuracy of such information. Starting with the child or adolescent being assessed, adequate compliance with questionnaire instructions may not be obtained—reflecting either inadequate language or reading skills, or lack of adequate motivation for the task. It is also likely that a youth may not want to share a personal history of maladaptive behavior and current internal discomfort with mental health professionals, although a negative presentation of parent adjustment and home conflict may be more readily provided. At times, youth may also be motivated to admit to problems and symptoms that are not present. These same motivations and conditions may also influence parent report. Indeed, there has been some concern that poor parent adjustment may compromise the validity of parent report (Achenbach, 1981). However, a subsequent review (Richters, 1992) and the specific analysis of this issue with the Personality Inventory for Children (PIC) (Lachar, Kline, & Gdowski, 1987) have found no empirical support for this consideration. Some instruments, however, incorporate validity scales to identify such response sets. These scales are designed to measure random or inadequate responses to scale statements, defensive denial of existing problems, as well as admission of symptoms that are unlikely to be present or exaggeration of actual problems. (Examples of the effect of informant defensiveness on profile pairs are presented in Table 5 and discussed in associated text for the Personality Inventory for Youth (PIY) and the Personality Inventory for Children, 2nd Edition (PIC-2). See also Lachar, in press.)

For the clinician there is a distinct pragmatic advantage in using an assessment system with comparable parent, teacher, and youth versions. Research may document one day that one source of descriptive report is superior to another in the assessment of a specific problem domain (only surveys of the opinions of clinicians and parents are currently available: Loeber, Green, & Lahey, 1990; Phares, 1997). Until that day, however, conditions regularly occur in the conduct of psychological evaluations that on some occasions make it necessary to collect one, and on other occasions another, of these parent-, teacher-, and youth-report questionnaires. Children may be too young, uncooperative, or language-impaired to complete a questionnaire. The evaluation may occur during the summer vacation, or the youth may have not consistently attended one school, or may have left school permanently, making collection of a teacher rating difficult or impossible. Parents may fail family appointments when their child is hospitalized, or children may be under agency guardianship rather than parent supervision. In such not infrequent cases, a "family" of objective questionnaires separately normed for parents, teachers, and youth provide the flexibility to facilitate the collection of data from the source or sources that are available.

4.13.1.4 Psychometric and Technical Considerations

The topics discussed in this section are unlikely to spark much enthusiasm in clinicians who also demonstrate little interest in reading the technical sections of test manuals. These issues, however, are important as they influence test validity and ease of interpretation and must be considered in the selection of instruments for specific clinical applications.

4.13.1.4.1 Scale stability and homogeneity

It may seem quite reasonable to apply the technical standards and statistical techniques first developed in the construction of ability and achievement measures to scales of adjustment and personality. For example, in the development of an age-normed measure of verbal IQ or a grade-normed measure of arithmetic calculation skill, certain assumptions are made. As each of these measures would be conceptualized as relatively homogenous and stable over at least weeks or months, item components would need to demonstrate substantial intercorrelation as well as little change over time (coefficent alphas and test–retest correlations of one week/three months of at least 0.80).

Measures of adjustment or psychopathology, in contrast, can vary in degree of homogeneity and may in fact in some cases be expected to be relatively unstable over time or sensitive to the effects of treatment. A measure of trait adjustment might consist of responses that describe previous behaviors and therefore demonstrate considerable temporal stability, while a measure of mood that consists of self-statements made in the present tense may provide a state measure that is expected to vary considerably over time. An empirically constructed scale in which items have been selected because they have statistically separated well-defined samples (e.g., normal adolescents vs. adolescents adjudicated delinquent) is likely to be multidimensional and limited in homogeneity. In contrast, a scale constructed by factor analysis consists of items that correlate highly with each other and such scales demonstrate substantial homogeneity. The importance of scale homogeneity as measured by coefficient alpha and scale stability represented by the test–retest correlation must be evaluated in light of the expressed meaning and application of each measure. Of importance in any case is the demonstrated validity of these scales; that is, evidence that each scale measures what it is supposed to and that interpretive guidelines are derived from empirically based relationships between scale values and independent observations.

4.13.1.4.2 Scale construction methods and scale character

It is important to be aware of both scale content and scale performance when applying scale standard scores in clinical assessment. Scale content has gained in judged importance and scale items have been selected frequently because of specific content or written to measure specific dimensions (see, for example, the Conners *DSM-IV* scales (Conners, 1997) and the Disruptive Behavior *DSM-IV* scales of the Student Behavior Survey (Lachar, Kline, Wingenfeld, & Gruber, 1999; Pisecco, Lachar, Gallen, Gruber, & Huzinec, 1998). This recent focus on item content is in sharp contrast to previously developed empirically keyed scales in which item performance took preference over item content. Current consensus is that scale content is necessary but insufficient to assure scale validity. It is recommended that clinicians first read the items on each scale and examine the content dimensions provided as well as the content equivalence for parent, teacher, and youth questionnaires before attempting any clinical application.

It is of course generally true that increased scale length will improve the depth or variety of expressed content. Against this reality must be balanced concerns about the time required of informants to complete a questionnaire, an issue that appears to be significant to psychologists who work in test development and those who work in test application. One must also consider the context in which evaluations are to be conducted. A baseline or initial evaluation would justify a more comprehensive evaluation than would a repeated evaluation or some approach to treatment evaluation in which daily or weekly observations are required.

Each clinician will need to judge the value of the information obtained vs. the time needed to generate these measures. In this author's experience, within a clinical (vs. research) context, parents view these questionnaires as appropriate and appreciate such an offer to make this sort of contribution to the diagnostic process. Teachers, unless asked to complete lengthy forms on a whole classroom of students, are quite willing to spend 15 or 20 minutes completing a form that either initiates a referral to a school psychologist or provides necessary classroom observation for an independent or agency-conducted evaluation. The instruments of the families of questionnaires presented in this chapter vary in length from a "short form" of 10 items that generates one standard score to a "short form" of eight 12-item shortened scales that is as long as the full-length versions of some other questionnaires.

Scale length (number of items) and rate of item endorsement (infrequently endorsed items in short scales) within normative and clinical samples directly influence scale performance, such as range of standard scores and the shape of score distributions. Response format is also relevant in that a scale score increment of one point may represent the endorsement of one additional symptom or characteristic (true/false; present/absent) or a subjective increase in symptom observed certainty or frequency (somewhat true to very true; sometimes true to often true). It may be useful to identify the minimum descriptive content necessary to generate a clinically significant standard score. It may also be useful to determine the amount of response difference necessary to move a scale value from normal limits to the clinical range. Note that a scale with a 0–18 raw score range can be generated from either six symptoms rated from 0 to 3, or from 18 true/false statements that vary in both symptom intensity and frequency. Although it will take longer to respond to 18 than to six items, it is obvious that the content basis for standard scores will vary substantially for these two measures. A comprehensive evaluation of any dimension is likely to require more than a handful of observations.

Although these scale scores are usually positively skewed (symptoms and dimensions of psychopathology are not normally distributed), scale distribution may be excessively restricted due to scale length, item response frequency, or normative sample character so that the scale mean and standard deviation that generate standard scores are very limited. (See Rowe & Rowe (1997) for a contemporary discussion of this issue with a focus on the Conners' Global Index.) For example, a scale with a normative mean of 1.13 and a standard deviation of 2.78 (Conners' Teacher Rating Scale-Revised (CTRS-R): Oppositional, female 12–14 years) would generate the following raw to T-score conversions: $0 = 46T$, $1 = 50T$, $2 = 53T$, $3 = 57T$, $4 = 60T$, $5 = 64T$, $6 = 69T$. Because the Oppositional scale consists of six items rated from 0 (not true at all/never, seldom) to 3 (very much true/very often, very frequent), the following interpretive issues are suggested. Considering that the scale raw scores for Oppositional can range from 0 to 18, responses of "not true" or "just a little true" predominate in the normative sample. Only a scale score of "0" extends below the mean. Of most interest, a clinically elevated score ($> 65T$) can be obtained by responding "just a little true" to each item, or "very much true" to only two statements. Please note that other CTRS-R scales actually obtain mean raw scores below 1.0 and standard deviations below 2.0, but the CTRS-R is far from the only questionnaire to exhibit these troublesome psychometric characteristics. Users are well advised to review a test's raw score-to-T score translation table or profile where the issue can be evaluated.

4.13.1.4.3 Evidence of scale and questionnaire validity

All too often test evaluations and reviews discuss internal consistency and reliability at length and then give short shrift to test validity. This is perhaps understandable in that it is easier to make comparisons and establish numerical guidelines for evaluating these characteristics. Unfortunately—and particularly so for instruments rating psychopathology or other infrequent behavior—the emphasis should be quite the opposite. Validity is far more difficult to define, measure, and establish, but it really is the only bottom line. Even a scale that demonstrates somewhat weak or even impaired reliability can be extremely valuable clinically if it can successfully identify a rare or transient characteristic such as violent labile aggression. It is necessary to establish the validity of specific applications of each scale before it can be interpreted with certainty. It should be understood that standard scores derived from the most representative and comprehensive of normative data simply provide the frequency or infrequency of a given score in this normative sample. Just because a pattern of scale standard scores is infrequent does not necessarily demonstrate that this pattern is clinically meaningful or significant. It is also important to acknowledge that scale validity is not a general attribute. The validity of a scale can only be demonstrated within a specific context for a specific purpose.

The construction or selection of a normative sample for each questionnaire is also a process requiring critical review—especially when questionnaires provide a variety of normative options (see, for example, Reynolds & Kamphaus, 1992). In the evaluation of school-aged children, scale raw scores may be converted to standard scores based upon normative (not receiving counseling or special education services), representative (incorporated an accurate proportion of all children), or referred (clinical or special education) samples. Normative samples also differ as to whether they are gender specific or combine the protocols of both boys and girls. It should also be noted that normative samples may be constructed to reflect child age. Such age-referenced samples may represent broadly defined groups such as preadolescent/adolescent, or provide norms that reflect a much more narrow age range (see Conners, 1997). The relative utility of any normative sample rests upon the demonstration that derived standard scores are comparable in validity. For example, the scales of the PIC have been found to be equivalent in their ability to predict external criteria for boys vs. girls and for children vs. adolescents (Kline & Lachar, 1992; Kline, Lachar, & Sprague, 1985). Test authors who promote the use of an atypical normative procedure should demonstrate that the resulting interpretive guidelines are as accurate as those obtained by more traditional means. For example, if a questionnaire provides both traditional gender-specific and novel mixed-gender norms, evidence should be provided that mixed-gender norms generate interpretive guidelines which are either more accurate or as accurate as more traditional gender-specific guidelines.

The construction of a new set of scales for a multidimensional inventory or set of inventories must follow a path in which fundamental evidence of validity must first be obtained. If obtained, additional study of more specific and focused evidence of accurate scale performance must be pursued. For example, when items are placed on dimensions on the basis of manifest content, the accuracy of this placement could be demonstrated through item-to-total correlations in which each item will be expected to obtain the greatest correlation with the dimension on which it has been placed. If the scales are constructed to measure adjustment or psychopathology, each scale could then be applied to regular education and special education and/or clinical samples to determine if meaningful separation occurs between these groups for each scale. Here "meaningful" is more than a statistically significant difference, which when samples are large could be so small as to be of no pragmatic value. Scales in this context should demonstrate at least a moderate effect size—generally about one-third of a standard deviation or greater (Cohen, 1988).

Unlike ability tests that gain initial construct validity by correlation with similar although established measures, scales of adjustment are less likely to demonstrate comparable concordance (0.70s–0.80s) because of differences in content and format. Indeed, substantial correlations obtained between pairs of scales from different tests completed by the same informant most likely reflect substantial shared item content. Adjustment scales for children and adolescents acquire validity through demonstration of external correlates that may take a variety of forms including independent ratings, special education placement, diagnoses, or specific treatment. For example, in students referred for evaluation, elevation of a parent-completed attention deficit hyperactivity scale should correlate with content-congruent teacher ratings, and students diagnosed with ADHD and assigned to receive stimulant therapy should obtain clinically elevated scale scores that subsequently fall within the normative range following successful treatment.

The accumulation of substantial construct validity and the development of empirically based interpretive guidelines are prerequisite for a questionnaire to achieve status as an established clinical measure. Unfortunately, the accumulation of a substantial bibliography provides an insufficient basis for the development of this status. Tests are far more often applied in the quantification of phenomena under study rather than applied in the study of scale score meaning. For example, contrast "How many children with mild mental retardation also demonstrate the *DSM-IV* criteria for a disruptive behavior disorder?" with "How do children with mild mental retardation who also obtain an elevated parent-informant hyperactivity scale differ in some way from children with mild mental retardation who do not obtain such an elevated scale?" That is, tests of adjustment are usually applied in the paradigm in which the investigator asks the question "What is the likelihood of a given test result given certain clinical criteria?" vs. "What is the likelihood of certain clinical phenomena given a certain test score, score range, or score pattern?" It is the latter question (and associated research design) that replicates the use of tests in clinical assessment and hence offers the better evidence of validity.

The clinician who reads that 95% of a sample of adolescents hospitalized on a psychiatric unit with a discharge diagnosis of major depression obtain an elevated score on a certain depression scale should understand that this information is insufficient to demonstrate either construct validity or provide an interpretive guideline. If, for example, a comparable proportion of the remaining adolescents on that psychiatric unit also received similar test results, these additional data would suggest that this "depression" scale measures general distress rather than depression. In addition, clinicians should pay attention to any evidence of a scale's ability to make clinically meaningful distinctions, not the separation of normal and referred subjects (see, for example, Forbes (1985) in which PIC profiles of children with the diagnosis of ADHD were compared to the PIC profiles of behaviorally disturbed children whose symptom presentation did not support such a diagnosis). It is also important to focus on studies that document the independent correlates of scale elevations or profile patterns (see, for example, LaCombe, Kline, Lachar, Butkus, & Hillman (1991) and the clinician, teacher, and self-report correlates of parent-report PIC-2 shortened scales presented in Table 4). In general, such studies assemble a sizable data set that includes both the test scores under study and other clinically meaningful independent information, with the goal of determining the degree to which test variables accurately predict clinically meaningful variables across all subjects. In this way a scale's convergent validity (scale elevation for patients afflicted with the problem to which the scale is directed) and discriminant validity (scale nonelevation for nonafflicted subjects or patients afflicted with unrelated problems) (Campbell & Fiske, 1959) are established.

4.13.2 MULTIDIMENSIONAL MULTISOURCE RATING SCALES

4.13.2.1 Overview

The second section of this chapter describes the four families of commercially available multidimensional, multisource questionnaires. Although considerable detail is provided for each set of scales, clinicians are urged to review all materials that may be appropriate for their specific applications. The conduct of this review has been difficult. Within these four sets of scales, one (Behavior Assessment System for Children; Reynolds & Kamphaus, 1992) was first published in 1992, two (Child Behavior Checklist and related measures: Achenbach, 1991; Conners Rating Scales (CRS): Conners, 1997) have recently been revised and expanded, while the fourth set of scales includes one recently published questionnaire (Personality Inventory for Youth: Lachar & Gruber, 1995), one questionnaire just revised (Personality Inventory for Children, 2nd Edition (PIC-2): Lachar, 1999), and one new measure (Student Behavior Survey (SBS): Lachar, Kline, Wingenfeld, & Gruber, 1999). Although the established materials have been collectively applied in approximately 3000 studies, no systematic review has been conducted to identify those studies that could form the basis for the development of interpretive guidelines (note in contrast the actuarial studies of profile scales and profile types for the PIC: Lachar & Kline, 1994).

4.13.2.2 Conners' Rating Scales-Revised

[Multi-Health Systems Inc., 908 Niagara Falls Boulevard, North Tonawanda, NY 14120-2060, USA]

The Conners scales were derived from teacher and parent measures first used in research conducted during the 1960s in the study of the pharmacological treatment of disruptive behaviors. The Conners' Rating Scale-Revised (CRS-R) emerged from a considerable research base generated over more than 30 years. An annotated bibliography of over 450 studies is available from the publisher (Wainwright & MHS Staff, 1996). Parent, teacher, and adolescent response to questionnaire items takes the same four-choice options: 0 = Not True at All (Never, Seldom), 1 = Just a Little True (Occasionally), 2 = Pretty Much True (Often, Quite a Bit), and 3 = Very Much True (Very Often, Very Frequent). CRS-R standard scores are linear *T* scores derived from contiguous three-year segments of the normative sample. The CRS-R hand scoring forms are easily completed and scored, and computer support for scoring and interpretation is available. The 1997 manual (Conners, 1997) and most current review chapter (Conners, in press) suggest that these scales continue to focus on ADHD and strengthen their assessment of related or comorbid disorders.

Conners in these publications emphasizes the relative importance of teacher observation in making these diagnostic determinations and suggests that the comparison of sources (teacher, parent, adolescent) can determine the reliability of the generated diagnostic hypotheses. Conners also suggests that the CRS-R scales should be augmented when their internalizing scales are elevated by using other measures, such as the Children's Depression Inventory (CDI) (Kovacs, 1992) and the Revised Children's Manifest Anxiety Scale (Reynolds & Richmond, 1985).

Conners details his concern for the accuracy of each protocol and the possible effect of response sets, such as defensiveness, although the CRS-R does not provide measures of such response sets. Conners suggests that it is useful to review the within-scale consistency of the content of items rated "3" (very much true/very often, very frequent) within elevated ($>64T$) scales/indices, although the efficiency of this process would be vastly improved through the development of a consistency scale. Although the manual suggests that a clinically elevated scale exceeds 64*T*, case examples present scale elevations below 60*T* as contributing to the diagnostic process. The manual presents and analyzes raw scores (mean, standard deviation) rather that providing *T*-score ranges and *T*-score frequencies within well-defined clinical or special education samples (e.g., establishing the scale *T*-score ranges for the dimensions that best define children diagnosed with ADHD).

4.13.2.2.1 Conners' Parent Rating Scale-Revised

The 80-statement form is the most comprehensive of the Conners' parent measures. Although the manual notes the reading level of this form as ninth grade, Conners' Parent Rating Scale-Revised (CPRS-R) items range in complexity from one-word stems such as "Irritable" to a 28-word descriptive phrase: "Does not follow through on instructions and fails to finish school work, chores or duties in the workplace (not due to oppositional behaviour or failure to understand instructions)." The CPRS-R generates seven factor-derived nonoverlapping scales apparently generated from the ratings of regular education students from the normative sample ages 3–17 years (Oppositional, Cognitive Problems, Hyperactivity, Anxious-Shy, Perfectionism, social problems, and psychosomatic), Conners' ADHD Index, Conners' Global Index, and *DSM-IV* scales Inattentive, Hyperactive-Impulsive, and Total. The ADHD Index consists of the 12 items empirically determined to best identify children at risk for an ADHD diagnosis (this 12-item set varies by informant source across this family of tests). The Conners' Global Index includes the

10 items historically labeled the "Hyperactivity Index" in previous published and unpublished versions of the Conners parent and teacher scales. This index is presented as the best general measure of behavioral adjustment and the one most sensitive to treatment effects.

Three additional forms are derived from these 80 items: (i) a shortened form of 27 items (CPRS-R:S) that generates the 12-statement ADHD Index and shortened forms of the Oppositional, Cognitive Problems, and Hyperactivity scales, (ii) a 26-item format that provides the ADHD Index and the three *DSM-IV* scales (CADS-P), and (iii) a sheet that only presents the 10-item Global Index. These three shortened forms are provided as options to monitor treatment effects.

Table 1 presents the pattern of scale and Index overlap (scales: Oppositional, 10 items; Anxious-Shy, 8 items; Perfectionism, 7 items; Social Problems, 5 items; and Psychosomatic, 6 items are each formed from unique items).

The *DSM-IV* Hyperactive–Impulsive and Hyperactivity scales exhibit considerable overlap as do the *DSM-IV* Inattentive and Cognitive Problems scales, suggesting the need for additional guidelines for application and interpretation of the full-length and shortened parent and teacher scales. It is also interesting to note that the Global Index, presented as the 10 items most sensitive to behavioral disturbance, almost never appears on any other scales (eight out of 10 items are unique).

4.13.2.2.2 Conners' Teacher Rating Scale-Revised

The full-length version of the Conners' Teacher Rating Scale-Revised (CTRS-R) consists of 59 items. Across-informant analysis may be facilitated in that the six CTRS-R factor-derived scales from three- to 17-year old normative children have been given the same names as the CPRS-R scales (there is not a teacher version of the CPRS-R Psychosomatic scale). The Conners' Global Index, Conners' ADHD Index, and three *DSM-IV* scales of the CPRS-R are also incorporated in this form. Three short teacher forms are also available: (i) a 28-item version that provides shortened Oppositional and Cognitive Problems scales, the Hyperactivity scale, and the ADHD Index, (ii) a 27-item version that provides the 12-item ADHD Index and the three *DSM-IV* scales, and (iii) a single sheet that presents the 10-item Global Index.

A comparison of CPRS-R and CTRS-R scales and index item content and scale length may shed some light on CRS-R construction and suggest issues in clinical application. Because the CTRS-R is shorter (59 items) than the CPRS-R (80 items), CTRS-R factor-derived scales are also shorter, with individual responses contributing greater to score derivation. The greatest shortening is present on factor scales Oppositional (10 to six items) and Cognitive Problems (12 to eight items). Assignment of identical titles to CPRS-R and CTRS-R scales should not lead the clinician necessarily to assume comparable item content for parent–teacher scale pairs. While the Global Index and *DSM-IV* scales do consist of items with identical or comparable content, the remaining CTRS-R score dimensions, except for the Cognitive Problems and Hyperactivity scales, consist of no more than 50% comparable content for similarly titled CPRS-R dimensions (six of the eight CTRS Cognitive Problems items have equivalents on the identically named 12-item CPRS dimension. Six of the seven CTRS Hyperactivity items have equivalents on the identically named nine-item CPRS dimension). Therefore, it was not unexpected to observe that the greatest parent–teacher agreement for normative cases occurred on the Global Index and *DSM-IV* measures (Connors, 1997, p. 126).

4.13.2.2.3 Conners–Wells' Adolescent Self-Report Scale

This 87-item questionnaire has been designed for students ages 12–17 with at least a sixth grade reading competence. Conner–Wells'

Table 1 CPRS-R item overlap.

Measure (no. of items/no. of unique items)	*B*	*C*	*H*	*K*	*L*	*M*
B:Cognitive problems (12/6)			3		5	
C:Hyperactivity (9/3)				1		5
H:Conners' ADHD Index (12/6)	3			1	2	2
K:Conners' Global Index (10/8)		1	1			
L:*DSM-IV* Inattentive (9/4)	5		2			
M:*DSM-IV* Hyperactive-Impulsive (9/2)		5	2			

Adolescent Self-Report Scale (CASS) includes six nonoverlaping factor-derived scales with lengths of either eight terms (Anger Control Problems; Hyperactivity) or 12 items (Family Problems; Emotional Problems, Conduct Problems, Cognitive Problems). This form can also be scored for a 12-item ADHD Index and two *DSM-IV* ADHD scales (as in the parent and teacher versions). Two shortened versions are available. In one, 27 items yield shortened versions of three factor-derived scales (Conduct Problems, Cognitive Problems, and Hyperactivity) and the 12-item ADHD Index, while the other provides the ADHD Index and the two *DSM-IV* scales Inattentive and Hyperactive-Impulsive in 30 items.

Comparison of item content between adolescent and parent report demonstrates excellent concordance for *DSM-IV* scales. Although reasonable agreement would be expected to be found for ADHD Indexes, Cognitive Problems and Hyperactivity scales, and between CPRS-S Oppositional and CASS Anger Control Problems, very few items from the parent scale can be identified in the self-report version.

4.13.2.2.4 Commentary

Table 2 displays average *T*-score group estimates derived from mean raw scores for parent, teacher, and self-report scales from three samples: general clinical, ADHD, and matched normative (Conners, 1997, pp. 135–136).

It would be very useful for the CRS-R Manual to provide for each sample the proportion of scale scores that equaled or exceeded the minimum *T*-score for the clinical range, whether this is 60*T* or 65*T*. Such lack of clinically relevant information is also demonstrated in a journal presentation of the six factor-derived CASS scales (Conners et al., 1997). In comparing 86 adolescents with a sole diagnosis of ADHD combined type to matched normative controls, an overall correct classification rate of 82.6% was obtained using all six scales. Although all six scales were statistically significant in group contrasts (demonstrating comorbidity and need for multidimensional assessment), no indication of the relative importance of each scale's contribution to this discrimination is given, nor are the proportion of elevated scores per scale provided for ADHD and contrast samples.

Table 2 documents substantial normative/clinical differences as well as the relative superiority of parent report when it is compared to the companion self-report form. Brevity and ease of application would recommend the CRS-R measures in monitoring the treatment of children with ADHD. Versions of *DSM-IV* scales for parent, teacher, and youth will allow the study of the accuracy of each informant source in diagnosing ADHD as well as the study of across-informant agreement.

The majority of CRS-R measures appear to have been developed out of distinctly different item pools, resulting in sets of scales that may complement rather than duplicate each other. It would be useful for the manual to explain why the CRS norms incorporate three-year intervals. The current profiles suggest in displayed raw-to-*T*-score conversion substantial and variable gender and age effects. The magnitude of this age effect is easily demonstrated by tracking on a profile the *T*-score equivalents for one scale raw score across the five sets of age norms. For example, a *DSM-IV* Hyperactive Impulsive raw score of 12 for males resulted in the following five *T* scores: 3–5, 60*T*; 6–8, 63*T*; 9–11, 66*T*; 12–14, 70*T*; 15–17, 81*T*. In contrast, the ADHD Index demonstrated far less age-related variation.

4.13.2.3 Child Behavior Checklist/4–18 and 1991 Profile (CBCL); Teacher's Report Form and 1991 Profile (TRF); Youth Self-report and 1991 Profile (YSR)

[University Associates in Psychiatry, 1 South Prospect Street, Burlington, VT 05401-3456, USA]

The CBCL, the first published measure of these three (other forms now cover preschool, direct observation, and young adults), was initially published in 1983 as a direct extension of the factor analytic study of child problems and symptoms first published in 1966 (Achenbach, 1966). The 1983 CBCL consisted of two parts, an assessment of competence and a series of factor-derived behavior problem scales constructed from 120 brief descriptions. Each checklist item is rated by a parent as either 0 = Not True (as far as you know), 1 = Somewhat or Sometimes True, or 2 = Very True or Often True, with several items requiring individual elaboration when these items are positively endorsed (scores of 1 or 2). Because of these 17–21 items in each checklist, the informant must be closely monitored and an experienced clinician must review these comments before each pamphlet is submitted to more automatic scoring procedures (Drotar, Stein, & Perrin, 1995). A companion teacher rating form became available in 1986 and a self-report checklist was published in 1987. These forms and subsequent revisions have been used in over 2000 citations (Vignoe & Achenbach, 1997). The popularity of the CBCL and related instruments in research application has had a significant influence on the study of child and adolescent psychopathology since 1986.

Table 2 Average approximate *T*-scores for ADHD, clinical (CLIN), and normative (NORM) samples.

	ADHD	*CLIN*	*NORM*
CPRS-R			
Oppositional	59	63	47
Cognitive Problems	67	57	46
Hyperactivity	69	58	48
Anxious-Shy	56	58	49
Perfectionistic	49	57	50
Social Problems	63	68	48
Psychosomatic	56	57	48
Global Index	67	64	46
ADHD Index	66	59	47
DSM-IV Total	69	59	47
CTRS-R			
Oppositional	61	64	47
Cognitive Problems	56	50	47
Hyperactivity	69	68	49
Anxious-Shy	62	62	46
Perfectionistic	56	62	49
Social Problems	51	58	53
Global Index	66	62	46
ADHD Index	65	58	48
DSM-IV Total	64	58	46
CASS			
Family Problems	53	53	46
Emotional Problems	56	56	47
Conduct Problems	57	54	45
Cognitive Problems	60	56	44
Anger Control Problems	57	57	47
Hyperactivity	57	52	45
ADHD Index	59	53	44

In each of these three forms competence has been measured through rationally developed scales that vary by informant source. The factor-derived dimensions of child problems and symptoms followed a similar general structure for each informant source (each designated an "axis" by Achenbach). Each form's profile recorded several narrow-band scales (designated "syndromes" by Achenbach), summative Internalizing and Externalizing scales derived from second-order factor analysis, and a Total Symptoms scale. A series of item analyses (principal components analysis followed by varimax rotation) was applied to independent clinical samples of boys or girls within several age-range defined samples, resulting in dimensions often unique for one gender or one age group, as well as scales with the same titles but differing in content. The products of what appears to be a direct (vs. guided) unreplicated application of a data reduction technique of scale construction proved to be problematic. These item analyses apparently identified at times either sample-specific or age-/gender-specific variance rather than broadly demonstrable dimensions. For example, only the profile for 12–16-year-old boys did not include a depression dimension. (Certainly this result could not be taken as support for the conclusion that adolescent boys do not demonstrate problematic depression!) Obvious problems presented themselves in applying these scales in research and clinical assessment. Repeated administration to the individual case over time could result in changes in scores as likely to be related to instrument or normative variation as to actual changes in subject status.

The 1991 revision, documented in five monographs totaling over 1000 pages, represented a major departure (Achenbach, 1991a, 1991b, 1991c, 1991d, 1993). This revision, in contrast to the earlier versions of these checklists, emphasized consistencies in scale dimensions and scale content across child age (4–18 years), gender, and respondent/setting. Because the actual checklist items remained essentially the same, this revision also provides substantial continuity, allowing either the continued use of the original sex-age-specific narrow-band scales, or the rescoring of previously obtained protocols to generate the new scales. (Readers may find it useful to first read Chapter 3 of Achenbach (1993) to gain an overall appreciation of these checklists.)

Of primary importance is that a series of within-instrument item analyses was conducted to obtain dimensions common across gender and age groups ("core syndromes"). Substantial samples of protocols from each form obtained from clinical and special education settings were evaluated (4455 CBCLs, 2815 TRFs, and 1272 YSRs). This current version also provides eight narrow-band scales for each informant derived from analysis of the 89 items common to parent-, teacher-, and self-report forms, as well as Internalizing, Externalizing, and Total Scales. The CBCL Internalizing scale score is derived from the responses of three narrow-band scales: Withdrawn, Somatic Complaints, and Anxious/Depressed. The Externalizing scale score includes the responses from two narrow-band scales: Delinquent Behavior and Aggressive Behavior. Three other common scales were not placed on these two broad-band dimensions: Social Problems, Thought Problems, and Attention Problems. Review of scale content reveals some item overlap across these eight scales within each form, substantial difference in scale length within a profile (e.g., TRF scales range from eight to 25 items), and scales that are identically named yet vary in content across informants. Close examination reveals that only one of these narrow-band scales ("cross-informant syndromes"), Somatic Complaints, presents the same nine items to parent, teacher, and youth.

The 1991 forms now rely on national vs. regional normative (nonreferred) samples. Although the CBCL and YSR are routinely self-administered in clinical application, the CBCL normative data and some undefined proportion of the YSR norms were obtained through interview of the appropriate informant. It may be that the interactive process of interview administration of the CBCL inhibited informant response to checklist items. For example, six of eight parent informant scales obtained average normative raw scores of less than 2 with associated restricted scale variance. It is important to note that increased problem behavior scale elevation reflects increased problems, although these scales do not consistently extend below 50*T*. Because of the idiosyncratic manner in which *T* scores are assigned to scale raw scores, it is useful to follow each manual's suggestion to use scale raw scores, rather than *T* scores, in research applications. It is difficult to determine the interpretive meaning of checklist *T* scores, the derivation of which has been of concern (Lachar, 1993; Kamphaus & Frick, 1996; Kline, 1994, 1995). Scale values under 70*T* are normalized and therefore compressed and unrepresentative of actual scale score distributions in the normative sample. Values above 70*T* have been arbitrarily assigned by the checklists' author with the goal of placing all possible scale scores evenly across available profile space and do not represent any specific probabilistic or statistical relationship with scores under 70*T*. The CBCL checklists do not provide any measure of informant response set or protocol validity. Not only is the identification of either symptom denial or random response problematic, but the *T* score assignment process above 70*T* can result in protocols that are the result of a response set to exaggerate problems but appear as valid estimates of child adjustment.

4.13.2.3.1 Child Behavior Checklist/4–18 (CBCL) and 1991 Profile

Parents complete a four-sided pamphlet and clinicians complete one of two two-sided profile forms (one for girls, one for boys). Norms are provided for two age ranges (4–11 and 12–18 for problem scales, 6–11 and 12–18 for competence scales) separate by gender. Rated competence items are organized by manifest content into three narrow scales (Activities, Social, and School) which are then summed into a total score. Parents are asked to list and then rate (frequency, performance level) a child's participation in sports, hobbies, organizations, and chores. Parents also describe the child's friendships, social interactions, performance in academic subjects, need for special assistance in school, and history of retention in grade. As standard scores for these scales increase with demonstrated ability, a borderline range is suggested at 30–33*T* and the clinical range is designated at $T < 30$. It has been suggested that a child's social and economic opportunities as well as race and ethnicity significantly affect these values (Drotar et al., 1995). These scales have also been compared to those of the PIC in their ability to predict adaptive level as defined by the Vineland Adaptive Behavior Scales (Pearson & Lachar, 1994).

The behavior problem scales consist of one to eight word descriptions ("Overtired," "Argues a lot," "Feels others are out to get him/her"). Appendix A of the 1991 CBCI provides scoring directions for the 21 items that incorporate informant individual comment. Clinicians must understand checklist content and child psychopathology to correctly score these items. For example, item 113 is "Please write in any problems your child has that were not listed above." The person who scores the checklist must review the up to three examples that can be elicited, determine if they duplicate other checklist content, and then score the one unique symptom that receives the highest rating (Very True or Often True rather than Somewhat or Sometimes True). The importance of this issue cannot be minimized as almost all items of the Thought Problems scale for all three informants are provided in this format. Syndrome T scores are interpreted as borderline at 67–70T and in the clinical range at $T > 70$. There is some evidence, however, which suggests that ADHD and comorbid conditions are more accurately identified with a lower cutting point (Biederman et al., 1993; Chen, Faraone, Biederman, & Tsuang, 1994; Faraone, Biederman, Weber, & Russell, 1998; Steingard, Biederman, Doyle, & Sprich-Buckminster, 1992). The Total Problem scale is interpreted as borderline at 60–63T and in the clinical range at $T > 63$. Although the summary Internalizing and Externalizing scales are discussed in terms of their relative elevation when compared to each other, absolute interpretive guidelines are not provided.

Profile and pamphlet formats facilitate review of response content at the item level. Although scales are developed to overcome the unreliability of item responses, CBCL content may be important for a number of reasons. On several scales a "very true or often true" response to as few as two or three statements places a scale T score in the clinical range. When scales are limited in length and relatively heterogeneous in content, the same scale elevation can be generated by quite different observational content. A casual review of scale content reveals some inconsistencies that follow into TRF and YSR dimensions. Obsessions and compulsions, often associated with anxiety disorders, join strange ideas, behaviors, and hallucinations on the Thought Problems dimension; "Nervous, highstrung, or tense" appears on Attention Problems as well as Anxious/Depressed; while "Unhappy, sad, or depressed" appears on Withdrawn as well as Anxious/Depressed. The 19–25 item Aggressive Behavior dimension appears to have a number of items that appear to be more readily associated with ADHD phenomena than any "aggressive" dimension: "Demands a lot of attention," "Showing off or clowning," "Talks too much," as well as two items only on the TRF: "Talks out of turn," and "Disrupts class discipline." These observations suggest that although a factor analysis can identify clinically meaningful dimensions, subsequent efforts may be necessary to assure the stability of item scale placement. (See Erford (1996) for an example of an attempt to replicate the structure of a short factor-derived teacher rating form.) It may also be useful to review the item by dimension correlation matrix that underlies each factor analysis to determine the appropriateness of each item's scale placement.

4.13.2.3.2 Teacher's Report Form (TRF) and 1991 Profile

The TRF emerged from the CBCL and the comparability of the 120 symptom/problem items is readily apparent, as are the eight syndrome, two broad-band, and one summary score. However, comparison of the competence sections of the TRF and the CBCL reveals a marked difference. TRF measures are based upon very limited data: an average rating of academic performance based on up to six subjects identified by the teacher, individual seven-point ratings on four topics (how hard working, behaving appropriately, amount learning, and how happy), as well as a summary score from these four items. The TRF designates a borderline interpretive range for the mean academic performance and the summary score of 40–37T, with clinical range of $T < 37$. Even a cursory review of the TRF pamphlet suggests that the equivalence of CBCL and TRF forms was given priority over collection of teacher observations that would be unique to the classroom and the school environment. If observations of social and academic skills were integrated into the TRF, the teacher could also provide estimates of adaptive behaviors necessary for success in the classroom.

4.13.2.3.3 Youth Self-report (YSR) and 1991 Profile

The elements of this self-report measure have most of the basic characteristics of the CBCL and TRF. The YSR requires a fifth grade reading ability and provides competence and problem behavior standard scores normed separately by gender for the age range of 11 through 18. Seven adaptive competency items are scored for Activities, Social, and a Total Competence scale. The YSR manual's Appendix A is necessary to score these items (the addition of scoring examples would be very useful). These seven multipart items tap competence and level of involvement in sports, activities, organizations, jobs, and chores. Items also provide self-report of academic achievement, interpersonal adjustment, and level of socialization.

The remainder of the pamphlet presents 120 statements in which 16 are socially desirable items endorsed by most students. These items replace potential YSR equivalents of CBCL items that presented either problem behaviors of very young children or school-related difficulties. The presence of these 16 unscored items represents an unused opportunity to develop a semantic inconsistency scale to measure random response or inadequate comprehension, or some sort of defensiveness measure. The lack of validity scales in this self-report of symptoms and problems is considered problematic by several reviewers (Kamphaus & Frick, 1996; Kline, 1995; Lachar, 1993). Two reviewers cite literature that documents the YSR's diagnostic insensitivity (Kamphaus & Frick, 1996; Merrell, 1994).

The eight problem scales range from 50 to $100T$, with $67–70T$ assigned the borderline range and $T > 70$ the clinical range. The Activities and Social Competency scales range from 55 to $20T$, while the Total Competence score is profiled from 10 to $80T$. Activities and Social are in the borderline range from 33 to $30T$, while the clinical range includes scales $< 30T$. On the Total Competence scale, T scores of 40–37 are designated as borderline, while T scores < 37 form the clinical range.

4.13.2.3.4 Commentary

The CBCL, TRF, and YSR format is responsible for both the broad success that these instruments have achieved as well as their shortcomings. Each form can be completed in 10–20 minutes and appears easily scored (if competence items and items requiring informant explanation are avoided). Brevity is especially attractive for research applications in which subject and human subject research committee tolerance for multiple measures is of specific concern. Scoring software provided by the CBCL/TRF/YSR author facilitates not only the processing of individual instruments, but provides estimates of agreement over multiple reports on the same instrument, as well as agreement across different instruments.

It is almost certain that the CBCL, TRF, and YSR common format will stimulate meaningful across-informant research that will document the effect of item and scale content as well as informant characteristic (gender, age, ethnicity, clinical status) on both degree of agreement and relative contribution to the diagnostic process. Such research should generate useful guidelines for the clinical application of these instruments. On the other hand, the primary focus on parallel dimensions and consistency in the problem behavior scale construction technique limits collection of informant-specific observation. The other instrument families presented in this chapter vary in dimensions assessed to some degree to facilitate the collection of such information.

The decision to revise scales without modification of the original problem statement pool assured continuity but at the same time restricted potential improvement. These forms continue to demonstrate their original limitations in psychometric character. It is difficult to understand why the author has taken an idiosyncratic and statistically unsupportable approach both in subjecting summations of symptoms to a normalization process and in providing profiles that assign T scores to all physically possible raw scores. Certainly problematic item performance in the form of restricted range in symptom frequency within these normative samples is at the core of attempts to modify score distributions to improve psychometric performance. The absence of validity scales continue to present random, defensive, and exaggerated protocols as profiles that should be interpreted rather than excluded from this process.

The most remarkable observation, in contrast to the continuing research growth of CBCL/TRF/YSR application, is the apparent absence of interpretive guidelines for the clinician (certainly percentiles derived from normative samples do not provide evidence of scale meaning beyond that suggested by manifest item content). These 1991 manuals present as primary evidence of validity that items and scales differentiate clinical and normative samples. Such evidence is only clinically useful if such a decision is frequently made in practice. (Clinicians know whether a student has or has not been referred for an evaluation.) Certainly there should be sufficient opportunity within 2000 published studies to document the answers to specific clinically relevant questions. (Indeed, it is quite clear that much of the results of research in child psychopathology over the last 15–20 years have been dependent on the structure and psychometric characteristics of the CBCL.) It should be quite easy to identify and summarize those studies that used a contrasted-groups design to identify those scales that are effective in making specific distinctions.

A review of CBCL citations would easily find carefully selected samples relevant to clinical practice and scale validity. These protocols could then be scored for 1991 dimensions and subsequent profile types to establish the proportion of each sample that is elevated on each dimension or obtains a specific classification for each instrument. For example, one investigation would be presented in tabular form by profile scales that would be represented by 11 columns (eight narrow-band, two broad-band, one summary) and samples that would be represented by individual rows. In this manner, one could determine not only if Anxious/Depressed is usually clinically elevated for students who receive solitary or combined anxiety and/or depression diagnoses, but that an Anxious/Depressed clinical elevation is infrequent in samples that are not characterized by either depression or anxiety. A review of this literature may also demonstrate which scales may be used to differentiate between conditions frequently addressed in the diagnostic process, as well as external correlates of these scales. (Apparently some of the best available published support for scale validity is confined to contrasted group analysis instead of continuous heterogeneous clinical samples and to the collection of nonindependent criteria from the same informant who completed the CBCL measure: see for example, Edelbrock & Costello, 1988; Weinstein, Noam, Grimes, Stone, & Schwab-Stone, 1990.)

4.13.2.4 Behavior Assessment System for Children

[American Guidance Service Inc., 4201 Woodland Road, Circle Pines, MN 55014-1796, USA]

The Behavior Assessment System for Children (BASC) differs from the three other systems presented in this chapter in that the entire system (Parent Rating Scales, PRS; Teacher Rating Scales, TRS; Self-report of Personality, SRP) was published at one time. (The BASC also includes a developmental history form and a classroom observation form, the Student Observation System). The BASC conveniently provides one integrated manual for all rating instruments (Reynolds & Kamphaus, 1992). Considering that these materials first became available in 1992 (with the advanced computer program first available in 1994), its application in the literature has been limited, although assessment applications continue to be developed. For example, the BASC Behavior Monitor for ADHD, brief parent and teacher forms intended for repeated use to demonstrate whether the behaviors of children already diagnosed as ADHD change as a result of treatment, became available in 1998.

In a manner similar to the Achenbach and the Conners' scales, the BASC ratings completed by parent, teacher, and youth are marked directly on self-scoring pamphlets or on one-page forms that allow the recording of responses for subsequent computer entry. Each of these forms is relatively brief (126–186 items) and can be completed in 10–30 minutes. PRS and TRS items are rated on a four-point frequency scale (never, sometimes, often, and almost always), while SRP items are rated as either true or false. Raw scale scores are easily obtained from the self-scoring form, although selecting and transferring scale and composite standard scores in the form of linear T scores from 77 pages of conversion tables appears to require significant effort and concentration. Computer scoring and interpretation greatly facilitate this process.

For all three informants, scales are provided that measure clinical dimensions and adaptive dimensions (positive, desirable). All items are placed on only one substantive scale following an iterative process in which item content, item to scale correlations, normative–clinical endorsement differences and factor structure determined scale structure. In addition, these scales are combined on the basis of their intercorrelations into broad-band measures or scale composites, Externalizing Problems, Internalizing Problems, and Adaptive Skills. The relative diagnostic utility of scales vs. composites remains to be demonstrated. Children of school age are evaluated by the use of parent (PRS) and teacher (TRS) forms designated for ages 6–11 or 12–18. Scale item composition differs by both informant and age group, although substantial similarity is demonstrated across age groups within each informant-specific form. The Self-report of Personality (SRP) is similarly divided into forms for ages 8–11 and 12–18. In contrast to the PRS and TRS that emphasize across-informant similarities, the SRP has been designed to complement parent and teacher reports as a measure focused upon emotions and self-perceptions rather than as reports of overt behaviors.

BASC forms are also notable for their incorporation of validity scales to estimate the accuracy of obtained standard scores as well as provision of a brief Critical Items set. (The BASC manual does not, however, explain the rationale for item selection or provide relevant item performance statistics for these critical items.) Three sets of norms are provided: gender-specific (including representative proportions of students receiving special education services), combined gender, and clinical.

The use of a combined gender normative sample to generate standard scores is the exception to general assessment practice, as most measures of adjustment and personality compare each student to same-gender norms. It would seem that such a procedure would further minimize the significance of disruptive behaviors in girls and internalizing symptoms in boys, as well as present boys as relatively lacking in adaptive competence when compared to girls (see BASC manual, Figure 11.2). One could easily calculate the effect of any cutting point, such as $T > 60$, on each BASC scale using both normative approaches within a sample of referred students. After students would be so identified, the accuracy of this classification could be established using information independent of BASC data.

In contrast to the BASC comprehensive normative sample, the sample used to generate clinical standard scores appears to both be limited in size (in some age/form samples less than 100 students) and restricted in composition. Yet test materials provide complete support for scoring results on clinical norms. Clinical norms are only routinely used for assessment measures that cannot be effectively applied within general population samples (e.g., a rating scale for severely disturbed hospitalized patients in which 99% of a regular population would obtain a summary score of 0). It is also difficult to define the parameters necessary to construct a "representative" clinical sample. Certainly TRS, PRS, and SRP samples of 30 or fewer children diagnosed as depressed or autistic would be inadequate for anything but initial exploration of scale or composite construct validity.

4.13.2.4.1 BASC Parent Rating Scales

PRS scale composites and associated narrow-band scales are Internalizing Problems (Anxiety, Depression, and Somatization), Externalization Problems (Hyperactivity, Aggression, and Conduct Problems), and Adaptive Skills (Adaptability (ages 6–11 only), Social Skills, and Leadership). Additional profiled scales include Atypicality, Withdrawal, and Attention Problems. The PRS also provides a Behavioral Symptoms Index, consisting of those clinical scales common to all age levels on both the PRS and TRS that load the highest on the first unrotated factor in a common-factor analysis. This combination of scales is presented in the manual as a measure of overall level of problem behavior. The BASC manual suggests that clinical scales should be considered of potential significance when $T > 60$, while adaptive scale scores should receive diagnostic attention when $T < 40$.

4.13.2.4.2 BASC Teacher Rating Scales

TRS scale composites and associated narrow-band scales include Internalizing Problems and Externalization Problems composites with the same narrow-band scales as for the PRS. The TRS Adaptive Skills composite includes the PRS scale components as well as a Study Skills scale that provides unique classroom observation (e.g., "Reads assigned chapters," "Studies with other students"). Both TRS forms provide a School Problems Composite that incorporates scales Attention Problems and Learning Problems ("Does not complete tests," "Makes careless errors"). In this manner the TRS provides 21 or 22 items that are unique to the classroom.

Both the PRS and TRS incorporate a "fake bad" (F) index to assess the possibility that a teacher or parent rated a child in an inordinately negative fashion. The F index includes 16–20 items per form that represent either maladaptive behaviors to which the respondent answered "Almost always" or adaptive behaviors to which the respondent answered "Never." These extreme scores (neither item nor scale statistics are provided in the manual) may represent either problem exaggeration or the presence of severe behavioral disturbance.

4.13.2.4.3 BASC Self-Report of Personality (SRP)

SRP scales and scale composites vary from both PRS and TRS dimensions. Narrow-band scales Attitude to School, Attitude to Teachers, and (only for ages 12–18) Sensation Seeking form the School Maladjustment Composite. Scales Atypicality, Locus of Control, Social Stress, Anxiety, and (only for ages 12–18) Somatization form the Clinical Maladjustment Composite. A Personal Adjustment Composite is formed by scales Relations with Parents, Interpersonal Relations, Self-Esteem, and Self-Reliance. Two other unassigned narrow-band scales are included: Depression and Sense of Inadequacy. An additional measure, the Emotional Symptoms Index (ESI) is, according to the BASC manual, the SRP's most global indicator of serious emotional disturbance, particularly internalized disorders. The ESI is composed of two scales from the Clinical Maladjustment composite (Social Stress and Anxiety), two scales from the Personal Adjustment composite (Interpersonal Relations and Self-Esteem), and the two clinical scales that do not appear on any SRP composite (Depression and Sense of Inadequacy). Interpretive guidelines presented in the BASC manual generally suggest that scores of 60–69T suggest an at-risk or mild/moderate range and scores above 69T a "clinical range" interpretation for clinical scales. Because the adaptive scales are worded and scored in the positive direction, low elevations are given interpretive significance, at comparable ranges of 31–40T and below 31T.

The SRP also provides three scales to measure the presence of response sets that are likely to compromise scale validity. As in the PRS and TRS, the SRP includes an F-Index composed of infrequent responses that may reflect symptom exaggeration ("Nobody likes me," "Nothing about me is right"). In addition, a 14-item L-Index has been incorporated to measure problem denial ("I always go to bed on time," "My parents are always right"). The third measure, V-Index, is not a scale but a small set of items that, when responded to in the True direction, are likely to reflect either lack of cooperation or inadequate comprehension ("I have never been in a car," "Television does not really exist").

4.13.2.4.4 BASC Monitor for ADHD

The BASC Monitor (Kamphaus & Reynolds, 1998) consists of brief parent and teacher forms that, along with the BASC Student Observation System and other variables defined by the clinician, provide a multimethod system designed to evaluate the effectiveness of treatments used with ADHD. The primary purpose is to demonstrate whether the behaviors of children already diagnosed as ADHD change as a result of treatment. These forms are intended for repeated use with the same child. Items are derived from 1992 PRS and TRS forms, other items collected during form development but not placed on a 1992 scale, as well as a few items written for this specific purpose that reflect *DSM-IV* diagnostic criteria. Norms were directly derived or extrapolated from BASC data and are organized into five age groups in order to be sensitive to age effects: 4–5, 6–7, 8–11, 12–14, and 15–18.

The Parent Monitor Ratings (PMR) consists of 46 items and the Teacher Monitor Ratings (TMR) consists of 47 items. All items are rated on the BASC four-choice frequency format of Never, Sometimes, Often, or Almost Always. Each form provides scales named Attention Problems, Hyperactivity, Internalizing Problems, and Adaptive Skills, and a listing of *DSM-IV* Items, some of which appear on other PRS and TRS scales, had been studied but not retained, or were written specifically to reflect *DSM* criteria. Item content was selected to be appropriate across the 4–18 years span, and some variation between PMR and TMR was allowed to reflect the unique opportunities for observation associated with each rater (e.g., TMR: "Bothers other children when working," PMR: "Fiddles with things while at meals").

4.13.2.4.5 Commentary

It is difficult for a newly published set of instruments to compete with those with a 20- to 30-year history of application. Certainly the BASC cannot draw upon an extensive research literature to demonstrate scale validity and to suggest appropriate clinical and research applications. The BASC family of instruments are attractive for several reasons, however, especially if all three ratings (parent-, teacher-, and self-report) are obtained for a student. Scale and composite distributions appear to be less skewed than many Achenbach and Conners' measures. Cluster analysis of the TRS normative sample, for example, resulted in seven clusters. In these clusters mean clinical scale *T* scores ranged from a low 40s to a high at or above 70, while mean adaptive scale *T* scores ranged from a low of 33 to a high of 60 (Kamphaus, Huberty, DiStefano, & Petoskey, 1997). The second reason, SRP *T* scores easily are obtained below 50*T*, in contrast to the YSR, and substantial item content is usually represented in *T* scores within the "clinical range." Another reason to consider these scales is that application of the BASC in school psychology settings in which ease of collection of both teacher and parent ratings is greater than in other settings would take advantage of the complementary nature of the SRP when the PRS and TRS are routinely collected. The final reason to consider the BASC is that the SRP scales appear especially attractive in the dimensions they generate, if parent and teacher ratings are available to document comorbid disruptive behavior phenomena. For example, it would be clinically meaningful in an adolescent described by parent and teacher as aggressive and conduct disordered if he would describe himself as seeing others as responsible for his problems (Locus of Control) and as focused on the experiences of thrill-seeking and excitement (Sensation Seeking).

As with any first manual, however, the psychologist will be left with many questions regarding application of BASC scales to the individual student. First, it is unclear from what source interpretive guidelines are derived, unless from scale item content. For example, elevation of SRP Sensation Seeking is presented to be associated with potential for alcohol and drug use or experimentation, while elevation of SRP Social Stress in young children with scores higher than 70*T* suggests children who "may turn inward in an unsuccessful attempt to cope with these tensions" (*BASC manual*, pp. 60–61). Perhaps these conclusions are not derived from scale correlates, but from an extrapolation of the phenomena associated with the clinical dimension measured by a given scale.

Although validity scales are provided, especially for SRP, the psychometric details of their construction and their performance in data presented in the manual are not provided. Does the L-Index correlate negatively with SRP clinical scales and positively with SRP adaptive scales? Were the clinical samples presented in the manual first screened by these scales to exclude potentially invalid protocols? The evidence of scale validity primarily takes the form of factor structure and correlations with other published measures. Readers of the manual will not find a series of case studies, nor are tabled and profiled mean scale values for a variety of referred samples presented in terms of proportion that exceed some value presented as clinically meaningful. In this regard, it is most interesting to contrast manual tables 12.28, 13.25, and 14.22 in that not one SRP clinical scale obtains a mean of at least $60T$, while only one adaptive scale score obtains a mean of $40T$ or less. Similar normal limits results were obtained for a relatively small sample of children in state custody due to abuse or neglect, although the L-Index was apparently not scored (Dalton, 1996).

A brief review of published studies using the PRS and TRS, in contrast, provides some encouraging evidence of scale validity. BASC parent and teacher scales demonstrated convergent validity with CBCL and TRF scales (Vaughn, Black, Hall, Hynd, & Riccio, 1995). PRS scales were found to demonstrate convergent validity in nonreferred kindergarten students with scales of the Social Skills Rating System for Children (Flanagan, Alfonso, Primavera, & Povall, 1996) and demonstrated similar validity in elementary school students at risk for conduct disorder with scales of the CBCL (Doyle, Ostrander, Skare, Crosby, & August, 1997). In a study of cognitive problem-solving, increasing PRF Hyperactivity Scale elevation was associated with more problem solutions with aggressive or noncompliant content as well as fewer reasons or consequences justifying the selection of the "best solution" (Bloomquist, August, Cohen, Doyle, & Everhart, 1997). Four TRS scales (Aggression, Conduct Problems, Depression, and Social Skills) significantly classified ADHD and nondisabled students, as well as ADHD students with and without a comorbid condition (Lett & Kamphaus, 1997). TRS Learning Problems correlated in referred elementary students with the Freedom from Distractibility factor, although no relation was obtained for the Attention Problems or Hyperactivity scales (Lowman, Schwanz, & Kamphaus, 1996). In addition, TRS Externalizing Problems correlated significantly with a peer rating dimension "aggressive–disruptive," while TRS Internalizing Problems correlated significantly with a peer rating dimension "sensitive–isolated" (Realmuto, August, Sieler, & Pessoa-Brandao, 1997). In a longitudinal study, teacher/parent composite measures of Hyperactivity/Attention Problems and Aggression/Conduct Problems were strongly intercorrelated and negatively related to adaptive functioning 3.5 years later (August, MacDonald, Realmuto, & Skare, 1996).

4.13.2.5 Personality Inventory for Children, 2nd Edition (PIC-2); Personality Inventory for Youth (PIY); Student Behavior Survey (SBS)

[Western Psychological Services, 12031 Wilshire Boulevard, Los Angeles, CA 90025-1251, USA]

Forty years ago two University of Minnesota psychologists began the development of a new inventory approach to the evaluation of children and adolescents. They assembled a 600-item administration booklet and named it the "Personality Inventory for Children. For use with children from six through adolescence." The directions stated that each item was to be answered as either "true" or "false" by the child's mother in order to describe both the child and family relationships. Professors Wirt and Broen accumulated administration booklet descriptive statements following an outline. To ensure comprehensive coverage of child behavior and adjustment, 50 items were sorted into 11 separate content areas: aggression, anxiety, asocial behavior, excitement, family relations, intellectual development, physical development, reality distortion, social skills, somatic concern, and withdrawal. To these 550 potential scale items, 50 items were added in an effort to strengthen or clarify the meaning of certain areas of concern. Following many of the general procedures employed in development of the Minnesota Multiphasic Personality Inventory, PIC scales were identified from this item pool over a span of 20 years.

The initial 1977 published profile included three validity scales, a general screening scale, and 12 measures of child ability and adjustment and family function developed through either empirical item selection techniques or through iterative content valid procedures. In 1981 the administration booklet was revised and items were sorted into one of four parts. Completion of part I (items 1–131) allowed the scoring of four broad-band factor-derived scales (Lachar, Gdowski, & Snyder, 1982). Completion of parts I and II (items 1–280) generated the entire clinical profile with "shortened scales" (Lachar, 1982), while parts I–III (420 items) allowed the scoring of the original length scales. The final 180 items of this booklet were eventually dropped from the administration booklet because they did not appear on any of the standard full-length profile scales.

From the beginning, the task of PIC scale and profile interpretation attributed relatively little importance to item content (except for the construction of a Critical Items list), but instead established external correlates and interpretive guidelines for individual profile scales (Lachar & Gdowski, 1979) and replicated profile patterns (Gdowski, Lachar, & Kline, 1985; Kline, Lachar, & Gdowski, 1987; LaCombe et al., 1991). A diagnostic procedure in which similarity coefficients are calculated between the individual PIC profile to be interpreted and the mean profiles of students receiving specific special education services has also been incorporated into profile scoring and interpretation software (Kline, Lachar, Grubber, & Boersma, 1994). Special effort has focused on determining that PIC scale validity is not restricted by a child's age, gender, or ethnicity status (Kline & Lachar, 1992; Kline et al., 1985).

The development efforts for the PIC-2 started in 1989 with the rewriting of the first 280 items of the PIC booklet into a self-report format for the PIY (Lachar & Gruber, 1993, 1995). Many dimensions were easily translated from the "my child" to the "I" format, although this process was either difficult or impossible for some dimensions. For example, accurate self-report of early developmental phenomena and severe developmental delay would be unlikely or inappropriate for students who can complete a comprehensive self-report inventory ("My child was difficult to toilet train," "My child can comb his/her own hair"). Such PIC items were replaced with other self-report statements likely to be used in validity scales or to supplement coverage of dimensions such as depression and defective reality testing (see PIY technical manual for an elaboration of this process). The PIY was developed with a subscale-within-scale structure that facilitates profile interpretation and has also incorporated a set of validity scales constructed from the responses of children and adolescents.

Development of the PIY facilitated concurrent critical review of the structure and content of PIC scales and the PIC profile. This review motivated substantial data collection using a PIC research edition administration booklet that allowed both the scoring of the PIC and the collection of data on revised and new inventory items. Data collected from application of this booklet have been subjected to considerable statistical analysis through which the PIC (now PIC-2) and PIY have achieved a great deal of similarity to facilitate comparison. The PIC validity scales, already a significant component of profile interpretation, have been improved in the PIC-2 and one consistent clinical scale and subscale construction methodology has been applied to the assessment of all revised clinical dimensions. In addition, a national normative sample has been collected concurrently with that for the third diagnostic component, a new teacher rating form, the Student Behavior Survey (SBS). Although these revisions have increased the focus on content validity in this family of parent, teacher, and self-report questionnaires, the 20-year tradition of deriving interpretive guidelines from the empirical relations between test measures and independent phenomena of adjustment and behaviour continues in the PIC-2 and SBS.

4.13.2.5.1 Personality Inventory for Children, 2nd Edition (PIC-2); Personality Inventory for Youth (PIY)

The PIC-2 and PIY profiles graph the linear *T* scores of three parallel validity scales and nine substantive clinical scales. PIC-2 norms are provided for students in grades K though 12 (a preschool version is under development), while PIY norms are available for students in grades 4 through 12. Clinical scales were constructed using an iterative process in which initial item placement based upon previous PIC scale placement and/or substantive item content was then empirically supported within item-to-scale correlation matrices based upon large clinical samples. Items retained in each final PIC-2 and PIY clinical scale demonstrated both statistical significance and the most substantial item-to-scale correlation among the nine potential scale placements. All other items, unless placed on an inventory validity scale or a critical item list, were dropped from the final version of each inventory. This process, in contrast to item selection through contrasted groups, placed an emphasis on item content and scale homogeneity.

The majority of PIY clinical scale items (86–100%, mean = 95%) appear on the identically named PIC-2 clinical scale, facilitating comparison of parent- and youth-report. The items of the PIY clinical scales are placed on only one scale and one subscale. Of the 264 items that comprise the nine PIC-2 clinical scales, only 16 (6.1%) appear on two scales. Each clinical scale is also divided into two or three nonoverlapping subscales. Guided by a series of item factor analyses, these subscales consist of item subsets that are more homogeneous than their associated total scales. Clinical and validity scale content remains the same across child age or gender, allowing expected age- and gender-related variance to appear. (For example, greater elevation on the Impulsivity and Distractibility scale will be found for boys in comparison to girls and for younger children in comparison to older children.) PIY clinical scales become interpretable at 60*T*, and subscales usually at 65*T*. (Analyses of PIC-2 scale standard scores suggest comparable guidelines.) Table 3 presents the scale/subscale structure, representative items, and internal consistency estimates of the PIC-2 parent-report and PIY self-report clinical scales.

Table 3 demonstrates that PIC-2 and PIY item content varies from the description of commonly occurring difficulties to statements suggestive of severe psychopathology. It should be noted that many PIC-2 and PIY subscales incorporate more descriptors than the full scales found in the other three families of objective questionnaires.

The PIY and PIC-2 both incorporate a screening or short assessment procedure. The first 80 items of the PIY comprise a 32-item screening scale chosen to provide an optimal identification of those regular education students who—when administered the full PIY—produced clinically significant results. These items also include three "scan items" for each scale. Scan items were selected in such a manner so that individuals who endorse two or more of each set of three items would be those with a high probability of scoring $> 59T$ on the corresponding clinical scale. Shortened versions of three validity scales can also be derived from these items.

The PIC-2 provides a short form to measure change in clinical status associated with therapeutic intervention. Although PIC scales may have demonstrated such sensitivity to change (see therapeutic case example in Lachar & Kline, 1994), a brief form tailored specifically for this purpose was constructed. Selected items were (i) written in the present tense, (ii) frequently endorsed in the context of clinical assessment, and (iii) described clinical phenomena often the focus of short-term intervention. Using these guidelines, the 12 most favourable items from each of eight PIC-2 clinical scales were selected. (Cognitive Impairment was excluded due to historical components, lack of appropriate therapeutic focus, and the global/stable nature of most descriptions on this dimension.) These 96 inventory statements have been placed at the beginning of the 275-item PIC-2 booklet to serve as both a short form and a method of efficient re-evaluation of a child following short-term intervention. These 96 items are also available in a self-scoring format. It is intended for these scale scores to be graphed on the same profile at baseline and at appropriate interim and post-treatment intervals to demonstrate both dimensions of change and stability.

The initial concurrent validity of these shortened scales was established through correlation of these scale scores with clinician ratings, teacher descriptions, and self-report descriptions. These data were drawn from an ongoing clinical project in which the PIY, SBS, clinician ratings, diagnoses, and individually administered psychometric results are being collected to some degree with over 1000 PIC-2 assessments. Even though an exploratory analysis, some care was taken in the selection of these correlates. Each obtained scale descriptor was placed on only the one scale with which it received the largest correlation, all being at least significant at $p < 0.01$. Table 4 summarizes the number of ratings identified in this manner from each source (clinician, teacher, and student) and provides three examples from each rating source for each of these eight scales. Correlations between these shortened scales and their full-length versions are also presented.

Table 4 demonstrates that these brief scales correlated substantially with full-length versions and obtained independent correlates from nonparent observers that matched expressed scale content and diagnostic intent. Clinicians provided the greatest support for ADH, DLQ, and DIS, focusing on problems of disruptive and noncompliant behavior and intense and dysphoric affect that often form the basis of clinical referral. These analyses demonstrate that ADH, as previously demonstrated for the PIC Hyperactivity scale (Lachar & Gdowski, 1979), assesses those behaviors most related to problems in classroom adjustment. In addition, observations obtained directly from the student under study provide those internal and subjective judgments that demonstrate the clinical value of PIC-2 dimensions which do not receive robust correlates from clinicians or teachers.

PIY and PIC-2 profiles incorporate three validity scales. The Inconsistency scale (INC) evaluates the likelihood that responses to items are random or reflect in some manner inadequate comprehension of inventory statements or compliance with test instructions. The Defensiveness scale (DEF) identifies profiles likely to demonstrate the effect of minimization or denial of current problems. The third validity scale, Dissimulation (FB), identifies profiles that may result from either exaggeration of current problems or a malingered pattern of atypical or infrequent symptoms.

PIY and PIC-2 Inconsistency scales measure semantic inconsistency through the classification of response to 35 pairs of highly correlated items drawn from all nine clinical scales (for example, "I have many friends/I have very few friends"; "My child has a lot of talent/My child has no special talents"). For each item pair, two response combinations are consistent and two are inconsistent. Each inconsistent pair

Table 3 PIC-2 and PIY Clinical Scale and Subscale Composition.

Cognitive Impairment PIC-2: 39 items, alpha = .881
PIY: 20 items, alpha = .740
Items in common: 20 (only 1 on PIC/2 COG3)

PIC-2 Subscales:
COG1/Inadequate Abilities (13 items, alpha = .792)
My child is rather absent-minded.
COG2/Poor Achievement (13 items, alpha = .785)
Reading has been a problem for my child.
COG3/Developmental Delay (13 items, alpha = .799)
My child could eat with a fork before age four years.

PIY Subscales:
COG1/Poor Achievement and Memory (8 items, alpha = .649)
School has been easy for me.
COG2/Inadequate Abilities (8 items, alpha = .673)
Other people think that I am talented.
COG3/Learning Problems (4 items, alpha = .441)
I have been held back a year in school.

Impulsivity and Distractibility PIC-2: 27 items, alpha = .921
PIY: 17 items, alpha = .773
Items in common: 17

PIC-2 Subscales:
ADH1/Disruptive Behavior (21 items, alpha = .917)
My child jumps from one activity to another.
ADH2/Fearlessness (6 items, alpha = .693)
My child will do anything on a dare.

PIY Subscales:
ADH1/Brashness (4 items, alpha = .535)
I like to show off.
ADH2/Distractibility/Overactivity (8 items, alpha = .613)
I cannot keep my attention on anything.
ADH3/Impulsivity (5 items, alpha = .543)
I often act without thinking.

Delinquency PIC-2: 47 items, alpha = .957
PIY: 42 items, alpha = .923
Items in common: 39

PIC-2 Subscales:
DLQ1/Antisocial Behavior (13 items, alpha = .881)
My child has been in trouble with the police.
DLQ2/Dyscontrol (17 items, alpha = .921)
When my child gets mad, watch out!
DLQ3/Noncompliance (17 items, alpha = .927)
My child often disobeys me.

PIY Subscales:
DLQ1/Antisocial Behavior (15 items, alpha = .831)
I have run away from home.
DLQ2/Dyscontrol (16 items, alpha = .839)
What people say often makes me angry.
DLQ3/Noncompliance (11 items, alpha = .828)
I tend to see how much I can get away with.

Family Dysfunction PIC-2: 25 items, alpha = .882
PIY: 29 items, alpha = .869
Items in common: 25

PIC-2 Subscales:
FAM1/Conflict Among Members (15 items, alpha = .847)
Out family argues a lot at dinner time.
FAM2/Parent Maladjustment (10 items, alpha = .781)
One of the child's parents drinks too much alcohol

Table 3 (continued).

PIY Subscales:
FAM1/Parent–Child Conflict (9 items, alpha = .817)
I am unhappy about my home life.
FAM2/Parent Maladjustment (13 items, alpha = .738)
My parents often argue.
FAM3/Marital Discord (7 items, alpha = .701)
My parents are now divorced or living apart.

Reality Distortion PIC-2: 29 items, alpha = .898
PIY: 22 items, alpha = .831
Items in common: 20

PIC-2 Subscales:
RLT1/Developmental Deviation (14 items, alpha = .842)
My child does not understand other people.
RLT2/Hallucinations and Delusions (15 items, alpha = .817)
My child sometimes sees things that are not there.

PIY Subscales:
RLT1/Feelings of Alienation (11 items, alpha = .768)
I am different from most kids.
RLT2/Hallucinations and Delusions (11 items, alpha = .709)
People secretly control my thoughts.

Somatic Concern PIC-2: 28 items, alpha = .829
PIY: 27 items, alpha = .853
Items in common: 24

PIC-2 Subscales:
SOM1/Psychosomatic Preoccupation (17 items, alpha = .793)
My child seems tired most of the time.
SOM2/Muscular Tension and Anxiety (11 items, alpha = .675)
Recently my child has complained of chest pains.

PIY Subscales:
SOM1/Psychosomatic Syndrome (9 items, alpha = .730)
I often have headaches.
SOM2/Muscular Tension and Anxiety (10 items, alpha = .740)
At times I have trouble breathing.
SOM3/Preoccupation with Disease (8 items, alpha = .596)
I am worried about disease.

Psychological Discomfort PIC-2: 39 items, alpha = .909
PIY: 32 items, alpha = .864
Items in common: 31

PIC-2 Subscales:
DIS1/Fear and Worry (13 items, alpha = .728)
My child is often afraid of little things.
DIS2/Depression (18 items, alpha = .881)
My child tends to feel sorry for himself/herself.
DIS3/Sleep Disturbance/Preoccupation with Death (8 items, alpha = .770)
My child's sleep is calm and restful.

PIY Subscales:
DIS1/Fear and Worry (15 items, alpha = .779)
I worry about things that adults worry about.
DIS2/Depression (11 items, alpha = .730)
The future looks good to me.
DIS3/Sleep Disturbance (6 items, alpha = .703)
I often get up at night.

Social Withdrawal PIC-2: 19 items, alpha = .796
PIY: 18 items, alpha = .800
Items in common: 18

Table 3 (continued).

PIC-2 Subscales:
WDL1/Social Introversion (11 items, alpha = .781)
  My child worries about talking to others.
WDL2/Isolation (8 items, alpha = .656)
  My child often stays in his/her room for hours.

PIY Subscales:
WDL1/Social Introversion (10 items, alpha = .781)
  Shyness is my biggest problem.
WDL2/Isolation (8 items, alpha = .591)
  I keep my thoughts to myself.

Social Skill Deficits PIC-2: 28 items, alpha = .915
PIY: 24 items, alpha = .855
Items in common: 23
PIC-2 Subscales:
SSK1/Limited Peer Status (13 items, alpha = .853)
  My child often brings friends home.
SSK2/Conflict with Peers (15 items, alpha = .888)
  Other children often get mad at my child.

PIY Subscales:
SSK1/Limited Peer Status (13 items, alpha = .788)
  I am sure of myself in a group.
SSK2/Conflict with Peers (11 items, alpha = .804)
  Other kids make fun of my ideas.

Note: coefficient alphas based upon 1178 PIY and 901 PIC-2 protocols obtained in the context of clinical evaluation.

identified in a given protocol contributes one point to the INC raw score. Application of a cutting raw score of > 12 resulted in correct identification of 90–95% of clinical protocols and 92–96% of a random sample. The DEF is an expanded version of the PIC Lie scale. DEF items represent denials of common problems ("Sometimes I put off doing a chore. False"; "My child almost never argues. True") and attributions of improbable positive adjustment ("My child always does his/her homework on time. True"; "I am almost always on time and remember what I am supposed to do. True"). DEF elevations above 59*T*, even in hospitalized psychiatric patients, result in profiles that either minimize current problems or consistently deny the presence of most or all problems in adjustment.

The third validity scale, Dissimulation (FB), was empirically constructed through item analysis comparing clinical protocols and two sets of protocols completed by nonreferred regular education students or their mothers. First the PIY or PIC-2 was completed with directions to provide an accurate description, then a second questionnaire was completed as if the student was in need of mental health counseling or hospitalization. Selected items in the scored direction were very infrequent in the accurate normal (PIY, 11%; PIC-2, 4%) and clinical (PIY, 18%; PIC-2, 15%) condition, while characteristic (PIY, 83%; PIC-2, 55%) of the "fake bad" or dissimulated condition. The pattern of these three scales readily identifies profiles in which caution must be applied to their interpretation. Table 5 presents PIC-2 and PIY profile scores from three hospitalized 12-year-old adolescents.

The evaluation of Case A resulted in PIC-2 and PIY validity scale results which did not suggest that the accuracy of these profiles had been compromised in any systematic manner. It is important to observe that both profiles present clinically elevated scale scores. First consider similarities in scale and subscale pattern and scale and subscale clinical elevations (e.g., apparent disagreement for COG3 represents the reality that PIC-2 COG3 consists of parent report of developmental delay, while PIY COG3 represents learning problems—a dimension much more similar to PIC-2 COG2). Case A is a 12-year-old in his fifth psychiatric hospitalization who had been placed in a self-contained special education classroom for behavioral maladjustment. He carries diagnoses of ADHD combined type, Oppositional Defiant Disorder, and Conduct Disorder. His behavior at home and in the hospital demonstrated serious behavioral dyscontrol. He had been noncompliant in taking psychotropic medication to improve his emotional and behavioral control. He had threatened to kill himself (see

Table 4 Correlates of PIC-2 Brief Clinical Scales.

Impulsivity and Distractibility (ADH12: .96[a])

Clinician Ratings: total = 14
Hyperactive, overactive
Restless
Frequently frustrated

Teacher Ratings: total = 34
Disobeys class or school rules
Impulsive, acts without thinking
Misbehaves unless closely supervised

Student Ratings: total = 9
Brags about being sent to the principal
School has sent notes home about bad behavior
Likes to show off

Delinquency (DLQ12: .93)

Clinician Ratings: total = 44
Argues, Oppositional
Disobedient to parents
Poorly modulated anger

Teacher Ratings: total = 11
Does not demonstrate polite behavior
Blames others for his/her problems
Becomes upset for little or no reason

Student Ratings: total = 24
Gives parents a lot of trouble
Sometimes swear at parents
Ran away from home

Family Dysfunction (FAM12: .93)

Clinician Ratings: total = 6
Conflict between parents
Parent divorce/separation
Emotionally abused

Teacher Ratings: total = 6
Uses alcohol or drugs
Strikes or pushes school personnel
Preoccupied with sex

Student Ratings: total = 18
Family does not enjoy being with each other more than other families
A lot of tension in the home
Parents often argue

Reality Distortion (RLT12: .95)

Clinician Ratings: total = 3
Auditory hallucinations
Inappropriate emotion, affect
Delusions, paranoia

Teacher Ratings: total = 0

Student Ratings: total = 13
Don't get along with others most of the time
Need a lot of help from others
Do strange or unusual things

Somatic Concern (SOM12: .92)

Clinician Ratings: total = 6
Excessive sleeping
Sexually abused
Continually tired, listless

Table 4 (continued).

Teacher Ratings: total = 0

Student Ratings: total = 33
Often get very tired
At times have trouble breathing
Don't have as much energy as most other kids

Psychological Discomfort (DIS12: .93)

Clinician Ratings: total = 22
Depressed, sad, unhappy
Moodiness
Inadequate self-esteem

Teacher Ratings: total = 4
Appears sad or unhappy
Doesn't take successes and failures in stride
Worries about little things

Student Ratings: total = 25
Shy with kids my own age
Not often in a good mood
Often afraid of little things

Social Withdrawal (WDL12: .96)

Clinician Ratings: total = 4
Withdrawn
Shy
Uncommunicative, seldom talks

Teacher Ratings: total = 2
Pessimistic about the future
Overcritical of himself/herself

Student Ratings: total = 12
Shyness is my biggest problem
Hardly ever talk
Stay in the house for days at a time

Social Skill Deficits (SSK12: .96)

Clinician Ratings: total = 5
Teased by peers
Isolated, few or no friends
Attends self-contained special education class

Teacher Ratings: total = 17
Does not seem to have fun
Avoids social interaction in class
Unaware of the feelings of others

Student Ratings: total = 28
Often rejected by other kids
Have very few friends
Not very popular with other kids

---

[a] Indicates correlation between total and 12-item version of PIC-2 clinical scale in a sample of 905 referred students. Note text for explanation of correlate selection procedure.

DIS3 scores) and had assaulted his mother who he said did not want him at home (see PIY FAM1 and FAM2). This patient attempted to escape from the hospital and required multiple time-outs and seclusions to control rages, threats, and aggressive and inappropriate behavior (see DLQ2). Off medication, A demonstrated an attention span of less that 15 minutes (see ADH values). He had a history of impulsive and disruptive behavior, fighting with peers, noncompliance with adults, verbal and physical aggression, and running away from home (elevated DLQ values for PIY and PIC-2).

Case B, in contrast to Case A, documents considerable parent/child disagreement. The

Table 5 The influence of respondent defensiveness on PIY/PIC-2 profile pairs.

	Case A		*Case B*		*Case C*	
Scale/subscaleNORM	*PIC-2*	*PIY*	*PIC-2*	*PIY*	*PIC-2*	*PIY*
Inconsistency	53	49	67	57	50	68
Dissimulation	60	72	81	48	47	72
Defensiveness	30	39	30	64	65	50
Cognitive Impairment	61	63	67	57	43	47
COG1	56	57	68	52	49	42
COG2	67	50	70	45	41	45
COG3	53	85	48	85	43	67
Impulsivity and Distractibility	79	69	81	39	44	64
ADH1	75	77	83	49	46	53
ADH2	81	62	64	37	41	62
ADH3		57		41		66
Delinquency	83	75	86	41	42	49
DLQ1	71	75	63	43	46	49
DLQ2	82	71	98	41	43	56
DLQ3	79	67	76	43	42	42
Family Dysfunction	60	69	58	57	67	49
FAM1	58	69	58	52	62	47
FAM2	60	74	54	57	72	49
FAM3		53		58		53
Reality Distortion	67	71	73	48	50	79
RLT1	69	64	83	53	55	74
RLT2	62	75	56	41	44	79
Somatic Concern	82	75	74	49	41	65
SOM1	82	66	76	38	42	59
SOM2	72	71	65	53	41	65
SOM3		73		59		65
Psychological Discomfort	90	59	97	63	53	68
DIS1	70	64	81	64	51	66
DIS2	88	38	92	52	49	58
DIS3	83	70	83	63	63	68
Social Withdrawal	47	83	85	41	46	59
WDL1	46	77	72	46	49	53
WDL2	50	78	88	38	42	65
Social Skill Deficits	57	43	86	56	49	57
SSK1	43	40	68	53	50	56
SSK2	74	50	97	59	48	55

Note: Table 3 provides the names and examples of PIC-2 amd PIY subscales.

only clear agreement as to problem area is in academic achievement (PIC-2 COG2, PIY COG3). History and psychometric assessment document retention in grade, special class placement, and achievement substantially below assessed ability. Clearly the PIC-2 most accurately describes a 12-year-old male who demonstrates multiple handicaps. The elevation of the PIY DEF scale ($T = 64$) is the most likely explanation for a PIY profile that is essentially within normal limits. Indeed, review of the medical record details this patient's repeated denial and minimizing of his problems during this hospitalization in an attempt to facilitate his early discharge from treatment. This child's psychiatric history was secondary to a traumatic motor vehicle accident. He felt lonely and scared, he frequently cried, sobbed, shook, avoided others, and was preoccupied with excessive worries (DIS, WDL). He externalized his problems (DLQ) and had difficulties with peers and had little insight into his role in these conflicts (SSK2). This pattern of youth defensiveness is fairly common in inpatient settings.

Case C is quite unusual in that the mother-completed PIC-2 is essentially within normal limits, with the exception of FAM1, FAM2, and DIS3 obtaining some elevation. This normal-limits profile is quite unusual in light of the fact that the mother referred her 12-year-old daughter for this hospitalization. C presented with suicidal ideation, low self-esteem, depression, crying spells, poor appetite, and associated weight loss (DIS, WDL2). She actively demonstrated somatic concern and somatic symptoms in response to conflict during this hospitalization (SOM) and said that she would not talk about her problems with her mother, because she was afraid that such discussion would distress her mother who was under psychiatric care (PIC-2 FAM2). Clinicians were sufficiently concerned with C's internalizing problems to assign discharge diagnoses of Generalized Anxiety Disorder and Depressive Disorder NOS and to prescribe antidepressant medication. Why was C's mother more defensive in describing her daughter's problems than C was herself? It was clearly documented in the medical record that C's mother was concerned that, because of her and C's psychiatric problems, she would be seen as an inadequate mother and consequently would lose custody of her child to another adult family member.

4.13.2.5.2 Student Behavior Survey

The Student Behavior Survey (SBS) differs from the other teacher rating scales described in this chapter in that no particular effort was made to develop the SBS into a parallel teacher version of the PIC-2 and PIY. Indeed, the development of SBS items consisted of several iterations in which the authors focused on content appropriate to teacher observation. SBS items and their rating options appear on a self-scoring form. These items are sorted into content meaningful dimensions under 11 scale headings rather than presented in random order. The resulting items demonstrate a school focus in that 58 of 102 items specifically refer to in-class or in-school behaviors and judgments that can only be made by school staff (Wingenfeld, Lachar, Gruber, & Kline, 1998). The SBS consists of 102 items profiled on to 14 scales that assess student academic status and work habits, social skills, parental participation in the educational process, and problems such as aggressive or atypical behavior and emotional stress. Norms that generate linear *T* scores are gender-specific and divided into two age groups: 5–11 and 12–18 years.

The SBS consists of three sections (Lachar et al., 1999). In the first section the teacher circles one rating out of five (Deficient, Below Average, Average, Above Average, Superior) for eight areas of achievement such as reading comprehension and mathematics that are summed for an estimate of current academic performance. The remaining SBS items are rated on a four-point frequency scale: Never, Seldom, Sometimes, and Usually. The next two scales consist of positively worded descriptions of student adaptive behaviors. Academic Habits (13 items; "Completes class assignments"; "Follows the teacher's directions") and Social Skills (eight items: "Helps other students"; "Participates in class activities") scales document positive behaviors. Ratings of parents are very school-specific in that the teacher is asked to judge the degree to which parents support the student's educational program (six items: "Parent(s) encourage achievement"; "Parent(s) meet with school staff when asked").

The seven Problems in Adjustment scales consist of negatively worded items organized into Health Concerns (six items: "Complains of headaches"; "Talks about being sick"); Emotional Distress (15 items: "Appears sad or unhappy"; "Worries about little things"); Unusual Behavior (seven items: "Seems lost or disoriented"; "Says strange or bizarre things"); Social Problems (12 items: "Criticized by other students"; "Avoids social interaction in class"); Verbal Aggression (seven items: "Insults other students"; "Teases or taunts other students"); Physical Aggression (five items: "Destroys property when angry"; "Starts fights with other students"); and Behavior Problems (15 items: "Disrupts class by misbehaving"; "Talks excessively"). In initial samples of 1173 regular education students and 601 students referred for evaluation or receiving special education services in grades K through 12, 99 of 102 items significantly ($p < 0.001$) separated these two samples. All items demonstrated that they had been placed on the scale with which each obtained the largest correlation. Scale scores of regular education and referred students obtained meaningful three-factor solutions.

Additional effort (Pisecco et al., 1998) resulted in three additional, 16-item nonoverlapping scales consisting of SBS items first nominated as fitting *DSM-IV* diagnoses of ADHD (combined type), Oppositional Defiant Disorder, and Conduct Disorder. Item-to-scale correlations and a three-factor solution of these 48 SBS items empirically supported the placement of these scale items.

4.13.2.5.3 Commentary

The revision of the PIC, the addition of a teacher rating scale, and the collection of a national representative normative sample for each have gone a long way to respond to the concerns raised by reviewers regarding the PIC's age (Kamphaus & Frick, 1996; Knoff, 1989; Merrell, 1994). The evaluations of the SBS and PIC-2 manuals and the review of their subsequent demonstrated ability to evaluate emotional adjustment at baseline and to quantify response to intervention will appear well into the twenty-first century.

The emphasis on evaluating response accuracy using validity scales and the empirical determination of interpretive guidelines continues to characterize these measures. Many psychologists unconvinced of the importance of these concepts may not value their contributions to assessment. Although the PIC has been reduced from 420 to 275 items, into which a set of subscales and a brief form have been incorporated, some clinicians may still judge the length of these questionnaires to be problematic. Although this chapter's author is obviously biased against such a position, it is certain that the breadth and depth of a measure's content establishes the potential boundaries for its utility. Even the 270 items of the PIY are easily completed in less than 45 minutes by children as young as students in the fourth grade. Efficiency has been improved by rejecting any item not actively used in the interpretive process as well as providing computer software for scoring and interpretation.

As either "new" or "improved" measures, the approximately 400 PIC-relevant publications (bibliography available from this author) at best now suggest diagnostic potential for these forms on the dimensions that retain the greatest similarity from original to revised formats. A great deal of effort will be necessary to establish the diagnostic utility of the new and revised forms to achieve the standards of the original inventory (Lachar & Kline, 1994). Such efforts have begun. For example, matched samples of adolescents with discharge diagnoses of either Conduct Disorder or Major Depression were correctly classified by PIY subscales in 83% of cases (Lachar, Harper, Green, Morgan, & Wheeler, 1996).

4.13.3 CONCLUSIONS

Clinician interest in applying objective questionnaires in the evaluation of children and adolescents continues to grow. These measures are efficiently obtained, easily scored, and at least appear to be relatively easy to interpret (although a sophisticated knowledge of both developmental psychopathology and psychometrics is clearly a prerequisite). The routine application of such measures in clinics, hospitals, and private practices will continue to be a growing trend. This chapter documents that a number of quality measures are currently available to meet this need. Clinicians should review each set of measures to select the specific family of instruments best suited for them to be applied as a routine procedure, or the specific instruments to be applied to a specific clinical activity.

Routine collection of information that is both multidimensional and multisource will influence future research and clinical practice. Efforts to document the effect of comorbid conditions and to establish interpretive guidelines for individual profiles and sets of profiles will continue. Interpretive handbooks, similar to those available for popular adult personality inventories, will be developed to support training and clinical applications. Projects will identify research designs in which the meaning of scale values and profile patterns can be established through the analysis of independent criteria. Given the infrequency of studies designed to evaluate tests, as compared to studies that use tests to evaluate clinical phenomena, efforts will be made to identify and collect substantial data that will allow such analyses. Indeed, the concurrent application of parent, teacher, youth, and clinician forms, as demonstrated in Table 4, will allow the establishment of external correlates and the determination of scale standard score interpretive ranges.

Studies will routinely investigate the relative efficacy of informant sources (parent, teacher, and youth) to support a given diagnostic decision, as well as compare the instruments across these four families. Some early efforts include Pearson and Lachar's (1994) comparison of CBCL and PIC scales in their ability to predict the standard scores of the Vineland Adaptive Behavior Scale, and the comparison of BASC and CBCL forms in identifying students with ADHD and in differentiating between ADHD types (Ostrander, Weinfurt, Yarnold, & August, in press; Vaughn et al., 1997). The former study established the relative value of the BASC PRS Attention Problem Scale for identifying students likely to demonstrate nonspecific ADHD, although both measures demonstrated limited ability to identify ADHD subtype. In the second study, both BASC parent and teacher scales demonstrated significant superiority over the CBCL and TRF in accurately classifying ADHD type.

The routine use of these measures will expand to repeated assessment paradigms. The interest in measuring treatment effectiveness as a clinical as well as a research priority will make a meaningful contribution to identifying optimal treatment strategies as well as assigning specific treatment options for the individual case. These trends will facilitate retrospective "file drawer" study within clinical settings and expand the investigation of clinically referred youth.

4.13.4 REFERENCES

Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: A factor-analytic study. *Psychological Monographs, 80* (Whole Issue No. 615).

Achenbach, T. M. (1981). A junior MMPI? (Review of multidimensional description of child personality: A manual for the Personality Inventory for Children and Actuarial assessment of child and adolescent personality: An interpretive guide for the Personality Inventory for Children profile.) *Journal of Personality Assessment, 45,* 332–333.

Achenbach, T. M. (1991a). *Integrative guide for the 1991 CBCL/4-18, YSR, and TRF profiles.* Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1991b). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile.* Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1991c). *Manual for the Teacher's Report Form and 1991 Profile.* Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1991d). *Manual for the Youth Self-Report and 1991 Profile.* Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1992). New developments in multiaxial empirically based assessment of child and adolescent psychopathology. In J. Rosen & P. McReynolds (Eds.), *Advances in psychological assessment* (Vol. 8, pp. 75–102). New York: Plenum.

Achenbach, T. M. (1993). *Empirically based taxonomy: How to use syndromes and profile types derived from the CBCL/4–18, TRF, and YSR.* Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101,* 213–232.

August, G. J., MacDonald, A. W., Realmuto, G. M., & Skare, S. S. (1996). Hyperactive and aggressive pathways: Effects of demographic, family, and child characteristics on children's adaptive functioning. *Journal of Clinical Child Psychology, 25,* 341–351.

August, G. J., Realmuto, G. M., MacDonald III, A. W., Nugent, S. M., & Crosby, R. (1996). Prevalence of ADHD and comorbid disorders among elementary school children screened for disruptive behavior. *Journal of Abnormal Child Psychology, 24,* 571–595.

Biederman, J., Faraome, S. V., Doyle, A., Lehman, B. K., Kraus, I., Perrin, J., & Tsuang, M. T. (1993). Convergence of the Child Behavior Checklist with structured interview-based psychiatric diagnoses of ADHD children with and without comorbidity. *Journal of Child Psychology and Psychiatry, 34,* 1241–1251.

Bloomquist, M. L., August, G. J., Cohen, C., Doyle, A., & Everhart, K. (1997). Social problem solving in hyperactive-aggressive children: How and what they think in conditions of automatic and controlled processing. *Journal of Clinical Child Psychology, 26,* 172–180.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Chen, W. J., Faraone, S. V., Biederman, J., & Tsuang, M. T. (1994). Diagnostic accuracy of the Child Behavior Checklist scales for Attention-Deficit Hyperactivity Disorder: A receiver-operating characteristric analysis. *Journal of Consulting and Clinical Psychology, 62,* 1017–1025.

Conners, C. K. (1997). *Conners' Rating Scales-Revised technical manual.* North Tonawanda, NY: Multi-Health Systems.

Conners, C. K. (in press). Conners' Rating Scales-Revised. In M. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (2nd ed.). Hillsdale, NJ: Erlbaum.

Conners, C. K., Wells, K. C., Parker, J. D. A., Sitarenios, G., Diamond, J. M., & Powell, J. W. (1997). A new self-report scale for assessment of adolescent psychopathology: Factor structure, reliability, validity, and diagnostic sensitivity. *Journal of Abnormal Child Psychology, 25,* 487–497.

Dalton, J. E. (1996). Juvenile male sex offenders: Mean scores on the BASC Self-Report of Personality. *Psychological Reports, 79,* 634.

Doyle, A., Ostrander, R., Skare, S., Crosby, & August, G. J. (1997). Convergent and criterion-related validity of the Behavior Assessment System for Children–Parent Rating Scale. *Journal of Clinical Child Psychology, 26,* 276–284.

Drotar, D., Stein, R. E., & Perrin, E. C. (1995). Methodological issues in using the Child Behavior Checklist and its related instruments in clinical child psychology research. *Journal of Clinical Child Psychology, 24,* 184–192.

Edelbrock, C., & Costello, A. J. (1988). Convergence between statistically derived behavior problem syndromes and child psychiatric diagnoses. *Journal of Abnormal Child Psychology, 16,* 219–231.

Erford, B. T. (1996). Analysis of the Conners' Teacher Rating Scale-28 (CTRS-28). *Assessment, 3,* 27–36.

Faraone, S. V., Biederman, J., Weber, W., & Russell, R. L. (1998). Psychiatric, neuropsychological, and psychosocial features of DSM-IV subtypes of Attention-Deficit/Hyperactivity Disorder: Results from a clinically referred sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 37,* 185–103.

Flanagan, D. P., Alfonso, V. C., Primavera, L. H., & Povall, L. (1996). Convergent validity of the BASC and SSRS: Implications for social skills assessment. *Psychology in the Schools, 33,* 13–23.

Forbes, G. B. (1985). The Personality Inventory for Children (PIC) and hyperactivity: Clinical utility and problems of generalizability. *Journal of Pediatric Psychology, 10,* 141–149.

Gdowski, C. L., Lachar, D., & Kline, R. B. (1985). A PIC profile typology of children and adolescents: I. An empirically-derived alternative to traditional diagnosis. *Journal of Abnormal Psychology, 94,* 346–361.

Greenbaum, P. E., Dedrick, R. F., Prange, M. E., & Friedman, R. M. (1994). Parent, teacher, and child ratings of problem behaviors of youngsters with serious emotional disturbances. *Psychological Assessment, 6,* 141–148.

Jensen, P. S., Watanabe, H. K., Richters, J. E., Roper, M., Hibbs, E. D., Salzberg, A. D., & Liu, S. (1996). Scales, diagnoses, and child psychopathology: II. Comparing the CBCL and the DISC against external validators. *Journal of Abnormal Child Psychology, 29,* 151–168.

Kamphaus, R. W., & Frick, P. J. (1996). *Clinical assessment of child and adolescent personality and behavior.* Boston, MA: Allyn & Bacon.

Kamphaus, R. W., Huberty, C. J., DiStifano, C., & Petoskey, M. D. (1997). A typology of teacher-rated child behavior for a national U.S. sample. *Journal of Abnormal Child Psychology, 25,* 453–463.

Kamphaus, R. W., & Reynolds, C. R. (1998). *BASC Monitor for ADHD manual.* Circle Pines, MN: American Guidance Service.

Kline, R. B. (1994). New objective rating scales for child assessment, I. Parent- and teacher-informant inventories of the Behavior Assessment System for Children, the Child Behavior Checklist, and the Teacher Report Form. *Journal of Psychoeducational Assessment, 12,* 289–306.

Kline, R. B. (1995). New objective rating scales for child assessment, II. Self-report scales for children and adolescents: Self-Report of Personality of the Behavior Assessment System for Children, the Youth Self-Report, and the Personality Inventory for Youth. *Journal of Psychoeducational Assessment, 13,* 169–193.

Kline, R. B., & Lachar, D. (1992). Evaluation of age, sex, and race bias in the Personality Inventory for Children (PIC). *Psychological Assessment, 4,* 333–339.

Kline, R. B., Lachar, D., & Gdowski, C. L. (1987). A PIC typology of children and adolescents: II. Classification rules and specific behavior correlates. *Journal of Clinical Child Psychology, 16,* 225–234.

Kline, R. B., Lachar, D., Gruber, C. P., & Boersma, D. C. (1994). Identification of special education needs with the Personality Inventory for Children (PIC): A profile-matching strategy. *Assessment, 1,* 301–313.

Kline, R. B., Lachar, D., & Sprague, D. J. (1985). The Personality Inventory for Children (PIC): An unbiased predictor of cognitive and academic status. *Journal of Pediatric Psychology, 10,* 461–477.

Knoff, H. M. (1989). Review of the Personality Inventory for Children, Revised Format. In J. C. Connolly & J. C. Kramer (Eds.), *The tenth mental measurements yearbook* (pp. 624–630). Lincoln, NE: Buros Institute of Mental Measurements.

Kovacs, M. (1992). *Children's Depression Inventory (CDI) manual.* Toronto: Multi-Health Systems.

Lachar, D. (in press). Personality Inventory for Children-2nd Edition (PIC-2), Personality Inventory for Youth (PIY), and the Student Behavior Survey (SBS). In M. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (2nd ed.). Hillsdale, NJ: Erlbaum.

Lachar, D. (1982). *Personality Inventory for Children (PIC) revised format manual supplement.* Los Angeles: Western Psychological Services.

Lachar, D. (1993). Symptom checklists and personality inventories. In T. R. Kratochwill & R. J. Morris (Eds.), *Handbook of psychotherapy with children* (pp. 38–57). New York: Allyn & Bacon.

Lachar, D. (1999). *Personality Inventory for Children-2nd Edition (PIC-2) manual.* Los Angeles: Western Psychological Services.

Lachar, D., & Gdowski, C. L. (1979). *Actuarial assesment of child and adolescent personality: An interpretive guide for the Personality Inventory for Children profile.* Los Angeles: Western Psychological Services.

Lachar, D., Gdowski, C. L., & Snyder, D. K. (1982). Broad-band dimensions of psychopathology: Factor scales for the Personality Inventory for Children. *Journal of Consulting and Clinical Psychology, 50,* 634–642.

Lachar, D., & Gruber, C. P. (1993). Development of the Personality Inventory for Youth: A self-report companion to the Personality Inventory for Children. *Journal of Personality Assessment, 61,* 81–98.

Lachar, D., & Gruber, C. P. (1995). *Personality Inventory for Youth (PIY) manual technical guide.* Los Angeles: Western Psychological Services.

Lachar, D., Harper, R. A., Green, B. A., Morgan, S. T., & Wheeler, A. C. (1996, August). *The Personality Inventory for Youth: Contribution to diagnosis.* Paper presented at the 104th Annual Convention, American Psychological Association, Toronto, Canada.

Lachar, D., & Kline, R. B. (1994). Personality Inventory for Children and Personality Inventory for Youth. In M. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment.* Hillsdale, NJ: Erlbaum.

Lachar, D., & Kline, R. B., & Gdowski, C. L. (1987). Respondent psychopathology and interpretive accuracy of the Personality Inventory for Children: The evaluation of a "most reasonable" assumption. *Journal of Personality Assessment, 51,* 165–177.

Lachar, D., Kline, R. B., Wingenfeld, S. A., & Gruber, C. P. (1999). *Student Behavior Survey (SBS) manual.* Los Angeles: Western Psychological Services.

LaCombe, J. A., Kline, R. B., Lachar, D., Butkus, M., & Hillman, S. B. (1991). Case history correlates of a Personality Inventory for Children (PIC) profile typology. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 678–687.

Lett, N. J., & Kamphaus, R. W. (1997). Differential validity of the BASC Student Observation System and the BASC Teacher Rating Scale. *Canadian Journal of School Psychology, 13,* 1–14.

Loeber, R., Green, S. M., & Lahey, B. B. (1990). Mental health professionals' perception of the utility of children, mothers, and teachers as informants on childhood psychopathology. *Journal of Clinical Child Psychology, 19,* 136–143.

Lowman, M. G., Schwanz, K. A., & Kamphaus, R. W. (1996). WISC-III third factor: Critical measurement issues. *Canadian Journal of School Psychology, 12,* 15–22.

Merrell, K. W. (1994). *Assessment of behavioral, social, & emotional problems. Direct & objective methods for use with children and adolescents.* New York: Longman.

Michael, K. D., & Merrell, K. W. (1998). Reliability of children's self-reported internalizing symptoms over short to medium-length time intervals. *Journal of the American Academy of Child and Adolescent Psychiatry, 37,* 194–201.

O'Connor, T. G., McGuire, S., Reiss, D., Hetherington, E. M., & Plomin, R. (1998). Co-occurrence of depressive symptoms and antisocial behavior in adolescence: A common genetic liability. *Journal of Abnormal Psychology, 107,* 27–37.

Ostrander, R., Weinfurt, K. P., Yarnold, P. R., & August, G. J. (in press). Diagnosing attention deficit disorders using the BASC and the CBCL: Test and construct validity analyses using optimal discriminant classification trees. *Journal of Consulting and Clinical Psychology.*

Pearson, D. A., & Lachar, D. (1994). Using behavioral questionnaires to identify adaptive deficits in elementary school children. *Journal of School Psychology, 32,* 33–52.

Phares, V. (1997). Accuracy of informants: Do parents think that mother knows best? *Journal of Abnormal Child Psychology, 25,* 165–171.

Pisecco, S., Lachar, D. Gallen, R. T., Gruber, C. P., & Huzinec, C. (1998). *Development of disruptive behavior DSM-IV scales from teacher ratings.* Manuscript submitted for publication.

Realmuto, G. M., August, G. J., Sieler, J. D., & Pessoa-Brandao (1997). Peer assessment of social reputation in community samples of disruptive and nondisruptive children: Utility of the revised class play method. *Journal of Clinical Child Psychology, 26,* 67–76.

Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior Assessment System for Children manual.* Circle Pines, MN: American Guidance Service.

Reynolds, C. R., & Richmond, B. O. (1985). *Revised Children's Manifest Anxiety Scale manual.* Los Angeles: Western Psychological Services.

Richters, J. E. (1992). Depressed mothers as informants about their children: A critical review of the evidence for distortion. *Psychological Bulletin, 112,* 485–499.

Rowe, K. S., & Rowe, K. J. (1997). Norms for parental ratings on Conners' Abbreviated Parent–Teacher Questionnaire: Implications for the design of behavioral rating inventories and analyses of data derived from them. *Journal of Abnormal Child Psychology, 25,* 425–451.

Steingard, R., Biederman, J., Doyle, A., & Sprich-Buckminster, S. (1992). Psychiatric comorbidity in attention deficit disorder: Impact on the interpretation of Child Behavior Checklist results. *Journal of the Academy of Child and Adolescent Psychiatry, 31,* 449–454.

Vaughn, M., Black, K., Hall, J., Hynd, G., & Riccio, C. A. (1995). *Use of the BASC and Achenbach for assessment and intervention planning*. Paper presented at the 1995 Convention of the National Association of School Psychologists, Chicago.

Vaughn, M. L., Riccio, C. A., Hynd, G. W., & Hall, J. (1997). Diagnosing ADHD (predominantly inattentive and combined subtypes): Discriminant validity of the Behavior Assessment System for Children and the Achenbach parent and teacher rating scales. *Journal of Clinical Child Psychology, 26,* 349–357.

Vignoe, D., & Achenbach, T. M. (1997). *Bibliography of published studies using the Child Behavior Checklist and related materials: 1997 edition*. Burlington, VT: University of Vermont, Department of Psychiatry.

Wainwright, A., & MHS Staff (1996). *Conners' Rating Scales: Over 25 years of research—An annotated bibliography*. Toronto: Multi-Health Systems.

Weinstein, S. R., Noam, G. G., Grimes, K., Stone, K., & Schwab-Stone, M. (1990). Convergence of DSM-III diagnoses and self-reported symptoms in child and adolescent inpatients. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 627–634.

Wingenfeld, S. A., Lachar, D., Gruber, C.P., & Kline, R. B. (1998). Development of the teacher-informant Student Behavior Survey. Manuscript submitted to *Journal of Psychoeducational Assessment*.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.14
Objective Personality Assessment with Adults

JAMES N. BUTCHER and JEANETTE TAYLOR
University of Minnesota, Minneapolis, MN, USA
and
G. CYNTHIA FEKKEN
Queen's University, Kingston, ON, Canada

4.14.1 INTRODUCTION 404
  4.14.1.1 Intelligence and Insight 404
  4.14.1.2 Veridicality of Self-report 405
  4.14.1.3 Personality Stability 405
    4.14.1.3.1 Definitions 405
    4.14.1.3.2 Influences on personality stability 406
4.14.2 ASSESSING ADULTS IN CLINICAL SETTINGS 407
  4.14.2.1 The Minnesota Multiphasic Personality Inventory-Revised 407
    4.14.2.1.1 Origin of the MMPI/MMPI-2 408
    4.14.2.1.2 Measurement dimensions 408
    4.14.2.1.3 Recent validity research for the MMPI-2 409
  4.14.2.2 Basic Personality Inventory 413
  4.14.2.3 Personality Assessment Inventory 415
4.14.3 SPECIALIZED OR FOCUSED CLINICAL ASSESSMENT MEASURES 416
  4.14.3.1 Millon Clinical Multiaxial Inventory 416
  4.14.3.2 The Beck Depression Inventory 418
  4.14.3.3 The State-Trait Anxiety Inventory 418
  4.14.3.4 Whitaker Index of Schizophrenic Thinking 419
4.14.4 NORMAL RANGE PERSONALITY ASSESSMENT 420
  4.14.4.1 Objective Personality Measures in Research 420
    4.14.4.1.1 The Five Factor Model (the Big Five) 420
    4.14.4.1.2 NEO Personality Inventory 420
    4.14.4.1.3 Multidimensional Personality Questionnaire 420
    4.14.4.1.4 Personality Research Form 421
    4.14.4.1.5 Sixteen Personality Factor Test 421
    4.14.4.1.6 California Psychological Inventory 421
  4.14.4.2 Objective Personality Measures in Educational/Vocational Assessment 422
    4.14.4.2.1 FFM 422
    4.14.4.2.2 NEO-PI 422
    4.14.4.2.3 CPI 422
    4.14.4.2.4 MMPI/MMPI-2 422
  4.14.4.3 Personnel Screening 423
  4.14.4.4 Other Personality Measures in Personnel Selection 423
  4.14.4.5 The 16PF 423
  4.14.4.6 FFM 423

4.14.4.7 NEO-PI 424
4.14.4.7.1 CPI 424
4.14.5 SUMMARY 424
4.14.6 REFERENCES 425

4.14.1 INTRODUCTION

Objective personality tests form a standard part of most applied psychologists' toolboxes when it comes to measuring personality. In contrast to projective techniques and to subjective approaches to personality assessment, objective personality questionnaires are made up of relatively unambiguous stimuli or items; offer the respondent relatively restricted response options; and present a scoring scheme that involves few, if any, scoring judgments, resulting in high scoring reliability (Wiggins, 1973). Most objective personality assessment instruments used by clinicians and researchers employ a self-report format. In contemporary psychology, the Minnesota Multiphasic Personality Inventory (MMPI) and the MMPI-2 are the most widely used personality assessment instruments in both clinical and research settings (Butcher and Rouse, 1996). The next most popular instruments used to assess personality, the Rorschach and the Thematic Apperception Test, are projective (not objective), but each still requires the clinician to obtain self-reported behavior. In essence, the clinician is dependent on the client's ability and willingness to make accurate self-reports when assessing personality with standardized instruments. It is, therefore, important to determine whether adults being assessed can competently reveal information about their personalities through self-report instruments. In order to address this question, the client's intelligence and insight in sharing self-information and the success of personality tests for eliciting self-information that can be externally verified need to be considered. We will examine each of these considerations below.

4.14.1.1 Intelligence and Insight

The existing literature on intelligence and personality is largely focused on the predictive relationship between the two constructs. However, clinicians are generally not concerned about the ability of a personality profile to predict an intelligence quotient or vice versa. Instead, most are concerned with the more basic question of how intellectual functioning impacts the ability of a client to self-disclose on an objective personality test. No literature exists which directly addresses the issue of the relationship between a client's level of intelligence and his or her ability to competently reveal information about personality through a self-report instrument.

Objective personality measures leave the issue of the client's intellectual functioning in the hands of the clinician. None of the widely used personality assessment instruments require a pre-screening for normal or above normal intelligence. However, most self-report personality instruments require that the examinee have a minimum reading level competency (often expressed as a minimum grade level). Thus, a minimum level of school achievement is implicit in the validity of a client's score, and it follows that the clinician should use care when administering a self-report instrument to a client with known or suspected intellectual impairment. It is also essential for the clinician to be aware of the educational range of the standardization sample for any self-report measure, and to use caution when interpreting protocols from clients who fall outside of that range.

The clinician needs to be aware of a client's intellectual functioning as it may influence the client's ability to provide information about his or her personality on a self-report measure. If a client appears to be low in intellectual functioning, then the results of self-report measures should be interpreted with caution and perhaps supplemented with reports from other sources.

While the capacity for insight cannot readily be extrapolated from an intelligence quotient in a linear fashion, insight into one's problems might be discernible from the personality profile itself. As Butcher and Rouse (1996) note in their recent review of personality research, several of the widely used objective personality measures have attitudinal measures that aid the clinician in evaluating the client's disclosure capability. Crookes and Buckley (1976) found a relationship between the Eysenck Personality Inventory Lie (L) scale score and diagnosis among both psychiatric inpatients and outpatients. They found that high L scale scores were positively related to diagnosis of disorders associated with low insight. They concluded that the L score (a validity indicator found on several objective personality instruments) is highly related to a person's awareness of his or her behavior. Moreover, this finding is in agreement with the interpretive guidelines for the MMPI-2 L scale, which suggests the general utility of this measure at appraising a client's self-awareness.

The quality of insight that a client brings to the assessment situation may greatly influence the validity of the assessment. Some instruments are designed to assess clients through behavioral items that require low insight or self-observation. These items are specifically selected because of their ability to predict relevant criteria and may appear at face value totally unrelated to the particular aspects of personality functioning intended to be predicted. This highlights one advantage of using objective personality measures in clinical settings: clients who may yield poor information during a clinical interview that requires self-evaluation and/or introspection may be fully capable of providing valuable clinical information when a personality instrument comprising empirically selected items is employed.

4.14.1.2 Veridicality of Self-report

Aside from the issues of intellectual functioning and insight the utility of self-report measures can be assessed by examining their success at corroborating independent facts about a client. For example, in an early study of the MMPI, Payne and Wiggins (1972) examined the relationship between content-based profiles and external descriptors of a large group of psychiatric inpatients. Interpretation of content-based profiles is based on combinations of "obvious" test items which correspond to traditional self-report instruments (e.g., "I am a high-strung person"). That is, the client endorses the item as a means of directly relaying information about him- or herself. The authors found that the patients' self-reported MMPI profiles matched quite well with interview report and external observation.

In a similar vein, Koss and Butcher (1973) found that psychiatric inpatients identified (through their observed behavior and presenting symptoms) as belonging to one of six operationally defined crisis situations could be discriminated from one another on the basis of their endorsed MMPI content. The authors interpreted their results as providing evidence for the competence and willingness of adults in clinical settings to reflect information accurately about their personality and psychological functioning through self-report test items. Similar evidence of a positive relationship between self-reported personality measures and external criteria has been found for certain scales from the California Psychological Inventory (CPI; e.g., Hindelang, 1972), the Millon Clinical Multiaxial Inventory (MCMI; e.g., Jaffe and Archer, 1987), and the Basic Personality Inventory (BPI; Holden, Fekken, Reddon, Helmes, & Jackson, 1988), to name three.

Although the research literature shows that people can competently disclose information about their personalities through standardized, objective, self-report measures, some individuals, however, may not be motivated to cooperate in their psychological assessment. That is, some respondents, because of a need or motivation to present themselves in a particular way, do not respond in a truthful, open manner. The most common examples of individuals being motivated to appear different from the way they actually are in today's assessment settings are applicants for employment, parents involved in custody disputes, and other individuals being evaluated as part of a court case. A self-report assessment instrument must contain an effective means of identifying respondents who are dissimulating. Perhaps the most common way of detecting whether people are not motivated to report accurately about themselves is via validity scales or control scales. In some assessment situations the validity scales provide the most important and useful information about the client. A self-report personality measure without effective validity scales is too limited to operate across a broad range of assessment situations. Furthermore, tests that possess validity indicators, such as the L scale on the MMPI/MMPI-2, offer the clinician information regarding both the client's willingness to cooperate with the assessment and his or her level of insight. These points argue in favor of the continued use of standardized, objective self-report measures in the assessment of personality in clinical settings.

4.14.1.3 Personality Stability

One of the assumptions of objective personality assessment is that the personality characteristic being measured is stable. Various major studies have come to the conclusion that personality changes little in adulthood, particularly after age 30 (Conley, 1985; Finn, 1986; McCrae & Costa, 1990; Schuerger, Zarrella, & Hotz, 1989). Typical levels of stability on personality tests are in the 0.5–0.7 range. This generalization bears some comment.

4.14.1.3.1 Definitions

What exactly is meant by stable responding? On an objective personality questionnaire, a set of responses that are consistent over time may be defined in at least three ways: (i) as an identical scale score, (ii) as a set of identical item responses, or (iii) as a scale score that signifies the same relative standing on the personality characteristic.

Conclusions about the stability of personality are most commonly based on scale scores. That is, the stability of personality estimates result from comparing scale scores obtained at two points in time. Presumably the same scale score could result even if the test respondent endorsed somewhat different subsets of items. The consistency of responses to individual test items has been studied as a meaningful individual differences variable and as an index for establishing the interpretability of individual response protocols. Item response consistency has not been examined as a definition of longitudinal personality stability *per se*.

Many researchers have focused on changes in average scale scores. Their goal is to understand normative changes in personality. On the other hand, some researchers examine personality stability by correlating scale scores. Thus, when coefficients of personality stability are reported to be in the 0.5–0.7 range, this should be understood to be a statement about relative, and not absolute, scale score stability.

4.14.1.3.2 Influences on personality stability

Although the typical stability of personality scale scores is quite high, there are nonetheless particular variables that affect score stability in predictable ways. The first consideration is the influence of instrument characteristics. Specifically, the average number of items per scale and the homogeneity of the scales are the important predictors of scale score stability (Finn, 1986; Schuerger et al., 1989). Variance associated with specific instruments (e.g., the MMPI vs. the Sixteen Personality Factor Test [16PF]) apparently does not contribute to greater personality stability beyond the effects of scale length and homogeneity (Schuerger et al., 1989). Such findings underscore the need for test users to select reliable measures, particularly if test data are intended to evaluate long-term personality change.

Second, estimates of personality stability are affected by the length of the retest interval. Not surprisingly, personality appears to more stable over short intervals than long ones. In particular, score stability drops over the first year before it levels off (Schuerger et al., 1989; Windle, 1954) and then stays stable at the same level for very long periods of time (McCrae & Costa, 1990).

One simple reason for relatively high short-term personality stability may be that people remember and repeat their responses to the items on a questionnaire. Lumsden (1977), however, proposed a substantive rather than a methodological explanation. He argued that particular personality characteristics "swell" into prominence over the short term and then fade again. For a student entering college, for example, "independence" may become a salient dimension and hence, self-reports on this dimension are likely to have exaggerated consistency over the (short) time period when "independence" is prominent.

A third influence on personality stability has to do with the operationalization of personality. Most objective personality questionnaires are intended to measure traits that are by definition stable and enduring. The construction of such questionnaires favors selection of items that will yield stable scores. In addition, the overall instructions as well as the wording of specific items prompt people to make broad generalizations about themselves and to downplay variations. Thus, personality may appear to be stable because the constructs of personality are explicitly conceptualized and measured as stable. There are certainly empirical studies that challenge the idea of personality stability by providing evidence of developmental change over long time periods based on personality constructs and measures that expressly incorporate the notion of change (e.g., Whitbourne, Zuschlag, Elliot, & Waterman, 1992).

Fourth, some personality constructs are associated with more stability than others. Schuerger et al.'s (1989) data show that measures of psychopathology have less scale score stability than measures of "normal" personality characteristics. Perhaps "normal" personality constructs are more crystallized than psychopathological constructs and thus less error is built into the construct itself. Similarly, psychopathological constructs may be more confounded with the effects of response style variance such as social desirability and acquiescence than normal personality constructs, again compounding the noisiness of the construct.

Many researchers have tried to highlight differences in stability among the specific constructs that they chose to study. For example, Finn (1986) demonstrated that mood-related constructs (e.g., depression) had comparatively low levels of consistency, whereas Helson and Moane (1987) found that constructs related to socialization had high stability. There are shortcomings to such an approach. The identification of those dimensions on which people are changeable (i.e., states) as distinct from the more permanent individual differences dimensions needs to be tackled systematically if a comprehensive list is to be obtained (Cattell, 1963).

The interpretation of differences in stability between constructs is further complicated by age and cohort effects. Studies such as that of Helson and Moane (1987) estimate personality stability using information obtained from a single group of subjects who have been assessed at different points in time. There are limits to the generalizability of this information to other groups of people. Other studies such as that of Finn (1986) use basically a cross-sectional approach. They assess two or more groups of subjects having different ages at a single time. Here it is difficult to attribute differences in the stability of particular personality constructs to age as opposed to generational factors. More complicated designs may be needed in order to understand such age and cohort effects (Conley, 1985). At the very least, any generalizations about which personality constructs are more or less stable need to be carefully qualified.

A fifth influence on personality stability has to do with features of the persons being evaluated. Test respondents exhibit remarkable individual differences in personality stability (Assendorp, 1992; McCrae & Costa, 1990). Do some subgroups of people show more personality stability? Across studies, patient and prisoner samples show less personality test score stability than "normals" (Schuerger et al., 1989). Perhaps "non-normals" have systematically higher scale scores than "normals" and, hence, their retest scores show regression toward the mean (Windle, 1954). Alternatively, "non-normal" groups also may be more likely than normal groups to seek actively to change personality.

The other two person variables that have received attention in the literature on personality stability are age and sex differences. Some researchers report that older adults are more stable than younger adults on many traits (Finn, 1986; Schuerge et al., 1989) although others argue that stability differences after about age 30 are minimal (McCrae & Costa, 1990). Finn (1986) has asked whether older adults might just be more rigid in their self-perceptions. However, the picture of personality stability that results from self-reports does tend to be substantiated by ratings on personality information collected from others, such as spouses (McCrae & Costa, 1990). Thus, personality may indeed be more stable in older adults.

Many studies in the area of personality stability sample exclusively men or women. Yet there seems to be little evidence of an overall difference in the level of personality stability for men and women (Schuerger et al., 1989). Nonetheless, researchers interested in normative change often argue that men (Finn, 1986) and women (Helson & Moane, 1987) may be expected to change on different constructs as a function of distinct life paths or societal pressures. Few studies have been explicitly designed to compare sex differences in personality stability.

Overall, personality shows considerable stability in adulthood. The objective measures traditionally used in personality assessment may predispose us to find evidence of stability as a function of the kinds of constructs, instructions, and test construction techniques associated with these measures. Aspects of the test instrument itself, such as length or homogeneity, may moderate stability estimates, as could the length of the interval between personality assessments. Samples of "normals" and of older adults may exhibit more personality stability; sex differences in general personality stability appear to be minimal. The assumption of personality stability, at least over the short run, is central to objective personality assessment if test scores are going to predict relevant behaviors and have an impact on planning how to manage that behavior.

4.14.2 ASSESSING ADULTS IN CLINICAL SETTINGS

In this section we will provide an overview of a number of objective personality inventories that have been designed to assess adults in clinical settings. More detail on the use of the MMPI-2 will be given because it is the most widely researched and used measure. Two other measures that have recently been published to measure essentially the same clinical problem areas as the MMPI-2, BPI (Jackson, 1989) and the Personality Assessment Inventory (PAI) (Morey, 1991) will be briefly described. In the next section, several specialized measures will then be surveyed that have been developed to measure more specific clinical problems or behaviors: the MCMI to assess personality disorders; the Beck Depression Inventory (BDI; Beck, Steer, & Garbin, 1988) to assess anxiety; the State–Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970); and the Whitaker Index of Schizophrenic Thinking (WIST; Whitaker, 1973).

4.14.2.1 The Minnesota Multiphasic Personality Inventory-Revised

The MMPI-2 is the most widely used personality test with adults (Lubin, Larsen, & Matarazzo, 1984; Watkins, 1996). Keilen and Bloom (1986) found that the MMPI/MMPI-2 was the most frequently used test in custody evaluations. Developed originally in the late

1930s by Starke Hathaway and J. C. McKinley and redeveloped by Butcher, Dahlstrom, Graham, Tellegen, and Kaemmer (1989), this instrument provides a comprehensive survey of personality characteristics and clinical problems. In the original MMPI, a strictly empirical scale construction approach was followed to develop scales that would assess a patient's probable "membership" in a clinical group. An extensive amount of research has been published on the effectiveness of these measures at predicting and describing problems in adults. Research on the original version of the MMPI covered a very broad range of people—psychiatric and medical patients, substance abusers, incarcerated felons, and many other clinical groups.

In addition, a very broad range of "normals" have been studied including applicants for various jobs such as airline pilots, US Navy submariner crew members, and police and security personnel. Moreover, the instrument came to be widely employed as a personality research instrument. Butcher and Rouse (1996), in a survey of 20 years of research in clinical assessment, found that the MMPI/MMPI-2 is the most widely researched instrument with nearly twice the number of articles as the seven next leading tests. What has made the MMPI/MMPI-2 the most widely used assessment technique in the personality area? We will examine the make-up, utility, and limitations of the instrument in assessing adults.

4.14.2.1.1 Origin of the MMPI/MMPI-2

A growing dissatisfaction with subjective methods of evaluating patients in clinical situations (such as interviews and projective tests) led the original MMPI developers, Hathaway (a psychologist) and McKinley (a psychiatrist), to experiment with an objective method of clinical problem assessment. They accumulated a large number of items (symptoms, beliefs, attitudes, etc.) from tests and case material and administered them to a large group of "normal" people to serve as a comparison group. They then administered the items to well-defined and homogeneous groups of clinical patients. They empirically contrasted the responses of the patient groups with the normals to obtain items that significantly separated the groups. These items were then combined into scales for the various clinical problem areas such as depression, schizophrenia, and so forth. Once derived, the clinical scales were normed on the non-patient sample and *T* scores were developed to enable the test interpreter to determine how extreme a particular person's score was on a given scale. All of the scales were normed so that they had a mean score of 50 and a standard deviation of 10. These standard score distributions then allowed the scores to be plotted on a profile so that the interpreter would have a visual picture of how extreme a particular score was when compared with the normal. Their empirical approach produced highly valid and effective scales that predicted or described the likelihood that a person's score on a scale was in the clinical range or similar to the patient groups.

The original MMPI clinical scales have undergone very substantial study and cross-validation since their publication. They have become a standard means of objective symptom classification since their development. The clinical scales and configurations of scales, referred to as code types, have undergone substantial documentation as an objective classification schema—often referred to as "cookbooks." Researchers, for example, have cataloged the behavioral characteristics associated with their test indices allowing for automatic interpretation. That is, when a particular scale score or cluster of scales are obtained, then a well-validated set of behavioral descriptions—known as descriptors—are applied. Their objective classification approach to interpretation fostered the development of computer-based interpretation methods so popular today (Butcher, 1995; Butcher et al., 1998).

4.14.2.1.2 Measurement dimensions

(i) Validity scales

As noted earlier, it is essential to any self-reported personality assessment to appraise carefully possible invalidating conditions or circumstances. Structural elements of the test administration require evaluation to determine if extratest factors influenced the item responses. For example, were the instructions clearly presented, was the individual able to read and comprehend the items?

The MMPI-2 contains a number of scales and indexes that provide the test interpreter with information on the individual's cooperativeness and honesty in responding to the items (see Table 1 for a summary of the validity or control scales for the MMPI-2). A clear picture is obtained as to whether the person has attempted to present a false picture on the test. These scales are of several types. First, noncontent-oriented measures can provide information on whether the person was inconsistent in responding (the True Response Inconsistency or TRIN and Variable Response Inconsistency or VRIN scales). Additionally, the Cannot Say or "?' scale provides information about the person's cooperation in completing the items. Two other indexes that can provide clues to uncooperative test-taking behavior are the percentage of true and percentage of false endorsement. Records with a nearly all true or all false response pattern suggest uncooperativeness in completing the task.

Three scales have been developed to provide information about the tendency on the part of some people to claim excessive virtue or to deny problems. The L scale assesses an unsophisticated tendency to claim excessive virtue; the K scale measures test defensiveness; and the S scale (Superlative Self-presentation) assesses the tendency of some people to present themselves in a highly favorable manner (Butcher & Han, 1995).

Three scales have been developed to assess the tendency that some respondents have to claim excessive mental health symptoms. The original F scale developed by Hathaway and McKinley (1940) assesses excessive symptom checking by attending to extreme or rare responses. Those who endorse a large number of rare items are thought to be presenting an unselective or exaggerated complaint pattern. High scores on F have been associated with malingering (Berry, Baer, & Harris, 1991; Schretlen, 1988).

The F(b) scale, an infrequency scale that covers extreme items in the back of the MMPI-2 items operates much like the original F scale (Butcher et al., 1989). The newest validity measure developed by Arbisi and Ben-Porath (1995) is an extremely valuable type of infrequency scale that differs from the original F scale in an important way: the F(p) scale assesses extreme responding in a clinical sample. The F(p) scale assesses infrequent responding in a clinical setting and thereby, when extreme, suggests malingering of psychiatric symptoms.

(ii) Clinical scales

The traditional MMPI clinical scales (Hathaway & McKinley, 1940) have been updated in the MMPI-2 (Butcher et al., 1989) and have been extensively revalidated in their revised form (Archer, Griffin, & Aiduk, 1995; Butcher & Williams, 1992; Graham & Ben-Porath, 1990; Graham & Butcher, 1988). The MMPI-2 clinical scales are empirically derived measures of several well-established clinical patterns such as hypochondriasis, depression, paranoia, and schizophrenia (see Table 2).

These scales have been well studied over more than 50 years. They were kept nearly intact in MMPI-2 because of their descriptive power and ability to generalize across groups.

Two scales on the clinical profile, Mf and Si, are not empirically derived scales as were the original clinical scales but were included on the instrument to assess personality characteristics of masculine–feminine interests and social introversion.

(iii) Content scales

The expanded item pool of the MMPI-2 allowed for the development of a new set of content-based scales to assess an expanded array of problems compared with the traditional clinical measures. The 15 content scales (see Table 3) were developed according to a multimethod/multistage strategy. The 567 items were subjected to a content analysis in order to derive homogeneous content groups rationally. This approach is not as theoretically blind as it seems, since the original MMPI content had been well delineated by Wiggins (1969) and many of these constructs were still available in a modified or reduced form in the revised instruments. Moreover, the MMPI-2 committee also wrote new items to cover contents that were limited or not available in MMPI-2, for example, suicide ideation, substance abuse, and type A behavior. It therefore became possible to develop scales for a number of problem areas that have clinical relevance and are sufficiently large to assess reliably. Following provisional derivation, other scale construction strategies such as internal consistency, item scale correlation, and external validation were followed to define and improve the item groups.

Content scales provide a different type of information to the clinical interpreter's resources. These scales are viewed as direct communications between the client and the clinician. The obvious content on the scales enables the client to address problems he or she has that are considered important to address in the clinical intervention.

In addition to their value as summaries or themes considered pertinent by the patient the content scales have clearly established external validity. A number of studies have provided data on the external validity of the content scales (Ben-Porath, Butcher, & Graham, 1991; Butcher, Graham, Williams, & Ben-Porath, 1990).

4.14.2.1.3 Recent validity research for the MMPI-2

An empirically derived and based instrument requires substantial research validation in order to be useful. The MMPI-2 has been substantially researched and validated. The fact that the MMPI-2 clinical scales are the same items as in the original instrument means that all of the research on the traditional measures applies with the revised form.

Table 1 Personality characteristics associated with validity indicator elevations.

? Cannot say score

The total number of unanswered items. A defensive protocol with possible attenuation of scale scores is suggested if the ? raw score is more than 35.

L (Lie) scale

A measure of an unsophisticated or self-consciously "virtuous" test-taking attitude. Elevated scores (above 70 *T*) suggest that the individual is presenting himself or herself in an overly positive light, attempting to create an unrealistically favorable view of his or her adjustment.

F (Infrequency) scale

A high score (*T* above 90) suggests an exaggerated pattern of symptom checking that is inconsistent with accurate self-appraisal and suggests confusion, disorganization, or actual faking of mental illness. *T* scores above 100 invalidate the profile.

F(B) scale

A second infrequency scale that appears toward the back of the MMPI-2 item pool is used to assess exaggerated responding at the end of the test. This scale operates much like the original MMPI F scale in detecting malingering, random responding, or motivation to exaggerate symptoms.

F(p) scale

The Psychopathology Infrequency Scale F(p) was developed by Arbisi and Ben-Porath (1995) to assess infrequent responding in psychiatric settings. This scale is valuable in appraising the tendency for some people to exaggerate mental health symptoms in the context of patients with genuine psychological disorder.

K (Defensiveness) scale

Measures an individual's willingness to disclose personal information and discuss his or her problems. High K scores (*T* above 65) reflect an uncooperative attitude and an unwillingness or reluctance to disclose personal information. Low scores (*T* below 45) suggest openness and frankness. This scale is positively correlated with intelligence and educational level, which should be taken into account when interpreting the scores.

Variable response inconsistency

The VRIN scale consists of 49 pairs of specially selected items. The members of each VRIN item pair have either similar or opposite content; each pair is scored for the occurrence of an inconsistency in responses to the two items. The scale score is the total number of item pairs answered inconsistently. High VRIN scores are a warning that a test subject may have been answering the items in the inventory in an indiscriminate manner, and raise the possibility that the protocol may be invalid and that the profile is essentially uninterpretable.

True response inconsistency

The TRIN is made up of 20 pairs of items that are opposite in content. If a subject responds inconsistently by answering True to both items of certain pairs, one point is added to the TRIN score; if the subject responds inconsistently by answering False to certain item pairs, one point is subtracted. A very high TRIN score indicates a tendency to give True answers to the items indiscrimately ("acquiescence"), and a very low TRIN score indicates a tendency to answer False indiscriminately ("nonacquiescence"). (Negative TRIN scores are avoided by adding a constant to the raw score.) Very low or very high TRIN scores are a warning that the test subject may have been answering the inventory indiscriminately so that the profile may be invalid and uninterpretable.

S scale

A new MMPI-2 scale was developed in an effort to improve the assessment of highly virtuous responding. The subjects used in the development of the scale were 274 male airline applicants and the MMPI-2 normative sample (*N* = 1138 men and 1462 women). The S scale was initially developed by examining item response differences between airline pilot applicants, who tend to engage in superlative self-description in order to impress examiners, and normative men from the MMPI-2 restandardization sample. The scale was refined by using internal consistency methods to ensure high-scale homogeneity. A factor analysis of the resulting 50 item scale (S) yielded five factors named: Beliefs in "Human Goodness"; Serenity; Contentment with Life; Patience and Denial of Irritability and Anger; and Denial of Moral Flaws. Linear *T* score conversion tables were computed for both men and women separately and combined (unisex) using the MMPI-2 restandardization data sets. The S scale was shown to have a number of behavioral correlates reflecting the presentation of oneself as a well-controlled, problem-free person.

Table 2 Empirical correlates for MMPI-2 clinical scales.

Scale 1 (Hypochondriasis)

High-scoring people show: excessive bodily concern; somatic symptoms that tend to be vague and undefined; fatigue, pain, weakness; manifest anxiety; selfish, self-centered, and narcissistic behavior; pessimistic, defeatist, cynical outlook on life; dissatisfied and unhappy; make others miserable with whining, complaining behavior; demanding and critical of others; expresses hostility indirectly; rarely act out; dull, unenthusiastic, unambitious; ineffective in oral expression, longstanding health concerns; function at a reduced level of efficiency without major incapacity; not very responsive to therapy, tend to terminate therapy when therapist is seen as not giving enough attention and support; tend to seek medical solutions to life problems.

Scale 2 (Depression)

High-scoring people show: depressed, unhappy, and dysphoric behavior; they are pessimistic and self-deprecating; tend to feel guilty; report being sluggish; have somatic complaints such as weakness, fatigue, and loss of energy; are agitated, tense, highly-strung, and irritable; prone to worry; lack self-confidence; feel useless and unable to function; feels like a failure at school or on the job; introverted, shy, retiring, timid, and seclusive; aloof, maintain psychological distance; avoids interpersonal involvement; cautious, and conventional; has difficulty making decisions; nonaggressive; overcontrolled, deny impulses; make concessions to avoid conflict; motivated for therapy.

Scale 3 (Hysteria)

High-scoring people show: poor response to stress; they avoid responsibility through development of physical symptoms; have headaches, chest pains, weakness, and tachycardia, anxiety attacks; symptoms appear and disappear suddenly; lack insight about causes of symptoms; lack insight about own motives and feelings; lack anxiety, tension, and depression; rarely report delusions, hallucinations, or suspiciousness; psychologically immature, childish, and infantile; self-centered, narcissistic, and egocentric; expect attention and affection from others; use indirect and devious means to get attention and affection; do not express hostility and resentment openly; socially involved; friendly, talkative, and enthusiastic; superficial and immature in interpersonal relationships; show interest in others for selfish reasons; occasionally act out in sexual or aggressive manner with little apparent insight; initially enthusiastic about treatment; respond well to direct advice or suggestion; slow to gain insight into causes of own behavior; resistant to psychological interpretations.

Scale 4 (Psychopathic deviate)

High-scoring people show: antisocial behavior; rebellious toward authority figures; stormy family relationships; blame parents for problems; history of under-achievement in school; poor work history; marital problems; impulsive; strive for immediate gratification of impulses; do not plan well; act without considering consequences of actions; impatient; limited frustration tolerance; poor judgment; take risks; do not profit from experience; immature, childish, narcissistic, self-centered, and selfish; ostentatious, exhibitionistic; insensitive; interested in others in terms of *how* they can be used; likeable and usually create a good first impression; shallow, superficial relationships, unable to form warm attachments; extroverted, outgoing; talkative, active, energetic, and spontaneous; intelligent; assert self-confidence; have wide range of interests; lack definite goals; hostile, aggressive; sarcastic, cynical; resentful, rebellious; act out; antagonistic; aggressive outbursts, assaultive behavior; little guilt over negative behavior; may feign guilt and remorse when in trouble; free from disabling anxiety, depression, and psychotic symptoms; likely to have personality disorder diagnosis (antisocial or passive–aggressive); prone to worry; dissatisfied; show absence of deep emotional response; feel bored and empty; poor prognosis for change in therapy; blame others for problems; intellectualize; may agree to treatment to avoid jail or some other unpleasant experience but are likely to terminate before change is effected.

Scale 5 (Masculinity–femininity)

MALES

Scores of T over 80: Show conflicts about sexual identity; insecure in masculine role; effeminate; aesthetic and artistic interests; intelligent and capable; value cognitive pursuits; ambitious, competitive, and persevering; clever, clear-thinking, organized, logical; show good judgment and common sense; curious; creative, imaginative, and individualistic in approach to problems; sociable; sensitive to others; tolerant; capable of expressing warm feelings toward others; passive, dependent, and submissive; peace-loving; make concessions to avoid confrontations; good self-control; rarely act out. (The interpretation of high 5 scores should be tempered for males with advanced academic degrees.)

High T score between 70 and 79: May be viewed as sensitive; insightful; tolerant; effeminate; showing broad cultural interests; submissive, passive. (In clinical settings, the patient might show sex role confusion; or heterosexual adjustment problems.)

Table 2 (continued)

Low T < 35: "Macho" self-image, present self as extremely masculine; overemphasize strength and physical prowess; aggressive, thrill-seeking, adventurous, and reckless; coarse, crude, and vulgar; harbor doubts about own masculinity; have limited intellectual ability; narrow range of interests; inflexible and unoriginal approach to problems; prefer action to thought; are practical and nontheoretical; easy-going, leisurely and relaxed; cheerful, jolly, humorous; contented; willing to settle down; unaware of social stimulus value; lack insight into own motives; unsophisticated.

FEMALES
Scores of T over 70: Reject traditional female roles and activities; masculine interests in work, sports, hobbies; active, vigorous, and assertive; competitive, aggressive, and dominating; coarse, rough, and tough; outgoing, uninhibited, and self-confident; easy-going, relaxed, balanced; logical, calculated; unemotional, and unfriendly.

Low T < 35: Describe self in terms of stereotyped female role; doubts about own femininity; passive, submissive, and yielding; defer to males in decision-making; self-pity; complaining, fault finding; constricted; sensitive; modest; idealistic. (This interpretation for low 5 females does *not* apply for females with postgraduate degrees.)

Scale 6 (Paranoia)

Extremely high-scoring people (T > 80) show: frankly psychotic behavior; disturbed thinking; delusions of persecution and/or grandeur; ideas of reference; feel mistreated and picked on; angry and resentful; harbor grudges; use projection as defense; most frequently diagnosed as schizophrenia or paranoid state.

Moderate elevations (= 65–79 for males; 71–79 for females): Paranoid predisposition; sensitive; overly responsive to reactions of others; feel they are getting a raw deal from life; rationalize and blame others; suspicious and guarded; hostile, resentful, and argumentative; moralistic and rigid; overemphasizes rationality; poor prognosis for therapy; do not like to talk about emotional problems; difficulty in establishing rapport with therapist.

Extremely low (T < 35): should be interpreted with caution. In a clinical setting, low 6 scores, in the context of a defensive response set, may suggest frankly psychotic disorder; delusions, suspiciousness, ideas of reference; symptoms less obvious than for high scorers; evasive, defensive, guarded; shy, secretive, withdrawn.

Scale 7 (Psychasthenia)

High-scoring people show: anxious, tense, and agitated; high discomfort; worried and apprehensive; high strung and jumpy; difficulties in concentrating; introspective, ruminative; obsessive, and compulsive; feel insecure and inferior; lack self-confidence; self-doubting, self-critical, self-conscious, and self-derogatory; rigid and moralistic; maintain high standards for self and others; overly perfectionistic and conscientious; guilty and depressed; neat, orderly, organized, and meticulous; persistent; reliable; lack ingenuity and originality in problem solving; dull and formal; vacillates; are indecisive; distort importance of problems, overreact; shy; do not interact well socially; hard to get to know; worry about popularity and acceptance; sensitive, physical complaints; shows some insight into problems; intellectualize and rationalize resistant to interpretations in therapy; express hostility toward therapist; remain in therapy longer than most patients; makes slow but steady progress in therapy.

Scale 8 (Schizophrenia)

Very high scorers (= over 80–90) show: blatantly psychotic behavior; confused, disorganized, and disoriented; unusual thoughts or attitudes; delusions; hallucinations; poor judgment.

High (65–79): schizoid lifestyle; do not feel a part of social environment; feel isolated, alienated, and misunderstood; feel unaccepted by peers; withdrawn, seclusive, secretive, and inaccessible; avoid dealing with people and new situations; shy, aloof, and uninvolved; experience generalized anxiety; resentful, hostile, and aggressive; unable to express feelings; react to stress by withdrawing into fantasy and daydreaming; difficulty separating reality and fantasy; self-doubts; feel inferior, incompetent, and dissatisfied; sexual preoccupation, and sex role confusion; nonconforming, unusual, unconventional, and eccentric; vague, long-standing physical complaints; stubborn, moody, and opinionated; immature, and impulsive; highly-strung; imaginative; abstract, vague goals; lack basic information for problem-solving; poor prognosis for therapy; reluctant to relate in meaningful way to therapist; stay in therapy longer than most patients; may eventually come to trust therapist.

Scale 9 (Hypomania)

High-scoring people (T > 80) show: overactivity; accelerated speech; may have hallucinations or delusions of grandeur; energetic and talkative; prefer action to thought; wide range of interest; do not utilize energy wisely; do not see projects through to completion; creative, enterprising, and ingenious; little interest in routine or detail; easily bored and restless; low frustration tolerance; difficulty in inhibiting expression of impulses; episodes of irritability, hostility, and aggressive outbursts; unrealistic, unqualified optimism; grandiose aspirations; exaggerates self-worth and self-importance; unable to see own limitations; outgoing, sociable, and

Table 2 (continued)

gregarious; like to be around other people; create good first impression; friendly, pleasant, and enthusiastic; poised, self-confident; superficial relationships; manipulative, deceptive, unreliable; feelings of dissatisfaction; agitated; may have periodic episodes of depression; difficulties at school or work, resistant to interpretations in therapy; attend therapy irregularly; may terminate therapy prematurely; repeat problems in stereotyped manner; not likely to become dependent on therapists; becomes hostile and aggressive toward therapist.

Moderately elevated scores (T > 65, LE 79): Over-activity; exaggerated sense of self-worth; energetic and talkative; prefer action to thought; wide range of interest; do not utilize energy wisely; do not see projects through to completion; enterprising, and ingenious; lack interest in routine matters; become bored and restless easily; low frustration tolerance; impulsive; has episodes of irritability, hostility, and aggressive outbursts; unrealistic, overly optimistic at times; shows some grandiose aspirations; unable to see own limitations; outgoing, sociable, and gregarious; like to be around other people; create good first impression; friendly, pleasant, and enthusiastic; poised, self-confident; superficial relationships; manipulative, deceptive, unreliable; feelings of dissatisfaction; agitated; view therapy as unnecessary; resistant to interpretations in therapy; attend therapy irregularly; may terminate therapy prematurely; repeat problems in stereotyped manner; not likely to become dependent on therapists; become hostile and aggressive toward therapist.

Low scorers (T below 35): Low energy level; low activity level; lethargic, listless, apathetic, and phlegmatic; difficult to motivate; report chronic fatigue, physical exhaustion; depressed, anxious, and tense; reliable, responsible, and dependable; approach problems in conventional, practical, and reasonable way; lack self-confidence; sincere, quiet, modest, withdrawn, seclusive; unpopular; overcontrolled; unlikely to express feelings openly.

Scale 10 (Social introversion)

High-scoring people (> 65) show: socially introversion more comfortable alone or with a few close friends; reserved, shy, and retiring; uncomfortable around members of opposite sex; hard to get to know; sensitive to what others think; troubled by lack of involvement with other people; overcontrolled; not likely to display feelings openly; submissive and compliant; overly accepting of authority; serious, slow personal tempo; reliable, dependable; cautious, conventional, unoriginal in approach to problems; rigid, inflexible in attitudes and opinions; difficulty making even minor decisions; enjoys work; gain pleasure from productive personal achievement; tend to worry; are irritable and anxious; moody, experience guilt feelings; have episodes of depression or low mood.

Low (T < 45): sociable and extroverted; outgoing, gregarious, friendly and talkative; strong need to be around other people; mix well; intelligent, expressive, verbally fluent; active, energetic, vigorous; interested in status, power and recognition; seeks out competitive situations; have problem with impulse control; act without considering the consequences of actions; immature, self-indulgent; superficial, insincere relationships; manipulative, opportunistic; arouses resentment and hostility in others.

Adapted from Butcher (1989).

In addition, prior to publication of the MMPI-2 there were a number of validity studies conducted on the revised form. For example, the MMPI revision committee collected personality ratings on more than 800 couples included in the normative sample. These personality ratings clearly cross-validated a number of the original scales. Moreover, validation research was conducted on a number of samples including schizophrenics and depressives (Ben-Porath, Butcher, & Graham, 1991); marital problem families (Hjemboe & Butcher, 1991); potential child-abusing parents (Egeland, Erickson, Butcher, & Ben-Porath, 1991); alcoholics (Weed, Butcher, Ben-Porath, & McKenna, 1992); airline pilot applicants (Butcher, 1994); military personnel (Butcher, Jeffrey et al., 1990).

Since the MMPI-2 was published in 1989 a number of other validation studies have been published (Archer, Griffin, & Aiduk, 1995; Ben-Porath, McCully, & Almagor, 1993; Blake et al., 1992; Clark, 1996; Husband & Iguchi, 1995; Keller & Butcher, 1991; Khan, Welch, & Zillmer, 1993).

4.14.2.2 Basic Personality Inventory

The BPI (Jackson, 1989) was published as an alternative to the MMPI-2 for the global assessment of psychopathology. The key aims in developing the BPI were to produce a broad-band measure of psychological dysfunctioning as measured by the MMPI that was: (i) relatively short, (ii) incorporated modern principles of test construction, and (iii) showed

Table 3 Description of the MMPI-2 content scales.

1. Anxiety (ANX)

High scorers report general symptoms of anxiety including tension, somatic problems (i.e., heart pounding and shortness of breath), sleep difficulties, worries, and poor concentration. They fear losing their minds, find life a strain, and have difficulties making decisions. They appear to be readily aware of these symptoms and problems, are willing to admit to them.

2. Fears (FRS)

A high score indicates an individual with many specific fears. These specific fears can include blood; high places; money; animals such as snakes, mice, or spiders; leaving home; fire; storms and natural disasters; water; the dark; being indoors; and dirt.

3. Obsessiveness (OBS)

High scorers have tremendous difficulties making decisions and are likely to ruminate excessively about issues and problems, causing others to become impatient. Having to make changes distresses them, and they may report some compulsive behaviors such as counting or saving unimportant things. They are excessive worriers who frequently become overwhelmed by their own thoughts.

4. Depression (DEP)

High scorers on this scale show significant depression. They report feeling blue, uncertain about their future, and uninterested in their lives. They are likely to brood, be unhappy, cry easily, and feel hopeless and empty. They may report thoughts of suicide or wishes that they were dead. They may believe that they are condemned or have committed unpardonable sins. Other people may not be viewed as a source of support.

5. Health concerns (HEA)

Individuals with high scores report many physical symptoms across several body systems. Included are gastrointestinal symptoms (e.g., constipation, nausea and vomiting, stomach trouble), neurological problems (e.g., convulsions, dizzy and fainting spells, paralysis), sensory problems (e.g., poor hearing or eyesight), cardiovascular symptoms (e.g., heart or chest pains), skin problems, pain (e.g., headaches, neck pains), respiratory troubles (e.g., coughs, hay fever, or asthma). These individuals worry about their health and feel sicker than the average person.

6. Bizarre mentation (BIZ)

Psychotic thought processes characterize individuals high on the BIZ scale. They may report auditory, visual, or olfactory hallucinations and may recognize that their thoughts are strange and peculiar. Paranoid ideation (e.g., the belief that they are being plotted against or that someone is trying to poison them) may be reported as well. These individuals may feel that they have a special mission or powers.

7. Anger (ANG)

High scorers tend to have anger control problems. These individuals report being irritable, grouchy, impatient, hotheaded, annoyed, and stubborn. They sometimes feel like swearing or smashing things. They may lose self-control and report having been physically abusive towards people and objects.

8. Cynicism (CYN)

High scorers tend to show misanthropic beliefs. They expect hidden, negative motives behind the acts of others, for example, believing that most people are honest simply for fear of being caught. Other people are to be distrusted, for people use each other and are only friendly for selfish reasons. They likely hold negative attitudes about those close to them, including fellow workers, family, and friends.

9. Antisocial practices (ASP)

High scorers tend to show misanthropic attitudes like high scorers on the CYN scale. The high scorers on the ASP scale report problem behaviors during their school years and other antisocial practices like being in trouble with the law, stealing or shoplifting. They report sometimes enjoying the antics of criminals and believe that it is all right to get around the law, as long as it is not broken.

10. Type A (TPA)

High scorers report being hard-driving, fast-moving, and work-oriented individuals, who frequently become impatient, irritable, and annoyed. They do not like to wait or be interrupted. There is never enough time in a day for them to complete their tasks. They are direct and may be overbearing in their relationships with others.

11. Low self-esteem (LSE)

High scores on LSE characterize individuals with low opinions of themselves. They do not believe that they are liked by others or that they are important. They hold many negative attitudes about themselves including beliefs that they are unattractive, awkward, and clumsy, useless, and a burden to others. They certainly lack self-confidence, and find it hard to accept compliments from others. They may be overwhelmed by all the faults they see in themselves.

Table 3 (continued)

12. Social discomfort (SOD)
High scorers tend to be very uneasy around others, preferring to be by themselves. When in social situations, they are likely to sit alone, rather than joining in the group. They see themselves as shy and dislike parties and other group events.

13. Family problems (FAM)
High scorers tend to show considerable family discord. Their families are described as lacking in love, quarrelsome, and unpleasant. They even may report hating members of their families. Their childhood may be portrayed as abusive, and marriages seen as unhappy and lacking in affection.

14. Work interference (WRK)
High scorers tend to show behaviors or attitudes that are likely to contribute to poor work performance. Some of the problems relate to low self-confidence, concentration difficulties, obsessiveness, tension and pressure, and decision-making problems. Others suggest lack of family support for the career choice, personal questioning of career choice, and negative attitudes towards co-workers.

15. Negative treatment indicators (TRT)
High scorers tend to show negative attitudes towards doctors and mental health treatment. High scorers do not believe that anyone can understand or help them. They have issues or problems that they are not comfortable discussing with anyone. They may not want to change anything in their lives, nor do they feel that change is possible. They prefer giving up, rather than facing a crisis or difficulty.

Adapted from: Butcher, Graham, Williams, & Ben-Porath (1990).

empirical evidence of being able to discriminate between normal and dysfunctional persons as well as being able to predict pathological behavior.

The BPI is made up of 240 items, grouped into 12 × 20 item scales. Neurotic tendencies are measured through the scales of Hypochondriasis, Depression, Anxiety, Social Introversion, and Self-depreciation. Aspects of sociopathy are measured by Denial, Interpersonal Problems, Alienation, and Impulse Expression scales. Psychotic behavior is assessed by scales labeled Persecutory Ideas, Thinking Disorder, and to some degree, by the Deviation scale. The Deviation scale comprises 20 critical items that are intended to serve as the basis for further clinical follow up. In contrast, the definitions of the constructs reflected in the other 11 items on the BPI are based on the results of a multivariate analysis of the content underlying the MMPI and the Differential Personality Inventory (Jackson & Messick, 1971).

The strong internal psychometric properties of the BPI attest to its careful construction (Jackson, 1989). Item properties and item factor analyses support the internal structure of the instrument. Both internal consistencies and test–retest reliability estimate fall in the 0.70 to 0.80 range. Various validity studies (some published in the manual, others in the literature) show that the BPI can indeed discriminate between normal and non-normal (e.g., delinquent) persons and can, within psychiatric populations, predict a variety of clinical criteria. Norms exist for adolescents and adults, spanning a variety of sample types such as community, psychiatric, college, and forensic. However, these norms are almost entirely based on white populations. Moreover, most of the normative sample was collected using nonstandard data collection procedures. For example, booklets were mailed to subjects instead of being administered under standard conditions. In addition, each of the subjects in the normative sample responded to one-third of the items in the booklet, an artifact that makes it difficult to perform some analyses on the normative sample (e.g., alpha coefficients).

Among the difficulties that have been associated with the BPI is a lack of validity scales for identifying invalid response protocols. The one content scale that would logically appear to have some bearing on the issue of protocol validity is the Denial scale, which is described in the manual as a measure of lack of insight and lack of normal affect. Unfortunately, Denial appears to be a relatively weak scale in terms of its reliability and validity in various empirical studies (Holden et al., 1988; Jackson, 1989). Two other complaints that have been voiced are a lack of an established link to diagnostic categories and a lack of work on profile interpretation.

Overall, the BPI has shown psychometric potential as a general measure of psychopathology. The basic developmental work on it is sound. What the BPI now needs are more extensive norms along with further work on its clinical applicability.

4.14.2.3 Personality Assessment Inventory

The PAI (Morey, 1991) is another inventory of general psychopathology. The PAI is a 344 item self-report measured designed to screen for approximately the same pathological domains as the MMPI/MMPI-2. It is used to collect information related to diagnosis, and to provide input on treatment planning.

The PAI includes four validity scales, 11 clinical scales, five treatment scales and two interpersonal scales. The clinical scales are Somatic Complaints, Anxiety-related Disorders, Depression, Mania, Paranoia, Schizophrenia, Borderline Features, Antisocial Features, Alcohol Problems, and Drug Problems. Treatment scales are Aggression, Suicidal Ideation, Stress, Nonsupport, and Treatment Rejection. Interpersonal scales are Dominance and Warmth. Of these scales, 10 are further divided into subscales that are intended to measure distinct constructs. A total of 27 items are designated critical items which, according to the author, should be followed up. This does seem to be a relatively large number of scales to interpret meaningfully in view of the total number of items.

Evidence for the reliability of the 22 scales is generally good. Across normative, clinical and college samples, median alpha coefficients are all in the 0.80–0.90 range. One month test–retest coefficients are reported to be in the 0.80s in the manual and in the 0.70s in the literature (Boyle & Lennon, 1994). Interestingly, some of the lowest reliabilities are found for the validity scales, a phenomenon that has been found for the validity scales on other instruments such as the MMPI (e.g., Fekken & Holden, 1991) and may well be a function of range restriction in the scores on such scales.

Normative PAI data for the USA are extensive. Standardization samples for a census-matched group, a clinical sample representing 69 sites, and a college sample drawn from seven different US universities each number over 1000 respondents. Normative data in the manual are also reported separately by age, education, gender, and race. Normative information for other countries (Canada, Australia, the UK, etc.) would be desirable.

In view of the recency of the publication of the PAI, there are only a handful of studies that bear on its validity. The manual reports evidence of the concurrent validity of the PAI in the form of correlations with other measures of psychopathology. A number of other studies show that the PAI can discriminate between diagnostic groups. Boyle and Lennon (1994) showed that the PAI can distinguish normals, alcoholics, and schizophrenics. Schinka (1995) was further able to use the PAI to develop a typology for alcoholics that had validity with several external variables. Finally, Alterman and colleagues (1995) demonstrated that methadone maintenance patients scored differently from the normative populations.

Generally, preliminary data suggest that the PAI is a well-constructed and brief measure of psychopathology. More research on its clinical validity needs to be completed before it can be considered to be a useful measure of psychopathology in clinical settings. There are no data to support the PAI's use in lieu of the MMPI-2 which has a more substantial empirical database.

4.14.3 SPECIALIZED OR FOCUSED CLINICAL ASSESSMENT MEASURES

In this section we will address several other measures that have been developed for clinical assessment and research to assess specific or more narrowly focused characteristics rather than omnibus instruments such as the MMPI-2. Several of these instruments will be examined to illustrate their application and potential utility for evaluating specific problems in clinical settings.

4.14.3.1 Millon Clinical Multiaxial Inventory

The MCMI (Millon, 1977, 1987, 1994) was developed by Theodore Millon for making clinical diagnoses on patients. The MCMI was intended to improve upon the long-established MMPI. In contrast to the MMPI/MMPI-2, the MCMI was designed with fewer items; is based on an elaborate theory of personality and psychopathology; and explicitly focuses on diagnostic links to criteria from the *Diagnostic and statistical manual of mental disorders* (*DSM*).

The MCMI was developed rationally rather than empirically. Millon has stated in his three test manuals (1977, 1987, 1994) as well as elsewhere (Millon & Davis, 1995) that development of the MCMI is to be an ongoing process. To keep the MCMI maximally useful for clinical diagnosis and interpretation, it must be continually updated in view of theoretical refinements, empirical validation studies, and evolutions in the official *DSM* classification systems. Most updated test manuals leave the test user with the impression that the developer considered test revision a necessary evil. Very rarely do you see continuous improvement as a test developer's goal, in part because this ever-changing process makes the accumulation of a solid research base difficult.

All three MCMI versions comprise 175 true/false items. However, across versions, the exact test items have evolved through revision or replacement. The number of scales and validity indices that can be calculated from these items has also increased. The original version, the MCMI-I, had 20 clinical scales and two validity scales. The MCMI-II yielded 22 clinical scales and three validity scales. The current MCMI-III has 24 clinical scales, three modifying indices and a validity index. Many items appear on several scales making for great item overlap.

On the MCMI-III, 14 clinical scales assess personality patterns that relate to *DSM-IV* Axis II disorders. Another 10 scales measure clinical syndromes related to *DSM-IV* Axis I disorders. The modifying indices, Disclosure, Desirability, and Debasement, are correction factors applied to clinical scale scores to ameliorate respondents' tendencies to distort their responses. The validity index comprises four bizarre or highly improbable items meant to detect careless, random, or confused responding.

The relationship between scales and items is explicated in detail in the manual. Millon started with a theory-based approach to writing items, followed by an evaluation of the internal structure of the items, and finally engaged in an assessment of the diagnostic efficiency of each item for distinguishing among diagnostic groups before final placement of an item on a scale. Millon departed from usual psychometric practice in a way that results in some unfortunate complications. Item overlap across scales is permitted: on average, items appear on three different scales with differential weights. This makes scoring inordinately complex, which makes assessment of scale homogeneity complex, and in turn this makes evaluation of the empirical structure underlying the scales complex. There are technical solutions for these problems but the result is that the MCMI is not an easy instrument with which to work.

Despite its psychometric drawbacks, the theory underlying the MCMI is generally agreed to be elegant and a substantial asset (McCabe, 1987; Reynolds, 1992). Each dimension measured by the test has a clear conceptual link to Millon's theory of psychopathology. Such a theory allows for the generation of clinical inferences based on a small number of fundamental principles (Millon & Davis, 1995). Not only do these inferences guide measurement, but also they enhance understanding of the constructs, bear on practical treatment decision, and produce research hypotheses.

One of the stated goals of the MCMI is to place patients into target diagnostic groups. To this end, the MCMI scales are directly coordinated with the *DSM* diagnostic categories. How well does the MCMI live up to its aim? The manuals report good evidence of diagnostic efficiency. However, the recent literature (available on the MCMI-I and MCMI-II) suggests the following three generalizations. First, the MCMI has only modest accuracy for assigning patients to diagnostic groups across a variety of clinical criteria (e.g., Chick, Martin, Nevels, & Cotton, 1994; Chick, Sheaffer, Goggin, & Sison, 1993; Flynn, 1995; Hills, 1995; Inch & Crossley, 1993; Patrick, 1993; Soldz, Budman, Demby, & Merry, 1993). Second, the MCMI may be better at predicting the absence than the presence of a disorder (Chick et al., 1993; Hills, 1995; Soldz et al., 1993). Third, the MCMI may be better at predicting some types of disorders than others but there is little agreement on which ones (Inch & Crossley, 1993; Soldz et al., 1993).

One source of the difficulty may be the base rate scores. Raw scores on scales are weighted and converted to base rate scores. The base rate scores reflect the prevalence of a particular personality disorder or pathological characteristic in the overall population. Their use is intended to maximize the number of correct classifications relative to the number of incorrect classifications when using the MCMI to make diagnoses (Millon & Davis, 1995). If the estimated base rates for the various diagnostic categories are poor, then the predictive accuracy of the MCMI can be expected to be poor (Reynolds, 1992).

One negative consequence of the type of norms used in the development of the MCMI inventories is that they do not discriminate between patients and normals. Use of the MCMI assumes that the subject is a psychiatric patient. Consequently, the MCMI overpathologizes individuals who are not actually patients. The MCMI should not be used where issues of normality need to be addressed. For example, if the test were used in family custody evaluations or personnel screening, the test interpretation would appear very pathological—it cannot do otherwise.

How does the MCMI compare to the MMPI? The MCMI publisher appears to emphasize that the MCMI and MMPI-2 measure different characteristics and the MCMI is shorter to administer to patients. Whereas the MMPI measures a broad range of psychopathology, the MCMI has its premier focus on the assessment of personality disorders. Consonant with its rational construction, the elaborate theoretical underpinnings of the MCMI are impressive. In contrast, however, the test literature that supports the validity of the MMPI/MMPI-2 is not available for the MCMI. Validation research on it has not proceeded at a very high pace. Whether the MCMI will have either the clinical utility or the heuristic value that the MMPI enjoys remains unanswered until more clinical research, and perhaps more refinements, are undertaken with the MCMI.

4.14.3.2 The Beck Depression Inventory

The BDI was first introduced in 1961, and it has been revised several times since (Beck et al., 1988). The BDI has been widely used as an assessment instrument in gauging the intensity of depression in patients who meet clinical diagnostic criteria for depressive syndromes. However, the BDI has also found a place in research with normal populations, where the focus of use has been on detecting depression or depressive ideation.

The BDI was developed in a manner similar to the MMPI: clinical observations of symptoms and attitudes among depressed patients were contrasted to those among nondepressed patients in order to obtain differentiation of the depressed group from the rest of the psychiatric patients. The 21 symptoms and attitudes contained in the BDI reflect the intensity of the depression; items receive a rating of zero to three to reflect their intensity and are summed linearly to create a score which ranges from 0 to 63. The 21 items included reflect a variety of symptoms and attitudes commonly found among clinically depressed individuals (e.g., Mood, Self-dislike, Social Withdrawal, Sleep Disturbance). The BDI administration is straightforward, and it can be given as an interview by the clinician or as a self-report instrument (requiring a fifth or sixth grade reading level).

The BDI is interpreted through the use of cut-off scores. Cut-off scores may be derived based on the use of the instrument (i.e., if a clinician wishes to identify very severe depression, then the cut-off score would be set high). According to Beck et al. (1988), the Center for Cognitive Therapy has set the following guidelines for BDI cut-off scores to be used with affective disorder patients: scores from 0 through 9 indicate no or minimal depression; scores from 10 through 18 indicate mild to moderate depression; scores from 19 through 29 indicate moderate to severe depression; and scores from 30 through 63 indicate severe depression.

Two important issues must be considered by clinicians regarding the results of the BDI. Unlike the MMPI/MMPI-2 and other major self-report instruments, the BDI has no safeguards against faking, lying, or variable response sets. Thus, clinicians are warned against this drawback of the BDI in assessing depressive thoughts and symptoms. In settings where faking or defensiveness are probable threats to the validity of the test, clinicians may need to reconsider their use of the BDI. The other issue pertains to the state–trait debate in assessment. The BDI is extremely sensitive to differences in the instructions given to an examinee such that certain instructions yield a state-like index of depressive thinking, whereas, other instructions yield a more trait-like index of depressive thinking. Again, clinicians are encouraged to use caution when administering the BDI and to tailor the administration instructions to the type of index (state or trait) that is desired.

In their review of the psychometric properties of the BDI, Beck et al. (1988) reported high internal consistency reliability of the instrument among both psychiatric and nonpsychiatric populations. The authors also reported that the BDI closely parallels the changes in both patient self-report and clinicians' ratings of depression (i.e., the BDI score accurately reflects changes in depressive thinking). Finally, they also presented evidence for the content, concurrent, discriminant, construct, and factorial validity of the BDI.

The acceptable reliability and validity of the BDI have helped make it a widely used objective index of depressive thinking among clinicians. Perhaps the most obvious use of the BDI is as an index of change in the level or intensity of depression. With an increasing focus on managed healthcare and accountability by psychotherapeutic service providers, the BDI offers a reliable and valid index of depressive symptoms and attitudes which can be used effectively to document changes brought about in therapy.

4.14.3.3 The State-Trait Anxiety Inventory

The STAI was developed by Spielberger et al. (1970) to measure anxiety from the perspective of states vs. traits. The state measurement assesses how the individual feels "right now" or at this moment. Subjects are asked to rate the intensity of their anxious feelings on a four point scale as to their experience of feelings in terms of: not at all, somewhat, moderately so, or very much so. The trait anxiety measure addresses how the individuals generally feel by rating themselves on a four-point scale: almost never, sometimes, often, or almost always.

Since it was developed in 1966 the STAI has been translated into over 48 different languages and has been widely researched in a variety of clinical and school settings (Spielberger, Ritterband, Sydeman, Reheiser, & Unger, 1995). The evidence for construct validity of the STAI comes from a variety of sources, for example, correlations with other anxiety measures (Spielberger, 1977), clinical settings (Spielberger, 1983), and medical and surgical patients (Spielberger, 1976). Nonetheless, there has been general debate in the literature about the conceptual and practical benefits of the trait vs. state distinction.

The STAI has become a very widely used measure in personality and psychopathology research in the USA and in other countries. However, the state-trait inventories have not been as broadly used as clinical assessment instruments.

4.14.3.4 Whitaker Index of Schizophrenic Thinking

The WIST (Whitaker, 1973, 1980) was developed to measure the type of thought impairment that differentiates between schizophrenic and "normal" thought processes. The WIST is intended to be individually administered as a screening tool or as one part of a battery of tests. Its multiple choice format makes the WIST a test that is easy to give and to score.

On the WIST, schizophrenic thought is defined as a discrepancy between actual and potential performance on cognitive reasoning tasks. This impaired thinking has three components: (i) a degree of illogicality, as reflected in inappropriate associations and false premises; (ii) a degree of impairment relative to previous performance, as reflected in slowness; and (iii) a degree of unwittingness, as reflected in unawareness of the incorrectness of responses. Whitaker's definition of schizophrenic thought is carefully explicated in the manual, making it possible for the test user to understand exactly what the WIST is measuring. Whitaker, however, has been criticized for basing his definition on a narrow reading of the literature on schizophrenic thought disorder (Payne, 1978).

The WIST itself is made up of 25 multiple choice items that are divided into three subtests: Similarities, Word Pairs, and New Inventions. Each item consists of a stimulus and five response options that differ in degree of illogicality. To illustrate, consider this sample item for the Similarities subtest: car: automobile, tires, my transportation, jar, smickle. The correct answer receives a score of 0; a loose association, reference idea, clang association, and nonsense association would receive scores of 1, 2, 3, and 4, respectively. The test administrator presents for a second time any items that the respondent answered incorrectly on the initial test taking.

Three scores are calculated for the WIST: total Score for all response alternatives selected either in the original or in the second enquiry phase; total Time for the initial completion of the WIST; and an overall Index that combines WIST Score and WIST Time. Presumably the WIST Score and Time components are added together because they both relate to schizophrenic thought disorder. However, a review of the empirical data supporting this conceptualization finds mixed results at best (Fekken, 1985). One view is that the predictive efficiency of the WIST is not enhanced by including the Time component.

There are two parallel forms of the WIST which differ in content. Items on Form A are intended to be stressful or anxiety provoking. They assess oral-dependency, hostility, and manifestly sexual content. The content of Form B is neutral. There has been little formal evaluation of the comparability of the two forms. Although some studies have reported similar rates of diagnostic efficiency for the two forms (Evans & Dinning, 1980; Leslie, Landmark, & Whitaker, 1984), overall validity evidence would suggest that Form A is stronger than Form B. Additional work to clarify the comparability of Forms A and B would have implications for selecting which form to use and for using these alternate WIST forms to assess changes in symptomatology.

There is relatively little information on reliability available on the WIST. The manual does report intratest reliabilities of the two WIST forms as Hoyt's reliability coefficients of around 0.80. Test–retest reliability data, either on the WIST subscores or subscales, are not provided in the manual nor do they appear to be readily available in the literature. Similarly, alternate forms of reliability, calculated as correlations between subscales, or subscores, are not readily available.

There is reasonable evidence for the convergent validity of the WIST. Two reviews report that WIST scores tend to have a 60–70% agreement with systems for diagnosing schizophrenia (Fekken, 1985; Grigoriadis, 1993). The WIST has been empirically associated with other indices of schizophrenia such as the MMPI/MMPI-2 Sc scale (Evans & Dinning, 1980; Fishkin, Lovallo, & Pishkin, 1977; Grigoriadis, 1993) although not with other schizophrenia indices including conceptually relevant SCL-90 scales (Dinning & Evans, 1977) and the New Haven Schizophrenic Index (Knight, Epstein, & Zielony, 1980).

Discriminant validity of the WIST has been harder to establish. The WIST has difficulty distinguishing among different psychiatric diagnostic groups (e.g., Burch, 1995; Pishkin, Lovallo, & Bourne, 1986). This is a serious shortcoming because the WIST is likely to be used in clinical settings precisely to help make such distinctions. A second problem with discriminant validity has been the tendency of the WIST to correlate negatively with measures of general cognitive ability. Based on such data one could share the view of at least one pessimistic reviewer and claim that the WIST has no demonstrated use (Payne, 1978).

Alternatively, the WIST may have a role in a comprehensive assessment battery. The WIST has never been promoted as a stand-alone test, nor has it ever been promoted as a comprehensive measure of the full range of schizophrenic symptomatology. Rather, the data on the WIST may be thought of as a general measure of cognitive deficit rather than cognitive deficit specific to schizophrenia. Thus, it may provide one objective source of data for accepting or rejecting a more general diagnosis of psychosis.

4.14.4 NORMAL RANGE PERSONALITY ASSESSMENT

Identification and description of psychological disorder is but one reason to administer a personality measure to an adult. Objective personality measures also are widely used outside the clinic or hospital setting for normal range assessment situations. Several objective personality measures will be discussed regarding their use in research, educational, vocational assessment, and personnel selection.

4.14.4.1 Objective Personality Measures in Research

4.14.4.1.1 The Five Factor Model (the Big Five)

One of the most popular current conceptualizations being studied today involves taxonomies of personality traits based on factor analytic methods and is commonly referred to as the "Big Five," or the Five Factor Model (FFM). As Butcher and Rouse (1996) note in their review, some personality researchers have rejected the FFM as an end-all to personality trait theory, while others in the area of personality research continue to embrace it. Among the proponents of the FFM are Costa and his colleagues, who have proposed the NEO Personality Inventory (NEO-PI) as a self-report measure of the FFM personality dimensions (Costa & McCrae, 1992). The FFM consists of five factor-analytically derived dimensions of personality. The five dimensions (are Neuroticism (N), Extroversion (E), Openness to Experience (O), Agreeableness (A), and Conscientiousness (C). Neuroticism has long been a familiar adjective among clinicians for describing people who tend to experience psychological distress. At the opposite end of the Neuroticism dimension is Emotional Stability, which represents the tendency to stay on a psychologically even keel. Extroversion encompasses the concepts of positive emotionality, sociability, and activity. The Extroversion dimension is rather broad in its scope, although lay people often use the term narrowly to refer to general "outgoingness." Openness to Experience represents, broadly, a person's level of constriction in their experiencing of the world; it is often associated with creativity (and even hypnotizability). Agreeableness represents the dimension of interpersonal behavior; the counterpart of Agreeableness is Antagonism. Finally, Conscientiousness represents a dimension of scrupulous organization of behavior.

4.14.4.1.2 NEO Personality Inventory

The NEO-PI is a 181-item inventory designed to index the FFM personality dimensions (N, E, O, A, and C); the NEO-PI also yields several subscales of the N, E, and O dimensions (Costa & McCrae, 1992). There is a self-report form of the NEO-PI as well as an observer-rating form. Although the authors of the NEO-PI argue for its utility in clinical settings, the instrument has been studied almost exclusively in nonclinical populations. One example of a recent investigation using the NEO-PI in research with psychiatric samples is the examination of differences in stimulus intensity modulation among older depressed individuals with and without mania features (Allard & Mishara, 1995). The NEO-PI was used to examine the hypothesis that stimulus intensity augmenters with unipolar depression would be introverted, whereas reducers with bipolar depression would be extroverted. Determinations of depression with and without mania features were based on scores on the MMPI. Costa and McCrae (1992) point to the NEO-PI as a useful assessment tool in aiding the clinician with understanding the client, selection of treatment, and even anticipating the course of therapy. However, the NEO-PI has not found wide use among clinicians to date (Butcher & Rouse, 1996).

Unfortunately, the NEO-PI is very susceptible to faking (Bailley & Ross, 1996) and does not contain validity indices to detect deviant response sets.

4.14.4.1.3 Multidimensional Personality Questionnaire

The Multidimensional Personality Questionnaire (MPQ) is a 300-item self-report instrument that was developed by Tellegen (unpublished manuscript) in an attempt to clarify the "self-view domain" in personality research. Tellegen used a classical iterative test construction approach involving several rounds of factor analysis to come up with the 11 primary scales, six validity scales, and three "higher-order" factors. The 11 primary scales (which include the dimensions of Social Potency, Control, Harm Avoidance, Well-being, Aggression, and others) load onto the three higher-order scales, which represent the familiar personality domains of Positive Affectivity, Negative Affectivity, and Constraint. The six validity indices include VRIN and TRIN, which are conceptually similar to those on the MMPI-2.

The MPQ is not as widely used as other objective personality measures in either the clinical or the research domain. However, recent investigations suggest a place for the MPQ in normal range and clinical personality assessment. For example, Kuhne, Orr, and Baraga (1993) demonstrated the utility of certain MPQ scales for discriminating among veterans with and without post-traumatic stress disorder. Krueger et al. (1994) administered the MPQ to adolescents in their community-based longitudinal study and found that certain MPQ scales were useful in distinguishing those who engaged in delinquency from those who abstained. These two reports suggest the utility of the MPQ both in clinical settings and in normal range personality assessment.

4.14.4.1.4 Personality Research Form

One well-known measure of normal personality is the Personality Research Form (PRF). Developed by Douglas N. Jackson (1984), the PRF is a true/false, multiscale measure of 20 of the psychosocial needs (e.g., achievement, aggression, sociability) originally defined by Henry Murray (1938). Many psychometrics texts hold up the PRF as a model of the construct approach to test construction. Indeed, much of the appeal of the PRF lies in its ability to measure a large number of normal personality characteristics while minimizing both scale intercorrelations and the influence of social desirability and acquiescence. Moreover, a variety of validity studies have been published in the 1980s and 1990s supporting the psychometric soundness of the PRF. Critics of the PRF complain that despite its technical elegance the PRF does not reflect an integrated model of personality, which limits its applicability to real-life testing situations. Recent research, however, shows that the content assessed by the PRF may be well described by the Five Factor structure (Paunomen, Jackson, Trzebinski, & Fosterling, 1992). In research the popularity of the PRF remains high as attested to by the number of references in bibliographies, such as the one produced by MacLennan in 1991, which lists over 375 studies featuring the PRF in the literature.

4.14.4.1.5 Sixteen Personality Factor Test

The 16PF was originally developed in the 1940s by Raymond Cattell to measure the primary factors of normal personality. At that time, Cattell's unique contribution was to apply factor analysis as a method for uncovering the full scope of personality. The fifth and current edition of the 16PF (Cattell, Cattell, & Cattell, 1993) measures the well-known 16 personality factors, plus it summarizes these factors into five global factors which again bear similarity to the well-known "Big Five." The global factors are: extroversion, anxiety, tough-mindedness, independence, and self-control. Relative to earlier editions, the fifth editions includes updated language; fewer items; and improved reliability and response style scales to measure impression management, infrequent responding, and acquiescence. Because of the recency of the publication of the fifth edition, there are few studies pertaining to its validity. However, a large database supports the validity of earlier editions of the 16PF. The 16PF has applicability in clinical, educational, organizational, and research settings. With the expansion of its interpretive reports and profiles, the 16PF would appear to be particularly useful in personal or vocational counseling settings.

4.14.4.1.6 California Psychological Inventory

The CPI developed by Gough (1957) is a multiscale, objective, self-report instrument used with normal range and psychiatric populations (Megargee, 1972). The CPI is similar in structure and content to the MMPI (and the MMPI-2). In fact, many of the items on the CPI are identical in wording to the items on the MMPI. The CPI focuses on "everyday" concepts about personality, such as dominance and responsibility among others. Eighteen scales (divided into four classes) are derived from the CPI; the results of the test are interpreted by reference to the plotted profile of standard scores. The body of literature using the CPI in normal range assessment is quite extensive in both quantity and breadth. Investigations include the use of the CPI to identify personality types, based on profiles, among a group of college students (Burger and Cross, 1979) and an examination of the underlying personality structure as measured by the CPI, again using a sample of college students (Deniston and Ramanaiah, 1993). The CPI also has been employed in the assessment of psychiatric samples. Especially noteworthy is the utility of the CPI with criminal samples (see Laufer, Skoog, and Day, 1982, for a relevant review of this literature).

Some specific applications of psychological tests in "normal range" settings are now described.

4.14.4.2 Objective Personality Measures in Educational/Vocational Assessment

The use of objective personality measures has become increasingly popular among professionals in the field of educational/vocational assessment. Research has identified the FFM, the NEO-PI, and the CPI as particularly useful in educational/vocational assessment.

4.14.4.2.1 FFM

The Big Five personality dimensions of the FFM have been studied in several contexts related to educational/vocational assessment. Moreover, because the FFM is a theoretical concept, various instruments have been used to assess the relationship between the Big Five personality dimensions and various educational/vocational assessment variables. Two recent investigations illustrate the utility of the Big Five personality dimensions in identifying candidates for admission to educational institutions and identifying characteristics among students in various university programs. Williams, Munick, Saiz, and Formy-Duval (1995) found that a mock graduate school admissions board (composed of graduate school faculty) favored for admission those hypothetical candidates whose applications reflected high ratings on the Big Five dimensions of Conscientiousness and Openness to Experience. Conversely, the Big Five dimensions of Extroversion and Agreeableness were not associated with a favorable impression of hypothetical candidates. Kline and Lapham (1992) assessed the Big Five personality dimensions among a group of college students to examine personality differences among the various fields of study (i.e., between different college majors). Students of various majors were not discriminated by levels of either Neuroticism or Extroversion. However, students in two fields of study (science and engineering) were marked by high ratings on the Big Five personality dimensions of Conscientiousness and Conventionality.

4.14.4.2.2 NEO-PI

The personality dimensions that underlie the NEO-PI appear to be related to at least one well-known typology of vocational personalities (which, in turn, correspond to vocational preferences). Gottfredson, Jones, and Holland (1993) found a relationship between the Big Five personality dimensions assessed by the NEO-PI and the six vocational personality dimensions proposed by Holland (assessed with the Vocational Preferences Inventory). In general, there was overlap among two to four of the significant factors extracted from each of the assessment instruments. However, the NEO-PI Neuroticism, Likability, and Control factors were not represented in the Holland vocational personality dimensions, which suggests a distinctive and qualitatively different role for objective personality assessment in vocational counseling. Other work with the NEO-PI has shown that two of the Big Five personality dimensions, Neuroticism and Agreeableness, are strongly related to occupational burnout among healthcare workers (Piedmont, 1993). Healthcare workers with higher ratings on the Neuroticism dimension were more likely to experience occupational burnout. Conversely, workers with higher ratings on the Agreeableness dimension were less likely to succumb to occupational burnout.

4.14.4.2.3 CPI

The CPI is one of the most widely used personality instruments in normal range assessment in educational/vocational settings. Walsh (1974) examined personality traits among college students identified as making hypothetical career choices that were either congruent or incongruent with their vocational personality type (as proposed by Holland). It was found that congruent students could be described by their CPI profiles as socially accepted, confident, and planful, whereas incongruent students could be described as impulsive, unambitious, and insecure. The well-known Strong Vocational Interest Blank (SVIB) used commonly among vocational and educational counselors is related to various personality traits as well. Johnson, Flammer and Nelson (1975) found a relationship among SVIB factors and CPI personality factors, especially those related to the global introversion/extroversion personality dimension.

4.14.4.2.4 MMPI/MMPI-2

Although it was not designed for educational research or placement purposes the MMPI/MMPI-2 has been among the most frequently employed instruments in this context. In a recent survey of test use in personnel and educational screening the MMPI has been employed effectively in a number of studies, for example: Anderson (1949), Appleby and Haner (1956), Applezweig (1953), Barger and Hall (1964), Barthol and Kirk (1956), Burgess

(1956), Centi (1962), Clark (1953, 1964), and Frick (1955) to mention only a few.

It appears evident that personality assessment contributes information independent of direct educational/vocational interest measures. There is extensive information available on the use of the MMPI/MMPI-2 in this setting. As demonstrated in the research with the NEO-PI and the CPI, objective personality assessment instruments yield information relevant to vocational assessment over and above that given by narrowly defined vocational personality assessments (like the one proposed by Holland). Moreover, elements of the FFM and independent personality constructs of the CPI have proven their utility in educational assessment in diverse areas such as admissions to educational institutions and choice of major fields of study. Clearly, objective personality assessment has carved out a valuable niche in normal range educational/vocational assessment.

4.14.4.3 Personnel Screening

Among the earliest and most extensive uses of personality tests with normals has been for the purpose of employment screening. The first formal North American, English language personality inventory, the Woodworth Personnel Data Sheet, was developed to screen out unfit draftees during World War I. A number of other personality questionnaires were developed in the 1930s to aid in personnel selection decisions.

The development and use of the MMPI during World War II provided a means for assessment psychologists to detect psychological problems that might make people unsuitable for key military assignments. Early research on the use of the MMPI centered around pilot selection and selection of nuclear submarine crewman. Following World War II, the MMPI came to be widely used in personnel selection particularly for occupations that required great responsibility or involved high stress such as air flight crews (Butcher, 1994; Cerf, 1947; Fulkerson, Freud, & Raynor, 1958; Fulkerson & Sells, 1958; Garetz & Tierney, 1962; Geist & Boyd, 1980; Goorney, 1970; Jennings, 1948); police and other law enforcement personnel (Beutler, Nussbaum, & Meredith, 1988; Beutler, Storm, Kirkish, Scogin, & Gaines, 1985; Butcher, 1991; Dyer, Sajwaj, & Ford, 1993; Hargrave & Hiatt, 1987; Saxe & Reiser, 1976; Scogin, & Beutler, 1986; Scogin & Reiser, 1976); and nuclear power employees (Lavin, Chardos, Ford & McGee, 1987).

Following the revision of the MMPI and publication of the MMPI-2, the revised instrument has been used in personnel screening (Butcher & Rouse, 1996) and currently is the most frequently used personality measure in personnel screening situations, particularly when the position is one that requires good mental health, emotional stability, and responsible behavior.

The MMPI-2 is usually employed in personnel selection to screen out candidates who are likely to have psychological problems from critical occupations such as police officer, airline pilot, and nuclear power control rooms, fire department, or air traffic control personnel.

4.14.4.4 Other Personality Measures in Personnel Selection

Objective measures of personality characteristics have a role in personnel selection similar to that in educational/vocational assessment. Namely, professionals in the field of personnel selection are interested in knowing which personality variables aid in the selection of quality employees. The literature suggests a valuable role for objective personality measures in the normal range assessment field of personnel selection.

4.14.4.5 The 16PF

One of the most widely used personality scales for employment screening is the 16PF. This inventory, with broad-ranging employment-relevant personality items, has been widely used in different contexts including: law enforcement (Burbeck & Furnham, 1985; Fabricatore, Azen, Schoentgen, & Snibbe, 1978; Hartman, 1987; Lawrence, 1984; Lorr & Strack, 1994; Topp & Kardash, 1986); pilots (Cooper & Green, 1976; Lardent, 1991); cabin crew personnel (Furnham, 1991); managers (Bartram, 1992; Bush, & Lucas, 1988; Chakrabarti, & Kundu, 1984; Henney, 1975); occupational therapists (Bailey, 1988); church counselors (Cerling, 1983) and teachers (Ferris, Bergin, & Wayne, 1988). The 16PF provides information about personality functioning and is typically employed to screen employees for positive personality features.

4.14.4.6 FFM

Barrick and Mount (1991) conducted a meta-analysis of the Big Five personality dimensions and their relationship to three criterion variables (job proficiency, training proficiency, and personnel data) within five occupational groups (professionals, police, managers, skilled/semiskilled and sales). Conscientiousness was related to each of the five occupational groups. Additionally, Conscientiousness was related to the three criterion variables. These findings lead the authors to conclude that Conscientiousness is a personality trait related to job performance across occupational types. In a similar vein, Dunn, Mount, Barrick, and Ones (1995) found that the Conscientiousness personality dimension was related to managers' ratings of applicant hireability. This held true for various job types (medical technologist, carpenter, secretary, etc.). Taken together, these two reports suggest that professionals in the field of personnel selection can gain valuable information from objective personality measures of the FFM. Specifically, the Conscientiousness dimension, when taken with other relevant application information, may be a valuable discriminator among prospective job applicants.

4.14.4.7 NEO-PI

While research suggests that the Big Five personality dimensions add qualitative information to the personnel selection process, it may be that the various instruments used to measure the Big Five have differential validity across populations. Schmit and Ryan (1993) administered a shortened version of the NEO-PI to a sample of college students and to a sample of government job applicants. The FFM structure fitted the student population better than it fitted the job applicant population, which suggests a note of caution in putting too much weight on the Big Five personality dimensions in personnel selection. The authors note that job applicants are under different situational demands, which may affect their approach to a personality questionnaire (e.g., they may adopt a defensive response style). The lack of validity scales to assess response styles clearly limits the NEO-PI for this application. A recent study by Bailley and Ross (1996) showed that the NEO-PI is quite vulnerable to faking and is limited in not having scales to detect deviant response attitudes.

4.14.4.7.1 CPI

The NEO-PI and other personality measures have been examined in personnel selection studies covering a wide range of occupational groups (managers, secretaries, carpenters, etc.). However, the CPI has a long tradition of use in the selection of a specific occupation: law enforcement officers. Hargrave and Hiatt (1989) examined the ability of the CPI scales to differentiate police cadets rated by their instructors as suitable or unsuitable for the job of law enforcement officer. Several scales (including Sense of Well-being, Sociability, and Social Presence) differentiated suitable from unsuitable cadets. In the second study of the report, the authors compared incumbent law enforcement officers who either had or had not experienced serious on-the-job problems (e.g., providing drugs to inmates, excessive use of force). The Socialization scale was among the best discriminators between groups. The authors concluded that the CPI is a useful aid to the selection of law enforcement officers.

In addition to the literature on law enforcement screening, the CPI may be useful as an index of work performance in other fields. Toward this end, Hoffman and Davis (1995) validated the Work Orientation and Managerial Potential scales for the CPI on groups of employees in an entertainment facility. However, several of the original CPI scales performed as well as the two new scales in predicting job performance, which questions the need for additional CPI scales in the selection of personnel.

In summary, the use of objective personality measures in the selection of job applicants has proven a worthwhile endeavor overall. The Big Five personality dimensions, especially Conscientiousness, might be useful to the personnel selection process especially in the assessment of conscientiousness. Professionals involved in the screening of law enforcement candidates can use the CPI to differentiate suitable from unsuitable candidates. The CPI may also be useful in discriminating among groups of job applicants (although its primary use has been in the law enforcement field).

The scales on the 16PF have been shown to be relevant to personality descriptions that are useful in personnel selection. Finally, when it comes to evaluating potential psychopathology in potential employees the MMPI-2 is usually the instrument most employed.

Although objective personality measures appear to have a place in personnel selection, all information on the candidate must be weighed, as several personality researchers have noted the problem of response sets associated with the situational demands of the application process. Specifically, job applicants may feel under pressure to make a good impression on their prospective employer and they may subsequently present themselves as overly virtuous or defensive.

4.14.5 SUMMARY

Human fascination with the concept of personality lead to the nineteenth century invention of the first objective personality test.

As with those very early personality measures, many of today's objective personality inventories are self-reports. The clinician's first concern when utilizing an objective personality measure is whether or not a client can accurately reveal information about his or her personality through a self-report instrument. Many objective personality instruments incorporate indices of test-taking attitudes (e.g., the Lie scale of the MMPI/MMPI-2) which allow the clinician to gauge a client's level of insight and willingness to self-disclose. Additionally, research shows that clients who do cooperate with testing produce personality profiles that match external criteria (e.g., clinician's notes and observations regarding the patient). Essentially, most clients are able to reveal their personalities competently through self-report measures.

Once a client is able to self-disclose information regarding his or her personality, the scale scores that are produced appear to be quite stable over time. Thus, most of the objective personality measures manage to capture trait (as opposed to state) characteristics. Five factors exhibit influence on the stability of personality: (i) instrument characteristics, (ii) length of retest interval, (iii) operationalization of personality as a stable construct for test construction, (iv) the extent to which a particular personality construct is associated with stability, and (v) person variables of the test-taker.

Several objective personality measures are designed to assess adults in clinical settings. The most well-known and widely used objective personality inventory in clinical settings is the MMPI-2, which provides a comprehensive survey of personality characteristics and clinical problems. The Basic Personality Inventory (BPI) and the Personality Assessment Inventory (PAI) are alternatives to the MMPI-2, but neither is as widely used as the MMPI-2. Several other objective measures have been developed for specialized or focused clinical use including the Millon Clinical Multiaxial Inventory (MCMI) designed for making personality diagnoses; the Beck Depression Inventory (BDI), which assesses depressive ideation; the State-Trait Anxiety Inventory (STAI) designed to assess both long-term and short-term anxiety features; and the Whitaker Index of Schizophrenic Thinking (WIST) designed to measure the type of thought impairment that differentiates between schizophrenic and "normal" thought processes.

While certain objective personality measures were developed for and have been used widely in clinical setting, several inventories have gained use in research and other normal range assessment settings. Many researchers utilize inventories that measure the Big Five (or Five Factor Model, FFM) personality traits: Neuroticism (N), Extroversion (E), Openness to Experience (O), Agreeableness (A), and Conscientiousness (C). The NEO Personality Inventory (NEO-PI) assesses N, E, and O and has been used as an index of the Big Five in research. The Multidimensional Personality Questionnaire (MPQ) assesses the broad personality domains of Positive Affectivity, Negative Affectivity, and Constraint. The MPQ contains two validity indexes (VRIN and TRIN) common to the MMPI-2, which make the MPQ appealing to many researchers. Other objective personality instruments that have been used widely in research include the Personality Research Form (PRF), a measure of several psychosocial needs; the 16PF, which measures global factors similar to the Big Five; and the California Psychological Inventory (CPI), designed to assess "everyday" concepts about personality.

Finally, professionals working in the fields of educational/vocational assessment and personnel selection have found use for several objective personality measures. The 16PF, the CPI, and, most commonly, the MMPI/MMPI-2 have been used in both of these settings. Objective personality instruments give professionals information about an individual's personality which would not be obtained through standard applications or interviews. Thus, professionals can use objective personality inventories as an efficient method of obtaining more information to aid them in their task of advising clients about educational/vocational decisions or advising employers in the selection of personnel.

Whether one needs a tool to assess adult personality in clinical, research, or industry settings, there is probably an objective personality inventory to fit the bill. Many of today's objective inventories offer the efficiency of providing a comprehensive assessment of personality functioning, and some even provide computerized interpretation of the personality profile. Perhaps most importantly, most of the objective personality inventories available commercially are standardized instruments that can assess adult personality validly and reliably, which means these instruments can be used over the course of a client's treatment to document goals for change—a feature that is becoming increasingly important to clinicians in this era of managed healthcare.

4.14.6 REFERENCES

Allard, C., & Mishara, B. L. (1995). Individual differences in stimulus intensity modulation and its relationship to two styles of depression in older adults. *Psychology and Aging, 10,* 395–403.

Alterman, A. I., Zaballero, A. R., Lin, M. M., Siddiqui, N., Brown, L. S., Jr., Rutherford, M. J., & McDermott, P. A. (1995). Personality Assessment Inventory scores of lower-socioeconomic African American and Latino methadone maintenance patients. *Assessment, 2,* 91–100.

Anderson, W. F. (1949). *Predicting success in nurses training*. Unpublished master's thesis, University of Nebraska, Lincoln, NE.

Appleby, T. L., & Haner, C. F. (1956). MMPI profiles of a college faculty group. *Proceedings of the Iowa Academy of Sciences, 53,* 605–609.

Applezweig, M. H. (1953). Educational levels and Minnesota Multiphasic profiles. *Journal of Clinical Psychology, 9,* 340–344.

Arbisi, P., & Ben-Porath, Y. S. (1995) An MMPI-2 infrequency scale for use with psychopathological populations: The Infrequency–Psychopathology Scale, F (p). *Psychological Assessment, 7,* 424–431.

Archer, R. P., Griffin, R., & Aiduk, R. (1995). Clinical correlates for ten common code types. *Journal of Personality Assessment, 65,* 391–408.

Assendorp, J. B. (1992). Beyond stability: Predicting inter-individual differences in intra-individual change. *European Journal of Personality, 6,* 103–117.

Bailey, D. M. (1988). Occupational therapy administrators and clinicians: Differences in demographics and values. *Occupational Therapy Journal of Research, 8,* 299–315.

Bailley, S. E., & Ross, S. R. (1996, May). *The effects of simulated faking on the five-factor (NEO-FFI)*. Paper given at the 68th Annual Meeting of Midwestern Psychological Association, Chicago.

Barger, B., & Hall, E. (1964). Personality patterns and achievement in college. *Educational and Psychological Measurement, 24,* 339–346.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44,* 1–26.

Barthol, R. P., & Kirk, B. A. (1956). The selection of graduate students in public health education. *Journal of Applied Psychology, 40,* 159–163.

Bartram, D. (1992). The personality of UK managers: 16PF norms for short-listed applicants. *Journal of Occupational and Organizational Psychology, 65,* 159–172.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8,* 77–100.

Ben-Porath, Y. S., Butcher, J. N., & Graham, J. R. (1991). Contribution of the MMPI-2 scales to the differential diagnosis of schizophrenia and major depression. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 634–640.

Ben-Porath, Y. S., McCully, E., & Almagor, M. (1993). Incremental validity of the MMPI-2 Content Scales in the assessment of personality and psychopathology by self-report. *Journal of Personality Assessment, 61,* 557–575.

Beutler, L. E., Nussbaum, P. D., & Meredith, K. E. (1988). Changing patterns of police officers. *Professional Psychology: Research and Practice, 19,* 503–507.

Beutler, L. E., Storm, A., Kirkish, P., Scogin, F., & Gaines, J. A. (1985). Parameters in the prediction of police officer performance. *Professional Psychology: Research and Practice, 16,* 324–335.

Blake, D. D., Penk, W. E., Mori, D. L., Kleespies, P. M., Walsh, S. S., & Keane, T. M. (1992). Validity and clinical scale comparisons between MMPI and MMPI-2 with psychiatric inpatients. *Psychological Reports, 70,* 323–332.

Boyle, G. J., & Lennon, T. J. (1994). Examination of the reliability and validity of the Personality Assessment Inventory. *Journal of Psychopathology and Behavioral Assessment, 16,* 173–187.

Burbeck, E., & Furnharm, A. (1985). Police officer selection: A critical review of the literature. *Journal of Police Science and Administration, 13,* 58–69.

Burch, J. W. (1995). Typicality range deficit in schizophrenics' recognition of emotion in faces. *Journal of Clinical Psychology, 51,* 140–150.

Burger, G. K., & Cross, D. T. (1979). Personality types as measured by the California Psychological Inventory. *Journal of Consulting and Clinical Psychology, 47,* 65–71.

Burgess, E. (1956). Personality factors of over- and under-achievers in engineering. *Journal of Educational Psychology, 47,* 89–99.

Bush, A. J., & Lucas, G. H. (1988). Personality profiles of marketing vs. R&D managers. *Psychology and Marketing, 5,* 17–32.

Butcher, J. N. (1989). *MMPI-2 scale correlates*. MMPI-2 Workshop Materials. Minneapolis MN: University of Minnesota Press.

Butcher, J. N. (1991). Screening for psychopathology: Industrial applications of the Minnesota Multiphasic Personality Inventory-2 (MMPI-2). In J. Jones, B. D. Steffey, & D. Bray (Eds.), *Applying psychology in business: The manager's handbook*. Boston: Lexington.

Butcher, J. N. (1994). Psychological assessment of airline pilot applicants with the MMPI-2. *Journal of Personality Assessment, 62,* 31–44.

Butcher, J. N. (1995). Clinical use of computer-based personality test reports. In J. N. Butcher (Ed.), *Clinical personality assessment: Practical approaches* (pp. 78–94). New York: Oxford University Press.

Butcher, J. N., Berah, E., Ellertsen, B., Miach, P., Lim, J., Nezami, E., Pancheri, P., Derksen, J., & Almagor, M. (1998). Objective personality assessment: Computer-based MMPI-2 interpretation in international clinical settings. In C. Belar (Ed.), *Comprehensive clinical psychology: Sociocultural and individual differences*. New York: Elsevier.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.

Butcher, J. N., Graham, J. R., Williams, C. L., & Ben-Porath, Y. S. (1990). *Development and use of the MMPI-2 Content Scales*. Minneapolis, MN University of Minnesota Press.

Butcher, J. N. & Han, K. (1995). Development of an MMPI-2 scale to assess the presentation of self in a superlative manner: The S Scale. In J. N. Butcher & C. D. Spielberger (Eds.), *Advances in personality assessment*, (Vol. 10, pp. 25–50). Hillsdale, NJ.: Erlbaum.

Butcher, J. N., Jeffrey, T., Cayton, T. G., Colligan, S., DeVore, J., & Minnegawa, R. (1990). A study of active duty military personnel with the MMPI-2. *Military Psychology, 2,* 47–61.

Butcher, J. N., & Rouse, S. V. (1996). Personality: Individual differences and clinical assessment. *Annual Review of Psychology, 47,* 87–111.

Butcher, J. N., & Williams, C. L. (1992). *MMPI-2 and MMPI-A: Essentials of clinical interpretation*. Minneapolis, MN: University of Minnesota Press.

Cattell, R. B. (1963). Personality, role, mood, and situation perception: A unifying theory of modulators. *Psychological Review, 70,* 1–18.

Cattell, R. B., Cattell, A. K. S., & Cattell, H. E. P. (1993). *Sixteen Personality Factor Questionnaire* (5th ed.). Champaign, IL: Institute for Personality and Ability Testing.

Centi, P. (1962). Personality factors related to college success. *Journal of Educational Research, 55,* 187–188.

Cerf, A. Z. (1947). Personality inventories. In J. P. Guilford

(Ed.), *Printed classification tests*. Washington, DC: Army Air Force Aviation Psychology Program Research Reports.

Cerling, G. L. (1983). Selection of lay counselors for a church counseling center. *Journal of Psychology and Christianity, 2,* 67–72.

Chakrabarti, P. K., & Kundu, R. (1984). Personality profiles of management personnel. *Psychological Studies, 29,* 143–146.

Chick, D., Martin, S. K., Nevels, R., & Cotton, C. R. (1994). Relationship between personality disorders and clinical symptoms in psychiatric inpatients as measured by the Millon Clinical Multiaxial Inventory. *Psychological Reports, 74,* 331–336.

Chick, D., Sheaffer, C. I., Goggin, W. C., & Sison, G. F. (1993). The relationship between MCMI personality scales and clinician generated *DSM-III* personality disorder diagnoses. *Journal Personality Assessment, 61,* 264–276.

Clark, D. L. (1964). Exploring behavior in men's residence halls using the MMPI. *Personnel Guidance Journal, 43,* 249–251.

Clark, J. H. (1953). Grade achievement of female college students in relation to non-intellective factors: MMPI items. *Journal of Social Psychology, 37,* 275–281.

Clark, M. E. (1996). MMPI-2 negative treatment indicators content and content component scales: Clinical correlates and outcome prediction for men with chronic pain. *Psychological Assessment, 8,* 32–47.

Conley, J. J. (1985). Longitudinal stability of personality traits: A multi trait–multi method–multi occasion analysis. *Journal of Personality and Social Psychology, 49,* 1266–1282.

Cooper, C. L., & Green, M. D. (1976). Coping with occupational stress among Royal Air Force personnel on isolated island bases. *Psychological Reports, 39,* 731–734.

Costa, P. T., Jr., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment, 4,* 5–13.

Crookes, T. G., & Buckley, S. J. (1976). Lie score and insight. *Irish Journal of Psychology, 3,* 134–136.

Deniston, W. M., & Ramanaiah, N. V. (1993). California Psychological Inventory and the five-factor model of personality. *Psychological Reports, 73,* 491–496.

Dinning, W. D., & Evans, R. G. (1977). Discriminant and convergent validity of the SCL-90 in psychiatric inpatients. *Journal of Personality Assessment, 41,* 304–310.

Dunn, W. S., Mount, M. K., Barrick, M. R., & Ones, D. S. (1995). Relative importance of personality and general mental ability in managers' judgments of applicant qualifications. *Journal of Applied Psychology, 80,* 500–509.

Dyer, J. B., Sajwaj, T. E. G., & Ford, T. W. X. (1993, March). *MMPI-2 normative and comparative data for nuclear power plant personnel who were approved or denied security clearances for psychological reasons*. Paper presented at the 28th Annual Symposium on Recent Developments in the Use of the MMPI/MMPI-2, St. Petersburg, FL.

Egeland, B., Erickson, M., Butcher, J. N., & Ben-Porath, Y. S. (1991). MMPI-2 profiles of women at risk for child abuse. *Journal of Personality Assessment, 57,* 254–263.

Evans, R. G., & Dinning, W. D. (1980). A validation of Forms A and B of the Whitaker Index of Schizophrenic Thinking. *Journal of Personality Assessment, 44,* 416–419.

Fabricatore, J., Azen, S. P., Schoentgen, S., & Snibbe, H. (1978). Predicting performance of police officers using the Sixteen Personality Factor Questionnaire. *American Journal of Community Psychology, 6,* 63–70.

Fekken, G. C. (1985). The Whitaker Index of Schizophrenic Thinking. In D. J. Keyser & R. C. Sweetland (Eds.), *Test critiques* (Vol. III, pp. 717–725). Kansas City, KS: Test Corporation Of America.

Fekken, G. C., & Holden, R. R. (1991). The construct validity of person reliability. *Personality and Individual Differences, 12,* 69–77.

Ferris, G. R., Bergin, T. G., & Wayne, S. J. (1988). Personal characteristics, job performance, and absenteeism of public school teachers. *Journal of Applied Social Psychology, 18,* 552–563.

Finn, S. E. (1986). Stability of personality self-ratings over 30 years: Evidence for an age/cohort interaction. *Journal of Personality and Social Psychology, 50,* 813–818.

Fishkin, S. M., Lovallo, W. R., & Pishkin, V. (1977). Relationship between schizophrenic thinking and MMPI for process and reactive patients. *Journal of Clinical Psychology, 33,* 116–119.

Flynn, P. M. (1995). Issues in the assessment of personality disorder and substance abuse using the Millon Clinical Multiaxial Inventory (MCMI-II). *Journal of Clinical Psychology, 51,* 415–421.

Frick, J. W. (1955). Improving the prediction of academic achievement by use of the MMPI. *Journal of Applied Psychology, 39,* 49–52.

Fulkerson, S. C., Freud, S. L., & Raynor, G. H. (1958, February). The use of the MMPI in psychological evaluation of pilots. *Aviation Medicine* 122–128.

Fulkerson, S. C., & Sells, S. B. (1958). Adaptation of the MMPI for aeromedical practice norms for military pilots. *USAF School of Aviation Medicine Reports,* 58–128.

Furnham, A. (1991). Personality and occupational success: 16PF correlates of cabin crew performance. *Personality and Individual Differences, 12,* 87–90.

Garetz, F. K., & Tierney, R. W. (1962). Personality variables in army officer candidates. *Military Medicine, 127,* 669–672.

Geist, C. R., & Boyd, S. T. (1980). Personality characteristics of Army helicopter pilots. *Perceptual Motor Skills, 51*(1), 253–254.

Goorney, A. B. (1970). MMPI and MMPI scores, correlations and analysis for military aircrew population. *British Journal of Social and Clinical Psychology, 9,* 164–170.

Gottfredson, G. D., Jones, E. M., & Holland, J. L. (1993). Personality and vocational interests: The relation of Holland's six interest dimensions to five robust dimensions of personality. *Journal of Counseling Psychology, 40,* 518–524.

Gough, H. G. (1957). *Manual for the California Psychological Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Graham, J. R., & Ben-Porath, Y. S. (1990, June). *Congruence between the MMPI and MMPI-2.* Paper given at the 25th Annual Symposium on Recent Developments in the Use of the MMPI/MMPI-2, Minneapolis, MN.

Graham, J. R., & Butcher, J. N. (1988, March). *Differentiating schizophrenic and major affective disorders with the revised form of the MMPI.* Paper presented at the 23rd Annual Symposium on Recent Developments in the Use of the MMPI, St. Petersburg, FL.

Grigoriadis, S. (1993). *Sources of inconsistency on tests of psychopathology*. Unpublished doctoral dissertation, Queen's University, Kingston, ON.

Hargrave, G. E., & Hiatt, D. (1987, May). *Use of the MMPI to predict aggression in law enforcement officer applicants.* Paper presented at the 22nd Annual Symposium on Recent Developments in the Use of the MMPI, Seattle, WA.

Hargrave, G. E., & Hiatt, D. (1989). Use of the California Psychological Inventory in law enforcement officer selection. *Journal of Personality Assessment, 53,* 267–277.

Hartman, B. J. (1987). Psychological screening of law enforcement candidates. *American Journal of Forensic Psychology, 5,* 5–10.

Hathaway, S. R., & McKinley, J. C. (1940). A multiphasic personality schedule (Minnesota): 1. Construction of the schedule. *Journal of Psychology, 10,* 249–254.

Helson, R., & Moane, G. (1987). Personality change in women from college to midlife. *Journal of Personality and Social Psychology, 53,* 176–186.

Henney, A. S. (1975). Personality characteristics of a group of industrial managers. *Journal of Occupational Psychology, 48,* 65–67.

Hills, H. A. (1995). Diagnosing personality disorders: An examination of the MMPI-2 and MCMI-II. *Journal of Personality Assessment, 65,* 21–34.

Hindelang, M. J. (1972). The relationships of self-reported delinquency to scales of the CPI and MMPI. *Journal of Criminal Law, Criminology, and Police Science, 63,* 75–81.

Hjemboe, S., & Butcher, J. N. (1991). Couples in marital distress: A study of demographic and personality factors as measured by the MMPI-2. *Journal of Personality Assessment, 57,* 216–237.

Hoffman, R. G., & Davis, G. L. (1995). Prospective validity study: CPI work orientation and managerial potential scales. *Educational and Psychological Measurement, 55,* 881–890.

Holden, R. R., Fekken, G. C., Reddon, J. R., Helmes, E., & Jackson, D. N. (1988). Clinical reliabilities and validities of the Basic Personality Inventory. *Journal of Consulting and Clinical Psychology, 56,* 766–768.

Husband, S. D., & Iguchi, M. (1995). Comparison of MMPI-2 and MMPI clinical scales and high point scores among methadone maintenance clients. *Journal of Personality Assessment, 64,* 371–375.

Inch, R., & Crossley, M. (1993). Diagnostic utility of the MCMI-I and MCMI-II with psychiatric outpatients. *Journal of Clinical Psychology, 49,* 358–366.

Jackson, D. N. (1984). *Personality Research Form manual* (3rd ed.). Port Huron, MI: Research Psychologists Press.

Jackson, D. N. (1989). *Basic Personality Inventory manual.* Port Huron, MI: Research Psychologists Press.

Jackson, D. N., & Messick, S. (1971). *The Differential Personality Inventory*. London, ON: Authors.

Jaffe, L. T. & Archer, R. P. (1987). The prediction of drug use among college students from MMPI, MCMI, and sensation seeking scales. *Journal of Personality Assessment, 51,* 243–253.

Jennings, L. S. (1948). Minnesota Multiphasic Personality Inventory; differentiation of psychologically good and poor combat risks among flying personnel. *Journal of Aviation Medicine, 19,* 222.

Johnson, R. W., Flammer, D. P., & Nelson, J. G. (1975). Multiple correlations between personality factors and SVIB occupational scales. *Journal of Counseling Psychology, 22,* 217–223.

Keilen, W. G., & Bloom, L. J. (1986). Child custody evaluation practices: A survey of experienced professionals. *Professional Psychology: Research and Practice, 17,* 338–346.

Keller, L. S., & Butcher, J. N. (1991). *Use of the MMPI-2 with chronic pain patients.* Minneapolis, MN: University of Minnesota Press.

Khan, F. I., Welch, T., & Zillmer, E. (1998). MMPI-2 profiles of battered women in transition. *Journal of Personality Assessment, 60,* 100–111.

Kline, P., & Lapham, S. L. (1992). Personality and faculty in British universities. *Personality and Individual Differences, 13,* 855–857.

Knight, R. A., Epstein, B., & Zielony, R. D. (1980). The validity of the Whitaker Index of Schizophrenic Thinking. *Journal of Clinical Psychology, 36,* 632–639.

Koss, M. P., & Butcher, J. N. (1973). A comparison of psychiatric patients' self-report with other sources of clinical information. *Journal of Research in Personality, 7,* 225–236.

Krueger, R. F., Schmutte, P. S., Caspi, A., Moffitt, T. E., Campbell, K., & Silva, P. A. (1994). Personality traits are linked to crime among men and women: Evidence from a birth cohort. *Journal of Abnormal Psychology, 103,* 328–338.

Kuhne, A., Orr, S., & Baraga, E. (1993). Psychometric evaluation of post-traumatic stress disorder: The Multidimensional Personality Questionnaire as an adjunct to the MMPI. *Journal of Clinical Psychology, 49,* 218–225.

Lardent, C. L. (1991). Pilots who crash: Personality constructs underlying accident prone behavior of fighter pilots. *Multivariate Experimental Clinical Research, 10,* 1–25.

Laufer, W. S., Skoog, D. K., & Day, J. M. (1982). Personality and criminality: A review of the California Psychological Inventory. *Journal of Clinical Psychology, 38,* 562–573.

Lavin, P. F., Chardos, S. P., Ford, W. T., & McGee, R. K. (1987). The MMPI profiles of troubled employees in relation to nuclear power plant personnel norms. *Transactions of the American Nuclear Society, 54,* 146–147.

Lawrence, R. A. (1984). Police stress and personality factors: A conceptual model. *Journal of Criminal Justice, 12,* 247–263.

Leslie, B. A., Landmark, J., & Whitaker, L. C. (1984). The Whitaker Index of Schizophrenic Thought (WIST) and thirteen systems for diagnosing schizophrenia. *Journal of Clinical Psychology, 40,* 636–648.

Lorr, M. & Strack, S. (1994). Personality profiles of police candidates. *Journal of Clinical Psychology, 50,* 200–207.

Lubin, B., Larsen, R. M., & Matarazzo, J. D. (1984). Patterns of psychological test usage in the United States: 1935–1982. *American Psychologist, 39,* 451–454.

Lumsden, J. (1977). Person reliability. *Applied Psychological Measurement, 1,* 477–482.

MacCabe, S. P. (1987). Millon Clinical Multiaxial Inventory. In D. J. Keyser & R. C. Sweetland (Eds.), *Test Critiques Compendium* (pp. 304–315). Kansas City, KS: Test Corporation Of America.

MacLennan, R. N. (1991). *Personality Research Form annotated research bibliography.* Regina, AL: University of Regina, Department of Psychology.

McCrae, R. R., & Costa, P. T., Jr. (1990). *Personality in adulthood.* New York: Guilford.

Megargee, E. I. (1972). *The California Psychological Inventory handbook*. San Francisco: Jossey-Bass.

Millon, T. (1977). *Manual for the Millon Clinical Multiaxial Inventory*. Minneapolis, MN: National Computer Systems.

Millon, T. (1987). *Manual for the Millon Clinical Multiaxial Inventory-II.* Minneapolis, MN: National Computer Systems.

Millon, T. (1994). *Manual for the Millon Clinical Multiaxial Inventory-III.* Minneapolis, MN: National Computer Systems.

Millon, T., & Davis, R. (1995). Putting Humpty Dumpty together again: Using the MCMI in psychological assessment. In L. E. Beutler & M. R. Berren (Eds.), *Integrative assessment of adult personality* (pp. 240–279). New York: Guilford.

Morey, L. (1991). *Personality Assessment Inventory: Professional Manual.* Odessa, FL: Psychological Assessment Resources.

Murray, H. A. (1938). *Explorations in personality*. Cambridge, MA: Harvard University Press.

Patrick, J. (1993). Validation of the MCMI-I Borderline Personality Disorder scale with a well-defined criterion sample. *Journal of Clinical Psychology, 49,* 28–32.

Paunonen, S. V., Jackson, D. N., Trzebinski, J., & Fosterling, F. (1992). Personality structures across cultures: A multi method evaluation. *Journal of Personality and Social Psychology, 62,* 447–456.

Payne, F. D., & Wiggins, J. S. (1972). MMPI profile types and the self-report of psychiatric patients. *Journal of Abnormal Psychology, 79,* 1–8.

Payne, R. W. (1978). Review of Whitaker Index of Schizophrenic Thought. In O. K. Buros (Ed.), *The eighth mental measurements yearbook* (pp. 1146–1147). Highland Park, NJ: Gryphon Press.

Piedmont, R. L. (1993). A longitudinal analysis of burnout in the health care setting: The role of personal dispositions. *Journal of Personality Assessment, 61,* 457–473.

Pishkin, V., Lovallo, W. R., & Bourne, L. E. (1986). Thought disorder and schizophrenia: Isolating and timing a mental event. *Journal of Clinical Psychology, 42,* 417–424.

Reynolds, C. R. (1992). Review of the Millon Clinical Multiaxial Inventory-II. In J. J. Kramer & J. C. Conoley (Eds.), *The eleventh mental measurements yearbook* (pp. 533–535). Lincoln, NE: Buros Institute of Mental Measurements.

Saxe, S. J., & Reiser, M. (1976). A comparison of three police applicant groups using the MMPI. *Journal of Police Science and Administration, 4,* 419–425.

Schinka, J. A. (1995). Personality Assessment Inventory scale characteristics and factor structure in the assessment of alcohol dependency. *Journal of Personality Assessment, 64,* 101–111.

Schmit, M. J., & Ryan, A. M. (1993). The Big Five in personnel selection: Factor structure in applicant and nonapplicant populations. *Journal of Applied Psychology, 78,* 966–974.

Schretlen, D. (1988). The use of psychological tests to identify malingered symptoms of mental disorder. *Clinical Psychology Review, 8,* 451–476.

Schuerger, J. M., Zarrella, K. L., & Hotz, A. S. (1989). Factors that influence the temporal stability of personality by questionnaire. *Journal of Personality and Social Psychology, 56,* 777–783.

Scogin, F., & Beutler, L. (1986). Psychological screening of law enforcement candidates. In P. A. Keller & L. G. Ritt (Eds.), *Innovations in clinical practice* (Vol. 5, pp. 317–330). Sarasota, FL: Professional Resources Exchange.

Scogin, F., & Reiser, M. (1976). A comparison of three police applicant groups using the MMPI. *Journal of Police Science and Administration, 4,* 419–425.

Soldz, S., Budman, S., Demby, A., & Merry, J. (1993). Diagnostic agreement between the Personality Disorder Examination and the MCMI-II. *Journal of Personality Assessment, 60,* 486–499.

Spielberger, C. D. (1977). Anxiety, theory and research. In B. Wolman (Ed.) *International encyclopedia of neurology, psychiatry, psychoanalysis, and psychology*. New York: Human Sciences Press.

Spielberger, C. D. (1976). Stress and anxiety and cardiovascular disease. *Journal of South Carolina Medical Association* (Supplement 15), *72,* 15–22.

Spielberger, C. D. (1983). *Manual for the State Trait Anxiety Inventory: STAI (form Y)*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. D. (1970). The STAI: Manual for the State–Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., Ritterband, L. M., Sydeman, S. J., Reheiser, E. C., and Unger, K. K. (1995). Assessment of emotional states and personality traits: Measuring psychological vital signs. In J. N. Butcher (Ed.), *Clinical personality assessment: Practical approaches* (pp. 43–58). New York: Oxford University Press.

Tellegen, A. *Brief manual for the Differential Personality Questionnaire*. Unpublished manuscript.

Topp, B. W., & Kardash, C. A. (1986). Personality, achievement, and attrition: Validation in a multiple-jurisdiction police academy. *Journal of Police Science and Administration, 14,* 234–241.

Walsh, W. B. (1974). Consistent occupational preferences and personality. *Journal of Vocational Behavior, 4,* 145–153.

Watkins, C. E. (1996). On Hunsley, Harangue, and Hoopla. *Professional Psychology: Research and Practice, 27,* 316–318.

Weed, N. C., Butcher, J. N., Ben-Porath, Y. S., & McKenna, T. (1992). New measures for assessing alcohol and drug abuse with the MMPI-2: The APS and AAS. *Journal of Personality Assessment, 58,* 389–404.

Whitaker, L. C. (1973). *The Whitaker Index of Schizophrenic Thinking*. Los Angeles: Western Psychological Services.

Whitaker, L. C. (1980). *Objective measurement of schizophrenic thinking: A practical and theoretical guide to the Whitaker Index of Schizophrenic Thinking*. Los Angeles: Western Psychological Services.

Whitbourne, S. K., Zuschlag, L. B., Elliot, L. B., & Waterman, A. S. (1992). Psychosocial development in adulthood: A 22-year sequential study. *Journal of Personality and Social Psychology, 63,* 260–271.

Wiggins, J. S. (1969). Content dimensions in the MMPI. In J. N. Butcher (Ed.), *MMPI: Research developments and clinical applications* (pp. 127–180).

Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment*. Reading, MA: Addison-Wesley.

Williams, J. E., Munick, M. L., Saiz, J. L., & Formy-Duval, D. L. (1995). Psychological importance of the "Big Five": Impression formation and context effects. *Personality and Social Psychology Bulletin, 21,* 818–826.

Windle, C. (1954). Test–retest effect on personality questionnaires. *Educational and Psychological Measurement, 14,* 617–633.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.15
Projective Assessment of Children and Adolescents

IRVING B. WEINER and KATHRYN KUEHNLE
University of South Florida, Tampa, FL, USA

4.15.1 INTRODUCTION 432

4.15.2 OBJECTIVITY AND SUBJECTIVITY IN PERSONALITY ASSESSMENT METHODS 432

4.15.3 STRUCTURE AND AMBIGUITY IN PROJECTIVE TECHNIQUES 433

4.15.4 VALUE OF PROJECTIVE ASSESSMENT 435
- *4.15.4.1 Conceptual Basis* 435
- *4.15.4.2 Empirical Basis* 436

4.15.5 UTILITY OF PROJECTIVE ASSESSMENT 436

4.15.6 APPLICABILITY OF PROJECTIVE METHODS TO CHILDREN AND ADOLESCENTS 437

4.15.7 REVIEW OF PROJECTIVE ASSESSMENT METHODS 439
- *4.15.7.1 Rorschach Inkblot Method* 439
  - *4.15.7.1.1 Administration and scoring* 439
  - *4.15.7.1.2 Psychometric foundations* 439
  - *4.15.7.1.3 Clinical utility* 440
- *4.15.7.2 Thematic Apperception Test* 440
- *4.15.7.3 Administration and scoring* 441
  - *4.15.7.3.1 Psychometric foundations* 441
  - *4.15.7.3.2 Clinical utility* 442
- *4.15.7.4 Children's Apperception Test* 442
  - *4.15.7.4.1 Administration and scoring* 443
  - *4.15.7.4.2 Psychometric foundations* 443
  - *4.15.7.4.3 Clinical utility* 443
- *4.15.7.5 Roberts Apperception Test for Children* 444
  - *4.15.7.5.1 Administration and scoring* 444
  - *4.15.7.5.2 Psychometric foundations* 444
  - *4.15.7.5.3 Clinical Utility* 445
- *4.15.7.6 Tell-me-a-story* 445
  - *4.15.7.6.1 Administration and scoring* 446
  - *4.15.7.6.2 Psychometric foundations* 446
  - *4.15.7.6.3 Clinical utility* 447
- *4.15.7.7 Draw-a-person* 447
  - *4.15.7.7.1 Administration and scoring* 448
  - *4.15.7.7.2 Psychometric foundations* 448
  - *4.15.7.7.3 Clinical utility* 449
- *4.15.7.8 House-tree-person* 449
  - *4.15.7.8.1 Administration and scoring* 450
  - *4.15.7.8.2 Psychometric foundations* 450
  - *4.15.7.8.3 Clinical utility* 450
- *4.15.7.9 Kinetic Family Drawing* 450
  - *4.15.7.9.1 Administration and scoring* 451

4.15.7.9.2 Psychometric foundations 451
4.15.7.9.3 Clinical utility 452
4.15.7.10 Sentence Completion Methods 452
4.15.7.10.1 Administration and scoring 452
4.15.7.10.2 Psychometric foundations 453
4.15.7.10.3 Clinical utility 453
4.15.8 FUTURE DIRECTIONS 454
4.15.9 SUMMARY 454
4.15.10 REFERENCES 455

4.15.1 INTRODUCTION

Projection, as first formulated by Freud (1962) and later elaborated in his presentation of the Schreber case (Freud, 1958), consists of attributing one's own characteristics to external objects or events without adequate justification or conscious awareness of doing so. Frank (1939) suggested that personality tests in which there is relatively little structure induce a subject to "project upon that plastic field . . . his private world of personal meanings and feelings" (pp. 395, 402). By linking the concept of projection to the response process in such measures as the Rorschach Inkblot Method and the Thematic Apperception Test (TAT), Frank gave birth to the so-called projective hypothesis in personality assessment. His observations about what he called "projection measures" led to the Rorschach, the TAT, and other assessment methods involving some ambiguity being routinely designated as projective tests.

This chapter begins with some observations concerning the nature of objectivity and subjectivity in personality assessment methods and the role of structure and ambiguity in different types of projective techniques. It then turns to the value and utility of projective assessment and the applicability of projective methods in clinical work with young people. Information is given on the composition, administration, scoring, psychometric foundations, and clinical utility of nine projective techniques widely used in assessing children and adolescents.

4.15.2 OBJECTIVITY AND SUBJECTIVITY IN PERSONALITY ASSESSMENT METHODS

As a legacy of Frank's projective hypothesis, tests designated as projective methods came to be regarded as subjective in nature and hence quite different from objective methods, such as self-report inventories. This presumed subjectivity of projective methods fostered a commonplace conviction that these methods are inherently less scientific and less valid than objective methods. In actuality, however, ambiguity is a dimensional rather than a categorical characteristic of tests, and there is little basis for regarding projective methods as inherently unscientific and invalid or as sharply distinct from objective measures.

Being scientific does not inhere in the nature of a method or instrument, whether subjective or not, but only in whether it can be studied scientifically. When projective tests are used to generate personality descriptions that can be independently and reliably assessed for their accuracy, they function as a scientific procedure.

Likewise, the validity of a test inheres not in its nature, but rather in the extent to which it generates significant correlations with personality characteristics or behaviors it can identify or predict. Abundant research attests that projective methods, when properly used, can yield valid inferences (Hibbard et al., 1994; Hibbard, Hilsenroth, Hibbard, & Nash, 1995; Parker, Hanson, & Hunsley, 1988; Weiner, 1996).

Regarding sharp distinctions between projective and objective measures, the subject's task on commonly used personality tests injects considerable objectivity into many projective methods and substantial subjectivity into most objective methods. For example, responses on the Rorschach, which is the most widely used projective measure, are routinely inquired by asking subjects "Where did you see it?," which is a concrete, unambiguous request for a specific item of objective information. When subjects reply that they used the whole blot for a percept, their response is given a location choice code of W, which is an objective and unambiguous procedure on which coders achieve virtually 100% agreement. Numerous other features of how subjects choose to look at the inkblots can also be objectively coded with good inter-rater reliability, such as the percentage of responses given to the multicolored cards (affective ratio), the number of commonly given percepts reported (populars), and the kinds of objects the blots are said to resemble (people, animals, etc.) (Exner, 1991, pp. 459–460; McDowell & Acklin, 1996; Seaton & Allen, 1996).

Elements of objectivity mark the interpretation as well as the coding of many Rorschach variables. In the case of whole responses, for example, an unusual preponderance of W in a record correlates with objectively observable tendencies to attend to experience in a global fashion; a low affective radio identifies an inclination to withdraw from emotionally charged situations; a small number of popular responses correlates with behavioral manifestations of unconventionality; and numerous human percepts is associated with an active interest in people. In these and many other ways, Rorschach responses identify personality characteristics through an objective process of coding response features and relating these coded features to their known corollaries in observable behavior.

There are aspects of Rorschach interpretation that may be highly subjective, especially when inferences are drawn from the thematic imagery subjects produce when they associate to the inkblots along with describing them. Moreover, most other projective methods have not been as extensively codified as the Rorschach and depend more on qualitative than quantitative analysis. The present point is merely that the basic nature of projective methods does not preclude their being codified and interpreted to some extent along objective lines.

Turning now to aspects of subjectivity in objective methods, consider the uncertainty that characterizes many items in the most widely used objective measure of personality, the adult and adolescent forms of the Minnesota multiphasic personality inventory (MMPI-2/MMPI-A). Although MMPI-2/MMPI-A instructions to respond true or false are unambiguous and the coding of these responses is completely objective, Weiner (1993) has previously called attention to the idiography that is embedded in asking subjects to interpret such items as "I often lose my temper." Items of this type provide no benchmarks for the frequency of "often" or for what constitutes loss of temper. In many instances, consequently, responses to self-report items involve subjectivity on the part of respondents, who must define for themselves what certain terms mean before they can decide how to answer.

Subjectivity influences the interpretation as well as the response process in objective assessment. Granted, the hallmark of objective tests is an extensive array of quantified scale scores having empirically demonstrated behavioral correlates. Nevertheless, the interpretation of self-report measures in clinical practice typically goes beyond identifying known corollaries of scale scores to include consideration of complex patterns of interaction among these scores and between the test profile and aspects of a subject's clinical history, interview behavior, and performance on other tests. Some of these complex interactions have been examined empirically, such as various two- and three-point codes on the MMPI-2/MMPI-A, but many have not.

This is not to say that the MMPI-2/MMPI-A and other self-report instruments are basically subjective in nature or that they derive their utility primarily from clinical judgment. The point is merely that, just as projective instruments are not entirely subjective, self-report methods are not completely objective, but instead involve some aspects of ambiguity in how subjects respond to them and how examiners interpret them.

It is for this reason that ambiguity is not a categorical function that characterizes some tests called projective measures, but not others called objective measures. Instead, ambiguity is a dimensional function that characterizes most tests to some degree, in relation to how structured they are. Generally speaking, objective tests are more structured than projective tests and therefore less ambiguous; projective tests are generally less structured than objective tests and hence more ambiguous; and there is no sharp objective/subjective dichotomy between relatively structured and relatively unstructured instruments.

4.15.3 STRUCTURE AND AMBIGUITY IN PROJECTIVE TECHNIQUES

Projective techniques comprise inkblot methods, story-telling methods, figure drawing methods, and sentence completion methods. In addition to being less structured than objective measures, these four types of projective methods differ from each other in their degree of ambiguity and in whether their ambiguity resides mainly in their stimuli or in their instructions.

Thus in the case of the Rorschach inkblot method, subjects are asked to look at relatively ambiguous stimuli but are given fairly specific instructions to indicate what they see, where they see it, and what makes it look as it does. Story-telling methods such as the TAT involve showing subjects real pictures that are much less ambiguous than inkblots; however, by using general instructions ("Tell me a story") and open-ended questions ("What will happen next?"), examiners provide only minimal guidance in how subjects should respond. If the Rorschach instructions were "Tell me a story about this inkblot," the Rorschach would be more ambiguous and more of a projective test than it is. If the TAT instructions were "Tell me what you see here," the TAT would lose most of its ambiguity and function only barely as a projective test.

Figure drawing techniques use no stimuli at all, save a blank piece of paper, and provide little guidance to subjects, other than some instructions concerning the figures to be drawn (e.g., yourself, a family). Sentence completion methods, in common with story-telling techniques, call for subjects to provide thematic content in response to real and relatively unambiguous test stimuli. Unlike story-telling techniques, however, sentence completion methods do not ordinarily involve querying subjects about their responses or encouraging them to elaborate those that are brief or unrevealing. Thus a stem of "I AM" may be completed with "a happy person," in which case some subjectivity has been allowed to enter the response, or simply with "here," in which case only a completely objective response has been given. On balance, figure drawing techniques are the most ambiguous of projective tests and sentence completion methods the least, with inkblot and story-telling techniques in between.

These differences in ambiguity among projective methods were originally noted by Stone and Dellis (1960), who proposed "a levels hypothesis" to take practical account of this variability. According to the levels hypothesis, the degree to which a test is structured is directly related to the level of conscious awareness at which it taps personality processes. The more structured and less ambiguous a test is, the more likely it is to yield information about relatively conscious and superficial levels of personality; conversely, the less structured and more ambiguous a test is, the more likely it is to provide information about deeper levels of personality and characteristics of which subjects themselves may not be consciously aware.

Research reported by Stone and Dellis (1960) and subsequently replicated by Murstein and Wolf (1970) provided empirical support for a relationship between the ambiguity of a test and its likelihood of measuring deeper levels of personality, especially in normally functioning persons. These findings mirrored the basic conception of TAT assessment articulated by Murray (1951), who regarded the virtue of the instrument as residing not in its revelations about what subjects are able and willing to say about themselves, but in what it conveys about personality characteristics: "the patient is unwilling to tell or is unable to tell because he is unconscious of them" (p. 577).

In addition to differing from objective tests and from each other in their degree of structure and ambiguity, individual projective measures typically include both relatively objective and relatively subjective elements. As elaborated by Weiner (1977), the objective elements of projective test data involve structural features of the manner in which responses are formulated, whereas the subjective elements consist of thematic features of the imagery with which responses are embellished.

When projective test data are being interpreted objectively, structural aspects of the subject's responses, such as focusing on wholes and seeing numerous human figures on the Rorschach, are taken as being directly representative of similar behavioral tendencies in the person's life, that is, attending to experience globally and paying close attention to people. When projective test data are being interpreted subjectively, thematic imagery is taken as being indirectly symbolic of a subject's underlying needs, attitudes, conflicts, and concerns. Thus the Rorschach response of "Two girls who are really mad at each other fighting over something they both want" may identify a subject's experiencing peer or sibling rivalry, viewing social interactions as aggressive confrontations in which people are only concerned with what they can get for themselves, or feeling angry or resentful about being in such situations.

On story-telling measures, an example of a structural response feature is giving long stories, which can be objectively scored (by counting the number of words) and which provides a representative indication of inclinations to be verbose. As for subjectively interpreted features, a TAT story in which two people are described as about to separate, leaving one of them sad and lonely for the rest of his or her life, exemplifies thematic imagery that appears to symbolize concerns about suffering the loss of love objects and facing an unhappy future.

On figure drawing measures, which as previously noted are the most ambiguous of projective tests, structural features of the data are limited. Some variables, such as the size of figures drawn, how complete they are, and whether they are clothed, are objective facts that can usually be coded with good agreement. However, interpretation of such objective characteristics of figure drawings, as well as of subjective impressions of drawing qualities, is based mostly on their being symbolic rather than representative of behavior. Interpreting the way figures are drawn or placed is thus primarily thematic. For example, unusual emphasis on a particular part of the body may be interpreted as suggesting concern about functions associated with that part of the body, and a family drawing in which the self is located on one side of the page and the other family members are closely grouped on the other side of the page may be interpreted as symbolizing feelings of isolation or rejection in the family setting.

In sentence completion responses, frequent self-referencing is an example of an objectively scorable, behaviorally representative structural index of tendencies to focus attention on oneself rather than others. Consider the difference between the completions "WHAT PAINS ME is seeing how many unfortunate people there are in the world" and "WHAT PAINS ME is not being able to get the things that I want." An accumulation of the latter as opposed to the former type of response is objectively representative of self-centeredness. At the same time, the thematic content of both completions suggests in a more subjective way certain underlying concerns, such as worries about the welfare of the human race in the first instances and feelings of being personally deprived in the second.

To bring these introductory observations full circle, the opportunities that projective methods create for subjects to project aspects of themselves into their responses has frequently led to their being associated with psychoanalytic theories of personality, in the context of which the notion of projection was first elaborated. However, there is no necessary relationship between psychoanalytic theory and projective testing, nor is there any reason for clinicians who conceptualize behavior in other ways to view projective methods as incompatible with their frame of reference. The basic principle underlying projective techniques is that something can be learned about people from sampling how they respond in ambiguous situations. This principle is not prisoner to any personality theory, and its utility transcends the theoretical persuasions of individual examiners. Inferences from projective data can be couched equally well in psychodynamic, behavioral, cognitive, and humanistic terms, and the use to which these inferences can be put depends less on theoretical differences in terminology than on the nature of the assessment issues being addressed.

4.15.4 VALUE OF PROJECTIVE ASSESSMENT

Projective test data provide valuable information about how people are likely to think, feel, and act that is difficult to obtain from objective assessment procedures. This contribution of projective methods to the personality assessment process has both a conceptual and an empirical basis.

4.15.4.1 Conceptual Basis

Because of their relatively unstructured nature, projective tests measure personality characteristics in subtle and indirect ways. Even those features of projective test data that can be objectively scored and interpreted involve responses that seldom have obvious meaning. Subjects in the process of responding usually have little awareness of the interpretive significance that attaches to their seeing numerous human figures on the Rorschach, giving long stories on the TAT, drawing themselves on the far side of the page from the rest of their family, or repetitively referring to "I" in their sentence completions; indeed, they may not even be aware of having responded in these ways.

By contrast, relatively structured objective tests measure personality characteristics in direct ways that often have obvious interpretive significance. Adolescents who answer "true" to such MMPI-A statements as "At times I feel like smashing things" and "I am easily downed in an argument" will usually have a good idea of what they are indicating about themselves.

The distinction between subtle, indirect measurement of personality characteristics with projective techniques and relatively direct assessment through questionnaire methods has been formulated by McClelland, Koestner, and Weinberger (1989) in terms of differences between self-attributed and implicit motives. According to McClelland et al., self-attributed motives are measured by self-report instruments and are influenced by social incentives in a person's external environment. Implicit motives, however, are measured by such indirect techniques as story-telling procedures and are influenced by the internal pleasure derived from various activities in which a person engages.

Whereas self-attributed motives are comparatively good predictors of immediate specific responses to structured situations, McClelland et al. continue, implicit motives are comparatively good predictors of long-term trends in behavior across various types of situations. Research findings described by McClelland et al. confirmed that indirect assessments of underlying motives have greater validity for predicting long-term trends in behavior than self-report assessments of motives that people directly attribute to themselves.

More recently, Bornstein (1995) has used a metaanalysis of 97 studies of measures of dependency to demonstrate further this difference between objective and projective assessment. With respect to differential prediction, according to Bornstein, available research indicates that objectively measured dependency correlates better with symptoms and the diagnosis of dependent personality disorder than does projectively measured dependency, whereas projectively measured dependency correlates better with dependency-related behaviors.

The conceptual analysis formulated by McClelland et al. and elaborated by Bornstein has direct bearing on the contribution of projective methods to personality assessment. As previously described, projective assessment taps implicit motives and underlying personality characteristics that may not be readily apparent and may not be within a subject's conscious awareness. Because these covert motives and characteristics exert a powerful influence on long-term behavioral trends, this type of indirect measurement adds an important dimension to personality evaluations that would not be tapped in its absence.

Finally, with respect to what projective methods contribute to assessment batteries, the relative ambiguity of these methods makes them less subjective than structured instruments to influence by test-taking attitudes. This is not to say that projective methods are immune to subjects' efforts to present themselves in a positive or negative light. The relatively open-ended nature of projective testing situations and the dialogue they frequently elicit give subjects abundant opportunity to voice attitudes toward the tests, the examiner, and being examined. However, as long as subjects continue to give responses, neither their attitudes nor their expression of them is likely to prevent their projective test responses from revealing their personality characteristics. Simply put, the limited face validity of projective measures makes them more difficult to fake than objective measures, which means that they can balance a test battery to particularly good effect when self-presentation effects are of concern.

4.15.4.2 Empirical Basis

Projective measures vary in the extent to which they have been examined in well-designed research studies, and in many instances adequate empirical support for these instruments has lagged behind the uses to which clinicians sometimes put them. Nevertheless, the two most frequently used projective methods, the Rorschach and the TAT, have for the past generation been among the three most frequently studied personality assessment instruments, exceeded in this respect only by the MMPI/MMPI-2 (Butcher & Rouse, 1996). For both the Rorschach and the TAT, substantial evidence has accumulated to attest their validity for describing aspects of personality structure and dynamics and applying these descriptions in differential diagnosis and treatment planning (Abraham, Lepisto, Lewis, & Schultz, 1994; Alvarado, 1994; Bornstein, 1995; Cramer & Blatt, 1990; Exner & Andronikoff-Sanglade, 1992; Ornduff & Kelsey, 1996; Ronan, Colavito, & Hammontree, 1993; Weiner, 1996; Weiner & Exner, 1991).

Later in this chapter, specific information is presented concerning the psychometric foundations and demonstrated corollaries of the projective measures most frequently used with young people. Suffice it to say in summary at this point that these measures prove valuable in personality assessment because they add information that would otherwise be unavailable and because they withstand relatively well efforts to exaggerate or conceal.

4.15.5 UTILITY OF PROJECTIVE ASSESSMENT

Whereas the value of projective techniques lies in the previously elaborated reasons *why* they should be used in personality assessment, their utility relates to decisions concerning *when* these methods should be included in a test battery. The kinds of information provided by projective test data indicate that projective measures should be used whenever a thorough personality assessment is considered relevant to formulating a differential psychodiagnosis or recommending alternative intervention strategies. Because of their relatively unstructured nature and indirect format, projective measures balance a test battery by tapping personality characteristics at a less conscious level than relatively structured measures. Assessments lacking such balance sample personality functioning from a limited perspective that will usually fail to paint a complete picture of the individual being examined. Batteries limited solely to projective techniques are similarly imbalanced and ill-advised in comprehensive personality assessments.

The previously mentioned conceptualization of McClelland et al. (1989) bears closely on the importance of a balanced test battery in clinical assessment. McClelland and his colleagues noted that measures of self-attributed and implicit motives seldom correlate with each other and should not be expected to do so, because they are measuring different aspects of personality. Moreover, given that directly and indirectly measured motives each predict certain kinds of behavior better than the other, they concluded that "Separate measures of self-attributed and implicit motives may be combined to yield a better understanding and prediction of certain types of behavior" (p. 692).

These formulations concerning different types of measures have subsequently been elaborated and confirmed for clinical purposes with respect to relationships between the Rorschach and the MMPI. Rorschach structural variables and MMPI scales have been found to show only a few modest correlations in both adult and adolescent samples (Archer & Krishnamurthy, 1993a, 1993b). At the same time, however, apparent contradictions between Rorschach and MMPI findings have been conceptualized by Weiner (1993, 1995b) not as invalidating either instrument or challenging the incremental utility of administering them in tandem, but rather as generative data.

Specifically, Weiner argues, apparently discrepant findings between personality assessment instruments of different kinds can be generative by virtue of complementing each other. Whereas findings on two tests that concur in suggesting the same personality characteristic are confirmatory and support definite conclusions, he continues, findings that diverge raise important questions to which they may also suggest helpful answers, especially if one of the tests is a relatively structured and the other a relatively unstructured instrument.

Consider, for example, a youngster who appears depressed on the Rorschach, with a high depression index (DEPI), but does not elevate on Scale 2 or the depression content scale of the MMPI-A. This divergent finding could well provide a useful clue to the adolescent's having an underlying or emerging depression that is not yet being keenly felt or manifest in well-structured situations. Alternatively, it could be that the subject is trying to deny or repress depressive affects and cognitions, or is making a conscious decision not to report manifestations of depression that nevertheless emerge in the absence of supportive structure or are revealed when the subject is uncertain how to conceal them.

The situation described in this example is familiar to assessment psychologists, who not infrequently work with psychologically troubled adolescents who can remain reasonably comfortable and controlled in relatively structured situations but become upset and disorganized in relatively unstructured situations and who may accordingly produce a benign MMPI-A protocol and a disturbed Rorschach. In such circumstances the objective measure has not erroneously overlooked psychopathology, nor has the projective measure mistakenly exaggerated it. Instead, the two types of test have combined in complementary fashion to provide valid information concerning the subject's likelihood of behaving in a relatively adaptive or maladaptive fashion, depending on the situational context in which the behavior appears.

Similarly, subjects in some circumstances may produce a clinically unremarkable Rorschach while showing numerous elevations on the clinical and content scales of the MMPI-A. Such divergence is best understood not as error variance, but as a possible clue to the psychological stance of subjects whose degree of disturbance is minimal but who, when asked about themselves in language they can understand, want to make sure that others fully appreciate whatever problems and concerns they do have. Further illustrations of the clinical utility of divergence as well as convergence between a projective measure, such as the Rorschach and an objective measure such as the MMPI-2, are provided by Finn (1996) and Ganellen (1996).

4.15.6 APPLICABILITY OF PROJECTIVE METHODS TO CHILDREN AND ADOLESCENTS

Except for an occasional example, this chapter has thus far made no specific reference to young people. This apparent oversight is warranted by the fact that the nature of projective methods, the way in which they function, the kinds of information they provide, and the reasons for using them, are identical for persons of almost all ages. Hence the discussion of projective assessment to this point is as applicable to children and adolescents as to adults, and requires no modification or qualification as our focus now shifts specifically to young people.

Indeed, assessors who have learned to interpret projective test data provided by adults do not need to learn any new ways of working with the data should they begin to examine children and adolescents. By and large, the basic interpretive conclusions and hypotheses that attach to projective test variables apply regardless of the age of the subject. Whether they are age 8, 18, or 80, subjects who see numerous human figures on the Rorschach are likely to be quite interested in people; those who give long TAT stories are likely to be verbose; those who refer frequently to themselves in sentence completions are likely to be self-centered; and those who draw grotesquely distorted human figures probably harbor some disturbing concerns about their own nature or that of other people.

However, in order to determine the implications of these and other personality characteristics suggested by projective test data, examiners assessing young people must take into account normative developmental expectations. For example, the data of developmental psychology indicate that children are more self-centered than adults, and subsequently become increasingly aware of and concerned about the needs of others as they grow through adolescence and approach maturity. Accordingly, test data that identify a high degree of self-centeredness may imply maladaptive narcissistic personality traits in an adult, but reflect normal development and adaptation in a child; conversely, minimal self-centeredness may indicate altruism and good adjustment in an adult but suggest deviant development and low self-esteem in a child.

Developmental psychology similarly provides some normative expectations for how children are likely to make drawings. Preschool age children commonly draw with what is called "intellectual realism," which means that they draw what they know to be there regardless of whether it would actually be visible. Thus, in x-ray fashion, young children often draw transparencies, such as people who are visible through walls (Di Leo, 1983). At about age 7 or 8, this intellectual realism gradually gives way to "visual realism," in which what is drawn resembles what realistically can be seen. Di Leo (1983, p. 38) observes that this developmental shift mirrors a metamorphosis in thinking from an egocentric to an increasingly objective view of the world. Hence a human figure drawing by a preschool child showing a belly button through clothing is much less likely to imply maladaptive functioning that the same drawing done by an adolescent.

As these examples indicate, familiarity with and adequate attention to normative expectation hold the key to valid and useful applications of projective methods in the assessment of young people. Ideally, projective methods manuals should include normative reference data that delineate quantitative as well as qualitative expectations for such developmental phenomena as maturational changes in self-centeredness. Regrettably, even though numerous projective test variables have been quantified in various ways, little progress has been made in generating age-graded norms for them.

The main exception to this dearth of normative developmental data for projective techniques is the Rorschach. Developmental trends in Rorschach responses from early childhood through adolescence were initially charted many years ago by Ames and her colleagues (Ames, Metraux, Rodell, & Walker, 1974; Ames Metraux, & Walker, 1971). More recently, the Rorschach comprehensive system has provided reference data for each of its codified variables on samples of approximately 100 nonpatient young people at each age from 5 to 16 (Exner & Weiner, 1995, chap. 3).

Aside from identifying needs for further research, an analysis of available data can guide clinicians in choosing which projective methods to include in a battery for assessing a young person's functioning. The more thorough and reliable the normative developmental data available for the instrument, the better the choice it will make. Similarly, the better established an instrument's correlates are in relation to behaviors that are central to the purpose of an assessment, the more reason there is to include it. Thus an instrument that has been demonstrated to be particularly helpful in identifying youthful depression may be a good choice in one case, whereas an instrument known to be especially sensitive in revealing family dynamics may be a good choice in another case.

Similarly, available empirical data and reported clinical experience should be drawn on to determine whether a particular instrument is likely to yield useful information concerning the personality functioning of individuals at certain ages. Thus the children's apperception test (CAT) depicting animal figures may be a more effective story-telling measure for a young child than the TAT, but certainly not for an adolescent (Bellak, 1993, p. 237).

As these observations indicate, projective methods provide sound clinical data only if they are employed in appropriate ways. First, examiners should have recourse to standardized procedures for administering and scoring any test they use. Lack of such standard methodology compromises the value of the data obtained, and inattention to standardized methods by examiners who opt instead for personalized approaches to administration and scoring is clinically disadvantageous and professionally questionable.

Second, clinical interpretations should be derived from test variables with demonstrated reliability and validity. Inadequate psychometric foundations limit the use to which test data can be put, and examiners who draw conclusions in the absence of supporting empirical evidence, without framing such conclusions as speculative hypotheses, are doing their patients and their methods a disservice. Third, the adequacy of projective assessment of young people will be limited in the absence of normative reference data for test responses of both adjusted and maladjusted children and adolescents and for developmental changes in these responses over time.

This chapter continues with reviews of the major inkblot, story telling, figure drawing, and sentence completion methods used in assessing young people. The composition, administration, and scoring of each of these measures are described; what is known about their reliability, validity, and normative database is reported; and the clinical purposes they are likely to serve are discussed. The specific measures reviewed are selected primarily on the basis of their emphasis and frequency of use in clinical and school settings, as reported in surveys by Archer, Imhof, Maruish, and Piotrowski (1991), Elbert and Holden (1987), Hutton, Dubes, and Muir (1992), Kennedy, Faust, Willis, and Piotrowski (1994), Piotrowski and Keller (1989), Stinnett, Havey, and Oehler-Stinnett (1994), and Watkins, Campbell, Nieberding, and Hallmark (1995).

4.15.7 REVIEW OF PROJECTIVE ASSESSMENT METHODS

4.15.7.1 Rorschach Inkblot Method

The Rorschach inkblot method comprises 10 cards that are inked in shades of black and gray (five cards); black, gray, and red (two cards); and various pastel colors (three cards). The cards are reproduced in standard fashion, but the inkblot stimuli were originally designed at random and do not portray any specific objects (Rorschach, 1942). When subjects respond to the Rorschach, they draw on the shape, shading, and color of the blots to form impressions of what they might be, and in so doing they treat the instrument as a cognitive–perceptual task (e.g., "It looks like a bat, because it's got a body here and wings here and it's black"). In addition, subjects frequently elaborate their responses beyond the stimulus properties of the blots, and in so doing they treat the instrument as an associational task (e.g., "This bird is flying around looking for something to eat").

The cognitive–perceptual aspects of responses constitute structural data in Rorschach assessment and provide representative indications of the resources and coping style that a person generally brings to bear in problem-solving situations. The associational aspects of responses constitute thematic data in Rorschach assessment and provide symbolic clues to the underlying needs, attitudes, conflicts, and concerns that are likely to influence a person's actions and state of mind. The basic nature of the Rorschach in these respects is discussed further by Exner and Weiner (1995, chap. 1) and Weiner (1986, 1994).

4.15.7.1.1 Administration and scoring

The Rorschach is introduced by telling subjects that the inkblots they are about to see are not anything in particular, but that people see many different things in them and that their task will be to indicate what the inkblots look like to them. The 10 cards are then given to subjects one at a time with the instruction "What might this be?" Requests for structure (e.g., "Can I turn the card?") are deflected back to the subject (e.g., "It's up to you"; "Any way you like"). The unguided responses to the 10 cards constitute the free association phase of the administration, following which there is an inquiry phase in which the examiner reads back each response and asks subjects where they saw it and what made it look as it did. The purpose of the inquiry is to facilitate coding of the structural features of the responses, and associations during this phase are not requested or encouraged. Responses are recorded verbatim, however, and the content of any spontaneous thematic elaborations is carefully noted.

Numerous approaches to codifying Rorschach responses have emerged during the long history of this instrument. However, for many years the comprehensive system of Exner (1993) has been by far the most widely used and researched (Piotrowski, 1996). Rorschach responses are coded in the comprehensive system for various aspects of where percepts are seen (location), why they look as they do (determinants), what they consist of (content), how commonly they occur (form level and populars), and whether they involve pairs of objects, organization of parts, or special kinds of elaborations, such as cooperative or aggressive interaction. These codes are than tallied and combined in various ways to yield a large number of indices, ratios, and percentages that guide the interpretive process, as elaborated in detail by Exner (1991, chaps 5–10).

4.15.7.1.2 Psychometric foundations

The psychometric foundations of an assessment instrument comprise the extent to which it can demonstrate adequate interscorer agreement, reliable measurement, valid correlates, and a representative normative database. The Rorschach inkblot method, as already indicated in part by examples used earlier in the chapter, rests on a solid psychometric basis. Interscorer agreement for the types of variables coded in the comprehensive system typically ranges from 80% to 100%. The reliability of Rorschach data has been demonstrated in a series of retest studies conducted over intervals ranging from seven days to three years and involving child, adolescent, and adult subjects. Most of the core variables associated with trait dimensions of personality show stability coefficients greater than 0.80 in these studies, and some, including the affective ratio and the egocentricity index, consistently hover around 0.90 (Exner, 1991, pp. 459–460; Exner & Weiner, 1995, pp. 21–27; McDowell & Acklin, 1996; Weiner, 1997).

The validity of Rorschach assessment was confirmed in a series of metaanalytic studies that led Parker, Hanson, and Hunsley (1988) to conclude that the Rorschach meets usual psychometric standards for validity and is comparable to the MMPI in this respect. Specifically, Parker et al. used the effect sizes reported in 411 studies to derive population estimates of convergent validity of 0.41 for the Rorschach and 0.46 for the MMPI. Subsequent further confirmations of the validity of this instrument are noted by Weiner (1996).

With respect to its normative database, available information for the comprehensive system includes data on 700 nonpatient adults demographically representative of the 1980 US census, 1390 nonpatient children and adolescents age 5 to 16, and large samples of schizophrenic, depressed, and character disordered patients (Exner, 1993, chap. 12). In addition, longitudinal data reported by Exner, Thomas, and Mason (1985) on a group of young people tested every two years from age 8 to 16 provide useful reference information concerning developmental stability and change in Rorschach variables during childhood and adolescence.

As implied by the nature of the normative data, the Rorschach comprehensive system is applicable to young people from age five. Preschool age children have ordinarily not yet matured sufficiently to deal with the cognitive–perceptual aspects of the Rorschach situation in ways that lend themselves to the codification that is central to the comprehensive system interpretive process. An unusually mature four-year-old might on occasion produce a useful record, and immature five- and six-year-olds may produce records that have limited interpretive significance within the framework of the comprehensive system. Working within other frameworks, Ames et al. (1974) discuss and provide some normative findings for Rorschach responses of young children, and Leichtman (1996) has recently presented a developmental rationale for deriving information from the records of preschoolers.

4.15.7.1.3 Clinical utility

In common with projective techniques in general, the Rorschach serves clinical purposes primarily as a result of the information it provides about the structure and dynamics of personality functioning. With respect to personality structure, the Rorschach has proved especially helpful in identifying and quantifying states of subjectively felt distress that combine elements of anxiety and depression and in reflecting trait dimensions of how people typically think, process information, handle emotions, manage stress, feel about themselves, and relate to others. Regarding personality dynamics, the thematic content of Rorschach responses, as previously noted, is often quite revealing of underlying needs, attitudes, conflicts, and concerns that influence how people are likely to think, feel, and act at particular points in time and in particular situations.

In addition, Rorschach data can frequently contribute to differential diagnosis in clinical settings. The comprehensive system provides indices for schizophrenia and depression (DEPI) that can help to identify these conditions in children and adolescents as well as in adults; for basic deficits in coping capacity that point to developmental arrest in young people; and for numerous features of conduct and anxiety and/or withdrawal disorders (Exner & Weiner, 1995, chaps 5–8; Weiner, 1986). Rorschach findings have also demonstrated considerable clinical utility in the treatment process by clarifying treatment targets, identifying potential obstacles to progress in therapy, and providing a basis for evaluating treatment change and outcome (Abraham et al., 1994; Weiner, 1994).

4.15.7.2 Thematic Apperception Test

The most widely known and used story telling technique is the TAT. It was developed by Morgan and Murray (1935) in the belief that the content of imagined stories would provide clues to the underlying dynamics of a subject's interpersonal relationships and self-attitudes. As elaborated by Murray (1943, 1971) and Bellak (1993, chap. 4), TAT data are expected to reveal the hierarchy of a person's needs and the nature of his or her dominant emotions and conflicts.

The TAT stimuli comprise 19 black-and-white illustrations of people or scenes and one blank card. The cards are intended for use with persons age five or older of both genders, and for nine of the cards there are alternate versions for use with adult and child/adolescent males and with adult and child/adolescent females. Because of the time required to administer the full set of TAT cards, examiners typically select a subset of 8–12 cards that they anticipate will elicit themes relevant to the assessment issues in a particular case. The themes usually elicited by the individual cards and the selection of subsets suited for children and adolescents are reviewed by Bellak (1993, chap. 3), Dana (1985), and Obrzut and Boliek (1986). Regrettably with respect to standardization, however, there are no specific short forms of the instrument, and how many and which cards are typically chosen vary from one examiner to another and from one examination to the next.

4.15.7.3 Administration and scoring

The TAT cards are given to subjects one at a time with instructions to make up a story for each picture that includes (i) what is happening at the moment, (ii) what the characters are thinking and feeling, (iii) what led up to the situation, and (iv) what the outcome will be. The narrated stories are recorded verbatim by the examiner.

Murray (1943) originally proposed a scoring scheme in which each TAT story is rated for the presence and strength of a long list of needs that are being experienced by the central figure in the story and presses that are being exerted by the environment. This scoring system proved too elaborate and time consuming for clinical work, and numerous alternative approaches to clinical interpretation were subsequently developed for the instrument. As reviewed by Chandler (1990), Murstein (1963), and Vane (1981), some of these interpretive approaches, like Murray's, have consisted of formal quantitative ratings of story characteristics. However, most interpretive approaches have eschewed quantitative scoring in favor of qualitative analyses of story content, and no scoring system has gained widespread use either clinically or in research studies.

The most commonly employed methods of interpreting the TAT in clinical practice appear to be variations of an "inspection technique" proposed by Bellak (1993, chap. 4). This technique consists simply of reading through subjects' stories to identify repetitive themes and recurring elements that appear to fall together in meaningful ways. Because this approach lacks any quantification and rests on the capacity of individual examiners to relate story themes and elements to aspects of personality functioning, Dana (1985) was moved to observe that "TAT interpretation has become a clinical art form" (p. 90).

Bellak's influential approach stresses 10 aspects of a story, each of which is taken to have implications for how subjects view and are likely to deal with interpersonal events and what they anticipate the future to hold for them. These include the main theme of the story, the identity of the central figure or hero, the main needs of the hero, the way in which the environment is conceived, the identity and intentions of other figures in the story, the nature of any anxiety or other affect that is being experienced, the nature of any conflict that is described or suggested, the ways in which conflicts and fears are defended against, the ways in which misbehavior is punished, and the level of ego integration.

With respect to research studies, the most productive utilization of the TAT has derived from quantitative scoring systems developed by McClelland, Atkinson, and their colleagues to measure needs for achievement, affiliation, and power (Atkinson & Feather, 1966; McClelland, Atkinson, Clark, & Lowell, 1953). Although scoring for achievement, affiliation, and power motivation has had little clinical impact, other schemes for coding specific personality characteristics reflected in TAT thematic content have subsequently emerged. These include scales for level of ego development (Sutton & Swenson, 1983), preferred defense mechanisms (Cramer, 1987), quality of interpersonal affect (Thomas & Dudek, 1985), problem-solving style (Ronan et al., 1993), and object relations capacities (Westen, Lohr, Silk, Kerber, & Goodrich, 1985).

Particularly promising among these is the use of the TAT to assess aspects of object relatedness through the Westen et al. (1985) measure, known as the social cognition and object relations scale (SCORS). The SCORS provides quantitative indices of the affective tone subjects ascribe to relationships, their capacity for emotional investment in relationships and social standards, their understanding of social causality, and the complexity of their representations of people. By including ratings of subjects along dimensions of maturity as well as normality/pathology, the SCORS is proving especially relevant to the assessment of young people (Westen et al., 1991).

4.15.7.3.1 Psychometric foundations

Efforts to demonstrate the reliability and validity of global approaches to interpreting the TAT have been handicapped by the previously noted proliferation of scoring systems, by clinicians' preferences for a strictly qualitative and uncoded approach to the data, and by enormous variation in how the test is administered, including which subset of cards is selected for use. Because of this long-standing lack of standardization, there has been little opportunity for systematic accumulation of data bearing on the reliability and validity of the TAT in general, nor has it been possible to develop any substantial normative database.

Accordingly, for both inspection techniques and overall scoring systems developed in the tradition of Murray, the psychometric literature on the TAT is generally acknowledged to comprise a mix of positive and negative findings that cannot easily be compared with one another. Hence, despite the widespread use of inspection techniques in clinical practice, neither these nor other global approaches have been demonstrated to show adequate psychometric properties.

However, research studies with TAT scales developed to measure specific personality characteristics have demonstrated that the instrument can generate reliable and valid findings when it is used in a standardized manner. The previously mentioned scales of Cramer (1987) and Westen et al. (1985) are cases in point. Cramer's scale reliably identifies three major mechanisms of defense—denial, projection, and identification—and has shown valid corollaries in changes observed in patients undergoing psychotherapy (Cramer & Blatt, 1990). The Westen et al. SCORS has been found to provide reliable identification of developmental variables related to disturbed object relations in children and, as already mentioned, is therefore especially relevant to the assessment of young people. Validation studies with SCORS have involved psychiatrically disturbed, borderline, physically abused, and sexually abused young people (Freedenfeld, Ornduff, & Kelsey, 1995; Ornduff, Freedenfeld, Kelsey, & Critelli, 1994; Westen, Ludolph, Block, Wixom, & Wiss, 1990; Westen, Ludolph, Lerner, Ruffins, & Wiss, 1990).

4.15.7.3.2 Clinical utility

The clinical utility of the TAT lies mainly in its potential for elucidating dynamic aspects of personality functioning, particularly with respect to the feelings and attitudes that subjects hold towards other people, themselves, and possible turns of fortune in their lives for better or worse. Based on the assumption that children and adolescents identify with the central figures in their TAT stories and project fantasies and realities regarding their own lives into the events and circumstances they describe, the obtained data can shed light on a broad range of underlying influences on how young people are likely to think, feel, and act.

As previously noted in commenting on the research of McClelland et al. (1989), the implicit types of motives measured by the TAT are more likely to correlate with persistent dispositions to behave in certain ways rather than with immediate actions or symptom formation. Accordingly, TAT findings will usually not add very much to structural diagnosis of adjustment problems in young people, but they can be extremely helpful in suggesting possible dynamic origins of adjustment problems. In this regard, the psychometrically sound SCORS may be a useful scale to include in forensic assessment batteries when issues of custody or adoption are being addressed. This TAT scale can frequently assist examiners in grasping a young person's representations of people and his or her capacities for emotional investment in relationships.

The thematic content of TAT stories has additional potential to facilitate planning and conducting psychotherapy with young people, particularly with respect to identifying treatment targets and monitoring progress in therapy. The TAT can also be used in treatment as a play therapy tool, as in Gardner's (1971) story telling technique. For example, after a youngster has told TAT stories, the therapist and child can act out the stories in play, or the therapist can create stories to the same picture stimuli for comparisons with the child's stories. Hoffman and Kupperman (1990) describe such an intervention with a 13-year-old boy in which both therapists wrote stories to the same TAT cards to which the patient had responded. As it turned out, Hoffman's stories emphasized the main character's maladaptive coping mechanisms, whereas Kupperman's stories emphasized positive and healthy aspects of the central character's coping capacities. Over a number of sessions, this boy and his therapist engaged in discussions concerning whose version of the story was most accurate.

4.15.7.4 Children's Apperception Test

Consistent with the purpose of the TAT, Bellak (1993, chap. 13) developed the CAT to facilitate understanding of personality processes in children, including their "dynamic way of reacting to and handling the problems of growth" (Bellak & Siegel, 1989, p. 102). The CAT pictures were designed to elicit fantasies about aggression, sibling rivalry, fears of being alone at night, attitudes toward parental figures, and eating problems.

The CAT-Animal (CAT-A) form, originally published in 1949 and designed for children 3–10 years old, consists of 10 pictures depicting animals in human situations. The use of animal figures was based on the assumption that young children identify more readily with animals than with people and will accordingly tell more meaningful stories about animal than human figures. Moreover, according to Bellak (1993, chap. 13), the use of animal figures makes the CAT-A a culture-free test that is equally applicable to Caucasian, African-American, and other minority group youngsters as well as to children from different countries, except where there is little familiarity with some of the inanimate objects depicted, such as bicycles.

There is also a human form of the CAT (the CAT-H) that was developed by Bellak and Hurvich (1966) in response to criticism of the assumption that children identify more easily with animal than with human figures. Studies reviewed by Bellak and Hurvich indicate little difference in stimulus value between the original CAT-A and the CAT-H, in which human figures are substituted for the animals in the CAT-A scenes. However, the CAT-H does not appear ever to have become much used in clinical practice.

4.15.7.4.1 Administration and scoring

Children being administered the CAT are told that they are going to take part in a game in which they will tell stories about pictures. Subjects who appear to regard the CAT as a test are informed that it is not the type of test in which they will be graded for correct or incorrect answers. Standard procedures call for all 10 CAT pictures to be administered in numerical order, from Card 1 to Card 10. Children are told to narrate what the animals are doing in the pictured scenes and are asked at appropriate points to say what went on previously and what will happen next. The examiner encourages and prompts the subject as necessary but avoids being suggestive or asking leading questions. Examiners may also query each story by asking the child to elaborate specific points such as the ages of characters and why they were given particular names.

In clinical work the CAT is typically interpreted along the lines proposed by Bellak for the TAT, that is, with an inspection technique used to form qualitative impressions of various dimensions of the subject's personality functioning (Bellak, 1993, chap. 14). There is an alternative but rarely used quantitative approach developed by Haworth (1965), called the schedule of adaptive mechanisms, in which CAT responses are rated numerically for the degree of adaptability or disturbance they reflect. Haworth also stressed the importance of recognizing that young children are highly reactive to the immediate circumstances in their lives and less likely than adolescents or adults to have formed well-established personality traits. She accordingly emphasized careful interpretation of CAT responses in the context of adequate information concerning subjects' home situation and the nature of any recent or impending crises in their lives.

4.15.7.4.2 Psychometric foundations

There has regrettably been little accumulation of empirical data bearing on the reliability and validity of the CAT. The widespread use of Bellak's qualitative inspection technique in CAT interpretation and a corresponding lack of quantification have precluded examination of the instrument's psychometric foundations. Although it has sometimes been suggested that the idiographic nature of CAT as well as TAT data makes traditional psychometric criteria difficult to apply or even irrelevant, there is nothing in the nature of the data generated by story telling techniques that prevents their being reliably coded for various types of feelings, motives, attitudes, and capacities that can in turn be validated against meaningful correlates. The previously noted development of psychometrically sound TAT scales for such specific aspects of personality as achievement motivation and social cognition proves the point that clinical interpretation of stories can go beyond being an art form and attain respectability as a scientific procedure as well.

Regarding what research is available concerning the CAT, Bellak (1993, chap. 16) provides a review of studies comparing the responses typical of children at different ages and examining special features in the stories of maladjusted, schizophrenic, speech disordered, retarded, brain-damaged, and chronically ill children. Almost all of these studies date from the 1950s, however, and none provides an adequate basis for developing any formal normative standards or diagnostic guidelines for the instrument.

4.15.7.4.3 Clinical utility

As in the case of the TAT, the CAT is useful in clinical assessment primarily as a source of hypotheses concerning subjects' personality dynamics, particularly with respect to how they view themselves and other important people in their lives, the nature of their hopes and fears, and what they expect will happen to them. Although not suitable for adolescents, the CAT is more useful than the TAT in work with younger children, who are likely to relate more easily to the familiar situations and youthful figures depicted in the CAT illustrations than to the primarily adult figures and unpopulated scenes shown in the TAT.

Like the TAT, the CAT may also contribute to treatment planning, by suggesting areas of concern on which to focus in the therapy, and it may itself serve as a play technique. In diagnostic assessments, however, hypotheses generated by this instrument require support from other data prior to being addressed to questions that necessitate empirical decision making.

4.15.7.5 Roberts Apperception Test for Children

The Roberts apperception test for children (RATC) is intended for use with young people of ages 6–15 and was designed to improve on the TAT and CAT by presenting familiar stimuli and employing a standardized scoring system (McArthur & Roberts, 1990).

Instead of illustrations primarily of adults, such as those used in the TAT, or illustrations of animals, as used in the CAT, the RATC primarily portrays child and adolescent figures engaged in everyday interactions, including scenes of parental affection, disagreement, school and peer relationships, and observation of nudity. There are 27 RATC cards, 11 of which are alternate versions for male or female subjects. There is an addition an alternate set of cards portraying African-American individuals in similar scenes.

4.15.7.5.1 Administration and scoring

The standard 16 cards that compose the RATC are administered individually in numerical order, using male or female versions as appropriate. Subjects are instructed to make up a story about each picture and to tell what is happening in the picture, what led up to the scene, how the story ends, and what the people are talking about and feeling. Responses are recorded verbatim.

If subjects tell an incomplete story or omit certain aspects, such as how the characters are feeling, additional inquiry may be used to help teach them to give scorable responses. This inquiry may be used liberally with the first two cards but not thereafter, which means that responses on cards 3 through 16 may at times have limited scorable data. The possibility of limited scorable data is the price to be paid for maintaining careful standardization of the RATC procedure. McArthur and Roberts sanction deviations from the standard procedures only in specific instances, such as the examination of severely disturbed children. When standard procedures are not or cannot be followed, they recommend considerable caution in using the scoring procedures.

The scoring system for the RATC consists mainly of coding the content of each story for the presence or absence of 13 profile scales that have implications for personality characteristics. Eight of these are adaptive scales that relate to thematic indications of reliance on others, giving support to others, supporting oneself, limit-setting by authority figures, identification of problem situations, and resolving problems in unrealistic, constructive, or particularly insightful ways. The other five are clinical scales that pertain to thematic manifestations of anxiety, aggression, depression, experiences of rejection, and inability to resolve problems. The test manual provides guidelines and examples that promote reliable scoring of each of these scales.

To facilitate interpretation of the data, the raw scores for each of the 13 profile dimensions are summed over the 16 cards and then plotted on a normatively scaled profile form to yield a visual representation of the data, much in the manner of an MMPI-A profile. The interpretive yield of the data is further enriched by use of an interpersonal matrix, which consists of a tabular representation of the frequency of convergence between particular scales and the figures identified (e.g., cooccurrence of themes of reliance on others with description of interaction with a maternal figure).

4.15.7.5.2 Psychometric foundations

As with the careful coding of specific variables on the TAT, the development of standardized administration and scoring procedures for the RATC has demonstrated that story telling projective methods can achieve psychometric respectability. McArthur and Roberts (1990) report inter-rater agreement ranging from 0.80 to 0.93 in the scoring of the various dimensions in their system and split-half reliabilities ranging from 0.44 to 0.86 on their profile scales, with half of these scales showing reliability coefficients of 0.73 or higher. The adaptive scales appear to work better in this regard than the clinical scales: only one of the five clinical scales (inability to resolve problems) shows a split-half reliability greater than 0.55, whereas all but one of the eight adaptive scales (reliance on others) shows a split-half reliability above 0.60.

Validity studies conducted with 200 nonpatient and 200 outpatient youngsters aged 6–15 indicate that subjects generally tell stories that fit the expectations for the cards for example, that a card intended to depict parental affection typically elicits stories involving affection. In addition to thus showing content validity, the RATC profiles have been found to distinguish between well-adjusted and clinic youngsters, to be more resistant to efforts at manipulation than self-report measures, and to correlate well with a behavior problem checklist completed by subjects' parents (McArthur & Roberts, 1990; Worchel, Rae, Olson, & Crowley, 1992). Significantly, however, in the comparisons between well-adjusted and clinic youngsters, significant differences were found on all eight of the adaptive scales and on two of the five clinical scales, but not on the clinical scales for anxiety, aggression, and depression.

With respect to its normative database, the RATC manual provides the mean and standard deviation for each of the profile scales for 200 nonpatient youngsters divided by gender and by four age categories (6–7, 8–9, 10–12, and 13–15).

4.15.7.5.3 Clinical utility

Similar to the TAT, CAT, and other storytelling measures, the RATC contributes to personality assessment of young people primarily by casting light on their underlying attitudes and concerns. Although the measure includes specifically designated clinical scales that might be expected to assist in clinical diagnosis as well as dynamic analysis of a youngster's personality functioning, its clinical scales as noted appear less reliable than its adaptive scales and less capable of discriminating between patient and nonpatient populations.

The RATC is especially valuable to clinicians as well as researchers by virtue of its careful standardization and adequate psychometric properties. Aside from facilitating the interpretive process and providing the basis for systematic accumulation of data, these features of the instrument make it particularly attractive to examiners conducting forensic evaluations. Having used the RATC in their projective assessment of a child or adolescent, rather than the TAT or CAT, psychologists will be better prepared to justify their procedures and conclusions when giving testimony as an expert witness.

However, examiners may sometimes question whether the information they get from the RATC warrants the amount of time required to administer it, which in our experience approximates an hour on the average. For some young children, moreover, there may be difficulties in sustaining their investment in the task. There is no short form of the RATC to use and no shortcut, such as selecting just a few cards to use or discontinuing a protracted administration before giving all 16 cards. Doing so eliminates the standardization of the instrument and prevents any meaningful scoring or comparisons with normative data, which means that the test loses its advantage over a TAT or CAT interpreted by inspection. Our view in this matter is that the RATC warrants the time required to administer and score it in the standardized manner and, more generally, that psychologists should if at all possible take whatever time is necessary to use standardized instruments when they are available.

The alternate form of the RATC involving African-American figures may prove useful in assessing African-American young people. However, there are as yet no data establishing the reliability and validity of this alternate form, nor have any representative normative data been published for it. Hence the RATC should be used cautiously in multicultural assessments, for which at present the most promising instrument is Tell-me-a-story.

4.15.7.6 Tell-me-a-story

The Tell-me-a-story (TEMAS) was designed as a multicultural story-telling test based on the concept that personality development occurs within a sociocultural system in which individuals internalize the cultural values of their family and society (Costantino, Malgady, & Rogler, 1988). The TEMAS is intended for use with African-American, Hispanic, and Caucasian children and adolescents aged 5–18 and comes in two parallel sets, one for minority and the other for nonminority youngsters. The two sets feature either predominantly Hispanic and African-American characters or predominantly nonminority characters all shown in urban environments. Each set comprises 23 cards, 11 of which have alternate versions for males and females and one of which has alternate versions for children and adolescents.

As a distinctive feature of the TEMAS, many of the pictures portray a split scene showing contrasting or conflicting intrapersonal and interpersonal situations that require some resolution, much in the manner of Kohlberg's (1976) moral dilemma stories. For example, one side of a scene may depict apparent delay of gratification, and the other side an inability to delay gratification. How subjects resolve this conflict in their stories speaks to the adaptiveness of their personality functioning and their stage of moral development.

The TEMAS provides quantitative scales for measuring the adequacy of subjects' adaptation with respect to nine aspects of personality functioning: interpersonal relations, aggression, anxiety/depression, achievement motivation, delay of gratification, self-concept, sexual identity, moral judgment, and reality testing. There are also quantitative scales for four cognitive functions related to how individual process information (reaction time, total time, fluency as reflected in the number of words used in a story, and omissions of relevant visual details) and for four affective functions as indicated by mood states attributed to the main characters in a story (happy, sad, angry, and fearful). These quantitative scales are supplemented by several qualitative indicators used to describe various other characteristics of the stories.

As stressed by Costantino, Malgady, and Rogler (1988), the TEMAS was developed to overcome limitations of traditional thematic apperception tests and differs from the TAT in several significant ways. These include a focus on interpersonal relationships rather than on intrapsychic dynamics; the use of personally relevant and culturally sensitive stimuli that emphasize meaning rather than ambiguity; the representation of both positive and negative poles of emotions, cognitions, and interpersonal functions, as opposed to the heavy weighting of the TAT stimuli with representations of depression, gloom, anger, and hostility; and the introduction of the joint depiction of contrasting circumstances to elicit expressions of conflict resolution and moral judgment.

4.15.7.6.1 Administration and scoring

Subjects are administered either the minority or nonminority form of the TEMAS, and examiners can choose between a long form, which comprises all 23 cards and requires approximately two hours to administer, and a standard short form, which consists of nine cards and requires one hour to give. In keeping with generally recommended practice in multicultural assessment (Dana, 1993, chap. 6, 1996), young people should be tested in their primary language, and those who are bilingual should be tested by a similarly bilingual examiner. Subjects are instructed to tell a complete story about each picture that indicates what is happening now, what happened before, and what will happen in the future. These instructions may be repeated as often as necessary and supplemented with structured inquiries to elicit information concerning who and where the characters are and what they are thinking and feeling.

The TEMAS is scored by rating stories for the presence of the various cognitive and affective functions and on a four-point scale for the level of adjustment reflected in thematic indications of the personality functions. The resulting scores are then totaled and translated into normalized T scores that are graphed to provide a readily interpretable visual profile. The TEMAS manual provides detailed guidelines and case examples to illustrate these scoring procedures.

4.15.7.6.2 Psychometric foundations

As described in the test manual (Costantino, Malgady, & Rogler, 1988), the TEMAS was carefully developed and standardized over several years prior to its publication. Unfortunately, except for continued work by the test's authors, much of it in the form of paper presentations, there has been little published research concerning the psychometric adequacy of the instrument. On balance, however, the data provided in the test manual appear to indicate a promising beginning in demonstrating its soundness.

With respect to inter-rater agreement, studies with Hispanic, African-American, and Caucasian subjects have indicated that both the minority and nonminority versions of TEMAS can be scored reliably, with agreements between trained examiners generally ranging from 75% to 95% across the various scales. Regarding issues of reliability, on the long form of the TEMAS 11 of the 17 personality, cognitive, and affective scales showed internal consistency (alpha) correlations of 0.74 or higher in the standardization data, and the median for all 17 was 0.83. Internal consistency was lower on the short form, with a median of 0.68. The internal consistency data for the short form must be considered preliminary, however, because in this analysis the short form scores were extracted from the protocols of subjects who had in fact completed the long form, rather than from actual administration of the short form. In an assessment of its retest reliability, the short form was administered twice to 51 behavior problem children over an 18 week interval. Very little stability over time was demonstrated in this study. Costantino, Malgady, and Rogler (1988) suggest several plausible explanations for this disappointing result, such as the narrow range of scores among the subjects. Nevertheless, test–retest reliability remains to be demonstrated for both the short and long forms of the instrument.

Turning to the validity of the TEMAS, ratings by psychologists of the types of personality functions pulled by the stimulus pictures have indicated good agreement with the intent in designing them, thus attesting that the test measures what it purports to measure. In terms of its criterion validity, the TEMAS has been found in several studies to discriminate between patient and nonpatient youngsters in Hispanic and African-American as well as Caucasian samples, and to accomplish this distinction in inner city as well as middle-class settings (Costantino, Malgady, Bailey, & Colon-Malgady, 1989; Costantino, Malgady, Colon-Malgady, & Bailey, 1992; Costantino, Malgady, Rogler, & Tsui, 1988). The TEMAS has not demonstrated any capacity to differentiate specific kinds of disorders, but there is some evidence to suggest that certain story characteristics, including the omission of main characters or events and failure to notice a conflict in the picture (due to lack of attention), may be sensitive to the presence of attention deficit hyperactivity disorder (Costantino, Colon-Malgady, Malgady, & Perez, 1991).

Initial studies reported in the TEMAS manual indicate further that many of its scales correlate significantly with behavior ratings by subjects' mothers and teachers and with behavioral observations of their inclinations toward aggressive or disruptive behavior and their capacities for self-confidence and delay of gratification. Preliminary data indicate further that TEMAS profiles of young people prior to their entering therapy can significantly predict aspects of their treatment outcome.

The TEMAS standardization sample comprised 281 male and 361 female youngsters from the New York City area who ranged in age from 5 to 13 years. This sample of 642 children and adolescents included groups of Caucasian, African-American, Puerto Rican, and other Hispanic subjects from predominantly lower- and middle-income families. The test manual lists the mean and standard deviation for each scale by gender and ethnicity for the age groups 5–7, 8–10, and 11–13.

4.15.7.6.3 Clinical utility

The TEMAS brings to clinical assessments the distinct advantages of a well conceptualized and quantitatively standardized story-telling technique that is also culturally sensitive and normed for minority groups of young people. As in the case of the RATC, its coded scales and profile graphs facilitate interpretation and provide the type of documentation that typically proves valuable in forensic cases. Moreover, this measure stands alone as a story-telling test proved applicable in the assessment of African-American and Hispanic children and adolescents. The TEMAS accordingly merits serious consideration for inclusion in a test battery for evaluating personality functioning in young people, especially if they come from an urban minority background and if there are forensic issues in the case.

At this stage in its development, however, the TEMAS has some drawbacks that examiners should keep in mind. First, although it appears to be useful for assessing level of adjustment and incorporation of societal norms among both minority and nonminority youth, TEMAS does not provide the breadth of information concerning specific types of concerns and relationships that emerge from TAT, CAT, and RATC analyses. Second, the normative data available thus far are limited to 5–13-year-old city dwellers and thus do not provide a psychometric basis for drawing conclusions about adolescents aged 14–18, or suburban and rural dwelling youngsters. Third, the length of time (two hours) required to administer the 23 card long-form of the test is often impractical in clinical evaluations. The nine card short-form is an attractive alternative but, as noted, adequate reliability has not yet been demonstrated for the short-form.

4.15.7.7 Draw-a-person

The use of human figure drawings as a projective method of personality assessment is based on the expectation that how subjects draw people will reveal aspects of how they perceive themselves and feel about others. Clinical use of drawing techniques originated in work with children, and the first formal drawing test, called the draw-a-man, was developed by Goodenough (1926) and later refined by Harris (1963) and Naglieri (1988) for use as a nonverbal measure of intellectual development. Machover (1948) introduced the notion of using human figure drawings as a projective device that generates nonverbal, symbolic messages concerning subjects' impulses, anxieties, and conflicts, and she proposed numerous possible meanings for structural features of drawings (e.g., where figures are placed on the page) and the manner in which various parts of the body are drawn (e.g., a disproportionaly large head).

Koppitz (1968, 1984) subsequently focused attention specifically on draw-a-person (DAP) assessment of young people and used on Machover's interpretive hypotheses to formulate 30 specific indicators of emotional disturbance involving the quality of drawings (e.g., asymmetry, transparencies), special features (e.g., teeth showing, arms clinging to body), and omission of body parts (e.g., no eyes, no arms). Naglieri and his colleagues (Naglieri, McNeish, & Bardos, 1991; Naglieri & Pfeiffer, 1992) have developed further coding refinements to produce the draw-a-person: screening procedure of emotional disturbance (DAP:SPED). The DAP:SPED is an actuarialy derived and normatively based system comprising 55 objectively scorable items, such as the measured dimensions and placement of figures. It is intended as a screening test for classifying young people aged 6–17 with respect to their likelihood of having adjustment difficulties that call for further evaluation.

4.15.7.7.1 Administration and scoring

The DAP is administered by giving subjects a plain 8.5 × 11 inch piece of paper and asking them to draw a person. When they have finished the drawing, they are given another piece of paper and asked to draw a person of the opposite sex from the one they have just drawn. Subjects are further instructed to draw the figure of a whole person rather than a cartoon or stick figure. In keeping with further suggestions by Machover (1951), subjects are also typically asked and to draw a picture of themselves and to provide some thematic content concerning the figures they have drawn.

As elaborated by Handler (1985), there are three alternative ways of eliciting such thematic content: by asking subjects to associate to their drawings, by asking them to make up a story about the people they have drawn, or by asking them specific questions about their drawings. Machover (1951) provided 31 specific questions to be used for this purpose with children, such as "What is their ambition?," "How happy are they?," "What do they worry about?," and "What are their good points?"

As with story-telling techniques, figure drawing methods are most commonly interpreted in clinical practice by an inspection in which personality characteristics are inferred primarily from subjective impressions of noteworthy or unusual features of the figures drawn. Items included in the Koppitz and DAP:SPED scales may often enter into these impressionistic assessments. However, these scales serve only to identify emotional disturbance without contributing in other ways to personality description, such as by indicating how individuals process information, handle emotion, and manage stress. They have consequently not been widely adopted clinically, and there are no other DAP scoring systems that have attracted much attention in the research literature.

4.15.7.7.2 Psychometric foundations

There is very little psychometric foundation for traditional applications of the DAP, and clinical use of this instrument in assessing young people frequently goes well beyond any empirical justification in empirical data. The influential inspection approach used by Machover is neither standardized nor codified, which precludes any systematic evaluation of its reliability or the accumulation of a normative database. Additionally, although it may be reasonable to expect that unusual emphasis on or omission of some body part will reflect some particular concern about the nature or functions of that body part, most of Machover's specific hypotheses concerning the symbolic significance of figure drawing characteristics lack consistent research support (Kahill, 1984; Roback, 1968; Swensen, 1957, 1968).

The Koppitz system is adequately codified and intended to comprise items that have demonstrably low occurrence in the normal population. However, there is no normative database for the system, its reliability is yet to be demonstrated, and there is some question as to whether it can differentiate between well-adjusted and emotionally disturbed children. In a carefully done study in which Tharinger and Stark (1990) compared groups of mood disordered, anxiety disordered, mood/anxiety disordered, and well-adjusted youngsters aged 9.5–14.75 years, the Koppitz signs showed good interscorer agreement but did not differentiate among the subject groups either in mean total score or in the frequency of any of the 30 individual items.

Such findings warrant concern that the DAP may be a faulty projective technique with questionable propriety for continued clinical use. However, it could be that the DAP is a potentially sound method for which there has not yet been sufficiently sophisticated development and evaluation to document its capacities. The previously mentioned work of Naglieri et al. (1991) on the DAP:SPED appears to speak to this point. The DAP:SPED was standardized on a representative national sample of 2355 children and adolescents aged 6–17 years; its objective scoring procedures have generated inter-rater agreements above 90%; its internal consistency (alpha) reliability estimates are 0.76 among 6–8-year-olds, 0.77 among 9–12-year-olds, and 0.71 among 13–17-year-olds; and its normalized total score has shown substantial capacity to differentiate nonpatient youngsters from those with identified behavioral or emotional problems (McNeish & Naglieri, 1991; Naglieri & Pfeiffer, 1992).

Also noteworthy is the work of Tharinger and Stark (1990), who paired their failure to validate the Koppitz system with an investigation of a proposed new system of their own, the DAP integrative system. The integrative DAP is based on a qualitative holistic scoring approach in which drawings are given an overall adjustment rating on a scale from 1 (absence of psychopathology) to 5 (severe psychopathology). Examiners are instructed to base their impressions on their integrated sense of four characteristics of a drawing, with the pathological end of the scale involving (i) inhumanness of the drawing suggesting feelings of being incomplete, grotesque, or monstrous; (ii) lack of agency as conveyed by a sense of powerlessness; (iii) lack of well-being as reflected in negative facial expressions; and (iv) a hollow, vacant, or stilted sense indicating lack of capacity to interact.

In the same study in which Tharinger and Stark failed to validate the Koppitz system, their integrative system total score significantly discriminated the mood disordered and anxiety/mood disordered subjects from the well-adjusted youngsters and also correlated significantly with the Coopersmith self-esteem inventory, thus attesting the capacity of the DAP to depict a youngster's sense of self. Contemporary literature abounds with sharply divided opinion concerning whether the DAP is a worthless test that should no longer be used (e.g., Gresham, 1993; Motta, Little, & Tobin, 1993) or is instead a potentially valuable clinical tool that has too often been carelessly used or inadequately researched (e.g., Bardos, 1993; Holtzman, 1993). Results to date with the DAP:SPED and the DAP integrative system give some reason to believe that improved methodology may yet establish sound psychometric foundations for carefully specified applications of the DAP.

4.15.7.7.3 Clinical utility

Despite its widespread use in describing subjects' personality characteristics, there is virtually no empirical evidence that traditional DAP interpretation has any clinical utility. Smith and Dumont (1995) asked a group of experienced psychologists and graduate students who had been trained in the use of the DAP to review and comment on a case file. They found that these clinicians routinely utilized specific symbolic representations in the drawings to draw inferences about the client's personality characteristics and diagnostic status—even though research does not support any such isomorphic correspondence of specific drawing characteristics to specific features of personality. Aside from the general ethical issues of using test instruments in unwarranted ways, examiners who are preparing forensic testimony jeopardize their credibility by employing the DAP in this manner.

This does not mean that human figure drawings are without utility in the clinical assessment of children and adolescents. First, by contrast with the practical difficulties of administering the full RATC and TEMAS, the DAP is an easily administered measure that requires no test stimuli and usually takes less than 10 minutes.

Second, as a nonverbal measure the DAP can prove especially useful in the evaluation of frightened, reticent, or otherwise uncommunicative young people, and it is unaffected by language difficulties or bilingualism. As noted by Cummings (1986), moreover, many young children may be more capable of expressing their thoughts and feelings in drawings than in words, and drawing pictures is a more familiar activity to most youngsters than most of the tasks that are set for them in a psychological examination.

Third, unusual characteristics of drawings can suggest avenues for further exploration, even in the absence of definitive conclusions about what these features signify. Examiners need to be circumspect in pursuing such avenues, however, lest they prematurely conclude that they are exploring in correct directions. The pursuit of speculative hypotheses is just as likely to lead into blind alleys as down fruitful paths.

Finally, there is ample indication in the contemporary literature that such approaches as the DAP:SPED and the DAP Integrative System can help to identify the presence and severity of emotional disturbance in general and can accordingly contribute to treatment recommendations, planning, and monitoring. In the future, codifiable and interpretable thematic content in subjects' stories about their drawings, comparable to the data derived from story telling techniques, may further enhance the utility of the DAP.

4.15.7.8 House-tree-person

The House-tree-person (HTP) test was devised by Buck (1948, 1985) as a means of tapping the concerns, interpersonal attitudes, and self-perceptions of young people more fully than is possible with human figure drawings alone. The HTP is intended for use with anyone over the age of three and was regarded by Buck as a nonthreatening instrument that can serve well to minimize a child's anxiety in testing situations and to assess personality functioning in multicultural and bilingual settings. As postulated by Buck and subsequently elaborated by Hammer (1958, 1985), subjects' drawings of these three objects are considered to provide symbolic representations of important aspects of their world.

Specifically, the house, as a dwelling place, is expected to arouse feelings toward the subject's home life and family relationships and, particularly for children, attitudes toward their parents and siblings. The tree is seen as encouraging projection of personal feelings about the self that would be more anxiety-provoking to express in drawing a person, because the latter is more obvious in its representation as a self-portrait. More specifically, the way the trunk of the tree is drawn is considered to portray a subject's feeling of basic power and inner strength; the branches are seen as depicting the subject's ability to derive satisfaction from the environment; and the overall organization of the drawing is taken as a reflection of the individual's feeling of intrapersonal balance. Finally, the drawing of the person is expected to reveal aspects of how subjects view themselves, how they would like to be, and what they think about significant other people in their lives.

4.15.7.8.1 Administration and scoring

Although Buck recommends a four-page booklet with pages measuring 7 × 8.5 inches, most HTP examiners use four sheets of standard 8.5 × 11 in paper on which subjects are asked first to draw "as good a picture of a house as you can," then a tree, and then a person of each sex. Subjects are told they can take us long as they wish and draw any kind of house, tree, or person they like. Completion of the drawings is followed by an interrogation phase in which numerous questions devised by Buck are used to encourage subjects to define, describe, and associate to their drawings (e.g., "About how old is that tree?," "What does that house make you think of?" "Is that person happy?" Buck also recommended a chromatic phase of the HTP in which subjects would do a second rendering of their drawings in crayon rather than pencil, followed by another interrogation. There are no data to indicate the frequency with which examiners conduct Buck's full HTP administration or instead limit the test to the pencil drawings, without either an inquiry or a chromatic phase.

Following the example of Goodenough, Buck originally proposed an elaborate quantitative system for objective coding of structural features of the HTP drawings, such as their size and proportions. As best as can be determined from the literature, however, quantitative coding of the HTP has rarely been employed in clinical practice. Instead, clinicians using this instrument typically rely on a qualitative inspection technique to identify symbolic implications of drawing characteristics for aspects of personality functioning. As in the case of interpreting the DAP in the tradition of Machover, many of the interpretive hypotheses for the HTP suggested by Buck and by Hammer are quite specific. A door that is tiny in relation to the size of the house is interpreted as indicating reluctance to make contact with the environment and an inhibited capacity for social relations, for example, and overemphasis on the roots of the tree where they make contact with the ground is taken as evidence of subjects' concerns about losing their grip on reality.

4.15.7.8.2 Psychometric foundations

The entire psychological literature contains only a handful of articles bearing on the psychometric foundations of the HTP. Most of these are over 25 years old, and none of them provides convincing supportive evidence for the interpretive uses of the instrument recommended by its leading proponents. Buck's (1985) 350-page revised manual contains extensive guidelines and case illustrations to facilitate interpretation, but neither reliability nor validity appears in the index. As intriguing as the rationale for the instrument may be to clinicians, especially those who are psychodynamically oriented, there is at present no empirical basis to warrant inferring personality characteristics from it.

4.15.7.8.3 Clinical utility

Similar to the DAP, the HTP offers the potential advantages in clinical practice of a brief, easily administered, nonverbal, and largely culture-free assessment instrument, along with perhaps being even less anxiety provoking than the DAP. Emerging refinements with the DAP suggest that the HTP as well might prove useful in identifying maladjustment in general, monitoring progress and change in psychotherapy, and even in pointing to possible specific areas of conflict and concern. Moreover, like the DAP inquiry, the HTP interrogation can produce story-telling content that in turn can be codified and suggest topics for further exploration. As matters presently stand, however, any such utility remains an unfulfilled potential, and examiners should be circumspect about including the HTP in their test batteries and basing any firm conclusions on it.

4.15.7.9 Kinetic Family Drawing

Machover (1948) and numerous other clinicians who pioneered in using the DAP to assess young people suggested that useful information might also be obtained by asking individuals to draw members of their family. This suggestion was formalized by Burns and Kaufman (1970, 1972) as the kinetic family drawing (KFD) technique, in which subjects are instructed to draw a picture of everyone in their family, including themselves, doing something. These drawings are then examined for such objective features as omissions of body parts or of members of the family and interpreted according to the actions, styles, and symbols represented in them.

Actions in the Burns and Kaufman approach refer to the ways in which the figures drawn are behaving toward each other, which are thought to provide clues to the intensity and emotional tone of their relationships. Styles concern barriers between family members that prevent them from interacting at all, which may be expressed by drawing some of them at a far distance from the others or encasing them in a circle or a box. Symbols comprise a list of specific items, such as beds, flowers, stoves, cats, and the like, the inclusion of which is considered to reflect various specific unconscious impulses or concerns.

Publication of the KFD was followed by a school adaptation of the technique by Prout and Phillips (1974), known as the kinetic school drawing (KSD), in which children are asked to draw a school picture of themselves, their teacher, and a friend or two in which everyone is doing something. The KSD was intended to provide information about peer relationships and about attitudes and concerns related to school in the same manner as the KFD does for family relationships and feelings about the home.

Knoff and Prout (1985a, 1985b) subsequently recommended combining the KFD and KSD and administering both measures for purposes of analysis and comparison. This combined approach, which they call the kinetic drawing system, is expected to identify adjustment difficulties both at home and in school, to clarify causal or reciprocal relationships between family and school-related issues, and to indicate which people in subjects' lives (e.g., father, sister, teacher) are sources of support of tension. In a further extension of this approach, Burns (1987) has developed a kinetic house-tree-person, in which subjects are instructed to draw on a single page a picture of a house, a tree, and a person in "some kind of action."

4.15.7.9.1 Administration and scoring

Similar to procedures followed in other drawing techniques, administration of the KFD consists of giving subjects an 8.5 × 11 inch piece of plain paper and asking them to draw a picture of everyone in their family, including themselves, doing something. Beginning with an interpretive approach modeled basically after Machover's qualitative system, Burns (1982) developed a long list of actions, style, and symbol characteristics for examiners to consider. As summarized by Knoff and Prout (1985a, 1985b), at least four objective methods for coding characteristics of kinetic drawings have also been proposed by various investigators, but none of these has become consistently visible in the literature. A review of the KFD literature by Handler and Habenicht (1994) indicates in general that cumulative knowledge concerning the adequacy and utility of this instrument has been limited by considerable variation in whether and how it has been scored in research studies and applied in clinical practice.

Also of note is a thematic elaboration of the KFD proposed by McConaughy and Achenbach (1994) and called the semistructured interview protocol. In this approach subjects are given the following question to answer after they have completed their family drawing: (i) what are they doing, (ii) what kind of person is [each member], (iii) what are three words that describe [each member], (iv) how does [each member] feel in this picture, (v) what is [each member] thinking, (vi) who do you get along with best, (vii) who do you get along with least, and (viii) what is going to happen next in your picture?

4.15.7.9.2 Psychometric foundations

Handler and Habenicht (1994) were able to reference a substantial number of publications concerning the KFD. However, they were forced to conclude that, despite its widespread use, this projective method has not yet been adequately developed with respect to its psychometric properties. Most of the systems previously proposed for coding KFD characteristics can achieve substantial inter-rater reliability, with percentages of agreement generally ranging well above 0.85 in various studies (Cummings, 1986; Handler & Habenicht, 1994). However, none of these scoring systems has demonstrated satisfactory retest reliability, even after very brief intervals, and there is little empirical basis for challenging the opinion that "the KFD still remains primarily a clinical instrument with inadequate norms and questionable validity" (Handler & Habenicht, 1994, p. 441).

With this in mind, Handler and Habenicht have recommended a holistic, integrative approach to KFD interpretation, much in the manner of the previously noted integrative DAP method used by Tharinger and Stark (1990), rather than the coding and summation of lists of signs and symbols. In the same study in which Tharinger and Stark demonstrated the superiority of their integrative DAP to the Koppitz scoring system, they also found that a holistic method of evaluating kinetic drawings discriminated adjusted from maladjusted children more effectively than a traditional scoring guide for KFD interpretation developed by Reynolds (1978).

In developing his scoring guide, Reynolds urged clinicians and researchers to avoid cookbook approaches that fail to go beyond positing direct and absolute links between individual drawing characteristics and specific personality features. Instead, scores for drawing characteristics should be only a first step in interpreting drawings as gestalts that take on meaning only in relation to a youngster's personal, family, and cultural context. It can be hoped that eventual attention to this sound advice, perhaps through the standardization of integrated coding systems, will establish for kinetic drawings the validity that is presently lacking for them to be used with confidence in clinical evaluations.

4.15.7.9.3 Clinical utility

Conceived as a means of understanding how children and adolescents conceptualize their family and perceive themselves within the family context, the KFD is a potentially useful instrument for elaborating the interpersonal dynamics of young people and orchestrating individual or family therapy for those with adjustment difficulties. By employing the kinetic drawing system, examiners can explore school-based as well as family-related issues for such purposes. Of further potential benefit, subjects from a variety of ethnic and minority group backgrounds have been found to express dimensions of their family culture in their KFDs (Handler & Habenicht, 1994), which suggests a role for this instrument in multicultural assessment.

Unfortunately, however, the previously noted psychometric limitations of the KFD makes its clinical use problematic at present, particularly with respect to basing any firm conclusions on what and how subjects draw. Until such time as adequate research has clarified what the KFD and other projective drawings methods can and cannot do, clinicians are well advised to regard their figure drawing findings as suggesting but not confirming any notions about the subject. We would endorse in this regard the conclusion of Knoff (1990) that "The hypothesis generating use of projective drawings and their ability to be interpreted within various psychological orientations remains both viable and defensible [but] the validity of hypotheses tied to specific drawing characteristics must still be determined" (p. 99).

4.15.7.10 Sentence Completion Methods

Sentence completion methods consist of initial words or phrases, called stems, that subjects are asked to extend either orally or in writing into complete sentences. Typically used stems vary in length and structure from just one or two words, such as "People ..." or "I wish ...," to detailed specification of people or situations, such as "If only my mother would ..." or "When he found he had failed the examination, he ..." As in the case of responses on other projective methods, the manner in which subjects complete sentences is expected to provide an indirect source of information concerning their underlying feelings, attitudes, and level of adjustment. In addition to eliciting general information in these respects, the specificity in many sentence completion stems encourages subjects to reveal their orientation toward particular events and circumstances in their lives.

Sentence completion methods originated in word association tests dating back to the 1890s, and, as reviewed by Haak (1990) and Lah (1989b), were developed during the 1940s and 1950s into a large number of formal and informal versions, of which the most noteworthy published scales were the Rohde, the Sacks, the Forer, and the Miale-Holsopple sentence completion tests and the Rotter incomplete sentences blank (RISB). Of these, the RISB has become the best known and most widely used and includes an adult, a college, and a high school form (Rotter, Lah, & Rafferty, 1992). Also available are more recently developed sentence completion forms constructed by Brown and Unger (1992) for use with adolescents and adults and by Hart (1986) for evaluating school age children. Further comments focus primarily on the RISB, as the predominant and most recently revised exemplar of the sentence completion method, and on the Hart sentence completion test (HSCT), because of its specificity for assessing young people.

4.15.7.10.1 Administration and scoring

The RISB is a 40-item test printed on the front and back of a one-page form and comprised mainly of brief stems (eight have just a single word, 21 have two words, six have three words, and the remaining five have four words). Subjects are given a pencil and asked to "complete these sentences to express your real feelings." The test can be administered either individually or in a group setting; however, Rotter et al. (1992) caution against taking oral rather than written responses, primarily because doing so injects an interpersonal component into the testing administration that can influence subjects' responses in various confounding ways.

The RISB and other sentence completion tests are typically interpreted in clinical practice by the inspection methods we have described previously; that is, examiners read the content of the items and from impressions of what they signify concerning a subject's probable personality characteristics. Rotter et al. (1992) also provide a scoring system for rating each item on a seven-point scale from 0 (most positive adjustment) to 6 (most indication of conflict). These ratings are totaled to yield an overall adjustment score. There is little indication that this or any other codification has received much attention in clinical practice or research studies, even though some work with college students indicated that the overall RISB score can discriminate those who are receiving counseling or psychotherapy from their nonpatient peers (Lah, 1989a).

The HSCT is also a 40-item measure in which the item stems were designed specifically to use with children and to sample family, social, school, and self dimensions of their lives (Hart, Kehle, & Davies, 1983). HSCT scoring involves rating each item as negative, neutral, or positive with respect to adjustment and also rating 10 scales composed of various clusters of items concerned with perceptions of self, family, and school on a five-point negative-to-positive continuum. Criteria are provided for each item to guide examiners in their ratings and thereby to enhance scoring objectivity and the prospects for achieving inter-rater agreement.

4.15.7.10.2 Psychometric foundations

Although there is a substantial RISB research literature, most of the published studies have used the instrument as a measure of adjustment but, as pointed out by Lah (1989a), they were not designed to evaluate its properties. Particularly with respect to its validity, there has been virtually no accumulation of empirical evidence to support any diagnostic or predictive inferences about personality functioning beyond the modestly demonstrated capacity of the RISB to identify maladjustment. Likewise, little systematic progress has been made in assessing the reliability of RISB data and establishing normative standards for them.

This is a regrettable state of affairs because, aside from being a psychodynamically compelling method of enriching personality assessment, the RISB is an eminently codeable instrument for which excellent interscorer agreement has been demonstrated using the Rotter et al. approach (Lah, 1989a). Moreover, as discussed in the introductory portions of this chapter, sentence completion responses have numerous objective, structural, and behavioral components (e.g., time to completion, length of sentences, frequency of self- vs. other-reference, items omitted) that have rarely been considered in codifying these methods or attempting to establish a psychometric foundation for them.

The HSCT coding system has also achieved substantial inter-rater agreement and some preliminary success differentiating between emotionally disturbed and well-adjusted children. However, Hart (1986, p. 269) observed that it remained to further research to demonstrate adequate reliability and criterion validity for his instrument and to develop sufficiently broad normative data. The available literature, circa 1996, suggests that not much if any, progress has been made in this regard.

4.15.7.10.3 Clinical utility

Sentence completion methods bring to the personality assessment battery an easily administered projective test that can be given to groups as well as individuals and requires only about 30 minutes on average for subjects to complete. The scorable 40-item RISB high school form is suitable for high school and most middle school youngsters, and the 40-item HSCT can be used comfortably with younger children. Barring a language or reading difficulty, the meaning of the sentence completion stems is clear, and subjects are not asked to explain, elaborate, or otherwise account for their responses. Hence, incomplete sentence methods are often less threatening or anxiety-provoking than inkblot, story-telling, and figure drawing procedures. However, the previously noted relative lack of ambiguity in sentence completion stems means that subjects are more aware than on other projective methods of how they are presenting themselves through their responses.

At this point in time, the RISB and HSCT can be used with justification to form general impressions of a young person's level of adjustment and to formulate hypotheses concerning possible conflicts or concerns the young person has and how he or she may feel about self, other people, and certain situations. Such hypotheses can justifiably be expressed only as speculations, however, not as conclusions, and they should be considered reasonably correct only to the extent that they are supported by other reliable sources of information.

Haak (1990) has asserted that sentence completion methods can be used effectively with children to rule out intellectual difficulties, attention deficit disorder, and stress and to rule in depression, anxiety, thought disturbance, and defensiveness, and provides detailed clinical guidelines for doing so. These suggestions seem sensible in each case, but they are entirely unverified by empirical data. Hence, like other impressions formed by experienced clinicians, Haak's diagnostic guidelines provide an agenda for confirmatory research, but they should not be elevated from speculations to the status of conclusions until that empirical confirmation becomes available.

4.15.8 FUTURE DIRECTIONS

Projective methods have been extensively used, taught, and studied for many years, and it seems likely that they will continue in the future to be regularly included in test batteries for assessing personality functioning in young people. Survey data cited in this chapter document that over the last 30 years psychologists have continued with much the same frequency to apply these instruments in practice and utilize them in research. During this same period the psychometric adequacy and clinical contributions of projective methods have been regularly and vigorously challenged, and this chapter indicates that the uses to which some of these instruments are put often go beyond available justification in empirical data.

How do we account for the persistent use by clinicians of many such as yet unvalidated methods? One could say that there just happen to be legions of uninformed or unethical clinicians in practice who do not hesitate to employ useless methods if doing so serves their purpose in some way. However, it seems doubtful that a vast segment of the profession deserves to be tarred with such a broad brush of evil. Rather, at least the majority of professionals who use projective methods must have good reason on the basis of their clinical and individual case experience to believe sincerely that these methods provide valid and useful information about personality functioning in ways that have not yet been translated into supporting research data. One could then say that such clinical confidence in projective methods is based solely on illusory correlation. However, it seems doubtful that a vast segment of professional personality assessors could be so thoroughly deluded.

These observations suggest that projective techniques have been and will continue to be used because they yield valuable information and generate fruitful hypotheses in clinical assessments. Hopefully, adequate research methods will lead the way toward improved standardization and codification of projective methods, that will in turn enhance the psychometric foundations on which they rest. As Kuehnle (1996) has pointed out, inappropriate use of projective instruments for purposes for which they have not been validated, such as identifying children as having been traumatized or sexually abused, violates ethical standards and risks causing harm to young people and their families.

More than any of the other projective methods, the Rorschach has had the benefit of rigorous attention to research methodology, as demonstrated in publications by Exner (1995) and Weiner (1995a) and, as indicated earlier in this chapter, this method presently rests on a solid psychometric foundation. What the future of projective methods needs is similar methodological attention to documenting the reliability of other techniques and their validation for various purposes. To say that sophisticated research methods are not applicable to projective methods is as unwarranted as asserting that these methods are by nature invalid. To say that projective test responses cannot be examined scientifically without detracting from their idiographic richness sells projective methods short and prevents them from realizing their full potential.

Along with these needs for improved research and a narrowed gap between data and practice, the most important future direction for projective testing of young people lies in developing adequately representative normative data that will facilitate age-specific and multicultural assessment. There is no lack of projective techniques, but there is a decided lack of reliable information concerning how children and adolescents of different ages, from diverse backgrounds, and with various kinds of personality strengths and weaknesses, should be expected to respond to them. In addition to the nearly empty coffers of cross-sectional normative data of these kinds, the cupboard is virtually bare with respect to longitudinal ata indicating how young people are likely to change over time and with maturation in how they respond to projective tests. Collecting and disseminating such data is the number one agenda item for the future development of projective assessment of children and adults.

4.15.9 SUMMARY

Projective tests are methods of personality assessment in which some degree of ambiguity in the test stimuli or instructions creates opportunities for subjects to structure their responses in terms of their individual personality characteristics, and thereby provide information about the nature of these characteristics. Although projective methods are accordingly more ambiguous and less structured than so-called objective methods, the differences between these methods are relative rather than absolute. All projective tests contain objective as well as subjective features and elicit responses that are representative as well as symbolic of behavior, and they differ from each other in the extent to which they are ambiguous.

Because of their relatively unstructured nature, projective tests measure personality functioning in subtle and indirect ways and tap underlying psychological characteristics at a less conscious level than relatively structured measures. Projective test data consequently provide valuable information about how people are likely to think, feel, and act that is difficult to obtain from objective assessment procedures, and they are also less susceptible than objective test data to the influence of test-taking attitudes.

Projective methods can be used to good effect with children and adolescents as well as adults. The basic interpretive conclusions and hypothesis that attach to projective test variables apply regardless of the age of the subject, provided that examiners determine the implications of their data in the light of normative developmental expectations. Surveys of clinical and school settings indicate that the projective instruments most frequently administered in evaluating young people are the Rorschach inkblot method, the thematic apperception test (TAT), the children's apperception test, the Roberts apperception test for children (RATC), the tell-me-a-story (TEMAS) test, the draw-a-person, the house-tree-person, the kinetic family drawing, and alternate forms of the sentence completion test.

For each of these nine projective tests, this chapter reviews their composition, administration, scoring, psychometric foundations, and clinical utility. Despite widespread utilization of these nine tests in clinical practice to draw conclusions about the personality characteristics, level of adjustment, and treatment needs of young people, only the Rorschach presently rests on a solid empirical foundation. Properly collected and interpreted, Rorschach data provide numerous demonstrably reliable and valid indices that facilitate differential diagnosis and treatment planning for children and adolescents with adjustment difficulties.

The RATC and TEMAS have shown through the development of standardized administration and scoring procedures that story-telling methods have the potential to achieve psychometric respectability. However, further research on the RATC and TEMAS is needed with regard to multicultural and adolescent norms, respectively. With regard to the TAT, recently emerging psychometrically sound schemes for coding specific personality characteristics reflected in thematic content, such as SCORS, appear to provide a basis for empirical decision making.

The other measures reviewed have in various ways also shown potential to be codified and refined on the basis of empirical data, and there is reason to be hopeful that advances in research methodology will eventually close a currently regrettable gap between what is known for sure about these projective methods and what is frequently assumed to be true about them in clinical practice. Until this gap is narrowed, and especially until such time as more extensive normative and multicultural data become available, most clinical inferences from projective data should be regarded as hypotheses to be confirmed rather than as facts on which to base conclusions and recommendations.

4.15.10 REFERENCES

Abraham, P. P., Lepisto, B. L., Lewis, M. G., & Schultz, L. (1994). Changes in Rorschach variables of adolescents in residential treatment: An outcome study. *Journal of Personality Assessment, 62*, 505–514.

Alvarado, N. (1994). Empirical validity of the Thematic Apperception Test. *Journal of Personality Assessment, 63*, 59–79.

Ames, L. B., Metraux, R. W., Rodell, J. L., & Walker, R. N. (1974). *Child Rorschach responses* (Rev. ed.). New York: Brunner/Mazel.

Ames, L. B., Metraux, R. W., & Walker, R. N. (1971). *Adolescent Rorschach responses* (Rev. ed.). New York: Brunner/Mazel.

Archer, R. P., Imhof, E. A., Maruish, M., & Piotrowski, C. (1991). Psychological test usage with adolescent clients: 1990 survey findings. *Professional Psychology, 22*, 247–252.

Archer, R. P., & Krishnamurthy, R. (1993a). A review of MMPI and Rorschach interrelationships in adult samples. *Journal of Personality Assessment, 61*, 277–293.

Archer, R. P., & Krishnamurthy, R. (1993b). Combining the Rorschach and MMPI in the assessment of adolescents. *Journal of Personality Assessment, 60*, 132–140.

Atkinson, J. W., & Feather, N. T. (1966). *A theory of achievement motivation.* New York: Wiley.

Bardos, A. N. (1993). Human figure drawings: Abusing the abused. *School Psychology Quarterly, 8*, 177–181.

Bellak, L. (1993). *The T.A.T., C.A.T., and S.A.T. in clinical use* (5th ed.). Boston: Allyn & Bacon.

Bellak, L., & Hurvich, M. (1966). A human modification of the Children's Apperception Test. *Journal of Projective Techniques, 30*, 228–242.

Bellak, L., & Siegel, H. (1989). The Children's Apperception Test (CAT). In C. S. Newmark (Ed.), *Major psychological assessment instruments* (Vol. II, pp. 99–127). Boston: Allyn & Bacon.

Bornstein, R. F. (1995). Sex differences in objective and projective dependency tests: A meta-analytic review. *Assessment, 2*, 319–331.

Buck, J. N. (1948). The H-T-P technique, a qualitative and quantitative method. *Journal of Clinical Psychology, 4*, 317–396.

Buck, J. N. (1985). *The House-Tree-Person technique: Revised manual.* Los Angeles: Western Psychological Services.

Burns, R. C. (1982). *Self-growth in families: Kinetic Family Drawings (K-F-D) research and application.* New York: Brunner/Mazel.

Burns, R. C. (1987). *Kinetic-House-Tree-Person drawings (K-H-T-P).* New York: Brunner/Mazel.

Burns, R. C., & Kaufman, S. H. (1970). *Kinetic Family Drawings (K-F-D): An introduction to understanding children through kinetic drawings.* New York: Brunner/Mazel.

Burns, R. C., & Kaufman, S. H. (1972). *Actions, styles, and symbols in Kinetic Family Drawings (K-F-D).* New York: Brunner/Mazel.

Butcher, J. N., & Rouse, S. V. (1996). Personality: Individual differences and clinical assessment. *Annual Review of Psychology, 47*, 87–111.

Chandler, L. A. (1990). The projective hypothesis and the development of projective techniques for children. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children* (Vol. 2, pp. 55–69). New York: Guilford Press.

Costantino, G., Colon-Malgady, G., Malgady, R. G., & Perez, A. (1991). Assessment of attention deficit disorder using a thematic apperception technique. *Journal of Personality Assessment, 57*, 97–95.

Costantino, G., Malgady, R. G., Bailey, J., & Colon-Malgady, G. (1989). *Clinical utility of TEMAS: A projective test for children.* Paper presented at the meeting of the Society for Personality Assessment, New York.

Costantino, G., Malgady, R. G., Colon-Malgady, G., & Bailey, J. (1992). Clinical utility of the TEMAS with non minority children. *Journal of Personality Assessment, 59*, 433–438.

Costantino, G., Malgady, R. G., & Rogler, L. H. (1988). *TEMAS (Tell-Me-A-Story) manual.* Los Angeles: Western Psychological Services.

Costantino, G., Malgady, R. G., Rogler, L. H., & Tusi, E. C. (1988). Discriminant analysis of clinical outpatients and public school children by TEMAS: A thematic apperception test for Hispanics and Blacks. *Journal of Personality Assessment, 52*, 670–678.

Cramer, P. (1987). The development of defense mechanisms. *Journal of Personality, 55*, 597–614.

Cramer, P., & Blatt, S. J. (1990). Use of the TAT to measure change in defense mechanisms following intensive psychotherapy. *Journal of Personality Assessment, 54*, 236–251.

Cummings, J. A. (1986). Projective drawings. In H. M. Knoff (Ed.), *The assessment of child and adolescent personality* (pp. 199–204). New York: Guilford Press.

Dana, R. H. (1985). Thematic Apperception Test (TAT). In C. Newmark (Ed.), *Major psychological assessment instruments* (pp. 89–134). Boston: Allyn & Bacon.

Dana, R. H. (1993). *Multicultural assessment perspectives for professional psychology.* Boston: Allyn & Bacon.

Dana, R. H. (1996). Culturally competent assessment practice in the United States. *Journal of Personality Assessment, 66*, 472–487.

Di Leo, J. H. (1983). *Interpreting children's drawings.* New York: Brunner/Mazel.

Elbert, J. C., & Holden, E. W. (1987). Child diagnostic assessment: Current training practices in clinical psychology internships. *Professional Psychology, 18*, 587–596.

Exner, J. E., Jr. (1991). *The Rorschach: A comprehensive system. Vol. 2. Interpretation* (2nd ed.). New York: Wiley.

Exner, J. E., Jr. (1993). *The Rorschach: A comprehensive system. Vol. 1. Basic foundations* (3rd ed.). New York: Wiley.

Exner, J. E., Jr. (Ed.) (1995). *Issues and methods in Rorschach research.* Mahwah, NJ: Erlbaum.

Exner, J. E., Jr., & Andronikoff-Sanglade, A. (1992). Rorschach changes following brief and short-term therapy. *Journal of Personality Assessment, 59*, 59–71.

Exner, J. E., Jr., Thomas, E. A., & Mason, B. (1985). Children's Rorschachs: Description and prediction. *Journal of Personality Assessment, 49*, 13–20.

Exner, J. E., Jr., & Weiner, I. B. (1995). *The Rorschach: A comprehensive system. Vol. 3. Assessment of children and adolescents* (2nd ed.). New York: Wiley.

Finn, S. E. (1996, March). *Assessment feedback integrating MMPI-2 and Rorschach findings.* Paper presented at the annual meeting of the Society for Personality Assessment, Denver, CO.

Frank, L. K. (1939). Projective methods for the study of personality. *Journal of Psychology, 8*, 389–413.

Freedenfeld, R., Ornduff, S., & Kelsey, R. M. (1995). Object relations and physical abuse: A TAT analysis. *Journal of Personality Assessment, 64*, 552–568.

Freud, S. (1958). Psycho-analytic notes upon an autobiographical account of a case of paranoia (dementia paranoides). *Standard edition* (Vol. XII, pp. 9–82). London: Hogarth. (Original work published in 1911.)

Freud, S. (1962). Further remarks on the neuro-psychoses of defence. *Standard edition* (Vol. III, pp. 162–185). London: Hogarth. (Original work published in 1896.)

Ganellen, R. J. (1996). *Integrating the Rorschach and the MMPI-2 in personality assessment.* Mahwah, NJ: Erlbaum

Gardner, R. A. (1971). *Therapeutic communication with children: The mutual storytelling technique.* Northvale, NJ: Aronson.

Goodenough, F. L. (1926). *Measurement of intelligence by drawings.* New York: Harcourt, Brace & World.

Gresham, F. M. (1993). "What's wrong in this picture?": Response to Motta et al.'s review of human figure drawings. *School Psychology Quarterly, 8*, 182–186.

Haak, R. A. (1990). Using the sentence completion to assess emotional disturbance. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children* (Vol. 2, pp. 147–167). New York: Guilford Press.

Hammer, E. F. (1958). *The clinical application of projective drawings.* Springfield, IL: Charles C. Thomas.

Hammer, E. F. (1985). The House-Tree-Person test. In C. S. Newmark (Ed.), *Major psychological assessment instruments* (pp. 135–164). Boston: Allyn & Bacon.

Handler, L. (1985). The clinical use of the Draw-a-person test (DAP). In C. S. Newmark (Ed.), *Major psychological assessment instruments* (pp. 165–216). Boston: Allyn & Bacon.

Handler, L., & Habenicht, D. (1994). The Kinetic Family Drawing technique: A review of the literature. *Journal of Personality Assessment, 62*, 440–464.

Harris, D. B. (1963). *Children's drawings as a measure of intellectual maturity.* New York: Harcourt, Brace & World.

Hart, D. H. (1986). The sentence completion techniques. In H. M. Knoff (Ed.), *The assessment of child and adolescent personality* (pp. 245–272). New York: Guilford Press.

Hart, D. H., Kehle, T. J., & Davies, M. V. (1983). Effectiveness of sentence completion techniques: A review of the Hart Sentence Completion Test. *School Psychology Review, 12*, 428–434.

Haworth, M. (1965). *A schedule of adaptive mechanisms in CAT responses.* Larchmont, NY: CPS.

Hibbbard, S., Farmer, L., Wells, C., Difillipo, E., Barry, W., Korman, R., & Sloan, P. (1994). Validation of Cramer's defense mechanism manual for the TAT. *Journal of Personality Assessment, 63*, 197–210.

Hibbard, S., Hilsenroth, M. J., Hibbard, J. K., & Nash, M. R. (1995). A validity study of two projective object representation measures. *Psychological Assessment, 7*, 432–439.

Hoffman, S., & Kupperman, N. (1990). Indirect treatment of traumatic psychological experiences: The use of TAT cards. *American Journal of Psychotherapy, 44*, 107–115.

Holtzman, W. H. (1993). An unjustified, sweeping indictment by Motta et al. of human figure drawings for assessing psychological functioning. *School Psychology Quarterly, 8*, 189–190.

Hutton, J. B., Dubes, R., & Muir, S. (1992). Assessment practices of school psychologists: Ten years later. *School Psychology Review, 21*, 271–284.

Kahill, S. (1984). Human figure drawings in adults: An update of the empirical literature. *Canadian Psychology, 25*, 269–292.

Kennedy, M. L., Faust, D., Willis, W. G., & Piotrowski, C. (1994). Social-emotional assessment practices in school psychology. *Journal of Psychoeducational Assessment, 12*, 228–240.

Knoff, H. M. (1990). Evaluation of projective drawings. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children* (Vol. 2, pp. 89–146). New York: Guilford Press.

Knoff, H. M., & Prout, H. T. (1985a). The Kinetic Drawing System: A review and integration of the Kinetic Family and Kinetic School Drawing techniques. *Psychology in the Schools, 22*, 50–59.

Knoff, H. M., & Prout, H. T. (1985b). *The Kinetic drawing system: Family and school*. Los Angeles: Western Psychological Services.

Kohlberg, L. (1976). Moral stage and moralization: The cognitive developmental approach. In T. Lickona (Ed.), *Moral development and behavior: Theory, research, and social issues* (pp. 31–53). New York: Holt, Rinehart & Winston.

Koppitz, E. M. (1968). *Psychological evaluation of children's human figure drawings*. New York: Grune & Stratton.

Koppitz, E. M. (1984). *Psychological evaluation of human figure drawings by middle school pupils*. New York: Grune & Stratton.

Kuehnle, K. (1996). *Assessing allegations of child sexual abuse*. Sarasota, FL: Professional Resource Press.

Lah, M. I. (1989a). New validity, normative, and scoring data for the Rotter Incomplete Sentences Blank. *Journal of Personality Assessment, 53*, 607–620.

Lah, M. I. (1989b). Sentence completion tests. In C. S. Newmark (Ed.), *Major psychological assessment instruments, Volume II* (pp. 133–163). Boston: Allyn & Bacon.

Leichtman, M. (1996). *The Rorschach: A developmental perspective*. Hillsdale, NJ: Analytic Press.

Machover, K. (1948). *Personality projection in the drawing of the human figure*. Springfield, IL: Charles C. Thomas.

Machover, K. (1951). Drawing of the human figure: A method of personality investigation. In H. H. Anderson & C. L. Anderson (Eds.), *An introduction to projective techniques* (pp. 341–369). New York: Prentice-Hall.

McArthur, D. S., & Roberts, G. E. (1990). *Roberts Apperception Test for Children manual*. Los Angeles: Western Psychological Services.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1953). *The achievement motive*. New York: Appleton-Century-Crofts.

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review, 96*, 690–702.

McConaughy, S. H., & Achenbach, T. M. (1994). *Manual for the semistructured clinical interview for children and adolescents*. Burlington, VT: University of Vermont Department of Psychiatry.

McDowell, C., & Acklin, M. W. (1996). Standardizing procedures for calculating Rorschach inter-rater reliability: Conceptual and empirical foundations. *Journal of Personality Assessment, 66*, 308–320.

McNeish, T. J., & Naglieri, J. A. (1991). Identification of the seriously emotionally disturbed using the Draw A Person: Screening Procedure for Emotional Disturbance. *Journal of Special Education, 27*, 115–121.

Morgan, C. D., & Murray, H. A. (1935). A method for investigating fantasies: The Thematic Apperception Test. *Archives of Neurology and Psychiatry, 34*, 289–306.

Motta, R. W., Little, S. G., & Tobin, M. I. (1993). The use and abuse of human figure drawings. *School Psychology Quarterly, 8*, 162–169.

Murray, H. A. (1943). *Thematic Apperception Test manual*. Cambridge, MA: Harvard University Press.

Murray, H. A. (1951). Uses of the Thematic Apperception Test. *American Journal of Psychiatry, 107*, 577–581.

Murray, H. A. (1971). *Thematic Apperception Test: Manual* (Rev. ed.). Cambridge, MA: Harvard University Press.

Murstein, B. I. (1963). *Theory and research in projective techniques (Emphasizing the TAT)*. New York: Wiley.

Murstein, B. I., & Wolf, S. R. (1970). Empirical test of the "levels" hypothesis with five projective techniques. *Journal of Abnormal Psychology, 75*, 38–44.

Naglieri, J. A. (1988). *Draw A Person: A quantitative scoring system*. New York: Psychological Corporation.

Naglieri, J. A., McNeish, T. J., & Bardos, A. N. (1991). *Draw-A-Person: Screening Procedure for Emotional Disturbance*. Austin, TX: ProEd.

Naglieri, J. A., & Pfeiffer, S. I. (1992) Performance of disruptive behavior disordered and normal samples on the Draw A Person: Screening Procedure for Emotional Disturbance. *Psychological Assessment, 4*, 156–159.

Obrzut, J. E., & Boliek, C. A. (1986). Thematic approaches to personality assessment with children and adolescents. In H. M. Knoff (Ed.), *The assessment of child and adolescent personality* (pp. 173–198). New York: Guilford Press.

Ornduff, S. R., Freedenfeld, R., Kelsey, R. M., & Critelli, J. (1994). Object relations of sexually abused female subjects: A TAT analysis. *Journal of Personality Assessment, 63*, 223–228.

Ornduff, S. R., & Kelsey, R. M. (1996). Object relations of sexually and physically abused female children: A TAT analysis. *Journal of Personality Assessment, 66*, 91–105.

Parker, K. C. H., Hanson, R. K., & Hunsley, J. (1988). MMPI, Rorschach and WAIS: A meta-analytic comparison of reliability, stability, and validity. *Psychological Bulletin, 103*, 367–373.

Piotrowski, C. (1996). The status of Exner's Comprehensive System in contemporary research. *Perceptual and Motor Skills, 82*, 1341–1342.

Piotrowski, C., & Keller, J. W. (1989). Psychological testing in outpatient facilities: A national study. *Professional Psychology, 20*, 423–425.

Prout, H. T., & Phillips, P. D. (1974). A clinical note: The kinetic school drawing. *Psychology in the Schools, 11*, 303–306.

Reynolds, C. R. (1978). A quick scoring guide to the interpretation of children's Kinetic Family Drawings (KFD). *Psychology in the School, 15*, 489–492.

Roback, H. B. (1968). Human figure drawings: Their utility in the clinical psychologist's armamentarium for personality assessment. *Psychological Bulletin, 70*, 1–19.

Ronan, G. F., Colavito, V. A., & Hammontree, S. R. (1993). Personal problem-solving system for scoring TAT responses: Preliminary validity and reliability data. *Journal of Personality Assessment, 61*, 28–40.

Rorschach, H. (1942). *Psychodiagnostics*. Bern, Switzerland: Hans Huber. (Original work published in 1921.)

Rotter, J. B., Lah, M. I., & Rafferty, J. E. (1992). *Manual—Rotter Incomplete Sentences Blank* (2nd ed.). Orlando, FL: Psychological Corporation.

Seaton, B., & Allen, J. (1996, March). *Interscorer reliability of Rorschach structural summary data.* Poster session presented at the annual meeting of the Society for Personality Assessment, Denver, CO.

Smith, D., & Dumont, F. (1995). A cautionary study: Unwarranted interpretations of the Draw-A-Person test. *Professional Psychology, 3*, 298–303.

Stinnett, T. A., Havey, J. M., & Oehler-Stinnett, J. (1994). Current test usage by practicing school psychologists: A national survey. *Journal of Psychoeducational Assessment, 12*, 331–350.

Stone, H. K., & Dellis, N. P. (1960). An exploratory investigation into the levels hypothesis. *Journal of Projective Techniques, 24*, 333–340.

Sutton, P. M., & Swenson, C. H. (1983). The reliability and concurrent validity of alternative methods for assessing ego development. *Journal of Personality Development, 47*, 468–475.

Swensen, C. H. (1957). Empirical evaluations of human figure drawings. *Psychological Bulletin, 54*, 431–466.

Swensen, C. H. (1968). Empirical evaluations of human figure drawings: 1957–1966. *Psychological Bulletin, 70*, 20–44.

Tharinger, D. J., & Stark, K. (1990). A qualitative versus quantitative approach to the Draw-a-Person and Kinetic Family Drawing: A study of mood- and anxiety-disorder children. *Psychological Assessment, 2,* 365–375.

Thomas, A. D., & Dudek, S. Z. (1985). Interpersonal affect in TAT responses: A scoring system. *Journal of Personality Assessment, 49*, 30–37.

Vane, J. R. (1981). The Thematic Apperception Test: A review. *School Psychology Review, 1*, 319–336.

Watkins, C. E., Campbell, V. L., Nieberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology*, *26*, 54–60.

Weiner, I. B. (1977). Projective tests in differential diagnosis. In B. B. Wolman (Ed.), *International encyclopedia of neurology, psychiatry, psychoanalysis, and psychology* (pp. 112–116). Princeton, NJ: Van Nostrand Reinhold.

Weiner, I. B. (1986). Assessing children and adolescents with the Rorschach. In H. M. Knoff (Ed.), *The assessment of child and adolescent personality* (pp. 141–171). New York: Guilford Press.

Weiner, I. B. (1993). Clinical considerations in the conjoint use of the Rorschach and the MMPI. *Journal of Personality Assessment*, *60*, 148–152.

Weiner, I. B. (1994). Rorschach assessment. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome evaluation* (pp. 249–278). Hillsdale, NJ: Erlbaum.

Weiner, I. B. (1995a). Methodological considerations in Rorschach research. *Psychological Assessment*, *7*, 330–337.

Weiner, I. B. (1995b). Psychometric issues in forensic applications of the MMPI-2. In Y. S. Ben-Porath, J. R. Graham, G. C. N. Hall, R. D. Hirschman, & M. S. Zaragoza (Eds.), *Forensic applications of the MMPI-2* (pp. 48–81). Thousand Oaks, CA: Sage.

Weiner, I. B. (1996). Some observations on the validity of the Rorschach Inkblot Method. *Psychological Assessment, 8,* 206–213.

Weiner, I. B. (1997). Current status of the Rorschach Inkblot Method. *Journal of Personality Assessment, 68,* 5–19.

Weiner, I. B., & Exner, J. E., Jr. (1991). Rorschach changes in long-term and short-term psychotherapy. *Journal of Personality Assessment*, *56*, 453–465.

Westen, D., Klepser, J., Ruffins, S. A., Silverman, M., Lifton, N., & Boekamp, J. (1991). Object relations in childhood and adolescence: The development of working representations. *Journal of Consulting and Clinical Psychology*, *59*, 400–409.

Westen, D., Lohr, N., Silk, K., Kerber, K., & Goodrich, S. (1985). *Object relations and social cognition TAT scoring manual.* Ann Arbor, MI: University of Michigan.

Westen, D., Ludolph, P., Block, M. J., Wixom, J., & Wiss, F. C. (1990). Developmental history and object relations in psychiatrically disturbed adolescent girls. *American Journal of Psychiatry*, *147*, 1061–1068.

Westen, D., Ludolph, P., Lerner, H., Ruffins, S., & Wiss, F. C. (1990). Object relations in borderline adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *29*, 338–348.

Worchel, F. F., Rae, W. A., Olson, T. K., & Crowley, S. L. (1992). Selective responsiveness of chronically ill children to assessments of depression. *Journal of Personality Assessment*, *59*, 605–615.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.16

Assessment of Schema and Problem-solving Strategies with Projective Techniques

HEDWIG TEGLASI
University of Maryland, College Park, MD, USA

4.16.1 INTRODUCTION 460

4.16.2 PATTERNS OF USE 460

4.16.3 BASIC ASSUMPTIONS OF PROJECTIVE TECHNIQUES AND RESEARCH FROM OTHER PSYCHOLOGY SUBFIELDS 461

4.16.3.1 Schema Theory and Projective Techniques 462
4.16.3.2 Unconscious Processing of Information 464
4.16.3.3 Inert Knowledge vs. Usable Knowledge 465
4.16.3.4 Affective and Motivational Influences on Cognition 465
4.16.3.5 Narrative Psychology and Thematic Apperception Techniques 466

4.16.4 LEVELS OF PERSONALITY ASSESSED BY PROJECTIVE AND QUESTIONNAIRE METHODS 467

4.16.5 SELF-REPORT VS. PROJECTIVE PERSONALITY TESTING 470

4.16.6 CONTRIBUTION OF PROJECTIVE TECHNIQUES TO DIAGNOSIS AND INTERVENTION 470

4.16.7 PROJECTIVE TECHNIQUES AS PERFORMANCE MEASURES OF PERSONALITY 472

4.16.7.1 Total Personality 474
4.16.7.2 Many Correct Solutions 476
4.16.7.3 Generalize to Different Criteria 476
4.16.7.4 Differences in Conditions of Learning and Performance 477

4.16.8 SPECIFIC PROJECTIVE TECHNIQUES 477

4.16.8.1 Stimulus 477
4.16.8.2 Response 477
4.16.8.3 Task Demand 478
4.16.8.4 Interpretation 478

4.16.9 THEMATIC APPERCEPTIVE TECHNIQUES 479

4.16.9.1 Response 479
4.16.9.2 Stimuli 479
4.16.9.3 Interpretation 480

4.16.10 RORSCHACH TECHNIQUE 481

4.16.10.1 Stimulus Features and Response Parameters 482
4.16.10.2 Interpretation 483

4.16.11 DRAWING TECHNIQUES 483

4.16.11.1 Interpretation 484

4.16.12 CASE ILLUSTRATION 484

4.16.13 VALIDATION ISSUES 487
4.16.13.1 Reliability 488
4.16.13.1.1 Scorer reliability 488
4.16.13.1.2 Decisional reliability 488
4.16.13.1.3 Test–retest reliability 488
4.16.13.1.4 Internal consistency 488
4.16.13.2 Construct Validation 489
4.16.13.3 Multitrait, Multimethod Validation 489
4.16.13.4 Criterion Validation 489
4.16.13.5 Part–Whole Relationships 490
4.16.13.6 Convergence of Psychometric and Conceptual Treatment of Data 490
4.16.13.6.1 Normative 491
4.16.13.6.2 Nomothetic 491
4.16.13.6.3 Case study 491
4.16.14 FUTURE DIRECTIONS 492
4.16.15 SUMMARY 493
4.16.16 REFERENCES 495

4.16.1 INTRODUCTION

Clinical assessment practices typically lag behind the conceptualizations psychologists use to understand adaptive and maladaptive patterns in human functioning. We recognize that personality is a functional whole, but cling to assessment procedures that resemble the proverbial blind men and the elephant. Each method of assessment pertains to one dimension or level of personality, yet we expect different sources of data to confirm one another rather than to reveal the different patterns contributing to the larger mosaic. A clarification of the part–whole relationships with regard to personality assessment promotes an understanding of the place of projective techniques in a comprehensive evaluation. The aims of this chapter are: (i) to facilitate the process of integration of projective techniques with self-report approaches to personality assessment; (ii) to reframe projective techniques as performance measures of personality that differ in their demands from other tasks within the battery; (iii) to show that the central tenets of projective techniques embodied in the projective hypothesis are supported by empirical research in various psychology subfields; and (iv) to demonstrate the utility of projective techniques for assessing schema (the internal representation of reality) and problem-solving strategies to meet the reality demands posed by the task.

That projective techniques do not belong to a particular theory of personality was asserted early on (Auld, 1954). Projective methods yield open-ended responses, and their interpretation can accommodate changing paradigms in various subdisciplines, such as the cognitive and neurosciences as well as in the study of emotion and social cognition. It is easy to say that the framework to understand and interpret responses to projective tasks should coincide with key concepts in the various subdisciplines, but the task of continuously updating the conceptual frameworks that are applied to projective methods remains an ongoing challenge.

Responses to projective tests yield information about the internal structure of personality and provide the context for understanding overt behavior and symptomatology. This emphasis on inner structure is currently shared by social, cognitive, and clinical psychology. Newer appreciation is emerging of the wide role of emotions and unconscious processes (e.g., Bargh, 1994; Clore, Schwarz, & Conway, 1994), as well as of meaning structures (e.g., Bruner, 1986) underlying behavioral expression. Projective methods reveal the structure and process of personality that are largely unconscious, and they differ in important ways from checklists or self-report interviews. In addition, the competencies required by projective tasks can be distinguished from those demanded by other measures in the typical assessment battery such as the Wechsler scales.

Therefore, projective techniques can be viewed as performance measures of personality that reveal inner structures or schema as resources needed in daily life. Each projective method sets a task that requires the application of these structures to respond adaptively. All factors contributing to personality development are involved in shaping these inner structures of meaning. Therefore, the interpretation of responses to projective techniques promotes the integration of knowledge from various subdisciplines.

4.16.2 PATTERNS OF USE

Despite criticisms of projective tests, there continues to be broad interest in the Rorschach and the Thematic Apperception Test (TAT;

Butcher & Rouse, 1996). Most clinical psychology doctoral training programs include formal instruction in the Rorschach (Piotrowski & Zalewski, 1993). In addition, the majority of internship sites approved by the American Psychological Association (APA) value knowledge in Rorschach technique (Durand, Blanchard, & Mindell 1988). The Exner Comprehensive System (Exner, 1993; Exner & Weiner, 1995) has become widely accepted because it provides a more reliable and objective basis for interpretation than was previously available. Surveys find the Rorschach, TAT, and various drawing methods among the top 10 most frequently used assessment techniques (Archer, Marnish, Imhof, & Piotrowski, 1991; Piotrowski & Keller, 1989; Piotrowski, Sherry, & Keller, 1985; Watkins, Campbell, & McGregor, 1988; Watkins, Campbell, Nieberding, & Hallmark, 1995). Interest in the Rorschach and TAT is also evident outside of the USA (Weiner, 1994).

The contrast between the negative evaluation of projective methods by researchers and their continued popularity with practitioners (Piotrowski, 1984) begs for a scrutiny of the assumptions and methods applied to judging the utility of these techniques. This state of affairs calls for an analysis not only of psychometric properties but also of the conceptual and methodological underpinnings pertaining to projective testing. Experienced practitioners continue to use these techniques because they find them helpful. For instance, an advantage of the Rorschach is that the validity of its administration is independent of the client's reading level (Archer et al., 1991). Shneidman's (1951) edited volume presented the conclusions of 16 clinicians who blindly interpreted the same TAT protocol using their own preferred method. Despite different approaches and foci, they were remarkably accurate and consistent with each other. Also included in this volume were Klopfer's conclusions based on a blind interpretation of the same individual's Rorschach record which was highly consistent with those based on the TAT protocols. It should be noted that the interpreters all had one thing in common—they were "experts." In light of what we know about how experts perform (e.g., Chi, Glaser, & Farr, 1987), we would not expect to replicate such findings with more recently trained individuals.

In evaluating the evidence of reliability and validity in the clinical use of the TAT, the training and experience of interpreters was a primary consideration (Karon, 1981). Research designs that isolate response variables and subject them to separate statistical analyses do not mirror the interpretation of projective methods in clinical practice. The clinician simultaneously evaluates multiple dimensions of the response in reference to a psychological construct rather than adding individual response variables. More sophisticated validation efforts are needed that focus-on the integrative approach rather than piecemeal summation of isolated response characteristics (e.g., Handler & Habenicht, 1994). Finally, more refined approaches to validation suggest that different facets of psychological constructs are measured by different techniques. For instance, a comprehensive meta-analysis indicated that projective measures of achievement motivation predict long term behavioral trends, whereas self-report measures predict immediate choices (Spangler, 1992). Refinements in conceptualization and approaches to validation are calling into question earlier criticism based on psychometric shortcomings.

4.16.3 BASIC ASSUMPTIONS OF PROJECTIVE TECHNIQUES AND RESEARCH FROM OTHER PSYCHOLOGY SUBFIELDS

The term "projective technique" was coined by Frank (1939) to describe tests using ambiguous stimuli and/or tasks that are less obvious in their intent and are therefore less subject to faking. The aim of any projective technique is to set a task that allows the individual to express characteristic ways of perceiving and organizing experiences. The uniqueness of the individual's style of responding is particularly evident when stimuli are ambiguous and there is no ready response. Under such conditions, there are many possible ways to approach the task or situation, and the person must actively organize the response. A central assumption of all projective methods is that stimuli from the environment are perceived and organized by the individual's specific needs, motives, feelings, perceptual sets, and cognitive structures, and that in large part this process occurs automatically and outside of awareness (Frank, 1948).

These assumptions about projective testing, first elaborated by Frank (1939) in his analysis of the projective hypothesis, are compatible with converging trends across various subfields in psychology pointing to the role of previously organized inner structures or mental sets in the interpretation of stimuli. Since the cognitive revolution of the 1960s, psychology has undergone a significant transformation by moving away from an emphasis on stimulus–response units to describe human perception and action to increasing emphasis on how humans impose meaning and organization on their experiences (Singer & Salovey, 1991). Within cognitive psychology, simple associational models (e.g., paired associate learning) have become less important in describing ongoing information processing than the organized meaning structures, such as schema or scripts that are brought to bear on the interpretation of experiences (Piaget, 1954; Taylor & Crocker, 1981). At the same time, developments in social psychology (Abelson, 1976; Festinger, 1957; Heider, 1958; Kelley, 1967) led to the exploration of organized meaning structures as the fundamental bases of social action (Wyer & Srull, 1994). In the psychodynamic paradigm, the emergence of object relations theory (Blatt & Wild, 1976) implicitly relies on internal structures similar to schema to guide information processing about the self and other. The considerable work in cognitive, social, personality, and clinical psychology bearing on schema theory and supporting the projective hypothesis is briefly highlighted next.

4.16.3.1 Schema Theory and Projective Techniques

Earlier cognitive work within clinical psychology focused on the role of self-statements and inner thoughts. But clinical scientists interested in the role of cognition have increasingly adopted an information-processing approach to clinical phenomena (Ingram, 1986). This approach focuses on the schema that provide rules guiding behavior in social relationships and appear to influence how information about relationships is stored in memory (Fiske, Haslam, & Fiske, 1991). Well-learned, organized knowledge structures (schema or scripts) have considerable impact on memory storage and retrieval. Projective tasks require the superimposition of previously acquired schematic structures to the perception of stimuli and the organization of the response. Acklin (1994) presented a reformulation of the responses to the Rorschach on the basis of schema theory and information processing. Within this framework, emphasis is placed on the nature and adaptiveness of the schema activated by the stimulus properties.

Memory structures such as schema or scripts that store information about situations, persons, and events guide the interpretation of experiences by providing criteria for regulating attention to lend focus to the process of encoding, storage, and retrieval of information in specific domains (Taylor & Crocker, 1981). Event schema refer to the type of script that organizes understanding of a sequence of events in a stereotyped or routine situation such as ordering dinner in a restaurant (Abelson, 1981). Such a "script is a set of expectations about what will happen next in a well understood situation" (Schank, 1990, p. 7). As long as we know what script others are following, we know how to act and to predict the actions of others. Upon entering a restaurant, we are seated, order from the menu, receive our meal, pay, and depart. This process of comparing experience with the internal schema enables us to perceive rapidly unstated information and to anticipate what will happen next. Some individuals may view situations for which they have no scripts (novel) as challenging, whereas others are stymied (e.g., "speechless") in such conditions.

Person schema include information about the self, others, and their interactions. They deal with the individual's rules for predicting, interpreting, responding to, and controlling affectively charged encounters (Tomkins, 1979). For Tomkins (1979, 1987, 1991), emotion is a key organizer of personal scripts. Once the script has been formed, the script organizes and modifies new experiences to fit the preexisting structure. Person schema develop through the individual's synthesis of past experiences according to the individual's style of processing information (Horowitz, 1991). Therefore the construct of person schema brings together models of perception, cognition, memory, affect, action, and feedback. Furthermore, it has the potential to integrate psychodynamic formulations and cognitive perspectives in understanding personality (Horowitz, 1991; Stein & Young, 1992). This process of interpreting new information according to previously acquired sets takes place without conscious awareness and may also lead to systematic distortions in perception, interpretation, and action in interpersonal encounters.

Causal schema give rise to attributions about the causes of events. Schema theory and attribution theory overlap in that attribution theory maintains that individuals are motivated to make causal inferences about experiences and that attributions are made in ways that are congruent with existing schema of self and other, and of the assumed relationships between causes and effects (Tunis, 1991). The relationship between such causal attributions and depression has been studied.

Evidence for the existence of schema is seen in their consequences for information processing, particularly in how information is remembered and retrieved. Schema-driven information processing involves the application of organized knowledge structures akin to theories that govern the perception and interpretation of facts. Thus, schema function to allow general knowledge to influence perception of specific experiences. Schema increase efficiency in identifying perceptions, organizing them into manageable units, filling in missing information, and selecting a strategy for how to obtain further information if needed. Because schema guide behavior in adaptive and maladaptive ways, their study pertains to clinical psychology as well as to the general study of personality, cognition, and social perception. Understanding the recurrence of maladaptive interpersonal patterns is facilitated by studying the operation of schema along with the realistic properties of the current situation (Horowitz, 1991).

The overlap between schema or script theory and the projective hypothesis is evident in their common features. The development of schema or scripts and their retrieval from memory represent an unconscious process through which past perceptions influence the interpretation of current situations. These are precisely the processes that are assessed with projective techniques. The focus of projective methods is on the application of schema to the task demands (e.g., how schema influence the response process) and, by extension, to life challenges requiring similar adaptations.

Even though schema are structures of the mind, they must be sufficiently malleable to adapt to new situations and new configurations of events (Rummelhart, Smolensbus, McClelland, & Hinton, 1986). Therefore, it is appropriate to view schema as emerging when they are needed "from the interaction of large numbers of much simpler elements all working in concert with one another" (p. 20). A distinction has been made between enduring schema and working models (Horowitz, 1991). The former are intrapsychic meaning structures containing generalized formats of knowledge that can be activated by other mental activities related to that knowledge such as motivational concerns (Fiske & Taylor, 1984). Working models combine internal and external sources of information, such as when the individual is actively contemplating an interpersonal situation or task such as the TAT. Working models actively integrate stimuli from a current situation with past knowledge through triggering of an associative network of ideas drawn from enduring person schema.

A person's repertoire may include several different enduring person schema in relation to a given type of relationship, situation, or activity. The working model may incorporate different elements from each of the person schema. The wider the repertoire of enduring person schema, the more flexibility in constructing working models of interpersonal situations. Discrepancies between working models and active, enduring schema may contribute to upsetting emotional experiences. Working models may not match the actual qualities of current social situations, leading to errors in judgment and behavior and to subsequent negative emotions. The aim of many psychotherapeutic techniques is to promote conscious awareness of unconscious, schematic functions. Such awareness may permit the individual to actively override some of the influence of the unconscious information processing.

Scripts or schema may be further organized into metascripts (Singer & Salovey, 1991) that reflect the individual's style of dealing with scripts. These superordinate schema may explain resilience because they influence how individuals reshape their schema when confronted with daily stress, unforeseen failure, or unexpected upheaval. Those with more complex schema may be more willing to replay negative information to explore a variety of alternative schema or scripts. More flexible schema involve the realization that current ideas that seem so clear and self-evident at the moment may change as a function of new experience. This process cannot be taught directly, but is a byproduct of experiencing changes in perception, understanding, and feeling in light of new information. Consequently, it is important to assess the style of synthesizing experience through the organization and reorganization of schema.

Analysis of responses elicited by projective techniques can be informed by models of schema development. Abelson (1976) assumes that schema begin as representations of single concrete examples and become more abstract. Rudimentary schema are compilations of single, concrete examples that are used to make snap judgments about seemingly similar instances. Stereotypic schema encompass the most representative features of events, persons, or groups. More abstract schema recognize inconsistencies between reality and the activated schema. The individual is aware of complexities and ambiguities rather than simply encoding schema–congruent information. Highly developed schema permit the individual to engage in complex information processing seemingly effortlessly and without awareness. The more abstract the schema, the more flexibly they can be applied because they include conditional and inferential concepts, abstract rules, and affective information. The progression from rudimentary to more complex schema is the product of experience and the degree of active, strategic effort brought to the structuring and organization of experience. Schema theory recognizes the role of affect, motivation, and biological factors (genetics, temperament), as well as environmental influences (stress and supports) as they interact with cognitive processes. The shaping of schema, especially social scripts by culture, is also acknowledged and has been studied by anthropologists and linguists (Quinn & Holland, 1987).

Responses to projective methods may be evaluated according to aspects of schema described above. When telling stories to picture stimuli, some individuals superimpose associative elements without an organizational network; others impose stereotypes that may or may not precisely fit the stimulus configurations; still others draw creatively on various elements of their experiences to tell a cohesive story that captures the gist of the stimulus and complies with the instructions. If an individual is presented with a TAT picture that cannot be readily explained by a stereotypic script or an intact schema, then the respondent is called upon to actively construct a script to fit the situation. Similarly, when faced with inkblots that are rough approximations of real objects, the individual must remain flexible in applying schema to answer the question "What might this be." Those who rely on more rudimentary schema will make associations directly from discrete portions of the stimuli. However, those who apply more cohesive schema find a meaningful framework with which to integrate disparate or complex stimuli.

Two aspects of long-term memory structures (schema), procedural and declarative knowledge (Anderson, 1983; Kihlstrom, 1984), are relevant to the interpretation of projective responses. Procedural knowledge refers to unconscious processes or skills such as the structure of language, the organization of music, or other implicit rules that order information or perception. Declarative schematic structures entail the recall of factual information such as names, locations, and historical events. Responses to projective techniques reveal the implicit organization of knowledge as well as the content that enters awareness in response to the stimuli presented. Interpretive approaches can focus on the organization of schema by examining the structural aspects of responses and the sequence of ideas expressed, as well as by analyzing the content.

4.16.3.2 Unconscious Processing of Information

The idea that human behavior is based on a considerable storehouse of organized knowledge structures operating outside of conscious awareness is now widely accepted by cognitive psychologists (Bargh, 1994). Since the 1970s, increasing evidence suggests that information not accessible to conscious awareness influences memory, perception, and thinking (Kihlstrom, 1990). Efforts to refine assumptions about unconscious processes continue across the various subfields of psychology and are not restricted to projective testing or even to clinical psychology.

Social cognitive psychology (social cognition) has become increasingly interested in unconscious thought processes of mental representations of situations, self, and others (schema) and in the interaction of cognition and affect (Wyer & Srull, 1994). Attitudes, expectations, or schema that are strong enough to be automatically activated have been described in terms of their "chronic accessibility" (Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Chronically accessible schema may be cued by emotions arising in ambiguous situations, including reactions to projective stimuli. According to Westen (1993), "Research on chronic schema accessibility has confirmed the projective hypothesis that in ambiguous situations enduring interests, concerns, needs, and ways of experiencing reality are likely to be expressed" (p. 381). Bargh (1994) suggests that temporarily accessible schema due to a recent experience can mimic chronic accessibility and constitute a pitfall for studies of enduring cognitive structures. Indeed, this pitfall is common in the use of personality tests, be they questionnaire or projective devices. The distinction between state and trait variables acknowledges states as temporary conditions of mentality, mood, levels of arousal or drive, and these states may be related to temporarily evoked schema. One of the advantages of projective measures is the opportunity to examine both the structure and content of schema even if temporarily evoked. The content of the schema may be more susceptible to influence by such variations in experience than the structure. However, data are needed to support this contention.

A dramatic demonstration of the effect of a recent experience occurred one semester when the author's students administered assessment batteries to incarcerated youths (generally between the ages of 18 and 22) to determine educational strategies. A great number of TAT stories ended by the characters "talking about it and solving their problem." After some investigation, it was learned that a communication program had been recently instituted to teach the value of resolving conflict by talking things out. A close inspection of the stories themselves revealed that the schema about communication were not well developed but verbalized as a stock ending apart from other components of the problem-solving process. For example, the problems were poorly understood, and characters' feelings, perspectives, and intentions were vaguely or inaccurately deciphered, if at all. The idea for talking about the problem became accessible to conscious awareness through training as declarative knowledge but without the understandings or procedural knowledge needed for its implementation. Conscious awareness of the benefits of communication was not combined with the implicit and largely unconscious understandings and affective reactions that promote such interactions. This example illustrates the importance of studying the inner logic and cohesiveness of schematic structures to discover how story content derived from recent experiences (including vicarious ones such as books or movies) is integrated with other aspects of the task demand such as accurately perceiving the stimuli and following instructions. Both declarative and procedural knowledge must be addressed to overcome the pitfall of attributing too much weight to content that is not meaningfully incorporated into guiding structures.

4.16.3.3 Inert Knowledge vs. Usable Knowledge

Inert knowledge has been described as information that the person knows, but does not apply unless explicitly cued or prompted to do so (Bransford, Franks, Vye, & Sherwood, 1989). The reason why a person fails to use relevant knowledge to solve particular problems is that this information does not spontaneously come to awareness in the context of that problem, or that the schema for applying the information are not sufficiently developed. The manner in which knowledge is encoded and organized in the first place determines its subsequent accessibility and applicability in other situations.

Effective learning in general requires active strategies to organize and recall information such as rehearsal or breaking down the task into manageable units. Torgesen (1977), who introduced the concept of the inactive learner, reported that learning-disabled students exhibit a learning style characterized by minimal planning and limited use of strategies. Individuals may process information from one domain more thoroughly than in others, or the inactive stance may be more pervasive. This inactive style is manifested by processing information exactly as it comes in (memorization) without organizing and restructuring it. In the absence of active, effortful, strategic processing of information, the knowledge may be dormant in memory and not available for use unless externally prompted.

Different types of tasks are needed to differentiate knowledge that is used spontaneously to solve problems from knowledge that is available only when cued by the context. Everyday behavior requires spontaneous use of prior knowledge to make social judgments or other decisions; the possession of factual knowledge is far removed from problem solving in everyday encounters. The TAT stories of the incarcerated youths referred to earlier shows how knowledge may be stated without being embedded in a usable framework. Projective measures differ from typical cognitive tasks and self-report measures of personality in ways that have relevance for distinguishing between inert and useful knowledge. Rather than being asked to provide specific information about the self, the respondent is provided with a relatively unstructured task and asked to give an open-ended response that is evaluated by a professional. Likewise, solving verbally communicated problems such as those on intelligence tests does not resemble the conditions of everyday situations. Such problems tend to contain words that act as cues for accessing relevant knowledge, whereas everyday problem-solving requires the individual to size up relevant features of the situation to notice that there is a problem (Sternberg, 1985). Real problems are accompanied by emotions that shape the way the individual thinks about the problem. Individuals may verbalize several acceptable alternatives for handling a situation, if prompted. Nevertheless, when faced with real-life dilemmas, they may be unable to translate this knowledge into appropriate action. The projective tests reveal the spontaneous accessibility of the schema that guide the application of what comes to awareness.

4.16.3.4 Affective and Motivational Influences on Cognition

Most studies in this area focus on the way mood or emotion influences memory and judgment (Isen, 1984, 1987) and on unconscious processing of information (Uleman & Bargh, 1989). Epstein (1994) describes the impact of emotion on thinking and argues for the existence of two fundamentally different but interactive modes of information processing: a rational system and an experiential system driven by emotion. Thinking in the experiential system is shaped by emotion and tends to occur without deliberate effort or conscious awareness. The rational system is characterized by a more deliberate process of thinking and of acquiring information such as through textbooks and direct teaching. Epstein explains that stories and anecdotes spice up an otherwise dry lecture because they increase emotional engagement and thereby appeal to the experiential system.

Emotions dramatically alter the thought process and influence behavior. Thinking is transformed by intense emotion in the direction of being categorical, personal, concrete, unreflective, and action oriented. The product of such thinking is likely to be considered as self-evident and not requiring proof (Epstein, 1994). Depressed affect is associated with preference for immediate gratification in place of delayed but more substantial rewards (Wertheim & Schwartz, 1983). Anxiety tends to bias the interpretation of ambiguous text (MacLeod & Cohen, 1993). Chronically experienced negative emotions appear to have a cumulative influence on the development of cognitive structures amenable to assessment with projective techniques. Children temperamentally prone to experience negative affect told stories to TAT stimuli that reflected less complex information processing and more reactive self-regulatory styles dominated by more immediate concerns than their less emotional peers (Bassan-Diamond, Teglasi, & Schmitt, 1995).

The following vignette shows how intricately problem solving is tied to emotions (*Readers Digest*, May, 1982, p. 79. Quoted from Jim Whitehead, quoted by Seymour Rosenberg in Spartanburg, SC *Herald*):

> Two hikers were walking through the woods when they suddenly confronted a giant bear. Immediately, one of the men took off his boots, pulled out a pair of track shoes and began putting them on. "What are you doing?" cried his companion. "We can't outrun that bear, even with jogging shoes." "Who cares about the bear?" the first hiker replied. "All I have to worry about is outrunning you."

Now what if the two were husband and wife, parent and child, or very close friends? The emotional investment in the relationship would change the problem to how can "we" escape or even how can "you" reach safety. A parent may suggest that the child run for it. The automatic influence of affect in defining the range of acceptable solutions to problems is evident. Westen (1993) observes that cognitive psychologists do not focus sufficiently on the motivational and emotional factors of unconscious processing of information. Affect and motivation can enhance or disrupt the degree of organization and resourcefulness in the acquisition and subsequent application of knowledge structures.

4.16.3.5 Narrative Psychology and Thematic Apperception Techniques

The analysis of narrative has been a rich source of data in studying the schema or scripts that individuals bring to current understanding of experiences. The social-cognitive approach examines the memory system by analyzing the individual's style of processing and organizing information. Memories such as self-defining experiences (Moffit & Singer, 1994) or incidents of being rejected (Baumeister, Wotman, & Stillwell, 1993) are examined to extract abstract rules and cognitive principles. A study on script formulation (Demorest & Alexander, 1992) first asked participants to generate autobiographical memories. Scripts extracted from these memories were subsequently (one month later) compared with those derived from participants' invented stories in response to affective stimuli. The scripts drawn from fictional stories were similar to those derived from autobiographical memories. These findings suggest that scripted knowledge structures are superimposed on new affective stimuli.

Schank (1990) addressed the question of how stories exchanged during the course of social conversation are accessed in memory. To understand how a person is reminded of a story to tell at the right moment in a conversation requires breaking it down into themes, plans, goals, actions, and outcomes (Schank, 1990). These elements constitute a filing system for indexing the relevant components of experiences to be stored in memory and subsequently retrieved. A higher level organization of these concepts involves a prediction, a moral, or a lesson learned. When a lesson repeatedly occurs, it becomes a type of structure that exists apart from the specific stories from which the lesson arose.

For Schank, telling a good story at the right time is a hallmark of intelligence because it represents an understanding of what is being talked about now through its connection with the lessons that the listener has stored in memory. Recalling an appropriate story to tell and understanding how it relates to what is currently being spoken about depends on how the stories are catalogued in memory. It seems logical that individuals who engage in more active and reflective classification of experience for storage into memory are also more likely to be reminded of an appropriate story at the right moment in a conversation. This framework is directly applicable to the interpretation of stories told to TAT-like stimuli. Respondents are reminded of a story by the stimulus, but the particular story details are shaped by convictions or lessons learned from experience.

One way to validate constructs is through the congruence of models drawn from various lines of research. Arnold (1962) proposed that individuals express their basic attitudes or convictions in their TAT stories, and that these convictions constitute their motivational set. Like Schank, she focused on goals, plans, actions, and outcomes as the basic ingredients of the story import or lesson learned.

The rise of narrative psychology has the potential to enrich story-telling methods of personality assessment. Analysis of narratives such as those elicited by the TAT is considered useful for clarifying the degree of complexity and organization of information processing pertaining to the synthesis of experience (Carlson & Carlson, 1984). These cognitive and perceptual processes can also be assessed through the Rorschach technique. It has been suggested that the Rorschach should combine its emphasis on perception with a focus on cognitive representation (Blatt, 1992).

4.16.4 LEVELS OF PERSONALITY ASSESSED BY PROJECTIVE AND QUESTIONNAIRE METHODS

Projective methods measure different dimensions or levels of personality than self-reports. One of the clearest ways to understand the unique contribution of projective techniques is in the context of understanding these levels. McAdams (1995) proposed three conceptually distinct levels of understanding personality: level one—traits or stylistic, habitual tendencies; level two—developmental or motivational constructs such as goals, plans, and strivings; and level three—the life narrative or the evolving story that is internalized to provide broader meaning, purpose, and cohesiveness to specific experiences. McAdams does not describe relationships among the levels, though it seems logical to assume that they are not independent since they are different perspectives (or units) on the functional whole. For Klopfer (1981), a trait itself can be understood at three levels. First, as viewed by significant others or the public image (with or without awareness); second, as viewed by the individual or conscious self-concept; and third as manifested (with or without awareness) in behaviors such as responses to projective tests. This emphasis on the source of the data is consistent with a fundamental distinction between two levels of personality, based on the perspectives of the actor and observer (Hogan, 1987). Projective tests were better able to predict peer rankings on four personality traits (publicly observed image) than the individual's self-report of those traits (McGreevy, 1962). Different techniques may be useful for elucidating different levels or dimensions of personality. However, a comprehensive understanding requires a framework for how the different levels relate to one another.

Measuring achievement motives through self-report of intentions, goals and reasons for actions (self-attributed) and inferences drawn from stories written to picture stimuli (implicit) are not equivalent (Koestner, Weinberger, & McClelland, 1991; McClelland, Koestner, & Weinberger, 1989). McClelland and colleagues propose that the general absence of significant correlations between measures of self-attributed motives derived through self-report, and implicit motives inferred from stories, cannot be ascribed to the worthlessness of projective measures nor to poorly designed questionnaires, but must be taken seriously as evidence that these are in fact different variables with the same name. Rather than seeking to remedy deficits in either type of measure, they suggest that it may be more fruitful to acknowledge them as reflecting two qualitatively different kinds of human motivation. They contend that implicit motives develop by intrinsic enjoyment generated when doing tasks or experiencing activities, and predict spontaneous goal-directed actions sustained over time even in the absence of specific social demands. However, self-attributed motives are built around explicit social incentives or demands of socializing others, and predict responses to situations structured to provide such incentives. The implicit achievement motivation does not give information about the area of life to which a person will direct efforts to succeed. Self-attributed motives, plans, and goals may express the person's conscious intentions but do not give information about the person's commitment and capacity to follow through. Under specific conditions, self-attributed motives act like implicit motives in that they energize, direct, and select behavior. However, the problem with predicting long-term behaviors from self-attributed motives is that the social incentives may not be salient enough to elicit the behavior. Implicit motives exert relatively greater influence because they drive activity that is inherently enjoyable even in the absence of specific social demands.

The acquisition of implicit and self-attributed motives such as proposed by McClelland and colleagues may relate, respectively, to the experiential and rational systems discussed earlier (Epstein, 1994). Learning in the experiential system takes place in an affective context where emotions and interests drive attention and information processing. Additionally, the development of these two types of motives may involve different principles of learning (Meissner, 1974, 1981; Raynor & McFarlin, 1986; Sandler & Rosenblatt, 1962). One pertains to the acquisition of cognitive skills that are functional and enlarge the individual's repertoire of adaptive capacities through the teaching and reinforcement of these competencies. The other relates to the development of the inner world or the self-system that provides relative independence from external incentives. The advantage of developing this inner structure is to shift from dependence on external supports to increased reliance on self-regulatory mechanisms (Meisner, 1981). These regulatory functions increase the individual's capacity to master the environment, endure stress, and delay gratification of impulses.

The development of increasingly complex and differentiated self-regulatory structures or schema takes place at a level of organization that is different from that of acquiring behavioral patterns through reinforcement. Such development relies on the individual's integration of life experience to seek consistency within the self (Raynor & McFarlin, 1986). However, emotional, cognitive, and attentional limitations interfere with synthesis of life experiences and detract from the development of inner self-regulatory structures. For example, the domination of awareness by that which is immediately striking or concrete is not compatible with the development of long range interests, values, or aims. The primary affective impediment to the internalization of such self-regulatory values are the limitations in the capacity to sustain emotional involvement in interests that are relatively remote from personal needs (Shapiro, 1965).

Motivation can be understood in relation to two necessary ingredients: the goal or intention and the self-regulatory capacity to sustain goal directed action (Arnold, 1962; Kuhl & Beckmann, 1994). Thus, a stated goal or intention such as playing violin in a symphony orchestra may be difficult to implement and sustain, regardless of musical talent because of the rigorous independent practice required. Such sustained effort is facilitated by the enjoyment of and interest in the activity. Otherwise, careful programming of external incentives is needed.

Understanding the gaps between a person's stated goals or intentions and the individual's self-regulative resources is crucial to planning interventions. A 19-year-old college student was referred by her therapist for an evaluation because no meaningful discussions were taking place after six months of regular attendance of sessions. Sandra was in therapy at the insistence of her family because "she had no direction in life." During the initial interview with Sandra, she described her erratic schedule of eating and sleeping, her general lethargy and boredom, as well as sporadic class attendance. She reported having similar problems in high school such as difficulty getting up, missing school once a week, not studying, and cheating to get by. She stated emphatically that she was not depressed.

When asked directly about drug use, Sandra reported significant involvement starting in the fifth grade. She did not discuss this issue with her previous therapist because "no one asked me." She described the therapist as not providing enough structure during the sessions so she just talked about meaningless topics. Sandra verbalized her desire to achieve and recognized that she was not accomplishing her goals. At the same time, she admitted that she found school work aversive.

As part of the evaluation, the TAT, the Wechster Adult Intelligence Scale (WAIS-R), and Rorschach were administered among other measures. The first five TAT stories that she told are given below to illustrate the lack of connection between Sandra's goal setting and means toward their attainment. The emphasis is not on developing a comprehensive clinical picture but on presenting general conclusions about each story that are pertinent to Sandra's motivational schema:

> *Card 1.* Don't like pictures. Not all of them . . . Umm . . . like story—like once upon a time. This is a boy . . . He's . . . OK, he . . . is it a violin? I don't know. Can I say two things? (Whatever you want). He's either frustrated because he can't play it or he's sad because he's not allowed to play it anymore. He wants to play it but for some reason he can't . . . (stares at picture) (TO) He becomes a great (did I say that was a violin?) violin player, and he's happy. (OK, I think you have the idea.)

Sandra is immediately put off by the task and vacillates in her interpretation of the picture as she does in her own intentions. The story content is appropriate to the stimulus, but Sandra cannot determine whether the boy is experiencing internal (frustration) or external (not allowed) barriers to achievement. In any case, the boy takes no heed of these problems and becomes a great violin player, seemingly without effort. This unrealistic connection among circumstances, intentions, actions, and outcomes suggests that Sandra's schema do not include a clear grasp of the obstacles to achievement nor of the steps needed to attain goals.

> *Card 2.* God! These things are … OK, there's … (stares) … This girl on her way to school, she's walking to school, and she sees this man working in the field, and that's his wife watching him. His wife's talking to him I guess, may be talking to him. He's working with his horse on the farm, and she's thinking that she doesn't want to look like that when she gets older (like the lady). She doesn't want to live on the farm. She feels sorry for that lady, for both, for the lady. So she goes to school to get an education so she's not like them.

The girl goes to school not because she really wants to learn but to avoid becoming what she fears. Sandra's schema about the importance of education does not include a positive affective quality nor the spontaneous accessibility of the effort involved. Sandra does not delve into the purposes and inner life of the characters but relies on the stimulus, on external appearances (doesn't want to look like the lady), and social convention (getting an education). Thus external structures and supports are important in guiding her reactions.

> *Card 3BM.* Don't know what this is … umm … guess it's a girl. A boy or girl, either sleeping, crying, or … umm … (frustrated) don't even know what. (repeat directions) I guess she's crying because … OK, this doesn't make sense though. (stares as if answer is in the card) I think they're car keys. No idea. This is dumb. She's crying because … what is that thing is … someone stole her car. She has no way home from this place. So then she calls the police and tells them that her car is missing. Eventually they find her car, and then she's happy. Can you tell me what that is? (It can look like different things to different people.)

Sandra initially vacillates before settling on a plot. She became frustrated, stated that the task was dumb, but responded to encouragement. This tendency to blame outside factors is consistent with her reliance on the environment to regulate her behavior. The outcome of getting the car back only deals with part of the problem. Having "no way home from this place" was not addressed. Sandra's incomplete processing of information (only dealing with part of the problem, leaving out details called for in the directions) and limited resources (no internal representation of possible support; indecision and disjointed approach to narrating the story) suggest why she cannot regulate her behavior.

> *Card 4.* OK, … this is a man and a woman who … they either just had a fight, or he wants to go somewhere, and she doesn't want him to go. OK, she's trying to persuade him either not to go or to forgive her. And then, I guess eventually he does it, he forgives her or he doesn't—or he might go, or he'll stay. And they live happily ever after.

Again, Sandra cannot commit to one story line, and the characters' intentions or reasons for acting are not examined. As with the previous stories, she seems unable to develop her ideas clearly beyond the initial identification of the tension depicted in the card, suggesting dependence on external stimuli. Her vacillation about what's going on in the picture (noted in several picture cards) suggests that she feels uncertain about her judgment in situations that are somewhat ambiguous. As in previous stories, the ending is unrealistically positive as Sandra does not make appropriate connections among intentions, actions, and outcomes.

> *Card 5.* OK, can you make up people who aren't there? This is a lady who is coming to tell her children it's time for their nap. So she sticks her head in and says it's dinner time, and all the kids come to dinner. Everybody eats dinner, and kids go back and play. In the library, den—I think it's supposed to be the library.

Sandra's query about introducing additional characters further points to her need for external guidance. The lady intends to tell the children that it's time to take their nap but ends up calling them to dinner. This twist in the story suggests that Sandra is not monitoring story details (also seen in other stories) and that she has similar difficulty keeping track of her own intentions.

Generally, Sandra is able to recognize the conflict or tension in the pictures but has difficulty sticking to one explanation. Sometimes, she does not monitor story details sufficiently, leaving gaps or contradictory details. She does not examine feelings, thoughts, or intentions and has difficulty connecting planful, realistic actions with outcomes. She cannot develop the themes that she introduces beyond a vague and Polyanna-like story. Sandra persists in staying in school because she believes that the way to ensure her future is to get an education but is not able to get invested in the process of learning. Sandra is floundering in the unstructured campus setting because she has not developed the inner structures (implicit motives) to regulate her attention and behavior.

When she received the results of the evaluation, Sandra easily identified with the example of someone who thinks an exercise program is needed and joins an aerobics class but does not attend for one reason or another (too tired, too cold). She agreed that even the thought of doing school work is aversive, and the key was in finding ways to side-step the aversive elements, possibly through social supports. She acknowledged that by joining a sorority, she would have peers to eat with on a regular basis, and the structure would help her keep a schedule and attend class. At first, she worked on seeking the external incentives that would motivate her behavior and through these initiatives began to feel she had some choices and control over her life. Subsequently, therapy sessions focused on ways to build these structures and prompts into her daily routines to gain increasing independence.

Sandra's schema as assessed with the TAT did not include connections among goals, plans, actions, and outcomes, and these resources were not available to guide her daily behavior. The therapeutic aim was to help Sandra deal with the factors that impeded the development of these connections. She could not have verbalized these impediments in an interview or self-report measure because she was not aware of them. Likewise, the more structured Wechsler scales did not reveal the sources of her difficulty. Therefore, sole reliance on such measures would have been insufficient.

4.16.5 SELF-REPORT VS. PROJECTIVE PERSONALITY TESTING

The foregoing discussion has already demonstrated that self-reports, even if candid and accurate, may predict behavior only when social incentives are operating or when the situation is structured to prompt the response. However, it is useful to contrast questionnaire and projective devices as direct and indirect methods of personality assessment (Vane & Guarnacia, 1989). Direct methods are comprised of observations of behavior, inventories, or interviews. The indirect method utilizes relatively unstructured stimuli such as the TAT, Rorschach, or drawings. In the latter, the information sought is not readily apparent and, therefore, difficult to manipulate. Vane and Guarnacia cite the following problems with the direct method: (i) limitations of self-knowledge preclude accurate self reports; (ii) test items are open to misinterpretation; (iii) real-life situations cannot be represented by pencil and paper items; (iv) desire of the respondent to manage impressions need to be overcome by including scales to detect faking or response sets; (v) inventories yield information without clues for understanding the reason; and (vi) the relationship between traits on inventories and behaviors is complex

Multiple choice or other highly constrained response formats do not give clues about the style of organizing the response. The respondent is asked to tell versus show. Thus, no evidence exists beyond self-report to judge response parameters despite availability of norms. Finally, responses to structured questions only tap conscious phenomena. Therefore, methods such as interviews or self-report inventories provide what the individual is capable of perceiving and is willing to share. However, responses to projective testing can be evaluated by analysis of various parameters.

Both direct and indirect sources of information are useful depending on the purpose. Self-reports are susceptible to conscious or unconscious distortion and reveal only one dimension of the personality. Nevertheless, it is important to obtain information that the respondent willingly discloses. Self-reports are easily scored and amenable to traditional methods for estimating reliability and validity. Projective methods require extensive training, are time consuming, and require more cumbersome methods to establish psychometric credibility. However, their purpose is not evident to the respondent, and they provide information that is different from other sources of data. The use of both direct and indirect methods is likely to yield the most accurate picture.

4.16.6 CONTRIBUTION OF PROJECTIVE TECHNIQUES TO DIAGNOSIS AND INTERVENTION

The process of responding to the projective task can be conceptualized as an attempt to impose meaning on the stimuli presented and to comply with the instructions given. Analysis of the response reveals the individual's schema or scripts, both in terms of the structural organization of experience and the content of awareness. Appreciation that the manner in which individuals acquire, organize, and represent knowledge into cognitive schema have a fundamental influence on the way they behave, guides the use of projective techniques. Rather than looking for one-to-one correspondence between test responses and inferences, such as searching for aggressive content in projective tests, it is more fruitful to explore processes that promote aggression in particular situations. The social-information-processing model of children's social behavior (Crick &

Dodge, 1994) details the social-information-processing steps that precede aggressive behavior in specific situations. According to this model, aggressive behavior can be understood in terms of how individuals encode social cues, how they interpret those cues, how they select and clarify goals, how they generate possible solutions to perceived problems or dilemmas, how they make decisions about the selection of responses, and how they execute and monitor the selected behavior. Information regarding each of these steps in social information processing can be gleaned from responses to projective techniques.

Thematic apperceptive methods can provide information about the attribution of intentions for others' reactions, the interpretation of social cues, the anticipated consequences of alternative actions, as well as about the cognitive flexibility and organization of the individual's meaning structures. Information about the execution of the behaviors is provided at two levels: how the characters follow through on intentions and resolve problems, and how the narrator develops and monitors the story. Any aggressive content that is expressed can then be better understood in light of the aforementioned processes.

Responses to the Rorschach also clarify social information processing that promotes aggression. For example, a preponderance of answers that simplifies the stimulus (e.g., high lambda) suggests that the individual may be involved in confrontational situations because of a tendency to make decisions without considering important cues (Exner, 1993). The connection between such responses to projective tasks and aggressive behavior is borne out by research suggesting that aggressive children selectively attend to and recall aggressive cues, partly because they respond prior to processing all of the available information (Dodge & Feldman, 1990).

A taxonomy of psychopathology comprised of a descriptive, atheoretical compilation of symptom categories (e.g., *Diagnostic and statistical manual of mental disorders*, 4th ed. [*DSM-IV*]) could be enhanced by a framework that relates these symptoms to basic structures of psychological organization. Within a diagnostic category such as autism, there exists a wide variability of impairment in functional capacities. Responses to projective techniques can range on a continuum from being flexibly adaptive to demonstrating various levels of impairment that can go beyond a specific diagnosis. The manner in which an individual uses the cues in the stimuli and responds in an organized way to the task demands is an index of the differentiation and organization of previously developed schema. Furthermore, responses reveal whether the person must be prompted to respond or actively restructures knowledge for adaptive use in stressful or ambiguous situations.

Horowitz (1991) envisions a future *DSM* Axis II as incorporating three interconnected components: (i) recurrent maladaptive patterns of self-regard and interpersonal behavior; (ii) person schematic characteristics explaining that behavior—this would include the developmental level of self-organization; and (iii) style of self-regulation comprised of habitual control processes leading to coping and/or defense. An example would be how a person regulates emotionality by inhibiting or facilitating the activation of schema. Horowitz's conceptualization implicitly recognizes that any Axis I diagnosis, such as attention-deficit hyperactivity disorder relates to the development of inner schema. Difficulty regulating the attentional process may interfere with the active and strategic synthesis of knowledge gained through experience. Without organized schematic structures, the individual may not be able to govern behavior according to internalized rules and standards, a prominent characteristic of individuals with attentional deficits (Barkley, 1990). Responses to open-ended projective tasks represent the convergence of all of the individual's traits on the organization of experience, including compensatory strategies that could mitigate against risk factors.

A changing emphasis within psychoanalytic theory favors the view of psychopathology as representing an impairment in the formation of psychic structures that are analogous to the schema of social and cognitive psychology. This focus has spurred diagnostic efforts to evaluate the quality of these structures (self-system, object relations). Blatt (1991) assumed that the symptomatic manifestations of various forms of psychopathology are associated with different types of impairments of cognitive-affective structures in the representational world. For example, the growing consensus about the existence of two basic subtypes of depression (empty and guilty) is based on the phenomenology of experience associated with each. One type involves exaggerated preoccupation with issues of interpersonal relatedness, feelings of depletion, dependency, helplessness, or loss. The second type occurring later developmentally is associated with issues of self-definition, autonomy, guilt, and feelings of failure. The two types of depressed individuals change in various ways in the treatment process and are differentially responsive to different forms of therapy (Blatt et al., 1988). Wilson (1988) showed how each of these forms of depression is demonstrated differently on the Rorschach and TAT protocols. For example, stories told to TAT pictures with single characters that describe the person as empty, worthless, or helpless without introducing other characters exemplify the first type of depression. Such stories suggest that the narrator is dependent on what is immediately evident in the surroundings without being able to use inner resources to represent the support needed. The vulnerability to depression of an individual lacking such resources may be moderated by a supportive environment.

The profession is in need of a diagnostic system that identifies variables useful for the treatment process (Leve, 1995). Such a system would emphasize characteristics that interact with methods of intervention and, therefore, would serve as guidelines for their selection. The variables that Leve tentatively proposed are grouped into three categories: cognitive, environmental, and social-emotional. The cognitive and social-emotional categories include the following characteristics, among others: adequacy of causal reasoning, social-emotional reality testing, moral reasoning, ability to form and maintain close relationships, ability to identify and express emotions, as well as internal or external locus of control. Leve (1995) reasoned that cause–effect thinking is essential to the success of psychodynamic therapies since developing insight is part of the process. Cognitive treatments, however, do not require as high a level of causal reasoning because the therapist is more actively training the client either by explaining explicitly the causal connections or bypassing explanations altogether through giving behavioral instructions. Likewise, different therapies require different degrees of social intimacy for success, although all therapies have an interpersonal component. The interface of psychological variables and therapy processes need to be refined through further research. It is evident, however, that projective techniques can provide important data about these types of variables.

A fundamental consequence of schematic processing is the perpetuation of the schema in the face of conflicting evidence. Distorted schema that bias attention and information processing to conform to the original conception are related to psychopathology (Horowitz, 1991). The assessment of schema can guide psychotherapy by identifying these unconscious schematic distortions (Blatt, 1991). Bringing to awareness different schema of self (e.g., actual self, ideal self, ought self, dreaded self) and other resulted in positive therapeutic changes in both psychodynamic and cognitive therapies (Singer & Salovey, 1991).

The clinician needs a flexible repertoire of assessment techniques including projective methods because the information that each one provides is not equivalent. No method supplants or takes priority over others but contributes its part to the functional whole. The contribution of projective techniques is best addressed by looking at their place among other measures. The content and structure of responses to projective techniques can be compared with information from self and other reports. The problem-solving strategies and organization of ideas as expressed in projective responses can be compared with how these processes are expressed in more structured tasks. Each source of data is conceptualized according to its contribution to the cohesive patterns that include different levels of personality and variability in functioning under different conditions. A synthesis of these multiple layers permits more precise insight about functioning in various contexts.

Projective techniques clarify the organization and structure of the inner world. Their use is warranted whenever inner life and complex self-regulatory functions are important considerations, particularly when respondents are lacking insight or would be motivated to distort self-report. The task itself involves the utilization of previously organized knowledge structures to provide a response that is not well rehearsed. The linkages among the various elements of the response promote an understanding of the individual's organizational framework. Projective techniques add a unique dimension to the assessment by revealing the respondent's strategies for accomplishing the task and, at the same time, showing the content and organization of ideas that occupy awareness.

4.16.7 PROJECTIVE TECHNIQUES AS PERFORMANCE MEASURES OF PERSONALITY

The projective hypothesis implies that every human action and reaction bears the characteristic features of individuality (Rappaport, Gill, & Schafer, 1968). Therefore, projection is basic to every response process rather than specific to a certain set of stimuli. Essentially, any test is a performance measure designed as an analog to a life task. Projective devices are simply stimuli with known qualities that are used to elicit samples of behaviors that correlate with other behaviors.

The question has been raised about what aspects of responses to the Rorschach stimuli reflect compliance to task demands and what elements entail projection. Exner (1989) suggested that most responses are simply best-fit answers, and that projection on the Rorschach occurs only in responses that deviate from the norm or are elaborated beyond the stimulus field. This view assumes that projection on the Rorschach is an attribute of the response and not inherent in the task. Exner explains that projection is not encouraged by the instructions which simply call for answering "what might this be" nor by the blots which are not completely ambiguous. He points out that techniques such as the TAT force projection, by asking the respondent to develop a story that goes well beyond the stimulus provided. Yet, the story-telling task also makes problem-solving demands (Holt, 1961); the narratives are expected to be organized productions rather than fantasies or random associations. Just as identification of blot contours are expected to be guided by reality constraints, so do perceptions of emotions and relationships as well as the sequence of events in stories told to TAT stimuli.

Exner claims that the operations contributing to the formulation of responses to Rorschach stimuli such as scanning, encoding, classifying, refining, evaluating, discarding, and selecting are cognitive and not projective. However, in keeping with the projective hypothesis, individuals are expected to superimpose unique styles of organization on each of these cognitive operations. The distinction between problem-solving and projection is not an either/or, but a matter of relative emphasis. All projective tasks have problem-solving elements, and all responses are influenced by individual sets or internalized schema. Those who favor the projective viewpoint insist that such inner structures are fundamental to the response process. The conceptualization of the Rorschach as a perceptual-cognitive-behavioral task has been criticized on the grounds that this approach does not give sufficient weight to the role of stimuli from the inner world in the interpretation of stimuli from the external world (Willock, 1992).

As discussed earlier, projection is an ongoing process that comes to the fore if the situation is novel. In a routine, overlearned situation, individuals will respond with familiar, scripted knowledge. However, even a highly structured situation can be misinterpreted. When ideas on the more structured cognitive tasks are associative or disorganized, these cognitive difficulties will be magnified in responding to the less structured performance tests. One youngster who demonstrated serious problems in the application of rote knowledge in a relatively structured task also showed impaired thought process on the projective tests. To the Wechsler Intelligence Scale for Children (WISC III) comprehension item, "what would you do if you saw thick smoke coming from your neighbor's home?" he responded: "Stop, drop, and roll." This response was an association cued by some of the words in the item without understanding the entire question. The fire was in the neighbor's home, yet the child produced a rote response as if he were personally experiencing the fire. He had rehearsed this scenario at school, but was unable to apply the concept even under relatively structured conditions.

The capacity to organize the inner world and deal effectively with distracting thoughts, emotions, or motives should translate into the ability to tolerate and to deal successfully with ambiguity, complexity, and apparent contradiction in a variety of situations. Therefore, the experience of ambiguity and complexity may stem not only from reality demands but also from the individual's emotions and motives (Blatt, Allison, & Feirstein, 1969). A framework is needed for conceptualizing how projective techniques such as the Rorschach and TAT shed light on cognitions in ways that differ from more structured tests. Differences and similarities in task demands of various projective techniques also need to be understood.

The response to the Rorschach task involves the matching of ambiguous stimuli with a memory trace (Exner, 1993). The goodness-of-fit or perceptual match between the blot contours and the object reported constitutes the basic element of reality testing. When telling a story about a picture portraying one or more people, what is demanded is not simply perceptual matching required by the Rorschach task (Beck, 1981) but experiential matching. The narrator searches for a relevant explanation for the picture, then marshalls possible details from memory to satisfy the instructions and meet criteria for an acceptable story. If the stimulus is highly stereotypic, then finding an exact story from memory would be sufficient to accomplish the task. However, if the narrator encounters a picture that cannot be explained by existing schema, then the individual actively constructs the response and fills in details from various schema. Schank (1990) suggests that what exists in memory is a database of partial stories or story elements rather than whole ones. When telling a story to a pictured scene, the narrator draws on this database to construct a set of events and inner processes (characters' thoughts and feelings) that captures the gist of the stimulus.

All tasks are on a continuum of how much organization is required in making meaning of the stimulus and developing the response. The more open-ended the task, the greater the demand to construct the response actively rather than to respond in rote fashion to the stimuli. Therefore, projective techniques are particularly useful when adjustment to routine circumstances is good, but the individual is experiencing problems in less structured situations. Task by task analysis of the assessment battery permits the identification of the specific processes or competencies required for each task and the life situations to which these competencies can generalize.

A wide range of human ability lies outside the domain of standard cognitive tests. It is these competencies in the broad sense that are within the purview of projective testing. The primary advantage in regarding projective methods as performance measures of personality is that various tasks and measures can be differentiated in their demands. Generalization from one test performance to another and to real life performance then can be based on similarity of the demands. If projective techniques are treated as performance measures of personality, then it is possible to delineate task requirements, norms, and expectations about the product as seen with the Rorschach. In addition, it is possible to delineate linkages of that product with the acquisition of prior knowledge structures and with the application of these resources to similar problem-solving situations.

The term "*performance test of personality*" is proposed because it captures two key dimensions of projective techniques: (i) the problem-solving aspect of meeting specified performance expectations such as form quality on the Rorschach; and (ii) the organization of inner resources or schema brought to bear on the production. This dual approach to understanding projective techniques is shown in Table 1.

The distinction between self-report and projective measures has been described earlier. Self-reports ask the respondent to tell about the self, whereas projective measures require the performance of a task from which psychological processes are inferred. If we want to know how a person solves certain types of math problems, rather than ask the individual to report on how well he or she can solve linear algebraic equations, we would present some sample problems. Similarly, if we want to assess personality, we could obtain information through self-report by asking questions such as how flexible are you? How do you interpret cues in an unstructured situation? In contrast, we could request that the individual perform a task that demonstrates how various psychological processes are applied to meet the problem-solving demand.

Tests such as the Wechsler Scales are generally seen as measures of cognition, whereas the TAT and Rorschach techniques are tasks that are thought to reveal personality. This distinction is not entirely a function of the task but includes how the performance is evaluated. Human figure drawings, for example, have been used to estimate both cognitive level and personality functioning. The essential characteristics of tasks that measure personality, in contrast to cognition, are described below.

4.16.7.1 Total Personality

As do life situations, tasks vary in terms of the amount of structure, cues, or prompts provided for responding. As the expression goes, "It's not what you know but when you know it that matters." To be useful, knowledge structures must be spontaneously available when circumstances warrant. In the psychological test battery, personality performance tasks provide minimal guidance to meet the problem-solving demands and, thereby, permit the assessment of spontaneous strategies.

Generally, cognitive measures such as tasks on Wechsler scales do not reproduce the conditions encountered in everyday situations. Sternberg (1985) distinguished between analytic intelligence that is assessed by more structured tasks and practical intelligence. Analytic problems provide all necessary information and have a single correct solution separate from the emotional and social context. Real life problems are less clearly defined and typically call for information seeking. Note the following question on the WAIS-R: "In a movie theater, you are the first person to notice smoke and fire. What should you do?" The respondent is cued that the expectation is to do something. If this were a real situation, the person would not know that he or she was the first to see the smoke and fire. The person may not take responsibility with others around and, therefore, delay action. The individual may consider specifics of the context such as proximity to the fire, its size, or the availability of a fire extinguisher. However, the individual's emotional reaction may disrupt thinking. The actual response would involve the total personality, not merely the intellectual component.

The limitations on generalizing from responses to structured tasks to real life conditions are evident when we consider the attributes that these measures do not assess. These include being aware that a problem exists or that a task needs to be done; setting priorities and planning toward their implementation; pacing and self-monitoring; seeking or utilizing feedback; sustaining long term interest in independent activities; taking necessary risks; organizational skills; and recognizing and

Table 1 Task analysis of performance measures of personality.

Technique	*Ideal performance*	*Competencies needed*	*Inner life (how experience is organized and internally represented)*
Drawing	Symmetry, coherence, realism, match with instructions	Plan execution of drawings within allotted space and limits on drawing ability; organize details with context, handle frustration; investment in the task	Reality testing and quality of thought process; use of schema and organizational strategies; specific preoccupations or concerns
Narrative	Story matches picture and incorporates instructions, appropriate transitions and cause–effect connections; synthesis of various dimensions of experience within individual characters such as integration of inner life with external circumstances; balance among views and needs of all characters depicted; balance between excessive detail and vagueness; cohesiveness among intentions, actions, and outcomes; appropriate time perspectives	Draw on personal experience to construct a story that adequately explains the picture; modulate affects and recognize tensions depicted in the stimulus; organize, plan, and monitor details of story for cohesiveness and inner logic and for compliance with instructions; direct attention from one aspect of experience such as feelings to actions and outcomes; initiative to transcend the pictured cues to describe intentions and realistic, purposeful actions to resolve tensions; bring general abstract principles to bear on the task for optimal integration of multiple dimensions of the stimulus and coordination of inner and outer aspects of experience (thought, feeling, action, and outcome); investment in the task	Reality testing and quality of thought process; understanding social cues; strategies to organize experiences; specific preoccupations and concerns; nature of inter- and intrapersonal schema or scripts brought to interpretation of experiences and picture stimuli (experiential matching)
Inkblot	Accuracy and specificity in matching percept to form; organizing various components of the blots; balance among form and other determinants as well as between precision and vagueness in form definitions; logical and responsive communication during inquiry	Adequate investment in responding; realistic perception and organization of stimulus components; understanding of hierarchical relationships implicit in balancing form with other determinants; comfort with matching precision of percepts to relative imprecision of blots; confidence with ambiguous stimuli; systematic approach to the task	Reality testing; strategic processing of information and connecting new input with previously acquired sets or schema (perceptual matching)

Note: The task is somewhat different for each specific set of instructions and stimuli (e.g., draw-a-person in the rain vs. draw-a-person). Similarly, each Rorschach card and or TAT picture presents a unique stimulus configuration.

responding appropriately to subtle interpersonal cues. These attributes are traditionally viewed as pertaining to the personality domain and reflect the interplay of cognitive and affective processes. A response to the TAT involves the interpretation of the emotions and tensions depicted to recognize a problem, whereas a verbally communicated question or problem often contains cues to facilitate the response, and such cues may not exist in a real-life version of the scenario. Likewise, the Rorschach presents the respondent with options for interpreting the blot contours, organizing the percepts, and communicating them to the examiner.

4.16.7.2 Many Correct Solutions

Just as real-life dilemmas are amenable to diverse resolutions, personality performance measures can be approached in several ways. Therefore, many correct solutions are possible. Responses can differ according to the individual's interpretation of the stimuli and organization of the response. Variability, rather than uniformity, is the expectation with personality performance tasks. Rather than imposing specific criteria for correct responses, such tasks set a general expectation to deal with the stimulus realistically, follow instructions, and logically organize the response. In contrast, most cognitive tasks have one correct solution regardless of whether they call for rote knowledge or a more flexible application of prior knowledge. These distinctions between cognitive and personality performance tests are not absolute; various cognitive tasks share some common processes with personality performance measures. Reading comprehension, for instance, bears a resemblance to personality performance measures when previous sets influence the understanding of the text. There can be little argument about the facts presented in a passage, but individuals can justifiably differ to some extent on their inferences based on prior learning.

4.16.7.3 Generalize to Different Criteria

The distinction between typical vs. maximal performance (Cronbach, 1970) has been used to differentiate between personality and ability measures. This dichotomy contrasts responses to immediate cues and external motivating structures from performance that is regulated and maintained by the individual's long-term investment and initiative. A structured situation such as an achievement or cognitive test would generally tap maximum performance, but such performance generalizes only to similarly structured situations. Because most real life tasks provide less structure, measures of typical performance are also needed in a comprehensive assessment. Correlations are low between measures tapping maximal and typical performance (Sackett, Zedeck, & Fogli, 1988). Therefore, each provides unique information.

The following conditions were suggested by Sackett and colleagues as generally yielding estimates of maximum performance: (i) there is a heightened level of effort and attention because the task is seen as important; (ii) expectations and performance standards are clear; and (iii) the observation takes place over a relatively short time where the individual can exhibit an uncharacteristic spurt of effort that could not be sustained over the long haul. In contrast, the characteristics of measures of typical performance are as follows: (i) individuals are unaware that they are being observed or evaluated, so they are not trying deliberately to perform to the best of their ability; (ii) performance is monitored over a long period of time; (iii) the performance tasks require skills that have to be learned through continuous past efforts—if the task is highly complex, the individual has to bring a great deal of past learning (typical performance) to the current effort; and (iv) performance guidelines are not clear and, therefore, individuals impose their characteristic way of organizing and dealing with the situation.

Performance measures of personality meet several criteria for assessing typical performance. Prior knowledge (schema) is superimposed on current task demands; individuals are unaware of what aspects of their performance are evaluated; and in the absence of structure, they are required to organize and plan their response according to their typical mode of functioning. Understanding different requirements of different tasks helps explain variations in the assessment of similar constructs with different methods. Johnston and Holzman (1979) designated indices of thought disorder in the WAIS and the Rorschach which they tested with a sample of schizophrenics, their parents, and controls. The IQ score and the Thought Disorder Index (TDI) derived from the WAIS were negatively correlated such that the higher the IQ score, the lower the TDI. In contrast, IQ scores and the TDI derived from the Rorschach were uncorrelated. The authors conclude that the tasks make different demands. The WAIS calls for "habitual reactions and the social frame of reference is clear" (p. 61). The social expectations on the Rorschach are less obvious. Therefore, the task sets different requirements and has different implications for the assessment of impaired thought process.

In their sample of schizophrenics, those who were more intelligent were able to limit expression of disordered thinking on the WAIS more easily than on the Rorschach, in part, because WAIS items can be answered with overlearned responses. The feeling of ambiguity versus security determines efficiency of functioning on problem-solving tasks and in real life. Individuals who exhibit thought problems only on unstructured tasks may be more capable of functioning with the supports of structures and clear social cues.

A criticism of laboratory research has been that results do not generalize to nonlaboratory situations (Fromkin & Streufert, 1976; Snow, 1974). This is the case because precise linkages between processes engaged by laboratory activities and other life domains are not specified. Likewise, predictions from test performance to real life adjustment can be accurate only if they are functionally similar. By understanding patterns of strength and weakness across various tasks, the professional can point to areas where the individual can and cannot function adaptively.

4.16.7.4 Differences in Conditions of Learning and Performance

Responses to performance measures of personality, as most tasks, require previously organized knowledge and strategies. However, we distinguish between learning that is promoted by direct teaching (e.g., lecture, textbook) from learning mediated by the individual's synthesis of experience (Epstein, 1994). Personality performance tasks are guided more by self-regulated learning than by formal education. Again, this distinction in the conditions of learning does not apply in a dichotomous fashion to personality versus cognitive measures. For example, an individual's general fund of information is a joint function of direct teaching and the individual's interest and active, effortful processing of information. Depending on the individual, knowledge is acquired through some combination of direct teaching and self-regulated synthesis of experience.

Task requirements also differ on the basis of the conditions of performance. These conditions pertain to the spontaneous versus cued accessibility of prior knowledge and degree of organization required to produce the response. Cognitive measures that assess general fund of information often elicit previously acquired knowledge in piecemeal fashion (highly structured). Personality performance measures such as the TAT and Rorschach set conditions of performance that demand spontaneous accessibility to prior knowledge to interpret the stimulus and formulate the response according to the directions. The more open-ended the response, the greater the need for self-regulated strategies to organize the product.

Performance measures of personality maximize the imprint of organization so that the principles by which experiences are structured and the inner organization of the personality are revealed. A comprehensive assessment battery provides tasks that vary on the continuum of structure provided. Given such variation, it is possible to relate competencies required in each task to performance in life situations requiring similar competencies. This linkage is accomplished by understanding how processes exhibited during test performance are carried over to everyday functioning in various situations and incentive conditions.

4.16.8 SPECIFIC PROJECTIVE TECHNIQUES

Three basic aspects of projective techniques have been recognized (Rabin, 1981). First, the task entails the presentation of an ambiguous set of stimuli and a request to give an open-ended response. The Rorschach and TAT techniques include both of these stimulus and response attributes. Therefore, the focus of interpretation is on the perception of the stimuli presented and the organization of the response. Drawing techniques also demand open-ended responses but usually do not provide a stimulus. Second, the response is shaped by processes that are outside of conscious awareness. An important factor making the response less amenable to conscious manipulation (faking) is that the respondent does not comprehend the meaning of the answers given. Third is the complexity of the interpretive process. Each of these components is briefly addressed below.

4.16.8.1 Stimulus

In general, projective techniques present stimuli that are amenable to various interpretations, and instructions that can be addressed in a variety of ways. The degree of ambiguity has been a prime consideration, although other features of stimuli are also important determinants of the response. A systematic accounting of how the respondent uses stimuli is the essence of the Rorschach technique and must be given greater weight in thematic apperceptive methods (Henry, 1956; Teglasi, 1993).

4.16.8.2 Response

All projective tasks require the individual to draw on internal images, ideas, and relationships to create a response. The respondent must dredge forth past experiences, direct or vicarious, and organize them to meet the task demands. The greater the stimulus ambiguity and the more open-ended the response, the greater the reliance on the organizational structures of the personality rather than on rote knowledge. Yet, the stimulus must have sufficient structure to permit evaluation of the plausibility of the respondent's interpretation. The projective task demand is an analogue of other unstructured tasks and situations where available cues are subject to interpretation. The manner in which the individual interprets the stimuli and organizes the response shows how they will respond under similar conditions (Bellak, 1975, 1993).

Responses to projective tasks involve complex, interrelated processes that have conscious and unconscious components. These include the interplay of cognition-emotion-action tendencies that coordinate perceptions of the outward world with experience of the inner. Projective techniques may reveal aspects of emotion, motivation, and cognition that a person may not wish to expose. These unconscious aspects of responses may relate to issues of invasion of privacy. Faking, malingering, or defensiveness are problematic for any form of assessment but are assumed to be less so with projective techniques. However, the issue of fakability of projective tests has been inadequately investigated (Rogers, 1997).

4.16.8.3 Task Demand

The stimuli presented together with the instructions set the task demands (Teglasi, 1993). Projective techniques impose task demands that cannot be met with a simple response such as a request for specific information. They require respondents to apply what they know to produce a story or a drawing or to identify an object that may fit the ambiguous contours of an inkblot. Various projective tasks have features in common yet differ in important ways. Projective methods have been understood as problem-solving tasks with designated performance expectations as well as measures of personality with emphasis on individual variation.

4.16.8.4 Interpretation

The interpretive task is complex even when the scoring categories and interpretive guidelines are straightforward as in the Comprehensive System for the Rorschach. Despite the relatively clear guidelines and availability of norms (not to mention computerized reports), the Rorschach requires the examiner's trained inference along with the more objective coding. Interpretation is in keeping with the clinician's theoretical framework and understanding of the task demand. Therefore, the interpretation of responses to a projective test cannot be more satisfactory than the adequacy of the theory informing the interpretation of the evidence and the examiner's skill in evaluating that evidence.

The complexity of the interpreter's job is maintained because conclusions rest on the understanding of meaningful patterns rather than isolated response elements. Personality performance tests yield products where the whole is more than the sum of the parts. The evaluation of the response must account for the organization and cohesiveness of the different components. Therefore, various units abstracted from the whole cannot be treated in a piecemeal fashion. The more open-ended the response, the more amenable to separate analysis of structure, form, and style. Although these aspects of the response can be conceptually separated, their interpretive value lies in their relationship to the task demand and to each other. Content, for example, cannot be properly understood apart from the manner in which it is organized. Finally, the examiner must be aware of connections between empirical findings and theory to avoid speculation. Yet, empirical support of theory is not always conceptualized in ways that are useful in making decisions about one individual.

The sign approach attempts to provide empirical evidence for interpretations through the identification of features that occur most frequently in specified clinical populations. However, using a list of signs in an atheoretical cook book fashion is inadequate because a particular sign derives its meaning from the context of other responses. For example, in a TAT protocol, a stereotyped approach or meticulous listing of details in story telling may be viewed as resistance or as representing the respondent's best efforts. Likewise, concern with minutiae that most respondents disregard may be viewed as an index of hypervigilance or of concrete functioning. The appropriate interpretation depends on the pattern of responses within the story telling task and across tests in the battery.

Inferences are drawn from the clients' behavior during the evaluation as well as from the product. Any changes in demeanor or emotional reactions in response to the various tasks are noted as are spontaneous comments, time elapsed, expressions of uncertainty about performance, or attempts to seek structure. Aspects of the response process are also considered. These refer to the manner of working, compliance with instructions, sequence of ideas (e.g., planful, trial and error, organized, haphazard), as well as to dysfluencies, pauses, or hesitations. The product itself is amenable to analysis in relation to the process of the response, content, and structure. The formal or structural aspects of the product, of course, reflect the response process.

Next is a brief review of three major projective techniques in terms of the following elements: (i) task demands that include the stimuli and instructions; (ii) response process; and (iii) general interpretive approaches and issues.

4.16.9 THEMATIC APPERCEPTIVE TECHNIQUES

Apperception tests generally use pictures and standard instructions to elicit stories. The TAT (Morgan & Murray, 1935) is the most popular of these. Although introduced as a method to assess a particular theory of personality (Murray, 1938), this technique did not remain wedded to the theory of personality from which it sprang. The thematic approach described in the standard manual (Murray, 1943) was amenable to use with a wide range of scoring methods and has been adapted for clinical and research purposes. Spin-offs of thematic apperceptive approaches have used different picture stimuli and diverse scoring approaches. A variety of nonclinical coding systems for content analysis of the TAT appears in an edited volume by Smith (1992a). More clinically oriented systems also abound (e.g., Bellak, 1975; Cramer, 1996; Henry, 1956; Karon, 1981; Rappaport, Gill, & Schafer, 1968; Teglasi, 1993; Tomkins, 1947; Westen, 1991; Wyatt, 1947). Typically, clinicians have used the TAT by applying broad units of inference in an idiographic, qualitative manner. In contrast, researchers have scored TAT responses for various personality characteristics using specific and narrow scoring criteria. As of this date, there is no widely agreed upon scoring system for the clinical use of thematic methods. However, it is generally acknowledged that classification into diagnostic categories is not a chief purpose of the TAT.

4.16.9.1 Response

When given TAT-like pictures, the scene sets the topic. The story telling directions call for a description of what is happening in the picture, what happened before, what people are thinking, how they're feeling, and how everything turns out at the end. These instructions require the narrator to attribute thoughts, feelings, and motives to the characters depicted and embed these inner aspects of experience into an appropriate time frame and sequence of events that coordinate inner life with appropriate actions and outcomes.

4.16.9.2 Stimuli

Picture stimuli play a major role in determining the story content (Kenny, 1964; Murstein, 1965), and reviews of frequency of themes specific to TAT stimuli are available (Bellak, 1975; Henry, 1956; Holt, 1978; Murstein, 1968; Stein, 1955). Ambiguity was a primary consideration in designing the TAT stimuli (Murray, 1938). However, the degree of ambiguity varies within the set of TAT cards. The manner of defining ambiguity has ranged from judges' estimates of the number of interpretations that can apply to each card (Kenny & Bijou, 1953) to actual degree of variability in responses (Campus, 1976). Many of the TAT pictures clearly show who the characters are, what they are doing, and the emotions they are experiencing (Murstein, 1965). Yet, they permit a great deal of variation in the style of expressing similar themes. While stimulus ambiguity is essential to the projective hypothesis, other issues pertinent to stimuli such as similarity of main character(s) to the narrator, emotional tone, complexity, and latent meaning have been researched (see review by Teglasi, 1993).

Broad conclusions about the advantages of various degrees of stimulus ambiguity cannot be drawn without specifying the nature of the scoring system, population, and purpose. Low ambiguity has been favored in the assessment of specific motives (Singer, 1981). However, cards with high ambiguity have been preferred to measure the relative strength of two motives (Atkinson, 1992). When studying hostility, it was found that aggressive intent was most effectively measured by a picture with low relevance for hostility, whereas guilt over hostility was best measured with a picture of high relevance for hostility (Saltz & Epstein, 1963). The authors concluded that pictures with low relevance for unacceptable behavior such as hostility measure drive toward its expression, and those with high relevance measure inhibition or guilt about its expression.

Three degrees of ambiguity within the TAT set were estimated by a group of judges—high, medium, and low (Kenny & Bijou, 1953). The richness of personality content obtained varied as a function of card ambiguity. Cards in the medium ambiguous set yielded stories with the most personality information. The authors discussed two dimensions of ambiguity in card stimuli: (i) the number of cue constellations available to guide the response; and (ii) the definitiveness of the cues available. Sets of stimuli with graded levels of ambiguity permit the evaluation of responses to situations along a continuum of structure.

Studies of stimulus ambiguity are inconclusive because investigators compare various degrees of ambiguity that are not on the same points on the ambiguity dimension. Furthermore, conclusions about stimulus variables have generally been drawn regarding their adequacy to elicit content germane to the assessment of specific motives or relatively narrow psychological processes rather than their effectiveness to assess formal qualities of organization that reflect the respondent's capacity to deal with the task demands.

For clinical purposes, stimuli must permit the simultaneous evaluation of multiple psychological processes and analysis of both structural and content properties. Some clinicians argue that stimuli are relatively unimportant as long as they are not highly structured (Arnold, 1962; Karon, 1981). Criteria for picture selection within the TAT set are available (Bellak, 1975; Birney, 1958; Henry, 1956; Teglasi, 1993). Haynes and Peltier's (1985) survey of clinicians in juvenile forensic settings indicated that the mean number of cards used was 10.25. The most frequently administered TAT cards were similar to those deemed as most productive in previous studies (Cooper, 1981; Hartman, 1970). The TAT pictures have been criticized for being out of date in relation to clothes and hairstyle (Henry, 1956; Murstein, 1968) and for their predominantly negative tone (Ritzler, Sharkey, & Chudy, 1980). These concerns are mitigated by findings that the stimuli effectively permit expression of convictions and conceptualizations of affect. The negative tone presents unfinished business or a dilemma to be resolved and provides the opportunity to observe how the respondent interprets the tensions depicted in the picture. Furthermore, the negative scenes make it possible to observe how the narrator moves beyond the sadness or conflict to a more positive resolution by noting appropriateness of transitions between the negative state of affairs and positive outcomes. It has been suggested that pictures representing relatively universal social situations that most people encounter in their lives are suitable across various age and subcultural groups (Veroff, 1992).

4.16.9.3 Interpretation

Karon (1981) suggests that only a complex scoring system can be clinically useful. At the same time, "an adequate scoring system is so complex that no one has time to score the protocol" (pp. 94–95). Simple scoring systems are inadequate because they "throw away most of the information in the process of scoring, and hence turn out not to be clinically useful" (p. 95). This point is well taken. A multiplicity of thematic methods are available that provide carefully described scoring systems to assess relatively narrow aspects of personality. Although these systems have demonstrated a high degree of accuracy and reliability, their narrowness limits their clinical utility. Karon's remedy is to rely on clinical judgment in a sentence by sentence interpretation of the protocol, much like a clinical interpretation of an open-ended interview. This approach is compatible with the view that projective techniques are not tests, but are clinical tools that rely on the skill of the practitioner (Anastasi, 1976). An alternative is to agree on key psychological processes and guidelines for their measurement that may be time consuming to learn but less cumbersome to use once mastered.

Historically, there have been two ways to designate units of analysis for TAT stories: product centered versus narrator (person) centered. The former method focuses on qualities of stories that differentiate between groups; the latter approach emphasizes the psychological constructs to which attributes of stories relate. The narrator-centered approach permits the hierarchical organization of elements identified in stories in reference to broader features of the personality. For example, Rappaport, Gill, and Schafer (1968) organized story qualities according to affective lability and looked for content shaped primarily by affective responses to the picture. Holt (1958) looked at clarity of thought (vagueness, overgeneralization, disjointedness of organization) and emotional inappropriateness (arbitrary turn of events, forced endings). With such a focus on psychological processes, patterns of story elements can be organized in relation to relevant constructs.

Holt and Luborsky also emphasized adequacy of hero. This variable, however, is not a psychological process but a quality of the story considered important because it correlated with supervisor's ratings of overall competence during psychiatric residency training. Adequacy of the hero was subsequently viewed by Bellak (1975) as an index of the narrator's competency. However, this interpretation may not hold if other characters depicted in the stimulus or introduced in the story are helpless. This conclusion is also problematic if the narrator has not accurately incorporated the cues provided in the stimulus. A narrator-centered approach to this variable would focus on various psychological processes indicative of the story teller's competency. Three points are relevant here. First, no part of the story can be interpreted separately from the others. Second, the emphasis must always be on the narrator rather than the narrative. Third, the demands set by stimulus properties must be considered.

Emphasis on qualities of the narrator rather than of the product permits consideration that these qualities can be expressed in different ways. Therefore, the interpreter does not seek one-to-one correspondences but focuses on the fit between the pattern of responses and the psychological processes to which they pertain. For example, the assessment of a construct such as cognitive integration with thematic apperceptive methods may be accomplished by exploring the manner of interpreting stimuli or the manner of coordinating different dimensions of experience such as thoughts, feelings, and actions. More complex units, such as self-regulation, subsume cognitive integration as well as affective-motivational processes (Teglasi, 1993). Such a focus on the psychological processes of the narrator avoids the sign approach and promotes an understanding of the constructs being measured. This occurs by organizing story qualities in terms of a clear conceptualization of their relationship to relevant psychological variables. Conceptualizations of the psychological variables can undergo continuous refinement in keeping with research across various subdisciplines.

Form and content in the TAT have been distinguished (Henry, 1956; Holt, 1958; Teglasi, 1993). Formal features focus on generalized properties of content such as the organization and coherence of the response. These formal units of analysis refer to how the details of the content relate to each other, to the evoking stimulus, and to the directions given. For example, inferences can be made from TAT stories about striving for long range goals (independently of the specific goals or concerns) based on connections among characters' purposes, actions, and anticipated outcomes. Internal attribution of feelings is a formal quality because it need not refer to any specific feeling. These formal qualities of the story provide information about the cohesiveness and reality base of the narrator's schema rather than about specific concerns. They are akin to procedural knowledge within schema theory and should be evident across stimulus cards within a task and even across various projective tasks. Formal features of stories lead to inferences about the structure of personality such as cognitive integration, affect maturity, or self-regulation. Given that picture stimuli evoke specific content, it is these generalized or formal properties of content that can reliably reveal clinically important psychological processes.

A focus on the formal elements involves the following general principles of TAT interpretation.

(i) An organized narrative, not a patchwork of associations, is expected. The story is, therefore, evaluated in terms of how the narrator meets problem-solving task demands. The individual's problem-solving approach can be examined from two vantage points. The first focuses on how the story is told in relation to the stimulus and instructions. Accordingly, the narrative is evaluated in terms of accuracy in capturing the tensions depicted in the stimuli, compliance with instructions, and the logical and realistic unfolding of events. The second emphasizes how the characters are described in reference to the stimulus cues, the cohesiveness of inner states, actions and outcomes, as well as the manner in which characters define and resolve the dilemmas set before them in the stimulus.

(ii) Units of inference based on interconnections of elements such as the links between causes and effects or actions and outcomes pertain to formal characteristics of content. The content is not taken at face value. Instead, units of inference are based on understanding of psychological processes and task demands. The instructions call for inclusion of various levels or dimensions of experience such as thoughts, feelings, and actions. These human tendencies to think, feel, perceive, and act are interrelated, and interpretive meaning is derived from an understanding of their patterns. However, the narrator's understanding and the clinician's framework for conceptualizing such relationships must be clearly distinct (Cramer, 1996). The professional determines the implications of the narrator's understanding of the linkages among sequences of events, thoughts, feelings, behaviors, and outcomes.

(iii) The interpretation seeks patterns that elucidate the schematic structure connecting the various story elements (e.g., the manner in which self and other are differentiated and expectations about causal sequences). These schema are products of the synthesis of past experience and provide the templates for organizing current experiences.

4.16.10 RORSCHACH TECHNIQUE

The Rorschach has achieved wide acceptance due in large part to the development of the comprehensive system (Exner, 1993), an integration of previously established methods. The availability of normative data for discrete coding variables establishes psychometric credibility. In addition, the comprehensive system provides interpretive strategies for synthesizing the complex data and encourages the interpreter to move back and forth between the formal data of the structural summary and the content and language of the response.

The Rorschach administration is conducted in two phases. First is the association phase during which the respondent is shown each of 10 inkblots and asked "What might this be?" The inquiry phase begins after the respondent reports what is seen on each of the cards. The examiner guides the respondent to clarify the determinants of each response to permit accurate coding. Although the Rorschach is described as a cognitive-perceptual task (Beck, 1981), what is interpreted is what the respondent verbalized. Hence, the examiner must be well trained in coding and administrative procedures to avoid the many potential pitfalls of conducting an inquiry and to maintain the delicate balance between seeking sufficient clarification and promoting response sets by pressing too far.

4.16.10.1 Stimulus Features and Response Parameters

The structural approach to interpreting responses to Rorschach cards creates a close link between the stimulus qualities and parameters of the responses that are coded. Stimuli vary in the degree to which they are solid or broken, colorful or achromatic. They also display variation in shading, empty (white) spaces, and contours that are sufficiently familiar to evoke popular responses. Reported perceptions are coded according to patterns of attending to stimulus qualities including choices of blot areas and use of various features of the stimuli such as form, color, shading, hue, white space, card symmetry, or any combination. Also important are the accuracy of the match between the reported percept and selected blot contours and the relationships among the percepts (organization). Each code assigned to a response represents a particular psychological process. However, specific response elements such as choice of blot areas are interpreted in relation to other variables such as degree of organization in relating various blot areas. A response involving the synthesis of parts into larger units reflects more complex thinking than a vague, undifferentiated holistic perception of the card.

Perception of movement is not attributable to a quality of the stimulus which is static, but an imposition of the perceiver. It has been argued that phenomena such as movement are not perceptions (because blots are static, not moving) but mental constructions based on perceptual experiences (Blatt, 1992). In general, movement responses attempt to put greater specificity on form qualities. When such responses involve human activity, they often represent the individual's use of inner resources to modify perceptions of the external world. However, this interpretation changes if these efforts are combined with inaccurate form identification. If the forms identified differ greatly from those reported by others in a culture or society, then the individual is likely to make other people uncomfortable (without necessarily knowing why) by engaging in behaviors that depart from their expectations. Use of white space as figure-ground reversal is an analog to other figure-ground reversals such as an oppositional stance toward the world. Conventionality is interpreted according to the extent to which thought process is like others or different. The shading is a relatively subtle aspect of the Rorschach stimulus so that not all respondents report them as determinants (although they may perceive them). Such responses require a certain level of sensitivity or perceptiveness to verbalize.

The style of communication and nuances of language are important considerations (Smith, 1994). Therefore, characteristics of verbalizations such as redundancies or logical inconsistencies are noted. Even commonly reported responses can be expressed in unique ways. Percepts on the Rorschach are at times given without commitment or reflection, and the examinee might in such instances be reluctant to engage in the inquiry just as he or she was removed from the response process. These stylistic qualities of the response process provide a context for the interpretation of the variables in the structural summary.

Content given is also considered. For example, those who give a broad range of content have a greater variety of interests, and those who provide many human associations are exhibiting a broader interest in people. However, what is emphasized in coding responses is the manner in which the respondent uses the stimulus.

The Holtzman Inkblot Technique (HIT) (Holtzman, Thorpe, Swartz, & Heron, 1961) was designed to overcome what had been perceived as psychometric limitations of the Rorschach. Two parallel sets of inkblots were constructed, each set containing 45 inkblots plus two practice blots that are identical for both sets. The HIT was designed to differ from the Rorschach in several ways besides the number of inkblots (Holtzman, 1981). Two important differences are that the respondent is instructed to give only one response per card rather than leaving it open and that a short and restricted inquiry is given after each response. This overcomes criticisms of the Rorschach arising from the variations in the style of inquiry and from the widely varying number of responses obtained which make it difficult to use norms. However, abandoning the principles basic to the Rorschach also has disadvantages. One cannot observe whether the respondent can see something different in an inkblot once it has been identified in a particular way (flexibility), and the number of responses *per se* is an important consideration. Sequential analysis of multiple responses to the same stimulus is not possible in the Holtzman format.

Clearly, in devising variations on techniques, there are trade-offs that need systematic attention. Given such acknowledgment, clinicians can make appropriate choices.

4.16.10.2 Interpretation

The process of interpreting the Rorschach has distinct phases. First, the examiner codes the responses and organizes them into patterns by calculating the variables in the structural summary. These variables are then compared with norms to designate deviations from expected patterns. A strategy for interpreting the protocol is selected based on constellations of responses that depart from the norms. The final step involves the synthesis of the norm referenced response patterns according to the psychological process associated with them. This step also involves analysis of the details of content and of the quality of the verbalized responses. The comprehensive system provides clear guidelines for each step of the interpretation process (Exner, 1993), but the examiner integrates information from various sources to refine conclusions and formulate recommendations.

The comprehensive system began with an almost exclusive emphasis on formal, structural characteristics emphasizing the perceptual-cognitive nature of the task to match blot areas with objects (Weiner, 1994). However, language and other associations are considered important (Rappaport, Gill, & Schafer, 1968) and have become increasingly incorporated into the interpretive procedure (Exner, 1993). Smith (1994) cautions against acceptance of the Rorschach as a test with one correct method of interpretation. An exclusive focus on a single approach fails to acknowledge the potential contribution of those with a different perspective. Content interpretation including recent attempts at codings of object relations (Lerner, 1992; Stricker & Healy, 1990) is difficult to reconcile with the emphasis on structural variables. Historically, these two types of data have been viewed through different theoretical perspectives. An alternative suggested by Weiner (1994) is to focus on distinctions within personality rather than on classes of data such as form versus content.

4.16.11 DRAWING TECHNIQUES

Drawing tasks do not provide a stimulus but do exhibit the dual nature of projective devices as problem-solving tasks that reflect internal representations. Drawing a person or a family requires the translation of a three-dimensional memory image into a two-dimensional graphic representation within the constraints of the respondent's artistic ability. This conversion of three dimensions into two and compliance with instructions are the chief problem-solving task demands. Performance of the drawing is affected by the conceptualization of the object drawn, motor execution, attention to detail and spatial relationships, as well as planning and organizing the production. Like other projective tasks, the drawing can be evaluated according to its structure (proportion, elaboration, consistency of detail), content (what is drawn), and style (line pressure, size, placement). Machover (1949) suggested that structural or stylistic elements of size, placement, quality of line, positioning, symmetry, elaboration, or shading are more reliable than contents such as body parts and clothing. Performance standards for drawings can be set such as respect for inside and outside boundaries of the persons drawn and coherence and balance among details. A stylistic quality such as perfectionism and frequent erasures can be compared to the eventual quality of the product. The sheet of paper sets the boundaries or limits for the drawing. Very small or very large drawings suggest that the individual has difficulty setting boundaries in relation to the environment. However, other influences such as impaired psychomotor functioning (e.g., arthritis) in older clients (Kahana, 1978) must be acknowledged. Furthermore, processes such as lack of planning contribute to the use of available space. The individual may draw a head that is too large (lack of planning) so that the drawing either is out of proportion or cannot be completed on the page.

Drawings have been used to estimate intellectual functioning and neurological status as well as to assess personality. The draw-a-person Test (DAP) widely used as a projective technique to assess personality (Machover, 1949) is also used to measure mental development (Goodenough, 1926). Its usefulness as an indicator of concept formation (intelligence) in the developing child was based on how closely the drawing approximated life-like proportions and details. Basically, the degree of realism was the criterion reflecting the child's conceptualization of the outward world through correct depiction of space and perspective. The Goodenough (1926) Draw-A-Man Test and the Harris (1963) revision of some criteria included items that appeared more frequently with increasing age. However, even as emphasizing age related trends, Goodenough (1926) noted the following individual variations in drawing that were not linked to development: (i) detail dominated but with few ideas; (ii) unique depictions that seem comprehensible only to the drawer; (iii) suggestive of a flight of ideas; and (iv) contradictory combination of primitive and mature characteristics.

Koppitz (1968) also differentiated qualities of children's drawings that were not a function of age. These aspects of drawings were stylistic: (i) qualitative aspects of integration, symmetry, and shading; (ii) presence of unexpected characteristics; and (iii) absence of expected characteristics beyond various ages. These qualities were viewed as manifestations of concept formation that were impacted by motivational and emotional aspects of the personality rather than chronological development. Drawing tasks are frequently given to adults to assess neurological status based on the observation that brain damage interferes with the integration of spatial, perceptual, and motor responses needed to execute drawings (Swindell, Holland, Fromm, & Greenhouse, 1988; Mendez, Ala, & Underwood, 1992).

Currently, the most frequently used projective drawing technique is the Draw-A-Person (DAP) as developed by Machover (1949) and expanded by others (Hammer, 1958; Handler, 1985; Koppitz, 1968, 1984; Urban, 1963). Also popular are the kinetic family drawing (KFD) technique (Burns, 1987; Burns & Kaufman, 1970, 1972); and House–Tree–Person (H–T–P) drawing task (Buck, 1948, 1987).

4.16.11.1 Interpretation

The dichotomy between person centered and product centered interpretation mentioned earlier is relevant here. Person centered approaches organize aspects of the production according to relevant psychological processes of the client. The interpreter looks for coherence in form, content, and style of the production with the hypothesized psychological variables. The professional also notes consistency with drawing instructions, nature of verbalizations about the figures drawn, and the amount of effort expended. Global interpretations of interrelated parts are more valid than isolated interpretation of specific details. The sign approach to validating drawings has been criticized (Kahill, 1984; Roback, 1968; Swensen, 1968) because there are multiple possible interpretations for any quality, depending on the overall pattern. For example, size and detail of drawings can signal preoccupation, differential valuing, conflict (Fisher, 1986) and/or quality of planning and organization. Similarly, qualities such as shading or erasures are said to signal anxiety. However, the nature of the shading is important since this feature may enhance artistic quality and not reflect anxiety. It should also be noted that people differ widely in how they cope with anxiety. Some anxious individuals will race through the task without shading or erasing. If the focus is on the psychological processes of the individual, then the configuration of response elements or signs would be expected to cohere around that process.

It has been suggested that the clinician ask what a particular sign could mean rather than what it does mean (Handler, 1985). A critical review of the literature on the KFD technique (Handler & Habenicht, 1994) emphasized the need to study more holistic, integrative approaches to the KFD rather than the interpretation of a series of single signs. The authors argue for the importance of focusing research on the interpretive approach of the clinician using the KFD rather than on the technique itself. The use of drawings for multiple purposes (e.g., concept formation, neurological status, or personality) demonstrates the multidimensional aspects of projective tasks which permit simultaneous evaluation of responses from multiple perspectives to reveal various interrelated facets of functioning. Broad performance expectations pertaining to structural features of the product can be delineated. As with other performance measures of personality, individuals are free to vary their approach to meeting these general problem-solving expectations. One caveat is that artistic quality of the drawings appears to influence clinician's interpretation (Feher, VandeCreek, & Teglasi, 1983).

4.16.12 CASE ILLUSTRATION

On the surface, the three projective techniques reviewed seem very different. Yet, they reveal similar information when compared in terms of structural qualities. The overlap between the TAT and Rorschach is illustrated by comparing formal aspects of Carl's TAT stories with conclusions drawn from the Rorschach. Carl, a 19-year-old man, had stopped attending school on a regular basis after the seventh grade and left school altogether during the 10th grade. At the time of this evaluation, Carl was incarcerated and attempting to continue his education in prison. He was doing poorly in his classes and was referred to determine if he would qualify for special education services. His WAIS-R IQ score was low average. The intent here is not to develop a comprehensive clinical picture but to point out the overlaps between the two projective techniques. Table 2 shows the consistency of the TAT variables across cards and also displays the relevant Rorschach data.

> *Card 1.* Little boy thinking, he's thinking how. I guess he's thinking a way how to work the violin. (Before?) He don't know how to work the violin. (TO?) Turns out he's still sitting there thinking.

Table 2 Corresponding aspects of Carl's TAT stories and Rorschach variables.

	Card									
TAT variables	1	2	3	4	5	6BM	8BM	12M	13MF	13B
Imprecise accounting of the stimulus		X	X	X		X	X		X	
Vague, concrete, or stereotypic story	X	X	X	X	X	X	X	X	X	X
Requires more than one query	X	X	X	X	X		X	X	X	X
Story ending is concretely tied to the stimulus	X	X	X	X	X	X	X	X	X	X
Relationship among characters are unclear, stereotypic, or unstated	NA	X	NA	X	NA	X	X	X	X	NA
Sense of helplessness (e.g., inaction, when it is warranted; lack of initiative or inertia)	X		X	X	X	X		X	X	X
Inner life such as intentions for actions or feelings is not sufficiently elaborated	X	X	X	X	X	X	X	X	X	X
Implausible sequence of events (e.g., cause–effect, timing, coherence of action with purposes)	X	X	X	X	X	X	X	X	X	X
Focus is dominated by the immediate circumstance or consideration	X	X	X	X	X	X	X	X	X	X
Insufficient integration of detail to comply with instructions	X	X	X	X	X	X	X	X	X	X

Rorschach variables

Resources and control

EB = 0:0; EA = 0; eb = 7:0; es = 7; *D = –2; *Adj. D = –2

Affect

FC:CF + C = 0:0; Pure C = 0; *Afr = .2143; S = 0; *Blends: R = 0:17; CP = 0

Interpersonal

* COP = 0; AG = 0; Food = 0; Isolate/R = .1764;* H:(H) + (Hd) + Hd = 0:0 *H + A:Hd + Ad = 13:0

Ideation

a:p = 4:3; Sum 6 = 0; M^a:M^p = 0:0; Lvl 2 = 0; 2 Ab + (Art + Ay) = 0; W Sum 6 = 0; M– = 0; M none = 0

Mediation

*P = 4 *X + % = .29; *F + % = .30; *X–% = .4117; S–% = 0; *Xu% = .29

Processing

Zf = 12; *Zd = –.8.5; *W:D:Dd = 13:1:3; *W:M = 13:0 *DQ + = 3; *DQv = 4

* Deviates from normative expectation; R = 17; Lambda = 1.428; PSV = 3; Positive indices: SCZI; CDI.

The boy in the story can only keep sitting there thinking how to work the violin. He does not have sufficient resources to deal with the task or to find alternatives. Likewise, the narrator cannot go beyond the picture cues provided (concrete) to produce an ending or to describe purposes or deliberate actions. When prompted to tell what happened before, the response (didn't know how to work the violin) stayed within the moment without considering larger purposes. The fact that the narrator does not introduce other characters to garner support is consistent with the absence of human content in the Rorschach. The helplessness displayed by the character and the narrator is consistent with the positive Coping Deficit Index (CDI) on the Rorschach, suggesting insufficient resources to formulate responses to demands. The short and simplistic story is devoid of inferential or interpretive processes as consistent with vague and concrete information processing on the Rorschach (Lambda, Zd, DQ+, DQv, W:M).

> *Card 2*. The man in the field working. Lady holding onto the stump because she looks like she's pregnant. Girl just come home from school. (T&F). Probably thinking about all the work they gotta do. And the man and the girl probably feeling the same pain, or sorta, for the lady that's pregnant. (Before?) The man started working. (TO?) Everything got done.

The story is closely anchored to the stimulus, yet the narrator does not give priority to the young woman in the foreground, nor does he indicate how the three people are related. Connections are vague and concrete such as the description of the lady as holding on to the stump because she's pregnant. Characters are differentiated only by external appearances, and they all feel the same. As in the previous story, inner life is not elaborated. Carl's Rorschach record contains no human movement, a variable that is associated with greater acceptance of inner thoughts and greater interpersonal awareness. None of the TAT stories suggests the availability of these resources.

> *Card 3*. Little kid was probably tired so he or she fell out the chair. (T/F?) Maybe, I guess they think that's where they gonna sleep at. And probably feeling stupid because he's on the floor. (TO?) That's where they gonna be.

The character falls because he's tired and, despite feeling stupid, just stays there. Again, there's no purposeful action or inner resources but an inertia that's consistent with vague and insufficient processing of information. Feeling stupid, yet doing nothing, is consistent with low self-worth shown on the Rorschach. Other Rorschach variables suggest that self-perceptions are naive and not guided by insight (no FD, no H) nor are these perceptions accompanied by dysphoric mood, anger, or irritation.

> *Card 4*. Lady trying to convince her husband that she love him, but he ain't trying to hear it. (Before?) They was fussing, probably. (TO?) He's about ready to walk away.

The story shows no attempt to understand the inner life or concerns of others but indicates detachment from relationships. The explanation of the stimulus is incongruous—the man looks very angry, but his wife is declaring her love. When asked what happened before, the couple was described as fussing, but the nature of their disagreement is unstated. The picture shows the man turning away, and the narrator does not have the inner resource to provide a solution to the conflict that departs from the stimulus. The Rorschach also suggests that social relationships are superficial, distant, and guarded. The oversimplifying style is likely to lower sensitivity to the needs and interests of others, and this tendency is coupled with a lack of interest or detachment from relationships (no H; no COP or AG). The affect cluster suggests an approach to the environment that minimizes affective engagement. Lack of resources rather than affective provocation is at the heart of Carl's difficulties. His reluctance to process emotional stimuli (Afr) is consistent with his general style of oversimplified processing of information (low blends, high lambda, under-incorporating approach).

> *Card 5*. The lady must have heard someone in the house, so she ran and hide in the closet. (T/F?) She probably thinking that she gonna get hurt. (TO?) That no one was there.

The woman runs and hides because she feels vulnerable. Later, when she realizes there's no danger, she makes no attempt to discover the source of her concern (noise). This absence of initiative and curiosity is consistent with non-reflection, simplified processing, and detachment evident in the Rorschach as described previously.

> *Card 6BM*. Looks like his mother told him some bad news. But the news didn't only hurt him; it hurt both of them. (TO?) That something did happen.

Again, the story offers vague, nonspecific descriptions such as bad news and little differentiation between characters. Since the nature of the bad event is not understood, it is not possible to deal with it adaptively. The absence of specific story details is consistent with the other cards and with vague, simplified processing suggested by the Rorschach. The story ends with the character convincing himself that something did happen.

> *Card 8BM.* Looks like they was trying to rob the lady, and she shot one of them. And the man's two buddies trying to get bullets out of him. (T/F?) Probably thinking that he ain't gonna make it. (TO?) That he made it.

The story presents an unrealistic sequence of events and poor integration of foreground and background components of the picture. There's no reasoning about intentions or consequences. Rather, the narrator gives simplistic associations to the stimulus without rule governed connections among likely sequences of events, between causes and effects, or between short and long term outcomes. This detachment from conventional thinking is consistent with poor reality testing on the Rorschach (poor form quality, low populars, incomplete processing of available information).

> *Card 12M.* This is a lady? (It can be whatever it looks like to you) The old man is praying for the lady because he thinks that if he prays, she'll get better. (Before?) She was sick. (TO?) That she didn't make it.

As with the other stories in this protocol, the story does not depart from the immediate cues of the stimulus. Connections between sequences of events and among characters depicted are vague or nonexistent. For example, we do not know what is the relationship between the old man and the lady. The narrator is powerless to introduce alternatives that are not cued by the stimulus such as consulting a physician. The lack of integration and insufficient organization of details is consistent with Rorschach patterns suggesting that Carl formulates decisions without sufficient processing of information (e.g., high lambda, low developmental quality, low blends, and underincorporating style).

> *Card 13MF.* The old man just got up for work, but his wife didn't have breakfast made for him. (T/F?) One of them feeling really tired. The other feeling left out. (TO?) Turns out that one of them sleep and one of them woke.

This story is also rather concrete and stereotypic. Each character has his and her roles to play, but the narrator cannot develop the story beyond what is seen in the stimulus (reliance on immediate external circumstances).

> *Card 13B.* Little boy ain't got no friends and wants something to drink bad. (Before?) Just sitting there. (T/F?) I guess he thinking why he's the only one there. (TO?) That he wasn't.

Just as in Card 5, the ending contradicts the initial premise of the character. Things are not the way they seem, but there's not enough initiative to investigate or deal with the discrepancy. Such lack of investment in processing information is compatible with the failure to process critical cues (underincorporation) on the Rorschach. Again, no purpose, interpersonal connection, or guiding principle is expressed.

The similarity of the conclusions derived from both techniques suggests that they are assessing common processes. Both methods suggest that resources and reality testing are not adequate to meet daily life demands. Furthermore, the focus on the immediate and the concrete, along with haphazard processing of information, hinder the development of schema that are sufficiently elaborated to provide rules that govern the synthesis of experience, the regulation of behavior, and the expression of affect. Without such inner guides to self-regulation, Carl exhibits tendencies toward impulsive, antisocial behavior and has limited capacity to delay gratification. Carl would benefit from a highly structured learning environment with highly structured tasks and frequent feedback and redirection.

4.16.13 VALIDATION ISSUES

Major criticism has been directed not only towards projective methods but also towards psychometric approaches to assessing personality which classify instances of experience, thought, or action into trait categories. The inference process, in general, has been maligned by the behaviorist movement. Compelling arguments have been made to exclude from consideration anything but overt behavior and objectively coded environmental variables (Skinner, 1953). Psychology has moved far from that position to a recognition of the importance of understanding the meanings that individuals assign to environmental events and to their own behaviors. Indeed, cognitive structures and processes are now considered to be at the heart of individual differences in experience, thought, and action, both adaptive and maladaptive (Cantor & Kihlstrom, 1987).

The process of inference is not restricted to the use of projective tests—it is what makes the psychologist a professional rather than a psychometrist. By giving priority to some scores or using qualitative information to increase understanding of the scores, the clinician engages in the process of making professional judgments (Groth-Marnat, 1990). These judgments cannot and should not be eliminated. It is important, however, for practitioners to reflect on the quality of their inferences and the usefulness of their decisions.

Evidence for validity of conclusions made on the basis of inferences drawn from projective methods needs to be convincing. Absolute standards that cut across all methods are essential and they must be precise enough to inform judgments about when the standard is or is not being met. However, in their application to projective techniques, general criteria should accommodate to the nature of the method. Validity and reliability are not established for a generic technique such as figure drawings, the TAT, or the Rorschach. Rather, the utility of each set of instructions and stimuli together with the scoring method is separately evaluated according to the accomplishment of its specific purpose. Ways of demonstrating validity and reliability for all techniques must be in tune with the logic and coherence of the measures and constructs under consideration.

4.16.13.1 Reliability

Establishing reliability for projective techniques and for questionnaire methods is fundamentally different. As Kelly (1958) pointed out: "When the subject is asked to guess what the examiner is thinking, we call it an objective test; when the examiner tries to guess what the subject is thinking, we call it a projective device" (p. 332). Reliability of items on a rating scale is essentially a matter of consistency in the respondent's interpretation of the items. Reliability of inferences based on projective methods rests on a combination of the stimulus, the response, the method of interpretation, and the skill of the interpreter. Psychometric procedures to establish reliability based on forced-choice test items can be carried out without any clinical expertise because reliability pertains primarily to the test-takers responses. With projective measures, reliability, in part, is an attribute of the interpreter because the scoring units are inferences of the professional. Attempts to designate cookbook procedures so that the validation process can be carried out quickly and easily are out of tune with the nature of projective methods.

The following is a brief overview of how traditional psychometric indicators must be modified to establish reliability of projective techniques:

4.16.13.1.1 Scorer reliability

The accuracy of two people looking for the same information requires adequate training and clear guidelines. When raters are well-trained and scoring systems are well documented, such reliabilities tend to be high. A related aspect of scorer reliability, particularly relevant to clinical inference is the consistency with which one rater codes the same protocol over time (Karon, 1981).

4.16.13.1.2 Decisional reliability

The consistency of decisions drawn from a protocol can be assessed when specific units are not the focus. For example, Shneidman (1951) showed that 16 clinicians using their own methods came to similar conclusions. Other influences on the reliability of decisions may relate to the number of performance samples needed. For example, in the measurement of particular motives with TAT cards, it is important to know if the specific motive emerges reliably in every card or only some. To assure adequate reliability, it has been recommended that at least six stories be obtained from each respondent (Lundy, 1985; Smith, 1992b).

Decisions must be based on clear conceptualizations. Therefore, decisional reliability must account for the theoretical appropriateness of the match between predictor and criteria. Test responses can be expected to correlate with criteria only if their meaning is functionally similar. Thus, adequate functioning in a structured situation can occur despite disorganized responses on a projective test.

4.16.13.1.3 Test–retest reliability

One important factor in demonstrating the reliability of the measure upon retesting is whether one is looking for similarity of content or of the inference (Karon, 1981). The reliability of the specific content is less relevant than consistency in the meaning of the response.

4.16.13.1.4 Internal consistency

Internal consistency of thematic content across cards would be an inappropriate measure of reliability because cards are designed to elicit different themes (Lundy, 1985). However, despite variability in specific content, the responses may yield similar inferences. More stylistic units representing psychological processes such as the accuracy with which the content captures the "gist" of the scene presented or linkages between causes and effects may be generalizable across different stimuli. Internal consistency based on the number of words per story (alpha = 0.96) was much higher than internal consistency of need for achievement (Atkinson, Bongort, & Price, 1977). Likewise, Rorschach cards present stimuli with important differences and, rather than estimating internal consistency, inferences are made on the basis of the entire coded protocol and not card by card. Responses to the white spaces in the blot have different meaning for each card, and not all white space responses represent figure ground reversal. Therefore, adding all space responses is a rough estimate of the trait in question.

Internal consistency expected within a measure should relate to the nature of the task and to the consistency inherent in the construct under consideration.

4.16.13.2 Construct Validation

Messick (1989) defines construct validity as "an integration of any evidence that bears on the interpretation or meaning of test scores" (p. 17). Because traditional indices of content or criterion validity contribute to the meaning of test scores, they too pertain to construct validity. Thus, construct validity subsumes all other forms of validity evidence.

This emphasis on construct validity as the overriding focus in test validation represents a shift from prediction to explanation as the fundamental focus of validation efforts. Construct validation emphasizes the development of models explaining processes underlying performance on various tests and their relationships to other phenomena. Accordingly, correlations between test scores and criterion measures contribute to the construct validity of both predictor and criterion. In Messick's words, "Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on test scores or other modes of assessment" (p. 13).

4.16.13.3 Multitrait, Multimethod Validation

Although the multitrait, multimethod approach is a standard technique for construct validation (Campbell, 1960; Campbell & Fiske, 1959), this technique is vulnerable to problems with the definition and measurement of constructs. This procedure seeks to establish higher correlations across diverse measures of the same trait (convergent evidence) and lower correlations among similar measures of different traits (discriminant evidence) to show that a construct is distinct from other constructs and that it is not uniquely tied to a particular measurement method. Yet, the possibility that the constructs may indeed be tied to the measurement tool must be considered. McClelland and colleagues showed that attempts to correlate self-report and projective measures of achievement motivation to validate either one are misguided. Likewise, Horowitz (1991) notes that schema, although important to measure, are not directly known by the subject and cannot be confirmed through self-report but by a consensus of two or more independent observers. Theoretical understanding of the construct must always be central to the choice of validation efforts.

It is important to acknowledge the potential for corroboration of constructs with various assessment measures. Yet, different measures of the same construct (e.g., ratings by actor versus observer) may be assessing qualitatively different aspects of that construct. Meaningful differences in the constructs are demonstrated by distinct patterns of relationships with various criterion measures (see criterion validation below).

Even highly similar measures (e.g., thematic approaches) used to assess identical constructs but with different procedures are not necessarily addressing the same qualities. For example, Arnold (1962) developed criteria for scoring TAT stories for achievement motivation by comparing groups of individuals known to differ in their job success. In contrast, the McClelland–Atkinson tradition (McClelland & Koestner, 1992) for developing criteria for scoring achievement motivation from stories told to picture stimuli was based on the comparison of groups given different instructional sets to arouse the achievement motive. Each of these two approaches to contrasting groups may be appropriate for given purposes with specific populations but clearly involve distinct conceptualizations of the achievement motive. One takes the position that the motive is present and, when aroused, energizes and directs behavior towards a particular class of goals or incentives. The other focuses on the achievement motive as being a relatively stable disposition reflected by complex patterns of cognitive–emotional processes that guide perceptions and behavior. If scoring units are empirically derived through contrasting groups, then the construct is defined by the nature of the groups and procedures used.

4.16.13.4 Criterion Validation

Criterion related evidence also belongs under the rubric of construct validity (Messick, 1989). Patterns of correlations with other variables contribute to the defining attributes of each construct. The same set of external correlates do not apply to achievement motivation measured through projective and self-report methods (Spangler, 1992). Projective measures of achievement motivation correlate with spontaneous effort sustained over time, whereas self-reports correlate with activities that are cued by the situation. Such variation in the match between the predictor and the target clarify the dimensions of the achievement motive assessed by each type of measure. This understanding of constructs in terms of the life conditions to which they generalize constitutes the simultaneous validation of the predictor and the criteria. Relationships between predictors and criteria derived from a construct must be tested for expected patterns of discriminant and convergent evidence. In doing so, both situational specificity and generalizability of constructs and their measures are addressed at the same time. Generalizability of the predictor–criterion relationships across different population groups, settings, task domains, ages, or gender must be empirically established.

4.16.13.5 Part–Whole Relationships

Personality is a functional whole comprised of patterns of interrelated components. Therefore, predictions from single elements of personality to more encompassing units, such as general adjustment, proceed from part to whole and can only account for a portion of the variance. This part–whole prediction is illustrated in Figure 1 showing the influence of temperament on adjustment as mediated by goodness-of-fit. Any one temperamental attribute would account for a small proportion of the variance in the fit between the person and various situational contexts or task demands. A configuration of such traits would improve the prediction. However, the actual goodness-of-fit would be a better predictor of adjustment because this broader construct subsumes the configuration of traits (even those not specifically addressed in the prediction), the environmental demands and supports, as well as coping mechanisms.

This understanding of the part–whole relationship is essential for the validation of projective techniques. Projective techniques look at individual differences at a more global level than other measures. The products reflect schema or inner structures that develop through the individual's synthesis of life experiences. These involve the interplay of all of the person's characteristics and reciprocal transactions with the environment. Therefore, part–whole patterns exist between various trait constructs and response parameters within the test. Part–whole predictions can relate specific neuropsychological or temperamental processes to individual differences in responses to projective tasks (Bassan-Diamond, Teglasi, & Schmitt, 1995). Part–whole relationships depend on conceptualizations about how the various dimensions of individual differences come together in the development of broader units of personality such as the influence of temperament on conscience development (Kochanska, 1993). Given the expected role of multiple variables, it is apparent that expectations about the size of correlations need to be in line with understanding of the phenomena.

The part–whole issue also applies to the relationships among responses within a specific projective technique. The functional significance of each part hinges on its interplay with other parts in relation to the constructs being considered. Various dimensions of the product such as form and content variables are cohesively related and need to be systematically coordinated into higher-order constructs. Psychometric and theory-based approaches need to be integrated to build useful frameworks for clinical use. Specific response parameters can be abstracted from the whole and pieced together to find conceptually meaningful patterns.

4.16.13.6 Convergence of Psychometric and Conceptual Treatment of Data

Normative, nomothetic, and case study approaches to data collection are interrelated. Conceptual linkages among these approaches are essential to the interpretation and validation of projective techniques, as described next.

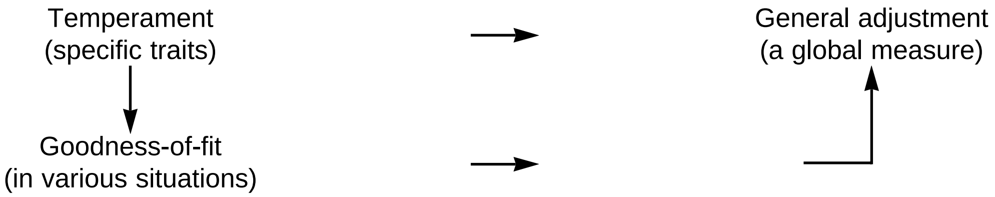

Figure 1 Illustration of part–whole prediction.

4.16.13.6.1 Normative

Before initiating the arduous process of collecting norms for projective methods, the units of analysis must be conceptually meaningful, theoretically relevant, and clinically useful. Furthermore, criteria for coding must be provided in sufficient detail to assure rater reliability. The Comprehensive System for the Rorschach has shown that the establishment of norms for meaningful units of analysis is possible, and that larger conceptual units can be designated through the use of multiple cutoffs to establish interpretively useful patterns of interrelated response parameters.

4.16.13.6.2 Nomothetic

It is important to base norms on units of interpretation that represent psychological processes that can serve as explanatory constructs. The nomothetic approach attempts to promote conceptual understanding through the study of general principles and functional relationships among variables. Such a conceptual approach to the Rorschach (Blatt & Berman, 1984; Weiner, 1977) has advocated organizing discrete response parameters into theoretically relevant clusters that contribute to a well defined construct. A strict empiricist would be satisfied with the prediction of patterns of behavior from various configurations of test responses (e.g., of Rorschach variables). However, establishing such empirical relationships is only a starting point for building conceptual frameworks for understanding and explaining behavior. The meaning of test response patterns emerges from a network of relationships with other theoretically relevant response patterns and theoretically appropriate external criteria.

The collection of normative data for projective techniques is embedded in the construct validation process. The constructs and units of inference for projective measures need to be defined *a priori* in the same way that test items in questionnaires are selected to represent constructs. An example of this approach is the designation of units of inference for TAT stories on the basis of constructs derived from research and theory on empathy (Locraft & Teglasi, 1997). Scores based on these units subsequently differentiated groups of children designated as high, medium, and low on empathy on the basis of teacher ratings. Such units require cross validation on numerous populations to assure wide applicability prior to attempting large scale normative studies. Unless the unit of inference is conceptually clear and represents meaningful psychological processes, normative data or group comparisons are not particularly useful.

4.16.13.6.3 Case study

The case study is most relevant to the practicing clinician because decisions for clinical purposes relate to understandings about one person. The adequacy of these decisions can be validated on a case by case basis. Although case study methods focus on the unique patterns of variation in responses of a single individual, the pattern of variables under consideration can be referenced to norms and to theoretical constructs. With projective methods, the respondent is allowed free expression within a specified context (stimuli and directions), and effective use of these techniques require a conceptual framework at three levels. First is an understanding of the psychological meaning of the response patterns within the projective measure. Second is an understanding of how the projective measure fits with other information in a comprehensive evaluation. Third is a conceptualization of how the patterns relate to competencies needed in the relevant life situations. Therefore, to conduct adequate case studies, the examiner needs a framework to understand how various types of data fit together within and across measures and how the emerging patterns apply to the various demands of the individual's environment. The attempt to integrate psychometric (empirical) and nomothetic (theoretical) data into the case study is typified in the attempt to synthesize structural and content features of Rorschach responses (Erdberg, 1993). The problem is that these efforts have applied different conceptual frameworks to different aspects of the data. A true integration of the empirical and theoretical perspectives requires a focus on personality constructs and the establishment of coherent patterns of data from various sources around the constructs.

The term "conceptual validity" has been applied to the process of psychological assessment (Maloney & Ward, 1976). Whereas construct validity focuses on confirming expected patterns of relationships across individuals, conceptual validity focuses on observing cohesive patterns within an individual. These patterns of expected relationships among observations constitute a working model of the individual being evaluated. According to Maloney and Ward, establishing such a model is a prerequisite for answering the referral question. When information from various sources is understood in terms of the constructs that explain an individual's difficulties and point to appropriate decisions, the assessment has conceptual validity.

4.16.14 FUTURE DIRECTIONS

The anticipation of future directions for projective techniques emerges from the clues gleaned from the scientific literature pertaining to the development and assessment of personality. However, the prognosticator's wishful thinking also influences this process. The intent of the prognosticator is to show how unfolding trends or even new spins on old ideas point toward desired alternatives for moving forward. Future possibilities for projective techniques are drawn from three conceptually distinct perspectives: (i) converging constructs from the various subfields of psychology that scaffold the use of projective techniques; (ii) refinements in the conceptualization of personality that point to a multidimensional view and to a unique role for projective techniques; and (iii) improved understanding and use of the specific projective tools.

Methods and constructs in other subfields of psychology provide a hospitable *zeitgeist* for projective techniques and bode well for their future development and utility. Converging evidence appears to validate the basic tenets of the projective hypothesis. Concepts from psychodynamic formulations such as transference have been redefined in terms of contemporary psychology, demonstrating that such phenomena are not unique to one theory but can be understood in several ways (Singer & Singer, 1994). The examination of clinically relevant phenomena in light of conceptualizations of memories, schema, and scripts from various subdisciplines support the work in each subfield (e.g., Horowitz, 1991; Stein & Young, 1992). Cognitive theories of perception, memory, and learning are increasingly emphasizing the unconscious or implicit social attitudes (Uleman & Bargh, 1989). Research on construct availability and accessibility (Higgins, 1990) as well as script or schema theory also recognizes the influence of cognitive processes that occur outside of awareness. Knowledge from other areas of psychology will increasingly inform the designation of interpretive units for measuring personality with projective techniques. Recent interest in narrative methods to assess schema and social cognitions is applicable to thematic apperceptive techniques (Cramer, 1996). Stories elicited through picture stimuli, like other narrative approaches, reflect the human orientation to perceive the world as stories or myths grounded in culture and personal experience.

Conceptions of psychopathology within psychoanalytic theory are placing increasing emphasis on functions of inner psychic structures and subjective meaning systems (Atwood & Stolorow, 1984). In this context, adjustment problems arise because the structures are insufficiently developed to permit adaptive coping rather than from conflict among the structures. In an analogous manner, clinical syndromes have been related to problems in the development of person schema (Horowitz, 1991). If psychopathology is a reflection of impairment in the formation of psychic structures, then it is reasonable for assessment to be geared to the evaluation of these structures (e.g., object relations, self-system, along with associated processes such as reality testing).

Levels of impairment based on the complexity and organization of inner structures will be identified along with a description of symptomatic behaviors. Projective techniques will figure prominently in the assessment of psychological variables for the purpose of intervention planning. The continued refinement of integrative therapies (Leve, 1995; Norcross, 1986) will increase the usefulness of identifying relevant processes to be targeted and matched with optimal intervention strategies.

Understanding personality according to different styles and levels of organization of experience as they relate to functioning will rest on increasing emphasis on part whole conceptualizations. For example, with a focus on inner structures, aspects of personality such as emotions will be increasingly recognized as organizing processes that shape adaptation and problem solving (Greenberg, Rice, & Elliott, 1993). Other personality processes, such as distractibility or inattention, will also be understood in relation to how they shape the development of inner psychic structures. Future research will focus on establishing linkages between various levels of personality constructs such as the role of various traits in the development of inner strivings and the organization of meaning structures. Research will also focus on clarifying relationships between discrete measures of neuropsychological processes, such as planning, organizing, retrieval from memory, or continuous attention, and the manifestation of these processes in the performance of more complex tasks analogous to real-life situations. Inherent in the part whole conceptualization of personality is the view of development as being propelled simultaneously by biological, psychological, and social forces. Within the psychological realm, perceptions and behaviors will be increasingly understood in terms of the interplay of experience with relevant affective and cognitive processes.

A multilevel understanding of personality will permit clearer conceptualizations of the utility of different types of techniques to assess different facets of personality. One example is the distinction between measures of self-attributed and implicit motives to achieve (McClelland et al., 1989). Story-telling measures of achievement motivation (implicit) were more effective in the prediction of long term outcomes such as career success, whereas self-reports (self-attributed) were better in predicting immediate choices (Spangler, 1992). The two types of measures of achievement motivation relate to different criteria, develop through different pathways, and can be understood in terms of different levels of personality.

In training programs, the professional courses are the places where the core psychological knowledge bases are integrated. Expertise in projective techniques includes the linking of the psychological processes assessed with appropriate strategies for therapeutic interventions. The complexity of interpreting projective instruments and of understanding their implications for guiding interventions requires no less than the systematic and flexible application of prior knowledge. Therefore, projective assessment can serve as a centerpiece for integrating concepts from core areas such as physiological, affective and cognitive bases of behavior. Training should emphasize the gradual development of the professional's schema to guide practice rather than the acquisition of knowledge applied in rote fashion. Trainees' schema should be sufficiently broad and flexible not only to incorporate information currently available but to accommodate continuous learning throughout the professional career. Measures are simply tools. Rather than emphasizing the assessment technique, training programs would do well to focus on the development of integrative frameworks that include how various sources of information fit together into meaningful patterns.

Increased acceptance of qualitative measures as having scientific merit has led to more frequent use of open-ended methods in the study of personality such as the study of early memories (Bruhn, 1992), thought sampling (Rubin, 1986), examination of the life story (McAdams, 1990), self-defining memories (Moffit & Singer, 1994), and analysis of therapy transcripts (Luborsky & Crits-Christoph, 1990). The psychometric challenges of projective methods are shared among all open-ended techniques. The effort to master these difficulties will require, above all, conceptual clarity and an emphasis on construct validation. These endeavors will promote shared conceptualizations and methods across disciplines, appreciation of the progress already made with projective techniques, and the spurring of further developments.

Frameworks for interpretation of projective techniques not only elucidate the inner structure of personality but reveal problem-solving strategies in relation to the task demands. Therefore, it may be fruitful to view these techniques, in part, as *performance measures of personality* that can be compared to other tests in an assessment battery in terms of what the task requires. The most significant difference is that performance tasks of personality maximize the use of spontaneous strategies to organize perceptions and responses. Variation in response patterns across different types of tasks permits distinctions between knowledge structures that are inert and require external prompts from those that are meaningfully organized and spontaneously accessible. Furthermore, the configuration of responses within and across diverse tasks foster explanations of situational variability and consistency in performance and behavior. For instance, social situations require the individual to organize perceptions, size up intentions, and anticipate reactions of others. These requirements are similar to the demands of projective tests to organize perceptions of the stimuli and coordinate the responses with the instructions.

The relationship between content and formal elements in the units of inference drawn from projective techniques still needs to be clarified. The most fruitful approach to integrating the analysis of form and content is through the identification of their linkages with important psychological processes. Formal structural analysis should receive greater emphasis in the apperceptive methods. Generalized properties of content and features of narrative structure are more durable indices of psychological processes than isolated content. Furthermore, the analysis of the subtext or underlying structure of narratives can yield significant information about the individual's schematic organization of experience. The struggle to apply valid and reliable methods to assessment will continue. The inevitable link between theory and method necessitates the simultaneous effort to refine both. Conceptualizations developed in other subfields of psychology and lessons learned from qualitative research methods will be applied to improve the reliability, validity, and clinical utility of projective measures.

4.16.15 SUMMARY

The chapter is summarized by presenting a model for understanding performance measures of personality shown in Figure 2. The interaction of the person and environment (boxes 1 and 2) as a fundamental unit of study has wide acceptance and applicability. However, the objective features of the environment and the subjective world are not the same. Lewin (1935) refers to the psychological environment as the inner experience of external reality. Persons participate in and influence their external environments as well as being shaped by them. The manner in which the individual encodes and stores encounters with the external world drives the construction of the inner psychological world (box 3). Therefore, one can only conceptually separate the person from the environment—they are embedded in one another.

Schema are memory structures that develop through the individual's constant and reciprocal transactions with the environment. The development of schema is influenced by (all factors said to influence personality) various interactive determinants including constitutional factors of temperament (see reviews by Emde, 1989; Plomin, 1986) as well as meaning structures transmitted by family and culture in conjunction with the individual's experiences throughout life. These schema are important to assess because they: (i) are mental sets that guide attentional processes and serve as filters for interpreting information; (ii) guide actions in situations that require complex resources; and (iii) shape the storage of new information in memory.

Certain conditions such as attention-deficit hyperactivity disorder are diagnosed through a careful examination of previous history as well as scrutiny of current behavioral patterns. However, these attentional processes shape the individual's prior and ongoing synthesis of experience through their impact on the information entering awareness, the feedback received from others and on the development of internalized schema. Therefore, the part–whole relationships between specific symptoms or psychological processes and larger units of the personality need to be understood. Projective testing relies on an appreciation that overt behavior is linked to inner meaning. Consequently, inner life is a mechanism for perceiving the outer world and for regulating behavior to adapt to these perceptions. The projective techniques (e.g., TAT and Rorschach) provide information about how prior experience has been organized and how inner resources (schema, inner structures) are applied to relatively unstructured situations. Responses to projective tasks are understood as cohesive products reflecting all of the processes involved in the synthesis of experience. Specific psychological processes such as attention or emotion can be studied as precursors and sequelae of the development of inner structures.

Projective techniques require a person to demonstrate the qualities of organization and strategic planning by performing a problem-solving task rather than by telling about the self in an interview or by responding to a questionnaire. Therefore, the term *performance measure of personality* aptly describes projective methods. These performance measures are less structured than the other tasks in the typical assessment battery and provide unique information. The manner in which the problem is solved

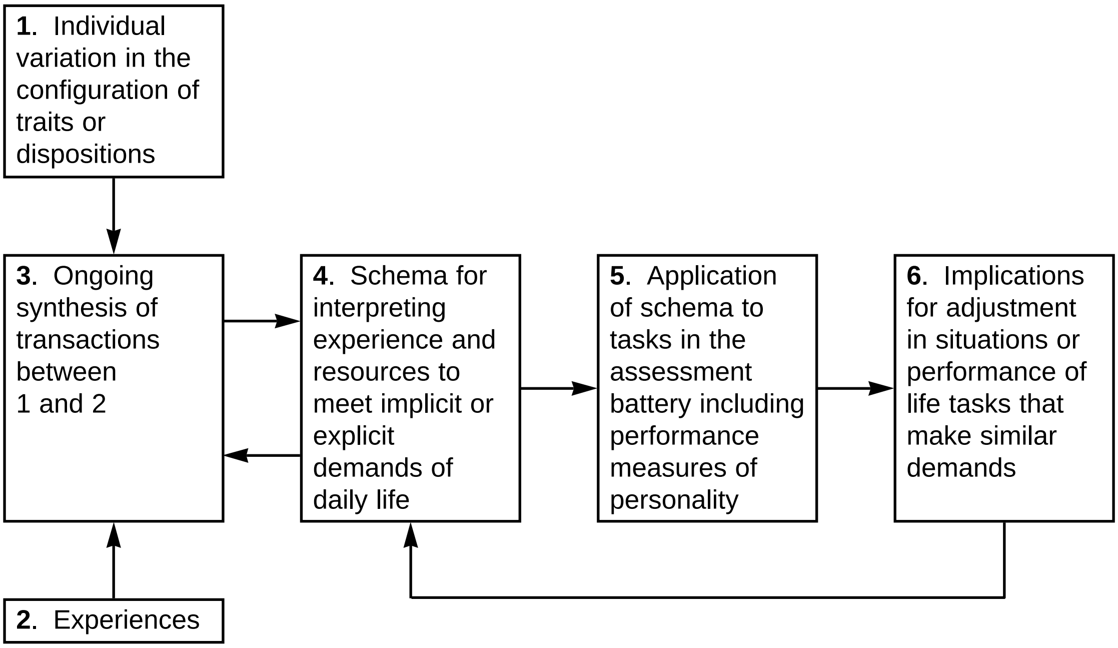
1. Individual variation in the configuration of traits or dispositions
3. Ongoing synthesis of transactions between 1 and 2
4. Schema for interpreting experience and resources to meet implicit or explicit demands of daily life
5. Application of schema to tasks in the assessment battery including performance measures of personality
6. Implications for adjustment in situations or performance of life tasks that make similar demands
2. Experiences

Figure 2 Understanding performance measures of personality.

or the task is accomplished reveals the individual's organizational strategies and resources to deal with similarly unstructured situations. In addition, the nature and organization of knowledge structures reveal the frameworks or convictions that guide responses in ambiguous situations. Furthermore, inferences drawn from performance measures and self-reports pertain to different facets of personality and predict different criteria (McClelland et al., 1989).

As Figure 2 shows (box 4), the individual's ongoing synthesis of experience leads to the development of knowledge structures. Various performance measures in an assessment battery can be examined in terms of what the task requires and what inner structures or resources are brought to bear on the performance (box 5). Actual adjustment in a given situation (box 6) depends not only on the person's resources but also hinges on environmental expectations and supports.

The person is assessed within a particular social context of the testing situation including expectations about testing, anticipation of the potential consequences of testing, and the relationship with the examiner. Therefore, these variables are considered along with the task demands. The job of the examiner is to organize multiple sources of data including reports of self and others, various performance measures, relevant past history, and current circumstances into a framework for understanding the referral issues. All of the information is sifted through an understanding of family and cultural background as pertains to the individual being assessed. The understanding is informed not only by theories of personality, human development, and psychopathology, but also by a conceptualization of the various tasks in the battery and how they relate to performance in life situations. Findings are not a listing of isolated facts but conclusions based on cohesive patterns that provide an understanding of the problem and point to appropriate action.

4.16.16 REFERENCES

Abelson, R. P. (1976). Script processing in attitude formation and decision making. In J. S. Carroll & J. W. Payne (Eds.), *Cognition and social behavior*. Hillsdale, NJ: Erlbaum.

Abelson, R. P. (1981). Psychological status of the script concept. *American Psychologist, 36,* 715–729.

Acklin, M. W. (1994). Some contributions of cognitive science to the Rorschach Test. In I. B. Weiner (Ed.), *Rorschachianna* (pp. 129–145). Seattle, WA: Hogrefe & Huber.

Alper, T. G., & Greenberger, E. (1967). Relationship of picture structure to achievement motivation of college women. *Journal of Personality and Social Psychology, 7,* 362–371.

Anastasi, A. (1976). *Psychological testing*. New York: Macmillan.

Anderson, J. R. (1990). *Cognitive psychology and its implications*. New York: Freeman.

Archer, R. P., Maruish, M., Imhof, E. A., & Piotrowski, C. (1991). Psychological test usage with adolescent clients: 1990 survey findings. *Professional Psychology: Research and Practice, 22,* 247–252.

Arnold, M. B. (1962). *Story sequence analysis: A new method of measuring motivation and predicting achievement*. New York: Columbia University Press.

Atkinson, J. W. (1992). Motivational determinants of thematic apperception. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 21–48). New York: Cambridge University Press.

Atkinson, J. W., Bongort, K., & Price, L. H. (1977). Explorations using computer simulation to comprehend TAT measurement of motivation. *Motivation and Emotion, 1,* 1–27.

Atwood, G., & Stolorow, R. (1984). *Structures of subjectivity*. Hillsdale, NJ: Analytic Press.

Auld, F., Jr. (1954). Contributions of behavior theory to projective testing. *Journal of Projective Techniques, 18,* 421–426.

Bargh, J. A. (1989). Conditioned automaticity: Varieties of automatic influence on social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 3–51). New York: Guilford Press.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 1–40). Hillsdale, NJ: Erlbaum.

Barkley, R. A. (1990). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment*. New York: Guilford Press.

Bassan-Diamond, L. E., Teglasi, H., & Schmitt, P. (1995). Temperament and a story-telling measure of self-regulation. *Journal of Research in Personality, 29,* 109–120.

Baumeister, R. F., Wotman, S. R., & Stillwell, A. M. (1993). Unrequited love: On heartbreak, anger, guilt, scriptness, and humiliation. *Journal of Personality and Social Psychology, 64,* 377–394.

Beck, S. J. (1981). Reality Rorschach and perceptual theory. In A. I. Rabin (Ed.), *Assessment with projective techniques* (pp. 23–46). New York: Springer.

Bellak, L. (1975). *The T.A.T., C.A.T., and S.A.T. in clinical use*. New York: Grune & Stratton.

Bellak, L. (1993). *The T.A.T., C.A.T., and S.A.T. in clinical use* (5th ed.). Needham Heights, MA: Allyn & Bacon.

Birney, R. C. (1958). Thematic content and the cue characteristics of pictures. In J. W. Atkinson (Ed.), *Motives in fantasy, action, and society* (pp. 630–643). New York: Van Nostrand.

Blatt, S. J. (1991). A cognitive morphology of psychopathology. *Journal of Nervous and Mental Disease, 179,* 449–458.

Blatt, S. J. (1992). The Rorschach: A test of perception on an evaluation of representation. In E. I. Megargee & C. D. Spielberger (Eds.), *Personality assessment in America: A retrospective on the occasion of the fiftieth anniversary of the Society for Personality Assessment* (pp. 160–169). Hillsdale, NJ: Erlbaum. (Reprinted from 1990 *Journal of Personality Assessment, 55,* 394–416).

Blatt, S. J., & Berman, W. H., Jr. (1984). A methodology for the use of the Rorschach in clinical research. *Journal of Personality Assessment, 48,* 226–239.

Blatt, S. J., Ford, R. Q., Berman, W., et al. (1988). The assessment of therapeutic change in schizophrenic and borderline young adults. *Psychoanalytic Psychology, 5,* 127–158.

Blatt, S. J., & Lerner, H. (1983). The psychological assessment of object representations. *Journal of Personality Assessment, 47,* 7–28.

Blatt, S. J., & Wild, C. M. (1976). *Schizophrenia: A developmental analysis.* New York: Academic Press.

Blatt, S. J., Allison, J., & Feirstein, A. (1969). The capacity to cope with cognitive complexity. *Journal of Personality, 37,* 269–288.

Bransford, J. D., Franks, J.J., Vye, N. J., & Sherwood, R. D. (1989). New approaches to instruction: Because wisdom can't be taught. In S. Voshiadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 470–497). Cambridge, UK: Cambridge University Press.

Bruhn, A. R. (1992). The early memories procedure: A projective test of autobiographical memory, Part 2. *Journal of Personality Assessment, 58,* 326–346.

Bruner, J. S. (1986). *Actual minds, possible worlds.* Cambridge, MA: Harvard University Press.

Buck, J. N. (1948). The H–T–P technique: A quantitative and qualitative scoring manual. *Clinical Psychological Monographs, 5,* 1–120.

Buck, J. N. (1987). *The House-Tree-Person technique: Revised manual.* Los Angeles: Western Psychological Services.

Burns, R. C. (1987). *Kinetic-House-Tree-Person drawings (KHTP).* New York: Brunner/Mazel.

Burns, R., & Kaufman, S. (1970). *Kinetic Family Drawings (K-F-D): An introduction to understanding children through kinetic drawings.* New York: Brunnel/Mazel.

Burns, R. C., & Kaufman, H. S. (1972). *Actions, styles, and symbols in Kinetic Family Drawings: An interpretive manual.* New York: Brunner/Mazel.

Butcher, J. N., & Rouse, S. V. (1996). Personality: Individual differences and clinical assessment. In J. T. Spence, J. M. Darley, & D. J. Foss (Eds.), *Annual Review of Psychology, 47,* 87–111.

Campbell, D. T. (1960). Recommendations for APA test standards regarding construct, trait or discriminant validity. *American Psychologist, 15,* 546–553.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Campus, N. (1976). A measure of needs to assess the stimulus characteristics of TAT cards. *Journal of Personality Assessment, 40,* 248–258.

Cantor, N., & Kihistrom, J. F. (1987). *Personality and social intelligence.* Englewood Cliffs, NJ: Prentice-Hall.

Carlson, L., & Carlson, R. (1984). Affect and psychological magnification. Deviations from Tomkins' script theory. *Journal of Personality, 52,* 36–45.

Chi, M., Glaser, R., & Farr, M. (1987). *Nature of expertise.* Hillsdale, NJ: Erlbaum.

Clore, G. L., Schwarz, N., & Conway, M. (1994). Affective causes and consequences of social information processing. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (pp. 323–417). Hillsdale, NJ: Erlbaum.

Cooper, A. (1981). A basic TAT set for adolescent males. *Journal of Clinical Psychology, 37,* 411–414.

Cramer, P. (1996). *Storytelling narrative and the Thematic Apperception Test.* New York: Guilford Press.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social-information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115,* 74–101.

Cronbach, L. J. (1970). *Essentials of psychological testing* (3rd ed.). New York: Harper & Row.

Demorest, A. P., & Alexander, I. E. (1992). Affective scripts as organizers of personal experience. *Journal of Personality, 60,* 645–663.

Dodge, K. A., & Feldman, E. (1990). Issues in social cognition and sociometric status. In S. R. Asher & J. D. Coie (Eds.), *Peer rejection in childhood* (pp. 119–155). New York: Cambridge University Press.

Durand, V. M., Blanchard, E. B., & Mindell, J. A. (1988). Training in projective testing: Survey of clinical training directors and internship directors. *Professional Psychology: Research and Practice, 19,* 236–238.

Emde, R. N. (1989). The infant's relationship experience: Developmental and affective aspects. In A. J. Sameroff & R. N. Emde (Eds.), *Relationship disturbances in early childhood: A developmental approach.* New York: Basic Books.

Epstein, S. (1994). Integration of the cognitive and psychodynamic unconscious. *American Psychologist, 49,* 709–724.

Erdberg, P. (1993). The U.S. Rorschach scene: Integration and elaboration. In I. B. Weiner (Ed.), *Rorschachianna XIX: Yearbook of the International Rorschach Society* (pp. 139–151). Seattle, WA: Hografe & Huber.

Exner, J. E. (1989). Searching for projection in the Rorschach. *Journal of Personality Assessment, 53,* 520–536.

Exner, J. E. (1993). *The Rorschach: A comprehensive system: Vol. 1. Basic processes.* New York: Wiley.

Exner, J. E., & Weiner, I. B. (1995). *The Rorschach: A comprehensive system: Vol. 3. Assessment of children and adolescents.* New York: Wiley.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50,* 229–238.

Feher, E., VandeCreek, L., & Teglasi, H. (1983). The problem of art quality in the use of the Human Figure Drawing Test. *Journal of Clinical Psychology, 39,* 268–275.

Festinger, L. (1957). *A theory of cognitive dissonance.* Evanston, IL: Row, Peterson.

Fisher, S. (1986). *Development and structure of the body image.* Hillsdale, NJ: Erlbaum.

Fiske, A. P., Haslam, N., & Fiske, S. T. (1991). Confusing one person with another: What errors reveal about the elementary forms of social relations. *Journal of Personality and Social Psychology, 60,* 656–674.

Fiske, S. T., & Taylor, S. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.

Frank, L. D. (1939). Projective methods for the study of personality. *Journal of Psychology, 8,* 389–413.

Frank, L. D. (1948). *Projective Methods.* Springfield, IL: Thomas.

Fromkin, H. L., & Streufert, S. (1976). Laboratory experiments. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 415–465). Chicago: Rand McNally.

Goodenough, F. L. (1926). *Measurement of intelligence by drawings.* New York: World Books.

Greenberg, L. S., Rice, L. N., & Elliott, R. (1993). *Facilitating emotional change: The moment by moment process.* New York: Guilford Press.

Groth-Marnat, G. (1990). *Handbook of psychological assessment* (2nd ed.). New York: Wiley.

Hammer, E. F. (1958). *The clinical application of projective drawings.* Springfield, IL: Charles C. Thomas.

Handler, L. (1985). The clinical use of the Draw-A-Person Test (DAP). In C. S. Newmark (Ed.), *Major psychological assessment instruments.* Newton, MA: Allyn & Bacon.

Handler, L., & Habernicht, D. (1994). The Kinetic Family Drawing Technique: A review of the literature. *Journal of Personality Assessment, 62,* 440–464.

Harris, D. B. (1963). *Children's drawings as measures of intellectual maturity.* New York: Harcourt Brace Jovanovich.

Hartman, A. A. (1970). A basic TAT set. *Journal of Projective Techniques, 34,* 391–396.

Haynes, J. P., & Peltier, J. (1985). Patterns of practice with TAT in juvenile forensic settings. *Journal of Personality Assessment, 49,* 26–29.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Henry, W. E. (1956). *The analysis of fantasy: The Thematic Apperception Test in the study of personality*. New York: Wiley.

Higgins, E. T. (1990). Personality, social psychology, and person-situation relations: Standards and knowledge activism as a common language. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research,* (pp. 301–338). New York: Guilford Press.

Hogan, R. (1987). Personality psychology: Back to basics. In J. Aronoff, A. J. Rabin, & R. A. Zucker (Eds.), *The emergence of personality* (pp. 79–105). New York: Springer.

Holt, R. R. (1958). Formal aspects of the TAT—A neglected resource. *Journal of Projective Techniques, 22,* 163–172.

Holt, R. R. (1961). The nature of TAT stories as cognitive products: A psychoanalytic approach. In J. Kagan & G. Lesser (Eds.), *Contemporary issues in thematic apperceptive methods* (pp. 3–40). Springfield, IL: Charles C. Thomas.

Holt, R. R. (1978). *Methods in clinical psychology: Vol. 1. Projective assessment*. New York: Plenum.

Holt, R., & Luborsky, L. (1958). *Personality patterns of psychiatrists: A study of methods for selecting residents*. New York: Basic Books.

Holtzman, W. H. (1981). Holtzman Inkblot Technique. In A. I. Rabin (Ed.), *Assessment with projective techniques* (pp. 47–83). New York: Springer.

Holtzman, W. H., Thorpe, J., Swartz, J., & Heron, E. (1961). *Inkblot perception and personality: Holtzman Inkblot Technique*. Austin, TX: University of Texas Press.

Horowitz, M. J. (1991). *Person schema and maladaptive interpersonal patterns*. Chicago: University of Chicago Press.

Ingram, R. E. (Ed.) (1986). *Information processing approaches to clinical psychology*. New York: Academic Press.

Isen, A. M. (1984). Toward understanding the role of affect in cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 3, pp. 179–236). Hillsdale, NJ: Erlbaum.

Isen, A. M. (1987). Positive affect, cognitive processes, and social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 20, pp. 203–253). New York: Academic Press.

Johnston, M. H., & Holzman, P. S. (1979). *Assessing schizophrenic thinking*. San Francisco: Jossey-Bass.

Kahana, B. (1978). The use of projective techniques in personality assessment of the aged. In M. Storandt, I. Siegler, & M. Elias (Eds.), *The clinical psychology of aging*. New York: Plenum.

Kahill, S. (1984). Human figure drawings in adults: An update of the empirical evidence, 1962–1982. *Canadian Psychology, 25,* 269–292.

Karon, B. P. (1981). The Thematic Apperception Test (TAT). In A. I. Rabin (Ed.), *Assessment with projective techniques: A concise introduction* (pp. 85–120). New York: Springer.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivations* (Vol. 15). Lincoln, NE: University of Nebraska Press.

Kelly, G. A. (1958). The theory and technique of assessment. In P. R. Farnsworth & Q. McNemar (Eds.), *Annual Review of Psychology, 9,* 323–352. Palo Alto: Annual Reviews.

Kenny, D. T. (1964). Stimulus functions in projective techniques. In B. A. Maher (Ed.), *Progress in experimental personality research* (pp. 285–354). New York: Academic Press.

Kenny, D. T., & Bijou, S. W. (1953). Ambiguity of pictures and extent of personality factors in fantasy responses. *Journal of Consulting Psychology, 17,* 283–288.

Kihlstrom, J. F. (1984). Conscious, subconscious and preconscious: A cognitive perspective. In K. S. Bower & D. Michenbaum (Eds.), *The unconscious reconsidered.* New York: Wiley.

Kihlstrom, J. F. (1987). The cognitive unconscious. *Science, 237,* 1445–1452.

Kihlstrom, J. F. (1990). The psychological unconscious. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 445–464). New York: Guilford Press.

Kihlstrom, J. F., & Cantor, N. (1984). Mental representations of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 17) New York: Academic Press.

Klopfer, W. G. (1981). Integration of projective techniques in the clinical case study. In A. I. Rabin (Ed.), *Assessment with projective techniques* (pp. 233–264). New York: Springer.

Kochanska, G. (1993). Toward a synthesis of parental socialization and child temperament in early development of conscience. *Child Development, 64,* 325–347.

Koestner, R., Weinberger, J., & McClelland, D. C. (1991). Task-intrinsic and social-extrinsic sources of arousal for motives assessed in fantasy and self-report. *Journal of Personality, 59,* 57–82.

Koppitz, E. (1968). *Psychological evaluation of children's human figure drawings*. New York: Grune & Stratton.

Koppitz, E. M. (1984). *Psychological evaluation of human figure drawings by middle school pupils*. New York: Grune & Stratton.

Kuhl, J., & Beckmann, J. (Eds.) (1994). *Volition and personality*. Seattle, WA: Hogrefe & Huber.

Lerner, P. (1992). Toward an experiential psychoanalytic approach to the Rorschach. *Bulletin of the Meninger Clinic, 56,* 451–464.

Leve, R. M. (1995). *Child and adolescent psychotherapy: Process and integration*. Boston: Allyn & Bacon.

Lewin, K. (1935). *A dynamic theory of personality*. New York: McGraw-Hill.

Locraft, C., & Teglasi, H. (1997). Teacher rated empathic behavior and children's T.A.T. stories. *Journal of School Psychology, 35,* 217–237.

Luborsky, L., & Crits-Christoph, P. (1990). *Understanding transference: The CCRT method*. New York: Basic Books.

Lundy, A. (1985). The reliability of the Thematic Apperception Test. *Journal of Personality Assessment, 49,* 141–145.

Machover, K. (1949). *Personality projection in the drawing of the human figure*. Springfield, IL: Charles C. Thomas.

MacLeod, C., & Cohen, I. L. (1993). Anxiety and the interpretation of ambiguity: A text comprehension study. *Journal of Abnormal Psychology, 102,* 238–247.

Maloney, M. P., & Ward, M. P. (1976). *Psychological assessment: A conceptual approach*. New York: Oxford University Press.

McAdams, D. P. (1990). Unity and purpose in human lives: The emergence of identity as a life story. In A. I. Rabin, R. Zucker, R. Emmons, & S. Frank (Eds.), *Studying persons and lives* (pp. 148–200). New York: Springer.

McAdams, D. P. (1995). What do we know when we know a person? *Journal of Personality, 63,* 365–396.

McClelland, D. C., & Koestner, R. (1992). The achievement motive. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 143–152). New York: Cambridge University Press.

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review, 96,* 690–702.

McGreevy, J. C. (1962). Interlevel disparity and predictive efficiency. *Journal of Projective Techniques, 26,* 80–87.

Meissner, W. W. (1974). Differentiation and integration of learning and identification in the developmental process. *Annual of Psychoanalysis, 2,* 181–196.

Meissner, W. W. (1981). Internalization in psychoanalysis. *Psychological Issues Monograph, 50.* New York: International Universities Press.

Mendez, M. F., Ala, T., & Underwood, K. L. (1992). Development of scoring criteria for the clock drawing task in Alzheimer's disease. *Journal of the American Geriatric Society, 40,* 1095–1099.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: Macmillan.

Moffit, K. H., & Singer, J. A. (1994). Continuity in the life story: Self-defining memories, affect, and approach/avoidance personal strivings. *Journal of Personality, 62,* 21–43.

Morgan, C. O., & Murray, H. A. (1935). A method for investigating fantasies: The Thematic Apperception Test. *Archives of Neurology and Psychiatry, 34,* 289–306.

Murray, H. A. (1938). *Explorations in personality.* New York: Oxford University Press.

Murray, H. A. (1943). *Thematic Apperception Test manual.* Cambridge, MA: Harvard University Press.

Murstein, B. I. (1965). The stimulus. In B. Murstein (Ed.), *Handbook of projective techniques.* New York: Basic Books.

Murstein, B. I. (1968). Efforts of stimulus, background, personality, and scoring system on the manifestation of hostility on the TAT. *Journal of Consulting and Clinical Psychology, 32,* 355–365.

Norcross, J. C. (Ed.) (1986). *Handbook of eclectic psychotherapy.* New York: Brunner/Mazel.

Piaget, J. (1954). *The construction of reality in the child.* New York: Basic Books.

Piotrowski, C. (1984). The status of projective techniques: Or, "wishing won't make it go away." *Journal of Clinical Psychology, 40*(6), 1495–1502.

Piotrowski, C., & Keller, J. W. (1989). Psychological testing in outpatient mental health facilities: A national study. *Professional Psychology: Research and Practice, 20,* 423–425.

Piotrowski, C., Sherry, D., & Keller, J. W. (1985). Psychodiagnostic test usage: A survey of the Society for Personality Assessment. *Journal of Personality Assessment, 49,* 115–119.

Piotrowski, C., & Zalewski, C. (1993). Training in psychodiagnostic testing in APA-approved PsyD and PhD clinical psychology programs. *Journal of Personality Assessment, 61,* 394–405.

Plomin, R. (1986). *Development, genetics, and psychology.* New York: Erlbaum.

Quinn, N., & Holland, D. (1987). *Cultural models in language and thought.* New York: Cambridge University Press.

Rabin, A. I. (1981). Projective methods: A historical introduction. In A. I. Rabin (Ed.), *Assessment with projective techniques* (pp. 1–22). New York: Springer.

Rappaport, D., Gill, M., & Schafer, R. (1968). *Diagnostic psychological testing* (Rev. ed.). New York: International University Press.

Raynor, J. D., & McFarlin, D. B. (1986). Motivation and the self system. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 315–349). New York: Guilford Press.

Ritzler, B. A., Sharkey, K. J., & Chudy, J. F. (1980). A comprehensive projective alternative to the TAT. *Journal of Personality Assessment, 44,* 358–362.

Roback, H. B. (1968). Human figure drawings: Their utility in the clinical psychologist's armamentarium for personality assessment. *Psychological Bulletin, 70,* 1–19.

Rogers, R. (Ed.) (1997). *Clinical assessment of malingering and deception.* New York: Guilford Press.

Rubin, D. C. (Ed.) (1986). *Autobiographical memory.* New York: Cambridge University Press.

Rummelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Schematic and sequential thought processes in PDP models. In J. L. McClelland & D. E. Rummelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2). Cambridge, MA: MIT Press.

Sackett, P. R., Zedeck, S., & Fogli, L. (1988). Relations between measures of typical and maximum job performance. *Journal of Applied Psychology, 73,* 482–486.

Saltz, G., & Epstein, S. (1963). Thematic hostility and guilt responses as related to self-reported hostility, guilt, and conflict. *Journal of Abnormal and Social Psychology, 67,* 469–479.

Sandler, J, & Rosenblatt, B. (1962). The concept of the representational world. *The Psychoanalytic Study of the Child, 17,* 128–145.

Schank, R. C. (1990). *Tell me a story: A new look at real and artificial memory.* New York: Charles Scribner's Sons.

Shapiro, D. (1965). *Neurotic styles.* New York: Basic Books.

Shneidman, E. S. (1951). *Thematic test analysis.* New York: Grune and Stratton.

Singer, J. L. (1981). Research applications of projective methods. In A. I. Rabin (Ed.), *Assessment with projective techniques* (pp. 297–331). New York: Springer.

Singer, J. L., & Salovey, P. (1991). Organized knowledge structures and personality. In M. J. Horowitz (Ed.), *Personal schemas and maladaptive interpersonal patterns* (pp. 33–80). Chicago: University of Chicago Press.

Singer, J. A., & Singer, J. L. (1994). Social-cognitive and narrative perspectives on transference. In J. M. Masling & R. F. Bornstein (Eds.), *Empirical perspectives on object relations theory.* Washington, DC: American Psychological Association.

Skinner, B. F. (1953). *Science and human behavior.* New York: Macmillan.

Smith, B. L. (1994). Object relations theory and the integration of empirical and psychoanalytic approaches to Rorschach interpretation. In I. B. Weiner (Ed), *Rorschachiana XIX: Yearbook of the International Rorschach Society* (pp. 61–77). Seattle, WA: Hogrefe & Huber.

Smith, C. P. (Ed.) (1992a). *Motivation and personality: Handbook of thematic content analysis.* New York: Cambridge University Press.

Smith, C. P. (1992b). Reliability issues. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 126–139). New York: Cambridge University Press.

Snow, R. E. (1974). Representative and quasi-representative designs for research on teaching. *Review of Educational Research, 44,* 265–291.

Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement. *Psychological Bulletin, 112,* 140–154.

Stein, D. J., & Young, J. E. (Eds.) (1992). *Cognitive science and clinical disorders.* San Diego, CA: Academic Press.

Stein, M. J. (1955). *The Thematic Apperception Test* (Rev. ed.). Cambridge, MA: Addison-Wesley.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence.* New York: Cambridge University Press.

Stinson, C. H., & Palmer, S. E. (1991). Parallel and distributed processing models of person schemas and psychopathologies. In M. J. Horowitz (Ed.), *Personal schemas and maladaptive interpersonal patterns* (pp. 339–377). Chicago: University of Chicago Press.

Stricker, G., & Healey, B. J. (1990). Projective assessment of object relations: A review of the empirical literature. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2,* 219–230.

Swensen, C. H. (1968). Empirical evaluations of human figure drawings: 1957–1966. *Psychological Bulletin, 70,* 20–44.

Swindell, C. S., Holland, A. L., Fromm, D., & Greenhouse, J. B. (1988). Characteristics of recovery of drawing ability in left and right brain-damaged patients. *Brain and Cognition, 7,* 16–30.

Taylor, S. E., & Crocker, J. (1981). Schematic bases of social information processing. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social Cognition. The Ontario symposium on personality and social psychology*. Hillsdale, NJ: Erlbaum.

Teglasi, H. (1993). *Clinical use of story telling: Emphasizing the TAT with children and adolescents*. Needham Heights, MA: Allyn & Bacon.

Thomas, A. D., & Dudek, S. Z. (1985). Interpersonal affect in Thematic Apperception Test responses: A scoring system. *Journal of Personality Assessment, 49,* 30–36.

Tomkins, S. S. (1947). *Thematic Apperception Test*. New York: Grune & Stratton.

Tomkins, S. S. (1979). Script theory: Differentiated magnification of affects. In H. E. Howe, Jr. & R. A. Dienstbier (Eds.), *Nebraska Symposium on Motivation* (Vol. 26). Lincoln, NE: University of Nebraska Press.

Tomkins, S. S. (1987). Script theory. In J. Aronoff, A. J. Rabin, & R. A. Zucker (Eds.), *The emergence of personality* (pp. 147–216). New York: Springer.

Tomkins, S. S. (1991). *Imagery, affect, consciousness* (Vol. 3). New York: Springer.

Torgesen, J. K. (1977). The role of non-specific factors in task performance of learning disabled children: A theoretical assessment. *Journal of Learning Disabilities, 10,* 27–35.

Tunis, S. L. (1991). Causal explanations in psychotherapy: Evidence for Target-and-Domain-Specific schematic patterns. In M. J. Horowitz (Ed.), *Personal schemas and maladaptive interpersonal patterns* (pp. 261–276). Chicago: University of Chicago Press.

Uleman, J. S., & Bargh, J. A. (Eds.) (1989). *Unintended thought*. New York: Guildford Press.

Urban, H. M. (1963). *The Draw-A-Person*. Los Angeles: Western Psychological Services.

Vane, J. R., & Guarnaccia, V. J. (1989). Personality theory and personality assessment measures: How helpful to the clinician? *Journal of Clinical Psychology, 45,* 5–19.

Veroff, J. (1992). In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 100–109). New York: Cambridge University Press.

Watkins, C. E., Campbell, V. L., & McGregor, P. (1988). Counseling psychologists' uses of and opinions about psychological tests: A contemporary perspective. *The Counseling Psychologist, 16,* 476–486.

Watkins, C. E., Campbell, V. L., Nieberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology Research and Practice*, *26,* 54–60.

Weiner, I. B. (1977). Approaches for Rorschach validation. In M. Rickers-Ovsiankina (Ed.), *Rorschach psychology*. Huntington, NY: Krieger.

Weiner, I. B. (Ed.) (1994). *Rorschchianna XIX: Yearbook of the International Rorschach Society*. Seattle, WA: Hografe & Huber.

Wertheim, E. H., & Schwartz, J. C. (1983). Depression, guilt, and self-management of pleasant and unpleasant events. *Journal of Personality and Social Psychology, 45,* 884–889.

Westen, D. (1991). The clinical assessment of object relations using the TAT. *Journal of Personality Assessment, 56,* 56–74.

Westen, D. (1993). Social cognition and social affect in psychoanalysis and cognitive psychology: From regression analysis to analysis of regression. In J. W. Barron, M. N. Eagle, & D. L. Wolitzky (Eds.), *Interface of psychoanalysis and psychology* (pp. 375–388). Washington, DC: American Psychological Association.

Willock, B. (1992). Projection, transitional phenomena, and the Rorschach. *Journal of Personality Assessment, 59,* 99–116.

Wilson, A. (1988). Levels of depression and clinical assessment. In H. D. Lerner & P. M. Lerner (Eds.), *Primitive mental states and the Rorschach* (pp. 441–462). Madison, CT: International Universities Press.

Wyatt, F. (1947). The scoring and analysis of the Thematic Apperception Test. *Journal of Psychology, 24,* 319–330.

Wyer, R. S., Jr., & Srull, T. K. (Eds.) (1994). *Handbook of social cognition* (2nd ed., Vols. 1 & 2). Hillsdale, NJ: Erlbaum.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.17

Computer Assisted Psychological Assessment

GALE H. ROID and W. BRAD JOHNSON
George Fox University, Newberg, OR, USA

4.17.1 INTRODUCTION 502
4.17.1.1 Definitions and Distinctions 502
4.17.1.2 Brief History of Computer Assisted Psychological Assessment 503
4.17.2 A TYPOLOGY OF COMPUTER ASSISTED PSYCHOLOGICAL ASSESSMENTS 504
4.17.2.1 Test Administration 505
4.17.2.2 Computer Scoring 506
4.17.2.3 Descriptive Interpretation 507
4.17.2.4 Narrative Interpretation 507
4.17.2.5 Statistical–Actuarial Programs 507
4.17.3 COMPUTER ASSISTED PSYCHOLOGICAL ASSESSMENT: ADVANTAGES 509
4.17.3.1 Improved Administration and Scoring 509
4.17.3.2 Objectivity 509
4.17.3.3 Speed 509
4.17.3.4 Reliability 509
4.17.3.5 Cost Effectiveness 509
4.17.3.6 Expert Consultation 510
4.17.3.7 Flexibility 510
4.17.4 COMPUTER ASSISTED PSYCHOLOGICAL ASSESSMENT: DISADVANTAGES 510
4.17.4.1 Excessive Generality: The Barnum Effect 510
4.17.4.2 Lack of Validity 511
4.17.4.3 Depersonalizing the Assessment Process 511
4.17.4.4 Potential for Misuse and Client Harm 511
4.17.4.5 Computer as Clinician 511
4.17.5 ETHICAL ISSUES 512
4.17.5.1 Test Development 512
4.17.5.2 Basis for Scientific and Professional Judgments 512
4.17.5.3 Describing the Nature of Psychological Services 513
4.17.5.4 Competence 513
4.17.5.5 Professional Context 514
4.17.5.6 CAPA with Special Populations 515
4.17.6 GUIDELINES FOR USERS OF COMPUTER-BASED TESTS AND INTERPRETATIONS 515
4.17.6.1 Administration 517
4.17.6.2 Evaluation and Selection of CBTIs 517
4.17.6.3 Interpretation 518
4.17.7 GUIDELINES FOR DEVELOPERS OF COMPUTER-BASED TEST SERVICES 519
4.17.7.1 Human Factors 519
4.17.7.2 Psychometric Properties 519
4.17.7.3 Classification Strategy 519
4.17.7.4 Validity of Computer Interpretations 519
4.17.7.5 Facilitation of Review 520

4.17.8 DIRECTIONS FOR THE FUTURE 520

4.17.9 CONCLUSIONS AND RECOMMENDATIONS 521

4.17.10 REFERENCES 522

4.17.1 INTRODUCTION

A common dilemma for clinical psychologists is the tension between the need for diagnosis and the complexities of the individual case. Although resources to help the client may be contingent on certain types of psychological disorders, concern for the client and the avoidance of stigmatizing labels clearly make assessment difficult when cases are unusual or complex. Also, time may be limited or resources of the client may be limited, making an extended assessment difficult to complete. In the midst of this dilemma, various methods of computer assisted psychological assessment (CAPA) have emerged and some have become prominent, especially with the advent of the personal computer in the 1980s, with both promise (e.g., Jackson, 1985) and potential problems (e.g., Matarazzo, 1986). This chapter surveys the definition, history, types of implementation, advantages and disadvantages, reliability and validity concerns, and issues of ethical and professional responsibility in the use of computers in assessment. No attempt will be made in this chapter to review or critique actual software or existing CAPA programs, except by example or reference, due to the rapidly changing nature of the technology and the continual updating of various programs. This chapter will, however, take a hard look at the proposed advantages of CAPA, and, instead of placing inordinate weight on the technical promise of computers, propose some firm limits on the use of CAPA. First, some definitions and important distinctions are made, followed by a brief history of the development of CAPA.

4.17.1.1 Definitions and Distinctions

CAPA is broadly defined as any application of computers to the development, administration, scoring or interpretation of tests, scales, inventories or questionnaires used in education or psychology.

Psychological texts such as Gregory (1996) have proposed similar definitions. Those appear consistent with early use of the term CAPA such as that employed by Fowler (1985). The present chapter will focus more narrowly on computer applications to cognitive or personality instruments in which administration, scoring, and interpretation are attempted. Methods of CAPA development and techniques of using computers in test development are beyond the scope of the present chapter. For a review of some of the technical methods proposed for the development of computer-based psychological tests, see Green (1991a) on adaptive testing; Guastello and Rieke (1994) on the expert-system approach; Snyder, Widiger, and Hoover (1990) for a review of test-interpretive program development; and the review by Roid (1986).

By "assessment" we refer to the process of examining the entire range of information, including standardized tests, scales, inventories, questionnaires, protective tests, observations, interview data, histories, and other observations by experienced psychologists who study an individual for purposes of diagnosis, description, classification, treatment planning, pre-therapy observation, or any other professional evaluation. Most psychometric textbooks make an important distinction between assessment and "testing" or "measurement." Testing usually involves the administration of standardized stimuli (e.g., test or inventory items) to clients whose responses are scored by objective or judgment-based scoring methods. Measurement is traditionally defined as the assignment of categories or scale values to such aspects of individuals as traits, attributes, attitudes, behavior, and preferences by a psychological testing instrument. The word "test" is typically reserved for tasks in which there are correct answers or problems to be solved, in contrast to "scale" or "inventory" or "questionnaire" which measure traits or preferences for which no single correct answer exists for all people.

Matarazzo (1986) made the most vivid distinction between testing and assessment by arguing that assessment should be reserved as a term to describe the process conducted by an experienced psychologist who gathers information for purposes of interpreting it and giving it meaning within the context of the examinee's total life history. Matarazzo's distinction would result in questions such as, "Can computers really provide assessments or only measurements which must be interpreted by professionals?" This penetrating question strikes at the heart of the ethical issues to be discussed in this chapter. Matarazzo was contrasting his definition of assessment with the technical, clerical, or computerized processing of test information. His position received immediate reactions such as that of Fowler and Butcher (1986) who argued for the view of the assessment process as one that could include both clinical and computerized statistical assistance in the interpretive process.

The standards for psychological testing (American Psychological Association (APA), 1985) make a distinction between test administration and test interpretation, and warn that test developers should inform users of any special training or expertise needed for either administration or interpretation. Thus, CAPA may be used for test administration, for administration and scoring, or for a full range of administration and interpretation. Typically, a higher level of training is assumed for the interpretation of test results, such as graduate-level measurement or psychometrics courses, supervised assessment and reporting, and knowledge of concepts such as error of measurement. Also, a higher level of validation is required for any interpretations of tests that impact clients. The distinction between administration and interpretation of tests is critical to the evaluation of CAPA and the role of computers because it highlights the continuing responsibility of the psychologist to proactively interpret the results of the assessment process. Thus, we define another category of programs, computer-based test interpretive (CBTI) programs, to delineate those that emphasize interpretation, often with narrative descriptions of results.

A final distinction of great importance was first discussed in depth by Meehl (1954) —clinical versus statistical (actuarial) prediction. By clinical prediction, common usage would normally suggest the processing of information by the trained clinician who makes a prediction, diagnosis, classification, or evaluation of a client based on clinical experience, "intuition," and professional judgment. By statistical prediction, common usage would suggest methods such as multiple regression or validated "cutting scores" being used to predict, diagnose, and classify an individual based on a relevant research database. Meehl's (1954) lengthy discussion of the issues pointed out important distinctions such as the fact that both clinical judgment and statistical methods may be predicting an individual case from trends in previous group data—the group of all previous clients in the instance of clinical judgment. Later researchers such as Goldberg (1968) showed that clinical judgments are not necessarily more complex or "configural" as compared to statistical predictions, since simple linear equations effectively modeled the behavior of skilled judges who were assessing, for example, psychosis vs. neurosis from the Minnesota Multiphasic Personality Inventory (MMPI). It would be impossible to summarize all of the careful distinctions among clinical vs. statistical methods, and the reader is referred to Meehl (1954), Goldberg (1968), or to the review by Garb (1994).

4.17.1.2 Brief History of Computer Assisted Psychological Assessment

Fowler (1985) and Moreland (1992) provide brief histories of CAPA. Highlights include the early attempts in the 1940s to computer score the Strong Vocational Interest Blank (SVIB), Meehl's (1954) landmark book, and early 1960s versions of mainframe computer programs to score and interpret MMPI responses. To say that the history of CAPA has followed the development of computers and scanning equipment is obvious, but several landmarks are instructive. Following the refinement of mainframe computers in World War II, and until approximately 1970, most efforts to study or implement CAPA were conducted on mainframe computers with several critical attributes: (i) access to the computer was typically limited to professional computer operators, (ii) programs were developed and operated in "batch" mode where little interaction occurred between the user and the computer—interactions were specified in advance with various control commands, and (iii) the responses of examinees had to be scanned by a separate document scanner or key entered. Thus, "dynamic" entry of data was not possible at most installations. The advanced development of high-capacity scanning equipment is often attributed to Lindquist and his colleagues at the University of Iowa (e.g., Flanagan & Lindquist, 1951), for use on various educational achievement tests, although similar developments were employed for the SVIB and the MMPI (Moreland, 1992).

In the 1970s (and, perhaps, earlier in experimental laboratories), "time sharing" computer systems proliferated that connected the user via Teletype machine or early versions of computer terminals. Some CAPA applications were developed on systems initially designed for computer-assisted instruction. These "real time" systems allowed for a degree of interaction between the programmer and the system, between the user and the output, and, in some experimental applications, between the examinee and the time-share computer (e.g., Klingler, Miller, Johnson, & Williams, 1977). Klingler et al. developed an automated assessment system for psychiatric inpatients at a veterans administration hospital in Utah. The potential cost-benefit of such applications was immediately apparent, and this stimulated discussions at psychological conventions about the ethical issues of on-line test administration and scoring.

Another important innovation in CAPA was the development of multistage "branching" or "adaptive" tests, which emerged from early psychometric studies in educational measurement (e.g., Angoff & Huddleston, 1958; Linn, Rock, & Cleary, 1969; Lord, 1968, 1971), from sequential methods in statistics (e.g., Cowden, 1946; Wald, 1947), and from the development of item-response theory (e.g., Birnbaum, 1968; Rasch, 1980). As early as the late 1960s, there was a considerable unpublished literature on adaptive testing methods and their application to psychological scales as well as educational tests (e.g., Bayroff & Sealy, 1967; Patterson, 1962; Roid, 1969). However, the first widely distributed applications of computerized adaptive testing emerged in the 1980s. Weiss (1983) was a key developer of methodology and applications in college and military settings. Operational programs on personal computers emerged in aptitude testing (e.g. McBride, 1988) and in conjunction with a large-scale project to computerize the Armed Services Vocational Aptitude Battery (ASVAB; Green, 1991a; Sands & Gade, 1983).

With the development of the first Apple computers and the release of the first IBM personal computer in approximately 1980, the widespread implementation and practicality of CAPA for the local psychologist was finally available. The early 1980s witnessed a flurry of rapid development and distribution of test scoring and interpretive programs for microcomputers (e.g., Roid & Gorsuch, 1984). In reaction to the swift proliferation of CAPA programs, Matarazzo (1983) published a stern warning about the lack of validation of narrative interpretations and the potential dangers of distribution to inexperienced users. Special series of articles appeared in the *Journal of Consulting and Clinical Psychology* (Butcher, 1985), and *Computers in Human Behavior* (Mitchell & Kramer, 1985.) These early articles were a mixture of praise about the potential of the methodology and warnings about the limitations and ethical consequences of irresponsible usage. In response to this outpouring of attention to CAPA, the APA (1986) published a booklet of guidelines for development and usage of "computer-based tests and interpretations" (see Table 2). Several resource books, cataloging various programs and user options, were published (e.g., Butcher, 1987; Krug, 1984).

Because of the rapid innovation in computer technology, and, in the mid-1990s, the development of the world-wide web, Internet, video conferencing, and multimedia desk-top computers, it is difficult to anticipate the developments in CAPA. Certainly, the addition of multimedia to assessment instruments will be more easily achieved and affordable by more test developers. Dynamic video segments within standardized tests could make real-life situational assessment more feasible. In any case, the potential and the new ethical concerns stimulated by the prospect of test processing by Internet or e-mail services looms large on the horizon.

In the next section, we review a typology of CAPA programs. Included is a review of some of the key literature that documents the current status of CAPA and further definitions and distinctions.

4.17.2 A TYPOLOGY OF COMPUTER ASSISTED PSYCHOLOGICAL ASSESSMENTS

Not all applications of CAPA have the same level of complexity or developmental sophistication to support them. Thus, it is important to have a typology of various programs so that the proper role of CAPA can be evaluated. It will be the overall recommendation of this chapter that the psychologist carefully evaluate the level and type of CAPA product considered for clinical use, and perhaps, "draw a line" of restriction in terms of the type, technical quality, and validity of CAPA results actually employed in the evaluation of clients. For these reasons, a typology of CAPA programs, derived from the literature and the previous work of the senior author (Roid & Gorsuch, 1984; Roid, 1986), is presented below. First, a few key distinctions and definitions are discussed.

As stated earlier in this chapter, a wide range of computer "assistance" is used in the development of psychological tests and assessments, and we have chosen to delete these programs from our typology in favor of emphasis on CAPA products that may be directly used in client assessments. Also, there are a potentially large number of CAPA products available in the "employment testing" industry that are not considered due to the clinical focus of the current discussion. Finally, computer-based products that are designed for on-line delivery of psychotherapeutic exercises or cognitive training are excluded, even though some of them include assessment components.

Significant differences exist between CAPA programs for cognitive, neuropsychological, and personality assessments. Some of the same statistical techniques have been applied to the profile scores of both ability and personality batteries, but, for the most part, each has a particular type and style of presentation. For example, it is now common in the software to score intelligence scales to include computations, frequencies, and probabilities of differences between two or more subtests or indexes (e.g., verbal IQ vs. performance IQ). Statistical contrasts of differences are more rare in personality assessment, where, instead, visual inspection of the profile of scores is more common. In computerized interpretive programs in the personality realm (e.g., Lachar, 1984), it is common to include "critical items" and correlates between clinician ratings and "profile elevations" (ranges of scores, such as 70–79T on the MMPI). However, in terms of the typology that follows, the collective statistical techniques can be classified together as "profile and statistical analysis."

Table 1 A typology of computer assisted assessment.

- *Test administration*
  - Conventional test administration
  - Multimedia or specialized test administration
  - Computerized adaptive testing (CAT)
- *Computer scoring*
  - Conventional scoring
  - Statistical scoring and profile analysis
  - Graphic displays
  - Database analysis
- *Descriptive interpretation*
- *Narrative interpretation (CBTI programs)*
  - Informal narrative
  - Clinician-modeled or expert system
- *Statistical–Actuarial interpretation*

4.17.2.1 Test Administration

Table 1 presents the typology. Some software is designed to administer tests and to record the results, which may be printed or briefly summarized on the computer screen. Administration can include highly sophisticated, multimedia programs with attached database capability, such as the MicroCog by Powell et al. (1993), an on-line testing system to detect cognitive impairment in older adults. Programs can vary from basic "page turning" software that simulates paper-and-pencil tests, to those such as MicroCog that include graphic displays, reaction time, optional sound, and delayed memory trials also timed. Of particular importance in this category of the typology, are programs such as the Conners' (1995) Continuous Performance Test (CPT), used to assess attentiveness, particularly for children and adults referred for attention-deficit evaluation. The CPT assesses visual vigilance in scanning and concentrating on a stimulus array—a type of performance that must have accurate, split-second timing of both stimuli and responses—a perfect application of computers in test administration.

A separate category of test administration is computer-adaptive testing (CAT). CAT requires specialized software to implement the sophisticated (usually item-response theory or Bayesian) statistical model on which it is based. Several prominent examples exist, including the newer versions of the computerized Graduate Record Exams (Educational Testing Service, 1995), the ASVAB (Sands & Gade, 1983), and, one of the first widely published commercially published tests, the Differential Aptitude Tests (DAT Adaptive) Computerized Adaptive Edition (Psychological Corporation, 1987). The essential attribute of CAT administration is that the difficulty of the test items is tailored dynamically (Lord, 1968) to the ability level of the examinee during the testing session. Research (e.g., Green, 1991a) has shown that modern implementations of CAT can reduce testing time by half while maintaining an acceptable level of measurement error. To implement such a system requires complex psychometric development in which each item in the pool of available items is given field trials and calibrated statistically. Programs such as MicroCAT (Assessment Systems, 1990) or other specialized computer programs are then used to implement the presentation of items. A broad range of "branching rules" and scoring methods are now available for such tests, depending on the type of item-response theory model used in the calibration of the items (e.g., Wainer, 1990).

While the statistical complexity of CAT applications may be beyond the level of psychometric training for many clinical psychologists, one can consult the professional reviews of such tests in the *Buros Mental Measurements Yearbooks*, or the current on-line electronic versions of these reviews. Alternatively, legitimate CAT programs should have an accompanying manual which can be evaluated for conventional reliability and validity data, just as any published test is evaluated. Psychological conventions and meetings often include exhibits or demonstrations of computer software, which can be a more cost-effective way to review these products, since publishers are often reluctant to distribute expensive software for review. As a final recourse, one should consult a psychometric psychologist at a local university who can help to evaluate the technical qualities of programs considered for purchase.

4.17.2.2 Computer Scoring

Conventional scoring software has typically processed answer sheets or key-entered item responses and printed basic raw scores and various derived scores such as percentiles or standard scores. For many years, such programs had a minimum of graphic display, due to the restricted nature of graphics possible until the advent of the high-speed laser printer. As mentioned in the brief history of CAPA, the earliest versions of conventional scoring were on mainframe computers for the SVIB or the MMPI. Many of these programs were quite creative and complex, relying on "typeface" characters to plot profiles or display histogram-type plots of scores.

At the next level in the typology are scoring programs that include sophisticated statistical computations and profile analysis. A proliferation of "scoring assistant" programs has emerged, such as the "compuscore" series for the Stanford-Binet and Woodcock-Johnson at Riverside Publishers, the PsyTest and Wechsler Intelligence Scale for Children-Third Edition (WISC-III) and Wechsler Individual Achievement Test (WIAT) programs of The Psychological Corporation, the Western Psychological Services Test Report series, the software systems developed by Psychological Assessment Resources, and the "Assist" software for the tests published by American Guidance Service, to name a few of the larger groupings of scoring programs. Although many of these programs also include interpretive functions, to be discussed in the next section, the scoring sections of these programs tend to be more graphic and more sophisticated than earlier, conventional scoring programs (Roid, 1985). For example, a much wider array of profile indexes, score-difference analyses, critical item comparisons, and profile-matching routines are included than previously available programs (with exceptions being some of the complex, early MMPI interpretive programs). Two examples, one in personality and one in cognitive assessment, may illustrate the characteristics of such programs. The Tennessee Self Concept Scale (Roid & Fitts, 1988) computerized scoring program includes a wide array of research-based profile indexes, checks on the validity of response patterns, faking-good scales, critical-item lists, and a multivariate profile matching method that is implemented on a complex, color-printed display. The profile-matching method, interval-banded profile analysis, was initially developed by Huba (1986) to provide a statistical test of the degree of match between prototypical profiles (stored within the computer program) and those of new profiles being scored. The essential part of the matching routine is a multivariate chi-square test of "fit" between the profile pattern (which is allowed to vary within "bands" or ranges of scores) and "target" prototypical patterns that have been established through statistical studies of clinical and normative cases. A second example of sophisticated profile analysis is in the Wechsler Scoring Assistant (Psychological Corporation, 1992), where profile-score differences among all the subtests of the WISC-III are analyzed statistically, and the regression–prediction method of calculating ability vs. achievement discrepancies (WISC-III vs. Wechsler Individual Achievement Test scores) are included for screening of learning disabilities.

As more and more technical advances are made in computer printers, compact-disk and computerized video displays, the more sophisticated the graphic display of test profiles and score patterns will become. In the final level of the typology are programs that allow for the archiving of multiple cases, set up and usage of extensive databases of test results, and statistical manipulation of these data. Examples of such programs are found among the offerings of all the publishers of the large achievement-test batteries, who typically offer software for analyzing trends in achievement data across an entire school district. The most sophisticated versions of these programs allow for the assessment of growth or change at the individual, classroom, and school levels of analysis.

Based on the experience of the senior author of this chapter, who participated in the development of several computer-scoring systems (e.g., Roid, 1985; Roid & Fitts, 1988; and the Wechsler Scoring Assistant, Psychological Corporation, 1992), these profile-analysis programs are typically based on extensive data analysis, not informal or subjective processes that often occur in the "informal narrative" category of the typology. Most of the profile analyses are based on actual data from the standardization and validity studies of the major tests, and extensive staff, resources and time are invested in the development, and (elaborate) cross-checking of program accuracies. Because it is well known that clerical errors in scoring standardized tests are all too common, these scoring programs provide a valuable service to clinical psychology by delivering accurate scores. Further, the complexity of profile analysis, difference-score computation and analysis, and profile-pattern matching would not be feasible with hand-scoring methods. Except for a few unusual cases, the statistical scoring programs tend to be well documented in the technical manuals published with them.

Thus, in summary, the sophisticated "scoring assistant" model of computer scoring software has distinct advantages for the clinical psychologist. Time saved by the clinician in using scoring software could be reinvested in more time with the client or additional personal interview or case follow-up.

4.17.2.3 Descriptive Interpretation

A descriptive type of program would generate sentences of explanation, such as "The client has a significantly higher score on subtest three as compared to the other subtests in the profile," along with the printed scores and profile. The distinction between this level of description and the narrative interpretations described in the next section of the typology is that description remains tied to the facts and does not indicate cause-effect relationships or connections with research or clinical findings. The descriptions would be analogous to the phrases used in the results section of a research article—reporting the findings before they are discussed or interpreted. As noted in the example above, the descriptions can be rooted in sophisticated statistical comparisons between scores (e.g., Silverstein, 1981).

Some of the first computer interpretive programs widely used in the 1980s had redundant, printed phrases that described score elevations for multiple scores, using exactly the same wording. The best examples of descriptive interpretation include sophisticated "sentence generators" that compose explanations using a variety of modifiers, sentence construction and style, similar to the variety present in good report writing. One example is that of the Barclay Classroom Assessment System (Barclay, 1983) which varied pronouns such as "he" and "she," and used research on the scaling of verbal phrases (e.g., Lichtenstein & Newman, 1967; Pohl, 1981) to compose explanations of score elevations. For example, scale values for descriptors such as "fairly often" vs. "very infrequently" can be used to anchor score-level descriptions, in a way more precise than an informal or subjective use of such language. Use of more precise methods of description, could increase the potential of such programs to maintain a more objective level of description, that is, sticking to the facts.

4.17.2.4 Narrative Interpretation

The worst examples of informal, unvalidated, narrative interpretations of psychological tests were the target of Matarazzo (1983, 1986) in his critiques of CAPA. By informal narrative we mean the wordy descriptions printed on computer reports of assessment results that attach clinical significance, often based on clinical lore or theoretical explanations, to the patterns or levels of scores. Such narratives have been faulted for lack of validation of specific sentences or phrases and for lack of attention to individual client variations. Informal narratives also have the impact of vague generalities that result in a "Barnum effect" in which enough truth lies within a complex of statements if one emphasizes the accurate parts and minimizes the inaccurate parts. In research on the accuracy of computerized reports, Adams and Shore (1976) found an inverse relationship between length of reports and their accuracy rated by clinicians. For these reasons, clinicians should be very cautious about the validity of informal narrative programs.

Basically, the key feature of such programs, that clinicians should be able to discern from the manual and advertising material published with the computer program, is whether or not the narrative sentences have been accumulated from empirical research or whether they were written by the author(s) of the program without validation studies. Another key is whether or not the statements were validated by empirical linking of clinician ratings and scores, where both were collected in the same research studies. Positive examples of proper validation of narrative programs are reviewed by Snyder et al. (1990). Published programs that include extensive empirical validation of narrative reports include Lachar (1984) and Snyder (1981), to name only two prominent examples.

By clinician-modeled programs we mean those that either (i) employ the process used by a renowned clinician within the logic of the computer program, or (ii) employ statistical models of the process used by expert clinicians, determined through research on clinical judgment (e.g., the methods described by Goldberg, 1968). Examples of the former include the WISC-III program of Kaufman (1996) that implements his documented approach to interpretation of the Weschler scales (Kaufman, 1994). Examples of the latter have never been implemented, as far as we know, although the methodology is particularly promising. For a review of expert systems and their application to CAPA, see Guastello and Rieke (1994).

4.17.2.5 Statistical–Actuarial Programs

The term "actuarial" as applied to psychological assessment was coined by Meehl (1954) in analogy to the actuary process in insurance-risk determination. Gregory (1996) presented a classic definition of an actuarial interpretation, attributed to J. O. Sines, as founded on "the empirical determination of the regularities that may exist between specified psychological test data and equally clearly specified socially, clinically, or theoretically significant nontest characteristics of the person tested" (p. 579) Thus, an actuarial approach is databased and must show a statistical link between information collected outside the test (e.g., clinician's observations) and the test scores or patterns of test results. At their best, narrative statements appearing in a truly actuarial program would not be based on clinical opinion, but rather on rigorous research linking test and nontest information.

A classic example of an actuarial approach to the MMPI is provided by Marks and Seeman (1963) who defined a 4–8–2 profile "code type" as follows:

(i) Scales 4, 8, and 2 over 70T
(ii) Scale 4 minus 2 less than 15T
(iii) Scale 7 not to exceed 4 by more than 4T
(iv) Scale 8 minus 2 less than 15T
(v) Scale 8 minus 7 more than 5T
(vi) Scale 8 minus 9 more than 10T
(vii) Scale 9 less than 70T
(viii) Scales L and K less than F, F less than 80T

Thus, detailed specifications are given for the entire profile pattern, not just the high scores on 4 (psychopathic deviate), 8 (schizophrenia), and 2 (depression). Even the validity scales, L (lie scale), F (frequency), and the suppressor scale K, are used to verify the accuracy of the profile. Marks and Seeman (1963) reported that patients obtaining this profile were mainly diagnosed psychotic (71% schizophrenic, paranoid type) though some were seen as personality disorders (e.g., 21% sociopathic). Note that these are percentages for a given sample of clinical patients, and that they do not predict with 100% accuracy, as with all probability relationships. Therefore, the most accurate statement that should appear in narrative form would be something like, "some research studies of clinical patients have shown a frequency of about 70% schizophrenic–paranoid for this pattern of scores," and the clinician should be careful to screen all such comments to be sure they apply to the current case.

Another type of actuarial approach called "empirical correlates" was developed by Lachar (1984) for the Personality Inventory for Children (PIC) computer interpretations. Lachar administered clinical checklists to psychologists who had interviewed children who were subsequently tested with the PIC and, previously, known to have certain diagnoses. In addition, a large sample of children in a normative sample were given the PIC and the checklist was completed by examiners for these children also. For certain scales (and certain scale elevations, e.g., scores greater than 79T), empirical correlates were those descriptive statements statistically associated with the "elevation" (score in the clinical range). For example, for the achievement scale, a percentage (e.g., 73%) of clinical cases would show symptoms such as behavioral adjustment difficulties or negative self concept (as indicated on the clinical checklist), when the achievement scale score exceeded 79T. To confirm such findings, the same correlational study would be repeated as a cross-validation. Since empirical findings always "shrink" upon cross validation, the percentage may reduce to 70% of cases in the above example.

The best statistical–actuarial interpretation programs are those that meet the standards suggested by Snyder et al. (1990), where cut-score rules, program logic, and narrative sentences have been subjected to empirical research and are well documented in a technical manual. Also, the best of such programs have been subjected to validation research that surveys clinical users of the reports, collects accuracy ratings, and assesses the impact on clinical decisions of various client reports in rigorous follow-up studies. The number of programs with such rigorous development are, unfortunately, few. Some systems that come close are the larger, well-researched MMPI programs, the PIC (Lachar, 1984), the Marital Satisfaction Inventory (Snyder, 1981), and the Wechsler programs developed by the Psychological Corporation (1994), to name only four examples. Even these programs can be faulted in that every phrase or descriptor may not have been subjected to empirical trials based on moderator variables, and all may not have been examined for report accuracy. Such research, as with all good construct validation, takes decades of accumulated research. One trend and positive aspect of some of the newer programs (e.g., WISC-III Writer, Psychological Corporation, 1994), is the provision for placing the clinician in control of the final collection of narrative statements that appear in the report. As will be discussed in later sections of this chapter, the individual clinician must maintain control and oversight over the selection and accuracy of all statements generated by computer interpretations that are employed in case reports.

With the typology of CAPA programs in mind, the following section reviews some of the advantages and disadvantages of CAPA. Key literature is cited in the next section.

4.17.3 COMPUTER ASSISTED PSYCHOLOGICAL ASSESSMENT: ADVANTAGES

Computerized approaches to psychological assessment have rapidly become standard practice in most mental health treatment settings. Not only have computer administration and scoring of various tests become commonplace, but CBTIs have become a booming industry in their own right. In 1989, a survey of 413 mental health facilities in the USA revealed that 53% of these major facilities employed some form of CBTI (Piotrowski & Keller, 1989) CAPA offers numerous advantages to users, organizations, and clients alike. The following are some of the more salient advantages.

4.17.3.1 Improved Administration and Scoring

Computer administration of psychological tests serves to enhance standardization and clinician control over the testing process. In addition to reduced time for administration and rapid availability of feedback, computer administration allows presentation of even complex testing stimuli (Krug, 1987) and early identification of examinees who misunderstand directions or test stimuli (Wise & Plake, 1990). Though some have expressed concerns about the equivalence of computer administered test scores, reviews suggest that when instructions are similar, computer-administered tests generally yield scores which are equivalent to paper and pencil versions (Finn & Butcher, 1991). Additionally, preliminary research indicates that computer-administered tests are not only acceptable to examinees but often preferred over conventional testing procedures (Burke & Normand, 1987; Finn & Butcher, 1991). Specifically, examinees report greater interest, less anxiety, and greater comfort in responding to computer generated test stimuli. Computer-administered tests are also quite useful for more disturbed clinical populations who, by virtue of their level of disorganization or inattention, may respond more effectively to a computer than another person (Bloom, 1992).

4.17.3.2 Objectivity

In his early call for a good actuarial "cookbook" for use in psychological assessment, Meehl (1956) noted the potential benefit of decreased distortion and bias, in the recording, storage, and retrieval of test data and interpretive material. Most appear to agree that CAPA can substantially reduce errors associated with ignorance, bias, or stereotyping on the part of clinicians (Butcher, 1987; Dahlstrom, 1993). Once interpretive rules are developed and programmed, they are automatically applied to protocols regardless of extraneous circumstances.

4.17.3.3 Speed

Perhaps the most pragmatic and obvious advantages of computer assisted psychological assessment is the potential for marked reduction in time required for administration, scoring and interpretation of psychological instruments (Butcher, 1987; Kleinmuntz, 1969, 1975; Wise & Plake, 1990). Administration time for both achievement and personality tests is significantly reduced by computer administration (Wise & Plake, 1990). Computerized scoring and interpretive report generation serve to enhance timely processing of test data by clinicians and subsequent delivery of feedback to consumers. In addition, these substantial reductions in time are associated with improvements both in consistency and accuracy.

4.17.3.4 Reliability

Computer assisted scoring and interpretation of psychological tests serves to radically reduce error variance. Butcher (1987) noted "The computer seldom has an off day as human test interpreters do" (p. 5). As a result, well developed test interpretation programs offer nearly perfect reliability for the functions of scoring and compiling preprogrammed interpretive statements (Burke & Normand, 1987; Graham, 1993; Krug, 1987). Once scale correlates have been identified and replicated, they can easily be stored and reliably recalled when particular scale elevations or score configurations appear in a protocol.

4.17.3.5 Cost Effectiveness

Computer scoring and interpretation of psychological tests generally result in marked reduction in clinician time and, therefore, expense to the consumer (Burke & Normand, 1987; Butcher, 1987; Graham, 1993). Though computerized systems will certainly require an initial outlay for equipment, software, and training, long-term savings to both users and consumers may be expected. With the advent of managed behavioral healthcare and increasing attention to issues such as cost-effectiveness, efficiency, and treatment utility, CAPA offers several ways for clinicians to reduce the time and expense required for a thorough assessment.

4.17.3.6 Expert Consultation

Many modern day CBTI programs appear to have effectively realized Meehl's (1956) hope for systems which employ large databases and widely representative samples on which to base predictions and diagnoses. Many well developed CBTIs systematically organize and access massive normative databases and extensive bodies of empirical research findings to bolster and undergird test interpretations (Krug, 1987). In spite of some methodological weaknesses, research on the equivalency of computer generated vs. clinician generated test interpretation suggests many CBTIs perform at least as well as clinicians (Bloom, 1992; Burke & Normand, 1987; Finn & Butcher, 1991). For example, Kleinmuntz (1969), in a large scale multi-site study, found that computer generated MMPI interpretations performed as well as expert MMPI interpreters and surpassed the performance of average clinicians in predicting a client's primary clinical problem.

Computer generated interpretations may ideally offer the skilled clinician a source of expert consultation. Due to the voluminous information available via the computer system, interpretations generated by a CBTI may markedly enhance both the accuracy and comprehensiveness of clinical decision making, treatment planning, and feedback to the client. Garb (1994) pointed out that even when a CBTI report appears to conflict with interview, historical, and observational data, the report can still be quite valuable in leading the clinician to consider alternative hypotheses and collect additional data. CBTIs may be most helpful when a case is ambiguous and the diagnosis unclear. Also, as a result of the "expert" and "objective" look and sound of computerized psychological reports, CBTIs may be particularly advantageous in forensic settings (Butcher, 1987) where such reports may be viewed as a form of outside and corroborating opinion. Finally, computer generated reports are particularly advantageous as a source of objective and expert opinion in the psychotherapy enterprise (Finn & Butcher, 1991). Here the test interpretive report is presented as a consultation by an outside expert and introduced for the purpose of clarification and discussion.

4.17.3.7 Flexibility

A final advantage of computer assisted psychological assessment is the flexibility possible in such systems (Graham, 1993; Kleinmuntz, 1975). Perhaps the most tangible example of this flexibility is the development and ongoing refinement of adaptive or tailored testing programs. Here the system is programmed to adjust the difficulty level or volume of material in discrete areas to an individual examinee. Examples include research with the ASVAB and the MMPI. Sympson, Weiss, and Ree (1982) found that an adaptive version of the ASVAB produced validity coefficients which were equivalent to conventional administration of the test in spite of the fact that the adaptive versions were typically one-half the length of their conventional counterparts. Similarly, Roper, Ben-Porath, and Butcher (1991) found that among 155 college age subjects, an adaptive version of the MMPI-2 (averaging 28% shorter in length) produced profiles which were equivalent to conventional MMPI-2 administration.

4.17.4 COMPUTER ASSISTED PSYCHOLOGICAL ASSESSMENT: DISADVANTAGES

In addition to the many advantages inherent in CAPA, there are also concerns and unresolved dilemmas in the development, implementation, and utilization of computerized systems. Below, are highlighted several of these potential disadvantages. These include the problem of excessive generality, scanty validity, potential for depersonalizing the assessment process, potential for misuse or client harm, and the danger inherent in viewing the computer as a competent clinician.

4.17.4.1 Excessive Generality: The Barnum Effect

Computer generated testing interpretive reports have been criticized soundly for their frequently broad and highly generalized narrative statements (Butcher, 1987; Groth-Marnat & Schumaker, 1989; Matarazzo, 1986). Butcher described this as the problem of excessive generality, otherwise known as the Barnum effect or the Aunt Fanny report. CBTIs are notorious for offering patient descriptions based on insufficient empirical research. These narrative statements then apply to most human beings and are merely modal statements vs. person-specific descriptors.

This problem appears to have consistent empirical demonstration. O'Dell (1972) found that subjects perceived Barnum reports to be more accurate descriptions of their own personality than actual computer generated interpretations of their MMPI profiles. O'Dell concluded that statements with very high base rates tend to be believed or concurred with. Guastello, Guastello, and Craft (1989) found that 58% of respondents rated Barnum reports for the Personality Profile Compatibility Questionnaire as quite accurate descriptions. Nonetheless, actual CBTI reports for this measure were rated as significantly more accurate than the Barnum reports. Less encouraging was a similar study employing a CBTI for the Exner Rorschach system. Prince and Guastello (1990) reported that the Exner report offered only 5% discriminating power for any one patient when compared with bogus Exner interpretive reports. Perhaps more concerning was the finding that approximately 60% of the CBTI statements contained statements which merely described characteristics of the outpatient population in general. Related to this concern regarding Barnum statements and excessive generality is the frequently expressed concern regarding an "aura" of credibility or accuracy potentially attributed to CBTI reports by consumers. The scanty research in this area suggests that this concern may be largely unfounded and consumers rate the quality, credibility, and accuracy of CBTI or clinician generated reports as essentially similar (Andrews & Gutkin, 1991).

4.17.4.2 Lack of Validity

One of the most glaring problems inherent in the development and use of CBTI report systems is the pervasive dearth of empirical evidence of their validity. Very few of the existing CBTI systems have been validated in even a rudimentary manner (Butcher, 1987; Finn & Butcher, 1991; Matarazzo, 1986). Rather than actuarial programs based exclusively on empirically derived base rate data, most CBTI reports are based on clinical lore or the conclusions and hypotheses of expert clinicians (Gregory, 1996; Groth-Marnat & Schumaker, 1989). Concerns regarding validity are often expressed as part of a larger concern about the wide range in product quality within the CBTI market. Butcher (1987) wondered at the "mind-boggling" (p. 6) array of computers and software packages from which clinicians might choose. While some systems are based on reasonably rigorous development procedures, most are not. As a result, many psychologists remain quite skeptical of CBTI programs (Burke & Normand, 1987) and avoid using them.

4.17.4.3 Depersonalizing the Assessment Process

Critics of computerized assessment systems often express concern that introduction of computers in the assessment process will serve to heighten the distance between the clinician and client and lead to a more sterile and perhaps dehumanizing assessment process (Burke & Normand, 1987; Krug, 1987). In fact, research does not substantiate this concern and instead, suggests quite the opposite. Most people appear to prefer computer administered tests and some are more truthful in response to computer administered questions (Fowler, 1985). Computer administered assessments appear particularly useful for more disturbed or anxious clients.

4.17.4.4 Potential for Misuse and Client Harm

Later in this chapter we will address a range of potential ethical problems inherent in the use of computer assisted assessment. The potential for misuse of CBTIs and resulting harm to clients is substantial and results from difficulties with their current use (Butcher, 1987; Groth-Marnat & Schumaker, 1989). First, because CBTI programs and services are widely available and sold to a wide range of professionals, it is likely that professionals without adequate awareness of the limitations of CBTIs will apply them to clients regardless of context or important mitigating factors. A related concern involves the potential for factors indigenous to computerized assessment, but irrelevant to the purposes of the test, to significantly alter test performance (Moreland; 1985). Second there may be a tendency for consumers to uncritically accept statements generated by a computer as more factual than those generated by a clinician. Finally, because CBTI reports are rarely signed, there are serious concerns regarding professional responsibility and legal culpability (Gregory, 1996). Without a qualified psychologist assuming responsibility for the service offered, potential for misuse is enhanced.

4.17.4.5 Computer as Clinician

The final potential disadvantage in employment of computers in the assessment process has to do with the danger that well trained psychologists might lose control (Krug, 1987) of the assessment process. Specifically, this would be a loss to mechanization, technology, and large marketing interests. Though clinicians clearly stand to benefit from the speed and reliability of computer driven assessments, there is concern that the expert clinician of the past will become the testing technician of the future (Butcher, 1987; Butcher, Keller, & Bacon, 1985; Groth-Marnat & Schumaker, 1989; Matarazzo, 1990). As as result of the highly professional appearance of many CBTI reports, both users and recipients of computerized narratives may confuse them with comprehensive assessments. To the extent that the skills of the human clinician are relegated to a position of diminished importance in generating assessment outcomes, and to the extent that clinicians perceive themselves as less responsible for these outcomes, there certainly exists greater risk to the profession and the consumer.

4.17.5 ETHICAL ISSUES

In light of the rapid proliferation of CAPA techniques, and CBTIs in particular, it is not surprising that this burgeoning area of research and practice has become the focus of a wide range of ethical concerns. The APA's Ethical Principles and Code of Conduct (APA, 1992) has as its primary goal "the welfare and protection of the individuals and groups with whom psychologists work" (p. 1599). Principle C from the APA code is perhaps most relevant to the ethical issues inherent in computerized assessment, "Psychologists uphold professional standards of conduct, clarify their professional roles and obligations, accept appropriate responsibility for their behavior and adapt their methods to the needs of different populations" (p. 1599). With these goals in view, we will consider several of the most salient ethical obligations for psychologists involved in CAPA. If neglected, each could serve as a source of potential harm to consumers of computerized services.

4.17.5.1 Test Development

The Ethics Code (APA, 1992) states clearly that psychologists involved in the development and provision of computerized assessment services accurately describe the purpose, development procedures, norms, validity, reliability, and applications of the service as well as particular qualifications or skills required for their use. Psychologists participating in such product development should attempt to clearly link interpretive statements to specific client scores or profiles, qualify narrative statements to minimize the potential for misinterpretation and perhaps provide some form of warning statement to alert users to the potential for misinterpretation (Hofer & Green, 1985). At the very least, the developer might note that the clinical interpretations offered in narrative printouts are not to serve as the sole basis on which important clinical decisions are made (Matarazzo, 1986). Adherence to the highest standard of the profession would also require developers to provide rather detailed information regarding the system's development and structure in a separate manual. Because individual users are responsible for determining the validity of any CBTI for individual test-takers, availability of such system information is critical. Bersoff and Hofer (1991) noted that in spite of the apparent conflict between the developer's proprietary interest in the product and the clinician's need to responsibly evaluate the service, open and critical review of tests and CBTIs is critical for ensuring the quality of such materials and upholding the profession's ethical code.

4.17.5.2 Basis for Scientific and Professional Judgments

Psychologists rely on scientifically and professionally derived knowledge when making both scientific and professional judgments. The Ethics Code (APA, 1992) also states explicitly that "Psychologists select scoring and interpretation services (including automated services) on the basis of evidence of the validity of the program and procedures as well as on other appropriate considerations" (p. 1604). While psychologists are clearly compelled to justify their professional statements and behavior with empirically derived evidence, the current state of CBTI development makes this a difficult task indeed. The overwhelming majority of automated interpretive programs lack even preliminary forms of established validity (Lanyon, 1984). With this concern in mind, Matarazzo (1986) insisted that until research establishes (even the most primitive) validity for CBTIs, "It is essential that they be used only as tools by the clinician trained in their use and not as equivalents of, and thus substitutes for, professional education and training" (p.14).

Empirical validation of CBTI systems is exceptionally difficult given the exhaustive range of potential test scores and profiles. As a result, there are no purely actuarial interpretive programs in existence (Butcher et al, 1985; Fowler, 1985). Instead, most CBTIs offer a form of automated clinical prediction (Graham, 1993) in which published research, clinical hypotheses, and clinical experience on the part of an "expert" clinician are integrated into interpretive narrative statements. Nonetheless, the validity of such statements cannot be assumed by users of such reports and must be demonstrated every bit as much as the validity of the test on which it is based. Psychologists who employ CBTI reports must clearly understand the basis for the statements offered by such services, their validity or lack thereof, and take reasonable steps to ensure that those with whom they work are not harmed by the irresponsible or uncritical use of such reports.

4.17.5.3 Describing the Nature of Psychological Services

In his review of the state of personality assessment for the *Annual Review of Psychology*, Lanyon (1984) offered a stern indictment of the CBTI industry. He noted that available literature regarding these programs appeared to come in three essential varieties. These included:

> (a) glossy promotional literature sometimes masquerading as scientific data and usually accompanied by sample records, (b) studies of customer and user satisfaction, which has never been much of a problem, and (c) an occasional paper giving actual information about the development or validation of an automated system. (p. 690)

Lanyon was particularly distressed by the fact that gross deficits in program validity appeared to have become the norm for CBTI systems. The primary ethical issue of concern here has to do with the manner in which psychologists describe and/or promote computerized assessment services. Two sections of the Ethics Code (APA, 1992) have particular relevance here. First, when describing the nature and results of psychological services, "Psychologists provide, using language that is reasonably understandable to the recipient of those services, appropriate information later about results and conclusions" (p. 1600). Second, the section on avoidance of false or deceptive statements emphasizes that psychologists refrain from making false, deceptive, or misleading statements either by virtue of what they convey or omit concerning their services and work activities. This emphasis on avoiding deception also extends to descriptions of the scientific or clinical basis for, or results or degree of success of psychologists services. The implication of these standards would suggest a rather clear mandate for psychologists to explicitly describe those CBTI services they participate in developing, promoting, or utilizing in their clinical work. This would include descriptions of the process by which the program was constructed, the manner in which it generates interpretive material and any existing evidence of validity. By the same token, psychologists must be proactive in highlighting deficits in system validity or performance such that consumers and users might avoid harmful outcomes. Finally, psychologists should avoid misuse of their influence in the promotion of CBTIs and other computerized assessment products. "They [Psychologists] are alert to and guard against personal, financial, social, organizational, or political factors that might lead to misuse of their influence" (APA, 1992, p. 1601). Endorsements by respected psychologists in the field of assessment are often coveted and highly promoted by product marketers. The Ethics Code clearly warns against irresponsible promotion of CBTI products in a manner which might compromise the profession or increase misuse of such materials by other professionals and consumers.

4.17.5.4 Competence

Establishing and maintaining an appropriate level of competence in the area of computerized psychological assessment may be one of the most significant areas of ethical risk for psychologists at this time. The Ethics Code (APA, 1992) makes numerous references to the importance of competence for psychologists. The first ethical principle in the Ethics Code relates to competence and stresses that psychologists maintain a high degree of competence in their work and clearly articulate the boundaries of their expertise. "Psychologists function only within boundaries of competence based on education, training and supervised experience or appropriate professional experience" (p. 1599). Further, they provide services in new areas of practice "only after first undertaking appropriate study, training, supervision and/or consultation from persons competent in those areas" (p. 1600). Initially establishing and demonstrating competence in the area of CAPA would not appear to be adequate, however, as psychologists are additionally enjoined by the Ethics Code to maintain their expertise as long as they practice as psychologists. "Psychologists maintain a reasonable level of awareness of current scientific and professional information in their fields of activity and undertake ongoing efforts to maintain competence in the skills they use" (p. 1600).

Development of clear standards and guidelines for use of CBTIs in particular will serve to protect consumers and the profession. They might also serve as critical guides to psychologists, judges, and the courts in determining the standard of practice in this area (Hofer & Green, 1985). Bersoff and Hofer (1991) noted that establishing a prevailing "standard of care" in the area of CBTIs is critical to determining whether a test program's user, developer, or publisher violated a prevailing standard and is therefore ethically out of compliance or legally culpable. Examples of such noncompliance might include negligent entry of data, selection of a scoring system the psychologist should know is inappropriate for a client or unreasonable reliance on interpretive material from a CBTI narrative report. Problematically, the wide availability of CBTI and other computerized assessment services to persons of varied professional and educational backgrounds (Fowler, 1985), has rendered development of guidelines for competence in this area quite difficult.

The current Ethics Code more explicitly addresses the practice of psychological assessment and offers a clearer picture of how "competence" in this area might be defined:

> Those who develop, administer, score, interpret or use psychological assessment instruments do so in a manner and for purposes that are appropriate in light of research or on evidence of the usefulness and proper application of the techniques ... Psychologists who perform interventions or administer, score, interpret, or use assessment techniques are familiar with the reliability, validation, and related standardization or outcome studies of and proper applications and uses of the techniques they use. (APA, 1992, p. 1603)

Although broad, the Ethics Code does suggest several primary areas in which psychologists should have reasonable expertise if they are to competently utilize computer generated assessment material in their work with clients.

As a result of the ethical requirement for practitioners to evaluate the soundness of CBTI reports, they must obviously be qualified to interpret the test themselves. This requires not only basic familiarity with psychometric principles but also a rather detailed understanding of the manner in which the particular system in question was developed and generates interpretive material. On the basis of this understanding, the psychologist might then be able to reject, modify, or expand reports for particular clients (Hofer & Green, 1985). Psychologists will need to be familiar with three components of the CBTI services they utilize in order to do this effectively. These include (i) the examinees score on the relevant test or scale, (ii) the test scale or combination of scales on which interpretations are based, and (iii) research or clinical evidence supporting the interpretation.

In addition to basic competence in psychometrics and interpretation of CBTI assessment findings for individual clients, psychologists must also demonstrate competence in the appropriate monitoring of CBTI data. The Ethics Code requires psychologists to make reasonable efforts to maintain the security of tests and other assessment techniques (APA, 1992). To do this effectively in the area of CAPA, psychologists must establish formal procedures for retaining, reviewing, and releasing computerized assessment data. Tranel (1994) noted that psychologists bear responsibility for determining whether those requesting test data are qualified to interpret it appropriately. The concern here relates to nonexperts drawing erroneous conclusions based on naive use of CBTI reports. It appears that psychologists must not only avoid irresponsible use of computerized test data themselves, but must also prevent the same on the part of others, "Psychologists do not misuse assessment techniques, interventions, results and interpretations and take reasonable steps to prevent others from misusing the information these techniques provide" (APA, 1992, p. 1603). Related to the release of test data is concern that test questions or CBTI system information may become part of the public domain, resulting in risk of invalidation of tests as well as potential violation of copyright laws and contractual obligations (APA, 1996; Tranel, 1994). Because many CBTI reports include printouts of the client's raw scores as well as those critical items endorsed, users must exercise the same approach to maintaining test security they might employ with any other form of test data.

4.17.5.5 Professional Context

Perhaps the most alluring and potentially dangerous property of CBTI narrative reports is their polished, professional, and thorough appearance. Psychologists, like other users of these services, may be tempted to rely excessively on information from such computerized narratives without adequate interaction with the individual client or reasonable consideration of the unique circumstances in which the client presents for evaluation. The Ethics Code (APA, 1992) rather clearly addresses this concern,

> Psychologists perform evaluations, diagnostic services or interventions only within the context of a defined professional relationship ... Psychologists' assessments, recommendations, reports and psychological diagnostic or evaluative statements are based on information and techniques (including personal interviews of the individual when appropriate) sufficient to provide appropriate substantiation for their findings. (p. 1603)

Most concur that computerized assessment narratives must be carefully reviewed for appropriateness and "fit" with the examinee in light of research, complete information about the examinee, and solid professional judgment (Carson, 1990; Fowler, 1985; Graham, 1993; Hofer & Green, 1985).

Matarazzo (1990) made the case that one of the primary sources of ethical and professional danger in this regard has been a rather subtle but progressive loss of distinction between psychological assessment as a professional activity and mere testing.

> Objective psychological testing and clinically sanctioned and licensed psychological assessment are vastly different, even though assessment usually includes testing ... Psychological assessment is engaged in by a clinician and a patient in a one-to-one relationship and has statutorily defined or implied professional responsibilities ... Specifically, it [assessment] is the activity of a licensed professional, an artisan familiar with the accumulated findings of his or her young science. (p. 1000)

Similarly, Carson (1990) called for defense of "clinicianship" (p. 437) within psychological assessment and highlighted many of the dangers inherent in considering CBTI data apart from other primary client information and without the benefit of a clear client–professional relationship. Those portions of the Ethics Code addressing utilization of assessment results and computerized scoring and interpretive services are clear that psychologists retain full responsibility for conducting competent and context appropriate assessment, versus context blind and potentially harmful psychological testing, "Psychologists retain appropriate responsibility for the appropriate application, interpretation and use of assessment instruments, whether they score and interpret such tests themselves or use automated or other services" (APA, 1992, p. 1604). A related but unresolved concern, however, is how psychologists are to retain such responsibility when most CBTI reports are not signed by a responsible psychologist. Matarazzo (1986) highlighted this problem:

> Although it is not yet part of psychology's code of ethics, my experience leaves no question that computerized clinical interpretations offered in a professional setting about a person's intellectual, personality, brain-behavior and other highly personal characteristics constitutes a legally and professionally significant invasion of privacy and requires at the least, that the individuals offering these clinical interpretations sign their names to such consultations just as is done in every other profession. (p. 21)

4.17.5.6 CAPA with Special Populations

Related to the foregoing concern regarding the professional context of assessment and the psychologist's responsibility for ensuring the accuracy and appropriateness of computer generated assessment material, is concern regarding the implications of CBTIs for special client populations. The APA Ethics Code requires sensitivity to and respect for human differences among examinees and clients. This includes sensitivity to differences across such domains as age, gender, race, ethnicity, national origin, religion, sexual orientation, disability, language, and socioeconomic status. The Ethics Code specifically states that psychologists "remain vigilant for situations in which adjustments must be made in administration or interpretation secondary to individual or contextual factors" (APA, 1992, p. 1603). The rather clear implication here is that computer-assisted interpretive system results must be passed through the clinician's own interpretive grid with an eye toward identification of demographic or contextual factors on the part of the client which might raise concern about the validity of the findings (Bersoff & Hofer, 1991). This of course demands that the psychologist understands, and has access to, differences in base rates for specific demographic groups with respect to both the test and the interpretive program in question.

As a result of the burgeoning of CBTI systems, the proliferation of unsatisfactory and typically invalidated systems and the widespread marketing of such systems to unqualified users, there have been frequent calls in the scholarly and professional literature for development of standards and regulations relative to computerized approaches to assessment (Burke & Normand, 1987). In 1986 the APA published a set of guidelines for practice by psychologists in this arena. *Guidelines for computer-based test interpretations* (APA, 1986) offered a set of brief and general aspirational guidelines for developers and users of computerized assessment techniques and products. Conscious of the foregoing summary of salient ethical concerns in the field of CAPA, we will now offer a synopsis of those guidelines with a focus on why such guidelines are significant and how they might be applied by psychologists.

4.17.6 GUIDELINES FOR USERS OF COMPUTER-BASED TESTS AND INTERPRETATIONS

The following guidelines are intended for those professionals who use computer-based testing and interpretive services with those to whom they provide services. Table 2 contains the APA Guidelines for users of computer-based tests and interpretations (APA, 1986) and will serve as an outline for the current discussion.

Table 2 Guidelines for users and developers of computer-based tests and interpretations.

Guidelines for users

Administration

1. Influences on test scores due to computer administration that are irrelevant to the purposes of assessment should be eliminated or taken into account in the interpretation of scores.
2. Any departure from the standard equipment, conditions, or procedures, as described in the test manual or administrative instructions, should be demonstrated not to affect test scores appreciably. Otherwise, appropriate calibration should be undertaken and documented (see Guideline 16).
3. The environment in which the testing terminal is located should be quiet, comfortable, and free from distractions.
4. Test items presented on the display screen should be legible and free from noticeable glare.
5. Equipment should be checked routinely and should be maintained in proper working condition. No test should be administered on faulty equipment. All or part of the test may have to be readministered if the equipment fails while the test is being administered.
6. Test performance should be monitored, and assistance to the test-taker should be provided, as is needed and appropriate. If technically feasible, the proctor should be signaled automatically when irregularities occur.
7. Test-takers should be trained on proper use of the computer equipment, and procedures should be established to eliminate any possible effect on test scores due to the test-taker's lack of familiarity with the equipment.
8. Reasonable accommodations must be made for individuals who may be at an unfair disadvantage in a computer testing situation. In cases where a disadvantage cannot be fully accommodated, scores obtained must be interpreted with appropriate caution.

Interpretation

9. Computer-generated interpretive reports should be used only in conjunction with professional judgment. The user should judge for each test-taker the validity of the computerized test report based on the user's professional knowledge of the total context of testing and the test-taker's performance and characteristics.

Guidelines for developers

Human factors

10. Computerized administration normally should provide test-takers with at least the same degree of feedback and editorial control regarding their responses that they would experience in traditional testing formats.
11. Test-takers should be clearly informed of all performance factors that are relevant to the test result.
12. The computer testing system should present the test and record responses without causing unnecessary frustration or handicapping the performance of test-takers.
13. The computer testing system should be designed for easy maintenance and system verification.
14. The equipment, procedures, and conditions under which the normative, reliability, and validity data were obtained for the computer test should be described clearly enough to permit replication of these conditions.
15. Appropriate procedures must be established by computerized testing services to ensure the confidentiality of the information and the privacy of the test-taker.

Psychometric properties

16. When interpreting scores from the computerized versions of conventional tests, the equivalence of scores from computerized versions should be established and documented before using norms or cutting scores obtained from conventional tests. Scores from conventional and computer administrations may be considered equivalent when the rank orders of scores of individuals tested in alternative modes closely approximate each other, and the means, dispersions, and shapes of the score distributions are approximately the same, or have been made approximately the same by rescaling the scores from the computer mode.
17. The validity of the computer version of a test should be established by those developing the test.
18. Test services should alert test users to the potential problems of nonequivalence when scores on one version of a test are not equivalent to the scores on the version for which norms are provided.
19. The test developer should report comparison studies of computerized and conventional testing to establish the relative reliability of computerized administration.
20. The accuracy of computerized scoring and interpretation cannot be assumed. Providers of computerized test services should actively check and control the quality of the hardware and software, including the scoring, algorithms, and other procedures described in the manual.
21. Computer testing services should provide a manual reporting the rationale and evidence in support of computer-based interpretation of test scores.

Table 2 (continued)

Classification

22. The classification system used to develop interpretive reports should be sufficiently consistent for its intended purpose (see Chapter 2 of the 1985 *Testing Standards*). For example, in some cases it is important that most test-takers would be placed in the same groups if retested (assuming the behavior in question did not change).
23. Information should be provided to the users of computerized interpretation services concerning the consistency of classifications, including, for example, the number of classifications and the interpretive significance of changes from one classification to adjacent ones.

Validity of computer interpretations

24. The original scores used in developing interpretive statements should be given to test users. The matrix of original responses should be provided or should be available to test users on request, with appropriate considerations for test security and the privacy of test-takers.
25. The manual or, in some cases, interpretive report, should describe how the interpretive statements are derived from the original scores.
26. Interpretive reports should include information about the consistency of interpretations and warnings related to common errors of interpretation.
27. The extent to which statements in an interpretive report are based on quantitative research vs. expert clinical opinion should be delineated.
28. When statements in an interpretive report are based on expert clinical opinion, users should be provided with information that will allow them to weigh the credibility of such opinion.
29. When predictions of particular outcomes or specific recommendations are based on quantitative research, information should be provided showing the empirical relationship between the classification and the probability of criterion behavior in the validation group.
30. Computer testing services should ensure that reports for either users or test-takers are comprehensible and properly delimit the bounds within which accurate conclusions can be drawn by considering variables such as age or sex that moderate interpretations.

Review

31. Adequate information about the system and reasonable access to the system for evaluating responses should be provided to qualified professionals engaged in a scholarly review of the interpretive service. When it is deemed necessary to provide trade secrets, a written agreement of nondisclosure should be made.

4.17.6.1 Administration

Guidelines 1–8 clearly address the responsibility of the computerized system's user in ensuring that examinees are not adversely affected by the computerized administration itself. The environment must be conducive to comfort and maximal performance on the part of the examinee and extraneous influences on his or her performance should be minimized. While computerized administration generally appears to enhance the ease of test taking while decreasing overall time required (Green, 1991b), such programs should routinely assess for evidence that the examinee understands tests and procedures. If an examinee demonstrates any reservation or apprehension regarding interacting with a computer, a conventional format version of the test should be substituted when possible. In the future, a good deal more research is needed to better understand the impact of computers on examinee experience of the assessment process, satisfaction with the experience, and response to the results generated (Finn & Butcher, 1991).

4.17.6.2 Evaluation and Selection of CBTIs

The potential for profound misuse of testing material exists when individuals without adequate training are granted access to automated interpretive systems (Butcher, 1987). When the CBTI user is not a psychologist or lacks the background to effectively and reliably interpret tests, danger exists that CBTIs will not be appropriately scrutinized and carefully selected for the client in question. Moreland, Eyde, Robertson, Primoff, and Most (1995) reported on an attempt to establish basic test user qualifications. The authors found that the 86 identified testing competencies could be distilled to 12 "minimum competencies" and that these could be further reduced to two major categories of competence. Test-users must (i) possess adequate knowledge of the test (or test scoring and interpretation program) and its limitations, and (ii) accept responsibility for the competent use of the test (or computerized test program). Competent use of CBTIs requires that the psychologist carefully evaluate and select appropriate and reasonably validated interpretive programs.

Although well-designed interpretive programs which utilize a careful compilation of empirical data and expert opinion often yield more valid interpretive reports than those generated by clinicians (Green, 1991b), users must become familiar with and conversant regarding the program's established validity. Nonetheless, determining acceptable levels of validity for automated reports poses two problems. First, many tests themselves lack adequate development research and second, most interpretive programs have no established validity themselves or very little (Butcher, 1987). Therefore, it is critical for the prospective user to review the test manual, related documentation, and examples of the computerized reports prior to utilizing them with clients. It is particularly important to evaluate the system rationale for interpreting scores and profiles. How well are the classification and decisional criteria operationalized? Potential users are also encouraged to evaluate the credentials of the system's authors (Lanyon, 1987) and consider the authors' standing as both an expert clinician and a scholar in the field of CBTI.

Another basis on which to evaluate the potential value of a CBTI is that of general utility. As a rule, the rationale or justification for conducting an assessment hinges on provision of information of value with respect to planning and executing treatment (Hayes, Nelson, & Jarrett, 1987). A CBTI will have utility to the extent there is evidence that it contributes to positive treatment outcome, or in some way enhances the status of individuals for whom it is employed. Those systems high in utility will generally be highly efficient and usable, time and cost-effective, and able to discriminate effectively such that real between-examinee differences are detected (Krug, 1987).

Finally, Ben-Porath and Butcher (1986) offered several questions which might be utilized by prospective users of CBTI systems determining which system to employ. We believe these questions offer a good summary of the major concerns expressed in these guidelines. They include (i) to what extent has the validity of the reports been established? (ii) to what extent do the reports rely on empirical findings in generating interpretations? (iii) to what extent do the reports incorporate all of the currently available validated descriptive information? (iv) do the reports take demographic variables into account? (v) are different versions of the report available for different referral questions? (vi) do the reports include practical suggestions? and (vii) are the reports periodically revised to reflect newly acquired information?

4.17.6.3 Interpretation

The APA guidelines (APA, 1986) are clear in warning that CBTIs should never be used as a singular indicator of a person's characteristics, psychological functioning, or diagnosis. Rather, professional judgment is required to determine the extent to which the computerized report is valid and appropriate in light of the total context of testing and test-taker's specific performance and characteristics. Many CBTI program developers and users express similar concern about the danger inherent in divorcing the skilled psychologist from the interpretation and use of computerized reports (Bersoff & Hofer, 1991; Butcher, 1987; Finn & Butcher, 1991; Groth-Marnat & Schumaker, 1989). The polished and objective appearance of narrative reports may lend to the temptation to accept them as valid without adequate scrutiny of their match with the examinee. Bersoff and Hofer (1991) noted

> There must be an interposition of human judgment between the CPTI report and decision making to ensure that decisions are made with full sensitivity to the nuances of test administration and interpretation, and that the unique constellation of attributes in each person is evaluated. (p. 241)

Responsible assessment will necessarily involve a multistep process (Finn & Butcher, 1991) in which computerized assessment results are skillfully integrated with other sources of information about the examinee. Currently, computers are not capable of offering a sophisticated synthesis of psychological test results. This integrative function appears to be squarely within the purview and professional responsibility of the psychological diagnostician. Without such integration, psychological reports are necessarily overly general, nonspecific, and questionably accurate. They would certainly fail to capture the test-taker's cognitive, affective, and behavioral functioning across a variety of situations (Bersoff & Hofer, 1991).

The prevailing standard ethically and legally appears to be that the computer-based report is a professional-to-professional consultation (Butcher, 1987). In this way, the computer is merely the equivalent of a library or consultant which might offer the most frequently indicated inferences or correlates for specific test scores or profiles. With the computer as librarian or "look up table" (Butcher, 1987, p. 9), the final determiner of the accuracy and adequacy of the computer-based report is the psychologist who receives the report.

An excellent example of adherence to this guideline is offered by Finn & Butcher (1991) who quote the disclaimer from the Clinical Interpretive Report for the MMPI. This disclaimer is printed on each interpretive report and highlights for users how the report should be utilized, "This MMPI interpretation can serve as a useful source of hypotheses about clients ... the personality descriptions, inferences, and recommendations contained herein need to be verified by other sources of clinical information since individual clients may not fully match the prototype" (p. 367).

4.17.7 GUIDELINES FOR DEVELOPERS OF COMPUTER-BASED TEST SERVICES

The following guidelines apply most directly to those involved in the construction, validation, and marketing of computerized assessment systems for use by others. Table 2 contains the guidelines for developers of computer-based test services (APA, 1986).

4.17.7.1 Human Factors

Guidelines 10–15 suggest developers of CAPA systems are broadly responsible to ensure that computerized administration of tests does not hamper the examinees' performance in any way. Examinees must also maintain reasonable control of the testing process and the testing service must assume full responsibility for establishing appropriate procedures for ensuring the confidentiality of data collected and the privacy of the examinee. This may be particularly challenging in light of the development of on-line test scoring and interpretation services. On-line access to client information must be carefully controlled by the testing service.

4.17.7.2 Psychometric Properties

Guidelines 16–21 in Table 2 require test or CBTI program developers to carefully evaluate and communicate the psychometric quality of the test or computerized interpretive system. When developing computerized versions of conventional tests, norms and cutting scores for the conventional test can only be used if the computer and conventional form are found to be equivalent (Green, 1991b). Not only should correlations between the two versions be high, but score distributions must be generally equivalent as well. Developers are expected to communicate results of equivalency studies to prospective users and to take initiative in "alerting" users to potential problems resulting from nonequivalence between forms.

4.17.7.3 Classification Strategy

When a computerized assessment system utilizes a classification system based on cutting scores, system developers must offer a convincing rationale for the particular system adopted and the cutting scores selected. In order for a CBTI system to be fully actuarial, system output is determined solely by statistical regularities that have been empirically demonstrated to exist between input and output data (Moreland, 1985). Instead, most systems combine actuarial and clinical expertise approaches. Users must be aware of the system strategy for integrating statistical and clinical prediction in the service of classifying examinees. Additionally, developers bear responsibility for communicating to users the consistency of the classification system and the meaning associated with changes between categories.

4.17.7.4 Validity of Computer Interpretations

As indicated by Guidelines 24–30, developers of computer-based test services are responsible for communicating information to users concerning the system's validity. Not only should validity data be made available to users, but developers must also clarify the extent to which interpretive statements are based on expert clinical opinion or quantitative research. Users must be shown the connection between such research or clinical opinion and specific interpretive statements and classification decisions. Further, developers are to warn users of common errors and potential pitfalls associated with the interpretive system. One of the reasons for the pervasive problem with establishing the validity of CBTI systems has to do with the practice of utilizing clinician generated reports as the criterion in validity research (Moreland, 1985). At times, computer reports may be more accurate than their clinician generated counterparts, thus falsely lowering validity coefficients. Developers should consider alternative criterion variables when possible.

Lanyon (1987) described six factors which should be considered by CBTI system developers in the service of increasing the validity of CBTI programs. First, reliability coefficients for both predictor and criterion variables have been unacceptably low for most CBTIs and should be increased. Second, in balancing the "bandwidth-fidelity" tension, CBTI developers should develop test indices that lead to single, focused predictions (narrow bandwidth/high-fidelity) versus trying to say too much from too little data. Third, departures from empirical data should be minimized. When test interpretations are based on data without attempts to polish, cluster, or otherwise alter them for the sake of appearance, errors in interpretation are minimized. Fourth, unwarranted generalizations from the standardization sample of CBTIs is a major source of invalidity. When the gap between the population from which an interpretive system was derived and the population on which it is used is substantial, validity declines and erroneous interpretations abound. Fifth, Lanyon recommends that an "unclassifiable" option be vigorously employed. Adding such a category (versus forcing predictive or interpretive statements for every profile) substantially enhances system validity. Finally, whenever possible, different base rates for characteristics being predicted or described should be employed in the interpretive system.

4.17.7.5 Facilitation of Review

The APA guidelines (APA, 1986) highlight the requirement for CBTI system developers to provide adequate information about the system as well as reasonable access to this information on the part of qualified professionals engaged in reviewing the interpretive system. Previous reviewers of this topic appear to concur that availability of detailed development information and data is critical to responsible evaluation of CBTI systems and their usefulness to clinicians (Green, 1991b; Lanyon, 1987; Roid & Gorsuch, 1984; Snyder et al., 1990). High quality interpretive programs clearly label the program using standardized descriptions which clarify the function of the system and the specific manner in which it generates interpretive statements. Such programs provide detailed data relative to development of the system and extensive references to the empirical basis for the decision rules used.

4.17.8 DIRECTIONS FOR THE FUTURE

Various authors have speculated about the future role and developmental course of CAPA. These have included highly optimistic outlooks such as that of Moreland (1985), "I am confident that the computer will eventually replace the professional for most, but not all, assessment functions ... and this may happen sooner rather than later" (p. 229), as well as more pessimistic perspectives based on fears about the dehumanization of the assessment process (Matarazzo, 1986).

A primary problem in the development in CAPA is the substantial lag in "psychotechnology" currently evident in the assessment field. Specifically, our understanding of assessment appears to lag behind available computer technology. Along these lines, there is a rather profound need for better research designs and demonstrations of program validity (Lanyon, 1984) among CBTI systems. Currently, a number of such programs offer little in the way of adherence to APA (1986) guidelines for development of CBTI software.

Hopefully, the future holds more promise for empirical research on the reliability and validity of CBTI systems. As more powerful desk-top (and lap-top) computers expand, increasingly complex statistical analyses are possible for test developers who have the expertise to use them. No longer is computer power located in a small number of central or "mainframe" facilities. Greater resources and more creativity in research designs for the study of CBTI are certainly needed. Creative methods of linking test and nontest data, in cost-effective formats, will also be crucial to the development and long-term refinement of CBTI. Test developers should study the classic examples given by Lachar (1984) and the methods reviewed by Snyder et al. (1990) and find ways to connect clinician's ratings, medical and treatment histories, and demographic data to the test-score data file. An important area of expansion in neuropsychological assessment will be the connection of imaging technology (e.g., magnetic resonance imaging [MRI] results) to psychological and cognitive test scores.

Alternatively, the field of CBTI should shift more toward a model of computer as "research assistant" (Roid, 1985), and provide methods for the clinician to access research findings, empirical correlates, and, perhaps, brief and verifiable narrative descriptions such as symptom lists from clinician checklists. However, in the "assistant" model, the final selection and control of narrative statements would be retained as a function of the clinician, not the computer program. As mentioned previously, the word-processing capabilities of systems such as WISC-III Writer (Psychological Corporation, 1994) should make correlated empirical findings available in a "scrapbook" but not be automatically printed in a report. Such systems could become even more elaborate in the future, employing extensive database functions and multimedia graphics, to supplement the conventional test-score results. Much could be done in the future to have even more demographic and historical data available for each client, as long as privacy is protected. The control of the assembly, selection, and composition of the final collection of, for example, data, statements, and graphs, should remain in the hands of the experienced clinician.

Another area of future development will surely be in the expansion of the multimedia capability of CAPA programs. All types of CAPA, from administration to scoring to interpretation could employ more sophisticated graphic and, perhaps, video capability. Gregory (1996) reports on the development of multimedia "situational" tests being developed at IBM that depict actual on-the-job scenes. The computer may briefly "pause" a video display and ask the examinee to answer questions, or predict the "next step" in the scenario, for purposes of assessing the examinee's sensitivity to interpersonal or technical concerns on the job. For example, the senior author of this chapter assisted in a research project to develop qualification tests for parole officers who were being screened for ability to perform "combative arrests." Film and audio depictions of scenarios with real parole officers were studied to identify critical behaviors such as advanced planning of back-up assistance, prediction of the parolee's potential route in the event of an escape attempt, and physical strategy for placing handcuffs on the subject. Extensive studies of physical movements involved in arrest scenarios were conducted to determine hand and arm strengths required. However, it became clear that many skills were "cognitive" rather than physical and included skills such as planning, prediction, and knowledge of typical parolee strategies of escaping arrest. Such an evaluation might be depicted on computerized video collections, shown to potential applicants, and present questions of strategy and planning. Clearly, extensive empirical validation with actual parole officers, including those who were previously judged to be "expert" in combative arrest, would be essential to the development of such a system. The key element of this "futuristic" system would be the video reenactments, presented on multimedia computer equipment—a method of simulation that has been available on motion-picture film for decades. In the future, more precise computer-adaptive testing may be possible for this medium.

One of the more promising areas for future development is that of adaptive testing. By administering only those test items necessary to draw supportable clinical conclusions, psychologists should be able to substantially reduce testing time while enhancing both the reliability and validity of findings (Krug, 1987). Computer-generated and tailor-made personality tests are a particularly interesting possibility.

Finally, the debate regarding the role of computers in the assessment process will certainly continue in the future around the issue of the role of the clinical psychologist. Butcher (1987) noted

> As interest in developing integrated clinical diagnostic reports broadens, more research or system adequacy will be stimulated, and, no doubt, more intense dialogue will be generated on the appropriateness of machines to perform what is believed by some to be an essentially human activity. (p. 11)

4.17.9 CONCLUSIONS AND RECOMMENDATIONS

In summary, the best recommendation for clinical practice, given the current state of validation of narrative computer-interpretive programs is for clinicians to "draw a line" between scoring programs and narrative-interpretive programs. Computer administration of tests and scoring (including elaborate statistical scoring and graphical display) can clearly assist the clinician in terms of efficient use of time and accuracy of calculation. Unless the following conditions are satisfied, it is recommended that CBTI programs with extensive narrative reports be used only for hypothesis generation and never as an unedited section of a psychological report:

(i) *Clinician retains control of word/sentence selection.* If the narrative program gives control to the clinician for assembly of narrative descriptions, then the unique situation of the client can be used to temper the assessment narrative. Automatically printed statements should never be used unedited. Thus, CBTI programs should be "research assistants," not "automated interpreters."

(ii) *Test developers apply actuarial research and document it.* The technical manual accompanying CBTI programs should clearly describe the details of the empirical validation of all descriptive statements, list all cutting scores and their validation studies, present classification accuracy statistics for all profile rules employed, and provide cautionary statements of study limitations.

(iii) *Researchers of CBTI should develop more cost-effective research designs.* Given that the cost of validation is a frequent excuse given for lack of empirical study, statisticians and research experts are encouraged to creatively design new types of studies in cooperation with CBTI developers so the field of CAPA can advance on a more scientific basis where possible.

(iv) *Psychologists should be vigilant in distinguishing assessment from testing.* In the interests of preventing the erosion of the meaning of clinical assessment, as emphasized by Matarazzo (1986, 1990), all psychologists should be alerted to the concern that the final evaluation of all relevant client information, including the context, history, and uniqueness of the client, be reserved for the experienced, trained clinician, not the computer.

4.17.10 REFERENCES

Adams, K. M., & Shore, D. L. (1976). The accuracy of an automated MMPI interpretation system in a psychiatric setting. *Journal of Clinical Psychology, 32,* 80–82.

American Psychological Association. (1985). *Standard for educational and psychological testing*. Washington, DC: Author.

American Psychological Association. (1986). *Guidelines for computer-based tests and interpretations*. Washington DC: Author.

American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47,* 1597–1611.

American Psychological Association. (1996). Statement on the disclosure of test data. *American Psychologist, 51,* 644–648.

Andrews, L. W., & Gutkin, T. B. (1991). The effects of human versus computer authorship on consumers' perceptions of psychological reports. *Computers in Human Behavior, 7,* 311–317.

Angoff, W. H., & Huddleston, E. M. (1958). The multi-level experiment: A study of two-stage system for the College Board SAT (Statistical Report No. 58–21). Princeton, NJ: Educational Testing Service.

Assessment Systems (1990). *MicroCAT testing system manual.* Minneapolis, MN: Author.

Barclay, J. R. (1983). *Barclay classroom assessment system manual.* Los Angeles: Western Psychological Services.

Bayroff, A. G., & Seeley, L. C. (1967, June). *An exploratory study of branching tests*. Technical Research Note 188, US Army Behavioral Science Research Laboratory.

Ben-Porath, Y. S., & Butcher, J. N. (1986). Computers in personality assessment: A brief past, an ebullient present, and an expanding future. *Computers in Human Behavior, 2,* 167–182.

Bersoff, D. N., & Hofer, P. J. (1991). Legal issues in computerized psychological testing. In T. B. Gutkin & S. L. Wise (Eds.), *The computer and the decision-making process* (pp. 225–244). Hillsdale, NJ: Erlbaum.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), *Statistical theories of mental test scores*. Reading, MA: Addison Wesley.

Bloom, B. L. (1992). Computer-assisted psychological intervention: A review and commentary. *Clinical Psychology Review, 12,* 169–197.

Burke, M. J., & Normand, J. (1987). Computerized psychological testing: Overview and critique. *Professional Psychology: Research and Practice, 18,* 42–51.

Butcher, J. N. (1985). Introduction to the special series. *Journal of Consulting and Clinical Psychology, 53,* 746–747.

Butcher, J. N. (1987). The use of computers in psychological assessment: An overview of practices and issues. In J. N. Butcher (Ed.), *Computerized psychological assessment: A practitioner's guide* (pp. 3–14). New York: Basic Books.

Butcher, J. N., Keller, L. S., & Bacon, S. F. (1985). Current developments and future directions in computerized personality assessment. *Journal of Consulting and Clinical Psychology, 53,* 803–815.

Carson, R. C. (1990). Assessment: What role the assessor? *Journal of Personality Assessment, 54,* 435–445.

Conners, K. (1995). *Conners' Continuous Performance Test computer program.* North Tonawanda, NY: Multi Health Systems.

Cowden, D. J. (1946). An application of sequential sampling to testing students. *Journal of the American Statistical Association, 41,* 547–556.

Dahlstrom, W. G. (1993). Tests: Small samples, large consequences. *American Psychologist, 48,* 393–399.

Educational Testing Service (1995). *Learning for tomorrow: 1995 Annual Report*. Princeton, NJ: Author.

Finn, S. E., & Butcher, J. N. (1991). Clinical objective personality assessment. In M. Hersen, A. E. Kazdin, & A. S. Bellack (Eds.), *The Clinical Psychology Handbook* (pp. 362–373). New York: Pergamon.

Flanagan, J. C., & Lindquist, E. F. (Eds.) (1951). *Educational measurement,* Washington, DC: American Council of Education.

Fowler, R. D. (1985). Landmarks in computer-assisted psychological assessment. *Journal of Consulting and Clinical Psychology, 53,* 748–759.

Fowler, R. D., & Butcher, J. N. (1986). Critique of Matarazzo's view on computerized psychological testing. *American Psychologist, 41,* 94–96.

Garb, H. N. (1994). Judgment research: Implications for clinical practice and testimony in court. *Applied and Preventive Psychology, 3,* 173–183.

Goldberg, L. R. (1968). Simple models or simple processes? Some research on clinical judgments. *American Psychologist, 23,* 483–496.

Graham, J. R. (1993). *MMPI-2: Assessing personality and psychopathology*. New York: Oxford University Press.

Green, B. F. (1991a). Computer based adaptive testing in 1991. *Psychology & Marketing, 8*(4), 243–257.

Green, B. F. (1991b). Guidelines for computer testing. In T. B. Gutkin & S. L. Wise (Eds.), *The computer and the decision-making process* (pp. 245–274). Hillsdale, NJ: Erlbaum.

Gregory, R. J. (1996). *Psychological testing: History, principles, and applications*. Boston: Allyn and Bacon.

Groth-Marnat, G., & Schumaker, J. (1989). Computer-based psychological testing: Issues and guidelines. *American Journal of Orthopsychiatry, 59,* 257–263.

Guastello, S. J., & Rieke, M. L. (1994). Computer-based test interpretations as expert systems. *Computers in Human Behavior, 10,* 435–455.

Guastello, S. J., Guastello, D. D., & Craft, L. L. (1989). Assessment of the Barnum effect in computer-based test interpretations. *The Journal of Psychology, 123,* 477–484.

Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1987). The treatment utility of assessment. *American Psychologist, 42,* 963–974.

Hofer, P. J., & Green, B. F. (1985). The challenge of competence and creativity in computerized psychological testing. *Journal of Consulting and Clinical Psychology, 53,* 826–838.

Huba, G. J. (1986). Interval banded profile analysis: A method for matching score profiles to "soft" prototypic patterns. *Educational and Psychological Measurement, 46,* 565–570.

Jackson, D. N. (1985). Computer-based personality testing. *Computers in Human Behavior, 1,* 255–264.

Kaufman, A. S. (1994). *Intelligent testing with the WISC-III.* New York: Wiley.

Kaufman, A. S. (1996). *Wechsler integrated interpretive system*. Odessa, FL: Psychological Assessment Resources.

Klinger, D. E., Miller, D., Johnson, J., & Williams, T. (1977). Process evaluation of an on-line computer-assisted unit for intake assessment of mental health patients. *Behavior Research Methods and Instrumentation, 9,* 110–116.

Kleinmuntz, B. (1969). Personality test interpretation by computer and clinician. In J. N. Butcher (Ed.), *MMPI: Research developments and clinical applications* (pp. 97–104). New York: McGraw-Hill.

Kleinmuntz, B. (1975). The computer as clinician. *American Psychologist, 30,* 379–387.

Krug, S. E. (1984). *Psychware: A reference guide to computer-based products.* Kansas City, MO: Test Corporation of America.

Krug, S. E. (1987). Microtrends: An orientation to computerized assessment. In J. N. Butcher (Ed.), *Computerized psychological assessment: A practitioner's guide* (pp. 15–25). New York: Basic Books.

Lachar, D. (1984). *WPS Test Report for the Personality Inventory for Children.* Los Angeles: Western Psychological Services.

Lanyon, R. I. (1984). Personality assessment. *Annual Review of Psychology, 35,* 667–701.

Lanyon, R. I. (1987). The validity of computer-based personality assessment products: Recommendations for the future. *Computers in Human Behavior, 3,* 225–238.

Lichtenstein, S., & Newman, J. R. (1967). Empirical scaling of common verbal phrases associated with numerical probabilities. *Psychonomic Science, 9,* 563–564.

Linn, R. L., Rock, D., & Cleary, A. (1969). The development and evaluation of several programmed testing methods. *Educational and Psychological Measurement, 29,* 129–146.

Lord, F. M. (1968). Some test theory for tailored testing. *Research Bulletin* RB-68–38. Princeton, NJ: Educational Testing Service.

Lord, F. M. (1971). The self-scoring flexilevel test. *Journal of Educational Measurement, 8,* 147–151.

Marks, P. A., & Seeman, W. (1963). The actuarial description of abnormal personality. Baltimore: Williams and Wilkins.

Matarazzo, J. D. (1983). Computerized psychological testing (Editorial). *Science, 221,* 323.

Matarazzo, J. D. (1986). Computerized clinical psychological test interpretations: Unvalidated plus all mean and no sigma. *American Psychologist, 41,* 14–24.

Matarazzo, J. D. (1990). Psychological assessment versus psychological testing. *American Psychologist, 45,* 999–1017.

McBride, J. R. (1988, August). *A computerized adaptive version of the Differential Aptitude Test.* Paper presented at the annual meeting of the American Psychological Association, Atlanta.

Meehl, P. E. (1954). *Clinical vs. statistical prediction.* Minneapolis, MN: University of Minnesota Press.

Meehl, P. E. (1956). Wanted: A good cookbook. *American Psychologist, 11,* 263–272.

Mitchell, J. V., & Kramer, J. J. (1985). Computer-based assessment and the public interest: An examination of the issues and introduction to the special issue. *Computers in Human Behavior, 1,* 203–305.

Moreland, K. L. (1985). Computer-assisted psychological assessment in 1986: A practical guide. *Computers in Human Behavior, 1,* 221–233.

Moreland, K. L. (1992). Computer-assisted psychological assessment. In M. Zeidner and R. Most (Eds.) *Psychological testing: An inside view.* Palo Alto, CA: Consulting Psychologists Press.

Moreland, K. L., Eyde, L. D., Robertson, G. J., Primoff, E. S., & Most, R. B. (1995). Assessment of test user qualifications. *American Psychologist, 50,* 14–23.

O'Dell, J. W. (1972). P. T. Barnum explores the computer. *Journal of Consulting and Clinical Psychology, 38,* 270–273.

Patterson, J. J. (1962). *An evaluation of the sequential method of psychological testing.* Unpublished doctoral dissertation, Michigan State University.

Piotrowski, C., & Keller, J. W. (1989). Use of assessment in mental health clinics and services. *Psychological Reports, 64,* 1298.

Pohl, N. F. (1981). Scale considerations in using vague quantifiers. *Journal of Experimental Education, 49,* 235–240.

Powell, D., Kaplan, E., Whitla, D., Weintraub, S., Catlin, R., & Funkenstein, H. (1993). *MicroCog: Assessment of cognitive functioning.* San Antonio, TX: Psychological Corporation.

Prince, R. J., & Guastello, S. J. (1990). The Barnum effect in a computerized Rorschach interpretation system. *The Journal of Psychology, 124,* 217–222.

Psychological Corporation (1987). *Differential aptitude tests computerized adaptive edition manual.* San Antonio, TX: Author.

Psychological Corporation (1992). *Weschler scoring assistant manual.* San Antonio, TX: Author.

Psychological Corporation (1994). *WISC-III Writer manual.* San Antonio, TX: Author.

Rasch, G. (1980). *Some probability models for aptitude and attainment tests.* Chicago: University of Chicago Press.

Roid, G. H. (1969). *Branching methods for constructing psychological test scales.* Unpublished doctoral dissertation, University of Oregon.

Roid, G. H. (1985). Computer-based test interpretation: The potential of quantitative methods of test interpretation. *Computers in Human Behavior, 1,* 207–219.

Roid, G. H. (1986). Computer technology in testing. In B. S. Plake & J. C. Witt (Eds.), *The future of testing* (pp. 29–69). Hillsdale, NJ: Erlbaum.

Roid, G. H., & Fitts, W. H. (1988). *Tennessee Self Concept Scale revised manual.* Los Angeles: Western Psychological Services.

Roid, G. H., & Gorsuch, R. L. (1984). Development and clinical use of test-interpretive programs on microcomputers. In M. D. Schwartz (Ed.), *Using computers in clinical practice* (pp. 141–149). New York: Haworth Press.

Roper, B. L., Ben-Porath, Y. S., & Butcher, J. N. (1991). Comparability of computerized adaptive and conventional testing with the MMPI-2. *Journal of Personality Assessment, 57,* 278–290.

Sands, W. A., & Gade, P. A. (1983). An application of computerized adaptive testing in Army recruiting. *Journal of Computer-Based Instruction, 10,* 37–89.

Silverstein, A. B. (1981). Reliability and abnormality of test score differences. *Journal of Clinical Psychology, 37,* 392–394.

Snyder, D. (1981). *Manual for the marital satisfaction inventory.* Los Angeles: Western Psychological Services.

Snyder, D., Widiger, T., & Hoover, D. (1990). Methodological considerations in validating computer-based test interpretations: Controlling for response bias. *Psychological Assessment, 2,* 470–477.

Sympson, J. B., Weiss, D. J., & Ree, M. J. (1982). *Predictive validity of conventional and adaptive tests in an Air Force training environment (AFHRL TR 81–40).* Brook Air Force Base, TX: Manpower and Personnel Division, Air Force Human Relations Laboratory.

Tranel, D. (1994). The release of psychological data to non experts: Ethical and legal considerations. *Professional Psychology: Research and Practice, 25,* 33–38.

Wainer, H. (Ed.) (1990). *Computerized adaptive testing: A primer.* Hillsdale, NJ: Erlbaum.

Wald, A. (1947). *Sequential analysis.* New York: Wiley.

Weiss, D. J. (Ed.) (1983). *New horizons in testing.* New York: Academic Press.

Wise, S. L., & Plake, B. S. (1990). Computer-based testing in higher education. *Measurement and evaluation in counseling and development, 23,* 3–9.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.18
Therapeutic Assessment: Linking Assessment and Treatment

MARK E. MARUISH
Strategic Advantage, Minneapolis, MN, USA

4.18.1 INTRODUCTION 526
4.18.2 THE CURRENT PRACTICE OF PSYCHOLOGICAL ASSESSMENT IN THE THERAPEUTIC ENVIRONMENT 528
4.18.3 PSYCHOLOGICAL ASSESSMENT AS A THERAPEUTIC ADJUNCT 530
  4.18.3.1 Psychological Assessment for Clinical Decision-making 530
  4.18.3.2 Psychological Assessment as a Treatment Technique 530
  4.18.3.3 Psychological Assessment for Outcomes Assessment 531
4.18.4 GENERAL CONSIDERATIONS FOR THE SELECTION AND USE OF PSYCHOLOGICAL TEST INSTRUMENTATION 531
  4.18.4.1 Types of Instrumentation for Therapeutic Assessment 532
    4.18.4.1.1 Psychological/psychiatric symptom measures 532
    4.18.4.1.2 Measures of general health status and role functioning 533
    4.18.4.1.3 Quality of life measures 534
    4.18.4.1.4 Service satisfaction measures 534
  4.18.4.2 Guidelines for Instrument Selection 535
    4.18.4.2.1 National Institute of Mental Health criteria 535
    4.18.4.2.2 Other criteria and considerations 536
4.18.5 PSYCHOLOGICAL ASSESSMENT AS A TOOL FOR SCREENING 538
  4.18.5.1 Research-based Use of Psychological Screeners 539
  4.18.5.2 Implementation of Screeners into the Daily Work Flow of Service Delivery 540
4.18.6 PSYCHOLOGICAL ASSESSMENT AS A TOOL FOR TREATMENT PLANNING 541
  4.18.6.1 Assumptions About Treatment Planning 541
  4.18.6.2 The Benefits of Psychological Assessment for Treatment Planning 542
    4.18.6.2.1 Problem identification 542
    4.18.6.2.2 Problem clarification 542
    4.18.6.2.3 Identification of important patient characteristics 543
    4.18.6.2.4 Monitoring of progress along the path of expected improvement 544
4.18.7 PSYCHOLOGICAL ASSESSMENT AS A THERAPEUTIC INTERVENTION 545
  4.18.7.1 What Is Therapeutic Assessment? 545
  4.18.7.2 The Impetus for Therapeutic Assessment 546
  4.18.7.3 The Therapeutic Assessment Process 546
    4.18.7.3.1 Step 1: The initial interview 547
    4.18.7.3.2 Step 2: Preparing for the feedback session 547
    4.18.7.3.3 Step 3: The feedback session 547
    4.18.7.3.4 Additional steps 548
  4.18.7.4 Empirical Support for Therapeutic Assessment 548
4.18.8 PSYCHOLOGICAL ASSESSMENT AS A TOOL FOR OUTCOMES MANAGEMENT 549
  4.18.8.1 What Are Outcomes? 549
  4.18.8.2 Outcomes Assessment: Measurement, Monitoring, and Management 550

4.18.8.3 The Benefits of Outcomes Assessment 550
4.18.8.4 The Therapeutic Use of Outcomes Assessment 550
4.18.8.4.1 Purpose of the outcomes assessment 551
4.18.8.4.2 What to measure 551
4.18.8.4.3 How to measure 552
4.18.8.4.4 When to measure 553
4.18.8.4.5 How to analyze outcomes data 554
4.18.9 FUTURE DIRECTIONS 555
4.18.9.1 What the Industry Is Moving Away From? 555
4.18.9.2 Trends in Instrumentation 556
4.18.9.3 Trends in Data Use and Storage 556
4.18.9.4 Trends in the Application of Technology 557
4.18.10 SUMMARY 558
4.18.11 REFERENCES 559

4.18.1 INTRODUCTION

The cost of health care in the USA has reached astronomical heights. In 1995, approximately $1 trillion, or 14.9% of the gross domestic product, was spent on health care, and a 20% increase is expected by the year 2000 (*Mental Health Weekly*, 1996a). The cost and prevalence of mental health problems and the accompanying need for behavioral health care services in the USA continue to rise at rates which give cause for concern. America's mental health bill in 1990 was $147 billion (*Mental Health Weekly*, 1996c). The Center for Disease Control and Prevention (1994) recently reported on the results of a survey of 45 000 randomly interviewed Americans regarding their quality of life. The survey found that one-third of the respondents reported they suffered from depression, stress, or emotional problems at least one day a month, and 11% percent of the sample reported having these problems more than eight days a month.

The American Psychological Association (APA; 1996) also reports statistics, summarized below, that bear attention.

(i) It is estimated that 15–18% of Americans suffer from a mental disorder; 14 million of these individuals are children.

(ii) Approximately eight million Americans suffer from depression in any given one-month period.

(iii) As many as 20% of Americans will suffer one or more major episodes of depression during their lifetime.

(iv) An estimated 80% of elderly residents in Medicaid facilities were found to have moderate to intensive needs for mental health services.

Moreover, information from various studies indicates that at least 25% of primary health care patients have a diagnosable behavioral disorder (*Mental Health Weekly*, 1996b).

The need for behavioral health care services is significant. In analyzing data from a 1987 national survey of 40 000 people in 16 000 households, Olfson and Pincus (1994a, 1994b) found that 3% of the population was seen for at least one psychotherapeutic session that year. Of these visits, 81% were to mental health professionals. Estimates provided by VandenBos, DeLeon, and Belar (1993) in the early 1990s indicated that in any year, 37.5 million Americans (or 15% of the population at that time) could benefit from mental health services.

What is the value of the services provided to those suffering from mental illness or substance abuse/addiction/dependency? Some might argue that the benefit is either minimal, or too costly to achieve if significant effects are to be gained. This is in the face of data which suggest otherwise. Numerous studies have demonstrated that treatment of mental health and substance abuse/dependency problems can result in substantial savings when viewed from a number of perspectives. This "cost offset" effect probably has been demonstrated most clearly in savings in medical care dollars over given periods of time.

Medical cost offset considerations are significant, given reports that 50–70% of usual primary care visits are for medical problems that involve psychological factors (APA, 1996). APA also reports that 25% of patients seen by primary care physicians have a disabling psychological disorder, and that depression and anxiety rank among the top six conditions seen by family physicians. Following are just a few of the findings supporting the medical cost benefits that can accrue from providing behavioral health care treatment.

(i) At least 25% or more of patients seen in a primary care setting have diagnosable behavioral disorders and use two to four times as many medical resources as those patients without these disorders (*Mental Health Weekly*, 1996b).

(ii) Sipkoff (1995) reported several conclusions, drawn from a review of numerous studies conducted between 1988 and 1994 and listed in the *Cost of addictive and mental disorders and effectiveness of treatment* report published by the Substance Abuse and Mental Health Services Administration (SAMHSA). One conclusion derived from a meta-analysis of offset effect was that treatment for mental health problems results in an approximately 20% reduction in the overall cost of health care. The report also concluded that while alcoholics were found to spend twice as much on health care as those without abuse problems, one-half of the cost of substance abuse treatment is offset within one year by subsequent reductions in the combined medical cost savings for the patient and his or her family.

(iii) Strain et al. (1991) found that screening a group of 452 elderly hip fracture patients for psychiatric disorders prior to surgery and providing mental health treatment to the 60% of the sample needing treatment reduced total medical expenses by $270 000. The cost of the psychological/psychiatric services provided to this group was only $40 000.

(iv) Simmons, Avant, Demski, and Parisher (1988) compared the average medical costs for chronic back pain patients at a multidimensional pain center (providing psychological and other types of intervention) during the year prior to treatment to those costs of the year following treatment. The pretreatment costs per patient were $13 284 while post-treatment costs were $5596.

The reader is referred to Friedman, Sobel, Myers, Caudill, and Benson (1995) for a detailed discussion of various ways in which behavioral interventions can both maximize care to medical patients and achieve significant economic gains.

APA (1996) has very succinctly summarized what appears to be the prevalent findings of the medical cost offset literature.

> (i) Patients with mental disorders are heavy users of medical services, averaging twice as many visits to their primary care physicians as patients without mental disorders.
>
> (ii) When appropriate mental health services are made available, this heavy use of the system often decreases, resulting in overall health savings.
>
> (iii) Cost offset studies show a decrease in total health care costs following mental health interventions even when the cost of the intervention is included.
>
> (iv) In addition, cost offset increases over time, largely because ... patients continue to decrease their overall use of the health care system, and don't require additional mental health services. (p. 2)

Medical cost offset effects are relatively obvious and easy to measure. Benefits, financial and otherwise, that accrue from the treatment of mental health and substance abuse/dependency problems also can come in forms that may not be so obvious. One area in which treatment can have a tremendous impact is that of the workplace. For example, note a few of the facts assembled by APA (1996).

(i) In 1985 behavioral health problems resulted in over $77 billion in lost income to Americans.

(ii) California's stress-related disability claims totaled $350 million in 1989.

(iii) In 1980, alcoholism resulted in over 500 million lost work days in this country.

(iv) Major depression cost an estimated $23 billion in lost work days in 1990. In addition, individuals with this disorder are three times more likely than nondepressed individuals to miss time from work and four times more likely to take disability days.

(v) Of all subjects from 58 psychotherapy effectiveness studies focusing on the treatment of depression, 77% received significantly better work evaluations than depressed subjects who did not receive treatment.

(vi) Treatment resulted in a 150% increase in earned income for alcoholics and a 390% increase in income for drug abusers in one study of 742 substance abusers.

In related findings, anxiety disorders accounted for one-third of America's $147 billion mental health bill in 1990 (*Mental Health Weekly,* 1996c). And on another front, the former director of the Office of the National Drug Control Policy reported that for every dollar spent on drug treatment, America saves seven dollars in health care and criminal justice costs (*Substance Abuse Funding News*, 1995).

Society's need for behavioral health care services provides an opportunity for trained providers of mental health services to become part of the solution to a major health care problem that shows no indication of decline. Each of the helping professions has the potential to make a particular contribution to this solution. Not the least of these contributions are those that can be made by clinical psychologists. As pointed out in an earlier volume (Maruish, 1994), the use of psychological tests in the assessment of the human condition is one of the hallmarks of clinical psychology. In fact, the training and acquired level of expertise in psychological testing distinguishes the clinical psychologist from other behavioral health care professionals probably more than anything else. Indeed, expertise in test-based psychological assessment can be said to be the particular and unique contribution that clinical psychologists make to the behavioral health care field.

For decades, clinical psychologists and other behavioral health care providers have come to rely on psychological assessment as a standard tool to be used with other sources of information for diagnostic and treatment planning purposes. However, changes that have taken place in the delivery of health care in general, and behavioral health care services in particular, during the past several years have led to changes in the way in which third-party payers and clinical psychologists themselves think about and/or use psychological assessment in day-to-day clinical practice. Some question the value of psychological assessment in the current time-limited, capitated service delivery arena where the focus has changed from clinical priorities to fiscal priorities (Sederer, Dickey, & Hermann, 1996). Others argue that it is in just such an arena that the benefits of psychological assessment can be most fully realized and contribute significantly to the delivery of cost-effective treatment for behavioral health disorders. Consequently, it could assist the health care industry in appropriately controlling or possibly reducing the utilization and cost of health care over the long term. It is this latter side of the argument that is supported by this author, and it provides the basis for this chapter.

In developing this chapter, the intent has been to provide students and practitioners of clinical psychology with an overview of how psychological assessment could and should be used in this era of managed behavioral health care. In doing so, this author discusses how psychological assessment is currently being used in the therapeutic environment and the many ways in which it might be used to the ultimate benefit of patients, providers, and payers.

As a final introductory note, it is important for the reader to understand that the term "psychological assessment," as it is used in this chapter, refers to the evaluation of a patient's mental health status using psychological tests or related instrumentation. This evaluation may be conducted with or without the benefit of patient or collateral interviews, review of medical or other records, and/or other sources of relevant information about the patient.

4.18.2 THE CURRENT PRACTICE OF PSYCHOLOGICAL ASSESSMENT IN THE THERAPEUTIC ENVIRONMENT

For a number of decades, psychological assessment has been viewed as a valued and integral part of the services offered by clinical psychologists. However, its popularity has not been without its ups and downs. Megargee and Spielberger (1992) have described a decrease in interest in assessment that began in the 1960s. This was due to a number of factors, including shifts in focus to those aspects of treatment for which assessment was thought to contribute little. Examples of these aspects included a growing emphasis on behavior modification techniques, the increasing use of psychotropic medications, and an emphasis in studying symptoms rather than personality syndromes and structures. Fortunately, Megargee and Spielberger also noted a number of factors that indicate a relatively recent resurgence in the interest in assessment, including a new realization of how psychological assessment can assist in interventions provided to mental health care patients.

But where does psychological assessment actually fit into the daily scope of activities for practicing psychologists? The results of two recent surveys provide inconsistent findings. The newsletter *Psychotherapy Finances* (1995) reported the results of a nationwide readership survey of 1700 mental health providers of various professions. In this survey, 67% of the participating psychologists reported that they provide psychological testing services. This represents about a 10% drop from a similar survey published in 1992 by the same publication. Also of interest in this survey is the percent of professional counselors (39%), marriage and family counselors (16%), psychiatrists (21%), and social workers (13%) offering these same services.

In a 1995 survey conducted by the APA's Committee for the Advancement of Professional Practice (Phelps, 1996), 14 000 practitioners responded to questions related to workplace settings, areas of practice concerns, and range of activities. Most of the respondents (40.7%) were practitioners whose primary work setting was an individual independent practice. Other general work settings, that is, government, medical, academic, group practice settings, were represented by fairly equal numbers of respondents from the remainder of the sample. The principal professional activity reported by the respondents was psychotherapy, with 43.9% of the sample acknowledging involvement in this service. Assessment was the second most prevalent activity, being reported by 14% of the sample.

Differences in the two samples utilized in the above surveys may account for the inconsistencies in their findings. Psychologists who are subscribers to *Psychotherapy Finances* may represent that subsample of the APA survey respondents who are more involved in the delivery of clinical services. Certainly the fact that only about 44% of the APA respondents offer psychotherapy services supports this hypothesis.

Regardless of the two sets of findings, psychological assessment does not appear to be utilized as much as in the past, and one does not have to look hard to determine at least one reason why. One of the major changes that has come about in the American health care system during the past several years has been the creation and proliferation of managed care organizations (MCOs). The most significant direct effects of managed care include reductions in the length and amount of service, reductions in accessibility to particular modalities (e.g., reduced number of outpatient visits per case), and profession-related changes in the types of services managed by behavioral health care providers (Oss, 1996). Overall, the impact of managed behavioral health care on the services offered by psychologists and other health care providers has been tremendous. In the APA survey reported above (Phelps, 1996), approximately 79% of the respondents reported that managed care had either a low, medium, or high negative impact on their work. How has managed care negatively impacted the use of psychological assessment? It is not clear from the results of this survey, but perhaps others can offer at least a partial explanation.

Ficken (1995) has provided some insight into how the advent of managed care has limited the reimbursement for (and therefore the use of) psychological assessment. In general, he sees the primary reason for this as being a financial one. In an era of capitated behavioral health care coverage, the amount of money available for behavioral health care treatment is limited. MCOs therefore require a demonstration that the amount of money spent for testing will result in a greater amount of treatment cost savings. In addition, Ficken notes that much of the information obtained from psychological assessment is not relevant to the treatment of patients within an MCO environment. Understandably, MCOs are reluctant to pay for the gathering of such information.

Werthman (1995) provides similar insights into this issue, noting that

> Managed care … has caused [psychologists] to revisit the medical necessity and efficacy of their testing practices. Currently, the emphasis is on the use of highly targeted and focused psychological and neuropsychological testing to sharply define the "problems" to be treated, the degree of impairment, the level of care to be provided and the treatment plan to be implemented.
>
> The high specificity and "problem-solving" approach of such testing reflects MCOs' commitment to effecting therapeutic change, as opposed to obtaining a descriptive narrative with scores. In this context, testing is perceived as a strong tool for assisting the primary provider in more accurately determining patient "impairments" and how to "repair" them. (p. 15)

In general, Werthman views psychological assessment as being no different from other forms of patient care, thus making it subject to the same scrutiny, demands for demonstrating medical necessity and/or utility, and consequent limitations imposed by MCOs on other covered services.

The foregoing representations of the current state of psychological assessment in behavioral health care delivery could be viewed as an omen of worse things to come. In this author's opinion, they are not. Rather, the limitations that are being imposed on psychological assessment and the demand for justification of its use in clinical practice represent part of the customers' dissatisfaction with the way things always have been done in the past. In general, this author views the tightening of the purse strings as a positive move for both behavioral health care and the profession of psychology. It is a wake-up call to those who have contributed to the health care crisis by either uncritically performing costly psychological assessments, being unaccountable to the payers and recipients of our services, and generally not performing our services in the most responsible, cost-effective and efficient way possible. It is telling us that we need to evaluate what we've done and the way we've done it, and to determine what is the best way to do it in the future. As such, it provides an opportunity for clinical psychologists to re-establish the valuable contributions they can make to improving the quality of behavioral health care delivery through their knowledge and skills in the area of psychological assessment.

In the sections that follow, this author will convey what he sees are the opportunities for psychological assessment in the behavioral health care arena, both in the present and the future, and the means of best achieving them. The views that are advanced are based on his knowledge of and experience in current psychological assessment practices as well as directions provided by the current literature. Some will probably disagree with the proposed approach, given their own experience and thinking on the matters discussed. However, it is hoped that even though in disagreement, the reader will be challenged to defend his or her position to themselves and as a result, feel more comfortable in their thinking about their approach to their psychological assessment practices.

4.18.3 PSYCHOLOGICAL ASSESSMENT AS A THERAPEUTIC ADJUNCT

The role of psychological assessment in the therapeutic environment traditionally has been quite limited. Those of us who did not receive our graduate clinical training within the past few years probably have been taught the value of psychological assessment only at the "front end" of treatment. We were instructed in the power and utility of psychological assessment as a means of assisting in the identification of symptoms and their severity, personality characteristics relevant to understanding the patient and his or her typical way of perceiving and interacting with the world, and other aspects of the individual (e.g., intelligence, vocational interests) that are important in arriving at a description of the patient at one particular point in time. Based on these data and information obtained from patient and collateral interviews, medical records and the individual's stated goals for treatment, a diagnostic impression was given and a treatment plan was probably formulated and placed in the patient's chart, hopefully to be reviewed at various points during the course of treatment. In some cases, the patient was assigned to another practitioner within the same organization or referred out, never to be contacted or seen again, much less be assessed again by the one who performed the original assessment.

Fortunately, during the past few years the usefulness of psychological assessment as more than just a tool to be used at the beginning of treatment has come to be recognized. Consequently, its utility has been extended beyond being a mere tool for describing an individual presenting themselves for treatment to being a means of facilitating the treatment and understanding of behavioral health care problems throughout the episode of care and beyond. Psychologists and others who employ it in their practices are now finding that psychological assessment can be used for a variety of purposes. Generally speaking, several psychological tests currently being marketed can be employed as tools for assisting in clinical decision-making, outcomes assessment and, more directly, as treatment techniques in and of themselves. Each of these uses can uniquely contribute incremental value to the therapeutic process.

4.18.3.1 Psychological Assessment for Clinical Decision-making

Traditionally, psychological assessment has been used to assist clinical psychologists and other behavioral health care clinicians in making important clinical decisions. The types of decision-making for which it has been used include those related to screening, treatment planning, and monitoring of treatment progress. Generally, screening may be undertaken to assist in either: (i) identifying the patient's need for a particular service, or (ii) determining the likelihood of the presence of a particular disorder or other behavioral/emotional/psychological problem. More often than not, a positive finding on screening leads to a more extensive evaluation of the patient in order to confirm with greater certainty the existence of the problem, or to further delineate the problem. The value of screening lies in the fact that it permits the clinicians to identify, quickly and economically, with a fairly high degree of confidence (depending on the particular instrumentation used), those who are and are not likely to need care or at least further evaluation.

In many instances, psychological assessment is performed in order to obtain information that is deemed useful in the development of a specific plan for treatment. Typically, it is the type of information that is not easily (if at all) accessible through other means or sources. It is information which, when combined with other information about the patient, aids in understanding the patient, identifying the most important problems and issues that need to be addressed, and formulating recommendations about the best means of addressing them.

Another way in which psychological assessment can play a role in clinical decision-making is in the area of treatment monitoring. Repeated assessment of the patient at regular intervals during the treatment process can provide the therapist with feedback regarding the progress which is being made in the therapeutic endeavor. Based on the findings, the therapist will be encouraged either to continue with the original therapeutic approach or, in the case of no change or exacerbation of the problem, to modify or abandon the approach in favor of an alternate one.

4.18.3.2 Psychological Assessment as a Treatment Technique

It is only recently that empirical studies and other articles addressing the therapeutic benefits that can be realized directly from discussing psychological assessment results with the patient have been published. Rather than just providing test feedback as directed by APA's *Ethical principles of psychologists* (APA, 1992), therapeutic use of assessment involves a presentation of assessment results (including assessment materials such as test protocols, profile forms, other assessment summary materials) directly to the patient; an elicitation of the patient's reactions to them; and an in-depth discussion of the meaning of the results in terms of patient-defined assessment goals. In essence, the assessment data can serve as a catalyst for the therapeutic encounter via the objective feedback that is provided to the patient, the patient self-assessment that is stimulated, and the opportunity for patient and therapist to arrive at mutually agreed upon therapeutic goals, based on impressionistic and objective data available to both parties.

4.18.3.3 Psychological Assessment for Outcomes Assessment

Currently, one of the most common reasons for conducting psychological assessment in the USA is to assess the outcomes of behavioral health care treatment. It is difficult to open a trade paper or health care newsletter or to attend a professional conference without being presented with a discussion on either how to "do outcomes" or what the results of a certain facility's outcomes study have revealed. The focus on outcomes assessment most probably can be traced to the "continuous quality improvement" (CQI) movement that was initially implemented in business and industrial settings. The impetus for the movement originally was a desire to produce quality products in the most efficient manner, resulting in increased revenues and decreased costs.

In the health care arena, outcomes assessment has multiple purposes, not the least of which is as a tool for marketing the organization's services. Related to this, those organizations vying for lucrative contracts from third-party payers to provide health care services to their covered lives frequently require outcomes data demonstrating the effectiveness of the services offered by the bidders. Equally important to those awarding contracts is how satisfied patients are with the provider's services. But probably the most important potential use of this data for provider organizations (although not always recognized as such) can be found in the knowledge it yields about what works and what doesn't. In this regard it can serve a program evaluation function. It is this knowledge that, if attended to and acted upon, can lead to improvement in the services the organization offers. When used in this manner, outcomes assessment can become an integral component of the organization's CQI initiative.

But more importantly for the individual patient, outcomes assessment provides a means of objectively measuring how much improvement he or she has made from the time of treatment initiation to the time of treatment termination. Feedback to this effect may serve to instill in the patient greater self-confidence and self-esteem, and/or a more realistic view of where he or she is (from a psychological standpoint) at that particular time in their life. Conversely, it may serve as an objective indicator to the patient of the need for continued treatment.

The purpose of the foregoing is to present a broad overview of psychological assessment as a multipurpose behavioral health care tool. Depending on the individual clinician or provider organization, it may be employed for one or more, or all, of the purposes just described. Knowing the various ways in which psychological assessment can be used in the service of therapeutic change should help the reader understand the more in-depth and detailed discussion about how these applications can facilitate or otherwise add value to the psychotherapeutic services offered by providers. This detailed discussion follows below.

Before beginning this discussion, however, it is important to briefly review the types of instrumentation most likely to be used in therapeutic psychological assessment, as well as the significant considerations and issues related to the selection and use of this instrumentation for the stated purposes. This should further facilitate the reader's understanding of the remainder of the chapter.

4.18.4 GENERAL CONSIDERATIONS FOR THE SELECTION AND USE OF PSYCHOLOGICAL TEST INSTRUMENTATION

New instrumentation for facilitating and evaluating behavioral health care treatment is released by the major test publishers annually. Thus, the availability of instrumentation for these purposes is not an issue. However, selection of the appropriate instrument(s) for one or more of the therapeutic purposes described above is a matter requiring careful consideration. Inattention to the instrument's intended use, its demonstrated psychometric characteristics, its limitations, and other aspects related to its practical application can result in misguided treatment and potentially harmful consequences for a patient.

Several types of instruments could be used for the general therapeutic assessment purposes described above. For example, neuropsychological instruments might be used to assess memory deficits that could impact the clinician's decision to perform further testing, the goals established for treatment, and the approach to treatment that is selected. Tests designed to provide estimates of level of intelligence might be used for the same purposes. It is beyond the scope of this chapter to address, even in the most general way, all of the types of tests, rating scales, and the instrumentation that might be employed in the therapeutic environment. Instead, the focus here will be on general classes of instrumentation that have the greatest applicability in the service of the therapeutic endeavor. To a limited extent, specific examples of such instruments will be presented. This will be followed by a discussion of criteria and considerations that will assist the clinician in selecting the best instrumentation for his or her intended purposes.

4.18.4.1 Types of Instrumentation for Therapeutic Assessment

The instrumentation required for any therapeutic application will depend on: (i) the general purpose(s) for which the assessment is being conducted, and (ii) the level of informational detail that is required for those purpose(s). Generally, one may classify the types of instrumentation that would serve the purpose(s) of the therapeutic assessment into one of four general categories. As mentioned above, other types of instrumentation are frequently used in clinical settings for therapeutic purposes. However, the present discussion will be limited to those more commonly used by a wide variety of clinical psychologists in their day-to-day practices.

4.18.4.1.1 Psychological/psychiatric symptom measures

Probably the most frequently used instrumentation for several therapeutic purposes are measures of psychopathological symptomatology. Besides the fact that these are the types of instruments on which the majority of the clinician's psychological assessment training has probably been focused, they were developed to assess the problems that typically prompt people to seek treatment.

There are several subtypes of these measures of psychological/psychiatric symptomatology. The first is the comprehensive multidimensional measure. This is typically a lengthy, multiscale instrument that measures and provides a graphical profile of the patient on several types of psychopathological symptom domains (e.g., anxiety, depression) or disorders (schizophrenia, antisocial personality). Also, summary indices sometimes are available to provide a more global picture of the individual with regard to his or her psychological status or level of distress. Probably the most widely used and/or recognized of these measures are the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1951) and its restandardized revision, the MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989), the Millon Clinical Multiaxial Inventory-III (MCMI-III; Millon, 1994), and the Personality Assessment Inventory (PAI; Morey, 1991).

Multiscale instruments of this type can serve a variety of purposes that facilitate therapeutic efforts. They may be used upon initial contact with the patient to screen for the need for service and, at the same time, yield information that is useful for treatment planning. Indeed, some such instruments (e.g., the MMPI-2) may make available supplementary, content-related, and/or special scales that are designed to assist the user in addressing specific treatment considerations (e.g., low motivation for treatment). Other multiscale instruments might be useful in identifying specific problems that may be unrelated to the patient's chief complaints (e.g., low self-esteem). They can also be administered at numerous times during the course of treatment to monitor the patient's progress toward achieving established goals and to assist in determining what adjustments (if any) must be made to the clinician's approach. In addition, use of the instrument in a pre- and post-treatment fashion provides information related to the outcomes of the treatment. Data obtained in this fashion can be analyzed with results from other patients to evaluate the effectiveness of an individual therapist as well as an organization.

Abbreviated multidimensional measures are quite similar to the comprehensive multidimensional measure in many respects. First, by definition, they contain multiple scales for measuring a variety of symptom domains and/or disorders. They also may allow for the derivation of an index of the patient's general level of psychopathology or distress. In addition, they may be used for screening, treatment planning and monitoring, and outcomes assessment purposes just like the comprehensive instruments. The distinguishing feature of the abbreviated instrument is its length. Again, by definition, these instruments are relatively short, and easy to administer and (usually) score. Their brevity does not allow for an in-depth assessment of the patient and his or her problems, but this is not what these instruments were designed to do.

Probably the most widely used of these brief instruments are Derogatis' family of symptom checklist instruments. These include the original Symptom Checklist-90 (SCL-90; Derogatis, Lipman, & Covi, 1973) and its revision, the

SCL-90-R (Derogatis, 1983). Both of these instruments contain a checklist of 90 psychological symptoms, most of which score on the instruments' nine symptom scales. For each of these instruments an even briefer version has been developed. The first is the Brief Symptom Inventory (BSI; Derogatis, 1992), which was derived from the SCL-90-R. In a health care environment that is cost-conscious and unwilling to make too many demands on patient time, this 53-item instrument is gaining popularity over its longer and more expensive 90-item parent instrument. Similarly, a brief form of the original SCL-90 has been developed. Titled the Symptom Assessment-45 Questionnaire (SA-45; Strategic Advantage, Inc., 1996), its development did not follow Derogatis' approach to the development of the BSI; instead, cluster analytic techniques were used to select five items each for assessing each of the nine symptom domains found on the three Derogatis checklists.

The major strength of the abbreviated multiscale instruments is their ability to broadly and very quickly survey several psychological symptom domains and/or disorders relative to the patient. Its value is most clearly evident in settings where both the time and dollars available for assessment services are quite limited. These instruments provide a lot of information quickly. Because of their brevity, they are much more likely to be completed by patients than their lengthier comprehensive counterparts. This last point is particularly important if one is interested in monitoring treatment or assessing outcomes, both of which require at least two or more assessments to obtain the desired information.

4.18.4.1.2 Measures of general health status and role functioning

During the past decade, there has been an increasing interest in the assessment of health status in health care delivery systems. Initially, this interest was shown mostly by those organizations and settings focusing primarily on the treatment of physical diseases and disorders. Within recent years, behavioral health care providers have recognized the value in assessing the patient's general level of health. It is important to recognize that the term "health" means more than just the absence of disease or debility; it also implies a state of well-being throughout the individual's physical, psychological, and social spheres of existence (World Health Organization [WHO], 1948). Dickey and Wagenaar (1996) point out how this view of health recognizes the importance of eliciting the patient's point of view in assessing health status. They also point to similar conclusions reached by Jahoda (1958) specific to the area of mental health. Here, an individual's self-assessment relative to how he or she feels they should be is an important component of "mental health."

Measures of health status and physical functioning can be classified into one of two groups: generic and condition-specific. Probably the most widely used and respected generic health status measures are the 36-item Medical Outcomes Study Short Form Health Scale (SF-36; Ware & Sherbourne, 1992; Ware, Snow, Kosinski, & Gandek, 1994) and the 39-item Health Status Questionnaire 2.0 (HSQ; Health Outcomes Institute, 1993; Radosevich, Wetzler, & Wilson, 1994). Aside from the minor variations in the scoring of one of the instruments' scales (i.e., Bodily Pain) and the HSQ's inclusion of three depression screening items, the two measures essentially are identical. Each assesses eight dimensions of health, four addressing mental health-related constructs and four addressing physical health-related constructs, that reflect the WHO concept of "health."

Role functioning has recently gained attention as an important variable to address in the course of assessing the impact of a physical or mental disorder on an individual. In devising a treatment plan and monitoring progress over time, it is important to know how the person's ability to work, perform daily tasks, or interact with others is affected by the disorder. The SF-36 and HSQ both address these issues with scales designed for this purpose.

Responding to concerns that even these relatively brief objective measures are too lengthy for regular administration in clinical and research settings, 12-item, abbreviated versions of each have been developed. The SF-12 (Ware, Kosinski, & Keller, 1995) was developed for use in large scale, population-based research where the monitoring of health status at a broad level is all that is required. Also, a 12-item version of the HSQ, the HSQ-12 (Radosevich & Pruitt, 1996), was developed for similar uses. Interestingly, given that the two abbreviated versions were derived from essentially the same instrument, there is only a 50% item overlap between the two shortened instruments. Both instruments are relatively new but the data supporting their use that has been gathered up to 1997 is promising.

Condition-specific health status and functioning measures have been utilized for a number of years. Most have been developed for use with physical rather than mental disorders, diseases, and conditions. However, condition-specific measures of mental health status and functioning are beginning to appear. A major source of this type of instrument is the Minnesota-based Health Outcomes Institute (HOI), a successor to the health care think tank InterStudy. In addition to the HSQ and the HSQ-12, HOI serves as the distributor/clearinghouse for the condition-specific "technology of patient experience (TyPE) specifications." The available TyPEs that would be most useful to clinical psychologists and other behavioral health care practitioners include those developed by a team of researchers at the University of Arkansas Medical Center for use with depressive, phobic, and alcohol and substance disorders. TyPEs for other specific psychological disorders are currently under development at the University of Arkansas for distribution through HOI.

4.18.4.1.3 Quality of life measures

In their brief summary of this area, Andrews, Peters, and Teesson (1994) indicate that most of the definitions of "quality of life" (QOL) describe a multidimensional construct encompassing physical, affective, cognitive, social, and economic domains. Objective measures of QOL focus on environmental resources required to meet one's needs and can be completed by someone other than the patient. The subjective measures of QOL assess the patient's satisfaction with the various aspects of his or her life and thus must be completed by the patient.

Andrews et al. (1994) draw other distinctions in the QOL arena. One has to do with the differences between QOL and health-related quality of life, or HRQL, and (similar to the case with health status measures) the other has to do with the distinction between generic and condition-specific measures of QOL. QOL measures differ from HRQL measures in that the former assess the whole "fabric of life," while the latter assess quality of life as it is affected by a disease or disorder, or by its treatment. Generic measures are designed to assess aspects of life that are generally relevant to most people; condition-specific measures are focused on aspects of the lives of particular disease/disorder populations. However, as Andrews et al. point out, generic and condition-specific QOL measures tend to overlap quite a bit.

4.18.4.1.4 Service satisfaction measures

With the exploding interest in assessing the outcomes of treatment for the patient, it is not surprising to see an accompanying interest in assessing the patient's and, in some instances, the patient's family's satisfaction with the services received. In fact, many professionals and organizations equate satisfaction with outcomes and frequently consider it the most important outcome. In a recent survey of 73 behavioral health care organizations, 71% of the respondents indicated that their outcomes studies included measures of patient satisfaction (Pallak, 1994).

Although some view service satisfaction as an outcome, it is this author's contention that it should not be classified as such. Rather, it should be considered a measure of the overall therapeutic process, encompassing the patient's (and at times, others') view of how the service was delivered, the capabilities and attentiveness of the service provider, the benefits of the service (if any), and any of a number of other selected aspects of the service he or she received. Patient satisfaction surveys don't answer the question "What was the result of the treatment rendered to the patient"; they do answer the question "How did the patient feel about the treatment he or she received?" Thus, they serve an important program evaluation/improvement function.

The number of questionnaires that are currently being used to measure patient satisfaction is countless. This reflects the attempts of individual health care organizations to develop customized measures that assess variables important to their particular needs, which in turn reflects a response to outside demands to "do something" to demonstrate the effectiveness of their services. Often, this "something" has not been evaluated to determine its basic psychometric properties. As a result, there exists numerous options that one may choose from, but very few that actually have demonstrated their validity and reliability as measures of service satisfaction.

Fortunately, there are a few instruments that have been investigated for their psychometric integrity. Probably the most widely used and researched patient satisfaction instrument designed for use in behavioral health care settings is the eight-item version of the Client Satisfaction Questionnaire (CSQ-8; Attkisson & Zwick, 1982; Nguyen, Attkisson, & Stenger, 1983). The CSQ-8 was derived from the original 31-item CSQ (Larsen, Attkisson, Hargreaves, & Nguyen, 1979), which also yielded two longer 18-item alternate forms, the CSQ-18A and CSQ-18B (LeVois, Nguyen, & Attkisson, 1981). The more recent work of Attkisson and his colleagues at the University of California at San Francisco is the Service Satisfaction Scale-30 (SSS-30; Greenfield & Attkisson, 1989), a 30-item multifactorial scale that yields information regarding different aspects of satisfaction with mental health service, such as perceived outcome and manner and skill of the clinician.

4.18.4.2 Guidelines for Instrument Selection

Regardless of the type of instrument one might consider using in the therapeutic environment, many clinical psychologists frequently must choose between many product offerings. But what are the general criteria for the selection of any instrument for psychological assessment? What should guide the clinician's selection of an instrument for a specific therapeutic purpose? As part of their training, clinical psychologists and professionals from related psychological specialties have been educated about the important psychometric properties that should be considered when determining the appropriateness of an instrument for its intended use. However, this is just one of several issues that should be taken into account in an evaluation of a specific instrument for a specific therapeutic use. The guidance that has been offered by experts with regard to instrument selection is worth noting here.

4.18.4.2.1 National Institute of Mental Health criteria

Probably the most thorough and clinically relevant guidelines for the selection of psychological assessment instruments comes from the National Institute of Mental Health (NIMH) supported work of Ciarlo, Brown, Edwards, Kiresuk, and Newman (1986). A synopsis of Newman and Ciarlo's (1994) updated summary of this NIMH work is presented here. Note that the criteria discussed below were originally developed for use in evaluating instruments for outcomes assessment purposes. However, most have relevance to the selection of instrumentation used for the other therapeutic assessment purposes described above. Exceptions and qualifications with regard to this issue will be noted when appropriate.

Newman and Ciarlo (1994) describe 11 criteria for the selection of outcomes assessment instruments, each of which can be grouped into one of five types of consideration. The first consideration is that of applicability. The issue here is the relevance of the instrument to the target population. The instrument should assess those problems, symptoms, characteristics, and so on, that are common to the group to whom the instrument will be administered. The more heterogeneous the population, the more chance that modifications will be required and that these will alter the standardization and psychometric integrity of the instrument. Another applicability issue to consider when the instrument is to be used for outcomes assessment purposes is its independence from the type of treatment to be offered to the population.

The second set of general considerations is that of methods and procedures (Newman & Ciarlo, 1994). Several selection criteria are related to this group. The first is that administration of the instrument is simple and easily taught. Generally, this is more of an issue with clinician-rating scales than self-report scales. In the case of rating scales, concrete examples, or objective referents, at each rating level should be provided to the user. Next, the instrument should allow input not only from the patient but also from other sources (e.g., the clinician, collaterals). The benefits of this include the opportunities to obtain a feel for the patient from many perspectives, to validate reported findings and observations, and to promote honesty in responding from all sources (given that all parties will know that others will also be providing input). The final methods and procedures criterion, though not necessarily as important for the instrument being used for screening or treatment planning purposes, is that the instrument provide information relevant to understanding how the treatment may have effected change in the individual.

Newman and Ciarlo's (1994) third set of considerations have to do with the psychometric strengths of the instruments. According to the NIMH panel of experts, outcomes measures should: (i) meet the minimum psychometric standards for reliability (including internal consistency, test–retest reliability, and as appropriate, interrater reliability) and validity (content, construct, and concurrent validity); (ii) be difficult to “fake bad” or “fake good”; and (iii) be free from response bias and not reactive or sensitive to factors unrelated to the constructs that are being measured (e.g., physical settings, behavior of the treatment staff). These criteria obviously also apply to other psychological instruments used for purposes other than outcomes assessment. However, for outcomes assessment purposes, the instrument also must be sensitive to change related to treatment.

The fourth group of considerations concerns the cost of the instruments. Newman and Ciarlo (1994) point out that the answer to the question of how much one should spend on assessment instrumentation and associated costs (e.g., staff time for administering, scoring, processing, and analyzing the data) will depend on how important the data gathered is to assuring a positive return on the functions they support. In the context of the NIMH undertaking, Newman and Ciarlo felt that the data obtained through treatment outcomes assessment would support screening/treatment planning, efforts in quality assurance and program evaluation, cost containment/utilization review activities, and revenue generation efforts. However, that may be considered the ideal. At this point, the number and nature of the purposes that would be supported by the obtained data will depend on the individual organization. The more purposes the data can serve, the less costly the instrumentation is likely to be, at least from a value standpoint. In terms of actual costs, Ciarlo et al. (1986) estimated that 0.5% of an organization's total budget would be an affordable amount for materials, staff training, data collection, and processing costs related to outcomes assessment. However, one should be mindful that the recommendation was made in 1986 and may not reflect changes in policies, requirements, and attitudes related to the use of psychological assessment instruments since that time.

The final set of considerations in instrument selection has to do with the utility of the instrument. Four criteria related to utility are posited by Newman and Ciarlo (1994). First, the scoring procedures and the manner in which the results are presented should be comprehensible to all with a stake in the treatment of the organization's patients. This would not only include the patient, his or her family, the organization's administrative staff and other treatment staff, but also third-party payers and (in the case of outcomes assessment or program evaluation) legislative and administrative policy makers. Related to this is the criterion that the results of the instrument be easily interpreted by those with a stake in them. Another utility-related criterion is that the instrument should be compatible with a number of clinical practices and theories that are employed in the behavioral health care arena. This should allow for a greater range of test applicability and greater acceptance by the various stakeholders in the patient's treatment.

Another important aspect of utility is that "the instrument support[s] the clinical processes of a service with minimal interference" (Newman & Ciarlo, 1994, p. 107). There are two issues here. The first has to do with whether the instrument can support the screening, planning, and/or monitoring activities in addition to the outcomes assessment activities. In other words, are multiple purposes served by the instrument's results? The second issue is one that has to do with the extent to which the organization's staff is burdened with the collection and processing of assessment data. How much will the assessment process interfere with the daily work flow of the organization's staff? Equally important is whether the benefits that accrue justify the cost of implementing an assessment program for whatever purpose(s).

4.18.4.2.2 Other criteria and considerations

Although the work of Ciarlo and his colleagues provides more extensive instrument selection guidelines than most, others who have addressed the issue have arrived at recommendations that serve to reinforce and/or complement those found in the NIMH document. For example, Gavin Andrews' work in Australia has led to significant contributions to the body of outcomes assessment knowledge. As part of this, Andrews et al. (1994) have identified six general "qualities of consumer outcome measures" that are generally in concordance with those from the NIMH study. First, the measure should meet the criterion of applicability. In other words,

> it should address dimensions which are important to the consumer (symptoms, disability, and consumer satisfaction) and useful for the clinician in formulating and conducting treatment, yet the measure should be one which can have its data aggregated in a meaningful way so that the requirements of management can be addressed. (p. 30)

Multidimensional instruments yielding a profile of scores on all dimensions of interest are viewed as a means of best serving the interests of all concerned.

Acceptability, that is, being both brief and user-friendly, is another desirable quality identified by Andrews et al (1994). Closely associated with this is the criterion of practicality. It might be viewed as a composite of those NIMH criteria related to matters of cost, ease of scoring and interpretation, and training in the use and interpretation of the measure. Again in agreement with the NIMH work, the final three criteria identified by Andrews et al. relate to reliability, validity, and sensitivity to change. With regard to reliability, Andrews et al. specify what they consider to be the minimum levels of acceptable internal consistency reliability (0.90 for long tests), interrater reliability (0.40), and construct and criterion validity (0.50). They also stress the importance of an instrument's face validity in helping to ensure cooperation from the patient, and of self-report instruments having multiple response options (rather then just "yes/no" options) for increasing sensitivity of an instrument to small but relevant changes in the patient's status over time.

In Ficken's (1995) discussion of the role of assessment in an MCO environment, he concludes that the difficulties clinicians are experiencing in demonstrating the utility of psychological assessment to payers lies in the fact the instruments and objectives of traditional psychological assessment are not in synch with the needs of MCOs. The solution to the problem appears simple:

> the underlying objectives of testing must be aligned with the values and processes of MCOs. In short, this means identifying decision points in managed care processes that could be improved with objective, standardized data. There are two avenues in which these can be pursued: through facilitation/objectification of clinical-decision processes and through outcome assessment. (p. 12)

In general, Ficken (1995) sees opportunities in areas that this author has previously identified as screening, treatment planning, and outcomes assessment, specifically in the areas of primary medical care and behavioral health care (see below). Requirements of instruments used for screening were noted to include:

(i) high levels of sensitivity and specificity to diagnostic criteria from the *Diagnostic and statistical manual of mental disorders* (4th ed., *DSM-IV*; American Psychiatric Association, 1994) or the most up-to-date version of the *International classification of diseases* (*ICD*);

(ii) a focus on hard-to-detect (in a single office visit) but treatable disorders that are associated with imminent harm to self or others, significant suffering, and a decrease in productivity;

(iii) an administration time of no more than 10 minutes; and

(iv) an administration protocol that easily integrates into the organization's work flow.

Cases testing "positive" on the screener would be administered one or more "second-tier" instrument(s) to establish severity and a specific diagnosis. Ficken feels that if they are to be accepted by MCOs, these second-tier instruments should meet the requirements of screeners and either specify or rule out a diagnosis.

According to Ficken (1995), successful outcomes assessment instruments also must possess certain qualities. Because the areas most important to assess for outcomes measurement purposes are symptom reduction, level of functioning, quality of life and patient satisfaction, the instrument should (i) focus on one of these areas, (ii) be brief, (iii) meet "traditional standards" for validity and reliability, and (iv) be sensitive to clinical change.

Based on the work of Vermillion and Pfeiffer (1993), Burlingame, Lambert, Reisinger, Neff, and Mosier (1995) recommended four criteria for the selection of outcomes measures. The first is acceptable "technical features," that is, validity and reliability. Specifically, these authors recommended that instruments have an internal consistency of at least 0.80, test–retest reliability of at least 0.70, and concurrent validity of at least 0.50. The second criterion is "practicality features." These include brevity, ease of administration and scoring, and simplicity of the reporting of results. Third, the instrumentation should be "suitable" for the patients that are seen within the setting. Thus, because of the nature of most presenting problems in mental health settings, it should assess symptomatology and psychosocial functioning. The fourth criterion is sensitivity to "meaningful" change over time, allowing for a differentiation of *symptomatic* change from interpersonal/social role *functional* change.

Schlosser (1995) proposed a rather nontraditional view of "outcomes assessment." In what he refers to as a "patient-centric" view, assessment information is gathered and used during the course of therapy to bring about change during therapy, not after therapy has ended. Essentially, this equates to what this author has referred to above (and discusses in more detail below) as treatment monitoring. In this model, Schlosser feels that this type of assessment requires "elements regarding very specific, theoretically derived, empirically validated areas of functioning" (Schlosser, 1995, p. 66). These would involve the use of both illness and well-being measures that assess the patient on emotional, mental/cognitive, physical, social, life direction, and life satisfaction dimensions.

Many of Schlosser's (1995) considerations for selection of such measures are not unique (i.e., having "acceptable' levels of reliability and validity, brief, low-cost, and sensitive). However, for the purposes described Schlosser also indicates that they should also: (i) have "paradigmatic sensibility" (i.e., key words have the same meaning across instruments); (ii) be designed for repeated administration for feedback or self-monitoring purposes; and (iii) provide actionable information.

In addition to some already mentioned criteria (acceptable validity, reliability, affordability, ease of administration, and ease of data entry and analysis), Sederer et al. (1996) discuss other considerations that warrant attention in selecting outcomes measures for specific situations. These include automation capabilities related to availability of software for data analysis and reporting, compatibility with the organization's existing information system, and the ability to enter data via an optical scanner. They also provide advice that should help guide the user in selecting the appropriate instrumentation:

> A plan should be developed that addresses the following questions: which patients will be included in the study? What outcomes will be most

> effected by the treatment? When will the outcomes be measured? Who is going to read (and use) the information provided by the outcomes study? The more specific the answer to these questions, the better the choice of outcome instrument. (p. 4)

This and other recommendations would appear equally applicable when selecting instruments for other therapeutic assessment purposes.

One final set of criteria should be considered in the light of the following section on screening. Screening for the likelihood of the presence of disorders or for the need for additional assessment requires considerations that do not necessarily apply to instruments when they are used for the other therapeutic assessment purposes addressed in this chapter. A major one here is a specific consideration relative to a screener's criterion validity. Although broadly encompassed by the construct of "validity" that was previously discussed, it demands particular attention when evaluating instruments for screening purposes. What is being referred to here is the instrument's classification accuracy or efficiency.

Classification efficiency is usually expressed in terms of the following statistics: *sensitivity*, that is, the proportion of those individuals with the characteristic of interest who are accurately identified as such; *specificity*, that is, the proportion of individuals not having the characteristic of interest who are accurately identified as such); *positive predictive power*, which is the proportion of a population identified by the instrument as having the characteristic who actually do have the characteristic, and *negative predictive power*, which is the proportion of a population identified by the instrument as not having the characteristic who actually do not have the characteristic. This information can provide the clinical psychologist and other evaluators with empirically based information that is useful in the type of decision-making requiring the selection of one of two choices. The questions answered are typically those of the "yes/no" type, such as "Is the patient depressed or not?" or "Does the patient have a psychological problem significant enough to require treatment?" The reader is referred to Baldessarini, Finkelstein, and Arana (1983) for a discussion of issues related to the use of these statistics.

In evaluating these statistics, one must consider a few very important issues. One is the degree to which the clinician is willing to accept false-positives or false-negatives. This will be a function of the importance of maximizing the correct identification of those with the particular characteristic of interest vs. the importance of maximizing the correct identification of those not having the characteristic vs. the importance of optimizing the identification of both groups. This in turn will be dependent on the cutoff score recommended by the developer of the instrument and/or the efficiency values that are available when other cutoff scores are applied. These and related issues are discussed more extensively in the next section.

4.18.5 PSYCHOLOGICAL ASSESSMENT AS A TOOL FOR SCREENING

One of the most significant ways in which psychological assessment can contribute to the development of an economic and efficient behavioral health care delivery system is by using it to screen potential patients for need for behavioral health care services, and/or to determine the likelihood that the problem being screened is a particular disorder of interest. Probably the most concise, informative treatment of the topic of the use of psychological tests in screening for behavioral health care disorders is provided by Derogatis and DellaPietra (1994). In this work, these authors turn to the Commission on Chronic Illness (1987) to provide a good working definition of health care screening in general, that being:

> the presumptive identification of unrecognized disease or defect by the application of tests, examinations or other procedures which can be applied rapidly to sort out apparently well persons who probably have a disease from those who probably do not. (Commission on Chronic Illness, 1987, p. 45)

Derogatis and DellaPietra (1994) further clarify the nature and the use of screening procedures, stating that:

> the screening process represents a relatively unrefined sieve that is designed to segregate the cohort under assessment into "positives" who presumptively have the condition, and "negatives" who are ostensibly free of the disorder. Screening is not a diagnostic procedure *per se*. Rather, it represents a preliminary filtering operation that identifies those individuals with the highest probability of having the disorder in question for subsequent specific diagnostic evaluation. Individuals found negative by the screening process are not evaluated further. (p. 23)

The most important aspect of any screening procedure is the efficiency with which it can provide information useful to clinical decision-making. In the area of clinical psychology, the most efficient and thoroughly investigated screening procedures involve the use of psychological assessment instruments. As implied by the foregoing, the power or utility of a psychological screener lies in its ability to determine, with a high level of probability, whether the respondent does or does not have a particular disorder or condition, or whether he or she is or is not a member of a group with clearly defined characteristics. In daily clinical practice, the most commonly used screeners are those designed specifically to identify some aspect of psychological functioning or disturbance or provide a broad overview of the respondent's point-in-time mental status. Examples of problem-specific screeners include the Beck Depression Inventory (BDI; Beck, Rush, Shaw, & Emery, 1979) and the State–Trait Anxiety Inventory (STAI; Spielberger, 1983). Examples of screeners for more generalized psychopathology or distress include the SA-45 and BSI.

4.18.5.1 Research-based Use of Psychological Screeners

The establishment of a system for screening for a particular disorder or condition involves determining what it is one wants to screen in or screen out, at what level of probability one feels comfortable in making that decision, and how many incorrect classifications or what percentage of errors one is willing to tolerate. Once it is decided what one wishes to screen for, one then must turn to the instrument's classification efficiency statistics, that is, sensitivity, specificity, positive predictive power (PPP), and negative predictive power (NPP), for the information necessary to determine whether a given instrument is suitable for the intended purpose(s).

Recall that sensitivity refers to the proportion of those with the characteristic of interest who are accurately identified as such by an instrument or procedure, while specificity refers to the proportion of those not having the characteristic of interest who are accurately identified. The cutoff score, index value, or other criterion used for classification can be adjusted to maximize either sensitivity or specificity. However, maximization of one will necessarily result in a decrease in the other, thus increasing the percentage of false-positives (with maximized sensitivity) or false-negatives (with maximized specificity). Stated differently, false-positives will increase as specificity decreases, while false-negatives will increase as sensitivity decreases (Elwood, 1993).

Another approach is to optimize both sensitivity and specificity, thus yielding a fairly even balance of true positives and true negatives. Although optimization might seem to be the preferable approach in all instances, there are situations in which a maximization approach is more desirable. For example, a psychiatric hospital with an inordinately high rate of inpatient suicide attempts begins to employ a screener designed to help identify patients with suicide potential as part of its admission procedures. The hospital adjusts the classification cutoff score to a level that identifies all suicidal patients in the screener's normative group. This cutoff score is then applied to all patients being admitted to the hospital for the purpose of identifying those requiring an extensive evaluation for suicide potential. This not only increases the number of true positives, but it also decreases the specificity and increases the number of false positives. However, the trade-off of identifying more suicidal patients early on with having more nonsuicidal patients receiving suicide evaluations would appear worthwhile for the hospital's purposes. Similarly, in other instances, maximization of specificity may be the preferred approach. For example, an MCO might wish to use a measure of overall level of psychological distress to identify those covered lives that are not in need of behavioral health care services. Sensitivity will decrease but, for the MCO's purposes, this might be quite acceptable.

Hsiao, Bartko, and Potter (1989) note that "a diagnostic test will not have a unique sensitivity and specificity. Instead, for each diagnostic test, the relationship between sensitivity and specificity depends on the cutoff point chosen for the test" (p. 665). The effect of employing individual classification cutoff points can be presented via the use of receiver operating characteristic (ROC) curves. These curves are nothing more than a plotting of the resulting true positive rate (sensitivity) against the false positive rate for each cutoff score that might be employed with a test used for classification purposes. The plotting allows for a graphical representation of what may be gained and/or lost by shifting cutoff scores. The resulting area underneath the curve provides an indication of how well the test performs. Development of ROC curves from available data for a test being considered for screening purposes is recommended. The reader is referred to Hsiao et al. and Metz (1978) for a more detailed discussion of ROC curves and their use.

In day-to-day clinical work, an instrument's PPP and NPP can provide information that is more useful than sensitivity and specificity. As Elwood (1993) has pointed out,

> Although sensitivity and specificity do provide important information about the overall performance of a test, their limitation in classifying

> individual subjects becomes evident when they are considered in terms of conditional probabilities. Sensitivity is $P\ (+/d)$, the probability (P) of a positive test result (+) given that the subject has the target disorder (d). However, the task of the clinicians in assessing individual patients is just the opposite: determining $P\ (d/+)$, the probability that a patient has the disorder given that he or she obtained an abnormal test score. In the same way, specificity expresses $P\ (-/-d)$, the probability that a patient will have a negative test result given that he or she does not have the disorder. Here again, the task confronting the clinician is usually just the opposite: determining $P\ (-d/-)$, the probability that the patient does not have the disorder given a negative test result. (p. 410)

A note of caution is warranted when evaluating the two predictive powers of a test. Unlike sensitivity and specificity, both PPP and NPP are affected and change according to the prevalence or base rate at which the condition or characteristic of interest (i.e., that which is being screened by the test) occurs within a given setting. As Elwood (1993) reports, the lowering of base rates results in lower PPPs while increasing base rates result in higher PPPs. The opposite trend is true for NPPs. He notes that this is an important consideration because clinical tests are frequently validated using samples in which the prevalence rate is 0.50, or 50%. Thus, it is not surprising to see a test's PPP drop in "real-life" applications where the prevalence is lower.

Derogatis and DellaPietra (1994) indicate that a procedure referred to as "sequential screening" may provide at least a partial solution to the limitations or other problems that low base rates may pose for the predictive powers of an instrument. Sequential screening essentially involves the administration of two screeners, each of which measures the condition of interest, and two-phase screening. In the first phase, one screener is administered to the low base rate population. The purpose of this is to identify those individuals without the condition, thus requiring relatively good specificity. These individuals are eliminated from involvement in the second phase, resulting in an increase in the prevalence of the condition among those who remain. This group is then administered another screener of equal or better sensitivity. With the increased prevalence of the condition in the remaining group, the false positive rate will be much lower. As Derogatis and DellaPietra point out,

> Sequential screening essentially zeros in on a high-risk subgroup of the population of interest by virtue of a series of consecutive sieves. These have the effect of eliminating from consideration individuals with low likelihood of having the disorder, and simultaneously raising the base rate of the condition in the remaining sample. (p. 45)

In summary, PPP and NPP can provide information that is quite valuable to those making important clinical decisions, such as determining need for behavioral health care services, assigning diagnoses, or determining appropriate level of care. However, these users must be cognizant of the manner in which the predictive powers may change with the population to which the test or procedure is applied.

4.18.5.2 Implementation of Screeners into the Daily Work Flow of Service Delivery

The utility of a screening instrument is only as good as the degree to which it can be integrated into an organization's daily regimen of service delivery. This, in turn, depends on a number of factors. The first is the degree to which the administration and scoring of the screener is quick and easy, and the amount of time required to train the provider's staff to successfully incorporate the screener into their day-to-day activities.

The second factor relates to its use. Here, the screener is not used for anything other than determining the likelihood that the patient does or does not have the specific condition or characteristic the instrument is designed to assess. Use for any other purpose (e.g., assigning a diagnosis based solely on screener results, determining the likelihood of the presence of other characteristics) only serves to undermine the integrity of the instrument in the eyes of staff, payers, and other parties with a vested interest in the screening process.

The third factor has to do with the ability of the provider to act on the information. It must be clear how the clinician should proceed based on the information available.

The final factor is staff acceptance and commitment to the screening process. This comes only with a clear understanding of the importance of the screening, the usefulness of the obtained information, and how the screening process is to be incorporated into the organization's business flow.

Ficken (1995) provides an example of how screeners can be integrated into an assessment system designed to assist primary care physicians to identify patients with psychiatric disorders. This system (which also allows for the incorporation of practice guidelines) seems to take into account the first three utility-related factors listed above. It begins with the administration of a screener that is highly sensitive and specific to *DSM*- or *ICD*-related disorders. Ficken indicates that screeners should require no more than 10 minutes to complete, and that "their administration must be integrated seamlessly into the standard clinical routine" (p. 13). Somewhat similarly to the sequence described by Derogatis and DellaPietra (1994), positive findings would lead to a second level of testing. Here, another screener that meets the same requirements as those for the first screener and also affirms or rules out a diagnosis would be administered. Positive findings would lead to additional assessment for treatment planning purposes. Consistent with standard practice, Ficken recommends confirmation of screener findings by a qualified psychologist or physician.

4.18.6 PSYCHOLOGICAL ASSESSMENT AS A TOOL FOR TREATMENT PLANNING

The administration of screeners is only one way in which psychological assessment can serve as a valuable tool for treatment planning. However, many would argue that it is the most limited way in which this tool can be used for planning a course of treatment. When employed by a trained clinician, psychological assessment can provide information that can greatly facilitate and enhance the planning of the therapeutic intervention for the individual patient.

The importance of treatment planning has received significant attention during recent years. The reasons for this were summarized previously by this author (Maruish, 1990) as follows:

> Among important and interrelated reasons ... [are] concerted efforts to make psychotherapy more efficient and cost effective, the growing influence of "third parties" (insurance companies and the federal government) that are called upon to foot the bill for psychological as well as medical treatments, and society's disenchantment with open-ended forms of psychotherapy without clearly defined goals. (p. iii)

The role that psychological assessment can play in planning a course of treatment for behavioral health care problems is significant. Butcher (1990) indicated that information available from instruments such as the MMPI-2 can not only assist in identifying problems (see above) and establishing communication with the patient (see below), it can also help ensure that the plan for treatment is consistent with the patient's personality and external resources. In addition, psychological assessment may reveal potential obstacles to therapy, areas of potential growth, and problems of which the patient may not be consciously aware. Moreover, both Butcher and Appelbaum (1990) viewed testing as a means of quickly obtaining a second opinion. Other benefits of the results of psychological assessment, identified by Appelbaum, include assistance in identifying patient strengths and weaknesses, identification of the complexity of the patient's personality, and establishment of a reference point or guide to refer to during the therapeutic episode.

The types of information that can be derived from patient assessment and the manner in which it is applied for this purpose are quite varied—a fact that will become evident below. Nevertheless, Strupp (see Butcher, 1990) probably provided the best summary of the potential contribution of psychological assessment to treatment planning, stating that "careful assessment of patient's personality resources and liabilities is of inestimable importance. It will predictably save money and avoid misplaced therapeutic effort; it can also enhance the likelihood of favorable treatment outcomes for suitable patients" (pp. v–vi).

4.18.6.1 Assumptions About Treatment Planning

The introduction to this section presented a broad overview of ways in which psychological assessment can assist in devising and successfully implementing plans of treatment for behavioral health care patients. These and other benefits will be discussed in greater detail below. However, it is important to first clarify what treatment planning is and some of the general, implicit assumptions that one typically can make about this important therapeutic activity.

For the purpose of this discussion, the term "treatment planning" indicates that part of a therapeutic episode in which the treatment provider develops a set of goals for an individual presenting with behavioral health care problems, and outlines the specific means by which he/she or other resources will assist the patient in achieving those goals in the most efficient manner. General assumptions underlying the treatment planning process are as follows.

(i) The patient is experiencing behavioral health problems that have been identified either by themself or by another party. Common external sources of problem identification include the patient's spouse, parent, teacher, employer, and the legal system.

(ii) The patient experiences some degree of internal and/or external motivation to eliminate or reduce the identified problems. An example of external motivation to change is the potential loss of job or marriage if problems are not resolved.

(iii) The goals of treatment are tied either directly or indirectly to the identified problems.

(iv) The goals of treatment have definable criteria for achievement, are indeed achievable by the patient, and are developed in collaboration with the patient.

(v) The prioritization of goals is reflected in the treatment plan.

(vi) The patient's progress toward achievement of the treatment goals can be tracked and compared against an expected path of improvement in either a formal or informal manner. This expected path of improvement may be based on the clinician's experience or (ideally) on objective data gathered on patients similar to the patient.

(vii) Deviations from the expected path of improvement will lead to a modification in the treatment plan, followed by subsequent monitoring to determine the effectiveness of the alteration.

These assumptions should not be considered exhaustive, nor are they reflective of what actually occurs in all situations. For example, some patients seen for therapeutic services may have no motivation to change. As may be seen in juvenile detention settings or in cases where children are brought to treatment by the parents, their participation in treatment is forced, and they may engage in intentional efforts to sabotage any therapeutic intervention. Also, it is likely that there are still clinicians who identify and prioritize treatment goals without the direct input of the patient. Nevertheless, the assumptions above represent this author's view of the aspects of treatment planning that have a direct bearing on the manner in which psychological assessment can best serve treatment planning efforts.

4.18.6.2 The Benefits of Psychological Assessment for Treatment Planning

As has already been touched upon, there are several ways in which psychological assessment can assist in the planning of treatment for behavioral health care patients. Following is a discussion of the more common and evident contributions that assessment can make to treatment planning efforts. These can be organized into four general categories: problem identification, problem clarification, identification of important patient characteristics, and monitoring of treatment progress.

4.18.6.2.1 Problem identification

Probably the most common use of psychological assessment in the service of treatment planning is for the purpose of problem identification. Often, the use of psychological testing *per se* is not needed to identify what problems the patient is experiencing. He or she will either tell the clinician directly without questioning, or they will readily admit to their problem(s) during the course of a clinical interview. However, this is not always the case.

The value of psychological testing becomes apparent in those cases where the patient is hesitant or unable to identify the nature of his or her problems. However, with a motivated and engaged patient who responds to items on a well validated and reliable test in an open and honest manner, the process of identifying what brought the patient to treatment also may be greatly facilitated. Cooperation shown during testing may be attributable to the nonthreatening nature of responding to questions presented on paper or a computer monitor (as opposed to those posed by another human being); the subtle, indirect, or otherwise nonthreatening nature of the questions (compared to those asked by the clinician); instrumentation that "casts a wider net" than the clinician in his or her interview with the patient; or any combination of these reasons.

In addition, the nature of some of the more commonly used psychological test instruments allows for the identification of secondary problems of significant severity that might otherwise be overlooked. Multidimensional inventories such as the MMPI-2 and the PAI are good examples of these types of instruments. Moreover, these instruments may be sensitive to other problems or patient traits or characteristics that may not necessarily be problems but which may exacerbate or otherwise contribute to the maintenance of the patient's problems.

Note that the type of problem identification described here is different from that conducted during screening (see above). Whereas screening is focused on determining the presence or absence of a single problem, problem identification generally takes a broader view and investigates the possibility of the presence of multiple problem areas. At the same time, there is also an attempt to determine the extent to which the problem area(s) affect the patient's ability to function.

4.18.6.2.2 Problem clarification

Psychological testing can often assist in the clarification of a known problem. Through tests designed for use with individuals presenting problems similar to the patient's, aspects of identified problems can be elucidated. This will improve the patient's and clinician's understanding of the problem and likely lead to a better treatment plan. The three most important types of information that can be gleaned for this purpose are the severity of the problems, the complexity of the problems, and the degree to which the problems impair the patient's ability to function in one or more life roles.

The manner in which a patient is treated depends a great deal on the severity of his or her problem. In particular, severity has a great bearing on the setting in which the behavioral health care intervention is provided. Those whose problems are so severe that they are considered a danger to themselves or others more often than not are best suited for inpatient treatment, at least until dangerousness is no longer an issue. Similarly, problem severity may be a primary criterion for an evaluation for a medication adjunct to treatment. Severity also may have a bearing on the type of psychotherapeutic approach that is taken by the clinician. For example, it may be more productive for the clinician to take a supportive role with severe cases; all things being equal, a more confrontational approach may be more appropriate with patients with problems in the mild to moderate range of severity.

As alluded to above, the problems of patients seeking behavioral health care services are frequently multidimensional. Patient and environmental factors that play into the formation and maintenance of a problem, along with the latter's relationship with other problems, all contribute to its complexity. Knowing the complexity of the target problems is invaluable in devising an effective treatment plan. Again, multidimensional instruments or batteries of tests measuring specific aspects of psychological dysfunction serve this purpose well.

As with problem severity, knowledge of the complexity of a patient's psychological problems can help the clinician and patient in many aspects of treatment planning, including determination of appropriate setting, therapeutic approach, need for medication, and other matters on which important decisions must be made. However, possibly of equal importance and concern to the patient and outside parties (spouse, employer, school, etc.) is the extent to which these problems affect the patient's ability to function in his or her role as parent, child, employee, student, friend, and so on. Data gathered from the administration of measures of role functioning can provide information that not only clarifies the impact of the patient's problems and serves to establish role-specific goals, but also identifies other parties that may serve as potential allies in the therapeutic process. In general, the most important role-functioning domains for assessment would be those related to work or school performance, interpersonal relationships, and activities of daily living (ADLs).

4.18.6.2.3 Identification of important patient characteristics

The identification and clarification of the patient's problems is of key importance in planning a course of treatment for the patient. However, there are numerous types of nonproblem-oriented patient information that can be useful in planning treatment and can be rather easily identified through the use of psychological assessment instruments. The vast majority of treatment plans are developed or modified with consideration of at least some of these other patient characteristics. The exceptions mostly are found with clinicians or programs that take a "one size fits all" approach to the treatment of general or specific types of disorders. It is beyond the scope of this chapter to provide an exhaustive list of what other types of information may be available to the clinician. However, a few are particularly worth mentioning.

Probably the most useful type of nonproblem-oriented information that can be gleaned from psychological assessment results is the identification of the patient characteristics or conditions that can serve as assets or areas of strength for the patient in working toward achieving the therapeutic goals. For example, Morey and Henry (1994) point to the utility of the PAI's Nonsupport scale in identifying whether the patient perceives an adequate social support network, this being a predictor of positive therapeutic progress. Other examples include "normal" personality characteristic information, such as that which can be obtained from Gough, McClosky, and Meehl's Dominance and Social Responsibility scales (1951, 1952) developed for use with the MMPI/MMPI-2. Greene (1991) indicates that those with high scores on the Dominance scale are described as "being able to take charge of responsibility for their lives. They are poised, self-assured, and confident of their own abilities" (p. 209). Gough and his colleagues interpreted high scores on the Social Responsibility scale as being indicative of individuals who, among other things, trust the world, are self-assured and poised, and stress the need for one to carry his or her share of duties. Thus, scores on these scales may reveal some important aspects of patient functioning that can be used in the service of affecting therapeutic change.

Similarly, knowledge of the patient's weaknesses or deficits may impact the type of treatment plan that is devised. Greene and Clopton (1994) provided numerous types of deficit-relevant information from the MMPI-2 Content Scales that have implications for treatment planning. For example, a clinically significant score ($T > 64$) on the Anger scale should lead one to consider the inclusion of training in assertiveness and/or anger control as part of the patient's treatment. On the other hand, uneasiness in social situations, as suggested by a significantly elevated score on either the Low Self-Esteem or Social Discomfort scale, suggests that a supportive approach to the intervention would be beneficial, at least initially.

Moreover, use of specially designed scales and procedures can provide information related to the patient's ability to become engaged in the therapeutic process. For example, the MMPI-2 Negative Treatment Indicators content scale developed by Butcher and his colleagues (Butcher, Graham, Williams, & Ben-Porath, 1989) may be useful in determining whether the patient is likely to be resistant to any form of "talk" therapy. Morey and Henry (1994) have supplied algorithms utilizing T scores for various PAI scales to make statements about the presence of positive characteristics, such as the presence of sufficient distress to motivate engagement in treatment, the ability to form a therapeutic alliance, and the capacity to utilize psychotherapy. The Therapeutic Reactance Scale (Dowd, Milne, & Wise, 1991) is yet another example of an instrument from which the clinician can be forewarned of potential resistance to therapeutic intervention.

Other types of patient characteristics that can be identified through psychological assessment have implications for the choice of the therapeutic approach and thus can contribute significantly to the treatment planning process. Beutler and his colleagues (Beutler & Clarkin, 1990; Beutler, Wakefield, & Williams, 1994; Beutler & Williams, 1995) have identified four patient characteristics that are thought to be important to matching patients and treatment approach for maximized therapeutic effectiveness. These include symptom severity, symptom complexity, coping style, and potential resistance to treatment. At different points in time, other patient variables also have been identified by these investigators as important considerations in the selection of the best treatment for a given patient. These include the problem-solving phase the patient has reached (Beutler & Clarkin, 1990), and subjective distress and social support (L.E. Beutler, personal communication, January 15, 1996).

Moreland (1996) points out how psychological assessment can assist in determining whether the patient deals with problems through internalizing or externalizing behaviors. All things being equal, internalizers would probably profit most from an insight-oriented approach rather than a behaviorally oriented approach. The reverse would be true for externalizers. In addition, cognitive factors also are important. Knowing that intelligence test results indicate an average or above IQ can assist the clinician in determining whether a patient will be able to benefit from a cognitive approach.

4.18.6.2.4 Monitoring of progress along the path of expected improvement

Information from repeated testing during the treatment process can help the clinician to determine if the treatment plan is appropriate for the patient at that particular point in time. Thus, many clinicians use psychological assessment to determine whether their patients are showing the expected improvement as treatment progresses. If not, adjustments can be made. These adjustments may reflect the need for a more intensive or aggressive treatment approach (e.g., increased number of psychotherapeutic sessions each week, addition of a medication adjunct) or for a less intensive approach (e.g., reduce or terminate medication, transfer from inpatient to outpatient care). Either way, this may require further retesting in order to determine whether the treatment revisions have impacted the course of change in the expected direction. This process may be repeated any number of times. In-treatment retestings also can provide information relevant to the decision of when to terminate treatment.

The goal of monitoring is to determine whether treatment is "on track" with the progress that is expected at a given point in time. When and how often one might assess the patient is dependent on a few factors. The first is the instrumentation. Many instruments are designed to assess the patient's status at the time of testing. Items on these measures are generally worded in the present tense (e.g., "I feel tense and nervous," "I feel that my family loves and cares about me"). Changes from one day to the next on the constructs measured by the instrument should be reflected in the test results.

Other instruments, however, ask the patient to indicate if a variable of interest has been present, or how much or to what extent it has occurred during a specific time period in the past. The items usually are asked in the context of something like "During the past month, how often have you . . ." or "During the past week, to what extent has ...” Readministration of a measure containing interval-of-time-specific items or subsets of items should be undertaken only after a period of time equivalent to or longer than the time interval to be considered in responding to the items has past. For example, an instrument which asks the patient to consider how much certain symptoms have been problematic during the past seven days should not be readministered for at least seven days. The responses elicited during a readministration that occurs less than seven days after the first administration would include the patient’s consideration of his or her status during the previously considered time period. This may make interpretation of the change of symptom status (if any) from the first to the second assessment difficult if not impossible.

Methods to determine whether clinically significant change has occurred from one point in time to another have been developed and can be used for treatment monitoring purposes. These are discussed in the outcomes assessment section of this chapter below. However, for monitoring purposes, another approach to evaluating therapeutic change may be superior. This approach may be referred to as the “glide path” approach, with the term referring to the narrow descent course or path that airplanes must follow when landing. Deviation from the flight glide path requires corrections in the plane’s speed, altitude, and/or attitude in order to return to the glide path and a safe landing.

R.L. Kane (personal communication, July 22, 1996) has indicated that just as a pilot has the instrumentation to alert him or her about the plane’s position on the glide path, the clinician may use psychological assessment instruments to track how well the patient is following the glide path of treatment. The glide path in this case represents expected improvement over time in one or more measurable areas of functioning (e.g., symptom severity, social role functioning, occupational performance). The expectations would be based on objective data obtained from similar patients at various points during their treatment and would allow for minor deviations from the path. The end of the glide path is one or more specific goals that are part of the treatment plan. Thus, “arrival” at the end of the glide path signifies the attainment of specific treatment goals.

4.18.7 PSYCHOLOGICAL ASSESSMENT AS A THERAPEUTIC INTERVENTION

The use of psychological assessment as a means of therapeutic intervention in and of itself has received more than passing attention during the past few years. “Therapeutic assessment” with the MMPI-2 has received particular attention primarily through the work of Finn and his associates (Finn, 1996a, 1996b; Finn & Martin, in press; Finn & Tonsager, 1992). Finn’s approach appears to be applicable with instruments or batteries of instruments that provide multidimensional information relevant to the concerns of patients seeking answers to questions related to their mental health status. The approach espoused by Finn thus will be presented here as a model for deriving direct therapeutic benefits from the psychological assessment experience.

4.18.7.1 What Is Therapeutic Assessment?

In discussing the use of the MMPI-2 as a therapeutic intervention, Finn (1996a) describes an assessment procedure whose goal is to “gather accurate information about clients ... and then use this information to help clients understand themselves and make positive changes in their lives” (p. 3). Elaborating on this procedure and extending it to the use of any test, Finn and Martin (in press) describe therapeutic assessment as

> collaborative, interpersonal, focused, time limited, and flexible. It is ... very interactive and requires the greatest of clinical skills in a challenging role for the clinician. It is unsurpassed in a respectfulness for clients: collaborating with them to address *their* concerns (around which the work revolves), acknowledging them as experts on themselves and recognizing their contributions as essential, and providing to them usable answers to their questions in a therapeutic manner.

The ultimate goal of therapeutic assessment is to provide an experience for the client that will allow him/her to take steps toward greater psychological health and a more fulfilling life. This is done by recognizing the client’s characteristic ways of being, understanding in a meaningful, idiographic way the problems the client faces, providing a safe environment for the client to explore change, and providing the opportunity for the client to experience new ways of being in a supportive environment.

Simply stated, therapeutic assessment may be considered an approach to the assessment of mental health patients in which the patient is not only the primary provider of information needed to answer questions, but also is actively involved in formulating the questions that are to be answered by the assessment. Feedback regarding the results of the assessment is provided to the patient and is considered a primary, and possibly the principal element of the assessment process. Thus, the patient becomes a partner in the assessment process; as a result, therapeutic and other benefits accrue.

The reader should note that in this section, the term "therapeutic assessment" is used to denote the specific approach advocated by Finn and his colleagues for using the psychological assessment process as an opportunity for therapeutic intervention. It should not be confused with the more general term "therapeutic psychological assessment" as it has been employed throughout this chapter; "therapeutic assessment" is but one aspect of therapeutic psychological assessment.

4.18.7.2 The Impetus for Therapeutic Assessment

To say that clinical psychologists performing psychological assessments in mental health settings traditionally have never shared much of their findings with their patients is probably not an overstatement. A common scenario throughout many mental health settings was that a patient being treated by a psychologist was evaluated by the latter, or a patient was referred to the psychologist by another mental health professional for assessment only. In the first instance, the degree to which the psychologist might directly share the results of the often lengthy and expensive evaluation would vary. Generally, a detailed review of the findings would be a rarity. In the latter instance, the patient would be evaluated, a report of the results dictated, and a copy of the report sent back to the referring clinician. In either instance, the purpose of the assessment probably would be to answer questions posed by the treating clinician. Unfortunately, the patient and his or her concerns as they related to the psychological assessment were typically of only secondary consideration, if any.

Fortunately, recent occurrences have begun to change the way in which assessment information is used. Consequently, the degree to which the patient is involved in the assessment process is changing. One reason for this is the relatively recent revision of the ethical standards of the APA (1992). This revision included a mandate for psychologists to provide feedback to clients whom they test. According to ethical standard 2.09:

> Unless the nature of the relationship is clearly explained to the person being assessed in advance and precludes provision of an explanation of results (such as in some organizational consulting, preemployment or security screenings, and forensic evaluations), psychologists ensure that an explanation of the results is provided using language that is reasonably understandable to the person assessed or to another legally authorized person on behalf of the client. Regardless of whether the scoring and interpretation are done by the psychologist, by assistants, or by automated or other outside services, psychologists take reasonable steps to ensure that appropriate explanations of results are given. (p. 8)

Many clinicians and other psychologists involved in assessment activities (e.g., counseling psychologists, neuropsychologists) have had to modify their practice routine to accommodate this requirement. Some view this requirement as resulting in an improvement in the quality of their services; others likely see it as nothing more than an inconvenience which, in the era of managed care and limited access to treatment, further limits the amount of time they have to work with a patient. However, most would agree that the patient has benefited from the required feedback.

Finn and Tonsager (1992) identified other factors that may have contributed to the recent interest in providing patients with assessment feedback. One is another external influence, that is, the recognition of the patient's right to see their medical and psychiatric health care records. However, they also point to several clinically and research-based findings and impressions that suggest that therapeutic assessment enhances patient care through the facilitation of patient–therapist rapport, cooperation during the assessment process, positive feelings about the process and the clinician, improvement in mental health status, and feelings of being understood by another. In addition, Finn and Tonsager refer to Finn and Butcher's (1991) summary of potential benefits that may accrue from providing test results feedback. The listed benefits, based on clinical experience, include increased feelings of self-esteem and hope, reduced symptomatology and feelings of isolation, increased understanding and self-awareness, and increased motivation to seek or be more actively involved in mental health treatment. Finally, Finn and Martin (in press) note that the therapeutic assessment process can lead to increased feelings of mastery and control and decreased feelings of alienation. At the same time, it can serve as a model for relationships that can result in mutual respect and the patient being seen for who he or she is.

4.18.7.3 The Therapeutic Assessment Process

Finn (1996a) has outlined a three-step procedure for therapeutic assessment using the MMPI-2. As indicated above, it should work equally well with other multidimensional instruments that one might select. Finn describes this procedure as one to be used in those situations in which the patient is seen only for assessment (i.e., the patient is not to be treated later by the assessing clinician). From the present author's standpoint, the procedures are equally applicable for use by clinicians who test patients whom they later treat. With these points in mind, the three-step procedure is summarized below.

4.18.7.3.1 Step 1: The initial interview

According to Finn (1996a), the initial interview with the patient serves multiple purposes. It provides an opportunity to build rapport, or to increase rapport if a patient–therapist relationship already exists. The assessment task is presented as a collaborative one, and the patient is given the opportunity to identify questions that he or she would like answered using the assessment data. Background information related to the patient-identified questions is subsequently gathered. Any reservations about participating in the therapeutic assessment process (e.g., confidentiality, previous negative experiences with assessment) are dealt with in order to facilitate maximal involvement in the process.

After responding to the patient's concerns, Finn (1996a) recommends that the clinician restate the questions posed earlier by the patient. This ensures the accuracy of what the patient would like to have addressed by the assessment. The patient also is encouraged to ask questions of the clinician, thus reinforcing the collaborative context or atmosphere that the clinician is trying to establish. Step 1 is completed as the instrumentation and its administration, as well as the responsibilities and expectations of each party, are clearly defined and the particulars of the process (e.g., date and time of assessment, date and time of the feedback session, clinician fees) are discussed and agreed upon.

4.18.7.3.2 Step 2: Preparing for the feedback session

Upon completion of the administration and scoring of the instrumentation used during the assessment, the clinician first outlines all results obtained from the assessment, including those not directly related to the patient's previously stated questions. Finn (1996a) presents a well-organized outline for the types of information that the trained user can extract from MMPI-2 data. These include response consistency, test-taking attitude, distress and disturbance, major symptoms, underlying personality, behavior in relationships, implications for treatment, diagnostic impression, and recommendations. Unfortunately, clinicians who do not or cannot use the MMPI-2 or other well-researched, multidimensional instruments will not have the same amount or type of data available to them. (This should not preclude them from identifying the types of valid and useful information that can be derived from the instruments and organizing it into a usable form for presentation to the patient.) This is followed by a determination of how to present the results to the patient. This can be guided by the clinician asking himself or herself the following questions:

> (i) How do the (test) findings relate to the client's goals?
> (ii) What are the most important findings of the (tests administered)?
> (iii) To what extent is the client likely to already know about and agree with the (test) findings?
> (iv) How much new information is the client likely to be able to integrate in the feedback session?
> (v) What is likely to happen if the client becomes overwhelmed or is presented with findings that are greatly discrepant from his/her current self-concept? (p. 34)

As a final point in this step, Finn (1996a) indicates that the clinician must determine what is the best way to present the information to the patient so that he or she can accept and integrate the information while maintaining his or her sense of identity and self-esteem. This also is a time when the clinician can identify information that he or she may not wish to reveal to the patient because it is not important to answering the patient's questions; doing so may negatively affect the collaborative relationship. In addition, the clinician may want to prepare for presenting those aspects of feedback that he or she feels will be most problematic for him or her (i.e., the clinician) by role-playing with a colleague.

4.18.7.3.3 Step 3: The feedback session

As Finn (1996a) states: "The overriding goal of feedback sessions is to have a therapeutic interaction with clients" (p. 44). Thus, the initial tasks of the feedback session are focused on setting the stage for this type of encounter. This is accomplished by allaying any anxiety the patient may have about the session, reaffirming the collaborative relationship, and familiarizing him or her with the presentation of the test results (e.g., explaining the profile sheet upon which the results are graphed, discussing the normative group to which he or she will be compared, providing an explanation of standard scores).

When the session preparation is completed, the clinician begins providing feedback to the patient (Finn, 1996a). This, of course, is centered on answering the questions posed by the patient during Step 1. Beginning with a positive finding from the assessment, the clinician proceeds to first address those questions that the patient is most likely to accept. He or she then carefully moves to the findings that are more likely to be anxiety-arousing for the patient and/or challenge his or her self-concept. A key element to this step is to have the patient verify the accuracy of each finding and provide a real-life example of the interpretation that is offered. Alternately, one should ask the patient to modify the interpretation to make it more in line with how he or she sees themselves and their situation. Finn (1996a) provides specific suggestions about how to deal with a rejection of a finding, the final suggestion being to allow the client to disagree with but not totally dismiss the finding. This leaves the door open for representing the finding at another time when the patient is more open to accepting it.

Finn (1996a) recommends that the clinician should end the session by responding to any additional questions the patient may have; confirming that the patient has accurately understood the information that was presented; giving permission for the patient to contact the clinician should further questions arise; and (in the assessment-only arrangement) termination of the relationship. Throughout the session, the clinician maintains a supportive stance with regard to any affective reactions to the findings.

4.18.7.3.4 Additional steps

Finn and Martin (in press) indicate two additional steps that may be added to the therapeutic assessment process. The purpose of the first additional step, referred to as an "assessment intervention session" essentially is to clarify initial test findings through the administration of additional instruments. For example, Finn and Martin explain how MMPI-2 findings can be further fleshed out through a nonstandard administration of an instrument such as the Thematic Apperception Test (TAT). Here, the clinician controls the patient's interpretation in order to draw out information relevant to the patient's questions. Also, solutions to problems elicited by the TAT cards are suggested to the patient.

The other additional step discussed by Finn and Martin (in press) is the provision of a written report of the findings to the patient. In addition to summarizing both the test results and the answers to the patient's questions, it also attempts to elicit feedback and reactions from the patient about the assessment.

The reader should note that the preceding summary presents only the key technical aspects of the therapeutic assessment procedures espoused by Finn and his associates. Much of the clinical/dynamic aspect of this approach has not been addressed because of the focus of this chapter. Those interested in incorporating the process into their clinical practice are encouraged to read Finn (1996a).

4.18.7.4 Empirical Support for Therapeutic Assessment

Noting the lack of direct empirical support for the therapeutic effects of sharing test results with patients, Finn and Tonsager (1992) investigated the benefits of providing feedback to university counseling center clients regarding their MMPI-2 results. A total of 32 subjects underwent therapeutic assessment and feedback procedures similar to those described above while on the counseling center's waiting list. Another 28 subjects were recruited from the same waiting list to serve as a control group. There were no significant differences between the two groups on any important demographic or examiner contact-interval variables.

Instead of receiving feedback, Finn and Tonsager's (1992) control group received nontherapeutic attention from the examiner. However, they were administered the same dependent measures as the feedback group at the same time as the experimental group received feedback. They also were administered the same dependent measures as the experimental group two weeks later (i.e., two weeks after the experimental group received the feedback) in order to determine if there were differences between the two groups on those dependent measures. These measures included a self-esteem questionnaire, a symptom checklist (i.e., the SCL-90-R), a measure of private and public self-consciousness, and a questionnaire assessing the subjects' subjective impressions of the feedback session. (Note that the control group was administered only that portion of the feedback assessment questionnaire that was relevant to them.)

The results of Finn and Tonsager's (1992) study indicated that compared to the control group, the feedback group demonstrated significantly less distress at the two-week post-feedback follow up, and significantly higher levels of self-esteem and hope at both the time of feedback and the two-week post-feedback follow up. In other findings, feelings about the feedback sessions were positively and significantly correlated with changes in self-esteem from testing to feedback, both from feedback to follow up and from testing to follow up among those who were administered the MMPI-2. In addition, the change in level of distress from feedback to follow up correlated significantly with private self-consciousness (i.e., the tendency to focus on the internal aspects of oneself) but not with public self-consciousness.

4.18.8 PSYCHOLOGICAL ASSESSMENT AS A TOOL FOR OUTCOMES MANAGEMENT

The 1990s have witnessed a positively accelerating growth curve reflecting the level of interest in and development of behavioral health care outcomes programs. Cagney and Woods (1994) attribute this to four major factors. First, behavioral health care purchasers are asking for information regarding the value of the services they buy. Second, there is an increasing number of purchasers who are requiring a demonstration of patient improvement and satisfaction. Third, MCOs need data that demonstrate that their providers render efficient and effective services. And fourth, outcomes information will be needed for the "quality report cards" that MCOs anticipate they will be required to provide in the future. In short, fueled by soaring health care costs, there has been an increasing need for providers to demonstrate that what they do is effective. And all of this has occurred within the context of the continuous quality improvement (CQI) movement, in which there have been similar trends in the level of interest and growth.

As this author has noted previously, the interest in and necessity for outcomes measurement in the era of managed care and accountability provides a unique opportunity for clinical psychologists to use their training and skills in assessment (Maruish, 1994). However, the extent to which the clinical psychologist becomes a key and successful contributor to an organization's outcomes initiative (whatever that might be) will depend on his or her understanding of what "outcomes" and their measurement and applications are all about.

4.18.8.1 What Are Outcomes?

Before discussing outcomes, it is important to have a clear understanding of what is meant by the term. Experience has shown that the meaning varies according to whom one may speak.

Donabedian (1985) has identified three dimensions of quality of care. "Structure" refers to the organization providing the care. It includes aspects such as how the organization is "organized," the physical facilities and equipment, and the number and professional qualifications of its staff. "Process" refers to the specific types of services that are provided to a given patient (or group of patients) during a specific episode of care. These might include various types of tests and assessments (e.g., psychological tests, lab tests, magnetic resonance imaging), therapeutic interventions (e.g., group psychotherapy, medication), and discharge planning activities. Treatment complications (e.g., drug reactions) are also included here. "Outcomes", on the other hand, refers to the results of the specific treatment that was rendered.

The outcomes, or results, of treatment should not refer to change in only a single aspect of functioning. Treatment may impact various facets of a patient's life. Stewart and Ware (1992) have identified five broad aspects of general health status: physical health, mental health, social functioning, role functioning, and general health perception. Treatment may affect these aspects of health in different ways, depending on the disease or disorder being treated and the effectiveness of the treatment. Some specific aspects of functioning related to these five areas of general health status that are commonly measured include: feeling of well-being, psychological symptom status, use of alcohol and other drugs, functioning on the job or at school, marital/family relationships, utilization of health care services, and ability to cope.

In considering the various types of outcomes that might be assessed in behavioral health care settings, a substantial number of clinicians probably would identify symptomatic change in psychological status as being the most important. Nevertheless, however important change in symptom status may have been in the past, clinical psychologists and other behavioral health care providers have come to realize that changes in many of the other aspects of functioning identified by Stewart and Ware (1992) are equally important indicators of treatment effectiveness. As Sederer et al. (1996) have noted:

> Outcome for patients, families, employers, and payers is not simply confined to symptomatic change. Equally important to those affected by the care rendered is the patient's capacity to function within a family, community, or work environment or to exist independently, without undue burden to the family and social welfare

> system. Also important is the patient's ability to show improvement in any concurrent medical and psychiatric disorder ... Finally, not only do patients seek symptomatic improvement, but they want to experience a subjective sense of health and well being. (p. 2)

A much broader perspective is offered in Faulker and Gray's *The 1995 behavioral outcomes and guidelines sourcebook* (Migdail, Youngs, & Bengen-Seltzer, 1995):

> Outcomes measures are being redefined from a vague "is the patient doing better?" to more specific questions, such as, "Does treatment work in ways that are measurably valuable to the patient in terms of daily functioning level and satisfaction, to the payor in terms of value for each dollar spent, to the managed care organization charged with administering the purchaser's dollars, and to the clinician charged with demonstrating value for hours spent?" (p. 1)

Thus, "outcomes" holds a different meaning for each of the different parties who have a stake in behavioral health care delivery. What is measured generally depends on the purposes for which outcomes assessment is undertaken. As will be shown, these vary greatly.

4.18.8.2 Outcomes Assessment: Measurement, Monitoring, and Management

Just as it is important to be clear about what is meant by outcomes, it is equally important to clarify the three general purposes for which outcomes assessment may be employed. The first is outcomes measurement. This involves nothing more than pre- and post-treatment assessment of one or more variables to determine the amount of change that has occurred (if any) in these variables as a result of therapeutic intervention.

A more useful approach is that of outcomes monitoring. This refers to "the use of periodic assessment of treatment outcomes to permit inferences about what has produced change" (Dorwart, 1996, p. 46). Like treatment progress monitoring used for treatment planning purposes, outcomes monitoring involves the tracking of changes in the status of one or more outcomes variables at multiple points in time. Assuming a baseline assessment at the beginning of treatment, reassessment may occur one or more times during the course of treatment (e.g., weekly, monthly), at the time of termination, and/or during one or more periods of post-termination follow up. Whereas treatment progress monitoring is used to determine how much the patient is on or off the expected course of improvement, outcomes monitoring focuses on revealing aspects about the therapeutic process that seem to affect change.

The third and most useful purpose of outcomes assessment is that of outcomes management. Dorwart (1996) defines outcomes management as "the use of monitoring information in the management of patients to improve both the clinical and administrative processes for delivering care" (pp. 46–47). In outcomes management, information is used to improve the quality of services offered to the patient population(s) served by the provider, not to any one patient. Information gained through the assessment of patients can provide the organization with indications of what works best with whom and under what set of circumstances, thus helping to improve the quality of services for all patients. In essence, outcomes management can serve as a tool for those organizations with an interest in implementing a CQI initiative.

4.18.8.3 The Benefits of Outcomes Assessment

The implementation of any type of outcomes assessment initiative within an organization does not come without effort from and cost to the organization. However, if it is implemented properly, all interested parties, that is, patients, clinicians, provider organizations, payers, and the health care industry as a whole, should find a substantial yield from the outlay of time and money. Cagney and Woods (1994) identify several benefits to patients, including enhanced health and quality of life, improved health care quality, and effective use of the dollars paid into benefits plans. For providers, the outcomes data can result in improved clinical skills, information related to the quality of care provided and local practice standards, increased profitability, and decreased concerns over possible litigation.

Outside of the clinical context, benefits also can accrue to payers and MCOs. Cagney and Woods (1994) see the potential payer benefits as including healthier workers, improved health care quality and worker productivity, and reduced or contained health care costs. As for MCOs, the benefits include increased profits, information that can help shape the practice patterns of their providers, and decisions that are based on quality of care.

4.18.8.4 The Therapeutic Use of Outcomes Assessment

The foregoing overview of outcomes assessment provides the background necessary for discussing the use of psychological outcomes assessment data in day-to-day clinical practice.

Whereas the focus of the above review was centered on both the individual patient and patient populations, it now will narrow to the use of outcomes assessment primarily in service to the individual patient. The reader interested in issues related to large, organization-wide outcomes studies conducted for outcomes management purposes (as defined above) is encouraged to seek other sources of information that specifically address that topic (see, for example, Migdail, Youngs, & Bengden-Seltzer, 1995; Newman, 1994).

There is no one system or approach to the assessment of treatment outcomes for an individual patient that is appropriate for all providers of behavioral health care services. Because of the specific type of outcomes one is interested in, the reasons for assessing them, and the manner in which they may impact the decisions made by the patient, payer and clinician, any successful and useful outcomes assessment approach must be customized. Customization should reflect the needs of the primary beneficiary of the information gained from the assessment (i.e., patient, payer, or provider), with consideration of the secondary stakeholders in the therapeutic endeavor. Ideally, the identified primary beneficiary would be the patient. Although this is not always the case, it would appear that only rarely would the patient not benefit, at least indirectly, from the gathering of outcomes data.

Following are considerations and recommendations for the development and implementation of an outcomes assessment initiative by behavioral health care providers. Although space limitations do not allow a comprehensive review of all issues and solutions, the information that follows can be useful to clinical psychologists and others with similar training wishing to begin to incorporate outcomes assessment into their standard therapeutic routine.

4.18.8.4.1 Purpose of the outcomes assessment

There are numerous reasons for assessing outcomes. For example, in a recent survey of 73 behavioral health care organizations, various reasons were identified by the participants as to why they had conducted outcomes studies (Pallak, 1994). Among the several indicated, the top five reasons (in descending order) were to: evaluate outcomes for patients, evaluate provider effectiveness, evaluate integrated treatment programs, manage individual patients, and support sales and marketing efforts. However, from the clinician's standpoint, a couple of purposes are worth noting.

In addition to monitoring the course of progress during treatment (see above), clinicians may employ outcomes assessment to obtain a direct measure of how much patient improvement has occurred as the result of the course of treatment intervention. Here, the findings are of more benefit to the clinician than to the patient himself because a pre- and post-treatment approach to the assessment is utilized. The information will not lead to any change in the patient providing the information, but the feedback it provides to the clinician could assist him in the treatment of other patients later on.

Another common reason for outcomes assessment is to demonstrate the patient's need for therapeutic services beyond that which is typically covered by the patient's medical and behavioral health care benefits. When assessment is conducted for this reason, the patient and clinician are only secondary beneficiaries of the outcomes data. As will be shown below, the type of information that a third party payer requires for authorization of extended benefits may not be the most relevant or useful to the patient or the clinician.

4.18.8.4.2 What to measure

The aspects or dimensions of patient functioning that are measured as part of outcomes assessment will depend on the purpose for which the assessment is being conducted. As discussed earlier, probably the most commonly measured variable is that of symptomatology or psychological/mental health status. After all, disturbance or disruption in this dimension is probably the most common reason why people seek behavioral health care services in the first place. However, there are other reasons for seeking help, including difficulties in coping with various types of life transitions (e.g., a new job, recent marriage or divorce, other changes in the work or home environment), inability to deal with the behavior of others (e.g., spouse, children), general dissatisfaction with life, or perhaps other less common reasons. Additional assessment of related variables therefore may be necessary, or may even take precedence over the assessment of symptoms or other mental health indicators.

In the vast majority of the cases seen for behavioral health care services, the assessment of the patient's overall level of psychological distress or disturbance will yield the most singularly useful information, regardless of whether it is used for outcomes measurement, outcomes monitoring, outcomes management, or to meet the requirements of third-party payers for authorization of additional benefits. Indices such as the Positive Symptom Total (PST) or Global Severity Index (GSI) that are part of the SA-45 or BSI can provide this type of information efficiently and economically.

For some patients, measures of one or more specific psychological disorders or symptom clusters are at least as important if not more important than overall symptom or mental health status. Here, if interest is in only one disorder or symptom cluster (e.g., depression), one may choose to measure only that particular set of symptoms using an instrument designed specifically for that purpose (e.g., use of the BDI with depressed patients). For those interested in assessing the outcomes of treatment relative to multiple psychological dimensions, the administration of more than one disorder-specific instrument or a single, multidimensional instrument which assesses all or most of the dimensions of interest would be required. Again, instruments such as the SA-45 or the BSI can provide a quick, broad assessment of multiple symptom domains. Although much lengthier, other multiscale instruments, such as the MMPI-2 or the PAI, permit a more detailed assessment of several disorders or symptom domains using one inventory.

In many cases, the assessment of mental health status is adequate for outcomes assessment purposes. There are other instances in which changes in psychological distress or disturbance either provide only a partial indication of the degree to which therapeutic intervention has been successful, are not of interest to the patient or a third-party payer, are unrelated to the reason why the patient sought services in the first place, or are otherwise inadequate or unacceptable as measures of improvement in the patient's condition. One may find that for some patients, improved functioning on the job, at school, or with family or friends is much more relevant and important than symptom reduction. For other patients, improvement in their quality of life or feeling of well-being is more meaningful.

It is not always a simple matter to determine exactly what should be measured. However, careful consideration of the following questions should greatly facilitate the decision.

(i) *Why did the patient seek services?* People pursue treatment for many reasons. The patient's stated reason for seeking therapeutic assistance may be the first clue in determining what is important to measure.

(ii) *What did the patient hope to gain from treatment?* The patient's stated goals for the treatment he or she is about to receive may be a primary consideration in the selection of outcomes to be assessed.

(iii) *What are the patient's criteria for the successful completion of the current therapeutic episode?* The patient's goals for treatment may provide only a broad target for the therapeutic intervention. Having the patient identify exactly what will have to happen to consider treatment successful and no longer needed will help in specifying the most important constructs and/or behaviors to assess.

(iv) *What are the clinician's criteria for the successful completion of the current therapeutic episode?* What the patient identifies as being important to accomplish during treatment may reflect a lack of insight into his or her problems, or it might not otherwise concur with what the clinician's experience would indicate.

(v) *What are the criteria of significant third parties for the successful completion of the current therapeutic episode?* From a strict treatment perspective, this should be given the least amount of consideration. From a more realistic perspective, one cannot overlook the expectations and limitations that one or more third parties have for the treatment that is rendered. The expectations and limitations set by the patient's behavioral health care plan, the guidelines of the organization in which the clinician practices, and possibly other external forces may significantly play into the decision about when to terminate treatment.

(vi) *What, if any, are the outcomes initiatives within the provider organization?* One cannot ignore any outcomes programs that have been initiated by the organization in which the therapeutic services are delivered. Regardless of the problems and goals of the individual patient, organization-wide studies of effectiveness may dictate the gathering of specific types of outcomes data from patients who have received services.

Note that the selection of the variables to be assessed may address more than one of the above issues. Ideally, this is what should happen. However, one needs to take care that the gathering of outcomes data does not become too burdensome. As a general rule, the more outcomes data one attempts to gather from a given patient or collateral, the less likely one is to obtain any data at all. The key is to identify the point at which the amount of data that can be obtained from a patient and/or collaterals, and the ease with which it can be gathered, is optimized.

4.18.8.4.3 How to measure

Once the decision concerning what to measure has been made, one must then decide how this should be measured. In many cases, the most important data will be that obtained directly from the patient through the use of self-report instruments. Underlying this assertion are the assumptions that valid and reliable instrumentation, appropriate to the needs of the patient, is available to the clinician; the patient can read at the level required by the instruments; and the patient is motivated to respond honestly to the questions asked. If this is not the case, other options are available.

Other types of data-gathering tools may be substituted for self-report measures. Rating scales completed by the clinician or other members of the treatment staff may provide information that is as useful as that elicited directly from the patient. In those cases in which the patient is severely disturbed, unable to give valid and reliable answers (e.g., younger children), unable to read, or is an otherwise inappropriate candidate for a self-report measure, clinical rating scales can substitute as useful means of gathering data. Related to these instruments are parent-completed inventories for child and adolescent patients. These are particularly useful in obtaining information about the child or teen's behavior that might not otherwise be known.

Collateral rating instruments can also be used to gather information in addition to that obtained from self-report measures. When used in this manner, these instruments provide a mechanism by which the clinician and other treatment staff can contribute data to the outcomes assessment endeavor. This not only results in the clinician or provider organization having more information upon which to evaluate the outcomes of therapeutic intervention, it also gives the clinician an opportunity to ensure that the perspective of the treatment provider is considered in the evaluation of the effects of the treatment given.

Another potential source of outcomes information is administrative data. In many of the larger provider organizations, this information is easily retrieved through their management information systems (MISs). Data related to the patient's diagnosis, dose and regimen of medication, physical findings, course of treatment, and other types of data typically stored in these systems can be useful to those evaluating the outcomes of therapeutic intervention.

4.18.8.4.4 When to measure

There are no hard and fast rules, guidelines, or accepted conventions related to when outcomes should be assessed. The common practice is to assess the patient at least at treatment initiation and treatment termination/discharge. Obviously, at the time of treatment initiation, the clinician should obtain a baseline measure of whatever variables will be measured at the termination. At the minimum, this allows for "outcomes measurement" as described above. As has been discussed, additional assessment of the patient on the variables of interest can take place at other points in time, that is, at other times during the course of treatment and upon post-discharge follow up.

Many would argue that postdischarge/post-termination follow-up assessment provides the best or most important indication of the outcomes of therapeutic intervention. Two types of comparisons may be made on follow-up. The first is a comparison of the patient's status on the variables of interest at the time of treatment initiation, or at the time of discharge or termination, to that of the patient at some point after treatment has ended. Either way, this follow-up data will provide an indication of the more lasting effects of the intervention. Generally, the variables of interest for this type of comparison include such things as symptom presence and intensity, feeling of well-being, frequency of substance use, and social and role functioning.

The second type of post-treatment investigation involves looking at the frequency at which some aspect(s) of the patient's life circumstances, behavior or functioning occurred during a given period prior to treatment, compared to that which occurred during an equivalent period of time immediately preceding the post-discharge assessment. This approach is commonly used in determining the cost-offset benefits of treatment. For example, the number of times a patient has been seen in an emergency room for psychiatric problems during the three-month period preceding the initiation of outpatient treatment can be compared to the number of emergency room visits during the three-month period preceding the postdischarge follow-up assessment. Not only can this provide an indication of the degree to which treatment has helped the patient deal with his problems, it can also demonstrate how much medical expenses have been reduced through the patient's decreased use of costly emergency room services.

In general, post-discharge outcomes assessment probably should take place no sooner than a month after treatment has ended. When feasible, one probably should wait three to six months to assess the variables. This should provide a more valid indication of the lasting effects of treatment. Assessments being conducted to determine the frequency with which some behavior or event occurs (as may be needed to determine cost-offset benefits) can be accomplished no sooner than the reference time interval used in the baseline assessment. Thus, suppose that the patient reports 10 emergency room visits during the three-month period prior to treatment. If one wants to know if the patient's emergency room visits have decreased after treatment, the assessment cannot take place any earlier than three months after treatment termination.

4.18.8.4.5 How to analyze outcomes data

There are two general approaches to the analysis of treatment outcomes data. The first is to determine whether changes in patient scores on outcomes measures are statistically significant. The other is to establish whether these changes are clinically significant. Use of standard tests of statistical significance is important in the analysis of group or population change data. Clinical significance is more relevant to change in the individual patient. As this chapter is focused on the individual patient, this section will center on matters related to determining clinically significant change as the result of treatment.

The issue of clinical significance has received a great deal of attention in psychotherapy research during the past several years. This is at least partially owing to the work of Jacobson and his colleagues (Jacobson, Follette, & Revenstorf, 1984, 1986; Jacobson & Truax, 1991) and others (e.g., Christensen & Mendoza, 1986; Speer, 1992; Wampold & Jenson, 1986). Their work came at a time when researchers began to recognize that traditional statistical comparisons do not reveal a great deal about the efficacy of therapy. In discussing the topic, Jacobson and Truax broadly define the clinical significance of a treatment as

> its ability to meet standards of efficacy set by consumers, clinicians, and researchers. While there is little consensus in the field regarding what these standards should be, various criteria have been suggested: a high percentage of clients improving...; a level of change that is recognizable by peers and significant others ...; an elimination of the presenting problem ...; normative levels of functioning at the end of therapy ... ; high end-state functioning at the end of therapy ...; or changes that significantly reduce one's risk for various health problems. (p. 12)

From their perspective, Jacobson and his colleagues (Jacobson, Follette, & Revenstorf, 1984; Jacobson & Truax, 1991) felt that clinically significant change could be conceptualized in one of three ways. Thus, for clinically significant change to have occurred, the measured level of functioning following the therapeutic episode would either:

(i) fall outside the range of the dysfunctional population by at least two standard deviations away from the mean of that population, in the direction of functionality;

(ii) fall within two standard deviations of the mean for the normal or functional population; or

(iii) be closer to the mean of the functional population than to that of the dysfunctional population.

Jacobson and Truax viewed the third option as being the least arbitrary and provided different recommendations for determining cutoffs for clinically significant change, depending upon the availability of normative data. Lambert (1994) demonstrated how the third option could be modified to allow for the inclusion of more than one categorization of dysfunction (e.g., mild, moderate, severe). This assumes, of course, that the necessary normative data needed to separate the gradations of dysfunction are available.

At the same time, these same investigators noted the importance of considering the change in the measured variables of interest from pre- to post-treatment, in addition to the patient's functional status at the end of therapy. To this end, Jacobson et al. (1984) proposed the concomitant use of a reliable change (RC) index to determine whether change is clinically significant. This index, modified on the recommendation of Christensen and Mendoza (1986), is nothing more than the pretest score minus the posttest score divided by the standard error of the difference of the two scores, expressed as:

$$\mathrm{RC} = (x_2 - x_1)/S_{\mathrm{diff}}$$

where x_1 is the pretest score, x_2 is the post-test score, and S_{diff} is the standard error of the difference. The standard error of the difference is computed as:

$$S_{\mathrm{diff}} = \sqrt{2\,(\mathrm{SEM})^2}$$

where SEM is the standard error of measurement for a functional group (e.g., normals, nonpatients) on the instrument. If the RC index is greater than 1.96, the change in scores is not likely to be due to chance ($p < 0.05$), but rather to reflect real change.

Speer (1992) recommended a different approach when regression to the mean has been demonstrated to contribute to the improvement in scores from pre- to post-test. The alternate approach, based on the combined work of Nunnally (1967) and Edwards, Yarvis, Mueller, Zingale, and Wagman (1978), involves developing a confidence interval of ± 2 SEMs around the estimated true pretest score. A post-test score falling outside of this confidence interval is considered significantly different from the initial pretest score at $p < 0.05$. Using this approach, more change is needed to show clinically significant improvement than to show clinically significant deterioration. Note that the criterion for determining whether regression to the mean is occurring is met when a negative correlation is found to exist between the pretreatment score and amount of change that has taken place. This implies the evaluation of group data, and for this reason this empirical criterion may not be of use for the individual patient unless the latter is a member of a sample for which test results are available.

Lambert (1994) proposes a modified recommendation for the dual criteria for clinically significant change (that is, RC greater than 1.96 and movement of the patient's score from the dysfunctional group's distribution to the functional group's distribution) such that movement from one degree of dysfunction to a lesser degree would also meet one of the two criteria for clinically significant change. In an example, Lambert illustrated that normative data for the Global Severity Index (GSI) found in the SCL-90-R literature can be used to empirically define four levels of symptom intensity: asymptomatic, mildly symptomatic, moderately symptomatic, and severely symptomatic. Assuming an RC of 1.96 or greater, clinically significant change can be said to have occurred if a patient's GSI score moves from severely to moderately or mildly symptomatic, or to asymptomatic; from moderately to mildly symptomatic, or to asymptomatic; or from mildly symptomatic to asymptomatic. Although this criterion is less stringent than having to move from being symptomatic (regardless of the severity) to asymptomatic, it still provides information that is quite useful for clinical decision making.

4.18.9 FUTURE DIRECTIONS

The ways in which clinical psychologists have conducted the types of psychological assessment described in this chapter have undergone dramatic changes during the 1990s. This should come as no surprise to anyone who spends a few minutes a day skimming the newspaper or watching television. The health care revolution started gaining momentum at the beginning of the 1990s and has not since slowed down. And there are no indications that it will subside in the foreseeable future. There was no real reason to think that behavioral health care would be spared from being a target of the revolution, and there is no good reason why it should have been spared. The behavioral health care industry certainly had contributed its share of waste, inefficiency, and lack of accountability to the problems that led to the revolution. Now, like other areas of health care, it is forced to "clean up its act." Some consumers of mental health or chemical dependency services have benefited from the revolution, others have not. In any case, the way in which health care is delivered and financed has changed, and clinical psychologists and other behavioral health care professionals must adapt to survive in the market.

Some of those involved in the delivery of psychological assessment services may wonder (with some fear and trepidation) where the revolution is leading the behavioral health care industry and, in particular, how their ability to practice will be affected. At the same time, others are eagerly awaiting the inevitable advances in technology and other resources that come with the passage of time. What will occur is open to speculation. However, close observation of the practice of psychological assessment and the various industries that support it (particularly the forms of therapeutic assessment described in this chapter) has led this author to arrive at some predictions as to where the field of therapeutic psychological assessment is headed and the implications these have for clinicians, provider organizations, and patients. What follows in this section are the most important of these predictions. Also included are what this author feels are the needs that must be met if psychological assessment is to continue to be a valued contributor to the delivery of efficient, cost-effective behavioral health care.

4.18.9.1 What the Industry Is Moving Away From?

One way of discussing what the field is moving toward is to first talk about what it is moving away from. In the case of therapeutic psychological assessment, two trends are becoming quite clear. First, starting at the beginning of this last decade of the twentieth century, the use of (and reimbursement for) psychological assessment has gradually been curtailed. This particularly has been the case with regard to indiscriminate assessment involving the administration of lengthy and expensive batteries of psychological tests. Payers began to demand evidence that the knowledge gained from the administration of these instruments contributed to the delivery of cost-effective, efficient care to mental health and substance abuse patients. There are no indications that this trend will stop.

Second, assessment has begun to move away from the use of lengthy, multidimensional objective instruments (e.g., the MMPI) or time-consuming projective techniques (e.g., Rorschach) that previously represented the standard of practice. When assessment is authorized now, it usually involves the use of inexpensive yet well-validated, problem-oriented instruments. This reflects modern behavioral health care's time-limited, problem-oriented approach to treatment. The clinician can no longer afford to spend a great deal of time in assessment activities when the patient has only a limited number of payer-authorized sessions with him or her. Thus, both now and in the foreseeable future, brief instruments will be used for problem identification or clarification, progress monitoring, and/or outcomes assessment.

4.18.9.2 Trends in Instrumentation

The move toward the use of brief, problem-oriented instruments for therapeutic psychological assessment purposes has just been identified. Another trend in the selection of instrumentation is the increasing use of public domain tests, questionnaires, rating scales, and other types of measurement tools. Previously, these free-use instruments were not developed with the rigor that is usually applied in the development of psychometrically sound instruments by commercial test publishers. Consequently, they typically lacked the validity and reliability data that are necessary to judge their psychometric integrity.

Recently, however, there has been a significant improvement in the quality and documentation of the public domain and other "for-free" tests that are available for therapeutic psychological assessment. Instruments such as the SF-36/SF-12 and HSQ/HSQ-12 health measures are good examples of such tools. These and instruments such as the Behavior and Symptom Identification Scale (BASIS-32; Eisen, Grob, & Klein, 1986) and the Outcome Questionnaire (OQ-45.1; Lambert, Lunnen, Umphress, Hansen, & Burlingame, 1994) have undergone psychometric scrutiny and have gained widespread acceptance. Although copyrighted, they may be used for a nominal one-time or annual licensing fee; thus, they generally are treated much like public domain assessment tools. One can expect that other quality, useful instruments will be made available for use at little or no cost in the future.

As for the types of instrumentation that will be needed and developed, one can probably expect some changes. Accompanying the increasing focus on outcomes assessment is a recognition by payers and patients that changes in several areas of functioning are at least as important as changes in level of symptom severity when evaluating the effectiveness of the treatment. For example, employers are interested in the patient's ability to resume the functions of his or her job, while family members may be concerned with the patient's ability to resume their role as spouse or parent. Increasingly, measurement of the patient's functioning in areas other than psychological/mental status has come to be included as part of behavioral health care outcomes systems. Probably the most visible indication of this is the incorporation of the SF-36 or HSQ into various behavioral health care studies, and the fact that two major psychological test publishers offer HSQ products in their catalogs of clinical products. One will likely see other public domain and commercially available nonsymptom-oriented instruments, especially those emphasizing social and occupational role functioning, in increasing numbers over the next several years.

Other types of instrumentation will also become prominent. These will include measures of variables that support the outcomes and other therapeutic assessment initiatives undertaken by provider organizations. What one organization or provider feels is important, or what it is told is important for reimbursement or other purposes, will dictate what is measured. Instrumentation will also include measures that will be useful for the prediction of outcomes for individuals seeking psychotherapeutic services from those organizations.

4.18.9.3 Trends in Data Use and Storage

There are two areas of application in which the valuable data obtained from therapeutic psychological assessment have heretofore been overlooked or underutilized. Indications are that this will change for both in the future. One area for which assessment data has potential application is that of clinical decision-making. This of course pertains only to the use of outcomes assessment data. Generally, data gathered solely for the purpose of outcomes assessment is used for just that: the assessment of the results of treatment. This is particularly the case in large, formal outcomes management programs. As has been discussed earlier in this chapter, data gathered at the beginning of treatment can be used immediately for treatment planning purposes while also serving as baseline data that can be compared to discharge data later on.

The other potential area of data application is in the development of local, regional, and national databases of therapeutic assessment data. Patient data gathered by various providers, organizations or programs within organizations, at one or more points during the therapeutic episode, can be pooled and used for various purposes. These databases can then serve as the bases for two highly beneficial (and probably profitable) endeavors. The first is the generation of sets of normative data for various populations delineated along any number of parameters. Norms for any number of instruments or health care variables could be generated "on demand" and continuously updated to reflect trends in behavioral health care. This author is aware of one large, national behavioral health care system where such a database already exists. He also is aware of efforts at establishing cross-organizational databases of this kind.

The second benefit afforded by the information contained in these databases is that of predictive modeling. For example, the behavioral health care organization mentioned above has taken advantage of the organizational data available to it to investigate the relationships between a number of treatment, demographic and other variables and the outcomes of treatment. Subjecting the large data sets available to it to sophisticated statistical analyses has allowed this organization to determine those types of patients requiring special care or attention in order to achieve desired outcomes at the time of treatment termination. Predictive modeling can also be used for identifying variables related to other aspects of patient care, such as patient satisfaction with the care received. The possibilities for the use of data in this manner are enormous.

4.18.9.4 Trends in the Application of Technology

Clinical psychologists have not been shy when it has come to taking advantage of the technological advances that have been achieved since the late 1970s. This is no more evident than in the extent to which the personal computer and the vast array of psychological assessment software have been incorporated into their delivery of clinical services. Automated administration, scoring, and interpretation and reporting of the results of nearly all major objective tests are currently available to the clinician through PC-based software. In addition, the availability of affordable desktop optical scanners allows the clinician to maintain the portability of the assessment instruments while retaining the scoring and interpreting power of the computer for processing the test data.

To speculate on how technology will be advanced in the service of therapeutic psychological assessment in the future is a risky businesss. As has been witnessed since the late 1970s, much can happen quickly. There are, however, three areas of technologic or technology-dependent advances on the horizon to which clinical psychologists should have access in the not too distant future. The first is the availability of online administration, scoring, and interpretation and reporting of tests via the Internet. In fact, an Internet version of the SA-45 is being beta tested at the time of writing (1997). To this author's knowledge, this represents the first use of the Internet for psychological assessment purposes. It is anticipated that the Internet version of the SA-45 will be commercially available in the very near future, and it will be quickly followed by the availability of Internet versions of other assessment instruments.

The second advance is actually a technology that has been around for a while but has undergone improvements, that is, the fax-back technology that is being used for scoring and reporting of objective, paper-and-pencil tests. Essentially, the fax machine replaces the optical scanner as a means of data entry. The electronic data is entered directly into the test publisher's computer for processing and report generation. However, instead of generating a hard-copy report of results at the processing site, the report is transmitted in electronic form and sent back to the clinician's fax machine within minutes of processing. At that point, the report is printed out just like any other fax transmission. Currently, this technology is being used on a somewhat limited basis. This is partially owing to the degree to which test publishers are making this form of automated scoring and reporting available to their customers. However, in the relatively near future, one should see more tests being offered to clinicians in this manner, particularly as the technology continues to improve.

The third area of technologic advancement has more to do with the application of technology than the development of new technology. L. E. Beutler and O. B. Williams (personal communication, January 15, 1996) have taken Beutler and Clarkin's (1990) Systematic Treatment Selection (STS) model of prescriptive treatment assignment and have developed specifications for software for automating the matching of treatments, therapists, and patients. The capability of subsequently tracking patients during the course of treatment is also included in these specifications. Originally entitled *STS for Windows*, this software is now under development through a behavioral health care publishing and consulting company. Driving the STS system is patient assessment data related to six variables: subjective distress, functional severity, problem complexity, potential for therapeutic resistance, coping style, and social support. Each of these variables may be assessed through either commercially available self-report instruments or clinician rating scales developed specifically for use with STS.

When fully developed, the STS should serve as the standard for in-office treatment–patient–therapist matching and patient-tracking software. According to the developers (L. E. Beutler & O. B. Williams, personal communication, January 15, 1996), the software will include numerous features, the most important of which will be: a comprehensive treatment planning report with up-to-date references to relevant research articles and treatment manuals; automatic entry of each patient's data into a growing database that is used for treatment planning and prediction; the ability to predict the amount of symptom reduction from a specific course of therapy; a report profiling the patient's symptom status over time; a report indicating individual clinician's ability to treat specific types of symptomatology; and the ability to incorporate case notes into the patient's electronic file. The major benefits of the system include the ability to: use different assessment means (self-report or clinician rating) to obtain the information needed to drive the system; develop treatment recommendations based on information that is optimal for the patient; easily monitor patient progress on a glide path developed from the treatment of similar patients and adjust the therapy plan (if necessary) on a timely basis; and determine a clinician's therapeutic strengths and weaknesses, thus permitting the most effective patient–therapist match.

All in all, when fully developed, the STS software will combine the knowledge and expertise of a leader in the field of psychotherapeutic research with state-of-the-art technology, thus yielding a powerful decision-making behavioral health care tool. One can be assured that similar products are likely to follow once the benefits of the STS software become widely known.

4.18.10 SUMMARY

The health care revolution has brought mixed blessings to those in the behavioral health care professions. It has resulted in limitations for reimbursement for services rendered and has forced many to change the way they practice their profession. At the same time, it has led to revelations about the cost savings benefits that can accrue from the treatment of mental health and substance use disorders. This has been the bright spot in an otherwise bleak picture for some behavioral health care professionals. For clinical psychologists, the picture appears to be somewhat different. They now have additional opportunities to contribute to the positive aspects of the revolution and to gain from the "new order" it has imposed. By virtue of their training in psychological assessment and through the application of appropriate instrumentation, they are uniquely qualified to support or otherwise facilitate multiple aspects of the therapeutic process. It is the clinical psychologist's contributions to aspects of "therapeutic psychological assessment" that this chapter has sought to identify and address in some detail.

Earlier, this author identified some of the types of psychological assessment instruments that are commonly used in the service of therapeutic endeavors. These included both brief and lengthy (multidimensional) symptom measures, as well as measures of general health status, quality of life, and patient satisfaction with the services received. Also identified were different sets of general criteria that can be applied when selecting instruments for use in therapeutic settings. The main intent of this chapter, however, was to present a detailed discussion of the various therapeutic uses of psychological assessment.

Generally, psychological assessment can assist the clinician in three important clinical activities: clinical decision-making, treatment itself (when used as a specific therapeutic technique), and treatment outcomes assessment. Regarding the first of these activities, three important clinical decision-making functions can be facilitated by psychological assessment: screening, treatment planning, and treatment monitoring. The first of these can be served by the use of "down and dirty" instruments to identify, within a high degree of certainty, the likelihood of the presence (or absence) of a particular condition or characteristic. Here, the diagnostic efficiency of the instrument used (as indicated by the PPP and NPP) is of great importance. Through their ability to identify and clarify problems as well as other important treatment-relevant patient characteristics, psychological assessment instruments can also be of great assistance in planning treatment. In addition, treatment monitoring, or the regular determination of the patient's progress throughout the course of treatment, can be served well by the application of psychological assessment instruments.

Secondly, assessment may be used as part of a therapeutic technique. In what Finn terms "therapeutic assessment," situations in which patients are evaluated via psychological testing are used as opportunities for the process itself to serve as a form of therapeutic intervention. This is accomplished through involving the patient as an active participant in the assessment process, not just as the object of the assessment.

Thirdly, psychological assessment can be employed as the primary mechanism by which the outcomes or results of treatment can be measured. However, the use of assessment for this purpose is not a cut-and-dried matter. As discussed, there are issues, pertaining to what to measure, how to measure, and when to measure, that require considerable thought prior to undertaking any standard (or even nonstandard) plan to assess outcomes. Guidelines for resolving these issues are presented, as is information pertaining to how to determine whether the measured outcomes of treatment are indeed "significant."

In the final section of the chapter, this author shares some thoughts about where psychological assessment is probably heading in the future. No radical revelations are presented since no signs really point in that direction. What is foreseen is the appearance of more quality assessment instruments that will remain in the public domain, and greater application of communications technology, fax and the Internet, in particular, as assessment delivery, scoring and reporting mechanisms. Also predicted is the application of tomorrow's computer technology to available assessment data for optimized treatment–patient–therapist matching. The innovative proposals of Beutler and Williams in this regard seem to represent the state-of-the-art thinking at this time.

There is no doubt that the practice of psychological assessment has been dealt a blow within recent years. However, as this chapter hopefully has shown, clinical psychologists have the skills to take this powerful tool, apply it in ways that will benefit those suffering from mental health and substance abuse problems, and demonstrate its benefits and their skills to patients and payers. Whether they will be successful in this demonstration will be determined in the near future. In the meantime, advances will continue to be made that will facilitate their work and improve its quality.

4.18.11 REFERENCES

American Psychological Association (1992). *Ethical principles of psychologists*. Washington, DC: Author.

American Psychological Association (1996). *The costs of failing to provide appropriate mental health care*. Washington, DC: Author.

Andrews, G., Peters, L., & Teesson, M. (1994). *The measurement of consumer outcomes in mental health*. Canberra, Australia: Australian Government Publishing Service.

Appelbaum, S. A. (1990). The relationship between assessment and psychotherapy. *Journal of Personality Assessment, 54*, 791–801.

Attkisson, C. C., & Zwick, R. (1982). The Client Satisfaction Questionnaire: Psychometric properties and correlations with service utilization and psychotherapy outcome. *Evaluation and Program Planning, 6*, 233–237.

Baldessarini, R. J., Finkelstein, S., & Arana, G. W. (1983). The predictive power of diagnostic tests and the effect of prevalence of illness. *Archives of General Psychiatry, 40*, 569–573.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.

Beutler, L. E., & Clarkin, J. (1990). *Systematic treatment selection: Toward targeted therapeutic interventions*. New York: Brunner/Mazel.

Beutler, L. E., Wakefield, P., & Williams, R. E. (1994). Use of psychological tests/instruments for treatment planning. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 55–74). Hillsdale, NJ: Erlbaum.

Beutler, L. E., & Williams, O. B. (1995). Computer applications for the selection of optimal psychosocial therapeutic interventions. *Behavioral health care Tomorrow, 4*, 66–68

Burlingame, G. M., Lambert, M. J., Reisinger, C. W., Neff, W. M., & Mosier, J. (1995). Pragmatics of tracking mental health outcomes in a managed care setting. *Journal of Mental Health Administration, 22*, 226–236.

Butcher, J. N. (1990). *The MMPI-2 in psychological treatment*. New York: Oxford University Press.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A. M., & Kaemmer, B. (1989). *MMPI-2: Manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.

Butcher, J. N., Graham, J. R., Williams, C. L., & Ben-Porath, Y. (1989). *Development and use of the MMPI-2 content scales*. Minneapolis, MN: University of Minnesota Press.

Cagney, T., & Woods, D. R. (1994). Why focus on outcomes data? *Behavioral health care Tomorrow, 3*, 65–67.

Center for Disease Control and Prevention (1994, May 27). Quality of life as a new public health measure: Behavioral risk factor surveillance system. *Morbidity and Mortality Weekly Report, 43*, 375–380.

Christensen, L., & Mendoza, J. L. (1986). A method of assessing change in a single subject: An alteration of the RC index [Letter to the editor]. *Behavior Therapy, 17*, 305–308.

Ciarlo, J. A., Brown, T. R., Edwards, D. W., Kiresuk, T. J., & Newman, F. L. (1986). *Assessing mental health treatment outcomes measurement techniques*. DHHS Pub. No. (ADM)86–1301. Washington, DC: US Government Printing Office.

Commission on Chronic Illness (1987). *Chronic illness in the United States: 1*. Cambridge, MA: Commonwealth Fund, Harvard University Press.

Derogatis, L. R. (1983). *SCL-90-R: Administration, scoring and procedures manual-II*. Baltimore: Clinical Psychometric Research.

Derogatis, L. R. (1992). *BSI: Administration, scoring and procedures manual-II*. Baltimore: Clinical Psychometric Research.

Derogatis, L. R., & DellaPietra, L. (1994). Psychological tests in screening for psychiatric disorder. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 22–54). Hillsdale, NJ: Erlbaum.

Derogatis, L. R., Lipman, R. S., & Covi, L. (1973). SCL-90: An outpatient psychiatric rating scale—preliminary report. *Psychopharmacology Bulletin, 9,* 13–27.

Dickey, B., & Wagenaar, H. (1996). Evaluating health status. In L. I. Sederer & B. Dickey (Eds.), *Outcomes assessment in clinical practice* (pp. 55–60). Baltimore: Williams & Wilkins.

Donabedian, A. (1985). *Explorations in quality assessment and monitoring: Vol. III. The methods and findings of quality assessment monitoring: An illlustrative analysis.* Ann Arbor, MI: Health Administration Press.

Dorwart, R. A. (1996). Outcomes management strategies in mental health: Applications and implications for clinical practice. In L. I. Sederer & B. Dickey (Eds.), *Outcomes assessment in clinical practice* (pp. 45–54). Baltimore: Williams & Wilkins.

Dowd, E. T., Milne, C. R., & Wise, S. L. (1991). The therapeutic Reactance Scale: A measure of psychological reactance. *Journal of Counseling and Development, 69,* 541–545.

Edwards, D. W., Yarvis, R. M., Mueller, D. P., Zingale, H. C., & Wagman, W. J. (1978). Test-taking and the stability of adjustment scales: Can we assess patient deterioration? *Evaluation Quarterly, 2,* 275–292.

Eisen, S. V., Grob, M. C., & Klein, A. A. (1986). BASIS: The development of a self-report measure for psychiatric inpatient evaluation. *The Psychiatric Hospital, 17,* 165–171.

Elwood, R. W. (1993). Psychological tests and clinical discrimination: Beginning to address the base rate problem. *Clinical Psychology Review, 13,* 409–419.

Ficken, J. (1995). New directions for psychological testing. *Behavioral Health Management, 20,* 12–14.

Finn, S. E. (1996a). *Manual for using the MMPI-2 as a therapeutic intervention.* Minneapolis, MN: University of Minnesota Press.

Finn, S. E. (1996b). Assessment feedback integrating MMPI-2 and Rorschach findings. *Journal of Personality Assessment, 67,* 543–557.

Finn, S. E., & Butcher, J. N. (1991). Clinical objective personality assessment. In M. Hersen, A. E. Kazdin, & A. S. Bellack (Eds.), *The clinical psychology handbook* (2nd ed., pp. 362–373). New York: Pergamon.

Finn, S. E., & Martin, H. (in press). Therapeutic assessment with the MMPI-2 in managed health care. In J. N. Butcher (Ed.), *Objective personality assessment in managed health care: A practitioner's guide.* Minneapolis, MN: University of Minnesota Press.

Finn, S. E., & Tonsager, M. E. (1992). Therapeutic effects of providing MMPI-2 test feedback to college students awaiting therapy. *Psychological Assessment, 4,* 278–287.

Friedman, R., Sobel, D., Myers, P., Caudill, M., & Benson, H. (1995). Behavioral medicine, clinical health psychology, and cost offset. *Health Psychology, 14,* 509–518.

Gough, H. G., McClosky, H., & Meehl, P. E. (1951). A personality scale for dominance. *Journal of Abnormal and Social Psychology, 46,* 360–366.

Gough, H. G., McClosky, H., & Meehl, P. E. (1952). A personality scale for social responsibility. *Journal of Abnormal and Social Psychology, 47,* 73–80.

Greene, R. L. (1991). *The MMPI-2/MMPI: An interpretive manual.* Boston: Allyn & Bacon.

Greene, R. L., & Clopton, J. R. (1994). Minnesota Multiphasic Personality Inventory—2. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 137–159). Hillsdale, NJ: Erlbaum.

Greenfield, T. K., & Attkisson, C. C. (1989). Progress toward a multifactorial service satisfaction scale for evaluating primary care and mental health services. *Evaluation and Program Planning, 12,* 271–278.

Hathaway, S. R., & McKinley, J. C. (1951). *MMPI manual.* New York: The Psychological Corporation.

Health Outcomes Institute (1993). *Health Status Questionnaire 2.0 manual.* Bloomington, MN: Author.

Hsiao, J. K., Bartko, J. J., & Potter, W. Z. (1989). Diagnosing diagnoses: Receiver operating characteristic methods and psychiatry. *Archives of General Psychiatry, 46,* 664–667.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15,* 336–352.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1986). Toward a standard definition of clinically significant change [Letter to the editor]. *Behavior Therapy, 17,* 309–311.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12–19.

Jahoda, M. (1958). *Current concepts of mental health.* New York: Basic Books.

Lambert, M. J. (1994). Use of psychological tests for outcome assessment. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 75–97). Hillsdale, NJ: Erlbaum.

Lambert, M. J., Lunnen, K., Umphress, V., Hansen, N. B., & Burlingame, G. M. (1994). *Administration and scoring manual for the Outcome Questionnaire (OQ-45.1).* Salt Lake City, UT: IHC Center for Behavioral health care Efficacy.

Larsen, D. L., Attkisson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning, 2,* 197–207.

LeVois, M., Nguyen, T. D., & Attkisson, C. C. (1981). Artifact in client satisfaction assessment: Experience in community mental health settings. *Evaluation and Program Planning, 4,* 139–150.

Maruish, M. (1990, Fall). Psychological assessment: What will its role be in the future? *Assessment Applications,* 7–8.

Maruish, M. E. (1994). Introduction. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 3–21). Hillsdale, NJ: Erlbaum.

Megargee, E. I., & Spielberger, C. D. (1992). Reflections on fifty years of personality assessment and future directions for the field. In E. I. Megargee & C. D. Spielberger (Eds.), *Personality assessment in America* (pp. 170–190). Hillsdale, NJ: Erlbaum.

Mental Health Weekly. (1996a, April 8). Future targets behavioral health field's quest for survival, 1–2.

Mental Health Weekly. (1996b, April 8). Leaders predict integration of MH primary care by 2000, 1–6.

Mental Health Weekly. (1996c, April 29). Anxiety disorders screening day occurs this week, 7.

Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine, 8,* 283–298.

Migdail, K. J., Youngs, M. T., & Bengen-Seltzer, B. (Eds.) (1995). *The 1995 behavioral outcomes & guidelines sourcebook.* New York: Faulkner & Gray.

Millon, T. (1994). *MCMI-III manual.* Minneapolis, MN: National Computer Systems.

Moreland, K. L. (1996). How psychological testing can reinstate its value in an era of cost containment. *Behavioral health care Tomorrow, 5,* 59–61.

Morey, L. C. (1991). *The Personality Assessment Inventory professional manual.* Odessa, FL: Psychological Assessment Resources.

Morey, L. C., & Henry, W. (1994). Personality Assessment Inventory. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 185–216). Hillsdale, NJ: Erlbaum.

Newman, F. L. (1994). Selection of design and statistical procedures for progress and outcome assessment. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 111–134). Hillsdale, NJ: Erlbaum.

Newman, F. L., & Ciarlo, J. A. (1994). Criteria for selecting psychological instruments for treatment outcome assessment. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 98–110). Hillsdale, NJ: Erlbaum.

Nguyen, T. D., Attkisson, C. C., & Stegner, B. L. (1983). Assessment of patient satisfaction: Development and refinement of a service evaluation questionnaire. *Evaluation and Program Planning, 6,* 299–313.

Nunnally, J. C. (1967). *Psychometric theory*. New York: McGraw-Hill.

Olfson, M., & Pincus, H. A. (1994a). Outpatient psychotherapy in the United States, I: Volume, costs, and user characteristics. *American Journal of Psychiatry, 151,* 1281–1288.

Olfson, M., & Pincus, H. A. (1994b). Outpatient psychotherapy in the United States, II: Patterns of utilization. *American Journal of Psychiatry, 151,* 1289.

Oss, M. E. (1996). Managed behavioral health care: A look at the numbers. *Behavioral Health Management, 16,* 16–17.

Pallak, M. S. (1994). National outcomes management survey: Summary report. *Behavioral health care Tomorrow, 3,* 63–69.

Phelps, R. (1996, February). Preliminary practitioner survey results enhance APA's understanding of health care environment. *Practitioner Focus, 9,* 5.

Psychotherapy Finances (1995, January). Fee, practice and managed care survey. 21(1), Issue 249.

Radosevich, D., & Pruitt, M. (1996). *Twelve-item Health Status Questionnaire (HSQ-12) version 2.0 user's guide*. Bloomington, MN: Health Outcomes Institute.

Radosevich, D. M., Wetzler, H., & Wilson, S. M. (1994). *Health Status Questionnaire (HSQ) 2.0: Scoring comparisons and reference data*. Bloomington, MN: Health Outcomes Institute.

Schlosser, B. (1995). The ecology of assessment: A "patient-centric" perspective. *Behavioral health care Tomorrow, 4,* 66–68.

Sederer, L. I., Dickey, B., & Hermann, R. C. (1996). The imperative of outcomes assessment in psychiatry. In L. I. Sederer & B. Dickey (Eds.), *Outcomes assessment in clinical practice* (pp. 1–7). Baltimore: Williams & Wilkins.

Simmons, J. W., Avant, W. S., Demski, J., & Parisher, D. (1988). Determining successful pain clinic treatment through validation of cost effectiveness. *Spine, 13,* 34.

Sipkoff, M. Z. (1995, August). Behavioral health treatment reduces medical costs: Treatment of mental disorders and substance abuse problems increases productivity in the workplace. *Open Minds,* 12.

Speer, D. C. (1992). Clinically significant change: Jacobson and Truax (1991) revisited. *Journal of Consulting and Clinical Psychology, 60,* 402–408.

Spielberger, C. D. (1983). *Manual of the State–Trait Anxiety Inventory: STAI (Form Y)*. Palo Alto, CA: Consulting Psychologists Press.

Stewart, A. L., & Ware, J. E., Jr. (1992). *Measuring functioning and well-being*. Durham, NC: Duke University Press.

Strain, J. J., Lyons, J. S., Hammer, J. S., Fahs, M., Lebovits, A., Paddison, P. L., Snyder, S., Strauss, E., Burton, R., & Nuber, G. (1991). Cost offset from a psychiatric consultation-liaison intervention with elderly hip fracture patients. *American Journal of Psychiatry, 148,* 1044–1049.

Strategic Advantage, Inc. (1996). *Symptom Assessment-45 Questionnaire manual.* Minneapolis, MN: Author.

Substance Abuse Funding News, (1995, December 22). Brown resigns drug post. 7.

VandenBos, G. R., DeLeon, P. H., & Belar, C. D. (1993). How many practitioners are needed? It's too early to know! *Professional Psychology: Research and Practice, 22,* 441–448.

Vermillion, J., & Pfeiffer, S. (1993). Treatment outcome and continuous quality improvement: Two aspects of program evaluation. *Psychiatric Hospital, 24,* 9–14.

Wampold, B. E., & Jenson, W. R. (1986). Clinical significance revisited [Letter to the editor]. *Behavior Therapy, 17,* 302–305.

Ware, J. E., Kosinski, M., & Keller, S. D. (1995). *SF-12: How to Score the SF-12 Physical and Mental summary scales* (2nd ed.). Boston: New England Medical Center, The Health Institute.

Ware, J. E., & Sherbourne, C. D. (1992). The MOS 36-Item Short Form Health Survey (SF-36). I. Conceptual framework and item selection. *Medical Care, 30,* 473–483.

Ware, J. E., Snow, K. K., Kosinski, M., & Gandek, B. (1993). *SF-36 Health Survey manual and interpretation guide.* Boston: New England Medical Center, The Health Institute.

Werthman, M. J. (1995). A managed care approach to psychological testing. *Behavioral Health Management, 15,* 15–17.

World Health Organization (1948). Constitution. In *Basic Documents*. Geneva, Switzerland: Author.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

4.19
Forensic Assessment

DAVID FAUST
University of Rhode Island, Kingston, RI, USA

4.19.1 INTRODUCTION 563
4.19.2 SOME KEY ISSUES AND CONSIDERATIONS IN PERSONAL INJURY CASE 564
  4.19.2.1 Four Central Issues 564
  4.19.2.2 The Four Elements and the Forensic Examiner's Task 565
  4.19.2.3 Treating Professionals Called into Court 568
    4.19.2.3.1 Maintaining the treatment alliance as primary 568
    4.19.2.3.2 Release of assessment or treatment records 569
4.19.3 ADMISSIBILITY 570
4.19.4 ASSESSMENT METHODS 573
  4.19.4.1 Informed Consent 573
  4.19.4.2 Clarifying Referral Questions and Deciding Whether they are Appropriate and Within the Clinician's Expertise 573
  4.19.4.3 Access to Information 574
  4.19.4.4 Design of Assessment Procedures: Some Do's and Don't Do's 575
    4.19.4.4.1 Use the best available methods 575
    4.19.4.4.2 Obtaining adequate information 578
    4.19.4.4.3 Conduct technically proficient evaluations 580
    4.19.4.4.4 Give adequate consideration to alternatives 582
  4.19.4.5 Interpretive Strategies 584
  4.19.4.6 Preparing Reports 587
4.19.5 LAWYERS' STRATEGIES AND TACTICS 588
  4.19.5.1 Credentials 589
  4.19.5.2 Bias 589
  4.19.5.3 Manner of Conducting the Examination 590
  4.19.5.4 Erroneous or Questionable Conclusions 590
  4.19.5.5 Scientific Status 590
4.19.6 DEPOSITIONS AND TRIAL TESTIMONY 591
  4.19.6.1 A Sampling of Deposition Topics 591
  4.19.6.2 Some Suggestions for Depositions 594
  4.19.6.3 Trial Testimony 596
4.19.7 REFERENCES 598

4.19.1 INTRODUCTION

Psychologists interface with the courts in various contexts and situations. A developmental psychologist may describe a 3 year old's recognition of risk in a product liability case in which a child was injured playing with a toy. A clinical psychologist may testify about the more suitable parent in a custody dispute in family court, or may opine about a defendant's mental state when he committed murder in a criminal case. A neuropsychologist may give evidence about cognitive or brain functions in a personal injury case. Other areas of testimony might include such diverse topics as the trustworthiness of eye witnesses reports, the appropriateness of a statistical analysis purportedly linking a toxin with a disorder, test bias in screening job applicants, and the characteristics of individuals who delay a considerable period of time before reporting traumatic events.

Such qualitative variation in courtroom activities precludes coverage of all these topics in one chapter and necessitates an alternative strategy. Much of the material to follow will have a generic quality, that is, many of the points that will be raised relate to a broad array of courtroom work. However, such generic discussion will be anchored by, or mainly directed towards, the civil arena, with an emphasis on distress claims (e.g., post-traumatic stress disorder) and brain injury claims. These topics provide suitable illustration of many of the more general points that will be raised. Additionally, distress and brain injury claims are very active areas in psychology and law, and hence of likely interest to a broad audience. Although assessment is the organizing theme of the chapter, this activity is not entirely separable from the context of courtroom opinions and other collateral issues (e.g., depositions), and thus the material that follows will not be rigidly limited to evaluative methodology. Finally, readers who are mainly interested in a highly specialized area (e.g., child custody, or multiple personality in criminal cases) may supplement the material presented here with additional sources.

The chapter will begin with coverage of basic issues that frame psychologists' courtroom involvement and that are often the focus of legal proceedings, such as the core elements of civil cases and whether psychologists participate as treating professionals or as retained experts. I will then discuss, in turn, the admissibility of psychological evidence; methods for conducting and improving courtroom evaluations; and case strategies and tactics, including lawyers' main areas of attack, depositions, direct testimony, and cross-examination.

The emphasis of this chapter is far more practical than philosophical or theoretical: the main intent is to provide materials that will enhance the quality of mental health professionals' legal work. Other psychologists may take a different view than my own on the level or type of scientific backing that is appropriate for legal work, although my views on this matter are considerably more moderate than some readers might suppose (e.g., I simply maintain that psychological testimony should normally rest on fundamentally sound scientific foundations [see Faust,1993]). Further, I believe that psychology is making substantial strides, and that there are an increasing number of areas in which testimony is likely to be of considerable scientific and practical merit in the courtroom (see the upcoming revision of Faust, Ziskin, & Hiers, 1991; Faust, Zisken, Hiers, & Miller, in press). Whatever differences there might be in viewpoint about the needed level of scientific backing for legal testimony, which, of course, is ultimately a matter for the courts to decide, presumably all psychologists would agree that there is merit in increasing the quality of courtroom work: it is with that spirit and intent that this chapter is written.

Some psychologists are hardened veterans of courtroom work and others are just beginning to consider courtroom involvement. This chapter is directed mainly towards those at an intermediate or beginning level. Further, some of the points I raise are based on my experience with legal work, particularly as a consultant. I am not assuming that my experience with legal cases is necessarily representative of such cases overall, and I am not using my observations to try to reach scientific conclusions or generalizations, but rather to provide what I hope will be practical guidance.

The guidelines and suggestions for the practice of forensic psychological assessment detailed in this chapter emerge from, and are directed towards, the USA court system, and their generalization and application to other court systems vary. Nevertheless, many of the main points likely will apply to forensic work in other countries, in particular the emphasis on methods and strategies that increase the likelihood of accurate conclusions.

4.19.2 SOME KEY ISSUES AND CONSIDERATIONS IN PERSONAL INJURY CASE

4.19.2.1 Four Central Issues

Personal injury cases, which involve civil (and not criminal) law, address situations in which it is claimed that one or more individuals have not carried out some duty owed to one or more other individuals, or have not exercised reasonable care. For example, Smith may have been negligent in failing to shovel the driveway, and Jones may have slipped on the ice and suffered a broken arm. Many personal injury cases involve lay individuals, not professionals. Thus, the legal issues frequently relate to responsibilities of the general citizenry, such as exercising reasonable care when driving a car or shooting a gun.

Individuals are held to obligations and standards commensurate with the role in which they are functioning and the assumed levels of knowledge and responsibility that are expected or required. For example, in a medical malpractice case, a neurosurgeon will almost surely be held to a much higher standard in recognizing a brain tumor than a clinical psychologist with a general practice. At the same time, both the neurosurgeon's and the psychologist's responsibilities are dictated by their roles as professionals. Once they take on a patient's case, they are not only obligated to the duties that bind the ordinary citizen, but in addition the responsibilities involved in the professional care of a patient. In contrast, and obviously, ordinary citizens are not obligated to properly diagnose and treat individuals for medical or psychological disorders. The citizen may be responsible for the broken arm resulting from the fall on his unshoveled driveway, but not for interpreting the x ray or setting the cast properly.

In most legal cases, in order for a plaintiff to prevail, or at least to make a monetary recovery, four elements, and all four elements, must be proven (this also applies to malpractice cases). First, there must be a duty owed. For example, we might owe a duty to maintain our premises free of foreseeable hazards, but not to maintain our neighbor's premises. Second, it must be proven that there was a breach of duty, or that reasonable care was not exercised. For example, if we do not shovel our sidewalk, we may well have breached our duty. But if an earthquake caused our sidewalk to crack and someone fell and hurt himself before the rumbling stopped and action could be taken, the plaintiff likely will not prevail.

Third, the breach of duty must be a cause of the harm. If the person fell, but it was because her feet got tangled in the dog's leash and had nothing to do with the tools left scattered on the stoop of the shop, the business owner is not liable. Depending on the circumstances, issues, and other possible factors, the potential agents in question may need to be the main cause of the harm, a substantial contributor, or may only have to have made any type of meaningful (versus trivial) contribution. Damages are sometimes awarded in relation to the relative contribution of the cause at issue. Thus, if the event is assumed to account for 10% of the outcome, one multiplies the award by .10. In other cases, if a factor made any significant contribution, the defendant is responsible for the total damage that ensued. For example, if the defendant's careless driving caused the accident, then it makes no difference that most of the plaintiff's injuries would have been avoided had she worn a seat belt.

The fourth and final issue is damages. Assuming the first three criteria are met and the defendant is hence found to be blameworthy (i.e., liable), plaintiffs are to be compensated for any damages stemming from the event. A duty that was owed may have been breached and caused an undesired event, but if that event resulted in no damages or minimal damages, there will be little or no award (and the case, in fact, will probably go nowhere in the courts). For example, a therapist may have engaged in an egregious action, such as throwing a suicidal patient out of treatment the moment he discovered the patient's insurance had run out. However, if no harm resulted, for example, the patient did not make a suicide attempt and through some odd quirk ended up benefiting from the termination, there are no damages for which the patient is to be compensated, although the courts or a professional board could still sanction the professional.

4.19.2.2 The Four Elements and the Forensic Examiner's Task

Depending on the type of civil case in which a psychologist is involved as an expert and the particulars, any or all of the four elements discussed may be pertinent to the professional's task. In malpractice cases, for example, the professional may address each of the elements. There may be debate about whether: (a) a brief or even casual contact with a troubled person created a professional duty or relationship, or about the nature of that duty; (b) whether the psychologist has followed reasonable care in assessing suicidal ideation or has met a certain standard of care; (c) whether failure to meet some practice standard caused the harm the person suffered; and (d) about the type and extent of harm that occurred. However, in many other cases, mental health professionals do not touch on at least the first two issues, that is, whether or what duty was owed and whether it was breached. For example, a clinical psychologist usually does not testify about proper driving practices.

The issue of causation may or may not be central to the psychologist's courtroom testimony. In some cases, another expert will testify about cause. For example, a neurologist may indicate that the plaintiff's cognitive disturbance is due to brain damage suffered in the accident at issue. In such a case, a neuropsychologist might primarily address the issue of damages, or the cognitive and behavioral impairments that are present. There may be little need for the neuropsychologist to link the damages to brain injury because another professional, the neurologist in this case, has already made the causal connection. The plaintiff's attorney may prefer to leave the issue of cause in the neurologist's hands, or the neuropsychologist may be restricted from providing causal testimony. In some cases there may be multiple causal elements to be considered, one or more of which the psychologist will address. For example, the issue of whether reckless driving caused the accident can be separate from an issue such as whether the accident caused brain damage. The first question about cause involves the assignment of blame for the event, and the second the link between the event and its possible consequences. The jury could decide that an improper left-hand turn caused the accident, and that the accident caused a chronic neck condition but not a brain injury.

In other cases, however, if the mental health professional does not testify to cause, there may be no other expert who can. For example, in a malpractice case, it will likely be up to the psychiatrist or psychologist to connect a therapist's breach of duty with the plaintiff's problems, or at least those problems linked to the event (such as depression, increased anxiety, or decreased responsiveness to therapeutic intervention). If there are plausible alternative causes for the problems (e.g., substance abuse or other emotional traumas), and if the professional herself cannot determine whether there is a causal link, how can one expect the jury to make a judgment about a matter the expert cannot decide? A decent cross-examiner can thus quickly show that with all the professional's presumed expertise, she cannot say there is a link between the event and adverse functioning. This may be devastating to the plaintiff's case. In some instances, if the expert's inability to address cause comes to light prior to trial (e.g., during a deposition that is taken as part of the discovery process), this may lead the judge to place tight restrictions on the psychologist's trial testimony, or even to exclude the expert or dismiss the entire case. For example, in a toxin case in which brain damage is being claimed, if the attorney can produce no expert to help with this core element—causality—it may not be possible to prove it legally and the case may be over.

The need to establish cause also illustrates some of the differences that can arise across clinical and courtroom situations. In clinical practice, a determination of cause or etiology may make little difference because it may not alter treatment. In the courtroom, the specific cause, whatever the treatment implications, may make all of the difference. A psychologist may fail to recognize how an issue that can be of such minimal relevance in a clinical situation, and, therefore, sometimes bypassed or set aside during the course of an evaluation, might have huge repercussions at deposition or trial. This is also why, in the legal arena, it is often not enough to identify a condition as present. It may be true that Smith has brain damage, but if it was not the car accident but a toxic exposure in a different place and circumstance that caused it, the driver of the other vehicle is not responsible. It follows that careful consideration of alternate causes, including procurement of relevant documents and information, is often crucial to the psychologist's legal work.

Psychologists usually should have a basic familiarity with key legal issues in a case before they undertake, or, preferably, before they decide whether to undertake, a courtroom evaluation, and should not assume that the lawyer will convey this information. The psychologist is typically better off knowing in advance whether cause is at issue and whether the lawyer does (or does not) want the psychologist to address this matter; and whether the standard applicable to the case is a substantial versus primary contributor. Such foreknowledge may allow psychologists to identify circumstances in which they are being asked to do something that is unreasonable or beyond their expertise, or may allow them to direct their efforts towards the matters at issue. I have seen many cases in which psychologists had limited or minimal acquaintance with key issues in a case and overlooked them in their assessments or reports. It is difficult to conduct an optimal or proper examination when one does not know the specific purposes to which it should be directed. For example, exhaustive cognitive testing may be of little use if there is really no issue in the case related to intellectual functioning, or occupational functioning may not be covered thoroughly although it is far and away the most important area of dispute.

Monetary awards are intended to cover past, present, and future expenses, and to compensate for losses that follow and are due to the injury. Expenses might include damaged property, needed equipment (e.g., a cervical collar), alterations of one's home setting (e.g., a wheelchair accessible ramp), and treatment. Treatment might be one of the largest, if not the largest, component of damages. For example, in a serious brain injury case, extended inpatient rehabilitation and outpatient services may cost hundreds of thousands of dollars.

Losses might include earnings, pain and suffering, and, at least theoretically separate from the latter and where allowed, "hedonic damages." Hedonic damages refer to lost opportunities to experience pleasurable or positive events (as opposed to experiencing negative events), such as the enjoyment of attending one's senior prom. Consortium claims are also common, in which the spouse claims that the injured partner is less able to fulfill marital roles (e.g., the injured person has decreased sexual interest or capacity). With lost earnings, one considers time missed from work to date and whether the person has returned, or likely will return, to gainful employment. If the individual has resumed working, decreased likelihood of promotions or ability to compete for higher paying jobs can be considered. In the case of younger persons, projections might involve the type of jobs the individual may have held in comparison to post-injury work capacities. In principle, damages should flow from the injury. An individual who previously had a poor marital relationship should not recover for the marital problems that pre-dated the event or were caused by other factors.

The above discussion should make it clear that individuals are generally not supposed to be compensated for having a condition per se, but from the consequences, problems, suffering, or dysfunction associated with the condition. For example, if two individuals both have Post-Traumatic Stress Disorder (PTSD), but one is usually able to control her symptoms, needs limited treatment, and is functioning well at work, she is likely to be compensated much less than the individual who has broad impairment and cannot hold competitive employment or maintain intimate relationships. This emphasis on the consequences of injury is one reason it is so important for an evaluator to go beyond a label and the immediate office setting and try to appraise, if scientifically viable methods are available, an individual's life functioning (see further below). Failure to do so can result in a serious over- or underestimation of damages. For example, an individual with orbital frontal brain damage may seem relatively unremarkable on cognitive testing and structured interview but may exhibit incapacitating problems in everyday behavior, whereas another person with markedly elevated scores on personality testing and many complaints may be functioning proficiently.

Although it may fall within the psychologist's purview to describe areas of positive and negative functioning and to formulate projections over time, certain technical aspects of damage appraisal, especially their conversion to monetary values, are usually better left to others (which does not necessarily or always mean that one endorses their methods). For example, life care planners may list the equipment a brain-injured individual needs and the cost of various items, or an economist may translate conclusions about occupational capacity to numbers. Psychologists should be very careful not to undertake tasks for which they are not well qualified or knowledgeable or for which underlying methodology is dubious. As will be discussed, effective cross-examiners have refined abilities to detect weak points and to start tearing at these very locations. The credibility of an otherwise exemplary evaluation may be quickly demolished when the lawyer starts asking the psychologist about the grades he received in business classes or technical questions about calculating interest rates and cost of living increases. It often only takes one or a few instances in which an expert has been shown to express opinions in areas she knows little about before the jury assumes the expert does not know much about anything.

Also, although I will not repeat the same point over and over throughout the chapter and instead will mainly emphasize practical consequences, there are compelling ethical reasons to restrain practice in relation to level of expertise and the availability of adequate scientific methodology/backing. Parties have much to gain and lose in legal proceedings, and the hope and expectation is that the expert in a branch of science will assist juries in reaching sounder decisions than they would otherwise, something that demands some reasonable level of knowledge and scientific underpinning.

A psychologist need not go to law school to become an effective forensic examiner, but should try to understand potential differences in the clinical and legal arenas, including the basic rules or principles of procedure and just what the issues are that define the scope of the expert's involvement and task. For example, in the clinical arena: (a) we are concerned first and foremost with the client's best interests (within the bounds of ethics) and we hope the client perceives us (or comes to perceive us) as such; (b) we usually assume that clients are not motivated to deceive us for purposes of self-gain, potentially at cost to innocent parties, and we usually do not obtain collateral records specifically for the purpose of checking on their veracity; (c) we may not be particularly concerned about cause, especially if we have narrowed possibilities down to the point that the alternatives do not change treatment; and, as follows, (d) we tailor our assessment to clinical needs and treatment questions.

In contrast, in the legal arena: (a) we presumably try to render an objective opinion, even if it might conflict with the examinee's self-interests, and we (should) assume that the examinee realizes we are a potential adversary (and ought to inform him of such); (b) we recognize that examinees might be motivated to deceive us and that we need to take active steps to check on their credence; (c) we often must focus on cause and may attempt relatively fine or subtle distinctions that may not impact on treatment at all; and, as follows, (d) we need to address, where appropriate and possible, key legal questions, such as everyday functional capacities, as opposed to what might otherwise be our main concerns (e.g., subjective beliefs and perceptions).

As is also evident, in a clinical context we do not typically expect our ideas or conclusions to be subjected to cross-examination designed to counter our view or to damage our believability or character; we might feel free to engage in hypothesizing and conjecture when trying to achieve greater understanding of our patient; and we do not expect the examinee to suffer potentially devastating consequences even if his claim is worthy but we cannot defend our opinions in court. Legal evaluators try to keep in mind the eventual uses of, and challenges to, their work. They apply this awareness to guide their efforts from the start, one hopes in the direction of performing high quality work that addresses the legally relevant issues and respects the bounds of scientifically sound method and opinion.

Finally, with all this said, juries' decisions may overlook technicalities that, in theory, should be decisive or at least highly influential. For example, although in principle the plaintiff usually must prove his case (the standard often being "more likely than not" or "by a preponderance of the evidence"), other considerations may prevail. For example, in one case, jurors interviewed after the trial agreed that they thought the plaintiff probably was not brain damaged, but that in case he might be they wanted to give him enough money to pursue the treatment he might require. In other situations, one or the other side may be perceived in such a negative light that judgments that should not be swayed by such reactions may nevertheless be altered. For example, the jury may be overly generous in deciding fault in favor of the defendant because they found her honest and likeable and thought that the plaintiff was an undeserving, devious individual. An expert who fares poorly or who is perceived as dishonest can in fact do considerable damage to her own side on issues that might seem only remotely related because she may cast doubt on others, such as the lawyer who retained her.

4.19.2.3 Treating Professionals Called into Court

4.19.2.3.1 Maintaining the treatment alliance as primary

My comments distinguishing the clinical and legal contexts assume, in the latter instance, that the psychologist has been asked to perform a legal assessment and knows from the start that this is the task. Treating psychologists may be called into court and, having served until then in a clinical role, often cannot be expected to have approached the case in a manner, or to have taken the steps, expected of a legal examiner. Further, the psychologist might find certain expectations for legal examinations (e.g., taking little or nothing the patient says for granted and performing careful checks on credibility) contrary to the role of treater. Psychologists who testify in such circumstances usually are best served by openly acknowledging any limits in their methods. Any decrement in credibility associated with such limits may be more than offset by the believability often accorded to treating professionals as opposed to experts hired by one or the other side, and by the jury's relative intolerance for a lawyer who is overly aggressive with a treater, especially one who is likeable and candidly discusses shortcomings. For example, if the lawyer criticizes the psychologist for failing to obtain extensive background information, the psychologist might comment that were he conducting a legal assessment he would have sought this information, and that it might have been useful, but when the patient presented to him with suicidal ideation he was not worried about how he would look in court but about preventing a tragedy. In contrast, if the psychologist becomes defensive and insists the lack of background information is irrelevant or that such records could not possibly change his opinion, his credibility is likely to be damaged and he might no longer be perceived as the type of doctor the jurors themselves would seek out for treatment.

Treating psychologists may want to try to avoid testifying altogether, at least in cases in which they are concerned about damaging the therapeutic relationship. The specter of court can also compromise treatment relationships from the beginning. For example, parents who are considering divorce may be more concerned with impressing the psychologist favorably, just in case a custody suit eventuates and they need a helpful opinion, than about solving their problems.

When initiating assessment or beginning therapy with possible courtroom involvement looming in the future, and especially when circumstances are such that it threatens to impede progress, it can be helpful to take an immediate position that elevates the therapeutic role as primary and minimizes, to the extent possible, the likelihood that the present clinician will impact on any subsequent legal proceedings. A psychologist cannot assure a patient she will not be called on to testify in court and should think twice before assuring the patient that, even if called, she will not comply (because this could lead to serious sanctions, such as jail time). However, the therapist can say, with all sincerity, that she considers the treatment role to be primary. She can indicate that if asked to serve as a potential witness, she will tell the attorney (or either attorney) that she has not attempted to form an objective or detached opinion and is likely to be biased or strongly influenced by the role as treater, that is, that she does not have, nor has she tried to develop, the type of impartiality desired of a testifying expert. She would therefore advise the attorney to retain an independent expert to perform an evaluation and address the issues relevant to the courtroom. Some potential clients, who, in fact, are primarily interested in using the clinician for courtroom leverage may abandon treatment immediately.

If the treating psychologist conveys such a therapeutically-oriented position to the attorney, he may well retain a separate expert. Attorneys have strong reservations, if not terror, about calling witnesses that are supposedly there to support their position but that they cannot control or predict. The damage caused by an adverse opinion is usually much greater if it comes from the attorney's own witness as opposed to that of the other side. The attorney may still call the therapist, not to state expert opinions, but simply to describe facts about such things as entries in the chart or treatment costs to date. Of course, it is difficult for the treating psychologist to arrive at a satisfactory solution sometimes, because a patient, although initially endorsing the psychologist's position of neutrality, ends up feeling disappointed or betrayed when the moment of truth arrives and abstraction becomes reality. Clients also need to realize that the therapist's commitment to treatment over legal involvement is not *carte blanche*, because certain behaviors or actions (e.g., child abuse) would not only likely alter obligations but are also associated with reporting requirements.

4.19.2.3.2 Release of assessment or treatment records

Although a series of generalizations can be provided about the release of records, only limited specifics are possible given both restrictions in the author's knowledge and the complexities and idiosyncrasies that often arise. For example, situations may occur regarding minors in state custody and who has the authority to release records; special questions may be raised when the mental competence of the patient is at issue, therapists may feel a strong obligation to review records in detail with former or present patients before releasing them; psychologists may be concerned about guarding the security of test materials, the clinician may be unsure what constitutes his file (e.g., does this also include records from other providers), or records may make references to others that are not party to a suit. Malpractice cases raise an additional set of concerns; for additional details and guidelines, the reader is referred to Bennett, Bryant, VandenBos, and Greenwood's helpful book (1990).

Some psychologists believe they can assert a right of confidentiality when their records are subpoenaed in legal cases, especially if they or their patients never anticipated courtroom involvement, the legal issues seem indirectly or minimally related to the therapeutic work, and the file contains personally sensitive materials. On the contrary, confidentiality is a right of the patient, not the therapist. If the individual wants his records released (and especially if he is aware of the contents), and if there is no reason to believe the person's decision-making powers are seriously impaired, the therapist is obligated to comply. (Issues related to collateral records that the therapist has not generated or copyrighted materials, such as test items, can become complex and will not be taken up here.) Further, in civil cases in which an individual has placed his mental status at issue, the confidentiality of past or current therapeutic materials is likely to be waived, no matter what the client prefers (unless she drops her case or at least certain damage claims). There are exceptions, however. For example, some states have exceptions for alcohol or drug treatment, and juvenile records are often protected. Further, records may contain very sensitive materials that seem irrelevant to the case at hand but that could be terribly embarrassing to the client or might damage her public image and occupational endeavors. In such situations, the responsibility usually resides with the plaintiff's attorney to protest the release, or use, of the materials, although the keeper of the records (i.e., the therapist) may need to alert the appropriate party to the presence of sensitive materials in the file. A judge might choose to review the records privately (in camera), and then decide whether to release all, none, or parts of the material. As a general guide, the psychologist should not release records unless he has the consent of the patient or a court order (not just a subpoena) demanding their release, although one should not simply ignore a subpoena (see further below).

It is mistaken to assure a patient when beginning an assessment or therapy that she has an absolute right to confidentiality or that there are only a few circumstances in which records can be obtained by third parties. There are, in fact, many exceptions to confidentiality in most states, often over a dozen, and patients may end up disclosing materials that damage their courtroom cases or reputations that they otherwise would have withheld had they been properly informed. For example, an individual in a custody case may not want to tell the therapist about occasional wild sexual fantasies that he can easily resist and would never pursue.

A few additional points can be noted. First, the psychologist should not just ignore a subpoena, and certainly not a court order. This does not mean that records must be released immediately, but if the psychologist is going to resist doing so or needs more information before acting, he needs to communicate the basis for the initial noncompliance to the relevant party in a timely and proper manner. In such cases it might be wise to consult an attorney. Second, it is usually wise to assume that treatment notes can be obtained in a legal case, or at least to recognize the real possibility that some other party might gain possession of the psychologist's records. Third, do not alter records. In a malpractice case, for example, it is not only obviously unethical and illegal to alter records, but it is often fatal to the case. Some psychologists seem naive about the mechanisms, and sometimes relative ease, with which altered records can be detected. Altered records demolish the clinician's credibility, and once this occurs the case is often, in effect, over. Fourth, if in doubt, contact an attorney, or at least another professional who is knowledgeable about legal matters. Issues relating to the release of records can quickly become complex, the adversarial nature of courtroom proceedings may be foreign to the professional, and the potential consequences of improper decisions can be serious, thereby calling for caution.

4.19.3 ADMISSIBILITY

The issue of admissibility, or what is allowed into evidence, has occupied many legal scholars and has many facets that often become extremely intricate. The issues that will involve us here is when, or whether, psychologists are allowed to testify as experts, and the extent to which their testimony might be constrained. To clarify the latter point, although a psychologist may be allowed to testify, the scope of testimony might be restricted. For example, in a criminal case, she might be allowed to address issues related to symptoms of PTSD that she believes the claimed victim of an assault manifests. However, she might not be allowed to address whether the examinee's presentation fits expectations for victims of violent rape, as such testimony might be seen as unreliable and as invading the province of the jury to decide matters of fact or ultimate issues. Various types of experts with various types of credentials testify in court. Psychological testimony is usually considered to fall within a branch of science, and thus the applicable standards are those relating to the admission of scientific evidence.

Standards for admissibility often vary across federal and state courts, and from state to state, and judges may vary considerably in the way standards are applied. What one court and one judge lets in, another might bar. There is little doubt, however, that the Supreme Court's recent ruling in Daubert v. Merrell Dow (1993), which involved the admissibility of scientific evidence, is having considerable impact. Many states are guided, in large or substantial part, by these new Federal guidelines. Briefly, prior to Daubert, admissibility of scientific evidence in Federal court was decided by two sets of standards: Frye (1923), which emphasized such considerations as acceptance within the scientific community, and a separate set of Federal standards. In Daubert, the court rejected Frye, elevated the Federal standards, and further elaborated upon them. Although the exact impact and interpretation of the Court's ruling will be clarified gradually as more and more cases are decided, at least two relatively clear trends have emerged.

First, Daubert explicitly acknowledges the judge's role as gatekeeper in deciding whether or not to admit scientific evidence. This means, in essence, that judges will likely feel freer to exclude expert evidence on scientific (or purportedly scientific) matters with less concern about reversal.

Second, in a number of cases, Daubert has lead to the exclusion of evidence that might otherwise have been admitted under Frye, particularly in instances in which scientific foundations are weak (see Bersoff, 1997). In deciding whether to admit evidence and in evaluating scientific status, Daubert directs the judge to consider such matters as demonstrations (or lack thereof) of scientific validity, whether findings have been published in peer-reviewed journals, and whether the method has a known error rate. None of these considerations are necessarily dispositive, and the court did not attempt to create an exhaustive list of criteria nor to place them in hierarchical order. Thus, Daubert is very clear in situations in which all scientific indicators are either positive or negative, and increasingly incomplete in its guidance when indicators conflict (although the same basic point could be made for methods and approaches that practicing scientists and philosophers of science use to try to settle scientific disputes—see Faust & Meehl, 1992). As would be expected in such a circumstance, there has been, and undoubtedly will continue to be, considerable variation in the application and interpretation of criteria for deciding whether evidence meets the Daubert test. However, the weaker the scientific support, especially in extreme cases, the greater the likelihood post-Daubert that evidence or testimony will be excluded. Consequently, in cases in which an expert, when challenged, cannot cite any decent scientific support for his opinion or assessment methods (e.g., there are no published studies), there is a very real chance that the psychologist will be prohibited from testifying. Overall, Daubert has reinvigorated challenges to the admission of scientific testimony, the potential testimony of mental health professionals and others is being placed under much greater scrutiny, there are an increasing number of decisions in which testimony has been excluded, and some reduction of junk science in the courtroom seems to be occurring.

What does all of this mean for the mental health professional? In an increasing number of situations, especially when the scientific foundation for testimony is weak, experts can expect challenges to the admission of evidence. In Federal court and in an increasing number of state courts, there is now more than a remote chance that testimony may be limited or excluded if there is minimal scientific support for the expert's opinions or methods, or if the literature is predominantly negative. Testimony might be allowed in areas in which scientific backing is stronger, but if foundations are weak across the board the psychologist may be barred from testifying entirely. Most commonly, an objection to the admission of testimony takes place before the trial. Written documents are submitted and, depending on the judge's discretion, there may be a pre-trial hearing at which arguments are heard and experts might testify.

In some areas there seems to be little question that psychological method or knowledge meets the Daubert standard. For example, many statistical methods have sound scientific foundations. There is also a good deal known about the potential impact of mild to severe brain injuries, at least in general (as opposed to the difficulties that might be involved in assessing specific impact in individual cases). In other areas, however, psychologists may be very susceptible to Daubert challenges. For example, many neuropsychologists construct their own idiosyncratic batteries, and there may be no peer-reviewed studies on the effectiveness of these batteries as a whole. Predictions of long-term outcomes may also lack scientific foundations, or studies of their accuracy may have produced consistently negative results.

Although a psychologist might view Daubert as an unenlightened or unjustified restriction, the hurdle that it creates—that evidence purported to be scientific should have a reasonable scientific basis—hardly seems outlandish. In fact, standards of this type would seem to be something that organized psychology, with its stated commitment to science, could support. Such standards might ultimately improve psychology's position with the courts, given the methodological sophistication of many members of the field and the potential to build or extend strong scientific foundations in key areas of legal interest. For now, a psychologist might want to think carefully before agreeing to perform an assessment or provide testimony in areas with weak scientific underpinnings. Of course, this same argument certainly could be made whether or not there is likely to be a Daubert challenge. I would note in this context that although psychologists might believe that experience can partly or fully compensate for scientific shortcomings, even serious ones, the extensive literature on experience and accuracy raises major questions about this conclusion. In fact, the negative results of many studies on experience and accuracy in the mental health field have been noted by diverse individuals (e.g., Faust, 1984; Garb, 1989; Brodsky, in press, the latter being a revision of his earlier views [1991]).

Even if the courts allow testimony on some topic, a psychologist may feel that the state of knowledge is insufficient to develop trustworthy opinions and thus may opt not to be involved. Just as the courts may err in excluding testimony, they may err in admitting it; and it is a dubious position for a professional in a branch of science to assert that so long as it's good enough for the courts it is good enough for me. Given psychologists' methodological knowledge, situations certainly can arise in which the professional believes that a judge has overestimated the scientific standing of a test, method, or assertion and that, instead, knowledge is very shaky or dubious. Most psychologists' courtroom involvement is voluntary, and there is usually no external authority that compels the psychologist to testify if she does not wish to do so, or that will invoke official sanctions for nonparticipation. In these situations, psychologists can apply internal or professional standards in deciding on their courtroom involvement. (For a more extensive discussion of dimensions a professional might consider when determining a method's readiness for the courtroom, see Faust, 1993).

The court does not decide admissibility solely on matters relating to science. In addition, testimony must be relevant to the issue at hand. For example, a psychologist might be a first-rate authority on test bias, but the topic might not arise in the case, or it might be so secondary or remote to any of the matters in dispute that the judge excludes the expert. Additionally, the expert must contribute something over and above what the jury knows or can determine on its own. For example, if the defendant's left foot has six toes and webbed feet and a plaster cast matches the pattern exactly, the jury probably does not need a footprint expert to provide commentary on the resemblance.

Experts also need to have adequate credentials. Usually, before opinions are expressed about issues in the case, the lawyer has the psychologist recite her credentials and then offers (proffers) the witness as an expert in some area (e.g., clinical psychology or PTSD). The expert should have credentials to support the offer. Psychologists may not realize that the standards judges apply when deciding whether to qualify the expert are often relatively minimal (e.g., a Ph.D. degree and some education and training in the area) rather than lofty, or that jurors may have difficulty separating, or care little about, differences in credentials that might strike the professional as weighty. A few publications might seem about as good as 20 to a jury, and a Ph.D. is likely to be an impressive degree whether or not it came from a big-name university. Ironically, some experts create major problems for themselves by failing to represent their credentials with a high degree of accuracy, when the difference often would matter little to a juror or to a judge evaluating qualifications. For example, an expert might exaggerate the number of patients seen with a particular problem, or might represent himself as a graduate of an APA-approved training program when, in fact, the program was accredited after the time of graduation and the listing is consequently inappropriate.

Take the last example. If the credential is listed accurately, the psychologist might be able to indicate, perhaps during his direct, that his program gained APA approval shortly after he graduated. The psychologist could further indicate that although the APA, in essence, approved the same program he attended, it would be technically incorrect for him to list himself a graduate of an APA-approved program, and he would not want to create a false impression. A juror is unlikely to give this a second thought, and it might even be a plus as it conveys honesty. In contrast, the following cross-examination illustration shows what can happen with even a seemingly small misrepresentation:

> Question (Q): Doctor, is there anything you'd like to correct in your direct testimony before I start asking you other questions?
> Answer (A): No.
> Q: Doctor, is there anything about your credentials you may have stated in error?
> A: No, not to my awareness.
> Q: Doctor, do you remember raising your right hand before you started to testify?
> A: Of course.
> Q: And during that oath, you gave your word that you would tell the whole truth, isn't that correct?
> A: Yes, I did.
> Q: And doctor, you told us, did you not, that you were a graduate of an APA-approved program. Wasn't that your testimony?
> A: Yes.
> Q: That was a misrepresentation, wasn't it?
> A: Absolutely not.
> A: And your status as a graduate of an APA-approved program, that's the same representation contained in your resume, isn't it?
> A: Yes.
> Q: And this is the resume you send to lawyers, and have presented in court many times, isn't that true?
> A: Well, I don't know about many times, but it is the information I have provided.
> Q: It is the same information you would present to patients, would you not, if they asked you whether you came from an APA-approved training program?
> A: I'm not sure a patient would ask such a question.
> Q: Doctor, I'm sorry if my question was unclear to you. If asked by a patient, you would tell them that you were a graduate of an APA-approved program, correct?
> A: Yes.
> Q: Then let me ask you one more time. You are telling the ladies and gentlemen of the jury that it is true, without qualification, that you are a graduate of an APA-approved program. You are giving us your word, is that correct, just like you promised to be honest about the other areas in which you testified?
> A: Yes.
> Q: And, doctor, you graduated in 1985, isn't that the case?
> A: Yes.

The lawyer can then produce an issue of the American Psychologist that includes the section indicating that a psychologist cannot describe his- or herself as a graduate of an APA-approved program unless the program was accredited at the time of graduation. Things can almost only go on a disastrous downhill course from there. For example:

> Q: Having read that, and seeing the listing for your program and the date of approval as 1990, the truth is you are not a graduate of an APA-approved program, isn't that the case?
> A: The listing is probably in error.
> Q: Is that right doctor? I'll tell you what, we can check on the listings from other years of the American Psychologist, which I have on

> the table here, and the sworn statement I've also obtained from the director of your graduate program. Do you think that will help clarify the truth of the matter?

A misrepresentation of this type can be fatal in not only the instant case, but in future cases. For example, in a subsequent case, a lawyer might ask the expert whether she has been busy calling all of the judges, lawyers, and patients with whom she's been involved to inform them that she misrepresented her credentials.

4.19.4 ASSESSMENT METHODS

In this section I will present various "tips" for performing legal assessments. Many of these guides apply similarly to conducting quality clinical assessments, although there are also a number of features distinct to forensic work.

4.19.4.1 Informed Consent

The fact that an evaluation is being performed for legal purposes typically does not alter the need to obtain informed consent before proceeding. The nature and purpose of the assessment should be presented, and it should be made clear that the clinician's primary obligation is not to advance the examinee's interests but to address certain questions, whether or not the conclusions help or hurt the individual's case. The expert should indicate who retained him, such as plaintiff or defense counsel, or judge. The expert should explain that his role is to perform an evaluation and not to provide treatment, and thus that there is no doctor-patient relationship. (This certainly does not mean the psychologist should not make treatment recommendations she deems appropriate or helpful.) At this point, the examinee may not consent to the procedure, which, in civil litigation, and often in criminal cases, is his perfect right.

There are cases in which experts are retained to perform evaluations, but with the understanding that they themselves may undertake treatment if they deem it appropriate. For example, the plaintiff's attorney may retain an expert she esteems to perform a neuropsychological evaluation of head trauma and, as needed, to provide rehabilitative services. The involvement in the dual role of forensic examiner and treater can create complications. For example, once engaged in a therapeutic relationship, it may be difficult, when it comes time to testify, to be as objective as one otherwise might be or to express views that will likely harm the plaintiff's case. Also, challenges can be raised about a financial incentive to diagnose disorder in order to engage the individual in treatment for which one can charge. For these and other reasons, it is probably best, when possible, to avoid serving as both a forensic examiner and treating psychologist.

4.19.4.2 Clarifying Referral Questions and Deciding Whether they are Appropriate and Within the Clinician's Expertise

Lawyers usually make initial contact with an expert on a specific case by written correspondence or phone. Such communication will commonly eventuate in a request to perform an evaluation and perhaps some commentary about the questions the lawyer wants the expert to consider. At times, such requests are vague. For example, the lawyer might ask the expert to evaluate psychological status but not even specify whether she is primarily interested in cognitive or emotional status, or both, or neither.

Vague questions may reflect the lawyer's lack of familiarity with issues in the mental health field, and the expert should attempt to clarify just what is at issue in the case, such as the possibility of brain damage, emotional disturbance in response to a traumatic event, present and future work capacity, malingering of mental disorder, or whatever. The lawyer may be unsure about what a psychologist, or the expert in question, does and does not do and can and cannot do; and the psychologist may need to gain some familiarity with the file before a more productive conversation is possible or greater clarification can be achieved. For example, the lawyer may not realize that psychologists cannot prescribe medicine, that not all clinical psychologists are trained to evaluate neuropsychological status, or that many psychologists are not fully prepared to evaluate members of minority groups. Or, they might not realize that psychologists often provide treatment services, that many clinical psychologists are very familiar with various matters involving psychotherapy, or that the condition in question has a grave prognosis and that this issue would seem highly relevant to the plaintiff's case. It is not that the psychologist should do the attorney's job, but rather that the psychologist's input may be needed to hone in on relevant issues in the case or what is and is not possible to accomplish through a psychological evaluation.

Referral questions also need to be unobtuse so that the clinician can determine whether the required scientific foundations exist to provide informed answers, and whether she has the requisite expertise. It is not only ethically questionable to take on issues or questions for which one lacks the needed knowledge, but it can easily lead to courtroom fiascoes. It is rather unnerving to have to admit on cross-examination in open court that this is only the second case of Fisbee's disease one has seen and that almost all of one's scientific knowledge on the topic was acquired through concentrated reading during the last week. This from the same expert who has contended, in effect, that the jury should defer to him over Dr. Jones, who has written the seminal work on the disorder and has acquired exhaustive knowledge about the condition through years of study. The psychologist is also likely to embarrass herself if a well-prepared and knowledgeable lawyer begins to ask about specifics. For example, someone who attempts to acquire quick knowledge about neuropsychology is unlikely to know what the abbreviations HEENT and GCS in the emergency room record mean, or may startle the jury when the response to the inquiry, "How many journal articles have you read devoted exclusively to mild head injury in the last 5 years," is, "None." There is a saying in aviation that there are old pilots, and there are bold pilots, but there are few old, bold pilots. This might be remembered as one engages in internal debate about whether or not to take a case. The psychologist might get by with questionable work for some time, but a single encounter with a skilled and well-prepared lawyer can inflict serious, long-term damage.

4.19.4.3 Access to Information

As will be seen, the quality of legal evaluations, and how well they stand up under scrutiny, often depend on the thoroughness of information gathering. Further, the time at which information is received can matter a great deal. For example, a psychologist who forms an opinion rapidly and in the face of inadequate information and who never revises his views in the slightest way, even after obtaining a great deal of additional information, may not seem credible. Also, a jury might look askance at an expert who reaches opinions and perhaps issues treatment recommendations, or even begins treatment, well before much of the information has been obtained, especially if that information was readily available. For example, many experts, who may have completed their reports a year or more earlier, indicate that they received or reviewed additional key information the morning of the deposition or the night before their courtroom appearance. It might be difficult for the jury to believe that the psychologist could be completely dispassionate in her use and weighting of the additional information under such circumstances.

Further, given the uncertainties and ambiguities that often exist in our field, it would be surprising if considerable amounts of new information did not contain at least some evidence contrary to our conclusions, or even evidence that required new areas of inquiry or raised serious questions about the initial opinion. Receiving such information at a late date can therefore create major difficulties on the stand. For example, the expert might be asked a line of questions like the following:

> Q: Don't you agree that mild head injury and the effects of serious alcohol abuse can look similar on cognitive testing? (If the expert disagrees, a statement he may have made in his deposition or in other cases confirming this assertion can be raised.)
> Q: And didn't you describe the level of drinking that can result in abnormal performance on cognitive testing?
> Q: And didn't you indicate that one of the ways you ruled out alcohol as a factor in Mr. Smith's testing results was his telling you that he had almost never touched a drop in his life?
> Q: Until yesterday, when attorney Jones gave you the records, you were unaware of Mr. Smith's previous treatment for alcohol abuse, weren't you?
> Q: So, when you formed your opinion and issued your report, you had not seen the treatment records describing Mr. Smith's history of alcohol abuse, had you?
> Q: Surely this is the type of information you would have preferred to have known about when you conducted your evaluation, isn't that correct? (The lawyer is unlikely to care much, and may even welcome it, if the psychologist wants to fight her on this point, because it will probably seem unreasonable to the jury and further compromise the expert's standing.)
> Q: And had you had this information, you would have asked additional questions about substance abuse, wouldn't you?
> Q: You would have asked questions to find out whether Mr. Smith drank to the point that you would expect abnormalities on the testing, isn't that right?
> Q: But it's too late to ask those questions now, isn't it?

An expert who denies the import of such questions may lose all credibility with the jury.

Many lawyers will not know what records a psychologist needs to perform a proper evaluation (see more on this below). Thus, it is up to the psychologist to inform the lawyer. The mechanics for obtaining information (as opposed to a decision about what information to obtain) can be worked out in different ways, although it is often appropriate for the lawyer to do the needed leg work. It can be very difficult or impossible to obtain certain types of information (e.g., earlier school records in the case of an elderly individual), but it may become obvious that a lawyer is not really motivated to carry out the psychologist's requests and rather wishes to control the flow of information. This reluctance and attitude may severely compromise the quality of the evaluation and place the psychologist in a vulnerable position. For example, on deposition, an expert might be asked about the type of information she routinely procures, or prefers to have, when performing a comparable evaluation. Procedures followed in similar cases in the past might be raised. Discrepancies between the usual and present case can be pointed out, and the expert can then be asked to explain the reasons why the current evaluation is so much less complete, the extent to which gaps in information compromise the trustworthiness of the conclusions, and who has controlled the flow of information. A careless response to the effect that such extensive information is unnecessary can lead to such questions as, "Are you telling me that when you gathered far more extensive information in your many other cases and charged thousands of dollars for the totality of your time reviewing those materials, all that was unnecessary?" An expert who is true to the oath and admits that the lawyer exercised control might be asked whether lawyers taught their assessment courses in graduate school or whether it is a routine procedure when performing a psychological evaluation to consult a lawyer in order to determine what information the professional needs.

An expert might also contemplate whether she wants to be involved with a lawyer who tries to exercise this type of control. There is a big difference between a lawyer pointing out legal technicalities that fall within his purview (e.g., what standards of evidence will be applied or what exactly is being disputed in the case), and one who tries to control how the expert performs activities that fall within the psychologist's professional domain. One way to handle possible differences in viewpoint on such matters is to tell the lawyer that one will feel compelled to express reservations about one's conclusions given the present restrictions the lawyer suggests. Both parties might then make more informed decisions about course of action, including whether to part ways.

This does not mean that every difference, even a very small one, necessitates an all-or-nothing stance, that the lawyer will never have a compelling counterargument, or that all such differences reflect a lawyer trying to control the expert. For example, a record a psychologist might ideally like to have may really be very remote to the issues in dispute and extremely difficult and expensive to obtain. Or, due to time limits imposed by the court, a psychologist may have to decide whether to perform an evaluation before all of the desired information is obtained. The psychologist might decide to go ahead, but to be sure the report contains an explicit statement describing the circumstances and indicating that opinions are subject to change or elaboration as more is learned. However, psychologists who take the long view probably should err on the side of caution. There are certainly some attorneys who wish to control access to information in order to increase the chances of a desired result, or who plainly want to manipulate the outcome. A psychologist will not only be on safer ethical ground but likely to have a much longer career as a forensic evaluator by avoiding such lawyers.

4.19.4.4 Design of Assessment Procedures: Some Do's and Don't Do's

4.19.4.4.1 Use the best available methods

It is important for psychologists to use the best available methods, not only to maximize the probability of reaching accurate conclusions, but because their work may be placed under intense scrutiny in legal proceedings. For example, the psychologist may be asked a series of pointed and specific questions about technical issues. Good attorneys can have a remarkable ability to pick up quasi-expertise rapidly, retain it, and generalize it to new contexts. An occasional attorney will ask questions about such matters as false-positive rates, standard error of measurement, or criterion-based validity with a proficiency that can shock an unprepared expert. Attorneys may also retain consultants to scrutinize records, educate them about technical matters, and provide them with background literature on the methods used in the case in order to better prepare for depositions and cross-examination.

Psychologists may use tests with poor normative information when better choices are available, or may use old or obsolete norms when more contemporary or complete normative information exists on the same test. For example, many neuropsychologists still use earlier norms for the Halstead-Reitan Battery (Reitan & Wolfson, 1993), which are decades old and which show a marked propensity to overdiagnose certain groups of individuals, such as older and less educated persons. More recent norms of the type that Heaton, Matthews, and Grant (1991) developed, which adjust for age, education, and gender, can decrease overdiagnosis considerably.

Additionally, norms developed for one group of individuals may not be wholly applicable or may create major problems when applied to other groups of individuals. For example, norms for measures of motor speed developed with young individuals may be applied to the elderly, or norms developed among members of the dominant culture in the USA may be applied to recent immigrants with very limited English proficiency. Psychologists may be unaware of specialized normative data that have been developed for differing sociodemographic groups or, when such information is unavailable, may act as if the application of norms developed with one group to another group with distinct differences could not possibly raise any particular concerns or issues. When these types of problematic practices occur, lawyers can introduce materials from such sources as the Standards for Educational and Psychological Testing (APA, 1985), which describe the need for normative data appropriate to the individual under consideration.

For a substantial number of psychological tests, there are inconsistent sets of norms that may lead to very different interpretations. For example, depending on the norms used for the Halstead-Reitan Battery, the number of errors on the Category Test that falls two standard deviations beyond the mean may be less than 50 or more than 100 (see Faust et al., 1991). In the case of the Auditory Verbal Learning Test (AVLT) (Rey, 1964), a list learning task, many psychologists rely on the original norms developed many years ago in France with relatively small samples. There have been many subsequent studies showing that those norms are too demanding for many groups (e.g., Bolla-Wilson & Bleecker, 1986; Wiens, Crossen, & McMinn, 1988). A psychologist may be asked questions like the following on this topic:

> Q: You told us that the plaintiff had problems on the Smith Memory Test, isn't that correct?
> Q: And the Smith test is the one with the list of words, like a grocery list, correct?
> Q: The standards you used to decide what was normal for the Smith Memory Test were developed more than 50 years ago, isn't that correct?
> Q: Those standards were developed in another country where they speak a language other than English, isn't that correct?
> Q: And there have been many studies published in the United States criticizing the use of those standards with today's Americans, isn't that true?
> Q: There are more than 10 recent studies that would classify Mr. Smith's very same performance as perfectly normal, isn't that true?

At times, it is clear that one set of norms has been better developed than another, or it is obvious that the findings from one study deviate markedly from those of most or all other studies. Unfortunately, at other times, the underlying basis for conflicting norms, or the most appropriate choice among competing norms, is unclear, creating something of an impasse. A psychologist might consider the merits (or lack there of) of using such tests, especially if tests with a better or clearer normative base are available that are designed to assess comparable areas.

Lawyers may also raise questions about other technical qualities of tests, such as reliability and validity. Some psychologists may see such questioning as meddlesome or something akin to badgering. However, standing on these technical qualities can be all important in assessing the likelihood of an accurate result, they are certainly fair game, and they often do need to be critically appraised. A psychologist may be surprised about the occasional lawyer's sophistication on these topics. For example, on deposition, one might be asked:

> Q: In the language of your field, test-retest reliability refers to the stability of test scores over time, isn't that true? (If the expert is evasive on this and other questions, the attorney may call her own testing expert to provide appropriate and clear answers and to show, in effect, that the opposing expert was being less than genuine.)
> Q: Isn't it a basic tenet in your field that tests that lack satisfactory reliability will also lack satisfactory validity or accuracy?
> Q: Test-test reliability is often measured by the correlation coefficient, isn't that true?
> Q: Correlation coefficients can range from .00 to plus or minus 1.00, isn't that correct?
> Q: And in general, higher correlation coefficients are better, or indicate a higher level of reliability, isn't that correct?
> Q: Isn't it true that (a prominent psychologist in the area of measurement will be named here) considers a test-retest reliability of .80 to represent a minimal standard? (If the expert does not acknowledge the author or source, there are other ways to get at the same thing, such as by asking what authorities on testing the expert respects, or about texts and

journal articles that were used in the expert's education and training.)

Q: You've not published your own text on psychological testing, have you?

At trial, after reintroducing these topics in clear and understandable language for the jury, the lawyer may proceed to show that various tests the psychologists used do not meet the standards for reliability the witness affirmed and thus, by the witness's own reckoning, cannot be trusted.

Of course, issues related to reliability are important not only because they can lead to courtroom embarrassment, but because adequate reliability is often necessary for reaching accurate conclusions. For example, deficient reliability can impede or cripple interpretive strategies that psychologists often consider essential, such as comparisons between test scores and pattern analysis. If reliability is low, the standard error of measurement of the difference $(\mathrm{SEM})_{\mathrm{diff}}$ may be so great that there are likely to be frequent false-negative and false-positive judgments about whether true contrasts exist across tests. When comparing test scores, and especially when analyzing test score patterns, error is usually additive. Stated differently, the level of measurement error will exceed, sometimes by a great margin, the average level of error per test (with the margin expanding rapidly as the number of scores considered together as part of pattern analysis grows).

Test validity is not a global quality, and hence there is little meaning to a statement like, "The Fisbee Test is highly valid." Rather, validity is a specific quality, and a test that is highly valid for one purpose and with a certain population may lack validity for other purposes and with other populations. That is, validity refers to the interpretations given to test scores and not to the test itself. Therefore, when selecting tests, it is important to go beyond general information and to become familiar with research on specifics, in particular the specifics involved in the current application.

Particularized questions about validity are often essential to legal work because many of the issues psychological tests were originally designed to answer are different, in subtle or not so subtle ways, from those commonly asked in the courtroom. For example, in a criminal case, it may be of much greater interest to determine an individual's mental state at a previous point in time rather than in the present. The possibility of malingering and the need to evaluate for it are often considerably greater in the legal context. The differences that can exist in the purposes of traditional psychological tests and the questions that arise in the courtroom have led a number of psychologists to develop instruments specifically designed to address legal issues, such as competency to stand trial (see Grisso, 1986) and malingering of psychological disorder (see Rogers, 1997a). Psychologists who conduct legal evaluations, especially in areas that traditional psychological tests are not designed to appraise and for which there is little or no research on these measures, might be well advised to become familiar with specialized methods, if they are not so already.

Along related lines, although one test may have usually outperformed another in comparative studies, important exceptions may exist in specific areas. For example, although Test A may be generally better than Test B, there may be much more extensive or positive research on the capacity of Test B to distinguish between the effects of alcohol abuse versus mild head injury. Such differences in level of background research might dictate a reversal of usual test selection. Such possibilities again suggest that specific knowledge about specific test qualities can greatly facilitate test selection.

Legal evaluations may also be weakened by the use of obsolete tests, particularly when more contemporary and improved versions are available. Some tests do not age well, and should the lawyer present some of the items in the cross-examination, jurors may be left shaking their heads and wondering, in fact, who these supposedly famous people that the test required the examinee to identify might have been.

Psychologists may substitute short, and clearly inferior, versions of full or standard forms of tests. Shortened versions commonly are far from perfect predictors of results on full versions of tests. As full versions of tests are usually rather imperfect predictors of the matters at issue, using shortened versions adds error to error (Ziskin, 1995). The lawyer might also be able to introduce statements by the test creator arguing vehemently against the use of short forms.

In some cases, alternate or parallel versions of a test are available, but one of the versions has been much more thoroughly researched than the other and there is inadequate research about the equivalence of the alternative forms. Especially if the test is administered only once, it would seem prudent to use the better known alternative. Similarly, for some tests, there are alternate administration formats, or alternate test materials or equipment, with one version being far better investigated than another. For example, a psychologist may use certain equipment when administering a test of finger tapping speed, but may depend on norms developed on other equipment when interpreting the results.

Although the use of one piece of equipment over another might seem like a trivial matter, research may show, as it does in this case, that different equipment can yield different results (e.g., Brandon, Chavez, & Bennett,1986). The psychologist may also have to admit that he ignored the stern admonition of Dr. Reitan, a co-creator of the very battery he used, to avoid alterations in original equipment lest one compromise the interpretive value of results (Reitan & Wolfson, 1993).

I have consulted on more than a few cases in which psychologists used badly deteriorated test materials or made up their own "duplicates" of tests that contained errors in instructions or in the reproduction of stimulus materials. For example, in one case, the stimulus sheets for a task requiring color discrimination were more than a little faded; and in another, a homemade form for an intelligence test contained alterations in standard test instructions. Still another case involved Part B of the Trail Making Test (Reitan & Wolfson, 1993), a paper-and-pencil task, which contains numbers and letters randomly arranged on the page that are to be connected in order (e.g., 1-A-2-B-3-C, etc.). The stimulus materials are arranged so that the continuing line to be drawn from one number or letter to another never crosses over itself, thereby keeping the spatial layout relatively clean or simple. However, the psychologist had mispositioned the stimuli such that the respondent now had to crisscross over the line repeatedly, a new variation that seemingly made the task much harder. In other cases, reproductions of materials from personality questionnaires contained errors: it must have been difficult for the examinee to answer unintended alterations of original items from the Minnesota Multiphasic Personality Inventory (Hathaway & McKinley, 1951, 1983) such as the following: "I sometimes feel that there is a tight bank around my head."

4.19.4.4.2 Obtaining adequate information

One of lawyers' most common cross-examination tactics is to bring up concrete facts that contradict, or seem to contradict, the psychologist's opinion. For example, in a case of purported PTSD, in which the psychologist has described the plaintiff's decided tendency to avoid reminders of the accident, the lawyer may ask if the psychologist knows that the plaintiff replaced the totalled vehicle with a new one of the same make and model and regularly drives on the road where the accident occurred when an alternative route is easily accessible. In general, courtroom opinions will likely be much more vulnerable to cross-examination if the psychologist lacks an adequate factual base. Reaching accurate conclusions about key matters also often demands such information. Consider again that legal evaluations often raise questions that differ from those that are most common in clinical settings and may well call for information gathering that differs in type and amount.

As discussed in Section 4.19.4.3, when information is received can be crucially important. For example, jurors are likely to react negatively if they find that the psychologist formed opinions very early and far before most information was reviewed, or waited until the eve of trial to perform a thorough analysis of records, years after issuing treatment recommendations intended to guide other professionals. The latter can look especially bad if the psychologist has insisted he places his role as a treater above that of courtroom expert and cares deeply about the examinee's welfare. The jury almost cannot help but wonder if this is impression management and disingenuous. After all, if the psychologist really is concerned about the individual, he would not have made important treatment recommendations on the basis of such incomplete information. Also, as noted, new records may call for entirely new lines of inquiry, and if materials are not examined until the 11th hour the opportunity may be lost. All of this again emphasizes the need for the psychologist to be thorough in information gathering in legal cases and, to the extent feasible, to obtain records earlier rather than later.

Information gathering usually must go beyond self-report and testing and include various types of collateral information, such as past and present medical records and reports, employment records, school records, and materials memorializing other's observations or reports about the plaintiff. It may be especially useful to access the observations of a neutral party who knew the plaintiff before and after the accident or event. These types of observations may be contained in background records (e.g., work evaluations) or in depositions, which may be as good or better than interviews. Information gathering will usually address at least four issues: a) prior functioning, b) current everyday functioning, c) possible causes or alternative explanations for the plaintiff's presentation or reported problems, and d) the accuracy and completeness of the plaintiff's report to the examiner.

To elaborate on these four areas, an analysis of prior functioning is important in order to determine the presence and possible extent of changes in functioning. In principle, plaintiffs are not to be compensated for problems that pre-date the injury. Typical psychological assessment methods, by themselves, are often limited tools for determining prior abilities or adjustment. For example, psychological testing results usually provide, at best, inferential or indirect evidence. Much more direct information about prior functioning often is available or can be made available. Thus, extensive occupational records showing excellent and steady job performance are likely to be considerably more helpful and trustworthy in developing an understanding of past work function than inferences based on an IQ test. There are also, often, similarly direct and independent sources of information about current functioning that can supplement and extend, and serve as means to check on, impressions formulated from the plaintiff's report and testing.

Using collateral records to help appraise the accuracy and completeness of the plaintiff's report is essential, not only because there is a heightened risk of misrepresentation in legal cases, but because inadvertent misrepresentations can occur and lead to over- or underestimates of adverse changes. For example, depressed individuals may understate, or individuals with serious brain injuries may overstate, their current level of functioning. Also, the plaintiff may simply have forgotten, or may be unaware of, important matters. For example, she probably will be unable to state precisely when her developmental milestones were achieved, may be uncertain about or unaware of performances on past standardized testing, or may be confused about prior medical diagnoses. A psychologist is in an especially compromised position when he does not try to obtain collateral information in areas in which the deficits he assumes the examinee manifests would seem to impede that person's capacity to provide the very information the professional seeks. For example, the expert might ask a person, who he believes has serious problems in long-term memory, about remote historical events.

Many sources of information are usually potentially available to psychologists conducting legal evaluations. Past testing frequently can be obtained through such sources as school records, and sometimes through military records and pre-employment evaluations. Past work records, medical records, pharmacy records, psychological and counseling records, and criminal records may also provide important information or leads about previous strengths and weaknesses and about possible causes for presenting problems. One can examine these materials for the facts they provide, and also for direct and indirect indicia of prior functioning and capacities. Forms completed in the past, for example, may allow the psychologist to determine whether his impression about decreased mechanical writing quality is accurate. Information about post-incident and present functioning may be available from similar and other sources, such as work records and work samples. Further, some individuals who stop working return to educational activities or develop new hobbies. Contemporaneous school records may suggest that a brain-damaged patient over-reported current academic performance. Alternatively, they may show abilities to learn new skills and to engage in problem solving that raise very serious questions about the accuracy or meaning of low scores on measures designed to tap such functions.

The plaintiff's deposition may provide an extended sample of various cognitive functions, such as language use and the ability to attend to questions. I have been involved in cases in which plaintiffs with purportedly severe attentional and language comprehension problems answered hundreds of questions without a single one having to be repeated and seemed to have no difficulty understanding the lawyer's inquires, despite the recurrent use of complex grammatical structure and relatively sophisticated vocabulary. Review of the plaintiff's deposition may also provide experts with important information that is not contained in other records, such as materials on past accidents or problems with the law.

Legal cases often involve questions relating to quality of life, day-to-day functioning, and work capacity. Psychologists may not habitually collect detailed information in areas that are directly relevant to these concerns. For example, an accident may have forced the injured individual to discontinue various enjoyable activities or hobbies. It is difficult to address impact on day-to-day functioning unless one develops a reasonably detailed description and chronology in this area. A courtroom evaluation in which the psychologist assigns a diagnosis, describes areas of maladjustment, often in abstract terms (e.g., Smith has elevated anxiety levels), and presents a general treatment plan may be of minimal help in understanding possible damages or the specific consequences of an injury. Too many legal evaluations are devoid of an adequate connection with the individual's day-to-day life and functions. It is usually much more helpful to know what the individual did before and can and cannot do now due to his reduced memory than to know the percentiles for his scores on memory testing.

In the area of work capacity, except in gross or obvious cases, it would often seem difficult to reach a well-grounded conclusion about possible return to former employment without a reasonable understanding of prior job requirements. A general label or description may not be sufficient, and one needs to understand just what the job required. In one case, the plaintiff indicated he operated a machine that made paper bags, which the psychologist assumed was a rather simple matter. However, it turned out that the equipment was incredibly complex and often required intricate adjustments involving dozens of steps. Psychologists need not feel obligated to address work capacity, even if it is an issue in the case, should it involve matters with which they do not feel sufficiently informed or if they do not believe the required scientific backing is available to perform this analysis. (It would be much better to inform the attorney of this limit in advance, however, so that counsel can consider retaining an additional expert and is not unexpectedly stuck at trial with no one to address occupational issues.) The point is that if a psychologist is going to address work capacity, certain information would seem essential.

A well-rounded picture of an individual and an understanding of her functional capacities is also likely to touch on such topics as self-help skills, household activities and chores, interpersonal relationships, activities outside of the home and job, the state of the marriage, child care responsibilities and capacities, and travel activities. The examiner may also want to know who handles other responsibilities of major importance, even if they come up only occasionally, such as large purchases. There is something potentially inconsistent about a claim of gross reductions in mental ability and the fact that this same individual, with the spouse's full blessing, handled delicate and complex negotiations for the acquisition of the new home.

In this vein, the psychologist might be especially alert to, and actively look for, circumstances in which the benefits and disadvantages of being capable and incapable are reversed. For example, it might be to the plaintiff's advantage to appear quite disabled during the psychological evaluation, but quite capable when applying for a business loan. Consistency in presentation across such circumstances, even when the individual has a great deal to lose, suggests something about the genuineness of incapacity. In one case, a plaintiff in a psychological injury case was also embroiled in a unrelated custody battle. The plaintiff had undergone independent psychological evaluations in each case. On the evaluation conducted for the personal injury case, he endorsed many items in the pathological direction on the MMPI, but reversed these answers in virtually every instance in the custody case: when it was to his advantage to claim disorder he claimed disorder, and when it was to his advantage to claim psychological health he claimed health.

4.19.4.4.3 Conduct technically proficient evaluations

It is not unusual for opposing counsel to obtain the psychologist's complete file, scrutinize it closely, and have her own expert review it as well. Technical problems with the examination may well be uncovered and cause difficulties.

Some of the more common problems I have observed in psychologists' courtroom work include errors in scoring psychological test items and in summing and transforming scores. I have reviewed cases in which psychologists made dozens of scoring errors or miscalculated standard scores by 30 or 40 points. In fact, the psychologist's errors may be close cousins to those he used to diagnose disorder in the plaintiff. The psychologist might then face a series of questions like the following:

> Q: You told us that Ms. Smith made errors on simple math problems, and that this entered into your conclusion that she is brain damaged, correct?
> Q: You made errors on simple math problems also, didn't you.
> Q: You told us that Ms. Smith sometimes did not attend to important details when performing paperwork, and that this was another piece of evidence suggesting brain damage, isn't that true?
> Q: You forgot to date a number of your test records, isn't that correct?
> Q: Dating a test record might be considered an important detail when performing paperwork, isn't that correct?
> Q: And you told us that Ms. Smith sometimes failed to follow fairly simple directions and that this was another indication of brain damage, correct?
> Q: Doctor, the instruction you had to follow, "Stop testing after five consecutive errors," is not complex, is it?
> Q: And you failed to follow this direction on multiple occasions, isn't that correct?
> Q: Let me see if I have this right doctor. When Ms. Smith makes errors, they indicate brain damage. When you make errors, they're just errors? Strike that, I withdraw the question.

In other cases, psychologists may violate standard testing procedures without any clear or compelling rationale. A psychologist in one case repeatedly terminated tests prematurely, calculated scores as if this had never occurred, failed to mention these alterations in procedure in her report, and had even published an article on the need to adhere exactly to standardized administration formats. Fortunately for her, the case settled before trial. There are certainly situations in which a test must be terminated prematurely or it is sensible to alter procedures (e.g., dividing a test that is preferably administered in one sitting into two sessions due to the onset of extreme fatigue part way through), but these departures should be recorded and described. What is hard to justify is altering standard procedures without any good reason. Psychology can be difficult enough without introducing unnecessary sources of error. Standard procedures may also be violated due to insufficient familiarity with prescribed methods. For example, in the cases I review, many experts seem unfamiliar with the precise rules for scoring the Visual Reproduction subtests from the Wechsler Memory Scale-Revised (Wechsler, 1987).

Failure to follow standardized testing procedures can make the psychologist look bad. Take, for example:

> Q: Doctor, you told us that Ms. Smith demonstrated problems in delayed memory, isn't that true?
> A: Yes, there were problems that seemed significant to me.
> Q: And you administered the delayed version of the Fisbee Memory Test, correct?
> A: Correct.
> The lawyer might next ask some questions to help the jury get a clear picture of the test and the way delayed memory is examined. The lawyer then asks:
> Q: And you told us that Ms. Smith scored in the borderline range on the test, didn't you?
> A: Yes, that's a label I used, although her percentile was rather low, at about the 9th percentile, if I remember correctly.
> Q: Borderline is the category between normal and abnormal, right?
> A: I'm not sure I would say it that way, but I would agree with the basic thrust of your question.
> Q: Had she gotten a few more points of credit on the test, the score would have been classified as low average, isn't that true?
> A: That's true, but she didn't.
> Q: And you have already agreed that according to school records, Ms. Smith was functioning at a low average level in a number of areas before the accident, isn't that true?
> A: Yes, but I'm not sure that applies to her memory functioning.
> Q: And one of the reasons you are not sure it applies to memory functioning is because we have no direct pre-accident memory tests, do we?
> A: Well, I think some of the past tests reflect on memory abilities, but it is true I have seen no records of specific memory tests that predate the accident.
> Q: Doctor, let's return to the Fisbee Memory Test that you gave to help determine the possible impact of the accident. Doesn't the test manual state that the delayed version is to be given about 30 minutes after completing the immediate version?
> A: Yes.
> Q: Doctor, do we usually remember things better over a longer or shorter period of time?
> A: Shorter may be better, but not always.
> Q: Doctor, you're not telling us, are you, that if I want to win a bet about when these jurors will remember the most about the trial, I should bet on next month rather than tomorrow?
> A: I don't accept your analogy, and psychological phenomena do not always follow what might be considered common sense.
> Q: So we're learning. In any case, when you administered the delayed version, you did not wait 30 minutes as the manual indicates, you waited 60 minutes instead, isn't that true?
> A: Yes, but I don't think it makes any difference.
> Q: Doctor, the instruction in the manual does not say it's fine to wait 60 minutes, does it?
> A: No it doesn't, but many studies suggest that the amount of memory loss that occurs after 30 minutes and 60 minutes is similar.
> Q: Doesn't the manual also state, and I quote, "The examiner should follow the instructions specified in this manual for test administration exactly in order to ensure proper comparison with the standardization group"?
> A: I can't recall exactly.
> Q: Well, I can show you the manual if you want.
> A: No, that's OK, I'll take your word for it.
> Q: Certainly it is possible, is it not, that a different result might have been obtained if delayed memory had been administered after 30 minutes versus waiting twice as long? (It probably will not matter what the psychologist says.)
> A: It is possible, but I doubt it.
> Q: And doctor, the reason to give the test as the manual specifies is so that we know how the individual performs when the test is given as designed, rather than having to guess how it would have come out if it had been given as the manual instructs, correct?

A: Those are your words.
Q: Doctor, I'm asking you. However, let's just move on. In a sense, doctor, you were measuring her on the 30 yard dash but made her run 60 yards, isn't that true?
A: No, I don't think you can make that comparison at all.
(The expert might think that but the jury is likely to believe that the analogy makes pretty good sense.)
Q: You waited twice as long as the prescribed time, and Ms. Smith obtained a borderline score, which you agreed was just a few points below the low average range, correct?
A: Yes.

I have also reviewed reports laden with factual inaccuracies relating to such matters as level of education, the date of the accident, the number of prior accidents, the number of children the plaintiff has, job history, etc. It may be difficult for a juror to believe that an expert can reach accurate conclusions about things that cannot be directly observed and that are complex—-such as the area in which the brain is damaged or about the inner workings of the mind—if the expert cannot get simple facts straight, such as the day the accident occurred.

4.19.4.4.4 Give adequate consideration to alternatives

Research on clinical judgment suggests that diagnosticians can increase accuracy by waiting longer before reaching conclusions and considering alternative possibilities more actively (Faust, 1984; Faust & Willis, in press). In courtroom cases, the opposing lawyer will often present causal theories or explanations that contrast to those the expert proposed. An expert who has already made a systematic effort to evaluate and consider alternatives is more likely to be correct in the first place and better prepared to defend his position. For example, when the lawyer asks, "Isn't it true that excess caffeine intake can also cause symptoms of anxiety? the response might be, "Although that is certainly true, the chart I constructed of caffeine intake and anxiety levels shows that anxiety levels often were high even when caffeine intake was low." The expert who has not carefully evaluated plausible alternatives might have to admit, instead, that she is not in a position to say whether, or the extent to which, caffeine might have contributed to the clinical presentation. The increased probability of an accurate conclusion also enhances the likelihood of providing true assistance to the trier of fact and may be critical in the examinee's care, whatever the courtroom implications. For example, if the plaintiff's cognitive dysfunction is due to the onset of a schizophrenic disorder as opposed to mild brain damage, very different treatment is likely to be indicated.

(i) Alternative diagnoses, including malingering

Given the overlap in symptomatology across conditions, or the lack of specificity of many symptoms (e.g., anxiety, sleep problems, difficulties concentrating), clinical presentations often raise multiple alternative possibilities that require careful analysis and reanalysis of the positive and negative evidence. The defense expert who is leaning towards adjustment disorder may need to reconsider the chronicity of the condition and the substantial level of maladjustment. The plaintiff's expert who diagnoses PTSD may need to re-examine the absence of physiological arousal and the infrequency of intrusive thoughts. A neuropsychologist who is quick to diagnose a mild brain injury may have mistaken it for a depression that started earlier following disturbing life events. A psychologist who identifies characterological disorder as a basis for what he believes are false or exaggerated perceptions of sexual harassment may need to abandon this conclusion when collateral records for the 10 years prior to the reported events nearly all suggest good adjustment and interpersonal relations. Along these lines, various diagnoses are packed with assumptions about previous functioning, and experts may miss an excellent opportunity to check on their impressions by examining whether records about these periods conform to expectation. For example, someone diagnosed with hypochondriasis, which is usually assumed to be chronic, would be expected to have voiced multiple medical complaints in the past, and not just since the injury 6 months ago. If various past medical records show select, delineated, and realistic medical complaints (e.g., the finger did turn out to be broken when x rayed), it may be time to rethink the diagnosis.

In many legal cases, psychologists should be more complete in the assessment of malingering. For example, they may not collect any collateral information and may limit appraisal of malingering to interview impressions and a single test that the literature shows to have poor sensitivity. I have reviewed many evaluations in which the Rey 15-Item Test (Rey, 1964) was the only measure specifically used to appraise malingering. The sensitivity of this test is so poor that in cases of malingering, it apparently stands a worse chance of detection than a coin toss (see Rogers, 1997a). Other research suggests that it is difficult to detect malingering through typical interview methods and clinical impression (see Faust & Ackley, 1998); and as Ziskin (1995) suggested, when a mental health professional asserts otherwise, the lawyer can ask, "Each time you've been fooled, you don't know it, do you?"

The psychologist who aspires to expertise in legal assessment should have a solid familiarity with the literature on malingering. Rogers' (1997a) edited book is a good starting point in this venture. The literature demonstrates the limits of traditional methods (e.g., interviews, many psychological tests), suggesting that: (a) individuals can manipulate results on a wide variety of tests, and in so doing may fool clinicians into overdiagnosing emotional or cognitive disorder; (b) it may be very difficult to detect face-to-face lies, especially by interview or impressionistic methods; (c) experience by itself does not ensure adequate detection capacities; and (d) the base rates for malingering may be higher than many psychologists believe (but lower than other psychologists believe) (see Faust & Ackley, 1998; Rogers, 1997a). For example, Reynolds (1998) asserts that "reasonable and thorough research indicates that at least 25% of cases of head injury in litigation involve malingering" (p. viii). Familiarity with the "negative" literature on clinicians' detection accuracy using more traditional methods is helpful because it directs us to take additional steps and because there has been such a recent growth in malingering detection methods—alternative or supplemental methods are available. This is a prime example of an area in which recognition of limitations in the field has had a very constructive impact by sparking intensive, productive research efforts.

The MMPI remains the most thoroughly researched test for malingering detection. MMPI indices and methods for malingering detection go beyond such traditional scales as F, L, and K, and are well described in various overviews of the topic (e.g., Greene, 1991, 1997). Differences in application with emotional distress versus brain damage claims are important to recognize, although the MMPI may be of considerable value in malingering detection with both types of cases.

There are also many specialized tests for malingering detection that are at varying stages of scientific development and of varying utility. One such example is forced-choice methods (Pankratz, 1988). Tasks are set up in which the examinee is instructed to select the correct answer from among two or more choices. For example, the examiner may read a string of digits to an examinee, then show her two written strings, and ask her to select the one that she just heard. A delay may be introduced before selections are made to increase the difficulty, or perceived difficulty, of the task. With a dichotomous choice format, even an individual with no memory capacity should achieve about a 50% level of accuracy through random guessing. Some malingerers overplay the role, failing to realize that performance that falls substantially below chance requires knowledge of correct answers. Although such types of forced-choice methods may have high valid-positive rates (correctly identifying malingering when test results are positive), valid-negative rates (correctly excluding malingering when results are negative) are often poor (Rogers, 1997a). There are, however, rapid ongoing developments with this and other specialized malingering detection methods, creating good reason for optimism and making it important to regularly update one's knowledge of the literature.

Interview methods are also available, in particular the Structured Interview of Reported Symptoms (SIRS), developed by Rogers, Bagby, and Dickens (1992). The SIRS uses various detection strategies, the majority of which seem to capitalize on false stereotypes or general beliefs about mental disorder. A variety of studies from independent researchers on a range of topics suggests that the SIRS may have strong properties for malingering detection (see Rogers, 1997b).

Other recent studies have investigated whether knowledge of disorder and knowledge of the strategies that underlie the design of malingering detection methods help the examinee escape detection (see Faust & Ackley, 1998). This research is still at a relatively early stage of development, although it suggests the following tentative generalizations. With testing methods that use something more than the very simple types of detection strategies (like the MMPI and unlike, say, forced-choice methods), knowledge of disorder seems to be of limited helpfulness. In contrast, knowledge of detection strategies increases—sometimes markedly—the likelihood of escaping detection. With typical or unstructured interviews, however, knowledge of disorder may well be useful, and perhaps the most helpful element in successful malingering. If these conjectures turn out to be correct, it seems likely that the chances of detecting better prepared or more sophisticated malingerers may be greatly enhanced by using a combination of specialized testing and specialized interview methods (such as the SIRS), although one must worry about inflating the false-positive error rate. Optimal use of these methods, especially if multiple data sources are obtained, will likely be achieved through the development and application of formally validated decision rules or actuarial strategies for data combination (Dawes, Faust, & Meehl, 1989; Grove & Meehl, 1997; Meehl, 1954). The idea that more information is better may already seem obvious, but considerable research in clinical decision making across multiple areas shows that intuitions about these matter are frequently misplaced and that there are many circumstances in which greater selectivity in data use and combination would result in greater accuracy (see Dawes et al., 1989; Faust, 1984; Faust & Willis, in press).

(ii) Considering alternative causative factors

Assuming a condition (e.g., brain damage or PTSD) is present, the question of what caused it can completely determine the outcome of a legal case, and hence is often one of the forensic evaluator's central concerns. Although some experts depend mainly or solely on self-report to determine cause, even the most honest plaintiff may not know what created her difficulties, or can make inadvertent errors in associating symptoms with events and conditions. If patients always knew what caused their problems they would not need doctors or mental health experts to perform differential diagnosis. Overattention to temporal sequence or focus on a salient event may lead to misattributions, as might faulty diagnoses by other providers. Factors may be operating that the individual could not possibly have been aware of, such as an unrecognized exposure to a toxin at the work site that causes a gradual, delayed reaction. Suppose such exposure leads to insidious decreases in reaction time and coordination, which in turn lead to a car accident that causes a very mild head injury. It is easy to see how causal confusion can arise. In other instances, of course, plaintiffs purposely mislead. This may be very difficult to detect, because their conditions may be real but may have arisen from another cause that they will hide or disguise. The chances of uncovering the deception may depend largely on obtaining sufficient background information.

Assessment results may also stem, in part or in whole, from transient or extraneous factors (for discussion of multiple possible confounding variables in brain damage cases, see Faust, 1995). Individuals are sometimes tested under poor conditions that obfuscate conclusions about more stable impairments. In one case, a nationally known expert, who had been retained by the defense, examined the plaintiff in a room in an airport. He concluded that problems seen on testing probably were not caused so much by the head injury in question but rather by exposure to jet fuel fumes during the examination. In another case, an individual set out very early in the morning to drive the 200 miles to the psychologist's office, and ended up so tired he could not stay awake continuously during the testing. I have even reviewed cases in which examinees told the psychologist that they literally had not slept at all the night before, and yet the testing proceeded. Children are sometimes tested when they are ill and cranky or suffering from ear infections. In other situations, psychologists continue testing for 8 or 9 or 10 hours, and then write a report indicating that one of the plaintiff's main difficulties is rapid fatigability (while maintaining they have obtained "valid" testing results or results that do not underestimate capacities).

Many psychologists have very busy schedules, and cancelling a case slotted for a half or full day can create real headaches, but sometimes there would seem to be no reasonable alternative. Otherwise, the psychologist may have to answer questions like, "Doctor, do you ever tell your patients, before an important examination, to try to stay up all night so that they can do their best?" In contrast, there may be circumstances in which these types of accompanying problems and maladies are stable features of an individual's condition. For example, it might be that the plaintiff's physical discomfort, versus the brain injury, was the greatest contributor to performance difficulties during the examination. However, if the accident caused painful orthopaedic problems that are unlikely to remit, the adverse effects on cognitive performance may well be characteristic and typical of the individual's day-to-day functioning.

The many factors that may cause extraneous or transient alterations in cognitive, emotional, or behavioral status include, to name a few, sleep deprivation, pain, caffeine abuse, alcohol and other forms of substance abuse, medication side effects, and a myriad of independent medical conditions. Stressors and problems separate from the accident may also impact greatly on emotional or cognitive functioning. In situations in which these types of factors may have altered results, it can be very helpful to re-examine the individual at a later time. For example, if there is a question about the relative contributions of mood disorder versus brain injury to performance on cognitive tests, retesting after the depression has been treated successfully may help to parse these factors.

4.19.4.5 Interpretive Strategies

A detailed discussion of interpretive strategies is well beyond the scope of this chapter, and only a few general guidelines will be suggested. Some of these guides may seem obvious and might be best viewed as reminders; I hope they will not strike the reader as too rudimentary. Other guides stem from the extensive literature on decision making, which provides many useful ideas and procedures. The decision making literature often has not infused writings on clinical practice. This is unfortunate because many of these principles, while very helpful, are counterintuitive, even for professionals, and hence are unlikely to be recognized or realized without direct exposure. Introductions to the decision making literature can be found in a variety of sources (e.g., Arkes, 1981; Faust, 1986; Faust & Wedding, 1989; Faust & Willis, in press; Wedding & Faust, 1989).

The recognition that interpretation is unlikely to be better than the data upon which it rests highlights the need for careful collection of information and adherence to proper examination procedures. Prudent data gathering can go a long way towards increasing the chances of reaching sound and defensible opinions.

It is hardly newsworthy to suggest that not all interpretive procedures and available methods are equally sound. Not uncommonly, various procedures are applicable to the same data and lead to conflicting conclusions. For example, there may be five or more procedures for judging cooperation with the examination or malingering, with some pointing in different directions. Further, there may be no simple way to combine or integrate these results, because they may directly contradict one another. If one decision rule indicates that there is brain damage and the other that there is not, both cannot be correct.

Obviously, when all decision rules coincide, there is nothing to resolve, but when they conflict directly, one must select one over the other (or defer judgment). At times this may not be too difficult, because the great bulk of the evidence, especially that of the highest quality, points in a certain direction; but in other cases the results are more ambiguous. For example, what does one do if the single best decision procedure argues for Conclusion A, but two other methods of more modest accuracy argue for Conclusion B, and one is not sure how these latter two methods operate in combination?

In these conflictual situations, a psychologist is much better off, to illustrate a relative extreme (and assuming all other things are equal), if the decision rule given preference has stronger research support, has provided a clearer result, is better suited to the examinee, and, when pitted against the other decision rule(s), has been shown to be correct more often. As the table tilts in the other direction, the chances of being correct decrease and vulnerability to informed and skilled cross-examination increases. These and other issues involved with the combination of data are addressed in such sources as Dawes et al. (1989); Meehl (1973); and Faust (1984) (see also Faust & Willis, in press). As these authors note, considerable research suggests that it is not necessarily best to try to combine all of the data or to analyze complex configural relations. Formal decision procedures (e.g., actuarial methods) that have been developed through proper scientific methods almost always equal or exceed the overall accuracy attained through these types of attempts at subjective or clinical data integration.

Interpretive methods may fail to consider base rates (see Meehl & Rosen, 1955; Wedding & Faust, 1989), or the frequency of events. For example, the accuracy of diagnostic signs and methods varies in relation to the frequency of a condition, and adjustments in the application of such indicators and in estimates of likelihood are likely to be needed as base rates change. To illustrate, as the frequency of malingering decreases across settings of application, the same score on a simulation index signals a lower probability of malingering, and hence cutting scores may need to be raised to avoid an unacceptable false-positive error rate. Along related lines, characteristics that are common among the general population may be described as indicators of pathology. Some such "signs" are part of clinical lore, and their use may not have been modified in accord with research. For example, a psychologist may assert that five points of subtest scatter on the Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981) is indicative of disorder, although research shows that most normal individuals equal or exceed this level (see Matarazzo, Daniel, Prifitera, & Herman, 1988). Given mental health professionals' skewed exposure to abnormal populations, it may be difficult to determine whether features seen commonly among patients are indicative of pathology or rather just common among individuals in general and thus poor discriminators of disorder.

I have reviewed many reports in which pedestrian human failings were stated as self-evident signs of disorder. When supposed pathological features include sometimes misplacing one's keys, occasionally forgetting where the car is parked, and irritability with misbehaving children, a good cross-examiner can have a field day. It is often necessary to consult epidemiological research and associated literature to determine just how often features occur among normal and abnormal populations, and also whether they occur with differential frequency among different forms of disorder. The latter type of knowledge can be essential to differential diagnosis.

Overdiagnosis, or the tendency to see pathology that is not present or to overestimate its severity, may result from various sources. Overestimation of prior functioning may lead to overestimation of changes in functioning. False conclusions about prior functioning may stem from such sources as faulty history that the plaintiff reported or faulty methodology. For example, some psychologists believe that the single highest, or few highest, test scores provide a good estimate of prior overall or general intellectual functioning (the so-called "best performance approach"). Given normal variation in test scores, this method is almost sure to lead to overestimates, and not uncommonly gross overestimates, of prior functioning, especially among unaffected individuals. Given the average seven point spread between the highest and lowest subtest scores on the Wechsler Intelligence Scales (Matarazzo et al., 1988), use of the single highest score with normal individuals will result in an average overestimate of 15 to 20 points in prior Full Scale IQ (FSIQ). The result will be an estimated loss of 15 to 20 FSIQ points in an individual with absolutely no adverse change in intellectual functioning. This might be the difference, for example, between a score near the 90th percentile and one at the 50th percentile. The absurdity of this situation is that, on average, an individual will have to have gained over 15 FSIQ points to be judged to have retained his prior capacities!

Another basis for overdiagnosis is the use of inappropriate norms that set overly demanding performance standards. This problem is especially common when norms are not adjusted for sociocultural and demographic features or test bias. For example, a test may show pronounced age and education effects and yet a clinician, perhaps unaware of these findings, might apply norms developed on young, highly educated individuals to much older and less educated individuals. Also, in at least some normative studies, individuals with virtually any significant risk factor are eliminated from the sample, resulting in supernormal groups, or groups whose performances well exceed the general population (this being one reason why average IQs are often so high in normative studies on various neuropsychological measures). The result can be a mistaken belief that unaffected individuals have suffered a loss attributable to the event in question (e.g., a head injury) because their test performances do not meet the inflated standard set by the "normative" group. This can also lead to the initiation of treatments which may carry risks for target problems that do not exist. Overdiagnosis can also stem from overattention to select weaknesses, insufficient awareness of the overlap between normal and abnormal populations, and failure to recognize normal variations within or across individuals, the latter of which was previously illustrated by reference to subtest scatter on intellectual tests. Normal individuals are rarely unremarkable or well adjusted or well functioning in all respects, and the tendency to conclude that an individual is aberrant due to even minor shortcomings creates a hurdle that few of us would pass.

Other psychologists underestimate or fail to recognize pathology, with explanatory factors sometimes representing the flip side of the same coin that leads to overdiagnosis. For example, psychologists may be quick to assume that problems pre-dated the accident. Weaknesses on cognitive or neuropsychological tests may be attributed to pre-existing learning disabilities, even without checking school records. Emotional problems may be assumed to represent Axis II disorders, and hence characterized as life-long difficulties as opposed to consequences of a more recent accident. In a case in which I consulted involving an individual with a moderate to severe brain injury, a psychologist passed off various potential indications of frontal lobe disorder, such as impulsiveness and marked difficulties with interpersonal relations, as indicators of a borderline personality disorder (BPD). However, this individual also presented with various features unrelated to BPD but positively associated with brain injury, only demonstrated characteristics of BPD that overlapped with those of brain damage, and demonstrated virtually nothing pre-injury to suggest any type of personality disorder.

Some experts are very quick to assume that individuals are malingering and that this explains their symptom presentation. In a toxic exposure case, a psychologist casually dismissed seemingly strong neurological findings as feigned symptoms. Ironically, as it turned out, although the psychologist was likely right that the toxin did not cause a problem, she was otherwise quite wrong: further medical work up yielded a nearly definitive diagnosis of multiple sclerosis and left virtually no doubt that the plaintiff had serious neurologic disorder all along. Underestimations of prior functioning may also lead to missed pathology. Finally, in some cases, experts misidentify chronic or permanent symptoms for transient ones. For example, in brain damage cases, experts may be too ready to assume that all of the plaintiff's problems are due to depression when, for example, the original injury was serious, the individual did not appear to have a low mood when evaluated, depression has fluctuated but impairment has not, and at least some of the observed problems (e.g., perseveration and aphasic errors) are much more strongly associated with brain damage than with depression.

4.19.4.6 Preparing Reports

Reports may or may not be introduced as evidence at a trial. When they are, every word becomes a possible target for cross-examination, and thus they should be prepared very carefully. Although I will leave detailed recommendations about report preparation to others, I do have a few suggestions. First, sloppy reports with factual inaccuracies can create a very bad impression, even if the errors are not substantive. In one case, due to a typographical error and poor notes, a psychologist could not say whether a plaintiff, earlier in life, had fallen out of a rocker or had been struck in the head with a rock. Reports should also strive to provide a balanced representation of the case and of positives and negatives. For example, the report of an expert that repeatedly glosses over negatives that seemingly are evident and should not be ignored may be difficult to defend. Alternatively, the report may contain nothing but negatives. For example, when describing responses on a questionnaire or a depression inventory, the examiner may list only the unhealthy responses. This can become especially problematic when, in fact, the majority of the responses are positive. This creates an easy target for the cross-examiner, to wit:

> Q: Doctor, in presenting before us today, you strive, do you not, to provide fair and balanced testimony?
> Q: In gaining a complete understanding of an individual, strengths can be just as important as weaknesses, isn't that true?
> Q: Interventions or approaches to helping a person often build on someone's strengths, isn't that correct?
> Q: When describing the results of the Method X Depression Inventory, you shared Mr. Smith's responses to three items, isn't that correct?
> Q: Each of these items suggested possible problems, you would agree with that, wouldn't you?
> Q: And each of us, doctor, is something of a mixture of positives and negatives, wouldn't you agree?
> Q: There are 20 items on the Method X Depression Inventory, isn't that right?
> Q: We haven't heard anything about the other 17 responses, have we?
> Q: On item number 3, doesn't Mr. Smith indicate that he gets as much enjoyment out of things as he used to?
> Q: You didn't mention that in your report or your testimony, did you?
> The lawyer than reads five more positive answers, and the expert may make some comment such as the following:
> A: You're only mentioning the positive responses.
> Q: And it would be very wrong to only present one side of the picture, wouldn't it?

Some experts write exceedingly long reports that address many minor or irrelevant matters. Again, every word and every comment is potential fodder for cross-examination, and thus including matters that are really not important to the task at hand is unlikely to be of much benefit and could create a problem. For example, such text might include various factual inaccuracies that can make the psychologist look foolish.

Reports may also include many extreme statements. A good cross-examiner usually welcomes extreme statements or exaggerations, because they create an easier target. For example, in one case, an expert did not just say that one method was better than another, but that there was a general consensus among psychologists that it was not just superior, but far superior to another method. This extreme claim was easy to deflate when literature was introduced by the very author of the test describing its limitations. In another case, a psychologist testified that even a very mild brain injury affects every single aspect of an individual's functioning and existence. Although the case was resolved before trial, in part because this psychologist's assertions were so vulnerable, it would have been a simple matter to confront the witness with the many normal test performances (which were entirely consistent with pre-accident measures) and to ask her whether there was any chance they could have been exceptions to her claim. She almost certainly would have said no, leading to a rapid self-destruction.

The exact wording of reports may be important and lead to unanticipated legal consequences. For example, in one case, an extremely bright neurologist wrote a conclusory sentence that included the phrase, "but for the accident." His intended meaning was that the condition would have occurred whether or not the accident occurred. However, in the legalistic world of the arbitrator, the conventional interpretation of this phrase was to the contrary, thas is, that in the absence of the accident, the condition likely would not have occurred. It is generally much safer to stay away from legal terminology unless it is necessary and one understands exactly what one is doing, and rather to just say things plainly and clearly. The wording of a report should also be checked carefully to avoid unintended meanings or interpretations, although it may be possible to clear these up at deposition or trial without creating a problem.

4.19.5 LAWYERS' STRATEGIES AND TACTICS

When dealing with adverse expert testimony, the lawyer's most basic task is to undermine credibility. Some lawyers will attempt to do this with a scalpel and others with a blunt stone, but the aim is the same in either case, and it is to the witness's advantage not to forget it.

Although an attack may be directed at personal matters (e.g., bias, financial incentives), this is often merely part of doing business and reflects nothing personal. The same lawyer who attempted to paint the witness as biased or incompetent may shake hands on the way down the courtroom steps and tell the expert he thinks she did an excellent job. On the one hand, most lawyers do not act in a blatantly hostile or obnoxious manner because they believe their case is best served if the jurors like them. A scowling, hateful, and abusive manner towards a witness, especially one who has done little or nothing to provoke it and who acts in a perfectly civilized and seemingly impartial manner, can hurt the attorney much more than the expert. For an expert who might be a bit thin-skinned, it can be helpful to perform a little personal Albert Ellis and tell oneself that such personal attacks often stem from a position of weakness and do not provide true commentary on one's human worth. A lawyer who is really loaded with ammunition may well prefer to take on an almost pained and solemn expression when bringing to light terrible problems with the expert's work (so that the jury does not end up feeling sorry for the doctor). On the other hand, arrogance and unwillingness to admit obvious points, and other similar demeanor may leave the jury hoping that the expert gets what he deserves and may give the lawyer license to deal out punishment. Obvious lack of preparation can also quickly alienate a jury.

If the lawyer is particularly successful in weakening an expert, she may then introduce materials and questions that are not so much intended to damage the expert's credibility further, for that job has already been accomplished, but rather to put on her own side of the case. For example, rather than the attorney attacking the credibility of a plaintiff directly, which, if overdone or too aggressive, could inflame a jury, it is often safer for an attorney to do so through an opposing expert. The lawyer might raise inaccuracies in the plaintiff's report to the expert that seem purposeful and self-serving. A weakened expert is often fairly helpless against adverse, or seemingly adverse, evidence; and it can have a much greater impact for the attorney to introduce affirmative elements of her case through an opposing, rather than her own, witness. One might consider the effect when the opposing expert cannot fend off attacks on the plaintiff's credibility, versus the dubious impressions that can arise when the plaintiff seems credible and the plaintiff's experts, who vouched for the plaintiff's honesty, were not really questioned on this score. Rather, the first person to raise serious concerns about truthfulness just happens to be the expert the attorney hired.

Cross-examination often does not follow the contour of the expert's direct examination. Although, in some jurisdictions, cross-examination is supposedly restricted to the content covered in direct, in practice it is usually a relative free-for-all with few topics off limits. Further, a good cross-examiner rarely wants or needs to return to the points the expert covered tit-for-tat. Rather, the attorney is selective, looking for weaknesses or vulnerabilities. Many lawyers would much prefer to win a cross-examination 3 to 0, rather than 10 to 2. This is one reason experts should aspire to avoid weak components in their assessment batteries or procedures.

The areas of cross may or may not touch on any of the points covered in direct. In one case, a plaintiff's neuropsychologist had already admitted on deposition that during his entire professional career he had never concluded with relative certainty that someone was not brain damaged, nor had he ever directly identified someone as malingering. He further admitted that in the instant case, he had presumed that the plaintiff was brain damaged nearly from the start, based solely on the very limited (and seemingly far from definitive) information he received at the time the referral was arranged. It would not matter very much what that expert said during his direct, because the lawyer was going to come back to these points, drive them home, and then stop, leaving the expert in shambles. In fact, a lawyer often is able to prepare a cross-examination without thinking all that much about what the expert will say on direct. Rather, the lawyer may consider the gist of the expert's testimony, and then focus most of her attention on the underlying bases or evidence for conclusions and on matters minimally related to the specific content of the direct.

Points of attack, to be covered in order, may include credentials, bias, flaws in the conduct of the examination, questionable conclusions, and weaknesses in scientific underpinnings. Again, almost all of these subjects come back to the expert's credibility. Many lawyers would rather stay away from scientific topics, but others are well prepared to enter this arena. Experts may be accustomed to getting by with cursory or questionable answers to inquiries about underlying scientific methods or potential scientific weaknesses, and may be shocked the first time they confront an attorney who starts asking very specific questions about one or another line of research and will not settle for incomplete or vague answers. Some experts have permanently damaged their credibility by placing on record a host of patently wrong answers to questions about science. The result is the appearance of incompetence or, even worse, dishonesty. I have read more than a few transcripts with claims such as the following: Basic psychometric principles that apply to psychological tests are not relevant to neuropsychological tests; there is no type of board certification in psychology; the (you name it) battery is nearly 100% effective in diagnosing brain injury; there is an absence of research showing a limited relation between experience and accuracy, and so on.

4.19.5.1 Credentials

As discussed, most experts have credentials that sound impressive to a jury. For example, completing a doctoral dissertation, receiving an advanced degree, and publishing (at all) are likely to be viewed favorably. More advanced accomplishments are not likely to make much of a difference. It is ironic, then, that some experts will puff and inflate, or even distort, small points that a juror is likely to find unimportant or trivial. As illustrated earlier in the cross-examination exchange on graduation from an accredited program (see the heading, "Admissibility"), these little points can expand into disasters if they are discovered, especially if the expert will not yield in the face of obvious contrary evidence. It is not unusual for attorneys to conduct background checks and to obtain information about an expert's credentials and past experiences.

4.19.5.2 Bias

A cross-examiner may start by asking the expert whether he endorses the scientific method, and whether it calls for impartiality in examining data, including fair consideration of evidence for and against a proposition. It is difficult to imagine an expert answering negatively to inquiries of this type. The lawyer may then try to show that the expert failed to live by such principles in the present case, or was not even-handed in considering and presenting evidence.

Some reports overemphasize or overattend to either the good or the bad, and may thereby make it easy for the lawyer to demonstrate bias. For example, an expert may list only the positive responses on an anxiety inventory. The lawyer might start reading the negative responses and after each one merely ask the expert, "Did I read the item correctly?" Alternatively, the expert may use dramatic terms and more lengthy narrative to describe weak cognitive performances, while downplaying, or even ignoring, strong performances. Problems may be referred to as "deficits," "impairments," or "serious shortcomings," but a remarkable performance is described as "essentially within normal limits."

Bias may also be suggested when experts fail to conduct sufficient investigation of alternative explanations for results. For example, an expert who quickly decided that the accident caused a brain injury may have been superficial or incomplete when looking into substance abuse as a possible alternative cause for reduced cognitive efficiency, despite records containing various suggestive references. Or, in a PTSD case, the expert may have made a perfunctory attempt to uncover and analyze other potential stressors. If the lawyer can show that alternative explanations were plausible but not pursued with anything approaching the zealousness displayed for the favored explanation, it can create a very negative impression.

Other experts come across as biased because they will not concede the obvious or will not give ground, even when they should. Similarly, some experts strike jurors as evasive. In one case in which I consulted, it was blatantly obvious that the plaintiff's signature was much poorer when produced at the psychologist's request then it was when he signed checks and documents in the course of his everyday life. When the lawyer merely asked, "Do these signatures appear to be of different quality?" the expert would not concede the point. This allowed the attorney to keep reframing the same basic question in a way that made the expert look increasingly outlandish, e.g.,

Q: Isn't there a clear difference in the quality?
Q: Isn't there some difference in the quality?
Q: Do these different signatures look exactly the same to you?

The lawyer concluded the cross by stating, "Doctor, what would you say if I asked you ... ah, never mind, I know what you'd say." Although there may have been a perfectly reasonable explanation for the difference in writing quality (e.g., the plaintiff was heavily medicated when seen by the psychologist), the expert's stubborn adherence to an unsupportable position convinced the jury that nothing he had said before, and nothing he might say after, was worth a second thought.

Other ways the lawyer might try to show bias include financial incentive (e.g., the expert charges much higher fees for legal than clinical work, or has performed numerous legal evaluations with the same attorney), or systematic error that consistently favors the expert's position. For example, a defense expert may have repeatedly overcredited a plaintiff on cognitive tests, thereby underestimating loss in functioning.

4.19.5.3 Manner of Conducting the Examination

Examination procedures have been covered at length previously and will not be repeated here. Errors of omission and commission can create easy fodder for attorneys and can lead to a complete disregard of the expert's testimony.

4.19.5.4 Erroneous or Questionable Conclusions

Erroneous conclusions can stem from various factors, such as mistakes in the scoring of tests or misapplication of scientific methods. In many instances, the correct answer is not cut and dry, and it may be difficult to show, conclusively, that an expert has made an error. Considering, however, the selectivity of cross-examiners, if even a few instances of clear-cut errors on nontrivial matters can be identified and brought out, it may greatly reduce the expert's impact. Concrete facts often provide the lawyer with the best opportunity to find such occurrences, and the search may be greatly facilitated if the expert's review of records has been incomplete or careless.

I have reviewed many cases in which, for example, an expert did not confirm educational level. For a neuropsychologist who adjusts most or all test scores in relation to education, this may lead to erroneous results on almost all measures. It can be difficult for a witness to regain his equilibrium with cross-examination surprises of this magnitude, especially if the normative tables are back at the office. I have also reviewed many cases in which the expert had taken a very limited occupational history. The psychologist might have testified that the plaintiff had never been fired from a job and had achieved a certain level of earnings, when work records show that these assumptions are plainly wrong. Many lawyers are careful in collecting and reviewing facts, are attuned to incomplete or faulty factual renditions, and are capable of using such material to raise challenges, that is, they are much more at home with this type of subject matter than with psychological concepts and research. An attorney is likely to be considerably more comfortable and effective arguing about former earnings than about whether the subject-to-variable ratio was sufficient to conduct a multivariate analysis.

Experts sometimes present conclusions that sound just plain silly or out of touch. For example, they may make a great deal out of normal human failings or overlook obvious, everyday explanations for events. I have read many reports in which supposed deficits are illustrated through examples that apply to most anyone, e.g., difficulty getting organized for a vacation, occasionally forgetting the exact day of the month, or a tendency to fatigue by the late afternoon (this in a person with a hectic job and four children). Financial incentives in cases may be too readily dismissed, or not acknowledged as potentially relevant. For example, a plaintiff who complains bitterly about problems but who has complied with almost no treatment recommendations, even those involving minimal effort and possible discomfort, may be described as giving no thought to the legal case and only wanting to get better. Some psychologists are so used to thinking in complex and abstract ways that they tend to overlook more common or seemingly mundane considerations. At other times, experts may fail to think through the implications of presumed problems or deficits, and whether the kinds of things one consequently expects to be present in the plaintiff's everyday life and lifestyle are present and those that seemingly should not be present are not present. For example, if an individual really has a severe problem with hostility and impulse control, it is doubtful his hunting buddies would continue their monthly get togethers at the cabin. Or, if the plaintiff really develops excruciating headaches when exposed to noise, one would not expect her to join a rock band. Everyday activities that fly in the face of the expert's conclusions can be decisive with juries.

4.19.5.5 Scientific Status

An attack on scientific status might be broad and aimed at the field in general, or narrow and more specifically targeted at the particular methods used in the case. Many lawyers shy away from challenging scientific foundations, but others have little or no hesitancy to do so, are conversant with the issues, and have retained a consultant to help them prepare. Further, unlike cross-examination, deposition inquiries about scientific status carry little risk, and helpful admissions or responses to even a small minority of questions may satisfy the attorney's eventual and basic purpose—to reduce or vitiate the expert's credibility at trial. An attorney may spend 2 or 3 hours asking many questions about the expert's methods and scientific backing, and then, at trial, focus on the one or few areas of questioning in this domain in which she feels she can make the most headway. Again, effective cross in just a few areas can inflict more general damage. Thus, other then wasted time and expense, it matters little if only one in three, or one in five, or one in ten areas of deposition questioning about research will be used at trial.

An expert makes the lawyer's search for damaging material on science much easier if she uses poor methods, lacks basic familiarity with pertinent research, makes grossly overblown claims, will not concede limitations that are well established in the literature, or repeatedly guesses when she is not sure about answers to questions. Much like certain forms of the martial arts, many points about science would probably carry little impact with the jury if experts did not take actions that gave them strength. Suppose, for example, there are a few negative studies on some otherwise well-supported method for appraising malingering. If asked, the expert could say, "Although most of the literature on the method is positive, one has to use it with some caution because isolated studies have not been supportive." However, on deposition, the expert might have argued that the literature is uniformly supportive. Once receiving that answer at the deposition, the lawyer may have nudged the expert out further and further on a limb. For example, the lawyer might ask about the importance of maintaining familiarity with literature on the methods one uses, how negative literature can bring a proposition into question, etc. Then, at trial, the lawyer can wave around the negative studies that the expert denied existed and recite the names of authors that the expert separately acknowledged as authorities.

As discussed, selection of the strongest possible methods is important not only in dealing with cross-examination, but in maximizing the chances of reaching correct conclusions and thereby fulfilling the presumed prescriptive (i.e., normative) role of a courtroom expert, to assist the jury in its deliberations. If there are no adequate methods available, or if the best available methods are questionable, the expert might decide not to undertake the assignment at all. Also, as noted, post-Daubert many courts are placing methods under increased scientific scrutiny and eliminating those found wanting.

4.19.6 DEPOSITIONS AND TRIAL TESTIMONY

As part of the discovery process, most states allow attorneys to depose opposing experts in order to learn what opinions they may express at trial and the underlying bases for their views. Attorneys vary greatly in their approaches to depositions, and the type and position of the case may dictate strategy. Some attorneys play dumb, ask many open-ended questions, and try to get as much helpful material as they can while revealing as little as possible about their trial strategy and anticipated lines of cross-examination. Their aim is to surprise the expert at trial. Others are far more aggressive and ask pointed, challenging questions that are intended to inflict damage, even should the element of surprise be reduced. If the case is almost certain to go to trial, many attorneys will take a more guarded posture, trying to save their best material for the cross. If it is a case the attorney wishes to settle, the deposition style may be more aggressive and aimed at showing the other side that its expert has problems and that monetary demands or offers need to be adjusted. Many cases call for some type of strategy in between, for example, one that exposes weaknesses in a few areas to convey a message that might aid in settlement negotiations, but that reserves some or most of the material for cross should the case go to trial.

4.19.6.1 A Sampling of Deposition Topics

Although the supposed aims of a deposition may be to uncover trial opinions and their bases, the attorney often already has a very good idea what conclusions the expert will express, especially if a detailed report has been prepared. For instance, basic elements of PTSD cases are frequently similar, e.g., that it was the event in question that caused the disorder, that a decrease in level of functioning resulted, and that the plaintiff was essentially forthright and cooperative during the examination. The attorney's main deposition aims will probably lie in other areas, and the scope of questioning is often wide ranging.

Although many experts prefer to talk in more general and abstract terms (e.g., this is a serious case of PTSD that has caused substantial distress and diminished functioning), many attorneys prefer to talk in more specific and concrete terms. In the area of damages, they want to know specifics. What is it exactly that the plaintiff can and cannot do? How long will therapy need to continue and at what frequency? Does the expert have an opinion about work capacity and about diminution in earnings? Exactly which problems can be attributed to the event and which pre-dated it? For problems that have supposedly been exacerbated, what is the precise extent of the change? Some psychologists become unnerved by these types of questions and react in ways that become problematical at trial. For example, they may fail to admit uncertainty, speculating about specifics that can be shown through concrete example and evidence to be wrong.

In addition to specific elements of damage, some of the topics that the expert can expect a well-prepared attorney to cover on deposition in a serious case can be discussed in turn. The attorney is likely to ask the expert about all her sources of information and when they were obtained. The attorney will probably request that the expert bring her complete file to the deposition, and may go through it document by document, asking when each was received. The attorney may also raise questions to determine who controlled the flow of information, e.g., did the expert request the document, did the attorney send it on his own, or did the plaintiff provide it? This type of questioning may translate into lines of cross-examination aimed at showing either that the expert formed opinions early and absent critical information; that the expert never has reviewed key documents; or that the lawyer, rather than the expert, determined what materials the expert reviewed. Further, knowing what documents the expert has not seen can give the lawyer a decided advantage. For example, if the attorney, but not the expert, knows that the plaintiff misrepresented her educational and occupational history, this can lead to embarrassing problems at trial.

As discussed already, the psychologist should not assume that the attorney who retained her will provide all relevant documents, or that it matters little when records were reviewed. Attorneys may not know which of the available documents an expert might want to see, or what new materials an expert might want to obtain. Some attorneys try to contain costs by limiting an expert's access to records or, in some cases, prefer to withhold certain documents for fear of their impact, hoping that the case will settle or that it will not create too big a problem at trial. Also, if a conclusion is reached weeks, months, or years earlier, and the expert does not review potentially critical documents until the eve of a deposition or trial, it can look very bad, especially if treatment recommendations had been issued. How can the expert explain why it took until 1997 to review documents that were available in 1994 and that may have altered the diagnosis and treatment recommendations? Additionally, on deposition, the opposing attorney commonly asks the expert whether he plans to do anything further on the case before the trial, and requests that any changes in opinion be disclosed. (Some jurisdictions require amended reports if new or altered opinions are to be introduced as evidence at trial.) This puts the expert in a very difficult position if he plans a last minute review of documents he should have examined much earlier.

The attorney may also ask whether any literature was used or consulted in the case. If the response is positive, detailed questioning may ensue. Some attorneys will ask knowledgeable questions about such matters as norms, reliability, and validity. Related questions may be asked about the specific assessment methods used in the case. For example, for the various tests the psychologist administered, the lawyer may want to know what normative standards were used, whether other norms were available, the basis for selecting one set over another, what results would have been produced with other norms, and whether the norms contained demographic corrections. The expert might also be asked about the existence of literature that raises questions about her assessment methods or demonstrates limitations in their use, and whether each method is supported by a body of research on accuracy or validity. Again, a positive response may be met with numerous specific questions, such as what studies exist and who published them; whether they involved the exact same methods, populations, and questions; and whether there is contrary literature. With questioning such as this, inaccurate answers or overblown claims, rather than concessions about relative weaknesses, often cause experts far greater problems at the time of trial.

An attorney usually cannot introduce literature during cross-examination unless the opposing expert has acknowledged it as authoritative or as something he relied on in forming his opinion. (The attorney may still have the option of introducing that literature through his own expert.) Generally, the lawyer wants to find out at deposition whether the needed acknowledgement can be obtained, because it is highly preferable to know what is feasible in advance of trial and plan accordingly. Discovering at trial that an expert will not acknowledge an article that was to serve as the centerpiece of a cross-examination may leave the attorney on a bridge that just lost its undergirding. Once an expert acknowledges an article as authoritative at deposition, there is usually no keeping it out at trial, even if the expert tries to backtrack. For example, should the expert say at trial, "Based on subsequent reading I no longer view that publication as authoritative," the lawyer can still refer back to the acknowledgement at the deposition and will probably be allowed to introduce the article and ask questions about it.

Some experts, due perhaps partly to uncertainty about the legal meaning of "authoritative" and because they wish to appear widely read, acknowledge a great range of literature as authoritative or provide very general endorsements. However, in the legal arena, authoritative means, in effect, that the expert defers to the source or considers it worthy of attention. This type of definition should be kept in mind when answering deposition questions about what is authoritative. Thus, for example, unless an expert believes that every article ever published in a particular journal is the definitive word on a topic or somewhere in this arena, it is probably a mistake to endorse that journal as a whole as authoritative. Rather, one might say something like, "The X journal contains a number of strong articles and others I do not think are as good, and I would really need to know what article or articles you are referring to in order to tell you whether I think they are authoritative or if I relied on them in this case." It is also reasonable to say that one respects a certain author, although one does not necessarily agree with everything the writer has said, and would need to know specifically what the attorney is referring to in order to answer a question about authoritativeness or possible use in the case.

Other experts will go to the other extreme, denying that anything is authoritative. Such experts can make a poor courtroom impression because they come across as pushing the view that there is only one person worth listening to on a topic, themselves naturally; jurors tend not to like individuals who act like self-appointed, self-anointed know-it-alls. Also, the lawyer often can still ask general questions about the literature that incorporate the gist of the findings, such as, "Isn't it true there are many studies showing (something contradictory to what the expert has described)? The lawyer might hold up a stack of articles to convey the impression that the assertion about the literature did not emerge from thin air. The expert may not acknowledge the authoritative status of the literature or the findings, and hence the lawyer might not be able to get further, or much further, into the specifics of the particular work. However, the lawyer has been able to bring the findings to the attention of the jury, and the expert's repeated rejections of contrary literature can create a strong suggestion of bias. The cross-examiner also has the option of putting on her own expert, who can acknowledge the existence or status of literature the opposing expert has denied. Finally, it is almost always possible to get certain literature introduced through one or another means. For example, the lawyer will almost always be permitted to ask questions about the manuals for the tests that the expert used.

The attorney will probably ask whether the expert has talked with anyone about the case. If the answer is affirmative, questions will almost surely follow about the identify of the other party or parties and what specifically was discussed; and the attorney might decide to depose one or more of these individuals to get their descriptions. For example, if the defendant psychologist in a malpractice case talked to his supervisor, that supervisor may well be deposed.

The attorney may conduct thorough questioning about credentials. The expert may be asked about courses taken in graduate school, whether an APA-approved internship was completed, and if the expert pursued a post-doctoral fellowship. Questions may also be raised about supervisors and their qualifications, any malpractice claims or ethics complaints, performance on the licensing examination, continuing education activities, and board certification. There are also likely to be questions about research activities and publications. Some experts exclude publications from their resume that might be embarrassing, which tends to make things worse if the attorney uncovers them.

Attorneys frequently ask about fees and fee arrangements, and how charges for legal work compare to those for other activities. In some cases, experts insist that an evaluation was conducted primarily or solely for clinical purposes, but have billed at the much higher rate used for their legal work. The expert may also be asked how often she has been retained by the same attorney in the instant case, and by the attorney's firm, and perhaps what percentage of her income comes from these cases or her legal work in general.

The expert should be cautious about circumstances in which she is unlikely to be paid unless "her side" wins. For example, an expert may have a multi-thousand dollar fee outstanding with an impoverished plaintiff. A cross-examining attorney can bring out the situation and ask the expert whether these financial arrangements might make it difficult to be fully objective. The lawyer probably will care little about what the expert answers because the seed has been planted in the jury's mind, and an expert who denies such a possibility might well appear to be showing the very bias he denies he manifests. It is thus often advantageous to have at least an understanding, if not a written agreement, that the attorney who retained the expert will be responsible for fees (although deposition time will usually be covered by the opposing attorney).

Experts may also be asked about possible alternative explanations for reported problems or difficulties, evidence that might exist for and against each possibility, and the process they followed in making their selections. They may be asked what type of evidence could be uncovered that might alter their selections or opinions. If the expert has not been thorough in reviewing background materials, the answer may contain elements that, in actuality, are present in the file. For example, suppose the expert answers that narcotic abuse could produce the symptoms he observed on his examination, and that the main way of distinguishing this possible cause from the one he identified is temporal sequence or association. It might just be that the plaintiff had obtained prescriptions for pain-killing narcotics from three separate doctors, was filling them simultaneously, and that the alterations the expert described started 3 months after the accident but just days after the visit to the last pharmacy.

The expert may be asked to critique the work and opinions of other experts, including those on her own side of the case. Sometimes the idea is to link a strong expert with a weak one, creating considerable difficulties for the former when the latter fares badly. If the expert says he endorses another expert's work completely, he had better be prepared to live with the approval not only in the present case, but perhaps in future cases as well. For example, failure to criticize weak methods when asked about views on the other expert's work may translate, in effect, into tacit endorsement, which can be brought up again in the next case and may also lead the expert into contradictory positions.

4.19.6.2 Some Suggestions for Depositions

Perhaps the most obvious suggestion, and one that applies similarly to depositions and cross-examination, is to prepare thoroughly. The required preparation varies and may include a review, or re-review, of the case records, relevant scientific publications, and, perhaps, test manuals and related materials. With large files, it would not be unusual to need a half day to refamiliarize oneself with the case, especially if it has been laid aside for some time, and for very large files, a full day might be required. With more complex files, constructing a time line or chronology that summarizes major events or findings can greatly reduce the time that is needed to become reacquainted with the details at a later date.

It is usually sensible to meet with the attorney with whom one is working in advance of the deposition. If not already accomplished, the attorney needs to gain a clear understanding of what the expert can and cannot say, the boundaries of the expert's opinions, and what materials have been reviewed (e.g., the expert may have examined literature on his own that was not part of the attorney's case file). The expert also may need to be informed about particular technical or legal issues. For example, certain materials may have been excluded or deemed inadmissible and should not be referred to when discussing opinions. The expert may also benefit from learning something about the opposing attorney. For example, some attorneys may try to be especially provocative during depositions, hoping the expert will respond angrily and say something foolish.

If a lawyer is not interested in a pre-deposition meeting, it might be time to start wondering about the attorney and the situation that the expert might have gotten himself into. Similarly, the attorney needs to allow (and be willing to pay for) adequate preparation time. For counsel, this is likely to be one case with one expert, and the attorney's more basic obligation is to the client; for the expert, most every case becomes part of his "permanent record." The attorney is unlikely to lose the next case because her (now abandoned) expert in the previous case was trounced, but the consequences for the expert may carry across cases and years. Some experts, however, go overboard, spending too much time on minor or secondary issues or requiring inordinate amounts of time to prepare for depositions. It is hard to say exactly where to draw the line, and matters become much more difficult if disagreements about how to handle and prepare for a case arise late. If the expert feels that the attorney is setting restrictions that compromise professional standards, and if this situation is understood from the beginning, she can turn down the case as outlined. Alternatively, the attorney can be told what limits this could place on the soundness of opinions, and that any such shortcomings would be openly conveyed on a deposition. The attorney can then decide whether to loosen restrictions or retain someone else, the latter option sometimes being best for both parties.

Deposition questions, especially those that are well articulated, frequently call for specific answers and not dissertations. For example, if the attorney asks, "What is contained in your file?" or, "Did you consult any scientific literature when working on this case?" an exegesis on the state of psychology is not required. It is often unwise to answer vague, unclear, or overly general questions. For example, the attorney may ask, "Do you believe Smith suffered a head injury?" It may be impossible to tell whether the question refers to any type of injury to the head (e.g., a facial laceration), or specifically to the brain. If one goes ahead and answers "yes," an attempt at trial to explain that one thought the question referred to any type of head injury, and not solely to a brain injury, may fall on deaf ears (assuming the expert ever gets a chance to attempt an explanation). Responding to overly general questions can cause similar problems. For example, the attorney may ask, "Isn't it true that individuals with PTSD show difficulties with interpersonal relationships?"

When problematical questions are raised, one can simply say that the question is difficult to answer as stated and explain why, e.g., " ... because the answer differs depending on the specifics." If the question was not worded artfully but the expert thinks she knows what was intended, it is reasonable to respond with something like, "I understand your question to mean .... Assuming this, then ..." The lawyer, of course, can stop the expert if the rephrasing distorts the intended meaning (although he might prefer the question the expert has created over his original one).

Sometimes single or seemingly small changes in wording can ruin a question. For example, the lawyer might start by asking, "Is it fair to say that post-traumatic stress syndrome ...," and one might not know if the reference is to the formal diagnostic category or to symptoms that can follow trauma. Depending on which meaning is intended, the answer can change entirely. There is often nothing wrong with explaining what the ambiguity is and why it may be important to clarify the exact reference; and usually a good attorney can quickly discern whether experts are trying to be cooperative and are requesting needed clarification, or rather are being difficult and evasive.

Although it is important to listen to questions very carefully to be sure one knows just what is being asked and to determine whether the question is answerable, one can go too far and make the process miserable for everyone. It can be very difficult for lawyers who are not expert in an area to ask questions with a high level of technical proficiency and exactitude. As long as the question is clear or the expert, through rephrasing, can check to make sure the question is understood (with this assumed meaning memorialized on the record), it should be possible to answer. The relative ease with which a deposition proceeds can depend in large part on the rapport between the opposing lawyer and the expert. If the lawyer is constantly rephrasing answers, attempting to get the expert to endorse a distortion that supports some contrary argument, or if the expert is really being nonresponsive in order to evade the import of questions, the whole process can bog down and become cumbersome, to say the least. If, however, the lawyer is doing his best to be clear and the expert her best to be responsive, and if the two can cooperate in clarifying ambiguities in questions, the process will usually go along reasonably.

The lawyer who has retained the expert will sometimes ask her to approach depositions parsimoniously, restricting herself to the question asked and volunteering no additional or extra information. As a somewhat hyperbolized example, if the lawyer asks, "Did you do anything to confirm allegation A?" the answer is, "Yes," not, "Yes, what I did was ..." Other lawyers will ask experts to answer more fully. They may fear that areas of testimony will be blocked if, upon questioning, opinions and their bases are not adequately explicated. For example, if cursory answers are provided to questions about the literature the expert relied on, she may be prohibited from discussing that literature at trial. After all, a fundamental purpose of a deposition is to learn the underlying bases for an expert's opinion, or to avoid courtroom surprises, so that the opposing lawyer can prepare for trial. Other lawyers want more complete answers because they hope or expect that a show of strength will give them settlement leverage.

Whatever the response style adopted, it is questionable to avoid an answer by exploiting trivial technical problems with the wording when the expert really knows what the opposing attorney is asking. I attended one trial in which an expert was asked, "Is there board status in neuropsychology?" The context of this question almost surely made it apparent to the expert that the lawyer meant board certification. However, because the lawyer did not ask the question exactly right and the expert wanted to avoid losing ground, the response was, "No." Aside from such a response style being disingenuous and arguably obstructing the legal process, getting caught can be costly. In the case in question, during a break, the lawyer recognized his mistake. He went back and made it very obvious to the jury that the expert had known exactly what the attorney was asking and had avoided answering on a technicality. This made the answer hurt 10 times more than it ever would have had it been surrendered earlier. Those less familiar with the trial process may be surprised by the magnitude of the consequences that can follow when an expert is caught in an intentional misrepresentation, whether it is a direct lie or the offspring of evasion.

Uncertainty about the answer to a deposition question should be readily conceded. One cannot know everything, and such a concession is usually far less injurious than a wrong guess, or repeated wrong guesses. Lawyers may try to bait experts by insinuating in some way, perhaps simply by tone of voice or suggestive words, that they are asking about something that is very basic and that any professional would be expected to know. "Doctor, are you familiar with the large number of studies on . . .?" "Isn't one of the most frequently demonstrated findings in your field . . ." The lawyer may be hoping that the expert will make guesses, because with each guess there is an increased risk of error. If an expert guesses 10 times and is wrong three times, it is not hard to anticipate which three topics will be raised on cross-examination. The expert should not feel that she has to be absolutely certain about everything she utters in a deposition, but if the level of uncertainty passes some relatively low threshold, it might at least be noted (e.g., "If I remember right . . ."), or one should just say that one is not sure and does not want to hazard a guess. (Of course, if these types of uncertainties apply to opinions, and especially if the expert is merely guessing or almost tossing coins, he probably does not belong in the courtroom in the case; and one hopes the attorney who has retained him will not learn about this for the first time during the deposition.) Similarly, it is extremely risky (and potentially unethical) to make some claim that cannot be supported, e.g., "Despite the many negative studies on experience and accuracy, there are far more studies showing a positive relation between the two."

Depositions require a high level of concentration and can strain endurance. If one is too tired to pay close attention to questions, it is at least time to take a break, or to call a halt to the process. Stopping can be cumbersome if expensive travel arrangements are involved, but poor answers on depositions can destroy cases and haunt experts for years, or forever. Experts also should not let their guard down, something that is more likely to occur with fatigue. An off-handed comment on a break may be the first topic raised when the deposition is resumed. It may be particularly hard to maintain vigilance with a highly personable, friendly attorney. Whether or not the attorney is an honest, good-willed, considerate, and kind individual, her job is to win the case and she is likely to use anything she can (within professional ethics and the law) against the expert. Thus, at least within the context of the litigation, that attorney may best be thought of as someone for whom you are the enemy.

Lawyers will sometimes spring new sources of information on experts during depositions, such as publications, medical reports, or documents relating to the plaintiff's everyday functioning, such as work records. If one is unfamiliar with the material, it calls for an open admission. If the lawyer wishes to ask questions, the expert should request the time needed to review the material. If it is not possible to perform an adequate review on the spot, the deponent should so indicate. Alternatively, if the expert believes she has acquired a reasonable grasp of the new material, she might state on the record that she will try to be helpful by answering questions about documents that she has just seen for the first time, although upon more careful study and reflection impressions or conclusions might change. Keep in mind that new materials might be presented out of context, which can lead to misimpressions. For example, one may come to learn that a letter describing insubordination as a basis for termination was written by a boss subsequently arrested for criminal behavior that endangered the public, and that it was the plaintiff who had heroically reported the misdeeds.

4.19.6.3 Trial Testimony

Prior to trial testimony, as with depositions, the expert typically should meet with the attorney. The meeting can address the topics to be covered on the direct examination and what might occur on cross-examination. Once again, the lawyer needs to understand what the expert can and cannot say and any reservations and uncertainties pertinent to the expert's views. Almost any attorney would want to know about these limitations and problems in advance, preferably as soon as possible, rather than discovering them during the expert's cross-examination. In one case, an expert testified that the plaintiff would have developed a particular disorder whether or not an accident had occurred. What he did not tell the attorney that retained him was that he believed it may have taken as long as 10 more years for the condition to develop if the traumatic accident in question had not occurred, something that came out at the end of the cross-examination. Had the lawyer known about this qualification in advance, he would have offered considerably more money to settle the case.

For each topic to be covered on direct, the lawyer needs to be able to ask the expert a question, or questions, that permit entry into the subject matter. If a sufficiently precise question is not articulated, the expert may not even understand what the lawyer is asking, or may have no way to connect the question to the intended material. The lawyer can try out questions in advance to see if they prove sufficient to elicit the intended topic. An expert also needs to prepare for the possibility that upon objection, the judge may preclude queries or entire lines of questioning, which again emphasizes the need for expert and attorney to have a good idea about the topics to be covered. If the attorney does not really have this appreciation and is operating more by rote, unexpected alterations in the direct due to a judge's rulings or small slip-ups can be extremely disruptive.

Like a decent lecture, testimony is usually more effective if it is straightforward, accessible, unfettered with needless and endless complications, and not too lengthy. The jury may already have had a long day by the time the expert goes on, and, often, the expert's testimony is just one piece, although perhaps a crucial one, of a much larger composite that may include long strings of witnesses and days of evidence. A complex, obtuse, and exceedingly detailed presentation may quickly bore the jury and lead to inattention. This is not to argue for glossing over crucial points or simplifying at the cost of distortion. It is often very important to present the underlying bases for opinions, rather than just the conclusions, and in some instances detailed analysis and explanation are needed. For example, when discussing the use of a psychological instrument in malingering detection, a relatively detailed description of scales and their rationale may be necessary. Nevertheless, careful analysis will show that in most cases one's testimony revolves around only a few major points, many details are not of particular importance, and the direct can usually be limited to 1 to 2 hours or less. Also, every point raised in direct is potential material for cross-examination. Thus, unneeded complications or content are not neutral and primarily have a down side.

Visual displays can strengthen an expert's direct considerably by increasing interest and clarity. One might consider using at least one visual aid for each major point covered. It is usually preferable to keep demonstratives basic. A rule of thumb is to ask whether a visual aid is nearly, or completely, self-interpreting and can be comprehended quickly or with minimal explanation. Visual materials do not need to be fancy or elaborate, only clear. For example, a memory disorder in which rehearsal yields minimal gains in new learning can be easily illustrated through a learning curve. For visual displays, I often use overheads because the equipment is readily available, materials are inexpensive, materials can be blown up to sufficient size with little distortion, and one can keep the lights up in the room, a real advantage late in the day.

Some experts try to impress juries with their technical knowledge. In particular, their presentations are laden with jargon. Although some technical terms may be necessary and have exact meanings that are otherwise difficult to capture succinctly (e.g., "post-traumatic amnesia"), other terms are pedantic and contribute little. Overuse of jargon is likely to alienate jurors.

On direct and cross-examination, the expert should remember which audience is the important one. It is not either counsel, but the trier of fact—the judge or jury—that will make the ultimate decisions and that the expert is there to address. If the opposing attorney hates you, or acts like you are an object of disdain, it does not influence the outcome of the trial an iota if the jury feels otherwise. Answers often should be directed at the jury, that is, one should look at the jurors and speak to them.

Many points about cross-examination have already been covered, and only a few additions will be provided here. Cross-examination, as noted, often does not follow the outline of the direct examination, so that the expert should not be surprised if completely different topics are raised. Depositions may give helpful clues about at least some of the upcoming points of attack, but there will almost always be some unanticipated questions at trial. As with depositions, or more so, preparation is extremely important. In addition to a pre-trial meeting with the attorney, one should be familiar, or very familiar, with the file, and with key background literature.

It is exceptional to get through an entire cross-examination untouched, and few cases that get to trial are so one-sided that at least some reasonable counterarguments to the expert's opinions cannot be raised. In the end, if the expert has done a decent job on direct, and if the attorney scores some points on the cross-examination but misses as or more often, it is the cross-examiner who has probably lost considerable ground. One of the worst characteristics of some witnesses is to not concede anything, even the obvious, as if the loss of even one exchange is intolerable and completely nullifying. An expert who is defensive and unreasonable, and who does not make what would seem to be required concessions loses credibility. If the attorney asks, in the context of *DSM-IV* criteria for malingering, "Would you agree that lying on the interview could be viewed as lack of cooperation with the examination?" the simple answer would seem to be, "Yes." The attorney has not asked whether the expert thinks the plaintiff lied or whether the plaintiff was malingering. An expert might try to rush in to make these points, but a good attorney will simply say something like, "Doctor, I don't think you've answered my question. Please listen carefully," and then will slowly repeat the exact same question. If the expert again tries to dance around the question, she will begin to look evasive. When it seems to the jury that the expert would vehemently disagree should the cross-examiner assert that 1 + 1 = 2, he becomes an object of ridicule.

The expert's responsibility is to conduct herself competently, professionally, and ethically, not to win or lose the case, and not to influence the outcome through sleight of hand. If honest concessions lead the jury to reject the expert's opinion, so long as these concessions do not stem from avoidable error or negligence (e.g., scoring errors), it does not mean that the expert failed to perform in a respectable and worthy manner. It may well be that the jury reached the right decision and that justice has been done. The issues at stake in the courtroom are often of great personal import, and one hopefully would much prefer to have one's opinion rejected than to prevail if, in truth, one was wrong. Almost every attorney wins and loses cases, and a "bad" outcome will not necessarily lead them to develop negative views towards an expert. This is unlike a situation in which an expert has been dishonest with the attorney, by withholding weaknesses in his opinions, for example, and gets caught on the stand. Unfortunately, in some such instances, the expert's actions may have destroyed the chances that the meritorious party will prevail.

4.19.7 REFERENCES

American Psychological Association (1985). *Standards for educational and psychological testing*. Washington, DC: Author.

Arkes, H. R. (1981). Impediments to accurate clinical judgments and possible ways to minimize their impact. *Journal of Consulting and Clinical Psychology, 49,* 323–330.

Bennett, B. E., Bryant, B. K., VandenBos, G. R., & Greenwood, A. (1990). *Professional liability and risk management*. Washington, DC: American Psychological Association Press.

Bersoff, D. N. (1997). *The application of Daubert to forensic and social science evidence*. Presentation to the Federal Judicial Center's National Workshop for Magistrate Judges, Denver, CO.

Bolla-Wilson, K., & Bleecker, M. L. (1986). Influence of verbal intelligence, sex, age, and education on the Rey Auditory Verbal Learning Test. *Developmental Neuropsychology 2,* 203–211.

Brandon, A. D., Chavez, E. L., & Bennett, T. L. (1986). A comparative evaluation of two neuropsychological finger tapping instruments: Halstead–Reitan and Western Psychological Services. *International Journal of Clinical Neuropsychology, 8,* 64–65.

Brodsky, S. L. (in press). *A hierarchical-conflict model of ethics in expert testimony*. Unpublished manuscript.

Brodsky, S. L. (1991). *Testifying in court. Guidelines and maxims for the expert witness*. Washington, DC: American Psychological Association Press.

Daubert v. Merrell Dow Pharmaceuticals, Inc. (1993). 509 US 579, 113 S. Ct. 2786, 125 L ed 2d 469.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243,* 1668–1674.

Faust, D. (1984). *The limits of scientific reasoning*. Minneapolis, MN: University of Minnesota Press.

Faust, D. (1986). Research on human judgment and its application to clinical practice. *Professional Psychology, 17,* 420–430.

Faust, D. (1993). Use and then prove, or prove and then use? Some thoughts on the ethics of mental health professionals' courtroom involvement. *Ethics and Behavior, 3,* 359–380.

Faust, D. (1995). Neuropsychological (brain damage) assessment. In J. Ziskin (Ed.), *Coping with psychiatric and psychological testimony* (5th ed., Vol. 2, pp. 916–1044). Los Angeles: Law and Psychology Press.

Faust, D, & Ackley, M. A. (1998). Did you think it was going to be easy? Some methodological suggestions for the investigation and development of malingering detection techniques. In C. R. Reynolds (Ed.), *Detection of malingering during head injury litigation* (pp. 1–54). New York: Plenum.

Faust, D., & Meehl, P. E. (1992). Using scientific methods to resolve questions in the history and philosophy of science: Some illustrations. *Behavior Therapy, 23,* 195–211.

Faust, D., & Willis, W. G. (in press). *Counterintuitive imperatives: A guide to improving clinical assessment and care by predicting more accurately*. Boston, Allyn & Bacon.

Faust, D., Ziskin, J., & Hiers, J. B., Jr. (1991). *Brain damage cases: Coping with neuropsychological evidence* (Vols. 1 & 2). Los Angeles: Law and Psychology Press.

Faust, D., Ziskin, J., Hiers, J. B., Jr., & Miller, W. J. (in press). Revision and update of Faust, D., Ziskin, J., & Hiers, J. B., Jr. (1991). *Brain damage cases: Coping with neuropsychological evidence* (Vols. 1 & 2). Los Angeles: Law and Psychology Press.

Frye v. US (DC Cir. 1923). 293 Fed. 1013, 1014.

Garb, H. N. (1989). Clinical judgment, clinical training, and professional experience. *Psychological Bulletin, 105,* 387–396.

Greene, R. L. (1991). *The MMPI-2/MMPI: An interpretive manual*. Boston: Allyn & Bacon.

Greene, R. L. (1997). Assessment of malingering and defensiveness by multiscale inventories. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 169–207). New York: Guilford.

Grisso, T. (1986). *Evaluating competencies*. New York: Plenum.

Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychology, Public Policy, and Law, 2,* 293–323.

Hathaway, S. R., & McKinley, J. C. (1951). *MMPI manual*. New York: Psychological Corporation.

Hathaway, S. R., & McKinley, J. C. (1983). *Manual for administration and scoring of the MMPI*. Minneapolis, MN: National Computer Systems.

Heaton, R. K., Matthews, C. G., & Grant, I. (1991). *Comprehensive norms for an expanded Halstead–Reitan Battery*. Odessa, FL: Psychological Assessment Resources.

Matarazzo, J. D., Daniel, M. H., Prifitera, A., & Herman, D. O. (1988). Inter-subtest scatter in the WAIS-R standardization sample. *Journal of Clinical Psychology, 44,* 940–950.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis, MN: University of Minnesota Press.

Meehl, P. E. (1973). *Psychodiagnosis. Selected papers.* Minneapolis, MN: University of Minnesota Press.

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin, 52,* 194–216.

Pankratz, L. (1988). Malingering on intellectual and neuropsychological measures. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 169–192). New York: Guilford.

Reitan, R. M., & Wolfson, D. (1993). *The Halstead–Reitan Neuropsychological Test Battery. Theory and clinical interpretation* (2nd. ed.). S. Tucson, AZ: Neuropsychology Press.

Rey, A. (1964). *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.

Reynolds, C. R. (1998). Preface to C. R. Reynolds (Ed.), *Detection of malingering during head injury litigation* (pp. vii–ix). New York: Plenum.

Rogers, R. (Ed.) (1997a). *Clinical assessment of malingering and deception* (2nd. ed.). New York: Guilford.

Rogers, R. (1997b). Structured interviews and deception. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 301–327). New York: Guilford.

Rogers, R., Bagby, R. M., & Dickens, S. E. (1992). *SIRS. Structured Interview of Reported Symptoms. Professional manual*. Odessa, FL: Psychological Assessment Resources.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale-Revised.* New York: Psychological Corporation.

Wechsler, D. (1987). *Manual for the Wechsler Memory Scale-Revised.* New York: Psychological Corporation.

Wedding, D., & Faust, D. (1989). Clinical judgment and decision making in clinical neuropsychology. *Archives of Clinical Neuropsychology, 4,* 233–265.

Wiens, A. N., Crossen, J. R., & McMinn, M. R. (1988). Rey Auditory–Verbal Learning Test: Development of norms for healthy young adults. *Clinical Neuropsychologist, 2,* 67–87.

Ziskin, J. (1995). *Coping with psychiatric and psychological testimony* (5th ed., Vols. 1–3). Los Angeles: Law and Psychology Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.01

Developmental-contextual Considerations: Person–Context Relations as the Bases for Risk and Resiliency in Child and Adolescent Development

RICHARD M. LERNER, MARY E. WALSH, and KIMBERLY A. HOWARD
Boston College, Chestnut Hill, MA, USA

5.01.1 INTRODUCTION 1
5.01.1.1 The Role of Developmental Contextual Theory 2
5.01.2 AN OVERVIEW OF DEVELOPMENTAL CONTEXTUALISM 3
5.01.2.1 Change and Relative Plasticity 4
5.01.2.2 Relationism and the Integration of Levels of Organization 4
5.01.2.3 Historical Embeddedness and Temporality 5
5.01.2.4 The Limits of Generalizability, Diversity, and Individual Differences 5
5.01.3 BASES AND IMPLICATIONS OF DEVELOPMENTAL CONTEXTUALISM AND CHILD AND ADOLESCENT DEVELOPMENT 6
5.01.3.1 Children's Influences on Their Own Development 6
5.01.3.2 Development as a Life-span Phenomenon 7
5.01.3.2.1 Circular functions and bidirectional socialization 8
5.01.3.2.2 The goodness-of-fit model 8
5.01.3.3 Development in its Ecological Context 9
5.01.3.3.1 Levels of embeddedness in the ecology of human development 10
5.01.4 IMPLICATIONS FOR METHOD AND APPLICATION 12
5.01.4.1 Clinical Application 14
5.01.4.2 The Need for Collaborative Research and Outreach 17
5.01.4.3 Implications for Policies and Programs 18
5.01.5 CONCLUSIONS AND FUTURE DIRECTIONS 19
5.01.6 REFERENCES 20

5.01.1 INTRODUCTION

Neither behavioral and emotional problems of children and adolescents, nor instances of positive functioning and healthy development that occur for these age groups, arise from biogenic, psychogenic, or sociogenic sources acting independently on the individual

(Lerner, 1986, 1995; Lerner, Hess, & Nitz, 1990, 1991). Rather, both problematic and healthy instances of behavior and development during childhood and adolescence—and indeed across the entire life-span—arise from the dynamic interaction (Lerner, 1978) or fusion (Tobach & Greenberg, 1984) of variables from multiple levels of organization. These levels include the biological (e.g., the genetic or hormonal), psychological (e.g., the cognitive or personological), social (e.g., peer group and family relational), and societal (e.g., educational, political, or economic institutional); and also involve the cultural, the designed and natural physical ecological, and the historical levels (Bronfenbrenner, 1979, in preparation; Elder, Modell, & Parke, 1993; Ford & Lerner, 1992; Lerner, 1986, 1991, 1995).

Thus, both the emergence of behavioral and emotional problems, and the risk for development of such problems, lies in the changing relations among multiple levels of organization that comprise the context, or the ecology, of human life (Bronfenbrenner & Morris, 1998; Lerner, 1995; Lerner et al., 1990). Similarly, development of positive, healthy behaviors (i.e., behaviors that meet, or fit, the behavioral, physical, or social demands of the context, and are thus able to be construed as "adaptive," or adjusted; Lerner & Lerner, 1983), and resistance to, or resiliency from, threats to development of such behaviors lie in integrations across time of biological-through-historical levels of organization, that is, in the micro and macro systems comprising the ecology of human behavior and development (Bronfenbrenner, 1979).

Accordingly, any attempts to describe or explain the problematic vs. healthy status of behavior, and/or the bases or actualization of risk or resiliency in human life, require reference to the relations between individuals and the contexts of human development. Simply, then, behavior or health problems, and risks and resiliency, require integrated consideration of people and the context within which they develop (Lerner, 1991, 1995, 1997; Sameroff, 1989). A "developmental contextual" (Lerner, 1986; Lerner & Kaufmann, 1985) theoretical frame is needed to understand, appraise, and devise policies and programs that promote positive human behavior and development (Lerner, 1995).

5.01.1.1 The Role of Developmental Contextual Theory

Developmental contextualism is a theory of human development that has direct importance for understanding the bases, assessment, and treatment of behavioral and emotional problems in childhood and adolescence (Lerner, 1986, 1995; Lerner et al., 1990, 1991). Behavioral/emotional problems involve the malfunctioning of the human psychosocial system (Magnusson & Öhman, 1987). Thus, development of behavioral/emotional problems involves the emergence of nonadaptive, nonfit, or inadequate characteristics in this psychosocial system. From a developmental contextual perspective, behavioral/emotional problems exist and develop when, and only when, an inadequate relation exists between psychological and social functioning, that is, between individuals and their context. It is a malfunctional person–context relation which defines problematic behavior and development; and it is from the changing nature of this relation—occurring normatively or through remedial, preventive, or enhancing interventions—that either healthy/positive or maladaptive, and hence negative, nonfit, or problematic behaviors develop (Sameroff, 1989).

We should emphasize that this relational, developmental contextual view is compatible with those of other developmentalists studying either clinical problems in childhood and adolescence and/or developmental psychopathology (e.g., Cicchetti, 1984, 1987; Magnusson, 1988; Magnusson & Öhman, 1987; Super, 1987; Wapner & Demick, 1988). To illustrate, Cicchetti (1987), writing about psychopathology during human infant development, notes that:

> it is inappropriate to focus on discrete symptomatology to infer the presence of nascent or incipient infant psychopathology. Rather, disorders in infancy are best conceptualized as relational psychopathologies, that is, as consequences of dysfunction in the parent–child–environment system. (p. 837)

A similar stress on the relational and developmental contextual character of psychopathology is made by Super (1987), who considers "disorders of development as a class of phenomena that reveal something of the nature of our children and also of the nature of the worlds we make for them" (p. 2). Indeed, he notes that:

> The settings, customs, and psychology of caretakers not only regulate the healthy emergence of human potential, they also shape the possibilities of disorder. At every stage of the etiology, expression, course, intervention, and final outcome of developmental problems, human culture structures the experience and adjusts the odds. (Super, 1987, p. 7)

In a corresponding vein, Öhman and Magnusson (1987) indicate that:

> any disease must be understood as resulting from a multiplicity of causal events spanning observational levels from biochemistry to sociology. Thus, any single-factor, linear causal model is unrealistic for understanding disease in general, and mental disorders in particular. Instead, multiple interacting causal factors that often change in their effect over time must be postulated. (p. 18)

(see also Magnusson, 1988; Magnusson & Allen, 1983; Super & Harkness, 1982, 1986).

Given, then, the apparent prominence of ideas pertinent to a developmental contextual perspective for understanding development across the life-span and, for present purposes, the pertinence of this perspective for conceptualizing the developmental, contextual, and individual-ecological relational bases of child and adolescent problem behaviors, or resiliency to the emergence of such problems, it is useful to discuss the features of developmental contextualism.

5.01.2 AN OVERVIEW OF DEVELOPMENTAL CONTEXTUALISM

Children and adolescents, and their families, communities, and societies develop; they show systematic and successive changes over time (Lerner, 1986). These changes are interdependent. Changes within one level of organization, for example, developmental changes in personality or cognition within the individual, are reciprocally related to developmental changes within other levels, such as those involving changes in caregiving patterns or spousal relationships within the familial level of organization (e.g., Hetherington, Lerner, & Perlmutter, 1988; Lerner, 1984; Lerner & Spanier, 1978; Lewis, 1997; Lewis & Rosenblum, 1974).

Moreover, reciprocal changes among levels of organization are both products and producers of reciprocal changes within levels. For example, over time, parents' "styles" of behavior and of rearing influence children's personality and cognitive functioning and development; in turn, interactions between personality and cognition constitute an emergent "characteristic" of human individuality that affects parental behaviors and styles and the quality of family life (e.g., Baumrind, 1995; Lerner, 1978, 1982; Lerner & Busch-Rossnagel, 1981; Lerner, Castellino, Terry, Villarruel, & McKinney, 1995; Lewis, 1997; Lewis & Rosenblum, 1974).

Within developmental contextualism, levels are conceived of as integrative organizations. That is:

> the concept of integrative levels recognizes as equally essential for the purpose of scientific analysis both the isolation of parts of a whole and their integration into the structure of the whole. It neither reduces phenomena of a higher level to those of a lower one, as in mechanism, or describes the higher level in vague nonmaterial terms which are but substitutes for understanding, as in vitalism. Unlike other "holistic" theories, it never leaves the firm ground of material reality ... The concept points to the need to study the organizational interrelationships of parts and whole. (Novikoff, 1945, p. 209)

Moreover, Tobach and Greenberg (1984) have stressed that:

> the interdependence among levels is of great significance. The dialectic nature of the relationship among levels is one in which lower levels are subsumed in higher levels so that any particular level is an integration of preceding levels ... In the process of integration, or fusion, *new* levels with their own characteristics result. (p. 2)

The course of human development is the product of the processes involved in the "dynamic interactions" (Lerner, 1978, 1979, 1984) among integrative levels.

In emphasizing that systematic and successive change (i.e., development) is associated with alterations in the dynamic relations among structures from multiple levels of organization, developmental contextualism provides a scope for contemporary developmental research and intervention that is not limited by (or, perhaps better, confounded by an inextricable association with) a unidimensional portrayal of the developing person (e.g., the person seen from the vantage point of only cognitions, or emotions, or stimulus–response connections, or genetic imperatives; e.g., see Piaget, 1970; Freud, 1949; Bijou & Baer, 1961; and Rowe, 1994, respectively). Rather, the power of the stress in developmental contextualism on processes of dynamic person–context relations is the "design criteria" imposed on research, method, and intervention pertinent to the study of any content area or dimension of the developing person. This power is constituted by four interrelated assumptive dimensions of developmental contextual theory (Lerner, 1998). Accordingly, it is useful to discuss these dimensions in order to illuminate the key theoretical and methodological (e.g., research design and measurement) issues pertinent to understanding how biological, psychological, and contextual processes combine to promote either healthy or problematic behavior and development across the life-span and also within the periods of childhood and adolescence.

5.01.2.1 Change and Relative Plasticity

The central idea in developmental contextualism is that changing, reciprocal relations (or dynamic interactions) between individuals and the multiple contexts within which they live comprise the essential process of human development (Lerner, 1986; Lerner & Kauffman, 1985). Accordingly, developmental contextualism stresses that the focus of developmental understanding must be on systematic change (Ford & Lerner, 1992). This focus is required because of the belief that the potential for change exists across the life-span (e.g., Baltes, 1987). Although it is also assumed that systemic change is not limitless (e.g., it is constrained by both past developments and by contemporary contextual conditions), developmental contextualism stresses that relative plasticity exists across life (Lerner, 1984). The resilient behavior in children and adolescents provides a concrete instance of such plasticity (Masten, 1989; Werner, 1990, 1995).

There are important implications of relative plasticity for the application of development science to address problems of human behavior and development. For instance, the presence of relative plasticity legitimates a proactive search across the life-span for characteristics of people and of their contexts that, together, can influence the design of policies and programs promoting positive development (Birkel, Lerner, & Smyer, 1989; Fisher & Lerner, 1994; Lerner & Hood 1986).

5.01.2.2 Relationism and the Integration of Levels of Organization

From a developmental contextual perspective, human behavior is both biological and social (Featherman & Lerner, 1985; Tobach & Schneirla, 1968). In fact, no form of life as we know it comes into existence independent of other life. No animal lives in total isolation from others of its species across its entire life-span (Tobach, 1981; Tobach & Schneirla, 1968). Biological survival requires meeting demands of the environment or, as we note later, attaining a goodness of fit (Chess & Thomas, 1984; Lerner & Lerner, 1983, 1989; Thomas & Chess, 1977) with the context. Because this environment is populated by other members of one's species, adjustment to (or fit with) these other organisms is a requirement of survival (Tobach & Schneirla, 1968).

Developmental contextualism emphasizes that the bases for change, and for both plasticity and constraints in development, lie in relations that exist among multiple levels of organization that comprise the substance of human life (Ford & Lerner, 1992; Schneirla, 1957; Tobach, 1981). These levels range from the inner biological level, through the individual/psychological level and the proximal social relational level (e.g., involving dyads, peer groups, and nuclear families), to the sociocultural level (including key macro-institutions such as educational, public policy, governmental, and economic systems) and the natural and designed physical ecologies of human development (Bronfenbrenner, 1979; Riegel, 1975). These levels are structurally and functionally integrated, thus requiring a systems view of the levels involved in human development (Ford & Lerner, 1992; Sameroff, 1983; Smith & Thelen, 1993; Thelen & Smith, 1994).

Developmental contextualism (Lerner, 1986, 1991, 1995) is one instance of such a developmental systems perspective and, as such, promotes a relational unit as a requisite for developmental analysis (Lerner, 1991): variables associated with any level of organization exist (are structured) in relation to variables from other levels; qualitative and quantitative dimensions of the function of any variable are also shaped by the relations that the variable has with variables from other levels. Unilevel units of analysis (or components of, or elements in, a relation) are not an adequate target of developmental analysis; rather, the relation itself—the interlevel linkage—should be the focus of such analysis (Lerner, 1991; Riegel, 1975).

Relationism and integration have a clear implication for unilevel theories of development and for approaches to intervention that emphasize a similarly unilevel (e.g., a psychogenic or biogenic) orientation to etiology or "treatment": at best, such approaches are severely limited, and inevitably provide a nonveridical depiction of development, due to their focus on what are essentially main effects embedded in higher order interactions (e.g., see Walsten, 1990); at worst, such approaches are neither valid nor useful. Accordingly, neither biogenic theories or interventions (e.g., ones based on genetic reductionistic conceptions such as behavioral genetics or sociobiology; Freedman, 1979; Plomin, 1986; Rowe, 1994; Wilson, 1975), psychogenic theories or interventions (e.g., behavioristic or functional analysis models; Baer, 1970, 1976; Bijou, 1976; Bijou & Baer, 1961; Skinner, 1938), nor sociogenic theories or interventions (e.g., "social mold" conceptions of socialization; e.g., Homans, 1961; and see Hartup, 1978, for a review) provide adequate frames for understanding or addressing problems of human development. Interventions based on such theories are insufficiently attentive to the potential plasticity that exists in the human development system, and thus the multiple levels of this system that provide "points of entry" for useful interventions (Birkel, Lerner, & Smyer, 1989; Lerner, 1984, 1995).

Developmental contextualism moves beyond simplistic division of sources of development into nature-related and nurture-related variables or processes. Instead, developmental contextualism sees multiple levels of organization that exist within the ecology of human development as part of an inextricably fused developmental system, one that affords formulation of a rich and multilevel array of strategies for developmental interventions into the life course (for example, see the chapters in Lerner, 1998).

5.01.2.3 Historical Embeddedness and Temporality

The relational units of analysis of concern in developmental contextualism are considered "change units" (Lerner, 1991). The change component of these units derives from the ideas that all levels of organization involved in human development are embedded in history, that is, they are integrated with historical change (Elder, 1980; Elder et al., 1993). Relationism and integration mean that no level of organization functions as a consequence of its own, isolated activity (Tobach, 1981; Tobach & Schneirla, 1968). Each level functions as a consequence of its structural integration with other levels (Tobach & Greenberg, 1984). History—change over time—is incessant and continuous, and it is a level of organization that is fused with all other levels. This linkage means that change is a necessary, an inevitable, feature of variables from all levels of organization (Baltes, 1987; Lerner, 1984); in addition, this linkage means that the structure, as well as the function, of variables changes over time.

The continuity of change that constitutes history can lead to both intra-individual (or, more generally, intralevel) continuity or discontinuity in development, depending on the rate, scope, and particular substantive component of the developmental system at which change is measured (Brim & Kagan, 1980; Lerner, 1986, 1988; Lerner & Tubman, 1989). Thus, continuity at one level of analysis may be coupled with discontinuity at another level; quantitative continuity or discontinuity may be coupled with qualitative continuity or discontinuity within and across levels; and continuity or discontinuity can exist in regard to both the processes involved in (or "explanations" of) developmental change and in the features, depictions, or outcomes (i.e., the "descriptions") of these processes (Cairns & Hood, 1983; Lerner, 1986).

These patterns of within-person change pertinent to continuity and discontinuity can result in either constancy or variation in rates at which different individuals develop in regard to a particular substantive domain of development. Thus, any pattern of intra-individual change can be combined with any instance of interindividual differences in within-person change, that is, with any pattern of stability or instability (Lerner, 1986; Lerner & Tubman, 1989). In other words, continuity–discontinuity is a dimension of intra-individual change and is distinct from, and independent of, stability–instability, which involves between-person change, and is, therefore, a group, and not an individual, concept (Baltes & Nesselroade, 1973; Lerner, 1976, 1986).

In sum, since historical change is continuous, temporality is infused in all levels of organization. This infusion may be associated with different patterns of continuity and discontinuity across people. The potential array of such patterns has implications for understanding the importance of human diversity, and design and implementation of interventions into the course of human development.

5.01.2.4 The Limits of Generalizability, Diversity, and Individual Differences

The temporality of the changing relations among levels of organization means that changes that are seen within one historical period (or time of measurement), and/or with one set of instances of variables from the multiple levels of the ecology of human development, may not be seen at other points in time (Baltes, Reese, & Nesselroade, 1977; Bronfenbrenner, 1979; Valsiner, 1998). What is seen in one data set may be only an instance of what does or what could exist. Accordingly, developmental contextualism focuses on diversity: of people, of relations, of settings, and of times of measurement (Lerner, 1991, 1995, 1997).

Individual differences within and across all levels of organization are seen as having core, substantive significance in understanding human development (Baltes et al., 1977; Lerner, 1991, 1995, 1997) and for designing interventions that fit specific needs and characteristics of diverse groups toward which such efforts may be directed. Indeed, diversity is the exemplary illustration of the presence of relative plasticity in human development (Lerner, 1984). Moreover, by inferring that it is possible to generalize from across-person variation that exists within time to within-person variation across time, diversity constitutes the best evidence of the potential for change in the states and conditions of human life (Brim & Kagan, 1980). It underscores the belief that interventions can make a difference in human development.

Moreover, individual structural and functional characteristics constitute an important source of people's development (Lerner, 1982; Lerner & Busch-Rossnagel, 1981). The individuality of each person promotes variation in the fusions he or she has with the levels of organization within which the person is embedded. For instance, people's distinct actions or physical features promote differential actions (or reactions) in others toward them (Lerner, 1987). These differential actions, which constitute feedback to the person, shape, at least in part, further change in the person's characteristics of individuality (Lerner & Lerner, 1989; Schneirla, 1957). For example, the changing match, congruence, or goodness of fit between people's developmental characteristics and their context provide a basis for consonance or dissonance in their ecological milieu; the dynamic nature of this interaction constitutes a source of variation in positive and negative outcomes of developmental change (Lerner et al., 1990, 1991; Lerner & Lerner, 1983; Thomas & Chess, 1977; Thomas, Chess, Birch, Hertzig & Korn, 1963).

The major assumptive dimensions of developmental contextualism—systematic change and relative plasticity, relationism and integration, embeddedness and temporality, generalizability limits and diversity—are very much intertwined facets of a common conceptual core. And, as is also the case with levels of organization that are integrated to form the substance of developmental change, the assumptive dimensions reflect the essential features of superordinate developmental systems views of human development (Ford & Lerner, 1992). As is the case with the several defining features of the life-span developmental perspective, which—according to Baltes (1987)—need to be considered as an integrated whole, the assumptive dimensions of developmental contextualism need to be appreciated simultaneously. Such appreciation is required to understand the breadth, scope, and implications for research and application of this conceptual framework.

5.01.3 BASES AND IMPLICATIONS OF DEVELOPMENTAL CONTEXTUALISM AND CHILD AND ADOLESCENT DEVELOPMENT

Since the early 1970s the study of children and their contexts has evolved in at least three significant directions: changes in the conceptualization of the nature of the person; the emergence of a life-span perspective about human development; and stress on the contexts of development. These trends were both products and producers of developmental contextualism. This perspective has promoted a rationale for a synthesis of research and outreach; a synthesis focused on the diversity of children and on the contexts within which they develop.

Much of the history of the study of human development prior to the mid-1970s was predicated on either organismic or mechanistic (reductionistic) models (Overton & Reese, 1973, 1981; Reese & Overton, 1970). In turn, it is accurate to say that since the 1970s, developmental contextual conceptions have been increasingly prominent bases of scholarly advances in human development theory and methodology (Dixon & Lerner, 1992; Dixon, Lerner, & Hultsch 1991; Lerner, Hultsch, & Dixon, 1983; Riegel, 1975, 1976a, 1976b; Sameroff, 1975, 1983). Indeed, the three above-noted themes in the study of human development define the place of developmental contextualism in theory and research since the late 1970s. Accordingly, it is useful to discuss each of these themes in some detail.

5.01.3.1 Children's Influences on Their Own Development

Children have come to be understood as active producers of their own development (Bell, 1968; Lerner & Spanier, 1978; Lewis & Rosenblum, 1974; Thomas et al., 1963). These contributions primarily occur through the reciprocal relations individuals have with other significant people in their context: for example, children with family members, caregivers, teachers, and peers.

The content and functional significance of the influences that people have on others and, in turn, on themselves, occur in relation to people's characteristics of individuality (Schneirla, 1957). Individual differences in people evoke differential reactions in others, reactions which provide feedback to people and influence the individual character of their further development (Schneirla, 1957). Accordingly, individuality—diversity among people—is central in understanding the way in which any given person is an active agent in his or her own development (Lerner, 1982, 1991; Lerner & Busch-Rossnagel, 1981). Thus, as we have emphasized, diversity has core, substantive meaning and, as such, implications for all studies of human development.

To illustrate these points, it is useful to recall the old adage "the child is father of the man." This means simply that people's characteristics when they are children relate to their characteristics during adulthood. However, there is another way of interpreting this quotation: How we behave and think as adults—and perhaps especially as parents—is very much influenced by our experiences with our children. Our children rear us as much as we do them. Because we are parents, we are different adults than we would be if we were childless. But, more importantly, the specific and often special characteristics of a particular child influence us in unique ways. How we behave toward our children depends to a significant extent on how they have influenced us to behave. Such child influences are termed "child effects."

The presence of such child effects constitutes the basis of bidirectional relations between parents and children. This bidirectional relation continues when the child is an adolescent and an adult. Corresponding relations exist between people and siblings, friends, teachers, and indeed all other significant people in their life. Indeed, this child–other relation is the basic feature of the developmental contextual relations that characterize the social creature we call a human being.

Child effects emerge largely as a consequence of a child's individual distinctiveness. All children, with the exception of genetically identical (monozygotic) twins, have a unique genotype, that is, a unique genetic inheritance. Similarly, no two children, including monozygotic twins, experience precisely the same environment. All human characteristics, be they behavioral or physical, arise from an interrelation of genes and environment (Anastasi, 1958; Lerner, 1986). Given the uniqueness of each child's genetic inheritance and environment, the distinctiveness of each child is assured (Hirsch, 1970; Feldman & Lewontin, 1975). In other words, every child is unique and therefore individually different from every other child.

This individuality may be illustrated by drawing on the study of temperament (Chess & Thomas, 1984; Thomas & Chess, 1977; Thomas et al., 1963). Temperament is a characteristic of children's behavior that describes how they act. For instance, all children eat and sleep. Temperament is the style of eating or sleeping shown by the child; if the child eats the same amount at every meal and/or gets hungry at the same time, then this child has, in regard to eating, a regular, or rhythmic, temperament. A child who gets hungry at different times of the day, or who may eat a lot or a little without any seeming predictability, would, in regard to eating, have an arrhythmic temperament. Similarly, obviously all children sleep. However, some children may sleep irregularly, that is, for seemingly unpredictable (at least to their parents) lengths of time, periods interspersed with wakeful periods of crying and fussing. Simply, one child may be an arrhythmic eater and sleeper, whereas another child might sleep and eat in a more regularly patterned way, and/or when awake may show more smiling than crying and fussing.

The importance of these individual differences arises when we recognize that, as a consequence of their individuality, children will present different stimulation to parents. Children with the above-noted temperamental styles present different stimuli to their parents as a consequence of their respective eating and sleep/wake patterns; the experience for a parent of having a pleasant, regularly sleeping child, who is predictable in regard to eating habits as well, is quite different from the experience for a parent who has a moody, irregularly sleeping and eating child. However, the effect of the child's stimulation of the parent depends in part on the parent's own characteristics of individuality. To explain this point it is useful to consider the second theme in the literature that helped crystallize the developmental contextual view of human development.

5.01.3.2 Development as a Life-span Phenomenon

The emergence of interest during the 1970s and 1980s in a life-span perspective about human development led to the understanding that development occurs in more than the childhood or adolescent years (Baltes, 1968, 1987; Block, 1971; Brim & Kagan, 1980; Elder, 1974, 1980; Featherman, 1983; Riley, 1979; Schaie, 1965). Parents as well as children develop as distinct individuals across life (Lerner & Spanier, 1978). Parents develop both as adults in general and, more specifically, in their familial and extrafamilial (for example, vocational or career) roles (Vondracek, Lerner, & Schulenberg, 1986). Indeed, the influence of children on their parents will depend in part on the prior experience the adult has had with the parental role and on the other roles in which the parent is engaged (e.g., worker, adult–child, and caregiver for an aged parent) (Hetherington &

Baltes, 1988). Thus, people's unique history of experiences and roles, as well as their unique biological (e.g., genetic) characteristics, combine to make them unique, and with time, given the accumulation of the influences of distinct roles and experiences, increasingly more so across the course of life (Lerner, 1988). This uniqueness is the basis of the specific feedback parents give to their individual children.

That is, parents who are stimulated differentially may be expected to differentially react to, or process (e.g., think and feel about) the stimulation provided by their child. A child with an irregular temperament might evoke feelings of frustration and exasperation and thoughts of concern in his or her parents (Brazelton, Koslowski, & Main, 1974; Lewis & Rosenblum, 1974). We might expect, however, that the parents of a child with a predictable and rhythmic style would be better rested than would the parents of an arrhythmic child. When the rhythmic child was awake, they would have a child with a more regularly positive mood, and this too would present less stress on them as parents and as spouses.

The individuality of parental reactions underscores the idea that parents are as individually distinct as their children. Not all parents of an irregularly eating and sleeping, moody child will react with concern and/or frustration. Similarly, some parents will be stressed by even the most regular, predictable, and positive of children. Such parental individuality makes child effects more complicated to study. However, at the same time, parental individuality underscores the uniqueness of each child's context. Simply, then, it may be expected that, as a consequence of the different stimulation received from their children, and in relation to their own characteristics of individuality, parents will provide differential feedback to their children.

Such differential feedback may take the form of different behavior shown to children by parents and/or of different emotional climates created in the home (Brazelton et al., 1974). For instance, parents of an irregular child might take steps to alter his or her eating and sleep/wake patterns. In regard to sleeping, they might try to cut naps short during the day in order that the child be more tired in the evening. In addition, during the time when they are appraising the success of their attempts to put the child on an imposed schedule a general sense of tenseness might pervade the household. "Will we have another sleepless night? Will we be too tired to be fully effective at work?" they might wonder.

The presence of differential feedback by parents of temperamentally distinct children becomes an important part of the child's experience, and this feedback is distinct in that it is based on the effect of the child's individuality on the parent. Thus, the feedback serves to promote the child's individuality further.

5.01.3.2.1 Circular functions and bidirectional socialization

The reciprocal child–parent relations involved in child effects constitute a circular function (Schneirla, 1957) in individual development: children stimulate differential reactions in their parents, and these reactions provide the basis of feedback to the children, that is, return stimulation which influences their further individual development. These circular functions underscore the point that children (and adolescents, and adults) are producers of their own development and that people's relations to their contexts involve bidirectional exchanges (Lerner, 1982; Lerner & Busch-Rossnagel, 1981). Parents shape their children, but part of what determines the way in which they do this is the children themselves.

Children shape their parents—as adults, as spouses, and of course as parents *per se*—and in so doing, children help organize feedback to themselves, feedback which contributes further to their individuality and thus the circular function starts all over again (that is, the child effects process returns to its first component). Characteristics of behavioral or personality individuality allow the child to contribute to this circular function. However, this idea of circular functions needs to be extended; that is, in and of itself the notion is silent regarding specific characteristics of the feedback (e.g., its positive or negative valence) children will receive as a consequence of their individuality. In other words, to account for the specific character of child–context relations the circular functions model needs to be supplemented; this is the contribution of the goodness-of-fit model.

5.01.3.2.2 The goodness-of-fit model

Just as children bring their characteristics of individuality to a particular social setting there are demands placed on them by virtue of the social and physical components of the setting. These demands may take the form of: (i) attitudes, values, or stereotypes that are held by others in the context regarding the person's attributes (either physical or behavioral characteristics); (ii) the attributes (usually behavioral) of others in the context with whom the child must coordinate, or fit, his or her attributes (also, in this case, usually behavioral) for adaptive interactions to exist; or, (iii) the physical characteristics of a setting (e.g., the presence or absence of access ramps for the motorically handicapped) which require the child to possess certain attributes (again, usually behavioral abilities) for the most efficient interaction within the setting to occur.

Children's individuality, in differentially meeting these demands, provides a basis for the specific feedback they get from the socializing environment. For example, considering the demand "domain" of attitudes, values, or stereotypes, teachers and parents may have relatively individual and distinct expectations about behaviors desired of their students and children, respectively. Teachers may want students who show little distractibility, but parents might desire their children to be moderately distractible, for example, when they require their children to move from television watching to dinner or to bed. Children whose behavioral individuality was either generally distractible or generally not distractible would thus differentially meet the demands of these two contexts. Problems of adjustment to school or to home might thus develop as a consequence of a child's lack of match (or goodness of fit) in either or both settings. In turn, interventions aimed at ameliorating or preventing such problems should focus on enhancing fit—on improving the child–context relation—and not on either child or context alone.

The importance of such an approach to intervention is underscored by the research of Thomas and Chess (1977, 1980, 1981) and Lerner and Lerner (1983, 1989). These researchers have found that if a child's characteristics of individuality provide a goodness of fit with demands of a particular setting, adaptive outcomes will accrue in that setting. Those children whose characteristics match most of the settings within which they exist receive supportive or positive feedback from the contexts and show evidence of the most adaptive behavioral development. In turn, of course, poorly fit, or mismatched, children—those whose characteristics are incongruent with one or most settings—appear to show alternative developmental outcomes.

In sum, then, the literature on child effects and on the life-span perspective promote a concern with individual differences; with variation in developmental pathways across life; and with the developmental contextual idea that changing relations (occurring "naturally" or through interventions) between people and their context provide the basis, across life, of the individual's unique repertoire of physical, psychological, and behavioral characteristics (Lerner, 1991). The recognition of this link between person and context was a product and a producer of the third theme emerging in the study of human development since the 1970s.

5.01.3.3 Development in its Ecological Context

The study of children and their parents became increasingly "contextualized," or placed within the broader "ecology of human development," during this period (Bronfenbrenner, 1977, 1979; Elder, 1974; Garbarino, 1992; Pepper, 1942). This focus has involved a concern with the "real life" situations within which children and families exist; it has also led to the study of the bidirectional relations between the family and other social settings within which children and parents function, for instance, the workplace, the welfare office, the day-care center, the Medicaid screening office, and the formal and the nonformal educational and recreational settings present in a neighborhood or a community (Lewis & Feiring, 1978; Lewis & Rosenblum, 1974). Thus, fit between youth and their contexts involves much more than relations between children and their parents and, as we have emphasized, interventions to enhance the congruence between person and context can involve any one or several levels within the ecology of human development.

To understand how the social context contributes to bidirectional person–context relations, we should reiterate that a child is not best thought of as merely similar to all other children, or as simply different from others with respect to only one, or even just a few, characteristics. Instead, individual differences exist with respect to numerous characteristics. We have illustrated how at least some of these changes occur through bidirectional relations children have with their parents. However, we have not yet stressed that parents too are made up of multiple dimensions of individuality which, as with the child, develop across time (e.g., see Baltes, 1987).

Another point not yet illustrated is that both the child and the parent have other social roles. These roles lead both children and parents into social relationships with other groups of people, that is, with other social networks. Parents are also spouses, adult children of their own parents, workers, and neighbors. Children also may be siblings and friends of other children; and as they progress through childhood and, later, adolescence, they become students and often at least part-time employees, respectively. The sorts of relationships in these other social networks in which children and parents engage when outside of their role of child or parent, respectively, can be expected to influence the parent–child relationship (Bronfenbrenner, 1977, 1979).

A child's poor performance at school may influence his or her behavior in the home, and especially, may alter the quality of the parent–

child relationship. In turn, a problematic home situation—as is experienced by children in families wherein parental abuse or neglect of the child occurs—will affect the child's relationships with peers, teachers, and other family members (Baca Zinn & Eitzen, 1993; Belsky, Lerner, & Spanier, 1984).

In regard to parents, strain in the spousal relationship can occur if adults spend too much energy in their parental caregiving role (J. V. Lerner, 1994). For instance, a child's unpredictable sleep pattern and negative mood when awake can severely tax a parents' energy. Energy and time needed for parents to be good spouses to each other therefore may not be available. In addition, the fatigue caused by the demands of parental roles can be expected to influence parents' performance in the workplace (J. V. Lerner, 1994). It is difficult for people who have been up all night caring for a crying infant to be at their best at work the next morning. In turn, of course, problems at work can be brought home. Parents whose energies are significantly depleted outside the home may not have the stamina during the evening to be attentive, patient parents to their children (or attentive, emotionally supportive spouses for their mates).

Thus, bidirectional relationships exist between child and parent (Bornstein & Tamis-LeMonda, 1990; Lerner & Lerner, 1987); and these relationships are in turn reciprocally related to other social networks within which the dyad exists and to the broader societal and cultural context. For example, studying mother–infant dyads in the United States, France, and Japan, Bornstein et al. (1992) found both culture-general and culture-specific patterns of maternal responsiveness to characteristics of infant individuality (e.g., in regard to activity level, exploration, and vocalization). In short, relationships among children, parents, and contexts constitute a complex state of affairs.

One approach to the conceptualization of these relations that we see as particularly useful has been proposed by Bronfenbrenner and his colleagues (e.g., Bronfenbrenner & Crouter, 1983; Bronfenbrenner & Morris, 1998; Garbarino, 1992). Bronfenbrenner's work has been a major catalyst in promoting the contextualization of human development, and in helping us understand why the study of development must move beyond its status during the 1970s, as "the science of the strange behavior of children in strange situations with strange adults for the briefest possible periods of time" (Bronfenbrenner, 1977, p. 513). Bronfenbrenner (1977, 1979, 1983) has argued that human development needs to be understood as it occurs in its real-world setting or "ecology." He believes that this ecology of human development is composed of four distinct although interrelated systems, or types of settings.

The first, the "microsystem" is composed of "the complex of relations between the developing person and environment in an immediate setting containing the person" (Bronfenbrenner, 1977, p. 515). Furthermore, an infant's microsystems may be interrelated. What occurs in day care may affect what happens in the family, and vice versa. The second ecological stratum, termed the "mesosystem," is defined as "the interrelations among major settings containing the developing person at a particular point in his or her life" (e.g., the interaction between day care and family: Bronfenbrenner, 1977, p. 515).

Bronfenbrenner labels the third system as the "exosystem" and defines it as:

> an extension of the mesosystem embracing . . . specific social structures, both formal and informal, that do not themselves contain the developing person but impinge upon or encompass the immediate setting in which the person is found, and thereby delimit, influence, or even determine what goes on there [e.g., the parent's social role as "worker"]. (Bronfenbrenner, 1977, p. 515)

Finally, Bronfenbrenner notes that there exists a "macrosystem" within the ecology of human development. This system is composed of cultural values and beliefs as well as historical events (e.g., wars, floods, famines), both of which may affect other ecological systems.

In short, then, Bronfenbrenner's model of the ecology of human development allows us to devise a means to represent the idea that the bidirectional socialization which occurs between children and parents is embedded in a still more complex system of social networks and of societal, cultural, and historical influences. This system represents a rich set of "targets of opportunity" for the design, delivery, and sustainability of interventions into the course of human development.

5.01.3.3.1 Levels of embeddedness in the ecology of human development

The core idea in developmental contextualism is that the organism (organismic attributes or, most generally, biology) and context cannot be separated (Gottlieb, 1991, 1992; Lerner, 1984; Tobach, 1981). Both are fused across all of life, and thus across history. One way to illustrate just what is involved in this relation, even for one person, is to consider the diagram presented in Figure 1 (see Lerner, 1984, 1986).

It is important to indicate at this point that we may speak of dynamic interactions between

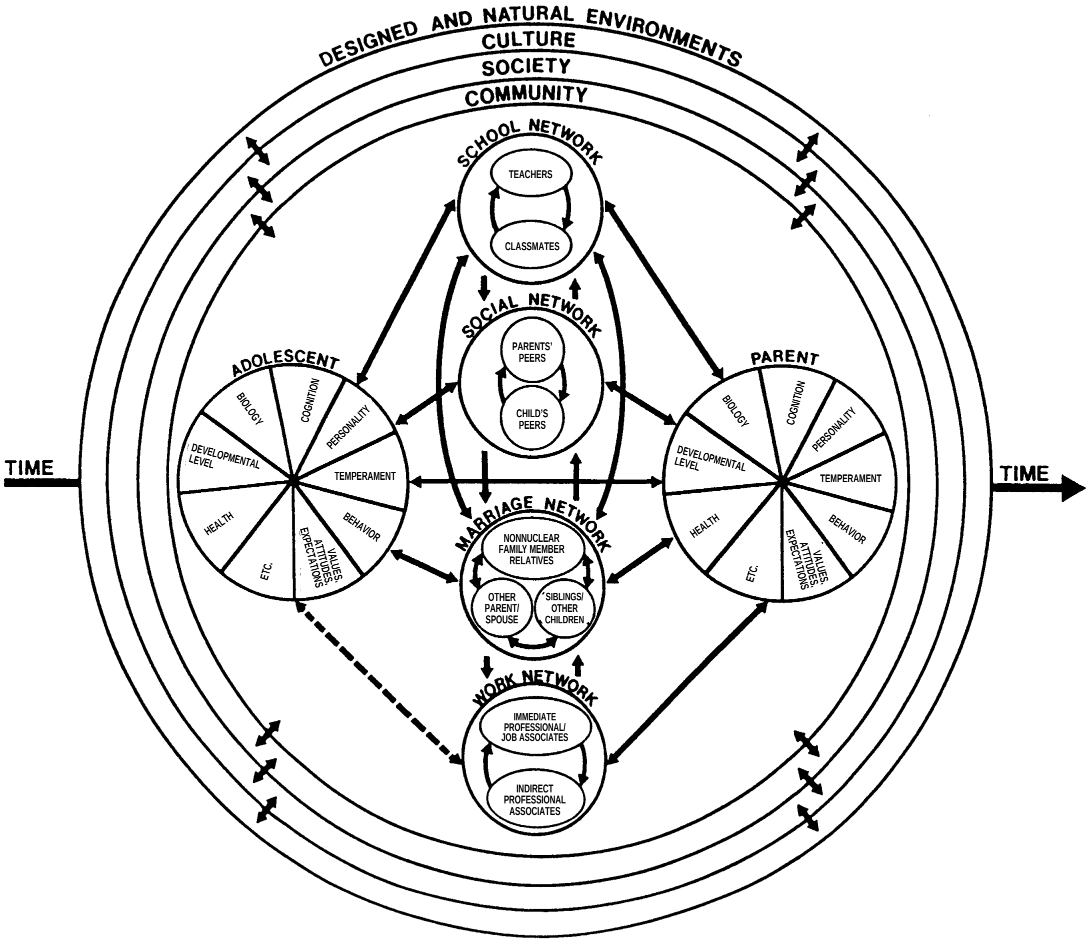
DESIGNED AND NATURAL ENVIRONMENTS
CULTURE
SOCIETY
COMMUNITY
SCHOOL NETWORK
TEACHERS
CLASSMATES
SOCIAL NETWORK
PARENTS' PEERS
CHILD'S PEERS
ADOLESCENT
BIOLOGY
COGNITION
PERSONALITY
DEVELOPMENTAL LEVEL
TEMPERAMENT
BEHAVIOR
VALUES, ATTITUDES, EXPECTATIONS
ETC.
HEALTH
TIME
PARENT
BIOLOGY
COGNITION
PERSONALITY
DEVELOPMENTAL LEVEL
TEMPERAMENT
BEHAVIOR
VALUES, ATTITUDES, EXPECTATIONS
ETC.
HEALTH
TIME
MARRIAGE NETWORK
NONNUCLEAR FAMILY MEMBER RELATIVES
OTHER PARENT/ SPOUSE
SIBLINGS/ OTHER CHILDREN
WORK NETWORK
IMMEDIATE PROFESSIONAL/ JOB ASSOCIATES
INDIRECT PROFESSIONAL ASSOCIATES

Figure 1 The developmental contextual view of human development: adolescent–child relations and interpersonal and institutional networks are embedded in and influenced by particular community, societal, cultural, and designed and natural environments, all changing across time (across history).

parent and child that pertain to either social or physical (for instance, biological or physiological) relations. For example, in regard to social relationships, the parent "demands" attention from the child, but the child does not show it; this "lights" the parent's "short fuse" of tolerance; the parent scolds the child, who then cries; this creates remorse in the parent and elicits soothing behaviors from him or her; the child is calmed, snuggles up to the parent, and now both parties in the relationship show positive emotions and are happy (see Tubman & Lerner, 1994, for data pertinent to such parent–child relationships).

In turn, we may also illustrate dynamic interactions which involve not only the exchange of "external" social behaviors but also involve biological or physiological processes. For example, parental religious practices, rearing practices, or financial status may influence the child's diet and nutritional status, health, and medical care. In turn, the contraction of an infectious disease by either parent or child can lead to the other member of the relationship contracting the disease. Moreover, the health and physical status of the child influences the parents' own feelings of well-being and thus hopes and aspirations regarding the child (Finkelstein, 1993).

Thus, Figure 1 illustrates that within and among each of the networks which is depicted, one may conceive of bidirectional relationships existing among the people populating the network. A child effect may function, in a sense, like a small pebble thrown into a quiet lake. It can prompt a large ripple. In turn, of course, the reverse of this possibility can occur. Events in settings lying far beyond the child–parent relationship can influence it. For instance, the resources in a community for child day care during the parent's working hours, the laws (e.g., regarding tax exemptions) or social programs available in a society supporting day care, and the cultural values regarding families who place their infants in day care, all exert an impact on the quality of the parent–child relationship.

Moreover, as we have just noted, the child–parent relationship, and the social networks in which it is located, are embedded in still larger community, societal, cultural, and historical levels of organization. These relations are illustrated also in Figure 1. Time (history) cuts through all the systems. This feature of the figure is introduced to remind us that, as with the people populating these social systems, change is always occurring. Diversity within time is created as change across time introduces variation into all levels of organization involved in the system depicted in Figure 1.

In other words, all changes are embedded in history (Baltes, 1987; Elder, 1974; Elder et al., 1993), that is, time "cuts through" all levels of organization. As such, the nature of parent–child relations, of family life and development, and of societal and cultural influences on the child–parent–family system are influenced by both normative and non-normative historical changes (Baltes, 1987) or, in other words, by "evolutionary" (i.e., gradual) and "revolutionary" (i.e., abrupt; Werner, 1957) historical changes. This system of multiple, interconnected, or "fused" (Tobach & Greenberg, 1984) levels comprises a complete depiction of the integrated organization involved in the developmental contextual view of human development (Lerner, 1986, 1991).

5.01.4 IMPLICATIONS FOR METHOD AND APPLICATION

In essence, (i) individuality (diversity); (ii) change, involving both the individual and the context; and, as a consequence; (iii) further individuality, are the essential features of human development within developmental contextualism. Given that the multiple levels of change involved in person–context relations may involve individuals at any point in their lives—whether they are infants or young children, on the one hand, or adults (and acting in roles such as parents, spouses, or teachers), on the other—it is possible to see why a developmental contextual perspective provides a useful frame for studying development across the life-span and a rich opportunity to design and implement both individual and group interventions into the course of human life.

The possibility that bidirectional relations exist across the life-span among all the levels shown in Figure 1 represents a formidable state of complexity, but one which behavioral and social science theory, research, and intervention must address. If the scholarship of research and application does not cope with this complexity, then neither research nor intervention will be adequate. That is, research inattentive to the complexity of person–context relations will be deficient in that either it will fail to appreciate the substantive nature of individual, familial, or relationship variation and/or it will mistakenly construe variation around some mean level as, at best, error variance (Lerner, 1991, 1997). In turn, applications—policies and/or intervention programs (that are at least ideally) derived from research (Lerner & Miller, 1993)—will insufficiently fit with the needs of the specific people intended to be served by these interventions, if it is the case that these activities are insufficiently informed by knowledge about the specific characteristics of individuality of these groups. However, developmental contextualism offers an alternative to this situation. It does so by stressing the importance of a focus on diversity and context for integrated research and application.

As we have emphasized, developmental contextualism stresses the bidirectional connections between individuals and the actual ("ecologically valid") settings within which they live. This emphasis has brought to the fore of concern in the social and behavioral sciences both "diversity" (individual differences) and "context" (of peoples and their sociocultural institutions). In addition, we have explained that developmental contextualism emphasizes the relation between individuals and their context, a focus that has resulted in the recognition that a synthesis of perspectives from multiple disciplines is needed to understand the multilevel (e.g., person, family, and community) integrations involved in human development. Furthermore, there has been a recognition that to understand the basic process of human development—the process involved in the changing relations between individuals and contexts—both descriptive and explanatory research must be conducted within the actual ecology of people's lives.

Descriptive research involves the depiction, or representation, of development as it exists for a given person or group, in one or more contexts, at one or more points in time. Explanatory research involves the introduction (through manipulation or statistical modeling) of variation into such person–context relations. These planned variations in the course of human life are predicated on: (i) theoretical ideas about the source of particular developmental phenomena (for specific combinations of people and contexts); or on (ii) theoretically guided interests about the extent to which a particular developmental phenomenon (e.g., cognitive development in the aged years) may show systematic change in structure and/or function, that is plasticity, across the life-span (Baltes, 1987; Lerner, 1984). In either of the above cases such researcher-introduced variation is an attempt to simulate the "natural" variation of life; if theoretical expectations are confirmed, the outcomes of such investigations provide an explanation of how developmental change occurs within a person or group.

Given the developmental contextual focus on studying person–context relations within the actual ecology of human development, explanatory investigations by their very nature constitute intervention research. In other words, the goal of developmental contextual explanatory research is to understand the ways in which variations in ecologically valid person–context relations account for the character of actual or potential trajectories of human development, that is, life paths enacted in the "natural laboratory" of the "real world." Therefore, to gain understanding of how theoretically relevant variations in such person–context relations may influence actual or to-be-actualized developmental trajectories, the researcher may introduce policies and/or programs as, if you will, "experimental manipulations" of the proximal and/or distal natural ecology. Evaluations of outcomes of such interventions become, then, a means to bring data to bear on theoretical issues pertinent to changing person–context relations and, more specifically, to the plasticity in human development that may exist, or that may be capitalized on, to enhance human life (Lerner, 1988). In other words, a key theoretical issue for explanatory research in human development is the extent to which changes—in the multiple, fused levels of organization comprising human life—can alter the structure and/or function of behavior and development.

Of course, independent of any researcher-imposed attempts to intervene in the course of human development, the naturally occurring events experienced by people constantly shape, texture, and help direct the course of their lives. That is, the accumulation of the specific roles and events a person experiences across life—involving normative age-graded events, normative history-graded events, and non-normative events (Baltes, 1987; Baltes, Reese, & Lipsitt, 1980)—alters each person's developmental trajectory in a manner that would not have occurred had another set of roles and events been experienced. The between-person differences in within-person change that exist as a consequence of these naturally occurring experiences attest to the magnitude of the systematic changes in structure and function—the plasticity—that characterizes human life (Lerner, 1984).

Explanatory research is necessary, however, to understand what variables, from what levels of organization, are involved in particular instances of plasticity that have been seen to exist. In addition, such research is necessary to determine what instances of plasticity may be created by science or society. In other words, explanatory research is needed to ascertain the extent of human plasticity or, in turn, the limits of plasticity (Lerner, 1984). From a developmental contextual perspective, conduct of such research requires scientists to alter the natural ecology of the person or group they are studying. Such research may involve either proximal and/or distal variations in the context of human development (Lerner & Ryff, 1978); but, in any case, these manipulations constitute theoretically guided alterations of the roles and events a person or group experiences at, or over, a portion of the life-span.

These alterations are indeed, then, interventions: they are planned attempts to alter the system of person–context relations constituting the basic process of change; they are conducted in order to ascertain the specific bases of, or to test the limits of, particular instances of human plasticity (Baltes, Dittmann-Kohli, & Dixon, 1984; Baltes, Smith, & Staudinger, 1992). These interventions are a researcher's attempt to substitute designed person–context relations for naturally occurring ones, a substitution done in an attempt to understand the process of changing person–context relations providing the basis of human development. In short, then, basic research in human development is intervention research.

Accordingly, the cutting-edge of theory and research in human development lies in the application of the conceptual and methodological expertise of human development scientists to the "natural ontogenetic laboratory" of the real world. Multilevel—and hence, qualitatively and quantitatively multivariate—and longitudinal research methods must be used by scholars from multiple disciplines to derive, from theoretical models of person–context relations, programs of research that involve the design, delivery, and evaluation of interventions aimed at enhancing—through scientist-introduced variation—the course of human development.

In short, in developmental contextualism there is a stress on ontological (and on epistemological, we would add) relationism and contextualization. These emphases have brought to the fore of scientific, intervention, and policy concerns issues pertinent to the functional importance of diverse person–context interactions. Examples are studies of the effects of variations in maternal employment on infant, child, and young adolescent development; the importance of differences in quality day care for the immediate and long-term development in children of healthy physical, psychological, and social characteristics; and the effects of variations in marital role strain and in marital stability–instability on the healthy development of children and adolescents.

Diversity of people and their settings means that one cannot assume that general rules of development either exist for, or apply in the same way to, all children and families. Moreover, one cannot assume, even if only small portions of the total variance in human behavior and development reflect variance that is unique to an individual or group, that this nonshared variance is not the most salient information we have when attempting to understand or enhance the quality of the lives for the person or group.

5.01.4.1 Clinical Application

With its emphasis on the relations between the individual and the context, and on the diversity of people and settings, developmental contextualism has important implications for clinical practice, including: (i) a focus on the dynamic character of behavior and pathology, (ii) specification of the impact of context on behavior, (iii) recognition of the role of strengths and assets, and (iv) utilization of multiple levels of assessment and intervention.

The developmental contextualist framework challenges clinicians to shift from a static focus on classification of the symptoms of psychopathology to a more dynamic orientation in which psychopathology is viewed as a product of the relationship between development and specific environments. The dominant diagnostic system now used by most clinicians, *Diagnostic and statistical manual of mental disorders* (*4th ed.*) (*DSM-IV*) (American Psychiatric Association, 1994), classifies problem behavior in terms of characteristic features of disorders at a single point in time, accounting for neither the development of the disorder nor the context in which it occurs (Achenbach, 1990; Cowan, 1988).

A developmental perspective requires a consideration of the meaning of behaviors at different developmental periods. Behaviors that are "normal" at one age, may be "abnormal" at another (Campbell, 1989; Sroufe, 1990; Sroufe & Rutter, 1984). For instance, enuresis is not uncommon in three- and four-year-old children but is unusual for a 10 year old. Knowing the child's developmental period may aid one in more thoroughly assessing the factors contributing to this behavior. Achenbach (1990) points out that:

> to improve our understanding of maladaptive behavior, it is helpful to view it in relation to normative sequences and achievements for particular ages. When this is done it is evident that many behavioral and emotional problems for which professional help is sought are not qualitatively different from those that most individuals display to some degree at some time in their lives. (p. 4)

A developmental perspective, therefore, encourages clinicians to consider not only what is abnormal behavior or development, but what is normal as well. An understanding of healthy development and the factors that influence both positive and negative outcomes contributes to our understanding of psychopathology (Cicchetti & Garmezy, 1993; Sroufe, 1990; Sroufe & Rutter, 1984).

The influence of development on the interpretation of behaviors is particularly evident in the area of cognitive development (Campbell, 1989). For example, if a five-year-old child whose grandfather has died says that he wants "to go where grandpa is," the clinician interprets the child's wish quite differently than a similar wish expressed by a 14 year old. The clinician recognizes that the 14 year old has a more mature concept of the meaning of death, requiring a different kind of clinical response. The behavior may look similar, but its meaning is very different at these two developmental periods. Similar characteristics displayed at various points in ontogeny may have different implications because they are embedded in a distinct developmental milieu. In short, the accurate assessment of behaviors requires an understanding of appropriate and inappropriate, adaptive and maladaptive, and expected and unexpected behaviors/skills/abilities throughout an individual's development.

An understanding of the development of a disorder, that is, how an individual found himself or herself on a particular developmental path, provides a more complete explanation of an individual's behaviors, thus allowing for a broader range of interventions. A developmental perspective provides guidance in the choice of interventions (Campbell, 1989). For instance, understanding the cognitive-developmental level of a four-year-old child, one would not be likely to use cognitive-behavioral reframing strategies during treatment. Clinicians must plan interventions that are appropriate for children at their current developmental level. Similarly, it is a widely understood phenomenon that older children can experience a greater number of differentiated emotions than a younger child. The child's developing ability to experience and identify emotions will guide practitioners in developing appropriate interventions.

A developmental contextualist framework involves not only a focus on development, but development in relation to context. The traditional *DSM-IV* approach to diagnosis/treatment does not view the behavior as the product of person–context relations. On the contrary, while it "does not claim that every category of disorder represents a mental illness, its classification system has its conceptual roots in the medical model of psychopathology, with its assumption that psychopathology is a disease located somewhere in the patient' (Cowan, 1988, p. 7). Within a developmental contextualist framework, the problem behavior is not located "within the child" (e.g., the child "has" a conduct disorder), but rather is viewed as a disturbance in the relationship or a lack of fit between the child and the settings which impact him or her. Rather than a focus on "fixing the child," developmental contextualism requires the clinician to address the fit or match between the child and the contexts which surround the child (e.g., family, school, community, culture, etc.). From this perspective, behavior is inexplicable when considered in isolation from the context in which it occurs. As an example, consider a five-year-old boy exhibiting behaviors characteristic of separation anxiety, that is, manifesting distress when away from his parents, worrying that he may get hurt if not with his parents, and refusing to sleep apart from his parents. An acontextual explanation may conclude that this behavior, while typical of a two-year-old child, is abnormal for a five-year-old child. However, if this child is living in a violent neighborhood where physical danger is a daily reality, the behaviors that he is exhibiting are not "abnormal," nor are they unexpected. To be understood appropriately and completely, this child's behavior would need to be considered as a function of his context. Spencer (1990), states that "the integration or consideration of culture and context should afford a better match between policy, practice, and positive family and child outcomes" (p. 269). From a developmental contextualist perspective, clinicians will be most effective when they understand and approach problem behavior in terms of a lack of fit between child and context.

A developmental contextual approach also acknowledges that behaviors that are adaptive in one context may be maladaptive in another context. For instance, a child's neighborhood environment may make aggressive behavior important to the child's survival, but this behavior may be maladaptive in the school environment. If one intervenes to change the maladaptive behavior witnessed at school without understanding the adaptability of this same behavior in the neighborhood, one could be doing this child a disservice. The child no longer exhibits aggressive behavior at school, but now is disadvantaged on the streets. An accurate assessment of children's difficulties requires an understanding of children within their contexts and their "problematic" behaviors within these contexts. In our example, the behavior of the child fits his neighborhood context, but does not fit well with his school context. The goodness-of-fit principle explained above implies that an intervention should work to enhance the fit of child and context. The problem does not lie solely in the environment or solely within the child. Problems result when there is a lack of fit between the individual and the context. Using a developmental contextual perspective, a clinician will gather information on the child's context to better understand the factors influencing and reinforcing the problem behavior.

Not only do children exist within their immediate contexts, but they also exist within a specific historical moment. The concept of temporality suggests that at different historical moments the person–context relation could have different meanings, that is, what is considered "normal" changes over time. For instance, a teacher living in the southern United States who is teaching poor black children to read would be viewed quite differently at different points in history. One hundred years ago such a teacher may have been viewed as a sociopath, as someone who goes against social conventions and social norms. Today that same teacher may be viewed as a dedicated, caring individual. The person–context relation not only exists in ontogenetic time, but in historical time as well.

The concept of temporality also argues for examining "tried and true" interventions such as individual counseling. Humans are embedded in the historical time in which they live. Not only do individuals change over time, but changes also occur in the wider society. Systems of assessment and intervention must be reevaluated to ensure that earlier models, developed in earlier historical moments, remain relevant and effective in the current time period. For example, the earlier exclusive reliance on individual psychotherapy has begun to give way to increased emphasis on prevention measures and group interventions. New models of treatment have begun to emerge, for example, group and family counseling and psychoeducational intervention. There is an increased emphasis on providing services through community schools. As our culture and needs change, so do our modes of intervention. In a similar manner, the types of problems presented to clinicians today are often different from those one would have found in past decades. While syndromes similar to what is now called Attention Deficit Hyperactivity Disorder have been discussed since the 1860s, the symptoms that have defined the syndrome have changed. In the 1960s, the emphasis was on hyperactivity, whereas today hyperactivity is only one aspect of the disorder; one that is not even necessary for diagnosis (Barkley, 1990).

The implicit suggestion in many of the above examples is that a child's context is not unidimensional. Children exist within a family, a peer group, a school system, a community, a culture, and a specific period in history. Similarly, their problems are not unidimensional. Their development is affected by experiences at the inner biological, the individual/psychological, the social relational, the sociocultural, and the ecological levels of organization. Given these multiple levels, accurate assessment requires integrating information about the child from multiple perspectives. Achenbach (1990) suggests five axes that are relevant to the assessment of children. These comprise parent reports, teacher reports, cognitive assessment, physical assessment, and direct assessment of the child. The axes from which one gathers information will depend upon the relevance of each axis to the child's current situation.

A developmental contextual perspective requires not only multiple-level assessment but also multiple-level interventions. Within this framework, psychotherapeutic interventions are expanded to include multiple levels of organization (i.e., biological, sociohistorical, cultural, psychological, etc.). This results in a shift in the clinician's orientation from a typical unilevel approach to intervention strategies involving multiple levels of organization. A child's development does not occur at the same rate and in a parallel fashion at each of the levels of organization. Therefore, an unevenness in strengths and weaknesses may be observed. By focusing on weaknesses only, one falls into a deficit model that focuses only on children and their "abnormal" behaviors without considering contextual factors. A developmental-contextual approach considers children within their contexts and identifies strengths as well as deficits. This approach uses an optimization framework, focusing on and building on strengths to maximize a child's full development. The clinical intervention, then, will focus on the relationships among the levels of organization and therefore will be multimodal. While a child may manifest a weakness at one level, for example, the inner-biological level, he or she may be functioning well at another level, for example, the social relational level. Intervention can occur at any level and, given the interconnectedness of the levels, will help to promote growth in other levels.

The implications of the interconnectedness of the levels (i.e., plasticity) are evident in the research on resilience. The resilience literature clearly demonstrates that children have "self-righting tendencies that move children toward normal development under all but the most persistent adverse circumstances" (Werner, 1990, p. 112). Werner (1990) identifies three areas of protective factors that may contribute to a child's resilience. These include:

> dispositional attributes of the child that elicit predominantly positive responses from the environment, such as physical robustness and vigor, an easy temperament, and intelligence; affectional ties and socialization practices within the family that encourage trust, autonomy and initiative; and external support systems that reinforce competence and provide children with a positive set of values. (p. 111)

These protective factors are found at various levels of organization, for example, dispositional attributes are characteristics of the child, and are, therefore, found at the individual/psychological level, while affectional ties are located at the social relational level of organization. Behavioral change, then, is the result of the impact of both contextual and organismic characteristics on the individual's current functioning. However, as the literature on risk and prevention demonstrates, contextual and organismic factors do not interact in an additive or linear manner, but in a multiplicative way (Sameroff, 1996). It is therefore imperative that a clinician appreciate the complex and continually changing interplay between individuals and their contexts when engaging in clinical interventions.

In essence, plasticity implies that interventions are not fruitless endeavors and that poorly functioning individuals are not doomed to remain poorly functioning if an intervention can be developed that addresses a strength. While it is true that development may be constrained by past developments and current contextual conditions, plasticity implies that humans do have the potential for change across the life-span (Sroufe, 1990; Sroufe & Rutter, 1984). Clinicians can help to remove the constraining aspects of the child's context. To optimize the effects of an intervention, the clinician must address the contextual constraints at all levels of organization. For example, a homeless child who acts out in school requires a more complex diagnosis than that provided by the diagnostic label "conduct disorder," and a treatment plan more comprehensive than that offered by a unilevel intervention such as "insight-oriented therapy," or "pharmacological treatment" (Walsh, 1992; Walsh & Buckley, 1994). This broader view of assessment and intervention planning requires that a single practitioner, for example, a psychologist, take into account the various perspectives of the child's behavior offered by different professionals. Cowen (1984) argued compellingly for a more integrated approach to practice:

> Although most mental health professionals are knowledgeable about the nature and assessment of human adjustment, they are not Renaissance persons with exhaustive knowledge of complex feeder-domains such as education, sociology, environmental planning, economics, family relations, architecture, public health, epidemiology, and social ecology, which contain significant knowledge about factors that relate to human adjustment. The points to stress are that: (a) the potential generative base for primary prevention work is much broader than the mental health fields as classically designed, and (b) in the long run, accessing and harnessing important bodies of generative information calls both for new alliances between mental health professionals and members of the other disciplines and new types of training that cut across traditional disciplinary lines. (p. 487)

The concept of multiple levels of organization, then, implies not only multiple levels of assessment and intervention by clinicians, but also collaboration among professionals. For instance, in the example of the homeless child who acts out in school, one might have the school nurse assess their nutritional and physical needs, the teacher assess their educational needs, the school psychologist explore their emotional needs, and the parents report on their behavior outside of school and the social and economic stressors placed on them as a family. Such a collaboration would lead to a more holistic understanding of this boy's needs than would a simple diagnosis of "conduct disorder." Accordingly, the resulting intervention would also be holistic in nature, involving all of the collaborators.

In short, from a developmental contextual perspective, integrated multidisciplinary and developmental research devoted to the design of intervention programs for the person or group should be moved to the fore of scholarly concern. This integrative research and outreach must involve two other considerations: first, this research and outreach must involve collaborations among disciplines and between scholarly and community interests; and second, this research and outreach must be integrated with policies and programs.

5.01.4.2 The Need for Collaborative Research and Outreach

To be successful, these endeavors require more than collaboration across disciplines. In addition, two other types of collaboration are necessary. First, multiprofessional collaboration is essential. Colleagues in the research, policy, and intervention communities must plan and implement their activities in a synthesized manner in order to successfully develop and extend this vision. All components of this collaboration must be understood as equally valuable, indeed as equally essential. The collaborative activities of colleagues in university extension and outreach; in program design and delivery; in elementary, middle (or junior high), and high schools; in policy development and analysis; and in academic research are vital to the success of this new agenda for science and outreach for children, adolescents, parents, and their contexts, for example, their extended families, their schools, their workplaces, and their communities (Brabeck et al., 1997; Weissbourd, 1996).

Second, given the contextual embeddedness of these synthetic research and outreach activities, collaboration must occur with the people we are trying both to understand and to serve. Without incorporation of the perspective of the community into our work—without the community's sense of ownership, of value, and of meaning for these endeavors—research and outreach activities cannot be adequately integrated into the lives we are studying.

Thus, from a developmental contextual perspective, research that "parachutes" into the community from the heights of the academy (i.e., that is done in a community without collaboration with the members of the community) is flawed fatally in regard to its ability to understand the process of human development. This is the case because human development does not happen at the general level (Lerner, 1988, 1991); it does not occur in a manner necessarily generalizable across diverse people and contexts (Bibace & Walsh, 1979). Development happens in particular communities, and it involves the attempts of specific children and families to relate to the physical, personal, social, and institutional situations found in their communities. Without bringing the perspective of the community into the "plan" for research, then, scholars may fail to address the correct problems of human development, the ones involved in the actual lives of the people they are studying. And if the wrong problem is being addressed, any "answers" that are found are not likely to be relevant to the actual lives of people. Not surprisingly, these answers will be seen all too often (and quite appropriately) as irrelevant by the community.

In turn, however, if community members collaborate in the definition of the problems of development that they are confronting, and if they participate in the research process, then the obtained answers are more likely to be considered significant by the community members. This will increase the likelihood of these answers being used to build community-specific policies and programs. Moreover, community empowerment and capacity building occur by engaging in a collaborative process wherein the community places value and meaning on, and participates in, the research and outreach being conducted within its boundaries (Dryfoos, 1990, 1994; Schorr, 1988).

5.01.4.3 Implications for Policies and Programs

In order to be complete, the integrative research promoted by a developmental contextual view of human development must be synthesized with two other focuses. Research in human development that is concerned with one, or even a few, instances of individual and contextual diversity cannot be assumed to be useful for understanding the life course of all people. Similarly, policies and programs derived from such research, or associated with it in the context of a researcher's tests of ideas pertinent to human plasticity, cannot be assumed to be applicable, or equally appropriate and useful, in all contexts or for all individuals. Accordingly, developmental and individual differences-oriented policy development and program (intervention) design and delivery would need to be integrated fully with the new research base for which we are calling (Lerner & Miller, 1993; Lerner et al., 1994).

As emphasized in developmental contextualism, the variation in settings within which people live means that studying development in a standard (for example, a "controlled") environment does not provide information pertinent to the actual (ecologically valid), developing relations between individually distinct people and their specific contexts (for example, their particular families, schools, or communities). This point underscores the need to conduct research in real-world settings, and highlights the ideas that: policies and programs constitute natural experiments, that is, planned interventions for people and institutions; and the evaluation of such activities becomes a central focus in the developmental contextual research agenda we have described.

In this view, then, policy and program endeavors do not constitute secondary work, or derivative applications, conducted after research evidence has been complied. Quite to the contrary, policy development and implementation, and program design and delivery, become integral components of this vision for research; the evaluation component of such policy and intervention work provides critical feedback about the adequacy of the conceptual frame from which this research agenda should derive (cf. Lanier, 1990). This conception of the integration of multidisciplinary research endeavors centrally aimed at diversity and context, with policies, programs, and evaluations is illustrated in Figure 2.

A vision of the integration between developmental research and policies and programs was articulated in the 1970s by Bronfenbrenner (1974), and more recently by Zigler (in press). Bronfenbrenner argued that engagement with social policy not only enhances developmental research but, consistent with the developmental contextual perspective, it also augments understanding of key theoretical issues pertinent to the nature of person–context relations. Bronfenbrenner (1974) noted that:

> In discussions of the relation between science and social policy, the first axiom, at least among social scientists, is that social policy should be based on science. The proposition not only has logic on its side, but what is more important, it recognizes our proper and primary importance in the scheme of things. The policymakers should look to us, not only for truth, but for wisdom as well. In short, social policy needs science.
>
> My thesis in this paper is the converse proposition, that, particularly in our field, science needs social policy—needs it not to guide our organizational activities, but to provide us with two elements essential for any scientific endeavor—vitality and validity ... I contend that the pursuit of [social policy] questions is essential for the further development of knowledge and theory on the process of human development. Why essential? ... [Because] issues of social policy [serve] as points of departure for the identification of significant theoretical and scientific questions concerning the development of the human organism as a function of interaction with its enduring environment—both actual and potential. (pp. 1–2,4)

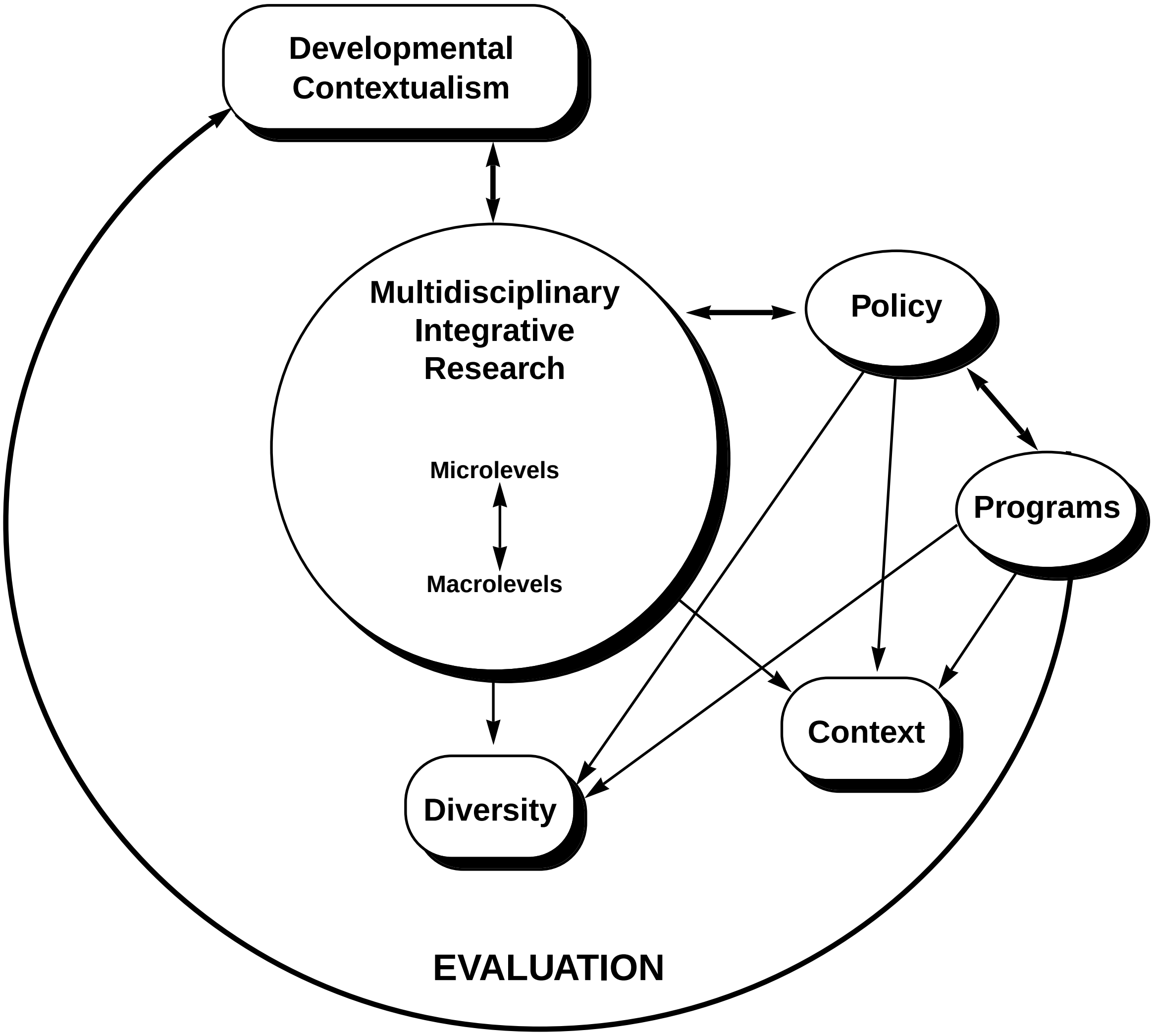
Developmental Contextualism
Multidisciplinary Integrative Research
Microlevels
Macrolevels
Policy
Programs
Context
Diversity
EVALUATION

Figure 2 A developmental contextual model of the integration of multilevel, multidisciplinary research, aimed at diversity and context, with policies, programs, and evaluations.

In short, developmental research informs and is informed by social policy. It is our contention that policy development, program design and delivery, and developmental research need to be integrated fully to optimally benefit the recipients of social programs.

5.01.5 CONCLUSIONS AND FUTURE DIRECTIONS

Developmental contextualism and the research and intervention associated with it take an integrative approach to the multiple levels of organization presumed to comprise the nature of human life; that is, "fused" (Tobach & Greenberg, 1984) relations among biological, psychological, and social and physical contextual levels comprise the process of developmental change in human life. Rather than approach variables from these levels of analysis in either a reductionistic or in a parallel-processing approach, developmental contextualism rests on the idea that variables from these levels of analysis are dynamically interactive: they are reciprocally influential over the course of human ontogeny (Lerner, 1986, 1991, 1995).

In turn, developmental contextualism provides an agenda not only for a developmental, dynamic, and systems approach to research about child and adolescent development. As well, such theory allows us to envision the possibility of promoting positive developmental trajectories in youth (Lerner, 1995). We, as clinicians, researchers, and policy makers, may actualize this vision if we remain assiduously committed to a developmental systems orientation; if we recognize the "double-edged sword" nature of plasticity that derives from the functioning of this system; and if we therefore create, through policies and programs, a "convoy of social support" (Kahn & Antonucci, 1980) across the life course of children and adolescents. Such a convoy would be a network encompassing the familial, community, institutional, and cultural components of the ecology that impacts people's behavior and development across their life.

The concepts of person, of context, and of the relations between the two found in a developmental contextual perspective are, as a set, quite complex, and impose formidable challenges on those who seek to understand the developmental bases of psychopathology and of nonpsychopathologic behavior, to derive feasible research from this perspective, and to use this conception and its associated research for the design and implementation of interventions. As we have argued, this developmental contextual perspective leads to an integrated, multilevel concept of development, one in which the focus of inquiry is the organism–environment dynamic interaction. Furthermore, such an orientation places an emphasis on the potential for intra-individual change in structure and function—for plasticity—across the life-span. Moreover, this perspective requires a temporal (historical)/longitudinal and a relational/comparative orientation to research and to intervention.

A major challenge of this perspective is the derivation and empirical testing of models reflecting the nature of dynamic, interlevel interactions across time. As we have indicated, one reasonably successful path we have taken involves the testing of the goodness-of-fit model of person–context relations. Of course, the goodness-of-fit model is not the only conception of person–context relations that may be derived from a developmental contextual orientation. There are perhaps an infinity of possible interlevel relations that may occur and a potentially similarly large array of ways to model them. Indeed, tests of other models derived from or consistent with a developmental contextual or life-span perspective have also found considerable empirical support (e.g., Baltes, 1987; Caspi & Elder, 1988; Featherman, 1985; Perlmutter, 1988).

The tests of such models will profit by triangulation of the constructs within each of the levels of analysis thought to interact dynamically within a given model. For instance, in regard to the biological/organismic-psychosocial interactions assessed within the goodness-of-fit model, biological changes could be simultaneously indexed both by more molecular (e.g., hormonal) changes, and by more molar, bodily measures. Similarly, contextual demands could be simultaneously appraised by assessing both attitudes/expectations about behavior held by a person's significant other and actual behavioral exchanges between a person and these others. Such triangulation would not only provide convergent and discriminant validation information but also better insight would be gained about whether all modalities of functioning within a level of analysis are of similar import for the development of psychopathological or nonpsychopathological functioning in particular person–context interactions.

In sum, the relative plasticity of human development across the life-span—a plasticity deriving from the dynamic interactions between organism and context which characterize human functioning—is already well documented (Baltes, 1987; Brim & Kagan, 1980; Featherman, 1983; Hetherington, Lerner, & Perlmutter, 1988; Lerner, 1984; Sorensen, Weinert, & Sherrod, 1986). Thus, a future including the sorts of directions we suggest should enrich greatly our understanding of the precise conditions promoting and constraining the development of psychopathology. Given, then, the present literature, and the promise we see for tomorrow, we believe there is reason for great optimism about the future scientific and clinical use of the developmental contextual view of the biological and social contextual bases of, on the one hand, functional, adaptive behaviors and, on the other, malfunctional, psychopathological ones.

5.01.6 REFERENCES

Achenbach, T. M. (1990). Conceptualization of developmental psychopathology. In M. Lewis & S. M. Miller (Eds.), *Handbook of developmental psychopathology* (pp. 3–14). New York: Plenum.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anastasi, A. (1958). Heredity, environment, and the question, "how?" *Psychological Review*, *65*, 197–208.

Baca Zinn, M., & Eitzen D. S. (1993). *Diversity in families* (3rd ed.). New York: Harper Collins College.

Baer, D. M. (1970). An age-irrelevant concept of development. *Merrill Palmer Quarterly of Behavior and Development*, *16*, 238–45.

Baer, D. M. (1976). The organism as host. *Human Development*, *19*, 87–98.

Baltes, P. B. (1968). Longitudinal and cross-sectional sequences in the study of age and generation effects. *Human Development*, *11*, 145–171.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology*, *23*, 611–626.

Baltes, P. B., Dittmann-Kohli, F., & Dixon, R. A. (1984). New perspectives on the development of intelligence in adulthood: Toward a dual-process conception and model of selective optimization with compensation. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (Vol. 6, pp. 33–76). New York: Academic Press.

Baltes, P. B., & Nesselroade, J. R. (1973). The developmental analysis of individual differences on multiple measures. In J. R. Nesselroade & H. W. Reese (Eds.), *Life-span developmental psychology: Introduction to research methodological issues* (pp. 219–251). New York: Academic Press.

Baltes, P. B., Reese, H. W., & Lipsitt, L. P. (1980). Life-span developmental psychology. *Annual Review of Psychology*, *31*, 65–110.

Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1977). *Life-span developmental psychology: Introduction to research methods.* Monterey, CA: Brooks/Cole.

Baltes, P. B., Smith, J., & Staudinger, U. M. (1992). Wisdom and successful aging. In T. B. Sonderegger (Ed.), *Nebraska Symposium on Motivation* (Vol. 39, 123–167). Lincoln, NE: University of Nebraska Press.

Barkley, R. A. (1990). Attention deficit disorders: History, definition, and diagnosis. In M. Lewis & S. M. Miller (Eds.), *Handbook of developmental psychopathology* (pp. 3–14). New York: Plenum.

Baumrind, D. (1995). *Child maltreatment and optimal caregiving in social contexts* (MSU Series on Children, Youth, and Families). New York: Garland.

Bell, R. Q. (1968). A reinterpretation of the direction of effects in studies of socialization. *Psychological Review*, *115*, 81–95.

Belsky, J., Lerner, R. M., & Spanier, G. B. (1984). *The child in the family*. Reading, MA: Addison-Wesley.

Bibace, R., & Walsh, M. (1979). Clinical developmental psychologists in family practice settings. *Professional Psychology*, *10*, 441–450.

Bijou, S. W. (1976). *Child development: The basic stage of early childhood.* Englewood Cliffs, NJ: Prentice-Hall.

Bijou, S. W., & Baer, D. M. (Ed.) (1961). *Child development: A systematic and empirical theory*. New York: Appleton-Century-Crofts.

Birkel, R., Lerner, R. M., & Smyer, M. A. (1989). Applied developmental psychology as an implementation of a life-span view of human development. *Journal of Applied Developmental Psychology*, *10*, 425–445.

Block, J. (1971). *Lives through time*. Berkeley, CA: Bancroft.

Bornstein, M. H., & Tamis-LeMonda, C. S. (1990). Activities and interactions of mothers and their firstborn infants in the first six months of life: Covariation, stability, continuity, correspondence, and prediction. *Child Development*, *61*, 1206–1217.

Bornstein, M. H., Tamis-LeMonda, C. S., Tal, J., Ludemann, P., Toda, S., Rahn, C., Pecheux, M., Azuma, H., & Vardi, D. (1992). Maternal responsiveness to infants in three societies: The United States, France, and Japan. *Child Development*, *63*, 808–821.

Brabeck, M., Cawthorne, J., Cochran-Smith, M., Gaspard, N., Green, C. H., Kenny, M., Krawczy, K. R., Lowery, C., Lykes, M. B., Minuskin, A. D., Mooney, J., Ross, C. J., Savage, J., Soifer, A., Smyer, M., Sparks, E., Tourse, R., Turillo, R. M., Waddock, S., Walsh, M., & Zollers, N. (1997). Changing the culture of the university to engage in outreach scholarship. In R. M. Lerner & L. A. K. Simon (Eds.), *Creating the new outreach university for America's youth and families: Building university–community collaborations for the twenty-first century (pp. 335–366)*. New York: Garland Press.

Brazelton, T. B., Koslowski, B., & Main, M. (1974). The origins of reciprocity: The early mother–infant interaction. In M. Lewis & L. A. Rosenblum (Eds.), *The effect of the infant on its caregivers* (pp. 49–76). New York: Wiley.

Brim, O. G., Jr., & Kagan, J. (Eds.) (1980). *Constancy and change in human development*. Cambridge, MA: Harvard University Press.

Bronfenbrenner, U. (1974). Developmental research, public policy, and the ecology of childhood. *Child Development*, *45*, 1–5.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, *32*, 513–531.

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.

Bronfenbrenner, U. (1983). The context of development and the development of context. In R. M. Lerner (Ed.), *Developmental psychology: Historical and philosophical perspectives* (pp. 39–83). Hillsdale, NJ: Erlbaum.

Bronfenbrenner, U. & Morris, P. A. (1998). The ecology of developmental processes. In R. M. Lerner (Ed.), *Theoretical models of human development*. Vol. 1. of the *Handbook of child psychology* (5th ed., pp. 993–1028). New York: Wiley.

Bronfenbrenner, U., & Crouter, A. C. (1983). The evolution of environmental models in developmental research. In W. Kersen (Ed.), *Handbook of child psychology, Vol 1: History, theories, and methods* (pp. 39–83). New York: Wiley.

Cairns, R. B., & Hood, K. E. (1983). Continuity in social development: A comparative perspective on individual difference prediction. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (Vol. 5, pp. 302–358). New York: Academic Press.

Campbell, S. B. (1989). Developmental perspectives. In T. H. Ollendick & M. Herson (Eds.), *Handbook of child psychopathology* (pp. 5–28). New York: Plenum.

Caspi, A., & Elder, G. H., Jr. (1988). Childhood precursors of the life course: Early personality and life disorganization. In E. M. Hetherington, R. M. Lerner, & M. Perlmutter (Eds.), *Child development in life-span perspective* (pp. 115–142). Hillsdale, NJ: Erlbaum.

Chess, S., & Thomas, A. (1984). *The origins and evolution of behavior disorders: Infancy to early adult life*. New York: Brunner/Mazel.

Cicchetti, D. (1984). The emergence of developmental psychopathology. *Child Development*, *55*, 1–7.

Cicchetti, D. (1987). Developmental psychopathology in infancy: Illustration from the study of maltreated youngsters. *Journal of Consulting and Clinical Psychology*, *55*, 837–845.

Cicchetti, D., & Garmezy, N. (1993). Prospects and promises in the study of resilience. *Development and Psychopathology*, *5*, 497–502.

Cowan, P. A. (1988). Developmental psychopathology: A nine-cell map of the territory. In E. D. Nannis & P. A. Cowan (Eds.), *New Directions for Child Development, No. 39*: *Developmental psychopathology and its treatment* (pp. 5–29). San Francisco, CA: Jossey-Bass.

Cowen, E. L. (1984, April). A general structural model for primary prevention program development in mental health. *Personnel and Guidance Journal*, 485–490.

Dixon, R. A., & Lerner, R. M. (1992). A history of systems in developmental psychology. In M. H. Bornstein & M. E. Lamb (Eds.), *Developmental psychology: An advanced textbook* (3rd ed., pp. 3–58). Hillsdale, NJ: Erlbaum.

Dixon, R. A., Lerner, R. M., & Hultsch, D. F. (1991). The concept of development in individual and social change. In P. Van Geert & L. P. Mos (Eds.), *Annals of theoretical psychology*, (Vol. 7, pp. 279–323). New York: Plenum.

Dryfoos, J. G. (1990). *Adolescents at risk: Prevalence and prevention*. New York: Oxford University Press.

Dryfoos, J. G. (1994). *Full service schools: A revolution in health and social services of children, youth and families.* San Francisco: Jossey-Bass.

Elder, G. H., Jr. (1974). *Children of the Great Depression: Social change in life experiences*. Chicago: University of Chicago Press.

Elder, G. H., Jr. (1980). Adolescence in historical perspective. In J. Adelson (Ed.), *Handbooks of adolescent psychology* (pp. 3–46). New York: Wiley.

Elder, G. H., Jr., Modell, J., & Parke, R. D. (Eds.) (1993). *Children in time and place: Developmental and historical insights*. New York: Cambridge University Press.

Featherman, D. L. (1983). Life-span perspectives in social science research. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (Vol. 5, pp. 1–57). New York: Academic Press.

Featherman, D. L. (1985). Individual development and aging as a population process. In J. R. Nesselroade & A. von Eye (Eds.), *Individual development and social change: Explanatory analyses* (pp. 213–241). New York: Academic Press.

Featherman, D. L., & Lerner, R. M. (1985). Ontogenesis and sociogenesis: Problematics for theory about development across the lifespan. *American Sociological Review*, *50*, 659–676.

Feldman, M. W., & Lewontin, R. C. (1975). The heritability hang-up. *Science*, *190*, 1163–1168.

Finkelstein, J. W. (1993). Familial influences on adolescent health. In R. M. Lerner (Ed.), *Early adolescence: Perspectives on research, policy and intervention* (pp. 111–126). Hillsdale, NJ: Erlbaum.

Fisher, C. B., & Lerner, R. M. (Eds.). (1994). *Applied developmental psychology*. New York: McGraw-Hill.

Ford, D. L., & Lerner, R. M. (1992). *Developmental systems theory: An integrative approach*. Newbury Park, CA: Sage.

Freedman, D. G. (1979). *Human sociobiology: A holistic approach*. New York: Free Press.

Freud, S. (1949). *Outline of psychoanalysis*. New York: Norton.

Garbarino, J. (1992). *Children and families in the social environment* (2nd ed.). New York: Aldine de Gruyter.

Gottlieb, G. (1991). The experiential canalization of behavioral development: Theory. *Developmental Psychology*, *27*, 4–13.

Gottlieb, G. (1992). *Individual development and evolution: The genesis of novel behavior*. New York: Oxford University Press.

Hartup, W. W. (1978). Perspectives on child and family interaction: Past, present, and future. In R. M. Lerner & G. B. Spanier (Eds.), *Child influences on marital and family interaction: A life-span perspective* (pp. 23–45). New York: Academic Press.

Hetherington, E. M., & Baltes, P. B. (1988). Child psychology and life-span development. In E. M. Hetherington, R. M. Lerner, & M. Perlmutter (Eds.), *Child development in life-span perspective* (pp. 1–19). Hillsdale, NJ: Erlbaum.

Hetherington, E. M., Lerner, R. M., & Perlmutter, M. (Ed.). (1988). *Child development in life-span perspective*. Hillsdale, NJ: Erlbaum.

Hirsch, J. (1970). Behavior-genetic analysis and its biosocial consequences. *Seminars in Psychiatry*, *2*, 89–105.

Homans, G. C. (1961). *Social behavior: Its elementary forms*. New York: Harcourt, Brace, & World.

Kahn, R. L., & Antonucci, T. C. (1980). Convoys over the life course: Attachment, roles, and social support. In P. B. Baltes & O. G. Brim (Eds.), *Life-span development and behavior* (Vol. 3, pp. 253–286). Hillsdale, NJ: Erlbaum.

Lanier, J. E. (1990). *Report to focus group colleagues on "teaching," National Academy of Education study on the future of educational research*. East Lansing, MI: College of Education, Michigan State University.

Lerner, J. V. (1994). *Working women and their families*. Thousand Oaks, CA: Sage.

Lerner, R. M. (1976). *Concepts and theories of human development*. Reading, MA: Addison-Wesley.

Lerner, R. M. (1978). Nature, nurture, and dynamic interactionism. *Human Development*, *21*, 1–20.

Lerner, R. M. (1979). A dynamic interactional concept of individual and social relationship development. In R. L. Burgess & T. L. Huston (Eds.), *Social exchange in developing relationships* (pp. 271–305). New York: Academic Press.

Lerner, R. M. (1982). Children and adolescents as producers of their own development. *Developmental Review*, *2*, 342–370.

Lerner, R. M. (1984). *On the nature of human plasticity*. New York: Cambridge University Press.

Lerner, R. M. (1986). *Concepts and theories of human development* (2nd ed.). New York: Random House.

Lerner, R. M. (1987). A life-span perspective for early adolescence. In R. M. Lerner & T. T. Foch (Eds.), *Biological-psychosocial interactions in early adolescence: A life-span perspective* (pp. 9–34). Hillsdale, NJ: Erlbaum.

Lerner, R. M. (1988). Personality development: A life-span perspective. In E. M. Hetherington, R. M. Lerner, & M. Perlmutter (Eds.), *Child development in life-span perspective* (pp. 21–46). Hillsdale, NJ: Erlbaum.

Lerner, R. M. (1991). Changing organism–context relations as the basic process of development: A developmental-contextual perspective. *Developmental Psychology*, *27*, 27–32.

Lerner, R. M. (1995). *America's youth in crisis: Challenges and options for programs and policies*. Thousand Oaks, CA: Sage.

Lerner, R. M. (1997). Relative plasticity, integration, temporality, and diversity in human development: A developmental contextual perspective about theory, process, and method. *Developmental Psychology, 32*(4), 781–786.

Lerner, R. M. (Ed.) (1998). *Theoretical models of human development*. Volume 1 of *The Handbook of Child Psychology* (5th ed., pp. 1–24). New York: Wiley.

Lerner, R. M., & Busch-Rossnagel, N. A. (Eds.) (1981). *Individuals as producers of their development: A life-span perspective*. New York: Academic Press.

Lerner, R. M., Castellino, D. R., Terry, P. A., Villarruel, F. A., & McKinney, M. H. (1995). A developmental contextual perspective on parenting. In M. H. Bornstein (Ed.), *Handbook of parenting: Vol. 2. Biology and ecology of parenting* (pp. 285–309). Hillsdale, NJ: Erlbaum.

Lerner, R. M., Hess, L. E., & Nitz, K. (1990). A developmental perspective on psychopathology. In M. Hersen & C. G. Last (Eds.), *Handbook of child and adult psychopathology: A longitudinal perspective* (pp. 9–32). New York: Pergamon.

Lerner, R. M., Hess, L. E., & Nitz, K. (1991). Towards the integration of human developmental and therapeutic change: A developmental contextual perspective. In P. R. Martin (Ed.), *Handbook of behavior therapy and psychological science: An integrative approach* (pp. 13–34). New York: Pergamon.

Lerner, R. M., & Hood, K. E. (1986). Plasticity in development: Concepts and issues for intervention. *Journal of Applied Developmental Psychology*, *7*, 139–152.

Lerner, R. M., Hultsch, D. F., & Dixon, R. A. (1983). Contextualism and the character of developmental psychology in the 1970s. *Annals of the New York Academy of Sciences*, *412*, 101–128.

Lerner, R. M., & Kauffman, M. B. (1985). The concept of development in contextualism. *Developmental Review*, *5*, 309–333.

Lerner, R. M., & Lerner, J. V. (1983). Temperament–intelligence reciprocities in early childhood: A contextual model. In M. Lewis (Ed.), *Origins of intelligence: Infancy and early childhood* (pp. 399–421). New York: Plenum.

Lerner, R. M., & Lerner, J. V. (1987). Children in their contexts: A goodness of fit model. In J. B. Lancaster, J. Altmann, A. S. Rossi, & L. R. Sherrod (Eds.), *Parenting across the lifespan: Biosocial dimensions* (pp. 377–404). Chicago: Aldine.

Lerner, R. M., & Lerner, J. V. (1989). Organismic and social contextual bases of development: The sample case of early adolescence. In W. Damon (Ed.), *Child development today and tomorrow* (pp. 69–85). San Francisco: Jossey-Bass.

Lerner, R. M., & Miller, J. R. (1993). Integrating human development research and intervention for America's children: The Michigan State University model. *Journal of Applied Developmental Psychology*, *14*, 347–364.

Lerner, R. M., Miller, J. R., Knott, J.H., Corey, K. E., Bynum, T. S., Hoopfer, L. C., McKinney, M. H., Abrams, L. A., Hula, R. C., & Terry, P. A. (1994). Integrating scholarship and outreach in human development research, policy, and service: A developmental contextual perspective. In D. L. Featherman, R. M. Lerner, & M. Perlmutter (Eds.), *Life-span development and behavior*, (Vol. 12, pp. 249–273). Hillsdale, NJ: Erlbaum.

Lerner, R. M., & Ryff, C. D. (1978). Implementation of the life-span view of human development: The sample case of attachment. In P. B. Baltes (Ed.), *Life-span development and behavior* (Vol. 1, pp. 1–44). New York: Academic Press.

Lerner, R. M., & Spanier, G. B. (Eds.). (1978). *Child influences on marital and family interaction: A life-span perspective*. New York: Academic Press.

Lerner, R. M., & Tubman, J. (1989). Conceptual issues in studying continuity and discontinuity in personality development across life. *Journal of Personality*, *57*, 343–373.

Lewis, M. (1997). *Altering fate*. New York: Guilford Press.

Lewis, M., & Feiring, C. (1978). A child's social world. In R.M. Lerner & G.B. Spanier (Eds.), *Child influences on marital and family interaction* (pp. 47–66). New York: Academic Press.

Lewis, M., & Rosenblum, L. A. (Ed.). (1974). *The effect of the infant on its caregivers*. New York: Wiley.

Magnusson, D. (1988). *Individual development from an interactional perspective: A longitudinal study*. Hillsdale, NJ: Erlbaum.

Magnusson, D., & Allen, V. L. (Ed.). (1983). *Human development: An interactional perspective*. New York: Academic Press

Magnusson, D., & Öhman, A. (Eds.). (1987). *Psychopathology: An interactional perspective*. Orlando, FL: Academic Press.

Masten, A. S. (1989). Resilience in development: Implications of the study of successful adaptation for developmental psychopathology. In D. Cicchetti (Ed.), *The emergence of a discipline: Vol. 1. Rochester symposium on developmental psychopathology* (pp. 261–294). Hillsdale, NJ: Erlbaum.

Novikoff, A. B. (1945). The concept of integrative levels of biology. *Science*, *62*, 209–215.

Öhman, A., & Magnusson, D. (1987). An interactional paradigm for research on psychopathology. In D. Magnusson & A. Öhman (Eds.), *Psychopathology: An interactional perspective* (pp. 3–19). Orlando, FL: Academic Press.

Overton, W. F., & Reese, H. W. (1973). Models of development: Methodological implications. In J. R. Nesselroade & H. W. Reese (Eds.), *Life-span developmental psychology: Methodological issues* (pp. 65–86). New York: Academic Press.

Overton, W. F., & Reese, H. W. (1981). Conceptual prerequisites for an understanding of stability–change and continuity–discontinuity. *International Journal of Behavioral Development*, *4*, 99–123.

Pepper, S.C. (1942). *World hypotheses*. Berkeley, CA: University of California.

Perlmutter, M. (1988). Cognitive development in life-span perspective: From description of differences to explanation of changes. In E. M. Hetherington, R. M. Lerner, & M. Perlmutter (Eds.), *Child development in life-span perspective* (pp. 191–214). Hillsdale, NJ: Erlbaum.

Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology* (Vol. 1, pp. 703–732). New York: Wiley.

Plomin, R. (1986). *Development, genetics, and psychology*. Hillsdale, NJ: Erlbaum.

Reese, H. W., & Overton, W. F. (1970). Models of development and theories of development. In L. R. Goulet & P. B. Baltes (Eds.), *Life-span developmental psychology: Research and theory* (pp. 115–145). New York: Academic Press.

Riegel, K. F. (1975). Toward a dialectical theory of development. *Human Development*, *18*, 50–64.

Riegel, K. F. (1976a). The dialectics of human development. *American Psychologist*, *31*, 689–700.

Riegel, K. F. (1976b). From traits and equilibrium toward developmental dialectics. In W. J. Arnold & J. K. Cole (Eds.), *Nebraska symposium on motivation* (pp. 348–408). Lincoln, NE: University of Nebraska Press.

Riley, M. W. (Ed.) (1979). *Aging from birth to death*. Washington, DC: American Association for the Advancement of Science.

Rowe, D. C. (1994). *The limits of family influence: Genes, experience, and behavior*. New York: Guilford Press.

Sameroff, A. (1975). Transactional models in early social relations. *Human Development*, *18*, 65–79.

Sameroff, A. (1983). Developmental systems: Contexts and evolution. In W. Kessen (Ed.), *Handbook of child psychology* (Vol. 1, pp. 237–294). New York: Wiley.

Sameroff, A. (1989). Models of developmental regulation: The environtype. In D. Cicchetti (Ed.), *The emergence of a discipline* (Vol. 1, pp. 41–68). Hillsdale, NJ: Erlbaum.

Sameroff, A. (1996, Fall). Democratic and republican models of development: Paradigms and perspectives. *Developmental Psychology Newsletter*, 1–9.

Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin*, *64*, 92–107.

Schneirla, T. C. (1957). The concept of development in comparative psychology. In D. B. Harris (Ed.), *The concept of development* (pp. 78–108). Minneapolis, MN: University of Minnesota Press.

Schorr, L. B. (1988). *Within our reach: Breaking the cycle of disadvantage*. New York: Doubleday.

Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton.

Smith, L. B., & Thelen, E. (Eds.). (1993). *A dynamic systems approach to development: Applications*. Cambridge, MA: MIT Press.

Sorensen, B., Weinert, E., & Sherrod, L. R. (Ed.). (1986). *Human development and the life course: Multidisciplinary perspectives*. Hillsdale, NJ: Erlbaum.

Spencer, M. B. (1990). Development of minority children: An introduction. *Child Development*, *61*, 267–269.

Sroufe, L. A. (1990). Considering normal and abnormal together: The essence of developmental psychopathology. *Development and Psychopathology*, *2*, 335–347.

Sroufe, L. A., & Rutter, M. (1984). The domain of developmental psychopathology. *Child Development*, *55*, 17–29.

Super, C. M. (1987). The role of culture in developmental disorder. In C. M. Super (Ed.), *The role of culture in developmental disorder* (pp. 2–7). San Diego, CA: Academic Press.

Super, C. M., & Harkness, S. (1982). The infant's niche in rural Kenya and metropolitan America. In L. L. Adler (Ed.), *Cross-cultural research at issue* (pp. 47–56). New York: Academic Press.

Super, C. M., & Harkness, S. (1986). The developmental niche: A conceptualization at the interface of child and culture. *International Journal of Behavioral Development*, *9*, 1–25.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Brunner/Mazel.

Thomas, A., & Chess, S. (1980). *The dynamics of psychological development*. New York: Brunner/Mazel.

Thomas, A., & Chess, S. (1981). The role of temperament in the contributions of individuals to their development. In R. M. Lerner & N. A. Busch-Rossnagel (Eds.), *Individuals as producers of their own development: A life-span perspective*. New York: Academic Press.

Thomas, A., Chess, S., Birch, H. G., Hertzig, M. E., & Korn, S. (1963). *Behavioral individuality in early childhood* (pp. 231–255). New York: New York University Press.

Tobach, E. (1981). Evolutionary aspects of the activity of the organism and its development. In R. M. Lerner & N. A. Busch-Rossnagel (Eds.), *Individuals as producers of their development: A life-span perspective* (pp. 37–68). New York: Academic Press.

Tobach, E., & Greenberg, G. (1984). The significance of T. C. Schneirla's contribution to the concept of levels of integration. In G. Greenberg & E. Tobach (Eds.), *Behavioral evolution and integrative levels* (pp. 1–7). Hillsdale, NJ: Erlbaum.

Tobach, E., & Schneirla, T. C. (1968). The biopsychology of social behavior of animals. In R. E. Cooke & S. Levin (Eds.), *Biologic basis of pediatric practice* (pp. 68–82). New York: McGraw-Hill.

Tubman, J. G., & Lerner, R. M. (1994). Affective experiences of parents and their children from adolescence to young adulthood: Stability of affective experiences. *Journal of Adolescence*, *17*, 81–98.

Valsiner, J. (1988). The development of the concept of development: Historical and epistemological perspectives. In W. Damon (Ed.), *Handbook of child psychology* (5th ed., Vol.1, pp. 189–232). New York: Wiley.

Vondracek, F. W., Lerner, R. M., & Schulenberg, J. E. (1986). *Career development: A life-span developmental approach*. Hillsdale, NJ: Erlbaum.

Walsh, M. E. (1992). *"Moving to nowhere": Children's stories of homelessness*. Westport, CT: Auburn House.

Walsh, M. E., & Buckley, M. A. (1994). Children's experiences of homelessness: Implications for school counselors. *Elementary School Guidance and Counseling*, *29*(1), 4–15.

Walsten, D. (1990). Insensitivity of the analysis of variance to heredity–environment interaction. *Behavioral and Brain Sciences*, *13*, 109–120.

Wapner, S., & Demick, J. (1988, October 25–27). *Some relations between developmental and environmental psychology: An organismic-developmental systems perspective*. Paper presented at the "Visions of Development, the Environment, and Aesthetics: The Legacy of Joachim Wohlwill" conference, The Pennsylvania State University, University Park.

Weissbourd, R. (1996). *The vulnerable child: What really hurts America's children and what we can do about it*. Reading, MA: Addison-Wesley.

Werner, E. E. (1990). Protective factors and individual resilience. In S. Meisels & J. Shonkoff (Eds.), *Handbook of early childhood intervention* (pp. 97–116). New York: Cambridge University Press.

Werner, E. E. (1995). Resilience in development. *Current Directions in Psychological Science*, *4*(3), 81–85.

Werner, H. (1957). The concept of development from a comparative and organismic point of view. In D. B. Harris (Ed.), *The concept of development* (pp. 125–148). Minneapolis, MN: University of Minnesota Press.

Wilson, E. O. (1975). *Sociobiology: The new synthesis*. Cambridge, MA: Harvard University Press.

Zigler, E. E. (in press). Child development and social policy. *Child Development*.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.02
Clinical Formulation

MARTIN HERBERT
University of Exeter, UK

5.02.1 INTRODUCTION 26
5.02.2 THE CLINICAL FORMULATION: THEORETICAL ISSUES 26
  5.02.2.1 Clinical Psychology as Applied Science 26
5.02.3 THE ASSESSMENT PROCESS 28
  5.02.3.1 The "What?" Question 28
  5.02.3.2 The "Which?" Question 28
  5.02.3.3 The "Why?" Question (Etiology) 28
  5.02.3.4 The "How?" Question 31
5.02.4 EXPLANATORY MODELS 31
  5.02.4.1 The Medical Model—Diagnostic Formulations 32
  5.02.4.2 The Statistical Model—Multivariate Analyses 34
  5.02.4.3 The Psychoanalytic Model—Psychodynamic Formulations 34
  5.02.4.4 The Personality Typology Model—Trait Formulations 35
  5.02.4.5 The Systemic Model—Family Therapy Approaches 36
    5.02.4.5.1 Behavioral family therapy 38
  5.02.4.6 The Behavioral Model—Functional Analysis 39
    5.02.4.6.1 Behavioral formulations 40
    5.02.4.6.2 Distal consequences (Figure 4) 40
    5.02.4.6.3 The formal assessment 42
  5.02.4.7 Developmental-behavioral Formulations 42
    5.02.4.7.1 A developmental knowledge base 43
  5.02.4.8 Social-learning Model—Cognitive-behavioral Approach 44
  5.02.4.9 Integrative Models 44
5.02.5 THE CLINICAL FORMULATION—PRACTICE ISSUES 45
5.02.6 CONDUCTING A FORMULATION 45
  5.02.6.1 The Intake Interviews—Preparation 45
  5.02.6.2 Engaging the Clients 45
  5.02.6.3 The First Contact 46
  5.02.6.4 The "What?" Question 46
  5.02.6.5 Unpacking the Problem 48
  5.02.6.6 Communication 48
  5.02.6.7 The "Which?" Question 48
5.02.7 THE FORMULATION 48
  5.02.7.1 The "Why?" Question 49
  5.02.7.2 Explaining to Parents 50
5.02.8 THE "HOW?" QUESTION—IMPLEMENTING AN INTERVENTION 51
  5.02.8.1 The Collaborative Approach 51
  5.02.8.2 Rigorous Evaluation 51
5.02.9 SUMMARY 52
5.02.10 REFERENCES 52

5.02.1 INTRODUCTION

The process of piecing together a clinical formulation lies at the very center of the practitioner's craft, bridging, as it does, the issues of assessment, etiology, and intervention which inform clinical work. The skills involved in specifying the nature of the presenting problem, formulating explanatory hypotheses, determining and evaluating a test of the formulation, and translating it into a treatment plan have their roots in psychological science and art.

The assertion that there is a place for art (imagination and lateral thinking) as well as science (empirical research) in the many complex activities that contribute to the final clinical formulation, might well be regarded as contentious. Indeed those who embrace a scientist-practitioner approach to clinical psychology practice may well see it as erroneous—a matter returned to later in this chapter. Generally speaking, the formulation is a means to an end, an intervention of some kind. However, it may be an end in itself: an expert report for the courts, or an assessment for the social, health, or educational services.

The way in which clinical formulations are conceptualized, and the consequent selection of data describing the client's problem, varies not only in terms of its purpose but also (as we shall see) in relation to the theoretical assumptions of the practitioner. The process usually involves four kinds of activity, summarized by the mnemonic ASPIRE (Sutton & Herbert, 1992), which represents the stages in working up a case study and, more specifically, a clinical formulation:

(i) *Stage 1, Assessment (AS)*:
  (a) focusing on the "What?" question—that is what is/are the problem(s)?
  (b) focusing on the "Which?" question—that is which of the problems are to be addressed, and in what order?
  (c) focusing on the "Why?" question—that is why have the problems arisen?
(ii) *Stage 2, Planning (P)*: focusing on the "How?" question—that is, how are we (practitioner and clients) going to address the problems?
(iii) *Stage 3, Implementation of the intervention (I)*
(iv) *Stage 4, Rigorous evaluation (RE)*.

This framework provides a process-guide which, like the experimental method it resembles, is well suited to problem-solving. It also serves as a means of organizing this chapter, which examines each of the stages: first within a theoretical and then within a practice-oriented context.

5.02.2 THE CLINICAL FORMULATION: THEORETICAL ISSUES

The process of formulating is, figuratively speaking, like the action of a funnel containing a series of filters which represent choice and decision points. They have the function of distilling a many-sided childhood problem into a relatively brief, formal statement about one's conclusions and recommendations. The resulting formulation may serve as an expert report of some kind, or as the groundplan for a hoped-for alleviation of the presenting difficulties. Whatever its purpose, it is likely to be an amalgam of analysis, explanation, and intentions that emerge at the end of the metaphorical funnel.

The broadly based strategic decisions will have been made at the opening to the funnel, namely the chosen ideological stance (e.g, scientist-practitioner or humanist) and the theoretical model (say, cognitive-behavioral, Rogerian, or psychodynamic) adopted for the casework. Of course, the choices may not be starkly either/or, as when an eclectic style of working is preferred.

5.02.2.1 Clinical Psychology as Applied Science

At the broad end of the figurative funnel referred to earlier, is a filter influencing the clinical information that is edited in (and out) of the formulation. It involves the choice or rejection of a science-based philosophy of clinical work.

Long and Hollin (1997) make the point that the history of the scientist-practitioner stance is marked by an enduring struggle between advocates and opponents of a research-based profession—which is to say, a practice that is based on an empirically based assessment of problems, and the application of validated treatment methods which are linked (through the formulation) to theoretically coherent causal hypotheses. The reliance of the "scientist-practitioner" on empirical methods generates a model for service provision that is competency based. For Gambrill (1983) this requires:

(i) a focus on the present, which may be related to events in the past, and requires clarification by an exploration of the past (e.g., the person's learning history, deeply rooted attributions and ideologies);

(ii) a focus on describing problems (assessing them by a process of surveying, selecting, and prioritizing difficulties) and operationalizing them by means of multidimensional measurable outcomes;

(iii) a focus on multiple response systems of an individual nature (e.g., cognition, affect, and physiology) and of an interactional kind (e.g., family, school, workplace);

(iv) a focus on positive behaviors and events;

(v) a focus on the relationship between behavior and events in the external environment that elicit or maintain it.

Criticisms of the validity of such an idealized model have centered, *inter alia*, on the inadequacies of positivist philosophies of science and experimental research methodology for addressing the complexities and subtleties of individual psychopathology, or the sheer irrelevance of research for clinical practice. An example would be the attitude to formulating held by psychologists in the tradition of Maslow (1954) and Rogers (1957) whose ideas have been seminal in humanist psychology. Individuals are viewed as having an inherent need to grow. This intrinsic motivation reveals itself in the process called self-actualization.

Such a positive, optimistic (critics might say romantic) view of human nature, has little in common with the kind of formulations which are so preoccupied with analyzing and cataloging the failures or deficiencies of parents and children that they neglect to see their strengths and potential for self-directed growth. Such analyses are regarded by the severest critics as mechanistic, demeaning, and dehumanizing.

This is a view that is strongly contested (e.g., Herbert, 1987a, 1987b) when it is used in an overinclusive manner to cover all science-based practice. For those clinical psychologists who see themselves as scientific practitioners, their appeal to science is not usually restricted to the results and methods of traditional experimentation alone. The limitations of experimental methods are well known in relation to the study of human behavior: its sensitivity to minor variations in sampling and procedure, leading to failures of replication (i.e., to findings which are not robust); its confusion of aggregate effects with generalized effects (i.e., tendency to confuse an aggregate statistical effect with a general or common effect); its tendency to prefer methodological rigor over conceptual or practical experience; its liability to experimenter effects and demand factors; its neglect of contextual factors; and finally, its presumption that the experimental method is superior to the other methods regardless of circumstance (see Bromley, 1986).

Bromley (1986), while acknowledging that the case study cannot replace other experimental methods, is of the opinion that certain problems, notably the issues faced by clinical psychologists, lend themselves to this method of investigation. He provides an invaluable guide to the effective use of the case study as a method of inquiry and formulation. He suggests that the findings of scientific case studies depend, not so much on the logic of statistical inference, but rather on the logic of substantive reasoning. Case studies, he proposes, can generate and test abstract and general theories; when used systematically they are capable of combining basic and applied science—not only psychological, but also legal science. Bromley argues plausibly for the application of what he calls a quasi-judicial approach to clinical problems, drawing on the knowledge base of legal science for guidance on (among other things) rules of evidence. Some practical implications are studied later in this chapter.

With regard to the other criticisms of the irrelevance of research, Cohen's (1977, 1979) surveys indicated that 40% of mental health professionals thought there was no relevant research to inform clinical practice, while the other 60% were of the opinion that less than one-fifth of research articles had any applicability to professional settings. It is perhaps surprising that research and practice should be thought of by so many as mutually unhelpful or, at worst, incompatible, given that problem-solving is an essential ingredient of both of them (see for different sides of the argument, Meehl, 1971; Peterson, 1976; Shakow, 1978; Watzlawick, Weakland, & Fisch, 1974).

Long and Hollin (1997) point out that the growth in clinical research which was a concomitant of the burgeoning of learning-theory-based (behavioral) methods of intervention, from the 1950s onwards, led to the development of innovative and flexible research designs and methods (see Barlow & Hersen, 1984). This meant that practitioners could link more confidently, and in logical fashion, their application of particular therapeutic methods/techniques to an empirically based theory of dysfunctional behavior. This relationship could then be clearly enunciated in a formulation and, furthermore, validated by means of $N=1$ designs. It has to be said that by no means was everyone convinced by such confidence in the logical linkage between learning theory and its therapeutic methodology (e.g., Erwin, 1979; Power, 1991). Another development that broadened the appeal of the scientist-practitioner viewpoint was the growing sophistication and respectability of qualitative research methods (e.g., Marshall & Rossman, 1989). Certainly, it has gone some way to mitigate the accusations of triviality aimed at the kind of research that is narrowly or obsessively "quantiphrenic" in its definition of legitimate methodology.

Despite such developments, reservations about the validity of a scientist-practitioner approach to generating clinical formulations continues to inform debate on the appropriate model for practice and, thus, training in the profession (see Andrews, 1989; Gambrill, 1990; Pilgrim & Treacher, 1992, for discussions of these issues).

5.02.3 THE ASSESSMENT PROCESS

The filters mentioned above play a critical role in determining what information is pertinent in the investigation of the client's problems, and which assessment methods will generate the clinical data necessary for an intervention. As the funnel narrows, so does the focus of the questions: "What?," "Which?," "Why?," and "How?" that give direction to the assessment.

The final "How?" query leads to the strategic formulation of a precise treatment plan, and the tactical specification of methods/techniques for bringing about change.

5.02.3.1 The "What?" Question

Hepworth and Larsen(1990, p. 155) define this question in terms of "gathering, analyzing and synthesizing salient data." But what constitutes salient data? Answers to this question are likely to vary (alarmingly, some would say) according to the clinician's therapeutic model and its theoretical assumptions. At its most basic is the issue of what constitutes a priority in any clinical investigation: description or categorization—a debate centering on the concepts of assessment and diagnosis. Clinical psychologists tend to prefer the term "assessment" to "diagnosis"—a word that means to distinguish, and is essentially about classification. An assessment conveys the broader descriptive implications of an investigation or analysis of the causes or nature of a condition, situation, or problem. These implications are supposed to be neatly encapsulated within a diagnostic label, but, in reality, seldom are.

The differentiation of the "What?" and "Why?" questions is somewhat artifical as they tend to merge as the assessment progresses. When the collection of sufficient data allows for it, there is a formal explication: the formulation of a set of explanatory hypotheses describing the determinants, historical and current, of the problem. These, in turn, generate possible solutions to the client's difficulties.

5.02.3.2 The "Which?" Question

One of the clinicians' first choices is where to concentrate their attention in what may present itself as a welter of conflicting claims, complaints, and accusations. Some of the problems will reside mainly within an individual (e.g., fear) but even then they will have repercussions. The individual's parents (or partner) worry about his or her suffering. Or the problems put constraints on the activities of the family, as when holidays are not possible because (say) a mother is agoraphobic, feeling anxious about leaving the house, and panicky in a crowded shopping center.

Other problems arise from the relationships and interactions (such as the give-and-take transactions) between people. For example:

(i) mother–father: disagreements about the children, quarrels over decisions, other marital difficulties;

(ii) parent(s)–child: management difficulties, disappointments over the child's achievements or lack of them;

(iii) child–parent(s): resentments about being "babied," complaints about unfairness, or favoritism;

(iv) child–child: sibling rivalry, jealousy.

As the assessment proceeds, the client(s) may be one of these persons (a school-refusing child), two of them (the embattled parents), or all of them (the family). The practitioner might even have to work more widely in the neighborhood, at the school, or with other agencies in the community. Who the client is, is in fact a vexed question, as can be seen when examining systemic ideologies.

5.02.3.3 The "Why?" Question (Etiology)

Etiology—the study of causes of disease—has served physical medicine well. Research firmly rooted in scientific philosophy and methodology has led to an impressive armamentarium of rational treatments which have done much to reduce human suffering. Not surprisingly, psychologists and psychiatrists entertained high hopes of emulating their physician colleagues by endeavoring to work within a scientific and (notably, for psychiatrists) a medical framework. The complexity of psychiatric and psychological problems has led to great confusion in the analysis of causes (Herbert, 1994a). There is an all-too-human tendency to oversimplify and think of causality in linear, univariate terms: *A* causes *B*; *B* is the effect of *A*. However, there is no limit to the analysis of causes. One finds not a single antecedent, not even a chain of antecedents, but a whole interlacing network of them, as indicated by this complexity in Fielding's formulation of an adolescent patient's obsessional anxiety (Fielding, 1983) (see Figure 1).

One point of agreement that appears to emerge from the plethora of findings about problems like these is their multicausality. Mayer-Gross, Slater, and Roth (1955) make the point that in medicine it is necessary to deal with causes of all kinds, not just those that are

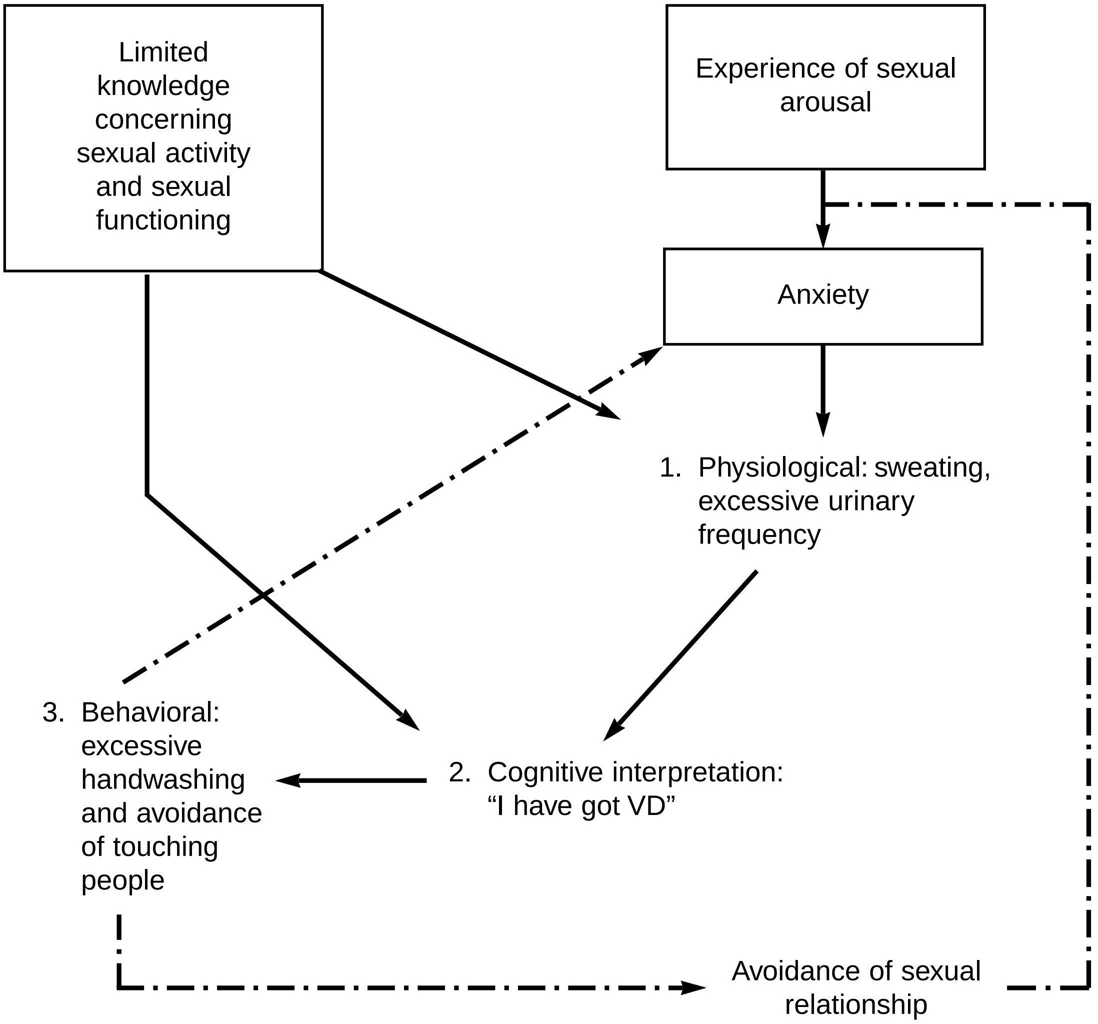
Limited knowledge concerning sexual activity and sexual functioning
Experience of sexual arousal
Anxiety
1. Physiological: sweating, excessive urinary frequency
2. Cognitive interpretation: "I have got VD"
3. Behavioral: excessive handwashing and avoidance of touching people
Avoidance of sexual relationship

Figure 1 Clinical formulation of obsessional hand-washing in a 16-year-old male adolescent. (The Practice of Clinical Psychology in Great Britain, by D. Fielding (1993) Chichester: Wiley Copyright 1994 by Wiley. Reprinted with permission.)

both necessary and sufficient. In searching for a cause of some phenomenon, we are really searching for a quantitative relationship. If *A* is the necessary and sufficient cause of *B*, then there is a one-to-one relationship between *A* and *B*. If *A* is a necessary but not sufficient cause of *B*, then there is no *B* without *A*, but *A* may be combined with *X* or *Y*, instead of with *B*. If the variety of these *X*s and *Y*s is great, the causal relationship, though it still exists, is thereby attenuated. If *A* is neither necessary nor sufficient, then there are *A*s without *B*s and *B*s without *A*s, and the strength of the causation will depend on the proportionate relationship between *A*/*B* and *A* on the one hand and *B* on the other.

Mayer-Gross et al. (1955) observe that it is not difficult to founder in a causal network. We are less likely to do so if we take into account the quantitative aspect of causation. It is at this point that so much psychological thinking loses cogency and direction. Quantitatively important causes are tangled in a knot of others with only a slight or entirely unknown quantitative relationship with the effects we are interested in.

The reality for psychologists and psychiatrists is the complex, multivariate nature of most of the phenomena they study. In their clinical formulations, they have to take into account both intrinsic (organismic) factors and extrinsic (environmental) influences. The latter may consist of direct (proximal) and indirect contemporary (contextual), as well as indirect predisposing (organismic) and historical (distal) determinants (see Figure 2).

The tendency for clinical child psychologists to work in multidisciplinary teams which include, *inter alia*, psychiatrists, speech therapists, and social workers, or in close collaboration with medical colleagues (e.g., GPs and

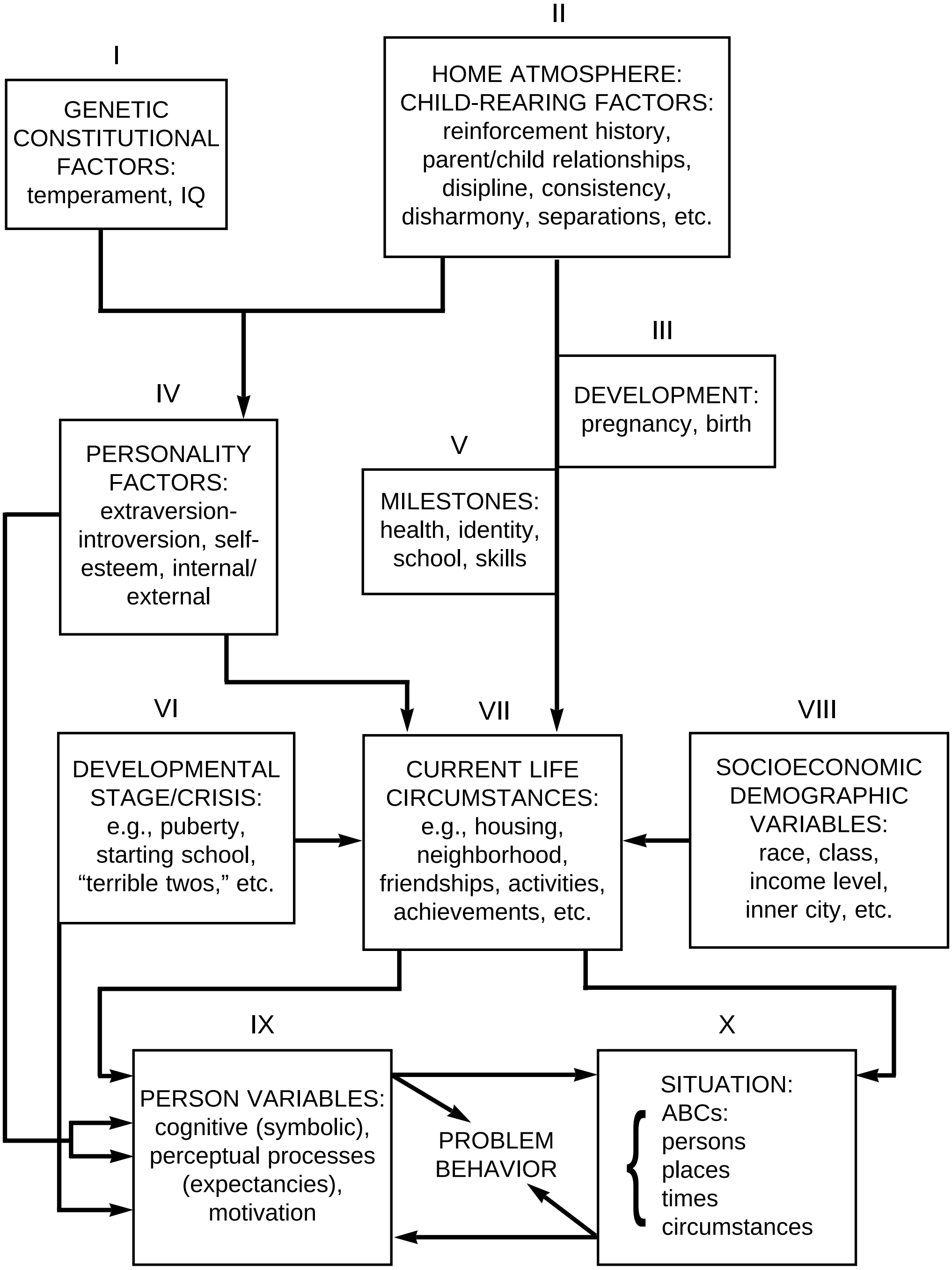
I
GENETIC CONSTITUTIONAL FACTORS: temperament, IQ
II
HOME ATMOSPHERE: CHILD-REARING FACTORS: reinforcement history, parent/child relationships, disipline, consistency, disharmony, separations, etc.
III
DEVELOPMENT: pregnancy, birth
IV
PERSONALITY FACTORS: extraversion-introversion, self-esteem, internal/ external
V
MILESTONES: health, identity, school, skills
VI
DEVELOPMENTAL STAGE/CRISIS: e.g., puberty, starting school, "terrible twos," etc.
VII
CURRENT LIFE CIRCUMSTANCES: e.g., housing, neighborhood, friendships, activities, achievements, etc.
VIII
SOCIOECONOMIC DEMOGRAPHIC VARIABLES: race, class, income level, inner city, etc.
IX
PERSON VARIABLES: cognitive (symbolic), perceptual processes (expectancies), motivation
PROBLEM BEHAVIOR
X
SITUATION: ABCs: persons places times circumstances

Figure 2 The 10-factor clinical formation of causation ("Psychology and crime," by R. G. V. Clarke, 1977; *Bulletin of the British Psychological Society, 30,* pp. 280–283. Copyright 1977 by the British Psychological Society. Reprinted with permission).

pediatricians) is understandable when we look at this complexity of physical, social, economic, and psychological influences on psychopathology. Behavior results from a very complex interaction of psychosocial and biological determinants (Herbert, 1991a). The important caveat in relation to clinical formulations is not to neglect, through a commitment to a psychogenic model or biogenic theory of psychopathology, the possibility of complications from the other side of the causal equation.

The trouble with the search for causal reasons is that there are several levels of explanation which may apply to a particular problem. This is illustrated by a young child who has a learning problem which is affecting her ability to read. On top of this, she has certain behavioral problems, as constant failure is causing her morale to sink lower and lower. Diagnostic tests carried out by the psychologist show that she has a visual-perceptual handicap. Remedial treatment is available to help her overcome her reading problem. The remedial teacher could teach her to recognize words and their meanings through the medium of her other sensory modalities. Here, the visual handicap is one explanation of the reading difficulty. Some clinicians would rest content with this level of causal explanation—the so-called instrumental level of explanation. The explanation is sufficiently precise to be instrumental in planning some therapeutic measures to mitigate the cause–effect sequence of events. Or it may simply be instrumental in providing the individual with a satisfactory account of some event. It is sufficient in that it explains, and raises no further questions in his or her mind.

While this explanation provides an account of an important antecedent condition in the problem—in this case the visual-perceptual handicap—other clinicians might claim that the diagnostician should search for "original" or "deeper" causes. In this case they would feel it necessary to determine whether the visual problem is due to organic causes, perhaps resulting from complications during her birth or a genetic predisposition.

5.02.3.4 The "How?" Question

The assessment of causal influence should, in an ideal world, inform the choice of therapeutic strategy or some broader-based community intervention. Only too often the assessment data, like the occupants of Procrustes' bed, are made to fit the favored therapeutic model. There is an implied suggestion that a particular approach can be applied to all problems. Sadly, faith, rather than evidence, is mostly what is on offer to support the more Panglossian prospectuses. The most durable and popular of the psychological therapies for children have been relationship therapy, play therapy, child analysis, behavior therapy/modification, and the relative newcomers cognitive behavior therapy, behavioral and nonbehavioral family therapy, and behavioral parent training. There is some overlap in process (if not in name) between the various psychological therapies, but they are essentially different, not only in their outward and visible manifestations, their "technology," but also, as one would expect, in their underlying theoretical rationale and, consequently, their goals.

Although there is no one coherent theory of therapeutic change, there is some overlap in goals. For Kazdin (1987, 1988), the goals consist of improving adjustment and functioning in both intrapersonal and interpersonal spheres and reducing maladaptive behaviors and various psychological (also physical) complaints. Even when there are differences of emphasis in therapeutic goals, all treatments share the objective of producing change.

Any form of therapy (be it psychotherapy or behavior therapy) involves psychological influence—implicit or explicit—deployed in the setting of a relationship between client and therapist. The dichotomy drawn between interpersonal and relationship factors on the one hand, and technique on the other, seems somewhat forced; indeed, it may be false, because it is doubtful whether these two major components in the therapeutic enterprise can ever be isolated in pure and separate cultures.

5.02.4 EXPLANATORY MODELS

One can say with confidence that there are considerable variations on the theoretical themes that inform similarly-named treatments. Different theories of family process generate different schools of family therapy (Herbert, 1997a; Vetere & Gale 1987). Play therapy also draws on different nuances of childhood "psychopathology" theory, and thus takes different forms (including active play therapy, release therapy, and the play interview) and is being applied on a dyadic, one-to-one basis (e.g., Axline's approach) or in a group format (e.g., Filial therapy).

Psychological therapies can be differentiated on two major dimensions:

(i) a passive vs. an active role on the part of the therapist;

(ii) the method (technique) vs. relationship as the main agent for change.

They can further be distinguished in terms of five attributes:

(i) the goals (or purpose) of therapy, for example, supportive, prescriptive, or exploratory;

(ii) the means by which goals are sought; this encompasses the theoretical rationale, for example, behavioral, cognitive, client-centered, or psychodynamic, (in child therapy these approaches may be mediated by talking to, playing with, rewarding the child or rehearsing skills with him or her);

(iii) the modality of therapy, for example, individual, group, couple or family;

(iv) the level of expertise of the therapist;

(v) the adoption of an expert or collaborative approach to clients.

The range of choices of individual psychotherapies—notably in the case of adolescents—constitutes a "menu" which almost rivals the feast available to adults: psychodynamic, client-centered, behaviorally oriented psychotherapy, communication, drama, role-construct, social system, multimodal, transactional, existential, gestalt, to mention a few. These are relatively broad approaches or perspectives. When we come down to techniques a conservative estimate suggests that there are more than 200 to choose from (Kazdin, 1988; Tuma, 1989).

What does such profusion mean? The sceptic might say that when so many treatments are in use, applied often to the same kinds of problem, it means that we lack a firm conceptual grasp of causation. In any event does it matter if the critics are correct in saying that psychotherapy has no adequate unifying theory? Why not simply consider treatment as a technology without recourse to any distinctive foundational theory? Erwin (1979) is of the opinion that we cannot afford to ignore the study of theoretical foundations. He states that an adequate theory might provide some important practical benefits. For example, if better clinical techniques are to be developed, it will be helpful to know exactly which features of the successful techniques are therapeutically productive and why, also which techniques have a greater antecedent likelihood of success than do others. This is a cue to look briefly at the major therapeutic models and the formulations which predicate their use.

5.02.4.1 The Medical Model—Diagnostic Formulations

Nowhere is the influence of medical ideologies in clinical child psychology more pronounced than in the reliance on classification and the diagnoses of children's psychological problems. Underlying the creation of classificatory systems (e.g., the *DSM-IV* (American Psychiatric Association, 1993) and *ICD-10* (World Health Organization, 1992) and taxonomies of "symptoms" is the assumption that specific syndromes (disease patterns) with identifiable and specific causes (etiologies) can be diagnosed. It might be added that the only justification for applying diagnostic categories to children is that not only should they imply reliable descriptive criteria and clear causal theories, but also that treatment implications should flow from the choice of label.

The particular issue of reliability in the assignment of diagnostic categories remains a contentious one (Zwick, 1983). This is a potentially major stumbling block as the validity of any system of measurement or classification depends, in large part, on its reliability. Achenbach and Edelbrock (1989) made the point that in order to be effective, formal diagnoses (and diagnostic formulations) require reliable taxonomic distinctions so that we can link cases that share useful similarities and to distinguish between cases that differ in important ways. The various *DSM* criteria for childhood psychopathology consist mainly of lists of behavioral-emotional problems rather than clearly identifiable physical abnormalities, and they are problems to be judged as present or absent. Unfortunately, as the authors state:

> most of these problems are not easily categorized as present or absent. Instead they vary with the child's age, situation, and interaction patterns. Whether the child is deemed to have a particular problem depends on when and how the child is assessed. Because many of the problems for which mental health services are sought occur outside the clinic setting, assessment typically requires information from people other than clinicians, such as parents and teachers. The different situations in which they see the child and the different roles they play with respect to the child are bound to affect what they report. (p. 55)

Achenbach, McConaughty, and Howell (1987) carried out a meta-analysis of 269 samples (in studies published from 1960 to 1986) of correlations between different informants (parents, teachers, trained observers, mental health workers, peers, and the children themselves) reporting children's behavioral-emotional problems. The mean correlations between informants playing similar roles with respect to the children, including pairs of parents, teachers, mental health workers, and trained observers ranged from 0.54 (mental health workers) to 0.64 (pairs of teachers) with an overall mean of 0.60. This substantial, although far from perfect, level of agreement dropped considerably when informants had different roles with the children. The mean correlations ranged from 0.24 (between parents and mental health workers) to 0.42 (between teachers and direct observers) with an overall mean of 0.28—statistically significant but low. With regard to the self-reports and others' reports of the same child, the mean correlation of 0.22 is statistically significant but even lower.

With regard to the face-validity of psychiatric categories for professionals, a survey of 1000 qualified or trainee American psychiatrists by Jampala, Sierles, and Taylor (1986) is of interest. A full 48% did not believe in the validity of the *DSM-III* criteria and 55% would have ceased to use it, had it not been required. They felt that the diagnostic categories failed to accommodate almost half of their patients. This is a serious indictment since the salient point about a classification-based formulation is that it needs to be comprehensive and to group individuals according to their distinguishing problematic behavioral, cognitive, and affective patterns. In the taxonomics of childhood disorder there is probably no better example of the difficulties, contradictions, and confusions than the criteria for attention deficit hyperactivity disorder and hyperkinesis provided in the *DSM-IV* and *ICD-10* systems, respectively.

Reviews of the literature suggest that despite interesting correlations between "problematic" childhood behaviors, the vast majority of "disorders" (it is difficult to avoid medical usage) cannot be conceptualized—except figuratively—as disease entities. Unlike many of the acute physical illnesses the structure of a presenting psychological problem, the etiology of the problem, and the implied treatment do not closely map on to each other (Sturmey, 1996). Achenbach and Edelbrock (1989) conclude a review of these issues by acknowledging that there is little agreement about the proper nature and role of diagnosis with respect to the psychopathology of children.

Whatever the reliability and validity of diagnostic schemes, there is an objection raised by many psychologists (and psychiatrists) to what they see as invidious "labeling" and stigmatizing of individuals by reducing complex "life problems" to a restrictive medical category. There is concern about how parents', teachers', and children's attributions, perceptions, and actions are affected by being told by an expert that the child has a disorder of some kind. Such reification may absolve parents from any sense that they can influence the difficulties. It may "tell" the child that they are different from other children, and furthermore, not in control of their actions—a pessimistic message to instill in someone at an impressionable stage of life. Most worring is the absence of a clear relationship between most global diagnostic classifications applied to particular children and any clear etiology or focused therapeutic/remedial program of action (Herbert, 1964).

The reality is that childhood signs of psychological abnormality are, by and large, manifestations of behavioral, cognitive, and emotional responses common to all children. Their quality of being dysfunctional lies in their inappropriate intensity, frequency, and persistence; most childhood disorders differ in degree, rather than kind, from normal behavior. The exceptions to this generalization are rare (see Herbert, 1994b). That generalization does, however, raise ethical problems for the practitioner. Leung (1975) observes that, whereas the medical model provides a reasonably unambigious ethical rationale for treatment by an *a priori* standards of "health" and "pathology," behavioral definitions tend to construe abnormality as not in itself different from normal behavior. If abnormal behavior is considered learned, therapists are involved in making, or at least concerning themselves with, value judgments that some other behavior would be preferable—a social, subjective, and therefore potentially prejudicial decision. This leads inexorably to the following questions, the answers to which have the status of ethical imperatives for the therapist as agent of change: "To whom is the behavior undesirable? Is it *really* in need of modification, and if so, to what must it reasonably change?"

Of course, the medical model has produced highly effective aids to the treatment of childhood psychopathology, but it also raises a moral dimension, particularly in the use of potentially harmful and/or addictive medication. As in the case of "owning" a diagnostic label, the attributions for a child's locus of control that follow from the regular use of drugs prescribed to make him or her "quieter" and "better behaved" are matters of concern.

A pitfall of using diagnostic classification in clinical formulations is the tempting but illusory impression it can give the practitioner of having explained the problem, when all it amounts to is a renaming process, and consequently a "thought-stopper." Take the boy who is always on the go, fidgets, is endlessly flitting from one activity to the next, and is very difficult to manage. The teacher who refers him to the clinic wishes to have an answer to the question: "Why does the child have such poor concentration, and why is he so disruptive—forever interfering and going "off task," defying my requests?" The report that comes back states that the child is ADHD (hyperactive). If asked to explain this diagnosis ("How do you conclude that the child is hyperactive?"), the clinician is quite likely to have to return to the observations they were asked to explain in the first place: "Because the child is inattentive, fidgetty, overactive, and noncompliant." Many of the explanatory constructs in the mental health field are tautologies of this kind.

Ross (1968) finds parallels between the practitioner's preoccupation with classification and the fairy tale in which the opportunity for a princess to live her life happily ever after depends on her discovering the name of an ill-tempered creature. She goes to great lengths, and upon achieving her goal earns her salvation. Ross calls the preoccupation with seeking a name for a child's problem in order to "save" it, the "Rumpelstiltskin fixation." It should not be the primary task of psychologists, he admonishes, to identify a disease by its signs and symptoms, but rather to get on with their real job of describing children and their difficulties in terms adequate enough to bring effective interventions into play, and with them, positive outcomes.

In the real world, of course, the clinician does not always have the luxury to be too pedantic (or precious) about such matters. While there may be misgivings about labeling children, there may be compelling policy, planning, service development, resource mobilizing, and research reasons for doing so. For the private practitioner there is the consideration that insurance companies insist on formal diagnoses.

5.02.4.2 The Statistical Model—Multivariate Analyses

Statistical (and particularly factor analytic) methods have proved popular in the search for clusters of problems which might lend support to the classifications built upon clinical observation and experience. What is advocated is an empirical approach which involves a minimum of assumptions and constructs regarding the causation of behavioral disorders. Despite the range of instruments that is available and the number of dimensions measured by these instruments, two major dimensions emerge from most analyses: problems of an *undercontrolled* type (variously referred to as conduct disorder, aggressive, externalizing, acting out) and an *overcontrolled* type (emotional disturbance, personality disorder, inhibited, internalizing, anxious) (Quay, 1986).

There is some limited overlap between behavioral dimensions (obtained by psychological measurement and factor analysis) and *DSM* criteria (see Quay, 1986). Their limitations with regard to implying causation or generating individual treatment options are much the same, although Achenbach and Edelbrock (1989) have illustrated ways in which the Child Behavior Checklist and Child Behavior Profile can be put to practical use, by providing quantitative descriptions of behavioral competencies and problems at various levels.

5.02.4.3 The Psychoanalytic Model—Psychodynamic Formulations

The "psychodynamic approach" is a superordinate title for a variety of therapeutic developments which had their origins in Freudian psychoanalysis (Lee & Herbert, 1970; Rycroft, 1966; Tuma, 1989). The term encompasses theories of personality and therapies which assume the existence of unconscious mental processes, and which concern themselves with the elucidation of motives and the transference relationship, that is, the transfer on to the therapist of attitudes, longings, and feelings attaching to significant others.

Farrell (1970), in his synoptic account of psychoanalytic doctrine, states that the classical Freudian theory is primarily a theory about mental energy. This takes two fundamental forms which, in the early and better-known version of the theory, are the sexual and the self-preservative; and this energy goes to determine the character of the stages (oral, anal, etc.) through which the mind develops. In the course of this development the mind becomes differentiated into the structures of id, ego, and superego, cut across by the critically important barrier separating the unconscious from the preconscious elements of the mind. The total balance and distribution of energy is controlled by the mechanisms of defense, such as repression, regression, displacement, and so on. Where this control had not been satisfactorily achieved in the past, the repression barrier keeping the unconscious instincts of the system at bay is apt to break down under frustration. The unconscious comes to the surface in the compromise form of neurotic or "acting out" symptoms.

The neo-Freudian ideas of Melanie Klein (1975) have a different emphasis. This is primarily a doctrine about the "object relations" an infant forms (especially with the objects of mother and her breast) and about the internal world of the infant's unconscious fantasies. From the outset, infants experience the anxieties of living and are faced, in particular, with the breast that both satisfies and frustrates them. They deal with these difficulties by means of the defenses of introjection, projection, and splitting. Thus, to begin with, infants introject the breast (i.e., incorporate it in fantasy) and split the good breast from the bad. The character of infants' object relations—their good or bad quality—and the extent to which they get through various early stages satisfactorily are all of crucial importance to them. Therefore, the liability of the grown-up man or woman to break down, and the character of his or her collapse when it comes, are critically dependent on what happens inside the infant during the first year of life.

Therapists working with a psychodynamic model give much more weight to intrapsychic events in their formulations than behavior therapists because their perspectives are characterized by different epistemologies. These differences are illustrated in case history formulations from the mid-1920s. Where Little

Hans, with his deeply symbolic fear of horses, became the much-quoted standard-bearer of the psychoanalytic paradigm, Little Peter (age 2 years 10 months at the beginning of his treatment for a fear of rabbits) became the exemplar of the behavioral paradigm (see Freud, 1926; Jones, 1924).

Freud was essentially a proponent of *verstehende* or "understanding" psychology; he wished to answer the "whys" of human conduct. Despite his training in medicine and research work in physiology he was not content with the "mere" observation of surface behavior or with the investigation of neurophysiology; he was determined to explain the mental mechanisms underlying complex sequences of behavior. He wished to account for the whole of human mentation at the psychological level of explanation—formulating the experience of the neurotic patient, in terms of the meaning of his/her symptoms.

Psychoanalytic formulations (whatever the present-day school of psychoanalysis) tend to be derived from extensive arrays of molar segments of self-report and verbal exchange with the therapist, collected (usually) over a considerable period of time. The records of the therapy would be copious. By way of contrast, the clinical child psychologist, in scientific mode, is concerned with a deliberately curtailed and simplified series of events in the life of his/her clients. Analysts see superficiality in such parsimony—a misdirection of "gaze" which will simply result, if only presenting "symptoms" are removed, in other symptoms (the outer and visible signs of the "real" underlying difficulties) coming to the surface. This is the disputed phenomenon called "symptom substitution."

There have been intense debates about the status of psychoanalysis and its many-faceted progeny, the "psychodynamic approach." It is difficult for the hermeneutic framework of the psychodynamic model to be integrated within that of the scientist-practitioner approach (see Farrell, 1970; Fisher & Green, 1996; Long & Hollin, 1996; Pumpian-Mindlin, 1952; Rachman, 1963; Wachtel, 1977; Webster, 1995). And yet Freud's aim was to establish a "scientific psychology," and his wish was to achieve this "metapsychology" by applying the same principles of causality as were at that time considered valid in physics and chemistry.

For Farrell (1966), the fact that no usable "correspondence rules" (i.e., operational definitions) had been constructed for any of the large numbers of concepts of psychoanalytic theory, was a major stumbling block in testing its validity. If patients ask what terms like "repression" mean or what phrases like Klein's "fantasized incorporation by the infant to the breast" indicate, there is likely in Farrell's words (p. 22) to be "elucidation by means of, in part, the concept of psychic energy . . . an explanation *per obscurius*, or an abrupt plunge into psychological absurdity."

Rycroft (1966) maintains that what Freud was engaged in was not explaining the patient's symptoms causally, but understanding them in order to give them meaning. This was not the scientific procedure of elucidating causes, but the semantic one of making sense of them. He says it can be argued that much of Freud's work was discovering that neurotic symptoms are meaningful, disguised communications, but that owing to his scientific training and allegiance, he formulated his findings in the conceptual framework of the physical sciences.

Psychoanalytically-oriented psychotherapy or play therapy with children is a less ambitious and demanding form of psychoanalysis than the form used with adults. Here the relationship between child and analyst is a means whereby therapeutic goals are achieved. The child is helped, by means of play and therapeutic conversations:

> to undo the various repressions, distortions, displacements, condensations, etc., which had been brought about by the neurotic defence mechanisms, until, with the active help of the child, the unconscious content of the material is laid bare. (Anna Freud, 1946, p. 71)

5.02.4.4 The Personality Typology Model—Trait Formulations

The personality structure approach to assessment postulates personality dimensions, in part inherited, that influence what and how children learn as they grow up. It has been observed that every person is, in certain respects, like all other persons, like some other persons, and like no other person. The nomothetic disciplines which favor the philosophy and methods of the exact sciences have been applied to the study of the highly generalized personality attributes (types and traits), while idiographic methods have been used to explore the aspect of uniqueness and individuality. According to one theory, the possession of particular combinations of personality characteristics (e.g., extraversion/introversion and neuroticism) determine whether certain children are predisposed to acquire dysthymic or conduct disorders (Eysenck, 1967). Given an extreme position on the continuum toward introversion or extraversion (which has implications for the individual's conditionability) and a labile or reactive autonomic nervous system (a high rating on neuroticism), the child is thought to be particularly susceptible to acquiring phobic anxiety through ease of social conditioning, or a conduct disorder because of its relative failure.

There are several limitations to the use of personality typing in the formulating process. The situations sampled are restricted in range: for example, behavior at home or at school. Furthermore, assessment methods tend to emphasize verbal behavior such as interviews, self-reports, and questionnaires. In the applications of such methods of assessment, personality differences in behavior across settings are seen as error variance in an imperfect measurement system. Greatest weight is placed on the initial assessment and diagnosis rather than on repeated, ongoing assessments. In personality-type assessments the person's behavior is compared with statistical norms.

The practical implications of an assigned personality "diagnosis" or profile for a treatment plan are limited (Herbert, 1965; Moore, Bobblitt, & Wildman, 1968). The issue of situation specificity makes it difficult to make predictions or arrive at precise conclusions of an individual kind, on the basis of generalized constructs. On the other hand, the early assessment of temperament—the constitutional aspects of personality—into its major categories "difficult," "easy," or "slow to warm up" and its behavioral referents, may allow preventive clinical work to be planned (Thomas, Chess, & Birch, 1968). A mismatch between parental attributes and the child's temperamental style can result in an extended series of mutually unrewarding interactions which are foundational for later relationships. They can lead to depression in mothers (in particular) and faulty or incomplete socialization in children (see Iwaniec, 1995).

The idiographic disciplines, such as biography, literature, and history, endeavor to understand and illuminate some particular event. The use of idiosyncratic and subjective projective techniques is an example of this tradition in clinical work. The case study is an exemplar of what Allport (1937) believes is an ideal combining of the idiographic and nomothetic approaches for the fully rounded appreciation of the person and their situation. In the view of Bromley (1986) it is the bedrock of clinical investigation, and part of the scientist-practitioner's armamentarium.

5.02.4.5 The Systemic Model—Family Therapy Approaches

An individual in a family system is affected by the activities of other members of the family. Whereas the traditional treatment model tended to identify the nominated client or patient as the unit of attention (e.g., the child referred to the child guidance clinic), the focus of assessment in the light of this interactional frame of reference is far more broadly conceived. Thus the focus of help is not necessarily prejudged as only the child who was referred; rather, the unit of attention becomes the family (or one of its subsystems).

The systemic perspective is well-illustrated in family therapy, a fashionable approach to childhood psychological disorder. There is no one therapeutic entity under the rubric "family therapy," but several schools or paradigms. For example, there is the structural school of Minuchin and his colleagues (Minuchin, 1974). Strategic family therapy has its origins in the Palo Alto research group led by Bateson (e.g., Bateson, Jackson, Haley, & Weakland, 1956; Jackson, 1957). Humanistic, existential therapies of the 1960s, such as Gestalt therapy, psychodrama, client-centered therapy, and the encounter group movement, influenced the theory and methods of various experiential family therapies—challenging the positivist tenets of the more problem-focused schools of family therapy (e.g., Satir, 1967). Behavioral family therapy and cognitive-behavioral therapy have their intellectual roots in social learning theory (Patterson, 1982). Psychodynamic and object relations perspectives, derived from psychoanalytic theory, found an exponent in Ackerman (1958). The intergenerational perspective, which looks beyond the immediate family circle and enlists the cooperation of others in resolving the family's distress, is associated with the names of pioneers such as Bowen (Bowen, 1966). If this were not a daunting enough choice, the clinican could look to the Milan Systemic Approach, the McMaster Family Model, Problem-solving Therapy, Brief Family Therapy, or Social Network Therapy (see Herbert, 1997a). Piercy, Sprenkle, and associates (1986) describe 54 technique skills for the major family therapies.

As with other therapeutic models the paths on the journey toward making a formulation are many and not always clearly signposted. What does unite most family therapies as they engage in their divergent formulations and treatment strategies, although at a high level of abstraction, is a perspective which requires that children's problems be understood as the consequence of the pattern of recursive behavioral sequences that occur in dysfunctional family systems. The systems approach, as it translates to family work, embraces the concept of reciprocal/circular causation in which each action can be considered as the consequence of the action preceding it and the cause of the action following it. No single element in the sequence controls the operation of the sequence as a whole, because it is itself governed by the operation of other elements in the system. Thus, any individual in a family system is affected by the activities of other members of the family, activities which his or her actions or decisions (in turn) determine. This recursive element makes the formulation of what is cause and effect extremely complex.

Family therapists, whatever their methods and theoretical underpinnings, tend to believe that behavior problems in children are symptomatic (or, indeed, artifacts) of dysfunctional family life; the goal of treatment is therefore the improvement of family functioning. The members are encouraged, by a variety of therapeutic strategies and homework tasks, to understand the alliances, conflicts, and attachments that operate within the family unit and to look at themselves from a fresh perspective and to seek alternative solutions to their dilemmas.

The family, as a small group, can be observed and assessed on a variety of dimensions: patterns of communication, cohesion and processes of decision making (see Vetere & Gale, 1987). The major concepts for an assessment and formulation in family therapy are listed by Dare (1985) as follows:

(i) seeing the family as having an overall structure;

(ii) understanding the symptom as having a potential function;

(iii) understanding the location of the family on the life cycle;

(iv) understanding the intergenerational structure of the family;

(v) making an overall formulation linking the preceding four features;

(vi) linking the formulation to appropriate interventions.

Typically, an assessment by a family therapist (see Lask, 1987) might concern itself with whether there is:

(i) too great a distance between members of the family, leading potentially to emotional isolation and physical deprivation;

(ii) excessive closeness between members of the family, leading potentially to overidentification and loss of individuality;

(iii) an inability to work through conflicts, solve problems, or make decisions;

(iv) an inability on the part of parents to form a coalition and to work together, with detrimental effects on the marriage and/or the children;

(v) an alliance across the generations, disrupting family life, as when a grandparent interferes with the mother's child-rearing decisions;

(vi) poor communication between members;

(vii) a failure to respond appropriately to each other's feelings.

Some of the ways in which a child may contribute (wittingly or unwittingly) to a family's inability to cope with conflict have been described by Lask (1987):

(i) *parent–child coalition*, where one parent attacks the other, using one of the children as an ally;

(ii) *triangulation*, where both parents attempt to induce a child to take *their* side;

(iii) *go-between*, where a child is used to transmit messages and feelings;

(iv) *whipping-boy*, where one parent, instead of making a direct attack on the other, uses their child as a whipping-boy;

(v) *child as weapon*, where one parent attacks the other using the child as a weapon;

(vi) *sibling transfer*, where the children agree to divert the parents' arguing.

There are family therapists who neglect to assess the influence of wider social systems (e.g., school, neighborhood, sociocultural network). The social context is played down as a source of influence. It is an article of faith that the outward and visible signs of the family's problems, as manifested to a group of observers/commentators behind a one-way screen, represent reliably what is going wrong. How representative of a family's repertoire the samples of interaction (sometimes distorted by observer effects) are in reality, is a moot point.

Bronfenbrenner (1979) recommends the assessment of the following systems when producing a formulation:

(i) *microsystems*, which are the daily settings of the client (e.g., home, school, work), the physical environments, and the people within these settings with whom the client interacts;

(ii) *mesosystems:* the links between microsystems (i.e., the name and quality of connections between the environments);

(iii) *exosystems:* the situation (e.g., mother working full time) which may impact on the client's life, but in which he or she is not directly involved;

(iv) *macrosystems:* the broad institutional patterns and value systems of the culture which affect the meso- and exosystems.

Whiting (1963) points out that the most significant actions taken by parents, significant in the sense of having the most impact on children, have to do with their assignment of children to various niches or settings in a society, from crèches and nurseries/kindergartens to private schools, apprenticeships, helping sell produce in the market, or tending animals. Each of these settings has flow-on effects upon the people children encounter, their opportunities for play with peers, and their opportunities to acquire informal cognitive skills. These many and varied influences give the child his or her unique quality.

The individual client is also conceptualized in systemic terms. For example, Hepworth and Larsen (1990) make a formulation on the basis of describing (say) the child as a system composed of six intrapersonal dimensions:

(i) biophysical functioning (e.g., physical attributes, physical health, use and abuse of drugs and alcohol);

(ii) cognitive and perceptual functioning (e.g., cognitive level and flexibility, self-concept, judgment);

(iii) emotional functioning (e.g., emotional control, depression, anxiety);

(iv) behavioral functioning (e.g., deficits, excesses, patterns of concern);

(v) motivation (e.g., precipitating events);

(vi) culturally determined patterns (e.g., norms, values).

What we now have is an assessment process which may be focused on an individual (e.g., child), a group (e.g., family) or an organization (e.g., school). The first is sometimes neglected in family therapy formulations.

The family is thought of as dynamic in the sense that it is subject to continuous change. Like individuals, families are believed to pass through developmental life-cycle stages. Each life-cycle stage is believed to have its own accompanying set of tasks to be accomplished if families are to make successful transitions from one stage to the next (Carter & McGoldrick, 1988). A concept like that of life stages highlights a problem for practitioners assessing families. How do they adopt a standard by which to evaluate functional ("healthy") as opposed to dysfunctional ("unhealthy") patterns of family organization, transitions, boundaries, homeostatic (equilibrium-seeking) strategies, rules, roles, decision-making processes, communications, or any of the other attributes family therapists believe to be significant? Obviously, there are diverse cultural "norms" or values in these matters.

What is not always clear in clinical formulations which highlight the types of causal influence listed by Lask (above) is the precise mechanism whereby family patterns relate functionally to the day-to-day specificities (i.e., with regard to persons, places, times, or situations) of the referred problem. The majority of nonbehavioral family therapy processes are formulated in somewhat global terms and pitched at a high level of abstraction. It becomes very difficult to define operationally the key independent treatment variables at this level of strategy generality (de Kemp, 1995). At the dependent variable end of the therapeutic equation, emphasis can be placed on symptom reduction/removal alone (called first-order change), or a more systemic level of family transformation (second-order change). Second-order change is empirically and conceptually difficult to identify and quantify.

The primary goals of family therapists, as defined by family therapists from several disciplines, tend to be divergent. Furthermore, there is little consensus about which factors constitute evidence of an efficacious outcome. A survey (*inter alia*) of family therapists' primary goals for therapy (Group for the Advancement of Psychiatry, 1970) elicited the following objectives expressed as percentages with regard to all families:

(i) improved communication (86%);

(ii) improved autonomy and individuation (56%);

(iii) improved empathy (56%);

(iv) more flexible leadership (34%);

(v) improved role agreement (32%);

(vi) reduced conflict (23%);

(vii) individual symptomatic improvement (23%);

(viii) improved individual task performance (12%).

Practitioners are not always clear about whether they are concerned with symptom reduction, family restructuring, or both.

Critical surveys (e.g., Herbert, 1987a; Kaslow, 1980; Reimers & Treacher, 1995) suggest that pragmatic, structured family therapies (i.e., those concerned principally with behavioral change) have produced the most convincing evidence of effective therapeutic outcome for the psychosocial problems of children. A particular criticism of the family therapy field is their antiempirical stance and consequent rejection of basic research (e.g., Watzlawick, Weakland, & Fisch, 1974).

5.02.4.5.1 Behavioral family therapy

The work of Patterson in Oregon was part of, and the inspiration for, a rapid growth in what might be called "behavioral family therapy" (see Alexander & Parsons, 1982; Griest & Wells, 1983; Herbert, 1994b; Patterson, 1982). This approach—sometimes also referred to as "systemic behavioral therapy"—is in large part concerned with the assessment of naturally occurring environmental influences, specifically those occurring within the family between parents, siblings, and child, in order to modify deviant behavior and teach new skills and behavior repertoires. It is an elaboration of learning-theory-based behavior therapy, and finds much of its intellectual sustenance in social learning theory. The basic level of formulation remains the functional analysis to which is turned to; it illustrates neatly the scientist-practitioner way of working, the approach of the applied psychologist.

5.02.4.6 The Behavioral Model—Functional Analysis

Haynes and O'Brien(1990) define functional analysis as: "the identification of important, controllable, causal, functional relationships applicable to a specified set of target behaviors for an individual client" (p. 654). In the behavioral canon, the formulation is directed towards the precise identification of the antecedent, outcome, and symbolic conditions which control the problem behavior. Behavior theorists often refer to their assessment—as a simple mnemonic for clients—in ABC terms (see Figure 3).

From a behavioral perspective, a child who indulges in frequent acts of defiance and outbursts of verbal abuse would be observed in order to determine the specific circumstances and conditions in which the unacceptable behaviors occur. Such conditions may involve antecedent organismic (O) and/or environmental stimulus events (A) which precipitate or set the scene as discriminative stimuli for the outbursts and the consequent events (C). The B term in the ABC analysis can stand for the behavior(s) being studied and the beliefs (attributions, perceptions, interpretations) of the child/parents about what is happening, why it is happening, and what the outcomes are likely to be.

The formulation involving an ABC linear analysis is, of necessity (given the complexity of behavior), elaborated into a recursive sequence such that Cs become As that generate new Cs, and so on. Learning occurs within a social nexus; rewards, punishments, and other events are mediated by human agents within attachment and social systems and are not simply the impersonal consequences of behavior. Children do not simply respond to stimuli; they interpret them. They are relating to, interacting with, and

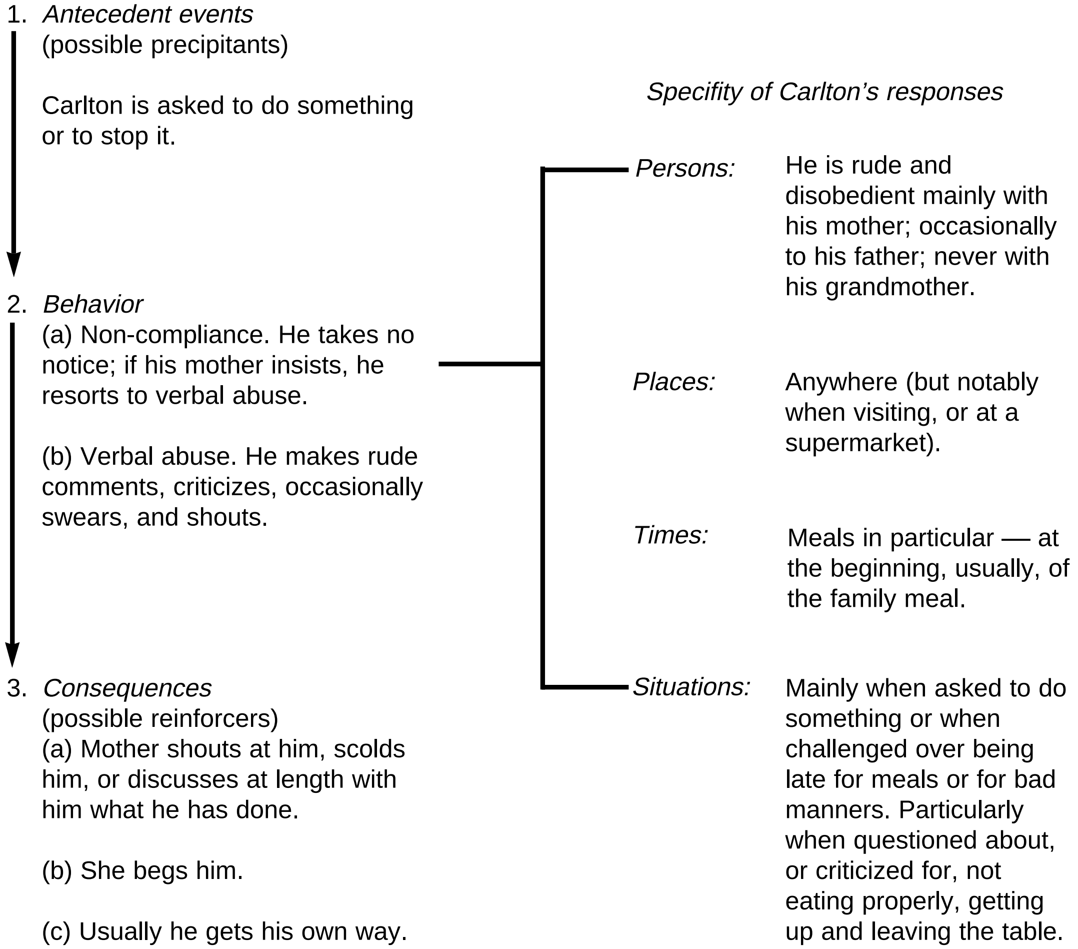

1. *Antecedent events*
(possible precipitants)

Carlton is asked to do something or to stop it.

2. *Behavior*
(a) Non-compliance. He takes no notice; if his mother insists, he resorts to verbal abuse.

(b) Verbal abuse. He makes rude comments, criticizes, occasionally swears, and shouts.

3. *Consequences*
(possible reinforcers)
(a) Mother shouts at him, scolds him, or discusses at length with him what he has done.

(b) She begs him.

(c) Usually he gets his own way.

Specifity of Carlton's responses

Persons: He is rude and disobedient mainly with his mother; occasionally to his father; never with his grandmother.

Places: Anywhere (but notably when visiting, or at a supermarket).

Times: Meals in particular — at the beginning, usually, of the family meal.

Situations: Mainly when asked to do something or when challenged over being late for meals or for bad manners. Particularly when questioned about, or criticized for, not eating properly, getting up and leaving the table.

Figure 3 The ABC of behavoir.

learning from people who have meaning and value for them.

Bandura (1977) suggests that stimuli influence the likelihood of particular behaviors through their predictive function, not because they are automatically linked to responses by occurring together; contingent experiences create expectations rather than stimulus–response connections. Not surprisingly, behavioral work based on this premise is increasingly systemic, dealing with caregivers and nominated child "patients" in their own right as individuals, and analyzing their relationships to each other (as dyadic attachment subsystems). Also studied are their communications, interactions, boundaries, and perceptions of one another within a holistic and dynamic family system.

5.02.4.6.1 Behavioral formulations

In identifying controlling variables, the contemporary causes of problem behavior may exist in the client's environment or in his or her own thoughts, feelings, or bodily processes (organismic variables), and they may exert their influence as proximal (current) influences which are direct in their effects and close in time to the actions they influence. They are functionally related to behavior and can thus be tested in therapy—as hypotheses about causation—using single-case experimental designs. The formulation is directed toward the precise identification of the antecedent, outcome, and symbolic conditions that control the problematic behavior or beliefs (see Figure 4).

A distinction is made between the direct and vicarious learning experiences that contribute to the acquisition of problematic behavior and those contemporary influences which determine whether the client will perform the behavior he or she has acquired. For example, in the conduct disorders, some antisocial action which has been acquired by imitation or modelling may not be performed, either because appropriate instigating conditions do not occur or because its consequences are likely to be unrewarding or unpleasant. Maladaptive actions that are reinforced in terms of favorable consequences tend to be repeated. The reinforcement which strengthens an adolescent's disruptive behaviour—to take an increasingly typical example from the contemporary classroom—may be in terms of direct external reinforcement, vicarious or observed reinforcement, or self-reinforcement.

5.02.4.6.2 Distal consequences (Figure 4)

While it is true to say that clinicians delineate certain "symptoms" as pathognomonic of particular psychiatric disorders, the fact is that in the area of disorders of childhood it is mainly social rather than medical criteria of what is problematic that are dealt with. It is important to look not only at the immediate consequences of a client's behavior, but also at the longer-term implications (distal outcomes). What are the likely consequences of nonintervention in the problem for the person and his family? Problem behaviors are usually so called because they have a variety of unfavorable short-term and long-term outcomes. They are therefore referred to as maladaptive actions or dysfunctional thoughts and feelings; they are inappropriate in terms of several criteria which are assessed by the therapist.

Ultimately, the professional judgment of a client's behavioral/psychosocial/mental status is made in individual terms, taking into account his or her particular circumstances. It involves an estimate of the consequences that flow from the client's specific thoughts, feelings, and behaviors and general approach to life, with particular reference to their personal and emotional well-being, their ability to form and maintain social relationships, their ability to work effectively and (in the case of children) learn academically, and their accessibility to socialization (Figure 4, category 7). All are subject to disruption in emotional and behavioral disorders, and are gravely affected in the conduct disorders of childhood and adolescence (Herbert, 1987a). Other factors to be considered are youngsters' self-esteem and competence.

Social psychology can contribute to the overall assessment (Durkin, 1995). Take, for example, the socially isolated child: immature, self-centered children are not always able to manage the give-and-take of friendship. Exchange theory gives us pointers to why this should be so; it provides one method of evaluating friendships. Those theorists who employ "exchange theory" in analyzing aspects of social behavior and friendship view social interactions as a social exchange somewhat analogous to economic exchange. People are influenced (consciously and unconsciously) by the ratio of "rewards" to "costs" incurred in the interaction.

In the assessment of the implications for social relationships, it might be useful to draw up a balance sheet. On the debit side, the term "cost" is applied to deterrents that may be incurred in contacts with another person—such as hostility, anxiety, embarrassment, and the like. For attraction to a potential friend to occur the reward–cost outcome must be above the "comparison level," a standard against which satisfaction is judged.

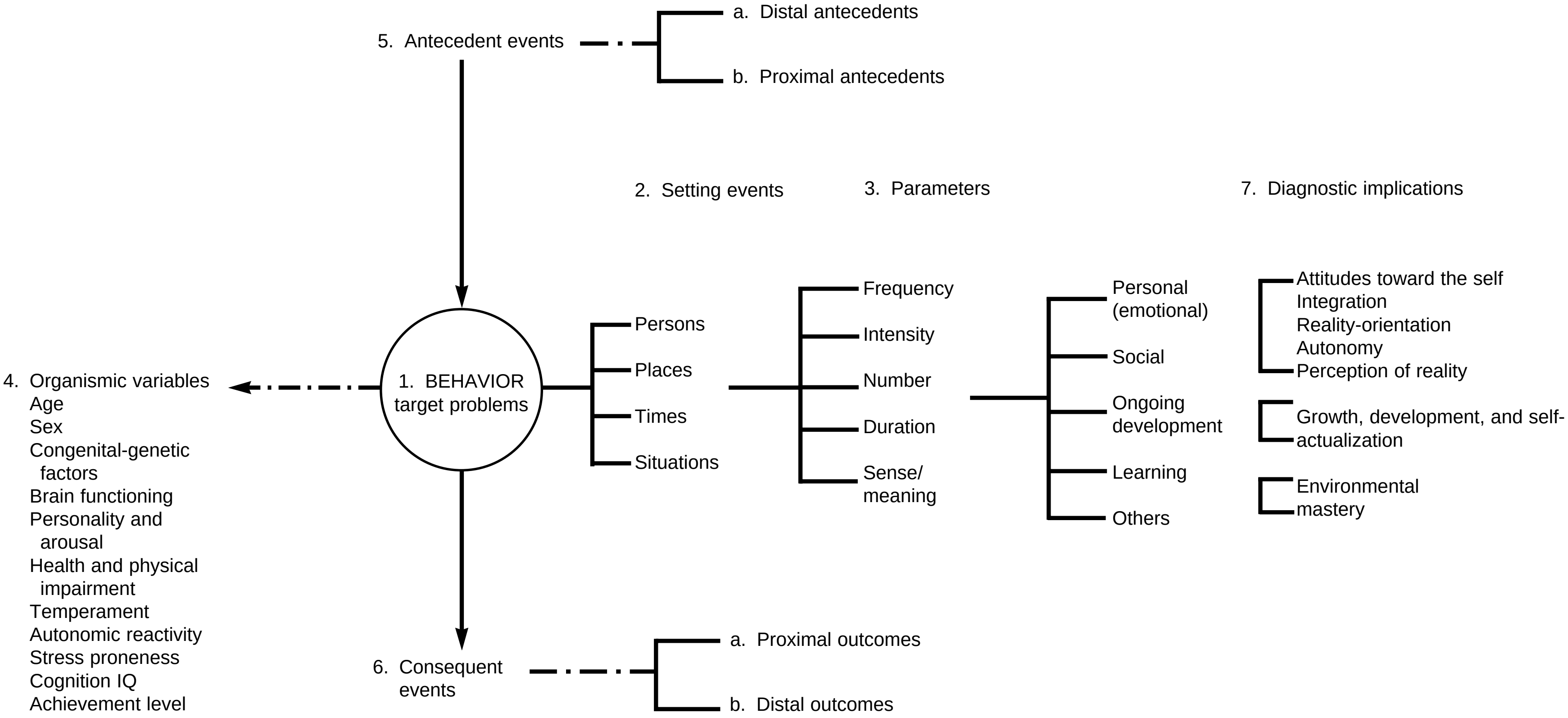
5. Antecedent events
a. Distal antecedents
b. Proximal antecedents
2. Setting events
3. Parameters
7. Diagnostic implications
1. BEHAVIOR target problems
Persons
Places
Times
Situations
Frequency
Intensity
Number
Duration
Sense/ meaning
Personal (emotional)
Social
Ongoing development
Learning
Others
Attitudes toward the self
Integration
Reality-orientation
Autonomy
Perception of reality
Growth, development, and self-actualization
Environmental mastery
4. Organismic variables
Age
Sex
Congenital-genetic factors
Brain functioning
Personality and arousal
Health and physical impairment
Temperament
Autonomic reactivity
Stress proneness
Cognition IQ
Achievement level
6. Consequent events
a. Proximal outcomes
b. Distal outcomes

Figure 4 A conceptual framework for an assessment of behavior problems. This figure represents the author's conceptual map of the essential data to be obtained for a clinical formulation. It takes account of the client, significant others (e.g., interactions), behaviors and beliefs, current and historical events. It also provides diagnostic criteria by examining the longer-term implications of these behaviors and beliefs.

Organismic variables include individual differences produced by age, sex, genetic constitution, physiology, and by cognitive representations (schema, attributions, etc.) based on past learning (Figure 4, category 4). Behavior therapists tend to adopt a transactional position—the view that behavior results from a (still poorly understood) interaction of the current situation and individual differences (biological and psychosocial).

5.02.4.6.3 The formal assessment

Assessment in the behavioral mode involves a recursive process both in repeatedly testing and in revising the model of the person's behavior. It is also repeatedly used to monitor changes in the person's behavior across time and settings. Repeated evaluation is also essential to evaluate and modify treatment. The assessment procedure is idiographic and tailored very carefully to the person's own unique configuration of strengths and weaknesses. Behavior therapists insist on carefully monitored pretreatment assessments. They believe that the client—which may include adults (e.g., parents) as well as children—is best described and understood by determining what he or she thinks, feels, and does in particular life situations. The goal is eventually to help the clients to control their own behavior and achieve self-selected goals. It is crucial in behavioral assessment and treatment to reformulate these complaints and goals in operational terms, that is, in overt terms of what the client says and does, and in a manner that lends itself to quantifying the problem (Ollendick & Hersen, 1984).

Wilson and Evans (1983) studied the reliability of target behavior selection in behavioral assessment by asking 188 members of the Association for Advancement of Behavior Therapy to assess three written case descriptions of (i) fearfulness, (ii) conduct disorders, and (iii) social withdrawal in children. Subjects were posed four open-ended questions: (i) whether treatment was warranted; (ii) to describe the child's major difficulty; (iii) state treatment goals; and (iv) specify and rank order treatment targets. Granted that the means of assessment is somewhat artificial there was nevertheless a surprisingly low agreement (38%) between clinicians in selecting a first-priority behavior for treatment. There was considerable variability in selecting behaviors for intervention and, perhaps surprisingly, a substantial (22%) tendency to introduce psychodynamic and intrapsychic terminology such as "internalized hostility," "poor self-concept," and "insecure child." Only approximately 20% of responses referred to specific target behaviors. Psychodynamic explanations were more likely in complex problems. This study is a matter of concern for advocates of functional analysis if its results are cross-validated. It indicates poor agreement between raters on the apparently simple task of identifying target behaviors.

Two questions regarding the validity of functional analysis are especially important: "What constitutes an adequate functional analysis?" and "When are the clinician's judgments valid during functional analysis?" Oliver and Head (1993, p. 14) reviewed published criteria for an adequate functional analysis and suggest that three criteria must be met: "that behavior (B) is considered in terms of the influence of events preceding the behavior, or its antecedents (A), and the events consequent (C) upon it...That the influences of antecedents and consequences should be empirically demonstrated...That all necessary and sufficient conditions for a behavior to occur should be considered."

Contemporary behavior analysts use and integrate many different assessment methods, including interviews with the client and third parties, psychometric assessments, idiographic self-reports, psychophysiological measures, and mini-experiments (Dougher, 1994; Sturmey, 1996). The purpose is to generate and evaluate hypotheses concerning the clinical problem. Information is collected which requires careful interpretation. At times this may be a highly inferential process. Although the functions of target behaviors may be readily apparent in some situations, this is not always the case.

Barrios (1988) notes that assessment, including behavioral assessments, may serve many different purposes. Behavioral assessment has been used in order to assist in diagnosis to evaluate treatment efficacy, to evaluate the most appropriate form of intervention for an individual client; also to assess whether clients are capable of change (e.g., in their parenting practices (Herbert, 1991b)).

5.02.4.7 Developmental-behavioral Formulations

"Development" is typically defined as a progressive series of orderly, coherent changes leading to the goal of maturity. This may be the grand design, but as many long-suffering parents and teachers know to their cost, children's progress through life is often disorderly and incoherent, and the changes (when change is not being resisted) are not always in the direction of maturity. Developmental frameworks for assessing individuals and families emerged from research and theory relating to their progress through life-cycle stages, and the therapeutic strategies they generate (e.g., Carter & McGoldrick, 1988). Several theorists (e.g. Achenbach, 1974, 1982; Cicchetti, 1984a, 1984b; Herbert, 1974) have urged the necessity for underpinning the study of child psychopathology with principles and findings from developmental psychology—an approach which has been referred to as the "developmental psychopathology movement" (Gelfand, Jensen, & Drew, 1988). This developmental framework has been applied to clinical practice in the form of a broad-based, interdisciplinary knowledge-base, incorporating works as diverse as that of ethologists, developmental neurologists, cognitive anthropologists, cognitive-development theorists, and social-learning theorists.

5.02.4.7.1 A developmental knowledge base

The difficulty for the clinician formulating his/her assessment within a developmental context is that there is no agreed, conceptually coherent theory of normal child development as a touchstone for understanding and assessing clinical abnormality, nor any single explanation of the processes of normal growth and development. Developmental theories range from the psychoanalytic (e.g., Freud's stages of psychosexual development), the ethological (e.g., Bowlby's attachment theory), behavioral (e.g., Bandura's social-learning theory) to the structural-developmental (e.g., Kohlberg's theory of moral development). Each of these theoretical perspectives provides a particular methodology and conceptual framework for viewing the specifics of developmental processes. Conceptions of individual characteristics also vary from those postulating stability to those assuming constant change. For example, trait theorists investigate nomothetically, relatively fixed relationships between characteristics (e.g., temperament, intelligence) across various ages, while those favoring an idiographic approach examine the individual's unique and dynamically changing organization of personality attributes.

A contemporary assessment framework leading to a formulation and treatment plan for childhood psychological disorders (see Herbert, 1987b; Ollendick & King, 1991) requires measures of behavioral, cognitive, and physiological responding as well as, notably, a determination of the developmental, social, and cultural context within which the problem occurs. The strategy is to begin with a broad-based assessment of the child and his or her environment (e.g., family, school, peers) and then to obtain information regarding specific stimulus features, response modes, antecedents, and consequences, severity, duration, and pervasiveness of the particular problems. The assessment utilizes a multimethod, problem-solving approach to obtain as complete a picture of the child and the family as is possible, one that generates in logical fashion, a treatment plan. The design of such treatment plans requires a multidimensional model in order to accommodate conceptually the different social contexts (home, school, and neighborhood) and the various levels of intervention that characterize child behavior therapy.

There are several key developmental questions to ask in formulating whether a problem merits serious concern:

(i) Do the child's problems interfere with the forward momentum and course of normal development?

(ii) Are they causing him/her to "regress," or slowing him/her down?

(iii) Do they represent a "cul-de-sac," taking him/her in the direction of a self-defeating lifestyle?

(iv) Is the child's adaptive behavior appropriate to his/her age, intelligence, and social situation?

(v) Is the environment making reasonable demands of the child?

(vi) Is the environment satisfying the crucial needs of the child, that is, the needs that are vital at his/her particular stage of development?

The family (or some other agent of socialization) may be dysfunctional. Thus, if the child's behavior is abnormal (unusual), troublesome, and self-defeating—and the answers to the last two questions are "No"—then you are still faced with a problem. However, it is more of a "problem situation" than an intrinsic disorder, or, more accurately, a systemic, interactional problem.

For Cicchetti and his colleagues (Cicchetti, Toth, & Bush, 1983) abnormal patterns of behavioral adaptation—as represented by clinical problems such as hyperactivity and failure-to-thrive—are most fruitfully construed and formulated as a series of stage—and age-related "tasks." The issue of "competence" is common to all of them and, as the child gets older, his or her self-esteem. If the child fails to develop skills and social competence, he or she is likely to suffer a sense of inadequacy which has spiralling ramifications. There is evidence, not only of the power of parents to facilitate the child's mastery of developmental tasks or hinder him or her, but to do this unwittingly (Cicchetti, Toth, & Bush, 1983). These matters become important themes in the final formulation.

The evaluation of normality and abnormality within a developmental framework requires a familiarity with general principles of development with particular reference to personality and behavior. It necessitates a comprehensive knowledge of children, how they think and talk, their skills and limitations at various ages, their typical repertoires of behavior, and the life-tasks and crises that they confront. Such a normative or comparative approach needs to be complemented by an idiographic assessment which takes into account biographical and social-contextual influences (Whiting, 1963).

The value of a developmental framework for clinical work does not cease after adolescence. We do not stop developing as adults. It is shortsighted to try to understand, say, an adolescent's misdemeanors without considering the preoccupations and attitudes characteristic of his or her parents' stage of life.

5.02.4.8 Social-learning Model—Cognitive-behavioral Approach

Behavior does not occur in a social vacum, nor is its acquisition a passive process. Human beings are problem solvers and active explorers of their world. Cognitive-behavior therapists (e.g., Hollon & Beck, 1994) subscribe to a model of human behavior and cognition which goes beyond the earlier epistemological position of radical behaviorists and focuses on the active processing of expectations, hypotheses, and theories. These have been added to that other "ghost in the machine"—the self—with which most behavior therapists now happily (and others, uneasily) cohabit.

An important feature of the social learning model, within which cognitive-behavior therapy finds much of its theoretical rationale, is the acknowledgment of this active element in learning and the part played by *understanding*. Not all of a child's learning (e.g., to ride a bike) involves understanding, but much learning involves knowing rather than simply doing. Stimuli, notably social stimuli, have meaning for the child, who acquires knowledge of social situations which is used in adapting to them. The failure to comprehend these meanings—as seen, for example, in the child with a Semantic-Pragmatic Disorder—has a devastating effect.

The cognitive-behavior therapy approach does not lend itself to any one simple definition (see Fishman, Rodgers, & Franks, 1988; Guidano & Liotti, 1983). Its focus, however, on goal setting and self-directive behavior by the client, and on divergent thinking as a means to inventive solutions to life's problems, is congruent with the philosophy of much contemporary casework with children and their families. Therapists assess, as primary data, children's problematic behavior, but also the phenomenology of their problems: their verbal reports of internal representations of events, experiences, and opinions, such as attitudes towards self, parents, and school. These attitudes, like theories, have, *inter alia*, a knowledge function. They provide a frame of reference within which individuals make interpretations of their world, creating a model—a figurative set of "goggles"—through which they construe and make sense of people and life-events. This personal construct system (Kelly, 1955) is accessible to investigation by means of the Kelly Repertory Grid.

The constructs of parents, some developed in the bosom of their family of origin, others elaborated with more recent experiences of parenting, give meanings to their offspring's behavior and can facilitate mutually-satisfactory interactions. Sadly, some constructs, if distorted by ignorance, misunderstandings, or paranoia, can mislead parents into self-defeating confrontations with their children.

There is a wide range of cognitive-behavioral methods; they share a common assumption that children (and, indeed, their parents) can be helped to eliminate dysfunctional attitudes and behaviors and to learn more adaptive strategies and skills (e.g., self-control, effective communication, negotiation), by reinforcing prosocial and socially sensitive behavior, challenging irrational beliefs and faulty logic, and by self-instruction and positive self-statements as well as the more traditional behavioral change techniques (Herbert, 1987b, 1989, 1993, 1997b; Kendall & Braswell, 1985).

5.02.4.9 Integrative Models

Given that there is no grand, general psychological theory of human behavior on offer in academic psychology departments, but rather a number of poorly integrated middle-range theories (e.g., learning theory, cognitive science, etc.), it should not be surprising that clinical psychologists find few convincing overarching theories or integrative models to make sense of psychopathology in childhood or adulthood.

The Kendall and Lochman (1994, p. 844) description of the cognitive-behavioral therapies is in terms that suggest there is a broader convergence of concepts in their rationale than in most of the other approaches. They state that:

> ...they integrate cognitive, behavioral, affective, social and contextual strategies for change. The cognitive-behavioral model includes the relationships of cognition and behavior to the affective state of the organism and the functioning of the organism in the larger social context.

A rationale for many practitioners is that all therapies contain an element of *learning*, something about which psychologists have considerable knowledge and experience of workable applications. Goldstein, Heller, and Sechrest (1966, p. 63) propose the following definition of psychotherapy:

> Whatever else it is, psychotherapy must be considered a learning enterprise. We need not specify too narrowly just what is to be learned in psychotherapy; it may be specific behaviors or a whole new outlook on life, but it cannot be denied that the intended outcome of psychotherapy is a change in an individual that can only be termed a manifestation of learning.

In the author's opinion, what gives cognitive-behavioral therapy its present ethos of an informed rather than ragbag eclectism, is its deep roots in empirical psychology, and a track record of encouragingly successful outcomes in a wide range of childhood behavior problems (Herbert, 1997a, 1997b; Kendall & Braswell, 1985; Kendall & Lochman, 1994).

5.02.5 THE CLINICAL FORMULATION—PRACTICE ISSUES

Rigorous assessment is the *sine qua non* of effective intervention. The formulation, which bridges these activities, depends upon the clinician's practice wisdom and skill in "mining" for the data with which it is constructed. The different stages of the ASPIRE process leading towards a formulation (described at the beginning of this chapter), involve a sometimes daunting array of practical choices, decisions, actions, and caveats. Not infrequently, incoming data—sometimes as far forward in client contact time as the intervention—demands a reformulation and possibly a "rethink" of the therapeutic program.

As any choice of key issues to discuss for their practice implications is likely to seems somewhat arbitary a few areas found to be important for clinical practice have been selected (see Tables 1 and 2).

5.02.6 CONDUCTING A FORMULATION

5.02.6.1 The Intake Interviews—Preparation

The interview is an essential method for gaining access to parents' and children's problems. It sounds deceptively simple and straightforward but requires careful preparation. The following questions illustrate some of the issues:

(i) What do I wish to find out (what are my objectives in this interview?)

(ii) Who do I need to speak to in order to fulfill my objectives? (All the family members; parents only; the child alone; his/her brothers and sisters?)

(iii) Do I invite them to meet me as a full family group? Do I speak to the mother first? Both parents together? Should the child be present initially?

(iv) How do I begin the interview?

(v) How do I introduce myself?

(vi) How do I let them know what to expect?

(vii) How do I best express some quite complex and potentially threatening ideas?

(viii) How do I reassure them about confidentiality?

(ix) What is the best way of eliciting reliable and relevant information?

(x) How do I deal with their tendency to digress or to set an agenda which avoids key issues?

(xi) Indeed, what *is* relevant (salient) information?

(xii) How do I terminate the interview, without leaving clients feeling "up in the air" or threatened?

5.02.6.2 Engaging the Clients

There is an important general consideration: the perspective, "style," or model of help to be adopted. In the expert model the professionals view themselves as very much in charge because of their monopoly (or near-monopoly) of expertise, responsibility, and, therefore, decision-making. The client is relatively passive as a recipient of advice, "prescriptions" (about health or how to behave), or possibly therapy of one kind or another. At the other extreme, clients (say parents) are viewed very much as consumers of the professionals' services, with the right to select what *they* believe is most appropriate to their needs. Decision-making is ultimately in the parents' control, with professionals acting as consultants, negotiation and discussion playing a large part in the client–practitioner relationship. The formulation is very much a self-directed, demystified process.

In between, lies the collaborative or partnership model. Here the professionals perceive themselves as having expertise but sharing it and imparting it to parents and other non- or paraprofessionals so that *they* facilitate much of the training/therapy of the child.

Cunningham and Davis (1985), in the context of work with mentally handicapped clients, devised a list of questions for estimating where one stands on the expert–consumer continuum vis-à-vis clients:

(i) Have I met the family?

(ii) Do I consider the child in the context of his/her family?

(iii) Do I have regular, two-way communication with the family?

(iv) Do I respect and value the child as a person?

(v) Do I respect and value the family?

(vi) Do I feel the family has strengths to help the child?

(vii) Have I identified the parents' abilities and resources?

(viii) Do I always act as honestly as possible?

(ix) Do I give them choices about what to do?

(x) Do I listen to them?

(xi) Have I identified their aims?

(xii) Do I negotiate with them?

(xiii) Do I adjust according to the joint conclusions?

(xiv) Do I assume they have some responsibility for what I do for their child?

(xv) Do I assume I have to earn their respect?

(xvi) Do I make the assumption that we might disagree about what is important?

(xvii) Do I believe they can change?

(xviii) Have I tried to identify the parents' perceptions of their child?

(A high proportion of "yes" responses suggests the generalized use of a consumer model; many "no" responses indicates an expert model).

5.02.6.3 The First Contact

The clinician conducts a guided conversation in which they ask questions and listen empathetically (with the "third ear" as it has been called) to the answers and to spontaneous comments. A collaborative style comes into effect from the very first encounter if this is the approach of choice. The practitioner, after explaining their role and what lies ahead for the client(s), attempts to engage willingness to cooperate and possibly change, by entering the client's concerns, experience, and feelings. The construction of a family tree (genogram) is a useful means—especially with children—of establishing rapport and getting to know who significant members of the family and friends are (Herbert, 1991b).

Parents are invited to explain their coping and disciplinary strategies, as well as their theories regarding possible causes of the child's problems which they have had the opportunity to describe. Ideally, they come to feel that the clinician is making a genuine effort to understand their internal reality. Other matters for discussion are the child's assets, and the parents', and, if present, the child's hopes and goals for therapy (see Herbert, 1987a, 1987b; Webster-Stratton & Herbert, 1994, for detailed accounts).

5.02.6.4 The "What?" Question

A failure to define the client's problem in precise terms, or to bring the circumstances in which it occurs into sharp focus, leads to the premature formulation of cause–effect relationships. To discuss causal factors before the "to-be-explained" phenomenon is adequately specified and measured is a reversal of scientific problem-solving.

Table 1 Beginning an assessment.

Initial screening	*Phase 1*
Step 1	Explain how you intend to work. Foster a good working relationship
Step 2	Obtain a general statement of the problem(s)
Step 3	Specify the problem(s) more precisely
Step 4	Elicit the desired outcome(s).
Step 5	Construct a problem profile (Herbert, 1987b)
Step 6	Introduce the idea of behavior sequences by looking at your client(s) typical day(s)
Step 7	Establish problem priorities
Further assessment	*Phase II*
Step 8	Work out a family life map (Herbert, 1988)
Step 9	Make an estimate of your client's assets
Step 10	Ask your adult clients about their goals (e.g., for their children). Ask the child about his or her wishes and requirements
	Phase III
	Collect more detailed "baseline" data. Give clients homework tasks, for example, keeping diaries, frequency charts, etc ...

Table 2 From planning to implementation.

Planning	*Phase 1*
Step 1	Consider informal and/or direct solutions
Step 2	Accentuate the positive
Step 3	Consider the wishes of those around the client
Step 4a	Begin with problems where success is likely
Step 4b	Select relevant areas of change
Step 5	Choose your intervention
Step 6	Monitor your client's progress
Implementation	*Phase II*
Step 7	Implement the intervention
Step 8	Review your intervention
Step 9	Gradually fade out the program
Follow-up	*Phase III*
Step 10	Conduct a follow-up exercise

This is not the place to consider the many aids to description and measurement for formulation purposes, that are available. (Useful guides to the psychometric and evaluative literature are available in Barkley, 1990, 1994; Berger, 1996; Ollendick & Hersen, 1984; Ollendick & King, 1991; Scott, 1996; Webster-Stratton & Herbert, 1994).

There are several "what" questions which, with others, generate precise hypotheses to explain the child's difficulties:

(i) What is the child doing and saying that is problematic?

(ii) Under what conditions are these behaviors emitted?

(iii) What tasks/problems is the child confronting in his/her life-situations?

(iv) What pro-social behaviors does the child emit?

(v) Under what conditions do these behaviors occur?

(vi) What situations are being avoided?

(vii) What are the implications for the child (and significant others) if s/he does not change?

(viii) What actions need to be encouraged and shaped up?

The preliminary information about the referral comes usually from parents. They tend to report their children's problems in terms of rather vague, global labels such as "aggressiveness," "tantrums," "defiance," or "rebelliousness." It is vital to encourage them to give descriptive examples of the problem, in other words to define what they mean in specific and observable terms. There follow two examples of a parent's complaint. The first is uninformative; the second provides a firm foundation for further assessment.

(i) Peter was playing with his cars when I asked him to get ready for school. Peter took no notice. I snatched away his toys and there was a scene. As a result, Peter had one of his tantrums and gave me a lot of abuse.

(ii) Peter was playing with his cars. I told him it was already 7.30 and time to get washed, dressed, and ready to leave for school. Peter did not answer or make a move. I told him the time again and repeated my instructions. Peter asked for a few minutes play. I said he was already late and told him to put the cars away. He now said "No!" just like that. I snatched his cars away. He lay down, screamed, and kicked his legs on the floor for a few minutes until I smacked him. He then went to change, muttering that he hated me and would leave home. I took no notice. His father took no interest in what went on; he ignored the entire thing.

After an initial report from parents (or teachers), the clinician needs to observe the child. It is important to confirm the parents' observations and especially those based on hearsay, for example, what she has heard about her child's behavior at school. A major source of error in the assessment and treatment of children in clinics or residential settings is the tendency to ascribe to the child properties which overemphasize the generality and invariance of personality traits and problem behavior. Such reification undermines effective formulating.

As an example, the problem could be a boy who bullies other children. The tendency is to make an entity of the bullying. It is done by describing the problem in the form of a noun or adjective ("there is a lot of violence in him"; "he is an aggressive boy"). The youngster is assessed in terms of what he is or has rather than in terms of what he does in particular situations. In the next stage the child is quite likely to be inferred to have "a need for aggression." From this point on, the distinction between what is observed and what is inferred, what is direct evidence and what is hearsay interpretation tends to become blurred. For all the popularity of trait theories, practitioners, as we have seen, make little use in therapy of the diagnostic information from personality typologies.

Behaviorally oriented psychologists tend to view behaviors with, say, aggressive or anxiety attributes as instances or samples of response classes rather than as the outward and visible signs of internal or underlying dispositions. They tend to avoid trait labels. Diverse attributes would be considered to be members of the same response class because it could be shown that such attributes enter into the same functional relationships with antecedent, concurrent, and consequent stimulus conditions, rather than because they coexist or covary in a group of persons.

5.02.6.5 Unpacking the Problem

In the case of complex problems it is helpful to have (or to construct) a conceptual framework—a guide to the factors that require investigation. Figure 5 presents an example of how, with a problem like school refusal, the clinician can unpack or unravel the processes underlying the problematic situation into their constituent elements, as an aid to the assessment and formulation. Thus with the school refusal problem the clinician might ask: "What does a young person have to do in order to go to school (successfully), remain at school, and then return home at the appropriate time?" "What incentives ('push–pull') or deterrent influences are operating in this child's situation?" The flow chart (Figure 5) illustrates the "push" and "pull" factors which function to impel or attract children to school, and hold them there, and the stages at which things can go wrong.

5.02.6.6 Communication

Young children tend to be talkative but may be limited in their ability to reflect insightfully about their experiences; adolescents are often introspective (reflective) but have a way of becoming monosyllabic when asked personal questions. This poses a problem for the would-be interviewer.

Children are not always very good at expressing their fears, frustrations, or uncertainties. They cannot always tell their parents, let alone a comparative stranger, how they feel, but they have a language that adults can learn to translate—the language of behavior and fantasy. What they do (in a direct sense in everyday life) and say (indirectly through play or storytelling) can be most revealing. The advantage of using projective techniques (which include play, puppets, dramatic creations, completing stories or sentences) for assessment, is that they involve relatively unstructured tasks that permit an almost unlimited variety of responses. The client has to fall back on his or her own resources rather than stereotyped, socially desirable answers. The techniques (as psychometric instruments) do have their critics, but are invaluable if used cautiously as aids to communicating with children (Babiker & Herbert, 1996).

Where the child is too loyal, too frightened or ashamed, or too inarticulate to speak about feelings (or painful events in the family) it may be possible to express these emotions in play, drawing, drama (with puppets or miniatures), or stories. They are undoubtedly an invaluable adjunct to work with children and involve the empathic imagination and creativity referred to at the beginning of this chapter (see Herbert, 1988).

5.02.6.7 The "Which?" Question

An issue which arises early in a clinical assessment involves prioritizing problems. Clinicians with mandatory obligations have their own guidelines to follow. The practitioner's and parents' hierarchy of tasks might be influenced by the following considerations with regard to the problem(s):

(i) their annoyance value;

(ii) their actual/potential dangerousness;

(iii) their interference with the life of the family or its individual members;

(iv) their accessibility to change (improvement) and an intervention;

(v) their frequency, intensity, and magnitude;

(vi) their disabling implications;

(vii) the "costs" of change in terms of resources (time, money, etc.) and other people's well-being;

(viii) the ethical acceptability of the desired outcome;

(ix) the availability of the necessary helping skills or resources.

5.02.7 THE FORMULATION

The formal or main phase of assessment can be said to be over when the clinician has a reasonably clear picture of the clients' difficulties, strengths, history, social setting, and background. They should have a reasonably clear picture of why the problems developed and how they "look" against a framework of normal development. Relevant intrinsic attributes of the child have been assessed. The problem is

Has to wake up on time → Get bathed, dressed, breakfasted → Leave the house on time → Make a journey → Enter the school gates → Go to assembly → Stay in the school → Go home at appropriate time →

(1) Does he wake on time? If not, why not?
(2) Does he get enough sleep? If not, why not (going to bed late; lying in bed unable to sleep because of morbid preoccupations, tense, depressed?)

(3) Anyone to structure his day at home, e.g. supervise his getting ready (“push”)?
(4) Is he sick/anxious/panicky?
(5) Any reason he needs to be at home (care for a sick member of the family, parents keep him at home to look after siblings, etc.)?
(6) Is he afraid to leave home because concerned about his mother’s health, afraid of an accident befalling her (preoccupation with death, separation anxiety)?
(7) Is he depressed, overwhelmed by apathy, helplessness, inertia?

(8) Is he teased/bullied on the way to, or at school?
(9) Claustrophobia/clothing (adequate for school?)/homework
(10) Is there anything to keep him at school (interests, friends, teacher)?
(11) Deviant models (peer group) for truanting
(12) Other “pull” factors absent? (Is he under-achieving grossly at school, bored?)
(13) Does anyone really know him or take an interest in him at school?

Figure 5 “Push–pull” influences in a child’s school attendance.

formulated usually in systemic terms, that is to say as it impinges on a family or school system rather than on an individual alone.

5.02.7.1 The “Why?” Question

The formulation provides a summary of the salient information collected; it is “formulated” in the sense that the practitioner puts forward an “explanatory story” (hopefully, a valid one) to impose meaning on the data collected. The formulation consists of a series of hypotheses which are statements about how the problem arose, when and where it occurs, and why it is maintained.

A particular skill is one of editing. It involves the identification of the important influences and controlling variables—among, potentially, so many—in order to plan an effective intervention. The principle of parsimony is helpful here; one invokes those causal factors that are sufficient to facilitate such an intervention. There is a balance to be struck between the wastefulness and risks of overinclusiveness (that is to say excessive complexity), on the one hand, and simplistic theorizing on the other.

Bromley (1986) describes the required steps for explication of the individual case.

(i) State clearly the problems and issues.

(ii) Collect background information as a context for understanding (i).

(iii) Put forward *prima facie* explanations (conjectures/hypotheses) and solutions (program formulation) with regard to the client’s personality and predicament—on the basis of information available at the time, and on the basis of the principle of parsimony. Examine the simple and obvious answers first. They may, of course, have to be rejected if they don’t stand up to critical examination. This guides (iv).

(iv) Search for further/additional evidence. New hypotheses/explanations will have to be formulated and examined.

(v) Search again and admit for consideration, sufficient evidence to eliminate as many of the suggested explanations (hypotheses) as possible; the hope is that one of them will be so close to reality as to account for all the evidence and be contradicted by none of it. The evidence may be direct or indirect; but it is vital that it should be admissable, relevant, and obtained from competent and credible sources.

(vi) Enquire critically into the sources of evidence, as well as the evidence itself. Bromley (1986) makes the point that in the case of personal testimony, this is analogous to cross-examinations in a court of law; otherwise it amounts to checking the consistency and accuracy of all items of evidence.

(vii) Examine carefully the internal logic, coherence, and external validity of the entire network of associations and hypotheses formulated to explain the client’s predicament and proposals to solve the problems.

(viii) Select the “most likely” interpretation, provided it is compatible with the evidence (some lines of argument will be obviously inadequate whereas others will be possible or even convincing).

(ix) Work out the implications of your explanations for intervention/treatment or some other action (or, indeed, inaction).

Always ask yourself about the implications (e.g., the risks involved) of making Type I as opposed to Type II errors in your assessment. Is it more damaging to your client if you risk Type I errors (i.e., asserting relationships falsely) than if you risk Type II errors, which deny relationships that actually exist? The academic psychologist tends to minimize errors of incautious assertion at the expense of relatively common Type II errors. The clinician often acts on the basis of weakly supported propositions because of the dangers of ignoring potentially significant associations. But of course there may also be some risks in presuming relationships which do not have any basis in reality. This step requires that you

(x) Work out the implications in specific terms.

Therapeutic change depends, *inter alia*, upon persuasion (providing reasons why the client should change his/her attitudes and behavior); teaching (all therapies, it might be said, involve new learning); and the provision of explanations (reframing; cognitive restructuring).

Insight is not rejected as a facilitative therapeutic agent. Yelloly (1972, p. 147) states that:

> Awareness can operate in a number of ways. The sheer provision of accurate information may correct a false and erroneous belief and bring about considerable change in behavior. Prejudice for instance, may be diminished by new information which challenges the prejudiced belief. And in human beings (pre-eminently capable of rational and purposive action) comprehension of a situation, knowledge of cause and effect sequences and of one's own behavior and its consequences, may have a dramatic effect on manifest behavior. Thus to ignore the role of insight is just as mistaken as to restrict attention wholly to it.

5.02.7.2 Explaining to Parents

The attempt to "make sense" of complicated problems for (say) parents and teachers, depends upon the particular circumstances of the case, but also on the ability of the practitioners to translate technical matters into plain, unpatronizing language. Although behavior therapy is not primarily an insight-based therapy, undoubtedly the understanding of its rationale and how it applies to the individual child facilitates the process of change, as mentioned above (Herbert, 1987b).

A functional analysis can be a rich source of insights for the client. A connotation of such an analysis is that the behaviors under consideration serve a purpose, that is, are functional (dysfunctional though they may appear to the professional eye) for the individual. Clinical psychology is full of examples of this kind of functionalism; therapists commonly talk of the secondary gains arising from clinical problems. Their implications for the individual's well-being and effective functioning provide a diagnostic guideline for the therapist and an insightful view of causation for the client. The meaning of the problems for the child and indeed for his or her family—the sense made of them and the payoff they provide—constitute a vital element of the formulation conveyed to the clients.

In behavioral work, in the author's opinion, explanations in terms of a functional analysis can operate at two levels. At its simplest, behavior is a function of certain contingent stimuli, originating in the person's internal and external environment. Here, the important questions are: "what triggers (elicits) the phobia?" or "what reinforcement does the child get for behaving in this antisocial way?" At a more interpretive level, the child's behavior may have the function of solving (or attempting to solve) a developmental or life problem. To make sense of it, one might ask (*inter alia*): "what immediate "solutions" (even if self-defeating in the longer term) do the child's actions provide for himself or herself?" Also, "what purpose does the child's behavior serve in terms of his or her family life and its psychological and social dynamics?"

A clinical formulation which is being explained to parents might sound something like this:

> Andrew has not yet learned to control his temper. He "lets go" in a frightening tantrum—banging his head, kicking, screaming, and yelling—when he cannot get his own way, also when you try to insist on his doing something, refuse his commands, or attend to people other than himself at a time when he wants your undivided attention. The result (usually) is that he achieves his goal, he coerces you both into giving way—a very rewarding state of affairs seen from his point of view; a very unrewarding (and sometimes humiliating) state of affairs seen from your perspective. Occasionally you stick to your guns which means that Andrew has been getting rather inconsistent messages; the consequences of his undesirable actions are not predictable. Sometimes (but rarely) you punish his unacceptable behavior, generally when there are visitors. Sometimes you ignore it, for example, at the supermarket. Generally you "reward" it and thus make it more likely to recur, by giving in to him. Andrew has learned to make this outcome

more likely by escalating the tantrums into very violent and therefore frightening episodes. It now has to be made quite clear to him that his bad behavior will have consequences that are not only unrewarding, but also unpleasant enough to make him relinquish his tantrum. You have tried smacking which you admit makes him worse and you miserable. A method called response-cost will be used—in essence, fining his tantrums. At the same time it will be made very beneficial for him to be more obedient, to ask nicely and to control his temper. His successes will be recorded on this chart and show him how he can earn treats when he has collected a certain number of stickers.

5.02.8 THE "HOW?" QUESTION—IMPLEMENTING AN INTERVENTION

Client commitment and resistance to change are among the most critical features to assess and manage when implementing a successful intervention of whatever kind. The collaborative approach as elaborated by Webster-Stratton and Herbert (1994) is designed, by means (*inter alia*) of self-empowering strategies, the use of relationship, respectful attention to clients' feelings and ideas, and the use of metaphor and humor, to stimulate their interest and incentive to change, and to overcome resistances. Herbert and Wookey (1997) have been able to confirm the effectiveness of such process variables in a succession of parent groups.

The self-empowering strategies are designed to enhance self-esteem and perceived self-efficacy. Perceived self-efficacy is a significant item in any formulation of the likely success of such an approach (see Webster-Stratton & Herbert, 1994). Bandura (1977) is of the opinion that human behavior is subject to two major categories of influence: efficacy expectations (the conviction that one can successfully carry out the behavior required to produce a given outcome) and outcome expectations (the estimate that one's behavior will lead to certain outcomes).

These constituent parts of this crucial notion of perceived self-efficacy are distinguished because a mother or father may believe that a particular course of action (e.g., a behavioral program) will produce certain outcomes, that is, an improvement in their child's behavior. However, they may have serious misgivings as to whether they have the wherewithal (e.g., patience and consistency) to bring about such a desirable outcome. All psychological procedures designed to bring about change, whatever the type, are thought (by some psychologists) to depend on beliefs about the level of skill required to bring about an outcome and the likely end result of a course of action.

Efficacy expectations are thought to be the most important component. The main effect on outcome expectations is through the strengthening of efficacy expectations ("I *am* able to do it!"). Successful helping thus depends, in this view, on the degree to which the interventions create or strengthen the client's expectations of personal efficacy. It appears that verbal persuasion has only relatively weak and short-lived effects on such expectations. Performance accomplishments, on the other hand, have the most powerful effects, hence the concentration on "doing" (e.g., behavior rehearsal, modeling, role-play, homework tasks) in behavioral parenting training (Webster-Stratton, & Herbert, 1994).

5.02.8.1 The Collaborative Approach

Webster-Stratton and Herbert (1993) are of the opinion that the collaborative approach to clinical work by empowering clients is highly effective in producing favorable outcomes with the conduct disorders (see Herbert, 1995; Herbert & Wookey, 1997; Patterson & Forgatch, 1985; Spitzer, Webster-Stratton, & Hollinsworth, 1991). In a collaborative relationship:

> ... the therapist works with parents by actively soliciting their ideas and feelings, understanding their cultural context, and involving them in the therapeutic experience by inviting them to share their experiences, discuss their ideas, and engage in problem-solving. The therapist does not set him/herself up as the "expert" dispensing advice nor lectures to parents about how they should parent more effectively; rather, the therapist invites parents to help write the "script" for the intervention program. The therapist's role as collaborator, then, is to understand the parents' perspectives, to clarify issues, to summarize important ideas and themes raised by the parents.... (Webster-Stratton & Herbert, 1994, p. 108)

The partnership between clients and practitioner, it is suggested, has the effect of giving back dignity, respect, and self-control to parents who are often seeking help for their children's problems at a vulnerable time of low self-esteem, self-blame, guilt, and learned helplessness (Herbert & Wookey, 1997).

5.02.8.2 Rigorous Evaluation

Data collection begun in the baseline period continues throughout the intervention period. It is vital to keep a close watch on what happens in the early stages to ensure that the program is being followed correctly and any unforeseen problems that arise are dealt with quickly. This is the rigorous evaluation component of the ASPIRE process.

In the light of the data collected relating to the target behavior(s), or any other objective information, it may be necessary to reformulate one's previous conclusions and modify the intervention in some way. In the case of behavioral work the program is flexible and if the chosen target behavior is not changing after a reasonable period of time it becomes essential to review all aspects of the work.

The following questions may be useful.

(i) Are there powerful competing reinforcers or interfering factors in the client's environment operating against the program? Can these be modified?

(ii) Are the reinforcers well chosen and effective?

(iii) Is the parent working effectively and being reinforced for his or her participation?

(iv) Are the behavioral objectives realistic? Are they within the client's repertoire?

(v) Have you proceeded too quickly? Is it necessary to go back to an earlier stage in the intervention plan?

The time for the program to end depends upon the goals established at the beginning of treatment although new (or elaborated) goals may emerge during the intervention.

5.02.9 SUMMARY

This chapter begins by stating that the piecing together of a clinical formulation is central to the practitioner's craft. The choice of the word "craft" was deliberate. The craft of the potter is an apt metaphor for the many-faceted psychological skills applied to the task of analyzing and synthesizing, into a relatively brief statement—a formulation—about complex human problems. Potters are *applied scientists* when they draw upon an empirical knowledge base for the different clays and glazes they use, and the parameters for firing the products they make. But they become artists as well, when they design and decorate their creations, and when they bring to bear a delicacy, intricacy, and sureness of touch as they shape the clay.

Undoubtedly there is something called a "clinical art," especially in the ingenuity with which an assessment and intervention are put into effect. Nevertheless, in the opinion of this author, the major part of the clinical endeavor remains a scientific one. The balance of these emphases differs according to the basic assumptions of the theoretical model adopted and the clientele one works with. Certainly therapy with children demands a good deal of imagination and "artistry," especially if they are to remain well motivated.

The chapter demonstrates how difficulties and, indeed, failures in the treatment of children's behavior problems can frequently be traced back to a lack of precision during the crucial assessment phase of a clinical contact. A scientist-practitioner trying to explain a particular phenomenon will be at pains to bring that phenomenon into "sharp focus" by describing and measuring it in all its manifestations as objectively as possible. The analysis of a problematic situation and the planning of an intervention require as much knowledge and skill as any other aspect of the clinical psychologist's input, possibly more. The ASPIRE procedures outlined for conducting such an assessment follow broadly the tried and tested scientific method of problem-solving.

The scientist-practitioner's science, however, is not, and because of the exigencies of clinical work cannot be, of the kind represented by what might be called "scientism"—a pharisaical insistence on the letter rather than the spirit of scientific method. There has to be room for intuition, insight, and creativity as long as they are checked and validated whenever possible by rigorous evaluation. It has been said that what cannot be measured cannot be changed, a proposition that has much merit as long as measurement does not generate into a sterile obsession with minutiae. The case study is the bedrock of clinical work, and an invaluable scientific method of enquiry. It provides us with an exemplar of that unique blending of idiographic and nomothetic approaches that have their parallels in the artistic and scientific facets of clinical psychology.

5.02.10 REFERENCES

Achenbach, T. M. (1974). *Developmental psychopathology*. New York: Ronald Press.

Achenbach, T. M. (1982). *Developmental psychopathology* (2nd ed.). New York: Wiley.

Achenbach, T. M., & Edelbrock, C. S. (1989). Diagnostic, taxonomic and assessment issues. In T. Ollendick & M. Hersen (Eds.), *Handbook of child psychopathology* (2nd ed.). New York: Plenum.

Achenbach, T.M., McConaughty, S. H., & Howell, C.T. (1987). Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101,* 213–232.

Ackerman, N. (1958). *The psychodynamics of family life*. New York: Basic Books.

Alexander, J. F., & Parsons, B. V. (1982). *Functional family therapy*. Monterey, CA: Brooks/Cole.

Allport, G. W. (1937). *Personality: A psychological interpretation*. London: Constable.

American Psychiatric Association (1993). *Diagnostic and statistical manual of mental disorders* (4th ed.) (*DSM-IV*). Washington, DC: American Psychiatric Association.

Andrews, D. A. (1989). Recidivism is predictable and can be influenced: using risk assessments to reduce recidivism. *Forums on Corrections Research, 1,* 11–18.

Axline, V. M. (1947). Play therapy procedures and results. *American Journal of Orthopsychiatry, 25,* 618–627.

Babiker, G., & Herbert, M. (1996). The role of psychological instruments in the assessment of child sexual abuse. *Child Abuse Review, 5,* 239–251.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84,* 191-215.

Barkley, R. A. (1990). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment.* New York: Guilford Press.

Barkley, R. A. (1994). *Taking charge of ADHD.* New York: Guilford Press.

Barlow, D., & Hersen, M. (1984). *Single case experimental designs: Strategies for studying behavior change.* New York: Pergamon.

Barrios, B. A. (1988). On the changing nature of behavioral assessment. In A. S. Bellack & M. Hersen (Eds.), *Behavioral Assessment: A practical handbook* (3rd ed.). New York: Pergamon.

Bateson, G., Jackson, D. D., Haley, J., & Weakland, J. (1956). Toward a theory of schizophrenia. *Behavioral Science, 1,* 251–264.

Berger, M. (1996). *Outcomes and effectiveness in clinical psychology practice.* Division of Clinical Psychology Occasional Paper, No. 1. Leicester: British Psychological Society.

Bowen, M. (1966). Theory in the practice of psychotherapy. In P. J. Guerin (Ed.), *Family therapy: Theory and practice.* New York: Gardner.

Bromley, D. (1986). *The case-study method in psychology and related disciplines.* Chichester, UK: Wiley.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Cambridge, MA: Harvard University Press.

Carter, B., & McGoldrick, M. (1988). *The changing family life cycle: A framework for therapy* (2nd ed.) New York: Gardner.

Cicchetti, D. (1984a). The emergence of developmental psychopathology. *Child Development, 55,* 1–7.

Cicchetti, D. (1984b). *Developmental psychopathology.* Chicago: University of Chicago Press.

Cicchetti, D., Toth, S., & Bush, M. (1983). Developmental psychopathology and incompetence in childhood: suggestions for intervention. In B. B. Lahey & A. E. Kazdin (Eds.), *Advances in clinical child psychology* (Vol. 11). New York: Plenum.

Clarke, R. G. V. (1977). Psychology and crime. *Bulletin of the British Psychological Society, 30,* 280–283.

Cohen, L. H. (1977). Factors affecting the utilisation of mental health evaluation and research findings. *Professional Psychology, 8,* 526–534.

Cohen, L. H. (1979). The research relationship and information source reliance of clinical psychologists. *Professional Psychology, 10,* 78–85.

Cunningham, C., & Davis, H. (1985). *Working with parents: Frameworks for collaboration.* Milton Keynes, UK: Open University Press.

Dare, C. (1985). Family therapy. In M. Rutter & L. Hersov (Eds.), *Child and adolescent psychiatry* (2nd ed.). Oxford: Blackwell Scientific.

de Kemp, R. (1995). *Interactions in family therapy: A process research.* Nijmeger, The Netherlands: Nederlands.

Dougher, M. J. (1994). Introduction. *The Behavior Analyst, 17,* 261.

Durkin, K. (1995). *Developmental social psychology: From infancy to old age.* Cambridge, MA: Blackwell.

Erwin, E. (1979). *Behavior therapy: scientific philosophical and moral foundations.* Cambridge: Cambridge University Press.

Eysenck, H. J. (1967). *The biological basis of personality.* Springfield, IL: C. C. Thomas.

Farrell, B. A. (1970). Psychoanalysis: The method. In S. G. Lee & M. Herbert (Eds.), *Freud and psychology.* Harmondsworth, UK: Penguin.

Fielding, D. (1983). Adolescent services. In A. Liddell (Ed.), *The practice of clinical psychology in Great Britain.* Chichester: Wiley/Leicester: British Psychological Society.

Fisher, S., & Green, R. (1996). *Freud scientifically reappraised: Testing the theories and the therapy.* Chichester, UK: Wiley.

Fishman, D. B., Rodgers, F., & Franks, C. M. (1988). *Paradigms in behavior therapy: Present and Past.* New York: Springer.

Freud, A. (1946). *The psychoanalytic treatment of children.* New York: International Universities Press.

Freud, S. (1926). *Inhibitions, symptoms and anxiety. Standard Edition* (Vol. 21, pp. 87–172). London: Hogarth Press (1959).

Gambrill, E. (1983). *Casework: A competency based approach.* Englewood Cliffs, NJ: Prentice-Hall.

Gambrill, E. (1990). *Critical thinking in clinical practice.* San Francisco: Jossey-Bass.

Gelfand, D. M., Jensen, W. R., & Drew, C. J. (1988). *Understanding childhood behavior disorders.* New York: Holt, Rinehart, and Winston.

Goldstein, A. P., Heller, H., & Sechrest, L. B. (1966). *Psychotherapy and the psychology of behavior change.* New York: Wiley.

Griest, D. L., & Wells, K. C. (1983). Behavioral family therapy with conduct disorders in children. *Behavior Therapy, 14,* 37–53.

Group for the Advancement of Psychiatry (GAP) (1970). *The field of family therapy.* New York: GAP.

Guidano, V. F., & Liotti, G. (1983). *Cognitive processes and emotional disorders.* New York: Guilford Press.

Haynes, S. N., & O'Brien, W. H. (1990). Functional analysis in behavior therapy. *Clinical Psychology Review,10,* 649–668.

Hepworth, D., & Larsen, J. (1990). *Direct social practice: Theory and skills* (3rd ed.). Belmont, CA: Wadsworth.

Herbert, M. (1964). The concept and testing of brain-damage in children: A review. *Journal of Child Psychology and Psychiatry,5,* 197–216.

Herbert, M. (1965). Personality factors and bronchial asthma. *Journal of Psychosomatic Research,8,* 353–356.

Herbert, M. (1974). *Emotional problems of development in children.* London: Academic Press.

Herbert, M. (1987a). *Conduct disorders of childhood and adolescence: A social learning perspective* (2nd ed.). Chichester, UK: Wiley.

Herbert, M. (1987b). *Behavioral treatment of children with problems.* London: Academic Press.

Herbert, M. (1988). *Working with children and their families.* Leicester: British Psychological Society/Routledge.

Herbert M. (1989). Aggressive and violent children. In K. Howells & C. R. Hollin (Eds.), *Clinical approaches to violence.* Chichester, UK: Wiley.

Herbert, M. (1991a). *Clinical child psychology: Social learning, development and behavior.* Chichester, UK: Wiley.

Herbert, M. (1991b). *Child care and the family: Resource pack.* Windsor, UK: National Foundation of Educational Research—Nelson.

Herbert, M. (1993). Foreword. In V. Varma (Ed.), *How and why children hate.* London: Jessica Kingsley.

Herbert, M. (1994a). Etiological considerations. In T. H. Ollendick, N. J. King, & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in children and adolescents.* New York: Plenum.

Herbert, M. (1994b). Behavioral methods. In M. Rutter, E. Taylor, & L. Hersov (Eds.), *Child and adolescent* *psychiatry* (3rd ed.). Oxford: Blackwell Scientific.

Herbert, M. (1995). A collaborative model of training for parents of children with disruptive behavior disorders. *British Journal of Clinical Psychology,34,* 325–342.

Herbert M. (1997a) *Family treatment*. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of child psychology* (3rd ed.). New York: Plenum.

Herbert, M. (1997b) *Cognitive-behavior therapy of adolescents with conduct disorders*. In P. Graham (Ed.), in press.

Herbert, M. & Wookey, J. A. (1997). *Child-wise parenting skills manual*. Exeter, UK: Impact Publications.

Hollon, S. D., & Beck, A. T. (1994). Cognitive and cognitive behavioural therapies. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed.). New York: Wiley.

Iwaniec, D. (1995). *Emotional abuse and neglect*. Chichester, UK: Wiley.

Jackson, D. (1957). The question of family homeostasis. *Psychiatric Quarterly Supplement,31,* 79–80.

Jampala, V. C., Sierles, F. S., & Taylor, M. B. (1986). Consumers' views of DSM-III: Attitudes and practices of U.S. psychiatrists and 1984 graduating residents. *American Journal of Psychiatry,143,* 148–153.

Jones, M. C. (1924). The elimination of children's fear. *Journal of Counseling Psychology,7,* 383–390.

Kaslow, F. W. (1980). History of family therapy in the United States: A kaleidoscopic overview. *Marriage and Family Review,3,* 77–111.

Kazdin, A. E. (1987). Treatment of antisocial behaviour in children: current status and future directions. *Psychological Bulletin,102,* 187–203.

Kazdin, A. E. (1988). *Child psychotherapy: developing and identifying effective treatment*. New York: Pergamon.

Kelly, G. A. (1955). *The psychology of personal constructs*. New York: Norton.

Kendall, P. C., & Braswell, L. (1985). *Cognitive-behavioral therapy for impulsive children*. New York: Guilford Press.

Kendall, P. C., & Lochman, J. (1994). *Cognitive-behavioral therapies*. In M. Rutter, E. Taylor, & L. Hersov (Eds.), *Child and adolescent psychiatry* (3rd ed.). Oxford: Blackwell Scientific.

Klein, M. (1975). *The writings of Melanie Klein* (4 vols.). London: Hogarth Press.

Lask, B. (1987). Physical illness, the family and the setting. In A. Bentovim, G. Gorell-Barnes, & A. Cooklin (Eds.), *Family therapy*. London: Academic Press.

Lee, S. G., & Herbert, M. (1970). *Freud and psychology*. Harmondsworth, UK: Penguin.

Leung, F. L. (1975). The ethics and scope of behavior modification. *Bulletin of the British Psychological Society,28,* 376–379.

Long, C. G., & Hollin, C. R. (1997). The scientist-practitioner model in clinical psychology: a critique. *Clinical Psychology and Psychotherapy, 4,* 75–83.

Marshall, C., & Rossman, G. B. (1989). *Designing qualitative research*. Newbury Park, CA: Sage.

Maslow, A. H. (1954). *Motivation and personality*. New York: Harper & Row.

Mayer-Gross, W., Slater, L., & Roth, M. (1955). *Clinical psychiatry*. London: Cassell.

Meehl, P. B. (1971). A scientific, non-researched doctorate for clinical practitioners: Arguments pro and con. In R. R. Holt (Ed.), *New horizons for psychotherapy*. New York: International University Press.

Minuchin, S. (1974). *Families and family therapy*. Cambridge, MA: Harvard University Press.

Moore, C., Bobblitt, W., & Wildman, R. (1968). Psychiatric impressions of psychological reports. *Journal of Clinical Psychology, 24,* 373–376.

Ollendick, T. H., & Hersen, M. (Eds.) (1984). *Child behavioral assessment principles and procedures*. New York: Pergamon.

Ollendick, T. H., & King, N. J. (1991). Fears and phobias of childhood. In M. Herbert (Ed.), *Clinical child psychology: Social learning development and behaviour*. Chichester, UK: Wiley.

Oliver, C., & Head, D. (1993). Self-injurious behavior: Functional analyses and interventions. In R. S. P. Jones & C.B. Eyeres (Eds.), *Challenging behavior and intellectual disability: A psychological perspective*. Clevedon, UK: BILD Publications.

Patterson, G. (1982). *Coercive family process*. Eugene, OR: Castalia.

Patterson, G. R., & Forgatch, M. (1985). Therapist behavior as a determinant for client noncompliance: A paradox for the behavior modifier. *Journal of Consulting and Clinical Psychology, 53,* 846–851.

Peterson, D. R. (1976). Need for the doctor of psychology degree in professional psychology. *American Psychologist, 31,* 792–798.

Piercy, F. P., Sprenkle, D., & associates (1986). *Family therapy sourcebook*. New York: Guilford Press.

Pilgrim, D., & Treacher, A. (1992). *Clinical psychology observed*. London: Tavistock/Routledge.

Power, M. J. (1991). Cognitive science and behavioral psychotherapy: where behavior was there shall cognition be. *Behavioral Psychotherapy, 19,* 20–41.

Pumpian-Mindlin, E. (Ed.) (1952). *Psychoanalysis as science*. New York: Basic Books.

Quay, H. C. (1986). Classification. In H. C. Quay & J. S. Werry (Eds.), *Psychopathological disorders of childhood* (3rd ed.). New York: Wiley.

Rachman, S. (Ed.) (1963). *Critical essays on psychoanalysis*. Oxford: Pergamon.

Reimers, S., & Treacher, A. (1995). *Introducing user-friendly family therapy*. London: Routledge.

Rogers, C. R. (1957). *Client centered therapy*. Dallas, TX: Houghton Mifflin.

Ross, A. O. (1968). Conceptual issues in the evaluation of brain damage. In J. L. Khanna (Ed.), *Brain damage and mental retardation: a psychological evaluation*. Springfield, IL: C. C. Thomas.

Ross, R. R., & Gendreau, P. (1980). *Effective correctional treatment*. Toronto, Canada: Butterworth.

Rycroft, C. (1966). Causes and meaning. In C. Rycroft (Ed.), *Psychoanalysis observed*. London: Constable.

Satir, V. (1967). *Conjoint family therapy*. Palo Alto, CA: Science and Behavior Books.

Scott, S. (1996). Measuring oppositional and aggressive behaviour. *Child Psychology and Psychiatry Review, 1,* 104–109.

Shakow, D. (1978). Clinical psychology seen some 50 years later. *American Psychologist, 33,* 148–158.

Spitzer, A., Webster-Stratton, C., & Hollinsworth, T. (1991). Coping with conduct-problem children: Parents gaining knowledge and control. *Journal of Clinical Child Psychology, 20,* 413–427.

Sturmey, P. (1996). *Functional analysis in clinical psychology*. Chichester, UK: Wiley.

Sutton, C., & Herbert, M. (1992). *Mental health: A client support resource pack*. Windsor, UK: National Foundation of Educational Research—Nelson.

Thomas, A., Chess, S., & Birch, H. G. (1968). *Temperament and behavior disorders in children*. London: University of London Press.

Tuma. J. M. (1989). Traditional therapies with children. In T.H. Ollendick & M. Hersen (Eds.), *Handbook of child psychopathology* (2nd ed.). New York: Plenum.

Vetere, A., & Gale, A. (1987). *Ecological studies of family life*. Chichester, UK: Wiley.

Wachtel, P. L. (1977). *Psychoanalysis and behavior therapy*. New York: Basic Books.

Watzlawick, P., Weakland, J., & Fisch, R. (1974). *Changes: principles of problem formulation and problem resolution*. New York: Norton.

Webster, R. (1995). *Why Freud was wrong: Sin, science and psychoanalysis*. London: Harper Collins.

Webster-Stratton, C., & Herbert, M. (1993). What really happens in parent training? *Behavior Modification, 17,* 407–456.

Webster-Stratton, C., & Herbert, M. (1994). *Troubled families, problem children: Working with parents—a collaborative approach*. Chichester, UK: Wiley.

Whiting, B. (1963). *Six cultures: Studies of child rearing*. New York: Wiley.

Wilson, F. E., & Evans, J.M. (1983). The reliability of target-behavior selection in behavioral assessment. *Behavioral Assessment, 5,* 15–23.

World Health Organization (1992). *The ICD-10 classification of mental and behavioral disorders: Clinical descriptions and diagnostic guidelines*. Geneva, Switzerland: World Health Organization.

Yelloly, M. (1972). The concept of insight. In D. Jehu (Ed.), *Behavior modification in social work*. Chichester, UK: Wiley.

Zwick, R. (1983). Assessing the psychometric properties of psychodiagnostic systems: How do research diagnostic criteria measure up? *Journal of Consulting and Clinical Psychology, 51,* 117–131.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.03

Process Issues in Child Psychotherapy

STEPHEN R. SHIRK
University of Denver, CO, USA
and
ROBERT L. RUSSELL
Loyola University Chicago, IL, USA

5.03.1 INTRODUCTION 57
5.03.2 REFERRAL TO CHILD PSYCHOTHERAPY 58
  5.03.2.1 Parent and Teacher Perceptions of Problem Behaviors 58
    5.03.2.1.1 Child age and referral 59
    5.03.2.1.2 Child gender and referral 60
    5.03.2.1.3 Referral and cultural background 60
    5.03.2.1.4 Referral and parental psychopathology 61
  5.03.2.2 Process Issues in Clinical Referral 62
5.03.3 READINESS AND PREPARATION FOR CHILD THERAPY 62
  5.03.3.1 Readiness for Therapy 62
    5.03.3.1.1 Problem acknowledgment 63
    5.03.3.1.2 Problem attributions 63
    5.03.3.1.3 Help-seeking expectations 64
  5.03.3.2 Preparation for Child Therapy 65
  5.03.3.3 Process Issues in Readiness and Preparation for Therapy 66
5.03.4 IN-SESSION TRANSACTIONS IN CHILD THERAPY 68
  5.03.4.1 Interpersonal Transactions in Child Therapy 68
    5.03.4.1.1 Research on interpersonal processes in child therapy 69
  5.03.4.2 Process Issues in Interpersonal Transactions 71
    5.03.4.2.1 Future directions for research 72
  5.03.4.3 Symbolic Transactions in Child Psychotherapy 74
    5.03.4.3.1 Future directions for research 76
5.03.5 THERAPEUTIC TRANSACTIONS AND TREATMENT OUTCOMES 76
  5.03.5.1 Promising Directions for Process–Outcome Research 78
5.03.6 CONCLUSIONS 79
5.03.7 REFERENCES 80

5.03.1 INTRODUCTION

According to Orlinsky and Howard (1986, pp. 311–312) the therapy process refers to "everything that can be observed to occur between and within the patient and therapist during their work together." This is indeed a broad domain of inquiry; nevertheless, research on therapeutic transactions, patient and therapist characteristics, and change events during treatment have been relatively neglected by child therapy investigators. As Kazdin, Bass, Ayers, and Rodgers (1990) have reported, less than 3% of all child therapy studies have investigated treatment processes, and few studies have examined either child or therapist characteristics that could moderate treatment outcome. Although there is some evidence to suggest increasing attention to these aspects of child treatment since the 1980s (cf. Russell & Shirk, in press; Shirk & Russell,1996), much of the child therapy literature retains a restricted focus on comparisons of treatment packages or, to a lesser degree, on the evaluation of specific techniques.

The relative neglect of child therapy processes is puzzling for a number of reasons. First, from a clinical perspective, the primary experience of practicing therapists is the ongoing transactions that comprise therapy sessions. In essence, social and symbolic exchanges constitute the principle phenomena of child psychotherapy. Viewed from this perspective, one might expect the study of process to be at the forefront of child therapy research. In fact, Kazdin et al. (1990) found that child practitioners regard studies of process and its relation to outcome to be a top research priority. Yet the investigation of process remains in the background of child therapy research. Second, growing evidence from the adult psychotherapy literature has shown that variation in processes within sessions, as well as characteristics of therapists and patients, have an important bearing on treatment outcome (Orlinsky, Grawe, & Parks, 1994). Nevertheless, process research is "perhaps the area of work that is most discrepant between child and adult therapy" (Kazdin, 1995, p. 268).

Given the restricted focus of child therapy research, it is proposed to cast a broad net across empirical literatures that could inform understanding of child therapy processes. It has been argued elsewhere (cf. Shirk & Russell, 1996) that much of the child psychotherapy literature has evolved in isolation from basic developmental research that entails important implications for understanding therapeutic processes. Therefore, the approach to process issues in child psychotherapy is developmental in two ways. First, child therapy, regardless of treatment type, can be viewed as developmental process. Psychotherapy process is fundamentally concerned with the course of change over time. What makes this process developmental is that the aim of therapeutic change involves the progressive transformation of the child's basic capacities to interact competently in the world. Second, research on developmental processes in everyday, nontherapeutic contexts can have an important bearing on how the process of child treatment is understood (Shirk, 1988; Shirk & Russell, 1996). Consequently, the chapter is not restincted to studies of child therapy process *per se* but integrates research from the developmental literature where appropriate.

A number of models of therapeutic change have been advanced in the adult therapy literature (cf. Howard, Lueger, Mahling, & Martinovich, 1993; Prochaska, DiClemente, & Norcross, 1992). Common to these models is a sequential perspective on the course of therapy that delineates different phases of treatment with characteristic treatment processes. Such a framework has yet to be developed for child treatment. As a start, this review of child therapy process is organized around salient phases of treatment. Specifically, this chapter follows the "flow" of therapeutic interaction from referral to termination. Along the way factors that influence the referral process, readiness or preparation for treatment, social and symbolic exchanges in sessions, and the relation between process and outcomes are considered. The primary aim is to identify critical process issues that have relevance for the practice of child therapy and for the conduct of future research.

5.03.2 REFERRAL TO CHILD PSYCHOTHERAPY

Unlike adults, children rarely refer themselves for psychological treatment. In fact, it is not uncommon for others to be more distressed by the child's problems than is the child. Referral to child therapy typically involves judgments and decisions by other people, usually parents and teachers, in the child's life. Thus, from the outset, the process of child therapy is embedded in a social context that goes beyond the transactions between patient and therapist. In order to understand the referral process in child therapy, it is essential to examine the reactions of "referring" adults to various forms of child problem behaviors. To this end, investigators have begun to examine factors that influence parent and teacher decision making about referral to child treatment.

5.03.2.1 Parent and Teacher Perceptions of Problem Behaviors

One of the well-worn paths to the child psychology clinic begins with the parent–teacher conference. In this context, adult perceptions and appraisals of child behavior problems influence decisions about referral. Because teachers are a primary referral source, a number of investigators have examined their evaluations of children's problem behaviors. Walker, Bettes, and Ceci (1984) asked preschool teachers to evaluate the severity, long-term outcome, stability, and need for referral of three vignettes that depicted children at two ages (3.5 and 5.5 years) who presented with aggression, hyperactivity, or withdrawal. The gender of the child was also varied. Among the central findings of the study, preschool teachers tended to view disruptive behavior problems as more severe, more stable, and portending poorer long-term outcomes than problems of withdrawal.

Consistent with these evaluations, teachers were more likely to agree that disruptive children should be referred for professional help than children showing signs of withdrawal. A similar pattern of results was obtained by Weisz et al. (1988), who found that elementary school teachers rated undercontrolled problems, relative to overcontrolled problems, as more serious, worrisome, and less likely to improve. Examination of referral problems in US clinics has revealed higher rates of undercontrolled problems than overcontrolled problems. Although it is likely such differences reflect variations in prevalence rates for the two broad classes of disorders, it is also possible that differences in problem appraisal among referring adults contributes to this difference.

One implication of this pattern—that undercontrolled problems are viewed as more serious than overcontrolled problems—is that children with symptoms of depression and anxiety may be under-referred for treatment because of beliefs that such problems are transient and will diminish over time. Research, however, has indicated that internalizing problems may be more persistent than once believed, and that untreated internalizing problems constitute a risk factor for future psychiatric disorders (Ollendick & King, 1994).

An alternative approach to evaluating the "referability" of various types of emotional and behavioral problems has been advanced by Weisz and Weiss (1991, p. 267), who operationalized referability as "the frequency with which a child problem is reported as a presenting complaint at mental health clinics, relative to the frequency of that problem in the general population." Epidemiological data were collected in both the USA and Thailand with a broad-band measure of child behavior problems. Clinic records from multiple sites were coded for specific presenting problems in both countries. Referability indices were computed for specific behavioral problems for children between the ages of 6 and 11 years. Results from the US sample indicated that problems with serious implications for self or others, such as attacking other people or deliberately harming the self, were among the most referable problems. And like the results from the vignette studies, overcontrolled problems, such as perfectionism and self-consciousness, were among the least referable problems. This pattern of results suggests that the lower number of internalizing problems in clinics is not just a function of differences in problem prevalence but may reflect parental judgments about the prognosis for specific problems and the degree to which such problems can be remediated or tolerated in the home.

The foregoing studies suggest that problem type is an important contributor to decisions about clinic referral. Overall, undercontrolled problems appear to have the greatest capacity to elicit referral. Other factors, however, may influence parents' and teachers' referral decisions. A number of investigators have examined other child and family characteristics that could affect the referral process, including: child age, gender, and cultural background. In addition, numerous studies have suggested that parental characteristics, specifically parental psychopathology, are associated with perceptions of child behavior problems.

5.03.2.1.1 Child age and referral

Results from several studies have shown that younger children in the classroom are referred for mental health services at a higher rate than their older classmates (Drabman, Tarnowski, & Kelly, 1987; Tarnowski, Anderson, Drabman, & Kelly, 1990; Wisniewski, Andrews, & Mulick, 1995). Although one might expect these children to be less mature and possibly less advanced academically, Tarnowski et al. did not find significant differences in cognitive functioning or academic achievement between younger and other children in the same classroom. This pattern of findings was replicated and extended by Wisniewski et al., who also found a disproportionately high referral rate for younger children, and did not find group differences in indices of cognitive functioning or in levels of behavior problems. Thus, the disproportionate rate of referral did not seem to be attributable to "objective" age differences in the rate or severity of behavioral or academic problems.

The lack of a systematic relationship between referral and "objective" reasons for referral has suggested the possible contribution of teacher bias in the referral process. In their analysis of child characteristics associated with referral, Wisniewski et al. (1995) found that as children deviated upward from the mean in height and weight, they were referred at disportionately higher rates. This tendency was particularly prominent among younger children in the classroom. As Wisniewski et al. observed, children's physical characteristics appear to be influencing teachers' referral decisions. Younger children who are relatively large for their age may be referred at a higher rate because they are misperceived to be similar to older children on academic dimensions. Thus, when these less mature children function at a level that is discrepant with teacher expectations for the average child in the class, but appropriate for their age level, they are viewed as needing remedial services (Wisniewski et al., 1995). In essence, perceived discrepancies based on nonacademic factors (physical size) appear to contribute to teacher referral decisions. This interpretation is consistent with the formulation advanced by Masten and Braswell (in press), who propose that referral decisions are linked to judgments about discrepancies between age and the successful negotiation of specific developmental tasks. Teachers may have clear expectations for grade level, and expect all children at that level to complete specific tasks regardless of variation in age within the classroom.

5.03.2.1.2 Child gender and referral

Despite developmental evidence suggesting that parents monitor their daughters' behavior more closely than their sons', and tolerate less behavioral disruption from female children than males (Zahn-Waxler, 1993), boys are more often referred to clinics than girls. Several explanations can be offered for this apparent inconsistency. First, boys may simply evince higher rates of emotional and behavioral problems than girls (McDermott, 1996), consequently differences in prevalence rates may account for differences in referral rates. Second, although parents monitor their daughters' behavior more closely than their sons', differences in the types of problems expressed by males and females could affect referral. Given evidence for more disruptive behavior problems among males than females (McDermott, 1996), and evidence indicating that such problems are viewed as more serious by teachers and parents, gender differences in type of problem could account for gender differences in referral. Parents might conclude that disruptive, particularly aggressive, problems cannot be handled without professional help. On the other hand, the tendency to view internalizing problems as more transient and less severe may result in the underreferral of females who have been shown to express more internalizing problems, such as depression, particularly during adolescence (Nolen-Hoeksema, 1990).

In this context, the finding of Weisz and Weiss (1991) that problems are more referable for girls than for boys, at least in the USA, is somewhat surprising. Although gender did not affect the relative referability of specific problems for males and females, fighting, for example was highly referable for both genders; in general, problems were more referable for girls than boys. Thus, when girls show a specific problem, they are more likely to be referred than boys—although it is important to keep in mind that a number of problems characteristic of depression have low referability.

One area where this is particularly the case, are problems that load on the Sex Problems factor of the Child Behavior Checklist (Achenbach & Edelbrock, 1981). Although girls do not show an overall higher prevalence rate on these types of problems, when they do show such problems they are more likely to be referred than boys. Taken together, evidence for greater referability for females than males, but higher rates of male than female referral to clinics, suggests that higher rates of behavior problems, particularly undercontrolled problems, among young males relative to females account for their overrepresentation in clinics.

In summary, although parents and teachers do not report differences in the seriousness of problems when they evaluate males and females in vignettes (Walker et al., 1984; Weisz et al., 1988), differences in referability suggest that their actual referral decisions may be influenced by the child's gender, and possibly by different sex-typed expectations for boys and girls. The overrepresentation of males in clinics appears to follow from higher rates of undercontrolled behavior problems relative to females, and possibly to the under-referral of females with internalizing problems.

5.03.2.1.3 Referral and cultural background

A series of studies has revealed that culture plays an important role in adults' perceptions of child emotional and behavioral problems (Lambert, Weisz, & Knight, 1989; Lambert et al., 1992; Weisz et al., 1988; Weisz & Weiss, 1991). For example, in comparing the perceptions of American and Jamaican adults, Lambert et al. (1992) found that cultural background was related to evaluations of undercontrolled behaviors but not overcontrolled problems, such that Jamaicans seemed more concerned about disruptive problems than American adults. Cultural differences were also reported by Weisz et al., who found that Thai adults expressed less concern about both over- and undercontrolled behaviors than American adults. This pattern was also partially supported in the "referability" study of Weisz and Weiss. Undercontrolled problems, but not overcontrolled problems, were more referable in the USA than Thailand.

These findings suggest that culture may set a threshold for adult tolerance for distress about childhood problems, and that tolerance levels may vary by type of problem (Lambert et al., 1992). For example, Weisz and Weiss (1991) found that problems such as fighting, swearing, and threatening were more likely to prompt referral in the USA than Thailand, whereas the opposite was true for such problems as strange behavior, fears of school, and liking to be alone. Thus, one factor that may influence the degree to which problems are regarded as serious enough for referral is the cultural context in which they occur. It is likely that ethnic variation within the same society could exert a similar influence; however, such findings have yet to be reported. Clinical reports also have suggested that members of different ethnic groups vary in their interest or willingness to utilize professional services for child behavior problems (cf. Tharp 1991).

5.03.2.1.4 Referral and parental psychopathology

Parental accounts of child behavior problems are often presented with a strong affective valence. Consequently, parents, who are often entangled in their child's presenting problems, are not necessarily neutral reporters. This has raised concerns that parents' own affective states could bias their reports of child behavior problems. This clinical concern appears to be validated by a growing body of evidence that suggests that parental depression is associated with ratings of child behavior problems (cf. Schaughency & Lahey, 1985; Webster-Stratton, 1988). Specifically, the association between maternal depression and ratings of externalizing problems has been established in multiple studies, and has prompted some to conclude that "parent's depression rather than the child's behavior is a significant factor in the referral of these children for treatment" (Rickard, Forehand, Wells, Greist, & McMahon, 1981, p. 204).

From a process perspective, depression, or other forms of parental distress, may narrow the parent's tolerance for variation in child behavior such that minor behavioral deviations are misperceived as abnormal. Depressed parents, relative to their nondepressed counterparts, may become more distressed by typical child misbehavior, and amplify the severity of child problems in their behavioral reports. To the degree that referral is associated with parental judgments about the abnormality of behavior, the narrowed band of tolerance found among depressed parents would result in a lowered threshold for referral. In essence, referral is based on a distorted perception of child behavior problems.

In his review of research on the link between maternal depression and distorted perception of child behavior problems, Richters (1992) raised serious concerns about the existing evidence for this connection. Obviously, inferences about perceptual distortions do not directly follow from group differences in behavioral reports of depressed and nondepressed mothers. High rates of behavior problems could stem from problematic interactions in families with a depressed parent rather than from parental misperceptions. According to Richters, the distortion hypothesis requires (i) depression-related discrepancies with criterion ratings such that depressed mothers show less concordance with validated criterion than do nondepressed mothers, and (ii) evidence that the criterion rating is more reliable or accurate than maternal ratings, themselves. Based on a review of 17 studies that reported a depression–distortion link, Richters concluded that the vast majority fail to meet the foregoing criteria. Consequently, he concluded that "there appears to be no empirical foundation for the widespread belief that depressed mothers have distorted perceptions of their children's problems" (Richters, 1992, p. 496). However, as Richters wisely noted, the absence of evidence should not be confused with evidence for absence of distortion. Although existing research has revealed a relation between parental depression and child problem behaviors, the causal relationship and relative contribution of parental perceptions have not been clearly established.

Parents are in a unique position to evaluate the nature and severity of child problem behaviors, for example, they are aware of the child's behavior across contexts and over time; however, parents are often participants in or recipients of the child's problem behaviors. Consequently, parents' perceptions of their child's behavior could be influenced not just by their own affective state but by their role in their child's problem. Research in this area is virtually nonexistent; however, Reid, Kavanagh, and Baldwin (1987) have observed that abusive parents appear to view their children's behavior as more problematic than suggested by extensive behavioral observations. As Reid et al. have noted, the perception of child behavior as highly problematic may function to rationalize maltreatment by abusive parents. Such findings serve as a reminder that parents are participant-observers of their children's behavior, and that such a standpoint may contribute to biased perceptions or interpretations of child behavior.

5.03.2.2 Process Issues in Clinical Referral

Adult perceptions of child problems are an integral part of the clinical referral of children. Existing research has revealed that problem type, child age, gender, and cultural context are associated with adult perceptions of children's problems. Although there is evidence for a relationship between maternal depression and reports of child behavior problems, the accuracy of depressed mothers' perceptions has not been adequately addressed and awaits further investigation.

One of the major limitations of research on referral to child therapy is the absence of an explicit process model. Fragments of a model can be found across studies, but an integrated perspective on the relative contribution of specific judgments, for example, about the severity, abnormality, prognosis, and etiology of child problems, to referral decisions has not been advanced. As in all process research, the guiding question should be *how* do parents arrive at a decision to refer their children for treatment. For example, it is possible that referral involves the interplay between cognitive processes, such as judgments about abnormality (itself a function of social comparison processes), and emotional processes, such as demoralization following repeated failures at resolving the problem within the family. Future studies could manipulate these factors in case vignettes, for example, problem severity and intransigence, in order to evaluate their contribution to referral decisions. Other factors, such as the availability of services and parental perceptions of the efficacy of such services, are likely to contribute to the referral process as well.

5.03.3 READINESS AND PREPARATION FOR CHILD THERAPY

Treatments vary in the degree to which child participation is critical for successful outcome. Traditional child treatments such as psychodynamic, experiential, and play therapies, as well as cognitive and cognitive-behavioral therapies, require active child involvement in therapy. Although treatment tasks vary substantially across these therapies, they share the common assumption that children's active participation is essential for successful outcomes. In contrast, some child treatments such as parent management training require only minimal, if any, participation by the child in therapy sessions. Instead, parent collaboration with treatment is critical for altering child behavior (Patterson & Chamberlain, 1994). It should be noted, however, that many child behavioral treatments, such as desensitization and social skills training, require high levels of child participation in sessions.

Research on adult psychotherapy has indicated that active participation in treatment is one of the best process predictors of treatment outcome (Frieswyk et al., 1986; Gomes-Schwartz, 1978). In this context it is interesting to note that research has indicated that nearly two-thirds of all children aged between 7 and 15 years would feel negatively if they were referred to see a psychologist (Sigelman & Mansfield, 1992). Given that children appear to be "reluctant consumers" of mental health services, it is surprising that few studies have addressed children's participation in treatment and its relation to outcome.

One of the few studies to address this issue, conducted by Braswell, Kendall, Braith, Carey, and Vye (1985), examined children's level of involvement in cognitive–behavioral therapy and found that a number of verbal indicators of active participation predicted therapeutic gains. Specifically, children's tendency to make spontaneous verbal elaborations about treatment tasks and therapists" encouragement and corrective feedback appeared to be among the best predictors of change in teacher ratings of self-control. Similarly, Colson et al. (1991) found that ratings of treatment difficulty, that is, the degree to which patients' were viewed as defiant and uncooperative in treatment, were negatively associated with ratings of patient progress in a sample of hospitalized adolescents. The findings of this study, although suggestive of a link between quality of participation and treatment outcome, are compromised by the absence of independent process and outcome ratings. Despite limited empirical findings on the relation between in-session participation and treatment outcome in the child literature, clinical theory and evidence from the adult therapy literature indicate that the study of treatment participation and the factors that promote or obstruct it represent a critical area for child psychotherapy process research.

5.03.3.1 Readiness for Therapy

Some years ago, Anna Freud (1965) recognized that a number of developmental factors could interfere with children's willingness to participate in therapy. A major obstacle, according to Freud, is the child's lack of awareness or acknowledgment of his or her own emotional or behavioral problems. As she noted, due to their immaturity, children often have "no insight into their abnormalities" and that, accordingly, "they do not develop the same wish to get well and the same type of treatment alliance" as adults (Freud, 1965, p. 28). In the absence of self-acknowledged emotional or behavioral problems, children do not experience a need for change, and, consequently, a need for treatment. Moreover, when problems are acknowledged, Freud observed that children typically locate the source of the problem in the external world and prefer environmental or situational solutions to their problems. This type of problem construal is incongruent with forms of therapy that focus on child-centered treatment processes such as altering affect regulation strategies, modifying cognitive patterns, or developing specific social skills. Furthermore, even when children recognize emotional or behavioral problems, and acknowledge a role in them, they do not necessarily view others as useful agents of change. Consequently, the degree to which children view interpersonal coping strategies as useful or effective could influence their willingness to utilize psychotherapy to deal with their difficulties. Given these potential obstacles, treatment engagement represents one of the major challenges of child therapy.

In an effort to integrate multiple processes that could influence children's willingness to participate in therapy, Shirk (Shirk, 1990; Shirk & Saiz, 1992) presented a developmental social–cognitive model of alliance formation. Elaborating on Anna Freud's clinical insights (Freud, 1965), Shirk hypothesized that developmental variations in children's cognitions about their own emotions and behaviors could contribute to their willingness to engage in the process of therapy.

5.03.3.1.1 Problem acknowledgment

The first source of social cognitive influence is hypothesized to be children's recognition and acknowledgment of emotional or behavioral problems. Children who do not acknowledge problems are expected to be more difficult to engage in a process that involves problem remediation. Moreover, from a developmental perspective, the ability to recognize one's emotional and behavioral difficulties should be related to the development of self-evaluative, especially social comparative, processes. Younger children, lacking in perspective-taking ability, find it difficult to distinguish between wishful images of self and actual characteristics (Leahy & Huard, 1976). In fact, research by Harter (1988) has indicated that younger children, especially those of preschool age, overrate competencies and minimize difficulties. Furthermore, there is some evidence to suggest that this tendency is amplified in clinic populations (Vondra, Barnett, & Cicchetti, 1989; Zimet & Farley, 1986). In fact, it is not uncommon to find significant discrepancies between parent and child evaluations of problem behaviors and emotions (Achenbach, McConaughy, & Howell, 1987; Kazdin, French, & Unis, 1983).

In a study with a heterogeneous sample of psychiatrically hospitalized children, Shirk, Saiz, and Sarlin (1993) found that children's level of problem acknowledgment at admission was predictive of therapists" ratings of the treatment alliance 3 weeks later. This finding is consistent with earlier research indicating that children's reluctance to participate in therapy is linked to problem denial (Taylor, Adelman, & Kaser-Boyd, 1985). Closer analysis of the alliance data indicated that problem acknowledgment predicted the affective quality of the therapeutic relationship but not children's level of collaboration with therapeutic tasks. The failure to find a relationship with task collaboration could be due to methodological limitations. First, task collaboration was narrowly defined as verbalization about problems and feelings; and, second, verbalization was assessed through global ratings by therapists. It is possible that specific coding of verbal interactions by observers would have yielded different results. In fact, research by Russell, Bryant, and Estrada (1996) has shown that child therapy sessions in which there is a positive therapeutic relationship are characterized by distinctive patterns of child–therapist discourse. However, the link between child problem acknowledgment and involvement in treatment processes, including verbal interaction, awaits further investigation.

5.03.3.1.2 Problem attributions

The second postulated source of social-cognitive influence involves children's attributions for their presenting problems. Research on adult psychotherapy has indicated that patients who view their problems as externally imposed are not likely to benefit from treatment (Kirtner & Cartwright, 1958; Salzman, Luetgert, Roth, Creasant, & Howard, 1976). Similarly, children may acknowledge the existence of problems but locate the source of their problems in the external environment. Although the child's attributions may be accurate in some cases, and consistent with certain forms of therapy, for example, family therapy, for child-centered treatments such an attributional pattern could undermine willingness to participate in child-focused interventions.

Developmental research indicates that children's understanding of the causes of behavior shows a general trend from external to internal attributions with increasing age (Shirk, 1988). As children acquire complex concepts of internal, psychological functioning, and develop the capacity to reflect on such processes, these concepts can be applied to observed behavior, including their own problematic behavior. As a result, internal attributions for behavior problems should increase with age during middle childhood. In fact, research on children's understanding of deviant behavior has revealed that children increasingly attribute such behavior to internal, psychological processes with increasing age (Coie & Pennington, 1976).

Two treatment outcome studies point to the potential relevance of children's attributions for treatment process. In the first, Bugental, Whalen, and Henker (1977) predicted that children's causal attributions for academic outcomes would moderate the effect of two types of treatment, a 6 week self-instructional training program (internal locus) or a 6 week social reinforcement program (external locus) for hyperactive boys. It was predicted that outcomes would be more positive for subjects with attribution-congruent treatments, for example, social contingency treatment for children with external attributions, than for those with incongruent treatments. Although the treatments did not produce changes in teacher ratings of disruptive behavior, subjects in the attribution-congruent treatments showed greater improvement on a measure of cognitive functioning. Thus, it appeared that treatment outcome was, in part, a function of the match between child attributional pattern and locus of treatment. It is possible that children in attribution-incongruent treatments were less engaged, or viewed the treatments as less acceptable, and therefore benefited less than children in congruent treatments; however, this process hypothesis was not evaluated.

Research by Weisz (1986) on the relationship between control-related beliefs and treatment outcome in child therapy carries implications for the link between problem attribution and treatment involvement. According to Weisz, children vary in the degree to which they view outcomes as contingent on their own behavior or efforts. In a therapeutic context, children who believe that problem resolution is contingent on altering their own behavior should be more willing to engage in therapeutic problem solving than children who view problem solutions as independent of their efforts. As Weisz noted, individual differences in contingency beliefs could lead to corresponding differences in therapeutic involvement, thereby resulting in differential rates of change during treatment. In a study with 8–17 year olds referred for outpatient therapy, Weisz found a significant relationship between contingency beliefs, measured prior to treatment, and total problem reduction. Although these results are consistent with the view that attributions for problems promote or obstruct therapeutic engagement, the level of treatment involvement was not directly assessed in this study. Consequently, the link between pretreatment attributional patterns and therapy process has yet to be investigated.

5.03.3.1.3 Help-seeking expectations

A third source of social cognitive influence involves children's beliefs about the usefulness of other persons in resolving emotional or behavioral difficulties. From a developmental perspective, beliefs about the responsiveness and effectiveness of potential caregivers, including therapists, is likely to be rooted in early experiences of care (Shirk & Saiz, 1992). One of the basic tenets of attachment theory is that individuals develop internal working models of the self in relation to attachment figures that reflect the history of caregivers' reactions to the child's emotional needs and distress (Bowlby, 1988; Main, Kaplan, & Cassidy, 1985). Repetitions of infant–caregiver interactions are presumed to establish core beliefs about the availability and responsiveness of others.

Approach or avoidance patterns in times of distress are hypothesized to be linked to these core beliefs (Kobak & Sceery, 1988). Research on children's coping strategies has revealed significant individual differences in the degree to which children utilize others when confronted with psychological stressors (Garber, Braafladt, & Zeman, 1991). Among some groups of troubled children, for example, children displaying symptoms of depression, the use of others for coping with distress appears to be relatively limited (Garber et al., 1991). In fact, Garber, Braafladt, and Weiss (1995) have found that depressed girls rated mother-initiated affect regulation strategies as less effective than did nondepressed girls. Van Horn (1996) found that beliefs about the effectiveness of interpersonal coping strategies was associated with preadolescents perceptions of prior maternal care. Taken together, these results suggest that early experiences of care could shape children's expectations about the usefulness of other persons, including therapists, in helping them deal with emotional or behavioral difficulties.

Although there is growing interest in the relationship between attachment related beliefs and psychotherapy process and outcome among adults (cf. Dozier, Cue, & Barnett, 1994; Fonagy et al., 1996), this potentially fruitful area remains unexplored with children in treatment.

In addition to pretreatment social cognitions, other child characteristics could influence children's willingness to engage in therapy. For example, patterns of interpersonal problems, such as those that characterize oppositional–defiant disorder, are likely to carry over into therapeutic interactions and obstruct treatment participation. In this connection, Eltz, Shirk, and Sarlin (1995) found that the number of interpersonal problems reported prior to treatment was negatively related to the development of a working alliance among adolescent inpatients. In addition, characteristics associated with specific disorders could constitute impediments to treatment. For example, the relatively low levels of social competence and reflectiveness found among children with externalizing problems could interfere with their involvement in therapy. Surprisingly, children with externalizing symptoms do not show poorer outcomes than children with internalizing symptoms (Weisz, Weiss, Han, Granger, & Morton, 1995); however, this result could reflect the use of parent-centered interventions, for example, parent management training, rather than child-centered interventions with externalizing disorders. Among the internalizing disorders, symptoms of depression such as hopelessness, lack of energy, or social withdrawal could undermine participation in treatment (Stark et al., 1991).

In summary, studies of differential treatment responsiveness are beginning to consider pretreatment child characteristics (cf. Kazdin, 1995); however, few studies have examined these characteristics in relation to treatment process. What is needed are studies that examine relations among child characteristics that are hypothesized to promote or obstruct children's willingness to participate in treatment, specific dimensions of treatment process such as participation in treatment tasks, and variations in treatment outcome. Unlike comparisons of treatment packages, where individual patterns are typically ignored, such studies could potentially reveal patterns of "fit" between child characteristics and treatment processes that are associated with improved outcomes.

5.03.3.2 Preparation for Child Therapy

Evidence from both the adult and child psychotherapy literatures has indicated that discrepancies between client's expectations about the roles and processes of therapy and the reality they encounter in treatment can have a variety of negative effects, including willingness to sustain participation in treatment (Coleman & Kaplan, 1990; Day & Reznikoff, 1980; Frank, 1974; Strupp & Hadley, 1977). Studies with normative samples of children have shown that young children often have limited knowledge about psychologists' roles, practices, and the process of psychotherapy, but that knowledge and expectations become increasingly accurate with age (Bonner & Everett, 1982; Sigelman & Mansfield, 1992). Interestingly, there is some evidence to suggest that children's perceptions of psychologists' traits, such as smart or perceptive, become less positive with age (Sigelman & Mansfield, 1992), though overall attractiveness of a depicted therapist remains high among normative samples (Bonner & Everett, 1980). Furthermore, it has been shown that appropriateness of children's expectations are associated with accuracy of parental expectations (Day & Reznikoff, 1980). Of central importance, however, is a finding indicating that inappropriate expectations about therapy are related to premature termination from child treatment (Day & Reznikoff, 1980). This pattern of findings has prompted investigators to design preparation techniques that could counter inappropriate expectations and sustain or enhance children's participation in treatment.

A number of preparation techniques have been devised, including pretherapy preparation interviews during which the purpose, process, and roles of therapy are explained (Holmes & Urie, 1975), audiotape presentations in which a "radio announcer" and child interview a therapist about the structure, roles, and processes of therapy (Bonner & Everett, 1986), and videotape presentations in which child participants are interviewed about their experiences and/or simulated therapy process is modeled (Coleman & Kaplan, 1990; Weinstein, 1988).

One of the most consistent findings to emerge from studies of therapy preparation has been that prepared children and parents evince greater knowledge and understanding about the structure and process of therapy than their nonprepared counterparts (Bonner & Everett, 1986; Coleman & Kaplan, 1990; Day & Reznikoff, 1980; Weinstein, 1988). However, the impact of this increased knowledge and understanding on therapy process and outcome is less clear. Results are mixed with regard to preventing premature drop-out from therapy. For example, Holmes and Urie (1975) found that prepared clients were less likely to terminate therapy prematurely than nonprepared clients, whereas two other studies failed to demonstrate this effect (Day & Reznikoff, 1980; Weinstein, 1988). Day and Reznikoff did find that prepared families had fewer failed or cancelled appointments than nonprepared families. Thus, the impact of these preparatory interventions on maintaining participation in treatment has yet to be clearly established. In fairness, research in this area has been hampered by small, heterogeneous samples that may obscure preparatory effects. However, as Weinstein has noted, the modest impact of preparatory interventions may reflect the fact that most interventions have not been guided by conceptual models of treatment engagement.

Another aim of preparatory interventions has been to enhance children's active participation in therapy. Several studies have examined relations between therapy preparation and indices of treatment process. Unfortunately, the results have been disappointing. For example, Day and Reznikoff (1980) examined the relationship between preparation and child verbalization in therapy. Preparation consisted of videotaped modeling of children talking in therapy. Despite the close relation between the preparation technique and the measure of process, comparisons of verbal samples from three sessions failed to show an effect for preparation. In a related study, Weinstein (1988) examined the relation between preparation and children's ability to adapt to the client role. Adaptation to role was assessed by child and therapist questionnaires that tapped children's comfort with therapy and their communication with the therapist. Again the preparatory intervention involved children modeling active participation in treatment. There was also no effect for preparation on either the quality of the therapeutic relationship or the character of communication. In one of the only studies to examine preparation in relation to outcome, Coleman and Kaplan (1990) found that mothers in the prepared group reported a significantly greater decrease in problem behaviors by the fourth session than did nonprepared mothers. Unfortunately, no results were reported for post-treatment outcomes; consequently, it is not clear whether these short-term effects reflected transitory changes in perceptions or enduring, enhanced outcomes. Consistent with these findings, Bonner and Everett (1986) found that preparation increased parents' and children's expectations for positive treatment outcomes. Thus, it is possible that preparation influences parents' prognostic expectations which are reflected in more positive perceptions of their children's problems relative to pretreatment. Alternatively, changes in child and parent outcome expectations could have an immediate impact on child behavior problems that may or may not endure over the course of treatment. The fact that prepared children did not report greater behavioral improvement than nonprepared children favors the interpretation that preparation may influence maternal perceptions of child problems. Research in this area would be advanced by evaluating the durability of these effects, and by the assessment of change by sources who are blind to preparation condition.

In summary, despite consistent evidence showing that treatment preparation enhances parents' and children's understanding of therapy, evidence for the impact of improved understanding on therapy process is quite limited. It is possible that other therapist, child, and family characteristics exert a greater influence on treatment participation than general knowledge of and expectations for therapy. The goodness of fit between child and therapist or parent and therapist is likely to be a powerful determinant of treatment involvement and continuation; however, research on therapist–client matching in child treatment lags far behind the adult literature (cf. Beutler, 1991). In fairness, few studies have examined the relation between preparation and treatment process; consequently, it is probably premature to dismiss such interventions as ineffective. However, research on treatment preparation could be improved by a closer analysis of child and parent cognitions that interfere with treatment participation. Targeting specific "interfering" cognitions could prove to be more effective than interventions aimed at increasing general knowledge about the structure and process of therapy.

5.03.3.3 Process Issues in Readiness and Preparation for Therapy

Initiation of child therapy is complicated by the fact that children rarely seek psychological services on their own. Youngsters often enter therapy with considerable reluctance that is frequently associated with (i) negative perceptions of therapy, (ii) lack of problem acknowledgment, and (iii) lack of choice in the decision to seek help (Taylor et al., 1985). These findings suggest that children's cognitions about therapy and their own difficulties, as well as their control-related beliefs and help-seeking expectations, may impact their willingness to engage in treatment. Unfortunately, research on readiness and preparation for child therapy is marked by a number of major limitations.

First, and perhaps most striking, is the dearth of validated measures of child treatment participation. Before investigators can identify predictors of treatment involvement, conceptual and operational definitions of positive participation must be advanced and assessed. It should be noted that a few studies have produced promising leads. For example, research by Braswell et al. (1985) suggests that children's verbalizations in sessions can be used as markers of active involvement. Of course, given the fact that verbal activity in therapy varies with age (Lebo & Lebo, 1957), and that different forms of therapy rely on verbal interaction to varying degrees (Russell, Greenwald, & Shirk, 1991), a single index of verbal participation is not likely to be an adequate measure of involvement for different forms of treatment and for children at different developmental levels. In related research, Estrada, Russell, McGlinchey, and Hoffman (1994) developed an observational rating scale for identifying dimensions of participation in child therapy sessions. Factor analysis of ratings of child behavior produced three factors that the authors labeled therapeutic relationship, therapeutic work, and readiness. The therapeutic work factor was defined by high loadings on items assessing productivity, understanding, exploration of feelings, and engagement. Given that the work factor was derived from observations of psychodynamic therapy, it is not clear whether similar factors would emerge from ratings of other types of treatments, or if other indicators of active participation would be more relevant in other forms of child therapy.

Although these initial attempts to operationalize therapeutic participation appear promising, they raise two critical issues. The first is that the character of treatment participation is likely to vary across types of child therapies. Consequently, research in this area will benefit from decomposition of the stream of therapy process into specifiable tasks (Greenberg & Newman, 1996). Identification of therapy tasks and their component features will provide an anchor for operational definitions of active participation or therapeutic collaboration.

Second, treatment tasks are likely to vary according to children's developmental level (if therapists are gearing their interventions to children's capacities). Thus, a brand-name treatment, such as cognitive-behavioral therapy, may involve different tasks for children at different ages, or if it is developmentally sensitive the components of specific treatment tasks could vary developmentally, for example, tasks such as self-instructional training might be simplified for younger children.

A second major limitation to research on readiness for therapy is that the conceptual and empirical focus has been restricted to the individual child. One of the distinctive features of child therapy is that it occurs in the context of ongoing relationships with other care-givers, typically parents. Research on parent management training, in which parents are the primary participants, has indicated that maternal depression and antisocial behavior predict initial resistance in therapy (Patterson & Chamberlain, 1994). Yet the contribution of parent or family factors to child treatment participation has been largely unexamined. In an exploratory study, Raney, Shirk, Sarlin, Kaplan, and During (1991) found that adolescent inpatients' perceptions of their parents' collaboration with the treatment team was associated with therapists' ratings of the individual treatment alliance later in treatment. Although preliminary, these results suggest that alliance formation in individual treatment may be influenced by the parents' relationship with the therapist, or at least by the child's or adolescent's perception of that relationship. Such potential effects are worth investigating, as they redirect attention to the broader social context of children's participation in treatment.

A final limitation is the virtual absence of studies of in-session transactions that could account for variation in child participation. Again, research on parent management training has revealed that pretreatment characteristics of parents account for initial levels of resistance, but that the nature of the transactions between parent and therapist predict changes in resistance (negative participation) over the course of treatment (Patterson & Chamberlain, 1994). Early 1970s child therapy research has shown that variation in therapist-offered levels of empathy, regard, and genuineness are associated with child verbal behavior in sessions (Siegel, 1972).

However, this type of research, linking therapist behavior with child involvement, has not developed up to 1997. Similarly, despite interest in child characteristics that predict participation in therapy, studies of therapist characteristics and their impact on child treatment participation are lacking. In the 1970s, Ricks (1974) demonstrated that child therapists trained in the same treatment orientation vary substantially in the way they relate to child patients, and that these differences are associated with differential treatment outcomes. Similarly, Alexander, Barton, Schiavo, and Parsons (1976) found that therapist relationship characteristics, such as warmth and humor, accounted for substantial variance (45%) in treatment outcome in functional family therapy. Despite these promising leads, research has remained focused on the evaluation of treatment packages while neglecting the therapist who delivers the treatments. Overall, then, research in this area should shift its focus from the individual child to encompass both child and therapist, and to the characteristics that define productive and unproductive therapeutic dyads.

5.03.4 IN-SESSION TRANSACTIONS IN CHILD THERAPY

A major focus of psychotherapy process research is on transactions that occur during face-to-face meetings between patient and therapist. Often these transactions are studied in their own right with the aim of discovering what actually comprises "therapy" of one type or another. Less frequently, aspects of transactions are measured and related to outcomes at the session or treatment level. Among the most salient aspects of child therapy transactions are interpersonal and symbolic exchanges between child and therapist. Both have been posited to contribute to variation in treatment outcome (cf. Shirk & Russell, 1996), and both have received some attention from child therapy investigators.

5.03.4.1 Interpersonal Transactions in Child Therapy

Child psychotherapy, regardless of treatment type, occurs in an interpersonal context. For some time, interpersonal processes have figured prominently in the adult psychotherapy literature, and reviews have indicated that characteristics of the relationship between patient and therapist are among the best process predictors of treatment outcome (Horvath & Luborsky, 1993). Although research on interpersonal processes in child psychotherapy is relatively limited, a review of child process reveals an enduring, but thin, line of inquiry on interpersonal processes in child treatment (cf. Russell & Shirk, in press). A review of this literature also indicates that interpersonal and relationship processes have been conceptualized in a number of ways.

Two basic models of the interpersonal processes in child treatment are prominent. The first conceptualizes relationship processes as a means to an end. The quality of the therapeutic relationship, and the transactions that contribute to relationship quality, are viewed as the foundation for the implementation of other (active) technical interventions. In essence, the relationship functions to promote or obstruct collaboration with treatment tasks that are posited as active change mechanisms. Interpersonal transactions that contribute to a positive relationship between child and therapist enable the child to work purposefully on resolving problems. This conceptualization is found in most models of the therapeutic alliance where a distinction is often made between the affective quality of the relationship and the quality of treatment collaboration (cf. Bordin, 1979; Horvath & Luborsky, 1993; Shirk & Saiz, 1992). Thus, despite substantial differences in the nature of treatment tasks, psychodynamic and cognitive-behavioral approaches appear to share this basic conceptualization of the therapeutic relationship. This conceptualization might be termed the indirect model of interpersonal process in child treatment. Although relationship processes are necessary for successful outcome, their influence on outcome is mediated through the child's active participation in treatment tasks. Relationship processes facilitate the operation of technical interventions, ranging from interpretation to practicing consequential thinking.

An alternative conceptualization views relationship processes as active change mechanisms. Thus, rather than positing an indirect relation between interpersonal processes and treatment effects, this perspective holds that interpersonal transactions have an unmediated impact on treatment outcome. This view might be termed the direct model of interpersonal process in child therapy, and takes its clearest form in treatments that emphasize "corrective relationship experiences."

Two child treatment approaches have emphasized the "corrective" nature of therapeutic transactions, but with somewhat different interpretations of the corrective process. The first can be found in client-centered models of child therapy. Here, the provision of a warm, accepting, and nonjudgmental relationship is viewed as a catalyst for self-propelled development (Axline, 1947). The second is found in psychodynamic, especially analytic, models that emphasize the role of transference in treatment process. A version of this perspective holds that children enter therapy with pathogenic beliefs that are "tested" in relation to the therapist (Gibbons & Foreman, 1990). Therapist actions in response to these interpersonal tests can either confirm or disconfirm the underlying pathogenic beliefs that prompt the behavior. The relationship serves a corrective function insofar as therapeutic transactions disconfirm pathogenic expectations. Although corrective processes are conceptualized differently in these two approaches, both assume that interpersonal processes are directly responsible for therapeutic change. It is interesting to note that in the adult literature, relationship processes have been shown to have a direct effect on treatment outcome even among treatments that emphasize technical operations, such as cognitive therapy (Burns & Nolen-Hoeksema, 1992). Thus, relationship processes might exert both direct and indirect effects on child outcomes, even among treatments that focus on collaboration with specific therapeutic tasks.

5.03.4.1.1 Research on interpersonal processes in child therapy

Although research on relationship processes is relatively limited, studies based, often implicitly, on either the direct or indirect model of interpersonal process can be found. One of the clearest examples of research based on the direct model comes from studies on the impact of variations in therapist-offered conditions of warmth, genuineness, and accurate empathy on treatment outcome. For example, Truax, Altman, Wright, and Mitchell (1973) compared the outcomes of a small sample of children described as mildly disturbed who received either relatively high or low levels of therapist-offered relationship conditions during treatment. Levels of therapist warmth, genuineness, and empathy were assessed during two sessions, and conditions were categorized by a median split. Outcomes, largely measures of adjustment, were assessed from multiple perspectives, including parent, therapist, and independent psychometrician. Regarding the impact of high or low levels of therapist conditions, the results were mixed. No group differences were found for psychometrist ratings; however, both therapist and parent ratings of overall change in adjustment showed significant effects for condition. Consistent with expectations, children who received high levels of therapist-offered conditions showed significantly greater benefits than their low-condition counterparts; in fact, children who received low levels of therapist-offered conditions showed deterioration on five times as many measures of specific outcomes as children receiving high levels. The authors interpreted these findings in a manner consistent with the direct model of interpersonal process, namely that relationship quality, or at least the type of relational response offered by the therapist, exerts a significant impact on treatment outcome. The study, however, was compromised in a number of ways. Children were not randomly assigned to therapists who were identified as high or low on the provision of warmth, genuineness, and regard prior to treatment. Consequently, differences in outcome might reflect unequal distributions of recalcitrant children across conditions. Furthermore, given that therapists were classified on the basis of their responses to treated children, difficult children might not have elicited high levels of therapist-offered warmth, genuineness, and empathy. Therapist-offered conditions, then, might be related to outcome, not because of children's differential responsiveness to variation in therapist-offered conditions but because children who are more or less difficult to change differentially elicit these conditions from therapists. Finally, it is important to note that interpersonal process is defined in terms of the behavior of only one of the interactants, that is, in terms of the type of therapist-offered relationship. The quality of the dyadic relationship as it develops between therapist and child, or even as it is experienced by the child, is not considered.

Many of these serious limitations are overcome in Kendall's (1994) randomized clinical trial for the treatment of anxiety disorders in children. As part of this manualized, cognitive-behavioral intervention, children's perceptions of the therapeutic relationship were assessed in order to examine possible relations with outcome. Items on the scale referred to children's feelings for the therapist and perceptions of the quality of their relationship with the therapist. In this controlled outcome study there was some, albeit limited, evidence for associations between perceived relationship quality and indices of treatment outcome. Approximately 30% of correlations between relationship quality and outcome were significant in the predicted direction; however, the majority of these involved shared source variance (child self-reports). The strongest pattern of results was between perceived relationship quality and child reports of changes in symptoms of anxiety and depression, with significant less improvement for those cases reporting poorer relationships than those reporting more positive relationships. Although the study did not provide strong evidence for the role of relationship factors, the tendency for most children to report positive relationships with their therapists may have limited variability and attenuated relationships with outcome.

Clearly, the study represents a significant advance over earlier work on relations between interpersonal process and outcome. Technical aspects of the treatment were controlled, children were randomly assigned to therapists who delivered the "same" treatment, and the quality of the relationship, rather than behaviors of one participant were assessed. With respect to interpersonal processes, however, the study is limited in several ways. The quality of the relationship was assessed from only one perspective which yielded limited variability, and relationship quality was assessed at only one point in time, at the end of treatment. Thus, it is possible that children's experience of symptom reduction could have contributed to their evaluation of the therapeutic relationship.

A study by Eltz et al. (1995) examined relationships among maltreatment experience, therapeutic relationship formation, and treatment outcome in a heterogeneous sample of psychiatrically hospitalized adolescents. Like the previous studies, this investigation is consistent with the direct model of interpersonal process; however, two advances are noteworthy. First, the quality of the therapeutic relationship was assessed from both therapists' and adolescents' perspectives; and, second, the relationship was measured at two points in time during the course of treatment. The latter enabled the investigators to examine patterns of relationship formation in a rudimentary manner. The study revealed that adolescents who showed more positive change in relationship quality, regardless of maltreatment status, evinced greater treatment gains across sources of outcome ratings relative to those showing less positive patterns of relationship development. These results suggest that changes in the pattern of relationship quality may be a better predictor of outcome than "one-shot" assessments. The more consistent pattern of associations between relationship quality and outcome in this study compared to Kendall's (1994) trial might reflect important differences in the types of children treated or in types of treatments delivered. In the latter study, patients were seriously disordered and showed marked variability in their capacity to form a therapeutic relationship. Furthermore, inpatient psychotherapy was not manualized, but instead focused on processing relationship issues. In this context, relationship quality may be the primary active ingredient of therapy, in contrast to treatments that rely on compliance with specific therapeutic tasks. Of course, the lack of control over the delivery of therapy raises the possibility that the association between relationship quality and outcome are confounded with unexamined technical differences across treatments.

In summary, these three studies provide relatively clear examples of existing research on interpersonal processes in child therapy based on the direct model. According to this perspective, the quality of the therapeutic relationship is assumed to have a direct bearing on variations in treatment outcome. Although progress can be detected across these studies, evidence for a direct relation between relationship quality and outcome in child treatment is rather limited.

Research on the indirect model of interpersonal process in child therapy is sparse. According to this perspective, relationship quality impacts outcome by promoting or obstructing treatment collaboration. Drawing on the work of Bordin (1979) and Luborsky (1976), Shirk and Saiz (1992) developed a measure of the therapeutic alliance for child therapy that distinguished between the affective quality of the relationship and collaboration with treatment tasks. Consistent with the indirect model, Shirk and Saiz hypothesized that a positive relationship between child and therapist would promote collaboration with therapy tasks. In a study with a diagnostically heterogeneous, hospitalized sample of 7–12 year old children who were seen in short-term dynamically oriented therapy, children and their therapists rated both dimensions of the working alliance. Reliabilities for both therapist and child ratings ranged from acceptable to excellent. Moreover, child and therapist perspectives on the affective quality of the relationship were moderately correlated. For both therapist and child reports, the affective quality of the relationship was positively associated with task collaboration, in this case, verbalization of feelings and problems. The results were consistent with the prediction that the affective quality of the relationship potentiates task collaboration. However, inferences of directionality are limited by the fact that the alliance was assessed at only one point in time; thus it is possible that children who establish a collaborative relationship feel positively about their relationship with their therapist, and are viewed by their therapist in a positive light. Confidence in the hypothesized relation between affective quality and task collaboration would be increased through the assessment of the alliance at multiple points in time, and by supplementing participant reports of task collaboration with independent observational ratings of involvement in specific therapeutic tasks. The major limitation of this study, however, was the absence of data on the relation between alliance and outcome; consequently, neither direct nor indirect effects could be assessed.

Along similar lines, Estrada et al. (1994) developed observational scales to assess dimensions of participation in child therapy. Factor analyses of 117 session segments revealed 3 therapist and 3 child factors. One of the child factors and one of the therapist factors concerned the quality of the therapeutic relationship. Items such as empathy, warmth, and positive relationship loaded on the therapist factor; and in a complementary manner items such as trust and openness loaded highly on the child relationship factor. There was significant covariation between independent ratings of child and therapist relationship quality. In addition, factors emerged for both child and therapist ratings that referred to involvement in the tasks of therapy, what Estrada et al. labeled therapeutic work. Again, child and therapist ratings were significantly correlated. Furthermore, in keeping with the indirect model, there was evidence to suggest that the quality of the relationship is associated with involvement in therapeutic work. Children with more positive relationships evinced higher levels of work than children who were rated as having less positive relationships. Moreover, ratings of technical lapses by the therapist, such as low responsiveness or inattentiveness, were negatively associated with the quality of the child's relationship with the therapist. These results suggest that relationship quality and task collaboration represent distinguishable aspects of the working alliance with children, and that they are associated in predictable ways. However, the directionality of this association remains ambiguous because process factors were assessed at the same point in time. In fact, the finding that therapist technical lapses are associated with the child's relationship factor is suggestive of an alternative account, namely that relationship quality is a function of the therapist's technical skill. For example, therapists who are adept at introducing treatment tasks and stimulating participation may be viewed quite positively by their child patients. Finally, like the Shirk and Saiz (1992) study, this investigation was aimed at scale development and not at evaluating process–outcome relations. As a result, the indirect model of interpersonal process has yet to be tested.

Other investigators have assessed aspects of the therapeutic relationship and/or involvement in treatment tasks. As mentioned, Braswell et al. (1985) have operationalized involvement in terms of child and therapist verbal behavior, but they did not examine this characteristic in connection with relationship quality. Smith-Acuna, Durlak, and Kaspar (1991) assessed both child and therapist experiences of therapy and found reliable dimensions corresponding to affective quality and interactive behavior; however, it was not clear that the latter dimension reflected therapeutic work. Moreover, no significant relations were found between therapist and child ratings of either dimension; however, associations within perspective were not reported. Finally, Patterson and his colleagues (Chamberlain, Patterson, Reid, Kavanagh, & Forgatch, 1984; Patterson & Chamberlain, 1994) have developed a microanalytic coding system to assess cooperative or resistant behavior by parents in parent management training. Although this line of research has yielded important findings on factors that influence resistance, and on the relation between resistance and outcome, as is discussed later in this chapter, quality of relationship is subsumed under resistant or cooperative behavior. As such, the parents' feelings about the therapist are not viewed as a separate component of the treatment relationship, although therapist behaviors that are often viewed as facilitators of relationship quality are examined separately.

In summary, research examining either the direct or indirect models of interpersonal process in child treatment is relatively sparse. One of the most glaring limitations in this area of investigation is the virtual absence of studies that have adequately assessed both process and outcome. Given the clinical emphasis on the importance of the therapeutic relationship for successful outcome across multiple types of treatments, research that addresses multiple dimensions of the relationship at multiple points in treatment, and relates these dimensions, not just with each other but with treatment outcome, is desperately needed.

5.03.4.2 Process Issues in Interpersonal Transactions

In addition to the virtual absence of research linking interpersonal processes with treatment outcome, a number of other limitations characterize this line of investigation. First, there has been a tendency to treat the therapeutic relationship as a static entity. Most studies evaluate the relationship at one point in time and fail to assess variations in relationship patterns over the course of treatment. There have been some noteworthy exceptions. For example, Siegel (1972) examined changes in child verbalizations as a function of high and low levels of therapist-offered conditions across multiple sessions of dynamically oriented therapy. But, like most studies in this area, relations with outcome were untested.

Second, almost no studies have conceptualized or assessed interpersonal processes as transactions consisting of linked sequences of exchanges between child and therapist. Instead, the focus has been on child or therapist experience of the relationship, an important aspect of interpersonal process, or on ratings of individual behavior. Interestingly, one of the earliest studies of child therapy process, conducted by Moustakas and Schalock (1955), examined contingencies, albeit in rudimentary way, between child and therapist behaviors. This approach to interpersonal process in child treatment remains underdeveloped.

Finally, child process investigators have largely failed to identify critical interpersonal change events or sequences that could be predictive of outcome. For example, sequences such as the "rupture and repair" of the therapeutic alliance (Safran & Muran, 1996, p. 447) or the resolution of therapeutic impasses (Diamond & Liddle, 1996) have been postulated as critical events that contribute to outcome. Close analyses of the interpersonal transactions that constitute these events could have important implications for the conduct of therapy.

5.03.4.2.1 Future directions for research

Programmatic research by Patterson and his colleagues on resistance in parent management training addresses many of the major limitations found in process studies of child-centered therapy, including the explication of a specific process model, the examination of interpersonal exchanges over multiple sessions, and the evaluation of process–outcome relations. Although this work focuses on parental resistance, and not on interactions between therapist and child, these studies constitute a research strategy that could be applied to interactional processes in child therapy. Among the many strengths found in this line of research are (i) the operationalization of a pivotal clinical construct, (ii) the development of a reliable coding system for tracking different types of resistant behavior in therapy, (iii) the systematic analysis of patterns of resistance over the course of therapy, (iv) the investigation of antecedents of resistance, including both client pretreatment characteristics and therapist actions, and (v) the analysis of the impact of resistance on treatment outcome. The approach also sets the stage for the study of therapeutic transactions, that is, patterns of interdependencies between therapist and client behavior over the course of treatment. Because this line of investigation goes beyond the scope of many studies of child therapy process, and could serve as a prototype for investigations in this area, this work will be reviewed in some detail.

As Patterson and Chamberlain (1994) note, resistance to change is a phenomenon acknowledged by virtually all the major therapeutic systems, yet relatively little empirical research has been conducted on this pivotal clinical problem. One difficulty has stemmed from the lack of a consensual definition of resistance and a corresponding method of operationalizing the concept. In many ways, resistance can be understood most simply as the inverse of the collaborative aspect of the therapeutic alliance. High levels of collaboration indicate that the client is working with the therapist on the specified tasks of therapy. Resistance, then, reflects the absence of this collaborative activity. This framework appears to be consistent with the conceptualization presented by Patterson and colleagues. Parent behavior in therapy is conceptualized as either cooperative or resistant. The former reflects a pattern of behavior that follows the direction set by the therapist, whereas the latter involves behavior that is oppositional, helpless, or tangential to the therapeutic task. Based on observations of treatment failures, five types of resistant responses were identified and reliably coded as part of the Client Resistance Coding System (Chamberlain et al., 1984). These included such behaviors as talking over or cutting off the therapist, expressing an unwillingness or an inability to follow the therapist's suggestions, challenging the therapist's qualifications, introducing new topics to shift the session away from the therapist's direction, and not tracking the therapist's verbalizations. Thus, unlike most of the previously reviewed studies, this approach to interpersonal process employed an observational system with well-defined behavioral markers. It is noteworthy that these markers of interpersonal process are most readily identified through parental language behavior. Given developmental differences in language use between children and adolescents, it is not clear if these markers would be as evident in child therapy.

Several studies have indicated that resistant behaviors in parent training sessions can be reliably coded (Chamberlain et al., 1984; Patterson & Forgatch, 1985), and that they cohere as a meaningful index of parental resistance. Correlational analyses suggest that there are two clusters of parental resistance, one characterized by challenging responses and the other by helpless responses, or what Patterson and Chamberlain (1994, p. 59) have dubbed, "I won't" or "I can't" patterns of resistance.

One advantage of conceptualizing resistance at the level of molecular events is that it allows for an analysis of patterns of resistance over the course of a session or across an entire treatment. It was hypothesized that a prototypical pattern of resistance would emerge in successful parent training cases (Chamberlain et al., 1984; Patterson & Forgatch, 1985). Resistance was expected to be low in the initial phase of treatment as the therapist attempts to build rapport and gather information about the presenting problem. Movement into the middle phase of treatment was expected to be marked by a rise in resistance as the therapist focused on parenting strategies and recommended changes in parenting behavior. During the final phase of treatment it was expected that resistance would decline as the therapist focused on consolidation of acquired skills. In essence, an explicit process model, labeled the "struggle hypothesis" (Patterson & Chamberlain, 1994), was formulated and tested.

Research by Chamberlain et al. (1984) confirmed this pattern, with parents showing the highest levels of resistance in the middle phase of treatment. Cases rated as more successful showed the expected decline in resistance between the middle and late phases of treatment. Parents who dropped out of treatment evinced higher levels of initial resistance to therapy than those that continued; however, among those completing therapy, initial levels of resistance were unrelated to ratings of treatment success. It should be noted that an important limitation of this study was the use of therapist ratings of success as the sole index of treatment outcome (although such ratings have subsequently been shown to be related to changes in child behavior). Nevertheless, the results of this study provide the foundation for two important hypotheses. First, the findings suggest that the pattern of resistance over the course of treatment may be a better predictor of treatment outcome than mean levels of resistance. Second, the results suggest that parental resistance is not exclusively the function of individual dispositional differences, but may be closely related to variations in therapist behavior. Both hypotheses represent significant departures from the type of questions typically addressed in studies of child therapy process. The former draws attention to patterns of change and the shape of trajectories over the course of treatment, and the latter underscores the importance of investigating process at the level of client–therapist exchanges where interpersonal interdependencies can be uncovered.

In an effort to understand the impact of therapist behavior on parent resistance, Patterson and Forgatch (1985) examined the covariation in therapist and client reaction sequences. They hypothesized that efforts to produce change would have an unintended "paradoxical" effect on clients by increasing levels of noncompliance. In the first study involving six therapist–client pairs, they predicted that therapist efforts to teach and confront would be associated with the greater conditional likelihood of resistant behavior compared to base rate levels. Conversely, therapist supportive and facilitative behaviors were expected to result in lower conditional probabilities of noncompliance relative to base rate levels. In order to evaluate covariation between therapist and client behavior, it was necessary to develop a therapist behavior coding system to complement the client codes. The Therapist Behavior Code includes seven mutually exclusive categories such as support, teach, and confront. Results indicated that in four of the six pairs, therapists' attempts to teach was associated with significantly higher than base rate levels of resistance, and in six of six pairs, confrontation, described as a "more intrusive effort to re-educate," were associated with higher levels than base rate. Supportive and facilitative therapist behaviors were consistently associated with lower than base rate probabilities of resistant behavior. As Patterson and Forgatch (1985) note, despite the fact that therapist behavior preceded the parent reaction, causality could not be inferred because of the possible operation of an unidentified third variable.

In order to address the causal issue, Patterson and Forgatch (1985) conducted an experimental manipulation of therapist behavior in a second study. Utilizing an ABAB design, they showed that systematic manipulation of therapist "teach" and "confront" behaviors resulted in corresponding changes in client resistant behavior. Increases in teaching and confronting were matched by increases in resistant behavior in B1 and B2, and a reduction of teaching and confronting in A2 produced a return to baseline in resistant behaviors. Taken together these results provide support for the hypothesis that variation in parental resistance is not merely a function of dispositional differences across parents, but is linked to what the therapist does during the session. This is not to say that pretreatment characteristics are unrelated to parental resistance. In fact, Patterson and Chamberlain (1994) report that stress, depression, and parental antisocial behavior contribute to initial levels of maternal resistance even after partialling out indices of social disadvantage. It is noteworthy that changes in level of resistance are not predicted by pretreatment characteristics, instead the pattern of resistance over the course of treatment appears to be determined by interactional processes in sessions.

Finally, evidence indicates that the pattern of resistance over the course of treatment is strongly predictive of outcome (Stoolmiller, Duncan, Bank, & Patterson, 1993). Patterson and Chamberlain (1994, p. 56) hypothesized that the prototypical pattern of resistance in successful cases follows and inverted V function, and have dubbed this pattern the "struggle hypothesis." Cases that show a rise in resistance during the instructional middle phase followed by a decline toward the baseline in the latter phase of treatment were expected to show positive outcomes. Utilizing growth curve analysis, Stoolmiller et al. showed that lack of negative curvature in resistance curves predicted child arrests over a 2 year period following treatment. In contrast, the more conventional approach to assessing process by measuring resistance early and late in treatment and then computing a change score failed to predict child treatment outcome. As Stoolmiller et al. (1993, p. 927) note, "what goes on in the therapy process appears to have better predictive validity than the assessed therapy endpoint or overall change from therapy initiation to therapy termination."

In summary, this line of process research provides a blueprint for future investigations of interpersonal processes in child therapy. Although such studies might examine other process dimensions, for example, patterns of emotion-regulating interactions, emulation of the programmatic character of research on parental resistance would increase the likelihood of useful results.

5.03.4.3 Symbolic Transactions in Child Psychotherapy

Much of what has been said about interpersonal transactions in child psychotherapy applies to symbolic transactions as well. In fact, strictly speaking, the symbolic character of communication processes *per se* suggests that interpersonal and other types of meaningful exchanges in child therapy are properly considered a subtype of symbolic transactions. Clearly, discourse, nonverbal gestures, and play are inherently symbolic, regardless of what their topical focus may be, and regardless of how they may differ in terms of their constitutive communication channel (e.g., verbal vs. nonverbal), their degree of rule-governedness, their type of referentiality (e.g., referring to actual vs. pretend worlds), and their degree of consequentiality (e.g., literal vs. play actions). In brief, child psychotherapy is a form of symbolic interaction, and to understand its processes and how they relate to outcomes, requires an understanding of what humans do most characteristically, namely communicate vis-à-vis symbolic exchange.

Although the above may seem painfully rudimentary, the field of psychotherapy research, child and adult alike, has, in the main, skirted the difficult tasks of studying symbolic processes during the therapeutic hour and of relating them to outcomes, often themselves different types of symbolic processes (e.g., social skills, academic performance, self-presentation, etc.). This is odd indeed if it is true that symbolic construals of self, other, and the world significantly determine conduct, coalitions, and feelings. Who we are and how we adjust may boil down to what we communicate in our ongoing discourses with others. It should come as no surprise that our status as "*Homo symbolicus*" or "*Homo narrans*" exacts such a price, or, rather, confers such an awesome privilege. The history of research on child psychotherapy seems to have avoided dealing systematically with the consequences of this anthropological truism.

At a simpler level, the lack of a systematically developed research tradition on symbolic processes in child psychotherapy is no less surprising. Experiential, psychodynamic, and cognitive–behavioral theories of treatment each stress aspects of verbalization or symbolization (i.e., what is termed "discourse" in the more current idioms) as constitutive of curative processes. According to experiential theory, symbolizing warded-off prereflective experiences, helps to integrate the client's self, or alternatively, reintroduces the client to his or her experience and to the complete sphere of his or her phenomenal world. In early psychodynamic theory, fully and completely describing early trauma in affect-laden language was thought to serve a cathartic function, just as in later psychodynamic theory, redescribing one's autobiographical experience in action language is thought to promote self-transformation and empowerment. In cognitive theory, rescripting event expectations or transforming narrative schemas of important autobiographical episodes corrects faulty causal theories of the world and appraisals of self-worth and blame. Cognitive–behavioral approaches often focus on self-directed language in the form of self-instructions or verbal problem-solving strategies. As central as symbolic processes have been in the theories that have informed child therapy practice, empirical research on them has been peripheral and sporadic.

However, there have been isolated attempts to study symbolic processes in child psychotherapy. These attempts can be conveniently described as exemplifying two very different research paradigms. These can be dubbed the monological and the dialogical paradigms. Basically, the monological paradigm promotes research strategies that tend to excise the individual's or dyad's symbolic communication from the temporal parameters of discourse and situations in which they are embedded in order to characterize therapeutic communication in terms of its internal features. The dialogical paradigm promotes research strategies that tend to contextualize the individual's or dyad's symbolic communication in the temporal flow of discourse and situations in which they are embedded in order to characterize communication in terms of its relational features.

By noting whether a study focuses on an individual's or dyad's symbolic communication (i.e., its speaker scope), and whether it uses monological or dialogical research strategies, a 2 × 2 matrix describing four methods for studying symbolic processes in child psychotherapy is devised. Rudimentary examples of two of these four investigative strategies were contained in the very first study of the process of child therapy. In it, Landisberg and Snyder (1946) reported the frequencies of certain types of child and therapist speech acts. For example, 30% of the utterances in the sessions sampled were therapist nondirective statements. Here, the focus is on an individual, the therapist, and aspects of his discourse have been excerpted from the flow of conversation and tallied in relative isolation from the context in which it had been embedded. These authors also reported dependent relationships between aspects of therapist and client discourse. For example, they reported that therapist reflection of feelings were most often followed by client giving of information or taking some positive action. Here the focus is on the interactive dyad, and aspects of the therapist's and client's discourse are described in relation to each other. The same strategy was employed by Moustakas and Schalock (1955), but they described dependencies running from therapist to child discourse and from child to therapist discourse. The history of research in child and adult symbolic processes in therapy, however, has contained far more studies influenced by the former, monological strategy than the latter, dialogical strategy.

Note that the studies described above are exclusively focused on describing process and relations in it. They do not relate the frequencies of, or the extent of certain types of dependencies between, aspects of therapist or client discourse to outcomes. Truax et al. (1973) provide an example of the use of the monological approach to relate process to outcomes. They reported that children who received the highest levels of therapist-offered relationship conditions evidenced more improvement on a host of outcome measures than children receiving the lowest levels of therapist conditions. This study would seem to support the idea that there is a dose–response relationship between therapist conditions and child outcomes, although the findings are ambiguous insofar as the therapists' behavior could have resulted from variation in child responsiveness (see Russell (1994) for a discussion of this relationship).

The other two cells of the 4 × 4 matrix represent two kinds of hybrids of the monological and dialogical approach. For example, Truax and Wittmer (1973) tabulated the degree to which a certain type of therapist–client exchange occurred across sessions of child treatment, and related these frequencies to outcomes. Basically, the exchange that they tabulated consisted of pairs of child defense–therapist confrontation. The study is focused on aspects of the dyadic exchange (a feature common to the dialogical approach), but exchanges are excerpted from the temporal flow of interactive discourse, and their frequency related to outcomes (a feature common to the monological approach). Interestingly, the authors reported that those clients whose defenses were most often confronted achieved significantly better outcomes than those whose defenses were less often confronted.

In the 1980s and 1990s the other hybrid began to emerge in process studies. For example, Mook (1982a, 1982b) and Russell et al. (1996) applied methods (the P technique, a form of factor analysis applied to within-subject data matrices) that enabled them to describe the underlying structure of covariation of sets of variables used to rate either the therapist's or client's utterances. Here the focus is relational and contextual, as it describes how sets of utterance characteristics (e.g., degrees of information seeking, interpretation, confrontation, empathy) covary over the temporal course of sessions (a feature common to the dialogical approach). In these studies the context is restricted to the corpus of therapist or client utterances, not the interaction between both (a feature common to the monological approach). Together, these studies suggested that therapist participation in treatment discourse is structured around at least two underlying tasks—information seeking alternating with attentive listening, and the application of therapist techniques that help to sustain topical focus across utterances.

The studies reviewed above are minimal instantiations of the four types of process studies defined by crossing methodological strategy (monological vs. dialogical) and speaker scope (one or more dimensions). More sophisticated instantiations have begun to appear. For example, in Shirk and Russell (1996), a study by Estrada, Russell and colleagues is described in which the P technique is used to examine the underlying interactive structure of child therapist and client discourse. Here the focus is on context and relation as incorporated in the P technique, but now both child therapist and child client discourse are entered into the analyses, thus representing the first fully dialogical study undertaken in child therapy. The authors compared high- with low-quality child therapy sessions, and found that the high-quality sessions contained more P-technique factors that indicated systematic conversational exchange between therapist and client across at least two turns of talk than the low-quality sessions.

Obviously, there are other types of studies that can be conducted on symbolic processes in child therapy. For example, Russell et al. (1991) attempted to describe types of therapist discourse taken in their totality. Three types were sufficient to describe the bulk of treatment studies. Therapist discourse tended to be either conversational, instructional, or behavioral as in operant or counterconditioning treatments. The intent of this meta-analytic study was to relate these types of discourse on symbolic tasks such as reading, spelling, and verbal interaction. The authors reported that the effect size associated with conversational discourse was over 10 times as large as that associated with instructional discourse and over four times as large as that associated with behavioral discourse. In addition, there was a relatively strong dose–response relationship between the number of sessions of treatment and the effect size associated with language outcome variables.

The conclusions of Russell et al. (1991, p. 918) to their meta-analysis are worth repeating: (i) "A potentially rich and clinically significant area of child functioning has been only sporadically studied (i.e., only about 20 out of 105 outcome studies contained a language measure)"; (ii) "Ratings or categorization of the *in vivo* therapist and child language interaction are [nearly] uniformly missing"; and (iii) "This domain of functioning seems sensitive to the effects of child treatment, providing an additional, nonreactive dimension for the assessment of outcome." As we have seen, since 1991, there have been both calls for more child process research (Kazdin, 1995) and several attempts to devise and apply dialogically oriented research strategies (e.g., Russell et al., 1996). The field is ripe to more fully plumb the complexities of symbolic interaction in child treatments.

5.03.4.3.1 Future directions for research

Several directions for future research seem worth pursuing. One is suggested by trends in the child process literature itself, and is reinforced by similar but earlier achieved advances in the adult process research. Over the course research since the 1940s, there has been a trend, in child and in adult process research, away from disproportionate reliance on monological research strategies and toward the development of more dialogical research strategies. These developments should continue as we better understand the bi- or multi-directional causality underlying the patterning of human interaction. As suggested in Section 5.03.4.2.1 in interpersonal processes, changes in the patternings of child and therapist discourse may be the best process predictors of treatment progress and outcomes. Second, there has been a dearth of process and outcome measures cut from the same conceptual cloth. Symbolic processes, evident everywhere in human interaction, can provide a rich resource for defining process and outcome variables that share considerable family resemblances. For example, despite the importance of language competence in skillful social behavior, problem-solving, and self-regulation, few studies measure language change, such as pragmatic competence, as an outcome of child treatment. Third, when the symbolic character of child therapy is taken seriously, its study will require familiarization not only with advanced quantitative and qualitative techniques of data analysis, but with theories of discourse and language analysis that have been developed in other disciplines. Thus, we recommend cross-disciplinary training for process researchers. Finally, following developments in the adult process literature, we recommend intensive study of sets of change episodes rather than random samples or whole therapy sessions. Such a focus may enable us to draw a finer bead on the mechanisms responsible for therapeutic change.

5.03.5 THERAPEUTIC TRANSACTIONS AND TREATMENT OUTCOMES

As Orlinsky et al. (1994, p. 270) have observed, the investigation of process–outcome relations is guided by the core question "What is effectively therapeutic about psychotherapy?" Given the relevance of this question for the practice of psychotherapy, one would assume that the investigation of process–outcome relations would represent one of the most prominent strategies among psychotherapy researchers. In fact, in the adult psychotherapy literature there has been rapid growth in this approach to treatment research. In their review of the process–outcome literature, Orlinsky and colleagues uncovered over 2300 findings, with more than half of them appearing since 1985. Studies of the therapeutic relationship alone accounted for nearly 1000 findings. Most importantly, process–outcome investigations have yielded promising results; Orlinsky et al. identified 11 robust process–outcome links, including global therapeutic bond, patient cooperation vs. resistance, and patient expressiveness.

In stark contrast to the adult literature, remarkably few studies of child psychotherapy have examined process–outcome relations. The magnitude of discrepancy cannot be overstated. In their review of 223 child treatment studies, Kazdin et al. (1990) identified only 6 that examined treatment process in relation to outcome! Our own review (Russell & Shirk, in press) revealed some growth in process–outcome research during the 1990s, but even a doubling of the number of studies represents a small effort in absolute terms. Not surprisingly, then, "robust" findings, or, for that matter, replicated findings, remain elusive.

How is the paucity of research on process–outcome relations in child psychotherapy to be accounted for? Surely child clinicians and researchers are no less interested than their adult counterparts in what is "effectively therapeutic." One clue to this remarkable gap can be found in what has been done by child therapy researchers. As Kazdin (1994, p. 577) has noted, much of child therapy research has utilized either a "treatment package strategy" or a "dismantling strategy." In the former, effects of whole treatments are evaluated, and in the latter, treatments are dissected and components are varied across groups. Common to both of these strategies is the assumption that broad technical procedures, for example, social skills training or problem-solving training, constitute the "effectively therapeutic" components of treatment. Thus, by examining the presence or absence of these components in relation to outcome, either by use of a control or dismantling comparison, one attains some perspective on treatment process, at least at a molar level. Although this is a reasonable starting point, the limitations of this perspective are worth mentioning.

First, even if we assume that the active ingredients of therapy are technical procedures—be they problem-solving training or mutual story-telling exchanges—such procedures can be instantiated in treatment in a variety of ways. For example, problem-solving training typically involves some combination of didactic, modeling, and role play activities with varying degrees of emphasis on hypothetical or "live" situations. Moreover, research has shown that the "same" treatment procedure or component delivered in alternative ways can have a differential effect on treatment outcome. For example, Schleser, Meyers, and Cohen (1981) varied the manner in which self-instructional training was presented to children with self-control problems.

In the traditional condition, the therapist provided the child with direct examples of self-instructions, such as, "What is it that I need to do"; in the directed-discovery condition, children were induced to discover basic self-instructions through dialogue with the therapist. For example, the therapist might ask "What is the first thing you need to ask yourself when you start a task?" Variation in this dimension of process, in this case discourse patterns, resulted in differences in skill generalization among more mature children. Although all treated children received the same treatment component, self-instructional training, the manner in which it was presented had an impact on outcome.

Alternatively, other investigators have found that variation in presentation of specific treatment procedures do not produce differential outcomes. Instead, what matters is the quality of the child's involvement with the treatment procedure, regardless of how it is presented. For example, Milos and Reiss (1982) assigned children who evinced high levels of separation anxiety to one of three thematic play conditions: free play, directed play, and modeled play. All three conditions produced comparable reductions in anxiety despite substantial variation in the nature of play interactions, that is, despite differences in the instantiation of thematic play as a treatment technique. However, variation in outcome was related to quality of treated children's play. Children who benefitted most displayed high levels of mastery of oriented play, regardless of technical variations in play interventions. Thus, what was "effectively therapeutic" could not be attributed to specific play techniques but to variations in children's play activity.

The focus on treatment packages or dismantled components of whole therapies restricts the search for effective ingredients of therapy in a second way. As mentioned, these approaches share the assumption that technical procedures are primarily responsible for therapeutic change. At one level of analysis, such an assumption seems to be warranted. Meta-analytic findings suggest that behavioral treatments produce larger treatment effects than nonbehavioral treatments (Weisz et al., 1995), and most certainly these two broad classes of treatment entail substantially different types of technical interventions. However, comparisons of treatments at a more molecular level, for example, comparisons of different types of behavioral treatments such as parent training vs. cognitive-behavioral treatments, do not reveal differential treatment effects. Thus, despite important differences in treatment techniques across subtypes of child behavior therapy, variations in treatment effectiveness are not clearly evident. Such equivalence raises the possibility that other processes which are common across technically diverse treatments could have a significant impact on treatment outcome.

Research on treatment of adult depression is an interesting case in point. Here the central focus has been on comparisons of treatment packages, for example, cognitive-behavioral vs. interpersonal therapy. Despite important differences in component techniques, several treatments appear to be equally effective (Robinson, Berman, & Neimeyer, 1990). Moreover, emerging evidence suggests that other processes, aside from technical procedures, contribute to recovery from depression. For example, analyses of outcomes for depressed adults treated with cognitive-behavioral methods have revealed that therapeutic empathy has a moderate to large effect on recovery, even after controlling for initial level of depression and technical factors (Burns & Nolen-Hoeksema, 1992). Our point is not that relationship processes are the active ingredient of psychotherapy, but rather that such processes are often overlooked when researchers narrowly focus on the evaluation or comparison of treatment packages. As this example suggests, important change processes can be detected by an analysis of process–outcome relations within a well-defined form of therapy. Our understanding of "what is effectively therapeutic" in child therapy will be advanced by broadening the scope of inquiry to include other, "nontechnical" dimensions of therapy process.

5.03.5.1 Promising Directions for Process–Outcome Research

Few studies have actually examined treatment processes in relation to various outcomes, at either the treatment or session level. The absence of findings makes it tempting to turn to the adult process–outcome literature for direction. As Kazdin (1995) has noted, one virtue of such an approach is that continuities and discontinuities between adult and child therapy could be investigated. Alternatively, one could adopt a developmental perspective and consider some of the features of child therapy that represent unique challenges for practicing child clinicians. Thus, analogous to the change events paradigm of process research (Greenberg, 1991), in which predefined change events are examined as an alternative to sampling the full flow of events during sessions, aspects of therapy could be selected because of their developmental significance in child treatment.

One of the most developmentally salient, and at times most problematic, aspects of child therapy process involves the referral process. Children rarely refer themselves for treatment, and often do not acknowledge stress or distress. Consequently, it is likely that children and adolescents will vary considerably in their interest in and motivation for treatment. Variation in motivation could have direct effects on treatment outcome, or could influence outcome indirectly by affecting collaboration with treatment tasks. In one of the few studies of process–outcome relations in child therapy, Braswell et al. (1985) found that verbal markers of active involvement in treatment were predictive of gains in cognitive-behavioral therapy. Thus, one area that merits further investigation is the relation between treatment collaboration/resistance and treatment outcome with children.

Although treatment collaboration/resistance may be related to various child pretreatment characteristics, it is also likely that patterns of therapist–child interaction contribute to treatment involvement. One area that has been ignored by researchers involves the establishment of treatment goals with child patients. Such goals are often defined by parents, teachers, or others who are affected by the child's difficulties. But to the degree that the child is to be an active participant in treatment, a shared definition of a problem (or situation to be changed) may be critical for treatment involvement. The early exchanges that contribute to the definition of goals for therapy represent a developmentally critical aspect of child therapy process, and a therapy event that is likely to vary substantially across children at different developmental levels. Language processes are likely to be pivotal in this aspect of therapy. For example, discrepancies between child and therapist narratives about presenting problems could influence the development of a collaborative treatment relationship.

A related issue involves the role of parents in child therapy. The dependency of children on adult care-givers carries a number of important implications for process–outcome research. First, parents make decisions about initiation and continuation in treatment. There are few consistent findings on factors affecting attrition from child therapy (Armbruster & Kazdin, 1994), but research indicates that a substantial portion of parents elect not to continue after an initial evaluation (Weisz, Weiss, & Langmeyer, 1987). One possibility is that continuation in child therapy may be predicted from the strength of alliance the therapist forms, not just with the child, but with the parents as well.

If this is the case, collateral parent consultation that keeps the parents informed and involved may be critical for effective child treatment. Relatively little is known about the type of parental involvement that could facilitate child treatment.

A second issue involves the role of ongoing parent–child interactions in relation to in-session interactions. For example, therapists may attempt to facilitate emotion expression in the context of therapy, only to have such efforts counteracted by sanctions against expressive behavior in the family. From a treatment perspective, collateral work with the family would be clearly indicated. From a process perspective, the link between in-session processes and treatment outcomes is likely to be moderated by parental acceptance and support of therapeutic activities. It is possible that treatment acceptability (Kazdin, French, & Sherick, 1981) will exert a significant influence on the strength of process–outcome relations. Similarly, variation in parental participation in programs aimed at generalization of in-session activities will undoubtedly influence process–outcome relations. Thus, the study of process–outcome relations in child therapy will need to be contextualized in order to account for the impact of facilitative or corrosive parent–child interactions.

In summary, process–outcome relations in child therapy remains a relatively uncharted area. Although various aspects of process have been investigated in child therapy (cf. Shirk & Russell, 1996), remarkably few studies have related these processes to treatment outcomes. Consequently, this area of investigation has yet to find a meaningful starting point. It is our recommendation that investigators consider developmentally salient process issues, such as interactions aimed at defining mutual treatment goals, as they begin to examine process–outcome relations.

5.03.6 CONCLUSIONS

A review of child therapy process research reveals some promising leads but many substantial gaps in the existing literature. Most striking is the dearth of research linking treatment processes with treatment outcomes. Linkages among referral processes, treatment preparations, treatment transactions, and outcomes are virtually unexplored. Although one might expect such studies to be the centerpiece of process research, their absence undermines our ability to answer the core question: "What is effectively therapeutic about child psychotherapy?" It is the contention that progress in the field hinges on a shift in research focus to an examination of therapeutic processes in relation to treatment outcomes. In order for such a shift to be fruitful, however, a number of changes must occur in the conduct of child process research.

First, the notion of therapy process implies a temporal dimension. Static operationalizations of process fail to capture the flow of therapy. "One-shot" assessments of critical processes, such as the therapeutic alliance, are not as likely to contribute to our understanding treatment process as patterns of processes across sessions or whole treatments. Furthermore, the investigation of transactions, that is, sequences of exchange that occur over time, should complement detailed characterizations of process. Such research holds promise for identifying specific patterns of interaction that constitute broad constructs such as the alliance or treatment involvement. Furthermore, the identification of such patterns could provide useful information to practicing child clinicians.

Second, process research itself has had a restricted focus. Most studies have examined the activities of one individual in isolation. Child process research must redirect attention to therapeutic exchanges and contingencies between child and therapist behavior. Ironically, few studies have examined interaction *per se*, the hallmark of therapeutic process. Instead, frequencies of individual behavior are typically tabulated without reference to dependencies between behaviors. Alternatively, few efforts have been made to examine process at the level of the dyad. Just as a secure attachment describes a type of relationship, shifting the unit of analysis to the dyad could reveal process characteristics that are not captured at the individual level. Finally, relatively little attention has been directed toward the broader social context in which child therapy process is embedded. For example, the contribution of parents and peers to children's involvement in therapy is relatively unexplored.

Finally, most child process research has been descriptive in nature, and atheoretical in focus. Part of the problem stems from the fact that child process research has evolved in isolation from the field of developmental psychopathology (Shirk & Russell, 1996). Treatment processes attain relevance to the degree to which they are specifically related to pathogenic processes. Thus, child process research cannot be dissociated from investigations of pathogenic mechanisms that contribute to child dysfunction. Understanding of pathogenic mechanisms will guide the search for change processes in child psychotherapy.

5.03.7 REFERENCES

Achenbach, T. M., & Edelbrock, C. S. (1981). Behavior problems and competencies reported by parents of normal and disturbed children aged 4 through 16. *Monographs of Society for Research in Child Development, 46* (1, No. 188).

Achenbach, T., McConaughy, S., & Howell, C. (1987). Child/Adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101,* 213–232.

Alexander, J., Barton, C., Schiavo, R., & Parsons, B. (1976). Systems–behavioral intervention with families of delinquents: therapist characteristics, family behavior, and outcome. *Journal of Consulting and Clinical Psychology, 44,* 656–664.

Armbruster, P., & Kazdin, A. (1994). Attrition in child psychotherapy. In T. Ollendick & R. Prinz (Eds.), *Advances in clinical child psychology* (Vol. 16, pp. 81–108). New York: Plenum.

Axline, V. (1947). *Play therapy.* New York: Ballantine.

Beutler, L. (1991). Have all won and must all have prizes? Revisiting Luborsky et al.'s verdict. *Journal of Consulting and Clinical Psychology, 59,* 226–232.

Bonner, B., & Everett, F. (1982). Influence of client preparation and therapist expectations on children's attitudes and expectations of psychotherapy. *Journal of Clinical Child Psychology, 11,* 202–208.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of working alliance. *Psychotherapy: Theory, Research, and Practice, 16,* 252–260.

Bowlby, J.*A secure base: parent–child attachment and healthy human development.* New York: Basic Books.

Braswell, L., Kendall, P., Braith, J., Carey, M., & Vye, C. (1985). "Involvement" in cognitive-behavioral therapy with children: Process and its relationship to outcome. *Cognitive Therapy and Research, 9,* 611–630.

Bugental, D., Whalen, C., & Henker, B. (1977). Cognitive attributions of hyperactive children and motivational assumptions of two behavior change approaches: evidence for an interactionist position. *Child Development, 48,* 874–884.

Burns, D., & Nolen-Hoeksema, S. (1992). Therapeutic empathy and recovery from depression in cognitive–behavioral therapy: a structural equation model. *Journal of Consulting and Clinical Psychology, 60,* 441–449.

Chamberlain, P., Patterson, G., Reid, J., Kavanagh, K., & Forgatch, M. (1984). Observation of client resistance. *Behavior Therapy, 15,* 144–155.

Coie, J., & Pennington, B. (1976). Children's perception of deviance and disorder. *Child Development, 47,* 407–413.

Coleman, D., & Kaplan, M. (1990). Effects of pretherapy videotape preparation on child therapy outcomes. *Professional Psychology: Research and Practice, 21,* 199–203.

Colson, D., Cornsweet, C., Murphy, T., O'Malley, F., Hyland, P., McParland, M., & Coyne, L. (1991). Perceived treatment difficulty and therapeutic alliance on an adolescent psychiatric hospital unit. *American Journal of Orthopsychiatry, 61,* 221–229.

Day, L., & Reznikoff, M. (1980). Preparation of children and parents for treatment at a children's psychiatric clinic through videotaped modeling. *Journal of Consulting and Clinical Psychology, 48,* 303–304.

Diamond, G., & Liddle, H. (1996). Resolving a therapeutic impasse between parents and adolescents in multidimensional family therapy. *Journal of Consulting and Clinical Psychology, 64,* 481–488.

Dozier, M., Cue, K., & Barnett, L. (1994). Clinicians as caregivers: role of attachment organization in treatment. *Journal of Consulting and Clinical Psychology, 62,* 793–800.

Drabman, R., Tarnowski, K., & Kelly, P. (1987). Are younger children disproportionately referred for childhood academic and behavior problems? *Journal of Consulting and Clinical Psychology, 55,* 907–909.

Eltz, M., Shirk, S., & Sarlin, N. (1995). Alliance formation and treatment outcome among maltreated adolescents. *Child Abuse and Neglect, 19,* 419–431.

Estrada, A., Russell, R., McGlinchey, K., & Hoffman, L. (1994, October). *The development of child and child therapist psychotherapy scales.* Poster presented at Kansas Conference on Clinical Child Psychology, Lawrence, Kansas.

Fonagy, P., Leigh, T., Steele, M., Kennedy, R., Mattoon, G., Target, M., & Gerber, A. (1996). The relation of attachment status, psychiatric classification, and response to psychotherapy. *Journal of Consulting and Clinical Psychology, 64,* 22–31.

Frank, J. (1974). Therapeutic components of psychotherapy. *Journal of Nervous and Mental Disease, 159,* 325–342.

Freud, A. (1965). *Normality and pathology in childhood: Assessments of development.* New York: International Universities Press.

Frieswyk, S., Allen, J., Colson, D., Coyne, L., Gabbard, G., Horwitz, L., & Newsome, G. (1986). Therapeutic alliance: its place as a process and outcome variable in psychotherapy research. *Journal of Consulting and Clinical Psychology, 54,* 32–38.

Garber, J., Braatladt, N., & Weiss, B. (1995). Attach regulation in depressed and nondepressed children and young adolescents. *Development and Psychopathology, 7,* 93–115.

Garber, J., Braafledt, N., & Zeman, J. (1991). The regulation of sad affect: An information processing perspective. In J. Garber & K. Dodge (Eds.), *The development of emotion regulation and dysregulation* (pp. 208–242). New York: Cambridge University Press.

Gibbons, J., & Foreman, S. (1989, June). *Advances in child psychotherapy research using the Mt. Zion Psychotherapy Research Model.* Panel presented at meetings of Society for Psychotherapy Research, Wintergreen, VA.

Gomes-Schwartz, B. (1978). Effective ingredients in psychotherapy: Prediction of outcome from process variables. *Journal of Consulting and Clinical Psychology, 46,* 1023–1035.

Greenberg, L. (1991). Research in the process of change. *Psychotherapy Research, 1,* 14–24.

Greenberg, L., & Newman, F. (1996). An approach to psychotherapy change process research: introduction to the Special Section. *Journal of Consulting and Clinical Psychology, 64,* 435–438.

Harter, S. (1988). Developmental and dynamic changes in the nature of the self-concept: implications for child psychotherapy. In S. Shirk (Ed.), *Cognitive development and child psychotherapy* (pp. 119–160). New York: Plenum.

Holmes, D., & Urie, R. (1975). Effects of preparing children for psychotherapy. *Journal of Consulting and Clinical Psychology, 43,* 311–318.

Howard, K., Lueger, R., Mahling, M., Martinovich, Z. (1993). A phase model of psychotherapy outcome: causal mediation of change. *Journal of Consulting and Clinical Psychology, 61,* 678–685.

Horvath, A., & Luborsky, L. (1993). The role of the therapeutic alliance in psychotherapy. *Journal of Consulting and Clinical Psychology, 61,* 561–573.

Kazdin, A. (1994). Psychotherapy for children and adolescents. In A. Bergin & S. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (pp. 543–596). New York: Wiley.

Kazdin, A. (1995). Bridging child, adolescent, and adult psychotherapy: directions for research. *Psychotherapy Research, 5,* 258–277.

Kazdin, A. E., Bass, D., Ayers, W. A., & Rodgers, A.

(1990). Empirical and clinical focus of child and adolescent psychotherapy research. *Journal of Consulting and Clinical Psychology, 58*(6), 729–740.

Kazdin, A., French, N., & Sherick, R. (1981). Acceptability of alternative treatments for children: evaluations by inpatient children, parents, and staff. *Journal of Consulting and Clinical Psychology, 49,* 900–907.

Kazdin, A., French, N., & Unis, A. (1983). Child, mother, father evaluations of depression in psychiatric inpatient children. *Journal of Abnormal Child Psychology, 11,* 167–180.

Kendall, P. (1994). Treating anxiety disorders in children: results of a randomized clinical trial. *Journal of Consulting and Clinical Psychology, 62,* 100–110.

Kirtner, W., & Cartwright, D. (1958). Success and failure in client-centered therapy as a function of initial in-therapy behavior. *Journal of Consulting Psychology, 22,* 267–281.

Kobak, R., & Sceery, A. (1988). Attachment in late adolescence: working models, affect regulation, and representations of self and other. *Child Development, 59,* 135–146.

Lambert, M., Weisz, J., & Knight, F. (1989). Over- and undercontrolled clinic referral problems of Jamaican and American children and adolescents: The cultural general and the culture specific. *Journal of Consulting and Clinical Psychology, 57,* 467–472.

Lambert, M., Weisz, J., Knight, F., Desrosiers, M., Overly, K., & Thesiger, C. (1992). Jamaican and American perspectives on child psychopathology: further explorations of the threshold model. *Journal of Consulting and Clinical Psychology, 60,* 146–149.

Landisberg, S., & Snyder, W. U. (1946). Non-directive play therapy. *Journal of Clinical Psychology, 2,* 203–213.

Leahy, R., & Huard, C. (1976). Role-taking and self-image disparity. *Developmental Psychology, 12,* 504–508.

Lebo, D., & Lebo, E. (1957). Aggression and age in relation to verbal expression in nondirective play therapy. *Psychological Monographs: General and Applied, 71*(20), 1–12.

Luborsky, L. (1976). Helping alliances in psychotherapy. In J. Cleghorn (Ed.), *Successful psychotherapy* (pp. 92–116). New York: Bruner/Mazel.

Main, M., Kaplan, N., & Cassidy, J. (1985). Security in infancy, childhood, and adulthood: A move to the level of representation. In I. Bretherton & E. Waters (Eds.), Growing points of attachment and research. *Monographs of the Society for Research in Child Development, 50,* (1–2, Serial No. 209).

Masten, A., & Braswell, L. (in press). Developmental psychopathology: An integrative framework for understanding behavior problems in chlldren and adolescents. In P. Martin (Ed.), *Handbook of behavior therapy and psychological science: An integrative approach.* New York: Pergamon.

McDermott, P. (1996). A nationwide study of developmental and gender prevalence in psychopathology in childhood and adolescence. *Journal of Abnormal Child Psychology, 24,* 53–66.

Milos, M., & Reiss, S. (1982). Effects at three play conditions on separation anxiety in children. *Journal of Consulting and Clinical Psychology, 50,* 389–395.

Mook, B. (1982a). Analyses of therapist variables in a series of psychotherapy sessions with two child clients. *Journal of Clinical Psychology, 38,* 63–76.

Mook, B. (1982b). Analyses of client variables in a series of psychotherapy sessions with two child clients. *Journal of Clinical Psychology, 38,* 263–274.

Moustakas, C. E., & Schalock, H. D. (1955). An analysis of therapist–child interaction in play therapy. *Child Development, 26*(2), 143–157.

Nolen-Hoeksema, S. (1990). *Sex differences in depression.* Stanford, CA: Stanford University Press.

Ollendick, T., & King, N. (1994). Diagnosis, assessment, and treatment of internalizing problems in children: the role of longitudinal data. *Journal of Consulting and Clinical Psychology, 62,* 918–927.

Orlinsky, D., Grawe, K., & Parks, B. (1994). Process and outcome in psychotherapy—Noch einmal. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 477–502). New York: Guilford Press.

Orlinsky, D. E., & Howard, K. I. (1986). Process and outcome in psychotherapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (pp. 311–381). New York: Wiley.

Patterson, G., & Chamberlain, P. (1994). A functional analysis of resistance during parent training therapy. *Clinical Psychology: Science and Practice, 1,* 53–70.

Patterson, G., & Forgatch, M. (1985). Therapist behavior as a determinant of client noncompliance: a paradox for the behavior modifier. *Journal of Consulting and Clinical Psychology, 53,* 846–851.

Prochaska, J., DiClemente, C., & Norcross, J. (1992). In search of how people change: applications to addictive behaviors. *American Psychologist, 47,* 1102–1114.

Raney, D., Shirk, S., Sarlin, N., Kaplan, D., & During, L. (1991, March). *Parent collaboration as a predictor of adolescent inpatient treatment process and progress.* Paper presented at meetings of the Society for Adolescent Medicine, Denver, CO.

Reid, J., Kavanagh, K., & Baldwin, D. (1987). Abusive parents' perceptions of child behavior problems: an example of parental bias. *Journal of Abnormal Child Psychology, 15,* 457–466.

Richters, J. (1992). Depressed mothers as informants about their children: a critical review of the evidence for distortion. *Psychological Bulletin, 112,* 485–499.

Rickard, K., Forehand, R., Wells, K., Greist, D., & McMahon, R. (1981). Factors in the referral of children for behavioral treatment: a comparison of mothers of clinic-referred deviant, clinic-referred nondeviant, and nonclinic children. *Behavior Research and Therapy, 19,* 201–205.

Ricks, D. (1974). Supershrink: Methods of a therapist judged successful on the basis of adult outcomes of adolescent patients. In D. Ricks, M. Rott, & A. Thomas (Eds.), *Life history research in psychopathology,* (Vol. 3, pp. 275–297). Minneapolis, MN: University of Minnesota Press.

Robinson, L., Berman, J., & Neimeyer, R. (1990). Psychotherapy for the treatment of depression: a comprehensive review of controlled outcome research. *Psychological Bulletin, 108,* 30–49.

Russell, R. (1994). *Empirical investigations of psychotherapeutic techniques: A critique of and prospects for language analysis.* Ann Arbor, MI: University Microfilms International.

Russell, R., Bryant, F., & Estrada, A. (1996). Confirmatory P-technique analyses of therapist discourse: High versus low quality child therapy sessions. *Journal of Consulting and Clinical Psychology, 64,* 1366–1377.

Russell, R. L., Greenwald, S., & Shirk, S. R. (1991). Language in child psychotherapy: a meta-analytic review. *Journal of Consulting and Clinical Psychology, 6,* 916–919.

Russell, R., & Shirk, S. (in press). Child psychotherapy process research. In T. Ollendick & R. Prinz (Eds.), *Advances in clinical child psychology* (Vol. 20). New York: Plenum.

Safran, J., & Muran, C. (1996). The resolution of ruptures in the therapeutic alliance. *Journal of Consulting and Clinical Psychology, 64,* 447–458.

Salzman, C., Luetgert, M., Roth, C., Creasant, J., & Howard, L. (1976). Formation of a therapeutic relationship: experiences during the initial phase of psychotherapy as predictors of treatment duration and outcome. *Journal of Consulting and Clinical Psychology, 44,* 546–555.

Schaughency, E., & Lahey, B. (1985). Mothers' and fathers' perception of child deviance: roles of child deviance, parental depression, and marital satisfaction. *Journal of Consulting and Clinical Psychology, 53,* 718–723

Schleser, R., Meyers, H., & Cohen, R. (1981). Generalization of self-instructions: effects of general versus specific content, active rehearsal, and cognitive level. *Child Development, 52,* 335–340.

Shirk, S. (Ed.) (1988). *Cognitive development and child psychotherapy,* New York: Plenum.

Shirk, S. (1990). Cognitive processes in child psychotherapy: where are the developmental limits? In J. Dewit, W. Slot, H. van Leeuwen, & M. Terwogt (Eds.), *Developmental psychopathology and clinical practice* (pp. 19–31). Amsterdam: Acco.

Shirk, S., & Russell, R. (1996). *Change processes in child psychotherapy: Revitalizing treatment and research.* New York: Guilford Press.

Shirk, S., & Saiz, C. (1992). Clinical, empirical, and developmental perspectives on the therapeutic relationship in child psychotherapy. *Development and Psychopathology, 4,* 713–728.

Shirk, S., Saiz, C., & Sarlin, N. (1993, June). *The therapeutic alliance in child and adolescent treatment: initial studies with inpatients.* Paper presented at meetings of the Society for Psychotherapy Research, Pittsburgh, PA.

Siegel, C. (1972). Changes in play therapy behavior over time as a function of differing levels of therapist-offered conditions. *Journal of Clinical Psychology, 28,* 235–236.

Sigelman, C., & Mansfield, K. (1992). Knowledge of and receptivity to psychological treatment in childhood and adolescence. *Journal of Clinical Child Psychology, 21,* 2–9.

Smith-Acuna, S., Durlak, J. A., & Kaspar, C. J. (1991). Development of child psychotherapy process measures. *Journal of Clinical Child Psychology, 20,* 126–131.

Stark, K., Rouse, L., & Livingston, R. (1991). Treatment of depression during childhood and adolescence: cognitive–behavioral procedures for the individual and the family. In P. Kendall (Ed.), *Child and adolescent therapy: Cognitive-behavioral procedures* (pp. 165–208). New York: Guilford Press.

Stoolmiller, M., Duncan, T., Bank, L., & Patterson, G. (1993). Some problems and solutions to the study of change: significant patterns in client resistance. *Journal of Consulting and Clinical Psychology, 61,* 920–928.

Strupp, H., & Hadley, S. (1977). A tripartite model of mental health and psychotherapy outcomes. *American Psychologist, 32,* 187–196.

Tarnowski, K., Anderson, D., Drabman, R., & Kelly, P. (1990). Disproportionate referrals for child academic/behavior problems: replication and extension. *Journal of Consulting and Clinical Psychology, 58,* 240–243.

Taylor, L., Adelman, H., & Kaser-Boyd, N. (1985). Exploring minors' reluctance and dissatisfaction with psychotherapy. *Professional Psychology: Research and Practice, 3,* 418–425.

Tharp, R. (1991). Cultural diversity and treatment of children. *Journal of Consulting and Clinical Psychology, 59,* 799–812.

Truax, C. B., Altman, H., Wright, L., & Mitchell, K. M. (1973). Effects of therapeutic conditions in child therapy. *Journal of Community Psychology, 1,* 313–318.

Truax, C. B., & Wittmer, J. (1973). The degree of therapist's focus on defense mechanisms and the effect on therapeutic outcome with institutionalized juvenile delinquents. *Journal of Community Psychology, 1,* 201–203.

Van Horn, M. (1996). *The relationship of interpersonal perceptions and depression to evaluations of supportive interactions and interpersonally-oriented coping strategies.* Unpublished doctoral dissertation, University of Denver, Denver, CO.

Vondra, J., Barnett, D., & Cicchetti, D. (1989). Perceived and actual competence among maltreated and comparison school children. *Development and Psychopathology, 1,* 237–255.

Walker, E., Bettes, B., & Ceci, S. (1984). Teachers' assumptions regarding the severity, causes, and outcomes of behavioral problems in preschoolers: implications for referral. *Journal of Consulting and Clinical Psychology, 52,* 899–902.

Webster-Stratton, C. (1988). Mothers' and fathers' perceptions of child deviance: role of parent and child behaviors and parent adjustment. *Journal of Consulting and Clinical Psychology, 56,* 909–915.

Weinstein, M. (1988). Preparation of children for psychotherapy through videotaped modeling. *Journal of Clinical Child Psychology, 17,* 131–136.

Weisz, J. (1986). Contingency and control beliefs as predictors of psychotherapy outcomes among children and adolescents. *Journal of Consulting and Clinical Psychology, 54,* 789–795.

Weisz, J., Suwanlert, S., Chaiyasit, W., Weiss, B., Walter, B., & Anderson, W. (1988). Thai and American perspectives on over- and undercontrolled child behavior problems: exploring the threshold model among parents, teachers, and psychologists. *Journal of Consulting and Clinical Psychology, 56,* 601–609.

Weisz, J., & Weiss, B. (1991). Studying the "referability" of child clinical problems. *Journal of Consulting and Clinical Psychology, 59,* 266–273.

Weisz, J., Weiss, B., Han, S., Granger, D., & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin, 117,* 450–468.

Weisz, J., Weiss, B., & Langmeyer, D. (1987). Giving up on child psychotherapy: who drops out? *Journal of Consulting and Clinical Psychology, 55,* 916–918.

Wisniewski, J., Andrews, T., & Mulick, J. (1995). Objective and subjective factors in the disproportionate referral of children for academic problems. *Journal of Consulting and Clinical Psychology, 63,* 1032–1036.

Zahn-Waxler, C. (1993). Warriors and worriers: gender and psychopathology. *Development and Psychopathology, 5,* 79–89.

Zimet, S., & Farley, G. (1986). Competence and self-esteem of emotionally-disturbed children beginning day treatment. *Journal of the American Academy of Child and Adolescent Psychiatry, 25,* 76–83.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.04

Outcome Findings and Issues in Psychotherapy with Children and Adolescents

JOHN R. WEISZ
University of California, Los Angeles, CA, USA

5.04.1 INTRODUCTION 84
5.04.2 SOCIAL, DEVELOPMENTAL, AND CULTURAL CONTEXT OF CHILD ASSESSMENT AND THERAPY 84
5.04.3 PREVALENCE OF CHILD DYSFUNCTION AND CHILD TREATMENT 84
5.04.4 TYPES OF TREATMENT AND PROBLEMS TARGETED IN TREATMENT 85
5.04.5 WHO CARES ABOUT OUTCOMES OF CHILD TREATMENT? 85
5.04.6 APPROACHES TO ASSESSING TREATMENT EFFECTS 85
5.04.7 POOLING OUTCOME STUDIES VIA META-ANALYSIS: METHODS AND ISSUES 86
5.04.8 FINDINGS FROM META-ANALYSES 88
5.04.8.1 Findings of Broad-based Meta-analyses 88
5.04.8.2 Findings of Specially Focused Meta-analyses 88
5.04.8.3 Therapy, Therapist, and Child Factors Related to the Magnitude of Treatment Effects 90
5.04.9 EXTERNAL VALIDITY AND RESEARCH THERAPY VS. CLINIC THERAPY 92
5.04.9.1 Evidence on the Effects of Clinic Therapy 92
5.04.9.2 Research Challenges 93
5.04.10 EXPORTABILITY OF EMPIRICALLY SUPPORTED TREATMENTS 95
5.04.11 CHILD PSYCHOTHERAPY AND DEVELOPMENTAL PSYCHOLOGY 96
5.04.11.1 Age 96
5.04.11.2 Cognitions About the Therapy Process and the Role of the Therapist 97
5.04.11.3 Ability to "Decenter" 97
5.04.11.4 Language Ability 97
5.04.11.5 Comprehension of Concepts or "Lessons" of the Treatment Program 97
5.04.11.6 Abstract Reasoning 98
5.04.11.7 Organization and the Orthogenetic Principle 98
5.04.11.8 Looking to the Future 98
5.04.12 RELATION BETWEEN PROCESS AND OUTCOME 98
5.04.12.1 Components of Treatment 99
5.04.12.2 Therapist Behavior 99
5.04.12.3 Child and Family Behavior 99
5.04.12.4 Therapeutic Relationship 100
5.04.13 CHILD AND FAMILY ETHNICITY AND CULTURE 100
5.04.13.1 How Effective are Tested Treatments with Ethnic Minority Children? 101
5.04.13.2 Is there Evidence of Treatment Type and Ethnic Group Interaction? 101

5.04.13.3 Outcomes of Therapist–Client Matching for Ethnicity or Language 101
5.04.13.4 Outcomes of Adapting Therapies Specifically for Ethnic Minority Children 101
5.04.14 CONTEXT SENSITIVITY OF TREATMENTS 102
5.04.15 REFERENCES 103

5.04.1 INTRODUCTION

After decades of child treatment guided by psychology's grand theories and rarely subjected to rigorous test, we are moving into an era in which claims of success may have little credence in the absence of outcome research. The forces ushering in this new era are both scientific and fiscal. On the scientific front, both the American Psychological Association (APA) and the American Psychiatric Association are moving toward identification of research-based principles and practices for intervention with adults, adolescents, and children (see, e.g., Task Force, 1995). On the fiscal front, the advent of managed care has brought increased attention to the issue of outcomes, as companies seek to limit the cost of care. In the area of child and adolescent treatment, the body of relevant evidence is now growing fast, with more than 300 child and adolescent treatment outcome studies already published. This chapter summarizes some of the major findings of these studies, and then becomes critical, noting limitations of the available evidence and suggesting issues that will need attention in the future. Throughout the chapter, the term "children" is used to refer to the age period from early childhood through adolescence, except where it is necessary to draw a distinction between children and adolescents.

5.04.2 SOCIAL, DEVELOPMENTAL, AND CULTURAL CONTEXT OF CHILD ASSESSMENT AND THERAPY

Although clinical work with children bears obvious similarities to work with adults, some important differences warrant emphasis. First, because children rarely perceive themselves as "disturbed" or as candidates for therapy, most referrals for treatment, up until late adolescence, tend to be made by parents, teachers, or other adults. These adults may thus be construed as "clients" in the sense that they commission the therapy, pay for it, and identify some or all of the goals the therapist is to pursue. Children may or may not participate in identifying target problems or setting treatment goals, and even when they participate, the adult input may be weighted more heavily. With therapy commissioned by adults, and its goals heavily influenced by adults, it is understandable that children may enter the process with little motivation for treatment or personal change, or with different objectives than those shared by the adults involved.

Given marked developmental differences in the self-awareness, psychological mindedness, and expressive ability of their clientele, child therapists must rely heavily on adults for information about the youngsters they treat, and this can present problems of several types. First, parents' and teachers' reports may be inaccurate, based on distorted samples of child behavior, influenced by their own adult agendas, calculated to conceal their own failings as parents (including neglect or abuse), or even biased by their own pathologies (see e.g., Kazdin, 1989); and levels of agreement among different adult informants reporting on the same child tend to be low (Achenbach, McConaughy, & Howell, 1987). Bias and distortion aside, adult reports of child behavior and adult identification of referral concerns are both apt to reflect the values, practices, and social ideals of their cultural reference group (see Weisz, McCarty, Eastman, Chaiyasit, & Suwanlert, 1997; Weisz et al., 1988; Weisz & Weiss, 1991).

Finally, children tend to be captives of their externally engineered environments to a much greater extent than do adults; one consequence may be that the "pathology" the child therapist treats may reside as much in a disturbed environment from which the child cannot escape as in the child himself or herself. This may limit the impact of interventions involving the child as solo or primary participant, and it may argue for involvement of others from the child's social context, but such significant others are not always willing or cooperative. So, in a number of ways, the child therapist faces challenges that are rather different from those confronted by one who treats only adults.

5.04.3 PREVALENCE OF CHILD DYSFUNCTION AND CHILD TREATMENT

Notwithstanding the difficulties of child assessment and child treatment, both practices are widespread, in both clinical practice and research contexts. Systematic efforts at assessment in the general population suggest that large numbers of children suffer from significant levels of disturbance, at least as defined within the current taxonomic systems. Epidemiologic studies in the late 1980s (summarized by Costello, 1989) indicated that at least 17% of children in the general population met criteria for at least one diagnosis in the *Diagnostic and statistical manual of mental disorders, Third Edition* (*DSM-III*; American Psychiatric Association, 1980); preliminary findings point to higher prevalence rates for the most recent edition of the diagnostic manual (*DSM-IV*). Of course, these rates of formal diagnosis overlook the many children who have very significant problems that fall short of full diagnosability, but may well need help, nonetheless.

Each year, many troubled children receive intervention in the form of child psychotherapy, but precise figures on the extent and cost are difficult to find. Figures for the late 1980s indicated that about 2.5 million American children received treatment each year (Office of Technology Assessment, 1986), and that the annual cost was more than $1.5 billion (Institute of Medicine, 1989). These figures are apt to be underestimates, given the likelihood that much child treatment is either not formally labeled "psychotherapy" (e.g., counseling in school) or not formally reported to government agencies (e.g., private practice paid out of pocket). Of course, much has changed since the 1980s, with some forces (e.g., the rise of managed care, general reductions in duration of inpatient stays) discouraging growth, but other forces (e.g., increased public awareness, legislated entitlements for handicapped children) encouraging growth, so updated statistics on use and costs are clearly needed.

5.04.4 TYPES OF TREATMENT AND PROBLEMS TARGETED IN TREATMENT

The most common types of treatment used with children, and the most common child conditions addressed in treatment, are described in other chapters of this volume, but a few facts germane to outcome assessment warrant emphasis here. As Kazdin (1988) noted, there are more than 230 named psychotherapies, but most of these have not been subjected to controlled tests of outcome. In general, outcome research with children has focused on a distinct subset of the treatments that bears little apparent relation to frequency of use in clinical practice. Among the psychological interventions (i.e., excluding pharmacological therapies), psychodynamic and family therapies are among the most widely practiced by clinicians but among the least studied by outcome researchers. By contrast, behavioral and cognitive-behavioral interventions are emphasized in outcome research to a degree that far exceeds their use in clinical practice.

There is somewhat closer correspondence between practice patterns and the outcome research with respect to problems targeted in treatment, but the correspondence is less than perfect. Most of the types of child problems encompassed in Chapters 13–27, this volume have been the focus of both clinical attention and some outcome research, although much of the research has addressed specific problems or clusters of problems rather than full-blown diagnostic categories; this may have been appropriate, given periodic changes in the taxonomy, but late 1990s trends incline toward formal diagnostic assessment of children in clinical trials. Of the diagnostic categories noted in this volume, the problems associated with conduct disorder (and oppositional defiant disorder), attention deficit hyperactivity disorder (ADHD), anxiety disorders (including particularly specific phobias), and possibly somatoform disorder (especially chronic headache) have been emphasized in clinical trials research, whereas other problems frequently treated by practitioners are less evident in clinical trials research.

5.04.5 WHO CARES ABOUT OUTCOMES OF CHILD TREATMENT?

Stakeholders with an interest in the outcome of child treatment are numerous. The treated child's need for help, not to mention his or her investment of time and energy in the process of treatment, arguably makes the child the major stakeholder. Add parents and other family members who seek relief for the child, and also frequently for the family. Teachers' interests, as well, may include both concern for the child and for the classroom of which the child is a part. Those who finance the child's treatment (e.g., family members, government agencies, insurance carriers) have a clear stake. And finally, the therapists, clinic staff, administrators, and others in the "provider" community have a clear interest in the outcomes of the care in which they invest their careers. All these parties to the process of treatment have a clear stake in the question, "How effective is child psychotherapy?"

5.04.6 APPROACHES TO ASSESSING TREATMENT EFFECTS

This question can be answered via several different methods, a few of which need to be noted here. The most widely recognized is the clinical trial, an outcome study in which the post-treatment adjustment of a group of children who received a candidate intervention is compared with that of one or more control groups who did not. It is these clinical trials studies that are most frequently pooled in reviews and meta-analyses (see below), and that thus constitute most of the evidence discussed in this chapter. However, in circumstances where all the children in a target group must receive an active treatment, multiple baseline designs, ABAB (sometimes called "reversal") designs, and simultaneous/alternating treatment designs are useful. Such approaches characterize much of the treatment research with ADHD youngsters (see, e.g., Pelham et al., 1993), for example, and they are frequently used in cases where an entire classroom is the target of an intervention (e.g., Wurtele & Drabman, 1984). These non-clinical trials designs can also be applied to cases (sometimes involving rare conditions) where only one or two children will be treated (e.g., McGrath, Dorsett, Calhoun, & Drabman, 1988; Tarnowski, Rosen, McGrath, & Drabman, 1987). These alternative outcome assessment designs have generated a rich body of data on treatment effects, data that await an enterprising reviewer. For now, though, the focus is on the clinical trials research, which has been reviewed rather thoroughly in the form of several meta-analyses, as described below.

5.04.7 POOLING OUTCOME STUDIES VIA META-ANALYSIS: METHODS AND ISSUES

Research findings on psychotherapy effects can be pooled via a technique called meta-analysis (see Mann, 1990; Smith, Glass, & Miller, 1980; but see also critiques, e.g., by Wilson, 1985). The building block of meta-analyses is the effect size (ES) statistic. The ES is an index of the size and direction of treatment effects. For typical clinical trials studies, ES is the difference between the post-treatment mean on an outcome measure for the treated group vs. the corresponding control group mean, with the difference divided by the SD of the outcome measure. Figure 1 is a guide to interpreting ES values. As the figure indicates, positive ES values indicate treatment benefit, and negative values indicate a harmful effect. Each ES value corresponds to a percentile standing of the average treated child on the outcome measure(s) if that child were placed in the control group after treatment; for example, an ES of 0.9 indicates that the average treated child scored better after treatment than 82% of the control group. As an aid to interpretation, Cohen's (1992) guidelines suggest that an ES of 0.20 may be considered a "small" effect, 0.50 a "medium" effect, and 0.80 a "large" effect. By averaging across the outcome measures used, a meta-analyst may compute a single mean ES for each study (or each treatment group) in the collection being reviewed. This permits computation of an overall mean ES for the entire collection of studies; it also permits comparison of mean ES across studies differing in potentially important ways: for example, in the type of therapy employed, the target problem being treated, or the age or gender of the children involved.

Methods have been proposed for assessing the quality of studies considered for inclusion in meta-analyses (see, e.g., Chalmers et al., 1981), but most meta-analysts in the psychotherapy research area have aimed for rather broad, representative collections of outcome studies, accepting a range of methods, provided that certain basic standards of experimental design are satisfied. Weiss and Weisz (1990) carried out a meta-analysis to assess the extent to which methodological variations in child outcome studies might be associated with differences in ES. Findings indicated that nine commonly noted internal and external validity factors (e.g., whether subjects were randomly assigned to groups, whether subjects were analog cases or clinical cases who would have been treated independently of the study) together accounted for 7% of the variance in ES, over a sample of 105 outcome studies. They also found no evidence that inclusion of methodologically weaker studies in meta-analyses has led to overestimates of ES; on the contrary, increased experimental rigor in the studies they sampled was generally associated with *larger* ES. Thus, it is possible that the most inclusive meta-analyses may tend to underestimate the true effect of therapy, compared to meta-analyses that apply strict methodological criteria to candidate studies.

Meta-analysts must make myriad decisions as they carry out their work, and many of these decisions have an impact on the ultimate ES values that will emerge. For example, decisions must be made about whether to (i) compute ES by dividing the treatment–control group difference by standard deviation (SD) of the control group or of the pooled control plus treatment groups; (ii) compute an average ES across all outcome measures within a study or keep these separate for overall ES calculation; (iii) combine ES computation across various treatment conditions, compute separate ES values for each treatment group in a study, or use some amalgam of these two approaches; and (iv) weight all studies equally or adjust ES according

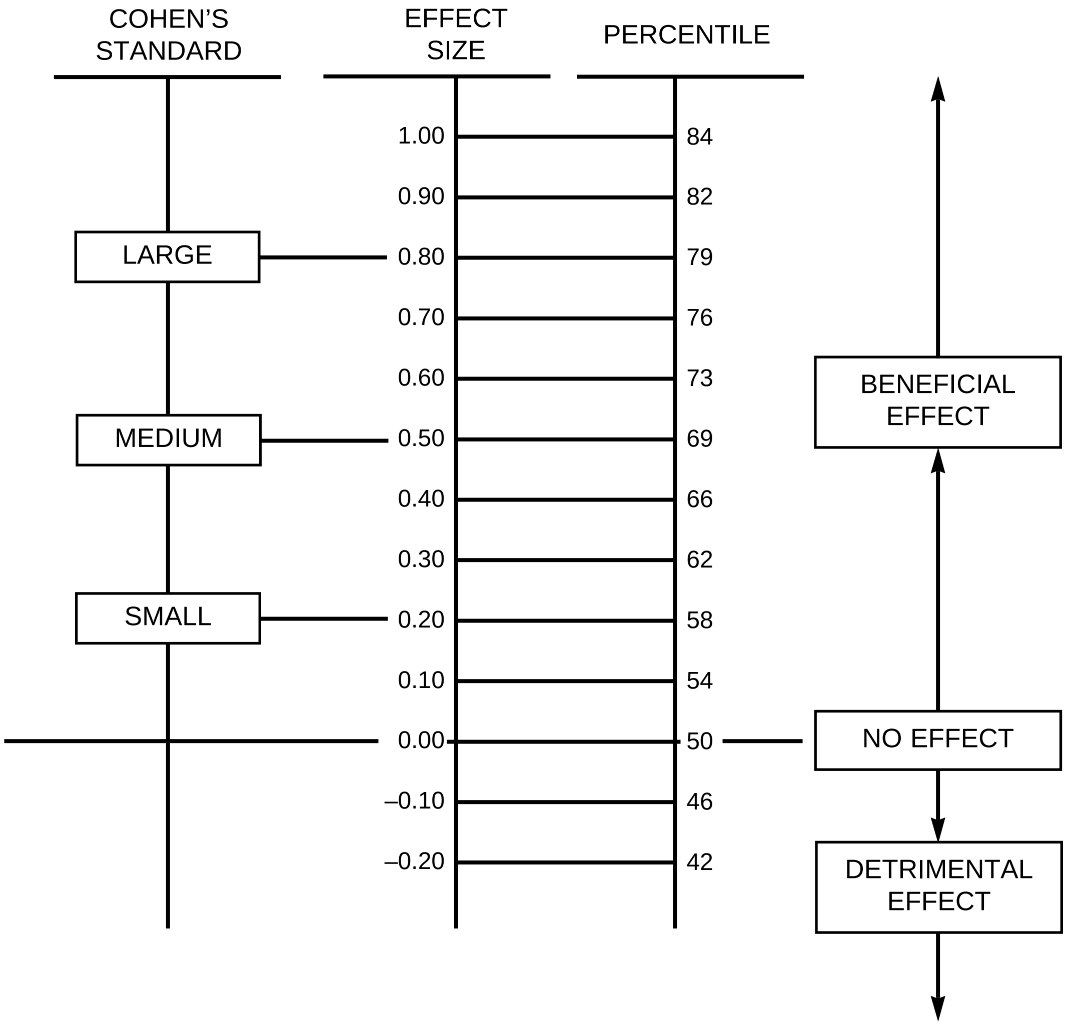

Figure 1 An aid to interpreting effect size (ES) statistics. ("Bridging the gap between laboratory and clinic in child and adolescent psychotherapy," by J. R. Weisz, G. R. Donenberg, S. S. Han, and B. Weiss, 1995, *Journal of Consulting and Clinical Psychology, 63,* pp. 688–701. Copyright 1995 by the American Psychological Association. Reprinted with permission.)

to sample size, ES variance, or some other leveling factor (for discussion of these and other decision-making issues, see Weisz & Weiss, 1993; Weisz, Weiss, Han, Granger, & Morton, 1995). Because scores of such decisions are made in any meta-analysis, it is essentially impossible that any two teams working independently will conduct their work in exactly the same way.

This fact, of course, highlights the need for caution in comparing findings across different meta-analyses, and thus in reading the results reported below. Differences in ES means across different meta-analyses may be influenced in part by differences in the ways the respective meta-analysts compiled studies, pooled them, computed ES, or averaged ES within or across studies. To assist the reader in interpreting the meta-analytic findings reported below, the four issues noted in the preceding paragraph have been addressed in the following way: For each meta-analysis, only *exceptions* to the most common approach to each issue are noted. That is, unless otherwise noted, the meta-analysis being described involved the following modal procedures for child meta-analysis: (i) ES was calculated by dividing the treatment–control group difference by SD of the pooled control plus treatment groups; (ii) within each study, ES values are averaged across all outcome measures; (iii) within each study, separate ES means are maintained for each treatment vs. control group comparison; and (iv) ES values reported are unadjusted.

5.04.8 FINDINGS FROM META-ANALYSES

This section considers some of the findings of meta-analytic reviews, first covering overall effect sizes found in broad-based analyses, then noting findings of specially focused analyses, and finally identifying some heuristically useful findings on therapy, therapist, and child factors related to the magnitude of treatment effect.

5.04.8.1 Findings of Broad-based Meta-analyses

To date, there have been at least four broad-based child psychotherapy meta-analyses: that is, meta-analyses imposing minimal limits on treated problems or types of intervention to be included. In all, these four meta-analyses encompass more than 300 separate treatment outcome studies. In the earliest of these, Casey and Berman (1985) surveyed those outcome studies published between 1952 and 1983 that involved children aged 12 and younger. Mean ES was 0.71 for those studies that included treatment–control comparisons; in percentile terms, the average treated child scored better after treatment than 76% of control group children, averaging across outcome measures. In a second meta-analysis, Weisz, Weiss, Alicke, and Klotz (1987) reviewed outcome studies published between 1952 and 1983, with children aged 4–18. Mean ES (computed using control group SD) was 0.79; after treatment, the average treated child was at the 79th percentile of control group peers.

In a third broad-based meta-analysis, Kazdin, Bass, Ayers, and Rodgers (1990) included studies published between 1970 and 1988, with youngsters aged 4–18. Kazdin et al. included multiple types of studies, but for the subset that compared treatment groups and no-treatment control groups, mean ES was 0.88; the average treated child scored higher after treatment than 81% of the no-treatment comparison group. For studies comparing treatment groups to active control groups, mean ES was 0.77, indicating that the average treated child was functioning better, post-treatment, than 78% of the control group.

In the fourth broad-based meta-analysis, Weisz, Weiss, et al. (1995) included studies published between 1967 and 1993, involving children aged 2–18. ES values were calculated using control group SD (for rationale, see Weisz, Weiss, et al., p. 455), and using the conservative procedure of collapsing across treatment groups up to the level of analysis; for example, except for analyses comparing different types of therapy, ES values were averaged across treatment groups within a study. An overall mean ES of 0.71 indicated that, after treatment, the average treated child scored higher on the outcome measures than 76% of control group children. When ES values were weighted by the inverse of their variance, the overall ES mean dropped to 0.54. (For more detailed descriptions of the procedures and findings of these broad-based meta-analyses, see Weisz & Weiss, 1993.)

The findings of these four broad-based meta-analyses show rather consistent beneficial treatment effects; ES values ranged from 0.71 to 0.84 (0.84 is an estimated overall mean for the treatment–control comparison studies in Kazdin et al., 1990), near Cohen's (1988) "large effect" index of 0.80. (Note, though, that the Weisz, Weiss, et al. (1995) finding noted in the preceding paragraph suggest that true population ES means, adjusting for heterogeneity of variance, may be closer to "medium" effects.) For comparative purposes, Figure 2 shows findings of the four child meta-analyses together with findings of two of frequently cited meta-analyses with older groups, that is, Smith and Glass's (1977) meta-analysis of primarily adult psychotherapy outcome studies, and Shapiro and Shapiro's (1982) meta-analysis of exclusively adult outcome studies. As the figure indicates, effects found in the child meta-analyses fall roughly within the range of effects found in these two adult meta-analyses.

5.04.8.2 Findings of Specially Focused Meta-analyses

Beyond the findings of these broad-based meta-analyses, there is a good deal to be learned from focused meta-analyses addressing rather specific questions about child treatment effects. Three such projects have dealt with cognitive-behavioral therapy in particular. For example, Baer and Nietzel (1991), reviewing 36 outcome studies, found quite positive treatment effects for cognitive and/or cognitive-behavioral treatments addressing child impulsivity. Mean ES was 0.90, calculated with control group ES, pooling only across similar outcome measures, and without weighting; the mean dropped to 0.77 when ES was weighted by study sample size. Dush, Hirt, and Schroeder (1989) focused on 48 child outcome studies using the cognitive-behavioral technique of self-statement modification; pooling outcome measures to produce one ES mean per study, Dush et al. found a mean ES of 0.47. Durlak, Fuhrman, and

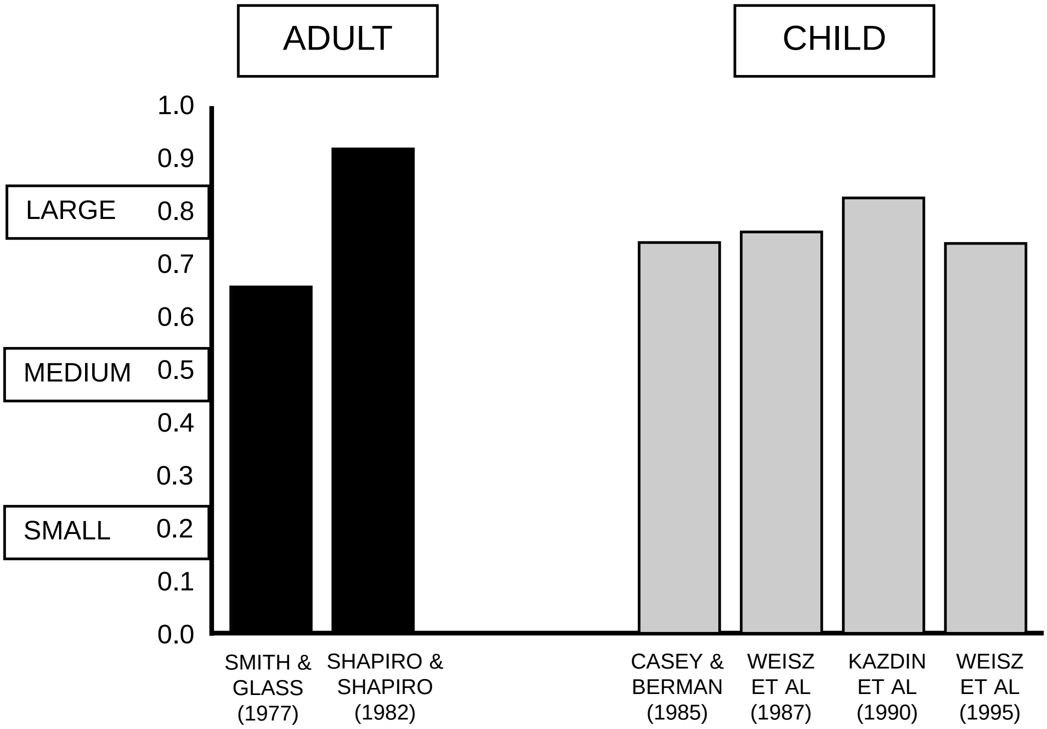

Figure 2 Mean effect sizes found in meta-analyses of psychotherapy outcome studies with adults and children. ("Bridging the gap between laboratory and clinic in child and adolescent psychotherapy," by J. R. Weisz, G. R. Donenberg, S. S. Han, & B. Weiss, 1995, *Journal of Consulting and Clinical Psychology*, *63*, pp. 688–701. Copyright 1995 by the American Psychological Association. Reprinted with permission.)

Lampman (1991) surveyed 64 studies of cognitive-behavioral treatment with children, applied across a range of target problems. With ES values pooled within studies to produce a single mean ES per study, and weighting by study sample size, Durlak et al. found a mean ES of 0.56, averaging across all age levels. However, as these authors had hypothesized, the mean was considerably higher for older children (0.92 for ages 11–13) than for younger groups (0.57 for ages 5–7, 0.55 for ages 7–11), suggesting that these cognitively oriented treatments may be a better fit to more cognitively mature youth. In a worrisome finding, Durlak et al. noted a negligible correlation (–0.22) between cognitive changes and behavioral changes, thus raising an important question about a key assumption underlying cognitive-behavioral therapy, namely the notion that the technique brings about behavioral change by inducing cognitive change.

Family therapy has been the subject of two meta-analyses. Hazelrigg, Cooper, and Borduin (1987) reviewed 20 family therapy outcome studies, averaging across multiple effects within studies to produce a single mean ES per study. For the seven studies that met their methodological criteria and included outcome measures of family interaction, the mean ES, weighted by study sample size, was 0.45; for the six methodologically appropriate studies that included behavior ratings of individual children in the family, the weighted mean ES was 0.50. Thus, although the number of relevant family therapy studies was small, Hazelrigg's meta-analysis pointed to beneficial effects. In another meta-analysis, Shadish et al. (1993) surveyed 101 studies involving family therapy (plus another 62 studies of marital therapy). Aggregating multiple ES values within studies to the study level, and weighting studies by the inverse of their sampling variance, Shadish et al. found a mean ES of 0.47 for the 44 family therapy studies that included treatment–control group comparisons. For comparison with the Hazelrigg et al. findings, Shadish also computed weighted ES means separately, for measures of family interaction and for ratings of individual family members' behavior; the family interaction ES was 0.31 (vs. 0.45 in Hazelrigg et al.), and the behavior rating ES was 0.66 (vs. 0.50 in Hazelrigg et al.).

Three additional meta-analyses illustrate the range of questions that can be addressed via this method of synthesis. Prout and DeMartino's (1986) meta-analysis, focused on psychotherapy in school settings, based ES computation on SD of the control groups, and pooled ES values only across groups of similar outcome measures; thus, 33 studies, involving 52 treatment–control comparisons, generated 119 individual ES values that were included in the analysis. The resulting (unadjusted) ES mean was 0.58. Russell, Greenwald, and Shirk (1991), computing ES using the control group SD in their meta-analysis of 18 child outcome studies, found that treatment was associated with significant improvement in children's language proficiency (unadjusted $ES = 0.39$). Treatments that emphasized spontaneous verbal interaction produced the biggest language improvements, compared to those involving more structured and constrained procedures. Finally, Saile, Burgmeier, and Schmidt (1988) carried out a meta-analysis of 75 controlled outcome studies involving psychological preparation of children for medical (including dental) procedures, ranging from minor procedures such as injections and drawing blood samples to such serious procedures as lumbar punctures. ES was computed using SD of study control groups. The overall mean ES was modest, 0.44, but there was considerable variation across types of intervention; for example, mean ES was only 0.22 for procedures in the "minor" category (e.g., injections) but 1.50 for procedures classified as "major" (e.g., cardiac catheterization). This report by Saile et al. (1988), like the article by Prout and DeMartino (1986), reminds us that a good deal of psychotherapy with children takes place in settings outside a therapist's office, and that evidence is needed on outcomes in these less traditional settings just as for office-based treatment.

5.04.8.3 Therapy, Therapist, and Child Factors Related to the Magnitude of Treatment Effects

In addition to overall mean ES values, meta-analyses can yield estimates of the impact of various therapy, therapist, and client factors on treatment outcome. Such estimates need to be interpreted with caution because of the confounding among factors (e.g., treatments and treated problems) that is common in meta-analyses. Some of the confounding can be addressed via statistical control and testing of interaction effects (see, e.g., Weisz et al., 1987; Weisz, Weiss, et al., 1995), although this is only a partial solution. In the two meta-analyses from the author's lab (Weisz et al., 1987; Weisz, Weiss, et al., 1995), studies involving behavioral treatments (e.g., behavioral contracting, modeling, cognitive-behavioral therapy) were found to produce larger effects than studies using nonbehavioral treatments (e.g., insight-oriented therapy, client-centered counseling). The Casey–Berman (1985) meta-analyses showed the same effect, at $p = 0.06$; Kazdin et al. (1990) did not make this comparison; general superiority of behavioral methods was also reported in Prout and DeMartino's (1986) meta-analysis of school-based treatment studies, and in meta-analyses not specific to children, by Nicholson and Berman (1983) and Shadish et al. (1993). By contrast meta-analyses have generally not found treatment outcomes to differ reliably for different types of treated problems (e.g., internalizing vs. externalizing; for one exception, see Casey & Berman, 1985, pp. 392–393).

The relation between age and treatment outcome has varied across meta-analyses. However, the meta-analysis involving the most recent collection of studies (Weisz, Weiss, et al., 1995) found mean ES to be larger for adolescents than for children. This main effect was qualified by the age and gender interaction depicted in Figure 3; for samples of predominantly or exclusively adolescent girls, mean ES was twice as large as for adolescent boys and for children of both genders. Of course, adolescent girls may be more likely to be treated for internalizing problems than are younger children or adolescent boys; but we found no reliable difference in mean ES for internalizing vs. externalizing problems, nor was the age and gender interaction in Figure 3 qualified by type of treated problem (internalizing vs. externalizing), type of treatment (behavioral vs. nonbehavioral), or level of therapist training (professional vs. clinical trainee vs. paraprofessional). Perhaps there is something about the 1990s generation of treatments reflected in this 1995 meta-analytic collection of studies that fits the characteristics and needs of adolescent girls particularly well, but one is left to speculate about what that elusive quality may be.

Three additional findings illustrate other potentially useful contributions of meta-analytic data. First, meta-analysis can be used to assess the holding power of intervention effects. Findings in both the meta-analyses (Weisz et al., 1987; Weisz, Weiss, et al., 1995) indicated that treatment effects assessed immediately after treatment are quite similar to effects measured at follow-up assessments, which average about six months after treatment termination. This suggests that treatment benefits tend to be durable, at least within typical follow-up time frames (for similar conclusions derived from different kinds of outcome studies, see Durlak et al. (1991); Nicholson & Berman (1983); Shadish et al. (1993)). Second, individually administered treatments have tended to show larger effects than group treatments, although not always significantly so (see Casey & Berman, 1985;

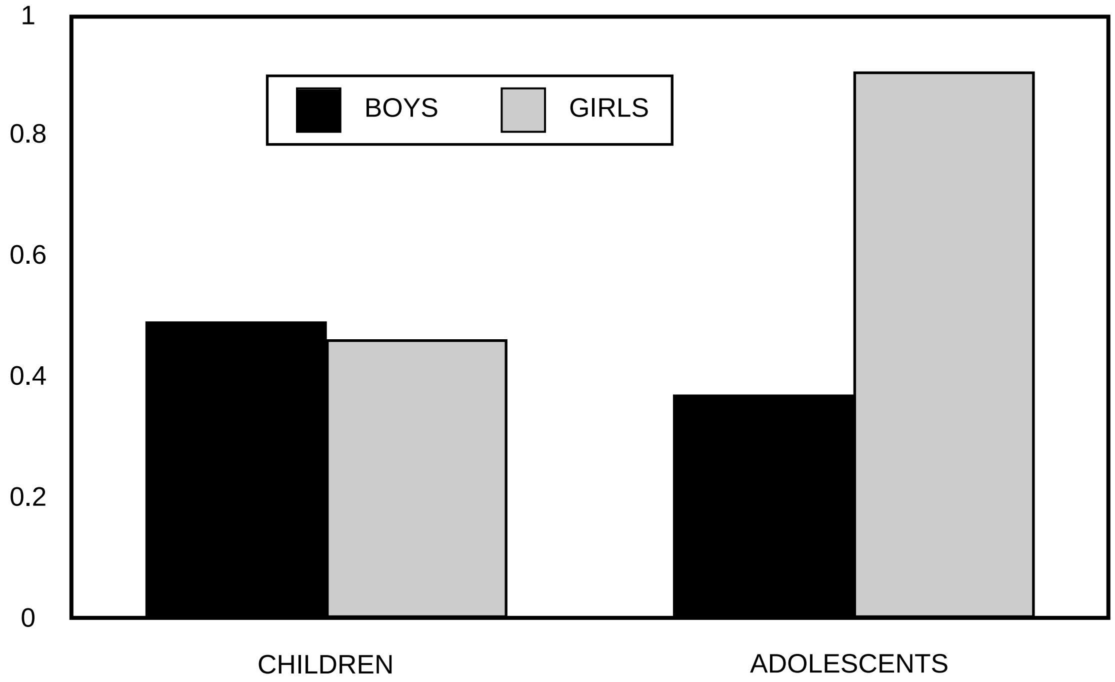

Figure 3 Mean effect size for samples of predominantly male and female children (11 years of age and younger) and adolescents (12 years and older). ("Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies," by J. R. Weisz, B. Weiss, S. S. Han, D. A. Granger, and T. Morton, 1995, *Psychological Bulletin, 117,* pp. 450–468. Copyright 1995 by the American Psychological Association. Reprinted with permission.)

Russell et al., 1991; Weisz et al., 1987; Weisz, Weiss, et al., 1995), and there is at least one exception in the area of cognitive-behavioral treatments for impulsivity (Baer & Nietzel, 1991).

A third finding (from Weisz, Weiss, et al., 1995) concerns the specificity of treatment effects. Obviously, children differ markedly from one another in the particular problems they manifest, and various therapies differ in the problems they are designed to address, but do these individual differences actually influence treatment outcome? Some (e.g., Frank, 1973) have proposed that psychotherapy has "nonspecific" effects, for example, helping people with diverse problems through such general means as promoting a feeling of being understood or encouraging an expectancy of relief. An alternative view is that therapies help in specific ways, having their strongest influence on the specific problems they are designed to address.

This controversy was addressed in Weisz, Weiss, et al. (1995) with a test of whether effects were larger for the specific problem domains targeted by a treatment than for other, more incidental domains. For example, we asked whether a treatment for anxiety produced bigger changes in anxiety than in related but more peripheral problems such as depression. Across multiple comparisons like these, we found that ES means were about twice as large for the specific problems addressed in treatment as for related problems that were not specifically addressed. This suggests that these psychotherapies were not merely producing global or nonspecific good feelings that influence diverse outcomes equally; instead, the treatments appeared to have rather precise, focused effects consistent with the particular objectives of the therapy.

In sum, evidence from broad-based meta-analyses of child psychotherapy outcome studies points to positive, problem-specific, and durable effects of mental health interventions for a variety of child problems. Clearly, child psychotherapy research is generating a number of encouraging findings. On the other hand, several conceptual and methodological issues require a good deal of attention in the future. Attention is now focused on some of these issues, and related suggestions for future research.

5.04.9 EXTERNAL VALIDITY AND RESEARCH THERAPY VS. CLINIC THERAPY

One issue is that of external validity. Most of the 300-plus studies in the meta-analyses reviewed above (particularly the recent and behavioral studies) involved participants, interventions, and/or treatment conditions that are actually not very representative of conventional clinical practice with referred children. In many of the studies, (i) children were recruited for treatment, were not actual clinic cases, and had lower levels of dysfunction than most referred children; (ii) child samples were homogenous, with therapy addressing only one or two focal problems (e.g., a specific phobia); (iii) therapists received considerable pretherapy training and session-by-session supervision in the specific intervention techniques they would use; and/or (iv) the therapy involved adherence to those specific techniques. In addition, (v) therapy was often highly structured and guided by a manual, with therapists monitored for fidelity to the treatment program.

These features of the outcome studies, although they are useful experimentally, tend to coalesce around an abstract category that the author and his colleagues (e.g., Weisz, Weiss, & Donenberg, 1992) have called research therapy, which we distinguish from conventional clinic therapy. Table 1 summarizes some illustrative differences between the two therapy genres. The two are best thought of as two poles of a multidimensional continuum; certainly no single feature shown in the table under Research therapy is present in all clinical trials studies, and few of the individual features listed under Clinic therapy are true of all clinic-based treatment. However, differences between child therapy in clinics and child therapy in clinical trials outcome studies are common enough for it to be reasonable to ask whether the positive outcomes generated in the research therapy studies, and summarized in the meta-analyses above, are representative of the outcomes achieved in actual clinical practice with children.

Table 1 Some common characteristics of research therapy and clinic therapy.

Research therapy	*Clinic therapy*
Recruited cases (less severe, study volunteers)	Clinic-referred cases (more severe, some coerced into treatment)
Homogeneous groups	Heterogeneous groups
Narrow or single-problem focus	Broad, multiproblem focus
Treatment in lab, school settings	Treatment in clinic, hospital settings
Researchers as therapists	Professional career therapists
Very small caseloads	Very large caseloads
Heavy pretherapy preparation for therapists	Little/light pretherapy preparation for therapists
Preplanned, highly-structured treatment (manualized)	Flexible, adjustable treatment (no treatment manual)
Monitoring of therapist behavior	Little monitoring of therapist behavior
Behavioral methods (e.g., operant, respondent, modeling, CBT)	Nonbehavioral methods (e.g., psychodynamic, eclectic

Source: Weisz, Donenberg, Han, & Weiss (1995). ("Bridging the gap between laboratory and clinic in child and adolescent psychotherapy," by J. R. Weisz, G. R. Donenberg, S. S. Han, & B. Weiss, 1995, *Journal of Consulting and Clinical Psychology*, *63*, pp. 688–701. Copyright 1995 by the American Psychological Association. Reprinted with permission.)

5.04.9.1 Evidence on the Effects of Clinic Therapy

To address this question, the author and colleagues carried out a search (described in Weisz, Donenberg, Han, & Weiss, 1995) for published outcome studies that focused on what might fairly be called clinic therapy. Studies were sought that involved (i) treatment of clinic-referred (i.e., not "analog" or recruited) youngsters, (ii) treatment in service-oriented clinics or clinical agencies, not in research settings (e.g., not university labs or public schools), (iii) therapy carried out by practicing clinicians (as opposed to trained research assistants), and (iv) therapy that was part of the usual service provided by the clinic, not a special treatment program designed specifically for research. It was required that the studies involve direct comparison between youngsters who received treatment and a control group receiving no treatment or a placebo intervention.

One of the first lessons learned was this: clinic studies that meet the criteria outlined above are very rare. The author and colleagues had carried out one such study (Weisz & Weiss, 1989), but found only eight others that fit the criteria (Ashcroft, 1971; De Fries, Jenkins, & Williams, 1964; Jacob, Magnussen, & Kemler, 1972; Lehrman, Sirluck, Black, & Glick, 1949; Levitt, Beiser, & Robertson, 1959; Shepherd, Oppenheim, & Mitchell, 1966; Smyrnios & Kirkby, 1993; Witmer & Keller, 1942), and these spanned 50 years, most having been published many years earlier. Although the studies all compared treatment and control groups, they used several different methodologies (for details, see Weisz, Donenberg, Han, & Kauneckis, 1995). To facilitate comparison of these nine studies with the meta-analytic findings reviewed above, an ES or ES estimate was computed for each of the nine studies (using, where necessary, estimation procedures described by Smith et al., 1980, and Glass, McGaw, & Smith, 1981). As shown in Figure 4, ES ranged from –0.40 to +0.29, with mean ES for the nine clinic studies (0.01) falling well below the mean ES of the four broad-based meta-analyses discussed earlier (0.77). Although these nine clinic studies provide only a very limited sample of clinical treatment, the disappointing findings certainly raise the possibility that outcomes of conventional clinical therapy may be less positive than the outcomes of research therapy.

It is useful to consider these findings in the light of evidence on "continuum of care" or "system of care" programs for children, that is, efforts to link multiple conventional mental health services for children with the services frequently organized and coordinated by individual child case managers (see, e.g., Stroul & Friedman, 1986). In one of the most recent (1995) and most ambitious of these efforts (see, Bickman, 1996; Bickman et al., 1995), the US Army spent $80 million to provide a continuum-of-care program for children of military personnel at Fort Bragg, NC, and to test the program's cost-effectiveness relative to more typical fragmented services in a matched comparison site. Considerable evidence indicates that the Fort Bragg program did produce a rich array of well-integrated mental health services (see details in Bickman et al., 1995). Indeed, the program was judged by the APA's section on Child Clinical Psychology and Division of Child, Youth, and Family Services Joint Task Force to be

> the most comprehensive program to date, integrating many of the approaches demonstrated by other service programs . . . integrated and flexibly constructed, yet comprehensive, [with] services available to be adapted to meet the needs of children and their families, rather than a simplistic application of a single approach. (Roberts, 1994, p. 215)

There is good evidence (see Bickman et al., 1995) that the program produced improved access to treatment, but it also cost more than services in the comparison site: $7777 per treated child vs. $4904. Unfortunately, despite the increased access to care and despite the increased costs, the Fort Bragg program produced clinical outcomes that were no better than those in the matched comparison site. Fort Bragg children received more mental health intervention, at greater cost, but their mental health outcomes were not improved by the increase.

Similar null findings have emerged from other studies designed to alter, link, or improve delivery of conventional clinical services (see, e.g., Evans et al., 1994; Lehman, Postrado, Roth, McNary, & Goldman, 1994). A number of alternative interpretations of such findings may be plausible, but one possible interpretation is that the various treatments that are linked and coordinated in these continua of care are simply not very effective, individually or in combination. In fact, Bickman and colleagues (Bickman, 1996; Bickman et al., 1995) appear to favor such an interpretation, after having tested the major rival interpretations they could identify. Put simply, there is no indication that the individual interventions employed in these various continua of care have been shown to be effective in clinical trials (or approximations thereto); thus, it is possible that the various interventions are simply not very helpful to children. An array of ineffective services may not produce much true benefit regardless of the extent to which they are multiplied, coordinated, or organized into continua of care.

So, the evidence reviewed on representative clinical interventions with children points to two conclusions: Evidence on conventional clinical treatments provides little support for their effectiveness; evidence on effects of integrating conventional interventions into systems of care also shows little evidence of treatment benefit.

5.04.9.2 Research Challenges

The findings on clinic- and community-based interventions suggest useful directions for future research. First, it is clear that we need more information on outcomes of treatment under representative clinical conditions. Research in this genre—known as "effectiveness" studies, as opposed to controlled "efficacy" studies—is difficult but certainly not impossible, as the nine studies cited in Figure 4 demonstrate (for a discussion of the pros and cons of various relevant methods, see Weisz & Weiss, 1993). The base of information on psychotherapy effects in public clinics is quite thin, but the situation is even worse for other treatment contexts. As best as can be determined there is, up to 1997, no methodologically sound treatment–control study of outcomes in such now-common treatment configurations as individual and group private practice and health

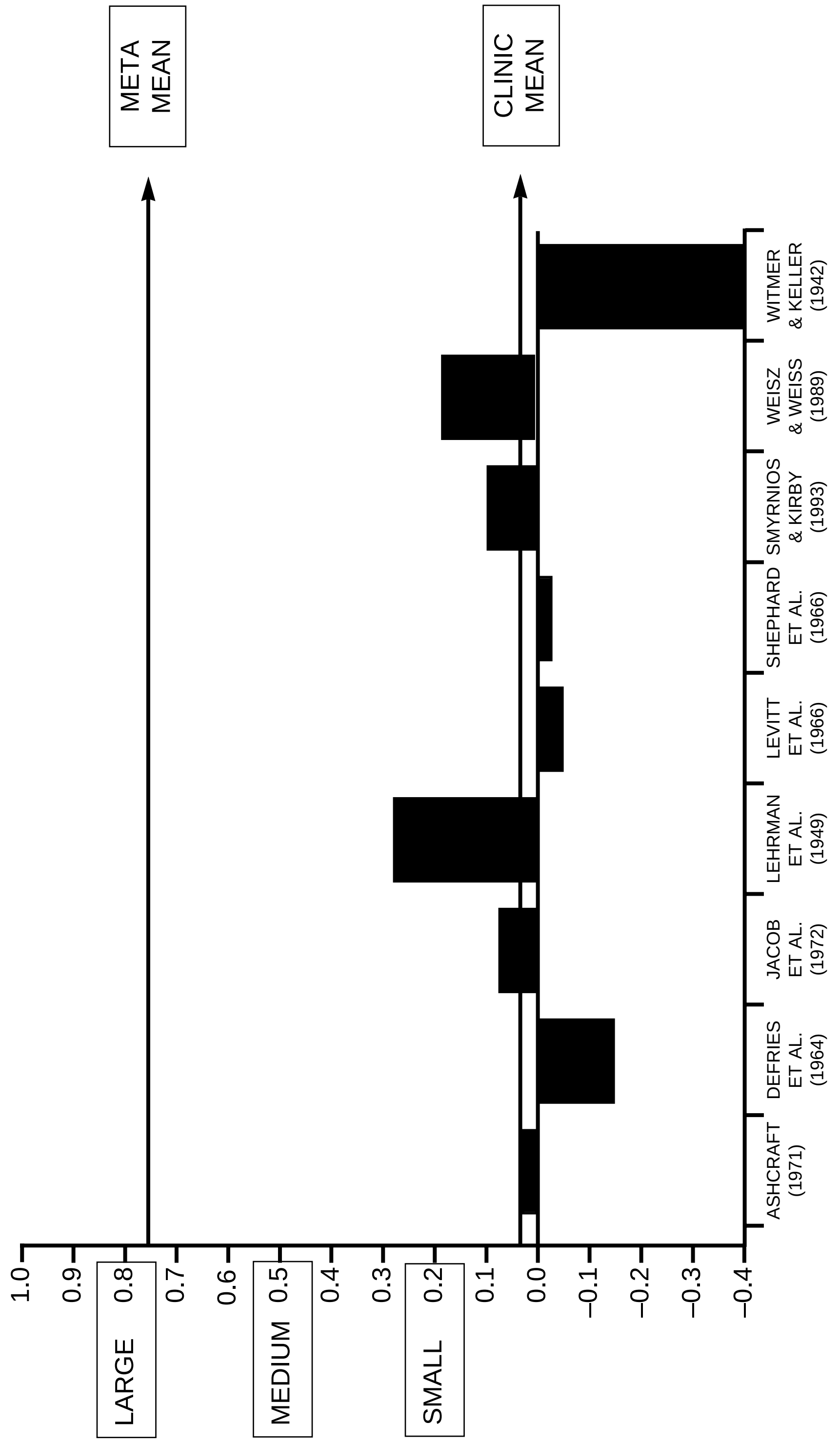

Figure 4 Estimated effect sizes for nine studies of clinic-based psychotherapy with children and adolescents. Horizontal arrows show mean effect size for four broad-based meta-analyses of laboratory outcome studies (top), and averaging across the nine clinic-based studies (bottom). ("Bridging the gap between laboratory and clinic in child and adolescent psychotherapy," by J. R. Weisz, G. R. Donenberg, S. S. Han, and B. Weiss, 1995, *Journal of Consulting and Clinical Psychology*, *63,* pp. 688–701. Copyright 1995 by the American Psychological Association. Reprinted with permission.)

maintenance organized (HMOs). Evidence is needed on outcomes in these forms of practice to ascertain how practice effects in the 1990s compare to outcomes of laboratory interventions. Moreover, without such information, there is a lack of baseline data for tracking the impact of potentially critical changes in service patterns, such as the introduction of managed care, or implementation of empirically validated treatments (discussed below).

Second, assuming that further research continues to show poor effects of conventional clinical treatment, it is necessary to identify factors that explain why therapy in clinical trials experiments produces strong positive effects and therapy in clinics does not. The author and colleague have made two attempts to address this issue (Weisz, Donenberg, Han, & Kauneckis, 1995; Weisz, Donenberg, Han, & Weiss, 1995). In both, they used meta-analytic data sets to investigate which, if any, of the factors that distinguish research therapy from clinic therapy (e.g., of those shown in Table 1) might account for significant variance in outcome. In a 1995 effort, using the most complete sample (Weisz, Donenberg, Han, & Weiss, 1995), we examined eight potentially relevant factors (from the list shown in Table 1), and two were found that were significantly related to treatment outcome: First, behavioral treatments were associated with better outcomes than nonbehavioral treatments, and second, analog cases showed better outcomes than clinic-referred children. The first finding suggests the possibility that clinic-based treatment might be more effective if more behavioral treatments were used; nonbehavioral interventions are widely used in clinical practice, and it appears that behavioral therapies are not the first choice of most practitioners (see Kazdin, Siegel, & Bass, 1990). The second finding suggests that even lab-tested treatment methods may be less successful with truly clinic-referred children than with the less disturbed children who are so often the subjects in lab studies. This issue is discussed in the next section.

5.04.10 EXPORTABILITY OF EMPIRICALLY SUPPORTED TREATMENTS

At first, the array of findings presented above might appear to argue for incorporation of empirically supported treatments into clinical practice. The logic would be this: (i) evidence from clinical trials research with children shows positive effects of numerous treatments; (ii) evidence on conventional treatments in clinical settings shows generally modest-to-negligible effects; thus, (iii) to generate beneficial effects in clinical settings, one should identify those treatments that have been supported in clinical trials and export them to clinics. This logic may be appealing to many who support the efforts of the APA Task Forces (adult and child, see Task Force, 1995) on empirically validated treatments, and the logic may prove to be valid in the long term. However, it may be a mistake to assume that the empirically supported treatments are ready for export. The subjects and treatment conditions involved in tests of the empirically supported treatments tend to differ so much from everyday clinic cases and conditions that it is not clear how workable or effective the supported treatments will be in clinical settings. At the very least, the differences mean that treatment developers cannot simply drop their manuals at the clinic doorstep and assume that beneficial treatment effects will follow.

Table 1 offers several reasons for concern. As one example, the fact that empirically supported treatments have most often been supported with recruited or analog samples rather than referred children may be important. Indeed, in the analysis described above (Weisz, Donenberg, Han, & Weiss, 1995, p. 695) mean ES in even the clinical trials research was found to be significantly lower for studies using clinic-referred children than for those using analog samples. This raises a question as to whether treatments supported in clinical trials with subclinical samples recruited from schools would be equally effective with seriously disturbed children referred to clinics. As a second example, clinical trials studies frequently focus on homogeneous groups, selected for the presence of one or two target problems, and with exclusionary criteria eliminating children who have additional unwanted problems. One cannot be sure that treatments supported with such samples will be equally effective with the heterogeneous groups of multiproblem children frequently seen in clinical practice. As a third example from Table 1, the simple fact that most empirically supported treatments are behavioral (including cognitive-behavioral) may make them difficult to implement in clinics where most of the therapists are strongly psychodynamic, or are unfamiliar with, or even hostile toward, the behavioral perspective.

Beyond the issues suggested by Table 1, numerous practical problems arise. A simple example concerns the number of sessions needed to produce beneficial effects. The author's research in child mental health clinics (Weisz & Weiss, 1989; Weisz, 1996), has found that the average outpatient receives 5–10 sessions. By contrast, most of the manuals used to guide empirically supported treatments require more than a dozen sessions, and these are typically devoted to a single focal problem. Given the specificity of most manuals, children referred for three or four major problems may well be candidates for three or four manuals, each requiring more than 12 sessions. In an era of managed care in the private sector and modest capitated budgets in the public sector, strict session limits are apt to be the rule rather than the exception. Thus, it is fair to ask whether the kinds of manualized treatments that now prevail in the empirical literature can be made to work in clinics with prevailing session limits.

Before manualized treatments derived from the empirical literature are implemented in clinics and clinical training programs, a new genre of treatment outcome research may be needed that involves taking empirically supported treatments out of the laboratories where they were developed and studying them in the crucible of clinical practice. The purpose would be to find out what modifications are needed to make the treatments effective with the clientele, and the real-life constraints, of clinical practice. Several investigators have taken steps in this direction, for example, by treating truly disturbed children in university-based lab clinics (see, e.g., Kendall, 1994; Lovaas, 1987). However, more extensive attempts may be needed to incorporate lab-tested treatments into actual clinical practice, and test their effects, before it can be known just how exportable the experimentally derived treatments are, and what changes will be needed to make them work with seriously disturbed children.

5.04.11 CHILD PSYCHOTHERAPY AND DEVELOPMENTAL PSYCHOLOGY

A third issue that needs attention is the relative isolation of child psychotherapy research from the base of theory and evidence on human development. Despite clear overlap in their populations of interest, and despite their shared emphasis on the study of change, research on developmental psychology and on child psychotherapy have remained surprisingly separate, insular enterprises. Saile et al. (1988) have expressed a similar concern regarding treatments designed to prepare children for medical procedures. In introductions to journal articles on the various child treatments, it is surprisingly unusual to see either theories or empirical findings of developmental psychology cited as a basis for the treatment program. The Results sections of treatment outcome articles are frequently notable for relative inattention to developmental factors (even age!) that may relate to treatment effects.

One cause of this state of affairs may be the very success these relatively adevelopmental treatment efforts have had in producing positive effects. Such success may foster confidence that one is on the right track, and may thus undermine motivation to seek input from other subdisciplines. But treatment effects, although positive in most studies, might well be improved through closer ties to the study of development. Treatment benefit might well be enhanced to the extent that treatments fit developmental characteristics of the treated individuals; a knowledge of relevant developmental literature should enhance treatment planners' capacity to produce a good fit, and thus to test this notion.

In addition to treatment planning, developmental concerns may be relevant to outcome assessment. Because outcome research provides much of the feedback needed to guide the development and refinement of treatment procedures, outcome researchers may need to consider ways to attune such research to developmental questions, so as to stimulate clinical–developmental cross-pollination. Accordingly, this section addresses developmental issues that bear on both child treatment planning and outcome assessment.

5.04.11.1 Age

Most outcome researchers report the age range and mean for their samples, but few assess the potential moderating effect of age. Indeed, the Kazdin et al. (1990) survey of 223 child treatment outcome studies showed that only 7% examined any child, family, and therapist variables in relation to outcome. Age is one of the most easily accessible child characteristics, and within-study assessment of its moderating effects could be useful heuristically, generating promising developmental hypotheses. Broad-based meta-analyses like those described earlier can address the general relation between sample age and outcome, across studies; but this pooled approach lacks the precision and control of within-study analyses, or even more narrowly focused meta-analyses, such as Durlak et al.'s (1991) meta-analysis of cognitive-behavioral therapy outcome studies, discussed earlier. The Durlak et al. findings suggest the potential importance of testing for a moderating role of age in each individual outcome study. On the other hand, it must also be noted that the age variable is, at best, a rough summary index of multiple, diverse developmental factors—cognitive, social, and contextual—each warranting attention in its own right.

5.04.11.2 Cognitions About the Therapy Process and the Role of the Therapist

Adult clients frequently begin treatment with some understanding of the nature and purpose of psychotherapy and its purpose, and some sense of the therapist's role. Children, by contrast, may have little awareness of what therapy is, what a therapist does, or what their relationship with the therapist should be. The concepts children at various developmental levels apply to the process, the therapist, and the relationship, are apt to reflect their cognitive developmental level, in interaction with their previous relevant experience, which will differ markedly from one child to another. Our field needs basic inquiry into children's cognitions about these concepts, each of which could well influence their response to treatment and their ultimate treatment outcome.

5.04.11.3 Ability to "Decenter"

Piaget (e.g., 1929, 1962) wrote extensively about the development of "decentration," the ability to detach from one's own point of view and perceive objects or events from an alternate perspective. Early work on this theme dealt with visual perception, but the notion was eventually extended to social contexts and the ability to recognize the perspective of others on activities, events, and even oneself. A limited ability to decenter may set limits on the impact of treatments (e.g., cognitive-behavioral, psychodynamic) that involve role-playing or other efforts to help children see events, conditions, and even themselves from the perspective of others. Accordingly, research on this cognitive ability as it relates to child therapy could shed new light on how children respond to various perspective-taking intervention techniques.

5.04.11.4 Language Ability

Developmental variations in ability to encode and decode language may set limits on the success of therapy, particularly when it relies heavily on verbal interaction. Low level encoding skills may limit a child's ability to convey thoughts and feelings to the therapist, and thus limit the therapist's ability to tailor interventions to the child's inner state. As a simple example, both cognitive-behavioral and psychodynamic treatments for anxiety may require young clients to describe their anxious state in terms of both physiological arousal (e.g., "I feel tense, and I have a knot in my stomach") and psychological state (e.g., "It feels like everyone is staring at me; if I make a mistake, they'll think I'm stupid"). Children who cannot generate such descriptions may not be able to help their therapist understand their anxious states fully enough to design appropriate interventions.

Limited ability to decode therapist comments may also be a problem, limiting how helpful those comments actually can be. Highly scripted, manual-driven therapies that are language-laden may fall outside the range of young clients' capabilities, thus limiting the impact of treatment. Assessment of children's language skills would facilitate detection of such problems.

Finally, considerable research, dating to the work of Piaget (1955/1923) and Vygotsky (1962/1934), points to developmental differences in the ability to use "private speech" or "inner speech" to inhibit or guide behavior (see also Kohlberg, Yaeger, & Hjertholm, 1968; and Zivin, 1979). Piaget and Vygotsky made different predictions about the developmental course of this phenomenon, and subsequent evidence has been equivocal and dependent on the specific type of inner speech involved (see, e.g., Meichenbaum & Goodman, 1979). However, it seems clear that developmental differences in the use of language to guide behavior may well foster differences in responsiveness to those therapies that stress language as a means of self-control. In cognitive-behavioral therapy, for example, children are taught to use "self-talk" to make themselves less impulsive, less anxious, less depressed, less aggressive, and/or more prosocial. Use of self-talk in these ways assumes a well-developed connection between language and action, an assumption that may need to be tested in child outcome research.

5.04.11.5 Comprehension of Concepts or "Lessons" of the Treatment Program

One cognitive phenomenon that cries out for attention in child psychotherapy research is the child's understanding of the conceptual content and central principles of the therapy program. Many child therapies are, in part, educational programs. For example, some treatment programs for depression are designed to teach children the basic components of depression, plus strategies for alleviating depressed mood; and some programs for anxiety try to teach the building blocks of anxiety (e.g., fear, physiological arousal, habitual avoidance, relief that rewards the avoidance), plus specific techniques for promoting exposure to the feared situation. Are there developmental differences in children's acquisition of these concepts and skills, and might such differences influence treatment outcome? For most child therapies, there is too little relevant evidence to provide an empirically respectable answer. Indeed, one of the most striking limitations of research in this field is the infrequency with which outcome researchers test which "lessons" of their treatment were actually learned. Beyond the value of such information for assessing goodness-of-fit between a treatment program to children at different developmental levels, the information would facilitate assessment of the extent to which grasping the various concepts and lessons is associated with clinical improvement.

5.04.11.6 Abstract Reasoning

Certain kinds of abstract thinking, including hypothetico-deductive reasoning, are consolidated in formal operations, typically in adolescence (Piaget, 1970). Yet some treatment programs for preadolescent children appear to require considerable abstract reasoning. For example, some social skills training programs ask children to generate hypothetical stressful social situations, think of various ways they might respond, envision ways that others might respond to their response, and imagine various possible outcomes of the hypothetical interactions that might ensue. As another example, some therapies for depression require movement from concrete instances to abstract categories when they try to teach children to recognize categories of depressogenic thinking such as overgeneralization and catastrophizing (see examples in Cicchetti, Rogosch, & Toth, in press). Developmentally sensitive investigators could help guide treatment developers, by assessing the extent to which children at various developmental levels can master such seemingly abstract reasoning.

5.04.11.7 Organization and the Orthogenetic Principle

The developmental theorist Heinz Werner (1957) proposed an *orthogenetic principle*, the notion that

> wherever development occurs it proceeds from a state of relative globality and lack of differentiation to a state of increasing differentiation, articulation, and hierarchic integration. (p. 126)

This principle overlaps partly with Piaget's (1970) concept of organization, the notion that the developing system integrates cognitive building blocks into an operating structure. In Piaget's view (supported by evidence), even memories of past experience can be reorganized to integrate such experience with newly developed cognitive structures. One implication of these two principles for child treatment researchers is that some of the skills and information taught in a treatment program may be integrated into the developing child's cognitive system over the course of later development, not necessarily within the time frame of the typical child outcome study.

The implications of this notion could be profound. Perhaps, instead of construing child treatment as a direct induction of change, treatment programs should be seen as perturbations planted in the fertile soil of the individual's developing system and destined to interact with that system over time to produce structural changes in the ways the individual processes and responds to experience. Because the nature and timing of this interactive, unfolding process might be strongly influenced by the natural pace of development, our thinking about the objectives and the timing of outcome assessment might also need to be re-examined. Instead of focusing on proximal outcomes induced by their treatment programs alone and seen at the end of those 8–24 week programs, treatment researchers might focus on therapeutic-developmental exchanges, some effects of which might not be measurable until months or even years after treatment termination. Extending outcome research beyond the traditional follow-up periods (which are now rarely longer than six to nine months) might permit detection of slow-blooming changes or "sleeper effects," as the child's developing system integrates the lessons of therapy with other developmental inputs. Of course, extending follow-up assessments would also help researchers detect a falling off of treatment benefit over time, when that occurs.

5.04.11.8 Looking to the Future

This brief sampling of developmental issues suggests that there is much work to be done. Breaking down the traditional insularity of child psychotherapy research and developmental research, though difficult, may be useful to both enterprises. Benefits for those who study child psychotherapy may include richer and more fully informed models of child dysfunction, readiness for treatment, and intervention-to-child fit, and an expanded picture of how outcome assessment may be conducted and employed in the process of treatment development.

5.04.12 RELATION BETWEEN PROCESS AND OUTCOME

A fourth issue that needs attention is the relation between therapy process and outcome.

In Chapter 3, this volume, Stephen Shirk and Robert Russell describe the state of process research in the child area; the present chapter considers a few basic notions relevant to outcome assessment. The most general point concerns the importance of assessing process in relation to outcome. In the absence of process information, broadly construed, information about treatment outcomes is of limited value. When treatment works well, we need to know what processes were involved, so we can repeat them; and when treatment falls, we need to know what to stop doing. Relatively few child outcome studies provide clear measures of the "active ingredients" of their treatments. The typical clinical trial in the child area provides a rather global comparison of a control group to a treatment program involving multiple techniques. Significant global group differences at post-treatment and follow-up are indeed important elements of treatment validation, and for cost-effectiveness probably the most useful first step in treatment development. However, the lack of specificity regarding which elements of the treatment are producing which aspects of client change has led to calls for a more fine-grained analysis of the processes of child psychotherapy that mediate and/or moderate treatment outcome (Kendall & Morris, 1991; Kazdin, 1989), and other aspects of the therapeutic process deserve attention as well. Kazdin et al. (1990) reported that a fourth of their sample of 223 outcome studies varied some therapy components across groups, and only 2% evaluated outcome in relation to treatment processes; this suggests that there is room for growth in research on what factors over the course of the therapy process actually influence change. There are at least four ways of construing such research, each potentially useful.

5.04.12.1 Components of Treatment

In cases where well-replicated effects show a composite treatment program to be effective, a logical next step is to test the effects of various components of the program, in combinations that make sense theoretically and logically. At least two variations on this theme are well-recognized in the field. A dismantling approach involves breaking a program down into components and varying these across groups; an additive approach involves progressive addition of new components, or combining two or more treatments into one. In the Kazdin et al. (1990) analysis of 223 child treatment outcome studies, 26% involved the first approach, to some extent, and 19% involved the second approach. If these numbers seem surprisingly large, it should be borne in mind that 60% of the Kazdin and co-workers' sample of studies involved comparison of two or more treatments, not necessarily with a no treatment or active control group. Nonetheless, the data suggests that research involving "unpacking" specific treatment components is underway in the child area.

5.04.12.2 Therapist Behavior

A second aspect of therapy process that deserves attention is the behavior and style of the therapist, independent of the specific components of an outline or manual. For example, a good deal of research has focused on therapist directness in communication with clients. In a series of studies, Patterson and his colleagues have noted that therapists' direct instructions to parents of antisocial children to change their parenting style are frequently met with noncompliance in session and resistance to change outside sessions (e.g., Patterson & Chamberlain, 1994; Patterson & Forgatch, 1985). On the other hand, Truax and Wittmer (1973) found evidence that directness by the therapist may have positive effects in confronting defense mechanisms in child therapy. As the findings of these two investigative teams suggest, any one aspect of therapist style or behavior may have differential effects depending on the content and objectives of the treatment; and, of course, therapist style may well interact with client style and personality to shape the ultimate effects. Despite the complexity of the task, and the infinite array of therapist style and behavior dimensions that might be addressed, research on this theme may be well worth the investment of time and intellectual resources it will require. Of all the research discussed in this chapter, it is among the most relevant to the task of training future therapists.

5.04.12.3 Child and Family Behavior

A third dimension of process that requires attention is the behavior of treated children and their families over the course of therapy. Research on this theme is illustrated by the work of Braswell, Kendall, Braith, Carey, and Vye (1985) and Gorin (1993) indicating that positive treatment outcomes are associated with clients' active participation in therapy. Causality is difficult to nail down in such research, and it is certainly possible that such client behavior as active participation may be either a cause of the ultimate treatment benefit, a signal that the treatment is resonating with the client and producing change, or both. But, causality aside, the identification of child and family behavior during therapy that can predict ultimate outcome can, in principle, provide a much-needed tool for therapists: a means of determining whether the treatment is working or not, and thus whether adjustments are needed, before the entire intervention program has ended.

5.04.12.4 Therapeutic Relationship

The therapeutic relationship, or working alliance, has been construed as involving two interrelated parts: the client's positive emotional connection to the therapist, and a shared conceptualization between the client and therapist of the tasks and goals of therapy (Bordin, 1979). In the adult literature, development of a therapeutic relationship has emerged as a particularly significant process correlate of positive outcome in several studies (Horowitz, Marmar, Weiss, DeWitt, & Rosenbaum, 1984; Luborsky, Crits-Christoph, Mintz, & Auerbach, 1988). Shirk and Saiz (1992) have argued that this process variable may be an even more significant contributor to outcome for children due to the "involuntary client" status of many children at the beginning of therapy, the nonverbal nature of many forms of client-centered and play therapy for children, and the social deficits that are hypothesized to be central in the development and maintenance of many serious child problems (e.g., aggression).

Child process researchers have studied the therapeutic relationship in play therapy (Howe & Silvern, 1981; Truax, Altmann, Wright, & Mitchell, 1973), family therapy (Friedlander, Wildman, Heatherington, & Skowron, 1994; Pinsoff & Catherall, 1986), individual psychodynamic therapy (Shirk & Saiz, 1992), cognitive-behavioral treatment (Kendall, 1994), child behavior therapy (Motta & Lynch, 1990; Motta & Tobin, 1992), and parent training (Webster-Stratton & Herbert, 1993). However, unlike the consistently positive findings in the adult process literature relating the working alliance to therapeutic success, evidence from child psychotherapy has been mixed, with Shirk & Saiz (1992) reporting positive associations between their measures of the therapeutic relationship and outcome in child psychodynamic psychotherapy, and both Kendall (1994) and Motta (Motta & Lynch, 1990; Motta & Tobin, 1992) finding no significant association between relationship quality and outcome with cognitive-behavioral treatment for anxiety and behavioral therapy for learning and disruptive behavior problems, respectively. The lack of agreement across studies is difficult to interpret, given the study-to-study differences in the way therapeutic relationship is assessed. The field needs a well-validated set of measures for assessing the relationship. Developing such measures for child therapy will be a challenging task. Among the difficulties confronted, two are particularly notable: (i) the complexity of the relationship concepts involved relative to the limited verbal and conceptual facility of many of the children who will be the targets of the assessment; and (ii) the need to encompass not only the child–therapist dyad, but also the parents and possibly additional family members, whose relationship to therapist and child may be critical to the success of treatment.

5.04.13 CHILD AND FAMILY ETHNICITY AND CULTURE

A fifth issue is that of culture and ethnicity in relation to child treatments. In the great majority of child outcome studies, the samples appear to be predominantly Caucasian, although authors have frequently failed to be explicit about sample composition. In the Kazdin et al. (1990) survey of 223 child outcome studies, 80% failed to identify the ethnic composition of their sample. While mere reporting of the relevant numbers is rare, actual tests of ethnicity as a moderator of outcome are even rarer. Some have suggested that treatments tested primarily with mainstream samples may not necessarily be optimal for troubled members of ethnic minority groups (Gibbs & Huang, 1989; Rogler, Malgady, Costantino, & Blumenthal, 1987; Spurlock, 1985; Sue, 1977; Tharp, 1991). The treatments may not, for example, take into account the language, values, customs, child-rearing traditions, expectancies for child and parent behavior, and distinctive stressors and resources associated with different cultural groups. In the treatment setting, such cultural factors may lead to miscommunication and misunderstanding between the therapist and the client and family, thus increasing the likelihood of premature termination and poor treatment outcome (Ho, 1992; Sue & Zane, 1987). Tharp (1991) has also suggested that even the therapy modality requires attention, and that family and group interventions may be more appropriate than individual treatment for many minority youth, and more likely, as well, to generate information for the therapist on cultural issues relevant to the child, family, and community.

The literature on culture and psychotherapy is rich in recommendations for how to treat specific ethnic groups, but poor in controlled empirical assessment on these issues, with most recommendations based on anecdotal and experiential reports (see Gibbs & Huang, 1989; Ho, 1992). The array of hypotheses now available should provide fertile ground for experimentation in the future. For the present, though, a modest base of evidence is available to suggest trends that bear further study. Weisz, Huey, and Weersing (1996) identified 19 treatment outcome studies (most found in the meta-analytic collections from Weisz et al., 1987 and Weisz, Weiss, et al., 1995) in which the majority of the sample were ethnic minority children or families. A few of the findings of these studies are considered below, and four key questions are raised about treatment outcome and culture that need attention in future research.

5.04.13.1 How Effective are Tested Treatments with Ethnic Minority Children?

In general, children in the 19 studies improved significantly more than children in control conditions. Treatment showed a significant beneficial effect on externalizing problems in 79% of the group comparisons, and for internalizing problems the figure was 86%. On other outcomes, such as self-esteem, social skills, and family functioning, 46% of the comparisons showed significant treatment benefit. These figures appear somewhat more modest than the general trends seen in meta-analyses that apparently included predominantly Caucasian majority samples, but such a comparison is problematic because of numerous study-to-study differences in treated problems, types of treatment, and so forth. Direct comparison is far superior, but only one such comparison was identified in the 19 studies; Henggeler, Melton, and Smith (1992) found that African-American and Caucasian delinquents responded equally well to multisystemic treatment.

5.04.13.2 Is there Evidence of Treatment Type and Ethnic Group Interaction?

In a word, no. But this may be only because no complete test of such an interaction has been carried out. In the author's sample of 19 minority youth outcome studies, one study involved an initial step toward such a test. Szapocznik and colleagues (Szapocznik et al., 1989; Szapocznik, Kurtines, Santisteban, & Rio, 1990) tested the efficacy of structural family therapy (SFT) against individual psychodynamic child therapy (IPCT) with Hispanic families of boys with behavioral and emotional problems. Following treatment, SFT-treated children had fewer behavior problems than IPCT children, but neither group differed significantly from children in a recreational control condition. At follow up, family functioning improved substantially following SFT relative to IPCT; indeed, with IPCT, family functioning actually worsened over time. The findings seem consistent with Tharp's (1991) contention that family-based therapies are more appropriate than individual treatment for minority children. However, Szapocznik et al. did not include comparison groups of non-Hispanic youth treated with SFT and IPCT. Thus, it appears that the field still awaits a full test of treatment type and ethnicity interaction.

5.04.13.3 Outcomes of Therapist–Client Matching for Ethnicity or Language

Although nine of the 19 studies in the collection involved some attempt at ethnicity and/or language match, none provided a direct test of the impact of the match, for example, by comparing matched and unmatched therapist–child pairs. Hayes, Cunningham, and Robinson (1977) reported that counselor race (African-American vs. Caucasian) did not appear to influence treatment outcomes with test anxiety and poor school motivation among African-American children, but they did not present substantiating analyses. In general, those studies in which some form of matching was carried out did not seem to show notably better success rates than the remainder of the 19. Again, such global comparison across studies lacks the precision of the direct, within-study tests that are most needed to address the "matching" question.

5.04.13.4 Outcomes of Adapting Therapies Specifically for Ethnic Minority Children

Ten of the 19 studies involved some sort of adaptation of the treatment program to cultural characteristics of the minority sample. The adaptations ranged from narrowly focused changes such as depicting minority figures in modeling tasks to such extensive change as essentially designing entire programs for a particular ethnic minority group. The fact that only one of the "unadapted" studies focused on internalizing problems ruled out comparisons on that dimension; however, the percentage of comparisons showing significant positive treatment effects was somewhat higher for externalizing problems (83% vs. 75%) and for problems in the "other" category (57% vs. 33%).

Direct comparisons of adapted and unadapted methods within the same study are difficult to find. A study by Constantino, Malgady, and Rogler (1986) did find that maladjusted Puerto Rican children in a culturally adapted modeling program involving cuento or folklore therapy showed less aggressive behavior following treatment than those in an art/play therapy control condition (APT). However, since APT subjects were even more aggressive than those in the no-therapy (NT) condition, the findings may reflect the ineffectiveness of APT as much as the superiority of cuento therapy. A more robust test of any culture-specific treatment approach would entail comparison with another empirically supported approach (e.g., social skills training for aggression). Where culture-calibrated changes have been made to a standard treatment program, it will be important to carry out direct comparison of outcomes for minority children receiving the altered program vs. outcome for minority children receiving the standard program.

Generalizing across the various themes and questions reviewed above, the state of affairs regarding culture and child psychotherapy can be summarized succinctly: there are intriguing hypotheses and important questions that need attention, but the numerous direct comparisons needed to test the hypotheses and answer the questions are lacking. With concern about culture and mental health growing steadily, it seems likely that the next generation of research will include increased attention to these matters.

5.04.14 CONTEXT SENSITIVITY OF TREATMENTS

A sixth issue arises because the interventions involved in most treatment research are rather removed from the contexts in which children live their lives. Most researchers would agree that children do not develop as solitary beings in a sterile environment, but rather as active participants in complex, multifaceted physical and social systems. Yet most treatment outcome research with children involves interactions with a single therapist, or sometimes with a small group of unfamiliar children, in the sterile environment of the therapist's office or therapy room. Pretherapy assessment and treatment planning typically involve very limited sampling of the child's life circumstances and behavior at home, at school, or with familiar peers. This may limit the capacity of the therapist to fit interventions precisely to conditions and context of the individual child's problems. In the worst cases, the problem may tilt treatment development in the direction of "one-size-fits-all" or "cookie-cutter" therapies.

Numerous theorists and researchers (e.g., Bronfenbrenner, 1979, 1986; Masten, Best, & Garmezy, 1991) have emphasized the context boundedness of development, and discussed implications for adaptation and dysfunction. Others (e.g., Cicchetti & Toth, in press; Mash & Dozois, 1996) have noted the diverse ways in which the child's contexts and ecological systems can influence the development and expression of dysfunctional behavior and emotional states. Still others (e.g., Forehand, Lautenschlager, Faust, & Graziano, 1986; Kazdin, 1989) have noted that even what parents report (e.g., to assessors and therapists) regarding deviance and dysfunction in their children can be influenced by such diverse factors as parental psychopathology, marital discord, stress in the home, and even an intent to conceal harmful parental practices (e.g., abuse or neglect). Finally, it seems self-evident that the impact of psychotherapy with children may vary depending on the extent to which significant others in the child's contexts (e.g., parents, teachers) are involved and supportive of the process. The power of all the influences noted in this paragraph may be felt disproportionately in childhood, in part because children have such limited ability to select the contexts in which their development unfolds.

The message for treatment planners and treatment outcome researchers is that contextual factors and key individuals in the child's social environment (e.g., parents, teachers, siblings, peers) may need to figure significantly in pretreatment assessment, in treatment planning, in treatment delivery, and in outcome assessment. In general, these steps have only been taken in very limited and tentative ways in child psychotherapy research to date. There are some exceptions to this generalization, however, and noting a few of these may help illustrate what is possible in future child psychotherapy research.

In one approach, Lewinsohn and colleagues (see Lewinsohn, Clarke, Hops, & Andrews, 1990; Lewinsohn, Rohde, Clarke, Hops, & Seeley, 1994; Rohde, Lewinsohn, & Seeley, 1994) have created a parent counterpart to their Adolescent Coping with Depression group intervention. The objective is to promote parental understanding and acceptance of what is being taught to the adolescents, and to reduce family conflict by teaching parents some of the same communication and problem-solving skills their adolescents are learning. In another approach, such investigators as Szapocznik and colleagues (e.g., Szapocznik et al., 1989) treat problem behavior in Hispanic boys by means of structural family therapy. Webster-Stratton and colleagues (e.g., Webster-Stratton, Kolpacoff, & Hollinsworth, 1988) have developed a self-administered videotape therapy for families with conduct-problem children. Henggeler and colleagues (e.g., Henggeler, Melton, & Smith, 1992; Borduin et al., 1995), in perhaps the most context-sensitive approach yet developed, send therapists into the settings where juvenile offenders live their lives, working with them to develop treatments tailored to the strengths and limitations of their family, school, peer group, and neighborhood. Each of the intervention approaches cited here has shown positive effects relative to control groups, although Lewinsohn and colleagues have not yet found that adding parent training alone to their adolescent intervention improves benefits over and above treatment of adolescents only. Perhaps procedures that delve more deeply into the child's social systems are required to produce benefits greater than those of individually administered interventions.

Moving in the directions illustrated by these research teams certainly will complicate the work of treatment development and outcome assessment. On the other hand, such efforts seem essential if we are to assess the benefits of context-sensitive treatment. We need to find out what can be gained if we push beyond a narrow focus on child characteristics and toward a broader focus on potentiating and inhibiting forces in the social systems within which the child is growing up.

ACKNOWLEDGMENTS

The research program described in this chapter was supported through a Research Scientist Award (K05 MH01161) and research grants (R01 MH34210, R01 MH38450, R01 MH38240) from the National Institute of Mental Health, which I gratefully acknowledge. I also appreciate the important contributions made to this work by faculty and graduate student colleagues, clinic administrators and therapists, and the many children and parents who have participated in the research before, during, and after treatment.

5.04.15 REFERENCES

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101,* 213–232.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: American Psychiatric Association Press.

Ashcroft, C. W. (1971). The later school adjustment of treated and untreated emotionally handicapped children. *Journal of School Psychology, 9,* 338–342.

Baer, R. A., & Nietzel, M. T. (1991). Cognitive and behavioral treatment of impulsivity in children: A meta-analytic review of the outcome literature. *Journal of Clinical Child Psychology, 2,* 400–412.

Bickman, L. (1996). A continuum of care: More is not always better. *American Psychologist, 51,* 689–701.

Bickman, L., Guthrie, P. R., Foster, E. M., Lambert, E. W., Summerfelt, W.T., Breda, C. S., & Heflinger, C. A. (1995). *Evaluating managed mental health services: The Fort Bragg experiment.* New York: Plenum.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research, and Practice, 16,* 252–260.

Borduin, C.M., Mann, B. J., Cone, L. T., Henggeler, S. W., Fucci, B. R., Blaske, D., & Williams, R. A. (1995). Multisystemic treatment of serious juvenile offenders: Long-term prevention of criminality and violence. *Journal of Consulting and Clinical Psychology, 63,* 569–578.

Braswell, L., Kendall, P. C., Braith, J., Carey, M. P., & Vye, C. S. (1985). "Involvement" in cognitive-behavioral therapy with children: Process and its relationship to outcome. *Cognitive Therapy and Research, 9,* 611–630.

Bronfenbrenner, U. (1979). *The ecology of human development.* Cambridge, MA: Harvard University Press.

Bronfenbrenner, U. (1986). Ecology of the family as a context for human development. *Developmental Psychology, 22,* 723–742.

Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin, 98,* 388–400.

Chalmers, T. C., Smith, H., Blackburn, B., Silverman, B., Schroeder, B., Reitman, D., & Ambroz, A. (1981). A method for assessing the quality of a randomized clinical trial. *Controlled Clinical Trials, 2,* 31–49.

Cicchetti, D., & Toth, S. L. (in press). Transactional ecological systems in developmental psychopathology. In S. S. Luthar, J. A. Burack, D. Cicchetti, & J. R. Weisz (Eds.), *Developmental psychopathology: Perspectives on adjustment, risk, and disorder.* New York: Cambridge University Press.

Cicchetti, D., Rogosch, F. A., & Toth, S. L. (in press). Ontogenesis, depressotypic organization, and the depressive spectrum. In S. S. Luthar, J. A. Burack, D. Cicchetti, & J. R. Weisz (Eds.), *Developmental psychopathology: Perspectives on adjustment, risk, and disorder.* New York: Cambridge University Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159.

Costantino, G., Malgady, R., & Rogler, L. (1986). Cuento therapy: a culturally sensitive modality for Puerto Rican children. *Journal of Consulting and Clinical Psychology, 54,* 639–645.

Costello, E. J. (1989). Developments in child psychiatric epidemiology. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 836–841.

DeFries, Z., Jenkins, S., & Williams, E. C. (1964). Treatment of disturbed children in foster care. *American Journal of Orthopsychiatry, 34,* 615–624.

Durlak, J. A., Fuhrman, T., & Lampman, C. (1991). Effectiveness of cognitive-behavior therapy for maladapting children: A meta-analysis. *Psychological Bulletin, 110,* 204–214.

Dush, D. M., Hirt, M. L., & Schroeder, H. E. (1989). Self-statement modification in the treatment of child behaviour disorders: A meta-analysis. *Psychological Bulletin, 106,* 97–106.

Evans, M. E., Armstrong, M. I., Dollard, N., Kuppinger, A. D., Huz, S., & Wood, V. M. (1994). Development and evaluation of treatment foster care and family-centered intensive case management in New York. *Journal of Emotional and Behavioral Disorders, 2,* 228–239.

Forehand, R., Lautenschlager, G. J., Faust, J., & Graziano, W. G. (1986). Parent perceptions and parent–child interactions in clinic-referred children: A preliminary investigation of the effects of maternal depressive moods. *Behavior Research and Therapy, 24,* 73–75.

Frank, J. D. (1973). *Persuasion and healing: A comparative study of psychotherapy.* Baltimore: Johns Hopkins University Press.

Friedlander, M. L., Wildman, J., Heatherington, L., & Skowron, E. A. (1994). What we do and don't know about the process of family therapy. *Journal of Family Psychology, 8,* 390–416.

Gibbs, J. T., & Huang, L. N. (1989). *Children of color: Psychological interventions with minority children.* San Francisco: Jossey-Bass.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Gorin, S. S. (1993). The prediction of child psychotherapy outcome: Factors specific to treatment. *Psychotherapy, 30,* 152–158.

Hayes, E. J., Cunningham, G. K., Robinson, J. B. (1977). Counseling focus: Are parents necessary? *Elementary School Guidance and Counseling, 12,* 8–14.

Hazelrigg, M. D., Cooper, H. M., & Borduin, C. M. (1987). Evaluating the effectiveness of family therapies: An integrative review and analysis. *Psychological Bulletin, 101,* 428–442.

Henggeler, S. W., Melton, G. B., & Smith, L. A. (1992). Family preservation using multisystemic therapy: An effective alternative to incarcerating serious juvenile offenders. *Journal of Consulting and Clinical Psychology, 60,* 953–961.

Ho, M. K. (1992).*Minority children and adolescents in therapy.* Newbury Park, CA: Sage.

Horowitz, M. J., Marmar, C. R., Weiss, D. S., DeWitt, K. N., & Rosenbaum, R. L. (1984). Brief psychotherapy of bereavement reactions: The relationship of process to outcome. *Archives of General Psychiatry, 41,* 438–448.

Howe, P. A., & Silvern, L. E. (1981). Behavioral observation of children during play therapy: Preliminary development of a research instrument. *Journal of Personality Assessment, 45,* 168–182.

Institute of Medicine (1989). *Research on children and adolescents with mental, behavioral, and developmental disorders.* Washington, DC: National Academy Press.

Jacob, T., Magnussen, M. G., & Kemler, W. M. (1972). A follow-up of treatment terminators and remainers with short-term and long-term symptom duration. *Psychotherapy: Theory, Research, and Practice, 9,* 139–142.

Kazdin, A. E. (1988). *Child psychotherapy: Developing and identifying effective treatments.* Elmsford, NY: Pergamon.

Kazdin, A. E. (1989). Developmental psychopathology: Current research, issues, and directions. *American Psychologist, 44,* 180–187.

Kazdin, A. E., Bass, D., Ayers, W. A., & Rodgers, A. (1990). Empirical and clinical focus of child and adolescent psychotherapy research. *Journal of Consulting and Clinical Psychology, 58,* 729–740.

Kazdin, A. E., Siegel, T. C., & Bass, D. (1990). Drawing on clinical practice to inform research on child and adolescent psychotherapy: Survey of practitioners. *Professional Psychology: Research and Practice, 21,* 189–198.

Kendall, P. C. (1994). Treating anxiety disorders in children: Results of a randomized clinical trial. *Journal of Consulting and Clinical Psychology, 62,* 100–110.

Kendall, P. C., & Morris, R. J. (1991). Child therapy: Issues and recommendations. *Journal of Consulting and Clinical Psychology, 59,* 777–784.

Kohlberg, L., Yaeger, J., & Hjertholm, E. (1968). The development of private speech: Four studies and a review of theories. *Child Development, 39,* 691–736.

Lehman, A. E., Postrado, L. T., Roth, D., McNary, S. W., & Goldman, H. H. (1994). Continuity of care and client outcomes in the Robert Wood Johnson Foundation program on chronic mental illness. *Milbank Quarterly, 72,* 105–122.

Lehrman, L. J., Sirluck, Black, B. J., & Glick, S. J. (1949). *Success and failure of treatment of children in the child guidance clinics of the Jewish Board of Guardians, New York City.* Jewish Board of Guardians Research Monographs, No. 1.

Levitt, E. E., Beiser, H. R. & Robertson, R. E. (1959). A follow-up evaluation of cases treated at a community child guidance clinic. *American Journal of Orthopsychiatry, 29,* 337–347.

Lewinsohn, P. M., Clarke, G. N., Hops, H., & Andrews, J. A. (1990). Cognitive-behavioral treatment for depressed adolescents. *Behavior Therapy, 21,* 385–401.

Lewinsohn, P. M., Rohde, P., Clarke, G. N., Hops, H., & Seeley, J. R. (1994). *Cognitive-behavioral treatment for depressed adolescents: Treatment outcome and the role of parental involvement.* Unpublished manuscript. Eugene, OR: Oregon Research Institute.

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55,* 3–9.

Luborsky, L., Crits-Christoph, P., Mintz, J., & Auerbach, A. (1988). *Who will benefit from psychotherapy*? New York: Basic Books.

Mann, C. (1990). Meta-analysis in the breech. *Science, 249,* 476–480.

Mash, E. J., & Dozois, D. J. A. (1996). Child psychopathology: A developmental-systems perspective. In E. J. Mash & R. A. Barkley (Eds.), *Child psychopathology* (pp. 3–60). New York: Guilford Press.

Masten, A., Best, K., & Garmezy, N. (1991). Resilience and development: Contributions from the study of children who overcome adversity. *Development and Psychopathology, 2,* 425–444.

McGrath, M. L., Dorsett, P. G., Calhoun, M. E., & Drabman, R. S. (1988). "Beat-the-buzzer": A method for decreasing parent–child morning conflicts. *Child and Family Behavior Therapy, 9,* 35–48.

Meichenbaum, D., & Goodman, S. (1979). Clinical use of private speech and critical questions about its study in natural settings. In G. Zivin (Ed.), *The development of self-regulation through private speech* (pp. 325–360). New York: Wiley.

Motta, R. W., & Lynch, C. (1990). Therapeutic techniques vs. therapeutic relationships in child behavior therapy. *Psychological Reports, 67,* 315–322.

Motta, R. W., & Tobin, M. I. (1992). The relative importance of specific and nonspecific factors in child behavior therapy. *Psychotherapy in Private Practice, 11,* 51–61.

Nicholson, R. A., & Berman, J. S. (1983). Is follow-up necessary in evaluating psychotherapy? *Psychological Bulletin, 93,* 261–278.

Office of Technology Assessment (1986). *Children's mental health: Problems and services—A background paper* (Publication No. OTA-BP-H-33). Washington, DC: US Government Printing Office.

Patterson, G. R., & Chamberlain, P. (1994). A functional analysis of resistance during parent training therapy. *Clinical Psychology: Science and Practice, 1,* 53–70.

Patterson, G. R., & Forgatch, M. S. (1985). Therapist behavior as a determinant for client noncompliance: A paradox for the behavior modifier. *Journal of Consulting and Clinical Psychology, 53,* 846–851.

Pelham, W. E., Carlson, C., Sams, S., Vallano, G., Dixon, J., & Hoza, B. (1993). Separate and combined effects of methylphenidate and behavior modification on boys with attention-deficit-hyperactivity disorder in the classroom. *Journal of Consulting and Clinical Psychology, 55,* 76–85.

Piaget, J. (1929). *The child's conception of the world.* Totowa, NJ: Littlefield, Adams.

Piaget, J. (1955). *The language and thought of the child.* New York: Meridian. (Originally published 1923).

Piaget, J. (1962). *Play, dreams, and imitation.* New York: Norton.

Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology* (Vol. 1, pp. 703–32). New York: Wiley.

Pinsoff, W. M., & Catherall, D. R. (1986). The integrative psychotherapy alliance: Family, couple, and individual therapy scales. *Journal of Marital and Family Therapy, 12,* 137–151.

Prout, H. T., & DeMartino, R. A. (1986). A meta-analysis of school-based studies of psychotherapy. *Journal of School Psychology, 24,* 285–292.

Roberts, M. C. (1994). Models for service delivery in children's mental health: Common characteristics. *Journal of Clinical Child Psychology, 23,* 212–219.

Rogler, L. H., Malgady, R. G., Costantino, G., & Blumenthal, R. (1987). What do culturally sensitive mental health services mean?: The case of Hispanics. *American Psychologist, 42*(6), 565–575.

Rohde, P., Lewinsohn, P. M., & Seeley, J. R. (1994). Responses of depressed adolescents to cognitive-behavioral treatment: Do differences in initial severity clarify the comparison of treatments? *Journal of Consulting and Clinical Psychology, 62,* 851–854.

Russell, R. L., Greenwald, S., & Shirk, S. R. (1991). Language change in child psychotherapy: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 59,* 916–919.

Saile, H., Burgmeier, R., & Schmidt, L. R. (1988). A meta-analysis of studies on psychological preparation of children facing medical procedures. *Psychology and Health, 2,* 107–132.

Shadish, W. R., Montgomery, L. M., Wilson, P., Wilson, M. R., Bright, I, & Okwumabua, T. (1993). Effects of family and marital psychotherapies: A meta-analysis. *Journal of Consulting and Clinical Psychology, 61,* 992–1002.

Shapiro, D. A., & Shapiro, D. (1982). Meta-analysis of comparative therapy outcome studies: A replication and refinement. *Psychological Bulletin, 92,* 581–604.

Shepherd, M., Oppenheim, A. N., & Mitchell, S. (1966). Childhood behavior disorders and the child-guidance clinic: An epidemiological study. *Journal of Child Psychology and Psychiatry, 1,* 39–52.

Shirk, S. R., & Saiz, C. S. (1992). Clinical, empirical, and developmental perspectives on the therapeutic relationship in child psychotherapy. *Development and Psychopathology, 4,* 713–728.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32,* 752–760.

Smith, M. L., Glass, G. V., & Miller, T. L. (1980). *Benefits of psychotherapy.* Baltimore: Johns Hopkins University Press.

Smyrnios, K. X., & Kirkby, R. J. (1993). Long-term comparison of brief versus unlimited treatments with children and their parents. *Journal of Consulting and Clinical Psychology, 61,* 1020–1027.

Spurlock, J. (1985). Assessment and therapeutic intervention of Black children. *Journal of the American Academy of Child Psychiatry, 24*(2), 168–174.

Stroul, B. A., & Friedman, R. (1986). *A system of care for children and youth with severe emotional disturbances* (Rev. ed.). Washington, DC: Georgetown University Child Development Center, CASSP Technical Assistance Center.

Sue, S. (1977). Community mental health services to minority groups: Some optimism, some pessimism. *American Psychologist, 32,* 616–624.

Sue, S., & Zane, N. (1987). The role of culture and cultural techniques in psychotherapy: A critique and reformulation. *American Psychologist, 42*(1), 37–45.

Szapocznik, J., Kurtines, W., Santisteban, D. A., & Rio, A. T. (1990). Interplay of advances between theory, research, and application in treatment interventions aimed at behavior problem children and adolescents. *Journal of Consulting and Clinical Psychology, 58,* 696–703.

Szapocznik, J., Rio, A., Murray, E., Cohen, R., Scopetta, M., Rivas-Vasquez, A., Hervis, O., Posada, V., & Kurtines, W. (1989). Structural family vs. psychodynamic child therapy for problematic Hispanic boys. *Journal of Consulting and Clinical Psychology, 57,* 571–578.

Tarnowski, K. J., Rosen, L. A., McGrath, M. L., & Drabman, R. S. (1987). A modified habit reversal procedure in a recalcitrant case of trichotillomania. *Journal of Behavior Therapy and Experimental Psychiatry, 18,* 157–163.

Task Force on Promotion and Dissemination of Psychological Procedures, Division of Clinical Psychology, American Psychological Association (1995, Winter). Training in and dissemination of empirically validated psychological treatments: Report and Recommendations. *The Clinical Psychologist, 48,* 3–24.

Tharp, R. G. (1991). Cultural diversity and treatment of children. *Journal of Consulting and Clinical Psychology, 59*(6), 799–812.

Truax, C. B., Altmann, H., Wright, L., & Mitchell, K. M. (1973). Effects of therapeutic conditions in child therapy. *Journal of Community Psychology, 1,* 313–318.

Truax, C. B., & Wittmer, J. (1973). The degree of therapist focus on defense mechanisms and the effect on therapeutic outcome with Institutionalized juvenile delinquents. *Journal of Community Psychology, 1,* 201–203.

Vygotsky, L. (1962). *Thought and language.* Cambridge, MA: MIT Press. (Originally published 1934).

Webster-Stratton, C., & Herbert, M. (1993). "What really happens in parent training?" *Behavior Modification, 17,* 407–456.

Webster-Stratton, C., Kolpacoff, M., & Hollinsworth, T. (1988). Self-administered videotape therapy for families with conduct-problem children: Comparison with two cost-effective treatments and a control group. *Journal of Consulting and Clinical Psychology, 56,* 558–566.

Weiss, B., & Weisz, J. R. (1990). The impact of methodological factors on child psychotherapy outcome research: A meta-analysis for researchers. *Journal of Abnormal Child Psychology, 18,* 639–670.

Weisz, J. R. (1996, in progress). *Studying clinic-based child mental health care.* Ongoing research project, University of California at Los Angeles.

Weisz, J. R., Donenberg, G. R., Han, S. S., & Kauneckis, D. (1995). Child and adolescent psychotherapy outcomes in experiments vs. clinics: Why the disparity? *Journal of Abnormal Child Psychology, 23,* 83–106.

Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between lab and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63*(5), 688–701.

Weisz, J. R., Huey, S. J., & Weersing, V. R. (in press). *Psychotherapy outcome research with children and adolescents.* In T. H. Ollendick & R. J. Prinz (Eds.), *Advances in Clinical Child Psychology.* New York: Plenum.

Weisz, J. R., McCarty, C. A., Eastman, K. L., Chaiyasit, W., & Suwanlert, S. (1997). Developmental psychopathology and culture: Ten lessons from Thailand. In S. S. Luthar, J. Burack, D. Cicchetti, & J. R. Weisz (Eds.), *Developmental psychopathology: Perspectives on adjustment, risk, and disorder* (pp. 568–592). New York: Cambridge University Press.

Weisz, J. R., Suwanlert, S., Chaiyasit, W., Weiss, B., Walter, B. R., & Anderson, W. (1988). Thai and American perspectives on over- and undercontrolled child behavior problems: Exploring the threshold model among parents, teachers, and psychologists. *Journal of Consulting and Clinical Psychology, 56,* 601–609.

Weisz, J. R., & Weiss, B. (1989). Assessing the effects of clinic-based psychotherapy with children and adolescents. *Journal of Consulting and Clinical Psychology, 57,* 741–746.

Weisz, J. R., & Weiss, B. (1991). Studying the referability of child clinical problems. *Journal of Consulting and Clinical Psychology, 59,* 266–273.

Weisz, J. R., & Weiss, B. (1993). *Effects of psychotherapy with children and adolescents.* Newbury Park, CA: Sage.

Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology, 55,* 542–549.

Weisz, J. R., Weiss, B., & Donenberg, G. R. (1992). The lab vs. the clinic: Effects of child and adolescent psychotherapy. *American Psychologist, 47,* 1578–1585.

Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A., & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin, 117,* 450–468.

Werner, H. (1957). The concept of development from a comparative and organismic point of view. In D. Harris (Ed.), *The concept of development.* Minneapolis: University of Minnesota Press.

Wilson, G. T. (1985). Limitations of meta-analysis in the evaluation of the effects of psychological therapy. *Clinical Psychology Review, 5,* 35–47.

Witmer, H. L., & Keller, J. (1942). Outgrowing childhood problems: A study of the value of child guidance treatment. *Smith College Studies in Social Work, 13,* 74–90.

Wurtele, S. K., & Drabman, R. S. (1984). "/Beat-the-buzzer" for classroom dawdling: A one-year trial. *Behavior Therapy, 15,* 403–409.

Zivin, G. (Ed.) (1979). *The development of self-regulation through private speech.* New York: Wiley.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.05

Applied Behavior Analysis

ALAN HUDSON
Royal Melbourne Institute of Technology, Bundoora, Vic, Australia

5.05.1 THEORETICAL UNDERPINNINGS 108

5.05.2 HISTORY OF THE APPROACH 110

5.05.3 ASSESSMENT AND CLINICAL FORMULATION 111

- *5.05.3.1 Specification of Behavior* 111
- *5.05.3.2 Measurement of Behavior* 111
- *5.05.3.3 Assessment of the Function of Behavior* 112
  - *5.05.3.3.1 Interviews* 113
  - *5.05.3.3.2 Written materials* 113
  - *5.05.3.3.3 Direct observation of the behavior* 113
  - *5.05.3.3.4 Developing a hypothesis regarding the function of the behavior* 116

5.05.4 DESCRIPTION OF INTERVENTION PROCEDURES AND PROPOSED MECHANISMS OF CHANGE 116

- *5.05.4.1 Procedures for Strengthening Behavior* 116
  - *5.05.4.1.1 Systematic use of reinforcement* 116
  - *5.05.4.1.2 Prompting* 118
  - *5.05.4.1.3 Shaping* 118
  - *5.05.4.1.4 Chaining* 118
  - *5.05.4.1.5 Modeling* 119
- *5.05.4.2 Procedures for Weakening Behavior* 119
  - *5.05.4.2.1 Differential reinforcement of incompatible behavior* 119
  - *5.05.4.2.2 Differential reinforcement of alternative behavior* 120
  - *5.05.4.2.3 Differential reinforcement of other behavior* 120
  - *5.05.4.2.4 Extinction* 120
  - *5.05.4.2.5 Timeout* 121
  - *5.05.4.2.6 Overcorrection* 122
  - *5.05.4.2.7 Punishment* 122
  - *5.05.4.2.8 The principle of the least restrictive alternative* 123
- *5.05.4.3 Complex Procedures for Changing Behavior* 123
  - *5.05.4.3.1 Token economies* 123
  - *5.05.4.3.2 Behavioral contracting* 123
  - *5.05.4.3.3 Problem solving training* 123
- *5.05.4.4 The Triadic Model in ABA with Children and Adolescents* 123

5.05.5 RESEARCH FINDINGS 124

- *5.05.5.1 Efficacy* 124
- *5.05.5.2 Effectiveness* 126

5.05.6 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 126

5.05.7 SUMMARY 127

5.05.8 REFERENCES 128

5.05.1 THEORETICAL UNDERPINNINGS

Applied behavior analysis (ABA) has it roots in the various learning theories of psychology. The theories of Pavlov, Watson, Thorndike, and Hull (Bry, 1975) provided the early impetus for the development of modern day ABA. Despite the substantial contributions of these scholars, the pre-eminent theorist in the field has clearly been Burrhus Frederick Skinner. The theoretical work by Skinner (1938, 1953, 1957, 1969) on operant conditioning has been the foundation on which ABA has been established. Skinner developed his early theoretical formulations of operant conditioning from research in laboratories with rats and pigeons. His later writings moved more toward applications of the theory with humans. One of his most read works, *Walden Two* (Skinner, 1948), was a novel about the ways in which operant conditioning could be used to create a human utopia.

Although operant conditioning is the foundation of ABA procedures, it is important to indicate that this does not mean that ABA clinicians are of the view that all problem behavior of human beings is the product of operant conditioning. On the contrary, ABA clinicians see all human behavior as resulting from the complex interaction of a number of variables. A useful conceptualization of the determinants of behavior has been offered by Ross (1976), who argued that the variables that effect an individual's behavior are (i) long-term genetic/constitutional factors, (ii) past learning history, (iii) current physiological state, and (iv) current environmental factors. The first three of these, often referred to as organismic variables (Ollendick & Cerny, 1981), are seen as being important but not easily manipulated to effect behavior change. In contrast, environmental factors that effect behavior are more amenable to change. It is the manipulation of these environmental variables, primarily in accordance with the principles of operant conditioning, that is the basis of ABA interventions in clinical psychology.

The fundamental elements in the theory of operant conditioning are the behaviors that people engage in and the antecedents and consequences of those behaviors. In regard to the relationship between behaviors and their consequences, the future occurrence of behaviors is dependent upon the current consequences of those behaviors. Behaviors that are followed by reinforcement will increase in frequency. Reinforcement can occur in two conceptually different ways. First, it can involve the contingent presentation of a "pleasant" stimulus, called a positive reinforcer, with the process being termed positive reinforcement. Alternatively it can involve the contingent removal of an "unpleasant" stimulus, called a negative reinforcer, with this process termed negative reinforcement. It is important to note that the property of a stimulus being a positive reinforcer is defined purely in terms of the effect of that stimulus when it is presented contingent upon the occurrence of the behavior. If contingent presentation leads to an increase in the frequency of the behavior, the stimulus is a positive reinforcer. In a similar fashion, a negative reinforcer is defined in terms of its capacity to increase behavior when it is removed contingent upon the occurrence of the behavior.

Behaviors that are followed by punishment will decrease in frequency. In a similar manner to reinforcement, punishment can occur in two ways: by the contingent presentation of a negative reinforcer (in these circumstances the term "punisher" may be used), or the contingent removal of a positive reinforcer. It might be expected that these processes would be called positive and negative punishment. This has not been the practice. Some authors refer to them as type 1 and type 2 punishment. Others use the term "punishment" only for the presentation of a negative reinforcer (punisher), and use the term response cost for the contingent removal of a positive reinforcer.

Finally, behaviors that are followed neither by reinforcement nor punishment will decrease in frequency, with this phenomenon being known as extinction.

A substantial amount of early operant conditioning research examined ways in which variations in the delivery of reinforcement effected the behavior it was contingent upon. For example, it was found that reinforcement delivered immediately after the occurrence of the behavior is more effective in increasing the rate of the behavior than reinforcement which is delayed for a short period of time. Of particular interest in regard to variation in delivery of reinforcement was the effect of not providing reinforcement for every occurrence of the behavior. Providing reinforcement for every occurrence became known as using a continuous schedule of reinforcement, whereas some variation of this is referred to as an intermittent schedule of reinforcement. The effects of various intermittent schedules on the rate of increase of behavior and subsequent extinction of the behavior were studied. Initial variations were made in the ratio of reinforcement (the number of responses that had to be emitted before reinforcement was provided), and the interval of reinforcement (the period of time that would be allowed to elapse before reinforcement was provided for the next response emitted). Ratio schedules could be fixed, in that reinforcement was provided after a set number of responses, or could be variable, in that reinforcement was provided after a varying number of responses. Similarly, interval schedules could involve a fixed or variable period of time. The effects of all types of schedules have been well documented in the literature, but one important finding for the ABA field is that to teach new behaviors it is better to begin with a continuous schedule. When the behavior is established, however, it is advisable to move to variable schedules as they are more resistant to extinction once learning has taken place.

As has been described, the effects of consequences on the future occurrence of behavior is a central component of the theory of operant conditioning. In addition, however, events or stimuli that precede the occurrence of the behavior are also important. These are usually referred to as the antecedents of the behavior. Antecedents go with the behavior and its consequences to make up the antecedent–behavior–consequence sequence that is so important in the use of ABA in applied settings.

One way in which antecedent stimuli are important in operant conditioning is that they can act as a signal that a behavior will be reinforced if the behavior is emitted. A young child might find that if he asks his mother for some candy, the answer will be no. If, however, he asks his father, the answer is yes. The child quickly learns to ask the father and not the mother. In these circumstances, discrimination learning has taken place, and the father is the discriminative stimulus for the behavior of asking for candy. The systematic reinforcement of a behavior under some stimulus conditions but not others is used to promote discrimination, and is referred to a discrimination training. Such training is an important component of many interventions based on ABA.

Traditionally, operant conditioning theory focused only on those antecedents and consequences that were temporally contiguous with the behavior. These are referred to as proximal antecedents and consequences. It became clear, however, that antecedents and consequences that are temporally removed from the behavior can also be important. These are referred to as distal antecedents and consequences.

In regard to distal consequences, they may exert more influence on a particular behavior than those that are more immediate. A student, for example, may engage in long hours of study during the weekend, resulting in fatigue and the missing of social activities (immediate punishers), for the sake of completing an assignment due on Monday morning (a powerful reinforcer).

Distal antecedents that are gaining increased attention in the use of ABA are referred to as setting events (Wahler & Fox, 1981). These are events that alter the probability that a proximal antecedent will be discriminative for the occurrence of a particular behavior. By way of example, a child at school may usually respond to a request from the teacher in a cheerful and compliant manner. However, on the days following domestic disruptions at home, the child is moody and non-compliant. Here the domestic disruptions act as a setting event that alters the probability of the child responding positively to the teacher.

An important point in relation to the antecedent–behavior–consequence sequence is that a single behavior may be associated with more than one antecedent and/or more than one consequence. For example, the behavior of one child hitting another child could be occasioned by the other child calling her a name, the other child having a toy that is wanted, or in response to a dare from a third child. Similarly, the aggressive behavior could have many consequences, for example the other child cries, the aggressive child gains the coveted toy, and also receives attention from an adult, all of which could be reinforcing.

One area of operant theory that is often not given much prominence in basic texts on ABA is that of rule-governed behavior. To illustrate this concept, let us take the example of a child who is usually slow with his work at school. One day the frustrated teacher tells him that if he completes all his mathematics problems in class, he will be permitted to go to play early. The child completes the problems and goes to play early. The behavior of completing the problems was not established by contingent reinforcement, but is an example of rule-governed behavior. Skinner defined a rule as a verbal statement that specifies the conditions under which a behavior is to occur, the behavior itself, and the consequence that will follow the behavior. The child in our example followed the rule "If you complete all problems in class, you will get to go to play early." However, such rule-governed behavior will occur only if certain conditions are met. First, the child must possess sufficient language skill to understand the rule. Second, the general response class of following similar rules must have been established. In this case, the child's learning history must be such that he has been contingently reinforced for following such rules in the past.

Skinner distinguished between two types of rules, one he called a command and the other advice. The word "command" is used when the person stating the rule controls the availability of the reinforcer. The situation given above is an example of a command rule. The teacher controlled the availability of early access to play. The word "advice" is used for those situations in which the person stating the rule does not control the availability of the reinforcer, but is aware of the contingency. An example of such a rule would be a parent telling her child "If you poke that stick into the electricity outlet, you will get a shock." A knowledge of the method of establishing or changing rule-governed behavior is progressively becoming more important in clinical applications of ABA.

It is critical that practitioners using ABA procedures in applied settings be thoroughly familiar with the fundamental theoretical concepts. A sound understanding of them is not only necessary for the development of effective treatment programs, it is impossible to effectively "trouble shoot" intervention programs if the clinician does not have mastery of the theory underlying the design of the program.

5.05.2 HISTORY OF THE APPROACH

As indicated in Section 5.05.1, ABA has derived primarily from the principles of learning, particularly operant conditioning. The refinement of these principles was effected through experimental research with animals, and was published through eminent journals such as the *Journal of the Experimental Analysis of Behavior*.

The application of ABA procedures with humans did not begin in any substantial way until the 1950s, when learning-theory-based approaches to intervention in human problems became a serious competitor to the traditional psychodynamic approaches derived from Freud and other psychoanalysts. Proponents of learning-theory-based approaches were dissatisfied with three major aspects of traditional approaches. The first was the heavy reliance on the use of putative intrapsychic constructs to explain behavior. Learning theorists preferred to study the behavior directly. The second was a propensity to explain abnormal behavior in terms of a medical or illness model. Learning theorists saw abnormality as being at the end of the continuum of normal development, and hence as being quantitatively, not qualitatively, different from normal behavior. The final point of dissatisfaction was with what was seen as unsubstantiated claims for success of traditional psychodynamic approaches. Hence, learning theorists placed great emphasis on experimental demonstrations of treatment efficacy.

Despite these common concerns with traditional therapies, learning-theory-based therapy did not emerge as a single entity. In the early days, two relatively different approaches were developed. ABA was one of these, and as indicated earlier, grew out of the operant conditioning work of Skinner. The other approach was founded more on classical conditioning, particularly the work of Pavlov (1927) and Hull (1943). Development of this approach for work with humans can probably be attributed to Eysenck (1959) and Wolpe (1958). A major difference between the two approaches was the importance placed on private or covert events within humans, such as thinking and feeling. Although Skinner recognized the existence of private events, he considered that there were serious limitations in being able to study them scientifically, and hence ABA clinicians initially focused on overt, observable behavior. In contrast, proponents of the other approach were quite comfortable working with constructs such as anxiety and fear. They considered that these processes were governed by the same principles of learning as were overt behaviors (Agras, Kazdin, & Wilson, 1979).

As time passed, further approaches developed within the framework of learning theory approaches to therapy. The 1960s and 1970s saw the evolution of what became known as cognitive behavior therapies. Perhaps the first of these was social learning theory. The principal theoretician in the development of social learning theory was Bandura (1969, 1977), who argued that cognitive processes played a mediational role between the antecedent and consequent stimuli that impinged on humans. Bandura's experimental work on observational learning has been particularly useful in explaining the acquisition of many human behaviors, but also is an important theoretical component of therapeutic modeling. Bandura's construct of self-efficacy has also proven to be popular in clinical practice. He argued that the knowledge by a person that a particular behavior would be reinforced was not necessarily enough to motivate the person to perform that behavior. What also was necessary was the person's belief that he or she had the ability to perform that behavior.

Other theoreticians who have contributed to the field of cognitive behavior therapy are Ellis (1979), Mahoney (1974), and Meichenbaum (1977). While the final products that they have developed are all a little different from each other, a key element that runs through all of them is what is referred to a cognitive restructuring. This process essentially involves teaching individuals to change their behavior by changing the ways in which they think about their behavior and its consequences.

While all of these developments within the framework of learning theory approaches to therapy have been important, the focus in this chapter is on ABA, although much of what is described as being the application of ABA incorporates elements from the other approaches.

Before moving on to the assessment and clinical formulation of the ABA approach, some comments about its applications could be informative. Because ABA first emerged as the application of principles derived from laboratory research to the resolution of human problems (Baer, Wolf, & Risley, 1968), many early applications of ABA were with psychiatrically and developmentally disabled people. The research by Ayllon and Michael (1959) on teaching psychiatric nurses to be behavioral engineers has been seen as the first publication in the emergence of ABA. The first issue of the *Journal of Applied Behavior Analysis* (*JABA*) in 1968 is probably a better-recognized marker of the beginning of the widespread usage of the approach. Many of the early *JABA* articles were in the disability field. As time has progressed, the principles of ABA have begun to be used in wide variety of fields. Engelmann and Carnine (1982) used ABA as a major plank in the development of their theory of instruction that is now the basis of many educational programs (Kameenui & Simmonds, 1990). Sulzer-Azaroff (1982) has used ABA in the promotion of better practices in the field of occupational health and safety, and there are many examples of the use of ABA in the organizational behavior and management literature (e.g., Brethower, 1972; Connellan, 1978). Finally, Biglan, Metzler, and Ary (1994) report on the use of ABA procedures to change behavior on a community-wide basis. In the remainder of this chapter, the focus will be on the use of ABA in the treatment of clinical problems of children and adolescents.

5.05.3 ASSESSMENT AND CLINICAL FORMULATION

There are several excellent texts available that give a full description of assessment and intervention procedures in ABA. To mention just a few authors, there is Cooper, Heron, and Heward (1987), Grant and Evans (1994), Martin and Pear (1996), and Sulzer-Azaroff and Mayer (1991). This chapter summarizes procedures that are more fully covered in these texts.

There are three steps involved in assessment and clinical formulation using the procedures of ABA: specification of the behavior, measurement of the behavior, and assessment of the function of the behavior.

5.05.3.1 Specification of Behavior

The first step in assessment and clinical formulation is the specification or description of the behavior of interest. It is essential that this specification be as precise as possible so that any confusion about what is meant is minimized. Terms such as "aggressive" or "self-injurious" are useful only as general descriptors of classes of behavior. They are not used in an ABA assessment as they can be interpreted in many different ways by observers. What is necessary is an accurate description of the behavior of interest; for example, the child "punches other children," or "slaps himself in the face." It is only when a precise description of the behavior is formulated that measurement, the second step in the formulation, can be reliably carried out.

5.05.3.2 Measurement of Behavior

Several methods of measuring behavior have been developed, the most commonly used being frequency counts, duration measures, counting of permanent products, and time sampling procedures, although other measures such as intensity and latency can be useful.

A frequency count involves, as the name suggests, counting the number of times the behavior occurs in a given time period, for example, during the course of one day. A duration measure involves measuring the total time the child spends engaged in the behavior of interest, again usually during the course of a day. The choice of measurement method is based on a consideration of which method will "give the best picture" of what is actually happening. To illustrate this, consider a child who punches other children. Each punch is quite discrete, and a frequency count will give an accurate assessment of the extent of the problem. In contrast, consider the child who sucks his thumb so much that physical damage is occurring. If this child puts his thumb in his mouth after breakfast and only takes it out to have lunch, a frequency count would record only one instance for that morning, not really reflecting the extent of the problem. A duration measure would clearly be better in this case.

Many behaviors result in relatively permanent products that can be counted, with the resulting sum being the measure of the occurrence of the behavior. The mathematics output of a child at school can be measured by the number of problems completed. A person who is trying to give up smoking could measure the amount of smoking by counting cigarette butts in an ashtray at the end of the day.

Time sampling is more complex than the frequency, duration, and permanent product methods, and involves breaking the observation period into relatively short time intervals (e.g., 15 seconds), and then noting whether or not the behavior occurs during that interval. There are variations of time sampling referred to as total interval, partial interval, and momentary time sampling. The term "total interval" is used when the behavior of interest has to occur for the complete duration of the time interval for it to be scored as an occurrence. "Partial interval" is used when the behavior only has to occur at some time during the interval, and not for the total interval. The term "momentary" is used when a behavior is scored only if it occurs at a particular instant in the interval, usually the end of the interval. While time sampling procedures are more labor-intensive than frequency or duration methods, they have some advantages in that several behaviors of several children can be observed simultaneously. The particular type of time sampling procedure used will depend on matters such as the type of behaviors being assessed and the time available to the observer.

Sometimes clinicians will be interested not only in how often or for how long a behavior occurs but also how strong or intense it is, and hence would introduce an intensity measure. A clinician treating a child who inappropriately calls out in public places might be interested in the volume of the calling out. Volume would be important if the objective of intervention was not to eliminate public conversation, but to reduce it to a socially acceptable level.

A final measure that might be used in assessing a behavior is that of latency, or the time that elapses between some important event and the onset of the behavior by the child. Latency measures are often used when the behavior of interest is compliance by the child with some instruction given by a parent or other caregiver.

There are two reasons why clinicians need to take precise measurements of the behaviors of interest as part of the assessment process. The first is related to making a decision about whether subsequent intervention is actually warranted. For example, take the case of a child who is referred by her parents because of a "low" level of compliance to instructions given by them. An initial assessment of the child's rate of compliance might indicate that her compliance is typical for a child of her developmental level. Under these circumstances, intervention to improve the child's compliance would not be indicated. Instead, the parents need assistance with improving their knowledge of child development, and perhaps some help with accepting the child as she is.

In most cases of the application of ABA, the initial measurement of the behavior of interest does indicate that some intervention is warranted. On some occasions, the measurement will show that the behavior is occurring more often than considered appropriate. This is referred to as having a behavioral excess, and hence the objective of intervention will be to reduce the occurrence of the behavior (weaken it), perhaps even to totally eliminate it. On other occasions, the measurement will show that behavior is not occurring as often as considered appropriate. This is referred to as having a behavioral deficit, and the objective of intervention will be to increase its occurrence (strengthen it). Once a behavior has been identified as being in need of either strengthening or weakening, it is referred to as a target behavior.

The second reason for taking precise measurements of the behaviors of interest is related to the need to evaluate the efficacy of any intervention that is proceeded with. Under these circumstances, the initial measurements form the baseline against which intervention efforts are evaluated. This means, of course, that measurement of the behaviors must continue during intervention. The issue of determining the efficacy of ABA intervention is dealt with more fully in Section 5.05.5.1.

5.05.3.3 Assessment of the Function of Behavior

The third step in clinical formulation is the assessment of the function of the target behavior, that is, to identify the antecedents and consequences which are functionally related to the behavior. Before a fuller discussion of this element is possible, a distinction needs to be made between the topography of a behavior and the function of a behavior. The topography of a behavior is what the person does, and can be readily seen by an observer. The function of the behavior is the purpose of the behavior, and this is much less obvious to the observer.

The distinction between topography and function is a very important one in ABA because the relationship between a behavior and its antecedents and consequences is to do with the function of the behavior, not the topography. To illustrate this, two scenarios can be considered.

The first considers two developmentally-delayed children who frequently slap their faces to the extent that physical damage is caused. One may do it, having learned that it is an effective way of getting attention from other people, that is, the behavior is being maintained by positive reinforcement. The second child may do it having learned that people who are making demands, stop it and go away, that is, the behavior is being maintained by negative reinforcement. The behaviors of both children may have exactly the same topography, but they have quite different functions. Hence, despite the similarity of the topography of the behaviors, the intervention programs will need to be very different and be based on the analysis of function of the behaviors.

In the second scenario, again consider two developmentally delayed children. Like the first child in the other scenario, one slaps his face because he has learned that it gains adult attention. The second child frequently takes his clothes off because he has found that this gains attention from other people. In both cases, the behaviors which are quite different topographically are performing the same function. Here, the intervention programs are likely to be quite similar.

A quite common mistake made by clinicians not skilled in ABA is that they design intervention programs around behavior topographies. Such an approach may achieve an occasional success, but in general it will produce failure.

In order to adequately identify the antecedents and consequences that are functionally related to the target behavior, information about the child and behavior must be gathered from an extensive range of sources. The most commonly used methods of information collection include the use of interviews, completion of written materials, and direct observation of the target behavior and its antecedents and consequences.

5.05.3.3.1 Interviews

When working with young children or children with developmental delay, the person initially interviewed is the child's caregiver, usually the parent. Protocols for the conduct of such interviews have been published (Table 1; Murphy et al., 1985; Webster-Stratton & Herbert, 1994). One focus of the interview is to gather as much information about the target behavior as possible. This will include the precise specification of the behavior, estimates of the occurrence of the behavior, and suggestions about potential antecedents and consequences that are functionally related to it. In addition, however, other information needs to be gleaned from parents about the skill levels of the child and the child's preferences (potential reinforcers). These could be important in the design of an intervention program. Another issue to be discussed is what treatments have been tried in the past, presumably ineffectively. Such information is critical to the detail of what is included in an intervention program. It is usual that all of these matters are discussed at an initial interview with caregivers, the final part of which typically involves the establishment of procedures to collect baseline data.

When working with older children and adolescents, an interview might be conducted with them. Additionally, interviews might be conducted with significant others in the child's life, for example, teachers. Many of the same issues are covered in these interviews as those with caregivers. These interviews do, of course, allow the clinician to hear a different, and probably useful, perspective on the issues.

5.05.3.3.2 Written materials

In addition to the use of interviews, information can be gathered using a range of written materials. These would typically include standardized checklists about the child, for example, the Child Behavior Checklist (Achenbach & Edelbrock, 1983). If relevant, they may also include checklists about the functioning of other members, for example, measures of parental relationships or affective status.

Such written material can be of value for several reasons. First, it often includes normative data and can provide information on the occurrence of the behavior of interest relative to the general population. This information should be relevant in deciding whether or not to try to change the behavior of interest. Second, the use of standardized checklists might bring the clinician's attention to some relevant matters that have not been mentioned in any interviews conducted. Finally, if an intervention is carried out, the checklist can also be administered after intervention, and hence add to other information collected to evaluate the efficacy of the intervention.

5.05.3.3.3 Direct observation of the behavior

The final general source of information is that provided by direct observation of the behavior. Mostly, this is done by caregivers who collect data on the target behavior using frequency counts or duration measures. Obtaining good-quality data from caregivers is not always an easy task. It is important for clinicians to realize that for this part of the exercise it is the caregiver's recording behavior that is important, and attention must be paid to the antecedents and consequences of that behavior. First, the clinician needs to provide structured data collection sheets, clear instructions, and

Table 1 An interview schedule for use in the behavioral assessment of children's problems.

The "client" referred to below is a significant person in the referred child's life, for example, a parent

1. The establishment

(i) Introduce yourself to the client (name and profession; how you like to be addressed; introduce any observer or colleague—name and role)
(ii) Engage in "small talk" to put the client at ease
(iii) Direct client to an appropriate chair if the interview is in an office (have chairs and furniture prearranged to facilitate communication, etc.)

Remember that this period is critical to the development of a good client–therapist relationship. Such a relationship is essential for therapeutic success

2. General

(i) State purpose(s) of interview
The client should be given a general statement on the purpose of the interview. This is important because some clients expect an immediate answer to their problem. Over the course of the interview it may be necessary to remind the client that management advice is not possible at this stage

(ii) Seek permission for note taking, etc.
Permission should be obtained from the client for note taking. Usually this is not a problem in terms of client permission. Should audio- or video-recording be thought to be an advantage, seek additional prior permission

(iii) Seek client's perception of problem
At this stage it is helpful to ask the client about the nature of the problem. This is simply to prompt clients to give their perception of why they have come to see you. The description given may not be satisfactory to you or the client at this stage, but continued exploration about the nature of the problem will follow later on in the interview. The idea of asking the question here is partially to set the stage for the next step, gathering general background information. Terminate this step by saying that you will come back to talking about the problem later, but first you want to obtain some other information about the child

(iv) Determine referral agency
Obtain or check details of the referring agency (professional's name, agency, phone number, etc.). Remember that good liaison with other agencies in the community is essential, but also that dual management can be wasteful, confusing, or deleterious

3. General background information

(i) Biographical information of child

(a) name,
(b) age (including date of birth),
(c) address, and
(d) telephone number

(ii) Family background

Here the interviewer collects information about:
(a) whether the mother and father (or others filling these roles) are present in the household,
(b) what employment, if any, parents (or other caretakers) have,
(c) number and ages of siblings in the family,
(d) any unusual family organizational pattern, such as an aunty or cousin living with the family, and
(e) quality of relationship between the target child and other family or household members

(iii) Educational background

Here careful detail of the child's educational background is recorded, including:
(a) number and type of schools attended,
(b) general progress at school including any grades repeated,
(c) any history of conflict with teachers,
(d) current teacher's name, and
(e) relationships with peers

(iv) Medical background

Here the interviewer lists any information regarding the child's medical background, particularly anything that is slightly out of the ordinary. Remember to question regarding:
(a) general development milestones and behavior patterns,
(b) childhood illnesses, and
(c) any medication the child is currently on

Table 1 (continued)

If the behavioral interviewer is unsure about "normal" developmental rates of infants and children, standardized developmental checklists should be used

The above areas need to be explored because (a) most clients receiving clinical service expect such areas to be addressed, and (b) the information helps the clinician to appreciate the context of the problem, assists the clinician's choice of target behaviors, guides selection of assessment devices to be used, and identifies possible barriers to particular treatment programs

4. Specific information build-up

(i) Investigate the precise nature of the problem
Here the interviewer seeks to have the client state precisely what the nature of the problem is. Some clients may need extra time to articulate the problem, and some may need prompting to be specific

(ii) Investigate the extent of the problem
Here the interviewer seeks to obtain some estimate from the client about the frequency or the duration of the previously described behavior(s). At this point, self-report information is relied upon; at a later stage, more objective methods of assessment will need to be introduced

(iii) Investigate the apparent antecedents and consequences of the problem
Here the interviewer seeks to draw from the client a rough idea of what he or she perceives to be the antecedents and consequences associated with the particular behavior(s)
Choice of language on the part of the clinician is very important. Phrases like "what happens before," "events that trigger" (antecedents), and "what happens during and after" (consequences) are better for most clients

(iv) Determine what has previously been tried
Clients should be assisted to state previous attempts to handle the problem, and advice received from other agencies previously or currently involved

5. Clarification of expectations

(i) Establish agreement on the goal of treatment
Here the interviewer needs to come to agreement with the client regarding the goal of treatment. The goal must be stated in specific terms

(ii) Determine suitability of agency
State whether or not this agency is suitable for helping to achieve that goal, and if not, give client referral options

(iii) Establish realistic timeline for improvement
Ensure that the client is aware that the therapist does not have a magic wand to wave over the problem child and be able to immediately convert him or her from being a problem to no problem at all

(iv) Involvement of client
Impress upon the client the need for him or her to be actively involved in the intervention program, and if they decide to proceed that this will necessarily require active participation by the client

6. Setting of homework (if feasible at this stage)

The interviewer here needs to set very precise homework for the client to carry out during the subsequent period of time. Usually 1 week elapses between the first and second interviews. The specific instructions would usually require the collection of baseline data and the identification of the antecedents and consequences associated with the target behavior. Appropriate antecedent stimulus control of the client's behavior is usually best achieved through the handing out of very structured sheets that client has to fill in during the week and return at the next interview. It is important at the conclusion of the interview to make sure that the client is able to repeat precisely what is expected of him or her by the interviewer. (In many cases this homework setting can take up to 1 hour, so a separate interview for this purpose may be necessary)

7. Scheduling of next appointment

8. Termination of interview

Again remember the importance of a good client–therapist relationship

Source: reproduced with permission from Murphy, Hudson, King, and Remenyi (1985).

perhaps some role play practice at filling in the sheets. Although baseline data are usually collected for a period of two weeks, many clinicians find that a check-up phone call 2 days into data collection is important. Such a call will help overcome any difficulties encountered in the first days and increase the probability that the caregiver will return with useful data. These are all examples of establishing antecedent stimulus control over the data-collecting behavior of the caregivers. As ongoing data collection is important for evaluating interventions, the behavior of the caregivers needs to be reinforced by the clinician.

Older children and adolescents might also participate in collecting direct observation data on their own behavior. This would be absolutely necessary if the target behavior was covert; for example, having obsessive thoughts.

The final way in which direct observation data can be collected is by the use of independent observers who enter into the natural environment of the child. A potential difficulty to be addressed here is the extent to which the presence of an outsider is reactive and has an influence on the behavior being observed. The use of independent observers in standard clinical practice is often a luxury which cannot be afforded, and hence their use is often restricted to research settings.

5.05.3.3.4 Developing a hypothesis regarding the function of the behavior

When a broad set of information has been collected by the various methods described above, the clinician is in a position to formulate hypotheses about the functions of target behaviors. This task is a little like putting together a jigsaw puzzle. Some relevant information will be gathered from interviews, some more will result from written material, and even more from direct observations. Not always will the information from different sources be consistent. When the clinician stands back, as it were, a picture will emerge that suggests a potential functional relationship between the target behavior and its antecedents and consequences.

The extent to which this hypothesized relationship is correct can only be tested by manipulation of those antecedents and consequences via a carefully designed intervention program, and the ongoing measurement of the target behavior. If the behavior changes in the desired direction, the hypothesis has been supported (perhaps not proven, but this matter will be discussed further in Section 5.05.5.1). If the program has been implemented as intended and the behavior does not change in the desired direction, the hypothesis is not supported, and the clinician needs to go back to the drawing board. The introduction of the intervention tests the hypothesis. This hypothesis-testing process is a defining feature of the ABA approach to clinical intervention.

5.05.4 DESCRIPTION OF INTERVENTION PROCEDURES AND PROPOSED MECHANISMS OF CHANGE

As has been described earlier, the logical outcome of the completion of an ABA assessment is the development of an intervention program that is designed to strengthen or weaken the target behavior. Such a program will involve the use of procedures developed from the basic principles of operant conditioning described earlier. It is important to point out that a comprehensive intervention program for a given child may involve more than one target behavior, and hence involve the simultaneous application of more than one procedure. For purposes of clarity, however, each procedure will be described separately here. First, there will be coverage of the relatively specific procedures used to strengthen behavior. Then there will be a discussion of the procedures for weakening behavior. Finally, there will be a discussion of complex procedures for strengthening and weakening behaviors.

5.05.4.1 Procedures for Strengthening Behavior

The most commonly used procedures for strengthening behavior are the systematic use of reinforcement, prompting, shaping, chaining, and modeling.

5.05.4.1.1 Systematic use of reinforcement

The most straightforward procedure for strengthening behavior is simply to wait for the behavior to occur, and then to reinforce it. However, there are some important issues related to this seemingly straightforward process. Perhaps the most important of these is to select an appropriate reinforcer. It must be remembered that a reinforcer is defined in terms of its effect on the behavior it is contingent upon. Clinicians can err in what they think will be reinforcing for particular children. The information collected in the assessment phase is critical for the right choice of an effective reinforcer, and this is not always an easy task. Some clinicians develop what are termed reinforcement menus, or lists of things that are potentially reinforcing, and then have the child or caregiver select from the menu. Another very helpful hint is to use what is known as the Premack principle. Premack (1959) noted that the activities children chose to engage in during their free time could be used to reinforce other, less-preferred activities. He technically defined this as using higher-probability behaviors to reinforce lower-probability behaviors. The importance of the principle is that if you are having trouble identifying what is reinforcing for children, observe what they do in their free time.

Reinforcers vary in nature and can be classified in a number of ways. One of the earliest distinctions made was between primary and secondary reinforcers. A primary reinforcer, sometimes referred to as an unconditioned reinforcer, is one that does not depend on learning to have taken place for it to develop its reinforcing properties. Food, water, and warmth are examples of such primary reinforcers. In contrast, learning has been involved in secondary or conditioned reinforcers developing their reinforcing properties. For example, some children find playing games on a computer to be very reinforcing. For this to have happened, some learning about the computer and the particular game must have taken place. It can be argued that there is limited clinical usefulness in the distinction between primary and secondary reinforcers as ethical considerations dictate that children have a fundamental right to unrestricted access to most primary reinforcers, and hence they should not be used in ABA programs. It is the case, however, that edible reinforcers such as candy have been used successfully to motivate children who do not respond to secondary reinforcers.

A second distinction that is made in regard to types of reinforcers is that between naturally occurring and artificial reinforcers. A naturally occurring reinforcer is one that is readily available and typically used in the setting in which the ABA program is being conducted. In most homes, reinforcers such as parent attention, pocket money, and access to television are available to children. Unfortunately, these reinforcers are often given to the children in an unsystematic or unplanned manner. An artificial reinforcer is one that is not usually available in the setting, but is brought in for the purposes of the program. Candy is not usually used as a reinforcer in most classrooms but could be introduced for use in a program for a special purpose. It is generally advisable for clinicians to use artificial reinforcers only when absolutely necessary, as the fading of such reinforcers can be more difficult than is the case for naturally occurring reinforcers.

A third distinction that is made is between extrinsic and intrinsic reinforcers. An extrinsic reinforcer is one that is quite separate from the behavior that it is reinforcing in that the behavior could occur and not be followed by the reinforcer. If a clinician is using praise to reinforce a child's behavior, the behavior can be allowed to occur without the reinforcement being administered. An intrinsic reinforcer can be viewed as one that is actually part of the behavior. The behavior is reinforcing by itself and does not need an extrinsic reinforcer to maintain it. Most children do not need to be extrinsically reinforced for watching their favorite show on television. They do it because they like it. The distinction between intrinsic and extrinsic reinforcement can be clinically important. Some behaviors, for example reading, may initially need to be extrinsically reinforced to get them going, but once established, can be maintained by their own intrinsic reinforcement.

Yet another way of distinguishing among reinforcers is to classify them based on the broad type of reinforcer. The types listed can vary among authors in the field, but a common list includes social, food, activity, and token reinforcers. Social reinforcers include praise, hugs, or attention from another person. As the name suggests, food reinforcers are things the child can eat. Rather than being basic food requirements, which, as mentioned earlier, are problematical from an ethical viewpoint, they are special treats that the child does not have regular access to. Activity reinforcers are those things that the child likes to do, for example, playing games and watching television. Finally, token reinforcers do not have value in themselves but derive their value from what they can be exchanged for. Many classroom teachers have children earning points for completion of academic work, and then allow the children to exchange those points for back-up reinforcers such as extra time at recess or reduced homework. Money is certainly the most frequently used token reinforcer.

A final point to be made in regard to the systematic use of reinforcement to strengthen a target behavior is that the fundamental principles derived from learning theory need to be adhered to. Reinforcement usually needs to be immediate, particularly with children at a low developmental level. If another behavior occurs between the target behavior and the delivery of the reinforcement, the more recent behavior might be the one that is strengthened. The second behavior under these circumstances is usually referred to as superstitious behavior. Also, attention needs to be paid to what is known about schedules of reinforcement. A continuous schedule needs to be used to maximize the rate of development of the new behavior, and then there should be a move to an intermittent schedule, referred to as a thinning of the reinforcement, to make the behavior more resistant to extinction.

5.05.4.1.2 Prompting

The use of systematic reinforcement to strengthen a behavior can only be effective if the behavior occurs in the first place. If it is not occurring, it needs to be prompted in some way, and it can then be reinforced. Prompts can take a variety of forms, but can be categorized as being physical, verbal, or visual prompts.

A physical prompt is one in which the child is given some physical assistance to perform the required response. Many parents use physical prompts to teach their children to learn feeding and dressing tasks. Care needs to be taken in determining the precise manner of using physical prompts. For example, when teaching a child to use a spoon, it is better to stand behind the child with the assisting hand over the child's hand. Such placement minimizes distraction for the child and allows for co-active movement of adult and child from the bowl to the mouth. As with all prompts, this prompt will need to be faded as the behavior becomes established over several trials, and this can be done by movement of the adult hand slowly up the arm of the child.

A verbal prompt is one that involves the use of spoken words or sounds to help the child perform the behavior. A boy learning to dress himself might be prompted to "hold your shorts in front of you, now put one leg in, etc.," or a child learning to read but stumbles on a word might be prompted to "say the sound of the first letter."

As the name suggests, a visual prompt is something the child sees that helps with the response. If a child is sorting blocks according to shape, the teacher might prompt by pointing to the correct container. This is often referred to as a gestural prompt. A picture placed on the refrigerator to remind a child to empty the trash would also be a visual prompt. Written notes can also be considered to be visual prompts, although some authors consider these to be verbal prompts.

All prompts need to be faded as the target behavior becomes stronger. The way in which this is done will vary depending upon the task in question, but in general the procedure is to start with a strong prompt and weaken it progressively, but not so fast that the target behavior is lost. It must be remembered that the purpose of prompting is to get the behavior to occur so that it can be reinforced, and reinforcement should not usually be thinned until the prompt is removed.

5.05.4.1.3 Shaping

Some behaviors are so complex that no amount of prompting will elicit the desired final form of it, resulting in no reinforcement, and hence no behavior change. Under these circumstances, a decision is made to reinforce only some approximation to the desired behavior. Progressively, closer approximations to the behavior are required before reinforcement is forthcoming, and eventually the final behavior will have been totally shaped.

5.05.4.1.4 Chaining

Many behaviors that are of interest to clinicians are complex behaviors consisting of a number of smaller behaviors linked together to form a behavior chain. Each link of the chain can be conceptualized within the usual antecedent–behavior–consequence sequence, and is taught using the procedures described thus far. Additionally, however, the completion of one link is viewed as the antecedent stimulus occasioning the performance of the next link, which in turn becomes the stimulus for the occurrence of the next. The process of developing the final chain can be of three types, forward chaining, backward chaining, and total task training.

When using forward chaining, the initial link of the chain is the first taught, then the second, then the third, and so on until the complete chain has been learned. Prompting, shaping, and reinforcing could all be used in the teaching of each separate link. While learning each link, the child is usually helped through all of the remaining links so that task completion can be reinforced. Teaching a child to put on a pair of trousers might be a typical skill taught by using forward chaining.

When using backward chaining (sometimes called reverse chaining) it is the final link in the chain that is taught first, then the others are proceeded with working in reverse order. The initial link thus becomes the last taught. This approach is considered desirable when the complex behavior leads to some naturally occurring reinforcer being available, for example a child learning to use a spoon might start with the final link of placing the food in his mouth. This contiguity of naturally occurring reinforcement and task completion is considered to provide better motivation to learn the earlier links in the chain.

Whereas forward and backward chaining are typically used when a child has mastered few if any of the separate behavioral links, total task training is more efficient for situations in which the child has mastered most or all of the links. A child who has mastered simple addition, subtraction, multiplication, and division would typically be taught the behavioral chain involved in completing a long division by using a total task method.

5.05.4.1.5 Modeling

Modeling is sometimes described as being an example of visual prompting, but modeling used properly as a therapeutic technique involves more than just a prompt. Modeling as a technique depends to a large extent on the fact that people can and do learn much from imitating others. Children often engage in incidental imitation of others, and, much to the chagrin of parents, they often pick up behaviors other than those preferred by parents. Imitation is important but is only one of the many components of therapeutic modeling. As Bandura (1969) has pointed out, modeling involves attentional, retentional, motor reproduction, and motivational components. In practical terms, the full modeling procedure includes the following steps:

(i) Explanation to the learner to pay attention to the relevant aspects of the behavior to be modeled (e.g., "Look at how the model does . . .").

(ii) Recall by the learner of the aspects of the behaviors to be observed.

(iii) Reinforcement of this recall and corrective feedback by the clinician as necessary.

(iv) The demonstration of the desired behavior by the clinician.

(v) Recall by the learner of what was observed.

(vi) Reinforcement of this recall and corrective feedback by the clinician as necessary.

(vii) Performance of the behavior by the learner.

(viii) Reinforcement of this performance and corrective feedback by the clinician as necessary.

(ix) Repeat of the above steps as necessary.

This procedure will vary depending on the particular case. The model can be live or on film.

5.05.4.2 Procedures for Weakening Behavior

The most commonly used procedures for weakening behavior are differential reinforcement of incompatible behavior (DRI), differential reinforcement of alternate behavior (DRA), differential reinforcement of other behavior (DRO), extinction, timeout, overcorrection, and punishment. As can be seen in Sections 5.05.4.2.1–5.05.4.2.3, several of these procedures for weakening behavior involve a concurrent strengthening of a nontarget behavior.

5.05.4.2.1 Differential reinforcement of incompatible behavior

Differential reinforcement of incompatible behavior (DRI) is a weakening procedure that is based on reinforcement. The behavior that is reinforced is not the target behavior, as this would clearly strengthen it, but rather is a behavior that is both desirable and topographically incompatible with the target behavior. The rationale is that the child cannot simultaneously engage in two incompatible behaviors, so as one is strengthened, the other must be weakened.

A DRI procedure could be used to intervene with a child who is continually out of his seat in the classroom. A desirable behavior that is topographically incompatible with roaming around the classroom is that of completing academic work. The teacher therefore could reduce the out-of-seat behavior by systematically reinforcing the child for completion of work. Another example of a behavior that could be treated with DRI is that of a young child who, while watching television, continually "twiddles" with her hair to the extent that bald patches are appearing. Such hair "twiddling" could be reduced by reinforcing the child for cuddling a teddy bear.

The big advantage of DRI is that the focus is on doing something positive, and not simply the removal of an unwanted behavior. It is for this reason that DRI is sometimes referred to as a constructional approach to weakening behavior. One potential procedural difficulty with DRI is that it is not always easy to find an appropriate behavior that is topographically incompatible with the target behavior. Another potential difficulty is that the child must be able to engage in the incompatible behavior for all of the time he would normally engage in the target behavior. This is usually referred to as complying with the 100% rule. Reinforcing piano playing could be used to treat the preschool child who hits other children, but the child could not be expected to play the piano for the total time in attendance at preschool.

It is important to note that when using DRI, some thought must to be given to what will be done if the target behavior does occur. There will be a need for a systematic response to that occurrence. The options will include strategies such as extinction, timeout, and punishment, all of which are discussed below.

5.05.4.2.2 Differential reinforcement of alternative behavior

In a similar fashion to DRI, the differential reinforcement of alternate behavior, or DRA, is a procedure that weakens the target behavior by reinforcing some specific other behavior. This is sometimes referred to as ALT-R. In contrast to DRI, the other behavior in DRA is not topographically incompatible with the target behavior. It is technically possible for the child to simultaneously engage in both the target behavior and the behavior being reinforced, but the expectation is that the reinforcement available for the new behavior will be of such strength that it will be preferred over the target behavior.

DRA also has the advantage of being a constructional approach to behavior reduction, and avoids the DRI difficulty of having to identify a topographically incompatible behavior.

As with DRI, when using DRA the target behavior can occur, and some systematic method of responding needs to be planned for.

5.05.4.2.3 Differential reinforcement of other behavior

Like DRI and DRA, the differential reinforcement of other behavior, or DRO, is a reinforcement-based procedure. Unlike the other procedures which reinforce some specific other behavior, DRO involves the reinforcement of any other behavior, that is, any behavior other than the target behavior. In effect, the DRO procedure translates into reinforcement being delivered for the absence of the target behavior for a predetermined period of time. Sometimes DRO is referred to as omission training. If DRO was being used to treat an aggressive behavior such as hitting other children, reinforcement would be delivered to the child after a set period time during which the child refrained from hitting.

There are many procedural variations of DRO (Vollmer & Iwata, 1992). First, there can be variations in the DRO time interval. The selection of this time is somewhat arbitrary, but it needs to be short enough to give the child a reasonable chance of securing the reinforcement. The time interval can be fixed or variable, but most commonly it is fixed for ease of implementation. One method of selecting the time is to make it the mean of the time intervals elapsing between responses during the collection of baseline data. This has been termed the inter-response time (IRT). Often in the use of DRO, a short time interval is started with, but this is progressively increased in length over time. This procedure allows the gradual fading of the use of the reinforcement.

Another procedural variation is to do with what is termed interval resetting when the target behavior occurs. One option is to reset the clock at zero, so if the target behavior occurs after 2 minutes of a 3 minute interval, the child has to abstain from that behavior for a further 3 minutes before reinforcement is forthcoming. A second option is to allow the current interval to finish (with no reinforcement), and then require another full interval to elapse before reinforcement. In the case above, the child would have to wait 4 minutes for reinforcement. This second option has some advantages in settings such as classrooms where it is easier for a teacher to attend to a child at set times, for example every 15 minutes. A potential problem with this method of resetting is that a "smart" child who transgresses during a set interval may realize that there is effectively no penalty for further transgressions during the remainder of that period, and take advantage of this. The reset to zero approach can also encounter some difficulties with the smart child. Take, for example, a child on a DRO intervention for hitting other children. Immediately following the receipt of reinforcement for abstaining from hitting, she may realize that now is the time to "get in a few good ones," as the effect of this is to delay the next reinforcement for only a few seconds. Clinicians have combated this by making the reinforcement available after successive intervals free of target behavior greater than that available after just one interval.

A final point in regard to the use of DRO, just as is the case with DRI and DRA, is that the target behavior can occur when using DRO, and some systematic method of responding (e.g., extinction, timeout, punishment) needs to be planned for.

5.05.4.2.4 Extinction

Extinction as a weakening procedure in clinical settings is based directly on the operant conditioning principle that a behavior which is never reinforced will decrease in frequency. It relies on the successful identification, during the functional assessment of the target behavior, of the reinforcers that are maintaining that behavior. Once identified, the contingent delivery of the reinforcers is stopped, and the behavior will eventually disappear.

On the face of it, extinction appears to be a straightforward procedure to implement. However, there are many procedural difficulties that need to be considered in developing an extinction-based intervention. First, extinction can only be used for behaviors that are extrinsically reinforced. It is simply not possible to remove intrinsic reinforcement. Unfortunately, many undesirable behaviors of children and adolescents are engaged in because they are fun.

The second procedural problem is related to the fact that the implementation of an extinction procedure is often followed by a rapid escalation in the target behavior, aptly called the extinction burst, before the behavior eventually weakens. It is critical that this burst is able to be tolerated. If extinction was being used to reduce the self-injurious behavior of a developmentally disabled child, considerable damage could be caused during the burst, making extinction a contraindicated treatment. Similarly, although extinction is very effective in treating night-time crying of children, there are some parents who become so anxious during the burst that the treatment is aborted.

A third procedural problem associated with extinction results from the fact that a given target behavior might be maintained by several reinforcers. If extinction is to be effective, all sources of reinforcement must be terminated. Take, for example, the situation where the target behavior is that of a preschool child hitting other children. The functional analysis of this behavior might show, *inter alia*, that this happens when the child wants a toy from the second child, and that the hit results in the toy being given up, hence reinforcing the behavior. A hastily conceived extinction program could dictate that whenever this happens the preschool teacher should admonish the offender and return the toy to the second child, hence cutting off that source of reinforcement. It might be, however, that the offending child finds adult attention to be very reinforcing, and this will still be readily available when an aggressive incident takes place. Additionally, incidents such as these attract attention from other children, which is also potentially reinforcing. To be effective, an extinction program needs to remove all sources of reinforcement from the target behavior. Some, such as teacher attention, might be relatively easy to remove, whereas others, for example, peer attention, can be more of a problem.

The final procedural problem associated with extinction is that the source of reinforcement needs to be removed on all, not just some or even most, of the occasions on which the behavior occurs. We know from basic operant conditioning principles that behaviors can be maintained on very thin schedules of reinforcement, particularly if the schedule is variable. Consider the parents who deems that their attention is maintaining the tantrums being engaged in by their toddler. They decide to use extinction to remove the tantrums, and ignore them for almost all of the time. Occasionally, they are a little stressed, cannot ignore the child's behavior, and pay attention to him. Here the parents have in fact placed the child, not on an extinction schedule, but on a thin variable schedule of reinforcement. Sadly, as is also know from operant principles, such a schedule makes behavior resistant to extinction.

5.05.4.2.5 Timeout

Timeout is perhaps the most misunderstood concept in ABA. Technically it refers to time out from reinforcement, literally meaning that reinforcement is not available for any behavior. Given that it is probably impossible to create an environment that is totally devoid of all sources of reinforcement, the term "timeout" is used to refer to the situation in which the reinforcement available in the environment is reduced whenever the target behavior occurs. The reduction is usually for a relatively brief period of time such as a few minutes. The contingent use of timeout leads to a weakening of the target behavior.

Frequently, timeout is taken to mean removal of the child from the environment she is in to another environment. If the new environment has less reinforcement available than the former, then this is timeout. If, however, as is often the case, the new environment has more reinforcement available, it is not timeout. It is not uncommon for children in schools to be taken from their classrooms and placed in "timeout" in the corridor, perhaps outside the principal's office. Under these circumstances there is likely to be more reinforcement available in the corridor than in the classroom. For a start, aversive work requirements are removed. And other activities such as chatting to children moving through the corridor are available. Under these circumstances, the use of "timeout" is unlikely to weaken the target behavior. In reality, it is "timein," and likely to strengthen the target behavior.

There are several procedural variations of timeout. First, there is the distinction between exclusionary and nonexclusionary timeout. Exclusionary timeout refers to the actual removal of the child to another location which is less reinforcing. Nonexclusionary timeout refers to making the current environment less reinforcing. This can be done, for example, by stopping interaction with other people in the room, or restricting access to certain activities in the room, or even using a special timeout chair. As with exclusionary timeout, the period of time of implementation is brief.

The amount of minimum time that is associated with timeout varies with the clinician, although it is usually related to the age of children. One useful rule of thumb is to begin with 2 minutes for 2 year olds, and then add a minute for each year of age to 10 minutes.

An important issue to do with timeout is related to what the child is doing when the time period is up. Let us take the example of a 3 year old child who has been placed in exclusionary timeout, but screams, shouts, and bangs on the door for the entire 3 minutes. If he is then permitted to leave timeout, it is possible that screaming and shouting will be reinforced and be more likely to occur the next time the child is placed in timeout. This, of course, would be counterproductive, and most clinicians require a period of quiet time before timeout is terminated, resulting in the timeout period potentially being longer than the established minimum. The length of the extension varies from clinician to clinician, some require only 15 seconds of quiet time, while others require a period equivalent to the established minimum. In practice, while some children will initially engage in resistive behaviors, after a few systematic trials they learn that resistive behaviors simply increase the length of the timeout and are quickly dropped.

Timeout has clearly proven to be one of the more powerful ABA procedures, and is widely applicable. The major limitation to its use is the availability of a timeout environment. In terms of nonexclusionary timeout, it is sometimes difficult to make the current environment less reinforcing. For exclusionary timeout, it may be difficult to access a suitable venue. An additional potential problem with exclusionary timeout relates to the child refusing to go. A small child can be carried, but a child who is large compared to the adult might successfully resist. Problems such as these would be contraindicative of the use of timeout.

A final comment on exclusionary timeout relates to its use in settings where there are policies regarding seclusion of children. Clinicians need to be thoroughly familiar with such matters. Additionally, regardless of statutory requirements, ethical considerations demand that the timeout venue meets minimum standards of warmth, light, and physical comfort. It is also a good idea for a record to be kept of the frequency of use and duration of each application of timeout. Such records are important for demonstrating that timeout is being used as a therapeutic technique, and not just a method for isolating a difficult child.

5.05.4.2.6 Overcorrection

The use of overcorrection to weaken a target behavior typically involves two components, restitution and positive practice. The restitution component involves the child in making good any damage caused by the target behavior. The positive practice component involves the child in repeatedly practising an alternative, acceptable behavior. Consider the use of overcorrection to treat disruptive behavior of children in the classroom. On entering the classroom, the child rushes down the aisle, knocks books off desks as he goes, and finally launches at a seat, pushing it across the classroom. The restitution component of overcorrection would require the child to return the scattered books to their desks and replace the chair to its proper place. The positive practice component would then require the child to leave the classroom, enter in an orderly fashion, and sit gently on the chair. Typically, several repetitions of the positive practice component would be required.

Overcorrection has had extensive use in treating disruptive behaviors, particularly those of children with a developmental disability, although its popularity does seem to have waned in recent years. This could be due the fact that the positive practice trials often involved physically forcing the child to engage in the behavior, and this has been seen as being unnecessarily aversive.

5.05.4.2.7 Punishment

Punishment as a therapeutic procedure has had two separate forms, the first involving the contingent removal of a positive reinforcer and the second involving the contingent delivery of a negative reinforcer.

The removal of a positive reinforcer is often referred to as response cost. The reinforcer would typically be one that is naturally occurring in the child's environment. Many families, for example, routinely restrict access to television or a favorite toy as punishment for a misdemeanor. Such use of punishment has rarely been controversial, provided that the reinforcers removed are viewed as privileges and not rights.

The negative reinforcer used in the second type of punishment is usually referred to as an aversive stimulus, and the use of such punishment procedures has been quite controversial. The controversy has centered not so much on the efficacy of such punishment, as there are well-documented examples of its successful use with seemingly intransigent behaviors such as self-injurious behavior exhibited by some people with developmental disabilities, but on ethical grounds, and the objections focus on the issue of whether punishment is ever really necessary. This matter will be addressed again in the discussion of the principle of the least restrictive alternative.

5.05.4.2.8 The principle of the least restrictive alternative

The procedures discussed for weakening behavior have been described with a particular order in mind. First covered were the differential reinforcement procedures, where the emphasis is on weakening the target behavior by strengthening other behaviors. These are generally considered to be minimally restrictive or intrusive for the child involved. At the other end of the spectrum was punishment using aversive stimuli, clearly a very intrusive procedure. It is generally accepted that clinicians should use interventions that are as minimally restrictive as possible, and this is referred to as the application of the principle of the least restrictive alternative (Turnbull, 1981).

While there is agreement that this principle ought be adhered to, what there is not agreement about is whether it is necessary to use aversive procedures at all. Some argue that all problematic behavior can be successfully treated if procedures such as the differential reinforcement procedures are used well. Others argue that there are not yet enough data available to warrant this conclusion, particularly in regard to behaviors such as the chronic self-injury exhibited by some people with developmental disabilities. An excellent selection of papers on this topic are included in Repp and Singh (1990).

5.05.4.3 Complex Procedures for Changing Behavior

The procedures described above are relatively specific, and are designed for either strengthening or weakening particular target behaviors. Sometimes the clinician is faced with a situation in which more generalized behavior change is required. In such situations, more complex procedures for changing behavior are appropriate. Examples of these are token systems, behavioral contracting procedures, and problem solving training.

5.05.4.3.1 Token economies

A token economy is a specifically constructed system in which desirable behaviors are reinforced with tokens and undesirable behaviors are punished by loss of tokens. Periodically, tokens can be exchanged for back-up reinforcers. Detailed planning is necessary for the successful establishment of a token economy. Care needs to be taken to ensure that back-up reinforcers are only available through the system, and circumvention of the contingencies through theft is not possible.

5.05.4.3.2 Behavioral contracting

Behavioral contracting involves the entering of a formal agreement with the child about what behaviors are required of him and what reinforcers will be consequently available. Contracts are usually written, and need to be constructed with a careful eye to specificity. From a theoretical perspective, behavioral contracts are a particular example of establishing rule-governed behavior in the child (Martin & Pear, 1996).

5.05.4.3.3 Problem solving training

The aim of this overall process is not to solve a particular problem for the child but to teach a general strategy that can be implemented for a variety of presenting problems. To do this, however, problems must be dealt with on an individual basis. This procedure involves selecting an individual problem and directively leading the child through a series of steps to solve it. The steps followed typically include (i) specification of the problem, (ii) generation of possible courses of action, (iii) selection of a course of action, and (iv) trying out the selected option. When one problem has been successfully dealt with, others are handled in turn, but with the clinician playing a less directive role on each subsequent problem. From a theoretical perspective, problem solving training, like contracting, is related to the establishment of rule governed behavior (Martin & Pear, 1996). However, unlike in contracting where the rules are command rules in that the clinician controls the reinforcers, in problem solving training the rules are more advice rules as the clinician is unlikely to control the relevant reinforcers.

5.05.4.4 The Triadic Model in ABA with Children and Adolescents

As has been indicated, a fundamental tenet of ABA is that the environment plays a major role in the development and maintenance of all behavior. From a clinical perspective, the reinforcers and punishers that naturally occur in the environment are essential to the maintenance of the problem behaviors. For children, the major sources of these reinforcers and punishers are the significant adults in the child's environment, that is, people such as parents, teachers, or child care personnel. It is quite logical, therefore, that intervention for the clinical problems of children is mediated through these significant others. The systematic manipulation of reinforcers and/or punishers that must be implemented to change the child's behavior is effected by changing the behavior of these significant others. Hence, the application of ABA procedures with children is often referred to as the triadic model, as it involves three persons, that is, the clinician, the significant other, and the child (Herbert, 1991; Tharp & Wetzel, 1969).

The implication of this requirement to change the contingencies of the child's behavior by altering the adult's behavior is that the adult's behavior becomes the focus of the intervention. For example, noncompliance in young children is treated by teaching the parents to give clear instructions and to reinforce differentially child compliance (Forehand & McMahon, 1981). In these situations, the clinician must think about the ways in which the antecedents and consequences of the parents' behavior can be changed in order to alter that behavior. Modeling might be used to teach the parents new child management skills, and then social reinforcement used contingent upon applying these skills in the home.

5.05.5 RESEARCH FINDINGS

5.05.5.1 Efficacy

A trademark of ABA is the systematic collection of data on the behaviors that are targeted for change. These data are then graphed to display the impact of the intervention in changing the baseline rates of the behaviors. Time series research designs such as reversal designs and multiple baseline designs are typically used to examine the functional relationship between the intervention and the target behavior (Hersen & Barlow, 1984; Kazdin, 1982). While designs such as these have respectable internal validity and are necessary to demonstrate a functional relationship between the treatment and the observed changes in target behaviors, they are not always used in standard clinical practice. Most practitioners rely upon the simple AB (baseline–treatment) design, hopefully with some form of follow-up. While this design is weak in terms of internal validity in that other plausible hypotheses might be offered to account for the behavior change, it is much easier to implement in standard clinical settings. The design will tell the clinician if the behavior has changed in the desired direction, it is just that the clinician cannot be absolutely certain that the behavior change was caused by the intervention. While demonstrations of functional control can only be demonstrated in individual interventions by the better-quality designs, some authors (Harris & Jensen, 1985) have argued that several replications of AB designs will produce the same strength of evidence of functional control as the more elaborate designs.

Regardless of the type of design, the traditional method for determining the existence of a treatment effect was that of visually analyzing the graphs constructed from the data collection. However, some authors (e.g., Crosbie, 1993) have warned against possible errors with this method, and have argued for the use of time series statistical analyses.

One point that supporters of visual analysis of data have made is that a statistically significant change might occur in the occurrence of the target behavior, but that the change might not be large enough to make a difference to the client's everyday functioning, that is, it might not be clinically significant. This issue of clinical significance of outcomes is part of a general construct that has become known as the social validity of the intervention (Kazdin, 1977; Wolf, 1978). Social validity is considered to have three components. The first of these is the social significance of the goals of the intervention, that is, have the right behaviors been selected as targets of change. The second is the social appropriateness of the intervention strategies, that is, if they are socially acceptable. The third is the social importance of the outcomes, that is, if the outcomes are clinically significant.

Notwithstanding the debate about the relative merits of visual and statistical analysis of data, ABA has been used either exclusively or in combination with other procedures to treat a broad range of clinical problems of children. These have included both externalizing disorders such as conduct disorders (Kazdin, 1987) and attention deficit hyperactivity disorder (Barkley, 1990), and internalizing disorders such as anxiety disorders (Siegel & Ridley-Johnson, 1985) and depression (Kazdin, 1990).

One matter that has received substantial attention in the use of ABA is the extent to which clinical effects are generalized. In order to examine generalization more fully, Drabman, Hammer, and Rosenbaum (1979) developed what they referred to as the generalization map. First, they discriminated among four basic types of generalization, namely generalization across settings, behaviors, subjects, and time. Setting generalization was considered to have occurred if there was change in the target behavior in a setting other than the one in which it was treated. Behavior generalization was considered to have occurred if a behavior not targeted changed along with the target behavior. Subject generalization was said to have occurred if the target behavior changed in an individual other than the child on whom treatment was focused. Finally, time generalization was said to have occurred if a change effected in the target behavior was still in place at a later time. In regard to generalization across time, a distinction is sometimes made between those occasions on which the treatment has been terminated and those on which it is still in place. Time generalization is used when the intervention has ceased, but the term maintenance is used if the treatment is still in place.

Each of the four basic types of generalization could occur alone or in combination with another. For example, generalization across settings and behavior would be said to occur when a nontargeted behavior changed in a setting other than the treatment setting. The four basic types can be combined in sets of two, three, or four to produce a map of 16 classes of generalization. Allen, Tarnowski, Simonian, Elliot, and Drabman (1991) reviewed 904 studies published in 28 journals over a 10 year period and found that only about half reported generalization data.

It may be the case that a lack of generalization is not a problem. Take the case of a child who exhibits problem behaviors in the home, but only at home. If an intervention strategy is put in place at home, and this successfully eliminates the behaviors, the issue of setting generalization is not relevant. If some form of generalization is required, however, it needs to be planned for. The intervention program should incorporate elements that promote the likelihood of generalization occurring. The need to actively program for generalization was initially called for by Stokes and Baer (1977), and more recently Stokes and Osnes (1989) presented a description of several programming tactics grouped into three general categories.

They described the first general category as the exploitation of current functional contingencies. This included tactics such as using naturally occurring reinforcers (and punishers) in the program. New behaviors may have to be taught, but, once learned, they will be maintained by the naturally occurring reinforcers. This will help with the fostering of time generalization. Another tactic in this category is to use naturally occurring reinforcers to systematically reinforce any examples of generalization that occur. If a child is on a program to improve behavior at home, reinforcement of better behavior at school by a teacher will promote setting generalization.

The second of Stokes and Osnes's categories was called training loosely. This refers to designing intervention programs that may begin by being tightly constrained, but move over time to being more diverse. An important tactic in such training is the use of what is referred to as multiple exemplars. A child learning a concept such as dog will be better ably to correctly identify all dogs if taught using a large range of examples of a "dog." A child undergoing compliance training will display better generalized compliance if taught using a range of instruction types and is reinforced using a range of reinforcement types.

The third of Stokes and Osnes's categories was referred to as the incorporation of functional mediators. A good example of a tactic in this category is seen in social skills training programs for children and adolescents. To increase the probability that skills learned in a clinical setting will be implemented in the community, the child carries a cue card which prompts the critical elements of the response required for a particular social situation.

As was indicated earlier, ABA interventions with children typically involve the triadic model in that behavior change in the child is effected through behavior change in a significant adult in the child's environment. When thinking from within the framework of the triadic model, the issue of generalization is complex because consideration must be given not only to generalization of the child's behavior but also to that of the adult. Sanders and James (1983) identified types of generalization that were relevant to parent behavior when training parents to implement intervention programs with their children. They referred to these as generalization across settings, behaviors, children, and time. Generalization across settings refers to the parent's ability to use the child management skills learned in one setting (e.g., the clinic) in a second setting (e.g., the home). Generalization across behaviors refers to the ability of the parent to use skills learned for one child behavior with a second behavior of the child. Generalization across children refers to the ability of the parent to use skills learned for managing one child for the management of a second child. Generalization across time refers to the parent's ability to continue to use learned skills after the training program has been concluded. These types of generalization of parent behavior are clearly an extension of the types of generalization cited by Drabman et al. (1979) in relation to child behavior.

Within the triadic model of treating children's clinical problems, it must be remembered that generalization of the behavior of the significant other in the child's environment must be planned for. The suggestions offered by Stokes and Osnes (1989) for promotion of generalization apply equally to the behavior of the adults as they do to the children who are the target of intervention. Sanders and Dadds (1993) provide an excellent set of suggestions for promoting the generalization of skills learned by parents in child management training programs.

5.05.5.2 Effectiveness

A distinction between efficacy and effectiveness is offered by Hoagwood, Hibbs, Brent, and Jensen (1995). Efficacy is considered to have been demonstrated when positive results are found in tightly controlled clinical trials. Effectiveness, however, is only demonstrated when efficacious procedures are successfully implemented with heterogeneous samples in natural settings (health clinics, homes, or schools) by practitioners other than research therapists. To assist in attempts to move from studies of efficacy to studies of effectiveness, Hoagwood et al. propose a three dimensional model, with the dimensions being labelled validity, intervention, and outcome.

For increased external validity, researchers are encouraged to conduct efficacy studies in more naturalistic settings. Because a fundamental premise of ABA is that the environment is a major determinant of behavior, most ABA interventions do occur in homes, or schools, and, as such, many of the reported studies already meet this requirement. The treatment of the presenting problem in the venue in which it occurs reduces the need to build in procedures for generalization across settings. Additionally, most ABA interventions with children involve the triadic model in which a significant other in the child's environment (parent or teacher) actually implements the intervention program. This training of the significant other to do this adds further to the validity of the intervention.

For improved performance on the outcome dimension, researchers are encouraged to broaden the range of outcome measures to include variables other than defining symptoms. ABA practitioners have taken substantial steps in this regard with the search for improved generalization across behaviors, and hence a variety of areas of child functioning are measured. Also, in a desire to achieve social validity of interventions, practitioners are assessing aspects such as consumer satisfaction with the outcomes of intervention.

For improved performance on the intervention dimension, researchers are urged to use less highly structured and manualized interventions to allow more access by practitioners in the field. However, as indicated by Hoagwood et al. (1995), care needs to be taken not to sacrifice treatment integrity. The move from high to low structure without loss of treatment integrity can only be achieved with highly trained practitioners. As indicated earlier, a thorough knowledge of the principles and procedures of ABA are necessary if a practitioner is to successfully implement, and potentially "trouble shoot," intervention programs.

5.05.6 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

ABA is now well accepted as a sound model for the treatment of the clinical problems of children and adolescents. As was mentioned earlier, generalization of behavior change was a major issue in ABA. The acceptance of the need to actively program for generalization and the development of tactics for this (Stokes & Baer, 1977; Stokes & Osnes, 1989; Sulzer-Azaroff & Mayer, 1991) has made it less of an issue. Where to from here? While efficacious and effective treatment models are important, it would be generally accepted that prevention of clinical problems is always a better option than cure. It is in the sphere of prevention that ABA has much to offer.

Caplan's (1964) distinction among primary, secondary, and tertiary prevention of mental health problems is useful for a discussion of a future role of ABA. Primary prevention is absolute prevention as commonly understood. Secondary prevention refers to early identification of a problem and intervening in a manner to restrict development of it and to ameliorate any long-term effects. Tertiary prevention is intervention with fully developed mental health problems. Most of the applications of ABA up to 1997 have been in the realm of tertiary prevention. A child presents with a clinical problem and an ABA intervention is developed and implemented. There have been some excellent examples of ABA being used in secondary prevention. The work of Lovaas with autistic children is very relevant here. The long-term intensive use of ABA with young children diagnosed as autistic led to improvement in all of them, some to the extent that they could not be discriminated from typically developing children (McEachin, Smith, & Lovaas, 1993).

Researchers such as Loeber (1990, 1991) have studied the developmental pathways of conduct disorders and have identified the key risk factors associated with the development of such disorders. FAST Track is a long-term program aimed at providing intensive family- and school-based interventions for young children identified as being at risk for the development of conduct disorder (Conduct Problems Research Group, 1992; McMahon et al., 1995). It is a multicomponent program being implemented in four different locations in the USA and involves 500 high-risk children in elementary school. In addition, there are 800 high-risk or normal comparison children. All children are followed through eight years of school.

The real challenge is to have the principles and procedures of ABA adopted in mainstream family, health, and educational services with a view to the promotion of sound mental health. A useful model for conceptualizing interventions with families has been provided by Sanders and Dadds (1993), who identify five levels of intervention in families. These are described as (i) written advice only, (ii) written advice plus minimal therapist contact, (iii) written advice plus active parent training, (iv) intensive behavioral parent training, and (v) behavioral family intervention. These levels reflect differing intensities of intervention. The parent training referred to is the behavioral parent training that has been used extensively and successfully (Forehand & McMahon, 1981; Sanders & Dadds, 1993). At level (iii) it is relatively focused and helps parents manage specific problems with their children. At level (iv) it is more general and helps parents deal with a broader range of presenting problems. At level (v), the behavioral family intervention includes the training of level (iv), but addresses other matters such as marital problems or parental anxiety or depression. The interventions of levels (iii)–(v) are conducted by mental health specialists, and are examples of Caplan's tertiary prevention procedures. Interventions at levels (i) and (ii), however, are much closer to primary and secondary prevention.

Work by Sanders and his colleagues at the University of Queensland in Australia involves the development of an extensive range of written material that is available to families via level (i) and (ii) interventions. The materials, known as "tip sheets," provide advice to parents on fostering healthy development in their children. The tip sheets focus on specific parenting issues such as the management of feeding, sleeping, and compliance problems in their children. They are available to the public through a wide variety of primary care service agencies such as child health nurses, local doctors, and kindergartens. The advice given in the tip sheets is based on extensive empirical research.

The development of broad-scale, community-based, preventative interventions, such as those being developed by Sanders, is the challenge for the future use of ABA with children and adolescents.

5.05.7 SUMMARY

Clinical interventions based on the ABA approach have their roots in learning theory, particularly operant conditioning. Clinical problems are conceptualized as specific behaviors occurring too often (behavioral excesses) or too infrequently (behavioral deficits). The focus of intervention is on increasing behavioral deficits or weakening behavioral excesses. This is effected by manipulation of the antecedents and consequences that are functionally associated with those behaviors.

An important feature of ABA is its strong commitment to systematic data collection and evaluation of the success of interventions. A corollary of this commitment to data-based evaluation is that ABA is a hypothesis-testing approach. Once a target behavior has been identified for change, a hypothesis is formulated about the antecedents and consequences that are functionally related to the behavior. If success in not achieved by a manipulation of these antecedents and consequences, a new hypothesis needs to be considered.

Interventions derived from ABA have been used for the successful treatment of a range of clinical problems of children and adolescents. Early applications were in the field of mental retardation, but there is now a substantial record of success in the use of ABA for treating the broad spectrum of externalizing and internalizing disorders.

The challenge for the future use of ABA is in the realm of prevention. The learning principles which form the foundation of ABA apply to all children. The opportunity is at hand, therefore, to use these principles to construct community-based programs that either prevent the development of clinical problems (primary prevention), or that involve the early identification and treatment of clinical problems before they become major difficulties (secondary prevention).

ACKNOWLEDGMENTS

Thank you to my colleagues Jay Birnbrauer, Jan Matthews, and Roger Miller for their valuable comments on an early draft of this chapter.

5.05.8 REFERENCES

Achenbach, T. M., & Edelbrock, C. S. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, CT: University Associates in Psychiatry.

Agras, W. S., Kazdin, A. E., & Wilson, G. T. (1979). *Behavior therapy: toward an applied clinical science*. San Francisco, CA: Freeman.

Allen, J. S., Tarnowski, J. K., Simonian, S. J., Elliot, D., & Drabman, R. S. (1991). The generalization map revisited: assessment of generalized treatment effects in child and adolescent behavior therapy. *Behavior Therapy*, *22*, 393–405.

Ayllon, T., & Michael, J. (1959). The psychiatric nurse as a behavioral engineer. *Journal of the Experimental Analysis of Behavior*, *1*, 323–334.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*, 91–97.

Bandura, A. (1969). *Principles of behavior modification*. New York: Holt, Rinehart, & Winston.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Barkley, R. A. (1990). *Attention-deficit hyperactivity disorder: a handbook for diagnosis and treatment*. New York: Guilford Press.

Biglan, A., Metzler, C. W., & Ary, D. V. (1994). Increasing the prevalence of successful children: the case for community intervention research. *The Behavior Analyst*, *17*, 335–351.

Brethower, D. M. (1972). *Behavior analysis in business and industry: a total performance system*. Kalamazoo, MI: Behaviordelia.

Bry, A. (1975). *A primer of behavioral psychology*. New York: New American Library.

Caplan, G. (1964). *Principles of preventative psychiatry*. New York: Basic Books

Conduct Problems Prevention Reseach Group (1992). A developmental and clinical model for the prevention of conduct disorders: The FAST Track Program. *Development and Psychopathology*, *4*, 509–527.

Connellan, T. K. (1978). *How to improve human performance: behaviorism in business and industry*. New York: Harper and Row.

Cooper, J., Heron, T., & Heward, W. (1987). *Applied behavior analysis*. Columbus, OH: Merrill.

Crosbie, J. (1993). Interrupted time-series analysis with single-subject data. *Journal of Consulting and Clinical Psychology*, *61*, 966–974.

Drabman, R. S., Hammer, D. A., & Rosenbaum, M. S. (1979). Assessing generalization in behavior modification with children: the generalization map. *Behavioral Assessment*, *1*, 203–219.

Ellis, A. (1979). Rational-emotive therapy as a new theory of personality and therapy. In A. Ellis & J. M. Whitely (Eds.), *Theoretical and empirical foundations of rational-emotive therapy*. Monterey, CA: Brookes/Cole.

Engelmann, S. & Carnine, D. (1982). *Theory of instruction: principles and applications*. New York: Irvington.

Eysenck, H. J. (1959). Learning theory and behavior therapy. *Journal of Mental Science*, *195*, 61–75.

Forehand, R., & McMahon, R. (1981). *The noncompliant child*. New York: Guilford Press.

Grant, L., & Evans, A. (1994). *Principles of behavior analysis*. New York: Harper Collins.

Harris, F. N., & Jensen, W. R. (1985). Comparisons of multiple-baseline across person designs with AB designs with replication: issues and confusions. *Behavioral Assessment*, *7*, 121–127.

Herbert, M. (1991). *Clinical child psychology: social learning, development, and behavior*. Chichester, UK: Wiley.

Hersen, M., & Barlow, D. (1984). *Single case experimental designs: Strategies for studying behavior change* (2nd ed.). New York: Pergamon.

Hoagwood, K., Hibbs, E., Brent, D., & Jensen, P. (1995). Introduction to the special edition: efficacy and effectiveness in studies of child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology*, *63*, 683–687.

Hull, C. L. (1943). *Principles of behavior*. New York: Appleton Century Crofts.

Kameenui, E., & Simmons, D. (1990). *Designing instructional strategies: the prevention of academic and learning problems*. Columbus, OH: Merrill.

Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior Modification*, *1*, 427–451.

Kazdin, A. E. (1982). *Single-case research designs: methods for clinical and research settings*. New York: Oxford University Press.

Kazdin, A. E. (1987). *Conduct disorder in children and adolescents*. Newbury Park, CA: Sage.

Kazdin, A. E. (1990). Childhood depression. *Journal of Child Psychology and Psychiatry*, *31*, 121–160.

Loeber, R. (1990). Development and risk factors of juvenile antisocial behavior and delinquency. *Clinical Psychology Review, 10*, 1–41.

Loeber, R. (1991). Antisocial behavior: more enduring than changeable? *Journal of American Child and Adolescent Psychiatry, 30*, 393–397.

McEachin, J. H., Smith, T., & Lovaas, O. I. (1993). Long term outcome for children with autism who received early intensive behavioral treatment. *American Journal of Mental Retardation, 97*, 359–372.

McMahon, R. J., Bierman, K., Cole, J., Dodge, K., Greenberg, M., & Lochman, J. (1995, July). *The prevention of conduct disorders in school-aged children: The FAST Track Project*. Paper presented at the World Congress of Behavioral and Cognitive Therapies, Denmark.

Mahoney, M. J. (1974). *Cognition and behavior modification*. Cambridge, MA: Ballinger.

Martin, G., & Pear, J. (1996). *Behavior modification: what it is and how to do it* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Mahoney, M. J. (1974). *Cognition and behavior modification*. Cambridge, MA: Ballinger.

Meichenbaum, D. (1977). *Cognitive behavior modification: an integrative approach*. New York: Plenum.

Murphy, G. C., Hudson, A. M., King, N. J., & Remenyi, A. (1985). An interview schedule for use in the behavioural assessment of children's problems. *Behaviour Change, 2*, 6–12.

Ollendick, T. H., & Cerny, J. A. (1981). *Clinical behavior therapy with children*. New York: Plenum.

Pavlov, I. P. (1927). *Conditioned reflexes*. London: Oxford University Press.

Premack, D. (1959). Toward empirical behavioral laws. 1: positive reinforcement. *Psychological Review, 66*, 219–233.

Repp, A. C., & Singh, N. N. (Eds.) (1990). *Current perspectives on the use of nonaversive and aversive interventions for persons with developmental disabilities*. New York: Sycamore.

Ross, A. O. (1976). *Psychological disorders of children: a behavioral approach*. New York: McGraw-Hill.

Sanders, M. R., & Dadds, M. R. (1993). *Behavioral family intervention*. Needham Heights, MA: Allyn & Bacon.

Sanders, M. R., & James, J. E. (1983). The modification of parent behavior: a review of generalization and maintenance. *Behavior Modification, 7*, 3–27.

Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1948) *Walden two*. New York: Macmillan.

Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.

Skinner, B. F. (1957). *Verbal behavior*. New York: Macmillan.

Skinner, B. F. (1969). *Contingencies of reinforcement: a theoretical analysis*. New York: Appleton-Century-Crofts.

Siegel, L. J., & Ridley-Johnson, R. (1985). Anxiety disorders of childhood and adolescence. In P. H. Bornstein & A. E. Kazdin (Eds.), *Handbook of clinical behavior therapy with children* (pp. 266–308). Homewood, IL: Dorsey Press.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349–367.

Stokes, T. F., & Osnes, P. G. (1989). An operant pursuit of generalization. *Behavior Therapy, 20*, 337–355.

Sulzer-Azaroff, B. (1982). Behavioral approaches to occupational safety and health. In L. Frederiksen (Ed.), *Handbook of organizational behavior management* (pp. 505–538). New York: Wiley.

Sulzer-Azaroff, B., & Mayer, G. R. (1991). *Behavior analysis for lasting change*. New York: Holt, Rinehart, & Winston.

Tharp, R. G., & Wetzel, R. J. (1969). *Behavior modification in the natural environment*. New York: Academic Press.

Turnbull, H. R. III (Ed.) (1981). *The least restrictive alternative: principles and practice*. Washington, DC: American Association on Mental Deficiency.

Vollmer, T. R., & Iwata, B. A. (1992). Differential reinforcement as treatment for behavior disorders: procedural and functional variations. *Research in Developmental Disabilities, 13*, 393–417.

Wahler, R. G., & Fox, J. J. (1981). Setting events in applied behavior analysis: Toward a conceptual and methodological expansion. *Journal of Applied Behavior Analysis, 14*, 327–338.

Webster-Stratton, C., & Herbert, M. (1994). *Troubled families—problem children: working with parents a collaborative process*. New York: Wiley.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*, 203–214.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.06
Cognitive-behavioral Therapy

PHILIP C. KENDALL, ABBE L. MARRS, and BRIAN C. CHU
Temple University, Philadelphia, PA, USA

5.06.1 THEORETICAL UNDERPINNINGS 131
5.06.1.1 Key Features of CBT 132
5.06.1.2 Behavior and Cognition in Context 132
5.06.1.3 The Nature of Cognition 133
5.06.1.4 Developmental Sensitivity 134
5.06.1.5 Role(s) of the CB Therapist 134
5.06.1.6 Realistic Goals for Therapy 134
5.06.2 HISTORY OF THE APPROACH 134
5.06.2.1 The Impact of Cognitive Psychology 135
5.06.2.2 Self-control Clinical Interventions 136
5.06.2.3 Cognitive Therapy 136
5.06.3 ASSESSMENT AND CLINICAL FORMULATION OF THIS APPROACH 136
5.06.3.1 Developmental Sensitivity Revisited 137
5.06.4 THE EFFICACY AND EFFECTIVENESS OF COGNITIVE-BEHAVIORAL TREATMENTS 138
5.06.5 PROCEDURES AND PROPOSED MECHANISMS OF CHANGE 141
5.06.5.1 Proposed Mechanisms of Change 142
5.06.6 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 143
5.06.7 REFERENCES 144

5.06.1 THEORETICAL UNDERPINNINGS

Cognitive-behavioral therapy (CBT) with children and adolescents, relative to other theoretical approaches to psychological therapy, is a young but quickly developing area of clinical psychology. In a survey published in 1982, it was found that 40% of the new faculty in clinical psychology programs approved by the American Psychological Assoication identified themselves as cognitive-behaviorally oriented with an additional 18% identifying themselves as behaviorally oriented (Kleges, Sanchez, & Stanton, 1982). The combination of these two groups represents a majority of those surveyed. Given this presence in the training programs for the field, it is reasonable to expect that this orientation will continue to flourish and expand in the future. More importantly, however, the continued development of CBT in children and adolescents will be determined by evaluation of outcomes applied to the wide variety of psychological problems in youth (e.g., Kendall, 1991).

The achievements of CBT may be attributed to an intelligent integration of several theoretical perspectives: most notably, learning theory (behavior therapy) and cognitive theory (and therapy). CBT combines the paradigms of learning theory with an understanding of, and attention, to the cognitive activities of the client. Greatest emphasis is placed on the learning process and the influence of contingencies and models in the environment, while still reserving a central role for the individual's mediating/information processing style in understanding the development and remediation of psychological distress (Kendall, 1985).

5.06.1.1 Key Features of CBT

In addition to defining CBT as a rational integration of behavioral and cognitive theories and therapies, it may also be useful to distinguish this approach from others by examining its key features. Following a cognitive prototypes construct (e.g., Rosch, 1973, 1975) from experimental cognitive psychology, the degree to which a category member is "related" to the category in general depends on the number of features that the member shares with the category. In this case, given a list, such as the one below, of potential features of the cognitive-behavioral (CB) approach, one can determine the degree to which one would classify any given example as cognitive-behaviorally oriented. Aside from the first two features listed, none of the other elements is required for an intervention to be considered cognitive-behavioral. The first two features in the following list from Ingram, Kendall, and Chen (1991) represent core characteristics of the approach.

(i) Cognitive variables are assumed to be important causal and maintaining mechanisms. This statement does not rule out that other causal mechanisms may be operating as well, but rather that cognitive variables are important in conceptualizing the onset and course of a disorder.

(ii) Given that cognitive variables can be causal and maintaining agents, at least some of the treatment methods and strategies are aimed at cognitive targets.

(iii) A functional analysis of the variables maintaining the disorder is undertaken. This analysis may include examination of cognitive, behavioral and affective variables.

(iv) Both cognitive and behavioral therapeutic strategies are employed. As evidenced by Beck's (Beck, Rush, Shaw, & Emery, 1979) approach to the treatment of adult depression, as well as Stark and Kendall's (1996) approach with depressed youth, this combination may involve using behavioral strategies to target cognitive objectives as well as cognitive strategies to effect behavioral change.

(v) Empirical verification is sought and emphasized in at least three different ways. First, empirical research is used to examine the effects of the therapeutic procedures and to study the causal mechanisms of the changes observed. Second, the use of objective assessment during therapy provides a means of assessing therapeutic progress. Third, clients are taught how to gather data to test their assumptions and/or beliefs.

(vi) Typically, interventions are time-limited. Booster and maintenance sessions may be applied as necessary.

(vii) The client and therapist operate in a collaborative relationship, working together toward the goal of alleviating dysfunctional thinking and behavior.

(viii) CB therapists are active and offer to help evaluate suggestions rather than being passive and nondirective.

(ix) Educating the client about the CB model of dysfunction, the role of their thinking in the maintenance of dysfunction, and the need to modify dysfunctional cognition and behavior is deemed important.

(x) CB therapists serve as coping models, providing demonstrations in role-plays of ways to manage dysfunction and different situations.

A few words on what CBT is not. Unlike the psychodynamic perspective which concerns itself in part with efforts to uncover unconscious early experiences and trauma, CBT focuses more on the present, but acknowledges that early experiences are influential in learning behavior patterns and cognitive styles. Similarly, CBT does not focus on biological, neurological, or genetic aspects although again it acknowledges that they play a role, to varying degrees, in several disorders.

5.06.1.2 Behavior and Cognition in Context

Although it is clear from the listed features that CBT places major emphasis on both the learning process and mediating/information-processing factors, CBT also considers the relationship of cognition and behavior to affective state and functioning in the larger social context. This aspect of CBT is especially important in the treatment of children and adolescents because of the centrality of the social context in the developmental trajectory of children. Specifically, adequate relations with peers is a crucial component of successful adjustment, and consequently an approach which is sensitive to these issues both during the assessment and intervention phases of treatment is favored.

Equally, if not more, important, in considering children in context, is an understanding of the family. Many of the rules and roles for later social interaction are first introduced in this social microcosm. Consideration of the modeling and reinforcement practices parents use with their child can provide clues to the therapist about factors that may be maintaining the targeted behavior. For example, a parent who has perfectionistic expectations of their child, and who only rewards the child for flawless performance, may be a contributing factor in the maintenance of the child's anxiety in testing situations or in the child's overly harsh criticisms of themselves. Additionally, parents may prove to be a valuable resource as diagnostic and therapeutic collaborators: first, by providing information about the nature of the problem as it manifests itself outside the clinic, and second, by reinforcing at home lessons taught and gains made in therapy.

5.06.1.3 The Nature of Cognition

Given the centrality of cognition, an account of CBT requires a discussion of the more detailed features of cognition: cognitive structures, cognitive content, cognitive processes, and cognitive products (Ingram & Kendall, 1986; Kendall, 1993; Kendall & Ingram, 1987, 1989). Cognitive structures can be thought of as mental templates that are based on an accumulation of experiences in memory which act to guide and filter the understanding of new experiences. Cognitive content refers to ongoing self-talk, as well as the information that is stored and organized in memory. Self-talk or internal speech are examples of surface cognitive content. Cognitive structures serve to trigger automatic cognitive content and information processing. Cognitive processes are the procedures by which the cognitive system operates—how we go about perceiving, recalling, and interpreting experience. Search and storage mechanisms and inferential and retrieval processes are examples of cognitive processes. Cognitive products are the end result of the operations of the system. These are the thoughts that an individual experiences as a result of the interacting content, process, and structure, and can be illustrated by causal attributions

Psychopathology (and psychological therapy) may be related to problems (and corrections) in any or all of the above areas of cognition, and the understanding of specific disorders and their treatment benefits from consideration of each of these factors for each client. For example, an anxiety-disordered youth who has a schema for threat (cognitive structure) about danger (evaluative and physical) would have automatic thoughts (questions) and an internal dialogue (cognitive content) that would be consistent with that schema (worry). The cognitive process that this individual engaged in would be influenced by the cognitive structures and would therefore be likely to consist of selective attention to cues for evaluation and risks of negative outcomes. Consequently, the individual would be predisposed to see danger in events and to hold demanding expectancies about the future.

Cognition can be divided further. Just as child behavior disorders can be broken down into internalizing and externalizing varieties where internalizing behavioral disorders involve problems within the self and externalizing disorders pertain to conflict with the outside world (Achenbach, 1906), so too can dysfunctional cognitions be subdivided. Two separate categories for cognitive dysfunctions have been described (Kendall, 1985, 1993): cognitive deficiencies and cognitive distortions. Deficiencies refers to an absence of thinking where it would be beneficial (i.e., acting before thinking). Children and adolescents with these deficits in information processing fail to engage in forethought and planning. Conversely, children and adolescents who have distorted cognition are not failing to engage in information processing but are rather doing so in a dysfunctional, "crooked" manner. On a historical note, the distinction between deficient and distorted cognition highlights the differences between the forerunners of adult CBT which focused primarily on distorted thinking, and early child CBT which was mainly intent upon remediation of cognitive deficits in children (Kendall, 1977; Meichenbaum & Goodman, 1971).

Although the distinction between distorted and deficient cognition often runs parallel to the distinction between internalizing and externalizing behavior (i.e., disorders associated with distorted cognition also tend to be associated with internalizing behavior, anxiety, and depression; and deficient cognition with externalizing disorders, attention-deficit hyperactivity disorder (ADHD), tend to group together), the associations are not perfect (Kendall & MacDonald, 1993). For example, aggressive behavior (externalizing) entails cognitive deficiencies and cognitive distortions (e.g., Dodge, 1980; Dodge & Frame, 1982; Lochman, 1987).

Some consider the frontier in current CBT conceptualizations of cognition to be the constructivist narrative. Proponents of this view (e.g., Epstein & Erskine, 1983; Guidano & Liotti, 1983; Harvey, Weber, & Orbach, 1990; Mahoney & Lyddon, 1988; McNamee & Gergen, 1992; Meichenbaum & Fitzpatrick, 1993; Neimeyer, 1993; Neimeyer & Feixas, 1990; White & Epston, 1990) see the human mind as a product of constructive symbolic activity and see reality as a product of personal meanings that individuals create. In contrast to a solely information-processing view of cognition, this perspective holds that there are many realities, and thus the distortion of a universal reality can no longer be seen as a viable metaphor for explaining the existence of psychopathology. This evolving CB approach maintains instead that some realities are more functional than others for given individuals in given circumstances. A goal of the therapist then becomes one of helping clients become aware of how they create their realities and the ensuing consequences of those constructions.

5.06.1.4 Developmental Sensitivity

CBT does *not* subscribe to what Kendall (1984) described as "the developmental uniformity myth." Rather, children are not homogeneous and CBT acknowledges that treatment may be differentially effective as a function of a child's developmental level. This developmental sensitivity also implies that each individual child is not "constant" from pre- to post-treatment. Relatedly, one must note that children with the same behaviors but of different ages are not necessarily alike. Avoiding the developmental uniformity myth means giving adequate consideration to age (age groups or developmental groups) as a potential moderator variable of treatment effects, and creating interventions which are intentionally geared specifically to the child's age or level of cognitive development. This issue is considered further in Sections 5.06.3 and 5.06.5.

5.06.1.5 Role(s) of the CB Therapist

Elsewhere, Kendall (1993) used the terms "consultant," "diagnostician," and "educator" to refer to the therapeutic posture of the therapist. As a consultant (collaborator), the therapist is viewed as having ideas worthy of trying but not as the one who has all of the answers. A goal is to give the client the opportunity to try new ways of thinking and behaving and to help them make sense of the experiences and the results. The therapist intentionally acts as a problem-solving model as one way of encouraging the client to become an independent and mature problem solver. As a diagnostician, the therapist integrates data gathered from many sources: synthesizing new information with knowledge about psychopathology, normal development, and psychologically healthy environments. The outcomes of this process are well reasoned decisions about the nature of the problem and the optimal strategy for its treatment. As an educator, the therapist attempts to present the information of therapy in a manner which facilitates the client's learning-specific behaviors, cognitive skills, and emotional development. Helping clients to think for themselves is an ultimate goal. Thinking for oneself may provide the maximum possibility for the generalization of skills learned in therapy to other problematic areas which may or may not be present at the time of treatment. In this way, although CBT is not aimed at providing a cure *per se* for a client's problems (Kendall, 1989), the changes produced by the intervention can be used in the management of psychopathology in areas beyond those where they were initially applied.

5.06.1.6 Realistic Goals for Therapy

In helping a client cope with problems (rather than striving to cure the problem), CBT provides the client with realistic expectations about potential outcomes. Psychological change is not linear, and lapses/relapses are likely. CBT prepares clients for the fact that booster sessions may be necessary and helps them view lapse in a constructive manner rather than as a failure. Implicit in this idea is the notion that all forms of psychopathology are not equally responsive to psychological interventions. This realistic appraisal helps the client and therapist focus on doable change that is likely to occur.

Although little written attention has been devoted to factors such as trust, respect, and relationship as part of the CBT process, this should not be taken as an indication that they are not important elements of this treatment. Our discussion has focused on elements of treatment that are specific to CBT and because those factors are essential elements of all forms of treatment, it was not deemed necessary to discuss them specifically here.

5.06.2 HISTORY OF THE APPROACH

Consideration of the origins of CBT is helpful in setting the current features of the approach in their proper context. However, all accounts of historical information can only represent a limited perspective on the events of the past. This account of the history of CBT with children was shaped in large part by the work of Meichenbaum (1995) and Meyers and Craighead (1984).

In some senses Witmer (1907) could be credited with the earliest example of contemporary behavior therapy. His psychological clinic, established in 1896, was the first of its kind. His interventions with childhood academic and mental health problems had many of the hallmarks of current CB approaches: they were directive, educational, based on principles from perception and learning, and empirically evaluated for effectiveness.

More traditional accounts of the birth of behavior therapy with children, however, usually begin with the work of Watson (1924) and his colleagues Rayner (Watson & Rayner, 1920) and Jones (1924a, 1924b). Together these pioneers in behavior therapy with children studied the effects of conditioning on the development and amelioration of fears. Somewhat later, although equally durable to the effects of time is the "bell-and-pad" procedure originally introduced by the Mowrers (Mowrer & Mowrer, 1938) for the treatment of enuresis.

These early endeavors were all but ignored until the 1960s, however, due to the prevalence of the psychoanalytic treatment of psychological problems at that time. When the psychoanalytic era waned, it was Rogers' (1951) client-centered therapy which emerged as an approach that was especially embraced by school psychologists. By the late 1950s and early 1960s, most psychologists, psychiatrists, social workers, and school psychologists were either using psychoanalytic or client-centered techniques with their child clients.

The therapeutic effectiveness of the prevailing approaches began to be questioned: in child-clinical circles, the critiques of psychodynamic psychotherapy (Levitt, 1957, 1963) were based on the lack of data to support the effectiveness of the psychodynamic approach. Early inroads by behavior therapists were seen in the treatment of severely disordered children. Perhaps because of the level of difficulty associated with treating these clients with the traditional methods of the time and perhaps because of the appropriateness of the behavioral approach, these early applications of behavior therapy to clinical cases were welcomed and successful. The majority of clinical applications of behavioral principles during this time focused on the application of operant procedures, with some ancillary use of modeling and systematic desensitization used specifically for anxiety and phobic disorders.

Fueled by empirical literature, clinical behaviorism took hold. The potential to predict and modify abnormal behavior through the study of the stimulus–response links thought to account for its occurrence created much excitement and anticipation in the field. The boom of behavior therapy was further bolstered by the creation of several scholarly journals (e.g., the *Journal of Applied Behavior Analysis*, created in 1968) and professional associations which specialized in behavior therapy and applied behavior analysis (e.g., the Association for the Advancement of Behavior Therapy).

Experimental and developmental psychologists were among the first to reconsider cognitive constructs and models as possible moderators of behavior. Clinical and social psychology also began to incorporate theories of cognition into what had previously been strictly behavioral models. Although the shift away from strict behavior therapy toward cognitive behavior therapy was marked by more than a few bumps in the road, this process should be considered more of an evolution than a revolution (Meichenbaum, 1995). According to Meyers and Craighead (1984), the shift toward CBT with children was influenced by three primary factors: cognitive psychology, self-control clinical interventions, and cognitive therapy.

5.06.2.1 The Impact of Cognitive Psychology

Cognitive psychology influenced behavior therapy with children in several ways. First, social learning theorists (e.g., Bandura, 1969; Kanfer & Phillips, 1970; Mischel, 1973; Rotter, 1966) and others provided explanations of phenomena which had previously been explained solely by behavioral principles. Perhaps the most long-lived of these has been the account of modeling effects. Bandura suggested that in addition to being explained by the principles of learning, observational learning was also influenced by the processes of attention, retention, motor reproduction, and incentive and motivation. Not only was this expansion into cognitive concepts, such as attention and retention, innovative and groundbreaking in its own right, but it also set the stage for more flexible and expansive theorizing than was previously the case under strict behaviorism.

A second avenue of impact for cognitive psychology came from research findings on the inferential and decision-making processes of individuals (e.g. Kahnemann, Slovic, & Tversky, 1982; Nisbett & Ross, 1980; Tversky & Kahnemann, 1974). These researchers and others drew the field's attention to the fact that individuals are often biased, inefficient, and incompetent information processors, with poor decision-making abilities. They mounted evidence for the fact that decision making was often influenced by habits of thinking, or mental heuristics, rather than being based on real events. This information provided a much needed metaphor for understanding psychopathology from a cognitive perspective.

Problem-solving and coping skills interventions were developed to put findings from cognitive psychology to work in clinical practice (e.g., D'Zurilla, 1986; D'Zurilla & Goldfried, 1971; Meichenbaum, 1985; Spivack & Shure, 1974). Programs were designed to teach and evaluate social problem-solving thinking: alternative thinking—generating multiple potential solutions to interpersonal problems; means–end thinking—considering the step-by-step process necessary to reach a desired goal; and, consequential thinking—focusing on the consequences that occur as a result of choosing a particular behavioral response (Spivak & Shure, 1974). The early success in teaching children problem-solving skills has led to much work in this area.

Based on the work of the cognitive developmental psychologists Vygotsky (1978) and Luria (1976), Meichenbaum and Goodman (1971) developed a self-instruction program to teach impulsive children how to control their behavior. Vygotsky and Luria suggested that during development, children's behavior is first under the verbal control of the social environment (most often by adults), and then gradually under personal control through overt speech and then slowly through covert speech. Applying these principles, Meichenbaum and Goodman's procedure entailed having an experimenter model overt behavior and associated self-statements for the impulsive child. The success of this early trial of self-instruction training has led to its application in a wide variety of other child and some adult psychopathology fields (e.g., Craighead, Wilcoxon-Craighead, & Meyers, 1978; Kendall & Williams, 1981; Meichenbaum, 1977).

5.06.2.2 Self-control Clinical Interventions

Self-control interventions, like many other influences on strict behavior therapy, were originally explained by the same operant models which formed the core of behavior therapy. During the late 1960s, however, cognitive explanations for self-control procedures began to develop and to spark researchers to think about changing internal thought processes as a means of teaching the individual self-control (e.g. Mahoney, 1974; Kendall & Braswell, 1982; Kanfer & Karoly, 1972; Craighead, Brownell, & Horan, 1981).

5.06.2.3 Cognitive Therapy

Unlike self-control programs, cognitive therapy developed outside and separate from the realm of behavior therapy. Working almost exclusively in the clinic as opposed to in the laboratory, Ellis and Beck, two major leaders in the development of cognitive therapy, were separately developing interventions aimed specifically at altering maladaptive cognitive processes. Ellis' (1962) rational-emotive therapy focused on the idea that individuals behave maladaptively because of irrational and illogical thought processes. It is what individuals think about the things they say and do that has the potential to upset them. Beck's (1976) approach to cognitive therapy also saw maladaptive cognitions as a culprit in the etiology of psychopathology. Beck argued that these maladaptive thought processes were a reflection of an individual's assumptions and beliefs about themselves and the world. Cognitive therapy focuses on modifying the fundamental assumptions and beliefs on which the maladaptive thoughts are based. In both cases, the success of these approaches in treating adults has greatly influenced the development of CBT.

5.06.3 ASSESSMENT AND CLINICAL FORMULATION OF THIS APPROACH

CBT's approach to assessment has two foci. First, the assessment process is viewed as a hypothesis-testing endeavor. These hypotheses then guide the process of understanding a given child in their social context so that appropriate treatment strategies can be implemented to ameliorate the child's current difficulties (Ollendick & Hersen, 1993). Unlike more traditional assessment methods which are focused on searching for underlying personality characteristics or traits that are assumed to be causal factors in current distress, the hypothesis-generating process places primary importance on clearly describing current behavior, thought patterns, affective understanding, and environmental conditions which are thought to be related to the presence and maintenance of the child or adolescent's distress. Also, in contrast to traditional assessment techniques, CBT assessment strives to take a sampling as opposed to a symbolic approach to choosing and interpreting the results of testing. That is to say, CBT assessment techniques, being based on observable behaviors and reliably measured cognitive and environmental variables require little inferential work. More traditional assessment procedures, in contrast, rely on viewing the testing results as an indirect manifestation of underlying personality traits which are thought to be causing the presenting problems. Accurate sampling of behavior and cognition which are similar if not identical to the behavior and cognition which are targeted by the intervention is therefore a cornerstone of the CB assessment process.

Second, CBT views the assessment process as a means of evaluating the effectiveness of specific intervention procedures. Given the importance CBT places on the empirical evaluation of its treatment procedures, this is an essential part of the assessment process. Being able to report that a given behavior or cognition changes from pretreatment to post-treatment requires careful measurement of these elements at both time periods. Establishing that the observed changes are due to the intervention and not to some other confounding variable, however, requires additional consideration and controls.

5.06.3.1 Developmental Sensitivity Revisited

In child and adolescent treatment, a potential threat to the internal validity of the treatment is the effects of maturation. Children and adolescents are changing organisms who may simply "grow out of" their original presenting problems. To assert that the observed changes from pre- to post-treatment are due to the intervention and not to maturation, one must not only accurately measure the target constructs at pre- and post-treatment, but one must also know what norms are appropriate for comparison purposes at each time. The norms used prior to treatment may no longer be appropriate following treatment. CBT assessment procedures strive to be developmentally sensitive by using age-appropriate data on which decisions about the efficacy of treatment will be based.

A method for treatment evaluation might be to use the "normative comparisons" approach to clinical significance (Kendall & Grove, 1988). In addition to using statistically significant changes as a barometer of effectiveness, this approach also evaluates interventions based on the child's level of functioning in comparison to a meaningful and representative nondisturbed reference group. Because of maturational factors, it is essential that normative reference groups match clients on age. When a client, after statistically meaningful treatment effects, has returned to symptom levels indistinguishable from normals, then the normative comparison method can be said to have provided evidence of clinically significant change.

CBT assessments are geared to be developmentally sensitive because, in deciding if treatment is needed for a given problem or in deciding what aspects of a child's presenting picture are outside of the normal range, a clinician needs to be aware of the appropriate group with which to compare a child's behavior and cognition. For example, it is more acceptable for a six-year-old to exhibit mild features of separation anxiety than for a 14-year-old to exhibit the same pattern.

A further challenge to the accurate assessment of children and adolescents is that manifested symptoms of a given disorder may vary with developmental level. This complicates the diagnostic picture for youth because the behavior and cognition targeted by a certain intervention may vary and other symptoms of the same problem may take their place. It is analogous to trying to hit a moving target. For instance, Strauss, Lease, Last, and Francis (1988) reported that the American Psychiatric Association's *Diagnostic and statistical manual of mental disorders (third edition, revised) (DSM-III-R)* category of overanxious disorder was most often typified by unrealistic worry about the appropriateness of past behavior and comorbidity with major depression and simple phobia in children and adolescents aged 12–19; in contrast, in children aged 5–11 this type of worry was often not endorsed and comorbidity was more often seen with separation anxiety and ADHD.

It is somewhat unclear whether the observed differences in symptom manifestation across developmental levels are due to changes in comprehension and/or expression abilities or in some other factor which changes with developmental stage. This being so, it is important for classification systems used to diagnose children to be sensitive to these issues. For this reason, among others, the practice of simply scaling down adult syndromes to be applied to youth is not considered appropriate or desirable. While our current system of classifying childhood psychopathology takes into account the fact that children possess the same capacities as adults in terms of: influencing others through social interaction; processing and interpreting information in their environments; and, making choices that shape their worlds; it has also evolved to include specific diagnostic categories which may be unique to childhood, and may be more congruent with how a child or adolescent would present with a given difficulty. Diagnostic systems need to continue to develop as more data about the reliability and validity of the childhood diagnostic categories become available (Callahan, Panichelli-Mindel, & Kendall, 1996).

The challenge of collecting assessment data from children and adolescents is formidable. One facet of this issue is that assessment instruments are created by adults, and although some progress has been made in developing assessment tools that are geared for younger clients, this is still an area in need of attention. For example, a very common way of assessing the degree to which someone experiences a specific symptom is through a rating scale with approximately five anchor points ranging from "never" to "always." Evans and Nelson (1986) presented data that call into question a young child's ability to make the fine-grain distinctions which are required by these five-point scales. Some investigators (e.g., Finch & Montgomery, 1973; Sheslow, Bondy, & Nelson, 1982) have used pictorial assessment instruments instead of verbal ones to address this problem.

To increase the likelihood of obtaining assessment data that is reliable and valid, CB theorists agree with the suggestion to use multiple modes of assessment. Typically, paper-and-pencil questionnaires, structured interviews, and behavioral observations are cornerstones of CB assessments. No single approach perfectly captures the complexity that is involved in developing a case profile for a child or an adolescent, but the hope is that through using multiple methods one can begin to triangulate on the truth. It is important to emphasize, however, that this combination of assessment modalities is only as strong as its weakest link, and therefore the compilation of multiple methods does not excuse the assessor from considering reliability and validity data for each assessment tool used.

Another consideration when attempting to formulate a treatment plan for a child or adolescent is that youth rarely refer themselves for treatment. Usually a concerned adult (teacher, parent, or other adult figure in the child's life) brings the child to treatment. This may mean that the child or adolescent does not acknowledge that they have a problem requiring treatment and either from obstinacy or ignorance may not be able to provide the assessor with enough information on which to formulate the case. With some disorders, the child or adolescent's inability to provide the needed information may be directly related to the presenting problem. For example, an anxious child with perfectionism and need for social acceptance may be especially apprehensive about sharing their problems with the assessor because of a fear of being perceived negatively.

For the parent or referring adult's part as well, there may be reasons not to rely solely on the information provided. For instance, consistent with Harris and Ferrari (1983), it has been speculated that parents cannot consistently evaluate their child's behavior objectively; they may hold distorted perceptions and/or unrealistic expectations for their child. This inaccuracy may be related to something as simple as unfamiliarity with what is appropriate behavior for children of certain ages or it may be related to issues and problems that the parent is experiencing. CB theorists do not believe that a child's behavior is entirely caused by underlying personality traits that are constant across situations, so it also stands to reason that a parent or teacher who is observing a child in only a limited number of environments could not be expected to provide complete information about the nature of a child's difficulties. Therefore, whether through lack of information or inappropriate expectations, assessment of children and adolescents is best when drawn from many sources. Typically these include self-, parent-, and teacher-reports, and shared experiences with the therapist.

5.06.4 THE EFFICACY AND EFFECTIVENESS OF COGNITIVE-BEHAVIORAL TREATMENTS

Two terms have appeared for evaluating a treatment's success: efficacy and effectiveness. Efficacy and effectiveness studies can be distinguished by the amount of control exerted by the experimenter in the study. Hoagwood, Hibbs, Brent, and Jensen (1995, p. 683) defined efficacy studies as those in which "considerable control has been exercised by the investigator over sample selection (usually recruited samples), over delivery of the intervention, and over the conditions under which the intervention or treatment occurred." The typical example of an efficacy study is a randomized clinical trial in a structured university clinic setting. An intervention's effectiveness is typically tested using a more heterogeneous sample of clients or a more naturalistic setting (e.g., home, school, or more general medical setting), or by having treatment provided by real-world practitioners rather than research therapists (Hoagwood et al., 1995). The primary purpose of effectiveness studies is to evaluate how viable, or how transportable, an efficacious treatment is in real-world settings which often include a wide array of environmental factors that may impact treatment success.

Randomized clinical trials of psychological therapy for children and adolescents have been scant in comparison to studies of adults. Nevertheless, compared to other child treatment modalities, CBT is one of the more frequently and rigorously evaluated therapeutic interventions. In one study, analyzing the characteristics of child and adolescent research from 1970 to 1988, behavior modification was found to be the focus of 49.5% of all therapy studies and cognitive-behavioral approaches the focus of 22.1% (Kazdin, Bass, Ayers, & Rodgers, 1990). The majority of the studies surveyed were conducted in research-oriented clinics. Recently, the transportability of treatment from well-controlled research experimentation to more real-world practicing clinical settings has become a research question (Hoagwood et al., 1995; Kendall & Southam-Gerow, 1995). Among the differences between clinical research and clinical practice, clients seen in research settings are thought to present with a restricted range of disorders with less severe symptomatology (though this has not been the case in recent research), due to exclusion criteria set by the experimenter. Therapy in research clinics is also often conducted by specially trained therapists with small caseloads in well-controlled environments (Weisz, Weiss, & Donenberg, 1992). These differential specifications and conditions under which treatment is provided can be a source of concern when considering evaluations of treatment methods and processes.

A growing list of well-controlled therapy outcome studies has contributed to establishing the efficacy of CBT for children. CB interventions designed for children with depression have produced reductions in self-reports and clinical ratings of depressive symptoms in comparison to waitlist controls (Reynolds & Coats, 1986; Stark, Reynolds, & Kaslow, 1987), accompanied by reductions in anxiety and improvement in academic self-concept (Reynolds & Coats, 1986). Similarly, in a comparison of CBT with and without parental involvement, Lewinsohn et al. (1990) found that compared to a waitlist condition, both treatment conditions resulted in greater reductions in depressive symptomatology and percentage of clients meeting diagnostic criteria for major depressive disorder (MDD) (i.e., 57.1% of children in individual treatment, and 52.4% of individual plus parent treatment met criteria for depression at post-treatment, vs. 94.7% of waitlist). CBT has also demonstrated some superiority over traditional counseling modalities. For example, Stark, Rouse, and Livingston (1991), in a direct comparison of CBT to a traditional counseling program (duration of 24–26 sessions), found that CBT led to greater improvements in depression and reductions in depressive thinking (see also Stark & Kendall, 1996; Stark et al., 1996).

With regard to long-term treatment effects for children with depression, treatment gains have been maintained over short-term follow-up periods of 5–8 weeks (Reynolds & Coats, 1986; Stark, Reynolds, & Kaslow, 1987). However, longer-term follow-up evaluations have yielded more inconsistent results. For example, Stark et al. (1991) noted that some treatment gains were not maintained at seven-month follow-up. Conversely, Lewinsohn et al. (1990) found that depressed adolescents who had received CBT treatment continued to show increased recovery rates over a 2-year follow-up period, with only 17% still meeting *DSM* criteria for MDD at 6, 12, and 24 months.

Cognitive-behavioral interventions with anxiety-disordered youth, a clinical population often overlooked because their behavior may be less disruptive than children with externalizing disorders, have received empirical support (Kendall, 1994; Kendall et al., in press). In comparison to an eight-week waitlist control, a 16-week CB intervention (see Kendall, Kane, Howard, & Siqueland, 1990) produced both statistically and clinically significant improvements on parent-, child-, and teacher-reported measures for 47 children diagnosed with a primary anxiety disorder (Kendall, 1994). These improvements were evident on measures of anxiety, internalizing behavior problems, and depressive symptomatology. Nearly two-thirds of participants no longer met diagnostic criteria for their primary anxiety disorder following treatment. Children participating in the CBT group showed clinically significant gains by returning to within nondeviant limits on many of these measures. Gains seen at post-treatment were also maintained at one-year (Kendall, 1994) and longer-term (3.35-year) follow-ups (Kendall & Southam-Gerow 1996). In a second randomized clinical trial, Kendall et al. (in press) replicated the efficacy of this protocol with 94 anxiety-disordered children. In addition to demonstrating similar treatment gains as the previous trial (Kendall, 1994), comorbid status was not shown to moderate outcome. The program was effective when applied with families (Howard & Kendall, 1996), and work done in Australia, using a modified version of this program (Barrett, Dadds, & Rapee, 1996), provides a cross-cultural and independent replication of favorable treatment outcomes.

CBT has been applied with children exhibiting aggression and antisocial behavior. In a randomized clinical trial of 112 children (aged 7–13) exhibiting antisocial behavior, Kazdin, Bass, Siegel, and Thomas (1989) found significantly greater therapeutic effects for two forms of CBT in comparison to relationship therapy (RT). The two CBT interventions were a problem-solving skills training (PSST) and a problem-solving skills training with *in vivo* practice (PSST-P). At post-treatment, outcome measures revealed significant treatment effects for parent-, teacher-, and child-completed measures. Although all three treatment groups led to some improvements over the course of treatment, the magnitude of the improvements and breadth across measures were greater for children in the PSST and PSST-P groups than for those in the RT group. The improvements of the two PSST groups were particularly superior to those of the RT group in the areas of prosocial measures (social competence and global adjustment) and on diverse scales that reflect problem behavior at school. In this study, the practice component appeared to enhance PSST. In another study (Kazdin, Siegel, & Bass 1992), a parent management component enhanced therapeutic effects over PSST in child and parent functioning and placed a greater proportion of youth within normative levels of functioning.

Other studies have found that CB therapeutic effects may lead to improvement in social-cognitive processes associated with aggressive behaviors (e.g., disciplining for fighting) or improvements in interpersonal skills, but may not lead to change on ratings on teacher reports (Baum, Clark, McCarthy, Sandler, & Carpenter, 1986; Kettlewell & Kausch, 1983; Lochman, White, & Wayland, 1991). The success of CBT in the treatment of aggressive youth is also tempered by difficulties with generalization of treatment effects and persistently elevated levels of disruptive behaviors in normative comparisons. For example, long-term follow-up assessments suggest that decreases in aggressive or delinquent behaviors are often not maintained over time (Kazdin, 1987; Kendall & Braswell, 1982; Lochman, 1992), or are not generalized across situations, for example, classroom behaviors (Baum et al., 1986; Lochman, Burch, Curry, & Lampson, 1984).

To aggregate findings across studies, researchers have used meta-analysis. Meta-analysis, with reference to therapy outcomes, provides a format for quantitatively assessing the effects of particular treatments across studies. In general, behavioral and CB therapies have compared favorably to other theoretical approaches. In meta-analyses including only controlled studies, the mean effect sizes for CBT have ranged from 0.56 to 0.81, for behavioral therapies (including studies of CBT) from 0.88 to 0.91, and for nonbehavioral therapies (including client-centered and psychodynamic therapies) from 0.40 to 0.44 (Casey & Berman, 1985; Durlak, Furhman, & Lampman, 1991; Weisz, Weiss, Alicke, & Klotz, 1987). In general, findings have supported the importance of the behavioral component of CBT (Durlak et al., 1991; Weisz et al., 1987). A meta-analysis that compared controlled outcome studies conducted in both research and practicing clinics found that only the use of behavioral interventions and the treatability of clients (severity of symptoms and presence of comorbidity) predicted treatment outcome (Weisz, Donenberg, Han, & Weiss, 1995).

One meta-analysis focused specifically on CBT. Durlak et al. (1991) reviewed 64 reports of controlled CBT studies conducted between the years of 1970 and 1987 involving children between the ages of five and 13. The overall effect size (ES) for CBT was 0.56. This overall ES represents a moderately beneficial treatment effect, but it should be noted that children who had reached the cognitive stage of formal operations (ages 11 to 13), displayed a large treatment effect (ES = 0.92) in response to CBT. The ES in older children was twice as large as that for children in less advanced cognitive stages (0.57 for ages five to seven; 0.55 for ages eight to 11). The analyses provided data that are consistent with two assumptions of CBT: first, that cognitive developmental level is a factor in a client's response to treatment; and second, that activating cognitive processes may be one of the specific treatment components of CBT.

Durlak et al.'s (1991) review also demonstrated that children undergoing CBT exhibited clinically meaningful gains. In studies comparing treated groups to normative samples, (see also normative comparisons; Kendall & Grove, 1988), treated children displayed clinically meaningful gains (i.e., changed from being beyond normal limits at pretest to being well within normal limits at post-treatment) on measures of personality functioning and cognitive tempo. Contrary to expectations, clinically significant changes were not obtained on behaviorally oriented measures in children who underwent treatment. Although scores on behavioral measurements did decrease in the predicted direction, they did not decrease enough to return youths to within normal ranges. In general, after age was taken into account, no other factors were significant in moderating treatment outcome.

A challenge facing "effectiveness" researchers is finding ways to employ interventions in a fashion that will reach wider populations without sacrificing treatment integrity. Clinical research has traditionally placed internal validity (the extent to which the treatment works) in a position of primary importance above external validity (a treatment's generalizability). This is not unreasonable because many interventions have yet to be empirically supported (Kazdin, 1988). Issues of internal and external validity are of interest to the effectiveness researcher.

Psychological treatment conducted with children in practicing-clinical settings has traditionally enjoyed less success than psychological interventions performed in research-based clinics (Weisz & Weiss, 1989). In Weisz et al.'s (1995) analysis of this discrepancy between research and clinical interventions, it was demonstrated that improvement differences could be accounted for by the use of behavioral interventions and the relative treatability of the client population (severity of symptoms, less presence of comorbidity) in research clinics. Thus, two possibilities for improvement in interventions in clinical settings might be to employ more behavioral techniques, and to specialize clinics so that treatment can focus on circumscribed problems. Research clinics must also progress towards developing interventions that lend themselves better to real-world clinical practice. This may include developing interventions capable of addressing multiple problems sequentially or simultaneously.

5.06.5 PROCEDURES AND PROPOSED MECHANISMS OF CHANGE

Cognitive-behavioral therapy is an active, structured approach to treatment that aims to produce change in an individual's dysfunctional cognition and maladaptive behavior. Its procedures are based on the assumption that cognitive variables are important causal and maintaining mechanisms in the onset and course of a disorder, but that affective and behavioral variables are interrelated and influential as well. Specifically, CBT has been defined as those sets of therapeutic procedures that: (i) place primary importance on the causal/maintaining role of cognitive processes; (ii) use at least some therapeutic procedures aimed at altering aspects of cognition; (iii) examine cognitive, behavioral, and affective variables maintaining the disorder; (iv) may use both behavioral and cognitive techniques to target cognitive objectives, (v) emphasize empirical verification in research, assessment, and treatment; (vi) tend to be time-limited with the possibility of maintenance sessions; (vii) promote a collaborative relationship between client and therapist; and (viii) recommend active therapist participation during sessions (Ingram, et al., 1991).

Given this framework, one therapeutic pathway lies in remedying the identified dysfunction, focusing on the provision of skills and opportunities to practice these skills, and beginning new learning trajectories. One strategy for reaching this goal is the "consultative" role (also called "collaborative empiricism"; Hollon & Beck, 1979), in which the client and therapist work together as active collaborators in the identification of problems, the design and execution of tests of specific hypotheses, and the reanalysis of beliefs (Kendall, Vitousek, & Kane, 1991). As therapy progresses, the client learns to become an active and skeptical social scientist who has the ability to evaluate critically their own ideas, beliefs, and resulting behavior. Cognitive-behavioral theorists believe that data generated by the client in unbiased experiments, more than a therapist's credibility, persuasiveness, or authenticity, are the instigators of change. Treatment emphasizes inductive strategies for the modification of beliefs, operating from the assumption that "an actual demonstration of a belief's invalidity will go further to change a belief than will a discussion with the therapist" (Piasecki & Hollon, 1987, p.123).

A second therapeutic pathway lies in changing the valence of the child's self-talk. In one study of the treatment of youth with anxiety disorders (Treadwell & Kendall, 1996), negative self-talk was found to play a role in both predicting anxiety, and mediating treatment gains. It was hypothesized that negative self-talk, as opposed to positive self-talk, is predictive of maladjustment (Kendall, 1985). The results (Treadwell & Kendall, 1996) indicated that higher functioning was linked to fewer negative self-statements, but not to greater positive self-talk. Moreover, as children reduced their anxious distress, the accompanying negative self-talk became less frequent. The reduction of negative self-talk was found to mediate gains from CBT.

Procedures used by many CBTs include: affective education, relaxation training, social problem-solving, cognitive restructuring/attribution retraining, contingent reinforcement, modeling, role plays, and training workbooks. The behavioral component of CBT (e.g., modeling, role plays) is designed to illustrate the connection between mood and behavior and to provide practice opportunities for new skill-building. The cognitive component (e.g., cognitive restructuring, problem solving) is designed to teach the child to identify, re-evaluate, and replace maladaptive thoughts. Some techniques necessarily draw from both behavioral and cognitive concepts emphasizing the connection between thought, affect, and action.

Affective education teaches the child to recognize, label, and self-monitor physiological and emotional cues and, eventually, control behavioral reactions that accompany these cues. It focuses on gaining a greater understanding of one's outward manifestations of emotions to become aware of the situations that pose particular difficulties. Relaxation training, which is often tied to affective education, may be considered a first line of defense for children against increased levels of arousal. An objective in relaxation training is to teach the child to manage the physiological responses to emotionally stressful situations. The tandem of affective education and relaxation training free the child from the initial distress of a stressful situation and allow them to process information in a more reasoned and less dysfunctional manner.

Cognitive restructuring consists of strategies aimed at altering cognitive distortions or teaching skills to overcome cognitive deficiencies. The child practices identifying and testing the legitimacy of their own self-talk, or automatic thoughts, and changing self-talk to see the behavioral consequences. Attribution retraining is similar in that it teaches the child to be aware of the nature of their explanations for positive and negative outcomes. Reward and reinforcement can be used as either a method of motivating behavioral activity and teaching cognitive flexibility (as in children with internalizing disorders) or as a way of suppressing behavioral activity (as in children with externalizing disorders).

Modeling and role-playing are two techniques in which a therapist demonstrates to a child how to analyze and evaluate affect, cognition, and behavior. In coping modeling, the therapist describes a situation and relates how they might react in such a situation. The therapist would verbally proceed through the steps of identifying the problem, gathering and isolating key pieces of information, devising alternative actions, and thinking through and acting out conclusions. By modeling the entire scenario, the therapist illustrates to the child that effective problem solving involves a step-by-step process. Role plays enable the child and therapist to learn to be prepared for problems before they arise, and provide practice in the skills learned in sessions before attempting them in real-life situations.

Generally, CBT protocols have two portions. The first portion of the treatment plan focuses on educating the child in the above skills so that the child will be prepared ahead of time when difficult or challenging situations arise. The second half of the treatment plan generally is used to provide the child with opportunities to practice the skills which he or she has just learned. This is accomplished by having the clinician set up a number of manageable scenarios in the real world and allowing the child to confront the ensuing problems on their own.

Not all of the strategies are used identically in every case. Rather, the therapist uses judgment in adjusting strategies to be best suited to the child's developmental sophistication, presenting problem, and extra-therapy environment. Assuming that all children or all therapists are the same, or that all outcomes will be of a uniform type, would be to fall prey to the "uniformity myths" (Kiesler, 1966).

When designing a treatment strategy it is important that the plan includes training for generalization across skills and settings (Kendall, 1989). That is, the therapist must be mindful that mastery of a skill within the confines of the therapy room does not necessarily translate into successful application in the child's natural environment. A child who has demonstrated competence in giving a speech in front of a group may still have a fear of attending a party. To facilitate generalization, therapy can focus on "training diversely" (i.e., exposing the child to a variety of scenarios) so that the child will not be caught off-guard by novel situations (Stokes & Baer, 1977). The therapist can be mindful of current functional contingencies in the child's natural environment: for instance, if a child has difficulties asserting themself, the therapist, under the assumption that most teachers would provide positively reinforcing experiences, might set up an experience where the child must ask the teacher for help. The goal of generalization is to make the child aware that the skills learned in session are not unique to any individual setting, but useful across settings.

5.06.5.1 Proposed Mechanisms of Change

As mentioned previously, childhood disorders can be broken down into internalizing and externalizing varieties, each with cognitive dysfunctions (e.g., deficiencies, distortions). Children who exhibit externalizing behavior tend to display impulsivity and deficits in thinking whereas children with internalizing disorders tend to show overcontrol and distortions in thought. Because of these differences, it is reasonable that the strategies used in treatment accommodate to the specific cognitive needs of the child. Likewise, our notions regarding the active mechanisms involved in change reflect our conceptualization of the client within these categories.

Children with ADHD, for example, characteristically exhibit heightened motor activity, impulsivity, deficits in attention, and deficits in rule-governed behavior (Barkley, 1990). Treatment for children with ADHD tends to involve response-cost contingencies within a context of social reward to reduce unwanted behavior and increase the rates of desirable behavior. The primary goal of the CB therapist is to help the impulsive child internalize steps that can guide effective problem solving: identifying the problem, reflecting on solutions, making a decision, and taking action (Kendall, Kortlander, Chansky, & Brady, 1992). Therapy operates on the idea that the child is deficient in a particular cognitive skill (e.g., rule-governed and linguistically guided behavior) and that training can provide the necessary information to shape more positive behavioral trajectories.

On the other hand, modifying existing cognitive distortions is a primary component and proposed instigator of change in children with internalizing disorders, such as anxiety. Clinically anxious children may have excessive concerns about separation from their caregivers, as in separation anxiety, or they may excessively fear evaluation or rejection by their peers, as in social phobia. Such anxiety manifests in physiological (e.g., excess perspiration, or blushing), behavioral (e.g., avoidance, tremors), and cognitive (e.g., self-doubt, misperceptions of environmental demands) symptoms and can lead to severe interference in the child's daily activities. These distorted misperceptions of the demands in the environment, along with distorted underestimates of personal ability, exacerbate unwanted anxious arousal. The strategies used in treating anxious children may resemble those used with externalizing populations (e.g., social problem solving, modeling, positive reinforcement) but, with internalizing children, the process is focused on altering dysfunctional thinking (distortions) and changing related maladaptive behavior. Thus, the anxious child is taught to evaluate the legitimacy of anxiety-producing automatic thoughts ("If I go to the party, I'll just embarrass myself") and to reconsider less disturbing ways of looking at the world. The therapist helps the child develop a coping template that involves correcting misinterpretations, and viewing events and the environment through an alternative lens based on coping.

In aggression, both cognitive deficiencies and distortions may be present (Kendall, Ronan, & Epps, 1991). This combination requires multifaceted treatment. Also, with comorbidity being the rule rather than the exception, therapists must often combine strategies. In such cases, therapy has the dual function of filling in the gaps where information was needed and aiding the child in constructing a coping template which promotes alternative, healthy ways of thinking. Both components are viewed as essential if the child is to successfully overcome the challenges of adjustment.

In returning to the client as a social scientist, we can conceptualize CBT as any therapy which involves teaching the client strategies that will help them to: (i) formulate specific predictions relevant to their dysfunctional beliefs; (ii) design and carry out extratherapy experiments that bear on these predictions; and (iii) re-evaluate their original hypotheses with respect to the data that are obtained. This formulation might also include the act of supplying a client with any information that would facilitate the process of hypothesis testing (e.g. teaching the problem-solving steps to an impulsive child).

The reader is reminded that the CBT model does not propose to provide a cure for childhood disorders. Rather, it imparts, and promotes through practice, the ingredients essential for successful coping. The ultimate goal of CBT is not to eradicate the disorder, but to provide the child with multiple lines of defense so that they will have a sense of competence and control when confronted with stress and emotional arousal.

5.06.6 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

With a rapidly growing proportion of researchers studying CBT (Kazdin et al., 1990; Klesges et al., 1982) and an expanding list of successful applications of CB to childhood disorders (see Kendall, 1993), there is reason to suggest that CBT will continue to undergo expansion. Attention to some key issues will assure the field's continued development. These issues include: (i) a growing need for development and study of standardized treatment manuals; (ii) an expanding emphasis on research aimed at understanding the active components of CBT, as well as understanding the underlying mechanisms maintaining childhood disorders; (iii) an increasing concern for the transportability of interventions from the research clinic to real-world service clinics; and (iv) a need for further study in the comparative and facilitative effects of pharmacological treatments.

Several randomized clinical trials evaluating CBT attest to its efficacy with some childhood disorders (e.g., anxiety, depression). More work needs to be done with disorders that have not been fully studied with CB protocols. The development and evaluation of manualized treatments, and the study of the role of flexibility in the use of manuals, deserves emphasis.

Demonstrating the efficacy of an intervention is a first step. Understanding the underlying behavioral and cognitive processes causing and maintaining a disorder as well as understanding the specific effects of therapeutic techniques is the second step. Which components are necessary? Which are helpful? Which are harmful? More clinical trials with active control groups (e.g., relationship therapy) may help examine the role of nonspecific factors of treatment. Such studies would allow us to begin to separate gains unique to the intervention from gains resulting from more general therapeutic factors.

Research is needed to determine the degree to which research clinic procedures can be transported to service clinics (Hoagwood et al., 1995; Kendall & Southam-Gerow, 1995). Are there client or therapist factors that contribute to differential outcomes? How are treatment-outcome studies received? Practicing clinicians may incorporate findings into their daily work, but research journals are not the practitioner's preferred method of receiving knowledge (Beutler, Williams, Wakefield, & Entwhistle, 1995). The transportability of research clinic interventions would benefit from progress on two fronts: first, research interventions should be tested in service clinics and any moderators of treatment should be examined; second, there needs to be more collaboration among research clinicians and service practitioners. Collaboration may help inform scientists on which areas of research are most relevant to practitioners and in what ways transportability could be improved. Such collaboration may help practitioners appreciate the value of data-supported interventions and inform practitioners how best to utilize findings from research clinics.

The recent pharmacological trend in treating mental health problems dictates that professionals within the field keep abreast regarding the usage and efficacy of pharmacological agents in treatment. Medications used in adult treatment are sometimes prescribed to children presenting similar symptoms but without the benefit of controlled, experimental evaluations. For example, in the treatment of depressed children and adolescents, various tricyclic antidepressants (TCA) (e.g., imipramine, desipramine, amitriptyline, and nortriptyline) and serotonin-specific reuptake inhibitors (SSRIs; e.g., fluoxetine, sertraline, paroxetine, and fluvoxamine) are commonly prescribed. However, despite what appears to be widespread general use of TCAs and SSRIs with children and adolescents, neither category of medication has consistently been demonstrated to be superior to a placebo (Jain et al., 1992; Simeon et al., 1990; Sommers-Flanagan & Sommers-Flanagan, 1996). Future research with pharmacological treatments with children needs to be directed toward determining the efficacy of the pharmacological agent alone or in combination with psychological interventions. Research addressing the comparative and facilitative effects of pharmacotherapy with youth is also required.

Much work remains in the application of CBT to childhood disorders. Applications of CBT for some disorders need to be developed, whereas in other disorders (e.g., anxiety, depression) research might focus on underlying cognitive processes and mechanisms of change. What aspects of cognitive processes and self-talk differentiate depressive from anxious self-talk? Which behaviors and cognitive processes show the greatest correlation with therapeutic change? Further research can also be done to enhance the transportability of CB interventions to practicing clinics. What are the factors that moderate outcome outside the research clinic's walls? What are the comparative and facilitative benefits of pharmacological treatments used in combination with CBT with children? Although CBT with youth is itself a relatively young approach, it has matured and continues to enjoy much activity and growth.

5.06.7 REFERENCES

Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: A factor analytic study. *Psycological Monographs*, *80* (whole No. 615).

Bandura, A. (1969). Social-learning theory of identification processes. In D. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 213–262). Chicago: Rand McNally.

Barkley, R. (1990). *Attention-deficit hyperactivity disorder. A handbook for diagnosis and treatment*. New York: Guilford Press.

Barrett, P., Dadds, M., & Rapee, R. (1996). Family treatment of child anxiety: A controlled trial. *Journal of Consulting and Clinical Psychology, 64,* 333–342.

Baum, J. G., Clark, H. B., McCarthy, W., Sandler, J., & Carpenter, R. (1986). An analysis of the acquisition and generalization of social skills in troubled youths: Combining social skills training, cognitive self-talk, and relaxation procedures. *Child and Family Behavior Therapy, 8,* 1–27.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford.

Beutler, L. E., Williams, R. E., Wakefield, P. J., & Entwhistle, S. R. (1995). Bridging scientist and practitioner perspectives in clinical psychology. *American Psychologist, 50,* 984–994.

Callahan, S. A., Panichelli-Mindel, S., & Kendall, P. C. (1996). DSM-IV and internalizing disorders: Modifications, limitations, and utility. *School Psychology Review, 25,* 298–308.

Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin, 98,* 388–400.

Craighead, W. E., Brownell, K. D., & Horan, J. J. (1981). Behavioral interventions for weight restriction and smoking cessation. In W. E. Craighead, A. E. Kazdin, & M. J. Mahoney (Eds.), *Behavior modification: Principles, issues, and applications* (2nd ed., pp. 288–312). Boston: Houghton Mifflin.

Craighead, W. E., Wilcoxon-Craighead, L. W., & Meyers, A. W. (1978). New directions in behavior modification with children. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification* (Vol. 6, pp. 159–201). New York: Academic Press.

Dodge, K. A. (1980). Social cognition and children's aggressive behavior. *Child Development, 51,* 162–170.

Dodge, K. A., & Frame, C. L. (1982). Social cognitive biases and deficits in aggressive boys. *Child Development, 53,* 620–635.

Durlak, J. A., Furhman, T., Lampman, C. (1991). Effectiveness of cognitive-behavior therapy for maladapting children: A meta-analysis. *Psychological Bulletin, 110,* 204–214.

D'Zurilla, T. (1986). *Problem-solving therapy: A social competence approach to clinical intervention*. New York: Springer.

D'Zurilla, T., & Goldfried, M. (1971). Problem solving and behavior modification. *Journal of Abnormal Psychology, 78,* 107–126.

Ellis, A. (1962). *Reason and emotion in psychotherapy*. New York: Lyle Stuart.

Epstein, S., & Erskine, N. (1983). The development of personal theories of reality. In D. Magnusson & V. Allens (Eds.), *Human development: An interactional perspective*. New York: Academic Press.

Evans, I. M., & Nelson, R. O. (1986). Assessment of children. In A. R. Ciminero, K. S. Calhoun, & H. E. Adams (Eds.), *Handbook of behavioral assessment* (2nd ed., pp. 601–630). New York: Wiley.

Finch, A. J., & Montgomery, L. E. (1973). Reflection-impulsivity and information seeking in emotionally disturbed children. *Journal of Abnormal Child Psychology, 1,* 358–362.

Guidano, V. F., & Liotti, G. (1983). Cognition in processes and emotional disorders. New York: Guilford Press.

Harris, S. L., & Ferrari, M. (1983). Developmental factors in child behavior therapy. *Behavior Therapy, 14,* 54–72.

Harvey, J. H., Weber, A. L., & Orbach, T. L. (1990). *Interpersonal accounts: A social psychological perspective*. Oxford, UK: Blackwell.

Hoagwood, K., Hibbs, E., Brent, D., & Jensen, P. (1995). Introduction to the special section: Efficacy and effectiveness in studies of child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63,* 683–687.

Hollon, S., & Beck, A. T. (1979). Cognitive therapy of depression. In P. C. Kendall and S. Hollon (Eds.), *Cognitive-behavioral interventions: Theory, research, and procedures* (pp. 153–204). New York: Academic Press.

Howard, B., & Kendall, P. C. (1996). Cognitive-behavioral family therapy for anxiety-disordered children: A multiple-baseline evaluation. *Cognitive Therapy and Research, 20,* 423–444.

Ingram, R. E., & Kendall, P. (1986). Cognitive clinical psychology: Implications of an information processing perspective. In R. Ingram (Ed.), *Information processing approaches to clinical psychology* (pp. 3–22) New York: Academic Press.

Ingram, R. E., Kendall, P. C., & Chen, A. H. (1991). Cognitive-behavioral interventions. In C. R. Snyder & D. R. Forsyth (Eds.), *Handbook of social and clinical psychology: The health perspective* (pp. 509–522). New York: Pergamon.

Jain, U., Birmaher, B., Garcia, M., et al. (1992). Fluoxetine in children and adolescents with mood disorders: A chart review of efficacy and adverse effects. *Journal of Child and Adolescent Psychopharmacology, 2,* 259–265.

Jones, M. C. (1924a). The elimination of children's fears. *Journal of Experimental Psychology, 7,* 382–390.

Jones, M. C. (1924b). A laboratory study of fear: The case of Peter. *Journal of Genetic Psychology, 31,* 308–315.

Kahnemann, D., Slovic, P., & Tversky, A. (1982). *Judgement under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.

Kanfer, F. H., & Karoly, P. (1972). Self-control: A behavioristic excursion into the lion's den. *Behavior Therapy, 3,* 398–416.

Kanfer, F., & Phillips, J. (1970). *Learning foundations of behavior therapy*. New York: Wiley.

Kazdin, A. E. (1987). *Conduct disorder in childhood and adolescence*. Beverly Hills, CA: Sage.

Kazdin, A. E. (1988). *Child psychotherapy: Developing and identifying effective treatments*. Elmsford, NY: Pergamon.

Kazdin, A. E., Bass, D., Ayers, W. A., & Rodgers, A. (1990). Empirical and clinical focus of child and adolescent psychotherapy research. *Journal of Consulting and Clinical Psychology, 58,* 729–740.

Kazdin, A. E., Bass, D., Siegel, T., & Thomas, C. (1989). Cognitive-behavioral therapy and relationship therapy in the treatment of children referred for antisocial behavior. *Journal of Consulting and Clinical Psychology, 57,* 522–535.

Kazdin, A. E., Siegel, T., & Bass, D. (1992). Cognitive problem-solving skills training and parent management training in the treatment of antisocial behavior in children. *Journal of Consulting and Clinical Psychology, 60,* 733–747.

Kendall, P. C. (1977). On the efficacious use of verbal self-instructions with children. *Cognitive Therapy and Research, 1,* 331–341.

Kendall, P. C. (1984). Social cognition and problem-solving: A developmental and child-clinical interface. In B. Gholson & T. Rosenthal (Eds.), *Applications of cognitive developmental theory* (pp. 115–148). New York: Academic Press.

Kendall, P. C. (1985). Toward a cognitive-behavioral model of child psychopathology and a critique of related interventions. *Journal of Abnormal Child Psychology, 13,* 357–372.

Kendall, P. C. (1989). The generalization and maintenance of behavior change: Comments, considerations, and the "no-cure" criticism. *Behavior Therapy, 20,* 357–364.

Kendall, P. C. (1991). Guiding theory for treating children and adolescents. In P. C. Kendall (Ed.), *Child and adolescent therapy: Cognitive-behavioral procedures* (pp. 3–24). New York: Guilford Press.

Kendall, P. C. (1993). Cognitive-behavioral therapies with youth: Guiding theory, current status, and emerging developments. *Journal of Consulting and Clincial Psychology, 61,* 235–247.

Kendall, P. C. (1994). Treating anxiety disorders in children: Results of a randomized clinical trial. *Journal of Consulting and Clinical Psychology, 62,* 100–110.

Kendall, P. C., & Braswell, L. (1982). Cognitive-behavioral self-control therapy for children: A components analysis. *Journal of Consulting and Clinical Psychology, 50,* 672–690.

Kendall, P. C., Flannery-Schroeder, E., Panichelli-Mindel, S. M., Southam-Gerow, M. A., Henin, A., & Warman, M. (in press). Therapy for anxiety-disordered youth: A randomized clinical trial and the effects of comorbidity and treatment segments. *Journal of Consulting and Clinical Psychology*.

Kendall, P. C., & Grove, W. M. (1988). Normative comparisons in therapy outcome. *Behavioral Assessment, 10,* 147–158.

Kendall, P. C., & Ingram, R. E. (1987). The future of the cognitive assessment of anxiety: Let's get specific. In L. Michelson & M. Ascher (Eds.), *Anxiety and stress disorders: Cognitive-behavioral assessment and treatment* (pp. 89–104). New York: Guilford Press.

Kendall, P. C., & Ingram, R. E. (1989). Cognitive-behavioral perspectives: Theory and research on depression and anxiety. In P. C. Kendall & D. Watson (Eds.), *Anxiety and depression: Distinctive and overlapping features*. New York: Academic Press.

Kendall, P. C., Kane, M., Howard, B., & Siqueland, L. (1990). *Cognitive-behavioral therapy for anxious children: Treatment manual*. Ardmore, PA: Workbook.

Kendall, P. C., Kortlander, E., Chansky, T. E., & Brady, E. U. (1992). Comorbidity of anxiety and depression in youth: Treatment implications. *Journal of Consulting and Clinical Psychology, 60,* 869–880.

Kendall, P. C., & MacDonald, J. P. (1993). Cognition in the psychopathology of youth and implications for treatment. In K. S. Dobson and P. C. Kendall (Eds.), *Psychopathology and cognition* (pp. 387–430). San Diego, CA: Academic Press.

Kendall, P. C., Ronan, K. R., & Epps, J. (1991). Aggression in children/adolescents: Cognitive-behavioral treatment perspective. In D. J. Pepler & K. H. Rubin (Eds.), *The development and treatment of childhood* *aggression* (pp. 341–360). Hillsdale, NJ: Erlbaum.

Kendall, P. C., & Southam-Gerow, M. A. (1995). Issues in the transportability of treatment. The case of anxiety disorders in youths. *Journal of Consulting and Clinical Psychology, 63,* 702–708.

Kendall, P. C., & Southam-Gerow, M. A. (1996) Long-term follow-up of a cognitive-behavioral therapy for anxiety-disordered youth. *Journal of Consulting and Clinical Psychology, 64,* 724–730.

Kendall, P. C., Vitousek, K. B., & Kane, M. (1991). Thought and action in psychotherapy: Cognitive-behavioral approaches. In M. Hersen, A. Kazdin, & A. Bellack (Eds.), *Clinical psychology handbook* (pp. 596–626). New York: Pergamon.

Kendall, P. C., & Williams, C. L. (1981). Behavioral and cognitive behavioral approaches to outpatient treatment with children. In W. E. Craighead, A. E. Kazdin, & M. J. Mahoney (Eds.), *Behavior modification: Principles, issues, and applications* (2nd ed., pp. 402–417). Boston: Houghton Mifflin.

Kettlewell, P. W., & Kausch, D. F. (1983). The generalization of the effects of a cognitive-behavioral treatment program for aggressive children. *Journal of Abnormal Child Psychology, 11,* 101–114.

Kiesler, D. J. (1966). Some myths of psychotherapy research and the search for a paradigm. *Psychological Bulletin, 65,* 110–136.

Klesges, R. C., Sanchez, V. C., & Stanton, A. L., (1982). Obtaining employment in academia: The hiring process and characteristics of successful applicants. *Professional Psychology, 13,* 577–586.

Levitt, E. (1957). The results of psychotherapy with children: An evaluation. *Journal of Consulting Psychology, 21,* 189–196.

Levitt, E. (1963). Psychotherapy with children: A further evaluation. *Behaviour Research and Therapy, 1,* 45–51.

Lewinsohn, P. M., Clarke, G. N., Hops, H., & Andrews, J. (1990). Cognitive-behavioral treatment for depressed adolescents. *Behavior Therapy, 21,* 385–401.

Lochman, J. E. (1987). Self- and peer perceptions and attributional biases of aggressive and nonaggressive boys in dyadic interactions. *Journal of Consulting and Clinical Psychology, 55,* 404–410.

Lochman, J. E. (1992). Cognitive-behavioral interventions with aggressive boys: Three-year follow-up and preventive effects. *Journal of Consulting and Clinical Psychology, 60,* 426–432.

Lochman, J. E., Burch, P. R., Curry, J. F., & Lampron, L. B. (1984). Treatment and generalization effects of cognitive-behavioral and goal-setting interventions with aggressive boys. *Journal of Consulting and Clinical Psychology, 52,* 915–916.

Lochman, J. E., White, K. J., & Wayland, K. K. (1991). Cognitive-behavioral assessment and treatment with aggressive children. In P. C. Kendall, (Ed.), *Child and adolescent therapy: Cognitive-behavioral procedures* (pp. 25–65). New York: Guilford Press.

Luria, A. R. (1976). *Cognitive development: Its cultural and social foundations.* Cambridge, MA: Harvard University Press.

Mahoney, M. J. (1974). *Cognition and behavior modification.* Cambridge, MA: Ballinger.

Mahoney, M. J., & Lyddon, W. J. (1988). Recent developments in cognitive approaches to counseling and psychotherapy. *The Counseling Psychologist, 16,* 190–234.

McNamee, S., & Gergen, K. J. (Eds.) (1992). *Therapy as social construction.* London: Sage.

Meichenbaum, D. (1977). *Cognitive-behavior modification: An integrative approach.* New York: Plenum.

Meichenbaum, D. (1985). *Stress inoculation training.* Elmsford, NY: Pergamon.

Meichenbaum, D. (1995). Cognitive-behavioral therapy in historical perspective. In B. Bongar & L. E. Beutler (Eds.), *Comprehensive textbook of psychotherapy* (pp. 140–158). New York: Oxford University Press.

Meichenbaum, D., & Fitzpatrick, D. (1993). A constructivist narrative perspective of stress and coping: Stress inoculation applications. In L. Goldberger & S. Breznitz (Eds.), *Handbook of stress* (pp. 706–723). New York: Free Press.

Meichenbaum. D. H., & Goodman, J. (1971). Training impulsive children to talk to themselves: A means of developing self-control. *Journal of Abnormal Psychology, 77,* 115–126.

Meyers, A. W., & Craighead, W. E. (1984). Cognitive behavior therapy with children: A historical, conceptual, and organizational overview. In A. W. Meyers & W. E. Craighead (Eds.), *Cognitive behavior therapy with children* (pp. 1–17). New York: Plenum.

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review, 80,* 221–234.

Mowrer, O. H., & Mowrer, W. M. (1938). Enuresis—A method for its study and treatment. *American Journal of Orthopsychiatry, 8,* 730–738.

Neimeyer, R. A. (1993). An appraisal of constructivist psychotherapies. *Journal of Consulting and Clinical Psychology, 61,* 221–234.

Neimeyer, R. A., & Feixas, G. (1990). Constructivist contributions to psychotherapy integration. *Journal of Integrative and Eclectic Psychotherapy, 9,* 4–20.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgement.* Englewood Cliffs, NJ: Prentice-Hall.

Ollendick, T. H., & Hersen, M. (1993). Child and adolescent behavioral assessment. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of child and adolescent assessment* (pp. 3–14). Boston: Allyn & Bacon.

Piaseki, J., & Hollon, S. D. (1987). Cognitive therapy for depression: Unexplicated schemata and scripts. In N.S. Jacobson (Ed.), *Psychotherapists in clinical practice: Cognitive and behavioral perspectives* (pp. 121–152). New York: Guilford Press.

Reynolds, W. M., & Coats, K. I. (1986). A comparison of cognitive-behavioral therapy and relaxation training for the treatment of depression in adolescents. *Journal of Consulting and Clinical Psychology, 54,* 653–660.

Rogers, C. R. (1951). *Client-centered therapy.* Boston: Houghton Mifflin.

Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. M. More (Ed.), *Cognitive development and the acquisition of language* (pp. 111–144). New York: Academic Press.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104,* 192–233.

Rotter, J. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 85.*

Sheslow, D. V., Bondy, A. S., & Nelson, R. O. (1982). A comparison of graduated exposure, verbal coping skills and their combination in the treatment of children's fear of the dark. *Child and Family Behavior Therapy, 4,* 33–45.

Simeon, J. G., Dinicola, V. F., Ferguson, B. H., et al. (1990). Adolescent depression: A placebo-controlled fluoxetine study and follow-up. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 14,* 791–795.

Sommers-Flanagan, J., & Sommers-Flanagan, R. (1996). Efficacy of antidepressant medication with depressed youth: What psychologists should know. *Professional Psychology: Research and Practice, 27,* 145–153.

Spivack, G., & Shure, M. B. (1974). *Social adjustment of young children: A cognitive approach to solving real-life* *problems*. San Franscico: Jossey-Bass.

Stark, K. D., & Kendall, P. C. (1996). *Treating depressed children: Therapist manual for "ACTION."* Ardmore, PA: Workbook.

Stark, K., Kendall, P. C., McCarthy, M., Stafford, M., Barron, R., & Thomeer, M. (1996). *ACTION: A workbook for overcoming depression*. Ardmore, PA: Workbook.

Stark, K. D., Reynolds, W. M., & Kaslow, N. J. (1987). A comparison of the relative efficacy of self-control therapy and a behavioral problem-solving therapy for depression in children. *Journal of Abnormal Child Psychology, 15,* 91–113.

Stark, K. D., Rouse, L. W., & Livingston, R. (1991). Treatment of depression during childhood and adolescence: Cognitive-behavioral procedures for the individual and family. In P. C. Kendall (Ed.), *Child and adolescent therapy: Cognitive-behavioral procedures* (pp. 165–206). New York: Guilford Press.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10,* 349–367.

Strauss, C. C., Lease, C. A., Last, C. G., & Francis, G. (1988). Overanxious disorder: An examination of developmental differences. *Journal of Abnormal Child Psychology, 16,* 433–443.

Treadwell, K. R. H., & Kendall, P. C. (1996). Self-talk with anxiety disorders: States of mind, content specificity, and treatment outcome. *Journal of Consulting and Clinical Psychology, 64,* 941–950.

Tversky, A., & Kahnemann, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Watson, J. B. (1924). *Behaviorism*. Chicago: University of Chicago Press.

Watson, J. B. & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3,* 1–14.

Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63,* 688–701.

Weisz, J. R., & Weiss, B. (1989). Assessing the effects of clinic-based psychotherapy with children and adolescents. *Journal of Consulting and Clinical Psychology, 57,* 741–746.

Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology, 55,* 542–549.

Weisz, J. R., Weiss, B., & Donenberg, G. R. (1992). The lab versus the clinic: Effects of child and adolescent psychotherapy. *American Psychologist, 7,* 1578–1585.

White, M., & Epston, D. (1990). *Narrative means to therapeutic ends*. New York: Norton.

Witmer, L. (1907). Clinical Psychology. *Psychological Clinic, 1,* 1–9.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.07
Family Systems Therapy with Children and Adolescents

MICHAEL S. ROBBINS and JOSÉ SZAPOCZNIK
University of Miami, FL, USA
JAMES F. ALEXANDER
University of Utah, Salt Lake City, UT, USA
and
JAMIE MILLER
Miami Institute of Professional Psychology, FL, USA

5.07.1 INTRODUCTION 150

5.07.2 THEORETICAL UNDERPINNINGS OF FAMILY SYSTEMS THERAPY 151
- *5.07.2.1 The Family System* 151
  - *5.07.2.1.1 Wholeness: the whole is more than the sum of its parts* 151
  - *5.07.2.1.2 Boundaries: individuals, subsystems, and the environment* 153
  - *5.07.2.1.3 Homeostasis, feedback, and change* 154
  - *5.07.2.1.4 Complementarity* 155
  - *5.07.2.1.5 Reciprocal vs. linear causality* 156
  - *5.07.2.1.6 Process vs. content* 156

5.07.3 HISTORY OF FAMILY SYSTEMS THERAPY 156
- *5.07.3.1 The Family's Role in the Etiology of Schizophrenia* 157
  - *5.07.3.1.1 Bateson and the Palo Alto Group* 157
  - *5.07.3.1.2 Lidz: the role of the marital relationship* 157
  - *5.07.3.1.3 Wynne: the family unit* 158
  - *5.07.3.1.4 Bowen: multigenerational influences and the development of family systems therapy* 158
  - *5.07.3.1.5 About schizophrenia* 159
- *5.07.3.2 Other Major Theoretical Figures in Family Systems Therapy* 159

5.07.4 EFFICACY AND EFFECTIVENESS RESEARCH OF FAMILY SYSTEMS THERAPY 159
- *5.07.4.1 Mood Disorders* 160
  - *5.07.4.1.1 Depression* 160
  - *5.07.4.1.2 Youth with other mood related symptoms* 160
- *5.07.4.2 Psychosomatic Disorders* 160
- *5.07.4.3 Conduct Disorders* 164
  - *5.07.4.3.1 Delinquency* 164
  - *5.07.4.3.2 Substance abuse* 165

5.07.5 ASSESSMENT AND CLINICAL FORMULATION 167
- *5.07.5.1 Conducting Assessments in Family Systems Therapy* 168
  - *5.07.5.1.1 Family genogram* 168
  - *5.07.5.1.2 Assessment through direct observation of family interactions: enactments* 169
- *5.07.5.2 The Underlying Nature of Family Interactions* 169

5.07.5.2.1 Functional family therapy 169
5.07.5.2.2 Structural family therapy 170
5.07.5.3 Clinical Formulation in Family Systems Therapy 171
5.07.6 MEDIATORS OF CHANGE: MECHANISMS OF ACTION IN FAMILY THERAPY WITH ACTING-OUT YOUTH 172
5.07.6.1 The Critical Role of the Therapeutic Alliance in Family Therapy 172
5.07.6.1.1 Building an alliance with behavior problem youth: specific recommendations 174
5.07.6.2 Reducing Within-family Negativity 174
5.07.6.2.1 Examining the impact of therapist interventions on within family negativity 175
5.07.6.3 Changing Parent–Adolescent Interactions 175
5.07.6.3.1 Changes in family interactions in successful and unsuccessful outcome cases 176
5.07.6.3.2 Family members' resistance to change 176
5.07.6.3.3 The relationship between changes in parenting behaviors and adolescent behavioral outcomes 177
5.07.7 SUMMARY AND FUTURE DIRECTIONS 177
5.07.8 REFERENCES 178

5.07.1 INTRODUCTION

Family systems therapy has emerged as a viable treatment for certain child and adolescent problems. In fact, for some clinical populations of youth, family systems therapy has been acknowledged as the "treatment of choice" (Gordon, Arbruthnot, Gustafson, & McGreen, 1988; Kazdin, 1993, 1994a; Liddle & Dakof, 1995a, 1995b). Despite its rapid growth in clinical and research settings, some continue to view family systems therapy as merely an adjunct to traditional individual modalities.

Family systems therapy, however, represents a "whole new way of conceptualizing human problems, of understanding behavior, the development of symptoms, and their resolution" (Sluzki, 1978, p. 366) in which the transformation of family interactions is its defining feature. Because interactions are viewed as pivotal in the etiology and maintenance of psychopathology, interactions are the ultimate target of intervention in family systems therapy.

Interactions are defined as the complex and interdependent patterns of behaviors and responses within the family. A fundamental assumption in family systems theory is that an individual's behavior is quite different from what it would be if it were possible for that individual to act in isolation. That is, the behaviors of individual family members are linked in an interdependent fashion, and each family member is viewed as being responsible for contributing to, maintaining, or changing the system's interactions (Szapocznik & Kurtines, 1989). The interdependent or linked behaviors among individuals tend to repeat, and ultimately become patterns of interaction that characterize the family system. It is these repetitive patterns of interaction that are the focus of family systems therapy.

While the family is at the core of the intervention strategy, family systems therapy appreciates the full range of systemic influences that affect the individual. For example, family systems therapy acknowledges the important influence of peers (Boivin & Begin, 1989; Dishion, French, & Patterson, 1995; Parker & Asher, 1987), school (Rutter, 1983), and the neighborhood (Brooks-Gunn, Duncan, Klebanov, & Sealand, 1993) on children and their families. A fuller appreciation of the range of systems that influence children and their families also includes biological processes (i.e., genetics) as well as individual psychological processes (i.e., cognitions). However, although the interface between the family and its context is acknowledged (Alexander, Barton, Waldron, & Mas, 1983; Szapocznik & Kurtines, 1993), it is the interactions that occur within the family that are the primary focus of intervention.

It is important to note that there is considerable heterogeneity within the field of family systems therapy in how different models of systemic intervention conceptualize and treat psychopathology. Some models focus relatively more on intrapersonal (i.e., psychodynamic, cognitive) processes, while others focus more on interpersonal (i.e., behavioral, multisystemic/contextual) processes. Such differential foci, however, tend to be a matter of emphasis rather than epistemology, and most models contain both intrapersonal as well as interpersonal elements. What cuts across family systems models, however, is a focus on patterns of interactions within the family. Thus, in this chapter, it is our intention to highlight the interactive and interdependent nature of family interactions that is ubiquitous to family systems intervention, rather than to draw attention to differences (often subtle) between the various systemic interventions.

In this chapter we present an overview of the development and current state of the science in family systems therapy with children and adolescents. By family systems therapy, we refer to clinical models that postulate underlying mechanisms which are based on the fundamental tenets of systems theory. To be included in this review, the intervention model had to contain a specific focus on processes (i.e., patterns of interaction) of the family system. By family, we refer to nuclear, blended, or other types of family arrangement. Intervention models more strongly rooted in behavioral theory are not reviewed in this chapter (for a review of these see Estrada & Pinsof, 1995).

5.07.2 THEORETICAL UNDERPINNINGS OF FAMILY SYSTEMS THERAPY

Considerable research indicates that social/contextual factors profoundly influence a child's developmental trajectory (Bronfenbrenner, 1995). Within the social context, the family has undoubtedly the most fundamental influence, having been referred to as the "bedrock" of child development (Bronfenbrenner, 1979; 1986; Szapocznik & Coatsworth, submitted for publication). The family system serves a critically important socialization influence for children and adolescents (Rutter & Cox, 1985); consequently, problems within the family system can dramatically influence youth. In fact, research has consistently demonstrated that problems within the family system are associated with a variety of psychological and behavioral problems in children and adolescents (Forehand, 1993; Kazdin & Kolko, 1986; Rutter & Garmezy, 1983).

So, what is the family system? How does it influence the child's development, including the development of psychological and behavioral problems? Is the family system amenable to change? And, if so, are these changes associated with improvements in the child's behavior? To answer these questions, it is important to understand some of the basic processes that underly the family system. In this section, some of the basic tenets of family systems theory are presented.

5.07.2.1 The Family System

The fundamental epistemology underlying family systems therapy is that of general systems theory (von Bertalanffy, 1968). In fact, the core assumption in family systems therapy is that the family operates according to basic rules and processes that govern all types of systems. From this perspective, a system is defined as a:

> circumscribed complex of relatively bounded phenomena, which, within these bounds, retains a relatively stationary pattern of structure in space or of sequential configuration in time in spite of a high degree of variability in the details of distribution and interrelations among its constituent units of lower order. (The system) responds to alterations of the environment by an adaptive redirection of its component processes in such a manner as to counter the external change in the direction of optimum preservation of its systemic integrity. (Weiss, 1969, pp. 11–12)

Chaos theory (Briggs & Peat, 1989; Gleik, 1987) can be used to explain the apparent paradox that a high degree of variability in the details of the behaviors of individual family members, in the context of the family, emerge as relatively stationary patterns of interaction over time and space. At the core of family systems theory is the notion that the family is a dynamic, self-regulating system comprised of interdependent parts that interact with one another in a somewhat predictable manner. The family system is dynamic in that it is comprised of individuals that interact and mutually influence one another (and their environment) continuously. And, it is self-regulating in that these interactions strive to maintain the integrity of the system. Integrity or self-preservation, however, are not equivalent to adaptive functioning. By maintaining integrity, we refer to the system's ability to maintain its characteristic organization regardless of whether this organization is adaptive or maladaptive.

The most important systemic principle that dominates family systems theory is that of interdependence. Interdependence refers to the complex connections and mutual influence of individual parts within a system. In families, interactions between and among family members are the primary means through which interdependence is defined. Understanding interactions—and consequently, interdependence—requires an understanding of the basic principles of systems theory. The following section briefly reviews some of the systemic principles that are critically relevant for assessment and intervention in the family context. This review is not exhaustive. For a more thorough review of systems principles see Broderick (1993), Buckley (1968), Goldenberg and Goldenberg (1996), Hanson (1995) and Steinglass (1987).

5.07.2.1.1 Wholeness: the whole is more than the sum of its parts

Basic to family systems theory is the assumption that the family is a dynamic system that can only be understood by overtly recognizing the complex interplay among all of its constituent parts. One cannot break the system down to its molecular parts, and then put the parts back together again to understand how the family operates. That is, achieving some level of understanding about an individual's internal (e.g., psychological) dynamics or a single aspect of family functioning (e.g., marital conflict) is useful, but it does not necessarily help the therapist predict (or even understand) how the family comes together and interacts as a system. It is only by examining the unique way in which the parts interact and reciprocally influence each other that the family system can be understood. This philosophy is represented in the systemic assumption that "the whole is more than the sum of its parts."

This is not to say that the individual parts of the family system are not important. In fact, we support attempts at formally integrating intrapersonal factors into family systems treatment (Breunlin, Schwartz, & Kune-Karrer, 1992; Liddle, 1994), and believe that future family systems model building must continue to build on such attempts to take into account the individual characteristics of family members. What we are trying to convey here, however, is that molecular examinations of the family do not capture the true complexity of the family system. And that even though "there is no phenomenon in a living system that is not molecular ... there is none that is only molecular, either" (Weiss, 1969, pp. 10–11). In other words, it is useful to understand how an individual family member interprets/processes information and how this processing influences their behavior; however, to understand the mutual influence and interdependent nature of family interactions such information must then be incorporated at the level of the family. Thoughts, feelings, and individual behaviors are thus considered to contribute to the interpersonal context of the family, but are not considered to be the defining features of the family. In family systems therapy, assessment and intervention occurs at both the molecular and molar levels. The therapist attends to the internal realities of individual family members and specific aspects of family relations (e.g., marital conflict), but also examines the unique way in which these aspects come together and influence the adaptive and maladaptive behaviors of the youth.

The essence of the concept of wholeness exists in the reality that although the family is comprised of multiple individuals with different emotions, beliefs, and behavioral repertoires, the family system develops its own unique identity that is not fully explained by these individual characteristics. As suggested in the biological sciences, the patterned processes that occur within systems do not occur because of some "prearranged, absolutely stereotyped mosaic of singletracked component performances" (Weiss, 1969, p. 9). Rather, these repetitive processes occur as individual parts in the system submit to the organizing rules of the system. In the family system, for example, although individual family members are capable of exhibiting an infinite number of interactions, they tend to interact in a manner that "fits" or is "consistent with" the governing rules of the family system.

The concept of "wholeness" has critical implications for treatment. For example, while individual family members' values, beliefs, and emotions, as well as particular relationships within the family (e.g., sibling, parent, parent–child) are considered to be important, systemic treatment must involve assessment and intervention at the level of the family as a whole. The family therapist is thus interested in changing the family's unique pattern of interaction. For example, an intervention designed to promote healthy connectedness between a parent and an adolescent must also take into account the impact that such an increase in connectedness will have on other aspects of the parent–adolescent relationship as well as on other individuals in the family. As is evident in this example, family systems therapists often target individuals or specific relationships within the family, but they do so with an explicit understanding of the potential reverberations that such interventions will have throughout the family system. It should also be noted that these interventions are typically utilized as a means to an end, and are often implemented as one aspect of a larger treatment plan that addresses changes in complex patterns of interaction at the level of the family system. Thus, in the scenario presented above, improving parent–adolescent connectedness may serve as a springboard for improving the parents' monitoring of the adolescent's behaviors. Moreover, this improved parent–adolescent bond may result in the formation of a new leadership subsystem (comprised of parent and adolescent) in which the adolescent is supported in taking over some of the developmentally appropriate leadership responsibilities within the family.

It is important to note that the concept of wholeness explains current family interactions, but does not explain why the family adopts one particular pattern of interaction over another. In fact, most family systems theories focus more on present interactions than on their historical antecedents. Many factors are likely to contribute to the emergence of specific patterns of interaction within the family, including individuals' predispositions (e.g., biological, personality) and personal history (e.g., family of origin, interpersonal relationships) as well as major developmental milestones (e.g., marriage, birth of children) and contextual factors (e.g., deteriorating conditions in neighborhoods, parent's loss of job).

When individuals first come together to form a relationship, their unique individual qualities probably define the possible range of interactions that may occur. Although there are an infinite number of possibilities that may occur within this range, specific patterns emerge and become relatively stable. As noted, the emergence of one pattern over another is probably influenced by many factors. While the relationship is partly determined by the persons involved, it cannot be predicted from the individual predispositions of each of its participants. The individuals involved in the interaction define the range of possibilities that may occur, but the specific pattern that emerges cannot be predicted from knowledge of each individual's potential.

Initially specific patterns may become repeated because they are personally gratifying to one or both of the participants or because it serves an important function in the relationship. As a particular interaction is repeated, it becomes increasingly stable and predictable, and, over time, may persist even if it no longer serves the important function it once served.

5.07.2.1.2 Boundaries: individuals, subsystems, and the environment

The systemic concept of "boundary" is a prominent aspect of family systems therapy. Conventionally, boundaries have been seen as the metaphorical "walls" that separate people, denoting where one person or group of persons ends and where the next begins. However, family systems theorists (Szapocznik & Coatsworth, submitted for publication) have recognized that just as boundaries signify separateness, they also denote connectedness. For example, the parent–child boundary not only signifies the distinction between the parenting subsystem and the child subsystem, but it also signifies the precise point at which they come together to interact as parent and child. Take for example an analogy from biology. Initially a cell membrane was viewed as a "wall" that helped contain the contents of the cell, and that separated the inside of the cell from the outside of the cell. This view provided the membrane with a certain "concreteness." A more contemporary view of a cell membrane, however, reveals that the "wall" is constantly negotiating transfer of biological components from inside to outside and vice versa. Thus, the cell membrane represents the cell's systematic mechanism for regulating transactions with its environment. This dynamic view of "boundaries" is reflected in the family systemic focus on the reciprocal process of interactions within the family, and between the family and its environment (Bronfenbrenner, 1979, 1986; Magnusson, 1995; Szapocznik & Coatsworth, submitted for publication). In families, boundaries are established by rules, both those that are discussed as well as those that are never discussed, but which everyone understands. Four important boundaries in family systems therapy are presented below.

The *family–environment boundary* denotes the separation/connectedness between the family and the environment. The simplest way of differentiating (or defining the family–environment boundary) a family from the environment is by examining biological relationships, or by examining who lives together and shares daily household responsibilities (Greenwood, Hervis, Mitrani, Taylor, & Szapocznik, 1995). Beyond these important distinguishing qualities, the family system is also characterized by its relatively stationary and unique pattern of interactions.

The boundary between the family and the environment is more or less permeable. That is, much like the membrane of a cell, this boundary regulates communication between the family and environment, and represents the precise point at which they intersect and reciprocally influence each other. (Note: Although we do not focus on it in this section, information also flows to and from the environment to specific individuals within the family. That is, individuals within the family also interact directly with the environment.)

The family–environment boundary may be critically related to the family's functioning, and is frequently an indication of potential problems in the family's ability to function adaptively. For example, Alexander (1988) notes that in families in which there is incest, the boundary between the family and environment is often rigid, resulting in increased insularity of the family, and an inability of the family to receive and respond adaptively to "feedback" from the environment.

The internal *family subsystem boundary* separates/connects subsystems within the family. The family system divides itself into various alliances to distribute its many roles and functions. These alliances are referred to as subsystems. Subsystems are typically organized by age or function. For example, the parental subsystem consists of the parents or, in single-parent families, the parent and another adult figure or an older child. The child subsystem consists of the siblings in the family (typically grouped together by age or gender). Subsystems are often utilized as the vehicle for completing the various roles and functions of the family.

The most important subsystem boundary in family systems therapy is the parent–child subsystem boundary. This boundary regulates the extent to which parents (or a parent and an older child, extended family member, or other parental adult) are able to work together as a unit to carry out the family's rules and responsibilities. Implementing rules, however, is only one way in which the parental subsystem interacts with the child subsystem. The parental subsystem is also responsible for caring, nurturing, and providing guidance to the youth. It should also be noted that influences at the parental subsystem and child subystem are bidirectional. The child subsystem also influences the parents. Take the case of a behavior problem youth as an example. The behaviors of these youths are frequently very intense and disruptive, and often elicit controlling responses by the parent(s). In fact, over time parents of behavior problem youth are likely to utilize many different strategies to attempt to manage or cope with their child's disruptive behaviors. This bidirectional influence is not limited to maladaptive interactions; in fact, in every family (beginning with the birth of children), children and parents continuously influence the behaviors of one another.

The *interpersonal boundary* separates/connects individual family members. This boundary represents the amount of separation/connection in each of the dyadic relationships in the family. At one extreme, this interpersonal boundary can be extremely rigid or impermeable (e.g., nothing gets through the boundary). When boundaries are so impermeable, the emotional and psychological distance between people is great. One strength in relationships in which there is an impermeable boundary is that individuals are often very autonomous and independent. However, a weakness in these relationships is that individuals do not benefit from supportive relationships, and they are more likely to experience existential crises and feelings of loneliness.

At the other extreme, the boundary can be far too tenuous. When boundaries are so tenuous, the emotional and psychological closeness between people is great. Individuals in relationships with highly permeable boundaries benefit from the supportive ties they have with other individuals. It is not uncommon for individuals in these relationships, however, to have difficulty establishing their own unique identity separate from other family members.

It is important to note that for any individual, the boundary separating/connecting them to other family members is unique to each particular relationship. For example, a father and his daughter may have a relatively permeable boundary in which there is a consistent, unobstructed flow of information. However, in the same family, there may be a rigid boundary between father and son in which communication is virtually nonexistent. Thus, the interpersonal boundary between individual family members is uniquely defined for each relationship in the family. In this sense, boundaries are considered to be characteristics of relationships rather than trait-like characteristics of individuals. (It should be noted that the interpersonal boundary is not unique to family relationships. Any relationship involves the continuous flow of information between its participants.)

The *interpersonal–intrapersonal boundary* represents the point at which interactions influence an individual's internal (i.e., psychological) world, and the point at which these internal processes influence the interaction. In family systems theory, the boundary between family interactions (interpersonal) and intrapersonal functioning is emphasized because of the family's profound influence on an individual's emotional and behavioral adjustment. All interpersonal relationships include emotional and cognitive intrapersonal components. These intrapersonal components influence the interpretation of information as well as the individual's behavioral response. Thus, the interpersonal–intrapersonal boundary regulates the flow of information from interactions with family members into the intrapersonal realm of the individual, where it is processed, and then ultimately influences the behavioral reactions of individual family members involved in the interaction.

5.07.2.1.3 Homeostasis, feedback, and change

As mentioned above, every system has built in mechanisms for ensuring its survival. The struggle for the survival of the system involves balancing between maintaining stability within the system and adapting to meet changing internal and external demands. Thus, there are two ways a system can function to maintain its viability. The first involves the process by which the system continuously monitors the activities of the family, and provides corrective feedback (information flow) to family members when they behave in ways that are not consistent with the family's rules. Such feedback operates to restore the family's equilibrium. This process, homeostasis, is critical for maintaining consistency and stability within the family over time.

However, because families are always in flux (e.g., children and parents' development, changing environmental conditions), there must also be mechanisms for ensuring that the family system adjust to meet the changing needs of its members or its environmental context. Similar to the processes involved in maintaining homeostasis, the process of change requires continuous feedback about the consequences of interactions within and outside of the family. As information is fed back to the system, the family is able to change those interactions that may have become obsolete or are leading towards the dissolution of the family. This process of the family system changing to meet changing internal or external demands is called recalibration.

A common example of how the recalibration process works and how it often goes awry in families is evident in the changing demands that are encountered as children mature during adolescence. The "task oriented" leadership style that may have worked when children were young often does not work when issues of autonomy and independence surface during adolescence. Parents must adjust their parenting style to fit the needs of the developing adolescent. Many parents continue to be very authoritarian (Baumrind, 1991a, 1991b) despite the fact that it is not working and that their child now requires a more flexible, negotiating parental style.

As the above example suggests, flexibility is a critical component in the recalibration process. That is, families must have the flexibility to adapt to changing internal and external circumstances. Flexibility is a measure of the system's ability to change and reorganize to meet changing needs. Flexibility also refers to the system's ability to adjust to the various situations confronting its members on a daily basis. To the extent that the system rigidly adheres to one pattern of interaction, the more likely it is that problems will arise. Thus, in a sense, flexibility refers to the inherent capabilities of the system and its ability to adaptively utilize a repertoire of interactions. More flexible systems are able to engage in a broader range of behaviors and are able to alter their patterns of interaction to meet the unique challenges they face on a continuous basis. Adaptive family functioning requires a balance between homeostasis and flexibility. Too much emphasis on stability prevents necessary adaptability, but excessive flexibility leads to chaotic family functioning (Minuchin & Fishman, 1981).

5.07.2.1.4 Complementarity

Family systems theory suggests that behaviors in an interaction are complementary (Minuchin & Fishman, 1981; Szapocznik & Kurtines, 1989), like a hand and glove are complementary: they fit together perfectly well. Thus, in addition to being inextricably linked to one another, family members' behaviors occur together as perfect counterparts. The basic assumption is that family members' behaviors mutually influence one another; consequently, any symptomatic behavior must be connected to other behaviors within the family. The principle of complementarity flows logically from the concept of interdependence (i.e., the connectedness and mutual influence of persons within the system). Thus, if a youth is using drugs, the question becomes: What are the complementary behaviors; that is, the behaviors that allow or encourage the youth to behave in this way? In other words, what are the social interconnections between family members that are conducive to drug use?

In addition to the complementary relationships among family members, family systems therapists are also concerned with the complementary interactions at the interface ("boundary") between the family and its social environment. For example, Steinberg and colleagues have shown that an authoritative parenting style is more strongly associated with academic achievement than nonauthoritative parenting style (Steinberg, Mounts, Lamborn, & Dornbush, 1991; Steinberg, Darling, Fletcher, Brown, & Dornbush, 1995). Thus, authoritative parenting behaviors are associated with/complement the child's academic achievement. However, this relationship appears to be moderated by ethnicity and peer group membership. European and Hispanic-American adolescents benefit more from authoritative parenting than African- or Asian-American adolescents (Steinberg et al., 1991; see also Florsheim, Tolan, & Gorman-Smith, in press). In part, this appears to be due to the nature of the peer groups with whom they affiliate. While in general adolescents from authoritative homes tend to associate with competent peers, for some ethnic groups it appears that the choice of which peer group one can associate with is restricted due to the neighborhood in which the child lives. Steinberg, Dornbush, and Brown (1992) found that Asian-Americans tended to associate with peers who support academic achievement, while the opposite was true for African-American adolescents in inner city neighborhoods. What these results suggest is that the nature of the peer group constrains the relationship between parenting style and academic achievement. It is also possible that parents may have an impact on the peer group. Thus, to fully understand the cross domain complementary (parent–peer) links associated with problem behaviors it is essential to consider the full range of intrafamily, adolescent–peer, and family–peer interactions that directly and indirectly influence the child's behaviors.

5.07.2.1.5 Reciprocal vs. linear causality

Family systems theory assumes reciprocal as opposed to linear causal explanations. Linear causality is based on the traditional stimulus–response behavioral view that one event "A" causes some response "B." This type of reductionistic thinking fails to capture the interdependence of individuals in any system. While it may be appropriate to highlight microanalytic sequences within the family (e.g., parents allow their adolescent daughter to stay out after curfew to avoid a confrontation with her), limiting our focus on these sequences fails to capture the true complexity of family interactions.

Circular causality assumes that any behavior in an interaction is simultaneously influenced by and influential on other behaviors in the interaction. In other words, causality is bidirectional. Thus, in the example of parents allowing their daughter to stay out after curfew to avoid a confrontation, a circular view of this sequence would begin with a recognition that the parent's behavior (e.g., avoiding confrontation) not only influenced the daughter's immediate likelihood of not being angry, but it also influenced the probability of her staying out beyond curfew in the future. Such a parental response may be the result of previous interactions in which the daughter responded with intense anger to the parent's confrontation. As a result, the parents' avoidance of confrontation in the current interaction may also influence the likelihood that the daughter will respond with intense anger to future parental confrontations of the daughter's behavior.

As is evident from this description, circular causality is informed by combinations of sequences of behavior. In this example only one particular sequence was highlighted. From a systemic perspective, however, it is not possible to understand a sequence outside of the context in which it occurs. Thus, to fully understand the mutual influence of parents and the adolescent in this example, the family systems therapist must also address other factors within the family (i.e., the parent's may not agree on basic rules within the house; consequently, they are unable to present a unified front to their daughter) or outside the family (i.e., the daughter may be associating with friends that foster an attitude of parental disrespect).

5.07.2.1.6 Process vs. content

Process can refer to the behaviors which are involved in an interaction; what and how things happen; as well as to the messages which are communicated by the nature of interactions or by the style of communication, including all that is communicated nonverbally such as feelings, tone, and power.

Content refers to the specific and concrete facts used in the communication. Content includes such things as the reasons that family members offer for a given interaction. For example, when a depressed mother says that her child is an obstacle to her completing school, it is important to attend both to the content (the specific issue raised) as well as to the process (the child being identified as the sole source of all of the family's problems).

Process and content may sometimes be contradictory. For example, an adolescent might say "I want help for my depression," while her behavior is that she disappears from home at the time agreed for family therapy sessions.

An emphasis on process (as opposed to content) is a predominate theme across family systems therapy models. The emphasis is on identifying the nature of the interactions in the family (diagnosis) and in changing those interactions that are maladaptive. Content is not ignored, particularly content that is emotionally valenced. However, the focus of therapy is on changing patterns of interaction. The therapist who responds primarily to content and loses sight of the process, will be unable to implement the kinds of family changes required to alleviate the problems experienced by the child.

5.07.3 HISTORY OF FAMILY SYSTEMS THERAPY

Most reviews cite the period directly following World War II as the most influential time in the evolution of family systems therapy (Goldenberg & Goldenberg, 1983). A variety of factors, including difficulties in the reunification process, the "baby boom," and changing social mores about familial roles and responsibilities created a large need for the development of a new wave of intervention that directly addressed family factors (Goldenberg & Goldenberg, 1983; Nichols & Schwartz, 1995). As a result of the large personal and societal demands, psychosocial interventions, including family therapy, flourished.

There are many excellent reviews that trace the development of family systems therapy, including landmark reviews by Guerin (1976), Goldenberg and Goldenberg (1983) Kaslow (1980), Nichols (1984), and Nichols and Schwartz (1995). Rather than summarize the work of these authors, this section focuses on those developments that directly contributed to our current conceptions of interactions, the cornerstone of the family systems approach. This review is not intended to be exhaustive. Our intention here is to chronicle the evolution of some of the critical events that led to our current conceptualization of the role of family interactions in the etiology and treatment of child and adolescent psychological problems.

5.07.3.1 The Family's Role in the Etiology of Schizophrenia

The 1950s was a critical decade in the development of the family systems approach. Four independent research groups proved to be the most influential in developing systemic theories of family functioning. Each of these groups were developing family theories about the development of schizophrenia.

5.07.3.1.1 Bateson and the Palo Alto Group

Bateson and his colleagues, the Palo Alto Group, were interested in developing a communication theory, which applies basic cybernetic principles within the family system, to explain the etiology of schizophrenia. The combined work of the Palo Alto Group provided the underpinnings for many of the current systemic theories of family functioning and family therapy. In particular, their systemic conceptualization of homoestasis, identified patienthood, and process vs. content is recognized in contemporary family systemic treatment models.

According to Bateson and colleagues (1972), whenever the integrity of the system is at stake, its natural tendency is to maintain homeostasis. One way in which the family system restores equilibrium is through a process in which one person becomes the symptom bearer. For example, an adolescent becomes severely depressed and suicidal whenever there is marital conflict, forcing the parents to stop fighting by refocusing on their child's depression. This process by which the child refocuses the problematic experience/event onto themselves labeled "identified patienthood," represented a shift in focus away from a purely intrapsychic conceptualization of the etiology of psychopathology towards a more interactional understanding of the onset of psychological problems.

Perhaps the greatest contribution of the Palo Alto Group to family systems theory was their conceptualization of the etiology of schizophrenia from a communication theory perspective. Bateson, Jackson, Haley, & Weakland (1956) introduced the concept of the double-bind communication as evidence of the family's role in the etiology of schizophrenia. It was their assumption that schizophrenia is caused by confusing communications within the family. In examining family conditions that precipitated schizophrenic episodes, they discovered that communication patterns in families with schizophrenic patients were characterized by confusion and contradiction. In particular, they noted that these communications occurred in the context of an important relationship, they contained two or more related but contradictory messages that also contained a threat (or perception of threat), and that the schizophrenic patient was forced to respond to these contradictory messages. For example, Bateson et al. (1956, p. 259) present the case of a young schizophrenic patient who was visited by his mother in the hospital. When she arrived, he put his arm around her and she stiffened. Sensing her discomfort he immediately withdrew his arm, and she responded by asking, "Don't you love me anymore?" The patient then blushed, prompting his mother to respond, "Dear, you must not be so easily embarrassed and afraid of your feelings." Following his mother's visit, the patient relapsed.

The concept of double-bind communications has influenced the development of several modern day family systemic concepts. For example, the concept of double-bind was the earliest articulation of the importance of process. Although content was critical in determining whether or not contradictory messages were in fact present, the primary emphasis was on the interactive process that was viewed as conducive to the development of schizophrenic symptomatology. For example, to engender schizophrenic behaviors, double-bind communications had to be a repetitive and stable characteristic of family interactions.

5.07.3.1.2 Lidz: the role of the marital relationship

Lidz may have been the first scientist interested in examining the role of the family in the etiology of schizophrenia. Lidz was strongly rooted in psychoanalytical therapy, and his primary interest was in understanding the way in which parents influenced the internal processes of identification and incorporation in their offspring (c.f. Nichols, 1984). However, through his intensive study of the families of schizophrenic patients, he developed a family theory of the origins of schizophrenia that brought together psychoanalytic theories about the important influence of parental personality on child development, and modern family systems theory.

In 1949, Lidz challenged Fromm-Reichmann's (1948) belief that a cold, domineering, and guilt-inducing mother ("schizophrenogenic mother") was responsible for the etiology of schizophrenia. In particular, Lidz documented the important contributions of fathers and—more importantly—the marital relationship, to the development of schizophrenia (Lidz & Lidz, 1949; Lidz, Cornelison, Fleck, & Terry, 1957; Lidz, Parker, & Cornelison, 1956). This transition from the individual to dyad to triadic relationships is a remarkable contribution to family systems theory.

Lidz is most recognized for his articulation of two types of marital relationships which contribute to the onset of schizophrenia. The first, "marital schism," involves a chronic failure in the marital relationship. In marital schism, spouses are in constant conflict and frequently undermine each other, openly competing for the loyalty of their children, and behaving in ways which fail to complement the behaviors of their partner. Quite often one partner, typically the father, is emotionally or physically disengaged from the family.

The second pattern, "marital skew," is characterized by severe psychopathology in one of the spouses. Spouses assume complementary roles which are pathological. One is dominant while the other is submissive. The dominant partner is typically the one with the serious psychopathology. Thus, even though on the surface they may appear strong and capable, their ability to provide direction and leadership within the family is severely compromised. The weaker partner goes along with and even supports the pathological tendencies of the superior parent.

Lidz et al., (1957) provided a complex rationale about the way in which "marital schism" and "marital skew" led to schizophrenia in girls and boys. What is important here, however, is his articulation of the relationship between family processes and the development of serious psychopathology. His observations and research went beyond the prominent linear cause–effect theories of the time, and prompted the development of a new wave of research and clinical development.

5.07.3.1.3 Wynne: the family unit

Wynne's theories of the role of the family in the development of schizophrenia profoundly influenced current family systemic theories' conceptualization of boundaries and alliances. According to Wynne (Wynne, Rycoff, Day, & Hirsch, 1958), interactions in families with schizophrenic members were characterized by two processes: pseudomutuality and pseudohostility.

Pseudomutuality is defined as interactions that give the impression that the family is functioning adaptively, but hide the fact that there are serious problems in the family. Boundaries in these families are very blurred, and individual autonomy and outside relationships are discouraged. He describes the boundary between the family and its environment as a "rubber fence" that stretches to allow some contact, but that is essentially impermeable. Thus, outside forces rarely precipitate a restructuring of family responsibilities; and, even if conflict between the family and its environment does emerge, the family responds by becoming even more insular.

Pseudohostility is similar to pseudomutuality, but involves a dramatic expression of conflict in response to minor disruptions in family alliances. Since the family is preoccupied with maintaining togetherness, a dramatic amplification of a minor problem actually operates as a "feedback" mechanism for the family. Family members move in quickly to repair the minor split and restore the system's togetherness.

Wynne's theories were critical in articulating the importance of boundaries and feedback mechanisms in the family system. His explanation of how a disturbance in family boundaries influences the development of major psychopathology represents one of the earliest structural interpretations of family functioning.

5.07.3.1.4 Bowen: multigenerational influences and the development of family systems therapy

Similar to Lidz, Bowen focused on disturbances in parental interactions, and their relationship to schizophrenia. In particular, Bowen (1960) coined the term "emotional divorce" to point out the emotional distance between the parents of schizophrenics. He claimed that children were often the only area in which these parents shared any common interest or were able to collaborate. Perhaps his most well-known contribution to the study of the family's role in the development of schizophrenia is his "three generation hypothesis."

According to Bowen (1960), the parents of schizophrenics are typically very disturbed individuals, who have themselves experienced serious problems with their parents which they are now replaying with their own children.

Bowen's work, however, goes well beyond his contributions regarding schizophrenic patients. In fact, Bowen is recognized as the father of family systems therapy (Papero, 1990). His theories influenced our current focus on triangulational, multigenerational interactions, and the sibling subsystem. He was also influential in adapting his family theories into a clinical model of family systems therapy. According to Bowen, family systems therapy is phasic, consisting of evaluation and intervention. To assist in the evaluative process, he developed the "genogram" as a method of depicting multigenerational patterns and influences (Papero, 1990).

5.07.3.1.5 About schizophrenia

Research of the family's role in the onset of schizophrenia was not merely coincidental. At this time, psychosocial interventions with schizophrenia were virtually ineffective and psychotropic medications were virtually nonexistent. The medical treatments that did exist (e.g., electroconvulsive therapy and lobotomy) were associated with severe side effects. There was obviously a strong demand for identifying causal pathways to the disease, and perhaps an even greater need for identifying effective interventions. Many of the assumptions of these early family theorists about the relationship between schizophrenia and pathological family relationships have since been refuted by contemporary theories and research of schizophrenia. Perhaps the error in these early theories about family process and schizophrenia was to not consider the possible impact of the child's schizophrenic functioning on parents' behaviors over the course of the child's development. Nevertheless, research with schizophrenia proved to be fertile ground for family therapy theorists to take their first steps.

5.07.3.2 Other Major Theoretical Figures in Family Systems Therapy

Among the most influential figures in the systemic movement were John Bell; Don Jackson—the founder of the Mental Research Institute—who was later joined by Virginia Satir, Jay Haley, John Weakland, Paul Watzlawick, Arthur Bordin, and Richard Fisch; Nathan Ackerman; Carl Whittaker; Ivan Boszormenyi-Nagy; Gerald Zuk; James Framo; Salvador Minuchin—the founder of the Philadelphia Child Guidance Clinic—who worked closely with Braulio Montalvo and Bernice Rosman; and Mara Selvini-Palazzoli at the Institute for Family Studies in Milan.

It is impossible to document all of the important contributions made by each of these leaders. Excellent reviews are found elsewhere (Guerin, 1976; Goldenberg & Goldenberg, 1983; Kaslow, 1980; Nichols, 1984; Nichols & Schwartz, 1995). Suffice it to say that nearly every current model of family therapy has strong theoretical ties to the work of one—and sometimes many—of these family theorists. These individuals were critical in bringing basic research and theory about family interactions to life in the clinical setting. They proliferated new theories about the family system and developed innovative techniques to modify family interactions. The seminal work of these investigators carried over into the 1970s as family therapists developed a new wave of techniques to change family interactions (Goldenberg & Goldenberg, 1983), including working with families at their home (Speck, 1964) and working in the community with extended family members, friends, neighbors and other persons (Speck & Attneave, 1973).

It is unfortunate, however, that the rush towards clinical practice was so profound that research into basic family processes and into the efficacy and effectiveness of family systems interventions was delegated to a lesser role. For example, in one comprehensive review of the literature Wells & Dezen (1978) identified only 20 well designed studies that examined the efficacy or effectiveness of family systems interventions. In the following section, we review evidence for the efficacy and effectiveness of family therapy models that have been influenced by the theoretical and clinical work of these early family systems theorists.

5.07.4 EFFICACY AND EFFECTIVENESS RESEARCH OF FAMILY SYSTEMS THERAPY

In reviewing the current empirical status of family systems therapy, our primary strategy involved identifying groups of studies or programs of research within a particular target population (e.g., depression, conduct disorder). Individual studies and case studies were excluded. The studies reported here are organized along three broad clinical domains: (i) mood disorders; (ii) psychosomatic disorders; and (iii) conduct disorders. Most of the studies cited here included randomization procedures for assigning cases to treatment, an appropriate comparison condition, and clear inclusion/exclusion criteria specifying a "clinical" child/adolescent population. Some of the effectiveness studies included in this review did not have a formal comparison group, but did have some degree of methodological rigor (e.g., pre–post assessments, procedures for monitoring of the delivery of intervention).

Although it was our intention to address the breadth of contexts in which family systems therapy has been applied, we attempted to identify those family models with rich empirical as well as clinical histories. It is our assertion that in a review chapter such as this it is important to include family interventions that have some level of empirical justification. We do not wish to downplay the potential utility of applying family systems therapy to a variety of clinical populations. However, we do wish to convey the importance of the interplay between theory, research, and clinical model building. Table 1 contains relevant information from selected studies in this review.

5.07.4.1 Mood Disorders

Similar to reviews by Kazdin (1994), Alexander, Holtzworth-Munroe, and Jameson (1994), and Estrada and Pinsof (1995), our review of the literature noted a lack of research on the efficacy and effectiveness of family systems therapy for internalizing disorders with children and adolescents. This section reviews the sparse research that does exist, and attempts to highlight promising programs underway at the time of writing.

5.07.4.1.1 Depression

We were not able to identify a single study that examined the efficacy of family systems therapy with depressed youth. In our discussions with family systems therapists and researchers, we were able to identify two significant depression studies. First, Brent and colleagues at Western Psychiatric Institute have made a large efficacy study (funded by the National Institutes of Mental Health) comparing the effects of different psychotherapy models, including family systems therapy, with depressed youth (Brent, personal communication February 8, 1997).

Second, Diamond (Philadelphia Child Guidance Center and Medical School of the University of Pennsylvania) and Siqueland (Medical School of the University of Pennsylvania) have adapted structural family therapy for treating depressed adolescents (Diamond & Siqueland, 1995). This approach integrates structural family therapy (Minuchin, 1974; Minuchin & Fishman, 1981; Stanton et al., 1982) and multidimensional family therapy principles (Liddle, 1991; Liddle, Dakof, & Diamond, 1991) as well as research on attachment. Therapy focuses on: (i) re-establishing or strengthening the parent–adolescent relationship to provide both structure and emotional support for the adolescent; (ii) decreasing marital or parent–adolescent conflict, and improving negotiation and communication within the family; and (iii) decreasing parental hostility and criticism. Diamond and Siqueland (1996) re port important changes in depressive symptoms, attachment, and family interaction in pilot cases. This model is under formal development (funded by the National Institute on Mental Health) at the Philadelphia Child Guidance Center.

5.07.4.1.2 Youth with Other Mood Related Symptoms

Henggeler and colleagues (in press) compared the effectiveness of multisystemic therapy (MST) to hospitalization for youth who presented with suicidal ideation, homicidal ideation, or threat of harm to self or others due to mental illness. MST (discussed further in Section 5.07.4.3.1), is a family-based, contextual approach that addresses the complexity of factors (family and environmental) that influence the youth's psychological and behavioral adjustment. In this pilot investigation, youth in multisystemic therapy showed a greater decrease in internalizing and externalizing symptoms than youth who had been hospitalized. MST was effective in reducing rates of hospitalization and total days hospitalized.

In a study conducted by Gutstein, Rudd, Graham, and Rayha (1988) a systemic crisis intervention was implemented with suicidal, depressed, or behavior problem youth who presented with an immediate crisis. The crisis intervention involved stabilizing family members' anxiety surrounding the crisis, mobilizing extended family members to become involved around the crisis, and restructuring family and extended family relationships to provide long-term solutions to the current crisis. Results demonstrate the effectiveness of the systemic crisis intervention in ameliorating family members' sense of crisis, limiting future suicidal attempts, and improving family functioning over time.

5.07.4.2 Psychosomatic Disorders

There are promising finding on the effectiveness of structural (e.g., Minuchin et al., 1975, 1978) and strategic (e.g., Palazzoli–Selvini,

Table 1 Review of selected publications.

Reference	*Model*	*Population*	*Design*	*Finding*
Mood disorders				
Diamond & Siqueland (1995)	Structural/ Multidimensional/ Attachment principles	Depression	Pilot data for 10 families participating in a treatment development study. Pre–post assessments of emotional problems, and intensive assessments of therapeutic processes included	Changes in depressive symptoms, attachment, and family interaction were noted
Henggeler et al. (in press)	Multisystemic (MST)	Suicidal or homicidal ideation, or psychosis	Participants ($n = 64$) were randomly assigned to MST or hospitalization, and completed assessments of emotional and behavioral problems at five time points. Monthly interviews monitoring utilization of mental health and juvenile justice services were also conducted	MST showed a greater decrease in internalizing and externalizing scores, and reduced rates of out-of-home placements and total days of stay
Gutstein et al. (1988)	Systemic crisis intervention (SCI)	Suicidal ideation or attempt, depression, violent behavior, or serious behavior problems	All participants ($n = 75$) received a brief systemic intervention ($<$ 20 hours), and completed four assessments of emotional and behavioral problems and family functioning	SCI effective in resolving crisis that precipitated treatment and decrease family dependence on institution
Psychosomatic disorders				
Minuchin et al. (1978)	Structural (SFT)	Anorexia	Participants ($n = 53$) were the first 53 anorectic patients treated. Data were gathered on the presenting characteristics of patients, the course of treatment, and follow-up data (approximately 2 years)	At follow-up, 86% of participants were rated as "recovered" from anorexia and its psychosocial components
Dare et al. (1990); Russell et al. (1987); Szmukler et al. (1985)	Structural/ strategic (SSFT)	Anorexia or bulimia	Participants ($n = 80$) were randomly assigned to SSFT or invidual therapy (IT), and completed assessments at intake and follow-up	SSFT and IT effective; however, SSFT was more effective when age of onset was before 18
Conduct disorder				
Alexander & Parsons (1973)	Functional (FFT)	Delinquent	Participants ($n = 86$) were randomly assigned to FFT, group therapy, psychodynamic family therapy, or control. Recidivism rates were obtained for 6–18 months following treatment	FFT more effective than other conditions in reducing recidivism rates

Table 1 continued

Reference	*Model*	*Population*	*Design*	*Finding*
Klein, Alexander, & Parsons (1977)	FFT	Siblings of youth in Alexander & Parsons (1973)	Arrests rates of the siblings of the delinquent youth in the original sample were compared at $2^1/_2$ year follow-up	Siblings of youth in FFT had lower rates of court involvement than siblings in other conditions
Barton et al. (1985)	FFT	Delinquent	Participants ($n = 74$) were referred to FFT or alternative treatment. Youth in the alternative treatment were selected to match youth in FFT. Number of arrests and recidivism rates were obtained for the 15 month period following incarceration	Youth in FFT were less likely to commit subsequent offenses than youth in alternative treatment. Youth in FFT also committed fewer offenses
Gordon et al. (1988); Gordon, Graves, & Arbuthnot (1995)	FFT	Delinquent	Participants ($n = 54$) received FFT or probation only. Recidivism rates were obtained for the 2–3 year period following intervention. In a follow-up study with these youth (now adults), court records from the next 3 years (5–6 years post intervention) were obtained	FFT reduced recidivism rates in the 2–3 year period following intervention (11% vs. 68%). Follow-up data show an arrest rate of 9% for FFT cases compared to 41% for the probation cases
Borduin et al. (1995)	Multisystemic (MST)	Delinquent	Participants ($n = 176$) were randomly assigned to MST or individual therapy. Assessments were administered at pre- and post-treatment	MST improved family functioning and adjustment, and prevented future arrests (4-year follow-up)
Henggeler, Melton, & Smith (1992)	MST	Delinquent	Participants ($n = 96$) were randomly assigned to MST or treatment as usual. Pre- and post-treatment assessments conducted, and criminal behavior and incarceration for 15 months following referral was obtained	MST reduced arrests and days incarcerated, and improved family functioning and peer relations
Henggeler et al. (1986)	MST	Delinquent	Participants ($n = 124$) received MST or treatment at a community agency. Participants were matched on important demographic characteristic. A nondeviant comparison group was included. Pre–post treatment assessments were completed	MST more effective in improving conduct, anxious-withdrawn behaviors, and family and peer relations
Szapocznik et al. (1989)	Structural (SFT)	Conduct disorder	Participants ($n = 69$) were randomly assigned to SFT, individual child dynamic psychotherapy (IT), or a no treatment control group. Assessments of behavior problems, family functioning, and individual dynamics were conducted at pre-, post-, and 1-year follow-up	SFT and IT reduced behavior problems. IT effective in improving individual dynamic functioning, but associated with a deterioration in family functioning, SFT effective in maintaining family functioning

Table 1 continued

Substance abuse				
Santisteban et al. (in preparation)	Structural (SFT)	Behavior problem (at risk) youth	Participants ($n=79$) were randomly assigned to SFT or a control condition administered in group format (CG). Assessments were conducted at pre-, post-, and at 1-year follow-up	SFT more effective than CG in improving behavior problems and in maintaining family functioning. Exploratory analyses suggest that improvements in family functioning may mediated treatment outcome in SFT; however, only in those cases identified as "poor family functioning" at intake
Szapocznik et al. (1983) Szapocznik, Kurtines et al. (1989)	One person family therapy (OPFT)	Behavior problem (at risk) youth	Participants ($n=75$; pooled sample) were randomly assigned to OPFT (individual sessions) or SFT (conjoint sessions). Assessments of behavior problems and family functioning were conducted at pre-, post- and follow-up	OPFT as effective as SFT in reducing behaviors problems and drug use, and in improving family functioning. Results demonstrate that a focus on family interactions in an individual treatment context yields results similar to those found in conjoint family sessions
Szapocznik et al. (1988)	Strategic structural systems engagement (SSSE)	Behavior problem (at risk) youth	Participants ($n=108$) were randomly assigned to SSSE or engagement as usual (EAU). EAU was designed to resemble community standards in engaging behavior problem youth in treatment. Treatment was monitored to ensure differential level of engagement interventions in the SSSE and EAU groups	SSSE was more effective than EAU in engaging families and youth into treatment. Participants in SSSE were also more likely to complete treatment than participants in EAU
Santisteban et al. (1996)	SSSE	Behavior problem (at risk) youth	Participants ($n=193$) were randomly assigned to engagement family therapy (SSSE + SFT), family therapy (SFT only), and group therapy (GT only). Treatment was monitored to ensure differential levels of engagement interventions across treatment groups	Results demonstrate the effectiveness of SSSE in engaging participants into treatment. Further analyses shed light on variables moderating the effectiveness of the specialized engagement interventions
Liddle et al. (1997)	MDFT	Substance abuse	Participants ($n=95$) were randomly assigned to MDFT, adolescent group therapy, and multifamily educational intervention. Assessments were conducted at intake, termination, and at 6 and 12 month follow-up	MDFT demonstrated lower drug use and higher grades at termination than the alternative treatments. Results were maintained at 12 month follow-up

Boscolo, Cecchin, & Prata, 1974) family systems therapy in the treatment of psychosomatic disorders (e.g., eating disorders, diabetes, asthma). Much of this work is represented in the widespread publication of numerous case studies demonstrating the positive effects of family therapy with psychosomatic youth. However, a solid empirical base for family systems therapy is also emerging.

Family systems intervention with psychosomatic youth is traced directly to the work of Minuchin and colleagues at the Philadelphia Child Guidance Center. In their initial work, Minuchin et al. (1978) described five characteristics of "psychosomatic families": (i) enmeshed subsystem boundaries; (ii) overprotectiveness; (iii) rigidity; (iv) lack of conflict resolution; and (v) the child's involvement in parental conflict. In an initial study of structural family therapy, Minuchin et al. reported an 86% recovery rate.

There is considerable controversy, however, about the notion of whether there is such an entity as the "psychosomatic family." Coyne and Anderson (1989) assert that over-protection and enmeshment may be adequate responses of a system to the crisis of the patients. Moreover, other researchers (Garfinke & Garner, 1982; Grigg & Friesen, 1989; Marcus & Wiener, 1989) have found no unique pattern of family interaction that characterizes these families. This controversy is further fueled by the fact that Minuchin's impressive findings have never been replicated in subsequent studies (Kog, Vandereycken, & Vertommen, 1989).

Nevertheless, family therapy appears to be beneficial for patients with eating disorders (e.g., Kog & Vandereycken, 1989). For example, in what is perhaps the best designed study in this area, Dare and his colleagues (Dare, Eisler, Russell, & Szmukler, 1990; Russell, Szmukler, Dare, & Eisler, 1987; Szmukler, Eisler, Russell, & Dare, 1985) compared the efficacy of structural/strategic family therapy and individual therapy with anorectic and bulimic patients. Results demonstrate that both family and individual therapy are effective treatments with anorexic and bulimic patients. However, certain pretreatment patient attributes interacted with treatment. For example, family therapy was more effective than individual therapy when the patient's age of onset for the eating disorder was below 18.

Taken together, these results suggest that family systems therapy may be a promising treatment for adolescents with eating disorders. Although there is considerable controversy about whether or not there is a distinct entity called the "psychosomatic family," the overall results of treatment research in this area are encouraging. However, further research is needed to investigate the assumed theoretical mechanisms of action: the potential mediating role that changes in family interaction may have on treatment outcome with psychosomatic youth.

5.07.4.3 Conduct Disorders

In contrast to research with internalizing disorders, there is a large body of treatment research on the efficacy and effectiveness of family systems therapy with conduct disorders. Reviews of the treatment outcome research literature have consistently documented the efficacy and effectiveness of family-based approaches with problem behavior youth (e.g., Alexander et al., 1994; Chamberlain & Rosicky, 1995; Gurman, Kniskern, & Pinsof, 1986; Kazdin, 1993, 1994; Liddle & Dakof, 1995a, 1995b; Nichols & Schwartz, 1995; Tolan, Cromwell, & Brasswell, 1986). In this section, we review family systems interventions for youth with delinquency and substance abuse.

5.07.4.3.1 Delinquency

Family systems therapy has received extensive support with delinquent and behavior problem adolescents. Two programs of research have emerged for working with these youth: Functional family therapy and multisystemic family therapy.

Alexander and colleagues at the University of Utah have developed, refined, and evaluated the effectiveness and efficacy of functional family therapy (FFT; Alexander & Parsons, 1982; Barton & Alexander, 1981) with delinquent youth. FFT is based upon early family systemic (e.g., Haley, 1976) and behavioral (e.g., Patterson, 1971) frameworks, and has clearly identified and manualized specific phases of treatment (Alexander & Parsons, 1982; Morris, Alexander, & Waldron, 1988). Each phase includes a specific description of goals, requisite therapist characteristics, and techniques (Alexander et al., 1983).

In the initial phase of treatment, the introduction phase, the focus is on therapist characteristics that enhance the family's expectations about therapy. Empathy, compassion, and professionalism are all critical features of this phase. The key feature of this phase is that the therapist attempts to maximize the likelihood that the family will return for additional sessions with "hope" that family therapy can be beneficial.

In the second phase of treatment, the assessment phase (described below, see Section 5.07.5.2.1), the main therapeutic task is to identify the interpersonal relatedness needs (i.e., functions) of family members, and the maladaptive interactions and behaviors that family members implement to meet these needs.

In the third phase of treatment, the therapy phase, the main goal of the therapist is to create a motivational context that is conducive for long-term changes to occur. As in the introduction phase, basic therapist relational skills are a requisite for this phase; however, to meet the goals of this phase, these skills include the extensive use of therapist reframing to disrupt negative, defensive family interactions.

In phase four, the behavior change phase, FFT prescribes a greater emphasis on therapist structuring characteristics. Although relational skills are still important, therapists utilize a battery of behavioral techniques to promote behavior change within the family (Alexander et al., 1983). Techniques fall into three major classes: communication training, assignment of specific activities, and use of technical aids. FFT further emphasizes that these specific techniques must be tailored to the unique characteristics of each family, and must be consistent with the capabilities and interpersonal relatedness needs of each family member (Alexander et al., 1994; Barton & Alexander, 1981). Resistance to behavior change is redefined as the failure of the therapist to adequately motivate family members to change, an inappropriate assessment of interpersonal needs (i.e., functions), or the implementation of behavior change techniques that are not consistent with the functional needs of individual family members. Hence, resistance is moved from an intrapersonal or intrafamilial to a therapeutic holon.

Finally, in phase five, the generalization phase, the therapist is focused on generalizing the gains made in therapy to real world settings. Problematic interactions in other systems are identified and targeted for change, and potential future problems are anticipated and strategies are generated to deal with these problems.

In their initial investigations (Alexander & Parsons, 1973; Parsons & Alexander, 1973), FFT was shown to be superior to client-centered group therapy, psychodynamic family therapy, placebo treatment condition, and no-treatment control groups in reducing recidivism with primarily "status" offending adolescents. In a two and a half year follow-up (Klein, Alexander, & Parsons, 1977), siblings of youth in the FFT condition had significantly lower rates of court involvement than siblings from youth in the other conditions. In a subsequent investigation, Barton et al. (1985) further demonstrated the effectiveness of FFT by extending these findings to multiply offending delinquents. Independent researchers have also documented the efficacy and effectiveness of FFT with behavior problem youth (Friedman, 1989; Gordon, Arbuthnot, Gustafson, & McGreen, 1988). In one investigation, examination of arrest rates during the three years following intervention, reveals that FFT cases (now adults) have an arrest rate of 9% for criminal offenses, compared to a 41% arrest rate for youth who had received probation services only (Gordon, Graves, & Arbuthnot, 1995).

Henggeler and colleagues' MST (Henggeler & Borduin, 1990) has also received extensive recognition for its effectiveness with delinquent youth. MST was developed for working with youth with severe delinquency and behavior problems. In MST, interactions in the family are critical; however, interactions with other important systems are also targeted. In particular, MST addresses interactions with school, peers, and juvenile justice. MST borrows from many models of family systems intervention to facilitate changes in the family system, and does not articulate a unique family systems component. MST has been shown to be effective in reducing delinquency, substance use, and other behavioral problems in several well-controlled randomized investigations (Borduin, Henggeler, Blaske, & Stein, 1990; Henggeler, Melton, & Smith, 1992; Henggeler et al., 1986). Perhaps more than any other family therapy model with adolescents, MST is having an impact at the policy level, as reflected by its adoption for public funding.

5.07.4.3.2 Substance abuse

Family systems therapy (FST) is also gaining recognition as an effective alternative for treating adolescent substance abuse. Liddle and Dakof (1995a, 1995b) note that a number of studies have been conducted since the mid-1980s examining the efficacy and effectiveness of FST with this clinical population (Azrin, Donohue, Besalel, Kogan, & Acierno, 1994; Friedman, 1989; Henggeler et al., 1991; Joanning, Thomas, Quinn, & Mullen, 1992; Lewis, Piercy, Sprenkle, & Trepper, 1990; Liddle & Dakof, 1994; Liddle et al., in press).

Liddle and Dakof (1995a, 1995b) also highlight Szapocznik and colleagues' program of research examining the efficacy of structural family therapy (SFT) with substance abusing youth (Szapocznik, Kurtines, Foote, Perez-Vidal, & Hervis, 1983, 1986; Szapocznik et al., 1988). This program of research has attempted to identify the needs of minority families and to develop, demonstrate, and investigate intervention strategies for substance abusing, behavior problem youth. The version of SFT that has emerged from this research

(Szapocznik & Kurtines, 1989) draws from structural (Minuchin, 1974; Minuchin & Fishman, 1981; Minuchin, Rosman, & Baker, 1978) and strategic (Haley 1976; Madanes, 1981) traditions in family therapy theory as well as from a rich clinical and research history working with inner city minority families.

Central to SFT theory is the concept that dysfunctional family structures (i.e., repetitive patterns of family interaction that characterize the interdependence of family members) give rise to or maintain the youth's problems (Minuchin Montalvo, Guerney, Rosman, & Schumer, 1967; Minuchin, 1974; Minuchin & Fishman, 1981; Szapocznik & Kurtines, 1989; Szapocznik & COSSMHO, 1994). Thus, the focus in SFT is on diagnosing and correcting repetitive patterns of interaction.

The three basic techniques in SFT are joining, diagnosing, and restructuring (Minuchin & Fishman, 1981; Szapocznik & Kurtines, 1989; Kurtines & Szapocznik, 1996). "Joining" refers to establishing a therapeutic relationship, in which the therapist is both a member and the leader, and that includes all family members. Joining, in addition to establishing rapport with each family member, requires identifying interactional patterns that guide family functioning; respecting such interactional patterns; and, entering the family by utilizing its established structure. As described in more detail below (see Section 5.07.5.2.2), diagnosis occurs along five dimensions: (i) structure, (ii) resonance, (iii) developmental stage, (iv) identified patienthood, and (v) conflict resolution.

"Restructuring" defines the broad range of interventions that the therapist undertakes to change maladaptive interactional patterns identified in the diagnostic process. These techniques represent a set of complementary tools that are not intended to be independent of each other, and include:

(i) *Working in the present*—the active ingredient is transforming family interactions in the here and now, which is distinguished from the typical use of "here and now" that refers to the intrapersonal process of experiencing feelings in the present. This present orientation is defined in terms of a focus on the process of family interactions, rather than on their content, or individual processes.

(ii) *Reframing*—represents a strategy of cognitive restructuring that promotes a new understanding or meaning for the family which in turn permits new and more adaptive interactions to emerge.

(iii) *Shifting boundaries and alliances*—involves breaking inappropriate alliances between family members and creating more adaptive and flexible alliances. For example, a common problem in families with a drug using adolescent is the tendency of parent figures to deflect dyadic conflict by involving the problem adolescent in the conflict. This process, labeled triangulation, often occurs when a lopsided alliance is formed between one parent figure and the problem youth against the other parent figure. By better balancing the overly strong and overly weak parent–youth alliances, and the overly weak parental alliance, more direct/uninterrupted communication between the parent figures gives rise to a new and stronger boundary around the parents.

(iv) *Redirecting communications*—involves promoting direct and specific communication between dyads, which sometimes requires blocking the interference of family members that do not permit such communications to occur.

In their initial investigations, Szapocznik and colleagues demonstrated that a family approach was effective with Hispanic families (Scopetta et al., 1977), and that family psychoeducational interventions rich with cultural content and aimed at transforming family interactional patterns could produce results as strong as family therapy (Szapocznik, Santisteban et al., 1986). In a subsequent study, combining a family approach with a cultural intervention, they demonstrated significant improvements in children's problem behaviors and family functioning (Szapocznik, Santisteban et al., 1989). In separate investigations, the efficacy of SFT with substance abusing and at risk adolescents was compared to a group therapy control condition (Santisteban et al., 1996, 1997) and individual dynamic psychotherapy (Szapocznik, Rio et al., 1989). In the former investigation, SFT produced significantly greater improvements in behavior problems than group counseling. Moreover, analyses of clinically significant change revealed that a substantially larger proportion of family therapy cases demonstrated reliable clinical improvement in behavior problems than did group cases. In the latter study, both SFT and the individual dynamic psychotherapy condition demonstrated significant improvements in behavioral problems (as compared to a no-treatment control group). However, SFT resulted in one-year follow-up improvements in family functioning, while individual dynamic psychotherapy resulted in a one-year follow-up deterioration in family functioning.

Szapocznik and colleagues have also developed and investigated two unique applications of SFT. The first application, one person family therapy (OPFT), is an intervention that is designed to target changes in maladaptive family interactions, but that attempts to do so without the whole family present in therapy (Szapocznik & Kurtines, 1989; Szapocznik, Kurtines, Perez-Vidal, Hervis, & Foote, 1989). OPFT has been shown to be as effective as the conjoint family approach in bringing about and maintaining significant improvements in: (i) behavior problems and drug use in youth, and (ii) family functioning (Szapocznik, et al, 1983; Szapocznik, Kurtines et al., 1986). For one-person interventions to produce desired changes in family interactions they had to explicitly target family change (Szapocznik, Rio et al., 1989).

The second application, strategic structural systems engagement (SSSE; Szapocznik & Kurtines, 1989; Szapocznik, Perez-Vidal, Hervis, Brickman, & Kurtines, 1989), was designed to more effectively engage adolescent drug users and their families in treatment. This approach was designed to overcome resistance by restructuring the interactions that permit resistance to occur. A core assumption in SSSE is that resistance is interaction-based; and consequently, therapists can strategically alter their behaviors to change the interactions that are interfering with the engagement process. In separate investigations with drug using Hispanic youth (Santisteban et al., 1996; Szapocznik et al., 1988), SSSE was significantly more effective than conventional engagement procedures in engaging in youth and their families into treatment. Moreover, the dropout rates in therapy were up to three times lower for families that had received the SSSE intervention than for those families that had not received this intervention.

Liddle and colleagues have developed and refined multidimensional family therapy (MDFT; Liddle; 1992, 1995) for inner city adolescent substance abusers. MDFT is an integrative family-based intervention that draws from structural family therapy, social learning theory, developmental theory, attachment theory, and empirical work (Liddle, 1994) to guide assessment and intervention within four interacting systems: (i) family system; (ii) parent–adolescent relations; (iii) parent subsystem—the self of the parent (parent as a parent and an individual adult); and (iv) adolescent subsystem—the self of the adolescent (adolescent as a son or daughter and as a member of a peer culture). Preliminary evidence for the efficacy of MDFT has been demonstrated in a controlled randomized clinical trial study comparing MDFT to adolescent group therapy and a multifamily educational intervention condition (Liddle et al., in press). Results demonstrated MDFT's superiority in reducing drug use and improving grades. The results from this study are currently in preparation. Data are currently being collected in a second National Institute on Drug Abuse funded controlled clinical trial study which compares MDFT to an individual-based cognitive therapy condition.

5.07.5 ASSESSMENT AND CLINICAL FORMULATION

The presenting problem or symptom exhibited by the child or adolescent (e.g., depression, substance abuse) is the family's ticket to coming to therapy. Family members typically enter therapy with the explicit goal of wanting to change the youth, and have minimal awareness of their own contribution to the problems experienced by the youth. For the family systems therapist, however, family interactions are the primary focus of treatment. The key to effective intervention thus begins with an accurate assessment and clinical formulation of family interactions that are related to the youth's problems. Assessment and clinical formulation (and ultimately intervention) always consider the general nature of family functioning as well as the specific relationship between the presenting problem and general family functioning.

Assessment and clinical formulation involves identifying strengths and weaknesses within the family. Weaknesses (i.e., problematic interactions) are presumed to be directly linked to the problem presented by the child. For example, parental conflict, parent–child relationship, triangulation, and a lack of subsystem differentiation, may all be related to the problem experienced by the child or adolescent. Family systems interventions reflect this focus on problematic interactions.

It is important to note, however, that the therapist's assessment of the family's strengths is perhaps as important as identifying weaknesses. Strengths refer to adaptive family interactions as well as to the particular capabilities of individual family members. This focus on strengths is not merely "lip service" or a reaction to "pathology-based" intervention models. Because the majority of family systems interventions are brief and problem-focused, therapists must utilize the family's resources to maximize their impact on the family. Although treatment does involve some skills building, it is not the primary focus of intervention. Rather, the primary focus is on utilizing the family's strengths to achieve the goals of treatment.

At a microlevel (moment by moment) in treatment, the therapist looks for intrapersonal and interpersonal strengths. Some examples of intrapersonal strengths include positive features such as love, commitment, and a desire to make things better. Interpersonal strengths include open and direct communication between family members, positive expressions of support, and the healthy expression of differences of opinion. These strengths are contrasted with microlevel intrapersonal and interpersonal weaknesses. Intrapersonal weaknesses at this level include hopelessness, anger (at worst, hate, or contempt), and negative attributions about self or other. Interpersonal weakness include the expression of contempt (as opposed to appropriate disagreements), intense and escalating levels of negativity, vague and indirect communication, developmentally inappropriate parent or child behavior, and conflict between parent figures.

At a molar level (assessment and clinical formulation of the family system), the strengths and weaknesses of complex patterns of interaction are considered. Some examples of strengths at this level include the family's flexibility to respond to changing internal (within family) and external conditions, their ability to negotiate and resolve their differences of opinion, and their ability to recognize that their problems are multifaceted as opposed to residing within one family member. Weaknesses at the molar level include a lack of parental cooperation in setting and enforcing rules, overly connected or disconnected interpersonal boundaries, the centralization of one family member, and the consistent denial, avoidance, or diffusion of family conflict.

5.07.5.1 Conducting Assessments in Family Systems Therapy

Family therapists are active and directive. They gather information about family interactions by asking directive questions about family interactions and by encouraging family members to interact in the treatment context. The process of assessment and clinical formulation, however, varies across family systems models. Thus, there is no standard, universally accepted procedure for conducting assessments and clinical formulation. Nevertheless, some of the strategies for assessing interactions are similar across models. In this section, we describe two general strategies for conducting assessments in family systems theory: (i) the family genogram, and (ii) enactments.

5.07.5.1.1 Family genogram

The genogram is a frequently used method for identifying who is in the family, and what role important family members play with respect to one another. The conventional genogram includes biological family; however, genograms may also include individuals based on the function they serve within the family. For example, in our own work with inner city African-American families (Greenwood, Herves et al., 1995; Greenwood, Samuels, et al., 1995), defining "family" solely through biological relationships fails to represent the complex family/kinship network of these families. What is required is a more flexible approach to identifying family/kin that examines: (i) blood relation and/or extended kinship; (ii) perceived strength and duration of relationships; and (iii) perceived support—in either direction—(including financial, emotional, and instrumental help, e.g., who provides transportation, child care, and other types of assistance). In our experience, "family" often includes all individuals living within a household (except for roommates and transient "live-ins") as well as other important significant others who may or may not be biologically related to one another.

The genogram is like a family tree in that it depicts all of the important relationships within the family. One of the most important reasons for constructing a genogram is to obtain a more complete picture of the family constellation. The genogram is a tool for systematically identifying who is in the family—getting around the common problem of some families not letting therapists know about key but conflicted or undesirable relations. Thus, the genogram can also be useful in helping the therapist depict adaptive and maladaptive family interactions (e.g., conflicted or supportive relations may be depicted using a jagged line or straight line).

There are many strategies for conducting genograms, and various symbols are used to denote meaningful relationships and/or roles within the family. Regardless of the strategy or symbols used, the genogram is an excellent method for organizing the complex interconnections between family members in a coherent, organized manner. The drawback of the genogram is that it is a "snapshot" of the family, and as such yields relatively little information about family interactions. In other words, the genogram tends to be static rather than dynamic. Genograms represent the first stage of the assessment process, and are often used to establish rapport and gather information about who are the potentially critical "players" in the youth's life. This information provides the context for further assessment into the dynamic interaction patterns that characterize the family members identified in the genogram.

5.07.5.1.2 Assessment through direct observation of family interactions: enactments

Minuchin and colleagues (Minuchin, 1974; Minuchin & Fishman, 1981) were the first to emphasize the importance of enactments in FST. In fact, they describe enactments as the primary vehicle through which interactions are identified and changed in therapy. Enactments are defined as family interactions in the treatment context that reflect the family's characteristic manner of interacting outside of therapy. That is, enactments reflect the family's overlearned behaviors that are present in most situations. Enactments allow the therapist to directly observe family interactions, and are thus critical for assessment and clinical formulation. The basic philosophy is that family members will enact "real world" interactions through their enactments in the treatment. Thus, by observing these enactments in therapy, the therapist can directly assess strengths and weaknesses within the family, including important changes in family interaction that occur as therapy progresses.

Enactments are promoted by a therapeutic focus of encouraging, helping, or allowing family members to behave/interact in their characteristic manner, that is, as they would naturally behave if the therapist were not present (Szapocznik & Kurtines, 1989). Very frequently family members will spontaneously enact in their typical way when they fight, interrupt, or criticize one another. However, because of the nature of therapy, it is not uncommon for family members to centralize therapists, in which case, the therapist will need to be more active in facilitating direct communication between family members (i.e., enactments). To facilitate enactments, the therapist systematically redirects communications to encourage interactions between family members.

Enactments permit the therapist to view the family directly as it really is, with its overlearned, rigidly repetitive patterns of interaction. In this case, the therapist observes problems directly and does not have to rely on stories about what typically happens when the therapist is not present. By doing so, the therapist gathers information about family functioning directly by facilitating family interactions in therapy. We believe that there can be considerable discrepancies between what individuals report and how they actually behave (enact) in the family context. Moreover, family reports tend to be influenced by attributions; whereas, observations of enactments permit the identification of circular/systemic processes.

5.07.5.2 The Underlying Nature of Family Interactions

In clinical work, there is a range of theory driven information that is obtained through assessment. In this section, we present assessment information from two theory driven interventions for which there is also considerable efficacy research.

5.07.5.2.1 Functional family therapy

In FFT, the therapist's assessment is designed to identify the "functions" of each family member in relation to one another. Functions represent a short-hand term for the relational needs, whether they be inborn or learned, that underly individuals' apparent motivations for distance/autonomy and connectedness/interdependency with others (Alexander & Parsons, 1982). These patterns become evident over time in relationships, and often differ for a given individual in different relationships. Thus, unlike the intrapersonal concept of "personality" which presumes a core underlying motivational structure, a functional family therapist's assessment of functions often identifies important motivational differences within one person. A parent, for example, may have a "favorite" child (with whom they might even be "enmeshed"), while at the same time have another child whom they reject and may even attempt to expel from the household. In such situations it is inappropriate to think of the parent in a unitary fashion (enmeshed or disengaged). Instead the functional family therapist understands that the motivational needs of the parent, and perhaps the children, with respect to each other are markedly different. As a result, behaviors that would be syntonic for the parent with respect to the "close" child could be quite dystonic with respect to the "distanced" child. Additional complexity occurs when the motivational structure differs for members of a dyad—for example, a parent retains strong affiliation (connectedness needs) towards a youth, while the youth is "emancipating" (distancing from the parent) and developing much stronger connectedness needs outside the family. In FFT, neither the parent's functional need with respect to the child or the child's functional need with respect to the parent are pathologized. Rather, the therapist identifies the maladaptive behaviors that a person is using to meet their functional needs. For example, adolescents who sneak out of the house after curfew to hang out with friends may be considered to have an interpersonal (i.e., functional) need for distance from their parents. However, the manner in which they are fulfilling this need is considered to be inappropriate (e.g., sneaking out of the house). In FFT, the goal of assessment and clinical formulation is to identify each family members' functional needs with respect to other family members, and to understand the adaptive and maladaptive behaviors that are being utilized by family members to maintain these interpersonal needs.

5.07.5.2.2 Structural family therapy

A different strategy for assessing family interactions is found in the work of Szapocznik and colleagues (Hervis, Szapocznik, Mitrani, Rio, & Kurtines, 1991; Szapocznik & Kurtines, 1989; Szapocznik et al., 1991). The core of assessment and diagnosis in SFT involves identifying repetitive interactional patterns within the family along five interrelated dimensions: structure, resonance, developmental stage, identified patienthood, and conflict resolution.

(i) Structure

This refers to the organizational aspects of the repetitive patterns of interaction within the family. Three specific categories of family organization are examined: leadership, subsystem organization, and communication flow.

Leadership assesses the distribution of authority and responsibility within the family. This category includes (i) hierarchy (Who takes charge of the family's directorship? Is leadership in the appropriate hands? Is it shared? Is hierarchy appropriate with respect to age, role and function?); (ii) behavior control (Who keeps order, if anyone? Are attempts to keep order successful or ignored?); and (iii) guidance (Who provides advice and suggestions? Does the advice provided have an impact on family interaction?).

Subsystem organization is concerned with the formal and informal organization of the family system, including (i) alliances (Who supports whom? Are dyad members closer to each other than to the rest of the family? Are alliances appropriate?); (ii) triangulations defined as an interference by family member C in a conflict between family members A and B, where C acts as an intermediary between A and B and where such interference causes a detour away from the original conflict; and (iii) subsystem membership (Who is a member of which subsystem? Are subsystems comprised appropriately with respect to age and function? Are subsystem boundaries clearly defined?).

Communication flow is concerned with communication flow within the family. The following aspects of the quality, quantity, and direction of communication are assessed: (i) directness of communication (Do individuals communicate directly with each other?); (ii) gatekeepers—switchboard operators (Is there a gatekeeper who controls, directs, or channels communication flow?); and (iii) spokesperson (Is there a family member who speaks for others in the family or for the whole family?).

(ii) Resonance

This is a measure of subsystem differentiation which takes into account the threshold of each family member's sensitivity to one another. At one extreme, boundaries can be either extremely rigid or impermeable (lack of connectedness). At the opposite extreme, they can be too overinvolved. Both of these extremes are considered maladaptive. (This view may be contrasted with functional family therapy in which the interpersonal relatedness needs, i.e., functions, are not considered to be maladaptive.) Ideally, there is a midpoint which allows for permeability, such that interaction and communication are possible at appropriate times, while retaining adequate differentiation and separateness.

In assessing resonance, allowances must be made for the cultural background of the family (Kurtines & Szapocznik, 1996). However, when behaviors result in symptoms, they are considered problematic whether or not they are culturally sanctioned. For example, some families or relationships may be characterized by machismo, manifested as an uncertainty of a man's feelings towards his wife and a nonverbal expression that replaces communication in the form of sexually forcing himself onto his wife. In this extreme and stereotyped example of machismo, even if machismo is cultural, this pattern of behavior would clearly be viewed as maladaptive.

(iii) Developmental stage

This refers to the appropriateness of family members' interactions with respect to roles and tasks assigned to various family members, taking into consideration their age and position within the family. The following sets of roles and tasks are considered in examining this dimension. Parenting roles and tasks: Are parents parenting at a level consistent with the age of the children? For example, are controlling and nurturant functions in accordance with child's age and stage of psychological development?

Spousal roles and tasks: Are spouses parenting at cooperative and equal levels of development? Child/sibling roles and tasks: Do the children function competently for their age and have appropriate rights and responsibilities? Extended family member's roles and tasks: In extended families, are parents able to assume proper parental position relative to their children in light of the role of their own parents and other relatives?

It is noteworthy that the major flare-ups in family stress and conflict are likely to emerge at times when developmental milestones are reached, because these occasions expose the families to unexpected demands which stress the family. Frequently, as the family responds in old ways to the new emergent situation represented by a developmental milestone, crises at the interface between homeostasis and flexibility often propel the family to seek help.

(iv) Identified patienthood

Identified patienthood (IPhood) refers to the extent to which the family is convinced that their primary problem is all the fault of the person exhibiting the symptom. Thus, IPhood refers to the extent to which the family places the identified patient (IP) and the patient's identified problem to be the sole problem of the family and uses that IPhood as a means of maintaining family homeostasis. There are five signs indicative of strong IPhood: (i) negativity about the IP—statements to the effect that the IP is seen as the sole cause of family pain and unhappiness; (ii) IP centrality—the IP is frequently the center of attention and topic of conversation and interactions; (iii) overprotection of the IPhood—family avoids confrontation with IP dysfunction by excusing it, explaining it away, minimizing it, or feeling that there is nothing that can be done about it ("That's just the way they are."); (iv) nurturance of IPhood—the IP dysfunction is supported or abetted by other family members, for example, a sibling does the homework for an IP who skips school; and (v) denial of other problems—statements implying that the IP is causing problems and pain coupled with statements that there are no other family problems ("We would be a happy family if it wasn't for the IP's problem.")

(v) Conflict resolution

This is a measure of the family's style in managing disagreements. It is important to observe the handling of differences and disagreements in a variety of situations since various situations, differing in the level of inherent conflict or level of emotionality, may be handled differently. There are five conflict resolution styles identified: (i) denial—disagreements are not allowed to emerge and situations are structured to prevent different opinions or critical opinions, or redefined to prevent the emergence of conflict; (ii) avoidance—disagreement begins to emerge but is stopped, masked, or strongly inhibited in some way, for example, "Let's not have a fight now," "You're so cute when you're mad," "That's not really important"; (iii) diffusion—moving from one disagreement to another without letting any emerge fully or making personal attacks which are not part of the conflict issue; (iv) emergence without resolution—separate accounts and opinions regarding one disagreement are clearly expressed but no solution is negotiated; and (v) emergence with resolution—separate accounts and opinions regarding a single disagreement are clearly expressed and a single final version or solution acceptable to all family members is negotiated.

There is considerable debate as to which of these conflict resolution styles is most adaptive and successful. From a family systems, communication theory, and psychoanalytic point of view, it might be possible to order these conflict resolution styles from least to most healthy as follows: denial, avoidance, diffusion, emergence without resolution, and emergence with resolution. However, from a more strategic perspective, it can be suggested that different styles may be adaptive at certain times. As demonstrated by Greenwood et al. (submitted for publication; see also Szapocznik & Kurtines, 1989), considerable judgement is recommended in determining what is an appropriate mode of conflict resolution given the overall circumstances confronted.

5.07.5.3 Clinical Formulation in Family Systems Therapy

"Assessment" refers to the process of conducting a systematic review of the molecular aspects of family interaction to identify specific qualities in the patterns of interaction of each family. That is, assessment identifies the strengths and weaknesses of family interactions. In contrast, clinical formulation refers to the process of integrating the information obtained through assessment into molar processes that characterize the family's interactions (c.f., Szapocznik & Kurtines, 1989). In individual psychodiagnostics, clinical formulation explains the presenting symptom in relationship to the individual's psychodynamics. Similarly, clinical formulation in family systems, therapy explains the presenting symptom in relationship to the family's characteristic patterns of interaction.

In FST therapy, assessment and clinical formulation are essential components of the intervention model, and are thus guided by a unitary clinical theory. For this reason, the way in which the information gathered from assessments is combined differs across systemic approaches. This means that, depending on the treatment model, certain aspects of family interaction will receive extensive consideration and other aspects will receive less attention. For example, in FFT, the therapist's clinical formulation will always include a description of the way in which the youth's problem behaviors are meeting interpersonal relatedness needs; whereas, in SFT the therapist's clinical formulation will always include an articulation of the way in which disturbances in family interactions (i.e., structure, resonance, etc.) give rise to, or maintain, problem behaviors.

A variety of factors influence clinical formulation. Although not included in the assessment, information that can help contextualize family interactions is required for clinical formulation. For example, in assessment the appropriateness of each family member's developmental level is evaluated. However, in clinical formulation, information on major family developmental transitions or events that occur within or outside the family are also considered because they help to explain why a symptom may bring a family into treatment at a particular time. Also important to clinical formulation is the chronicity of the patterns encountered, or conversely, the family's flexibility in adapting to internal or external changes. Such flexibility is an important indicator of the level of effort that will be required to bring about changes in the family system.

Individual as well as external systemic factors must also be considered in the clinical formulation. At the individual level, psychological factors such as beliefs, attitudes, intelligence, and psychopathology, as well as biological factors must be considered when evaluating the impact of family interactions on the problems experienced by the youth. Moreover, other systems may have a profound impact on the family, and consequently, must be considered in the clinical formulation (Dishion et al., 1995; Szapocznik, Kurtines et al., 1997; Tolan, Guerra, & Kendall, 1995). For example, the youth's interactions at school or with peers, or the nature of the neighborhood may serve as a powerful risk or protective factor. In addition, parents' extended family, friends, or work may serve as a source of strength or stress that may or may not contribute to the problems experienced by the youth (Szapocznik & Coatsworth, submitted for publication).

5.07.6 MEDIATORS OF CHANGE: MECHANISMS OF ACTION IN FAMILY THERAPY WITH ACTING-OUT YOUTH

One of the most interesting contemporary theoretical and research questions in FST, and psychotherapy research in general, is "what makes therapy effective?" It is of considerable interest that across the various models reviewed there appear to be certain common ingredients that contribute to their efficacy/effectiveness. This section highlights these common ingredients by reviewing existing in-session family therapy process research on the "mechanisms of action" involved in the successful treatment of behavior problem youth. In particular, the primary goal is to identify specific therapist and family processes that contribute to outcome.

In this review, we focus on existing empirical research rather than on theoretical articulations of the therapeutic process. Unfortunately, only a handful of studies have been conducted to identify the critical processes that are essential to effective family therapy. Because so much of this work has been conducted with problem behavior adolescents, this section is organized to present the presumed mediators of successful and unsuccessful outcomes with behavior problem adolescents. Whenever possible, however, we have attempted to integrate research and clinical theory from other work with children and adolescents.

Process research findings from three important domains of family systems therapy are presented in this section: (i) establishing a therapeutic alliance with adolescents and their family members, (ii) managing highly conflicted family interactions; and (iii) changing parent–adolescent and family interactions. Table 2 summarizes this research, and includes clinical recommendations to family systems therapists working with behavior problem youth and their families.

5.07.6.1 The Critical Role of the Therapeutic Alliance in Family Therapy

It is now generally recognized that the therapeutic alliance is an important variable in psychotherapy research and a strong

Table 2 Clinical recommendations for family systems therapy with behavior problem youth.

Clinical recommendation	*Reference*	*Future research questions*
Building alliance with behavior problem youth	Diamond & Liddle (1996)	(i) What strategies are most effective in building alliances with parent figures and other family members? (ii) Are individual or conjoint sessions more effective in building alliances with adolescents (or parent figures)? (iii) How does the alliance change over the course of therapy? (iv) What is the relationship between the adolescent alliance, the parent alliance, and changes in behavioral outcomes?
Present as an ally by showing support and by demonstrating ways in which you can assist in helping the adolescent meet her/his treatment (or personal) goals		
Collaborate with adolescent to formulate personally meaningful treatment goals		
Attend to the adolescent's experiences by reflecting their thoughts and feelings		
Reducing within-family negativity		
Selectively respond to family member negativity by asking questions about exceptions to negative attributions or statements	Melidonis & Bry (1995)	(i) What is the relationship between negativity in the treatment context and changes in family functioning or behavioral outcomes? (ii) Does reframing improve the therapeutic alliance? (iii) Does reframing produce similar results over the course of therapy? (iv) What is the relationship between intrapersonal (i.e., attributional) changes and interpersonal (i.e., interactions) changes?
Reframe negative attributions and expectations in a positive manner	Robbins, Alexander, Newell, & Turner (1996); Robbins, Alexander & Turner (1996)	
Reframe in the context of negative family interactions to minimize the escalation of negativity in the session		
In conjoint sessions, direct statements that orient family to treatment toward parent figures		
Handling resistance		
Approach resistance as interaction based, and adjust interventions during the engagement phase by using structural family therapy principles to overcome resistance to treatment	Szapocznik et al. (1989); Santisteban et al. (1996)	(i) What is the relationship between level of resistance and short-term (i.e., dropout) and long-term (i.e., behavioral changes) outcomes? (ii) Do different family members respond in a similar manner to therapist attempts to manage resistance? (iii) What are the therapist's structural maneuvers that encourage or discourage the family's (and specific family members') involvement in therapy?
Manage resistance by minimizing teaching or confrontive interventions, and by facilitating and supporting family members	Patterson & Forgatch (1985)	
Resolve therapeutic impasse by blocking and working through negative affect, and by amplifying thoughts and feelings that promote constructive dialogue	Diamond & Liddle (1996)	
Changing parent–adolescent interactions		
Focus on improving parental leadership, family subsystem organization, and communication flow	Robbins, Mitrani et al. (1996)	(i) What therapist interventions are most effective in changing parent–adolescent interactions? (ii) Are individual sessions with parent figures or the adolescent necessary to facilitate changes in parent–adolescent interactions? (iii) What is the relationship between the therapeutic alliance and changes in family interaction?
Focus on facilitating the expression and resolution of within-family disagreements		
Focus on improving parenting practices, including positive discipline, monitoring, limit setting and communication, and consistency between parents	Schmidt et al. (1996)	

predictor of therapeutic outcome (Bordin, 1979; Dierick & Lietaer, 1990; Gaston, 1990; Hatcher, Barends, Hansell, & Gutfreund, 1995; Horvath & Luborsky, 1993; Lambert, Shapiro, & Bergin, 1986; Luborsky, Barber, & Crits-Christoph, 1990; Safran, Crocker, McMain, & Murray, 1990; Waterhouse & Strupp, 1984; Weinberger, 1995). While most conceptualizations of common factors in psychotherapy suggest that the therapeutic bond is important in all psychotherapies (Orlinsky, Grawe, & Parks, 1994), most of the research on alliance to date has focused on individual therapy. The individual psychotherapy literature indicates that the alliance itself has a direct therapeutic effect (Henry, Strupp, Schacht & Gaston, 1994; Lambert, 1992; Strupp & Hadley, 1979) in both long and short-term therapies, and appears to operate within treatments from many therapeutic orientations, including the more behavioral therapies in which the alliance is not as likely as in analytic therapies to be overtly mentioned as a therapeutic goal (Eaton, Abeles, & Gutfreund, 1988; Gaston, Marmar, Gallagher, & Thompson, 1991; Gomes-Schwartz, 1978; Horvath & Symonds, 1991; Koss & Shiang, 1994; Krupnick, Sotsky, Simmens, & Moyer, 1992; Muran et al., 1995).

The therapeutic alliance also appears to be an important predictor of the outcome of clients in non-individual treatment modalities (Gaston & Schneider, 1992; Henry et al., 1994; Pinsof & Catherall, 1986). In specific reference to family therapy, Gurman & Kniskern (1978) concluded that "the ability of the therapist to establish a positive relationship ... receives the most consistent support as an important outcome-related therapist factor in marital and family therapy" (p. 875). In their review of family therapy research, Friedlander, Wildman, Heatherington, and Skowrow (1994) assert that components of alliance in family therapy have been predictive of session effectiveness, continuation in treatment, and therapeutic outcome. Although these reviews assert that the therapeutic alliance is critical in family therapy, few studies have been specifically designed to examine the impact of alliance in family therapy. In fact, DiGiuseppe, Linscott, and Jilton (1996) note that process research on the therapeutic alliance in therapy with youth in general is alarmingly scarce.

5.07.6.1.1 Building an alliance with behavior problem youth: specific recommendations

Diamond and Liddle (1996) examined the relationship between specific therapist engagement interventions and the strength of the therapist–adolescent alliance in multidimensional family therapy with drug using adolescents. In this study, therapist interventions and family alliance were rated in first and third sessions of five cases that showed the most improvement in alliance and five cases showing the least improvement in alliance. Results indicate that differences between the improved and unimproved groups do not appear until the third session of therapy. In session one, therapists try equally hard to engage reluctant adolescents. By session three, however, therapists in the improved alliance group were "formulating goals" "presenting themselves as an ally," and "attending to the adolescent's experience" significantly more than therapists in the unimproved group. Therapists in the unimproved group were "orienting adolescents to the collaborative nature of therapy" more than therapists in the improved group, though in both groups this intervention decreased over the first three sessions. These results are consistent with the assertions of other researchers who assert that although these factors are important, adolescents are particularly concerned with being involved in establishing the goals and tasks of therapy (DiGiuseppe, et al., (1996).

5.07.6.2 Reducing Within-family Negativity

Negativity, in the context of family functioning, has been heavily implicated in adolescent drug use and behavior problems. Research has shown that high levels of negativity interfere with effective problem solving and communication within the family (Alexander, Waldron, Barton, & Mas, 1989; Margolin, Burman, & John, 1989). These processes appear to play a similar role in the process of therapy itself, with numerous examples linking negativity in the therapy session to poor therapy outcome. Alexander, Barton, Schiavo, and Parsons (1976), for example, found that the ratio of negative to supportive statements was significantly higher in cases that dropped out of therapy than among cases that completed treatment. In turn, premature termination predicted recidivism in adolescents. In the same study, among the families that completed therapy, a reduction in family member negative behaviors during treatment was associated with positive outcome. Greater rates of within family disagreements have also been shown to be characteristic of noncompleters in structural-strategic family therapy (Shields, Sprenkle, & Constantine, 1991). A comprehensive review of the process of family therapy concluded that, in general, negativity may be predictive of premature termination (Friedlander et al., 1994).

5.07.6.2.1 Examining the impact of therapist interventions on within family negativity

Melidonis & Bry (1995) examined the effects of therapist questions about exceptions to family member statements on rates of blaming and positive statements in four families seen in a crisis unit for disruptive adolescent problem behaviors. They demonstrated that the therapist's use of exceptions questions was associated with a decline in blaming statements and an increase in positive statements.

Robbins, Alexander, Newell, & Turner (1996) examined the impact of specific types of therapist intervention (e.g., reframing, reflection, and organizational) on family members' negative attitude during the initial session of family therapy with a delinquent adolescent. Results revealed that: (i) across all therapist's interventions, adolescents were significantly more negative than mother's following therapist behaviors, but (ii) adolescents' attitudes improved from negative to neutral following therapist reframes but not other therapist interventions.

In a follow-up investigation, Robbins, Alexander, & Turner (submitted for publication) further examined the impact of therapist interventions on family processes. Similar to the previous study, the immediate effects of therapist reframing, reflection, and structuring interventions on family member behaviors were compared; however, only those therapist interventions that followed family member negative statements were included in the analysis. Results demonstrate that for this sample: (i) reframing is more effective than reflection and structuring statements in reducing family member negative behaviors; (ii) the results expanded the previous finding that adolescents respond more favorably to reframes than do mothers and fathers, now providing evidence of the effectiveness of reframing interventions for all family members; and (iii) adolescents responded more favorably to therapists (regardless of intervention) than did family members during negative exchanges.

These results demonstrate that highly negative and conflicted family interactions may be modified by therapist interventions in the treatment context. In particular, in the context of highly conflicted interactions, interventions that are positive and nonblaming appear to have a powerful effect on negative family interactions. These results are particularly evident for the adolescents in these samples, which is particularly promising given their highly negative and oppositional style when interacting with other family members and the therapist. Future research must attempt to link these changes in in-session conflict and negativity to other therapeutic changes and to the ultimate outcome of therapy.

5.07.6.3 Changing Parent–Adolescent Interactions

One of the core factors that consistently emerges in most etiological explanations of adolescent behavior problems is poor parenting and/or parent–child relationships. Aspects of the parent–child relationship that increase vulnerability to adolescent behavior problems include: lack of warmth and involvement, poor communication, lack of support, emotional distance, rejection, and lack of parental management (Kazdin, 1993, 1995; Kumpfer, 1989; Rutter & Guller, 1983; Werner & Smith, 1992; Widom, 1989). Research has also demonstrated that other parent and family factors, such as a lack of parental monitoring (Baumrind, 1991a, 1991b; Patterson & Dishion, 1985), harsh discipline, observed marital conflict, reciprocated negative communications, chronic parent–child conflict, and intense and frequent negative emotional expression (Alexander, 1973; Block, Block, & Keyes, 1988; Kandel, Kessler, & Margulies, 1978; Krinsley & Bry, 1991; Liddle & Dakof, 1995b; Rutter, 1980), play an important role in placing adolescents at risk for behavior problems (Hawkins, Catalano, & Miller, 1992; Jessor, 1976; Patterson, 1982; Simcha-Fagan, Gersten, & Langer, 1986). There is also an extensive literature indicating that changes in parenting practices positively impacts children's problem behaviors (for review see Estrada & Pinsof, 1995).

Support for the utility of interventions targeting the parent–child relationship in the context of family therapy comes from several sources. Parsons and Alexander (1973) found that positive modifications in parent–adolescent communications were associated with reduced recidivism (Alexander & Parsons, 1973) as well as a marked reduction in future offending by siblings (Alexander et al., 1976). In one study, Schmidt, Liddle, and Dakof (1996) linked improvements in parenting to reductions in drug abuse and behavior problems.

5.07.6.3.1 Changes in family interactions in successful and unsuccessful outcome cases

Robbins, Mitrani, Zarate, Coatsworth, & Szapocznik (1996) examined family interactions over the course of therapy in successful and unsuccessful cases to determine if changes in theoretically meaningful family interactions mediated treatment outcome. Four cases were selected that met the following criteria: case 1 showed improvement on measures of conduct disorder and had "good" pretreatment family functioning; case 2 showed improvement in conduct disorder and had "poor" pretreatment family functioning; case 3 showed no improvement and had "good" pretreatment family functioning; and case 4 showed no improvement and had "poor" pretreatment family functioning.

Results indicate that for families that showed improvement on measures of conduct disorder (cases 1 and 2), regardless of pretreatment family functioning, positive changes in family interaction were observed over the course of treatment; while for families that showed no improvement (cases 3 and 4), family functioning did not change, or worsened, over the course of treatment. These results provide limited but important support for the family therapy tenet that changes in family interaction mediate outcome. Moreover, these results highlight the importance of examining the therapeutic processes to understand those factors that contribute to successful therapy.

5.07.6.3.2 Family members' resistance to change

The results presented above suggest that family interactions play a critical mediating role in therapy with substance abusing, behavior problem youth. Attempts to change family interactions are often met with considerable resistance in therapy. For example, therapists appear to encounter the most "resistance" during the phase of treatment in which they are attempting to restructure family interactions (Chamberlain, Patterson, Reid, Kavanagh, & Forgatch, 1984). Therapists must successfully negotiate in-session "resistance" to achieve positive treatment outcomes.

Several important therapy traditions have viewed the therapeutic system's role in resistance and examined what therapists can do to prevent or manage resistance in family therapy. Resistance can be considered as a family characteristic or as a family response to an interaction with the therapist (Szapocznik & Kurtines, 1989; Szapocznik et al., 1989). Hence, at least in family systems thinking, the term "family resistance" fails to capture the systemic nature of the therapist–family relationship (Santisteban et al., 1996). This theoretical work and the empirical research of Patterson and colleagues, Alexander and colleagues, and Szapocznik and colleagues clearly demonstrate that so called family "resistant" behaviors appear to be elicited in interactions with the therapist, and as a result can be modified by altering the therapist's behaviors. For example, Patterson and Forgatch (1985) examined the immediate impact of therapist interventions on client noncompliance in families with a conduct disorder youth. The results of this study showed that therapist teach and confront were associated with increased noncompliance, whereas therapist facilitate and support were associated with decreased noncompliance. Reframing yielded mixed results.

Diamond and Liddle (1996) examined patterns of therapist's interventions and family interactions necessary to resolve therapeutic impasse in multidimensional family therapy with substance abusing youth. Impasse was defined as negative exchanges that thwart the therapist's attempts to facilitate parent–adolescent negotiations. This study demonstrated that families (in general) declined in functioning from the impasse phase to the intervention phase; however, families who successfully resolved the impasse improved their level of functioning following the impasse. Unsuccessful families, on the other hand, were initially more conflicted and pessimistic than successful families, and declined in functioning after the impasse. Results also indicated that therapists played a critical role in resolving parent–adolescent impasses. In successful resolutions, therapists were able to create an emotional treaty among family members by blocking and working through families' negative affect, and by amplifying thoughts and feelings that promoted constructive dialogue. Finally, adolescents became more cooperative and engaged in therapy when parents shifted from trying to control them to trying to understand them.

Taken together these results provide specific recommendations for handling "resistant" behaviors in family therapy. In particular, resistant behaviors must first be reconceptualized as part of an interaction with the therapist. Adopting this view permits the therapist to examine their own contribution to resistance and opens up possibilities for strategically altering their own behavior to influence family members. Therapist facilitative, supportive, and reframing interventions may be particularly useful when "working through" so-called resistant interactions. Likewise, therapist interventions that block negative affect and/or promote constructive dialogue appear to decrease resistant interactions in family therapy.

5.07.6.3.3 The relationship between changes in parenting behaviors and adolescent behavioral outcomes

Schmidt et al. (1996) examined in-session processes in multidimensional family therapy to determine: (i) the nature and extent of change in the behavioral, affective, and cognitive features of parenting; and (ii) the link between parental subsystem changes, and reduction in adolescent substance use and behavior problems. Results demonstrated significant decreases in negative parenting practices and significant increases in positive parenting features. Analyses with adolescent drug use and behavior problem scores indicated a significant relationship between parent improvement and reduction in drug use and acting out. These results demonstrate that not only are parenting behaviors amenable to therapist interventions, but that changes in parenting behaviors in family therapy are associated with changes in relevant adolescent behavioral outcomes. In fact, this relationship was evident in almost 60% of the families in this sample.

5.07.7 SUMMARY AND FUTURE DIRECTIONS

Although FST has emerged as a treatment of choice in the treatment of some child and adolescent clinical populations, the field lacks a solid empirical base demonstrating the efficacy and effectiveness of family systems interventions for a variety of clinical disorders of youth. With the exception of work with behavior problem youth, including substance abuse, family systems treatments have flourished in clinical rather than research settings. Thus, there is an obvious need to bridge the clinical–research gap in FST with childhood disorders.

Research also must be conducted to elucidate the important mechanisms of action in FST. At a molar level, this research should examine universal processes across theoretical models and clinical populations to shed light on what may be the "core" ingredients of FST. However, at a micro level, research must also provide therapists with specific clinical recommendations about how to work with youth and their families. That is, because therapists must respond to family members in the context of their present interactions in therapy, process research can assist clinical model developers in generating or testing the positive or negative impact of particular therapeutic interventions on immediate and long-term outcomes. With respect to immediate and long-term outcomes, such research can also help describe the shape or trajectory of change in FST. The development of new statistical procedures, such as hierarchical linear modeling (Bryk & Raudenbush, 1992; Raudenbush, 1993) and latent growth curve analysis (see Burchinal, Bailey, & Snyder, 1994; Speer & Greenbaum, 1995), will permit research to examine complex clinical theories about the nature of change.

Other challenges face family systems theorists and researchers. For example, family systems therapists are increasingly acknowledging the interface between the family and the environment, and the complexity of factors that influence individual development (Dishion et al., 1995; Henggeler & Borduin, 1990; Szapocznik et al., 1997; Tolan et al., 1995). This recognition has led to the development of a new wave of multisystemic/ecosystemic interventions targeting interactions in (and between) many systems. The fact that most of these interventions have emerged in inner city contexts has illuminated a variety of other challenges that family systems therapists and researchers must address, such as: (i) working with ethnically diverse and the urban poor populations, (ii) confronting the deteriorating neighborhoods in which these families reside, and (iii) challenging conventional methods of service delivery (e.g., location of treatment). Such challenges will require considerable perseverance and ingenuity from the therapists who work in these contexts, and for the researchers who must attempt to shed light on the efficacy and effectiveness of this work and to understand how the change process unfolds.

A final challenge involves the considerable clinical skill and sophistication that is required to implement FST. It should be apparent from the discussion of FST in this chapter, that the therapist is required to recognize the nature of the interactions at every moment in therapy—and to track the evolving process as it unfolds. Moreover, the therapist must be able to intervene by strategically selecting, from a complex set of interventions, the intervention that will most effectively move the family toward the therapeutic goal of the intervention model at that point in time (e.g., engaging, building alliances, reducing negativity, or changing parenting). Traditionally, it has not been fully recognized that the teaching and learning of FST is a rather laborious process that requires education about the key concepts of the family process paradigm and the complex interpersonal skills required to facilitate change in the FST context. Moreover, training also requires an extensive apprenticeship in which new therapists can see families under intensive supervision. Recognition of the challenge of training and learning FST is critical if we will ever be able to transfer research findings into clinical practice. One of the great factors impeding this research to practice translation has been the lack of recognition of the considerable work involved when training and supervising therapists who participate in research projects; training which must be replicated at the service level if we are ever to translate our promising research findings into widespread clinical practice.

ACKNOWLEDGMENTS

This work supported by grants from the National Institute on Drug Abuse (# 1RO1 DA 10574) and the National Institute of Mental Health (# 1RO1 MH 55795) to José Szapocznik, Principal Investigator.

5.07.8 REFERENCES

Alexander, J. F. (1973). Defensive and supportive communications in family systems. *Journal of Marriage and the Family, 35*(4), 613–617.

Alexander, J. F. (1988). Phases of family therapy process: A framework for clinicians and researchers. In L. Wynne (Ed.), *The state of the art in family therapy and research: Controversies and recommendations* (pp. 175–187). New York: Family Process Press.

Alexander, J. F., Barton, C., Schiavo, R. S., & Parsons, B. V. (1976). Systems-behavioral intervention with families of delinquents: Therapist characteristics, family behaviors, and outcome. *Journal of Consulting and Clinical Psychology, 44,* 656–664.

Alexander, J. F., Barton, C., Waldron, H. B., & Mas, C. H. (1983). Beyond the technology of family therapy: The anatomy of intervention model. In K. D. Craig & R. J. McMahon (Eds.), *Advances in clinical behavior therapy* (pp. 48–63). New York: Brunner/Mazel.

Alexander, J. F., Holtzworth-Munroe, A., & Jameson, P. B. (1994). Research on the process and outcome of marriage and family therapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 595–630). New York: Wiley.

Alexander, J. F., & Parsons, B. V. (1973). Short-term behavioral intervention with delinquent families: Impact on family process and recidivism. *Journal of Abnormal Psychology, 81,* 219–225.

Alexander, J. F., & Parsons, B. V. (1982). *Functional family therapy*. Monterey, CA: Brooks/Cole.

Alexander, J. F., Waldron, H. B., Barton, C., & Mas, C. H. (1989). Minimizing blaming attributions and behaviors in delinquent families. *Special series on marital and family disorders. Journal of Consulting and Clinical Psychology, 57*(1), 19–24.

Alexander, P. C. (1985). A systems theory conceptualization of incest. *Family Process, 24,* 79–88.

Azrin, N. H., Donohue, B., Besalel, V. A., Kogan, E. S., & Acierno, R. (1994). Youth drug abuse treatment: A controlled outcome study. *Journal of Child and Adolescent Substance Abuse, 3*, 1–16.

Barton, C., & Alexander, J. F. (1981). Functional family therapy. In A. Gurman & D. Kniskern (Eds.), *Handbook of family therapy* (pp. 403–443). New York: Brunner/Mazel.

Barton, C., Alexander, J. F., Waldron, H., Turner, C. W., & Warburton, J. (1985). Generalizing treatment effects of functional family therapy: Three replications. *The American Journal of Family Therapy, 13*(3), 16–26.

Bateson, G. (1972). *Steps to an ecology of mind*. New York Dutton.

Bateson, G., Jackson, D. D., Haley, J., & Weakland, J. H. (1956). Toward a theory of schizophrenia. *Behavioral Science, 1,* 251–264.

Baumrind, D. (1991a) Effective parenting during the early adolescent transition. In P. A. Cowan & M. Heatherington (Eds.). *Family transitions* (pp. 111–163). Hillsdale, NJ: Erlbaum.

Baumrind, D. (1991b). The influence of parenting style on adolescent competency and substance abuse. *Journal of Early Adolescence, 11*(1), 56–95.

Block, J., Block, J. H., & Keyes, S. (1988). Longitudinally foretelling drug usage in adolescence: Early childhood personality and environmental precursors. *Child Development, 59, 336–355.*

Boivin, M., & Begin, G. (1989). Peer status and self-perception among early elementary school children: The case of the rejected children. *Child Development, 60,* 591–596.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy, 16,* 252–260.

Borduin, C. M., Henggeler, S. W., Blaske, D. M., & Stein, R. J. (1990). Multisystemic treatment of adolescent sexual offenders. *International Journal of Offender Therapy and Comparative Criminology, 34*(2), 105–113.

Bowen, M. A. (1960). A family concept of schizophrenia. In D. D. Jackson (Ed.), *The etiology of schizophrenia* (pp. 341–372). New York: Basic Books.

Breunlin, D. C., Schwartz, R. C., & Kune-Karrer, B. M. (1992). *Transcending the models of family therapy*. San Francisco: Jossey-Bass.

Briggs, J., & David Peat F. (1989). Turbulent mirror: An illustrated guide to chaos theory and the science of wholeness. New York: Harper and Row.

Broderick, C. B. (1993). *Understanding family process: Basics of family systems theory*. Newbury Park, CA: Sage.

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.

Bronfenbrenner, U. (1986). The ecology of the family as a context for human development. *Developmental Psychology, 22,* 723–742.

Bronfenbrenner, U. (1995). Developmental ecology through space and time: A future perspective. In P. Moen, G. H. Elder, & K. Luscher (Eds.), *Examining lives in context: Perspectives on the ecology of human development* (pp. 619–647). Washington, DC: American Psychological Association.

Brooks-Gunn, J., Duncan, G. J., Klebanov, P. K., & Sealand, N. (1993). Do neighborhoods influence child and adolescent development? *American Journal of Sociology*, *99*, 353–395.

Bryk, A. S., Raudenbush, S. W. (1992). *Hierarchical linear models: applications and data analysis methods.* Newbury Park, CA: Sage.

Buckley, W. (Ed.) (1968). *Modern systems research for the behavioral scientist*. Chicago: Aldine.

Burchinal, M. R., Bailey, D. B., & Snyder, P. (1994). Using growth curve analysis to evaluate child change in longitudinal investigations. *Journal of Early Intervention, 18,* 403–423.

Chamberlain, P., Patterson, G. R., Reid, J. B., Kavanagh, K., & Forgatch, M. S. (1984). Observation of client resistance. *Behavior Therapy, 15,* 144–155.

Chamberlain, P., & Rosicky, J. G. (1995). The effectiveness of family therapy in the treatment of adolescents with conduct disorders and delinquency. *Journal of Marital and Family Therapy, 21*(4), 441–460.

Coyne, J. C., & Anderson, B. J. (1989). The "psychosomatic family" reconsidered II: Recalling a defective model and looking ahead. *Journal of Marital and Family Therapy, 15,* 139–148.

Dare, C., Eisler, I., Russell, G. F. M., & Szmukler, G. I. (1990). The clinical and theoretical impact of a controlled trial of family therapy in anorexia nervosa. *Journal of Marital and Family Therapy, 16,* 39–57.

Diamond, G., & Liddle, H. A. (1996). Resolving a therapeutic impasse between a parent and an adolescent in family therapy. *Journal of Consulting and Clinical Psychology, 64*(3), 481–488.

Diamond, G., & Siqueland, L. (1995). Family therapy for the treatment of depressed adolescents. *Psychotherapy, 32*(1), 77–90.

Diamond, G. M., & Siqueland, L. (1996, June). *A family based treatment approach for adolescent major depression: Model and pilot results.* Paper presented at the 27th annual meeting of the Society for Psychotherapy Research, Amelia Island, FL.

Dierick, P., & Lietaer, G. (1990). Member and therapist perceptions of therapeutic factors in therapy and growth groups: Comments on a category system. In G. Lietaer, J. Rombauts, & R. Van Balen (Eds.), *Client-centered and experiential psychotherapy in the nineties* (p. 741–770). Leuven, Belgium: Leuven University Press.

DiGiuseppe, R., Linscott, J., & Jilton, R. (1996). Developing the therapeutic alliance in child–adolescent psychotherapy. *Applied & Preventive Psychology, 5*(2), 85–100.

Dishion, T. J., French, D. C., & Patterson, G. R. (1995). The development and ecology of antisocial behavior. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology. Vol. 2: Risk, disorder and adaptation* (pp. 421–471). New York: Wiley.

Eaton, T. T., Abeles, N., & Gutfreund, M. J. (1988). Therapeutic alliance and outcome: Impact of treatment length and pretreatment symptomatology. *Psychotherapy, 25,* 536–542.

Estrada, A. U., & Pinsof, W. M. (1995). The effectiveness of family therapies for selected behavioral disorders of childhood. *Journal of Marital and Family Therapy, 21*(4), 403–440.

Florsheim, P., Tolan, P. H., & Gorman-Smith, D. (In press). Family processes and risk for externalizing behavior problems among African-American and Hispanic boys. Submitted to *Journal of Consulting and Clinical Psychology.*

Forehand, R. (1993). Family psychopathology and child functioning. *Journal of Child and Family Studies, 2,* 79–85.

Friedlander, M. L., Wildman, J., Heatherington, L., & Skowrow, E. A. (1994). What we do and don't know about the process of family therapy. *Journal of Family Psychology, 8,* 390–416.

Friedman, A. S. (1989). Family therapy vs. parent groups: Effect on adolescent drug abusers. *The American Journal of Family Therapy, 17,* 335–347.

Fromm-Reichmann, F. (1948). Notes on the development of treatment of schizophrenics by psychoanalytic psychotherapy. *Psychiatry, 11,* 263–274.

Garfinkel, P. E., & Garner, D. M. (1982). *Anorexia nervosa: A multidimensional perspective.* New York: Brunner/Mazel.

Gaston, L. (1990). The concept of the alliance and its role in psychotherapy: Theoretical and empirical considerations. *Psychotherapy, 27,* 143–153.

Gaston, L., Marmar, C. R., Gallagher, D., & Thompson, L. W. (1991). Alliance prediction of outcome beyond in-treatment symptomatic change as psychotherapy processes. *Psychotherapy Research, 1,* 104–113.

Gaston, L., & Schneider, J. (1992, August). *The Alliance in Group Psychotherapy.* Paper presented at the meeting of the International Association of Group Psychotherapy, Montreal, Canada.

Gleick, J. (1987). *Chaos: making a new science.* New York: Viking.

Goldenberg, I., & Goldenberg, H. (1983). Historical roots of contemporary family therapy. In B. B. Wolman & G. Stricker (Eds.), *Handbook of family and marital therapy* (pp. 77–89). New York: Plenum.

Goldenberg, I. & Goldenberg, H. (1996). *Family therapy: An overview* (4th ed.). Pacific Grove, CA: Brooks/Cole.

Gomes-Schwartz, B. (1978). Effective ingredients in psychotherapy: Prediction of outcome from process variables. *Journal of Consulting and Clinical Psychology, 46,* 1023–1035.

Gordon, D. A., Arbuthnot, J., Gustafson, K. E., & McGreen, P. (1988). Home-based behavioral-systems family therapy with disadvantaged juvenile delinquents. *The American Journal of Family Therapy, 16*(3), 243–255.

Gordon, D. A., Graves, K., & Arbuthnot, J. (1995). The effect of functional family therapy for delinquents on adult criminal behavior. *Criminal Justice & Behavior, 22*(1), 60–73.

Greenwood, D., Hervis, O. E., Mitrani, V. B., Taylor, D., & Szapocznik, J. (1995, August). *Family definition in low income, HIV-1 seropositive African American women.* Paper presented at the meeting of the American Psychological Association, APA 103rd Annual Convention, New York.

Greenwood, D., Robbins, M. S., Tejeda, M., Samuels, D., Ironson, G., Antoni, M., & Szapocznik, J., (submitted for publication). Family functioning and social support moderate and relationship between average monthly CD4 cell change and psychological distress in HIV-positive African American women. *Journal of Consulting and Clinical Psychology.*

Greenwood, D., Samuels, D., Sorhaindo, L., McIntosh, S., Ironson, G., Antoni, M., Blake, R., & Szapocznik, J. (1995, February). *Social support networks and satisfaction in African American, post-partum, HIV positive women.* Paper presented at the First National Scientific Meeting on HIV infection in Adult and Adolescent Women, Washington, DC.

Grigg, D. N., & Friesen, J. D. (1989). Family patterns associated with anorexia nervosa. *Journal of Marital and Family Therapy, 15,* 29–42.

Guerin, P. J. (1976). Family therapy: The first twenty-five years. In P. J. Guerin (Ed.), *Family therapy: Theory and practice* (pp. 2–22). New York: Gardner.

Gurman, A. S., & Kniskern, D. P. (1978). Research on marital and family therapy: Progress, perspective and prospect. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis* (2nd ed., pp. 817–902). New York: Wiley.

Gurman, A. S., Kniskern, D. P., & Pinsof, W. M. (1986). Research on the process and outcome of marital and family therapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed., pp. 565–624). New York: Wiley.

Gutstein, S. E., Rudd, M. D., Graham, J. C. & Rayha, L. L. (1988). Systemtic crisis intervention as a response to adolescent crises: An outcome study. *Family Process, 27,* 201–211.

Haley, J. (1976). Development of theory: A history of a research project. In C. E. Sluzki & D. C. Ransom (Eds.), *Double bind: The foundation of a communications approach to the family.* New York: Grune & Stratton.

Hanson, B. G. (1995) *General systems theory: Beginning with wholes.* Washington, DC: Taylor & Francis.

Hatcher, R. L., Barends, A., Hansell, J., & Gutfreund, M.

J. (1995). Patients' and therapists' shared and unique views of the therapeutic alliance: An investigation using confirmatory factor analysis in a nested design. *Journal of Consulting and Clinical Psychology, 63,* 636–643.

Hawkins, J. D., Catalano, R. F., & Miller, J. Y. (1992). Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: Implications for substance abuse prevention. *Psychological Bulletin, 112,* 64–105.

Henggeler, S. W., & Borduin, C. W. (1990). *Family therapy and beyond: A multisystemic approach to treating the behavior problems of children and adolescents.* Pacific Grove, CA: Brooks/Cole.

Henggeler, S. W., Borduin, C. M., Melton, G. B., Mann, B. J., Smith, L. A., Hall, J. A., Cone, L., & Fucci, B. R. (1991). Effects of multisystemic therapy on drug use and abuse in serious juvenile offenders: A progress report from two outcome studies. *Family Dynamics Addiction Quarterly, 1,* 40–51.

Henggeler, S. W., Melton, G. M., & Smith, L. A. (1992). Family preservation using multisystemic therapy: An effective alternative to incarcerating serious juvenile offenders. *Journal of Consulting and Clinical Psychology, 60,* 953–961.

Henggeler, S. W., Rodick, J. D., Borduin, C. M., Hanson, C. L., Watson, S. M., & Urey, J. R. (1986). Multisystemic treatment of juvenile offenders: Effects on adolescent behavior and family interaction. *Developmental Psychology, 22,* 132–141.

Henggeler, S. W., Rowland, M. D., Pickrel, S. G., Miller, S. L., Cunningham, P. B., Santos, A. B., Schoenwold, S. K., Randall, E., & Edwards, J. E. (in press). Investigating family based alternatives to institution-based mental health services for youth. Lessons learned from a pilot study of a randomized field trial.

Henry, W. P., Strupp, H. H., Schacht, T. E., & Gaston, L. (1994). Psychodynamic approaches. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 467–508). New York: Wiley.

Hervis, O. E., Szapocznik, J., Mitrani, V. B., Rio, A. T., Kurtines, W. (1991). *Structural family sytstems ratings: a revised manual.* Unpublished document. Miami, FL: University of Miami, Spanish Family Guidance Center, Department of Psychiatry and Behavioral Sciences.

Horvath, A. O., & Luborsky, L. (1993). The role of the therapeutic alliance in psychotherapy. *Journal of Consulting and Clinical Psychology, 64,* 561–573.

Horvath, A. O., & Symonds, D. B. (1991). Relationship between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology, 38,* 139–149.

Jessor, R. (1976). Predicting time of onset of marijuana use: A developmental study of high school youth. *Journal of Consulting and Clinical Psychology, 44,* 125–134.

Joanning, H., Thomas, F., Quinn, W., & Mullen, R. (1992). Treating adolescent drug abuse: A comparison of family systems therapy, group therapy, and family drug education. *Journal of Marital and Family Therapy, 18,* 345–356.

Kandel, D., Kessler, R., & Margulies, R. (1978). Antecedents of adolescent initiation into stages of drug use: A developmental analysis. *Journal of Youth and Adolescents, 7,* 13–40.

Kaslow, F. W. (1980). History of family therapy in the United States: A kaleidoscopic overview. *Marriage and Family Review, 3,* 77–111.

Kazdin, A. E. (1993). Treatment of conduct disorder: Progress and directions in psychotherapy research. *Development and Psychopathology, 5,* 277–310.

Kazdin, A. E. (1994a). Psychotherapy for children and adolescents. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 543–594). New York: Wiley.

Kazdin, A. E. (1995). *Conduct disorders in childhood and adolescence* (2nd ed.). Thousand Oaks, CA: Sage.

Kazdin, A. E., & Kolko, D. J. (1986). Parent psychopathology and family functioning among childhood firesetters. *Journal of Abnormal Child Psychology, 14,* 315–329.

Kazdin, A. E., Stolar, M. J., & Marciano, P. L. (1995). Risk factors for dropping out of treatment among white and black families. *Journal of Family Psychology, 9*(4), 402–417.

Klein, N. C., Alexander, J. F., & Parsons, B. V. (1977). Impact of family systems intervention on recidivism and sibling delinquency: Preliminary findings. *Journal of Consulting and Clinical Psychology, 45,* 469–474.

Kog, E., & Vandereycken, W. (1989) The facts: A review of research data on eating disorder families. In W. Vandereycken, E. Kog, & J. Vanderlinden (Eds.), *The family approach to eating disorders. New York: PMA.*

Kog, E., Vandereycken, W., & Vertommen, H. (1989). Multimethod investigation of eating disorder families. In W. Vandereycken, E. Kog, & J. Vanderlinden (Eds.), *The family approach to eating disorders.* New York: PMA.

Koss, M. P., & Shiang, J. (1994). Research on brief psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 664–700). New York: Wiley.

Kinsley, K., & Bry, B. H. (1991). Sequential analyses of adolescent, mother, and father behaviors in distressed and nondistressed families. *Child and Family Behavior Therapy, 13*(4), 45–62.

Krupnick, J., Sotsky, S., Simmens, S., & Moyer, J. (1992, June). *The role of the therapeutic alliance in psychotherapy and pharmacotherapy outcome: Findings from the NIMH Treatment of Depression Collaborative Research Program.* Paper presented at the annual meeting of the Society for Psychotherapy Research, Berkeley, CA.

Kumpfer, K. L. (1989). Family function factors associated with delinquency. In K. L. Kumpfer (Ed.), *Literature review of effective parenting projects: Strategies for high risk youth and families.* (Research Report No. 2). Washington, DC: Office of Juvenile Justice Delinquency Prevention.

Kurtines, W. & Szapocznik, J. (1996). Family interaction patterns: Structural family therapy within contexts of cultural diversity. In E. D. Hibbs & P. S. Jensen (Eds.), *Psychosocial treatments for child and adolescent disorders: Empirically based strategies for clinical practice.* Washington, DC: American Psychological Association.

Lambert, M. J. (1992). Implications of outcome research for psychotherapy integration. In J. C. Norcross & M. R. Goldstein (Eds.), *Handbook of psychotherapy integration.* New York: Basic Books.

Lambert, M. J., Shapiro, D. A., & Bergin, A. E. (1986). The effectiveness of psychotherapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed., pp. 157–211). New York: Wiley.

Lewis, R. A., Piercy, F. P., Sprenkle, D. H., & Trepper, T. S. (1990). Family based interventions for helping drug-abusing adolescents. *Journal of Adolescent Research, 5,* 82–95.

Liddle, H. A. (1991). Empirical values and the culture of family therapy. *Journal of Marital and Family Therapy, 17,* 327–348.

Liddle, H. A. (1992). A multidimensional model for the adolescent who is abusing drugs and alcohol. In W. Snyder & T. Ooms (Eds.), *Empowering families, helping adolescents: Family-centers treatment of adolescents with alcohol, drug and other mental health problems.* Washington, DC: U.S. Public Health Service, U.S. Government Printing Office.

Liddle, H. A. (1994). The anatomy of emotions in family therapy with adolescents. *Journal of Adolescent Research, 9,* 120–157.

Liddle, H. A. (1995). Conceptual and clinical dimensions of a multidimensional, multisystems engagement strategy in family-based adolescent treatment (Special issue: Adolescent Psychotherapy). *Psychotherapy: Theory, Research and Practice, 32,* 39–58.

Liddle, H. A., & Dakof, G. A. (1994, February). *Effectiveness of family-based treatments for adolescent substance abuse*. Paper presented at the 1994 Society for Psychotherapy Research Conference, Santa Fe, NM.

Liddle, H. A., & Dakof. G. A. (1995a). Efficacy of family therapy for drug abuse: Promising but not definitive. *Journal of Marital and Family Therapy, 21*(4), 511–544.

Liddle, H. A., & Dakof, G. A. (1995b). Family-based treatment for adolescent drug use: State of the science. In E. Rahdert & D. Czechowicz (Eds.), *Adolescent drug abuse: Clinical assessment and therapeutic interventions* (pp. 218–254). NIDA Research Monograph No. 156, NIH Publication No. 95–3908. Rockville, MD: National Institute on Drug Abuse.

Liddle, H. A., Dakof, G. A., & Diamond, G. (1991). Multidimensional family therapy with adolescent substance abuse. In E. Kaufman & P. Kaufman (Eds.), *Family therapy with drug and alcohol abuse* (pp. 120–178). Boston: Allyn & Bacon.

Liddle, H. A., Dakof, G. A., Parker, K., Diamond, G. S., Barrett, K., & Tejeda, M. (in press). Multidimensional family therapy of adolescent substance abuse.

Lidz, T., Cornelison, A., Fleck, S., & Terry, D. (1957). Intrafamilial environment of the schizophrenic patient. I: The father. *Psychiatry, 20,* 329–342.

Lidz, R. W., & Lidz, T. (1949). The family environment of schizophrenic patients. *American Journal of Psychiatry, 106,* 332–345.

Lidz, T., Parker, B., & Cornelison, A. R. (1956). The role of the father in the family environment of the schizophrenic patient. *American Journal of Psychiatry, 113,* 126–132.

Luborsky, L., Barber, J., & Crits-Christoph, P. (1990). Theory-based research for understanding the process of dynamic psychotherapy. *Journal of Consulting and Clinical Psychology, 58,* 281–287.

Magnusson, D. (1995). Individual development: A holistic, integrated model. In P. Moen, G. H. Elder, & K. Luscher (Eds.) *Examining lives in context,* (pp. 19–61). Washington, DC: American Psychological Association.

Madanes, C. (1981). *Strategic family therapy*. San Francisco: Jossey-Bass.

Marcus, M. A., & Wiener, M. (1989). Anorexia nervosa reconceptualized from a psychosocial transactional perspective. *American Journal of Orthopsychiatry, 59,* 346–354.

Margolin, G., Burman, B., & John, R. S. (1989). Home observations of married couples reenacting naturalistic conflicts. Special issue: Coding marital interaction. *Behavioral Assessment, 11*(1), 101–118.

Melidonis, G. G., & Bry, B. H. (1995). Effects of therapist exceptions questions on blaming and positive statements in families with adolescent behavior problems. *Journal of Family Psychology, 9*(4), 451–457.

Minuchin, S. (1974). *Families and family therapy*. Cambridge, MA: Harvard University Press.

Minuchin, S., Baker, L., Rosman, B. L., Liebman, R., Milman, L., & Todd, T. C. (1975). A conceptual model of psychosomatic illness in children. *Archives of General Psychiatry, 32,* 1031–1038.

Minuchin, S., & Fishman, H. C. (1981). *Family therapy techniques*. Cambridge, MA: Harvard University Press.

Minuchin, S., Montalvo, B., Guerney, B. G., Rosman, B. L., & Schumer, F. (1967). *Families of the slums*. New York: Basic Books.

Minuchin, S., Rosman, B. L., & Baker, L. (1978). *Psychosomatic families: Anorexia nervosa in context*. Cambridge, MA: Harvard University Press.

Morris, S., Alexander, J. F., & Waldron, H. (1988). Functional family therapy: Issues in clinical practice. In I. R. H. Falloon (Ed.), *Handbook of behavioral therapy* (pp. 109–127). New York: Guilford.

Muran, J. C., Gorman, B. S., Safran, J. D., Twining, L., Samstag, L. W., & Winston, A. (1995). Linking in-session change to overall outcome in short-term cognitive therapy. *Journal of Consulting and Clinical Psychology, 63,* 651–657.

Nichols, M. P. (1984). *Family therapy: Concepts and methods*. New York: Gardner.

Nichols, M. P., & Schwartz, R. C. (1995). *Family therapy: Concepts and methods* (3rd ed.). Boston: Allyn & Bacon.

Orlinsky, D. E., Grawe, K., & Parks, B. K. (1994). Process and outcome in therapy—noch einmal. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 270–376). New York: Wiley.

Palazzoli-Selvini, M., Boscolo, L., Cecchin, G. F., & Prata, G. (1974). The treatment of children through brief therapy of their parents. *Family Process, 13,* 429–442.

Papero, D. V. (1990). *Bowen family systems theory*. Boston: Allyn & Bacon.

Parker, J. G., & Asher, S. R. (1987). Peer relations and later personal adjustment: Are low-accepted children at risk? *Psychological Bulletin, 102,* 357–389.

Parsons, B. V., & Alexander, J. J. (1973). Short-term family intervention: A therapy outcome study. *Journal of Consulting and Clinical Psychology, 41,* 195–201.

Patterson, G. R. (1971). *Families: Applications of social learning to family life*. Champaign, IL: Research Press.

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia Publishing.

Patterson, G. R., & Dishion, T. J. (1985). Contributions of families and peers to delinquency. *Criminology, 23,* 63–79.

Patterson, G. R., & Forgatch, M. (1985). Therapist behavior as a determinant for client non-compliance: A paradox for the behavior modifier. *Journal of Consulting and Clinical Psychology, 53*(6), 846–851.

Pinsof, W. M., & Catherall, D. R. (1986). The integrative psychotherapy alliance: Family, couple, and individual therapy scales. *Journal of Marital and Family Therapy, 12,* 137–151.

Raudenbush, S. W. (1993). Hierarchical linear models and experimental data. In L. Edwards (Ed.), *Applied analysis of variance in behavioral sciences* (pp. 459–496). New York: Marcel Dekker.

Robbins, M. S., Alexander, J. F., Newell, R. M., & Turner, C. W. (1996). The immediate effect of reframing on client attitude in family therapy. *Journal of Family Psychology, 10*(1), 28–34.

Robbins, M. S., Alexander, J. F., & Turner, C. W. (submitted for publication). The impact of therapist reframing on reciprocal pejorative family member interactions in family therapy with delinquent youth.

Robbins, M. S., Mitrani, V., Zarate, M., Coatsworth, D., & Szapocznik, J. (1996, June). *Linking process to outcome in structural family therapy with drug using youth: An examination of the process of family therapy* in successful and unsuccessful outcome cases. Paper presented at the 27th Annual meeting of the Society for Psychotherapy Research, Amelia Island, FL.

Russell, G. F. M., Szmukler, G. I., Dare, C., & Eisler, I. (1987). An evaluation of family therapy in anorexia nervosa and bulimia nervosa. *Archives of General Psychiatry, 44,* 1047–1056.

Rutter, M. (1980). *Changing youth in a changing society*. Cambridge, MA: Harvard University Press.

Rutter, M. (1983). School effects on pupils progress: Research findings and policy implications. *Child Development, 54,* 1–29.

Rutter, M., & Cox, A. (1985). Other family influences. In M. Rutter & L. Hersov (Eds.), *Child and adolescent psychiatry*. Boston: Blackwell Scientific Publication.

Rutter, M., & Garmezy, N. (1983). Developmental psychopathology. In P. H. Mussen (Ed.) *Handbook of child psychology* (Vol. IV). New York: Wiley.

Rutter, M., & Guiller, H. (1983). *Juvenile delinquency: Trends and perspectives*. New York: Penguin.

Safran, J., Crocker, P., McMain, S., & Murray, P. (1990). Therapeutic alliance rupture as therapy event for empirical investigation. *Psychotherapy, 27,* 154–165.

Santisteban, D. A., Coatsworth, J. D., Perez-Vidal, A., Mitroni, V. B., Jean-Gilles, M., & Szapocznik, J. (1997). Brief structural strategic family therapy with African American and Hispanic high risk youth: A report of outcome. *Journal of Community Psychology, 25*(5), 453–471.

Santisteban, D. A., Szapocznik, J., Perez-Vidal, A., Kurtines, W. M., Coatsworth, J. D., LaPerriere, A., & Lyons-Howie, E. (1996). *The efficacy of brief strategic/structural family therapy in modifying behavior problems and an exploration of the role that family functioning plays in behavior change*.

Schmidt, S., Liddle, H. A., & Dakof, G. (1996). Changes in parenting practices and adolescent drug abuse during multidimensional family therapy. *Journal of Family Psychology, 10*(1), 12–27.

Scopetta, M. A., Szapocznik, J., King, O. E., Ladner, R., Alegre, C., & Tillman, W. S. (1977). *The Spanish Drug Rehabilitation Research Project* (Final Report to NIDA Grant No. H81 DA 01696-03). Miami, FL: University of Miami Spanish Family Guidance Center.

Shields, C. G., Sprenkle, D. G., & Constantine, J. A. (1991). Anatomy of an initial interview: The importance of joining and structuring skills. *The American Journal of Family Therapy, 19*(1), 3–18.

Simcha-Fagan, O., Gersten, J. C., & Langner, T. S. (1986). Early precursors and concurrent correlates of patterns of illicit drug use in adolescence. *Journal of Drug Issues, 16,* 7–28.

Sluzki, C. E. (1978). Marital therapy from a systems theory perspective. In T. J. Paolino & B. S. McCrady (Eds.), *Marriage and marital therapy*. New York: Brunner/Mazel.

Speck, R. V. (1964). Family therapy in the home. *Journal of Marriage and the Family, 26,* 72–76.

Speck, R. V., & Attneave, C. (1973). *Family networks: Rehabilitation and healing*. New York: Pantheon.

Speer, D. C., & Greenbaum, P. E. (1995). Five methods for computing significant individual client change and improvement rates: support for an individual growth curve approach. *Journal of Consulting and Clinical Psychology, 63*(6), 1044–1048.

Stanton, M. D., Todd, T. C., and associates. (1982). *The family therapy of drug abuse and addiction*. New York: Guilford.

Steinberg, L., Darling, N., Fletcher, A. C., Brown, B. B. & Dornbush, S. M. (1995). Authoritative parenting and adolescent adjustment: An ecological journey. In P. Moen, G. H. Elder & K. Luscher (Eds.) *Examining lives in context* Washington, DC: American Psychological Association.

Steinberg, L., Dornbush, S. M., & Brown, B. B. (1992). Ethnic differences in adolescent achievement: An ecological perspective. *American Psychologist, 47,* 723–729.

Steinberg, L. Mounts, N., Lamborn, S. & Dornbusch, S. (1991). Authoritative parenting and adolescent adjustment across varied ecological niches. *Journal of Research on Adolescence, 1*(1), 19–36.

Steinglass, P. (1987). A systems view of family interaction and psychopathology. In T. Jacob (Ed.), *Family interaction and psychopathology: Theories, methods, and findings,* (pp. 25–65). New York: Plenum.

Strupp, H. H., & Hadley, S. W. (1979). Specific versus nonspecific factors in psychotherapy: A controlled study of outcome. *Archives of General Psychiatry, 36,* 1125–1136.

Szapocznik, J., & Coatsworth, J. D. (submitted for publication). Structural ecosystems theory: An ecodevelopmental framework for organizing risk and protection for drug abuse. *American Psychologist*.

Szapocznik, J., & COSSMHO (1994). Structural family therapy. In J. Szapocznik (Ed.), *A Hispanic/Latino family approach to substance abuse prevention* (pp. 41–74). Rockville, MD: Center for Substance Abuse Prevention.

Szapocznik, J., & Kurtines, W. M. (1989). *Breakthroughs in family therapy with drug abusing and problem youth*. New York: Springer.

Szapocznik, J., & Kurtines, W. M. (1993). Family psychology and cultural diversity: Opportunities for theory, research and application. [Invited Article] *American Psychologist, 48*(4), 400–407.

Szapocznik, J., Kurtines, W. M., Foote, F., Perez-Vidal, A., & Hervis, O. E. (1983). Conjoint versus one person family therapy: Some evidence for effectiveness of conducting family therapy through one person. *Journal of Consulting and Clinical Psychology, 51,* 881–899.

Szapocznik, J., Kurtines, W. M., Foote, F., Perez-Vidal, A., & Hervis, O. E. (1986). Conjoint versus one person family therapy: Further evidence for the effectiveness of conducting family therapy through one person. *Journal of Consulting and Clinical Psychology, 54*(3), 395–397.

Szapocznik, J., Kurtines, W. M., Perez-Vidal, A., Hervis, O. E., & Foote, F. (1989). One person family therapy. In R. A. Wells & V. J. Gianetti (Eds.), *Handbook of brief psychotherapies* (pp. 493–510). New York: Plenum.

Szapocznik, J., Kurtines, W., Santisteban, D. A., Pantin, H., Scopetta, M., Mancilla, Y., Aisenberg, S., McIntosh, S., Perez-Vidal, A., & Coatsworth, J. D. (1997). The evolution of structural ecosystemic theory for working with Latino families. In J. Garcia & M. C. Zea (Eds.) *Psychological interventions and research with Latino populations* (pp. 166–189). Boston: Allyn & Bacon.

Szapocznik, J., Perez-Vidal, A., Brickman, A. L., Foote, F. H., Santisteban, D., Hervis, O. & Kurtines, W. M. (1988). Engaging adolescent drug abusers and their families in treatment: A strategic structural systems approach. *Journal of Consulting and Clinical Psychology, 56,* 552–557.

Szapocznik, J., Perez-Vidal, A.., Hervis, O. E., Brickman, A. L., & Kurtines, W. M. (1989). Innovations in family therapy: Strategies for overcoming resistance to treatment. In R. A. Wells & V. J. Giannetti (Eds.), *Handbook of brief psychotherapies* (pp. 93–114). New York: Plenum.

Szapocznik, J., Rio, A. T., Hervis, O. E., Mitrani, V. B., Kurtines, W. M., & Faraci, A. M. (1991). Assessing change in family functioning as a result of treatment: The Structural Family Systems Rating Scale (SFSR). *Journal of Marital and Family Therapy, 17*(3), 295–310.

Szapocznik, J., Rio, A. T., Murray, E., Cohen, R., Scopetta, M. A., Rivas-Vasquez, A., Hervis, O. E., & Posada, V. (1989). Structural family versus psychodynamic child therapy for problematic Hispanic boys. *Journal of Consulting and Clinical Psychology, 57*(5), 571–578.

Szapocznik, J., Santisteban, D., Rio, A., Perez-Vidal, A., & Kurtines, W. M. (1989). Family Effectiveness Training: An intervention to prevent drug abuse and problem behavior in Hispanic adolescents. *Hispanic Journal of Behavioral Sciences, 11*(1), 3–27.

Szapocznik, J., Santisteban, D., Rio, A., Perez-Vidal, A., Kurtines, W. M., & Hervis, O. E. (1986). Bicultural effectiveness training (BET): An intervention modality for families experiencing intergenerational/intercultural conflict. *Hispanic Journal of Behavioral Sciences, 6*(4), 303–330.

Szmukler, G. I., Eisler, I., Russell, G. F. M., & Dare, C. (1985). Anorexia nervosa, parental "expressed emotion" and dropping out of treatment. *British Journal of Psychiatry, 147,* 265–271.

Tolan, P. H., Cromwell, R. E., & Brasswell, M. (1986). Family therapy with delinquents: A critical review of the literature. *Family Process, 25,* 619–650.

Tolan, P. H., Guerra, N. G., & Kendall, P. C. (1995). A developmental-ecological perspective on antisocial behavior in children and adolescents: Toward a unified risk and intervention framework. *Journal of Consulting and Clinical Psychology, 63,* 579–584.

Von Bertalanffy, L. (1968). *General systems theory*. New York: Braziller.

Waterhouse, G. J., & Strupp, H. H. (1984). The patient–therapist relationship: Research from the psychodynamic perspective. *Clinical Psychology Review, 4,* 77–92.

Weinberger, J. (1995). Common factors aren't so common: The common factors dilemma. *Clinical Psychology: Science and Practice, 2,* 45–69.

Weiss, P. A. (1969). The living system: Determinism stratified. In A. Koestler & J. R. Smythies (Eds.), *Beyond reductionism* (pp. 3–42). Boston. Beacom Press.

Wells, R. A., & Dezen, A. E. (1978). The results of family therapy revisited: The nonbehavioral methods. *Family Process, 17,* 251–274.

Werner, E. E., & Smith, R. S. (1992). *Overcoming the odds: High risk children from birth to adulthood.* Ithaca, NY: Cornell University Press.

Widom, C. S. (1989). Does violence beget violence? A critical examination of the literature. *Psychological Bulletin, 106,* 3–28.

Wynne, J. C., Ryckoff, I., Day, J., & Hirsch, S. I. (1958). Pseudo-mutuality in the family relationships of schizophrenics. *Psychiatry, 21,* 205–220.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.08

Parent Training for Child Conduct Problems

CAROLYN WEBSTER-STRATTON and CAROLE HOOVEN
University of Washington, Seattle, WA, USA

5.08.1 INTRODUCTION 186
  5.08.1.1 General Theoretical Underpinnings 186
  5.08.1.2 Contributions from Social Learning Theory 187
  5.08.1.3 Contributions from Cognitive Psychology 187
  5.08.1.4 Contributions from Relational Theories 188
  5.08.1.5 Contributions from Family Systems Theory 188
  5.08.1.6 Emerging Perspectives 188
5.08.2 CHILD CONDUCT PROBLEMS 189
  5.08.2.1 Prevalence 189
  5.08.2.2 Etiological Models 190
    5.08.2.2.1 Biological model 190
    5.08.2.2.2 Attachment model 191
    5.08.2.2.3 Family functioning 192
    5.08.2.2.4 Sociocognitive model 195
    5.08.2.2.5 Environmental model 196
5.08.3 DESCRIPTION OF INTERVENTION PROCEDURES AND PROPOSED MECHANISMS OF CHANGE 197
5.08.4 RESEARCH FINDINGS REGARDING TRADITIONAL PARENT TRAINING 200
  5.08.4.1 Short- and Long-term Follow-up and Program Generalizeability 200
  5.08.4.2 Comparison of PT with Other Types of Family Therapy 201
  5.08.4.3 Factors Contributing to Program Success or Failure 201
    5.08.4.3.1 Parental depression 202
    5.08.4.3.2 Marital conflict 202
    5.08.4.3.3 Insularity and socioeconomic disadvantage 203
    5.08.4.3.4 Child characteristics 204
    5.08.4.3.5 Therapy characteristics 204
    5.08.4.3.6 Therapist characteristics 204
  5.08.4.4 Limitations of Studies Related to Treatment Outcome 205
5.08.5 BROADENING THE FOCUS OF PT 206
  5.08.5.1 PT with Marital and Other Parent-focused Adjuncts 206
  5.08.5.2 PT with Child Training Adjuncts 207
  5.08.5.3 Future Directions for Research and for Practice 208
  5.08.5.4 Expansion Across Multiple Settings 209
  5.08.5.5 Teachers and Other Professionals 209
  5.08.5.6 Addressing Extrafamilial Change 209
  5.08.5.7 Engaging Low-income Families, Understanding Obstacles and the Importance of Culture 210
  5.08.5.8 More Complex Models 212
  5.08.5.9 Understanding Therapeutic Process Variables 212
5.08.6 SUMMARY 213
5.08.7 REFERENCES 213

5.08.1 INTRODUCTION

Prior to 1965, most attempts to treat childhood behavior problems focused exclusively on the child. Approaches included outpatient play therapy and inpatient child therapy. Since the mid-1970s, however, there has been a shift in treatment philosophy for child conduct problems from an exclusive focus on the child to recognition of the primary social context in which the child lives—that is, the family. As a result, parent training (PT) has become an integral part of services for many childhood disorders including autism and developmental disabilities (Schreibman, Kaneko, & Koegel, 1991); academic, learning and language delays (Laski, Charlop, & Schreibman, 1988); externalizing problems such as attention deficit disorder (Anastopoulos, Barkley, & Sheldon, 1996); oppositional and conduct disorders (Forehand, Steffe, Furey, & Walley, 1983; Patterson, Chamberlain, & Reid, 1982; Webster-Stratton, Kolpacoff, & Hollinsworth, 1988); and internalizing disorders such as fears and anxiety disorders (Barrett, Dadds, & Rapee, 1993). PT is also widely used with abused and neglected children as a component of a multifaceted intervention (Lutzker, 1992), as well as for maritally distressed parents (Dadds, Schwartz, & Sanders, 1987) and for parents who are divorced (Grych & Finchman, 1992). This extensive empirical base supports the hypothesis that when parents are trained to implement behavior change strategies, there is a corresponding improvement in their parenting interactions, which in turn results in improvements in children's social and emotional adjustment.

PT is a generic term used herein to refer to forms of intervention for child behavior problems in which parents are trained to work with their children. In this approach parents are not only included in the therapeutic process, they are utilized as cotherapists working under the guidance of the therapist to invoke change in their child's behavior. The training focuses on altering parents' perceptions, attributions, affect, and parenting behaviors toward their children. The term "parent training" is used here in preference to other terms such as "family management training," "parent–child interactional training," or "functional family therapy" to more accurately reflect the fact that in PT it is the parents who are the targets of the treatment and prevention efforts. PT is an empirically supported, therapeutic process based on the twin assumptions that, since parents are the primary socializing agents in their children's lives, parenting practices are important—sometimes the critical determinants—in the development and maintenance of many maladaptive child behaviors and, more importantly, that positive parenting practices are the key to facilitating and maintaining positive changes in child behavior.

There are many versions of PT, but they all share several characteristics in common. First, it is the parents who are trained; there is no direct intervention with the children. Second, parents are taught to identify, observe and define behaviors they want to increase or decrease. Third, parents learn social learning principles including positive reinforcement (e.g., praise and incentive programs) and discipline strategies (e.g., ignoring, loss of privileges, time out). Fourth, they are taught these approaches by means of interactive discussion, modeling, role playing, videotape demonstrations, home assignments, and direct feedback.

A number of theories about the socialization of behavior and emotion have shaped the specific methods, goals, content, and therapeutic process of different PT programs. A parent's function may be formulated as a teacher, coach, model, relational partner, attachment figure, or proximal gatekeeper for the larger environmental conditions; and each shift in emphasis brings a somewhat different focus for PT. Moreover, as current theories of child psychopathology have become more developmental in perspective and ecological in scope, PT has drawn from and integrated a number of theoretical perspectives, from general theories of learning and development to specific research areas and agendas. The evolution of PT from an operant behavioral approach to a model which includes relational, affective, and even broader ecological factors has brought a broader theoretical base, one that incorporates multiple theoretical approaches.

5.08.1.1 General Theoretical Underpinnings

PT draws from and includes theories of learning, theories of therapeutic change, and theories of child social and emotional development. The theoretical mainstays of PT can be, and often have been, subsumed under three general areas: (i) social learning theory, which includes behavioral and cognitive-behavioral views; (ii) "relational theories" such as attachment and psychodynamic theories; and (iii) a family systems perspective which includes the parent–child and other familial relationships. In addition, extensive research in child development has provided PT programs with a framework of empirically-based models of normal and pathological child development.

5.08.1.2 Contributions from Social Learning Theory

Contemporary PT has its roots in applied behavior analysis, models of operant behavior (Baer, Wolf, & Risely, 1968), and social learning theory (Bandura, 1977). A key assumption is that children's behaviors are learned from their interactions with significant persons in their lives, particularly their parents. Child problem behaviors—be they internalizing problems, such as fears and anxieties, or externalizing problems, such as defiance and aggression—are thought to be maintained by environmental reinforcements. The focus of PT from this perspective is on changing maladaptive child behaviors by changing the environmental contingencies which maintain these problem behaviors. For example, to take an internalizing problem such as a social phobia or separation anxiety, research has suggested that family interactions play a role in the development and maintenance of these fears (Kendall, 1992); parents may inadvertently reward anxious behavior by their attention as well as by removing aversive stimuli such as household chores or permitting a child to stay home from school (King, Hamilton, & Ollendick, 1988).

Therefore, this approach emphasized the importance of parents as "behavior change agents" for their children (Patterson, 1982; Patterson, Cobb, & Ray, 1973; Wahler, 1976). Parents were taught specific parenting strategies such as child-directed play, praise and tangible reinforcement, and discipline approaches such as ignoring misbehavior, time out and response cost. These parenting behaviors and strategies helped parents avoid negative reinforcement traps and coercive interactions with their children (Patterson, 1982, pp. 27, 28) as well as positive reinforcement traps (Wahler, 1975) which were thought to lead to escalating child behavior problems. Likewise, this conceptualization has been particularly useful and relevant for child externalizing problems. Research has demonstrated that parents of antisocial children do engage in particular parenting practices which promote aggressive behavior (through attention and compliance to child coercive strategies) and suppress prosocial behavior (by ignoring or even providing aversive consequences) (Hinshaw & Anderson, 1996; Patterson, 1982, 1986).

Social learning theory (Bandura, 1977) posits that behavior is learned not only by experiencing the direct consequences of those behaviors, but also by observing similar behavior and its consequences. Thus, Bandura emphasized the importance of parents "modeling" appropriate social interactions for their children, since children learn by watching their parents interact (with each other and with others), and not just by their responses to the child's own behaviors. Behavioral models are particularly potent when they are persons of higher social status, such as parents, teachers, and older peers, and when their acts are perceived to be rewarded or accepted (Bandura, 1977). Again, the research supports this modeling theory in studies which have shown that children with high levels of fears and anxieties are more likely than nonfearful children to have anxious or fearful parents (Kendall, 1992), as well as studies showing that aggressive children are more likely than nonaggressive children to have parents who use aggressive discipline or who are antisocial themselves (Patterson, DeBaryshe, & Ramsey, 1989).

In accordance with this social learning model, then, PT programs are aimed at helping parents identify and isolate their children's prosocial (or appropriate) and maladaptive (or inappropriate) behaviors, change the reinforcement contingencies, and implement rewards for positive behavior and sanctions for negative behavior. The general purpose of PT from this perspective is to alter the interactions between parents and children so that prosocial behavior rather than coercive behavior is modeled and reinforced by parents (Kazdin, 1987). Social learning theory informs PT's methods as well as content, using live and videotape modeling, role play and rehearsal of new skills, home practice assignments, and direct reinforcement (social and tangible) of parents for their achievements during training.

5.08.1.3 Contributions from Cognitive Psychology

Other theoretical perspectives helped expand our understanding of not only the importance of training parents to change their behavior responses to children but also the importance of influencing parents' perceptions, affect, and cognitions about their children. Cognitive psychology, in particular, focuses on the meanings we make of life events, and how those meanings influence our motivation, behavior, and ability to change. Cognitive events such as attributional biases, self-efficacy beliefs, and negative self-statements influence the course of psychopathology as well as the ability to parent effectively. Bandura proposed that central to parents' ability to change their behavior is their belief about their own self-efficacy (Bandura, 1982)—that is, their belief about the degree to which they can be successful with their children. Attribution theory (Smith & O'Leary, 1995) and learned helplessness theory (Abramson, Seligman, & Teasdale, 1978; Folkman & Lazarus, 1988; Seligman, 1975) are other cognitive theories that posit the importance of beliefs and expectations in influencing parents' interactions with their children.

These expanded theoretical perspectives led to a blending of cognitive and behavioral theories and a concomitant shift in PT known as the cognitive-social learning approach. The cognitive-social learning approach treats cognitions as covert behaviors, subject to the same laws of change as more overt behaviors, and hence appropriately examined, isolated, and differentially reinforced (Meichenbaum & Turk, 1987) as part of the PT process. For example, if parents perceive their children as more powerful than themselves, as motivated and able to hurt them, perhaps even predicting when their child is young that their child will become a juvenile delinquent, those parents are less likely to respond to their child's misbehavior in a positive and consistent manner. Furthermore, parents' negative expectations for a child may become a self-fulfilling prophesy, influencing the child's own self-perceptions and expectations. Parents are often unaware of how their beliefs about their own powerlessness relative to their child can undermine their parenting interactions—resulting in poor limit setting, withdrawal from conflict, lack of structure, and further loss of control. Out of this school of thought have come adjuncts to PT including training for parents in coping via self-statements, problem-solving techniques, stress management, and self-reinforcement strategies.

5.08.1.4 Contributions from Relational Theories

In the category of relational theories attachment theory and psychoanalytic theory are included because of their central concern with emotion, affective processes, and the quality of relationships. Although a more positive parent–child relationship is expected to result from improved parenting strategies, an articulated relational theory did not initially inform PT. Qualitative aspects of parent–child interactions were only indirectly acknowledged. For instance, even in those few behavioral programs which did include training for parents in child-directed play skills, the purpose of this training was to increase the reinforcing power of parental attention; improving relationships was merely a secondary benefit (Forehand & McMahon, 1981). However, as PT expanded beyond its operant focus, child-directed play therapy became a technique and a source of theory concerning the affective and relational aspects of parenting (Eyberg & Boggs, 1989), as separate from behavioral management. Within a relational focus, PT aims to increase parent–child bonding by teaching parents techniques for increasing their positive interactions with their child—as a goal in itself. The emphasis is on increased expression and communication of positive affect, including love, affection, acceptance, enjoyment, and empathy as well as increasing parents' pleasure in their children and enjoyment of play with them. Another aspect to PT that grows out of relational theory is training parents in how to respond to children's expression of emotions as well as how to manage their own emotions. This renewed interest in affective processes reflects a growing recognition from the research that a parent's emotional expression and regulation are likely to affect the quality of their children's emotional language, which in turn affects the quality of their social relationships and ability to self-regulate in the face of conflict.

5.08.1.5 Contributions from Family Systems Theory

Current PT programs have also been influenced by family systems theory. In this theoretical perspective, family factors are recognized as the larger context for parent–child interactions. From a family systems perspective, this means the focus of concern in PT is enlarged to include such factors as family roles, rules, and communication patterns, and those factors are conceptualized as family structures (Minuchin, 1974) and processes (Haley, 1976). A family systems perspective emphasizes the impact of these structures and processes on the parent–-child relationship and behaviors. Inherent in the treatment of families from the systems perspective is an understanding of how family dysfunction (e.g., marital conflict) impacts both parent and child functioning, and PT targets not just parenting behaviors but family interaction patterns (including parent–child, marital, and sibling relationships), family perceptions of individual members, as well as parents' memories of their family of origin and what they learned from that experience.

5.08.1.6 Emerging Perspectives

More recently, in the early 1990s, functional contextualism (Biglan, 1993) has led to our understanding of how social environment—such as stressful life events, socioeconomic status, racism, lack of support systems—influence parents' ability to maintain behavior change. At this point the influence of this perspective on PT programs is still slight, but therapists are beginning to consider factors such as these when designing programs for socioeconomically deprived families and families experiencing considerable life stress.

5.08.2 CHILD CONDUCT PROBLEMS

Despite some evidence suggesting that PT may be a useful treatment for children's internalizing disorders (see Sanders, 1996 for review), this chapter focuses on PT as a treatment approach for young preschool and school-age children with oppositional defiant disorder (ODD) and conduct disorder (CD), defined generically in this paper as conduct problems. We have confined our review of PT to the treatment of conduct problems because there exists a large body of programmatic research regarding the etiology and treatment of conduct problems and because comparatively less is known about the causes of internalizing disorders. Conduct problems frequently co-occur with other disorders such as attention deficit disorder, learning disabilities, language delays, and internalizing problems such as anxiety and somatization. Moreover, children with conduct problems frequently come from families who are experiencing considerable marital discord, depression, and distress. What we know about PT as an intervention for children with conduct problems may provide a paradigm for other less well-researched childhood disorders, including internalizing disorders.

5.08.2.1 Prevalence

Childhood conduct problems are the most frequently occurring disorder in both clinic-referred and general populations (Quay, 1986). Results of epidemiological studies have indicated that the percentage of young children meeting the criteria for the clinical diagnoses of ODD and CD ranges from 7% to 25%, with the prevalence varying according to the population surveyed (Campbell, 1995; Landy & Peters, 1991; Richman, Stevenson, & Graham, 1982). Even for preschoolers, relatively high proportions of parents complain of noncompliance, limited self-control, aggression, and poor relations with siblings and peers (Richman et al., 1982). In a study involving over 400 high-risk families in Seattle, Washington, Webster-Stratton (1995) reported that more than 25% of children scored above the 90th percentile for aggressive behaviors as reported by their mothers. In a large-scale screening study of day-care attendees (2–4 years old) in rural Vermont, Crowther, Bond, and Rolf (1981) found that at least 20% of the children exhibited high frequencies of aggressive and disruptive behaviors, with the greatest severity observed among preschool boys. Boys are two to three times more likely than girls to manifest conduct problems (Quay, 1986).

"Externalizing" behavior problems are characterized by high rates of hyperactivity, aggression, impulsivity, defiance, and noncompliance (Campbell, 1995; Campbell & Ewing, 1990; Robins, 1981). The greatest stability in preschool child behavior problems occurs among externalizing problems in contrast to internalizing disorders, which are age-specific and usually but not always remit over the course of development (Fisher, Rolf, Hasazi, & Cummings, 1984). A variety of studies have shown high continuity between behavior problems in the early preschool years and conduct disorders in adolescence (Egeland, Kalkoske, Gottesman, & Erickson, 1990; Richman et al., 1982; Rose, Rose, & Feldman, 1989). Campbell's (1991) review of a series of longitudinal studies of hard-to-manage preschoolers reveals a surprising convergence of findings. At least 50% of preschool children with moderate to severe externalizing problems continued to show some degree of disturbance at school age, with boys doing more poorly than girls. Of those with continuing behavior problems, 67% met diagnostic criteria for attention deficit hyperactivity disorder (ADHD), ODD, or CD at age 9. Moreover, Eyberg (1992) points out that this percentage may be an underestimate, since many of the most dysfunctional families were lost to follow-up.

Developmental theorists have suggested that there may be two developmental pathways related to conduct disorders: the "early-onset" vs. "late-onset" pathways (Patterson et al., 1989). The hypothesized "early-onset" pathway begins formally with the emergence of aggressive and oppositional disorders (ODD) in the early preschool period, progresses to aggressive (e.g., fighting) and nonaggressive (e.g., lying and stealing) symptoms of conduct disorders in middle childhood, and then develops into the most serious symptoms by adolescence, including interpersonal violence and property violations (Lahey, Loeber, Quay, Frick, & Grimm, 1992). In addition, there is an expansion of the settings in which the problem behaviors occur, from home to daycare or preschool, then to school, and finally to the broader community. For "late-onset" (adolescent-onset) conduct disorders, the prognosis seems more favorable. Adolescents who are most likely to be chronically antisocial are those who first evidenced symptoms of aggressive behaviors in the preschool years, followed by early-onset ODD or CD (White & Bailey, 1990). Children with early-onset ODD also account for a disproportionate share of delinquency in adolescence. Thus ODD is a sensitive predictor of subsequent CD, and the primary developmental pathway for serious conduct disorders in adolescence and adulthood appears to be established in the preschool period (Loeber, 1991).

The next section discusses theories regarding the etiology of conduct problems. These are important in terms of their implications for intervention, including PT.

5.08.2.2 Etiological Models

Theories regarding the causes of child conduct problems posit five different types of influence on child behavior: biological, attachment, family and parent–child interactions, sociocognitive, and environmental. Due to space limitations we will focus primarily on parental and familial models because of their clear implications for PT. However, it should be clear that by limiting our discussion of the biological, attachment, and sociocognitive models we do not mean these should be ignored and should not be enlisted when appropriate nor do we mean to imply that these are necessarily competing models. For instance, research indicates that biological and environmental factors are interrelated and transactional (Caspi & Moffitt, 1995). A more complete review of these models is available from the first author.

5.08.2.2.1 Biological model

Psychobiological research on child conduct disorders has lagged behind similar work on adult criminality and antisocial behavior, but is gaining recognition—particularly research concerning the temperamental, psychophysiological, and neurophysiological aspects of externalizing disorders. Still, many studies suffer from small or incompletely specified samples, lack of diagnostic specificity, and lack of longitudinal follow-up; there is a need for replication and cross-validation. The biological model argues that some abnormal aspect of the child's internal organization at the physiological, neurological, and/or neuropsychological level (which may be genetically transmitted) is at least partially responsible for the development of conduct problems. The most overt evidence for a biological component for children with ODD/CD includes the disproportion of boys to girls and the apparent heritability of antisocial traits, based on the high number of criminals in some families (Hutchings & Mednick, 1977). Adherents of this view have proceeded somewhat cautiously, lest they appear to be endorsing biological determinism. However, since the late 1980s a substantial body of research has been accumulating which supports the inclusion of biological factors in models of development and maintenance of conduct disorder (see Hinshaw & Anderson, 1996 for a review).

Biological explanations for antisocial behavior need not imply that an antisocial personality is biologically predetermined. Even biology is hypothesized to be subject to environmental impact. Biology in itself is only suggestive of a vulnerability to other risk factors (e.g., parental socialization influences) for antisocial behaviors. Furthermore, biological markers of antisocial behavior have thus far been isolated only for subsets of children with more extreme deficits, such as undersocialized aggressive conduct disorder, who are often doubly impaired with ADHD. We do not know what, if anything, biology has to contribute to our understanding of milder deficits. Lahey, Hart, Pliszka, Applegate, and McBurnett (1993) state clearly that investigation of biological variables such as neurotransmitters, brain imaging techniques, skin conductance, or hormonal influences does not imply that psychosocial factors have no role in the maintenance of childhood conduct problems. On the contrary, they postulate "that a socio-environmental event (e.g., abnormal infant experience) could be one of the causes of aggression, but that the effect of this experience on aggression is mediated by alterations in neurotransmitter activity" (Lahey et al., 1993, p. 142). Indeed, the amount of variance accounted for by biological factors appears to be relatively small. Factors such as family support, quality of parent management strategies and socioeconomic status appear to interact with the child's biology or temperament to influence outcome, again suggesting the mediating role of the child's microsystems.

The research findings regarding children with conduct problems who have such difficulties as low autonomic reactivity (Quay, 1993) have implications for how biological deficits may be amplified over time and for how parenting can exacerbate or minimize inherent deficits. For example, children with biological deficits are not as responsive to their parents' normal efforts to praise, reward, set limits, and impose negative consequences; they may even appear to their parents to be indifferent to their discipline attempts. This lack of responsiveness on the part of the child not only makes the child less reinforcing to the parent but also leads parents to feel the effects of their parenting approaches are unpredictable. As parents continue to experience a lack of success in parenting, they develop a low sense of self-efficacy, become more inconsistent in their parenting responses and sometimes resort to increased use of spanking and harsh punishment.

This theory suggests the need for PT programs which help parents understand their children's biological deficits (their unresponsiveness to aversive stimuli and heightened interest in novelty), and support parents in their use of effective parenting approaches so that they can continue to be positive and provide consistent responses. Parents who are concerned that their child is not responsive to praise or appears not to care about consequences can learn that this apparent invulnerability is a coping response for dealing with negative feedback about his or her behavior. The data regarding autonomic underarousal theory suggests that these children will require over teaching (i.e., repeated learning trials) in order to learn to inhibit undesirable behaviors and to manage emotion. They will need consistent, clear, specific limit setting that utilizes simple language and concrete cues and reminders. Their parents will need to rely on tangible reinforcement and novel reward systems, rather than punishment, and will need to help their children anticipate possible consequences. The research regarding these children's possible language and reading delays suggests the need to train parents in ways to help their children with verbal skills such as self-talk management, problem solving, and communication, as well as reading skills.

5.08.2.2.2 Attachment model

Bowlby's (1980) observations of neglected infants introduced us to a new way of understanding children's emotional development. Both his method and thinking have led to attachment theory as a model for describing and accounting for the effects of parent–child relationships. Studies of attachment relationships have focused on mothers and infants; young school-age children have been studied much less frequently. Rarely has the research concerned itself with father/child attachment relationships. Specific research regarding the connections between children's attachment status and the development of conduct problems is very much in its infancy.

Since the late 1980s, methods of measuring attachment beyond toddlerhood have been developed, with two new attachment paradigms targeted for preschool and early primary age children (Cassidy & Marvin, 1992; Main & Cassidy, 1987). Several studies utilizing those measures of attachment with young children have found relationships between insecure attachment and child aggressive behaviors (Greenberg, Speltz, & DeKlyen, 1993). However, attachment status on its own probably does not directly relate to conduct disorder, nor to any specific psychological disorder. The differences between high- and low-risk samples suggest that secure attachment serves as a protective factor in the presence of other risk factors for conduct problems (Dishion, French, & Patterson, 1995).

Finally, in postulating an association between early attachment and psychopathology, one must exercise caution. For attachment status interacts not only with parenting practices, but also emerges as a result of reciprocal interactions with other risk factors such as child biology and temperament, parent psychopathology and social stress. For example, an interaction of biology and attachment may operate as follows: for a certain subset of children with early-onset conduct problems, biology—that is, a difficult temperament—may underlie their difficulties in social learning and communication, both of which may influence mutual acceptance, leading to negative affect and attachment difficulties. In addition, parental depression has been shown to be linked to child attachment status (Radke-Yarrow, 1991; Speiker & Booth, 1988). Furthermore, both attachment and depression are related to parents' parenting skills. For example, in one study both insecurely attached mothers and depressed mothers were likely to retreat when their attempts to limit set with their children were met with resistance; they also showed more inappropriate use of affection. Finally, severe maternal depression, like other severe stressors, appears to be related to disorganized attachment status (Teti, Gelfand, & Messinger, 1995).

The concept of internal working models is helpful for PT programs because it focuses attention on the quality of parent–child relationships and the need for security and predictability in relationships. In the case of the insecurely attached child and parent, the model suggests the need for a more complex PT program involving a focus on relationship building, communications skills, cognitive techniques, and affect management skills. Along with parenting techniques, the insecurely attached parent will need to learn to control negative affect in the parent–child relationship and to communicate positive emotion such as empathy; beyond that, the parent may need help developing empathy and positive feelings for the child. PT can work with parents' perceptions of their child so that parents can recognize how their responses to their child's behavior may in a large part be based on their own past relationships with their parents and not on the child himself (Stroufe & Fleeson, 1986). The therapist's frequent articulation and expression of empathy during training for the parent and child as well as education about child development—specifically, the needs of difficult children—can help parents to understand their child's point of view. This change in perspective can ultimately lead to change in the parents' internal working model of the child, which can then lead to changes in the internal working model of the relationship for both the parent and child. PT for the insecurely attached child must also involve training the parent in attachment-promoting interactions, such as "child-directed play," a method of play which involves listening, following the child's lead, describing and reflecting on the child's actions, and praising and encouraging the child's ideas and behaviors. Daily child-directed play sessions influence the child's working model of the parent as well as the parent's working model of the child. Finally, the attachment perspective on PT suggests that the therapist's role is to provide a secure base for parents who are exploring new perspectives and models of relationships. By being consistent, predictable, empathic, accepting, and reinforcing, the therapist can provide the parent with a new working model, namely that of a supportive relationship.

5.08.2.2.3 Family functioning

Research findings have provided the impetus to expand our models to include other family and contextual variables, reflecting a broader systems view of family influences on children's behavior. Clear connections have been established between parental psychopathology (including antisocial behavior and depression), marital conflict and parenting interactions and child conduct problems.

(i) Antisocial behavior

Intergenerational linkages with respect to criminal behavior have been established for some time (Frick & Jackson, 1993), and there is considerable evidence associating certain types of parental psychopathology with childhood aggression and conduct problems (Patterson, 1982). Specifically, depression in the mother, alcoholism and substance abuse in the father (Frick et al., 1992), and antisocial and aggressive behavior in either the mother or the father (Faraone, Biederman, Keenan, & Tsuang, 1991; Frick, Kuper, Silverthorn, & Cotter, 1995) have been implicated as risk factors. An association has also been found between children's aggression and their parents' aggression at the same age (Huesmann, Eron, Lefkowitz, & Walder, 1984). In a prospective study of 171 clinic boys, parental antisocial personality disorder (APD) was a significant correlate of conduct problems, but the interaction of parental APD and the boys' verbal intelligence predicted the persistence of conduct problems over time (Lahey, Loeber, Hart, & Frick, 1995).

(ii) Depression

Depression—maternal depression in particular—has received a great deal of attention in terms of its influence on child adjustment. Children of depressed parents have been shown to be at increased risk for conduct problems (Downey & Coyne, 1990). Clinic mothers of children with conduct problems consistently rate themselves as more depressed (on the Beck Depression Inventory) than nonclinic mothers of typically developing children, suggesting a relationship between maternal depression and reported child behavior problems (Griest & Wells, 1983; Webster-Stratton & Hammond, 1988). Studies have suggested that depressed mothers make more negative appraisals of their children's behaviors (Schaughency & Lahey, 1985) than their spouses or nondepressed mothers (Webster-Stratton & Hammond, 1988). However, research (using direct observation) evaluating the relationship between depressed and nondepressed mothers' parenting behavior and their children's behavior has yielded conflicting conclusions. Some researchers (Hops et al., 1987; Rickard, Forehand, Wells, Griest, & McMahon, 1981) have found that depressed mothers of clinic-referred children use not more but fewer commands and criticisms than nondepressed mothers, while others have found maternal depression to be related to increased criticisms, spankings, and unsupportive or inconsistent parenting (Cunningham, Bennes, & Siegel, 1988; Webster-Stratton & Hammond, 1988). Still others have found no relationship between maternal depression and parenting behavior (Rogers & Forehand, 1983). Furthermore, the possible linkage between fathers' depression, parenting behaviors, and child conduct problems has not been explored in the research. Finally, it is important to recognize that depression is associated with other stress variables such as poverty, marital distress, and negative life events, making it difficult to determine what effects are attributable to depression *per se*.

The relationship between maternal depression and conduct problems has been conceptualized as follows: maternal depression results in negative appraisals of the child's behavior and contributes to a mother's negative responses (increased irritability, hostility, and punishment) to her child's misbehavior (Patterson, 1982). It has also been theorized that in response to an increase in maternal criticisms, the child displays an increase in noncompliance and other misbehaviors (McMahon & Forehand, 1984; Webster-Stratton & Hammond, 1988). These child misbehaviors further exacerbate the mother's depression, negative appraisals of the child and negative parenting behavior. Not only does the mother reinforce the child's misbehavior (through her negative attention), but the child learns more hostile and irritable behaviors through modeling the behavior of the depressed mother (Downey & Coyne, 1990). It has also been argued that depression contributes to the "coercive" cycle by resulting in decreased maternal attention for positive behaviors, decreased positive reinforcement, and lack of consistent supervision. In other words, depression results in the mother being emotionally unavailable and inattentive to the child. Thus maternal depression may indirectly lead to conduct problems as a result of negative reinforcement of inappropriate child behaviors, inconsistent limit setting, and emotional unavailability.

However, the association between maternal depression and child conduct problems is not adequate justification for concluding that depression "causes" child conduct problems. An alternative model could be postulated reversing the directionality—that is, the child's difficult and unpredictable negative behaviors are minimally reinforcing the mother's efforts, resulting in inconsistent and negative parenting responses; as the mother fails in her efforts to control the child's behavior, she becomes increasingly depressed. The research finding that maternal depression can improve with PT programs suggests that there is a three-way reciprocal interaction between the child's conduct problems, negative parenting style, and maternal depression. And as noted earlier it is highly probable that depression is linked to attachment status and that these two factors operate in some kind of synergy.

(iii) Marital conflict

In addition to depression in the parents, another family characteristic that has been found associated with the development of conduct problems is interparental conflict and divorce (Kazdin, 1987). In particular, boys appear to be more apt to show significant increases in antisocial behavior following divorce. However, some single parents and their children appear to do relatively well over time postseparation, whereas others are chronically depressed and report increased stress levels. Once researchers began to differentiate between parental divorce, separation, and conflict, they began to understand that it was not divorce *per se* that was the critical factor in the child's subsequent behavior and socioemotional development, but rather the degree and intensity of parental conflict (O'Leary & Emery, 1982). For example, children whose parents divorce but whose homes are conflict-free are less likely to have problems than children whose parents stay together but experience a great deal of conflict (McCord, McCord, & Thurber, 1962). Children whose parents divorce and continue to have conflict have more problems than children whose parents experience conflict-free divorce (Hetherington, Cox, & Cox, 1982) Moreover, intact families who seek help for child behavior problems commonly present high levels of marital strife (Johnson & Lovitz, 1974; Oltmanns, Broderick, & O'Leary, 1977). In Webster-Stratton's studies of over 400 families with conduct problem children, 75% of the parents reported having been divorced at least once and/or described their current marriage as distressed; half of the married couples reported experiences with spouse abuse and violence. (Webster-Stratton, 1996b). These findings highlight the role of parents' marital conflict and spousal aggression as key factors influencing children's behavior problems. This is corroborated by the earlier work of Rutter et al. (1974), who reported that marriages characterized by tension and hostility were more closely associated with children's behavior disturbances than marriages characterized as apathetic and indifferent.

The role played by marital conflict, as well as parents' expression of negative affect, is further emphasized in studies of couples with children in laboratory situations requiring interpersonal negotiation or conflict resolution (Grych & Fincham, 1990). Studies suggest that factors such as children's exposure to marital conflict (Grych & Fincham, 1990; Porter & O'Leary, 1980), spousal physical aggression (Jouriles et al., 1991; Jouriles, Murphy, & O'Leary, 1989), and child-rearing disagreements (Dadds et al., 1987) account for variance beyond that of general marital stress in a control sample. Katz and Gottman (1994) found an association between high levels of "mutual contempt" or belligerence in couples' relationships and high levels of anger, physical aggression, and noncompliance in the children of those couples. Cummings and his colleagues have demonstrated that exposure to verbal anger and violence between adults is associated with anger and physical aggression in children (J. S. Cummings, Pellegrinia, Notarius, & E. M. Cummings, 1989). When children become sensitized to conflict as a result of repeated exposure to their parents' fights, they are more prone to emotional and behavioral dysregulation (i.e., greater distress and anger) (Cummings & Zahn-Waxler, 1992). These studies indicate the importance of moving beyond the study of individual risk factors (such as marital dissatisfaction and conflict) to the study of risk mechanisms—that is, the underlying processes that cause the risk factor to lead to conduct problems. The important questions then become: which aspect of conflict carries the risk? and, can it be mediated by an affectionate relationship between parent and child or by enlightened parenting?

However, like depression, marital variables appear to interact reciprocally in complex ways with other family and biological factors. For example, the effects of marital conflict and hostility on conduct problems appear to be mediated by the quality of parenting and degree of parental availability (Cummings, Simpson, & Wilson, 1993). Furthermore, Lahey et al. (1988) discovered that the effects of divorce on conduct problems were almost entirely related to parental diagnosis of antisocial personality disorder. Rutter (1986) has called attention to the relationship between marital discord and parental depression, and the possible causal role of marital discord in child behavioral outcomes. One possible hypothesis is that parents' depression affects their children by influencing marital functioning (increasing marital distress); marital functioning, in turn, affects parenting behaviors and parent–child interactions. An alternative hypothesis is that marital distress increases the likelihood of ineffective parenting; this contributes to child conduct problems, which in turn contribute to parental depression and further marital distress.

Certainly, the data suggest that parental depression covaries with marital distress and anger (Hay, Zahn-Waxler, Cummings, & Iannotti, 1992). Both overt and covert anger between parents are more frequent in homes of depressed parents. Depressed mothers have been shown to have more difficulty resolving conflicts than nondepressed mothers (Kochansaka, Kuczynski, Radke-Yarrow, & Welsh, 1987). Instead of attempting to reach compromises as a means of ending conflicts, depressed mothers have been reported to withdraw. However, impaired conflict resolution skills are not limited to the depressed mother. High levels of negative affect and hostility in mothers or fathers certainly can disrupt a couple's ability to resolve conflicts. Moreover, it is further hypothesized that a couple's chronic failure to resolve conflicts leads to increasing negative affect in both parents, which may well contribute to escalating anger in parenting as well as marital interactions. In turn, children who are repeatedly exposed to their parents' negative affect and inability to resolve issues are said to be sensitized, whereby their negative reactions are intensified (E. M. Cummings & Davies, 1994).

A number of investigators are postulating that marital distress (conflict, hostility, etc.) is the primary mediator in the transmission of conduct problems, rather than depression *per se* (Cox, Puckering, Pound, & Mills, 1987). On the basis of their extensive review of the literature on children with depressed parents, Downey and Coyne (1990, p. 68) concluded that "marital discord is a viable alternative explanation for the general adjustment difficulties of children with a depressed parent" (see also Cummings & Davies, 1994). Nonetheless, the precise mechanisms underlying the relationship between marital conflict, depression, parenting, and child adjustment have been little investigated and are poorly understood. This is partially due to the fact that research has consisted largely of global self-report measures of marital satisfaction (or of depression) rather than measures of specific areas of disagreement and close observation of communication and affect. The prediction of child outcomes may be improved substantially by greater specification of how conflict is expressed within families. In a study of 120 families involving detailed observation of marital communication and problem-solving as well as observation of children's interactions with a best friend, Webster-Stratton (1996b) reported a correlation between, on the one hand, high levels of collaboration, engagement, and positive communication in couples and, on the other hand, lower levels of child noncompliance and externalizing problems at home, as well as a correlation between higher levels of positive social skills and conflict management skills and lower levels of negative interactions with peers. Alternatively, negative marital communication, low levels of communication, and poor parental problem-solving between couples were correlated with children's conduct problems at home and their poor conflict management skills with peers. These data provide some support for the notion that a couple's communication patterns in general—and their conflict-resolution style in particular—may be a key variable in the development of their child's conduct problems and peer relationships. This dovetails with the work of Cummings and colleagues (Cummings, Ballard, El-Sheikh, & Lake, 1991; Cummings & Davies, 1994), who found that helping families resolve their conflicts and providing children with explanations of their conflicts (i.e., absolving children from any blame) is helpful for children's emotional adjustment. They conclude that parents' successful resolution of conflicts results in a significant reduction in their children's anger and distress.

(iv) Parenting interactions

Parenting interactions are clearly the most well-researched and most important proximal cause of conduct problems. Research has indicated that parents of children diagnosed as ODD/CD lack certain fundamental parenting or behavior management skills. For example, those parents have been reported to exhibit fewer positive behaviors, to be more violent and critical in their use of discipline, to be more permissive, erratic, and inconsistent, to be more likely to fail to monitor or supervise their children's behaviors, and to be more likely to reinforce inappropriate behaviors and to ignore or punish prosocial behaviors (Patterson, 1982). The most influential developmental model for describing the family dynamics that underlie early antisocial behavior is Patterson's theory of the "coercive process" (Patterson, 1982), a process whereby children learn to escape or avoid parental criticism by escalating their negative behaviors, which in turn leads to increasingly aversive parent interactions. These negative responses, in turn, directly reinforce the child's deviant behaviors. Such mutual training in aversive responding intensifies both the child's aggressive behavior and the parents' hostile, nonresponsive behavior. These findings point also to the importance of the affective nature of the parent–child relationship. There is considerable evidence that a warm, positive bond between parent and child leads to more positive communication and parenting strategies and a more socially competent child (Baumrind, 1971).

Research on family socialization related to aggression increasingly recognizes the bi-directionality of child and parent behavior (Lytton, 1990). It is conceivable that negative parenting behavior is in part a reaction to difficult, oppositional, aggressive child behavior. In a study by Anderson, Lytton, and Romney (1986), mothers of boys with CD were compared with another group of mothers interacting with their own sons, with another boy who had a diagnosis of CD, and with a boy who had no diagnosis of CD. Mothers in both groups were more negative when interacting with the boys with CD, supporting the child-to-parent model. However, mothers of sons with CD were the most negative, suggesting that a history of negative interactions plays an important role (Anderson et al., 1986). The most conclusive evidence for the causal role of parenting interactions in promoting aggressive behavior arises from the research on PT interventions designed to decrease negative parenting behavior. Randomized studies have consistently shown that the risk for conduct problems decreases in groups of parents who received training in more effective parenting skills. These studies will be reviewed in a subsequent section of this paper.

This research regarding family functioning and parent–child interactions argues for parent training programs which attend not only to parenting skills but also to the other interpersonal and relational difficulties of family members. While positive parenting may buffer some of the effects of marital conflict, antisocial behavior, and depression, it would seem logical that if the intervention could also address parental psychopathology and reduce marital conflict, then there might be a greater likelihood for longer-lasting results from intervention.

5.08.2.2.4 Sociocognitive model

One way in which biological and family factors may exert effects on children's behavior problems is through the child's perceptions and evaluations of his or her social circumstances. The seminal work of Dodge and his colleagues has been extraordinarily helpful in suggesting ways in which children process and encode their social experiences (Crick & Dodge, 1994). A series of studies has revealed that aggressive children display deficits in social problem-solving skills (Asarnov & Callan, 1985), define problems in hostile ways, seek less information, generate fewer alternative solutions to social problems, and anticipate fewer consequences for aggression (Rubin & Krasnor, 1986; Slaby & Guerra, 1988). Aggressive behavior in children is correlated with their low empathy for others across a wide age range (Feshbach, 1989). It has also been suggested that children with conduct disorders distort social cues during peer interactions (Milich & Dodge, 1984), including attributing hostile intent to neutral situations. Aggressive children search for fewer cues or facts (i.e., they underutilize cues) when determining another's intentions (Dodge & Newman, 1981) and focus more on aggressive cues (Goutz, 1981). There is also a suggestion in the literature that attributional distortions and underutilization of cues pertain specifically to the subgroup of aggressive children with comorbid ADHD (Milich & Dodge, 1984). It is theorized that the impulsive cognitive style of these children limits their scanning of pertinent social cues before responding, so that they have difficulty perceiving or understanding another person's point of view and are unable to interpret interpersonal situations accurately. This would explain both their lack of social competence and their antisocial behavior. Negative exchanges with peers or with parents then contribute to rejection, further negative responses, and further negative attributions.

These findings argue for the need for PT programs to include ways that parents can teach their children more appropriate social skills as well as how to problem-solve, to control angry responses, and to anticipate negative consequences of aggression. They need to help their children learn to reframe negative appraisals, to detect and express their feelings appropriately, and to learn to take the perspective of another—i.e., develop empathy skills. Parents also need help finding ways to foster more positive peer relationships for their children, such as by inviting friends over and setting up structured home activities, by enrolling them in appropriate community programs, and so forth. These data also suggest the need for PT to be combined with child training (CT), a process whereby children are directly trained in social skills and problem-solving by teachers, school counselors, and parents.

5.08.2.2.5 Environmental model

(i) Insularity and lack of support

Parents' relationships with others outside the family have also been linked to child conduct problems. Studies have assessed both the quantity and quality of maternal contacts outside the home in clinic-referred families of children with conduct problems. Wahler (1980) has used the term "insular," defined as "a specific pattern of social contacts within the community that are characterized by a high level of negatively perceived social interchanges with relatives and/or helping agency representatives and by a low level of positively perceived supported interchanges with friends" (Wahler & Dumas, 1984), to characterize a subgroup of mothers of children with conduct problems. These mothers have been observed to be more aversive and use more aversive consequences with their children than "noninsular" mothers (Webster-Stratton, 1985a, 1985b). Wahler found that on days when mothers reported a high number of contacts with friends (days we might characterize as "noninsular"), maternal aversive behavior and oppositional child behavior was consistently lower than on days when the number of contacts with friends was low (Wahler, 1980).

"Insular" families feel isolated not only from other parents in their communities, but also from their children's schools. They report frequent negative encounters with teachers concerning their children's behavior problems. Such encounters only add to parents' feelings of incompetence, their sense of helplessness regarding appropriate strategies to solve the problems, and their alienation from the school. Parents' inability to collaborate successfully with teachers (and thereby to reinforce them) contributes to teachers' perceptions of parents as uninterested or uncaring in regard to their children's problems. Such blaming attitudes on the part of teachers further escalates parents' sense of isolation and stigmatization regarding their child's problems. The child, in turn, is constantly observing his or her parents' mounting frustration and inability to communicate, as well as his or her teachers' lack of commitment to his family and his learning. This spiraling pattern of child negative behavior, parent demoralization and withdrawal, and teacher reactivity disrupts the connection between home and school and produces a dissonance between the socialization activities of the school and home, further contributing to the child's disengagement from the teacher and from school, and making eventual school drop out a likelihood.

(ii) Socioeconomic disadvantage

Isolation covaries with other important environmental factors that adversely affect family functioning, including economic disadvantage and an aggregate of related risk factors: unemployment, perceived stress, low education, and number of stressful life events. These have been shown to have negative effects on parenting, including the development of abusive disciplinary practices, and on child behavior, including early-onset conduct problems (Hawkins, Catalano, & Miller, 1992; Kazdin & Kolko, 1986; Rutter & Giller, 1983). Low-income parents are more at risk than are middle-income parents for high levels of psychological stress (Gecas, 1979), "power assertive" (i.e., coercive) discipline strategies (Daro, 1988; Trickett, Aber, Carlson, & Cicchetti, 1991), and a tendency to rationalize as legitimate even inappropriate parenting decisions (Hoffman, 1984). It has been argued that "negative parenting behavior is the primary pathway through which poverty undermines children's socioemotional functioning" (McLoyd, 1990). Clearly, child abuse occurs at all income levels; however, there is a higher incidence of physical abuse in families below the poverty line (Straus & Gelles, 1986), even when reporting bias is taken into account. Offord and colleagues found that low income was one of the most significant risk factors for conduct disorder, but that it had its effect on young children, not on adolescents (Offord, Alder, & Boyle, 1986). Moreover, they showed that rates of conduct disorder were significantly higher for children from low-income welfare families than for children from low-income nonwelfare families (Offord, Boyle, & Szatmari, 1987).

Families of children with conduct problems report major life stressors at a rate two to four times greater than families with "normal" children (Webster-Stratton, 1990b). Moreover, parents of children with conduct problems indicate that they experience more day-to-day difficulties than nonclinic families. An accumulation of minor but chronic day-to-day stressors appears to disrupt parenting behaviors and leads to increased rates of coercive behavior and irritability in parents' interactions with their children (Forgatch, Patterson, & Skinner, 1988). Additionally, parents who are depressed, stressed, or demoralized are less likely to be able to provide the cognitive stimulation, emotional support, and social learning necessary to foster a child's positive behavior. Wahler and Sansbury (1990) proposed that highly stressed mothers are less able to screen out extraneous information (e.g., negative interchanges with neighbors or spouse), causing them to react inconsistently or indiscriminately to their children's behavior.

A key issue is whether these environmental factors directly contribute to child behavior problems or whether their effects are mediated by variables such as parenting skills. A synthesis by Capaldi and Patterson (1994) examined a wide array of factors for their predictive relationships to conduct disorders, testing for direct and indirect effects. They found that unemployment, low socioeconomic status, multiple transitions, and high levels of family adversity were related to early-onset conduct problems, but not to late-onset; however, the relationship was less strong when parenting variables were taken into account. The direct effects of low socioeconomic status, in particular, were erased when parenting variables were included; the roles of family transitions, stress, and unemployment also appeared to be indirect. They proposed that environmental variables were related to conduct disorders in a "chain reaction" fashion, whereby unemployment, for example, increased the family stress and the numbers of family transitions, which in turn decreased parental involvement and monitoring and increased the amount of "coercive" parenting. In another study, it was shown that family variables served as a protective factor against risk incurred by high-frequency encounters with significant violence in the neighborhood (Richters & Martinez, 1993). Thus it would seem that the effects of environmental factors on conduct problems are primarily indirect, not direct. However, this is not to discount their importance in explaining the development of conduct problems or their potential impact on parents' ability to sustain consistent and positive parenting.

Again, we see the powerful buffering effects of parenting even in situations of poverty and high levels of stress. But these data also seem to suggest that PT programs which help parents cope with environmental stressors will be more effective and far reaching.

This discussion of etiology relating to conduct problems is concluded by arguing for a complex theoretical model which considers the strong likelihood of transactional, reciprocal relationships among the biological, psychological, familial, sociocognitive, and environmental factors influencing child behavior. In all likelihood, antisocial behavior is not the sum of a certain number of risk factors, but the result of the interaction of multiple risk factors (Hinshaw & Anderson, 1996). Indeed some variables, for example, depression, may be consequences rather than precursors of conduct problems. Furthermore, the interplay of risk is particularly potent for "early-onset" conduct disorder, the course of which leads to more serious delinquency (Quay, 1986). Such a complex model will necessitate a family intervention which takes into consideration all of these intertwining factors.

5.08.3 DESCRIPTION OF INTERVENTION PROCEDURES AND PROPOSED MECHANISMS OF CHANGE

This review, emphasizes interventions geared towards younger preschool and school-aged children (as opposed to adolescents) because of the emerging evidence that children who begin antisocial behavior early are at significantly greater risk than those who become antisocial later, both for chronic offending during adolescence and for careers as antisocial adults (Patterson et al., 1989). The following discussion highlights several PT programs which were selected on the basis of their widespread availability, detailed descriptions of training procedures, and extensive evaluation of both the short-term and long-term effectiveness of the intervention.

The most highly influential PT program was developed by Patterson, Reid, and their colleagues at the Oregon Social Learning Center (Patterson et al., 1982; Patterson, Reid, Jones, & Conger, 1975). Spanning two decades of research with more than 200 families, their work provides an exemplary model for outcome research with conduct-problem children. Although directed toward parents of preadolescent and adolescent children who were engaged in overt conduct disorders, their program will be described here because it has provided the foundation for numerous other PT programs.

The program starts with having parents read a programmed text, either *Living with children* (Patterson, 1976) or *Families* (Patterson, 1975). Afterwards, they complete a test on the reading material. The therapist then works with each parent individually in a step-by-step approach wherein each newly learned skill forms the foundation for the next new skill. Five behavior management practices form the core content of the program. First, parents are taught to pinpoint the problem behaviors of concern and to track them at home (e.g., compliance versus noncompliance). Second, they are taught how to use social and tangible reinforcement techniques (e.g., praise, point systems, privileges, treats) and to shift from tangible to social reinforcers over time. Third, they are taught discipline procedures: 5-minute time-outs, short-term privilege removal, response cost, and work chores. Fourth, they are taught to "monitor'—that is, to provide close supervision for their children even when the children are away from home. This involves parents knowing where their children are at all times, what they are doing, and when they will be home. In the final phase of treatment, parents are taught problem-solving and negotiation strategies and become increasingly responsible for designing their own programs. The treatment content has been described in a manual by Patterson (1975) and elaborated upon by Reid (1987).

Like all PT programs, this intervention is based on a model which posits the primacy of ineffective parenting skills in the development of conduct problems: parents inadvertently teach their children noncompliance and aggression by modeling and reinforcing those behaviors in their daily interactions with their children. Here, the focus of intervention is the particular parent behaviors which support the child's negative behaviors. For example, instead of backing down (and rewarding) their child's coercive response to commands, or using harsh punitive parenting to gain compliance, parents are taught to follow through with commands in a reasonable and consistent manner. Instead of ignoring prosocial efforts and attending to negative behaviors, they are taught to reward appropriate behavior that they had hitherto overlooked.

Another influential PT program was developed originally by Hanf (1970) at the University of Oregon Medical School and later modified and evaluated extensively by McMahon and Forehand (1984). It was designed to treat noncompliance in young children, ages 3–8 years. As described by Forehand and McMahon (1981) in their book, *Helping the noncompliant child,* the first phase of this comprehensive PT program involves teaching parents how to play with their children in a nondirective way and how to identify and reward children's prosocial behaviors through praise and attention. The objective is for parents to learn to break the coercive cycle by increasing their social rewards and attention for positive behaviors and reducing their commands, questions, and criticisms. Parents also learn to use social and tangible rewards for child compliance and to ignore inappropriate behaviors. The second phase of the program involves teaching parents how to give direct commands in such a way as to gain more compliance and how to use 3-minute time-outs for noncompliance. Progression to each new skill in the treatment program is contingent on the parent's achieving an acceptable degree of competence in the previously presented skill.

The program is conducted in a clinic setting where the therapist works with individual parents and children together. Treatment methods include role-playing, modeling, and coaching. The clinic utilizes a playroom equipped with one-way mirrors for observation and "bug-in-the-ear" devices through which the therapist can prompt and give feedback to parents while they play with their child. Homework is assigned in the form of daily 10-minute play sessions with the child using the strategies learned in the clinic.

In this intervention the focus is on the reinforcement value of parental attention. Parents are taught to use "descriptive commenting" with appropriate behaviors—that is, to describe their children's behavior when they are acting in appropriate, positive ways—and to praise those behaviors. Parental attention reinforces and thus promotes the replacement behaviors, the positive behaviors which serve to replace the negative behaviors. Only after mastering the use of positive attention are parents taught discipline techniques to limit negative behavior. Reorienting the pattern of parent–child interactions from negative to positive reinforcement is presumed to interrupt the overlearned negative cycles in which many parents of noncompliant children find themselves, and which only promote further noncompliance.

The emphasis of this parent training (Hanf, 1970) on the relational aspects is also found in the intervention developed by Eyberg (1988) called "parent–child interaction therapy" (Hembree-Kigin & McNeil, 1995). While the emphasis on behavior management is maintained, the skills for child-directed play are elaborated in great detail, composed of "DRIP skills": describe, reflect, imitate, praise. Eyberg (1988) presents this program as an integration of traditional play skill values and current behavioral thinking about child management (Eyberg & Boggs, 1989). It is felt that parents' nondirective play with their children improves children's frustration tolerance, helps reduce the anger level of oppositional children, and offers more opportunities for prosocial behavior to occur (Hembree-Kigin & McNeil, 1995). Moreover, engaging in play with their children helps parents recognize their children's positive qualities. As parents learn nondirective play skills, they learn how to respond in a sensitive and genuine manner, how to relate to their child's level of development, and how to stimulate their learning. The primary goal of this training is to strengthen attachment and establish a warm, loving relationship between the parent and child.

Another example of a PT program for young children with conduct problems was developed by Webster-Stratton (1981b, 1982a, 1982b, 1984). Based on the early theoretical work of Patterson (1975) and Hanf (1970) regarding key parenting and relationship skills and behavioral principles to be learned in order to reduce conduct problems, the program utilizes videotape modeling methods. The content of the BASIC program incorporates Patterson's (1982) nonviolent discipline components concerning time-out, logical and natural consequences, and monitoring, components of Hanf's (1970) "child-directed play" approaches and the strategic use of differential-attention, encouragement and praise, and effective use of commands. This content has been embedded in a relational framework including parent group support, mutual problem-solving, self-management, and a collaborative relationship with the therapist. This approach is designed to promote parental self-efficacy and engagement with the program and to reduce parental resistance and drop-out (Webster-Stratton & Hancock, in press). Parents are taught to examine and alter the irrational thinking which derails their implementation of effective parenting. The ADVANCE PT program includes content on problem-solving, communication (D'Zurilla & Nezu, 1982; Spivak, Platt, & Shure, 1976), and self-control techniques. In 1990, another new program was developed, entitled *Supporting your child's education*, to address risk factors related to children's poor academic readiness and weak home/school connection (Webster-Stratton, 1992b). In this program, parents learn how to prepare their preschool children so they can be successful in school. (Webster-Stratton and colleagues have developed a videotape program for teachers based on the same principles; teachers are trained in strategies for managing classroom behavior and for working collaboratively with parents.)

This program of research has been concerned with methods of presenting the parenting program—that is, developing the most cost-effective, widely applicable, and sustaining methods for training parents. Based on Bandura's (1982) modeling theory, the program utilizes videotape-modeling methods. The series of 10 videotape programs of modeled parenting skills (250 vignettes, each of which lasts approximately 1–2 minutes) are shown by a therapist to groups of parents (8–12 parents per group) in 13 sessions (26 hours). The vignettes show parent models in natural situations (unrehearsed) with their children "doing it right" and "doing it wrong" in order to demystify the notion that there is "perfect parenting" and to illustrate how one can learn from one's mistakes. After each vignette, the therapist leads a group discussion of the relevant interactions and encourages parents' ideas. Vignettes of parenting mistakes provide a wonderful opportunity to engage parents in analyses and critical thinking about what makes one technique effective while another is not. Group discussion provides a way for parents to gain acknowledgment for their thoughts and experiences. It also provides a safe forum for disclosure of past mistakes as well as fears, regrets, misgivings, etc. Therapists have the opportunity to model positive self-talk as they discuss vignettes or relate similar experiences, and such talk encourages parents to refrain from self-blame about their own parenting mistakes, and instead to believe in their future successes. Efforts are made to promote the modeling effects for parents by encouraging identification with the models shown on the videotapes; the videotapes show parents of differing sexes, ages, cultures, socioeconomic backgrounds, and temperaments demonstrating different levels of parenting skill, so that parents will perceive the models as similar to themselves and their children. Home activities are given out each week which include daily practice exercises as well as a weekly chapter to read or listen to on audiotape from the book, *The incredible years: A trouble-shooting guide for parents* (Webster-Stratton, 1992a).

The group discussion and collaborative format were chosen to ensure that the intervention would be sensitive to individual cultural differences and personal values, as well as to enhance parents' commitment to parental self-management. The program is also designed to help parents understand and learn to accept normal variation in children's developmental abilities, emotional reactions, and temperament style. The self-management aspect of the program is emphasized by asking each parent to identify the particular positive behaviors they want to see more of in their children and the negative behaviors they want to decrease. These targeted behaviors then become the focus for them to apply the "parenting principles" which they learn in the program. In this sense the program is "tailored" to each family's individual needs and goals as well as to each child's abilities.

The program utilizes group process not only because it is more cost effective but also because it addresses an important family risk factor—namely, the family's sense of isolation and stigmatization. The parent groups enhance and promote parent support networks and mutual collaboration (Webster-Stratton & Herbert, 1993, 1994). This support is also extended during the week by means of weekly assignments which include "buddy calls," a process whereby parents call each other to share their experiences regarding the assignments. However, the program has also been given to over 80 parents of conduct problem children as a completely self-administered intervention—that is, parents view the videotapes and complete the homework assignments without therapist feedback or group support.

5.08.4 RESEARCH FINDINGS REGARDING TRADITIONAL PARENT TRAINING

5.08.4.1 Short- and Long-term Follow-up and Program Generalizeability

The use of PT as an intervention for child conduct problems has been extensively researched, and there are a number of excellent reviews (Henggler, Borduin, & Mann, 1993; Kazdin, 1987; Miller & Prinz, 1990; Patterson, Dishion, & Chamberlain, 1993). Programs have reported high parental ratings of acceptability and consumer satisfaction (Cross Calvert & McMahon, 1987; McMahon & Forehand, 1984; Webster-Stratton, 1989). The success of short-term treatment outcome has been verified by significant changes in parents' and children's behavior and in parental perceptions of child adjustment (Kazdin, 1985; McMahon & Forehand, 1984; Patterson & Reid, 1973; Webster-Stratton, 1981a, 1982a, 1984, 1990a; Webster-Stratton, Hollinsworth, & Kolpacoff, 1989) in comparison to waiting-list control families. Home observations have indicated that parents who have undergone training have been successful in reducing children's levels of aggression by 20–60% (Patterson et al., 1982; Webster-Stratton, 1985b). While the majority of studies have been conducted with Caucasian mothers, there is some evidence that parent training is also effective with fathers (Webster-Stratton, 1985a, 1990a) and with ethnic minorities (Strayhorn & Weidman, 1989; Webster-Stratton, 1995).

Generalization of behavior improvements has been demonstrated from the clinic setting to the home (Patterson & Fleischman, 1979; Peed, Roberts, & Forehand, 1977; Webster-Stratton, 1984) over reasonable follow-up periods (1–4 years) and to untreated child behaviors (Arnold, Levine, & Patterson, 1975; Fleischman, 1981; Forehand & Long, 1986; Webster-Stratton, 1982a; Webster-Stratton & Hammond, 1990). A 14 year follow-up of parents who attended the Forehand and McMahon program indicated that treated families were performing comparably to non-referred "normative" samples on measures of internalizing and externalizing behaviors, social competence, relationship with parents and academic progress. Unfortunately, the data are mixed regarding generalization of child behavior improvements from the home to the school setting.

Several studies have found that even when a child's behavior improved at home, his or her teacher did not necessarily report improvements in conduct problems and peer relationships in the classroom (Breiner & Forehand, 1982; Forehand, Breiner, McMahon, & Davies, 1981; Forehand, Rogers, McMahon, Wells, & Griest, 1981; Forehand et al., 1979). The Webster-Stratton (1981a) program found significant improvements in child adjustment at school as reported by teachers immediately post-treatment, but a year later these were not maintained. In another randomized study of 43 aggressive boys (grades 2–6), when treatment families were compared with a waiting-list control group, they showed significant increases in family cohesion, empathy, problem-solving efficiency, total family relationships, and positive child behaviors at home and school, with corresponding decreases in family conflict and negative/aggressive child behaviors at home and at school (Sayger, Horne, Walker, & Passmore, 1988). In a follow-up study of this sample, Sayger, Horne, and Glaser (1993) found that these positive changes in behavior at school, in marital satisfaction, in maternal depression, and in family relationships in general were maintained.

5.08.4.2 Comparison of PT with Other Types of Family Therapy

A few studies have compared the basic PT approach with other forms of family therapy. Patterson compared families randomly assigned to his PT program with those who received treatment from various clinicians in the community representing a range of therapeutic approaches. Results indicated the superiority of PT in producing reductions in child deviance (Patterson et al., 1982). In a subsequent study, Patterson and Chamberlain (1988) randomly assigned parents to their PT program or to a community agency employing eclectic family therapy. Findings indicated significant reductions in child deviant behaviors for PT, but no significant reductions for children in the family therapy condition. Wells and Egan (1988) also evaluated a study comparing PT with family systems therapy. Results indicated that their program was more effective than family systems therapy based on observations of parent and child behaviors, but both programs showed significant decreases in child problem behaviors as reported by parents.

In 1996, Taylor, Schmidt, and Hodgins (1996) reported on a study where they randomly assigned 110 families to a waiting-list control group, to Webster-Stratton's PT program (videotape modeling therapist-led group approach), or to a children's mental health center whose usual treatment approach to conduct problems was eclective, based on individual intervention tailored according to the therapist's personal assessment of a family's needs. Results indicated that the PT program was superior to eclectic treatment in terms of mother's reports of child improvements at home, mothers' confidence levels, and overall consumer satisfaction.

In addition to comparing PT with other forms of family therapy, there have been some attempts to determine which aspects of PT account for the improvements resulting from treatment. For example, to what extent is therapeutic change accounted for by the content, the therapy methods (i.e., videotape modeling), the group support, or the therapist's interpersonal skills? In this vein, Webster-Stratton and colleagues have attempted to determine which methods of PT are most effective. In their first randomized study, they showed that the therapist-led group discussion videotape modeling method (GDVM) was equally good if not more effective than a PT method based on the highly individualized one-to-one "bug-in-the-ear" approach involving direct coaching of the parent and child (Webster-Stratton, 1984). Next they showed that PT based on the GDVM approach was somewhat more effective than either PT based on a therapist-led group discussion approach—without videotape modeling (GD)—or a completely self-administered videotape modeling approach—without therapist feedback or group discussion (IVM) (Webster-Stratton et al., 1988, 1989). All three approaches showed significant improvements in comparison to the waiting-list control condition—even the self-administered videotape modeling program had better treatment adherence than the group discussion approach without videotape modeling. At 3 year follow-up, the improvements in conduct problems (as reported by parents) were maintained only in the group trained by GDVM. In contrast, parents in both the IVM and GD conditions reported significant increases in their children's conduct problems from the 1 year to 3 year follow-up (although still below baseline levels). This component analysis of PT methods seems to suggest that PT methods based on videotape modeling plus parent group discussion and support will produce more sustained and long-term effects than programs which do not use videotape modeling or group discussion and mutual support methods. Moreover, the group approach represents a cost-effective alternative to conventional family therapy (i.e., individual therapy with a single family).

5.08.4.3 Factors Contributing to Program Success or Failure

Despite the general overall success of these programs in producing statistically significant changes in parent and child behaviors, there is also evidence that some families do not respond to treatment; these children continue to have clinically significant behavior problems after treatment (Eyberg & Johnson, 1974). In long-term follow-up studies, 30–40% of treated parents and 25% of teachers report children to have behavior problems in the deviant or clinical range (Forehand, Furey, & McMahon, 1984; Schmaling & Jacobson, 1987; Webster-Stratton, 1990a; Webster-Stratton & Hammond, 1990). In the Forehand and Long (1986) review, according to behavior observation measures, 4 of the 12 programs failed to produce changes in behavior that were maintained over time; and follow-ups were beset with high attrition rates (50%). In an effort to account for these treatment failures, researchers have focused on identifying variables other than parenting skills which may be directly or indirectly related to the continuation of conduct problems. In particular, there is an emerging literature concerned with family variables such as parents' personal problems, marital relationships, and relationships with others outside the family. Parent and family characteristics such as parental depression, marital conflict, spouse abuse, lack of a supportive partner, poor problem-solving skills, and high life stress have been shown to be associated with fewer treatment gains (Forehand et al., 1984; Forgatch, 1989; Webster-Stratton, 1985b).

5.08.4.3.1 Parental depression

There are somewhat conflicting results regarding the influence of depression on parents' ability to benefit from PT (Forehand et al., 1984; Wahler, 1980; Webster-Stratton, 1988; Webster-Stratton & Hammond, 1988). Researchers have long noted that pretreatment levels of maternal depression (self-reported symptoms, not formal diagnosis) are significant predictors of attrition rates during PT (McMahon, Forehand, Griest, & Wells, 1981) and failure to take part in follow-up assessments. However, research has rarely evaluated paternal depression as a predictor of treatment outcome. In the only study to be found concerning the role of fathers' depression, Webster-Stratton and Hammond (1990) evaluated the relative contribution of maternal and paternal depression (along with other variables related to parents' mental and emotional states) to short-term and long-term treatment outcome for 101 families who had been treated for their children's conduct problems via PT. They found that pretreatment levels of maternal and paternal depression (as measured by the Beck Depression Inventory) were significant predictors of maternal and paternal reports of their children's maladjustment (i.e., failure to fall into the normal range on standardized parent report measures) immediately post-treatment; however, reported depression was not related to clinically significant behavioral improvements following treatment—either in parenting behaviors (i.e., a 30% drop in criticisms from baseline) or in the child's interactions with mothers or fathers (i.e., a 30% drop in total negative behaviors from baseline) (Webster-Stratton, 1994). At 1 year follow-up, pretreatment levels of depression still predicted mother and father reports of child adjustment, but again were not related to parent and child behavioral improvements at home (as independently observed) or to child behavior at school (as reported by teachers). On the other hand, marital status and socioeconomic status for mothers were strong predictors of continuation of mothers' critical behaviors and child conduct problems post-treatment, while marital dissatisfaction and negative life events were strong predictors of fathers' critical behaviors and child conduct problems at follow-up. This study confirmed an earlier study by Forehand and Furey (1985) which found an association between marital satisfaction and parent and child behaviors, and an association between depression and mothers' perceptions or attributions concerning their children, but not mothers' behaviors.

Studies have also shown that maternal depression significantly improves following PT when subjects are compared with waiting-list control mothers (Forehand, Wells, & Griest, 1980). In a study involving 85 families of children with conduct problems, Webster-Stratton (1994) reported that both mothers and fathers showed significant reductions in depressive symptoms immediately after PT, levels which were maintained 2 years later. Of the total sample, 31.2% of mothers and 22.4% of fathers indicated depressive symptoms in the mild to moderate range pretreatment; after PT, 54.2% of the depressed mothers and 83.3% of the depressed fathers had changed into the normal range. Further analysis indicated that those mothers who showed significant improvements in their depression scores (into the normal range) reported more significant improvements in child adjustment than mothers whose depression scores remained abnormal, and their children showed a higher ratio of positive to negative social skills (30% increase) than children of mothers with abnormal depression scores. For fathers, no significant relationships were found between an improvement in depression level and improvements in child behavior. These data suggest that the best predictor of long-term treatment outcome for children may be mothers' post-treatment levels of depression, rather than pretreatment levels.

5.08.4.3.2 Marital conflict

Next, consider the impact of marital conflict on treatment outcome (Grych & Fincham, 1990). Earlier it was noted that the relationship between the intensity of marital conflict and levels of childhood behavior problems (Fantuzzo et al., 1991). Unfortunately, only a few studies have examined the relationship of marital variables to treatment outcome for parents of children with conduct disorders. Some of the early studies (Oltmanns et al., 1977) found no relationship between pretreatment levels of marital adjustment and treatment outcome, and found no changes in marital satisfaction following treatment. However, others have found marital distress to be an important predictor of treatment failure (Dadds et al., 1987; Webster-Stratton, 1994). Dadds found that marital conflict was predictive of poor treatment outcome assessed at 6 month follow-up (Dadds & McHugh, 1992). In Webster-Stratton and Hammond's (1990) study, marital status and marital conflict (for mothers and fathers, respectively) were factors that made the greatest contribution to the prediction of child behavior problems immediately post-treatment. One year later, marital status was the strongest predictor of treatment outcome as measured by teachers' reports of child behavior; single-parent families accounted for significantly more nonresponders to treatment (Webster-Stratton, 1990a). Mothers of children classified as nonresponders also reported significantly lower income and more depression than mothers of children classified as responders. However, at 1 year follow-up there were no significant differences in levels of marital conflict reported by parents of children classified as responders and parents of nonresponders. Moreover, Webster-Stratton found that marital conflict did not improve as a result of PT, unlike mothers' depression. Consequently, it appears that while PT can improve parent–child interactions and maternal depression, it has no impact on marital satisfaction or conflict; conversely, marital conflict and marital satisfaction do appear to have an impact on the outcome of PT.

These findings suggest a model in which the key causal factor behind child conduct problems is not lack of parenting skills *per se*, but rather parents' lack of conflict management skills. Parents who have more difficulties with communication, conflict resolution, and affect regulation find it harder to cope with life stressors, marital disagreements, and common child misbehaviors. As a result, not only do they fall into the negative reinforcement trap when parenting, but they also model their troubled communication and affective patterns for their children, thereby further contributing to their children's difficulties with peer relationships including poor communication and poor problem-solving, as well as escalating anger and aggression.

5.08.4.3.3 Insularity and socioeconomic disadvantage

How do insularity, family socioeconomic disadvantage, and high levels of life stress contribute to treatment success or failure? Recruitment rates for PT interventions with low-income families of conduct-disordered children are low (Spoth & Redmond, 1995). Families who are socioeconomically disadvantaged have often been reported to be more likely to drop out of treatment (Eyberg, 1992) and more likely to relapse or fail to make clinically significant improvements following treatment or to maintain treatment effects over time (Dumas & Wahler, 1983; Wahler & Afton, 1980; Wahler & Dumas, 1984; Wahler & Fox, 1982; Webster-Stratton, 1985b). In a study of 257 clinic children and their families who initiated treatment for aggressive behavior, it was shown that parental stress (total stress and life events) predicted early drop-out from treatment but not late drop-out (Kazdin & Mazurick, 1994). Wahler (1980) found that mothers who reported having had negative interactions with their community before treatment were less likely to maintain treatment effects over time.

Socioeconomic disadvantage coupled with social isolation or insularity resulted in a steady increase in the probability of PT treatment failure (Dumas & Wahler, 1983). Finally, in another study it was shown that those clinic families who had a partner or father involved in treatment (presumably providing more support) were more likely to sustain treatment effects 2–3 years later (Webster-Stratton, 1985a). A family's ability to maintain treatment effects at one year follow-up has been correlated with their degree of negative life crisis and environmental stresses (e.g., move to new neighborhood, death in family, unemployment), degree of insularity or support, and poverty (Webster-Stratton, 1985b).

All of these findings taken together suggest that PT programs need to be broadened to emphasize partner involvement (see earlier discussion regarding marital support and communication training), parent support, marital communication, problem-solving and coping skills. Moreover, the failure of many programs to result in child behavior improvements which generalize to school settings suggests the need for some form of training that equips parents to collaborate with teachers in order to improve their children's behavior at school as well as at home. This broadened view of PT will more accurately reflect the more complex model concerning the etiology of conduct problems.

5.08.4.3.4 Child characteristics

With respect to child characteristics, such as the nature of the child conduct problems, the child's age, sex, race, and particular biology, there is a lack of information concerning the important child predictors of long-term treatment outcome. Outcome studies rarely separate results for preschoolers from results for school-aged children. However, there is a suggestion from several follow-up studies of PT programs for parents of children with conduct disorders that the younger the child at the age of treatment, the more positive the child's behavioral adjustment at home and at school (Strain, Steele, Ellis, & Timm, 1982). Dishion, Patterson, and Kavanagh (1992) reported that children aged 2–6 years had the most significant improvements. Neither sex nor race have predicted outcome, although there is a suggestion in the literature that there may be different etiological factors leading to ODD/CD for girls and boys (Webster-Stratton, 1996a). There is scant research concerning how child biological factors affect PT outcome, although it has been suggested that the prognosis is worse for children with conduct problems combined with ADHD (unless medication is added) (Barkley, 1996; Walker, Lahey, Hynd, & Frame, 1987). Nonetheless, this research was carried out with older children, leaving it unclear how child biological factors influence PT for young children. Webster-Stratton's studies did not find that child age, sex, or hyperactivity predicted treatment success; instead, it was family factors such as life stress, socioeconomic status, and marital conflict that predicted long-term outcome in terms of the continuation of child antisocial behavior following her PT program (Webster-Stratton, 1985b; Webster-Stratton & Hammond, 1990).

5.08.4.3.5 Therapy characteristics

It is clear that not all PT programs are equally effective. Therapy factors, such as the format and methods of training, the dose and length of treatment received, and therapist characteristics, undoubtedly play an important role in predicting treatment outcome. However, these therapy factors have rarely been evaluated. For example, length of treatment has rarely been assessed as an independent variable. Kazdin (1987) has suggested that PT programs less than 10 hours in duration are less likely to be effective with conduct-disordered children. Certainly the programs which we have highlighted above have been lengthy, ranging from 20 to 45 hours. Related to the length of the treatment program is the dose of therapy which the family actually received. There is some suggestion in the literature that families who attend more sessions (greater than 50%) have a more successful outcome than families with poor attendance (Strain et al., 1982). However, this finding has not been reported by other researchers (Dumas & Albin, 1986). Another factor related to treatment dose is whether the father (or some other family member) participated in the training along with the mother. In one study, it was shown that those clinic families who had a partner or father involved in treatment (presumably providing more support) were more likely to sustain treatment effects 2–3 years later (Webster-Stratton, 1985a). The existing evidence would suggest that the involvement of another supportive family member leads to better generalization and maintenance of treatment effects over time (Webster-Stratton, 1985a; Webster-Stratton et al., 1988). Finally, Webster-Stratton's studies suggest that treatment methods such as group support and videotape modeling add substantially to the long-term effectiveness of PT programs (Webster-Stratton, 1984, 1990a).

5.08.4.3.6 Therapist characteristics

There is even less research concerning the role of the therapist and the therapeutic process in predicting treatment outcome. Webster-Stratton and Herbert c(1993, 1994) have described their PT programs as based on a collaborative or partnership model, whereby the therapist works *with* parents by actively soliciting their ideas and jointly involving them in the process of problem-solving, sharing, discussing, and debating ideas and solutions (also see Webster-Stratton & Herbert, 1994). It is suggested that this approach empowers parents by giving back dignity, respect, and self-control to parents who are often seeking help for their children's problems at a vulnerable time of low self-confidence and intense feelings of guilt and self-blame (Spitzer, Webster-Stratton, & Hollinsworth, 1991). On the other hand, some behavioral PT programs seem to be based on an "expert" or hierarchical model, whereby the therapist is more directive—suggesting solutions and dispensing advice for the parents. To this author's knowledge, there has been no research comparing PT programs in terms of their different helping models and different sets of assumptions about the cause of family problems.

While there are no comparison studies, one of the most comprehensive microsocial analyses of therapist–client interchanges was conducted by Patterson and Forgatch (1985). They showed that directive therapist behaviors such as "teach" and "confront" increased the likelihood of parental resistance and noncooperativeness in a session. On the other hand, therapist behaviors such as "facilitate" and "support" led to reliable decreases in client noncompliance. Patterson and Chamberlain (1988) have proposed a therapy model which postulates that therapist behaviors play a secondary role to extrafamilial, interpersonal, and child factors in predicting parent response to the early stages of treatment, but play a primary role in predicting client noncompliance in the later stages of therapy. In another study, Alexander and Parsons (1973) examined the role of therapist characteristics in predicting outcome (as defined by completion of treatment and recidivism rate) for families that participated in family functioning therapy (FFT). They found that relationship characteristics (affect, warmth, humor) accounted for 45% of the variance in predicting outcome, whereas structuring characteristics (directiveness and self-confidence) only accounted for an additional 15% of the variance.

5.08.4.4 Limitations of Studies Related to Treatment Outcome

The research regarding treatment outcome is beset with methodological problems. First, there is a lack of unanimity as to what is meant by "successful" treatment or "improvement" in conduct problems. Different studies have utilized different criteria for evaluating treatment outcome as well as different measures of those criteria. Because of the expense involved in observational methods, some studies have used mothers' reports of improvement in the child's behaviors as the definitive outcome criterion; however, global ratings by parents may reflect how they feel about themselves and their child, rather than any real change in child behavior. Further, it appears that the validity of raters' evaluations varies according to characteristics of the raters. For example, Patterson has shown that the more deviant or socially disadvantaged the parents, the less convergence there is between their ratings and teacher ratings or observational ratings (Patterson et al., 1993). Studies of clinical samples have shown maternal ratings of problem children to be correlated more highly with their self-ratings of depression than with the observed behavior of their children (Forehand et al., 1984; Griest, Forehand, Wells, & McMahon, 1980). Patterson and Narrett (1990) have suggested that mothers' ratings of child behavior reflect improvement even when there is no treatment—resulting in the impression that anything and everything works! On the other hand, neither is independent observation of behavior a wholly unbiased method of evaluating change, though it appears to be less susceptible to some of the problems of self-report measures.

Even when researchers agree on the criterion for evaluating treatment outcome, there may be no agreement on what variable represents that criterion. For instance, if observable behavior is the criterion, the question then becomes whether the relevant outcome variable is parent behavior or child behavior. Whether one expects parent and child behavior to change simultaneously depends on one's theoretical model. Those who ascribe to the PT model believe that changes in parenting behavior (the most proximal variable being influenced by the treatment) eventually will cause changes in child behaviors (the more distal variable), which implies that the appropriate criterion for evaluating the success of PT is, at least for the short term, parenting behavior. Support for this theory is found in one study which showed that the magnitude of changes in parental discipline correlated significantly with changes in antisocial behavior (Dishion et al., 1992). Patterson and Forgatch (1995) examined separately the contributions of baseline, termination, and change scores from parent and teacher reports, as well as observations and process variables to determine their validity as predictors of future child outcomes. The results indicated that termination scores for parents' behaviors post-treatment and the amount of change in parental practices (i.e., positive problem-solving) were powerful predictors of long-term outcome for children, defined as future arrests or removal from home. Neither termination ratings nor change ratings of teacher or parent reports nor termination scores of child adjustment immediately post-treatment predicted future child arrests or out-of-home placements. This study supports the use of parent behavior as predictors of child outcomes—and yet there is no consensus that child arrests or out-of-home placements are criteria for treatment success.

Another issue regarding treatment outcome is: how much change constitutes significant improvement? Often observational measures lack normative data on comparable nonclinic samples with which to compare our clinic samples. As a result, researchers typically utilize percentage improvement from baseline to indicate clinically significant change. But there is no consensus regarding the degree of improvement—the percentage of change—that is necessary to declare behavior "improved." Moreover, in a case where a child does show a 30% or even a 50% reduction in aggressive or noncompliant behaviors from baseline, but is still in the clinical range, is he or she defined as improved or not?

The next limitation to existing research concerns risk factors. As can be seen, in this review, studies typically have analyzed a single risk factor by means of global measures (such as depression, marital satisfaction, family insularity, or life stress) and correlated it with some aspect of treatment outcome. Just as there is a problem with the use of global self-report measures of treatment outcome, there is a problem with the use of global self-report measures as predictors. Global reports do not elucidate the precise aspects of, say marital distress or depression, nor the mechanisms by which these problems maintain the problem behaviors—i.e., contribute to the failure of treatment. In addition to this problem, most studies have been based on short-term results of treatment; they have told us little about the more important concern, namely how these family processes influence long-term treatment outcome. Moreover, very few studies have compared the relative weight or contribution of individual risk factors (such as depression in combination with marital distress) with treatment outcome. In comparison to the large number of individual treatment outcome studies, there is a paucity of studies containing any information about predictors of treatment success, in most cases because the study sample size was too small to form any meaningful conclusions.

Another problem with the research on risk factors is that those studies which have looked at predictors of treatment outcome in terms of children's conduct problems have tended to combine both boys and girls in their samples. It is highly probable that the predictors will be different for boys and girls. Webster-Stratton (1996a) found not only that the predictors differed depending on whether the outcome variable assessed was child externalizing problems at home (according to independent observations) or at school (according to teachers), but also that there were gender differences in predictors: mother negativity and depression and father negativity and life stress significantly predicted girls' externalizing problems 2 years post-treatment and accounted for 49% of the variance, whereas for boys no parenting or family variables emerged as predictors—rather independent observations of externalizing behaviors at baseline were the best predictors of boys' continuing externalizing problems at home 2 years later.

5.08.5 BROADENING THE FOCUS OF PT

Despite the limitations in research studies and differences in outcome measures employed, there has been remarkable consistency in the findings and they seem to hold true even in the face of these problems. Research results regarding the characteristics of families who failed to respond to PT pointed to the necessity of expanding our etiological models as well as broadening our PT focus. Since the mid-1980s, efforts have been made to improve the effectiveness of PT by adding other therapy components to the basic programs: training in social learning principles, problem-solving, self-control and emotional regulation, marital communication, anger management skills, "synthesis training," and ways to give and get support. These components include more sophisticated affective and cognitive skill training. Evaluation of the effectiveness of these adjuncts to PT helps test models regarding the important factors contributing to the maintenance of conduct problems.

5.08.5.1 PT with Marital and Other Parent-focused Adjuncts

Based on the studies indicating associations between marital conflict and poor child-treatment outcomes, a number of investigators have evaluated the effects of adding specific marital treatment components to standard PT. Greist et al. (1982) compared PT with a program which combined PT with a focus on marital adjustment and extrafamilial relationships. The combined program showed greater improvements in terms of child deviance at short-term follow-up. Dadds et al. (1987) examined the specific effects of a brief marital communication intervention, partner-supported training, as an adjunct to PT for parents with a child who had conduct problems. Results indicated that the adjunctive marital component led to enhanced treatment outcomes for couples experiencing discord (but not for maritally satisfied couples) compared with the couples experiencing discord who received only PT (Dadds & McHugh, 1992; Dadds et al., 1987).

Two other consistently documented factors related to treatment failure or drop-out and failure to maintain change following treatment are social disadvantage and isolation (Dumas & Albin, 1986; Webster-Stratton, 1985b). A number of studies have developed PT programs addressing these issues. Wahler (1980) and Wahler and Dumas (1989) defined "insular" parents as those who experience low levels of positively perceived social interactions and high levels of negatively perceived coercive exchanges within the community. They hypothesize that insular mothers have disruptions in their monitoring processes which interfere with their ability to adequately monitor their children and result in indiscriminate responding. Consequently, these researchers developed an adjunct treatment called "synthesis training" whereby parents are taught how their external events or circumstances influence their parenting behavior in order to help them be more effective at monitoring and providing discriminate responses. Synthesis training has been evaluated with multistressed mothers of children with conduct problems (Wahler, Cartor, Fleischman, & Lambert, 1993) and compared with PT alone, and with PT plus friendship liaison. Results indicated that parents who received the synthesis training had reductions in their indiscriminate parenting and their children demonstrated behavioral improvements. Mothers in the basic PT did not change their behavior, nor did their children.

Dadds and McHugh (1992) assessed the role of social support in the outcome of PT for single parents of conduct problem children and assessed the impact of "adjunctive ally support" training (AST) on treatment outcome. Single parents with children diagnosed with ODD or CD were randomly assigned to PT or PT plus AST. Improvements were found in both groups, but AST produced no extra gains. However, responders from both groups were more likely than nonresponders to report high levels of social support from friends, a finding which is consistent with that mentioned earlier regarding father involvement in the PT (Webster-Stratton, 1985a).

In an attempt to address a number of family issues including social support, marital issues and maternal depression, Webster-Stratton (1994) developed a lengthy, 14 week adjunct program (ADVANCE) involving therapist-led group discussions of videotape vignettes on all the following topics: strengthening social support and self-care, personal self-control, family communication skills, problem-solving between adults (and with teachers), and teaching children to problem-solve. Emotional communication is a major emphasis of this advanced program. This is an important skill for families who have until recently felt overwhelmed by negative affect, and who may have indirectly or directly communicated to their children that negative emotions are too frightening or too aversive to share. Parents learn to talk to children about emotion and to respond to their children's emotions appropriately and contingently as a way of increasing their children's emotional regulation (Denham, 1989; Denham, Zoller, & Couchold, 1994).

Families of children with conduct problems were randomly assigned to either the basic PT (GDVM) or GDVM plus ADVANCE. Results immediately post-treatment showed significant improvements in parents' problem-solving, communication and collaboration skills, as well as in children's problem-solving, for families in the combined treatment. Both mothers and fathers from the combined program reported increased consumer satisfaction, and there were no drop-outs from the ADVANCE program, suggesting the perceived usefulness of the skills taught in this program. Moreover, fathers who showed improvements in marital communication and problem-solving post-treatment also showed significantly greater improvements in parenting skills (specifically, reduced hostility and criticisms with children). However, children of ADVANCE parents failed to show significantly enhanced improvements over the children of GDVM parents in terms of child deviance. Both groups of children showed significant reductions in child deviance at home. It was hypothesized that the effects of the ADVANCE parents' improved communication and problem-solving skills may have been delayed, as children need to be exposed to repeated modeling over a sustained period of time. This suggests the need for longer-term follow-up of the two groups.

In summary, studies which have investigated the potential contribution of adjuncts to the standard PT intervention have generally indicated that components which address the social context of families, marital issues, and parental depression can improve the short- and long-term outcomes over and above basic PT (Dadds et al., 1987; Griest et al., 1982; Webster-Stratton, 1994 1997b).

5.08.5.2 PT with Child Training Adjuncts

Studies have indicated that children with ODD/CD may have some sociocognitive deficits in social problem-solving, self-control, and perceptions; these may be of biological origin, may be the result of parental modeling, or some combination of the two. This has led to a few studies evaluating the effects of combining child training (CT) with PT. Kazdin, Siegel, and Bass (1992) compared three interventions: PT, problem-solving skills training for children (PSST), and PT combined with PSST. Each treatment reduced overall child conduct problems and increased social competence. However, the combined treatment led to more durable changes in child behavior, placing a higher proportion of children in the range of normal functioning. The combined treatment also resulted in greater changes in parent functioning, including reductions in parental stress, depression, and overall symptoms. These results were maintained at 1 year follow-up. In another study, Baum, Reyna-McGlone and Ollendick (1986) compared three different interventions: PT, PT combined with child self-control training, and a control condition in which parent discussion was combined with a child attention control condition (control). Results indicated that the PT intervention (with or without the CT component) was superior to the control condition and, for the most part, the two treatment conditions did not differ from each other immediately post-treatment. However, at follow-up the combined treatment (PT plus child self-control training) was superior to PT alone and to the control condition on observational measures of child deviant behavior and on child and mother reports of adjustment.

Finally, Webster-Stratton (1997b) compared four conditions: PT, CT, CT + PT, and a waiting-list control group. Post-treatment assessments indicated that all three treatment conditions had resulted in significant improvements in comparison with controls, as measured by mother and father reports, daily observations of targeted behaviors at home and laboratory observations of interactions with a best friend. Comparisons of the three treatment conditions indicated that CT and CT + PT children showed significant improvements in problem solving as well as conflict management skills, as measured by observations of their interactions with a best friend; differences among treatment conditions on these measures consistently favored the CT condition over the PT condition. As for parent and child behavior at home, PT and CT + PT parents and children had significantly more positive interactions in comparison with CT parents and children. One year follow-up assessments indicated that all the significant changes noted immediately post-treatment were maintained over time. Moreover, child conduct problems at home had significantly lessened over time. Analyses of the clinical significance of the results suggested that the combined CT + PT condition produced the most significant improvements in child behavior at 1 year follow-up.

In summary, these studies on child adjuncts to PT support the hypothesis that supplementing PT with a component that addresses the child directly increases the efficacy of PT. Additionally, these data point to the necessity for long-term assessment of interventions, since the benefits of the adjunct treatment often were not evident until the follow-up evaluation.

5.08.5.3 Future Directions for Research and for Practice

Our review of the research regarding PT has focused on programs for children with ODD/CD because this area comprises the largest body of research regarding PT. Second, we have focused on programs that target preschool and school-age children rather than adolescents because we feel adolescent conduct disorders necessitate other interventions such as individual therapy in addition to or alongside PT. We have not reviewed PT programs specifically targeted at other childhood disorders such as autism, anxiety disorders, learning disabilities, and ADHD.

Despite the limitations of existing research, the evidence appears to support a more comprehensive, multisystem-based model in which conduct problems are not seen merely as evidence of inept parenting skills which need to be changed via treatment, but instead where parenting behavior is seen to be impacted by the events and conditions within and outside the family. We believe the findings to date argue for broader-based treatments which address multiple risk factors and strengthen multiple protective factors such as parenting skills, support networks, and problem-solving and communication as a means of preventing early-onset conduct problems. By strengthening parenting skills and problem-solving strategies and ensuring adequate support (through broader-based family training support groups) for families of young children with conduct problems, we can help families buffer some of the potentially disruptive effects of ongoing depression, marital conflict, negative life stress, and poverty on children's adjustment. Indeed, initial studies of broader-based interventions suggest they promote more enduring child improvements which generalize across settings and over time, as well as fewer treatment drop-outs. Combining PT with CT appears to lead to even greater maintenance of improvements in parent and child behaviors.

Nonetheless, the original formulation of Patterson et al. (1975) is still compelling, particularly as it relates to young children with conduct problems: parenting skills are the most important proximal cause of early-onset conduct problems, because parents are still the primary agents of socialization of the child at this age. Given that this formulation is correct, the success of treatment programs in changing the parenting behavior of parents of young children should predict the child's future adjustment. For children aged 3–7 years, peer rejection, deviant peer groups, academic failure, and negative reputations have not stabilized, whereas for the older school-age child these factors play a key role in maintaining the conduct problems and must be taken into account in intervention. Moreover, parents of younger children are still hopeful and their patterns of parenting are more malleable than those parents who have spent years of coping unsuccessfully with an oppositional child at home. Support for this idea comes from studies which have suggested that PT is more effective in improving both parenting skills and child adjustment when offered to parents of younger children versus older school-age children and adolescents (where presumably more will be required in terms of direct intervention with the child and his or her peer group) (Strain et al., 1982). Further support for the importance of developing comprehensive interventions that can strengthen parenting skills and thereby influence the long-term outcome for children's adjustment is provided by Patterson and Forgatch's (1995) study.

5.08.5.4 Expansion Across Multiple Settings

Nonetheless, while PT holds much promise for treating children with conduct problems, there are several important limitations to this approach. The first important limitation is the inability of this approach to consistently produce child behavior improvements which generalize beyond the home to day-care or school settings and to peer relationships. While the majority of children improve their social behavior at home, 30–50% continue to have significant school problems such as social acceptance, conduct problems, and academic underachievement. Intervening with children's teachers as well as their parents would seem to offer far better possibilities of generalizing improved social skills from the home to preschool and school settings. In particular, schools may enhance the effects of PT by providing teachers not only with classroom behavior management training, but also with the curriculum and support to directly offer training in social skills and problem-solving.

A second limitation is that, despite the documented links between academic underachievement, language delays, reading and learning disabilities, ADHD, and conduct disorders, PT programs rarely, if ever, have included an academic skills enhancement program for parents. Parents need to know not only how to help their children with their antisocial problems but also how to teach and support them regarding their academic difficulties. They need to know how to work with teachers and schools in order to foster a supportive relationship between the home and school settings. Parents and teachers alike need to understand how to collaborate with one another, to build supportive teams to help these children both academically and socially. Such a coordinated effort between the home and school regarding social and academic goals will offer the possibility for more consistent generalization of child improvements across settings. There is little or no research comparing PT with PT offered in conjunction with teacher training or, more generally, evaluation of intervention efforts to span home and school. Research which evaluates the added effects of combining teacher training and parents' academic skills training to our traditional PT approaches should be encouraged.

5.08.5.5 Teachers and Other Professionals

Since preschoolers living in more dysfunctional or disrupted family environments or in poverty are at particular risk for persistent problems, the need for strong educational and therapeutic preschool programs with nurturing and well-qualified teachers is critical. This means preschool teachers and child care workers should be well trained in child development and behavior management skills. Stable and nurturing caregivers may be able to buffer the effects of parental unavailability or psychopathology, at least for some children. However, all too often these are the very children who are spending large amounts of time in the least adequate facilities with untrained caregivers, where child-to-caretaker ratios make it impossible to provide the attention, affection, and emotional support they need. Finally, since parents bring their first questions about common developmental issues and problem behaviors to their pediatrician or family physician, these professionals also need training that allows them to understand parental perspectives within a developmental framework—so that they neither pathologize common problems nor minimize parents' concerns.

5.08.5.6 Addressing Extrafamilial Change

Interventions which have focused on adjuncts to PT reflect a paradigm shift from individual change (i.e., parenting skills) to within-family change (i.e., marital communication and problem-solving skills). But there has been no further shift to what might be called extrafamilial or interfamilial change, such as a family's need to form stronger and more supportive connections with other families and with the community in general. We theorize that, particularly for low-income families, there is an urgent need to broaden PT programs so as to focus on building community and parent support networks. Indeed there is evidence from the "buffering" interpretation of social support (Cohen & Willis, 1985) that particularly for low-income families, the greater the number of people parents felt they could rely on for informal assistance and the more satisfied they felt with their social support, the more likely they were to be nurturing and positive in their parenting interactions and the less likely they were to report problematic behavior (Hashima & Amato, 1994) compared with low-income mothers who felt isolated and dissatisfied with their social support (Jennings, Stagg, & Conners, 1991).

Many interventions have been based on the individual one-to-one counseling model which fosters reliance on the therapist. Those programs which have used a group approach (Webster-Stratton, 1994) have suggested that fostering supportive social networks was a primary purpose of the group approach. The therapeutic group can decrease families' isolation, promote a feeling of support and involvement, and build a sense of community. However, little reference is made to the specific therapeutic strategies used to achieve this goal. Indeed, compared with our well-developed research methods of measuring parent behavior change, few studies report outcome measures having to do with change in social networks, parents' sense of support, or their involvement in their communities (including their children's school) as a result of PT. If we are going to develop more effective interventions targeting these factors, we will need suitable outcome measures (Webster-Stratton, 1997a).

5.08.5.7 Engaging Low-income Families, Understanding Obstacles and the Importance of Culture

The PT literature has suggested that PT is less effective with disadvantaged parents—particularly low-income single mothers (Wahler, 1980). Such families have been described as unmotivated, resistant, unreliable, disengaged, chaotic, in denial, disorganized, uncaring, dysfunctional, and unlikely candidates for this kind of treatment—in short, "unreachable." However, these families might well describe traditional clinic-based programs as "unreachable." Clinical programs are often too far away from home, too expensive, insensitive, distant, inflexible in terms of scheduling and content, foreign in terms of language (literally or figuratively), blaming or critical of their lifestyle. A cost–benefit analysis would, in all likelihood, reveal that the costs to these families of receiving treatment far outweigh the potential benefits—even though they do genuinely want to do what is best for their children. Perhaps this population has been "unreachable" not because of their own characteristics, but because of the characteristics and obstacles of the interventions they have been offered.

The paradox is that while, on the one hand, we decry the lack of efficacy of therapy (i.e., PT) with economically disadvantaged families, on the other hand, we also maintain the belief that if we could just do *more psychodynamic therapy* focusing more broadly (i.e., on family dysfunction and parental psychopathology), we would be more effective. But we believe the problem may not lie in the focus of the therapy (i.e., parenting skills vs. family dynamics), but rather in the therapeutic model or approach—namely, a traditional clinic-based model of PT. Before abandoning PT as an intervention with this population, we should examine alternative models of delivering PT which take the broader social context into consideration. If PT is offered in nonstigmatizing locations (as opposed to mental heath centers) and made readily available in communities through churches, schools, and community centers, the programs might well attract more families. PT programs which provide for transportation, day care, meals, flexible scheduling, and attend to some of the obstacles for families getting to and staying involved with programs are likely to be more successful. More research is needed on how programs can be made culturally sensitive and tailored to the specific needs and priorities of the particular parents involved. In general, much more research is needed on ways to promote parental involvement and engagement in PT programs.

Socioeconomically disadvantaged families and rural families are less likely to seek mental health services for children's behavioral problems. This may be due to psychological obstacles such as the stigma associated with seeking assistance for psychological problems. Or it may be due to differences in priorities, or to cultural differences in how behavior problems are viewed and solved within a family. For example, in some cultures there is shame attached to disclosing family problems outside the family which extends to the entire family, relatives, and ancestors. Or it may be the result of an attitude that little can be done about the problem. As we noted earlier, failure to access such services may result from structural obstacles such as difficulty accessing and paying for services, lack of transportation, the distance involved in getting to the service, lack of babysitting for other children, inflexible clinic hours, etc. It is necessary to understand the obstacles that disadvantaged groups experience when accessing services and to develop more effective ways of reaching these families and attracting them to such programs.

The growing ethnic diversity among families in the United States also makes it imperative that we not only recognize and understand different cultural attitudes and practices regarding parenting, but more importantly develop interventions that are culturally informed and appropriate for different ethnic groups. It is important to be cautious about generalizations regarding specific cultural groups, for every parent is different and has his or her own unique history. Nevertheless, it is essential to understand the cultural values and historical factors that are the context for parenting, influencing parents' behavior with their children (Harkness & Super, 1995). For example, parents from certain cultural groups may put a high priority on their children developing habits of obedience, humility, respect for the elders, reliance on family networks, and family interdependence, while parents from other groups may value creativity, personal self-control, and emotional independence. In some cultural groups communication styles are open, volatile, and expressive, while in other groups communication is more indirect and infrequent. Attitudes toward discipline range from strict discipline practices of cultures which value obedience to authority to the permissive discipline practices of cultures which place emphasis on personal autonomy and self-respect. PT research encompassing cultural issues is almost nonexistent (Forehand & Kotchick, 1996). Apart from anecdotal information, we do not know whether certain behavior management strategies (e.g., play, time out, ignoring) are more easily accepted in some cultural groups than others or if a certain therapeutic style would be more effective than another. Finally, mistrust of the majority culture by ethnic minorities can be a barrier to PT. Ethnic minorities walk a fine line between maintaining their cultural history and traditions and adopting the strategies of the European–American culture that are typically associated with success (Coughlin, 1995). This is yet another reason why parenting programs which are collaborative are more likely to be accepted by parents: because parents set their own goals for therapy based on their personal values. Collaborative therapy is inherently more accommodating of cultural differences, whereas noncollaborative therapies are more likely to be shaped by the cultural values and norms of the therapist.

Webster-Stratton has conducted a randomized study to evaluate a prevention program for a multiethnic group of 250 Head Start families, 90% of whom were on welfare. This 9 week, 2 hour parenting program was community-based (held in housing units, churches, and schools) and addressed many of the issues of economically disadvantaged families by providing transportation, day care, dinners, afternoon and evening groups, and parent group support. Since a major theme of the program (in addition to learning parent management skills) was to decrease family isolation by learning effective ways to give and get support (inside and outside the family), parents were paired up to make weekly "buddy" calls. Group leaders called parents weekly. Teachers were also trained in the program and encouraged to provide extra support to families. Visits to classrooms were arranged as well as helping parents write positive notes to teachers.

Preliminary results are promising, with over 85% of those who participated attending over two-thirds of the sessions. In comparison with mothers in the Head Start centers which did not offer this program, mothers made significantly fewer critical remarks, used less physically negative discipline, and were more positive, appropriate, and consistent in their discipline style. Intervention mothers perceived their family service workers as more supportive than did comparison mothers; furthermore, teachers reported that intervention mothers were more involved in their children's education than nonintervention mothers. In turn, intervention children were observed at home to exhibit significantly fewer negative behaviors, less noncompliance, more positive affect, and more prosocial behaviors than children in Head Start centers which did not offer the parent intervention.

Consumer satisfaction with the program was high, with 89% reporting "positive" to "very positive" overall satisfaction, 91% reporting they expected positive results, and 95% saying they would "highly recommend" the program to others. Over 85% of the parents in the intervention condition wanted the program to be longer and to continue into the kindergarten year (Webster-Stratton, 1995). While it cannot be determined precisely what the active ingredient of this intervention was because there was no comparison PT intervention, the data are suggestive of the value of directly targeting social support both within and outside the family as well as eliminating barriers regarding transportation, day care, and food for low-income families.

In another randomized study (Cunningham, Bremner, & Boyle, 1995), families were randomly assigned to either a 12 week clinic-based individual PT (Clinic/Individual) or a 12 week community-based large group PT (Community/Group) or a waiting-list control condition. Families who were significantly more likely to enroll in the Community/Groups were characterized as immigrant families, those using English as a second language, and parents of children with severe behavior problems. Parents in Community/Groups reported greater improvements in behavior problems at home and better maintenance of these gains at 6 month follow-ups. Moreover, Community/Group interventions were six times more cost effective than Clinic/Individual training. This important study suggests the potential for group community PT programs as a valuable resource which may reduce the cultural, linguistic, and family barriers which prevent participation in clinic-based programs.

5.08.5.8 More Complex Models

It has become evident that the etiology of and treatment model for young children's conduct problems are complex and transactional and must involve not only attention to family functioning and parent–child mechanisms but also the other domains of influence we have mentioned—namely, child biological and sociocognitive factors, attachment issues, and the wider contextual influences (Caspi & Moffitt, 1995). While there have been some efforts to study the interrelationship between various types of PT and family contextual factors, there is comparatively little research looking at the interrelationships between PT and child biology, sociocognitive factors, and parent–child attachment.

These interrelationships are important to understand, for the length, breadth, setting, and focus of training needs to be matched to the risk domains affecting the child's problems. For example, if a child has biological and sociocognitive risk factors as well as insecure attachment with his or her parent, and lives in a stressful family environment, he or she will need a far more comprehensive intervention (involving parent and child training as well as teacher training to alter the peer context in the classroom) than a child who has no biological or family risk factors and has not yet developed pervasive conduct problems which have generalized from home to school. The child who has no biological factors and nonpervasive conduct problems (that is, setting specific) may be adequately treated with a basic PT program. On the other hand, the child coming from a family where there are family problems and social context factors such as drug abuse, marital distress, depression, and attachment problems but who has no biological factors will need an intervention that focuses on family interpersonal problems as well as parenting skills. Depending on the parents' willingness and/or ability to benefit from such a comprehensive family intervention, we may need to supplement the family intervention with teacher training and child training. We should undertake research where we identify risk groups and then evaluate levels of intervention—ranging from relatively basic PT programs to more comprehensive programs involving parent, teacher, and child training.

5.08.5.9 Understanding Therapeutic Process Variables

It is clear that there is a need for research evaluating competing models regarding the etiology of childhood conduct problems. There is a further need for research regarding the critical therapeutic change processes (e.g., comparing interventions which focus on promoting parent–child attachment, interpersonal relationships, and the collaborative process vs. those which are more parent behavior skill-based). While PT research in recent years has focused on expanding the content for PT to include broader family issues, it has largely ignored the role of variables having to do with the therapeutic process itself. Yet these variables affect a family's level of engagement in therapy, their level of resistance and the acceptability of parenting strategies by different cultural and ethnic groups. For example, in one of the earliest studies of therapeutic process, Patterson and Forgatch (1985) have highlighted the critical importance of the manner in which the therapist deals with resistance. They showed that when therapists met parent resistance with confrontation and teaching, there was an increase in parental resistance; when the therapist was nonconfrontive, supportive, and nondirective, the resistance decreased. These findings are important, since parental resistance is a major cause of parental drop-out from therapy and failure to make changes. In another study, Patterson and Forgatch (1995) showed a relationship between fathers' confrontational behavior in therapy and lack of change in mothers' discipline strategies, and a relationship between mothers' hopeless and defending behavior and fathers' discipline strategies. In addition, Alexander and his colleagues have examined the influences of therapist characteristics (e.g., gender) and processes on family behavior (Alexander, Waldron, Newberry, &

Liddle, 1988). They have provided extensive work on the notion of therapist "reframing," suggesting that it is a more effective therapeutic process than therapist teaching and reflecting for reducing defensiveness in parents and adolescents (Robbins, Alexander, Newell, & Turner, 1996). Webster-Stratton and Herbert (1994) have emphasized the importance of a collaborative relationship and interpersonal process with parents and a group approach to PT. This collaboration encompasses 10 domains of therapist skills: (i) building supportive and caring relationships; (ii) empowering parents through behaviors such as reinforcing and validating parental insights; (iii) using active teaching skills (e.g., videotape modeling, role-play, rehearsal, parent group and therapist feedback, home assignments); (iv) interpreting (analogies and metaphors), reframing, and persuading; (v) leading and challenging; (vi) prophesizing (anticipating set-backs and successes); (vii) individualizing, generalizing, and contextualizing; (viii) preparing for the long-term; (ix) fostering PT groups as support systems; and (x) building parent support outside the group. Further research regarding the use of these therapy skills is worth pursuing, in order to understand their relationship to parent engagement in PT and to compare their effects with PT programs which do not utilize these strategies.

Patterson and Forgatch (1995) showed that parental social disadvantage, antisocial behavior, and depression were associated with increased parental resistance throughout PT. They proposed that parental resistance could result from four factors: (i) a history of failure in discipline attempts with the child; (ii) parental psychopathology, depression, and antisocial behavior; (iii) stress and social disadvantage; and (iv) the therapist's clinical skills in teaching and confronting parents. These findings suggest the complex interplay between family, contextual, and therapy process factors.

5.08.6 SUMMARY

Despite the evidence concerning the stability of conduct problems stemming from preschool into later life and the effectiveness of family intervention programs, there is an appalling lack of availability of comprehensive training programs for parents of young children with behavior problems. While funds are available for various mental health screening programs for preschoolers, they are rarely available to provide intervention once children with conduct problems are identified. These children have largely been neglected by the mental health community, which has targeted school-age and adolescence for intervention efforts. Preschoolers are offered help primarily if they have more serious cognitive or developmental delays such as mental retardation, or marked psychosocial disadvantage. When young children with conduct problems come to the attention of professionals, they are frequently ineligible for programs or their problems are dismissed as age-appropriate behaviors which are likely to be outgrown.

Not only are there few programs available, even when children have been identified as having ODD or CD, but there are even fewer prevention programs designed to identify high-risk families and to offer training and support when children are toddlers, before conduct problems develop. Gross, Fogg and Tucker (1995) evaluated the 10 week Webster-Stratton basic parenting program with parents of toddlers judged by their parents as behaviorally difficult. Results suggested that PT led to significant increases in maternal self-efficacy, decreases in maternal stress, and improvements in the quality of mother–toddler interactions. Improvements were maintained 1 year later.

Because it is clear that relatively marked behavior problems in preschoolers have long-term developmental consequences for some proportion of children and their families, the need for early intervention programs is obvious—and pressing. Although comprehensive treatment programs for preschoolers may seem costly in the current economic climate, reducing conduct problems in young children will ultimately decrease the need for mental health and educational services later on. The cost of not treating such early problems far outweighs the cost of instituting appropriately targeted intervention programs when children are young.

5.08.7 ACKNOWLEDGMENTS

This research was supported by the NIH National Center for Nursing Research Grant #5 RO1 NR01075–10 and Research Scientist Development Award MH00988–05 from NIMH.

5.08.8 REFERENCES

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87,* 49–74.

Alexander, J. F., & Parsons, B. V. (1973). Short-term behavioral intervention with delinquent families: Impact on family process and recidivism. *Journal of Abnormal Psychology, 1,* 219–225.

Alexander, J. F., Waldron, H., Newberry, A. M., & Liddle,

N. (1988). Family approaches to treating adolescents. In F. M. Cox, C. Chilman, & E. Nunnally (Eds.), *Families in trouble*. New York: Sage.

Anastopoulos, A. D., Barkley, R. A., & Sheldon, T. L. (1996). Family based treatment: Psychosocial intervention for children and adolescents with attention deficit hyperactivity disorder. In E. D. Hibbs & P. S. Jensen (Eds.), *Psychosocial treatments for child and adolescent disorders: Empirically based strategies for clinical practice* (pp. 267–285). Washington, DC: American Psychological Association.

Anderson, K. E., Lytton, H., & Romney, D. M. (1986). Mothers' interactions with normal and conduct-disordered boys: Who affects whom? *Developmental Psychology, 22,* 604–609.

Arnold, J. E., Levine, A. G., & Patterson, G. R. (1975). Changes in sibling behavior following family intervention. *Journal of Consulting and Clinical Psychology, 43,* 683–688.

Asarnov, J. R., & Callan, J. W. (1985). Boys with peer adjustment problems: Social cognitive processes. *Journal of Consulting and Clinical Psychology, 53,* 80–87.

Baer, D. M., Wolf, M. M., & Risely, T. R. (1968). Some current dimensions of applied behavior analyses. *Journal of Applied Behavior Analyses, 1,* 91–97.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1982). Self-efficacy mechanisms in human agency. *American Psychologist, 84,* 191–215.

Barkley, R. A. (1996). Attention deficit/hyperactivity disorder. In E. J. Mash & R. A. Barkley (Eds.), *Child psychopathology* (pp. 63–112). New York: Guilford Press.

Barrett, P. M., Dadds, M. R., & Rapee, R. M. (1993). *Family Intervention for Childhood Anxiety: A Controlled Trial*. Paper presented at the 27th Annual Convention of the Association for Advancement of Behavior Therapy, Atlanta, GA.

Baum, C. G., Reyna-McGlone, C. L., & Ollendick, T. H. (1986). *The efficacy of behavioral parent training: Behavioral Parent Training Plus Clinical Self-control Training, and a Modified STEP Program for Children Referred with Noncompliance*. Paper presented at the meeting of the Association for Advancement of Behavior Therapy, Chicago.

Baumrind, D. (1971). Current patterns of parental authority. *Psychology Monographs, 1,* 1–102.

Biglan, A. (1993). *A functional contextualist framework for community intervention*. Reno, NV: Context Press.

Bowlby, J. (1980). *Attachment and loss: Loss, sadness, and depression*. New York: Basic Books.

Breiner, J., & Forehand, R. (1982). Mother–child interactions: A comparison of a clinic-referred developmentally delayed group and two non-delayed groups. *Applied Research in Mental Retardation, 3*(2), 175–183.

Campbell, S. (1995). Behavior problems in preschool children: A review of recent research. *Journal of Child Psychology and Psychiatry & Allied Disciplines, 36*(1), 113–149.

Campbell, S. B. (1991). Longitudinal studies of active and aggressive preschoolers: Individual differences in early behavior and outcome. In D. Cichetti & S. L. Toth (Eds.), *Rochester symposium on developmental psychopathology* (pp. 57–90). Hillsdale, NJ: Erlbaum.

Campbell, S. B., & Ewing, L. J. (1990). Follow-up of hard-to-manage preschoolers: Adjustment at age 9 and predictors of continuing symptoms. *Journal of Child Psychology and Psychiatry, 31*(6), 871–889.

Capaldi, D. M., & Patterson, G. R. (1994). Interrelated influences of contextual factors on antisocial behavior in childhood and adolescence for males. In D. C. Fowles, P. Sutker, & S. H. Goodman (Eds.), *Progress in experimental personality and psychopathology research* (pp. 165–198). New York: Springer.

Caspi, A., & Moffitt, T. E. (1995). The continuity of maladaptive behavior: From description to understanding in the study of antisocial behavior. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology* (Vol. 2, pp. 472–511). New York: Wiley.

Cassidy, J., & Marvin, R. S. (1992). *Attachment organization in preschool children: Procedures and coding manual*. Unpublished manuscript.

Cohen, S., & Willis, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin, 98,* 310–357.

Coughlin, E. K. (1995). American's dilemma. *The Chronicle of Higher Education, 42*(2), 10–11.

Cox, A. D., Puckering, C., Pound, A., & Mills, M. (1987). The impact of maternal depression in young people. *Journal of Child Psychology and Psychiatry, 28,* 917–928.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information processing mechanisms in children's social adjustment. *Psychological Bulletin, 115,* 74–101.

Cross Calvert, S., & McMahon, R. J. (1987). The treatment acceptability of a behavioral parent training program and its components. *Behavior Therapy, 18,* 165–179.

Crowther, J. K., Bond, L. A., & Rolf, J. E. (1981). The incidence, prevalence, and severity of behavior disorders among preschool-aged children in day care. *Journal of Abnormal Child Psychology, 9,* 23–42.

Cummings, E. M., Ballard, M., El-Sheikh, M., & Lake, M. (1991). Resolution and children's responses to interadult anger. *Developmental Psychology, 27,* 462–470.

Cummings, E. M., & Davies, P. (1994). *Children and marital conflict*. New York: Guilford Press.

Cummings, E. M., Simpson, K. S., & Wilson, A. (1993). Children's responses to interadult anger as a function of information about resolution. *Developmental Psychology, 29,* 978–985.

Cummings, E. M., & Zahn-Waxler, C. (1992). *Emotions and the socialization of aggression: Adults' angry behavior and children's arousal and aggression*. New York: Springer.

Cummings, J. S., Pellegrinia, D. S., Notarius, C. I., & Cummings, E. M. (1989). Children's responses to angry adult behavior as a function of marital distress and history of interparent hositility. *Child Development, 60,* 1035–1043.

Cunningham, C. E., Bennes, B. B., & Siegel, L. (1988). Family functioning, time allocation, and parental depression in families of normal and ADDH children. *Journal of Clinical Child Psychology, 17,* 169–179.

Cunningham, C. E., Bremner, R., & Boyle, M. (1995). Large group community-based parenting programs for families of preschoolers at risk for disruptive behaviour disorders: Utilization, cost effectiveness, and outcome. *Journal of Child Psychology and Psychiatry, 36,* 1141–1159.

Dadds, M. R., & McHugh, T. (1992). Social support and treatment outcome in behavioral family therapy for child conduct problems. *Journal of Consulting and Clinical Psychology, 60,* 252–259.

Dadds, M. R., Schwartz, M. R., & Sanders, M. R. (1987). Marital discord and treatment outcome in behavioral treatment of child conduct disorders. *Journal of Consulting and Clinical Psychology, 16,* 192–203.

Daro, D. (1988). *Confronting child abuse: Research for effective program design*. London: Free Press.

Denham, S. A. (1989). Maternal affect and child's social emotional competence. *American Journal of Orthopsychiatry, 59,* 368–376.

Denham, S. A., Zoller, D., & Couchold, E. A. (1994). Socialization of preschoolers' emotion understanding. *Developmental Psychology, 30,* 928–936.

Dishion, T. J., French, D. C., & Patterson, G. R. (1995).

The development and ecology of antisocial behavior. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology. Vol 2: Risk disorder and adaptation* (pp. 421–471). New York: Wiley.

Dishion, T. J., Patterson, G. R., & Kavanagh, K. (1992). *An experimental test of the coercion model: Linking theory, measurement, and intervention.* New York: Guilford Press.

Dodge, K. A., & Newman, J. P. (1981). Biased decision-making processes in aggressive boys. *Journal of Abnormal Psychology, 90*(4), 375–379.

Downey, G., & Coyne, J. C. (1990). Children of depressed parents: An integrated review. *Psychological Bulletin, 108,* 50–76.

Dumas, J. E., & Albin, J. B. (1986). Parent training outcome: Does active parental involvement matter? *Behavior Research and Therapy, 24*(2), 227–230.

Dumas, J. E., & Wahler, R. G. (1983). Predictors of treatment outcome in parent training: Mother insularity and socioeconomic disadvantage. *Behavioral Assessment, 5,* 301–313.

D'Zurilla, T. J., & Nezu, A. (1982). Social problem-solving in adults. In P. C. Kendall (Ed.), *Advances in cognitive behavioral research and therapy* (Vol. 1). New York: Academic Press.

Egeland, B., Kalkoske, M., Gottesman, N., & Erickson, M. F. (1990). Preschool behavior problems: Stability and factors accounting for change. *Journal of Child Psychology and Psychiatry, 31,* 891–909.

Eyberg, S. M. (1988). Parent–child interaction therapy: Integration of traditional and behavioral concerns. *Child and Family Behavior Therapy, 10,* 33–46.

Eyberg, S. M. (1992). Assessing therapy outcome with preschool children: Progress and Problems. *Journal of Clinical Child Psychology, 21*(3), 306–311.

Eyberg, S.M., & Boggs, S. R. (1989). Parent training for oppositional-defiant preschoolers. In C. E. Schaeffer & J. M. Brienmeister (Eds.), *Handbook of parent training* (pp. 105–132). New York: Wiley.

Eyberg, S. M., & Johnson, S. M. (1974). Multiple assessment of behavior modification with families. *Journal of Consulting and Clinical Psychology, 42,* 594–606.

Fantuzzo, J. W., DePaola, L. M., Lambert, L., Martino, T., Anderson, G., & Sutton, S. (1991). Effects of interpersonal violence on the psychological adjustment and competencies of young children. *Journal of Consulting and Clinical Psychology, 59,* 258–265.

Faraone, S. V., Biederman, J., Keenan, K., & Tsuang, M. T. (1991). Separation of *DSM-III* attention deficit disorder and conduct disorder: Evidence from a family genetic study of American child psychiatry patients. *Psychological Medicine, 21,* 1091–1121.

Feshbach, N. (1989). The construct of empathy and the phenomenon of physical maltreatment of children. In D. Cicchetti & V. Carlson (Eds.), *Child maltreatment: Theory and research on the causes and consequences of child abuse and neglect* (pp. 349–373). Cambridge, MA: Cambridge University Press.

Fisher, M., Rolf, J. E., Hasazi, J. E., & Cummings, L. (1984). Follow-up of a preschool epidemiological sample: Cross-age continuities and predictions of later adjustment with internalizing and externalizing dimensions of behavior. *Child Development, 55,* 137–150.

Fleischman, M. J. (1981). A replication of Patterson's "Intervention for boys with conduct problems." *Journal of Consulting and Clinical Psychology, 49,* 342–351.

Folkman, S., & Lazarus, R. S. (1988). Coping as a mediator of emotion. *Journal of personality and Social Psychology, 54*(3), 466–475.

Forehand, R., Breiner, J., McMahon, R. J., & Davies, G. (1981). Predictors of cross setting behavior change in the treatment of child problems. *Journal of Behavior Therapy and Experimental Psychiatry, 12*(4), 311–313.

Forehand, R., & Furey, W. M. (1985). Predictors of depressive mood in mothers of clinic-referred children. *Behavior Research and Therapy, 23*(4), 415–421.

Forehand, R., Furey, W. M., & McMahon, R. J. (1984). The role of maternal distress in a parent training program to modify child noncompliance. *Behavioral Psychotherapy, 12,* 93–108.

Forehand, R., & Kotchick, B. A. (1996). Cultural diversity: a wake -up call for parent training. *Behavior Therapy, 27,*187–206.

Forehand, R., & Long, N. (1986). *A long-term follow-up of parent training participants.* Unpublished conference proceedings, Chicago.

Forehand, R. L., & McMahon, R. J. (1981). *Helping the noncompliant child: A clinician's guide to parent training.* New York: Guilford Press.

Forehand, R., Rogers, T., McMahon, R. J., Wells, K. C., & Griest, D. L. (1981). Teaching parents to modify child behavior problems: An examination of some follow-up data. *Journal of Pediatric Psychology, 6*(3), 313–322.

Forehand, R., Steffe, M. A., Furey, W. A., & Walley, P. B. (1983). Mother's evaluation of a parent training program completed three and one-half years earlier. *Journal of Behavior Therapy and Experimental Psychiatry, 14,* 339–342.

Forehand, R., Sturgis, E. T., McMahon, R. J., Aguar, D., Green, K., Wells, K., & Breiner, J. (1979). Parent behavioral training to modify child noncompliance: Treatment generalization across time and from home to school. *Behavior Modification, 3,* 3–25.

Forehand, R., Wells, K., & Griest, D. L. (1980). An examination of the social validity of a parent training program. *Behavior Therapy, 11,* 488–502.

Forgatch, M. (1989). Patterns and outcome in family problem solving: The disrupting effect of negative emotion. *Journal of Marriage and the Family, 5*(1), 115–124.

Forgatch, M., Patterson, G., & Skinner, M. (1988). A mediational model for the effect of divorce in antisocial behavior in boys. In E. M. Hetherington & J. D. Arasteh (Eds.), *The impact of divorce, single parenting and stepparenting on children* (pp. 135–154). Hillsdale, NJ: Lawrence Erlbaum Associates.

Frick, P. J., & Jackson, Y. K. (1993). Family functioning and childhood antisocial behavior: Yet another reinterpretation. *Journal of Clinical Child Psychology, 22,* 410–419.

Frick, P. J., Kuper, K., Silverthorn, P., & Cotter, M. (1995). Antisocial behavior, somatization, and sensation-seeking behavior in mothers of clinic-referred children. *Journal of American Academy of Child and Adolescent Psychiatry, 34*(6), 805–812.

Frick, P. J., Lahey, B. B., Loeber, R., Stouthamer-Loeber, M., Christ, M. A., & Hanson, K. (1992). Familial risk factors to oppositional defiant disorder and conduct disorder: Parental psychopathology and maternal parenting. *Journal of Consulting and Clinical Psychology, 60,* 49–55.

Gecas, V. (1979). The influence of social class on socialization. In W. Bum (Ed.), *Contemporary theories about the family.* New York: Free Press.

Goutz, K. (1981). Children's initial aggression level and the effectiveness of intervention strategies in moderating television effects on aggression. In K. R. Goutz (Ed.), *Attention and social problem-solving as correlates of aggression in preschool males* (pp. 181–197).

Greenberg, M. T., Speltz, M. L., & DeKlyen, M. (1993). The role of attachment in the early development of disruptive behavior problems. *Development and Psychopathology, 5,* 191–213.

Griest, D. L., Forehand, R., Rogers, T., Breiner, J., Furey, W., & Williams, C. A. (1982). Effects of parent enhancement therapy on the treatment of outcome and generalization of a parent training program. *Behaviour Research and Therapy, 20*(5), 429–436.

Griest, D. L., Forehand, R., Wells, K. C., & McMahon, R. J. (1980). An examination of differences between nonclinic and behavior-problem clinic-referred children and their mothers. *Journal of Abnormal Psychology, 89*(3), 497–500.

Griest, D. L., & Wells, K. C. (1983). Behavioral family therapy with conduct disorders in children. *Behavior Therapy, 14,* 37–53.

Gross, D., Fogg, L., & Tucker, S. (1995). The efficacy of parent training for promoting positive parent–toddler relationships. *Research in Nursing and Health, 18,* 489–499.

Grych, J. H., & Finchman, F. D. (1990). Marital conflict and children's adjustment: A cognitive contextual framework. *Psychological Bulletin, 108,* 267–290.

Grych, J. H., & Finchman, F. D. (1992). Interventions for children of divorce: Toward greater integration of research and action. *Psychological Bulletin, 111,* 434–454.

Haley, J. (1976). *Problem-solving therapy*. New York: Colophon Books.

Hanf, C. (1970). *Shaping mothers to shape their children's behavior*. Portland, OR: University of Oregon Medical School.

Harkness, S., & Super, C. M. (1995). Culture and parenting. In M.H. Bornstein (Ed.), *Handbook of parenting: Biology and ecology of parenting* (Vol. 2, pp. 211–234). Mahwah, NJ: Lawrence Erlbaum.

Hashima, P. Y., & Amato, P. R. (1994). Poverty, social support, and parental behavior. *Child Development, 65,* 394–403.

Hawkins, J. D., Catalano, R. F., & Miller, Y. (1992). Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: Implications for substance abuse prevention. *Psychological Bulletin, 112,* 64–105.

Hay, D. F., Zahn-Waxler, C., Cummings, E. M., & Iannotti, R. J. (1992). Young children's views about conflict with peers: A comparison of the daughters and sons of depressed and well women. *Journal of Child Psychology and Psychiatry, 33,* 669–684.

Hembree-Kigin, T. L., & McNeil, C. B. (1995). *Parent–child interaction therapy*. New York: Plenum.

Henggler, S. W., Borduin, C. M., & Mann, B. J. (1993). Advances in family therapy: Empirical foundations. *Advances in Clinical Child Psychology, 15,* 207–241.

Hetherington, E. M., Cox, M., & Cox, R. (1982). *Effects of divorce on parents and children*. Hillsdale, NJ: Erlbaum.

Hinshaw, S. P., & Anderson, C. A. (1996). Conduct and oppositional defiant disorders. In E.J. Mash & R.A. Barkley (Eds.), *Child Psychopathology*. New York: Guilford Press.

Hoffman, L. W. (1984). Work, family, and socialization of the child. In R. Parke, R. Emde, H. McAdoo, & G. Sackett (Eds.), *Review of child development research: The family,* (Vol. 7). Chicago: University of Chicago Press.

Hops, H., Biglan, A., Sherman, L., Arthur, J., Friedman, L., & Osteen, R. (1987). Home observations of family interactions of depressed women. *Journal of Consulting and Clinical Psychology, 55,* 341–346.

Huesmann, L. R., Eron, L. D., Lefkowitz, M. M., & Walder, L. O. (1984). Stability of aggression over time and generations. *Developmental Psychology, 20,* 1120–1134.

Hutchings, B., & Mednick, S.A. (1977). Criminality in adoptees and their adoptive and biological parents: A pilot study. In S. A. Mednick & K. O. Christiansen (Eds.), *Biosocial bases of criminal behavior*. New York: Gardner Press.

Jennings, K. D., Stagg, V., & Conners, R. E. (1991). Social networks and mother interactions with their preschool children. *Child Development, 62,* 966–978.

Johnson, S. M., & Lovitz, G. K. (1974). The personal and marital adjustment of parents as related to observed child deviance and parenting behaviors. *Journal of Abnormal Child Psychology, 2,* 193–207.

Jouriles, E. N., Murphy, C. M., Farris, A. M., Smith, D. A., Richters, J. E., & Waters, E. (1991). Marital adjustment, parental disagreements about child rearing, and behavior problems in boys: Increasing the specificity of the marital assessment. *Child Development, 62,* 1424–1433.

Jouriles, E. N., Murphy, C. M., & O'Leary, K. D. (1989). Interspousal aggression, marital discord, and child problems. *Journal of Consulting and Clinical Psychology, 57*(3), 453–455.

Katz, L. F., & Gottman, J. M. (1994). Patterns of marital interaction and children's emotional development. In R. D. Parke & S. G. Kellam (Eds.), *Exploring family relationships with other social contexts* (pp. 49–74). Hillsdale, NJ: Lawrence Erlbaum.

Kazdin, A. (1985). *Treatment of antisocial behavior in children and adolescents*. Homewood, IL: Dorsey Press.

Kazdin, A. E. (1987). Treatment of antisocial behavior in children: Current status and future directions. *Psychological Bulletin, 102*(2), 187–203.

Kazdin, A. E., & Kolko, D. J. (1986). Parent psychopathology and family functioning among childhood firesetters. *Journal of Abnormal Child Psychology, 14*(2), 315–329.

Kazdin, A., & Mazurick, J. L. (1994). Dropping out of child psychotherapy: Distinguishing early and late dropouts over the course of treatment. *Journal of Consulting and Clinical Psychology, 62,* 1069–1074.

Kazdin, A. E., Siegel, J. C., & Bass, D. (1992). Cognitive problem-solving skills training and parent management training in the treatment of antisocial behavior in children. *Journal of Consulting and Clinical Psychology, 60,* 733–747.

Kendall, P. C. (1992). Childhood coping: Avoiding a lifetime of anxiety. *Behavior Change, 3,* 70–73.

King, N. J., Hamilton, D. I., & Ollendick, T. H. (1988). *Children's fears and phobias: A behavioral perspective*. Chichester, UK: Wiley.

Kochansaka, G., Kuczynski, L., Radke-Yarrow, M., & Welsh, J. D. (1987). Resolution of control episodes between well and affectively ill mothers and their young child. *Journal of Abnormal Child Psychology, 15,* 441–456.

Lahey, B. B., Hart, E. L., Pliszka, S., Applegate, B., & McBurnett, K. (1993). Neurophysiological correlates of conduct disorder: A rationale and review of current research. *Journal of Clinical Child Psychology, 22,* 141–153.

Lahey, B. B., Hartdagen, W. E., Frick, P. J., McBurnett, K., Connor, R., & Hynd, G. W. (1988). Conduct disorder: Parsing the confounded relationship between parental divorce and antisocial personality. *Journal of Abnormal Psychology, 97,* 334–337.

Lahey, B. B., Loeber, R., Hart, E., & Frick, P. J. (1995). Four-year longitudinal study of conduct disorder in boys: Patterns and predictors of persistence. *Journal of Abnormal Psychology, 104*(1), 83–89.

Lahey, B. B., Loeber, R. L., Quay, H. C., Frick, P. J., & Grimm, J. (1992). Oppositional defiant and conduct disorders: Issue to be resolved for DSM-IV. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*(3).

Landy, S., & Peters, R. D. (1991). Understanding and treating the hyperaggressive toddler. *Zero to Three, February,* 22–31.

Laski, K., Charlop, M. H., & Schreibman, L. (1988). Training parents to use the natural language paradigm to increase their autistic children's speech. *Journal of Applied Behavioral Analyses, 21,* 391–400.

Loeber, R. (1991). Antisocial behavior: More enduring than changeable? *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 393–397.

Lutzker, J. R. (1992). Developmental disabilities and child abuse and neglect: The ecobehavioral imperative. *Behavior Change, 9,* 149–156.

Lytton, H. (1990). Child and parent effects in boys' conduct disorder: A reinterpretation. *Developmental Psychology, 26,* 683–697.

Main, M., & Cassidy, J. (1987). *Reuion-based classifications of child–parent attachment organization at six-years of age*. Unpublished manuscript, University of California at Berkley.

McCord, J., McCord, W., & Thurber, E. (1962). Some effects of paternal absence on male children. *Journal of Abnormal and Social Psychology, 64,* 361–369.

McLoyd, V. C. (1990). The impact of economic hardship on black families and children: Psychological distress, parenting, and socioemotional development. *Child Development, 61,* 311–346.

McMahon, R. J., & Forehand, R. (1984). Parent training for the noncompliant child: Treatment outcome, generalization, and adjunctive therapy procedures. In R. F. Dangel & R. A. Polster (Eds.), *Parent training: Foundations of research and practice* (pp. 298–328). New York: Guilford Press.

McMahon, R. J., Forehand, R., Griest, D. L., & Wells, K. (1981). Who drops out of treatment during parent behavioral training? *Behavioral Counseling Quarterly, 1,* 79–85.

Meichenbaum, D., & Turk, D. (1987). *Facilitating treatment adherence: A practitioner's guidebook*. New York: Plenum.

Milich, R., & Dodge, K. A. (1984). Social information processing in child psychiatric populations. *Journal of Abnormal Child Psychology, 12*(3), 471–489.

Miller, G. E., & Prinz, R. J. (1990). Enhancement of social learning family interventions for childhood conduct disorder. *Psychological Bulletin. 108,* 291–307.

Minuchin, S. (1974). *Families and family therapy*. Cambridge, MA: Harvard University Press.

Offord, D. R., Alder, R. J., & Boyle, M. H. (1986). Prevalence and sociodemographic correlates of conduct disorder. *American Journal of Social Psychiatry, 6,* 272–278.

Offord, D. R., Boyle, M. H., & Szatmari, P. (1987). Ontario Child Health Study II: Six month prevalence of disorder and rates of service utilization. *Archives of General Psychiatry, 44,* 832–836.

O'Leary, K. D., & Emery, R. E. (1982). Marital discord and child behavior problems. In M. D. Levine & P. Satz (Eds.), *Middle childhood: Developmental variation and dysfunction* (pp. 345–364). New York: Academic Press.

Oltmanns, T. F., Broderick, J. E., & O'Leary, K. D. (1977). Marital adjustment and the efficacy of behavior therapy with children. *Journal of Consulting and Clinical Psychology, 45,* 724–729.

Patterson, G. R. (1975). *Families: Applications of social learning to family life*. Champaign, IL: Research Press.

Patterson, G. R. (1976). *Living with children: New methods for parents and teachers*. Champaign, IL: Research Press.

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia.

Patterson, G. R. (1986). Performance models for antisocial boys. *American Psychologist, 41*(4), 432–444.

Patterson, G. R., & Chamberlain, P. (1988). Treatment process: A problem at three levels. In L. C. Wynne (Ed.), *The state of art in family therapy research: Controversies and recommendations* (pp. 189–223). New York: Family Process Press.

Patterson, G. R., Chamberlain, P., & Reid, J. B. (1982). A comparative evaluation of a parent training program. *Behavior Therapy, 13,* 638–650.

Patterson, G. R., Cobb, J. A., & Ray, R. S. (1973). A social engineering technology for retraining the families of aggressive boys. In H. E. Adams & I. P. Unike (Eds.), *Issues and trends in behavior therapy* (pp. 139–210). Springfield, IL: Thomas.

Patterson, G. R., DeBaryshe, B. D., & Ramsey, E. (1989). A developmental perspective on antisocial behavior. *American Psychologist, 44*(2), 329–335.

Patterson, G. R., Dishion, T. J., & Chamberlain, P. (1993). *Outcomes and methodological issues relating to the treatment of antisocial children*. New York: Plenum.

Patterson, G. R., & Fleischman, M. J. (1979). Maintenance of treatment effects: Some considerations concerning family systems and follow-up data. *Behavior Therapy, 10,* 168–185.

Patterson, G. R., & Forgatch, M. S. (1985). Therapist behavior as a determinant for client noncompliance: A paradox for the behavior modifier. *Journal of Consulting and Clinical Psychology, 53*(6), 846–851.

Patterson, G. R., & Forgatch, M. S. (1995). Predicting future clinical adjustment from treatment outcome and process variables. *Psychological Assessment, 7*(2), 275–285.

Patterson, G. R., & Narrett, C. M. (1990). The development of a reliable and valid treatment programs for aggressive young children. *International Journal of Mental Health, 19,* 19–26.

Patterson, G. R., & Reid, J. B. (1973). Intervention for families of aggressive boys: A replication study. *Behaviour Research and Therapy, 11,* 383–394.

Patterson, G. R., Reid, J. B., Jones, R. R., & Conger, R. W. (1975). *A social learning approach to family intervention* (Vol. 1). Eugene, OR: Castalia.

Peed, S., Roberts, M., & Forehand, R. (1977). Evaluations of the effectiveness of a standardized parent training program in altering the interaction of mothers and their noncompliant children. *Behavior Modification, 1,* 323–350.

Porter, B., & O'Leary, K. D. (1980). Marital discord and childhood behavior problems. *Journal of Abnormal Child Psychology, 16,* 97–109.

Quay, H. C. (1986). Conduct disorders. In H. C. Quay & J. S. Werry (Eds.), *Psychopathological disorders of childhood* (pp. 35–72). New York: Wiley.

Quay, H. C. (1993). The psychopathology of undersocialized aggressive conduct disorder: A theoretical perspective. *Development and Psychopathology, 5,* 165–180.

Radke-Yarrow, M. (1991). Attachment patterns in children of depressed mothers. In C. M. Parke, J. Stevenson-Hinde, & P. Marris (Eds.), *Attachment across the life cycle*. London: Routledge.

Reid, J. B. (1987). *Therapeutic interventions in the families of aggressive children and adolescents,* Paper presented at the meeting of the Oreganizzato dalle Cattedre di Psicologia Clinica e delle Teorie di Personalita dell'Universita di Roma, Rome.

Richman, N., Stevenson, L., & Graham, P. J. (1982). *Preschool to school: A behavioural study*. London: Academic Press.

Richters, J. E., & Martinez, P. E. (1993). Violent communities, family choices, and children's chances: An algorithm for improving the odds. *Development and Psychopathology, 5,* 609–627.

Rickard, K. M., Forehand, R., Wells, K. C., Griest, D. L., & McMahon, R. J. (1981). Factors in the referral of children for behavioral treatment: A comparison of mothers of clinic-referred deviant, clinic-referred non-deviant and non-clinic children. *Behavior Research and Therapy, 19*(3), 201–205.

Robbins, M. S., Alexander, J. F., Newell, R. M., & Turner,

C. W. (1996). The immediate effect of reframing on client attitude in family therapy. *Journal of Family Psychology, 10*(1), 28–34.

Robins, L. N. (1981). Epidemiological approaches to natural history research: Antisocial disorders in children. *Journal of the American Academy of Child Psychiatry, 20,* 566–580.

Rogers, T. R., & Forehand, R. (1983). The role of depression in interactions between mothers and their clinic-referred children. *Cognitive Therapy and Research, 46,* 315–324.

Rose, S. L., Rose, S. A., & Feldman, J. (1989). Stability of behavior problems in very young children. *Development and Psychopathology, 1,* 5–20.

Rubin, K. H., & Krasnor, L. R. (1986). Social-cognitive and social behavioral perspectives on problem-solving. In M. Perlmutter (Ed.), *Cognitive perspectives on children's social and behavioral development, The Minnesota Symposia on Child Psychology* (Vol. 18, pp. 1–68). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rutter, M. (1986). *The developmental psychopathology of depression: Issues and perspectives.* New York: Guilford Press.

Rutter, M., & Giller, H. (1983). *Juvenile delinquency: Trends and perspectives.* Harmondsworth, UK: Penguin.

Rutter, M., Yule, B., Quinton, D., Rowlands, O., Yule, W., & Berger, M. (1974). Attainment and adjustment in two geographic areas: Some factors accounting for area differences. *British Journal of Psychiatry, 126,* 520–533.

Sanders, M. R. (1996). New directions in behavioral family intervention with children. In T. H. Ollendick & R. J. Prinz (Eds.), *Advances in clinical child psychology* (Vol. 18, pp. 283–320). New York: Plenum Press.

Sayger, T. V., Horne, A. M., & Glaser, B. A. (1993). Marital satisfaction and social learning family therapy for child conduct problems: Generalization of treatment effects. *Journal of Marital and Family Therapy, 19,* 393–402.

Sayger, T. V., Horne, A. M., Walker, J. M., & Passmore, J. L. (1988). Social learning family therapy with aggressive children: Treatment outcome and maintenance. *Journal of Family Psychology, 1,* 261–285.

Schaughency, E. A., & Lahey, B. B. (1985). Mothers' and fathers' perceptions of child deviance: Roles of child behavior, parental depression, and marital satisfaction. *Journal of Consulting and Clinical Psychology, 53,* 718–723.

Schmaling, K. B., & Jacobson, N. S. (1987). *The clinical significance of treatment gains resulting from parent training interventions for children with conduct problems: An analysis of outcome data.* Paper presented at the meeting of the Association for the Advancement of Behavior Therapy, Boston.

Schreibman, L., Kaneko, W. M., & Koegel, R. L. (1991). Positive effects of parents of autistic children: A comparison across two teaching techniques. *Behavior Therapy, 22,* 479–490.

Seligman, M. E. P. (1975). *Helplessness.* San Francisco: Freeman.

Slaby, R., & Guerra, N. (1988). Cognitive mediators of aggression in adolescent offenders: 1. Assessment. *Development Psychology, 24,* 580–588.

Smith, A. M., & O'Leary, S. G. (1995). Attributions and arousal as predictors of maternal discipline. *Cognitive Therapy and Research, 19*(4), 459–471..

Speiker, S. J., & Booth, C. L. (1988). Maternal antecedants of attachment quality. In J. Belsky & T. Nnezworski (Eds.), *Clinical implications of attachment* (pp. 95–134). Hillsdale, NJ: Erlbaum.

Spitzer, A., Webster-Stratton, C., & Hollinsworth, T. (1991). Coping with conduct-problem children: Parents gaining knowledge and control. *Journal of Clinical Child Psychology, 20,* 413–427.

Spivak, G., Platt, J. J., & Shure, M. B. (1976). *The problem solving approach to adjustment.* San Francisco: Jossey-Bass.

Spoth, R., & Redmond, C. (1995). Parent motivation to enroll in parenting skills programs: A model of family context and health belief predictors. *Journal of Family Psychology, 9,* 294–310.

Strain, P. S., Steele, P., Ellis, T., & Timm, M. A. (1982). Long-term effects of oppositional child treatment with mothers as therapists and therapist trainers. *Journal of Applied Behavior Analysis, 15,* 1163–1169.

Straus, M. A., & Gelles, R. (1986). Societal change and change in family violence from 1975–1985 as revealed by two national surveys. *Journal of Marriage and the Family, 48,* 465–479.

Strayhorn, J. M., & Weidman, C. S. (1989). Reduction of attention deficit and internalizing symptoms in preschoolers through parent–child interaction training. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 888–896.

Stroufe, L. A., & Fleeson, J. (1986). Attachment and the construction of relationships. In W. Hartup & Z. Rubin (Eds.), *Relationships and development* (pp. 51–71). Hillsdale, NJ: Erlbaum.

Taylor, T. K., Schmidt, F., & Hodgins, C. (1996). *A comparison of two treatments on young conduct problem children.* Paper presented at the American Psychological Association, Toronto, Canada.

Teti, D. M., Gelfand, D. M., & Messinger, D. S. (1995). *Developmental Psychology, 31,* 364–376.

Trickett, P. K., Aber, J. L., Carlson, V., & Cicchetti, D. (1991). Relationship of socioeconomic status to the etiology and developmental sequelae of physical child abuse. *Developmental Psychology, 27,* 148–158.

Wahler, R. G. (1975). Some structural aspects of deviant child behavior. *Journal of Applied Behavioral Analysis, 8,* 27–42.

Wahler, R. G. (1976). Deviant child behavior within the family: Developmental speculations and behavior change strategies. In H. Leitenberg (Ed.), *Handbook of behavior modification and behavior therapy.* Englewood Cliffs, NJ: Prentice-Hall.

Wahler, R. G. (1980). The insular mother: Her problems in parent–child treatment. *Journal of Applied Behavior Analysis, 13*(2), 207–219.

Wahler, R. G., & Afton, A. D. (1980). Attentional processes in insular and noninsular Mothers. *Child Behavior Therapy, 2,* 25–41.

Wahler, R. G., Cartor, P. G., Fleischman, J., & Lambert, W. (1993). The impact of synthesis teaching and parent training with mothers of conduct disordered children. *Journal of Abnormal Child Psychology, 12,* 425–440.

Wahler, R. G., & Dumas, J. E. (1984). Changing the observational coding styles of insular and noninsular mothers: A step toward maintenance of parent training effects. In R. F. Dangel & R. A. Polster (Eds.), *Parent training: Foundations of research and practice* (pp. 379–416). New York: Guilford Press.

Wahler, R. G., & Dumas, J. E. (1989). Attentional problems in dysfunctional mother–child interactions: An interbehavioral model. *Psychological Bulletin, 105,* 116–130.

Wahler, R. G., & Fox, J. J. (1982). Response structure in deviant child–parent relationships: Implications for family therapy. *Nebraska Symposium on Motivation, 29*(1), 1–46.

Wahler, R. G., & Sansbury, L. E. (1990). The monitoring skills of troubled mothers: Their problems in defining child deviance. *Journal of Abnormal Child Psychology, 18*(5), 577–589.

Walker, J. L., Lahey, B. B., Hynd, G. W., & Frame, C. L. (1987). Comparison of specific patterns of antisocial behavior in children with conduct disorder with or without coexisting hyperactivity. *Journal of Consulting and Clinical Psychology, 55*(6), 910–913.

Webster-Stratton, C. (1981a). Modification of mothers' behaviors and attitudes through videotape modeling group discussion program. *Behavior Therapy, 12,* 634–642.

Webster-Stratton, C. (1981b). Videotape modeling: A method of parent education. *Journal of Clinical Child Psychology, 10*(2), 93–98.

Webster-Stratton, C. (1982a). The long term effects of a videotape modeling parent training program: Comparison of immediate and 1-year followup results. *Behavior Therapy, 13,* 702–714.

Webster-Stratton, C. (1982b). Teaching mothers through videotape modeling to change their children's behaviors. *Journal of Pediatric Psychology, 7*(3), 279–294.

Webster-Stratton, C. (1984). Randomized trial of two parent-training programs for families with conduct-disordered children. *Journal of Consulting and Clinical Psychology, 52*(4), 666–678.

Webster-Stratton, C. (1985a). The effects of father involvement in parent training for conduct problem children. *Journal of Child Psychology and Psychiatry, 26*(5), 801–810.

Webster-Stratton, C. (1985b). Predictors of treatment outcome in parent training for conduct disordered children. *Behavior Therapy, 16,* 223–243.

Webster-Stratton, C. (1988). Mothers' and fathers' perceptions of child deviance: Roles of parent and child behaviors and parent adjustment. *Journal of Consulting and Clinical Psychology, 56*(6), 909–915.

Webster-Stratton, C. (1989). Systematic comparison of consumer satisfaction of three cost-effective parent training programs for conduct problem children. *Behavior Therapy, 20,* 103–115.

Webster-Stratton, C. (1990a). Long-term follow-up of families with young conduct problem children: From preschool to grade school. *Journal of Clinical Child Psychology, 19*(2), 144–149.

Webster-Stratton, C. (1990b). Stress: A potential disruptor of parent perceptions and family interactions. *Journal of Clinical Child Psychology, 19*(4), 302–312.

Webster-Stratton, C. (1992a). *The incredible years: A trouble-shooting guide for parents of children ages 3–8 years.* Toronto, Canada: Umbrella Press.

Webster-Stratton, C. (1992b). Supporting your child's education: parents and children videotape—program 8. Seth Enterprises, Seattle, WA.

Webster-Stratton, C. (1994). Advancing videotape parent training: A comparison study. *Journal of Consulting and Clinical Psychology, 62*(3), 583–593.

Webster-Stratton, C. (1995). *Preventing conduct problems in Head Start children: Preliminary short term results.* Paper presented at the Society for Research in Child Development, Indianapolis, IN.

Webster-Stratton, C. (1996a). Early onset conduct problems: Does gender make a difference? *Journal of Consulting and Clinical Psychology, 64,* 540–551.

Webster-Stratton, C. (1996b). *Relationship between marital problem-solving and communication skills and child behaviors.* Unpublished manuscript, University of Washington, Seattle.

Webster-Stratton, C. (1997a). From parent training to community building. *Journal of Contemporary Services, 78,* 156–171.

Webster-Stratton, C. (1997b). Treating children with early-onset conduct problems: A comparison of child and parent training interventions. *Journal of Consulting and Clinical Psychology, 65*(1), 93–109.

Webster-Stratton, C., & Hammond, M. (1988). Maternal depression and its relationship to life stress, perceptions of child behavior problems, parenting behaviors and child conduct problems. *Journal of Abnormal Child Psychology, 16*(3), 299–315.

Webster-Stratton, C., & Hammond, M. (1990). Predictors of treatment outcome in parent training for families with conduct problem children. *Behavior Therapy, 21,* 319–337.

Webster-Stratton, C., & Hancock, L. (in press). Parent training: Content, methods and processes. In E. Schaefer (Ed.), *Parent training*. New York: Wiley.

Webster-Stratton, C., & Herbert, M. (1993). What really happens in parent training? *Behavior Modification, 17*(4), 407–456.

Webster-Stratton, C., & Herbert, M. (1994). *Troubled families—Problem children: Working with parents: A collaborative process.* Chichester, UK: Wiley.

Webster-Stratton, C., Hollinsworth, T., & Kolpacoff, M. (1989). The long-term effectiveness and clinical significance of three cost-effective training programs for families with conduct-problem children. *Journal of Consulting and Clinical Psychology, 57*(4), 550–553.

Webster-Stratton, C., Kolpacoff, M., & Hollinsworth, T. (1988). Self-administered videotape therapy for families with conduct-problem children: Comparison with two cost-effective treatments and a control group. *Journal of Consulting and Clinical Psychology, 56*(4), 558–566.

Wells, K. C., & Egan, J. (1988). Social learning and systems family therapy for childhood oppositional disorder: Comparative treatment outcome. *Comprehensive Psychiatry, 29*(2), 138–146.

White, A. G., & Bailey, J. S. (1990). Reducing disruptive behaviors of elementary physical education students with sit and watch. *Journal of Applied Behavior Analyses, 23*(3), 353–359.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.09
Play Therapy

SANDRA W. RUSS
Case Western Reserve University, Cleveland, OH, USA

5.09.1 INTRODUCTION 221
5.09.2 THEORETICAL UNDERPINNINGS 222
5.09.2.1 Conceptualizations from Child Therapy 222
5.09.2.2 Conceptualizations from Child Development 223
5.09.3 HISTORY OF PLAY THERAPY 223
5.09.4 ASSESSMENT AND CLINICAL FORMULATION 224
5.09.4.1 Assessing Children's Play 225
5.09.5 INTERVENTION PROCEDURES AND PROPOSED MECHANISMS OF CHANGE 226
5.09.5.1 Issues in Using Play in Therapy 226
5.09.5.1.1 How to get started 226
5.09.5.1.2 Kinds of play materials 226
5.09.5.1.3 How much to engage in play 227
5.09.5.1.4 How much to interpret 227
5.09.5.1.5 How to set limits 227
5.09.5.2 Mechanisms of Change in Therapy 227
5.09.5.3 Mechanisms of Change in Different Types of Therapy 228
5.09.5.3.1 Psychodynamic play therapy 228
5.09.5.3.2 Cognitive-behavioral play therapy 230
5.09.5.4 Change Processes in Play in Child Development 231
5.09.5.4.1 Play and cognitive processes 231
5.09.5.4.2 Play and affective processes 232
5.09.5.4.3 Play and adjustment 235
5.09.5.4.4 Meta-analysis 235
5.09.6 RESEARCH FINDINGS 236
5.09.6.1 Psychotherapy Outcome Research 236
5.09.6.2 Play Intervention Research 237
5.09.7 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 239
5.09.8 REFERENCES 240

5.09.1 INTRODUCTION

Play therapy is in a state of transition. Play has been a part of therapy with children since Melanie Klein (1927) and Anna Freud (1927) first began using play techniques in child therapy in the 1930s. As of 1992, play in some form was used in child therapy by a majority of clinicians, as reported by Koocher and D'Angelo (1992). They stated that "play-oriented therapy remains the dominant and most enduring approach to child treatment . . . practiced by clinicians" (p. 458). At the same time, play therapy is being increasingly criticized, along with many types of child therapy, for not having demonstrated treatment efficacy. In truth, there is not a strong empirical base for the effectiveness of play therapy (Phillips, 1985).

Suffice it to state, the empirical base of play therapy has not grown along with the clinical application of play therapy.

Nevertheless, there is an empirical literature in various areas of child psychology that supports the importance of play in child development. Play is involved in the development of many cognitive, affective, and personality processes that are important for adaptive functioning. Unfortunately, the field has not yet learned how to apply this accumulating knowledge base to play therapy. We know much about the role of play in child development, but have not utilized this knowledge in developing intervention and prevention programs that can be evaluated empirically. Research programs need to be developed that expand our knowledge about how play effects change and how we can facilitate play. The main point in this chapter is the thesis that play is an extremely worthwhile intervention approach with children and should be extensively investigated so that we can help children utilize this resource more effectively.

5.09.2 THEORETICAL UNDERPINNINGS

The theoretical underpinnings of play therapy come from two different sources. The first is the theory from the child psychotherapy literature. The second source is the theory and empirical work from the child development literature.

5.09.2.1 Conceptualizations from Child Therapy

In the child therapy literature, three broad functions of play emerge as important in therapy. First, play is a natural form of expression in children. Chethik (1989) refers to the language of play. Children use play to express feelings and thoughts. Chethik states that play emerges from the child's internal life and reflects the child's internal world. Therefore, children use play to express affect and fantasy and, in therapy, to express troubling and conflict-laden feelings. The expression of feelings itself, sometimes termed catharsis, is thought to be therapeutic (Axline, 1947; A. Freud, 1965; Moustakas, 1953). The therapist facilitates this process by giving permission for feelings to be expressed and by labeling the affect. By labeling the affect, the therapist helps to make the feeling less overwhelming and more understandable. In addition, the child feels more accepted as a whole person by the therapist and, in turn, is thought to become more self-accepting.

The child also uses this language of play to communicate with the therapist. It is essential that the therapist understand these communications, so that the therapeutic relationship can develop (Chethik, 1989). The therapist actively labels, empathizes and interprets the play, which in turn, helps the child feel understood (Russ, 1995). For many children, this feeling of empathy from the therapist facilitates change in their interpersonal representations and interpersonal functioning. The importance of expression through play and communication with the therapist is thought to be important by both psychodynamic and client-centered or person-centered approaches.

A second major function of play is as a vehicle for the occurrence of insight and working through. The conceptualization of this function of play is a psychodynamic one. Psychodynamic theory views the emotional resolution of conflict or trauma as a major mechanism of change in child therapy. Children reexperience major developmental conflicts or situational traumas in therapy. Many of these conflicts are expressed in play. The play process itself has been thought of as a form of conflict resolution. For example, Waelder (1933) has described the play process as one in which the child repeats an unpleasant experience over and over until it becomes manageable.

Freedheim and Russ (1992) described the slow process of gaining access to conflict-laden material and playing it out until the conflict is resolved. Erickson (1963) presented the concept of mastery, in which the child uses play to gain mastery over traumatic events and everyday conflicts. During this process, the therapist labels and interprets the play. Although there is controversy in the psychodynamic literature about how much interpretation should occur (Klein, 1955; A. Freud, 1966/1937), there is general agreement that working through and mastery are important mechanisms of change in play therapy.

A third major function of play in therapy is that of providing opportunities to practice with a variety of ideas, behaviors, interpersonal behaviors, and verbal expressions. Because play is occurring in a safe environment, in a pretend world, with a permissive, nonjudgmental adult, the child can try out a variety of expressions and behaviors without concern about real-life consequences. In some forms of play therapy, the therapist is quite directive in guiding the child to try new behaviors. For example, Knell (1993) developed a cognitive behavioral play therapy approach that actively uses modeling techniques and a variety of cognitive behavioral techniques.

It is important to point out that although these three major functions of play occur in normal play situations, the therapist builds on these normal functions by enhancing the play experience. The therapist creates a safe environment, gives permission for play to occur, actively facilitates play, and labels the thoughts and feelings expressed. For the psychodynamic therapist, interpretation specifically aids conflict-resolution. Because there are so many individual differences in play skills and abilities in children, there are many different kinds of play therapy techniques that are utilized.

5.09.2.2 Conceptualizations from Child Development

Conceptualizations from child development theory and research are relevant to the area of play therapy. Play is thought to facilitate both cognitive and emotional development. Children's play is related to a number of cognitive and affective processes that are important in child development (Russ, 1995). There is evidence that play facilitates cognitive and emotional development (see D. Singer & J. L. Singer, 1990, and Russ, 1993, 1995 for reviews). In play therapy, the therapy process uses these change processes, although not yet in any kind of systematic fashion. Research in this area is reviewed later in this chapter.

Play has been linked to the development of cognitive processes important in logical problem solving and in creative problem solving (Singer & Singer, 1990). For example, play has been related to and has facilitated insight ability (Vandenberg, 1980) and divergent thinking ability (Dansky, 1980; Russ & Grossman-McKee, 1990). Both cognitive and affective processes, important in creative thinking, are facilitated through play (Russ, 1993, 1995). In therapy, as both types of problem solving abilities develop, problem solving in everyday life can be enhanced.

Play is also thought to be related to general adjustment. Again, both cognitive and affective processes, important in adjustment, are developed in play. Play helps children learn to express and regulate their emotions. Play helps develop cognitive structures that, in turn, help children to integrate and modulate their emotions.

Children also use play to practice with and try out various coping strategies for different situations (Christiano & Russ, 1996). Children role play different scenarios and generate different solutions to everyday problems and stressful situations. For example, a child might try out different ways to handle teasing from other children or prepare for a frightening situation. In the Christiano and Russ (1996) study, the ability to express emotion and use good fantasy in play was related to the amount and variety of coping efforts children made in a dental situation. Good players also reported less distress.

Traditionally, the play therapist has utilized all of these conceptualizations and mechanisms of change in working with children. The choice of specific techniques and mechanisms of change has depended upon the conceptualization of the case. What is needed for the future are play intervention strategies based on research that are matched to specific types of child populations.

5.09.3 HISTORY OF PLAY THERAPY

The history of play therapy is intertwined with the history of child therapy, and play has been used in different ways in different theoretical approaches (Kessler, 1988). Play was first used in child therapy by therapists in the psychoanalytic tradition. Tuma and Russ (1993) reviewed the psychoanalytic literature in detail. Psychoanalytic techniques were adapted to children by Hug-Hellmuth (1921, 1924), A. Freud (1927), and Burlingham (1932). Play was used to substitute for free association. In addition, the therapist was more responsive and gratifying to the child than the therapist would be with adults, and the child therapist actively worked to develop a positive attachment (A. Freud, 1946). Melanie Klein saw the importance of the communication value of play. She suggested that play for the child was the same as free association for the adult. She advocated active and direct interpretation of the unconscious processes expressed in play. Therefore, the therapist would continually interpret the child's play. A. Freud (1966/1937) also viewed play as a direct expression of fantasy and instincts in a less disguised and more accessible form than in adults. A. Freud was more cautious in her interpretations of the meaning of the child's play. She was also more respectful of the child's defenses than was Klein, and encouraged greater participation by the child. She thought that it was important that therapy be a positive experience for the child and that the child want to come to therapy.

The psychoanalytic approach has evolved into the psychodynamic approach. Psychodynamic therapists base their interventions and techniques on psychoanalytic principles, but therapy is shorter, less frequent (once a week rather than 4 or 5 times a week), has more focused and immediate goals, and is more flexible in incorporating a variety of therapeutic techniques (Tuma & Russ, 1993). Play remains a core part of the therapy process. The use of play in therapy and in child development is a great legacy of the psychoanalytic tradition.

The client-centered and relationship approaches to child treatment also utilized play. Axline (1947), in her nondirective approach, has the therapist focus on play as a major form of communication for and with the child. The therapist strives to understand and empathize with the child's issues. Interpretation of underlying dynamics or impulses is rare, however. The therapist trusts the child's developmental process and striving for self-development. A wonderful introduction to the use of play in child therapy, for beginning therapists of any theoretical perspective, is Axline's *Dibbs in Search of Self* (1964). Moustakas (1953) was another lending theorist in this area who stressed the importance of expression of feelings in play and the importance of the relationship between the child and therapist. Moustakas discussed the importance of the child and therapist experiencing each other (Moustakas, 1992). Genuineness in the relationship is an important aspect of therapy for Moustakas.

Although not usually associated with cognitive behavior therapy approaches, play has been used as a tool within that framework (Kessler, 1988). Meichenbaum (1974) thought that imagery and fantasy could be used to teach children self-control. Knell (1993) has developed cognitive behavioral play therapy techniques that use principles of modeling and reinforcement.

In addition, play techniques have been used in a variety of more specific approaches, such as Gardener's (1971) mutual story-telling techniques and Levy's (1938) release therapy for children who have experience trauma.

Today, as Koocher and D'Angelo (1992) stated, play approaches are the dominant treatment approaches in working with children. Play is used with a variety of child populations, in short-term and long-term approaches, and by a variety of theoretical approaches. Trends in child psychotherapy apply to play therapy as well. Therapy is becoming more specific, more focused, more active, and more theoretically integrated (Freedheim & Russ, 1992).

5.09.4 ASSESSMENT AND CLINICAL FORMULATION

One purpose of psychological assessment is to obtain a developmental picture of the child in specific areas of cognitive and personality functioning. Assessment determines how the child is functioning in areas of basic cognitive processes, object relations, capacity for empathy, internalized conflicts, capacity for delay of gratification, self-esteem, and coping resources (Russ, in press). Ultimately, assessment should help with choosing the optimal form of intervention. Play is frequently a part of the assessment process.

When one has the luxury of time, a battery of psychological tests combined with a play or interview session will usually give the fullest view of a child. Of course, information from the individual assessment of the child should be combined with multiple sources of information from the parents, school, and community. Because managed care has resulted in pressures to move more quickly in clinical practice, one must select carefully which assessment methods to use. Only the most reliable and empirically validated measures should be used.

The fact that most child psychologists use tests that tap different psychological processes is reflected in the results of different surveys of child practitioners. Tuma and Pratt (1982) found that most psychologists in their survey used intelligence tests (primarily the Wechsler Intelligence Scales), achievement measures, and projective tests such as the Rorschach, TAT, and drawings. Archer, Maruish, Imhof, and Piatrowski (1991), in a survey focusing on psychological assessment of adolescents, found that the most frequently used tests were the Wechsler Intelligence Scales, the Rorschach (ranked second overall), Bender-Gestalt, TAT, Sentence Completion, and the MMPI. Of those respondents who used a standard assessment battery (139 of 165), the most frequently used measures were the Wechsler, Rorschach, and TAT.

When using a standard battery, different processes are measured by different tests (Tuma & Russ, 1993). Measures of intelligence and perceptual motor functioning (such as the Bender-Gestalt) give information about basic cognitive functioning. Even if there is not a specific question about school performance, the intelligence test is valuable because it assesses how well the child is functioning in a variety of cognitive and perceptual areas. Intelligence tests provide an anchor for a developmental assessment.

Projective tests give information about the personality structure and personality dynamics of the child. Play and interview information are useful for this purpose as well. Frequently, the Rorschach and a story-telling projective technique (e.g., the TAT, Children's Apperception Test or CAT, and Roberts Apperception Test for Children) will be used together.

The Rorschach is especially valuable in describing personality structure (Exner, 1986). Exner discussed how the Rorschach reveals the organizing principles that the individual uses to deal with the world. This is true of children as well as adults. The Rorschach assesses reality testing, the capacity to perceive what is commonly perceived in situations, and how affect influences cognitive functioning. It also assess how the child experiences the world in terms of availability of affect, capacity to experience negative affects, and internal representations of other people.

Projective tests that involve telling stories complement the Rorschach by assessing personality dynamics. The child's wishes, fears, conflicts, and specific areas of emotional distress can emerge. Increasingly available are projective tests containing stimuli appropriate for younger children and ethnic minorities. Drawings and fantasy play can also be used as diagnostic instruments which are especially valuable in revealing personality dynamics.

Projective tests are being rediscovered with the growing focus on the development of object relations as a key element in a child's development. Projective tests are especially suited to assessing object relations. In a review of the literature, Stricker and Healey (1990) defined object relations as concerning "the cognitive, affective, and emotional processes that mediate interpersonal functioning in close relationships" (p. 219). Understanding the developmental level of the child's object relations gives clues to the interpersonal world of the child and offers a framework for explaining interpersonal difficulties the child may be having. Stricker and Healey (1990) also reviewed the major Rorschach-based and TAT-based object relations scales. These scales are relatively new and are still being validated. Interpersonal content has always been part of projective tests, but these new scales and the growing theoretical literature on object relations enable us to develop a more comprehensive understanding of this area by means of a more systematic evaluation of this variable.

Projective tests have been criticized for their lack of demonstrated validity (Dawes, 1994). However, a growing body of construct validity studies on the Rorschach (Weiner, 1996) and story-telling techniques (Westen, 1991) are available and should serve as a guide for when projective tests are appropriate.

Observer rating scales and self-report instruments round out the picture by offering well-validated diagnostic information about the child, compared with their same-aged peers.

5.09.4.1 Assessing Children's Play

Assessing children's play is a valuable addition to a standard test battery. Although many young children cannot readily engage in an interview, they can engage in play. Until the early 1990s there were very few standardized measures of children's play available. Schaefer, Gitlen, and Sangrund (1991) have reviewed the growing number of current play measures. As these play measures become better known and as their reliability and validity is ascertained, we can begin to assess play in a systematic way prior to, during, and after treatment. The development of play assessment instruments is an important step for the field. A formal assessment of play could tell us how well the child plays; assess how well the child could use play in therapy; identify the important conflict areas; assess how well affect is expressed and is regulated; determine how constricted the child's fantasy is; assess the development of internal representations; and assess the child's overall cognitive organization. Some of the information would supplement other test information and some would be unique. For example, the best way to determine how well a child can use play in therapy is to observe the child at play.

One possible instrument for assessing play is the Affect in Play Scale (Russ, 1987, 1993). In order to study affect in play and its correlate the Affect in Play Scale was developed by Russ and her students. This scale consists of a standardized puppet play task and specific criteria for a rating scale. The play task utilizes two neutral looking puppets, one boy and one girl and three small blocks that are laid out on a table. It is administered individually to the child. The play task is appropriate for children from 5 to 10 years of age. The instructions ask the child to play with the puppets any way he or she likes for five minutes. The play is videotaped. The instructions can be altered to pull for specific affective themes such as aggression or happiness. An extensive manual has been developed for rating the play. The Affect in Play Scale measures the amount of affect, type of affect, variety of affect, intensity of affect, comfort, quality of fantasy, imagination, and affective integration in the play. There are 11 possible affective categories: Happiness/Pleasure; Anxiety/Fear; Sadness/Hurt; Frustration/Displeasure; Nurturance/Affection; Aggression; Oral; Oral Aggression; Anal; Sexual; and Competition. Any verbal or physical expression between the puppets that reflects an affective category is rated as one affective unit. For example, if one puppet says to the other "I don't like you" it would be scored for one unit of aggressive content.

The Affect in Play Scale has good interater reliability, internal consistency, and good stability of scores across different samples of children. The scale has been used with public school, parochial school, middle-class suburban and inner city populations and has a growing body of validity studies. A review of the Affect in Play Scale and the manual can be found in Russ (1993). This scale could also be used as a measure of change in play intervention research and in play psychotherapy.

It is important to note that for many of the play measures available, including the Affect in Play Scale, play does not relate to traditional IQ measures (Russ & Grossman-McKee, 1990; Singer, 1973). Most of this research has been with children of normal intelligence. Children with major cognitive deficits should also show serious deficits in play skills. Nevertheless, it is important to have measures of children's fantasy and cognitive organization that are relatively independent of intelligence.

Play assessment is well-suited to a repeated measure design, (Russ, in press-b). Changes in the quality of play or the themes in play should occur during therapy. Different types of changes should occur for different types of children. For example, the constricted child should express more affect in play after therapy. The acting-out child should show better modulation of affect in play after treatment. Play measures should be sensitive to change processes in psychotherapy. Shirk and Russell (1996) have stressed the importance of developing measures that are sensitive to specific processes that are being addressed in psychotherapy. Play assessment can focus on cognitive, affective, and interpersonal processes important in psychotherapy.

The general assessment process should help the therapist formulate a diagnostic picture and assess the child's motivation for treatment, capacity to establish a good relationship, and ability to use play and/or verbalization. It should help decide what type of intervention to use, what specific process or symptom should be targeted, and treatment goals. Major mechanisms of change and techniques to bring about those changes should be identified. This leads to the discussion of intervention procedures and mechanisms of change.

5.09.5 INTERVENTION PROCEDURES AND PROPOSED MECHANISMS OF CHANGE

The concept of play psychotherapy is a bit of a misnomer because most psychotherapy with children is a mix of play and talk (Russ, 1995). Play is a tool which can be used in therapy for a variety of purposes. If and how play is used with a particular child in therapy depends upon the child's ability to use play, developmental level, age, ability to verbalize, and the overall treatment approach. In general, play therapy occurs within the context of the overall treatment plan.

In most forms of play therapy, the child and therapist meet individually once a week for 45–50 minute sessions. The mutual agreement between the child and the therapist is that the therapist is there to help the child express feelings and thoughts, understand causes of behavior, and form a relationship with the therapist (Freedheim & Russ, 1992). Traditionally, in both psychodynamic and client-centered approaches, the child structures the therapeutic hour by choosing the topics, forms of play, and setting the pace of therapy. In most cases, individual work with the child is only one part of the treatment program. Parent guidance and education, family sessions, and work with the school usually occur simultaneously with individual child therapy.

5.09.5.1 Issues in Using Play in Therapy

A number of practical issues arise in working with children in play therapy. Ideally, a psychological assessment will have been carried out so that the therapist has identified treatment goals and has developed a treatment plan. The therapist should have determined how much play will be utilized in the therapy, as well as the nature of the play to be encouraged. Usually, play approaches should be considered for children from 4 to 10 years of age.

5.09.5.1.1 How to get started

Many children need help initiating the play process. The therapist usually starts by telling the child that they can play or talk, and shows the child the toys and play materials that are available. Although many children go right to it, many others are reticent. The therapist might tell the child to pick one thing, and start with that or the therapist could pick something for the child. Clay and drawing material are good starters. As a last resort, the therapist might start with something themselves, and have the child join in.

5.09.5.1.2 Kinds of play materials

Most therapists agree that unstructured play materials that can encourage the use of fantasy and imagination are ideal. Toys that leave much room for individual expression are most appropriate for play therapy. Examples of relatively unstructured toys and material are: clay, crayons, cars, trucks, puppets, dolls, doll houses, and legos. Games such as checkers do not encourage free expression, and are not ideal for traditional play therapy. The child is often asked for suggestions of other toys that might be useful and encouraged to bring toys in from home. Different children use different media for expression and it is important to have a variety of items and to individually tailor what is available.

5.09.5.1.3 How much to engage in play

One recurring dilemma is how much the therapist should engage in the child's play. The amount of direction and activity by the therapist depends upon the general theoretical approach. In more traditional client-centered and psychodynamic approaches, the therapist tries not to play with the child, but rather to observe and comment on the child's play. Many children will eventually become comfortable playing in this way. Some children need more engagement by the therapist. With those children, the therapist tries to follow the lead of the child. The therapist might put on one of the puppets and play, but follow the lead of the child in choosing the topic and setting the dialogue.

In more directive approaches, such as cognitive-behavioral play therapy (Knell, 1993), the therapist will play in a very directive fashion. The therapist might put on a play, modeling adaptive coping strategies for an issue that the child is dealing with.

5.09.5.1.4 How much to interpret

An optimal amount of interpretation of the child's play would facilitate the play process and help the child understand his or her thoughts, feelings, and behavior. Too much interpretation will stop the play process. Most therapists today would agree with A. Freud (1976) that the child's defenses should be respected. Timing of interpretations is as important in child therapy as in adult therapy.

5.09.5.1.5 How to set limits

For play to be a safe mode of expression, limits about how to play are essential. If rules need to be set, then they should be set. Toys cannot be broken and the therapist cannot be a target for affective expression. Alternative modes of expression through play should be facilitated. For example, the child can be encouraged to verbalize anger at the therapist or to have the puppets fight it out in a pretend mode.

5.09.5.2 Mechanisms of Change in Therapy

How one uses play in the therapy depends upon how one conceptualizes the specific mechanisms of change to be used in the therapy. Freedheim and Russ (1983, 1992) identified six major mechanisms of change that occur in individual child psychotherapy. These mechanisms were based on those identified by Applebaum (1978) and Garfield (1980) in the adult literature. Different mechanisms of change are utilized in different types of psychodynamic psychotherapy with different types of childhood disorders. However, there is rarely a pure type of psychotherapy and frequently all of these mechanisms may occur in any one case. These mechanisms of change are thought to be universal and to cut across various theoretical approaches to psychotherapy. The specific role of play in these mechanisms was discussed by Russ (1995).

(i) Catharsis and Labeling of Feeling. Through talk and play, children express feelings and release emotion. This release of emotion has long been thought to be therapeutic (Axline, 1947, A. Freud, 1965; Moustakas, 1953). In addition, by labeling affect, the therapist helps to make the feeling less overwhelming and more understandable. Often, the labeling of affect occurs during pretend play. By saying to the child that the puppet is feeling angry, the therapist connects a label to a feeling state. Words help to put feelings into a context for the child, thus making the feelings less overwhelming.

(ii) Corrective Emotional Experience. The therapist accepts the child's feeling and thoughts. Often, the child's learned expectations are not met and a corrective emotional experience occurs (Kessler, 1966). For example, the automatic connection between the child's angry feelings towards father and resultant anxiety should gradually decrease as the therapist helps the child accept the feeling and understand the reasons for the anger. The therapist is not punitive, and is accepting of the child's having angry feelings. Often, these feelings are expressed through play.

(iii) Insight and Working Through. The emotional resolution of conflict or trauma is a major mechanism of change in play psychotherapy. One goal of the therapist when utilizing this mechanism of change is to help the child reexperience major developmental conflicts or situational traumas in therapy. Frequently, play is the vehicle for this working through process. Cognitive insight into origins of feelings and conflicts, causes of symptoms, links between thoughts, feelings, and actions is a goal of psychotherapy when underlying conflicts are a major issue (Sandler, Kennedy, & Tyson, 1980; Shirk & Russell, 1996). Verbal labeling of unconscious impulses, conflicts, and causes of behavior helps lend higher order reasoning skills to understanding problems. However, in many cases, especially with young children, cognitive insight does not occur. Rather, emotional reexperiencing, emotional working through, and mastery do occur and result in symptom reduction and healthy adjustment. This is an important point and is an often overlooked mechanism of change in child treatment. Messer and Warren (1995) also state that the goal of making the unconscious conscious needs to be modified in child play therapy with many children. In Erikson's (1963) concept of mastery, the child uses play to gain mastery over traumatic events and everyday conflicts. Resolving conflicts through play is part of normal child development. Waelder (1933) described the play process as one in which the child repeats an unpleasant experience over and over until it becomes manageable. As he puts it, the child "digests" the event. Freedheim and Russ (1992) describe the slow process of gaining access to conflict-laden material and playing it out until the conflict is resolved. The therapist helps guide the play, labels thoughts, feelings, and events, and makes interpretations to facilitate conflict resolution and the working through process. Because cognitive insight is not necessary for conflict resolution to occur, the amount of interpretation should be carefully considered by the therapist. This is especially true of interpretation of symbols in the play or of deeply forbidden wishes. Mild interpretations that link feelings and thoughts to behavior and that spell out cause and effect are especially helpful.

(iv) Learning Alternative Problem-Solving Techniques and Coping Strategies. The therapist, in a more directive approach, helps the child think about alternative ways of viewing a situation and generate problem-solving strategies. Role-playing and modeling of coping strategies are used. Practice with a variety of verbal expressions and interpersonal behaviors can occur in a safe, pretend play situation. For example, D. Singer (1993) gives examples of modeling techniques during therapy. Knell (1993) teaches the child new strategies for coping with feelings and situations.

(v) Development of Internal Structure. Many children have structural deficits that result in problems with self/object differentiation, interpersonal functioning, self-esteem regulation, impulse control, object constancy, and separation of fantasy from reality. In these children, there are major deficits in underlying cognitive, affective, and interpersonal processes. Structure-building approaches are based upon conceptualizations by Mahler (1968) and Kohut (1977) and view the therapist as being a stable, predictable, caring and empathic figure. Development of good object relations is a major goal of play therapy with these children. Gilpin (1976) stressed that the role of the therapist is to become an internalized object. The relationship between the therapist and child is probably the most important aspect of therapy in helping this process to occur. Genuine understanding and expression of empathy by the therapist is a major technique that enables the child to develop. Play is important here as a form of communication in that the therapist can empathize with the child's expressions. The major change that occurs is through the relationship with the therapist that facilitates developing interpersonal representations in the child.

(vi) Nonspecific Variables. Nonspecific variables function in child therapy as they do in adult therapy. Expectation of change, hope, awareness of parental concern, no longer feeling so alone, are all factors that contribute to change in therapy.

5.09.5.3 Mechanisms of Change in Different Types of Therapy

5.09.5.3.1 Psychodynamic play therapy

Different types of therapy emphasize different mechanisms of change. Different techniques are used to foster the different mechanisms. The form of therapy most associated with the psychodynamic approach is insight-oriented therapy, and it is most appropriate for the child with anxiety and internalized conflicts (Tuma & Russ, 1993). This approach is appropriate for children who have age-appropriate ego development, show evidence of internal conflicts, have the ability to trust adults, have some degree of psychological-mindedness, and can use play effectively. Insight-oriented therapy is most often recommended for internalizing disorders including many of the anxiety disorders and depressive disorders. Children with internalizing disorders often experience internal conflicts and have good ego development and good object relations. An insight-oriented approach with a focus on conflict resolution is most appropriate for internalizing disorders. Many children in this broad category are good players and can easily engage in play in the therapy situation.

The goals of insight-oriented therapy are to help the child resolve internal conflicts and master developmental tasks. The major mechanism of change is insight and working through. Through the use of play and interpretation from the therapist, the child "calls forth forbidden fantasy and feelings, works through and masters developmental problems, and resolves conflicts" (Freedheim & Russ, 1983). Active interpretation of the child's play, expressions, and resistances is a major technique. For example, the therapist might interpret a child's stealing from mother's purse as an expression of anger at feeling neglected by her after the birth of a baby brother.

Insight and working through can also be helpful for a child with good inner resources who has experienced a specific trauma (such as the loss of as a parent). Altschul (1988) described the use of psychoanalytic approaches in helping children to mourn the loss of as a parent. In this application, Webber (1988) stressed that the therapist must first address the question of whether the child can do his or her own psychological work. If not, therapy can be a major aid in the mourning process.

A second major form of psychotherapy is the structure-building approach, which is used with children with structural deficits and major problems in developing good object relations. For children with impaired object relations, self/other boundary disturbances, and difficulty distinguishing fantasy from reality, the therapist uses techniques that foster the development of object permanence, self/other differentiation, and modulation of affect. The major mechanism of change is the building of internal structure such as object relations. Anna Freud (1965) described the development of object relations through a continual process of separation from the significant adult, usually the mother. Mahler (1975) elaborated on the separation-individuation process and described the development of object constancy and object representations. As Blank and Blank (1986) stressed, Mahler's concept of separation-individuation represented a new organizing principle to development. Object relations is not just another ego function, but plays a major role in the organization of intrapsychic processes.

The growing theory on the development of object relations is a new phase in psychoanalytic theory construction (Tuma & Russ, 1993). Good object relations involve well-developed object representations. The child must invest in the mental representation of the loved external object. Children who have inadequately developed object relations have structural deficits that impair a variety of functions. This impairment is evident with psychotic and characterological disorders. Children with severely impaired object relations, such as borderline children, have early developmental problems with a mix of severe dysfunction in the family and in the case of borderline children, perhaps a genetic predisposition. These children require a structure-building psychotherapeutic approach.

In this approach, empathy on the part of the therapist (as a general relationship factor) (Kohut & Wolfe, 1978) is a much more important intervention than is interpretation (Russ, 1995). Chethik (1989) provided an excellent discussion of psychotherapy with borderline children and narcissistically disturbed children. He pointed out that many of the therapeutic techniques are supportive in that they "shore up" defenses. The problems characteristic of borderline and narcissistic children are early developmental problems usually stemming from severe disturbance in the parent-child interaction. Kohut and Wolfe (1978) discussed the failure of empathy from the parent that is a major issue in the faulty parent-child interaction. Because of the frequency of this defect in parent-child relations, empathy from the therapist around the history of empathic failure becomes an important part of therapy. Frequently, help with problem solving and coping is also used with these children. Therapy with these children is usually long-term (one to two years) to be effective.

A third form of psychodynamic therapy is supportive psychotherapy, most appropriate for children with externalizing disorders. These children frequently act-out, have antisocial tendencies, and are impulse-ridden. The broad syndrome of externalizing disorders includes labels of acting-out, antisocial, character disorders, attention deficit disorders, and conduct disorders. Theoretically, psychodynamic theory views these children as having major developmental problems. These children have not yet adequately developed the processes necessary for delay of gratification. In addition, these children are frequently egocentric, demonstrate an absence of shame and guilt, and their ability to empathize with others is impaired. Kessler (1988) has recommended that structured, supportive therapy is more helpful to these children than any other kind of psychodynamic therapy. Therapy focuses on the here and now and on the development of problem-solving skills and coping resources. For example, the therapist might role-play with the child about how to handle teasing at school or how to be assertive with parents.

At this point, given the effectiveness of behavioral and cognitive-behavioral approaches in working with externalizing disorders, it appears that supportive psychodynamic psychotherapy is not the treatment of choice (Russ, 1995). It should only be used as a supplement to other treatment approaches in order to work on a specific issue.

Chethik (1989) has a thorough discussion of the use of play within the psychodynamic approach. As reviewed by Tuma and Russ (1993), he describes four stages of play development within psychotherapy (see Chethik, 1989, pp. 48–66 for a more detailed description).

(i) Initial period of nonengagement. Setting the stage-developing expectations, structure, and limits. The therapist first defines how play will be used for communication and how the child's internal life combined with play materials will express and replay the child's internal life for them both. "Meaningful play" must be developed, sometimes by varying the structure. This means that the overinstinctualized child (e.g., impulsive, fast to react) may require more structure, whereas the underinstinctualized child (e.g., obsessive, slow to react) may need to be encouraged to express instinctual life in play.

(ii) Early phase of affective engagement. As play develops, the therapist begins to share metaphors that emerge, and the child becomes attached to both the process and the therapist. When this happens, the therapist can then permit regressions by becoming a player in the play (by doing what the child asks him or her to do). The child can then express his or her instinctual life more freely because he or she identifies with the therapist and the therapist's sanctions. The safety the child feels in expressions are further ensured by imposing boundaries (e.g., by keeping forbidden expression in the room or having clean-up time). The unstructured quality, the accepting attitudes, and the boundaries all foster early "regression in the service of the ego." As the child feels more comfortable and masters anxiety, his or her play becomes more open. Those expressions at first defensively avoided are now displayed in full view of the therapist.

(iii) Emergence of central fantasies. As the process intensifies, the child elaborates highly invested fantasies in play. Repetitive play (characterized by the "compulsion to repeat") begins to deal with past traumatic and difficult situations. In the therapy process, however, the past has a changed outcome: acceptance of the play and interpretations by the therapist permit new solutions, either verbally or through play. Now the situation is in the control of the child.

(iv) Period of working through. Specific symptoms or behaviors often have more than one meaning. A working through period is necessary where a series and variety of interpretations are made to bring about change in a symptom. Symptoms are discussed in different contexts until all the meanings are worked out.

5.09.5.3.2 Cognitive-behavioural play therapy

Principles of cognitive-behavioral therapy and play therapy have been integrated into one approach. Knell (1993) introduced the concept of Cognitive-behavioral play therapy (CBPT) in which cognitive and behavioral interventions are incorporated within a play therapy paradigm. Play itself is used to resolve problems.

Knell (1993) identified six properties of CBPT:

(i) The child is involved in the treatment through play.

(ii) CBPT focuses on the child's thoughts, fantasies, and environment.

(iii) CBPT provides a strategy or strategies for developing more adaptive thoughts and behaviors. The child is taught coping strategies for feelings and for situations.

(iv) CBPT is structured, directive, and goal oriented, rather than open-ended.

(v) CBPT incorporates empirically demonstrated techniques, such as modeling.

(vi) CBPT allows for an empirical examination of treatment.

Knell (1993) also identified similarities and differences between CBPT and more traditional psychodynamic or client-centered approaches. Similarities included: the importance of the therapeutic relationship; communications occur through play; therapy is a safe place; and play provides "clues" to understanding the child.

Differences suggested that in CBPT therapy is more directed and goal oriented; the therapist is involved in choosing play materials; play is used to teach skills and alternative behaviors; interpretations are given by the therapist (similar to psychodynamic but different from client-centered) and praise is a crucial component. It is also more empirically based.

These similarities and differences are true for the "classic" forms of psychodynamic play therapy. However, new forms of psychodynamic play therapy are also being developed that are more goal oriented. For example, Chethik's (1989) focal therapy is focused on specific problems and is of short duration. Messer and Warren (1995) have called for psychodynamic approaches to be adapted to short-term goal oriented frameworks. They state that short-term therapy is a frequent form of psychodynamic intervention and that play is a good vehicle for change. Still, Knell's (1993) approach is a thoughtful integration of different theoretical approaches and techniques and is an excellent model for how to integrate treatment approaches.

Bodiford-McNeil, Hembree-Kigin, and Eyberg (1996) have developed a CBPT approach that is focused on a specific population of children. Their short-term play therapy for disruptive children utilizes the principles of CBPT but tailors the approach for disruptive children. Their approach is also set for 12 sessions with very specific goals and objectives per session. For most of the 12 sessions, the first half of the session is task-oriented and the second half is child-directed play. The therapist uses a variety of techniques such as praise, reflection, imitation of play, questions, interpretation, reinforcement, and contingent attention. The therapist follows the child's lead in the play, but also tries to move it as quickly as possible. By the eighth session, the parent is included in the process and is coached to facilitate the child's play. Parents practice their play sessions at home with the child. Parents are taught most of the play facilitation skills, but not that of interpretation, which belongs in the domain of psychotherapy. Bodiford-McNeil et al.'s approach integrates techniques from a variety of theoretical approaches and chooses those that are most effective for a specific population, that is disruptive children. It is a good example of developing an integrated approach for a specific population.

5.09.5.4 Change Processes in Play in Child Development

Play facilitates change and growth in a variety of cognitive, affective, and personality processes. Play fosters growth in these areas in normal child development. These change mechanisms are thought to operate in the play therapy process, although not usually in a targeted and systematic fashion. The empirical literature in the play and child development area is relevant to the play psychotherapy literature. Russ (in press) has a more detailed review of the play and child development literature, with a specific focus on creativity.

5.09.5.4.1 Play and cognitive processes

Research has supported a relationship between play and creative cognitive processes (Dansky, 1980; Fein, 1981; Russ, in press-a). Saltz, Dixon, and Johnson (1977) found that fantasy play facilitated cognitive functioning on a variety of measures. They theorized that fantasy play is related to cognitive development because of the involvement of representational skills and concept formation. Singer and Singer (1976) concluded that the capacity for imaginative play is positively related to divergent thinking, verbal fluency, and cognitive functioning in general.

Early research on play and creative problem solving investigated play and insight ability. In a series of studies, Sylva, Bruner, and Genova (1976) concluded that play in children 3–5 years of age facilitated insight in a problem solving task. In one study, they had three groups of children. One group played with the objects which were later used in the problem-solving task. A second group observed the experimenter solve the problem. A third group was exposed to the materials. Significantly more children in the play and observation groups solved the problem than in the control group. The play group was more goal oriented in their efforts on the task and was more likely to piece together the solution than the other groups.

Vandenberg (1978) refined the experimental methodology of the Sylva et al. studies and used a wider age group, 4–10 years of age. The experimental group played with the materials to be used in the problem solving task and the control group was asked questions about the material. Children were also given hints as to the solution. The play group did significantly better on one of the two insight tasks following the intervention. Six and seven-year-olds benefited most from the play experience. Vandenberg concluded that the relationship between play and insightful tool-use was mediated by age and task characteristics. In a review of the insight and play studies, Vandenberg (1980) concluded that all studies in the area had the consistent finding that play facilitated insightful tool use and enhanced motivated task activity.

There is a substantial body of studies that has found a relationship between play and divergent thinking. Divergent thinking is the generation of a variety of ideas and alternative solutions to problems (Guilford, 1968). Divergent thinking is especially important in creativity. Singer and Singer (1990) viewed play as a way of practicing divergent thinking ability. Singer and Rummo (1973) found a relationship between play and divergent thinking in kindergarten boys. Pepler and Ross (1981) found that play was related to divergent thinking. Feitelson and Ross (1973) found that thematic play facilitated creative thinking. Experience with a divergent thinking task facilitated performance on divergent thinking tasks in a study by Pepler (1979). In that study, performance on the divergent thinking task could be predicted from the expression of symbolic and representational play. Hughes (1987) studied four- and five-year-olds and reported that manipulative play with objects facilitated divergent thinking, but only for the number of nonstandard responses on the Alternate Uses Test.

Wallach (1970) stressed the importance of the relationship between divergent thinking and fantasy. Subjects who scored well on divergent thinking tests produced novel stories on the TAT (Maddi, 1965) and engaged in daydreaming activity (Singer, 1973). Wallach (1970) proposed that breadth-of-attention deployment is the underlying variable involved in divergent thinking tasks. As Kogan (1983) pointed out, breadth-of-attention deployment refers to a scanning of the environment and memory in an associational manner. Both creativity and fantasy may share breadth-of-attention deployment. From a cognitive perspective, this variable could also account for the play-creativity link.

In several important studies, play facilitated divergent thinking in preschool children (Dansky, 1980; Dansky & Silverman, 1973). In particular, Dansky and Silverman found that children who played with objects during a play period gave significantly more uses for those objects than did control subjects. In a later study, Dansky (1980) found that make-believe play was the mediator of the relationship between play and divergent thinking. Free play facilitated divergent thinking, but only for children who engaged in make-believe play. Also, in this second study, play had a generalized effect in that the objects in the play period were different from those in the test period. These two studies are important because they are experimental studies that show a direct effect of play on divergent thinking.

Dansky's (1980) theoretical rationale for hypothesizing that play would facilitate divergent thinking was that the process of free combination of objects and ideas involved in play is similar to the elements involved in creative thinking. Dansky (1980) speculated that the free symbolic transformations inherent in pretend play helped create a temporary cognitive set toward the loosening of old associations. These ideas are consistent with the work of Sutton-Smith (1966, 1992) who stressed the role of play in the development of flexibility in problem solving. Play provides the opportunity to explore new combinations of ideas and to develop new associations for old objects. The object transformations that occur in play help to develop the capacity to see old objects in new ways. The capacity to see old objects and ideas in new ways should also aid in developing transformation abilities (i.e., the ability to break out of an old set and see a new solution to a problem). Kogan (1983) also suggested that children's play behavior involves a search for alternate modes of relating to the object, a process similar to searching for alternate uses for objects in divergent thinking tasks.

Pellegrini (1992) also identified flexibility as one link between play and creativity. In a study of third and fifth grade boys, flexibility in rough and tumble play was predictive of a variety of prosocial problem solving responses. Pellegrini proposed that in play, children recombine behaviors and develop flexible strategies. A varied problem solving repertoire aids in social competence. Saracho (1992) obtained results that also support a link between play and flexibility. She found that field independent children engaged more in play than did field dependent children and concluded, from observing the children's play, that the field independent children were exhibiting cognitive flexibility.

Until 15 years ago, the research on play and creativity had focused on cognitive variables as the explanatory mechanisms underlying the observed relationship. As discussed, explanations have included practice with divergent thinking, the recombination of objects and ideas, symbolic transformations, breadth-of-attention deployment, and the loosening of old cognitive sets or cognitive flexibility.

5.09.5.4.2 Play and affective processes

Play also facilitates affective development. Affect may also be involved in the link between play and creative problem solving. Play and affective processes are just beginning to be studied, mainly because of the difficulty in measuring affective processes in play. In general, the empirical study of affective expression in children is a young area (Masters, Felleman, & Barden, 1981). The theoretical explanations for the links between affect, play, and creativity have been in existence for some time, although research that tests theory is relatively recent.

The first major theoretical explanation for the relationship between affect and creativity was that of psychoanalytic theory, which proposed that controlled access to primary process thinking facilitated creativity. Primary process thinking was first conceptualized by S. Freud (1915/1958) as an early, primitive system of thought that was drive-laden and not subject to rules of logic or oriented to reality. Another way to view primary process thought is as affect-laden cognition. Russ (1987, 1993, 1996) proposed that primary process thought is a subtype of affect in cognition. Primary process content is material around which the child had experienced early intense feeling states (oral, anal, aggressive, etc.) According to psychoanalytic theory, primary process thinking facilitates creativity (Kris, 1952). Children and adults who have controlled access to primary process thinking should have a broader range of associations and be better divergent thinkers than individuals with less access to primary process. S. Freud's (1926/1959) formulation that repression of "dangerous" drive-laden content leads to a more general intellectual restriction predicts that individuals with less access to affect-laden cognitions would have fewer associations in general. Thus, children who are more expressive of and open to affective content would develop a richer, more complex store of affect-laden memories. This richer store of memories would facilitate divergent thinking and transformation abilities because it provides a broader range of associations and more flexible manipulation of images and ideas.

Primary process content can be expressed in play. As Waelder (1933) has noted, play is a "leave of absence from reality" (p. 222), and is a place to let primary process thinking occur. Play can be important in the development of primary process thought and, in turn, foster creative thinking.

Primary process theory is consistent with Bower's (1981) conceptualization of affect and memory processes (see Russ, 1993). The work on mood and memory suggests that the search process for associations is broadened by the involvement of emotion. Russ (1993) proposed that if primary process is thought of as mood-relevant cognition, then it could fit into a mood and memory theoretical framework. When stirred, primary process content could trigger a broad associative network. Primary process content was stored into the memory system when emotion was present. Access to this primary process content would activate emotion nodes and associations, thus broadening the search process.

Fein (1987) views affect as intertwined with pretend play and creativity. Fein (1987) viewed play as a natural form of creativity. She studied 15 children who were master players and concluded that good pretend play consists of five characteristics.

(i) *referential freedom*. The "as if," concept is important in that one object is treated as if it were another, one person functions as if they were another, time and place is as if it were different. Object substitutions and transformations occur.

(ii) *denotative license*. The child takes a divergent stance with respect to actual experience. There are pretend events, not just object substitution in an accurate account of events.

(iii) *affective relations*. Symbolic units represent affective relationships such as "fear of, love for, anger at." Fein proposed an affective symbol system that represents real or imagined experience at a general level. These affective units constitute affect-binding representational templates which store salient information about affect-laden events. The units are "manipulated, interpreted, coordinated and elaborated in a way that makes affective sense to the players" (p. 292). These affective units are a key part of pretend play and are especially important for creative thinking. Fein stated further that divergent thinking abilities like daydreams, pretend play, or drawing can activate the affective symbol system.

(iv) *sequential uncertainty*. The sequence of events in pretend play has a nonlinear quality.

(v) *self-mirroring*. Children are aware of the pretend, nonreal quality of the play. The self is observed from a distance through the play.

One of Fein's (1987) major conclusions is that creative processes cannot be studied independently of an affective symbol system An affective symbol system is activated in pretend play and is probably facilitated through pretend play. The concept of affective symbols is consistent with the concept of primary process. Primary process content is proposed to be stored in the affect symbol system.

Vygotsky (1930/1967) is also a major theoretician in the area of affect, play, and creativity. He presented a rich conceptualization of play and creativity. Smolucha (1992) has translated and integrated Vygotsky's major papers on the topic of creativity. From her review of his work, a major premise in Vygotsky's theory is that imagination develops out of children's play. He stated: "The child's play activity is not simply a recollection of past experience but a creative reworking that combines impressions and constructs from them new realities addressing the needs of the child" (1930/1967, p. 7). Through play, children develop combinatory imagination (i.e., the ability to combine elements of experience into new situations and new behaviors). Combinatory imagination is significant because it can contribute to artistic and scientific creativity.

Morrison (1988) placed cognitive-affective development within an interpersonal framework. The cognitive integration of affect occurs within safe interactions with parents. Representations of self and others are fused with affect. In play, the child reconstructs past experience and explores definitions of the self. Old metaphors are constantly reworked. In this way, the child develops reflective thought. Conflicts from early interpersonal experience can be a major source of creative thinking in that the metaphors of early experience are reworked in creative acts. Santostefono (1988) also stressed the importance of play in metaphor construction. The process of constructing and negotiating metaphors is creative and can lead to later creativity.

Russ (1993) reviewed the affect and creativity literature and proposed five categories of affect that emerged as important in the creative process: access to affect-laden thoughts; openness to affect states; affective pleasure in challenge; affective pleasure in problem solving; and cognitive integration and modulation of affect. The question of whether or not these five categories are truly separate dimensions of affect needs to be investigated systematically. In Russ' model of affect and creativity, these affective processes are related to specific creative cognitive processes. The links between the processes are based on theory and the empirical literature. Two broad affective processes are access to affect-laden thoughts and openness to affect states.

Access to affect-laden thoughts is the ability to think about thoughts and images that contain emotional content. Affective fantasy in daydreams and in play are examples of affect-laden thoughts. Thoughts involving emotional themes such as aggressive and sexual ideation illustrate this kind of blending of affect and cognition. The psychoanalytic concept of primary process thinking is also an example of emotion-laden thinking (see Kris, 1952; Dudek, 1980; Rothenberg, 1990).

Openness to affect states is the ability to feel the affect itself. Comfort with intense affect, the ability to experience and tolerate anxiety, and passionate involvement in a task are examples of openness to affect states.

Two other more specific affective processes important in creativity are affective pleasure in challenge and affective pleasure in problem solving. The capacity to enjoy the excitement and tension in the challenge (Runco, 1994) and the capacity to take deep pleasure in problem solving (Amabile, 1990) are important in the creative process.

Finally, cognitive integration and modulation of affect is important in producing good and adaptive products. Cognitive control of affect is essential for the critical evaluation of ideas and products (Arieti, 1976; Kris, 1952).

There is a growing body of research that supports the link between play, affect, and creativity. Lieberman's (1977) work supports a relationship between affect in play and divergent thinking. She focused on the variable of playfulness which included the affective components of spontaneity and joy. She found that playful kindergarten children did better on divergent thinking tasks than nonplayful children. Singer and Singer (1990) reported that positive affect was related to imaginative play. Christie and Johnson (1983) also concluded that there is a relationship between playfulness and creativity. Singer and Singer (1981) found that preschoolers related as high imagination players showed significantly more themes of danger and power than children with low imagination.

In a series of studies by Russ and her colleagues using the Affect in Play Scale (APS), affect in play was found to be related to creativity. Russ and Grossman-McKee (1990) investigated the relationship among the APS and divergent thinking and primary process thinking on the Rorschach in 60 first and second grade children. As predicted, affective expression in play was significantly positively related to divergent thinking, as measured by the Alternate Uses Test. All major scores on the APS were significantly correlated with divergent thinking, with correlations ranging from 0.23 between comfort and divergent thinking to 0.42 between frequency of affective expression and divergent thinking. All correlations remained significant when IQ was partialed out; IQ had low correlations with the APS. The fact that intelligence did not relate to any of the play scores is consistent with the theoretical model for the development of the scale and is similar to the results of Singer (1973). Also, there were no gender differences in the pattern of correlations between the APS and divergent thinking.

The finding of a relationship between affect in play and divergent thinking (Russ & Grossman-McKee, 1990) was replicated by Russ and Peterson (1990; Russ, 1993) who used a larger sample of 121 first and second grade children. Once again, all of the APS scores were significantly and positively related to the Alternate Uses Test, independent of intelligence. Again, there were no gender differences in the correlations. Thus, with this replication, more confidence in the robustness of the relationship between affect in pretend play and creativity in young children is possible.

An important question about the APS is whether it is indeed measuring two separate dimensions of play—an affective dimension and a cognitive dimension—or is measuring one dimension an affect in fantasy dimension. The results of two separate factor analyses with the scale suggest two separate dimensions. Russ and Peterson (1990) carried out a principal component analysis with oblique rotation. It yielded two separate factors as the best solution. The first and dominant factor appeared to be a cognitive factor. Imagination, organization, quality of fantasy, and comfort loaded on this first factor. The second factor appeared to be an affective factor. Frequency of affective expression, variety of affect categories, and intensity of affect loaded on this second factor. Although separate factors, there was a significant amount of shared variance ($r = 0.76$), suggesting that the factors overlapped. A study by D'Angelo (1995) replicated the finding of two factors, one cognitive and one affective, with a sample of 95 first, second, and third grade children. This is an important finding because it lends support to the idea that cognitive expression and affective expression in play are related, but separate processes. Future studies with the play scale should use factor scores on the cognitive and affective factors as predictors of creativity.

Another interesting finding in D'Angelo's study (1995) was a significant relationship between the APS and Singer's (1973) imaginative play predisposition interview. Good players on the APS reported that they preferred activities that required using imagination.

A 1995 study (Russ, Robins, & Christiano, 1995), was a follow-up study of the first and second graders in the 1990 investigation of Russ and Peterson (1990). Fifth and sixth grade children were used for the follow-up. Thirty-one children agreed to participate in the follow-up study. This was a longitudinal study that explored the ability of the APS to predict creativity over a four year period (five years in some cases because the study took two years to complete). The major finding of the study was that quality of fantasy and imagination on the APS was predictive of divergent thinking over a four year period. However, the correlation between variety of affect and divergent thinking did not reach significance possibly due to the small sample size. In this study, an adapted version of the play task was administered to the older children. The instructions were altered so that the children were asked to put on a play with the puppets. The task was then scored based on the scoring criteria for the APS. Good stability was found in the dimensions being measured by the APS. For example, the size of the correlation between the two frequency of affect scores was $r = 0.33$, between the two variety of affect scores was $r = 0.38$, and between the two frequency of positive affect scores was $r = 0.51$. In general, the size of the correlations is respectable for a period of four and five years and lends support for enduring, stable constructs of affective expression in fantasy that are predictive over time of creative thinking. These findings also suggest an enduring quality to the affective and cognitive dimensions of the APS over a five year period.

5.09.5.4.3 Play and adjustment

There is some empirical work that suggest that play is related to adjustment. Singer and Singer (1990) concluded that imaginative play in children is related to academic adjustment and flexibility of thought. They also found that toddlers and preschoolers who engage in make-believe play are better adjusted across different situations. Burstein and Meichenbaum (1979) found that children who voluntarily played with stress-related toys prior to surgery demonstrated less distress and anxiety following surgery than children who avoided the toys. One might speculate that those children were accustomed to using play to deal with stress and problems. In a study of 4 to 11 year olds, Kenealy (1989) investigated strategies that children use when they are feeling depressed and found that 50% of the children's responses included play strategies. Also, indirectly, if play facilitates flexible problem solving, there is evidence that flexible problem solving, in turn, aids the coping process (Follick & Turk, 1978).

In a study of urban children from 4 to 5 years of age, Rosenberg (1984) found that the quality of fantasy play for children playing in dyads was positively related to measures of social competence and ego resilience (Block-Q sort). Frequency of positive themes and relationship themes in the play was also related to ego resilience and social competence. In general, children with behavior problems and attachment problems had fewer positive and negative themes in play, with the exception of diffuse hostility.

Grossman-McKee (1990) found, using the APS with first- and second-grade boys, that boys who expressed more affect in play had fewer pain complaints than boys with less affect in play. Good players were also less anxious on the State-Trait Anxiety Inventory for Children (Spielberger, 1973).

In a study of 7 to 9 year olds, Christiano and Russ (1996) found a positive relationship between play and coping, and a negative relationships between play and distress. Children who were "good" players on the APS implemented a greater number and variety of cognitive coping strategies (correlations ranging from 0.52 to 0.55) during an invasive dental procedure. In addition, good players reported less distress during the procedure than children who expressed less affect and fantasy in play.

5.09.5.4.4 Meta-analysis

Fisher (1992) conducted a meta-analysis of 46 studies in the play and child development area up to 1987. He investigated the impact of play on cognitive, affective-social, and linguistic processes. Both correlational and experimental studies were included. He found a mean effect size of 0.347 which is of low to modest size. He concluded that the major finding was that play does result in moderately large to noteworthy improvements on children's development. He made special note of the larger effect size for divergent thinking and perspective taking criteria.

In summary, play helps the child to:

(i) Practice with the free flow of associations that is important in divergent thinking.

(ii) Practice with symbol substitution, recombining of ideas, and manipulation of object representations. These processes are important in developing insight abilities.

(iii) Express and experience both positive and negative affect.

(iv) Express and think about affect themes in fantasy. Emotional content and primary process content is permitted to surface and be expressed through play. Over time, the child develops a variety of memories and associations. This broad repertoire of associations helps in creative problem solving. Fein's (1987) affective symbol system could be developed in this way.

(v) Resolve conflicts and master the traumas of daily life.

(vi) Develop cognitive structure that could help the child contain and modulate affect.

(vii) Practice with a variety of coping skills and interpersonal behavior important in everyday life.

5.09.6 RESEARCH FINDINGS

Russ (1995) reviewed the current state of play therapy research. In that review, a distinction between psychotherapy outcome research in general and play intervention research was made. The broad category of psychotherapy outcome research makes sense because most of those studies did not focus specifically on play. Play is frequently embedded in the therapy intervention. Therefore, the results of those studies provide an important context in which to think about play.

5.09.6.1 Psychotherapy Outcome Research

In general, the early reviews of child therapy outcome studies concluded that there was little or no support for child therapy. More recent work has concluded that there is support for the effectiveness of child psychotherapy if the research is well designed.

A classic early review by Levitt (1957) concluded that the mean improvement rate for children was not significantly better than the baseline improvement rate of 72.5% for untreated controls. In later work, in a review of 47 reports of outcome studies, Levitt (1963, 1971) concluded that approximately two-thirds of treated children in therapy were improved, but again treated children were no better off than untreated controls. Levitt's conclusions were taken seriously by the field of child psychotherapy.

A number of researchers responded to Levitt's conclusions (Halpern, 1968; Heinicke & Goldman, 1960; Hood-Williams, 1960). One of the major methodological issues was that so many of the untreated controls were defectors from treatment. Defectors were children who were evaluated and recommended for treatment but who had not entered treatment. Therefore, there may have been a number of confounding variables operating here to account for the results.

Barrett, Hanpe, and Miller (1978) and Hartmann, Roger, and Gelsand (1977) took a closer look at Levitt's reviews and at the research literature in general. They concluded that there was still no solid empirical evidence for the effectiveness of psychotherapy. Barrett et al. (1978) stated that the global nature of the research was a major problem and concluded that most of the research studies were not specific enough or focused enough to enable research questions to be answered. There was too much of a mix of populations, therapeutic approaches, and interventions in these studies. Often, the outcome measures were unrefined or nonexistent. This led to their often quoted conclusion that the question in psychotherapy research should not be "Does psychotherapy work?" but rather "Which set of procedures is effective when applied to what kinds of patients with which sets of problems and practiced by which sorts of therapists?" (p. 428)

A number of other methodological issues important for research in the child therapy area have been identified: namely, the importance of classification according to developmental level (Heinicke & Strassman, 1975); controlling for maturational effects (Koocher & Broskowski, 1977); the need for homogeneous treatment groups (Achenbach, 1978; Hartmann et al., 1977); the need to control for sex and age variables (Cass & Thomas, 1979); and the need for adequate outcome measures given at appropriate intervals (Kazdin, 1990, 1993b).

The field of child therapy research has followed these research guidelines and the research studies have become more focused and methodologically sophisticated. In addition, the technique of meta-analysis has enabled the field to arrive at a more systematic evaluation of outcome studies. As Weisz and Weiss (1993) noted, meta-analysis is a technique that enables the pooling and statistical summarizing of the results of outcome studies. The effect size (ES) is the statistical summary of the treatment efficacy across studies. Use of this systematic procedure helps avoid reviewer subjectivity in coming to conclusions.

Weisz and Weiss (1993) reviewed the major meta-analytic studies in the field of child psychotherapy. Casey and Berman (1985) calculated the effect of psychotherapy across 64 studies and found a mean effect size of 0.71. A slightly higher effect size of 0.79 was found by

Weisz, Weiss, Alicke, and Klotz (1987) in a review of 163 treatment-control comparisons. Both studies concluded that the average treated child functioned better after treatment than three-fourths of the untreated controls. In the Casey and Berman review, effect sizes did not differ as a function of whether play was used. In a meta-analyses by Kazdin, Bass, Ayers, and Rodgers (1990), for 64 studies involving treatment vs. no-treatment comparisons, the mean effect size was 0.88. Weisz and Weiss (1993) concluded that "the mean effect sizes reported in child meta-analyses are quite comparable to those of adult meta-analyses and that findings in both categories point to quite positive effects of therapy" (p. 46).

As Kazdin (1990) has pointed out, the results of these meta-analyses have contributed to the field in that they offer evidence that psychotherapy is more effective than no treatment with children. This conclusion is more encouraging than the conclusions based on the reviews in the 1950s, 1960s, and 1970s. Although these child therapy outcome studies did not focus on play *per se*, one might infer that play is an effective form of treatment since it is so frequently part of the therapy process. The Casey and Berman (1985) review found no difference in effectiveness between those studies that used play and those that did not.

Weisz and Weiss (1989, 1993) pointed out that most of the research studies in the meta-analyses involved controlled laboratory interventions. In many of these studies children were recruited for treatment and were not clinic-referred; samples were homogeneous; there was a focal problem; therapy focused on the target problem; therapists were trained in the specific treatment approaches to be used; and the therapy relied primarily on those techniques. In essence, this was good research that followed many of the methodological guidelines for adequate research design. On the other hand, Weisz and Weiss (1993) cautioned that the evidence for the effectiveness of psychotherapy is based on studies that are not typical of conventional clinical practice. Thus the findings may not be generalizable to real clinical work.

The results of the meta-analysis point to the need for specificity and precision. Weisz and Weiss (1993) concluded that the studies that showed positive results tended to "zoom in" on a specific problem with careful planning of the intervention. Freedheim and Russ (1983, 1992) stated that we needed to become very specific and ask "Which specific interventions affect which specific cognitive, personality, and affective processes? How are these processes related to behavior and practical clinical criteria?" (1983, p. 988). Shirk and Russell (1996) also call for the targeting of specific cognitive, affective, and interpersonal processes in child therapy.

In a review of the play therapy research, Phillips (1985) also called for a systematic program of research with well-controlled studies. He concluded that the play therapy research that found positive results were those studies of a cognitive-behavioral nature that were carefully designed. Phillips speculated that the specificity of treatment goals and focused methods of the cognitive-behavioral studies partially account for the positive results. He recommended that all forms of play therapy be investigated with the precision of the cognitive-behavioral studies. In general there are not many play therapy studies in the literature and little exploration of variables that leads to change (Faust & Burns, 1991).

It should be relatively easy to apply the principles of specificity and focus to the play area. Play interventions and the cognitive and affective processes that they effect can be broken down into discrete units in controlled conditions. The large body of research in the play and child development literature offers a wealth of ideas and research lines that could be followed.

5.09.6.2 Play Intervention Research

As Russ has reviewed (1995), studies exist that have investigated the effect of play on specific types of problems and in specific populations. These studies are a good bridge between empirical laboratory studies of play and specific processes, like creativity, and more global clinical practice outcome studies. These studies have been labeled play intervention rather than play therapy, because the focus is highly specific. Usually, they involve only a few sessions with no emphasis on forming a "relationship" with a therapist. On the other hand, these studies differ from specific process research in child development in that they are problem focused and are not as fine-tuned as they would be in laboratory research. These play intervention research studies seem to fit some of Weisz and Weiss's (1993) criteria by including children who are not clinic-referred, by having homogeneous samples, and by having a focal problem that the therapy focused on.

Phillips (1985) reviewed two studies that would fall into this play intervention research category. Both involved the use of puppet play to reduce anxiety in children facing surgery. Johnson and Stockdale (1975) measured Palmer Sweat Index level before and after surgery. Puppet play in this study involved playing out the surgery. Johnson and Stockdale found less anxiety for the puppet-play group before and after surgery. The one exception was immediately before surgery, when the increased information may have elevated their anxiety. Cassell (1965) used puppets with children undergoing cardiac catheterization and found that anxiety was reduced before surgery for the puppet-play group compared with the no treatment control. There were no differences after surgery. The treatment group was less disturbed during the cardiac catheterization and expressed more willingness to return to the hospital for further treatment. Rae, Worchel, Upchurch, Sanner, and Daniel (1989) investigated the effects of play on the adjustment of 46 children hospitalized for acute illness. Children were randomly assigned to one of four experimental groups:

(i) Therapeutic play condition in which the child was encouraged to play with medical and nonmedical materials. Verbal support, reflection, and interpretation of feelings were expressed by the research assistant.

(ii) A diversionary play condition in which children were allowed to play with toys but fantasy play was discouraged. The toys provided did not facilitate fantasy, nor did the research assistant.

(iii) A verbally oriented support condition in which children were encouraged to talk about feelings and anxieties. The research assistant was directive in bringing up related topics and would ask about procedures.

(iv) A control condition in which the research assistant had no contact with the child.

All treatment conditions consisted of two 30-minute sessions. The main result of this study was that children in the therapeutic play group showed significantly more reduction in self-reported hospital-related fears than children in the other three groups. There were no differences among the groups for parent ratings.

Another specific problem area that lends itself to focused intervention research is that of separation anxiety. In an excellent example of a well-designed play intervention study, Milos and Reiss (1982) used play therapy for preschoolers who were dealing with separation anxiety. They identified 64 children who were rated as high-separation-anxious children by their teachers. The children were randomly assigned to one of four groups. Three play groups were theme related: the free-play group had appropriate toys; the directed-play group had the scene set with a mother doll bringing the child to school; the modeling group had the experimenter playing out a separation scene. A control group also used play with toys irrelevant to separation themes (blocks, puzzles, crayons). All children received three individual 10 minute play sessions on different days. Quality of play was rated. The results showed that all three thematic play conditions were effective in reducing anxiety around separation themes when compared to the control group. An interesting finding was that, when the free-play and directed-play groups were combined, the quality of play ratings were significantly negatively related ($r = -0.37$) to a posttest anxiety measure. High-quality play was defined as play that showed more separation themes and attempts to resolve conflicts. One might speculate that the children who were already good players used the intervention to master their separation anxiety. Milos and Reiss concluded that their results support the underlying assumption of play therapy, that play can reduce anxiety associated with psychological problems. The finding that quality of play was related to effectiveness of the intervention is consistent with the finding of Dansky (1980); that free play facilitated creativity only for those children who used make-believe well.

A well-designed study by Barnett (1984) also looked at separation anxiety and expanded upon work by Barnett and Storn (1981) in which free play was found to reduce distress in children following a conflict situation. In the 1984 study, a natural stressor, the first day of school, was used. Seventy-four preschool children were observed separating from their mothers and were rated anxious or nonanxious. These two groups were further divided into play or no play conditions. The play condition was a free play condition. The no-play condition was a story-listening condition. For half of the play condition, play was solitary. For the other half, peers were present. The story condition was also split into solitary and peers present segments. Play was rated by observers and categorized in terms of types of play. Play significantly reduced anxiety in the high-anxious group. Anxiety was measured by the Palmer Sweat Index. There was no effect for low-anxious children. For the anxious children, solitary play was best in reducing anxiety. High-anxious children spent more time in fantasy play than did low-anxious children, who showed more functional and manipulative play. They engaged more in fantasy play when no other children were present. Barnett interpreted these results to mean that play was used to cope with a distressing situation. The findings supported her idea that it is not social play that is essential to conflict resolution, but rather imaginative play qualities that the child introduces into playful behavior. Actually, the presence of peers increased anxiety in the high-anxious group.

These play intervention studies are a few examples of the kind of studies that tell us about how play can be helpful in dealing with specific problems. The results of these studies suggest that play helps children deal with fears and reduce anxiety and that something about play itself is important and serves as a vehicle for change. The play experience is separate from the experience of a supportive and empathic adult. Results also suggest that children who are already good players are more able to use play opportunities to resolve problems when these opportunities arise. Teaching children good play skills would provide children with a resource for future coping.

5.09.7 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

The main tasks of the next 10 years, up to 2007, in the play therapy area are to: (i) apply the body of knowledge from the play and child development research literature to play intervention and (ii) carry out well-designed play intervention studies with specific populations of children. These are major tasks that require a coordinated set of research studies and programs.

Play therapy itself will become more focused with specific populations of children. This is already occurring. For example, Chethik (1989) described "focal therapy" as therapy that focuses on a specific problem and is of short duration. Play is frequently used in this therapy. Focal stress events such as death in the family, divorce, hospitalization, or illness are examples of problems that would be dealt with Chethik stated that focal therapy is effective when the child has accomplished normal developmental tasks before the stressful event occurs.

Mann and McDermott (1983) discussed play therapy with abused and neglected children. Often, these children must be taught how to play. Gil (1991) provided a good review and discussion of post-traumatic play. An active approach in which she brings in outside material and events is often necessary. Gil stressed the importance of the therapist being an active participant in the play with abused children. If the child is a disorganized player, the therapist must focus the play choices in a relatively restricted setting. If the child has been understimulated and cannot play well, then the therapist should bring in appropriate toys and model play. For many abused children, the therapist needs to actively facilitate self-expression by such techniques as presenting the child with cartoon figures in different situations with the child filling in their words or pulling secrets from a "secrets" bag. For many children post-traumatic play is repetitive, devoid of pleasure, and can remain fixed (Gil, 1991). Gil attempts to intervene in this repetitive play by making verbal statements, having the child take a specific role, or encouraging the child to differentiate between the traumatic event and current reality in terms of safety and what has been learned. Gil stated that the goal of interrupting the play is to generate alternatives that can lead to a sense of control, help feelings be expressed and orient the child toward the future. Gil's work illustrates how specific play techniques need to be developed for specific populations.

D. Singer (1993) described the active use of modeling, imagery techniques, and teaching adaptive skills within the play therapy context. Along with traditional play therapy techniques, she utilizes imagery techniques that help the child visualize people, conflicts, and resolutions. She also incorporates behavior modification techniques that reinforce positive behavior and lead to self-reward. Knell's (1993) cognitive-behavior play therapy and Bodiford-McNeil et al's. (1996) short-term play therapy are examples of focused therapies with a very active therapist.

Play therapy will also need to integrate a variety of theoretical approaches and techniques. As it becomes evident that a variety of theoretical approaches can use play in therapy, a larger variety of play techniques will become available. If we can combine this increased repertoire of play techniques with good research then the field of play therapy will move forward.

Short-term play intervention modules may evolve as effective interventions. Kazdin (1993a) discussed the possibility of having different modules of intervention for different problems. This concept would work for the play therapy area. 6–12 week play modules could be developed for different types of problems. Bodiford-McNeil et al. (1996) are already doing this with disruptive children. Children who have experienced trauma might benefit from the opportunity to play-out the issues in a focused approach. Constricted children could benefit from play modules directed at increasing affective expression. Play assessment could identify what types of play experiences could be most beneficial.

Research programs in play therapy need to be at both the micro and macro levels. First, laboratory research on play and cognitive and affective processes must continue. Concurrent, longitudinal, and experimental studies on specific dimensions of play and specific cognitive, affective, and personality processes need to be carried out in a systematic fashion. We especially need to identify what dimensions of play most relate to specific processes. For example, does affect expression increase divergent thinking? Does positive affect have a different effect than negative affect? Are there different effects with different age groups and different populations of children?

Second research needs to be carried out with focused play interventions. Russ (1995) outlined different types of play intervention research.

(i) Specific Play Interventions with Specific Populations and Specific Situations. The Barnett (1984) study with children who were experiencing the first day of school is a good example of this type of study. There are a variety of natural stressors that could be used to investigate play intervention. Divorce, natural disasters, dental visits, presurgery, and loss of a parent are all situations in which play intervention is used. We need to develop an empirical base for play intervention in these situations.

(ii) Refining Specific Play Techniques. The general question of what kinds of intervention by the therapist best facilitate play needs to be studied empirically. There are many guidelines in the clinical literature about how to facilitate play, but few are based on empirical work. How do we best encourage affect in play? When is modeling more effective than a less directive approach.? When is it less effective? For example, Gil (1991) pointed out that it is frequently important with sexually abused children to be nondirective, so that the child does not feel intruded upon. What kinds of intervention most enhance the working-through process and conflict resolution? These kinds of research questions can be posed in well-controlled experimental studies and in psychotherapy-process research.

Third, dissemination of research results should inform clinical practice and prevention programs. For example, guidelines about how to best facilitate affective expression in play could be better incorporated into play therapy. Treatment manuals using these guidelines can and should be developed.

In the future, large scale play training and prevention programs could reach large numbers of children. Play centers in schools might be one way to reach children (Russ, 1993). Smilanksy (1968) has demonstrated that play skills can effectively be taught to children by teachers in a relatively short time. More training groups for parents on how to help children play at home, perhaps through Head Start programs would also be worthwhile. Play preparation for hospital procedures might increase. Teaching children to use play as a major coping resource might be a trend of the future.

One thing is certain, children will continue to play. Above all else, it is fun and reinforcing. Children naturally use play for a variety of purposes. It is up to us to learn how children use play and to help them use play more effectively.

5.09.8 REFERENCES

Achenbach, T. (1978). Psychopathology of childhood: Research problems and issues. *Journal of Consulting and Clinical Psychology, 46,* 759–776.

Amabile, T. (1990). Within you, without you: The social psychology of creativity and beyond. In M. Runco & R. Albert (Eds.), *Theories of creativity* (pp. 61–69). Newbury Park, CA: Sage.

Applebaum, S. (1978). Pathways to change in psychoanalytic therapy. *Bulletin of Menniger Clinic, 42,* 239–251.

Archer, R., Maruish, M., Imhof, E., & Piatrowski, C. (1991). Psychological test usage with adolescent clients: 1990 survey findings. *Professional Psychology Research and Practice, 22,* 247–252.

Arieti, S. (1976). *Creativity: The magic synthesis.* New York: Basic Books.

Axline, V. M. (1947). *Play therapy.* Boston: Houghton-Mifflin.

Axline, V. M. (1964). *Dibs: In search of self: Personality development in play therapy.* Boston: Houghton Miffllin.

Barnett, I. (1984). Research note: Young children's resolution of distress through play. *Journal of Child Psychology and Psychiatry, 25,* 477–483.

Barnett, I., & Storn, B. (1981). Play, pleasure and pain: The reduction of anxiety through play. *Leisure Science, 4,* 161–175.

Barrett, C., Hampe, T. E., & Miller, L. (1978). Research on child psychotherapy. In S. Garfield & A. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (pp. 411–435). New York: Wiley.

Blank, R., & Blank, G. (1986). *Beyond ego psychology: Developmental object relations theory.* New York: Columbia University Press.

Bodiford-McNeil, C., Hembree-Kigin, T. L., & Eyberg, S. (1996). *Short-term play therapy for disruptive children.* King of Prussia, PA: The Center for Applied Psychology.

Bower, G. H. (1981). Mood and memory. *American Psychologist, 36,* 129–148.

Burlingham, D. (1932). Child analysis and the mother. *Psychoanalytic Quarterly, 4,* 69–92.

Burstein, S., & Meichenbaum, D. (1979). The work of worrying in children undergoing surgery. *Journal of Abnormal Child Psychology, 7,* 121–132.

Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin, 98,* 388–400.

Cass, L., & Thomas, C. (1979). *Childhood pathology and later adjustment.* New York: Wiley.

Cassell, S. (1965). Effect of brief puppet therapy upon the emotional response of children undergoing cardiac catheterization. *Journal of Consulting Psychology, 29,* 1–8.

Chethik, M. (1989). *Techniques of child therapy: Psychodynamic strategies.* New York: Guilford.

Christiano, B., & Russ, S. (1996). Play as a predictor of coping and distress in children during an invasive dental procedure. *Journal of Clinical Child Psychology, 25,* 130–138.

Christie, J., & Johnson, E. (1983). The role of play in social-intellectual development. *Review of Educational Research, 53,* 93–115.

D'Angelo, L. (1995). Child's play: The relationship between the use of play and adjustment styles. Unpublished dissertation. Case Western Reserve University, Cleveland, OH.

Dansky, I. (1980). Make-believe: A mediator of the relationship between play and associative fluency. *Child Development, 15,* 576–579.

Dansky, J., & Silverman, F. (1973). Effects of play on associative fluency in preschool-aged children. *Developmental Psychology, 9,* 38–43.

Dawes, R. (1994). *House of cards: Psychology and psychotherapy, truth or myth*. New York: Free Press.

Dudek, S. (1980). Primary process ideation. In R. H. Woody (Ed.), *Encyclopedia of clinical assessment* (Vol. 1, pp. 520–539). San Francisco: Jossey-Bass.

Erikson, E. N. (1963). *Childhood and society*. New York: Norton.

Exner, J. (1986). *The Rorschach: A comprehensive system* (Vol. 1, 2nd ed). New York: Wiley.

Faust, J., & Burns, W. (1991). Coding therapist and child interaction: Progress and outcome in play therapy. In C. Schaefer, K. Gitlin, & A. Sandgrund (Eds.), *Play diagnosis and assessment* (pp. 663–689), New York: Wiley.

Fein, G. (1981). Pretend play in childhood: An integrative review. *Child Development, 52,* 1095–1118.

Fein, G. (1987). Pretend play: Creativity and consciousness. In P. Gorlitz & J. Wohlwill (Eds.), *Curiosity, imagination, and play* (pp. 231–304). Hillsdale, NJ: Erlbaum.

Feitelson, D., & Ross, G. (1973). The neglected factor-play. *Human Development, 16,* 202–223.

Fisher, E. (1992). The impact of play on development: A meta-analysis. *Play and Culture, 5,* 159–181.

Follick, M., & Turk, D. (1978). *Problem specification by ostomy patients*. Paper presented at the 12th Annual Convention for the Advancement of Behavior Therapy. Chicago, IL.

Freedheim, D. K., & Russ, S. W. (1983). Psychotherapy with children. In C. E. Walker & M. E. Roberts (Eds.), *Handbook of clinical child psychology* (2nd ed., pp. 978–994). New York: Wiley.

Freedheim, D. K., & Russ, S. W. (1992). Psychotherapy with children. In E. Walker & M. Roberts (Eds.), *Handbook of clinical child psychology* (2nd ed., pp. 765–780). New York: Wiley.

Freud, A. (1927). Four lectures on child analysis. In *The Writings of Anna Freud* (Vol. 1, pp. 3–69). New York: International Universities Press.

Freud, A. (1946). *The psychoanalytic treatment of children*. New York: International Universities Press.

Freud, A. (1966). The ego and the mechanisms of defense. In *The Writings of Anna Freud* (Vol. 2). New York: International Universities Press.

Freud, A. (1965). Normality and pathology in childhood: Assessments of development. In *The Writings of Anna Freud* (Vol. 6). New York: International Universities Press.

Freud, A. (1976). Changes in psychoanalytic practice and experience. In *The Writings of Anna Freud* (Vol. 7, pp. 176–185). New York: International Universities Press.

Freud, S. (1958). The unconscious. In S. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 14, pp. 159–215). London: Hogarth Press (original work published 1915).

Freud, S. (1959). Inhibition symptoms, and anxiety. In J. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 20, pp. 87–172). London: Hogarth Press (original work published 1926).

Garfield, W. (1980). *Psychotherapy: An eclectic approach*. New York: Wiley.

Gardner, R. (1971). *Therapeutic communication with children: The mutual storytelling technique*. New York: Aronson.

Gil, E. (1991). *The healing power of play*. New York: Guilford.

Gilpin, D. (1976). Psychotherapy of borderline psychotic children. *American Journal of Psychotherapy, 30,* 483–496.

Grossman-McKee, A. (1990). The relationship between affective expression in fantasy play and pain complaints in first and second grade children. *Dissertation Abstracts International, 50–09B,* 4219.

Guilford, J.P. (1968). *Intelligence, creativity and their educational implications*. San Diego, CA: Knapp.

Heinicke, C. & Goldman, A. (1960). Research on psychotherapy with children: A review and suggestions for further study. *American Journal of Orthopsychiatry, 30,* 483–494.

Heinicke, C. & Strassman, L. (1975). Toward more effective research on child psychotherapy. *Journal of Child Psychiatry, 14,* 561–588.

Hood-Williams, J. (1960). The results of psychotherapy with children. *Journal of Consulting Psychology, 24,* 84–88.

Hug-Hellmuth, H. (1921). On the technique of child-analysis. *International Journal of Psychoanalysis, 2,* 287–305.

Hug-Hellmuth, H. (1924). *New paths to the understanding of youth*. Leipzig-Wien, Germany: Franz Deuticki.

Hughes, M. (1987). The relationship between symbolic and manipulative (object) play. In D. Gorlitz & J. Wohwill (Eds.), *Curiosity, imagination, and play* (pp. 247–257). Hillsdale, NJ: Erlbaum.

Johnson, P. A., & Stockdale, D. E. (1975). Effects of puppet therapy on palmar sweating of hospitalized children. *Johns Hopkins Medical Journal, 137,* 1–5.

Kazdin, A. E. (1990). Psychotherapy for children and adolescents. In M. R. Rosenweig & L. W. Porter (Eds.), *Annual review of psychology* (pp. 21–54). Palo Alto, CA: Annual Review.

Kazdin, A. (1993a, August). Child and adolescent psychotherapy: Models for identifying and developing effective treatments. In S. Eyberg (Chair), *Psychotherapy for children and adolescents*. Symposium conducted at the meeting of the American Psychological Association, Toronto.

Kazdin, A. (1993b). Evaluation in clinical practice: Clinically sensitive and systematic methods of treatment delivery. *Behavior Therapy, 24,* 11–45.

Kazdin, A. Bass, D., Ayers, W., & Rodgers, A. (1990). Empirical and clinical focus of child and adolescent psychotherapy research. *Journal of Consulting and Clinical Psychology, 58,* 729–740.

Kenealy, P. (1989). Children's strategies for coping with depression. *Behavior Research Therapy, 27,* 27–34.

Kessler, J. (1966). *Psychopathology of childhood*. Englewood Cliffs, NJ: Prentice-Hall.

Kessler, J. (1988). *Psychopathology of childhood* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Klein, M. (1927). Symposium on child analysis. *International Journal of Psychoanalysis, 8,* 339–370.

Klein, M. (1955). The psychoanalytic play technique. *American Journal of Orthopsychiatry, 25,* 223–237.

Knell, S. (1993). *Cognitive-behavioral play therapy*. Northvale, NJ: Aronson.

Kogan, N. (1983). Stylistic variation in childhood and adolescence: Creativity, metaphor, and cognitive styles. In P. Mussen (Ed.), *Handbook of child psychology* (Vol. 3, pp. 631–706). New York: Wiley.

Kohut, H. (1977). *The restoration of the self*. New York: International Universities Press.

Kohut, H., & Wolfe, E. R. (1978). The disorders of the self and their treatment: An outline. *International Journal of Psychoanalysis, 59,* 413–424.

Koocher, G., & Broskowski, A. (1977). Issues in the evaluation of mental health services for children. *Professional Psychology, 8,* 583–592.

Koocher, G., & D'Angelo, E. J. (1992). Evolution of practice in child psychotherapy. In. D. K. Freedheim (Ed.), *History of psychotherapy* (pp. 457–492). Washington, DC: American Psychological Association.

Kris, E. (1952). *Psychoanalytic exploration in art*. New York: International Universities Press.

Lieberman, J. N. (1977). *Playfulness: Its relationship to imagination and creativity*. New York: Academic Press.

Levitt, E. E. (1957). The results of psychotherapy with children: An evaluation. *Journal of Consulting Psychology, 21,* 189–196.

Levitt, E. E. (1963). Psychotherapy with children: A further evaluation. *Behavior Research and Therapy, 1,* 45–51.

Levitt, E. E. (1971). Research in psychotherapy with children. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis* (pp. 474–484). New York: Wiley.

Levy, D. (1938). Release therapy in young children. *Psychiatry, 1,* 387–390.

Maddi, S. (1965). Motivational aspects of creativity. *Journal of Personality, 33,* 330–347.

Mahler, M. S. (1968). *On human symbiosis and the vicissitudes of individuation*. New York: International Universities Press.

Mahler, M. S. (1975). On human symbiosis and the vicissitudes of individuation. *Journal of American Psychoanalytic Association, 23,* 740–763.

Mann, E., & McDermott, J. (1983). Play therapy for victims of child abuse and neglect. In C. E. Schaefer & K. J. O'Connor (Eds.), *Handbook of play therapy* (pp. 283–307). New York: Wiley.

Masters, J., Felleman, E., & Barden, R. (1981). Experimental studies of affective states in children. In B. Lahey & A. Kazdin (Eds.), *Advances in clinical child psychology* (pp. 91–118). New York: Plenum.

Meichenbaum, D. (1974). *Cognitive-behavior modification*. New York: Plenum.

Messer, S. B., & Warren, C. S. (1995). Models of brief psychodynamic therapy. New York: Guilford.

Milos, M., & Reiss, S. (1982). Effects of three play conditions on separation anxiety in young children. *Journal of Consulting and Clinical Psychology, 50,* 389–395.

Morrison, D. (1988). The child's fist ways of knowing. In D. Morrison (Ed.), *Organizing early experience: Imagination and cognition in childhood* (pp. 3–14). Amityville NY: Baywood.

Moustakas, C. (1953). *Children in play therapy*. New York: McGraw-Hill.

Moustakas, C. (1992). *Psychotherapy with children: The living relationship*. Greeley, CO: Carron (original work published 1959).

Pelligrini, A. (1992). Rough and tumble play and social problem solving flexibility. *Creativity Research Journal, 5,* 13–26.

Pepler, D. (1979). Effects of convergent and divergent play experience on preschoolers problem-solving behaviors. Unpublished doctoral dissertation, University of Waterloo.

Pepler, D., & Ross, H. (1981). The effects of play on convergent and divergent problem solving. *Child Development, 52,* 1202–1210.

Phillips, R. (1985). Whistling in the dark?: A review of play therapy research. *Psychotherapy, 22,* 752–760.

Rae, W., Worchel, F., Upchurch, J., Sanner, J., & Daniel, C. (1989). The psychosocial impact of play on hospitalized children. *Journal of Pediatric Psychology, 14,* 617–627.

Rosenberg, D. (1984). *The quality and content of preschool fantasy play: Correlates in concurrent social-personality function and early mother-child attachment relationships*. Unpublished dissertation, University of Minnesota, Minneapolis.

Rothenberg, A. (1990). Creativity, mental health, and alcoholism. *Creativity Research Journal, 3,* 179–201.

Runco, M. A. (1994). Creativity and its discontents. In M. Shaw & M. A. Runco (Eds.), *Creativity and affect* (pp. 102–123). Norwood, NJ: Ablex.

Russ, S. W. (1987). Assessment of cognitive affective interaction in children: Creativity, fantasy and play research. In J. E. Butcher & C. Spielberger (Eds.), *Advances in personality assessment* (Vol. 6, pp. 141–155). Hillsdale, NJ: Erlbaum.

Russ, S. (1996). Psychoanalytic theory and creativity: Cognition and affect revisited. In J. Masling & R. Bornstein (Eds.), *Psychoanalytic perspectives on developmental psychology* (pp. 69–103) Washington DC: APA Books.

Russ, S. (in press-a) Creativity and play. In M. Runco (Ed.), *Creativity research handbook* (Vol. 3). Creskill, NJ: Hampton Press.

Russ, S. (in press-b). Teaching child assessment from a developmental-psychodynamic perspective. In L. Handler & M. Hilsenroth (Eds.), *Teaching and learning personality assessment*. Hilldale, NJ: Erlbaum.

Russ, S. W. (1993). *Affect and Creativity: The role of affect and play in the creative process*. A Volume in The Personality Assessment Series. Hillsdale, NJ: Erlbaum.

Russ, S. W. (1995). Play psychotherapy research: State of the Science. In T. Ollendick & R. Prinz (Eds.), *Advances in clinical child psychology, 17* (pp. 365–391). New York: Plenum.

Russ, S. W., & Grossman-McKee, A. (1990). Affective expression in children's fantasy play, primary process thinking on the Rorschach, and divergent thinking, *Journal of Personality Assessment, 54,* 756–771.

Russ, S. W., & Peterson, N. (1990). *The Affect in Play Scale: Predicting creativity and and coping in children*. Manuscript submitted for publication.

Russ, S., Robins, D., & Christiano, B. (1995, March). The Affect in Play Scale: Longitudinal Predication. Paper presented at the meeting of the Society for Personality Assessment, Atlanta.

Saltz, E., Dixon, D., & Johnson, J. (1977). Training disadvantaged preschoolers on various fantasy activities: Effects on cognitive functioning and impulse control. *Child Development, 48,* 367–380.

Sandler, J., Kennedy, H., & Tyson, R. L. (1980). *The technique of child psychoanalysis: Discussion with Anna Freud*. Cambridge, MA: Harvard University Press.

Santostefano, S. (1988). Process and change in child therapy and development: The concept of metaphor. In D. Morrison (Ed.), *Organizing early experience: Imagination and cognition in childhood* (pp. 139–172). Amityville NY: Baywood.

Saracho, O. (1992). Preschool children's cognitive style and play and implications for creativity. *Creativity Research Journal, 5,* 35–47.

Schaefer, C., Gitlin, K., & Sandgrund, A. (1991). *Play diagnoses and assessment*. New York: Wiley.

Shirk, S. R., & Russell, R. (1996). *Change processes in child psychotherapy: Revitalizing treatment and research*. New York: Guilford.

Singer, D. (1993). *Playing for their lives*. New York: Free Press.

Singer, D. L., & Rummo, J. (1973). Ideational creativity and behavioral style in kindergarten-age children. *Developmental Psychology, 8,* 154–161.

Singer, D. & Singer, J. L. (1990). *The house of make-believe*. Cambridge, MA: Harvard University Press.

Singer, J. L. (1973). *Child's world of make-believe*. New York: Academic Press.

Singer, J. L., & Singer, D. L. (1976). Imaginative play and pretending in early childhood: Some experimental approaches. In A. Davids (Ed.), *Child personality and psychopathology* (Vol. 3, pp. 69–112). New York: Wiley.

Singer, J. L., & Singer, D. L. (1981). *Television, imagination, and aggression*. Hillsdale, NJ: Erlbaum.

Smilansky, S. (1968). *The effects of sociodramatic play on disadvantaged preschool children*. New York: Wiley.

Smolucha, F. (1992). A reconstructiion of Vygotsky's theory of creativity. *Creativity Research Journal, 5,* 49–67.

Spielberger, C. D. (1973). *State-trait anxiety inventory for children.* Palo Alto, CA: Consulting Psychological Press.

Stricker, G., & Healey B. (1990). Projective assessment of object relations: A review of the empirical literature. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2,* 219–230.

Sutton-Smith, B. (1966). Piaget on play—a critique. *Psychological Review, 73,* 104–110.

Sutton-Smith, B. (1992). The role of toys in the investigation of playful creativity. *Creativity Research Journal, 5,* 3–11.

Sylva, K., Bruner, J., & Genova, P. (1976). The role of play in the problem solving of children 3–5 years old. In J. Bruner, A. Jolly, & K. Sylva (Eds.), *Play*. New York: Basic Books.

Tuma, J., & Pratt (1982). Clinical child psychology practice and training: A survey. *Journal of Clinical Child Psychology, 11,* 27–34.

Tuma, J., & Russ, S.W. (1993). Psychoanalytic psychotherapy with children. In T. Kratochwill & R. Morris (Eds.), *Handbook of psychotherapy with children and adolescents* (pp. 131–161). Boston: Allyn & Bacon.

Vandenberg, B. (1978). The role of play in the development of insightful tool-using abilities. Paper presented at the American Psychological Association Meeting, Toronto.

Vandenberg, B. (1980). Play, problem-solving, and creativity. *New Directions for Child Development, 9,* 49–68.

Vandenberg, B. (1988). The realities of play. In D. Morrison (Ed.), *Organizing early experience: Imagination and cognition in childhood* (pp. 198–209). Amityville, NY: Baywood.

Vygotsky, L.S. (1967). Vaobraszeniye I tvorchestvo v deskom voraste [Imagination and creativity in childhood]. Moscow: Prosvescheniye. Original work published 1930.

Waelder, R. (1933). Psychoanalytic theory of play. *Psychoanalytic Quarterly, 2,* 208–224.

Wallach, M. (1970). Creativity. In P. Mussen (Ed.), *Carmichael's manual of child psychology* (Vol. 1, pp. 1211–1272). New York: Wiley.

Webber, (1988). Diagnostic intervention with children at risk. In S. Altschul (Ed.), *Childhood bereavement and its aftermath* (pp. 77–105). Madison, WI: International Universities Press.

Weiner, I.C. (1996). Some observations on the validity of the Rorschach Inkblot Method. *Psychological Assessment, 8,* 206–213.

Weisz, J.R., & Weiss, B. (1989). Assessing the effects of clinical-based psychotherapy with children and adolescents. *Journal of Consulting and Clinical Psychology, 57,* 741–746.

Weisz, J., & Weiss, B. (1993). *Effects of psychotherapy with children and adolescents.* Newbury Park, CA: Sage.

Weisz, J.R., Weiss, B., Alick, M.D., & Klotz, M.L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology, 55,* 542–549.

Westen, D. (1991). Social cognitions and object relations. *Psychological Bulletin, 109,* 429–455.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.10
Psychodynamic Therapy

MARY TARGET and PETER FONAGY
University College London, UK

5.10.1 THE PSYCHOANALYTIC ROOTS OF PSYCHODYNAMIC THERAPY 245
- *5.10.1.1 Definition of Psychodynamic Therapy* 245
- *5.10.1.2 Historical Overview of Psychodynamic Developmental Theory* 246
- *5.10.1.3 Evolution of the Psychodynamic Technique with Children* 248
- *5.10.1.4 The Common Assumptions of Psychodynamic Therapy* 250

5.10.2 CONTEMPORARY TECHNIQUE OF PSYCHODYNAMIC PSYCHOTHERAPY WITH CHILDREN 251
- *5.10.2.1 Indications* 251
- *5.10.2.2 Forms of Intervention for Neurotic Children* 252
- *5.10.2.3 Treatment Technique with Non-neurotic Disorders* 254
- *5.10.2.4 Working with the Parents* 255

5.10.3 THE OUTCOME OF PSYCHODYNAMIC TREATMENT FOR CHILDREN AND ADOLESCENTS 256
- *5.10.3.1 Reviews and Meta-analyses of Studies of Efficacy* 256
- *5.10.3.2 Review of Effectiveness in the Clinic* 256
- *5.10.3.3 Individual Studies of Outcome of Psychodynamic or Interpersonal Therapies* 257
- *5.10.3.4 Evaluation of Modifications of Psychodynamic Therapy* 259

5.10.4 LIMITATIONS OF PSYCHODYNAMIC APPROACHES 262

5.10.5 REFERENCES 262

5.10.1 THE PSYCHOANALYTIC ROOTS OF PSYCHODYNAMIC THERAPY

5.10.1.1 Definition of Psychodynamic Therapy

The term psychodynamic therapy covers psychotherapeutic approaches which share the assumption that psychological disorders are rooted in conflicting motivational states, often unconscious, which the individual responds to with a variety of habitual strategies (psychiatric symptoms). Most psychodynamic formulations specify that such conflict is "intrapsychic" (e.g., Brenner, 1982); others include interpersonal conflict, but even there the implication remains that conflict occurs between an internal state and the internal meaning of an external situation (Sullivan, 1953). Psychological intervention is conceived of as assisting individuals to use and develop their inherent capacities for understanding, learning, and emotional responsiveness, in response to the therapeutic relationship and especially the therapist's interpretations of the patient's motivations and strategies, to arrive at more adaptive resolutions. There is no fixed set of techniques to be used in this task, and different therapeutic orientations emphasize different, although substantively overlapping, procedures. Distinctions between types of psychodynamic therapy can be made along several lines; for example, we distinguish psychodynamic individual therapy from group therapy (Rose, 1972) or family approaches (Selvini Palazzoli, Boscolo, Cecchin, & Prata, 1978), expressive from supportive techniques (Luborsky, 1984), Freudian from Kleinian psychoanalytic orientation (King & Steiner, 1991), and therapies may be distinguished according to the relative emphasis of adjuncts such as play (Schaefer & Cangelosi, 1993), art (Simon, 1992), or drama (Johnson, 1982).

This chapter will principally concern itself with individual psychodynamic psychotherapy, using the verbal or play technique.

5.10.1.2 Historical Overview of Psychodynamic Developmental Theory

We will briefly summarize psychoanalytic theories of development, which are considered in more depth Chapter 14, Volume 1.

Freud was the first to give meaning to mental disorder by linking it to childhood experiences (S. Freud & Breuer, 1895), and to the vicissitudes of the developmental process (S. Freud, 1900). One of S. Freud's significant contributions was undoubtedly the recognition of childhood sexuality (S. Freud, 1905). Few, inside or outside psychoanalysis, now doubt that children experience sexual feelings from the earliest days (Green, 1985). S. Freud's discoveries radically altered our perception of the child from one of somewhat idealized naïveté and innocence to that of a human being struggling to adapt his biological, constitutional characteristics to social circumstances. S. Freud's image of the child is of a person in turmoil (S. Freud, 1933), struggling to achieve control over their biological needs, and make them acceptable to society at large through the microcosm of their family (S. Freud, 1930).

Ego psychologists balanced this view by focusing on the evolution of children's adaptive capacities (Hartmann, 1939), which they bring to bear on the struggle with their biological needs. Hartmann's model (Hartmann, 1939) attempted to take a wider view of the developmental process, to link drives and ego functions, and show how negative interpersonal experiences beyond the expectable range could jeopardize evolution of the psychic structures essential to adaptation. He also showed that the reactivation of earlier structures (regression) was the most important component of psychopathology. Hartmann (1995, p. 221) was also among the first to point to the complexity of the developmental process, stating that the reasons for the persistence of particular behavior are likely to be different from the reasons for the original appearance of the behavior, earlier in development. Among the great contributions of ego psychologists are the identification of the ubiquitous nature of intrapsychic conflict throughout development (Brenner, 1982), and the recognition that genetic endowment, as well as interpersonal experiences, may be critical in determining the child's developmental path. The latter idea has echoes in the epidemiological concept of resilience (Garmezy & Masten, 1991; Rutter & Quinton, 1984).

Child analysts (e.g., Fraiberg, 1969, 1980; Anna Freud, 1965) taught us that symptomatology is not a fixed formation, but rather a dynamic entity superimposed upon, and intertwined with, an underlying developmental process. A. Freud's study of symptomatic and asymptomatic children under great social stress led her to formulate a relatively comprehensive developmental theory, where the child's emotional maturity could be identified independently of diagnosable pathology. Particularly in her early work in the war nurseries (A. Freud, 1941–1945), she identified many of the characteristics which later research linked to the so-called resilient child (Rutter, 1990). For example, her observations spoke eloquently of the importance of social support which children could give one another in situations of extreme stress (concentration camps), which could ensure their physical and psychological survival. More recent research on youngsters experiencing trauma has confirmed her assertion of the protective power of sound social support (Garmezy, 1983; MacFarlane, 1987; O'Grady & Metz, 1987; Werner, 1989). Anna Freud's work stayed so close to the external reality of the child that it lent itself to a number of important applications (Goldstein, Freud, & Solnit, 1973, 1979).

A. Freud was also a pioneer in identifying the importance of an equilibrium between developmental processes (A. Freud, 1965). Her work is particularly relevant in explaining why children, deprived of certain capacities by environment or constitution, appear to be at greater risk of psychological disturbance. Epidemiological studies substantiate that harmonious development is threatened by constitutionally acquired weaknesses. Anna Freud's work is remarkable in that she was the first psychoanalyst to place the process and mechanisms of development at the center-stage of psychoanalytic thinking. Her approach is truly one of developmental psychopathology, insofar as she defines abnormal functioning in terms of its deviation from normal development, while at the same time using the understanding gained from clinical cases to illuminate us about the ontogenetic progress of the normal child (Cicchetti, 1990a; Sroufe, 1990).

Mahler (1968) drew attention to the paradox of self-development, that a separate identity implies the giving up of a highly gratifying closeness with the caregiver. Her observations of the "ambitendency" of children in their second year of life were helpful in understanding individuals with chronic problems of consolidating their individuality. Mahler's framework highlights the importance of the caregiver in facilitating separation, and helps explain the difficulties experienced by children whose parents fail to perform a social referencing function for the child, evaluating for them the realistic danger associated with unfamiliar environments (Feinman, 1991; Hornik & Gunnar, 1988). A traumatized, troubled parent may hinder rather than help a child's adaptation (Terr, 1983). An abusive parent may altogether inhibit the process of social referencing (Cicchetti, 1990b; Hesse & Cicchetti, 1982). The pathogenic potential of the withdrawing object, when confronted with the child's wish for separateness, was further elaborated by Masterson (1972) and Rinsley (1977), and is helpful in accounting for the transgenerational aspects of psychological disturbance (see Baron, Gruen, & Asnis, 1985; Links, Steiner, & Huxley, 1988; Loranger, Oldham, & Tullis, 1982).

Sandler's development of Anna Freud's and Jacobson's work (e.g., Sandler & Rosenblatt, 1962) is a coherent integration of the developmental perspective with psychoanalytic theory. His comprehensive psychoanalytic model has permitted developmental researchers (Emde, 1983, 1988a, 1988b; Stern, 1985) to integrate their findings with a psychoanalytic formulation, which clinicians have also been able to use. At the core of Sandler's formulation lies the representational structure which contains both reality and distortion, and is the driving force of psychic life. A further important component of his model is the notion of the background of safety (Sandler, 1987), closely tied to Bowlby's concept of secure attachment.

Klein and her followers constructed a developmental model (Klein, Heimann, Isaacs, & Riviere, 1946) which at the time met great opposition because of the apparently extravagant assumptions these workers were ready to make about the cognitive capacities of infants. However, developmental research appears to be consistent with many of Klein's claims concerning the perception of causality (Bower, 1989) and causal reasoning (Golinkoff, Hardig, Carlson, & Sexton, 1984). Kleinian developmental concepts have become popular because they provide helpful descriptions of the clinical interaction between both child and adult patient and analyst. For example, projective identification depicts the close control that primitive mental function can exert over the analyst's mind. Modern Kleinian psychoanalysts (Bion, 1962, 1963; Rosenfeld, 1971) have been particularly helpful in underscoring the potential impact of emotional conflict on the development of cognitive capacities.

The early relationship with the caregiver emerged as a critical aspect of development from studies of severe character disorders by psychoanalysts in Britain. Fairbairn's focus on the individual's need for the other (Fairbairn, 1952) helped shift psychoanalytic attention from structure to content, and profoundly influenced both British and North American psychoanalytic thinking. As a consequence of this development, the self as a central part of the psychoanalytic model emerged in the work of Balint (1937, 1968) and Winnicott (1971). The concept of the caretaker or false self, as a defensive structure created to master trauma in a context of total dependency, has become an essential developmental construct. Winnicott's (1965) notions of primary maternal preoccupation, transitional phenomena, the holding environment, and the mirroring function of the caregiver, provided a clear research focus for developmentalists interested in individual differences in the development of self-structure. The significance of the parent–child relationship is consistently borne out by developmental studies of psychopathology. These studies in many respects support Winnicott's assertions concerning the traumatic effects of early maternal failure, particularly maternal depression (see Cummings & Davis, 1994), and the importance of maternal sensitivity for the establishment of a secure relationship (Ainsworth, Blehar, Waters, & Wall, 1978; Belsky, Rovine, & Taylor, 1984; Bus & Van IJzendoorn, 1992; Grossmann, Spangler, Suess, & Unzer, 1985).

The central developmental idea of Kohut's formulation (Kohut, 1971, 1977, 1984) was the need for an understanding caretaker to counteract the infant's sense of helplessness in face of biological striving for mastery. Kohut emphasizes the need for such understanding objects throughout life and these notions are consistent with accumulating evidence for the powerful protective influence of social support identified across a wide range of epidemiological investigations (Brown & Harris, 1978; Brown, Harris, & Bifulco, 1986). The mirroring object becomes a "self-object," and the need for empathy drives development, which culminates in the attainment of a cohesive self. Drive theory becomes secondary to self-theory, in that the failure to attain an integrated self-structure both leaves room for, and in itself generates, both aggression and isolated sexual fixation. However, the self remains problematic as a construct; in Kohut's model, it is both the person (the patient) and the agent which is assumed to drive the person (Stolorow, Brandchaft, & Atwood, 1987). However, Kohut's descriptions of the narcissistic personality have been powerful and influential examples of the use of developmental theory in psychoanalytic understanding. Kohut's focus on self-esteem has been shown by empirical research to be appropriate and somewhat lacking in previous psychoanalytic formulations.

Kernberg's systematic integration of structural theory and object relations theory (Kernberg, 1976, 1982, 1987) is probably the most frequently used psychoanalytic model, particularly in relation to personality disorders. His model of psychopathology is developmental, in the sense that personality disturbance is seen to reflect the limited capacities of the young child to address intrapsychic conflict. Neurotic object relations show much less defensive disintegration of the representation of self and objects into libidinally invested part–object relations. In personality disorder, part–object relations are formed under the impact of diffuse, overwhelming emotional states, ecstatic, or of an equally overwhelming but terrifying and painful nature which signals the activation of aggressive or persecutory relations between self and object. Kernberg's models are particularly useful because of their level of detail and his determination to operationalize his ideas far more than has been traditional in psychoanalytic writing.

Bowlby's (1969, 1973, 1980) work on separation and loss also focused developmentalists' attention on the importance of the security (safety and predictability) of the earliest relationships. His cognitive-systems model of the internalization of interpersonal relationships (internal working models), consistent with object relations theory (Fairbairn, 1952; Kernberg, 1975) and elaborated by other attachment theorists (Bretherton, 1985; Crittenden, 1990; Main, Kaplan, & Cassidy, 1985), has been particularly influential.

According to Bowlby, the child develops expectations regarding a caregiver's behavior and the child's own behavior.These expectations are based on the child's understanding of experiences of previous interaction, and are helpful in organizing the child's behavior with the attachment figure and (through a process of extrapolation with others) with others. The concept has had very broad application. Bowlby's developmental model highlights the transgenerational nature of internal working models: our view of ourselves depends upon the working model of relationships which characterized our caregivers. Empirical research which places this intergenerational model under systematic scrutiny has begun, and the robustness of the findings is encouraging. There is an accumulating body of data which confirms that there is some kind of intergenerational transmission of attachment security and insecurity (see review by van IJzendoorn, 1995) and that parental mental representations pertinent to this process may be assessed before the birth of the first child (e.g., Fonagy, Steele, Moran, Steele, & Higgitt, 1991a; Fonagy, Steele, & Steele, 1991b; Ward & Carlson, 1995).

Stern's (1985) work represents a milestone in psychoanalytic theories of development. His work is remarkable within psychoanalysis for being normative rather than pathomorphic, and prospective rather than retrospective. His focus is the reorganization of subjective perspectives on self and other as these occur with the emergence of new maturational capacities. He is perhaps closest to Sandler in his psychoanalytic model of the mind, but his formulation of object relations also has much in common with those of Bowlby and Kernberg.

5.10.1.3 Evolution of the Psychodynamic Technique with Children

Specific therapeutic techniques which could help psychoanalysts in addressing psychiatric problems of childhood did not appear until the 1920s. Application of the insights gained from psychoanalytic treatment of adults to the treatment of children was prepared by Freud's observations concerning the psychology of young children. The most notable ones include Anna Freud's wishful dream for strawberries (S. Freud, 1900), and his grandson's separation game (S. Freud, 1920). Best known is S. Freud's case study of Hans (S. Freud, 1909), a five-year-old with an animal phobia, whose treatment by his physician father was supervised by S. Freud. All these observations were used by S. Freud principally to confirm his assumptions about infantile instinctual life through the direct observation of children and to prevent him (and other adult psychoanalysts) from making false developmental assumptions (S. Freud, 1926, p. 215).

Hug-Helmuth (1920, 1921) was the first clinician to introduce play therapy into the psychoanalyst's armamentarium. Her pioneering work, which is now largely forgotten, combined an insight-oriented technique, focused primarily on the child's unconscious sexual fantasies, with a powerful developmental perspective, whereby she saw the young person as needing to be "strengthened" in the mental capacities needed for his developmental tasks. This latter emphasis has echoes in present-day cognitive therapy and psychoanalytic approaches. Up to that point, most analysts shared the Hungarian analyst Ferenczi's skepticism about offering psychoanalysis to children who do not wish to be analyzed as they solely wish to play (see Group for the Advancement of Psychiatry, 1982, for a comprehensive history of the field).

The other great pioneers of the field, A. Freud (1946) and Klein (1932), independently (but frequently with reference to one another), evolved techniques to enable clinicians to take a psychoanalytic therapeutic approach to children. Klein's approach, working under the influence of Abraham, a Berlin based psychoanalyst, was to regard children's play as essentially the same as the free association of adults, motivated by unconscious fantasy activated principally by the relationship with the analyst and requiring verbalization (interpretation) if the child's anxiety was to be adequately addressed. Interpretation of the child's deep anxieties concerning destructive and sadistic impulses was the principal focus of child analytic work. She advocated that therapists establish an interpretative relationship, even with preschool children, from the beginning. The emphasis placed on the relationship with the analyst meant that work with parents and other adults in the child's life (e.g., teachers) was not seen as central.

Direct interpretation of the projective processes was seen by Klein as critical. In *Notes on some schizoid mechanisms* (Klein, 1946) and in her paper on *The origins of transference* (Klein, 1952), Klein discussed the importance of the common childhood unconscious fantasy of placing part of the self into another person and perceiving unwanted qualities in them rather than oneself in order to relieve oneself of unwanted feelings (such as greed or envy). This form of fragmentation of the sense of self was regarded by her as part of normal infantile development but when persisting beyond infancy was the cause of pathology and the key focus for interpretation. She assumed that while projective identification was distressing to the child it was also helpful in creating the fantasy, not only that others were the "containers" of the child's own unwelcome traits, but also that (as these unattractive attributes still partly belonged to the self) the child could control the other person (object). The clinician's understanding was enriched by the assumption that a child's perception of the clinician gave clues about conflictual aspects of the child's experience of his self. Bion (1959) showed how such projective experiences could be expected to have an impact on "the container" and how the capacity of that individual to "metabolize" (understand and accept) the projection may be critical in development as well as in the success of therapy. Thus, the therapist's subjective experience (countertransference) could be a clue to the child's unconscious fantasies, and her capacity to understand and tolerate these, the key component of successful treatment.

Modern Kleinians (e.g., DeFolch, 1988; O'Shaughnessy, 1988) have, to some degree, modified the classical Kleinian position; early interpretations of assumed deeply unconscious material are less frequently offered, and there is greater attention to defensive qualities of many manifestations of the child's nonconscious processes. The immediate interaction with the analyst, however, remains the core focus of therapeutic work, and the underlying conceptualization remains based on the notion of fragmentation of the self-representational structure which may be undone through verbalization and interpretation. The countertransference experience of the therapist is the central guide.

Drawing on the work of Klein, Winnicott (1965, 1971) reinforced Klein's emphasis on the influence of early life on childhood pathology while introducing additional techniques into child analysis. For example, specific drawing techniques were used, while Winnicott also emphasized nonverbal aspects of the therapist's stance, in particular the importance of a "holding" environment and the central role of play. One of his major contributions was the concept of a transitional area between self and object where the subjective object and the truly objective object are simultaneously recognized (Winnicott, 1971). This insight was critical in developing an appreciation of the interpersonal nature of therapeutic interaction in child analysis (Altman, 1994).

By contrast, A. Freud's approach placed more emphasis on the child's developmental struggle with adaptation to a social as well as an intrapsychic environment. Her training as a nursery school teacher led her to be very concerned with the child's actual external circumstances, as well as their unconscious world. She made fewer assumptions concerning the meaning of the child's play, approached therapeutic work far more gradually, recommended working in collaboration with parents and teachers (particularly in communicating understanding derived from the therapy), and focused far more on the complications and conflicts arising from the child's libidinal (sexual) impulses than on innate aggression. Her focus was on the child's wish to protect their fragile internal world from conflict by adopting psychological strategies (mechanisms of defense, A. Freud, 1936) such as denial, repression, or identification with the aggressor. The focus of her technique therefore was the interpretation of defenses and through this, the anxieties which motivated them.

The Anna Freudian approach to child analysis invariably takes into consideration the limitations upon the child's "ego functioning" imposed by development, and has as its primary focus the support of the development of the ego and the restoration of the child to a normal developmental path. Pine (1985) stressed that the analyst, like the parent, creates a supportive environment for the child's incompletely developed ego. He saw such techniques as mutative in their own right and considered interpretations in the context of support to be qualitatively different from interpretations in the context of abstinence, as is recommended by Kleinian child analysts. The pressure for the analyst to be an active and "real" participant in the therapeutic situation has thus been growing apace (see also Altman, 1992; Warshaw, 1992). However, many of these reconceptualizations lack coherent theoretical rationale and specific technical recommendations as to how departures from abstinence may be put to good therapeutic effect.

Anna Freudian child psychoanalysis became popular as a treatment in the US in the first half of the twentieth century, and was systematically described by A. Freud and her colleagues (Sandler, Kennedy, & Tyson, 1980); it has also influenced many forms of psychodynamic treatment of children and families. Individual child psychodynamic psychotherapy, based on these principles, has been shown to be a form of treatment frequently used and highly regarded among child psychiatrists and psychologists in the US (Kazdin, Bass, Siegel, & Thomas, 1990a). In the UK, as well as in Latin America, Klein's model proved to be more popular.

5.10.1.4 The Common Assumptions of Psychodynamic Therapy

Psychodynamic child psychotherapy shares many of its basic assumptions with adult psychoanalysis. Numerous summaries of these are available (see Chapter 14, Volume 1, Fonagy, Target, Steele, & Gerber, 1995; Gabbard, 1994, 1995; Kernberg, 1995; Summers, 1994, for more comprehensive accounts) and here we shall only provide a brief overview.

Psychodynamically oriented child therapists assume that mental disturbance may be usefully studied at the level of *psychological causation*, that the representation of past experience, its interpretation, and meaning, largely unconscious, determine the child's reaction to his external world and capacity to adapt to it. The emphasis on psychic causation does not imply either a lack of respect for or inattention to other levels of analysis of childhood psychiatric problems such as the biological, the family, or broader social factors. Nevertheless, psychiatric problems, whether at the root genetic, constitutional, or socially caused, are seen by the psychodynamically oriented child clinician also as the meaningful consequence of the child's unconscious beliefs, thoughts, and feelings, and therefore, potentially accessible in psychotherapy.

Complex unconscious mental processes are assumed to be responsible for the content of thinking and behavior. In particular, unconscious fantasies associated with wishes for instinctual pleasure or safety (Sandler, 1987) motivate the child's behavior, affect regulation, and the ability to cope with the social environment.

The experience of the self with others is internalized and leads to representational structures of interpersonal interactions which, at the simplest level, determine expectations of others and more elaborately determine the "shape" of self and other representations, and the nature of the internal world of the child. This general idea is formulated somewhat differently by psychodynamic clinicians from different traditions. Bowlby's concept of "internal working models" of self–other relationships based on the infant–mother relationship is perhaps the closest to formulations from other areas of clinical psychology (Bowlby, 1973, 1980). In essence, all so-called "object-relations" theories posit that the emotional life of the child (and adult) is organized around mental representations, however partial, of the self in relation to an important figure, imbued with a specific affect (e.g., Kernberg, 1976).

It is assumed that psychic conflict is ubiquitous and a cause of suffering (or a felt lack of safety) (Brenner, 1982). Adverse environments of childhood may create intrapsychic conflicts of overwhelming intensity and/or fail to equip the child adequately to deal with conflicts within the normal range of early experience (Winnicott, 1965). Trauma (such as loss of a caregiver) or long-term abuse thus undermine personality development by intensifying incompatible wishes or reducing the child's capacity to resolve conflicts through mental work.

The child is predisposed to modify unacceptable unconscious wishes through a range of mental mechanisms aimed at reducing the sense of conflict. Defense mechanisms form a developmental hierarchy which reflects the degree of pathology experienced by the individual; developmentally early defenses, such as splitting, or projective identification, are normally associated with more severe disturbances. A neurotic symptom, such as phobic anxiety, may be understood as a result of displacement of fear from the representation of one model of interaction (e.g., father–child) to another (e.g., teacher–child). Considerably more resistant to therapy is the more primitive defense of splitting of affect, whereby a child alternately derogates and idealizes the caregiving figure. Immature or early defenses are assumed to reflect the absence of higher level, integrative capacities (Pine, 1985).

Psychodynamic therapists assume that the child's communication in the session has meaning beyond that intended by the child and, by analogous mechanisms, that the child's symptoms carry multiple meanings and reflect the nature of internal representations of others and their relationship to the child as the child perceives it. The therapist is able to bring the child's attention to aspects of his behavior which are ego-dystonic and hard to understand, and by making appropriate links illustrates to the child that they may be seen as rational in terms of unconscious mental experience and psychic causation (Sandler et al., 1980).

The relationship to the therapist has primacy, in that it provides a window on the child's expectations of others and can come to be a vehicle for the unwanted and disowned aspects of the child's thoughts and feelings. Transference displacement may include such aspects of past relationships or past fantasies about these as well as conflictual aspects of current relationships to parents, siblings, or important others. The child's verbal and nonverbal behavior (reenactments) have an impact on the clinician, and through exploring the role he or she has been placed in by the child, the therapist is able to achieve a better understanding of the child's representation of role relationships and feelings about them (Tyson & Tyson, 1986).

Modern psychodynamic child psychotherapy emphasizes the current state of the child in relation to his environment, history of past relationships, and adaptations to these. Psychotherapists generally recognize that the therapy they offer has an important holding or containing function in the child's life which, beyond interpretation and consequent insight, creates the possibility of a reintegration or reorganization of the child's internal world that in turn facilitates the child's adaptive development. The child therapist takes a "whole person" perspective, encompassing all aspects of the child's unfolding concerns (biological, environmental, intrapsychic). The establishment of a relationship with an adult which is as far as possible open and nonexploitative may serve as the basis of new internalizations, bringing about a healthier resolution of pathogenic experiences.

5.10.2 CONTEMPORARY TECHNIQUE OF PSYCHODYNAMIC PSYCHOTHERAPY WITH CHILDREN

5.10.2.1 Indications

There has been general agreement on the ideal indications for child psychotherapy (Dowling & Naegele, 1995). These have traditionally included: (i) good intelligence and capacity to verbalize; (ii) a relatively supportive environment; (iii) diagnostic assessment indicating the primacy of internal conflict underlying symptomatology; (iv) the absence of gross interferences between parents and child (indicating adequate internal–external object relations); (v) the presence of anxiety; and (vi) the absence of pervasive developmental disorder, psychosis, major deficiencies in psychological capacities, or family constellations likely to interfere with treatment (chaotic environments and uncooperative parents). Although there have always been a number of child psychotherapists who have attempted to work with psychotic and autistic children (e.g., Alvarez, 1993; Klein, 1936), the majority have felt that they cannot help these children using psychoanalytically oriented work.

Child therapists have also addressed other disorders, often called borderline (Rosenfeld & Sprince, 1963). This group has been described on the basis of records at the Anna Freud Centre (Fonagy & Target, 1996a, 1996b). Others have taken a more traditional, descriptive approach and arrived at very similar characterizations (Cohen, Towbin, Mayes, & Volkmar, 1994; Towbin, Dykens, Pearson, & Cohen, 1993). These children appear to suffer from a variety of deficiencies of psychological capacities, shown by, for instance, lack of control over affect, lack of stable self- and other-representations, and diffusion of the sense of identity. As we shall see, such children require substantial modifications of traditional technique (Fonagy & Moran, 1991). Whereas the less severe, so-called neurotic, cases are well-served by a therapeutic process aimed at addressing distortions in their mental representations of others or of their relationships to them, we have suggested that the more severely disturbed group of patients require developmental therapy which can strengthen and support their adaptive defenses, and assist them in labeling and verbalizing their thoughts and feelings (Fonagy & Target, 1996a). This distinction is most readily conceptualized psychologically in terms of the classical distinction between mental representations and the mental processes which generate them. Whereas so-called classical technique has its main effect on the organization and "shape" of the child's mental representations (Sandler & Rosenblatt, 1962), developmental therapy aims to help develop mental processes which have been inhibited during early development (Fonagy, Moran, Edgcumbe, Kennedy, & Target, 1993).

5.10.2.2 Forms of Intervention for Neurotic Children

Psychodynamic child psychotherapy involves the elaboration and correction of distorted, and often unconscious mental representations. The therapist, on the bases of the child's verbalization, nonverbal play, and other behaviors, attempts to construct a model of the child's conscious and unconscious thoughts and feelings and, using this model, helps the child obtain insight into why they may seem irrational and inappropriate. Such understanding may result in the integration of developmentally earlier modes of thinking into a more mature, age-appropriate framework.

This treatment involves the use of toys, games, and other devices to engage children in a process of self-exploration with an adult who is friendly but tries, within a trusting relationship, to draw attention to the unconscious determinants of the child's behavior. The therapist uses the child's fantasy, games, and associations in conjunction with other sources of data from the family, school etc., to construct a hypothetical picture of the child's unconscious mental life and current emotional concerns. In this way the therapist aims to help the child to understand his emotional responses, confusions, concerns about his body, the unconscious meaning of his presenting symptoms for the child, anxieties about unconscious aggressive or sexual impulses, and worries about his relationships with caregivers, siblings, and peers.

The techniques used by psychodynamic child therapists have been systematized, on the basis of empirical observation, by P. Kernberg (1995). In her studies, she was able to reliably distinguish a number of interventions on the part of the therapist, in addition to interpretive work, which may facilitate change in the patient's mental representations. In particular, she has delineated supportive interventions aimed at allaying the child's anxiety, and increasing a sense of competence and mastery through information, suggestions, reassurance, and empathy. In addition, she has observed facilitative statements which enhance and maintain therapeutic exchange with the child, when the therapist reviews, summarizes, or paraphrases the child's communication to that point. Related to this are clarifications aiming to either prepare the child for an interpretation, or direct his attention to noticeable aspects of his behavior. Although these are important distinctions, they are not placed within a theoretical context, and the therapeutic impact of the process is still seen as chiefly related to interpretations.

Analysts consider a number of aspects of a child's material when they try to interpret it. These include: (i) the main things the child talks about, especially what they say about themself, others, and the analyst, and the nature of the relationship these suggest between themself and objects in their mind; (ii) what is left out, such as not referring to arguments between the parents, to school, or to other important events in the child's life of which the analyst is aware; (iii) their nonverbal behavior, such as infantile gestures, unusual levels of activity, or avoiding entering the consulting room; (iv) their play, including roles assigned to the child, analyst, or play figures; (v) signs of the child's emotional state; and (vi) dreams that the child tells the analyst about.

The analyst is informed by, but does not necessarily mention, the child's past history and the contents of the previous sessions. While the analyst may often be able to make important connections between the child's analytic material and their past experiences, interpretations early on are generally restricted to the current conflicts which the child seems to be experiencing, in the immediate context through which the child brings these into the analysis. However, ultimately the analyst will aim to help the child develop an emotionally meaningful understanding of the connections between past experience and current conflict.

Kernberg (1995) distinguishes between three aspects of interpretations: interpretations of defenses show children how the therapist thinks they protect themselves from unacceptable thoughts, feelings, and actions. For instance, the therapist might suggest to the child "You often say that you are no good at things, perhaps because you are worried that people will think that you are big headed." Drawing the child's attention to these serves a dual function in that what is defended against is brought into awareness and, by implication, the child is also prompted to consider alternative strategies for coping with or expressing the ideas, impulses, or feelings previously warded off.

A second class of interpretation may address the child's wishes, which the therapist thinks may underlie his behavior. Most frequently, these two classes of interpretation are offered together, with the therapist leading the patient to see the unconscious wish by showing the child how he protects himself from it. The therapist might say "I think you secretly wish to feel much better than everyone else, and that is why you are frightened of seeming to be big-headed."

Some child therapists address the child's past experience as part of a causal account of the child's current state of mind. These reconstructions may be helpful, although few psychodynamic child therapists would nowadays feel that the actual past experience needs to be uncovered in order for the child's problems to be adequately addressed. More commonly, interpretations may make reference to current experience, particularly in relation to the therapist. The latter, so-called transference interpretations aim to identify the role in which the therapist experiences being placed by the child. The analyst further assumes that this role corresponds to an internal representation of a relationship in the child's mind. Thus, for example, a child who places the therapist in the role of a critic is assumed to be externalizing an internal representation of a figure who constantly undermines the self, and bombards it with disparagement and criticism. "I think you are so worried about me criticizing you because there is a terrible voice inside your head which says that you are such a naughty child that nobody could love you." It is important to note that analysts do not assume that such a figure actually exists in the child's life; an "internal object" may be a separated off part of the child's self-representation, which has been felt to be too destructive and aggressive for the child to be able to like himself. Interpretation eventually addresses this facet as well, by showing the child how unacceptable he finds the small amount of residual envy or destructiveness which has remained part of the self-structure, and through this gradually modifies the child's internal standards, allowing it to experience his destructiveness as less catastrophic.

Whether the content of the interpretation concerns defenses, wishes, or past or current experience, the underlying structure of child analytic interpretations is largely similar. The analyst focuses on the child's anxiety in relation to the above domains, in particular the emotional states they generate. Thus, destructive impulses are most likely to be taken up in terms of the anxiety the child feels about his angry feelings. A further important focus is explicit recognition that the child is actively struggling against certain wishes, and that aspects of current and past experience have necessitated the use of defenses. In psychoanalytic shorthand, this is referred to as "interpreting the conflict." Thus, rather than saying "You seem to be feeling very angry with me today," an analyst might say "You seem anxious, maybe because you feel angry." Naturally, this could be an inappropriate interpretation if the child was consciously aware of feeling angry, and not in conflict about this. However, in this case too, the analyst attempts to find the source of any current anxiety, for example, that the anger might destroy the otherwise valued relationship with the analyst.

The middle phase of analysis focuses particularly on the systematic use of transference interpretations to begin the "working through" process.

Transference is facilitated by: (i) the therapist's implicit offer of help and relatedness coupled with a basic stance of acceptance, encouragement to express thoughts and feelings, attention to resistances, and attunement to the child's predicaments; and (ii) a therapeutic structure that emphasizes regularity, consistency, and the specialness of the sessions (Chethik, 1989). Every action of the therapist conveys to the child that this time and space is set aside to better understand what bothers the child.

Signs of readiness for termination of the therapy are found both within and outside the sessions. Of course, sustained symptomatic improvement is important. Improvements in family and peer relationships are noted. The child's ability to make use of parents, teachers, and other people as sources of help and advice, and as models for identification, suggests that the analysis is becoming less necessary. Perhaps the most important indication for termination is evidence that the child is able to take advantage of normal developmental opportunities, to cope more effectively with stress and conflict, and to respond with greater freedom and mastery to demands for adaptation, coming from inside and outside.

Kernberg (1995) proposed the following specific features as showing readiness to end the therapy: (i) the child's statements about the therapist during the session convey an internalization of the analytic function, for example, comments about their sense of having been listened to or helped to make connections; (ii) the therapist finds more opportunities for interpretation, as opposed to clarifications or confrontations; (iii) the child introduces more material from current reality, suggesting a healthy decrease of investment in the analysis; (iv) there is more reflectiveness and capacity to use interpretations and resolve conflicts; (v) greater freedom, expressiveness, and pleasure in play; (vi) increased modulation of feelings, which permits discussion of experiences of sadness, concern, and gratitude; (vii) evidence of new adaptive interests; (viii) insight, humor, and a healthy capacity in the child to laugh at himself; (ix) flexible and adaptive coping and defensive operations, permitting the child to assume responsibility for their own actions; and (x) predominance of age-appropriate behavior, posture, and appearance.

5.10.2.3 Treatment Technique with Non-neurotic Disorders

More recently, the psychodynamic approach has been extended to apply not only to so-called neurotic disorders, but also to the understanding and treatment of borderline, narcissistic, delinquent, and conduct disordered youngsters (Bleiberg, 1987, 1994a, 1994b; Marohn, 1991; O'Brien, 1992; Rinsley, 1989), as well as schizoid and even psychotic children (e.g., Cantor & Kestenbaum, 1986). Clinical work with these more severely disturbed children quickly highlights the limitations of the "classical" analytic strategy. For example, the child's anxiety may not be accessible to awareness, and referring to it simply leaves the child confused. Also, there may be little evidence that the child is "struggling" against wishes, even when these are clearly socially unacceptable. In other children, defenses, as normally conceived, may be hard to identify. All these cases remind us of the need to reconsider child psychoanalytic techniques.

Bleiberg, Fonagy, and Target (1997) identify two clusters of youngsters who may be regarded as suitable for modified psychodynamic therapy, notwithstanding the severity of their disorder. They suggest that these clusters have in common the presence of at least one emotional disorder, for example, depression, dysthymia, generalized anxiety disorder (GAD), separation anxiety disorder, and social phobia; however, these symptoms must be seen in the context of a broader disturbance of social and emotional development, including marked impairment of peer relationships, affect regulation, frustration tolerance, and poor self-esteem.

One subgroup, designated cluster A, are characterized by fragile reality contact and thought organization, idiosyncratic magical thinking, ideas of reference, suspiciousness, and deep discomfort in social situations. They can neither make full sense of the social world, nor communicate adequately their internal states. They resemble the cluster A *DSM-IV* personality disorder diagnoses for adults and children described by Cohen et al. (1994), and Towbin et al. (1993) as multiplex personality disorder. By contrast, cluster B children show (if referred below school age) a hunger for social response, intense often dramatic affect, clinginess, hyperactivity, and temper tantrums. School-age children may meet axis 1 criteria for attention deficit hyperactivity disorder (ADHD), conduct disorder, or mood disorder, but their lack of adequate affect regulation is the major feature of the picture: a sense of elation and blissful merger with others close to the child seems to alternate with rage and self-hatred. They have been linked to adult "dramatic" personality disorders (cluster B), and have been described by Bleiberg (1994a, 1994b) and by Petti and Vela (1990). We have linked both clusters to an impairment of reflective function (Bleiberg et al., 1997; Fonagy & Target, 1996a). As psychodynamic therapists have extended the scope of their work to include these groups of patients, their therapeutic approaches have increasingly emphasized aspects of interaction between child and therapist which promote opportunities for playing with ideas. To some degree, this is a shift of emphasis within existing therapeutic techniques, at the same time as including some new techniques which have previously been systematically excluded from psychodynamic work with neurotic children, because of the expected interference with therapeutic neutrality. All these strategies have in common an emphasis on the therapeutic value of play, not just in the sense of pretending, but of creating a safe opportunity for playing with alternative meanings of the child's experiences.

The main forms of intervention include: (i) enhancing reflective processes by encouraging the observation and verbalization of what it is the child is feeling, focusing on somatic and psychological experiences in the immediate situation; (ii) enhancing impulse control, by looking for ways of channeling impulses into socially acceptable forms of behavior, and increasing control over expression of feeling in action; (iii) enhancing cognitive self-regulation, both through strategies involving symbolization or metaphor, and through demonstrating the therapist's own capacities for reflection and moderation of experience through thinking and talking, rather than physical action or coercion; (iv) encouraging awareness of the feelings and thoughts of others, often initially focusing on the child's perceptions of the therapist's mental states, as a precursor to self-reflection; and (v) developing play, perhaps first with objects, then another person, ultimately with ideas, with the aim of showing the child that habitual thoughts and feelings are not the only way of seeing things, or necessarily the way that others see the world, and that not all adults relate to them as they may have been related to before.

Techniques with developmental disturbances does not involve working with the transference in the sense of expecting the child to "transfer" their thoughts and feelings about their parents onto the analyst. However, the child's feelings about the therapist remain central, because the clarification of the child's feelings about themself and about the therapist is sometimes the most effective route towards acquiring mentalizing capacity. This requires the therapist to convey that the child's affect can be understood and managed by another person, rather than compounding what might have been the child's original experience, of a parent who has to twist or cut off emotional contact with the child to make the closeness of an intensive relationship more bearable. Late in the therapy with a developmentally disturbed child, it may also become helpful to make interpretations of neurotic conflict, which would previously have been experienced as meaningless or invasive.

The approach described is based on the assumption that reflective or mentalizing function is a key, biologically prepared human capacity triggered by specific aspects of normal parenting. A number of children—which we group into two developmental and clinical clusters—suffer from combinations of constitutional vulnerability and environmental disadvantage that create significant deficits in their ability to grasp mental states and find symbolic meaning in human behavior. This crucial deficit, as we have described, generates a host of problems in subsequent development and adjustment, including difficulties in relationships, impulse control, self-esteem, and overall difficulty in taking advantage of life's opportunities for growth.

We propose for these children a systematic psychodynamic psychotherapeutic approach aimed at enhancing the capacity to mentalize. While achieving such a goal is a by-product of traditional child psychoanalytic technique—and "psychological mindedness" is considered a precondition for child analysis proper—we suggest that a well-defined and deliberate focus on enhancing mentalization can lead to more effective and more specifically tailored approaches for a broad range of children with serious disturbances.

A therapeutic focus on mentalization, of course, does not preclude a variety of other therapeutic approaches, such as pharmacotherapy, family therapy, and cognitive-behavioral treatment. In effect, we propose that a psychoanalytic approach which promotes mentalization provides a conceptual glue to hold together a variety of interventions in a coherent treatment plan of greater clinical effectiveness (Bleiberg, 1994a, 1994b; O'Brien, Pilowsky, & Lewis, 1992).

Recasting psychoanalytic child therapy, from a predominantly insight-oriented, conflict-solving modality to a developmentally-based, mentalization-enhancing approach, brings it closer to the mainstream of contemporary developmental research and other, more empirically supported therapies. In particular, commonalities—and differences—between the approach described here and cognitive therapy warrant brief discussion. Both cognitive therapy and this approach aim at effecting a change in the ways people organize and structure experiences. Cognitive therapy, however, is both more "active" and directive and more narrowly focused on specific maladaptive cognitive schemas, for example, the particular cognitions underlying low self-esteem. By contrast, enhancing mentalization—while more targeted and "active" than traditional child psychoanalyses—aims at promoting a broad set of capacities, or mental processes, rather than correcting specific cognitive distortions.

The most important differences between the two approaches arise from the psychoanalytic origins of the psychotherapeutic strategy described in this chapter.

(i) These approaches are based on the assumption that the child's symptoms, and the entire organization of their subjective world, primarily represents an unconscious adaptation to the child's perceptions of reality. Child analysts therefore anticipate that children will try to resist the therapist's efforts, and the possibility of change: both of these, though well intentioned, threaten the child's best effort at adaptation thus far.

(ii) Child analysts have been trained to look for, respect, and address the child's deepest anxieties, which are likely to include anxiety about exercising the particular capacity that the treatment aims to promote, the recognition of mental experiences, and habitually thinking in terms of these to give the social world meaning and predictability.

(iii) Psychoanalytic training, supervision, and personal treatment should have helped the therapist to use her own thoughts and feelings to understand and empathize with the child's inner world, without getting stuck in response to the child's "concrete" ways of relating to themself and others.

5.10.2.4 Working with the Parents

Regardless of the child's predominant disturbance, it is generally helpful for the therapist to have meetings with the parents or other caregivers, aiming to enhance their sensitivity to the child's worries, limitations, and needs. This may involve work on aspects of the child's environment or family which seem to impinge on the child's development, and the exploration of ways of freeing the child from avoidable stresses. Sometimes this involves offering parents therapy for themselves, but much more often it is enough to help the parent to think more objectively and productively about the child, and to find together ways of overcoming obstacles to the child's positive development. Interventions range from concrete suggestions and strategies to helping the parents to be more understanding of their child's thoughts and feelings. The analyst may show the parent distortions affecting the relationship, for instance, where the mother identifies her son with her husband, and then unconsciously expresses the anger she feels towards her husband to the child, generally an easier and more available target.

Discussion of the child's symptoms aims to help the parents to develop more awareness of the nature of the child's difficulties and feelings, and (often) insight into possible reasons for these to have arisen and become entrenched. Parents may tend to minimize the difficulties facing the child. For example, in one family the parents needed to be helped to recognize their child's distress when they compared the child's ability to make friends with that of his much brighter and more attractive brother; the parents had felt that this comparison would encourage the child to try harder, not understanding that the child's difficulty went far beyond what could be corrected by greater effort. This sort of discussion may be combined with thinking about better ways of managing practical problems, for instance, major tantrums in response to separation. The parents can also be shown connections between external events and stresses and symptoms, whether somatic, emotional, or behavioral. Some stresses (e.g., an inappropriate school setting) can be removed, while others (such as having a physically handicapped sibling) of course cannot, but even here ways of limiting pressure and damage may be found.

5.10.3 THE OUTCOME OF PSYCHODYNAMIC TREATMENT FOR CHILDREN AND ADOLESCENTS

Research in the field of child therapies has very much lagged behind the adult literature, in terms of both number of studies and quality of design (e.g., Barnett, Docherty, & Frommelt, 1991). This lag has been greater still in the area of psychodynamic therapies, partly because of a traditional antagonism among psychodynamic clinicians towards empirical evaluation, but also because of specific methodological difficulties in the design of studies of long-term, intensive treatments, where choosing a suitable control group poses practical and ethical problems, and the treatment goals are relatively difficult to operationalize (usually what is sought is a change in unconscious psychic structure shown through subtle but pervasive personality changes, as opposed to change in more specific symptoms). For these reasons, evaluations of nonbehavioral therapies for children have most often been conducted by behaviorally-oriented clinicians, as a comparison condition for a trial of behavior or cognitive-behavior therapy (Shirk & Russell, 1992); they have also generally been administered in the form of brief, group therapy (to maintain comparability with the other treatment being evaluated) rather than longer-term, individual sessions as would be used in clinical practice (Kazdin, Siegel, & Bass, 1990b; Silver & Silver, 1983).

5.10.3.1 Reviews and Meta-analyses of Studies of Efficacy

Possibly partly for the reasons given above, meta-analyses of child therapy outcome studies have found psychodynamic and interpersonal therapies to be less effective than behaviourally-oriented techniques (Casey & Berman, 1985; Weisz, Weiss, Alicke, & Klotz, 1987). Weisz et al. (1987), for example, found an effect size exactly twice as great for behavioral therapies (0.88 vs. 0.44, $p < 0.05$). A further factor contributing to the difference has been identified by the authors of these meta-analyses, who found in both cases that the difference was reduced to nonsignificance once outcome measures very close to techniques practiced in the behavior therapy sessions were excluded from the comparison. Nevertheless, it remains a challenging fact that across the three studies which included some variant of brief, psychodynamic therapy, the mean effect size was negligible (0.01). If this were due to the use of therapeutic procedures in research studies which would never be expected to be effective in clinical practice, as persuasively argued by Shirk and Russell (1992), then it is imperative that the superior effectiveness of the procedures which would be used clinically be demonstrated. Thus far, the evidence from reviews of routine clinical work is dispiriting for psychodynamic therapy, as it is for child therapies of all orientations (see below), but the very small number of studies of outcome for this sort of therapy in a research setting which *do* use normal clinical procedures give more reason to be optimistic (see Section 5.10.3.3).

5.10.3.2 Review of Effectiveness in the Clinic

Serious doubt has been cast on the generalizablility of the lab-based studies of all forms of child treatment included in meta-analyses.

Weisz and Weiss (1989) observed that the meta-analyses to date had involved studies which did not represent normal clinic populations or practices. In particular, subjects were recruited rather than referred, samples were chosen for specific problems (e.g., phobias) on which the therapy was focused, therapists were specially and recently trained in the techniques to be evaluated. It was therefore felt necessary to see whether the good outcomes reported in these studies were also seen in the clinic.

The strategy was to compare 93 referred children who had received at least five sessions of therapy with 60 dropouts from the same clinics who had been offered treatment but did not attend. There were no significant differences on demographic or clinical variables between these groups when referred. The findings were very negative. There were no differences in outcome between treated and untreated children; both improved to an extent equivalent to the improvement found in untreated controls in other studies.

This first study of clinic-based treatment has generated increasing interest in the problem of applying research findings to clinical practice (see, for example, Hoagwood, Hibbs, Brent, & Jensen, 1995; Kendall & Southam-Gerow, 1995; Weisz, Weiss, Morton, Granger, & Han, 1992). Weisz, Donenberg, Han, and Weiss (1995) have reviewed evidence relating to the gulf between efficacy (in lab studies) and effectiveness (in clinic settings) in an attempt to disentangle the reasons behind the difference. The clinic-based studies included in their survey were few (only nine could be found in a thorough search, three of them carried out before 1960) and methodologically very limited, and one important point made by the authors is that well-designed effectiveness studies are desperately needed. They then attempted to test eight plausible reasons for the much larger effect sizes found in lab-based outcome studies, by conducting comparisons between over 100 of these studies, according to the presence or absence of possible mediating factors. They found that two did appear to be relevant (greater severity of clinical problems in clinic samples and greater likelihood of behavioral approach in lab studies). This, then, clearly suggests that the structured approach of behaviorally-oriented treatment tends to lead to greater effectiveness, and while it is clear that better-quality studies of the nonbehavioral therapies are badly needed, it may also be appropriate for psychodynamic clinicians to be considering whether their therapies can be made more focused on explicit goals, although these are likely to be formulated in terms of unconscious, intrapsychic structure and mechanisms.

5.10.3.3 Individual Studies of Outcome of Psychodynamic or Interpersonal Therapies

There have been very few studies of intensive, long-term psychoanalytic psychotherapy for children and adolescents.

A chart review of the outcome of 763 cases in child psychoanalysis has been carried out at the Anna Freud Centre in London (Fonagy & Target, 1994, 1995b; Target & Fonagy, 1994a, 1994b). While this retrospective methodology has severe limitations, the study reached a number of robust conclusions which need to be explored further in controlled, prospective investigations. The main findings were as follows.

(i) Attrition was low compared to reports of other treatment approaches. Overall, 18% of patients withdrew from treatment within six months, and as this made it unlikely that an analytic process would have been established and resolved, these children were excluded from the following analyses (further studies of attrition were, however, carried out).

(ii) Children with pervasive developmental disorders (e.g., autism) or mental retardation did not do well, even with prolonged, intensive treatment. Children with serious disruptive disorders also had relatively poor outcomes.

(iii) Younger children improved significantly more during psychodynamic treatment, and gained additional benefit from four to five times weekly sessions.

(iv) Anxiety disorders, particularly specific rather than pervasive symptoms, were associated with a good prognosis, even if the primary diagnosis was of a different type (e.g., disruptive disorder).

(v) Among children with emotional disorders, severe or pervasive symptomatology responded very well to intensive treatment (four to five sessions per week), but did not show satisfactory rates of improvement in nonintensive psychotherapy.

(vi) Predictors of improvement varied considerably between subgroups of the full sample, and by subdividing the sample according to diagnostic group and developmental level, it was possible to predict a majority of the variance in outcome within the subgroups.

We will now turn to smaller-scale, but prospective studies of psychoanalytic psychotherapy with children, beginning with the earliest systematic studies in the 1960s. After that, we will look at evidence on the efficacy of some forms of intervention based on modifications of psychodynamic psychotherapy.

Heinicke (1965) reported a study of a group of children aged 7–10 with developmental reading disorders linked to emotional symptoms. These children received psychoanalytic psychotherapy, either one or four sessions per week for two years. Greater improvement was found in the group receiving more frequent therapy. Heinicke and Ramsey-Klee (1986) extended this study, hoping to maximize the impacts of different treatment frequencies. They added a third group, matched to the first two, who received therapy once a week for the first year and four times a week for the second. Outcome was measured in terms of the referral problem (reading level) and general academic performance, together with a standardized psychodynamic diagnostic profile. At termination of therapy, there were no significant differences between the three treatment conditions, but at one-year follow-up, the groups which had received either four times a week treatment throughout, or once a week followed by four times a week, showed continued improvements on all measures, greater than those of the once a week group. It therefore seemed that more intensive treatment in the second year of treatment had a beneficial effect.

This study attempted to do three important things: (i) to evaluate the effectiveness of intensive and nonintensive psychodynamic therapy for emotional and learning disturbances; (ii) to isolate the impact of treatment frequency, which is of interest in a variety of therapies, and particularly—for practical and theoretical reasons—in psychodynamic treatment; and (iii) to measure change in both objective, service relevant ways, and ways consistent with the theoretical perspective. In all these respects, the study was innovative. There were difficulties with it, reflecting the methods prevalent in all outcome research 30 years ago when the basic design was planned. Thus, diagnostic characteristics of the sample were poorly described, the projective tests and diagnostic interview were of unknown reliability, and the therapy was not described in full detail, just being stated to be analytically oriented. The study does, however, clearly raise the possibility that for certain children, perhaps those with comorbid emotional and developmental disorders, intensive psychotherapy may offer more profound benefits than the same type of treatment offered once a week.

Moran and Fonagy have carried out a series of studies which overcame the problems of diagnosis and outcome measurement by choosing a group of physically ill children where difficulty in treating the medical condition was thought to be due to severely self-destructive behavior, sabotaging the treatment. The group chosen suffered from so-called brittle diabetes, and the studies assessed the effectiveness of adding psychoanalytic treatment to the existing medical care.

The first study (Moran & Fonagy, 1987) explored the relationship between metabolic control and the content of psychoanalytic sessions, in a single case study of an adolescent girl. Process reports were rated for the presence of psychodynamic themes; the association of certain themes with subsequent improvement, demonstrated through independent measures of diabetic control, was shown using time-series analysis. The second study by these authors (Moran, Fonagy, Kurtz, Bolton, & Brook, 1991) compared two matched groups of 11 diabetic children with highly abnormal blood glucose profiles and histories of regular admissions to the hospital. All patients were offered inpatient treatment, for the treatment group 15 weeks of intensive (three to four sessions per week) psychoanalytic psychotherapy was added. The children in the treatment group showed considerable improvements in diabetic control, maintained at one year follow-up; those in the comparison group had returned to pretreatment levels of metabolic control within three months of discharge from the hospital.

The third study (Fonagy & Moran, 1990) involved of a series of experimental single case investigations. The effect of brief psychoanalytic treatment on growth rate (measured by changes in height and bone age) was examined in three children whose height had fallen below the fifth percentile for age. In each case, treatment was linked to accelerated growth and a substantial increase in predicted adult height.

A different, ongoing study should be mentioned, in spite of methodological drawbacks, because of its attempt to assess outcome in clinically meaningful ways, and for its attention to a particularly needy (and costly) group. Lush, Boston, and Grainger (1991) compared 35 children in psychotherapy who were fostered or adopted with 13 similar children for whom psychotherapy had been recommended but did not start. The children were aged 2–18 years, and they mostly received weekly sessions for at least one year. For ethical reasons, children could not be randomly allocated to treated and untreated groups; the study was naturalistic. A further drawback was that, because no measures suitable for the assessment of psychodynamic change existed, measures were developed specially for the study and used without prior evidence of reliability or validity. However, there are indications that the therapy did benefit these deprived children. Preliminary results have been reported on the first 20 children to be treated. Sixteen cases made "good progress," as judged by therapists ratings and generally confirmed by parents' and external clinicians' opinions. An informal comparison was made with seven similar (but not matched) control children; none of these had improved during the same period.

Smyrnios and Kirkby (1993) have reported a comparison between 12 sessions of focal psychodynamic psychotherapy, and open-ended, nonfocused psychodynamic treatment with children aged 5–9 years. The mean number of child sessions was 10 for the time-limited condition and 27 for the unlimited condition. The parents of children in each treatment group were seen by a different therapist for guidance intended to support the child's treatment within a psychodynamic/family systems orientation. As with the individual therapy for the children, this parent work was much more focused on the identified problems and goals in the brief therapy condition. A further, minimal contact control condition involved assessment and feedback sessions only, with encouragement to the families to work at the goals by themselves. The groups were small, beginning with 10 children per condition, which had fallen to between six and eight by the time of the four year follow-up. The authors report that although pre- to post-treatment gains were significant for each condition, by follow-up, children in the time-unlimited condition no longer showed a significant improvement over the pretreatment levels of symptomatology as rated by the parents. Goals were found to be equally well attained at post-treatment and at follow-up in each of the conditions. Teachers' ratings of social adjustment did not show significant improvement for children in any condition, either at post-treatment or follow-up (there were trends towards improvement for all conditions, and the insignificant results may well reflect the small numbers involved). Essentially, children in the minimal contact control group showed improvements almost as great as those in the two treatment groups, and on one measure the time-limited treatment group did better than the open-ended one. The authors conclude that it may be most cost-effective to offer assessment and feedback rather than treatment to children such as those in this study.

There are several points that need to be made in interpreting the results of this study, some acknowledged by the authors.

The time-limited and minimal contact groups were both treated or assessed by one therapist, while the time-unlimited group were treated by a different therapist. This may well mean that therapist skill or commitment was not equal in the different conditions.

Although the assessments at pre- and post-treatment were carried out by an independent clinician, assessments at the four year follow-up were carried out by one of the authors who had also been the therapist for the time-limited and minimal treatment conditions. This would be likely to introduce a bias in favor of finding greater improvement in those conditions, not only because of therapist expectations, but also because parents interviewed by their child's therapist might be more likely to be positive than those seen by a stranger. Follow-up assessments by the children's teachers did not confirm the differences between groups found by the clinician interviewer.

It is mentioned that families in the minimal contact condition were in fact offered regular therapy sessions (presumably for child and parents separately, as for the other families) after a 12-week wait, although we are not told how many took this up. Thus, although the authors make an attractive argument that assessment only may be the most useful form of intervention, it seems unsafe to regard this group as untreated. In addition, it is stated that the majority of families across conditions sought further treatment during the follow-up period, which obscures the pattern of findings at follow-up.

The small sample size (down to an average of seven per group at follow-up) is unfortunate in that it casts doubt on generalizability and may have prevented certain trends from reaching statistical significance.

Despite these cautions, the study is a useful attempt to examine highly important questions, of whether evaluation alone may be as helpful as treatment, and whether offering more treatment produces better outcome. Although the authors intended to investigate the importance of the number of treatment sessions, it may be that the goal-oriented approach of both the time-limited and the minimal contact groups was helpful, and was the major difference between these groups and the open-ended one. It may also be, as the authors suggest, that the minimal contact condition inadvertently mobilized a useful sense of agency in the children and parents by helping them to clarify their goals and then encouraging the family to work towards them themselves. The open-ended treatment group therefore not only had a different therapist from the other two, but also had a far less focused and enabling approach.

5.10.3.4 Evaluation of Modifications of Psychodynamic Therapy

We will now more briefly survey some studies which have included forms of therapy related to psychodynamic therapies, or which have found some evidence of efficacy when it has been included as a control condition. Again, we will start with the earlier studies.

The pioneering study of Kolvin et al. (1981) will be included for what it can tell us about the amenability of different groups of children to treatment, with its careful design and relatively large sample size, even though it concerned therapy delivered through the school, and therefore difficult to apply directly in clinical psychology practice.

The study compared four forms of school intervention for nearly 600 children, found to be showing either neurotic or disruptive behavioral problems. These children were selected by screening 4300 children. The aim was to assess ways of helping maladjusted children within school, without recourse to psychiatric referral. The children were dichotomized in two major ways: younger vs. older (either 7–8 or 11–12 years old), emotional vs. disruptive disorders. These disorders were not defined by psychiatric diagnosis, but by extreme scores on measures of psychological adjustment in school, personality functioning, and peer acceptance. Children were randomly assigned to one of the four treatment conditions: (i) parent counseling—teacher consultation (three terms): social workers offered casework to parents and consultation to teacher; (ii) psychodynamic group therapy (one term only): the younger group played and were encouraged to reflect on feelings, while the older children discussed emotional issues; (iii) (a) behavior modification, for the older children only (two terms): attempted to alter behavior through systematic reinforcement, working towards specified individual goals; (b) nurture work, for younger children only (five terms): enrichment of social experience in small group with leader selected for warmth and motherliness, some shaping of desirable behavior; (iv) no treatment control condition. There were between 60 and 90 children in each age × treatment group.

Outcome was assessed at 18 months after the beginning of treatment, and at three years. In line with other intervention studies, children with emotional disorders were found to have significantly better outcome on the whole than those with disruptive behavior, irrespective of age and treatment type. It was also found that girls had significantly better outcome across groups. Generally, psychodynamic group therapy had the best outcome, and behavior modification was also helpful for the older children. Both were significantly more effective than counseling parents and teachers, or no treatment. Interestingly and unexpectedly, these differences between treatment and control groups widened over time, so that the effect of intervention appeared to be increasingly evident, even two years after it had ended. As Kolvin et al. (1981) note, the most effective interventions were in fact the briefest, and therefore easily the most cost-effective.

Psychodynamic group therapy was carried out over just one term, but the effects were very clear two and a half years later. This was a surprising finding, and it is not clear to what extent it might have been specific to the school setting; the intervention might not have had the same impact in a different context. School is a very important and familiar social context for children, one which may be able to facilitate and sustain considerable therapeutic impact. Similarly encouraging results have been reported in other smaller scale, more recent studies (e.g., Lochman, Coie, Underwood, & Terry, 1993). It is also unfortunate that the length of treatment was not consistent across the conditions; it is just conceivable that the briefer treatments were more effective because they were briefer, certainly this is a factor which needs to be considered more seriously in research on the impact of psychotherapies for children, as Smyrnios and Kirkby (1993) have indicated.

Another study in which psychodynamic therapy has been used as a comparison treatment has been reported by Szapocznik et al. (1989), who treated disruptive adolescent boys from Hispanic families, using either structural family therapy or individual psychodynamic therapy, compared with a recreational control group. All groups were seen for 12–24 sessions over up to six months. On parent-rated measures of child behavior, child psychodynamic functioning and family functioning, both family therapy and individual psychodynamic treatment led to significant gains, and these gains were not significantly different from each other. However, at one-year follow-up, although child behavior was still improved in both treatment groups, family functioning had deteriorated in the individual therapy group, continued to improve in the family therapy group, and stayed the same in the control group. If this finding proved to be generalizable beyond Hispanic families with conduct-disordered adolescent boys, then it would point to a very important problem in using individual psychodynamic therapy with children or adolescents, that although the child might get better, even sustaining the improvement a year later, the family might function less well over time.

Another study, which like the Kolvin report found a supportive psychotherapy group to be unexpectedly beneficial, was carried out by Fine, Forth, Gilbert, and Haley (1991). They were interested in evaluating the efficacy of social skills training in the treatment of major depression or dysthymia. The patients were 47 adolescents who had been referred for psychiatric treatment, and were assigned to a 12-week therapy group using either supportive therapy or social skills training. Forty-one percent of the patients were receiving concurrent treatment, either psychotherapy, medication, or both. Against expectations, adolescents in the therapeutic support group were significantly less likely than those given social skills training to be diagnosed as depressed following treatment ($p < 0.03$), and had significantly lower scores on the Child Depression Inventory (CDI) self-report measure of depression ($p < 0.05$). These group differences were no longer significant at follow-up, but this appeared to be a result of diminishing sample size rather than a change in relative rates of depression.

As the authors point out, in the absence of an untreated control group, it is not possible to be confident that the two treatments improved on the spontaneous remission rate; 68% of adolescents in the study were no longer diagnosable at nine months follow-up, and in Kovacs, Gatsonis, Paulauskas, and Richards (1989) study the median time to recovery was 9.5 months. This suggests some beneficial impact of therapy, but it is difficult to be sure how comparable the samples in the different studies were. The high proportion of adolescents receiving other treatments in the study by Fine and co-workers also raises a major problem of interpretation. Nevertheless, the authors acknowledge the unexpected superiority of the therapeutic support group over the "active" treatment, and provide an interesting discussion of this. There were indications from process measures that adolescents in the social skills training group were considerably less engaged and more avoidant. The authors suggest that social skills training may be too demanding for clinically depressed adolescents, but that they may become able to use this approach as they become less depressed. They suggest that a fruitful approach might be to offer therapeutic support which introduces social skills training at a later stage.

This is reminiscent of some other studies where a behaviorally-oriented treatment has apparently been perceived by children in clinical samples as less acceptable. For instance, Apter Bernhaut, and Tyano (1984) reported that hospitalized adolescent patients with obsessive compulsive disorder (OCD) refused to cooperate with behavioral treatment, but did accept and improve with supportive psychotherapy. Given that most outcome studies with children and adolescents have used nonreferred, convenience samples, this introduces a note of caution into the common conclusion that cognitive-behavioral treatment is more effective; if it were less likely to be accepted by children with a clinical level of disorder, then there would probably be a higher rate of attrition, and a lower overall improvement rate than the outcome research would suggest. This probably underlines the need for the development of cognitive behavior therapy (CBT) manuals appropriate to referred children rather than to the milder and/or less complex cases recruited for many outcome studies. Some such manuals have been written and successfully tested for clinical samples (Kendall, 1994; March, Mulle, & Herbel, 1994).

An important development in the area of relatively brief, manualized but nonbehavioral therapies for children and adolescents is Interpersonal Psychotherapy for Adolescents (IPT-A: Moreau, Mufson, Weissman, & Klerman, 1991). This is an adaptation of a treatment approach developed for depressed adult patients (IPT: Klerman, Weissman, Ronsaville-Chevron, & Chevrones, 1984), which focuses on a specified range of interpersonal problems which may underlie the individual's affective state. Although those who have developed this treatment approach distinguish it clearly from psychodynamic therapy, in its focus on interpersonal as opposed to intrapsychic difficulties, and its model of depression as rooted primarily in disrupted attachments rather than aggression (see Mufson, Moreau, Weissman, & Klerman, 1993), the contrast they make applies more to classical than to contemporary psychoanalytic thinking, and we include IPT-A here as an approach closely related to certain strong trends in psychodynamic work.

IPT for adults has been tested in several clinical trials, notably the National Institute of Mental Health collaborative study of the treatment of depression (Elkin et al., 1989). This adult therapy was tested for a group of 38 adolescent in-patients diagnosed as suffering from major depression in an open trial (Robbins, Alessi, & Colfer, 1989). Robbins and co-workers reported that 47% of patients showed clinically significant improvement with IPT alone. (Lack of response was associated with dexamethasone nonsuppression and "melancholic subtype.") The nonresponders were then treated with a combination of tricyclic antidepressant and this version of IPT; 92% then responded. Unfortunately, since there appears to have been a 50% placebo response rate (see Mufson et al., 1993), the 47% improvement is not impressive, the approach obviously needs further evaluation, and there is a strong case for introducing modifications to the adult manual, as have the developers of IPT-A (Moreau et al., 1991). The authors of the adapted treatment approach have conducted an open trial using this modified, manualized approach with 14 adolescents (Mufson et al., 1993). By week 3 of the 12 week trial, two of the 14 patients had dropped out and one had been excluded for noncompliance. Of the remaining 11, 9 showed a significant decrease in depressive symptomatology, and 10 no longer met criteria for major depression. One could therefore regard this as a 71% success rate, which is encouraging. A randomized controlled trial is now in progress.

5.10.4 LIMITATIONS OF PSYCHODYNAMIC APPROACHES

A series of limitations render psychodynamic approaches vulnerable to being seen as impractical, unscientific, and uneconomical.

(i) Little operationalization in most descriptions of the treatment, whose value is attested to largely by selected material from individual cases.

(ii) Little attention to what forms of technique may be appropriate for children with a particular diagnosis or psychodynamic formulation.

(a) Very limited evidence of efficacy, especially support from randomized controlled trials.

(b) A prolonged treatment, with rather vague or global goals, raising considerable doubts about cost-effectiveness.

Psychodynamic child psychotherapists are therefore facing great challenges to their convictions and, as managed care restricts payment for such lengthy and poorly supported treatments, to their professional survival. Nevertheless, we hope to have indicated in this chapter that these real and legitimate challenges are now being fully recognized and actively addressed by psychoanalytically-oriented child therapists.

5.10.5 REFERENCES

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the Strange Situation*. Hillsdale, NJ: Erlbaum.

Altman, N. (1992). Relational perspectives on child psychoanalytic psychotherapy. In N. J. Skolnick & S. C. Warshaw (Eds.), *Relational perspectives in psychoanalysis* (pp. 175–194). Hillsdale, NJ: Analytic Press.

Altman, N. (1994). The recognition of relational theory and technique in child treatment. Special Issue: Child analytic work. *Psychoanalytic Psychology, 11,* 383–395.

Alvarez, A. (1993). *Live company*. London: Routledge.

Apter, A., Bernhout, E., & Tyano, S. (1984). Severe obsessive compulsive disorder in adolescence: A report of eight cases. *Journal of Adolescence, 7,* 349–358.

Balint, M. (1937/1965). Early developmental states of the ego, primary object of love. *Primary love and psychoanalytic technique* (pp. 90–108). London: Tavistock.

Balint, M. (1968). *The basic fault*. London: Tavistock.

Barnett, R. J., Docherty, J. P., & Frommelt, G. M. (1991). A review of psychotherapy research since 1963. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 1–14.

Baron, J., Gruen, R., & Asnis, L. (1985). Familial transmission of schizotypal and borderline personality disorders. *American Journal of Psychiatry, 142,* 927–934.

Belsky, J., Rovine, M., & Taylor, D. G. (1984). The Pennsylvania Infant and Family Development Project. III: The origins of individual differences in infant–mother attachment: Maternal and infant contributions. *Child Development, 55,* 718–728.

Bion, W. R. (1959). Attacks on linking. *International Journal of Psycho-Analysis, 40,* 308–315.

Bion, W. R. (1962/1967). A theory of thinking. *Second thoughts* (pp. 110–119). London: Heinemann.

Bion, W. R. (1963). *Elements of psycho-analysis*. London: Heinemann.

Bleiberg, E. (1987). Stages in the treatment of narcissistic children and adolescents. *Bulletin of the Menninger Clinic, 51,* 296–313.

Bleiberg, E. (1994a). Borderline disorders in children and adolescents: The concept, the diagnosis, and the controversies. *Bulletin of the Menninger Clinic, 58,* 169–196.

Bleiberg, E. (1994b). Neurosis and conduct disorders. In J. M. Oldham, & M. B. Riba (Eds.), *American Psychiatric Press Review of Psychiatry* (Vol. 13, pp. 493–518). Washington, DC: Psychiatric Press.

Bleiberg, E., Fonagy, P., & Target, M. (1997). Child psychoanalysis: Critical overview and a proposed reconsideration. *Psychiatric Clinics of North America, 6,* 1–38.

Bower, T. R. (1989). *The rational infant: Learning in infancy*. New York: Freeman.

Bowlby, J. (1969). *Attachment and loss* (Vol. 1). London: Hogarth Press and the Institute of Psycho-Analysis.

Bowlby, J. (1973). *Attachment and loss* (Vol. 2). London: Hogarth Press and Institute of Psycho-Analysis.

Bowlby, J. (1980). *Attachment and loss* (Vol. 3). London: Hogarth Press and Institute of Psycho-Analysis.

Brenner, C. (1982). *The mind in conflict*. New York: International University Press.

Bretherton, I. (1985). Attachment theory: Retrospect and prospect. In I. Bretherton & E. Waters (Eds.), *Growing points of attachment theory and research. Monographs of the Society for Research in Child Development* (Vol. 50, pp. 3–35). Chicago: University of Chicago Press.

Brown, G. W., & Harris, T. O. (1978). *Social origins of depression: A study of psychiatric disorders in women.* London: Tavistock.

Brown, G. W., Harris, T. O., & Bifulco, A. (1986). Long-term effects of early loss of parent. In M. Rutter, C. E. Izard, & P. B. Read (Eds.), *Depression in young people: Developmental and clinical perspectives* (pp. 251–296). New York: Guilford Press.

Bus, A. G., & Van IJzendoorn, M. H. (1992). Patterns of attachment in frequently and infrequently reading mother–child dyads. *Journal of Genetic Psychology, 153,* 395–403.

Cantor, S., & Kestenbaum, C. (1986). Psychotherapy with schizophrenic children. *Journal of the American Academy of Child Psychiatry, 25,* 623–630.

Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin, 98,* 388–400.

Chethik, M. (1989). *Techniques of child therapy: Psychodynamic strategies*. New York: Guilford Press.

Cicchetti, D. (1990a). An historical perspective on the discipline of developmental psychopathology. In J. Rolf, A. Masten, D. Cicchetti, & S. Weintraub (Eds.), *Risk and protective factors in the development of psychopathology* (pp. 2–28). New York: Cambridge University Press.

Cicchetti, D. (1990b). The organization and coherence of socioemotional, cognitive, and representational development: Illustrations through a developmental psychopathology perspective on Down syndrome and child maltreatment. In R. Thompson (Ed.), *Socioemotional development. Nebraska symposium on motivation* (pp. 259–279). Lincoln, NE: University of Nebraska Press.

Cohen, D. J., Towbin, K. E., Mayes, L., & Volkmar, F. (1994). Developmental psychopathology of multiplex developmental disorder. In S. L. Friedman & H. C. Haywood (Eds.), *Developmental follow-up: Concepts, domains and methods* (pp. 155–182). New York: Academic Press.

Crittenden, P. M. (1990). Internal representational models of attachment relationships. *Journal of Infant Mental Health, 11,* 259–277.

Cummings, E. M., & Davies, P. T. (1994). Maternal depression and child development. *Journal of Child Psychology and Psychiatry, 35,* 73–112.

DeFolch, T. E. (1988). Guilt bearable or unbearable: A problem for the child in analysis. Special Issue: Psychoanalysis of children. *International Review of Psycho-Analysis, 15,* 13–24.

Dowling, A. S., & Naegele, J. (1995). Child and adolescent psychoanalysis. In B. E. Moore & B. D. Fine (Eds.), *Psychoanalysis: The major concepts* (pp. 26–44). New Haven, CT: Yale University Press.

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., Glass, D. R., Pilkonis, P. A., Lever, W. R., Docherty, J. P., Fiester, S. J., & Parloff, M.B. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatment. *Archives of General Psychiatry, 46,* 971–982.

Emde, R. N. (1983). Pre-representational self and its affective core. *The Psychoanalytic Study of the Child, 38,* 165–192.

Emde, R. N. (1988a). Development terminable and interminable. I. Innate and motivational factors from infancy. *International Journal of Psycho-Analysis, 69,* 23–42.

Emde, R. N. (1988b). Development terminable and interminable. II. Recent psychoanalytic theory and therapeutic considerations. *International Journal of Psycho-Analysis, 69,* 283–286.

Fairbairn, W. R. D. (1952). *An object-relations theory of the personality*. New York: Basic Books.

Feinman, S. (1991). *Social referencing and the social construction of reality in infancy*. New York: Plenum.

Fine, S., Forth, A., Gilbert, M., & Haley, G. (1991). Group therapy for adolescent depressive disorder: A comparison of social skills and therapeutic support. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 79–85.

Fonagy, P., & Moran, G. S. (1990). Studies on the efficacy of child psychoanalysis. *Journal of Consulting and Clinical Psychology, 58,* 684–695.

Fonagy, P., & Moran, G. S. (1991). Understanding psychic change in child analysis. *International Journal of Psycho-Analysis, 72,* 15–22.

Fonagy, P., Moran, G. S., Edgcumbe, R., Kennedy, H., & Target, M. (1993). The roles of mental representations and mental processes in therapeutic action. *The Psychoanalytic Study of the Child, 48,* 9–48.

Fonagy, P., Steele, H., Moran, G., Steele, M., & Higgitt, A. (1991a). The capacity for understanding mental states: the reflective self in parent and child and its significance for security of attachment. *Infant Mental Health Journal, 13,* 200–217.

Fonagy, P., Steele, H., & Steele, M. (1991b). Maternal representations of attachment during pregnancy predict the organization of infant–mother attachment at one year of age. *Child Development, 62,* 891–905.

Fonagy, P., & Target, M. (1994). The efficacy of psychoanalysis for children with disruptive disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 45–55.

Fonagy, P., & Target, M. (1996a). A contemporary psychoanalytical perspective: Psychodynamic developmental therapy. In E. Hibbs & P. Jensen (Eds.), *Psychosocial treatments for child and adolescent disorders: Empirically based approaches* (pp. 619–638). Washington, DC: American Psychological Association and National Institute of Health.

Fonagy, P., & Target, M. (1996b). Predictors of outcome in child psychoanalysis: a retrospective study of 763 cases at the Anna Freud Centre. *Journal of the American Psychoanalytic Association, 44,* 27–77.

Fonagy, P., Target, M., Steele, M., & Gerber, A. (1995). Psychoanalytic perspectives on developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology* (Vol. 1, pp. 504–554). New York: Wiley.

Fraiberg, S. (1969). Libidinal object constancy and mental representation. *The Psychoanalytic Study of the Child, 24,* 9–47.

Fraiberg, S. (1980). *Clinical studies in infant mental health.* New York: Basic Books.

Freud, A. (1936/1946). *The ego and the mechanisms of defense*. New York: International Universities Press.

Freud, A. (1941–1945/1974). Reports on the Hampstead Nurseries. In *The writings of Anna Freud.* New York: International Universities Press.

Freud, A. (1946). *The psychoanalytic treatment of children.* London: Imago Publishing.

Freud, A. (1965). *Normality and pathology in childhood.* Harmondsworth: Penguin .

Freud, S. (1900). The interpretation of dreams. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vols. 4 and 5, pp. 1–715). London: Hogarth Press.

Freud, S. (1905). Three essays on the theory of sexuality. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 7, pp. 123–230). London: Hogarth Press.

Freud, S. (1909). Analysis of a phobia in a five-year-old boy. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 10, pp. 1–147). London: Hogarth Press.

Freud, S. (1920). Beyond the pleasure principle. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 18, pp. 1–64). London: Hogarth Press.

Freud, S. (1926). The question of lay analysis. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 20, pp. 77–172). London: Hogarth Press.

Freud, S. (1930). Civilization and its discontents. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 21, pp. 57–146). London: Hogarth Press.

Freud, S. (1933). New introductory lectures on psychoanalysis. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 22, pp. 1–182). London: Hogarth Press.

Freud, S., & Breuer, J. (1895). Studies on hysteria. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 2, pp. 1–305). London: Hogarth Press.

Gabbard, G. O. (Ed.). (1994). *Psychodynamic psychiatry in clinical practice: The DSM-IV edition*. Washington, DC: American Psychiatric Press.

Gabbard, G. O. (1995). Psychoanalysis. In H. I. Kaplan & B. J. Sadock (Eds.), *Comprehensive textbook of psychia-*

try (6th ed., pp. 431–478). Baltimore, MD: Williams & Wilkins.

Garmezy, N. (1983). Stressors of childhood. In N. Garmezy & M. Rutter (Eds.), *Stress, coping, and development in children* (pp. 43–84). Baltimore, MD: Johns Hopkins University Press.

Garmezy, N., & Masten, A. (1991). The protective role of competence indicators in children at risk. In E. M. Cummings, A. L. Greene, & K. K. Karraker (Eds.), *Life-span developmental psychology: Perspectives on stress and coping* (pp. 151–174). Hillsdale, NJ: Erlbaum.

Goldstein, J., Freud, A., & Solnit, A. J. (1973). *Beyond the best interests of the child*. New York: Free Press.

Goldstein, J., Freud, A., & Solnit, A. J. (1979). *Before the best interests of the child*. New York: Free Press.

Golinkoff, R. M., Hardig, C. B., Carlson, V., & Sexton, M. E. (1984). The infant's perception of causal events: The distinction between animate and inanimate objects. In L. P. Lipsitt & C. Rovee-Collier (Eds.), *Advances in infancy research*. Norwood, NJ: Ablex.

Green, R. (1985). Atypical psychosexual development. In M. Rutter & L. Hersov (Eds.), *Child and adolescent psychiatry: Modern approaches* (pp. 638–649). Oxford: Blackwell.

Grossmann, K. E., Spangler, G., Suess, G., & Unzer, L. (1985). Maternal sensitivity and newborn orienting responses as related to quality of attachment in Northern Germany. In I. Bretherton & E. Waters (Eds.), *Growing points in attachment theory and research. Monographs of the Society for Research in Child Development* (Vol. 50, pp. 233–256). Chicago: University of Chicago Press.

Group for the Advancement of Psychiatry. (1982). *The process of child therapy (GAP Report 111)*. New York: Brunner/Mazell.

Hartmann, H. (1939/1958). *Ego psychology and the problem of adaptation*. New York: International Universities Press.

Hartmann, H. (1955/1964). Notes on the theory of sublimation. *Essays on ego psychology* (pp. 215–240). New York: International University Press.

Heinicke, C. M. (1965). Frequency of psychotherapeutic session as a factor affecting the child's developmental status. *The Psychoanalytic Study of the Child, 20,* 42–98.

Heinicke, C. M., & Ramsey-Klee, D. M. (1986). Outcome of child psychotherapy as a function of frequency of sessions. *Journal of the American Academy of Child Psychiatry, 25,* 247–253.

Hesse, P., & Cicchetti, D. (1982). Perspectives on an integrated theory of emotional development. *New Directions for Child Development, 16,* 3–48.

Hoagwood, K., Hibbs, E., Brent, D., & Jensen, P. (1995). Introduction to the special section: Efficacy and effectiveness in studies of child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63,* 683–687.

Hornik, R., & Gunnar, M. R. (1988). A descriptive analysis of infant social referencing. *Child Development, 59,* 626–634.

Hug-Helmuth, H. (1920). Child psychology and education. *International Journal of Psycho-Analysis, 1,* 316–323.

Hug-Helmuth, H. (1921). On the technique of child analysis. *International Journal of Psycho-Analysis, 2,* 287–303.

Johnson, E. (1982). Principles and techniques in drama therapy. *International Journal of Arts and Psychotherapy, 9,* 83–90.

Kazdin, A. E., Bass, D., Siegel, T., & Thomas, C. (1990a). Cognitive behavioral treatment and relationship therapy in the treatment of children referred for antisocial behavior. *Journal of Consulting and Clinical Psychology, 58,* 76–85.

Kazdin, A. E., Siegel, T. C., & Bass, D. (1990b). Drawing upon clinical practice to inform research on child and adolescent psychotherapy: A survey of practitioners. *Professional Psychology: Research and Practice, 21,* 189–198.

Kendall, P. C. (1994). Treating anxiety disorders in children: Results of a randomized clinical trial. *Journal of Consulting and Clinical Psychology, 62,* 100–110.

Kendall, P. C., & Southam-Gerow, M. A. (1995). Issues in the transportability of treatment: The case of anxiety disorders in youths. *Journal of Consulting and Clinical Psychology, 63,* 702–708.

Kernberg, O. F. (1975). *Borderline conditions and pathological narcissism*. New York: Aronson.

Kernberg, O. F. (1976). *Object relations theory and clinical psychoanalysis*. New York: Aronson.

Kernberg, O. F. (1982). Self, ego, affects and drives. *Journal of the American Psychoanalytic Association, 30,* 893–917.

Kernberg, O. F. (1987). An ego psychology-object relations theory approach to the transference. *Psychoanalytic Quarterly, 51,* 197–221.

Kernberg, P. F. (1995). Child psychiatry: Individual psychotherapy. In H. I. Kaplan & B. J. Sadock (Eds.), *Comprehensive textbook of psychiatry* (6th ed., pp. 2399–2412). Baltimore, MD: Williams & Wilkins.

King, P., & Steiner, R. (1991). *The Freud–Klein controversies*. London: Routledge.

Klein, M. (1932). *The psycho-analysis of children*. London: Hogarth Press.

Klein, M. (1936/1964). The psychotherapy of the psychoses, In *Contributions to psychoanalysis, 1921–1945*. New York: McGraw-Hill.

Klein, M. (1946). Notes on some schizoid mechanisms. In M. Klein, P. Heimann, S. Isaacs, & J. Riviere (Eds.), *Developments in psychoanalysis* (pp. 292–320). London: Hogarth Press.

Klein, M. (1952/1975). The origins of transference, *The Writings of Melanie Klein* (pp. 48–56). London: Hogarth Press.

Klein, M., Heimann, P., Isaacs, S., & Riviere, J. (Eds.) (1946). *Developments in psychoanalysis*. London: Hogarth Press.

Klerman, G. F., Weissman, M. M., Ronsaville-Chevron, B., & Chevrones, J. (1984). *Interpersonal psychotherapy of depression*. New York: Basic Books.

Kohut, H. (1971). *The analysis of the self*. New York: International Universities Press.

Kohut, H. (1977). *The restoration of the self*. New York: International Universities Press.

Kohut, H. (1984). *How does analysis cure*? Chicago: University of Chicago Press.

Kolvin, I., Garside, R. F., Nicol, A. E., MacMillan, A., Wolstenhome, F., & Leitch, I. M. (1981). *Help starts here: The maladjusted child in the ordinary school*. London: Tavistock.

Kovacs, M., Gatsonis, C., Paulauskas, S. L., & Richards, C. (1989). Depressive disorders in childhood IV. A longitudinal study of comorbidity with and risk for anxiety disorders. *Archives of General Psychiatry, 46,* 776–782.

Links, P. S., Steiner, M., & Huxley, G. (1988). The occurrence of borderline personality disorder in the families of borderline patients. *Journal of the Personality Disorders, 2,* 14–20.

Lochman, J. E., Coie, J. D., Underwood, M. K., & Terry, R. (1993). Effectiveness of a social relations intervention program for aggressive and nonaggressive, rejected children. *Journal of Consulting and Clinical Psychology, 61,* 1053–1058.

Loranger, A., Oldham, J., & Tullis, E. (1982). Familial transmission of DSM-III borderline personality disorder. *Archives of General Psychiatry, 39,* 795–799.

Luborsky, L. (1984). *Principles of psychoanalytic psy-*

chotherapy: A manual for Supportive-Expressive (SE) treatment. New York: Basic Books.

Lush, D., Boston, M., & Grainger, E. (1991). Evaluation of psychoanalytic psychotherapy with children: Therapists' assessments and predictions. *Psychoanalytic Psychotherapy, 5,* 191–234.

MacFarlane, A. C. (1987). Post-traumatic phenomena in a longitudinal study of children following a natural disaster. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 764–769.

Mahler, M. (1968). *On human symbiosis and the vicissitudes of individuation*. New York: International Universities Press.

Main, M., Kaplan, N., & Cassidy, J. (1985). Security in infancy, childhood, and adulthood: A move to the level of representation. In I. Bretherton & E. Waters (Eds.), *Growing points in attachment theory and research. Monographs of the Society for Research in Child Development* (Vol. 50, pp. 66–104). Chicago: University of Chicago Press.

March, J. S., Mulle, K., & Herbel, B. (1994). Behavioral psychotherapy for children and adolescents with obsessive-compulsive disorder: An open trial of a new protocol-driven treatment package. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 333–341.

Marohn, R. C. (1991). Psychotherapy of adolescents with behavioral disorders. In M. Slomowitz (Ed.), *Adolescent psychotherapy* (pp. 145–161). Washington, DC: American Psychiatric Press.

Masterson, J. F. (1972). *Treatment of the borderline adolescent: A developmental approach*. New York: Wiley Interscience.

Moran, G. S., & Fonagy, P. (1987). Psychoanalysis and diabetic control: A single case study. *British Journal of Medical Psychology, 60,* 357–372.

Moran, G. S., Fonagy, P., Kurtz, A., Bolton, A. M., & Brook, C. (1991). A controlled study of the psychoanalytic treatment of brittle diabetes. *Journal of the American Academy of Child Psychiatry, 30,* 241–257.

Moreau, D., Mufson, L., Weissman, M. M., & Klerman, G. L. (1991). Interpersonal psychotherapy for adolescent depression: Description of modification and preliminary application. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 642–651.

Mufson, L., Moreau, D., Weissman, M. M., & Klerman, G. L. (1993). *Interpersonal psychotherapy for depressed adolescents*. New York: Guilford Press.

O'Brien, J. (1992). Children with attention-deficit hyperactivity disorder and their parents. In J. O'Brien, D. Pilowsky, & O. Lewis (Eds.), *Psychotherapies with children and adolescents: Adapting the psychodynamic process* (pp. 109–124). Washington, DC: American Psychiatric Press.

O'Brien, J., Pilowsky, D. J., & Lewis, O. (Eds.) (1992). *Psychotherapies with children and adolescents: Adapting the psychodynamic process*. Washington, DC: American Psychiatric Press.

O'Grady, D., & Metz, J. R. (1987). Resilience in children at high risk for psychological disorder. *Journal of Pediatric Psychology, 12,* 3–23.

O'Shaughnessy, E. (1988). W. R. Bion's theory of thinking and new techniques in child analysis. In E. B. Spillius (Ed.), *Melanie Klein today: Developments in theory and practice*. (Vol. 2, pp. 177–190). London: Routledge.

Petti, T. A., & Vela, R. M. (1990). Borderline disorders of childhood: An overview. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 327–337.

Pine, F. (1985). *Developmental theory and clinical process*. New Haven, CT: Yale University Press.

Rinsley, D. B. (1977). An object relations view of borderline personality. In P. Hartocollis (Ed.), *Borderline personality disorders: The concept, the syndrome, the patient* (pp. 47–70). New York: International Universities Press.

Rinsley, D. B. (1989). Notes on the developmental pathogenesis of narcissistic personality disorder. *Psychiatric Clinics of North America, 12,* 695–707.

Robbins, D. R., Alessi, N. E., & Colfer, M. V. (1989). Treatment of adolescents with major depression: Implications of the DST and the melancholic clinical subtype. *Journal of Affective Disorders, 17,* 99–104.

Rose, S. D. (1972). *Treating children in groups*. London: Jossey-Bass.

Rosenfeld, H. (1971/1988). Contribution to the psychopathology of psychotic states: The importance of projective identification in the ego structure and object relations of the psychotic patient. In E. B. Spillius (Ed.), *Melanie Klein today* (pp. 117–137). London: Routledge.

Rosenfeld, S., & Sprince, M. (1963). An attempt to formulate the meaning of the concept borderline. *The Psychoanalytic Study of the Child, 18,* 603–635.

Rutter, M. (1990). Psychosocial resilience and protective mechanisms. In J. Rolf, A. S. Masten, D. Cicchetti, & S. Weintraub (Eds.), *Risk and protective factors in the development of psychopathology*. New York: Cambridge University Press.

Rutter, M., & Quinton, D. (1984). Long-term follow-up of women institutionalized in childhood: Factors promoting good functioning in adult life. *British Journal of Developmental Psychology, 18,* 225–234.

Sandler, J. (1987). *From safety to the superego: Selected papers of Joseph Sandler*. New York: Guilford Press.

Sandler, J., Kennedy, H., & Tyson, R. (1980). *The technique of child analysis: Discussions with Anna Freud*. London: Hogarth Press.

Sandler, J., & Rosenblatt, B. (1962). The concept of the representational world. *The Psychoanalytic Study of the Child, 17,* 128–145.

Schaefer, C. E., & Cangelosi, D. M. (Eds.) (1993). *Play therapy techniques*. Northway, NJ: Aronson.

Selvini Palazzoli, M., Boscolo, L., Cecchin, G., & Prata, G. (1978). *Paradox and counter-paradox*. New York: Aronson.

Shirk, S. R., & Russell, R. L. (1992). A reevaluation of estimates of child therapy effectiveness. *Journal of the American Academy of Child Adolescent Psychiatry, 31,* 703-709.

Silver, L., & Silver, B. (1983). Clinical practice of child psychiatry: A survey. *Journal of the American Academy of Child Psychiatry, 22,* 573–579.

Simon, M. R. (1992). *The symbolism of style: Art as therapy*. London: Routledge.

Smyrnios, K. X., & Kirkby, R. J. (1993). Long-term comparison of brief versus unlimited psychodynamic treatments with children and their parents. *Journal of Consulting and Clinical Psychology, 61,* 1020–1027.

Sroufe, L. A. (1990). An organizational perspective on the self. In D. Cicchetti & M. Beeghly (Eds.), *The self in transition: Infancy to childhood* (pp. 281–307). Chicago: University of Chicago Press.

Stern, D. N. (1985). *The interpersonal world of the infant: A view from psychoanalysis and developmental psychology*. New York: Basic Books.

Stolorow, R., Brandchaft, B., & Atwood, G. (1987). *Psychoanalytic treatment: An intersubjective approach*. Hillsdale, NJ: Analytic Press.

Sullivan, H. S. (1953). *The interpersonal theory of psychiatry*. New York: Norton.

Summers, F. (1994). *Object relations theories and psychopathology: A comprehensive text*. Hillsdale, NJ: Analytic Press.

Szapocznik, J., Rio, A., Murray, E., Cohen, R., Scopetta, M., Rivas-Vazquez, A., Hervis, O., & Posada, V. (1989). Structural family versus psychodynamic child therapy for problematic Hispanic boys. *Journal of Consulting and Clinical Psychology, 57,* 571–578.

Target, M., & Fonagy, P. (1994a). The efficacy of psychoanalysis for children with emotional disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 361–371.

Target, M., & Fonagy, P. (1994b). The efficacy of psychoanalysis for children: Prediction of outcome in a developmental context. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 1134–1144.

Terr, L. C. (1983). Chowchilla revisited: The effects of psychic trauma four years after a school-bus kidnapping. *American Journal of Psychiatry, 140,* 1543–1550.

Towbin, K. E., Dykens, E. M., Pearson, G. S., & Cohen, D. J. (1993). Conceptualizing "borderline syndrome of childhood" and "childhood schizophrenia" as a developmental disorder. *Journal of the American Academy Child and Adolescent Psychiatry, 32,* 775–782.

Tyson, R. L., & Tyson, P. (1986). The concept of transference in child psychoanalysis. *Journal of the American Academy of Child Psychiatry, 25,* 30–39.

van IJzendoorn, M. H. (1995). Adult attachment representations, parental responsiveness, and infant attachment: A meta-analysis on the predictive validity of the Adult Attachment Interview. *Psychological Bulletin, 117,* 387–403.

Ward, M. J., & Carlson, E. A. (1995). Associations among Adult Attachment representations, maternal sensitivity, and infant–mother attachment in a sample of adolescent mothers. *Child Development,* 66, 69–79.

Warshaw, S. C. (1992). Mutative factors in child psychoanalysis: A comparison of diverse relational perspectives. In N. J. Skolnick & S. C. Warshaw (Eds.), *Relational perspectives in psychoanalysis* (pp. 141–173). Hillsdale, NJ: Analytic Press.

Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63,* 688–701.

Weisz, J. R., & Weiss, B. (1989). Assessing the effects of clinic-based psychotherapy with children and adolescents. *Journal of Consulting and Clinical Psychology, 57,* 741–746.

Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: Meta-analytic findings for clinicians. *Journal of Consulting and Clinical Psychology, 55,* 542–549.

Weisz, J. R., Weiss, B., Morton, Granger, D., & Han, S. (1992). *Meta-analysis of psychotherapy outcome research with children and adolescents.* Los Angeles: University of California.

Werner, E. E. (1989). Children of the garden island. *Scientific American, April,* 106–111.

Winnicott, D. W. (1965). *The maturational process and the facilitating environment.* London: Hogarth Press.

Winnicott, D. W. (1971). *Playing and reality.* London: Tavistock.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.11

Pharmacological Therapies

NIRBHAY N. SINGH and CYNTHIA R. ELLIS
Virginia Commonwealth University, Richmond, VA, USA

5.11.1 INTRODUCTION 268

5.11.2 HISTORY OF PEDIATRIC PSYCHOPHARMACOLOGY 269

5.11.3 PREVALENCE OF PSYCHOPHARMACOLOGY 271

- *5.11.3.1 Developmental Disabilities* 271
- *5.11.3.2 Attention-deficit Hyperactivity Disorder* 271
- *5.11.3.3 Serious Emotional Disturbance* 272
- *5.11.3.4 Other Disorders* 272

5.11.4 GENERAL PRINCIPLES OF PHARMACOLOGY 272

- *5.11.4.1 Pharmacokinetics* 273
- *5.11.4.2 Pharmacodynamics* 273
- *5.11.4.3 Pharmacogenetics* 273

5.11.5 FACTORS AFFECTING THE OUTCOME OF PHARMACOTHERAPY 274

- *5.11.5.1 Behavioral Teratogenesis* 274
- *5.11.5.2 Behavioral Toxicity* 274
- *5.11.5.3 Compliance* 275
- *5.11.5.4 Comorbidity* 275
- *5.11.5.5 Placebo Effects* 275
- *5.11.5.6 Reciprocal Interactions* 276

5.11.6 ASSESSMENT AND CLINICAL FORMULATION 276

- *5.11.6.1 Comprehensive Assessment* 276
- *5.11.6.2 Baseline Assessments* 276
- *5.11.6.3 Disorder-specific Assessment Algorithms* 276
- *5.11.6.4 Baseline Assessments Specifically for Psychopharmacologic Treatment* 277
  - *5.11.6.4.1 Physical examination* 278
  - *5.11.6.4.2 Laboratory tests and diagnostic procedures* 278
  - *5.11.6.4.3 Behavioral observations and rating scales* 278
- *5.11.6.5 Clinical Formulation* 278

5.11.7 PSYCHOTROPIC MEDICATION IN CHILDHOOD DISORDERS 279

- *5.11.7.1 Antipsychotics* 279
  - *5.11.7.1.1 Mechanism of action* 279
  - *5.11.7.1.2 Indications* 281
  - *5.11.7.1.3 Dosing and adverse effects* 283
- *5.11.7.2 Antidepressants* 283
  - *5.11.7.2.1 Mechanisms of action* 283
  - *5.11.7.2.2 Indications* 283
  - *5.11.7.2.3 Dosing and adverse effects* 285
- *5.11.7.3 Antimanics* 285
  - *5.11.7.3.1 Mechanism of action* 285
  - *5.11.7.3.2 Indications* 288
  - *5.11.7.3.3 Dosing and adverse effects* 288
- *5.11.7.4 Anxiolytics* 288
  - *5.11.7.4.1 Mechanism of action* 288
  - *5.11.7.4.2 Indications* 288
  - *5.11.7.4.3 Dosing and adverse effects* 289

5.11.7.5 Stimulants 289
5.11.7.5.1 Mechanism of action 289
5.11.7.5.2 Indications 289
5.11.7.5.3 Dosing and adverse effects 289
5.11.7.6 Antiepileptics 289
5.11.8 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 290
5.11.9 SUMMARY 291
5.11.10 REFERENCES 291

5.11.1 INTRODUCTION

Psychopharmacology as a treatment modality for children and adolescents is a relatively young discipline that is based more on the collective experiences of clinicians than on empirical verification of the effects of various classes of psychoactive medications on childhood disorders. This is one of the reasons why, until the mid-1990s, pharmacological therapy was seen as the treatment of last resort. Indeed, it was rarely considered during initial treatment formulation until all other treatment modalities, such as psychoanalytic and psychosocial interventions, had been exhausted. Child and adolescent psychiatrists followed two simple guidelines when considering the use of medication. First, because the rationale for using psychotropic medication for specific disorders was not well established in children, they were concerned that they may actually do more harm than good in prescribing medication before trying other therapies. Thus, they were adhering to one of the fundamental principles of medicine: *prima non nocere* ("first do no harm"). Of course, in some cases, appropriate use of medication may have been withheld for several weeks or months while alternative, but eventually ineffective, therapies were used. Second, they followed another principle of medicine: when in doubt, don't. In so doing, they may have actually not prescribed medication to some children who would have benefited from it. Even as late as in the 1970s, few child and adolescent psychiatrists prescribed medication for childhood disorders (Towbin, 1995) and, when they did, it was typically for an early variant of what is now called attention-deficit hyperactivity disorder (ADHD).

Few would argue that the situation is vastly different today. Although there is still a paucity of empirically based studies on the use of psychotropic drugs in the treatment of childhood disorders, medication is considered, along with other treatment modalities, an integral part of a treatment plan (Kaplan, Simms, & Busner, 1994; Olfson, Pincus, & Sabshin, 1994). Given our current understanding that medication alone is rarely sufficient in the treatment of a childhood disorder, pharmacotherapy is now considered an important, and often necessary, part of a multimodal treatment plan. Modern child and adolescent psychiatrists who espouse the biopsychosocial model of childhood disorders understand that psychopharmacology, like any other single modality treatment, is limited in scope because it focuses on only one aspect of the child. Given that children develop in an interactional biological, social, and cultural matrix, effective interventions must be multimodal, multifocused, and multidisciplinary in nature (Singh, Parmelee, Sood, & Katz, 1993; Werry, 1993).

Undoubtedly, psychopharmacology is a major component in the treatment of childhood disorders and, as more research is conducted on its use, it is likely to solidify its role as the cornerstone of a comprehensive treatment for childhood disorders. Approximately 12–22% of American children (i.e., up to 14 million children) have a diagnosable mental illness, and it is very likely that many of these children may benefit from psychopharmacologic treatment (Institute of Medicine, 1990). This means that unless all professionals who treat childhood disorders have a working knowledge of psychopharmacology, they will not be able to make informed judgements about a necessary and important component of treatment for many children.

Although we often assume that psychiatric treatment and the prescription of psychotropic medication is solely in the domain of psychiatrists, the fact is that the majority of prescriptions for psychotropic medication written for psychiatric disorders are by family practice and primary care physicians who have limited training in mental health (Beardsley, Gardocki, Larsen, & Hidalgo, 1988). Psychologists assume a major role in the baseline assessment of the child and his family and contribute to the case formulation. In addition, they are often responsible for monitoring the child's response to medication (Ellis, Singh, & Singh, 1997). Because they are the primary therapist for the child's psychosocial treatment, psychologists are in an ideal position to monitor symptomatic improvement, observe the emergence of side effects, and check for medication compliance. Having a good knowledge of psychopharmacology enables the psychologist to play an active collaborative role in recommending medication and dosage changes to the child's physician.

In this chapter, we begin with a brief historical introduction to child psychopharmacology, followed by sections on the prevalence of psychopharmacology of childhood disorders, general principles of pharmacotherapy, factors that affect the outcome of psychopharmacologic treatment, assessment and clinical formulation, and psychopharmacologic treatment of childhood disorders. Finally, we present our view of future directions for research and practice. We use the term psychotropic drug in this chapter to refer to any pharmacological agent that is prescribed for the explicit purpose of bringing about behavioral, cognitive, or emotional changes in the child. The term psychoactive drug is used in a more general sense to refer to any pharmacological agent that may have these effects regardless of the physician's intent when prescribing this type of medication. We use the term child to include children and adolescents, and he or his in a generic sense to include reference to boys and girls.

5.11.2 HISTORY OF PEDIATRIC PSYCHOPHARMACOLOGY

Pediatric psychopharmacology began with Bradley's (1937) serendipitous finding that benzedrine affected the behavior disorders of children. However, little progress was made in the general field of psychopharmacology until the biological revolution in adult psychiatry began in the early 1950s. While conducting animal experiments, the Australian psychiatrist Cade (1949) noted that lithium carbonate made the animals lethargic, prompting him to use the drug with agitated psychiatric patients for its sedative effects. This led to the search for other drugs that could be used with adult psychiatric patients. The explosive growth of psychotherapeutic drugs in the 1950s changed the nature of American psychiatry and psychopharmacology became the mainstay of psychiatric treatment, especially for patients with severe mental illness.

The findings in adult psychopharmacology were extrapolated for use in child and adolescent psychiatry. From the 1950s, and especially during the 1960s, case reports and experimental studies began to appear that reported the effects of antipsychotic, antidepressant, "minor tranquilizer," and stimulant medications for a number of childhood psychiatric disorders. While these case reports and studies had serious methodological flaws, often being devoid of any experimental controls, soft evidence accumulated on the effects of certain drugs for specific disorders. For example, stimulants were generally found to be effective with many children with hyperactivity, imipramine (an antidepressant) was found to be effective for treating enuresis, and several antipsychotic medications were found to be effective for psychosis and aggressive behaviors.

The 1970s heralded an increasing interest in the scientific evaluation of existing drugs as well as in the development of related compounds in each drug class. Several advances were made in this period, including: (i) the development of diagnostic criteria for psychiatric disorders (e.g., research diagnostic criteria, *DSM*) so that the same disorders could be targeted across drug studies; (ii) increasingly sophisticated methodologies for drug evaluation studies; (iii) development of reliable and valid instruments sensitive to drug effects; (iv) increasing knowledge of pharmacokinetics of psychiatric drugs used with both children and adults; and (v) increasing monitoring of intended and adverse effects of drugs, both in the short- and long-term. Further, children were found to be drug responsive to a number of conditions that had previously not been the focus of pharmacotherapy, including eating disorders, aggressive conduct disorders, separation anxiety disorders (school phobia), bipolar disorders, depressive disorders, and Tourette's syndrome.

New drugs were added during this period of explosive growth in the use of pharmacotherapy. For example, anticonvulsant medications (e.g., carbamazepine, valproate) were found to be effective not only for seizure disorders but also for treating some children with bipolar disorder. The selective serotonin reuptake inhibitors (e.g., fluoxetine, sertraline, paroxetine) and the serotonergic-specific tricyclic drug, clomipramine, were found to be useful for a number of conditions, particularly depression and anxiety disorders, including obsessive-compulsive disorder. Further, dopamine receptor antagonists (e.g., clozapine, risperidone, olanzapine) that have far fewer neurological adverse effects than the traditional neuroleptics (e.g., chlorpromazine, thioridazine), and two new antidepressants (i.e., venlafaxine, nefazodone) were added to our pharmacotherapeutic armamentarium. Indeed, the search for new drugs that are more effective and have a better adverse effects profile than those currently available is one of the most vigorous and exciting areas of research in child psychiatry today.

Two other developments that occurred during this period should be noted. First, there was an increasing sophistication in the research methodology used to assess drug effects.

According to Werry (1993), because child psychiatry was preoccupied with psychoanalytic theory until the late 1970s, there was a lack of research in pediatric psychopharmacology. To a large extent, this void was filled by experimental and clinical psychologists who had better training than child psychiatrists in research methodology. They brought to the field of pediatric psychopharmacology experimental designs that could be used for clinical trials, and they were able to undertake complex statistical analyses of the data from experimental studies. Further, and in contrast to adult psychiatry, they emphasized the use of reliable and valid measures, including laboratory measures and rating scales, and the need for assessing the effects of medication on behavior, learning, and academic performance. However, one of the consequences of psychologists being more active than psychiatrists in psychopharmacological research was that they emphasized psychosocial rather than biological variables.

Second, beginning in the late 1980s, there was much negative media attention given to the use of Ritalin for the treatment of hyperactivity and attention deficit disorder (ADD: Brotman, 1992) and this negative attention is evident even in the late 1990s. One of the problems noted by the media was that this drug was prescribed excessively and probably inappropriately for many children. For example, Safer and Krager (1992) found that in Baltimore County, MD, the use of Ritalin increased fivefold from 1981 through 1987—far faster than one would have predicted from the figures on the prevalence of ADD in children. The media and litigation blitz which lasted about two years had a major effect on the prescribing practices of physicians in this particular county. Although the prescriptions for Ritalin decreased by 40% during this period, Safer and Krager reported that 36% of the children who discontinued Ritalin experienced major academic problems (i.e., failing grades, suspension from school) and another 47% had mild to moderate academic problems (e.g., lower grades, in-school detention). Further, as the number of prescriptions for Ritalin decreased during this period, there was concurrent fourfold increase in the number of prescriptions for tricyclic antidepressants among children with ADD. As the tricyclics have more troublesome side effects than Ritalin, the outcome for these children was not positive at all.

More recently, consumer groups and the Church of Scientology have attacked the use of fluoxetine (Prozac) on the basis of a report by Teicher, Glod, and Cole (1990) which reported the (re)emergence of severe suicidal preoccupation in six patients with severe depressive disorders following treatment with Prozac. However, none of the six patients actually attempted suicide while being treated with this drug. Further, studies have shown that Prozac is no more likely to induce suicidal feelings than any other antidepressant (Fava & Rosenbaum, 1991) and, at least in one study, it was found that the incidence of suicidal ideations was greater in patients treated with imipramine (a tricyclic antidepressant) or placebo than with Prozac (Beasley & Dornseif, 1991). Because research shows that the risk of Prozac-induced suicide is extremely low and that the suicide rate in untreated major depression is as high as 15%, the Federal Drug Administration (FDA) did not support the contentions of consumer groups or the Church of Scientology and refused to withdraw Prozac from the market (Burton, 1991). Although this controversy is related to adults, fluoxetine is also used with adolescents in the treatment of major depressive disorders, obsessive-compulsive disorder, and anxiety disorders (Popper, 1995). The effects of such controversies will affect the use of this drug with adolescents in the future.

Finally, the contribution of psychopharmacological research in developmental disabilities to the general field of child and adolescent psychopharmacology needs to be mentioned. From the mid-1960s until the 1980s, most child psychiatrists had an overwhelming interest in hyperactivity and ADHD and focused their research on the effects of stimulants on this disorder. Their interest in other disorders and other classes of drugs was minimal during this period. However, spurred by the methodological review of Sprague and Werry (1971), psychologists began testing the effectiveness of neuroleptics and other classes of drugs that were the mainstay of psychopharmacological treatment in developmental disabilities, especially for treating aggressive, self-injurious, and stereotypic behaviors (Aman & Singh, 1988; Singh, Ellis, & Singh, 1994; Singh, Guernsey, & Ellis, 1992; Singh & Winton, 1989). Much of the research on the psychopharmacology of the developmental disabilities since the mid-1970s has been methodologically innovative and experimentally rigorous. What we have learned from this endeavor has contributed much to the development of high-quality research in the psychopharmacology of child and adolescent disorders generally.

Since the 1980s, there has been an exponential growth in the number of studies published on the effects of medication in childhood disorders. The field of psychopharmacology of childhood disorders is rapidly achieving respectability and scientific credibility, and the launching of a specialist journal in 1990 (*Journal of Child and*

Adolescent Psychopharmacology) is an indication of the coming of age of this field.

5.11.3 PREVALENCE OF PSYCHOPHARMACOLOGY

Prevalence studies of pharmacotherapy for childhood disorders tell us something about the nature and patterns of drug usage in the treatment of psychiatric disorders in this population. Prevalence figures can also be useful in determining whether medication is being overused for certain disorders, or whether given drugs or classes of drugs are being overused for given symptoms or disorders. For example, the continuing debate about the excessive use of Ritalin is based in part on the number of prescriptions written each year for children with ADHD (Safer & Krager, 1992). Further, prevalence data also provide a useful guide to the way specific drugs are prescribed, evaluated, and monitored in clinical practice (Kaplan et al., 1994). As we have observed in the field of developmental disabilities (Ellis et al., 1997), prevalence studies are useful for examining the nature of the variables that are associated with drug use. These include, but are not limited to, characteristics of the child, physicians prescribing the medication (e.g., child psychiatrists, primary care physicians), the cultural and social milieu of the child, and psychosocial variables associated with drug use. Thus, prevalence studies provide us with a global picture of the patterns of psychopharmacological treatment for childhood disorders in general, and specific disorders in particular.

5.11.3.1 Developmental Disabilities

The best prevalence studies of pharmacotherapy have been published in the field of developmental disabilities (see Singh, Ellis, & Wechsler, in press). Following the seminal work of Lipman (1970) on the use of psychotropic drugs in residential facilities, a large number of prevalence studies have investigated the use of psychoactive medication in residential and community facilities for individuals with developmental disabilities. Overall, institutional surveys show that psychotropic drug usage ranges from a low of 19% to a high of 86%, with the majority of studies reporting prevalence rates between 30% and 50%. The use of seizure medication in institutions ranges from a low of 24% to a high of 56%, although typically it is between 25% and 45%. Together, psychotropic and antiepileptic medication usage in institutions ranges from 50% to 70%, with some individuals being prescribed both types of drugs concurrently.

Overall, there are fewer community surveys and these show that psychotropic drug usage with children ranges from 2% to 7%, and for adults it ranges from 14% to 36%. The use of seizure medication in community settings ranges from 12% to 31% for children and from 18% to 24% for adults. The combined prevalence of psychotropic and antiepileptic medications ranges from 19% to 33% for children and from 36% to 48% for adults.

Singh et al. (in press) have noted that various factors influence psychotropic drug use in individuals with developmental disabilities. For example, increasing dosages have been found to correlate highly with increasing age and decreasing intellectual impairment. There is a strong relationship between the type of residential facility and the use of medication to control behavior problems, with more medication being used in larger facilities and in facilities with very restrictive environments. Further, there is a strong positive correlation between the number and severity of an individual's behavior and psychiatric problems and the use of medication (Aman & Singh, 1988). For example, those who exhibit aggression, hyperactivity, self-injury, screaming, or anxiety are more likely to receive drug treatment than those who exhibit milder and less disruptive problems, such as noncompliance.

5.11.3.2 Attention-deficit Hyperactivity Disorder

Since the late 1960s, there has been increasing interest in the use of psychotropic drugs, especially psychostimulants, in the treatment of attention-deficit hyperactivity disorder (ADHD). Most of the early estimates came from the media rather than from scientific prevalence studies. For example, it was reported in the media that between 5% and 10% of students in Omaha public schools in 1970 were receiving stimulant medication for hyperactivity (Maynard, 1970), although later it was found that these figures actually were for children with learning disabilities. In any case, media estimates of the use of stimulants have continued to be published because of the concern that stimulant medication is prescribed for children who may not actually have ADHD and is being used in lieu of effective academic instruction (e.g., Gates, 1987; Hancock, 1996; Johnson, 1988; Landry, 1996).

In one of the earliest prevalence studies, Stephen, Sprague, and Werry (1973) estimated that 2–4% of school children in the Chicago area were receiving medication for hyperactivity in the 1970–1971 school year. Gadow's (1981) review of the use of drug treatment for hyperactivity suggested that the prevalence rate was between 1% and 2%, with considerable variability across geographic regions. Safer and Krager's (1988) longitudinal community surveys in Baltimore County, MD, showed that the prevalence of drug prescriptions doubled every 4–7 years, with about 6% of children in public elementary schools in this county receiving medication for hyperactivity by 1987. Further, almost all the prescriptions (99%) were for stimulants (93% methylphenidate, 4% pemoline, and 3% dextroamphetamine), and the ratio of boys to girls receiving the medication was 5:1. Sherman and Hertzig (1991) reviewed the state prescription records in 1986 and found that 0.4% of all 3–17 year olds in Suffolk County (Long Island, NY) received at least one prescription for methylphenidate for hyperactivity. The medication was typically prescribed by local pediatricians, and the ratio of boys to girls receiving methylphenidate was 6:1.

5.11.3.3 Serious Emotional Disturbance

Children with serious emotional disturbance (SED) typically have comorbid psychiatric disorders for which psychotropic medication is often indicated. They are prescribed more psychotropic medication than their peers without SED, regardless of whether the mental health services are provided in a psychiatric hospital (Singh, Landrum, Donatelli, Hampton, & Ellis, 1994) or in the community (Landrum, Singh, Nemil, Ellis, & Best, 1995). Further, the prevalence of psychotropic medication used with children with SED receiving inpatient psychiatric services or community-based services is much higher when compared to the 11% usage in school-identified students with SED (Cullinan, Epstein, & Sabornie, 1992; Cullinan, Gadow, & Epstein, 1987). For example, Silver et al. (1992) reported 43%, Epstein, Cullinan, Quinn, and Cumbald (1994) and Epstein, Cullinan, Quinn, and Cumbald (1995) reported 27% and 30%, respectively, and Barber, Rosenblatt, Harris, and Attkisson (1992) reported a prevalence of 34% for just antipsychotic medication. Singh et al. (1994) reported that 64% of the children with SED at a psychiatric hospital received at least one psychotropic medication.

5.11.3.4 Other Disorders

There is a paucity of prevalence data on the use of medication for children with other psychiatric disorders. The best data we have on individuals with autism comes from a large study by Aman, Van Bourgondien, Wolford, and Sarphare (1995). The mean age of their sample of 859 individuals with autism was about 16 years. Of this sample, about 34% were taking some psychotropic medication or vitamin for autism or associated behavioral/psychiatric problems. Although about 19% reported having epilepsy, only 13% were actually taking anticonvulsant drugs. In total, over 50% of the sample were on psychoactive medication. There are no recent studies on the prevalence of psychopharmacological studies of children with depression. In two studies using selected samples of children being treated for depression, Kovacs, Feinberg, Crouse-Novak, Paulauskas, and Finkelstein (1984) reported that of the 65 children in their study, only two were on antidepressants, and another was treated with an antianxiety agent. In the other study, Keller, Lavori, Beardslee, Wunder (1991) reported that only one of 38 adolescents diagnosed as having an episode of major depression was treated with medication, in this case an antianxiety agent. As for enuresis, Foxman, Valdez, and Brook (1986) reported that 2% of a sample of 1724 5–13 year olds received medication (typically imipramine) for bedwetting. Although no drug prevalence data on children with Tourette's syndrome are available, one study in which a majority of the individuals with Tourette's syndrome were under 20 years of age reported that 71% of 1034 individuals had taken medication for Tourette's syndrome at some time, and 43% were taking medication at the time of the survey (Stefl, Bornstein, & Hammond, 1988). There are no reliable estimates of the use of medication in other childhood disorders.

These data provide a general view of the use of psychotropic and antiepileptic medication in a number of childhood disorders. However, many of the prevalence studies were poorly controlled, using ill-defined variables, small and selected samples, and so on, thus limiting the usefulness of the data that were reported. It is likely that the current use of medication in all childhood disorders is actually increasing because we have better-controlled studies than in the past indicating the effectiveness of newer medications, such as the selective serotonin reuptake inhibitors (SSRIs), with a number of these disorders. Obviously, there is a need for better controlled studies that examine the prevalence and patterns of drug use in children with various childhood disorders.

5.11.4 GENERAL PRINCIPLES OF PHARMACOLOGY

Psychopharmacology is a rapidly evolving area in child and adolescent psychiatry and it is important for clinicians working with children with psychiatric disorders to remain current with the research literature in this field. While we do not think that it is critical for nonmedical clinicians to be experts in the pharmacology of psychiatric drugs, it may be helpful for them to have a working knowledge of basic concepts in the field so that their collaboration with physicians can be more meaningful. For example, it may help them to better understand the theoretical bases of the choices made by a psychiatrist in the selection of specific agents for treating single or multiple childhood disorders. We will briefly mention three areas that are important for clinical psychologists as well as other nonmedical clinicians to be familiar with. However, we emphasize that this is a very complex, technical, and specialized field that cannot be adequately covered here, and we suggest that interested readers begin with general overviews (e.g., Clein & Riddle, 1995; Paxton & Dragunow, 1993) before delving deeper into the field.

5.11.4.1 Pharmacokinetics

Pharmacokinetics describes the time course and effects of drugs and their metabolites on the body. Thus, it describes what the body does to a drug. The most important pharmacokinetic factors that affect the time course and the effects of drugs on patients include: (i) absorption, which is the process that determines how a drug travels from the site of administration to the site of measurement (e.g., plasma, whole blood); (ii) first-pass effect, which is the hepatic extraction of orally administered drugs before they reach the systemic circulation; (iii) distribution, which indicates how much of a drug is distributed to the various organs or sites of action throughout the body; (iv) steady-state concentration, which indicates the concentration of the drug when the amount administered is equal to the amount eliminated per unit time; (v) half-life, which is the time required for the concentration of the drug in plasma or whole blood to fall by one-half; (vi) elimination rate constant, which is the proportion of the drug in the body that is eliminated per unit time; (vii) clearance, which provides a measure of elimination of the drug from the body and is calculated by multiplying the amount of drug in the body by the elimination rate constant; (viii) first-order kinetics, which occurs when the amount of the drug eliminated per unit time is directly proportional to its plasma concentration; and (ix) zero-order kinetics, which occurs only when a fixed amount of the drug is eliminated per unit time regardless of plasma concentration. Data on these factors can be used to draw a unique kinetic profile of a given drug.

Together with pharmacodynamics, pharmacokinetics provides the bases for understanding the relationship of drug concentrations to clinical effects in children. Since the 1950s pharmacologists have developed the theoretical pharmacokinetic bases for most types of dosage regimens and have identified the therapeutic plasma concentration ranges for a number of psychotropic drugs that are currently used to derive rational dosage regimens (Evans, Schentag, & Jusko, 1986). Indeed, child psychiatrists rely on pharmacokinetic principles to optimize the effects of pharmacotherapy in children (Clein & Riddle, 1995; Greenblatt & Shader, 1987).

5.11.4.2 Pharmacodynamics

Pharmacodynamics deals with the relationship between drug dosage or concentration in the body and its drug effects, both desirable and undesirable. Thus, it deals with the mechanism(s) of drug action and generally describes what a drug does to the body. The most important pharmacodynamic considerations include: (i) receptor mechanism, which describes how the drug binds at the cellular level and initiates its pharmacodynamic effects; (ii) dose–response curve, which provides a plot of the drug concentration against the effects of the drug and allows comparison of the efficacy and potency of drugs; (iii) therapeutic index, which provides a relative measure of the toxicity and safety of a drug, and is calculated by dividing the median toxic dose by the median effective dose; (iv) lag time, which is the time taken for the full therapeutic effects of a given drug to appear, and the reasons for a delay in effects may be pharmacokinetic, pharmacodynamic, or both; and (v) tolerance, which refers to the responsiveness of a child to a particular drug as it is administered over time.

Children differ widely in terms of the drug dose that produces a given effect. Thus, clinicians must have a good knowledge of current pharmacodynamic principles to understand a child's response to psychotropic drugs (Dingemanse, Danhof, & Breimer, 1988; Paxton & Dragunow, 1993).

5.11.4.3 Pharmacogenetics

Pharmacogenetics deals with idiosyncratic or unusual drug responses that have a hereditary basis (Kalow, 1990). Responses to psychotropic and other drugs in children are modulated by their genetic predisposition. This is because "genes encoding enzymes or proteins that play a role in the drug response differ in some respect from one individual to the next" (Nebert & Weber, 1990, p. 469). When all other variables are held constant, the individual's pharmacogenetic response reflects a genetic difference in the metabolic rate when compared to that of a control subject. For example, Zhou, Koshakji, Silberstein, Wilkinson, and Wood (1989) compared the physiologic effects and the pharmacologic disposition of propranolol in a group of Caucasian Americans and a group of Chinese Orientals. They reported that the Chinese subjects were more responsive to the drug, displaying a larger reduction in the heart rate and blood pressure, because they metabolized the drug more efficiently than the subjects in the Caucasian group. According to Nebert and Weber (1990, p. 473), idiosyncratic drug responses that may be due to genetic variation in any of the subcellular steps involved in pharmacokinetics include the following mechanisms: (i) transport (absorption, plasma protein binding); (ii) transducer mechanism (receptors, enzyme induction, or inhibition); (iii) biotransformation; and (iv) excretory mechanisms (renal and biliary transport). Typically, clinical observations, family or twin studies, protein polymorphisms, animal modeling, and DNA polymorphism characterizations are methods used to discover new atypical drug responses that may have a pharmacogenetic basis. The clinical import of pharmacogenetics is that clinicians should be aware of the possibility of differences in drug response and dose requirements among children from various ethnic and racial groups.

5.11.5 FACTORS AFFECTING THE OUTCOME OF PHARMACOTHERAPY

The effects of psychotropic drugs are seldom determined solely by their pharmacological properties, such as those discussed above. Indeed, psychotropic medication is prescribed and used within a transactional system between the child, his family, physician and other clinicians, as well as others (e.g., teachers, peers) who are significant in the life of the child (Moerman, 1979; Singh & Aman, 1990). Of course, we must remember that children rarely, if ever, seek assistance from physicians directly; their parents and other caretakers do and, almost always, without asking for input from the child. The child's views of the treatment process, the child–physician therapeutic alliance, and the child's attributions of the effects of medication are also important and may be responsible for at least some of the effects of pharmacotherapy (Towbin, 1995). Further, other factors, such as compliance, placebo effects, and sociocultural and ecobehavioral factors, are important in treatment outcome. However, the majority of studies in child psychopharmacology have focused on biological variables (e.g., pharmacokinetics, pharmacodynamics, and pharmacogenetics) to the exclusion of virtually all others.

Singh and Aman (1990) noted that our knowledge of the intended effects of psychotropic medication in children is often based on an assessment of the changes in the target behavior, symptoms, and syndrome as a result of a controlled trial of a given drug. They criticized this approach on the grounds that it does not include the study of environment × behavior × drug interactions and inter-relationships. The majority of current psychopharmacological research in general focuses on the consequences of specific drugs or drug combinations on the target behavior or diagnosis, without regard to the setting events or contextual variables that may influence the effects of the medication. A transactional analysis would require that the effects of pharmacotherapy be explained in terms of the relevant contextual variables that transact to determine treatment outcome.

There are a large number of contextual variables that may contribute to the outcome of psychopharmacotherapy. While all variables cannot be covered here, we present diverse examples to show that assessment of treatment outcome is much more complex than has hitherto been noted in psychopharmacological research.

5.11.5.1 Behavioral Teratogenesis

Girls of child-bearing age are always cautioned about the potential interaction between pregnancy and nursing, the course of a psychiatric disorder, and psychotropic treatment (Coyle, Wayner, & Singer, 1976; Schou, 1990). With some drugs (e.g., lithium, valproic acid), there is a real possibility of a drug × pregnancy interaction which increases the risk of malformations in the unborn child and the potential for later developmental anomalies. For example lithium may increase the risk of fetal macrosomia, premature delivery, and perinatal mortality (Schou, 1990), and the teratogenic effects of valproic acid on exposed fetuses includes a 1–2% incidence of spina bifida (Centers for Disease Control, 1983)

5.11.5.2 Behavioral Toxicity

Behavioral toxicity refers to the undesirable effects of essentially therapeutic levels of medication clinically indicated for a given disorder (DiMascio, Soltys, & Shader, 1970). These undesirable effects may include anticholinergic effects, alpha-adrenergic blockade, and dopaminergic effects, among others. In behavioral terms, behavioral toxicity can be classified as a concurrent stimulus event that affects the ongoing interactions of the child. For example, take the case of a child who is on medication for a seizure disorder. The drowsiness produced by his anticonvulsant medication (e.g., phenobarbital) may cause him to function more poorly than he might in its absence. Thus, in this case, behavioral toxicity may operate concurrently with classroom instruction or activities at home. In assessing the effects of psychopharmacotherapy, clinicians should include measures of behavioral toxicity because it will assist the child's treatment team to arrive at risk-to-benefit treatment decisions on the basis of data.

5.11.5.3 Compliance

Compliance is the degree to which a child adheres to the recommended treatment plan of the treating physician. In children, compliance with medication is typically partial because they tend to take medication irregularly or less often than is prescribed. The reasons for medication noncompliance can be categorized as "rational, capricious, absolute refusal, confusion, or iatrogenic" (Janicak, Davis, Preskorn, & Ayd, 1993). Often, noncompliance occurs because of unpleasant side effects of the medication and the perceived stigma associated with taking medication in school. Some children believe that taking medication somehow lessens their control of their lives, or that the medication determines what they do. Medication compliance in children can be increased by: (i) having the child actively participate in treatment planning; (ii) explaining to the child before treatment begins the expected side effects of the medication; (iii) emphasizing the positive aspects of taking the medication as prescribed, and (iv) having the simplest drug regimen (e.g., taking medication once a day for long-acting drugs). Further, compliance can be increased by involving the family in the treatment process. Clearly, these psychosocial factors associated with medication compliance affect the outcome of psychopharmacologic treatment.

5.11.5.4 Comorbidity

Comorbidity, which is the presence of disorders other than the one targeted for treatment, affects treatment outcome because the additional disorders change the simple drug × disorder × environment interaction to a more complex drug × disorder_1 × environment × disorder_2 interaction, if there is just one comorbid disorder. This interaction may mean that a second treatment, which could be pharmacotherapy or an alternative, is necessary to control a comorbid disorder, resulting in drug_1 × disorder_1 × environment × drug_2 × disorder_2 interaction. Of course, other interactions may also occur, such as drug_1 × drug_2 interaction, and changes in the physiologically determined susceptibilities of the drugs to interact with the environment to produce certain unintended, positive or adverse effects (i.e., side effects). This scenario suggests that the child must be adequately assessed not only for the target symptoms, syndrome, or disorder but also for comorbid disorders. Further, the interactive effects of treatments × environments × disorders must be assessed during clinical treatment and under baseline, placebo, and treatment in research protocols. This is a critical consideration for many childhood disorders. For example, comorbid axis I psychiatric disorders are commonly associated with childhood depression and ADHD (Kye & Ryan, 1995). The clinical import of this variable is that comorbidity may contribute to differential response to psychopharmacologic treatment.

5.11.5.5 Placebo Effects

The placebo effect in pharmacology is the phenomenon that is observed when a child exhibits a clinically significant response to a physiologically inert substance. This is not a rare phenomenon as it may account for 30–70% of the therapeutic responses observed in psychopharmacologic treatment (White, Tursky, & Schwartz, 1995). The placebo response is not due to the psychopharmacological properties of the drug; that is, the outcome is not due to the effects of the prescribed intervention. It results from purely psychological effects as a consequence of expectancy of an outcome from the treatment. While a majority of the placebo effects are established simply by asking the child if he is feeling better or by having the child complete a rating scale that provides a measure of his subjective well-being, often placebo effects can also be measured objectively. For example, individuals given placebo treatment for anxiety disorder often show less autonomic nervous system arousal (e.g., decrease in heart rate, lower blood pressure). Although there is some evidence of sustained placebo response in adults, children relapse fairly rapidly, suggesting that continued monitoring of a placebo response is absolutely necessary in this population.

5.11.5.6 Reciprocal Interactions

In behavioral terms, reciprocal interactions are concurrent stimulus–response contexts that set the occasion for further interactions. This is an important variable that affects the course of psychopharmacotherapy because the initial effects of drugs on a child may also have an effect on the reactions of others who interact with him. This phenomenon was clearly illustrated in the stimulant treatment of children with ADHD in which (i) the drug increased a child's compliance with parent demands, and (ii) the parents, in turn, reduced their commands and disciplinary behaviors, while increasing their positive interactions with the child (Barkley & Cunningham, 1979; Cunningham & Barkley, 1978). This highlights the importance of assessing drug effects on reciprocal interactions of the child with significant others in his environment (e.g., parents, siblings, peers, teachers).

Although we have selectively discussed some factors that affect the outcome of psychopharmacologic treatment, there are many factors that should be monitored for their impact on the outcome of pharmacologic intervention.

5.11.6 ASSESSMENT AND CLINICAL FORMULATION

5.11.6.1 Comprehensive Assessment

A comprehensive psychiatric assessment is necessary before a child is given pharmacotherapy. The clinician's goal is to understand the child's presenting problems or symptoms within the context of the highest level of diagnostic sophistication that can be reached on the basis of a comprehensive multidisciplinary assessment. There are four levels of diagnostic sophistication: (i) symptomatic, which includes isolated symptoms (e.g., auditory hallucinations) that provide an indication of a possible diagnosis (e.g., psychosis, not otherwise specified); (ii) syndromic, which includes a constellation of signs and symptoms that have been present for a given time, and standardized inclusionary and exclusionary criteria can be used to derive a diagnosis (e.g., depression); (iii) pathophysiologic, which includes structural or biochemical changes (e.g., elevated thyroid function tests, lowered thyroid-stimulating hormone) that indicate the diagnosis (e.g., hyperthyroidism); and (iv) etiologic in which the diagnosis is based on known causative factors (Janicak et al., 1993). With children, most of the psychiatric diagnoses are at the syndromic level of sophistication because we do not currently have a good understanding of the etiologies of many childhood disorders. Thus, it is not uncommon to find wide variability in treatment outcome in children diagnosed with the same syndrome because they have similar presentations but with substantially different underlying mechanisms. This means that sometimes we treat children's behavioral symptoms or psychiatric disorders without fully appreciating their biological and genetic underpinnings or how these factors transact with the children's physical and cultural environment.

5.11.6.2 Baseline Assessments

A baseline assessment provides the basis for determining the child's psychopathological condition, indications for treatment, and the nature of the proposed treatment. It is multidimensional and multidisciplinary, incorporating assessments of the child's symptoms and functioning across multiple domains, as well as an evaluation of his family history and his physical and cultural environment. Typically, a baseline assessment includes: (i) source and the reason for referral, including the target symptoms that may be the focus of treatment; (ii) history of the presenting symptoms; (iii) psychiatric history and current mental status; (iv) developmental and medical history; (v) family and education/school history; and (vi) cultural context of the family (e.g., determine if there are any specific religious or cultural beliefs of the family that may interact with the child's psychiatric treatment in general, and psychopharmacological treatment in particular). Different mental health professionals may take the lead role in obtaining the baseline assessment data. Because of their broad-based training, child psychiatrists and clinical psychologists often assume a major role in these assessments, obtaining information on the child through structured interviews, behavioral observations, and standardized rating scales.

5.11.6.3 Disorder-specific Assessment Algorithms

Typically, clinicians develop a working hypothesis regarding the child's diagnosis during the initial assessments, or the child may be referred by another clinician for treatment of a specific disorder, such as ADHD. In such cases, clinicians may use an assessment algorithm designed to confirm the diagnosis and to provide data that can be used as the baseline for assessing treatment effects. For example, Table 1 presents the standard algorithm for screening children with ADHD at our clinic (see Garfinkel & Amrami, 1992). Assessment algorithms provide a distillation of current knowledge from multiple sources and disciplines regarding the assessment of a specific disorder. Further, they provide a useful framework for interdisciplinary assessments.

5.11.6.4 Baseline Assessments Specifically for Psychopharmacologic Treatment

The initial assessments can be used to determine the psychiatric diagnosis of the child, as well as to identify target symptoms best treated with pharmacotherapy (Sood & Singh, 1996). If a child is referred specifically for psychopharmacological treatment, he will need a baseline psychiatric assessment prior to initiation of medication. Premedication workup typically includes a complete medical history, and physical and neurological examinations. Further, if data from initial baseline assessments are unavailable, other data may have to be gathered prior to formulating a treatment plan.

A psychiatric interview with the child, alone and together with his parents, is undertaken to determine a psychiatric diagnosis or diagnoses for which medication may be indicated. Further, this assessment enables the clinician to identify the target symptoms for treatment. While it is often thought that medication is prescribed for a target disorder (e.g., bipolar disorder, ADHD), in reality this is not the case. The clinician targets specific symptoms that serve as markers for the underlying psychopathology and the changes in these symptoms over the course of treatment are used to evaluate the effects of medication. Multiple symptoms are used as markers because some may respond to medication before others. For example, if we take a child with bipolar disorder, his sleep and appetite disturbances will improve earlier than his mood. Further, the clinician is careful to determine the phase of the child's illness (i.e., acute, relapse, recurrence) because of its importance in clinical decision making with regard to initial treatment as well as the need for and duration of maintenance and prophylactic therapy. Information on the temporal patterns

Table 1 Screening algorithm for attention-deficit hyperactivity disorder.

Initial referral may be either from school or home
Begin teacher–parent–child observational rating scale (e.g., Conners)
Refer to pediatrician or child psychiatrist or both:
  Complete physical examination, including:
    screen for lead toxicity and iron-deficiency anemia
    screen for other nutritional deficiencies
    detailed neurologic assessment, including perceptual motor tests to rule out focal neurological deficit
    thyroid panel (if indicated)
  Complete history, including:
    antenatal and perinatal history
    developmental milestones
    sleep history
    social and family history
    school history utilizing school records, grade-level screening medical history, especially trauma, infection, neurologic history, and allergies
  Use cognitive testing to assess intelligence and learning disabilities and organic brain syndromes
    Wechsler Intelligence Scale for Children-III
    Woodcock–Johnson
    Other tests, as indicated
  Psychometric testing
    Gordon diagnostic system
    Continuous performance test
  Evaluation of home situation
    If the evaluation is consistent with ADHD, a trial of sympathomimetic medication may be initiated
    Evaluation to monitor side effects, effectiveness of medication, four times yearly; alternative pharmaceutical or other therapies may be considered

of drug response and the phase of a child's illness is critical in the overall pharmacologic management of the child (Janicak et al., 1993).

5.11.6.4.1 Physical examination

The standard physical examination includes recording the child's temperature, pulse rate, respiratory rate, blood pressure, height, and weight. Because drugs may have adverse effects on the developing fetus, girls of childbearing age should be given a pregnancy test.

5.11.6.4.2 Laboratory tests and diagnostic procedures

Many of these tests would have been done during the baseline medical evaluation and need not be repeated. According to Green (1995, p. 26), standard laboratory tests for a pre-medication work-up may include the following:

(i) Complete blood cell count (CBC), differential, and hematocrit.

(ii) Urinanalysis.

(iii) Blood urea nitrogen (BUN) level.

(iv) Serum electrolyte levels for sodium (Na^+), potassium (K^+), chloride (Cl^-), calcium (Ca^{2+}), and phosphate ($PO_4{}^{3-}$) and carbon dioxide content (CO_2).

(v) Liver function tests: aspartate aminotransferase (AST) or serum glutamic oxaloacetic transaminase (SGOT), alanine aminotransferase (ALT) or serum glutamic pyruvic transaminase (SGPT), alkaline phosphatase, lactic dehydrogenase (LDH), and bilirubin (total and direct).

(vi) Serum lead level determination in children under seven years of age and in older children, when indicated.

In addition, other laboratory tests may be indicated prior to using specific psychoactive medications. For example, thyroid function tests are recommended prior to using tricyclic antidepressants and lithium; kidney function tests prior to using lithium carbonate; baseline and periodic electrocardiogram (ECG) for tricyclic antidepressants and lithium; an electroencephalogram (EEG) for antipsychotics, tricyclic antidepressants, or lithium for those who "have a history of seizure disorder, who are on an antiepileptic drug for a seizure disorder, or who may be at risk for seizures" (Green, 1995, p. 27).

5.11.6.4.3 Behavioral observations and rating scales

Except in emergency situations (e.g., violent, explosive behavior), a child is observed for 7–10 days for patterns in his general behavior as well as the target symptoms before initiating pharmacotherapy. In addition to the psychiatrist, clinical psychologists are in the best position to provide valid and reliable observation data on the child's behavior from multiple settings. These observations are useful in determining the stability of the child's problems as well as the need for medication, should the child's condition improve substantially during the observation period.

Behavioral observations are labor-intensive and often impractical in clinical practice. Fortunately, there are various rating scales that can be used to obtain reliable data on a child across multiple settings and by multiple raters (Sood and Singh, 1996). For example, the child behavior checklist (Achenbach, 1991a) is used widely for initial assessment of children and adolescents referred for psychiatric assessment, and the various Conners rating scales are used extensively with children with ADHD (Conners, (1990). Interested readers should consult Aman (1993) and Sood and Singh (1996) who have reviewed an extensive array of rating scales that can be used for this purpose.

5.11.6.5 Clinical Formulation

The diagnostic formulation follows the standard biopsychosocial model, which means that biological, psychological, and social factors are taken into account in developing a comprehensive framework for understanding the child's symptomatic presentation. Once the relevant data are gathered, the clinician synthesizes the information and reaches a working diagnosis. However, before a psychiatric diagnosis is reached, any general medical condition that may account for the child's problems are ruled out. In addition, it is critical that deviation from the normal range of development is considered. Further, a child may have symptoms of multiple conditions and all relevant diagnoses need to be considered (e.g, ADHD and learning disorders).

In the diagnostic formulation for psychopharmacological interventions, the clinician pays particular attention to signs and symptoms associated with psychiatric disorders that have a biological basis. These disorders have the potential for being responsive to psychoactive agents. Much of this information is obtained from the clinician's use of a descriptive, phenomenological assessment of the child's feelings, emotions, and behaviors because it provides important insights into the child's disorder(s). Information on the clinical course helps the clinician to place the disorder in a specific context because many pharmacologically responsive disorders have "characteristic beginnings, occur during particular stages of development, or are associated with particular sequelae" (Walkup, 1995, p. 26). For example, attention and concentration difficulties at age 5 or 6 years may be related to ADHD but at age 15 or 16 years, these symptoms are more likely to be associated with a mood disorder (Walkup, 1995). Detailed family history is also useful in diagnostic formulation for a number of reasons, such as providing additional support for the clinician's working hypothesis. The clinician uses all of this information to formulate a treatment plan that includes pharmacotherapy as well as some type of educational and psychosocial intervention because pharmacotherapy alone is rarely sufficient for complete recovery.

5.11.7 PSYCHOTROPIC MEDICATION IN CHILDHOOD DISORDERS

The main psychopharmacologic agents used in children are presented in Table 2 by class, generic name, and trade name. What follows in each section is a discussion of a given class of medication, mechanism of action, indications, dosing, and adverse effects. Only general principles and procedures of psychopharmacologic therapies are covered because applications of this treatment modality to specific psychopathologies are discussed elsewhere. However, as a general rule, it must be remembered that psychopharmacologic treatment is typically used as one component of a broad therapeutic approach to help alleviate psychiatric symptoms in children. Medication provides symptomatic relief, allowing the child to function more fully at school and at home, and in his community. Further, medication usually relieves the child's symptoms of psychiatric illness but it does not remove vulnerability to its recurrence because the environmental and constitutional stressors that gave rise to the illness are not affected by the medication.

5.11.7.1 Antipsychotics

The antipsychotics have been called major tranquilizers because of an early observation that chlorpromazine, the prototypical drug in this class, produced somnolence and relaxation. This led to the mistaken belief that the major action of antipsychotics was sedative. Antipsychotics have also been called neuroleptics because they are noted for producing signs of neurological dysfunction, principally Parkinson's syndrome, and other extrapyramidal reactions. Typically, drugs that produce both antipsychotic and extrapyramidal effects are called neuroleptic drugs. While the term neuroleptic continues to be used interchangeably with antipsychotic, the discovery of clozapine, an antipsychotic that is not a neuroleptic, means that this practice may now be inappropriate.

The antipsychotics consist of a group of eight classes of drugs, including phenothiazines (e.g., chlorpromazine [Thorazin], fluphenazine [Prolixin], thioridazine [Mellaril]), thioxanthenes (e.g., thiothixene [Navane]), dibenzoxazepines (e.g., loxapine [Loxitane]), dihydroindoles (e.g., molindone [Moban]), butyrophenones (e.g., haloperidol [Haldol], droperidol [Inapsine]), diphenylbutylpiperidines (e.g., pimozide [Orap]), benzamides (e.g., sulpiride [Dogmatil]), and benzisoxazole (e.g., risperidone [Risperdal]). Pharmacologically, the major classes of antipsychotics are remarkably similar and, when given in equivalent doses, they produce comparable benefits and induce a similar range of adverse effects (see Table 3).

5.11.7.1.1 Mechanism of action

The antipsychotics produce their therapeutic effects by blocking the D_2 subtype of dopamine receptors, although some of them may block the D_1 receptors as well. Blockade of the dopamine receptors is also responsible for most of the neurological and endocrinological adverse effects of antipsychotics. The difference in the adverse effects profiles of the antipsychotics is due to the fact that some of the antipsychotics also block noradrenergic, cholinergic, and histaminergic receptors. For example, blocking of the alpha-noradrenergic receptors by some antipsychotics probably accounts for their hypotensive action. When compared to the low potency antipsychotics (e.g., chlorpromazine, thioridazine), those with high potency (e.g., haloperidol, pimozide) are typically less sedating, produce less hypotension, and they generally have less effect on seizure threshold, fewer anticholinergic effects, less cardiovascular toxicity, less weight gain, and minimal effect on bone marrow and liver. Risperidone, a new antipsychotic medication, has D_2 as well as serotonergic (5-HT2) antagonist properties and lacks the acute extrapyramidal adverse effects of the traditional antipsychotics. Clozapine, an atypical antipsychotic, has weak D_2 antagonist properties but a high affinity for D_4 dopamine receptors, as well as a relatively strong antagonistic interaction with central alpha$_1$-adrenergic, cholinergic, histaminic (H1), and serotonergic (5-HT2) receptors (Baldessarini & Frankenburg, 1991; Lieberman, Kane, & Johns, 1989). Like risperidone, clozapine also lacks acute extrapyramidal adverse effects.

Table 2 Psychoactive and psychotropic drugs by class.

Drug class/subclass	*Generic name*	*Trade name*
A. Antipsychotics		
Phenothiazines	chlorpromazine	Thorazine
	fluphenazine	Prolixin, Modecate
	mesoridazine	Serentil
	perphenazine	Trilafon
	thioridazine	Mellaril
	trifluoperazine	Stelazine
Thioxanthenes	chlorprothixene	Taractan
	thiothixene	Navane
Butyrophenones	haloperidol	Haldol
	pipamperon	Dipiperon
Rauwolfia alkaloids	reserpine	Rauloydin, Reserpoid, Sandril
Other antipsychotics	clozapine	Clozaril
	loxapine	Loxitane
	risperidone	Risperdal
	olanzapine	Zyprexa
B. Antidepressants		
Tricyclics	amitriptyline	Elavil, Amitril, Endep
	clomipramine	Anafranil
	desipramine	Norpramin, Pertofrane
	doxepin	Sinequan, Adapin
	imipramine	Tofranil
	nortriptyline	Pamelor
Monoamine oxidase inhibitors	isocarboxazid	Marplan
	phenelzine	Nardil
	tranylcypromine	Parnate
Selective serotonin reuptake inhibitors	fluoxetine	Prozac
	fluvoxamine	Luvox
	paroxetine	Paxil
	sertraline	Zoloft
Other antidepressants	bupropion	Wellbutrin
	trazodone	Desyrel
	venlafaxine	Effexor
C. Antimanics		
	lithium carbonate	Eskalith, Lithane, Lithobid
D. Anxiolytics		
Benzodiazepines	alprazolam	Xanax
	chlordiazepoxide	Librium
	diazepam	Valium
	lorazepam	Ativan
	oxazepam	Serax
	prazepam	Verstran
	temazepam	Restoril
	triazolam	Halcion
Antihistamines	diphenhydramine	Benadryl
	hydroxyzine	Atarax
	promethazine	Phenergan
Atypical anxiolytics	buspirone	BuSpar
E. Stimulants		
	amphetamine sulfate	Benzedrine
	dextroamphetamine	Dexedrine

Table 2 (continued)

Drug class/subclass	*Generic name*	*Trade name*
	methamphetamine	Desoxyn
	methylphenidate	Ritalin
	pemoline	Cylert
	mixed salts product[a]	Adderall
F. Antiepileptics		
	carbamazepine	Tegretol
	clonazepam	Klonopin
	diazepam	Valium
	ethosuximide	Zarontin
	felbamate	Felbatol
	gabapentin	Neurontin
	phenobarbital	Luminal, Gardenal
	phenytoin	Dilantin
	primidone	Mysoline
	sodium valproate	Depakene, Depakote, Epilim
	sulthiame	Ospolot
G. Others		
Beta-blockers	propranolol	Inderal
Alpha-2 blockers (nonspecific)	clonidine	Catapres
Alpha-2A blockers (selective)	guanfacine	Tenex
Opioid antagonists	naloxone	Narcan
	naltrexone	Trexan
Sympathomimetic amines	fenfluramine	Pondimin, Ponderax
Anticholinergics	benztropine	Cogentin
	biperiden	Akineton
	ethopropazine	Parsidol
	procyclidine	Kemadrin
	trihexyphendiyl	Artane

[a] Combination of dextroamphetamine sulfate, dextroamphetamine saccharate, amphetamine sulfate, and amphetamine aspartate.

5.11.7.1.2 Indications

In children, the antipsychotics are used mainly in the treatment of psychotic disorders, typically the positive symptoms, as well as other behavioral symptoms, such agitation, aggression, tics, and stereotypies (Green, 1995; Kutcher, 1997). The positive symptoms include hallucinations, delusions, thought disorder (i.e., incoherency), and catatonic symptoms (e.g., stupor, negativism, rigidity, excitement, and posturing) or bizzare affect. The negative symptoms, which include affective blunting, poverty of speech and thought, apathy, anhedonia, and poor social functioning, respond well to the atypical antipsychotics (e.g., clozapine, risperidone). Among the antipsychotics, haloperidol and pimozide are the drugs of choice for treating children with Tourette's disorder. The low-potency drugs (e.g., chlorpromazine, thioridazine) are used to treat mania and, as a last resort, to treat aggressive behavior and severe agitation in children. The high-potency drugs (e.g., fluphenazine, perphenazine) are used for severe insomnia and severe self-injury. Clozapine is useful in treating refractory psychosis in adolescence, and risperidone is useful in treating both positive and negative symptoms of psychosis, as well as some symptoms associated with pervasive developmental disorder. Until recently, antipsychotics have been used widely in the treatment of self-injury, aggression, and agitation in children with developmental disorders (Ellis et al., 1997).

In terms of effects, there is little data to support the use of one antipsychotic drug over another because the action of these drugs is related to their structural similarities (Janicak et al., 1993). Thus, clinicians typically base their choice of one antipsychotic over another with regard to the adverse effect profile of the drugs.

Table 3 Antipsychotic drugs, trade names, potencies, and potential adverse effects.

Antipsychotic drug	*Trade name(s)*[a]	*Potency*[b]	*Potential adverse effects profile*			
			Sedation	*Anticholinergic*	*Hypotension*	*Extrapyramidal*
Chlorpromazine	Thorazine, Largactil	100 (low)	strong	strong	strong	weak
Fluphenazine	Prolixin, Modecate	1.5–3 (high)	moderate	weak	weak	strong
Haloperidol	Haldol	2–5 (high)	weak	weak	weak	strong
Loxapine	Loxitane	10–15 (medium)	moderate	moderate	moderate	strong
Mesoridazine	Serentil	50 (low)	moderate	strong	moderate	moderate
Molindone	Moban	6–10 (medium)	moderate	moderate	weak	strong
Perphenazine	Trilafon	10 (medium)	weak	weak	weak	strong
Pimozide	Orap	1 (high)	weak	weak	weak	strong
Risperidone	Risperdal	2–3 (low)	weak	weak	weak–moderate	weak
Thioridazine	Mellaril	97 (low)	strong	strong	strong	weak
Thiothixene	Navane	2–5 (high)	weak	weak	weak	weak
Trifluoperazine	Stelazine	3–5 (high)	moderate	weak	weak	strong
Triflupromazine	Vesprin	25–50 (low)	strong	moderate	strong	moderate

[a] Not all drugs are approved for use in children or in all countries. [b] mg of drug equivalent to 100 mg chlorpromazine.

For example, because children should maintain a relatively high level of activity for their developmental and psychological well-being, the clinician may choose haloperidol (which has fewer sedative and cardiovascular effects) over chlorpromazine (which has fewer extrapyramidal complications). Further, our history of clinical experience suggests that some children do in fact respond to one antipsychotic but not to another. It may be that this is due to differences between the drugs in terms of absorption, distribution, accumulation at receptor sites, pharmacodynamic actions, metabolism of their derivatives, and adverse cognitive effects (e.g., as a result of sedation). In addition, the drugs may appear to act differentially if their effects are not compared at optimum dosages for the child.

5.11.7.1.3 Dosing and adverse effects

The recommended doses for antipsychotic drugs typically used with children are presented in Table 4. In general, oral dosage of the low-potency antipsychotics range between 3.0 and 6.0 mg/kg/day. On average, the dosage of high-potency antipsychotics range between 0.1 and 0.5 mg/kg/day, although dosages up to 1.0 mg/kg/day are not uncommon. The dosage of risperidone is about 1–3 mg/day or up to 85 μg/kg/day, the dosage for clozapine is between 3.0 and 5.0 mg/kg (Spencer, Wilens, & Biederman, 1995).

The side effects of the antipsychotics are presented in Table 5.

5.11.7.2 Antidepressants

As the name implies, antidepressants are used to treat depression in children and adults. The serendipitous finding that iproniazid, a drug used to treat tuberculosis, had mood-elevating properties and inhibited the enzyme monoamine oxidase led to the development of a class of antidepressants called the monoamine oxidase inhibitors (MAOIs). Some years later, the drug imipramine (Tofranil) was developed, followed by a number of others with a similar structure—the tricyclic antidepressants. A new generation of antidepressants have become available, with some of them (e.g., fluoxetine) enjoying remarkable clinical success in the treatment of depression.

5.11.7.2.1 Mechanisms of action

The different classes of antidepressants can be categorized according to their structure and/or presumed mechanism of action. For example, Janicak et al. (1993, p. 242) categorized them as follows:

(i) Heterocyclics (HCAs, e.g., imipramine, Tofranil), which block the neuronal reuptake of norepinephrine and/or serotonin.

(ii) Selective serotonin reuptake inhibitors (SSRIs, e.g., fluoxetine [Prozac]), which modify serotonin neurotransmission by inhibiting the reuptake carrier.

(iii) Monoamine oxidase inhibitors (MAOIs; e.g., isocarboxazid [Marplan]), which increase the concentration of several biogenic amines.

(iv) Aminoketones (e.g., bupropion [Wellbutrin]), whose most potent known effect is neuronal dopamine reuptake blockade, but whose mechanism of action is unknown.

(v) Triazolopyridines (e.g., trazodone [Desyrel]), which have mixed effects on the serotonin system, with the predominant effect being 5-HT2 receptor blockade.

(vi) 5-HT_{1A} receptor partial agonists (e.g., buspirone [BuSpar]), marketed as anxiolytics, but which also have proven antidepressant properties, especially with higher doses.

Although the different classes of antidepressants have specific mechanisms of action, in general, the antidepressants block the reuptake of one or more catecholamines (i.e., norepinephrine, serotonin, and dopamine), thereby decreasing or downregulating the number of postsynaptic receptors. The MAOIs have a different mechanism of action; they increase the availability of catecholamines by blocking the enzyme monoamine oxidase.

5.11.7.2.2 Indications

In general, the antidepressants are used with children to treat ADHD, obsessive–compulsive disorder (OCD), enuresis, depression, anxiety, and tic disorders. The tricyclic antidepressants (e.g., imipramine, desipramine, nortriptyline) are used to treat major depressive disorders (MDD), enuresis, ADHD, tic disorders, and anxiety disorders. Clomipramine has been found to be effective with OCD. The SSRIs (e.g., fluoxetine, sertraline, paroxetine, fluvoxamine) are used to treat MDD, OCD, and anxiety disorders. Bupropion is used to treat ADHD and MDD; trazodone for MDD, aggression, and insomnia; and venlafaxine for MDD, OCD, and anxiety disorders. The MAOIs are now rarely used with children because of potential problems with medication compliance and adverse effects in this population (Viesselman, Yaylayan, Weller, & Weller, 1993).

Table 4 Recommended doses for the various classes of psychotropic drugs.

Drug	*Average daily dose* *Children*	*Adolescents*
A. Antipsychotics		
Chlorpromazine (> 6 months of age)[a]	30–200 mg (2.5–6 mg/kg/day)	40–400 mg (3–6 mg/kg/day)
Thioridazine (> 2 years of age)[a]	75–200 mg (0.5–3 mg/kg/day)	10–200 + mg
Trifluoperazine (> 6 years of age)[a]	1–15 mg	1–20 mg
Thiothixene (> 12 years of age)[a]	2–10 mg	5–30 mg
Haloperidol (> 3 years of age)[a]	0.5–4 mg (0.05–0.15 mg/kg/day)	2–16 mg
Reserpine	0.02–0.25 mg	0.1–1.0 mg
Clozapine (> 16 years of age)[a]		50–200 mg (3–5 mg/kg/day)
Loxapine (> 16 years of age)[a]	5–50 mg	20–100 mg
B. Antidepressants		
Amitriptyline (> 12 years of age)[a]	30–100 mg (1–5 mg/kg/day)	50–100 mg (1–5 mg/kg/day)
Bupropion (> 18 years of age)[a]	25–150 mg	75–300 mg (3–6 mg/kg/day)
Clomipramine (> 10 years of age)[a]	25–100 mg	50–150 mg (2–3 mg/kg/day)
Desipramine (> 12 years of age)[a]	10–150 mg (1–5 mg/kg/day)	50–150 mg (1–5 mg/kg/day)
Fluoxetine (> 18 years of age)[a]	5–20 mg	10–60 mg (0.5–1 mg/kg/day)
Fluvoxamine (> 8 years of age)[a]	25–200 mg	50–300 mg
Imipramine (> 6 years of age)[a]	10–150 mg (1–5 mg/kg/day)	50–200 mg (1–5 mg/kg/day)
Notriptyline (> 12 years of age)[a]	10–100 mg	50–100 mg (1–3 mg/kg/day)
Phenelzine (> 16 years of age)[a]		15–45 mg (0.5–1 mg/kg/day)
Sertraline (not in children)[a]	25–100 mg	50–200 mg (1.5–3 mg/kg/day)
C. Antimanics		
Lithium carbonate[b] (> 12 years of age)[a]	300–900 mg	900–1200 mg (10–30 mg/kg/day)
D. Anxiolytics		
Alprazolam (> 18 years of age)[a]	0.25–2 mg	0.75–5 mg (0.02–0.06 mg/kg/day)
Chlordiazepoxide (> 6 years of age)[a]	10–30 mg	20–60 mg
Diazepam (> 6 months of age)[a]	1–10 mg	2–20 mg (max 0.8 mg/kg/day)
Lorazepam (> 12 years of age)[a]	0.25–3 mg	0.05–6 mg (0.04–0.09 mg/kg/day)
Diphenhydramine	25–200 mg	50–300 mg (1–5 mg/kg/day)
Hydroxyzine	25–100 mg	40–150 mg (2 mg/kg/day)
Buspirone (> 18 years of age)[a]	2.5–15 mg	5–30 mg (0.2–0.6 mg/kg/day)

Table 4 (continued)

Drug	*Average daily dose* *Children*	*Adolescents*
E. Stimulants		
Dextroamphetamine (>3 years of age)[a]	2.5–15 mg (0.15–0.5 mg/kg/dose)	5–40 mg (0.15–0.5 mg/kg/dose)
Methylphenidate (>6 years of age)[a]	2.5–30 mg (0.3–1 mg/kg/dose)	10–60 mg (0.3–1 mg/kg/dose)
Pemoline (>6 years of age)[a]	18.75–75 mg (1–3 mg/kg/day)	37.5–112.5 mg (1–3 mg/kg/day)
Mixed salts[c] (>3 years of age)[a]	2.5–30 mg	5–40 mg
F. Antiepileptics		
Carbamazepine[b] (>6 years of age)[a]	200–800 mg (5–20 mg/kg/day)	400–1000 mg (10–30 mg/kg/day; max dose 1000 mg/day)
Ethosuximide[b]	250–800 mg (20–30 mg/kg/day)	500–1500 mg (20–40 mg/kg/day)
Phenobarbital[b]	<250 mg (4–8 mg/kg/day)	75–250 mg (1–3 mg/kg/day)
Phenytoin[b]	<300 mg (7.5–9 mg/kg/day)	300–500 mg (6–7 mg/kg/day)
Primidone[b]	150–750 mg	750–1500 mg
Sodium valproate[b]	250–1000 mg	500–2000 mg
G. Others		
Propranolol	5–80 mg	20–140 mg (max 2 mg/kg/day)
Clonidine (not in children)[a]	0.25–0.3 mg (3–6 μg/kg/day)	0.3–0.4 mg (3–6 μg/kg/day)
Naltrexone (>18 years of age)[a]	10–50 mg (0.5–1.5 mg/kg/day)	40–120 mg (1–2 mg/kg/day)
Fenfluramine (>12 years of age)[a]	30–60 mg	40–100 mg (1–2 mg/kg/day)
Benztropine (>3 years of age)[a]	0.5–4 mg	0.5–6 mg (43–86 μg/kg/day)

[a] Recommended US Food and Drug Administration guidelines. [b] Dosage titrated by serum levels. [c] Combination of dextroamphetamine sulfate, dextroamphetamine saccharate, amphetamine sulfate, and amphetamine aspartate.

5.11.7.2.3 Dosing and adverse effects

The recommended doses for antidepressant drugs typically used with children are presented in Table 4. The dose of antidepressant drugs for children has to be individualized because of interchild variability in metabolism and elimination. In general, the typical dose for tricyclic antidepressants in children ranges from 2.0 to 5.0 mg/kg, with 1.0–3.0 mg/kg for nortriptyline. The potential cardiotoxic effects of the tricyclic antidepressants, and particularly the reports of sudden death associated with desipramine therapy (Riddle et al., 1991), have reduced the use of tricyclic antidepressants by some clinicians. The suggested daily dose for fluoxetine is 0.5–1.0 mg/kg; for paroxetine 0.25–0.70 mg/kg, for sertraline 1.5–3.0 mg/kg; and for bupropion 4.0–6.0 mg/kg/day in divided doses.

The side effects of the antidepressants are presented in Table 5.

5.11.7.3 Antimanics

Lithium carbonate is the main antimanic drug used to treat children who have bipolar disorder, with the anticonvulsants, carbamazepine (Ballenger & Post, 1980), and valproic acid (McElroy, Keck, & Pope, 1987) being used as alternative antimanic agents.

5.11.7.3.1 Mechanism of action

Little research has been done on the use of these antimanic agents with children, and their therapeutic mechanism(s) is virtually unknown.

Table 5 Psychiatric and behavioral indications and side effects of various classes of drugs.

A. Antipsychotics	
Indications:	Psychotic states, schizophrenia (exacerbations and maintenance); mania (in conjunction with lithium); behavior disorders with severe agitation, aggressivity, and self-injury; and dyskinetic movement disorders (e.g., Tourette's disorder and juvenile Huntington's disease)
Side effects:	Anticholinergic effects, including dry mouth, constipation, blurred vision, and urinary retention (most common with low-potency phenothiazines); extrapyramidal reactions, including acute dystonia, akathisia, and tremor (particularly with high-potency phenothiazines); neuroleptic malignant syndrome; tardive dyskinesia (lower risk with clozapine and risperidone); other central nervous system effects, including sedation, fatigue, cognitive blunting, psychotic symptoms, confusion, and excitement; orthostatic hypotension and cardiac conduction abnormalities; endocrine disturbances (e.g., menstrual irregularities and weight gain); gastrointestinal distress; skin photosensitivity; granulocytopenia and agranulocytosis (clozapine); and allergic reactions.
B. Antidepressants	
Indications:	Enuresis; attention-deficit/hyperactivity disorder; major depressive disorder; and anxiety disorders (including school phobia, separation anxiety disorder, panic disorder, and obsessive-compulsive disorder)
Side effects:	*Tricyclics:* Anticholinergic effects, including dry mouth, constipation, blurred vision, and urinary retention; cardiac conduction slowing (treatment requires EKG monitoring), mild increases and/or irregularity in pulse rate, and mild decreases or increases in blood pressure; confusion or the induction of psychosis; seizures; rash; and endocrine abnormalities *Monoamine oxidase inhibitors:* Mild decreases or increases in blood pressure; drowsiness; weight gain; insomnia; and hypertensive crisis with nonadherance to dietary restrictions (necessary to eliminate high tyramine foods from diet) or with certain drugs *Selective serotonin reuptake inhibitors:* Irritability; gastrointestinal distress; headaches; and insomnia *Other antidepressants:* Irritability (bupropion, venlafaxine); insomnia (bupropion, venlafaxine); drug-induced seizures (bupropion, with high doses); changes in blood pressure (trazodone, venlafaxine); priapism (trazodone); sedation and sleepiness (trazodone, venlafaxine); gastrointestinal distress (venlafaxine); and headache (venlafaxine)
C. Antimanics	
Indications:	Manic episodes of bipolar disorder; unipolar depression/adjunct treatment in major depressive disorder; and behavior disorders with extreme aggression
Side effects:	Kidney abnormalities leading to increased urination and thirst; gastrointestinal distress; fine hand tremor, weakness, and ataxia; possible thyroid abnormalities (with long-term use), weight gain, and electrolyte imbalances; sedation, confusion, slurred speech, irritability, headache, and subtle cogwheel rigidity; skin abnormalities; orthostatic hypotension and pulse rate irregularities; and allergic reactions
D. Anxiolytics	
Indications:	Anxiety disorders; seizure control; night terrors; sleepwalking; insomnia and acute management of severe agitation; adjunct treatment in mania and refractory psychosis; and Tourette's disorder
Side effects:	Headache, sedation, and decreased cognitive performance; behavioral disinhibition, including overexcitement, hyperactivity, increased aggressivity, and irritability; gastrointestinal distress; central nervous system disinhibition resulting in hallucinations, psychotic-like behavior, and depression; physical and psychological dependence (particularly with long-acting benzodiazepines); rebound or withdrawal reactions (particularly with short-acting benzodiazepines); blood abnormalities; anticholinergic effects, including dry mouth, constipation, and blurred vision (antihistamines); and allergic reactions
E. Stimulants	
Indications:	Attention-deficit/hyperactivity disorder (including those with mental retardation, fragile X syndrome, Tourette's disorder, head trauma, pervasive developmental disorder, or other comorbid disorders); narcolepsy; and adjunctive treatment in refractory depression

Table 5 (continued)

Side effects:	Decreased appetite; weight loss; abdominal pain; headache, insomnia, irritability, sadness and depression; mild increases in pulse rate and blood pressure; possible temporary suppression of growth (with long-term use); choreoathetosis (pemoline) and, rarely, tic disorders; and elevated liver functions tests (pemoline)
F. Antiepileptics	
Indications:	Seizure control; bipolar disorder; adjunct treatment in major depressive disorder; and severe behavior problems (e.g., aggression and self-injury)
Side effects:	Sedation, weakness, dizziness, disturbances of coordination and vision, hallucinations, confusion, abnormal movements, nystagmus, slurred speech, and depression; blood abnormalities; gastrointestinal distress; skin rashes, alterations in pigmentation, and photosensitivity reactions; increased or decreased blood pressure and congestive heart failure; abnormalities of liver functions (sodium valproate, carbamazepine – rare); genitourinary tract dysfunction; coarsening of facial features, enlargement of the lips, gingival hyperplasia, and excessive hair growth (phenytoin); and bone marrow suppression (carbamazepine, sodium valproate)
G. Others	
Propranolol	
Indications:	Behavior disorders with severe aggression, self-injury, or agitation; Tourette's disorder; and akathisia
Side effects:	Decreased heart rate, peripheral circulation, and blood pressure; fatigue, weakness, insomnia, nightmares, dizziness, hallucinations, and mild symptoms of depression; shortness of breath and wheezing (especially in patients with asthma); gastrointestinal distress; and rebound hypertension on abrupt withdrawal
Clonidine	
Indications:	Attention-deficit/hyperactivity disorder; Tourette's disorder; behavior disorders with severe aggression, self-injury, or agitation; adjunct treatment of schizophrenia and mania; and possible use in anxiety disorders
Side effects:	Sedation; decrease in blood pressure; rebound hypertension; dry mouth; confusion (with high doses); and depression
Guanfacine	
Indications:	Attention-deficit/hyperactivity disorder and Tourette's disorder
Side effects:	Sedation (less than with clonidine); decrease in blood pressure (less than with clonidine); rebound hypertension; dry mouth; confusion (with high doses); and depression
Opioid antagonists	
Indications:	Self-injury and reversal of narcotic depression
Side effects:	Drowsiness; dizziness, dry mouth, sweating; nausea, abdominal pain; and loss of energy
Fenfluramine	
Indications:	Management of obesity and possible use in the control of some behavior problems in pervasive developmental disorder
Side effects:	Anorexia, weight loss; drowsiness, dizziness, confusion, headache, and incoordination; mood alterations, anxiety, insomnia, weakness, agitation, and slurred speech; gastrointestinal distress; increased or decreased blood pressure and palpitations; skin rashes; dry mouth; eye irritation; and muscle aches
Anticholinergics	
Indications:	Treatment of extrapyramidal reactions (dystonia, rigidity, tremor, and akathisia); sleep disorders; and agitation
Side effects:	Sedation; cognitive impairment; and anticholinergic effects, including dry mouth, constipation, and blurred vision

What we know is that they interact with various neurotransmitters, such as the catecholaminergic, indolaminergic, cholinergic, and γ-aminobutyric acid systems. It is likely that they affect both the pre- and postneuronal receptors as well as the postreceptor activity of these neurotransmitters.

5.11.7.3.2 Indications

Lithium is currently indicated for the treatment of bipolar disorder, aggression, impulsivity, and temper tantrums in children (Alessi, Naylor, Ghaziuddin, & Zubieta, 1994). The drug works reasonably well in children with behavior problems which are characterized by impulsivity, aggressiveness, rage, or emotional lability (Popper, 1995). However, clinicians are cautioned that most of this understanding comes from clinical experience and not from double-blind, placebo-controlled studies. Carbamazepine and valproic acid are indicated for bipolar disorder and as an adjunct treatment in refractory MDD.

5.11.7.3.3 Dosing and adverse effects

The recommended starting dosage for lithium ranges from 10 to 30 mg/kg in divided doses once or twice a day. Although there are no empirical data on this, Spencer et al. (1995) suggest serum levels of 0.8–1.5 $\mathrm{mEq/L^{-1}}$ for acute episodes of mania and 0.6–0.8 $\mathrm{mEq/L^{-1}}$ for maintenance therapy. The dosage and serum level needs to be individualized for optimum effectiveness. Given the approximately 18 h elimination half-life, lithium may take 5–7 days to reach steady state (Alessi et al., 1994). Daily carbamazepine dose in children ranges between 10 and 20 mg/kg, given twice a day. The recommended starting dose of valproic acid is 15 mg/kg/day, which can be gradually titrated up to 60 mg/kg/day, given three times a daily.

The side effects of the antimanics are presented in Table 5.

5.11.7.4 Anxiolytics

The anxiolytics are used to treat pathological anxiety states. Two general classes of drugs, antidepressants (mainly the tricyclics and the MAOIs) and the antianxiety agents (most notably the benzodiazepines), are used to treat the various anxiety disorders seen in children (e.g., separation anxiety disorder, panic disorder, agoraphobia, school phobia, and generalized anxiety disorder). The benzodiazepines are also known as sedative-hypnotics because they act as hypnotics in high doses, as anxiolytics at moderate doses, and as sedatives in low doses. In general, the sedatives reduce daytime activity, temper excitement, and generally calm the child, and hypnotics produce drowsiness and facilitate the onset and maintenance of sleep. The following benzodiazepine drugs are currently labeled as antianxiety agents for use in the United States: chlordiazepoxide (Librium); clorazepate (Tranxene); diazepam (Valium); lorazepam (Ativan); oxazepam (Serax); prazepam (Centrax); halazepam (Paxipam); and alprazolam (Xanax). Flurazepam (Dalmante) and temazepam (Restoril) are two other benzodiazepine drugs but these are marketed for insomnia. Buspirone, a nonbenzodiazepine anxiolytic, is an azaspirodecanedione that lacks sedative, anticonvulsant, or muscle relaxant properties.

5.11.7.4.1 Mechanism of action

The benzodiazepines are closely associated with γ-aminobutyric acid (GABA), the most prevalent inhibitory neurotransmitter system in the brain (Zorumski & Isenberg, 1991). The recognition sites for GABA are coupled to chloride ion channels, and when GABA binds to its receptors, these channels open, thereby increasing the flow of chloride ions into the neurons. There are two subtypes of central nervous system benzodiazepine receptors, BZ_1 and BZ_2. The BZ_1 receptors are involved in the mediation of sleep, and the BZ_2 receptors are involved in cognition, memory, and motor control. Theoretically, a benzodiazepine hypnotic can have more affinity for one or the other receptor, thereby producing differential response.

The mechanism of action of buspirone does not involve interaction with a benzodiazepine-GABA receptor. It interacts with presynaptic dopamine receptors and with postsynaptic 5-HT_{1A} (serotonin) receptors, for which it is a partial agonist. However, despite its affinity for dopamine receptors, buspirone is not a neuroleptic and it does not have adverse effects or risk profile of neuroleptics.

5.11.7.4.2 Indications

The high-potency benzodiazepines (e.g., clonazepam [Klonopin], alprazolam [Xanax], lorazepam [Ativan]) are used in the treatment of anxiety disorders, as an adjunct in the treatment of refractory psychosis and mania, severe agitation, Tourette's disorder, severe insomnia, and major depressive disorder with anxiety. The atypical anxiolytic, buspirone (BuSpar), is also used for anxiety disorders and as an adjunct in the treatment of refractory OCD. It is also used for treating aggressive behaviors in children with mental retardation and autism.

5.11.7.4.3 Dosing and adverse effects

The recommended dosage for the long-acting, high-potency benzodiazepine (i.e., clonazepam) is a one to three times a day administration of a total daily dose of 0.01–0.04 mg/kg (Spencer et al., 1995). The short-/intermediate-acting, high-potency benzodiazepines are given three to four times a day because they present a higher risk for rebound and withdrawal reactions. In terms of total daily doses, alprazolam is 0.02–0.06 mg/kg, lorazepam is 0.04–0.09 mg/kg, and oxazepam is 0.05–1.7 mg/kg (Spencer et al., 1995). The total daily dose of buspirone ranges from 0.3 to 0.6 mg/kg.

The side effects of the anxiolytics are presented in Table 5.

5.11.7.5 Stimulants

Stimulants are probably the most widely used psychotropic drugs with children. They are sympathomimetic drugs that are structurally similar to the endogenous catecholamines. The most commonly used stimulants include dextroamphetamine (Dexedrine), methylphenidate (Ritalin), and magnesium pemoline (Cylert).

5.11.7.5.1 Mechanism of action

The stimulants act in the central as well as in the peripheral nervous system by blocking the reuptake of catecholamines into the presynaptic nerve endings and by blocking degradation of the nerve endings by the monamine oxidase enzyme. It is likely that each of the stimulants has a distinct mode of action and therefore have differential clinical effects. Indeed, the precise mechanism of action of these compounds is still poorly understood.

5.11.7.5.2 Indications

The stimulants are indicated for ADHD, ADHD with comorbid disorders, ADHD in children with developmental disabilities, and as adjunctive therapy for refractory depression. Generally, stimulants are effective in controlling the symptoms of ADHD, especially hyperactivity, impulsivity, distraction, and inattention (Barkley, DuPaul, & Costello, 1993). Further, they are effective in improving parent–child interactions, peer relationships, academic productivity, and classroom behavior (Barkley, 1990). However, academic achievement is minimally affected by the stimulants but they do enhance performance on measures of vigilance, impulse control, fine motor coordination, and reaction time.

5.11.7.5.3 Dosing and adverse effects

For the short-acting stimulants, the recommended daily dosage is 0.3–1.5 mg/kg for dextroamphetamine and 1.0–2.0 mg/kg for methylphenidate. The onset of action for these compounds is within 30–60 min, with peak clinical effect being evident for up to three hours. The longer-acting magnesium pemoline is given in a single daily dose of 1–3 mg/kg. At therapeutic doses, the onset of action for this drug is very rapid, producing a clinical response almost immediately. Generally, the stimulants are given with or after meals because of their anorexigenic effects. Because of increased risk of liver toxicity with pemoline, it is generally not considered a first-line stimulant medication.

The side effects of the stimulants are presented in Table 5.

5.11.7.6 Antiepileptics

There is increasing indication that carbamazepine, clonazepam, and valproic acid are being used alone and in combination with standard psychotropic medications to treat adults with various psychiatric disorders, such as mania, depression, treatment-resistent or atypical panic disorder, and bulimia, among others (Post, 1987). As discussed in Section, Au:X-ref.OK?, carbamazepine and valproic acid are used as alternative mood-stabilizing antimanic agents in children with bipolar disorders. Case reports and uncontrolled trials suggest that carbamazepine may be useful in treating ADHD, intermittent explosive disorder, aggressive disorders, conduct disorder, and some mood disorders in children (Carpenter & Vining, 1993). However, there are few well-controlled studies attesting to the efficacy of carbamazepine in childhood disorders (Evans, Clay, & Gualtieri, 1987; Remschmidt, 1976), and it has not been approved by the Food and Drug Administration for neuropsychiatric indications in children.

Antiepileptics have a history of use in the developmental disabilities, especially for controlling intractable behavior problems. The primary psychiatric use of antiepileptics in children with developmental disabilities is for affective disorders, such as mania, bipolar disorder, or schizoaffective disorder, particularly if these disturbances have been resistant to standard psychotropic medication. Early studies strongly suggested that phenytoin (Dilantin) may have some behavioral effects on individuals with developmental disabilities, but these claims have not been validated in well-controlled studies (see Aman & Singh, 1991). Carbamazepine, valproic acid, and, to a much lesser extent, clonazepam are currently being used for this purpose (Ellis et al., 1997). Much remains to be done with regard to establishing the efficacy of the antiepileptics in childhood psychiatric disorders, their mechanism(s) of action, dosage, and adverse effects.

5.11.8 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

It is clear that we have come far since Bradley's serendipitous finding of the effects of benzedrine on the behavior of children. For several decades, psychopharmacology of childhood disorders was a neglected area of clinical practice because of a lack of knowledge. Research in the field was almost nonexistent because of a shortage of well-trained researchers as well as a lack of adequate scientific methodologies for studying drug effects in children. Indeed, for many decades, our information on drugs and dosage was derived from the adult literature even though we knew that children are not miniature adults. Indeed, studies showing that, unlike adults, adolescents do not reliably respond better to tricyclic antidepressants than to placebo attests to the fact that we cannot ever again assume that the clinical and research findings with adults will also be true for adolescents. That this assumption is passé has been endorsed by the FDA which now requires that all drugs intended for children to be tested with children so that the effects of these drugs and dosages can be appropriately determined.

Since the mid-1980s, child and adolescent psychopharmacology has achieved the status of a field worthy of scientific study in its own right. Yet much remains to be achieved. In general, virtually every area of psychopharmacology of childhood disorders needs further research. Indeed, we know so very little about the pharmacokinetics and pharmacodynamics of most major psychoactive medications as they pertain to children. Because developmental factors play such an important role in the outcome of pharmacotherapy, we need research that takes into account the interaction between the child's developmental level, environment, and different classes of drugs.

As for pharmacogenetics, we know little about the effects of hereditary factors on drug response with regard to adult populations and almost nothing in children (Nebert & Weber, 1990). For example, most of our knowledge of drug effects is with white populations, yet the majority of the world's population is of other ethnic origins and there is very little data on the interaction between ethnicity and psychotropic drugs. Current research strongly suggests that, at least in adults, ethnicity accounts for some of the interindividual variations in drug response to many psychotropic and other drugs (Lin, Poland, & Nakasaki, 1993). It is timely for us to consider ethnic differences in pharmacokinetics and pharmacodynamics in child populations. Other factors that impact drug outcome, such as compliance and placebo effects, are also influenced by cultural beliefs and expectations. Because these sociocultural determinants of treatment outcome are also an important factor in psychological therapy, this provides clinical psychologists a unique opportunity to study the effects of these variables in both psychological and psychopharmacologic treatment.

A majority of children with mental health problems have comorbid psychiatric disorders (Achenbach, 1991b). Comorbidity presents interesting challenges to clinicians because not only does it have important implications for pharmacologic treatment, it also has implications for other therapies. Of interest is the use of targeted combined psychotropic drugs for comorbid disorders in children. While we have traditionally eschewed the use of polypharmacy, especially intraclass polypharmacy, our current thinking is that there is indeed a place for combined psychotropic treatments. However, much of this thinking is based on clinical experience and not on well-controlled outcome studies. Further, given that psychopharmacologic treatment is only one component of a treatment regimen that includes psychosocial as well as educational intervention, more research is needed to assess the individual, combined, and interactive effects of these treatments.

In terms of clinical practice, there has been growing interest in individualizing care within standardized treatment protocols. For example, clinicians and researchers have been developing psychopharmacologic treatment algorithms for the various childhood disorders (Vitiello, 1997; Walkup, 1995). These algorithms are based mainly on our clinical experience and take into account children with comorbid disorders as well as interindividual variations in drug response. Because these algorithms have been developed on the basis of global characteristics of children with specific disorders, it assumes that the physician will use good clinical judgement with individual cases based on thorough baseline assessments. Further, because these algorithms are multistep procedures, an empirical evaluation of a particular algorithm for a specific disorder will necessitate the evaluation of the entire algorithm rather than its component steps. This is an exciting development in child psychopharmacology as it integrates clinical practice and research, where the clinical concerns of the child drives the research protocol.

5.11.9 SUMMARY

Psychopharmacology of childhood disorders is a growing area of clinical practice and research. For most of its 60 years, childhood psychopharmacology has played second fiddle to adult psychopharmacology but this is rapidly changing. The biological revolution has affected both child and adult psychiatry equally. Further, experimental and clinical psychologists have brought to this field useful research methodologies that have anchored psychopharmacological research in the context of the child's life rather than just on the child's biology. However, much of our current knowledge is still based on experience rather than on science. While child psychopharmacology is both a science and an art, we are well short on the science.

5.11.10 REFERENCES

Achenbach, T. M. (1991a). *Manual for the CBCL/4-18 and profile*. Burlington, VT: Department of Psychiatry, University of Vermont.

Achenbach, T. M. (1991b). Comorbidity in child and adolescent psychiatry: Categorical and quantitative perspectives. *Journal of Child and Adolescent Psychopharmacology, 1,* 271–278.

Alessi, N., Naylor, M. W., Ghaziuddin, M., & Zubieta, J. K. (1994). Update on lithium carbonate therapy in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 291–304.

Aman, M. G. (1993). Monitoring and measuring drug effects. II. Behavioral, emotional, and cognitive effects. In J. S. Werry & M. G. Aman (Eds.), *Practitioner's guide to psychoactive drugs for children and adolescents* (pp. 99–159). New York: Plenum.

Aman, M. G., & Singh, N. N. (1988). *Psychopharmacology of the developmental disabilities*. New York: Springer-Verlag.

Aman, M. G., & Singh, N. N. (1991). Psychopharmacological intervention. In J. L. Matson & J. A. Mulick (Eds.), *Handbook of mental retardation* (2nd ed., pp. 347–372). New York: Pergamon.

Aman, M. G., Van Bourgondien, M. E., Wolford, P. L., & Sarphare, G. (1995). Psychotropic and anticonvulsant drugs in subjects with autism: Prevalence and patterns of use. *Journal of the American Academy of Child & Adolescent Psychiatry, 34,* 1672–1681.

Baldessarini, R., & Frankenburg, F. R. (1991). Clozapine: A novel antipsychotic agent. *New England Journal of Medicine, 324,* 746–754.

Ballenger, J. C., & Post, R. M. (1980). Carbamazepine in manic-depressive illness: A new treatment. *American Journal of Psychiatry, 137,* 782–790.

Barber, C. C., Rosenblatt, A., Harris, L. M., & Attkisson, C. C. (1992). Use of mental health services among severely emotionally disturbed children and adolescents in San Francisco. *Journal of Child and Family Studies, 1,* 183–207.

Barkley, R. A. (1990). *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment*. New York: Guilford Press.

Barkley, R. A., DuPaul, G. J., & Costello, A. (1993). Stimulants. In J. S. Werry & M. G. Aman (Eds.), *Practitioner's guide to psychoactive drugs for children and adolescents* (pp. 205–237). New York: Plenum.

Barkley, R. A., & Cunningham, C. (1979). The effects of Ritalin on the mother–child interactions of hyperactive children. *Archives of General Psychiatry, 36,* 201–208.

Beardsley, R. S., Gardocki, G. J., Larsen, D. B., & Hidalgo, J. (1988). Prescribing of psychotropic medication by primary care physicians and psychiatrists. *Archives of General Psychiatry, 45,* 1117–1119.

Beasley, C. M., & Dornseif, B. E. (1991). Fluoxetine and suicide: A meta-analysis of controlled trials of treatment of depression. *British Medical Journal, 303,* 685–692.

Bradley, C. (1937). The behavior of children receiving benzedrine. *American Journal of Orthopsychiatry, 9,* 577–585.

Brotman, A. (1992). *Practical reviews in psychiatry*. Birmingham, AL: Educational Reviews.

Burton, T. M. (1991). Antidepressant drug of Eli Lilly loses sales after attack by sect. *Wall Street Journal, April 19,* A1–A2.

Cade, J. (1949). Lithium salts in the treatment of psychotic excitement. *Medical Journal of Australia, 36,* 349.

Carpenter, R. O., & Vining, E. P. G. (1993). Antiepileptics (anticonvulsants). In J. S. Werry & M. G. Aman (Eds.), *Practitioner's guide to psychoactive drugs for children and adolescents* (pp. 321–346). New York: Plenum.

Centers for Disease Control (1983). Valproate: A new cause of birth defects—Report from Italy and follow-up from France. *MMWR, 32,* 438–439.

Clein, P. D, & Riddle, M. A. (1995). Pharmacokinetics in children and adolescents. *Child and Adolescent Psychiatric Clinics of North America, 4,* 59–75.

Conners, C. K (1990). *Conners' rating scales manual, Conners' teacher rating scales, Conners parent rating scales. Instruments for use with children and adolescents*. North Tonawanda, NY: Multi-Health Systems.

Coyle, I., Wayner, M. J., & Singer, G. (1976). Behavioral teratogenesis: A critical evaluation. *Pharmacology, Biochemistry and Behavior, 4,* 191–200.

Cullinan, D., Epstein, M. H., & Sabornie, E. J. (1992). Selected characteristics of a national sample of seriously emotionally disturbed adolescents. *Behavioral Disorders, 17,* 273–280.

Cullinan, D., Gadow, K. D., & Epstein, M. H. (1987). Psychotropic drug treatment among learning-disabled, educable mentally retarded, and seriously emotionally disturbed students. *Journal of Abnormal Child Psychology, 15,* 469–477.

Cunningham, C., & Barkley, R. (1978). The effects of methylphenidate on the mother–child interactions of hyperactive identical twins. *Developmental Medicine and Child Neurology, 20,* 634–642.

DiMascio, A., Soltys, J. J., & Shader, R. I. (1970). Psychotropic drug side effect in children. In R. I. Shader & A. DiMascio (Eds.), *Psychotropic drug side effects* (pp. 235–269). Baltimore: Williams and Wilkins.

Dingemanse, J., Danhof, M., & Breimer, D. D. (1988). Pharmacokinetic-pharmacodynamic modeling of CNS drug effects: An overview. *Pharmacology and Therapeutics, 38,* 1–52.

Ellis, C. R., Singh, Y. N., & Singh, N. N. (1997). Use of behavior-modifying drugs. In N. N. Singh (Ed.), *Prevention and treatment of severe behavior problems: Models and methods in developmental disabilities* (pp. 149–176). Pacific Grove, CA: Brooks/Cole Publishing Co.

Epstein, M. H., Cullinan, D., Quinn, K. P., & Cumbald, C. (1994). Characteristics of children with emotional and behavioral disorders in community-based programs designed to prevent placement in residential facilities. *Journal of Emotional and Behavioral Disorders, 2,* 51–57.

Epstein, M. H., Cullinan, D., Quinn, K. P., & Cumbald, C. (1995). Personal, family, and service use characteristics of young people served by an interagency community-based system of care. *Journal of Emotional and Behavioral Disorders, 3,* 55–64.

Evans, R. W., Clay, T. H., & Gualtieri, C. T. (1987). Carbamazepine in pediatric psychiatry. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 2–8.

Evans, W. E., Schentag, J. J., & Jusko, W. J. (1986). *Applied pharmacokinetic principles of therapeutic drug monitoring* (2nd ed.). San Francisco: Applied Therapeutics.

Fava, M., & Rosenbaum, J. F. (1991). Suicidality and fluoxetine: Is there a relationship? *Journal of Clinical Psychiatry, 52,* 108–111.

Foxman, B., Valdez, R. B., & Brook, R. H. (1986). Childhood enuresis: Prevalence, perceived impact, and prescribed treatments. *Pediatrics, 77,* 482–487.

Gadow, K. D. (1981). Prevalence of drug treatment for hyperactivity and other childhood behavior disorders. In K. D. Gadow & J. Loney (Eds.), *Psychosocial aspects of drug treatment for hyperactivity* (pp. 13–76). Boulder, CO: Westview Press.

Garfinkel, B. D., & Amrami, K. K. (1992). Assessment and differential diagnosis of attention-deficit hyperactivity disorder. *Child and Adolescent Psychiatry Clinics of North America, 1,* 311–324.

Gates, D. (1987). Just saying no to Ritalin. *Newsweek, November 23,* 6.

Green, W. H (1995). *Child and adolescent clinical psychopharmacology* (2nd ed.). Baltimore: Williams and Wilkins.

Greenblatt, D. J., & Shader, R. I. (1987). Introduction: pharmacokinetics in clinical psychiatry and psychopharmacology. In H. Y. Meltzer (Ed.), *Psychopharmacology: The third generation of progress* (p. 1339). New York: Raven Press.

Hancock, L. (1996). Mother's little helper. *Newsweek March 18,* pp. 51–56.

Institute of Medicine (1990). *Research on children and adolescents with mental, behavioral, and developmental disorders*. Bethesda, MD: National Institutes of Mental Health.

Janicak, P. G., Davis, J. M., Preskorn, S. H., & Ayd, F. J. (1993). *Principles and practice of psychopharmacology*. Baltimore: Williams and Wilkins.

Johnson, P. (1988). Family: Remedy led to "hell." *USA Today, April 27.*

Kalow, W. (1990). Pharmacogenetics: Past and future. *Life Sciences, 47,* 1385–1397.

Kaplan, S. L., Simms, R. M., & Busner, J. (1994). Prescribing practices of outpatient child psychiatrists. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 35–44.

Keller, M. B., Lavori, P. W., Beardslee, W. R., & Wunder, J. (1991). Depression in children and adolescents: New data on "undertreatment" and a literature review on the efficacy of available treatments. *Journal of Affective Disorders, 21,* 163–171.

Kovacs, M., Feinberg, T. L., Crouse-Novak, M. A., Paulauskas, S. L., & Finkelstein, R. (1984). Depressive disorders in childhood: A longitudinal prospective study of characteristics and recovery. *Archives of General Psychiatry, 41,* 229–237.

Kutcher, S. P. (1997). *Child and adolescent psychopharmacology*. Philadelphia: W. B. Saunders.

Kye, C., & Ryan, N. (1995). Pharmacologic treatment of child and adolescent depression. *Child and Adolescent Psychiatric Clinics of North America, 4,* 261–281.

Landrum, T. J., Singh, N. N., Nemil, M. S., Ellis, C. R., & Best, A. M. (1995). Characteristics of children and adolescents with serious emotional disturbance in systems of care. Part II: Community-based services. *Journal of Emotional and Behavioral Disorders, 3,* 141–149.

Landry, S. (1996). Uncertainty grows over use of Ritalin. *Richmond Times-Dispatch, March 28,* A2.

Lieberman, J. A., Kane, J. M., & Johns, C. A. (1989). Clozapine: Guidelines for clinical management. *Journal of Clinical Psychiatry, 50,* 329–338.

Lin, K. M., Poland, R. E., & Nakasaki, G. (1993). *Psychopharmacology and psychobiology of ethnicity*. Washington, DC: American Psychiatric Press.

Lipman, R. S. (1970). The use of psychopharmacological agents in residential facilities for the retarded. In F. J. Menolascino (Ed.), *Psychiatric approaches to mental retardation* (pp. 387–398). New York: Basic Books.

Maynard, R. (1970). Omaha pupils given "behavior" drugs. *Washington Post, June 29.*

McElroy, S. L., Keck, P., & Pope, H. G. (1987). Sodium valproate: Its use in primary psychiatric disorders. *Journal of Clinical Psychopharmacology, 7,* 16–24.

Moerman, D. E. (1979). Anthropology of symbolic healing. *Currents in Anthropology, 20,* 59–80.

Nebert, D. W., & Weber, W. W. (1990). Pharmacogenetics. In W. B. Pratt & P. Taylor (Eds.), *Principles of drug action* (pp. 469–531). New York: Churchill Livingstone.

Olfson, M., Pincus, H. A., & Sabshin, M. (1994). Pharmacotherapy in outpatient psychiatric practice. *American Journal of Psychiatry, 151,* 580–585.

Paxton, J. W., & Dragunow, M. (1993). Pharmacology. In J. S. Werry & M. G. Aman (Eds.), *Practitioner's guide to psychoactive drugs for children and adolescents* (pp. 23–55). New York: Plenum.

Popper, C. W. (1995). Balancing knowledge and judgement: A clinician looks at new developments in child and adolescent psychopharmacology. *Child and Adolescent Psychiatric Clinics of North America, 4,* 483–513.

Post, R. M. (1987). Mechanisms of action of carbamazepine and related anticonvulsants in affective illness. In H. Y. Meltzer (Ed.), *Psychopharmacology: The third generation of progress* (pp. 567–594). New York: Raven Press.

Remschmidt, H. (1976). The psychotropic effects of carbamazepine in non-epileptic patients, with particular reference to problems posed by clinical studies in children with behavioral disorders. In W. Birkmayer (Ed.), *Epileptic seizures, behavior, pain* (pp. 253–258). Bern: Hans Huber.

Riddle, M. A., Nelson, J. C., Kleinman, C. S., Rasmusson. A., Leckman, J. F., King, R. A., & Cohen, D. J. (1991). Sudden death in children receiving Norpramin: A review of three reported cases and commentary. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 104–108.

Safer, D. J., & Krager, J. M. (1988). A survey of medication treatment for hyperactive/inattentive children. *Journal of the American Medical Association, 260,* 2256–2258.

Safer, D. J., & Krager, J. M. (1992). Effect of a media blitz and a threatened lawsuit on stimulant treatment. *Journal of the American Medical Association, 268,* 1004–1007.

Schou, M. (1990). Lithium treatment during pregnancy, delivery, and lactation: An update. *Journal of Clinical Psychiatry, 51,* 410–412.

Sherman, M., & Hertzig, M. (1991). Prescribing practices of Ritalin: The Suffolk County, New York study. In L. L. Greenhill & B. B. Osman (Eds.), *Ritalin: Theory and patient management* (pp. 187–193). New York: Mary

Ann Liebert.

Silver, S. E., Duchnowski, A. J., Kutash, K., Friedman, R. M., Eisen, M., Prange, M. E., Brandenburg, N. A., & Greenbaum, P. E. (1992). A comparison of children with serious emotional disturbance served in residential and school settings. *Journal of Child and Family Studies, 1,* 43–59.

Singh, N. N., & Aman, M. G. (1990). Ecobehavioral assessment of pharmacotherapy. In S. Schroeder (Ed.), *Ecobehavioral analysis in developmental disabilities* (pp. 182–200). New York: Springer-Verlag.

Singh, N. N., Ellis, C. R., & Singh, Y. N. (1994). Medication management. In E. Cipani & F. Spooner (Eds.), *Curricular and instructional approaches for persons with severe handicaps* (pp. 403–423). Boston: Allyn & Bacon.

Singh, N. N., Ellis, C. R., & Wechsler, H. A. (in press). Psychopharmacoepidemiology in mental retardation. *Journal of Child and Adolescent Psychopharmacology.*

Singh, N. N., Guernsey, T. F., & Ellis, C. R. (1992). Drug therapy for persons with developmental disabilities: Legislation and litigation. *Clinical Psychology Review, 12,* 665–679.

Singh, N. N., Landrum, T. J., Donatelli, L., Hampton, C., & Ellis, C. R. (1994). Characteristics of children and adolescents with serious emotional disturbance in systems of care. Part 1: Partial hospitalization and inpatient psychiatric services. *Journal of Emotional and Behavioral Disorders, 2,* 13–20.

Singh, N. N., Parmelee, D. X., Sood, A., & Katz, R. C. (1993). Collaboration of disciplines. In J. L. Matson (Ed.), *Handbook of hyperactivity in children* (pp. 305–322). Boston, MA: Allyn & Bacon.

Singh, N. N., & Winton, A. S. W. (1989). Behavioral pharmacology. In J. K. Luiselli (Ed.), *Behavioral medicine and developmental disabilities* (pp. 152–179). New York: Springer-Verlag.

Sood, A., & Singh, N. N. (1996). Diagnostic instruments. In D. X. Parmelee (Ed.), *Child and adolescent psychiatry* (pp. 19–31). St. Louis, MO: Mosby.

Spencer, T., Wilens, T., & Biederman, J. (1995). Psychotropic medication for children and adolescents. *Child and Adolescent Psychiatric Clinics of North America, 4,* 97–121.

Sprague, R. L., & Werry, J. S. (1971). Methodology of psychopharmacological studies in the retarded. *International Review of Research in Mental Retardation, 5,* 147–219.

Stefl, M. E., Bornstein, R. A., & Hammond, L. (1988). *The 1987 Ohio Tourette survey.* Milford, OH: Tourette Syndrome Association of Ohio.

Stephen, K., Sprague, R. L., & Werry, J. (1973). *Drug treatment of hyperactive children in Chicago.* Urbana, IL: Children's Research Center.

Teicher, M. H., Glod, C., & Cole, J. O. (1990). Emergence of intense suicidal preoccupation during fluoxetine treatment. *American Journal of Psychiatry, 147,* 207–210.

Towbin, K. E. (1995). Evaluation, establishing the treatment alliance, and informed consent. *Child and Adolescent Psychiatric Clinics of North America, 4,* 1–14.

Viesselman, J. O., Yaylayan, S., Weller, E. B., & Weller, R. A. (1993). Antidysthymic drugs (antidepressants and antimanics). In J. S. Werry & M. G. Aman (Eds.), *Practitioner's guide to psychoactive drugs for children and adolescents* (pp. 239–268). New York: Plenum.

Vitiello, B. (1997). Treatment algorithms in child psychopharmacology research. *Journal of Child and Adolescent Psychopharmacology, 7,* 3–8.

Walkup, J. T. (1995). Clinical decision making in child and adolescent psychopharmacology. *Child and Adolescent Psychiatric Clinics of North America, 4,* 23–40.

Werry, J. S. (1993). Introduction: A guide for practitioners, professionals, and public. In J. S. Werry & M. G. Aman (Eds.), *Practitioner's guide to psychoactive drugs for children and adolescents* (pp. 3–21). New York: Plenum.

White, L. Tursky, B., & Schwartz, G. E. (1995). *Placebo: Theory, research and mechanisms.* New York: Guilford Press.

Zhou, H. H., Koshakji, R. P., Silberstein, D. J., Wilkinson, G. P., & Wood, A. J. (1989). Racial differences in drug response: Altered sensitivity to and clearance of propranolol in men of Chinese descent as compared with American whites. *New England Journal of Medicine, 320,* 565

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.12
Preventive Interventions

SUSAN H. SPENCE
University of Queensland, Brisbane, Qld, Australia

5.12.1 INTRODUCTION 295

5.12.2 APPROACHES TO PREVENTION 296

5.12.3 RISK FACTORS 297
- *5.12.3.1 Genetic and Biological Factors* 297
- *5.12.3.2 Family Environment* 298
- *5.12.3.3 Negative Life Events* 299
- *5.12.3.4 Intrinsic Child Characteristics* 300

5.12.4 PROTECTIVE FACTORS 301

5.12.5 IMPLICATIONS FOR PREVENTION 302

5.12.6 THE EFFECTIVENESS OF PREVENTIVE PROGRAMS 303
- *5.12.6.1 Universal Preventive Interventions* 303
- *5.12.6.2 Selective Preventive Interventions* 304
  - *5.12.6.2.1 Children exposed to socioeconomic deprivation* 304
  - *5.12.6.2.2 Assisting children to deal with school transition* 306
  - *5.12.6.2.3 Children of parents with psychiatric disorders* 306
  - *5.12.6.2.4 Children who have experienced trauma* 307
  - *5.12.6.2.5 Children whose parents separate or divorce* 307
  - *5.12.6.2.6 Children who show multiple risk factors* 308
- *5.12.6.3 Indicated Preventive Interventions* 308
  - *5.12.6.3.1 Early intervention to prevent conduct disorder* 309
  - *5.12.6.3.2 Early intervention to prevent anxiety and depression* 310
  - *5.12.6.3.3 Early intervention to prevent psychotic episodes* 311

5.12.7 DIFFICULTIES AND CHALLENGES 311

5.12.8 CONCLUSIONS 312

5.12.9 REFERENCES 312

5.12.1 INTRODUCTION

The past decade has seen rapid developments in the area of prevention of psychopathology. Mental health policymakers and planners have gradually come to recognize the value of prevention rather than intervention after problems have become well established. The prevention of psychological problems is typically emphasized as a target of many mental health policies across the world (Spence, 1996a). The increased focus upon prevention is reflected in the numerous published studies that describe preventive programs in many parts of the world. The literature reveals a wide spectrum of preventive efforts, ranging from generic attempts to enhance psychological well-being to those that focus on prevention of specific psychological disorders or a particular high-risk group. Nevertheless, the funds spent by health care providers on prevention of psychopathology represent a very small proportion of the total mental health budget (Mrazek & Haggerty, 1994; Spence, 1996a). In consequence, the vast majority of mental health professionals focus their work upon the treatment of existing psychopathology, rather than upon its prevention.

There are many reasons why prevention of psychopathology makes sense. Although estimates vary considerably, epidemiological studies suggest that, in any given year, around 20–30% of adults (Robins & Regier, 1991) and 10–15% of children and adolescents (Department of Health and Human Services, 1991; Institute of Medicine, 1989) will exhibit a psychological problem of sufficient magnitude to interfere significantly in daily functioning and that meets diagnostic criteria for a clinical disorder. The cost to the community of psychological disorders is enormous. For example, in 1993 the cost of psychological disorders in Australia (with a population of around 17 million) was estimated as being between A$3 and A$6 billion per annum (Human Rights and Equal Opportunity Commission, 1993, p. 835). This figure included the cost of consultations with mental health professionals, hospitalization, medications, housing, and pension payments. Estimating the annual cost of mental illness, drug and alcohol abuse in the USA for 1990, Eisenberg (1995) and Mrazek and Haggerty (1994) cite evidence to suggest an overall cost to the community of US$212 billion. Thus, successful prevention of mental health problems would produce huge economic savings to the community, over and above the obvious benefits of reduced personal suffering.

5.12.2 APPROACHES TO PREVENTION

Before proceeding, it is important to discuss the various approaches to prevention and the terminology used to describe these methods. Traditionally, three levels of prevention were described, namely primary, secondary, and tertiary (Caplan, 1964). The first level of primary prevention referred to efforts to reduce the incidence of new cases through intervention before disorders occur. Secondary prevention referred to interventions that aimed to reduce the prevalence of disorders through early identification of established cases. Tertiary prevention referred to efforts to reduce the duration and degree of disability associated with an existing disorder, through treatment and prevention of relapse.

The disadvantage of the primary, secondary, and tertiary distinction was that it could be said that secondary and tertiary levels related more to treatment than to prevention. Only the primary level related specifically to intervention prior to the development of a disorder, but the term provided minimal information about the type of intervention being used. More recently, the prevention literature adopted an alternative approach to classifying preventive efforts that provides more detail about the type of prevention approach (Gordon, 1987). Gordon proposed three types of preventive approaches based on the presence and extent of risk factors related to the development of the disorder of concern among individuals who did not already manifest a disorder. These approaches were labeled universal, selective, and indicated. Universal preventive interventions are those that are provided to entire populations. Selective prevention is targeted to those individuals who are members of a group, the membership of which places them at increased risk for the development of a mental disorders. Indicated prevention focuses on asymptomatic but high-risk individuals who are found to manifest a risk factor, condition, or abnormality that identifies them as being at high risk for the future development of psychopathology (Gordon, 1987).

Although the universal, selective, and indicated distinction represented an advance upon the primary, secondary, and tertiary approach, this method of categorization is still subject to criticism. Such an approach assumes that psychological disorders are discrete, present–absent phenomena. In reality, most forms of psychopathology lie on a continuum from mild or few symptoms to more severe and/or numerous symptoms. The decision to classify a person as experiencing a mental disorder depends on meeting some arbitrary cut-off for the number and severity of symptoms. Furthermore, early, mild symptomatology is a risk factor for later development of many forms of psychopathology. For example, conduct disorder in young people tends to develop along a trajectory in which early symptoms are predictive of later development of more severe symptomatology (Farrington, 1995). Once on this trajectory, many children show increasing levels of symptomatology with increasing age. Thus, indicated prevention must also target those individuals who show early, mild symptomatology that represents a risk factor for later development of a more serious psychological disorder. Any method of categorizing preventive interventions needs to take into account the continuum of psychological difficulties and the trajectory in the development of many disorders. Implicit in Gordon's (1987) approach to prevention is the assumption that evidence is available regarding risk factors for the development of mental disorders. Risk factors refer to biological, environmental, and psychological factors that increase the probability of development of a psychological problem. Effective prevention requires not only that we are able to identify risk factors for particular forms of psychopathology, but also that we have reliable and valid methods of identifying those individuals who manifest these risk features and that we have effective means of altering these variables. As described below, there is now considerable evidence regarding those factors that increase the risk of development of psychological problems. Similarly, evidence has shown that many of these variables can be changed effectively.

One further development that should be mentioned in the area of prevention is the focus not only on identification and alteration of risk variables but also on protective factors. Protective factors refer to those variables that produce a resilience to the development of psychological difficulties in the face of adverse risk factors. Clearly there are many young people who emerge from the face of adversity relatively well-adjusted and unaffected by the risk factors to which they have been exposed. Increasingly, research has focused on identifying variables associated with resilience which buffer the impact of risk factors. This shift in emphasis is important, as there are many risk factors that cannot be easily altered or removed, and an alternative or additional strategy for prevention is to build up protective factors.

5.12.3 RISK FACTORS

Efforts to identify causal factors relating to the development of child and adolescent psychopathology have yielded a great deal of information concerning risk and protective factors. This knowledge has emerged from longitudinal studies that allow us to entangle some of the interrelationships between psychopathology and the myriad of correlated variables that are associated with psychological disorders. Although we are a long way from fully understanding the etiological factors in child psychopathology, the evidence is sufficiently well developed to allow us to draw some tentative conclusions. These points are discussed in more detail below but can be summarized as follows:

(i) It is clear that child psychopathology results from a complex interaction of numerous environmental and intrinsic child variables.

(ii) There are many different sources of influence and no one factor is sufficient to account for any specific emotional and behavioral disorder.

(iii) Risk factors appear to have a multiplicative, not just additive effect so that the probability of developing a behavioral or emotional disorder increases considerably as the number of risk factors affecting a child increases.

(iv) Some risk factors are common to a wide range of emotional and behavioral disorders, whereas others have a more specific influence upon particular forms of psychopathology.

5.12.3.1 Genetic and Biological Factors

It is well established that parents with psychopathology are more likely than parents without psychopathology to have children who exhibit emotional and behavioral problems. Similarly, children with psychiatric disorders are more likely than their well-adjusted peers to have parents who meet diagnostic criteria for a mental disorder. These familial factors hold for internalizing and externalizing disorders alike (Rutter et al., 1990). However, such results tell us little about the relative contribution of genetic and family environmental factors to the development of psychopathology. Parental psychopathology is frequently associated with greater family stress, financial difficulties, marital discord, and poor parenting skills. Thus, in addition to genetic contributions, we must also consider the environmental influences that accompany parental psychopathology. Although strong familial associations are evident for childhood depression, anxiety, conduct disorder, and attention deficit disorder, the evidence has not yet clarified the degree to which these reflect genetic or environmental influences (Rutter et al.). It is only for disorders such as autism, multiple tics, Tourette's syndrome, bipolar disorder, and schizophrenia that strong genetic influences have been clearly established.

Research into the heritability of disorders such as depression, anxiety, and disruptive behavior problems have produced more conflicting results. Where genetic influences have been implicated for these disorders, it appears that heritability relates to a genetic predisposition to develop a range of psychological problems, rather than the heritability of a specific disorder such as anxiety or depression. Furthermore, where strong heritability influences have been found, it is not clear what is inherited. For example, it is possible that children inherit a particular temperament style or other intrinsic characteristics (e.g., learning disability) that predisposes them to develop a range of psychological problems.

In addition to genetic influences, there are a range of biological factors that are associated with increased prevalence of child psychopathology. These include prenatal, birth, and early childhood influences, such as poor prenatal health care (including smoking and drug–alcohol abuse), preterm delivery, birth delivery complications, and toxins. Again the mechanism of action of these biological variables is unclear. It is possible that these biological factors impact upon child psychopathology through their role in the development of learning disabilities, or irritable, difficult temperament style. Alternatively, it is possible that the crucial variables associated with the development of child psychopathology are correlates of the biological factors, rather than biological factors *per se*. For example, it is feasible that parents who engage in high-risk prenatal behaviors, and who are therefore at greater risk of premature births, are more likely to be those parents who subsequently show poor parenting skills. Thus, the causal factor in the development of child psychopathology could be poor parenting skills rather than premature birth. Researchers are still in the process of unraveling these complex causal relationships.

Although the evidence indicates genetic and biological influences upon many forms of child psychopathology, it is important to note that many children with a strong family history of mental disorder or exposure to biological risk factors do not proceed to develop psychopathology. Similarly, many children without a family history of psychopathology or exposure to biological trauma do develop emotional and behavioral disorders. Thus, we need to examine the other characteristics of family environments and individuals that may increase or decrease the risk of developing child psychopathology.

5.12.3.2 Family Environment

There are many family environment factors that increase the chance of developing emotional or behavioral problems. Many of these factors appear to increase the chance of developing a wide variety of psychological disorders, rather than any one form of psychopathology. This may explain the high level of comorbidity between behavioral disorders of children and adolescents. For example, one of the most frequently noted correlates of many forms of child psychopathology is socioeconomic disadvantage. However, in trying to identify risk factors, it is often difficult to determine whether a particular variable has a direct effect upon the development of psychological problems or is simply correlated with some other variable that impacts upon psychopathology. We know that socioeconomic disadvantage is associated with many factors such as poor housing conditions, poor educational resources, unemployment, single-parent status, increased financial stress, parental psychopathology (particularly depression and drug–alcohol abuse), residential areas with increased crime rates and impoverished social and recreational facilities. These variables are, in turn, all associated with elevated rates of child psychopathology (Dodge, Pettit, & Bates, 1994). Efforts to control for these variables in longitudinal research suggest that economic disadvantage, unemployment and single–parent status while being correlates of child psychopathology may not have direct effects. Rather, evidence suggests that factors such as parental psychopathology and poor parenting skills are likely to be the mechanism through which societal factors have their influence upon emotional and behavioral problems of childhood (Dodge et al., 1994; Patterson, 1996). Thus, many children from poor, unemployed, or single-parent families may emerge as relatively well-adjusted young people if the parents are mentally healthy and have strong parenting skills.

Parental anxiety, depression, and alcohol abuse have been shown to be associated with a range of child and adolescent mental health problems (Rutter et al., 1990). Typically, the association between these forms of parent disorder and child psychopathology is not limited to a specific disorder across generations. Rather, parental anxiety, depression, and substance abuse is linked to a range of psychological problems in the offspring. It is important, therefore, that we determine the mechanisms through which parental psychopathology has its action. There is a complex interrelationship between the factors of parental psychopathology, marital discord, lack of partner support, divorce and separation, single-parent status, increased exposure to life stressors (e.g., financial hardship), poor parenting skills, and negative patterns of communication within the family. All these variables have been identified as risk factors for the development of child psychopathology (Mrazek & Haggerty, 1994; Spence, 1996a). It is difficult to tease apart these interrelationships. Parental psychopathology is likely to increase marital difficulties, and similarly marital problems may increase disorders such as depression, anxiety, and substance abuse. Effective parenting is likely to decline when parents experience emotional problems, and the rate of negative communications between parents and children is likely to increase (Dadds, in press). Marital conflict may lead to divorce and separation, and the financial and emotional stress that frequently accompanies single-parent status. Such stressors place an additional strain upon effective parenting, and may also exacerbate parental psychopathology. Overall, a complex interplay between risk factors operates to increase the chance that a child will develop psychological difficulties. Evidence suggests that the crucial variables in this network are the parenting skills and communication patterns between parent and child which play a mediating role in determining the influence of parental psychopathology upon child behavior problems (Harnish et al., 1995).

The variables of poor parenting skills and negative family communication patterns have been associated with a range of child and adolescent problems, including conduct disorder (Dadds, in press), depression (Kaslow & Racusin, 1994), attention deficit disorder (Barkley, 1989), and anxiety disorders (Dadds, Barrett, & Rapee, 1996). For some families, poor parenting skills and negative communication patterns may reflect long-term skill deficits, rather than a response to life circumstances. In either case, preventive interventions should ensure the development and use of effective parenting skills and positive family communication patterns. It is important to note, however, that the direction of causality is not always a simple one in which aversive parental communication patterns cause problematic child behavior. There is some evidence that the relationship may be a reciprocal one in which difficult child behavior may trigger negative parenting responses (Barkley, 1989). Thus we need to take into account the reciprocal nature of influences between parent and child behavior.

Much of the research relating to parental behavior and child psychopathology has focused upon the development of conduct disorder. However, recent studies have also demonstrated that parental behavior is probably an important contributing factor in the development of children's anxiety and depression (Sanders, Dadds, & Barrett, 1995). To date, these studies have been cross-sectional or have involved retrospective reports from anxious or depressed adults about their parent's behavior. Thus, it is not possible to draw firm conclusions about the direction of causality. However, the evidence points towards a significant role of parental factors in the development of childhood anxiety and depression. Rapee (1997), in a recent review of the literature, concluded that anxiety in children was associated with overcontrolling/overprotective parenting styles which inhibit the development of problem-solving skills in children and limits their developing sense of self efficacy. Faced with stressful or novel situations, these children are then more likely to experience anxiety. In contrast, Rapee (1997) concluded that depression in children is associated with excessively critical and negative communications from parents.

5.12.3.3 Negative Life Events

There have been many studies investigating the impact of negative and traumatic life events in children. The results have tended to produce conflicting findings. Generally, the evidence suggests that higher rates of emotional and behavioral difficulties are associated with exposure to major negative life events (Goodyer & Altham, 1991). However, many young people who experience major trauma, such as death of a parent, family break up, war, and disasters, do not go on to develop clinically significant psychopathology. For example, children vary markedly in their adjustment following parental divorce, with many children coping well with the transition (Amata & Keith, 1991; Forehand, 1992). In contrast, some children experience considerable adjustment problems, particularly if the postdivorce environment is characterized by ongoing fighting between parents (Hess & Camara, 1979). Thus, negative life events clearly do not, on their own, provide an adequate explanation for the development of child psychopathology. Rather, it appears that the impact of negative life events depends on the characteristics of the child and the additional impact of other negative environmental factors. Furthermore, we need to understand more about the mechanisms and processes through which negative life events influence child psychopathology. Factors which appear to reduce the negative impact of aversive life events on children include social support and problem-focused coping strategies of the child (Compas, 1987). We also need to consider the impact of negative life events, such as parental divorce or death of a parent, upon the parenting style of the major caretaker. Traumatic life events may set in place a sequence of changes which culminate in ineffective parenting behavior and impaired family communication patterns. Thus, the relationship between negative life events and child psychopathology may be mediated by parenting behavior of the child's caretaker. There is some evidence to support this proposition from studies of children who experience significant trauma. McFarlane (1987) demonstrated that a strong predictor of children's adjustment following a severe bushfire was the mother's level of overprotective parenting.

In some instances, factors may interact to exacerbate or trigger the negative impact of adverse life events. For example Rutter (1987) described the interaction effect of maternal loss combined with other major, negative life events. Maternal loss on its own did not produce an increased risk of psychopathology, but when combined with some other major negative life event, the combined factors produced a threefold increase in the chance of developing depression, far exceeding the impact of major negative life events alone. These results also demonstrated the delayed impact of maternal loss, in that this factor created vulnerability to psychiatric disorder only later in life and only when combined with other direct risk variables.

5.12.3.4 Intrinsic Child Characteristics

In addition to the environmental risk factors of child psychopathology, it is important to consider the intrinsic characteristics of the child that place him or her at greater risk (Coie et al., 1993). Child temperament and intellectual functioning have been shown in many studies to be risk factors for the development of child psychopathology. An early childhood temperament characterized by irritability, high levels of crying, irregular sleep, and eating problems is predictive of externalizing behavior problems in later childhood (Prior, 1992). Temperament influences can also be identified for internalizing disorders. The work of Kagan and colleagues (Kagan, Reznick, & Gibbons, 1989; Kagan & Snidman, 1991) has been influential in identifying a relatively stable temperament style which they term "behavioral inhibition." The characteristic features of behavioral inhibition include initial timidity, shyness, withdrawal, distress, and emotional restraint when exposed to unfamiliar people, places, or contexts (Asendorpf, 1993). This behavioral pattern is associated with elevated physiological indices of arousal and has been shown to have a strong genetic component (Plomin & Stocker, 1989). Children who show a stable temperament style of behavioral inhibition are more likely to develop anxious behavior and anxiety disorders during childhood (Biederman et al., 1993; Hirshfeld et al., 1992; Rosenbaum et al., 1993).

Although the evidence points to a significant contribution of early childhood temperament to the development of psychopathology, Prior (1992) pointed out that not all children who show these temperament patterns will progress to show behavioral difficulties later in childhood. Thus, we need to explain how childhood temperament interacts with other personal and environmental characteristics to determine the occurrence of psychopathology. In particular, it is important to identify those parenting patterns which buffer the impact of a difficult early childhood temperament and those parenting factors which may exacerbate the influence of high-risk temperament characteristics.

Low IQ and learning difficulties have also been linked to the greater prevalence of a wide range of emotional and behavioral problems in children, including conduct disorder, attention deficit–hyperactivity disorder, and depression (Farrington, 1995). The mechanisms by which low IQ and learning difficulties have an impact upon psychopathology is unclear. It is possible that school failure and negative school experiences set in train a trajectory towards conduct problems, low self-esteem, and depression. Alternatively, the biological factors that underlie low IQ and learning difficulties could also be responsible for attention problems, impulsivity and "difficult" temperament for some children. A further possible explanation could lie in heritable characteristics through which the parents of children who have low IQ and learning problems may have restricted cognitive abilities and consequently poor parenting skills. Much more research is needed to determine the mechanisms by which low IQ and learning difficulties impact upon child psychopathology. Meanwhile, these variables should be regarded as risk factors towards which preventive interventions should be directed.

Children's cognitive style is a further intrinsic child characteristic that has been found to predict later psychopathology, with particular patterns and styles being associated with specific forms of psychopathology. For example, a pessimistic cognitive style for interpreting information has been shown to predict later depression in children (Jaycox, Reivich, Gillham, & Seligman, 1994). Although longitudinal studies have yet to be completed, cross-sectional studies have shown an association between a cognitive style characterized by increased attention to and interpretation of threat stimuli among clinically anxious children (Kendall, 1991). Similarly, aggressive children are characterized by a bias towards processing information in a confronting and threatening manner (Lochman & Dodge, 1994).

Early childhood behaviors that represent mild symptomatology have also been found to significantly predict later clinical levels of psychopathology. For example, many studies have shown that early aggressive behavior is the best single predictor of risk for later aggression and delinquency (Guerra, Huesmann, Tolan, & Van Acker, 1995; Loeber, 1990). Therefore, within the "high-risk" category of young people to whom preventive interventions could be directed, we must include children who show early behavior problems.

To summarize, there are clearly a wide range of environmental and intrinsic child characteristics that interact as risk factors to increase the chance of developing emotional and behavioral problems. The challenge for researchers is how to identify clusters of risk factors that collectively increase the chance of the development of psychological problems. Kazdin and Kagan (1994) pointed out that individual risk factors may have only a minimal effect on their own, but may have a multiplicative, interactive influence upon psychopathology. It is also important to determine why some individuals who experience particular risk factors do not progress to show psychological difficulties. For example, not all children who manifest a stable temperament of behavioral inhibition proceed to develop an anxiety disorder. Similarly, not all children who show a "difficult" irritable early temperament style go on to show disruptive behavior problems. Thus, it is likely that certain childhood temperaments interact with particular parenting practices to increase the chance of psychological difficulties. In the case of children with a temperament style of behavioral inhibition, it may be only those children whose parents are overprotective and overcontrolling who go on to show anxiety problems. Those initially behaviorally inhibited youngsters whose parents encourage them to experience new situations, and who model a calm, problem-focused coping style, may be less likely to show anxious, fearful behavior as they grow older (Spencer, 1996b). Therefore, future research needs to clarify the way in which risk factors interact to elevate the risk of developing child psychopathology.

5.12.4 PROTECTIVE FACTORS

Although we now have a great deal of knowledge concerning the risk factors for the development of psychopathology, many of these factors are extremely hard to change. High levels of family stress, bereavements, family break-ups, parental psychopathology, single-parent status, and socioeconomic disadvantage, while important targets for reduction, are likely to continue to occur with relatively high frequency in our communities. Thus, it is important to identify those factors that protect children from the adverse consequences of such risk factors. We do know that many children emerge from high-risk environments without showing emotional and behavioral disorders. However, only recently have researchers started to investigate systematically those protective influences that produce a resilience against adverse risk factors. Moreover, Rutter (1987) discussed confusion within the literature regarding what is meant by protective or resilience factors. He pointed out that many of the variables that had been proposed as protective factors merely represented the opposite end of continuum of a risk factor. For example, generally a high level of parenting skills, high intellectual functioning, and high socioeconomic status have been found to be associated with low levels of risk. Rutter proposed that protective factors do not refer to these opposite poles of risk factors, but rather they relate to variables that reduce the impact of a risk factor upon the development of psychopathology, instead of or in addition to having a direct effect of its own. Protective factors also include events which produce a turning point in a trajectory in the development of psychopathology towards a more adaptive long-term outcome.

Although the research is in its infancy, we can make some tentative suggestions about protective factors. Within the family, secure and harmonious parent–child relationships, social support from elders and peers, and a structured, coherent family environment have been shown to buffer the negative impact of social adversity (Rutter, 1987). Parental supervision and regulation of childrens' peer group activities outside the home have also been identified as factors that reduce the risk of delinquency for children reared in a high-risk environment (Patterson & Stouthamer-Loeber, 1984).

There are also intrinsic child characteristics that appear to serve a protective function against the development of a wide variety of emotional and behavioral problems. These include positive self-concept and a range of skills (problem solving, planning, skills to remove oneself from the risk environment, social skills), and academic success (Rutter, 1987; Werner, 1987). In a recent study, O'Donnell, Hawkins, and Abbott (1995) identified several factors that appeared to inhibit the risk of subsequent involvement in serious delinquency in a sample of boys who were identified as aggressive by teacher report at ages 10–11 years. Measures were taken at ages 12–13 years of prosocial skills, school achievement and bonding, norms against substance abuse, and interaction with adults and peers who were involved in antisocial behavior. These factors significantly discriminated between boys who were and were not involved in serious delinquent behavior one year later.

If we are able to identify protective factors, it is important that we examine the processes through which these variables modify the impact of risk factors. Again, the research is in its infancy, but we can speculate about likely mechanisms. For example, Rutter (1987) reported that children with "easy" temperaments are less likely to be scapegoated or become the targets for parental criticism and negative communication, thereby reducing the negative impact of risk factors such as parental depression or marital discord. Similarly, adolescents with strong planning skills were less likely to have teenage pregnancies nor to enter into marriage with a dysfunctional partner, thereby reducing the negative impact of early institutional care (Rutter, 1987). We can also speculate that children with good social skills are more likely to develop positive relationships with supportive parents, teachers, and peers, providing a buffer against a range of risk factors.

5.12.5 IMPLICATIONS FOR PREVENTION

Our knowledge regarding risk and protective factors for the development of psychopathology provides an indication as to the variables that should be targeted in preventive interventions. Effective prevention of child psychopathology requires not only that we are aware of risk and protective factors, but that effective methods are available for producing change in these variables. The evidence regarding risk and protective factors suggests that there are some variables that impact upon the development of a wide range of emotional and behavioral problems. However, many of these factors are difficult to change. For example, current technology does not permit alteration of genetic influences. Similarly, many children in our society will continue to be exposed to situations that place them at increased risk of psychopathology, such as parental mental illness, divorce, abuse, marital conflict, death of a parent, and other negative life events. Although an important goal for prevention is to reduce children's exposure to such high-risk situations, inevitably a significant number of children will still experience these adverse life circumstances. For these children, the aim of prevention must be to reduce the impact of risk factors by increasing childrens' coping skills and enhancing parenting skills.

There are many practical difficulties that lie in the way of the implementation of preventive programs. Frequently the families who could most benefit from preventive interventions are those who are most difficult to access in terms of willingness and ability to participate in the program. The children most at risk are likely to be living within impoverished life circumstances or within families who are experiencing significant stress (e.g., parental psychopathology, bereavements, single-parent status, and family break-up). The evidence reviewed above suggests that two crucial factors that mediate the impact of environmental risk factors upon many forms of child psychopathology are parenting skills and parent–child communication patterns. Thus, it makes sense to suggest that training in parenting and family communication skills should be fundamental components of effective prevention. One of the goals of preventive programs must therefore be to identify ways of engaging parents from high-risk groups. The evidence also suggests that there are a range of intrinsic child variables that may mediate or moderate the impact of risk factors in the development of child psychopathology. A further approach to prevention may therefore be to focus on the development of protective factors in children. For example, programs could focus on the teaching of problem solving, planning, social, coping, and academic skills, all of which have been identified as protective factors. Similarly, it may be feasible to develop positive, supportive social networks with other key adults in the child's life, as a further protective factor.

The evidence regarding risk factors also suggested that the probability of developing psychological problems increases markedly as the number of risk factors increased. Indeed, there is some indication that this is not simply an additive effect, but rather the probability of developing psychopathology increases multiplicatively with increasing numbers of risk factors (Coie et al., 1993). This suggests that children most at risk will be those who experience multiple risk factors. Policymakers and mental health planners need to decide where best to invest funds in preventive programs. So far, there are many unanswered questions.

(i) Is it more cost-effective to focus preventive efforts on the few children with multiple risk factors, or upon a larger number of children with only one or two risk factors?

(ii) For children with multiple risk factors, are intensive interventions that tackle multiple etiologies required in order to prevent psychopathology, or can prevention with these children be achieved by tackling only one or two risk or preventive factors?

(iii) Should children be singled out for prevention programs through some process of assessment or identification, or should prevention be applied to whole populations in order to avoid specific targeting and potential effects of labeling?

(iv) Should preventive funds be invested in efforts to minimize one single disorder or, given the many common etiologies across disorders, into broad-spectrum projects that aim to prevent multiple forms of psychopathology?

(v) Are prevention programs best geared towards efforts to minimize environmental risk factors, or to providing children with the skills to cope with adverse life circumstances?

(vi) To which age group should preventive efforts be directed; to older children who are more able to understand the concepts being taught, or to younger children who have been exposed to risk factors for a shorter period of time?

Certainly, there are more questions than answers.

5.12.6 THE EFFECTIVENESS OF PREVENTIVE PROGRAMS

Prevention programs come in many forms and evaluative studies have suffered from a wide range of methodological limitations, making it difficult to draw general conclusions about their effectiveness. Some programs have taken a broad population approach, whereas others have targeted a specific high-risk group. The goal of prevention for some interventions has been general enhancement of mental health, whereas others have aimed to prevent some specific disorder. It is beyond the scope of this chapter to provide extensive descriptions of the many preventive interventions that have been reported in the literature and interested readers are referred to Mrazek and Haggerty (1994) for a more detailed review. What follows is a summary of some of the major preventive studies, as a means of illustrating some of the key points within the prevention area. In keeping with current methods of categorization, the following review will divide preventive interventions into three main types: universal, selective, and indicated approaches. The literature describing preventive interventions varies markedly in terms of the methodological rigor of evaluation methods. Some studies are simply descriptive or report case illustrations. Others have used quasi-experimental designs, with weak control procedures and/or lack of random assignment to conditions. We also need to bear in mind that selective publishing is likely to occur, with studies that produce positive, significant findings possibly being more likely to reach publication status. Nevertheless, there is convincing evidence within the literature to suggest that some interventions have been effective in preventing, or at least attenuating, the development of child psychopathology.

5.12.6.1 Universal Preventive Interventions

Universal interventions are those that are applied to whole populations. In some instances, universal preventive interventions have been designed to enhance general mental health, whereas others have targeted one specific disorder such as conduct disorders, anxiety, depression, or substance abuse. The majority of universal preventive projects have been applied within school settings. For example, Kellam and Rebok (1992) described two classroom-based interventions that were designed to enhance children's reading skills and prosocial behavior as a means of preventing subsequent conduct and substance abuse problems. Preparatory research had identified reciprocal relationships between children's academic achievement and shy, depressive, and aggressive behavior. The first intervention involved a Mastery Learning curriculum for the enhancement of reading skills. The second intervention was based on the Good Behavior Game that consisted of a team-based, behavior management approach targeted at reducing aggressive behavior. In trials across 19 schools in Baltimore, Kellam, Rebok, Ialongo, and Mayer (1994) reported that the Good Behavior Game had a short-term impact upon aggression, but no direct or indirect impact upon academic achievement. In contrast, the Mastery Learning approach had a direct effect upon enhanced academic achievement, but also produced an indirect reduction in aggressive behavior. The longer term results from first grade intervention upon behavior at sixth grade revealed interesting results. Although there were no overall significant benefits of the Good Behavior Game in comparison to control children, a differential response was found for different groups of children. For males, those who showed the highest rates of aggression in first grade showed the greatest benefit from the Good Behavior Game at sixth grade follow-up. However, the approach was not effective in preventing children who were not aggressive in first grade from subsequently developing problems of aggression. These findings suggest that classroom-based behavior management approaches may offer promise in the prevention of aggressive behavior in children who manifest early disruptive behavior problems.

Reductions in conduct disorder and substance abuse are also goals of the Seattle Social Development Project. This intervention is based on the theoretical premise that strong bonds to family and school serve as protective factors against delinquency and drug abuse. Increases in feelings of attachment, commitment, and beliefs regarding participation in undesirable behavior are considered to be important aspects of the protective process. The project involves a multicomponent, universal preventive intervention that includes teacher training in classroom management and instructional methods, class-based interpersonal problem solving skills, training sessions, and a parent training component (Hawkins et al. 1992). The effects of the program were evaluated in schools within high crime areas. Those children who received the intervention were less likely to have initiated delinquent behavior and alcohol use, and showed higher school achievement four years after participating in the project, in comparison to children from schools or classrooms that did not receive the intervention.

Not all universal preventive studies have focused on conduct disorder as the outcome goal. In some instances the aim has been to prevent the onset of internalizing problems such as depression. For example, Clarke, Hawkins, Murphy, and Sheeber (1993) described two very brief preventive interventions for adolescent depression, neither of which produced significant benefits. Both interventions were conducted with a large number of youngsters within the context of regular classroom education. The first intervention involved three, 50 minute sessions of educational intervention, over three consecutive days, covering the nature of depression, its treatment, the benefits of seeking treatment when needed, and the need to increase pleasant events. No specific behavioral skills were trained. The intervention group was compared with a group of youngsters attending the usual health education class. The intervention had little impact on depressive symptoms for girls, and boys showed a small short-term improvement which was not maintained at 12 week follow-up. In the second intervention, Clarke et al. (1993) altered the program so that it involved five, 50 minute sessions. The first session included a shortened educational component covering the symptoms, causes, and treatments of depression. The remaining four sessions involved a behavioral program to increase participation in pleasant events. No significant benefits were found for either the sample as a whole, or for those adolescents with initial high levels of depression.

The poor results of these programs may make us pessimistic about the value of universal interventions in the prevention of depression among adolescents. However, such a conclusion may be premature. Clarke et al. (1993) criticized their own programs for failing to include a skills training component to rectify those skills deficits (e.g., social skills deficits) associated with depression in young people. In addition, the studies were excessively brief and the follow-up period of 12 weeks was too short to detect long-term benefits. Clarke et al. (1993) proposed that future prevention studies should focus on those factors that have been demonstrated to be risk factors for the development of depression, such as poor social and interpersonal problem-solving skills, overly critical self-evaluation, low rates of self-reinforcement, and pessimistic thinking style.

Drug and alcohol abuse have been the targets of several preventive interventions. Price (1995) provided a valuable outline of the issues facing researchers and practitioners working in this area. Price discussed a variety of approaches to prevent substance abuse, including school-based, family-focused, and media-based programs. The prevention of substance abuse highlights the many different perspectives that can be taken. For example, substance use is influenced by economic factors such as the availability and price of drugs and alcohol. One approach to reducing substance abuse is therefore to manipulate the supply and cost of substances. Of course, this action could also produce negative side effects relating to criminal activities that may act as a means of obtaining money to purchase drugs. Thus, no simple solutions are presented. Although we are still a long way from finding solutions to the prevention of substance abuse in young people, preliminary studies of universal interventions are emerging that provide some encouraging findings relating to the prevention of early stage substance use, such as tobacco smoking (e.g. Caplan et al., 1992; Ellickson & Bell, 1990; Perry et al., 1989).

5.12.6.2 Selective Preventive Interventions

Selective prevention efforts are applied to those individuals who are members of a group, the membership of which places them at increased risk for the development of mental disorders. Over the past 20 years preventive interventions have been applied to a wide range of high-risk groups, focusing on parents and families who are dealing with major life transitions and/or high-risk situations. These include groups characterized by parental psychopathology, premature births, single parents, socioeconomic deprivation, teenage pregnancies, divorce, bereavement of a close family member, exposure to trauma and school transitions.

5.12.6.2.1 Children exposed to socioeconomic deprivation

Several preventive programs have been targeted towards children from high-risk geographical areas with elevated levels of poverty and societal risk factors. In some instances these studies have used random assignment and control groups in order to determine the long-term efficacy of the interventions. For example, Olds, Henderson, Tatelbaum, and Chamberlin (1988) described the Prenatal/Early Infancy Project which aimed to reduce mothers' health-damaging prenatal behaviors (i.e., diet and substance abuse), enhance parenting skills, and provide social support, assistance with continuing education, and contraceptive counseling. Women who received home nurse visitation during pregnancy and until the child was two years old showed significant reductions in preterm births and verified cases of child abuse in comparison to women who did not receive such an active intervention. A home visiting approach was also beneficial in the Early Intervention for Preterm Infants project (Field, Widmayer, Stringer, & Ingatoff, 1980) in enhancing physical and cognitive development among preterm babies born to young, single mothers from low socioeconomic backgrounds. This program focused on teaching mothers about caretaking practices and enhancing parent–infant interactions. Home visiting plus preschool developmental activities have also produced positive results in terms of physical and cognitive development of low birth weight babies (Kraemer & Fendt, 1990).

Some programs have included a wider spectrum of intervention approaches, of which home visiting is just one component. Lally, Mangione, Honig, and Wittner (1988), in a quasi-experimental design, demonstrated that a home visiting approach, in association with child care and early educational enrichment was associated with significant reductions in delinquent behavior at 10 year follow-up. This project involved primarily low-income, single mothers who participated in the intervention from the last trimester of pregnancy through to age five of the child. Home visiting is also an important focus within the recently established Better Beginnings, Better Futures Project in Ontario, Canada, that is aimed at families in socioeconomically disadvantaged areas. This project involves a series of programs, some of which include home visiting as a component. Other aspects of the project include classroom enrichment for teaching of academic and social skills, child care enrichment, parent training, parent support groups, and increasing community resources (Peters & Russell, 1994). The Healthy Start in Hawaii service is also worth mentioning here (US Advisory Board on Child Abuse and Neglect, 1990, cited in Mrazek & Haggerty, 1994). This service was not established as a research study but as a community intervention to reduce child abuse and neglect. The program is designed to identify at-risk children in the first three months of life and aims to screen around 90% of the total population. Following early identification, families are allocated a lay-person home visitor, linkage to a primary health care service, parent education, and child care services. Since the establishment of the Healthy Start program, there has been a gradual decrease in child abuse and neglect cases in Hawaii, although it must be noted that there are no empirical control groups with which to compare the current rates of abuse and neglect of those families who received the intervention.

There are many other programs that have included home visiting as part of their intervention. Other components of multifaceted interventions have included parent education classes, preschool/nursery school facilities to promote social and academic skills, and home-based social and cognitive skills training (e.g., Houston Parent–Child Development Center, Johnson, 1991; Mother–Child Program of Verbal Interaction, Levenstein, 1992). These studies demonstrated that preventive interventions can be effective in enhancing the social and cognitive skills of children from socioeconomically deprived backgrounds. One of the most well-known, preventive projects is the Perry Preschool Program that combined home visiting and intensive preschool education. This selective intervention focused on children and their families who experienced severe poverty and included daily preschool attendance for two years and weekly home visiting by a trained teacher. Long-term follow-ups showed superior academic performance and social adjustment and lower rates of delinquent behavior through to late adolescence, compared to the randomly assigned control group (Berruteta-Clement, Schweinhart, Barnett, Epstein, & Weikart, 1984; Schweinhart, 1987).

Some studies have focused preventive efforts upon school-based interventions, with minimal involvement of parents. For example, Shure and Spivack (1980, 1982, 1988) reported the benefits of an interpersonal problem-solving program that was delivered to small groups of preschool children from economically deprived families. The curriculum ran over approximately three months, and involved several lessons per week to teach problem-solving skills. Children who received the intervention were reported to show enhanced social competence and lower rates of behavior problems in comparison to untrained control children. However, whether preschool interventions of this sort will be beneficial in producing long-term improvements through to adolescence remains to be demonstrated.

In summary, there are now a wide range of evaluative studies demonstrating that preventive interventions can be effective in enhancing social adjustment and decreasing psychopathology (particularly conduct disorder) in children from socioeconomically impoverished backgrounds. However, it is important to note that many of these studies involved relatively small numbers of families and very large numbers of intervention hours. Some studies required a very large investment of time from parents which precluded inclusion of those who were in employment. Furthermore, many studies do not report the cost of the intervention per family. Clearly, the cost of providing weekly home visits for 3–5 years, in addition to daily preschool facilities with low staff–child ratios, is enormous. Thus, although these interventions can clearly produce significant benefits, it remains to be determined whether it is feasible in financial and practical terms to apply such intensive interventions on a large-scale basis. A further problem in existing studies is the relatively high drop-out rate that is often reported, particularly from families with high levels of mobility. Thus, it is important to take into account the drop-out problem in interpreting the results of existing studies, and that efforts are put in place to maximize retention rates in future interventions.

5.12.6.2.2 Assisting children to deal with school transition

Children are required to deal with many changes throughout their lives, but the transition to a new school is suggested to be one of the most stressful (Soussignan, Koch, & Montagner, 1988). Children are required to adapt to different physical, social, and academic environments, with new sets of rules and regulations. Not surprisingly, many children find the transition very difficult and a variety of behavior problems have been linked with school transitions. These include peer relationship difficulties, school refusal, somatic complaints, academic failure, increased substance abuse, delinquency, and school drop-out (Hightower & Braden, 1991).

As a means of preventing adjustment problems associated with transition to secondary school, Felner and Adan (1988) developed the School Transition Environment Project (STEP). The STEP program consists of three components. The first organizes the physical plan of the school into units with home rooms in order to facilitate familiarity with the school environment. Children allocated to the program attend core academic subjects together in these home rooms. This process is suggested to make the school transition less overwhelming and stressful to incoming students. The second aspect of STEP includes a "home room" staff member who has the responsibility for taking the daily attendance list, following up non-attendances, and counseling of pupils regarding academic or school adjustment problems. The final component involves a coordinated liaison between teaching and school counseling staff. The overall aims of the program are to increase personal relationships between pupils and staff and to create subenvironments within the overall large school environment.

Evaluations of the STEP approach have demonstrated that children who participated in the program showed greater improvements in academic performance and self-esteem, better school attendance, and lower school drop-out rates compared to control children who did not take part (Felner & Adan, 1988). Furthermore, the benefits have been replicated in a variety of different school settings with children from various social backgrounds. Overall, the approach appears to be a valuable method of reducing problems associated with school transition and provides a model which could be routinely adopted by school authorities.

5.12.6.2.3 Children of parents with psychiatric disorders

Given the convincing evidence that parental psychopathology is a risk factor for the development of psychological problems among children and adolescents, it is not surprising that researchers and practitioners are beginning to examine the value of applying preventive interventions with children of mentally ill parents. To date, however, there are few controlled evaluations to determine the effectiveness of such efforts. There is a strong theoretical case to suggest that preventive efforts should be directed towards the children of parents who experience chronic problems of depression, anxiety or psychosis. In particular, it is important that these parents are taught to use positive parenting skills and that support mechanisms are put in place to increase the chance that parents will continue to use these skills during episodes of psychological problems. However, we must also acknowledge that the children who live with a mentally ill parent may also need to learn coping and survival skills to assist them in dealing with traumatic life circumstances and phases of poor parenting. In one of the few experimental evaluations of a prevention program with children of severely depressed parents, Beardslee, Salt, Porterfield, and Rothberg (1993) reported positive results for two approaches. The first intervention involved a clinician-led psychoeducational method, whereas the second approach was lecture-based. However, as the study did not include a no-intervention comparison group, it is difficult to draw firm conclusions about the benefits of the interventions. Clearly, there is an urgent need for methodologically rigorous examinations of preventive efforts with children of mentally ill parents.

5.12.6.2.4 Children who have experienced trauma

Some aversive events occur unexpectedly and are of sufficient magnitude that they constitute a traumatic experience to children. Such traumas include environmental disasters (e.g., earthquakes, war, storms, or fires), serious accidents, and death of a close family member. Family break-up through parental separations and divorce could also be regarded as a traumatic event, although this area is covered separately below. There is now considerable evidence that children who experience traumatic events are at increased risk of developing emotional and behavioral problems. For example, following trauma, children show increased levels of fears relating to stimuli associated with the traumatic event, avoidance behaviors, somatic complaints, depression, sleep disturbance, and intrusive thoughts and images (Dollinger, 1986; Dollinger, O'Donnell, & Staley, 1984; Terr, 1981; Yule & Williams, 1990). Although such problems ameliorate relatively quickly for most children, a significant proportion of children show persistent anxiety symptoms for many months after the trauma (Terr, 1981). There is therefore a strong case for rapid intervention following trauma in order to prevent the development of persistent psychological problems.

Although much has been written about the need for interventions to prevent long-term psychological difficulties following trauma in children, we still have very little empirical evidence to permit us to determine the most effective approaches to prevention. Sugar (1989) emphasized the need for initial interventions with the parents of children who experience a traumatic event, given the evidence that parent reaction plays a highly significant role in determining child response (McFarlane, 1987). In large-scale disasters which affect whole communities, Sugar (1989) also stressed the need for community counseling and crisis-oriented group approaches in order to provide mutual support, decrease isolation and provide a larger perspective for the individual. Sugar pointed out, however, that the cornerstone of therapy should be individual counseling with children, which takes into account their developmental level and specific details of the traumatic situation experienced. The approach taken to intervention should therefore be tailored to the needs of each child.

A variety of approaches have been used in what is commonly termed the "debriefing" of trauma victims. Yule (1991) suggested that this procedure should commence somewhere between 7 and 14 days following the event, with most survivors being too numb to benefit from counseling prior to this time. Nevertheless, Yule acknowledged that there is a lack of data to validate this proposal and we really do not know what type of intervention should best be used nor at what time after the occurrence of a traumatic event. Yule (1991) proposed a theory-driven approach to trauma counseling with children, based on Rachman's (1980) model of emotional processing which emphasized the need for vivid and prolonged exposure to disaster-related cues in order to facilitate emotional processing. Children may be trained in relaxation skills and gradually exposed to stimuli and memories relating to the trauma event. They are encouraged to describe their reactions and given reassurance that such reactions are understood and are a normal response to an abnormal experience.

To date, there is a lack of controlled experimental evaluations to enable us to identify the most effective means of preventing child psychopathology following trauma. There are many case studies and descriptive reports in the literature, but we now need carefully designed experimental evaluations.

5.12.6.2.5 Children whose parents separate or divorce

Parental divorce has been suggested to represent one of the most common and severe life stressors that confront children and adolescents (Hightower & Braden, 1991; Hodges, 1991). For some children, the emotional consequences may include depression, anger, disruptive behavior, anxiety and/or withdrawal, particularly in the first two years following the divorce or separation (Hess & Camara, 1979). Fortunately, evidence suggests that many children adapt well following family break-ups (Amata & Keith, 1991; Forehand, 1992). To a large extent, the adverse emotional consequences of parental separation and divorce are influenced by the quality of the relationships of the family members involved (Hess & Camara, 1979). Hess and Camara (1979), for example, reported that the negative effects of divorce were considerably less when positive relationships between the parents were maintained. Positive relationships between the child and both the custodial and the noncustodial parents were also found to be important in protecting the child from the negative effects of divorce. This study highlighted the importance of government and social service policies that facilitate positive relationships between ex-partners following separation and divorce. Mediation may be beneficial in reducing disputes between parents regarding custody, access, and property settlements.

We need to be able to identify those children who are most at risk of developing adverse reactions to parental divorce and separation and target preventative programs in their direction. Several programs have been developed for use with children whose parents have recently divorced (e.g., Hodges, 1991; Pedro-Carroll & Cowen, 1985). Pedro-Carroll and Cowen outlined the Children of Divorce Intervention Project (CODIP) which was developed for use on a small group basis within schools. The project aimed to prevent or ameliorate academic, behavioral, and emotional problems that children often experience during or after their parents' divorce. Components of the program included: (i) development of a supportive group environment; (ii) facilitation of the identification and expression of divorce related feelings; (iii) promotion of understanding of divorce-related concepts and rectifying misconceptions; (iv) teaching of coping skills, including social problem-solving skills; and (v) enhancement of children's positive perceptions of themselves and their families. The program involved 12–16 sessions, with methods including discussion, role-play, skills training, and home-based tasks.

Evaluations of the program demonstrated its effectiveness in producing reductions in anxiety, fewer behavioral problems, greater gains in school competencies, decreased feelings of self-blame, and better ability to solve divorce-related problems in comparison to children of divorce who did not participate in the project. Follow-up information suggested that the benefits were maintained for the majority of children at two years follow-up, although only around half the children continued to show the gains they had made over the comparison group at three year follow-up. Although independent replications using this approach are required, the CODIP program appears to offer promise in the prevention of adverse psychological consequences in children following parental divorce.

5.12.6.2.6 Children who show multiple risk factors

In view of the convincing evidence that risk factors have a multiplicative effect, not just an additive one, there is a strong case for focusing preventive efforts on young people who show a combination of risk factors. For example, Dishion and Andrews (1995) reported a study designed to prevent conduct disorder and substance abuse among "high-risk" adolescents who showed at least four risk factors. These risk factors concerned a range of factors such as poor parent–child relationships, lack of school engagement, problem behavior, early substance use, family substance use, and stressful life events. A parent management program involving 12, 90 minute, weekly sessions was beneficial in reducing family conflict and behavioral problems at school in comparison to the no-intervention group. There was also a trend towards longer-term reduction in substance use for the parent management group. Interestingly, the two comparison interventions that included grouping of high-risk youths together in sessions actually showed evidence of negative influence upon school behavior problems and substance use. This problem is discussed further below.

5.12.6.3 Indicated Preventive Interventions

As defined above, indicated prevention approaches are those that are applied to individuals who, on examination, are found to manifest a risk factor, condition, or abnormality that identifies them as being at high risk for the future development of psychopathology. The main distinction between selective and indicated prevention refers to the degree to which the risk factor is intrinsic to the individual rather than the group. Thus, selective interventions target persons who are judged to be at risk as the result of being a member of some group, the membership of which is known to be a risk factor for the development of psychopathology. In contrast, indicated prevention targets persons who possess some characteristic intrinsic to the individual that places them at risk for the development of psychopathology. Given our knowledge of early, mild symptomatology as a risk factor for later clinically significant psychopathology, programs that target mild psychopathology (alternatively known as early intervention) will be regarded as a form of indicated prevention. These approaches aim to identify early symptoms of emotional and behavioral disorders through screening programs, usually during early childhood. Such programs aim to break the trajectory typically involved in the development of many forms of psychopathology, including conduct disorder, depression, anxiety, and psychotic episodes.

5.12.6.3.1 Early intervention to prevent conduct disorder

Several programs have been developed for early identification with children who show subclinical levels of conduct disorder and aggression. A variety of approaches have been used in attempts to break the progression from early disruptive behavior problems to a full-blown conduct disorder. These have included parent training approaches, child-focused interventions, and multicomponent programs that involve a variety of parent- and child-focused methods. The rationale behind these interventions is the proposition that poor parenting skills, academic failure, and poor social skills are risk factors for the development of conduct disorder. Thus preventive interventions have focused on enhancing factors such as parenting, social, and academic skills.

For example, Tremblay et al. (1991, 1992, 1995) examined the outcome of a program designed to prevent serious conduct problems among a group of disruptive kindergarten boys from inner-city, low socioeconomic neighborhood schools. One hundred and seventy two boys, who were identified at the end of preschool as showing high levels of disruptive behavior, were randomly assigned to a preventive intervention, observational control, or no contact. The preventive intervention was conducted over a two-year period and involved parent and child interventions. The parent component involved up to 46 individualized home-based training sessions, with a mean number of 17.4 sessions focusing on monitoring of children's behavior, use of behavioral parenting methods, and crisis management techniques. The child-focused component included small group prosocial skills training, problem-solving, and self-control methods. These sessions were conducted over two school years, and included a small number of prosocial peers within each group.

Up to age 12, the boys who had earlier received the intervention showed better school attendance, were rated by their teachers as less disruptive in the classroom, and were less likely to have required special education experiences (Tremblay et al., 1992). However, the longer-term results reported by Tremblay et al. (1995) are extremely interesting. By age 15, the difference between preventive intervention and control boys had disappeared in terms of school placements and behavior in the classroom. Generally, there was a trend for behavior problems to decline with age for boys, irrespective of whether they participated in the preventive program. Juvenile court records also showed no significant differences between the intervention and control groups in terms of official court sanctions between ages 12 and 15. Furthermore, there were no differences between groups in terms of the boys' perceptions of their parents' punishment behavior, although the boys in the preventive group perceived their parents as engaging in a higher level of supervision than did the control group. Thus, despite a costly intervention over a two-year period, the long-term results of this study were disappointing. Positive benefits were noted in terms of classroom behavior and academic performance throughout elementary school, but the impact had largely disappeared by age 15. Tremblay et al. (1995) proposed that future studies should include booster sessions for boys and parents in order to maintain the benefits that are evident during elementary school years. They suggested that booster units are needed that focus on problem-solving skills, life skills, study skills, conflict resolution, self-control, and academic skills during early adolescence.

The value of the Tremblay et al. (1995) results is that they highlight the importance of extensive long-term follow-up. It will be several years before many of the recently established early intervention studies for the prevention of conduct disorders obtain adequate follow-up data to permit valid conclusions to be drawn about their long-term efficacy. For example, the FAST Track Project (Conduct Problems Research Group, 1992) is being conducted in multiple sites across the USA. Participants are preschoolers and their families where the youngsters have been identified as showing high levels of noncompliance and aggression at age four and therefore are considered to be at risk for future development of conduct disorders. The intervention continues over several years, with components focusing on the training of social skills, increasing peer acceptance, remedial educational tuition, and parent training. Parents are paid a small amount to attend the sessions and transport and child care is provided for the sessions. The long-term effectiveness of this approach is yet to be determined.

Lochman, Coie, Underwood, and Terry (1993) reported a controlled evaluation of a program targeted at fourth grade children who were identified as socially rejected and aggressive at school. The intervention included social skills training, anger management, and cognitively-based components to teach interpersonal problem solving and appropriate perception and interpretation of interpersonal cues. The school-based curriculum included small group and individual sessions over 12–18 sessions. By one year follow-up, there was a problem of drop out which restricts the conclusions that can be drawn. However, of those children who were both aggressive and rejected and who remained in the study at one year follow-up, those who received the intervention were rated by their teachers as being significantly less aggressive than those in the randomly assigned control condition. Whether these benefits are maintained at 5–10 year follow-ups remains to be demonstrated.

An alternative approach to early intervention with children who show disruptive and aggressive behavior during preschool years is the Triple P (Positive Parenting of Preschoolers) program described by Sanders and Markie-Dadds (1996). This project involves a multilevel approach in which the intensity of the intervention is tailored to the severity of the presenting problems of the child and family. Level 1 involves a community-based self-help approach involving tip sheets for parents and a manualized program concerning parenting skills for commonly occurring behavior problems. The Level 2 intervention includes minimal contact with a therapist by telephone in addition to the written self-help materials. Active training in parenting skills is then included at Levels 3 and 4, being applied to those families for whom the child behavior problems are not complicated by additional family problems such as marital discord or parental depression. In Level 5, intensive behavioral parent training is enhanced with treatments to reduce parental depression and marital discord. Short-term results suggest that this approach appears to offer promise in reducing subsequent conduct problems. However, it is not yet clear how effective the multilevel intervention strategy will be in producing long-term benefits through to adolescence, nor when applied on a large-scale, community basis.

5.12.6.3.2 Early intervention to prevent anxiety and depression

Recently, researchers have started to examine the benefits of preventive interventions with children who show mild depressive symptoms, again trying to break the trajectory in the development of clinical depression in later adolescence. Clarke et al. (1995) reported the results of a preventive program with high school adolescents who were assessed as showing elevated depressive symptomatology. These youngsters were randomly assigned to a 15-session cognitive group intervention or a "usual care" control condition. Survival analyses indicated a significant 12-month advantage for the prevention program, with affective disorder total incidence rates of 14.5% for the active intervention and 25.7% for the control condition. Jaycox, Reivich, Gillham, and Seligman (1994) also reported positive findings in the prevention of depression of primary school children who reported mild symptoms of depression and parental conflict in the home. The intervention program involved teaching children to shift their cognitive style to more optimistic, rather than pessimistic, interpretations and training in interpersonal skills. At six-month follow-up, the preventive intervention group reported fewer symptoms of depression compared to a nonintervention control group, an effect that was still evident at two-year follow-up (Gillham, Reivich, Jaycox, & Seligman, 1995).

An early identification and intervention approach is also being trialed in the prevention of anxiety disorders in children and adolescents. Dadds, Spence, Holland, Barrett, and Laurens (1997) described an intervention for children who were identified as being at risk for the development of anxiety disorders by virtue of showing mild levels of anxiety and fearfulness. The aim of the project is to reduce symptoms of anxiety and enhance protective factors, such as child coping skills and appropriate parenting skills, in order to reduce the later development of anxiety disorders. After recruitment and diagnostic interviews with 1786 7–14 year olds, 128 children were selected and assigned to a 10-week school-based child and parent-focused psychosocial intervention, or a monitoring group.

At postintervention, approximately 10% of children in both groups had developed a full *DSM-IV* anxiety diagnosis, with no significant differences between groups. However, at six months follow-up, 54% of the children in the monitoring group had developed an anxiety disorder, compared with 16% in the intervention group (Dadds et al., 1997). The finding that over half of the children in the monitoring group who were "at risk" only, progressed into a formal anxiety disorder at the six month monitoring period, highlights the importance of late childhood and early adolescence as a critical time in the development of anxiety disorders (Spence & Dadds, 1996). The long-term benefits of this preventive intervention are still being examined.

To date, research into the prevention of anxiety and depressive disorders in children who manifest early symptomatology in these areas is only in an embryonic state. However, the results so far are encouraging and justify continued investigation of methods to interrupt the trajectory from preclinical symptomatology to clinical depression and anxiety.

5.12.6.3.3 *Early intervention to prevent psychotic episodes*

One of the most exciting developments in the area of indicated prevention concerns early intervention with individuals who show early signs of psychosis. For example, the Early Psychosis Prevention and Intervention Centre (Jackson, McGorry, Edwards, & Hulbert, 1996) provides early intervention for 16–30 year olds with an emerging psychosis. A cognitively-oriented psychotherapy approach is taken in the multiphase program. Intervention methods include development of a sense of hope, identifying personal strengths, developing coping mechanisms, education about psychosis, and establishment of social support networks. Preliminary results suggest that participants with first episode psychosis show lower negative symptoms and reduced disability scores compared to those who refused to participate or dropped out of the program. Although the lack of random assignment to groups represents a methodological limitation, the preliminary results suggest that early detection and intervention programs for psychosis warrant further empirical examination. The program is also limited to individuals who have already experienced their first episode of psychosis. Some interesting research reported by Birchwood (1995) suggested that prodromal signs (such as loss of mental agility, racing thoughts, disjointed thinking and confusion, feelings of derealization and heightened perception) can reliably predict the imminent onset of new episodes of psychosis. Birchwood used case studies to illustrate how cognitive–behavioral interventions and/or medication can be used to prevent subsequent episodes of psychosis in young people who manifest these prodromal signs. The programs described by Birchwood (1995) and Jackson et al. (1996) involved individuals who have already manifested their first episode of psychosis. Monitoring and identification of prodromal signs is used to cue the use of techniques to prevent the onset of subsequent episodes of psychosis. However, there is also some tentative evidence that it may be possible to prevent first episode by early detection of presymptoms of psychosis and intervention with adolescents and young adults (e.g. Falloon, 1992).

5.12.7 DIFFICULTIES AND CHALLENGES

There are an enormous number of practical difficulties in implementing preventive interventions and in evaluating their efficacy. The following section attempts to outline just a few of these issues for consideration. For example, it is important to bear in mind the possibility that preventive efforts may, in some instances, be counterproductive. We need to consider the possible detrimental effects of labeling individuals as being "at risk" of developing psychopathology. Any preventive efforts must ensure that the intervention does not produce changes in the behavior of significant others, such as peers, teachers, or family members that further increases the chance that the young person will develop emotional or behavioral difficulties. Similarly, preventive efforts must be cautious about the potential negative side effects of bringing together children who share common risk patterns. This is particularly relevant to programs that bring together adolescents who share risk factors relating to the development of substance abuse or severe conduct disorder. In some instances, the aggregation of high-risk youths has produced an iatrogenic effect. For example, Dishion and Andrews (1995) found that interventions that aggregated high-risk youths into groups produced escalations in tobacco use and problem behavior at school. This effect was not evident in the intervention that focused on parenting skills training and did not include group attendance by the adolescents.

A second issue for preventive interventions is to whom should preventive efforts be targeted in order to produce the most cost-effective benefits. One of the most significant areas of research for the future will be identification of combinations of risk factors that accurately predict the development of psychopathology and for which there are effective methods of change. This will enable us to target preventive efforts more specifically to those at greatest risk. Lochman et al. (1995) proposed that the most useful risk markers for a prevention program are those that are potentially modifiable as a result of intervention, or those that are markers for other changeable processes that mediate negative outcome Lochman et al. (1995) described how a multiple gating procedure can be used to identify children at risk of developing conduct problems. This study showed how parent ratings of early childhood behavior made a significant contribution over and above teacher ratings of child behavior in the prediction of subsequent conduct problems at school. Studies of this type will be valuable in steering future directions for prevention research and policy.

Although it may make sense to suggest that limited funds for prevention should be spent on those children most at risk rather than upon universal interventions, we do not have the evidence to back up this assumption. We need to determine whether universal interventions applied to whole populations can produce sufficiently strong results to impact upon the development of psychopathology among children with multiple risk factors. Given the relatively disappointing long-term effects of those intensive interventions that have been focused upon the prevention of behavior problems among children with mild symptomatology (e.g., Tremblay et al., 1995), it seems unlikely that a universal intervention could produce significant long-term benefits for these children. Nevertheless, this possibility warrants examination. An alternative approach is to follow the multilevel example outlined by Sanders and Markie-Dadds (1996). This system requires an assessment and prevention strategy that provides relatively cheap community-based initiatives that are likely to be helpful to the majority of families. More intensive approaches are reserved for those families and children who show persistent behavior problems or multiple risk factors. These multilevel preventive strategies require that systems are in place to identify those families who are in need of more complex interventions than those available at a community, universal level.

A third area of concern is the need to provide cost–benefit analyses, in terms of the cost of the preventive intervention vs. the reduction in costs of treatment of psychopathology. If mental health policymakers and planners are to be convinced that there is a case for investing large amounts of funding in preventive efforts, they need to be shown evidence of the likely economic benefits to the community. In the area of physical health it is easier to provide estimates regarding the cost savings of prevention programs. For example, it has been estimated that the cost–benefit ratio of the mumps–measles–rubella vaccine program in the USA is 14:1. For example, the 1983 immunization scheme produced cost savings of US$1.4 billion (White, Koplan, & Orenstein, 1985). The benefits of prevention in mental health are much harder to estimate. However, it should be possible to determine the cost of preventive interventions compared to the cost of long-term treatment of full-blown clinical cases through indicators such as the cost of mental health consultations, time off work, hospitalization, and pension payments.

5.12.8 CONCLUSIONS

Considerable advances have been made over the past 20 years in the area of prevention of child and adolescent psychopathology. The evidence presented in the present chapter suggests that we have made considerable progress in identifying risk and protective factors relating to the development of child and adolescent psychopathology. It is now feasible to identify major risk and protective factors for the development of conduct disorders and internalizing problems in young people and many of these are potentially modifiable through psychological interventions. There have now been many studies of universal, selective, and indicated preventive interventions that have reported positive effects in reducing the risk for development of emotional and behavioral problems. However, most of the published papers relating to the efficacy of prevention of psychopathology are limited by relatively short follow-up periods or have failed to include a valid comparison group against which to compare changes over time. It is important therefore to bear these limitations in mind when interpreting the literature. There are still many unanswered questions. A great deal more research is needed before we can draw firm conclusions about which psychological problems can be prevented, what methods are most effective in preventing specific mental health problems, to whom should preventive programs be directed, and whether prevention is best directed at whole populations, or those with specific risk factors. Many methodologically rigorous and theoretically sound preventive studies are currently underway and the next decade should provide answers to many of these questions.

5.12.9 REFERENCES

Amata, P. R., & Keith, B. (1991). Parental divorce and the well-being of children: A meta analysis. *Psychological Bulletin, 110,* 26–46.

Asendorpf, J. B. (1993). Beyond temperament: A two-factor coping model of the development of behavioral inhibition during childhood. In K. H. Rubin & J. B. Asendorpf (Eds.), *Social withdrawal, inhibition and shyness in childhood* (pp. 265–269). Hillsdale, NJ: Lawrence Erlbaum Associates.

Barkley, R. A. (1989). Hyperactive girls and boys: Stimulant drug effects on mother–child interactions. *Journal of Child Psychology and Psychiatry, 30,* 379–390.

Beardslee, W. R., Salt, P., Porterfield, K., & Rothberg, P. C. (1993). Comparison of preventive interventions for families with parental affective disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 254–263.

Berrueta-Clement, J. R., Schweinhart, L. J., Barnett, W. S., Epstein, A. S., & Weikart, D. P. (1984). *Changed lives: The effects of the Perry Preschool Program on youths through age 19* (High/Scope Educational Research Foundation, Monograph 8). Pysilanti, MI: High/Scope Press.

Biederman, J., Rosenbaum, J. F., Bolduc-Murphy, E. A., Faraone, S. V., Chaloff, J., Hirshfeld, D. R., & Kagan, J. (1993). A 3-year follow-up of children with and without behavioral inhibition. *Journal of the American Academy of Child and Adolescent Psychiatry, 32*(4), 814–821.

Birchwood, M. (1995). Early intervention in psychotic relapse: Cognitive approaches to detection and management. *Behaviour Change, 12,* 2–19.

Caplan, G. (1964). *Principles of preventive psychiatry*. New York: Basic Books.

Caplan, M., Weissberg, R. P., Grober, J. S., Sivo, P. J., Grady, K., & Jacoby, C. (1992). Social competence promotion with inner-city and suburban young adolescents: Effects on social adjustment and alcohol use. *Journal of Consulting and Clinical Psychology, 60,* 56–63.

Clarke, G. N., Hawkins, W., Murphy, M., & Sheeber, L. (1993). School-based primary prevention of depressive symptomatology in adolescents: Findings from two studies. *Journal of Adolescent Research, 8,* 183–204.

Clarke, G. N., Hawkins, W., Murphy, M., Sheeber, L., Lewinsohn, P., & Seeley, J. R. (1995). Targeted prevention of unipolar depressive disorder in an at-risk sample of high school adolescents: A randomized trial of group cognitive intervention. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 312–321.

Coie, J. D., Watt, N. F., West, S. G., Hawkins, J. D., Asarnow, J. R., Markman, H. J., Ramey, S. L., Shure, M. B., & Long, B. (1993). The science of prevention: A conceptual framework and some directions for a National Research Program. *American Psychologist, 48,* 1013–1022.

Compas, B. (1987). Coping with stress during childhood and adolescence. *Psychological Bulletin, 101,* 393–403.

Conduct Problems Research Group. (1992). A developmental and clinical model for the prevention of conduct disorder: The FAST Track Program. *Development and Psychopathology, 4,* 509–527.

Dadds, M. R. (in press). Conduct disorder. In R. T. Ammerman & M. Hersen (Eds.), *Handbook of prevention and treatment of children and adolescents: Interventions in the real world context*. New York: Wiley .

Dadds, M. R., Barrett, P. M., & Rapee, R. M. (1996). Family process and child anxiety and aggression: An observational analysis. *Journal of Abnormal Child Psychology, 24,* 715–734.

Dadds, M. R., Spence, S. H., Holland, D. E., Barrett, P. M., & Laurens, K. R. (1997). Prevention and early intervention for anxiety disorders: A controlled trial. *Journal of Consulting and Clinical Psychology, 65,* 627–635.

DHHS (Department of Health and Human Services) (1991). *Healthy People 2000*. Washington, DC: Government Printing Office; DHHS Pub. No. (PHS) 91–50212.

Dishion, T. J., & Andrews, D. W. (1995). Preventing escalation in problem behaviors with high-risk young adolescents: Immediate and 1-year outcomes. *Journal of Consulting and Clinical Psychology, 63,* 538–548.

Dodge, K. A., Pettit, G. S., & Bates, J. E. (1994). Socialization mediators of the relation between socioeconomic status and child conduct problems. Special Issue: Children and poverty. *Child Development, 65,* 649–665.

Dollinger, S. J. (1986). The measurement of children's sleep disturbances and somatic complaints following a disaster. *Child Psychiatry and Human Development, 16,* 148–153.

Dollinger, S. J., O'Donnell, J. P., & Staley, A. A. (1984). Lightning-strike disaster: Effects on children's fears and worries. *Journal of Consulting and Clinical Psychology, 52,* 1028–1038.

Eisenberg, L. (1995). Social policy and the reality of prevention. In B. Raphael & G. D. Burrows (Eds.), *Handbook of studies on preventive psychiatry* (pp. 31–50). Amsterdam: Elsevier Science.

Ellickson, P. L., & Bell, R. M. (1990). Drug prevention in junior high: A multi-site longitudinal test. *Science, 247,* 1299–1305.

Falloon, I. R. (1992). Early intervention for first episode schizophrenia: A preliminary exploration. *Psychiatry, 55,* 4–15.

Farrington, D. P. (1995). The challenge of teenage antisocial behavior. In M. Rutter (Ed.), *Psychosocial disturbances in young people: challenges for prevention* (pp. 83–130). New York: Cambridge University Press.

Felner, R. D., & Adan, A. M. (1988). The School Transition Environment Project: An ecological intervention and evaluation. In R. H. Price, E. L. Cowen, R. P. Lorion, & J. Ramos-McKay (Eds.), *Fourteen ounces of prevention: A case book for practitioners* (pp. 111–122). Washington DC: American Psychological Association.

Field, T. M., Widmayer, S. J., Stringer, S., & Ignatoff, E. (1980). Teenage, lower-class, black mothers and their preterm infants: An intervention and developmental follow-up. *Child Development, 51,* 426–436.

Forehand, R. (1992). Parental divorce and adolescent maladjustment. *Behaviour Research and Therapy, 30,* 319–327.

Gillham, J. E., Reivich, K. J., Jaycox, L. H., & Seligman, M. E. P. (1995) Prevention of depressive symptoms in school children: Two year follow-up. *Psychological Science, 6,* 343–351.

Goodyer, I. M., & Altham, P. M. (1991). Lifetime exit events and recent social and family adversities in anxious and depressed school-aged children. *Journal of Affective Disorders, 21,* 219–228.

Gordon, R. (1987). An operational classification of disease prevention. In J. A. Steinberg & M. M. Silverman (Eds.), *Preventing mental disorders* (pp. 20–26). Rockville, MD: Department of Health and Human Services.

Guerra, N. G., Huesmann, L. R., Tolan, P. H., & Van Acker, R. (1995). Stressful events and individual beliefs as correlates of economic disadvantage and aggression among urban children. *Journal of Consulting and Clinical Psychology, 63,* 518–528.

Harnish, J. D., Dodge, K. A., Valente, E., & Conduct Problems Prevention Research Group (1995). Mother–child interaction quality as a partial mediator of the roles of maternal depressive symptomatology and socioeconomic status in the development of child behavior problems. *Child Development, 66,* 739–753.

Hawkins, J. D., Catalano, R. F., Morrison, D. M., O'Donnell, J., Abbott, R. D., & Day, L. E. (1992). The Seattle Social Development Project: Effects of the first four years on protective factors and problem behaviors. In J. McCord & R. Tremblay (Eds.), *The prevention of antisocial behavior in children* (pp. 139–161). New York: Guilford Press.

Hess, R. D., & Camara, K. A. (1979). Post-divorce family relationships as mediating factors in the consequences of divorce for children. *Journal of Social Issues, 35,* 70–95.

Hightower, A. D., & Braden, J. (1991). Prevention. In T. R. Kratochwill & R. J. Morris (Eds.), *The practice of child therapy* (pp. 410–440). New York: Pergamon.

Hirshfeld, D. R., Rosenbaum, J. F., Biederman, J., Bolduc, E. A., Faraone, S. V., Snidman, N., Reznick, J. S., & Kagan, J. (1992). Stable behavioral inhibition and its association with anxiety disorder. *Journal of American Academy of Child and Adolescent Psychiatry, 31,* 103–111.

Hodges, W. F. (1991). *Interventions for children of divorce*. New York: Wiley.

Human Rights and Equal Opportunity Commission (1993). *Human rights and mental illness: report of the national inquiry into the human rights of people with mental illness*. Canberra: Australian Government Publishing Service.

IOM (Institute of Medicine) (1989). *Research on children and adolescents with mental, behavioral, and developmental disorders*. Washington, DC: National Academy Press.

Jackson, H. J., McGorry, P. D., Edwards, J., & Hulbert, C. 1996). Cognitively oriented psychotherapy for early psychosis (COPE). In P. Cotton & H. Jackson (Eds.), *Early intervention and prevention in mental health* (pp. 131–154). Melbourne: The Australian Psychological Society.

Jaycox, L. H., Reivich, K. J., Gillham, J., & Seligman, M. E. P. (1994). Prevention of depressive symptoms in school children. *Behaviour Research and Therapy, 32,* 801–816.

Johnson, D. L. (1991). Primary prevention of behavior problems in young children: The Houston Parent–Child Development Center. In R. Price, E. L. Cowen, R. P. Lorion, & J. Ramos-McKay (Eds.), *Fourteen ounces of prevention: A casebook for practitioners* (pp. 44–52). Washington, DC: American Psychological Association.

Kagan, J., Reznick, J. S., & Gibbons, J. (1989). Inhibited and uninhibited types of children. *Child Development, 60,* 838–845.

Kagan, J., & Snidman, N. (1991). Infant predictors of inhibited and uninhibited profiles. *Psychological Science, 2,* 40–43.

Kaslow, N. J., & Racusin, G. R. (1994). Family therapy for depression in young people. In W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 345–363). New York: Plenum .

Kazdin, A. E., & Kagan, J. (1994). Models of dysfunction in developmental psychopathology. *Clinical Psychology: Science and Practice, 1,* 35–52.

Kellam, S. G., & Rebok, G. W. (1992). Building developmental and etiological theory through epidemiologically based preventive intervention trials. In J. McCord & R. E. Tremblay (Eds.), *Preventing antisocial behavior: interventions from birth through adolescence* (pp. 162–195). New York: Guilford Press.

Kellam, S. G., Rebok, G. W., Ialongo, N., & Mayer, L. S. (1994). The course and malleability of aggressive behavior from early first grade into middle school: Results of a developmental epidemiologically-based preventive trial. *Journal of Child Psychology and Psychiatry, 35,* 259–281.

Kendall, P. C. (1991). Considering cognition in anxiety-disordered children. Special Issue: Assessment of childhood anxiety disorders. *Journal of Anxiety Disorder, 5,* 167–185.

Kraemer, H. C., & Fendt, K. H. (1990). Random assignment in clinical trials: Issues in planning. *Journal of Clinical Epidemiology, 43,* 1157–1167.

Lally, J. R., Mangione, P. L., Honig, A. S., & Wittner, D. S. (1988). More pride, less delinquency: Findings from the ten-year follow-up study of the Syracuse University Family Development Research Program. *Zero to three, 8*(4) 13–18.

Levenstein, P. (1992). The Mother–Child Home Program: Research methodology and the real world. In J. McCord & R. C. Tremblay (Eds.), *Preventing antisocial behavior: Interventions from birth through adolescence* (pp. 43–66). New York: Guilford Press.

Lochman, J. E., Coie, J. D., Underwood, M. K., & Terry, R. (1993). Effectiveness of a social relations intervention program for aggressive and nonaggressive, rejected children. *Journal of Consulting and Clinical Psychology, 6,* 1053–1058.

Lochman, J. E., & Dodge, K. A. (1994). Social-cognitive processes of severely violent, moderately aggressive, and nonaggressive boys. *Journal of Consulting and Clinical Psychology, 62,* 366–374.

Lochman, J. E. & The Conduct Problems Prevention Research Group (1995). Screening of child behavior problems at school entry. *Journal of Consulting and Clinical Psychology, 63,* 549–559.

Loeber, R. (1990). Development and risk factors of juvenile antisocial behavior and delinquency. *Clinical Psychology Review, 10,* 1–41.

McFarlane, A. C. (1987). Posttraumatic phenomena in a longitudinal study of children following a natural disaster. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 764–769.

Mrazek, P. J., & Haggerty, R. J. (Eds.) (1994). *Reducing the risks for mental disorders: Frontiers for preventive intervention research.* Washington, DC: National Academy Press.

O'Donnell, J., Hawkins, J. D., & Abbott, R. D. (1995). Predicting serious delinquency and substance use among aggressive boys. Special section: Prediction and prevention of child and adolescent antisocial behavior. *Journal of Consulting and Clinical Psychology, 63*(4), 529–537.

Olds, D. L, Henderson, C. R., Tatelbaum, R., Chamberlin, R. (1988). Improving the life-course development of socially disadvantaged mothers: A randomized trial of nurse home visitation. *American Journal of Public Health, 78,* 1436–1445.

Patterson, G. R. (1996). Some characteristics of a developmental theory for early-onset delinquency. In M. F. Lenzenweger & J. J. Haugaard (Eds.), *Frontiers of developmental psychopathology* (pp. 59–71). New York: Oxford University Press.

Patterson, G. R, & Stouthamer-Loeber, M. (1984). The correlation of family management practices and delinquency. *Child Development, 55,* 1299–1307.

Pedro-Carroll, J. L., & Cowen, E. L. (1985). The Children of Divorce Intervention Program: An investigation of the efficacy of a school-based prevention program. *Journal of Consulting and Clinical Psychology, 53,* 603–611.

Perry, C. L., Grant, M., Ernberg, G., Florenzano, R. U., Langdon, M. C., Myeni, A. D., Waahlberg, R., Berg, S., Anderson, K., Fisher, K. J., Bla-Temple, D., Cross, D., Saunders, B., Jacobs, D. R., & Schmid, T. (1989). WHO Collaborative Study on Alcohol Education and Young People: Outcomes of a four-country pilot study. The International Journal of the Addictions, 24, 1145–1171.

Peters, R. D., & Russell, C. C. (1994). *Better Beginnings, Better Futures Project: Model, program and research overview.* Ontario: Queen's Printer for Ontario.

Plomin, R., & Stocker, C. (1989). Behavioral genetics of emotionality. In S. Reznick (Ed.), *Perspectives on behavioral inhibition* (pp. 219–240). Chicago, IL: University of Chicago Press.

Price, J. (1995). Substance abuse: Risks and opportunities for prevention. In B. Raphael & G. D. Burrows (Eds.), *Handbook of studies on preventive psychiatry* (pp. 531–548). Amsterdam: Elsevier Science.

Prior, M. (1992). Childhood temperament. *Journal of Child Psychology and Psychiatry, 33,* 249–281.

Rachman S. J. (1980). Emotional processing. *Behaviour Research and Therapy, 18,* 51–60.

Rapee, R. M. (1997). The potential role of childrearing practices in the development of anxiety and depression. *Clinical Psychology Review, 17,* 47–67.

Robins, L. H., & Regier, D. A. (1991). *Psychiatric disorders in America: The Epidemiological Catchment Area Study.* New York: The Free Press.

Rosenbaum, J. F., Biederman, J., Bolduc-Murphy, E. A., Faraone, S. V., Chaloff, J., Hirshfeld, D. R., & Kagan, J. (1993). Behavioral inhibition in childhood: A risk factor for anxiety disorders. *Harvard Review of Psychiatry, May/June,* 2–6.

Rutter, M. (1987). Psychosocial resilience and protective mechanisms. *American Journal of Orthopsychiatry, 57,* 316–331.

Rutter, M., Macdonald, H., Le Couteur, A., Harrington, R., Bolton, P., & Bailey, A. (1990). Genetic factors in child psychiatric disorders–II. Empirical findings. *Journal of Child Psychology and Psychiatry, 31,* 39–83.

Sanders, M. R., Dadds, M. R., & Barrett, P. M. (1995). The prevention of anxiety and stress related disorders in children. In B. Raphael & G. D. Burrows (Eds.), *Handbook of studies on preventive psychiatry* (pp. 399–421). Amsterdam: Elsevier Science.

Sanders, M. R., & Markie-Dadds, C. (1996). Triple P: A multilevel family intervention program for children with disruptive behavior disorders. In P. Cotton & H. Jackson (Eds.), *Early intervention and prevention in mental health* (pp. 59–85). Melbourne: The Australian Psychological Society.

Shure, M. B., & Spivack, G. (1980). Interpersonal problem solving as a mediator of behavioral adjustment in preschool and kindergarten children. *Journal of Applied Developmental Psychology, 1,* 29–44.

Shure, M. B., & Spivack, G. (1982). Interpersonal problem-solving in young children: A cognitive approach to prevention. *American Journal of Community Psychology, 10,* 341–356.

Shure, M. B., & Spivack, G. (1988). Interpersonal cognitive problem solving. In R. H. Price, E. L. Cowen, R. P. Lorion, & J. Ramos-McKay (Eds.), *Fourteen ounces of prevention: A casebook for practitioners* (pp. 69–82). Washingon, DC: American Psychological Association.

Soussignan, R., Koch, P., & Montagner, H. (1988). Behavioral and cardiovascular changes in children moving from kindergarten to primary school. *Journal of Child Psychology and Psychiatry, 29,* 321–333.

Spence, S. H. (1996a). A case for prevention. In P. Cotton & H. Jackson (Eds.), *Early intervention and prevention in mental health* (pp. 1–19). Melbourne: The Australian Psychological Society.

Spence, S. H. (1996b). The prevention of anxiety disorders in childhood. In P. Cotton & H. Jackson (Eds.), *Early intervention and prevention in mental health* (pp. 87–107). Melbourne: The Australian Psychological Society.

Spence, S. H., & Dadds, M. R. (1996). Preventing childhood anxiety disorders. *Behaviour Change, 13,* 241–249.

Sugar, M. (1989). Children in a disaster: An overview. *Child Psychiatry and Human Development, 19,* 163–179.

Terr, L. C. (1981). Psychic trauma in children: Observations following the Chowchilla school-bus kidnapping. *American Journal of Psychiatry, 138,* 14–19.

Tremblay, R. E., Bitaro, F., Bertrand, L., LeBlanc, M., Beauchesne, H., Boileau, H., & David, L. (1992). Parent and child training to prevent early onset of delinquency: The Montreal Longitudinal-Experimental study. In J. McCord & R. Tremblay (Eds.), *Preventing antisocial behavior: Interventions from birth through adolescence* (pp. 117–138). New York: Guilford Press.

Tremblay, R. D., McCord, J., Boileau, H., Charlebois, P., Gagnon, C., LeBlanc, M., & Larivee, S. (1991). Can disruptive boys be helped to become competent? *Psychiatry, 54,* 148–161.

Tremblay, R. E., Pagani-Kurtz, L., Vitaro, F., Masse, L. C., & Pihl, R. O. (1995). A bimodal preventive intervention for disruptive kindergarten boys: Its impact through mid-adolescence. *Journal of Consulting and Clinical Psychology, 63,* 560–568.

US Advisory Board on Child Abuse and Neglect. (1990). *Child abuse and neglect: Critical first steps in response to a national emergence*. Washington DC: Government Printing Office.

Werner, E. E. (1987). Vulnerability and resiliency in children at risk for delinquency: A longitudinal study from birth to young adulthood. In J. D. Burchard & S. N. Burchard (Eds.), *Prevention of delinquent behavior* (pp. 16–43). Beverly Hills, CA: Sage.

White, C. C., Koplan, J. P., & Orenstein, W. A. (1985). Benefits, risks, and costs of immunization for measles, mumps, and rubella. *American Journal of Public Health, 75,* 739–744.

Yule, W. (1991). Work with children following disasters. In M. Herbert (Ed.), *Clinical child psychology: A social learning approach to theory and practice* (pp. 223–241). Chichester, UK: Wiley.

Yule, W. & Williams, R. (1990). Post-traumatic stress reactions in children. *Journal of Traumatic Stress, 3,* 279–295.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.13

Anxiety Disorders

MARK R. DADDS and PAULA M. BARRETT
Griffith University, Qld, Australia
and
VANESSA E. COBHAM
University of Queensland, Australia

5.13.1 INTRODUCTION 317

5.13.2 ANXIETY DISORDERS 318
- *5.13.2.1 Phenomenology* 318
- *5.13.2.2 Diagnostic Features/Differential Diagnosis* 319
- *5.13.2.3 Prevalence* 321
- *5.13.2.4 Etiology* 322
  - *5.13.2.4.1 Psychodynamic* 322
  - *5.13.2.4.2 Biological* 323
  - *5.13.2.4.3 Learning mechanisms* 324
  - *5.13.2.4.4 Stressful events* 324
  - *5.13.2.4.5 Family processes* 324
  - *5.13.2.4.6 Recent studies of attachment processes* 325
  - *5.13.2.4.7 Family systems theories* 326
- *5.13.2.5 Developmental Course* 326

5.13.3 CONCEPTUALIZATION AND CLINICAL FORMULATION 327
- *5.13.3.1 Multimodal Assessment* 327
- *5.13.3.2 Treatment Formulation* 330

5.13.4 TREATMENT 331
- *5.13.4.1 Psychosocial Treatments* 331
- *5.13.4.2 Pharmacological Treatments* 333

5.13.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 335

5.13.6 SUMMARY 336

5.13.7 REFERENCES 337

5.13.1 INTRODUCTION

Since Freud's description of separation fears in Little Hans at the turn of the century, anxiety disorders in children have received scarce but increasing attention from clinical researchers. Only recently, however, have controlled studies into the epidemiology, development, and treatment of these children gathered momentum. There are a number of reasons for this neglect. Across most classes of psychopathology, children receive less attention than adults (Kazdin, 1985), and children with anxiety disorders often do not attract the attention of mental health services because their behavior may be relatively cooperative and unobtrusive compared with disruptive peers. Despite this, there is evidence to suggest that some forms of childhood anxiety are severely debilitating and predictive of chronic mental health problems.

Through this century, a number of critical issues regarding anxiety in children have emerged. Because far more research has been conducted into anxiety in adults, there has been a tendency for child researchers to borrow conceptual models and treatment techniques from the adult literature. The value of this has been assumed rather than supported by hard data and thus the field is becoming increasingly careful to incorporate and evaluate developmental factors into its borrowed models. Other issues are common to most forms of child psychopathology. The reliability, validity, and clinical utility of diagnostic systems remain controversial issues despite much effort to refine the various categorical systems currently in use. While this is true for all forms of psychopathology, the use of formal diagnoses with children has many unique problems raised by the rapid changes children go through and the diversity of stakeholders such as the child, parents, and teachers that are involved in the child's well-being. Further, the childhood anxiety disorders continue to elude attempts to find discrete categories, with overlap and comorbidity being the rule rather than the exception.

The relative importance and interaction of temperamental vs. learned factors has moved beyond simplistic nature–nuture debates to models that emphasize multiple causal variables that interact in predictable development sequences. Thus, a disorder is seen as taking a developmental trajectory through predictable changes and stable phases. For conduct disorder in children, much progress has been made in mapping common trajectories; for anxiety disorders, relatively little is known. Thus, the links between child and adult anxiety remain an area of speculation. The twentieth century has also heralded the growth of behavioral and cognitive sciences and family therapies as major clinical and scientific movements. Both are important to anxiety in children. Contemporary conceptual and treatment models emphasize the interactions of learning factors, cognitive style, and family process. Recently, exciting and important work has shown that these can be studied under controlled conditions and used with great effect to bring about clinically significant improvements in anxious children.

The aim of this chapter is to review current knowledge of the nature, development, and treatment of the major classes of childhood anxiety. Specifically, it will focus on the more general forms of anxiety disorder such as separation anxiety, social fears, and generalized anxiety states. Simple phobias, obsessive compulsive disorder, and traumatic stress disorder are covered elsewhere in this volume.

5.13.2 ANXIETY DISORDERS

5.13.2.1 Phenomenology

Marks (1969) described fear as "a normal response to active or imagined threat in higher animals ... (comprising) an outer behavioral expression, an inner feeling, and accompanying physiological changes" (p. 1). Anxiety, a closely related construct, has been defined as "a dysphoric, aversive feeling similar to fear that arises without an obvious external threat" (Miller, 1983, p. 338) and "apprehension, tension, or uneasiness related to the expectation of danger, whether internal or external" (Kendall et al., 1992, p. 1). Although distinctions are typically made between the two constructs, the prevailing diagnostic classification system, the *Diagnostic and statistical manual of mental disorders* (*DSM*) (American Psychiatric Association [APA], 1994), conceptualizes both fear and anxiety as being manifested in very similar physiological, cognitive, and affective patterns (Kashani, Dandoy, & Orvaschel, 1991). Based on Lang's (1968) tripartite model, anxiety is typically viewed as a multidimensional construct involving physiological, behavioral, and cognitive components or symptoms.

Fear and anxiety are common, typically transitory, and, in many situations, adaptive in that they protect the individual from harm. This being the case, anxiety and anxiety disorders are typically conceptualized as varying quantitatively along a continuum, with the degree of distress, impairment in functioning, and/or interference with daily life, discriminating between what is "normal" and adaptive and what is problematic. The difficulty, of course, comes in deciding where normal anxiety ends, and clinical anxiety begins. This difficulty is especially marked in work with children owing to developmental considerations that must be taken into account. Given that the rate of development is far greater during childhood and adolescence than at any time in life, knowledge of what are "normal" fears and anxieties at each developmental stage is vital when attempting to ascertain whether or not a child suffers from clinical anxiety. A fear that is normative for a preschool-aged child may be maladaptive if experienced by an adolescent.

As a consequence of children's developmental experiences and increasing cognitive abilities, the content of their fears and anxieties changes over time, with the focus generally moving with age from concrete, external fears to increasingly internalized and abstract anxieties. Thus, infants tend to fear strangers, loud noises, and strange and unexpected objects (Jersild & Holmes, 1935), while children of one and two years of age fear separation from their parents, loud noises, animals, the dark, and the toilet (Miller, Barrett, & Hampe, 1974). Of these, fears about animals, the dark, and separation from parents frequently persist into a child's fifth year. Between the ages of four and six, predominant fears focus on kidnappers and robbers and supernatural beings, such as ghosts and monsters. At the age of approximately six years, fears of bodily injury, death, and achievement develop (Bauer, 1976; Croake & Knox, 1973), and for many children, these continue into early adolescence. The final developmental stage appears to begin at approximately 10 or 11 years of age, when fears regarding social comparison, physical appearance, personal conduct, and school examinations predominate (Croake, 1969; Miller, Barrett, Hampe, & Noble, 1972). Factor analyses have provided additional information regarding the organization of the content of childhood fears, with the most commonly accepted solution yielding five factors consisting of fear of: failure and criticism, the unknown, injury and small animals, danger and death, and medical procedures (Ollendick, Matson, & Helsel, 1985).

McFarlane, Allen, and Honzik (1954), for instance, reported that 90% of the normal children aged two to 14 years in their sample reported at least one specific fear. The vast majority of children report experiencing multiple fears, and although the number of self-reported fears appears to decline with age, this reduction does not seem to be as great as might be expected. Thus, Ollendick et al. (1985) reported an average of 14.2 fears in children aged seven to nine years and an average of 11.6 fears in their sample of 16 to 18-year-olds. Another consistent finding is that girls typically report more fears than boys (although whether this is due to actual gender differences or sociocultural factors remains unclear). King et al. (1989) reported the average numbers of fears for males and females aged between eight and 16 years to be 10 and 18, respectively. Across cultures, total scores on the Revised Fear Survey Schedule for Children (Ollendick, 1983) appear to be similar for children and adolescents in English-speaking countries (Ollendick, King, & Frary, 1989; Yule, Udwin, & Murdoch, 1990), although Dutch children have been found to score comparatively lower (Oosterlaan, Prins, & Sergeant, 1992) and Portuguese children to score comparatively higher (Fonseca, 1993) than their American, Australian, and English counterparts. The most common fears reported by children and adolescents, however, appear to be stable across different countries and cultures.

In summary, fears and anxieties are common in children and adolescents and in the majority of cases, they are age-appropriate, mild, frequently adaptive, and transitory. Unfortunately, for some children the reverse is the case. The fear and anxiety they experience is so severe and/or pervasive that it can cause intense distress, functional impairment, and interference. It is upon this group that the remainder of this chapter will focus.

5.13.2.2 Diagnostic Features/Differential Diagnosis

The revised third edition of the *DSM* (*DSM-III-R*) (APA, 1987) listed three anxiety disorders that typically begin in childhood or adolescence. These were:

(i) Separation anxiety disorder, defined as undue anxiety regarding separation from significant figures in the child's life. The child's reaction to such separations is beyond that expected for his or her developmental level, and may at times approach the level of panic (Gittelman-Klein, 1988).

(ii) Avoidant disorder, defined, as extreme withdrawal from and/or avoidance of, contact with unfamiliar people. In order to meet diagnostic criteria for avoidant disorder, a child must manifest anxiety regarding new people and situations that is sufficiently severe to interfere with social functioning, while at the same time demonstrating a desire and a capacity for warm and close relationships with familiar people (such as family members).

(iii) Overanxious disorder, defined as exaggerated or undue anxiety or worry. This diagnostic category is characterized by generalized anxiety that may include: self-consciousness; excessive worry about future events (e.g., going to see a doctor), or about past events (e.g., something the child said that may have been taken the wrong way); and anxiety about performance and competence.

In addition to these three anxiety disorders of childhood and adolescence, children could also be diagnosed with anxiety disorders of "adulthood." These included agoraphobia, panic disorder with or without agoraphobia, specific or simple phobia, social phobia, obsessive compulsive disorder, post-traumatic stress disorder, acute stress disorder, generalized anxiety disorder, anxiety disorder due to a general medical condition, substance-induced anxiety disorder, and anxiety disorder not otherwise specified.

With the introduction of the *DSM-IV* (APA, 1994), categories have been changed so that only one anxiety disorder of childhood or adolescence remains—separation anxiety disorder. Avoidant disorder and overanxious disorder have been effectively subsumed into the diagnoses of social phobia and generalized anxiety disorder respectively. In addition, a child may still be diagnosed as suffering from any of the anxiety disorders of adulthood.

As it pertains to childhood anxiety, the classification system used in the *10th edition of the international classification of disorders* (*ICD-10*) (World Health Organization, 1992) is very similar to that used in *DSM-IV*. The differences between the *DSM-IV* and the *ICD-10* adult anxiety categories are minor and insignificant. For the childhood disorders, as is the case in *DSM-IV*, all anxiety diagnoses are grouped together in a single category, which is then split into a number of subcategories. In *ICD-10* these include: separation anxiety disorder; phobic anxiety disorder; social anxiety disorder of childhood; sibling rivalry disorder; and other childhood emotional disorders (including identity disorder and overanxious disorder).

It is important to note that very little empirical research has been conducted to validate the diagnostic categories for child anxiety disorders (Silverman, 1992; Werry, 1994). Although successive versions of the *DSM* have placed increasing emphasis on empirical research, Werry contended that, as yet, the major field trials aimed at validating the childhood anxiety disorders have not been undertaken. However, as Spence (1996) notes, the little evidence that has been available has typically been utilized effectively in developing the *DSM* system. For example, the diagnostic category of avoidant disorder was omitted from *DSM-IV* after research indicated that it did not differ sufficiently from the category of social phobia to justify its inclusion (Francis, Last, & Strauss, 1992). In a similar way, overanxious disorder was subsumed by generalized anxiety disorder on the basis of empirical research (Beidel, 1991).

The controversy over whether the childhood anxiety disorders are indeed distinct clinical entities is further fueled by the overlap in diagnostic criteria across the disorders (for instance, fears of poor performance and negative evaluation are common among children with both social phobia and generalized anxiety disorder), and by the consistent failure of research to differentiate between the different childhood anxiety disorders in terms of etiology. Further, comorbidity among the childhood anxiety disorders is common, much as it is for adults (Sanderson, Di Nardo, Rapee, & Barlow, 1990). Klein and Last (1989) report that almost all anxiety disorder children meet criteria for at least one other anxiety disorder, most commonly overanxious disorder. Similarly, over 50% of children diagnosed with overanxious disorder report comorbid anxiety disorders.

Depression is also commonly reported as a concomitant condition in the childhood anxiety disorders (Brady & Kendall, 1992). The precise figures have varied somewhat, with two small studies reporting either no anxious children meeting criteria for depression (Hershberg, Carlson, Cantwell, & Strober, 1982) or over 80% meeting criteria (Bernstein & Garfinkel, 1986). These discrepancies may be due to a confound with age since older children with overanxious disorder or social anxiety disorder have been found to be more likely to meet criteria for comorbid depression than younger children (Last, Hersen, Kazdin, Finkelstein, & Strauss, 1989; Strauss, Lease, Last, & Francis, 1988). Last et al. (1987) reported that approximately 33% of their population of overanxious disorder and social anxiety disorder children also met criteria for major depression, a figure more in line with the adult literature (Sanderson et al., 1990).

Recently, research has been conducted that provides some support for the *DSM* classification of anxiety disorders. Spence (1996) employed confirmatory factor analysis to test whether the pattern of anxiety symptoms among a community sample of children is in keeping with the *DSM-IV* classification. Using two samples of 698 children aged between eight and 12 years, a model comprising six discrete but correlated factors (reflecting the areas of panic/agoraphobia, social phobia, separation anxiety, obsessive-compulsive problems, generalized anxiety, and physical fears) provided an impressive fit of the data, while the high correlations among latent factors were more than adequately explained by a higher-order model in which each of the six first-order factors loaded on a single, second-order factor (Spence, 1996). Thus, Spence concluded that the *DSM-IV* diagnostic categories of anxiety disorders in children have validity in terms of a strong second-order factor related to anxiety in general, within which the specific categories of anxiety can be identified.

A number of studies have assessed inter-diagnostician reliability with childhood anxiety disorders. Two issues related to reliability are important. First, there is the issue of the degree of agreement between two interviewers who have information from the same source. Second, given that children rarely go for treatment without the urging of a guardian, there is the issue of the degree of agreement between two different sources of information, most commonly, parent and child.

Looking first at the issue of agreement between two interviewers, a number of studies have been conducted showing somewhat mixed results, with kappas ranging from 0.22 to 1.00 (see Silverman, 1991, for a review). Many of these differences are likely to be a result of methodological issues so that data are currently inconclusive with respect to the reliability of childhood anxiety disorders (Silverman, 1991). Nevertheless, some suggestions can possibly be made. First, it seems that the use of structured clinical interviews can greatly enhance reliability (Last, 1988). Second, reliability appears to be considerably better when the second rating is based on a recording of the first, rather than having two clinicians separately interviewing the same source (Silverman, 1991). This latter finding suggests that while the criteria for childhood anxiety diagnoses may be quite reliable, the informants are not.

Rapee, Barrett, Dadds, and Evans (1994) compared inter-rater reliability using the Anxiety Disorders Interview for Children (ADIS), parent and child versions (Silverman & Nelles, 1988), to assess the diagnostic status of 161 children referred to an anxiety disorders clinic. It should be noted that Rapee et al. (1994) reported difficulties in differentiating social phobia from avoidant disorder and so these were grouped together as social phobia. First, interviewers were provided with either live or videotaped information from both parents and the child, yielding kappa coefficients of 0.63–0.82 across the different anxiety disorders. Similar agreement levels were found based on either parent interviews or child interviews alone. However, the agreement levels calculated by comparing diagnoses based on parent interviews with diagnoses based on child interviews were low, especially for overanxious disorder which was much more frequently reported by parents than children. Overall, the results indicate that the use of a structured interview can produce moderately to highly reliable diagnoses using the *DSM-III-R* system for childhood anxiety disorders, and that utilizing different sources of information is important in reaching an accurate final consensus diagnosis. Other research investigating the agreement between different sources has indicated that information derived about the child from the parents is often quite different from that reported by the child (see Klein, 1991, for a review) and different again from teachers' reports (Stavrakaki, Vargo, Roberts, & Boodoosingh, 1987). At present there is no evidence to indicate that one source of evidence is any more valid than another, thus the recommendation is that multiple informants are used in the diagnosis of childhood disorders.

5.13.2.3 Prevalence

The results of recently conducted epidemiological studies indicate that anxiety disorders are probably the most common category of childhood and adolescent disorders (Bernstein & Borchardt, 1991). In a large epidemiological study employing a semistructured diagnostic interview schedule, Kashani and Orvaschel (1990) found that anxiety disorders were the most common form of psychopathology reported by a community sample of children and adolescents across three age groups (eight, 12, and 17 years). These researchers reported that 25.7% of the eight-year-old children, 15.7% of the 12-year-olds, and 21.4% of the 17-year-olds met diagnostic criteria for an anxiety disorder. It is important to note, however, that the estimates reported in this study are typically criticized as being excessively high relative to those figures reported in other epidemiological investigations. For instance, in a general sample of nearly 800 pediatric patients aged seven to 11 years, Costello (1989) found that 8.9% met diagnostic criteria for at least one anxiety disorder, while Kashani and Orvaschel (1988) found that in a representative sample of 150 adolescents aged 14 to 16 years, 17.3% met diagnostic criteria for at least one anxiety disorder, and 8.7% met diagnostic criteria *and* exhibited dysfunction warranting treatment. In another major epidemiological study, conducted in New Zealand, McGee et al. (1990) reported the prevalence rate of any anxiety disorder among their sample of almost 1000 adolescents was 12.6%, making anxiety disorders the most prevalent disorder for this age group.

With respect to specific diagnoses, Kashani and Orvaschel (1990) reported that 12.9% of all children and adolescents in their nonreferred community sample met diagnostic criteria for separation anxiety disorder. Across the three age groups, the prevalence rates were: 18.6% of the eight-year-olds, 8.6% of the 12-year-olds, and 11.4% of the 17-year-olds. Costello (1989) found that approximately 4% of children aged seven to 11 met criteria for separation anxiety disorder, and Anderson, Williams, McGee, and Silva (1987) reported a prevalence rate of 3.5% for the 11-year-olds in their sample. In a follow-up of this same sample four years later, McGee et al. (1990) found that 2% of 15-year-olds met diagnostic criteria for separation anxiety disorder, while Kashani and Orvaschel (1988) reported a rate of 0.7% for the adolescents in their sample. Within a sample of children referred to an anxiety disorders clinic, separation anxiety disorder was the most common primary diagnosis, with a prevalence rate of approximately 33% (Last et al., 1987). Within this diagnosis, females tend to be overrepresented (across their three age groups, Kashani and Orvashel (1990) reported that 21.0% of girls, compared with 4.8% of boys, met diagnostic criteria for separation anxiety disorder), as do younger children, and Caucasian children from lower socioeconomic backgrounds. Further, an epidemiological study conducted by Velez, Johnson, and Cohen (1989) identified lower socioeconomic status as a risk factor for the development of separation anxiety disorder.

Prevalence rates for avoidant disorder tend to be somewhat lower than for either of the other two anxiety diagnoses of childhood and adolescence. Costello (1989), for instance, reported that 1.6% of children aged seven to 11 met diagnostic criteria for avoidant disorder. Also relevant here are prevalence rates for social phobia, which replaces avoidant disorder in *DSM-IV*. Kashani and Orvaschel (1990) reported that 1.0% of their total sample met criteria for this diagnosis. Across the three age groups, the rates were as follows: 0.0% of the 8-year-olds, 1.4% of the 12-year-olds, and 1.4% of the 17-year-olds. These authors also reported an even gender ratio for the diagnosis of social phobia—across the three age groups, 1.0% of boys and 1.0% of girls met criteria.

For overanxious disorder in children, the prevalence rates range from 2.9% (Anderson et al., 1987) to 4.6% (Costello, 1989), while 7% of adolescents have been reported to meet criteria for the diagnosis (Kashani & Orvaschel, 1988). Kashani and Orvaschel (1990) reported that 12.4% of their total sample met criteria for overanxious disorder, with the age breakdown being as follows: 8.6% of eight-year-olds, 11.4% of 12-year-olds, and 17.1% of 17-year-olds. As can be seen from these figures, this disorder seems to occur more frequently in older children. Within a clinically anxious sample, approximately 15% had a primary diagnosis of overanxious disorder (Last et al., 1987). In terms of the gender ratio, males and females appear to be equally represented within this diagnosis, until adolescence when there is female predominance (Werry, 1991). Finally, it appears that Caucasian children from middle and upper class families may be overrepresented in overanxious disorder (Last et al.), although this finding may well reflect referrals rather than prevalence.

Thus, while available data on the prevalence of anxiety disorders are somewhat discrepant in terms of specific rates of the various subcategories of anxiety disorders, they are unanimous in pointing to ubiquity and significance of anxiety problems in general throughout childhood and adolescence. Clearly, anxiety problems emerge as the most common form of psychological distress when the children and adolescents are asked themselves.

5.13.2.4 Etiology

It would be impossible to review all theories of all types of childhood anxiety in this chapter. Thus, we limit the scope in order to cover the major etiological positions that have been translated into contemporary clinical practice. These include psychodynamic, social learning, family systems, and biological theories of anxiety.

5.13.2.4.1 Psychodynamic

Freud's (1909) description of a phobic disorder in a five-year-old boy was the first time psychoanalysis was applied to a childhood disorder, and remains a seminal documentation of the psychodynamic approach to anxiety disorders. The child had phobias of horses and of leaving the family home and Freud interpreted these as stemming from Oedipal conflicts. The boy had witnessed a horse fall violently in the street. To Freud, the horse's violent struggles brought the boy's unconscious wishes to be violent toward his father to the conscious mind, thus provoking considerable anxiety. Thus, in the psychoanalytic approach the feared stimuli are seen as being secondary to the internal stimuli that they represent. Any number of unconscious conflicts can be associated with anxiety problems; the Oedipal conflict is but one example. The common element is that anxiety results from the conscious mind being overwhelmed by uncontrollable, threatening stimuli. While these stimuli can emanate from external sources, these rarely cause clinical anxiety states. Anxiety that arises from uncontrollable, threatening stimuli coming from the unconscious mind is more likely to cause disabling clinical anxiety states. For the child, focusing on external stimuli as feared objects is seen as functional in eliminating the true source of the anxiety from consciousness.

Uncontrolled case studies have claimed effectiveness using psychodynamic methods with anxious children (e.g., Le Roy & Derdeyn, 1976). However, controlled studies have produced equivocal results (e.g., Milos & Reiss, 1982) and, overall, these methods have received little empirical support. Other problems with the psychodynamic approach have been thoroughly discussed elsewhere (e.g., King, Hamilton, & Ollendick, 1988) including difficulty evaluating theories and therapies because of a lack of specificity of definitional constructs, and the length of time needed to implement treatments. Notwithstanding, there are aspects of psychodynamic theories of anxiety that warrant recognition and further attention. First, they alert us to the common developmental aspects of fear. That is, certain fears and anxieties come to the forefront of experience at specific biologically determined phases of life. For example, separation from caregivers is one of the first fear-provoking situations that both children and parents must successfully negotiate. Second, many contemporary psychodynamic models emphasize the role of interpersonal processes in the development and maintenance of anxiety states. For example, Kohut's (1971) and Mahler's (1968) work on object–self relations emphasize the importance of early parent–child relationships in the development of stable and robust self structures that promote psychological health. Bowlby's work on attachment processes provides a well-developed theoretical and methodological model describing attachment/separation processes throughout life development, and will be discussed later in this chapter. Third, in terms of processes of internalization leading to healthy self structures, contemporary psychodynamic models share much in common with modern information-processing models of anxiety in their emphasis on intrapsychic, even unconscious, processes that lead the sufferer to attend to and overinterpret threat stimuli, both internal (e.g., physiological arousal, dangerous thoughts) and in the external environment (e.g., criticism) (see Brewin, 1987).

5.13.2.4.2 Biological

A good deal is known about the neurophysiology and endocrinology of immediate "fight or flight" reactions to noxious stimuli (Gray, (1988). However, very little is known about the biological mechanisms involved in anxiety states that are more diffuse over stimulus conditions and time. At present, there are several lines of inquiry that may shed light on possible biological bases of anxiety. A recent review by Rutter et al. (1990) of genetic transmission studies indicated that a familial loading is evident for adult anxiety disorders. However, it is impossible to discern from available evidence whether the transmission is via genetic or environmental factors. For childhood anxiety disorders, the picture is even less clear and less adequately researched. While there appears to be a familial transmission of anxiety disorders in children, it is not clearly a genetic phenomenon and it does not show disorder specificity. That is, anxious children often have parents with a range of psychiatric disturbance, and anxiety problems in parents are associated with a similarly broad range of disturbance in their children (Rutter et al.).

Research has indicated that vulnerability to anxiety varies across individuals and is stable across time. With rhesus monkeys, Suomi (1986) has shown that certain individuals are more likely than others to respond to "challenge" situations with fear and behavioral inhibition. This vulnerability appears to be traitlike, existing at relatively high frequencies in genetic as opposed to environmental (adopted) siblings, and stable over time. Further, Suomi, Seaman, Lewis, Delizio, and McKinney (1978) demonstrated that imipramine reduces these fear reactions in vulnerable individuals during separation from parental figures. Similarly, Kagan, Reznick, and Snidman (1988) have shown that the traits of "behaviorally inhibited" infants are stable over time and consistently emerge when the infant is placed in a stressful situation. More recently, Biederman et al. (1990) compared two samples of children who had been classified as either inhibited or not inhibited/uninhibited, with a view to investigating whether behavioral inhibition is associated with anxiety in children. The results indicated that a wide range of internalizing and externalizing disorders were more prevalent in inhibited children compared with the not inhibited and healthy control children. Although this study was exploratory, the data suggest that behavioral inhibition to the unfamiliar may be one of the multiple risk factors implicated in the development of childhood anxiety disorders. More recent research has examined factors that may determine which behaviorally inhibited children are at risk of developing childhood anxiety disorders. Hirshfield et al. (1992), for instance, demonstrated that the association between childhood anxiety disorders and behavioral inhibition appears to be accounted for by those children who have stable behavioral inhibition (i.e., children who remained behaviourally inhibited at each of the follow-up points). More specifically, Rosenbaum et al. (1992) reported that greater anxiety loading in the parents increased the risk for anxiety disorders in behaviorally inhibited children.

A number of theories of anxiety that attempt to bridge behavioral and biological dimensions have been put forward, particularly with regard to inherited autonomic reactivity and conditionability (e.g., Gray, 1988); however the nature and function of biological mechanisms underlying anxiety disorders remain elusive and controversial. Further research cooperation between biologists and behaviorists is sorely needed.

5.13.2.4.3 Learning mechanisms

Modern behavioral conceptions of anxiety disorders emphasize conditioning and social learning processing in both the origins and maintenance of anxiety (Davey, 1992; Graziano, De Giovanni, & Garcia, 1979; King et al., 1988; Morris & Kratochwill, 1983; Mowrer, 1960; Thyer & Sowers-Hoag, 1988). Classical conditioning processes are invoked to explain how previously neutral stimuli (e.g., public transport) become threatening through association with unconditioned aversive stimuli (e.g., being lost, abandoned, fainting, or dying). Avoidance behavior, which is negatively reinforced by the reduction in anxiety it produces, maintains the fear by not allowing any natural extinction of the fear response to occur. Operant factors, such as reinforcement from significant others and avoidance of unpleasant responsibilities may also help maintain the fear. Such learning can be direct or vicarious, that is, acquired simply by observing the person's experience of fear, and is cognitively mediated. Cues that elicit fear can be regarded as information that is processed by individuals. Thus, cognitive processing mechanisms can be seen as an important link in the fear response. In adults, measures of cognitive processing of threat information discriminate between people with anxiety disorders and others (Rapee, 1991), and more recent research has indicated similar findings with anxious children (Barrett, Rapee, Dadds, & Ryan, 1996; Vasey, 1993). Anxiety disorders, especially the social types, may be in part maintained by social skills deficits in the child. If social interactions are repeatedly unsuccessful or unpleasant for the child, they are increasingly likely to avoid or to feel anxious in such situations. Hence, social skills interventions have become an important part of behavioral approaches to child anxiety (Graziano et al., 1979; King et al., 1988).

5.13.2.4.4 Stressful events

Anxiety states in children can be associated with exposure to both acute and chronic stressors such as loss of a parent, divorce, school failure, sexual and physical abuse, and physical illness. However, the increased occurrence of environmental stressors appear to precede the onset of nearly all psychiatric disorders and even physical illness (Goodyer, 1990). Much of the effect of environmental stressors are mediated through their effects on parent–child relations. For example, many of the deleterious effects of divorce on children are associated with changes in daily routine, discipline practices, and parent–child communication (Emery, 1982). While it is clear that stressors can produce anxiety in children (e.g., Dadds & Powell, 1991) it is not clear whether this also applies to clinically significant anxiety disorders in children who already had some vulnerability or disturbance predating the stressor (Goodyer, 1990).

5.13.2.4.5 Family processes

In research up until the 1970s, reviewed by Hetherington and Martin (1979), mothers of anxious children were consistently described as domineering, overprotective, and overinvolved with the child. However, as Hetherington and Martin also noted, this image of the overprotective mother and the highly structured family system were based on "rather indirect sources of data" (p. 61). Empirical studies were more limited in the support they provided for family factors, but included elevated levels of anxiety in the parents of anxious children (Bernstein & Borchardt, 1991), and marital conflict as a correlate of anxiety in nonclinic children (Dadds & Powell, 1991). However, more recent samples of anxious children found they were not different from other groups on marital conflict or dimensions of family functioning measured by the FACES scale (emeshment, adapatability). Further, Green, Loeber, and Lahey (1992) found no evidence for the presence of deviant family hierarchies, as predicted by family systems theory (Haley, 1976; Minuchin, 1974), in the families of overanxious boys. A study by Stark, Humphrey, Crook, and Levis (1990) showed that anxious children saw their families as less supportive, less sociable, more conflictual, and more enmeshed than nonclinic children. However, the results were even more striking for depressed and mixed depressed/anxious children and so these factors are unlikely to be specific correlates of childhood anxiety.

More impressive data has come from direct observational studies of social learning processes. Krohne and Hock (1993) argued that a child's competencies are related to the parents' tendency to help the child develop problem-solving and coping skills, and, inversely, to overly control the child (restriction). The idea of parental overcontrol has often been related to anxiety problems in children. Solyom, Silberfeld, and Solyom (1976) reported that mothers of agoraphobic patients scored significantly higher on measures of maternal control than mothers of normals. Krohne and Hock (1991) found that the mothers of high-anxious girls were judged by independent observers to be more restrictive than mothers of low-anxious girls during a common problem-solving task. Similarly, Dumas, LaFreniere, and Serketich (1995) observed parent interactions among aggressive, anxious, and nondistressed dyads. The anxious dyads were characterized by high parental control and aversiveness. According to Krohne and Hock, consequence expectancies on the other hand are related to the consistency of positive and negative feedback from parents. Kohlmann, Schumacher, and Streit (1988) found significant positive associations between parental inconsistency and trait anxiety in the child.

A more recent study has indicated that parents of anxious children may differ from other parents in terms of the way they teach their children to interpret and respond to ambiguous threat cues. Barrett et al. (1996) demonstrated that anxious children and their parents make relatively high numbers of threat interpretations and, hence, choose avoidant solutions when faced with ambiguous hypothetical social problems. As part of their experimental procedure, Barrett et al. also asked the families of anxious, aggressive, and nonclinic children to discuss how their child should deal with these ambiguous situations. For anxious children only, the likelihood that the child would devise an avoidant solution increased after the family discussion. It appears as if the family process can facilitate expression of the child's vulnerabilities. In a follow-up paper, Dadds, Barrett, Rapee, and Ryan (1996) analyzed the contingent stream of family behaviors that had been videotaped in the family discussions in the Barrett et al. study. Results showed that parents of anxious children were more likely to reciprocate avoidant solutions and less likely to encourage prosocial solutions to ambiguous social situations, than parents of nonclinic and aggressive children.

5.13.2.4.6 Recent studies of attachment processes

Attachment refers to the establishment of early intimate relationships and the internalization of these relationships as stable cognitive representations of the relation of the self to intimates throughout the lifespan (Bowlby, 1969). In the late twentieth century a number of strategies for operationalizing attachment processes have been developed, including observational measures of parent–infant bonds (Ainsworth, 1989), and self-report and structured interview measures of current adult relationships and past relationships to parents (Main, 1996). A number of manifestations of insecure attachment have been described and empirically verified, including dismissive (or avoidant, e.g., avoiding or failing to seek out intimate contact) and preoccupied (or anxious/ambivalent styles, e.g., showing distress at separation, clinging, failure to show independent exploration). Using various methods of categorizing attachment, numerous studies have been published showing that different attachment processes characterize psychologically healthy versus distressed adults and children. Clinic-referred samples of children and adults show relatively low rates of secure attachment, that is, intimate relationships marked by predictable, stable, and generous levels of care and support, compared with nonclinic samples (Main).

In general, however, it has been difficult to find specific relationships between particular forms of psychopathology and specific types of insecure attachment (see van Ijzendoom & Bakermans-Kranenburg, 1996, for a meta-analytic review). Recent improvements in the design of studies may overcome this lack of specificity. For example, Rosenstein and Horowitz (1996) found that psychiatrically hospitalized adolescents with a "preoccupied" attachment style were much more likely to have anxiety and depression problems than those with a dismissive style, who tended to show conduct problems. Clearly these results are encouraging and careful research is worth pursuing in this area. It should be noted, however, that the area still tends to be characterized by reliance on global typology categories and absence of specific descriptions of behavioral processes mediating attachment and psychopathology. This is highlighted by the DeKlyen (1996) study which showed that maternal attachment added no predictive power to the child's behavioral problems once direct observational measures of parent–child interactions were considered. Gender effects also need to be considered and more effort is needed to include fathers in this area. Cowan, Cohn, Pape-Cowan, and Pearson (1996), for example, found that mothers' and fathers' attachment styles were differentially related to internalizing vs. externalizing problems in their children. Further, few treatment developments from an attachment perspective have been subjected to controlled outcome and process studies.

5.13.2.4.7 Family systems theories

While there are several varieties of theoretical approaches that fall under the general banner of family systems theory, the common thread is the notion that the family system can be viewed as a functional or homeostatic unit. Disturbance in one family member (the "identified patient") is seen as part of a disturbed family system (Barnes, 1985). At extremes, the identified patient's problems are seen as a symptom that serves to maintain homoeostatic balance in an otherwise dysfunctional family. As an example, a child's separation anxiety may be seen as providing a parent with the feeling that they are needed, thereby reducing the likelihood that marital conflict will erupt over a distant, nonintimate marital relationship. Family systems theories have been quickly translated into clinical practice and have arguably been the "boom" therapy in the late twentieth century. Unfortunately, research into the conceptual validity and treatment efficacy of these approaches has lagged dramatically behind their enthusiastic translation into clinical practice. Despite repeated calls for these approaches to be evaluated, very little empirical work has been undertaken, and thus it is difficult to support any further expansion of these family systems models at this time.

5.13.2.5 Developmental Course

As already noted, for the majority of children and adolescents, fears and anxieties are typically common, functional, and transitory. For this reason, both clinicians and parents often view childhood anxiety as a passing complaint (Keller et al., 1992). Unfortunately, for a significant proportion of all children and adolescents, this is not the case and the prognosis, if they do not receive treatment, is much less encouraging than has previously been thought. For instance, in one study, 70.6% of children aged six to 12 years who had been diagnosed with overanxious disorder were found to have retained this diagnosis at two-year follow-ups (Pfeffer, Lipkins, Plutchik, & Mizruchi, 1988). Keller et al. assessed past and present psychopathology in 275 children and adolescents aged six to 19 years who had been recruited as part of a risk study. Fourteen percent of the children were found to have a history of an anxiety disorder, and of these children only 34% were anxiety diagnosis-free at the time of the assessment. The average duration of the disorder to the time of interview was four years, but the Kaplan-Meier estimates indicated that 46% of the children with an anxiety disorder would still be ill eight years after the onset of the disorder. Keller et al. concluded by noting that "anxiety disorders in childhood may be more chronic than previously believed . . . anxiety disorders in children may be severe and enduring" (p. 598).

The chronicity of childhood anxiety would seem to be due, at least in part, to the tendency of anxiety to be associated with general social problems such as negative self-image, dependency on adults in social situations, comparatively poor problem-solving skills, unpopularity, and low rates of interaction with peers (Kashani & Orvaschel, 1990; Messer & Beidel, 1994; Panella & Henggeler, 1986; Rubin & Clark, 1983; Strauss, Frame, & Forehand, 1987). In clinical samples, anxious children are less successful in peer relationships than their nonreferred counterparts (Edelbrock, 1985; Puig-Antich et al., 1985; Strauss, Lahey, Frick, Frame, & Hynd, 1988). As a whole, this research indicates that anxiety in children and adolescents (in both clinical and nonreferred samples) is associated with significant psychosocial difficulties. Regardless of the casual direction of the relationship between anxiety and its psychosocial correlates, it makes sense to hypothesize that for many anxious children and adolescents, this relationship is one in which anxiety continually strengthens the psychological factors and vice versa, thus contributing to the chronicity of anxiety in children and adolescents who have not received treatment.

A final point worth making in this section is that although the connection between childhood and adult anxiety is not yet clearly understood, evidence from a variety of studies indicates that vulnerability to adult anxiety may manifest itself in childhood. In support of this contention, there is increasing empirical evidence that many adult anxiety disorders originally began in childhood or adolescence (Rapee & Barlow, 1992). Retrospective accounts of adult anxiety patients frequently describe the onset of anxious symptomatology in late childhood or adolescence (Gittelman, 1984; Mattison, 1992; Thyer, Parrish, & Curtis, 1985). Klein (1964), for instance, reported that at least half of a sample of severely impaired adult agoraphobic patients had had separation anxiety and difficulty adjusting to school during childhood, while Gittelman found that a notable proportion of hospitalized agoraphobic adults reported an early history of severe separation anxiety.

5.13.3 CONCEPTUALIZATION AND CLINICAL FORMULATION

5.13.3.1 Multimodal Assessment

Available assessment procedures in childhood anxiety disorders include the use of structured clinical interviews, self-report measures, direct behavioral observations, self-monitoring of behaviors and accompanying cognitions, and physiological assessment.

The structured clinical interview used most commonly with anxious children is the ADIS (Silverman & Nelles, 1988), with a form to be used for the child and a parallel form for the parents (ADIS-C and ADIS-P). The ADIS has questions relevant to the latest *DSM* categories and facilitates acceptable inter-rater reliability of anxiety diagnoses in children (see Rapee et al., 1994; Silverman, 1991). Many other structured interviews are available (see Silverman, 1991 for a review). It should be noted that the reliability achieved from these interviews varies considerably according to the source of the information, the skills of the interviewers, the age of the interviewees, the base rate of the target disorder, and other factors. Silverman (1991), in recognizing the variability of inter-rater reliability, advises researchers to speak of reliability levels in terms specific to particular cohorts rather than as a characteristic of particular disorders or interviews. With regard to the validity of diagnoses achieved from structured interviews, there is some evidence that convergence is acceptable with standardized checklists such as the CBCL (see below). Generally, however, little research has systematically addressed the many ways in which validity can be assessed, and, at present, the validity of the interviews must be considered unknown (Silverman, 1991).

A variety of self-report measures is available. The ones that have been most utilized both in research and clinical work include the Revised Children's Manifest Anxiety Scale (RCMAS) (Reynolds & Paget, 1981), the Revised Fear Survey Schedule for Children (FSSC-R) (Ollendick, 1983), the modified State-trait Anxiety Inventory for Children (STAIC) (Fox & Houston, 1983), and parent/teacher/child ratings of general adjustment across multiple dimensions of dysfunction on screening checklists such as the CBCL (Achenbach & Edelbroch, 1991). None of the anxiety checklists produces data that correspond to the current *DSM* or *ICD* categories. The FSSC-R is used for assessing the content and spread of children's fears, the RCMAS and the STAIC produce measures of felt anxiety symptoms (e.g., physiological, worry). While the reliability of these checklists is generally good, recent research has questioned their discriminative validity. That is, they cannot discriminate reliably between different groups of clinically referred children and the differentiation of anxious and attention deficit children seems particularly problematic (Perrin & Last, 1992). Further development of self-report measures is needed, exemplified by Spence's (1996) recent attempts to develop a self-report measure with symptom domains that correspond to current diagnostic categories.

It is possible to measure fear/anxiety indirectly by measuring the behaviors that people typically engage in when faced with specific threat-related stimuli. For example, measuring the amount of time a child spends at school will provide an indirect estimate of fear of school. Used under controlled conditions, these measures are usually referred to as behavioral avoidance or approach tests (BAT). The aim of the BAT is to collect objective information about how the person reacts (dependent variable) to fear stimuli (independent variable). For example, a person who has a crippling fear of dogs may be asked to approach and stroke a dog under controlled conditions (the dog is restrained). A range of dependent measures can be taken including proximity to the dog and other behaviors hypothesized to reflect the anxiety level. BATs can be standardized or individualized. The former are most commonly used for research purposes and involve a standard series of tasks to which all subjects are exposed. In an individualized BAT, the stimuli to which each child is exposed are selected on the basis of the child's own specific fears. Naturally, this would be more useful in a clinical situation but it is also valuable in research since it would maximize the fear ratings achieved.

There are many variations on the possible uses of BATs and we present a sample to give some indication of the adaptability of these tests in their use with children. Lang and Lazovic (1963) asked subjects to approach a harmless caged snake in a laboratory and scored the subject's behavior on a three-point scale corresponding to looking, touching, and holding behavior. In other studies, the subject was asked to remain still while the feared stimulus was brought closer in graded steps by an experimenter (Murphy & Bootzin, 1973). In these tests, the dependent measure is usually the distance from the child before the child conveys that he or she does not want the feared stimulus to come any closer. Although the child is always given the power to terminate these tests, active BATs are usually preferred owing to the unpleasantness and artificiality of asking a fearful child to remain passive while another person controls the feared stimulus, and to the advantages of practicing approach behavior that comes with the use of active BATs.

While most BATs have been used with animal phobias, other creative uses have been reported, including fear of blood (Van Hasselt, Hersen, Bellack, Rosenblum, & Lamparski, 1979), the dark (e.g., Giebenhain & O'Dell, 1984; Kelley, 1976), heights (Van Hasselt et al., 1979), medical procedures (Freeman, Roy, & Hemmick, 1976), water (Lewis, 1974), and strangers (Matson, 1981). In social fears and generalized anxiety conditions, it is difficult to specify the phobic stimuli in such a way as to allow their manipulation under controlled conditions. For example, a child may avoid school owing due to a fear of having to interact with large groups of unsupervised children as would be found in a school playground. Children with generalized anxiety are fearful of new events in their lives such as taking tests, going to camps, changing classes, or having to perform, which are all difficult to operationalize for use with BATs. This limitation is important and should be the attention of creative thinking and research. In the late twentieth century, anxiety research with both adults and children has increasingly focused on the more generalized forms of anxiety disorder (generalized anxiety, overanxious disorder, panic, social fear) rather than the simple of specific phobias. Owing to the difficulty of operationizing the objects of these fears, the salience of BATs in the research literature has decreased and progress in the development and refinement of these measurement strategies has slowed to a crawl.

A number of variables may affect the validity of BATs (Borkovec, Weert, & Bernstein, 1977). Clinicians and researchers rarely use existing procedures, and thus, comparison across studies is difficult. Furthermore, a BAT is a social situation in which many variables can influence the child's behavior and yet, very little research has been conducted into the role of contextual variables (cf., Kelley, 1976; Sheslow, Bondy, & Nelson, 1982). More attention needs to be given to contextual factors in childrens' performance in fearful settings. This should be done not just so we can control these as unwanted influences. Rather, we should try to understand the role of contextual factors as a means of further understanding the developmental origins of fear itself. The BAT, as well as being used as a standardized procedure for producing fear responses, could be used for assessing the role of social influences, such as the family, on childrens' anxiety.

Direct observations of childrens' anxiety have typically been designed to record discrete child behaviors in natural settings that elicit anxiety such as during medical procedures, public speaking, or school exams. Detailed reviews of these systems can be obtained from Barrios and Hartmann (1988), Dadds, Rapee, and Barrett (1994), and King et al. (1988). The Observer Rating Scale of Anxiety (Melamed & Siegel, 1975) was designed for observing children undergoing surgery and consists of 28 discrete child behaviors which the observer scores for frequency of occurrence. These include behaviors such as "trembling hands" and "clings to mother," as well as more general distress behaviors such as "crying" and "kicks." Similarly, the Behavior Profile Rating Scale (BPRS) (Melamed, Yurcheson, Fleece, Hutcherson, & Hawes, 1978) contains 27 behaviors for use in dental settings and, thus, includes relevant behaviors such as refusing to open one's mouth. For both systems, the frequency of discrete behaviors can be combined into an overall anxiety score, although it should be noted that the overall score reflects distress and noncompliant behavior as well as anxious behavior. Further, the BPRS contains items that are scored according to the dentist's behavior, for example, the "dentist uses loud voice." Reliability and discriminative validity data have been collected and are supportive of the systems. Another observational system that has been applied to stressful medical procedures and undergone extensive development and evaluation is the Procedure Behavioral Rating Scale and its derivatives (Jay & Elliott, 1984; Jay, Ozolins, Elliott, & Caldwell, 1983; Katz, Kellerman, & Siegel, 1980; Le Baron & Zeltzer, 1984). These systems involve an observer rating the frequency of a number of child fear and distress behaviors during bone marrow procedures.

Glennon and Weisz (1978) developed a system for observing anxiety in preschool children. Their main rationale was to develop a measurement system for use with young children that avoided problems of self-report and global adult ratings. The Preschool Observation Scale of Anxiety (POSA) contains 30 behaviors that were selected to reflect fearful behavior in young children. These include both direct signs of fear ("child complains of being afraid," "trembling," "fearful facial expression"), indirect signs of fear ("cry," "gratuitous movement," "nail biting"), and other behavior selected to reflect fear in this population ("masturbation," "silence," "rigid body posture"). Glennon and Weisz evaluated the POSA in observations of preschool children's behavior during a forced separation from their mothers at their regular preschool. However, the system could be used in any setting. Inter-rater reliabilities ranged from quite low to very high, as did the frequency of the observed behaviors themselves, and further work may be required to refine the sample of 30 behaviors. The POSA represents one of the few attempts to operationalize child anxiety into an observation system that can be used in a range of natural settings. Milos and Reiss (1982) used an observation system that scores the frequency of a single global behavior as a reflection of anxiety during separation. Neisworth, Madel, and Goeke (1975) designed a system that measures the duration of crying and screaming in preschoolers during their daily preschool routine.

By far the most research and theorizing that has occurred with regard to separation anxiety has occurred within the social attachment literature. The original impetus for this area came from the intensive observations of Harry and Margaret Harlow, using infant rhesus monkeys (Harlow & Harlow, 1962), Bowlby's (1973) observations of orphaned children, and Ainsworth's observations of separated infants (Ainsworth, Blehar, Waters, & Wall, 1978). Typically, the child is observed during separation from the mother, exposure to an unfamiliar adult, and reuniting with the mother (Ainsworth et al., 1978). Research by Kagan and colleagues has shown that the infant's behavior during such challenges is predictive of the later development of anxiety and social withdrawal problems. Specifically, infants that show "behavioral inhibition" (long latency to interact with and retreat from unfamiliar adults, cessation of play and clinging to the mother) are more likely to develop social anxiety and withdrawal later in childhood (Kagan, Reznick, & Snidman, 1988).

A number of systems have also been designed to observe children's responses to test performance and public speaking. For example, Wine (1979) developed an observational system that comprises 22 discrete behaviors that are recorded over a 15-minute interval immediately preceding a classroom examination. The behaviors are grouped into attending, task-related, activity, communication, and interactional behaviors. Few of these are directly reflective of the expression of anxiety but rather focus on the indirect effects of anxiety on the child's ongoing activities. Paul's (1966) timed behavior checklist was designed to measure the presence of overt anxiety in adolescents in public-speaking situations. It has been adapted for use with younger children's reactions to public speaking (Fox & Houston, 1981) and darkness (Giebenhain & O'Dell, 1994). The original contains 20 items that are recorded for occurrence or nonoccurrence over a series of observational intervals and can be reduced to an overall anxiety score for the observation session.

Another method for measuring anxiety is to have observers use global rating scales. As examples, an observer may be asked to observe a child for a five-minute period and then rate the child's level of anxiety on a five-point scale; and, a parent of a child with fear of school may be asked to provide a similar rating every minute for the 30-minute period at home immediately prior to leaving for school. These systems are generally easy to use, but have specific advantages and disadvantages. First, research has indicated that people are generally very consistent in their observations and descriptions of the effect in other people. Even across different cultures, people tend to agree on the facial expressions that signal fear (Ekman, 1980). Thus, inter-rater reliability is generally quite high for global ratings of anxiety. A number of studies of anxious children have used such global ratings and the reader is referred to the reviews by Barrios and Hartmann (1988) and King et al. (1988). The problem with global ratings is that it is not clear exactly what is being measured and thus they generally have little to contribute to our understanding of anxiety. This restricts their use to situations in which an inexpensive and simple measure of anxiety is needed, and as with all single measures, they should be used as part of a more comprehensive assessment package (King et al., 1988).

The development of observational strategies for measuring child anxiety mostly took place in the 1960s and 1970s and, since that initial burst of activity, little has occurred in terms of the development and refinement of these methodologies. Most existing observational strategies focus solely on the child. Creative attempts to operationalize the social context of the child's anxiety and examine its functional relationship to the development of anxiety are sorely needed. Questions remain unanswered about the validity of these systems in terms of the content of the behavioral categories used. With children, most of the observational systems contain a mix of direct expressions of anxiety as an effect, as well as protest behaviors that are thought to result from anxiety. Using observation systems that lump all effect and protesting behaviors together to produce one overall "anxiety" score will not differentiate between these children. More work is needed to examine whether it is psychometrically and conceptually sound to treat these various components as unitary reflections of anxiety; quite possibly they are functionally very different when one considers how they affect and are affected by their social context. The discriminative validity of these observational systems remain unclear. Very little evidence is available to indicate that these systems can differentiate anxious from nonanxious children. Further, there is no evidence to show that children with different types of anxiety problems can be differentiated. Finally, the targets of these systems are mainly restricted to simple fears that are expressed in highly structured settings. Little work has been done attempting to use observational measures with children with complex, generalized anxiety.

We believe that there is enormous unrealized potential in the use of direct observations of anxiety in young children. Two major areas for development suggest themselves. There is the need to understand the child's anxiety in terms of its social context. Thus, observation systems that score both the child's anxiety as well as other antecedent and consequent stimuli need to be developed. Also, observations need to be designed and used in a way that allows for conceptual integration with other aspects of anxiety such as cognitive processing style, family interaction, and physiological reactivity. For example, much progress is currently being made in understanding the way anxious people process information about threat in their environment (e.g., Butler & Mathews, 1983). That is, anxious people show an exaggerated tendency to perceive, attend to, and respond to threat in their environments. We may ask whether this processing style is learned through interaction with significant others, namely anxious parents. A methodology that integrates information processing and direct observation of the child in interaction with others may yield important insights into this process. The work by Krohne and colleagues (Krohne, 1992; Krohne & Hock, 1991) and Barrett et al. (1996), reviewed above, are examples of fruitful attempts in this direction.

5.13.3.2 Treatment Formulation

The treatment formulation represents the integration of the theory of etiology, a theoretically guided and multimodal assessment of the individual case, and the existing literature on the efficacy of available intervention strategies. While this will in part depend on the skills and theoretical approach of the particular clinician, a number of guiding principles can be extracted from existing knowledge.

The conceptualization of the case should include a diagnostic formulation in terms of current *DSM* or ICD systems with particular reference to the developmental and cultural characteristics of the child and the normative expression of fears and anxieties that exist in all children. As well as the normative and cultural aspects of fear, preliminary formulations should consider the likelihood that the child's fears and anxieties are realistic in context. We have assessed a number of cases of "separation" and "social" fears in which the referred child was being exposed on a daily basis to threats of removal from the family home, and bullying at school, respectively. As previously stated, multiple informants should be used in gathering these data.

Factors hypothesized to be implicated in the development and maintenance of the anxiety problem should include a thorough examination of the following:

(i) Social learning factors: the extent to which the child experienced a learning history predictive of conditioned or traumatic responses to previously neutral stimuli, and what operant factors are currently maintaining a pattern of avoidance and anxious responding.

(ii) Cognitive factors: the extent to which the child shows a pattern of negative competency and outcome expectancies, low self-esteem, and over- or selective attention to potential threat in the environment.

(iii) Physiological factors: the extent to which the child exhibits physiological arousal (muscle tension, problem breathing) associated with avoidance and anxiety.

(iv) Family factors: the extent to which the child has stable, nurturing family relationships, and the extent to which the other family members have anxiety problems.

(v) Peers and social skills: the extent to which the child has the skills and opportunities to develop appropriate and pleasing peer relationships.

Using an assessment and therapy procedure that empowers the child and family in terms of acquired knowledge, skills, and decision making, the above factors should be summarized into a concise formulation and treatment plan that clearly prioritizes and addresses the relevant causal factors. The treatment plan will thus usually involve some combination of social learning interventions such as: exposure, modeling, and reinforcement for overcoming avoidance; cognitive skills for the child such as positive self-talk, developing positive expectancies, attention to nonthreat stimuli, and reinforcing oneself for success; relaxation, breathing retraining, and other skills for overcoming uncontrollable and unpleasant physiological arousal; a family intervention to establish other family members as positive models of courage and social competence, as sources of stable care and nurturance, and sources of reinforcement for success; and social skills and school interventions to facilitate positive peer relationships and remove any sources of (realistic) fear from the school setting.

The treatment formulation should specify clear goals for improvement that are graduated from easy to difficult so the child has early success and a growing sense of mastery. The active involvement of family members is important, especially for children in the preteenage years.

5.13.4 TREATMENT

5.13.4.1 Psychosocial Treatments

Psychodynamic and systemic therapies have been used in the treatment of childhood anxiety disorders; however, the lack of controlled studies prohibits conclusions about their usefulness. However, there have been some preliminary studies of psychodynamic treatments for childhood disorders. For example, although the studies by Milos and Reiss (1982) and Target and Fonagy (1994) do not allow conclusions about differential efficacy, they offer some hope for the use of empirical procedures with psychodynamic therapies. Hopefully, the early years of the twenty-first century will see these methods come under the scrutiny of careful research designs. In the following, we focus on cognitive-behavioral, family, and pharmacological treatments that have been evaluated, followed by a review of the main treatment outcome studies related to specific diagnostic groups.

In terms of cognitive-behavioral interventions, King et al. (1988) and Ollendick and Francis (1988) recommend an integrative treatment approach, encompassing exposure variants, modeling, operant, and cognitive-based procedures. Exposure techniques include the use of systematic desensitization proper (imaginal representation of the fear producing stimuli and the use of muscular relaxation as the competing inhibiting response), *in vivo* and imaginal exposure, and emotive imagery (the imagery component is used as the fear-inhibiting agent). An advantage in using these techniques is the clarity of procedural guidelines available to the therapist. Still controversial is the issue of what underlying mechanisms account for the success of exposure-based procedures. Originally, Wolpe (1958) explained it as reciprocal inhibition of anxiety through relaxation. However, numerous case studies have been reported wherein the relaxation had little effect on treatment outcome. More recent views conceptualize the changes during exposure treatments in terms of information-processing changes (Foa & Kozak, 1986). Implosion and flooding-type treatments are not usually used with children owing to the potential aversiveness of the procedures involved.

Modeling procedures include the use of films, *in vivo*, or participant-modeling variants. The child is prompted to imitate the performance of the model who is demonstrating nonfearful behaviors and is reinforced for closer approximation to the coping behavior. Modeling can provide an opportunity for both the vicarious extinction of a child's anxiety, and an acquisition of more accurate information and effective coping skills relative to the anxiety-provoking situation. The efficacy of modeling procedures in the treatment of childhood phobias and in the preparation of children for stressful events has been clearly documented (Ollendick & Francis, 1988).

Operant-based programs (based on contingency management principles) have been especially successful in the treatment of school refusal, socially avoidant children, and specific phobias (King & Ollendick, 1989). Operant procedures are based on the premise that fear and anxiety will reduce when nonfearful behaviors are reinforced and fearful behaviors are placed on an extinction schedule. Within this approach, it would not be considered necessary to use methods such as relaxation to reduce anxiety directly. The child is simply reinforced for approaching the feared stimuli and rewards for fearful behaviors (e.g., parental comforting, substitution of aversive activities by pleasant play time) are removed. The evidence for the efficacy of operant programs comes mainly from case studies with phobic children, and in general therapeutic success is enhanced when operant procedures are combined with other techniques (King et al. 1988; Ollendick & Francis, 1988).

Self-instruction training is the most commonly used cognitive-based procedure in the treatment of childhood anxiety (Graziano & Mooney, 1980; Meichenbaum & Genest, 1980; Meichenbaum & Goodman, 1971). It is aimed at changing the child's self-talk on the premise that self-doubting and negativistic self-talk increases anxiety and inhibits appropriate responses to feared stimuli. Basically, the child is taught to use competence-mediating statements in the presence of the anxiety-provoking stimuli. However, cognitive procedures alone are often not effective in overcoming anxiety reactions in children owing to the multiplicity of factors involved, such as environmental contingencies, skills competencies, individual developmental factors regarding awareness and identification of self-statements, and the strong physiological component of anxiety disorders.

The integration of cognitive and behavioral procedures in the treatment of childhood anxiety disorders seems to be largely supported in the literature. However, most of the studies previously described have focused on the treatment of specific phobias using single case designs. Evaluative studies on the treatment of separation anxiety, avoidant, and overanxious disorders were less evident. However, in recent years, a number of controlled trials have demonstrated the effectiveness of protocol-based psychosocial interventions that are applicable across various categories of child and adolescent anxiety disorders. Blagg and Yule (1984), working with a total of 66 children presenting with fear of school and separation from parents, examined treatment outcome associated with behavioral therapy, hospitalization, and psychotherapy plus home schooling. The behavior therapy program incorporated exposure to the school setting and contingency management implemented by both parents and teachers. Both at post-treatment and one year follow-up, separation anxiety symptomatology was absent in the behavioral therapy group but still present in the hospitalized (33.3%) and psychotherapy (94%) groups. Average duration of treatment was two weeks for behavior therapy, 45 weeks for hospitalization, and 72 weeks for psychotherapy plus homeschooling.

In the first controlled trial for mixed groups of anxious children, Kendall (1994) evaluated the effectiveness of a 16-session cognitive-behavioral therapy (CBT) program for a group of children with overanxious, separation, and social anxiety problems. The treatment centers around having the child develop an individualized FEAR plan: *F*, for feeling good by learning to relax; *E*, for expecting bad vs. good things to happen through positive self-talk; *A*, for approaching actions to take in the face of fear; and *R*, for rewarding oneself for efforts to overcome fear or worry. Compared with a waitlist control, the treated children showed clinically significant gains that were maintained at one year follow-up. Follow-up studies have replicated these positive outcomes (Kendall et al., in press) and shown that the treatment effects were maintained to follow-up assessments conducted an average of 3.5 years after treatment (Kendall & Southam-Gerow, 1996).

An issue raised by Kendall (1994) concerned the role of the family in these CBT programs. His recommendation was that active parental involvement is useful for positive child outcomes. More family systemic and psychodynamic theorists would argue that change in the family system would be critical to real change in the child's problem and, at extremes, that solely changing anxiety symptoms in the child without focusing on underlying family dynamics, would be counterproductive. This issue was addressed by Heard, Dadds, and Conrad (1992) in a study researching the effects of a cognitive-behavioral intervention for childhood anxiety on family variables. They treated three adolescents presenting with simple phobias using contingency management, exposure, and self-instructional training. A multiple baseline design across subjects was used. Results showed marked improvement (at post-treatment and three months follow-up) in anxiety both at overbehavioral and cognitive levels for all three adolescents. Despite sole focus on the child's anxiety symptoms, no negative or compensatory effects on family process, or individual or dyadic adjustment in other family members were observed. Thus, they found no evidence for the family systems hypothesis that treating anxious children using solely child-focused behavioral techniques leads to negative consequences for the rest of the family system.

A recent controlled trial assessed the efficacy of involving parents of anxious children in a CBT program. Barrett, Dadds, and Rapee (1996) compared a CBT intervention based on Kendall's (1990) anxiety management program to an intervention that included the CBT intervention plus a family intervention, again for a mixed group of overanxious, separation anxiety, and socially phobic nine to 14-year-olds. Parental sessions introduced parents-to-child management skills (reinforcement skills, planned ignoring, giving, and backing up clear instructions) and how to use these skills to manage their child's anxiety, explained what the children were learning in the CBT program and illustrated how parents could model and encourage the use of strategies learned in the session, and showed parents how they could use the same strategies, that is, Kendall's FEAR plan, to manage their own anxiety. Evaluation of the program at the end of treatment revealed that 61% of children in the CBT group no longer met a *DSM-III-R* diagnosis, compared with 88% in the combined treatment. Less than 30% were diagnosis free in the waitlist control group. At 12 months follow-up, the relative superiority of the CBT plus family condition was maintained. Thus, it appears that the modification of parental skills can have a significant effect on anxiety in children when combined with a CBT program for the child. However, it is not clear whether a change in parental skills on its own will result in an improvement in child behavior, and it is not clear which aspect of the parental intervention was associated with the extra improvements.

Barrett (1995) showed that similar success rates could be achieved by presenting the combined CBT-family treatment in a group format to anxious children and their parents, thereby significantly reducing costs of intervention. Cobham and Dadds (1995) used the same group intervention to assess the role of parental anxiety in treatment outcome, and the extent to which the second component of Barrett et al.'s (1996) family treatment (parents skills for managing their own anxiety) could alleviate putative poorer treatment outcomes associated with high parental anxiety. Results indicated that high parental anxiety was a risk factor for poorer treatment outcomes for anxious children, and that specifically targeting parental anxiety for intervention could overcome this risk factor in the context of a cognitive-behavioral program for the child.

The above brief review indicates that anxiety disorders in late childhood and early adolescence can be effectively treated in the majority of referred cases. However, early intervention and prevention programs aimed at larger cohorts of children have the potential to be even more cost effective in reducing the incidence of childhood disorders and their cost to the community. A number of authors (e.g., King, Hamilton, & Murphy, 1983; Spence, 1994) have discussed the potential of such programs for children at risk for anxiety disorders. The Queensland Early Intervention and Prevention of Anxiety Project (Dadds, Spence, Holland, Barrett, & Laurens, 1997) utilized a school-based sample of 1786 seven- to 14-years olds. 9.7% of children were identified by teachers as having anxiety problems. A separate group (10.5%) scored 20 or above on the RCMAS. Only 2.0% were both nominated by teachers and scored above 19 on the RCMAS. *Post hoc* analyses showed that the low convergence between teachers' and children's reports may have been due in part to the tendency of some children not to report their anxiety, possibly owing to social desirability factors.

At the conclusion of a diagnostic interview, those children who fell within the "at risk" selection criteria were allocated to the intervention or monitoring condition on the basis of treratment vs. monitoring schools. The intervention was based on the Barrett study (1995) except that it was presented in group format in 10 sessions. On average, children attended over eight of the 10 child intervention sessions, mothers attended 1.7, and fathers 0.63 of three parent sessions. At pretreatment, approximately 75% of children interviewed met criteria for an anxiety diagnosis. At postintervention, improvement was noted for both intervention and monitoring groups. However, children who received the intervention emerged with lower rates of anxiety disorder at six months follow-up compared with those who were monitored only. Of those who had features of disorder, but no full disorder at pretreatment, 54% progressed to a diagnosable disorder at six months follow-up in the monitoring group compared with 16% in the intervention group. These results indicated that the intervention was successful in reducing rates of disorder in children with mild to moderate anxiety disorders, as well as preventing the onset of anxiety disorders in children with early features of a disorder.

5.13.4.2 Pharmacological Treatments

Medication is rarely used as the sole treatment for anxiety disorders in children and adolescents, but instead is typically one part of a treatment program (Allen, Leonard, & Swedo, 1995; Bernstein, 1994). Psychopharmacological interventions tend to be implemented more commonly with older children and adolescents, and also in cases where anxiety symptoms are particularly severe and resistant to other forms of treatment. Unfortunately, there are comparatively few studies that have attempted to evaluate the efficacy of psychopharmacological interventions with anxious children and adolescents. In addition, methodological shortcomings often make results equivocal or difficult to interpret, including small sample sizes, use of samples with mixed diagnoses based on noncurrent diagnostic systems (McDaniel, 1986), concurrent treatments which are not controlled, absence of placebo control groups, and inadequate medication dosages (Bernstein, 1994). These limitations must be kept in mind when evaluating outcome studies for each class of anxiolytics to be reviewed below.

A number of studies have suggested the possible benefit of benzodiazepines in the treatment of anxiety in children and adolescents. Simeon and Ferguson (1987) found that seven of 12 participants (aged eight to 16 years) with a diagnosis of avoidant and/or overanxious disorder manifested at least moderate improvements in terms of clinician-rated symptomatology when they received between 0.5 and 1.5 mg alprazolam (tradename Xanax) per day over a four-week period. Self-report, however, did not depict any improvement in anxiety symptoms. A subsequent, four week, double-blind, placebo-controlled study with a group of 30 children and adolescents (aged eight to 16 years) diagnosed with either avoidant or overanxious disorder and treated with a mean dose of 1.6 mg alprazolam found greater improvement on global improvement ratings compared with participants in a placebo group, although these results were not statistically significant (Simeon et al., 1992). More recently, Graae, Milner, Rizzotto, and Klein (1994) conducted a double-blind crossover trial with 12 children (aged seven to 13 years) with various anxiety diagnoses, in which four weeks of placebo were compared with four weeks of clonazepam (tradename Clonopin; up to 2 mg per day). Although 50% of the children no longer met diagnostic criteria at the end of the trial, clonazepam was not found to produce superior results to the placebo on clinical rating scales completed at the end of the study. In other uncontrolled studies, benzodiazepine treatments have been associated with a decrease in the frequency and/or intensity of symptomatology in children and adolescents with panic disorder (Kutcher & MacKenzie, 1988), and pediatric cancer patients prior to anxiety-provoking medical procedures (Pfefferbaum et al., 1987).

Clearly, the limited research that has examined the efficacy of benzodiazepine treatment of childhood anxiety is difficult to interpret. Less equivocal is the research on side effects related to this class of anxiolytics. The consensus in clinical studies appears to be that benzodiazepines are tolerated by children with minimal side effects (Bernstein, Garfinkel, & Borchardt, 1990; Biederman, 1987; Pfefferbaum et al., 1987; Simeon & Ferguson, 1987; Simeon et al., 1992), and that when side effects do occur, they tend to be dose-related (Kutcher, Reiter, Gardner, & Klein, 1992).

Reported side effects include: ataxia (unsteady gait), diplopia (blurred/double vision), reduced mental acuity, sedation, slurred speech, tremor, drowsiness (Biederman, 1991; Kutcher et al.), and behavioral disinhibition as evidenced by oppositional behavior, liability, and irritability (Graae et al., 1994). Although this class of medication seems to be comparatively safe in overdose (Kutcher et al.), as Biederman (1991) notes, the risks associated with tolerance and dependence in children treated with benzodiazepines are unknown at this time.

As is the case with the benzodiazepines, studies of the efficacy of tricyclic antidepressants with anxious children have produced conflicting results. Four placebo-controlled studies have investigated the outcome of tricyclic antidepressant interventions in the treatment of separation anxiety or school refusal associated with separation anxiety. In the first of these, Gittelman-Klein and Klein (1971, 1973) conducted a six-week trial in which 45 children were assigned randomly to either 100–200 mg of imipramine (tradename Tofranil) per day or a placebo group. Both groups also received behaviorally oriented psychotherapy. Compared with the placebo group, the imipramine group was superior—both in terms of the number of children returning to school and symptom improvement. Using an average dose of 153 mg of imipramine per day, Klein, Koplewicz, and Kanner (1992) were unsuccessful in their attempt to replicate these results in 20 children (aged six to 15 years) with separation anxiety. In another recent study, Bernstein et al. (1990) reported that there were no significant differences among groups of children and adolescents (aged seven to 17 years) who received either alprazolam (mean dosage of 1.8 mg per day), imipramine (mean dosage of 164.3 mg per day), or a placebo. Similarly disappointing results have been reported for the use of clomipramine (tradename Anafranil) in facilitating either a return to school or a reduction in anxious symptomatology with 51 anxious children and adolescents aged nine to 15 years (Berney et al., 1981). Although only one of the four studies reviewed above provides support for the efficacy of the tricyclic antidepressants in the treatment of separation anxiety (with or without school refusal), as Bernstein (1994) notes, one of the three negative studies used an arguably subtherapeutic medication dosage (Berney et al.), while the remaining two had very small sample sizes (Bernstein et al.; Klein et al.).

The most frequent side effects of using tricyclic antidepressants with children and adolescents include blurred vision, sedation, orthostatic-induced lightheadedness, dry mouth, urinary retention, and constipation. More uncommon side effects include jaundice, agitation, lethargy, heart rhythm problems, sleep disturbance, and allergic reactions. An additional problem associated with use of the tricyclic antidepressants is that overdosage may result in severe medical complications, including delirium, seizures, and heart rhythm problems.

Recently, there has been some suggestive evidence that buspirone (tradename Buspar) may be effective in the treatment of generalized anxiety in children and adolescents. In an uncontrolled case study, Kranzler (1988) reported that a 13-year-old boy with overanxious disorder and associated school refusal was successfully treated with 10 mg per day of buspirone. More recently, however, Kutcher et al. (1992), in a study of adolescents, diagnosed with either overanxious or generalized anxiety disorder, reported that following six weeks of buspirone therapy (between 15 mg and 30 mg per day), the adolescents' ratings of anxiety on self-report measures had decreased significantly. On the basis of these studies, Allen et al. (1995) suggested that controlled trials of buspirone with this clinical group may be merited. In addition to this suggestive evidence regarding its possible efficacy in the treatment of anxious children and adolescents, buspirone is also noteworthy for its lack of major side effects, its limited potential for abuse, and its low probability of producing withdrawal symptoms following cessation (Kutcher et al.).

Fluoxetine (tradename Prozac) has shown promise in the treatment of childhood obsessive compulsive disorder (Riddle, Hardin, King, Scahill, & Woolston, 1990; Riddle et al., 1992). More recently, however, an open study examining the efficacy and safety of fluoxetine in the treatment of more generalized anxiety in children has been conducted by Birmaher et al. (1994). Twenty-one treatment-resistant children and adolescents aged 11–17 years with diagnoses of overanxious disorder, separation anxiety disorder, or social phobia received a mean fluoxetine dose of 25.7 mg per day for an average of 10 months. At the end of the trial, 17 of the 21 subjects (81%) showed moderate to marked improvement in their anxiety symptoms. Combined with the fact that there were no significant side effects, these results prompted the authors to suggest that fluoxetine may have potential with nondepressed children and adolescents with anxiety disorders other than obsessive compulsive disorder and panic disorder (Birmaher et al., 1994)

As Coffey (1990) notes, the efficacy of beta-blockers in children and adolescents suffering from anxiety has yet to be clearly established, and as a consequence, this class of drug is not commonly used with children. There is, however, some evidence that beta-blockers may be effective in blocking the physiological symptoms of arousal typically associated with anxiety disorders. Famularo, Kinscherf, and Fenton (1988) conducted a four-week trial of propranolol (tradename Inderal; target dosage of 2.5 mg kg^1 per day) with 11 children (mean age of 8.5 years) with post-traumatic stress disorder. The results indicated that compared with a placebo, the active medication condition was associated with a significant improvement in symptoms, and that this improvement disappeared when the medication was discontinued. Side effects of the beta-blockers include constipation, sedation, nausea, hypotension, diarrhoea, bradycardia, increased airway resistance, vomiting, and depression. Kutcher et al. (1992) caution that this class of anxiolytics should be used only with treatment-resistant patients who are experiencing situational anxiety, and that if the beta-blockers are to be considered for use, the child or adolescent for whom they have been prescribed should be strictly monitored by medical staff.

Overall there is a lack of well-controlled and well-conducted research study into psychopharmacological interventions, and there is little support for them as sole treatments for anxiety disorders in children and adolescents. There may be a place for them in the treatment of this clinical population as an adjunct to other programs (such as behavioral, cognitive-behavioral, or psychotherapeutic interventions) but more research is needed. Of the various classes of anxiolytics reviewed here, the benzodiazepines have received the most empirical support, thus warranting their position as the medication most commonly prescribed for anxiety-disordered children and adolescents.

5.13.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

So much needs to be done in the area of childhood anxiety disorders that we will limit our suggestions to a few broad areas. First, the relationship between child development and anxiety disorders remains relatively unexplored and the developmental course of anxiety problems needs to be clarified using longitudinal studies of large cohorts. Second, the validity of current diagnostic systems is uncertain and this makes definitive research into etiology and prognosis difficult. Third, measures that can clearly differentiate between anxiety and other childhood disorders, and between different types of child anxiety disorder are sorely needed. Fourth, experimental studies that integrate different putative causal mechanisms are needed. That is, studies need to measure social learning, cognitive, interpersonal, and biological aspects of anxiety concurrently before more comprehensive models of etiology can be developed and tested. Each of these aspects of anxiety seems important, but with researchers only considering one or two at a time, it is difficult to discern their relative and interactive importance in etiology. Fifth, evidence is emerging that psychosocial treatments are quite efficacious with clinically referred anxious children. Further, a recent study indicated that such treatments have the potential to be used in school settings as preventive interventions. Further research is needed to explore the effectiveness of these treatments across a range of clinical and community settings. Sixth, existing research into pharmacological interventions is scarce, poorly controlled, and rarely assessed concurrently with psychosocial interventions. Their efficacy needs to be assessed as does their community effectiveness and social validity.

Finally, as noted earlier three theoretical models of the development of anxiety disorders have credible explanatory power and at least limited empirical support: social learning theory (SLT), attachment theory (AT), and biological models such as Kagan's inhibition model. There may be a number of potential benefits to be made by contrasting and possibility integrating their complementary strengths and weaknesses. For example, in its purest forms, SLT takes all stimuli as equal in potential to elicit fear and has little to say about the biological and development relevance of various stimulus settings. By contrast, AT recognizes that particular settings may be potent for learning fear vs. courage, such as the parents' and child's first experiences of separation and reunion. Conversely, AT places little emphasis on the microprocesses that establish and maintain behaviors, or the methodologies by which these microprocesses can be examined under controlled conditions. SLT has a rich and successful history in these domains. The two models make predictions that are seemingly contradictory, at least in terms of our current understanding with respect to the development and expression of fear. For example, SLT would predict that contingent parental soothing should reinforce fear displays, and punishment of fear responses should suppress them. By contrast, AT predicts that parental soothing will facilitate the child's skills in self-soothing, leading to a decrease in future fear displays, whereas punishment of fear responses will lead to an insecure attachment and an increase in fear behavior. This raises the exciting possibility that critical experiments could be designed to compare and contrast these predictions. Attempts at integration have potential beyond simply watering down each model to find facile similarities between the two. Interesting attempts in this direction have already begun: for example, Brewin's (1987) argument that Bowlby's attachment model is complementary to contemporary cognitive models of how self-structures are internalized in anxiety and depressive disorders appears fruitful.

5.13.6 SUMMARY

Humans have an inherent capacity to react to threat stimuli with fear and anxiety. The strength of this reaction varies across individuals and with the nature of the stimulus. Important aspects of anxiety disorders include developmental factors, in that normal developmental tasks predispose humans to different anxieties at different times of life. Within this developmental framework, considerable explanatory power is afforded by classical, operant, and social learning processes in both the origins and maintenance of child anxiety; and by variability in individual's susceptibility to react to stressful challenges with behavioral inhibition, anxiety and fear vs. disinhibition, impulsivity and aggression. The latter appear to be reasonably stable traits; however, the genetic and/or biological mechanisms underlying these processes are not understood. Recent developments in cognitive theory and research have shown that anxious people may process information in a way that increases the salience of threatening cues and catastrophic implications in their environment, but most of this work has been conducted with adults. With children, social contexts may be especially important to adjustment, and converging evidence suggests that high levels of control, restriction, attention to social threat, and endorsement of avoidant coping strategies are characteristic of families of anxious children.

Epidemiological studies indicate that anxiety disorders are the most common form of child and adolescent psychological problem. More tentative evidence indicates that these problems may, in their severest forms, be disabling and chronic and may have some continuity with adult anxiety problems. Anxiety is a very difficult phenomenon to measure in ecologically meaningful ways. To operationalize a person's ability to approach a snake in a laboratory is reasonably simple. However, to operationalize more complex forms of anxiety, such as generalized anxiety and panic, and then measure them in real-world situations is currently beyond our scope of knowledge and technology. Thus, the common but generalized forms of anxiety remain elusive to observational strategies, and the discriminative validity of existing self-report measures is poor. Progress has occurred in diagnostic systems but the validity of these remains controversial.

Social learning, cognitive, interpersonal, and biological models of etiology all seem to have some explanatory power but their relative and interactive importance is unknown as most research has considered them in isolation. Perhaps the most salient progress that has occurred over in the late twentieth century is the development and evaluation of cognitive-behavioral and family treatments. These appear to be highly efficacious with a broad range of clinically referred anxious children. The efficacy of drug treatment with any of the common childhood anxiety disorders remains a speculative issue. The broader community effectiveness and social validity of existing treatments remains unknown; however, there is some evidence that social learning therapies may have the potential to be used as early and preventive interventions in school settings.

5.13.7 REFERENCES

Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the child behaviour checklist and revised child behaviour profile* (2nd ed.). Department of Psychiatry, University of Vermont.

Ainsworth, M. D. S. (1989). Attachments beyond infancy. *American Psychologist, 44,* 709–716.

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation.* Hillsdale, NJ: Erlbaum.

Allen, A. J., Leonard, H, & Swedo, S. E. (1995). Current knowledge of medications for the treatment of childhood anxiety disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 976–986.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. rev.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anderson, J. C., Williams, S., McGee, R, & Silva, P. A. (1987). DSM-III disorders in preadolescent children: Prevalence in a large sample from the general population. *Archives of General Psychiatry, 44,* 69–76.

Barnes, G. G. (1985). Systems theory and family theory. In M. Rutter & L. Hersor (Eds.), *Child and adolescent psychiatry: Modern approaches* (2nd ed., pp. 216–229) Oxford: Blackwell.

Barrett, P. M. (1995, July). *Group treatment of anxious children.* Paper presented at the Fifth World Congress of Cognitive and Behavioral Therapies, Copenhagen, Denmark.

Barrett, P. M., Dadds, M. R., & Rapee, R. M. (1996). Family treatment of childhood anxiety: A controlled trial. *Journal of Consulting and Clinical Psychology, 64,* 333–342.

Barrett, P. M., Rapee, R. M., Dadds, M. R., & Ryan, S. (1996). Family enhancement of cognitive styles in anxious and aggressive children: The FEAR effect. *Journal of Abnormal Child Psychology, 24,* 187–203.

Barrios, B. A., & Hartmann, D. P. (1988). Fears and anxieties. In E. J. Mash & L. G. Terdal (Eds.), *Behavioral assessment of childhood disorders* (2nd ed., pp. 196–26). New York: Guilford Press.

Bauer, D. H. (1976). An exploratory study of developmental changes in children's fears. *Journal of Child Psychology and Psychiatry, 17,* 69–74.

Beidel, D. C. (1991). Social phobia and overanxious disorder in school-age children. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 545–552.

Berney, T., Kolvin, I., Bhate, S. R., Garside, R. F., Jeans, J., Kay, B., & Scarth, L. (1981). School phobia: A therapeutic trial with clomipramine and short-term outcome. *British Journal of Psychiatry, 138,* 110–118.

Bernstein, G. A. (1994). Psychopharmacological interventions. In T. H. Ollendick, N. J. King, & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in children and adolescents* (pp. 439–452). New York: Plenum.

Bernstein, G. A., & Borchardt, C. M. (1991). Anxiety disorders of childhood and adolescence: A critical review: *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 519–532.

Bernstein, G. A., & Garfinkel, B. D. (1986). School phobia: The overlap of affective and anxiety disorders. *Journal of the American Academy of Child Psychiatry, 25,* 235–241.

Bernstein, G. A., Garfinkel, B. D., & Borchardt, C. M. (1990). Comparative studies of pharmacotherapy for school refusal. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 773–781.

Biederman, J. (1987). Clonazepam in the treatment of prepubertal children with panic-like symptoms. *Journal of Clinical Psychiatry, 48,* 38–42.

Biederman, J. (1991). Psychopharcology. In J. M. Wiener (Ed.), *Textbook of child and adolescent psychiatry* (pp. 545–570). Washington, DC: American Psychiatric Press.

Biederman, J., Rosenbaum, J. F., Hirshfeld, D. R., Faraone, S. V., Bolduc, E. A., Gersten, M., Meminger, S. R., Kagan, J., Snidman, N., & Reznick, S. (1990). Psychiatric correlates of behavioral inhibition in young children with and without psychiatric disorders. *Archives of General Psychiatry, 47,* 21–26.

Birmaher, B., Waterman, G. S., Ryan, N., Cully, M., Balach, L., Ingram, J., & Brodsky, M. (1994). Fluoxetine for childhood anxiety disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 993–999.

Blagg, N. R., & Yule, W. (1984). The behavioral treatment of school refusal: A comparative study. *Behavior Research and Therapy, 22,* 119–127.

Borkovec, T. D., Weerts, T. C., & Bernstein, D. C. (1977). Assessment of anxiety. In A. R. Ciminero, K. S. Calhoun, & H. E. Adams (Eds.), *Handbook of behavioral assessment* (pp. 367–428). New York: Wiley.

Bowlby, J. (1969). *Attachment and Loss: Vol. 2. Separation, anxiety and anger.* New York: Basic Books.

Bowlby, J. (1973). *Attachment and loss.* (Vol. 2). New York: Basic Books.

Brady, E. U., & Kendall, P. C. (1992). Comorbidity of anxiety and depression in children and adolescents. *Psychological Bulletin, 111,* 244–255.

Brewin, C. R. (1987). *Cognitive foundations of clinical psychology.* London: Erlbaum.

Butler, G., & Mathews, A. (1983). Cognitive processes in anxiety. *Advances in Behaviour Research and Therapy, 5,* 51–62.

Clement, P. W., & Miller, D. C. V.(1967). Group play therapy and tangible reinforcers used to modify the behavior of eight year old boys. *Behavior Research and Therapy, 5,* 301–312.

Cobham, V. E., & Dadds, M. R. (1995, July). *Group treatment of anxious children: The role of parental anxiety.* Paper presented at the Fifth World Congress of Cognitive and Behavioral Therapies, Copenhagen., Denmark.

Coffey, B. J. (1990). Anxiolytics for children and adolescents: Traditional and new drugs. *Journal of Child and Adolescent Psychopharmacology, 1,* 57–83.

Costello, E. J. (1989). Child psychiatric disorders and their correlates: A primary care pediatric sample. *Journal of American Academy of Child and Adolescent Psychiatry, 28,* 851–855.

Cowan, P. A., Cohn, D. A., Pape-Cowan, C. P., & Pearson, J. L. (1996). Parent's attachment histories and children's externalizing and internalizing behaviors. *Journal of Consulting and Clinical Psychology, 64,* 53–63.

Croake, J. W. (1969). Fears of children. *Human Development, 12,* 239–247.

Croake, J. W., & Knox, F. H. (1973). The changing nature of children's fears. *Child Study Journal, 3,* 91–105.

Dadds, M. R., Barrett, P. M., Rapee, R. M., & Ryan, S. (1996). Family process and child anxiety and aggression: An observational analysis. *Journal of Abnormal Child Psychology, 24,* 715–734.

Dadds, M. R., & Powell, M. B. (1991). The relationship of interparental conflict and marital adjustment to aggression, anxiety and immaturity in aggressive and nonclinic children. *Journal of Abnormal Child Psychology, 19,* 553–567.

Dadds, M. R., Rapee, R. M., & Barrett, P. M. (1994). Behavioral observation. In T. H. Ollendick, N. J. King & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in children and adolescents* (pp. 349–364). New York: Plenum.

Dadds, M. R., Spence, S. H., Holland, D., Barrett, P. M., & Laurens, K. (1997). Early intervention and prevention of anxiety disorders: A controlled trial. *Journal of Consulting and Clinical Psychology, 65,* 627–635.

Davey, G. C. L. (1992). Classical conditioning and the acquisition of human fears and phobias: A review and synthesis of the literature. *Advances in Behavior Research and Therapy, 14,* 29–66.

DeKlyen, M. (1996). Disruptive behavior disorder and intergenerational attachment patterns. *Journal of Consulting and Clinical Psychology, 64,* 357–365.

Dumas, J. E., LaFreniere, P. J., & Serketich, W. J. (1995). "Balance of power": A transactional analysis of control in mother–child dyads involving socially competent, aggressive, and anxious children. *Journal of Abnormal Psychology, 104,* 104–113.

Edelbrock, C. (1985, June). *Teachers' perceptions of childhood anxiety and school adjustment.* Paper presented at a conference for "Anxiety Disorders in Children: Implications for School Adjustment," Cape Cod, MA.

Ekman, P. (1980). *The face of man.* New York: Garland.

Emery, R. E. (1982). Interparental conflict and the children of divorce and discord. *Psychological Bulletin, 92,* 310–330.

Famularo, R., Kinscherf, R., & Fenton, T. (1988). Propranolol treatment for childhood post-traumatic stress disorder, acute type: A pilot study. *American Journal of Diseases in Children, 142,* 1244–1247.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99,* 20–35.

Fonseca, A. C. (1993). Fears in children and adolescents: A study on the Portuguese population. *Revista Portuguesa de Pedagogia, 27,* 75–92.

Fox, J. E., & Houston, B. K. (1981). Efficacy of self instructional training for reducing children's anxiety in evaluation situations. *Behavior Research & Therapy, 19,* 509–515.

Fox, J. E., & Houston, B. K. (1983). Distinguishing between cognitive and somatic trait and state anxiety in children. *Journal of Personality and Social Psychology, 45,* 862–870.

Francis, G., Last, C. G., & Strauss, C. C. (1992). Avoidant disorder and social phobia in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 1086–1089.

Freeman, B. J., Roy, R. R., & Hemmick, S. (1976). Extinction of a phobia of physical examination in a 7-year-old mentally retarded boy. *Behavior Research & Therapy, 14,* 63–64.

Freud, S. (1909). Analysis of a phobia in a five-year-old boy. *Standard Edition, 10,* 3–149. London: Hogarth Press, 1955.

Giebenhain, J. E., & O'Dell, S. L. (1984). Evaluation of a parent training manual for reducing children's fears of the dark. *Journal of Applied Behavior Analysis, 17,* 121–125.

Gittelman, R. (1984). Anxiety disorders in children. In L. Grinspoon (Ed.), *Psychiatry update* (Vol. III, pp. 410–418). Washington, DC: American Psychiatric Association.

Gittelman-Klein, R. (1988). Childhood anxiety disorders. In C. J. Kestenbaum & D. T. Williams (Eds.), *Handbook of clinical assessment of children and adolescents* (Vol. 2). New York: NYU Press.

Gittelman-Klein, R., & Klein, D. F. (1971). Controlled imipramine treatment of school phobia. *Archives of General Psychiatry, 25,* 204–207.

Gittelman-Klein, R., & Klein, D. F. (1973). School phobia: Diagnostic considerations in the light of imipramine effects. *Journal of Nervous Mental Disorders, 156,* 199–215.

Glennon, B., & Weisz, J. R. (1978). An observational approach to assessment of anxiety in young children. *Journal of Consulting and Clinical Psychology, 46,* 1246–1257.

Goodyer, I. N. (1990). Family relationships, life events and childhood psychopathology. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 31* (1), 161–192.

Graae, F., Milner, J., Rizzotto, L., & Klein, R. G. (1994). Clonazepam in childhood anxiety disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 372–376.

Gray, J. (1988). The neuropsychological basis of anxiety. In C. Last & M. Hersen (Eds.), *Handbook of anxiety disorders.* New York: Pergamon.

Graziano, A. M., De Giovanni, I. S., & Garcia, K. A. (1979). Behavioral treatment of children's fears: A review. *Psychological Bulletin, 86,* 804–830.

Graziano, A. M., & Mooney, K. C. (1980). Family self-control instructions for children's night time fear reduction. *Journal of Consulting and Clinical Psychology, 48,* 206–213.

Green, S. M., Loeber, R., & Lahey, B. B. (1992). Child psychopathology and deviant family hierarchies. *Journal of Child and Family Studies, 1,* 341–350.

Haley, J. (1976). *Problem-solving therapy.* San Francisco: Jossey-Bass.

Harlow, H., & Harlow, M. (1962). Social deprivation in monkeys. *Scientific American, 207,* 136–146.

Heard, P. M., Dadds, M. R., & Conrad, P. (1992). Assessment and treatment of simple phobias in children: Effects on family and marital relationships. *Behavior Change, 9,* 73–82.

Hershberg, S. G., Carlson, G. A., Cantwell, D. P., & Strober, M. (1982). Anxiety and depressive disorders in psychiatrically disturbed children. *Journal of Clinical Psychiatry, 43,* 358–361.

Hetherington, E. M., & Martin, B. (1979). Family interaction. In H. C. Quay & J. S. Werry (Eds.), *Psychopathological disorders of childhood* (pp. 30–82). New York: Wiley.

Hirshfeld, D. R., Rosenbaum, J. F., Biederman, J., Bolduc, E. A., Faraone, S. V., Snidman, N., Reznick, J. S., & Kagan, J. (1992). Stable behavioral inhibition and its association with anxiety disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 103–111.

Jay, S. M., & Elliott, C. (1984). Behavioral observation scales for measuring children's distress: The effects of increased methodological rigor. *Journal of Consulting and Clinical Psychology, 52,* 1106–1107.

Jay, S. M., Ozolins, M., Elliott, C., & Caldwell, S. (1983). Assessment of children's distress during painful medical procedures. *Journal of Health Psychology, 2,* 133–147.

Jersild, A. T., & Holmes, F. B. (1935). *Children's fears.* New York: Teachers' College, Columbia University.

Kagan, J., Reznick, J. S., & Snidman, N. (1988). Biological bases of childhood shyness. *Science, 240,* 167–171.

Kashani, J. H., Dandoy, A. C., & Orvaschel, H. (1991). Current perspectives on anxiety disorders in children and adolescents: An overview. *Comprehensive Psychiatry, 32,* 481–495.

Kashani, J. H., & Orvaschel, H. (1988). Anxiety disorders in mid adolescence: A community sample. *American Journal of Psychiatry, 145,* 960–964.

Kashani, J. H., & Orvaschel, H. (1990). A community study of anxiety in children and adolescents. *American Journal of Psychiatry, 147,* 313–318.

Katz, E. R., Kellerman, J., & Siegel, S. E. (1980). Behavioral distress in children with cancer undergoing medical procedures: Developmental considerations. *Journal of Consulting and Clinical Psychology, 48,* 356–365.

Kazdin, A. E. (1985). Recent advances in child behavior therapy. In S. I. Pfeiffer (Ed.), *Clinical child psychology:*

Introduction to theory, research and practice. Orlando, FL: Grune & Stratton.

Keller, M. B., Lavori, P., Wunder, J., Beardslee, W. R., Schwarts, C. E., & Roth, J. (1992). Chronic course of anxiety disorders in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 595–599.

Kelley, C. K. (1976). Play desensitization of fear of darkness in preschool children. *Behavior Research & Therapy, 14,* 79–81.

Kendall, P. C. (1990). *Coping cat workbook*. Philadelphia, PA: Department of Psychology, Temple University.

Kendall, P. C. (1994). Treating anxiety disorders in youth: Results of a randomised clinical trial. *Journal of Consulting and Clinical Psychology, 62,* 100–110.

Kendall, P. C., Chansky, T. E., Kane, M. T., Kim, R. S., Kortlander, E., Ronan, K. R., Sessa, F. M., & Siqueland, L. (1992). *Anxiety disorders in youth*. Boston: Allyn & Bacon.

Kendall, P. C., Flannery-Schroeder, E., Panichelli-Mindel, S. M., Southam-Gerow, M., Henin, A., & Warman, M. (in press). Treatment of anxiety disorders in youth: A second randomized clinical trial. *Journal of Consulting and Clinical Psychology*.

Kendall, P. C., & Southam-Gerow, M. A. (1996). Long-term follow-up of a cognitive-behavioral therapy for anxiety-disordered youth. *Journal of Consulting and Clinical Psychology, 64,* 724–730.

King, N. J., Hamilton, D. J., & Ollendick, T. H. (1988). *Children's phobias: A behavioral perspective*. Chichester, UK: Wiley.

King, N. J., Hamilton, D. J., & Murphy, G. C. (1983). The prevention of children's maladaptive fears. *Child and Family Behvaiour Therapy, 5,* 43–57.

King, N. J., & Ollendick, T. H. (1989). School refusal: Graduated and rapid behavioral treatment strategies. *Australian and New Zealand Journal of Psychiatry, 23,* 213–223.

King, N. J., Ollier, K., Iacuone, R., Schuster, S., Bays, K., Gullone, E., & Ollendick, T. H. (1989). Fears of children and adolescents: A cross-sectional Australian study using the revised Fear Survey Schedule for Children. *Journal of Child Psychology and Psychiatry, 30,* 775–784.

Klein, D. F. (1964). Delineation of two-drug responsive anxiety syndromes. *Psychopharmacologia, 5,* 397–408.

Klein, R. G. (1991). Parent–child agreement in clinical assessment of anxiety and other psychopathology: A review. *Journal of Anxiety Disorders, 5,* 187–198.

Klein, R. G., Koplewicz, H. S., & Kanner, A. (1992). Imipramine treatment of children with separation anxiety disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 21–28.

Klein, R. G., & Last, C. G. (1989). *Anxiety disorders in children*. London: Sage.

Kohlmann, C. W., Schumacher, A., & Streit, A. (1988). Trait anxiety and parental child rearing behaviour: Support as moderator variable. *Anxiety Research, 1,* 53–64.

Kohut, H. (1971). *The analysis of the self*. New York: International Universities Press.

Kranzler, H. (1988). Use of buspirone in an adolescent with overanxious disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 789–790.

Krohne, H. W., & Hock, M. (1991). Relationships between restrictive mother–child interactions and anxiety of the child. *Anxiety Research, 4,* 109–124.

Krohne, H. W., & Hock, M. (1993). Coping dispositions, actual anxiety, and the incidental learning of success- and failure-related stimuli. *Personality and Individual Differences, 15,* 33–41.

Krohne H. W. (1992). Developmental conditions of anxiety and coping. A two process model of child rearing effects. In K. A. Hagtvet (Ed.), *Advances in test anxiety research* (Vol. 7, pp. 143–155). The Netherlands: Swets & Zeitlinger.

Kutcher, S. P., & MacKenzie, S. (1988). Successful clonazepam treatment of adolescents with panic disorder. *Journal of Clinical Psychopharmacology, 8,* 299–301.

Kutcher, S. P., Reiter, S., Gardner, D. M., & Klein, R. G. (1992). The pharmacotherapy of anxiety disorders in children and adolescents. *Pediatric Clinics of North America, 15,* 41–67.

Lang, P. J., & Lazovic, A. D. (1963). Experimental desensitization of a phobia. *Journal of Abnormal and Social Psychology, 66,* 519–525.

Lang, P. J. (1968). Fear reduction and fear behavior. Problems in treating a construct. In J. M. Schleen (Ed.), *Research in psychotherapy* (pp. 000–000). Washington, DC: American Psychological Association.

Last, C. G. (1988). Anxiety disorders in childhood and adolescence. In C. G. Last & M. Hersen (Eds.), *Handbook of anxiety disorders* (pp. 531–540). New York: Pergamon.

Last, C. G., Hersen, M., Kazdin, A. E., Finkelstein, R., & Strauss, C. C. (1987). Comparison of DSM-III separation anxiety and overanxious disorders: Demographic characteristics and patterns of comorbidity. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 527–531.

Le Baron, S., & Zeltzer, L. (1984). Assessment of acute pain and anxiety in children and adolescents by self-reports, observer reports, and a behavior checklist. *Journal of Consulting and Clinical Psychology, 52,* 729–738.

Le Roy, J. B., & Derdeyn, A. (1976). Drawings as a therapeutic medium: The treatment of separation anxiety in a 4 year-old boy. *Child Psychiatry and Human Development, 6,* 155–169.

Lewis, S. A. (1974). A comparison of behavior therapy techniques in the reduction of fearful avoidant behavior. *Behavior Therapy, 5,* 648–655.

McDaniel, K. D. (1986). Pharmacologic treatment of psychiatric and neurodevelopmental disorders in children and adolescents (I). *Clinical Pediatrics, 25,* 65–71.

Mahler, M. (1968). *On human symbiosis and the vicissitudes of individuation*. New York: International Universities Press.

Main, M. (1996). Overview of the field of attachment. *Journal of Consulting and Clinical Psychology, 64,* 237–243.

Marks, I. M. (1969). *Fears and phobias*. New York: Academic Press.

Matson, J. (1981) Assessment and treatment of clinical fears in mentally retarded children. *Journal of Applied Behavior Analysis, 14,* 287–294.

Mattison, R. E. (1992). Anxiety disorders. In S. R. Hooper, G. W. Hynd, & R. E. Mattison (Eds.), *Child psychopathology: Diagnostic criteria and clinical assessment* (pp. 179–202). Hillsdale, NJ:Erlbaum.

McFarlane, J. W., Allen, L., & Honzik, M. P. (1954). *A developmental study of the behavior problems of normal children between 21 months and 14 years*. Berkeley, CA: University of California Press.

McGee, R., Feehan, M., Williams, S., Partridge, F., Silva, P. A., & Kelly, J. (1990). DSM-III disorders in a large sample of adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 611–619.

Meichenbaum, D. H., & Genest, M. (1980). Cognitive behavior modification: an integration of cognitive and behavioral methods. In F. H. Kanfer & A. P. Goldstein, (Eds.), *Helping people change: A textbook of methods*. New York: Pergamon.

Meichenbaum, D. H., & Goodman, J. (1971). Training impulsive children to talk to themselves: A normal range of developing self-control. *Journal of Abnormal Psychology, 77,* 115–126.

Melamed, B. & Siegel, L. (1975). Reduction of anxiety in children facing hospitalization and surgery by use of filmed modeling. *Journal of Consulting and Clinical Psychology, 46,* 1357–1367.

Melamed, B., Yurcheson, R., Fleece, E. L., Hutcherson, S., & Hawes, R. (1978). Effects of film modeling on the reduction of anxiety-related behaviors in individuals varying in level of previous experience in the stress situation. *Journal of Consulting and Clinical Psychology, 46,* 1357–1367.

Messer, S. C., & Beidel, D. C. (1994). Psychosocial correlates of childhood anxiety disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 975–983.

Miller, L. C. (1983). Fears and anxiety in children. In C. E. Walker & M. C. Roberts (Eds.), *Handbook of clinical child psychology* (pp. 337–380). New York: Wiley.

Miller, L. C., Barrett, C. L., & Hampe, E. (1974). Phobias of childhood in a prescientific era. In A. Davids (Ed.), *Child personality and psychopathology: Current topics* (Vol. 1., pp. 89–134). New York: Wiley.

Miller, L. C., Barrett, C. L., Hampe, E., & Noble, H. (1972). Factor structure of childhood fears. *Journal of Consulting and Clinical Psychology, 39,* 264–268.

Milos, M. E., & Reiss, S. (1982). Effects of three play conditions on separation anxiety in young children. *Journal of Consulting and Clinical Psychology, 50,* 389–395.

Minuchin, S. (1974). *Families and family therapy*. Cambridge, MA: Harvard University Press.

Morris, R. H., & Kratochwill, T. R. (1983). *Treating children's fears and phobias.* New York: Pergamon.

Mowrer, O. H. (1960). *Learning theory and behavior.* New York: Wiley.

Murphy, C. M., & Bootzin, R. R. (1973). Active and passive participation in the contact desensitization of snake fear in children. *Behavior Therapy, 4,* 203–211.

Neisworth, J. T., Madle, R. A., & Goeke, D. E. (1975). "Errorless" elimination of separation anxiety: A case study. *Journal of Behavior Therapy and Experimental Psychiatry, 6,* 79–82.

Ollendick, T. H. (1983). Reliability and validity of the Revised Fear Survey Schedule for Children (FSSC-R). *Behavior Research & Therapy, 21,* 685–692.

Ollendick, T. H., & Francis, G. (1988). Behavioral assessment and treatment of childhood phobias. *Behavior Modification, 12,* 165–204.

Ollendick, T. H., King, N. J., & Frary, R. B. (1989). Fears in children and adolescents: Reliability and generalizability across gender, age and nationality. *Behavior Research & Therapy, 27,* 19–26.

Ollendick, T. H., Matson, J. L., & Helsel, W. J. (1985). Fears in children and adolescents: Normative data. *Behavior Research and Therapy, 23,* 465–467.

Oosterlaan, J., Prins, P. J. M., & Sergeant, J. (1992). *A Dutch translation of the Fear Survey Schedule for Children.* Paper presented at the SECAP meeting in Saratosa, FL.

Panella, D., & Henggeler, S. W. (1986). Peer interactions of conduct-disordered, anxious-withdrawn, and well-adjusted black adolescents. *Journal of Abnormal Child Psychology, 14,* 1–11.

Paul, G. L. (1996) *Insight versus desensitization in psychotherapy*. Stanford, CA: Stanford University Press.

Perrin, S., & Last, C. G. (1992). Do childhood anxiety measures measure anxiety? *Journal of Abnormal Child Psychology, 20,* 567–578.

Pfeffer, C. R., Lipkins, R., Plutchik, R., & Mizruchi, M. (1988). Normal children at risk for suicidal behavior. A two-year follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 34–41.

Pfefferbaum, B., Overall, J. E., Boren, H. A., Frankel, L. S., Sullivan, M. P., & Johnson, K. (1987). Alprazolam in the treatment of anticipatory and acute situational anxiety in children with cancer. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 532–535.

Puig-Antich, J., Lukens, E., Davies, M., Goetz, D., Brenn-Quattrock, J., & Todak, G. (1985). Psychosocial functioning in prepubertal major depressive disorders: 1. Interpersonal relationships during the depressive episode. *Archives of General Psychiatry, 42,* 511–517.

Rapee, R. M. (1991). Psychological factors involved in generalized anxiety. In R. M. Rapee & D. H. Barlow (Eds.), *Chronic anxiety, generalized disorder, and mixed anxiety-depression* (pp. 76–96). New York: Guilford Press.

Rapee, R. M., & Barlow, D. H. (1992). Generalized anxiety disorder, panic disorder and the phobias. In P. B. Sutker & H. E. Adams (Eds.), *Comprehensive handbook of psychopathology* (2nd ed.). New York: Plenum.

Rapee, R. M., Barrett, P. M., Dadds, M. R., & Evans, L. (1994). Reliability of the DSM-III-R childhood anxiety disorders using a structured interview. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 984–992.

Reynolds, C. R., & Paget, K. D. (1981). Factor analysis of the revised children's manifest anxiety scale for blacks, males and females with national innovative sample. *Journal of Consulting and Clinical Psychology, 49,* 352–359.

Riddle, M. A., Hardin, M. T., King, B., Scahill, L., & Woolston, J. L. (1990). Fluoxetine treatment of children and adolescents with Tourette's and obsessive compulsive disorder. Preliminary clinical experience. *Journal of American Academy of Child and Adolescent Psychiatry, 29,* 45–48.

Riddle, M. A., Scahill, L., King, R. A., Hardin, M. T., Anderson, G. M., Ort, S. I., Smith, J. C., Leckman, J. F., & Cohen, D. J. (1992). Double-blind, cross-over trial of fluoxetine and placebo in children and adolescents with obsessive compulsive disorder. *Journal of American Academy of Child and Adolescent Psychiatry, 31,* 1062–1069.

Rosenbaum, J. F., Biederman, J., Bolduc, E. A., Hirshfield, D. R., Faraone, S. V., & Kagan, J. (1992). Comorbidity of parental anxiety disorders as risk for childhood-onset anxiety in inhibited children. *American Journal of Psychiatry, 149*(4), 475–481.

Rosenstein, D. S., & Horowitz, H. A. (1996). Adolescent attachment and psychopathology. *Journal of Consulting and Clinical Psychology, 64,* 244–253.

Rubin, K. H., & Clark, M. L. (1983). Preschool teachers' ratings of behavioral problems: Observational, sociometric, and social-cognitive correlates. *Journal of Abnormal Child Psychology, 11,* 273–286.

Rutter, M., Macdonald, H., LeCouteur, A., Harrington, R., Bolton, P., & Bailey, A. (1990). Genetic factors in child psychiatric disorders-II. Empirical findings. *Journal of Child Psychology and Psychiatry, 31,* 39–83.

Sanderson, W. C., Di Nardo, P. A., Rapee, R. M., & Barlow, D. H. (1990). Syndrome comorbidity in patients diagnosed with a *DSM-III*-revised anxiety disorder. *Journal of Abnormal Psychology, 99,* 308–312.

Sheslow, D. W., Bondy, A. S., & Nelson, R. O. (1982). A comparison of graduated exposure, verbal coping skills, and their combination in the treatment of children's fear of the dark. *Child and Family Behavior Therapy, 4,* 33–45.

Silverman, W. K. (1991). Diagnostic reliability of anxiety disorders in children using structured interviews. *Journal of Anxiety Disorders, 5,* 105–124.

Silverman, W. K. (1992). Taxonomy of anxiety disorders in children. In G. D. Burrows, R. Noyes, & S. M. Roth

(Eds.), *Handbook of anxiety* (Vol. 5, pp. 281–308). Amsterdam: Elsevier.

Silverman, W. K., & Nelles, W. B. (1988). The Anxiety Disorders Interview Schedule for Children. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 772–778.

Simeon, J. G., & Ferguson, H. B. (1987). Alprazolam effects in children with anxiety disorders. *Canadian Journal of Psychiatry, 32,* 570–574.

Simeon, J. G., Ferguson, H. B., Knott, V., Roberts, N., Gauthier, B., Dubois, C., & Wiggins, D. (1992). Clinical, cognitive and neurophysiological effects of alprazolam in children and adolescents with overanxious and avoidant disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 29–33.

Solyom, L., Silberfeld, M., & Solyom, C. (1976). Maternal overprotection in the aetiology of agoraphobia. *Canadian Psychiatric Association Journal, 21,* 109–113.

Spence, S. H. (1994). Preventative strategies. In T. H. Ollendick, N. J. King, & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in children and adolescents* (pp. 453–474). New York: Plenum.

Spence, S. H. (1996). *The structure of anxiety symptoms amongst children: A confirmatory factor analytic study*. Manuscript submitted for publication.

Stark, K. D., Humphrey, L. L., Crook, K., & Lewis, K. (1990). Perceived family environments of depressed and anxious children. *Journal of Abnormal Child Psychology, 22,* 33–51.

Strauss, C. C., Frame, C. L., & Forehand, R. L. (1987). Psychosocial impairment associated with anxiety in children. *Journal of Clinical Child Psychology, 16,* 235–239.

Strauss, C. C., Lahey, B. B., Frick, P., Frame, C. L., & Hynd, G. W. (1988). Peer social status of children with anxiety disorders. *Journal of Consulting and Clinical Psychology, 56,* 137–141.

Strauss, C. C., Lease, C. A., Last, C. G., & Francis, G. (1988). Overanxious disorder. An examination of developmental differences. *Journal of Abnormal Child Psychology, 16,* 433–443.

Stavrakaki, C., Vargo, B., Roberts, N., & Boodoosing, L. (1987). Concordance among sources of information for ratings of anxiety and depression in children. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 733–737.

Suomi, S. J. (1986). Anxiety-like disorders in young nonhuman primates. In R. Gittelman (Ed.), *Anxiety disorders of childhood* (pp. 1–23). New York: Guilford Press.

Suomi, S. J., Seaman, S. F., Lewis, J. K., DeLizio, R. D., & McKinney, W. T. (1978). Effects of imipramine treatment of separation induced social disorders in rhesus monkeys. *Archives of General Psychiatry, 35,* 321–327.

Target, M., & Fonagy, P. (1994). Efficacy of psychoanalysis for children with emotional disorders. *Journal of American Academy of Child and Adolescent Psychiatry, 33,* 361–371.

Thyer, B., Parrish, R. T., & Curtis, G. C. (1985). Ages of onset of *DSM-III* anxiety disorders. *Comprehensive Psychiatry, 26,* 113–122.

Thyer, B. A., & Sowers-Hoag, K. M. (1988). Behavior therapy for separation anxiety disorder. *Behavior Modification, 12,* 205–233.

Van Hasselt, V. B., Hersen, M., Bellack, A. S., Rosenblum, N. D., & Lamparski, D. (1979). Tripartite assessment of the effects of systematic desensitization in a multiphobic child: An experimental analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 10,* 51–55.

van Ijzendoom, M. H., & Bakermans-Kranenburg, M. J. (1996). Attachment representations in mothers, fathers, adolescents and children: A meta-analytic search for normative data. *Journal of Consulting and Clinical Psychology, 64,* 8–21.

Vasey, M. W. (1993). Development and cognition in childhood anxiety. The example of worry. In T. H. Ollendick & R. Prinz (Eds.), *Advances in clinical child psychology* (Vol. 15, pp. 1–39). New York: Plenum.

Velez, C. N., Johnson, J., & Cohen, P. (1989). A longitudinal analysis of selected risk factors for childhood psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 861–864.

Werry, J. S. (1991). Overanxious disorder: A review of taxonomic properties. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 533–544.

Werry, J. S. (1994). Diagnostic and classification issues. In T. H. Ollendick, N. J. King, & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in children and adolescents* (pp. 21–42). New York: Plenum.

Wine, J. D. (1979). Test anxiety and evaluation threat: Children's behavior in the classroom. *Journal of Abnormal Child Psychology, 7,* 45–59.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press.

World Health Organization (1992). *The ICD-10 classification of mental and behavioral disorders: Clinical descriptions and diagnostic guidelines*. Geneva: Author.

Yule, W., Udwin, O., & Murdoch, K. (1990). The Jupiter sinking: Effects on children's fears, depression and anxiety. *Journal of Child Psychology and Psychiatry, 31,* 1052–1061.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.14
Specific Phobias

GOLDA S. GINSBURG
University of Baltimore, MD, USA
and
WENDY K. SILVERMAN
Florida International University, Miami, FL, USA

5.14.1 DEFINITION AND PHENOMONOLOGY 344
  5.14.1.1 Developmental Variations 344
  5.14.1.2 Gender Variations 344
  5.14.1.3 Race and Cultural Variations 345
5.14.2 ONSET AND STABILITY 346
5.14.3 DIAGNOSIS 347
  5.14.3.1 Differential Diagnosis 348
  5.14.3.2 Comorbidity 348
5.14.4 ETIOLOGY 348
  5.14.4.1 Biological 348
  5.14.4.2 Psychosocial Factors 350
    5.14.4.2.1 Psychoanalytic 350
    5.14.4.2.2 Behavioral 350
    5.14.4.2.3 Cognitive 351
    5.14.4.2.4 Family 353
5.14.5 EPIDEMIOLOGY 354
  5.14.5.1 Prevalence of Phobia Subtypes 354
  5.14.5.2 Sociodemographic Risk Factors 354
    5.14.5.2.1 Socioeconomic status 354
    5.14.5.2.2 Racial/ethnic variations 355
    5.14.5.2.3 Gender 355
5.14.6 ASSESSMENT 355
  5.14.6.1 Cognitive/Subjective 355
  5.14.6.2 Behavioral 357
  5.14.6.3 Physiological 357
5.14.7 TREATMENT 358
  5.14.7.1 Transfer of Control Model 359
5.14.8 CASE STUDY 360
  5.14.8.1 Presenting Problem 360
  5.14.8.2 Assessment and Diagnosis 360
  5.14.8.3 Treatment 360
  5.14.8.4 Post-treatment and Follow-up 361
5.14.9 SUMMARY 361
5.14.10 REFERENCES 361

5.14.1 DEFINITION AND PHENOMONOLOGY

A specific phobia (SP) is a persistent and excessive fear in the presence of, or in anticipation of, a circumscribed object or event (*DSM*; American Psychiatric Association, 1994). Unlike a "normal" nonclinical fear, an SP usually leads to extreme avoidance of the fear-provoking object or event, resulting in severe distress and interference in the daily functioning of children and their families. For example, children with an SP of dogs may not leave their home to avoid a possible encounter with a dog; children with an SP of sleeping alone in the dark may insist that their parents sleep along with them; children with an SP of medical procedures (such as injections) may refuse to go to doctors' appointments or to hospitals. Also, in contrast to children's normal fears, SPs are usually out of proportion to reality and do not dissipate by reassurance, reason, or information.

In referring to phenomonology, it has become common to focus on the three-response system: behavioral, cognitive, and physiological (e.g., Lang, 1977). In terms of the behavioral system, avoidance is most common, and, as noted, typically children go to great lengths to avoid the feared object or situation. Additional behavioral manifestations include tantruming, crying, or "freezing up" when faced with the feared object or event. In terms of the cognitive system, often children report catastrophic beliefs that involve feared outcomes, particularly of physical or personal injury (e.g., "the dog will bite me," "I'll get hit by lightening"). Finally, in terms of the physiological system, manifestations include somatic complaints (e.g., stomach aches, headaches), accelerated heart rate, sweating, shakiness, and muscle tension (Last, 1991). A vasovagal fainting response occurs in a majority of children with phobias of blood-injection-injury (American Psychiatric Association, 1994; Page, 1994).

5.14.1.1 Developmental Variations

Most of what is known about developmental variations is based on research of nonpathological fears using nonclinic samples of children. Findings from this research generally show that the number and intensity of children's fears decrease with age (e.g., Eme & Schmidt, 1978; Ollendick, King, & Frary, 1989). For example, in Ollendick et al. (1989), self-ratings of fear, among a school sample of 1185 children (aged 7–16 years) residing in the US and Australia, indicated that children in the youngest age group (aged 7–10 years) reported a greater number of fears (mean number = 17) compared to children in the middle (aged 11–13 years; mean number = 13) or oldest age group (aged 14–16 years; mean number = 12). Children in the youngest group also reported a higher intensity of fears as reflected by their total fear score (mean = 138.83) compared with children in the middle (mean = 133.44) and oldest age group (mean = 129.46) (Ollendick et al., 1989).

In addition to the number and intensity of fears, research findings suggest developmental variations in the content of children's fears in nonclinical samples (e.g., Bauer, 1976; Bell-Dolan, Last, & Strauss, 1990; McGee, Feehan, Williams, & Anderson, 1992; Ollendick et al., 1989). Earlier reviews of this literature (e.g., Morris & Kratochwill, 1983; Ollendick & Francis, 1988; Silverman & Nelles, 1990) indicated that loss of support, loud noises, and unfamiliar people predominate in infancy; imaginary creatures, small animals, and the dark predominate in the middle childhood years; and achievement/performance, social-evaluation, and bodily injury predominate in early and late adolescence.

Longitudinal investigations corroborate these patterns. For instance, McGee et al. (1992), in a study of community children (N = 925) assessed at age 11 and again at age 15 (using structured diagnostic interview and parent ratings), found that the content of children's fears changed over time. At age 11, children's fears were likely to be of the dark, heights, or animals; at age 15, the fears were likely to be of social and agoraphobic situations.

Although research using clinic samples of children with anxiety disorders is sparse (e.g., Last, Perrin, Hersen, & Kazdin, 1992; Strauss & Last, 1993), extant findings suggest similar developmental patterns in the content of childrens' fears. For instance, children with a primary diagnosis of social phobia (SOP; social-evaluation/performance fears) tend to be older than children with a primary diagnosis of separation anxiety disorder (SAD) or SP (animals, dark) (see Section 5.14.2).

5.14.1.2 Gender Variations

Research on gender variations using clinic samples has been conducted primarily with respect to prevalence rates (see Section 5.14.5). Research on gender variations in children's fears using nonclinic populations, and using self-, parent-, teacher-, and peer-ratings, shows that girls report more fears, and obtain higher fear ratings from others, than do boys (both frequency and intensity) (e.g., Croake, 1969; Ollendick, 1983; Ollendick & King, 1994; Ollendick, Yang, Dong, Xia, & Lin, 1995; Ollendick, Yule, & Ollier, 1991; Silverman &

Nelles, 1988b). One common explanation is that gender role expectations or attributions may be operating. That is, it is more acceptable for girls than for boys to act "feminine," which includes admitting to fears or behaving in a fearful manner (Maccoby, 1980). Consistent with this notion is the finding that children of both genders tend to ascribe more fear to girls than to boys (Ollendick et al., 1995; Silverman & Nelles, 1988b).

Although a positive relation between "femininity" and a negative relation between "masculinity" (as measured by self-report measures) and fear/anxiety have been found in adult populations (Carey, Dusek, & Spector, 1988; Dillon, Wolf, & Katz, 1988), these hypotheses have not been investigated directly in children. However, preliminary findings from a study on gender and children's fears (Ginsburg & Silverman, 1997) suggest that self-ratings of fear may indeed be associated with gender role orientation. Based on a clinic sample of children ($N = 92$; mean age = 10.3 years) diagnosed with a primary *DSM-III-R* anxiety disorder, a significant negative correlation ($r = -0.21$; $p < 0.05$) was found between children's total scores on the FSSC-R and self-ratings of masculinity on the Children's Sex Role Inventory (Boldizar, 1991).

5.14.1.3 Race and Cultural Variations

Studies of children's fears using nonclinic samples suggest that fears (as measured by child self-rated fear surveys) may be more frequent and intense in American minority groups, such as African-Americans, than in Caucasians (Lapouse & Monk, 1959; Nalven, 1970; Neal & Turner, 1991). Reviews of these studies also suggest that level of fearfulness as well as the presence of specific fears may vary by culture (see Fonesca, Yule & Erol, 1994). For example, a study by Ollendick, Yang, King, Dong, and Akande (1996) examined fears (using the FSSC-R) of 1200 children, (aged 7–17 years), from the US, Australia, China, and Nigeria (300 children from each country), and found several differences in number, level, and content, as well as differences in developmental and gender patterns across countries. One finding, for instance, was that Nigerian children reported more and higher levels of fear compared to their American, Australian, or Chinese counterparts, who did not differ from each other. Moreover, girls and boys in Nigeria reported similar levels of fear; girls in each of the other countries reported higher levels than boys. Finally, the Nigerian children were more likely to report intense fears associated with physical danger (e.g., snakes, guns, getting shock from electricity); the American and Australian children reported more intense fears of personal safety (e.g., burglars, getting lost). In explaining these cultural variations in fears, Ollendick et al. (1996) point to cultural differences in socialization practices. That is, countries such as Nigeria tend to emphasize socialization practices such as inhibition, compliance, and obedience (and thus children with more fear), and countries such as the US and Australia do not have this emphasis.

Despite these cultural variations, several studies have found similarities in the types of objects and situations feared by children across cultures. For example, in comparing the top 10 fears endorsed on the FSSC-R by children from five different countries (Portugal, UK, Turkey, Australia, and US), seven of the fears were the same across countries (Fonesca et al., 1994). These included: fears of being hit by a car or truck; not being able to breathe; bomb attacks; fire/being burned; burglars; falling from a height; and serious illness. Similarly, in the Ollendick et al. (1996) study cited above, the most common intense fears (i.e., the 10 fears rated "a lot" with the greatest frequency) across countries showed considerable overlap, and reflected fears of danger and/or death (e.g., not being able to breath, being hit by a car or truck, bombing attacks, failing from high places, death/dead people) and failure and/or criticism (e.g., failing a test, having parents argue, and getting poor grades).

Taken all together, the results from cross-cultural studies using nonclinic samples of children suggest that although certain types of fears appear to be culturally meaningful or specific, there also appear to be certain types of fears that are experienced uniformly by children across various cultures. Further research on the specific nature of the socialization practices (e.g., parenting practices) that influence the development of childhood fears across cultures awaits further systematic study.

In terms of racial/cultural variations using clinic samples of youth with anxiety disorders, findings, although sparse, suggest that youth with SP may be more alike than different. For instance, Last and Perrin (1993) compared the sociodemographic background, clinical characteristics, and life-time prevalence rates of *DSM-III-R* anxiety disorders (not only SP) of African-American ($N = 30$) and White children ($N = 139$) aged 5–17 years who presented to a childhood anxiety specialty clinic. Relevant to SP, Last and Perrin found that although the African-American children scored higher than the White children on the FSSC-R, the difference did not reach statistical significance.

Ginsburg and Silverman (1996) compared various characteristics of Hispanic ($N = 90$) and Caucasian children ($N = 143$) who presented to a childhood anxiety specialty clinic. Similar to the findings of Last and Perrin (1993), the children were more similar than different on the variables examined. One significant difference relevant to SP, however, was that the Hispanic parents rated their children as having more fears than did the parents of Caucasian children (using a parent version of the FSSC-R).

In sum, as far as we know, only two studies have examined systematically race and cultural variations of fears in clinic samples that included children with SP (Ginsburg & Silverman, 1996; Last & Perrin, 1993). Although both studies found more similarities than differences between the samples studied, further research is needed before firm conclusions are drawn.

5.14.2 ONSET AND STABILITY

Research has shown that the onset of many SPs occurs during childhood or adolescence (e.g., Ost, 1985, 1987; Ost & Hugdahl, 1983; D. V. Sheehan, K. E. Sheehan, & Minichiello, 1981; Thyer, Parrish, Curtis, Nesse, & Cameron, 1985), and that some SPs are stable over time. For example, one of the most extensive studies on the onset of SP, based on the retrospective reports of adult patients with SP ($N = 370$), reported that the average age of onset for animal phobia was seven, age nine for blood phobia, and age 12 for dental phobia (Ost, 1987).

Other studies, using retrospective reports of adult patients with SP have found similar patterns with respect to age of onset. For instance, Marks and Gelder (1966) found that the onset of phobias of darkness, insects, and blood-injury began before age seven; the onset of animal phobias was somewhat earlier, age 4.4 years. Liddell and Lyons (1978) found that the age of onset for blood phobia was 8.8 years, 10.8 years for dental phobia, and 11.9 years for thunderstorm phobias.

Another way to glean information about the onset and stability of SP in children is to examine the ages in which children present for treatment for SP and to assess the disorder's duration. Last et al. (1992) examined sociodemographic and clinical characteristics of youth ($N = 188$; aged 6–18 years) referred to a child anxiety specialty clinic and reported on developmental variations in primary *DSM-III-R* anxiety diagnoses. The average age at intake for children with a primary diagnosis of SOP ($N = 61$) was found to be 14.4 years, 10.3 years for SAD ($N = 84$), and 12.1 years for SP ($N = 80$). With respect to age of onset, the ages were 11.3 years for SOP, 7.5 years for SAD, and 8.4 years for SP.

Strauss and Last (1993), in a comparison of clinical characteristics of children who presented to a childhood anxiety clinic with primary diagnoses of either SP or SOP (aged 4–17 years), found that of the children with SP (N = 38; mean age = 11.1 years) the mean age of onset was 7.8 years and the average duration of the children's SP was 3.3 years.

In examining the average age at intake for children who presented to our research clinic, the Child Anxiety and Phobia Program at Florida International University, Miami, it was found also that the average age of the children ($N = 340$; aged 5–18 years) with primary *DSM-III-R* anxiety diagnoses varied. For instance, the average age of children with SAD ($n = 51$) was 8.7 years, 11.7 years for SOP ($n = 42$), and 9.9 years for SP ($n = 143$).

Finally, results obtained from longitudinal studies are yet another way to glean information about onset and stability of SP in children. Perhaps the most widely cited study is that of Agras, Chapin, and Oliveau (1972). They followed a community sample of 30 untreated phobic individuals (10 children under the age of 20 years and 20 adults) over a five-year period. After five years, 100% of the children were viewed as "improved" compared to 43% of the adults. Although the conclusion drawn by Agras et al. was that many phobic conditions improve without treatment, particularly in children, a reinterpretation of these data (Ollendick, 1979) revealed that the improved children were not symptom-free, and most continued to exhibit symptoms of sufficient intensity to be rated between "no disability" and "maximum disability" at the follow-up assessment. Ollendick's reinterpretation thus suggested that some types of phobias persist over time for a proportion of youngsters.

Another study conducted by Milne et al. (1995) examined the frequency of clinical phobias (SP, SOP, and agoraphobia (AG) combined) and subclinical fears in a community sample of youth over a one-year period as part of a large, two-stage epidemiological study on adolescent depression. In the first stage of the study (screening stage), a depressive symptom questionnaire was administered to a large community sample ($N = 3283$; 80% were White and the majority were 7th-graders). In the second stage (diagnostic stage), a selected sample of youth and their mothers ($N = 487$) were assessed with a structured diagnostic interview. These stages were repeated annually for up to three years, the number of children assessed at all three time periods was 112. Relevant here is the finding that of the 112 youth, 11% (with any type of phobia; $n = 12$) continued to meet diagnosis for that same type of phobia one year later.

5.14.3 DIAGNOSIS

A summary of the criteria for an SP diagnosis using the *DSM-IV* (APA, 1994) and *ICD-10* (World Health Organization, 1992) systems are presented in Table 1. In both systems, a diagnosis of SP is assigned when children display a persistent, recurrent, and/or excessive fear toward an object or event. Exposure to the phobic stimulus generally provokes an immediate anxiety reaction, often manifested in children as crying, having tantrums, freezing, or clinging. Generally, the phobic situation is avoided or experienced with severe distress. A key factor in assigning a diagnosis of SP is that the symptoms must impair functioning significantly.

Differences in the *ICD-10* and *DSM-IV* pertain to their views about the classification of the phobia as an adult or child disorder (based on whether the fear is developmentally age-appropriate), about the persistence of symptoms, and the classification of phobia subtypes. Specifically, according to *ICD-10*, phobias may be classified as either of one of two disorders, "phobic anxiety disorder of childhood" (under the broader category, "Emotional Disorders with Onset Specific to Childhood") or "phobic anxiety disorder" (under the broader category, "Neurotic, Stress-related, and Somatoform Disorders"), generally used for adults. The criteria for the childhood-onset disorder requires that the fears be age-appropriate but excessive and impairing. In cases in which fears are not age-specific, the adult disorder is assigned to that child as a diagnosis. In contrast, *DSM-IV* has only one classification for phobias, and fears need not be age-appropriate to be assigned the diagnosis. For *ICD-10*, duration is

Table 1 Diagnostic criteria for specific phobias in children.

DSM-IV

A. A marked, persistent fear that is excessive or unreasonable, prompted by the presence or anticipation of an object or situation (e.g., flying, heights, animals)
B. Exposure to the feared object or situation usually provokes an immediate anxiety response. In children this may be expressed by crying, tantrums, freezing, or clinging
C. A recognition that the fear is excessive or unreasonable (this feature may be absent in children)
D. The feared object or situation(s) is avoided or endured with intense anxiety or distress
E. The avoidance or distress toward the feared object or situation significantly interferes with the person's normal routine, occupational (or academic) functioning, or social activities or relationships, or causes marked distress
F. Fear is present for at least six months for individuals under age 18 years
G. The individual does not meet criteria for another mental disorder that better accounts for their anxiety, panic attacks, or phobic avoidance associated with the specific object or situation (e.g., obsessive compulsive disorder, separation anxiety disorder)

ICD–10 (includes both child and adult descriptions)

A. Phobic anxiety disorders. Disorders in which anxiety is evoked only, or predominantly, in well-defined situations that are not currently dangerous. Consequently, these situations are avoided or endured with dread. The patient's concern may be focused on individual symptoms like palpitations or feeling faint and is often associated with fears of dying, losing control, or going mad. Contemplating entry to the phobic situation usually generates anticipatory anxiety
B. Specific (isolated) phobias. Phobias that are restricted to highly specific situations such as proximity to particular animals, heights, thunder, darkness, flying, closed spaces, etc. Types of specific phobias are classified as: acrophobia, animal phobia, claustrophobia, simple phobia
C. Phobic anxiety disorder specific to childhood. Fears displayed during childhood, that arise in a majority of children, but that are abnormal in degree. Fears displayed in childhood that are not a normal part of psychosocial development (e.g., agoraphobia) are coded under the adult categories
D. Excludes generalized anxiety disorder and when associated with conduct disorders

not part of the diagnostic criteria; for *DSM-IV*, the children's symptoms must be continuous for six months to receive a diagnosis. Finally, *ICD-10* childhood onset does not specify subtypes of SP; *ICD-10* adult-onset phobia subtypes include: acrophobia, animal phobia, claustrophobia, simple phobia. The subtypes of *DSM-IV* are: animal (e.g., cats, dogs, insects); natural environment (e.g., storms, heights, water); blood-injection-injury (e.g., seeing blood, needles); situational (e.g., public transportation, elevators, flying); and other (e.g., loud noises, costume characters).

5.14.3.1 Differential Diagnosis

A diagnosis of SP should only be given once other anxiety conditions have been ruled out. For example, according to the *DSM*, a diagnosis of SP is assigned only when the fear is not restricted to situations that involve separation (as in SAD), social-evaluation and embarrassment (as in SOP), dirt/contamination (as in obsessive compulsive disorder, OCD), or fears of having a panic attack and/or being unable to escape (as in panic with or without AG). It is also essential that the fear not be a part of a larger reaction to a traumatic event, as in post-traumatic stress disorder, or a more pervasive pattern of anxiety, as in generalized anxiety disorder (GAD).

Finally, although children who refuse to attend school are often labeled as having a "school phobia," research has revealed that these children refuse school for a myriad of reasons (Bernstein & Garfinkel, 1986; Burke & Silverman, 1987; Kearney & Silverman, 1997) which may or may not be related to having SP. For instance, a child may refuse to attend school because of fears of speaking in class (SOP), being away from a parent (SAD), or hearing the bell ring (SP, situational subtype). Thus, to derive a differential diagnosis for children who present with school refusal, it is important to carefully assess the specific reasons for the school refusal.

5.14.3.2 Comorbidity

Clinic and community studies indicate that up to one-half of all children present with more than one comorbid disorder (see Caron & Rutter, 1991 for a review). Several studies examining the comorbid patterns among samples of clinic-referred children with anxiety disorders have been conducted (e.g., Last, Strauss, & Francis, 1987; Last et al., 1992; Strauss & Last, 1993). Last et al. (1987) examined the comorbid *DSM-III* diagnoses of a sample of 73 children (aged 5–18 years) with primary diagnoses of SAD ($n = 24$), overanxious disorder (OAD; $n = 11$), SP ($n = 11$), and major depression ($n = 11$). Findings with respect to SP indicated that 64% of the children in the SP group had a comorbid diagnosis. The most common comorbid anxiety disorder was SAD (39%).

In a subsequent study, Last et al. (1992) examined comorbid lifetime history diagnoses in children with anxiety disorders ($N = 188$; aged 5–18 years). Relevant to SP were the findings that 60 of the 80 children (75%) with a primary diagnosis of SP had a lifetime history of any anxiety diagnosis; 32.5% had a history of any depressive disorder; and 22.5% had a history of any disruptive disorder. The most common additional anxiety disorder was again SAD ($n = 31$; 38.8%).

Hammond-Laurence and Silverman (1997) examined current rates of comorbid diagnoses in a large sample of children referred to an anxiety specialty clinic ($n = 356$; aged 6–17 years. With respect to SP, the findings indicated that this was the most common diagnosis of the sample and, typically, children with a primary diagnosis of SP met criteria for a secondary diagnosis of another SP (29%) or SAD (18%). Also of interest is the finding that children with only one anxiety disorder tended to be rated lower (on average) in clinical severity relative to children with multiple disorders, indicating that there is "clinical significance" to having multiple comorbid disorders.

Although the high comorbid rates reported in the studies just cited could be due in part to the Berkson effects, that is, both the Last et al. (1987, 1992) and the Hammond-Laurence and Silverman (1997) studies were conducted at anxiety specialty clinics that attract the most severely anxious youth, these findings are in accord with the growing body of literature concerned with comorbidity (Caron & Rutter, 1991). However, a great deal more work is needed to understand the meaning and implications of comorbidity, not only in terms of clinical presentation, but also in terms of prognosis and treatment.

5.14.4 ETIOLOGY

5.14.4.1 Biological

Evidence for the role of biological factors is based largely on findings from family aggregate and twin studies (e.g., Burglass, Clarke, Henderson, Kreitman, & Presley, 1977; Davey, Forster, & Mayhew, 1993; Fyer et al., 1990; Fyer, Mannuzza, Chapman, Martin, & Klein, 1995; Kendler, Neale, Kessler, Heath, & Eaves, 1992; Last, Hersen, Kazdin, Orvaschel, & Perrin, 1991; Moran & Andrews, 1985; Noyes et al., 1986; Reich & Yates, 1988; Solyom, Beck, Solyom, & Hugel, 1974; Torgersen, 1979). Not all of these studies focused exclusively on SP, but examined various types of phobic and anxiety disorders. Overall, however, these studies provide support for a positive familial contribution to these disorders.

A representative study that focused on SP was conducted by Fyer et al. (1990). In this study, the probands of adults diagnosed with SP (*N* = 49) were assessed, including all their first-degree relatives, using the Schedule for Affective Disorders and Schizophrenia (SADS; Endicott & Spitzer, 1978). First-degree relatives of a normal proband group (*N* = 119) served as the control group. The results indicated that relatives of individuals with SP evidenced a higher rate of SP themselves (31%) compared with relatives of those in the comparison group (11%). In addition, 15% of the children whose parents had SP, compared to 8% of children whose parents did not have SP, evidenced SP.

Another study by the same research group (Fyer et al., 1995) compared the familial aggregation of three different types of phobic disorders: SP (*N* = 15), SOP (*N* = 39), and panic with AG (*N* = 49). All participants were of European descent, with one living first-degree relative, and were recruited from ongoing studies that were being conducted at two different anxiety disorders research clinics. None of the participants who comprised each of the three phobic disorder groups had comorbid diagnoses (present or lifetime). The control participants, who were recruited from a randomly selected patient acquaintance list, were 77 "not ill" individuals, that is, no past or current psychiatric diagnoses. *DSM-III-R* diagnoses were assessed using a combination of a direct interview (i.e., SADS-Lifetime Anxiety Version) and family informants' reports.

The results indicated that relatives of individuals in the SP group had significantly higher rates of SP (31%) than the relatives of individuals in the control group (9%). Interestingly, the rates of SP in the relatives of individuals with panic/AG (10%) and SOP (13%) did not differ significantly from the rates found in the relatives of individuals in the control group (9%). In interpreting these results, Fyer et al. (1995) suggested that there may be an "existence of separable, etiologically discrete syndromes within the group of cases identified by each of the three *DSM-III-R* criteria sets" (p. 572). However, the authors caution that this notion may apply to "pure" diagnostic cases (i.e., those with no present or life-time comorbid diagnoses). Also not clear is whether it applies to each subtype of SP (e.g., if relatives of blood phobics were more at risk for blood phobias than for dog phobias) as this issue was not examined due to limited sample sizes. In addition, although the findings provide evidence for some type of familial risk for SP, what is not clear is whether this familial risk is due to shared genes, common family environmental factors, or an interaction of the two (Foley & Hay, 1992; Silverman, Cerny, & Nelles, 1988).

Additional evidence of a biological contribution comes from twin studies (e.g., Abe, Oda, Ikenaga, & Yamada, 1993; Kendler et al., 1992; Slater & Shields, 1969; Stevenson, Batten, & Cherner, 1992; Torgersen, 1979; Young, Fenton, & Lader, 1971). For example, in one of the largest studies examining genetic and common familial environmental influences in the etiology of phobic disorders (AG, SOP, and SP), Kendler et al. administered diagnostic and clinical interviews to 2163 adult female twins recruited from a population-based twin registry. Findings relevant here indicated that concordance rates of animal phobia for monozygotic and dizygotic twins were 25.9% and 11.0%, respectively. Concordance rates for situational phobia for monozygotic and dizygotic twins were 22.2% and 23.7%, respectively. According to Kendler et al., such findings provide support for the contribution of genetic factors in the etiology of SP, with a higher contribution of familial environmental factors operating for situational phobias.

Another avenue for investigating biological contributions to SP has been in the area of behavioral inhibition. Behavioral inhibition (BI) refers to the temperament of approximately 10–20% of Caucasian infants who are predisposed to being irritable, shy, and fearful as toddlers, and cautious, quiet, and introverted as school-aged children (e.g., Biederman et al., 1990, 1993a, 1993b; Hirshfeld et al., 1992; Kagan, Reznick, & Gibbons, 1989; Kagan, Reznick, & Snidman, 1987). Although studies generally support a positive association between anxiety disorders and BI in children, it is premature to conclude that BI is linked uniquely with SP, or with any other specific anxiety disorder.

With this caveat in mind, emerging evidence from this line of research suggests that BI in infancy may be a risk factor for phobic disorders (e.g., Biederman et al., 1990, 1993a, 1993b). For example, a study conducted by Biederman et al. (1990) used a structured interview (DICA-parent version) to examine the prevalence of *DSM-III-R* internalizing disorders (i.e., major depression, SAD, OAD, avoidant disorder [AVD], phobic disorders, and OCD) as well as externalizing disorders (i.e., oppositional defiant, conduct, attention deficit disorder, encopresis, enuresis) in three different cohorts of children. The first cohort, examined cross-sectionally, was the at-risk group and comprised offspring of parents with psychiatric diagnoses. Children in this cohort were classified as either inhibited ($N = 18$; mean age = 5.9 years) or uninhibited ($N = 12$; mean age = 4.9 years). The second cohort (Kagan's original cohort; Kagan, Reznick, & Snidman, 1987, 1988) was an epidemiologically derived sample of children who, at age 21 months, had been classified as either inhibited ($N = 22$) or uninhibited ($N = 19$). The mean age of the children at the time of the reassessment was 8 years. The third cohort of children had no medical or psychiatric disorders ($N = 20$; mean age = 7.8 years) and was obtained from primary pediatric care referrals.

Most relevant here are the findings that were found with the children in Kagan's original cohort. Specifically, the findings indicated that the inhibited compared with the uninhibited children had significantly higher rates of phobic disorders (31.8% vs. 5.3%, respectively). Inhibited compared with uninhibited children also had higher rates of OAD, SAD, and AVD, but this finding was not statistically significant. In exploratory analyses of the specific fears of the children who met criteria for a phobic disorder ($n = 9$), the results indicated that the average number of fears per child was 3.4, and that the most common fears were: standing up and speaking in front of class ($n = 5$); animals or bugs ($n = 5$); strangers ($n = 4$); the dark ($n = 4$); being called on in class ($n = 3$); crowds ($n = 3$); elevators ($n = 2$); and physicians ($n = 2$).

In a review of the research literature on BI and anxiety disorders, Turner, Beidel, and Wolff (1996) concluded that children with stable BI have higher rates of anxiety disorders (any type) and phobias compared with uninhibited children. With respect to phobias, BI appears to be associated most closely with those phobias that are social or social-evaluative in nature. Turner et al. also pointed out that although the research suggests that BI is neither necessary nor sufficient for developing an anxiety/phobic disorder, its presence may serve to increase children's vulnerability to anxiety disorders, including SP.

5.14.4.2 Psychosocial Factors

The emphasis that may be placed on specific psychosocial factors is likely to be influenced in part by the theoretical model or orientation to which the researcher or practitioner ascribes (e.g., psychoanalytic, behavioral, cognitive, familial). These models are summarized briefly below.

5.14.4.2.1 Psychoanalytic

According to the psychoanalytic perspective (Freud, 1909/1955), phobias develop in response to unconscious forbidden sexual or aggressive impulses. The anxiety experienced by these impulses are "defended" against via displacement of these feelings on to less threatening objects or situations. Hence, in Freud's famous case of Little Hans, Hans' phobia of horses was attributed to his attempt to defend against or avoid anxiety aroused by his sexual desire for his mother and his wish to do away with his father, the competitor. The child defended against the anxiety of these impulses by displacing his feelings on to an external object that can be avoided (i.e., horses). It has been difficult to verify empirically the psychoanalytic conceptualization of SP because of the unconscious nature of the main, hypothesized processes, namely, displacement. Moreover, subsequent analysis of this case by Wolpe and Rachman (1960) revealed no direct evidence for Hans' wish to sleep with his mother, for his hatred or fear of his father, nor for the supposed relation between horses and Hans' father.

5.14.4.2.2 Behavioral

Behavioral perspectives emphasize the role of conditioning (classical, operant, and vicarious) in the acquisition of phobias. For instance, case reports, such as Little Peter (Jones, 1924a; 1924b) and Little Albert (Watson & Rayner, 1920), demonstrated how a child could acquire a phobia of a previously neutral stimulus (e.g., a rat) by contiguous pairing with an aversive event. Empirical and case reports have supported both direct (classical conditioning) and indirect (vicarious exposure) conditioning models (e.g., Liddell & Lyons, 1978; Merckelbach, De Ruiter, Van De Hout, & Hoekstra, 1989).

Subsequent research has extended initial, conditioning theory to "neoconditioning" theory, reflected by the writings of Rescorla (1988) and Davey (1992). In Rescorla's and Davey's reformulation of conditioning, the emphasis is not on contiguous pairing of stimuli but on one's expectancy that the conditioned stimulus predicts the unconditioned stimulus and by one's evaluation of the unconditioned stimulus. According to this reformulation, learning that a stimulus predicts another occurs separately from learning that the predicted stimulus is aversive. For example, by often playing near a neighbor's house in which a big dog sits in the yard, a child may learn that the house is associated with the presence of a big dog. If the child is later bitten by a dog in some other location, the child may come to fear the neighbor's house because it reliably predicts a stimulus (a dog) that has acquired aversive properties. Reiss (1991), similarly, delineated a neoconditioning theory, one that also emphasizes a cognitive expectancy account that involves certain individual personality differences, such as anxiety sensitivity.

Also representing an important extension of conditioning theory was Mowrer's (1939) two-factor model which postulated Pavlovian conditioning in the initial acquisition of fear, and operant conditioning in the maintenance of fear in that fear serves as an acquired source of motivation or drive, (i.e., behaviors that reduced fear are reinforced). The adequacy of Mowrer's two-factor model in explaining the persistence of phobic fear and avoidant behavior has been questioned, however, because many individuals with phobias can confront the phobic stimulus without either covert avoidance or a subsequent aversive unconditioned stimulus (Menzies & Clarke, 1995). That is, according to Mowrer, this should lead to reductions in fear and avoidant behavior, but this does not always occur (Clarke & Jackson, 1983). Modified conditioning explanations consequently have been developed to account for this, such as the partial irreversibility of fear account (Solomon & Wynne, 1954), the preparedness account (Seligman, 1970, 1971), the incubation account (Eysenck, 1979), the safety signal account (Rachman, 1983), and the serial conditioned stimulus account (Stampfl, 1991).

Yet another way to conceptualize phobia acquisition (or reacquisition) is the notion of dishabituation of formerly mastered fears. More specifically, this is the notion that fears are not unlearned but are tempered or assuaged via habituation or extinction in certain contexts. Bouton (1991) reviewed evidence in support of this view, suggesting that extinction of fear responses do not involve unlearning of previously learned associations; rather, it involves learning to discriminate contexts in which the conditioned stimulus–unconditioned stimulus relationship exists from contexts in which the relationship does not exist. Hence, from this perspective, reinstatement of children's fears can occur when children are reexposed to fearful stimuli in situations that are not associated with safety.

Rachman (1977) espoused the importance of additional pathways in "indirect" fear acquisition, namely, vicarious exposure and the transmission of information or instruction. Support for these pathways has been provided in the subclinical fears of adults (e.g., Hekmat, 1987; Murray & Foote, 1979; Ost, 1985; Rimm, Janda, Lancaster, Nahl, & Dittmar, 1977), and in the fears of nonclinic referred children (Milgrom, Mancle, King, & Weinstein, 1995; Ollendick, King & Hamilton, 1991).

Taken all together, empirical support for the various conditioning reformulations and extensions have been uneven across laboratories, samples, disorders, and so on. This unevenness helped, in part, to spur the "cognitive revolution" that occurred in behavioral psychology in the 1970s. This perspective is discussed next.

5.14.4.2.3 Cognitive

Maladaptive cognitions and cognitive processes are believed to play an important role in the etiology and/or maintenance of SP. It is unclear, however, whether such cognitive activity is the cause or consequence of the fear response. Also unclear is how cognitions and/or cognitive processes vary with development. In contrast to the research literature on cognitive explanations of SP in adults, literature on this issue in children is sparse. Moreover, the extant studies are based on either nonclinic samples of children with different types of fears, or on clinic samples of children with different types of anxiety disorders. Nevertheless, these studies provide some useful information which may shed light on the role of cognitions and cognitive processes in children with SP.

In terms of cognitions, Kendall and colleagues (e.g., Kendall & Chansky, 1991) have suggested a heuristic distinction between cognitive deficiencies (i.e., a lack of cognitive activity such as problem solving, planning, or perspective taking) and cognitive distortions (i.e., cognitive activity that is characterized by misperceptions, exaggerations, or overattending to environmental threat). Children with excessive fear and anxiety are hypothesized to present with the latter. In support of this hypothesis, early research on the cognitive activity of highly anxious children indicated that these children are more likely than their nonanxious counterparts to engage in cognitive distortions, such as thoughts that focus on negative evaluations or unfavorable social comparisons (e.g., Francis, 1988; Prins, 1985; 1986; Zatz & Chassin, 1985).

Prins (1986), for example, examined children's self-speech and self-regulation during a fear-provoking task. Forty-four children (aged 8–12 years) with varying levels of fear of water were asked to perform three tasks in a pool that provoked differential levels of fear (e.g., jumping or diving from a low board, diving from a high board). The children were classified into three groups (high, moderate, and low anxious) based on their performance during a regular swimming lesson at school, observed by the researchers. Children's self-statements (positive, neutral, negative) were assessed via interviews immediately following each task. Relevant here were the findings that the children in the high anxious group engaged in significantly more negative self-statements (e.g., "I'll fall hard," "It will go wrong," "Last time I fell on my belly,") relative to the children in the moderate and low anxious groups. Interestingly, however, no differences were observed among the groups in the frequency of positive or neutral self-statements.

In a study using adolescents recruited from high schools, King, Mietz, Tinney, and Ollendick (1995) compared the cognitions of adolescents who were low and high in test anxiety. Classification of low (*N* = 25; mean age = 15.6 years) and high (*N* = 22; mean age = 15.75 years) test anxiety was based on the adolescents' scores (i.e., the bottom and top 5%) on the Test Anxiety Scale for Children (TASC; Sarason, Davidson, Lighthall, Waite, & Ruebush, 1960). The two groups were compared on psychiatric diagnoses, levels of fear, anxiety, depression, and cognitions. Cognitions were assessed after a brief math and language test using the Children's Cognitive Assessment Questionnaire (CCAQ; Zatz & Chassin, 1983, 1985). The CCAQ has 50 statements about different types of thoughts: on-task, off-task, positive and negative self-evaluations, and coping self-statements. Findings with respect to diagnoses indicated that of the high test anxious youth that were interviewed (*n* = 18), 61% had an anxiety or phobic disorder (four youth had SP) compared to 2% of children in the low anxious group. With respect to cognitions, the adolescents in the high test anxiety group reported more off-task thoughts (e.g., "I wish this were over"), more negative self-evaluations (e.g., "I'm doing poorly on this"), more coping thoughts (e.g., "try to relax"), and fewer positive self-evaluations (e.g., "I do well on tests like this") than the adolescents in the low test anxious group. According to King et al., these results are consistent with the notion that nonnegative thinking may be related more strongly to adjustment than positive thinking (i.e., Kendall's notion of the "power of non-negative thinking"). Clinically, the implications are that it might be important to teach children with high test anxiety to control their off-task and negative self-talk rather than teaching the development of coping self-statements.

There is also a growing body of research on the factors that influence the thinking of anxious children. Some of these studies have included subsamples of youth with SP (Barrett, Rapee, Dadds, & Ryan, 1996), while some have not specified the specific types of anxiety disorders comprising the samples (e.g., Chorpita, Albano, & Barlow, 1996). Some of the factors that have been examined include: cognitive interpretive style (Barrett et al. 1996; Chorpita et al., 1996); causal attribution (Bell-Dolan & Wessler, 1994); and attentional bias (e.g., Vasey, Daleiden, Williams, & Brown, 1995). For example, in Barrett et al.'s study on children's cognitive interpretive style, children (aged 7–14 years) with anxiety disorders (*N* = 152; 27 had SP) were presented with several ambiguous situations and were asked to provide interpretations of the situations and to problem-solve a plan of action. One-half of the situations were social in nature (e.g., "You are going to school and a group of children are following you ... ") and the other half were physical (e.g., "On the way to school you feel funny in your tummy ... "). The comparison groups were nonclinic-referred children (*n* = 26) and clinic-referred children with *DSM-III-R* diagnoses of Oppositional Defiant Disorder (ODD; *n* = 27). The results indicated that the children with anxiety disorders perceived more threat in the ambiguous situations than the nonclinic children, but less threat relative to children with ODD. In addition, relative to both comparison groups, the children with anxiety disorders generated more avoidant plans as solutions to these situations. A closer examination of the responses of children with different primary anxiety diagnoses indicated that children with SP and OAD gave more threat interpretations to the physical situations than children with SAD. Moreover, children with SP gave more avoidant solutions to physical situations than children with any other type of anxiety disorder.

In terms of attentional bias, Martin, Horder, and Jones (1992) compared the response time for naming the colors of written words using a modified Stroop task of nonclinical school children (aged 6–13 years) with and without fears of spiders. Fearfulness was based on children's affirmative or negative responses to both of two questions: (i) whether the child liked spiders and (ii) whether the child would pick up a spider. All children completed a Stroop task that included spider-related and neutral words. The children with fears of spiders were found to be slower than the children without fears of spiders in their naming the color of spider-related words (e.g., web, crawl, hairy), supporting the notion of attentional bias in fearful children. Vasey et al. (1995) found similar results when they examined modified Stroop test performance using a sample of children (aged 9–14 years) with and without anxiety disorders ($N = 12$ in each group), supporting the hypothesis that children with anxiety disorders exhibit an attentional bias toward emotionally threatening stimuli. Vasey et al. further discuss how anxious children's narrowing of attention (i.e., hyperattention) may increase a child's focus on danger and threat stimuli, increase the encoding of threatening material, and thus maintain or increase levels of fear and anxiety.

In sum, empirical support is beginning to accumulate for the role of cognitions and certain cognitive processes in the etiology and/or maintenance of anxiety disorders. However, few studies have focused exclusively on the role of cognitions and cognitive processes in clinic samples of children diagnosed specifically with SP. This issue thus awaits further investigation.

5.14.4.2.4 Family

Familial contributions are also believed to play an important role in the development of SP. As noted in section 5.14.4.1, considerable evidence has documented that anxiety/fear "run in families." Specifically, findings from top-down studies generally reveal that children whose parents have an anxiety disorder are at risk of developing an anxiety disorder themselves (e.g., Silverman, Cerny, Nelles, & Burke, 1988; Turner, Beidel, & Costello, 1987; Weissman, Leckman, Merikangas, Gammon, & Prusoff, 1984). However, the specific mechanisms of transmission remain unclear. Two processes, parental modeling and parental child management (e.g., use of rewards and punishments) of fearful/avoidant behavior, have received the most attention by researchers. It is noted, however, that most of the studies conducted have not used clinic samples of children with SP. Most have used clinic samples of children with different types of anxiety disorders or nonclinic samples of children with different types of fears.

In terms of parental modeling, Milgrom et al. (1995) examined parental modeling as part of a larger study on origins of childhood dental fears. Children ($N = 895$; aged 5–11 years) recruited from public schools during dental exams, and their mothers or female guardians, were interviewed and administered several questionnaires to assess direct conditioning, parental modeling, and control variables (e.g., socioeconomic status (SES)). Most relevant here are the findings that among the parental modeling variables, maternal level of dental fear and maternal oral health were significant predictors of dental fears in the children, even after controlling for gender, age, and other sociodemographic variables.

Although not containing children with SP, another study conducted by Muris, Steerneman, Merckelbach, and Meesters (1996) provided further evidence for an association between parental modeling, particularly by mothers, and levels of child fear. These authors made use of the notion of "social referencing" in interpreting their findings: that is, children's tendencies to "look to" their parents, particularly their mothers, for emotional information as a way to appraise the dangers of a situation (Muris et al., 1996).

In discussing the contribution of parental modeling to child fear and anxiety problems, Silverman et al. (1988) suggested that it is not parental fear *per se* that is critical, but the extent to which parental fear is associated with parental avoidant behavior. Capps, Sigman, Sena, Henker, and Whalen (1996) provided evidence for this notion in a study in which they compared 16 offspring (aged 8–14 years) of parents with AG with 16 offspring of nonpsychiatric parents who were matched on age, gender, SES, and ethnicity. No significant correlation was found between parents' and children's self-ratings of fearfulness. However, the number of situations avoided by parents with AG was correlated significantly with the children's anxiety symptom scores, assessed via a structured interview.

In addition to parental modeling, research suggests that certain types of parental child management approaches may be associated with child fear and anxiety problems (e.g., Bush, Melamed, Sheras, & Greenbaum, 1986; Zabin & Melamed, 1980). For example, in Zabin and Melamed (1980), parents completed a questionnaire in which they endorsed the child management approach (e.g., positive reinforcement, punishment) that they used most often to encourage their child to deal with various fearful situations (e.g., getting a haircut, going to the dentist). Parental scores on this questionnaire were correlated with several indices of child trait anxiety (based on child and parent ratings) and state anxiety (based on child, parent, observer, and physiological measures) during their hospitalization for a fearful pediatric procedure.

Although measures of parent discipline varied in the strength of their relation to each of the child anxiety measures, and at each assessment period, overall, the patterns of results revealed that parental disciplinary techniques were associated with measures of child anxiety. More specifically, children whose parents endorsed using positive reinforcement and modeling of nonfearful behavior displayed low levels of child anxious behaviors (based on observer, child, and physiological measures). In contrast, children whose parents endorsed using punishment, physical force, or reinforcement of dependency displayed high levels of anxious behaviors (based on observer, parent, child, and physiological measures). Research findings obtained via direct observations of parent–child interactions of clinic samples of children corroborate this link between parental modeling as well as certain types of parental child management approaches and children's fear/anxiety problems (Barrett et al., 1996; Chorpita et al., 1996; Dadds, Barrett, Rapee, & Ryan, 1996). As noted, however, these studies contain too few children with SP (or the numbers are unspecified) to draw firm conclusions about the role of the family and family processes in the development of this particular disorder in children.

5.14.5 EPIDEMIOLOGY

Estimates of SP in childhood vary depending on the methodology used. Some of these methodological differences include variations in the informant (child, parent, both), the assessment method (structured vs. unstructured interviews), the sample (clinic or community), the age and gender of the participants, and whether an impairment index was included as part of the definition. Despite these differences, epidemiological studies provide a clearer picture than ever before of the prevalence of SP.

Table 2 presents findings from community studies, which have estimated rates of SP as ranging between 2 and 9% (e.g., Anderson, Williams, McGee, & Silva, 1987; Bird et al., 1988; Costello et al., 1988; Fergusson, Harwood, & Lynskey, 1993; McGee et al., 1990). The rate of SP among children who present for treatment has been estimated as being approximately 7% of referrals at outpatient community mental health clinics (Graziano & DeGiovanni, 1979; Silverman & Kearney, 1992), and approximately 30–35% of referrals at outpatient child anxiety specialty clinics (e.g., Last et al., 1987, 1992).

5.14.5.1 Prevalence of Phobia Subtypes

Information on the prevalence of each of the subtypes of SP among children is based largely on studies of youth who present to child anxiety specialty clinics. For example, in Strauss and Last's (1993) study, described earlier, in which the clinical characteristics of children with SOP and SP were compared, the most common subtypes of phobias of the 38 children (mean age 11.1 years) with a primary diagnosis of SP were: darkness 29%, school 24%, dogs 16%, other animals 8%, heights 8%, insects 8%, elevators 5%, and 3% for swimming, needles, drains, taxidermy.

In the Ginsburg and Silverman (1996) study, also described earlier, the most common subtypes of phobias were the same for the Hispanic and Caucasian children, namely, darkness/sleeping alone and small animals (i.e., dogs and cats). Silverman and Kearney (1992), in their survey of practicing psychologists, found that the most common types of fears treated were those relating to separation from caregivers, school, social situations, tests, and the dark.

5.14.5.2 Sociodemographic Risk Factors

Information detailing sociodemographic risk factors for SP is sparse. In this section, findings from community- and clinic-based studies are summarized briefly. These studies have tended to focus on SES, race, and gender.

5.14.5.2.1 Socioeconomic status

Most of what is known about SES as a risk factor for SP is based on samples of children who present at anxiety disorders specialty clinics. For example, in the Strauss and Last (1993) study in which the sociodemographic and clinical characteristics of children with SOP and SP were compared, Hollinghead index was used to assess SES. Of the subsample of children with SP, 22% were classified as levels I (major professionals) and II (medium business, minor professional, technical), 59% were classified as levels III (skilled craftsmen, clerical, and sales workers) and IV (machine operators and semiskilled workers) and 19% were classified as level V (unskilled workers, menial service workers). Last et al. (1992), in their comparison of the sociodemographic and clinical characteristics of a large sample of children with anxiety disorders ($N = 188$), found that most of the children with SP ($n = 80$) came from middle- to upper-class families.

In another study, annual family income was obtained for a sample (ages 6–17 years; mean age 10.1 years) of Caucasian ($N = 143$) and Hispanic ($N = 99$) families of youth with anxiety disorders (Ginsburg & Silverman, 1996). Of the Caucasian families, 8.5% had incomes of $15 000 or less, 22.6% had incomes of $15 000–30 000, and 68.9% had incomes over $30 000. Among the Hispanics, 29.2% had incomes of $15 000 or less, 28.1% had incomes of $15 000–30 000, and 42.7% had incomes over $30 000.

Table 2 Prevalence of specific phobias in epidemiological studies.

Study	*N*	*Age*	*M:F ratio*	*Rate of SP*
Anderson et al. (1987)	792	11	1:1.7	2.4
McGee et al. (1990)	943	15	1:2.9	3.6
Bird et al. (1988)	777	4–16		2.6
Costello et al. (1988)	789	7–11	1:2.0	9.2
Fergusson et al. (1993)	961	15		5.1
Kashani et al. (1987)	150	14–16	G > B	4.7
Kashani and Orvaschel (1990)	210	8,12,17	1:6	3.3

Source: Anderson (1994) and Silverman and Ginsburg (1998).

5.14.5.2.2 Racial/ethnic variations

Rates of SP and other psychiatric disorders were assessed in a community sample of youth in Puerto Rico using a multistage assessment procedure consisting of the parent and teacher versions of the Child Behavior Checklist (CBCL; Achenbach & Edelbrock, 1983) at the first stage (for screening), and a structured diagnostic interview at the second stage (for diagnosing) (Bird et al., 1988). As can be seen from Table 2, prevalence rates for SP in Puerto Rican youth (2.6%) were similar to those found in Caucasian youth.

Similar rates of SP in samples of African-American and White children (Last & Perrin, 1993) and in samples of Hispanic and White children (Ginsburg & Silverman, 1996), obtained from child anxiety specialty clinics, have also been reported. Specifically, Last and Perrin (1993) found that SP was among the "top" two primary diagnoses in both of their groups, both at intake (26.7% for African-Americans and 18.0% for Whites) and for lifetime prevalence (50% for African-Americans and 40.3% for Whites). Ginsburg and Silverman (1996) found SP to be the top primary diagnosis at intake in both of their groups, with a rate of 38.4% for Hispanics and 47.6% for Whites.

5.14.5.2.3 Gender

With respect to gender, several epidemiological studies have found rates of SP to be higher in girls than boys (see Table 2). This is consistent with the studies on nonclinical fears. In contrast, studies examining rates of SP in clinic samples of children generally have found rates of SP to be equal for boys and girls (Ginsburg & Silverman, 1996; Last & Perrin, 1993).

5.14.6 ASSESSMENT

The most widely used methods for assessing children's cognitive/subjective, behavioral, and physiological response systems (Lang, 1977) are summarized in this section. Choosing the "best" assessment procedure depends largely on the specific goal or desired outcome (e.g., self-rating scales may be best for screening, interview schedules for diagnosing) (Silverman & Kurtines, 1996). A protypical approach to assessment involves eliciting information from multiple perspectives (child, parent, teacher) and examining each of the three response systems. It is noted, however, that manifestations of fear/anxiety measured by these three systems do not necessarily covary (i.e., desynchrony).

5.14.6.1 Cognitive/Subjective

The most commonly used strategies to assess the cognitive/subjective system include clinical interviews, self-report questionnaires, and self-monitoring procedures, although these strategies may also be used to assess the behavioral (e.g., "do you try as hard as you can to stay away from dogs?") and physiological response systems (e.g., "does your heart beat fast when you see a dog?").

Clinical interviews are characterized as either unstructured (or nonstandardized) or structured (or standardized) (Richardson, Dohrenwend, & Klein, 1965). In the unstructured interview, the specific questions and the sequence of the questions asked vary depending on the interviewees' responses and on the interviewers' subjective judgements and training. Structured interviews contain a set of questions that are asked of all respondents in the same sequence (see Silverman, 1991, 1994, for reviews). Because of problems with the reliability of diagnoses derived from unstructured interviews (Edelbrock & Costello, 1984), the use of structured and semistructured interview schedules has increased through the years. Of the structured interviews available (e.g., the Schedule for Affective Disorders and Schizophrenia in School-Age Children; Puig-Antich & Chambers, 1978, the Diagnostic Interview for

Children and Adolescents; Herjanic & Reich, 1982, the Diagnostic Interview Schedule for Children; Costello, Edelbrock, Dulcan, Kalas, & Klaric, 1984), only the Anxiety Disorders Interview Schedule for Children (ADIS-C/P; Silverman & Nelles, 1988a; Silverman & Albano, 1996) was designed specifically for use with children with SPs and other anxiety disorders. Its utility for diagnosing SPs in youth has been studied extensively (Rapee, Barrett, Dadds, & Evans, 1994; Silverman & Eisen, 1992; Silverman & Nelles, 1988a; Silverman & Rabian, 1995), and have yielded excellent reliability coefficients (Kappa coefficients ranging from 0.84 to 1.0, depending on the reliability paradigm used (i.e., interrater, test–retest).

To assess the presence of SP, the ADIS-C/P contains a list of objects/events (e.g., dark, dogs, needles) to which the child/parent is instructed to rate the child's degree of fear and avoidance using a Feelings Thermometer. The Feelings Thermometer is useful in that it simplifies the rating task for children and removes some of the variability attributed to language skills that occur when young children respond to questionnaires (Barrios, Hartmann, & Shigetomei, 1981). In addition to fear ratings, the ADIS-C/P provides for an assessment of degree of interference by obtaining such ratings from the child and parent alike, also using a modified Feelings Thermometer. Additional sections of the ADIS-C/P contain questions addressing onset, course, and etiology of each specific phobia.

Another frequently used method for assessing the cognitive/subjective system are self-rating scales. The most widely used scale for assessing fears in children is the Revised Fear Survey Schedule for Children (FSSC-R; Ollendick, 1983). This measure provides information on the number, severity, and types of fears that children experience in 80 different situations. Factor analysis has revealed five factors including: (i) fear of the unknown, (ii) fear of failure and criticism, (iii) fear of minor injury and small animals, (iv) fear of danger and death, and (v) medical fears (Ollendick, 1983).

Other self-rating scales have been developed to assess a broad range of children's fears, including the Louisville Fear Survey (Miller, Barrett, Hampe, & Noble, 1972) and the Children's Fear Survey Schedule (Ryall & Dietiker, 1979). In addition, self-rating scales have been developed to assess children's specific fears, such as dental procedures (Children's Fear Survey Schedule; Melamed & Lumley, 1988), darkness (Nighttime Coping Response Inventory and Nighttime Fear Inventory; Mooney, 1985), medical procedures/hospital (Fears Schedule; Bradlyn, 1982), snakes (Snake Attitude Measure; Kornhaber & Schroeder, 1975), and tests (Test Anxiety Scale for Children; Sarason et al., 1960).

Despite the ease in administration and scoring of children's self-rating scales, as well as their apparent face validity, their utility in providing differential diagnoses, particularly of SP, awaits further study. Last, Francis and Strauss (1989), used the FSSC-R to compare the fears of 111 children (aged 5–18 years; mean age not provided) who presented at a childhood anxiety clinic and who met *DSM-III-R* criteria for a range of anxiety disorders. Specifically, 44 children had a primary diagnosis of SAD, 33 had a primary diagnosis of OAD, and 34 had a primary diagnosis of either SOP or SP (combined together as a "phobic group," with the most common subtype being school-related phobias). No differences were found across these diagnostic groups in either their total FSSC-R scores or their factor scores. Interestingly, however, differences among the groups were observed with respect to the total number of intense fears endorsed (i.e., number of items that children rated "a lot"), with the phobic group endorsing the least. Comparisons of the frequency of certain individual FSSC-R items indicated that children in the phobic group were more likely to endorse the school item than children in the OAD group, but were not different from children in the SAD group. Also, less than one-half of the children in the phobic group endorsed "a lot" of fear on the school item, despite the primary reason for being classified in the phobic group was based on reports of school-related fears obtained during a structured interview.

In another study, Perrin and Last (1992) examined the FSSC-R as well as other self-report measures of anxiety using a sample of 213 boys (aged 5–17 years); 105 with primary diagnoses of anxiety disorders (22 with SP), 56 with primary diagnoses of Attention Deficit Hyperactivity Disorder (ADHD), and 49 with no psychiatric diagnoses. With respect to children's fears, there were no differences among the three groups on the total or five subscale scores of the FSSC-R. There were also no significant differences in the FSSC-R scores of the boys with anxiety disorders who were also school refusers with the scores of the boys with ADHD as well as the normal controls.

In summary, although self-rating scales such as the FSSC-R are used widely, questions remain about their discriminative ability between youth with SP and normal controls as well as among various diagnostic groups. However, there may be some utility in a qualitative analysis of the individual items endorsed.

Another way to assess children's cognitive/subjective systems is via self-monitoring procedures. Self-monitoring requires that children self-observe and systematically record the occurrence of their fearful thoughts and behaviors. Self-monitoring data also provide information about the antecedents and consequences of fearful reactions as well as the severity and frequency of these reactions. In the authors' work, children engage in self-monitoring via a "daily diary." This requires the children to keep a record of the situations in which fear/anxiety was experienced, whether they confronted or avoided the situation, accompanying cognitions, and a rating of fear.

A more structured procedure described by Beidel, Neal, and Lederer (1991) requires children to mark on a checklist whether certain events occurred and whether they experienced certain feelings. In a community sample of elementary school children involved in an investigation of test-anxiety, Beidel et al. (1991) examined the reliability of the children's self-monitoring, and also compared the frequency and severity of anxiety-producing events during a two-week period for a text-anxious ($n = 17$) and nontest-anxious ($n = 17$) subsample. Compliance was also evaluated. As Beidel et al. point out, assessing reliability for an event-orient assessment procedure is difficult as events are variable and fluctuating. Not surprisingly, therefore, the resulting reliability coefficient (based on an n of 13) was relatively modest and not significant ($r = 0.50$). As preliminary support for validity, however, the test-anxious children reported significantly more emotional distress and more negative behaviors (e.g., crying, somatic complaints) than the nontest-anxious children. Compliance was found to be mixed, with 56% of the test-anxious children completing the forms on 10 of the 14 days, and with the percentage of children completing the forms for all 14 days dropping to 31%.

Similar types of analyses of self-monitoring assessment data using clinic samples of children with SP need to be done. In addition, reactivity effects and whether self-monitoring data are sensitive in showing clinical change from pre- to post-treatment warrant investigation. Methods to increase children's compliance with and accuracy of self-monitoring should also be studied.

5.14.6.2 Behavioral

Behavioral observation procedures provide useful information about the frequency and severity of symptoms/behaviors and their antecedents and consequences. In working with children with SP, the types of symptoms/behaviors that might be identified/assessed may be classified as (i) verbal behavior, such as rate of speech, positive/negative self-references, verbal social skills, etc., and (ii) overt behavior, such as head/gaze aversions, head nods, posture, eye contact, etc. Although there are a number of ways to observe these behaviors, the most frequent ways are on the basis of discrete response occurrences (i.e., response rate or frequency) or the amount of time that the response occurs (i.e., response duration).

Several observational coding procedures have been developed for assessing fear and anxiety in children. These coding procedures were developed to record discrete child behaviors in specific anxiety-provoking settings, such as in dental, medical, or child–parent separation settings (e.g., Glennon & Weisz, 1978; Melamed & Siegel, 1975). Such behavior observation coding schemes can be adapted for use with children with SP. They can also be adapted for use in analogous situations, which might be more feasible for clinicians.

Unfortunately, however, the psychometric properties of behavioral observation procedures in analogous situations have been insufficiently studied. An exception is a study conducted by Hamilton and King (1991) who examined the temporal stability of an analogous observation procedure using 14 children (aged 2–11 years; mean age not provided) with an SP of dogs. The analogous procedure comprised a series of 14 performance tests. Each test required the child to engage in increasingly more difficult interactions with a dog. Two test administrations were conducted seven days apart, prior to treatment. High test–retest reliability was obtained ($r = 0.97$), with 10 of the 14 children obtaining identical approach scores and with the other four children obtaining increased approach scores, (i.e., they displayed less avoidant behavior from Time 1 to Time 2). These results provide support for the temporal stability of approach scores obtained on analogous observation procedures. However, because of the small sample employed in this study and the wide age range, replication of this study would be worthwhile. Additional questions worth examining in the future include the role of differential levels of initial fear, the specific instructions employed (e.g., high demand vs. low demand), and children's developmental level.

5.14.6.3 Physiological

The most common strategies used to assess the physiological system typically involve measuring children's heart rate, blood pressure, and electrodermal activity (e.g., finger or palm sweat; see King, Hamilton, & Ollendick, 1988). These measurements are usually taken before, during, and after child exposure to the feared object or event. There is a paucity of data on the utility of these measures using samples of children with anxiety disorders, including SP. This is likely due to difficulties involved in using these procedures, including children's frequent movements, children's fears of the test instrumentation (e.g., the electrodes may shock them), and the instruments' sensitivity to nonanxiety influences (e.g., lighting, temperature, noise, novelty). There are also difficulties involved in the interpretive process with children, particularly as they relate to developmental level. For example, in relation to tonic heart rate in children, both heart rate and heart rate variability decrease with age (Shinebourne, 1974). Normative data for psychophysiological measurement with children are also lacking (Venables, 1981).

Despite these difficulties, evidence does suggest that psychophysiological assessment may be a reliable strategy for the assessment of children's anxious symptoms (e.g., Beidel, 1991; Beidel, Fink, & Turner, 1996; Murphy, Alpert, Wiley, & Somes, 1988). Unfortunately, however, the utility of psychophysiological assessments in clinical outcome studies of SP remains to be demonstrated adequately. In a treatment outcome study of childhood SP (described below), the results pertaining to heart rate tended to be uneven, similar to what has been reported in single case-design studies (e.g., Van Hasselt, Hersen, Bellack, Rosenblum, & Lamparski, 1979).

5.14.7 TREATMENT

The research literature shows that the one procedure that is essential to include in a fear or anxiety treatment program, if it is to be successful, is "exposure" (i.e., having the child confront or approach the feared object or event (see Barlow, 1988, for various accounts as to why exposure may work). Exposure may be live (*in vivo*) or imaginal, and with children is generally conducted in a gradual fashion.

The studies reviewed in this section often include an exposure component along with additional cognitive and behavioral strategies. Before discussing this literature, however, several methodological limitations that characterize most of the early studies conducted in this area are worth noting. First, the majority of the treatment outcome studies did not use controled experimental procedures, clinic referred samples, structured diagnostic procedures, multimethod–multisource assessment procedures, or systematic follow-ups (Silverman & Rabian, 1993). In addition, research has insufficiently addressed the relative efficacy of one treatment strategy over another (e.g., cognitive vs. behavioral) or one format over another (e.g., individual vs. group). There has also been very limited research conducted on potential predictors of treatment outcome (e.g., comorbidity, parental anxiety symptoms, treatment duration). Also absent is any type of conceptual framework to help guide clinicians' and researchers' thinking in how to go about developing and implementing intervention approaches for children with SP. With these limitations in mind, a brief overview is presented of the major cognitive and behavioral strategies that have been used in the treatment of childhood SP. A "transfer of control" model is then presented which might serve as a useful framework for developing and implementing psychosocial interventions for children with SP.

The most commonly used cognitive and behavioral procedures that have been used to treat child SP include: systematic desensitization, flooding, modeling, contingency management, relaxation training, and self-control. Descriptions of each of these strategies, along with evidence supporting their effectiveness, have been presented elsewhere (e.g., see King et al., 1988; Silverman & Eisen, 1993), and will not be detailed here.

Among these strategies, contingency management, modeling, and self-control have received the most attention and appear to be the most promising (Morris & Kratochwill, 1983). Briefly, contingency management, based on the principles of operant conditioning, stresses the importance of the causal relation between stimuli and behavior (Morris & Kratochwill, 1983), and relies on the therapist and the parents to rearrange the environment to ensure that positive consequences follow a child's exposure to the fearful stimulus, and to ensure that positive consequences (e.g., attention) do not follow avoidant behavior. Modeling involves the child learning to be less fearful or anxious by observing others handle the fearful object or event. The models observed by the child may be actual, or "live" models, or they may be observed on films or videotapes, or "symbolic" models. Finally, self-control stresses the important contribution of cognitive processes to behavior change and involves the child directly regulating her or his own behavior by applying specific thinking styles when confronted with the fearful stimulus.

All of the above strategies (i.e., contingency management, modeling, self-control), alone or in combination with other strategies, have been found to be effective in treating children who display a wide range of fears including those related to: nighttime (Graziano, Mooney, Huber, & Ignasiak, 1979; Graziano & Mooney, 1980; King, Cranstoun, & Josephs, 1989); darkness (Giebenhain & O'Dell, 1984; Heard, Dadds, & Conrad, 1992; Jackson & King, 1981; Kanfer, Karoly, & Newman, 1975); loud noises (Tasto, 1969; Wish, Hasazi, & Jurgela, 1973; Yule, Sacks, & Hersov, 1974); public speaking (Cradock, Cotler, & Jason, 1978; Fox & Houston, 1981); heights (Croghan & Musante, 1975; Holmes, 1936; Ritter, 1968); animals (Bandura, Blanchard, & Ritter, 1969; Glasscock & MacLean, 1990; Obler & Terwilliger, 1970; Sreenivasan, Manocha, & Jain, 1979); water (Lewis, 1974; Menzies & Clarke, 1993; Ultee, Griffioen, & Schellekens, 1982); menstruation (Shaw, 1990); telephones (Babbitt & Parrish, 1991); newspapers (Goldberg & Weisenberg, 1992); bowel movements (Eisen & Silverman, 1991); needles (Rainwater et al., 1988); medical procedures (Heard et al., 1992; Peterson & Shigetomi, 1981); dental procedures (Krochak, Slovin, Rubin, & Kaplan, 1988; Siegel & Peterson, 1980); traveling in cars (Stedman & Murphy, 1984; Thankachan, Mishra, & Kumaraiah, 1992); and illness/diseases (Hagopian, Weist, & Ollendick, 1990).

In an early controlled group treatment study that focused on treating childhood phobia, Miller et al. (1972) randomly assigned phobic children ($N = 67$; aged 6–15 years) (67% had a primary fear of school) to one of three treatment conditions: reciprocal inhibition ($n = 21$), play therapy ($n = 23$), and a waitlist control ($n = 23$). Treatment was provided three times a week for eight weeks. Outcomes, based on parent and clinician reports, revealed no significant differences between the two treatment conditions, but children in both treatment conditions improved relative to the children in the waitlist control. Moreover, an examination of children in both active treatments revealed that younger children (aged 6–10 years) appeared to evidence significantly greater improvements relative to older children (aged 11–15 years) in both conditions.

Menzies and Clarke (1993) compared three fear-reduction strategies for children ($N = 48$; aged 3–8) with water phobia. Children were assigned randomly to one of four groups: (i) *in vivo* exposure plus vicarious exposure (watching a model), (ii) vicarious exposure only, (iii) *in vivo* exposure only, and (iv) assessed only. All children received three treatment sessions and assessments were based on the following: behavioral tasks in water (children asked to engage in different tasks of varying difficulty level such as stepping into a pool, putting their face under water, etc.); self-ratings of water fears; and a combined observer and clinician rating of the children's overall reactions to the phobic situation while the child was engaged in completing tasks in the pool. On all measures, the results revealed that the children in both exposure groups (groups i and iii) improved relative to the assessed only and vicarious only (groups ii and iv). No differences were found between the exposure groups (group i and iii).

The authors, in collaboration with others at the Child and Family Psychology Social Research Center, completed a controlled clinical trials study for children with phobic disorders. This study compared the relative effectiveness of an exposure-based contingency management (CM) treatment, an exposure-based self-control (SC) treatment, and an education-support (ES) comparison condition in which exposure was discussed but was not directly prescribed (such as through CM or SC procedures). A total of 99 children (aged 6–16 years; mean age = 9.9 years) and their parents participated in the study. Fifty-two were boys and 47 were girls. All children met criteria for a primary *DSM-III-R* diagnosis for phobic disorder, including SP ($n = 82$), SOP ($n = 10$), or AG ($n = 7$), based on a structured interview administered separately to the child and parent.

Each of the 99 children and their parents were assigned randomly to one of the three conditions. There were 37 children in the CM condition, 40 in the SC condition, and 22 in ES. Each condition was a 10-week treatment program in which the children and parents were seen in separate treatment sessions with the therapist, followed by a brief conjoint meeting. Results indicated that all the procedures that were used to facilitate the occurrence of exposure, the CM condition, the SC condition, as well as the ES condition, produced effective therapeutic change on all of the main outcome measures (child, parent, and clinicians), including the indices of clinically significant change (i.e., percentage of children who no longer met diagnostic criteria). More impressively, these gains were maintained at three, six and 12 months follow-up. The findings are discussed within the context of the transfer of control model, described in Section 5.14.7.1.

5.14.7.1 Transfer of Control Model

According to the transfer of control model, the therapist is viewed as an expert consultant who possesses knowledge of the skills and methods necessary to produce long-term child therapeutic change. Effective change involves a gradual transfer of the knowledge, skills, and methods from the therapist to the parent to the child (Silverman & Kurtines, 1996). In treating children with SP, the focus of treatment is on controlling the occurrence and successful implementation of exposure to the feared object or situation. Based on findings from our treatment outcome study, it is believed that there are several strategies that might be used to facilitate an adequate transfer of control. These include a more parent-focused treatment, such as CM; a more child-focused treatment, such as SC; a combination (CM followed by SC); a less prescribed strategy, such as ES; a group treatment strategy (Ginsburg, Silverman, & Kurtines, 1995a), and even, perhaps, a format that involves a more thorough and systematic intervention on the child and parent relationship or to what is referred to as "dyadic treatment" (Ginsburg, Silverman, & Kurtines, 1995b). The task ahead for clinicians and researchers is to delineate the particular pathway (and thus, treatment strategy or format) that might be most useful for a child with SP and their family.

5.14.8 CASE STUDY

5.14.8.1 Presenting Problem

Jody, a 12-year-old Caucasian boy, was presented by his mother because of his extreme reactions to thunderstorms. She described that on cloudy or rainy days her son would repeatedly ask questions about the weather, watch only the weather channel on TV or listen to the weather radio, report somatic complaints, and generally appear tense and nervous. During thunderstorms, (a common occurrence in South Florida), Jody would scream, cry, and run and hide in his closet. In addition, Jody would refuse to go outside (e.g., to the mailbox, car, mall, etc.) and refused to sleep alone in his room. Although Jody's mother reported that he was always a worrier, his reactions to thunderstorms (and apprehensions about other natural disasters such as tornadoes) had increased dramatically within the past year and were significantly interfering with his and the family's functioning.

5.14.8.2 Assessment and Diagnosis

Jody and his mother were administered respective child and parent versions of the ADIS-C and ADIS-P. Jody also completed several self-rating scales, and he was administered an *in vivo* behavioral task during which time his heart rate was monitored. As noted above, Jody's mother reported that his fearful and avoidant behavior was interfering with various aspects of his and the family's life. For instance, his school performance had begun to decline because he was having difficulty concentrating in school when storms were approaching. Socially, Jody's fears prevented him from going on outings (e.g., fishing, camping) and he had begun to run off the playing field during his little league games if he saw lightning or heard thunder. At home, Jody's fear often prevented the family from completing errands or attending activities if a storm was approaching.

Based on Jody and his mother's responses to the interviews, Jody was assigned a diagnosis of SP, natural environment subtype. On the ADIS-C/P, both Jody and his mother reported that his fear of thunderstorms was a "7" on the 8-point fear thermometer. His scores on the questionnaires were also elevated. Jody's *in vivo* behavioral exposure test, which required him to walk outside with the therapist during an afternoon thunderstorm for five minutes, was terminated because he became too upset.

5.14.8.3 Treatment

Jody was enrolled in a 10-week exposure-based treatment program in which parent and child were seen individually and then briefly together. The program, based on the transfer of control model, involved first teaching Jody's mother the principles of contingency management and then teaching both Jody and his mother child self-control strategies. These strategies were then applied in gradual exposure tasks that involved Jody gradually approaching going outside during storms.

Exposure tasks were both imaginal and *in vivo* and were based on a hierarchy that varied how close to being outside Jody would get during thunderstorms. Initial *in vivo* steps on the hierarchy, during which Jody remained outside of the closet but inside the house during a storm, were successful, but as the exposures became more difficult, requiring him to go outside, Jody was unable to complete the exposures and explained to the therapist that storms were just too unpredictable. In discussing these difficulties with Jody's mother, she indicated that she felt Jody's fears were not totally irrational. Many Florida residents had lost their homes during Hurricane Andrew in 1992 and there had been recent incidents of tornadoes in the news.

To deal with these concerns, subsequent treatment sessions addressed some of the common obstacles encountered with treating SP. For instance, emphasis was placed on reviewing with Jody's mother the importance of parental modeling of courageous behavior. In part, this was accomplished by discussing common parental "protective behaviors" and her own fears related to hurricanes. In addition, Jody and his mother were helped to think of alternative ways of handling their fears, including specific steps they could take so that they could be maximally prepared for hurricanes. This included obtaining accurate information about hurricane preparedness and warnings (e.g., that although hurricanes are uncontrollable, there is usually plenty of advance warning and time to prepare). This approach helped to empower the family and enabled them to have an "action plan" if Jody's worst fears were to materialize (e.g., that a hurricane was forecast). All of these methods were helpful and subsequent contracts were successfully carried out, including the one that required him to go fishing with his grandfather on a rainy day.

In the final phase of the transfer of control, wherein self-control strategies were used (including self-reward to replace parental rewards), Jody practiced examining the specific cognitions which played a role in maintaining his fear (e.g., "the hurricane will blow my house away"). Jody learned to recognize when he was afraid or worried, to employ more adaptive coping thoughts and behaviors, and to praise himself for doing so.

5.14.8.4 Post-treatment and Follow-up

At post-treatment assessment, readministration of the relevant sections of the ADIS-C and ADIS-P revealed that Jody no longer met criteria for SP and there was no longer any interference in his daily functioning. In addition, his scores on the questionnaires decreased markedly and Jody was able to do the behavioral exposure test. These gains were maintained at three and 12-month follow-up assessments.

5.14.9 SUMMARY

This chapter presented an overview of what is known about specific phobias in children, beginning with a discussion about phenomenology, including variations which may be associated with age, gender, and race/ethnicity. Information about the onset and stability, diagnostic criteria, and issues related to differential diagnosis was presented. Thinking about etiology, both biological and psychosocial perspectives, was presented next. This was followed by a summary of the epidemiological data which indicate an average prevalence rate of about 2–9% (based primarily on Caucasian community samples). The chapter discussed assessment and treatment procedures that have been found useful for working with children with specific phobias, including the evolving transfer of control model as a possible framework to help guide clinicians' and researchers' thinking in how to develop and implement interventions. A case study was presented to illustrate the issues discussed.

5.14.10 REFERENCES

Abe, K., Oda, N., Ikenaga, K., & Yamada, T. (1993). Twin study on night terrors, fears and some physiological and behavioral characteristics in childhood. *Psychiatric Genetics, 3,* 39–43.

Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile.* Burlington, VT: University of Vermont, Department of Psychiatry.

Agras, W. S., Chapin, H. N., & Oliveau, D. C. (1972). The natural history of phobia. *Archives of General Psychiatry, 26,* 315–317.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Press.

Anderson, J. (1994). Epidemiological issues. In T. H. Ollendick, N. J. King, & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in children and adolescents* (pp. 43–66). New York: Plenum.

Anderson, J. C., Williams, S., McGee, R., & Silva, P. A. (1987). *DSM-III* disorders in preadolescent children. *Archives of General Psychiatry, 44,* 69–76.

Babbitt, R. L., & Parrish, J. M. (1991). Phone phobia, phact or phantasy? An operant approach to a child's disruptive behavior induced by telephone usage. *Journal of Behavior Therapy and Experimental Psychiatry, 22,* 123–129.

Bandura, A., Blanchard, E. B., & Ritter, B. (1969). Relative efficacy of desensitization and modeling approaches for inducing behavioral affective and attitudinal changes. *Journal of Personality and Social Psychology, 13,* 179–199.

Barlow, D. H. (1988). *Anxiety and its disorders: The nature and treatment of anxiety and panic.* New York: Guilford Press.

Barrett, P. M., Rapee, R. M., Dadds, M. M., & Ryan, S. M. (1996). Family enhancement of cognitive style in anxious and aggressive children. *Journal of Abnormal Child Psychology, 24,* 187–203.

Barrios, B. A., Hartmann, D. P., & Shigetomi, C. (1981). Fears and anxieties in children. In E. J. Mash & L. G. Terdal (Eds.), *Behavioral assessment of childhood disorders* (pp. 259–304). New York: Guilford Press.

Bauer, D. H. (1976). An exploratory study of developmental changes in children's fears. *Journal of Child Psychology and Psychiatry, 17,* 69–74.

Beidel, D. C. (1991). Social phobia and overanxious disorder in school-age children. *Journal of the American Academy of Child and Adolescent Psychiatry, 3,* 545–552.

Beidel, D. C., Neal, A. M., & Lederer, A. S. (1991). The feasibility and validity of a daily diary for the assessment of anxiety in children. *Behavior Therapy, 22,* 505–517.

Beidel, D. C., Fink, C. M., & Turner, S. M. (1996). Stability of anxious symptomatology in children. *Journal of Abnormal Child Psychology, 24,* 257–270.

Bell-Dolan, D. J., Last, C. G., & Strauss, C. C. (1990). Symptoms of anxiety disorders in normal children. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 759–765.

Bell-Dolan, D. J., & Wessler, A. E. (1994). Attributional style of anxious children: Extensions from cognitive theory and research on adult anxiety. *Journal of Anxiety Disorders, 8,* 79–96.

Bernstein, G. A., & Garfinkel, B. D. (1986). School phobia: The overlap of affective and anxiety disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 25,* 235–241.

Biederman, J., Rosenbaum, J. F., Bolduc-Murphy, E. A., Faraone, S. V., Chaloff, J., Hirshfeld, D. R., & Kagan, J. (1993a). A three-year follow-up of children with and without behavioral inhibition. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 814–821.

Biederman, J., Rosenbaum, J. F., Bolduc-Murphy, E. A., Faraone, S. V., Chaloff, J., Hirshfeld, D. R., & Kagan, J. (1993b). Behavioral inhibition as a temperamental risk factor for anxiety disorders. *Child and Adolescent Psychiatric Clinics of North America, 2,* 667–684.

Biederman, J., Rosenbaum, J. F., Hirshfeld, D. R., Faraone, V., Bolduc, E., Gersten, M., Meminger, S., & Reznick, S. (1990). Psychiatric correlates of behavioral inhibition in young children of parents with and without psychiatric disorders. *Archives of General Psychiatry, 47,* 21–26.

Bird, H. R., Canino, G., Rubio-Stipec, M., Gould, M. S., Ribera J., Sesman, M., Woodbury, M., Huertas-Goldman, S., Pagan, A., Sanches-Lacay, A., & Moscoso, M. (1988). Estimates of the prevalence of childhood maladjustment in a community survey in Puerto Rico. *Archives of General Psychiatry, 45,* 1120–1126.

Boldizar, J. P. (1991). Assessing sex typing and androgyny in children: The Children's Sex Role Inventory. *Developmental Psychology, 27,* 505–515.

Bouton, M. E. (1991). Sources of relapse after extinction in Pavlovian and instrumental learning. *Clinical Psychology Review, 11,* 123–140.

Bradlyn, A. S. (1982). *The effects of a videotape preparation package in reducing children's arousal and increasing cooperation during cardiac catheterization.* Unpublished doctoral dissertation, University of Mississippi.

Burglass, D., Clarke, J., Henderson, A. S., Kreitman, M., & Presley, A. S. (1977). A study of agoraphobic housewives. *Psychological Medicine, 7,* 73–86.

Burke, A. E., & Silverman, W. K. (1987). The prescriptive treatment of school refusal. *Clinical Psychology Review, 29,* 570–574.

Bush, J. P., Melamed, B. G., Sheras, P. L., & Greenbaum, P. E. (1986). Mother-child patterns of coping with anticipatory medical stress. *Health Psychology, 5,* 137–157.

Capps, L., Sigman, R., Sena, B., Henker, B., & Whalen, C. (1996). Fear, anxiety and perceived control in children of agoraphobic parents. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 37,* 445–452.

Carey, M. P., Dusek, J. B., & Spector, I. P. (1988). Sex roles, gender, and fears: A brief report. *Phobia Practice and Research Journal, 1,* 114–120.

Caron, C., & Rutter, M. (1991). Comorbidity in child psychopathology: Concepts, issues, and research. *Journal of Child Psychology and Psychiatry, 32,* 1063–1080.

Chorpita, B. F., Albano, A. M., & Barlow, D. H. (1996). Cognitive processing in children: Relationship to anxiety and family influences. *Journal of Clinical Child Psychology, 25,* 170–176.

Clarke, J. C., & Jackson, J. A. (1983). *Hypnosis and behavior therapy: The treatment of anxiety and phobias.* New York: Springer.

Costello, E. J., Costello, A. J., Edelbrock, C. S., Burns, B. J., Dulcan, M. J., Brent, D., & Janiszewski, S. (1988). DSM-III disorders in pediatric primary care: Prevalence and risk factors. *Archives of General Psychology, 45,* 1107–1116.

Costello, A. J., Edelbrock, C. S., Dulcan, M. K., Kalas, R., & Klaric, S. H. (1984). *Report to NIMH on the NIMH diagnostic interview schedule for children (DISC).* Washington, DC: National Institute of Mental Health.

Cradock, C., Cotler, S., & Jason, L. A. (1978). Primary prevention: Immunization of children for speech anxiety. *Cognitive Therapy and Research, 2,* 389–396.

Croake, J. W. (1969). Fears of children. *Human Development, 12,* 239–247.

Croghan, L. M., & Musante, G. J. (1975). The elimination of a boy's high building phobia by in vivo desensitization and game playing. *Journal of Behavior Therapy and Experimental Psychiatry, 6,* 87–88.

Dadds, M. R., Barrett, P. M., Rapee, R. M., & Ryan, S. (1996). Family process and child anxiety and aggression: An observational analysis. *Journal of Abnormal Child Psychology, 24,* 715–734.

Davey, G. C. L. (1992). Classical conditioning and the acquisition of human fears and phobias: A review and synthesis of the literature. *Advances in Behaviour Research and Therapy, 14,* 29–66.

Davey, G. C. L., Forster, L., & Mayhew, G. (1993). Familial resemblances in disgust sensitivity and animal phobias. *Behaviour Research and Therapy, 31,* 41–50.

Dillon, K. M., Wolf, E., & Katz, H. (1988). Sex roles, gender, and fear. *The Journal of Psychology, 119,* 355–359.

Edelbrock, C. S., & Costello, A. (1984). Structured psychiatric interviews for children and adolescents. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment,* (pp. 276–290). New York: Pergamon.

Eisen, A. R., & Silverman, W. K. (1991). Treatment of an adolescent with bowel movement phobia using self-control therapy. *Journal of Behavior Therapy and Experimental Psychiatry, 22,* 45–51.

Eme, R., & Schmidt, D. (1978). The stability of children's fears. *Child Development, 49,* 1277–1279.

Endicott, J., & Spitzer, R. L. (1978). A diagnostic interview: The schedule for affective disorders and schizophrenia. *Archives of General Psychiatry, 35,* 837–844.

Eysenck, H. J. (1979). The conditioning model of neurosis. *Behavioral and Brain Sciences, 2,* 155–199.

Fergusson, D. M., Horwood, L. J., & Lynskey, M. T. (1993). Prevalence and comorbidity of DSM-III-R diagnoses in a birth cohort of 15-year-old. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 1127–1134.

Foley, D., & Hay, D. A. (1992). Genetics and the nature of anxiety disorders. In G. F. Burrows, M. Roth, & R. Noyes, Jr. (Eds.), *Handbook of anxiety: Contemporary issues and prospects for research in anxiety disorders* (pp. 21–56). Amsterdam: Elsevier.

Fonesca, A. C., Yule, W., & Erol, N. (1994). Cross-cultural issues. In T. H. Ollendick, N. J. King, & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in children and adolescents* (pp. 67–84). New York: Plenum.

Fox, J. E., & Houston, B. K. (1981). Efficacy of self-instructional training for reducing children's anxiety in an evaluative situation. *Behaviour Research and Therapy, 19,* 509–515.

Francis, G. (1988). Assessing cognitions in anxious children. *Behavior Modification, 12,* 267–280.

Freud, S. (1909/1955). Analysis of a phobia in a five-year-old boy. In J. Strachey (Ed. and Trans.), *Standard edition of the complete psychological works of Sigmund Freud* (Vol. 10, pp. 3–149). London: Hogarth Press.

Fyer, A. J., Mannuzza, T. F., Chapman, L. M., Martin, L. Y., & Klein, D. F. (1995). Specificity in familial aggregation of phobic disorders. *Archives of General Psychiatry, 52,* 564–573.

Fyer, A. J., Mannuzza, S., Gallops, M. P., Martin, L. Y., Aaronson, C., Gorman, J. M., Liebowitz, M. R., & Klein, D. F. (1990). Familial transmission of simple phobias and fears. *Archives of General Psychiatry, 47,* 252–256.

Giebenhain, J. E., & O'Dell, S. L. (1984). Evaluation of a parent-training manual for reducing children's fear of the dark. *Journal of Applied Behavioral Analysis, 17,* 121–125.

Ginsburg, G. S., & Silverman, W. K. (1996). Phobic and anxiety disorders in Hispanic and Caucasian children. *Journal of Anxiety Disorders, 10,* 517–528.

Ginsburg, G. S., & Silverman, W. K. (1997). *Gender role orientation and fears in youth with anxiety disorders.* Unpublished manuscript.

Ginsburg, G. S., Silverman, W. K., & Kurtines, W. M. (1995a). Cognitive-behavioral group therapy. In A. R. Eisen, C. A. Kearney, & C. E. Schaefer (Eds.), *Clinical handbook of anxiety disorders in children* (pp. 521–549). Northwale, NJ: Jason Aronson.

Ginsburg, G. S., Silverman, W. K., & Kurtines, W. M. (1995b). Family involvement in treating children with anxiety and phobic disorders: A look ahead. *Clinical Psychology Review, 15,* 457–473.

Glasscock, S. E., & MacLean, W. E. (1990). Use of contact desensitization and shaping in the treatment of dog phobia and generalized fear of the outdoors. *Journal of Clinical Child Psychology, 19,* 169–172.

Glennon, B., & Weisz, J. R. (1978). An observational approach to the assessment of anxiety in young children. *Journal of Consulting and Clinical Psychology, 46,* 1246–1257.

Goldberg, J., & Weisenberg, M. (1992). The case of a newspaper phobia in a 9-year-old child. *Journal of Behavior Therapy and Experimental Psychiatry, 23,* 125–131.

Graziano, A. M., & DeGiovanni, I. S. (1979). The clinical significance of childhood phobias: A note on the proportion of child-clinical referrals for the treatment of children's fears. *Behaviour Research and Therapy, 17,* 161–162.

Graziano, A. M., & Mooney, K. C. (1980). Family self-control instruction for children's nighttime fear reduction. *Journal of Consulting and Clinical Psychology, 48,* 206–213.

Graziano, A. M., Mooney, K. C., Huber, C., & Ignasiak, D. (1979). Self-control instructions for children's fear-reduction. *Journal of Behavior Therapy and Experimental Psychiatry, 10,* 221–227.

Hagopian, L. P., Weist, M. D., & Ollendick, T. H. (1990). Cognitive-behavior therapy with an 11-year-old girl fearful of AIDS infection, other diseases, and poisoning: A case study. *Journal of Anxiety Disorders, 4,* 257–265.

Hamilton, D. I., & King, N. J. (1991). Reliability of a behavioral avoidance test for the assessment of dog phobic children. *Psychological Reports, 69,* 18.

Hammond-Laurence, K., & Silverman, W. K. (1997). *Comorbidity among childhood psychiatric disorders in clinic- referred children and adolescents with anxiety disorders.* Manuscript submitted for publication.

Heard, P. M., Dadds, M. R., & Conrad, P. (1992). Assessment and treatment of simple phobias in children: Effects on family and marital relationships. *Behaviour Change, 9,* 73–82.

Hekmat, H. (1987). Origins and development of human fear reactions. *Journal of Anxiety Disorders, 1,* 197–218.

Herjanic, B., & Reich, W. (1982). Development of a structured psychiatric interview for children: Agreement between child and parent on individual symptoms. *Journal of Abnormal Child Psychiatry, 10,* 307–324.

Hirshfeld, D. R., Rosenbaum, J. F., Biederman, J., Bolduc, E. A., Faraone, S. V., Snidman, N., Reznick, J. S., & Kagan, J. (1992). Stable behavioral inhibition and its association with anxiety disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 103–111.

Holmes, F. B. (1936). An experimental investigation of a method of overcoming children's fears. *Child Development, 7,* 6–30.

Jackson, H. J., & King, N. J. (1981). The emotive imagery treatment of a child's trauma-induced phobia. *Journal of Behavior Therapy and Experimental Psychiatry, 12,* 325–328.

Jones, M. C. (1924a). The elimination of children's fears. *Journal of Experimental Psychology, 1,* 383–390.

Jones, M. C. (1924b). A laboratory study of fear: The case of Peter. *Pedagogical Seminar, 31,* 308–315.

Kagan, J., Reznick, J. S., & Gibbons, J. (1989). Inhibited and uninhibited types of children. *Child Development, 60,* 838–845.

Kagan, J., Reznick, J. S., & Snidman, N. (1987). The physiology and psychology of behavioral inhibition. *Child Development, 58,* 1459–1473.

Kagan, J., Reznick, J. S., & Snidman, N. (1988). Biological bases of childhood shyness. *Science, 240,* 167–171.

Kanfer, F. H., Karoly, P., & Newman, A. (1975). Reduction of children's fear of the dark by confidence-related and situation threat-related verbal cues. *Journal of Consulting and Clinical Psychology, 43,* 251–258.

Kashani, J. H., Beck, N. C., Hoeper, E. W., Fallahi, C., Corcoran, C. M., McAllister, J. A., Rosenberg, T. K., & Reid, J. C. (1987). Psychiatric disorders in a community sample of adolescents. *American Journal of Psychiatry, 144,* 584–589.

Kashani, J. H., & Orvaschel, H. (1990). A community study of anxiety in children and adolescents. *American Journal of Psychiatry, 147,* 313–318.

Kearney, C. A., & Silverman, W. K. (1997). The evolution and reconciliation of taxonomic strategies for school refusal behavior. *Clinical Psychology: Science and Practice, 3,* 339–354.

Kendall, P. C., & Chansky, T. E. (1991). Considering cognition in anxiety-disordered children. *Journal of Anxiety Disorders, 5,* 167–185.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992). The genetic epidemiology of phobias in women. *Archives of General Psychiatry, 49,* 273–281.

King, N. J., Cranstoun, F., & Josephs, A. (1989). Emotive imagery and children's night-time fears: A multiple baseline design evaluation. *Journal of Behavior Therapy and Experimental Psychiatry, 20,* 125–135.

King, N. J., Hamilton, D. I., & Ollendick, T. H. (1988). *Children's phobias: A behavioral perspective.* New York: Wiley.

King, N. J., Mietz, A., Tinney, L., & Ollendick, T. H. (1995). Psychopathology and cognition in adolescents experiencing severe test anxiety. *Journal of Clinical Child Psychology, 24,* 49–54.

Kornhaber, R. C., & Schroeder, H. E. (1975). Importance of model similarity on extinction of avoidance behavior in children. *Journal of Consulting and Clinical Psychology, 43,* 601–607.

Krochak, M., Slovin, M., Rubin, J. G., & Kaplan, A. (1988). Treatment of dental phobia: A report of two cases. *Phobia Practice and Research Journal, 1,* 64–72.

Lang, P. J. (1977). Imagery in therapy: An information processing analysis of fear. *Behavior Therapy, 8,* 862–886.

Lapouse, R., & Monk, M. A. (1959). Fear and worries in a representative sample of children. *American Journal of Orthopsychiatry, 29,* 803–815.

Last, G. C. (1991). Somatic complaints in anxiety disordered children. *Journal of Anxiety Disorders, 5,* 125–138.

Last, C. G., Francis, G., & Strauss, C. C. (1989). Assessing fears in anxiety-disordered children with the Revised

Fear Survey Schedule for Children (FSSC-R). *Journal of Clinical Child Psychology, 18,* 137–141.

Last, C. G., Hersen, M., Kazdin, A., Orvaschel, H., & Perrin, S. (1991). Anxiety disorders and their families. *Archives of General Psychiatry, 48,* 928–934.

Last, C. G., & Perrin, S. (1993). Anxiety disorders in African-American and White children. *Journal of Abnormal Child Psychology, 2,* 153–164.

Last, C. G., Perrin, S., Hersen, M., & Kazdin, A. E. (1992). DSM-III-R anxiety disorders in children: Sociodemographic and clinical characteristics. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 1070–1076.

Last, C. G., Strauss, C. C., & Francis, G. (1987). Comorbidity among childhood anxiety disorders. *Journal of Nervous and Mental Disease, 175,* 726–730.

Lewis, S. (1974). A comparison of behavior therapy techniques in the reduction of fearful avoidant behavior. *Behavior Therapy, 5,* 648–655.

Liddell, A., & Lyons, M. (1978). Thunderstorm phobias. *Behaviour Research and Therapy, 16,* 306–308.

Maccoby, E. E. (1980). *Social development: Psychological growth and the parent-child relationship.* New York: Harcourt Brace Jovanovich.

Marks, I. M., & Gelder, M. G. (1966). Different onset ages in varieties of phobia. *American Journal of Psychiatry, 123,* 218–221.

Martin, M., Horder, P., & Jones, G. (1992). Integral bias in naming of phobia-related words. *Cognition and Emotion, 6,* 479–486.

McGee, R., Feehan, M., Williams, S., & Anderson, J. (1992). DSM-III disorders from age 11 to age 15 years. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 50–59.

McGee, R., Feehan, M., Williams, S., Partridge, F., Silva, P. A. & Kelly, J. (1990). DSM-III disorders in a large sample of adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 611–619.

Melamed, B. G., & Lumley, M. A. (1988). Dental subscale of the Children's Fear Survey Schedule. In M. Hersen & A. S. Bellack (Eds.), *Dictionary of behavioural assessment techniques* (p. 171). Oxford: Pergamon.

Melamed, B. G., & Siegel, L. J. (1975). Reduction of anxiety in children facing hospitalization and surgery by use of filmed modeling. *Journal of Consulting and Clinical Psychology, 43,* 511–521.

Menzies, R. G., & Clarke, J. C. (1993). A comparison of in vivo and vicarious exposure in the treatment of childhood water phobia. *Behaviour Research and Therapy, 31,* 9–15.

Menzies, R. G., & Clarke, J. C. (1995). The etiology of phobias: A nonassociative account. *Clinical Psychology Review, 15,* 23–48.

Merckelbach, H., De Ruiter, C., Van De Hout, M. A., & Hoekstra, R. (1989). Conditioning experiences and phobias. *Behaviour Research and Therapy, 27,* 657–662.

Milgrom, P., Mancle, L., King, B., & Weinstein, P. (1995). Origins of childhood dental fear. *Behavior Research Therapy, 33,* 313–319.

Miller, L. C., Barrett, C. L., Hampe, E., & Noble, H. (1972). Comparison of reciprocal inhibition, psychotherapy, and waiting list control for phobic children. *Journal of Abnormal Psychology, 79,* 269–279.

Milne, J. M., Garrison, C. Z., Addy, C. L., McKeown, R. E., Jackson, K. L., Cuffe, S. P., & Waller, J. L. (1995). Frequency of phobic disorder in a community sample of young adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 1202–1211.

Mooney, K. C. (1985). Children's nighttime fears: Ratings of content and coping behaviors. *Cognitive Therapy and Research, 9,* 309–319.

Moran, C., & Andrews, G. (1985). The familial occurrence of agoraphobia. *British Journal of Psychiatry, 146,* 262–267.

Morris, R. J., & Kratochwill, T. R. (1983). *Treating children's fears and phobias: A behavioral approach.* New York: Pergamon.

Mowrer, O. H. (1939). A stimulus-response theory of anxiety and its role as reinforcing agent. *Psychological Review, 46,* 553–565.

Muris, P., Steerneman, P., Merckelbach, H., & Meesters, C. (1996). The role of parental fearfulness and modeling in children's fear. *Behaviour Research and Therapy, 34,* 265–268.

Murphy, J. K., Alpert, B. S., Willey, E. S., & Somes, G. S. (1988). Cardiovascular reactivity to psychological stress in healthy children. *Psychophysiology, 25,* 144–152.

Murray, E. J., & Foote, F. (1979). The origins of fear of snakes. *Behaviour Research and Therapy, 17,* 489–493.

Nalven, F. B. (1970). Manifest fears and worries of ghetto versus middle class suburban children. *Psychological Reports, 27,* 285–286.

Neal, A. M., & Turner, S. M. (1991). Anxiety disorders research with African-Americans: Current status. *Psychological Bulletin, 109,* 400–410.

Noyes, R., Crowe, R. R., Harris, E. L., Hamra, B. J., McChesney, C. M., & Chaudhry, D. R. (1986). Relationship between panic disorder and agoraphobia. *Archives of General Psychiatry, 43,* 227–232.

Obler, M., & Terwilliger, R. F. (1970). Pilot study on the effectiveness of systematic desensitization with neurologically impaired children with phobic disorders. *Journal of Consulting and Clinical Psychology, 2,* 314–318.

Ollendick, T. H. (1979). Fear reduction techniques with children. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification* (Vol. 8, pp. 127–168). New York: Academic Press.

Ollendick, T. H. (1983). Reliability and validity of the revised Fear Survey Schedule for Children (FSSC-R). *Behaviour Research and Therapy, 21,* 395–399.

Ollendick, T. H., & Francis, G. (1988). Behavioral assessment and treatment of childhood phobias. *Behavior Modification, 12,* 165–204.

Ollendick, T. H., & King, N. J. (1994). Diagnosis, assessment, and treatment of internalizing problems in children: The role of longitudinal data. *Journal of Consulting and Clinical Psychology, 62,* 918–927.

Ollendick, T. H., King, N. J., & Frary, R. B. (1989). Fears in children and adolescents: Reliability and generalizability across gender, age and nationality. *Behaviour Research and Therapy, 27,* 19–26.

Ollendick, T. H., King, N. J., & Hamilton, D. I. (1991). Origins of childhood fears: An evaluation of Rachman's theory of fear acquisition. *Behaviour Research and Therapy, 29,* 117–123.

Ollendick, T. H., Yang, B., Dong, Q., Xia, Y., & Lin, L. (1995). Perceptions of fear in other children and adolescents: The role of gender and friendship status. *Journal of Abnormal Child Psychology, 23,* 439–452.

Ollendick, T. H., Yang, B., King, N. J., Dong, Q., & Akande, A. (1996). Fears in American, Australian, Chinese, and Nigerian children and adolescents: A cross-cultural study. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 37,* 213–220.

Ollendick, T. H., Yule, W., & Ollier, K. (1991). Fears in British children and their relationship to manifest anxiety and depression. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 32,* 321–331.

Ost, L. G. (1985). Ways of acquiring phobias and outcome of behavioral treatments. *Behaviour Research and Therapy, 23,* 683–689.

Ost, L. G. (1987). Age of onset in different phobias. *Journal of Abnormal Psychology, 96,* 123–145.

Ost, L. G., & Hugdahl, K. (1983). Acquisition of agoraphobia, mode of onset and anxiety response patterns. *Behaviour Research and Therapy, 21,* 623–631.

Page, A. C. (1994). Blood-injury phobia. *Clinical Psychology Review, 14,* 443–461.

Perrin, S., & Last, C. G. (1992). Do childhood anxiety measures measure anxiety? *Journal of Abnormal Child Psychology, 20,* 567–578.

Peterson, L., & Shigetomi, C. (1981). The use of coping techniques to minimize anxiety in hospitalized children. *Behavior Therapy, 12,* 1–14.

Prins, P. J. M. (1985). Self-speech and self regulation of high and low anxious children in the dental situation: An interview study. *Behaviour Research and Therapy, 23,* 641–650.

Prins, P. J. M. (1986). Children's self-speech and self-regulation during a fear-provoking behavioral test. *Behaviour Research and Therapy, 24,* 181–191.

Puig-Antich, J., & Chambers, W. (1978). *The Schedule for Affective Disorders and Schizophrenia for School-Aged Children*. New York: New York State Psychiatric Institute.

Rachman, S. (1977). The conditioning theory of fear acquisition: A critical examination. *Behaviour Research and Therapy, 15,* 375–387.

Rachman, S. (1983). Irrational thinking, with special reference to cognitive therapy. *Advanced Behaviour Research and Therapy, 5,* 63–68.

Rainwater, N., Sweet, A. A., Elliott, L., Bowers, M., McNeil, J., & Stump, N. (1988). Systematic desensitization in the treatment of needle phobias for children with diabetes. *Child and Family Behavior Therapy, 10,* 19–31.

Rapee, R. M., Barret, P. M., Dadds, M. R., & Evans, L. (1994). Reliability of the *DSM-III-R* childhood anxiety disorders using structured interview: Interrater and parent-child agreement. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 984–992.

Reich, J., & Yates, W. (1988). Family history of psychiatric disorders in social phobia. *Comprehensive Psychiatry, 29,* 72–75.

Reiss, S. (1991). Expectancy model of fear, anxiety and panic. *Clinical Psychology Review, 11,* 141–154.

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist, 43,* 151–160.

Richardson, S. A., Dohrenwend, B. S., & Klein, D. (1965). *Interviewing: Its forms and functions*. New York: Basic Books.

Rimm, D. C., Janda, H. L., Lancaster, D. W., Nahl, M., & Dittmar, K. (1977). An exploratory investigation of the origin and maintenance of phobias. *Behaviour Research and Therapy, 15,* 231–238.

Ritter, B. (1968). The group desensitization of children's snake phobias using vicarious and contact desensitization procedures. *Behaviour Research and Therapy, 6,* 1–6.

Ryall, M. R., & Dietiker, K. E. (1979). Reliability and clinical validity of the Children's Fear Survey Schedule. *Journal of Behavior Therapy and Experimental Psychiatry, 19,* 303–310.

Sarason, S. B., Davidson, K. S., Lighthall, F. F., Waite, R. R., & Ruebush, B. K. (1960). *Anxiety in elementary school children*. New York: Wiley.

Seligman, M. E. P. (1970). On the generality of the laws of learning. *Psychological Review, 77,* 406–418.

Seligman, M. E. P. (1971). Phobias and preparedness. *Behavior Therapy, 2,* 307–320.

Shaw, J. (1990). Menstruation phobia treated by cognitive correction: A case report. *Journal of Behavior Therapy and Experimental Psychiatry, 21,* 49–51.

Sheehan, D. V., Sheehan, K. E., & Minichiello, W. E. (1981). Age of onset of phobic disorders: A reevaluation. *Comprehensive Psychiatry, 22,* 544–553.

Shinebourne, E. A. (1974). Growth and development of the cardiovascular system. In J. A. Davis & J. Dobbing (Eds.), *Scientific foundations of pediatrics* (pp. 198–213). London: Heinemann.

Siegel, L. J., & Peterson, L. (1980). Stress reduction in young dental patients through coping skills and sensory information. *Journal of Consulting and Clinical Psychology, 48,* 785–787.

Silverman, W. K. (1991). Diagnostic reliability of anxiety disorders in children using structured interviews. *Journal of Anxiety Disorders, 5,* 105–124.

Silverman, W. K. (1994). Structured diagnostic interviews. In T. H. Ollendick, N. J. King, & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in children and adolescents* (pp. 293–315). New York: Plenum.

Silverman, W. K. & Albano, A. M. (1996). *The Anxiety Disorders Interview Schedule for Children–IV (Child and Parent Versions)*. San Antonio, TX: Psychological Corporation.

Silverman, W. K., Cerny, J. A., & Nelles, W. B. (1988). The familial influence in anxiety disorders: Studies on the offspring of parents with anxiety disorders. In B. B. Lahey & A. E. Kazdin (Eds.), *Advances in clinical child psychology* (Vol. II, pp. 223–248). New York: Plenum.

Silverman, W. K., Cerny, J. A., Nelles, W. B., & Burke, A. E. (1988). Behavior problems in children of parents with anxiety disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 779–784.

Silverman, W. K., & Eisen, A. R. (1992). Age differences in the reliability of parent and child reports of child anxious symptomatology using a structured interview. *Journal of American Academy of Child and Adolescent Psychiatry, 31,* 117–124.

Silverman, W. K., & Eisen, A. R. (1993). Phobic disorders. In R. T. Ammerman, C. G. Last, & M. Hersen (Eds.), *Handbook of prescriptive treatments for children and adolescents* (pp. 17–197). Boston: Allyn & Bacon.

Silverman, W. K., & Ginsburg, G. S. (1998). Anxiety disorders. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of child psychopathology* (3rd ed.), New York: Plenum.

Silverman, W. K., & Kearney, C. A. (1992). Listening to our clinical partners: Informing researchers about children's fears and phobias. *Journal of Behavior Therapy and Experimental Psychiatry, 23,* 71–76.

Silverman, W. K., & Kurtines, W. K. (1996). *Anxiety and phobic disorders: A pragmatic approach*. New York: Plenum.

Silverman, W. K., & Nelles, W. B. (1988a). The Anxiety Disorders Interview Schedule for Children. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 772–778.

Silverman, W. K., & Nelles, W. B. (1988b). The influence of gender on children's ratings of fear in self and same-aged peers. *Journal of Genetic Psychology, 149,* 17–22.

Silverman, W. K., & Nelles, W. B. (1990). Simple phobia in childhood. In M. Hersen & C. G. Last (Eds.), *Handbook of child and adult psychopathology: A longitudinal perspective* (pp. 183–196). New York: Pergamon.

Silverman, W. K., & Rabian, B. (1993). Simple phobias. *Child and Adolescent Psychiatric Clinics of North America, 2,* 603–622.

Silverman, W. K., & Rabian, B. (1995). Test-retest reliability of the DSM-III-R anxiety childhood disorders symptoms using the Anxiety Disorders Interview Schedule for Children. *Journal of Anxiety Disorders, 9,* 1–12.

Slater, E., & Shields, J. (1969). Genetical aspects of anxiety. In M. J. Lader (Ed.), *Studies of anxiety* (pp. 62–71). London: Royal Medico-Psychological Association.

Solomon, R. L., & Wynne, L. C. (1954). Traumatic avoidance learning: The principles of anxiety conservation and partial irreversibility. *Psychological Review, 61,* 358–385.

Solyom, L., Beck, P., Solyom, C., & Hugel, R. (1974). Some etiological factors in phobic neurosis. *Canadian Psychiatric Association Journal, 19,* 69–78.

Sreenivasan, V., Manocha, S. N., & Jain, V. K. (1979). Treatment of severe dog phobia in childhood by flooding: A case report. *Journal of Child Psychology and Psychiatry, 20,* 255–260.

Stampfl, T. G. (1991). Analysis of aversive events in human psychopathology: Fear and avoidance. In M. R. Denny (Ed.), *Fear, avoidance, and phobias: A fundamental analysis.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Stedman, J. M., & Murphy, J. (1984). Dealing with specific child phobias during the course of family therapy: An alternative to systematic desensitization. *Family Therapy, 11,* 55–60.

Stevenson, J., Batten, N., & Cherner, M. (1992). Fears and fearfulness in children and adolescents: A genetic analysis of twin data. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 33,* 977–985.

Strauss, C. C., & Last, C. G. (1993). Social and simple phobias in children. *Journal of Anxiety Disorders, 2,* 141–152.

Tasto, D. L. (1969). Systematic desensitization, muscle relaxation and visual imagery in the counterconditioning of a four-year-old phobic child. *Behaviour Research & Therapy, 7,* 409–411.

Thankachan, M. V., Mishra, H., & Kumaraiah, V. (1992). Behavioral intervention with phobic children. *Nimhans Journal, 10,* 95–99.

Thyer, B. A., Parrish, R. T., Curtis, G. C., Nesse, R. M., & Cameron, O. G. (1985). Ages of onset of DSM-III anxiety disorders. *Comprehensive Psychiatry, 26,* 113–122.

Torgersen, S. (1979). The nature and origin of common phobic fears. *British Journal of Psychiatry, 134,* 343–351.

Turner, S. M., Beidel, D. C., & Costello, A. (1987). Psychopathology in the offspring of anxiety disorders patients. *Journal of Consulting and Clinical Psychology, 55,* 229–235.

Turner, S. M., Beidel, D. C., & Wolff, P. L. (1996). Is behavioral inhibition related to the anxiety disorders? *Clinical Psychology Review, 16,* 157–172.

Ultee, C. A., Griffioen, D., & Schellekens, J. (1982). The reduction of anxiety in children: A comparison of the effects of systematic desensitization *in vitro* and systematic desensitization *in vivo*. *Behaviour Research and Therapy, 20,* 61–67.

Van Hasselt, V. B., Hersen, M., Bellack, A. S., Rosenblum, N. D. & Lamparski, D. (1979). Tripartite assessment of the effects of systematic desensitization in a multi-phobic child: An experimental analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 10,* 51–55.

Vasey, M. W., Daleiden, E. L., Williams, L. L., & Brown, L. (1995). Biased attention in childhood anxiety disorders: A preliminary study. *Journal of Abnormal Child Psychology, 23,* 267–279.

Venables, P. H. (1981). Automatic reactivity. In M. Rutter (Ed.), *Scientific foundations of developmental psychiatry* (pp. 165–175) Baltimore: University Park Press.

Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3,* 1–14.

Weissman, M. M., Leckman, J. F., Gammon, G. D., Merikangas, K. R., & Prusoff, B. A. (1984). Depression and anxiety disorders in parents and children. *Archives of General Psychiatry, 41,* 845–852.

Wish, P. A., Hasazi, J. E., & Jurgela, A. R. (1973). Automated direct deconditioning of a childhood phobia. *Journal of Behavior Therapy and Experimental Psychiatry, 4,* 279–283.

Wolpe, J., & Rachman, S. (1960). Psychoanalytic evidence: A critique based on Freud's case of Little Hans. *Journal of Nervous and Mental Disease, 131,* 135–147.

World Health Organization (1992). *International classification of mental and behavioral disorders, clinical descriptions and diagnostic guidelines* (10th ed.). Geneva: WHO.

Young, J. P. R., Fenton, G. W., & Lader, M. J. (1971). The inheritance of neurotic traits: A twin study of the Middlesex Hospital Questionnaire. *British Journal of Psychiatry, 119,* 393–398.

Yule, W., Sacks, B., & Hersov, L. (1974). Successful flooding treatment of noise phobia in an eleven-year-old boy. *Journal of Behavior Therapy and Experimental Psychiatry, 5,* 209–211.

Zabin, M. A., & Melamed, B. G. (1980). Relationship between parental discipline and children's ability to cope with stress. *Journal of Behavioral Assessment, 2,* 17–38.

Zatz, S., & Chassin, L. (1983). Cognitions of test-anxious children. *Journal of Consulting and Clinical Psychology, 51,* 524–534.

Zatz, S., & Chassin, L. (1985). Cognitions of test-anxious children under naturalistic test-taking conditions. *Journal of Consulting and Clinical Psychology, 53,* 393–401.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.15
Obsessive-compulsive Disorder

DEREK BOLTON
Institute of Psychiatry, London, UK

5.15.1 INTRODUCTION 367

5.15.2 OBSESSIVE-COMPULSIVE DISORDER 368
  5.15.2.1 Phenomenology 368
  5.15.2.2 Diagnostic Features 369
  5.15.2.3 Differential Diagnoses 370
  5.15.2.4 Prevalence 371
  5.15.2.5 Etiology 371
  5.15.2.6 Developmental Course 377

5.15.3 CONCEPTUALIZATION AND CLINICAL FORMULATION 378
  5.15.3.1 Multimodal Assessment 378
  5.15.3.2 Treatment Foundation 379

5.15.4 TREATMENT 380
  5.15.4.1 Psychosocial Treatments 380
  5.15.4.2 Pharmacological Treatments 381
  5.15.4.3 Effectiveness of Treatments 381

5.15.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 383

5.15.6 SUMMARY 384

5.15.7 REFERENCES 384

5.15.1 INTRODUCTION

The development of the current conception and classification of what we know as "obsessive-compulsive disorder" (OCD) has been full of controversy. While this fact in itself is not extraordinary in the area of psychopathology, it is remarkable that many of the historical controversies are still alive and some are even kicking. The disorder has proved particularly difficult to define and classify. This has to do with the fact that it involves, or is associated with, various functions and dysfunctions of mind and behavior. It evidently involves the intellect, with thoughts or doubts at or beyond the border of sanity; it frequently involves the emotions, particularly fear and melancholia; it typically involves an apparent impairment of the power of the will, with distressing thoughts and actions persisting against the person's wishes; the behavioral aspects of the disorder are in some cases akin to complex tics, or even to motor movements induced by brain seizure. The development of the current concept of OCD reflects all the various possibilities which this complex phenomenology opens up. Many early psychopathologists tackled the problem, including, for example, Esquirol, Morel, Dagonet, du Saulle, and Ball, in France, Krafft-Ebbing, Westphal, and Wille, in Germany, and Mickle, Tuke, and Jackson, in Britain (Berrios, 1995). During the nineteenth century particularly, what we now know as OCD was alternatively conceived as a disorder of intellect, volition, or emotions, and classified variously along with psychoses, neurologically based motor disorders, and anxiety disorders. The conception of OCD as primarily a disorder of the emotions, and its classification along with typical anxiety disorders, were consolidated during the early twentieth century and were embraced by the major nosological systems, where they remain today enshrined in the most recent editions, *DSM-IV* (American Psychiatric Association, 1994) and *ICD-10* (World Health Organization, 1988). The conception of OCD as an anxiety disorder was consistent with and reinforced by psychoanalytic theory, and it is equally well-suited to behavioral and cognitive behavioral theorizing.

However, major controversies have appeared, or more accurately reappeared, concerning the nature and appropriate classification of OCD. There has been a resurgence of interest in the notion of a spectrum of OCDs within which typical OCD is one kind of case, but which includes conditions other than anxiety disorders. Further, there is a view which has assumed something of the status of an orthodoxy, particularly in the United States, to the effect that the primary dysfunction in OCD is more to do with motor control than with emotional conflict, and that both motor and psychological aspects of the disorder probably cannot be understood in terms of psychological principles, whether psychoanalytic or cognitive-behavioral, since they are due to an hypothesized structural or biochemical neural lesion.

Notwithstanding theoretical uncertainty, progress on the treatment front has been substantial. OCD had been found to be refractory to psychiatric treatments, though some cases showed reasonably good response to psychosurgery, and also resistant to psychoanalysis (Rachman & Hodgson, 1980). The first studies of the course of the disorder showed reasonably good recovery in as little as a third of cases. The position was transformed, however, in the late 1970s and early 1980s, with the development of two new forms of treatment: behavior therapy—specifically exposure with response–prevention—and pharmacotherapy using agents which specifically inhibited uptake of the neurotransmitter serotonin. Trials of these two new forms of treatment, separately or in combination, found good outcome in up to two-thirds of cases. There is thus reason now to be fairly optimistic about prognosis. There is, on the other hand, ample scope for more improvement in a number of areas: in cases where the person has insufficient motivation for behavioral treatment, where there is poor response to treatment, and where there is subsequent serious relapse. In other words the current forms of treatment constantly come up against their limitations, and this requires not only continuing attention to therapeutic details and refinements, but also continuing work on the complex and controversial models of the etiology of the disorder.

5.15.2 OBSESSIVE-COMPULSIVE DISORDER

5.15.2.1 Phenomenology

As a young child, Christine used to worry occasionally about germs, afraid that she would infect other people or vice versa. At age 12 following the death of a family friend Christine was very upset and began washing her hands excessively. A few months later she became preoccupied with thoughts about lines which had to be unwoven by retracing her steps, and this inhibited her from going out. Age 13, Christine's paternal grandmother died of cancer, and she began menstruation during the same year. Over this time Christine became increasingly preoccupied with intrusive, distressing thoughts which usually had violent or sexual content. If a sexual or violent thought occurred in her mind, she would feel compelled to return to the place she had just been, or stage in a behavior sequence, and then move forward again without having the thought, which procedure she felt neutralized having the thought in the first place. In practice, this meant that she would take up to several hours to walk a short distance or get dressed. Christine would spend much time telling her mother of her thoughts, or being helped by her mother to restart activity when she got stuck.

John was referred at two and a half years to a child guidance clinic because of parental concern over his failure to socialize with other children, speech delay, and concern about possible mental retardation. He was clingy with his mother and was never able to dress himself fully. Later at school John was reported to be polite, cooperative, and of slightly above average intelligence. However, he was said to be "painfully slow" and a backward reader with barely legible writing. John was socially isolated, teased, and harassed by his peers. Following entry into secondary school John's difficulties worsened, and at age 11 he was admitted for several weeks to an inpatient psychiatric unit. He presented as a prepubertal boy with faint and hesitant speech. He was severely slow in all his personal functions, requiring the nurses to dress, wash, bathe, and help him eat. If left on his own, he took six to eight hours to dress and often stood motionless for long periods of time being distracted by something irrelevant to the task. He was frequently incontinent due to his slowness. He walked with a skipping gait, and had multiple tics and mannerisms such as squinting his eyes, tapping his hands, touching his right cheek with his left hand, licking his lips, and producing repetitive writhing neck movements, and jerking action of his upper limbs. When leaving a room John would repeatedly tap the door to ensure it was closed, he checked his knife and fork for dirt, and turned his head before starting to walk. John was often preoccupied with the correct manner of performing a task, being fearful he would make a mistake and thus feel compelled to return to the beginning of the sequence. He said he did not resist his habits and felt more anxious when prevented from performing them. He used "magical numbers," often having to tap his hands and feet in sixes and sevens.

These two short descriptions of actual cases illustrate some of the variety of phenomenology in children and adolescents suffering from OCD. There are common features, captured in general terms by the standard diagnostic systems, to be discussed below, but within these broad guidelines there are quite wide variations. With children and adolescents, as for adults with OCD, cleaning compulsions, linked to obsessions about contamination, and checking compulsions, linked to fear of more or less well-specified disasters, are most common. Other common compulsive activities include putting things in a set order, pattern or sequence, either physical or mental, touching, and hoarding (Foa & Kozak, 1995; Riddle et al., 1992; Swedo, Rapoport, Leonard, Lenane, & Cheslow, 1989; Thomsen & Mikkelsen, 1991).

Broadly speaking the phenomenology of OCD is fairly constant through the age range from childhood to adulthood, and this is reflected in the fact that the diagnostic criteria apply across the age range. While the form of the disorder remains essentially the same, some variations in content can be seen. These are generally understandable in biological or psychosocial terms. For example, children and adolescents commonly present with compulsive checking of homework, or with persistent, distressing requests for parents to reassure them that they are safe, and while sexual obsessions are common in adults (Foa & Kozak, 1995; Stern & Cobb, 1978), they are uncommon in children.

As to cultural factors, there is a reasonable amount of data on cultural factors affecting normal childhood fears, but there is little in relation to pathological fear, including obsessional anxiety (Fonseca, Yule, & Erol, 1994). Research on cultural factors affecting OCD has been on adults, and generally the findings are that the disorder presents in similar ways in various cultures, though it remains possible that the content of obsessions and the nature of compulsions may be sensitive in some degree to cultural norms and preoccupations. Most attention has focused on the sometimes very striking apparent similarities between OCD and religious beliefs and rituals, but this alongside the fact that within religions it is explicitly acknowledged that observance to the law can go, as it were, over the top, being no longer piety, but eccentricity or sickness, and in any case usually something to be avoided. This kind of picture has been noted in many countries and religions including Judaism (Greenberg, 1984; Suess & Halpern, 1989), Islam (Mahgoub & Abdel-Hafeiz, 1991; Pfeiffer, 1981), Catholicism (Suess & Halpern; 1989), and Hinduism (Chakraborty & Banerji, 1975; Sharma, 1968). There is, on the other hand, so far no evidence that the extent of religious belief or ritual in a culture affects the risk for developing OCD, at least in the sense that studies in a fairly wide variety of cultures converge on similar figures for prevalence of the disorder (see Section 5.15.2.4).

5.15.2.2 Diagnostic Features

The diagnostic criteria for OCD as specified by *DSM-IV* (American Psychiatric Association, 1994) may be summarized as follows. The person must have either or both of *obsessions* or *compulsions*. "Obsessions" are defined as recurrent, intrusive, distressing mental events, which are not just excessive real-life worries, which the person tries to ignore or suppress or neutralize, and which the person regards as the product of their own mind. "Compulsions" are defined as repetitive behaviors or mental acts carried out in response to obsessions, or according to rigid rules, aimed at reducing distress or averting some dreaded event in the absence of a realistic connection between the activities and the intended consequences. *DSM-IV* specifies that at some point these must have been recognized as excessive or unreasonable, but adds that this criterion need not apply in the case of children. As usual in *DSM-IV*, there is an additional criterion that the symptomatology causes marked distress, is time-consuming, or causes impaired functioning. There is a standard rule specifying that the diagnosis should be made only if the disturbance is not due to the direct physiological effects of a substance or a general medical condition. There is also a further rule excluding the diagnosis of OCD if there is another Axis I disorder present and if the content of obsessions or compulsions is restricted to it, for example, preoccupation with food in the presence of an eating disorder; hair-pulling in the presence of trichotillomania;

concern with appearance in the presence of body dysmorphic disorder; preoccupation with having a serious illness in the presence of hypochondriasis; preoccupation with sexual urges in the presence of paraphilia; or guilty ruminations in the presence of major depressive disorder.

This type of exclusion rule is legitimate, and indeed inevitable, as a means of defining precisely the boundaries of a particular diagnostic category, but of course they might be artificial. Such rules cannot preempt the question whether OCD as so defined is a distinct biopsychological entity compared with the other disorders to which it is apparently related. This bears on the question of "obsessive-compulsive spectrum disorders" to be discussed in Section 5.15.2.5.

ICD-10 uses similar diagnostic criteria (World Health Organiztion, 1988, p. 143). It specifies minimum duration of two weeks, and it defines subtypes of predominantly obsessions, predominantly compulsions, mixed obsessions and compulsive acts, other OCDs, and unspecified.

Foa and her colleagues conducted a field trial involving 431 patients diagnosed as having OCD in seven sites with the aim of examining some empirical and conceptual assumptions of the *DSM-III-R* criteria for OCD and making recommendations for *DSM-IV* and comparisons with *ICD-10* (Foa & Kozak, 1995; with author list correction in Foa et al., 1995). Three issues were examined: (i) the requirement that symptoms of OCD be viewed by the patient as excessive or unreasonable, (ii) the presence of mental compulsions in addition to behavioral compulsions, and (iii) validity of *ICD-10* subcategories. Primary measures were the Yale–Brown Obsessive-Compulsive Scale and face-valid questions about fixity of obsessive-compulsive beliefs. A wide range of views about the unreasonableness of obsessional fears was found among 250 participants: 13% were certain that feared consequences of not performing their compulsions would not occur; 27% were mostly certain; 30% were uncertain; 26% were mostly certain that the feared consequence would occur, and 4% were completely certain that it would occur. Eight percent were classified as currently lacking in insight, and 5% had never recognized the senselessness of their symptoms. The investigators noted that the 5% of patients who reported never having had insight, and perhaps also the 25% who were mostly certain that their symptoms were reasonable, would not formally meet *DSM-III-R* criteria for OCD, and inferred that the participating clinicians with extensive experience with OCD were compensating for imperfections in those criteria. It was subsequently proposed that the requirement for insight should be de-emphasized in *DSM-IV*, and a specifier "with poor insight" has been included. It was found that almost 80% of the patients reported having both mental and behavioral compulsions, and the *DSM-III-R* wording has been clarified in *DSM-IV* to make it clearer that compulsions can be mental and not just behavioral. The authors note that the identification of mental compulsive rituals is important for at least two reasons. First, it can eliminate confusion between OCD and other disorders involving ruminations, such as generalized anxiety disorder or depression; only individuals with OCD perform mental rituals. Second, behavior therapy uses very different techniques for obsessions on the one hand and mental rituals on the other; prolonged exposure is used for the former, and blocking techniques for the latter. Findings on the validity of the *ICD-10* subcategories of OCD were equivocal.

5.15.2.3 Differential Diagnoses

It would be fair to say that most of the time diagnosis of OCD is relatively straightforward, but there are some crucial and interesting kinds of cases in which differential diagnoses have to be considered, and the following are the most common and important.

(i) Ruminative brooding with depressed content may be part of a depressive syndrome rather than obsessions in the sense of OCD. Mood congruence so far suggests depressive disorder.

(ii) Anxious mental preoccupation may indicate generalized anxiety disorder (GAD) rather than OCD. Here a critical feature is that in GAD, worrying is seen as excessive but realistic, while in OCD mental preoccupation is unrealistic. However, the line may be difficult to draw, for example, over worry about parental health.

(iii) There is a fine line, at least in theory, between obsessions and delusions. Both are unreasonable, and may be bizarre, with the critical difference traditionally being that obsessions as opposed to delusions are recognized as such (Berrios, 1995). However, this condition is absent in many adult patients with OCD (Foa & Kozak, 1995; Stern & Cobb, 1978), and as already indicated *DSM-IV* has dropped it altogether for children. Differential diagnoses of OCD and psychoses specifically in adolescents is discussed by Thomsen (1992).

(iv) Stereotyped, repetitive behaviors of individuals with very low intelligence, tic disorders, or pervasive developmental disorders may resemble the compulsive behaviors characteristic of OCD. Some typical differences in the content of repetitive thoughts in autism and OCD have been reported in adults (McDougle et al., 1995), although content is not relevant to diagnosis. So far as concerns repetitive behaviors, the differences critical to diagnosis are that obsessive-compulsions are carried out in response to obsessions or according to strict rules, they relieve anxiety, and they generate distress.

One or more of these features is absent in nonobsessional, stereotyped, repetitive behaviors associated with autism (Baron-Cohen, 1989), mental retardation (Bodfish et al., 1995; Vitiello, Spreat, & Behar, 1989), and tics (Cath et al., 1992; Miguel et al., 1995a).

Decision as to appropriate diagnosis in these various kinds of cases requires careful investigation of phenomenology and the pattern of symptomatology as a whole. If two sets of diagnostic criteria are satisfied, then both diagnoses can be appropriately given, though this is subject to priority and exclusion rules in particular (versions of) diagnostic systems. Comorbidity is common in OCD, in children and adolescents as in adults. Other psychopathology has been reported in up to three-quarters of child and adolescent cases, with other anxiety disorders and affective disorders being the most common (Flament et al., 1988; Riddle et al., 1990; Swedo, Rapoport, Leonard, Lenane, & Cheslow 1989; Valleni-Basile et al., 1994).

5.15.2.4 Prevalence

Early estimates of the prevalence of childhood and adolescent OCD ranged between 0.2% and 1.2% of hospital clinic populations (Berman, 1942; Hollingsworth, Tanguay, Grossman, & Pabst, 1980; Judd, 1965) while the first community study found a 0.3% prevalence of "mixed obsessional/anxiety disorders" in 10 and 11 year-olds (Rutter, Tizard, & Whitmore, 1970). Community studies in the late 1980s and early 1990s estimated life-time prevalence of the disorder by adolescence at 1.9% (Flament et al., 1988) and 3% (Valleni-Basile et al., 1994; with correction in Garrison et al., 1995). Adult community studies have estimated 2–3% for life-time prevalence of the disorder in the United States (Karno, Golding, Sorenson, & Burnam, 1988; Robins et al., 1984), with comparable estimates in some other countries and cultures including Uganda (Orley & Wing, 1979) and Canada (Bland, Newman, & Orn, 1988), though with a lower estimate (approximately 1%) in Taiwan (Hwu, Yeh, & Chang, 1989).

As to age of onset, OCD can start very early in childhood, with cases of children below six and as young as three years described in the literature (Adams, 1973; Janet, 1903; Rettew, Swedo, Leonard, Lenane, & Rapoport, 1992). Childhood OCD typically onsets between five and eight years of age or during adolescence (Swedo, Leonard, & Rapoport, 1990), and the age of onset is below 25 years and mostly during adolescence in approximately two-thirds of adult cases (Kringlen, 1965; Pollitt, 1957; Rasmussen & Eisen, 1990). Rasmussen and Tsuang (1986) found two peaks of maximal incidence of the disorder, one at 12–14 years and the other at 20–22 years.

The fact that OCD can often have onset in young children is relevant in several ways to a developmental perspective on etiology. The apparent implication is that whatever maturation has to be reached for the disorder to appear, biological or psychological or both, has already been reached relatively early in childhood. At the same time OCD is apparently not a "developmental disorder": it does not obviously involve any specific developmental delay, and moreover, it can have late onset, in teenage or early adulthood, or beyond. It is also noteworthy that childhood obsessive-compulsive symptomatology appears at around the same time as normal childhood magical thinking and behavior (Gesell, Ames, & Ilg, 1974; Harris, 1994), and the two sets of phenomena—the one psychopathological and the other belonging to normal psychological development—have at least some superficial resemblances; both involve irrational beliefs and the carrying out of rituals aimed, for example, at averting feared outcomes. The exact relevance of these superficial resemblances to the nature of OCD remains controversial, however, and it is discussed further along with other developmental issues in Section 5.15.2.5.

5.15.2.5 Etiology

Accompanying the vigorous interest in OCD there has been much research and theorizing about etiology. The research programmes are of various types, described as psychological, neurological, biological, ethological, or as various compounds of these.

"Psychological" models are characterized by appeal to learning processes involving conditioning or the more complicated information-processing envisaged in the newer cognitive psychological paradigm. There have been many psychological models of OCD, critically reviewed by Jakes (1996), and what follows is a partial review of some main points. For the most part the psychological theories involve variations and elaborations on the themes of neurotic anxiety in combination with active but maladaptive anxiety reduction. This approach to the disorder is obviously linked closely to its typical phenomenology and its classification among the anxiety disorders. A significant exception to this approach is the hypothesis that the core psychological process in OCD is a memory dysfunction leading specifically to chronic doubting and checking (Boone, Ananth, Philpott, Kaur, & Djenderedjian, 1991; Reed, 1977; Sher, Frost, & Otto, 1983). Evidence relating to this hypothesis is conflicting (Martin, Wiggs, Altemus, Rubenstein, & Murphy, 1995; Tallis, 1995), but it plays a role in neuropsychological theorizing which is discussed later in this section.

Otherwise most psychological theories of OCD have emphasized anxiety and anxiety reduction. Obsessions are regarded as expressions of extreme and inappropriate anxiety, and compulsions as strategies for reducing it. Questions then immediately arise as to (i) why the person has such excessive anxiety, and (ii) why an apparently nonsensical, repetitive procedure should reduce it. These are the etiological questions formulated from within this conception of OCD. From the beginning of psychological theorizing about anxiety disorder, there was a clear assumption in both the psychoanalytic and conditioning paradigms that excessive anxiety is generated by trauma and resulting conflict among cognitive-affective states. Applied to obsessional neurosis, the implication is that the patient has experienced or is experiencing some form of trauma, the likely candidate being experience of being out of control over highly salient events (Freud, 1913; Rachman & Hodgson, 1980). As to ritualizing, Freud went on to observe that compulsive anxiety-reducing strategies apparently draw on prerational, magical styles of thinking (Freud, 1913). The basic idea is that, in the absence of real control over salient matters (such as illness and death), helplessness and anxiety can be avoided by taking control over something more within one's power, such as the way food is prepared, or counting in a set pattern. Symbolic power to reduce anxiety would be reinforced by real power according to simple principles of operant conditioning. Cognitive-behavioral theorists have envisaged these kinds of mechanisms as being also involved in the generation and maintenance of OCD (Franzblau, Kanadanian, & Rettig, 1995; Pitman, 1987; Rachman & Hodgson, 1980).

It is a relatively weak claim, though one which can be challenged, as will be considered below, that compulsions reduce the anxiety associated with obsessions and are thereby maintained. It is a much stronger and more controversial hypothesis to suppose that, notwithstanding appearances, obsessions signify anxiety which is in some sense realistic, specifically fear of being unable to control biopsychologically salient events. For what are these hypothesized events? Pitman (1993) notes evidence of raised rates of OCD among high-combat-exposed Vietnam veterans and discusses the case of a man without previous psychopathology who under combat stress developed persistent OCD as well as post-traumatic stress disorder. He observes that if one focuses on counting the feet of the war dead—one of the patient's counting compulsions—at least one avoids looking at their faces (Pitman, 1993, p. 106). However, it is plain that such cases are not the rule, that in most cases of OCD there is no obvious trauma which might help explain onset in this kind of way. Often no precipitating event is reported by the person or family, or reported precipitants are common life events (Rettew et al, 1992; Valleni-Basile et al., 1995), although common stressful events typically exacerbate symptoms in children and adolescents (Swedo, Rapoport, Leonard et al., 1989), and in adults (Rasmussen & Tsuang, 1986). Nor is it clear that adverse parenting experiences are implicated in the development of OCD. Speculation that parents of children who develop OCD are overstrict has not been borne out in any clear way (Judd, 1965; Merkel, Pollard, Wiener, & Staebler, 1993; Valleni-Basile et al., 1995). Hibbs et al. (1991) found raised levels of "expressed emotion" in families of children with OCD, but this phenomenon is known to be associated with a variety of disorders.

The upshot of all this is that the best candidates for highly anxiogenic experiences associated with OCD are likely to be in large part intrapsychic, to do with the individual's interpretation of otherwise ordinary events. One possibility which remains popular since Freud is that the person who develops OCD has particular difficulty managing high levels of negative emotion, such as is linked to loss or aggression, and hence feels out of control when faced with these emotions, and distracts from and temporarily remedies this threatening helplessness by obsessions and compulsions. The findings most relevant here have been on the excessive worrying characteristic of generalized anxiety, with indications that worrying can function as a mechanism for avoiding accessing and processing emotionally relevant information, in adults (Borkovec & Inz, 1990) and in children (Keppel-Benson & Ollendick, 1993; Vasey, 1993). It is likely that intrapsychic avoidance in the child would reflect avoidance of similar emotionally relevant material in the family (Bolton, 1995). Insofar as excessive worrying is systematically associated with obsessional symptomatology (Tallis & de Silva, 1992), it is possible that obsessional thinking and behavior may serve a similar function as worrying, serving to avoid negative emotional states (Schut & Castonguay, 1996).

The view of OCD as fundamentally involving anxiety and anxiety reduction, particularly insofar as the anxiety has predominantly intrapsychic origins, can be elaborated in various ways in the cognitive paradigm. Carr (1974) emphasized the unrealistic threat appraisals in the disorder. McFall and Wollersheim (1979) highlighted various kinds of dysfunctional assumptions such as that one should be perfect, that mistakes should result in condemnation, that the self is powerful enough to prevent disasters. This approach explores the important possibility that what is abnormal in OCD is not the manifest symptomatic thoughts, but is rather the sometimes less apparent assumptions behind, or appraisals of, these thoughts. Another approach along these lines proposes that an inflated sense of responsibility plays a critical role in the generation of OCD (Salkovskis, 1985). This model begins with the fact that everyone has normal, unwanted, intrusive thoughts, and supposes that people with OCD place special significance on these thoughts, perhaps because OCD sufferers have an exaggerated sense of responsibility. This leads to feeling responsible for producing the thought, to feeling oneself to be a bad person, and to excessive guilt. It may also prompt the feeling that they must act to prevent a feared disaster. The "neutralizing" action results in recurrence of the thought. Thus, there arises a vicious circle involving unwanted, intrusive thoughts, feelings of responsibility and guilt, attempts to neutralize the thoughts, and hence, inadvertently, their recurrence (Salkovskis, 1985). Feelings of responsibility and guilt might be exacerbated particularly in the absence of a clear distinction between having a negative thought, and carrying out the negative action itself, a blurring which has been called "thought–action fusion" (Rachman, 1993). Because these cognitive approaches implicate self-worth, guilt, and failure, they can account for the high association between OCD and depression (Van Oppen & Arntz, 1994).

The earlier psychological models of OCD were developmental in various ways. Insofar as conditioning theory implied early traumatic experience, it was in a primitive sense developmental. In Freud's developmental stage theory, obsessionality was linked particularly to the so-called anal stage, in which anxiety about a particular kind of aggression was prominent, namely, aggressive being out of control of one's own action and thereby making a mess (Freud, 1908). Freud (1913) also construed the irrationality of OCD as a pathological expression of normal childhood magical thinking, as belonging to a prerational stage of thought which is normally later replaced by rational thought. Research has shown that rational grasp of everyday causal principles is evident from very early on, from infancy, while appeal to magic can play a role in the young child's explanations of apparent violations of these principles (e.g., Harris, 1994). Freud's developmental approach to the irrationality in OCD can, however, survive this shift in emphasis away from "stage theories" of cognitive development. Indeed, the developmental approach might if anything be strengthened by it; the fact that in children (as also in so-called "magical cultures") magical thinking can coexist with and to some extent interact with rational thinking is so far compatible with this same pattern in OCD.

The possible link between OCD and childhood magical thinking is also relevant to the cognitive formulations of OCD discussed above. Cognitive styles involving, for example, exaggerated responsibility and thought–action fusion may have their origins in apparently related ways of thinking and states of mind typical of childhood, specifically in the tendency to overestimate the self's power, alongside a not yet firmed-up distinction between thought and action (Piaget, 1937, 1960; Vygotsky 1934).

It is also possible to link a psychological developmental model of OCD with an ethological model. The possibility of viewing at least some obsessive-compulsive behaviors as displacement behaviors in the sense of ethological theory was proposed some time ago by Holland (1974). Displacement behaviors are fixed action patterns carried out in the absence of a normal releasing stimulus, typically in situations of drive conflict or frustration, while fixed action patterns are innate and adaptive motor acts or sequences normally triggered by a so-called releasing stimulus, and include behavioral routines such as pecking, grooming, hoarding, and digging (Lorenz, 1966, 1981; Tinbergen, 1953). An explanation of why fixed action patterns occur in the absence of a normal releasing stimulus, typically in situations of drive conflict or frustration, may be that they have a physiologically low initiation threshold, but it is likely that they then acquire tension reducing, reinforcing functions (Eibl-Eibesfeld, 1970; Lorenz, 1966). Such a process is of course fundamental in evolution and learning; what starts as random, meaningless (nonfunctional) behavior turns out to produce benefits for the organism, typically drive-reduction, and is maintained, intensified, and shaped because of those properties. It seems plausible to suggest that the performance of displacement behaviors when faced with being unable to do anything else functions as an "illusion of control" for the animal. This points towards the area of the developmental cognitive account, and specifically towards the functional benefits of ritualized thinking and behavior already discussed.

It is possible in this way to piece together a psychological model of OCD which integrates a reasonable amount of both early and contemporary theorizing, and which provides that strong kind of biopsychological developmental model which runs seamlessly through phylogenesis and ontogenesis. This kind of model of OCD—sketched in some more detail in Bolton (1996)—is a reasonably good fit with some of the data but not with others. As already remarked, the basic view of obsessions and compulsions fits well because it is based on typical phenomemology. As is also well known, the core of this kind of psychological formulation of OCD—which emphasizes anxiety and its reduction—can provide a reasonable model for the effectiveness of behavior therapy methods, to be discussed later, which consist of methods of anxiety-reduction and the acquisition of more adaptive coping. None of these superficial good fits with the data are at all conclusive support, however. Other aspects of the psychological model as sketched above are doubtful and at best unproven, including, as discussed above, the hypothesized experience of being out of control over negative salient events, the link between OCD and normal developmental childhood magical thinking (Leonard 1989; Leonard, Goldberger, Rapoport, Cheslow, & Swedo, 1990), and the role of exaggerated sense of responsibility (Bogert, Katz, Amir, & Foa, 1996; Ladouceur et al., 1995; Van Oppen & Arntz, 1994). A further general problem for developmental accounts of any disorder is how to handle late onset, which in the case of OCD can be well into adulthood.

Relatively good or poor fits with the data are common enough in models of psychopathology, but what is particularly interesting in the case of OCD is the appearance of radical questioning about the orthodox way of looking at things. In its sharpest distinction, the alternative view stands on its head the conventional interpretation of the relation between obsessions and compulsions. It regards compulsive behaviors as primary, as triggered by some kind of neurological deficit, either structural or functional (biochemical) or both, and obsessions as being either mental versions of compulsive behaviors, or as *post hoc* rationalizations of—confabulations about—compulsive behaviors. Versions of this kind of model of OCD have been proposed by Rapoport and her colleagues (Rapoport, 1990; Rapoport, Swedo & Leonard, 1992; Rapoport & Wise 1988), and have been discussed, for example, by Insel and Winslow (Insel, 1988, 1992; Winslow & Insel, 1991). These models propose that the neurological basis of OCD involves cortico-striatal-thalamocortical circuitry, consistent with various lines of evidence including functional brain imaging (Baxter et al., 1988; Lucey et al., 1995; Luxenberg et al. 1988; Swedo et al., 1992; McGuire et al., 1994; Rubin, Villanueva-Meyer, Ananth, Trajmar, & Mena, 1992). The models postulate that structural and/or biochemical lesions in these systems may underlie OCD, and they combine the postulate of neural deficit with the ethological approach to compulsive behaviors, regarding them as akin to displacement activities. However, in the context of neurological deficit models the hypothesis is that these behaviors are triggered meaninglessly, in contrast to the suggestion within psychological models that they are regulated by information, and specifically that they function in the reduction of helplessness and anxiety.

Both Insel and Rapoport also refer to the apparently distinct psychological approach to OCD mentioned earlier, which focuses not so much on the ritualistic resolution of anxiety but more on the chronic doubting which characterizes the disorder. They propose that the basal ganglia, as well as regulating displacement/compulsive behaviors, may also be implicated in the resolution of doubt, and hence that lesions in these same structures may lead to the chronic and extreme skepticism characteristic of OCD (Insel, 1988; Rapoport, 1990).

Various kinds of evidence have been cited in support of the view that a neurological deficit underlies OCD. Association between obsessive-compulsive symptomatology and known as cerebral disease has long been known, particularly the case of encephalitis lethargica (von Economo, 1931), though also Sydenham's chorea (Grimshaw, 1964; Swedo, Rapoport, Cheslow et al., 1989), and multiple sclerosis (Miguel et al., 1995b). There is also association between obsessive-compulsive symptomatology and suspected signs of cerebral disorder, specifically tics (Como, 1995; Pauls, Alsobrook, Goodman, Rasmussen, & Leckman, 1995; Santangelo et al., 1994), and it occasionally appears following head injury (Jenike & Brandon, 1988; Max et al., 1995; McKeon, McGuffin, & Robinson, 1984). OCD sometimes responds to psychosurgery (Chiocca & Martuza, 1990; Hay et al., 1993). Further evidence of abnormal neurological involvement in OCD is well-replicated above normal frequency of neurological soft signs, and not so well replicated below normal performance on neuropsychological tasks such as tests of complex spatial and switching ability (Abbruzzese, Ferri, & Scarone, 1995; Alarcon, Libb, & Boll, 1994; Christensen, Kim, Dysken, & Hoover, 1992; Cox, Fedio, & Rapoport, 1989; Denckla 1989; Head, Bolton, & Hymas, 1989; Hollander et al., 1990; Hymas, Lees, Bolton, Epps & Head, 1991; Max et al., 1993). The most direct evidence of abnormal structural neurological involvement OCD is structural abnormalities detected by computed tomography (CT) or magnetic resonance imaging (MRI), though the findings here are conflicting (Hoehn-Saric & Benkelfat, 1994; Peterson, 1995; Robinson et al., 1995).

Apart from the equivocal structural scan results, the evidence for neurological deficit in OCD can be seen to be circumstantial and something of a mixture. Nor is it the case that the various signs of neurological disorder hang together. For example, in the above cited studies, the extent of neurological soft signs typically do not correlate with the extent of impairments on neuropsychological tests. In an interesting study of various commonly accepted signs of organic dysfunction, such as birth complications and neurological signs, in 61 children and adolescents with OCD and a psychiatric control group, Thomsen and Jensen (1991) found that a latent class analysis assigned fewer OCD children than control patients to the "organic" class, and concluded that the disorder did not depend on cerebral disturbance as detected by conventional neuropediatric methods.

As already indicated, there is the possibility of a specific biochemical abnormality in OCD. The disorder is highly selective in its response to psychotropic medication, suggesting a relatively circumscribed underlying neurochemistry. Most medications which are effective in the treatment of anxiety, depression, or psychosis are generally ineffective for OCD. The only drugs shown so far in group studies of various designs to be effective in the treatment of OCD are the potent inhibitors of synaptic uptake of serotonin (SRIs) (Piccinelli, Pini, Bellantuono, & Wilkinson, 1995; Stein, Spadaccini, & Hollander, 1995). The role of serotonergic fibers in mediating specific behavior remains speculative, however, though there is some evidence that serotonin has a prominent role among the brain mechanisms which may modulate aggressive behavior (Insel & Winslow, 1992). It is also notable that SRIs have been shown to be effective in the treatment of anxiety disorders other than OCD, in adults and in children (Birmaher et al., 1994; Schneier, Chin, Hollander, & Liebowitz, 1992) and this is consistent with the possibility that their effect in OCD is mediated by an effect on anxiety.

In this context a general point can be noted about the links between abnormal neurological function associated with a disorder, such as abnormal biochemistry or regional glucose metabolism, and "neurological deficit." The main point is that the former does not entail the latter. In brief, abnormal neurological function (in the sense of difference from that associated with no psychopathology) is so far quite compatible with the hypothesis that the disorder involves and is best explained in terms of psychological processes (intentionality, learning, information-processing, etc.). There is no legitimate inference from neurochemical or functional imaging findings to "neurological deficit," if this is understood as meaning that psychological processes and explanation have come to end (as they have in epilepsy, for example). There are complicated issues here to do with the logic of explanation of mental order and disorder, discussed in detail in Bolton and Hill (1996).

To summarize, research programs have focused on various aspects of OCD, including basic psychological processes involved, precursors in phylogenesis, models of how these map onto and are realized by neural structures and functions, and cutting across those topics there is apparent disagreement as to whether etiology can be understood in terms of biopsychological principles (such as adaptation to perceived loss of control) or whether the disorder is mainly a matter of a neurological lesion causing meaningless compulsions and confabulating obsessions. It would be fair to conclude that the etiology of OCD remains in fundamental ways uncertain. The traditional model of OCD and its various expressions in the broad cognitive psychological paradigm have a great deal of plausibility and evidence on their side. They do not, however, predict the array of findings which seem to implicate a neurological deficit. On the other hand, neurological deficits have not yet been definitely identified, and the possible interactions with psychological processes and symptomatology remain speculative.

In the background is the question whether OCD as a nosological entity will correspond to a "kind" in terms of psychological, biological, or neurological theories. It is possible from any or all those points of view that, on the one hand, subtypes of OCD may be significantly different, and on the other hand, that OCD may not be fundamentally different from some other disorders. Concerning neurobiological aspects of OCD, Insel and Winslow (1992) conclude that subgrouping based on such symptom types, character style, or comorbid diagnoses will be increasingly important for understanding the results of neuropharmacological and functional imaging studies. In the OCD family studies and genetics literature it has been suggested that a subtype of OCD is defined by association with tic disorders (Leonard et al., 1992; Pauls et al., 1995), and phenomenological differences have been reported (George, Trimble, Ring, Sallee, & Robertson, 1993). It has also been suggested that there may be an infection-triggered, autoimmune subtype of pediatric OCD and Tourette's syndrome (TS) (Allen, Leonard, & Swedo, 1995). As to the question whether the underlying processes in OCD are similar to those in other disorders, there has been an increasing interest in the concept of "OCD spectrum disorders" (Hollander & Wong, 1995; Tynes, White, & Steketee, 1990). Particular interest has focused on comorbidity and possible etiological linkages between OCD and eating disorder (de Silva, 1993; Halmi et al., 1991; Jarry & Vaccarino 1996; Rubenstein, Altmeus, Piggott, Hess, & Murphy, 1995; Solyom, Freeman, & Miles, 1982; Thiel, Broocks, Ohlmeier, Jacoby, & Schussler, 1995). Hollander and Wong (1995) propose an inclusive model which links hypothesized subtypes of OCD to many other kinds of disorders including somatoform disorders, dissociative disorders, eating disorders, schizo-obsessive disorder, tics, neurological disorders, and impulse disorders.

On the basis of studies, the role of genetic determination of OCD also remains unclear. The standard methodology of twin/adoption studies has not been readily applicable to OCD because of its rarity. Alternatively, particularly when the prevalence of the disorder was thought to be much rarer than it now is, discovery of identical twins both of whom had OCD was thought to provide evidence of genetic determination (Woodruff & Pitts, 1964), and there have been many reports of such twins in the literature (reviewed by Carey & Gottesman, 1981; Macdonald, Murray, & Clifford, 1991; and Pauls, Raymond, & Robertson, 1991). Generally, these findings are inconclusive in the absence of control for environmental factors, standardly by comparison with discordant monozygotic twins.

In the absence of sufficient numbers of twins one or both of whom have OCD, an alternative methodology involves investigation of obsessional traits in community sample twins. Such studies have suggested heritability of a general neurosis factor. Clifford, Murray, and Fulker (1984) gave the Leyton Obsessional Inventory (LOI) and the Eysenck Personality Questionnaire (EPQ) to 419 pairs of nonclinic referred twins. Heritability for LOI trait and symptom scales were estimated as 47% and 44%, respectively and comparison with the EPQ suggested a strongly inheritable neurosis factor. This finding was consistent with studies by Young, Fenton, and Lader (1971) and Torgerson (1980). Clifford et al. (1984) also found, however, an independent obsessional trait factor relating to "incompleteness, cleanliness, and checking."

Again in the absence of a direct twin methodology for studying the genetics of OCD, an alternative, weaker approach examines incidence of the disorder in the families of individuals who have it. Family studies of relatives of OCD probands have consistently suggested rates of OCD or "subthreshold" OCD higher than population base rates, ranging between 4% and 35% (Bellodi, Sciuto, Diaferia, Ronchi, & Smeraldi, 1992; Lenane et al., 1990; Leonard et al., 1992; Riddle et al., 1990; Thomsen, 1995b). This large range reflects at least the vagueness or variable use of the notion of "subthreshold" OCD, together with the fact that the base rate of this condition in the community is unknown; furthermore, the absence of nonclinic control groups in these studies is a significant methodological weakness. Another problem is absence of control for anxiety. Two family studies which controlled for anxiety by including an anxious non-OCD group failed to find raised rates of obsessive-compulsive symptomatology in first-degree relatives (Clark & Bolton, 1985; Last, Hersien, Kazdin, Orvaskhel, & Perrin, 1991). OCD family studies by Black and his colleagues bear on both these points. Black, Noyes, Goldstein, and Blum (1992) found that the rates of anxiety disorder were higher among the relatives of the probands with OCD compared with relatives of psychiatrically normal controls, but the rate of OCD was not; though they did find higher rates of subdiagnostic threshold obsessive-compulsive symptomatology in the parents of probands with OCD than in parents of the control subjects. The findings suggested that anxiety disorder diathesis is transmitted in families with OCD but that its expression is variable, and they were interpreted as consistent with the classification of OCD as an anxiety disorder. Subsequent analysis comparing psychiatric disorders in the relatives of probands with OCD and comorbid major depression or GAD suggested that GAD may represent an alternative expression of the genetic factor(s) contributing to OCD (Black, Goldstein, Noyes, & Blum, 1995).

A different line of evidence relating to genetic factors in OCD comes from studies of the disorder in relatives of patients with tics or TS. Several studies have suggested that TS is inherited as an autosomal dominant trait (Comings, Comings, Devor, & Cloninger, 1984; Kidd & Pauls, 1982; Pauls & Leckman, 1986; Price, Kidd, Cohen, Pauls, & Leckman, 1985), and the inclusion of relatives with OCD improves the fit of the single gene solution (Pauls & Leckman, 1986). These results suggest that in TS families, OCD is part of the TS spectrum. This hypothesis is supported also by the finding that rates of obsessive-compulsive symptoms are raised in relatives of patients with TS compared with relatives of patients with TS who do not themselves have TS (Comings, 1994). This suggests that the genes for TS may also carry risk for OCD. Other studies have shown that OCD occurs at similar rates in families of TS probands whether probands have OCD or not (Eapen, Pauls, & Robertson, 1993; Pauls, Towbin, Leckman, Zahner, & Cohen, 1986.)

A further study by the Yale group (Pauls et al., 1995) takes the converse approach, examining tics and TS and obsessive-compulsive symptomatology in first degree relatives of probands with OCD and nonclinical controls. Results concerning obsessive-compulsive symptomatology replicate previous findings of higher rates in relatives, though absence of an anxiety control group is a methodological weakness as noted above. As to rates of tics and TS, there was a higher rate among relatives of OCD probands than controls (4.6% vs. 1.0%), consistent with the findings of Leonard et al. (1992). This suggests a relationship between the two disorders, and that individuals with OCD and tics are more likely to have the form of OCD related to TS. It was also found that relatives were more likely to have tics/TS if the proband did, indicating that OCD and tics/TS are not always related and suggesting that in OCD families, TS is not an alternative expression of etiology responsible for OCD. This linkage may exist, however, in TS families as indicated above. The study found that relatives of early onset OCD cases (< 18 years) were more likely to have both OCD and tics, with greatest risk in case of onset between five and nine years. This suggests that childhood onset OCD may be the more heritable form (associated with tics/TS). Finally, the Yale group study found that rates of TS/tics were higher in relatives of OCD probands with a family history of OCD compared with those without family history, suggesting a nonfamilial form of OCD. Relative methodological weaknesses of this study include its use of family history as opposed to twin methodology, the absence of an anxiety control group, and use of a clinic sample, which may increase comorbidity and severity. On this last point, two studies of early onset OCD in community cohorts have failed to find association between tics and the disorder (Douglass, Moffitt, Dar, McGee, & Silva, 1995; Flament et al., 1988).

In summary, the role of genetic determination of OCD and whether there are genetically distinct subtypes remains unclear. Macdonald, Murray, and Clifford (1991) conclude their review by proposing that studies of obsessional symptomatology in larger community twin samples are needed, and this methodology would also test the validity of the Yale group's findings and hypotheses with a relatively stronger design.

5.15.2.6 Developmental Course

Studies of the course of OCD may be divided into those conducted before, and those after, development in the late 1970s of the two main current treatments, behavior therapy and pharmacotherapy with SRIs. It would be fair to say that this time period coincided roughly with increasing precision in diagnosis and increasing sensitivity of symptom measurement. A number of earlier long-term follow-up studies using various methodologies reported reasonable recovery in something the order of half of the cases (Lewis, 1936; Pollit, 1957; Rudin, 1953), though a subsequent large study indicated a somewhat less optimistic picture, however, with only one-quarter of the patients showing much improvement over the long term (Kringlen, 1965). The more rigorous studies since the 1970s are generally of the course of the disorder under more or less controlled conditions of effective behavioral and/or pharmacological treatments, and these are discussed in Section 5.15.4.

Other aspects of the course of the disorder concern continuity of presentation. While the content of obsessions and compulsions may change within individuals over time (Rettew et al., 1992), there is, as already noted, remarkable constancy in the phenomenology of OCD through the age range from childhood to adulthood. Further, the disorder usually "runs true" in the sense that disorder substitution is not common. This applies in retrospect, considering the past history of adult cases (Zeitlin, 1986), and in prospect, considering the future long-term course of child and adolescent cases (Bolton, Luckie, & Steinberg, 1995; Leonard et al., 1993; Thomsen & Mikkelsen, 1995).

More complicated, and so far hardly addressed questions, concern precursors of the disorder. As outlined above, one kind of current model of OCD postulates that it is fundamentally a neurological disorder, and it raises the question whether there is a neurodevelopmental pathway to it. Assuming that neurological injury is not sustained after birth, the broad possibilities are that there is a genetic abnormality predisposing to the disorder, or there is an early environmental impact on the brain, *in utero*, or in the form of obstetric complications. Neural abnormalities derived in one or more of the above ways, early in development, would be in principle detectable. They may have no manifestations prior to onset, but could show up as precursors of the disorder, or, particularly in the case of genetic predisposition, as risk factors evident not only in cases but also in relatives who may or may not have the disorder. The hypothesis of a neurodevelopmental pathway would be likely to predict that early onset, in childhood and early adolescence, would be associated with more signs of abnormal neurological involvement, with being male rather than female, with poorer treatment response and prognosis, and possibly with greater familial incidence of related symptomatology.

Such evidence as there is tends to be limited or nonsupportive. There is no clear support for the prediction that early onset cases, particularly males, have poorer prognosis (Bolton et al., 1995; Leonard et al., 1993; Thomsen and Mikkelsen, 1995). In addition, as discussed above, on the basis of studies so far the role of genetic determination remains unclear. So far as the author is aware there has been only one large-scale community study which has examined precursors of OCD in a prospective longitudinal design, the study was part of the Dunedin Multidisciplinary Health and Development Study based in the province of Otago, New Zealand (Douglass et al., 1995). Douglass and her colleagues report longitudinal data from birth on individuals with OCD at age 18 in a cohort of 930 people. The one year prevalence rate of OCD at age 18 was found to be 4% ($N = 37$). The majority of OCD cases met criteria for comorbid disorder, most commonly depression, social phobia, and psychoactive substance dependence. To examine correlates and predictors of OCD, the group with OCD ($N = 37$) was compared with three other groups defined as follows: without any psychopathology ("healthy group") ($N = 590$), conduct disorder without OCD ($N = 45$), and a mixed affective/anxiety disorder group (any anxiety disorder except OCD, and/or major depression or dysthymia) ($N = 215$). No significant differences between the four groups were found on measures of a number of factors implicated by current neurological deficit models of OCD, including measures of perinatal problems/complications, tics (assessed at ages seven, nine, and 11), and various neuropsychological tests of executive functioning (at age 13). The absence of association between OCD and tics/TS in the Dunedin sample replicates the same negative result found by Flament and her colleagues in their US community sample (Flament et al., 1988). Interestingly, the Dunedin OCD group was distinguished in an apparently quite different way, by reporting more anxiety and depression in late childhood and adolescence. The OCD group reported more depressive symptomatology than any other of the three comparison groups at ages 15 and 18, and at age 11 more than the "healthy group." The OCD group also reported significantly more symptoms of anxiety at age 15 than the healthy group, and at age 18 more than any other group. The investigators concluded that depression in early adolescence was a predictor for OCD in young adulthood, and that the OCD cases at age 18 may be cases of severe, persistent, internalizing disorder.

5.15.3 CONCEPTUALIZATION AND CLINICAL FORMULATION

5.15.3.1 Multimodal Assessment

Clinical diagnosis of OCD should be primarily by clinical interview, with special attention paid to issues of differential diagnosis. Clinical mental state assessment should also screen for any comorbidity, such as depression, which may require separate therapeutic attention. A structured diagnostic interview aids reliable diagnosis and differential diagnosis of OCD, and several such interviews applicable to children are available (Silverman, 1991), such as the Diagnostic Interview Schedule for Children, 2nd edition (DISC-2.1, Fisher et al., 1993), and the Anxiety Disorders Interview Schedule for Children and Parents (ADIS-C/P, Silverman & Nelles, 1988). Care should be taken to supplement such interviews if necessary with questions needed to deliver diagnosis using the most recent editions of the standard diagnostic systems.

Following clinical diagnostic interview, scaled measures of symptomatology may be used, to make assessment of symptoms more precise, and to allow for estimation of change. For children and adolescents these include the Leyton Child Obsessional Inventory (Berg, Rapoport, & Flament, 1986). This 44-item inventory measures the number of obsessive symptoms with yes/no responses as well as degree of resistance and interference with daily activities. The inventory has reasonable test-retest reliability and discriminative validity (Berg et al., 1986), and it is sensitive to symptomatic improvement (Flament et al., 1985). The Leyton Child Obsessional Inventory has also been adapted for use in surveys (Berg, Whitaker, Davies, Flament, & Rapoport, 1988; King, Inglis, Jenkins, Myerson, & Ollendick, 1995). Currently widely used with adults is the Yale–Brown Obsessive Compulsive Scale (Y-BOCS), a semistructured interviewer-based measure with reasonable reliability, validity, and sensitivity to change (Goodman et al., 1989a, 1989b; Van Oppen, Emmelkamp, van Balkom, & van Dyck, 1995; Woody, Steketee, & Chambless, 1995). In the interview, common obsessions and compulsions are screened for, and dominant ones selected for detailed assessment including estimates of severity and interference. Evidence that Y-BOCS scores for obsessions and for compulsions in fact correspond to two factors is conflicting (Coles, Amir, & Foa, 1996; McKay, Danyko, Neziroglu, & Yaryura-Tobias, 1995), and correlation between the Y-BOCS and self-report measures of obsessive-compulsive symptomatology has been found to be relatively low (Coles et al., 1996). The Children's Yale–Brown Obsessive Compulsive Scale (CY-BOCS) consists of virtually the same items and format as the adult version, with minor wording modification for younger children (Berg, 1989; Goodman, Price, Rasmussen, Riddle, & Rapoport, unpublished).

The extent of obsessive-compulsive symptomatology can also be assessed using general observational methods such as self-monitoring and monitoring by parents or nurses (Wolff & Wolff, 1991). There has also been interest in behavioral measures of severity of some aspects of obsessive-compulsive symptomatology (Steketee, Chambless, Tran, Worden, & Gillis 1996). A number of self-report measures are available for assessing unwanted, intrusive thoughts, reviewed by Clark and Purdon (1995), though generally they have been used so far with adults rather than with children.

Apart from diagnostic and quantitative assessment of symptomatology, assessment for clinical purposes depends on the model of etiology being used and on the intended methods of treatment.

5.15.3.2 Treatment Foundation

In preparing the child or adolescent and their family for treatment much depends on the treatment which is to be used. There are currently two main treatments: behavior therapy, specifically exposure with response prevention, and pharmacotherapy, specifically using serotonin reuptake inhibitors, and the two treatments can be used in combination.

If behavior therapy is to be used, the standard method of introduction is to frame OCD as an anxiety disorder, with compulsions described as methods of reducing anxiety prompted by obsessional thoughts. Behavioral treatment, specifically occurrence of the obsessional thought (perhaps cued by exposure to an appropriate external stimulus) followed by prevention of the compulsion, can then be described as a desensitization process. In this approach to OCD, preparatory assessment in effect proceeds as for any other anxiety disorder, whether treatment is to be with the individual (Kendall et al., 1991) or the family (Bolton, 1994). This preparatory assessment would include functional analysis of the anxiety—in the case of OCD of the obsessions which express it—and of coping strategies, in this case the compulsions, though perhaps also avoidance, including the role played in these by other family members. Preparation for behavioral treatment for OCD for adults is discussed by, for example, Rachman and Hodgson (1980), and Riggs and Foa (1993), and for children/adolescents by Piacentini, Gitow, Jaffer, Graae, and Whitaker (1994) and Albano, Knox, and Barlow (1995). The model of OCD as an anxiety disorder also recommends investigation of what realistic aspects of the child's life may be a source of anxiety, what coping skills they have, and what coping skills are being modeled by the parents (Bolton, 1994). If pharmacotherapy is to be used, a view of OCD as linked to biochemical neural imbalance would be proposed.

It is possible also to prepare for treatment by framing OCD in terms of what might be called a "medical model" of the disorder. This has become popular following the neurological and in particular the "neurological deficit" models of the disorder, in which obsessions or compulsions or both are regarded as senselessly triggered by abnormal brain activity. Several OCD treatment manuals for children recommend this approach, employing terms like "hiccoughs of the brain' or "brain short circuits" (Foster, 1994; March, Mulle, & Herbel, 1994; Wever & Phillips, 1994). This is clearly compatible with exclusive use of pharmacotherapy but makes less sense when combined with behavioral therapy. Behavior therapy like any other form of psychological therapy presupposes that some aspect of the symptom is operating under psychological principles, and hence can be learned or at least unlearned.

On the other hand, it might be argued that framing OCD in terms of the "medical model" may provide a useful device in the sense of strategic therapy (Haley, 1976). In this connection it is significant that March et al. (1994) cite as a rationale for their treatment introduction a book on therapy using narrative methods (White & Epston, 1990).

5.15.4 TREATMENT

5.15.4.1 Psychosocial Treatments

As already indicated, the standard psychological treatment for OCD at any age is behavioral therapy, specifically exposure with response prevention. Exposure is to feared stimuli, and these are specifically the obsessional thoughts. This exposure may occur naturally, with the passage of time, or in contact with things in daily life, but may for treatment purposes be artificially provoked by exposure to objects of fear, such as objects thought to be contaminated. Obsessional thoughts generate anxiety, and response prevention is then a matter of not carrying out the compulsion which normally serves to reduce it. At first the person's anxiety rises, but eventually it subsides. With repetition and consolidation, the anxiety associated with not performing the compulsion, and eventually the anxiety associated with the obsessional thought itself, attenuates. In psychological theory the process is fundamentally one of extinction or relearning (Foa & Kozak, 1986; Marks & Tobena, 1990; Rachman & Hodgson, 1980).

There are many crucial principles and technical points to be observed when carrying out exposure with response prevention. They are extensively discussed in the adult literature, for example, Foa and Kozak (1986), Rachman and Hogdson (1980), and Riggs and Foa (1993). Behavioral treatment for children and adolescents with OCD follows basically the same principles as for adults, though there are obviously amendments in procedure to achieve age appropriateness, for example, in explaining treatment rationale, in encouraging engagement and motivation, and in obtaining consent and authority for treatment (Albano et al., 1995; Piacentini et al., 1994). Frequently, the child or adolescent with OCD involves one or both parents in obsessive rituals, and in these circumstances a family therapy approach may be required, with response-prevention being a matter of encouraging the parents in kind but firm refusal to join in (Bolton, Collins, & Steinberg, 1983; Lenane, 1989). Evidence has appeared that adults with OCD also commonly involve family members in rituals (Calvocoressi et al., 1995; Freshman, Amir, Katz, & Foa, 1996). This may mean that family therapy is indicated in these adult cases as well, but this conclusion is not obviously correct given the high effectiveness of individual treatment for adults. The position is probably that family therapy is useful for children with OCD not simply because the family is involved in the symptoms, but more because of issues of parent–child authority and dependency; that, it is to say, for the same reasons that it is frequently best to involve parents in their child's treatment whatever the presenting problem. This implies that for the older adolescent with OCD, individual behavioral therapy requiring their consent, involvement, motivation and effort, may be indicated more than (or at least as well as) family therapy. This kind of age-related indication for individual as opposed to family therapy has been suggested in the case of eating disorders (Russell, Szmukler, Dare, & Eisler, 1987).

Cognitive models of OCD have naturally suggested cognitive therapy approaches. On the basis of their model McFall and Wollersheim (1979) recommended rational emotive therapy (Ellis, 1962). The development of cognitive therapy for depression and anxiety by Beck and his colleagues (Beck, 1976; Beck, Emery, & Greenberg, 1985) has naturally led to applications of this to OCD. As discussed in Section 5.15.2.5, the cognitive formulations of OCD implicate assumptions and styles linked to both anxiety and depression, and accordingly cognitive therapy for OCD is based on methods already established in the treatment of affective and anxiety disorders, such as challenging negative self-appraisals or overestimation of risk or responsibility (Salkovskis & Warwick, 1985; Van Oppen & Arntz, 1994). All the above work has been with adults, and its applicability to children and younger adolescents is not yet clear. It may be expected that the more the therapeutic approach requires facility with abstract or theoretical styles of thinking the more it is likely to exceed the cognitive abilities of the younger child (Durlak, Fuhrman, & Lampman, 1991; Inhelder & Piaget, 1958). The cognitive therapy approaches used with children emphasize identification and modification of maladaptive self-talk, though in the case of treatment of anxiety specifically they include also focus on appraisal of realistic likelihoods of feared outcomes (Kendall, Kane, Howard, & Siqueland, 1989). These methods have been applied to children with OCD. Henin et al. (1996) reported treatment of two girls aged nine and 10 with OCD following cognitive behavioral principles closely modeled on the methods developed by Kendall and his co-workers for treatment of anxiety in children; in the first phase of treatment the child was presented with ways of coping with situations which habitually triggered obsessive-compulsive symptoms, with use of strategies such as modeling and role-play; once the skills were mastered, exposure and response-prevention were implemented.

The kinds of case in which exposure with response-prevention can most readily be used are those in which there are readily accessible behavioral compulsions in response to obsessions. The technique does not apply in cases where the person has obsessions alone, without compulsive responses to prevent. These cases are, however, rare (Foa & Kozak, 1995). But the technique is also difficult where the compulsive responses to obsessions are themselves mental, because the therapist has no straightforward way of helping the patient stop what they are doing. Cases of mental ritualizing are not so rare (Foa & Kozak, 1995). Both kinds of case—obsessions alone, and where compulsions are predominantly mental—can be brought under the heading of obsessive-compulsive "ruminations." Treatment approaches include behavioral methods such as habituation and thought-stopping (Foa, Steketee, & Ozarow, 1985; Rachman, 1985) as well as the cognitive approaches referred to above.

Another kind of presentation of OCD which requires methods other than straightforward exposure with response-prevention is so-called primary obsessional slowness, described by Rachman (1974). Its status as a separate syndrome is questionable, however, and a linked point is that the indicated treatment—prompting, pacing, and shaping—can be construed as a form of response-prevention, since in effect it prevents the patient from maintaining desired order or exactness (Veale, 1993). In general, notwithstanding the variety of obsessive-compulsive symptomatology and the corresponding need for flexibility and innovation in behavioral treatment approaches, all or most of these approaches can plausibly be regarded as variations on the theme of response-prevention, for example, the treatment of hoarding compulsions by encouraging the patient to throw away unnecessary hoarded items (Ball, Baer, & Otto, 1996; Tallis & de Silva, 1992).

5.15.4.2 Pharmacological Treatments

As noted in Section 5.15.2.5, the only drugs shown so far in group studies of various designs to be effective in the treatment of OCD are the serotonin reuptake inhibitors (Piccinelli et al., 1995; Stein et al., 1995). They are effective not only in adults (Montgomery, 1994) but also in children and adolescents with OCD (DeVane & Sallee, 1996; DeVeaugh-Geiss et al., 1992; Flament et al., 1985; Geller, Biederman, Reed, Spencer, & Wilens, 1995; Leonard et al., 1989; Riddle et al., 1992). In the case of OCD, as for other disorders, pharmacological treatment of children and adolescents should be accompanied by attention to social and family factors (Rapoport, Leonard, Swedo, & Lenane, 1993). In the context of the demonstrated effectiveness of SRIs for OCD, much remains unclear about their mode of action, including the nature of any abnormalities in the serotonergic system of patients prior to treatment, as well as the links between serotonergic activity and psychological functions (Goodman, Price, Woods, & Charney, 1991; Hoehn-Saric, McLeod, Zimmerli, & Hipsley, 1993; Insel & Winslow, 1992; Rauch & Jenike, 1993).

5.15.4.3 Effectiveness of Treatments

Effectiveness of behavior therapy in approximately 60–80% of cases has been shown in a number of studies of adults (Marks, Hodgson, & Rachman, 1975; Rachman & Hodgson, 1980), as has effectiveness of SRIs (Clomipramine Collaborative Study Group, 1991; Fineberg, Bullock, Montgomery, & Montgomery, 1992), and of combinations of the two treatments (Cox, Swinson, Morrison, & Lee, 1993; Marks, Stern, Mawson, Cobb, & McDonald, 1980; Rachman & Hodgson, 1980; Stanley & Turner, 1995). Treatment gains following behavior therapy in adulthood have been shown to maintain in the medium- to long-term (O'Sullivan & Marks, 1990; O'Sullivan, Noshirvani, Marks, Monteior, & Lelliott, 1991), while long-term preservation of gains under pharmacotherapy may require maintenance dosage for several years (Fineberg, Bullock, Montgomery, & Montgomery, 1992; Stanley & Turner, 1995).

While behavior therapy for OCD is highly effective for most people who participate, about 25–30% of patients suitable for this form of treatment do not comply with it (McDonald, Marks, & Blizard, 1988). Among those who do stay in treatment, up to 80% respond well as noted above, but conversely 20% or more do not. In sum, then, about 40–50% of referred people with OCD are not helped by behavior therapy, and it is important to bear this in mind alongside the impressive figures for its effectiveness. Considerable effort has gone into identifying factors associated with poor response to behavioral treatment in adults. These may include over-valued ideation, comorbid personality disorder, and depression, though these and other possibilities have not as yet been consistently confirmed (Basoglu, Lax, Kasvikis, & Marks, 1988; Buchanan, Ko, & Marks, 1996; Castel et al., 1994; Fals-Stewart & Lucente, 1993; Foa, 1979; Foa, Kozak, Steketee, & McCarthy, 1992; Keijsers, Hoogduin, & Schaap, 1994; O'Sullivan et al., 1991; Reich & Vasile, 1993). The causes of nonacceptance of behavioral treatment have been relatively less explored, but low motivation is apparently involved, both in adults (Hansen, Hoogduin, Schaap, & de Haan, 1992), and in children and adolescents, for example, in the form of comorbid oppositional defiance (Bolton & Turner, 1984; Wever & Rey, in press).

Motivation in the form of determination to change and belief in self-efficacy may be a crucial factor mediating both treatment compliance and effectiveness in OCD. Cognitive techniques have been used with adults to try to enhance motivation for and engagement in behavior therapy (Salkovskis & Warwick, 1985; Tallis, 1995). It is interesting to note in this connection that the two approaches to behavioral treatment for children and adolescents with OCD referred to earlier (March et al., 1994; Wever & Phillips, 1994), as well as framing OCD in terms of the medical model, also both make liberal use of methods reminiscent of treating habit disorders such as smoking where motivation for change is crucial. Such methods typically involve acknowledging the habit as bad, and emphasizing that the person, not the behavior, is in charge. The whole problem of motivation and the person's attitude to their OCD has received relatively little attention so far in the psychological theory of the disorder. A sophisticated approach would be required, such as the highly influential model of motivation for change constructed by Prochaska and DiClemente which has been applied to treatment seeking and compliance including in relation to habitual and addictive behaviors (Prochaska & DiClemente, 1982; Prochaska, DiClemente, & Norcross, 1992).

As noted in Section 5.15.4.1, obsessional ruminations can be treated by behavioral methods such as thought stopping and habituation and by cognitive approaches such as modification of abnormal risk assessment. They are notoriously difficult to treat, however, and reported results using such methods are mixed (Arts, Hoogduin, Schaap, & de Haan, 1993; Foa et al., 1985; Rachman, 1985; Salkovskis & Westbrook, 1989).

Initial case reports of cognitive therapy for OCD in adults tended to be uncontrolled and inconclusive (James & Blackburn, 1995), but several controlled studies have shown cognitive methods to be at least as effective as exposure plus response prevention. In comparisons with rational emotive therapy, Emmelkamp and his co-workers found that the two had similar effects, with no incremental effect when combined (Emmelkamp & Beens, 1991; Emmelkamp, Visser, & Hoekstra, 1988). Van Oppen et al. (1995) compared cognitive therapy with exposure and response prevention. The cognitive therapy was based on the methods of Beck (1976) and Salkovskis (1985) and focused mainly on estimation of catastrophic risk and of personal responsibility. Seventy-one adult patients were randomly assigned to one of the two treatment conditions, and seven subsequently dropped out of each. Both treatments led to significant improvement, with more participants rated as "recovered" in the cognitive therapy group. It has been also suggested that cognitive therapy techniques might be particularly useful, and certainly worth trying, in cases of failure of behavioral treatment (Salkovskis & Warwick, 1985). As to application to children, Henin et al. (1996) have reported cognitive-behavioral treatment of two children with OCD, as already noted (Section 5.15.4.1), and the outcome was successful in both cases. The role of cognitive change in mediating behavioral change during cognitive therapy, with adults or children, is not yet clear, and indeed the same question arises in the case of behavior therapy (Durlak et al., 1991; Ladouceur, Rheaume, & Leger 1996; Van Oppen, Emmelkamp, et al., 1995).

The treatment of obsessional slowness involving prompting, shaping, and pacing is generally effective while the therapist is participating, but generalization and maintenance outside of sessions is poor (Clark, Sugrim, & Bolton, 1982; Veale, 1993).

Returning to the main treatments currently in use, exposure with response prevention and medication with SRIs, broadly similar results have been found for children and adolescents with OCD as for adults. Positive, though generally partial, response to SRIs has been found in the majority of young patients with OCD, though as for adults there tends to be relapse if medication is withdrawn (DeVane & Sallee, 1996; DeVeaugh-Geiss et al., 1992; Flament et al., 1985; Geller et al., 1995; Leonard et al., 1989; Rasmussen, Eisen, & Pato, 1993; Riddle et al., 1992). There is also usually a good response to behavior therapy in this age group. An early study of 15 adolescents treated mainly by exposure with response prevention found complete or near complete symptom relief in 87% of cases (Bolton et al., 1983). Similar results have been obtained in two studies using well-defined behavior therapy programs combined with pharmacological (SRI) treatment, with significant improvement in approximately 80% ($N = 14$) (March et al., 1994) and 68%

($N = 57$) (Wever & Rey, in press) of cases under these combined treatments.

The question of the course of OCD presenting in childhood and adolescence, and the long term effectiveness of treatments, is important clinically and theoretically. Early follow-up studies of clinically referred children and adolescents with OCD indicated poor medium- to long-term outcome. Warren (1965) followed up 15 young people aged 19–24 years who had been admitted to hospital with obsessive-compulsive states when aged 12–17 years. No treatment was specified except for psychosurgery in one case. Five of the 15 young people still had severe obsessional symptomatology at follow up, and a further four were still handicapped by obsessional symptoms. Four of the 15 tended to develop mild symptoms under stress, and the remaining two were considered completely recovered. Hollingsworth, Tanguay, Grossman, and Pabst (1980) evaluated 10 of 17 pediatric outpatients and inpatients with severe obsessional neurosis treated by intensive psychotherapy in all cases, one-and-a-half to 14 years later. Seven of the 10 cases reported continuation of obsessive-compulsive behavior though at less than pretreatment levels.

The findings from these earlier follow-up studies of reasonable recovery in a third or less of cases of pediatric OCD is comparable with other findings. In the first of a pair of studies from the National Institute of Mental Health (NIMH), Flament and her colleagues followed up 27 children and adolescents assessed as having OCD at NIMH between 1977 and 1983, between two and seven (mean 4.4) years later and at age 13–24 (mean 18.8) years (Flament et al., 1990). Of the 27 subjects, 17 (68%) still qualified for diagnosis of OCD at follow-up. Treatment during the follow-up period was described as diverse, sparse, and uncontrolled. A somewhat better prognosis was found in a Danish cohort of 23 children and adolescents followed up for between 18 months and five years after treatment which did not involve a "structured treatment program" but included use of SRIs in about 25% of cases and behavior therapy in all; about 50% of cases still qualified for diagnosis at follow-up (Thomsen & Mikkelsen, 1995). A similar recovery rate in a community sample of adolescents with OCD has been observed over a two-year period (Berg et al., 1989). Leonard and colleagues followed up a second cohort of 54 children and adolescents with OCD seen at NIMH between 1984 and 1988, between two and seven (mean 3.4) years later (Leonard et al., 1993). At follow-up 43% still qualified for diagnosis of OCD. Treatment during the follow-up period was active, in contrast to the cohort seen by Flament and her colleagues, and included medication in nearly all cases and behavior therapy in a third. The predicted difference in outcome in the two NIMH cohorts was found; improvement to below diagnostic threshold was found in 57% of cases in the Leonard et al. sample, as opposed to 32% of cases in the Flament et al. sample.

Similar results in a smaller cohort over a longer time period were reported by Bolton et al. (1995, 1996). Fourteen young adults treated for the disorder in adolescence, mainly with behavior therapy and family therapy, were followed up between nine and 14 years later: six (43%) were found to satisfy *DSM-III-R* criteria for OCD at long-term follow-up, while eight (57%) were not. Sustained recovery from OCD was associated with normal levels of social functioning, consistent with the pattern found in Thomsen's Danish sample (Thomsen, 1995a). Recovered participants were not currently taking medication, a finding which contrasted with the report of Leonard and her colleagues that, in their sample, high proportions, up to 80%, of various groups of participants who no longer met diagnostic criteria for OCD continued to receive SRI medication. One possibility is that effective behavior therapy during adolescence may obviate the need for long-term maintenance pharmacotherapy.

5.15.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

It would be fair to say that the enormous research effort which has gone into understanding OCD since the early 1980s has raised as many questions as have been answered about the nature and etiology of the disorder. There is an apparent contrast between the traditional view of OCD as fundamentally an anxiety disorder and what has emerged as a new orthodoxy, that it is fundamentally a motor disorder caused by some form of neurological lesion. No one kind of evidence, let alone a single test, could ever settle this kind of issue, but evidence from a variety of methodologies—as reviewed in this chapter—will continue to accrue until a coherent pattern or number of patterns emerge. The various methodologies exist at several levels, from the cultural to the neurological, and the questions they address are also important in their own right. Key outstanding issues include the nature of and interaction between the psychological processes implicated in the disorder, such as anxiety, magical belief, exaggerated sense of responsibility, and dysfunctional memory. There are outstanding issues concerning the neurological basis of the disorder, such as concerning the role of serotonin, and the site of structural lesions in some cases. Important questions in the genetics of the disorder remain unclear in the absence of a large community twin study.

A substantial question in the background, interweaving with etiological issues, is the extent to which OCD as understood in the 1990s, is in fact a well-defined psychobiological entity. It may turn out that significant subtypes exist within OCD and/or that causal pathways criss-cross OCD and associated or related psychopathologies. In the present state of theoretical uncertainty this is a particularly tempting line of enquiry, although as yet attempts to subcategorize OCD in useful, predictive ways have so far been speculative. One possibility specifically interesting in relation to childhood presentations of the disorder is whether there a distinctive neurodevelopmental pathway to and subtype of the disorder, and future research will no doubt cast light on this issue.

From the treatment point of view both pharmacotherapy with SRIs and behavior therapy using exposure with response-prevention have both been shown to be highly effective for many patients in the short- and long-term. Outstanding issues include how to improve outcome in treatment resistant cases, and how to manage relapse effectively. Cognitive therapy may play an increasingly important role in these respects.

5.15.6 SUMMARY

OCD is a relatively common disorder in children and young people, with current estimates suggesting a life-time prevalence by adolescence of around 2%. Comorbidity with other disorders is common, including depression, other anxiety disorders, and tics. Psychological theories of OCD mainly involve variations on the themes of unresolved anxiety in combination with active but maladaptive anxiety reduction, and they can be linked in one direction to phylogenesis in terms of displacement activities and in the other to ontogenesis in terms of styles of cognitive appraisal. An alternative view is that OCD may be a neurological motor disorder, with obsessional thought being like confabulation. Neuro-imaging findings on the neuroanatomical basis of the disorder are fairly consistent, and the specific effectiveness of SRIs indicates a relatively circumscribed neurochemical basis. On the basis of studies so far the role of genetic determination remains unclear. OCD often begins in childhood or adolescence, and a variety of useful assessment measures have been developed for these early onset cases. Without current, effective treatments the disorder can run a chronic course. Currently indicated treatments are behavior therapy, specifically exposure therapy with response-prevention, and pharmacotherapy with SRIs. Pharmacotherapy is less time-consuming but behavior therapy may protect more against relapse. Cognitive therapy for OCD has been shown to be as effective as behavior therapy with adults, and is just beginning to be applied with children.

5.15.7 REFERENCES

Abbruzzese, M., Ferri, S., & Scarone, S. (1995). Wisconsin Card Sorting Test performance in obsessive-compulsive disorder: No evidence for involvement of dorsolateral prefrontal cortex. *Psychiatry Research, 58,* 37–43.

Adams, P. L., (1973). *Obsessive children.* New York: Penguin.

Alarcon, R. D., Libb, J. W., & Boll, T. J. (1994). Neuropsychological testing in obsessive-compulsive disorder: a clinical review. *Journal of Neuropsychiatry, 6,* 217–228.

Albano, A. M., Knox, L. S., & Barlow, D. H. (1995). Obsessive compulsive disorder. In A. R. Eisen, C. A. Kearney, & C. E. Schaefer (Eds.) *Clinical handbook of anxiety disorders in children and adolescents* (pp. 282–316). Northvale, NJ: Jason Aronson.

Allen, A. J., Leonard, H. L., & Swedo, S. E. (1995). A new infection-triggered, autoimmune subtype of paediatric OCD and Tourette's syndrome. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 307–311.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.), Washington, DC: American Psychiatric Association.

Arts, W., Hoogduin, K., Schaap, C., & de Haan, E. (1993). Do patients suffering from obsessions alone differ from other obsessive-compulsives? *Behaviour Research and Therapy, 31,* 119–23.

Ball, S. G., Baer, L., & Otto, M. W. (1996). Symptom subtypes of obsessive-compulsive disorder in behavioral treatment studies: a quantitative review. *Behaviour Research and Therapy, 34,* 47–51.

Baron-Cohen, S. (1989). Do autistic children have obsessions and compulsions? *British Journal of Clinical Psychology, 28,* 193–200.

Basoglu, M., Lax, T., Kasvikis Y., & Marks, I. M. (1988). Predictors of improvement in obsessive-compulsive disorder. *Journal of Anxiety Disorders, 2,* 299–317

Baxter, L., Schwartz, J., Mazziotta, J., Phelps, M., Pahl, J., Guze, B., & Fairbanks, L. (1988). Cerebral glucose metabolic rates in nondepressed patients with obsessive-compulsive disorder. *American Journal of Psychiatry, 145,* 1560–1563.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders.* New York: International Universities Press.

Beck, A. T., Emery, G., & Greenberg, R. L. (1985). *Anxiety disorders and phobias: A cognitive perspective.* New York: Basic Books.

Bellodi, L., Sciuto, G., Diaferia, G., Ronchi, P., & Smeraldi, E. (1992). Psychiatric disorders in the families of patients with obsessive-compulsive disorder. *Psychiatry Research, 42,* 111–120.

Berg, C. Z. (1989). Behavioral assessment techniques for childhood obsessive-compulsive disorder. In J. L. Rapoport (Ed.), *Obsessive–compulsive disorder in children and adolescents* (pp. 41–70). Washington, DC: American Psychiatric Press.

Berg, C. Z., Rapoport, J. L, & Flament, M. (1986). The Leyton Obsessional Inventory Child Version. *Journal of Child Psychology and Psychiatry, 25,* 84–91.

Berg, C. Z., Rapoport, J. L., Whitaker, A., Davies, M., Leonard, H., Swedo, S. E., Braiman, S., & Lenane, M. (1989). Childhood obsessive-compulsive disorder: a two-year prospective follow-up of a community sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 528–533.

Berg, C. Z., Whitaker, A., Davies, M., Flament, M. F., & Rapoport, J. L. (1988). The survey form of the Leyton Obsessional Inventory–Child Version: norms from an epidemiological study. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 759–763.

Berman, L. (1942). Obsessive compulsive neurosis in children. *Journal of Nervous and mental Diseases, 95,* 26–39.

Berrios, G. E. (1995). Obessive–compulsive disorder. In G. E. Berrios & R. Porter (Eds.), *A history of clinical psychiatry: The origin and history of psychiatric disorders* (pp. 573–592). London: Athlone.

Birmaher, B., Waterman, G., Ryan, N., Cully, M., Balach, L., Ingram, J., & Brodsky, M. (1994). Fluoxetine for childhood anxiety disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 993–999.

Black, D. W., Goldstein, R. B., Noyes, R., & Blum, N. (1995). Psychiatric disorders in relatives of probands with obsessive-compulsive disorder and co-morbid major depression or generalized anxiety. *Psychiatric Genetics, 5,* 37–41.

Black, D. W., Noyes, R., Jr., Goldstein, R. B., & Blum, N. (1992). A family study of obsessive-compulsive disorder. *Archives of General Psychiatry, 49,* 362–368.

Bland, R. C., Newman, S. C., & Orn, H. (1988) Life-time prevalence of psychiatric disorders in Edmonton. *Acta Psychiatrica Scandinavica, 77* (Suppl. 338), 24–32.

Bodfish, J. W., Crawford, T. W., Powell, S. B., Parker, D. E., Golden, R. N., & Lewis, M. H. (1995). Compulsions in adults with mental retardation: Prevalence, phenomenology, and comorbidity with stereotypy and self-injury. *American Journal of Mental Retardation, 100,* 183–192.

Bogert, K. V., Katz, K., Amir, N., & Foa, E. B. (1996). Exaggerated responsibility for harm in obsessive--compulsive disorder. Poster presented at the 30th Annual Convention of the Association for the Advancement of Behavior Therapy, New York, November 1996.

Bolton, D. (1994). Family systems interventions. In T. H. Ollendick, W. Yule, & N. King (Eds.), *International handbook of phobic and anxiety disorders of children* (pp. 397–414). New York: Plenum.

Bolton, D. (1995). Self-knowledge, error and disorder. In M. Davies & A. Stone (Eds.), *Mental simulation: Evaluations & applications* (pp. 209–234). Oxford: Blackwell.

Bolton, D. (1996). Annotation: Developmental issues in obsessive-compulsive disorder. *Journal of Child Psychology and Psychiatry, 37,* 131–137.

Bolton, D., Collins, S., & Steinberg, D. (1983). The treatment of obsessive-compulsive disorder in adolescence: A report of fifteen cases. *British Journal of Psychiatry, 142,* 456–464.

Bolton, D., & Hill, J. (1996). *Mind, meaning, and mental disorder: The nature of causal explanation in psychology and psychiatry*. Oxford: Oxford University Press.

Bolton, D., Luckie, M., & Steinberg, D. (1995). Long-term follow-up of obsessive-compulsive disorder treated in adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 1441–1450.

Bolton, D., Luckie, M., & Steinberg, D. (1996). Obsessive–compulsive disorder treated in adolescence: 14 long-term case histories. *Clinical Child Psychology and Psychiatry, 1,* 411–432.

Bolton, D., & Turner, T. (1984). Obsessive–compulsive neurosis with conduct disorder in adolescence: A report of two cases. *Journal of Child Psychology and Pychiatry, 25,* 133–139.

Boone, K., Ananth, J., Philpott, L., Kaur, A., & Djenderedjian, A. (1991) Neuropsychological characteristics of nondrepresed adults with obsessive-compulsive disorder. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 4,* 96–109.

Borkovec, T. D., & Inz, J. (1990). The nature of worry in generalized anxiety disorder: a predominance of thought activity. *Behaviour Research and Therapy, 28,* 153–158.

Buchanan, A. W., Ko, S. M., & Marks, I. M. (1996). What predicts improvement and compliance during the behavioral treatment of obsessive-compulsive disorder? *Anxiety, 2,* 22–27.

Calvocoressi, L., Lewis, B., Harris, M., Trufan, S., Goodman, W. K., McDougle, C. J., & Price, L. H. (1995) Family accomodation in obsessive-compulsive disorder. *American Journal of Psychiatry, 152,* 441–443.

Carey, G., & Gottesman, I. (1981). Twin studies of anxiety, phobic, and obsessive disorders. In D. F. Klein & J. Rabkins (Eds.), *Anxiety: new research and changing concepts* (pp. 117–133). New York: Raven Press.

Carr, A. T. (1974). Compulsive neurosis: A review of the literature. *Psychological Bulletin, 81,* 311–318.

Castle, D. J., Deale, A., Marks, I. M., Cutts, F., Chadhoury, Y., & Stewart, A. (1994). Obsessive–compulsive disorder: Prediction of outcome from behavioral psychotherapy. *Acta Psychiatrica Scandinavica, 89,* 393–398.

Cath, D. C., van-de-Wetering, B. J., van-Woerkom, T. C., Hoogduin, C. A., Roos, R. A., & Rooijmans, H. G. (1992). Mental play in Gilles de la Tourette's syndrome and obsessive-compulsive disorder. *British Journal of Psychiatry, 161,* 542–545.

Chakraborty, A., & Banerji, G. (1975). Ritual, a culture-specific neurosis, and obsessional states in Bengali culture. *Indian Journal of Psychiatry, 17,* 211–216.

Chiocca, A. E., & Martuza, R. L. (1990). Neurosurgical therapy of obsessive-compulsive disorder. In M. A. Jenike, L. Baer, & W. E. Minichiello (Eds.), *Obsessive compulsive disorders: Theory and management* (pp. 283–294). Chicago: Yearbook Medical Publishers.

Christensen, K. J., Kim, S. W., Dysken, M. W., & Hoover, K. M. (1992). Neuropsychological performance in obsessive-compulsive disorder. *Biological Psychiatry, 31,* 4–18.

Clark, D. A., & Bolton, D. (1985). Obsessive–compulsive adolescents and their parents: A psychometric study. *Journal of Child Psychology and Psychiatry, 26,* 267–276.

Clark, D. A., & Purdon, C. L. (1995). The assessment of unwanted intrusive thoughts: A review and critique of the literature. *Behaviour Research and Therapy, 33,* 967–976.

Clark, D. A., Sugrim, I., & Bolton, D. (1982). Primary obsessional slowness: A nursing treatment programme with a 13 year old male adolescent. *Behaviour Research and Therapy, 20,* 289–292.

Clifford, C. A., Murray, R. M., & Fulker, D. W. (1984). Genetic and environmental influences on obsessional traits and symptoms. *Psychological Medicine, 14,* 791–800.

Clomipramine Collaborative Study Group (1991). Clomipramine in the treatment of patients with obsessive-compulsive disorder. *Archives of General Psychiatry, 48,* 730–738.

Coles, M. E., Amir, N., & Foa, E. B. (1996). *Factor structure of the Yale–Brown Obsessive–Compulsive Scale (Y-BOCS)*. Poster presented at the 30th Annual

Convention of the Association for the Advancement of Behavior Therapy, New York, November 1996.

Coles, M., Bogert, K. V., Krause, M. S., Amir, N., Salkovskis, P., Kozak, M., & Foa, B. (1996). *The Obsessive Compulsive Inventory*. Poster presented at the 30th Annual Convention of the Association for the Advancement of Behavior Therapy, New York, November 1996.

Comings, D. E. (1994). Tourette syndrome: A heriditary neuropsychiatric spectrum disorder. *Annual Clinical Psychiatry, 6,* 235–247.

Comings, D. E., Comings, B. G., Devor, E. J., & Cloninger, C. R. (1984). Detection of a major gene for Gilles de la Tourette syndrome. *American Journal of Human Genetics, 36,* 704–709.

Como, P. G. (1995). Obsessive–compulsive disorder in Tourette's syndrome. In W. J. Weiner & A. E. Lang (Eds.), *Behavioral neurology of movement disorders* (pp. 281–291), *Advances in Neurology, 65*. New York: Raven Press.

Cox, B. J., Swinson, R. P., Morrison, B., & Lee, P. S. (1993). Clomipramine, fluoxetine, and behavior therapy in the treatment of obsessive-compulsive disorder: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 24,* 149–153

Cox, C., Fedio, P., & Rapoport, J. L.(1989). Neuropsychological testing of obsessive-compulsive adolescents. In J. L. Rapoport (Ed.), *Obsessive–compulsive disorder in children and adolescents* (pp. 173–85). Washington, DC: American Psychiatric Press.

Denckla, M. (1989). Neurological examination. In J. L. Rapoport (Ed.), *Obsessive–compulsive disorder in children and adolescents* (pp. 107–115). Washington, DC: American Psychiatric Press.

DeVane, C. L., & Sallee, F. R. (1996). Serotonin selective reuptake inhibitors in child and adolescent psychopharmacology: a review of published experience. *Journal of Clinical Psychiatry, 57,* 55–66.

DeVeaugh-Geiss, J., Moroz, G., Biederman, J., Cantwell, D., Fontaine, R., Greist, J., Reichler, R., Katz, R., & Landau, P. (1992). Clomipramine hydrochloride in childhood and adolescent obsessive-compulsive disorder: a multicenter trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 45–49.

de Silva, P. (1993) Anorexia and obsessive-compulsive disorder. *Neuropsychiatrie de l'Enfance, 41,* 269–272.

Douglass, H. M., Moffitt, T. E., Dar, R., McGee, R., & Silva, P. (1995). Obsessive–compulsive disorder in a birth cohort of 18-year-olds: prevalence and predictors. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 1424–1431.

Durlak, J. A., Fuhrman, T., & Lampman, C. (1991). Effectiveness of cognitive-behavior therapy for maladapting children: A meta-analysis. *Psychological Bulletin, 110,* 204–214.

Eapen, V., Pauls, D. L., & Robertson, M. M. (1993). Evidence for autosomal dominant transmission in Tourette's syndrome—United Kingdom Cohort Study. *British Journal of Psychiatry, 162,* 593–596.

Eibl-Eibesfeld, I. (1970). *The biology of behavior* (Trans. E. Klinghammer). New York: Holt, Rinehart and Winston.

Ellis, A. (1962) *Reason and emotion in psychotherapy*. New York: Lyle-Stuart.

Emmelkamp, P., & Beens, H. (1991) Cognitive therapy with obsessive-compulsive disorder: a comparative evaluation. *Behaviour Research and Therapy, 29,* 293–300.

Emmelkamp, P., Visser, S., & Hoekstra, R. (1988). Cognitive therapy vs. exposure *in vivo* in the treatment of obsessive-compulsives. *Cognitive Therapy and Research, 12,* 103–114.

Fals-Stewart, W., & Lucente, S. (1993). An MCMI cluster typology of obsessive-compulsives: A measure of personality characteristics and its relationship to treatment participation, compliance and outcome in behaviour therapy. *Journal of Psychiatric Ressearch, 27,* 139–154.

Fineberg, N. A., Bullock, T., Montgomery, D. B., & Montgomery, S. A. (1992). Serotonin reuptake inhibitors are the treatment of choice in obsessive-compulsive disorder. *International Clinical Psychopharmacology, 7* (Suppl. 1), 43–47.

Fisher, P. W., Shaffer, D., Piacentini, J. C., Lapkin, J., Kafantaris, V., Leonard, H., & Herzog, D. B. (1993). Sensitivity of the Diagnostic Interview Schedule for Children, 2nd. ed. (DISC-2.1) for specific diagnoses of children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 666–673.

Flament, M. F., Koby, E., Rapoport, J. L., Berg, C. J., Zahn, T., Cox, C., Denckla, M., & Lenane, M. (1990). Childhood obsessive-compulsive disorder: a prospective follow-up study. *Journal of Child Psychology and Psychiatry, 31,* 363–380.

Flament, M. F., Rapoport, J. L., Berg, C. Z., Sceery, W., Kilts, C., Mellstrom, B., & Linnoila, M. (1985). Clomipramine treatment of childhood obsessive-compulsive disorder: A double-blind controlled study. *Archives of General Psychiatry, 42,* 977–983.

Flament, M. F., Whitaker, A., Rapoport, J. L., Davies, M., Berg, C. Z., Kalikow, K., Sceery, W., & Shaffer, D. (1988). Obsessive compulsive disorder in adolescence: an epidemiological study. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 764–771.

Foa, E. B. (1979). Failures in treating obsessive-compulsives. *Behaviour Research and Therapy, 17,* 169–176.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: exposure to corrective information. *Psychological Bulletin, 99,* 20–35.

Foa, E. D., & Kozak, M. J. (1995). *DSM-IV* field trial: Obsessive-compulsive disorder. *American Journal of Psychiatry, 152,* 90–96.

Foa, E. D., Kozak, M. J., Goodman, W. K., Hollander, E., Jenike, M. A., & Rasmussen, S. A. (1995). DSM-IV field trial: Obsessive–compulsive disorder. Correction. *American Journal of Psychiatry, 152,* 654.

Foa, E. B., Steketee, G. S., & Ozarow, B. J. (1985). Behavior therapy with obsessive-compulsives: From theory to treatment. In M. Mavissakalian, S. M. Turner, & L. Michelson (Eds.), *Obsessive-compulsive disorder: Psychological and pharmacological treatment* (pp. 49–129). New York: Plenum.

Foa, E. B., Kozak, M. J., Steketee, G. S., & McCarthy, P. R. (1992). Treatment of depressive and obsessive-compulsive symptoms in OCD by imipramine and behaviour therapy. *British Journal of Clinical Psychology, 31,* 279–292.

Fonseca, A. C., Yule, W., & Erol, N. (1994). Cross-cultural issues. In T. H. Ollendick, W. Yule, & N. King (Eds.), *International handbook of phobic and anxiety disorders of children* (pp. 67–44). New York: Plenum.

Foster, C. (1994). *Polly's Magic Games*. Maine: Dilligraff Publications.

Franzblau, S. H., Kanadanian, M., & Rettig, E. (1995). Critique of reductionistic models of obsessive-compulsive disorder: Toward a new explanatory paradigm. *Social Science and Medicine, 41,* 99–112.

Freshman, M. S., Amir, N., Katz, K., & Foa, E. B. (1996). *Family accommodation and rejection in relatives of OCD patients*. Poster presented at the 30th Annual Convention of the Association for the Advancement of Behavior Therapy, New York, November 1996.

Freud, S. (1908). Character and anal eroticism. In J. Strachey (Ed.), *Standard edition of the complete psychological works of Sigmund Freud* (Vol. 9, pp. 169–175). London: Hogarth Press, 1966.

Freud, S. (1913). Totem and taboo. In J. Strachey (Ed.), *Standard edition of the complete psychological works of*

Sigmund Freud, (Vol. 13, pp. 1–62). London: Hogarth Press, 1966.

Garrison, C. Z., Valleni-Basile, L. A., Jackson, K. L., Waller, J. L., McKeown, R. E., Addy, C. L., & Cuffe, S. P. (1995). Frequency of obsessive-compulsive disorder in a community sample of young adolescents: Errata. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 128–129.

Geller, D. A., Biederman, J., Reed, E., Spencer, T., & Wilens, T. (1995). Similarities in response to fluoxetine in the treatment of children and adolescents with obsessive-compulsive disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 36–44.

Gesell, A., Ames, L., & Ilg, F. (1974). *Infant and child in the culture today*. New York: Harper and Row.

George, M. S., Trimble, M. R., Ring, H. A., Sallee, F. R., & Robertson, M. M. (1993). Obsessions in obsessive-compulsive disorder with and without Gilles de la Tourette's syndrome. *American Journal of Psychiatry, 150,* 93–97.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Delgado, P., Heninger, G. R., & Charney, D. S. (1989b). The Yale–Brown Obsessive Compulsive Scale (Y-BOCS): Part II. Validity. *Archives of General Psychiatry, 46,* 1012–1016.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischman, R. L., Hill, C. L., Heninger, G. R., & Charney, D. S. (1989a). The Yale–Brown Obsessive Compulsive Scale (Y-BOCS): Part I. Development, use, and reliability. *Archives of General Psychiatry, 46,* 1006–1011.

Goodman, W. K., Price, L. H., Woods, S. W., & Charney, D. S. (1991). Pharmacologic challenges in obsessive-compulsive disorder. In J. Zohar, T. Insel, & S. Rasmussen (Eds.), *The psychobiology of obsessive-compulsive disorder* (pp. 162–186). New York: Springer.

Greenberg, D. (1984). Are religious compulsions religions or compulsive: A phenomenological study? *American Journal of Psychotherapy, 38,* 524–532.

Grimshaw, L. (1964). Obsessional disorder and neurological illness. *Journal of Neurology, Neurosurgery and Psychiatry, 27,* 229–231.

Haley, J. (1976). *Problem-solving therapy*. San Fransisco: Jossey-Bass.

Halmi, K. A., Eckert, E., Marchi, P., Sampugnaro, V., Apple, R., & Cohen, J. (1991). Comorbidity of psychiatric diagnoses in anorexia nervosa. *Archives of General Psychiatry, 48,* 712–718.

Hansen, A. M., Hoogduin, C. A., Scaap, C. A., & de Haan, E. (1992). Do drop-outs differ from successfully treated obsessive-compulsives? *Behaviour Research and Therapy, 30,* 547–550.

Harris, P. L. (1994). Unexpected, impossible and magical events: children's reactions to causal violations. *British Journal of Developmental Psychology, 12,* 1–7.

Hay, P., Sachdev, P., Cumming, S., Smith, J. S., Lee, T., Kitchener, P., & Matheson, J. (1993). Treatment of obsessive-compuylsive disorder by psychosurgery. *Acta Psychiatrica Scandinavica, 87,* 197–207.

Head, D., Bolton, D., & Hymas, N. (1989). Deficit in cognitive shifting ability in patients with obsessive-compulsive disorder. *Biological Psychiatry, 23,* 323–327.

Henin, A., Kendall, P. C., Flannery-Schroeder, E., Warman, M., Sugarman, A., & Marrs, A. (1996). *Cognitive behavioral therapy of OCD in children: A multiple baseline analysis.* Poster presented at the 30th Annual Convention of the Association for the Advancement of Behavior Therapy, New York, November 1996.

Hibbs, E. D., Hamburger, S. D., Lenane, M., Rapoport, J. L., Kruesi, M. J. P., Caesar, C. S., & Goldstein, M. J. (1991). Determinants of expressed emotion in families of normal and disturbed children. *Journal of Child Psychology and Psychiatry, 32,* 757–770.

Hoehn-Saric, R., & Benkelfat, C. (1994). Structural and functional brain imaging in obsessive-compulsive disorder. In E. Hollander, J. Zohar, D. Marazziti, & B. Olivier (Eds.), *Current insights in obsessive-compulsive disorder* (pp. 183–211). Chichester, UK: Wiley.

Hoehn-Saric, R., McLeod, D., Zimmerli, W., & Hipsley, P. (1993). Symptoms and physiologic manifestations in obsessive-compulsive patients before and after treatment with clomipramine. *Journal of Clinical Psychiatry, 54,* 272–276.

Holland, H. C. (1974). Displacement activity as a form of abnormal behaviour in animals. In H. R. Beech (Ed.), *Obsessional states* (pp. 161–173). London: Methuen.

Hollander, E., Schiffman, E., Cohen, B., Rivera-Stein, M. A., Rosen, W., Gorman, J. M., Fryer, A. K., Papp, L., & Liebowitz, M. R. (1990). Signs of central nervous system dysfunction in obsessive-compulsive disorder. *Archives of General Psychiatry, 47,* 27–32.

Hollander, E., & Wong, C. M. (1995). Obsessive–compulsive spectrum disorders. *Journal of Clinical Psychiatry, 56,* 3–6.

Hollingsworth, C. E., Tanguay, P. E., Grossman, L., & Pabst, P. (1980). Long-term outcome of obsessive-compulsive disorder in childhood. *Journal of the American Academy of Child and Adolescent Psychiatry, 19,* 134–144.

Hwu, H. G., Yeh, E. K., & Chang, L. Y. (1989). Prevalence of psychiatric disorders in Taiwan. *Acta Psychiatrica Scandinavica, 79,* 136–147.

Hymas, N., Lees, A., Bolton, D., Epps, K., & Head, D. (1991). The neurology of obsessional slowness. *Brain, 114,* 2203–2233.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolesence*. New York: Basic Books.

Insel, T. R. (1988). Obsessive compulsive disorder: neuroethological perspective. *Psychopharmacology Bulletin, 24,* 365–369.

Insel, T. R. (1992). Neurobiology of obsessive-compulsive disorder: a review. *International Clinical Psychopharmacology, 1,* 31–33.

Insel, T. R., & Winslow, J. T. (1992). Neurobiology of obsessive-compulsive disorder. *Psychiatric Clinics of North America, 15,* 813–824.

Jakes, I. (1996). *Theoretical approaches to obsessive-compulsive disorder*. Cambridge: Cambridge University Press.

James, I. A., & Blackburn, I. M. (1995). Cognitive therapy with obsessive-compulsive disorder. *British Journal of Psychiatry, 166,* 444–450.

Janet, P. (1903). *Les obsessions et la psychiasthenie* (Vol. I), Paris: Felix Alan.

Jarry, J. L., & Vaccarino, F. J. (1996). Eating disorder and obsessive-compulsive disorder: neurochemical and phenomonological abnormalities. *Journal of Psychiatry and Neuroscience, 21,* 36–48.

Jenike, M. A., & Brandon, A. D. (1988). Obsessive–compulsive disorder and head trauma: a rare association. *Journal of Anxiety Disorders, 2,* 353–359.

Judd, L. L. (1965). Obsessive compulsive neurosis in children. *Archives of General Psychiatry, 12,* 136–143.

Karno, M., Golding, J. M., Sorenson, S. B., & Burnam, M. A. (1988). The epidemiology of obsessive-compulsive disorder in five US communities. *Archives of General Psychiatry, 45,* 1094–1099.

Keijsers, G., Hoogduin, C., & Schaap, C. (1994). Predictors of treatment outcome in the behavioral treatment of obsessive-compulsive disorder. *British Journal of Psychiatry, 165,* 781–786.

Kendall, P. C., Chansky, T. E., Friedman, M., Kim, R., Kortlander, E., Sessa, F. M., & Siqueland, L. (1991). Treating anxiety disorders in children and adolescents. In P. C. Kendall (Ed.), *Child and adolescent therapy*:

Cognitive-behavioral procedures (pp. 131–164). New York: Guildford Press.

Kendall, P. C., Kane, M., Howard, B., & Siqueland, L. (1989). *Cognitive-behavioral therapy for anxious children: Treatment manual*. Philadelphia: Psychology Department, Temple University.

Keppel-Benson, J. M., & Ollendick, T. H. (1993). Post traumatic stress disorders in children and adolescents. In C. F. Saylor (Ed.), *Children and disasters* (pp. 29–44). New York: Plenum.

Kidd, K. K., & Pauls, D. L. (1982). Genetic hypotheses for Tourette syndrome. In T. N. Chase & A. J. Friedhoff (Eds.), *Gilles de la Tourette syndrome* (pp. 243–249). New York: Raven Press.

King, N., Inglis, S., Jenkins, M., Myerson, N., & Ollendick, T. H. (1995). Test–retest reliability of the survey form of the Leyton Obsessional Inventory–Child Version. *Perceptual Motor Skills, 80,* 1200–1202.

Kringlen, E. (1965). Obsessional neurotics: A long-term follow-up. *British Journal of Psychiatry, 111,* 709–722.

Ladouceur, R., Rheaume, J., & Leger, E. (1996). *Cognitive change during cognitive treatment and behavioral treatment* of *checking behaviors*. Communication presented at the 30th Annual Convention of the Association for the Advancement of Behavior Therapy, New York, November 1996.

Ladouceur, R., Rheume, J., Freeston, M., Aublet, F., Jean, K., Lachance, S., Langlois, F., & de Pokomandy, K., (1995). Experimental manipulations of responsibility: an analogue test for models of obsessive-compulsive disorder. *Behaviour Research and Therapy, 33,* 937–946.

Last, C. G., Hersen, M., Kazdin, A. E., Orvaschel, H., & Perrin, S. (1991). Anxiety disorders in children and their families. *Archives of General Psychiatry, 48,* 928–934.

Lenane, M. C. (1989). Families and obsessive-compulsive disorder. In J. L. Rapoport (Ed.), *Obsessive–compulsive disorder in children and adolescents* (pp. 237–249). Washington, DC: American Psychiatric Press.

Lenane, M. C., Swedo, S. E., Leonard, H., Pauls, D. L., Sceery, W., & Rapoport, J. L. (1990). Psychiatric disorders in first degree relatives of children and adolescents with obsessive-compulsive disorder. *Journal of American Academy of Child Adolescent Psychiatry, 29,* 407–412.

Leonard, H. L. (1989). Childhood rituals and superstitions: developmental and cultural perspective. In J. L. Rapoport (Ed.), *Obsessive–compulsive disorder in children and adolescents* (pp. 289–309). Washington, DC: American Psychiatric Press.

Leonard, H. L., Goldberger, E. L., Rapoport, J. L., Cheslow, D. L., & Swedo, S. E. (1990). Childhood ruituals: Normal development of obsessive-compulsive symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 17–23.

Leonard, H. L., Lenane, M. C., Swedo, S. E., Rettew, D. C., Gershon, E. S., & Rapoport, J. L. (1992). Tics and Tourette's disorder: A 2 to 7 year follow-up of 54 obsessive-compulsive children. *American Journal of Psychiatry, 149,* 1244–1251.

Leonard, H. L., Swedo, S. E., Lenane, M. C., Rettew, D. C., Hamburger, S. D., Bartko, J., & Rapoport, J. L. (1993). A 2- to 7-year follow-up study of 54 obsessive-compulsive children and adolescents. *Archives of General Psychiatry, 50,* 429–439.

Leonard, H. L., Swedo, S. E., Rapoport, J. L., Koby, E. V., Lenane, M. C., Cheslow, D. L., & Hamburger, S. D. (1989). Treatment of obsessive-compulsive disorder with Clomipramine and Desipramine in children and adolescents. *Archives of General Psychiatry, 46,* 1088–1092.

Lewis, A. (1936). Problems of obsessional illness. *Proceedings of the Royal Society of Medicine, 29,* 325–336.

Lorenz, K. Z. (1966). *On aggression*. New York: Bantam Books.

Lorenz, K. Z. (1981). *The foundations of ethology*. New York: Springer.

Lucey, J. V., Costa, D. C., Blanes, T., Busatto, G. F., Pilowsky, L. S., Takei, N., Marks, I. M., Ell, P. J., & Kerwin, R. W. (1995). Regional cerebral blood flow in obsessive-compulsive disordered patients at rest: differential correlates with obsessive-compulsive and anxious avoidant-dimensions. *British Journal of Psychiatry, 167,* 629–634.

Luxenberg, J., Swedo, S. E., Flament, M., Friedland, R., Rapoport, J. L., & Rapoport, S. (1988). Neuroanatomical abnormalities in obsessive-compulsive disorder detected with quantitative X-ray computed tomography. *American Journal of Psychiatry, 145,* 1089–1093.

Macdonald, A., Murray, R. M., & Clifford, C. A. (1991). The contribution of heredity to obsessional disorder and personality: a review of family and twin study evidence. In K. Kendler, A. M. Tsuang, & M. Lyons (Eds.), *Genetic issues in psychosocial epidemiology* (pp. 191–212). NJ: Rutgers University Press.

Mahgoub, O. M., & Abdel-Hafeiz, H. B. (1991). Patterns of obsessive-compulsive disorder in Eastern Saudi-Arabia. *British Journal of Psychiatry, 158,* 840–842.

March, J. S., Mulle, K., & Herbel, B. (1994). Behavioral psychotherapy for children and adolescents with obsessive-compulsive disorder: An open trial of a new protocol-driven treatment package. *Journal of the American Academy of Child & Adolescent Psychiatry, 33,* 333–341.

Marks, I., Hodgson, R., & Rachman, S. (1975). Treatment of chronic obsessive-compulsive neurosis by in vivo exposure. *British Journal of Psychiatry, 127,* 349–364.

Marks, I. M., Stern, R. S., Mawson, D., Cobb, J., & McDonald, R. (1980). Clomipramine and exposure for obsessive-compulsive rituals. *British Journal of Psychiatry, 136,* 1–25.

Marks, I. M., & Tobena, A. (1990). Learning and unlearning fear: A clinical and evolutionary perspective. *Neuroscience and Biobehavioral Reviews, 14,* 365–384.

Martin, A., Pigott, T. A., Lalonde, F. M., Dalton, I., Dubbert, B., & Murphy, D. L. (1993). Lack of evidence for Huntington's disease-like cognitive dysfunction in obsessive-compulsive disorder. *Biological Psychiatry, 33,* 345–353.

Martin, M., Wiggs, C. L., Altemus, M., Rubenstein, C., & Murphy, D. L. (1995). Working memory as assessed by subject ordered tasks in patients with obsessive-compulsive disorder. *Journal of Clinical Experimental Neuropsychology, 17,* 786–792.

Max, J. E., Smith, W. L., Lindgren, S. D., Robin, D. A., Mattheis, P., Stierwalt, J., & Morrisey, M. (1995). Case study: Obsessive compulsive disorder after severe traumatic head injury in an adolescent. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 45–49.

McDonald, R., Marks, I. M., & Blizard, R. (1988). Quality assurance outcome in mental health care: A model for routine use in clinical settings. *Heath Trends, 20,* 111–114.

McDougle, C. J., Kresch, L. E., Goodman, W. K., Naylor, S. T., Volkmar, F. R., Cohen, D. J., & Price, L. H. (1995). A case-controlled study of repetitive thoughts and behaviour in adults with autistic disorder and obsessive-compulsive disorder. *American Journal of Psychiatry, 152, 772–777.*

McFall, M. E., & Wollersheim, J. P. (1979) Obsessive–compulsive neurosis: A cognitive behavioral formulation and approach to treatment. *Cognitive Therapy and Research, 3,* 333–348.

McGuire, P. K., Bench, C., Frith, C. D., Marks, I. M., Frackowiak, R. S. J., & Dolan, R. J. (1994). Functional anatomy of obsessive-compulsive phenomena. *British Journal of Psychiatry, 164,* 459–468.

McKay, D., Danyko, S., Neziroglu, F., & Yaryura-Tobias, J. A. (1995). Factor structure of the Yale–Brown Obsessive–Compulsive Scale: a two dimensional measure. *Behaviour Research and Therapy, 33,* 865–869.

McKeon, J., McGuffin, P., & Robinson, P. (1984). Obsessive–compulsive neurosis following head injury: a report of four cases. *British Journal of Psychiatry, 144,* 190–192.

Merkel, W. T., Pollard, C. A., Wiener, R. L., & Staebler, C. R. (1993). Perceived parental characteristics of patients with obsessive-compulsive disorder, depression, and panic disorder. *Child Psychiatry and Human development, 24,* 49–57.

Miguel, E. C., Coffey, B. J., Baer, L., Savage, C. R., Rauch, S. L., & Jenike, M. A. (1995a). Phenomenology of intentional repetitive behaviours in obsessive-compulsive disorder and Tourette's disorder. *Journal of Clinical Psychiatry, 56,* 246–255.

Miguel, E. C., Stein, M. C., Rauch, S. L., O'Sullivan, R. L., Stern, T. A., & Jenike, M. A. (1995b). Obsessive–compulsive disorder in patients with multiple sclerosis. *Journal of Neuropsychiatry of Clinical Neuroscience, 7,* 507–510.

Montgomery, S. (1994). Pharmacological treatment of obsessive-compulsive disorder. In E. Hollander, J. Zohar, D. Marazziti, & B. Olivier (Eds.), *Current insights in obsessive-compulsive disorder* (pp. 215–225). Chichester, UK: Wiley.

O'Sullivan, G., & Marks, I. (1990). Long-term outcome of phobic and obsessive-compulsive disorders after treatment. In R. Noyes, M. Roth, & G. D. Burrows (Eds.), *Handbook of anxiety, Vol. 4: The treatment of anxiety* (pp. 87–108). Amsterdam: Elsevier Science.

O'Sullivan, G., Noshirvani, H., Marks, I., Monteior, W., & Lelliott, P. (1991). Six year follow-up after exposure and clomipramine therapy for obsessive-compulsive disorder. *Journal of Clinical Psychiatry, 52,* 150–155.

Orley, J., & Wing, J. K., (1979). Psychiatric disorders in two African villages. *Archives of General Psychiatry, 36,* 513–520.

Pauls, D. L., Alsobrook, J. P., Goodman, W., Rasmussen, S., & Leckman, J. F. (1995). A family study of obsessive-compulsive disorder. *American Journal of Psychiatry, 152,* 76–84.

Pauls, D. L., & Leckman, J. (1986). The inheritance of Gilles de la Tourette syndrome and associated behaviors: Evidence for autosomal dominant transmission. *New England Journal of Medicine, 315,* 993–997.

Pauls, D. L., Raymond, C. L., & Robertson, M. (1991). The genetics of obsessive-compulsive disorder: a review. In J. Zohar, T. Insel, & S. Rasmussen (Eds.), *The psychobiology of obsessive-compulsive disorder* (pp. 89–100). New York: Springer.

Pauls, D. L., Towbin, K., Leckman, J., Zahner, G., & Cohen, D. (1986). Gilles de la Tourette syndrome and obsessive-compulsive disorder: Evidence supporting an etiological relationship. *Archives of General Psychiatry, 43,* 1180–1182.

Peterson, B. S. (1995). Neuroimaging in child and adolescent neuropsychiatric disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 1560–1576.

Pfeiffer, W. M. (1981). Culture-bound syndromes. In I. Al-Issa (Ed.) *Culture and psychopathology* (pp. 201–218). Baltimore: University Park Press.

Piacentini, J., Gitow, A., Jaffer, M., Graae, M. D., & Whitaker, M. D. (1994). Outpatient behavioral treatment of child and adolescent obsessive-compulsive disorder. *Journal of Anxiety Disorders, 8,* 277–289.

Piaget, J. (1937). *The construction of reality in the child.* Trans. M. Cook. New York: Basic Books, 1954.

Piaget, J. (1960). *The child's conception of physical causality.* London: Routledge and Kegan Paul.

Piccinelli, M., Pini, S., Bellantuono, C., & Wilkinson, G. (1995). Efficacy of drug treatment in obsessive-compulsive disorder: A meta-analytic review. *British Journal of Psychiatry, 166,* 424–443.

Pitman, R. K. (1987). A cybernetic model of obsessive-compulsive psychopathology. *Comprehensive Psychiatry, 28,* 334–343.

Pitman, R. K. (1993). Post-traumatic obsessive-compulsive disorder: A case study. *Comprehensive Psychiatry, 34,* 102–107.

Pollitt, J. (1957). Natural history of obsessional states—a study of 150 cases. *British Medical Journal, 26,* 194–198.

Price, R. A., Kidd, K. K., Cohen, D. J., Pauls, D. L., & Leckman J. F. (1985). A twin study of Tourette syndrome. *Archives of General Psychiatry, 42,* 815–820.

Prochaska, J., & DiClemente, C., (1982). Transtheoretical therapy: toward a more integrative model of change. *Psychotherapy: Theory, Research and Practice, 19,* 276–288.

Prochaska, J., DiClemente, C. & Norcross, J. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist, 47,* 1102–1114.

Rachman, S. (1974). Primary obsessional slowness. *Behaviour Research and Therapy, 12,* 9–18.

Rachman, S. (1993). Obsessions, responsibility and guilt. *Behaviour Research and Therapy, 31,* 149–154.

Rachman, S. J. (1985). An overview of clinical and research issues in obsessional-compulsive disorders. In M. Mavissakaliah, S. M. Turner, & L. Michelson (Eds.), *Obsessive–compulsive disorder: Psychological and pharmacological treatment* (pp. 1–47). New York: Plenum Press.

Rachman, S., & Hodgson, R. (1980). *Obsessions and compulsions.* Englewood Cliffs, NJ: Prentice-Hall.

Rapoport, J. L. (1990). Obsessive compulsive disorder and basal ganglia dysfunction. *Psychological Medicine, 20,* 465–469.

Rapoport, J. L, Leonard, H. L., Swedo, S. E., & Lenane, M. C. (1993). Obsessive compulsive disorder in children and adolescents: Issues in management. *Journal of Clinical Psychiatry, 54,* 27–29.

Rapoport, J. L., Swedo, S. E., & Leonard, H. L. (1992). Childhood Obsessive compulsive disorder. *Journal of Clinical Psychiatry, 56,* 11–16.

Rapoport, J. L., & Wise, S. (1988). Obsessive–compulsive disorder: Evidence for basal ganglia dysfunction. *Psychopharmacology Bulletin, 24,* 380–384.

Rasmussen, S. A., & Eisen, J. L. (1990). Epidemiology and clinical features of obsessive-compulsive disorder. In M. A. Jenike, L. Baer, & W. E. Minichiello (Eds.), *Obsessive compulsive disorders: Theory and management* (pp. 10–27). Chicago: Yearbook Medical Publishers.

Rasmussen, S. A., Eisen, J. L., & Pato, M. T. (1993). Current issues in the pharmacologic management of obsessive-compulsive disorder. *Journal of Clinical Psychiatry, 54,* 4–9.

Rasmussen, S. A., & Tsuang, M. T. (1986). DSM-III obsessive-compulsive disorder: clinical characteristics and family history. *American Journal of Psychiatry, 143,* 317–322.

Rauch, S. L., & Jenike, M. A. (1993). Neurobiological models of obsessive-compulsive disorder. *Psychosomatics, 34,* 20–32.

Reed, G. F. (1977). Obsessional personality disorder and remembering. *British Journal of Psychiatry, 130,* 184–185.

Reich, J. H., & Vasile, R. G. (1993). Effect of personality disorders on the treatment outcome of axis I conditions: An update. *Journal of Nervous and Mental Diseases, 181,* 475–484.

Rettew, D. C., Swedo, S. E., Leonard, H. L., Lenane, M. C., & Rapoport, J. L. (1992). Obsessions and compulsions across time in 79 children and adolescents with obsessive-compulsive disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 1050–1056.

Riddle, M. A., Scahill, L., King, R. A., Hardin, M. T., Anderson, G. M., Ort, S. I., Smith, J. C., Leckman, J. F., & Cohen, D. J. (1992). Double blind crossover trial of fluoxetine and placebo in children and adolescents with obsessive-compulsive disorder: a multicenter trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 1062–1069.

Riddle, M. A., Scahill, L., King, R., Hardin, M. T., Towbin, K. E., Ort, S. I., Leckman, J. F., & Cohen, D. J. (1990). Obsessive compulsive disorder in children and adolescents: Phenomenology and family history. *Journal of American Academy of Child Adolescent Psychiatry, 29,* 766–772.

Riggs, D. S., & Foa, E. B. (1993). Obsessive compulsive disorder. In D. H. Barlow (Ed.), *Clinical handbook of psychological disorders* (pp. 189–239). New York: Guildford Press.

Robins, L. N., Helzer, J. E., Weissman, M. M., Orvaschel, H., Gruenberg, E., Burke, J. D., & Regier, D. A. (1984). Life-time prevalence of specific psychiatric disorders in three sites. *Archives of General Psychiatry, 41,* 958–967.

Robinson, D., Wu, H., Munne, R. A., Ashtari, M., Alvir, J., Lerner, G., Koreen, A., Cole, K., & Bogerts, B. (1995). Reduced caudate nucleus volume in obsessive-compulsive disorder. *Archives of General Psychiatry, 52,* 393–398.

Rubenstein, C., Altmeus, M., Piggott, T., Hess, A., & Murphy, D. (1995). Symptom overlap between OCD and bulimia nervosa. *Journal of Anxiety Disorders, 9,* 1–9.

Rubin, R. T., Villanueva-Meyer, J., Ananth, J., Trajmar, P. G., & Mena, I. (1992). Regional xenon 133 cerebral blood flow and cerbral technetium ^{99m}Tc-HMPAO uptake in umedicated patients with obsessive-compulsive disorder and matched normal control subjects: determination by high resolution single-photon emission computed tomography. *Archives of General Psychiatry, 49,* 695–702.

Rudin, E. (1953). Ein Beitrag zur Frage der Zwangskrankenheit insebesondere inhrere hereditaren Beziehungen. *Archives Psychiatrische Nervenkrankheit, 91,* 14–54.

Russell, G., Szmukler, G., Dare, C., & Eisler, I. (1987). An evaluation of family therapy in anorexia nervosa and bulimia nervosa. *Archives of General Psychiatry, 44,* 1047–1056.

Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, health and behaviour*. London: Longmans.

Salkovskis, P. (1985). Obsessional–compulsive problems: cognitive-behavioural analysis. *Behaviour Research and Therapy, 23,* 571–583.

Salkovskis, P., & Warwick, H. (1985). Cognitive therapy of obsessive-compulsive disorder: Treating treatment failures. *Behavioural Psychotherapy, 13,* 243–255.

Salkovskis, P., & Westbrook, D. (1989). Behaviour therapy and obsessional ruminations: Can failure be turned into success? *Behaviour Research and Therapy, 27,* 149–160.

Santangelo, S. L., Pauls, D. L., Goldstein, J. M. Faraone, S. V. Tsuang, M. T., & Leckman, J. F. (1994). Tourette's syndrome: What are the influences of gender and comorbid obsessive-compulsive disorder? *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 794–804.

Schneier, F. R., Chin, S. J., Hollander, E., & Liebowitz, M. R. (1992). Fluoxetine in social phobia. *Journal of Clinical Psychopharmacology, 12,* 62–64.

Schut, A. J., & Castonguay, L. G. (1996). *Alexithymia in generalized anxiety disorder: The role of compulsive checking behaviors*. Poster presented at the 30th Annual Convention of the Association for the Advancement of Behavior Therapy, New York, November 1996.

Sharma, B. P. (1968). Obsessive–compulsive neurosis in Nepal. *Transcultural Research Review, 5,* 38–41.

Sher, K., Frost, R., & Otto, R. (1983). Cognitive deficits in compulsive checkers: An exploratory study. *Behaviour Research and Therapy, 21,* 357–363.

Silverman, W. K. (1991). Diagnostic reliability of anxiety disorders in children using structured interviews. *Journal of Anxiety Disorders, 5,* 105–124.

Silverman, W. K., & Nelles, W. B. (1988). The anxiety disorders interview schedule for children. *Journal of American Academy of Child and Adolescent Psychiatry, 27,* 772–778.

Solyom, L., Freeman, R. J., & Miles, J. E. (1982). A comparative psychometric study of anorexia nervosa and obsessive neurosis. *Canadian Journal of Psychiatry, 27,* 282–286.

Stanley, M. A., & Turner, S. M. (1995). Current status of pharmacological and behavioral treatment of obsessive-compulsive disorder. *Behavior Therapy, 26,* 163–186.

Stein, D. J., Spadaccini, E., & Hollander, E. (1995). Meta-analysis of pharmacotherapy trials for obsessive-compulsive disorder. *International Clinical Psychopharmacology, 10,* 11–18.

Steketee, G., Chambless, D. L., Tran, G. Q., Worden, H., & Gillis, M. M. (1996). Behavioral avoidance test for obsessive-compulsive disorder. *Behaviour Research and Therapy, 34,* 73–83.

Stern, R. S., & Cobb, J. P. (1978) Phenomenology of obsessive-compulsive neurosis. *British Journal of Psychiatry, 132,* 233–239.

Suess, L., & Halpern, M. S. (1989). Obsessive–compulsive disorder: a religious perspective. In J. L. Rapoport (Ed.), *Obsessive–compulsive disorder in children and adolescents* (pp. 311–325). Washington, DC: American Psychiatric Press.

Swedo, S. E., Leonard, H. L., & Rapoport, J. L., (1990). Childhood onset obsessive-compulsive disorder. In M. A. Jenike, L. Baer, & W. E. Minichiello (Eds.), *Obsessive-compulsive disorders: theory and management* (pp. 28–38). Chicago: Medcal Yearbook Publishers.

Swedo, S. E., Pietrini, P., Leonard, H. H., Schapiro, M. B., Rettew, D. C., Goldberger, E. L., Rapoport, S. I., Rapoport, J. L., & Grady, C. L. (1992). Cerebral glucose metabolism in childhood-onset obsessive-compulsive disorder. Revisualization during pharmacotherapy. *Archives of General Psychiatry, 49,* 690–694.

Swedo, S. E., Rapoport, J. L., Cheslow, D. L., Leonard, H. L., R. Ayoub, E. M., Hosier, D. M., & Wald, E. (1989). High prevalence of obsessive-compulsive symptoms in patients with Sydenham's chorea. *American Journal of Psychiatry, 146,* 246–249.

Swedo, S. E., Rapoport, J. L., Leonard, H., Lenane, M., & Cheslow, D. (1989). Obsessive–compulsive disorder in children and Adolescents: Clinical phenomenology of 70 consecutive cases. *Archives of General Psychiatry, 46,* 335–341

Tallis, F. (1995). *Understanding obsessions and compulsions*. London: Sheldon Press.

Tallis, F., & de Silva, P. (1992). *Worry and obsessional symptoms: A correlational analysis. Behaviour Research and Therapy, 30,* 103–105.

Thiel, A., Broocks, A., Ohlmeier, M., Jacoby, G., & Schussler, G. (1995). Obsessive–compulsive-disorder among patients with anorexia nervosa and bulimia nervosa. *American Journal of Psychiatry*, 152, 72–75.

Thomsen, P. H. (1992). Obsessive–compulsive disorder in adolescence. Differential diagnostic considerations in relation to schizophrenia and manic-depressive disorder: A comparison of phenomenology and sociodemographic characteristics. *Psychopathology, 25,* 301–310.

Thomsen, P. H. (1995a). Obsessive–compulsive disorder in children and adolescents: A 6-22 year follow-up study of social outcome. *European Child and Adolescent Psychiatry, 4,* 112–122.

Thomsen, P. H. (1995b). Obsessive–compulsive disorder in children and adolescents: A study of parental psychopathology and precipitating events in 20 consecutive Danish cases. *Psychopathology, 28,* 161–167.

Thomsen, P. H., & Jensen, J. (1991). Latent class analysis of organic aspects of obsessive-compulsive disorder in children and adolescents. *Acta Psychiatrica Scandinavica, 84,* 391–395.

Thomsen, P. H., & Mikkelsen, H. U. (1991). Children and adolescents with obsessive-compulsive disorder: the demographic and diagnostic characteristics of 61 Danish patients. *Acta Psychiatrica Scandinavica, 83,* 262–266.

Thomsen, P. H., & Mikkelsen, H. U. (1995). Course of obsessive-compulsive disorder in children and adolescents: A prospective follow-up study of 23 Danish cases. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 1432–1440.

Tinbergen, N. (1953). *Social behaviour in animals.* London: Chapman and Hall.

Torgerson, S. (1980). The oral, obsessive and hysterical personality-syndromes: A study of hereditary and environmental factors by means of the twin method. *Archives of General Psychiatry, 37,* 1272–1277.

Tynes, L., White, K., & Steketee, G. (1990). Toward a new nosology of obsessive-compulsive disorder. *Comprehensive Psychiatry, 31,* 465–480.

Valleni-Basile, L. A., Garrison, C. Z., Jackson, K. L., Waller, J. L., Mckeown, R. E., Addy, C. L., & Cuffe, S. P. (1994). Frequency of obsessive-compulsive disorder in a community sample of young adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 782–91.

Valleni-Basile, L. A., Garrison, C. Z., Jackson, K. L., Waller, J. L., McKeown, R. E., Addy, C. L., & Cuffe, S. P. (1995). Family and psychosocial predictors of obsessive-compulsive disorder in a community sample of young adolescents. *Journal of Child and Family Studies, 4,* 193–206.

Van Oppen, P., & Arntz, A. (1994). Cognitive therapy for obsessive-compulsive disorder. *Behaviour Research and Therapy, 32,* 79–87.

Van Oppen, P., de Haan, E., van Balkom, A., Spinhoven, P., Hooghuin, K., & van Dyck, R. (1995). Cognitive therapy and exposure in vivo in the treatment of obsessive-compulsive disorder. *Behaviour Research and Therapy, 33,* 379–390.

Van Oppen, P., Emmelkamp, P., van Balkom, J., & van Dyck, R. (1995). The sensitivity to change of measures for obsessive-compulsive disorder. *Journal of Anxiety Disorders, 9,* 241–248.

Vasey, M. W. (1993). Development and cognition in childhood anxiety: The example of worry. In T. H. Ollendick & R. J. Prinz (Eds.), *Advances in clinical child psychology* (Vol. 15, pp. 1–39). New York: Plenum.

Veale, D. (1993). Classification and treatment of obsessional slowness. *British Journal of Psychiatry, 162,* 198–203.

Vitiello, B., Spreat, S., & Behar, D. (1989). Obsessive-compulsive disorder in mentally retarded patients. *Journal of Nervous and Mental Disease, 177,* 232–236.

von Economo, C. (1931). *Encephalitis lethargica: its sequelae and treatment.* Oxford: Oxford University Press.

Vygotsky, L. (1934). *Thought and Language,* Trans. E. Hanfmann & G. Vakar. Cambridge, MA: MIT Press, 1962.

Warren, W. (1965). A study of psychiatric inpatients and the outcome six or more years later: II. The follow-up study. *Journal of Child Psychology and Psychiatry, 6,* 141–160.

Wever, C., & Phillips, N. (1994). *The secret problem.* Sydney, NSW: Shrink-Rap Press.

Wever, C., & Rey, J. (in press). Outcome of a combined treatment in juvenile obsessive-compulsive disorder. *Australian and New Zealand Journal of Psychiatry.*

White, M., & Epston, D. (1990). *Narrative means to therapeutic ends.* New York: Norton.

Winslow, J. T., & Insel, T. (1991). Neuroethological models of obsessive-compulsive disorder. In J. Zohar, T. Insel, & S. Rasmussen (Eds.), *The psychobiology of obsessive-compulsive disorder* (pp. 208–226). New York: Springer.

Wolff, R. P., & Wolff, L. S. (1991). Assessment and treatment of obsessive-compulsive disorder in children. *Behavior Modification, 15,* 372–393.

Woodruff, R., & Pitts, F. (1964). Monozygotic twins with obsessional illness. *American Journal of Psychiatry, 120,* 1075–1080.

Woody, S. R, Steketee, G., & Chambless, D. L. (1995). Reliability and validity of the Yale–Brown Obsessive–Compulsive Scale. *Behaviour Research and Therapy, 33,* 597–605.

World Health Organization (1988). *Mental, behavioral, and developmental disorders: Clinical descriptives and diagnostic guidelines (ICD-10).* Geneva, Switzerland: World Health Organization, Division of Mental Health.

Young, J., Fenton, F., & Lader, M. (1971). The inheritance of neurotic traits. *British Journal of Psychiatry, 119,* 393–398.

Zeitlin, H. (1986). *The natural history of psychiatric disorder in children.* Oxford: Oxford University Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.16
Post-traumatic Stress Disorder

SEAN G. PERRIN, PATRICK A. SMITH, and WILLIAM YULE
Institute of Psychiatry, London, UK

5.16.1 INTRODUCTION 393

5.16.2 PTSD 394

5.16.2.1 Phenomenology of Stress Reactions in Children and Adolescents 394
5.16.2.1.1 Age and gender-related effects of trauma on children 395
5.16.2.1.2 Cultural factors in the expression of PTSD 396
5.16.2.2 Diagnostic Features 396
5.16.2.2.1 Prevalence 398
5.16.2.2.2 Etiology 400
5.16.2.2.3 A learning theory account of PTSD 400
5.16.2.2.4 An emotional processing model 400
5.16.2.2.5 A neurobiological model of PTSD 401
5.16.2.2.6 Characteristics of the trauma and other risk factors for PTSD 402
5.16.2.2.7 Developmental course 403

5.16.3 CONCEPTUALIZATION AND CLINICAL FORMULATION 404

5.16.3.1 Multimodal Assessment 404
5.16.3.1.1 Interviewing the child 404
5.16.3.1.2 Childhood measures of post-traumatic distress 405
5.16.3.1.3 Interviewing the parent 406
5.16.3.1.4 Discrepancies in parent and child reports 406
5.16.3.2 Treatment Formulation 406

5.16.4 TREATMENT 407

5.16.4.1 Psychosocial Approaches 407
5.16.4.1.1 Preventative measures and contingency planning 407
5.16.4.1.2 Debriefing and crisis intervention 408
5.16.4.1.3 Group treatment 409
5.16.4.1.4 Individual treatment 410
5.16.4.1.5 Community approaches 411
5.16.4.2 Pharmacological Approaches 412
5.16.4.3 Efficacy 413

5.16.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 413

5.16.6 SUMMARY 414

5.16.7 REFERENCES 414

5.16.1 INTRODUCTION

The period from the 1970s to the 1990s has shown that children react to life-threatening stressors with various forms of distress, anxiety, fears, depression, and bereavement reactions, as well as post-traumatic stress disorder (PTSD). The form of PTSD varies somewhat according to the age of the child and, sadly, the internal distress often goes unrecognized for long periods by parents and teachers. Indeed, one reason for the initial doubt whether PTSD occurred in children at all was that few investigators had asked children directly how they were affected.

Although post-traumatic stress reactions have been known to occur for centuries, our understanding of people's reactions to life-threatening experiences have been advanced in the aftermath of the major wars of the twentieth century. It was not until the persisting problems of Vietnam veterans were better documented that it was realized that three major groups of symptoms exist: distressing recurring recollections of traumatic events, avoidance of trauma-related stimuli, and a range of symptoms associated with increased arousal. These came to be known as PTSD (American Psychiatric Association, 1980).

Since its inclusion in the third edition of the *Diagnostic and statistical manual for mental disorders* (*DSM-III*) (American Psychiatric Association, 1980), PTSD has been the subject of much debate regarding its validity, etiology, placement among the anxiety disorders, and even its status as a "true" mental disorder (Yehuda & McFarlane, 1995). Nowhere has the controversy over the validity of PTSD as a diagnostic entity been greater than in the child trauma literature (Pynoos, 1994). This is due in part to the observation that the reactions commonly observed in traumatized children are so much wider than the narrow confines described under PTSD in the various revisions of the *DSM* and the *International classification of diseases* (*ICD-10*) (World Health Organization, 1992; Yule, 1994). Moreover, some have suggested that children are so resilient that even when they experience an acutely stressful event, they develop, at worst, some sort of transient reaction disorder (Garmezy, 1984; Garmezy & Rutter, 1985).

Despite the controversy, PTSD is a very useful framework for describing and understanding childrens' reactions to a wide range of life-threatening experiences. Research since the mid-1970s suggests that significant and often chronic symptoms of PTSD are found among children following exposure to war (Arroyo & Eth, 1985; Kinzie, Sack, Angell, Manson, & Rath, 1986; Saigh, 1989a, 1991), natural disaster (Earls, Smith, Reich, & Young, 1988; Frederick, 1985a; Galante & Foa, 1986; Goenjian; 1993; Green et al., 1991; Lonigan, Shannon, Finch, Daugherty, & Saylor, 1991; McFarlane, 1987; Milne, 1977; Newman, 1976), manmade and transportation disasters (Handford et al., 1986; Parry-Jones, 1996; Terr, 1979, 1983; Yule; 1992), violence and violent crime (Malmquist, 1986; Pynoos & Eth, 1988; Schwarz & Kowalski, 1991), sexual abuse (Goodwin, 1988; McLeer, Deblinger, Henry, & Orvaschel, 1992; Wolfe, Gentile, & Wolfe; 1989), and serious medical illness (Nir, 1984; Stuber, Nader, Yashuda, Pynoos, & Cohen, 1991). Still, there remain considerable gaps in our understanding of the disorder in children, particularly in regard to its developmental, neurobiological, and physiological manifestations across those of various ages and abilities.

In this chapter, the literature on childhood PTSD is highlighted and reviewed with a view to increasing the reader's understanding of the phenomenology, assessment, and treatment of the disorder. However, an exhaustive review of the issues pertinent to the manifestation of PTSD following specific types of traumas (e.g., war, torture, natural disaster, and sexual assault) is beyond the scope of this chapter. The reader should be aware that significant differences exist in the clinical presentation and treatment of children with specific trauma histories.

5.16.2 PTSD

5.16.2.1 Phenomenology of Stress Reactions in Children and Adolescents

From their experience working with child and adolescent survivors of the "Herald of Free Enterprise" car ferry sinking (Yule & Williams, 1990), the sinking of the cruise ship, "Jupiter" (Yule, 1992; Yule, Udwin, & Murdoch, 1990), war-affected children of the former Yugoslavia (Smith, Yule, Perrin, & Schwarz, 1996), and childhood survivors of serious road traffic accidents in Britain (Perrin, Yule, & Smith, 1996b; Yule, Perrin, & Smith, 1996), Yule and colleagues have noted a number of reactions that are common in children exposed to traumatic events.

Most children are troubled by repetitive, intrusive thoughts about the trauma. Such thoughts can occur at any time, but particularly when the children are otherwise quiet, as when they are trying to fall asleep. At other times, the thoughts and vivid recollections are triggered off by reminders in their environment. Sleep disturbances are very common, particularly within the first few weeks of the event. Fears of the dark, bad dreams, nightmares, and waking through the sleep cycle are widespread. Separation difficulties are frequent, even among teenagers. For the first few days, children may not want to let their parents out of their sight, even reverting to sleeping in the parents' bed. Many children become much more irritable and angry with parents and peers alike.

Although child survivors often experience a pressure to talk about their experiences, paradoxically they also find it very difficult to talk with parents and peers. Often they do not want to upset the adults, and so parents may not be aware of the full extent of their childrens' suffering. Peers may hold back from asking what happened lest they upset the child further; the survivor often takes this as a subtle form of rejection.

Children report a number of cognitive changes. Many experience difficulties in concentration, especially in school work. Others report memory problems, both in mastering new material and in remembering previously learned skills such as reading music. They become very alert to danger in their environment, being adversely affected by reports of other related traumatic events. Survivors often learn that life is quite fragile, particularly if they sustained any permanent form of injury or witnessed the death of another. This can lead to a sense of a foreshortened future, often expressed as a loss of thoughts about the future. Life priorities frequently change in response to trauma, with some feeling that they should live each day to the full and not plan for the future. Others realize that they have been overconcerned with materialistic or petty matters and resolve to rethink their values, frequently taking on the image of themselves as a helper to others. Such ways of thinking may reflect a change in the child's "assumptive world" (cf., Janoff-Bulman, 1992).

Not surprisingly, many children develop fears associated with specific aspects of the traumatic event(s), with phobic levels of avoidance to trauma-related stimuli and reminders. Many experience "survivor guilt" about others dying or being seriously injured, or think they should have done more to help others to survive. Adolescent survivors report significantly high levels of depression, some becoming clinically depressed, having suicidal thoughts, and taking overdoses in the year after a disaster. A significant number become very anxious after the trauma, although the appearance of panic attacks is sometimes considerably delayed.

In summary, children and adolescents surviving a life threatening trauma show a wide range of symptoms which tend to cluster around symptoms of re-experiencing, avoidance of trauma-related cues and emotions, and increased physiological arousal. There may be considerable comorbidity with depression, generalized anxiety, or pathological grief reactions as well.

5.16.2.1.1 Age and gender-related effects of trauma on children

While there have been relatively few systematic studies examining the impact of trauma at different stages of development, there do appear to be important age-related differences in the expression of PTSD (see Lyons, 1987 and Terr, 1991 for reviews). Schwarz and Perry (1994) have suggested that infants and very young children are more sensitive to emotional states and behavior of parental caregivers than to assessments of danger. As such, a trauma to either the child or primary caregiver may result in global dysfunction, apathy and failure to thrive, excessive crying, eating and sleeping, psychophysiological lability, and overstimulation (Schwarz & Perry, 1994).

Among toddlers and preschool children, high levels of dependency (clinginess) and separation anxiety are frequent, with regressive behaviors such as bed-wetting, baby-talk, and thumb-sucking also observed. Diarrhea and frequent somatic complaints may be noted, as well as irritability, sleep problems, and fearfulness. Some children may show delays in cognitive and language development (Law & Conway, 1992) and mutism (Eth & Pynoos, 1985). Re-enactment of the trauma through play and difficulty with verbal descriptions of the trauma are common. Many preschool children show a loss of interest in activities and hobbies that previously gave them pleasure. Although some have suggested that emotional numbing may be rare in this age group, this may be an artifact of the difficulty researchers have found eliciting evidence of such symptoms in young children (Frederick, 1985a).

In reviewing data from the Buffalo Creek Disaster, Green et al. (1991) noted that preschoolers had fewer PTSD symptoms overall, and in particular avoidance symptoms, than older children. They hypothesized that it may be more difficult to observe avoidant symptoms (as well as dissociation and emotional numbing) in this age group. Consistent with Schwarz and Perry's hypothesis (1994), the preschool child's perception of the dangerousness of the trauma (i.e., life threat) was not correlated with PTSD symptoms, while life threat was positively associated with PTSD symptoms in school-aged children and adolescents (Green et al.).

Relative to preschool children, those of school age may show an increased frequency of trauma-related nightmares, preoccupation with the trauma, vacillation between withdrawal and aggressive outbursts, and avoidant behaviors (Lyons, 1987). Somatic complaints may be less frequent, and this may be related to an increased verbal ability to describe the trauma. Children of this age can begin to appreciate the irreversibility of tragic events and may exhibit more overt signs of depressed mood and hopelessness.

Given that concentration difficulties are characteristic of PTSD (American Psychiatric Association, 1994), one might expect that children's school work would suffer after a traumatic event. However, Martin and Little (1986) found no differences in scholastic attainment across three groups of 9–10 year-old children with varying degrees of exposure, two months after a tornado. The authors speculated that this may have been the result of the community pulling together in the face of the trauma, and because children were relocated to new schools where teachers were instructed to discuss the children's experiences of the tornado (Martin & Little, 1986).

Among adolescents, there is an increased frequency of adult-like PTSD symptoms, problems with compliance and withdrawal, aggressiveness, substance abuse, and increased feelings of guilt. Problems of identity, eating, self-mutilation, and personality may also be noted (Schwarz & Perry, 1994). In contrast to younger children, there is generally an absence of traumatic play. In addition, there is clearer evidence of impaired academic performance post-trauma in adolescents. Girls who survived the sinking of the Jupiter who were above average in performance during the three years before the sinking, plummeted significantly to merely average levels one year after the disaster (Yule, 1992). Two years after the accident, the 16-year old survivors still performed less well in national school exams, although the difference with pretrauma performance was no longer as marked (Tsui, 1990).

With respect to gender, girls consistently score higher across measures of post-traumatic distress, anxiety, and depression than boys in a number of studies, including children exposed to a shipping disaster (Yule, 1992), a hurricane (Lonigan et al., 1991), flood (Green et al., 1991), an earthquake (Pynoos et al., 1993), and war in former Yugoslavia (Smith et al., 1996). One study (Burke, Moccia, Borus, & Burns, 1986) did find boys' distress to be higher than girls five months after a flood; however, this pattern reversed at a follow-up assessment five months later.

5.16.2.1.2 Cultural factors in the expression of PTSD

There has been some debate over the appropriateness of using PTSD to characterize the traumatic reactions of children from non-Western countries (see Eisenberg, 1996 and Summerfield, 1996). Unfortunately, there are few empirical investigations from which to draw firm conclusions. Studies of children's fears more generally suggest that the most common fears are the same across different countries and cultures (see Fonesca, Yule, & Erol, 1994 for a review). Indeed, the current literature suggests that there are remarkably few psychiatric disorders which are culture-specific (Leff & Bhugura, 1993). However, the question remains whether the traumatic reactions observed in children across different cultures are more similar than different and when such reactions should be considered abnormal or pathological.

Working with a core set of self-report instruments to screen children in exceptionally difficult circumstances, Yule, Pynoos, and others have begun to study this issue more directly. The Impact of Events Scale (IES) (Horowitz, Wilner, & Alvarez, 1979) has been modified for use with children in both civilian and war situations. Examining data from studies of war affected children in Iraq, Kuwait, and the former Yugoslavia, it was found that the IES factor structures obtained from children in Iraq (Dyregrov & Yule, 1995) and Yugoslavia (Perrin, Smith, Schwartz, & Yule 1996a) were identical to that obtained from British children who had survived the Jupiter shipping disaster. In addition, the number of war traumas reported by children from Yugoslavia was positively correlated with levels of distress on translated versions of the IES, Birleson Depression Inventory (Birleson, 1981), and Revised Children's Manifest Anxiety Scale (Reynolds & Richmond, 1978) (Perrin et al., 1996a). Moreover, this battery of instruments proved sensitive to pre–post-treatment changes in Yugoslavian children who participated in cognitive behavioral treatment (Dyregrov & Yule, 1995).

Finally, Pynoos and colleagues observed that child survivors of the Armenian earthquake not only show the classic symptoms of PTSD as measured with structured interview and self-report questionnaires, but also show greatly increased salivary cortisol levels and disturbed basal heart rates (Pynoos et al., 1993). The presence of these physiological markers of hyperarousal, which are presumably less affected by cultural influences than scores on translated self-report inventories, lends considerable support to the argument for PTSD symptoms manifesting across cultures. Of course, assessments and interventions should be culture-sensitive. As suggested by Parry-Jones (1996), labeling clinically significant stress reactions need not be stigmatizing and indeed may ensure that appropriate treatment services are offered.

5.16.2.2 Diagnostic Features

PTSD first entered the diagnostic nomenclature with its appearance in *DSM-III* (American

Psychiatric Association, 1980), appearing in the anxiety disorders section. The criteria were based largely on a then emerging body of research with adult trauma survivors, particularly combat veterans. Under *DSM-III*, PTSD was characterized by: (i) the experience of a traumatic event outside the range of usual human experience; (ii) distressing and recurring recollections of the traumatic event; (iii) numbing of responsiveness and/or reduced involvement with the external world; and (iv) at least two of a variety of symptoms including survivor guilt, sleep disturbances, hyperarousal, avoidance of trauma-related stimuli, and impaired concentration (Lyons, 1987). However, few and vague descriptions of the manifestations of PTSD in children of different ages were noted, owing to the paucity of methodologically sound research available on children's reactions to traumatic events.

The third revised edition of the *DSM* (*DSM-III-R*) (American Psychiatric Association, 1987) was the first to note the occurrence of PTSD in children but the specific criteria remained largely unchanged. Rather, the revisions involved a reorganization of the existing symptoms under re-experiencing, numbing and avoidance, and increased arousal categories (Lyons, 1987). However, there were two exceptions. First, a "sense of a foreshortened future" was added based on research with the children of the Chowchilla bus kidnaping (Lyons, 1987). Second, and unfortunately, "survival guilt" was removed, although this symptom is found frequently in both children and adults presenting with PTSD (Yule, 1994).

Under *DSM-IV* (American Psychiatric Association, 1994), the requirements for meeting a diagnosis of PTSD have changed considerably from its predecessors. The *DSM-IV* criteria for PTSD are as follows:

(i) The person has been exposed to a traumatic event in which both of the following have been present:

(a) The person has experienced, witnessed, or been confronted with an event or events that involve actual or threatened death or serious injury, or a threat to the physical integrity of oneself or others.

(b) The person's response involved intense fear, helplessness, or horror. Note: In children, it may be expressed by disorganized or agitated behavior.

(ii) The traumatic event is persistently re-experienced in at least one of the following ways:

(a) Recurrent and intrusive distressing recollections of the event, including images, thoughts, or perceptions. Note: In young children, repetitive play may occur in which themes or aspects of the trauma are expressed.

(b) Recurrent distressing dreams of the event. Note: In children, there may be frightening dreams without recognizable content.

(c) Acting or feeling as if the traumatic event were recurring (includes a sense of reliving the experience, illusions, hallucinations, and dissociative flashback episodes, including those that occur upon awakening or when intoxicated). Note: In young children, trauma specific re-enactment may occur.

(d) Intense psychological distress at exposure to internal or external cues that symbolize or resemble an aspect of the traumatic event.

(e) Physiological reactivity upon exposure to internal or external cues that symbolize or resemble an aspect of the traumatic event.

(iii) Persistent avoidance of stimuli associated with the trauma and numbing of general responsiveness (not present before the trauma), as indicated by at least three of the following:

(a) Efforts to avoid thoughts, feelings, or conversations associated with the trauma.

(b) Efforts to avoid activities, places, or people that arouse recollections of the trauma.

(c) Inability to recall an important aspect of the trauma.

(d) Markedly diminished interest or participation in significant activities.

(e) Feelings of detachment or estrangement from others.

(f) Restricted range of affect (e.g., unable to have loving feelings).

(g) Sense of a foreshortened future (e.g., does not expect to have a career, marriage, children, or a normal life span).

(iv) Persistent symptoms of increased arousal (not present before the trauma), as indicated by at least two of the following:

(a) Difficulty falling or staying asleep.

(b) Irritability or outbursts of anger.

(c) Difficulty concentrating.

(d) Hypervigilance.

(e) Exaggerated startle response.

(v) Duration of the disturbance (symptoms in (ii), (iii), (iv) are more than one month)

(vi) The disturbance causes clinically significant distress or impairment in social, occupational, or other important areas of functioning.

Specify if:

Acute: If duration of symptoms is less than three months.

Chronic: If duration of symptoms is three months or more.

With delayed onset: Onset of symptoms at least six months after the stressor.

In particular, criterion (i) has been clarified greatly such that more emphasis is placed on the threatening nature of the traumatic event to the person rather than it being outside the range of normal human experience. Also, the ways in which several symptoms manifest themselves in children are defined more clearly. Similar to *DSM-III*, *DSM-IV* notes that children's re-experiencing can be expressed as repetitive and trauma-thematic play, nightmares with/without traumatic content, and re-enactment of the trauma. Interestingly, the fourth edition no longer specifies "loss of previously acquired developmental skills" as a potential avoidance item and this may affect case identification (Ribbe, Lipovsky, & Freedy, 1995).

The inclusion of PTSD among the anxiety disorders in the *DSM* has led to considerable controversy among trauma and anxiety researchers alike (cf., Pynoos, 1994; Saigh, 1989b). While characterized by predominant symptoms of fear, avoidance, and hyperarousal, PTSD differs from the other anxiety disorders in that it requires a "putative" event (i.e., the trauma). Thus, there is an assumed causal relationship between an external event and the onset of the disorder. In this way, PTSD more resembles an adjustment disorder. Indeed, the notion that PTSD is "a normal reaction to an abnormal situation" has led some to question whether it is in fact a psychiatric disorder at all (cf., O'Donohue & Elliot; 1992).

PTSD also differs from the anxiety disorders in other important ways. Foa, Steketee, & Olasov-Rothbaum (1989) observed that trauma giving rise to PTSD violated more of the survivor's safety assumptions than did events leading to other forms of anxiety. Also, there was a much greater likelihood of fear generalization and subjective re-experiencing of the anxiety-eliciting event in PTSD than other anxiety disorders (Foa et al., 1989). Some have suggested that it is this internal, subjective experience of the traumatic event that most seems to set PTSD apart from the other anxiety disorders (Jones & Barlow, 1992; Keppel-Benson & Ollendick; 1993). Others have argued that it is the psychogenic amnesia and numbing of general responsiveness which separate PTSD clearly from anxiety, and places it more appropriately among the dissociative disorders. Still others (Pitman et al., 1990) have pointed out that exposure to trauma cues in combat veterans was just as likely to produce responses of anger and grief than fear, and thus PTSD might arguably be placed among the affective disorders.

To add to the controversy, it is widely recognized that the PTSD criteria, as currently defined, poorly reflect the wide variety of major stress responses reported in the literature. Thus, it can be argued that many seriously traumatized individuals might not meet criteria for PTSD under the *DSM*. As mentioned above, "survivor guilt" and grief reactions have largely been lost in each subsequent revision since *DSM-III*. Nor is there any mention in *DSM-IV* of the profound sexual, interpersonal, and identity problems often found in survivors of repeated sexual assault.

This has led some to ask whether a spectrum of disorders, including PTSD, should be delineated to capture more accurately the range of traumatic responses seen by clinicians and researchers (Kolb, 1988). In this regard, some have argued for an additional grouping to be recognized—Disorders of Extreme Stress Not Otherwise Specified (DESNOS) (Herman, 1992). The distinction between PTSD and DESNOS is similar to the distinction drawn by Terr (1991) between Type I (characterized by intrusive recollections) and Type II (characterized by denial, numbing, and dissociation) PTSD in children (Yule, 1994).

PTSD is also included in the tenth edition of the *International classification of diseases* (*ICD-10*) (World Health Organization, 1992). However, its inclusion among the adjustment disorders has avoided the controversy surrounding PTSD in the *DSM*. According to *ICD-10*: "[PTSD] arises as a delayed or protracted response to a stressful event or situation ... of an exceptionally threatening or catastrophic nature, which is likely to cause pervasive distress in almost anyone (e.g., natural or man made disaster, combat, serious accident, witnessing the violent death of others, or being a victim of torture, terrorism, rape, or other crime)" (World Health Organization, 1992, p. 147).

ICD-10 lists symptoms similar to those in *DSM-III-R*, but emphasizes the presence of repetitive, intrusive memories characteristic of the disorder rather than the avoidance symptoms. Specifically, the manual states: "A conspicuous emotional detachment, numbing of feeling, and avoidance of stimuli that might arouse recollections of the trauma are often present but *are not essential* for the diagnosis. The autonomic disturbances, mood disorder, and behavioral abnormalities are all contributory to the diagnosis but not of prime importance" (World Health Organization, 1992, p. 148). It is also worth noting that the *ICD-10* states that PTSD should only be diagnosed when it arises within six months of a major trauma, whereas *DSM-IV* allows the diagnosis to be made when onset occurs after six months post-trauma (as noted by the specification Delayed Onset (American Psychiatric Association, 1994).

5.16.2.2.1 Prevalence

Epidemiological investigations of adults in the general population estimate a lifetime prevalence of PTSD between 1 and 9.2% (see Fairbank, Schlenger, Saigh, & Davidson, 1995 for a review of epidemiological investigations of PTSD). The prevalence of PTSD among children in the community is unknown. There have been no epidemiological investigations of the disorder to date. However, there have been several studies of PTSD in children exposed to war, natural disaster, violent crime, and sexual abuse ("at-risk" groups).

Prevalence rates from studies of at-risk children vary widely (from 0 to 100%). This extreme variability is consistent with findings from studies of at-risk adults (Fairbank et al., 1995), and is reflected in the estimates given for at-risk groups in *DSM-IV* (3–58%) (American Psychological Association, 1994). Therefore, the following studies must be interpreted with caution. Specific diagnostic criteria for PTSD were not widely available until 1980. The term "psychic trauma" which included extreme fear reactions, grief, agitation, and depression was more often found in the literature. Also, few, if any, reliable screening instruments for childhood PTSD had been developed which severely undermined reliable estimation of the prevalence of PTSD (Galante & Foa; 1986; Yule & Williams, 1990).

In one of the earliest studies of at-risk children, Terr (1979) noted that the incidence of psychic trauma was 100% among 26 children involved in the Chowchilla bus-kidnaping. This rate was independent of the child's background and developmental history (Terr, 1979). Interestingly, in a separate study, the prevalence of "extreme fright" and psychic trauma was 40% in an age, race, and sex-matched "normal" control group chosen from the same school district as the kidnap victims (Terr, 1983), an unexpectedly high rate among so-called "normal" school-aged children.

Two additional studies provide estimates for PTSD among children exposed to civilian acts of violence. Pynoos et al. (1987) examined the occurrence of PTSD in 159 children one month after an attack by a sniper on their school playground. Seventy-seven percent of the children under direct threat (i.e., those on the playground) had moderate or severe PTSD as measured by the PTSD Reaction Index. Moderate or severe PTSD was also high (67%) among children not exposed to the sniper but at the school on the same day (Pynoos et al.).

More recently, Schwarz and Kowalski (1991) assessed 64 preadolescent school children six months after a shooting spree by a woman in their school. Prevalence rates were presented separately for the three diagnostic categorizations of PTSD under *DSM-III*, *DSM-III-R*, and *DSM-IV*, and the PTSD Reaction Index. Using "conservative" symptom thresholds (i.e., symptoms occurring "much" or "most of the time"), the prevalence of PTSD under *DSM-III* was 16%. By contrast, the prevalence of PTSD under *DSM-III-R* was only 8%, and 9% under *DSM-IV* (Schwarz & Kowalski, 1991). Using "liberal" symptom threshold levels (i.e., symptoms occurring at least "a little of the time"), prevalence rates ranged from 91% under *DSM-III* to 50% for *DSM-III-R*, and 26% for *DSM-IV*. This study underlines the difficulty with comparing PTSD prevalence rates across studies which use different versions of the *DSM*. Arguably, it also points out the risk of not identifying true cases of PTSD when using the subsequent revisions of the criteria under *DSM-III-R* and *DSM-IV*.

Several investigations have examined the prevalence of PTSD among children exposed to war. Kinzie et al. (1986) conducted one of the first investigations using operationally defined diagnostic criteria to study the effects of war on 40 Cambodian refugees living in Oregon. Nearly four years after leaving refugee camps in Cambodia as adolescents, 50% met current *DSM-III* criteria for PTSD based on structured clinical interviews. More severe psychiatric effects were found in those who had been separated from family members (Kinzie et al., 1986). Sack et al. (1994) reported much lower rates (lifetime = 21.5%; current = 18.0%) in 200 Cambodian refugees in the United States. Similarly, in a study of 840 war-affected children from Lebanon, Saigh (1991) observed a much lower rate of current PTSD (32.5%). Interestingly, levels of PTSD were consistent across routes of exposure (direct, observation, and indirect) (Saigh, 1991).

Several studies have investigated the prevalence of traumatic reactions in children following natural and manmade disasters. Handford et al. (1986) found higher than normal anxiety scores on children and adolescents from 35 households 1.5 years after the "Three Mile Island" nuclear disaster. None of the children, however, met criteria for a diagnosis of PTSD based on structured interviews. Similarly, Earls et al. (1988), using structured diagnostic interviews, found no cases of PTSD in 32 children and adolescents one year after a flash flood, although many met several criteria for the disorder.

By contrast, McFarlane, Policansky, and Irwin (1987) found a high frequency of PTSD symptom clusters (94%) obtained from parent and teacher versions of the Rutter Scales in 808 children two months after large-scale brushfires in Australia. Similarly, Milgram, Toubiana, Klingman, Raviv, and Goldstein (1988) observed a 40% prevalence rate for PTSD in children one month after a bus accident. Bradburn (1991) reported prevalence rates of 27% for moderate PTSD and 36% for mild PTSD in 22 children aged 10–12 years, some 6–8 months after the San Francisco earthquake. Of 179 children aged 2–15 years who were examined two years after the Buffalo Creek disaster, 37% received probable PTSD diagnoses based on retrospective examination of records (Greene et al., 1991).

A number of large-scale investigations of PTSD in children were conducted following Hurricane Hugo which struck South Carolina in 1989. Based on self-report data, Lonigan et al. (1991) reported PTSD prevalence rates of 5.06% in a no-exposure group, 10.35% in a mild, 15.54% in a moderate, and 28.95% in a high-exposure group of hurricane survivors. The overall prevalence rate for PTSD in the sample was 5% ($n = 5687$) nearly three months after the hurricane (Shannon, Lonigan, Finch, & Taylor, 1994). Similarly, Garrison, Weinrich, Hardin, Weinrich, and Wang (1993) reported current prevalence rates ranging from 1.5 to 6.2% one year after the hurricane in 11–17 year-olds.

From the above studies it can be seen that reliable estimation of the prevalence of PTSD in children and adolescents is extremely difficult, particularly from data collected before the operationalizing of PTSD in *DSM*. Moreover, there is considerable variability in prevalence estimates depending upon the age of the child, time since the trauma, assessment methods used, and under which version of *DSM* caseness was established.

5.16.2.2.2 Etiology

In contrast to most of the phenomenological classifications in *DSM-IV* (American Psychiatric Association, 1994) and *ICD-10* (World Health Organization, 1992), PTSD is one of the few psychiatric disorders which specifies an etiological agent. However, *DSM-IV* also recognizes that, at worst, only slightly more than half (58%) of those exposed to a trauma go on to develop PTSD (American Psychiatric Association, 1994). Thus, while trauma may be a necessary agent in the development of PTSD, it is clearly not a sufficient one (Yehuda & McFarlane, 1995). It is with this background that etiological theories have emerged to explain the development of PTSD since the early 1980s. A complete review of the etiological theories of PTSD is beyond the scope of this chapter and the interested reader should consult Foa et al. (1989) and Friedman, Charney, and Deutch (1995) for details. This chapter focuses on three models which have received considerable attention from researchers and for which there is limited empirical support or ongoing investigation. However, it should be noted that none of the etiological models put forward to date can explain adequately the full range of symptoms seen in PTSD (Yehuda & McFarlane, 1995).

5.16.2.2.3 A learning theory account of PTSD

Numerous researchers have proposed the use of Mowrer's (1960) two-factor learning theory to explain the development of PTSD (see Foa et al., 1989 for a review). Mowrer's is a conditioning model, incorporating elements of both Pavlovian and operant conditioning. As applied to PTSD, the traumatic event is an unconditioned stimulus that elicits certain involuntary responses: (i) orientation to the threat cue; (ii) autonomic arousal; (iii) fear; and (iv) the behavioral "flight or fight' response (Southwick, Bremner, Krystal, & Charney, 1994). Because of their proximity to the threatening event (the unconditioned stimulus (UCS)), previously neutral stimuli become conditioned stimuli (CS), capable of eliciting the same involuntary responses (conditioned response (CR)) as the trauma itself. For example, a child struck while crossing the street by a yellow taxicab (UCS) becomes overwhelmed with fear and intrusive thoughts of the accident (CR) upon any subsequent exposure to yellow taxicabs (CS).

Under normal conditions, repeated exposure to yellow cabs (CS) without being involved in an accident (UCS) should lead to extinction of the conditioned anxiety response (CR). However, with PTSD (Keane, Zimmering, & Caddell, 1985), this extinction fails to occur because the traumatized individual is negatively reinforced (i.e., their anxiety goes down) for avoiding the CS (e.g., yellow cabs). In addition, other previously neutral stimuli associated with the CS, such as the color yellow itself, may also elicit the conditioned anxiety response through the process of stimulus generalization. Finally, the CS themselves may go on to condition other nontrauma-related stimuli (higher order conditioning). Thus, traumatized persons may find themselves confronted with a wide array of anxiety-eliciting cues in their environment that cause persistent hyperarousal, intrusive recall of the traumatic event, and behavioral avoidance of the conditioned cues (Keane et al., 1985).

5.16.2.2.4 An emotional processing model

Observations from work on experimentally conditioned neuroses have led to the development of an emotional processing account of

PTSD derived from Lang's (1985) information processing theory of anxiety which, in contrast to conditioning models, takes into account subjective appraisals of the traumatic event in the development of PTSD. According to the emotional processing model (Foa & Kozak, 1986), traumatic events lead to the development of fear structures in memory that act to prevent further traumatization. These fear structures contain information about the objective characteristics of the trauma, the individual's responses to the trauma, and their subjective appraisal of the dangerousness and meaningfulness of the event (Foa & Kozak, 1986). Activation of this fear structure leads to increases in arousal, reexperiencing, and avoidant behaviors.

The meaning the individual attaches to the trauma, and their responses to it, are as important as the objective characteristics of the trauma. Specifically, the emotional processing model postulates that traumatic events can lead to changes in one's fundamental assumptions about safety and self-efficacy. This in turn leads to the development of new assumptions which attach increased threat value to a variety of previously neutral or even safety-signaling cues (Foa & Kozak, 1986). Consequently, many stimuli in the environment can now activate the fear structure, making it more readily accessible in memory than other potentially adaptive emotion networks, and diverting attentional resources from other nontrauma-related cues.

Ultimately, the fear structure can only be consolidated into long-term memory when corrective information (both cognitive and affective) is incorporated to develop a new information structure (Foa & Kozak, 1986). Foa et al. (1989) suggest that *in vivo* and imaginal exposure provide the type of corrective feedback necessary for consolidation of the fear structure. Specifically, increased contact with feared stimuli leads to habituation of the anxiety response, which they argue is a form of corrective feedback, as well as changes in the individual's beliefs about the likelihood of being retraumatized in the presence of neutral cues and safety signals.

5.16.2.2.5 A neurobiological model of PTSD

Numerous interrelated brain systems are activated during stressful events and have been implicated in the development of PTSD (see Friedman et al., 1995 for a review). The relationships among these systems are complex and poorly understood (Southwick et al., 1994). Given the complex interrelationships among brain systems, it is extremely unlikely that any one-to-one correspondence exists between PTSD symptoms and particular neurotransmitters or hormones (Southwick et al., 1994). Still, three systems have received considerable attention: the noradrenergic system, the hypothalamic pituitary (HPA) axis, and the endogenous opiates. These systems assist our understanding of the phenomenology of PTSD and have important implications for pharmacological treatment of the disorder (Sutherland & Davidson, 1994).

First, it has been suggested that these persistent symptoms of arousal reflect sympathetic nervous system dysregulation resulting from changes in the noradrenergic system (Southwick et al., 1994). During stressful events, central and peripheral levels of the neurotransmitter norepinephrine (NE) are increased leading to increases in heart rate and blood pressure and the experience of fear. These increases appear to serve an adaptive function as they arouse the individual and activate centers of the brain important to orienting to external stimuli, selective attention, and memory encoding of the event. However, repeated or uncontrollable stress may lead to chronic elevations in NE and alterations in NE receptors in the locus coereleus, resulting in the persistent physiological arousal seen in PTSD. There is an emerging body of evidence suggesting that veterans with PTSD have increased levels of NE and changes in noradrenergic receptivity compared with veterans without PTSD and veterans with other disorders (see Southwick et al., 1994 for a review). In addition to its effects on arousal, Pitman (1989) has suggested high states of NE-mediated arousal may "supercondition" traumatic events in memory, leading to persistent re-experiencing symptoms.

Second, there is increasing evidence that chronic or uncontrollable stress produces changes in activity of the corticotropin-releasing factor (CRF) and the hypothalamic–pituitary–adrenal (HPA) axis. The HPA system is known to play an important role in the stress response: releasing cortisol and adrenocorticotropic hormone (ACTH) assists in energy metabolism and tissue repair. Glucocorticoids exert negative feedback on the HPA axis to help maintain these hormones at nontoxic levels. Additionally, it appears that the NE and HPA systems stimulate each other, primarily through the action of CRF, leading to a mutually reinforcing feedback loop (Southwick et al., 1994).

Again, there is evidence that exposure to repeated or uncontrollable stress appears to cause the HPA system to become highly sensitized, leading to greater fluctuation in cortisol levels over a 24 h period (Southwick et al., 1994). A number of studies with combat veterans (see Southwick et al., 1994 and Yehuda & McFarlane, 1995 for a review) have found decreased levels of urinary cortisol in those with PTSD compared with psychiatric and normal controls, as well as changes in the number and receptivity of glucocorticoid receptors in the brain. Interestingly, Southwick et al. (1994) note that in animal studies, chronic elevation of glucocorticoids has been associated with damage to the hippocampal region and learning deficits, which may help explain the short-term memory impairments found in patients with PTSD.

Third, endogenous opiates are also released during periods of acute stress, leading to increased pain thresholds which may allow the individual to survive life-threatening events (Southwick et al., 1994). Again, however, it has been hypothesized that exposure to chronic and uncontrollable stress leads to increased sensitization of the opioid system such that lower levels of stress continue to produce the analgesic response (Maier, 1986). This response has been shown in Vietnam veterans during exposure to trauma-related cues (van der Kolk, Greenberg, Orr, & Pittman, 1989) and may contribute to development of avoidance and numbing symptoms (Southwick et al., 1994). Moreover, it has been suggested that the heightened sensitivity of the opioid system may play a role in intrusive recall of the trauma through its effects on NE and the consolidation of memory (McGaugh, 1990).

While other areas of the central nervous system are certainly involved in the pathogenesis of PTSD, the above findings are pertinent. First, dysregulation of the HPA and noradrenergic systems may explain why individuals with PTSD frequently suffer from chronic levels of hyperarousal, hyper-reactivity to trauma-related stimuli, emotional numbing, and difficulties with concentration and memory (including intrusive imagery) (Southwick et al., 1994). Second, these findings further our understanding of why not all individuals develop PTSD. Rather than viewing PTSD as a normal reaction to an abnormal event, it may be a complex psychological and biological response to trauma involving the phenotypic expression of vulnerabilities to the disorder (Yehuda & McFarlane, 1995). Finally, the ability to distinguish individuals with and without PTSD based on the role of HPA and noradrenergic dysregulation provides further evidence of the validity of the disorder, and has significant implications for the assessment and treatment of PTSD in both adults and children.

5.16.2.2.6 Characteristics of the trauma and other risk factors for PTSD

Regarding the trauma itself, revisions to Criterion A under *DSM-IV* aim to provide a tighter objective definition of the event, and recent research with children has concerned the relationship between the objective characteristics of the stressor and subsequent symptoms. A number of studies have reported the presence of a dose–response relationship among children involved in a sniper attack on a school (Pynoos et al., 1987), a shipping accident (Yule et al., 1990), an earthquake (Pynoos et al., 1993), a hurricane (Lonigan et al., 1991), and during war (Kuterovac, Dyregrov, & Stuvland, 1994). Similarly, attention to refining the descriptions of the nature of children's traumatic experiences has led to a more precise identification of the specific features of traumatic events which are associated with risk of PTSD. For example, exposure to direct life threat, witnessing mutilating injury or grotesque death (especially of family members or friends), and hearing unanswered screams for help were associated with more severe pathology among children and adolescents after a shipping disaster (Yule & Williams, 1990) and an earthquake (Pynoos et al., 1993).

The relationship between trauma and response has been described in detail by Terr (1991) who distinguishes between Type I and Type II traumatic events. Type I traumas are single, sudden, severe, acute stressors (e.g., natural disasters). By contrast, Type II traumas, such as childhood sexual or physical assault, are typically repeated, chronic, and expected. Terr (1991) suggests that Type I traumas are more likely to result in classic PTSD symptoms, whereas children exposed to Type II traumas are more likely to show characterological problems. While this distinction is of potential clinical relevance, it has yet to be empirically investigated and should not be overly applied. Indeed, traumatic events experienced during war are often repeated and chronic, but the associated psychopathology contains a high prevalence of classic PTSD symptoms (Kuterovac et al., 1994).

It is recognized, however, that stressors cannot be completely defined in objective terms (Rachman, 1980), and there is increasing interest in cognitive threat appraisal and attributional processes among children and adolescents (Joseph, Brewin, Yule, & Williams, 1991, 1993). The revised diagnostic criteria for the defining stressor in *DSM-IV* now include immediate responses of intense fear, helplessness, or horror (or disorganized or agitated behavior in children) in the face of actual or threatened life threat or injury to self or others (American Psychiatric Association, 1994). This stands in important contrast to previous formulations in which the stressor was seen as one that would cause significant distress to almost anyone.

There is evidence that attributional processes can mediate symptoms in child survivors. In young survivors of a shipping accident, Joseph, Brewin, Yule, and Williams (1993) found that more internal and controllable attributions were associated with intrusive thoughts and depressive feelings one year after the accident. Yule et al. (1990) have presented some evidence of a subjective dose–response relationship in survivors of the same accident. Children who had wanted to go on the cruise but had not obtained places were far less badly affected than those who went on the trip, but more badly affected than those who had never wanted to go.

Whatever the nature of the traumatic event and the child's appraisal of it, there are clearly a number of risk and protective factors which influence the development of disorder in children. Girls have been found to be more vulnerable than boys in a number of studies including a shipping disaster (Yule, 1992), a hurricane (Lonigan et al., 1991), and an earthquake (Pynoos et al., 1993) across a variety of self-report measures. Interestingly, Burke et al. (1986) found boys' distress to be higher than girls' five months after a flood, but reported that 10 months after the flood, girls showed more distress.

Age has been thought to be a factor in the development of PTSD. However, following the earthquake in Armenia, age was not found to correlate with Stress Reaction Index scores (Pynoos et al., 1993). Keppel-Benson and Ollendick (1993) argued that younger children may be protected from strong emotional reactions because of their limited cognitive capacity, particularly their awareness of threats of harm. Preliminary evidence of this hypothesis was found by Green et al. (1991) who noted that preschoolers' perceptions of threat were unrelated to symptomatology following the Buffalo Creek disaster. However, there are insufficient studies to determine whether children of different ages are at different risk of developing PTSD.

Studies of premorbid risk factors in children are also very limited. As yet, no studies have investigated systematically pretrauma abuse histories, psychological functioning, or familial psychiatric history in relation to the child's subsequent risk of PTSD. However, regarding premorbid ability and academic attainment, Yule and Udwin (1991) found that among survivors of the Jupiter shipping accident, an increased risk of PTSD was associated with low pretrauma attainment at school.

In addition to pretrauma risk factors, events and processes following exposure have been shown to influence development of the disorder. Family reactions to a disaster may be important moderating factors in the development of childhood PTSD. Studying families after the Australian bushfires, McFarlane (1987) found that parents' adjustment was an important determinant of the adjustment of their children: families who found it difficult to share their immediate reactions had more trouble with long-term adjustment. Equally important, the child's reaction to the fires affected the family's adjustment, demonstrating the reciprocal interactions among family members.

Pynoos (1994) argues that there are many interacting critical factors which play a role in the progression from traumatic exposure to subsequent psychopathology. These include social structures, community organization, family functioning, and the individual stressful challenges to the child. Pynoos' (1994) developmental model provides a useful framework for understanding this pathway, and implicates potential areas for prevention and treatment.

5.16.2.2.7 Developmental course

Being explicitly developmental, Pynoos' (1994) model (see above) provides a useful framework for understanding the course of post-trauma reactions in children. Nevertheless, there are very few systematic studies which have examined the long-term effects of exposure to traumatic events on children.

Two months after the Australian bushfires, children who had been exposed were rated as less disturbed on Rutter's screening scales than a nonexposed comparison group studied elsewhere (McFarlane et al., 1987). However, by eight months, both parents and teachers reported significant increases in the numbers of children at high risk of psychiatric disorder, and these high rates were maintained at 26 months. Around a third of this sample was still preoccupied with the event more than two years later (McFarlane et al., 1987).

Fourteen months after a school sniper attack in Los Angeles, Nader, Pynoos, Fairbanks, and Frederick (1991) found that the majority of children felt afraid when they thought about the attack, and most were specifically afraid of recurrence. In her follow-up of child survivors of a bus hijacking, Terr (1979, 1983) found very few children who were fear free some five years after the event. Thus it appears that fear may be a long-lasting consequence of traumatic events in children.

Assessing Cambodian refugees who had emigrated to America four years earlier as adolescents, Kinzie et al. (1986) found a range of intrusive and avoidance phenomena (50% of the sample suffered from PTSD), but also major depressive disorder. Prolonged depressive symptoms were found in this sample of adolescents and young adults who had been exposed to massive and repeated trauma.

Results from a six to seven-year follow-up of survivors of the Jupiter shipping accident (Yule & Udwin, 1991; Boyle et al., 1995) revealed soberingly high rates of psychiatric disorder. Preliminary results from a cohort of 98 survivors showed that of the 51% who met criteria for PTSD at some time since the disaster, 60% had PTSD which lasted for up to two years, with a further 18% having PTSD for five years. Rates of other disorder, including panic, separation anxiety, OCD, substance abuse, dysthymia, and hypomania, were also elevated compared with a control group. Thus, a range of symptoms may persist over several years. Children do improve over time, however. After a school bus crash, high levels of acute stress reactions during the first week dropped markedly over the following nine months, although initial levels of distress predicted later ones (Milgram et al., 1988).

Evidence from the Jupiter follow up (Boyle et al., 1995) and other studies suggest that the severity of the stressor may be one of the most important factors in determining the severity of later pathology. However, while Kinzie et al. (1986) found no direct relationship between extent and type of exposure and psychiatric diagnosis four years later among Cambodian refugees, their current living situation showed a strong relationship to diagnosis. McFarlane's (1987) data show family factors to be an important variable moderating outcome. It seems that the course of post-traumatic stress symptoms such as intrusive re-experiencing and avoidance of feelings follow a different time course than symptoms of depression and anxiety, both of which increase after the first six months postdisaster. Clearly, there will be a complex interaction over time of factors which influence the course and chronicity of stress reactions (Pynoos, 1994) and further research is needed to delineate these factors.

5.16.3 CONCEPTUALIZATION AND CLINICAL FORMULATION

5.16.3.1 Multimodal Assessment

As stated above, PTSD is a cluster of symptoms overlapping several domains of functioning. As such, the interview should be geared towards assessing all aspects of client functioning. This can be accomplished through a direct interview with the child as well as the family and teachers, self-report inventories and diaries, and psychophysiological recording. In addition, the therapist should try to gather all available information possible about the exact nature of the trauma from other witnesses, news reports, and police/rescue personnel. Finally, establishing the nature and extent of impairment through review of available medical and school reports is important. Experience has shown that these additional sources of information are important to help the interviewer understand the sequence of events during the trauma and often reveal symptoms that either are not readily admitted or remembered by the child and his or her family.

5.16.3.1.1 Interviewing the child

Pynoos and Eth (1986) have developed a technique for interviewing young children shortly after the witness of a traumatic or violent event, although it has broad applicability to all types of direct and indirect trauma experienced by children. The interview is broken down into three components: opening, trauma, and closure. "Opening" refers to the establishment of a focused but free discussion of the traumatic event. This is accomplished by: (i) communicating to the child that the interviewer has experience with talking to children about their traumatic experiences; (ii) giving the child pencil and paper and asking them to "draw whatever you would like but something you can tell a story about"; and (iii) encouraging the child to elaborate further (verbally or through drawing) on the story.

During the second stage of the interview, aspects of the child's drawing are linked explicitly to some aspect of the traumatic event by the therapist to facilitate emotional release. For example, the therapist might say, "I bet you wish that your father could have been saved in the end like the clown." Once the therapist has judged that the child is ready to proceed emotionally, he or she may say, "Now is a good time to tell what happened and what you saw." The child is supported in their focus on the traumatic event by asking them to draw the traumatic event, focusing and describing the sensory aspects of the event, the worst moment, any violence or mutilation, and how they coped with these experiences. Pynoos and Eth (1986) have also suggested that the child be encouraged to describe any plans or thoughts they had or have about ways in which they could have remedied the situation, feelings of blame, the desire for retaliation and fears of counter-retaliation, and feelings of control. Then the therapist inquires about previous traumas, recent dreams that may be remembered, their concerns about the future, and current stresses.

In the third and final stage of the interview, the process of termination is begun. This is accomplished by having the child review and summarize the session, normalizing the child's reactions to the traumatic event, and linking the trauma to the original drawing or story (Pynoos & Eth, 1986). In addition, the therapist may discuss with the child information they have about the expected course of the disorder. This is done by helping them identify potentially fear-eliciting, trauma-related stimuli and their potential reactions to these stimuli outside the session. To bolster the child's self-esteem, the therapist may remark on the child's courage during the event and in retelling the story during the interview. Finally, the child is asked to describe what was helpful or disturbing about the interview process and the therapist expresses his or her respect for the child's views and the privilege of having shared the interview experience with them. While the above interview technique does not specifically include any standardized assessment of PTSD symptomatology, it has proven useful in raising awareness of the sensitive nature of the trauma interview and the importance of therapist skill in obtaining accurate information.

5.16.3.1.2 Childhood measures of post-traumatic distress

Since the mid-1980s, a variety of interview techniques and instruments have been developed to assess PTSD in children directly. The majority have been adapted from semi-structured psychiatric interviews or self-report questionnaires originally developed with adults. In addition, several pre-existing measures, originally designed to assess childhood anxiety, fear, depression, and behavioral disturbance have been applied to the study of childhood PTSD. This section will introduce the reader to some of the more widely used instruments. Readers interested in a more comprehensive review of the available measures of childhood PTSD and their psychometric properties are referred to Nader (1996) and McNally (1991).

Perhaps the most widely used measure of PTSD symptomatology is the IES (Horowitz et al., 1979). It assesses the cardinal symptoms of PTSD (intrusion and avoidance), and has been used with children aged eight years and over (Yule & Udwin, 1991; Yule & Williams, 1990). It has proven particularly useful in the assessment of post-traumatic stress when used in conjunction with measures of childhood anxiety, fear, and depression (Stallard & Law, 1993; Yule & Udwin, 1991). Such measures include the Birleson Depression Inventory (Birleson, 1981), the Children's Manifest Anxiety Scale (Reynolds & Richmond, 1978), and the Fear Survey Schedule for Children (Ollendick, 1983). However, it should be noted that factor analyses suggest that some items of the IES are misunderstood by children and should be replaced (Yule et al., 1990).

Another widely used measure of childhood PTSD is the Children's Post Traumatic Stress Reaction Index (CPTS-R) (Frederick, 1985b). In a number of major studies of child trauma (Nader et al., 1991; Pynoos et al., 1987, 1993; Pynoos & Nader, 1988), the CPTS-R has been shown to have good internal consistency and to relate well to clinical judgement of PTSD severity (Yule & Udwin, 1991). Saigh's (1989) Children's Posttraumatic Stress Disorder Inventory is a structured inventory designed to explore *DSM-III* PTSD symptoms and has good psychometric properties. Finally, the Children's Impact of Traumatic Events Scale-Revised (CITES-R) (Wolfe, Gentile, Sas, & Wolfe, 1991) was developed to assess the impact of sexual abuse on children.

There are a number of semistructured interviews based directly on *DSM* criteria available that allow the interviewer to diagnose PTSD in the child. These include: the Clinician Administered PTSD Scale for Children (CAPS-C; Nader, Kriegler, Keane, Blake, & Pynoos, 1994); the Diagnostic Interview Schedule for Children (DISC; Shaffer et al., 1996); and the Diagnostic Interview for Children and Adolescents Revised (DICA; Reich, Shakya, & Taibelson, 1991). All require the interviewer to be trained in the administration of the scale, and have been shown to possess adequate validity and reliability (cf., Nader, 1995).

Unfortunately there has been a dearth of standardized approaches to the assessment of stress reactions in children less than eight years of age. While children between three and eight years can often give adequate verbal responses to questions from standardized measures used with older children, there is a need to develop measures particularly suited for young children. This is particularly important in light of the legal implication of PTSD interviews and emerging data on suggestibility in children.

Still, it has been argued that the rush to develop semi-structured interviews to elicit and quantify PTSD in children is somewhat misguided in that they simply rephrase *DSM* criteria into questionnaire format (Yule, 1994). As discussed above, many of the *DSM* criteria are irrelevant to children, and thus semi-structured interviews may have only limited validity in assessing the complete range of traumatic reactions observed in traumatized children. Moreover, it should be recognized that the interview is itself both a potentially threatening and therapeutic experience, requiring considerable skill and sensitivity on the part of the interviewer, and a firm understanding of child development. The establishment of rapport and a safe environment in which the child can discuss painful and angry feelings are very important to acquiring accurate information. The use of semistructured interviews and self-report questionnaires must always be placed within this context.

Finally, physiological measures such as galvanic skin response, heart rate, and respiration may prove useful in the assessment of PTSD, although they have received limited attention in the childhood PTSD literature (Finch & Daugherty, 1993).

5.16.3.1.3 Interviewing the parent

While the importance of interviewing children directly has been clearly established (Nader & Pynoos, 1989; Reich & Earls, 1987; Terr, 1979), children may not report the full range of PTSD symptoms (Nader & Fairbanks, 1994). Thus, a comprehensive psychiatric assessment should include a direct interview with the parents or primary caregivers. Most of the semistructured diagnostic interviews available include both child and parent versions which allow the interviewer to cover all symptoms systematically with parents and child. This is particularly important where the symptom may be of a particularly embarrassing nature, as with enuresis, thumb-sucking, and crying, or where the child may be unaware of their behavior, as in "clinging" or sleep disturbances.

Beyond a comprehensive coverage of the child's developmental history and current symptom state, it is important to establish the time-line of events leading up to, including, and following the trauma as this may help the therapist in assessing the child's accuracy of recall (Nader, 1995). Young children (less than eight years of age) are prone to time distortions because of age-appropriate levels of cognitive maturation. However, even older children may confuse the sequence of events because of extreme emotional arousal and chaos experienced during the traumatic event.

5.16.3.1.4 Discrepancies in parent and child reports

Children often minimize the extent and severity of their own disruptive behaviors, whereas parents often overemphasize these symptoms while remaining unaware of particular fears or negative emotions (Nader & Pynoos, 1989). An additional problem with parental reports is that parents may be inclined to minimize the effect of a trauma on their child and thus under-report symptoms (Sternberg et al., 1993). This may occur as a function of the parents' own abhorrence to the trauma and its recall. It is important then that the therapist be attentive to the parents' emotional state and level of arousal when recalling the trauma.

5.16.3.2 Treatment Formulation

From the above outline of multimodal assessment, it follows that treatment formulation will draw on a number of sources. Conventionally, treatment planning will be based on an assessment of current difficulties, the child's (and his or her family's) premorbid functioning, and currently available resources. Additionally, in trauma work, the kind of intervention offered will depend on how recently the traumatic event occurred.

Where children are seen within days or weeks after the traumatic event, the aim of intervention is to facilitate recovery rather than to treat symptoms. In general, recovery will be facilitated by retelling the facts of the incident, expressing feelings associated with it, normalizing reactions, educating and informing about the expected course, and by the enhancement of coping strategies (see Section 5.16.4.1.2 and Pynoos & Eth 1986). Nevertheless, effective intervention at this stage will depend also on a thorough assessment of the child's living situation and available resources. Work with carers is necessary to assist them to help their children, especially when they have themselves been directly involved in the incident. Practical issues of social welfare may arise, for instance, in cases where there is ongoing community violence, or where the child has suffered a parental bereavement. Here, multidisciplinary work is required so that as secure and safe an environment as possible is available to the child in the immediate aftermath and short term following the incident.

Treatment formulation will differ when children are referred months to years after the event. A comprehensive pretrauma history from the child's carers will help in understanding the child's usual reactions to stress and their ways of coping: whether current symptomatology is in part a function of prior temperament; what changes in behavior the traumatic event precipitated; and in setting appropriate treatment goals. Considering the child's developmental level will allow one to assess which if any of the presenting complaints are normal age-appropriate behaviors (e.g., separation anxiety), as well as guiding the choice of treatment strategies. Given the important mediating role of parental reactions and fears (McFarlane, 1987), assessment of all family members is usually necessary, both to determine whether adults or siblings also require intervention, and to advise carers on management of their children. Since the severity of the trauma is an important factor in determining the severity of later pathology (Yule, 1994), assessing the extent and type of exposure is necessary, and to address multiple episodes within a single traumatic event (Pynoos & Nader, 1993).

Treatment goals will differ according to the pattern of presenting problems, but common goals normally will include helping the child to understand the nature and effects of trauma, facilitating his or her ability to think and talk about the event without debilitating anxiety, and reducing the intensity and frequency of intrusive, avoidant, and hyperarousal symptoms (Ribbe et al., 1995). Prioritizing goals will be done collaboratively with the child and their carers, and will address those symptoms that are impacting most severely on the child's day-to-day functioning.

Common to most treatment approaches is the notion that the traumatic event requires adequate emotional processing (Rachman, 1980), and at the core of most psychosocial treatment methods is imaginal exposure to the event to facilitate such processing (see below). Also, given the range of reactions to critical incidents, a flexible approach to treatment is called for by which the child's coping strategies to deal with distressing symptoms are developed and enhanced. To take a common example of symptom management, sleep hygiene techniques can often be therapeutically beneficial while addressing the central trauma. While clinicians must be alerted to their own avoidance of the most difficult aspects of the child's experience, prioritizing areas of particular difficulty with the child and then bolstering coping strategies may increase a sense of mastery and self-efficacy while reducing feelings of helplessness and victimization.

Children with PTSD may present with other Axis I comorbidity and chronic pain reactions resulting from injury, and in this case, treatment formulation will be modified accordingly. In traumatic bereavement, Dyregrov (1991) recommends that the traumatic nature of the death needs to be addressed before grieving can proceed.

Given the cascade of events that may follow a traumatic incident, treatment formulation will also address secondary adversities, with the aim of preventing further complications. Often, these may be traumatic reminders, ongoing legal processes, and anniversaries. It is common, for example, that academic attainment drops following exposure (e.g., Yule, 1994), and this in turn may potentially impact on the child's functioning in other areas and later into adulthood. Therefore, a comprehensive treatment plan may include "pulsed intervention" at times of expected stress such as changing schools and anniversaries (e.g., James, 1989).

5.16.4 TREATMENT

While there is a growing number of single case studies and detailed clinical descriptions of treating traumatized children, there are as yet no randomized controlled outcome trials. In the main, however, treatment approaches are broadly cognitive behavioral: the guiding principle is that exposure to trauma-related cues and memories in tolerable doses is needed to reduce PTSD symptoms. Professionals can help by creating a relatively safe environment in which such recounting can take place. Given the nature of many traumatic events, early intervention is often best done in groups and through existing structures such as schools. Work has also focused on contingency planning and disaster management using a public health approach which involves community resources. There is little available evidence for the effectiveness of pharmacological treatment of PTSD in children.

5.16.4.1 Psychosocial Approaches

5.16.4.1.1 Preventative measures and contingency planning

Traumatic events are by their nature unexpected and uncontrollable. Nevertheless, it is well recognized that much can be done to alleviate distress in children if the unexpected is planned for. Outside the family, school is a primary source of social support for children, and a number of authors (Klingman, 1993; Pynoos & Nader, 1988; Yule & Gold; 1993) have provided guidelines for how schools can plan to help child survivors of catastrophes. Particular emphasis is placed on schools and mental health professionals working together.

Planning for disasters will not prevent them from happening, but by anticipating many of the actions needed, schools may reduce the impact of critical incidents. Thus, Yule and Gold (1993) recommend a number of short-, medium-, and long-term tasks for school staff to be alerted to. School staff should be made aware of the likely psychological effects of disasters on children, parents, and teachers before disaster strikes through liaison with local child mental health services. In the event of a traumatic incident, immediate action should include ensuring that correct information can get into and out of the school, ensuring that several staff have access to next-of-kin lists, setting up strategies to deal with enquiries, informing parents about what has happened with care and sensitivity and reuniting parents and children, nominating a press officer to deal with the media, informing staff and pupils in an appropriate and careful way, attempting to stick to normal routines as much as possible, and planning for attendance at funerals.

School staff should be briefed about what has happened, and after contacting relevant outside agencies, debriefing meetings can be set up for directly involved staff and children. Procedures for monitoring both staff and children should be set up, and strategies for allowing young people who want to express their feelings about the situation should be activated.

In the medium term, it will be important for children who remain at home or in hospital to be contacted, and arrangements made either for their return to school or for alternative teaching. Affected staff and children may continue to need support, and this will be achieved more easily if ground rules for working with outside agencies have been set up beforehand. Close liaison with parents at this stage will be necessary.

In the longer term, introducing strategies to continue to monitor vulnerable pupils and staff is useful, to decide how to mark anniversaries, and to be aware that legal processes and media attention may bring back distressing memories and cause temporary upsets within the school. Schools should not have to deal with such enormously difficult issues such as death and bereavement for the first time following a crisis, and it is recommended that schools plan the curriculum to work with rites of passage as a matter of course.

Both Klingman (1993) and Yule and Gold (1993) advise that classroom interventions by teachers can be effective in defusing and normalizing emotions, not least by addressing rumors and myths about the disaster. Classroom interventions will include monitoring for high-risk children, since in themselves, they are no substitute for individual, group, or family interventions when needed (Pynoos & Nader, 1993).

5.16.4.1.2 Debriefing and crisis intervention

When schools are prepared for the unexpected, they are in an optimal position to provide early crisis intervention, with support from mental health professionals. Pynoos and Nader (1993) describe a number of "psychological first aid" strategies for common symptoms which can help to normalize the recovery process. They point out that these must be developmentally appropriate and provide a number of examples of techniques or strategies for use by teachers in schools.

A more widely used technique is Critical Incident Stress Debriefing (CISD) (Mitchell, 1983; Mitchell & Everly, 1993). Developed initially for use with emergency service personnel, debriefing has been adapted and used successfully with children following disasters (Dyregrov, 1991). Within a few days of an incident, the survivors are brought together in a group with an outside leader. During the introductory phase, the leader sets the ground rules for the meeting, emphasizing that they are there to share feelings and help each other and that what goes on in the meeting is private. The information should not be used to tease other children. No one has to talk, although all are encouraged to do so. The leader clarifies the facts of what actually happened in the incident. This allows for clarifications of any misunderstandings about how the traumatic event unfolded. They are asked about what they thought when they realized something was wrong, and this leads naturally to discussions of how they felt and their current emotional reactions. In this way, children share the various reactions they have experienced and usually learn that others have felt similarly. The leader labels their responses as normal (understandable) reactions to an abnormal situation.

Many children are relieved to learn that they are not the only ones experiencing strange feelings, that there is an explanation and that they are not going mad or crazy. The leader summarizes the information arising in the group, elicits and reinforces the existing coping strategies of the children, and educates them in what simple steps they can take to control some of their reactions. They are also told of other help available should their distress persist. The aim of debriefing is to facilitate the recovery process rather than to treat existing symptoms. The combination of normalizing reactions and teaching basic coping strategies can do much to enhance children's sense of mastery and control over their feelings. There is promising evidence from two studies which report reductions in distress among children following debriefing meetings.

Following the sinking in Athens of the cruise ship "Jupiter," which was carrying a party of British school children, Yule and Udwin (1991) described how they carried out a debriefing session with 24 survivors 10 days after the event. The debriefing meeting was carried out at school with the involvement of the survivors' teachers. As part of the meeting, the girls completed a screening battery consisting of the IES (Horowitz et al., 1979), the Birleson Depression Inventory (Birleson, 1981), and the revised Children's Manifest Anxiety Scale (Reynolds & Richmond, 1978). The girls were rescreened between 5 and 10 months later, and results were compared with the same self-report measures from a group of 15 girls from a different school who did not participate in any debriefing (Yule, 1992). The debriefing group showed significantly lower IES scores, and this effect was strongest for intrusive thoughts. The debriefing group also showed significantly reduced fears (Fear Survey Schedule) (Ollendick, Yule, & Ollier, 1991) compared with the control group. Scores on the depression and anxiety measures were slightly but not significantly lower. While these results are promising, some girls in the debriefing group went on to receive group treatment, making it impossible to isolate the effects of debriefing from subsequent therapy.

More convincing evidence is reported by Stallard and Law (1993), who carried out debriefing sessions with a group of seven young survivors of a minibus accident. The two debriefing meetings were carried out one week apart in school, and six months after the accident. The same screening battery as above was used before the first session, and children were subsequently rescreened three months later. Significant reductions were found in the group scores on all measures (PTSD, depression, and anxiety). Although no control data were obtained, given the stability of the scores in the six months before intervention, the authors argue that the improvement was related to the intervention.

Following a recent spate of well-publicized traumatic incidents involving school children, the call for debriefing and crisis counseling for all survivors is increasing. Although the available evidence is promising, there are no controlled studies demonstrating the efficacy of debriefing in reducing later distress. Given the debate about the usefulness of CISD among adults (Meichenbaum, 1994; Mitchell, 1996), there is a clear need for further evaluation of its efficacy with children. It is not known, for example, whether all survivors benefit, nor indeed when best to offer such debriefing.

5.16.4.1.3 Group treatment

Where critical incidents involve relatively large numbers of children, it makes sense to direct some therapeutic support through naturally existing groupings that are present in communities and schools. As with debriefing, the aims of such therapeutic groups will be to share experiences and feelings, to boost children's sense of mastery and control, and to share ways of solving common problems. Pynoos and Nader (1993) point out that providing a forum for the expression of feelings is not sufficient for groups: this may only renew feelings of anxiety unless a constructive, therapeutic approach is taken. A variety of approaches to group treatment have been described following earthquakes (Galante & Foa, 1986; Goenjian, 1993), a school sniper attack (Gillis, 1993), and shipping disasters (Yule & Udwin, 1991; Yule & Williams, 1990).

After a devastating earthquake in Italy in 1980, Galante and Foa (1986) carried out a group treatment program with 62 children from the most severely affected village. They describe a relatively structured approach, holding roughly monthly meetings throughout the academic year. Sessions lasted for an hour and each group consisted of four children from the same grade. Sessions focused on communicating about feelings (particularly about the earthquake), discussion and normalization of fear reactions, "active discharge of feelings about the earthquake," focusing on the future and rebuilding of their village, and sharing active coping strategies. Pre- and postgroup evaluation using the Rutter Children's Behaviour Questionnaire for Teachers (Rutter, 1967; Zimmerman-Tansella, Minghetti, Tacconi, & Tansella, 1978) showed a significant reduction in distress which compared favorably with children from other villages where children did not receive such support. Additionally, frequency counts were taken during sessions of repeating earthquake stories and the expression of related and unrelated fears. It is interesting that after an initial rise in these scores in the first few groups, they showed a dramatic decline only after a session where children looked to the future and were engaged in play to rebuild the village and provide safer housing.

A somewhat less structured approach was described by Yule and Williams (1990) with child and adult survivors of the Herald of Free Enterprise shipping disaster. Groups were open to any survivor to attend and met initially at monthly intervals over about a year, tailing off to every six to eight weeks thereafter. The children seemed to appreciate the opportunity to share their feelings with others without having to explain everything to outsiders. They took advantage of the cognitive behavioral problem solving approach by which efforts were made to help them identify things they were finding difficult and to use the resources of the group to develop solutions. The therapist working with the children also acted as a go-between with the parents' group. Yule and Williams (1990) suggested that an important part of group work with children following a disaster is that they have time together without the pressure of having to talk about the trauma.

Gillis (1993) worked with children following a school sniper attack. In small groups, both structured and more open-ended elements were used, including role play and re-enactment of the shooting. Gillis emphasizes that when groups are less structured, setting clear limits is important, for if behavior is out of control, it is likely to result in an increase in the children's anxiety. He found that running separate groups for boys and girls was better because of the different reactions they had to the attack, boys showing more externalizing reaction, and girls more internalizing ones. Groups comprised children of the same level of maturity, and Gillis suggests that it is optimal to work with groups of six to eight children.

It is notable that while the group treatment approaches reported above are well described and appear promising, there are no controlled studies providing evidence for the efficacy of such groups. There is a consensus that group interventions will be effective for some, but not all, children. However, one function of therapeutic groups is to screen for high-risk children who may need individual therapy. More generally, children whose problems persist despite group help should be treated individually.

5.16.4.1.4 Individual treatment

Broadly-based cognitive behavioral treatment methods aim to help survivors make sense of what has happened and to master feelings of anxiety and helplessness after a traumatic incident. Most survivors recognize that eventually they must "face up to the traumatic event." The problem for the clinician is how to help the survivor remember and re-experience the event and the emotions that it engenders so that the distress can be mastered rather than magnified. This will depend foremost on the establishment of a safe and trusting environment in which the traumatic event can be remembered and discussed. Therapists must be prepared to ask children about the most difficult aspects of the traumatic experience, but at the same time ensure that exposure to traumatic memories is paced so that the child does not experience overwhelming anxiety. For many children, talking directly about the traumatic event may be too difficult and other means of accessing traumatic memories must be found.

Asking children to draw their experiences often assists in the recall of both the event and the accompanying emotions. Pynoos and Eth (1986) provide a detailed account of how drawings may be used, not as projective techniques, but as a means of communicating about the experience (see Section 5.16.3.1.1). In younger children, play may be used similarly. Misch, Phillips, Evans, and Berekowitz (1993) gave children models and dolls representing the various components of the traumatic event which the children then used to answer questions about their experience.

At the core of cognitive-behavioral treatments with children is the use of imaginal and *in vivo* exposure techniques within a safe therapeutic environment. Results of controlled trials with adults generally support the use of exposure therapy for PTSD (e.g., Boudewyns & Hyer, 1990; Cooper & Clum, 1990; Keane, Fairbank, & Caddell, 1989), but trials of these techniques are still needed with children. Nevertheless, single case studies with children have shown very promising results.

Saigh (1987a) was the first to show that, as Rachman (1980) had predicted, longer exposure sessions than normal are needed if desensitization is to occur. He employed flooding techniques to treat a 14-year-old Lebanese boy who met criteria for PTSD after having been abducted and tortured. The imaginal flooding process consisted of 10 min of deep muscle relaxation, followed by 60 min of therapeutic stimulation during which the boy was instructed to imagine the particular details of four anxiety-evoking scenes. SUDS ratings were elicited at two min intervals during the aversive scene presentations. After seven flooding sessions, SUD levels decreased appreciably. Post-treatment and at four months follow-up, there were clinically significant treatment gains in self-reported anxiety (STAI; Spielberger, Gorsuch, & Luschene, 1970), depression (BDI; Beck, Rush, Shaw, & Emery, 1979), and misconduct, and virtually no arousal to the anxiety provoking scenes. The boy could complete all activities of a Behavioral Avoidance Test (BAT) involving a 10 min walk to the place where he had been kidnapped. The efficacy of such procedures was confirmed in a series of clinical replications involving multiple baselines across traumatic scenes (Saigh 1987b, 1989a, 1989b); and Saigh (1992) subsequently has summarized a five-stage intervention process involving education, imagery training, relaxation training, the presentation of anxiety provoking scenes, and debriefing.

Yule (1992) has described a different approach in the treatment of a 16-year-old boy who had survived a shipping accident two years earlier. The boy was asked to imagine and describe accurately what he had observed, heard, felt, smelled, and thought during the accident. The session was audiotaped and the boy went on to listen to the tape at home between sessions. Following treatment, significant clinical gains were observed on the standardized measures of anxiety and depression, and on the IES. After treatment, the boy could travel on a boat again for the first time in two years.

Therapeutic exposure under supportive circumstances seems to deal well with intrusion symptoms and behavioral avoidance. However, a flexible treatment approach which uses additional techniques to treat other common reactions such as sleep disturbance, separation anxiety, anger and conduct problems, prolonged grief reactions, generalized anxiety, and depression will be necessary.

Sleep disturbance is a common problem among children who have lived through traumatic experiences. A careful analysis will reveal whether the problem is one of getting off to sleep or in waking due to trauma-related nightmares. In the former case, implementing relaxation routines before bed and masking thoughts with music may help. In the latter, there are now positive cognitive behavioral techniques for alleviating nightmares, including dream restructuring approaches. In younger children, post-traumatic regressive behavior may include night time enuresis, and this may be treated using star charts and enuresis alarm techniques while addressing the central trauma.

Given the nature of traumatic events, children are frequently bereaved as a consequence. Pynoos and Nader (1988) emphasize the need to help children to distinguish their trauma-related responses from those related to grief, and suggest that where several children are bereaved, small groups can be beneficial in the initial stages. Dyregrov's (1993) account of the distinguishing features of traumatic bereavement implies that the traumatic nature of the death and post-trauma reactions need to be addressed before grieving can begin. Black (1993) uses a variety of techniques, including the use of drawings and play in her work with children who have been bereaved because of one parent killing another. Importantly, she also describes how family work—including that with new carers—is necessary in cases where a child has lost a parent.

Guilt, self-blame, helplessness, and vulnerability are common themes among older children and adolescents who have survived traumatic events. While these may be related to depressive features, they are often seen to have a different quality to them. Children may talk of omens, relating perhaps to a search for meaning as well as being related to issues of culpability. In work with younger children, it is important that any misunderstandings about the causes of the event are clarified, while with older children, cognitive restructuring may be used to counter misattributions of predictability, causality, and responsibility. Janoff-Bulman's (1985) work with adults about the shattering of key assumptions regarding self-invulnerability, life meaningfulness, and world justice are useful here.

Pynoos' (1994) developmental framework for understanding trauma reactions has a number of implications for therapeutic intervention. The model suggests that the interaction over time of many critical factors (including the complexity of the traumatic experience and the interactions of traumatic reminders, secondary stress, post-traumatic distress, and development) play a role in the progression from traumatic exposure to psychopathology. Each of these factors can be seen as a potential focus for intervention. Similarly, James (1989) describes the use of early intervention with subsequent "pulsed intervention" over time using multifaceted treatment approaches.

Finally, the most recent addition to the armamentarium of individual treatment techniques for PTSD is Eye Movement Desensitization and Reprocessing (EMDR) (Shapiro, 1989). The technique employs dual-attention tasks, most notably saccadic eye movements, to facilitate accelerated processing of traumatic memories and is fully described elsewhere (Shapiro, 1995). There is considerable controversy about the efficacy of the procedure in adults with PTSD, and no controlled trials with children. Some authors have described how EMDR can be adapted successfully and used with children (e.g., Greenwald, 1994). Cocco and Sharpe (1993) report the successful use of an auditory variant of EMDR with a 4-year-old boy diagnosed as suffering from PTSD following a violent robbery. Further research is clearly needed, but given the poor understanding of the rationale behind EMDR and the lack of any convincing evidence for its efficacy in work with children, the procedure should be used with care until outcome evaluation data and reliable protocols are available.

5.16.4.1.5 Community approaches

Following large-scale disasters, whole communities may be affected and this clearly requires a different kind of approach. The type of intervention will depend, broadly, on the nature of the disaster and the kind of community affected.

Early intervention is designed to help reduce levels of distress that may later lead to psychopathology and will allow for screening of vulnerable individuals. Often, the needs of the community will be for short-term therapeutic assistance rather than a traditional mental health service. Therefore, and because of the chaos that ensues following mass disaster, outreach services are recommended in the first place. Joyner and Swenson (1993) provide an account of Project Hugo, implemented after Hurricane Hugo struck South Carolina in 1989. Outreach teams of counselors quickly provided support and information to thousands of affected people. Telephone hotlines were set up, and fliers and information booklets produced. Joyner and Swenson stress the need for networking among emergency and support teams, as well as for building on the existing pastoral resources of the community (i.e., churches or schools).

Concentrating on the needs of children after the Australian bushfires, McFarlane (1991) has described how a preventative outreach program was implemented through schools using workbooks for pupils. The workbooks provided information about the bushfires as well as exercises designed to label and normalize common reactions, and to help children develop coping strategies with help from parents and teachers.

Where pre-existing community resources are either scarce or completely devastated by the disaster, the needs are different, and a modified approach is needed. Pynoos et al. (1993) characterized the psychological effects on children of the Armenian earthquakes of 1988 as severe, pervasive, and chronic. In this case, it was not possible to set up telephone hotlines or draft in teams of counselors from elsewhere: following an evaluation of existing services and a pilot screening of children, the emphasis here was on training local therapists to carry out brief psychotherapy and behavioral therapy (Goenjian, 1993). Therapists were selected carefully before training, and worked in pairs or teams such that they could support each other. Makeshift clinics were set up, and group and individual therapy, which included play, drawing, and storytelling, was implemented wherever possible including parents and teachers. Two therapeutic booklets, one for adults and one for children, were prepared and distributed. Importantly, considerable attention was paid to helping manage the stress of community leaders and the newly trained carers, themselves also affected by the devastation.

The trend in the twentieth century for war increasingly to affect civilians (particularly women and children) directly (UNICEF; 1996) has led to attempts by international relief agencies to meet the psychological needs of children in areas of recent and ongoing conflict (Kuterovac et al., 1994; Macksoud, Dyregrov, & Raundalen, 1993; Smith & Yule, 1995). The emphasis has been on rebuilding scarce local resources by training health workers, teachers, religious leaders, and others to deliver services at a variety of levels. Useful manuals that can be used by teachers and others in recognizing the signs of traumatic stress and providing initial treatment have been written after conflicts in Mozambique (Raundalen, Dyregrov, & Gronvold-Bugge, 1990; Richman, 1991, 1993), Lebanon (Macksoud et al., 1993), and Sri Lanka (Nikapota & Samarasinghe, 1993), but it should be noted that none of these has yet been evaluated in the field.

5.16.4.2 Pharmacological Approaches

There is little available data on the pharmacological treatment of PTSD in children and adolescents. It has barely been studied, and usually as an adjunct to psychosocial treatments (Terr, 1996). Still, it is being increasingly recognized as part of a multidimensional treatment approach to PTSD in adults (Sutherland & Davidson, 1994). As discussed above, PTSD is a disorder of various symptoms including severe anxiety, dissociation and disorganization, depression, and anger, and with potential underlying neurobiological mechanisms. As such, pharmacological treatment in children cannot be ruled out.

A variety of pharmacological agents have been used in the treatment of adults with PTSD to varying degrees of success including: tricyclic antidepressants (TCAs); monoamine oxidase inhibitors (MAOIs; phenelzine); selective serotenergic reuptake inhibitors (SSRIs; fluoxetine, sertraline, and fluvoxamine); anticonvulsant agents (carbamazepine and valproic acid); benzodiazepines (aprazolam, clonazepam, lorazepam); clonidine; and beta-blockers (propranolol). While inferring from adult studies that such agents would be appropriate with children or effective is not possible, a brief discussion on this topic is instructive. In several controlled trials with adult PTSD patients (Davidson et al., 1990; Frank, Kosten, Giller, & Dan, 1988; Shestatzky, Greenberg, & Lerer, 1988), drugs with serotonergic action (TCAs, MAOIs, and SSRIs) have achieved the most positive effects with the best results achieved after longer treatment trials (Sutherland & Davidson, 1994). TCAs and SSRIs are often prescribed for children with severe anxiety and depressive reactions and their potential use with children and adolescents with PTSD should be explored.

Again, as discussed above, many of the symptoms of PTSD have been hypothesized to result from excessive activity of the central adregeneric system, leading some clinicians to use adrenergic blocking agents such as propranolol and clonidine (Sutherland & Davidson, 1994). In one of the few studies to directly approach this issue in children, Kinzie and Leung (1989) found that clonidine (a centrally acting drug sometimes used for treatment of Tourette's syndrome) may help children who have significant avoidant symptoms, startle responses, or trauma-related depression. Pynoos and Nader (1993) note preliminary results suggesting that clonidine may reduce persistent arousal in children. In an uncontrolled, nonblind study, Famularo, Küscheff, & Fenton, (1988) found that propranolol (a β-blocking agent) relieved the arousal symptoms of children who were exposed to sexual assault. The results of these studies are consistent with previous trials of clonidine and propranolol in adult combat veterans (see Kolb, Burris, & Griffith, 1984).

While carefully controlled drug-treatment trials in children and adolescents with PTSD are awaited, use of pharmacological agents shows some promise. While many psychologists express an appropriate level of caution (if not aversion) when considering referral of children for pharmacological treatment, the presence of chronic and severely debilitating symptoms of PTSD in a child may require thoughtful consideration of pharmacotherapy.

Still, some precautions are worth mentioning here. First, Lundin (1994) has suggested that drugs be chosen carefully so that they do not undermine the individual's capacity for working through the trauma. Thus, he suggests that making regular use of benzodiazepines is generally inadvisable. Second, the issue of comorbid diagnoses should be evaluated thoroughly as some agents may be useful for targeting particular PTSD symptoms but may affect certain comorbid conditions. Finally, as noted by Sutherland and Davidson (1994) in their review of the literature, response to pharmacotherapy in PTSD may be relatively slow, perhaps reflecting underlying structural changes in the brain, and trials of eight weeks or longer may be necessary to achieve the maximum benefits.

5.16.4.3 Efficacy

There have been no randomized, controlled treatment trials for PTSD in children, and no long-term treatment follow up studies. However, regarding debriefing, Yule (1992) observed that children who attended debriefing meetings fared better on a range of outcome measures than children who were not offered such help. Inferences drawn from this study must be limited, however, since the debriefing group also received additional treatment. Stallard and Law's (1993) uncontrolled trial of debriefing showed significant improvements on standardized self-report measures in a small group of young children involved in a road traffic accident.

Galante and Foa's (1986) evaluation of structured group therapy showed significant teacher-rated improvements on the Rutter Questionnaire (Rutter, 1967) in children from one earthquake-affected village. By comparison, in only one out of five affected villages where therapeutic groups were not held did children improve significantly. Concerning individual therapy, Saigh's series (1987a, 1987b, 1989a, 1989b) of single case multiple baseline across traumatic scene studies is promising for behaviorally based exposure treatments. In an uncontrolled trial of individual cognitive behavior therapy with 19 sexually abused and traumatized children, Deblinger, McLeer, & Henry (1990) found that all symptoms showed some improvement: each major category of PTSD symptoms improved, and although some children were still symptomatic, none met criteria post-treatment.

5.16.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

The study of post-traumatic stress reactions in children remains wide open with regard to areas in need of additional research. Five main areas appear particularly important for the area of child traumatic stress: (i) descriptive studies of young people's reactions across trauma types and age/developmental stages; (ii) neurobiological correlates of PTSD; (iii) neuropsychological and physiological assessment of PTSD; (iv) follow-up studies of children with PTSD; and (v) response to psychological and pharmacological treatment.

The paucity of data available looking at post-traumatic reactions at different age levels has been highlighted. While PTSD can be diagnosed even in very young children, it is inaccurate to assume that PTSD, as currently defined, is sufficient to capture the breadth of children's post-traumatic reactions. Additional research using age-appropriate measures and standardized behavioral observation methods are needed. Furthermore, children's post-traumatic perceptions of the trauma, themselves, and their world have yet to be evaluated in any systematic fashion. The adaptation of belief scales based on the work of Janoff-Bulman (1992) to children may be useful in this regard.

The neurobiological models of PTSD have not been rigorously tested in children. Studies employing the hormonal, neuronal, and brain-imaging techniques used with adults need to be developed for use with children and adolescents. Similarly, psychophysiological measures have only been used in the 1990s with children, however, with promising findings (cf., Pynoos et al., 1993). Additional studies employing salivary cortisol tests and ambulatory measures of hyperarousal may be of particular use in this regard. Moreover, studies using neuropsychological measures developed with children to increase our understanding of the memory and concentration impairments associated with PTSD in children are needed.

Finally, and perhaps most important, is the need for long-term follow-up and treatment studies of PTSD in children and adolescents. Studies employing both referred and nonreferred children with PTSD and carefully selected control groups are necessary before any firm conclusions about the chronicity of PTSD, and the effectiveness of currently available treatments, can be reached.

5.16.6 SUMMARY

In this chapter the literature on the nature and treatment of PTSD in children and adolescents has been highlighted. This literature is growing but clearly lags behind its adult counterpart. However, current research supports the conclusion that children and adolescents exposed to severe and life-threatening traumas can develop PTSD as well as other disorders.

While PTSD as a diagnostic category has significant limitations for capturing all the effects of trauma in both children and adults, it has proved a useful guideline for both clinical and research purposes. In particular, there is an emerging body of work suggesting that across cultures and traumatic events, children exhibit similar types of stress reactions, which may reflect post-traumatic changes in the central nervous system as well as emotional and cognitive disturbances. Unfortunately, our understanding of the effects of trauma at different ages and their chronicity remains poor, and like other areas of child psychopathology, controlled treatment-outcome studies of childhood PTSD are desperately needed.

5.16.7 REFERENCES

American Psychiatric Association (1980). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed. Rev.). Washington, DC: American Psychiatric Association.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Arroyo, W., & Eth, S. (1985). Children traumatized by Central American warfare. In S. Eth & R. S. Pynoos (Eds.), *Posttraumatic stress disorder in children*. Washington, DC: American Psychiatric Press.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*, New York: Guilford Press.

Birleson, P. (1981). The validity of depressive disorder in childhood and the development of a self-rating scale: A research report. *Journal of Child Psychology and Psychiatry*, *22*, 73–88.

Black, D. (1993). When father kills mother. In G. Forrest (Ed.), *ACPP Occasional Papers*, Number 8. London: Association for Child Psychology and Psychiatry.

Boudewyns, P. A., & Hyer, L. (1990). Physiologic response to combat memories and preliminary treatment outcome in Vietnam veteran PTSD patients treated with direct therapeutic exposure. *Behavior Therapy*, *21*, 63–87.

Boyle, S., Bolton, D., Nurrish, J., O'Ryan, D., Udwin, O., & Yule, W. (1995). *The Jupiter sinking follow up: Predicting psychopathology in adolescence following trauma*. Paper presented at International Society for Traumatic Stress Studies Annual Meeting (November), Boston, MA.

Bradburn, L. S. (1991). After the earth shook: children's stress symptoms 6–8 months after a disaster. *Advances in Behavioural Research & Therapy*, *13*, 173–179.

Burke, J., Moccia, P., Borus, J., & Burns, B. (1986). Emotional distress in fifth grade children ten months after a natural disaster. *Journal of the American Academy of Child and Adolescent Psychiatry*, *25*, 536–541.

Cocco, N., & Sharpe, L. (1993). An auditory variant of eye movement desensitization in a case of childhood posttraumatic stress disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, *24*, 373–377.

Cooper, N. A., & Clum, G. A. (1990). Imaginal flooding as a supplementary treatment for PTSD in combat veterans: a controlled study. *Behaviour Therapy*, *20*, 381–391.

Davidson, J. R. T., Kudler, H. S., Smith, R., Mahoney, S. L., Lipper, S., Hammett, E., Saunders, W. B., & Cavenar, J. O. (1990). Treatment of posttraumatic stress disorder with amitriptyline and placebo. *Archives of General Psychiatry*, *47*, 259–266.

Deblinger, E., McLeer, S. V., & Henry, D. (1990). Cognitive behavioral treatment for sexually abused children suffering posttraumatic stress: Preliminary findings. *Journal of the American Academy of Child and Adolescent Psychiatry*, *19*, 747–752.

Dyregrov, A. (1991). *Grief in children: A handbook for adults*. London: Jessica Kingsley.

Dyregrov, A. (1993). The interplay of trauma and grief. In G. Forrest (Ed.), *ACPP Occasional Papers*, Number 8.

Dyregrov, A., & Yule, W. (1995, May). *Screening measures: The development of the UNICEF screening battery*. Paper presented at The Fourth European Conference on Traumatic Stress, Paris.

Earls, F., Smith, E., Reich, W., & Jung, K. G. (1988). Investigating psychopathological consequences of disaster in children: A pilot study incorporating a structured diagnostic approach. *Journal of the American Academy of Child and Adolescent Psychiatry*, *27*, 90–95.

Eisenberg, C. (1996, August). *Mental health of people and* *the effects of war on children*. Plenary paper presented at the Tenth World Congress of Psychiatry, Madrid, Spain.

Eth, S., & Pynoos, R. S. (1985). Psychiatric interventions with children traumatized by violence. In D.H. Schetky & E. P. Benedek (Eds.), *Emerging issues in child psychiatry and the law* (pp. 285–309). New York: Brunner/Mazel.

Fairbank, J. A., Schlenger, W. E., Saigh, P. A., & Davidson, J. R. T. (1995). An epidemiologic profile or post-traumatic stress disorder: Prevalence, comorbidity, and risk factors. In M. J. Friedman, D. S. Charney, & A. Y. Deutch (Eds.), *Neurobiological and clinical consequences of stress: From normal adaptation to PTSD* (pp. 415–427). Philadelphia: Lippincott-Raven.

Famularo, R., Kinscherff, R, & Fenton, T. (1988). Propranolol treatment for childhood posttraumatic stress disorder, acute type. *American Journal of the Disordered Child*, *142*, 144–1247.

Finch, A. J., & Daugherty, T. K. (1993). Issues in the assessment of posttraumatic stress disorder in children. In C. F. Saylor (Ed.), *Children and disasters: Issues in clinical child psychology* (pp. 45–66). New York: Plenum.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin*, *99*, 220–35.

Foa, E. B., Steketee, G., & Olasov-Rothbaum, B. (1989). Behavioral/cognitive conceptualizations of post-traumatic stress disorder. *Behavior Therapy*, *20*, 155–176.

Fonseca, A. C., Yule, W., & Erol, N. (1994). Cross-cultural studies of children's fears. In T. H. Ollendick, N. King, & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in children and adolescents* (pp. 67–84). New York: Plenum.

(1988). A randomized clinical trial of phenelzine and imipramine for post traumatic stress disorders. *American Journal Psychiatry*, *145*, 1289–1291.

Frederick, C. (1985a). Children traumatized by catastrophic situations. In S. Eth & R. S. Pynoos (Eds.), *Posttraumatic stress disorder in children* (pp. 73–99). Washington, DC: American Psychiatric Press.

Frederick, C. (1985b). Selected foci in the spectrum of post traumatic stress disorders. In L. J. Murphy (Ed.), *Perspectives on disaster recovery* (pp. 110–130). East Norwalk, CT: Appleton-Century-Crofts.

Freidman, M. J., Charney, D. S., & Deutch, A. Y. (1995). *Neurobiological and clinical consequences of stress: From normal adaptation to post-traumatic stress disorder*. Philadelphia: Lippincott-Raven.

Galante, R., & Foa, D. (1986). An epidemiological study of psychic trauma and treatment effectiveness after a natural disaster. *Journal American Academy of Child Psychiatry*, *25*, 357–363.

Garrison, C. Z., Weinrich, M. W., Hardin, S. B., Weinrich, S., & Wang, L. (1993). Post-traumatic stress disorder in adolescents after a hurricane. *American Journal of Epidemiology*, *138*, 52–53.

Garmezy, N. (1984). Stress-resistant children: The search for protective factors. In J. Stevenson (Ed.), *Recent research in developmental psychology. Journal of Child Psychology and Psychiatry* (Suppl. 4, pp. 213–233). Oxford: Pergamon.

Garmezy, N., & Rutter, M. (1985). Acute reactions to stress. In M. Rutter & L. Hersov (Eds.), *Child and adolescent psychiatry: Modern approaches* (2nd ed., pp. 152–176). Oxford: Blackwell Scientific Publications.

Gillis, H. M. (1993). Individual and small-group psychotherapy for children involved in trauma and disaster. In C. F. Saylor (Ed.), *Children and disasters* (pp. 165–186). New York: Plenum.

Goenjian, A. (1993). A mental health relief programme in Armenia after the 1988 earthquake. *British Journal of Psychiatry*, *163*, 230–239.

Goodwin, J. (1988). Post traumatic symptoms in abused children. *Journal of Traumatic Stress*, *4*, 475–488.

Green, B. L., Korol, M., Grace, M. C., Vary, M. G., Leonard, A. C., Glesser, G. C., & Smithson-Cohen, S. (1991). Children and disaster: Age, gender, and parental effects on PTSD symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry*, *30*, 945–951.

Greenwald, R. (1994). Applying eye movement desensitization and reprocessing (EMDR) to the treatment of traumatized children: Five case studies. *Anxiety Disorders Practice Journal*, *1*, 83–97.

Handford, H. A., Mayes, S. O., Mattison, R. E., Humphrey, F. J., Bagnato, S., Bixler, E. O., & Kales, J. D. (1986). Child and parent reaction to the TMI nuclear accident. *Journal of American Academy of Child Psychiatry*, *25*, 346–355.

Herman, J. L. (1992). Complex PTSD: A syndrome in survivors of prolonged and repeated trauma. *Journal of Traumatic Stress*, *5*, 377–391.

Horowitz, M. J., Wilner, N., & Alvarez, W. (1979). Impact of event scale: A measure of subjective stress. *Psychosomatic Medicine*, *41*, 209–218.

James, B. (1989). *Treating traumatized children: New insights and creative interventions*. Lexington, MA: Lexington Books.

Janoff-Bulman, R. (1992). *Shattered assumptions: Towards a new psychology of trauma*. New York: Free Press.

Jones, J. C., & Barlow, D. H. (1992). A new model for posttraumatic stress disorder: Implications for the future. In P. A. Saigh (Ed.), *Post-traumatic stress disorder: A behavioral approach to assessment and treatment* (pp. 147–165). New York: Macmillan.

Joseph, S. A., Brewin, C. R., Yule, W., & Williams, R. (1991). Causal attributions and psychiatric symptoms in survivors of the Herald of Free Enterprise disaster. *British Journal of Psychiatry*, *159*, 542–546.

Joseph, S. A., Brewin, C., Yule, W., & Williams, R. (1993). Causal attributions and psychiatric symptoms in adolescent survivors of disaster. *Journal of Child Psychology Psychiatry*, *34*, 247–253.

Joyner, C. D., & Swenson, C. C. (1993). Community level intervention after a disaster. In C. F. Saylor (Ed.), *Children and Disasters* (pp. 211–239). New York: Plenum.

Keane, T. M., Fairbank, J. A., & Caddell, J. M. (1989). Implosive (flooding) therapy reduces symptoms of PTSD in Vietnam combat veterans. *Behavior Therapy*, *20*, 245–260.

Keane, T. M., Zimmering, R. T., & Caddell, J. M. (1985). A behavioral formulation of PTSD in Vietnam veterans. *The Behavior Therapist*, *8*, 9–12.

Keppel-Benson, J. M., & Ollendick, T. H. (1993). Posttraumatic stress disorders in children and adolescents. In C. F. Saylor (Ed.), *Children and disasters* (pp. 29–43). New York: Plenum.

Kinzie, J. D., & Leung, P. (1989). Clonidine in Cambodian patients with posttraumatic stress disorder. *Journal of Nervous and Mental Diseases*, *177*, 546–550.

Kinzie, J. D., Sack, W. H., Angell, R. H., Manson, S., & Rath, B. (1986). The psychiatric effects of massive trauma on Cambodian children: I. The children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *25*, 370–376.

Klingman, A. (1993). School-based intervention following a disaster. In C. F. Saylor (Ed.), *Children and disasters* (pp. 187–210). New York: Plenum.

Kolb, L. C. (1988). A critical survey of hypotheses regarding PTSD in light of recent research. *Journal of Traumatic Stress*, *3*, 291–304.

Kolb, L. C., Burris, B., & Griffith, S. (1984). Propranolol and Clonidine in the treatment of the chronic posttraumatic stress disorders of war. In B. Van der Kolk (Ed.), *Posttraumatic stress disorder: Psychological and biologi-*

cal sequelae (pp. 97–107). Washington, DC: American Psychiatric Association.

Kuterovac, G., Dyregrov, A., & Stuvland, R. (1994). Children in war: A silent majority under stress. *British Journal of Medical Psychology*, *67*, 363–375.

Lang, P. J. (1985). The cognitive psychophysiology of emotion: Fear and anxiety. In A. H. Turner & J. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 131–170). Hillsdale, NJ: Lawrence Erlbaum Associates.

Law, J., & Conway, J. (1992). Effect of abuse and neglect on children's speech and language. *Developmental Medicine and Child Neurology*, *34*, 943–948.

Leff, J., & Bhugra, D. (1993). *Principles of social psychiatry*. Oxford: Blackwell.

Lonigan, C. J., Shannon, M. P., Finch, A. J., Daugherty, T. K., & Saylor, C. M. (1991). Children's reactions to a natural disaster: Symptom severity and degree of exposure. *Advances in Behaviour Research and Therapy*, *13*, 135–154.

Lundin, T. (1994). The treatment of acute trauma: Post-traumatic stress disorder prevention. In D. A. Tomb (Ed.), *Posttraumatic Stress Disorder: Psychiatric Clinicas of North America* (Vol. 17, pp. 385–392). Washington, DC: American Psychiatric Association.

Lyons, J. L. (1987). Posttraumatic stress disorder in children and adolescents: A review of the literature. *Developmental and Behavioral Pediatrics*, *8*, 349–356.

Macksoud, M. S., Dyregrov, A., & Raundalen, M. (1993). Traumatic war experiences and their effects on children. In J. P. Wilson & B. Raphael (Eds.), *International handbook of traumatic stress syndromes* (pp. 625–633). New York: Plenum.

Maier, S. F. (1986). Stressor controllability and stress induced analgesia. *Annals of the New York Academy of Science*, *467*, 55–72.

Malmquist, C. P. (1986). Children who witness parental murder: Posttraumatic aspects. *Journal of the American Academy of Child and Adolescent Psychiatry*, *25*, 320–325.

Martin, S., & Little, B. (1986). The effects of a natural disaster on the academic abilities and social behavior of school children. *British Columbian Journal of Special Education*, *10*, 167–182.

McFarlane, A. C. (1987). Family functioning and over-protection following a natural disaster: The longitudinal effects of post-traumatic morbidity. *Australia and New Zealand Journal of Psychiatry*, *21*, 210–218.

McFarlane, A. C. (1994). Helping victims of disasters. In J. R. Freedy & S. E. Hobfoll (Eds.), *Traumatic stress: From theory to practice* (pp. 287–314). New York: Plenum.

McFarlane, A. C., Policansky, S., & Irwin, C. P. (1987). A longitudinal study of the psychological morbidity in children due to a natural disaster. *Psychological Medicine*, *17*, 727–738.

McGaugh, J. L. (1990). Significance and remembrance: The role of neuromodulatory systems. *Psychological Science*, *1*, 15–25.

McLeer, S. V., Deblinger, E., Henry, D., & Orvaschel, H. (1992). Sexually abused children at high risk for post-traumatic stress disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*, 875–879.

McNally, R. J. (1991). Assessment of posttraumatic stress disorder in children. *Psychological Assessment*, *3*, 531–537.

Meichenbaum, D. (1994). *A clinical handbook/practical therapist manual for assessing and treating adults with post-traumatic stress disorder (PTSD)*. Ontario: Institute Press.

Milgram, N. A., Toubiana, Y., Klingman, A., Raviv, A., & Goldstein, I. (1988). Situational exposure and personal loss in children's acute and chronic reactions to a school bus disaster. *Journal of Traumatic Stress*, *1*, 339–532.

Milne, G. (1977). Cyclone Tracy II: The effects of Darwin on Children. *Australian Psychologist*, *12*, 55–62.

Misch, P., Phillips, M., Evans, P., & Berekowitz, M. (1993). Trauma in preschool children: A clinical account. *Association of Child Psychology and Psychiatry: Occasional Papers*, *8*, 11–18.

Mitchell, J. T. (1983). When disaster strikes: The critical incident stress debriefing process. *Journal of Emergency Medicine*, *8*, 36–39.

Mitchell, J. T. (1996, March). *Critical Incident Stress Management: Current status and future directions*. Paper presented at the annual meeting of the European Society for Traumatic Stress Studies conference, Sheffield, UK.

Mitchell, J. T., & Everly, G. S. (1993). *Critical incident stress debriefing: An operations manual for the prevention of trauma among emergency service and disaster workers*. Baltimore, MD: Chevron Publishing.

Mowrer, O. H. (1960). *Learning theory and behavior*. New York: Wiley.

Nader, K. O. (1995). Assessing traumatic experiences in children. In J. P. Wilson & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD: A handbook for practioners* (pp. 217–244). New York: Guilford Press.

Nader, K. O., & Fairbanks, L. (1994). The suppression of reexperiencing: Impulse control and somatic symptoms in children following traumatic exposure. *Anxiety, Stress, and Coping: An International Journal*, *7*, 229–239.

Nader, K. O., Kriegler, J., Keane, T., Blake, D., & Pynoos, R. S. (1994). *Clinician Administered PTSD Scale: Child and Adolescent Version (CAPS-C)*. White River Junction, VT: National Center for PTSD.

Nader, K. O., & Pynoos, R. S. (1989). *Child posttraumatic stress disorder inventory: Parent Interview*. Unpublished manuscript, University of California, Los Angeles.

Nader, K. O., Pynoos, R. S., Fairbanks, L., & Frederick, C. (1991). Childhood PTSD reactions one year after a sniper attack. *American Journal of Psychiatry*, *147*, 1526–1530.

Newman, C. J. (1976) Children of disasters: Clinical observations at Buffalo Creek. *American Journal of Psychiatry*, *133*, 306–312.

Nikapota, A., & Samarasinghe, D. (1993). *Training manual for helping children in situations of armed conflict*. Colombo, Sri Lanka: UNICEF.

Nir, Y. (1984). Post traumatic stress disorder in children with cancer. In S. Eth & R. S. Pynoos (Eds.), *Post traumatic stress disorder in children* (pp. 121–132). Washington, DC: American Psychiatric Press.

O'Donohue, W., & Elliot, A. (1992). The current status of posttraumatic stress disorder as a diagnostic category: Problems and proposals. *Journal of Traumatic Stress*, *5*, 421–439.

Ollendick, T. H. (1983). Reliability and validity of the Revised Fear Survey Schedule for Children (FSSC-R). *Behaviour Research & Therapy*, *21*, 685–692.

Ollendick, T. H., Yule, W., & Ollier, K. (1991). Fears in British children and its relationship to manifest anxiety and depression. *Journal of Child Psychology and Psychiatry*, *32*, 321–331.

Parry-Jones, W. (1996, August). *Planning treatment programmes for traumatised school children*. Paper presented at ACPP 3rd European Conference, Glasgow, Scotland.

Perrin, S., Smith, P., Schwartz, D., & Yule, W. (1996a, November). *Self-report of stress reactions: A cross-cultural validation study*. Paper presented at the 12th Annual Convention of the International Society for Traumatic Stress Studies, San Francisco.

Perrin, S., Yule, W., & Smith, P. (1996b, November). *Post traumatic stress in children as a function of sibling and parental exposure to the trauma*. Paper presented at the 12th Annual Convention of the International Society for Traumatic Stress Studies, San Francisco.

Pitman, R. (1989). Posttraumatic stress disorder, hormones, and memory. *Biological Psychiatry*, *226*, 221–223.

Pitman, R., Orr, S., Forgue, D. F., Altman, B., de Jong, J. B., & Herz, L. R. (1990). Psychophysiologic responses to combat imagery of Vietnam veterans with posttraumatic stress disorder versus other anxiety disorders. *Journal of Abnormal Psychology*, *99*, 49–54.

Pynoos, R. S. (1994). In R. S. Pynoos (Ed.), *Posttraumatic stress disorder: A clinical review*. Lutherville, MD: Sidran Press.

Pynoos, R. S., & Eth, S. (1986). Witness to violence: The child interview. *Journal of the American Academy of Child and Adolescent Psychiatry*, *25*, 306–319.

Pynoos, R. S., Frederick, C., Nader, K. O., Arroyo, W., Steinberg, A., Eth, S., Nunez, F., & Fairbanks, L. A. (1987). Life threat and posttraumatic stress in school-age children. *Archives of General Psychiatry*, *44*, 1057–1063.

Pynoos, R. S., Goenjian, A., Karakashian, M., Tashjian, M., Manjikian, R., Manoukian, G., Steinberg, A. M., & Fairbanks, L. A. (1993). Posttraumatic stress reactions in children after the 1988 Armenian earthquake. *British Journal of Psychiatry*, *163*, 239–247.

Pynoos, R. S., & Nader, K. O. (1988). Psychological first aid and treatment approaches for children exposed to community violence: Research implications. *Journal of Traumatic Stress*, *1*, 243–267.

Pynoos, R. S., & Nader, K. O. (1993). Issues in the treatment of posttraumatic stress in children and adolescents. In J. Wilson & B. Raphael (Eds.), *International handbook of traumatic stress syndromes* (pp. 535–549). New York: Plenum.

Rachman, S. (1980). Emotional processing. *Behaviour Research and Therapy*, *18*, 51–60.

Raundalen, M., Dyregrov, A., & Gronvold Bugge, R. (1990). *Reaching children through teachers: Helping the war traumatized child: A manual*. Maputo: Red Cross.

Reich, W., & Earls, F. (1987). Rules for making diagnosis on the basis of multiple sources of information: Preliminary strategies. *Abnormal Child Psychology, 15*(4), 601–606.

Reich, W., Shakya, J. J., & Taibelson, C. (1991). *Diagnostic Interview for Children and Adolescents (DICA)*. St. Louis, MO: Washington University.

Reynolds, C. R., & Richmond, B. O. (1978). What I think and feel: A revised measure of children's manifest anxiety. *Journal of Abnormal Child Psychology*, *6*, 271–280.

Ribbe, D. P., Lipovsky, J. A., & Freedy, J. R. (1995). Posttraumatic Stress Disorder. In A. R. Eisen, C. A. Kearny, & C. E. Schaefer (Eds.), *Clinical handbook of anxiety disorders in children and adolescents* (pp. 317–356). Northvale, NJ: Jason Aronson.

Richman, N. (1991). *Helping children in difficult circumstances: A teacher's manual*. London: Save the Children.

Richman, N. (1993). *Communicating with children: Helping children in distress*. London: Save the Children.

Rutter, M. (1967). A children's behaviour questionnaire for completion by teachers: Preliminary findings. *Journal of Child Psychology and Psychiatry*, *8*, 1–11.

Sack, W. H., McSharry, S., Clarke, G. N., Kinney, R., Seeley, M. S., & Lewinsohn, P. (1994). The Khmer Adolescent Project: I. Epidemiologic findings in two generations of Cambodian refugees. *Journal of Nervous and Mental Diseases*, *182*, 387–395.

Saigh, P. A. (1987a). *In-vitro* flooding of an adolescent's posttraumatic stress disorder. *Journal of Clinical Child Psychology*, *16*, 147–150.

Saigh, P. A. (1987b). *In-vitro* flooding of a childhood posttraumatic stress disorder. *School Psychology Review*, *16*, 203–211.

Saigh, P. A. (1989a). The use of *in-vitro* flooding in the treatment of a traumatized adolescent. *Journal of Behavioural and Developmental Paediatrics*, *10*, 17–21.

Saigh, P. A. (1989b). The validity of *DSM-III* posttraumatic stress disorder classification as applied to children. *Journal of Abnormal Psychology*, *198*, 189–192.

Saigh, P. A. (1991). The development of posttraumatic stress disorder following four different types of traumatization. *Behaviour Research and Therapy*, *29*, 213–216.

Saigh, P. A. (1992). The behavioral treatment of child and adolescent posttraumatic stress disorder. *Advances in Behaviour Research and Therapy*, *14*, 247–275.

Shaffer, D., Fisher, P., Dulcan, M., & Davies, M. (1996). The NIMH Diagnostic Interview Schedule for Children (DISC-2.3): Description, acceptability, prevalences, and performance in the MECA study. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*(7), 865–877.

Shannon, M. P., Lonigan, C. J., Finch, A. J., & Taylor, C. M. (1994). Children exposed to disaster. I: Epidemiology of posttraumatic symptoms and symptom profiles. *Journal of the American Academy of Child and Adolescent Psychiatry*, *33*, 80–93.

Shapiro, F. (1989). Eye movement desensitization: a new treatment for post traumatic stress disorder. *Journal of Behaviour Therapy and Experimental Psychiatry*, *20*, 211–217.

Shapiro, F. (1995). *Eye movement desensitization and reprocessing: Basic principles protocols and procedures*. New York: Guilford Press.

Schestazky, M. Greenberg, D., & Lerer, B. (1988). A controlled trial of phenelzine in posttraumatic stress disorder. *Psychiatry Research*, *24*, 149–155.

Schwarz, E. D., & Kowalski, J. M. (1991). Posttraumatic stress disorder after a school shooting: Effects of symptom threshold selection and diagnosis by *DSM-III, DSM-III-R,* or proposed *DSM-IV*. *American Journal of Psychiatry, 148,* 592–597.

Schwarz, E. D., & Perry, B. D. (1994). The post-traumatic response in children and adolescents. *Psychiatric Clinics of North America*, *17*, 311–326.

Smith, P., & Yule, W. (1995, May). *The Mostar Programme*. Paper presented at the fourth annual European Conference on Traumatic Stress, Paris, France.

Smith, P., Yule, W., Perrin, S., & Schwartz, D. (1996). *Maternal reactions and child distress following the war in Bosnia*. Paper presented at the 12th Annual Convention of the International Society for Traumatic Stress Studies (November), San Francisco.

Southwick, S. M., Bremner, D., Krystal, J. H., & Charney, D. S. (1994). Psychobiologic research in post-traumatic stress disorder. *Psychiatric Clinics of North America*, *17*, 251–264.

Spielberger, C. D., Gorsuch, R. L., & Luschene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Stallard, P., & Law, F. (1993). Screening and psychological debriefing of adolescent survivors of life threatening events. *British Journal of Psychiatry*, *163*, 660–665.

Sternberg, K. J., Lamb, M. E., Greenbaum, C., Cichetti, D., Dawud, S., Cortes, R. M., Krispin, O., & Lorey, F. (1993). Effects of domestic violence on children's behavior problems and depression. *Developmental Psychology*, *29*, 44–52.

Stuber, M. L., Nader, K. O., Yashuda, P., Pynoos, R. S., & Cohen, S. (1991). Stress response after pediatric bone marrow transplantation: Preliminary results of a prospective longitudinal study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *30*, 952–959.

Summerfield, D. (1996). *The impact of war and atrocity on civilian populations: Basic principles for NGO interventions and a critique of psychosocial trauma projects* (Paper 14). London: Overseas Development Institute Network.

Sutherland, S. M., & Davidson, R. T. (1994). Pharmacotherapy for post-traumatic stress disorder. *Psychiatric Clinics of North America*, *17*, 409–423.

Terr, L. C. (1979). The children of Chowchilla. *Psychoanalytic Study of the Child*, *34*, 547–623.

Terr, L. C. (1983). Chowchilla revisited: The effects of psychic trauma four years after a schoolbus kidnapping. *American Journal of Psychiatry*, *140*, 1543–1550.

Terr, L. C. (1991). Childhood traumas: An outline and overview. *American Journal of Psychiatry*, *148*, 10–20.

Terr, L. C. (1996). Acute responses to external events and posttraumatic stress disorder. In M. Lewis (Ed.), *Child and adolescent psychiatry: A comprehensive textbook* (pp. 753–763). Baltimore, MD: Williams & Wilkins.

Tsui, E. P. (1990). *The "Jupiter" sinking disaster: Effects on teenager's school performance*. Unpublished master's dissertation, Institute of Psychiatry, University of London.

UNICEF (1996). *Annual state of the world's children report*. New York: United Nations Childrens Relief Fund.

Van der Kolk, B. A., Greenberg, M. S., Orr, S. P., Pittman, R. K. (1989). Endogenous opioids, stress induced analgesia, and posttraumatic stress disorders. *Psychopharmacology Bulletin*, *25*, 417–421.

Wolfe, V. V., Gentile, C., Sas, L., & Wolfe, D. A. (1991). Children's Impact of Traumatic Events Scale: A measure of post-sexual abuse PTSD symptoms. *Behavioral Assessment*, *13*, 359–338.

Wolfe, V. V., Gentile, C., & Wolfe, D. A. (1989). The impact of sexual abuse on children: A PTSD formulation. *Behavior Therapy*, *20*, 215–228.

World Health Organization (1992). *International classification of diseases: 10th edition*. Geneva: WHO.

Yehuda, R., Boisoneau, D., Lowry, M. T., & Giller, E. L. (1993). Relationship between lymphocyte glucocorticoid receptor number and urinary-free cortisol excretion in mood, anxiety, and psychotic disorder. *Biological Psychiatry*, *34*, 18–25.

Yehuda, R., Kahana, B., Binder-Byrnes, K., Southwick, S. M., Mason, J. W., & Giller, E. L. (1995). Low urinary cortisol excretion in Holocaust survivors with posttraumatic stress disorder. *American Journal of Psychiatry*, *152*, 982–986.

Yehuda, R., & McFarlane, A. C. (1995). Conflict between current knowledge about posttraumatic stress disorder and its original conceptual basis. *American Journal of Psychiatry*, *152*, 1705–1713.

Yule, W. (1992). Post traumatic stress disorder in child survivors of shipping disasters: The sinking of the "Jupiter." *Psychotherapy and Psychosomatics*, *57*, 200–205.

Yule, W. (1994). Post traumatic stress disorders. In M. Rutter, E. Taylor, & L. Hersov (Eds.), *Child and adolescent psychiatry: Modern approaches* (3rd ed., pp. 392–406). Oxford: Blackwells.

Yule, W., & Gold, A. (1993). *Wise before the event: Coping with crises in schools*. London: Calouste Gulbenkian Foundation.

Yule, W., Perrin, S., & Smith, P. (1996, August). *The effects on pre-school children of catastrophic events*. Paper presented at the Third European Conference of The Association for Child Psychology and Psychiatry, Glasgow, Scotland.

Yule, W., & Udwin, O. (1991). Screening child survivors for post-traumatic stress disorders: Experiences from the "Jupiter" sinking. *British Journal of Clinical Psychology*, *30*, 131–138.

Yule, W., Udwin, O., & Murdoch, K. (1990). The "Jupiter" sinking: Effects on children's fears, depression and anxiety. *Journal of Child Psychology and Psychiatry*, *31*, 1051–1061.

Yule, W., & Williams, R. (1990). Post traumatic stress reactions in children. *Journal of Traumatic Stress*, *3*, 279–295.

Zimmerman-Tansella, C. H., Minghetti, S., Tacconi, A., & Tansella, M. (1978). The children's behavioural questionnaire for completion by teachers in an Italian sample: Preliminary results. *Journal of Child Psychology and Psychiatry*, *19*, 167–173.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.17
Depression

WILLIAM M. REYNOLDS
University of British Columbia, Vancouver, BC, Canada

5.17.1 INTRODUCTION 420
- *5.17.1.1 Depression as an Internalizing Disorder* 420
- *5.17.1.2 Depression as a Problem in Children and Adolescents* 421
- *5.17.1.3 Contemporary Interest* 422

5.17.2 BIOPSYCHOSOCIAL MODEL OF DEPRESSION IN CHILDREN AND ADOLESCENTS 423
- *5.17.2.1 Tripartite Model* 424
- *5.17.2.2 Biopsychosocial Domains and Interactions* 425
  - *5.17.2.2.1 Biological influences* 426
  - *5.17.2.2.2 Psychological influences* 426
  - *5.17.2.2.3 Social influences* 426
- *5.17.2.3 Summary* 427

5.17.3 DIAGNOSTIC FORMULATIONS OF DEPRESSION IN CHILDREN AND ADOLESCENTS 427
- *5.17.3.1 Systems and Approaches in the Diagnosis of Depression* 428
- *5.17.3.2* Diagnostic and Statistical Manual of Mental Disorders 428
  - *5.17.3.2.1 Major depressive disorder* 428
  - *5.17.3.2.2 Dysthymic disorder* 429
  - *5.17.3.2.3 Bipolar disorder* 429
  - *5.17.3.2.4 Cyclothymic disorder* 430
- *5.17.3.3 Alternatives to* DSM*: Clinical Severity, Subsyndromal Depression, and Depressive Symptoms* 430

5.17.4 ASSESSMENT OF DEPRESSION IN CHILDREN AND ADOLESCENTS 431
- *5.17.4.1 Self-report Measures of Depression for Children and Adolescents* 432
  - *5.17.4.1.1 Children's Depression Inventory* 432
  - *5.17.4.1.2 Reynolds Child Depression Scale* 433
  - *5.17.4.1.3 Children's Depression Scale* 434
  - *5.17.4.1.4 Reynolds Adolescent Depression Scale* 434
  - *5.17.4.1.5 Beck Depression Inventory* 435
- *5.17.4.2 Clinical Interview Severity Measures of Depression* 436
  - *5.17.4.2.1 Children's Depression Rating Scale-Revised* 436
  - *5.17.4.2.2 Hamilton Depression Rating Scale* 436
- *5.17.4.3 Assessment Summary* 437

5.17.5 PSYCHOLOGICAL TREATMENT OF DEPRESSION IN CHILDREN AND ADOLESCENTS 437
- *5.17.5.1 Range of Therapeutic Approaches* 437
- *5.17.5.2 Control-group Investigations* 438
- *5.17.5.3 Psychotherapy Treatment Studies with Children* 438
- *5.17.5.4 Psychotherapy Treatment Studies with Adolescents* 442

5.17.6 PHARMACOTHERAPY FOR DEPRESSION IN CHILDREN AND ADOLESCENTS 444
- *5.17.6.1 Pharmacotherapy for Depression in Children and Adolescents* 445
- *5.17.6.2 Summary of Pharmacological Interventions* 445

5.17.7 COMORBIDITY OF DEPRESSION AND OTHER PSYCHOPATHOLOGY IN CHILDREN AND ADOLESCENTS 446
- *5.17.7.1 Comorbidity and Biological Communalities* 447
- *5.17.7.2 The Range of Disorders Comorbid with Depression* 447
- *5.17.7.3 Comorbidity with Other Internalizing Disorders and Problems* 448
  - *5.17.7.3.1 Anxiety disorders* 448

5.17.7.3.2 Eating disorders 448
5.17.7.4 Comorbidity with Externalizing Disorders and Problem Behaviors 449
5.17.7.4.1 Conduct disorder 449
5.17.7.4.2 Attention deficit hyperactivity disorder 450
5.17.7.4.3 Substance abuse 450
5.17.7.5 Diagnostic Overshadowing 451
5.17.7.6 Comorbidity Summary 451
5.17.8 SUMMARY AND CONCLUSIONS 452
5.17.9 REFERENCES 452

5.17.1 INTRODUCTION

It can be estimated that depression affects onwards of two million children and adolescents in the USA and Canada. As an internalizing disorder in young people it is a source of intense misery and distress to those affected. By its nature, depression is often a silent disorder that generally affects few persons other than the depressed youngster. Because of its internalizing nature, we may view depression as an insidious psychological disturbance in young people. Thus, for some youngsters depression may be inconspicuous yet have deleterious effects on the child.

Depression is a disorder and disturbance of mood that has been described from ancient times with continued efforts to clarify the nature of this aspect of the human condition. Although we no longer ascribe the etiology of depression, or melancholia, to the function of bodily humors such as black bile or "perturbations of the soul" (Jackson, 1986), there are many commonalties between ancient and contemporary descriptions of the phenomenology of depression. In the early part of the twentieth century, a number of psychoanalysts sought to explain depression in light of the prominent analytic theories of the time, as seen in the work of Karl Abraham, Sigmund Freud, Sandor Rado and, somewhat later, Melanie Klein. Although not specific to depression, of particular note is the pioneer work of Adolf Meyer, whose formulations of psychopathology may be seen as approximating a biopsychosocial perspective.

More recently, the study of depression has encompassed many diverse fields of scientific inquiry. Psychology and psychiatry are those fields that typically come to mind, but a wide-ranging review of the fields that have examined depression include: biochemistry, genetics, social work, nursing, counseling, business, epidemiology, pharmacology, neurology, pediatrics, general medicine, biology, zoology, sociology, education, physiology, endocrinology, and other disciplines. There are well over 1000 research and clinical articles on depression in children and adolescents that have been published in English language journals, with hundreds of new studies appearing each year. The range of investigations and domains covers a vast spectrum of the biopsychosocial sciences, from research on the activity of killer cells in depressed adolescents and the relationship between killer cell activity and severity of self-reported depression to studies of maternal attachment behavior and depressive affect in preschool children.

As described in this chapter, the wide range of scientific disciplines that have focused on depression in children and adolescents supports an integrated perspective for the understanding of this disorder in young people. This perspective may be illustrated by a biopsychosocial model of depression. The biopsychosocial model provides a platform or framework for exploring the nature and enhancing our understanding of depression by examining the complex interplay between the biological endowment and characteristics of the individual, the person's psychological competence, and the psychosocial stressors that may affect the individual. The biopsychosocial perspective is not unique to the study of depression and has been applied to a wide range of human conditions and problems.

5.17.1.1 Depression as an Internalizing Disorder

In the conceptualization of psychopathology in children and adolescents, a predominant theme over the past two decades has been the differentiation of disorders as a function of their general phenotypic expression. This has lead to the categorization of disorders as primarily internalizing or externalizing, based on their general symptom manifestation (Reynolds, 1992a). Internalizing disorders are those that tend to be composed of relative covert symptoms, many of which are cognitive, internal to the youngster, or relatively difficult to observe. Externalizing disorders are overt, highly observable either directly or indirectly (such as in the outcome of certain antisocial behaviors), and typically expressed as behavioral excess.

Depression and depressive (mood) disorders in children and adolescents may be viewed as internalizing disorders. Depression as a psychological disorder is replete with symptom characteristics that are internal to the individual. These features include symptoms of cognitive, emotional, behavioral, and physiological impairment or dysfunction. Depression is more than just a low or dysphoric mood such as feeling sad, blue, or down in the dumps. Depression typically affects multiple areas of individual functioning, including behavioral, emotional, somatic, and cognitive domains. Depressive disorders represent serious forms of psychopathology of childhood and adolescence, given that some depressive episodes may lead to potentially life-threatening or negative outcomes. Previous notions of "adolescent turmoil" or the perspective of the adolescent who is "just going through a moody stage" are no longer viable conceptualizations (Offer & Schonert-Reichl, 1992). This is amply evident when one considers the large numbers of depressed and suicidal youth, a significant number of whom do not survive to adulthood or do so with significantly reduced psychosocial competence or functioning.

Overt behavior problems in youngsters have typically been characterized by disorders of behavioral excesses, and as such are often viewed as externalizing problems. Externalizing disorders, such as conduct disorders and attention-deficit hyperactivity disorder, represent classes of disorders that frequently come to the attention of classroom teachers or parents. Symptoms of these disorders are typically represented by behavioral excesses that interfere with normal classroom behavior or social interactions. Contrary to this, internalizing disorders such as depression are often undetected due to the more covert nature of their symptoms. For instance, feelings of subjective misery, hopelessness, fatigue, anhedonia, difficulty sleeping, and other symptoms of depression may not be readily observable, particularly in youngsters who may be characterized as shy or somewhat withdrawn. Depression differs from other behavior disorders of childhood and adolescence in its comorbidity with suicidal risk.

5.17.1.2 Depression as a Problem in Children and Adolescents

Depression in children and adolescents, within the scope of current nomenclature and systems of classification, is considered to be a mood disorder similar in many respects to this disorder in adults. From another perspective, depression in children and adolescents may be considered as significant perturbation of mood. A depressed youngster may experience a range of symptoms, some of which may be overt, such as irritability or distinctly sad appearance, and others may be covert, as illustrated by feelings of low self-worth, hopelessness, suicidal thoughts, and guilt. In young people, depression represents a negative psychological state, that may be characterized as one of intense, subjective misery, despondency, and in some youngsters, irritability. Depression in children and adolescents can cause significant impairment in daily functioning and personal and social involvement (Puig-Antich et al., 1985). Children who are depressed are often withdrawn from the world around them, finding little pleasure in positive events or settings. Treats and previously fun activities are no longer desirable or sufficient to lighten their spirits. In young children, this affective state may be characterized as being moody or distant from others and slow to respond to social interactions with peers and family. Adolescents who are depressed often feel alone, helpless, and frustrated with friends and family who do not seem to understand them. Although this may characterize a great many adolescents, the depth and quality of these characteristics in depressed adolescents, and in particular, their lack of positive response to previously reinforcing events or activities generally distinguishes depression in adolescents from more normative trials and tribulations associated with the normative course of adolescence.

Some children and adolescents who are depressed also manifest suicidal thoughts and behaviors. This is particularly problematic and in some cases leads to death or physically disabling outcomes. Epidemiological data suggests that suicidal behaviors are a serious problem among children and adolescents (Reynolds & Mazza, 1994). Although some youngsters who demonstrate suicidal thoughts and behaviors do not manifest clinical levels of depression, the majority of suicidal youngsters are depressed, with many who view suicide as a way of ending their intense psychological distress.

Depression can take a variety of courses and chronocity, with relapse and recurrence relatively common over the lifespan (Keller, 1994; Kovacs, 1996a). Kovacs and colleagues, in a noteworthy longitudinal study of depression in children, found that for some youngsters, particularly those with a dysthymic disorder, their depression may last for many years, and in some children represents a precursor or risk factor for other psychiatric disorders (Kovacs, 1985; 1996b; Kovacs, Feinberg, Crouse-Novak, Paulauskas, Pollock, & Finkelstein, 1984a). Similarly, in a retrospective study of children who attended the Maudsley Hospital in London, UK, Harrington, Fudge, Rutter, Pickles, and Hill (1990) investigated the adult psychiatric status of 52 participants who initially demonstrated a significant number of depressive symptoms and 52 matched nonpsychiatric controls. They found that 40% of the childhood depressed group met Research Diagnostic Criteria (Spitzer, Endicott, & Robins, 1978) for depression within five years of entering adulthood, with 58% demonstrating a history of an affective disorder at the time of follow-up compared to 25% of the control sample. These investigations suggest that depression in children and adolescents may have long-term negative consequences.

Wamboldt (1994) examined the nature of medical and psychiatric disorders reported in child and adolescent patients seen at nine medical-psychiatric settings that are considered Type III or IV (moderate to severe medical and psychiatric problem treatment) child and adolescent medical-psychiatric units in the USA. With regard to psychiatric disorders, Wamboldt found a greater prevalence of internalizing disorders than externalizing disorders, with 32% of the patients demonstrating mood disorders, primarily depression. These data are based on a sample of 768 youngsters admitted during 1991 and who manifested both medical and psychiatric problems.

A number of studies suggest what appears to be a secular trend in the age of onset of depressive disorders in children and adolescents, with first episodes of depression found in increasingly younger children (Kovacs & Gatsonis, 1994; Lewinsohn, Rohde, Seeley, & Fisher, 1993). This is a significant point of concern and buttresses the need to recognize depression in young people at early stages and provide intervention or prevention.

5.17.1.3 Contemporary Interest

Contemporary interest in the study of depression in children and adolescents began in the late 1970s and early 1980s. Prior to this, the predominant perspectives in the literature were that depression in children was: nonexistent, masked, or expressed in symptoms that were significantly different from those found in depressed adults (Reynolds, 1985). Depression in children was also conceptualized by the construct of depressive equivalents (Glaser, 1967; Hollon, 1970). In this manner, depression in children was inferred by overt, negative behaviors such as acting-out, aggression, hyperactivity, and delinquency, as well as psychosomatic and hypochondriacal disorders (Lesse, 1981). Some exceptions to these perspectives were postulated (e.g., Sandler & Jaffe, 1965) although for a number of reasons, did not significantly influence the existing status of depression in childhood (Bemporad, 1994).

Early research and clinical interest in depression in children and adolescents also focused on bipolar disorder. This was generally seen as an early onset of a typically adult disorder. The extreme nature of symptoms in manic depressive illness and mania, and the similarity between symptoms of these disorders in children and adolescents to those found in adults, led to a number of clinical reports (e.g., Beres & Alpert, 1940; Campbell, 1952). Carlson (1994) notes that Kraepelin in 1921 was one of the first to describe manic depressive illness in a child. However, research reports on manic depressive illness in children and adolescents are scarce before the 1970s.

In the 1970s, a small but significant number of clinical and research reports appeared that suggested the importance of studying depression in young people. An important article that shed light on the existence of depression in childhood was published in 1970 by Pozanaski and Zrull who described clinical cases gleaned from the outpatient files of the Children's Psychiatric Hospital at the University of Michigan Medical Center from the years 1964 to 1968. Based upon an examination of records of 1788 children, 98 cases were identified that showed significant evidence of depression, with 14 cases providing sufficient file data for in-depth examination and reporting. Children ranged in age from 3 to 12 years. Primary symptoms of depression in these youngsters included: sad, unhappy, depressed appearance; withdrawal; expressed feelings of rejection and being unloved; and autoerotic behavior. Poor self-image and self-worth were predominant as a depressive symptom.

Weinberg and his colleagues are frequently associated with the development of the first contemporary diagnostic criteria for depression in children (Ling, Oftedal, & Weinberg, 1970; Weinberg, Rutman, Sullivan, Penick, & Dietz, 1973). The Weinberg criteria were an outcome of the observation of depressive symptoms in children that were phenomenologically similar to symptoms of the disorder as specified for adults. An examination of the symptoms and criteria delineated by Weinberg et al. (1973) shows a great deal of communality with contemporary diagnostic formulations. In the late 1970s, Petti (1978) developed the Bellevue Index of Depression (BID), a semistructured clinical interview of depression designed for use with children based on Weinberg symptomatology.

In the early 1980s researchers also began to test the viability of unique conceptualizations of depression in childhood. Christ, Adler, Isacoff, and Gershansky (1981) investigated the concept of masked depression by examining clinical records of over 10 000 psychiatrically hospitalized children and adolescents. They compared diagnoses of depression (psychotic and neurotic) with: symptoms of depression (suicidal behavior, sleep disturbance, appetite changes, fatigue, etc.) specified by contemporary classification systems; and hypothesized masked features of depression as indicated by serious aggression and misbehavior. The authors found that youngsters with a diagnosis of a depressive disorder demonstrated greater numbers of contemporary symptoms of depression, while masked features were negatively related to this diagnosis, and occurred at a lower frequency in depressed youngsters as compared to youngsters with other psychopathology.

Around this time period, a number of reports began to emerge as to the clinical use of adult antidepressant medications with children. Antidepressant medications were relatively new but, within 10 years of early reports with adults, Frommer (1967) described the use of antidepressants with children. Additional reports by Kuhn and Kuhn (1971) and Stack (1971), among others, soon followed.

In 1980, the American Psychiatric Association (APA) published the third edition of the *Diagnostic and statistical manual of mental disorders* (*DSM-III*; APA 1980). Depressive (mood) disorders where listed in the adult disorders section, differential symptom criteria and duration specifications for major depression and dysthymic disorder, respectively were provided for these disorders in children and adolescents as compared to the disorders in adults. The publication of *DSM-III*, as well as research findings such as those by Christ et al. (1981), led to a focus on symptoms of depression in children and adolescents as similar in many regards to adults.

The acceptance of the largely adult specifications for depression in the *DSM* is not universal nor, as suggested by Nurcombe and colleagues (Nurcombe, 1994; Nurcombe et al., 1989), strongly supported by research evidence. Nurcombe (1994) examined the validity of major depression as a diagnosis in children and adolescents using as criteria: natural history, psychobiological markers, genetic studies, response to treatment, and construct validity. The last was based on multivariate analyses of various checklists and measures. As presented, much of the evidence is mixed, suggesting that depressive disorders in children may not be as currently defined by *DSM* with more work needed for the presentation of diagnostic validity of the disorder.

Based on diagnostic and epidemiological surveys of children and adolescents it is evident that depression is a significant problem among youngsters (Reynolds & Johnston, 1994). Epidemiological studies suggest that between 3% and 5% of children and 7–10% of adolescents manifest clinical levels of depression (Fleming & Offord, 1990; Lewinsohn, Hops, Roberts, Seeley, & Andrews, 1993; Poznanski & Mokros, 1994; Whitaker et al., 1990), with lifetime rates in young people significantly higher (Lewinsohn et al., 1993). These data, and the general undertreatment of depression in children and adolescents (Keller, Lavori, Beardslee, Wunder, & Ryan, 1991) point to the critical need for study, identification, and treatment of depression in young people.

5.17.2 BIOPSYCHOSOCIAL MODEL OF DEPRESSION IN CHILDREN AND ADOLESCENTS

In psychiatry, the biopsychosocial perspective represents a shift from a biomedical model, with its focus primarily on the biology of disease or disorders, to one of integration of psychological and dynamic characteristics, social and environmental forces, and biological bases. The biopsychosocial model is useful for the study and understanding of depression in children and adolescents, as well as adults. The biopsychosocial perspective has also been applied to numerous other aspects of human behavior and disorders (Paris, 1994; Shalev, 1993).

The psychobiological approach to the study of human behavior and problems has a long history (Dewsbury, 1991), and is illustrated in its earlier presentation by the integrated, commonsense perspective of Adolf Meyer. Meyer's perspective of psychiatric disturbance as due in part to, or exacerbated by environmental causes, can be viewed as a precursor to the current biopsychosocial perspective. Meyer's integrated approach to psychiatry, which he called psychobiology (Meyer, 1957), was also interactive, in that psychiatric illness was viewed as an interaction (or reaction) between the psychobiological and life stressors. Likewise, the biobehavioral model of depression proposed by Akiskal (1979) is consistent with this perspective and integrates biological, psychosocial, developmental, and environmental stressors as a basis for the development of depression.

The biopsychosocial model for the provision of general medical services was espoused by

Engel (1978, 1980). Presented as a systems approach in which the person does not exist in isolation, the biopsychosocial model was proposed in part to deal with the general over-reliance of physicians on the biomedical model in which the focus was on disease with limited concern about the patient as a person. In this manner, the biopsychosocial model or perspective is similar in its general focus to the approach of Meyer in psychiatry, and provides a useful heuristic for developing an understanding of depression in children and adolescents.

The biopsychosocial perspective, although not extensively practiced, has been a relatively popular concept within the psychiatric treatment literature because it suggests an integrated systems approach for therapy. Karasu (1990) in his comparison and integration of models of psychotherapy (dynamic, cognitive, and interpersonal) for depression, has provided suggestions for the indication of pharmacotherapy, psychotherapy, or their combination for the treatment of this disorder. The combinatory approach to treatment conforms to the "biopsychological perspective" of childhood depression described by Kashani and Cantwell (1983).

The basic applications of this perspective are not new to the study of depression in children and adolescents. Lewis and Lewis (1979) described a psychobiological perspective on depression that recognized psychoanalytic, genetic, and biochemical bases for the development and subsequent treatment of depression in children. Vasile et al. (1987) described a "biopsychosocial approach" for treating persons with depression that implements a consultation model for the evaluation and recommendation of psychological, pharmacological, or combined therapies. Although specific to adults, this integrative and evaluative approach appears useful for the treatment of depressed children and adolescents.

The biopsychosocial model of depression delineated by Reynolds (1997a) and briefly described below provides a perspective for understanding depression in children and adolescents as well as a framework for intervention and the provision of services. Within this model, biological, as well as psychosocial, factors play a role in child and adolescent depression. In some cases, there is a greater emphasis exerted by either the biological or psychosocial, indicating a primary modality for treatment of the disorder. However, the interactive nature of the model suggests that multimodal treatments may show the greatest efficacy, both in initial treatment gains and potential reduction of reoccurrence.

5.17.2.1 Tripartite Model

The biopsychosocial perspective on depression in children and adolescents may be represented as a tripartite dynamic model (Reynolds, 1997a). This model is illustrated in Figure 1. Similar to the model of reciprocal determinism postulated by Bandura (1978) for understanding human behavior, the biopsychosocial prospective (described below) considers reciprocal pathways. The strength of these pathways in the pathogenesis, maintenance, or exacerbation of depression or depressive disorders may vary as function of developmental level, individual strengths and deficits within each domain, nature of the reciprocal interactions, and the nature of the depressive outcome.

In the model, biological, psychological, and social–environmental attributes and influences are closely intertwined with intersecting pathways unique to each individual. Given the myriad combinations and interactions, the untangling of these main effects, interactions, and reciprocal paths for an individual is well beyond the scope of current science and practice. However, it is likely that, with sufficient scrutiny, major sources and some level of interactions across the three primary domains can be investigated to gain insight into the nature, potential etiology, exacerbating and maintaining factors, and the possible course and outcome of the individual's affective status.

Two aspects that are not depicted, yet which are critical, are that this model is dynamic and developmental. There is evidence to infer that depression in children is, in a number of respects, different from depression in adults, although there are probably more communalities than differences. There is also the likelihood that some differences are not specific to diagnostic symptomatology, but in its internal structure. Likewise, there are a number of possible developmental trajectories that may relate to differences in the nature as well as expression of depression in youngsters as compared to adults. It may be that biological components of depression become more pronounced as the human organism matures. This may be most evident in findings from neuroendocrine research that has focused on the relationship between hypothalamic–pituitary–adrenal axis function and depression in children (Puig-Antich et al., 1989). Overall, research examining neurobiological correlates in depressed children compared to adults has been mixed (Emslie, Weinberg, Kennard, & Kowatch, 1994).

The psychological domain includes cognitive and behavioral competencies and deficits that are linked to the ontogenesis of depression. As

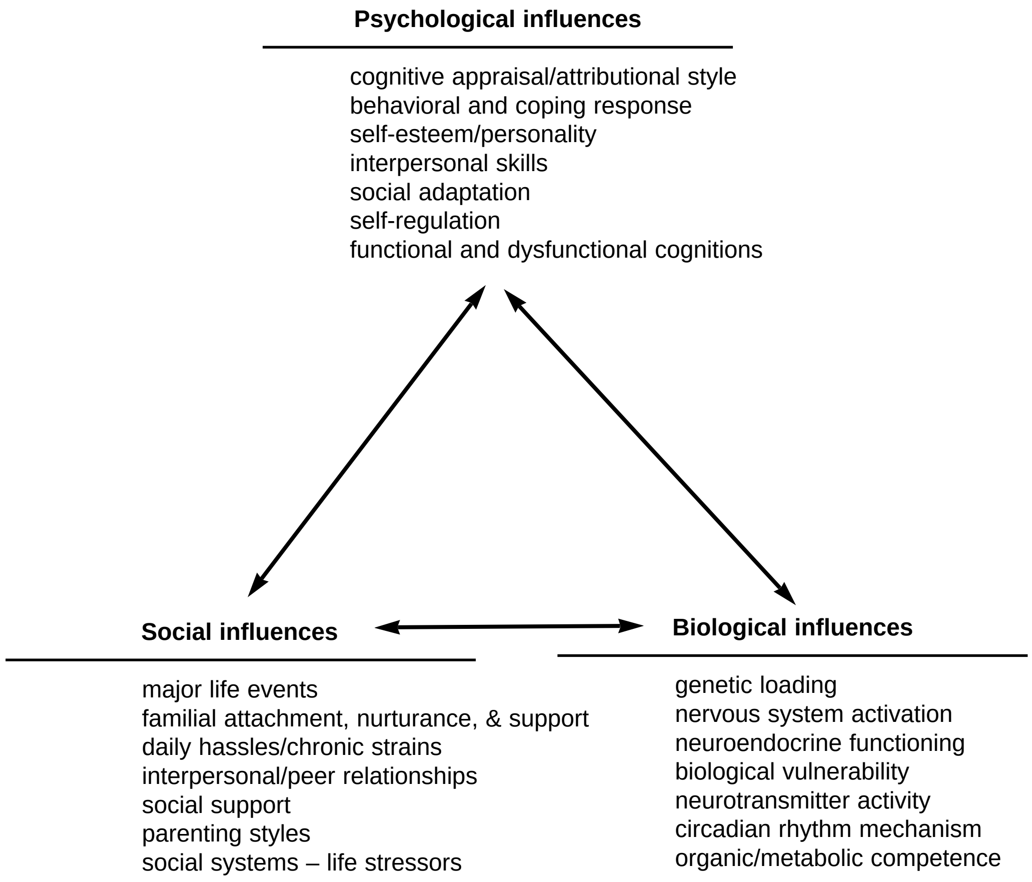
Psychological influences
cognitive appraisal/attributional style
behavioral and coping response
self-esteem/personality
interpersonal skills
social adaptation
self-regulation
functional and dysfunctional cognitions
Social influences
major life events
familial attachment, nurturance, & support
daily hassles/chronic strains
interpersonal/peer relationships
social support
parenting styles
social systems – life stressors
Biological influences
genetic loading
nervous system activation
neuroendocrine functioning
biological vulnerability
neurotransmitter activity
circadian rhythm mechanism
organic/metabolic competence

Figure 1 Biopsychosocial model of depression (Reynolds, 1997a).

models of depression suggest, these components do not act in isolation, but are set in motion by the experience of socio-environmental factors, such as stress, nurturance, and attachment, along with a biological propensity or interaction. However, the cognitive and behavioral/coping functioning of the individual will determine to some extent the impact of the social system (i.e., interpersonal stressors, minor/major life event, etc.) on affective competence and biological/neuroendocrine response/interaction. Likewise, dysfunctional cognitions and behavioral response can be a function of neuroendocrine influences (Wright & Thase, 1992).

Magnusson (1985) has described an interactional paradigm for examining aspects of human functioning that integrates psychology and biology as important person factors in examining person by environment interactions and is consistent with the biopsychosocial perspective. As such, a longitudinal approach is needed to understand the mechanism of human behavior. In this regard, development plays a critical role in the biopsychosocial perspective and can be considered to be an important component of psychology, biology, and the social domain. Magnusson and colleagues (Magnusson & Torestad, 1993; Stattin & Magnusson, 1989) have delineated this interactional approach with a major focus on the developmental and biological nature of person, and in particular adolescent and environment interactions, cumulating in the description of a holistic, dynamic model of personality. This approach also has utility for the study and understanding of depression in young people.

5.17.2.2 Biopsychosocial Domains and Interactions

The biopsychosocial model presented in Figure 1 is a simplified representation of the primary domains and their potential interactions. In reality, the level of interactions of both direct and indirect effects of individual components within and between domains is, in most situations, quite complex and, given the potential numbers and measurability of these variables, is beyond the scope of most research investigations. Remschmidt (1992) notes that the interactions between biological and psychosocial variables that influence the development of psychopathology in children and adolescents may be multiplicative as well as additive.

It is important to recognize that both strengths and weaknesses may be present within each domain. Thus, a youngster may show a deficit in the interpersonal skills domain as a psychological influence, or may demonstrate competence in interpersonal skills that may reduce vulnerability for depression. In a similar vein, personality may enhance psychological resilience or act as a factor that exacerbates biological or social–environmental events. Likewise, biological competence may provide a protective factor for adverse or stressful major life events.

Within the scope of this chapter it is not possible to provide delineation of the main effects of factors associated with the biological, psychological, and social–environmental domains, and the researched interactions between these domains and depression in children and adolescents. Research shows that all three domains play a role in the manifestation of depression in children and adolescents. Described below are a few examples of each domain of the model.

5.17.2.2.1 Biological influences

Biological bases of depressive disorders have been extensively examined in adults and to a more limited extent children and adolescents (Emslie et al. 1994). Domains of inquiry range from brain neurochemistry, such as serotonin mechanism, adrenal-hyper-cortical functioning and activity, and sleep architecture to genetics. Also included in this research domain is the study of antidepressants. For example, a great deal of research in psychiatry has focused on serotonergic mechanisms, specifically the role of serotonin (5-HT) in the regulation of psychological and biological functioning. This research has had a major impact on the development of new classes of pharmacological agents for the treatment of depression, such as antidepressant medications that are selective in their serotonin reuptake properties. These latter agents include such drugs as fluoxetine (Prozac), sertraline (Zoloft), and paroxetine (Paxil). Within the model, biology affects as well as interacts with psychosocial factors to increase vulnerability for the development of maladaptive psychological processes and adaptation.

Likewise, a great deal of interest and research has focused the hypothalamic–pituitary–adrenocortical (HPA) axis as a biological focal point in the environmental stress–depressive response pathway. A number of biological systems and factors, such as prolactin, growth hormone, melatonin metabolism, sleep, and neurotransmitter activity have been linked as of potential etiological relevance to depression (Leibenluft, Fiero, & Rubinow, 1994). There is also evidence to suggest a genetic loading for various depressive disorders, such as major depression and bipolar disorder, with bipolar disorder having the greater genetic component (Herdman, Gough, Liskowski, & Hall, 1995).

5.17.2.2.2 Psychological influences

The domain of psychological influences may best be described as the psychological characteristics, both strengths and weaknesses, as well as vulnerabilities that make up the youngster's emotional and mental health competence. In Figure 1 a number of individual characteristics are listed to provide a general description of this domain. In reality, there are many other psychological influences, some of which are well defined by research and theory, and others that are less well defined or understood. These factors may be either protective or risk factors.

A number of cognitive and behavioral processes are particularly relevant to the understanding of the development of depression. Some of these components are linked to theories of depression. For example, the role of self-regulation (Kanfer, 1970) as it relates to problems in self-control (self-monitoring, self-evaluation, self-reinforcement) has been postulated by Rehm (1977) as model of depression. Problems or deficits in cognitive appraisal, maladaptive behavioral response to reinforcement, dysfunctional cognitions, and erroneous attributional style have been identified by various contemporary theories and models as linked to the development of depression (Abramson, Seligman, & Teasdale, 1978; Beck, 1967; Ellis, 1962; Lewinsohn, 1975).

5.17.2.2.3 Social influences

The social component represents the social–environmental domain and by its nature includes a very wide range of influences. Included here are: stressful events, including major negative life events as well as daily hassles and chronic strains; parental influences, including attachment, nurturance, marital discord, support, and parenting style; social support; peer and interpersonal relationships; and a host of social and environmental influences such as exposure to violence, victimization, maltreatment, social disadvantage, peer and social pressures, and school and work stress among others.

As an example, parental nurturance is a broad domain that includes parenting behavior such as the nature and use of discipline (e.g., consistent/inconsistent, reasonable/harsh, etc.); communication style; over/under controlling; parental warmth; and other behaviors, characteristics, and competencies. Straus and Kantor (1994), in a retrospective study of adults queried as to having received corporal punishment during their teen years, found that corporal punishment was significantly related to increased risk for depression, even when controlling for factors such as socioeconomic status (SES), gender, age, and parental marital violence. Oakley-Browne et al. (1995) also found that adverse parenting in childhood had an increased risk for depression in adulthood. It is also important to note that, during adolescence, there is often a transition in the relationship that youngsters have with parents and peers, with peers becoming increasingly important as a source for support and attitude development.

5.17.2.3 Summary

In the biopsychosocial model, the three primary domains (biology, psychology, and social) interact to produce depression in persons. The nature, course, severity, treatment viability, and other characteristics of depression in young people as well as adults is a function of the nature of the interactions and the domain-specific characteristics and vulnerabilities. This model is also reciprocal, in that factors within the domains may interact as well as affect other domains in a reciprocal manner.

For example, some psychosocial stressors have been noted to be responsible for suppressed T-lymphocyte function (Brosschot et al., 1994; Syvalahti, 1987) and a host of hormonal changes. Such abnormalities in immune function and endocrine changes can increase vulnerability to depression. In adolescents, major negative life events have been linked to reduced immune function and neuroendocrine changes (Birmaher et al., 1994).

As the interactional approach may suggest, it is virtually impossible to separate out the variances associated with any one domain, given the potential complexities of the interactions within as well as between domains. However, researchers are beginning to recognize the multivariate nature of these domains as a basis for the study of depression. In a study examining the prediction of major depression in a sample of 1360 women (680 pairs of twins), Kendler, Kessler, Neale, Heath, and Eaves (1993) examined the main effects (both direct and indirect) of a number of biological (e.g., genetic), social (e.g., stressful life events), and psychological (e.g., neuroticism) factors.

The biopsychosocial model suggests there are multiple pathways to the development of depression in children and adolescents. From this perspective, psychological, social and biological models have validity but are individually insufficient to explain the range of etiologies associated with depression.

5.17.3 DIAGNOSTIC FORMULATIONS OF DEPRESSION IN CHILDREN AND ADOLESCENTS

The diagnosis of depression involves the evaluation and comparison of the characteristics and duration of an individual's symptoms with criteria and classification rules for one or more disorders. In this manner, diagnosis constitutes an assessment or measurement procedure. Thus, measurement error, or the reliability of the diagnosis is an important consideration. Sources of potential measurement error in making a diagnosis include error due to the informant, the diagnostic instrument (e.g., clinical interview), the diagnostician, as well as the classification system. In some respects, a diagnosis may be viewed as a hypothesis that represents "the clinician's best judgment" (Rush, 1990, p. 5) of the individual's condition. Research suggests that the diagnosis of psychopathology and, more specifically, depressive disorders in children is not perfect (APA, 1980; Vitiello, Malone, Buschle, Delaney, & Behar, 1990). For the evaluation of depression in children and adolescents, diagnosis is one method with the outcome being a characterization of specific disorder or disorders based on a formal classification scheme. Other measurement procedures, such as the assessment of severity of depressive symptoms, either by clinical interview or self-report without comparison to a classification system, are often used in clinical and research applications with children and adolescents. The latter may be viewed as providing a focus on the assessment of depression as a syndrome rather than delineating categorically distinct depressive disorders. As noted by Nurcombe (1994), the perspective of depression as a syndrome is postulated on the specifications of symptoms and signs that are predominated by dysphoric mood.

The classification systems currently used in formulating psychiatric diagnoses are continually being refined. In this sense, such systems are not static, immutable, or complete in their description of inclusion and exclusion criteria for specific disorders. This iterative process is exemplified by ongoing revisions in the *DSM* by the APA and the *International statistical classification of diseases and related health problems* (*ICD*) by the World Health Organization (WHO, 1992). These are the primary but not the only systems or approaches to the classification of depressive disorders. It is also important to recognize that depressive symptoms may occur in a number of medical and neurological conditions (Klerman, 1987; Rush, 1990) that need to be considered in the diagnosis of a depressive disorder. Several systems that have been used to diagnose depression in children and adolescents are described below.

5.17.3.1 Systems and Approaches in the Diagnosis of Depression

There are a number of classification systems that have been developed or provide for the diagnosis of depressive disorders in children and adolescents. Most of the current systems are based on a taxonomic approach for the differentiation of type and subtypes of depressive disorders based on particular clusters of symptoms and symptom duration. In North America the primary system used for the classification of depressive disorders in children, adolescents, and adults is the *DSM-IV* (APA, 1994). The *DSM-IV* criteria for depression in children and adolescents are described in a later section.

For certain research applications, the Research Diagnostic Criteria (RDC), developed by Spitzer et al. (1978) is used for the diagnosis of depressive subtypes. The RDC provides greater specificity of subtypes of depression and may be evaluated using a formal diagnostic clinical interview, the Schedule for Affective Disorders and Schizophrenia (SADS; Endicott & Spitzer, 1978) for adults as well as adolescents, and the kiddie SADS for children (K-SADS; Puig-Antich & Chambers, 1978), which has also been used for adolescents (e.g., Lewinsohn, Rohde, & Seeley, 1995). These interview schedules are specific to the RDC in the diagnoses they provide, although research investigators have developed algorithms for deriving *DSM* diagnoses from SADS interviews.

As noted earlier, Poznanski and Zrull (1970) using a chart-review method, provided descriptions of clinical symptoms constituting depression in children. As such, these symptoms were viewed as a basis for the diagnosis of depression in youngsters, with the general premise being that depression as a disorder did manifest itself in childhood (Poznanski, Krahenbuhl, & Zrull, 1976, Poznanski & Zrull, 1970). Later, Poznanski and her colleagues (Poznanski, 1982; Poznanski, Mokros, Grossman, & Freeman, 1985) reformulated their clinical description of symptomatology for depression in children, and developed a clinical interview for the assessment of depression in children (Poznanski, Cook, & Carroll, 1979; Poznanski et al., 1984).

In the early 1970s, Weinberg and colleagues proposed a set of criteria for depression in children (Weinberg et al. 1973) with Weinberg and Brumback (1976) providing diagnostic criteria for mania in children. The Weinberg criteria for depression include (i) the presence of both dysphoric mood and self-deprecatory thoughts; (ii) two or more of: aggressive behavior, sleep disturbance, change in school performance, diminished socialization, change in attitude toward school, somatic complaints, loss of usually energy, and change in appetite and/or weight; and (iii) symptoms need to be present for more than one month and represent a change in the child's usual behavior. Although the Weinberg criteria are rarely used today, they are noteworthy for their early impact as well as their similarity to symptom specifications for current *DSM-IV* criteria for depressive disorders in young people.

5.17.3.2 *Diagnostic and Statistical Manual of Mental Disorders*

Depressive disorders in the *DSM-IV* fall under the category of mood disorders, and include major depression, bipolar disorders, cyclothymic disorder, and dysthymic disorder as primary mood disorders. Disorders are differentiated by specific clusters of symptoms, age of onset, and the duration of symptoms. Disorders that are most likely to be found in school-aged youngsters are major depression and dysthymia, although adjustment disorder with depressed mood is also diagnostically relevant. Bipolar disorders also demonstrate an onset during adolescence in a significant number of individuals who develop this disorder (Loranger & Levine, 1978; McGlashan, 1988). The depressive disorders subsumed under the domain of mood disorders in the *DSM-IV* and their associated course and subtype specifiers are presented in Table 1. It should also be noted that diagnostic categories of: mood disorder due to medical condition; substance-induced mood disorder; and mood disorder not otherwise specified may also be diagnosed in children and adolescents.

5.17.3.2.1 Major depressive disorder

Major depressive disorder (MDD) in *DSM-IV* is presented as a relatively severe, acute form of depression. Major depressive disorder may be classified as a single episode or recurrent, the

Table 1 *DSM-IV* Mood Disorders in Children and Adolescents.

- Major depressive disorder
  - single episode[1–6]
  - recurrent[1–8]
- Dysthymic disorder (early onset)[2]
- Depressive disorder not otherwise specified (NOS)
- Bipolar disorder I
  - single manic episode[3,4,6]
  - most recent episode depressed[1–9]
  - most recent episode hypomanic[7–9]
  - most recent episode manic[3,4,6–9]
  - most recent episode mixed[3,4,6–9]
  - most recent episode unspecified[7–9]
- Bipolar disorder II
  - hypomanic[7–9]
  - depressed[1–9]
- Cyclothymic disorder
  - May also include:
    - [1] with melancholic features
    - [2] with atypical features
    - [3] with catatonic features
    - [4] with postpartum onset
    - [5] chronic
    - [6] severity/psychotic/remission specifiers
  - course specifiers:
    - [7] with seasonal pattern
    - [8] with rapid cycling
    - [9] with/without interepisode recovery

latter specifying a history of previous MDD episode(s). As a disorder, MDD is defined by the following criteria and symptomatology: the presence of five of the following symptoms, one of which is either dysphoric or irritable mood, or loss of interest or pleasure in all or almost all activities. The other symptoms include: sleep problems as manifested by insomnia or hypersomnia; complaints or other evidence of diminished ability to think, make decisions, or difficulty concentrating; loss of energy or general fatigue; eating problems as manifested by decreased or increased appetite or significant weight loss or gain (in young children failure to make expected weight gains is symptomatic); psychomotor retardation or agitation; suicidal or morbid ideation, death wishes, or suicide attempts; and feelings of self-reproach, worthlessness, or excessive or inappropriate guilt (which may be delusional). Symptoms need to be present nearly every day for a period of at least two weeks. Exclusion criteria are also specified that preclude a number of other pathologies, such as organic mental disorder and other etiologies (i.e., a normal response to bereavement) as concomitant problems, for a diagnosis of major depression. The *DSM-IV* provides information to assist the clinician in a differential diagnosis of major depression, as opposed to several conditions that may include significant depressive symptoms.

As shown in Table 1, major depression may be specified as a single episode (without a history of manic or hypomanic episode) or recurrent, the latter describing a history of two or more major depressive episodes separated by at least two months of normal functioning. There is also a range of additional features that may characterize the nature of the depressive episode, including: with melancholic features, catatonic features, atypical features, postpartum onset (adolescents), and indications of the severity (mild, moderate, severe), with or without psychotic features, and in partial or full remission. A seasonal pattern, often referred to as seasonal affective disorder, may be indicated.

5.17.3.2.2 Dysthymic disorder

In the *DSM-IV* dysthymic disorder represents a relatively chronic depressive disorder that is typically less severe in symptom distress than major depression, but typically of greater duration. Diagnostic criteria for dysthymic disorder in children and adolescents includes a depressed or irritable mood for most of the day, and manifested most of the time over a period of at least one year, although there may be periods of up to two months during which symptoms are not present. In addition to symptoms of depressed or irritable mood, at least two of the following symptoms must be present when depressed: appetite loss or gain, insomnia or hypersomnia, fatigue or low energy level, low self-esteem, difficulty concentrating or problems making decisions, and feelings of hopelessness. Exclusion criteria include: showing no evidence of major depressive episode during the initial year of the disorder; never having a manic or hypomanic episode; and not being superimposed on a chronic psychosis, or due to or maintained by an organic factor. Diagnostic criteria for dysthymia in children and adolescents differ from those for adults for whom symptoms of this disorder must be present for two years for a diagnosis.

5.17.3.2.3 Bipolar disorder

Bipolar disorder is a severe and typically long-term psychiatric condition that often represents a diagnostic challenge for clinicians. A number of researchers have noted the difficulties in this diagnosis as well as the propensity for misdiagnosis (Carlson, 1994; Joyce, 1984; Weller, Weller, Tucker, & Fristad, 1986). The essential feature of bipolar disorder is the occurrence of one or more episodes of mania or hypomania, typically with a history of one or more major depressive episodes. The *DSM-IV* delineates two forms of bipolar disorder, bipolar disorder I and bipolar disorder II, that differ as a function of the level of mania, with bipolar II associated with a hypomanic episode. Table 1 provides a description of the types of episodes of mania and depression associated with bipolar disorders. Of note is the mixed episode in which criteria for a manic episode and major depressive episode are present nearly every day with the youngster rapidly cycling between disorders. As noted in Table 1, bipolar disorder may also conform to a seasonal pattern.

The manic episode, which may last from one week to several months, is characterized by a clear-cut period of mood elevation, irritability, or expansiveness. In addition, three (or four if the manic episode is characterized by irritability) of the following symptoms must be present, including: grandiosity or inflated self-esteem, reduced need for sleep, more or excessively talkative, flights of ideas or perception that thoughts are racing, very distractible, increased time spent on goal-directed activities or psychomotor agitation, and overinvolvement in pleasurable, potentially risk-taking activities. Lastly, the overall clinical picture is sufficiently severe to cause impairment in functioning or require hospitalization. A somewhat milder variant that does not include the latter criteria of significant impairment in functioning nor delusions, is specified as a hypomanic episode.

5.17.3.2.4 Cyclothymic disorder

Cyclothymic disorder is a chronic mood disorder, symptomatically similar to bipolar disorder but of less severity. Cyclothymic disorder is analogous to dysthymic disorder in that depressive symptoms are not sufficient for a diagnosis of major depression, and the symptoms must be present for one year with periods of up to two months in which symptoms are not present. During the course of the disorder, there are numerous hypomanic episodes and periods of depressed mood or loss of interest or pleasure.

5.17.3.3 Alternatives to *DSM*: Clinical Severity, Subsyndromal Depression, and Depressive Symptoms

There have been a number of suggestions that criteria for *DSM* depressive disorders are too stringent and results in an underestimate of persons with significant levels of depressive symptoms. Thus, a person may show a level of depressive symptoms that is not sufficient to meet criteria for a *DSM* disorder yet, by its symptom severity level, may be considered clinically relevant and a valid target for intervention. Bland (1994) noted the problem of threshold in the use of *DSM*-specific criteria for diagnosis, in that a certain number of symptoms are required for a diagnosis and the patient who falls one symptom short of this does not receive a diagnosis, even when presenting symptoms are severe.

Gotlib, Lewinsohn, and Seeley (1995) found that youngsters who had clinically significant levels of depressive symptoms but did not meet criteria for major depression were twice as likely to develop a psychiatric disorder as compared to a nondepressed group, and also demonstrated significantly poorer psychosocial functioning than the nonsymptomatic group. Youngsters with clinically significant depressive symptoms were similar in psychosocial functioning to a comparison group of youngsters with major depression. These results suggest the clinical importance of attending to youngsters who may not meet criteria for major depression yet show significant levels of depressive symptomatology.

Judd, Rapoport, Paulus, and Brown (1994) examined depressive symptom data from the National Institute of Mental Health epidemiological catchment area (ECA) study and developed a clinical condition that they described as "subsyndromal symptomatic depression." Similarly, Horwath, Johnson, Klerman, and Weissman (1994) focused on ECA incidence data specific to persons who developed an episode of major depression over a one year interval between interviews. They found that persons with prior symptoms of depression or a dysthymic disorder were at relative risk for depression, with 50% of individuals who developed a major depressive episode demonstrating a history of depressive symptoms. The authors highlight the clinical importance of depressive symptoms by noting that "if depressive symptoms could be identified and treated before major depression first developed, many first onset cases of major depression could potentially be prevented" (p. 333), and that "measures designed to identify this group prior to the onset of major depression may represent an important step in primary prevention" (p. 333). Although the data from Judd et al. (1994) and Horwath et al. (1994) is specific to adults, it can be generalized to children and adolescents. Toward this end, a multiple-gate assessment procedure has been presented by Reynolds (1986a) for the identification of children and adolescents with significant clinical levels of depressive symptomatology.

In the diagnostic classification of mood disorders in the WHO *ICD-10*, a depressive episode may be classified as a severe depressive episode or a mild depressive episode, the latter different from dysthymia. However, in the worldwide field trials of the *ICD-10*, relative low interrater agreement was found for the mild depressive episode (code F31.1) as well as dysthymia (F34.1) (Sartorius et al. 1993). Sartorius et al. suggest that clinicians had difficulty distinguishing between dysthymia and mild depressive episode and recurrent mild depressive disorder (F33.1), as well as generalized anxiety disorder (F41.1), mixed anxiety and depressive disorder (F41.2) and adjustment disorder (F43.2). It is also useful to note that although interrater reliability for these disorders was low for the worldwide field trials, interrater agreement was even lower for field trials in the USA and Canada (Regier, Kaelber, Roper, Rae, & Sartorius, 1994). Although these data are based on adult samples, it does demonstrate that our diagnostic formulations of depressive disorders are still in a state of evolution.

5.17.4 ASSESSMENT OF DEPRESSION IN CHILDREN AND ADOLESCENTS

Assessment is a procedure basic to the identification of depression in children and adolescents for clinical purposes, as well as for use in the study of biological, social, and psychological characteristics; sequaele; outcome; determination of treatment efficacy; and other research applications. There are two primary categories of assessment measures: those designed for obtaining a diagnosis, and measures developed to evaluate the severity of depressive symptomatology. Although diagnostic measures may provide severity scores, their primary utilization is as diagnostic tools. In general clinical practice and in a multitude of research applications, the assessment of depression in children and adolescents is typically accomplished using self-report severity measures or clinical interviews, the latter including diagnostic measures and severity measures. Reports by significant others such as parents, teachers, and peers have been used to a limited extent and are not discussed in this chapter. Readers interested in assessment of depression via reports of others, as well as the issue of informant variability, are directed to Clarizio (1994) and Kazdin (1994).

There are several characteristics of depression that support the use of self-report assessment procedures, given linguistic and metacognitive competence in the child. Depression as an internalizing disorder includes primary symptoms that are internal to the youngster and are not easily observable. Cognitive symptoms of guilt, self-deprecation, suicidal ideation, hopelessness, and feelings of worthlessness are depressive symptoms that are subjective to the child. Some vegetative symptoms such as insomnia, appetite loss, and other problems are sometimes difficult for others to observe and may go undetected by parents and significant others.

This chapter presents information on several self-report measures of depression for children and adolescents. Although a number of structured and semistructured clinical interviews exist, the application of these interviews is generally limited to various research investigations. The primary tools used in the clinical assessment of depression in children and adolescents in clinical and school settings are self-report measures. Diagnostic interviews are briefly noted below and several semistructured clinical interviews that provide for the assessment of the severity of depressive symptomatology in children and adolescents are described.

The manner in which we assess depression in youngsters guides how we organize our perspective of depressive phenomena. For the determination of depression according to formal classification criteria that includes both inclusion and exclusion components, the standard practice in research investigations has been the use of diagnostic clinical interviews. Structured clinical interviews are typically formulated to evaluate all symptoms, their duration, and potential exclusion criteria as specified by a formal set of diagnostic criteria. In this manner, the K-SADS (Puig-Antich & Chambers, 1978) was designed to provide diagnostic information on disorders in youngsters according to the RDC (Spitzer et al., 1978). Likewise, *DSM* criteria for psychiatric disorders are assessed by other interviews such as the Child Assessment Schedule (Hodges, Kline, Stern, Cytryn, & McKnew, 1982), Diagnostic Interview for Children and Adolescents-Revised (Reich & Welner, 1988), and the Diagnostic Interview Schedule for Children-Revised (Shaffer et al., 1993); see Hodges' (1994) and Reynolds' (1992b) reviews of these and other clinical interview measures.

Self-report severity measures typically evaluate a range of depressive symptoms with the assessment format specific to the depth of symptom expression (e.g., frequency of occurrence, severity, etc.). Most of these measures generally provide a total score that may be compared to a derived cutoff score. In this manner, many severity measures of depression allow for the evaluation of a clinical level of depressive symptomatology, assuming an adequate coverage of the symptom domain (Reynolds, 1994). Using the nomenclature of Compas, Ey, and Grant (1993), single-scale self-report depression measures assess depressed mood in youngsters, and self-report measures that are components of a comprehensive assessment measure or that were developed to assess all symptoms of classification system specified depressive disorders allow for the assessment of depressive syndromes. Thus, self-report scales are not designed as diagnostic measures.

Limitations on the use of self-report are primarily developmental. Linguistic competence, including reading and comprehension ability, are important considerations when using self-report measures. These task demands suggest that most youngsters aged eight years and above who do not have learning problems can respond to self-report measures that include language at a third-grade reading level. To avoid confounding with reading, it is recommended that measures be read to youngsters below 10 years of age. Likewise, it is important to consider metacognitive demands of the measure. Thus, task demands go beyond reading to include the ability to recognize one's emotional and behavioral characteristics as they apply to a specified time-frame and evaluate the frequency, severity, or nature of these characteristics.

The ability of the youngster to self-monitor their behavior and emotions as well as self-evaluate these facets is a requirement of most measures. There are a number of additional developmental aspects specific to assessment of depression in children that are of primary concern to the assessment of young children rather than adolescents and address both the reliability and validity of measures of depression in youngsters. Many of these issues are generic to the use of self-report methods with children and are described in more detail elsewhere (Reynolds, 1992b).

As a function of the limitations noted above, there are few self-report measures that can be used with children below the age of eight years, although there is some research to suggest that self-report measures of depression can be used with children as young as ages 5 or 6 years. (Ialongo, Edelsohn, Werthamer-Larsson, Crockett, & Kellam, 1993). For children aged six and above, clinical interviews, such as the Children's Depression Rating Scale-Revised (CDRS-R; Poznanski et al. 1984) may be used, although some youngster's metacognitive ability for emotions may limit their response. Thus, for young children, parent reports using clinical interviews are a primary source of information, with parent's reports also of value for school-age children, assuming reliable and valid reporting on the part of the parent.

5.17.4.1 Self-report Measures of Depression for Children and Adolescents

Since the mid-1970s, a number of self-report measures have been developed for the assessment of depression in children and adolescents, with some researchers and clinicians using measures designed for use with adults. The use of self-report questionnaires for the assessment of depressive symptomatology in young people has witnessed a rapid growth in research and clinical applications. From a few reports in the early 1980s, to the routine use of self-report depression measures in a wide range of child and adolescent mental health disciplines, self-report measures provide a simple and direct method for the evaluation of the clinical severity of depressive symptoms in children and adolescents. Table 2 lists the more common self-report measures used in North America. These measures are among the most frequently used self-report scales for the assessment of depressive symptomatology in children and adolescents (Archer, Maruish, Imhof, & Piotrowski, 1991). A brief description of several of these measures is provided below. More information on these and other self-report measures of depression in children and adolescent is given by Reynolds (1994).

5.17.4.1.1 Children's Depression Inventory

The Children's Depression Inventory (CDI; Kovacs, 1979, 1992) as described by Kovacs and Beck 1977) was originally developed as a downward revision of the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). The CDI is one of the most frequently used research measures of depression in children and consists of 27 items, each evaluating a symptom of depression or related affect. Items are presented as three statements of varying symptom severity. Several of the items may be confusing to some children (e.g., I do not do what I am told most of the time).

The CDI manual (Kovacs, 1992) presents normative data based on a sample of 1266 children and young adolescents from an unspecified number of schools in Florida. Data appear to be a subsample of youngsters from a previously published study by Finch, Saylor, and Edwards (1985). Data on the race or ethnicity of the standardization sample is not

Table 2 Description of self-report measures of depression used with children and adolescents in North America.

Measure/authors	*Development target population*	*Number of items*	*Response format*	*Manual available*
Children's Depression Inventory Kovacs (1979, 1992)	children/adolescents	27	3-alternative	Yes
Children's Depression Scale Lang and Tisher (1987)	children/adolescents	66	card sort	Yes
Reynolds Child Depression Scale Reynolds (1989)	children	30	Likert 4-point	Yes
Reynolds Adolescent Depression Scale Reynolds (1986b)	adolescents	30	Likert 4-point	Yes
Beck Depression Inventory Beck, Ward, Mendelsohn, Mock, and Erbaugh (1961)	adults	21	4-alternative	Yes[a]

[a] The manual for the Beck depression inventory is specific to the use and characteristics of the BDI with adults.

reported, and data is provided for children in grades 2 through 8. Norms are not provided for high school students. Although not indicated in the manual, the normative sample data was based on an oral administration of the CDI as indicated in the original article by Finch et al. (1985). The CDI manual has been reviewed by Reynolds (1994).

The psychometric characteristics of the CDI have been reported in many studies. Researchers typically report internal consistency reliability coefficients in the low to upper 0.80s (Cole & Carpentieri, 1990; Kovacs, 1983; Reynolds, Anderson, & Bartell, 1985; Smucker, Craighead, Craighead, and Green, 1986) and test–retest reliability coefficients ranging from 0.38 to 0.87 (Finch, Saylor, Edwards, & McIntosh, 1987; Kovacs, 1980/1981, 1983; Nelson & Politano, 1990; Saylor, Finch, Spirito, & Bennett, 1984; Smucker et. al., 1986; Weiss et al., 1991). The variability in test–retest reliability is in part a function of the time interval between assessments and the sample characteristics.

One line of research has examined the utility of the CDI in distinguishing between depressed and nondepressed youngsters (e.g., Curry & Craighead, 1990; Fristad, Weller, Weller, Teare, & Preskorn, 1988; Moretti, Fine, Haley, & Marriage, 1985; Saylor et al., 1984). The results of these studies are mixed and tend to show a lack of diagnostic efficacy.

Although not designed as a diagnostic measure the CDI is useful in assessing the severity of depressive symptomatology. One question that has emerged is whether a particular cutoff score can represent a clinically relevant level of depressive symptoms on the CDI. A cutoff score of between 17 and 19 appears to be appropriate, given the score distributions for the CDI found with various school-based and clinical samples of children. In samples of school-based or nondepressed youngsters a mean CDI score of 8–10, and a standard deviation of 7–8 points is typically reported (Cole, 1991; Cole & Carpentieri, 1990; Craighead, Curry, & Ilardi, 1995; Finch et al., 1985; Ollendick, Yule, & Ollier, 1991; Smucker et al., 1986; Saylor et al., 1984). In large samples of nonreferred children a score of 19 identifies approximately 10% of the sample.

5.17.4.1.2 Reynolds Child Depression Scale

The Reynolds Child Depression Scale (RCDS; Reynolds, 1989) is a 30-item self-report measure designed for use with children aged 8–13. On the RCDS 29 items are rated as to the frequency of their occurrence during the past two weeks using a four-point "almost never" to "all the time" response format. Items are worded to evaluate current symptom status (e.g., item 8: I feel like crying; item 14: I feel like hurting myself). Seven items are reverse keyed (e.g., item 25: I feel like having fun) with a negative response indicative of depressive symptom status. Item 30 consists of five smiley-type faces ranging from sad to happy, that the child completes by placing an X over the face that indicates how he or she feels and represents an evaluation of dysphoric mood.

The RCDS has been used in a number of descriptive and intervention studies of depression in children (e.g., Bartell & Reynolds, 1986; Huebner, Scott, & Alderman, 1993; Rawson, 1992; Reynolds et al., 1985; Reynolds & Graves, 1989; Roseby & Deutsch, 1985; Smoller & Brosgole, 1993). The manual for the RCDS provides normative information on over 1600 children representing heterogeneous ethnic and socioeconomic backgrounds, along with procedures for administration, scoring and interpretation, and data on reliability and validity. The internal consistency reliability of the RCDS with the standardization sample was 0.90. Internal consistency reliability of the RCDS with samples of normal and depressed youngsters have ranged from 0.87 to 0.92 (Bartell & Reynolds, 1986; Crowley & Emerson, 1996; Reynolds, 1989; Reynolds et al., 1985; Reynolds & Graves, 1989). Reynolds and Graves (1989) report a test–retest reliability coefficient of 0.85 for the RCDS using a four-week interval between testing with an ethnically diverse sample of 220 children from grades 3 through 6, and a mean difference in RCDS scores between the two assessments of less than two points.

Validity data for the RCDS have been reported in the form of high correlations with other self-report and clinical interview measures of childhood depression, as well as content validity, factor analysis, discriminant validity, and clinical utility (Reynolds, 1989). Researchers have reported correlations between the RCDS and the CDI ranging from 0.70 to 0.79 across several studies (e.g., Bartell & Reynolds, 1986; Crowley & Emerson, 1996; Reynolds, 1989; Reynolds et al., 1985). Reynolds (1989) reports a correlation between the RCDS and CDRS-R of 0.76. Using the CDRS-R cutoff score to provide a base rate for a clinical level of depression, Reynolds (1989) reported a specificity rate of 97%, sensitivity of 73%, a Yule's Y of 0.81 ($P < 0.0001$) and a phi coefficient of 0.73 ($P < 0.0001$) for the RCDS cutoff score with a sample of 82 children. A number of studies (e.g., Rawson & Tabb, 1993; Stark, Reynolds, & Kaslow, 1987) have shown the RCDS to be sensitive to treatment outcome as a measure of depression.

5.17.4.1.3 Children's Depression Scale

The Children's Depression Scale (CDS; Tisher & Lang, 1983) was developed in Australia and uses a modified card-sort response format with its 66 items individually listed on cards and the child sorting the cards into one of five boxes labeled: very wrong, wrong, don't know/not sure, right, and very right. The CDS evaluates two broad factors, with 18 items comprising a "positive affective experience" (P score) subscale, and the remaining 48 items a depression subscale (D score). The positive affect scale includes one subscale of eight items (pleasure and enjoyment), with the remaining items on the P scale described as miscellaneous (Lang & Tisher, 1987). The depression subscale includes five subscales, (affective response, social problems, self-esteem, sickness and death, guilt, and pleasure and enjoyment) along with nine additional miscellaneous items. A manual by Lang and Tisher (1987) for use in North America was published with norms based on 37 normal children from Australia. There is also an adult-respondent form of the CDS (Tisher, 1995).

Several studies have described the psychometric characteristics of the CDS. Tisher, Lang-Takac, and Lang (1992) report on a test–retest study by Tonkin and Hudson (1981) who found a one week test–retest reliability coefficient of 0.74 for the CDS with a sample of 60 children. Internal consistency reliability has been found to be high, ranging from 0.92 to 0.94 (e.g., Bath & Middleton, 1985; Kazdin, 1987). Kazdin (1987) with a clinical sample of 185 children, reported correlations of 0.48 and 0.51 between child report on the CDS D scale and scores on the CDI and a modified form of the BID interview, respectively.

The CDS appears to be a reliable child self-report measure, although the number of items and format make administration and scoring a bit cumbersome. The lack of sufficient normative data is a limitation for the clinical use of this measure.

5.17.4.1.4 Reynolds Adolescent Depression Scale

The Reynolds Adolescent Depression Scale (RADS; Reynolds, 1986b) is designed to evaluate the severity of depressive symptoms in adolescents aged 12–18. The RADS consists of 30 items written to reflect symptoms delineated by the *DSM-III* (APA, 1980) for major depression and dysthymic disorder, as well as additional symptoms delineated by the RDC. Items on the RADS require a third-grade reading level. Although written at the time *DSM-III* was in effect, there have been few changes in symptoms and criteria for depressive disorders between *DSM-III*, *DSM-III-R* (APA, 1987), and *DSM-IV* (APA, 1994). Items on the RADS evaluate a range of symptom domains, including somatic, motivational, cognitive, mood, and vegetative components of depression.

The RADS was normed on an ethnically diverse sample of 2460 adolescents ages 12–19 years. The RADS norms appear to be relatively robust, with similar mean scores reported in a number of other studies of school-based youngsters in the USA and Canada (Dalley, Bolocofsky, Alcorn, & Baker, 1992; Reynolds & Miller, 1989; Schonert-Reichl, 1994; Weist, Proescher, Freedman, Paskewitz, & Flaherty, 1995). The RADS has been used as a measure of depressive symptomatology in numerous studies with school-based and clinical samples of adolescents (Adams & Adams, 1993; Adams, Overholser, & Spirito, 1994; Brand, King, Olson, Ghaziuddin, & Naylor, 1996; King et al., 1996; King, Hovey et al., 1997; King, Radpour, Naylor, Segal, & Jouriles, 1995; King, Raskin, Gdowski, Butkus, & Opipari, 1990; King, Segal et al., 1995; King, Segal, Naylor, & Evans, 1993; Licitra-Kleckler & Waas, 1993; Nieminen & Matson, 1989; Sadowski & Kelley, 1993; Wurzbacher, Evans, & Moore, 1991).

The reliability of the RADS is generally reported as high, ranging from 0.91 to 0.96 with samples of normal and depressed adolescents ranging in size from 62 to 2120 (Dalley et al., 1992; Reynolds, 1987, 1990; Reynolds & Coats, 1982; Reynolds & Miller, 1989; Schonert-Reichl, 1994). Test–retest reliability based on a sample of 104 adolescents who were tested with a six week interval between testings was 0.80. A 12 week test–retest reliability coefficient of 0.79 was also found for the RADS with a sample of 415 high school students (Reynolds, 1987). More recently, Reynolds and Mazza (in press) reported a test–retest reliability coefficient of 0.87 in a sample of 89 young adolescents from an inner-city school with an average retest interval of three weeks.

The validity of the RADS has been shown in relationships with other depression scales and measures of related constructs. The RADS has demonstrated strong correlations with other self-report measures of depression, including the BDI (Beck et al., 1961), and Center for Epidemiological Studies depression scale (Radloff, 1977) (r = 0.71–0.89). Kahn, Kehle, and Jenson (1987) in a study of 349 adolescents found a correlation of 0.75 between the RADS and CDI. Shain, Naylor, and Alessi (1990) in a sample of 45 adolescent inpatients with major depression, reported a correlation coefficient of 0.87 between the RADS and CDI.

Criterion-related validity of the RADS has been shown using semistructured clinical interviews of depression. In a study of 111 adolescents, Reynolds (1987) reported a correlation of 0.83 between the RADS and the Hamilton Depression Rating Scale (HDRS; Hamilton, 1967). In their study of adolescents with major depression, Shain et al. (1990) reported a correlation of 0.73 between the RADS and the HDRS. Reynolds and Mazza (in press) found a correlation of 0.76 between the RADS and HDRS. Overall, in studies of clinical samples of adolescents, investigators have reported strong evidence for criterion-related validity for the RADS with diagnostic and semistructured clinical interview measures of depression (King, Katz et al., 1997; Reinecke & Schultz, 1995).

The RADS uses a cutoff score to describe a clinical level of depressive symptom severity. The RADS cutoff score has been validated in several studies, using both HDRS scores as well as formal diagnosis based on the SADS as criterion measures (Reynolds, 1987). Using the RADS cutoff score of 77 and a HDRS cutoff of 15 with a subsample of 32 adolescents from a psychiatric center, Carey, Kelley, and Carey (1991) reported that 78.1% were correctly classified using the cutoff scores on these two measures.

In a study of adolescents with major depression and an age and sex matched normal control group, Shain et al. (1991) found significant differences in RADS scores between youngsters with major depression (M = 85.5) and normal controls (M = 45.5) t = 9.02, $p < 0.0001$. Reynolds and Mazza (in press) reported a hit rate of 94% and a kappa coefficient of 0.71 using the RADS cutoff score in comparison to a clinical cutoff score of 16 on the HDRS. The RADS has demonstrated clinical efficacy in its use in a number of adolescent depression treatment studies (Hains, 1992; Kahn, Kehle, Jenson, & Clark, 1990; Reynolds & Coats, 1986).

5.17.4.1.5 Beck Depression Inventory

The BDI (Beck et al., 1961) was developed for adults and consists of 21 items, each consisting of four statements of varying degrees of symptom severity. A number of investigators have used the BDI with adolescents, although the reading requirements and format of the scale present difficulty for some youngsters. Users of the BDI often apply cutoff scores used with adults to adolescents without adequate validation.

Internal consistency reliability of the BDI with adolescents has been reported in the 0.70s to 0.80s (Kashani, Sherman, Parker, & Reid, 1990; Reynolds & Coats, 1982; Roberts, Lewinsohn, & Seeley, 1991; Strober, Green, & Carlson, 1981). Strober et al. (1981) reported a five-day test–retest reliability of 0.69 in a sample of 78 psychiatric inpatient adolescents, similar to the short-term test–retest reliability of 0.67 reported by Roberts et al. (1991) with a large sample of school-based youngsters. In a study of 37 older adolescents, Kutcher & Marton (1989) found a one-week test–retest reliability coefficient of 0.83. Overall, the reliability of the BDI with adolescents is lower than that found in other self-report measures of depression developed specifically for youngsters.

Validity evidence for the BDI has been mixed in clinical and school-based samples. Barrera and Garrison-Jones (1988) with a sample of hospitalized psychiatric patients (*n* = 65) and school-based adolescents (*n* = 49) reported reasonable sensitivity and specificity in the school sample but poor results for the psychiatric sample. Kashani et al. (1990) noted a relatively low sensitivity rate (48%) using the BDI cutoff score of 16 in a sample of 100 clinic referred adolescents. Roberts et al. (1991) in a study of 1704 adolescents who completed the BDI as well as a diagnostic interview, found a correlation of 0.50 between the BDI and a 14-item form of the HDRS, which is relatively low, with low kappa coefficients for the BDI in identifying cases of major depression ($\kappa = 0.14$) and dysthymia ($\kappa = 0.03$). The somewhat lower reliability, adult language and comprehension demands, and issues of validity of the BDI cutoff score with adolescents are issues that should be recognized in the use of the BDI with adolescents.

5.17.4.2 Clinical Interview Severity Measures of Depression

A number of diagnostic and severity interview measures have been developed for the evaluation of depression in children and adolescents. As noted earlier, diagnostic measures that assess the wide range of psychopathology found in young people will not be discussed in this chapter. Described below are two semistructured clinical interviews that evaluate the severity of depressive symptoms, in youngsters. These measures, the CDRS-R (Poznanski et al., 1984) and the HDRS (Hamilton, 1967) have been used in research and treatment outcome applications.

5.17.4.2.1 Children's Depression Rating Scale-Revised

The CDRS-R (Poznanski et al. 1984) was originally developed in 1979 (Poznanski et al., 1979) and is a semistructured clinical interview measure of the severity of symptoms of depression designed for use with children aged 6–12. The CDRS-R consists of 17 items, each representing a depressive symptom, and rated on a 1 (normal) to 7 (severe) scale. On the CDRS-R, 14 of the items are presented in an interview format, and three items (depressed affect, tempo of language, and hypoactivity) are observational. The CDRS-R requires 20–30 minutes to administer.

Poznanski and her colleagues (Poznanski et al., 1979; Poznanski, Cook, Carroll, & Corzo, 1983; Poznanski, Carroll, Banegas, Cook, & Grossman, 1982) have reported acceptable levels of reliability and evidence of validity of the CDRS in differentiating RDC-diagnosed youngsters. Poznanski, Freeman, & Mokros (1985) recommends a score of 40 and above on the CDRS-R as a clinical cutoff for depression in children. Using a cutoff score of 44, Jensen, Burke, and Garfinkel (1988) reported sensitivity of 50% and specificity of 79% for a diagnosis of major depression in a sample of 35 youngsters aged 8–18, 12 of whom had a diagnosis of major depression. Heiligenstein and Jabobsen (1988) modified the CDRS-R with the removal of somatic items for the evaluation of depression in medically ill youngsters. The CDRS and CDRS-R have been used as severity measures in research on depression in children and adolescents (e.g., Mokros, Poznanski, Grossman, & Freeman, 1987; Shain et al., 1990; Weller, Weller, Fristad, Cantwell, & Preskorn, 1986), and the CDRS-R was used as an outcome measure in treatment studies of depression in children by Liddle and Spence (1990) and Stark et al. (1987).

5.17.4.2.2 Hamilton Depression Rating Scale

The HDRS (Hamilton, 1960, 1967) is a clinical interview measure of the severity of depressive symptoms and one of the most frequently used outcome measures in psychiatric studies of depression in adults. There are several versions of the HDRS, the most popular being the 17 item form. Items are evaluated by the clinician as to their severity with higher scores indicative of greater depressive pathology. Two items that assess psychomotor retardation and psychomotor agitation, are observational. Items on the HDRS are scored from 0 to 4 or 0 to 2 depending on the symptom assessed, with high scores indicating greater symptom pathology. Researchers have utilized a number of cutoff scores to delineate a clinical level of depressive symptomatology, with such scores typically ranging from 15 to 20.

There are a number of caveats to the use of the HDRS. It should be noted that a number of contemporary symptoms of depression are not evident on the HDRS (e.g., low self-esteem, self-deprecation, etc.). Likewise, several items dealing with loss of insight, hypocondriasis, and several anxiety items are tangentially related to depression. Nevertheless, as a severity measure of depression, the HDRS, with minor adjustments for age differences, works quite well with adolescents.

Research using the HDRS has included both clinical and nonclinical samples of adolescents

(Alessi, McManus, Grapentine, & Brickman, 1984; King, Naylor, Hill, Shain, & Greden, 1993; King, Segal et al., 1995; King et al., 1996; Reynolds, 1987; Shain et al., 1990). Reynolds (1987) found an internal consistency reliability of 0.91 for the HDRS with a sample of 111 school-based adolescents, and Reynolds and Mazza (in press) reported a reliability of 0.85 with young adolescents. Robbins, Alessi, Colfer, and Yanchyshyn (1985) obtained an interrater reliability of 0.90 with a sample of adolescent psychiatric inpatients. King, Naylor, Segal, Evans, & Shain (1993) reported an interrater reliability of 0.97 in a sample of 30 depressed and nondepressed adolescents.

5.17.4.3 Assessment Summary

A number of interviews and rating scales developed for the assessment of depression in children and adolescents have been described. A positive outcome since the mid-1980s in the study of depression in children and adolescents has been the development of diverse methodologies and procedures for the assessment of this disorder in youngsters. Although the role of child self-report in the evaluation process has been stressed, the availability of other rich sources of information can only serve to increase our perspectives regarding the mental health of children and adolescents. An examination of measures for the evaluation of depression in youngsters indicates that there is no one right or correct measure. Rather, there are different measures designed for different assessment purposes. Furthermore, the variance in sources of information about a youngster's depression, although disconcerting to researchers and scientists who often seek parsimony among measures of the same construct, may suggest that there is also no one right or correct source of information. The issue of differences in reports among sources of information, or informant variability is a complex issue that has developmental as well as psychometric underpinnings. Additional information on this issue as it relates to the assessment of depression in children and adolescents is given by Kazdin (1994).

It is recommended that the evaluation of depression in a child, particularly a young child, involves multiple assessment measures and, where possible, multiple assessment methods (e.g., self-report, clinical interview, reports by significant others). Waller and Rush (1983) have provided an interesting approach to assessment and diagnosis of depression in children which incorporates formal psychological assessment, medical procedures, and response to treatment for distinguishing primary depression, organic affective syndrome, and situational depression in children. As found in numerous research studies, disagreements occur across methods and measures. In such cases, a close examination of the specific assessment information is warranted, with the clinician deciding how best to combine these data. In the above discussion of depression assessment measures, there is an evident gap in procedures for the evaluation of depression in children below the age of 6 or 7 years.

5.17.5 PSYCHOLOGICAL TREATMENT OF DEPRESSION IN CHILDREN AND ADOLESCENTS

It is evident that depression is a serious psychological disturbance that is experienced by a very large number of children and adolescents. Surveys and studies of service utilization suggest that most of these youngsters do not receive treatment of any type (Keller et al., 1991). The high prevalence rate combined with the potential range of long-term negative consequences (e.g., Kovacs, 1996b) supports the need and requirement for effective treatment procedures. Treatment approaches for depression in children and adolescents are, for the most part, modifications or direct applications of procedures developed for use with adults. Depressed youngsters represent a particularly vulnerable population. Unlike children and adolescents with externalizing problems, such as conduct disorder, for whom treatment may be viewed by the youngster as undesirable or even punitive, depressed youngsters often see treatment as a potential avenue to reduce or alleviate their severe distress.

The empirical evidence for the efficacy of contemporary psychological treatments for depression in children and adolescents is limited to fewer than a dozen studies that have been conducted since 1980. Thus, for practical purposes, our evidence for the use of contemporary psychotherapies for the treatment of depression in youngsters is predicated on the results of these investigations. The use of psychoanalytically based treatments has a somewhat longer history, although there is limited empirical evidence to support therapeutic efficacy.

5.17.5.1 Range of Therapeutic Approaches

There are a multitude of treatment modalities across the two primary domains of psychological and somatic approaches. The latter include the use of antidepressant medications, such as tricyclics, monoamine oxidase inhibitors, specific serotonin reuptake inhibitors, and lithium carbonate. These procedures have generated a body of empirical research that allows for the evaluation of their clinical efficacy although, as in the case with psychotherapy, the number of studies is small in comparison to research with adults.

As noted, there have been fewer than a dozen empirical demonstrations of the efficacy of psychological treatments designed for the amelioration of affective disorders and depressive symptoms in children and adolescents. The small number of experimental studies on the psychological treatment of depression in children and adolescents is perplexing, particularly given the large number of empirical outcome studies on the treatment of depression in adults. However, a similar situation exists with regard to the number of pharmacological interventions for depression in young people in comparison to the large number of studies of antidepressant efficacy with adults.

A brief overview of control-group studies of psychotherapy for the treatment of depression in children and adolescents is presented in this chapter. Case reports, including single-subject designs, have been reported and been reviewed previously (Reynolds, 1992b).

5.17.5.2 Control-group Investigations

For the demonstration of clinical efficacy and generalizability, control-group experimental treatment studies that provide data on the efficacy of psychological treatments for depression in children and adolescents are of significance. The majority of the published empirical, control group investigations have utilized a range of cognitive-behavioral procedures, and have typically compared youngsters receiving such treatment with groups of youngster who receive another potentially active treatment, therapeutic support, and/or are assigned to a wait-list condition and do not receive treatment during the period of study. A number of contemporary psychotherapy investigations for depression in children and adolescents are briefly described below and summarized in Tables 3 and 4. In the majority of controlled group investigations reported below, structured therapy manuals have been developed that provide evidence for the combination of specific treatment components.

In addition to studies reported in peer-reviewed publications, a number of investigators have provided partial descriptions of treatment outcome studies within books or chapters in books. Results of a cognitive-behavioral intervention for depression with children were described in Stark (1990) and in a chapter by Stark, Rouse, and Livingston (1991) who reported on treatment outcome with 24 children who were randomly assigned to a cognitive-behavioral (self-control, social skills, relaxation training, and cognitive restructuring) or traditional counseling intervention (therapeutic discussion and support). Family therapy, consistent with treatment components of each condition was also provided. Treatments consists of 24–26 sessions of 45–50 minutes over a $3^1/_2$ month period administered in small groups of four youngsters to a group. Both conditions showed significant decreases in depressive symptomatology in participants at posttreatment with treatment gains maintained at a seven month follow-up. On one outcome measure, the cognitive-behavioral group was superior to the counseling condition, although this difference was not found at follow-up. Actual test score means for the groups at various assessment points were not provided by the authors.

5.17.5.3 Psychotherapy Treatment Studies with Children

One of the first control group treatment studies for depression in children was reported by Butler, Miezitis, Friedman, and Cole (1980) who compared a role play condition that emphasized role play, social skills and social problem solving and a cognitive restructuring treatment based in part on rational-emotive therapy. Control groups consisted of an attention and waitlist condition. Participants were a school-based sample of 54 moderately depressed fifth and sixth grade children. The active treatments and attention condition consisted of 10 one-hour sessions conducted in small groups of six or seven children. A pretest–post-test design was used, although the authors did not conduct a follow-up evaluation to determine the maintenance of treatment gains. Multiple outcome measures were used, with the CDI the only measure of depression. At the post-test, significant decreases in CDI scores were found for the two active treatment conditions, as well as for the waitlist control group. The attention group remained relatively unchanged in their self-reported level of depressive symptomatology.

The results of the Butler et al. (1980) investigation were somewhat mixed. Although decreases were found for the two active treatments, the level of mean CDI scores for these groups at post-test (14 and 16 points) may be considered elevated. Likewise, the decrease in scores seen in the waitlist, but not the attention placebo control is puzzling. A significant methodological problem was the failure to include a follow-up assessment in the research design. In this sense, the extent to

Table 3 Summary of controlled group psychological interventions for depression in children and adolescents.

Study	Treatment groups	Sample size	Sessions		Depression measures	Outcome postest	Follow-up length	Follow-up outcome
			Number	Duration				
Butler et al. (1980)	cognitive-restructuring (CR)	14	10	60 min	CDI	[PR = CR = WL] > AC		
	role-play (RP)	14						
	attention control (AC)	13						
	waitlist control (WL)	13						
Reynolds and Coats (1986)	cognitive-behavioral (CB)	9	10	50 min	RADS	[CB = RT] > WL	5 weeks	CB = RT = WL
	relaxation training (RT)	11			BDI	[CB = RT] > WL		[CB = RT] > WL
	waitlist control (WL)	10			BID	[CB = RT] > WL		[CB = RT] > WL
Stark, Reynolds, and Kaslow (1987)	cognitive-behavioral (CB)	9	12	50 min	CDI	PS = CB, CB > WL	8 weeks[b]	CB = PS
	problem solving (PS)	10			RCDS	[CB = PS] > WL		CB = PS
	waitlist control (WL)	9			CDRS-R			CB > PS
Kahn et al. (1990)	cognitive-behavioral (CB)	17	12	50 min	RADS	[CB, RT, SM] > WL	4 weeks	[CB = RT = SM]
	relaxation training (RT)	17	12	50 min	CDI	[CB, RT, SM] > WL		
	self-modeling (SM)[a]	17	12	12 min	BID	RT = CB, SM CB, SM > WL		
	Waitlist control (WL)	17						
Lewinsohn et al., (1990)	cognitive-behavioral (CB)	21	14	120 min	BDI	[CB = CBP] > WL	4 weeks[c]	CB = CBP
	CB + parent (CBP)	19	14	120 min	CES-D	[CB = CBP] > WL		CB = CBP
	waitlist control (WL)	19			CBCL-D	CBP > CB, CBP = WL		CB = WL, CBP > CB
Liddle and Spence (1990)	cognitive-behavioral (CB)	11	8	60 min	CDI	CB = AC = WL	12 weeks	CB = AC = WL
	attention control (AC)	10	8	60 min	CDRS-R			
	waitlist control (WL)	10						
Marcotte and Baron (1990)	rational-emotive (RET)	12	12	60 min	BDI	RET = WL	8 weeks	RET > WL
	waitlist control (WL)	13						
Fine et al. (1991)	social skills (SS)	30	12	50 min	CDI	TS > SS[d]	9 months	TS = SS
	therapeutic support (TS)	36			K-SADS	TS > SS		TS = SS
Haines (1992)	cognitive-behavioral (CB)	9	9[e]	40 min	RADS	[CB = RT] > WL	11 weeks	CB = RT
	relaxation training (RT)	8						
	waitlist control (WL)	8						

[a] The self-modeling treatment was individually administered. [b] Follow-up assessment was for CB and PS groups. WL group received treatment following posttesting. [c] Follow-up assessment also completed at 6, 12, and 24 months. [d] $P < 0.10$. Sessions included three group and six individual sessions.

Table 4 Summary of noncontrolled group psychological interventions for depression in children and adolescents.

Study	*Treatment groups*	*Sample size*	*Sessions*		*Depression measures*	*Outcome postest*	*Follow-up length*	*Follow-up outcome*
			Number	*Duration*				
Mufson et al. (1994)	interpersonal (IPT-A)	11	12	45 min	BDI	$t = 7.03^{***}$		
					HDRS	$t = 9.16^{***}$		
Belsher et al. (1995)	cognitive therapy (CT)[a]	15	12	variable	BDI	$t = 6.51^{***}$	10 weeks[b]	gain maintenance
					IDS-C	$t = 5.96^{***}$		gain maintenance

> [a] Approximately 40% of participants also received medication. [b] Second follow-up conducted at approximately 20 weeks.
$^{***} P < 0.001$.

which the groups maintained treatment gains and evaluation of treatment generalization is unknown. This study is noteworthy in that it appears to be the first published experimental group intervention for depression in children, and it used an attention control condition.

Stark et al. (1987) compared a cognitive-behavioral treatment that was largely predicated on self-control and attribution retraining components, and a behavioral problem-solving treatment that utilized components of interpersonal problem-solving, increased engagement in pleasant activities, and self-monitoring to a waiting list control condition. Participants at post-testing were 28 moderately depressed children, aged 9–12. Assessment measures included the RCDS, CDI, and CDRS-R along with measures of related domains. The intervention was school-based and consisted of 12 50-minute sessions administered in small groups over a five-week period.

At post-testing, both active treatments demonstrated significant improvement in depressive symptomatology while the difference for the waiting list control group was nonsignificant. An eight-week follow-up showed a general maintenance of the treatment effects, with some children in the self-control condition showing continued improvement from post-test to follow-up. Children in the waiting list condition were treated after the post-test and were not included in the between-group comparisons at follow-up. Results indicated that the treatments were reasonably effective, with 88% of the children in the self-control group and 67% of those in the behavioral group reporting nondepressed levels of depressive symptomatology at the follow-up assessment. On the basis of this study and further development of active components, Stark (1990) has provided a more comprehensive treatment program for depression in children that is expanded in its scope, content, and duration.

In a study conducted in Australia, Liddle and Spence (1990) compared social competence training (SCT) intervention that included cognitive and behavioral treatment components, to an attention placebo control group, and a waiting list control group with a sample of 31 children aged 7–11. The authors indicated that children were identified based on a multiple-source multiple-stage screening procedure recommended by Reynolds et al. (1985). Children were selected from the population of regular class youngsters attending two Catholic schools in Sydney, Australia, and for whom parental permission for assessment was provided. Children were administered the CDI, which was read to younger-age children and those with reading problems. A cutoff score of 19 was used to select children for the second stage of the assessment procedure. Of the 380 children tested, 70 (18.4%) obtained a score of 19 or higher on the CDI. Two weeks following the initial screening, the CDRS-R was individually administered to the 70 youngsters. Using a cutoff score of 40 on the CDRS-R, 33 children were identified as demonstrating a clinical level of depression.

Of the 33 children identified, permission from parents for child participation in treatment was obtained for 31 youngsters. It is interesting to note that there were 10 girls and 21 boys, suggesting a gender difference of 2 to 1, with boys demonstrating a higher rate of depression, assuming that there was minimal difference in the proportion of males and females in the total sample. The mean age of the sample was 9.2 years (SD = 1.15). For the total treatment sample the mean CDI score was 21.32, with a mean of 42.03 on the CDRS-R, suggesting a moderate to severely depressed sample. Participants were randomly assigned to either a SCT condition ($n = 11$), an attention placebo control (APC) ($n = 10$), or a waiting list control group ($n = 10$).

Treatment for the SCT and APC consisted of eight one-hour sessions administered once a week to youngsters in groups of 4–6 children. The waiting list group did not receive any treatment and were evaluated at pretest, post-test, and at a three month follow-up evaluation. The SCT intervention focused on teaching social skills, interpersonal problem solving, and included components of cognitive restructuring. Self-monitoring of negative thoughts was also included. There appear to be similarities between treatment components of the SCT program and the behavioral problem-solving intervention developed by Stark et al. (1987). The attention/placebo control consisted of teaching drama to participants and appears to be more of an attention control condition rather than a placebo treatment.

At post-testing, all groups showed a modest decrease in CDI scores with no difference in treatment efficacy between groups. Unfortunately the CDI was the only depression measure used to evaluate treatment effects. Similar findings were reported for the follow-up assessment. Overall, the SCT intervention failed to produce meaningful treatment effects beyond those found in the attention and waitlist control conditions. This is an interesting study that appears to be the first to include younger children (i.e., age seven years and above). The lack of clinical efficacy is somewhat unexpected given previous studies and may be in part a function of the younger age of participants.

5.17.5.4 Psychotherapy Treatment Studies with Adolescents

In the first published control group treatment study for depression in adolescents, Reynolds and Coats (1986) compared a cognitive-behavioral treatment, a relaxation training group, and a waiting list control group with a sample of 30 moderately to severely depressed high school students. The two active treatment conditions were administered in 10 50-minute sessions twice a week for a five-week period. A follow-up assessment was conducted after four weeks from the end of treatment. Assessment measures included the RADS, BDI, and BID, along with measures of related constructs (e.g., anxiety, self-esteem). A school-based screening procedure was used to identify participants with the pretest representing a second assessment point. Both active therapies were presented in a structured format, with detailed treatment manuals established for each condition. The cognitive-behavioral therapy used procedures found in similar treatments of depression in adults, with modifications made for developmental and situational (e.g., home, school, peer-group, etc.) differences of adolescents. Components included teaching self-control procedures (i.e., self-monitoring, self-evaluation, and self-reinforcement), cognitive restructuring consistent with aspects of Beck's (1967, 1976) cognitive therapy, and procedures for increasing involvement in pleasant activities (Lewinsohn, Munoz, Youngren, & Zeiss, 1978). The relaxation training condition involved training in progressive relaxation procedures, along with guided imagery procedures. Both treatments were matched for therapist time and amount of homework assignments.

At post-testing, significant treatment gains for both active treatment conditions were found and maintained at the four week follow-up evaluation. Significant decreases in depression were reported by youngsters in the cognitive-behavioral and relaxation groups at post-test, with minimal change in the waiting list youngsters. Therapeutic effects of the interventions on participants' self-concept and level of anxiety were also reported by Reynolds and Coats (1986), with the relaxation training condition demonstrating the greatest reduction in self-reported anxiety.

Several of the treatment conditions of the Reynolds and Coats (1986) study were replicated by Kahn, Kehle, Jenson, and Clark (1990) who reported a treatment outcome study with 68 moderately depressed young adolescents. Treatment conditions for depression included cognitive-behavioral, relaxation training, and a self-modeling, along with a waitlist control group. The cognitive-behavioral treatment included pleasant activity scheduling and other behavioral components consistent with the approach of Lewinsohn et al. (1978), along with cognitive restructuring, self-control, and social skills training. The relaxation training group was similar to that of Reynolds and Coats (1986). Cognitive-behavioral and relaxation treatments were group administered in 12 50-minute sessions. The self-modeling treatment was individual with each youngster developing a three-minute videotape of themselves modeling nondepressive behaviors (e.g., smiling, positive verbalizations, and appropriate eye contact). Outcome measures included the RADS, CDI, and BID.

The authors reported significant treatment gains for the three active therapy conditions at post-testing and at a four week follow-up. The greatest therapeutic gains at post-testing and follow-up were found for the cognitive-behavioral and relaxation training conditions. The results of this investigation are consistent with Reynolds and Coats (1986), and suggest treatment generalizability as well as the efficacy of these procedures with young adolescents.

Lewinsohn, Clarke, Hops, and Andrews (1990) reported on the downward extension of the coping with depression course (Lewinsohn et al., 1978), a multimodal psychoeducational intervention for unipolar depression (Clarke & Lewinsohn, 1989). The sample consisted of 59 adolescents who met *DSM-III* criteria for major depression or RDC criteria for current minor or intermittent depressive disorder. Treatment conditions included: cognitive-behavior therapy group administered to adolescents only and consisting of 14 two hour sessions over a seven week period; the same cognitive-behavioral treatment administered to another group of adolescents along with group-training for parents (seven two-hour sessions) of these adolescents; and a waitlist control group.

The cognitive-behavioral intervention utilized a social learning approach to the treatment of depression, and included: pleasant activities scheduling; social skills training; cognitive restructuring; relaxation training; and a component focusing on communication, negotiation, and conflict resolution skills. The parent training included discussion of the components and skills taught to the adolescents and training in negotiation, communication and conflict-resolution skills.

Treatment outcome was reported for post-treatment, and follow-up assessments at 1, 6, 12 and 24 months. The waitlist condition was only available at post-testing. At post-testing, both active treatments demonstrated significant change as compared to the waitlist group, but were not significantly different from each other in their treatment efficacy. Treatment gains were maintained with a trend toward further improvement at follow-up assessments. The lack of the waitlist group as a comparison condition following post-treatment limited the long-term generalization of follow-up results.

In one of the few experimental studies conducted in a psychiatric setting, Fine, Forth, Gilbert, and Haley (1991) compared a social skills training group with a therapeutic support group for the treatment of depression in a group of 66 adolescent outpatients with a *DSM-III-R* diagnosis of major depression or dysthymia. The social skills condition consisted of self-monitoring of self and other's feelings, assertiveness training, conversation skills, social problem solving and conflict resolution. The therapeutic support treatment was designed as a therapeutic and supportive milieu for discussing problems and enhancing self-concept and self-worth. Outcome data (CDI and others measures) were reported for 47 youngsters who completed eight or more of the 12 treatment sessions. Within-group (pretest to post-test) differences indicated significant improvement for both treatments on a 12 item symptom scale derived from the K-SADS, with only the therapeutic support group showing significant post-test change on the CDI. Between-group analyses indicated the therapeutic support condition was superior on the K-SADS, with no difference between groups on the CDI. At a nine month follow-up there were no differences between groups on the outcome measures.

The results of this study are mixed. The social skills group showed minimal change in self-reported depressive symptoms on the CDI (pretest $M = 20.3$, post-test $M = 18.4$). On the K-SADS change from clinical to nonclinical levels was found in 50% of the therapeutic support group and 40% of the social skills group. This study suggests that social skills training alone may have limited therapeutic benefits. The superiority of the therapeutic support group suggests the potential utility of more interpersonal problem-solving therapies with adolescents.

In a study of 25 adolescent boys, 15–16 years of age, seeking training in stress management, Hains (1992) examined, in comparison to a waitlist control group, the efficacy of: a cognitive-behavioral intervention for stress inoculation based on the work of Meichenbaum (1985); an anxiety management procedure predicated on the work of Suinn (1986) with a strong relaxation training component. Although the focus of the study was on anxiety and stress management, depressive symptomatology was also examined at pretest, post-test, and follow-up. Interventions incorporated both group (three sessions) and individual (six sessions) therapy, with an 11 week follow-up conducted for the two active treatment conditions. Because of the selection criteria, group means on the RADS of participants suggested mild levels of depressive symptomatology, below means reported in other investigations (e.g., Kahn et al., 1990; Reynolds & Coats, 1986). However, given the small number of studies reported in the literature, this investigation is of interest, although it may be viewed as treating a subclinical level of depression. At post-testing, significant treatment gains on the RADS were found for the two active conditions in comparison for the waitlist group, with treatment groups maintaining treatment gains at follow-up.

In a study conducted in Canada with francophone youngsters, a cognitive therapy for depression was compared to a waiting list control group in a study of 25 depressed adolescents conducted by Marcotte and Baron (1993). The authors used a cognitive treatment package based on the rational-emotive model of depression. The treatment was administered in 12 60-minute sessions over a six week period. Participants were ages 14–17 years and were identified by the three-stage school-based screening procedure suggested by Reynolds (1986a). Significant treatment gains were found on the BDI in the cognitive group in comparison to the waitlist condition at post-test and at eight week follow-up.

As noted earlier, there have been several open trial studies of psychotherapies for the treatment of depression in adolescents. Although not described in detail in this chapter they are noted and present promising evidence for treatment efficacy. These studies are listed in Table 4. The study by Mufson et al. (1994) was based on a modification of the interpersonal psychotherapy of Klerman and Weissman (Klerman & Weissman, 1982). The modification of this intervention for adolescents was originally described by Moreau, Mufson, Weissman, and Klerman (1991) and more fully in Mufson, Moreau, Weissman, and Klerman (1993). Significant posttreatment gains in depressive symptomatology scores were found on the BDI and the HDRS.

A noncontrol group treatment study using cognitive therapy for depression in adolescents was reported by Belsher et al. (1995) with 15 adolescents who met criteria for *DMS-III-R* major depression. The authors reported significant reduction in depressive symptomatology at post-testing and two follow-ups (10 and 20 weeks) on the BDI and a clinician rating form of the Inventory of Depressive Symptomatology (Rush et al., 1986). It is also noteworthy that, similar to Mufson et al. (1993), the authors and colleagues have presented a more detailed description of their cognitive therapy for depression in adolescents in a book (Wilkes, Belsher, Rush, & Frank, 1994).

In both of the open trials noted above, significant reductions in depressive symptomatology were reported. However, as has been found in controlled group investigations, it is not uncommon to find symptom reduction also to occur in waitlist or control conditions, such that the primary focus is the demonstration of greater clinical efficacy of the active condition in comparison to control conditions. The same design issues apply to the evaluation of pharmacotherapies.

A significant finding from the majority of studies was the relative treatment efficacy of cognitive-behavioral therapies for the treatment of depression, with promising results reported for interpersonal psychotherapy. It is also noteworthy that treatment effectiveness was reported across therapies based on between 10 and 15 sessions, with treatment gains maintained at follow-up for most studies. This suggests that brief, structured contemporary psychotherapies are potentially efficacious for the treatment of depression in children and adolescents. Given what is known regarding the clinical efficacy of pharmacotherapies, contemporary psychotherapies should be considered as a general first line of treatment unless clinical data shows strong support for antidepressant treatment. For example, in a descriptive study of melancholic and nonmelancholic subtype adolescent psychiatric inpatients with major depression who were "significantly incapacitated by their symptoms," Robbins, Alessi, and Colfer (1989, p. 100) found that 68% of youngsters with nonmelancholic major depression responded to an intensive six week, broad-based psychosocial treatment without the use of medication. Adolescent inpatients with melancholic type major depression showed minimal response to psychotherapy, although 26% of these youngsters did show significant improvement with psychotherapy alone.

Given the lack of significant side-effects and the slow, but growing literature to suggest efficacy, contemporary psychotherapies for depression in children and adolescents need to be integrated into the standard of practice for the treatment of affective disorders in young people. It is also useful to note that in a study of treatment acceptability for interventions for depression in children, Tarnowski, Simonian, Bekeny, and Park (1992) found that mothers view psychological interventions to be significantly more acceptable than pharmacological, with cognitive-behavioral and behavioral interventions found somewhat more acceptable than social skills interventions. At this point, it appears that there is sufficient evidence to support the use of cognitive-behavioral procedures for the treatment of depression by clinicians. Based on the evidence of studies cited above, it does not appear that interventions with social skills enhancement or training are particularly useful as a stand-alone treatment. Further empirical study of interpersonal psychotherapy and other procedures should be undertaken to provide clinicians with a range of tested psychotherapeutic treatment approaches. Lastly, the efficacy of relaxation training has been noted in several studies, with its inclusion as a component noted in others, suggests that this approach may be viable for treating depression in some youngsters.

5.17.6 PHARMACOTHERAPY FOR DEPRESSION IN CHILDREN AND ADOLESCENTS

The use of medications for the treatment of mood disorders in children and adolescents has a long history compared to the application of psychotherapy. An examination of the pharmacological literature suggests that most empirical investigations of antidepressant medications for the treatment of depression (as well as a host of other disorders such as obsessive compulsive disorder [OCD], enuresis, school phobia, anxiety problems, etc.) have been conducted with children, with few investigations that focus on the treatment of depressed adolescents.

It is important to recognize that the administration of antidepressant medications does not result in a cure of the depression. Antidepressant drugs will show therapeutic effects in some youngsters; however, typically the symptoms of depression will return if the treatment is discontinued during the course of the depressive illness (Johnston & Fruehling, 1994).

The research on the efficacy of antidepressants with children and adolescents is growing as newer drugs and classes of antidepressants are placed on the market. Rather than review each of the dozen or so controlled studies of antidepressant therapy with children and adolescents, this chapter provides a brief review of the general findings and some of the issues inherent in the use of pharmacological agents with children and adolescents. It should be noted that in addition to pharmacotherapy, reports of electroconvulsive therapy (Bertagnoli

& Borchard, 1990) and light/phototherapy for seasonal affective disorder (seasonal pattern) (Sonis, Yellin, Garfinkel, & Hoberman, 1987) with depressed youngsters have been published. Light therapy has also been suggested as an adjunct in the treatment of depressive symptoms in bipolar disorder in adolescents (Papatheodorou & Kutcher, 1995). This chapter does not review studies of pharmacotherapy for treatment of bipolar disorder in children and adolescents. For bipolar disorder, lithium is one of the more commonly used drugs along with carbamazepine. The treatment of bipolar disorder in youngsters as well as adults is a difficult undertaking and frequently involves the testing of various drugs, singly and in combination. Data from research and clinical studies suggest that pharmacological agents used with adults for the treatment of bipolar disorder work about as effectively with young people (Carlson, 1994).

5.17.6.1 Pharmacotherapy for Depression in Children and Adolescents

A large number of case reports and a limited number of experimental studies have been published examining the efficacy of tricyclic antidepressants (TCAs), monoamine oxidase inhibitors, selective serotonin reuptake inhibitors (SSRIs), and other drugs for the treatment of unipolar depression, and lithium carbonate and other compounds for the treatment of bipolar disorder in children and adolescents. Pharmacotherapy of depression in children and adolescents is somewhat controversial due to the limited treatment response over placebo reported in numerous double-blind clinical trials. Reviews of the pharmacological studies for depression in young people can be found in Alessi, Naylor, Ghaziuddin, and Zubieta (1994), Jensen, Ryan, and Prien (1992), and Johnston and Fruehling (1994).

Since the mid-1980s, a number of newer antidepressants have appeared on the market with demonstrated efficacy for adults. Of particular note have been the SSRIs including fluoxetine, fluvoxamine, sertraline, and paroxetine. There have been reports of fluoxetine use with youngsters (e.g., Dech & Budow, 1991; Naylor & Grossman, 1991), although at the time of this chapter the author could not find published controlled group outcome studies using these agents for depression in youngsters. Initial research with fluoxetine by Riddle, Hardin, King, Seahill, & Woolston (1990), with a sample of 10 youngsters with Tourette's syndrome and OCD, found mixed results specific to treatment of obsessive compulsive symptoms. Significant agitation was noted as a response to fluoxetine in four of the 10 participants, but there were minimal adverse biological side-effects. The latter may be expected, given the general lack of anticholinergic and antihistaminic side-effects with fluoxetine. Riddle et al. (1989) also reported a case of a fluoxetine overdose in an adolescent. Open trials of SSRIs have been reported by Apter et al. (1994), Boulos, Kutcher, Gardner, and Young (1992), and others. Initial reports suggest significant clinical efficacy for fluoxetine, even with youngsters who previous failed to respond to TCAs (Boulos et al., 1992).

By far, the majority of published, control group investigations of antidepressants for the treatment of depression in children and adolescents has utilized tricyclic antidepressants such as imipramine, amitriptyline, nortriptyline, and desipramine. Researcher reports on TCAs for the treatment of depression in children began to appear in the late 1960s and early 1970s (Frommer, 1967; Kuhn & Kuhn 1971; Ling et al., 1970; Weinberg et al., 1973). Stack (1971) reported on 490 depressed children, including 116 preschool children, treated with antidepressants, most of which were TCAs.

A general finding across controlled antidepressant treatment studies of depression in children and adolescents has been the lack of superiority of medications over placebo control conditions. Rationales for this lack of treatment efficacy have ranged from invalidity of the diagnosis of depression in children using current classification systems to errors in design methodology and insufficient sample size (Conners, 1992; Jensen et al., 1992). With regard to the latter, most of the published studies utilized relatively small sample treatment groups, such that power to detect a significant difference is often compromised. However, in a recent study by Kutcher et al. (1994) of desipramine for major depression in adolescents, 60 adolescents were studied after using a placebo washout phase prior to treatment. In this well-controlled investigation, antidepressant and placebo groups did not differ across a wide range of outcome measures including psychometric as well as laboratory assessments. A similar finding was reported by Kye et al. (1996) in a controlled study of amitriptyline for major depression in 31 adolescents (22 completers) with major depression, using a two-week placebo-washout period.

5.17.6.2 Summary of Pharmacological Interventions

The overriding conclusion from the majority of double-blind placebo-controlled studies is that the scientific data indicates antidepressants are not significantly more effective than placebo conditions in the treatment of unipolar depression in children and adolescents (Ambrosini, Bianchi, Rabinovich, & Elia, 1993; Elliot, 1992; Johnston & Fruehling, 1994). This has caused clinical and research dissonance, given the large number of youngsters who are treated with antidepressants by clinicians who report individual clinical efficacy (Fisher & Fisher, 1996). Johnston and Fruehling (1994) provided a comprehensive review of studies of antidepressant medications for the treatment of depression in children and adolescents up to that time. These authors note the overwhelming use of antidepressants in clinical practice and the lack of scientific evidence to support efficacy. With so many clinicians using these agents, Johnston and Fruehling (1994) suggest that clinicians may be seeing a level of clinical usefulness not evident in the drug studies. They also recommend that for most youngsters with depression “antidepressant medication should not be the ‘first-line’ treatment.” The dilemma of scientific evidence and clinical use of antidepressants has been debated (Eisenberg, 1996; Fisher & Fisher, 1996). It is evident that the jury is out as to whether antidepressants can be recommended for use with young people. This supports the need for further study of psychosocial interventions as well as the examination of the utility of combined pharmacotherapy–psychotherapy treatments for depressed young people. The evaluation of pharmacological treatments for depression in children and adolescents is complex, and the issue of side-effects, although not dealt with in this review, is significant and a major consideration in the use of antidepressant medication.

Lastly, as noted earlier, several case reports and an open trial have examined the combined efficacy of pharmacotherapy and psychotherapy. Given that psychotherapy is often seen as an important component of treatment along with antidepressant medications, there needs to be greater focus on the clinical outcome of combined pharmacotherapy and psychotherapy with depressed children and adolescents. The combinatory approach to treatment, that is, the use of both pharmacotherapy and psychotherapy, is consistent with the biopsychological perspective of childhood depression. The biopsychosocial approach suggested by Vasile et al. (1987) makes use of a consultation model for the recommendation of psychological, pharmacological, or combined therapies. This integrative and evaluative approach appears useful for the treatment of depressed children and adolescents.

5.17.7 COMORBIDITY OF DEPRESSION AND OTHER PSYCHOPATHOLOGY IN CHILDREN AND ADOLESCENTS

The study of the coexistence of other psychopathology along with depression in children, adolescents, and adults has emerged as a significant field of inquiry (Anderson & McGee, 1994; Caron & Rutter, 1991; Kendall & Clarkin, 1992). Comorbidity is a term originally used in medicine to describe the existence of a distinct, separate disease or disorder that is present during the course of a primary disorder or disease that is the focus of attention.

In the study of psychiatric disorders in children and adolescents, the finding of comorbidity is becoming increasingly evident. In a study of 476 new admissions by adolescents (mean age 15.6 years) to inpatient and outpatient settings in 29 clinical settings in 21 states, multiple diagnoses (two or more clinical disorders) at intake were found in 67% of the sample, with three or more concurrent disorders found in 26% of the sample (Reynolds, 1997b). Pfeffer and Plutchik (1989) in a study of comorbidity in children, found an average of 2.38 diagnoses in a sample of 106 inpatients and a mean of 2.35 diagnoses in a sample of 101 outpatients. The authors found 39 children with a primary diagnosis of major depression, only three (7.6%) of whom did not have a coexisting *DSM-III* Axis I or Axis II disorder. The most frequent comorbid disorder was dysthymic disorder (56.4%), followed by borderline personality disorder (46.2%), specific developmental disorder (38.5%), and conduct disorder (28.2%).

Alessi and Magen (1988), in a study of 160 child psychiatric inpatients, found that 48 (30%) children had three or more Axis I disorders. A higher rate of co-occurrence of three or more disorders was found among the 25 youngsters with a depressive disorder (52%) compared to the 135 children in the nondepressed psychiatric group (26%). The depressed youngsters included 16 children with major depression (MDD) and nine children with dysthymic disorder. Although the comorbidity of other Axis 1 and several Axis II disorders did not differ in proportion to that found in children with other diagnoses, there were some differences in comorbidity between youngsters with major depression and dysthymic disorder. Nearly 89% of dysthymic children also manifested an anxiety disorder, as compared to 25% of MDD children. The children with major depression also demonstrated a similar frequency of comorbidity with attention deficit disorder and conduct disorder. The high comorbidity between dysthymic disorder and anxiety disorders suggests that there may be some degree of nosological similarity between these forms of psychopathology.

The comorbidity with major depression was examined by Mitchell, McCauley, Burke, and Moss (1988) separately for children ($n = 45$) and adolescents ($n = 50$) who were seen in a clinical setting. Diagnoses were based on RDC using the K-SADS. The most frequent comorbid disorder in both children and adolescents was separation anxiety (42% and 44%, respectively). Conduct disorder was comorbid in 16% of children and 14% of adolescents. However, unlike separation anxiety that did not show significant gender differences in its presentation, conduct disorders were only present in preadolescent boys, whereas in adolescents, a somewhat greater proportion of girls (17%) than boys (10%) had conduct disorder comorbid with major depression.

One of the most noteworthy investigations of psychopathology comorbidity in adolescents is a study conducted by Lewinsohn and colleagues (Lewinsohn et al., 1995; Rohde, Lewinsohn, & Seeley, 1991) in a western Oregon community. This study used diagnostic methods to evaluate a community sample of adolescents for *DSM-III-R* psychiatric disorders, incorporating two assessment periods approximately 14 months apart. Rohde et al. (1991) examined prevalence of current and lifetime depression and comorbidity in a sample of 1710 adolescents. There were 50 adolescents with a current diagnosis of major depression and 347 with lifetime prevalence of depression. There were 21 (42.0%) of the depressed youngsters who demonstrated one or more additional psychiatric disorders, with anxiety disorders (18%), disruptive behaviors (conduct disorder and oppositional defiant disorder) (8%), substance use (14%), and eating disorders (4%) present. For the lifetime estimates of comorbidity, 42.9% of adolescents with a lifetime diagnosis of major depression also had experienced one or more additional psychiatric disorders. In their 1995 (Lewinsohn et al., 1995) report on a sample of 1507 youngsters aged 14 to 18, there were 342 with a lifetime diagnosis of major depression. Of these adolescents, there were 67 with comorbid anxiety disorders, 78 with substance abuse disorder, and 36 with disruptive behavior. It should be noted that the authors did not include cases with other comorbid disorders, such as eating disorders and schizophrenia. Of the group with major depression, 205 youngsters did not demonstrate other psychopathology. Lewinsohn et al. (1995) examined the proportion of youngsters who were receiving treatment by the number of diagnoses. Of youngsters with a psychiatric diagnosis without comorbidity (i.e., pure cases) 24.8% were receiving treatment. There was a significant increase in the proportion of adolescents receiving treatment as a function of comorbidity (50.4% for two disorders and 65.9% for three or more disorders).

Another major study of comorbidity in children with depressive disorders is the longitudinal investigation of Kovacs and colleagues (Kovacs, 1990; Kovacs et al., 1984a, 1984b; Kovacs, Akiskal, Gatsonis, & Perrone, 1994; Kovacs, Gatsonis, Paulauskas, & Richards, 1989; Kovacs, Paulauskas, Gatsonis, & Richards, 1988) who studied a cohort of 104 children aged 8–13 presenting diagnoses of major depressive disorder ($n = 46$), dysthymic disorder ($n = 23$), combined major depressive disorder and dysthymic disorder (double depression; $n = 16$), and adjustment disorder with depressed mood. In this cohort, significant comorbid conduct and anxiety disorders were also noted.

5.17.7.1 Comorbidity and Biological Communalities

The mechanisms for comorbidity are probably quite complex, and are beginning to be hypothesized. For instance, Lam, Solyom, and Tompkins (1991) found similarities in seasonal symptom variation among bulimic patients to those found among persons with seasonal affective disorder and hypothesized that there may be a common neurobiological abnormality, possibly involving serotonergic dysfunction. Genetic predisposition has also been noted as a potential explanation for some patterns of comorbidity. Walters et al. (1992), using evidence from twin studies suggests that there may be a partial shared common genetic basis for bulimia nervosa and major depression. Mufson, Weissman, and Warner (1992) examined comorbid depression and anxiety in children of parents with major depression, major depression and panic disorder, and other disorders, finding evidence for familial transmission. Similarly, Paul (1988), in examining the comorbidity between anxiety disorders and depression and a range of biological aspects of these disorders, suggests that anxiety and depression may represent different phenotypes of a common neurobiological substrate.

5.17.7.2 The Range of Disorders Comorbid with Depression

Case reports have illustrated a wide-range of disorders that may co-occur with depression in children and adolescents, including: trichotillomania (Naylor & Grossman, 1991; Weller, Weller, & Carr, 1989); autism (Komoto, Usui, & Hirata, 1984); borderline personality disorder (Stone, 1981); and conversion symptom (Weller & Weller, 1983), the latter obscuring the depressive disorder.

Most research on the comorbidity of depression and other psychiatric disorders in children and adolescents has been based on clinical samples with a few noteworthy community-based studies, such as that of Lewinsohn and colleagues (1995). In the study of depression and its comorbidity, a particularly ubiquitous relationship with anxiety disorders has been evident and generated a tremendous amount of research and debate (Brady & Kendall, 1992; Kendall, Kortlander, Chansky, & Brady, 1992; King, Gullone, & Ollendick, 1990; Weissman, 1990). In addition to anxiety disorders, researchers have reported comorbidity of depression in youngsters with attention deficit disorder (Alessi & Magen, 1988; Biederman, Munir et al., 1987; Jensen et al., 1988), conduct disorders (Alessi & Magen, 1988; Kovacs, Paulauskas, et al., 1988; Puig-Antich, 1982), substance abuse (De Milio, 1989; Levy & Deykin, 1989), and eating disorders (Alessi, Krahn, Brehm, & Wittekindt, 1989).

5.17.7.3 Comorbidity with Other Internalizing Disorders and Problems

Depressive disorders, by their symptom definitions, share a number of features with other disorders such as eating disorders and anxiety disorders. By its definition, comorbidity refers to a different case as may occur when a relationship between two disorders is found as a function of shared symptoms alone. However, it may be that certain disorders also share various precursors or etiological factors that result in an increased risk for their co-occurrence. The disorders presented below were selected in large part as a function of the research literature. It is evident that other disorders may also show some level of comorbidity, but by their nature or incidence/prevalence have not been a major focus of study.

5.17.7.3.1 Anxiety disorders

Anxiety disorders represent a major domain of psychopathology and a wide range of clinical disorders. One of the earliest reports of childhood depression and other psychopathology focused on the relationship between school phobia and childhood depression. Agras (1959) described seven children aged 6–12 years who demonstrated school phobia, six of whom showed significant depressive symptoms. It is interesting to note that the comorbidity of depression and symptoms of anxiety was viewed by Agras (1959, p. 534) as similar "phenomenologically to the depressive disorders of adults."

School refusal or school phobia has continued to be linked or found comorbid with depression (Bernstein & Garfinkel, 1986; Naylor, Staskowski, Kenney, & King, 1994). In a study of 35 children with school phobia, Waldron, Shrier, Stone, and Tobin (1975) found that a significantly greater number of these youngsters manifested depression (56%) than a comparison group of children diagnosed as having other neuroses (26%). With a matched sample of 28 inpatient and 28 outpatient youngsters who were evaluated as school refusers, Borchardt, Giesler, Bernstein, and Crosby (1994) found a diagnosis of major depression in 85.7% of the inpatients and 46.4% of the outpatients, the only comorbid diagnosis that was significantly different in proportions between the two groups.

In a study of the comorbidity of depressive disorders with anxiety disorders, Kovacs, Gatsonis et al. (1989) reported that 41% of their sample of 104 youngsters with major depression and/or dysthymic disorder had an anxiety disorder at the initial (index) study episode of depression, with separation anxiety and overanxious disorder the most frequent comorbid diagnoses. In 19 of 30 participants with major depression the anxiety disorder predated the onset of depression. Strauss, Last, Hersen, and Kazdin (1988) found that 28% of 106 children and adolescents with a *DSM-III* diagnosis of an anxiety disorder had a concomitant diagnosis of major depression.

5.17.7.3.2 Eating disorders

A great deal of discussion and research has focused on the comorbidity and relationship between depression and eating disorders (Swift, Andrews, & Barklage, 1986). Rastam (1992) in a study comparing 51 adolescents with anorexia nervosa to a matched control group of school-based adolescents found that 69% of the former also met criteria for major depression or dysthymic disorder, compared to 10% of the school-based comparison group. They did not find clear evidence to suggest that depression was an antecedent of anorexia. Smith and Steiner (1992) found comorbid major depression in 53% of a sample of adolescent girls seen in an eating disorders clinic. In a six year follow-up of 23 anorexic youngsters, Smith, Feldman, Nasserbakht, and Steiner (1993) found a lifetime prevalence of major depression in 86.9% of the sample.

A significant study of comorbidity was presented by Herzog, Keller, Sacks, Yeh, and Lavori (1992) with 229 females seen in a treatment setting with anorexia nervosa ($n = 41$), bulimia nervosa ($n = 98$) and mixed anorexia nervosa and bulimia nervosa ($n = 90$) using RDC criteria for diagnosis. They found that major depression was the most common comorbid disorder. For the total sample, at least one affective disorder was found in 63% of the cases, with 41% (94 cases) manifesting major depression. Patients with both anorexia and bulimia demonstrated the highest rate of comorbidity with major depression (53%) in comparison to the other eating disorder groups, and 76% of this mixed group had at least one affective disorder. It is also of interest that the mixed group was younger in age than the other groups at onset of the eating disorder ($M = 17.5$, versus 19.1 for anorexia and 18.8 for bulimia).

In a school-based sample of 3287 adolescents in France, Ledoux, Choquet, and Manfredi (1993) found a relatively high rate of depressive symptom endorsement among youngsters who engaged in binge eating (20.2%) compared to non-binge eaters (6.3%) $\chi^2 = 45.06, p < 0.001$. There was also a gender difference, with a higher proportion of binge-eating girls reporting high levels of depressive symptoms (26.6%) compared to binge-eating boys (10.6%).

In addition to clinic and school-based studies, case studies have also reported comorbidity between eating disorders and depression. In a somewhat unique case report, Alessi et al. (1989) reported on a case of a nine-year-old girl with anorexia nervosa and major depression. The child was hospitalized for 10 weeks and received psychotherapy. Significant decreases in scores on the CDRS-R (from 76 at intake to 20 at week nine of treatment) and increases in weight from 52.5 pounds to 70 pounds at discharge were reported with the youngster maintaining this weight for two weeks. Although this case is considered rare, it does show that comorbidity of these disorders may have an early onset in some youngsters.

5.17.7.4 Comorbidity with Externalizing Disorders and Problem Behaviors

The comorbidity of depression with externalizing disorders is a significant area of research, with questions as to why such diverse forms of psychological disorders co-occur and what underlying mechanism may be responsible for this outcome. Such questions remain to be answered. For the present, the comorbidity of depression and conduct disorders, substance abuse, and attention-deficit hyperactivity disorder (ADHD) are examined.

5.17.7.4.1 Conduct disorder

There is a substantial literature base on the coexistence of depression and conduct disorder, the latter sometimes operationalized as juvenile delinquency. In the early 1970s, a number of investigators were identifying depression among adolescents seen in clinical settings for delinquent youth (e.g., Weinberg & van Den Dungen, 1971). Some of the studies discussed below were based on youngsters adjudicated as delinquent, the majority of whom we may assume meet criteria for conduct disorder, although this may not apply to all youngsters.

In a study using RDC diagnosis of depression in a sample of juvenile delinquents, Chiles, Miller, and Cox (1980) found 28 (23%) of a sample of 120 adolescents met criteria for major depression. Somewhat lower rates were reported by Alessi et al. (1984) who found 15.5% of a sample of 71 incarcerated juvenile offenders had a RDC diagnosis of major depressive episode and another 15.5% had a diagnosis of minor depressive disorder. An additional 8.5% were diagnosed having a major depressive disorder in remission. Kovacs, Paulauskas, Gatsonis, and Richards (1988) found that in their sample of 104 children, 17 (16%) had a comorbid conduct disorder at the time of the study initial (index) depressive episode, with another seven youngsters who received a diagnosis of conduct disorder during the course of the study and after remission of the index episode of depression. There were no differences in proportion of comorbidity between types of depressive disorders.

Kutcher, Marton, and Korenblum (1989) reported on the psychiatric morbidity of adolescents admitted to an inpatient psychiatric unit. After excluding four youngsters with conduct disorder without comorbid Axis I disorders, the authors found 26 of the remaining 96 adolescents had a conduct disorder comorbid with other Axis I disorders. Comorbid disorders to the conduct disorder included depression (five cases), bipolar disorder (eight cases), psychotic disorders (seven cases) and other disorders (six cases).

Within the same domain of disruptive behaviors, researchers have also found oppositional defiant disorder (ODD) comorbid with depression. Wenning, Nathan, and King (1993) found that 14 of 28 children (aged 6–13) seen in a clinical setting who had a diagnosis of ODD also met criteria for dysthymia. None of the children had episodes of major depression. The authors, in a retrospective examination of intake diagnoses, indicated that only one of the 14 children with ODD and dysthymia were so diagnosed at intake. The remaining 13 children were found to also have had dysthymia at intake, that had been overlooked and only identified by interviews with therapists who were given the opportunity to review *DSM-III-R* mood disorder criteria.

In a study of 56 child and adolescent psychiatric inpatients with major depressive disorder, Meller and Borchardt (1996) found 41% with comorbid conduct disorder. An examination of symptoms that differentiated between depression only and comorbid groups indicated that anxious mood was present to a greater extent in youngsters with depression only.

5.17.7.4.2 Attention deficit hyperactivity disorder

The coexistence of hyperactivity and depression in children has been noted since the 1970s, with a wide range in rates of comorbidity (Biederman, Newcorn, & Sprich, 1991). Brumback and Weinberg (1977) studied 223 children aged 6–12 years who were referred to an educational diagnostic center. Using Weinberg criteria for the diagnosis of depression, 136 children were depressed with 86 (63%) of these children also demonstrating hyperactivity. The authors noted that for some of the youngsters with depression and hyperactivity, their hyperactivity was somewhat episodic with depression, worsening during depressive episodes and decreasing as depression remitted.

Weinberg and Emslie (1991) described the diagnostic status of 100 children and adolescents referred to a pediatric neurology clinical setting, finding a diagnosis of ADHD in 63 youngsters. Of these 63 youngsters, 46 (73%) also had a diagnosis of depression, with the majority of these youngsters ($n = 40$) also manifesting developmental-specific learning disorders. The authors also noted that in a group of 65 children recently referred to their clinical setting who had a *DSM-III-R* diagnosis of major depression, 21 (32%) also had a diagnosis of ADHD, with the ADHD preceding the depressive disorders in all cases. Similarly, Biederman et al. (1994) in a study of 140 boys found major depression comorbid in 39 (27.9%) youngsters.

ADHD has also been noted in adolescents with bipolar disorder, with West, McElroy, Strakowski, Keck, and McConville (1995) finding eight (57%) of 14 bipolar patients with *DSM-III-R* mania or hypermania demonstrating a comorbid ADHD. In all cases, the ADHD predated the onset of bipolar disorder, by an average of six years. The prior presentation of ADHD and the time between onset of disorders is consistent with ADHD typically identified in childhood, and bipolar disorder often emerging during adolescence.

5.17.7.4.3 Substance abuse

When studying comorbidity in adolescents, a relatively common finding is the coexistence of substance abuse and psychiatric illness. Depressive disorders are among the most studied disorders comorbid with substance abuse (e.g., Bukstein, Brent, & Kaminer, 1989), although conduct disorders and substance abuse tend to show higher rates of comorbidity (Kaminer, Tarter, Bukstein, & Kabene, 1992). The study of substance abuse and depression in young people is a very important domain of research as well as clinical attention. Substance abuse is a major problem in adolescents. In the study of psychopathology accompanying substance abuse, depression and conduct disorders appear to be the most concomitant disorders (Kaminer, 1991).

As noted, there have been numerous reports of depressive disorders concurrent with substance abuse. Kashani, Keller, Solomon, Reid, and Mazzola (1985) reported on 16 cases of double depression as well as another 12 youngsters with either major depressive disorder or dysthymic disorder in a sample of 100 adolescents with substance abuse disorders. The comorbidity of affective disorders and other psychiatric disorders in an inpatient psychiatric sample of 156 adolescents aged 13–18 with substance abuse was examined by Bukstein, Glancy, and Kaminer (1992) who found an affective disorder in 51.3% of the sample, second to conduct disorder which was found in 70.5% of the youngsters. Of the mood disorders identified, 30.7% demonstrated major depression, 5.1% dysthymia, 7.7% bipolar, and there were 9.0% of the sample who were comorbid with adjustment disorder with depressed mood. A significant number (22 of the 48) of the youngsters with major depression and substance abuse also had an additional diagnosis of conduct disorder. Overall, nearly 27% of the sample of substance abusers also had concurrent diagnoses of conduct disorder and an affective disorder. In determining the timing of the onset of disorders, in the major depression–substance abuse only group, the substance abuse preceded the onset of major depression by one year.

In a study of 52 inpatient adolescent substance abusers, Hovens, Cantwell, and Kiriakos (1994) found major depression and dysthymia present in 25% and 33% of youngsters, respectively. Of particular interest was the examination of the temporal sequence of substance abuse and psychiatric disorders. The authors found that substance abuse preceded the cases with major depression by about $3^1/_2$ years. A different picture was evident for youngsters with dysthymia, with dysthymia emerging prior to the onset of substance abuse in approximately 53% of the cases.

Stowell and Estroff (1992) examine the discharge diagnosis of 226 consecutive adolescent admissions (aged 12–18 years) to two inpatient treatment settings for substance abuse problems. Mood disorders were found in 61% of substance abusers, with 18% comorbid with major depression, 34% dysthymia, 9% bipolar, and 12% double depression with both major depression and dysthymia. Based on self-report data, the authors suggested that depression preceded the regular use of substances in 68% of the sample. However, this is not to suggest a causal pathway. Stowell and Estroff also noted that significant psychosocial stressors—such as sexual and physical abuse, parental substance abuse, divorce and separation, death of family members or close friends, and other stressors—may be contributory factors. In such cases, substance abuse became greater after such psychosocial stressors and exacerbated depressive symptoms.

Similar rates of depression among substance abusers were reported by Deykin, Buka, and Zeena (1992) with a sample of 223 chemically dependent adolescents in a residential treatment setting. They found 24.7% of the sample met *DSM-III-R* criteria for major depression, with a significantly higher proportion of females (48.2%) compared to males (16.9%) with major depression. With regard to temporal order of onset, 35.2% of those with depression developed chemical dependence following the depression, with chemical dependence preceding depression in 43.1% of the sample and occurring at the same time in 21.6% of the sample of comorbid youngsters. Other investigators studying youngsters with substance abuse have reported rates of major depression of 35% (DeMilio, 1989) with presence of any affective disorder noted as high as 78% (McKay, Murphy, Maisto, & Rivinus, 1992).

5.17.7.5 Diagnostic Overshadowing

A significant issue in the study of comorbidity is that of diagnostic overshadowing. In the Wenning et al (1993) investigation, the majority (93%) of the children with comorbid ODD and depression were not identified as being depressed as well as oppositional at intake. Likewise, other externalizing disorders such as ADHD may mask a depressive disorder or, as Weinberg and Emslie (1991) suggest, children referred for ADHD may not be systematically evaluated for the presence of additional psychopathology. A similar recommendation was made by McClellan, Rubert, Reichter, and Sylvester (1990) who found that a significant number of children who had an anxiety or depressive disorder also demonstrated attention deficit disorder (ADD). These authors suggested that clinicians evaluating children referred for ADD should consider the potential for a comorbid or primary affective disorder.

The above are useful examples of diagnostic overshadowing. Such cases may most frequently occur when an internalizing disorder such as depression is masked (in an overt sense) by the extremes and behavioral excesses found in most externalizing disorders.

5.17.7.6 Comorbidity Summary

The extensive findings of comorbidity among children and adolescents with psychiatric disorders, including depressive disorders, has important implications for our understanding of depressive disorders in young people, as well as approaches to identification, treatment, and prevention.

It is evident that when we study comorbidity within select clinical samples, such as in groups of youngsters referred to an eating disorders clinic or substance abusers in a chemical dependency treatment setting, rates of comorbidity will differ from those found in large-sample studies of depressed youngsters. Furthermore, data regarding comorbidity based on clinic samples may be biased and represent somewhat inflated values as a function of increased treatment seeking due to greater distress caused by two or more disorders, increased likelihood of referral, and other selection biases (du Fort, Newman, & Bland, 1993). Likewise, certain types of comorbidity may be seen to a greater extent in treatment and other clinical settings. For instance, depression comorbid with externalizing disorders, such as conduct disorder, ADHD, or substance abuse, is at increased risk for referral or identification by others. To more accurately understand the comorbidity between two disorders, data from large, representative samples drawn from the general population are needed.

A question in the study of comorbidity is that of the temporal relationship between disorders. That is, do consistent findings suggest that one disorder commonly predates another, and if so, does this indicate a predisposition of one disorder to another? As described in this chapter, a few studies have examined this by investigating the onset of disorders. For the most part, findings suggest that in the cases of certain comorbid disorders that may develop during certain developmental periods, there is little evidence to suggest that the initial disorder is a precursor to the emergence of the second disorder. For instance in the case of bipolar disorder with comorbid ADHD, the ADHD which typically is diagnosed in childhood predates the onset of bipolar disorder which tends to emerge during adolescence and young adulthood. For other disorders, the onset in reference to the index episode of depression may occur prior, during, or subsequent to the depressive episode.

From the above studies, it is quite clear that many youngsters with depression also manifest other forms of psychopathology, both of internalizing and externalizing phenomenology. Toward understanding the nature of comorbidity, investigators have begun to examine the course and potential causal pathways of depression and comorbid psychopathology (Capaldi, 1991; Caron & Rutter, 1991; Loeber, Russo, Stouthamer-Loeber, & Lahey, 1994).

The examination of the coexistence of other psychopathology and depression in children and adolescents is important for a number of reasons. As Alessi and Magen (1988) noted, the existence of multiple disorders may in some individuals lead to the development of a depressive disorder. Likewise, the co-occurrence may create an increased stress response with psychological and biological interactions. Clearly the biopsychosocial approach is critical when examining comorbidity.

5.17.8 SUMMARY AND CONCLUSIONS

In this chapter a biopsychosocial perspective was presented as a framework for understanding some of the complexities inherent in our knowledge of depression in children and adolescents. The biopsychosocial model provides a scaffold for examining the features, correlates, and etiologies of depression in young people. In this regard, it can be seen and understood that depression in children and adolescents is a complex interaction of social and environmental factors, predispositions, cognitive and behavioral skills and competencies, biological vulnerability factors, and other individual difference variables.

The importance for understanding the nature, identification, and treatment of depression in children and adolescents is best understood by the basic fact that depressed youngsters feel intense psychological distress and misery. Depressed children and adolescents face daily impairments in psychological functioning and often perceive that there is no hope for improvement. Research suggests that for many depressed youngsters, the long-term outcome is poor, with the likely development of more severe pathology in adulthood (Harrington, 1996; Kovacs, 1996a, 1996b).

The past two decades have provided professionals with a remarkable amount of information on depression as an internalizing disorder in children and adolescents. This has resulted in the development of measures that are now used by clinicians for the identification of depressed youngsters, as well as the publication of empirically tested treatment manuals for the provision of psychotherapy. Advances in the understanding of psychobiological correlates of depression in children and adolescents, and advances in pharmacological interventions also provide insight into the basic nature of this disorder in young people. Of domains in which basic and applied research needs to be conducted, the understanding of the developmental nature of depression in children remains an area of critical importance. Research and theoretical descriptions on the developmental nature of depression in children have begun to appear. However, there are numerous questions specific to the expression and syndromal description of depression in children and adolescents that still need to be answered.

Likewise, we are just beginning to understand the course of depression, although much of this is based on a few well-conceived and executed studies. However, given the heterogeneity of depressive disorders and the myriad etiological and interactive variables that come into play, we will require extensive investigation across numerous samples to determine how depressive disorders, both syndromal and subsyndromal, affect children and adolescents.

5.17.9 REFERENCES

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87,* 49–74.

Adams, D. M., Overholser, J. C., & Spirito, A. (1994). Stressful life events associated with adolescent suicide attempts. *Canadian Journal of Psychiatry, 39,* 43–48.

Adams, J., & Adams, M. (1993). Effects of a negative life event and negative perceived problem-solving alternatives on depression in adolescents: A prospective study. *Journal of Child Psychology and Psychiatry, 34,* 743–747.

Agras, S. (1959). The relationship of school phobia to childhood depression. *American Journal of Psychiatry, 116,* 533–536.

Akiskal, H. S. (1979). A biobehavioral approach to depression. In R. A. Depue (Ed.), *The psychobiology of the depressive disorders* (pp. 409–437). New York: Academic Press.

Alessi, N. E., Krahn, D., Brehm, D., & Wittekindt, J. (1989). Prepubertal anorexia nervosa and major depressive disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 380–384.

Alessi, N. E., & Magen, J. (1988). Comorbidity of other psychiatric disturbances in depressed psychiatrically hospitalized children. *American Journal of Psychiatry, 145,* 1582–1584.

Alessi, N. E., McManus, M., Grapentine, W. L., & Brickman, A. (1984). The characterization of depressive disorders in serious juvenile offenders. *Journal of Affective Disorders, 6,* 9–17.

Alessi, N., Naylor, M. W., Ghaziuddin, M., & Zubieta, J. K. (1994). Update on lithium carbonate therapy in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 291–304.

Ambrosini, P. J., Bianchi, M. D., Rabinovich, H., & Elia, J. (1993). Antidepressant treatments in children and adolescents: l. Affective disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 1–5.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. Rev.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anderson, J. C., & McGee, R. (1994). Comorbidity of depression in children and adolescents. In W. M. Reynolds and H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 581–601). New York: Plenum.

Apter, A., Ratzoni, G., King, R. A., Weizman, A., Iancu, I., Binder, M., & Riddle, M. A. (1994). Fluvoxamine open-label treatment of adolescent inpatients with obsessive-compulsive disorder or depression. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 342–348.

Archer, R. P., Maruish, M., Imhof, E., & Piotrowski, C. (1991). Psychological test usage with adolescent clients: 1990 survey findings. *Professional Psychology: Research and Practice, 22,* 247–252.

Bandura, A. (1978). The self-system in reciprocal determinism. *American Psychologist, 33,* 344–358.

Barrera, M., & Garrison-Jones, C. V. (1988). Properties of the Beck Depression Inventory as a screening instrument for adolescent depression. *Journal of Abnormal Child Psychology, 16,* 263–273.

Bartell, N. P. & Reynolds, W. M. (1986) Depression and self-esteem in academically gifted and nongifted children: A comparison study. *Journal of School Psychology, 24,* 55–61.

Bath, H. I., & Middleton, M. R. (1985). The Children's Depression Scale: Psychometric properties and factor structure. *Australian Journal of Psychology, 37,* 81–88.

Beck, A. T. (1967) *Depression: Causes and treatment*. Philadelphia: University of Pennsylvania Press.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.

Beck, A. T., Ward, C., Mendelson, M., Mock, J., & Erbaugh, J. (1961) An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561–571.

Belsher, G., Wilkes, T. C. R., & Rush, A. J. (1995). An open, multisite pilot study of cognitive therapy for depressed adolescents. *Journal of Psychotherapy Practice and Research, 4,* 52–66.

Bemporad, J. R. (1994). Dynamic and interpersonal theories of depression. In W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 81–95). New York: Plenum.

Beres, D., & Alpert, A. (1940). Analysis of a prolonged hypomanic episode in a five-year-old child. *American Journal of Orthopsychiatry, 10,* 794–800.

Bernstein, G. A., & Garfinkel, B. D. (1986). School phobia: The overlap of affective and anxiety disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 25,* 235–241.

Bertagnoli, M. W., & Borchardt, C. M. (1990). A review of ECT for children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 302–307.

Biederman, J., Munir, K., Knee, D., Armentano, M., Autor, S., Waternaux, C., & Tsuang, M. (1987). High rate of affective disorders in probands with attention deficit disorder and in their relatives: A controlled family study. *American Journal of Psychiatry, 144,* 330–333.

Biederman, J., Lapey, K. A., Milberger, S., Faraone, S. V., Reed, E. D., & Seidman, L. J. (1994). Motor preference, major depression and psychosocial dysfunction among children with attention deficit hyperactivity disorder. *Journal of Psychiatric Research, 28,* 171–184.

Biederman, J., Newcorn, J., & Sprich, S. (1991). Comorbidity of attention deficit hyperactivity disorder with conduct, depressive, anxiety, and other disorders. *American Journal of Psychiatry, 148,* 564–577.

Birmaher, B., Rabin, B. S., Garcia, M. R., Jain, U., Whiteside, T. L., Williamson, D. E., Al-Shabbout, M., Nelson, B. C., Dahl, R. E., & Ryan, N. D. (1994). Cellular immunity in depressed, conduct disorder, and normal adolescents: Role of adverse life events. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 671–678.

Bland, R. C. (1994). Introduction. *Acta Psychiatrica Scandinavica, (Suppl. 376),* 5–6.

Borchardt, C. M., Giesler, J., Bernstein, G. A., & Crosby, R. D. (1994). A comparison of inpatient and outpatient school refusers. *Child Psychiatry and Human Development, 24,* 255–264.

Boulos, C., Kutcher, S., Gardner, D., & Young, E. (1992). An open naturalistic trial of fluoxetine in adolescents and young adults with treatment-resistant major depression. *Journal of Child and Adolescent Psychopharmacology, 2,* 103–111,

Brady, E. U., & Kendall, P. C. (1992). Comorbidity of anxiety and depression in children and adolescents. *Psychological Bulletin, 111,* 244–255.

Brand, E. F., King, C. A., Olson, E., Ghaziuddin, N., & Naylor, M. (1996). Depressed adolescents with a history of sexual abuse: Diagnostic comorbidity and suicidality. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 34–41.

Brosschot, J. F., Benschop, R. J., Godaert, G. L. R., Olff, M., De Smet, M., Heijnen, C. J., & Ballieux, R. E. (1994). Influence of life stress on immunological reactivity to mild psychological stress. *Psychosomatic Medicine, 56,* 216–224.

Brumback, R. A., & Weinberg, W. A., (1977). Relationship of hyperactivity and depression in children. *Perceptual and Motor Skills, 45,* 247–251.

Bukstein, O. G., Brent, D. A., & Kaminer, Y. (1989). Comorbidity of substance abuse and other psychiatric disorders in adolescents. *American Journal of Psychiatry, 146,* 1131–1141.

Bukstein, O. G., Glancy, L. J., & Kaminer, Y. (1992). Patterns of affective comorbidity in a clinical population of dually diagnosed adolescent substance abusers. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 1041–1045.

Butler, L., Miezitis, S., Friedman, R., & Cole, E. (1980).

The effect of two school-based intervention programs on depressive symptoms in preadolescents. *American Educational Research Journal, 17,* 111–119.

Campbell, J. D. (1952). Manic depressive psychosis in children: Report of 18 cases. *Journal of Nervous and Mental Disease, 116*, 424–439.

Capaldi, D. M. (1991). Co-occurrence of conduct problems and depressive symptoms in early adolescent boys: I. Familial factors and general adjustment at grade 6. *Development and Psychopathology, 3,* 277–300.

Carey, T. C., Kelley, M. L., & Carey, M. P. (1991, August). *The relation of cognitions, emotions and behavior to depressive symptomatology*. Paper presented at the Annual Meeting, American Psychological Association, San Francisco.

Carlson, G. A. (1994). Adolescent bipolar disorder: Phenomenology and treatment implications. In W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 41–60). New York: Plenum.

Caron, C., & Rutter, M. (1991). Comorbidity in child psychopathology: Concepts, issues, and research strategies. *Journal of Child Psychology and Psychiatry, 32*, 1063–1080.

Chiles, J. A., Miller, M. L., & Cox, G. B. (1980). Depression in an adolescent delinquent population. *Archives of General Psychiatry, 37,* 1179–1184.

Christ, A. E., Adler, A. G., Isacoff, M., & Gershansky, I. S. (1981). Depression: Symptoms versus diagnosis in 10,412 hospitalized children and adolescents (1957–1977). *American Journal of Psychotherapy, 35*, 400–412.

Clarizio, H. F. (1994). Assessment of depression in children and adolescents by parents, teachers, and peers. In W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 235–248). New York: Plenum.

Clarke, G., & Lewinsohn, P. M. (1989). The coping with depression course: A group psychoeducational intervention for unipolar depression. *Behaviour Change, 6,* 54–69.

Cole, D. A. (1991). Preliminary support for a competency-based model of depression in children. *Journal of Abnormal Psychology, 100,* 181–190.

Cole, D. A., & Carpentieri, S. (1990). Social status and the comorbidity of child depression and conduct disorders. *Journal of Consulting and Clinical Psychology, 58,* 748–757.

Compas, B. E., Ey, S., & Grant, K. E. (1993). Taxonomy, assessment, and diagnosis of depression during adolescence. *Psychological Bulletin, 114,* 323–344.

Conners, C. K. (1992). Methodology of antidepressant drug trials for treating depression in adolescents. *Journal of Child and Adolescent Psychopharmacology, 2,* 11–22.

Craighead, W. E., Curry, J. F., & Ilardi, S. S. (1995). Relationship of Children's Depression Inventory factors to major depression among adolescents. *Psychological Assessment. 7,* 171–176.

Crowley, S. L., & Emerson, E. N. (1996). Discriminant validity of self-reported anxiety and depression in children: Negative affectivity or independent constructs? *Journal of Clinical Child Psychology, 25,* 139–146.

Curry, J. F., & Craighead, W. E. (1990). Attributional style in clinically depressed and conduct disordered adolescents. *Journal of Consulting and Clinical Psychology, 58,* 109–115.

Dalley, M. B., Bolocofsky, D. N., Alcorn, M. B., & Baker, C. (1992). Depressive symptomatology, attributional style, dysfunctional attitude, and social competency in adolescents with and without learning disabilities. *School Psychology Review, 21,* 444–458.

Dech, B., & Budow, L. (1991). The use of fluoxetine in an adolescent with Prader–Willi Syndrome. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 298–302.

DeMilio, L. (1989). Psychiatric syndromes in adolescent substance abusers. *American Journal of Psychiatry, 146,* 1212–1214.

Dewsbury, D. A. (1991). "Psychobiology". *American Psychologist, 46,* 198–205.

Deykin, E. Y., Buka, S. L., Zeena, T. H. (1992). Depressive illness among chemically dependent adolescents. *American Journal of Psychiatry, 149,* 1341–1347.

du Fort, G. G., Newman, S. C., & Bland, R. C. (1993). Psychiatric comorbidity and treatment seeking: Sources of selection bias in the study of clinical populations. *Journal of Nervous and Mental Disease, 181,* 467–474.

Eisenberg, L. (1996). Commentary: What should doctors do in the face of negative evidence? *Journal of Nervous and Mental Disease, 184,* 103–105.

Elliott, G. R. (1992). Dilemmas for clinicians and researchers using antidepressants to treat adolescents with depression. *Journal of Child and Adolescent Psychopharmacology, 2,* 7–9.

Ellis, A. (1962). *Reason and emotion in psychotherapy*. New York: Lyle Stuart.

Emslie, G. J., Weinberg, W. A., Kennard, B. D., Kowatch, R. A. (1994). Neurobiological aspects of depression in children and adolescents. In W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 143–165). New York: Plenum.

Endicott, J., & Spitzer, R. L. (1978) A diagnostic interview: The schedule for affective disorders and schizophrenia. *Archives of General Psychiatry, 35,* 837–844.

Engel, G. L. (1978). The biopsychosocial model and the education of health professionals. *Annals of the New York Academy of Science, 310,* 169–181.

Engel, G. L. (1980). The clinical application of the biopsychosocial model. *American Journal of Psychiatry, 137,* 535–544.

Finch, A. J., Saylor, C. F., & Edwards, G. L. (1985). Children's Depression Inventory: Sex and grade norms for normal children. *Journal of Consulting and Clinical Psychology, 53,* 424–425.

Finch, A. J., Saylor, C. F., Edwards, G. L., & McIntosh, J. A. (1987). Children's Depression Inventory: Reliability over repeated administrations. *Journal of Clinical Child Psychology, 16,* 339–341.

Fine, S., Forth, A., Gilbert, M., & Haley, G. (1991). Group therapy for adolescent depressive disorder: A comparison of social skills and therapeutic support. *Journal of the American Academy of Child Psychiatry, 30,* 79–85.

Fisher, R. L., & Fisher, S. (1996). Antidepressants for children: Is scientific support necessary? *Journal of Nervous and Mental Disease, 184,* 99–102.

Fleming, J. E., & Offord, D. R. (1990). Epidemiology of childhood depressive disorders: A critical review. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 571–580.

Frommer, E. A. (1967). Treatment of childhood depression with antidepressant drugs. *British Medical Journal, 1,* 729–732.

Glaser, K. (1967) Masked depression in children and adolescents. *American Journal of Psychotherapy, 21,* 565–574.

Gotlib, I. H., Lewinsohn, P. M., & Seeley, J. R. (1995). Symptoms versus a diagnosis of depression: Differences in psychosocial functioning. *Journal of Consulting and Clinical Psychology, 63,* 90–100.

Hains, A. A. (1992). Comparison of cognitive-behavioral stress management techniques with adolescent boys. *Journal of Counseling and Development, 70,* 600–605.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry, 23,* 56–62.

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology, 6,* 278–296.

Harrington, R. C. (1996). Adult outcomes of childhood and adolescent depression: Influences on the risk for adult depression. *Psychiatric Annals, 26,* 320–325.

Harrington, R., Fudge, H., Rutter, M., Pickles, A., & Hill, J. (1990). Adult outcomes of childhood and adolescent depression: I. Psychiatric status. *Archives of General Psychiatry, 47,* 465–473.

Heiligenstein, E., & Jacobsen, P. B. (1988). Differentiating depression in medically ill children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 716–719.

Herdman, R. C., Gough, M., Liskowski, D. R., & Hall, L. L. (1995). *New developments in the biology of mental disorders.* Piscataway, NJ: Research & Education Association.

Herzog, D. B., Keller, M. D., Sacks, N. R., Yeh, C. J., & Lavori, P. W. (1992). Psychiatric comorbidity in treatment-seeking anorexics and bulimics. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 810–818.

Hodges, K. (1994). Evaluation of depression in children and adolescents using diagnostic clinical interviews. In W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 183–208). New York: Plenum.

Hodges, K., Kline, J., Stern, L., Cytryn, L., & McKnew, D. (1982). The development of a child assessment interview for research and clinical use. *Journal of Abnormal Child Psychology, 10,* 173–189.

Hollon, T. H. (1970). Poor school performance as a symptom of masked depression in children and adolescents. *American Journal of Psychotherapy, 24,* 258–263.

Horwath, E., Johnson, J., Klerman, G. L., & Weissman, M. M. (1994). What are the public health implications of subclinical depressive symptoms? *Psychiatric Quarterly, 65,* 323–337.

Hovens, J. G., Cantwell, D. P., & Kiriakos, R. (1994). Psychiatric comorbidity in hospitalized adolescent substance abusers. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 476–483.

Huebner, E. Scott, & Alderman, G. L. (1993). Convergent and discriminant validation of a children's life satisfaction scale: Its relationship to self- and teacher-reported psychological problems and school functioning. *Social Indicators Research, 30,* 71–82.

Ialongo, N., Edelsohn, G., Werthamer-Larsson, L., Crockett, L., & Kellam, S. (1993). Are self-reported depressive symptoms in first-grade children developmentally transient phenomena? A further look. *Development and Psychopathology, 5,* 433–457.

Jackson, S. W. (1986). *Melancholia and depression: From Hippocratic times to modern times.* New Haven, CT: Yale University Press.

Jensen, J. B., Burke, N., & Garfinkel, B. D. (1988). Depression and symptoms of attention deficit disorder with hyperactivity. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 742–747.

Jensen, P. S., Ryan, N. D., & Prien, R. (1992). Psychopharmacology of child and adolescent major depression: Present status and future directions. *Journal of Child and Adolescent Psychopharmacology, 2,* 31–45

Johnston, H. F., & Fruehling, J. J. (1994). Pharmacotherapy for depression in children and adolescents. In W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 365–397). New York: Plenum.

Joyce, P. R. (1984). Age of onset in bipolar affective disorder and misdiagnosis as schizophrenia. *Psychological Medicine, 14,* 145–149.

Judd, L. L., Rapoport, M. H., Paulus, M. P., & Brown, J. L. (1994). Subsyndromal symptomatic depression: A new mood disorder. *Journal of Clinical Psychiatry, 55*(Suppl 4), 18–28.

Kahn, J. S., Kehle, T. J., & Jenson, W. R. (1987, April). *Depression among middle schools students: Descriptive and correlational analyses.* Paper presented at the Annual Meeting, National Association of School Psychologists, New Orleans, LA.

Kahn, J. S., Kehle, T. J., Jenson, W. R., & Clark, E. (1990). Comparison of cognitive-behavioral, relaxation, and self-modeling interventions for depression among middle-school students. *School Psychology Review, 19,* 196–211.

Kaminer, Y. (1991). The magnitude of concurrent psychiatric disorders in hospitalized substance abusing adolescents. *Child Psychiatry and Human Development, 22,* 89–95.

Kaminer, Y., Tarter, R. E., Bukstein, O. G., & Kabene, M. (1992). Comparison between treatment completers and noncompleters among dually diagnosed substance-abusing adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 1046–1049.

Kanfer, F. H. (1970). Self-regulation: Research, issues, and speculations. In C. Neuringer & J. L. Michael (Eds.), *Behavior modification in clinical psychology* (pp. 178–220). New York: Appleton-Century-Crofts.

Karasu, T. B. (1990). Toward a clinical model of psychotherapy for depression, II: An integrative and selective treatment approach. *American Journal of Psychiatry, 147,* 269–278.

Kashani, J. H., & Cantwell, D. P. (1983). Etiology and treatment of childhood depression: A Biopsychological perspective. *Comprehensive Psychiatry, 24,* 476–486.

Kashani, J. H., Keller, M. B., Solomon, N., Reid, J. C., & Mazzola, D. (1985) Double depression in adolescent substance users. *Journal of Affective Disorders, 8,* 153–157.

Kashani, J. H., Sherman, D. D., Parker, D. R., & Reid, J. C. (1990). Utility of the Beck Depression Inventory with clinic-referred adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 278–282.

Kazdin, A. E. (1987). Children's Depression Scale: Validation with child psychiatric inpatients. *Journal of Child Psychology and Psychiatry, 28,* 29–41.

Kazdin, A. E. (1994). Informant variability in the assessment of childhood depression. In W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 249–271). New York: Plenum.

Keller, M. B. (1994). Depression: A long-term illness. *British Journal of Psychiatry, 165*(Suppl. 26), 9–15.

Keller, M. B., Lavori, P. W., Beardslee, W. R., Wunder, J., & Ryan, N. (1991). Depression in children and adolescents: New data on "undertreatment" and a literature review on the efficacy of available treatments. *Journal of Affective Disorders, 21,* 163–171.

Kendall, P. C., & Clarkin, J. F. (1992). Introduction to special section: Comorbidity and treatment implications. *Journal of Consulting and Clinical Psychology, 60,* 833–834.

Kendall, P. C., Kortlander, E., Chansky, T. E., & Brady, E. U. (1992). Comorbidity of anxiety and depression in youth: Treatment implications. *Journal of Consulting and Clinical Psychology, 60,* 869–880.

Kendler, K. S., Kessler, R. C., Neale, M. C., Heath, A. C., & Eaves, L. J. (1993). The prediction of major depression in women: Toward an integrated etiological model. *American Journal of Psychiatry, 150,* 1139–1148.

King, C. A., Ghaziuddin, N., McGovern, L., Brand, E., Hill, E., & Naylor, M. (1996). Predictors of comorbid alcohol and substance abuse in depressed adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 743–751.

King, C. A., Hovey, J. D., Brand, E., Wilson, R. & Ghaziuddin, N. (1997). Suicidal adolescents post-hospitalization: Parent and family impacts on treatment follow-through. *Journal of the American Academy of*

Child and Adolescent Psychiatry, 36, 85–93.

King, C. A., Katz, S. H., Ghaziuddin, N., Brand, E., Hill, E., & McGovern, L. (1997). Diagnosis and assessment of depression and suicidality using the NIMH Diagnostic Interview Schedule for Children (DISC-2.3). *Journal of Abnormal Child Psychology, 25,* 173–181.

King, C. A., Naylor, M. W., Hill, E. M., Shain, B. N., & Greden, J. F. (1993). Dysthymia characteristics of heavy alcohol use in depressed adolescents. *Biological Psychiatry, 33,* 210–212.

King, C. A., Naylor, M. W., Segal, H. G., Evans, T., & Shain, B. N. (1993). Global self-worth, specific self-perceptions of competence, and depression in adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 745–752.

King, C. A., Radpour, L., Naylor, M. W., Segal, H. G., & Jouriles, E. N. (1995). Parents' marital functioning and adolescent psychopathology. *Journal of Consulting and Clinical Psychology, 63,* 749–753.

King, C. A., Raskin, A., Gdowski, C. L., Butkus, M., & Opipari, L. (1990). Psychosocial factors associated with urban adolescent female suicide attempts. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 289–294.

King, C. A., Segal, H., Kaminski, K., Naylor, M. W., Ghaziuddin, N., & Radpour, L. (1995). A prospective study of adolescent suicidal behavior following hospitalization. *Suicide and Life-Threatening Behavior, 25,* 327–338.

King, C. A., Segal, H. G., Naylor, M., & Evans, T. (1993). Family functioning and suicidal behavior in adolescent inpatients with mood disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 1198–1206.

King, N. J., Gullone, E., & Ollendick, T. H. (1990). Childhood anxiety disorders and depression: Phenomenology, comorbidity, and intervention issues. *Scandinavian Journal of Behaviour Therapy, 19,* 59–70.

Klerman, G. L. (1987). Depression associated with medical and neurological diseases, drugs, and alcohol. In A. J. Marsella, R. M. A. Hirschfeld & M. M. Katz (Eds.), *The measurement of depression* (pp. 20–29). New York: Guilford.

Klerman, G. L., & Weissman, M. M. (1982). Interpersonal psychotherapy: Theory and research. In A. J. Rush (Ed.), *Short-term psychotherapies for depression: Behavioral, interpersonal, cognitive, and psychodynamic approaches* (pp. 88–106). New York: Guilford.

Komoto, J., Usui, S., & Hirata, J. (1984). Infantile autism and affective disorder. *Journal of Autism and Developmental Disorders, 14,* 81–84.

Kovacs, M. (1979). *Children's Depression Inventory*, University of Pittsburgh School of Medicine, PA: Author.

Kovacs, M. (1980/1981). Rating scales to assess depression in school-aged children. *Acta Paedopsychiatrica, 46*, 305–315.

Kovacs, M. (1983). *The Children's Depression Inventory: A self-rating scale for school-aged youngsters.* Unpublished manuscript.

Kovacs, M. (1985). The natural history and course of depressive disorders in childhood. *Psychiatric Annals, 15*, 387–389.

Kovacs, M. (1990). Comorbid anxiety disorders in childhood-onset depression. In J. D. Maser & C. R. Cloninger (Eds.), *Comorbidity of mood and anxiety disorders* (pp. 271–281). Washington DC: American Psychiatric Press.

Kovacs, M. (1992). *Children's Depression Inventory manual.* North Tonawanda, NY: Multi-Health Systems.

Kovacs, M. (1996a). Presentation and course of major depressive disorder during childhood and later years of the life span. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 705–715.

Kovacs, M. (1996b). The course of childhood-onset depressive disorders. *Psychiatric Annals, 26*, 326–330.

Kovacs, M., Akiskal, H. S., Gatsonis, C., & Perrone, P. L. (1994). Childhood-onset dysthymic disorder. Clinical features and prospective naturalistic outcome. *Archives of General Psychiatry, 51*, 365–374.

Kovacs, M., & Beck, A. T. (1977). An empirical-clinical approach toward a definition of childhood depression. In J. G. Schulterbrandt & A. Raskin (Eds.) *Depression in childhood: Diagnosis, treatment and conceptual models,* (pp. 1–25). New York: Raven Press.

Kovacs, M., Feinberg, T. L., Crouse-Novak, M., Paulauskas, S. L., Pollock, M., & Finkelstein, R. (1984a). Depressive disorders in childhood: II. A longitudinal study of the risk for a subsequent major depression. *Archives of General Psychiatry, 41,* 643–649.

Kovacs, M., Feinberg, T. L., Crouse-Novak, M., Paulauskas, S. L., & Finkelstein, R. (1984b). Depressive disorders in childhood: I. A longitudinal prospective study of characteristics and recovery. *Archives of General Psychiatry, 41,* 229–237.

Kovacs, M., & Gatsonis, C. (1994). Secular trends in age at onset of major depressive disorder in a clinical sample of children. *Journal of Psychiatric Research, 28,* 319–329.

Kovacs, M., Gatsonis, C., Paulauskas, S. L., & Richards, C. (1989). Depressive disorders in childhood, IV: A longitudinal study of comorbidity with and risk for anxiety disorders. *Archives of General Psychiatry, 46,* 776–782.

Kovacs, M., Paulauskas, S. L., Gatsonis, C., & Richards, C. (1988). Depressive disorders in childhood: III. A longitudinal study of comorbidity with and risk for conduct disorders. *Journal of Affective Disorders, 15, 205–217.*

Kuhn, V. & Kuhn, R. (1971). Drug therapy for depression in children. Indications and methods. In A. Annell (Ed.), *Depressive states in childhood and adolescence* (pp. 455–459). Stockholm, Sweden: Almqvist & Wiksell.

Kutcher, S., Boulos, C., Ward, B., Marton, P., Simeon, J., Ferguson, H. B., Szalai, J., Katic, M., Roberts, N., Dubois, C., & Reed, K. (1994). Response to desipramine treatment in adolescent depression: A fixed-dose, placebo-controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 686–694.

Kutcher, S. P., & Marton, P. (1989). Utility of the Beck Depression Inventory with psychiatrically disturbed adolescent outpatients. *Canadian Journal of Psychiatry, 34,* 107–109.

Kutcher, S. P., & Marton, P., & Korenblum, M. (1989). Relationship between psychiatric illness and conduct disorder in adolescents. *Canadian Journal Psychiatry, 34,* 526–529.

Kye, C. H., Waterman, G. S., Ryan, N. D., Birmaher, B., Williamson, D. E., Iyengar, S., & Dachille, S. (1996). A randomized, controlled trial of amitriptyline in the acute treatment of adolescent major depression. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 1139–1144.

Lam, R. W., Solyom, L., & Tompkins, A. (1991). Seasonal mood symptoms in bulimia nervosa and seasonal affective disorder. *Comprehensive Psychiatry, 32,* 552–558.

Lang, M., & Tisher, M. (1987). *Children's Depression Scale (CDS) (9–16 years) manual* (North American Edition). Palo Alto, CA: Consulting Psychologists Press.

Ledoux, S., Choquet, M., & Manfredi, R. (1993). Associated factors for self-reported binge eating among male and female adolescents. *Journal of Adolescence, 16,* 75–91.

Leibenluft, E., Fiero, P. L., & Rubinow, D. R. (1994). Effects of the menstrual cycle on dependent variables in mood disorder research. *Archives of General Psychiatry, 51,* 761–781.

Lesse, S. (1981). Hypochondriacal and psychosomatic disorders masking depression in adolescents. *American Journal of Psychotherapy, 35,* 356–367.

Levy, J. C., & Deykin, E. Y. (1989). Suicidality, depression, and substance abuse in adolescence. *American Journal of Psychiatry, 146,* 1462–1467.

Lewinsohn, P. M. (1975). The behavioral study and treatment of depression. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification* (pp. 19–64). New York: Academic Press.

Lewinsohn, P. M., Clarke, G. N., Hops, H., & Andrews, J. (1990). Cognitive-behavioral treatment for depressed adolescents. *Behavior Therapy, 21,* 385–401.

Lewinsohn, P. M., Hops, H., Roberts, R. E., Seeley, J. R., & Andrews, J. (1993). Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. *Journal of Abnormal Psychology, 102,* 133–144.

Lewinsohn, P. M., Munoz, R. F., Youngren, M. A., & Zeiss, A. M. (1978). *Control your depression.* Englewood Cliffs, NJ: Prentice-Hall.

Lewinsohn, P. M., Rohde, P., & Seeley, J. R. (1995). Adolescent psychopathology: III. The clinical consequences of comorbidity. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 510–519.

Lewinsohn, P. M., Rohde, P., Seeley, J. R., & Fisher, S. A. (1993). Age-cohort changes in the lifetime occurrence of depression and other mental disorders. *Journal of Abnormal Psychology, 102,* 110–120.

Lewis, M., & Lewis, D. O. (1979). A psychobiological view of depression in childhood. In A. French & I. Berlin (Eds.), *Depression in children and adolescents* (pp. 29–45). New York: Human Sciences Press.

Licitra-Kleckler, D. M., & Waas, G. A. (1993). Perceived social support among high-stress adolescents: The role of peers and family. *Journal of Adolescent Research, 8,* 381–402.

Liddle, B., & Spence, S. H. (1990). Cognitive-behaviour therapy with depressed primary school children: A cautionary note. *Behavioural Psychotherapy, 19,* 85–102.

Ling, W., Oftedal, G., & Weinberg, W. (1970). Depressive illness in childhood presenting as severe headache. *American Journal of Diseases in Children, 120,* 122–124.

Loeber, R., Russo, M. F., Stouthamer-Loeber, M., & Lahey, B. B. (1994). Internalizing problems and their relation to the development of disruptive behaviors in adolescence. *Journal of Research on Adolescence, 4,* 615–637.

Loranger, A. W., & Levine, P. M (1978). Age at onset of bipolar affective illness. *Archives of General Psychiatry, 35,* 1345–1348.

Magnusson, D. (1985). Implications of an interactional paradigm for research on human development. *International Journal of Behavioral Development, 8,* 115–137.

Magnusson, D., & Torestad, B. (1993). A holistic view of personality: A model revisited. *Annual Review of Psychology, 44,* 427–452.

Marcotte, D., & Baron, p. (1993). L'efficacité d'une stratégie d'intervention émotivo-rationnelle auprès, d'adolescents dépressifs de milieu scolaire. *Revue Canadienne de Counseling, 27,* 77–92.

McClellan, J. M., Rubert, M. P, Reichler, R. J., & Sylvester, C. E. (1990). Attention deficit disorder in children at risk for anxiety and depression. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 534–539.

McGlashan, T. H. (1988). Adolescent versus adult onset of mania. *American Journal of Psychiatry, 145,* 221–223.

McKay, J. R., Murphy, R. T., Maisto, S. A., & Rivinus, T. R. (1992). Characteristics of adolescent psychiatric patients who engage in problematic behavior while intoxicated. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 1031–1035.

Meichenbaum, D. (1985). *Stress inoculation training.* New York: Pergamon.

Meller, W. H., & Borchardt, C. M. (1996). Comorbidity of major depression and conduct disorder. *Journal of Affective Disorders, 39,* 123–126.

Meyer, A. (1957). *Psychobiology.* Springfield, IL: Charles Thomas.

Mitchell, J., McCauley, E., Burke, P., & Moss, S. J. (1988). Phenomenology of depression in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 12–20.

Mokros, H. B., Poznanski, E., Grossman, J. A., & Freeman, L. N. (1987). A comparison of child and parent ratings of depression for normal and clinically referred children. *Journal of Child Psychology and Psychiatry, 28,* 613–627.

Moreau, D., Mufson, L., Weissman, M. M., & Klerman, G. L. (1991). Interpersonal psychotherapy for adolescent depression: Description of modification and preliminary application. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 642–651.

Moretti, M. M., Fine, S., Haley, G., & Marriage, K. (1985). Childhood and adolescent depression: Child-report versus parent-report information. *Journal of the American Academy of Child Psychiatry, 24,* 298–302.

Mufson, L., Moreau, D., Weissman, M. M., Wickramaratne, P., Martin, J., & Samoilov, A. (1994). The modification of interpersonal psychotherapy with depressed adolescents (IPT-A): Phase I and Phase II studies. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 695–705.

Mufson, L., Moreau, D., Weissman, M. M., & Klerman, G. L. (1993). *Interpersonal psychotherapy for depressed adolescents.* New York: Guilford Press.

Mufson, L., Weissman, M. M., & Warner, V. (1992). Depression and anxiety in parents and children: A direct interview study. *Journal of Anxiety Disorders, 6,* 1–13.

Naylor, M. W., & Grossman, M (1991). Trichotillomania and depression (letter to the Editor). *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 155–156.

Naylor, M. W., Staskowski, M., Kenney, M. C., & King, C. A. (1994). Language disorders and learning disabilities in school-refusing adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 1331–1337.

Nelson, W. M., & Politano, P. M. (1990). Children's Depression Inventory: Stability over repeated administrations in psychiatric inpatient children. *Journal of Clinical Child Psychology, 19,* 254–256.

Nieminen, G. S., & Matson, J. L. (1989). Depressive problems in conduct-disordered adolescents. *Journal of School Psychology, 27,* 175–186.

Nurcombe, B. (1994). The validity of the diagnosis of major depression in childhood and adolescence. In W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 61–77). New York: Plenum.

Nurcombe, B., Seifer, R., Scioli, A., Tramontana, M. G., Grapentine, W. L., & Beauchesne, H. C. (1989). Is major depressive disorder in adolescence a distinct diagnostic entity? *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 333–342.

Oakley-Browne, M. A., Joyce, P. R., Wells, J. E., Bushnell, J. A., & Hornblow, A. R. (1995). Adverse parenting and other childhood experience as risk factors for depression in women aged 18–44 years. *Journal of Affective Disorders, 34,* 13–23.

Offer, D., & Schonert-Reichl, K. A. (1992). Debunking the myths of adolescence: Findings from recent research. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 1003–1014.

Ollendick, T. H., Yule, W., & Ollier, K. (1991). Fears in British children and their relationship to manifest anxiety and depression. *Journal of Child Psychology and Psychiatry, 32*, 321–331.

Papatheodorou, G., & Kutcher, S. (1995). The effect of adjunctive light therapy on ameliorating breakthrough depressive symptoms in adolescent-onset bipolar disorder. *Journal of Psychiatry and Neuroscience, 20*, 226–232.

Paris, J. (1994). The etiology of borderline personality disorder: A biopsychosocial approach. *Psychiatry, 57*, 316–325.

Paul, S. M. (1988). Anxiety and depression: A common neurobiological substrate? *Journal of Clinical Psychiatry, 49*(Suppl. 10), 13–16.

Petti, T. A. (1978). Depression in hospitalized child psychiatry patients: Approaches to measuring depression. *Journal of the American Academy of Child Psychiatry, 17*, 49–59.

Pfeffer, C. R., & Plutchik, R. (1989). Co-occurrence of psychiatric disorders in child psychiatric patients and nonpatients: A circumplex model. *Comprehensive Psychiatry, 30*, 275–282.

Poznanski, E. O. (1982). The clinical phenomenology of childhood depression. *American Journal of Orthopsychiatry, 52*, 308–313.

Poznanski, E. O., Carroll, B. J., Banegas, M. C., Cook, S. C., & Grossman, J. A. (1982). The dexamethasone suppression test in prepubertal depressed children. *American Journal of Psychiatry, 139*, 321–324.

Poznanski, E. O., Cook, S. C., & Carroll, B. J. (1979). A depression rating scale for children. *Pediatrics, 64*, 442–450.

Poznanski, E. O., Cook, S. C., Carroll, B. J., & Corzo, H. (1983). Use of the Children's Depression Rating Scale in an inpatient psychiatric population. *Journal of Clinical Psychiatry, 44*, 200–203.

Poznanski, E. O., Freeman, L. N., & Mokros, H. B. (1985). Children's Depression Rating Scale–Revised. *Psychopharmacology Bulletin, 21*, 979–989.

Poznanski, E. O., Grossman, J. A., Buchsbaum, Y., Banegas, M., Freeman, L., & Gibbons, R. (1984). Preliminary studies of the reliability and validity of the Children's Depression Rating Scale. *Journal of the American Academy of Child Psychiatry, 23*, 191–197.

Poznanski, E. O., Krahenbuhl, V., & Zrull, J. P. (1976). Childhood depression: A longitudinal perspective. *Journal of the American Academy of Child Psychiatry, 15*, 491–501.

Poznanski, E. O., & Mokros, H. B. (1994). Phenomenology and epidemiology of mood disorders in children and adolescents. In W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 19–39). New York: Plenum.

Poznanski, E., Mokros, H. B., Grossman, J., & Freeman, L. N. (1985). Diagnostic criteria in childhood depression. *American Journal of Psychiatry, 142*, 1168–1173.

Poznanski, E., & Zrull, J. P. (1970). Childhood depression: Clinical characteristics of overtly depressed children. *Archives of General Psychiatry, 23*, 8–15.

Puig-Antich, J. (1982). Major depression and conduct disorder in prepuberty. *Journal of the American Academy of Child Psychiatry, 21*, 118–128.

Puig-Antich, J., & Chambers, W. (1978). *The Schedule for Affective Disorders and Schizophrenia for School-age Children (Kiddie-SADS)*. New York: New York State Psychiatric Institute.

Puig-Antich, J., Dahl, R., Ryan, N., Novacenko, H., Goetz, D., Goetz, R. M., Twomey, J., & Klepper, T. (1989). Cortisol secretion in prepubertal children with major depressive disorder: Episode and recovery. *Archives of General Psychiatry, 46*, 801–809.

Puig-Antich, J., Lukens, E., Davies, M., Goetz, D., Brennan-Quattrock, J., & Todak, G. (1985). Psychosocial functioning in prepubertal children with major depressive disorders, I. Interpersonal relationships during the depressive episode. *Archives of General Psychiatry, 42*, 500–507.

Radloff, L. S. (1977) The CES-D Scale: A self-report scale for research in the general population. *Applied Psychological Measurement, 1*, 385–401.

Rastam, M. (1992). Anorexia nervosa in 51 Swedish adolescents: Premorbid problems and comorbidity. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*, 819–829.

Rawson, H. E. (1992). The interrelationship of measures of manifest anxiety, self-esteem, locus of control, and depression in children with behavior problems. *Journal of Psychoeducational Assessment, 10*, 319–329.

Rawson, H. E., & Tabb, B. A. (1993). The effects of therapeutic intervention on childhood depression. *Child and Adolescent Social Work Journal, 10,* 39–52.

Regier, D. A., Kaelber, C. T., Roper, M. T., Rae, D. S., & Sartorius, N. (1994). The ICD-10 clinical field trial for mental and behavioral disorders: Results in Canada and the United States. *American Journal of Psychiatry, 151*, 1340–1350.

Rehm, L. P. (1977). A self-control model of depression. *Behavior Therapy, 8*, 787–804.

Reich, W., & Welner, Z. (1988). *Revised version of the Diagnostic Interview for Children and Adolescents (DICA-R)*. St. Louis, MO: Washington University School of Medicine.

Reinecke, M. A., & Schultz, T. M. (1995). Comparison of self-report and clinician ratings of depression among outpatient adolescents. *Depression, 3*, 139–145.

Remschmidt, H. (1992). The interaction of biological and psychosocial influences in developmental psychopathology. In H. Remschmidt & M. H. Schmidt (Eds.), *Developmental psychopathology* (pp. 17–25). Lewiston, NY: Hogrefe & Hubner.

Reynolds, W. M. (1985). Depression in childhood and adolescence: Diagnosis, assessment, intervention strategies and research. In T. R. Kratochwill (Ed.), *Advances in school psychology*. (Vol., 4. pp. 133–189). Hillsdale, NJ Lawrence Erlbaum.

Reynolds, W. M. (1986a). A model for the screening and identification of depressed children and adolescents in school settings. *Professional School Psychology, 1*, 117–129.

Reynolds, W. M. (1986b). *Reynolds Adolescent Depression Scale*. Odessa, FL: Psychological Assessment Resources.

Reynolds, W. M. (1987) *Reynolds Adolescent Depression Scale: Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Reynolds, W. M. (1989) *Reynolds Child Depression Scale: Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Reynolds, W. M. (1990). Development of a semi-structured clinical interview for suicidal behaviors in adolescents. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2*, 382–390.

Reynolds, W. M. (1992a). The study of internalizing disorders in children and adolescents. In W. M. Reynolds (Ed.), *Internalizing disorders in children and adolescents* (pp. 1–18). New York: Wiley.

Reynolds, W. M. (1992b). Depression in children and adolescents. In W. M. Reynolds (Ed.), *Internalizing disorders in children and adolescents* (pp. 149–253). New York: Wiley.

Reynolds, W. M. (1994). Assessment of depression in children and adolescents by self-report questionnaires. In W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 209–234). New York: Plenum.

Reynolds, W. M. (1997a). *A biopsychosocial model of* *depression in children and adolescents*. Unpublished manuscript.

Reynolds, W. M. (1997b). *Comorbidity of psychiatric disorders at intake among adolescents seen in inpatient and outpatient clinical settings*. Unpublished manuscript.

Reynolds, W. M., Anderson, G. & Bartell, N. (1985). Measuring depression in children: A multimethod assessment investigation. *Journal of Abnormal Child Psychology, 13*, 513–526.

Reynolds, W. M. & Coats, K. I. (1982, July). *Depression in adolescents: Incidence, depth and correlates*. Paper presented at the 10th International Congress of the International Association for Child and Adolescent Psychiatry, Dublin, Ireland.

Reynolds, W. M. & Coats, K. I. (1986). A comparison of cognitive-behavioral therapy and relaxation training for the treatment of depression in adolescents. *Journal of Consulting and Clinical Psychology, 54*, 653–660

Reynolds, W. M. & Graves, A. (1989). Reliability of children's reports of depressive symptomatology. *Journal of Abnormal Child Psychology, 17*, 647–655.

Reynolds, W. M., & Johnston, H. F. (1994). The nature and study of depression in children and adolescents. In W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 3–17). New York: Plenum.

Reynolds, W. M. & Mazza, J. J. (1994). Suicide and suicidal behaviors in children and adolescents. In W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents* (pp. 525–580). New York: Plenum.

Reynolds, W. M., & Mazza, J. J. (in press). Reliability and validity of the Reynolds Adolescent Depression Scale with young adolescents. *Journal of School Psychology*.

Reynolds, W. M., & Miller, K. L. (1989). Assessment of adolescents' learned helplessness in achievement situations. *Journal of Personality Assessment, 53*, 211–228.

Riddle, M. A., Brown, N., Dzubinski, D, Jetmalani, A. J., Law, Y. & Woolston, J. L. (1989). Fluoxetine overdose in an adolescent. *Journal of the American Academy of Child and Adolescent Psychiatry, 28*, 587–588.

Riddle, M. A., Hardin, M. T., King, R., Scahill, L., & Woolston, J. L. (1990). Fluoxetine treatment of children and adolescents with Tourette's and obsessive compulsive disorders: Preliminary clinical experience. *Journal of the American Academy of Child and Adolescent Psychiatry, 29*, 45–48.

Robbins, D. R., Alessi, N. E., Colfer, M. V. (1989). Treatment of adolescents with major depression: Implications of the DST and the melancholic clinical subtype. *Journal of Affective Disorders, 17*, 99–104.

Robbins, D. R., Alessi, N. E., Colfer, M. V., & Yanchyshyn, G. W. (1985). Use of the Hamilton Rating Scale for Depression and the Carroll Self-Rating Scale in adolescents. *Psychiatry Research, 14*, 123–129.

Roberts, R. E., Lewinsohn, P. M., & Seeley, J. R. (1991). Screening for adolescent depression: A comparison of depression scales. *Journal of the American Academy of Child and Adolescent Psychiatry, 30*, 58–66.

Rohde, P., Lewinsohn, P. M., & Seeley, J. R. (1991). Comorbidity of unipolar depression: II. Comorbidity with other mental disorders in adolescents and adults. *Journal of Abnormal Psychology, 100*, 214–222.

Roseby, V. & Deutsch, R. (1985). Children of separation and divorce: Effects of a social role-taking group intervention on fourth and fifth graders. *Journal of Clinical Child Psychology, 14*, 55–60.

Rush, A. J. (1990). Problems associated with the diagnosis of depression. *Journal of Clinical Psychiatry, 51*(Suppl.), 15–22.

Rush, A. J., Giles, D. E., Schlesser, M. A., Fulton., C. L., Weissenberger, J., & Burns, C. (1986). The Inventory of Depressive Symptoms (IDC): Preliminary findings. *Psychiatry Research, 18*, 65–87.

Sadowski, C., & Kelley, M. L. (1993). Social problem solving in suicidal adolescents. *Journal of Consulting and Clinical Psychology, 61*, 121–127.

Sandler, J., & Jaffee, N. G. (1965). Notes on childhood depression. *International Journal of Psychoanalysis, 46*, 88–96.

Sartorius, N., Kaelber, C. T., Cooper, J. E., Roper, M. T., Rae, D. S., Gulbinat, W., Ustun, T. B., & Regier, D. A. (1993). Progress toward achieving a common language in psychiatry: Results from the field trial of the clinical guidelines accompanying the WHO classification of mental and behavioral disorders in ICD-10. *Archives of General Psychiatry, 50*, 115–124.

Saylor, C. F., Finch, A J., Spirito, A., & Bennett, B. (1984). The Children's Depression Inventory: A systematic evaluation of psychometric properties. *Journal of Consulting and Clinical Psychology, 52*, 955–967.

Schonert-Reichl, K. A. (1994). Gender differences in depressive symptomatology and egocentrism in adolescence. *Journal of Early Adolescence, 14*, 49–64.

Shaffer, D., Schwab-Stone, M., Fisher, P., Cohen, P., Piacentini, J., Davies, M., Conners, C. K., & Reiger, D. (1993). The Diagnostic Interview Schedule for Children-Revised Version (DISC-R): I. Preparation, field testing, interrater reliability, and acceptability. *Journal of the American Academy of Child and Adolescent Psychiatry, 32*, 643–650.

Shain, B. N., Kronfol, Z., Naylor, M., Goel, K., Evans, T., & Schaefer, S. (1991). Natural killer cell activity in adolescents with major depression. *Biological Psychiatry, 29*, 481–484.

Shain, B. N., Naylor, M., & Alessi, N. (1990). Comparison of self-rated and clinician-rated measures of depression in adolescents. *American Journal of Psychiatry, 147*, 793–795.

Shalev, A. Y. (1993). Post-traumatic stress disorder: A biopsychosocial perspective. *Israeli Journal of Psychiatry and Related Sciences, 30*, 102–109.

Smith, C., Feldman, S., Nasserbakht, A., & Steiner, H. (1993). Psychological characteristics and *DSM-III-R* diagnosis at 6-year follow-up of adolescent anorexia nervosa. *Journal of the American Academy of Child and Adolescent Psychiatry, 32*, 1237–1245.

Smith, C., & Steiner, H. (1992). Psychopathology in anorexia nervosa and depression. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*, 841–843.

Smoller, A. S., & Brosgole, L. (1993). Visual and auditory affect recognition in depressed and nondepressed latency aged children. *International Journal of Neuroscience, 70*, 29–38.

Smucker, M. R., Craighead, W. E., Craighead, L. W., & Green, B. J. (1986). Normative and reliability data for the Children's Depression Inventory. *Journal of Abnormal Child Psychology, 14*, 25–39.

Sonis, W. A., Yellin, A. M., Garfinkel, B. D., & Hoberman, H. H. (1987). The antidepressant effect of light in seasonal affective disorder of childhood and adolescence. *Psychopharmacology Bulletin, 23*, 360–363.

Spitzer, R. L., Endicott, J., & Robins, E. (1978). Research diagnostic criteria: Rationale and reliability. *Archives of General Psychiatry, 35*, 773–782.

Stack, J. J. (1971). Chemotherapy in childhood depression. In A. Annell (Ed.), *Depressive states in childhood and adolescence* (pp. 460–466). Stockholm, Sweden: Almqvist & Wiksell.

Stark, K. D. (1990). *Childhood depression: School-based intervention*. New York: Guilford Press.

Stark, K. D., Reynolds, W. M., & Kaslow, N. J. (1987). A comparison of the relative efficacy of self-control therapy and behavioral problem-solving therapy for depression in children. *Journal of Abnormal Child Psychology, 15*, 91–113.

Stark, K. D., Rouse, L. W., & Livingston, R. (1991). Treatment of depression during childhood and adolescence: Cognitive behavioral procedures for the individual and family. In P. C. Kendall (Ed.), *Child and adolescent therapy: Cognitive-behavioral procedures* (pp. 165–206). New York: Guilford Press.

Stattin, H., & Magnusson, D. (1989). Social transition in adolescence: A biosocial perspective. In A. de Ribaupierre (Ed.), *Transition mechanisms in child development: The longitudinal perspective* (pp. 147–190). Cambridge, UK: Cambridge University Press.

Stone, M. H. (1981). Depression in borderline adolescents. *American Journal of Psychotherapy, 35*, 383–399.

Stowell, R. J. A., & Estroff, T. W. (1992). Psychiatric disorders in substance-abusing adolescent inpatients: A pilot study. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*, 1036–1040.

Straus, M. A., & Kantor, G. K. (1994). Corporal punishment of adolescents by parents: A risk factor in the epidemiology of depression, suicide, alcohol abuse, child abuse, and wife beating. *Adolescence, 29*, 534–561.

Strauss, C. C., Last, C. G., Hersen, M., & Kazdin, A. E. (1988). Association between anxiety and depression in children and adolescents with anxiety disorders. *Journal of Abnormal Child Psychology, 16*, 57–68.

Strober, M., Green, J., & Carlson, G. (1981). Utility of the Beck Depression Inventory with psychiatrically hospitalized adolescents. *Journal of Consulting and Clinical Psychology, 49*, 482–483.

Suinn, R. M. (1986). *Anxiety management training manual.* Fort Collins, CO: Rocky Mountain Behavioral Science Institute.

Swift, W. J., Andrews, D., & Barklage, N. E. (1986). The relationship between affective disorders and eating disorders: A review of the literature. *American Journal of Psychiatry, 143,* 290–299.

Syvalahti, E. (1987). Endocrine and immune adaptation in stress. *Annals of Clinical Research, 19,* 70–77.

Tarnowski, K. J., Simonian, S. J., Bekeny, P., & Park, A. (1992). Acceptability of interventions for childhood depression. *Behavior Modification, 16,* 103–117.

Tisher, M. (1995). Teacher's assessments of prepubertal childhood depression. *Australian Journal of Psychology, 47,* 93–96.

Tisher, M., & Lang, M. (1983). The Children's Depression Scale: Review and further developments. In D. P. Cantwell & G. A. Carlson (Eds.), *Affective disorders in childhood and adolescents—An update* (pp. 181–203). Jamaica, NY: Spectrum Publications.

Tisher, M., Lang-Takac, E., & Lang, M. (1992). The Children's Depression Scale: Review of Australian and overseas experience. *Australian Journal of Psychology, 44,* 27–35.

Tonkin, G., & Hudson, A. (1981). The Children's Depression Scale: Some further psychometric data. *Australian Council for Educational Research Bulletin for Psychologists, 30,* 11–18

Vasile, R. G., Samson, J. A., Bemporad, J., Bloomingdale, K. L., Creasey, D., Fenton, B. T., Gudeman, J. E., & Schildkraut, J. J. (1987). A biopsychosocial approach to treating patients with affective disorders. *American Journal of Psychiatry, 144,* 341–344.

Vitiello, B., Malone, R., Buschle, P. R., Delaney, M. A., & Behar, D. (1990). Reliability of DSM-III diagnoses of hospitalized children. *Hospital and Community Psychiatry, 41,* 63–67.

Waldron, S., Shrier, D. K., Stone, B., & Tobin, F. (1975). School phobia and other childhood neuroses: A systematic study of the children and their families. *American Journal of Psychiatry, 132,* 802–808.

Waller, D. A., & Rush, A. J. (1983). Differentiating primary affective disease, organic affective syndromes, and situational depression on a pediatric service. *Journal of the American Academy of Child Psychiatry, 22,* 52–58.

Walters, E. E., Neale, M. C., Eaves, L. J., Heath, A. C., Kessler, R. C., & Kendler, K. S. (1992). Bulimia nervosa and major depression: A study of common genetic and environmental factors. *Psychological Medicine, 22,* 617–622.

Wamboldt, M. Z. (1994). Current status of child and adolescent medical-psychiatric units. *Psychosomatics, 35,* 434–444.

Weinberg, J., & van den Dungen, M. (1971). Depression and delinquency in adolescence. In A. Annell (Eds.), *Depressive states in childhood and adolescence* (pp. 296–302). Stockholm, Sweden: Almqvist & Wiksell.

Weinberg, W. A., & Brumback, R. A. (1976). Mania in childhood: Case studies and literature review. *American Journal of Diseases in Children, 130,* 380–385.

Weinberg, W. A., & Emslie, G. J. (1991). Attention deficit hyperactivity disorder: The differential diagnosis. *Journal of Child Neurology, 6*(Suppl.), S23–S36.

Weinberg, W. A., Rutman, J., Sullivan, L., Penick, E. C., & Dietz, S. G. (1973). Depression in children referred to an educational diagnostic center: Diagnosis and treatment. *Journal of Pediatrics, 83,* 1065–1072.

Weiss, B., Weisz, J. R., Politano, M., Carey, M., Nelson, W. M., & Finch, A. J. (1991). Developmental differences in the factor structure of the Children's Depression Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 38–45.

Weissman, M. M. (1990). Evidence for comorbidity of anxiety and depression: Family and genetic studies of children. In J. D. Maser & C. R. Cloninger (Eds.), *Comorbidity of mood and anxiety disorders* (pp. 349–365). Washington, DC: American Psychiatric Press.

Weist, M. D., Proescher, E. L, Freedman, A. H., Paskewitz, D. A., & Flaherty, L. T. (1995). School-based health services for urban adolescents: Psychosocial characteristics of clinic users or nonusers. *Journal of Youth and Adolescence, 24,* 251–265.

Weller, E. B., & Weller, R. A., (1983). Case report of conversion symptom associated with major depressive disorder in a child. *American Journal of Psychiatry, 140,* 1079–1080.

Weller, R. A., & Weller, E. B. (1986). Tricyclic antidepressants in prepubertal depressed children: Review of the literature. *Hillside Journal of Clinical Psychiatry, 8,* 46–55.

Weller, E. B., Weller, R. A., & Carr, S. (1989). Imipramine treatment of trichotillomania and coexisting depression in a seven-year-old. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 952–953.

Weller, R. A., Weller, E. B., Fristad, M. A., Cantwell, M. L., & Preskorn, S. H. (1986). Dexamethasone suppression test and clinical outcome in prepubertal depressed children. *American Journal of Psychiatry, 143,* 1469–1470.

Weller, R. A., Weller, E. B., Tucker, S. G., & Fristad, M. A. (1986). Mania in prepubertal children: Has it been underdiagnosed? *Journal of Affective Disorders, 11,* 151–154.

Wenning, K., Nathan, P., & King, S. (1993). Mood disorders in children with oppositional defiant disorder: A pilot study. *American Journal of Orthopsychiatry, 63,* 295–299.

West, S. A., McElroy, S. L., Strakowski, S. M., Keck, P. E., & McConville, B. J. (1995). Attention deficit hyperactivity disorder in adolescent mania. *American Journal of Psychiatry, 152,* 271–273.

Whitaker, A., Johnson, J., Shaffer, D., Rapoport, J. L., Kalikow, K., Walsh, B. T., Davies, M., Braiman, S., & Dolinsky, A. (1990). Uncommon troubles in young people: Prevalence estimates of selected psychiatric disorders in a nonreferred adolescent population. *Archives of General Psychiatry, 47*, 487–496.

Wilkes, T. C. R., Belsher, G., Rush, A. J., & Frank, E. (1994). *Cognitive therapy for depressed adolescents*. New York: Guilford Press.

World Health Organization (1992). *International statistical classification of diseases and related health problems* (10th Rev.). Geneva, Switzerland: Author.

Wright, J. H., & Thase, M. E. (1992). Cognitive and biological therapies: A synthesis. *Psychiatric Annals, 22*, 451–458.

Wurzbacher, K. V., Evans, E. D., & Moore, E. J. (1991). Effects of alternative street school on youth involved in prostitution. *Journal of Adolescent Health, 12*, 549–554.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.18
Suicide and Suicide Attempts During Adolescence

ANTHONY SPIRITO and DEIDRE DONALDSON
Rhode Island Hospital and Brown University School of Medicine, Providence, RI, USA

5.18.1 INTRODUCTION 464
5.18.2 PHENOMENOLOGY 464
5.18.3 DIAGNOSTIC FEATURES 464
  5.18.3.1 Ideation 464
  5.18.3.2 Method of Attempt 465
  5.18.3.3 Precipitant of the Attempt 465
  5.18.3.4 Intent 465
5.18.4 PREVALENCE 465
5.18.5 ETIOLOGY: RISK FACTORS 466
  5.18.5.1 Gender 466
  5.18.5.2 Race 467
  5.18.5.3 Psychiatric Disorders 467
  5.18.5.4 Prior Suicide Attempts 468
  5.18.5.5 Biological Factors 468
  5.18.5.6 Method of Suicide: Firearms 469
  5.18.5.7 Exposure 469
5.18.6 DEVELOPMENTAL COURSE 470
5.18.7 CONCEPTUALIZATION AND CLINICAL FORMULATION OF ATTEMPTED SUICIDE 471
  5.18.7.1 Multimodal Assessment 471
  5.18.7.2 Attempt Characteristics 473
  5.18.7.3 Emotional State 473
  5.18.7.4 Cognitive Variables 473
  5.18.7.5 Maladaptive Coping Strategies 474
  5.18.7.6 Physical and Sexual Abuse 475
  5.18.7.7 Family Functioning 475
5.18.8 PSYCHOSOCIAL TREATMENTS 476
  5.18.8.1 Suicide Ideators 476
  5.18.8.2 Suicide Attempters 476
  5.18.8.3 Secondary Prevention of Suicidal Behavior in High-risk Groups 478
  5.18.8.4 Primary Prevention of Suicidal Behavior 479
5.18.9 PSYCHOPHARMACOLOGY 479
5.18.10 EFFICACY/EFFECTIVENESS OF TREATMENTS 479
5.18.11 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 480
5.18.12 SUMMARY 481
5.18.13 REFERENCES 481

5.18.1 INTRODUCTION

Although suicide is rare in childhood, the rates of adolescent (15–19 year-olds) suicide almost tripled from 1950 to 1980. From 1980 to 1992, the rate increased by another 28.3%, from 8.5 to 10.9 per 100 000 (Centers for Disease Control [CDC], 1995a). The sharp rise in adolescent suicide is of significant concern. Beyond the emotional toll on the survivors and the tragic loss of a life so young, there are also considerable economic costs related to suicide. It has been estimated that each youth suicide results in approximately $432 000 worth of lost economic productivity (Weinstein & Saturno, 1989). Suicide attempts that do not result in death are also a significant public health problem. They occur at a very high rate (7–10% of adolescents report that they have attempted suicide) and a significant number (approximately 2–3%) of suicide attempters receive costly medical care following their attempts (CDC, 1995a).

This chapter summarizes many of the factors that contribute to adolescent suicide and suicide attempts. Adolescent suicide completers and attempters form separate, but overlapping groups. Thus, it is necessary to describe both. The first portion of this chapter discusses factors associated with completed and attempted suicide. The second portion of the chapter describes the clinical evaluation of suicide attempters. A bio-psychosocial model is utilized to synthesize the multiple factors that may affect outcome among suicide attempters. Understanding these factors also aids in the development of comprehensive and effective treatment plans for this high-risk group of adolescents. The final section discusses treatments for adolescents with suicidal ideation and adolescents who have attempted suicide.

5.18.2 PHENOMENOLOGY

Suicidal behavior includes both thoughts and actions. For some adolescents, a suicide attempt is the result of an impulsive act in the face of significant stress. For others, suicide attempts follow a course from passive suicidal ideation, to serious consideration of suicide as an option, followed by active planning. Suicidal thoughts may vary in frequency, intensity, and duration. The time course prior to a suicide attempt can be quite variable, ranging from many months of suicidal ideation to only several days or even hours. A suicide attempt may hold a particular meaning for either an individual or family; for example, it may be an indirect way of communicating distress or a way to receive attention within a family. Completed suicides may occur in adolescents who did not actually want to die but rather made an impulsive attempt using lethal means. Thus, the psychological characteristics associated with adolescent suicidality vary significantly, and no one pattern of psychopathology adequately describes adolescents who experience suicidal ideation, make a suicide attempt, or complete suicide.

5.18.3 DIAGNOSTIC FEATURES

Suicide and attempted suicide are not psychiatric disorders, but rather, symptoms of emotional distress which may be associated with certain psychiatric disorders such as depression. Consequently, there are no true diagnostic criteria. However, there is a great deal of confusion in the literature regarding what constitutes a "true" suicide attempt, and even suicidal ideation. What degree of ideation is necessary before a clinician or researcher considers an adolescent a suicide ideator? Similarly, does there have to be a certain level of suicidal intent or sufficient medical lethality for a suicide attempt to be considered a "true" suicide attempt? A variety of terms have been used to describe suicidal behavior. Standard nomenclature for suicidal behavior does not currently exist, which makes it difficult to interpret empirical findings. Reference to terms such as self-injury, self-mutilation, parasuicide, and suicide threats is common, and the term "suicidal gesture" is used in clinical notes and sometimes in research reports. The latter term typically refers to a suicide attempt of low medical lethality. This chapter uses the terms "completed suicide," "attempted suicide" and "suicidal ideation."

Several characteristics of a suicide attempt are pertinent to determining its severity. These include the intensity of suicidal ideation prior to the attempt, method of attempt, the precipitant of the attempt, and suicidal intent accompanying the attempt.

5.18.3.1 Ideation

The degree of suicidal ideation accompanying an attempt may be a marker of the seriousness of a suicide attempt. Spirito, Overholser, and Vinnick (1995) found that more than half of a sample of 203 adolescent suicide attempters demonstrated clinically significant levels of ideation on the Suicide Ideation Questionnaires (Reynolds, 1987): 13.8% of the sample scored in the 99th percentile and 43.3% scored in the 95th percentile or above. However, a significant number of adolescents did not report very high levels of suicide ideation, suggesting either an impulsive attempt or denial of their suicidal state.

5.18.3.2 Method of Attempt

The method of attempting suicide is not typically a useful indication of severity because the large majority of adolescents attempt suicide by drug overdose (Spirito, Brown, Overholser, & Fritz, 1989). Compared to adults, adolescents have limited access to other more lethal means of suicide which may account for the overrepresentation of overdoses. Because of the availability of medical care in most Western nations, overdoses are rarely fatal. The potential for a medically serious suicide attempt increases when prescription drugs are available, regardless of whether or not the drugs have been prescribed for the adolescent or another family member. Suicide attempts by other means, for example, hanging or firearms, often warrant greater concern and consideration by clinicians. However, clinicians should not minimize the seriousness of suicide attempts by overdose before investigating other associated features of the suicide attempt.

5.18.3.3 Precipitant of the Attempt

The precipitant of a suicide attempt is another important factor which seems to be comparable across nations. In the United States, stress related to parents, school, friends, and boyfriend/girlfriend relationships are the most common precipitants (Spirito, Overholser & Stark, 1989). In Britain, very similar findings have been reported (Hawton & Fagg, 1992). It is, therefore, important for clinicians not to minimize a suicide attempt because the precipitating problem is a common everyday event. Adolescents should be interviewed regarding the precipitant to the suicide attempt in order to thoroughly understand the significance of the particular stressor in the adolescent's life.

5.18.3.4 Intent

The intent of a suicide attempt appears to be important in terms of understanding differences in the severity of suicide attempts. Spirito et al. (1995) reported that only a minority of suicide attempters report extensive planning (4%), writing a suicide note (14%), or performed final acts in anticipation of their death prior to the attempt as assessed using the Suicide Intent Scale (Beck, Schuyler, & Herman, 1974). Many adolescents expect their attempt to cause their death (31%) and a significant number are ambivalent or sorry to have survived their attempt (57%).

A factor analysis of the Suicide Intent Scale revealed three aspects of suicidal intent specific to adolescent suicide attempters (Spirito, Sterling, Donaldson, & Arrigan, 1996). One factor, labeled Expected Outcome, included items that address the expectations, purpose, seriousness, and lethality, as well as ambivalence about living following a suicide attempt. A second factor, Isolation Behaviors, consisted of isolation at the time of the attempt, timing of attempt, precautions against discovery, and acting to gain help. A third factor, Planning Activities, included items about degree of premeditation, planning, and final acts. The findings of the factor analysis suggest that it may be more useful, when evaluating adolescent suicide attempters, to examine subjective factors, particularly Expected Outcome, and selected objective factors, such as those found in the Planning Activities factor, rather than the overall level of suicidal intent.

The distinction between impulsive, unplanned attempts, and planned attempts is important in assessing the seriousness of suicide attempts. Brent (1987) found that a group of suicide attempters with high levels of hopelessness had made premeditated and highly lethal attempts compared to the unplanned and impulsive attempts made by adolescents who did not report high levels of hopelessness. In a general hospital sample, Brown, Overholser, Spirito, and Fritz (1991) found that about one-third of the suicide attempts could be classified as planned and two-thirds could be described as impulsive. Those adolescents who planned a suicide attempt were more hopeless and depressed and experienced greater suicidal ideation than the other impulsive attempters. Slap, Vorters, Chaudhuri, and Centor (1989) found that a lack of impulse control differentiated adolescent suicide attempters from a group of nonpsychiatrically hospitalized adolescents. There also appear to be gender differences in impulsivity in that male adolescent suicide attempters are more often described as displaying impulsive behavior than female suicide attempters (Marks & Haller, 1977). Because impulsivity does not characterize every adolescent suicide attempter, it can be used as a variable to classify a subgroup of these attempters.

5.18.4 PREVALENCE

Internationally, the youth suicide rate varies but is consistently one of the highest causes of death during adolescence. Diekstra and Golbinat (1993) reviewed statistics on suicide rates reported to the World Health Organization's (WHO) mortality data bank. The countries with the highest youth (15–24 years) suicide rates included Sri Lanka (62.3 per 100 000 in 1986), Finland (25.1 per 100 000 in 1991), Switzerland and Austria (16.2 per 100 000 in 1991), and Canada (15.0 per 100 000 in 1990). Suicide is the third leading cause of death among youth age 15–24 years old in the United States (National Center for Health Statistics [NCHS], 1993). The absolute number of deaths by suicide among adolescents in the US is fairly small (usually between 2000 and 2500 adolescents per year).

The statistics on suicide rates have generally been considered to be an underestimate of the true incidence due to misclassification of deaths that might be suicides (e.g., single car crashes, as well as failure to report deaths as suicides due to social stigma). The magnitude of the discrepancy between actual suicide rates and the reported statistics has been a subject of considerable debate (Males, 1991). The reporting of a suicide is related to several factors including whether mental illness preceded the death, who certifies the death, and the means of death (e.g., single car crash vs. gunshot wound). If there has been a true increase in the rate of adolescent suicide, as many suggest, one reason may be that adolescents from larger birth cohorts appear to be at higher risk for suicide than those in smaller cohorts (Holinger & Offer, 1987). This may reflect the fact that fewer opportunities and increased competition are more evident in larger cohorts than smaller ones, and may result in hopelessness and eventual suicide.

Systematic data collection of suicide attempt rates is rare. Estimates of adolescent suicide attempts in the United States range from approximately 8% in a study conducted in Oregon (Lewinsohn, Rhode, & Seeley, 1993) to 8.6% in a national survey of a representative sample of students in grades 9–12 collected by the Centers for Disease Control (1995a) using the Youth Risk Behavior Surveillance System. Based on 1988 data, Holinger (1990) estimated that over 15 000 adolescents between the ages of 10 and 19 attempted suicide and received medical attention. Several studies have attempted to calculate rates of suicide attempts seen in emergency departments. A rate of 140 per 100 000 was calculated based on suicide attempts in a suburban Chicago area (Christoffel, Marcus, Sagerman, & Bennett, 1988), while a rate of 316.8 per 100 000 was calculated in a suburban county of Atlanta (Birkhead, Galvin, Meehan, O'Carroll, & Mercy, 1993). Diekstra (1993) estimates a lifetime prevalence of attempted suicide for adolescents in the international community ranging from 2% to 20%. Because of the manner in which these statistics are collected, and because there is no central registry of suicide attempts, it is difficult to compare rates across countries, and changes in the rates over time are also difficult to interpret.

The WHO/EURO multicenter study collected data on attempted suicide in 15 European centers (Platt et al., 1992). Suicide attempt rates ranged from 72 per 100 000 in Leiden to 426 per 100 000 in Helsinki for men between the ages of 15 and 24 years. For women between the ages of 15 and 24 years, the rates ranged from 112 per 100 000 in Guipuzeon to 763 per 100 000 in Pontoise. The overall European rate for 15 to 19 year-old males was 106 per 100 000 and 323 per 100 000 for 15 to 19 year-old females. The highest adolescent rate was 911 per 100 000 in Pontoise followed by 791 per 100 000 in Szeged.

As regards suicide ideation, females report much higher rates than males in most countries of the world. Data collected in 1993 using the Youth Risk Behavior Surveillance Survey in the United States supported these findings (CDC, 1995b). Across all racial and ethnic groups, females reported more suicidal ideation than males.

5.18.5 ETIOLOGY: RISK FACTORS

There is no definitive underlying cause for suicide. However, there are a number of demographic (e.g., gender and race), psychiatric (e.g., certain mental disorders such as depression and conduct disorder; prior suicide attempts), biological (e.g., dysfunction of the serotoninergic system), and socio-environmental (access to lethal methods, e.g., firearms; exposure to suicidal behavior) factors which increase the risk of suicide. These risk factors will be reviewed in turn.

5.18.5.1 Gender

In all age groups, females attempt suicide more often than males, but males complete suicide more often than females. According to Centers for Disease Control data collected in the United States in 1992 (1995a), the suicide rate for adolescent males was 18.4 per 100 000 compared to a rate of 3.7 per 100 000 for adolescent females. Diekstra (1993) examined the suicide rates in 19 countries between 1970 and 1986. The rate for men increased across all age groups with a mean increase of 70% in the suicide rates in adolescents and young adults. For young women the increase was closer to 40%. One potential reason for the difference in suicide rates between males and females may be method preference. Males are more likely to use highly lethal suicide attempt methods whereas females are more likely to use less lethal methods, particularly drug overdoses (Sigurdson, Staley, Matas, Hildahl, & Squair, 1994). Although some drug overdoses can be lethal in high doses, most suicide attempts by drug overdose are reversible if brought to medical attention within a reasonable amount of time. In India and in Southeast Asia, where medical care is not always readily available, the majority of suicides are committed by women by overdose (Shaffer, 1988).

The difference in suicide methods between males and females has been linked to stereotypic sex role behavior. That is, men tend to delay expressing their distress and during this period of delay, find more lethal methods for suicide attempts, while females express their affect in a more immediate and sometimes impulsive fashion. There is some evidence to support this contention. When male adolescent suicide attempters were compared to female adolescent suicide attempters on measures of depression, hopelessness, suicidal ideation, and suicidal intent, the only differences between sexes were that males reported more final acts in anticipation of death and a greater degree of planning than the female attempters (Spirito et al., 1993).

5.18.5.2 Race

In the United States, adult Caucasians commit suicide more frequently than adult African-Americans, typically at a 3:1 ratio (NCHS, 1993). However, since the mid-1960s the suicide rate for adult African-American males has nearly tripled, and has doubled for adult African-American females. This increase has also been noted in adolescents, particularly males. For African-American adolescent males, the rate increased from 5.5 per 100 000 in 1980 to 14.8 per 100 000 in 1992, an increase of 164%. The suicide rate rose from 1.6 per 100 000 in 1987 to 1.9 per 100 000 in 1992 for African-American adolescent females, an increase of 19% (CDC, 1995a).

These differences in suicide rates by race have not been well studied. Gibbs and Hines (1989) hypothesized that there are protective factors in the African-American population, such as strong maternal family ties or church and fraternal organizations that increase social support and cohesion, thus decreasing the rate of suicide. Others note that African-American males die by other violent means besides suicide more frequently than whites, which may account for the larger rates of suicide in Caucasians than African-Americans (Gibbs & Hines, 1989). The dramatic increase in the 1980s of African-American adolescent suicide rates has led some to speculate that if culture is a protective factor, the rates of suicide may be equalizing because social factors which promote deculturalization exist (Shaffer, Gould, & Hicks, 1994).

Hispanic youth have lower rates of suicide than Caucasians, although suicide was the third leading cause of death for Hispanic youth in 1990 (NCHS, 1993). Gender differences in suicidal behaviors seem to be less prominent among Hispanic teenagers (Spirito et al., 1993).

Very high rates of suicide have been estimated among Native American adolescents (26.3 per 100 000) (US Congress, Office of Technology Assessment, 1990). This is due, at least in part, to the much larger proportion of adolescents and young adults among Native Americans than Caucasians. In addition, the suicide rates vary greatly among different tribes (McIntosh, 1983–84). One troubling example of the problem of suicide among Native American adolescents is the Zuni tribe. The suicide rate among Zuni adolescents rose steadily from 5.6 per 100 000 in 1957 to 32.2 per 100 000 in 1989.

5.18.5.3 Psychiatric Disorders

Certain psychiatric disorders are common among adolescents who complete suicide. Approximately 90% of adolescents in the New York City area who completed suicide in one study had a prior psychiatric diagnosis (Shaffer et al., 1996). These findings are comparable to those reported previously in Pittsburgh (Brent et al., 1988) and in Finland (Marttunen, Aro, Henriksson, & Lonnqvist, 1991). The most common psychiatric diagnosis in the New York City study (Shaffer et al., 1996) was Mood Disorder, occurring at a rate (61%) somewhat higher than that reported in the Finnish national study (51%) and in Pittsburgh (49%). Bipolar disorder was uncommon in both the Finnish study and the New York City study. In the New York City study, there was also a much higher percentage of Disruptive Behavior Disorders among the adolescents than the Finnish national study.

Bipolar Disorder, Affective Disorder with comorbidity (especially Attention Deficit Disorder, Substance Abuse, and Conduct Disorder), a lack of previous mental health treatment, and availability of firearms in the home differentiated 27 adolescents who had committed suicide from 56 psychiatric inpatients who had seriously considered or attempted suicide (Brent et al., 1988). Blumenthal and Kupfer

(1986) believe that the coexistence of the antisocial and depressive symptoms are a particularly high risk combination for adolescent suicide. A high rate of substance use in the male adolescents who completed suicide was found in the New York City study (Shaffer et al., 1996), where approximately 60% of the male adolescent suicide victims had one of the following risk factors: prior suicide attempt, mood disorder, and substance/alcohol use. Depression and substance abuse were the most common risk factors in Canadian youths age 15 to 24 years old (Sigurdson et al., 1994). Alcohol plays a prominent role in suicides among Alaskan natives (Hlady & Middaugh, 1988) and is particularly prevalent in tribes which have high adolescent suicide rates.

Some adolescents evidence no psychiatric disorder prior to suicide completion (Apter et al., 1993). These adolescents have been described as perfectionistic and rigid (Apter et al.), and more likely to have prior suicidal behavior, legal, or disciplinary problems, and family psychopathology (Brent, Perper, Moritz, Baugher, et al., 1993). Thus, it seems likely that pathology exists in this group, but is more covert and/or denied by the family.

There is limited research regarding personality characteristics and disorders among suicidal adolescents. In Finland, 32% of adolescents in a sample of completed suicides were diagnosed with a personality disorder (Marttunen et al., 1991). Antisocial Personality Disorder among males and Borderline Personality among both males and females were the most common diagnoses reported. One-third of 58 suicides between the ages of 15 to 19 years old in Sweden were diagnosed with Borderline Personality Disorder by psychological autopsy methods (Runeson & Beskow, 1991).

Psychiatric disorders have also been investigated among suicide attempters and ideators. There has been interest in determining whether adolescent suicide attempters have a higher rate of Borderline Personality Disorder than other adolescents. However, there have only been a few studies in this area using adolescent samples. Pfeffer, Newcorn, Kaplan, Mizruchi, and Plutchik (1988) examined factors associated with suicidal behavior in a chart review of the clinical diagnoses of 200 consecutively admitted adolescent psychiatric inpatients and found a positive association between suicidal behavior in adolescent females and Borderline Personality Disorder. Brent et al. (1993) reported on increased evidence of borderline personality disorder or traits among psychiatrically hospitalized suicide attempters compared to psychiatric controls who had never been suicidal.

Depression has also been implicated in suicidal ideation and attempts (Garrison, Jackson, Addy, McKeown, & Waller, 1991). Mood disorders, particularly depression, may be precursors to different types of suicidal behavior. However, it appears that other related difficulties (e.g., substance abuse or conduct problems), are often seen in combination with depressed mood. Substance use is also related to suicide attempts (Pfeffer et al., 1988) and is considered a risk factor for both attempted and completed suicide because it can affect multiple domains of functioning, (i.e., affective, cognitive, social, and behavioral; Shaffer, 1988). For example, alcohol may lower inhibition and lead to more impulsive and dangerous behaviors that result in death. When combined with depressed mood, substance abuse significantly increases the risk of suicide in vulnerable groups.

5.18.5.4 Prior Suicide Attempts

One of the best predictors of completed suicide is a previous suicide attempt (Spirito, Brown, et al., 1989; Diekstra, 1989). In their review of the literature, Shaffer and colleagues (1988) concluded that in approximately 20% of all suicides completed by male adolescents and 33% of all suicides completed by females, a history of a prior suicide attempt is evident. Among adolescents, male suicide attempters are much more likely than females to eventually complete suicide (Goldacre & Hawton, 1985; Otto, 1972). A previous attempt by active means, such as hanging or shooting, has also been found to be predictive of later completed suicide (Otto, 1972). In addition, it has been estimated that adolescents who attempted suicide are 8.1 times more likely to reattempt suicide than those who have never made a suicide attempt (Lewinsohn et al., 1993).

5.18.5.5 Biological Factors

The endocrine and neurotransmitter systems have been the major focus of research to determine whether biological correlates of suicide exist. Most of this research has been conducted with adults and is quite complex and methodologically complicated. Findings have been based on postmortem tissue analyses, brain chemistry, cerebrospinal fluid, and neuroendocrine correlates. Major findings are available in detailed reviews (Rifai, Reynolds, & Mann, 1992; Roy, 1994) and are briefly outlined in this chapter.

Serotonin is the major neurotransmitter that has been linked to suicidal behavior. Low levels of serotonin (5-HIAA) and 5-hydroxytryptophan (5-HTP) have been reported among suicide attempters and found to be related to impulsivity and aggressive behavior in both depressed and non-depressed samples (Roy, 1994; Traskman-Bandz et al., 1993). Serotonergic changes appear to be most strongly related to impulsive aggression and violent suicides (Mann et al., 1992). In another study (Nordstrom et al., 1994), low CSF levels of 5-HIAA were related to increased suicide risk following attempted suicide in depressed patients.

Several studies report that suicidal behavior is associated with low CSF concentrations of the dopamine metabolite, HVA. However, these findings seem to be applicable to depressed patients only. In a five year follow-up study (Roy, DeJong, & Linnoila, 1989), dopaminergic activity was shown to be related to repeat suicide attempts among depressed patients. The mechanism by which serotonin levels affect suicidal behavior is unknown, but environmental factors and other variables are most likely involved because many individuals with low CSF concentrations do not attempt suicide.

Although more research is needed to clarify the differential roles of various receptor sites and their relation to suicide, reduced serotonin levels are related to repeat suicide attempts, impulse control, and aggression in depressed and nondepressed samples. It also appears that the reduced activity of serotonin may be related to a higher risk of death by suicide following an unsuccessful suicide attempt. Although this area has promise, a lumbar puncture is necessary to gather a specimen of CSF, which limits its applicability to clinical care. There is a serotonin metabolite level that can be measured by urine sample, which may ultimately provide useful information regarding suicide risk. However, at present biological factors are not routinely assessed when evaluating suicidal behavior.

5.18.5.6 Method of Suicide: Firearms

In the United States, suicides are most commonly completed by using firearms (Holinger, 1990). In 1992, 64.9% of all suicides among people less than 25 years old were by firearms (CDC, 1995a). Hanging was the second most common suicide method. For 15–19 year-olds, the rate of suicide by firearms increased 81% between 1980 and 1992, probably reflecting greater access to firearms by adolescents. Guns have been shown to be twice as prevalent in the homes of adolescents who complete suicide compared to adolescents who either attempt suicide or have other psychiatric disorders (Brent et al., 1991). However, the method of storing guns has no effect on suicide (Brent et al., 1991). When firearms are available in the home, suicidal adolescents will actively kill themselves at a rate 75 times higher than when firearms are not available at home (Brent et al.). The most common method of attempting suicide among adolescents is by drug overdose. Estimates range between 75% and 95% of subjects across a number of studies conducted in the United States (Spirito, Brown, 1989) and Britain (Hawton & Fagg, 1992).

5.18.5.7 Exposure

Another potential risk factor for completed suicide is previous exposure to suicidal behavior in the family or in the peer group. Shaffi, Carrigan, Whittinghill, and Derrick (1985) found that adolescents who complete suicide are significantly more likely to have been exposed to suicidal behavior in parents, relatives, or friends. The risk of completing suicide increases with having a family member who has committed suicide. Higher rates of suicide and suicide attempts have been noted in the relatives of suicide completers compared to community controls (Brent et al., 1994).

Not only does exposure lead to increased risk of completed suicide, but the risk of attempting suicide increases significantly if there is a family history of suicide attempts (Sorenson & Rutter, 1991). Adolescents who report that a peer attempted or completed suicide display higher levels of depression and more frequent suicidal behavior themselves than those with no such exposure (Hazell & Lewin, 1993). However, Brent, Perper, Moritz, Baugher et al. (1993) examined the psychiatric sequelae of the suicide of a peer and found that suicide attempt rates did not differ between those exposed and those unexposed. The exposed group demonstrated more suicidal ideation and depression than the unexposed group. At three-year follow-up, the exposed group had an increased incidence of depression, anxiety, and Post-traumatic Stress Disorder compared to the unexposed group, but the two groups did not differ on suicidal behavior (Brent, Moritz, Bridge, Perper, & Canobbio, 1996).

The findings with respect to familial clustering of suicide and suicide attempts indicate that a genetically linked biological mechanism such as the serotonergic system may be involved. Alternatively, exposure may allow suicidal behavior to be modeled by family members who attempt or complete suicide. There are a number of studies (discussed in Section 5.18.7.7) Au:X-ref.OK?

in which family dysfunction increases the risk of suicidal behavior. Studies designed to differentiate the role of genetic and/or psychological mechanisms have yet to be conducted.

Adolescents may also be exposed to suicide indirectly via television or the news media. Several studies have demonstrated that a prominent report of a suicide in the media leads to an increase in suicides, particularly among young people, for one to two weeks following the news coverage (Shaffer, 1988; Phillips & Carstensen, 1986). In addition, exposure to a violent death, regardless of the means (e.g., homicide or car crash) increases the likelihood of subsequent suicides among youth in the community. It has been found that victims of high social status, deaths motivated by revenge, and failure of the media to report a victim's prior psychiatric disturbance increase the risk of imitation (Davidson, 1989; Davidson & Gould, 1989). Consequently, the Centers for Disease Control (1988) have developed guidelines for the media on reporting suicides.

There is also conflicting data regarding the relationship between fictional suicide stories and subsequent imitative suicides. Gould and Shaffer (1986) found increased suicide rates among adolescents in the New York City area two weeks before and two weeks after four television movies about suicide were broadcast. However, when the same movies were shown in California and Pennsylvania, albeit by only one station, which limited the audience and the impact, the findings were not replicated (Phillips & Paight, 1987).

5.18.6 DEVELOPMENTAL COURSE

The rates of suicide vary across development. Although rare, suicide rates among children 10 to 14 years old in the United States have been estimated to be on the rise, increasing two and a half times between 1970 and 1986, from 0.6 to 1.5 per 100 000 (US Bureau of the Census, 1990). As stated previously, the rate for adolescents increased from 8.5 to 10.9 per 100 000 between 1980 and 1992 (CDC, 1995a). Various investigators have emphasized the need to examine developmental risks and protective factors for suicide. Characteristic of this approach is an emphasis on the relation between suicidal behavior and developmental factors, for example, sociocognitive functioning, as opposed to chronological age. The range of developmental skills within a particular age group may vary greatly or fluctuate as a function of emotional distress (Jacobson et al., 1995).

Up to 1997, limited empirical work has examined suicide from a developmental perspective. However, McDowell and Stillion (1994) delineated several commonalities and differences in suicide across the lifespan. The commonalities include biological, psychological, cognitive, and environmental factors. With respect to biological factors, the serotonin hypothesis and gender differences described have been shown to transcend all ages (McDowell & Stillion, 1994). Psychological factors such as depression, low self-esteem, and hopelessness also contribute to suicide across the life span. Cognitive factors related to suicide across development include depressogenic thought patterns, such as cognitive rigidity, all-or-none thinking, and overgeneralization. In addition, perfectionism has been found to be a risk factor regardless of age (Hewitt, Flett, & Turnball-Donovan, 1992; Hewitt, Newton, Flett, & Callander, 1997). Regarding environmental factors, exit events (e.g., parental divorce), all types of abuse, and family stressors have been found to play a role in suicide for both children and adults, as has the presence of a means for completing suicide. Moreover, the precipitants of suicidal acts in young children (Pfeffer, Conte, Plutchik, & Jerrett, 1980) and adolescents (Spirito et al., 1989) are similar.

McDowell and Stillion (1994) also outlined suicide risk factors unique to young children and to adolescents. Risk factors for children younger than 15 years of age include impulsivity, a sense of inferiority, the expendable child syndrome, an immature view of death, and extreme levels of persistent family turbulence. With respect to impulsivity, suicidal behavior in young children has been found to be more violent and impulsive; for example, jumping or running out in front of a car (Joffe & Offord, 1983). Pfeffer et al. (1980) reported that jumping was the most common means of attempt among 6 to 12 year-old outpatient psychiatry patients. A sense of inferiority was used by McDowell and Stillion (1994) to refer to the Eriksonian concept of industry, or the need to develop confidence and positive self-regard. This concept is relevant to the findings in the literature regarding the relation between suicide and low self-esteem (e.g., Stillion, McDowell, & Shamblin, cited in McDowell & Stillion, 1994). The "expendable child syndrome" is a concept that refers to children whose parents communicate low regard and even hostility toward them to the extent that children believe they are expendable, which is related to the wish to die.

Concrete operational thinking, common in latency-aged children, has implications for the meaning and understanding of suicide in young children. Carlson et al. (1987) reported that although children between 8 and 12 years of age contemplate suicide, they have difficulty conceiving of the circumstances that might lead one to commit suicide and plan to carry it out. Similarly, young children's understanding of death, causality, and reversibility have important implications for suicidal behavior in this group (Jacobson et al., 1995). In addition, family problems have long been reported to be a more important factor for younger than older suicidal children because of the dependence of younger children on adult caregivers for survival (Otto, 1972). A particular family factor identified as a risk factor for children under 12 years of age is the presence of suicidal behavior in their parents (Pfeffer et al., 1980). With respect to suicide attempts, attempters younger than 14 years of age have been found to be more likely to leave a suicide note prior to an attempt and use hanging vs. ingestion as a means to attempt suicide (Garfinkel, Froese, & Hood, 1982).

Risk factors specific to adolescent suicide include the adolescent's mood state, the inability to develop a sense of identity, egocentrism, risk-taking behaviors, and social isolation (McDowell & Stillion, 1994). As discussed earlier, psychiatric diagnoses, particularly depression, conduct disorder, and substance abuse, are significant risk factors for suicide during adolescence. In addition, adolescents who have no sense of identity may not have the necessary level of self-competence to cope with the stressors that confront them during this period. Self-esteem has been found to be inversely related to suicidal ideation and behavior in both adolescent psychiatric patients and normal high school controls (Overholser, Adams, Lehnert, & Brinkman, 1995). Similarly, cognitive changes characterized by a renewed period of egocentrism may increase feelings of self-consciousness and consequently overgeneralization, such as that reported in depressive cognition patterns.

It has been suggested that a sense of invulnerability characterizes adolescents, and this has been hypothesized to lead to increased risk taking. Empirical studies of the relation between risk-taking behaviors and suicide have not yet been reported. Some studies have examined the role of interpersonal relationships and suicidal behavior. Peer relationship difficulties are one of the most common precipitants of suicide attempts in adolescence (Spirito et al., 1989), and suicidal adolescents have been found to be socially isolated and have poor peer relationships (King, Raskin, Gadowski, Butkus, & Opipari, 1990; Spirito et al. 1989). Adolescents are also more likely than younger children to mutilate themselves and make a planful suicide attempt (McDowell & Stillion, 1994).

5.18.7 CONCEPTUALIZATION AND CLINICAL FORMULATION OF ATTEMPTED SUICIDE

There is no standardized treatment for adolescent suicide ideators and attempters. Consequently, it is important to perform a comprehensive assessment in order to target in treatment those deficit areas that may trigger further survival behavior.

5.18.7.1 Multimodal Assessment

There are a number of important factors to consider when evaluating an adolescent who has just attempted suicide. One assessment model can be seen in Figure 1. This conceptual model is similar to a number of other models that have been hypothesized in the literature, and is based on the assumption that suicidal individuals display cognitive, behavioral, and emotional deficits which increase their risk of suicidal behavior at times of increased stress. Both the adolescent and his/her parents/guardians should be interviewed to complete the assessment.

Only one assessment model to date has been tested empirically. Based on epidemiologic data collected with a large sample of high school students, Lewinsohn, Rohde, and Seeley (1996) postulated a model which includes the constructs of psychopathology, environmental factors, interpersonal problems, and physical illness which affect cognitions/coping, leading to suicidal behavior. Using structural equation modeling, psychopathology had the strongest (0.34) direct effect on suicidal behavior. Physical health and environmental factors had small, but statistically significant, direct effects. All four constructs had statistically significant indirect effects on suicidal behavior via cognitive/coping style. Cognitive/coping style also had a direct effect on suicidal behavior (0.20). Tests of this model with clinical populations would be a useful contribution to the literature.

Assessment of the suicide attempter at the time of the attempt is critical in formulating an adequate treatment plan. Assessment should be conducted with both the adolescent and family members. Suicidal ideation preceding the attempt, as well as the reasons for the suicide attempt, must be investigated. The first goal of this evaluation is to determine whether the adolescent requires a psychiatric admission, usually based on continued suicidal ideation, a major psychiatric disorder that cannot be treated initially as an outpatient, or psychological or family factors that may place the adolescent at continued risk for immediate self-harm or a reattempt. The following areas are particularly important to assess: suicidal intent

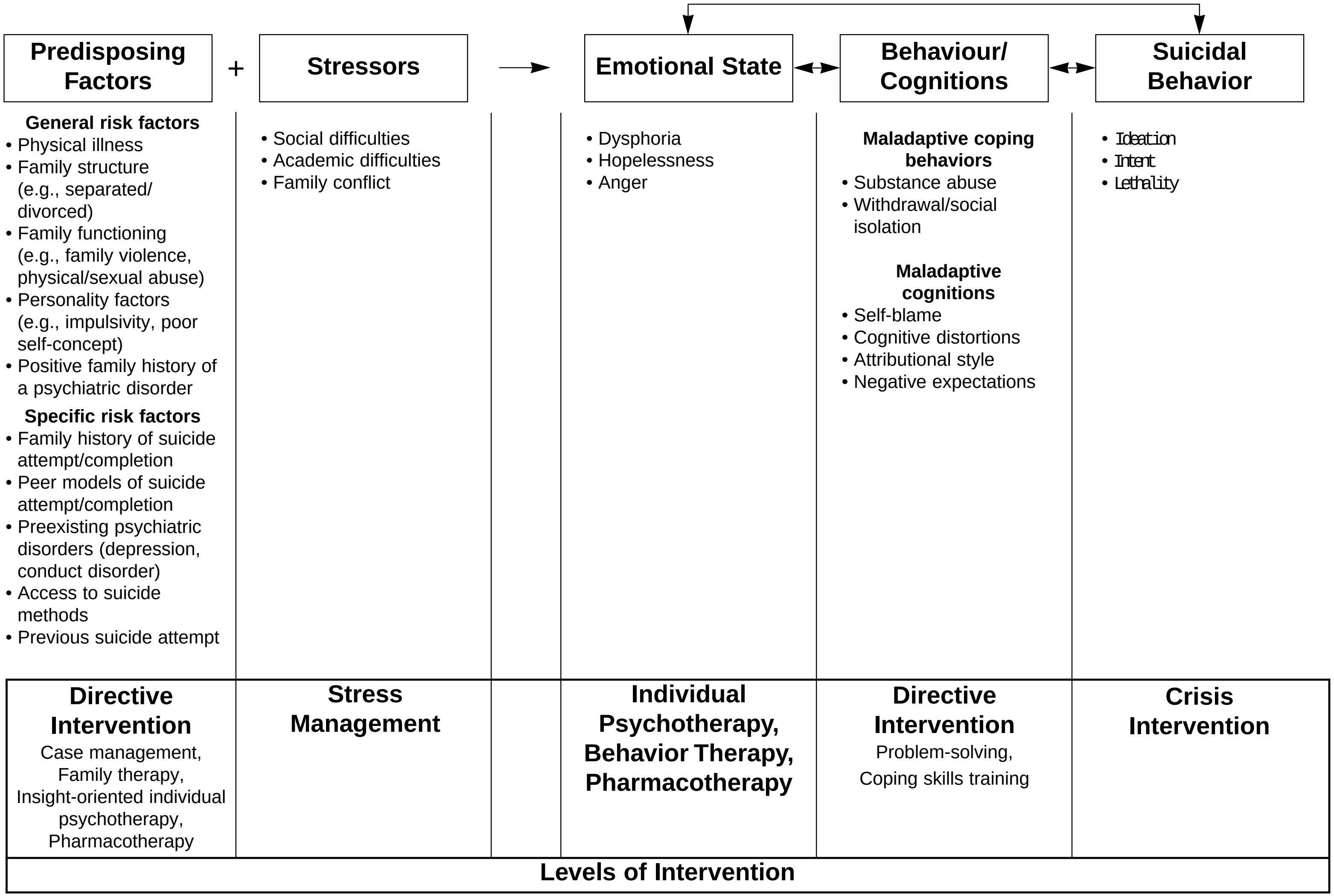
Predisposing Factors
+
Stressors
Emotional State
Behaviour/ Cognitions
Suicidal Behavior
General risk factors
• Physical illness
• Family structure (e.g., separated/ divorced)
• Family functioning (e.g., family violence, physical/sexual abuse)
• Personality factors (e.g., impulsivity, poor self-concept)
• Positive family history of a psychiatric disorder
Specific risk factors
• Family history of suicide attempt/completion
• Peer models of suicide attempt/completion
• Preexisting psychiatric disorders (depression, conduct disorder)
• Access to suicide methods
• Previous suicide attempt
• Social difficulties
• Academic difficulties
• Family conflict
• Dysphoria
• Hopelessness
• Anger
Maladaptive coping behaviors
• Substance abuse
• Withdrawal/social isolation
Maladaptive cognitions
• Self-blame
• Cognitive distortions
• Attributional style
• Negative expectations
• Ideation
• Intent
• Lethality
Directive Intervention
Case management, Family therapy, Insight-oriented individual psychotherapy, Pharmacotherapy
Stress Management
Individual Psychotherapy, Behavior Therapy, Pharmacotherapy
Directive Intervention
Problem-solving, Coping skills training
Crisis Intervention
Levels of Intervention

Figure 1 An assessment model (*Adolescent Health Problems* (p.103), by Wallarden and Siegel, 1995, New York: Guilford Press. Copyright 1995 by Guilford Press, Reprinted with permission).

or reason for the suicide attempt; current coexisting stressors; preexisting psychiatric disorders; family history, particularly of suicidal behavior; previous suicidal behavior; self-stated coping strategies and social support; an evaluation of the risk of an immediate reattempt; and attitudes of the adolescent and family towards further mental health treatment (Hawton, 1986). These factors are elaborated on in the following sections.

5.18.7.2 Attempt Characteristics

The intent of the suicide attempt maybe the most important characteristic of the attempt to assess in clinical formation. Data from adolescents (Spirito et al., 1989) suggest that two aspects of suicidal intent, making plans prior to a suicide attempt and expected negative outcome, are the most important factors in determining high risk adolescents.

5.18.7.3 Emotional State

An adolescent's mental status at the time of the attempt, as well as premorbid psychological factors (particularly depression, hopelessness, and anger) are important to assess, both in terms of their individual impact and their contribution to a diagnosis of a specific psychiatric disorder. Depressed mood has been implicated in many studies of adolescent suicide attempters. Spirito et al. (1989) reviewed the literature and found very high rates of depressed affect among attempters, ranging from 30% to 82%, across studies in psychiatric inpatients, outpatients, and emergency department settings. The lowest rates were found in suicide attempters who were medically hospitalized.

Hopelessness may contribute to suicidal ideation and behavior above and beyond depression. Hopelessness refers to a negative attitude about future events, including a global expectation of negative future occurrences. Spirito, Williams, Stark, and Hart (1988) found higher levels of hopelessness among adolescent suicide attempters than among either nonsuicidal adolescents in an outpatient psychiatric clinic or normal control subjects. In a separate study, Spirito, Overholser, and Hart (1991) found higher levels of hopelessness in adolescent suicide attempters than in nonsuicidal psychiatrically hospitalized adolescents. Some studies (e.g., Robbins & Alessi, 1985) have reported that hopelessness among adolescents is associated with the number of suicidal gestures, the seriousness of suicidal intent, and the medical lethality of a suicide attempt. Hopelessness has also been shown to have a stronger relationship to suicidal behavior than other factors, such as family dysfunction (Levy, Jurkovic, & Spirito, 1995). Gender effects on hopelessness (i.e., females report greater levels of hopelessness than males) have also been reported (Cole, 1989; Spirito, 1991). However, several studies have found comparable levels of hopelessness among adolescent suicide attempters and depressed adolescents (Brent, Kolko, Allan, & Brown, 1990; deWilde, Kienhorst, Diekstra, & Wolters, 1993).

Another important variable related to adolescent suicide attempts is anger. Anger has been shown to be more strongly related to suicidal ideation and behavior than depression (Cairns, Peterson, & Neckerman, 1988; Choquet & Menke, 1989). Repressed anger has also been shown to be highly correlated with hopelessness (Brown et al., 1991). When compared to normal adolescents, suicide attempters report an increased incidence of experiencing anger as well as higher levels of internalized anger (Lehnert, Overholser, & Spirito, 1994). Many adolescents report significant anger prior to their attempt (Hawton, O'Grady, Osborn, & Cole, 1982; Withers & Kaplan, 1987). These angry feelings may be expressed via verbal outbursts, but at other times the anger is repressed. Anger has also been associated with the seriousness of the suicide attempt (Gispert, Wheeler, Marsh, & Davis, 1985). In several studies, the strongest predictor of future suicidal tendencies and suicide reattempts was the level of anger at the time of the initial attempt (Myers, McCauley, Calderon, & Treder, 1991). At times, the anger and irritability so prominent in adolescent suicide attempters may be a symptom of depressive illness. At other times, anger and irritability may also be indicative of a conduct disorder. Indeed, a large proportion of childhood suicide attempters have been shown to exhibit assaultive and aggressive behavior independent of depression (Pfeffer, Plutchik, & Mizruchi, 1983). Thus, the origin of the aggressive and angry feelings must be differentiated on an individual basis.

5.18.7.4 Cognitive Variables

A negative attributional style (i.e., likelihood of attributing undesirable events to oneself or others, to unstable factors, or to global causes) has been found to be related to suicidal behavior in adults. Evidence in adolescents is more limited. The results of two studies of adolescents showed that suicide attempters were more likely than the nonsuicidal psychiatric patients to attribute good events to global causes and negative events to stable aspects of the environment (Rotheram-Borus, Trautmen, Dopkins, & Shrout, 1990; Spirito, Overholser et al., 1991). Although further study is needed, these data suggest that a maladaptive attributional style may be more related to overall level of psychopathology than specifically to suicidal behavior among adolescents.

Another cognitive factor which has been seen studied in suicidal adolescents is problem solving. Problem-solving deficits are important to investigate because the data have direct implications for the types of interventions that may be appropriate for this high-risk group. Female adolescent suicide attempters have been distinguished from nonsuicidal psychiatrically disturbed adolescents by their poorer interpersonal problem-solving ability and maladaptive attributional style (Rotheram-Borus et al., 1990). Sadowski and Kelley (1993) found that suicide attempters reported poor problem-solving orientation as well as less adequate problem-solving skills in general, than a comparison group of psychiatric patients and normal high school students. In addition, these adolescent suicide attempters reported less cognitive responding and deficits in initiating the problem-solving process than the other groups when confronted with problem situations. Suicidal children have also been found to use fewer active cognitive coping strategies than controls (Asarnow, Carlson, & Guthrie, 1987). Orbach, Rosenheim, and Mary (1987) compared psychiatrically hospitalized children with suicidal ideation and behavior to children with chronic medical illnesses and normal controls on the ability to generate alternative solutions to situations involving life and death. The suicidal children generated fewer alternatives than the normal or the chronically ill children. Among the suicidal children, those most attracted to death were least capable of generating effective alternative solutions. In a study comparing inpatient suicide attempters to controls, there was no difference in their ability to generate solutions to standardized personal problems (Wilson et al., 1995). However, when asked about how they coped with their most recent severe stressor, suicide attempters reportedly used fewer coping strategies and more maladaptive coping strategies than the controls.

Perfectionism is another cognitive characteristic which has been investigated in regard to adolescent suicidal behavior. Both self-oriented and socially prescribed perfectionism were found to be positively related to hopelessness in adolescent psychiatric inpatients (Hewitt et al., 1997). However, only the relation between socially prescribed perfectionism and suicidal ideation attained significance. Moreover, only socially prescribed perfectionism contributed unique variance to suicidal ideation when the effects of age, gender, self-oriented perfectionism, and hopelessness were controlled.

In summary, cognitive deficits such as poor problem solving and cognitive distortions appear more frequently in a significant portion of adolescent suicide attempters compared to non-suicidal adolescents. It is unclear whether these are temporary deficits evident at the time of acute stress, or more enduring characteristics. In either case, it is important for clinicians to identify and intervene in these deficit areas when they present.

5.18.7.5 Maladaptive Coping Strategies

Social isolation and avoidance behavior are maladaptive coping strategies reported by adolescent suicide attempters. More adolescent suicide attempters report using social withdrawal when faced with a problem than normal control subjects (Spirito Overkolser, et al., 1989). However, the use of social withdrawal among adolescent suicide attempters is similar to that found among nonsuicidal psychiatrically hospitalized adolescents (Spirito, Francis, Overholser, & Frank, 1996). Both groups reported using social withdrawal as a maladaptive coping strategy. These findings suggest that it is important for clinicians to address avoidance behavior. For example, adolescents who attempt suicide maybe similar to nonsuicidal adolescents in the sense that they attempt to use a number of coping strategies prior to the suicide attempt. However, they may use more maladaptive cognitive strategies and at a certain point, withdraw from social contact, which may increase the risk of a suicide attempt. This may be a sequence effect; adolescent suicide attempters may initially use adaptive coping strategies, then move on to maladaptive strategies, and finally withdraw prior to an attempt.

Attachment behavior has also been investigated as a precursor to social withdrawal. In a sample of both inpatient and outpatient adolescent psychiatric patients, 86% had attachment-related trauma when they were young (Adam, Sheldon-Keller, & West, 1996). These traumatic family experiences, including separation and abuse, may affect the development of attachment behavior, which may be a mediating mechanism for suicidal behavior, especially among female adolescents (Adam et al., 1996).

Substance use and abuse can be considered another maladaptive coping behavior. Studies in the US report significant rates of substance abuse ranging from 23% to as high as 42% among adolescent suicide attempters (Spirito, Brown et al., 1989). Work by Hawton & Fagg (1992), found elevated rates (39% males, 21% females) of consumption of alcohol a few hours before the attempt among British suicide attempters, but rates were not as high as typical rates reported in the US. In a large survey of high school students, alcohol use was associated with an increased suicide attempt rate, particularly among Caucasian adolescents (Adcock, Nagy, & Simpson, 1991). For Caucasian adolescents, the suicide attempt risks increased from 9% to 27% while among African-Americans it increased from 10% to 21%. Substance use, if it is of a chronic nature, can lead to significant psychosocial problems which eventually result in a suicide attempt. If substance use precedes a suicide attempt, it may interfere with problem solving and thus be a proximal factor to the suicide attempt.

5.18.7.6 Physical and Sexual Abuse

A history of physical or sexual abuse has been associated with suicidal behaviors in a number of studies. In a survey of almost 4000 predominantly white adolescents in a suburban/rural area, attempted suicide was reported by 4.7% of the physically-abused students, 5.1% of the sexually-abused students, and 9.2% of students who reported both physical and sexual abuse (Hibbard, Ingersoll, & Orr, 1990). Garnefski, Dieskta, and de Heus (1992) compared a sample of predominantly Caucasian high school students with a history of suicidal behavior to high school students with no such history. Of the female attempters, 51% reported a history of physical abuse and 32% had experienced sexual abuse. Of the females who did not attempt suicide, 24% reported a history of physical abuse and 7% reported sexual abuse. Of the male attempters, 22% had experienced sexual abuse compared to 2% of the nonattempters. This study did not include truant or dropout students, who may have an even higher incidence of physical and sexual abuse. A comparison of adolescent suicide attempters with depressed and nondepressed adolescents with no history of suicide attempts revealed that suicidal adolescents were more likely to report a history of sexual abuse (33%) than depressed (21%) or nondepressed adolescents (5%) who had not attempted suicide (deWilde et al., 1992). More depressed individuals (29%) reported physical abuse than suicide attempters (23%) or nondepressed adolescents (7%).

Deykin, Alpert, and McNamara (1985) examined adolescent suicide attempters who were identified following an emergency department visit and a matched comparison group of adolescents treated for medical conditions unrelated to suicide attempts. The suicide attempters were three to six times more likely than the comparison group to have had contact with state social services departments for an abuse-related incident. No information regarding reasons for social service involvement was available. The relationships between prior physical or sexual abuse and subsequent health-risk behaviors and suicide attempts were examined in a nonclinical sample of 600 high school students (Riggs, Alario, & McHorney, 1990). Those students who were victims of previous sexual maltreatment were three times more likely to have attempted suicide and those students with a history of physical maltreatment were five times more likely than those without a history of maltreatment.

Psychiatrically-hospitalized adolescents categorized as attempters, ideators, or nonsuicidal (i.e., those admitted for reasons other than suicide attempts and with no current suicide ideation) have been studied as to the effects of abuse (Shaunesey, Cohen, Plummer, & Berman, 1993). Females were more likely than males to have histories of both types of abuse, or sexual abuse alone, and females with past sexual abuse had a higher frequency of suicide attempts than females with no history of sexual abuse. In addition, adolescents with a high frequency of abuse (i.e., 10 or more events) had made significantly more suicide attempts.

5.18.7.7 Family Functioning

A number of studies have shown higher rates of family disturbance in suicide attempters compared to psychiatric controls (Mitchell & Rosenthal, 1992) and the severity of suicidal intent has been related to the degree of family dysfunction (Brent et al., 1990). Suicidal psychiatric inpatient adolescents and suicidal high school students report more distress and family dysfunction than nonsuicidal high school students (Adams, Overholser, & Lehnert, 1994). In a sample of high school students, family dysfunction influenced suicidal behavior indirectly through depression (Martin, Rozanes, Pearce, & Allison, 1995). In a large study of high school students, adolescents who reported a history of a suicide attempt also reported more frequent stresses related to their parents and more physical abuse, runaways, and family disruption than those who had not had previous attempts (Wagner, Cole, & Schwartzman, 1995).

Stressful life events as triggers of suicidal behavior have been of interest to many clinicians and researchers. Adams, Overholser, and Spirito (1994) examined both intrafamilial and extrafamilial life events. Hospitalized adolescent suicide attempters reported the loss of a family member due to death or divorce more often than high school controls (Adams, Overholser, & Spirito, 1994). Adolescent suicide attempters also reported a greater number of "exit events" such as parental divorce and chronic life stressors compared to nonsuicidal adolescents. Suicidal ideation was directly related to stressful life events among the female suicide attempters. The number of exit events was positively related to hopelessness and suicidal ideation among adolescent females. Although suicide risk seems to increase when elevated levels of stressful events exist, a high degree of family cohesion may mediate the effects of these stressful events (Rubenstein, Heeren, Housman, Rubin, & Stechler, 1989).

5.18.8 PSYCHOSOCIAL TREATMENTS

Treatment for adolescent suicidal behavior is quite varied and there is no one standard treatment approach for this group. A number of clinical articles have addressed the treatment of suicide ideators, focusing mainly on the management of the acute suicidal episode with little emphasis on ongoing treatment. A limited number of articles have been written on the treatment of adolescents following a suicide attempt. The majority of this work has also focused on short-term management of the adolescent following the suicidal crisis. Finally, interventions that focus on adolescents that have psychiatric disorders common to suicidal behavior, such as depression (e. g., Lewinsohn, Clarke, Hops, & Andrews, 1990), might also be considered as treatment approaches for adolescent suicidal behavior.

5.18.8.1 Suicide Ideators

Fremouw, de Perczel, and Ellis (1990) have outlined in great detail the principles for managing a suicide ideator. In particular, four principles are advocated: therapeutic activism, restoring hope, delaying of suicidal impulses, and environmental interventions. In regard to therapeutic activism, increasing the availability of the therapist during the suicidal crisis, working directly with the adolescent on problem solving and, when necessary, directing the adolescent and parents in how to manage a suicidal crisis are recommended (Fremouw et al., 1990). By actively working with the adolescent on how to solve problems and by taking one problem at time and generating solutions, the therapist can help to restore hope in what appears to be a hopeless situation. With adolescents, there are often acute conflicts with a family member, particularly parents. It is important for the therapist to actively assist in resolving the conflicts among family members at least for the short term and emphasize that the adolescent can make it through the current stress no matter how painful it seems. Environmental interventions are often necessary, such as working with the family on coping with adolescent suicidal behavior. It is important to involve parents in actively monitoring the adolescent's behavior. At the same time, removing pills and/or firearms from the home is very important. Finally, prior to the completion of a session, it is important for the therapist to work with the adolescent and their family on structuring the time between sessions to minimize the possibility of repeat attempts.

One particularly helpful problem-solving technique in managing suicide ideators has been described by Schneidman (1985). When confronted with a suicidal client, Schneidman suggests that the therapist and the patient write down a list of what might be done in the situation, including committing suicide. Then the therapist asks the patient to rank order the options from the most difficult and fatal (i.e., suicide) to the least difficult solutions. The therapist then begins to discuss the least difficult and painful solutions until the adolescent and the therapist find a solution, even a partial solution, that is acceptable to the adolescent. Schneidman contends that this approach has a calming effect on the adolescent because the therapist is helping to outline alternatives that the adolescent, due to cognitive restriction, was unable to generate. By discussing various options and their pros and the cons, progress is made because suicide is no longer the only solution.

5.18.8.2 Suicide Attempters

Adolescent suicide attempters have a significantly higher reattempt rate and risk of eventual completed suicide than adolescents who have not previously attempted suicide (Lewinsohn, Rohde, & Seeley, 1994). In 1997, there are no standard interventions for adolescent suicide attempters. Rather, intervention is typically developed on an individual basis, according to the symptoms occurring at the time of the suicide attempt.

The large majority of adolescents who attempt suicide can be treated effectively as outpatients. However, treatment noncompliance is particularly problematic among adolescent suicide attempters. Spirito et al. (1992) followed adolescent suicide attempters for three months after they were evaluated in an emergency department and discharged to outpatient care and found that 18% never attended an outpatient appointment, 11% went only once, 3% went twice, and 23% only went to three or four sessions. Even more striking are the results obtained by Trautman, Stewart, and Morishima (1993) indicating that adolescent suicide attempters drop out of outpatient psychotherapy at a much faster rate than nonsuicidal adolescents. Suicide attempters attended a median of three sessions before dropout, while nonsuicidal adolescents attended approximately 11 sessions.

This high rate of dropout of adolescent suicide attempters from psychotherapy suggests that clinicians who are evaluating adolescents immediately after a suicide attempt should take certain steps to ensure that outpatient therapy is accomplished following discharge. These include setting a specific date and time for the follow-up appointment; thoroughly explaining the reasons for continued outpatient care; explaining why therapy is important in the long run; and reviewing the likely goals and length of outpatient therapy (Jellinek, 1978). The adolescent's and their parents'/guardians' expectations for therapy should also be reviewed. Where resistance or hesitancy is noted, the clinician should address these problems. For example, if there is concern that the parents will not be able to get the adolescent to attend therapy, then problem-solving steps and potential alternatives to ensure the adolescent receives therapy should be reviewed with parents. In order to enhance compliance, it is preferable to discuss brief treatment (e.g., four sessions) with the option of continuation based on progress during that time. Asking adolescents and their parents to attend four treatment sessions before making a decision about further ongoing psychotherapy may be the most realistic approach to enhancing treatment compliance. Setting a specific time limit on psychotherapy has been shown to be more useful in ensuring treatment attendance than more open-ended treatment contracts (Sledge, Moras, Hartley, & Levine, 1990). As a final step, prior to discharge after a suicide attempt, the clinician might generate several potential stressors and ask how the adolescent would cope with those stressors if they were encountered. A directive, problem-solving approach will help the adolescent, particularly when discharged into a stressful situation. This type of "emergency plan" should focus on practical steps the adolescent should take if any suicidal feelings reemerge.

In developing a treatment plan for adolescent suicide attempters, therapists must rely on guidelines derived from either theoretical considerations or clinical care. Empirically-based guidelines specifically for suicide attempters are not available. In the initial therapy sessions with someone who has recently attempted suicide, a crisis intervention plan needs to be activated. The therapist must be continuously available, limit access to suicide methods, reestablish a no-suicide agreement if necessary, and ensure that the adolescent is not socially isolated, either from peers or from family. While managing the crisis using these guidelines, the therapist must also try to understand the reasons for the suicidal behavior and establish a working alliance with the adolescent in order to both reduce the risk of further suicidal behavior and maintain the adolescent in treatment (Berman & Jobes, 1991). Adolescents should be prepared, for example via role playing, for situations that may arise in the future that may elicit suicidal feelings. This approach can help prepare them for difficult circumstances as well as reduce the possibility of relapse.

The exact nature of treatment will vary according to the adolescent's presenting problems, the training and theoretical orientation of the therapist, and the adolescent's reactions to therapy. In general, however, there are a number of principles of psychotherapy with suicidal clients which should be observed no matter what the specific nature of treatment. In addition to those described above, these principles include maintaining an empathic but objective view of the suicidal adolescent, thoroughly understanding the motives of the suicidal behavior, continual reassessment and modification of hopelessness, assessment and modification of cognitive rigidity which impairs problem-solving, and education of the adolescent's family regarding suicide and suicidal behavior (Blumenthal & Kupfer, 1986). Once the crisis has passed, the feelings of depression and hopelessness associated with the suicidal behavior may begin to be addressed. The therapist can move from understanding the emotions that accompanied this suicidal behavior to specific problem-solving techniques that help to generate more effective ways of coping with life stressors as well as additional suicidal feelings that may arise in the future. In addition, after a short course of treatment, occasional booster sessions to help suicidal adolescents have some continuing contact with a therapist and review coping skills may be helpful in preventing additional suicidal behavior.

Although empirical data does not exist, several specific treatment modalities have been suggested for use with suicide attempters. First, cognitive therapy has been suggested as a possible treatment of choice for suicide attempters because it is highly structured and systematic. The cognitive distortions often evident in adolescent suicide attempters can be addressed directly using such a treatment approach. One goal of treatment is to help the adolescent become more aware of the relationships between their thoughts, emotions, and behaviors. Adolescents can begin to monitor their thoughts that lead to depressive cognitions and begin to challenge the validity of negative thought patterns (Bedrosian & Epstein, 1984). A number of cognitive techniques specifically useful for suicidal persons have been described (Freeman & Reinecke, 1993). These include reattribution of blame for the suicide precipitants, "decatastrophizing" (re-evaluating whether the adolescent is overestimating the "catastrophic" nature of their problems), making a list of pros and cons in regards to the circumstances precipitating the suicidal behavior, listing options and alternatives, "scaling" the severity of the events (e.g., from one to 100) to prevent the adolescent from conceptualizing everything as "black and white," and thought stopping.

One cognitive approach advocated by numerous clinicians as being helpful for suicidal patients is problem solving. As discussed above, therapy with suicidal adolescents often involves active teaching of new coping skills. Problem-solving therapy can help teach adolescents new ways of conceptualizing their problems and generating alternatives, because these adolescents often are unable to generate alternatives on their own. A description of the problem-solving approach with adolescent suicide attempters can be found in Brent and Kolko (1990). One focus of problem-solving is on interpersonal areas. Adolescents may at times use suicide attempts in a coercive manner for interpersonal gain. Thus, interpersonal problem solving to improve peer and family relationships can be effective. One behavioral approach to complement problem-solving efforts is to use social skills training for those adolescents who are particularly deficient in basic social skills which affect their peer relationships. Group therapy has been suggested as a potentially useful modality for treating adolescent suicide attempters. The group format can be helpful in teaching social skills and problem-solving skills as well as in reducing isolation and stigma associated with suicidal behavior.

Family therapy is often advocated as a means to deal with the intense conflict within families which can lead to maladaptive actions, including suicidal behavior (Richman, 1979). Family therapy itself helps to shift focus away from the individual adolescent, which in turn allows conflictual issues which have contributed to the development of the suicidal behavior to be investigated (Richman, 1979). Family therapy programs emphasizing communication and problem-solving skills that have been used with other high-risk adolescents (e.g., Robin, 1981) are applicable to work with families of adolescent suicide attempters. A six session cognitive-behavioral family therapy approach for adolescent suicide attempters has been described by Rotheram-Borus, Piacentini, Miller, Graae, and Castro-Blanco (1994). This treatment approach is very structured and includes a number of activities designed to create a positive atmosphere among family members, teach problem-solving strategies to all family members, shift the emphasis of the family to a concept of difficult situations as opposed to difficult family members (i.e., the adolescent), and enhance the confidence of the family members in mental health treatment. A number of cognitive techniques such as restructuring and reframing are used in this treatment approach as well as behavioral contracting, therapist modeling, and role playing. Outcome data are still being collected on this promising treatment approach.

5.18.8.3 Secondary Prevention of Suicidal Behavior in High-risk Groups

Another approach to reducing suicidal behavior is to study well-defined groups that are known to be at higher risk for suicide. These groups, such as runaways, Native Americans, adolescents who come from families with a history of suicidal behavior, adolescents who have been diagnosed with major depression, and surviving friends of adolescents who have completed suicide, often have significantly higher rates of suicidal behavior in their immediate or more distant history compared to other adolescents. Thus, interventions designed to help the overall functioning of high-risk groups may also have an effect on suicidal behavior. Unfortunately, few studies have specifically observed the effects of interventions on suicidal behavior. Rotheram-Borus and Bradley (1991) reported on an intervention with suicidal runaways seen in community shelters that resulted in reduced rates of suicide attempts when compared to the agency rate prior to implementation of the program.

Intervention programs specifically tailored for Native Americans must be based on an understanding of the cultural values and social factors that affect suicidal behaviors. For example, an intervention for Native American adolescents who attempted suicide and were living at a boarding school has been described (Ward, 1984). Tribal elders conducted group discussions about leadership in an attempt to reduce suicidal behavior. Interventions have also been conducted with Native Americans who have been incarcerated, in residential alcohol treatment programs, or school dropouts. However, these programs have not been empirically evaluated nor was suicidal behavior a specific outcome measure.

There have been several empirical treatment trials for other psychiatric disorders relevant to suicidal youth (e.g., depressed, substance abusing, and conduct disordered youth). However, suicidal behavior has not been typically examined at follow-up. It is important for researchers conducting interventions in these areas to investigate the effects of the interventions on suicidal behavior. Another area of treatment research that may have relevance for the treatment of adolescent suicidal behavior is cognitive-behavioral treatment of Borderline Personality Disorder. Among suicide attempters, especially adults, self-injury is common. Conversely, individuals diagnosed with Borderline Personality Disorder very often display suicidal behavior. Thus, for adolescent suicide attempters who demonstrate repetitive suicidal behaviors, treatment approaches advocated for Borderline Personality Disorder may be useful. However, these interventions have yet to be tested with adolescent populations.

5.18.8.4 Primary Prevention of Suicidal Behavior

Suicide awareness programs have been advocated as a way to help adolescents become more sensitive to identifying a peer in suicidal crisis and helping the suicidal teen get the professional assistance that is needed. Thus, these programs may promote effective coping strategies in adolescents. Despite their intuitive appeal, the efficacy of these suicide prevention programs is, at best, modest. Shaffer et al. (1990) studied the effects of three different school-based prevention programs and found a significant increase in the proportion of students who admitted to personal difficulties, suicidal thoughts, or the need for professional help. A study of a 14-hour experiential program in Israel (Orbach & Bar-Joseph, 1992) reported a mild reduction of suicidal feelings and hopelessness. Girls seemed to benefit more than boys, and those students with greater suicidal tendencies improved more than their peers. Other studies suggest that knowledge may improve from such educational efforts, but that attitudes about suicide and behavior are largely unaffected immediately following the education (Shaffer et al., 1990; Spirito, Overholser, Ashworth, Morgan, & Benedict-Drew, 1988) and at 18-month follow-up (Vieland, Whittle, Garland, Hicks, & Shaffer, 1991). Kalafat and Elias (1994) did find a positive change in attitudes towards help-seeking and intervening with trusted peers. However, suicide awareness programs may be viewed most negatively by adolescents who have made a prior suicide attempt (Shaffer, Garland, Underwood, & Whittle, 1987). Another intervention program helped high-risk students but was no more effective than an assessment-only condition (Shaffer et al., 1987).

To be most helpful, specific programs targeting particular groups of adolescents (e.g., males vs. females, attempters vs. nonattempters, depressed vs. nondepressed) may need to be developed (Overholser, Hemstreet, Spirito, & Vyse, 1989). Alternatively, prevention programs that focus on general mental health interventions such as social skills training, problem solving, crisis management, and anger control may be more beneficial than programs focusing only on suicide (Berman & Jobes, 1991). Shaffer et al. (1988) reviewed the literature and found that only a minority of school-based programs were designed to improve teenagers' coping skills by training in stress management. For adolescents with no psychiatric disorders, the examination of maladaptive methods used when faced with difficult life situations and establishment of programs to improve problem-solving and coping skills may be helpful (Hawton, 1986).

5.18.9 PSYCHOPHARMACOLOGY

Psychotropic medication should be considered as a possible treatment modality for all adolescent suicidal behavior. At times, medication is clearly indicated for a particular psychiatric disorder such as major depression or bipolar affective disorder. At other times, medications may be helpful in reducing acute symptomology which, in turn, may improve the adolescent's acceptance and receptivity to other psychotherapies (Berman & Jobes, 1991). Of course, prescribing medications for suicidal adolescents requires careful monitoring, particularly for those adolescents thought to be at higher risk for reattempts. However, there have not been specific drug trials that have examined effects of medication on the functioning of suicide attempters.

5.18.10 EFFICACY/EFFECTIVENESS OF TREATMENTS

There have been several empirical treatment studies examining different therapy approaches with adult suicide attempters. Outpatient therapy and outreach following a suicide attempt have been found to result in fewer reattempts (Welu, 1977). Other studies have not found differences in reattempts after outreach programs (Gibbons, Butler, Urwin, & Gibbons, 1978). Lieberman and Eckman (1981) compared behavior therapy to insight-oriented psychotherapy in adult psychiatric inpatients following a suicide attempt. There was no difference in the reattempt rate between the two groups, although the behavior therapy group had significantly less suicide ideation and fewer suicide threats. In a 1990 study, cognitive-behavioral treatment was found to be slightly superior to standard care in reducing repeat attempts during a 6-month treatment course of adult suicide attempters (Salkovkis, Atha, & Storer, 1990). A similar treatment protocol found a reduction in hopelessness at follow-up, but no differences between groups that received a cognitive-behavioral treatment and nondirective treatment on suicidal ideation and intention at follow-up (Patsiokas & Clum, 1985).

There have been very few empirical trials of treatment with adolescent suicide attempters. Deykin, Hsieh, Joshi, and McNamara (1986) conducted randomized outreach intervention with adolescent suicide attempters after they had been seen in an emergency department. Social work follow-up with the families was offered to help the adolescents keep their follow-up appointments, to provide support, and explore additional services for the adolescents. The suicide attempters in the experimental group were twice as likely as controls to comply with recommendations for other services. In addition, they had fewer visits in the emergency department for suicidal behavior and follow-up. However, there was no difference in repeat suicide attempts after adjusting for prior history of suicidal behavior. The use of group therapy with adolescents with a history of suicidal behavior has also been described. Ross and Motto (1984) conducted a 35 session group over a 40-week period. The goals of the group treatment included providing a peer and professional support system, teaching ways to deal with stress and dysphoria, improving self-image, and teaching group members how to befriend others. At two-year follow-up, none of the 14 members of the group contacted had made further suicide attempts.

5.18.11 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

The significant international differences in adolescent suicide rates are not fully understood and need to be investigated further. Attempted suicide in adolescence is a much more frequent phenomena than completed suicide, but the epidemiologic knowledge base on attempted suicide is quite limited, both nationally and internationally. The World Health Organization is attempting to establish a more systematic data collection method regarding suicide attempts, and similar systems have been advocated by the Centers for Disease Control. Greater knowledge of the epidemiology of attempted suicide is important because it offers the potential to reach adolescents demonstrating suicidal behavior early in the spectrum of suicidality, and thus potentially having a greater effect on future lethal suicidal behavior.

The risk factors for completed suicide have been investigated fairly well. One area that deserves more attention is the influence of family and peers on adolescent suicidal behavior. A greater understanding in this area may lead to more primary prevention strategies and a subsequent reduction in suicide and suicidal behavior. As reviewed above, it is clear that the majority of adolescents who complete suicide have a diagnosable psychiatric disorder. The mental health field is in the early stages of its development of well-tested and effective treatment programs for the major psychiatric disorders related to adolescent suicidal behaviors, such as depression and substance use. Both psychosocial and psychopharmacologic treatments used for these psychiatric disorders are adopted from adult studies. As the mental health field advances and conducts more empirical research specifically looking at child and adolescent psychiatric disorders, the benefits of learning how to treat these disorders effectively will undoubtedly apply not only to the major symptoms of the disorder, but also to the associated symptoms such as suicidality.

Specific interventions designed to reduce adolescent suicidal behavior are few and far between. Although suicide attempters are a high-risk group in need of treatment, treatment proves particularly difficult with this population. Not only is treatment compliance a problem, but there is no such thing as standard psychotherapeutic care for a suicide attempter. This is probably due to the heterogeneous nature of the suicide attempters in this population. However, it would appear that future research should focus on testing at least some minimum guidelines for the treatment of suicide attempters in the months following a suicide attempt.

Secondary prevention of suicidal behavior is another area where clinical research is important. Intervention programs that might be tested in future research include services for populations at high risk for suicide, such as runaways, depressed adolescents, friends of adolescents who have completed suicide, and adolescents in families with a strong history of suicidal behavior. Primary prevention programs aimed at improving self-esteem and problem-solving skills in schools have the potential to improve mental health and possibly reduce the onset of suicidal behavior.

It is clear that approaches to both intervention and prevention of suicidal behavior among adolescents must incorporate a number of different perspectives. Such multi-pronged approaches will be necessary to advance clinical treatment and research in this area. Collaboration between researchers and clinicians in the field of adolescent suicidal behavior may improve our understanding of suicidal behavior and the interaction of environmental, psychological, and biological factors that need to be addressed in these adolescents.

5.18.12 SUMMARY

This chapter has provided an overview of the factors which contribute to both completed suicide in adolescence and/or attempted suicide in adolescence. Females attempt suicide much more frequently than males, while males complete suicide much more frequently than females. Caucasians attempt and complete suicide more often than members of other races. The most common psychiatric disorders implicated in suicide are depression, conduct disorder, and substance abuse. Comorbidity is often evident. Adolescents who have made a prior suicide attempt are at much greater risk for eventual completed suicide than those adolescents without a prior attempt. Use of a firearm in a suicide attempt greatly increases the risk of completing suicide. Prior exposure to suicidal behavior also increases the risk of suicidal behavior.

For adolescents who attempt suicide, a model for comprehensive assessment was presented. This model includes emotional (anger, depression, and hopelessness), cognitive (attributional style, problem solving, and perfectionism), and family factors. Treatments for suicidal behavior are in their infancy. Efficacy of the existing treatments has not been established. Interventions that focus on problem solving and affect management appear promising and in need of future research.

5.18.13 REFERENCES

Adam, K., Sheldon-Keller, A., & West, M. (1996). Attachment, organization and history of suicidal behavior in clinical adolescents. *Journal of Consulting and Clinical Psychology, 64,* 264–272.

Adams, D., Overholser, J., & Spirito, A. (1994). Stressful life events related to adolescent suicide attempts. *Canadian Journal of Psychiatry, 39,* 43–48.

Adams, D., Overholser, J., & Lehnert, R. (1994). Perceived family functioning and adolescent suicidal behavior. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 498–507.

Adcock, A., Nagy, S., & Simpson, J. (1991). Selected risk factors in adolescent suicide attempts. *Adolescence, 26,* 818–828.

Apter, A., Bleich, A., King, R., Kron, S., Fluch, A., Kotter, M., & Cohen, D. J. (1993). Death without warning? A clinical postmortem study of suicide in 43 Israeli adolescent males. *Archives of General Psychiatry, 50,* 138–142.

Asarnow, J., Carlson, G., & Guthrie, D. (1987). Coping strategies, self-perceptions, hopelessness, and perceived family environments in depressed and suicidal children. *Journal of Consulting and Clinical Psychology, 53,* 361–366.

Beck, A. T., Schuyler, D., & Herman, I. (1974). Development of suicidal intent scales. In A. T. Beck, H. L. P. Resnik, & D. J. Lettiere (Eds.), *The prediction of suicide* (pp. 45–56). Philadelphia: Charles Press.

Bedrosian, R., & Epstein, N. (1984). Cognitive therapy of depressed and suicidal adolescents. In H. Sucak, A. Ford, & N. Rushforth (Eds.), *Suicide in the young* (pp. 345–366). Boston: L. John Wright.

Berman, A. L., & Jobes, P. A. (1991). The treatment of the suicidal adolescent. *Adolescent suicide. Assessment and intervention* (pp. 163–225). Washington, DC: American Psychological Association.

Birkhead, G., Galvin, V., Meehan, P., O'Carroll, P., & Mercy, J. (1993). The emergency department in the surveillance of attempted suicide: Findings and methodological considerations. *Public Health Reports, 108,* 323–331.

Blumenthal, S. J., & Kupfer, D. J. (1986). Generalizable treatment strategies for suicidal behavior. *Annals of the New York Academy of Science, 487,* 327–340.

Brent, D. (1987). Correlates of the medical lethality of suicide attempts in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 87–89.

Brent, D., & Kolko, D. (1990). The assessment and treatment of children and adolescents at risk for suicide. In S. Blumenthal & D. Kupfer (Eds.), *Suicide over the life cycle: Risk factors, assessment, and treatment of suicidal patients* (pp. 253–302). Washington, DC: American Psychiatric Press.

Brent, D., Kolko, D., Allan, N., & Brown, R. (1990). Suicidality in affectively disordered adolescent inpatients. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 586–593.

Brent, D., Moritz, G., Bridge, J., Perper, J., & Canobbio, R. (1996). Long-term impact of exposure to suicide: A three-year controlled follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 646–653.

Brent, D., Perper, J., Allman, C., Moritz, G., Wartella, M., & Zelenak, J. (1991). The presence and availability of firearms in the homes of adolescent suicides: A case-control study. *Journal of the American Medical Association, 266,* 2989–2995.

Brent, P. A., Perper, J. A., Goldstein, C. E., Kolko, D. J., Allan, M. J., Allman, C. J., & Zelenak, J. P. (1988). Risk factors for adolescent suicide: A comparison of adolescent suicide victims with suicidal inpatients. *Archives of General Psychiatry, 45,* 581–588.

Brent, P. A., Perper, J. A., Moritz, G., Allman, C., Shweers, J., Roth, C., Balach, L., Canobbio, R., & Liotus, L. (1993). Psychiatric sequelae to the loss of an adolescent peer to suicide. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 509–517.

Brent, P. A., Perper, J., Moritz, G., Baugher, N., & Allman, C. (1993). Suicide in adolescents with no apparent psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 494–500.

Brent, D., Peyser, J., Moritz, G., Liotus, L., Shweers, J., & Roth, C. (1994). Suicide in affectively ill adolescents: A case-control study. *Acta Psychiatrica Scandinavica, 89,* 52–58.

Brown, L., Overholser, J., Spirito, A., & Fritz, G. (1991). The correlates of planning in adolescent suicide attempts. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 95–99.

Cairns, R. B., Peterson, G., & Neckerman, H. J. (1988). Suicidal behavior in aggressive adolescents. *Journal of Clinical Child Psychology, 17,* 298–309.

Carlson, G. A., Asarnow, J. R., & Orbach, I. (1987). Developmental aspects of suicidal behavior in children. *Journal of the American Academy of Child and Adolescent Psychiatry, 26*(2), 186–192.

Centers for Disease Control. (1988) CDC recommendation for a community plan for the prevention and containment of suicide attempts. *Morbidity and Mortality Weekly Report, 37* (Suppl. 5–6), 1–12.

Centers for Disease Control (1995a). Suicide among children, adolescents, and young adults—United States, 1980–1992. *Morbidity and Mortality Weekly Report, 44,* 289–291.

Centers for Disease Control. (1995b). Youth Risk Behavior Surveillance—United States, 1993. *Morbidity and Mortality Weekly Reports, 44,* SS-1.

Choquet, M., & Menke, H. (1989). Suicidal thoughts during early adolescence: Prevalence, associated troubles and help-seeking behavior. *Acta Psychiatrica Scandinavia, 81,* 170–177.

Christoffel, K. K., Marcus, D., Sagerman, S., & Bennett, S. (1988). Adolescent suicide and suicide attempts: A population study. *Pediatric Emergency Care, 4,* 32–40.

Cole, D. A. (1989). Psychopathology of adolescent suicide: hopelessness, coping beliefs, and depression. *Journal of Abnormal Psychology, 98,* 248–255.

Davidson, L. (1989). Suicide clusters and youth. In C. Pfeffer (Ed.), *Suicide among youth: Perspectives on risk and prevention* (pp. 83–99). Washington, DC: American Psychiatric Press.

Davidson, L., & Gould, M. S. (1989). Contagion as a risk factor for youth suicide. *Report of the Secretary's Task Force on Youth Suicide: Vol. 2. Risk factors for youth suicide* (DHHS Publication No. (ADM) 89–1622, pp. 88–109). Washington, DC: US Government Printing Office.

deWilde, E., Kienhorst, C., Diekstra, R., & Wolters, W. (1992). The relationship between adolescent suicidal behavior and life events in childhood and adolescence. *American Journal of Psychiatry, 149,* 45–51.

deWilde, E., Kienhorst, C., Diekstra, R., & Wolters, W. (1993). The specificity of psychological characteristics of adolescent suicide attempters. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 51–59.

Deykin, E. Y., Alpert, J. J., & McNamara, J. J. (1985). A pilot study of the effect of exposure to child abuse or neglect on adolescent suicial behavior. *American Journal of Psychiatry, 142,* 1299–1303.

Diekstra, R. F. (1989). Suicidal behavior in adolescents and young adults: The international picture. *Crisis, 10,* 16–35.

Diekstra, R. F. (1993). The epidemiology of suicide and parasuicide. *Acta Psychiatrica Scandinavia, 371*(Suppl.), 9–20.

Diekstra, R. F., & Golbinat, W. (1993). The epidemilogy of suicidal behavior: A review of three continents. *World Health Statistics Quarterly, 46,* 52–68.

Eggert, L., Thompson, E., Herting, J., & Nicholas, L. (1995). Reducing suicide potential among high-risk youth: Tests of a school-based prevention program. *Suicide and Life-Threatening Behavior, 25,* 276–296.

Freeman, A., & Reinecke, M. (1993). *Cognitive therapy of suicidal behavior: A manual for treatment.* New York: Springer.

Fremouw, W., de Perczel, M., & Ellis, T. (1990). *Suicide risk: Assessment and response guidelines.* New York: Pergamon.

Garfinkel, B. D., Froese, A., & Hood, J. (1982). Suicide attempts in children and adolescents. *American Journal of Psychiatry, 139,* 1257–1261.

Garnefski, N., Dieksta, R. F., & de Heus, P. (1992). A population-based survey of the characteristics of high school students with and without a history of suicidal behavior. *Acta Psychiatria Scandinavica, 86,* 189–196.

Garrison, C., Jackson, K., Addy, C., McKeown, R., & Waller, J. (1991). Suicidal behaviors in young adolescents. *American Journal of Epidemiology, 133,* 1005–1014.

Gibbons, J. S., Butler, J., Urwin, P., & Gibbons, M. (1978). Evaluation of a social work service for self-poisoning patients. *British Journal of Psychiatry, 133,* 11–118.

Gibbs, J. T., & Hines, A. M. (1989). Factors related to sex differences in suicidal behavior among black youth: Implications for intervention and research. *Journal of Adolescent Research, 4,* 152–172.

Gispert, M., Wheeler, K., Marsh, L., & Davis, M. S. (1985). Suicidal adolescents: Factors in evaluation. *Adolescence, 20,* 753–762.

Goldacre, M., & Hawton, K. (1985). Repetition of self-poisoning and subsequent death in adolescents who take overdoses. *British Journal of Psychiatry, 146,* 395–398.

Gould, M. S., & Shaffer, D. (1986). The impact of suicide in television movies: Evidence of imitation. *New England Journal of Medicine, 315,* 690–694.

Hawton, K. (1986). *Suicide and attempted suicide among children and adolescents.* Beverly Hills, CA: Sage.

Hawton, K., & Fagg, J. (1992). Deliberate self-poisoning and self-injury in adolescents: A study of characteristics and trends in Oxford, 1976–1989. *British Journal of Psychiatry, 161,* 816–823.

Hawton, K., O'Grady, J., Osborn, M., & Cole, D. (1982). Adolescents who take overdoses: Their characteristics, problems, and contacts with helping agencies. *British Journal of Psychiatry, 140,* 118–123.

Hazell, P., & Lewin, T. (1993). Friends of adolescent suicide attempters and completers. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 76–81.

Hewitt, P. L., Flett, G. L., & Turnball-Donovan W. (1992). Perfectionism and suicide potential. *British Journal of Clinical Psychology, 31,* 181–190.

Hewitt, P. L., Newton, J., Flett, G. L., & Callander, L. (1997). Perfectionism and suicide ideation in adolescnt psychiatric patients. *Journal of Adnormal Child Psychology, 25,* 95–101.

Hibbard, R. A., Ingersoll, G. M., & Orr, D. P. (1990). Behavioral risk, emotional risk, and child abuse among adolescents in a nonclinical setting. *Pediatrics, 86,* 896–901.

Hlady, W. G., & Middaugh, J. P. (1988) Suicides in Alaska: Firearms and alcohol. *American Journal of Public Health, 78,* 179–180.

Holinger, P. C. (1990). The causes, impact, and preventability of childhood injuries in the United States. *American Journal of Diseases of Children, 144,* 670–676.

Holinger, P. C., & Offer, D. (1987). Suicide and homicide in the United States: An epidemiologic study of violent death, population changes, and the potential for prediction. *American Journal of Psychiatry, 144,* 215–219.

Jacobson, L., Rabinowitz, I., Popper, M., Solomon, R., Sokol, M., & Pfeffer, C. (1995). Interviewing prepubertal children about suicidal ideation and behavior. In M. E. Hertzig & E. A. Farber (Eds.), *Annual progress in child psychiatry and development* (pp. 295–321). New York: Brunner/Mazel.

Jellinek, M. (1978). Referrals from a psychiatric emergency room. Relationship of compliance to demographic and interview variables. *American Journal of Psychiatry, 135,* 209–213.

Joffe, R. T., & Offord, D. R. (1983). Suicidal behavior in childhood. *Canadian Journal of Psychiatry, 28,* 57–63.

Kalafat, J., & Elias, M. (1994). An evaluation of school-based suicide awareness intervention. *Suicide and Life-Threatening Behavior, 24,* 224–233.

King, C., Raskin, A., Gadowski, C., Butkus, M., & Opipari, L. (1990). Psychosocial factors associated with urban adolescent female suicide attempts. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 289–294.

Lehnert, R., Overholser, J., & Spirito, A. (1994). Internalized and externalized anger in adolescent suicide attempters. *Journal of Adolescent Researcher, 9,* 105–119.

Levy, S., Jurkovic, G., & Spirito, A. (1995). A multisystem analysis of adolescent suicide attempts. *Journal of Abnormal Child Psychology, 23,* 221–234.

Lewinsohn, P., Clarke, G., Hops, H., & Andrews, J. (1990). Cognitive-behavioral treatment for depressed adolescents. *Behavior Therapy, 21,* 385–401.

Lewinsohn, P., Rohde, P., & Seeley, J. (1993). Psychosocial characteristics of adolescents with a history of suicide attempt. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 60–68.

Lewinsohn, P., Rohde, P., & Seeley, J. (1994). Psychosocial risk factors for future adolescent suicide attempts. *Journal of Consulting & Clinical Psychology, 62,* 297–305.

Lewinsohn, P., Rohde, P., & Seeley, J. (1996). Adolescent suicidal ideation and attempts: Prevelance, risk factors, and clinical implications. *Clinical Psychology: Sense and Practice, 3,* 25–46.

Lieberman, R., & Eckman, T. (1981). Behavior therapy vs. insight oriented therapy for repeated suicide attempters. *Archives of General Psychiatry, 38,* 1126–1130.

Males, M. (1991). Teen suicide and changing cause of death certification, 1953–1987. *Suicide and Life-Threatening Behavior, 21,* 245–259.

Mann, J., McBride, P., Brown, R., Linnoila, M., Leon, A., Demeo, M., Mieczkowski, T., Mysers, J., & Stanley, M. (1992). Relationship between central and perceptual serotonin indexes in depressed and suicidal psychiatric inpatients. *Archives of General Psychiatry, 49,* 442–446.

Marks, P. A., & Haller, D. L. (1977). Now I lay me down for keeps: A study of adolescent suicide attempts. *Journal of Clinical Psychology, 33,* 390–400.

Martin, G., Rozanes, P., Pearce, C., & Allison., S. (1995). Adolescent suicide, depression, and family dysfunction. *Acta Psychiatrica Scandinavica, 92,* 336–344.

Marttunen, M., Aro, H., Henriksson, M., & Lonnqvist, T. (1991). Mental disorders in adolescent suicide. *Archives of General Psychiatry, 48,* 834–839.

McDowell, E., & Stillion, J. (1994). Suicide across the phases of life. In G. G. Noam & S. Borst (Eds.), *Children, Youth, and Suicide: Developmental Perspective* (pp. 7–22). San Francisco: Jossey-Bass.

McIntosh, J. (1983–84). Suicide among Native Americans: Further tribal data and considerations. *Journal of Death and Dying, 14,* 215–229.

Mitchell, M., & Rosenthal, D. (1992). Suicidal adolescents: Family dynamics and the effects of lethality and hopelessness. *Journal of Youth and Adolescence, 21,* 23–33.

Myers, K., MacCauley, E., Calderon, R., & Treder, R. (1991). The 3-year longitudinal course of suicidality and predictive factors for subsequent suicidality in youths with major depressive disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 804–810.

National Center for Health Statistics. (1993). Advance report of final mortality statistics, 1990. *Monthly Vital Statistics Report, 41, 7,* 1–44.

Nordstrom, P., Samuelsson, M., Asberg, M., Traksman-Bendz, L., Aberg-Wistedt, A., Nordin, C., & Bertilson, L. (1994). CSF 5-HIAA predicts suicide risk after attempted suicide. *Suicide and Life-Threatening Behavior, 24,* 1–9.

Orbach, I., & Bar-Joseph, H. (1992). The impact of a suicide prevention program for adolescents on suicidal tendencies, hopelessness, ego identity, and coping. *Suicide and Life-Threatening Behavior, 23,* 120–129.

Orbach, I., Rosenheim, E., & Mary, E. (1987). Some aspects of cognitive funcitoning in suicidal children. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 181–185.

Overholser, J., Adams, D., Lehnert, K., & Brinkman, D. (1995). Self-esteem deficits and suicidal tendencies among adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 919–928.

Overholser, J. C., Hemstreet, A., Spirito, A., & Vyse, S. (1989). Suicide awareness programs in the schools. Effects of gender and personal experience. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 925–930.

Otto, U. (1972). Suicidal acts by children and adolescents. *Acta Pediatrica Scandinavica, 232*(Suppl.), 7–123.

Patsiokas, A., & Clum, G. (1985). Effects of psychotherapeutic strategies in the treatment of suicide attempters. *Psychotherapy, 22,* 281–290.

Pfeffer, C., Conte, H., Plutchik, R., & Jerrett, I. (1980). Suicidal behavior in latency-age children: An outpatient population. *Journal of the American Academy of Child Psychiatry, 19,* 703–710.

Pfeffer, C., Newcorn, J., Kaplan, G., Mizruchi, M., & Plutchik, R. (1988). Suicidal behavior in adolescent psychiatric inpatients. *Journal of the American Academy of Child Psychiatry, 27,* 357–361.

Pfeffer, C. R., Plutchik, R. L., & Mizrucki, M. S. (1983). Suicidal and assaultive behavior in children: Classification, measurement, and interrelations. *American Journal of Psychiatry, 140,* 154–157.

Philips, D. P., & Carstensen, L. L. (1986). Clustering of teenage suicides after television news about suicides. *New England Journal of Medicine, 315,* 685–689.

Philips, D. P., & Paight, D. (1987). The impact of televised movies about suicide; A replicative study. *New England Journal of Medicine, 317,* 809–811.

Platt, S., Bille-Brahe, U., Kerkhof, A., Schmidtke, A., Bjerke, T., Crepet, P., DeLeu, D., Haring, C., Lonnqvist, J., Michel, K., Philippe, A., Pommereau. X., Querejeta, I., Salander-Renberg, E., Temesuary, B., Wasserman, D., & Sampaio-Fario, J. (1992). Parasuicide in Europe: The WHO/EURO multicentre study on parasuicide. I. Introduction and preliminary analysis for 1989. *Acta Psychiatrica Scandinavica, 85,* 97–104.

Reynolds, W. (1987). *Suicide Ideation Questionnaire.* Odessa, FL: Psychological Assesment Resources.

Richman, J. (1979). The family therapy of attempted suicide. *Family Process, 18,* 131–142.

Rifai, A., Reynolds, C., & Mann, J. J. (1992). Biology of elderly suicide. *Suicide and Life-Threatening Behavior, 22,* 48–61.

Riggs, S., Alario, A. J., & McHorney, C. (1990). Health risk behaviors and attempted suicide in adolescents who report prior maltreatment. *Journal of Pediatrics, 116,* 815–820.

Robin, A. L. (1981). A controlled evaluation of problem-solving communication training with parent–adolescent. *Behavior Therapy, 12,* 593–609.

Robbins, D., & Alessi, M. (1985). Depressive symptoms and suicidal behavior in adolescents. *American Journal of Psychiatry, 142,* 588–592.

Ross, C., & Motto, J. (1984). Group counseling for suicidal adolescents. In H. Sudak, A. Ford, & N. Rushforth (Eds.), *Suicide in the young*. Littleton, MA: John Wright PSG.

Rotheram-Borus, M. J., & Bradley, J. (1991). Triage model for suicidal runaways. *American Journal of Orthopsychiatry, 61,* 122–127.

Rotheram-Borus, M. J., Piacentini, J., Miller, S., Graae, F., & Castro-Blanco, D. (1994). Brief cognitive behavioral treatment for adolescent suicide attempters and their families. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 508–517.

Rotheram-Borus, M. J., Trautman, P., Dopkins, S., & Shrout, P. (1990). Cognitive style and pleasant activities among female adolescent suicide attempters. *Journal of Consulting and Clinical Psychology, 58,* 553–561.

Roy, A. (1994). Recent biological studies on suicide. *Suicide & Life-Threatening Behavior, 24,* 10–14.

Roy, A., DeJong, J., & Linnoila, M. (1989). Cerebrospinal fluid momoamine metabolites and suicidal behavior in depressed patients. *Archives of General Psychiatry, 46,* 609–612.

Rubenstein, J. L., Heeren, T., Housman, D., Rubin, C., & Stechler, G. (1989). Suicidal behavior in "normal" adolescents: Risk and protective factors. *American Journal of Orthopsychiatry, 59*(1), 59–71.

Runeson, B., & Beskow, J. (1991). Borderline personality disorder in youth Swedish suicides. *Journal of Nervous and Mental Disease, 179,* 153–156.

Sadowski, C., & Kelley, M. L. (1993). Social problem solving in suicidal adolescents. *Journal of Consulting and Clinical Psychology, 61*(1), 121–127.

Salkovkis, P., Atha, C., & Storer, D. (1990). Cognitive-behavioral problem-solving in the treatment of petients who repeatedly attempt suicide. *British Journal of Psychiatry, 157,* 871–876.

Schneidman, E. (1985). *Definition of suicide*. New York: Wiley.

Shaffer, D. (1988). The epidemiology of teen suicide: An examination of risk factors. *Journal of Clinical Psychiatry, 49,* 36–41.

Shaffer, D., Garland, A., Gould, M., Fisher, P., & Trautman, P. (1988). Preventing teenage suicide: A critical review. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 675–687.

Shaffer, D., Garland, A., Underwood, M., & Whittle, B. (1987). *An evaluation of three youth suicide prevention programs in New Jersey*. Report prepared for the New Jersey State Department of Health and Human Services.

Shaffer, D., Gould, M. Fisher, P., Trautman, P., Moreau, D., Kleinman, M., & Flory, M. (1996). Psychiatric diagnosis in child and adolescent suicide. *Archives of General Psychiatry, 53,* 339–348.

Shaffer, D., Gould, M., & Hicks, R. (1994). Worsening suicide rate in black teenagers. *American Journal of Psychiatry, 131,* 1810–1812.

Shaffer, D., Vieland, V., Garland, A., Rojas, M., Underwood, M., & Busner, C. (1990). Adolescent suicide attempters: Response to suicide prevention programs. *Journal of the American Medical Association, 264,* 3151–3155.

Shaffi, M., Carrigan, S., Whittinghill, J., & Derrick, A. (1985). Psychological autopsy of completed suicide in children and adolescents. *American Journal of Psychiatry, 142,* 1061–1064.

Shaunesey, K., Cohen, J. L., Plummer, B., & Berman, A. (1993). Suicidality in hospitalized adolescents: Relationship to prior abuse. *American Journal of Orthopsychiatry, 63,* 113–119.

Sigurdson, E., Staley, D., Matas, M., Hildahl, K., & Squair, K. (1994). A five year review of youth suicide in Manitoba. *Canadian Journal of Psychiatry, 39,* 397–403.

Slap, G. B., Vorters, D. F., Chaudhuri, S., & Centor, R. M. (1989). Risk factors for attempted suicide during adolescence. *Pediatrics, 84,* 762–772.

Sledge, W., Moras, K., Hartley, D., & Levine, M. (1990). Effect of time-limited psychotherapy on patient dropout rates. *American Journal of Psychiatry, 147,* 1341–1347.

Snyder, R. (1981). The first offender program: Children and our future. In I. N. Berlin (Ed.), *The International Year of the Child, 1979–1980*. Albuquerque, NM: University of New Mexico Press.

Sorenson, S. B., & Rutter, C. M. (1991). Transgenerational patterns of suicide attempt. *Journal of Consulting and Clinical Psychology, 59,* 861–866.

Spirito, A., Bond, A., Kurkjian, J., Devost, L., Bosworth, T., & Brown, L. (1993). Gender differences among adolescent suicide attempters. *Crisis: International Journal of Suicide and Crisis Studies, 62,* 464–468.

Spirito, A., Brown, J., Overholser, J., & Fritz, G. (1989). Attempted suicide in adolescence. A review and critique of the literature. *Clinical Psychology Review, 9,* 335–363.

Spirito, A., Francis, G., Overholser, J., & Frank, N. (1996). Coping, depression, and adolescent suicide attempts. *Journal of Clinical Child Psychology, 25,* 147–155.

Spirito, A., Overholser, J., Ashworth, S., Morgan, J., & Benedict-Drew, C. (1988). Evaluation of a suicide awareness curriculum for high school students. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 705–711.

Spirito, A., Overholser, J., & Hart, K. (1991). Cognitive characteristics of adolescent suicide attempters. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 604–608.

Spirito, A., Overholser, J., & Stark, L. (1989). Common problems and coping strategies, II: Findings with adolescent suicide attempters. *Journal of Abnormal Child Psychology, 17,* 213–221.

Spirito, A., Overholser, J., & Vinnick, L. (1995). Adolescent Suicide Attempters in General Hospitals: Psychological Evaluation and Disposition Handling. In J. L. Wallander & L. J. Siegel (Eds.), *Adolescent Health Problems: Behavioral Perspectives* (pp. 97–116). New York: Guilford Press.

Spirito, A., Plummer, B., Gispert, M., Levy, S., Kurkjian, J., Lewander, W., Hagberg, S., & Devost, L. (1992). Adolescent suicide attempts: Outcomes at follow-up. *American Journal of Orthopsychiatry, 62,* 464–468.

Spirito, A., Sterling, C., Donaldson, D., & Arrigan, M. (1996). Factor analysis of the Suicide Intent Scale with adolescent suicide attempters. *Journal of Personality Assessment, 67,* 90–101.

Spirito, A., Williams, C., Stark, L. J., & Hart, K. (1988). The Hopelessness Scale for Children: Psychometric properties and clinical utility with normal and emotionally disturbed adolescents. *Journal of Abnormal Child Psychology, 16,* 445–458.

Traskman-Bandz, L., Alling, C., Alsen, M., Regnell, G., Simonsson, P., & Ohman, R. (1993). The role of monomines in suicidal behavior. *Acta Psychiatrica Scandinavica Suppl. 371,* 45–47.

Trautman, P. D., Stewart, N., & Morishima, A. (1993). Are adolescent suicide attempters noncompliant with outpatient care? *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 89–94.

US Bureau of the Census (1990). *Statistical Abstract of the United States: 1989* (110th ed.) Washington, DC: Government Printing Office.

US Congress, Office of Technology Assessment. (1990).

Indian adolescent mental health (OTA-H 446). Washington, DC: Government Printing Office.

Vieland, V., Whittle, B., Garland, A., Hicks, R., & Shaffer, D. (1991). The impact of curriculum based suicide prevention programs for teenagers: An 18-month follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 811–815.

Wagner, B., Cole, R., & Schwartzman, P. (1995). Psychosocial correlates of suicide attempts among junior and senior high school youth. *Suicide and Life-Threatening Behavior, 25,* 358–372.

Ward, J. A. (1984). Preventive implications of a Native Indian mental health program: Focus on suicide and violent death. *Journal of Preventive Psychiatry, 2,* 371–385.

Weinstein, M. C., & Saturno, P. J. (1989). Economic impact of youth suicide and suicide attempts. *Report of the Secretary's Task Force of Youth Suicide: Vol. 4. Strategies for the prevention of youth suicide* (DHHS Publication No. (ADM) 89–1624, pp. 82–93). Washington, DC: Government Printing Office.

Welu, T. C. (1977). A follow-up program for suicide attempters. *Suicide and Life-Threatening Behavior, 7,* 17–30.

Wilson, K., Seltzer, J., Bergman, J., Kral, M., Inayatullah, M., & Elliott, C. (1995). Problem-solving, stress, and coping in adolescent suicide attempters. *Suicide and Life-Threatening Behavior, 25,* 241–252.

Withers, L. E., & Kaplan, D. W. (1987). Adolescents who attempt suicide. A retrospective clinical chart review of hospitalized patients. *Professional Psychology: Research and Practice, 18,* 391–393.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.19

Attention-deficit Hyperactivity Disorder in Children and Adolescents

C. KEITH CONNERS
Duke University Medical Center, Durham, NC, USA
and
DREW ERHARDT
Pepperdine University, Malibu, CA, USA

5.19.1 INTRODUCTION 488

5.19.2 PHENOMENOLOGY 490
- *5.19.2.1 Core Features* 490
  - *5.19.2.1.1 Inattention* 491
  - *5.19.2.1.2 Hyperactivity* 491
  - *5.19.2.1.3 Impulsivity* 492
- *5.19.2.2 Associated Features and Comorbidity* 492
- *5.19.2.3 Variations Related to Gender and Development* 494
  - *5.19.2.3.1 Gender* 494
  - *5.19.2.3.2 Developmental variation* 494

5.19.3 DIAGNOSIS 495

5.19.4 PREVALENCE 497
- *5.19.4.1 Caseness Criteria* 498
- *5.19.4.2 National and Cultural Factors* 498
- *5.19.4.3 Comorbidity* 499
- *5.19.4.4 Developmental Factors* 499

5.19.5 ETIOLOGY AND RISK FACTORS 499
- *5.19.5.1 Genetic Risk Factors* 500
- *5.19.5.2 Medical and Environmental Risk Factors* 501
- *5.19.5.3 Metabolic and Nutritional Factors* 501
- *5.19.5.4 Lead* 502
- *5.19.5.5 Vitamins and General Nutrition* 502
- *5.19.5.6 Temperament as a Risk Factor for ADHD* 503
- *5.19.5.7 Overview of Etiology and Risk Factors* 503

5.19.6 DEVELOPMENTAL COURSE AND OUTCOME 504

5.19.7 CONCEPTUALIZATION AND CLINICAL FORMULATION 504
- *5.19.7.1 Multimodal Assessment* 504
- *5.19.7.2 Treatment Formulation* 507

5.19.8 TREATMENT 508

5.19.8.1 Psychosocial Interventions 508
  5.19.8.1.1 Parent-based interventions 508
  5.19.8.1.2 School-based interventions 509
  5.19.8.1.3 Child-based interventions 510
5.19.8.2 Pharmacologic Interventions 511
  5.19.8.2.1 Patient selection 512
  5.19.8.2.2 Choice of drugs 512
  5.19.8.2.3 Dose–response and time–action 512
  5.19.8.2.4 Side effects 513
  5.19.8.2.5 Measurement of drug treatment effects 515
5.19.9 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 515
5.19.9.1 Future of Assessments of ADHD 515
5.19.9.2 System Issues in Treatment 516
5.19.9.3 Research Issues in ADHD 516
5.19.10 SUMMARY 516
5.19.11 REFERENCES 517

5.19.1 INTRODUCTION

The childhood disorder now known as attention-deficit hyperactivity disorder (ADHD) has a long and controversial history. A number of milestones are commonly cited in the evolution of the concept in its present form. The behavioral symptoms of inattention, impulsive behavior, and hyperactivity have remained fairly constant across many decades, although the interpretation of their origins and meaning has shifted with changes in the scientific paradigms of human behavior (Table 1).

At the end of the nineteenth century, mental afflictions were commonly divided into "moral" and physical causes. The attribution of moral carried different connotations from its usage today, meaning the impairment of judgment or restraint in the face of overwhelming stimulation, rather than mere unethical or immoral behavior. External and internal stimuli, including excesses of emotion and pain, were considered moral causes as opposed to purely physical causes. Because of their intensity they could overwhelm efforts to restrain impulse and thus prevent acting in acceptable ways. Thus, an influential text of the day stated that, "Every impression on the sensorium, through the external senses, and every passion in excess, may become a moral cause of insanity" (Burrows, 1828, p. 9).

Twenty years before Still (1902) provided his account of the impulsive, hyperactive child suffering from lapses in moral restraint, James provided a penetrating analysis of behavioral inhibition (James, 1890). Although his discussion of attention is deservedly well-known, few realize that James anticipated the modern linkage of attentional processes with behavioral inhibition and discharge of impulse into motor behavior. In adopting the traditional notion of a "moral" aspect of behavioral self-control, James argued that every overt action involves a hesitation requiring effortful attention to an "idea" that provides the impetus to action. A failure to restrain action is the consequence of a failure to hold the idea of a desired action in the focus of attention. This failure of attention is therefore the ultimate basis of an explosive and impulsive behavior pattern: "There is a normal type of character, for example, in which impulses seem to discharge so promptly into movements that inhibitions get no time to arise" (p. 537). Thus, James's influential opus established behavioral inhibition as exclusively a mental, as opposed to a physical, process—a process which linked impulsivity with inattention.

This was the prevailing paradigm when the physician Still (1902) provided an account of 20 youngsters whom he described as suffering from a defect of moral control. His clinical description of impulsive, inattentive, and overactive children anticipated much of what is now described as ADHD, although it also included features we would currently include under conduct, antisocial, or oppositional categories. These lapses in self-restraint did not discount the role of genetic and constitutional factors in the individual which might alter the threshold for succumbing to stimulation.

A shift of paradigms to a physical basis for behavioral disinhibition began three decades later. The catastrophic medical phenomena of the world wars, the pandemic of influenza and von Economo's encephalitis, along with the growing awareness of the long-term consequences of "reproductive casualties," provided the impetus to a more medical paradigm—specifically a neurological one—for explaining the deviant behavioral excesses and cognitive limitations of children. Although neuroanatomy was well-developed at this time, the neurological paradigms were largely those of a static system involving cortical inhibition over

Table 1 Some historical landmarks in the evolution of ADHD.

Author/event	*Contribution*
James (1890)	Linking of defects in voluntary inhibition of action to attentional phenomena: "Effort of attention is thus the essential phenomenon of will" (p. 562)
Still (1902)	Accurate clinical description of a syndrome of involving loss of "volitional inhibition" and "moral control"
Postencephalitis, in 1920	A syndrome described by numerous authors, involving emotional instability, irritability, lying, thieving, impaired memory and attention, poor motor control and general hyperactivity, and requiring milieu, cognitive, and behavioral therapies
Kahn and Cohen (1934)	Described the hyperkinetic syndrome as one of "organic-drivenness" due to brain-stem defects or (later) disinhibition due to cortical damage
Bradley (1937)	Administered benzedrine to a heterogeneous group of hospitalized hyperkinetic and behaviorally disordered children, noting an immediate calming effect on disinhibited behavior as well as an increased zest for learning
Goldstein (1942)	Based on observations of after effects of brain injuries in war, he called attention to many behavior patterns which were commonly considered "functional" as really being due to an "organic" etiology
Strauss and Lehtinen (1947)	Description of a brain damage syndrome with characteristic perceptual, motor, and emotional characteristics
Laufer, Denhoff, and Solomons (1957)	Comprehensive clinical description of a "hyperkinetic impulse disorder," with experiments implicating diencephalic (thalamic) mechanisms
Pasamanick and Knobloch (1959)	Described a "continuum of reproductive casualty" which was responsible for a syndrome of minimal cerebral damage in infancy
Clements and Peters (1962)	A classic paper describing the concept of minimal brain dysfunction, later formalized by Clements (1966) in a consensus statement, and linked to attentional dysfunctions by others from the Arkansas group (Dykman, Ackerman, Clements, & Peters, 1971)
American Psychiatric Association (1980)	Formalization of a diagnostic category whose core disorder is attentional in nature, and in which separate subtypes are evident involving hyperactivity, impulsivity, and inattention

more primitive appetitive systems. Organic explanations of behavior were largely based on a conception of the brain along strict localization of functions within separate structures. But Goldstein's (1942) studies of penetrating head wounds and war injuries broadened the range of behavioral phenomena from localized motor impairments to subtle disturbances of cognitive function such as "concrete thinking." These early postwar writings provided a way of understanding subtle behavioral disturbances as organically based phenomena linked to specific types of brain damage. "Organicity" came to mean a style of behavior as much as a statement of etiology.

During this same period, a very different paradigm was provided by Gestalt psychology, with its focus on the role of perceptual phenomena such as visual illusions. Here the emphasis was on the way the brain organizes sensory input and regulates output according to certain principles involving goodness of fit and whole–part relationships among sensory elements. Motor acts reflected control by organized perceptual structures. Werner, Strauss, and their colleagues applied these paradigms to the plentiful cases of children with identifiable neurological conditions such as cerebral palsy, providing an impetus to the perceptual-motor re-education programs initiated by Cruikshank and others in the early and mid-1940s (Kessler (1980) gives a good account of this period).

These studies of relatively extreme medical phenomena gradually filtered into thinking about behaviors that were not readily explained by obvious neurological antecedents. Eventually, however, it became clear that the paradigm of neurological damage had to be abandoned because many cases of apparently organic behavior had no identifiable neurological insults. But the notion lingered on that the perceptual, motor, and learning features associated with known brain damage were indicators of underlying brain dysfunction, but dysfunction that was minimal compared to the usual neurological catastrophes. The concept of minimal brain dysfunction (MBD) was the immediate progenitor of current descriptive classifications of the disruptive behavior disorders, including ADHD. In the definitive conference paper describing the syndrome of MBD, Clements (1966) proposed that:

> The term "minimal brain dysfunction syndrome" refers ... to children of near average, average, or above average intelligence with certain learning or behavioral disabilities ranging from mild to severe which are associated with deviations of function of the central nervous system. [The disabilities may be shown by] various combinations of impairment in perception, conceptualization, language, memory and *control of attention, impulse, or motor function* [italics added]. (pp. 9–10)

This concept incorporates much of the language and thinking of the previous 70 years, in which learning and behavioral disabilities were linked to the central nervous system through characteristic perceptual, cognitive, and motor patterns, and highlights the role of control of attention, impulse and motor function.

It is generally believed that the term MBD was eventually replaced for lack of direct evidence regarding brain dysfunction (Barkley, 1996). In fact, the concept received considerable support. The prospective perinatal collaborative study by the National Institute of Neurological Diseases and Blindness of 56 000 children followed from birth, demonstrated that symptoms of overactivity and learning disabilities were clearly predictable from adverse perinatal events, including maternal smoking and alcohol use (Nichols & Chen, 1981). As fundamental neuroscience has advanced, there has been unexpected further support for the concept of MBD. As noted by Rapoport and Castellanos (1996), "Since the publication of *DSM-III-R* in 1987 ... a number of anatomic MRI studies ... have in fact documented minimal but significant brain anatomical deviations—ironically, minimal brain dysfunction now takes on new meaning!" (p. 267).

However, the boundaries of the MBD concept were simply too large, allowing as it did for various combinations of impairment across many types of behavior. Anticipating the integrative review by Douglas and Peters (1978), the Arkansas group (Clements & Peters, 1962; Dykman, Ackerman, Clements, & Peters, 1971) began to narrow the core processes to attention. At the same time, the second edition of the *Diagnostic and statistical manual of mental disorders* (*DSM-II*) defined the hyperkinetic reaction of childhood as a disorder "characterized by overactivity, restlessness, distractibility, and short attention span." (American Psychiatric Association, 1968, p. 50.) This definition is etiological in that it characterizes the syndrome as a reaction, in keeping with then-prevailing psychoanalytic doctrine. Between the introduction in 1980 of *DSM-III* and of *DSM-IV* in 1994, opinion vacillated regarding the centrality of attention and hyperactivity/impulsivity in the syndrome, but etiological explanations were no longer included as necessary in a system which attempts to be purely descriptive.

This brief review highlights the evolutionary nature of the ADHD concept, and its dependence upon the prevailing scientific paradigms used to explain behavior and mental functioning. Evolution of the concept has proceeded from the mentalistic formulations of the nineteenth century, to the organic concepts of the early twentieth century, and finally to a purely behavioral and descriptive formulation. This chapter depicts the current state of the art with respect to assessment, diagnosis, and treatment, and points to the clinical and research goals which are yet to be achieved in understanding the disorder and its treatment.

5.19.2 PHENOMENOLOGY

5.19.2.1 Core Features

Despite great variation with respect to diagnostic labels, etiologic theories, and conceptions of primary deficits, clinical descriptions of children with the neurobehavioral syndrome now referred to as ADHD have been remarkably stable. The doggerel verse used to describe *Fidgety Phil* in 1863 by the German pediatrician Heinrich Hoffman might today be replaced by a downloaded video clip of a hyperactive child, but the symptoms depicted by any medium will include some combination of developmentally inappropriate levels of inattention, impulsivity, and overactivity that emerge early in life and persist with time and across situations. Indeed, despite comprising a heterogeneous set of behavioral and cognitive problems, it has been parsimonious to consider the symptoms of ADHD as emerging from deficient regulation in three areas: attention, response inhibition, and activity level. This familiar tripartite division of core features has not always been supported by factor analytic studies of parent and teacher behavior ratings, from which two rather than three stable dimensions consistently emerge (Hinshaw, 1994). One factor comprises inattention, cognitive impulsivity and, in some instances, restlessness, whereas the other is formed by behavioral impulsivity and motoric overactivity (Bauermeister, Alegria, Bird, Rubio-Stipec, & Canino, 1992; Lahey et al., 1988). However, in a study of a national normative sample, confirmatory factor analysis was used to demonstrate that a three-factor model was the best fit to the 18 *DSM-IV* symptoms (Conners, Parker, & Sitarenios, 1996). Nevertheless, the high correlation of the impulsivity and hyperactivity factors justifies their lumping together for practical purposes. Hence, the current diagnostic system employed in the USA (discussed below) defines inattention and hyperactivity–impulsivity as the core broad dimensions of ADHD.

It has long been recognized that the core features of ADHD, however partitioned, do not represent unitary constructs and precise specification of symptoms is lacking. The following are brief discussions of the heterogeneity and phenomenology of these features.

5.19.2.1.1 Inattention

Although coordinated in function, multiple dimensions of attention have been identified and evidence is emerging that each type reflects a separate underlying neural network (Mesulam, 1990; Posner, 1988; Voeller, 1991 (review)). Components of attention most consistently described include (i) arousal or alertness, (ii) selective or focused attention (the ability to attend to particular stimuli while ignoring competing, irrelevant stimuli), (iii) attentional capacity or span of apprehension (the amount of information to which the child can attend at one time), and (iv) sustained attention or vigilance (the persistence of focus over time). Clinically, children with ADHD may show fairly specific attentional problems, but more often have attentional difficulties that are multifaceted. Common manifestations include an inability to remain engaged in schoolwork, chores (and, in some instances, games), high rates of off-task behavior, seeming to be in another world when being spoken to, difficulties recalling and following through on instructions with multiple steps, poor task completion, distractibility by stimuli that are more appealing than the task at hand, and a disregard for details that often leads to careless mistakes.

Laboratory evidence compiled by Douglas in the 1980s (Douglas, 1983) led to the predominant view that poor sustained attention or vigilance represented the core deficit in children with ADHD. Indeed, the frequency with which these children drift off-task and fail to persist with effortful activities of low inherent interest suggests a vigilance decrement. However, results of laboratory studies aimed at illuminating this attention deficit have been conflicting. Whereas some investigations reveal deficient vigilance in ADHD samples (e.g., Seidel & Joschko, 1990), many have failed to support a sustained attention deficit (e.g., Schachar, Logan, Wachsmuth, & Chajczyk, 1988; Sergeant & Scholten, 1985; van der Meere, Wekking, & Sergeant, 1991). Part of the problem is that the vigilance decrement is an age-related phenomenon which is much less evident in older children and adolescents (Seidel & Joschko, 1990). In light of such findings, some have argued that the clinical phenomena commonly interpreted as a sustained attention deficit actually reflect a problem with the preparation, regulation, speed, and consistent execution of motor output (Barkley, 1995b; van der Meere, van Baal, & Sergeant, 1989). Barkley (1994) has reframed attentional problems in children with ADHD as deficits in goal-directed persistence, meaning that the persistence of their effortful behavior is highly dependent on the nature of the task context (e.g., intrinsic interest or novelty) and on immediately available reinforcement rather than being shaped by distal rewards and future goals. Clearly, the nature and primacy of attentional problems in ADHD remain elusive. Contributing to this state of affairs is the complex intertwining of cognitive, motivational, affective, and behavioral variables, along with varying task parameters, in the responses regarded as reflecting attention, both in the laboratory and the natural environment.

5.19.2.1.2 Hyperactivity

Characteristic symptoms in the hyperactive domain vary in both nature and degree. Features ranging from excessive talking and noise-making, to an inability to remain seated, are regarded as reflecting hyperactivity. The category encompasses both the child who taps his feet, fidgets, and manipulates objects when seated, to the whirling dervish who blazes chaotically through space with a driven, accelerated quality. Common to many of the symptoms in this domain is that they lack direction or purpose and are extraneous to the task at hand. As with inattentive symptoms, there is significant situational variability in hyperactivity (Luk, 1985), suggesting a deficient capacity to regulate activity levels to different setting or task demands (Routh, 1978). Thus, the child who stands out so painfully in the structured classroom setting may be far less distinguishable from peers on the playground. Despite the modulation of hyperactive symptoms by environmental factors, it is noteworthy that youngsters with ADHD generally exhibit greater levels of motor activity than comparison children throughout the day, including during sleep (Porrino et al., 1983). Moreover, whereas certain measures of inattention have not been found to distinguish children with ADHD from those with other psychiatric conditions, hyperactivity may represent a more specific marker of the disorder (Halperin, Matier, Bedi, Sharma, & Newcorn, 1992; Roberts, 1990).

5.19.2.1.3 Impulsivity

Although current *DSM* criteria in the impulsivity domain are limited to blurting out answers prematurely, difficulty waiting one's turn, and interrupting or intruding upon others, other impulsive features are commonly observed in children with ADHD, including engaging in physically dangerous activities without considering the possible consequences, deficient inhibition of intense emotional responses, and a hasty, careless response style when working on tedious or cognitively demanding tasks (Brown & Quay, 1977). A proclivity towards impulsive acts may also underlie the higher than average accident rate seen among youngsters with ADHD (Hartsough & Lambert, 1985; Szatmari, Offord, & Boyle, 1989a).

As with inattention and hyperactivity, the category of impulsivity is often conceptualized in an overly inclusive manner which makes it difficult to operationalize. Convergent validity is problematic as various measures of impulsivity (e.g., continuous performance tests, matching familiar figure tests, and parent or teacher ratings) tend to correlate poorly with one another (Milich & Kramer, 1984). There does, however, appear to be justification for distinguishing cognitive and behavioral aspects of impulsivity as they appear to be distinct factors (Taylor, 1994), with differing diagnostic and predictive significance. For example, measures of cognitive impulsivity have been found to distinguish ADHD from psychiatric comparison children (Halperin et al., 1992). In a longitudinal study of 400 boys conducted by White et al. (1994), both cognitive and behavioral impulsivity predicted academic achievement problems, but only the latter was associated with increased risk for antisocial behavior.

In contrast to prior conceptualizations of the disorder that identified hyperactivity or inattention as primary deficits, Barkley (1994) has argued that an impaired capacity to inhibit or delay responding is at the core of ADHD. This theory contends that the chief and associated features of ADHD can be regarded as secondary to poor inhibitory control. For instance, the capacity to inhibit responses is said to permit self-regulation which, in turn, limits task-irrelevant motor activity and makes persistence in the face of boredom, fatigue, or competing stimuli possible. Thus, Barkley reverses the causal direction of the relationship between impulsivity and inattention as interpreted by James.

5.19.2.2 Associated Features and Comorbidity

The phenotypic heterogeneity of ADHD is a product not only of the variegated nature of its core symptoms, but also multiple associated clinical features and co-occurring psychiatric conditions. The most common secondary features among children with ADHD include academic underachievement, noncompliant and aggressive behavior, family discord, and poor peer relationships. Adequate functioning in these domains has significance for mastering many critical developmental tasks (Richters et al., 1995). Conversely, poor functioning in these areas is highly predictive of later maladjustment (Hinshaw, 1994), making these secondary features of ADHD particularly pernicious. The following are brief comments regarding the particular difficulties experienced in these domains by children with ADHD.

The symptoms of ADHD typically compromise a child's ability to meet academic as well as behavioral demands at school. An overwhelming majority of children with the disorder demonstrate academic performance problems relative to their classmates (Cantwell & Baker, 1991), as indexed by lower rates of on-task behavior (Abikoff, Gittelman-Klein, & Klein, 1977) and academic task completion (Pfiffner & Barkley, 1990), and by higher rates of academic underachievement, placement in special education services, grade retention, and dropping out of school (Barkley, DuPaul, & McMurray, 1990; Barkley, Fischer, Edelbrock, & Smallish, 1990; Hinshaw, 1992b). Learning disabilities are highly correlated with externalizing or disruptive behavior disorders in general, but the relationship appears to be strongest with ADHD (DuPaul & Stoner, 1994; Hinshaw, 1992b). Although prevalence estimates of learning disorders among youngsters with ADHD vary widely, most recent estimates employing conservative criteria for defining learning disorders range from 10% to 25% (Hinshaw, 1992a; Semrud-Clikeman et al., 1992). Despite their far above chance co-occurrence and complex bidirectional influences, consensus has emerged that ADHD and learning disabilities represent distinct disorders (Shaywitz, Fletcher, & Shaywitz, 1995).

Children with ADHD commonly manifest a variety of externalizing features. These typically involve willful noncompliance or defiance in response to adult directives or rules, disruption of ongoing activities, argumentativeness, temper outbursts, verbal hostility, and physical aggression (Loney & Milich, 1982). Such behaviors are often of sufficient severity and duration to warrant a diagnosis of oppositional defiant disorder (ODD), which pertains to recurrent, age-inappropriate patterns of negative, defiant, and hostile behaviors, often accompanied by an irritable mood (American Psychiatric Association, 1994). Recurrent patterns of the more serious antisocial behaviors (e.g., stealing, property destruction, aggression towards persons or animals, deceit, truancy) that comprise conduct disorder (CD) (American Psychiatric Association, 1994) are less common than ODD features but still co-occur with ADHD at rates that far exceed chance. The rates of ODD or CD found in ADHD samples have ranged from 30% to over 50% (Anderson, Williams, McGee, & Silva, 1987; Beiderman, Newcorn, & Sprich, 1991). Despite the high degree of overlap between ADHD and both ODD and CD, research, examining group differences in variables including environmental background, family psychiatric history, cognitive impairments, family interactions, and course, supports the partial independence of these syndromes (Abikoff & Klein, 1992; Hinshaw, 1987). Moreover, those ADHD children with comorbid CD may represent a qualitatively distinct subgroup from ADHD youngsters without a comorbid disruptive behavior disorder (Hinshaw, 1994).

The intra- and extrafamilial interpersonal functioning of ADHD children is frequently disturbed. Elevated rates of negative and conflictual parent–child interactions characterize families with an ADHD child (Barkley, Karlsson, & Pollard, 1985; Mash & Johnston, 1982). In addition, these families are marked by higher levels of parental stress and psychopathology and lower levels of marital harmony and self-perceived parenting competence than those found in families without a disturbed child, although some of these findings apply to ADHD children with comorbid ODD or CD (Anastopoulos, Guevremont, Sheton, & DuPaul, 1992; Anderson, Hinshaw, & Simmel, 1994; Johnston, 1996; Schachar & Wachsmuth, 1990). The interactions of ADHD children with their teachers, like those with parents, are often characterized by lower than average rates of positive exchanges and higher rates of commands and negativity (Pfiffner & Barkley, 1990). Although the child and parent characteristics are bidirectional in their effects, it is worth emphasizing that many negative parent characteristics disappear with corrective intervention with the child, an important therapeutic point in reducing the guilt of parents prone to blame themselves for the appearance of the child's behavioral symptoms.

Peer relationships are also typically disrupted as children with ADHD have difficulties initiating and maintaining friendships (Guevremont, 1990). In fact, studies employing sociometric measures with ADHD samples have consistently revealed them to be poorly accepted or actively rejected by peers at much higher rates than normal children and children with other psychiatric disorders (Asarnow, 1988; Milich & Landau, 1982; Whalen & Henker, 1992). Rather than arising from deficient prosocial skill deficits, the social failure experienced by ADHD children appears to relate most strongly to their interpersonally noxious behaviors (e.g., aggression, noncompliance, disruption, and intrusiveness), affective dysregulation, and impaired capacity to modulate their social responses to shifting environmental demands. Moreover, the peer rejection suffered by children with ADHD emerges after very brief periods of interaction (Erhardt & Hinshaw, 1994; Pelham & Bender, 1982), is stable over time, and represents a robust predictor of a variety of adverse long-term outcomes (e.g., academic failure, school dropout, criminality or psychopathology) (Parker & Asher, 1987).

With respect to formal psychiatric diagnoses, ADHD is most commonly comorbid with ODD, CD, and learning disorders (see previous discussion). However, other disorders also co-occur with ADHD at above-chance levels. Indeed, it is more common for ADHD to present with one or more comorbid conditions than in the absence of another diagnosis (Newcorn & Halperin, 1994). These include anxiety disorders, which may occur in 25% or more of children with ADHD (Biederman et al., 1991), expressive language disorders (Hamlett, Pelligrini, & Conners, 1987; Zentall, 1989), Tourette's and other tic disorders (Cohen & Leckman, 1994), and motor coordination deficits (Barkley, 1990). The evidence with respect to co-occurring mood disorders is mixed, with some studies finding substantially elevated prevalence rates (Biederman, et al., 1991) and others reporting no significantly enhanced risk for affective illness in ADHD samples (Gittleman, Mannuzza, Shenker, & Bonagura, 1985; Lahey et al., 1988). Consideration of comorbid conditions is critical as they may have important implications for etiology, severity, course, and response to both pharmacologic and psychosocial treatments (March, Wells, & Conners, 1995a; Newcorn & Halperin, 1994).

5.19.2.3 Variations Related to Gender and Development

5.19.2.3.1 Gender

The expression, associated features, and course of ADHD in females is receiving increasing attention. However, research in this area remains scarce and what exists is plagued by small sample sizes which limit statistical power and preclude definitive conclusions regarding gender differences. The following comments on what is currently known have been culled from Arnold's (1996) summary of the findings emerging from a conference convened in late 1994 by the Child and Adolescent Disorders Branch of the National Institute of Mental Health to review the literature on gender differences in ADHD.

A preponderance of available evidence suggests a basic similarity between males and females with respect to the qualitative expression of the disorder; quantitatively, however, females with ADHD may, on average, show less hyperactive–impulsive behavior and less inattention than do their male counterparts. It is important to bear in mind that similar gender differences with respect to these behaviors exist in the normal population (Achenbach & Edelbrock, 1983; Goyette, Conners, & Ulrich, 1978), which raises the thorny issue of whether gender-based norms for establishing deviance should be incorporated into diagnostic thresholds. Diagnostically, the Predominantly Inattentive subtype of ADHD appears to be overrepresented in females compared to the Combined subtype (Lahey et al., 1994). Gender differences have emerged for comorbid conditions, as males with ADHD exhibit a higher rate of antisocial-spectrum difficulties, including aggression, oppositional-defiant behavior, and conduct disorder symptoms. Although some data suggest more severe intellectual, learning, and language deficits in females with ADHD, findings related to these variables are few and sufficiently mixed that conclusions would be premature. Sex-based differences in response to psychostimulant medications do not appear to exist. Insufficient evidence exists to address whether the outcomes of males and females with ADHD differ, although the greater severity of core symptoms and higher levels of associated antisocial problems may presage a more negative outcome for males. Concerns continue to be raised that females with ADHD represent an underdiagnosed and underserved population.

5.19.2.3.2 Developmental variation

The diagnosis of ADHD and the delineation of a coherent clinical picture of the disorder across the life-span are complicated by the developmental nature of the constituent behaviors. Many preschool-aged children are highly active and exhibit both inattentive and impulsive behaviors. Yet it is often difficult to distinguish normal variation in such behaviors that will dissipate as development proceeds, from the initial expression of a pathological syndrome that is likely to be chronic, impairing, and deserving of treatment (Campbell, 1990). Nevertheless, early onset behavior problems prove to be chronic and serious in a proportion of children who show such difficulties, particularly when a constellation of overactive, inattentive, impulsive, and defiant behaviors are fairly severe, persist for a year or longer, and are accompanied by negative parenting practices and family adversity (Campbell, 1990, 1995; Richman, Stevenson, & Graham, 1982). As discussed later in this chapter, it is now well-established that symptom patterns of sufficient severity and duration to warrant a diagnosis of ADHD in early childhood tend to persist at least into adolescence in a substantial majority of cases (Weiss & Hechtman, 1993).

Another way in which ADHD represents a developmental disorder is that its expression typically varies with age. Despite a trend toward overall symptom reduction with maturity (Barkley, 1990), both core and associated symptoms tend to be present in each age group, but with an altered topography. March et al. (1995a) summarize how ADHD symptoms manifest at different ages; Weiss & Hecthman (1993) provide a detailed discussion. For example, hyperactivity declines with age and minor extraneous motor behaviors (e.g., fidgeting or tapping) or an internal sense of restlessness may replace the excessive running and climbing of early and middle childhood (American Psychiatric Association, 1994). Some researchers have failed to find similar developmental declines in inattentive symptoms (Hart, Lahey, Loeber, Applegate, & Frick, 1995; Parker, Sitarenios, & Conners, 1996). Changes in symptom patterns are likely to reflect not only physiological maturation but also shifting social ecologies and environmental demands (Whalen & Henker, 1980). During adolescence, increased demands to manage multiple classes, to complete multistaged projects within a certain time-frame, and to structure one's own time, tend to increase the salience of organizational and executive functioning deficits in ADHD (Wender, 1995). Conversely, many symptoms may become less notable once individuals escape the unique demands of school and exercise greater freedom to determine working and living environments which provide a better match for their particular profile of strengths and deficits.

5.19.3 DIAGNOSIS

As part of the comprehensive assessment process described later in this chapter, the clinician will typically determine whether a child satisfies official diagnostic criteria for ADHD. Although limited in scope and utility, arriving at an accurate categorical diagnosis does represent a legitimate part of the assessment process as it serves a number of important functions. These include providing a common language for communication among professionals for both research and clinical purposes, conveying useful information on likely associated features, course, and treatment response and, on a more pragmatic level, facilitating the payment for clinical services by insurers and managed care providers (Volkmar, 1991).

The evolution of diagnostic systems for ADHD has been both incited and accompanied by constant debate regarding definitional boundaries, labels, and primary symptoms. The frequent changes in classification schemes and marked differences between those used in the USA and Europe have obstructed progress in our understanding of the disorder because they have compromised the legitimacy of cross-study comparisons. As noted by Hinshaw (1994), the ultimate aim of any categorical diagnostic scheme is to identify subgroups of individuals who are homogeneous with respect to presentation, pathogenesis, family history, course, and treatment response. Although this still represents a distal goal with respect to psychiatric nosology, progress is evident in the increasing reliance on empirical field trials of prospective criteria along with expert consultation to inform diagnostic systems (Lahey et al., 1994).

One by-product of such progress is the growing rapprochement between the two major taxonomies for psychopathology: the *DSM-IV* in the USA (American Psychiatric Association, 1994) and the *International statistical classification of diseases and related health problems, tenth revision* (*ICD-10*) (World Health Organization, 1992a). Conceptually, these diagnostic systems are quite similar in that the behavioral symptoms each lists as comprising ADHD are identical or show only minor variations in wording. The *DSM-IV* criteria for ADHD are listed in Table 2.

In addition to their overlap in symptom content, both criteria sets stipulate that symptoms be present to developmentally inappropriate degrees, persistent for at least six months, pervasive across settings, emergent before (*DSM-IV*) or by (*ICD-10*) the age of seven, and associated with significant impairment in social, academic, or occupational functioning (although significant distress can substitute for impairment in *ICD-10*). Despite their gradual convergence, significant differences remain between the diagnostic approaches to ADHD contained in *DSM-IV* and *ICD-10*. The *DSM-IV* system, drawing upon the results of factor analytic studies and extensive field trials (e.g., Bauermeister et al., 1992; Lahey et al., 1984, 1988), groups symptoms into the two dimensions of inattention and hyperactivity–impulsivity and defines three diagnostic subtypes on the basis of an individual's profile on these dimensions. Specifically, "ADHD Combined Type" requires at least six of the nine symptoms in both the inattention and hyperactive–impulsive domains, whereas the "ADHD, Predominantly Inattentive Type" and the "ADHD, Predominantly Hyperactive–impulsive Type" each require six or more symptoms in one problem area but fewer than six symptoms in the other domain.

The distinction of an inattentive subtype of ADHD (which is also commonly referred to as "attention deficit disorder without hyperactivity") is well supported by factor and cluster analytic studies and by data showing that family history and comorbid symptom patterns in such a subgroup diverge from those found in other children who are diagnosed with ADHD (Cantwell & Baker, 1992; Goodyear & Hynd, 1992). However, the validity of the unprecedented predominantly hyperactive–impulsive subgroup is less well established. This subtype has been found to apply disproportionately to children below school age (Lahey et al., 1994) who have had little opportunity to exhibit inattentive symptoms which become most evident in demanding, structured environments (e.g., careless errors, disorganization, or poor task completion). Thus, some suggest that the predominantly hyperactive–impulsive group may represent a developmental precursor of "ADHD, combined type" rather than a distinct subtype (Barkley, 1995a). Nonetheless, the *DSM-IV* criteria for ADHD is generally regarded to be an improvement over its predecessors, yielding diagnoses that are more reliable and more clearly associated with functional impairment (McBurnett, Lahey, & Pfiffner, 1993).

In *ICD-10*, the term "hyperkinetic disorder" is adopted rather than ADHD, and inattention, hyperactivity, and impulsivity are listed as separate problem areas, with diagnosis requiring symptoms in all three domains. However, inattention and hyperactivity are regarded as cardinal features and are emphasized over impulsive symptoms (World Health Organization, 1992b). Thus, the diagnosis of "hyperkinetic disorder" requires the presence of six of

Table 2 *DSM-IV* diagnostic criteria for attention-deficit hyperactivity disorder.

A. Either (1) or (2):

(1) Six (or more) of the following symptoms of **inattention** have persisted for at least six months to a degree that is maladaptive and inconsistent with development level:

Inattention

- (a) often fails to give close attention to details or makes careless mistakes in schoolwork, work, or other activities
- (b) often has difficulty sustaining attention in tasks or play activities
- (c) often does not seem to listen when spoken to directly
- (d) often does not follow through on instructions and fails to finish schoolwork, chores, or duties in the workplace (not due to oppositional behavior or failure to understand instructions)
- (e) often has difficulty organizing tasks and activities
- (f) often avoids, dislikes, or is reluctant to engage in tasks that require sustained mental effort (such as schoolwork or homework)
- (g) often loses things necessary for tasks or activities (e.g., toys, school assignments, pencils, books, or tools)
- (h) is often easily distracted by extraneous stimuli
- (i) is often forgetful in daily activities

(2) Six (or more) of the following symptoms of **hyperactivity–impulsivity** have persisted for at least six months to a degree that is maladaptive and inconsistent with developmental level:

Hyperactivity

- (a) often fidgets with hands or feet or squirms in seat
- (b) often leaves seat in classroom or in other situations in which remaining seated is expected
- (c) often runs about or climbs excessively in situations in which it is inappropriate (in adolescents or adults, may be limited to subjective feelings of restlessness)
- (d) often has difficulty playing or engaging in leisure activities quietly
- (e) is often on the go or often acts as if driven by a motor
- (f) often talks excessively

Impulsivity

- (g) often blurts out answers before questions have been completed
- (h) often has difficulty awaiting turn
- (i) often interrupts or intrudes on others (e.g., butts into conversations or games)

B. Some hyperactive–impulsive or inattentive symptoms that cause impairment were present before age seven years

C. Some impairment from the symptoms is present in two or more settings (e.g., at school/work and at home)

D. There must be clear evidence of clinically significant impairment in social, academic, or occupational functioning

E. The symptoms do not occur exclusively during the course of a pervasive developmental disorder, schizophrenia, or other psychotic disorder and are not better accounted for by another mental disorder (e.g., mood disorder, anxiety disorder, dissociative disorder, or a personality disorder)

Code based on type:

314.01 Attention-Deficit Hyperactivity Disorder, Combined Type: if both criteria A1 and A2 are met for the past six months

314.00 Attention-Deficit Hyperactivity Disorder, Predominantly Inattentive Type: if criterion A1 is met but criterion A2 is not met for the past six months

314.01 Attention-Deficit Hyperactivity Disorder, Predominantly Hyperactive–Impulsive Type: if criterion A2 is met but criterion A1 is not met for the past six months

Diagnostic and statistical manual of mental disorders (4th ed.), American Psychiatric Association, 1994, Washington, DC: Author. Copyright 1994 by American Psychiatric Association. Reprinted with permission.

nine symptoms of inattention, three of five hyperactive features, and one of four impulsivity symptoms. Of note is the fact that the item pertaining to excessive talking, which is grouped with the hyperactivity features in *DSM-IV*, is placed among the impulsivity items in *ICD-10*.

Other differences between the *DSM-IV*, and *ICD-10* systems pertain to subtyping, exclusionary criteria, and comorbidity. For example, *ICD-10* regards the predominantly inattentive and predominantly hyperactive–impulsive subtypes of ADHD, recognized by *DSM-IV*, to be subthreshold conditions for which separate diagnostic codes are not provided. Rather, the primary division within the *ICD-10* hyperkinetic disorders is between "disturbance of activity and attention," which corresponds closely to the *DSM-IV* ADHD, combined type, and "hyperkinetic conduct disorder," a mixed category that is applied when the criteria for both hyperkinetic disorder and conduct disorder are met.

The existence of the hyperkinetic conduct disorder diagnosis in *ICD-10* illustrates one difference in the way in which the two nosologies manage the comorbidity so common to ADHD. As discussed by Taylor (1994), *DSM-IV* represents a multiple diagnosis scheme in which comorbid diagnoses are readily assigned if the criteria for each is met, provided that the symptoms of one disorder are not better accounted for by a coexisting condition. Contrasting with this approach is the single diagnosis scheme adopted by *ICD-10*. Although multiple diagnoses are permitted, the system steers the diagnostician towards choosing a single disorder on the basis of the best available match between the child's presentation and the prototypic descriptions offered in *ICD-10*. Supporting this more conservative approach to multiple diagnoses is the inclusion in *ICD-10* of precedence rules (e.g., depressive and anxiety disorders generally take precedence over hyperkinetic disorders) and listing of exclusionary criteria which are more restrictive than those found in *DSM-IV*. One limitation common to both classification schemes is the developmental insensitivity of both the symptoms listed and the diagnostic thresholds (Barkley, 1995a; Hinshaw, 1994). Thus, preschool children and adults are judged by the same items and diagnosed using the same cutpoints, despite the changing topography and severity of symptoms with maturation. Taylor (1994) provides a more detailed consideration of the *DSM* vs. *ICD* approach to diagnosing ADHD.

Regardless of which classification scheme is used, a number of additional factors complicate the clinical judgment involved in assigning an ADHD diagnosis and create the potential for both overdiagnosis and misclassification. First, many of the symptoms occur frequently in nondisordered youngsters and reflect normal developmental variation rather than an underlying pathology. Second, the heterogeneous presentation of ADHD children and the variability in performance and symptom expression within a given child as a function of time, task, and setting variables can make the disorder an elusive target. Third, many of the symptoms are not specific to ADHD but can be associated with other forms of psychopathology (e.g., anxious or depressive disorders) and with adverse environmental circumstances (e.g., chaotic families, ineffectual classrooms, or psychosocial stressors). These potential pitfalls in diagnosing ADHD place a premium on the clinician's assessment of the developmentally deviant and impairing nature of presenting symptoms, reliance on a comprehensive multimodal, cross-informant assessment battery, and command of child psychopathology to aid in differential diagnosis.

Detailed differential diagnosis is especially important in ADHD due to both the frequent co-occurrence of ADHD with other disorders (as discussed above) and to the potential for other syndromes to produce symptoms that overlap with the behavioral criteria for ADHD. Examples of the potential clinical implications of differential diagnosis include the markedly different patterns of impairment and outcomes that typify ADHD with and without conduct disorder (Hinshaw, 1992a; Walker, Lahey, Hynd, & Frame, 1987), the divergent treatment plans indicated for ADHD, learning disorders, and their combination (DuPaul & Stoner, 1994), and the less favorable stimulant response among ADHD children with comorbid anxious disorders (Pliszka, 1989). Although a detailed discussion of differential diagnosis in ADHD is beyond the scope of this chapter (see Barkley, 1990; Sharma, Matier, & Halperin, 1994), the process typically requires a thorough knowledge of the full spectrum of childhood disorders and reference factors such as age and type (acute versus insidious) of onset, course (episodic versus chronic), symptom pervasiveness, family psychiatric history, and the child's environmental context (e.g., psychosocial stressors, family relations, and school climate).

5.19.4 PREVALENCE

The reported prevalence of any disorder is strongly affected by many variables, including the definition of "caseness," sampling methods, the sampling frame, age, and comorbidity. These factors probably account for the range of prevalence of 2–6.3% in recent studies of

ADHD (Szatmari, Offord, & Boyle, 1989b), and 20% or more in some older studies.

5.19.4.1 Caseness Criteria

Slight differences in symptom criteria can markedly affect recorded prevalence of the disorder. When the *DSM* shifted from the hyperactivity–impulsivity–inattention criteria of *DSM-III* to the single list of 14 symptoms of *DSM-III-R*, ADHD became 14.4% more prevalent than attention deficit disorder/hyperactivity (Lahey, Loeber, Stouthamer, Christ, 1990). It had been observed for some time that English clinicians tended to record far fewer cases of their *ICD-9* equivalent of attention deficit disorder (ADD) (hyperkinetic disorder), than their American counterparts. The causes of this unequal prevalence of hyperactivity in the UK and the USA were investigated by British and American clinicians and research panels using both *ICD-9* and *DSM-III* criteria. Although there was poor clinician reliability, the specially trained research panels reliably found that *ICD-9* generated fewer diagnoses of hyperkinetic syndrome than did *DSM-III* for attention deficit disorder with hyperactivity. This difference was greatest for UK clinicians, and both the diagnostic scheme and clinician training contributed significant variance to the difference in reported rates (Prendergast et al., 1988). When using standard rating scales, the mean hyperactivity score of English schoolchildren is actually higher than their American peers (Taylor & Sandberg, 1984).

Questionnaires can yield any prevalence rate, depending upon the choice of cutoff score. But other factors affect prevalence obtained by ratings. The format of questionnaires used for assessing hyperactivity or ADHD can have an impact on the obtained prevalence rates. In a study of 2000 Australian and English schoolchildren, a rate of 5.6% was found using the Conners teacher rating scale, 7.5% using a scale which compared behavior to a normal standard, and 8.9% using a scale with yes–no format. When all three scales were required for inclusion, the rate dropped to 3.5%, whereas the rate for identification by any one of the three scales rose to 12% (Holborrow, Berry, & Elkins, 1984).

The symptom informant has a strong effect on recorded prevalence. One of the largest studies, carried out on 5000 children in the Bay area of San Francisco, demonstrated that hyperactivity varied from 1% to 12% depending upon whether one, two, or three "system definers" of parent, teacher, and clinician made the diagnosis (Lambert, Sandoval, & Sassone, 1978). One study, using a structured diagnostic interview with mid-Western nine-year-olds and a parent, found a rate of 3.5% by child report, 12% by parent report, and 1.7% when both were required (Shekim et al., 1985).

5.19.4.2 National and Cultural Factors

In identifying the setting and task parameters which affect the expression of ADHD symptoms, researchers have been sensitive to context in a narrow sense. However, less attention has been paid to the impact of broader contextual issues such as culture, socioeconomic status, and ethnicity, on the expression, course, treatment response, and outcome of ADHD. The behavioral features that comprise ADHD appear in diverse cultures, although reported prevalence rates vary (Bhatia, Nigam, Bohra, & Malik, 1991; Szatmari et al., 1989b; Tao, 1992). Such variation may reflect true differences in prevalence but is more often thought to arise from differences with respect to assessment procedures, classification systems, standards for demarcating normal from disordered behavior, and culturally biased perceptions of behaviors (American Psychiatric Association, 1994; Chandra, 1993; Reid, 1995; Taylor, 1989). Although some studies have suggested a lack of uniformity in the rates of ADHD-related behaviors across socioeconomic and ethnic lines (e.g. Goyette, et al., 1978; Stevens, 1981; Waechter, Anderson, Juarez, Langsdorf, & Madrigal, 1979), findings in this area have been mixed.

Cross-national studies have been too few and methodologically different to warrant firm findings, but some evidence suggests that both the rates of the disorder and its definition may vary as a function of child-rearing values and gender roles in different cultures. A rating scale study of 344 second through fourth graders in Italian schools showed a markedly higher rate of hyperactivity in Italian boys (20%) than comparable US norms. The rate for girls was only 3%. Although Italian children exhibited rates of hyperactive behaviors similar to those of children in New Zealand, Spain, and portions of the USA, factor structures of the ratings for boys and girls were different, suggesting the need for separate factor analyses for males and females in cross-national research. Hyperactivity and conduct problems, in fact, were separate factors for girls, whereas such behaviors were subsumed under a single factor for boys. There were also important similarities and differences with the factor structures for boys and girls in Italy and Spain. These differences were interpreted as reflecting differences in the cultural acceptability of certain behaviors in Italy and Spain (O'Leary, Vivian, & Nisi, 1985). Notable cultural and ethnic differences in prevalence have been reported for Mexican and African-American studies of hyperactivity and ADD (Juarez, Madrigal, & Anderson, 1981; Lansdorf, 1979; Solis & Solis, 1988; Waechter et al., 1979). A teacher rating scale study was carried out in 2770 Chinese children from elementary schools in urban, suburban, and mountain communities. Prevalence rates of hyperactivity varied from 3.1% in urban areas, to 7.8% in suburban, and 7% in mountain communities (Yu, Yu, & Xiao, 1985). The suburban rates are closely comparable to the rate of 8.76% found in comparably aged recent Chinese immigrants (Yao, Solanto, & Wender, 1988). The marked regional differences contrast with relatively minor regional effects in national surveys within the USA (Achenbach, Conners, Quay, Verhulst, & Howell, 1989; Achenbach, Howell, Quay, & Conners, 1991) perhaps reflecting a greater range of regional differences in economic and social circumstances in the Chinese sample.

5.19.4.3 Comorbidity

Community-based studies show that many of the socioeconomic, gender, and family dysfunction factors associated with ADHD disappear when influences of comorbidity are statistically controlled (Szatmari et al., 1989a). Children of parents with diagnoses of panic disorder, major depressive disorder, or no diagnosis show marked differences in the rates of diagnosed attention-deficit disorders. The comorbidity of ADD and anxiety and/or depressive disorders was examined in 163 7–17-year-old offspring of parents with panic, depression, or no diagnoses. Diagnoses of ADD by self-report, parent report, and best-estimate consensus were all higher in the offspring of affected parents. Much higher rates of ADHD were reported when anxiety and/or depression were present in the child as well (McClellan, Rubert, Reichler, & Sylvester, 1990).

5.19.4.4 Developmental Factors

A major problem with categorical diagnostic schemes such as *DSM* is that fixed numbers of symptoms are required, and the normative rate of symptoms may change with age. For this reason, authoritative sources recommend supplementing *DSM* categorical criteria with norm-based symptom criteria (Barkley, 1990). The problem of syndromal definition is further complicated by differences in age trends for hyperactive–impulsive symptoms and inattention symptoms. Biological parents of hyperactive–impulsive children, compared with adoptive parents, show clear evidence of inattention but no hyperactivity–impulsivity, suggesting that the familial or genetic trait of hyperactivity-impulsivity tends to disappear in adolescence or adulthood (Alberts, Firestone, & Goodman, 1986). This finding is consistent with long-term follow-up studies which also suggest age-related changes in symptomatology (Hechtman, 1989; Weiss & Hechtman, 1986).

Although there is little doubt regarding age-related changes in the level of symptomatology, a more important issue is whether the syndromal structure is similar across ages. If not, then one could reasonably argue that a different appellation, such as "residual state," might be needed instead of the same diagnostic label that is used in childhood. In a recent study of a large national sample, both parents and teachers rated the 18 *DSM-IV* symptoms on a four-point Likert scale (not at all, just a little, pretty much, very much). Three different models of the factor structure were tested: a single linear factor; a hyperactivity–impulsivity and inattention two-factor model; and a hyperactivity–impulsivity–inattention three-factor model. For both parents and teachers the three-factor model provided the best fit to the data across all age groups (3–17 years), using stringent goodness-of-fit criteria. However, for practical purposes the two-factor model is probably sufficient, given a high correlation between the hyperactivity and impulsivity item clusters (Conners et al., 1996). This study clearly showed strong age-related declines in hyperactivity–impulsivity, whereas inattention tended to remain constant across the age groups. This cross-sectional study replicates a similar finding from the study of 7–12-year-old males with ADHD diagnoses followed longitudinally over a four year period (Hart et al., 1995).

5.19.5 ETIOLOGY AND RISK FACTORS

Modern concepts of disease emphasize that diseases are best conceptualized as a product of accumulated risks. There are many pathogens that require certain factors in the host before the disease will manifest itself. There are many protective factors, both biological and social, that operate in the host to ward off disease. Various contributing factors, such as genetic influences, neonatal and perinatal insults to the brain, normal variability of temperament patterns, and environmental risks and protective factors, interact in complex ways to produce the final behavioral outcome in a child. Considerable research demonstrates that many psychological and psychosocial phenomena are the result of biological risks (Rutter & Casaer, 1991). This concept is important because it frames the nature of research questions—

instead of asking: What is the cause of ADHD? with its implicit assumption of a single etiological pathway, it directs researchers to ask: What are the various combinations of risks and protections that produce the syndrome of ADHD?

The concept of risk factors is important for prevention, assessment, and treatment. Some environmental risks can be eliminated altogether, such as eliminating air-borne lead or other pollutants, preventing maternal smoking or drinking during pregnancy, or providing adequate nutrition to the developing brain. An understanding of biological risks is also important in the assessment and treatment process because knowledge regarding biological sources of ADHD helps parents by ameliorating guilt, develops rapport during assessment, and promotes positive attitudes towards the child and the family by minimizing blame and scapegoating. The role of biological risks in producing the syndrome has been important in gaining acceptance for school-based interventions for ADHD as one of the legitimate handicapping conditions.

Protective factors can offset or ameliorate the impact of risks, even when those risks are at the severe end of the spectrum (Werner & Smith, 1967; Werner et al., 1968). Good parenting, achievement orientation and the value of effort, socioeconomic advantage, educational opportunity, good diets, antistress techniques, and positive lifestyles, can all minimize otherwise disastrous effects of biological and environmental risks. The most important protective factors are social, parental, and educational. These form the basis of therapeutic efforts with ADHD children.

Among the most important risk factors for ADHD are various comorbid conditions that may occur along with ADHD. These conditions—anxiety, aggression, oppositionality, depression—may occur as genetically linked risks, secondary consequences, or simply as independently occurring processes. Some conditions, such as low self-esteem and demoralization, are almost inevitable secondary consequences of ADHD itself. Demoralization, depression, and anxiety frequently become part of the adolescent and adult picture of the disorder as the result of cumulative insults to self-esteem and failure in social and educational settings.

5.19.5.1 Genetic Risk Factors

Although no single marker for genetic transmission of ADHD has been found, comprehensive review of the neurobiological literature strongly suggests genetic and biological explanation for the syndrome (Zametkin, 1989). Overviews of the literature on genetic, biochemical, neurobehavioral, and neuroanatomical correlates of ADHD suggest that motor-regulatory systems involving subcortical and frontal systems are associated with ADHD. These studies also suggest a possible genetic basis because they implicate deficiencies in several neurotransmitters and because family studies show some concordance between affected parents and their children (Hynd, Hern, Voeller, & Marshall, 1991; McMahon, 1981).

Controlled case studies provide strong suggestions of genetic influences on hyperactivity and disruptive behavior (Heffron, Martin, & Welsh, 1984). Furthermore, more recent research applying powerful multiple regression techniques strongly suggests that ADHD is highly heritable (Gillis, Gilger, Pennington, & DeFries, 1992), with heritability estimates of 0.87–0.98. These estimates were similar to more recent data from hyperactive twins studied in London. Data from mixed-sex pairs did not support a two-threshold genetic explanation for the male excess of hyperactivity. The link between adverse family factors and hyperactivity was weak (Stevenson, 1992). Evidence suggests that only some forms of ADHD may be genetic, with the milder (situational) types showing less genetic influence and more association with antisocial behavior. These findings cast doubt on the validity of combining situational and pervasive hyperactivity into a single diagnostic category of ADHD, and also support a frequent finding that antisocial behaviour carries a different genetic pattern than ADHD (Goodman & Stevenson, 1989a, 1989b).

Family-genetic studies show that there are a number of associated risks shared with ADD, including conduct disorders, anxiety, depression, antisocial behavior, alcoholism, maternal personality disorders, and proneness to substance abuse in the first-degree relatives of ADHD (Cantwell, 1988). Relatives of ADD probands have a higher morbidity risk for ADD, antisocial disorders, and mood disorders than relatives of psychiatric and normal controls. The increased risk for ADD, however, cannot be accounted for by gender or generation of relative, the age of proband, social class, or the intactness of the family (Biederman, Faraone, Keenan, Knee, & Tsuang, 1990). Similar risks are found in ADHD girls as found in boys (Faraone, Biederman, Keenan, & Tsuang, 1991a). There are particularly high risks for offspring of parents with conduct, oppositional, or antisocial disorders (Faraone, Biederman, Keenan & Tsuang, 1991b). Further evidence of the genetic link to behavior problems comes from studies using parent ratings rather than diagnoses (O'Connor, Foch, Sherry, & Plomin, 1980). A classical twin analysis compared intraclass correlations for 54 pairs of identical twins and 38 pairs of same-sex fraternal twins, using a revised form of a parent symptom rating questionnaire (Conners, 1970). For seven of eight factor scales, correlations were significantly greater for the identical than for the fraternal twins. The identical twin correlations were similar to the test–retest reliabilities of the scales, suggesting that nearly all of the reliable variance for these behavioral problems was shared by identical twins.

There is a surprising link between familial ADHD and anxiety disorders (Biederman, Faraone, Keenan, Steingard, & Tsuang, 1991). Risk of anxiety disorders was twice as high in relatives of ADD when accompanied by an anxiety disorder than in those of ADD without anxiety disorder; there was also a tendency for relatives who had ADD to have a greater risk for anxiety disorders than relatives without ADD. Similar findings hold for affective disorders (Biederman, Newcorn, & Sprich, 1991). Some genetic risks may involve specific influences upon the early development of the fetus, as evidenced by the incidence of dysmorphology in the first-degree relatives of ADHD (Deutsch, Matthysse, Swanson, & Farkas, 1990). These studies suggest genetic heterogeneity for the disorder, and the possible value of dividing the disorder into more than one biological subtype.

Comorbidity of ADHD and learning disorders is high, with estimates ranging from 20% to 50%. It is not yet clear whether ADHD and reading disorders (RD) are genetically related or independent. Monozygotic and dizygotic twin pairs, in which at least one member of each pair was RD, were assessed for ADHD. Trends in the data and subtype analyses suggest that in some cases RD and ADHD may occur together because of a shared genetic etiology, and that a genetically mediated comorbid subtype may exist (Gilger, Pennington, & DeFries, 1992). Moreover, the specific type of learning disorder may be genetically determined, with phonological coding (i.e., sound-based reading) showing high heritability, while orthographic coding (i.e., visually based reading) shows no heritability. Phonological coding was substantially lower for most ADHD children with reading disabilities, indicating a developmental deficit in phonological coding rather than an equal developmental lag across all component reading skills (Olson, Rack, Conners, DeFries, & Fulker, 1991; Olson, Wise, Conners, & Rack, 1990; Olson, Wise, Conners, Rack, & Fulker, 1989).

5.19.5.2 Medical and Environmental Risk Factors

Medical factors are sometimes important because of their intertwined relationship to parenting variables and other aspects of social development. For example, low birthweight influences not only physical and cognitive competencies, but the behavior of mothers towards the child (Landry, Chapieski, Richardson, Palmer, & Hall, 1990). In a similar study (Lewis & Bendersky, 1989), independent impacts of severe intraventricular hemorrhage, other common medical complications of prematurity, and socioeconomic status, on the development in the second year of life were moderated by both the socioeconomic circumstances and the degree of medical complications.

Medical services are highly overused by ADHD children and families, even though mental health services are underused by ADHD children (Offord et al., 1987; Szatmari et al., 1989a). One reason may be that ADHD children have more actual psychosocial events leading to medical care than normal children. In one study they had more than twice as many actual medical and psychosocial events involving significant threat to the child. The causal direction of some of these risks is uncertain since many of the events may reflect the impulsivity and risk-taking behaviors of the children, and adverse environmental variables may be correlated through their relationship with comorbid conditions (Jensen, Shervette, Xenakis, & Bain, 1988).

5.19.5.3 Metabolic and Nutritional Factors

Metabolic and nutritional explanations of hyperkinesis have been around at least since 1929 when a report described a 12-year-old boy who,

> had been termed by psychiatrists and psychologists a "nervous hyperkinetic child." To use his own words, "I have to keep on going and be doing something ... " His diet was meager in meats and vegetables and rich in candy and cookies. This inadequacy of diet allied with hyperkinesis suggested the treatment. (Seham & Seham, 1929, p. 1)

Although comprehensive reviews of the role of nutrients, food additives, and sugar generally indicate a somewhat limited basis for inferring relationship to the etiology of ADHD (Milich, Wolraich, & Lindgren, 1986; Wolraich, Wilson, & White, 1995), current evidence suggests a complex intertwining of nutritional variables, such as sugar intake, exposure to environmental toxins, and other risk factors. Consider, for example, the role of cadmium and zinc. Cadmium is frequently inhaled from sidestream smoke by children (i.e., secondary smoke from smoking adults in the near environment). A large study of hair cadmium on the eastern shore of Maryland (Thatcher, McAlaster, Lester, & Cantor, 1984) found that children with high cadmium levels had lower IQ and poorer reading ability than those with low cadmium levels. However, this effect did not hold for children who regularly ate whole-wheat bread. Statistical controls ruled out demographic variables as causes. This unlikely phenomenon was found to depend upon the intake of zinc in the wheat kernel. The cadmium effect was inversely related to the levels of zinc measured in the children. The authors concluded that dietary zinc protects against the toxic effects of cadmium. Dietary intake of carbohydrate was also inversely related to the levels of zinc, a replicated finding in several studies for both micronutrients and vitamins (see below for further discussion of dietary risk factors). Thus, excess dietary sugar was found to have an important negative effect via its impact upon intake of needed dietary zinc. Even with socioeconomic status (SES) and other background factors controlled statistically, a multiple regression indicated that 20% of variance in reading level was accounted for by the combination of cadmium and the dietary factors.

5.19.5.4 Lead

The Centers for Disease Control (CDC) have published a major revision of their 1985 statement on prevention of lead poisoning in children (CDC, 1991). New evidence supports DHHS Secretary Louis Sullivan's assertion that lead is the number one environmental health threat to American children [Letters to the Editor, 1992]. The evidence linking low levels of lead exposure (10 μg or more per deciliter of whole blood), and a constellation of neurotoxic and other adverse effects in young children, is compelling. One study found that each 10 μg dl^{-1} increase in lead lowered IQ by eight points. The importance of lead as a subtyping factor in hyperactive children has been somewhat controversial, but considerable evidence suggests a relationship (David, Hoffman, Sverd, & Clark, 1977). Lead becomes especially important as the range of sociodemographic factors is increased and the risks of exposure are greater. Living in a crowded, urban neighborhood with high lead levels in older houses close to the increased lead deposition from automobile fumes, in combination with lower levels of nutrients such as calcium and iron, combines multiple risks. Calcium and iron protect against lead assimilation into the brain and gut, but are more likely to be absent from the child's diet if mothers (especially single mothers) start them earlier on cow's milk in order to return to work. Since the iron in cow's milk is less assimilable than breast milk, socioeconomic and nutritional risks interact. This form of complex interaction among risk factors may be the rule rather than the exception, making specific etiological or causal factors in ADHD difficult to disentangle.

5.19.5.5 Vitamins and General Nutrition

Several studies suggest a complex path relating hyperactivity, inattention, and cognitive function to nutritional status. Benton and Roberts (1988) reported that taking a vitamin and mineral supplement over an eight month period increased the nonverbal intelligence of 12-year-old British schoolchildren. The study was a well-designed double-blind placebo-controlled trial. This trial attracted considerable attention, and several replication studies appeared. A large study involving 615 US schoolchildren was carried out over a 12 month period of supplementation. The tablets contained either placebo or vitamins and minerals at 50%, 100%, and 200% of the US recommended daily allowance. Significant improvement in nonverbal intelligence was found for the 100% supplementation group compared with placebo. Equally important is the replicated finding of decline in performance IQ in the placebo-treated group (Benton & Buts, 1990). The cause of the apparent increase in nonverbal IQ with supplementation, and decline with the passage of time, is unclear.

Benton and Cook (1991) reported a study of 47 six-year-old children with a randomized double-blind procedure aimed at finding the nutritional and behavioral factors responsible. They found a 7.6 point IQ gain in the supplement group compared with a decline of 1.7 points in the placebo controls. Nutritional analyses revealed a strong positive relationship ($r = 0.55$, $P < 0.004$) between the amount of sugar in the diet and the change in IQ scores for three different IQ tests. The authors interpreted the findings to mean that high sugar intake is associated with vitamin deficiencies, which in turn result in lower IQ test results, mediated by effects upon attention and distractibility during testing. Supplementation is therefore more effective in children with the worst diets, that is, those who ingest excessive levels of carbohydrates and sugar. The argument was supported by direct observation by trained observers blind to the experimental conditions, showing that fidgeting and frustratibility during IQ testing showed a significant improvement in the supplemented group but not the placebo group.

The fact of excess sugar intake among children presenting for clinical services is indisputable. Its interpretation, however, is problematical. Dietary factors (e.g., high carbohydrate intake) may account for poor nutrition, but might also represent some correlated effect from social class or family characteristics. For example, the fact that children in the lowest quartile of attention are in the highest quartile of refined sugar intake was found to be related to parental license in giving access to the refrigerator (Prinz & Riddle, 1986). Available evidence of this type does not make it possible to determine (i) whether parents who are poor limit-setters produce uncontrolled children whose diets simply reflect another area in which the children lack self-control, or (ii) whether lack of dietary control leads to poor nutrition and thence poor cognitive and behavioral control.

These issues assume particular importance in patients from lower income groups since it is well known that poorer education and lower income are associated with poorer dietary status. We believe that there is a subgroup of ADHD children who are defined by the facts that they come from the lower part of the SES distribution. They routinely have poor dietary intake of vitamins and micronutrients; have relatively high carbohydrate to protein intake, and are lower in performance IQ than equivalent SES patients with good nutritional status. Nutritionally defined subgroups within the lower SES category may be an important predictor of developmental course and stimulant treatment response.

5.19.5.6 Temperament as a Risk Factor for ADHD

It has long been suspected that normal variations in temperament can dispose to later problems of ADHD. Certain extremes of normal temperament might become disorders when they conflict with parental styles or environmental requirements, despite being normal biological variations. There is little doubt that activity level is highly heritable and is normally distributed in the population (Conners & Kronsberg, 1985). In one study, Boudreault and Thivierge (1986) examined the relationship between temperament and externalizing- or internalizing-type behavior problems in school in 295 male and 301 female seven-year-olds. A temperament questionnaire was filled out by parents and the Conners Teacher Questionnaire was completed by teachers twice over a four week interval. The results indicated that a temperamental pattern of high activity and high approach with new stimuli predisposed boys to present externalizing behavior problems in school, whereas the temperamental traits of low adaptability or withdrawal with new stimuli predisposed girls to internalizing behavior problems in school. Others report relationships between parent ratings of adjustment problems and early temperament (Brody, Stoneman, & Burke, 1988). Long-term follow-up studies suggest that the adult outcome of ADHD may be related to early temperament styles (Hechtman, 1991). Predictions of long-term outcome based upon early temperament holds up in follow-up studies of epidemiological samples (Lambert, 1988).

One novel hypothesis considers vigilant attention as a temperamental trait. Weinberg and Brumback. (1990) suggest that the trait of vigilance has behaviors which overlap those of attention deficit–hyperactivity disorder. Vigilance is the state of being watchful, awake, and alert. Motor restlessness (fidgeting and moving about, yawning and stretching, talkativeness, or a combination of these) may be a method to improve alertness when sitting or standing still or when involved in tasks requiring continuous mental performance. In order to stay awake when prevented from being active, persons with lowered vigilance will stare off, daydream, show minor hyperactivity, and finally may fall asleep. They will also complain of boredom and monotony and usually avoid or lose interest in structured or repetitive activities. This trait is viewed as a dominantly inherited condition with onset in early childhood and worsening with age. Persons with this "primary disorder of vigilance" are described as having a remarkably kind and caring temperament. This description could be seen as simply another way of viewing ADHD; however, casting it as an inherited temperament trait has heuristic value since it shifts the search for etiology to extremes of normal temperament variation, rather than to a pathological process.

Short sleep durations and high ADHD symptom ratings appear to be features of children with a difficult temperament or high motor activity when awake and asleep. Thus, some children with colic (which is a common finding in the early history of ADHD children), may represent an extreme of a normal temperament trait (Weissbluth,1983, 1984).

5.19.5.7 Overview of Etiology and Risk Factors

It seems obvious that many different genetic, medical, temperamental and environmental risks can produce symptoms indistinguishable from ADHD. If extreme enough, virtually any single factor can, by itself, lead to the syndrome. However, it is seldom the case that single risk factors occur; rather, it is not uncommon to see family-genetic, medical, and other risks occur in combination. Development is transactional, so that risk factors create effects which then become a part of the ADHD picture; this new picture then leads to further complications, and so on. The high risks associated with maternal smoking or alcohol use (Nichols & Chen, 1981) are themselves linked to certain family psychiatric disorders as well as to the family environment of the mother and the toxic smoking environment of the child. In our view, then, the task of assessment in ADHD must focus on the understanding of the past and present risks, protective factors, and their interaction. The task of treatment is to maximize the protective factors while applying available remedies aimed at the biological substrata of behavior.

5.19.6 DEVELOPMENTAL COURSE AND OUTCOME

Results of longitudinal studies reported in the 1980s and early 1990s dramatically altered the prevailing view of ADHD as a disorder with symptoms that remit following puberty. In contrast to this putative benign course, these prospective investigations have revealed ADHD to be a typically chronic condition, marked by the persistence of symptoms into adolescence and often into adult life for the majority of children diagnosed with the disorder (Klein & Mannuzza, 1991; Weiss & Hecthman, 1993).

More than two-thirds of adolescents who were diagnosed with ADHD (by whatever label) in childhood continue to meet diagnostic criteria for the disorder, although the manifestations of the syndrome are likely to change over time, as discussed above (Barkley et al., 1990; Gittelman et al., 1985; Hart et al., 1995). Beyond the continuity of ADHD features, this group is characterized by rates of academic underachievement, school failure, low self-esteem, peer rejection, substance abuse, delinquency, and family disharmony that exceed those found among control subjects (Barkley et al., 1990; Faigel, Sznajderman, Tishby, Turel & Pinus, 1995; Fischer, Barkley, Edelbrock, & Smallish, 1990; Gittelman et al., 1985).

This pattern of persistent symptoms and heightened risk for negative outcomes across a broad spectrum of functional domains also appears to apply to at least half of the adults who received ADHD diagnoses as children. Investigations of this population are fewer and results more variable than for adolescents. Nevertheless, the emerging literature indicates the persistence of impairing ADHD symptomatology in a majority of cases along with generally lower levels of educational and occupational attainment and higher levels of job instability, substance abuse, antisocial behavior, and psychological distress than found in matched control groups (Barkley, Murphy, & Kwasnik, 1996; Biederman et al., 1995; Mannuzza, Gittelman-Klein, Bessler, Malloy, & LaPadula, 1993; Weiss & Hecthman, 1993). However, a number of carefully conducted studies have not found elevated rates of mood or anxiety disorders in adults with histories of ADHD (Gittelman et al., 1985; Mannuzza et al., 1993).

Despite this sobering outcome for many children with ADHD, it is important to note that a substantial proportion do indeed experience symptom remission and develop into well-adjusted, productive adults. Weiss and Hechtman (1993) estimate that up to one-third or more of children with ADHD may become indistinguishable from normal controls during adulthood. Thus, the heterogeneity that characterizes so many aspects of ADHD also applies to its course and outcome. Predictors of outcome, however, defy easy identification and are likely to involve complex interactions between a host of child, family, and ecological variables. The confounding of various predictors in many ADHD children can lead to misleading conclusions. For example, evidence suggests that the aforementioned link between childhood ADHD and later substance abuse is indirect; ADHD behaviors only increase the risk for substance abuse when they are accompanied by antisocial behavior patterns in childhood and it is the antisocial behavior rather than the early ADHD that predicts substance abuse (Lynskey & Fergusson, 1995). Similarly, aggression or other signs of conduct disorder may mediate the association between ADHD and later criminality and antisocial spectrum disorders (Hinshaw, 1994; Klein & Mannuzza, 1991). Clearly, the links between childhood externalizing symptoms and a variety of outcomes may be highly specific yet difficult to elucidate in the context of complex, multivariate causal chains.

5.19.7 CONCEPTUALIZATION AND CLINICAL FORMULATION

5.19.7.1 Multimodal Assessment

There has been no dearth of assessment research in ADHD. A recent synthesis of this literature commissioned by the Office of Special Education evaluated over 1300 references

(McKinney, Montague, & Hocutt; 1993). Along with many others (Angold, 1989; Barkley, Fischer, Newby, & Breen, 1988; Luiselli, 1991; McConaughy, 1993a; Morris & Collier, 1987; Ostrom & Jenson, 1988; Schaughency & Rothlind, 1991a), the authors of this synthesis assert that ADHD assessment should involve multiple informants, rely upon multiple technologies (ratings, interviews, direct observations, laboratory measures, and self-reports), cover multiple domains (home, school, peers, intra- and interpersonal functioning), and allow for a functional analysis that specifies treatment needs.

The major problem with these multilevel assessments is how to integrate information across informants and domains in such a way that the needs of both diagnosis and treatment formulation are served. Achenbach and colleagues made an attempt to formalize such a system using a multiaxial, cross-informant typology (Achenbach, 1992; McConaughy, 1993a). That approach has spawned vigorous debate (Macmann, Barnett, & Lopez, 1993; Martens, 1993; McConaughy, 1993b) and empirical validation has not yet occurred. In general, a best-estimate diagnosis is recommended by most authors, though rules for decision making and integration of information are usually left unspecified.

Experienced clinicians typically recommend a comprehensive evaluation that includes (i) a history of pregnancy, delivery, and developmental milestones from infancy; (ii) a family history of psychiatric disorders; (iii) assessment of specific symptoms, including their severity and frequency, situational specificity, and duration; (iv) an educational assessment which covers both academic functioning and classroom behavior; (v) intrapsychic processes, including self-image and sense of efficacy with family, peers, and school; (v) child and family interaction patterns and family structure; and (vii) neurological status when indicated by other evidence (Weiss & Hechtman, 1993). Although such an elaborate process provides a wealth of data, it does little to set boundaries around the disorder or to specify which treatments, in which combinations should be given; and it provides no clear operational methods for determining when treatment works.

Concerns about operational measures and lack of norms regarding attentional processes, and lack of consensual diagnostic criteria have prompted some to caution school psychologists against the diagnostic enterprise altogether (Ostrom & Jenson, 1988). They argue that there is a lack of a consensual definition of the disorder; that empirical research and theory have not isolated the particular attentional processes in question; and that there are no adequately normed measures of attentional processes alleged to cause the disorder. Instead, they recommend identification of the pattern of strengths and weaknesses, direct observations in the classroom with a normal control (response discrepancy), and a multigating approach (teacher identification, selection of the most severe cases, followed by school psychologist's evaluation of on-task behavior in the classroom and positive and negative social interactions at recess). This represents a more practical approach within schools and is responsive to treatment needs. However, from a broader perspective, this approach ignores the chronic long-term nature of ADHD, the role of medical treatment needs, and the treatments relating to home environment and social processes outside the school. It makes the assumption that all that is going on with these children is available to the naked eye, and that neurocognitive processes can be readily managed by strictly behavioral means. Throwing away everything we have learned about the biosocial nature of this disorder seems a heavy price to pay for the clarity of a strictly behavioral approach.

McKinney et al. (1993) were also disappointed at the paucity of studies assessing educationally relevant variables that specify how inattention, impulsivity, and overactivity impair learning on specific instructional tasks and in different educational settings. In response to this challenge, DuPaul and Stoner (1994) recommend a five-stage process of school-based assessment which starts with a teacher complaint involving one or more of the three main symptom categories (inattention, impulsivity, or overactivity): (i) use of normed teacher ratings of ADHD symptoms, followed by a multimethod assessment involving parent and teacher interviews; (ii) review of school records; (iii) behavior rating scales; (iv) observations of inclass behavior; and (v) academic performance data. A diagnostic formulation is made based on the number of ADHD symptoms, their deviance from age and gender norms, age of onset and chronicity, pervasiveness across situations, degree of functional impairment, and the presence of other disorders. A treatment plan is then developed which takes into account the severity of the symptoms, a functional analysis of behavior, the presence of associated disorders, response to prior treatment, and community-based resources. This is then followed by periodic collection of assessment data and a revision of the plan as necessary.

This approach is comprehensive, incorporates most current knowledge regarding ADHD, and reflects state-of-the-art clinical practice. However, it is both costly and potentially burdensome for the school system to undertake and inevitably extend services into the broader mental health system. An unanswered question is whether a straightforward approach which starts with the teacher's complaint, observes the behavior in the classroom, and immediately applies behavioral interventions without a broader diagnostic process involving labeling works as well as the more comprehensive five-stage process outlined by DuPaul and Stoner. Here, the assessment would be limited strictly to procedures which impact upon intervention, such as rates of deviant behaviors or academic functioning using curriculum-based measurement. An example is an eight-year-old boy with ADHD who attended regular third grade classes at his local elementary school. During academic instruction, he frequently displayed disruptive behaviors that rarely occurred during the rest of the school day. A brief functional analysis showed that disruptive behaviors were maintained by escape from task demands. Curriculum-based assessment revealed that disruptive behavior occurred most frequently during specific seating and grouping arrangements. Finally the effectiveness of an intervention derived from the assessments showed an immediate reduction in disruptive behavior and an increase in appropriate behavior that lasted throughout the data collection period (Umbreit, 1995).

Schaughency and Rothlind (1991b) also recommend a multimethod approach to assessment directed at answering four questions: Does the child meet *DSM* diagnostic criteria? Does an alternative diagnosis or conceptualization account for his or her difficulties (differential diagnosis)? Does the child display the behaviors to a developmentally inappropriate extent? Do these behaviors impair the child's functioning in school, social relationships, and/or in the home? They suggest a structured diagnostic interview with multiple informants to answer the first and second questions (along with IQ and achievement data, if needed), and information on developmental course and onset; behavior rating scales with multiple informants to answer the third question; and sociometrics and archival data (referrals, grades, classroom performance, etc.) to answer the final question. They recommend a best-estimate diagnostic approach, and also agree that we do not yet know what aspects of attention are disordered or how to measure them. They argue that the better performance of ADHD children in structured situations limits the utility of structured tasks for assessment. Their recommendations seem appropriate for younger children, but omits self-report information which is necessary in older children and adolescents (Conners & Wells, 1985). Self-report becomes increasingly important as children get older and develop more comorbid internalizing symptomatology. Their suggestion of using a lay interviewer instrument such as the Diagnostic Interview for Children and Adolescents or Diagnostic Interview Schedule for Children has the major disadvantage that these instruments significantly overdiagnose psychiatric disorders. This would be likely to significantly overburden the already-strained resources of the schools. The use of sociometrics (Hops & Lewin, 1984), although ecologically valid, has enormous practical problems, not the least of which is the necessity of getting informed consent from parents of both the peers and the target child.

Although multimethod/multisource information could theoretically prevent some of the excesses of diagnosis which occur with single-instrument methods, the question of whether the limited incremental validity and costs of these approaches (Homatidis & Konstantareas, 1981) can be justified in school settings and managed-care environments remains to be seen. In any case, most would agree that assessment for diagnosis, and assessment for treatment, although overlapping, are not the same thing and therefore require complementary methods. Structured interviews provide information needed to determine the diagnosis, but provide little guidance regarding target behaviors or modes of intervention. To address the latter, specific classroom behavior and academic problems must be identified, along with the nature of the home environment, particularly with regard to parenting practices and skills. If more extended contact is impossible, as often happens for either economic or geographic reasons, one provides parents with a list of resources which includes information regarding token management systems, self-monitoring skills, behavior contracting, schools and the law, books, legal resources, classroom modifications which work with most ADHD children, resources for parenting skills training, and parent support groups.

Our own clinical practice in a university medical school setting typically proceeds by collecting as much "free" information ahead of time for which the client is not billed, including a biographical history filled out by parents, parent and teacher rating scales which are computer-scored and analyzed (Conners, 1989); a developmental history and temperament scale (Conners & March, 1993); a multidimensional anxiety scale (March, Sullivan, & Conners, 1996); and a child depression scale (Kovacs' Child Depression Inventory, 1993). Office interviews are then conducted with the parents using a semistructured *DSM-IV* interview for ADHD, ODD, and CD if screening questions indicate that the latter is relevant (Conners, Erhardt, & Sparrow, in press); and an interview regarding family-genetic, medical, temperamental, and environmental risk factors (Conners, in press). While this interview is being carried out, computerized testing is usually conducted, including a continuous performance test (Conners, 1994a, 1994b; Conners, March, Fiore, & Butcher, 1994). This process takes 1–2 hours of clinical time. A separate session is usually held to provide feedback and initiate treatment plans. Whether or not the child meets any criteria for a *DSM* disorder, specific behaviors or academic problems which triggered the referral must be assessed and treated. Sometimes this amounts to simple reassurance in overanxious parents; but often the nonpharmacologic interventions will be the same in these very mild cases as that for someone meeting full diagnostic criteria. Some authors argue that diagnosis is therefore a waste of time (Ostrom & Jenson, 1988).

5.19.7.2 Treatment Formulation

The heterogeneity of the ADHD population with respect to presenting symptoms and comorbid conditions dictates that intervention strategies be individually tailored and based on results of a comprehensive clinical assessment. Indeed, despite its methodological shortcomings, the multimodal treatment study conducted by Satterfield and colleagues suggests that long-term, highly individualized treatment plans devised on the basis of detailed assessments may improve long-term outcome for ADHD youngsters (Satterfield, Satterfield, & Schell, 1987). As emphasized by Mash and Terdel (1988), assessment and treatment should be highly integrated and mutually informing clinical activities. Formulations regarding the contexts within which interventions need to occur (e.g., home, school, or peer group) and the nature of those interventions (e.g., pharmacologic, behavioral, or pedagogic) should develop during the course of the evaluation. After intervention has begun, formal periodic assessments should take place as a means of monitoring the child's status relative to the baseline evaluation in order to inform decisions about the need to modify, add, or terminate treatments. For example, an intensification of school-based behavioral programs may be indicated on the basis of an assessment documenting an increase in teacher-reported classroom misbehavior.

The selections of specific treatments for ADHD children cannot at present be informed by any known cure, certain or simple etiology, or any consensus regarding the underlying nature of the disorder. Thus, clinicians must rely whenever possible on interventions with empirically validated efficacy for the primary treatment targets identified through the comprehensive assessment. Hence psychostimulant medications and behavior therapy delivered through parents, teachers, or direct contingency management approaches are the likely cornerstones of treatment programs for ADHD. The need for adjunctive interventions to complement these core treatment components should be anticipated for many children with ADHD. Moreover, clinicians should endeavor to clearly delineate the relevant targets for each treatment. For example, improved attention to tasks may be identified as a target of pharmacotherapy, whereas reduced oppositionality may represent a goal of parent training. The specific foci of treatment should reflect the chief presenting concerns of the parents and/or child and the deficits or problem areas uncovered by the assessment, with particular emphasis on those features that are most impairing to current functioning and most highly associated with poor prognosis. Thus, reducing aggressive behavior or redressing academic deficits may be prioritized as treatment goals despite such problems representing associated rather than primary features of ADHD.

Common to each planned treatment component should be programmatic efforts to promote generalization of gains across behaviors, settings, and time (Conners et al., 1994). Extending the duration of treatments (e.g., years versus months), which is consistent with the chronic nature of ADHD, may contribute to temporal generalization. Improved generalization may also result from incorporating significant others into treatment (e.g., parents, teachers, or peers), using diverse training experiences, and providing booster sessions.

Decisions related to the sequencing and cessation of various pharmacologic and psychosocial treatment components are complex and cannot as of yet be informed by a body of evidence from controlled studies. Ultimately, such decisions must be made on the basis of clinical judgment, following consideration of relevant factors such as symptom severity, age and developmental level of the child, comorbid symptoms or conditions, the palatability of various treatments to the child and family, and the extent of available resources, both within the family and the community. Some argue that behavioral interventions should always be initiated prior to medications (Hoza, Vallano,

& Pelham, 1995) but the demonstrated short-term efficacy of stimulant treatment in conjunction with its low cost and ease of administration relative to psychosocial interventions may provide some justification for introducing pharmacotherapy as the initial treatment component. Moreover, a favorable stimulant response may enhance the efficacy of subsequent interventions, although this has yet to be demonstrated empirically.

A final point related to treatment planning involves the coordination of clinical services. The adequate treatment of children with ADHD is typically a complex and lengthy process, the many facets of which are unlikely to be within any given practitioner's scope of competence. It is thus necessary for the individual clinician to be familiar with community resources that may be of benefit to ADHD children and their families and to establish the relationships with professionals from other disciplines that are necessary to develop coordinated, multimodal treatment plans.

5.19.8 TREATMENT

Two primary goals should guide the selection and implementation of interventions for children with ADHD: (i) optimizing current functioning across behavioral, academic, family, and social domains by reducing the severity and impact of core and associated symptoms; and (ii) prevention of the adverse outcomes known to result for many of these youngsters by intensively targeting those problem areas (such as aggression and other antisocial behaviors, academic underachievement or failure, family dysfunction, or peer rejection) most strongly associated with poor prognosis. In the light of these overarching goals and the multifaceted problems experienced by most children with ADHD, consensus has emerged that no single treatment is likely to be sufficient for addressing the acute and chronic clinical needs of this population. Clinical and research efforts have moved towards the identification of multimodal treatment packages with the hope that active components will interact in a complementary or synergistic fashion to maximize therapeutic effects (Abikoff & Hechtman, 1996; March, Wells, & Conners, 1995b; Pelham & Murphy, 1986; Richters et al., 1995).

5.19.8.1 Psychosocial Interventions

Despite the well-established short-term clinical efficacy of psychostimulant medications (March, Erhardt, Johnston, & Conners, 1995; Swanson et al., 1993), alternative or adjunctive psychosocial interventions are invariably required for one or more of the following reasons: (i) medications are not an option for a minority of children with ADHD due to nonresponse, intolerable side effects, or parent refusal; (ii) gains achieved by medication often fail to normalize functioning in the short term and have yet to be associated with long-term benefits; (iii) many associated symptoms of ADHD (e.g., academic skill deficits or family disharmony) are not appropriate targets for pharmacotherapy (as reflected in the adage "pills don't teach skills"); (iv) the restriction of stimulant-related benefits to times when the child is actively medicated often results in a need for other interventions to be in place during evenings, weekends, and vacations; and (v) the use of combination treatments may permit lower intensities or dosages of each component (e.g., lower medication dosages resulting from the presence of concurrent contingency management programs).

Prior to discussing a selection of the wide range of psychosocial treatments which have been applied to children with ADHD, we wish to highlight that behaviorally oriented interventions enacted through parents, teachers, or direct contingency management programs in the classroom represent the most widely applied and empirically supported psychosocial interventions for ADHD (Hinshaw & Erhardt, 1993; Pelham & Hinshaw, 1992). Although less common and less efficacious with ADHD populations, cognitive-behavioral interventions, social skill training programs, and individual child treatment also warrant comment due to their potential utility as adjunctive components in multimodal treatment packages. There are a number of additional interventions with limited or no established efficacy for ADHD populations, including dietary management, electroencephalographic biofeedback, relaxation training, ocular exercises, chiropractic methods, megavitamins, and sensory integration training.

5.19.8.1.1 Parent-based interventions

Following their successful application to children with oppositionality, aggression, and other conduct problems (Wells, 1995), parent training (PT) approaches were extended to ADHD populations in the early 1980s and have emerged as the most prevalent psychosocial treatment for the disorder. Two objectives which characterize PT programs for ADHD are educating parents about the disorder and teaching a comprehensive set of behavior management skills so that parents may act as change agents to reduce negative behavior and promote adaptive functioning in the natural environment. Efforts are also geared towards improving the general family emotional climate and altering coercive parent–child interactions which may contribute to the exacerbation of ADHD symptoms and oppositional behavior (Barkley, 1990; Campbell, 1990).

Detailed, session by session descriptions of PT programs are available (Anastopoulos & Barkley, 1990; Barkley, 1987; Hembree-Kigin & McNeil, 1995; McMahon & Forehand, in press; Newby, Fischer, & Roman, 1991). Despite some variation in format, duration, and content, PT programs for ADHD generally rely heavily on therapist modeling, behavior rehearsal (e.g., role playing), and homework and include the following components: (i) psychoeducation on ADHD; (ii) introduction to social learning theory and general behavior management principles; (iii) positive reinforcement skills (e.g., positive attending, "catch your child being good," promoting independent play, or effective ignoring); (iv) effective commands; (v) point or token economy systems; (vi) nonphysical punishment procedures (e.g., response cost or time out); and (vii) managing behavior outside of the home. (A valuable augmentation to the psychoeducational component of PT is parent involvement in an organization designed to support and provide information to families with children with ADHD and related disorders (e.g., Children and Adults with Attention Deficit Disorders).) Innovative additions to some recent PT programs have included stress and mood management for parents, interventions to enhance marital functioning, academic support skills for the home, inputs to improve peer functioning, and training parents to advocate effectively for their child at school (Wells, in press). Specialized interventions incorporating aspects of behavioral, cognitive, and family systems approaches have been developed for families with ADHD adolescents (Robin & Foster, 1989) and found to yield improvements to family functioning (Anastopoulos, Barkley, & Shelton, 1996).

Controlled studies of PT have found it to reduce primary and associated symptoms of ADHD and to have favorable effects on parenting stress and parenting self-esteem (Anastopoulos et al., 1996; Anastopoulos, Shelton, DuPaul, & Guevremont, 1993; Wells, in press). Despite their documented efficacy for ADHD, PT approaches do not appear to achieve the same magnitude of behavioral effects as pharmacotherapy, typically fail to normalize functioning, and place high demands on parents that can reduce compliance and maintenance of treatment effects (Hinshaw, 1994; Pelham & Hinshaw, 1992).

5.19.8.1.2 School-based interventions

Some form of classroom intervention is almost invariably indicated for children with ADHD due to the adverse impact of inattentive, overactive, impulsive, and associated symptoms on learning, performance of acquired skills, and adherence to class rules and routines. School-based treatment for ADHD children may comprise both academic (e.g., tutors or special education services) and behavioral interventions. The former (DuPaul & Stoner, 1994) are not reviewed here but are often indicated due to scholastic skill deficits and possible comorbid learning disorders. Rather, we focus on behavioral interventions designed to improve the child's ability to meet both comportment and academic demands at school.

A host of classroom interventions for ADHD address antecedents of behavior by making modifications to the classroom environment, materials, and instructional techniques in an effort to create a "prosthetic" environment that helps to minimize the impact of ADHD symptoms (Abramowitz, 1994; Pfiffner & O'Leary, 1993). Beneficial changes to classroom structure may involve establishing clear and salient class rules, arranging seating so that the child is distanced from likely distracters and surrounded by well-behaved peers, instituting highly consistent routines (e.g., for turning in completed homework), and insisting that students use a daily assignment book. Students with ADHD often show higher rates of on-task behavior and task completion through the use of timers, assignments that are shortened or divided into smaller parts, checklists outlining the steps involved in a task, and alternative response methods (e.g., oral reports and videotapes). Moreover, the effectiveness of instructional methods is likely to be enhanced when they incorporate higher levels of activity, stimulation, and novelty than is typical for most classrooms.

Another group of school-based treatments for ADHD focuses on manipulating the consequences of behavior. In direct contingency management approaches, a behavior therapist establishes formal reinforcement systems with clear guidelines for behaviors that both earn and lose tokens, which may subsequently be exchanged for tangible reinforcers (e.g., access to a rewarding classroom computer). Numerous studies have documented the effectiveness of such methods for improving both academic productivity and classroom behavior (Ayllon & Rosenbaum, 1977; Hinshaw & Erhardt, 1993; Pelham & Hinshaw, 1992).

More common, however, are models in which therapists consult with teachers in order to educate them about ADHD and teach effective instructional and behavior management strategies. Much of the content in such interventions parallels that found in PT, albeit with adaptations to ensure applicability to the school environment. Thus, teachers learn to make positive attention contingent upon desirable behavior, issue effective commands and reprimands, provide high rates of prompts and feedback, establish token economies specifying individual or group-based contingencies, use response cost and time out as punishments, and enact many of the antecedent modifications discussed above. Teachers, like parents, are encouraged to provide a high ratio of positive to negative consequences but also cautioned that prudent negative consequences, such as response cost and time out, are often essential components of behavioral programs for youngsters with ADHD (Abramowitz, 1994). Detailed discussions of these and additional school-based interventions for students with ADHD, including daily report cards (DRC), behavior contracts, self-monitoring and self-reinforcement, study and organizational skill training, and group-based contingencies can be found in DuPaul and Stoner (1994), Parker (1992), and Pfiffner and O'Leary (1993). Additionally, Burcham, Carlson, and Milich (1993) review key features of educational practices for ADHD judged to be most promising in a project of the US Department of Education.

DRC systems merit comment because they represent a widely used and highly effective intervention wherein daily rewards are issued or withheld by parents at home contingent upon the child's behavior at school. DRC systems begin with teachers and parents collaboratively identifying behavioral and academic goals which are tailored to the unique needs of the child (e.g., "follows teacher instructions promptly," "completes and turns in homework," and "remains seated when required"). Next teachers complete a brief checklist (or DRC) at predetermined times throughout the school day that indicates whether the child met his/her goals. The DRC is sent home with the child daily and parents either provide or withhold rewards depending on the child's performance. When used alone or in combination with classroom reinforcement programs, the home–school contingencies enacted through a DRC represent a powerful means of improving both academic performance and classroom behavior (Abramowitz, 1994). Additional advantages include their low cost, ease and brevity of use, incorporation of more potent reinforcers than may be available at school, and provision of regular communication between teachers and parents.

Despite the significant benefits associated with the implementation of these behavioral school interventions in research protocols (Abikoff & Gittelman, 1984; Abramowitz, 1994), they continue to be underutilized in clinical practice. Obstacles to implementation include inadequately funded school systems, lack of training among both school and mental-health personnel to provide behavioral consultation to teachers, and demands on teachers' time which limits the practicality of some interventions (DuPaul & Stoner, 1994; Reid, Vasa, Maag, & Wright, 1994). Efforts to improve the ease of implementation and palatability for such interventions need to receive high priority in light of the typically unmet needs of students with ADHD and the legal rights of these children to adjunctive educational services (Davila, Williams, & MacDonald, 1991).

5.19.8.1.3 Child-based interventions

(i) Cognitive-behavioral methods

The prospect of addressing ADHD with cognitive-behavioral (CB) interventions initially inspired much enthusiasm due to the theoretical appeal of matching the disorder with a treatment approach designed to foster planful self-regulation. Investigations of many types of CB procedures, including verbal self-instructions, problem solving strategies, attribution retraining, self-evaluation, self-reinforcement, and anger control training, have emerged since the mid-1970s. As concluded in earlier reviews (Abikoff, 1991; Hinshaw & Erhardt, 1991), this line of research has generally failed to yield strong support for the application of CB methods to ADHD populations, particularly with respect to the much hoped for generalization of treatment gains across time and settings. Results for procedures based on self-instructional training methods have been particularly dismal.

Nonetheless, some positive outcomes have been found for anger management training and for procedures where the child is taught to monitor, evaluate, and reinforce themselves with respect to particular target behaviors (Hinshaw & Erhardt, 1991). Anger-coping programs, which are typically conducted in a group format, focus on teaching children to recognize their unique anger cues, to develop cognitive or behavioral plans for coping with anger adaptively, and to practice implementing those plans under realistic, provocative circumstances, such as being taunted by peers. However, at least two preconditions need to be met in order for these or any other cognitive intervention to yield clinically significant and generalized gains in ADHD youngsters: (i) key adults in the child's life (parents and teachers) must be taught to model, prompt, and reinforce targeted skills; and (ii) extrinsic reinforcement programs must be initiated to reward the child for the display of cognitive skills. Another rarely mentioned but useful application of a CB approach for this population involves employing cognitive restructuring methods to challenge and modify maladaptive beliefs about ADHD held by parents and teachers which may hinder their implementation of other treatments. Braswell and Bloomquist (1991) offer a detailed description of the clinical applications of CB procedures with ADHD children.

(ii) Social skill training

Despite the often glaring social failures of youngsters with ADHD, traditional social skill training programs, emphasizing the acquisition of prosocial behaviors, do not appear to be generally effective in improving the peer functioning of this population (DuPaul & Stoner, 1994; Whalen & Henker, 1992). This may well relate to the fact that youngsters with ADHD, as a group, do not appear to be deficient in the targeted social skills *per se*, but rather show problems in skill performance in natural social settings along with elevated rates of socially noxious behaviors.

Hinshaw (1992c) notes the broad array of potential foci for interventions designed to improve the social competence of children with ADHD, including decreasing negative behaviors known to be associated with poor peer status (e.g., aggression), teaching deficient social skills, enhancing self-monitoring of social behavior, altering social goals (e.g., emphasizing affiliation over dominance), teaching social problem-solving skills, promoting more accurate social information processing, and improving academic performance. He argues that meaningful and durable change in the social arena is only likely to result from multifaceted, individually tailored interventions that are delivered in naturalistic settings over long periods of time in combination with pharmacologic treatments and behavioral reinforcement programs.

Guevremont (1990) has developed a program for enhancing peer functioning that is more comprehensive than those typically employed. The approach combines social and CB skill training with programming to promote generalization and strategic peer involvement. Also promising are the multipush efforts to address peer problems contained in the intensive summer day camp treatment programs developed by Pelham (Pelham & Hoza, 1996). In addition to behavioral instruction of specific social skills (e.g., communication and conflict resolution), children are taught anger management strategies, trained in sports and game skills, and continuously reinforced for the display of appropriate social behavior in highly naturalistic classroom and playground settings. Moreover, a "buddy system" encourages the development of dyadic friendships, which may buffer some of the consequences of being unpopular in the larger peer group.

(iii) Psychotherapy

There is no compelling evidence to support the use of individual psychotherapy to address the core symptoms of ADHD (Hinshaw, 1994). Nevertheless, the experience of repeated failures across domains and frequent negative feedback from both adults and peers can lead to low self-esteem, demoralization, and negative self-image in many children with ADHD. Such problems do represent legitimate foci of expressive and CB psychotherapies which thus should be considered as adjunctive treatments for some youths with ADHD. Indeed, the multimodal treatment study conducted by Abikoff and Hecthman (1996) included a psychotherapy component for which a specific manual was developed. Greenfield, Gottlieb, and Weiss (1992) discuss some of the indications, challenges, and adaptations associated with conducting individual psychotherapy with children with ADHD.

5.19.8.2 Pharmacologic Interventions

Stimulant drugs such as Ritalin (methylphenidate) or dexedrine (dextroamphetamine) have been well-documented in their ability to bring about almost instantaneous behavioral changes in a wide array of symptomatic behaviors in ADHD. It is this remarkable alteration of behaviors, usually impervious to ordinary blandishments by parents and teachers, that helped make ADHD one of the most researched disorders in the behavioral sciences (as well as controversial in the public sphere). Contrary to popular belief, however, the pharmacologic treatment of ADHD is not a simple matter. It is true that a few drugs are highly efficacious against a wide array of symptoms, and have few side effects or adverse long-term consequences. But there are many variables that affect results in individual cases, and the deceptive ease of use of these medicines often leads to casual management without concern for the overall treatment needs of the child and family. Detailed examination of pharmacologic strategies with

ADHD shows that physicians must engage in a complex algorithm for deciding which drug to use, at what dose, and in what form (Greenhill et al., in press). Comprehensive reviews of efficacy and safety of stimulant drugs with ADHD are available (Swanson, McBurnett, Christian, & Wigal, 1995), and the interrelationship among genetics, neuroimaging and the neuropharmacology of ADHD have been well-described (Ernst & Zametkin, 1995). Here, we will focus on the essential clinical phenomena that need to be understood as part of a comprehensive treatment approach to ADHD. Important variables to consider are patient selection, choice of drugs, dose–response, time–action, side effects, and measurement issues.

5.19.8.2.1 Patient selection

Several meta-analyses of drug studies are in agreement in showing that stimulant drugs are most efficacious for symptoms of hyperactivity and other externalizing behavior problems, and somewhat less so for cognitive and achievement problems (Kavale, 1982; Ottenbacher & Cooper, 1983; Thurber & Walker, 1983). The effect sizes in those studies are very large—about one standard deviation for behavioral symptoms, and about half that for cognitive and academic problems. Although it is generally thought that the stimulants are specifically indicated only for hyperactivity, comprehensive reviews show that in addition to the primary symptoms of motor restlessness, concentration, effortful attention to tasks, and self-regulation; stimulants also improve general deportment and compliance, reduce verbal and physical aggression, decrease negative social behaviors, and improve academic productivity and accuracy (Swanson, et al. 1995). Comorbid anxiety may reduce the level of effect from stimulants (Pliszka, 1989) and be better treated with antidepressants such as imipramine (Pliszka, 1987). The fact that normal children also show improved function with stimulants warns against the misconception that the response to stimulants is diagnostic (Rapoport et al., 1980).

5.19.8.2.2 Choice of drugs

There are three stimulants approved by the Food and Drug Administration to treat ADHD patients of all ages: dextroamphetamine (Dexedrine, manufactured by SmithKline-Beachem), methylphenidate (Ritalin, manufactured by CIBA-Geigy), and magnesium pemoline (Cylert, manufactured by Abbot Laboratories). Dextroamphetamine and methylphenidate are structurally related to the catecholamines, dopamine and norepinephrine, whereas pemoline exerts its action primarily through dopaminergic mechanisms (Greenhill, 1992). As noted in recent reviews (March, Erhardt, et al., 1995, March, Wells, & Conners, 1995b), these compounds tend to increase central nervous system (CNS) activity in some but not all brain regions. They increase the activity of striatum and connections between orbitofrontal and limbic regions, while having an inhibitory effect on the neocortex. Stimulants act by enhancing the release of catecholamines into the synaptic cleft, decreasing their reuptake and slowing down their degradation by inhibiting monoamine oxidase. Direct postsynaptic agonist activity is also postulated for dextroamphetamine and methylphenidate. Although catacholaminergic mechanisms are undoubtedly involved in the mechanism of action in ADHD—perhaps by favorably altering the processes by which the CNS allocates attentional resources and/or decides stimulus salience—the exact processes by which these compounds work remain uncertain (Zametkin & Rapoport, 1987).

5.19.8.2.3 Dose–response and time–action

Two essential aspects of stimulant treatment are constructing a dose–response curve and analyzing time–action effects. The former refers to the relationship between dose of drug and the presence of benefits and adverse effects. Establishing these relationships depends on understanding the relationship between the timing of administration, onset of response, point of maximum responsiveness, and loss or offset of drug effect. Empirical evidence suggests that dose and response are independent of body weight and that there is wide intersubject variability in response to the same dose (Rapport & Kelly, 1991). The time–action curve for attentional functioning is somewhat different from that of motor behavior (Solanto & Conners, 1982). Departures from a linear dose–response pattern are common, with some children showing a threshold effect (no response below a threshold level) and others a quadratic response (linear at lower doses and degradation at higher doses). Clinically then, each child requires an individually constructed dose–response curve, taking into consideration the time–action effects of the drug at each dose before making dosage adjustments (Erhardt & Conners, 1995). When titration is accomplished in the traditional open fashion, one starts with the lowest possible dose, working upward every 5–7 days until benefits are maximized or the patient shows prohibitive side effects.

Sprague and Sleator (1977) first suggested that the dose of methylphenidate required to achieve optimal effect for social behavior is higher than the dose required for cognitive function. They found that a high dose in a demanding memory task actually lowered performance below that of the placebo condition. Although there have been several failures to replicate findings of curvilinear dose–response effects with effortful cognitive tasks, some well-designed studies nevertheless continue to find that key attentional and academic measures show a quadratic relationship of dose and effect (Rapport, DuPaul, Stoner, & Jones, 1986). More recent reviews express concern regarding "cognitive toxicity" of high doses, particularly when the laboratory or academic task is difficult or complex (Aman & Rojahn, 1992; Swanson, Cantwell, Lerner, McBurnett, & Hanna, 1991). Interestingly, some studies (Rapport & Kelly, 1991) find that higher doses of stimulants are more effective as the task becomes more complex and demanding. Rapport and Kelly summarized a large number of cognitive tasks used in drug studies and reported that (i) in 100% of the studies the higher dose was better than the lower dose, and (ii) those judged more complex responded, in general, better to the higher dose (though individual group curves sometimes show curvilinear effects). In any case, the findings generally support the proposition that each child requires an individualized assessment in order to define the appropriate dose for the particular target behavior in question, whether behavioral or cognitive.

Dextroamphetamine is usually started at 2.5–5 mg a day and gradually increased in 2.5–5 mg increments; methylphenidate is usually started at 5 mg twice a day and then titrated in 5 mg per dose increments every 4–7 days. Peak behavioral effects are noted 1–3 hours after ingestion and dissipate in 3–6 hours for both methylphenidate and dextroamphetamine. Pemoline is usually started at 37.5 mg each morning and titrated in 18.75 mg increments to maximum clinical effectiveness. Most children respond at a total daily dose of 56.5–75 mg. Although stepped-dose open-label trials of stimulants are the rule in clinical practice, double-blind placebo-controlled titration trials, using parent and teacher ratings and objective measures of performance and academic productivity, can be very helpful in distinguishing a true drug effect from a placebo response and in establishing the best dose for an individual child.

Clinical experience suggests that sustained-release methylphenidate is variably absorbed and shows variable behavioral effects, whereas the dextroamphetamine spansule may show a more even pattern of bioavailability. However, since both tablet and sustained release stimulants are swiftly absorbed after oral administration, it is not surprising that controlled studies show no benefit for sustained release methylphenidate over the standard formulation (Pelham et al., 1987). The sustained release preparations are often helpful when the standard preparation dissipates too rapidly, or when there is stigma associated with going to the teacher or nurse for a midday dose.

5.19.8.2.4 Side effects

Frequent stimulant-related side effects and their management are summarized in Table 3. Since many side effects diminish after several weeks of treatment, patience during stepped-dose titration is important in handling mild to moderate side effects, such as mild anorexia or sleep disturbance. Sometimes gastrointestinal difficulties can be ameliorated by administration of stimulants before, during, or after meals, depending on the child. (In the child on pemoline, prominent gastrointestinal side effects, including but not limited to anorexia, should prompt consideration of pemoline-induced hepatitis since, though rare—less than 1%—it requires drug discontinuation.) Insomnia can sometimes be avoided by reducing the last dose or by earlier administration. Exacerbation of ADHD symptoms and/or irritable dysphoric mood during stimulant offset (stimulant rebound) is uncommon but can be disruptive, especially near the end of the day during homework periods. Rebound requires adjusting the dose or perhaps switching to a sustained release preparation or to a different stimulant.

Stimulants do not appear to adversely influence ultimate growth and height in most children, although a minority of children may be affected either because of calorie restriction or direct effects on growth hormone. Thus, routine drug holidays are not indicated; when necessary, periods off medication may bring about a rebound in growth rate. Caution is in order when using stimulants in children with a personal or family history of hallucinations or thought disorder, especially since attentional abnormalities are common in children at risk for schizophrenia. Close monitoring of blood pressure and pulse is not mandatory with stimulant treatment unless there is a pre-existing cardiovascular disorder, such as tachyarrythmias or hypertension. Because of the potential for transaminitis (an elevation of liver function enzymes), pemoline requires liver function test

Table 3 Side effects of stimulant drugs.

Side effect	*Possible management strategies*
For all side effects	Unless severe, allow 7–10 days for tolerance to develop. Evaluate dose–response relations Evaluate time–action effects and then adjust dosing intervals or switch to a sustained-release preparation Evaluate for concurrent conditions, including comorbidities and environmental stressors Consider switching stimulant drugs
Anorexia/dyspepsia	Administer before, during, or after meals If on pemoline, consider drug-induced hepatitis
Dysphoric mood/emotional construction	Reduce dose or switch to a long-acting preparation Consider comorbidity requiring alternative or adjunctive treatment Switch stimulants
Insomnia/nightmares	Omit or reduce last dose Administer medication earlier in the day If taking sustained-release preparation, switch to tablet drug Consider adding a sedating antihistamine or clonidine at bedtime only
Tics	Firmly establish correlation between tics and pharmacotherapy by examining dose–response relationships, including a no medication condition If tics are very mild or abate after 7–10 days on medication, reconsider risks and benefits of continued stimulant treatment and renew informed consent Switch stimulants Consider nonstimulant treatment, eg clonidine or a tricyclic antidepressant (TCA) If tic disorder and ADHD are severe, consider combining stimulant with a high-potency neuroleptic
Dizziness	Monitor blood pressure and pulse Encourage adequate hydration If associated only with peak effects, change to a sustained-release preparation Switch stimulants, especially to pemoline
Weight loss	Give drug after breakfast and after lunch Implement calorie enhancement strategies, especially a scheduled fourth meal before bedtime Brief drug holidays
Slowed growth	Apply weight loss remedies Weekend and vacation (longer) drug holidays Consider another stimulant or a nonstimulant drug
True rebound	Switch to sustained-release preparation or to pemoline Combine long- and short-acting preparations Consider combined psychostimulant (at a lower dose) with a TCA or switching to bupropion
Psychosis	Discontinue stimulant treatment Assess for comorbid thought disorder Consider alternative treatments

In J. Jefferson & J. Geist (Eds.) *Psychopharmacology clinic of North America*, Pharmacology therapy of attention-deficit hyperactivity disorder by March et al., 1996, New York, Saunders. Copyright 1996 by The Journal of Practical Psychiatry and Behavioral Health. Reprinted with permission.

at baseline and every six months and should be discontinued if double-normal transaminases arise.

One of the major decisions in initiating drug therapy is which stimulant to prescribe in individual cases. In an early study, Millichap and Fowler (1967) randomly assigned 75 6–12-year-old children to dextroamphetamine, methylphenidate, or a placebo and found no differences between the two drugs on 11 behavioral laboratory measures or clinical rating measures. Although for every side effect examined dextroamphetamine had a higher frequency and severity, neither drug had a medically significant frequency of intolerable side effects requiring drug discontinuation. An early review of nine studies by Millichap and Fowler (1967) showed 84% efficacy for methylphenidate compared with 69% for dextroamphetamine, whereas Barkley's (1977) review of 14 methylphenidate and 15 dextroamphetamine studies found approximately equal efficacy with about 70% improvement rate for both stimulants (Arnold, Heustis, & Smeltzer, 1978). Although reasonably consistent, these results are all between-subject effects. Several studies which compare crossovers of the stimulants within subjects, show approximately equal efficacy overall, but not always for the same subjects (Pelham, et al., 1990; Winsberg, Press, Bialer, & Kupietz, 1974). Using a wide variety of social, academic, and behavioral measures, Pelham et al. (1990) also found that dextroamphetamine and methylphenidate were generally equivalent but that some subjects responded better to one drug than to the other. When dosages are raised to their maximum without side effects, the two stimulants are quite comparable in efficacy, but some children who respond to one do not respond to the other (Elia, Borcherding, Rapoport, & Keysor, 1991; Sallee, Stiller, & Perel, 1992). These studies imply that in clinical practice one should try alternative medications within the same class before going on to other classes of drugs. Pemoline shows comparable efficacy to the other stimulants but has had a perhaps unjustified reputation as being prone to produce toxic liver effects (Conners & Taylor, 1980; Conners, Taylor, Meo, Kurtz, & Fournier, 1972; Knights & Viets, 1975; Pelham et al., 1990), and should be considered along with other stimulants as a first-line drug for ADHD (March, Johnston, & Conners, 1996). Antidepressant drugs represent another class of drugs whose efficacy is satisfactory, but which carry higher risks for serious cardiovascular side effects (Gickling & Thompson, 1985; Ross, Poidevant, & Miner, 1995).

5.19.8.2.5 Measurement of drug treatment effects

Parent and teacher ratings have been the most widely used methods for assessing stimulant drug effects, and are considered an essential aspect of treatment monitoring. However, the considerations of dose and time–action effects described above suggests that multilevel assessment of drug effects is likely to be necessary in an optimal treatment regimen. Relying on behavioral ratings alone may jeopardize the cognitive function of the child if dosage is titrated upwards without regard to possible adverse effects on more subtle attentional and learning processes. Curriculum-based measurement uses the rate of learning in standard reading, spelling, or arithmetic tasks, using brief timed samples of performance. These methods have been successfully used in assessing and measuring treatment progress of ADHD children (Atkins & Pelham, 1991; Atkins, Pelham, & Licht, 1985). A number of classroom-based measures of academic productivity have been described which show good discriminative ability in differentiating ADHD children from controls. These methods appear to be sensitive to drug treatment and are certainly more ecologically valid than conventional standardized tests of academic function which tend to be insensitive to treatment (Balthazor, Wagner, & Pelham, 1991; Pelham, 1986; Pelham, Bender, Caddell, Booth, & Moorer, 1985; Pelham et al., 1987).

5.19.9 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

5.19.9.1 Future of Assessments of ADHD

Considerable advances have been made in our understanding of both the processes of attention and inhibition (Conners, March, Erhardt, Butcher, & Epstein, 1995). As yet, however, these newer approaches have not been accompanied by methods that are easily transported from the laboratory to the clinic. Nevertheless, computerized tests based upon neurocognitive models are likely to be available for use once adequate normative studies and validity studies are undertaken. These new approaches will be necessary to meet the justified criticisms of those who argue that the core features of ADHD have never been adequately operationalized. Equally important as assessments at the level of microprocesses of attention and inhibition, are larger family, social, and developmental issues, which at present are poorly assessed. Until assessments have an adequate treatment target, however, they had, perhaps, better be left alone.

5.19.9.2 System Issues in Treatment

One of the most common failings of physicians in medication management is the failure to insist upon the other treatments which are known to be important for the long-term success of ADHD. The overprescription of stimulant drugs, though undoubtedly magnified by public hysteria and the media, represents a real problem. At the same time, practitioners in schools and clinics often fail to provide physicians with the support they need in terms of feedback regarding the effect of dosage and drug changes both in the home and school setting. This problem represents failure of communication across professional boundaries, failure to develop a *modus vivendi* between professional disciplines, and a lack of rational distribution of assessment and treatment resources. We have the sad fact of teachers referring children for medication because they cannot be managed in the classroom, partly because the school system itself is bogged down in meetings and assessments required by law, with few resources left to address immediate behavioral concerns in the classroom, and partly because school psychologists and other professionals have little training in actual intervention strategies, but have plenty of training in modes of assessment which are both outmoded, time consuming, and inappropriate for the school setting.

Underprescription of stimulants is also a problem. There are many children who go unrecognized as having a legitimate disorder with profound consequences for their own and their family's development. This represents a failure of programs for screening with tools that can identify the threshold for administration of drugs, as well as an inadequate understanding by the public of the importance of early recognition and intervention.

5.19.9.3 Research Issues in ADHD

The current diagnostic criteria are based entirely upon observable symptoms or behaviors directed at the two concepts of overactivity–impulsivity and inattention. Yet there is clear evidence that activity and attention are both multidimensional concepts. Activity that involves motor overflow and fidgeting is quite different from activity that is disrupted because action schemas, both short- and long-term, flounder due to lapses of working memory or lack of meaningful goals. The loss of attention due to lack of arousal is clearly different than that involving a highly aroused, overly vigilant state in which every stimulus has excitatory value and salience. The brain is too complex and subtly geared to the richness of the environmental transactions it must make to have one activity center or attention center, yet our assessments and treatments appear to be predicated on nineteenth-century concepts of the way the brain works. If there are several different circuits involved in alerting to, switching to, and sustaining focus on things, biology dictates that if it can go wrong, at some point it will, for some of the people some of the time. Therefore, we can expect there to be several different types of attention disorder.

As James (1890) so presciently observed, the key to behavioral inhibition is closely related to attention, yet our criteria make no distinction between disinhibition in the classroom due to the inattention of the poor reader faced with an impossible task, and disinhibition which comes from an extremely low threshold for detecting movement in peripheral vision. The major research task, then, in our opinion is the resolution of the heterogeneous syndrome we now call ADHD into subtypes based on a clear understanding of the underlying brain mechanisms which regulate attentional and inhibitory states. Subtypes based solely on overt symptom criteria leave us at a stage of nosology comparable to that when medicine included fever as a major class of disorder. (A fuller discussion of the relationship between neurocognitive models and ADHD is given by Conners et al., 1995.)

5.19.10 SUMMARY

ADHD has evolved from nineteenth-century concepts of impairment of will, to an organically based neurological condition, to a behavioral syndrome of uncertain etiology. Despite limitations in knowledge about the etiology of the condition, considerable agreement regarding the clinical features, developmental course, response to treatment, and adult outcome have accumulated, allowing reasonable preventive, ameliorative, and therapeutic interventions. The core features of inattention and hyperactivity–impulsivity are readily identified using clinical interview, rating scales, and developmental history. The presence of several comorbid symptoms, including conduct disturbance, oppositional behavior, learning disorders, anxiety, and depression, complicates assessment and enlarges the boundaries of treatment needs for most ADHD children beyond the core symptoms themselves. ADHD appears to encompass a spectrum that runs from near-normal to very severe, posing difficult problems of determining a level of functional impairment that warrants diagnosis and intervention. This leads to broad variations in estimates of prevalence. There is still uncertainty as to whether inattention and hyperactivity are part of the same disorder, or two separate disorders. The fact that many different risk factors may lead to ADHD, including genetic, medical, environmental, temperamental, metabolic, and nutritional variables, suggests that ADHD may have several subtypes depending upon the pattern of insults to the developing brain. Protective factors enter in as variables that can reduce impairment below the threshold level of expression required for diagnosis. It is relatively clear that in the absence of intervention and protective factors such as high IQ or good family circumstances, untreated ADHD results in an adult condition of considerable consequence in terms of overall adjustment, vocational and academic accomplishment, and productive lifestyle.

Except for relatively mild versions of ADHD, multimodal treatment is almost always indicated. The sequence, degree, and length of these treatments may vary as a function of developmental level, emergent secondary symptoms such as problems in peer relationships, and parental and community resources. Most treatment packages will include a consideration of medication, school interventions for both behavior and academic competence, and interventions with the parents as the main therapeutic vehicles. Child-based interventions, such as cognitive therapy and psychotherapy, will be less useful, although these can sometimes be quite helpful in individual cases. The use of behaviorally based strategies such as token systems, contracting, and self-monitoring will prove to be more helpful. Pharmacologic management is also a multimodal therapy in the sense that all elements of the social system—school, child, parent, and physician—need to be involved for successful implementation of this adjunct to the total treatment package.

Future research with ADHD will probably focus upon continued efforts to identify the particular nature of the anatomic, neurochemical, and molecular bases of the disorder, with more precise assessment tools being linked to better understanding of the brain-related processes involved. The resolution of ADHD into more specific subtypes, identifiable by behavioral methodologies, would seem to be a particularly important clinical goal. On the treatment side, ongoing studies of the relative merits of single and combined therapies should elucidate which treatment strategies are effective, although such findings will ultimately require developmental level-specific versions to be useful.

5.19.11 REFERENCES

Abikoff, H. (1991). Cognitive training in ADHD children: Less to it than meets the eye. *Journal of Learning Disabilities, 24,* 205–209.

Abikoff, H., & Gittelman, R. (1984). Does behavior therapy normalize the classroom behavior of hyperactive children? *Archives of General Psychiatry, 41,* 449–454.

Abikoff, H., Gittelman-Klein, R., & Klein, D. (1977). Validation of a classroom observation code for hyperactive children. *Journal of Consulting and Clinical Psychology, 45,* 772–783.

Abikoff, H., & Hechtman, L. (1996). Multimodal therapy and stimulants in the treatment of children with attention deficit hyperactivity disorder. In E. D. Hibbs & P. S. Jensen (Eds.), *Psychosocial treatments for child and adolescent disorders: Empirically-based strategies for clinical practice* (pp. 341–369). Washington, DC: American Psychological Association.

Abikoff, H., & Klein, R. G. (1992). Attention-deficit hyperactivity and conduct disorder: Comorbidity and implications for treatment. *Journal of Consulting and Clinical Psychology, 60,* 881–892.

Abramowitz, A. J. (1994). Classroom interventions for disruptive behavior disorders. *Child and Adolescent Clinics of North America, 3,* 343–360.

Achenbach, T. (1992). New developments in multiaxial empirically based assessment of child and adolescent psychopathology. In J. Rosen & P. McReynolds (Eds.), *Advances in psychological assessment* (Vol. 8, pp. 75–102). New York: Plenum.

Achenbach, T. M., Conners, C. K., Quay, H. C., Verhulst, F. C., & Howell, C. T. (1989). Replication of empirically derived syndromes as a basis for taxonomy of child/adolescent psychopathology. *Journal of Abnormal Child Psychology, 17*(3), 299–323.

Achenbach, T. M., & Edelbrock, C. S. (1983). *Manual for the Child Behavior Profile and Child Behavior Checklist.* Burlington, VT: Author.

Achenbach, T. M., Howell, C. T., Quay, H. C., & Conners, C. K. (1991). National survey of problems and competencies among four- to sixteen-year-olds: Parents' reports for normative and clinical samples. *Monographs of the Society For Research In Child Development, 56*(3), 1–131.

Alberts, C. J., Firestone, P., & Goodman, J. T. (1986). Attention and impulsivity characteristics of the biological and adoptive parents of hyperactive and normal control children. *American Journal of Orthopsychiatry, 56*(3), 413–423.

Aman, M. G., & Rojahn, J. (1992). Pharmacological intervention. In N. N. Singh & I. L. Beale (Eds.), *Learning disabilities: Nature, theory, and treatment* (pp. 478–525). New York: Springer-Verlag.

American Psychiatric Association (1968). *Diagnostic and statistical manual of mental disorders* (2nd ed.). Washington, DC: Author.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anastopoulos, A. D., & Barkley, R. A. (1990). Counseling and training parents. In R. A. Barkley, *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment* (pp. 397–431). New York: Guilford Press.

Anastopoulos, A. D., Barkley, R. A., & Shelton, T. L. (1996). Family-based treatment: Psychosocial intervention for children and adolescents with attention deficit hyperactivity disorder. In E. D. Hibbs & P. S. Jensen (Eds.), *Psychosocial treatments for child and adolescent* *disorders: Empirically-based strategies for clinical practice* (pp. 267–284). Washington, DC: American Psychological Association.

Anastopoulos, A. D., Guevremont, D. C., Shelton, T. L., & DuPaul, G. J. (1992). Parenting stress among families of children with attention deficit hyperactivity disorder. *Journal of Abnormal Child Psychology, 20,* 503–520.

Anastopoulos, A. D., Shelton, T. L., DuPaul, G. J., & Guevremont, D. C. (1993). Parent training for attention deficit hyperactivity disorder: Its impact on parent functioning. *Journal of Abnormal Child Psychology, 21,* 581–596.

Anderson, C. A., Hinshaw, S. P., & Simmel, C. (1994). Mother–child interactions in ADHD and comparison boys: Relationships with overt and covert externalizing behavior. *Journal of Abnormal Child Psychology, 22,* 247–265.

Anderson, J. C., Williams, S., McGee, R., & Silva, P. A. (1987). DSM-III disorders in preadolescent children: prevalence in a large sample from the general population. *Archives of General Psychiatry, 44,* 69–76.

Angold, A. (1989). Structured assessments of psychopathology in children and adolescents. In C. Thompson (Ed.), *The instruments of psychiatric research.* New York: Wiley.

Arnold, L. E. (1996). Sex differences in ADHD: Conference summary. *Journal of Abnormal Child Psychology, 24,* 555–569.

Arnold, L. E. Heustis, R., & Smeltzer, D. J. (1978). Methylphenidate versus dextroamphetamine versus caffeine in minimal brain dysfunction. *Archives of General Psychiatry, 35,* 463–473.

Asarnow, J. R. (1988). Peer status and social competence in child psychiatric inpatients: A comparison of children with depressive, externalizing, and concurrent depressive and externalizing disorders. *Journal of Abnormal Child Psychology, 16,* 151–162.

Atkins, M. S., & Pelham, W. E. (1991). School-based assessment of attention deficit hyperactivity disorder. *Journal of Learning Disabilities, 24,* 197–204.

Atkins, M. S., Pelham, W. E., & Licht, M. H. (1985). A comparison of objective classroom measures and teacher ratings of attention deficit disorder. *Journal of Abnormal Child Psychology, 13,* 155–167.

Ayllon, T., & Rosenbaum, M. (1977). The behavioral treatment of disruption and hyperactivity in school settings. In B. Lahey & A. Kazdin (Eds.), *Advances in clinical child psychology* (Vol. 1, pp. 83–118). New York: Plenum.

Balthazor, M. J., Wagner, R. K., & Pelham, W. E. (1991). The specificity of the effects of stimulant medication on classroom learning-related measures of cognitive processing for attention deficit disorder children. *Journal of Abnormal Child Psychology, 19,* 35–52.

Barkley, R. (1977). A review of stimulant drug research with hyperactive children. *Journal of Child Psychology and Psychiatry, 18,* 137–165.

Barkley, R. A. (1987). *Defiant children: A clinician's manual for parent training.* New York: Guilford Press.

Barkley, R. A. (1990). *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment.* New York: Guilford Press.

Barkley, R. A. (1994). Impaired delayed responding: A unified theory of attention deficit hyperactivity disorder. In D. K. Routh (Ed.), *Disruptive behavior disorders in childhood* (pp. 11–58). New York: Plenum.

Barkley, R. A. (1996). Attention-deficit hyperactivity disorder. In E. J. Mash & R. A. Barkley (Eds.), *Child psychopathology* (pp. 63–112). New York: Guilford Press.

Barkley, R. A., DuPaul, G. J., & McMurray, M. B. (1990). A comprehensive evaluation of attention deficit disorder with and without hyperactivity as defined by research criteria. *Journal of Consulting and Clinical Psychology, 58,* 775–789.

Barkley, R. A., Fischer, M., Edelbrock, C. S., Smallish, L. (1990). The adolescent outcome of hyperactive children diagnosed by research criteria: I. An 8-year prospective follow-up study. *Journal of the American Academy of Child & Adolescent Psychiatry, 29,* 546–557.

Barkley, R. A., Fischer, M., Newby, R. F., & Breen, M. J. (1988). Development of a multimethod clinical protocol for assessing stimulant drug response in children with attention deficit disorder. *Journal of Clinical Child Psychology, 17*(1), 14–24.

Barkley, R. A., Karlsson, J., & Pollard, S. (1985). Effects of age on the mother–child interactions of hyperactive children. *Journal of Abnormal Child Psychology, 13,* 631–638.

Barkley, R. A., Murphy, K., & Kwasnik, D. (1996). Psychological adjustment and adaptive impairments in young adults with ADHD. *Journal of Attention Disorders, 1,* 41–54.

Bauermeister, J. J., Alegeria, M., Bird, H., Rubio-Stipec, M., & Canino, G. (1992). Are attentional-hyperactivity deficits unidimensional or multidimensional syndromes? Empirical findings from a community survey. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 423–431.

Benton, D., & Buts, J. (1990). Vitamin/mineral supplementation and intelligence. *The Lancet,* 1158–1160.

Benton, D., & Cook, R. (1991). Vitamin and mineral supplements improve the intelligence scores and concentration of six-year-old children. *Personality and Individual Differences, 12,* 1151–1158.

Benton, D., & Roberts, G. (1988). Effect of vitamin and mineral supplementation on intelligence of a sample of schoolchildren. *The Lancet,* 140–143.

Bhatia, M. S., Nigam, V. R., Bohra, N., & Malik, S. C. (1991). Attention deficit disorder with hyperactivity among pediatric outpatients. *Journal of Child Psychology and Psychiatry, 32,* 297–306.

Biederman, J., Faraone, S. V., Keenan, K., Knee, D., & Tsuang, M. T. (1990). Family-genetic and psychosocial risk factors in DSM-III: Attention deficit disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 526–533.

Biederman, J., Faraone, S. V., Keenan, K., Steingard, R., & Tsuang, M. T. (1991). Familial association between attention deficit disorder and anxiety disorders. *American Journal of Psychiatry, 148,* 251–256.

Biederman, J., Newcorn, J., & Sprich, S. (1991). Comorbidity of attention deficit hyperactivity disorder with conduct, depressive, anxiety, and other disorders. *American Journal of Psychiatry, 148,* 564–577.

Biederman, J., Wilens, T., Mick, E., Milberger, S., Spencer, T. J., & Faraone, S. V. (1995). Psychoactive substance use disorders in adults with attention deficit hyperactivity disorder (ADHD): Effects of ADHD and psychiatric comorbidity. *American Journal of Psychiatry, 152,* 1652–1658.

Boudreault, M., & Thivierge, J. (1986). The impact of temperament in a school setting: An epidemiological study. Special Issue: Canadian Academy of Child Psychiatry: A Canadian perspective. *Canadian Journal of Psychiatry, 31*(6), 499–504.

Bradley, C. (1937). The behavior of children receiving benzedrine. *American Journal of Psychiatry, 94,* 577–585.

Braswell, L., & Bloomquist, M. (1991). *Cognitive-behavioral therapy with ADHD children: Child, family, and school intervention.* New York: Guilford Press.

Brody, G. H., Stoneman, Z., & Burke, M. (1988). Child temperament and parental perceptions of individual child adjustment: an intrafamilial analysis. *American Journal of Orthopsychiatry, 58,* 532–542.

Brown, R. T., & Quay, L. C. (1977). Reflection-impulsivity of normal and behavior disordered children. *Journal of Abnormal Child Psychology, 5,* 457–462.

Burcham, B., Carlson, L., & Milich, R. (1993). Promising school-based practices for students with attention deficit disorder. *Exceptional Children, 60,* 174–180.

Burrows, G. M. (1828). *Commentaries on the causes, forms, symptoms and treatment, moral and medical, of insanity*. London: Thomas and George Underwood.

Campbell, S. B. (1990). *Behavior problems in preschool children: Clinical and developmental issues*. New York: Guilford Press.

Campbell, S. B. (1995). Behavior problems in preschool children: A review of recent research. *Journal of Child Psychology and Psychiatry, 36,* 113–149.

Cantwell, D. P. (1988). Families with attention deficit disordered children and others at risk. *Journal of Chemical Dependency Treatment, 1*(2), 163–186.

Cantwell, D. P., & Baker, L. (1991). Association between attention-deficit hyperactivity disorder and learning disorders. *Journal of Learning Disabilities, 24,* 88–95.

Cantwell, D. P., & Baker, L. (1992). Attention deficit disorder with and without hyperactivity: A review and comparison of matched groups. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 432–438.

Centers for Disease Control. (1991). *Preventing lead poisoning in children*. Atlanta, GA: Author.

Chandra, P. S. (1993). Cross-cultural psychiatry and children with defiant behaviors. *American Journal of Psychiatry, 150,* 1279–1280.

Clements, S. (1966). Minimal brain dysfunction in children—terminology and identification. *National Institute of Neurological Diseases and Blindness Monograph No. 3*. Bethesda, MD: US Public Health Service.

Clements, S., & Peters, J. (1962). Minimal brain dysfunctions in the school-age child. *Archives of General Psychiatry, 6,* 185–197.

Cohen, D. J., & Leckman, J. F. (1994). Developmental psychopathology and neurobiology of Tourette's syndrome. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 2–15.

Conners, C. K. (1970). Symptom patterns in hyperkinetic, neurotic, and normal children. *Child Development, 41,* 667–682.

Conners, C. K. (1972). Symposium: behavior modification by drugs. II. Psychological effects of stimulant drugs in children with minimal brain dysfunction. *Pediatrics, 49,* 702–708.

Conners, C. K. (1989). *Manual for Conners' Rating Scales*. North Tonawanda, NY: Multi-Health Systems.

Conners, C. K. (1994a). *The Conners Continuous Performance Test*. Toronto, Canada: Multi-Health Systems.

Conners, C. K. (1994b). *The continuous performance test (CPT): Use as a diagnostic tool and measure of treatment outcome*. Paper presented at the Annual Meeting of the American Psychological Association, Los Angeles, CA.

Conners, C. K. (in press). *Risk-factors checklist for ADHD*. Toronto, Canada: Multi-Health Systems.

Conners, C., Erhardt, D., & Sparrow, E. (in press). *CADDI: A structured interview for ADHD and other externalizing disorders*.

Conners, C. K., & Kronsberg, S. (1985). Measuring activity level in children. *Psychopharmacology Bulletin, 21,* 893–897.

Conners, C., & March, J. (1993). *Developmental history form for ADHD and related disorders*. Toronto, Canada: Multi-Health Systems.

Conners, C. K., March, J. S., Erhardt, D., Butcher, T., & Epstein, J. (1995). Assessment of attention deficit disorders (ADHD): Conceptual issues and future trends. *Journal of Psychoeducational Assessment*. Monograph series, Advances in psychoeducational assessment, 185–204.

Conners, C. K., March, J. S., Fiore, C., & Butcher, T. (1994). *Information processing deficits in ADHD: Effect of stimulus rate and methylphenidate*. Paper presented at the Annual Meeting of the American College of Neuropsychopharmacology, Honolulu, HI.

Conners, C. K., Parker, J., & Sitarenios, G. (1996). Parent and teacher ratings of DSM-IV ADHD symptoms: A test of the subtype model. Submitted.

Conners, C. K., & Taylor, E. (1980). Pemoline, methylphenidate, and placebo in children with minimal brain dysfunction. *Archives of General Psychiatry, 37,* 922–930.

Conners, C. K., Taylor, E., Meo, G., Kurtz, M. A., & Fournier, M. (1972). Magnesium pemoline and dextroamphetamine: a controlled study in children with minimal brain dysfunction. *Psychopharmacologia, 26*(4), 321–336.

Conners, C. K., & Wells, K. C. (1985). ADD-H Adolescent Self-Report Scale. *Psychopharmacology Bulletin, 21*(4), 921–922.

Conners, C. K., Wells, K. C., Erhardt, D., March, J. S., Schulte, A., Osborne, S., Fiore, C., & Butcher, A. T. (1994). Multimodality therapies: Methodologic issues in research and practice. *Child and Adolescent Psychiatric Clinics of North America: Disruptive Disorders, 3,* 361–377.

David, O., Hoffman, S., Sverd, J., & Clark, J. (1977). Lead and hyperactivity: Lead levels among hyperactive children. *Journal of Abnormal Child Psychology, 5,* 405–416.

Davila, R. R., Williams, M. L., & MacDonald, J. T. (1991). *Clarification of policy to address the needs of children with attention deficit hyperactivity disorders within general and/or special education* (memo). Washington, DC: US Department of Education.

Deutsch, C. K., Matthysse, S., Swanson, J. M., & Farkas, L. G. (1990). Genetic latent structure analysis of dysmorphology in attention deficit disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 189–194.

Douglas, V. I. (1983). Attention and cognitive problems. In M. Rutter (Ed.), *Developmental neuropsychiatry* (pp. 280–329). New York: Guilford Press.

Douglas, V., & Peters, K. (1978). Toward a clearer definition of the attentional deficit of hyperactive children. In G. Hale & M. Lewis (Eds.), *Attention and the development of cognitive skills* (pp. 173–248). New York: Plenum.

DuPaul, G. J., & Stoner, G. (1994). *ADHD in the schools: Assessment and intervention strategies*. New York: Guilford Press.

Dykman, R., Ackerman, P., Clements, S., & Peters, J. (1971). Specific learning disabilities: An attentional deficit syndrome. In H. Myklebust (Ed.), *Progress in learning disabilities* (Vol. 2, pp. 56–94). New York: Grune and Stratton.

Elia, J., Borcherding, B. G., Rapoport, J. L., & Keysor, C. S. (1991). Methylphenidate and dextroamphetamine treatments of hyperactivity: Are there true nonresponders? *Psychiatry Research, 36,* 141–155.

Erhardt, D., & Conners, C. K. (1995). Methodological and assessment issues in pediatric psychopharmacology. In J. M. Wiener (Ed.), *Diagnosis and psychopharmacology of childhood and adolescent disorders* (pp. 97–147). New York: Wiley.

Erhardt, D., & Hinshaw, S. P. (1994). Initial sociometric impressions of attention-deficit hyperactivity disorder and comparison boys: Predictions from social behaviors and from nonbehavioral variables. *Journal of Consulting and Clinical Psychology, 62,* 833–842.

Ernst, M., & Zametkin, A. (1995). The interface of genetics, neuroimaging and neurochemistry in attention-deficit hyperactivity disorder. In F. E. Bloom & D. J. Kupfer (Eds.), *Psychopharmacology: The fourth generation of progress* (pp. 1643–1652). New York: Raven.

Faigel, H. C., Sznajderman, S., Tishby, O., Turel, M., & Pinus, U. (1995). Attention deficit disorder during adolescence: A review. *Journal of Adolescent Health, 16,* 174–184.

Faraone, S. V., Biederman, J., Keenan, K., & Tsuang, M. T. (1991a). A family-genetic study of girls with *DSM-III* attention deficit disorder. *American Journal of Psychiatry, 148,* 112–117.

Faraone, S. V., Biederman, J., Keenan, K., & Tsuang, M. T. (1991b). Separation of DSM-III attention deficit disorder and conduct disorder: Evidence from a family-genetic study of American child psychiatric patients. *Psychological Medicine, 21,* 109–121.

Fischer, M., Barkley, R. A., Edelbrock, C. S., Smallish, L. (1990) The adolescent outcome of hyperactive children diagnosed by research criteria: II. Academic, attentional, and neuropsychological status. *Journal of Consulting and Clinical Psychology, 58,* 580–588.

Gickling, E. E., & Thompson, V. P. (1985). A personal view of curriculum-based assessment. Special Issue: Curriculum-based assessment. *Exceptional Children, 52,* 205–218.

Gilger, J. W., Pennington, B. F., & DeFries, J. C. (1992). A twin study of the etiology of comorbidity: attention-deficit hyperactivity disorder and dyslexia. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 343–348.

Gillis, J. J., Gilger, J. W., Pennington, B. F., & DeFries, J. C. (1992). Attention deficit disorder in reading disabled twins: Evidence for a genetic etiology. *Journal of Abnormal Child Psychology, 20,* 303–315.

Gittelman, R., Mannuzza, S., Shenker, R., & Bonagura, N. (1985). Hyperactive boys almost grown up. *Archives of General Psychiatry, 42,* 937–947.

Goldstein, K. (1942). *After-effects of brain injuries in war.* New York: Grune and Stratton.

Goodman, R., & Stevenson, J. (1989a). A twin study of hyperactivity: I. An examination of hyperactivity scores and categories derived from Rutter teacher and parent questionnaires. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 30,* 671–689.

Goodman, R., & Stevenson, J. (1989b). A twin study of hyperactivity: II. The aetiological role of genes, family relationships and perinatal adversity. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 30,* 691–709.

Goodyear, P., & Hynd, G. W. (1992). Attention-deficit disorder with (ADD/H) and without (ADD/WO) hyperactivity: Behavioral and neuropsychological differentiation. *Journal of Clinical Child Psychology, 21,* 273–305.

Goyette, C. H., Conners, C. K., & Ulrich, R. F. (1978). Normative data on Revised Conners Parent and Teacher Rating Scales. *Journal of Abnormal Child Psychology, 6,* 221–236.

Greenfield, B., Gottlieb, S., & Weiss, G. (1992). Psychosocial interventions: Individual psychotherapy with child, and family and parent counseling. *Child and Adolescent Psychiatric Clinics of North America, 1,* 481–503.

Greenhill, L. (1992). Pharmacologic treatment of attention deficit hyperactivity disorder. In D. Shaffer (Ed.), *The psychiatric clinics of North America: Pediatric psychopharmacology* (pp. 1–27). Philadelphia, PA: Saunders.

Greenhill, L., Abikoff, H., Arnold, E., Conners, C., Wells, K., Elliott, G., Hechtman, L., Hinshaw, S., Hoza, B., Jensen, P., March, J., Newcorn, J., Pelham, W., Severe, J., Swanson, J., & Vitiello, B. (1996). Medication treatment strategies in the MTA study: Relevance to clinicians and researchers. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 1304–1313.

Guevremont, D. C. (1990). Social skills and peer relationship training. In R. A. Barkley, *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (pp. 540–572). New York: Guilford Press.

Halperin, J. M., Matier, K., Bedi, G., Sharma, V., & Newcorn, J. H. (1992). Specificity of inattention, impulsivity, and hyperactivity to the diagnosis of attention-deficit hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 190–196.

Hamlett, K. W., Pellegrini, D. S., & Conners, C. K. (1987). An investigation of executive processes in the problem-solving of attention deficit disorder-hyperactive children. *Journal of Pediatric Psychology, 12,* 227–240.

Hart, E. L., Lahey, B. B., Loeber, R., Applegate, B., & Frick, P. J. (1995). Developmental change in attention-deficit hyperactivity disorder in boys: a four-year longitudinal study. *Journal of Abnormal Child Psychology. 23*(6), 729–749.

Hartsough, C. S., & Lambert, N. M. (1985). Medical factors in hyperactive and normal children: Prenatal, developmental, and health history findings. *American Journal of Orthopsychiatry, 55,* 190–210.

Hechtman, L. (1989). Attention-deficit hyperactivity disorder in adolescence and adulthood: An updated follow-up. *Psychiatric Annals, 19,* 597–603.

Hechtman, L. (1991). Resilience and vulnerability in long term outcome of attention deficit hyperactive disorder. Special issue: Child and adolescent psychiatry. *Canadian Journal of Psychiatry, 36,* 415–421.

Heffron, W. A., Martin, C. A., & Welsh, R. J. (1984). Attention deficit disorder in three pairs of monozygotic twins: A case report. *Journal of the American Academy of Child Psychiatry, 23,* 299–301.

Hembree-Kigin, T. L., & McNeil, C. B. (1995). *Parent–child interaction therapy.* New York: Plenum.

Hinshaw, S. P. (1987). On the distinction between attentional deficits/hyperactivity and conduct problems/aggression in child psychopathology. *Psychological Bulletin, 101,* 443–463.

Hinshaw, S. P. (1992a). Academic underachievement, attentional deficits, and aggression: Comorbidity and implications for intervention. *Journal of Consulting and Clinical Psychology, 60,* 893–903.

Hinshaw, S. P. (1992b). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin, 111,* 127–155.

Hinshaw, S. P. (1992c). Intervention for social competence and social skill. *Child and Adolescent Clinics of North America, 2,* 539–552.

Hinshaw, S. P. (1994). *Attention deficits and hyperactivity in children.* Thousand Oaks, CA: Sage.

Hinshaw, S. P., & Erhardt, D. (1991). Attention-deficit hyperactivity disorder. In P. C. Kendall (Ed.), *Child and adolescent therapy: Cognitive-behavioral procedures* (pp. 98–128). New York: Guilford Press.

Hinshaw, S. P., & Erhardt, D. (1993). Behavioral treatment. In V. B. Van Hasselt & M. Hersen (Eds.), *Handbook of behavior therapy and pharmacotherapy for children: A comparative analysis* (pp. 233–250). Needham Heights, MA: Allyn & Bacon.

Holborow, P. L., Berry, P., & Elkins, J. (1984). Prevalence of hyperkinesis: A comparison of three rating scales. *Journal of Learning Disabilities, 17,* 411–417.

Homatidis, S., & Konstantareas, M. M. (1981). Assessment of hyperactivity: Isolating measures of high discriminant ability. *Journal of Consulting and Clinical Psychology, 49,* 533–541.

Hops, H., & Lewin, L. (1984). Peer sociometric forms. In T.

Ollendick & M. Hersen (Eds.), *Child behavioral assessment: Principles and procedures* (pp. 124–147). New York: Pergamon.

Hoza, B., Vallano, G., & Pelham, W. E. (1995). Attention-deficit hyperactivity disorder. In R. T. Ammerman & M. Hersen (Eds.), *Handbook of child therapy in the psychiatric setting* (pp. 181–198). New York: Wiley.

Hynd, G. W., Hern, K. L., Voeller, K. K., & Marshall, R. M. (1991). Neurobiological basis of attention-deficit hyperactivity disorder (ADHD). *School Psychology Review, 20,* 174–186.

James, W. (1890). *The principles of psychology*. New York: Gryphon Editions.

Jensen, P. S., Shervette, R. E., Xenakis, S. N., & Bain, M. W. (1988). Psychosocial and medical histories of stimulant-treated children. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 798–801.

Johnston, C. (1996). Parent characteristics and parent–child interactions in families of nonproblem children and ADHD children with higher and lower levels of oppositional-defiant behavior. *Journal of Abnormal Child Psychology, 24,* 85–104.

Juarez, L. J., Madrigal, J., & Anderson, R. P. (1981). Social and cultural aspects of behavior disorders among minority children: Reflections upon contemporary literature. *Spanish Speaking Mental Health Research Center Monographs 8*.

Kahn, E., & Cohen, L. (1934). Organic drivenness: A brain stem syndrome and an experience. *New England Journal of Medicine, 210,* 748–756.

Kavale, K. (1982). The efficacy of stimulant drug treatment for hyperactivity: A meta-analysis. *Journal of Learning Disabilities, 15,* 280–289.

Kessler, J. W. (1980). History of minimal brain dysfunction. In H. E. Rie & E. D. Rie (Eds.), *Handbook of minimal brain dysfunctions: A critical view* (pp. 18–52). New York: Wiley.

Klein, R. G., & Mannuzza, S. (1991). Long-term outcome of hyperactive children: A review. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 383–387.

Knights, R. M., & Viets, C. A. (1975). Effects of pemoline on hyperactive boys. *Pharmacology, Biochemistry and Behavior, 3,* 1107–1114.

Kovacs, M. (1993). *The Children's Depression Inventory* (*CDI*). Toronto, Canada: Multi-Health Systems.

Lahey, B. B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L., Hynd, G. W., Barkley, R. A., Newcorn, J., Jensen, P., Richters, J., Garfinkel, B., Kerdyk, L., Frick, P. J., Ollendick, T., Perez, D., Hart, E. L., Waldman, I., & Shaffer, D. (1994). *DSM-IV* field trials for attention deficit hyperactivity disorder in children and adolescents. *American Journal of Psychiatry, 151,* 1673–1685.

Lahey, B. B., Loeber, R., Stouthamer-Loeber, M., & Christ, M. A. (1990). Comparison of *DSM-III* and *DSM-III-R* diagnoses for prepubertal children: Changes in prevalence and validity. *Journal of the American Academy of Child and Adolescent Psychiatry, 29*(4), 620–626.

Lahey, B. B., Pelham, W. E., Schaughency, E. A., Atkins, M. S., Murphy, H. A., Hynd, G. W., Russo, M., Hartdagen, S., & Lorys-Vernon, A. (1988). Dimensions and types of attention deficit disorder with hyperactivity in children: A factor analytic and cluster analytic approach. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 330–335.

Lambert, N. M. (1988). Adolescent outcomes for hyperactive children: Perspectives on general and specific patterns of childhood risk for adolescent educational, social, and mental health problems. *American Psychologist, 43,* 786–799.

Lambert, N. M., Sandoval, J. H., & Sassone, D. M. (1978). Prevalence estimates of hyperactivity in school children. *Pediatric Annals, 7,* 68–86.

Landry, S. H., Chapieski, M. L., Richardson, M. A., Palmer, J., & Hall, S. (1990). The social competence of children born prematurely: effects of medical complications and parent behaviors. *Child Development, 61,* 1605–1616.

Lansdorf, R. (1979). Ethnicity, social class, and perception of hyperactivity. *Psychology in the Schools, 16,* 293–298.

Laufer, M., Denhoff, E., & Solomons, G. (1957). Hyperkinetic impulse disorder in children's behavior problems. *Psychosomatic Medicine, 19,* 39–49.

Letters to the Editor. (1992). *Science, 256,* 294–295.

Lewis, M., & Bendersky, M. (1989). Cognitive and motor differences among low birth weight infants: impact of intraventricular hemorrhage, medical risk and social class. *Pediatrics, 83,* 187–192.

Loney, J., & Milich, R. (1982). Hyperactivity, inattention, and aggression in clinical practice. In M. Wolraich & D. K. Routh (Eds.), *Advances in developmental and behavioral pediatrics* (Vol. 2, pp. 113–147). Greenwich, CT: JAI Press.

Luk, S. (1985). Direct observation studies of hyperactive behaviors. *Journal of the American Academy of Child Psychiatry, 24,* 338–344.

Lynskey, M. T., & Fergusson, D. M. (1995). Childhood conduct problems, attention deficit behaviors, and adolescent alcohol, tobacco, and illicit drug use. *Journal of Abnormal Child Psychology, 23,* 281–302.

Macmann, G., Barnett, D., & Lopez, E. (1993). The child behavior checklist/4-18 and related materials: Reliability and validity of syndromal assessment. *School Psychology Review, 22,* 308–312.

Mannuzza, S., Gittelman-Klein, R., Bessler, A., Malloy, P., & LaPadula, M. (1993). Adult outcome of hyperactive boys: Educational achievement occupational rank, and psychiatric status. *Archives of General Psychiatry, 50,* 565–576.

March, J. S., Erhardt, D., Johnston, H., & Conners, C. K. (1995). Pharmacotherapy of attention-deficit hyperactivity disorder. In J. W. Jefferson & J. H. Greist (Eds.), *The psychiatric clinics of North America: Annual of drug therapy*. Philadelphia: Saunders.

March, J. S., Johnston, H., & Conners, C. K., (1996). Pharmacotherapy of attention-deficit hyperactivity disorder. In J. Jefferson & J. Greist (Eds.), *Psychopharmacology Clinic of North America*. New York: Saunders.

March, J. S., Sullivan, K., & Conners, C. K. (1996). The Multidimensional Anxiety Scale for Children (MASC): Reliability and validity. *Journal of Abnormal Child Psychology,* submitted.

March, J. S., Wells, K., & Conners, C. K. (1995a). Attention-deficit/hyperactivity disorder: Part I. Assessment and diagnosis. *Journal of Practical Psychiatry and Behavioral Health, 1,* 219–228.

March, J. S., Wells, K., & Conners, C. K. (1995b). Attention-deficit/hyperactivity disorder: Part II. Treatment strategies. *Journal of Practical Psychiatry and Behavioral Health, 2,* 23–32.

Martens, B. (1993). Social labeling, precision of measurement, and problem solving: Key issues in the assessment of children's emotional problems. *School Psychology Review, 22,* 308–312.

Mash, E. J., & Johnston, C. (1982). A comparison of the mother–child interactions of younger and older hyperactive and normal children. *Child Development, 53,* 1371–1381.

Mash, E. J., & Terdel, L. G. (1988). Behavioral assessment of child and family disturbance. In E. J. Mash & L. G. Terdel (Eds.), *Behavioral assessment of childhood disorders* (2nd ed., pp. 3–69). New York: Guilford Press.

McBurnett, K., Lahey, B. B., & Pfifner, L. J., (1993).

Diagnosis of attention deficit disorders in DSM-IV: Scientific basis and implications for education. *Exceptional Children, 60,* 108–117.

McClellan, J. M., Rubert, M. P., Reichler, R. J., & Sylvester, C. E. (1990). Attention deficit disorder in children at risk for anxiety and depression. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 534–539.

McCaughy, S. (1993a). Advances in empirically-based assessment of children's behavioral and emotional problems. *School Psychological Review, 22,* 285–307

McConaughy, S. (1993b). Responses to commentaries on advances in empirically based asessment. *School Psychology Review, 22,* 334–342.

McKinney, J., Montague, M., & Hocutt, A. (1993). *Synthesis of research on the assessment and identification of students with attention deficit disorder*. Washington, DC: Office of Special Education Programs, Office of Special Education and Rehabilitative Services, US Department of Education.

McMahon, R. (1981). Biological factors in childhood hyperkinesis: A review of genetic and biochemical hypotheses. *Journal of the American Academy of Child and Adolescent Psychiatry, 37,* 12–21.

McMahon, R., & Forehand, R. (in press). *Helping the noncompliant child: A clinician's guide to parent training*. New York: Guilford Press.

Mesulam, M. M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology, 28,* 597–613.

Milich, R., & Kramer, J. (1984). Reflections on impulsivity: An empirical investigation of impulsivity as a construct. In K. Gadow & I. Bialer (Eds.), *Advances in learning and behavioral disabilities* (Vol. 3, pp. 57–94). Greenwich, CT: JAI Press.

Milich, R., & Landau, S. (1982). Socialization and peer relations in hyperactive children. In K. D. Gadow & I. Bialer (Eds.), *Advances in learning and behavioral disabilities* (Vol. 1, pp. 283–339). Greenwich, CT: JAI Press.

Milich, R., Wolraich, M., & Lindgren, S. (1986). Sugar and hyperactivity: A critical review of empirical findings. *Clinical Psychology Review, 6,* 493–513.

Millichap, J., & Fowler, G. W. (1967). Treatment of "minimal brain dysfunction" syndrome. *Pediatric Clinics of North America, 14,* 767–777.

Morris, R. J., & Collier, S. J. (1987). Assessment of attention deficit disorder and hyperactivity. In C. L. Frame & J. L. Matson (Eds.), *Handbook of assessment in childhood psychopathology: Applied issues in differential diagnosis and treatment evaluation. Applied clinical psychology*.

Newby, R. F., Fischer, M., & Roman, M. A. (1991). Parent training for families of children with attention deficit-hyperactivity disorder. *School Psychology Review, 20,* 252–265.

Newcorn, J. H., & Halperin, J. M. (1994). Comorbidity among disruptive behavior disorders: Impact on severity, impairment, and response to treatment. *Child and Adolescent Psychiatric Clinics of North America, 3,* 227–252.

Nichols, P. L., & Chen, T. C. (1981). *Minimal brain dysfunction: A prospective study*. Hillsdale, NJ: Erlbaum.

O'Connor, M., Foch, T., Sherry, T., & Plomin, R. (1980). A twin study of specific behavioral problems of socialization as viewed by parents. *Journal of Abnormal Child Psychology, 8*(2), 189–199.

Offord, D. R., Boyle, M. H., Szatmari, P., Rae-Grant, N. I., Links, P. S., Cadman, D. T., Byles, J. A., Crawford, J. W., Blum, H. M., & Byrne, C. (1987). Ontario Child Health Study: II. Six-month prevalence of disorder and rates of service utilization. *Archives of General Psychiatry, 44,* 832–836.

O'Leary, K. D., Vivian, D., & Nisi, A. (1985). Hyperactivity in Italy. *Journal of Abnormal Child Psychology, 13,* 485–500.

Olson, R. K., Rack, J. P., Conners, F. A., DeFries, J. C., & Fulker, D. W. (1991). Genetic etiology of individual differences in reading disability. In L. V. Feogans, E. J. Short, & L. J. Meltzer (Eds.), *Subtypes of learning disabilities: Theoretical perspectives and research* (pp. 113–135). Hillsdale, NJ: Erlbaum.

Olson, R., Wise, B., Conners, F. A., & Rack, J. (1990). Organization, heritability, and remediation of component word recognition and language skills in disabled readers. In T. H. Carr & B. A. Levy, (Eds.), *Reading and its development: Component skills approaches* (pp. 261–322). San Diego, CA: Academic Press.

Olson, R., Wise, B., Conners, F. A., Rack, J., & Fulker, D. W. (1989). Specific deficits in component reading and language skills: Genetic and environmental influences. *Journal of Learning Disabilities, 22,* 339–348.

Ostrom, N. N., & Jenson, W. R. (1988). Assessment of attention deficits in children. *Professional School Psychology, 3*(4), 253–269.

Ottenbacher, J., & Cooper, H. (1983). Drug treatment of hyperactivity in children. *Developmental Medicine and Child Neurology, 25,* 358–366.

Parker, H. C. (1992). *The ADD hyperactivity handbook for schools*. Plantation, FL: Impact.

Parker, J. D. A., Sitarenios, G., & Conners, C. K. (1996). Abbreviated Conners' Rating Scales revisited: A confirmatory factor analytic study. *Journal of Attention Disorders, 1,* 55–62.

Parker, J. G., & Asher, S. R. (1987). Peer relations and later personal adjustment: Are low-accepted children at risk? *Psychological Bulletin, 102,* 357–389.

Pasamanick, B., & Knobloch, H. (1959). Syndrome of minimal cerebral damage in infancy. *Journal of the American Medical Association, 170,* 1384–1387.

Pelham, W. E. (1986). The effects of psychostimulant drugs on learning and academic achievement in children with attention-deficit disorders and learning disabilities, In J. K. Torgesen & B. Wong (Eds.), *Psychological and Educational Perspectives on Learning Disabilities* (pp. 259–295). New York: Academic Press.

Pelham, W. E., & Bender, M. E. (1982). Peer relationships in hyperactive children: Description and treatment. In K. Gadow & I. Bialer (Eds.), *Advances in learning and behavioral disabilities* (Vol. 1, pp. 365–436). Greenwich, CT: JAI Press.

Pelham, W. E., Bender, M., Caddell, J., Booth, S., & Moorer, S. (1985), The dose–response effects of methylphenidate on classroom academic and social behavior in children with attention deficit disorder. *Archives of General Psychiatry. 42*(10), 948–952.

Pelham, W. E., Greenslade, K. E., Vodde-Hamilton, M., Murphy, D. A., Greenstein, J. J., Gnagy, E. M., Guthrie, K. J., Hoover, M. D., & Dahl, R. E. (1990). Relative efficacy of long-acting stimulants on children with attention deficit–hyperactivity disorder: A comparison of standard methylphenidate, sustained-release methylphenidate, sustained-release dextroamphetamine, and pemoline. *Pediatrics, 86*(22), 226–237.

Pelham, W. E., & Hinshaw, S. P. (1992). Behavioral intervention for ADHD. In S. M. Turner, K. S., Calhoun, & H. E. Adams (Eds.), *Handbook of clinical behavior therapy* (2nd ed., pp. 259–283). New York: Wiley.

Pelham, W. E., & Hoza, B. (1996). Intensive treatment: A summer treatment program for children with ADHD. In E. D. Hibbs & P. S. Jensen (Eds.), *Psychosocial treatments for child and adolescent disorders: Empirically-based strategies for clinical practice,* (pp. 311–340). Washington, DC: American Psychological Association.

Pelham, W. E., & Murphy, H. A. (1986). Behavioral and pharmacological treatment of attention deficit and conduct disorders. In M. Hersen (Ed.), *Pharmacological and behavioral treatment: An integrative approach* (pp. 108–148). New York: Wiley.

Pelham, W. E., Sturges, J., Hoza, J., Schmidt, C., Bijlsma, J. J., Milich, R., & Moorer, S. (1987). Sustained release and standard methylphenidate effects on cognitive and social behavior in children with attention deficit disorder. *Pediatrics, 80,* 491–501.

Pfiffner, L. J., & Barkley, R. A. (1990). Educational placement and classroom management. In R. A. Barkley, *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (pp. 498–539). New York: Guilford Press.

Pfiffner, L. J., & O'Leary, S. G. (1993). School-based psychological treatments. In J. L. Matson (Ed.), *Handbook of hyperactivity in children* (pp. 234–255). Boston: Allyn & Bacon.

Pliszka, S. R. (1987). Tricyclic antidepressants in the treatment of children with attention deficit disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 127–132.

Pliszka, S. R. (1989). Effect of anxiety on cognition, behavior, and stimulant response in ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 882–887.

Porrino, L. J., Rapoport, J. L., Behar, D., Sceery, W., Ismond, D. R., & Bunney, W. E. (1983). A naturalistic assessment of the motor activity of hyperactive boys: I. Comparison with normal controls. *Archives of General Psychiatry, 40,* 681–687.

Posner, M. (1988). Structures and function of selective attention. In M. Dennis, E. Kaplan, M. Posner, D. Stein, & R. Thompson (Eds.), *Clinical neuropsychology and brain function: Research, measurement, and practice* (pp. 169–201). Washington, DC: American Psychological Association.

Prendergast, M., Taylor, E., Rapoport, J. L., Bartko, J., Donnelly, M., Zametkin, A., Ahearn, M., Dunn, G., Wieleberg, H. (1988). The diagnosis of childhood hyperactivity: A US–UK cross-national study of *DSM-III* and *ICD-9*. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. *29*(3), 289–300

Prinz, R., & Riddle, D. (1986). Association between nutrition and behavior in five-year old children. *Nutrition Reviews, 44,* 151–158.

Rapoport, J. L., Buchsbaum, M. S., Weingartner, H., Zahn, T. P., Ludlow, C., & Mikkelsen, E. J. (1980). Dextroamphetamine. Its cognitive and behavioral effects in normal and hyperactive boys and normal men. *Archives of General Psychiatry, 37,* 933–943.

Rapoport, J., & Castellanos, X. (1996). Attention-deficit/hyperactivity disorder. In J. M. Weiner (Ed.), *Diagnosis and psychopharmacology of childhood and adolescent disorders* (2nd ed., pp. 265–292). New York: Wiley.

Rapport, M. D., DuPaul, G. J., Stoner, G., & Jones, J. T. (1986). Comparing classroom and clinic measures of attention deficit disorder: Differential, idiosyncratic, and dose–response effects of methylphenidate. *Journal of Consulting and Clinical Psychology, 54,* 334–341.

Rapport, M. D., & Kelly, K. L. (1991). Psychostimulant effects on learning and cognitive function: Findings and implications for children with attention deficit hyperactivity disorder. *Clinical Psychology Review, 11,* 61–92.

Reid, R. (1995). Assessment of ADHD with culturally different groups: The use of behavioral rating scales. *School Psychology Review, 24,* 537–560.

Reid, R., Vasa, S. F., Maag, J. W., & Wright, G. (1994). An analysis of teachers' perceptions of attention deficit-hyperactivity disorder. *Journal of Research and Development in Education, 27,* 195–202.

Richman, N., Stevenson, J., & Graham, P. (1982). *Preschool to school: A behavioral study*. London: Academic Press.

Richters, J. E., Arnold, L. E., Jensen, P. S., Abikoff, H., Conners, C. K., Greenhill, L. L., Hechtman, L., Hinshaw, S. P., Pelham W. E., & Swanson, J. M. (1995). NIMH collaborative multisite multimodal treatment study of children with ADHD: I. Background and rationale. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 987–1000.

Roberts, M. A. (1990). A behavioral observation method for differentiating hyperactive and aggressive boys. *Journal of Abnormal Child Psychology, 18,* 131–142.

Robin, A. L., & Foster, S. L. (1989). *Negotiating parent–adolescent conflict*. New York: Guilford Press.

Ross, P. A., Poidevant, J. M., & Miner, C. U. (1995). Curriculum-based assessment of writing fluency in children with attention-deficit hyperactivity disorder and normal children. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 11,* 201–208.

Routh, D. K. (1978). Hyperactivity. In P. Magrab (Ed.), *Psychological management of pediatric problems* (pp. 3–48). Baltimore: University Park Press.

Rutter, M., & Casaer, P. (1991). *Biological risk factors for psychosocial disorders*. New York: Cambridge University Press.

Sallee, F. R., Stiller, R. L., & Perel, J. M. (1992). Pharmacodynamics of pemoline in attention deficit disorder with hyperactivity. *The Journal of the Academy of Child and Adolescent Psychiatry, 31,* 244–251.

Satterfield, J. H., Satterfield, B. T., & Schell, A. M. (1987). Therapeutic interventions to prevent delinquency in hyperactive boys. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 56–64.

Schachar, R., Logan, G., Wachsmuth, R., & Chajczyk, D. (1988). Attaining and maintaining preparation: A comparison of attention in hyperactive, normal, and disturbed control children. *Journal of Abnormal Child Psychology, 16,* 361–378.

Schachar, R., & Wachsmuth, R. (1990). Hyperactivity and parental psychopathology. *Journal of Child Psychology and Psychiatry, 31,* 381–392.

Schaughency, E. A., & Rothlind, J. (1991). Assessment and classification of attention deficit hyperactive disorders. *School Psychology Review, 20*(2), 187-202.

Seham, M., & Seham, G. (1929). The relation between malnutrition and nervousness. *American Journal of Diseases of Children, 37*(1), 1–5.

Seidel, W. T., & Joschko, M. (1990). Evidence of difficulties in sustained attention in children with ADDH. *Journal of Abnormal Child Psychology, 18,* 217–229.

Semrud-Clikeman, M., Biederman, J., Sprich-Buckminster, S., Lehman, B. K., Faraone, S. V., & Norman, D. (1992). Comorbidity between ADHD and learning disability: A review and report in a clinically referred sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 439–448.

Sergeant, J., & Scholten, C. A. (1985). On resource strategy limitations in hyperactivity: Cognitive impulsivity reconsidered. *Journal of Child Psychology and Psychiatry, 26,* 97–109.

Sharma, V., Matier, K., & Halperin, J. M. (1994). Disruptive behavior disorders: Assessment and differential diagnosis. *Child and Adolescent Psychiatric Clinics of North America, 3,* 253–270.

Shaywitz, B. A., Fletcher, J. M., & Shaywitz, S. E. (1995). Defining and classifying learning disabilities and Attention-Deficit/Hyperactivity Disorder. *Journal of Child Neurology, 10*(Suppl. 1), S50–S57.

Shekim, W. O., Kashani, J., Beck, N., Cantwell, D. P., Martin, J., Rosenberg, J., & Costello, A. (1985). The prevalence of attention deficit disorders in a rural midwestern community sample of nine-year-old children.

Journal of the American Academy of Child Psychiatry, 24(6), 765–770.

Solanto, M. V., & Connors, C. K. (1982). A dose–response and time–action analysis of autonomic and behavioral effects of methylphenidate in attention deficit disorder with hyperactivity. *Psychophysiology, 19*(6), 658–667.

Solis, C. R. P., & Solis, C. V. P. (1988). Estimacion del grado de actividad de los escolares segun los padres y maestras, y su relacion con la impulsividad cognoscitiva. (Estimation of the activity level of schoolchildren according to parents and teachers, and its relation with cognitive impulsivity.). *Salud Mental, 11,* 30–39.

Sprague, R. L., & Sleator, E. K. (1977). Methylphenidate in hyperkinetic children: Differences in dose effects on learning and social behavior. *Science, 198,* 1274–1276.

Stevens, G. (1981). Bias in the attribution of hyperkinetic behavior as a function of ethnic identification and socioeconomic status. *Psychology in the Schools, 18,* 99–106.

Stevenson, J. (1992). Evidence for a genetic etiology in hyperactivity in children. *Behavior Genetics, 22*(3), 337–344.

Still, G. (1902). Some abnormal psychical conditions in children. *Lancet, 1,* 1008–1012, 1077–1082, 1163–1168.

Strauss, A., & Lehtinen, L. (1947). *Psychopathology and education of the brain-injured child.* New York: Grune and Stratton.

Swanson, J. M., Cantwell, D., Lerner, M., McBurnett, K., & Hanna, G. (1991). Effects of stimulant medication on learning in children with ADHD. *Journal of Learning Disabilities, 24,* 219–230.

Swanson, J. M., McBurnett, K., Christian, D. L., & Wigal, T. (1995). Stimulant medications and the treatment of children with ADHD. *Advances in Clinical Child Psychology, 17,* 265–322.

Swanson, J. M., McBurnett, K., Wigal, T., Pfiffner, L., Lerner, M. A., Williams, L., Christian, D. L., Tamm, L., Willcutt, E., Crowley, K., Clevenger, W., Khouzam, N., Woo, C., Crinella, F. M., & Fisher, T. D. (1993). Effect of stimulant medication on children with attention deficit disorder: A review of reviews. *Exceptional Children, 60,* 154–162.

Szatmari, P., Offord, D. R., & Boyle, M. H. (1989a). Correlates, associated impairments and patterns of service utilization of children with attention deficit disorder: Findings from the Ontario Child Health Study. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 30,* 205–217.

Szatmari, P., Offord, D. R., & Boyle, M. H. (1989b). Ontario Child Health Study: Prevalence of attention deficit disorder with hyperactivity. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 30,* 219–230.

Tao, K. T., (1992). Hyperactivity and attention deficit disorder syndromes in China. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 1165–1166.

Taylor, E. (1989). On the epidemiology of hyperactivity. In T. Sagnolden & T. Archer (Eds), *Attention deficit disorder: Clinical and basic research* (pp. 31–52). Hillsdale, NJ: Erlbaum.

Taylor, E. (1994). Similarities and differences in *DSM-IV* and *ICD-10* diagnostic criteria. *Child and Adolescent Clinics of North America, 3,* 209–226.

Taylor, E., & Sandberg, S. (1984). Hyperactive behavior in English schoolchildren: A questionnaire survey. *Journal of Abnormal Child Psychology, 12,* 143–155.

Thatcher, R., McAlaster, R., Lester, M., & Cantor, D. (1984). Comparisons among EEG, hair minerals and diet predictions of reading performance in children. *Annals of the New York Academy of Sciences, 425,* 421–423.

Thurber, S., & Walker, C. (1983). Medication and hyperactivity: A meta-analysis. *The Journal of General Psychology, 108,* 79–86.

Umbreit, J. (1995). Functional assessment and intervention in a regular classroom setting for the disruptive behavior of a student with attention deficit hyperactivity disorder. *Behavioral Disorders, 20*(4), 267–278.

van der Meere, J., van Baal, M., & Sergeant, J. (1989). The additive factor method: A differential diagnostic tool in hyperactivity and learning disability. *Journal of Abnormal Child Psychology, 17,* 409–422.

van der Meere, J., Wekking, E., & Sergeant, J. (1991). Sustained attention and pervasive hyperactivity. *Journal of Child Psychology and Psychiatry, 32,* 275–284.

Voeller, K. K. S. (1991). What can neurological models of attention, intention, and arousal tell us about attention-deficit hyperactivity disorder? *Journal of Neuropsychiatry, 3,* 209–216.

Volkmar, F. R. (1991). Classification in child and adolescent psychiatry: Principles and issues. In M. Lewis (Ed.), *Child and adolescent psychiatry: A comprehensive textbook* (pp. 415–421). Baltimore: Williams & Wilkins.

Waechter, D., Anderson, R. P., Juarez, L. J., Langsdorf, R., & Madrigal, J. F. (1979). Ethnic group, hyperkinesis, and modes of behavior. *Psychology in the Schools, 16,* 435–439.

Walker, J. L., Lahey, B. B., Hynd, G. W., & Frame, C. L. (1987). Comparison of specific patterns of antisocial behavior in children with conduct disorder with or without coexisting hyperactivity. *Journal of Consulting and Clinical Psychology, 55,* 910–913.

Weinberg, W. A., & Brumback, R. A. (1990). Primary disorder of vigilance: A novel explanation of inattentiveness, daydreaming, boredom, restlessness, and sleepiness. *Journal of Pediatrics, 116,* 720–725.

Weiss, G., & Hechtman, L. T. (1986). *Hyperactive children grown up.* New York: Guilford Press.

Weiss, G. & Hechtman, L. T. (1993). *Hyperactive children grown up: ADHD in children, adolescents, and adults* (2nd ed.). New York: Guilford Press.

Weissbluth, M. L. K. (1983). Sleep patterns, attention span, and infant temperament. *Journal of Developmental and Behavioral Pediatrics, 4,* 34–36.

Weissbluth, M. (1984). Sleep duration, temperament, and Conners' ratings of three-year-old children. *Journal of Developmental and Behavioral Pediatrics, 5,* 120–123.

Wells, K. C. (1995). Parent management training. In G. P. Sholevar (Ed.), *Conduct disorders in children and adolescence* (pp. 213–236). Washington, DC: American Psychiatric Press.

Wells, K. C. (in press). Adaptations for specific populations. In R. McMahon & R. Forehand (Eds.), *Helping the noncompliant child: A clinician's guide to parent training.* New York: Guilford Press.

Wender, E. H. (1995). Attention deficit hyperactivity disorders in adolescence. *Developmental and Behavioral Pediatrics, 16,* 192–195.

Werner, E., Bierman, J., French, F., Simonian, K., Connor, A. Smith, R., & Campbell, M. (1968). Reproductive and environmental casualties: A report on the 10-year follow-up of the children of the Kauai pregnancy study. *Pediatrics, 42,* 112–127.

Werner, E., & Smith, R. (1967). *Kauai's children come of age.* Honolulu, HI: University of Hawaii Press.

Whalen, C. K., & Henker, B. (1980). *Hyperactive children: The social ecology of identification and treatment.* New York: Academic Press.

Whalen, C. K., & Henker, B. (1992). The social profile of attention-deficit hyperactivity disorder: Five fundamental facets. *Child and Adolescent Psychiatric Clinics of North America, 1,* 395–410.

White, J. L., Moffitt, T. E., Caspi, A., Bartusch, D. J., Needles, D. J., & Stouthamer-Loeber, M. (1994). Measuring impulsivity and examining its relation to delinquency. *Journal of Abnormal Psychology, 103,* 192–205.

Winsberg, B., Press, M., Bialer, I., & Kupietz, S. (1974). Dextroamphetamine and methylphenidate in the treatment of hyperactive/aggressive children. *Pediatrics, 53,* 236–241.

Wolraich, M. L., Wilson, D. B., & White, J. W. (1995). The effect of sugar on behavior or cognition in children. *Journal of the American Medical Association, 274,* 1617–1621.

World Health Organization (1992a). *International statistical classification of diseases and related health problems* (10th revision). Geneva, Switzerland: Author.

World Health Organization. (1992b). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva, Switzerland: Author.

Yao, K. N., Solanto, M. V., & Wender, E. H. (1988). Prevalence of hyperactivity among newly immigrated Chinese–American children. *Journal of Developmental and Behavioral Pediatrics, 9,* 367–373.

Yu, C. S., Yu, F. W., & Xiao, L. Y. (1985). An epidemiological investigation of minimal brain dysfunction in six elementary schools in Beijing. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 26*(5), 777–787.

Zametkin, A. J. (1989). The neurobiology of attention-deficit hyperactivity disorder: A synopsis. *Psychiatric Annals, 19,* 584–586.

Zametkin, A. J., & Rapport, J. L. (1987). Noradrenergic hypothesis of attention deficit with hyperactivity: A critical review. In H. Y. Meltzer (Ed.), *Psychopharmacology: The third generation of progress* (pp. 837–842). New York: Raven.

Zentall, S. S. (1989). Production deficiencies in elicited language but not in the spontaneous verbalizations of hyperactive children. *Journal of Abnormal Child Psychology, 16,* 657–673.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.20
Conduct Disorders

RONALD J. PRINZ
University of South Carolina, Columbia, SC, USA

5.20.1 INTRODUCTION 527
5.20.2 CONDUCT DISORDERS 527
  5.20.2.1 Phenomenology 527
  5.20.2.2 Diagnostic Considerations 528
  5.20.2.3 Prevalence 529
  5.20.2.4 Etiology 529
  5.20.2.5 Developmental Course 530
5.20.3 CONCEPTUALIZATION AND CLINICAL FORMULATION 531
  5.20.3.1 Multimodal Assessment 531
  5.20.3.2 Treatment Formulation 531
5.20.4 TREATMENT 532
  5.20.4.1 Psychosocial Treatments 532
  5.20.4.2 Pharmacological Treatments 533
  5.20.4.3 Efficacy/Effectiveness of Treatments 533
5.20.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 534
5.20.6 SUMMARY 535
5.20.7 REFERENCES 535

5.20.1 INTRODUCTION

The constellation of conduct disorders in children and adolescents presents a major challenge for society. Conduct disorders account for about half of all youth referrals for mental health services, tend to remain stable or worsen over time, and often cross generations (Costello, 1989; Loeber, 1990). Youth exhibiting various forms of conduct disorders victimize others, disrupt their families, fail at school, commit crimes in the community, and abuse alcohol and other drugs. In some cases, conduct disordered behaviors persist throughout the individual's life. These disorders inflict untold psychological, social, and economic costs on youth, families, victims, and communities, in the form of remedial education, law enforcement, heavy utilization of outpatient and inpatient mental health services, physical and emotional harm to others, property destruction, substance abuse, teen parenthood, school dropout, and adult criminality (Robins, 1981).

5.20.2 CONDUCT DISORDERS

5.20.2.1 Phenomenology

Conduct disorders are characterized by a recurring and persistent pattern of antisocial behavior involving the violation of others' basic rights and major societal norms or rules. The term which used to be singular is now plural in recognition of the multiplicity of profiles exhibited by youth with a conduct disorder. Some youth manifest conduct disorder in terms of overt aggressive and hostile acts towards others, while other youth show a pattern of covert, deceitful acts (e.g., stealing, lying, truancy) without accompanying interpersonal aggression, and still others show a combination of these two patterns of antisocial behavior. For some youth, conduct problems begin around school entry, and for others these problems do not emerge until the start of puberty. Conduct disorder can be found in youth with poorly developed relationships with peers and in youth who have well-developed social networks albeit with deviant peers. In sum, although many of these youth have in common some aspects of antisocial functioning, there is considerable variation in terms of expression, age of onset, and social relations.

Developmental context is important in considering whether aggressive or otherwise antisocial behaviors constitute a conduct disorder. It is common for preschool-age children to engage in some behaviors that bear some resemblance to features of conduct disorder. For example, many four-year olds on at least a sporadic basis take things that do not belong to them, lie to get their way, and display tantrum outbursts that include hitting others or destroying property. Engaging in these behaviors occasionally at age four is sufficiently common to be considered a usual aspect of development. The same behaviors at age ten are developmentally out of step with agemates and are more likely to be indicative of the conduct-disorder constellation. Similarly but at the other end of childhood, the developmental norms of adolescence need to be considered. Adolescence is a period of development marked by rebelliousness punctuated by some acting out behavior. Occasional rule or law breaking, such as skipping school or writing graffiti in a public place, is sufficiently commonplace to be considered almost normative developmentally in adolescence. To distinguish between normative misconduct and antisocial disorder, the frequency and seriousness of the behaviors need to be considered. A competing consideration is that conduct problems including serious misdeeds might be increasing among adolescents to the point that such acts are becoming almost normative in frequency (Coie & Jacobs, 1993).

5.20.2.2 Diagnostic Considerations

In the *Diagnostic and statistical manual of mental disorders* (*DSM-IV*; American Psychiatric Association, 1994), Conduct disorder (CD) falls within a larger domain called attention-deficit and disruptive behavior disorders. CD shares this domain with somewhat related diagnoses of oppositional defiant disorder (ODD) and attention-deficit hyperactivity disorder (ADHD). The *DSM-IV* diagnosis of CD requires the presence of three or more manifestations of repeated and persistent antisocial behavior over a 12-month period before age 18, with the behavioral disturbances resulting in significant impairment in social, academic, or job functioning.

Specifically, the three (or more) manifestations of antisocial behavior can be any of 15 criteria that are distributed across four domains: aggression towards people or animals, destruction of property, deceitfulness or theft, and serious violations of rules. Aggression includes bullying or threatening, fighting, use of dangerous weapons, physical cruelty to people or animals, confrontational theft, and coercive sexual activity. Property destruction includes vandalism or firesetting. Deceitfulness or theft includes breaking and entering, lying to obtain goods or favors or to avoid obligations, and nonconfrontive theft. Serious violations of rules includes staying out late even against parental prohibitions (beginning before age 13), running away, and school truancy.

A major departure by *DSM-IV* over previous versions is the specification of two CD types based on age at onset. *DSM-IV* distinguishes between childhood and adolescent onset. Childhood-onset CD begins prior to age 10 and requires at least one of the criterion manifestations of antisocial behavior before that age. Adolescent-onset CD begins in adolescence but also requires that none of the criterion manifestations occurred prior to age 10. The distinction between early vs. later onset is consistent with differing developmental patterns that have emerged in longitudinal and etiological research.

The *ICD-10* classification system (World Health Organization, 1992) also references conduct disorders in children and adolescence but uses an organizing scheme different from that found in *DSM-IV*. The core definition of conduct disorder in the two systems is essentially the same. However, instead of making the distinction between early and later onset, *ICD-10* identifies different subtypes. According to *ICD-10*, CD can be confined to the family context, which means that antisocial and aggressive behavior is confined primarily to interactions with family members. The unsocialized and socialized types of CD, which were present in earlier *DSM* versions, are part of *ICD-10* and distinguish between youth who have seriously disturbed relationships with other children and youth who are well integrated into their peer group. *ICD-10* categorizes ODD as one type of conduct disorder, whereas *DSM-IV* treats ODD as a category distinct from CD. However, both systems define ODD as the presence of markedly defiant, disobedient, and noxious behavior without the more severe aggressive or antisocial acts that violate the law or others' rights. Furthermore, both systems note that ODD is typically seen in children below the age of 9 or 10 and can be a precursor to CD in some children at a young age.

5.20.2.3 Prevalence

Estimates for the prevalence of CD have varied considerably. Costello (1989) reported that CDs affect approximately 5–8% of the general child population. Kazdin's (1995) review of prevalence indicated that the estimated rate of conduct disorders in children aged 4–18 has ranged from 2% to 6%. *DSM-IV* (American Psychiatric Association, 1994) reported that CD rates for youth under age 18 range from 6% to 16% for males and from 2% to 9% for females, although the basis or source of these estimates is not provided.

It is unclear specifically what has accounted for the wide variation in prevalence estimates. The actual obtained rate, however, can be greatly influenced by the threshold criteria for each of the symptom categories. For example, the symptom category pertaining to initiation of physical fights can be stretched or contracted depending on the frequency and intensity one requires for the fighting criterion. Kazdin (1995) noted that adolescents in the general population report very high rates of theft, assault, and vandalism, which implies that prevalence rates of CD diagnosis based on adolescent self-report could be quite high. In other words, the source of diagnostic information could significantly affect estimate of prevalence. The high variation in prevalence estimates could also reflect different time periods and sampling methods. Finally, there is some evidence to suggest that CD is becoming more prevalent in females over time (Loeber, 1990; Viemero, 1992) while not becoming less prevalent in males. Nonetheless, it is clear that CD is a highly prevalent problem that will not go away on its own.

5.20.2.4 Etiology

Much of the etiological work related to CD has focused on the study of risk factors, and to a lesser extent protective factors, that may impinge on the development of antisocial behavior. A number of significant clusters of risk factors have emerged over many years of research via longitudinal and cross-sectional studies. Some of the main categories of risk for CD include temperament, familial factors, school and peer factors, and sociological factors.

Although there is tremendous variability in developmental outcome, toddlers who show difficult temperaments are at heightened risk of developing conduct problems at school entry which might escalate over time (Campbell, 1991; Campbell, Breaux, Ewing, & Szumowski, 1986). Young children who exhibit greater impulsivity, emotional lability, and irritability than their agemates often face adverse environmental reactions if their behavioral style persists through early childhood. Given that most toddlers exhibit occasional outbursts and impatience, useful delineation of this risk factor emphasizes the frequent, chronic, and persistent nature of temperament difficulties.

Temperament difficulties do not typically persist or escalate independent of environmental context. Here is where familial risk factors may operate to potentiate temperamental risk. It is reasonable to hypothesize that parenting difficulties and family adversity exacerbate the adverse impact of child temperament difficulties, and that child temperament difficulties also add to parental and familial stress (Bates, Bayles, Bennett, Ridge, & Brown, 1991; Campbell et al., 1986; Dumas, 1992; Loeber & Stouthamer-Loeber, 1986; Moffitt, 1990; Offord, Boyle, & Racine, 1991). Such a transactional perspective does not preclude the possibility that there are multiple pathways to conduct disorder. For example, at the extreme ends of the spectrum, in children who do not exhibit early temperament difficulties, marked parenting problems such as severe child abuse can still lead to the development of conduct disorder. Conversely, for children showing intense difficulties with early temperament, positive parenting may not be sufficient to prevent the emergence of conduct disorder. However, any adequate etiological model for conduct disorder also must account for what happens to the many children between those two extremes.

A strong body of evidence exists supporting the contention that parenting difficulties during early and middle childhood significantly influence the development of early conduct problems (Patterson, 1982; Patterson, Reid, & Dishion, 1992; Wahler & Dumas, 1987). Relevant parenting dimensions include poor discipline (Farrington & West, 1981; McCord, 1988; Patterson et al., 1992), lack of positive parental response (Loeber & Stouthamer-Loeber, 1986; Wadsworth, 1980), and coercive family process (Patterson, 1982). Other risk factors in the family context, which may disrupt parenting and contribute to risk for CD, include alcohol and illicit drug abuse by caregivers (West & Prinz, 1987), parental social insularity and social disadvantage (Dumas, 1986; Wahler & Dumas, 1987), parental criminality and psychopathology (Offord, 1982; Robins, 1981), and marital discord (Offord & Boyle, 1986; Rutter & Giller, 1983).

Children already showing disruptive and inattentive behaviors in class tend to develop learning difficulties as well (Moffitt, 1990, 1993a, 1993b), which adds to their frustration. Learning difficulties plus the aversive responses from teachers provoked by aggressive and disruptive behaviors help to create negative school experiences that accelerate conduct problems (Campbell, 1991; Hawkins & Lishner, 1987; Hawkins & Weis, 1985; Hirschi, 1969; Walker & Buckley, 1973). Academic difficulties in elementary school and school failure leading to truancy and dropout during adolescence are both potential contributors to the development of serious antisocial behavior (Loeber, 1990). In addition, there is some evidence that certain school environments, such as those that are disorganized with large classes and nonrewarding practices, can exacerbate antisocial behavior (Rutter, Maughan, Mortimore, Ouston, & Smith, 1979; Rutter et al., 1974).

Disturbed relations with peers is another common source of risk, in the form of rejection by mainstream peers followed by attraction to deviant peer groups (Dishion, Patterson, Stoolmiller, & Skinner, 1991; Elliott, Huizinga, & Ageton, 1985; Snyder, Dishion, & Patterson, 1986; West & Farrington, 1977). A common assertion in the literature is that peers exert a strong influence that can promote antisocial behavior including substance abuse (Dishion & Andrews, 1995; Elliott et al., 1985). However, there is some disagreement about whether the peer influence is a primary causal variable or just a concomitant outcome. For example, Block, Block, and Keyes (1988) demonstrated that early family socialization practices exert a more cause influence than peer variables, particularly on substance abuse. On the other hand, adolescent-onset CD seems to be more closely tied to exposure to deviant peer influences than childhood-onset CD (Capaldi & Patterson, 1994).

Another perspective on possible etiological mechanisms derives from a social information processing framework. Specifically, Dodge and others have identified deficiencies in critical social-cognitive skills that are related to aggressive tendencies (Crick & Dodge, 1994; Dodge, 1985; Dodge, Bates, & Pettit, 1990; Dodge, Pettit, McClaskey, & Brown, 1986). Key features include an attributional bias towards presuming hostile intent in others (Dodge & Frame, 1982; Lochman, 1987), failure to attend to relevant social cues (Dodge et al., 1990), low skill level in social problem-solving (Asarnow & Callan, 1985; Lochman, Lampron, & Rabiner, 1989), and a tendency to view aggression as leading to positive outcomes (Perry, Perry, & Rasmussen, 1986; Dodge et al., 1986).

Yet another category of etiological theory about CD derives from a set of approaches inadequately identified as sociological in nature. A number of researchers from different perspectives have recognized that the larger social context, including neighborhood, economic stratum, community, and cultural reference group, has some connection to the development and desistance of antisocial patterns of behavior in children and adolescents. For example, social interactionists such as Patterson, Wahler, and their colleagues have demonstrated how high levels of family adversity associated with economic disadvantage exert indirect effects on youth antisocial behavior via disrupted parenting and related consequences (Capaldi & Patterson, 1994; Wahler & Hann, 1987). Others have shown how community and peer influences coalesce to impact youth behavior (Sampson & Groves, 1989), how exposure to community violence can increase youth antisocial behavior and other adjustment problems (Richters & Martinez, 1993), and how forces of prejudice and racism manifesting in exclusion from the mainstream, threats to self-identity, and demoralizing conditions of hopeless poverty can induce expressions of anger and outrage taking the form of antisocial acts (Boyd-Franklin, 1989; McLoyd, 1990; Ogbu, 1988).

5.20.2.5 Developmental Course

For childhood-onset CD, children who ultimately develop the disorder show prodromal signs at school entry or even earlier. The early starters often exhibit pronounced cross-situational aggressive behavior between ages four and eight (Dumas, 1992; Loeber, 1988, 1990; Loeber & LeBlanc, 1990; Patterson, DeBaryshe, & Ramsey, 1989), which is quite likely to meet the criteria for a diagnosis of ODD. Early aggression is predictive not only of later delinquency (Loeber, 1990) but also subsequent substance abuse (Hawkins, Catalano, & Miller, 1992). Oppositional-aggressive behavior interacts with parenting difficulties to place such children at further risk as they enter school (Patterson et al., 1992). During the early years of elementary school, children at risk for early-onset CD experience learning difficulties (Offord & Waters, 1983; Rutter & Giller, 1983), particularly in reading and language-related skills (Moffitt, 1993a, 1993b). Such children often engage in high rates of inappropriate classroom behavior that includes ignoring teacher instructions, hitting classmates, disrupting class, and destroying property (Prinz, Connor, & Wilson, 1981). Aversive reactions from teachers (Campbell, 1991; Walker & Buckley, 1973) and rejection from peers (Bierman, 1986; Cantrell & Prinz, 1985; Dodge, 1989; Parker & Asher, 1987) often further escalate the children's conduct problems and set the stage for full-blown CD.

Adolescent-onset CD has a somewhat different developmental course. Unlike youth with childhood-onset CD, the late starters develop at least minimal social skills and adequate peer relations, do not show antisocial behavior problems in early elementary school, and seem to acquire delinquent tendencies more from a deviant peer group than from coercive family interactions (Patterson et al., 1989, 1992). Late starters are not necessarily free from family problems, and their antisocial behavior can precipitate family difficulties (or aggravate existing ones). It is presumed that in comparison with early starters, late starters are more likely to desist by early adulthood with respect to delinquent behavior (Patterson et al., 1992).

5.20.3 CONCEPTUALIZATION AND CLINICAL FORMULATION

5.20.3.1 Multimodal Assessment

Appraisal of core CD symptomatology is a significant facet of assessment. For younger children, the main sources of information about conduct problems are parents and teachers via behavioral rating scales. Direct observation of child behavior in social settings (e.g., home, classroom, playground) can also provide useful data about interactive style, particularly with respect to interpersonal aggression. Structured diagnostic interviews of the caregiver have potential utility with younger and older children. For older children, youth report via self-report instruments (behavioral rating scales, behavioral incident checklists) and structured diagnostic interviews add breadth to the assessment of conduct problems.

Several rating scales have acceptable psychometric properties and adequate coverage of content pertinent to conduct disorder. The Child Behavior Checklist, a 113-item instrument, covers youth ranging from 2–16 years of age and has versions for parent, teacher, and youth report (Achenbach & Edelbrock, 1983). The Conners Parent and Teacher Rating Scales are a collection of instruments varying in length and specificity which are particularly useful for assessing comorbid ADHD as well as nuances of conduct problems and antisocial behavior (Conners, 1969, 1970, 1973; Goyette, Conners, & Ulrich, 1978). The Eyberg Child Behavior Inventory, a 36-item instrument designed specifically to assess aspects of conduct and oppositional problems in youth aged 2–16 years, is particularly useful for pinpointing problem areas and assessing treatment effects (Eyberg, 1992; Eyberg & Robinson, 1983; Eyberg & Ross, 1978).

Some intervention programs rely on direct observation, either in the home or in a clinic analogue setting, as a basis for characterizing the specific nature of the problems and for driving treatment implementation (Patterson, 1982; Sanders, Dadds, & Bor, 1989). With younger children, such observations may involve family tasks that involve free play by children and transition tasks that place demands on them. With older children, the observation tasks tend to take the form of problem-discussion tasks for families to attempt. In general, observation seems to be particularly useful for assessing key family processes and other interpersonal interactions (e.g., playground interactions) that might bear on the maintenance of CD, rather than for trying to establish rates and extent of conduct problems.

A major challenge in assessing conduct problems is the variability in types of symptoms that youth with CD can exhibit. While most of the younger children early-onset CD group exhibit demonstrative aggression towards others in public settings, the older youth are more variable. Some adolescents who clearly qualify for a CD diagnosis exhibit little or no overt aggression towards others but instead commit covert acts such as vandalism, cruelty to animals, or nonconfrontational theft. Others only commit aggressive acts in specific circumstances, such as in coercive sexual encounters. The implication of this variability is that there is no single constellation of symptoms to monitor for all youth with CD.

5.20.3.2 Treatment Formulation

Treatment strategies for CD differ somewhat for pre-adolescents vs. adolescents, and the distinction is not simply between early-onset vs. later-onset CD. By definition, pre-adolescent children exhibiting CD symptomatology belong to the early-onset group. However, adolescents in treatment for CD may have either early-onset or later-onset histories but this difference may not manifest in behavioral differences during adolescence. Consequently, CD treatment for CD may vary as a function of severity and setting constraints rather than onset history.

Historically, treatment for CD focused on intrapsychic approaches including play therapy and individual psychotherapy with children, which have met with less recent acceptance in the absence of evidence for efficacy or promise (Kazdin, 1987). Current treatment focuses much more on the social contexts in which children misbehave, namely in the family, in school, and with peers. Much of the current treatment for CD is predicated on the assumption that children with conduct problems need a concentrated regimen of accelerated positive socialization. Treatment strategies emanating from this assumption enlist parents, teachers, and even peers as therapeutic agents. Essentially, the more ecologically oriented interventions attempt to restructure the youth's social environment to shape and reinforce positive behaviors and limit antisocial behavior.

An alternative formulation focuses on the internal cognitive set characteristic of youth who engage in high rates of antisocial behavior. The proponents of the cognitive approach argue that changes in the social environment, in the absence of addressing beliefs and attitudes that lead to antisocial behavior, will not produce lasting positive gains.

5.20.4 TREATMENT

5.20.4.1 Psychosocial Treatments

One of the most common and well-researched psychosocial treatments for pre-adolescent youth with CD is parent training in behavioral management skills. Although there are several variants of parent training (Miller & Prinz, 1990), most approaches have in common a core set of principles and goals: (i) Parent trainers help parents pinpoint, and then monitor, specific positive and negative behaviors that they want to change. (ii) Positive reinforcement, which is a central construct in parent training, is emphasized as more important than punishment. (iii) Parents are encouraged to replace corporal punishment with time-out from reinforcement for the younger children or loss of privileges for the older ones. (iv) The principle of shaping is promoted as a strategy for starting with children's existing behavioral repertoires and expanding their capabilities. (v) Systematic attending and ignoring are key parenting skills that are particularly effective with younger children. (vi) Many of the parent training programs emphasize positive communications by parents to motivate children and strengthen family relationships. (vii) Programs attempt to increase the probability of positive parental attention for child prosocial behaviors (e.g., cooperation, calm talking, appropriate play behavior, sharing, complying with adult requests, staying out of fights) and to assist parents in coping more effectively with child antisocial, aggressive, and noncompliant behaviors.

Parent training is administered in group or individual format (or some combination of both). There are advantages and disadvantages to each format. With the group approach, parents receive support and modeling from other parents, they get to benefit from hearing other parents' questions and the group leader's responses, and the process is more cost efficient. Disadvantages of the group format include: (i) a fixed program that does not permit individual variation in content or pacing, (ii) fewer opportunities to handle each parent's concerns and difficulties, and (iii) inflexibility with respect to scheduling (or rescheduling). The individual format (i.e., therapist works with one family at a time) has a number of advantages: (i) a functional analysis of the particular family can be conducted; (ii) the pace, content, and relative emphasis in therapy can be tailored to the family's needs; (iii) crises can be handled more readily; and (iv) the therapist develops a closer relationship with youth and family that can sustain engagement in treatment. Conversely, therapy with individual families does not engender the same level of social support and normalizing of experiences that the group approach affords.

With adolescents, family-based treatment approaches tend to take the form of family therapy over parent training, with some exceptions (see Patterson, Chamberlain, & Reid, 1982). In contrast to parent training in which the therapist works primarily with the parents, family therapy involves at least the parent(s) and adolescent plus other members of the household such as siblings and extended family. Some of the more prominent family therapy approaches include structural family therapy (Minuchin, 1974; Minuchin & Fishman, 1982; Minuchin, Montalvo, Guerney, Rosman, & Schumer, 1967), functional family therapy (Alexander & Parsons, 1982; Barton & Alexander, 1981; Jameson & Alexander, 1994), and behavioral family therapy (Miller & Prinz, 1990; Tolan & Mitchell, 1989). Although adopting somewhat varied tactics, these three therapy approaches share in common the recognition that youth antisocial behavior is bound up to some degree in family interactions. Furthermore, these approaches all attempt to help families attain their goals and improve relationships and communication within the family.

A relatively new treatment approach has emerged called "multisystemic therapy," which is a hybrid of traditional family therapy and behavior therapy with a strong emphasis on ecological context (Henggeler & Borduin, 1990).

In multisystemic therapy, the therapist works with the family and the youth in the home and elsewhere to alter systemic processes known to be related to antisocial behavior. Multisystemic therapists address parental discipline, family communication and affect, peer associations, school issues, and other potentially significant domains in the larger contexts in which the youth operates.

In addition to interventions that focus primarily on larger contexts such as school and home, there are also youth-focused approaches that involve therapists working one-on-one or in small groups with children or adolescents. The youth-focused approaches arise primarily out of the social-skills and cognitive-behavioral areas (Dush, Hirt, & Schroeder, 1989; Kazdin, 1987; Michelson, 1987). Youth-focused treatments for CD include social skills training, problem-solving training, self-control training, anger management, or some combination of these. Social-skills approaches typically include individual and group reinforcement, coaching and shaping of specific interpersonal behaviors, positive group activities, friendship skills, and listening/speaking skills (Bierman, 1989; Bierman, Miller, & Staub, 1987; Michelson, Sugai, Wood, & Kazdin, 1983; Prinz, Blechman, & Dumas, 1994). Other approaches have combined interpersonal problem-solving and self-instruction (self-talk) in an effort to alter the cognitive bases assumed to underlie conduct problems (e.g., Kazdin, Bass, Siegel, & Thomas, 1989; Kazdin, Esveldt-Dawson, French, & Unis, 1987). Lochman and colleagues have focused on a cognitive-behavioral variant, primarily with young adolescents, that has focused on strategies for anger management (Lochman, 1992; Lochman, Burch, Curry, & Lampron, 1984; Lochman & Curry, 1986).

5.20.4.2 Pharmacological Treatments

Several psychoactive medications have been tried with children or adults who exhibit various forms of aggressive behavior. For example: (i) methylphenidate, D-amphetamine, pemoline citrate, and clonidine have been used with youth with comorbid CD and ADHD (Hinshaw, 1991; Hinshaw, Heller, & McHale, 1992); (i) diphenhydramine, clonazepam, and buspirone have been used with children showing CD combined with anxiety; (iii) fluoxetine, imipramine, desipramine, nortriptyline, and bupropion have been used for comorbid CD and depression; and (iv) lithium carbonate, neuroleptics, propranolol, carbamazepine, antidepressants, and buspirone have been used to address aggressive symptomatology (including self-injurious behavior) in individuals with CD (Campbell & Spencer, 1988; Zavodnick, 1994). However, evidence has not yet emerged in support of pharmacological treatment for the general population of youth with conduct disorder (Stewart, Myers, Burket, & Lyles, 1990). Zavodnick recommends medication primarily as an adjunct to behavioral and parent management treatments, particularly when conduct disorder co-occurs with a diagnosable disorder such as ADHD or depression.

CD and associated problems such as substance abuse, criminality, risky sexual behavior, and school failure are complex constellations of behaviors that are embedded in social contexts. Pharmacological treatment is predicated on an assumption of biological determination (or at least biological influence) that runs orthogonal to a social-contextual analysis. It is not clear at the present time that pharmacological treatment should be the primary approach of choice for many youth whose patterns of antisocial behavior are extricably connected to family, peer, school, and community influences.

5.20.4.3 Efficacy/Effectiveness of Treatments

Social-learning-based parent training for families with a conduct-problem child or young adolescent has met with some success by focusing on parental monitoring and supervision, discipline (e.g., compliance training, timeout, elimination of coercive parenting methods), and quality of parent–child interaction (Dumas, 1989; Kazdin, 1987; Miller & Prinz, 1990). A number of studies conducted by Patterson and colleagues have demonstrated that the majority of children exhibiting pronounced aggressive behavior and related conduct problems respond favorably to systematic parent training interventions (Fleischman, 1981; Patterson, 1974, 1982; Patterson, Chamberlain, & Reid, 1982; Patterson & Fleischman, 1979). Webster-Stratton and her colleagues also found that parenting training, primarily in group-administered formats, produced significant gains for conduct problems in preadolescents over untreated controls that were maintained at one-year follow-up (Webster-Stratton, 1984; Webster-Stratton, Hollinsworth, & Kolpacoff, 1989; Webster-Stratton, Kolpacoff, & Hollinsworth, 1988).

Parent-training investigators have acknowledged that optimum results are not always obtained and that there are obstacles to intervention (Hawkins, Catalano, Morrison, et al., 1992; Kazdin, 1987; Miller & Prinz, 1990).

One challenge is the difficulty in producing lasting changes in families with severe socioeconomic disadvantage and social isolation (Dumas, 1984; Dumas & Wahler, 1983; Patterson & Fleischman, 1979). Another problem is keeping families in treatment without premature attrition, or for that matter getting families into treatment to begin with (McMahon, Forehand, Griest, & Wells, 1981; Miller & Prinz, 1990; Prinz & Miller, 1994). In view of these kinds of difficulties, there is growing support for home-based family interventions (Fraser, Pecora, & Haapala, 1991; Henggeler, Melton, & Smith, 1992; Tharp, 1991) as a way of addressing the needs of culturally and economically diverse populations, avoiding the perceived stigma of clinic-administered treatments, and overcoming the practical obstacles of transportation, childcare, and management of appointments.

Multisystemic therapy (MST; Henggeler & Borduin, 1990) has shown promise with youthful offenders, a most challenging population. Henggeler et al. (1986) compared MST with usual community treatment for juvenile offenders and their families and found that MST was associated with a decrease in reported conduct and related adjustment problems and an improvement in quality of family interaction. Mann, Borduin, Henggeler, and Blaske (1990) randomly assigned chronic juvenile offenders either to MST or individual counseling and found that the youth participating in MST showed significantly lower recidivism over a four-year period (22% vs. 71%). Similar results were reported by Borduin, Henggeler, Blaske, and Stein (1990) in a study applying MST to adolescent sexual offenders. Henggeler, Melton, and Smith (1992) found that MST proved superior to incarceration with respect to subsequent behavior of serious juvenile offenders.

Child cognitive-behavioral therapies have led to at least short-term benefits in observed classroom behavior, academic performance, and teacher/parent reports of aggression and social adjustment (Dush, Hirt, & Schroeder, 1989). Two notable studies targeted truly antisocial samples and evaluated long-term impact (Kazdin, Bass, Siegel, & Thomas, 1989; Lochman, 1992). Kazdin et al. (1989) found that interpersonal problem-solving led to major improvements (immediate and at one-year follow-up) in prosocial behavior and in reductions of aggressive behavior at home and school but cautioned that child-focused intervention may not be sufficient for sustained, clinically significant improvement. Lochman (1992) found that a school-based cognitive-behavioral intervention had significant effects at three-year follow-up in terms lower rates of substance use (alcohol and marijuana) and higher rates of self-esteem in mid-adolescence, compared with untreated peers. However, the intervention had only limited effects on the adolescents' general level of conduct problems, causing Lochman (1992) to suggest that family intervention may be crucial to use in combination with youth-focused interventions.

Social skills interventions have also shown some degree of promise. Michelson et al. (1983) found that a social-skills training program with children aged 8–12 years who had conduct problems produced one-year maintenance of improved social adjustment. Bierman (Bierman, 1989; Bierman, Miller, & Staub, 1987) tested a 10 session social skills training with aggressive-rejected children aged 7–10 years using heterogeneous training groups composed of high-risk and competent children, individual and group reinforcement, coaching and shaping of specific interpersonal behaviors, and enjoyable group activities. Bierman et al. (1987) reported a reduction in antisocial behavior and some improvement in peer acceptance at one-year follow-up. Prinz et al. (1994) evaluated a peer coping-skills intervention that focused on enhancing speaking and listening skills for handling social challenges in small heterogeneous groups (i.e., equal proportions of children with high-risk and socially competent status in the same group). Children aged 7–9 years with high rates of aggressive behavior who received the peer coping-skills program made significant improvement over controls at post-intervention and next-school-year follow-up in terms of a reduced level of aggressive behavior and greater communication effectiveness as judged by teachers (Prinz et al., 1994). Of additional interest is that inclusion of socially competent children without aggressive profiles did not result in iatrogenic effects on those children and actually produced significant gains in communication effectiveness (over socially competent children who did not receive the peer intervention).

5.20.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

From a treatment perspective, one of the most challenging aspects of conduct disorder work has to do with engaging youth and families in the treatment process and maintaining engagement and full participation. Youth with antisocial behavior problems tend to come from families where many of the adults in the household are antisocial themselves and somewhat reticent to participating in treatment. Motivating caregivers to participate meaningfully in the treatment process is as yet an unsolved problem. In a free society, we cannot realistically force youth or parents to participate in treatment. The challenge is to find strategies for motivating and supporting prosocial treatment participation.

A related issue concerns matching of interventions and treatment strategies with particular youth and families. Although subgroups of youth with conduct disorder have been identified, these subgroups are not sufficiently specific or useful to match with treatment strategies. Treatment for conduct disorders often fails, and there are many contributing factors to such failure that need to be studied in order to arrive at clusters of characteristics and circumstances under which one strategy might be more successful than another. Cultural, familial, developmental, and topographical (i.e., type of conduct problems) variables may all play a role.

The future treatment strategies for conduct disorders that have the greatest likelihood for success will probably look like this. (i) Multiple participants, including parents, siblings, teachers, peers, and targeted youth, will contribute to the breadth of interventions. (ii) Interventions will begin early in the trajectory of antisocial behavior. (iii) Greater emphasis will be placed on preventive interventions before children enter the judicial system or are expelled from school. (iv) Interventions will be longer (sometimes lasting for several years), with phases of periodic maintenance and booster contacts. (v) Service delivery systems will rely more heavily on ecologically oriented models that include home visits, neighborhood programs, and family-support approaches.

5.20.6 SUMMARY

The conduct disorders are characterized by a recurring and persistent pattern of antisocial behavior and violation of major societal norms. Antisocial behaviors typically fall into four domains: aggression towards others, destruction of property, theft and other deceit, and serious rule violations (e.g., chronic school truancy). *DSM-IV* distinguishes between childhood and adolescent onset of conduct disorder. Childhood-onset conduct disorder that persists throughout adolescence is associated with greater severity and long-term adverse consequences (e.g., adult criminality, higher rate of violence) than is found with adolescent-onset conduct disorder. Etiological factors include parenting difficulties during early and middle childhood, other familial factors, early temperamental difficulties, school effects, and disturbed peer relations. Assessment methods for conduct disorder include direct observation with younger children, diagnostic interview schedules with youth or caregivers, and rating scales with parent, teacher, or youth informants. The more promising treatment approaches include social-learning-based approaches to parent training and family intervention, an ecological family approach called multisystemic therapy, and child cognitive-behavioral therapies that include social skills interventions, interpersonal problem-solving, and cognitive attributional retraining. Psychopharmacological intervention may be helpful primarily as an adjunctive treatment for youth who have comorbid ADHD or depression. Future research will need to address the challenge of motivating youth and families in treatment and in developing better strategies for matching approaches and incentives to individual cases.

5.20.7 REFERENCES

Achenbach, T. M., & Edelbrock, C. S. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, University of Vermont, Department of Psychiatry.

Alexander, J. F., & Parsons, B. V. (1982). *Functional family therapy: Principles and procedures*. Carmel, CA: Brooks/Cole.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Asarnow, J. R., & Callan, J. W. (1985). Boys with peer adjustment problems: Social cognitive processes. *Journal of Consulting and Clinical Psychology, 53,* 80–87.

Barton, C., & Alexander, J. F. (1981). Functional family therapy. In A. S. Gurman & D. P. Kniskern (Eds.), *Handbook of family therapy* (pp. 403–443). New York: Brunner/Mazel.

Bates, J. E., Bayles, K., Bennett, D. S., Ridge, B., & Brown, M. M. (1991). Origins of externalizing behavior problems at eight years of age. In D. J. Pepler & K. H. Rubin (Eds.), *The development and treatment of childhood aggression* (pp. 93–119). Hillsdale, NJ: Erlbaum.

Bierman, K. (1986). The relationship of social aggression and peer rejection in middle childhood. In R. J. Prinz (Ed.), *Advances in behavioral assessment of children and families* (pp. 51–178). Greenwich, CT: JAI Press.

Bierman, K. (1989). Improving the peer relationships of rejected children. In B. B. Lahey & A. E. Kazdin (Eds.), *Advances in clinical child psychology* (Vol. 12, pp. 53–84). New York: Plenum.

Bierman, K., Miller, C. M., & Staub, S. (1987). Improving the social behavior and peer acceptance of rejected boys: Effects of social skill training. *Journal of Consulting and Clinical Psychology, 55,* 194–200.

Block, J., Block, J. H., & Keyes, S. (1988). Longitudinally foretelling drug usage in adolescence: Early childhood peer and environmental precursors. *Child Development, 59,* 336–355.

Borduin, C. M., Henggeler, S. W., Blaske, D. M., & Stein, R. (1990). Multisystemic treatment of adolescent sexual offenders. *International Journal of Offender Therapy and Comparative Criminology, 34,* 105–113.

Boyd-Franklin, N. (1989). *Black families in therapy: A* *multisystems approach*. New York: Guilford Press.

Campbell, M., & Spencer, E. K. (1988). Psychopharmacology in child and adolescent psychiatry: A review of the past five years. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 267–279.

Campbell, S. B. (1991). Longitudinal studies of active and aggressive preschoolers: Individual differences in early behavior and outcome. In D. Cicchetti & S. L. Toth (Eds.), *Rochester Symposium on Developmental Psychopathology, Vol. 2: Internalizing and externalizing expressions of dysfunction* (pp. 57–90). Hillsdale, NJ: Erlbaum.

Campbell, S. B., Breaux, A. M., Ewing, L. J., & Szumowski, E. K. (1986). Correlates and predictors of hyperactivity and aggression: A longitudinal study of parent-referred problem preschoolers. *Journal of Abnormal Child Psychology, 14,* 217–234.

Cantrell, V. L. & Prinz, R. J. (1985). Multiple perspectives of rejected, neglected, and accepted children: Relationship between sociometric status and behavioral characteristics. *Journal of Consulting and Clinical Psychology, 53,* 884–889.

Capaldi, D. M., & Patterson, G. R. (1994). Interrelated influences of contextual factors on antisocial behavior in childhood and adolescence for males. In D. C. Fowles, P. Sutker, & S. H. Goodman (Eds.), *Progress in experimental personality and psychopathology research* (pp. 165–198). New York: Springer.

Coie, J. D., & Jacobs, M. R. (1993). The role of social context in the prevention of conduct disorder. *Development and Psychopathology, 5,* 263–275.

Conners, C. K. (1969). A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry, 126,* 884–888.

Conners, C. K. (1970). Symptom patterns in hyperkinetic, neurotic, and normal children. *Child Development, 41,* 667–682.

Conners, C. K. (1973). Rating scales for use in drug studies with children. *Psychopharmacological Bulletin, 9,* 24–84.

Costello, E. J. (1989). Developments in child psychiatric epidemiology. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 836–841.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information processing mechanisms in children's social adjustment. *Psychological Bulletin, 115,* 74–101.

Dishion, T. J., & Andrews, D. W. (1995). Preventing escalation in problem behaviors with high-risk young adolescents: Immediate and 1-year outcomes. *Journal of Consulting and Clinical Psychology, 63,* 538–548.

Dishion, T. J., Patterson, G. R., Stoolmiller, M., & Skinner, M. L. (1991). Family, school, and behavioral antecedents to early adolescent involvement with antisocial peers. *Developmental Psychology, 27,* 172–180.

Dodge, K. A. (1985). Attributional bias in aggressive children. In P. C. Kendall (Ed.), *Advances in cognitive-behavioral research and therapy* (Vol. 4, pp. 75–110). Hillsdale, NJ: Erlbaum.

Dodge, K. A. (1989). Enhancing social relationships. In E. J. Mash & R. J. Barkley (Eds.), *Behavioral treatment of childhood disorders* (pp. 222–244). New York: Guilford Press.

Dodge, K. A., Bates, J. E., & Pettit, G. S. (1990). *Mechanisms in the cycle of violence. Science, 250,* 1678–1683.

Dodge, K. A., & Frame, C. L. (1982). Social cognitive biases and deficits in aggressive boys. *Child Development, 53,* 620–635.

Dodge, K. A., Pettit, G. S., McClaskey, C. L., & Brown, M. (1986). Social competence in children. *Monographs of the Society for Research in Child Development, 51* (2, Serial No. 213).

Dumas, J. E. (1984). Child, adult-interactional, and socioeconomic setting events as predictors of parent training outcome. *Education and Treatment of Children, 7,* 351–364.

Dumas, J. E. (1986). Indirect influence of maternal social contacts on mother–child interactions: A setting event analysis. *Journal of Abnormal Child Psychology, 14,* 205–216.

Dumas, J. E. (1989). Treating antisocial behavior in children: Child and family approaches. *Clinical Psychology Review, 9,* 197–222.

Dumas, J. E. (1992). Conduct disorder. In S. M. Turner, K. S. Calhoun, & H. E. Adams (eds.), *Handbook of clinical behavior therapy* (2nd ed., pp. 285–316). New York: Wiley.

Dumas, J. E., & Wahler, R. G. (1983). Predictors of treatment outcome in parent training: Mother insularity and socioeconomic disadvantage. *Behavioral Assessment, 5,* 301–313.

Dush, D. M., Hirt, M. L., & Schroeder, H. E. (1989). Self-statement modification in the treatment of child behavior disorders: A meta-analysis. *Psychological Bulletin, 106,* 97–106.

Elliott, D. S., Huizinga, D., & Ageton, S. (1985). *Explaining delinquency and drug use*. Beverly Hills, CA: Sage.

Eyberg, S. M. (1992). Assessing therapy outcome with preschool children: Progress and problems. *Journal of Clinical Child Psychology, 221,* 306–311.

Eyberg, S. M., & Robinson, E. A. (1983). Conduct problem behavior: Standardization of a behavioral rating scale with adolescents. *Journal of Clinical Child Psychology, 12,* 347–354.

Eyberg, S. M., & Ross, A. W. (1978). Assessment of child behavior problems: The validation of a new inventory. *Journal of Clinical Child Psychology, 7,* 113–116.

Farrington, D. P., & West, D. J. (1981). The Cambridge study in delinquent development (United Kingdom). In S. A. Mednick & A. E. Baert (Eds). *Prospective longitudinal research: An empirical basis for the primary prevention of psychosocial disorders* (pp. 183–201). New York: Oxford University Press.

Fleischman, M. J. (1981). A replication of Patterson's "Intervention for boys with conduct problems." *Journal of Consulting and Clinical Psychology, 49,* 342–351.

Fraser, M. W., Pecora, P. J., & Haapala, D. A. (1991). *Families in crisis: The impact of intensive family preservation services*. Hawthorne, NY: Aldine de Gruyter.

Goyette, C. H., Conners, C. K., & Ulrich, R. F. (1978). Normative data on revised Conners Parent and Teacher Rating Scales. *Journal of Abnormal Child Psychology, 12,* 421–436.

Hawkins, J. D., Catalano, R. F., & Miller, J. Y. (1992). Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: Implications for substance abuse prevention. *Psychological Bulletin, 112,* 64–105.

Hawkins, J. D., Catalano, R. F., Morrison, D. M., O'Donnell, J., Abbott, R. D., & Day, L. E. (1992). The Seattle Social Development Project: Effects of the first four years on protective factors and problem behaviors. In J. McCord & R. Tremblay (Eds.), *The prevention of antisocial behavior in children* (pp. 139–161). New York: Guilford Press.

Hawkins, J. D., & Lishner, D. (1987). Etiology and prevention of antisocial behavior in children and adolescents. In D. H. Crowell, I. M. Evans, & C. R. O'Donnell (Eds.), *Childhood aggression and violence: Sources of influence, prevention, and control* (pp. 263–282). New York: Plenum.

Hawkins, J. D., & Weis, J. G. (1985). The social development model: An integrated approach to delinquency prevention. *Journal of Primary Prevention, 6,* 73–97.

Henggeler, S. W., & Borduin, C. M. (1990). *Family therapy and beyond: A multisystemic approach to treating the behavior problems of children and adolescents*. Pacific Grove, CA: Brooks/Cole.

Henggeler, S. W., Melton, G. K., & Smith, L. (1992). Multisystematic treatment of juvenile offenders: An effective alternative to incarceration. *Journal of Consulting and Clinical Psychology, 60,* 953–961.

Henggeler, S. W., Rodick, J. D., Borduin, C. M., Hanson, C. L., Watson, S. M., & Urey, J. R. (1986). Multisystemic treatment of juvenile offenders: Effects on adolescent behavior and family interaction. *Developmental Psychology, 22,* 132–141.

Hinshaw, S. P. (1991). Stimulant medication and the treatment of aggression in children with attentional deficits. *Journal of Clinical Child Psychology, 20,* 301–312.

Hinshaw, S. P., Heller, T., & McHale, J. P. (1992). Covert antisocial behavior in boys with attention-deficit hyperactivity disorder: External validation and effects of methylphenidate. *Journal of Consulting and Clinical Psychology, 60,* 274–281.

Hirschi, T. (1969). *Causes of delinquency*. Berkeley, CA: University of California Press.

Jameson, P. B., & Alexander, J. F. (1994). Implications of a developmental family systems model for clinical practice. In L. L'Abate (Ed.), *Handbook of developmental family psychology and psychopathology* (pp. 392–411). New York: Wiley

Kazdin, A. E. (1987) Treatment of antisocial behavior in children: Current status and future directions. *Psychological Bulletin, 102,* 187–203.

Kazdin, A. E. (1995). *Conduct disorders in childhood and adolescence*. Thousand Oaks, CA: Sage.

Kazdin, A. E., Bass, D., Siegel, T., & Thomas, C. (1989). Cognitive-behavioral therapy and relationship therapy in the treatment of children referred for antisocial behavior. *Journal of Consulting and Clinical Psychology, 57,* 522–536.

Kazdin, A. E., Esveldt-Dawson, K., French, N. H., & Unis, A. S. (1987). Problem-solving skills training and relationship therapy in the treatment of antisocial child behavior. *Journal of Consulting and Clinical Psychology, 55,* 76–85.

Lochman, J. E. (1987). Self and peer perceptions and attributional biases of aggressive and nonaggressive boys in dyadic interactions. *Journal of Consulting and Clinical Psychology, 55,* 404–410.

Lochman, J. E. (1992). Cognitive-behavioral intervention with aggressive boys: Three-year follow-up and preventive effects. *Journal of Consulting and Clinical Psychology, 60,* 426–432.

Lochman, J. E., Burch, P. R., Curry, J. F., & Lampron, L. B. (1984). Treatment and generalization effects of cognitive-behavioral and goal-setting interventions with aggressive boys. *Journal of Consulting and Clinical Psychology, 52,* 915–916.

Lochman, J. E., & Curry, J. F. (1986). Effects of social problem-solving training and self-instruction training with aggressive boys. *Journal of Clinical Child Psychology, 15,* 159–164.

Lochman, J. E., Lampron, L. B., & Rabiner, D. L. (1989). Format differences and salience effects in assessment of social problem-solving skills of aggressive and nonaggressive boys. *Journal of Abnormal Child Psychology, 18,* 230–236.

Loeber, R. (1988). The natural history of juvenile conduct problems, delinquency, and associated substance use: Evidence for developmental progressions. In B. B. Lahey & A. E. Kazdin (Eds.), *Advances in clinical child psychology* (Vol. 11, pp. 73–124). New York: Plenum.

Loeber, R. (1990). Development and risk factors of juvenile antisocial behavior and delinquency. *Clinical Psychology Review, 10,* 1–42.

Loeber, R., & LeBlanc, M. (1990). Toward a developmental criminology. In M. Tonry & N. Morris (Eds.), *Crime and justice* (Vol. 12, pp. 375–473). Chicago: University of Chicago Press.

Loeber, R., & Stouthamer-Loeber, M. (1986). Family factors as correlates and predictors of juvenile conduct problems and delinquency. In N. Morris & M. Tonry (Eds.), *Crime and justice: An annual review of research* (Vol. 7, pp. 29–149). Chicago: University of Chicago Press.

Mann, B. J., Borduin, C. M., Henggeler, S. W., & Blaske, D. M. (1990). An investigation of systemic conceptualizations of parent–child coalitions and symptom change. *Journal of Consulting and Clinical Psychology, 58,* 336–344.

McCord, J. (1988). Parental behavior in the cycle of aggression. *Psychiatry, 51,* 14–23.

McLoyd, V. C. (1990). The impact of economic hardship on Black families and children: Psychological distress, parenting, and socioemotional development. *Child Development, 61,* 311–346.

McMahon, R. J., Forehand, R., Griest, D. L., & Wells, K. C. (1981). Who drops out of therapy during parent behavioral training? *Behavioral Counseling Quarterly, 1,* 79–85.

Michelson, L. (1987). Cognitive-behavioral strategies in prevention and treatment of antisocial disorders in children/adolescents. In J. D. Burchard & S. N. Burchard (Eds.), *Prevention of delinquent behavior* (pp. 190–219). Beverly Hills, CA: Sage.

Michelson, L., Sugai, D., Wood, R., & Kazdin, A. E. (1983). Comparative outcome study of behavioral social skills training, cognitive problem-solving, and Rogerian treatments for child psychiatric outpatients: Process, outcome, and generalization effects. *Behavior Research and Therapy, 21,* 545–556.

Miller, G. E., & Prinz, R. J. (1990). Enhancement of social learning family interventions for childhood conduct disorder. *Psychological Bulletin, 108,* 291–307.

Minuchin, S. (1974). *Families and family therapy*. Cambridge, MA: Harvard University Press.

Minuchin, S., & Fishman, H. (1982). *Techniques of family therapy*. Cambridge, MA: Harvard University Press.

Minuchin, S., Montalvo, B., Guerney, B., Rosman, B., & Schumer, F. (1967). *Families of the slums*. New York: Basic Books.

Moffitt, T. E. (1990). Juvenile delinquency and attention deficit disorder: Boys' developmental trajectories from age 3 to age 15. *Child Development, 61,* 893–910.

Moffitt, T. E. (1993a). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100,* 674–701.

Moffitt, T. E. (1993b). The neuropsychology of conduct disorder. *Development and Psychopathology, 5,* 135–151.

Offord, D. R. (1982). Family backgrounds of male and female delinquents. In J. Gunn & D. P. Farrington (Eds.) *Delinquency and the criminal justice system* (pp. 120–131). New York: Wiley.

Offord, D. R., & Boyle, M. H. (1986). Problems in setting up and executing large-scale psychiatric epidemiological studies. *Psychiatric Developments, 3,* 257–272.

Offord, D. R., Boyle, M. C., & Racine, Y. A. (1991). The epidemiology of antisocial behavior in childhood and adolescence. In D. J. Pepler & K. H. Rubin (Eds.), *The development and treatment of childhood aggression* (pp. 31–54). Hillsdale, NJ: Erlbaum.

Offord, D. R. & Waters, B. G. (1983). Socialization and its failure. In M. D. Levine, W. B. Carey, A. C. Crocker, & R. T. Gross (Eds.), *Developmental-behavioral pediatrics* (pp. 244–259). Philadelphia: Saunders.

Ogbu, J. U. (1988). Cultural diversity and human development. In D. T. Slaughter (Ed.), *Black children and poverty: A developmental perspective* (pp. 11–28). San Francisco: Jossey Bass.

Parker, J. G., & Asher, S. R. (1987). Peer relations and later personal adjustment: Are low-accepted children at risk? *Psychological Bulletin, 102,* 357–389.

Patterson, G. R. (1974). Interventions for boys with conduct problems: Multiple settings, treatments, and criteria. *Journal of Consulting and Clinical Psychology, 42,* 471–481.

Patterson, G. R. (1982). *Coercive family process.* Eugene, OR: Castalia.

Patterson, G. R., Chamberlain, P., & Reid, J. B. (1982). A comparative evaluation of parent training procedures. *Behavior Therapy, 13,* 638–650.

Patterson, G. R., DeBaryshe, B. D., & Ramsey, E. (1989). A developmental perspective on antisocial behavior. *American Psychologist, 44,* 329–335.

Patterson, G. R., & Fleischman, M. J. (1979). Maintenance of treatment effects: Some considerations concerning family systems and follow-up data. *Behavior Therapy, 10,* 168–185.

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *A social interactional approach: IV. Antisocial boys.* Eugene, OR: Castalia.

Perry, D. G., Perry, L. C., & Rasmussen, P. (1986). Cognitive social learning mediators of aggression. *Child Development, 57,* 700–711.

Prinz, R. J., Blechman, E. A., & Dumas, J. E. (1994). An evaluation of peer coping-skills training for childhood aggression. *Journal of Clinical Child Psychology, 23,* 193–203.

Prinz, R. J., Connor, P. A., & Wilson C. C. (1981). Hyperactive and aggressive behaviors in childhood: Intertwined dimensions. *Journal of Abnormal Child Psychology, 9,* 191–202.

Prinz, R. J., & Miller, G. E. (1994). Family-based treatment for childhood antisocial behavior: Experimental influences on dropout and engagement. *Journal of Consulting and Clinical Psychology, 62,* 645–650.

Richters, J. E., & Martinez, P. E. (1993). Violent communities, family choices, and children's chances: An algorithm for improving the odds. *Development and Psychopathology, 5,* 609–627.

Robins, L. N. (1981). Epidemiological approaches to natural history research: Children's antisocial disorders. *Journal of the American Academy of Child Psychiatry, 20,* 566–580.

Rutter, M., & Giller, H. (1983). *Juvenile delinquency: Trends and perspectives.* New York: Penguin.

Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (1979). *Fifteen thousand hours: Secondary schools and their effects on children.* Cambridge, MA: Harvard University Press.

Rutter, M., Yule, B., Quinton, D., Rowlands, O., Yule, W., & Berger, M. (1974). Attainment and adjustment in two geographical areas: III—Some factors accounting for area differences. *British Journal of Psychiatry, 125,* 520–533.

Sampson, R. J., & Groves, W. B. (1989). Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology, 94,* 774–802.

Sanders, M. R., Dadds, M. R., & Bor, W. (1989). A contextual analysis of oppositional child behavior and maternal aversive behavior in families of conduct disordered children. *Journal of Clinical Child Psychology, 18,* 72–83.

Snyder, J., Dishion, T. J., & Patterson, G. R. (1986). Determinants and consequences of associating with deviant peers. *Journal of Early Adolescence, 6,* 29–43.

Stewart, J. T., Myers, W. C., Burket, R. C., & Lyles, W. B. (1990). A review of the psychopharmacology of aggression in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 269–277.

Tharp, R. G. (1991). Cultural diversity and treatment of children. *Journal of Consulting and Clinical Psychology, 59,* 799–812.

Tolan, P. H., & Mitchell, M. E. (1989). Families and the therapy of antisocial and delinquent behavior. *Journal of Psychotherapy and the Family, 6,* 29–48.

Viemero, V. (1992). Changes in patterns of aggression among Finnish girls over a decade. In F. Bjorkqvist & P. Niemela (Eds.), *Of mice and women: Aspects of female aggression* (pp. 99–106). San Diego, CA: Academic Press.

Wadsworth, M. E. J. (1980). Early life events and later behavioral outcomes in a British longitudinal study. In S. B. Sells, K. Crandell, M. Roff, J. S. Strauss, & W. Pollin (eds.), *Human functioning in longitudinal perspective* (pp. 168–180). Baltimore: Williams & Wilkins.

Wahler, R. G., & Dumas, J. E. (1987). Stimulus class determinants of mother–child coercive interchanges in multidistressed families: Assessment and intervention. In J. D. Burchard & S. N. Burchard (Eds.), *Prevention of delinquent behavior* (pp. 190–219). Beverly Hills, CA: Sage.

Wahler, R. G., & Hann, D. M. (1987). An interbehavioral approach to clinical child psychology: Toward an understanding of troubled families. In D. H. Ruben & D. J. Delprato (Eds.), *New ideas in therapy: Introduction to an interdisciplinary approach (pp. 53–78).* New York: Greenwood Press.

Walker, H. M. & Buckley, N. K. (1973). Teacher attention to appropriate and inappropriate classroom behavior: An individual case study. *Focus on Exceptional Children, 5,* 5–11.

Webster-Stratton, C. (1984). Randomized trial of two parent-training programs for families with conduct-disordered children. *Journal of Consulting and Clinical Psychology, 52,* 666–678.

Webster-Stratton, C., Hollinsworth, T., & Kolpacoff, M. (1989). The long-term effectiveness and clinical significance of three cost-effective training programs for families with conduct-problem children. *Journal of Consulting and Clinical Psychology, 57,* 550–553.

Webster-Stratton, C., Kolpacoff, M., & Hollinsworth, T. (1988). Self-administered videotape therapy for families with conduct problem children: Comparison with two cost-effective treatments and a control group. *Journal of Consulting and Clinical Psychology, 56,* 558–566.

West, D. J., & Farrington, D. P. (1977). The delinquent way of life. London: Heinemann.

West, M. O., & Prinz, R. J. (1987). Parent alcoholism and childhood psychopathology. *Psychological Bulletin, 102,* 204–218.

World Health Organization (1992). *The ICD-10 classification of mental and behavioural disorders.* Geneva, Switzerland: Author.

Zavodnick, J. M. (1994). Pharmacotherapy. In G. P. Sholevar (Ed.), *Conduct disorders in children and adolescents* (pp. 269–298). Washington, DC: American Psychiatric Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.21
Substance Abuse Disorders

HOLLY B. WALDRON
The University of New Mexico, Albuquerque, NM, USA

5.21.1 INTRODUCTION 540

5.21.2 SUBSTANCE ABUSE DISORDERS 540

- *5.21.2.1 Phenomenology* 540
  - *5.21.2.1.1 Problems in defining abuse* 540
  - *5.21.2.1.2 Adult vs. adolescent use* 541
  - *5.21.2.1.3 Comorbidity* 542
- *5.21.2.2 Diagnostic Features* 542
- *5.21.2.3 Prevalence* 544
  - *5.21.2.3.1 Substance abuse disorders* 544
  - *5.21.2.3.2 Substance use survey data* 544
  - *5.21.2.3.3 Prevalence by age, gender, race, and ethnicity* 545
- *5.21.2.4 Etiology* 545
  - *5.21.2.4.1 Risk and protective factors* 545
  - *5.21.2.4.2 Theory development* 546
- *5.21.2.5 Developmental Course* 547

5.21.3 CONCEPTUALIZATION AND CLINICAL FORMULATION 547

- *5.21.3.1 Multimodal Assessment* 547
  - *5.21.3.1.1 Methods for assessing substance use* 548
  - *5.21.3.1.2 Diagnosis* 549
  - *5.21.3.1.3 Screening instruments* 549
  - *5.21.3.1.4 Comprehensive assessment instruments* 549
  - *5.21.3.1.5 Other substance-related questionnaires* 549
- *5.21.3.2 Treatment Formulation* 552

5.21.4 TREATMENT 552

- *5.21.4.1 Psychosocial Treatments* 552
  - *5.21.4.1.1 Traditional and 12-step models* 552
  - *5.21.4.1.2 Cognitive-behavioral models* 553
  - *5.21.4.1.3 Family therapy models* 553
- *5.21.4.2 Pharmacological Treatments* 553
  - *5.21.4.2.1 Pharmacotherapy for substance abuse* 553
  - *5.21.4.2.2 Pharmacotherapy for dual diagnosis* 553
- *5.21.4.3 Efficacy and Effectiveness of Psychosocial Treatments* 554
  - *5.21.4.3.1 Traditional interventions* 554
  - *5.21.4.3.2 Behavioral outpatient interventions* 555
  - *5.21.4.3.3 Family-based interventions* 555
  - *5.21.4.3.4 Treatment engagement and retention* 557

5.21.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 557

5.21.6 SUMMARY 559

5.21.7 REFERENCES 559

5.21.1 INTRODUCTION

Only a half century ago the use of alcohol and other drugs among adolescents was quite rare, with the number of youth reporting ever drinking or using drugs estimated at one in five and frequent drinking or drug use essentially nonexistent (McCarthy & Douglass, 1949; Prosser, 1954). Over the years drinking and drug use have escalated sharply, to proportions many consider epidemic (Johnston, O'Malley, & Bachman, 1993). Similar increases in problems associated with alcohol and other illicit drug use have also been observed, including premature school drop out, crime, injuries and fatalities due to accidents and violence, sexual promiscuity and risk for sexually transmitted diseases, and teen pregnancy.

Although substance use and related problems have received greater attention in mental health fields in recent years, researchers and clinicians continue to struggle with fundamental conceptual and empirical issues. Some of the most perplexing issues include the complex and varied factors influencing the development of the disorders, the heterogeneous presentation of substance abuse disorders, the similarities and differences between adolescent and adult abuse, and the comorbidity of substance abuse with other disorders among adolescents, complicating diagnosis, assessment, and treatment.

This chapter provides an overview of adolescent substance abuse disorders, highlighting the challenges facing researchers and clinicians. The chapter is divided into three main sections focusing on (i) the nature of substance abuse disorders, (ii) methods for assessing adolescent substance use, and (iii) intervention approaches employed and the evidence for treatment effectiveness. Recommendations for further study of adolescent substance abuse disorders are also presented.

5.21.2 SUBSTANCE ABUSE DISORDERS

5.21.2.1 Phenomenology

Considerable ambiguity exists regarding the definition of substance abuse. A major obstacle in defining the concept of abuse is the array of terms such as recreational, experimental, problematic, pathological, risky, harmful, occasional, episodic, regular, light, moderate, and heavy use, employed to characterize substance use (Kaminer, 1994). On one point, however, theorists generally agree: substance abuse is not a unitary concept, but a complex phenomenon with heterogenous patterning and multiple determinants.

A unique characteristic of substance abuse disorders, unlike other mental disorders, is that they vary depending on factors such as the immediate social environment or the availability of substances. Moreover, classification depends on the presence of a "willing host [the abuser] who is an active participant in generating these disorders" (Newcomb, 1995, p. 7). Yet, substance use is a covert behavior not easily seen by parents, teachers, or other health professionals whose suspicion of use can be low. Failure to detect signs of use, in the absence of valid, reliable, practical definitions for the disorders, contributes to substance abuse being a commonly missed or misdiagnosed problem and hampers clinical efforts to identify adolescents in need of assessment, preventive care, and treatment. In addition, these problems have probably slowed research effort on prevalence, history, genetics, and treatment outcome (Bukstein, 1995).

5.21.2.1.1 Problems in defining abuse

The peak reporting of substance use occurred at the end of the 1970s and early 1980s when as many as 90% of high school seniors reported some alcohol use and nearly 50% marijuana use. Moreover, the majority of substance-involved youth do not develop substance abuse problems and do not go on to abuse as adults, with most drug use dropping sharply after age 21 (Blane, 1976; Kandel & Logan, 1984). In light of high rates of involvement, together with evidence for the discontinuity of use from youth into adulthood, the reality of normative adolescent substance use cannot be ignored. Some researchers have even suggested that, to a degree, drug use has developmental, adaptational utility for adolescents (Labouvie, 1986; Shedler & Block, 1990). Adolescents may use alcohol and other drugs to assert independence from their parents, identify with peers, or oppose societal norms and values, all part of the normal exploration of identity issues.

At the other extreme, any level of substance use by adolescents could be considered pathological. Use, or even possession, of alcohol and other drugs by underage youth constitutes illegal behavior in the USA and is often accompanied by truancy, curfew violations, shoplifting, and a host of other problem behaviors (Donovan & Jessor, 1985). Substance use also increases exposure to risky situations such as driving while intoxicated, engaging in sexual behavior that increases risk for AIDS, and engaging in violent exchanges (Farrell, Danish, & Howard, 1992; FeCaces, Stinson, & Hartford, 1991). Further, substance use can interfere with crucial developmental tasks, such as prosocial identity formation, interpersonal and educational skill acquisition, and family and work responsibility assumption (Baumrind, 1985; Bentler, 1992).

Certainly, diagnosing the vast majority of teenagers with a substance abuse disorder is problematic and a definition by which to identify adolescents abusing substances and in need of treatment is critical. Baumrind (1985) argued that drug abuse should be defined empirically, using frequency and problem consequences as determinants. Similarly, Pandina (1986) offered an operational definition on the basis of intensity (quantity and/or frequency) of use, presence of criteria-based symptoms, and negative consequences. However, no clear cut-offs have been established for quantity and frequency. The substance-related symptoms exhibited by adolescents have rarely been the focus of investigation and, as a result, are still poorly understood (Blum, 1987; Filstead, Parrella, & Conlin, 1988; Stewart & Brown, 1995). And even when a pattern of negative consequences is present, whether they result from substance use or stem from other preexisting or cooccurring factors such as another psychiatric disorder is often unclear (Bukstein, 1995).

Some have argued that substance abuse reflects only one aspect of a syndrome of problem behavior (Jessor & Jessor, 1977) or of general deviance (McGee & Newcomb, 1992). From such a perspective, substance abuse is not conceptualized as a disorder, but as a cluster of norm-violating attitudes and activities such as unconventionality, criminal behavior, and low academic orientation (Donovan & Jessor, 1985; McGee & Newcomb, 1992). Thus, a definition of substance abuse may need to address the broader context of problems intrinsically related to drug and alcohol use.

The widely varied patterning of use among adolescents must also be taken into account in any definition of substance abuse. Types of drugs used should certainly be considered. Alcohol, marijuana, stimulants, cocaine, opioids, sedatives, hallucinogenics, inhalants, and even anabolic steroids (for athletic enhancement) are all potential agents of abuse with specific pharmacologic effects. In addition, the number of drugs used and the combinations of drugs, age of the user, quantity and frequency of use of each substance, negative consequences stemming from use, presence of related problems, and developmental progression of use may all be factors that bear on the distinction between nonproblem use and abuse.

Tolerance, for example, develops rapidly to the euphoric effects of stimulants. With the larger doses of stimulants necessary to reproduce the euphoria, the potential for negative side effects increases and toxic overdose can be fatal. By contrast, tolerance develops more slowly to alcohol and the toxicity of marijuana is generally low. Thus, a consumption of one or two beers or smoking marijuana with friends on the weekend might not be considered abuse for a 17-year-old high school senior, whereas snorting crystal methadrine (i.e., crack) at the same rate might be judged more serious. On the other hand, an 11-year-old drinking one or two drinks or smoking marijuana with any regularity would probably be considered for treatment of substance abuse, given the quantity/frequency of use at such an early age of onset and the potential for rapid developmental progression of drug use.

5.21.2.1.2 Adult vs. adolescent use

Although most conceptualizations of adolescent abuse have been based on adult theories, models, and findings, the appropriateness of this approach has long been questioned and evidence is accumulating that adolescent use differs from adult use in important ways. Some differences are probably due to the restricted age range of the adolescent period. Adolescents have typically been experiencing substance problems for a shorter length of time (Brown, Mott, & Myers, 1990) and have had less time to exhibit deterioration in role functioning (Blum, 1987). Compared to adults, adolescents have also encountered fewer and different serious consequences related to their use, including physical evidence of dependence (Filstead, 1982; Filstead et al., 1988; Vingilis & Smart, 1981).

The patterning of use also appears to differ for adults and adolescents. For example, White and Labouvie (1989) found that adolescents drink less frequently than adults and consume less alcohol overall, but consume larger amounts at one time. Given the smaller body size and weaker tolerance of many youth, teens who drink can experience dangerous physical effects and alcohol-related problems at a lower level of consumption than adults. Brown et al. (1990) reported that 70–98% of adolescents in treatment abuse multiple substances, whereas reports of adult polysubstance use have been considerably lower. Also, the types and patterns of withdrawal symptoms among adolescents appear unique (Brown, Vik, & Creamer, 1989; Martin, Kaczynski, Maisto, Bukstein, & Moss, in press; Stewart & Brown, 1995). Adolescents also have a higher likelihood of certain negative social consequences, given their position of dependence and lower status in family and social systems (Barnes, 1984; White & Labouvie, 1989).

However, researchers have also noted similarities between adolescents and adults with respect to the pharmacologic effects of drugs, reasons for use, and the settings in which use occurs (Barnes, 1981). Also, similarities have been found with respect to patterns of problem drinking, including psychological dependence, binge drinking, and social consequences of drinking (Andersson & Magnusson, 1988; Hughes, Power, & Francis, 1992).

The question of whether adolescent abuse is a distinct category from adult abuse may be intertwined with the issues of nonproblem use versus abuse. Some researchers have found greater continuity from the adolescent period into adulthood for particular typologies of early-onset problem drinking (Babor et al., 1992). However, the majority of adolescents do not fit the severe category. All along the continuum of quantity/frequency and negative consequences of use there are adolescents who still appear to "mature out" of problem use when they reach adulthood (Blum, 1987; Winters & Henly, 1988). As evidence accumulates, we may yet be able to identify different levels or forms of use among adolescents, such as nonproblem use, problem use with relatively good prognosis for spontaneous remission, and problem use that persists into adulthood.

5.21.2.1.3 Comorbidity

Substance abuse disorders in adolescents are widely held to have a high prevalence of comorbidity with other psychological disorders. The most common disorder coexisting with adolescent substance involvement is conduct disorder. DeMilio (1989) found that 42% of adolescents presenting for treatment of substance abuse also met criteria for conduct disorder; Bukstein, Glancy, and Kaminer (1992) found 62% of adolescents receiving inpatient treatment were dually diagnosed. The interactive effects of conduct disorder and substance use and abuse are not well understood, but evidence indicates that dual diagnosis increases the likelihood of continued behavioral disturbance and substance use. Although attention deficit hyperactivity disorder (ADHD) has been hypothesized as a comorbid condition, Alterman and Tarter (1986) concluded that the apparent relationship between substance abuse and ADHD was an artifact of the high coincidence of ADHD and conduct disorder. This finding has been confirmed by Barkley and his colleagues who found that adolescents with only ADHD used substances no more frequently than normals (Barkley, Fischer, Edelbrock, & Smallish, 1990).

Affective disorders, anxiety, and bulimia also commonly coexist with adolescent substance abuse. For example, DeMilio (1989) reported a dual diagnosis of major depressive disorder for 35% of a treatment sample; Bukstein et al. (1992) found 30% in an inpatient setting. Studies examining the coexistence of anxiety and other disorders with substance abuse have been more common with adults but, like depression, the pattern of the relationship is probably similar for adolescents. Possible mechanisms for the connection between mood and substance abuse disorders are the depressive pharmacological effects of some drugs and the negative social consequences, such as school failure and peer rejection, experienced by adolescents using alcohol and other drugs. Similarly, the actual or expected pharmacologic effects of substances in reducing anxiety is an obvious explanation for the comorbidity with anxiety disorders. The relationship between substance abuse and bulimia is less clear, and the high comorbidity between bulimia and mood disorders may account for some of the comorbidity between bulimia and substance abuse.

The comorbidity issue adds a level of complexity to understanding substance abuse. Research examining adolescent substance abuse disorders often specifically excludes adolescents with dual diagnoses in order to limit potential confounds. However, even when the coexistence of disorders is the question of interest, the lack of valid, reliable definitions of substance abuse and other disorders of childhood and adolescence has hampered investigations. Moreover, the concept of comorbidity is inherently fraught with problems such as identifying comorbid disorders using sets of symptoms which result from the same underlying cause. For example, Crowley and Riggs (1995) observed that adolescents dually diagnosed with substance abuse and conduct disorders epitomize the general deviance syndrome, in which substance use and other deviant behaviors are viewed as stemming from common causes. Whether substance abuse is primary or occurs secondary to another disorder, and how the interaction of coexisting disorders influence the onset, identification, course, and treatment of substance abuse, problems remain in question.

5.21.2.2 Diagnostic Features

Formal diagnostic classification systems for substance use disorders have primarily evolved since the mid-1970s and have been developed exclusively for adults, with no distinction made between adolescent and adult users. The first criteria-based classifications appeared in the third edition of the *Diagnostic and statistical manual of mental disorders* (*DSM-III*; American Psychiatric Association, 1980), which identified two categories of substance use disorders: substance abuse and substance dependence. Substance abuse was defined as pathological use associated with impairment of social or occupational functioning with a one-month duration, and dependence as the presence of tolerance or withdrawal and either pathological use or impairment in social or occupational functioning due to such use. Prior to the third edition, the *DSM* viewed substance use problems such as alcoholism and drug dependence as subsets of sociopathic personality disturbance. The designation of a separate category for substance use disorders in *DSM-III* differentiated substance use from antisocial behavior more generally and recognized the importance of research on the nature of the disorders (Nathan, 1991).

A number of serious limitations of *DSM-III* were noted, however, including the insensitivity of the criteria in addressing heterogeneity of problems, the lack of theoretical and empirical basis for some criteria, and inconsistencies across criteria in various substance categories (cf. Nathan, 1991; Rounsaville, 1987). In addition, the reliance on tolerance or withdrawal for the dependence diagnosis determined that, in most cases, adolescents could only be diagnosed as abusers, since they rarely exhibit many of the tolerance or withdrawal symptoms for alcohol or other drugs (Hughes et al., 1992; Vingilis & Smart, 1981).

The *DSM-III-R* (American Psychiatric Association, 1987) was designed to correspond more closely to the *International classification of disease, ninth edition* (*ICD-9*), which had effected a conceptual shift in classifying substance use disorders. In keeping with the conceptual shift, the definition of dependence was expanded to include clinically significant behaviors, cognitions, and symptoms indicative of use and no single symptom was required for a diagnosis of dependence. Marked tolerance and withdrawal symptoms remained in the list of criteria, but other criteria—such as spending large amounts of time on activities to acquire substances, symptoms of intoxication when expected to fulfill major role obligations at work, school, or home, and giving up important activities because of substance use—made it easier for adolescents to be diagnosed with dependence. The changes also allowed for more variability in patterns of use.

By default, abuse was diagnosed when sufficient criteria for dependence were not met. Thus, abuse was conceptualized as a milder or predependence stage in the cluster of disorders. However, some research suggests that abuse and dependence disorders have distinct courses (Hasin, Grant, & Endicott, 1990), questioning the notion of a presumed continuum of severity.

The revision of the system for *DSM-IV* (American Psychiatric Association, 1994) was intended to reflect empirical research more closely and to increase compatibility with the *ICD-10* system. Accordingly, three essential changes were made to the classification of dependence: the criteria regarding the inability to fulfill major role obligations and the criteria for duration were dropped, and subtypes of tolerance and withdrawal symptoms were added. However, adolescents with substance use problems commonly experience difficulty in meeting role obligations. With regard to educational expectations, truancy, attending classes while under the influence of alcohol or drugs, school failure, and premature dropout are common. Similarly, failures to meet role obligations are frequently observed in family contexts. Thus, the elimination of the role obligation criterion will probably reduce the number of adolescents who are diagnosed as dependent.

The criteria for substance abuse in *DSM-IV* were expanded, requiring the presence of clinically significant impairment or distress, including the failure to fulfill major role obligations and recurrent substance-related legal problems, as part of the maladaptive pattern of substance use. The addition of the role obligation criteria to this category could facilitate diagnostic classification. Nevertheless, adolescents who show a maladaptive pattern of use must now also exhibit at least minimal impairment or distress. This moves classification of abuse away from a topographical definition of substance use, such as quantity, frequency, or patterning of use, toward a definition based on dysfunction associated with use, and essentially raises the bar for adolescents to meet criteria, since adolescent behaviors tend to be more detectable than distress resulting from the behaviors.

The most disturbing aspect of the evolution of classification systems for substance use is the complete lack of evidence either to support or reject the applicability of the systems with adolescents. To illustrate, consider the nature of the prevailing diagnostic systems. The *DSM* and *ICD* systems are categorical rather than dimensional, classifying substance abuse on the basis of clustering of symptoms as opposed to viewing symptoms as occurring along a continuum. Within these systems, then, it is possible for adolescents to exhibit high quantity and frequency of use, yet not meet criteria for a substance abuse disorder if they manage to avoid negative consequences or other signs of impairment considered part of the discrete cluster of symptoms that currently define abuse. The confirmation of the reliability and validity of the current approach for adolescents or, alternatively, the formulation of an age-specific substance abuse definition and diagnostic criteria appropriate for adolescents could play a key role in the development and evaluation of prevention and treatment programs.

5.21.2.3 Prevalence

5.21.2.3.1 Substance abuse disorders

The epidemiology of substance abuse disorders in adolescents has rarely been investigated, with the vast majority of research focused on prevalence of substance use more generally. Moreover, comparison across studies has been impeded by factors such as nonuniform definitions of abuse, the particular diagnostic system employed, and the type and age of sample selected. Using the *DSM-III* system, for example, Kashani et al. (1987) found a prevalence of 5.3% for alcohol abuse or dependence disorders and a 3.3% for drug abuse or dependence disorders in a community sample of 15-year-olds. Deykin, Levy, and Wells (1986), using the same system, found a prevalence rate of 8.2% for alcohol disorders and 9.4% for drug disorders in a young college population ranging from 16 to 19 years of age. Also focusing on older youth, aged 17–19 years, but using the *DSM-IIIR,* Reinherz, Giaconia, Lefkowitz, Pakiz, and Frost (1993) found a prevalence of 34% for alcohol abuse disorders and 9.8% for drug abuse disorders. Although the rates of alcohol disorders seem particularly high in this study, Cohen et al. (1993) examined two age groups, 14- to 16-year olds, and 17- to 20-year olds, with reported rates of 4.1% and 3.1% for younger boys and girls, respectively, and 20.3% and 8.9% for older boys and girls, respectively. The estimates vary widely and more uniformity in methodology is needed for developing a foundation of empirical work before any conclusions about epidemiology of substance abuse disorders in adolescents can be drawn.

5.21.2.3.2 Substance use survey data

Compared to substance abuse disorders, prevalence of substance use among adolescents has been much more intensively studied. Data are regularly obtained through two national surveys, the National Household Survey on Drug Abuse, a federal project studying illegal drug use in the USA by adults and adolescents since 1971, and the University of Michigan Monitoring the Future project, which focuses on high school seniors. These surveys have revealed a general pattern of increasing substance use among adolescents through the late 1970s and early 1980s, followed by marked declines through 1991. The senior survey data (Johnston et al., 1993), for example, shows that peak lifetime use of alcohol (i.e., having ever used) among high school seniors was 93.2% in 1980, but use in the past month declined from 72.1% in 1978 to 54.0% in 1991. Binge drinking (i.e., having an episode of five or more drinks on one occasion in the past two weeks) among seniors peaked in 1981 at 41.4%, but had dropped to 29.8% by 1991, and daily drinking peaked in 1979 at 6.9%, but had dropped to 3.6% by 1991. Similarly, lifetime use of marijuana peaked in 1980 at 60.3%, with highest daily use in 1978 at 10.7% and lowest use reported for 1991 at 2.0%. Stimulant and cocaine use peaked later, with peak past year stimulant use at 20.3% in 1982 and cocaine use at 13.1% in 1985. The percentage of past month stimulant use dropped from 10.7 in 1982 to 3.7 in 1990, and cocaine use dropped from 6.7 in 1985 to 1.9 in 1990. Other drugs use patterns have shown similar patterns of peak and decline, although it is interesting to note that lifetime prevalence of alcohol use has not shown the same sharp declines as for other drugs.

Recent survey data suggest that drug use may be on the rise again. According to the National Household Survey (Substance Abuse and Mental Health Services Administration [SAMHSA], 1996), the rate of past month illicit drug use among adolescents aged 12–17 years increased from 8.2% to 10.9% between 1994 and 1995, a rate which has doubled since 1992. Past month use of alcohol in 12- to 17-year-olds had fallen from 25.2% in 1988 to 15.7% in 1992, and was up to 21.6% and 21.1% in 1994 and 1995, respectively. Significant increases were also reported for past month marijuana use, from 6% in 1994 to 8.2% in 1995, continuing an upward trend since 1992. Similarly, between 1994 and 1995, cocaine use increased from 0.3% to 0.8% and hallucinogen use increased from 1.1% to 1.7%.

Tobacco use appears to be on the rise as well. The Youth Risk Behavior Survey reported current cigarette use among students in grades 9–12 increasing from 27.5% in 1991 to 34.8% in 1995 (Centers for Disease Control and Prevention, 1996). These findings are consistent with the Monitoring the Future project, which reported rates of smoking by 10th and 12th graders increasing, respectively, from 20.8% and 28.3% in 1991 to 27.9% and 33.5% in 1995. These rates are alarming when considering that roughly 11% of students reportedly smoke a half-pack per day or more (Johnston et al., 1993) and this percentage may be rising as well.

The rate of first use of marijuana, cocaine and crack, inhalants, and other substances, including adolescent alcohol use, has also risen sharply in the same period. First use has shown utility in predicting prevalence rates several years later. Increases in this index now could indicate continuing rise in prevalence for several years.

5.21.2.3.3 Prevalence by age, gender, race, and ethnicity

In general, prevalence rates for the use of alcohol and other substances increase with age. In 1993, for example, the Monitoring the Future project found that past month drinking to the point of intoxication was 28.9% for 12th graders, 19.8% for 10th graders, and 7.8% for 8th graders. Similarly, use of marijuana was 10.9% among 10th graders, but only 5.1% among eighth graders. Inhalant use is one exception, with past month use reported at 3.3% for 10th graders and 5.4% for eighth graders, based on 1994 data. This exception may result from the decreased availability of drugs in junior high and middle schools, compared to high schools, leaving glue or paint cans more readily available for getting high. Despite lower rates overall, trajectories of use for the younger age groups suggests there is reason for concern. That is, during the same period of time the senior survey revealed declines in cocaine, alcohol, and other drug use among seniors, use of these same substances increased among eighth graders.

Comparisons for gender have consistently shown that males tend to use alcohol and drugs in higher proportions than females, especially at higher levels of use (Johnston et al., 1993). The exceptions are sedatives, stimulants, and smoking tobacco, where rates are similar for males and females (Johnston et al., 1993; SAMHSA, 1996). Also, the fall and rise in prevalence rates over time appear to have been the same for male and female adolescents (SAMHSA, 1996).

According to Bachman et al., (1991), Native American adolescents had the highest prevalence of use for cigarettes, alcohol, and most other drugs; Anglo/European Americans (i.e., non-hispanic whites) had higher rates of use than Hispanic Americans, African-Americans, or Asian-Americans. However, National Household Survey data for 1994–1995 revealed no ethnicity differences in smoking (SAMHSA, 1996). With respect to changes in rates over time, during the period of overall declining use (i.e., through 1992) rates of use among non-hispanic whites showed dramatic declines, whereas rates for black and for Hispanic youth remained relatively steady. Prevalence rates across culture and race should continue to receive careful scrutiny. If these patterns hold, research will be needed to understand the mechanisms by which substance use became entrenched in these groups and to examine the necessity of developing treatment programs individually tailored to the needs of particular groups.

5.21.2.4 Etiology

5.21.2.4.1 Risk and protective factors

Although a broad spectrum of factors has been associated with the initiation and maintenance of substance use, no single etiological theory or model adequately accounts for the marked heterogeneity of patterns of substance abuse disorders. A number of researchers have adopted a risk factor perspective as a way to understand the multiple determinants and antecedents of drug use and abuse, including biological, cultural and societal, interpersonal, and intrapersonal factors (Hawkins, Catalano, & Miller, 1992). Although biological influences such as genetic, physiological, and biochemical factors have been shown to affect susceptibility to drug use, these factors have not accounted for large proportions of the variance. Because substance abuse is a complex phenomenon, any single-factor biological cause or theory is unlikely (Bukstein, 1995). Within the broad cultural and social context, substance use can be influenced by factors such as poverty, the existence of laws and social norms favorable toward substance use, and the general availability of alcohol and drugs. Intrapersonal influences include a wide range of individual cognitive and behavioral factors (e.g., positive attitudes toward drug use), aggression and other disruptive behavior problems, and low commitment to education. General findings on interpersonal factors have included familial influences such as: parent and sibling use, parent attitudes toward use, parenting practices, poor communication and parent–adolescent conflict, and parent–child bonding. In addition, poor peer relationships or having friends who use drugs increase the risk for use.

Researchers have also identified a variety of protective factors across the cultural/social, interpersonal, and intrapersonal domains that influence the likelihood of substance use initiation or reduce level of drug use. These factors may moderate the relationship between risk factors and use/abuse, or may have a direct effect on reducing drug involvement (Newcomb, 1995). Sanctions against drug use, family support, self-acceptance, and religiosity are but a few of the protective factors that have been identified (cf Hawkins et al., 1992).

The factors influencing substance abuse are likely to be interdependent and bidirectional. For example, Dishion, Patterson, and Reid (1988) found that parental drug use had both a direct effect, believed to result from modeling and opportunities for use, and an indirect effect, resulting from impaired parental control when parents were under the influence of drugs or alcohol. Similarly, findings of increased stress, increased conflict, highly charged negative affect, lack of openness, and poor cohesion in families, as well as the tendency of adolescents to seek support in relationships outside the family, may all have reciprocal influence. One model suggests that adolescents may not directly imitate their parents' use of specific drugs, but they may be exposed to observing that drug use is one method of coping with psychological stress (Kandel, Kessler, & Margulies, 1978). Effective problem solving may, in turn, attenuate drug use by providing an alternative for coping with stressful life events and family disturbance. Examining another aspect of influence, Duncan, Duncan, and Hops (1994) found that the largest effects of family cohesion on adolescent substance use occur in middle and later adolescence. Taking their results together with other research (Dishion et al., 1988; Patterson, Reid, & Dishion, 1992), they suggested that family influence may vary in a curvilinear fashion over time, with a low point in early adolescence when peer influences peak. Such research begins to tease apart the complex interactions among risk and protective factors associated with substance abuse and holds promise for much needed model development for adolescent substance abuse.

Although the mechanism by which risk and protective factors exert causal influence is poorly understood, Bry, McKeon, and Pandina (1982) found evidence that substance abuse derives from the sheer number of risk factors rather than the presence of any one factor or particular set of factors. These findings have been corroborated by Newcomb and his associates who also found that exposure to more risk factors was not only correlated with use, but predicted increasing drug use over time, providing support for a multiple etiological pathway model of drug use (McGee & Newcomb, 1992; Newcomb, Maddahian, & Bentler, 1986). The risk and protective factor model is a useful one for predicting youth who may develop substance abuse problems and for identifying specific targets for intervention. A limitation of the risk and protective factor model is that, in the absence of longitudinal research, we cannot distinguish between etiological factors, through which we may discover avenues of control, and factors which are merely antecedents or correlates of abuse and have common causes.

5.21.2.4.2 Theory development

Within the substance abuse literature, conceptual models are abundant (Lettieri, Sayers, & Pearson, 1980). In addition to the risk or vulnerability model, models proposed to explain the development of substance abuse disorders range from biological to spiritual and from intrapersonal to sociocultural. Two quite narrow models are the educational model, which views abuse as resulting from a lack of accurate information, and the disease model, which views abuse as irreversible and requires total abstinence to manage the disease. Although these perspectives may not be incompatible with other models, they fail to account for much of what is currently known about the correlates of abuse.

More comprehensive theories can be organized within the broader framework of intrapersonal, cognitive, learning, conventionality/social attachment, and interactional models (Petraitis, Flay, & Miller, 1995). Intrapersonal perspectives place more emphasis on characteristics of adolescents themselves, including personality traits, affective states, and/or behavioral skills, in the etiology of substance use. Although the self-medication hypothesis has been articulated in cognitive-behavioral frameworks as well, the role of substance use as a coping mechanism is an example of an intrapersonal (psychodynamic) perspective. Low self-esteem and self-derogation are also central intrapersonal constructs in theories of etiology.

Cognitive theories focus on the role of adolescents' perceptions of use in decision making related to the initiation and continuation of substance abuse. Attitudes, normative beliefs, and substance use expectancies, as well as self-efficacy, influence the evaluation of the benefits of use (e.g., peer approval) and the costs of use (e.g., health dangers) in decision making.

Operant learning perspectives view substance use in the context of the antecedents and consequences surrounding the behavior. In the presence of stress, for example, substance use may result in tension reduction, regulation of negative affect, or social enhancement, any of which may serve to reinforce use and thereby increase the likelihood that substance use will recur in the future. The social learning model is expanded to allow for cognitively mediated learning through observation and imitation of models (e.g., parents, siblings, or peers) who use substances and it includes self-efficacy beliefs and coping as key concepts.

The conventionality/social attachment approaches stem from sociological theories of control which argue that strong commitment to society, families, schools, and religions inhibits the expression of deviant impulses that all individuals share. Substance use, then, is but one manifestation of a broader cluster of unconventional or problem behaviors which develops in the context of weak conventional bonds, and are maintained through social learning and contingencies in the environment.

Family models view problems such as substance abuse as maladaptive behaviors expressed by one or more family members, but reflecting dysfunction in the system as a whole. The behavior may be viewed as serving an important function in the family, allowing the family to cope with internal or external stressors or maintain other processes that have become established in the organization of the system.

Integrative theories are increasingly complex, but account for more of the findings in the research literature (Petraitis et al., 1995). For example, family interaction theory (Brook, Brook, Gordon, Whiteman, & Cohen, 1990) incorporates family factors, social learning, and intrapersonal characteristics known to affect substance use. Problem-behavior theory (Jessor & Jessor, 1977) emphasizes unconventionality, but also incorporates cognitive, intrapersonal, social learning concepts and accounts for substance abuse and a variety of other deviant behaviors such as juvenile delinquency, poor school performance, and precocious sexual activity.

Flay and Petraitis have attempted to organize the array of theories focusing on substance use etiology into one comprehensive macromodel (Flay & Petraitis, 1994; Petraitis et al., 1995). Their model of triadic influence assumes that health-related behaviors such as substance use are most immediately controlled by decisions or intentions that are a function of three streams of influence: (i) cultural–environmental factors which influence attitudes, (ii) social situation–contextual factors which influence social learning and normative beliefs, and (iii) intrapersonal factors which influence self-efficacy. These streams have different origins and flow through causal factors which are organized into levels ranging from very proximal to distal to ultimate. The stream of influence, independently and in interaction with one another, are recognized as having both direct and indirect effects on behavior (Flay & Petraitis, 1994). This work highlights the complementarity of existing theories and may provide a useful structure for identifying gaps in our knowledge base with respect to etiology.

5.21.2.5 Developmental Course

One avenue which may hold promise for differentiating nonproblem use from more serious substance involvement is research examining the developmental trajectory associated with initiation and progression of use. Before substance use becomes problematic, adolescents pass through a stage during which they are not committed to continuing use and have not integrated substance use as a regular part of life (Clayton, 1992). Typically, adolescents who become abusers follow a relatively invariant sequence for drug use, beginning with alcohol and/or cigarette use, followed by marijuana, and then other hard drugs. This progression from "soft" to "hard" drug use, the "gateway" phenomenon (Kandel, 1975), has been confirmed in a number of studies (Andrews, Hops, Ary, Lichtenstein, & Tildesley, 1991; Kandel, Yamaguchi, & Chen, 1992; Welte & Barnes, 1985). Moreover, Ellickson, Hays, and Bell (1992) found that weekly alcohol use followed marijuana use and preceded use of all other illicit drugs, and that weekly smoking formed a distinct stage between initial use of pills and other hard drugs.

According to Kandel et al. (1978), different factors may predominate each stage of use or predicting the progression to other stages. Their findings suggested that parent modeling may have greater influence on the use of alcohol and tobacco, whereas peers may have greater influence on marijuana use. The use of other illicit drugs was more related to family relationship disturbance and general psychology dysfunction. Thus, although dysfunctional use of drugs can be generally characterized by increasing quantities and varieties of drugs used, understanding substance abuse disorders may involve not only identifying the types and number of substances adolescents have used and the quantity and frequency of use, but also the sequencing or patterning of use, as well as the factors predominating each stage. As Dishion et al. (1988) noted, different models may be required to explain the variations in patterning and progression of use.

5.21.3 CONCEPTUALIZATION AND CLINICAL FORMULATION

5.21.3.1 Multimodal Assessment

Assessment approaches range from narrow inquiry into substance use behavior to broad, multifaceted approaches which address use patterns, problem consequences, other aspects of use (e.g., drug use expectancies, normative beliefs, or readiness to change), and functioning across multiple domains (e.g., school, family, or leisure time). The purpose of the assessment, or the question the assessment is intended to answer, should determine the type of information obtained, how the information is obtained, and who provides the information. A self-report screening instrument summarizing substance use quantity and frequency may be sufficient to determine need for further evaluation among youth at risk for substance abuse problems. However, more comprehensive approaches are often used in establishing a diagnosis, evaluating treatment needs (i.e., inpatient or detoxification), developing a treatment formulation, or evaluating treatment effectiveness.

The main source of information is usually the adolescent, although collateral information is often obtained from parents or others in the social environment (i.e., teachers, siblings, or peers). Methods for assessing substance abuse include interviewing, questionnaires, checklists, rating scales, and physiological measures (e.g., urine toxicology screens). Observational measures are sometimes used to examine particular aspects of adolescent functioning, such as family and peer interactions (Hops, Tildesley, Lichtenstein, Ary, & Sherman; 1990), and self-monitoring has been used with adults (Sobell & Sobell, 1995).

In general, assessment of adolescents requires some recognition of the cognitive developmental level of the adolescent, attention to issues of engagement in the process and the need for establishing rapport, and careful consideration of issues of confidentiality. Although the validity of self-report of use has been questioned, adult studies have shown that direct self-report measures have high sensitivity in detecting substance use problems, and compare favorably to biomedical measures such as blood and urine tests (Sobell & Sobell, 1995). Similarly, some support has been found for adolescents (Bailey, Flewelling, & Rachal, 1992; Needle, McCubbin, Lorence, & Hochhauser, 1983), but more systematic examination is needed to address this issue in adolescent populations (Winters & Stinchfield, 1995).

5.21.3.1.1 Methods for assessing substance use

Almost all measures of substance use behavior focus on some type of direct self-report of frequency (e.g., how many days per month use occurred) and intensity of use (e.g., number of drinks consumed). The majority of adolescents referred for treatment use multiple substances, necessitating inquiry about the quantity/frequency of use of each specific drug type. Two general approaches to assessing frequency and quantity of substance use have been described: quantity/frequency self-report questionnaires and calendar-based interviewing methods. Many self-report instruments measuring use of alcohol and other substances provide summaries of use behavior or are scaled in a way that precludes detailed assessment of use (e.g., 0 times per week, 1–2 times per week, 3 times or more per week). As a result, such measures fail to capture the patterning of use with any degree of precision that would allow for the calculation of peak blood alcohol concentration, a potentially important indicator of dangerous or even life-threatening alcohol use. Summary measures have also been shown to produce significantly lower estimates of use (Sobell & Sobell, 1995) and may also fail to detect reductions in use among very heavy users who, for example, could decrease their use from daily to three times per week.

An alternative method is the time line follow back procedure (Sobell & Sobell, 1995), a structured interview technique that samples a specific time period, using a monthly calendar and memory anchor points to reconstruct daily consumption during the period of interest. A number of studies have compared the methods, including quantity/frequency measures and time line follow back, finding generally similar estimates. In principle, the time line follow back may offer the most sensitive assessment for adolescents, having the advantage of assessing the widely variable drinking patterns that often characterize teen drinking and that might not be modeled adequately by the averaging approaches (Leccese & Waldron, 1994). In a recent study, initial evidence for the internal reliability and convergent validity of the time line method for adolescent drinkers was found (Waldron, 1996). Moreover, a group of adolescents who were classified as nonproblem drinkers on the basis of the *Adolescent Drinking Index* (Harrell & Wirtz, 1989), a standardized alcohol screening instrument, were identified with substantial amounts of binge drinking. Of 67 nonabstaining youth who scored below the cut-off on the standardized instrument, half were reporting peak blood alcohol concentrations between 0.10 and 0.48 ($M=0.11$), levels which suggest these youth may have been good candidates for harm reduction interventions. These findings support the use of time line follow back with adolescents and point to one direction for future research.

Urine toxicology screens and other biomedical markers can be valuable for corroborating self-report (O'Farrell & Maisto, 1987). When multiple measures, such as self-report, urinalyses, and collateral reports attained by interviewing parents or peers, converge to give the same results, there is greater confidence in the validity of measures. Another useful aspect of biomedical tests is the potential for a "bogus pipeline" effect, creating the impression that there is an accurate check on self-report and thereby increasing honesty in descriptions of one's substance use (Miller, Westerberg, & Waldron, 1995).

5.21.3.1.2 Diagnosis

Several interview instruments have been designed specifically to determine whether an individual meets currently established criteria for disorders of substance abuse and substance dependence according to a specific taxonomic system. The *Adolescent Diagnostic Interview* (Winters & Henly, 1993) is one such system which covers *DSM-III-R* symptoms of psychoactive substance use disorders and thereby assumes the applicability of adult syndromes to adolescents. This structured interview explores the adolescent's drug use history and signs of abuse or dependence for several drug categories, and also covers levels of functioning and psychosocial stressors. Diagnosis, as a binary (present vs. absent) process, represents a gross indicator of problem use with limited utility when used in isolation or as an outcome measure in treatment research. As a treatment issue, however, adolescents whose use is prolonged and heavy without periods of abstinence must be assessed for dependence, since physiological withdrawal may require specific medical or psychological interventions (Kaminer, 1994).

5.21.3.1.3 Screening instruments

Screening instruments are designed to identify the potential presence of a particular problem. As such, they are typically used as a preliminary step in assessment, as a way of determining if further, more comprehensive assessment is necessary. As noted in the discussion of time line follow back, one limitation of screening instruments is their reliance on cut-off scores for decision making. With lower cut-off points, instruments can inappropriately identify adolescents as needing more extensive assessment or intervention, whereas higher cut-off can fail to identify adolescents who are in need of additional services. Moreover, the reduction of a continuously distributed variable to a binary classification probably sacrifices predictive power. Conversely, screening instruments are characteristically easy and inexpensive to administer and are especially useful in settings where large numbers of adolescents must be screened in a short period of time, such as within the juvenile justice system, adolescent medicine units, or general mental health agencies serving adolescents with varied presenting problems. Table 1 presents screening instruments and basic information about each.

5.21.3.1.4 Comprehensive assessment instruments

In addition to assessing substance use behavior, comprehensive assessments also take into account the consequences of substance use, the contexts within which use occurs, the functions which the use serves in the adolescent's life, and the co-occurrence of other problem behaviors. As noted above, many other classes of variables have been found to be related in complex ways to level of substance use, including age, comorbid disorders, peer factors, personality factors, biological and genetic factors, family functioning, and level of social skills. The existence of these relationships suggests that a comprehensive assessment of adolescent substance use could be enriched by information in these areas as well. More comprehensive assessment procedures usually allow for the examination of adolescents' scores along a continuum or along several continua. These instruments are designed to be given after a screening instrument or other process indicates the need for further assessment of alcohol and drug use. Table 2 displays the main characteristics of six comprehensive adolescent assessment instruments.

5.21.3.1.5 Other substance-related questionnaires

To the extent that substance abuse is multiply determined and is associated with diverse problem behaviors presumably generated by common causes, broad, multidimensional assessment is vital. Newcomb and Bentler (1989) maintained that any focus on substance use, in isolation from the associated antecedents, concurrent events, or consequences constituting the more general problem-behavior syndrome of which substance abuse is a part, will be distorted. A number of other instruments have been used to assess various aspects of alcohol and drug use in adolescents. Among them are measures of alcohol, use attitudes toward alcohol, unconventionality of behavior, and expectancies about alcohol and other drugs (cf Leccese & Waldron; 1994; Winters & Stinchfield, 1995). These instruments can be combined with other measures to assess specific concerns the clinician may have in the formulation of the treatment plan or may be used for research purposes.

Table 1 Characteristics of screening instruments.

Instrument	*Administration method*	*Items*	*Time to administer*	*Psychometrics*	*Information provided*
Drug and Alcohol Problem Quick Screen (DAP; Schwartz & Wirtz, 1990)	Self-report	30	10	Limited validity data	Cut-off for substance use risk; indicates the need for further assessment
Drug Use Screening Inventory (DUSI; Tarter & Hegedus, 1991)	Self-report or computer administered	149	20	Adequate content validity	Functioning measured across 10 domains; absolute problem density, overall problem index, and relative problem density scores
Perceived Benefit of Drinking Scale (Petchers & Singer, 1987)	Self-report	5	1	Adequate internal consistency and divergent validity	Scale scores reflect frequency, intoxication, and problems related to substance use
Personal Experience Screening Questionnaire (PESQ; Winters, 1992)	Self-report	40	10	High internal consistency, adequate discriminant and predictive validity	Indicates the need for further assessment
Problem Oriented Screening Instrument (POSIT; Rahdert, 1991)	Self-report	139	30–45	None reported	Functioning measured across 10 domains, "red-flag" items signal the need for further assessment
Teen Health Advisor (Paperny, Aono, Lehman, Hammer, & Riusser, 1990)	Computer administered	32–62	10	None reported	High-risk behavior evaluation and advice or referral information across several domains
Substance Abuse Subtle Screening Inventory-Adolescent (SASSI-A; Miller, 1990)	Self-report	81	20	Limited validity data, no reported reliability data	Face valid alcohol, face valid other drug, obvious attributes, subtle attributes, defensiveness (two scales), and two experimental scales

Table 2 Characteristics of comprehensive assessment instruments.

Instrument	*Administration method*	*Items*	*Time to administer*	*Psychometrics*	*Information provided*
Adolescent Diagnostic Interview (ADI; Winters & Henley, 1993)	Structured or computerized interview	varies	45–60	Adequate interrater, test–retest reliability, and criterion-related validity	*DSM-III-R* substance use disorders, functioning across several domains, psychosocial stressors, and screens for memory and orientation
Adolescent Drug Abuse Diagnosis (ADAD; Friedman & Utada, 1989)	Structured interview	150	45–60	High interrater, test–retest reliability, adequate concurrent, convergent, and discriminant validity	Problem severity ratings and need for treatment across nine domains
Adolescent Problem Severity Index (APSI; Metzger, Kushner, & McLellan, 1991)	Semistructured or computerized interview	varies	45–60	None reported	Composite score (number of risk factors) and severity rating (need for treatment) across seven domains
Adolescent Self Assessment Profile (ASAP; Wanberg, 1991)	Self-report	203	45–60	Adequate internal consistency	Profile of adjustment, functioning, and substance use across multiple domains
Personal Experience Inventory (PEI; Winters & Henley, 1989)	Self-report, computer administered	276	45–60	Adequate internal consistency, test–retest reliability, concurrent, discriminant and predictive validity	Substance use frequency and severity, personal and environmental risk factors, problem screening and faking
Teen Addiction Severity Index (T-ASI; Kaminer, Bukstein, & Tarter, 1991)	Semistructured interview	134	30–45	Adequate interrator reliability	Problem severity and need for treatment across seven domains rated by both client and interviewer

5.21.3.2 Treatment Formulation

How assessment findings should be used in treatment planning has seldom been addressed in the literature. The National Institute on Drug Abuse has developed the Adolescent Assessment/Referral System which advocates an assessment battery approach for identifying potential problems across multiple domains, including substance use, diagnosis, family functioning, and psychological functioning. Using this system, treatment planning determining whether the individual needed substance abuse treatment, family therapy, or other interventions is guided on the problem areas identified in the assessment. The Drug Use Screening Inventory (Tarter, 1990) represents a single, brief screening instrument that provides an assessment across 10 domains of functioning. This instrument was developed as a treatment formulation instrument yielding assessment findings that lead directly to treatment planning. However, evaluations of the effectiveness of assessment systems for treatment formulation and subsequent adolescent treatment outcomes have not been studied.

More often, treatment formulation is guided by the therapeutic orientation of agency or treatment provider. In addition to assessing the topography of substance use and related behaviors for the purpose of evaluating change brought about by intervention, behavioral treatment models might include a functional analysis of the behavior which is incorporated into treatment (e.g., modifying contingencies in the environment on the basis of the functional analysis). Family treatment models may include standardized family assessment instruments and measures of family interaction, but, for treatment formulation, most often conduct assessments of family functioning in an ongoing fashion in the context of family therapy sessions (Alexander, Barton, Waldron, & Mas, 1983).

Some obvious circumstances, such as evaluating the need for medical supervision of detoxification or inpatient treatment for suicidal substance abusers, influence treatment matching for adolescents. Beyond such examples, however, strategies for matching adolescents to treatments on the basis of assessment findings are premature (Project MATCH, 1997). Identifying environmental treatments for medical problems with quite different etiology (e.g., dietary restriction for phenylketonuria) may apply to substance abuse as well, challenging how we think about treatment formulation. And, as Henly and Winters (1989) noted, a variety of paths have been found leading to substance involvement and a variety of paths may also lead out. Given our current level of understanding of substance abuse disorders, treatment formulation should be determined on the basis of what has been shown empirically to work (Wagner & Kassel, 1995).

5.21.4 TREATMENT

5.21.4.1 Psychosocial Treatments

Psychosocial treatment models for substance abuse include the broadest spectrum of approaches in settings ranging from inpatient intervention to street-based harm reduction programs (e.g., needle exchange), from residential living to outpatient primary-care interventions, and from juvenile detention to school- and community-based preventive services provided for secondary and tertiary care. Inpatient programs typically provide a detailed admission assessment, psychosocial evaluation, exploration of chemical use history and assessment of educational, legal, and psychological status (Wheeler & Malmquist, 1987). Program components include group therapy, lectures, work assignments, a therapeutic milieu, attendance at support meetings, family therapy, and recreational activities. Therapeutic communities, used for adolescents with more severe drug abuse, have similar orientations to inpatient treatment but are longer term, with an average length of stay of 6–9 months. Outpatient treatment can include intensive day-treatment programs, services provided by individual, group, and family therapists, and education-based interventions usually offered in larger contexts.

5.21.4.1.1 Traditional and 12-step models

Modalities of treatment offered across the range of settings have included individual, family based, and group therapies based on a wide variety of orientations such as traditional 12-step philosophies, cognitive-behavioral, family, and educational models as well as recreation therapies and vocational counseling. Inpatient and residential treatment programs have traditionally used a milieu approach, a multimodal treatment often including an education component, individual, group, and family therapy, recreational activities. In many clinical settings, approaches to adolescent treatment for substance use problems are based on the philosophy and principles of Alcoholics Anonymous and/or Narcotics Anonymous (AA/NA; Wheeler & Malmquist, 1987). The philosophy holds that individuals cannot effect change on their own. Rather, individuals must recognize a higher spiritual power to attain a healthy (and abstinent) lifestyle, drawing upon resources from within and support from others and following the 12-step process for recovery.

5.21.4.1.2 Cognitive-behavioral models

Behavioral and cognitive-behavioral approaches have been the focus of many preventive intervention programs (Gilchrist & Schinke, 1985; Pentz, 1985) and are increasingly being used to treat adolescent substance abuse (Turner, Liddle, & Dakof 1996; Waldron & Slesnick, 1997). These approaches, based on principles of learning, often incorporate a stress-coping model of substance abuse and focus on teaching adolescents appropriate skills to avoid substance use. Treatment components may include self-monitoring, social skills training (e.g., problem solving or assertiveness), mood regulation (e.g., relaxation training, anger management, or modifying cognitive distortions), and relapse prevention (e.g., drug and alcohol refusal skills). Modeling, behavior rehearsal, feedback, and homework assignments are characteristic of treatment techniques.

5.21.4.1.3 Family therapy models

Family therapy has been a widely accepted treatment for substance abuse, frequently implemented with adolescent abusers (Stanton & Shadish, 1997; Waldron, 1997). Family systems models view substance abuse as a reflection of relationship dysfunction, and treatment is aimed at restructuring the interactional patterns associated with the abuse, theoretically making the abuse unnecessary in the maintenance of system functioning (Stanton & Todd, 1982). Techniques focus on developing the therapeutic relationship, helping family members understand the interrelatedness of their behaviors and the symptomatic behavior in a way consistent with family change (e.g., reframing), and restructuring or shifting family interaction patterns and establishing new behaviors.

Behavioral family models emphasize training parents in contingency management aimed to increase prosocial behaviors incompatible with substance use and decrease substance use behavior. Communication and conflict-resolution skills training are commonly used to enhance families' behavioral repertoires and allow them to resolve problems independently. The increased positive interactions they experience when putting such skills to use are presumed to reinforce the likelihood that the new behaviors will become established patterns.

5.21.4.2 Pharmacological Treatments

5.21.4.2.1 Pharmacotherapy for substance abuse

Mechanisms for pharmacological treatments of substance abuse include treatments which make drug use aversive, provide an alternative substitute for the drug, block the reinforcing effects of the drug, or relieve drug craving or withdrawal (Kaminer, 1994). However, these approaches are rarely used with adolescents because of the particular topography of substance use among adolescents and the potential problems associated with improper use of pharmacological treatments, as well as a host of other practical and ethical considerations. For example, disulfiram (Antabuse) causes intense nausea and other adverse effects when combined with alcohol and is typically used with regular, heavy drinkers who are motivated to achieve abstinence. Noncompliance (i.e., drinking while taking disulfiram) can be medically dangerous. Because adolescents tend to drink more impulsively, consuming alcohol in episodic binges, the potential for noncompliance and subsequent medical complications is high. Thus, apart from the general ethical considerations concerning aversion therapies with adolescents and issues related to consent for treatment, disulfiram would be inappropriate for use with adolescents.

Nevertheless, no pharmacological evidence suggests that adolescents should be treated differently than adults. No specific contra-indications exist for nicotine gum or patch, heroin substitutes such as l-acetyl-methadol and buprenorphine, methylphenidate or other stimulants for cocaine abuse with adolescents (Kaminer, 1994). Substitution of methadone for heroin use may also be helpful in eliminating craving or blocking the euphoric effects of the drug in heroin-addicted adolescents for whom detoxification is contra-indicated (e.g., pregnant teens who abuse heroin). Although access to methadone programs for adolescents is more difficult than for adults, methadone may also reduce harm among adolescent heroin users by eliminating the danger of contracting AIDS from contaminated needles (Kaminer, 1994).

5.21.4.2.2 Pharmacotherapy for dual diagnosis

The most common pharmacotherapy for adolescent substance abusers is the use of psychoactive medications to treat symptoms or disorders coexisting with the abuse. For example, adolescents with a dual diagnosis of substance abuse disorder and major depressive disorder may be treated with antidepressants such as imipramine, amitriptyline, or other tricyclics, or with sertraline or other serotonin reuptake inhibitors. Obviously, in the presence of suicide risk, the cardiotoxicity of any medications less cardiotoxic must be considered. By the same token, use of medications with addictive properties such as benzodiazepines, which may be used for adolescents diagnosed with anxiety disorders, would not be appropriate for adolescents with comorbid substance abuse disorder.

In general, Bukstein (1995) has advised against the use of pharmacological agents for substance-abusing adolescents without significant coexisting psychopathology. The main issue with pharmacotherapy for adolescent substance use problems is lack of any empirical attention. In a rare study, Kaminer (1992) reported on the facilitation of cocaine abstinence in adolescent using tricyclic antidepressant desipramine, with six-month follow-up findings confirming continued abstinence. Future research is needed to validate pharmacological treatments.

5.21.4.3 Efficacy and Effectiveness of Psychosocial Treatments

Compared to the number and variety of treatment approaches implemented in clinical settings for adolescent substance abuse, remarkably little research has examined treatment effectiveness. Prior to 1980, most of the evidence was based on findings from several large-scale studies examining treatment outcomes across different modalities. Rush (1979), for example, found a number of patient characteristics (i.e., older onset, being in school or having a job, absence of other behavioral disturbance, monodrug use) were associated with better outcomes for 2938 adolescents in primarily in outpatient settings, but that treatment variables such as length of treatment were not. The Drug Abuse Reporting Program (Sells & Simpson, 1979), another large-scale study, compared outcomes for 5405 adolescents in four treatment settings: methadone maintenance, therapeutic community, outpatient, and detoxification. Reductions in substance use were found during treatment, although at follow-up only decreases in opiate use were observed. Another large scale study, the Treatment Outcome Prospective Study (Hubbard, Cavanaugh, Craddock, & Rachal, 1985), evaluated treatment outcomes for 240 adolescents participating in residential or outpatient programs. Although reduced substance use was found for both settings, residential treatment clients fared better than outpatient clients. Unfortunately, no control or comparison groups were included in these studies.

5.21.4.3.1 Traditional interventions

One study (Amini, Zilberg, Burke, & Salasnek, 1982), has compared a psychodynamically oriented inpatient drug treatment with an outpatient probation service intervention in a sample of 74 substance abusing adolescents. Subjects were randomly assigned to treatment condition and both groups showed significant decreases in drug use. No between-group differences were found. A number of other studies, however, have examined treatments based on AA/NA philosophy. Friedman, Schwartz, & Utada (1989) evaluated the outcome of a day treatment intervention program using a pre/post design. The program used a modified version of the AA 12 steps, daily meetings, guided group interaction and group therapy, ration-emotive and reality therapy approaches. The mean number of months for follow-up evaluation was 14.6. Of 222 participants, with a mean age of 16.6 years, 85% reported less drug use than when they started in the program. Another pre/post only study was conducted by Alford, Koehler, & Leonard (1991), who examined the effectiveness of AA/NA programs. They evaluated outcomes for 157 substance abusing adolescents treated in an inpatient setting and found that 71% of male treatment completers and 79% of female completers were abstinent six months following discharge. After two years, abstinence among males had dropped to 40%, among females to 61%. Grenier (1985) compared outcome of a residential treatment program based on AA principles with a waiting list control group. The treatment group showed significantly greater abstinence (65.5% vs. 14.3%) two years after treatment.

Multimodal programs have also been examined. Feldman (1983) found significant reductions in substance use for 71 black and Latino adolescents who participated in an intervention consisting of recreational activities, counseling, family intervention, and legal services. Barrett, Simpson, and Lehman (1988) evaluated four community-based drug abuse intervention programs providing individual, group, and family therapy, recreational and community service activities, and substance abuse education. Three months after intake, significant improvements for all problem behaviors were revealed, but no differences were found between programs. Braukmann et al. (1985) compared delinquent adolescents in Teaching Family group homes with living in a treatment-as-usual group home and an untreated group. The Teaching Family group homes, applying skills training, self-government, motivation, relationship development and youth advocacy, showed significant during-treatment decreases in substance use. However, these decreases did not persist at 12-month follow-up.

5.21.4.3.2 Behavioral outpatient interventions

Davidge and Forman (1988) reviewed eight case studies published between 1967 and 1985 which provided limited support for behavior therapy with adolescent substance abusers. These studies were limited by small sample sizes, lack of control or comparison conditions and, in many cases, only anecdotal reporting of substance use behavior. Two of the cases, however, used urine analysis to validate self-report of substance use, lending more strength to study claims. These cases used contingency contracting with the adolescent and parents, focusing on school, work, and relationships (Cook & Petersen, 1985; Fredericksen, Jenkins, & Carr, 1976). The adolescents in both cases had drastically reduced substance use or remained abstinent at post-treatment and at follow-up through one year in one case and three years in the other.

In other behavioral investigations, Duehn (1978) evaluated the effectiveness of aversive conditioning for adolescents with severe drug-use histories. At 6- and 18-month follow-up, completed abstinence was reported for six of the seven youth who had participated in the 14 week covert sensitization training program. DeJong and Henrich (1980) conducted a two year follow-up study of 89 young addicts in a behavior modification program who attended a rehabilitation center for at least seven days. They reported that one-third of the total sample remained drug free. In a set of two studies using a pre/post design, Iverson and colleagues (Iverson, Jurs, Johnson, & Rohen, 1978; Iverson & Roberts, 1980) conducted a community education-based program for parents and adolescents and found significant reductions in drug use at post-treatment and six-month follow-up.

Using a between-groups design, Smith (1983) compared an eight-session school-based group treatment, focusing on skills training in the areas of problem solving, social skills, and self-monitoring, with a no treatment control condition for adolescent marijuana users. Self-reported use was reduced and academic and peer functioning was enhanced for the youth in the treatment. In a residential treatment setting, Hawkins, Catalano, and Miller (1992) found improved role play performance involving avoidance of drug use and interpersonal problem solving in an applied behavioral social skills training intervention in adolescents in a residential treatment center.

For both traditional and behavioral interventions, wide variation with respect to control groups, random assignment, outcome measures, and number and latency of follow-up assessments can be seen. However, none of the studies comparing across two or more intervention conditions found evidence that any treatment approach examined was better than any other. In their review of nonfamily-based treatments, Catalano, Hawkins, Wells, Miller, & Brewer (1990–91) addressed the limitations of these studies and could only conclude that some treatment is better than no treatment. In the absence of more randomized clinical trials including appropriate alternative comparison conditions, no more definitive conclusions can be made. Several federally funded controlled trials of promising treatment approaches are currently underway and are likely to improve the knowledge base with respect to empirically validated treatments for adolescent substance abuse.

5.21.4.3.3 Family-based interventions

In the last decade or so, considerable advances have been made in evaluating the effectiveness of family treatments for adolescent substance abuse, with a dozen or more controlled trials completed or underway (Stanton & Shadish, 1997; Waldron, 1997). Much of the impetus for these advances stemmed from the early controlled studies of family therapy conducted by Szapocznik and his colleagues. Scopetta, King, Szapocznik, and Tillman (1979) investigated whether family interventions need to be ecologically focused, involving multiple systems, or whether intervention within the conjoint family context is sufficient. Although both groups showed improved outcomes following treatment, no differences were found between the two interventions. In another set of studies, Szapocznik, Kurtines, Foote, Perez-Vidal, and Hervis (1983, 1986) compared two variations of structural family systems therapy, conjoint family therapy and one-person family therapy, examining whether family therapy outcome varies on the basis of who participates in the sessions. Families in both intervention conditions showed significant improvement at the end of therapy. Moreover, compared to conjoint therapy, one-person family therapy led to slightly greater improvement for some outcome measures. These studies were limited by small sample sizes, high attrition, inadequate substance use measurement, and the absence of a nonfamily comparison condition. However, like other early outcome studies, the findings of pre/post changes were promising.

Other randomized trials comparing systems family therapy to alternative family-based interventions have also been conducted. Friedman (1989) found pre/post differences comparing the effectiveness of Functional Family Therapy (Alexander & Parsons, 1982) to a parenting skills group intervention. Although no between-group differences were found, 93% of the families in family therapy engaged in treatment compared to only 67% of the parenting group families. In a reanalysis of these data which included treatment dropouts as failures, Stanton and Shadish (1997) reported that a significant difference between the two groups emerged, highlighting the importance of considering differential attrition in conducting treatment outcome research. Lewis, Piercy, Sprenkle, and Trepper (1990) also evaluated an integrative family therapy model, comparing Purdue Brief Family Therapy, to a didactic, family-oriented parenting skills intervention. Both interventions were found to significantly reduce drug use at post-treatment. However, a greater percentage of youth receiving family therapy decreased their use, compared to the parenting group intervention. No follow-up data were reported.

A number of randomized trials have included nonfamily comparison interventions in their designs. Joanning, Thomas, Quinn, and Mullen (1992) compared a family systems intervention to two comparison conditions, a family drug education intervention and adolescent group therapy. The results at post-treatment revealed that family therapy was most effective in decreasing problem behaviors and drug use, with 54% of youth in family therapy abstinent, compared to 29% in the family education condition and 17% in the group therapy condition. Using a more complex ecological intervention, Henggeler et al. (1991) conducted two studies, comparing multisystemic family therapy to individual therapy in the first study and to a treatment-as-usual juvenile probation supervision intervention in the second. As had been found in treating delinquents, multisystemic family therapy produced significantly better outcomes at post-test. Liddle et al. (1993) compared a multisystemic, multidimensional family therapy to two alternative treatments: adolescent group therapy and a family education group. All three conditions were associated with reduced substance use, with greatest and most consistent improvement in the family therapy condition, in which youth decreased their use from daily to about once a week for alcohol and marijuana, with 30% abstinent. Group therapy also showed gradual decline in drug use from pretreatment to follow-up, not apparent at post-treatment, and the authors suggested that the effect may have resulted from the delay between adolescents putting newly acquired skills into use and deriving benefit from implementing them.

Two behavioral family therapy trials have also been conducted. Azrin, Donohue, Besalel, Kogan, and Acierno (1994) compared behavioral family therapy to a process-oriented, nondirective (supportive) adolescent group therapy intervention. Compared to the supportive group intervention, behavioral family therapy produced greater reductions in substance use and improvements in other areas of functioning. In another behavioral family therapy evaluation, Krinsley and Bry (1997) examined the impact of a school-based behavioral family therapy intervention compared to a school-based intervention only condition. Relative to pretreatment, students who participated in family therapy showed greater reductions than students in the school intervention alone during the 15 months following treatment. The differences between groups appeared even greater over time, illustrating the importance of conducting follow-up assessments.

In the first controlled study of its kind, Turner, Liddle, and Dakof (1996) recently reported results of a randomized clinical trial comparing cognitive-behavior therapy and multisystemic family therapy for adolescent substance abuse to a treatment-as-usual condition. The cognitive-behavioral and family interventions both included individual and conjoint family sessions. The cognitive-behavioral condition included self-monitoring, communication and problem-solving skills training, contingency contracting, and substance refusal skills, while the family condition was a multidimensional family intervention (Liddle et al., 1993). Preliminary results indicated that both of the experimental treatments produce greater reductions in substance use relative to treatment-as-usual. However, some findings suggest that although there were no differences between the two active conditions through six month follow-up, the cognitive-behavioral treatment appeared more effective through a 12 month follow-up. However, data analysis is incomplete and no conclusions can be drawn until the final results are examined.

Taken together, the results of controlled clinical trials evaluating family-based interventions suggest that family therapy is an effective treatment for adolescents substance abuse. Only one study (Henggeler et al., 1991) reported finding no differences between family therapy and a nonfamily intervention. Every study demonstrated significant reductions in substance use from pre- to post-treatment. In seven of eight studies comparing family therapy to a nonfamily-based intervention, youth receiving family therapy showed greater reductions in substance use than those receiving adolescent group therapy (Azrin et al., 1994; Joanning et al., 1992; Liddle et al., 1993), family education (Joanning et al., 1992; Liddle et al., 1993), and individual therapy, individual tracking through schools, or juvenile justice systems (Henggeler et al., 1991; Krinsley & Bry, 1997). When compared to parenting skills training programs, family therapy fared better in the Lewis et al. (1990) study and in the Friedman (1989) study when dropouts were included in analysis. Stanton and Shadish (1997) give effect sizes in the their meta-analysis of family therapy outcome studies.

5.21.4.3.4 Treatment engagement and retention

Treatment retention in alcohol and drug abuse populations is a widely recognized problem. Adolescents are at even greater risk for dropping out of treatment, compared with adults, because they are less motivated to change, do not perceive treatment as suitable for themselves, and are more likely to be in treatment because of external pressures (Jainchill, Bhattacharya, & Yagelka, 1995). Kaminer and Frances (1991) prospectively compared treatment noncompleters with completers between the ages of 13 and 18, in an inpatient dual disorder substance abuse treatment program designed specifically for adolescents, hypothesizing that treatments which simultaneously address psychiatric issues would improve treatment outcome. Fifty adolescents completed treatment, 14 did not. Involuntary admission, past treatment, and parental psychopathology had no impact on dropout. A trend towards a higher prevalence of mood disorder and adjustment disorder was found in treatment completers. In contrast, the treatment dropout group had a higher prevalence of conduct disorder. Jainchill et al. (1995) also examined factors related to treatment retention in two residential treatment centers, comparing clients age 17 and under with clients 18 to 20 years old. They found that older clients felt more motivated and ready for treatment and perceived treatment to be more suitable for them. The most consistent predictor of treatment retention was age.

A recent research project evaluating the effectiveness of residential treatment for adolescent substance abuse has received funding through the National Institute on Drug Abuse (Jainchill et al., 1995). Data collection has been completed for 938 adolescents, but follow-up data are not yet available. However, several variables correlated significantly with retention. Consistent with Kaminer and Francis (1991), antisocial lifestyle and poorer psychological functioning may predict early dropout. Other predictors, consistent with Jainchill et al. (1995), included measures of self-esteem, motivation and readiness for change, environmental risk factors, and the interviewer's impressions of the client's likelihood of staying in the program.

Recognizing the problem of engaging drug abusers and their families in treatment, Szapocznik and his colleagues turned their research efforts toward evaluating ways to enhance patient retention in family therapy. In one study, Szapocznik et al. (1988) compared their family systems therapy in combination with an intensive engagement intervention to family systems therapy alone (with engagement as usual). The results were quite pronounced: 93% of the substance abusers and their families in the intensive engagement condition completed intake compared to 47% of the engagement as usual group. Once engaged in treatment, there were no further attrition differences between the groups.

In the second study, Santisteban et al. (1996) replicated the findings of the first study, and even with stricter criteria for treatment engagement they found 81% of the substance abusers and their families in the intensive engagement condition completing intake compared to 60% in the comparison conditions. Moreover, they found that ethnicity was associated with a differential engagement rate in the intensive condition: 97% of the non-Cuban Hispanics in the intensive condition were engaged vs. 64% of the Cuban Hispanics, allowing them to explore further the role of ethnicity in family therapy process. Taken together, motivational issues and their role in retention and recovery should continue as a focus study.

5.21.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

Research on adolescent substance abuse has been characterized by disparate definitions of abuse, wide variations in subject sampling strategies, and little systematic attention to heterogeneity among adolescents with substance abuse disorders. Such methodological problems may represent primary sources for inconsistency of findings and noncomparability across studies. However, important advances in assessing adolescent substance use have been made in recent years, with the development of a variety of instruments specifically designed for adolescents (Winters & Stinchfield, 1995). Establishing the psychometric properties of these instruments and developing more uniform assessment procedures for research and practice is needed to facilitate the development of treatment formulation procedures and facilitate treatment effectiveness comparisons. Greater collaboration among researchers in this field would also improve comparisons across studies and could yield greater convergence and consistency of findings.

Research is also needed to evaluate assessment strategies more sensitive to patterning of use over time. Some research has shown, for example, that family factors relate more powerfully to stage of substance use than to frequency of use (Brook, Whiteman, & Gordon, 1982). Also, using latent growth modeling to detect similarities and differences in developmental trajectories for substance use, Duncan et al. (1994) found that family cohesion may be more central in delaying escalation of cigarette use than escalation of other substances. With a more detailed and precise assessment of quantity, frequency, and patterning of use, researchers may be able to evaluate the effectiveness and efficacy of family therapies in more sensitive or meaningful ways. For example, forestalling adolescents' progression through the drug use sequence or initiating their return back through the gateway may reflect treatment effectiveness better than overall amount or frequency of use.

Liddle et al. (1993) have attempted to develop a developmentally sensitive treatment model for families of adolescent substance abusers. With respect to gender, different models may also be needed, with girls' substance abuse perhaps best understood in the context of factors such as reasons for using, negative and positive consequences, and expectancies related to use, or family relationships. Considerably more attention has been given to examining the influence of ethnicity on therapy process and outcome (Santisteban et al., 1996; Szapocznik et al., 1988). However, broad-based culturally sensitive treatment models have yet to be developed. Future research should focus on examining the components of treatment models, evaluating the importance of the adolescent development, gender, and ethnicity in relation to treatment effectiveness.

Treatment effectiveness studies have lagged far behind other areas of adolescent substance abuse. Moreover, the vast majority of studies conducted have been plagued with methodological problems. In addition to design and substance use assessment issues, at least half of the studies did not include follow-up assessment, a critical limitation given that several studies conducting follow-up assessments found that patterns of substance use reductions changed over time (Krinsley & Bry, 1997; Liddle et al., 1993; Szapocznik et al., 1983, 1986) and different conclusions could be drawn about treatment effectiveness on the basis of point in time of measurement. Moreover, differential attrition across conditions may have created confounds in interpreting the results of the outcome studies (Snow, Tebes, & Arthur, 1992; Stanton & Shadish, 1997). Some research has also revealed differential dropout rates in substance abusers with dual diagnoses. Thus, comorbid conditions may play a role in the course of treatment as well. More effort needs to be exerted to follow treatment dropouts and examine patterns of outcome related to client heterogeneity.

None of the studies reviewed reported examining therapy process variables which may have been associated with change. Within the broader psychotherapy literature, therapy process and outcome research has shown great promise in isolating therapist, client, and interactional mediators of outcome. The equivocal findings in the clinical trials for adolescent substance abuse highlight the need for incorporating a process perspective in this area of endeavor as well.

A number of family-based clinical trials compared two or more interventions to one another. However, no clear pattern emerged for the superiority of one treatment model over another. Few studies indicated use of a treatment manual, and fewer still used actual adherence ratings to ensure that the components described were actually being delivered. It is possible that no differences on the basis of theoretical model were found because, despite what techniques and procedures treatment researchers say are being implemented, therapists engaged in essentially the same behaviors across models. Alternatively, therapists may have been true to their models, but the effects of common factors present across all interventions such as therapist empathy or warmth and/or treatment expectancies or motivation may have been critical mediators of change.

Creating a motivational context for change represents a major phase of treatment presumed to produce higher engagement and retention rates in some models of family therapy (Alexander & Parsons, 1982). Similarly, Miller and Rollnick (1991) have developed a motivational enhancement approach that has been shown to be effective as an add-on to traditional treatments and as a stand-alone brief intervention for adult problem drinkers. The central feature of this intervention, like the motivational phase of family therapy, is the nonblaming, nonjudgmental stance taken by the therapist. This focus is wholly consistent with the observation of Stanton and Shadish that a nonconfrontational, nonblaming perspective appears to be essential in the treatment of families with substance abusing members and this component of therapy process merits close examination in future treatment studies with adolescents.

5.21.6 SUMMARY

The lack of uniform definition of substance abuse in adolescents has impeded all areas of research including epidemiology, etiology, diagnosis, assessment, treatment development, and treatment outcome evaluation. Studies of risk and protective factors have identified a variety of correlates of abuse that facilitate prediction of who may develop problems or be appropriate candidates for prevention or intervention. The appearance of well-designed clinical trials evaluating interventions in recent years begins to address the issue of optimal treatment strategies for adolescent substance abuse. Research efforts have led to the identification of several treatment approaches that appear to reduce adolescent substance use, relative to no treatment or alternative treatment comparisons. Family therapy, in particular, has engendered substantial empirical support. However, the mechanisms associated with treatment effectiveness have yet to be examined. Researchers need to channel more efforts into prospective model building, incorporating empirical work from the broader substance abuse literature and relying on more uniform, comprehensive, detailed measurement of use. Moreover, psychotherapy process research is needed to examine aspects of the change process associated with better outcomes.

5.21.7 REFERENCES

Alexander, J. F., Barton, C., Waldron, H. B., & Mas, C. H. (1983). Beyond the technology of family therapy: The anatomy of intervention model. In K. D. Craig & R. J. McMahon (Eds.), *Advances in clinical behavior therapy* (pp. 48–73). New York: Brunner/Mazel.

Alexander, J. F., & Parsons, B. V. (1982). *Functional family therapy*. Monterey, CA: Brooks/Cole.

Alford, G. S., Koehler, R. A., & Leonard, J. (1991). Alcoholics Anonymous–Narcotics Anonymous model inpatient treatment of chemically dependent adolescents: A 2-year outcome study. *Journal of Studies on Alcohol, 52,* 118–126.

Alterman, A. I., & Tarter, R. E. (1986). An examination of selected topologies: Hyperactivity, familial and antisocial alcoholism. In M. Galanter (Ed.), *Recent developments in alcoholism* (Vol. 4, pp. 169–189). New York: Plenum.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. Rev.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: Author.

Amini, F., Zilberg, N. J, Burke, E. L., & Salasnek, S. (1982). A controlled study of inpatient vs. outpatient treatment of delinquent drug abuse adolescents: One year results. *Comprehensive Psychiatry, 23,* 436–444.

Andersson, T., & Magnusson, D. (1988). Drinking habits and alcohol abuse among young men: A prospective longitudinal study. *Journal of Studies on Alcohol, 49,* 245–252.

Andrews, J. A., Hops, H., Ary, D., Lichtenstein, E., & Tildesley, E. (1991). The construction, validation, and use of a Guttman scale of adolescent substance use: An investigation of family relationships. *Journal of Drug Issues, 21,* 557–572.

Azrin, N. H., Donohue, B., Besalel, V. A., Kogan, E. S., & Acierno, R. (1994). Youth drug abuse treatment: A controlled outcome study. *Journal of Child and Adolescent Substance Abuse, 3,* 1–16.

Babor, T. F., Hofmann, M., DelBoca, F. K., Hesselbrock, V., Meyer, R. E., Dolinsky, Z. S., & Rounsaville, B. (1992). Types of alcoholics. I. Evidence for an empirically derived typology based on indicators of vulnerability and severity. *Archives of General Psychiatry, 49,* 599–608.

Bachman, J. G., Wallace, J. M., O'Malley, P. M., Johnston, L. D., Kurth, C. L., & Neighbors, H. W. (1991). Racial/ethnic differences in smoking, drinking and illicit drug use among American high school seniors, 1976–89. *American Journal of Public Health, 81,* 372–377.

Bailey, S. L., Flewelling, R. L., & Rachal, J. V. (1992). The characterization of inconsistencies in self-reports of alcohol and marijuana use in a longitudinal study of adolescents. *Journal of Studies on Alcohol, 53,* 636–657.

Barkley, R. A., Fischer, M., Edelbrock, C. S., & Smallish, L. (1990). The adolescent outcome of hyperactive children diagnosed by research criteria: I. An 8-year prospective follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 546–557.

Barnes, G. M. (1981). Drinking among adolescents: A subcultural phenomenon or a model of adult behaviors? *Adolescence, 16*(61), 211–229.

Barnes, G. M. (1984). Adolescent alcohol abuse and other problem behaviors: Their relationships and common parental influences. *Journal of Youth and Adolescence, 13,* 329–348.

Barrett, M. E., Simpson, D. D., & Lehman, W. E. K. (1988). Behavioral changes of adolescents in drug abuse intervention programs. *Journal of Clinical Psychology, 44,* 461–473.

Baumrind, D. (1985). Familial antecedents of adolescent drug use: A developmental perspective. In C. L. Jones and R. J. Battjes (Eds.), *Etiology of drug abuse: Implications for prevention,* (pp. 13–44). Rockville, MD: National Institutes of Health.

Bentler, P. M. (1992). Etiologies and consequences of adolescent drug use: Implications for prevention. *Journal of Addictive Diseases, 11,* 47–61.

Blane, H. (1976). Middle-aged alcoholics and young drinkers. In H. Blane & M. Chafetz (Eds.), *Youth, alcohol and social policy* (pp. 5–38). New York: Plenum.

Blum, R. W. (1987). Adolescent substance abuse: Diagnostic and treatment issues. *Pediatric Clinics of North America, 34,* 523–531.

Braukmann, C. J., Bedlington, M. M., Belden, B. D., Braukmann, B. P. D., Husted, J. J., Ramp, K. K., & Wolf, M. M. (1985). Effects of a community-based group-home treatment program on male juvenile offenders use and abuse of drugs and alcohol. *American Journal of Drug and Alcohol Abuse, 11,* 249–278.

Brook, J. S., Brook, D. W., Gordon, A. S., Whiteman, M.,

& Cohen, P. (1990). The psychosocial etiology of adolescent drug use: A family interactional approach. *Genetic, Social, and General Psychology Monographs, 116,* 111–267.

Brook, J. S., Whiteman, M., & Gordon, A. S. (1982). Qualitative and quantitative aspects of adolescent drug use: Interplay of personality, family, and peer correlates. *Psychological Reports, 51,* 1151–1163.

Brown, S. A., Mott, M. A., & Myers, M. G. (1990). Adolescent alcohol and drug treatment outcome. In R. R. Watson (Ed.), *Drug and Alcohol Abuse Prevention, Drug and Alcohol Abuse Prevention, Drug and Alcohol Abuse Reviews* (pp. 373–403). Clifton, NJ: Humana.

Brown, S. A., Vik, P. W., & Creamer, V. A. (1989). Characteristics of relapse following adolescent substance abuse treatment. *Addictive Behaviors, 14,* 291–300.

Bry, B. H., McKeon, P., & Pandina, R. J. (1982). Extent of drug use as a function of number of risk factors. *Journal of Abnormal Psychology, 91,* 273–279.

Bukstein, O. G. (1995). *Adolescent substance abuse: Assessment, prevention, and treatment*. New York: Wiley.

Bukstein, O. G., Glancy, L. G., & Kaminer, Y. (1992). Patterns of affective comorbidity in a clinical population of dually diagnosed adolescent substance abusers. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 1041–1045.

Catalano, R. F., Hawkins, J. D., Wells, E. A., Miller, J., & Brewer, D. (1990–91). Evaluation of the effectiveness of adolescent drug abuse treatment, assessment of risks for relapse, and promising approaches for relapse prevention. *The International Journal of the Addictions, 25, 1085–1140.*

Centers for Disease Control and Prevention. (1996). Tobacco use and sources of cigarettes among high school students—United States, 1995. *Morbidity and Mortality Weekly Report, 45*(20).

Clayton, R. R. (1992). Transitions in drug use: Risk and protective factors. In M. Glantz & R. Pickens (Eds.), *Vulnerability to drug abuse* (pp. 15–51). Washington, DC: American Psychological Association.

Cohen, P., Cohen, J., Kasen, S., Velez, C. M., Hartmark, C., Johnson, J., Rojas, M., Brook, J., & Streuning, E. L. (1993). An epidemiological study of disorders in late childhood and adolescence. I. Age and gender-specific prevalence. *Journal of Child Psychology and Psychiatry, 34, 851–867.*

Cook, P. S., & Petersen, D. (1985). Individualizing adolescent drug abuse treatment. In A. S. Friedman & G. Beschner (Eds.), *Treatment services for adolescent substance abusers* (pp. 164–177). Rockville, MD: National Institutes of Health.

Crowley, T. J., & Riggs, P. D. (1995). Adolescent substance use disorder with conduct disorder and comorbid conditions. In E. Rahdert & D. Czechowicz (Eds.), *Adolescent drug abuse: Clinical assessment and therapeutic interventions,* (pp. 49–111). Rockville, MD: National Institutes of Health.

Davidge, A. M., & Forman, S. G. (1988). Psychological treatment of adolescent substance abusers: A review. *Children and Youth Services Review, 10,* 43–55.

DeJong, R., & Henrich, G. (1980). Follow-up results of a behavior modification program for juvenile drug addicts. *Addictive Behaviors, 5,* 49–57.

DeLeon, G. (1984, April). *Adolescent substance abusers in the therapeutic community: Treatment outcomes.* Paper presented at the Research Conference on Juvenile Offenders with Serious Alcohol, Drug Abuse, and Mental Health Problems.

Deykin, E., Levy, J. C., & Wells, V. (1986). Adolescent depression, alcohol and drug abuse. *American Journal of Public Health, 76,* 178–182.

DiMileo, L. (1989). Psychiatric syndromes in adolescent substance abusers. *American Journal of Psychiatry, 146,* 1212–1214.

Dishion, T. J., Patterson, G. R., & Reid, J. R. (1988). Parent and peer factors associated with drug sampling in early adolescence: Implications for treatment. In E. R. Rahdert & J. Grabowski (Eds.), *Adolescent drug abuse: Analyses of treatment research* (pp. 69–93). Rockville, MD: National Institutes of Health.

Donovan, J. E., & Jessor, R. (1985). Structure of problem behavior in adolescence and young adulthood. *Journal of Consulting and Clinical Psychology, 56,* 890–904.

Duehn, W. D. (1978). Covert sensitization in group treatment of adolescent drug abusers. *International Journal of the Addictions, 13,* 485–491.

Duncan, T. E., Duncan, S. C., & Hops, H. (1994). The effects of family cohesiveness and peer encouragement on the development of adolescent alcohol use: A cohort-sequential approach to the analysis of longitudinal data. *Journal of Studies on Alcohol, 55,* 588–599.

Ellickson, P. L., Hays, R. D., & Bell, R. M. (1992). Stepping through the drug use sequence: Longitudinal scalogram analysis of initiation and regular use. *Journal of Abnormal Psychology, 101,* 441–451.

Farrell, A. D., Danish, S. J., & Howard, C. W. (1992). Relationship between drug use and other problem behaviors in urban adolescents. *Journal of Consulting and Clinical Psychology, 60,* 705–712.

FeCaces, M., Stinson, F. S., & Hartford, T. C. (1991). Alcohol use and physically risky behavior among adolescents. *Alcohol Health and Research World, 15,* 228–233.

Feldman, H. W. (1983, November). *A summary of the Youth Environment Study Final Reports*. State of California, Department of Alcohol and Drug Programs, Sacramento, CA.

Filstead, W. F. (1982). Adolescence and alcohol. In E. M. Pattison & E. Kaufman (Eds.), *Encyclopedic handbook of alcoholism,* (pp. 156–178). New York: Gardner.

Filstead, W. F., Parrella, D. P., & Conlin, J. M. (1988). Alcohol use and dependency in youth: Examining DSM-III diagnostic criteria. *Drugs & Society, 3,* 145–170.

Flay, B. R., & Petraitis, J. (1994). The theory of triadic influence: A new theory of health behavior with implications for preventive interventions. *Advances in Medical Sociology, 4,* 19–44.

Frederiksen, L., Jenkins, J., & Carr, C. (1976). Indirect modification of adolescent drug abuse using contingency contracting. *Journal of Behavior Therapy and Experimental Psychiatry, 7,* 377–378.

Friedman, A. S. (1989). Family therapy vs. parent groups: Effects on adolescent drug abusers. *The American Journal of Family Therapy, 17,* 335–347.

Friedman, A. S., Schwartz, R., & Utada, A. (1989). Outcome of a unique youth drug abuse program: A follow-up study of clients of Straight, Inc. *Journal of Substance Abuse Treatment, 6,* 259–268.

Friedman, A. S., & Utada, A. (1989). A method for diagnosing and planning the treatment of adolescent drug abusers: The Adolescent Drug Abuse Diagnosis (ADAD) insrument. *Journal of Drug Education, 19*(4), 285–312.

Gilchrist, L. D., Schinke, S. P. (1985). Preventing substance abuse with children and adolescents. *Journal of Consulting and Clinical Psychology, 53,* 121–135.

Grenier, C. (1985). Treatment effectiveness in an adolescent chemical dependency treatment program: A quasi-experimental design. *International Journal of the Addictions, 20,* 381–391.

Harrell, A. V., & Wirtz, P. W. (1989). *Adolescent Drinking Index Test and Manual.* Odessa, FL: Psychological Assessment Resources.

Hasin, D. S., Grant, B. F., & Endicott, J. (1990). The natural history of alcohol abuse: Implications for definitions of alcohol use disorders. *American Journal of Psychiatry, 147,* 1537–1541.

Hawkins, J. D., Catalano, R. F., & Miller, J. Y. (1992). Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: Implications for substance abuse prevention. *Psychological Bulletin, 112,* 64–105.

Henggeler, S. W., Borduin, C. M., Melton, G. B., Mann, B. J., Smith, L. A., Hall, J. A., Cone, L., & Fucci, B. R. (1991). Effects of multisystemic therapy on drug use and abuse in serious juvenile offenders: A progress report from two outcome studies. *Family Dynamics of Addition Quarterly, 1,* 40–51.

Henly, G. A., & Winters, K. C. (1989). Development of psychosocial scales of the assessment of adolescents involved with alcohol and drugs. *International Journal of the Addictions, 24,* 973–1001.

Hops, H., Tildesley, E., Lichtenstein, E., Ary, D., & Sherman, L. (1990). Parent–adolescent problem-solving interactions and drug use. *American Journal of Drug and Alcohol Abuse, 16,* 239–258.

Hubbard, R. L., Cavanaugh, E. R., Craddock, S. G., & Rachal, J. V. (1985). Characteristics, behaviors, and outcomes for youth in the TOPS. In A. S. Friedman & G. M. Beschner (Eds.), *Treatment services for adolescent substance abusers* (pp. 49–65). Rockville, MD: National Institutes of Health.

Hughes, S. O., Power, T. G., & Francis, D. J. (1992). Defining patterns of drinking in adolescence: A cluster analytic approach. *Journal of Studies on Alcohol, 53,* 40–47.

Iverson, D. C., Jurs, S., Johnson, L., & Rohen, R. (1978). The effect of an education intervention program for juvenile drug abusers and their parents. *Journal of Drug Education, 8,* 101–111.

Iverson, D. C., & Roberts, T. E. (1980). The juvenile intervention program: Results of the process, impact, and outcome evaluation. *Journal of Drug Education, 10,* 289–301.

Jainchill, N., Bhattacharya, G., & Yagelka, J. (1995). Therapeutic communities for adolescents. In E. Rahdert & D. Czechowicz (Eds.), *Adolescent drug abuse: Clinical assessment and therapeutic interventions* (pp. 190–217). Rockville, MD: National Institutes of Health.

Jessor, R., & Jessor, S. (1977). *Problem behavior and psychosocial development: A longitudinal study of youth.* New York: Academic Press.

Joanning, H., Thomas, F., Quinn, W., & Mullen, R. (1992). Treating adolescent drug abuse: A comparison of family systems therapy, group therapy, and family drug education. *Journal of Marital and Family Therapy, 18,* 345–356.

Johnston, L. D., O'Malley, P. M., Bachman, J. G. (1993). *National survey results on drug use from the Monitoring the Future Study, 1975–1992.* Rockville, MD: National Institutes of Health.

Kaminer, Y. (1992). Desipramine facilitation of cocaine abstinence in an adolescent. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 312–317.

Kaminer, Y. (1994). *Adolescent substance abuse: A comprehensive guide to theory and practice.* New York: Plenum.

Kaminer, Y., Bukstein, O., & Tarter, R. E. (1991). The Teen-Addiction Severity Index: Rationale and reliability. *International Journal of the Addictions, 26*(2), 219–226.

Kaminer, Y., & Frances, R. J. (1991). Inpatient treatment of adolescents with psychiatric and substance abuse disorder. *Hospital and Community Psychiatry, 42,* 894–896.

Kandel, D. B. (1975). Stages in adolescent involvement in drug use. *Science, 190,* 912–914.

Kandel, D. B., Kessler, R. C., & Margulies, R. S. (1978). Antecedents of adolescent initiation into stages of drug use: A developmental analysis. *Journal of Youth and Adolescence, 7,* 13–40.

Kandel, D. B., & Logan, J. A. (1984). Patterns of drug use from adolescence to young adulthood: I. Periods of risk for initiation, continued use, and discontinuation. *American Journal of Public Health, 74,* 660–666.

Kandel, D. B., Yamaguchi, K., & Chen, K. (1992). Stages of progression in drug involvement from adolescence to adulthood: Further evidence for the gateway theory. *Journal of Studies on Alcohol, 53,* 447–457.

Kashani, J. H., Beck, N. C., Hoeper, E. W., Fallahi, C., Corcoran, C. M. McAllister, J. A., Rosenberg, T. K., & Reid, J. C. (1987). Psychiatric disorders in a community sample of adolescents. *American Journal of Psychiatry, 144,* 584–589.

Krinsley, K. E., & Bry, B. H. (1997). *Decreasing school failure and early substance use in high risk adolescents through co-ordinated behavioral family and school intervention.* Manuscript submitted for publication.

Labouvie, E. W. (1986). The coping function of adolescent alcohol and drug use. In R. K. Silbereisen, K. Eyferth, & G. Rudinger (Eds.), *Development as action in context: Problem behavior and normal youth development* (pp. 227–240). New York: Springer-Verlag.

Leccese, M., & Waldron, H. B. (1994). Assessing adolescent substance abuse: A critique of current measurement instruments. *Journal of Substance Abuse Treatment, 11,* 553–563.

Lettieri, D., Sayers, M., & Pearson, H. (Eds.) (1980). *Theories on drug abuse: Selected contemporary perspectives.* Rockville, MD: National Institutes of Health.

Lewis, R. A., Piercy, F. P., Sprenkle, D. H., & Trepper, T. S. (1990). Family-based interventions for helping drug-abusing adolescents. *Journal of Adolescent Research, 5,* 82–95.

Liddle, H. A., Dakof, G. A., Parker, K., Barrett, K., Diamond, G. S., Garcia, R., & Palmer, R. (1993). *Effectiveness of family therapy vs. multi-family therapy and group therapy: Results of the Adolescents and Families Project—A randomized clinical trial.* Paper presented at the Annual Meeting of the Society for Psychotherapy Research, Pittsburgh, PA.

Martin, C. S., Kaczynski, N. A., Maisto, S. A., Bukstein, O. G., & Moss, H. B. (in press). Patterns of alcohol abuse and dependence symptoms in adolescent drinkers. *Journal of Studies on Alcohol.*

McCarthy, R. G., & Douglass, E. M. (1949). *Alcohol and social responsibility.* New York: Thomas Y, Crowell C, and Yale Plan Clinic.

McCrady, B. S., Noel, N. E., Abrams, D. B., Stout, R. L., Newlson, H. F., & Hay, W. M. (1986). Comparative effectiveness of three types of spouse involvement in outpatient behavioral alcoholism treatment. *Journal of Studies on Alcohol, 47,* 459–467.

McGee, L., & Newcomb, M. D. (1992). General deviance syndrome: Expanded hierarchical evaluations at four ages from early adolescence to adulthood. *Journal of Consulting and Clinical Psychology, 60,* 766–776.

Metzger, D. S., Kushner, H., & McLellan, A. T. (1991). *Adolescent Problem Severity Index Administration Manual.* Philadelphia, PA: Biomedical Computer Research Institute.

Miller, G. (1990). *Substance Abuse Subtle Screening Inventory-Adolescent (SASSI-A).* Bloomington, IN: SASSI Institute.

Miller, W. R., & Rollnick, S. (1991). *Motivational interviewing.* New York: Guilford Press.

Miller, W. R., Westerberg, V., & Waldron, H. B. (1995). In R. K. Hester & W. R. Miller (Eds.), *Handbook of alcoholism treatment approaches: Effective alternatives* (2nd ed., pp. 17–53). New York: Allyn & Bacon.

Nathan, P. E. (1991). Substance use disorders in DSM-IV. *Journal of Abnormal Psychology, 100,* 356–361.

Needle, R., McCubbin, H., Lorence, J., & Hochhauser, M. (1983). Reliability and validity of adolescent self-reported drug use in a family-based study: A methodological report. *International Journal of Addition, 18,* 901–912.

Newcomb, M. D. (1995). Identifying high-risk youth: Prevalence and patterns of adolescent drug abuse. In E. Rahdert & D. Czechowicz (Eds.), *Adolescent drug abuse: Clinical assessment and therapeutic interventions,* (pp. 7–37). Rockville, MD: National Institutes of Health.

Newcomb, M. D., & Bentley, P. M. (1989). Substance abuse and abuse among children and teenagers. *American Psychologist, 44,* 242–248.

Newcomb, M. D., Maddahian, E., & Bentler, P. M. (1986). Risk factors for drug use among adolescents: Concurrent and longitudinal analyses. *American Journal of Public Health, 76,* 525–531.

O'Farrell, T. J. & Maisto, S. A. (1987). The utility of self report and biological measures of alcohol consumption in alcoholism treatment outcome studies. *Advances in Behaviour Research and Therapy, 9,* 91–125.

Pandina, R. J. (1986). Methods, problems, and trends in studies of adolescent drinking practices. *Annals of Behavioral Medicine, 8,* 20–26.

Paperny, D. M., Aono, J. Y., Lehman, R. M., Hammar, S. L., & Riusser, J. (1990). Computer-assisted detection and intervention in adolescent high-risk health behaviors. *Journal of Pediatrics, 116,* 456–462.

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial boys.* Eugene, OR: Castalia.

Pentz, M. A. (1985). Social competence skills and self-efficacy as determinants of substance use in adoelscents. In S. Shiffman & T. A. Wills (Eds.), *Coping and substance use* (pp. 117–142). New York: Academic Press.

Petchers, M. K., & Singer, M. I. (1987). Perceived Benefit of Drinking Scale: Approach to screening for adolescent alcohol abuse. *Journal of Pediatrics, 110,* 977–981.

Petraitis, J., & Flay, B. R., & Miller, T. Q. (1995). Reviewing theories of adolescent substance use: Organizing pieces in the puzzle. *Psychological Bulletin, 117,* 67–86.

Project MATCH (1997). Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol, 58,* 7–29.

Prosser, W. L. (1954). The narcotic problem. *UCLA Law Review, 1,* 405–546.

Rahdert, E. R. (Ed.) (1991). *The Adolescent Assessment/Referral System Manual.* Rockville, MD: US Department of Health and Human Services.

Reinherz, H. Z., Giaconia, R. M., Lefkowitz, E. S., Pakiz, B., & Frost, A. K. (1993). Prevalence of psychiatric disorders in a community population of older adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 369–377.

Rounsaville, B. J. (1987). An evaluation of the DSM-III substance-use disorders. In G. Tischler (Ed.), *Treatment and classification in psychiatry* (pp. 175–194). New York: Cambridge University Press.

Rush, T. V. (1979). Predicting treatment outcomes for juvenile and young-adult clients in the Pennsylvania substance-abuse system. In G. M. Beschner (Ed.) *Youth drug abuse: Problems, issues and treatments* (pp. 629–656). Lexington, MA: Lexington Books.

Santisteban, D. A., Szapocznik, J., Perez-Vidal, A., Kurtines, W. M., Murray, E. J., & LaPerriere, A. (1996). Efficacy of interventions for engaging youth/families into treatment and some factors that may contribute to differential effectiveness. *Journal of Family Psychology, 10,* 35–44.

Schwartz, R. H., & Wirtz, P. W. (1990). Potential substance abuse detection among adolescent patients using the drug and alcohol (DAP) quick screen, a 30-item questionnaire. *Clinical Pediatrics, 29*(1), 38–43.

Scopetta, M. A., King, O. E., Szapocznik, J., & Tillman, W. (1979). *Ecological structural family therapy with Cuban immigrant families.* Report to the National Institute on Drug Abuse: Grant #H81DA 01696.

Sells, S. B., & Simpson, D. D. (1979). Evaluation of treatment outcome for youths in the Drug Abuse Reporting Program (DARP): A followup study. In G. M. Beschner (Ed.) *Youth drug abuse: Problems, issues and treatments* (pp. 571–628). Lexington, MA: Lexington Books.

Shedler, J., & Block, J. (1990). Adolescent drug use and psychological health: A longitudinal inquiry. *American Psychologist, 45,* 612–630.

Smith, T. E. (1983). Reducing adolescent's marijuana abuse. *Social Work in Health Care, 9,* 33–44.

Snow, D. L., Tebes, J. K., & Arthur, M. W. (1992). Panel attrition and external validity in adolescent substance use research. *Journal of Consulting and Clinical Psychology, 60,* 804–807.

Sobell, L. C., & Sobell, M. B. (1995). Alcohol consumption measures. In J. P. Allen & M. Columbus (Eds.), *Assessing alcohol problems: A guide for clinicians and researchers* (pp. 55–74). Rockville, MD: National Institutes of Health.

Stanton, M. D., & Shadish, W. R. (1997). Outcome attrition and family/couples treatment for drug abuse: A meta-analysis and review of the controlled, comparative studies. *Psychological Bulletin, 122,* 170–191.

Stanton, M. D., & Todd, T. C. (1982). *The family therapy of drug abuse and addiction.* New York: Guilford.

Stewart, D. G., & Brown, S. A. (1995). Withdrawal and dependency symptoms among adolescent alcohol and drug abusers. *Addiction, 90,* 627–635.

Substance Abuse and Mental Health Services Administration. (1996). *Preliminary estimates from the 1995 National Household Survey on Drug Abuse.* Washington, DC: US Government Printing Office.

Szapocznik, J., Kurtines, W. M., Foote, F. H., Perez-Vidal, A., & Hervis, O. (1983). Conjoint versus one-person family therapy: Some evidence for the effectiveness of conducting family therapy through one person. *Journal of Consulting and Clinical Psychology, 51,* 889–899.

Szapocznik, J., Kurtines, W. M., Foote, F. H., Perez-Vidal, A., & Hervis, O. (1986). Conjoint versus one-person family therapy: Further evidence for the effectiveness of conducting family therapy through one person with drug-abusing adolescents. *Journal of Consulting and Clinical Psychology, 54,* 395–397.

Szapocznik, J., Perez-Vidal, A., Brickman, A. L., Foote, F. H., Santisteban, D., Hervis, O, & Kurtines, W. M. (1988). Engaging adolescent drug abusers and their families in treatment: A strategic structural systems approach. *Journal of Consulting and Clinical Psychology, 56,* 552–557.

Tarter, R. E. (1990). Evaluation and treatment of adolescent substance abuse: A decision tree method. *American Journal of Drug and Alcohol Abuse, 16,* 1–46.

Tarter, R. E., & Hegedus, A. M. (1991). The Drug Use Screening Inventory: Its application in the evaluation and treatment of alcohol and other drug abuse. *Alcohol Health and Research World, 15*(1), 65–75.

Turner, R. M., Liddle, H., & Dakof, G. (1996, November). *Experimental evaluation of cognitive-behavior therapy for adolescent substance abuse.* Paper presented at the annual meeting of the Association for Advancement of Behavior Therapy, New York.

Vingilis, E., & Smart, R. G. (1981). Physical dependence on alcohol in youth. In Y. Israel, F. B. Gleser, & H. Kalant (Eds.), *Research advances in alcohol and drug problems* (Vol. 6, pp. 197–215). New York: Plenum.

Wagner, E. F., & Kassel, J. D. (1995). Substance use and abuse. In R. T. Ammerman, & M. Hersen (Eds.), *Handbook of child behavior therapy in the psychiatric setting* (pp. 367–388). New York: Wiley.

Waldron, H. B. (1996, November). *Developments in substance abuse assessment with adolescents*. Paper presented at the annual meeting of the Association for Advancement of Behavior Therapy, New York.

Waldron, H. B. (1997). Adolescent substance abuse and family therapy outcome: A review of randomized trials. In T. H. Ollendick & R. J. Prinz (Eds.), *Advances in clinical child psychology* (Vol. 19, pp. 199–234). New York: Plenum.

Waldron, H. B., & Slesnick, N. (1997, November).

Wanberg, K. W. (1991). *Adolescent Self-Assessment Profile*. Arvada, CO: Center for Alcohol/Drug Abuse Research and Evaluation.

Welte, J. W., & Barnes, G. M. (1985). Alcohol: The gateway to other drug use among secondary-school students. *Journal of Youth and Adolescence, 14,* 487–498.

Wheeler, K., & Malmquist, J. (1987). Treatment approaches in adolescent chemical dependency. *Chemical Dependency, 34,* 437–447.

White, H. R., & Labouvie, E. W. (1989). Towards the assessment of adolescent problem drinking. *Journal of Studies on Alcohol, 50,* 30–37.

Winters, K. C. (1992). Development of an adolescent alcohol and other drug abuse screening scale: Personal Experience Screening Questionnaire. *Addictive Behaviors, 17,* 479–490.

Winters, K. C., & Henly, G. A. (1988). Assessing adolescents who misuse chemicals: The Chemical Dependency Assessment Project. In E. R. Rahdert & J. Gradowski (Eds.), *Adolescent drug abuse: Analyses of treatment research* (pp. 4–18). Rockville, MD: National Institutes of Health.

Winters, K. C., & Henly, G. A. (1989). *Personal Experience Inventory Test and Manual.* Los Angeles, CA: Western Psychological Services.

Winters, K. C., & Henly, G. A. (1993). *Adolescent diagnostic interview schedule and manual.* Los Angeles: Western Psychological Services.

Winters, K. C., & Stinchfield, R. D. (1995). Current issues and future needs in the assessment of adolescent drug abuse. In E. Rahdert & D. Czechowicz (Eds.), *Adolescent drug abuse: Clinical assessment and therapeutic interventions* (pp. 146–171). Rockville, MD: National Institutes of Health.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.22
Elimination Disorders

C. EUGENE WALKER
University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

5.22.1 INTRODUCTION 566

5.22.2 ENURESIS 569
  5.22.2.1 Phenomenonology 569
  5.22.2.2 Diagnostic Features 569
  5.22.2.3 Prevalence 570
  5.22.2.4 Etiology 571
    5.22.2.4.1 Biological factors 571
    5.22.2.4.2 Emotional factors 572
    5.22.2.4.3 Learning factors 573
  5.22.2.5 Developmental Course 573
  5.22.2.6 Conceptualization and Clinical Formulation 574
    5.22.2.6.1 Multimodal assessment 574
    5.22.2.6.2 Treatment formulation 574
  5.22.2.7 Treatment 575
    5.22.2.7.1 Psychosocial treatments 576
    5.22.2.7.2 Pharmacological treatments 577
    5.22.2.7.3 Other treatments 577
  5.22.2.8 Effectiveness of Treatments 578
  5.22.2.9 Future Directions for Research and Practice 578

5.22.3 ENCOPRESIS 579
  5.22.3.1 Phenomenonology 579
  5.22.3.2 Diagnostic Features 579
  5.22.3.3 Prevalence 580
  5.22.3.4 Etiology 580
    5.22.3.4.1 Biological factors 581
    5.22.3.4.2 Emotional factors 581
    5.22.3.4.3 Learning factors 581
  5.22.3.5 Developmental Course 581
  5.22.3.6 Conceptualization and Clinical Formulation 581
    5.22.3.6.1 Multimodal assessment 581
    5.22.3.6.2 Treatment formulation 582
  5.22.3.7 Treatment 583
    5.22.3.7.1 Psychosocial treatments 583
    5.22.3.7.2 Pharmacological treatments 584
  5.22.3.8 Effectiveness of Treatments 584
  5.22.3.9 Future Directions for Research and Practice 585

5.22.4 SUMMARY 585

5.22.5 REFERENCES 586

5.22.1 INTRODUCTION

Toilet training is a ritual of considerable interest and concern to parents in all cultures, primitive and advanced (Hindley, Filliozat, Klackenberg, Nicolet-Meister, & Sand, 1965; Whiting & Child, 1964). Over the years, toilet training has been high on the list of concerns of parents with respect to the care of their young children (Ilg & Ames, 1962; Mesibov, Schroeder, & Wesson, 1977; Seim, 1989). It has generally ranked second or third in the list of concerns expressed by parents. The only area to consistently outrank toilet training has been the area of discipline and negative behavior in children.

Perusal of parents magazines and bookshelves with titles of interest to parents indicates a continuing and sustained interest in this area. Most parents spend a significant amount of time and effort in the process of toilet training and regard it as an important accomplishment when the process is complete. In many cultures the process is begun shortly after birth, even though physiological evidence would suggest that children do not have sufficient maturity or control of the muscles and nerves involved to be successful until approximately two years of age.

Although it may be argued that parents are more concerned about this process than they need be, proper toilet training does serve numerous functions. Most obviously, parents wish to have their children toilet trained as early as possible in order to avoid the unpleasant task of changing diapers and sheets. Almost equal in importance, however, is the idea held by many parents that early completion of toilet training indicates that they are good parents and that their child is superior. Beyond this, the importance of proper toileting to good hygiene and prevention of illness should be noted. Finally, many personality theorists point to the importance of toilet training as an early experience in the socialization of the child (Lieberman, 1972). Parents regard the achievement of toilet training as evidence of a major step in their child's progress towards independence and that their child is normal.

For children, toilet training is the first major developmental challenge they face that is more important to others than to them. The child suddenly discovers that his or her ability to control what was once an involuntary internal process is a matter of extreme interest, as well as social significance, to others, and that failures in this area merit the displeasure of parents and peers alike. Freud saw this process as fundamental to the child's development of respect for authority and social mores.

In spite of the importance parents ascribe to toilet training, they may take widely different views about how to accomplish the process. Some parents take the attitude that it is not necessary to directly toilet train children. Their view being that when the time comes, the child will naturally develop (presumably by example of other family members) the idea that he or she should care for eliminative processes and do so on his or her own. Others, however, believe that children will not learn unless they as parents take an active role in the process. As a result, a great deal of time and energy is invested.

One of the most elaborate procedures for toilet training has been outlined in the book *Toilet training in less than a day* by Azrin and Foxx (1976). In this very ingenious approach to toilet training, the child is first helped to toilet train a doll of the type that wets when squeezed. The child and the parent make a game of training the doll. The doll is given fluid after which the parent and child squeeze the doll to produce wetting accidents. If the wetting takes place in an area designated as the toilet, the doll is praised. If the doll wets in the pants, it is scolded mildly. The child helps the doll check the pants and go to the toilet properly. During the time the child is training the doll, the child is given large quantities of favorite beverages to drink in order to produce a full bladder and the need to urinate. Having trained the doll, the parent and child begin a game in which the child does what was demonstrated on the doll. Dry pants' inspections are made every five minutes and a reward of praise or a small treat is given if the pants are dry. Every 15 minutes the child participates in a potty trial in which an effort is made to urinate, whether the child actually has an urge to do so or not. Azrin and Foxx recommend that a child's potty chair be used for these trials since it is easier for the child to mount and sit on such a chair. In addition, they suggest using the type of potty chair that plays a tune when the child urinates. This alerts the parent and child to the fact that urination is occurring and serves as an added incentive to the child. Eventually, on one of the potty trials, the child will be able to urinate. The parent praises the child for this accomplishment and gives the child a small reward. The child is then encouraged to return to the potty chair at any time he or she feels a need to urinate.

At the beginning of the training, the parent assumes a major part of the responsibility by directing the child in the activities prescribed by the training program. The parent begins to fade involvement by reducing direct involvement until the parent is providing only minimal prompts and reminders that the pants should be dry and that the child should go to the potty if necessary. As the child begins to assume responsibility for the process, the potty practice trials are stopped and checks are made for dry pants every 15 minutes. However, if an accident occurs during this time, the parent shows mild verbal displeasure, has the child feel the wet pants, and then instructs the child to carry out a series of "positive practice trials." These trials are carried out by first having the child start from the place in the house where the accident occurred, walk rapidly to the toilet, remove clothing, and sit on the toilet. The child then does this same procedure starting from other points in the house until a total of 10 trials are completed. The child then cleans himself or herself and is given fresh clothing. The parent reminds the child that pants should be kept dry and wetting should be done in the toilet. This procedure continues for several hours (it can generally be accomplished in a long morning or afternoon if the parent and child are not distracted by other family members or other duties) until the child is using the toilet appropriately and independently while maintaining dry pants. Should accidents occur, following the treatment, these are handled by additional positive practice trials.

A couple of other features of this program are also interesting. First, parents are encouraged to give the child a written certificate at the end of the training to mark the accomplishment. Second, a list of friends and relatives, whom the child admires, is developed. During the training, the parent mentions various individuals from this list as individuals who will be happy when the child has learned and who wish the child success. After the child is successful in the training, the child is permitted to call some individuals from the list to report the accomplishment and receive positive feedback. According to Azrin and Foxx, this method is highly successful and most children can be trained in less than four hours. However, it should be pointed out that some studies of this method (Butler, 1976; Matson & Ollendick, 1977) have reported that most parents do better with this training if they have professional consultation in addition to the book.

A much simpler approach to toilet training was developed by the pediatrician Brazelton (1962) who has made numerous other contributions to our knowledge and understanding of the care of children. Brazelton developed this approach over 30 years ago but it is very sound advice and still has much use today. He recommends that the pediatrician begin a discussion of toilet training with the parents at the nine month visit. He notes that it is about this time that grandparents, friends, and relatives begin to put subtle pressure on the parents to see that the child is toilet trained. At this time, Brazelton simply discusses with the parents their attitudes about toilet training, attempts to reduce their anxiety, and encourages them to approach this in a relaxed, unpressured manner. He recommends that toilet training not begin at nine months. At around the age of 18 months, the first attempt is made to train the child.

The training begins by obtaining a small potty chair of the size used by children and identifying this as the child's own chair. An association is also made between this chair and the parents larger toilet chair. The parent then has the child sit on the chair with clothes on so that the chair will not feel cold or unfamiliar at the first sitting. The parent reads to the child and gives the child a small treat such as a cookie while the child sits on the chair. The child is not coerced and is free to leave the chair when they wish. After a week or so of getting the child familiar with sitting on the chair, the child is encouraged to sit on the chair without clothes or diapers. When the child is comfortable doing this, he or she is taken to the chair after a diaper has been soiled. The diaper is changed on the seat and the dirty diaper is dropped in to the pot under the child. The parent points out to the child that this is the eventual function of the chair. After doing this several times over a period of days, all diapers and pants are removed from the child for short periods of time when the parent thinks the child is likely to need to use the toilet. The toilet chair is placed in the room or play area near the child and the child is encouraged to use the potty chair. When the child follows through and uses the chair, the parent expresses approval and encourages the child to do this in the future. At this point, the child can be dressed in loose fitting training pants and taught to remove the training pants in order to use the potty chair when needed.

Nap and night training are not attempted until well after the child has mastered daytime control and shows an interest in staying clean and dry. This may coincide with achievement of daytime control or it may be as much as a year or two later. When the child shows an interest in having a dry bed, the parent can offer to help by waking him or her before the parent goes to bed so that the child can empty the bladder. In some cases, the parent may also have to wake the child a little earlier in the morning for toileting. The potty chair is placed beside the bed (Brazelton suggests painting the pot with luminous paint as an added incentive) and the child is encouraged to use the potty chair at any time during the night if needed. Brazelton stresses that these procedures should be carried out in a positive and supportive manner with the child. Where there are failures or problems, the parent is urged to stop the process and reassure the child that he or she is not bad for the failure. Optimism is expressed about success at some future time. Brazelton reported extremely good success with this method and, as a result, it has been widely used by pediatricians for many years.

The timing of toilet training is particularly important. As Walker (1978) has pointed out, toilet training is accomplished relatively easily when the child is ready to learn that skill. On the other hand, almost no amount of effort will produce the desired result if the child has not matured sufficiently and is not ready. Walker recommends attempting procedures similar to those outlined by Brazelton for a period of a week or two, when the child appears to be ready. If the learning takes place relatively effortlessly and with little power struggle, the process is complete and has been successful. However, if the child does not appear to be learning and if conflict is beginning to develop around the process, it is best to cease the training and to make another attempt after a period of about 3–6 months. This process is attempted every 3–6 months until the child is successful. It is extremely unwise to get into bitter power struggles and excessive conflict with the child over the process of toilet training. The disruption of the parent–child relationship is a much more serious problem than toilet training itself.

Signs of readiness that may be employed to determine when the first attempt at toilet training can be made have been outlined by Azrin and Foxx (1976). They suggest that the child be evaluated for physical readiness (does the child have enough finger and hand coordination to pick up objects easily and does the child walk easily from room to room without assistance) and instructional readiness (can the child point to the nose, eyes, mouth, or hair; can he or she follow directions to sit down on a chair, stand up, walk with you to a particular place, imitate you in a simple task such as playing patty cake, bring you a familiar object, and place one familiar object with another, for example, put the doll in the truck). A child is considered ready for toilet training when able to accomplish both physical readiness tasks and at least eight out of 10 instructional readiness tasks. Most children achieve this at around 20 months of age.

Brazelton suggests use of similar indices of physical maturation sufficient for toilet training and discusses what he calls "psychologic readiness," which is associated with a desire to control the impulses to defecate and urinate. This desire for control is derived from environmental influences including a secure and gratifying relationship with the parents, resulting in a desire to please them, a wish to identify with and imitate parents and other important adult figures, and the wish to develop autonomy and mastery of the primitive impulses to urinate and defecate. Physical and psychological readiness, according to Brazelton, is achieved by most children at around 18–30 months of age.

There is significant variation in readiness among children, due to the fact that children mature at different rates and often show growth spurts. In general, boys mature more slowly than girls. Individual differences as well as sex differences in developmental rates should be taken into account and children should not be hurried or rushed when they are not ready. Some of the claims of parents who assert that they have trained their children from birth or that their children achieved toilet training skill well before the age of two years should be taken with a grain of salt. Reports of early training are more likely situations in which it is the parent who has been trained to anticipate the child's elimination and to place the child on a potty chair. There is also an understandable exaggeration by parents of their child's accomplishments.

All of the above notwithstanding, the average age at which toilet training is completed varies somewhat from one culture to another. In some countries, there is considerable effort expended to toilet train and children appear to accomplish this goal at a somewhat younger age. Other cultures, which take a more relaxed approach, have statistics which indicate that successful toilet training occurs later in life. The same can be said for socioeconomic levels within a culture. Individuals of higher socioeconomic status generally spend a great deal more effort at toilet training their children and report younger ages for achievement than do individuals of lower socioeconomic status. Unfortunately, precise data on the amount of variability are not available. Therefore, although it is possible to influence the age of toilet training to an extent, by aggressive involvement with the child, the process can only be influenced within certain limits and it is questionable whether the stress and conflict involved are worth it. For example, Whiting and Child (1964), in their classic cross-cultural study of the effects of child rearing on personality, noted that of 22 societies compared there was a wide range of attitudes and expectations regarding toilet training. At one extreme are the Tamala of Madagascar where bowel training begins between the ages of two and three months and the child is expected to be continent by the age of six months. The other extreme is represented by the Dena of Africa where toilet training is not even begun until the child is almost five years of age. Most societies studied began toilet training somewhere between the ages 1.5 and 2.5 years. Those favoring early training found it necessary to use frequent and severe punishment. American culture tends to be mid-range in terms of the age at which toilet training is begun and in the severity of the measures used in training. In the early part of this century, relatively severe and early training practices were the mode. Currently, following the suggestions of Brazelton and Spock, more gentle and gradual approaches are endorsed by most parents.

Given the interest and concern of many parents in the toilet training process, it is not surprising that failures in this process result in numerous requests for professional help. Pediatricians, psychologists, and others who provide health care for children are frequently consulted for advice in treating children who are enuretic (unable to control the process of urination) or encopretic (unable to control bowel function). These two topics and certain related disorders will be considered in the remainder of this chapter.

5.22.2 ENURESIS

5.22.2.1 Phenomenonology

Children who experience difficulty with controlling urine have a number of personal and emotional reactions to this difficultly. First, they are often teased and ridiculed by their peers when the peers learn about it either by seeing a daytime accident or hearing about one that occurred at night. The child's parents generally show considerable anxiety about the fact that their child has not learned to control urine and, as a result, there is often a great deal of displeasure and even disgust directed toward the child. This frequently results in parent–child conflict and, in extreme cases, can even result in child abuse. For a child, it is often a mystery that all of a sudden so many people should be so concerned about a process that was virtually ignored previously. Thus, a two- or three-year-old child who is being toilet trained by the parents may wonder why there has been such a change in attitude and rules regarding this behavior. If bedwetting accidents continue into later childhood years, or if a child who has been dry for a while begins to wet again, he or she may be similarly mystified as to what is happening and why there should be so much social disgrace attached to it. Generally, the child feels that this process is uncontrollable and does not understand why people consider it a fault. Biology is on the side of the young child, at least, because the process of urination begins as a reflex in the infant and then must be controlled by the child as he or she develops. Nevertheless, the parental disapproval, peer ridicule, and seeming inability of the child to do anything about the problem results in considerable anxiety and loss of self esteem. George Orwell expressed this very well in his autobiography. He was a student at an English boarding school and records:

> It was looked upon as a disgusting crime which the child committed on purpose and for which the proper cure was a beating. For my part I did not need to be told it was a crime. Night after night I prayed with a fervor never previously attained in my prayers, "please God, do not let me wet my bed! Oh, please God, do not let me wet my bed!", but it made remarkably little difference. Some nights the thing happened, others not. There was no volition about it, no consciousness. You did not properly speaking *do* the deed: you merely woke up in the morning and found that the sheets were wringing wet. (Orwell, 1970 p. 379)

Orwell did eventually receive two severe beatings for this behavior at the hands of the principal of his school. After these beatings, he did not wet the bed again but the emotional scars remained.

In addition to the significant emotional consequences surrounding the problem of enuresis and the misguided attempts to correct it, there are other social consequences. For example, children often cannot participate in overnight parties, camping trips, and similar events due to the fact that they would possibly wet the bed. There are also significant problems when staying in hotels or motels or visiting relatives and friends, should the child wet the bed. As a result, children are often left out of significant activities because of their "problem."

DeGraaf, a chemical engineer, has described his frustrating experiences with enuresis and unsuccessful treatment in a paper entitled "Forty years of being treated for nocturnal enuresis" (1992). Eventually, desmopressin acetate (DDAVP) solved the problem in his case. Studies have demonstrated improvement in social relationships and self esteem following successful treatment of enuresis (Moffatt, 1989; Moffatt, Kato, & Pless, 1987).

5.22.2.2 Diagnostic Features

The 10th edition of the *International Classification of Diseases* (ICD-10) (World Health Organization [WHO], 1992a) contains the category "non-organic enuresis" (F 98.0) which is described as involuntary voiding of urine during the day or night which is abnormal in relation to the individual's mental age and not due to any organic or structural cause. This condition may have been present from birth (primary enuresis) or may develop following a period during which the child previously acquired bladder control (secondary enuresis). There may or may not be widespread emotional or behavioral disorder associated with this condition. The volume on *Clinical Description and Diagnostic Guidelines* discusses the possible relationship between the two and concludes that there is no simple way to decide if emotional problems result from, cause, or simply coexist with enuresis. The clinician is instructed to make a judgement in his or her own opinion for each case as to what constitutes the main problem. This same volume amplifies the age criteria by stating that there is no clear cut off, but generally the diagnosis would not be made unless the child has reached a chronological age of five years and a mental age of four years. The descriptors employed in this definition are rather vague and very general (the volume on research diagnostic criteria, 1993, is more precise and fairly similar to the American Psychiatric Association's *DSM-IV* [1994]). The *DSM-IV Diagnostic and statistical manual of mental disorders* contains the diagnostic category "307.6 Enuresis" (not due to a general medical condition). This diagnosis is defined much more precisely than that in *ICD-10* in that it requires: (i) there be repeated voiding of urine into bed or clothes, either involuntary or intentional; (ii) the behavior must be clinically significant in the sense that it is either of a frequency of at least twice a week for three consecutive months or the presence of the wetting causes significant distress or impairment in social, academic, or other important areas of functioning; (iii) the child is chronologically or developmentally at least five years of age; and (iv) the disorder is not due to a physiological effect from a substance (such as a medication) or from any general organic medical condition. In the *DSM-IV* system the diagnostician is instructed to specify the type which can be nocturnal only, diurnal only, or nocturnal and diurnal (*ICD-10* permits this but does not require it). The essence of both of these diagnoses revolves around the concept that, in the absence of organic disorder, the child is not controlling the process of urination when he or she should be able to do so (approximately five years of age) and is experiencing distress or impairment of functioning as a result of this condition. It is interesting to note that *ICD-10* mentions primary and secondary enuresis, whereas *DSM-IV* currently makes no mention of that distinction; however, *DSM-IV* is explicit about the need to specify whether wetting is day (diurnal) or night (nocturnal) or both. The decision by the writers of *DSM-IV* to specify "clinically significant" as a part of the definition is admirable, but their choice is certainly problematic at times. The requirement that wetting be at least twice a week for at least three consecutive months; or causing distress or impairment in important areas of functioning is a case of comparing apples with oranges. To start with, the frequencies chosen are fairly arbitrary and overly specific, while the term "distress or impairment" is overly vague. Strictly speaking, if the child wet only once (rather than twice) in any given week during the three month period, the child would not qualify as enuretic. On the other hand, if the child wet only once in a similar period but did so under circumstances that were very embarrassing or upsetting, the child could be diagnosed as enuretic.

5.22.2.3 Prevalence

There have been numerous studies of the prevalence of enuresis and the results vary quite widely from one source to another. The major reasons for these discrepencies appear to be differences in the definition of enuresis, differences in populations studied, sampling variations, and methods of data collection. In spite of somewhat conflicting results, there does appear to be some consensus regarding the rate of enuresis.

Physiologically, all children wet during the day and night at birth. By definition, enuresis is not diagnosed until the age of five. Although *DSM-IV* states that the rate for enuresis at age five is 7% for males and 3% for females, the most commonly reported estimate is that approximately 15–20% of five-year-old children are enuretic (Fergusson, Hons, Horwood, & Shannon, 1986; Oppel, Harper, & Rider, 1968). Rates steadily decline until, by the midteens, only approximately 1–2% of males and less than 1% of females would be diagnosable as enuretic. The spontaneous annual remission rate has been estimated at approximately 15% (deJonge, 1973; Forsythe & Redmond, 1974). In some instances, enuresis continues into adulthood. Studies of military recruits suggest that there is a 1–2% prevalence in young adult males (Plag, 1964). Boys are more likely to have problems with enuresis than girls. Some reports suggest that boys may be twice as likely to be involved as girls; however, these data frequently come from clinic-referred patients, and the higher rate of referral of boys for behavior problems may bias these statistics.

Studies in the community at large show some higher incidence in males but not remarkably higher (Fergusson et al., 1986; McGee, Makinson, Williams, Simpson, & Silva 1984; Oppel et al., 1968). Given that girls mature more rapidly than boys, a slightly higher rate in boys would be expected.

Diurnal enuresis is considerably less prevalent than nocturnal enuresis. For example, Hjalmas (1992a) found that only 3.1% of seven-year-old boys and 2.1% of seven-year-old girls experienced day wetting at least once per week. No doubt due to the fact that it is much less common, considerably less is known about diurnal enuresis. Some have speculated that diurnal enuresis may be more common in females than in males; that it may be more likely to occur in children who show signs of emotional disturbance; and may be more difficult to treat, especially if it occurs in older children (deJonge, 1973; Lovibond & Coote, 1970; Shaffer, 1973). Some clinicians are of the opinion that bedwetting in older teenagers and adults may frequently be related to prior child abuse, especially sexually abuse. Although only a small percentage of night-time wetters also wet during the day, a significant percentage (possibly as much as 50–60%) of those with daytime wetting problems also wet the bed at night (Hallgren, 1956). With respect to the primary vs. secondary distinction regarding enuresis, approximately 85% of all cases are considered primary. There is a tendency, as would be expected, for this to decrease with age. Secondary enuresis generally appears between the years of five and eight (APA, 1994). Enuresis is more prevalent in lower socioeconomic levels and in families where the mother has less education (Bakwin & Bakwin, 1972; MacKeith, Meadow, & Turner, 1973). Although there has not been sufficient research to clarify the reason for this, it is no doubt related to different attitudes regarding hygiene in general and toilet training in particular.

5.22.2.4 Etiology

There are three basic approaches to understanding the etiology of enuresis: biological factors, emotional factors, and learning. Each of these areas is briefly examined to determine what it can contribute to the overall understanding of the etiology of this problem.

5.22.2.4.1 Biological factors

It should be emphasized that urinary incontinence can be caused by numerous organic factors, ranging from anomalies of the nervous system, the bladder, or other structures of the urinary system to various types of bladder infections to secondary effects from chronic disease or other illnesses. However enuresis is by definition a functional disorder and, therefore, not the result of organic factors. Enuresis is functional urinary incontinence. Although specific organic pathology is explicitly ruled out of consideration with respect to the etiology of enuresis, various approaches to understanding the etiology presuppose some sort of general biological or physiological process underlying the problem. A wide range of factors have been proposed in this realm. Two that have a reasonable amount of support from the research literature are genetics and developmental delay. It has long been known that enuresis runs in families and tends to show the same general trends that are characteristic of disorders that are genetically determined (Bakwin, 1961; Elian, 1991; Fergusson et al., 1986).

Although there is fairly general support for the notion of a genetic influence in enuresis, there is very little sophisticated genetic research on this problem and no specific mechanism or gene has been determined to explain the inheritance of enuresis. For example, it is not possible to predict the inheritance of enuresis for specific individuals within a family. The most that can be said is that it runs in families.

Numerous clinicians and researchers over the years have suspected some sort of developmental delay as the basis for enuresis. Developmental delay is a very vague and nonspecific concept. In the present instance, it would no doubt refer to some delay in neurological development. Some studies have indicated that children who are enuretic are slow in achieving certain developmental milestones (MacKeith, 1972) but other studies have failed to replicate such findings. Developmental delays, when noted, are generally mild and it is questionable whether they would be sufficient to account for the presence of a condition such as enuresis, especially since physiologically the ability should be in place by 2–5 years of age.

Closely related to the developmental delay hypothesis has been the idea that a mild form of organic brain pathology may be involved in enuresis. The usual suspect is some sort of epileptic seizure during the night. Although it is true that individuals who have epileptic seizures often lose control of their urine, the majority of enuretic children do not suffer diagnosable seizure disorders and careful research has failed to produce significant pathology in terms of electroencephalograms (EEGs) in enuretic children (Mikkelsen, Brown, Minichiello, Millican, & Rapoport, 1982; Salmon, Taylor, & Lee, 1973).

A great deal of interest has centered recently on the possibility that the antidiuretic hormone (ADH) may be the cause of enuresis for some children. George, Messerli, and Genest (1975) have noted a circadian rhythm variation in the excretion of the ADH vasopressin in normal subjects in which there was an increase in secretion during the night. Other studies indicated that this increase did not occur in enuretic children (e.g., Rittig, Knudsun, Norgaard, Pedersen, & Djurhuus, 1989). Enuretic children with lower levels of ADH have been shown to produce larger quantities of dilute urine than normal children with the consequence that the probability of bed wetting is increased. However, later research has failed to confirm the earlier findings of a relatively direct path from low ADH levels to increased production of urine, polyuria and bedwetting (Evans & Meadow, 1992; Watanabe, Kawauchi, Kitamori, & Azuma, 1994). This approach seems to have some merit; however, considerably more research is needed before it will be fully understood.

Another possible factor of a biological nature in producing enuresis is what has been referred to as low functional bladder capacity. A child with low functional bladder capacity feels the need to urinate with lower amounts of urine present in the bladder than other children, and consequently voids smaller quantities of urine more frequently. Some investigators have reported smaller functional bladder capacity in enuretic children and have speculated that this limited bladder capacity results in night time wetting (Starfield, 1967; Zaleski, Gerrard, & Shokeir, 1973). However, there is evidence that the functional bladder capacities for enuretics and nonenuretics show considerable overlap (Rutter, 1973) and increases in functional bladder capacity have not been shown to be reliably related to improvement in enuretic children (Starfield & Mellits, 1968).

In addition, sleep disorder may be a biological factor involved in some cases of enuresis. Certainly, parents who present their children for treatment of enuresis almost uniformly report that their child is an exceedingly deep sleeper. Unfortunately, studies that have attempted to measure depth of sleep of bed wetters have failed to show consistent results (Mikkelsen et al., 1980). Although it is plausible to assume that some sort of sleep disorder is related to enuresis, no firm conclusion can be drawn until further research is done. Somewhat related to the sleep etiology of enuresis are various reports that have appeared in the literature, suggesting that sleep apnea may be related to the problem. For example, Weider and Hauri (1985), as well as Timms (1989), have reported that use of surgery or mechanical devices designed to improve breathing while asleep have resulted in reduction or cessation of bedwetting in a small number of children where this was a problem. Unfortunately, these reports, to date, are all anecdotal and lacking in scientific control. Even if further research were to substantiate this, it would account for only a small proportion of the total number of cases of enuresis.

Finally, food allergies have occasionally been implicated as a biological factor contributing to enuresis (e.g., Egger, Carter, Soothill, & Wilson, 1992). The idea that diet or food allergies may play some part in production of enuresis is certainly intriguing. Various foods are known to affect numerous other body systems; however, research demonstrating a link between food allergies or diet and enuresis, as well as an explanatory mechanism, is currently lacking. Clearly, more research is needed in this area. Should such factors be determined, they may be useful as a first approach to treatment for some cases and as a supplementary approach in others. Since most individuals overcome enuresis by treatment or spontaneous remission without any significant change in diet, it is almost certainly not primarily due to dietary factors.

5.22.2.4.2 Emotional factors

The idea that emotional factors may play a major role in the development of enuresis has a long history. In fact, in previous years many mental health workers were convinced that enuresis was synonymous with some sort of emotional disorder or neurotic conflict. Enuresis was often thought to be related to depression and that wetting was a type of "weeping through the bladder" (Imhof, 1956) or that it might be associated with some sort of sexual conflict in which the urination was symbolic of ejaculation or lubrication (Fenichel, 1946). Others thought that anger might be the major emotional factor and that the wetting was a type of passive aggressiveness against the mother (Soloman & Patch, 1969). Inappropriate or severe toilet training efforts were often implicated. It was frequently noted that a significant amount of family dysfunction, conflict, and other forms of pathology were present in cases where a child was enuretic. Clinicians and researchers thought they saw certain personality characteristics in enuretic children, most commonly immaturity. However, careful research has failed to support any of these notions. Most of these ideas appear to have resulted from a bias in referral patterns to mental health professionals vs. pediatricians. Most enuretic children are treated by pediatricians. Those who show signs of emotional disturbance are generally referred to mental health professionals. As a result, mental health professionals and their literature developed around the idea that some form of emotional disturbance was involved. Studies comparing large unselected samples of enuretic children vs. nonenuretic children failed to find significant relationships between emotional disturbance and enuresis, or to determine any enuretic personality profile (Cullen, 1966; Tapia, Jekel, & Domke, 1960).

The following points appear to be supported by the bulk of the literature and summarize the relationship between emotional disturbance and enuresis. First, the majority of enuretic children are not emotionally disturbed. Second, most emotionally disturbed children are not enuretic. Third, there is, however, a higher incidence of enuresis among emotionally disturbed children than there is among the general population. This area of overlap is relatively modest but has led to considerable misunderstanding. There is no specific diagnosis, emotional condition, or personality pattern associated with enuresis. Although clinical folklore suggests that secondary enuresis is more likely than primary enuresis to be associated with emotional disturbance, research has failed to confirm this (Fritz & Anders, 1979; Cho, 1984). Although data are minimal and somewhat conflicting, there may be a slight increase in the likelihood that emotional disturbance will be present along with enuresis if the enuretic is female, older, and if daytime wetting is present (Moilanen, Jarvelin, Vikevainen-Tervonen, & Huttunen, 1987; Rutter, Tizard, & Whitmore, 1970: Wagner, Smith, & Norris, 1988). Basically, enuresis and emotional disturbance must be thought of as separate categories, which can overlap. The overlap may be accounted for by the fact that emotional distress inhibits learning the necessary skills to control urine, or the distraction of emotional problems may result in a loss of performance of effective urine control techniques that were previously learned. The conflict and rejection suffered by enuretic children may also create emotional distress. At any rate, in most circumstances, the connection between the two is incidental rather than inherent.

5.22.2.4.3 Learning factors

The third and probably most useful explanation for the etiology of enuresis has to do with learning principles. This approach assumes that enuresis is a learning problem. The basic rationale of the approach is as follows. At birth, the process of urination is essentially reflexive. When a sufficient amount of urine has collected in the bladder, the bladder reflexively empties. Through experience, individuals learn to delay and inhibit this reflex for relatively long periods of time. However, even adults are often acutely aware of the fact that urination is basically reflexive and that they may not be able to inhibit this reflex indefinitely. At these times, finding a toilet facility is essential. During their developmental years, children are attempting to master the basic learning task of controlling this reflexive behavior. Enuretic children may be regarded as children who are experiencing difficulty in learning this behavior.

Children master many different tasks at different rates, and therefore it is not unexpected that some children will be delayed in developing this skill. That boys have more trouble with this than girls fits with the fact that males mature physically more slowly than females. Genetic factors, neurological factors such as developmental delay, and other biological factors may well play a part in predisposing the enuretic child to have difficulty learning this skill. The learning approach can also readily incorporate many features of the theory of an emotional basis for the etiology of enuresis. It is well known that emotionality interferes with learning or disrupts performance based on previous learning. Emotional distress may, thus, result in enuresis or delay the learning of control. In addition, family dysfunction or laxity in toilet training as well as conflict regarding failures in this respect could readily be seen as contributing to poor achievement. Ineffective or inappropriate toilet training methods would also be expected to result in poor learning. Thus, a learning theory explanation of the problem of enuresis fits best with the facts and encompasses all known information effectively. It also leads to effective treatment strategies, as described later in this chapter.

5.22.2.5 Developmental Course

The developmental course of enuresis is relatively easily described. Urinary incontinence is 100% at birth and steadily declines by approximately 15% per year until midteens when it reaches a level of around 1–2% which continues through adulthood. Most children achieve urinary continence between the ages of three and five with girls achieving this task somewhat more easily than boys. Those who do not learn urinary continence by age five or who lose this ability following a period of continence, are labeled enuretic.

5.22.2.6 Conceptualization and Clinical Formulation

5.22.2.6.1 Multimodal assessment

The assessment of enuretic children has been discussed in detail elsewhere (Walker & Shaw, 1988). The first step should always be a physical examination by a competent physician. Since enuresis is defined as a functional problem, the possibility of organic disease or etiology should be ruled out with care. Typically, a pediatrician or family physician would be able to accomplish this by relatively simple and noninvasive procedures involving an examination of the external genitalia, medical history, and urine analysis. Should a medical condition be diagnosed, this should be treated first. Frequently, after the medical condition is resolved, enuresis continues and may require further treatment. The family physician or pediatrician may do more extensive tests or, on rare occasions, refer the patient to a urologist for further testing. Medical tests that are sometimes, but rarely, required involve x ray, urodynamic testing, uroflowmetry, electromyography, cystometrography, urethral pressure profilimetry, and cystoscopy, for example. These tests are described in considerable detail by Maizels, Gandhi, Keating, and Rosenbaum (1993). In the majority of cases where a child is wetting, explanation for the problem will not reside in organic pathology. Nevertheless, when a medically diagnosable and treatable condition is present that should be tended to first.

Once general medical pathology has been ruled out or corrected, an in-depth psychological assessment is in order. This should begin with a careful history taken by interview that will include aspects of a typical intake interview to determine whether or not psychopathology of any sort is present in the child. Should there be suggestions of psychopathology, these should be explored through standard psychological testing and other assessment procedures. In some instances, where there is significant psychopathology, it is clinically advisable to treat the more serious pathology rather than attempting to treat enuresis in the presence of such pathology. The presence of severe pathology interferes with the treatment of enuresis and makes the outcome problematic. It is important to keep this in mind because practitioners who treat enuretic children often receive referrals from physicians that ostensibly are for the purpose of enuresis but, upon examination, reveal severe personal and/or family pathology that is the real reason for referral. Many physicians prefer to avoid confrontation with patients regarding their emotional difficulties and use enuresis as a token for referral.

If severe psychopathology is not present, treatment for enuresis may be undertaken. In cases where there is mild to moderate psychopathology present, a decision will need to be made as to whether the enuresis should be tackled early on or later in treatment. Often, achievement of success in control of urine enhances self-esteem and encourages progress of the child in other treatment. On the other hand, control of urine may not be possible until other problems are dealt with successfully.

Specific to the problem of enuresis, there should be a complete functional analysis including such things as onset of the problem, frequency, amount of urine expelled, number of times the child uses the bathroom during the day, whether or not urgency is present, reaction of the child and others in the family to the problem, discussion of past attempts at treatment (and their success or failure), family history of enuresis, effects on the child's functioning (especially socialization), power struggles that may be going on in the family regarding the problem, motivation of the child and family to overcome the problem, sleeping arrangements and habits of the child and the family, times at which the wetting accidents occur (day or night), whether or not child abuse (physical or sexual) has occurred, and other factors unique to the child that may play a part in the problem. Following this evaluation, the clinician may find it useful to use a scale developed by Morgan and Young (1975) which is designed to evaluate parental reactions to enuresis. In addition to a careful assessment of the motivation to participate in treatment of the problem, the clinician should carefully evaluate the acceptability of the various approaches to treatment that are described later in this chapter. Since there are a wide variety of approaches to the treatment of enuresis, it is wise to involve the parents and child in the selection of the treatment options.

5.22.2.6.2 Treatment formulation

As noted in the section on etiology, there is a wide range of possible factors that appear to have some relationship to enuresis. These factors involve everything from organic illness through biological/physiological predisposing factors through emotional difficulties to learning problems. Formulation of the case when dealing with an enuretic child will involve careful consideration of all possibilities and blending those factors that appear relevant into a treatment conceptualization and strategy. Depending on the particular blend of factors present, different treatment strategies may be expected to have more or less success and may be more or less acceptable to the child and family. As noted earlier, the general framework of a behavioral or learning theory approach has been found most useful by the majority of clinicians and readily incorporates factors from the other two major approaches to etiology. Thus, an overall strategy of viewing the problem as a failure to learn or to achieve competence in performing the task of controlling urine is clinically useful. The clinician can see his or her role as serving as consultant/coach to the child and family as they work on a resolution of the problem.

5.22.2.7 Treatment

Over the centuries, hundreds of different treatments have been attempted for the problem of enuresis. Glicklich (1951) has reviewed the history of the treatment of enuresis in what is now a classic article. From this review, we know that enuresis has been considered a problem since at least 1515 BC, when a medical text of that era recommended juniper berries, cypress, and beer as medication for enuresis. In subsequent times, numerous concoctions were prepared for the treatment of enuresis. Frequently, these concoctions were ground-up organs or tissue from animals such as the flesh of a hedgehog, the claws of a goat, or the trachea of a cock. At other times, they involved ingestion of a wide variety of chemicals, some of which may actually have been dangerous. Likewise, various substances were injected or applied to the sacral area. These applications and injections were theorized to stimulate the nerves in this area, thus producing continence. One of the more barbaric treatments involved burning blisters in the area. It was thought that the blistering would stimulate the sacral nerves. Cauterization of the opening of the urethra with silver nitrate was employed because it was thought that making this portion of the urethra more tender to the passage of urine would force the child to retain the urine. Occasionally, attempts were even made to seal the opening to the urethra with some substance during the night to prevent the passage of urine. Some thought that the warmth from the bed on the sacral nerves might cause bedwetting, or that the position of lying on the back was conducive to wetting. As a result, there were a number of instruments such as spikes or knotted ropes that were wrapped around the child to prevent the child from lying on their back during sleep.

Various other kinds of rotation of the body at different angles (such as elevating the feet, pelvis, or genitals) were also thought to be beneficial in preventing bedwetting. An apparatus popular during the nineteenth century involved insertion into the rectum of a metallic bung shaped like an elongated olive or small hens egg. Soldered to this bung was a piece of metal that simultaneously exerted pressure on the perineum and the coccygeal region of the spine. Whether a child could sleep comfortably with this device in place is open to serious question; however, it did appear to cut down on bedwetting. Various primitive tribes have employed a variety of measures for the treatment of enuresis, most involving punishment or ridicule. For example, one West African tribe pours a mixture of water and ashes on the head of the child. The child is then driven out into the street while other children of the village clap their hands and run after him or her singing "urine everywhere, urine everywhere." The Navajo indians were known to use a treatment during which the child stood naked with his or her legs spread over the burning nest of a Phoebe. This was thought to help because birds don't wet their nest. Mowrer and Mowrer (1938), in their classic article on the development of the pad-and-bell technique for treatment of enuresis, noted that the following treatments had been employed from time to time:

> innumerable drugs and hormones; special diets (including fresh fruit, caviar, and bacilli); restriction of fluids; voluntary exercises in urinary control; injections of physiological saline, sterile water, paraffin, and other inert substances; real and sham operations; passage of a bougie; public application of cantharides plasters; cauterization of the neck of the bladder; spinal punctures; tonsillectomy; circumcision, clitoridotomy, etc.; high-frequency mechanical vibration and electrical stimulation of various parts of the body; massage; bladder and rectal irrigations; Roentgen and other forms of irradiation; chemical neutralization of the urine; sealing or constriction of the urinary orifice; hydrotherapy; local freezing of the external genitalia with ice or chloratyl; elevation of the foot of the patients bed; sleeping on the back; not sleeping on the back; and use of a hard mattress.

Although some of these treatments may be around in one form or another today, most of them are completely without merit and are no longer used. In fact, some are dangerous. The admonition in the hippocratic oath to, "first do no harm," is very applicable to the case of enuresis. One must seriously consider the advantage of eliminating wet clothes or sheets, as opposed to the danger or risk of a treatment contemplated. Whether the danger is the possibility of physical harm to the child or emotional harm and disruption of the parent–child bond, the potential benefits of treatment may not warrant the risk. This is especially true for younger children since time is a definite ally in elimination of enuresis.

5.22.2.7.1 Psychosocial treatments

(i) Support and encouragement

Most books providing advice to parents and most textbooks for clinicians working with children suggest that the initial attempt to deal with enuresis involves talking with the child in an effort to give the child support and encouragement in his or her attempt to overcome the problem. Often, this will involve a brief medical examination so that the child can be reassured that there is nothing physically wrong or dangerous. The parents are advised to avoid conflict or punishment and to praise the child or provide some reward for success. This is often coupled with restriction of fluids, elimination of any drinks containing caffeine (coffee, tea, cola, soft drinks, etc.), and emptying the bladder just before retiring.

(ii) Periodic awaking

A number of different attempts have been made to use waking the child at night as a means of preventing bedwetting (Creer & Davis, 1975; Young, 1964). The simplest form of this intervention is for the parent to wake the child before the parent goes to bed, so that the child can urinate one more time before attempting to sleep till morning. In other versions, parents attempt to determine the time at which the child wets. The child is then woken 30 minutes prior to the anticipated time of the wetting. Over a period of time the waking time is moved in increments closer and closer to the actual bedtime of the child until the last voiding coincides with going to bed. Other versions employ the opposite strategy if the wetting typically occurs just before getting up in the morning. In these cases, the child is woken earlier than normal and prior to the wetting. The waking time is then gradually made later and later until it coincides with the normal time for getting up in the morning. Generally, waking is left up to the parents. However, some clinicians, especially with older children, employ an alarm clock which the child uses to wake up with.

(iii) Retention control and sphincter exercises

As noted earlier in this chapter, many children who have bedwetting problems appear to have a small functional bladder capacity. These children are noted to urinate frequently during the day and to have a great deal of urgency when they need to urinate (Sorotzkin, 1984; Starfield, 1967). The length of time the child is in bed and asleep exceeds the functional capacity of their bladder, resulting in bedwetting.

Consideration of this led to the development of two approaches to treatment that involve exercises (Miller, 1973). The first exercise involves urine retention. The child is instructed to go to the bathroom and to prepare to urinate, but retain urine as long as possible before beginning the urination process. The child is then encouraged to hold the urine longer each time the urge to urinate occurs. Eventually the child develops control to the point where it is measured in minutes. At this time, the child is allowed to leave the bathroom area and return at a later time to void. A delay of between 30 and 45 minutes is generally encouraged between the first urge to urinate and actual voiding. The approach is based on the rationale that urine retention exercises will increase the tolerance of the bladder for larger and larger volumes of urine.

Sphincter control exercises involve having the child start and stop the stream of urine once urination has begun. The sphincter control exercises are based on the rationale that these exercises increase the voluntary control of the sphincter, thus enabling the child to prevent urination.

(iv) Psychotherapy and family therapy

Use of psychotherapy and/or family therapy for the treatment of enuresis is based on the concept that some personality problem or family dysfunction may be the cause of the problem. As indicated earlier in this chapter there is little empirical evidence to support this position. Nevertheless, individual psychotherapy, family therapy, neurolinguistic reprogramming and use of metaphors have appeared in the literature as possible treatments (Crowley & Mills, 1986; Protinsky & Dillard, 1983; Rydzinski & Kaplan, 1985; Selig, 1982; Wood, 1988).

(v) Hypnosis

A variety of case reports and studies have appeared in the literature employing hypnosis in the treatment of enuresis (Banerjee, Srivastav, & Palan, 1993; Collison, 1970; Kohen, Olness, Corwell, & Heimel, 1980). Following hypnotic induction, the child is encouraged to visualize the urinary system and is given suggestions about holding back urine while asleep, or waking to use the toilet if not able to hold the urine back.

(vi) Cognitive interventions

With the increased interest in cognitive approaches to psychotherapy, attempts have been made to develop treatment programs for enuresis based on cognitive interventions. One such approach (Ronen & Wozner, 1995; Ronen, Wozner, & Rahav, 1992) involves four steps: modification of maladaptive concepts, understanding the enuresis process, increasing bladder control, and developing self-control. This appears to be a well designed protocol involving concepts common in a number of approaches to cognitive therapy. In addition, the bladder control phase involves retention control exercises.

(vii) Pad-and-bell treatment

Treatment of enuresis using a urine alarm was discovered by accident in the early part of the twentieth century by a German physician named Pfaunder. Pfaunder was working in a children's institution where the staff were instructed to change diapers promptly for children who were wet, in order to prevent skin irritation and infection. Pfaunder developed a device that was arranged so that when the child wet, a bell rang to signal to the staff that changing of the diaper was needed.

Pfaunder and the staff at the home were surprised to discover that under these conditions the number of wetting episodes decreased and in many cases ceased (Lovibond & Coote, 1970). Later, Mowrer and Mowrer (1938) experimented with variations of this procedure and demonstrated that it was highly effective in eliminating bedwetting in children. Following the work by the Mowrers, numerous devices were marketed that involved placing a pad in the bed of the child that was connected to a bell and/or light.

When the child urinated on to the pad, this completed an electrical circuit resulting in the ringing of a bell and/or turning on of a light to wake the child to go to the toilet.

Thousands of simple units of this sort were sold through the medical aids sections of major department store catalogs. In addition, numerous companies marketed devices through newspaper, magazine, and door-to-door sales efforts. Currently, a variety of simplified devices are available on the market that accomplish the same thing. These simplified devices involve sewing a small electrode into the undergarments of the child which responds when wet and activates a buzzer that is attached via a velcro strip to the shoulder of the child's pajamas or which is worn as a wristwatch. The buzzer wakes the child.

(viii) Multiple intervention package programs

A number of books and protocols have appeared in the literature that incorporate multiple components in the treatment of enuresis. These often involve such components as retention control and sphincter exercises, support and encouragement, pad-and-bell, positive practice, rewards, and periodic waking. Two of the most widely known and more carefully researched programs of this sort are dry-bed training (Azrin, Sneed, & Foxx, 1973) and full-spectrum home training (Houts, Liebert, & Padawer, 1983). An excellent description of the latter is available in *Bedwetting: A guide for parents and children* (Houts & Liebert, 1984). These programs are elaborate and require considerable training and supervision of the parents in order to complete them. Some work best when the treatment is carried out by trained assistants.

5.22.2.7.2 Pharmacological treatments

Although virtually every medication or chemical in existence has been administered to children in attempt to control bedwetting, there are currently only three medications that are used with any frequency (Miller, Atkin, & Moody, 1992). The most commonly prescribed medication is imipramine (Tofranil). The exact mechanism by which this medication is helpful in the treatment of enuresis is not clearly understood. A second medication often used is oxybutynin chloride (Ditropan). This drug reduces spasms of the bladder and appears to make the bladder less sensitive, resulting in an increase in functional bladder capacity. The third medication that has aroused a great deal of interest is desmopressin acetate (DDAVP). This medication is a synthetic form of vasopressin, which is a hormone that stimulates the kidneys to concentrate urine. Concentration of the urine reduces the amount of urine produced during the night, and appears to thereby enable some individuals to sleep through the night without bedwetting.

5.22.2.7.3 Other treatments

A wide variety of other approaches to treatment of enuresis have been employed and advocated at one time or another. For example, acupuncture has been used (Huo, 1988; Tuzuner, Kecik, Ozdemir, & Canakei, 1990; Xu, 1991). Chiropractic physicians have recommended spinal manipulations designed to reduce pressure on various nerves in order to prevent bedwetting (Gemmell & Jacobson, 1989). Dietary restrictions and manipulations have been attempted (Egger et al., 1992). Various surgical procedures, including dental surgery, have also been used (Timms, 1989; Weider & Hauri, 1985).

5.22.2.8 Effectiveness of Treatments

Since there is a spontaneous remission rate of approximately 15% each year and since virtually all people are able to control urination by the teenage years, it is often difficult to know whether a particular treatment has been successful or not. There have been thousands of studies in this area, however, and the following conclusions are justified. First, there is no doubt that the most effective form of treatment is the pad-and-bell (Djurhuus, Norgaard, Hjalmas, & Wille, 1992; Forsythe & Butler, 1989; Scott, Barclay, & Houts, 1992). Application of the pad-and-bell generally results in cessation of bedwetting for between 75% and 90% of cases over an 8–12 week period. Unfortunately, relapse rates are substantial and can be as high as 40% (Doleys, 1977). Reapplication of the procedure upon relapse is generally successful and in less time than the original course of treatment. Occasionally a third or fourth application may be necessary. Two other approaches that have been suggested to deal with the relapse problem are the use of overlearning (Young & Morgan, 1972) and intermittent schedules of reinforcement (Finley, Rainwater, & Johnson, 1982).

The various medications that have been employed have had only modest success. Although medications result in some improvement in most children, they often do not produce total cessation of the wetting and relapse rates are exceedingly high when medications are withdrawn. This, coupled with the ongoing expense and possibly dangerous side effects of the medications, makes these not the treatment of first choice except for short term interventions (Maizels et al., 1993; Thompson & Rey, 1995).

Other than a few anecdotal and case reports, psychotherapy and family therapy have not been shown to have significant effectiveness for enuresis (DeLeon & Mandell, 1966). The initial reports on cognitive psychotherapy indicate that this approach may have a higher success rate than more traditional approaches to psychotherapy (Ronen et al., 1992). Hypnosis has been shown to be effective when properly employed (Olness, 1975). Retention control and sphincter exercises initially showed good results but later research failed to confirm earlier findings (Doleys, Schwartz, & Ciminero, 1981; Harris & Purohit, 1977). Nevertheless, there is still interest in this approach and the sound physiological rationale would suggest that, properly employed, this treatment might be effective. Such approaches as restricting fluids, support and encouragement, awakening schedules, and so forth, are modestly effective with relatively simple cases. However, they are not sufficient for more difficult cases. Numerous reports have made excessive claims for the effectiveness of diet, acupuncture, chiropractic manipulation and so forth. Additional and more sophisticated research is needed on these approaches.

Probably the most effective approaches available are the multiple component programs by Azrin and Foxx, and Houts and his colleagues (Walker, 1995). These have much to recommend them in terms of careful research and demonstrated effectiveness. Their main drawback is the elaborate procedures required by the protocols and the amount of training, supervision, and effort required for successful execution of these programs. Some researchers doubt that the increase in efficiency of these programs over the simple use of a pad-and-bell justifies the amount of effort required.

5.22.2.9 Future Directions for Research and Practice

Walker (1995) has outlined a research agenda for enuresis. Given the amount of interest in this problem and the many centuries of effort to correct it, it is surprising how little is actually known about the process of developing urinary continence. There is currently very excellent and basic research underway in the UK by Fielding, (e.g. Fielding, 1982; Fielding, Berg, & Bell, 1978) and numerous excellent studies in Sweden and Denmark (Djrhuus, Norgaard, & Ritting, 1992; Hansson, 1992; Hellstrom, Hanson, Hansson, Hjalmas, & Jodal, 1990; Hjalmas, 1992b; Norgaard, Hansen, Nielsen, Rittig, & Djurhuus, 1989).

Walker (1995) notes that the various subtypes of enuresis have not been carefully evaluated with respect to etiology and treatment. Thus, we know that there is primary and secondary enuresis, diurnal and nocturnal enuresis, regular and intermittent enuresis, and so forth but we do not clearly understand the differences between these and whether there are treatment implications based on these differences.

Walker also notes that the most effective treatment which has been clearly demonstrated for several decades is some variation of the pad-and-bell. However, this is the least frequently employed treatment. This would appear to be an interesting research area in the sociology of science and medicine.

The third area Walker has outlined has to do with relapse rates for treatment of enuresis. All forms of treatment appear to have significant problems with relapse. More research is needed to develop ways of dealing with this problem.

Fourth, the relationship between bladder capacity and continence should be explored for possible treatment implications. Not all children who wet have small functional bladder capacity and, surprisingly, increases in functional capacity have not been shown to be directly related to improvement following treatment. Why this is so is not understood.

Fifth, the relationship between sleep and sleep patterns is not clearly understood with respect to enuresis. Although it is very expensive and time consuming to do research in this area, more work is indicated.

Sixth, the relative contributions of the components of the various package programs should be carefully evaluated with respect to their necessity. The package programs are successful but extremely elaborate and might be simplified with clarifying research.

Finally, more research is needed on the effects of diet on the etiology and management of enuresis. The same could be said of some of the alternative treatments, such as acupuncture and chiropractic manipulation. In some cases these are not known to be effective or ineffective because data are lacking. In particular, the cognitive and self-control approaches are deserving of more research.

5.22.3 ENCOPRESIS

5.22.3.1 Phenomenonology

The experience of the encopretic child is even more distressing than that of the enuretic child. This is for several reasons. First, the soiling accidents are more messy and more unpleasant in odor than are urination accidents. In addition, soiling accidents occur almost exclusively during the day, whereas wetting accidents are more likely to occur at night in one's own bedroom. As a result, the encopretic child lives in considerable terror that at any time he or she may have an accident that will result in an odoriferous mess as well as a great deal of embarrassment and ostracism by others.

The distress is heightened even more because, as discussed later in this chapter, the child generally is unable to control the encopretic episodes. They occur with very little warning and can not be managed successfully by the child without treatment. The child is frequently ridiculed by peers and referred to as "stinky" or some similar, even more derogatory, nickname. There is considerable conflict between parent and child because parents generally regard failure to control the bowels as a lack of motivation or slovenliness. Sometimes they even think it is an outright affront to the parents competence and authority. The perception, by the child, that all of his or her peers appear to be able to control bowels and that he or she is the only one unable to do so, makes the child feel considerably alienated from humanity and to feel totally incompetent at even basic life processes. This results in a significant lowering of self esteem and a feeling of defeatism. Many of these children feel that if they cannot control a simple thing such as bowel movements, so that they do not offend others around them, how can they control anything, be competent, or achieve in any area. Often, they are mystified as to why they cannot control their bowels and why others become so upset with them when the accidents occur.

In his book, *Worry: How to kick the serenity habit in 98 easy steps*, Robert Morley (1981) humorously notes that his worry in this area began with the requirement in the British school he attended that he and all the other students "write their name in the school register each morning after bowel evacuation."

Beyond these factors, bowel movements tend to be painful for children with encopresis; there is frequently irritation in the anal region from the fecal material; and the day-to-day living of the child may be marked by abdominal pain, gas, and other intestinal discomfort. Occasionally they are frightened by blood in the feces when hard stools scratch the colon and anus. One result of this is that these children often use denial and refuse to acknowledge that there is a problem. A second is that they may become aggressive and hostile toward those around them as a defense for the rejection that they feel from others. Children also frequently resort to subterfuge in hiding their soiled underclothes to avoid conflict with their parents.

James Roosevelt has provided an interesting insight into the problem of encopresis in his autobiographical writing:

> When I first got to Groton I had led so sheltered a life and was so shy that I was afraid to ask where the bathroom was. Hard as it may be to believe, I did what had to be done in my pants or in the bushes. My soiled pants were stuffed in a box, which I buried in the bushes. I suppose I was surrounded by foul smells. Holding back as long as I could, I started to suffer pains. Not knowing the cause, school officials suspected homesickness (well, I was homesick; at home I know where the bathrooms were) and asked my parents to take me home for a few days to ease my adjustment. (1976, pp. 60–61)

5.22.3.2 Diagnostic Features

DSM-IV defines encopresis as the repeated passage of feces into inappropriate places at least once per month for at least three months in a child who is chronologically four years of age (or developmentally the equivalent), in which the behavior is not due to the physiological effects of some substance such as a laxative or a medical illness. The soiling can be involuntary or intentional and the diagnostician has to determine whether or not constipation is present. When constipation is present the condition is referred to as overflow incontinence. The description of encopresis in *DSM-IV* is in very close agreement with the research data and clinical experience of those working in the field. Fecal incontinence is a general term that refers to a lack of control over bowel movements, whatever the cause. Encopresis is a term reserved for cases in which there is no organic basis other than constipation. Indicating that there is no organic basis other than constipation may be splitting hairs a bit; however, constipation is present in the vast majority of encopretic children and contributes to the problem, although it does not signal any serious medical disorder or illness. The constipation is generally behavioral in origin, involving diet and toilet habits. The requirement that the child be four years of age would seem to be unnecessary. Should a child become constipated at any age with resultant encopresis, it would appear to be a significant clinical problem worthy of treatment. This, of course, should be distinguished from the normal soiling or lack of toilet training in very young children (e.g., before the age of two). The *ICD-10* description of nonorganic encopresis describes the condition as repeated voluntary or involuntary passage of feces usually of normal or near normal consistency in places not appropriate for that purpose. *ICD-10* also notes that the condition may be a continuation of normal infantile incontinence beyond the age when bowel control should have developed; it may involve a loss of continence following successful toilet training; or, it may be a deliberate act of defecation in spite of the fact that the child is capable of bowel control. *ICD-10* notes that the symptom may be singular and isolated, or it may be a part of a wider disorder especially an emotional disorder or conduct disorder. The *ICD-10* description notes the possibility that constipation may be present but does not emphasize this as much as *DSM-IV*, and leans rather heavily toward the idea that an emotional disorder may be associated in many cases. Unfortunately, there is little evidence that emotional disturbance plays a major role in encopresis and considerable evidence that the great majority of the cases do involve constipation.

The discussion of encopresis in *ICD-10* also has other problems. For example, describing the feces as usually normal or near normal consistency does not square with the facts. Typically, a large number of the "accidents" consist of stains and rather pasty deposits or streaks on the underwear of these children caused by seepage around the impacted feces in the colon. These accidents are punctuated by large bowel movements of unusual diameter and/or by small, hard stools resembling rocks. Abnormal stools are the result of constipation.

5.22.3.3 Prevalence

As was the case with enuresis, estimates of the prevalence of encopresis vary widely, in large part due to differences in definition employed, populations studied, and methods of gathering data. There are currently no reliable data that indicate the differential rate of various forms of encopresis, such as retentive vs. nonretentive, primary vs. secondary, and so forth. By definition, functional encopresis does not begin before the age of four years. Overall estimates of the occurrence of encopresis range from 1.5–7.5% of children (Doleys et al., 1981). Some of the lower estimates that have been reported in the literature may be due to the fact that many parents are reluctant to discuss or report this problem with respect to their children. Clinicians in pediatric and mental health settings note that parents do not generally volunteer this information and it is obtained only by careful questioning (Walker, 1978). Boys are approximately four to five times more likely to be encopretic than girls (Bellmon, 1966; Rutter et al., 1970). Approximately 30% of encopretics are also found to be enuretic (Levine, 1975).

5.22.3.4 Etiology

Following an extensive review of the literature, Walker (1978) identified three subtypes of encopresis. First, he noted that there are children who can be described as "manipulative soilers." Children of this category soil when it is to their benefit to do so. Thus, some children soil if there is an examination to be given in school that they wish to avoid. The soiling makes it necessary for their parent to come to school and take them home for cleaning. On occasion, soiling appears to be related to anger or desire to distress the parent and represents a form of oppositional behavior. Or, there may be other environmental reasons that determine when the soiling occurs. Both *DSM-IV* and *ICD-10* note that soiling is sometimes intentional.

A second category of soiling is a form of diarrhea or loose bowels caused by excessive stress and expressed through the colon. A possibly related disorder to this is recurrent abdominal pain which may include diarrhea and is thought to be a precursor of irritable bowel syndrome in adults.

The third category is retentive encopresis based on chronic constipation. This is by far the most common etiology of encopresis. It is estimated that 80–90% of the cases fall into this category (Fitzgerald, 1975; Levine, 1975). As with enuresis, there are three general conceptualizations within which the etiology of encopresis may be considered: biological, emotional, and learning.

5.22.3.4.1 Biological factors

As the most common underlying feature of encopresis is constipation, it is quite likely that there are some underlying biological or physiological factors involved. First, some people have a higher constitutional likelihood for developing constipation; this is no doubt genetic (Petti & Davidson, 1991). Likewise, many individuals appear to be genetically predisposed to develop diarrhea under stress. It has also been speculated that some sort of subtle developmental delay may be involved in the development of encopresis (Bellmon, 1996).

5.22.3.4.2 Emotional factors

Early literature on encopresis often assumed a psychodynamic origin for this problem. The symptom was sometimes thought to be a result of an unconscious conflict, anger, depression, or some similar emotional factor. Encopretic children were frequently considered to have a particular personality, possibly involving immaturity, passive aggressiveness, anger, insecurity, anal fixation, anxiety, or low self esteem (Bemporad, Kresch, Asnes, & Wilson, 1978; Hoag, Norriss, Himeno, & Jacobs, 1971). Mothers were often depicted as rigid and masochistic, whereas fathers were thought to be weak, unaffectional, and uninvolved (Bemporad et al., 1978).

Other theories suggested that overly harsh or early attempts at toilet training may be the cause for encopresis (e.g. Hushka, 1942). Although there may be some merit to the idea that emotional factors play a role in the development of encopresis, most studies have failed to indicate such a connection (Fritz & Armbrust, 1982; Levine, 1975). It is probable that if such factors are involved, they are a secondary contributing factor rather than the major factor in most instances.

5.22.3.4.3 Learning factors

The third general rubric under which the development of encopresis can be understood has to do with learning theory. Obviously, manipulative soiling would follow a reinforcement model. If soiling is successful at manipulating the environment and results in reward for the child, the child will use this behavior in the future under appropriate circumstances. Likewise, chronic diarrhea and irritable bowel syndrome may be understood as symptoms of stress and anxiety which have been conditioned to selected stimuli. Finally, as noted, the most common form of encopresis is based on constipation. The development of constipation is based on behavioral factors such as poor dietary choices, failure to establish good toilet habits, withholding of stools, and similar behaviors which result in constipation.

It is likely that any of the three general categories outlined above (biological, emotional, and learning) may be instrumental in the development of encopresis. In many cases they may all be present to some degree or another. In other cases, one or more may be predominant and the main source of the problem. The most general and all encompassing of the three is the learning formulation because it readily incorporates the others and presents a complete description.

5.22.3.5 Developmental Course

The developmental course of encopresis is relatively simple to describe. At birth, fecal continence does not exist. The bowel reflexively empties when sufficient material is present and when the nerve endings in the rectum are stimulated by the presence of a full bowel. As the child develops, it becomes possible for toilet training to occur and most children develop bowel control relatively easily, prior to urinary control. Problems with encopresis tend to develop after control has been obtained and are generally based, as noted, on constipation which develops in the child. With or without treatment, most cases of encopresis cease by the midteen years. Encopresis persisting into adulthood does occur (Rex, Fitzgerald, & Goulet, 1992) but is rare. The reasons for this are not understood. That is, why would encopresis be a common childhood problem but disappear with or without treatment in the late teens? Whether this is related to improved toileting habits or bodily changes in the adult is not fully understood.

5.22.3.6 Conceptualization and Clinical Formulation

5.22.3.6.1 Multimodal assessment

The assessment of the encopretic child must begin with a medical examination. Typically, the physician will obtain a thorough medical history and do a basic physical examination including a rectal examination and palpation of the intestine for presence of fecal mass. The physician will also generally test for blood in the stool and, given the fact that females often have urinary tract infections if they are encopretic, may perform a urine analysis in females. There are a number of more invasive procedures, such as anorectal manometry, fecal studies, rectal biopsies, and barium enemas, which may be needed depending on what the physician obtains in the history and the initial examination. In the vast majority of cases these tests are not necessary. The basic physical examination can be performed by any competent physician familiar with child health. More advanced testing should be done by a pediatric gastroenterologist or other specialist. There are some organic pathologies that occur frequently enough that they must be ruled out before behavioral treatment is attempted. These include Hirschprung's disease (in which a segment of the colon has reduced, or a total absence of, neuralinnervation), congenital hypothyroidism, and pathology of the colon and surrounding structures such as anal stenosis, fissures, and so forth.

Psychological assessment must be broad and multifaceted (Walker & Shaw, 1988). First, there should be a screening assessment of the child's mental and emotional status. If significant psychopathology exists or significant family dysfunction is present, it may be necessary and advisable to deal with these prior to attempting treatment of the encopresis. Encopresis protocols tend to be relatively rigorous and demanding in terms of the amount of effort and the precision with which things must be done. If the child or the family is sufficiently disturbed that they are unlikely to be successful with the protocol, it is not appropriate to add the stress of this approach to the already existing stress. The existing difficulties should be dealt with followed by treatment for the encopresis. Should there be evidence of current psychopathology, standard psychological testing and assessment instruments may be needed to evaluate further the mental and emotional condition of the child or the degree of family disorganization.

The examination should include a thorough clinical interview of the parent and the child to determine the basic characteristics of the soiling in terms of frequency, timing, size of stools, staining of underwear, passage of blood during bowel movements, and other characteristics of toilet habits. In addition, the clinician would be interested in the manner and time at which toilet training progressed and whether or not the child ever achieved control of the bowel. The reaction of the child and the family to the problem and its effect on their interactions, as well as the motivation of each for treatment should be carefully assessed. The clinician should discuss, with the family, previous attempts to deal with the problem and their success or failure. Since diet is often involved in the development of constipation, there should be some exploration of eating habits. The ability and willingness of both the parents and the child to follow a carefully specified protocol are crucial to the success of treatment of encopresis. This should be thoroughly explored and emphasized in the assessment process.

Since the amount of conflict between parent and child regarding soiling is often significant and the family may be dysfunctional, the possibility of child abuse must be considered (Boon, 1991). The abuse may be physical, sexual, or psychological in nature. It is important to emphasize that the majority of children who are encopretic have not been abused (Clayden, 1988). However, in some cases where there has been abuse, the child will soil seemingly as a "cry for help." It is also possible in some cases, where there may be repeated anal rape, for there to be sufficient damage to the anus and musculature in that area that soiling results. This is not a frequent occurrence but is possible. It is important for the clinician to be alert to the possibility of child abuse in cases of encopresis, but it is equally important for the clinician to realize that abuse in these cases is definitely the exception, not the rule.

5.22.3.6.2 Treatment formulation

As noted in the previous section, there appear to be three basic types of conditions that result in soiling (Walker, 1978). One is the manipulative soiler who is using the soiling as an effort to manipulate the environment. This type of case would require a formulation involving family structure and dynamics.

The second category is the child who responds to stress with incontinence. Formulation in these instances will involve an assessment of the internal and external stressors that the child is experiencing in order to plan behavioral as well as other forms of psychotherapy to reduce stress, improve coping, and promote insight or self understanding. Environmental changes may also be required to reduce sources of stress. The overall formulation in these cases will include attempts to improve coping skills of the patient and to simultaneously reduce external stressors such that the problem will be resolved.

The most common etiology for soiling is constipation. The case formulation in these instances involves an evaluation of the factors that led to the constipation and an assessment of the willingness and ability of the child and family to follow the prescribed protocol. The way in which constipation contributes to the problem of encopresis must be understood. There are a number of circumstances through which a child can become constipated. Most of these have to do with diet and toileting behavior. For example, a diet that is lacking in roughage and/or that contains significant amounts of dairy products tends to lead to constipation. Likewise, withholding stools (which children at play often do when they do not wish to interrupt their play) can lead to encopresis. In addition, some individuals appear to have a physiological tendency to become constipated more easily than others. This is probably genetic in origin. There is some suggestion in the literature that certain emotional reactions such as depression may reduce motility of the intestine and eventually produce constipation. The amount of liquid consumed and the amount of dehydration that children experience while playing in warm weather may also be a factor. Lack of exercise is known to be associated with constipation. A number of other factors might also be mentioned, but the main point is that there can be many reasons for the child to become constipated. Unfortunately, when the child becomes constipated, this is frequently not observed by either the child or the parent. The child simply does not take notice of it and the parent assumes that the child has used the toilet at some time when they were not around. Whereas wetting accidents are readily observable, constipation in the early stages is not. Eventually the constipation becomes chronic and results in severe impaction in the colon. This severe impaction and chronic constipation result in stretching the colon which is referred to as psychogenic megacolon. When the colon is stretched and impacted, it ceases to function normally. The peristalic movement that moves material through the intestine and expels it through the anus is muted and unable to perform the task of evacuating the bowel. With this impaction, there is chronic pressure on the anus which results in a loss of the sensation that the child needs to evacuate. Thus, these children frequently report that they do not have the urge to evacuate. Another consequence of chronic impaction is that the liquid material in the intestine that has been emptied from the stomach makes its way only to the point of impaction and is not able to proceed further. As a result, it remains there in a pool. As the child plays or moves around, the wall of the intestine is momentarily separated from the impacted feces allowing a small space through which the liquid material in the pool seeps gradually around the impacted mass and eventually leaks out the anus. This results in the staining of underwear that is very common with encopretic children and which they claim not to be able to predict.

Encopretic children are correct when they say that they did not know it was happening until it was too late, because there was no sensation of contraction and evacuation. The seepage is a passive process. Material simply leaks out the anus without the child having any warning or control over it. In addition to the staining, large bowel movements occur in the clothing or in the toilet when significant amounts of the impacted material are expelled under pressure. These can be extremely embarrassing and upsetting to a child when they happen in public, and at home they are occasionally even large enough that the plumbing facilities cannot accommodate the diameter of the stools. When this happens stools have to be broken into smaller pieces in order to be flushed down the toilet.

These facts are a major part of the formulation and should be communicated to the parents. That is, they should be made aware of the fact that the child has an enlarged colon which is simply not functioning properly and cannot be controlled in the normal way. They should further be informed that the child has lost the sensation of needing to defecate and that the seepage and staining that occurs happens without warning to the child because it is a passive process. Explaining this to the parents often reduces their anger and their disgust because they previously believed that the child was simply defying their authority and refusing to try.

5.22.3.7 Treatment

5.22.3.7.1 Psychosocial treatments

Depending on the formulation of the case, treatment will vary. In cases of manipulative or intentional soiling, the most appropriate treatment would appear to be behavioral psychotherapy to assist the child in coping more effectively with the environment without inappropriate manipulative behavior (Keat, 1979; Walker, Hedberg, Clement, & Wright, 1981). Family therapy may also be very effective in dealing with behavior within the family system that is causing or supporting use of manipulation via soiling (Henggeler & Borduin, 1990).

In cases where diarrhea may be the origin of the soiling, psychotherapy, especially behaviorally oriented psychotherapy, may be helpful in reducing stress and increasing coping skills (Cohen & Reed, 1968; Hedberg, 1973). Anxiety reducing and antidiarrheal medications may also be used (Angelides & Fitzgerald, 1981).

Since the majority of cases of encopresis are based on constipation, the remainder of the remarks in this section are devoted to treatment of this form of encopresis. There have been occasional reports of the treatment of such children with traditional psychotherapy (McTaggert & Scott, 1959; Pinkerton, 1958) as well as family therapy and related approaches (Knights & Pandey, 1990; Margolies & Gilstein, 1983–1984; Wells & Hinkle, 1990; White, 1984).

Numerous case studies and research reports based on adequate samples have reported on the treatment of encopresis using behavioral strategies (Doleys, 1989; Houts & Abramson, 1990; Liebert & Fischel, 1990). Most of these approaches deal with a combination of behavioral principles based on reinforcement and/or punishment along with dietary changes, toileting skills development, changes in toilet habits, and often use of laxatives or enemas for cleansing the colon (Wright & Walker, 1977). Although there is wide variation in the exact components employed in these treatment approaches, a typical approach might involve cleansing the colon via laxatives on the first day of the treatment program, changing the child's diet to make it high in fiber content, and providing rewards for successful toileting attempts as well as for having clean underwear at the end of the day.

Several reports can be found in the literature on the use of biofeedback for bowel control problems (e.g., Whitehead, 1992). Biofeedback procedures are frequently used to assess the functioning of sphincters in children experiencing problems with encopresis. In addition, biofeedback equipment may be used to train a child in proper control of the muscles employed in elimination. Three components are generally employed in these programs: exercise of external sphincter muscles, training in discrimination of rectal sensations, and training in synchronization of the internal and external sphincter muscles so that proper elimination can occur (Latimer, Campbell, & Kasperski, 1984).

There have been reports of the use of hypnosis in the treatment of encopresis. Most of the reports (Olness, 1976; Tilton, 1980) indicate that suggestions are given to the child about the nature of the activity of the colon and of the child's ability to control the process. Finally, there have been attempts to use acupuncture and related procedures in the treatment of encopresis (An et al., 1986).

5.22.3.7.2 Pharmacological treatments

The typical treatment by physicians for constipation-based encopresis involves a regime in which the bowel is evacuated and thoroughly cleared. This is done through enemas and laxatives. The regime is then continued for a period of weeks or months, following which the laxative is either abruptly ceased or phased out. Most commonly, mineral oil (1–3 teaspoons, 2–3 times a day) is employed as the laxative (Abrahamian & Lloyd-Still 1984; Behrman, Vaughan, & Nelson, 1987; Levine, Carey, Crocker, & Gross, 1983). The rationale for this treatment is that the severe impaction has produced psychogenic megacolon, thus compromising the functioning of the colon. Treatment begins with a thorough evacuation of the bowel which is generally readily accomplished by laxatives and enemas; however, there have been occasions when impacted material had to be manually removed prior to beginning treatment. Next, daily use of mineral oil as a laxative produces regularity of bowel movements and produces stools that are easily passed. Maintaining a clear colon over a period of time allows the colon to shrink back to its normal shape, re-establish muscle tone, and begin more normal peristalic movement. When the laxative regime is removed, regular bowel movements frequently continue.

5.22.3.8 Effectiveness of Treatments

Medical and behavioral treatments are the most effective treatments for encopresis. In most reports it is found that 80% or more of children become continent with medical protocols employing mineral oil or other laxatives (Levine & Bakow, 1976; Sondheimer & Gervaise, 1982). Unfortunately, there is a significant rate of relapse following these treatments. Abrahamian and Lloyd-Still (1984) did a five-year follow-up study of their patients and found that 47% were completely symptom free, while another 36% found it necessary to use laxatives intermittently to control the symptoms.

Stark, Spirito, Lewis, and Hart (1990) found that mineral oil oriented programs were unsuccessful with a substantial number of children, especially those with behavioral problems. Referral to a psychologist for treatment in these cases appears to be warranted (Stark, Owens-Stively, Spirito, Lewis, & Guevremont, 1990).

Behavioral treatments for encopresis have been reported by a large number of people in both case studies and controlled research. Very high success rates have been reported for these programs (Doleys, 1989; Houts & Abramson, 1990; Liebert & Fischel, 1990; Walker, 1995).

Although follow-up studies have not been as frequent or as extensive as might be wished, the behaviorally oriented programs do not seem to have as bad a record of relapse as do those based solely on mineral oil or laxatives. The difference may be the greater emphasis on skills development in terms of bowel control, toilet habits, and related areas that are included in the behaviorally oriented programs. Most of the behaviorally oriented programs include laxatives and dietary manipulation, but add a significant dimension of behavioral psychotherapy and habit change.

Biofeedback training is sometimes helpful in bowel training and remediation of encopresis. It is particularly helpful in cases where an organic problem underlies the difficulty or in cases where simple behavioral programs are not effective due to underlying organic or physiological impairments (Whitehead, 1992). For example, biofeedback training has been found to be effective with patients following rectal surgery that impaired the functioning of the anal sphincters. It does not appear to improve success rates in the more routine and uncomplicated cases, however (Loening-Baucke, 1995).

Although there have been intriguing reports about the effectiveness of hypnosis and family therapy in the treatment of encopresis, there is not sufficient data to indicate that these are likely to be successful as a primary or singular treatment for the disorder. They would appear to be effective as adjuncts to an overall treatment program involving laxatives, diet manipulation, and careful toilet training. The same could be said for traditional psychotherapy (Achenbach & Lewis, 1971; Berg & Jones, 1964). Evidence for acupuncture and other related approaches is simply not sufficient to determine whether they are effective or not. It is, however, unlikely that they are.

5.22.3.9 Future Directions for Research and Practice

Walker (1995) proposed a research agenda for encopresis. He noted that, compared to enuresis, relatively little research is done on encopresis. Thus, there is a need for more information on virtually every point regarding encopresis. There is a need for considerably more research on the effects of diet and different foods on the process of digestion and elimination. For example, are there foods that are likely to produce constipation other than dairy products? Additionally, are there foods that promote regularity beyond the simple effects of fiber content?

The information on encopresis related to deliberate or manipulative soiling is very limited. Given the nature of this problem, it is difficult to get a group of patients together at one time for study. However, single case study methods and small group studies can be done and there is a need for considerably more work regarding children with encopresis of this sort. Likewise, whereas there is a considerable amount of information about the effects of stress producing diarrhea, there is relatively little known about precise treatments for this.

Areas involving associated features of encopresis need much more delineation, such as socioeconomic factors, ethnic factors, possible genetic influences, emotional status, and family status of encopretic children. None of these are well understood and better understanding might prove valuable in terms of planning treatment.

Wide variability in the size, composition, odor, and color of stools produced by encopretic as well as normal children have been noted. More information about the relationship between these variables and the function of the colon would be valuable.

With respect to various programs that have successfully employed a combination of laxatives, enemas, diet manipulation, and behavioral training, there has been very little research on the importance of the different components in the package and of possible variations in the components based on presenting symptoms, developmental age, or other factors related to the child.

There are a number of cases, that every clinician in this area has encountered, that seem to be resistant to all efforts employed by a variety of professionals. Careful study of some of these cases on an inpatient basis might reveal the source of the recalcitrance.

Finally, there are a number of related problems to encopresis that have little or no research to date, such as smearing of faces, coprophagia, and coprophilia.

Increased sophistication in treatment protocols can be accomplished only as more data are accumulated in the areas outlined above. Were more information available, it should readily be possible to treat children in a shorter amount of time, with fewer relapses and fewer failures.

5.22.4 SUMMARY

Toilet training is an area of great interest and concern to parents. Parents often regard the development of bowel and bladder control in their children as an important developmental milestone for the child and an affirmation of the parents' competence. Unfortunately, more emphasis is often placed on this process than is warranted. Experts generally recommend assessing the readiness of the child for toilet training and then employing a gentle and patient approach based on encouragement and support. Beginning the process at too early an age and employing excessively harsh approaches can often be harmful to the child.

Failure to achieve proper control of the process of urination is referred to as enuresis. Failure in bowel control is termed encopresis. Both enuresis and encopresis are, by definition, functional disorders of incontinence. Thus, organic bases such as specific disease processes must be ruled out prior to a diagnosis of enuresis or encopresis. A variety of biological and emotional factors have been proposed as causitive or contributing factors to the development of enuresis. However, the disorder can best be understood within the framework of learning theory. Development of urinary continence is a process that involves learning to inhibit reflexive emptying of the bladder, and of emptying the bladder at the appropriate time and place. Biological and emotional factors may inhibit the learning process or interfere with performance of skills already learned; however, they are best regarded as contributory factors. The most successful treatments are use of the pad-and-bell (which appears to be effective due to conditioning) and a variety of package programs which employ the pad-and-bell along with other learning and training experiences for the child. Psychotherapy and a variety of other approaches to the treatment of enuresis have not proved sufficiently successful to be regarded as useful. Medication is successful in some cases but has not been demonstrated to be as effective as learning approaches to the treatment of enuresis and suffers from the problem of side effects as well as the ongoing expense of the treatment.

Although numerous biological and emotional factors have been suggested as causes of encopresis, the most common cause is related to constipation and the resultant psychogenic megacolon. Effective treatment requires thorough evacuation of the bowel which permits the tissue to return to its normal shape, tone, and function. Along with restoration of normal functioning of the tissue, retraining of toilet functions is required. Numerous protocols are in the literature both in the pediatric medical literature and in the psychological literature which describe treatment for encopresis. These involve the use of laxatives or enemas for evacuation of the colon, dietary changes to enhance bowel function, and retraining of toileting skills. These programs have been found to be highly effective. Psychotherapy and other forms of treatment are generally ineffective for treatment of encopresis.

5.22.5 REFERENCES

Abrahamian, F. P., & Lloyd-Still, J. D. (1984). Chronic constipation in childhood: A longitudinal study of 186 patients. *Journal of Pediatric Gastroenterology and Nutrition, 3,* 460–467.

Achenbach, T. M., & Lewis, M. (1971). A proposed model for clinical research and its application for encopresis and enuresis. *Journal of the American Academy of Child Psychiatry, 10,* 535–554.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

An, X. C., Zhoo, H. S., Li, X. C., Tang, Y. Y., Sun, J., Liu, J., Yi, J. Q., Yu, Y. X., Zhang, G. C., & Hao, G. (1986). Acupuncture treatment for disturbances in urination and defecation from sacral cryptorachischisis. *Journal of Traditional Chinese Medicine, 6*(2), 95–98.

Angelides, A., & Fitzgerald, J. F. (1981). Pharmacologic advances in the treatment of gastrointestinal diseases. *Pediatric Clinics of North America, 28,* 95–112.

Azrin, N. H., & Foxx, R. M. (1976). *Toilet training in less than a day*. New York: Pocket Books.

Azrin, N. H., Sneed, T. J., & Foxx, R. M. (1973). Dry bed: A rapid method of eliminating bedwetting (enuresis) of the retarded. *Behaviour Research and Therapy, 11,* 427–434.

Bakwin, H. (1961). Enuresis in children. *Journal of Pediatrics, 58,* 806–819.

Bakwin, H., & Bakwin, R. M. (1972). *Behavior disorders in children*. Philadelphia: Saunders.

Banerjee, S., Srivastav, A., & Palan, B. M. (1993, October). Hypnosis and self-hypnosis in the management of nocturnal enuresis: A comparative study with imipramine therapy. *American Journal of Clinical Hypnosis, 36*(2), 113–119.

Behrman, R. E., Vaughan, V. C., & Nelson, W. E. (1987). *Textbook of pediatrics* (13th ed.). Philadelphia: Saunders.

Bellmon, M. (1966). Studies on encopresis. *Acta Paediatrica Scandinavica, 170*(Suppl.), 1–137.

Bemporad, J. R., Kresch, R. A., Asnes, R., & Wilson, A. (1978). Chronic neurotic encopresis as a paradigm of a multifactorial psychiatric disorder. *Journal of Nervous and Mental Disease, 166,* 472–479.

Berg, I., & Jones, K. V. (1964). Functional faecal incontinence in children. *Archives of Disease in Childhood, 39,* 465–472.

Boon, F. (1991, May). Encopresis and sexual assault. *Journal of the American Academy of Child Adolescent Psychiatry, 30*(3), 509–510.

Brazelton, T. B. (1962). A child-oriented approach to toilet training. *Pediatrics, 29,* 121–128.

Butler, J. F. (1976). The toilet training success of parents after reading "Toilet Training in Less Than A Day." *Behavior Therapy, 7,* 185–191.

Cho, S. C. (1984). Clinical study on childhood enuresis. *Seoul Journal of Medicine, 25,* 599–608.

Clayden, G. S. (1988). Reflex anal dilatation associated with severe chronic constipation in children. *Archives of Disease in Childhood, 63,* 832–836.

Cohen, S. I., & Read, J. L. (1968). The treatment of "nervous diarrhea" and other conditioned autonomic disorders by desensitization. *British Journal of Psychiatry, 114,* 1275–1280.

Collison, D. L. (1970). Hypnotherapy in the management of nocturnal enuresis. *Medical Journal of Australia, 1,* 52–54.

Creer, T. L., & Davis, M. H. (1975). Using a staggered-wakening procedure with enuretic children in an institutional setting. *Journal of Behavior Therapy and Experimental Psychiatry, 6,* 23–25.

Crowley, R. J., & Mills, J. C. (1986). The nature and construction of therapeutic metaphors for children. *British Journal of Experimental and Clinical Hypnosis, 3*(2), 69–76.

Cullen, K. J. (1966). Clinical observations concerning behavior disorders in children. *Medical Journal of Australia, 1,* 712–715.

DeGraaf, M. J. M. (1992, October 17). 40 years of being treated for nocturnal enuresis. *The Lancet, 340,* 957–958.

deJonge, G. A. (1973). Epidemiology of enuresis: A survey of the literature. In I. Kolvin, R. C. MacKeith, & S. R. Meadow (Eds.), *Bladder control and enuresis*. Philadelphia: Lippincott.

DeLeon, G., & Mandell, W. (1966). A comparison of conditioning and psychotherapy in the treatment of functional enuresis. *Journal of Clinical Psychology, 22,* 326–330.

Djurhuus, J. C., Norgaard, J. P., Hjalmas, K., & Wille, S. (Eds.) (1992). Nocturnal enuresis. *Scandinavian Journal of Urology and Nephrology, 143*(Suppl.), 3–29.

Djurhuus, J. C., Norgaard, J. P., & Ritting, S. (1992). Monosymptomatic bedwetting. *Scandinavian Journal of Urology and Nephrology, 141*(Suppl.), 7–17.

Doleys, D. M. (1977). Behavioral treatments for nocturnal enuresis in children: A review of the recent literature. *Psychological Bulletin, 84*(1), 30–54.

Doleys, D. M. (1989). Enuresis and encopresis. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of child psychopathology* (2nd ed., pp. 291–314). New York: Plenum.

Doleys, D. M., Schwartz, M. S., & Ciminero, A. R. (1981). Elimination problems: Enuresis and encopresis. In E. J. Mash & L. G. Terdal (Eds.), *Behavioral assessment of childhood disorders* (pp. 679–710). New York: Guilford Press.

Egger, J., Carter, C. H., Soothill, J. F., & Wilson, J. (1992). Effect of diet treatment on enuresis in children with migraine or hyperkinetic behavior. *Clinical Pediatrics, 31*(5), 302–307.

Elian, M. (1991). Treating bed wetting. *British Medical Journal, 302*(6778), 729.

Evans, J. H. C., & Meadow, S. R. (1992). Desmopressin for bed wetting: Length of treatment, vasopressin secretion, and response. *Archives of Disease in Childhood, 67,* 184–188.

Fenichel, O. (1946). *The psychoanalytictic theory of neurosis*. London: Routledge & Kegan Paul.

Fergusson, D. M., Hons, B. A., Horwood, L. J., & Shannon, F. T. (1986). Factors related to the age of attainment of nocturnal bladder control: An eight year longitudinal study. *Pediatrics, 78*(5), 884–890.

Fielding, D. (1982). An analysis of the behavior of day- and night-wetting children: Towards a model of micturition control. *Behaviour Research and Therapy, 20,* 49–60.

Fielding, D. M., Berg, I., & Bell, S. (1978). An observational study of postures and limb movements of children who wet by day and at night. *Developmental Medicine and Child Neurology, 20,* 453–461.

Finley, W. W., Rainwater, A. J., & Johnson, G. (1982). Effect of varying alarm schedules on acquisition and relapse parameters in the conditioning treatment of enuresis. *Behaviour Research and Therapy, 20,* 69–80.

Fitzgerald, J. F. (1975). Encopresis, soiling, constipation: What's to be done? *Pediatrics, 56,* 348–349.

Forsythe, W. I., & Butler, R. J. (1989). Fifty years of enuretic alarms. *Archives of Disease in Childhood, 64,* 879–885.

Forsythe, W. I., & Redmond, A. (1974). Enuresis and spontaneous cure rate: Study of 1129 enuretics. *Archives of Disease in Childhood, 49,* 259–263.

Fritz, G. K., & Anders, T. F. (1979). Enuresis: The clinical application of an etiologically-based classification system. *Child Psychiatry of Human Development, 10,* 103–113.

Fritz, G. K., & Armbrust, J. (1982). Enuresis and encopresis. *Pediatric Clinics of North America, 5,* 283–296.

Gemmell, H. A., & Jacobson, B. H. (1989). Chiropractic management of enuresis: Time-series descriptive design. *Journal of Manipulative and Physiological Therapeutics, 12*(5), 386–389.

Glicklich, L. B. (1951). An historical account of enuresis. *Pediatrics, 8,* 859–876.

George, C. P, Messerli, F. H., Genest, J. et al. (1975). Diurnal variation of plasma vasopressin in man. *Journal of Clinical Endocrinology Metabolism, 41,* 332–338.

Hallgren, B. (1956). Enuresis: I. A study with reference to morbidity, risk and symptomatology. II. A study with reference to certain physical, mental and social factors possibly associated with enuresis. *Acta Psychiatrica Scandinavica, 31,* 379–436.

Hansson, S. (1992). Urinary incontinence in children and associated problems. *Scandinavian Journal of Urology and Nephrology, 141*(Suppl.), 47–54.

Harris, L. S., & Purohit, A. P. (1977). Bladder training and enuresis: A controlled trial. *Behaviour Research and Therapy, 15,* 485–490.

Hedberg, A. G. (1973). The treatment of chronic diarrhea by systematic desensitization: A case report. *Journal of Behavior Therapy and Experimental Psychiatry, 4,* 67–68.

Hellstrom, A. L., Hanson, E., Hansson, S., Hjalmas, K., & Jodal, U. (1990). Micturition habits and incontinence in 7-year-old Swedish school entrants. *European Journal of Pediatrics, 149,* 434–437.

Henggeler, S. W., & Borduin, C. M. (1990). *Family therapy and beyond.* Pacific Grove, CA: Brooks/Cole.

Hindley, C. B., Filliozat, A. M., Klackenberg, G., Nicolet-Meister, D., & Sand, E. A. (1965). Some differences in infant feeding and elimination training in five European longitudinal samples. *Journal of Child Psychology and Psychiatry, 6,* 179–201.

Hjalmas, K. (1992a). Functional daytime incontinence: Definitions and epidemiology. *Scandinavian Journal of Urology and Nephrology, 141*(Suppl.), 39–44.

Hjalmas, K. (1992b). Urinary incontinence in children: Suggestions for definitions and terminology. *Scandinavian Journal of Urology and Nephrology, 141* (Suppl.), 1–6.

Hoag, J. M., Norriss, N. G., Himeno, E. T., & Jacobs, J. (1971). The encopretic child and his family. *Journal of the American Academy of Child Psychiatry, 10,* 242–256.

Houts, A. C., & Abramson, H. (1990). Assessment and treatment for functional childhood enuresis and encopresis: Toward a partnership between health psychologists and physicians. In S. B. Morgan & T. M. Okwumabua (Eds.), *Child and adolescent disorders: Developmental and health psychology perspectives* (pp. 47–103). Hillsdale, NJ: Erlbaum.

Houts, A. C., & Liebert, R. M. (1984). *Bedwetting: A guide for parents and children.* Springfield, IL: Charles C. Thomas.

Houts, A. C., Liebert, R. M., & Padawer, W. (1983). A delivery system for the treatment of primary enuresis. *Journal of Abnormal Psychology, 11,* 513–520.

Huo, J. S. (1988). Treatment of 11 cases of chronic enuresis by acupuncture and massage. *Journal of Traditional Chinese Medicine, 8*(3), 195–196.

Hushka, M. (1942). The child's response to coercive bowel training. *Psychosomatic Medicine, 4,* 301–308.

Imhof, B. (1956). Bettwasser in der erziehungsberatung. *Heilpaedagogische Werkblaetter, 25,* 122–127.

Ilg, F. L., & Ames, L. B. (1962). *Parents ask*. New York: Harper & Brothers.

Keat, D. B. (1979). *Multimodal therapy with children*. Elmsford, NY: Pergamon.

Knights, B., & Pandey, S. K. (1990). Paradoxical approach to the management of faecal incontinence in normal children. *Archives of Disease in Childhood, 65*(6), 598–600.

Kohen, D., Olness, K., Corwell, S., & Heimel, A. (1980, November). *Five hundred pediatric behavioral problems treated with hypnotherapy*. Paper presented at the annual meeting of the American Society of Clinical Hypnosis, Minneapolis, MN.

Latimer, P. R., Campbell, D., & Kasperski, J. (1984). A components analysis of biofeedback in the treatment of fecal incontinence. *Biofeedback and Self Regulation, 9,* 311–324.

Levine, M. D. (1975). Children with encopresis: A descriptive analysis. *Pediatrics, 56*(3), 412–416.

Levine, M. D., & Bakow, H. (1976). Children with encopresis: A study of treatment outcome. *Pediatrics, 58,* 845–852.

Levine, M. D., Carey, W. B., Crocker, A. C., & Gross, R. T. (1983). *Developmental—behavioral pediatrics*. Philadelphia: Saunders.

Lieberman, L. (1972). The changing ideology of socialization: Toilet training, mass media, and society. *International Journal of Contemporary Sociology, 9*(4), 188–199.

Liebert, R. M., & Fischel, J. E. (1990). The elimination disorders—enuresis and encopresis. In M. Lewis & S. M. Miller (Eds.), *Handbook of developmental psychopathology: Perspective in developmental psychology* (pp. 421–429). New York: Plenum.

Loening-Baucke, V. (1995, July). Biofeedback treatment for chronic constipation and encopresis in childhood: Long-term outcome. *Pediatrics, 96*(1), 105–110.

Lovibond, S. H., & Coote, M. A. (1970). Enuresis. In C.G. Costello (Ed.), *Symptoms of psychopathology: A handbook* (pp. 373–396). New York: Wiley.

MacKeith, R. C. (1972). Is maturation delay a frequent factor in the origins of primary nocturnal enuresis? *Developmental Medicine and Child Neurology, 14,* 217–223.

MacKeith, R. C., Meadow, R., & Turner, R. K. (1973). How children become dry. In I. Kolvin, R. C. MacKeith, & S. R. Meadow (Eds.), *Bladder control and enuresis* (pp. 3–21). Philadelphia: Lippincott.

Maizels, M., Gandhi, K., Keating, B., & Rosenbaum, D. (1993, November/December). Diagnosis and treatment for children who cannot control urination. *Current Problems in Pediatrics, 23,* 402–450.

Margolies, R., & Gilstein, K. W. (1983–1984). A systems approach to the treatment of chronic encopresis. *International Journal of Psychiatry in Medicine, 13,* 141–152.

Matson, J. L., & Ollendick, T. H. (1977). Issues in toilet training normal children. *Behavioral Therapy, 8,* 549–553.

McGee, R., Makinson, T., Williams, S., Simpson, A., & Silva, P. A. (1984). A longitudinal study of enuresis from five to nine years. *Australian Paediatric Journal, 20,* 39–42

McTaggert, A., & Scott, M. (1959). A review of twelve cases of encopresis. *Journal of Pediatrics, 54,* 762–768.

Mesibov, G. B., Schroeder, C. S., & Wesson, L. (1977). Parental concerns about their children. *Journal of Pediatric Psychology, 2,* 13–17.

Mikkelsen, E. J., Brown, G. L., Minichiello, M. D., Millican, F. K., & Rapoport, J. L. (1982). Neurologic status in hyperactive, enuretic, encopretic and normal boys. *Journal of the American Academy of Child Psychiatry, 21,* 75–81.

Mikkelsen, E. J., Rapoport, J. L., Nee, L., Gruenau, C., Mendelson, W., & Gillin, J. C. (1980). Childhood enuresis: I. Sleep patterns and psychopathology. *Archives of General Psychiatry, 37,* 1139–1144.

Miller, P. M. (1973). An experimental analysis of retention control training in the treatment of nocturnal enuresis in two institutionalized adolescents. *Behavior Therapy, 4,* 288–294.

Miller, K., Atkin, B., & Moody, M. L. (1992). Drug therapy for nocturnal enuresis: Current treatment recommendations. *Practical Therapeutics, 44*(1), 47–56.

Moffatt, M. E. K. (1989). Nocturnal enuresis: Psychologic implications of treatment and nontreatment. *The Journal of Pediatrics, 114*(4), 697–704.

Moffatt, M. E. K., Kato, C., & Pless, I. B. (1987). Improvements in self-concept after treatment of nocturnal enuresis: Randomized controlled trial. *The Journal of Pediatrics, 110*(4), 647–652.

Moilanen, I., Jarvelin, M. R., Vikevainen-Tervonen, L., & Huttunen, N. P. (1987). Personality and family characteristics of enuretic children. *Psychiatria Fennica, 18,* 53–61.

Morgan, R. T. T., & Young, G. C. (1975). Case histories and shorter communications. *Behaviour Research and Therapy, 13,* 197–199.

Morley, R. (1981). *Worry: How to kick the serenity habit in 98 easy steps*. New York: G. P. Putnam's Sons.

Mowrer, O. H., & Mowrer, W. M. (1938). Enuresis: A method for its study and treatment. *American Journal of Orthopsychiatry, 8,* 436–459.

Norgaard, J. P., Hansen, J. H., Nielsen, J. B., Rittig, S., & Djurhuus, J. C. (1989). Nocturnal studies in enuresis: A polygraphic study of sleep EEG and bladder activity. *Scandinavian Journal of Urology and Nephrology, 125*(Suppl.), 73–78.

Olness, K. (1975). The use of self-hypnosis in the treatment of childhood nocturnal enuresis. *Clinical Pediatrics, 14,* 273–279.

Olness, K. (1976). Autohypnosis in functional megacolon in children. *American Journal of Clinical Hypnosis, 19,* 28–32.

Oppel, W. C., Harper, P. A., & Rider, R. V. (1968). The age of attaining bladder control. *Pediatrics, 42,* 614–626.

Orwell, G. (1970). Such, such were the joys. In S. Orwell & I. Angus (Eds.), *The collected essays, journalism and letters of George Orwell: Vol. 4 In front of your nose, 1945–1950* (p. 379). Washington, DC: Penguin.

Petti, M. J., & Davidson, M. (1991). Idiopathic constipation. In W. A. Walker, P. R. Durie, J. R. Hamilton, J. H. Walker-Smith, & J. B. Watkins (Eds.), *Pediatric gastrointestinal disease: Vol. I Pathophysiology, diagnosis, management* (pp. 818–829). Philadelphia: Decker.

Pinkerton, P. (1958). Psychogenic megacolon in children: The implications of bowel negativism. *Archives of Disease in Childhood, 33,* 371–380.

Plag, J. A. (1964). *The problem of enuresis in the naval service*. US Navy Medical Neuropsychiatric Research Unit, Report 64–3. Washington, DC: US Government Printing Office.

Protinsky, H., & Dillard, C. (1983). Enuresis: A family therapy model. *Psychotherapy: Theory, Research and Practice, 22*(1), 81–89.

Rex, D. K., Fitzgerald, J. F., & Goulet, R. J. (1992). Chronic constipation with encopresis persisting beyond 15 years of age. *Diseases of the Colon and Rectum, 35*(3), 242–244.

Rittig, S., Knudsen, U. B., Norgaard, J. P., Pedersen, E. B., & Djurhuus, J. C. (1989). Abnormal diurnal rhythm of plasma vasopressin and urinary output in patients with enuresis. *American Journal of Physiology, 256,* F664–F671.

Ronen, T., & Wozner, Y. (1995). A self-control intervention package for the treatment of primary nocturnal enuresis. *Child & Family Behavior Therapy, 17*(1), 1–21.

Ronen, T., Wozner, Y., & Rahav, G. (1992). Cognitive intervention in enuresis. *Child & Family Behavior Therapy, 14*(2), 1–14.

Roosevelt, J. (1976). *My parents: A differing view*. Chicago: Playboy Press.

Rutter, M. (1973). Indication for research. In I. Kolvin, E. C. MacKeith, & S. R. Meadow (Eds.), *Bladder control and enuresis*. Philadelphia: Lippincott.

Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, health and behaviour*. London: Longmans, Green.

Rydzinski, J. W., & Kaplan, S. L. (1985). A wolf in sheep's clothing? Simultaneous use of structural family therapy and behavior modification in a case of encopresis and enuresis. *Hillside Journal of Clinical Psychiatry, 7*(1), 71–81.

Salmon, M. A., Taylor, D. C., & Lee, D. (1973). On the EEG in enuresis. In I. Kolvin, R. C. MacKeith, & S. R. Meadow (Eds.), *Bladder control and enuresis*. Philadelphia: Lippincott.

Scott, M. A., Barclay, D. R., & Houts, A. C. (1992). Childhood enuresis: Etiology, assessment, and current behavioral treatment. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification* (Vol. 28, pp. 83–117). Pacific Grove: CA: Sycamore.

Seim, H. C. (1989). Toilet training in first children. *Journal of Family Practice, 29,* 633–636.

Selig, A. L. (1982). Treating nocturnal enuresis in one session of family therapy: A case study. *Journal of Clinical Child Psychology, 11*(3), 234–237.

Shaffer, D. (1973). The association between enuresis and emotional disorder: A review of the literature. In I. Kolvin, R. C. MacKeith, & S. R. Meadow, (Eds.), *Bladder control and enuresis*. Philadelphia: Lippincott.

Solomon, P., & Patch, V. D. (1969). *Handbook of psychiatry*. Los Altos, CA: Lange.

Sondheimer, J. M., & Gervaise, E. P. (1982). Lubricant versus laxative in the treatment of chronic functional constipation of children: A comparative study. *Journal of Pediatric Gastroenterology and Nutrition, 1*(2), 223–226.

Sorotzkin, B. (1984). Nocturnal enuresis: Current perspectives. *Clinical Psychology Review, 4,* 293–315.

Starfield, B. (1967). Functional bladder capacity in enuretic and nonenuretic children. *Journal of Pediatrics, 70,* 777–781.

Starfield, B., & Mellits, E. D. (1968). Increase in functional bladder capacity and improvements in enuresis. *Journal of Pediatrics, 72,* 483–487.

Stark, L. J., Owens-Stively, J., Spirito, A., Lewis, A., & Guevremont, D. (1990). Group behavioral treatment of retentive encopresis. *Journal of Pediatric Psychology, 15*(5), 659–671.

Stark, L. J., Spirito, A., Lewis, A. V., & Hart, K. J. (1990). Encopresis: Behavioral parameters associated with children who fail medical management. *Child Psychiatry and Human Development, 20*(3), 169–179.

Tapia, F., Jekel, J., & Domke, H. R. (1960). Enuresis: An emotional symptom? *Journal of Nervous and Mental Disease, 130,* 61–66.

Thompson, S., & Rey, J. M. (1995). Functional enuresis: Desmopressin the answer? *Journal of the American Academy of Child Adolescent Psychiatry, 34*(3), 266–271.

Tilton, P. (1980). Hypnotic treatment of a child with thumbsucking, enuresis and encopresis. *American Journal of Clinical Hypnosis, 22,* 238–240.

Timms, D. J. (1989). Rapid maxillary expansion in the treatment of nocturnal enuresis. *Angle Orthodontist, 60*(3), 229–239.

Tuzuner, F., Kecik, Y., Ozdemir, S., & Canakei, N. (1990). Electro-acupuncture in the treatment of enuresis nocturna. *Acupuncture and Electro-Therapeutics Research International Journal, 14,* 211–215.

Wagner, W. G., Smith, D., & Norris, W. R. (1988). The psychological adjustment of enuretic children: A comparison of two types. *Journal of Pediatric Psychology, 13*(1), 33–38.

Walker, C. E. (1978). Toilet training, enuresis, and encopresis. In P. Magrab (Ed.), *Psychological management of pediatric problems* (Vol. 1, pp. 129–189). Baltimore: University Park Press.

Walker, C. E. (1995). Elimination disorders: Enuresis and encopresis. In M. C. Roberts (Ed.), *Handbook of pediatric psychology* (2nd ed.). New York: Guilford.

Walker, C. E., Hedberg, A. G., Clement, P. W., & Wright, L. (1981). *Clinical procedures for behavior therapy*. Englewood Cliffs, NJ: Prentice-Hall.

Walker, C. E., & Shaw, W. (1988). Assessment of eating and elimination disorders. In P. Karoly (Ed.), *Handbook of child health assessment* (pp. 434–469). New York: Wiley.

Watanabe, H., Kawauchi, A., Kitamori, T., & Azuma, Y. (1994). Treatment system for nocturnal enuresis according to an original classification system. *European Urology, 25,* 43–50.

Weider, D. J., & Hauri, P. J. (1985). Nocturnal enuresis in children with upper airway obstruction. *International Journal of Pediatric Otorhinolaryngology, 9,* 173–182.

Wells, M. E., & Hinkle, J. S. (1990). Elimination of childhood encopresis: A family systems approach. *Journal of Mental Health Counseling, 12*(4), 520–526.

White, M. (1984). Pseudo-encopresis: From avalanche to victory, from vicious to virtuous cycles. *Family Systems Medicine, 2*(2), 150–160.

Whitehead, W. E. (1992). Biofeedback treatment of gastrointestinal disorders. *Biofeedback and Self-Regulation, 17*(1), 59–76.

Whiting, J. W. M., & Child, I. L. (1964). *Child training and personality: A cross-cultural study*. New Haven, CT: Yale University Press.

Wood, A. (1988). King tiger and the roaring tummies: A novel way of helping young children and their families change. *Journal of Family Therapy, 10,* 49–63.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical description and diagnostic guidelines*. Geneva, Switzerland: Author.

Wright, L., & Walker, C. E. (1977). Treatment of the child with psychogenic encopresis: An effective program of therapy. *Clinical Pediatrics, 16,* 1042–1045.

Xu, B. (1991). 302 cases of enuresis treated with acupuncture. *Journal of Traditional Chinese Medicine, 11*(2), 121–122.

Young, G. C. (1964). A staggered-wakening procedure in the treatment of enuresis. *Medical Officer, 111,* 142–143.

Young, G. C., & Morgan, R. T. T. (1972). Overlearning in the conditioning treatment of enuresis: A longterm follow-up study. *Behaviour Research and Therapy, 10,* 419–420.

Zaleski, A., Gerrard, J. W., & Shokeir, M. H. K. (1973). Nocturnal enuresis: The importance of a small bladder capacity. In I. Kolvin, R. C. MacKeith, & S. R. Meadow (Eds.), *Bladder control and enuresis*. Philadelphia: Lippincott.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.23
Eating Disorders

DAVID M. GARNER
Toledo Center for Eating Disorders, Bowling Green State University, and University of Toledo, OH, USA
and
LINDA E. MYERHOLTZ
Bowling Green State University, OH, USA

5.23.1 INTRODUCTION 592

5.23.2 DIAGNOSTIC AND DEFINING FEATURES 592

- *5.23.2.1 Phenomonology* 592
- *5.23.2.2 Diagnosis* 593
  - *5.23.2.2.1 Anorexia nervosa* 593
  - *5.23.2.2.2 Bulimia nervosa* 593
  - *5.23.2.2.3 Eating disorders not otherwise specified* 594
  - *5.23.2.2.4 Relationship between different eating disorders* 594
  - *5.23.2.2.5 Differential diagnosis* 594

5.23.3 EPIDEMIOLOGY 594

- *5.23.3.1 Incidence and Prevalence* 594
- *5.23.3.2 Gender Considerations* 595
- *5.23.3.3 Developmental Considerations* 595
- *5.23.3.4 Have Eating Disorders Changed Over Time?* 596

5.23.4 COMORBIDITY 596

- *5.23.4.1 Depression* 597
- *5.23.4.2 Anxiety* 597
- *5.23.4.3 Personality Disorders* 598
- *5.23.4.4 Sexual Abuse* 598

5.23.5 ETIOLOGY 599

- *5.23.5.1 Genetic Contributions* 599
- *5.23.5.2 Developmental Theory* 599
- *5.23.5.3 Familial Factors* 601
- *5.23.5.4 The Addiction Model* 601
- *5.23.5.5 Cultural Factors* 602

5.23.6 MULTIMODAL ASSESSMENT 602

- *5.23.6.1 Targets of Assessment* 603
  - *5.23.6.1.1 Body weight history* 603
  - *5.23.6.1.2 Binge eating* 603
  - *5.23.6.1.3 Extreme weight losing behaviors* 603
  - *5.23.6.1.4 Psychopathology related to weight or shape* 603
  - *5.23.6.1.5 General psychopathology relevant to eating disorders* 604
  - *5.23.6.1.6 Psychopathology secondary to starvation* 604
- *5.23.6.2 Assessment methods* 604
  - *5.23.6.2.1 Clinical interview* 604

5.23.6.2.2 Self-report measures 604
5.23.7 CLINICAL DECISION MAKING IN TREATMENT 605
5.23.7.1 Integration and Sequencing of Treatments 605
5.23.7.2 Hospitalization and Day Treatment 605
5.23.7.3 Outpatient Medical Management 607
5.23.7.4 Self-help and Guided Self-help 607
5.23.7.5 Psychoeducation 608
5.23.7.6 Family Therapy 608
5.23.7.7 Cognitive Behavioral Therapy 609
5.23.7.8 Interpersonal Therapy 610
5.23.7.9 Psychodynamic Therapy 610
5.23.7.10 Pharmacotherapy 611
5.23.8 PSYCHOLOGICAL TREATMENT RESEARCH 611
5.23.8.1 Treatment Research on Anorexia Nervosa 611
5.23.8.2 Treatment Research on Bulimia Nervosa 613
5.23.9 PSYCHOPHARMACOLOGICAL TREATMENT RESEARCH 618
5.23.9.1 Treatment Research on Anorexia Nervosa 618
5.23.9.2 Treatment Research on Bulimia Nervosa 619
5.23.10 FUTURE DIRECTIONS 621
5.23.11 REFERENCES 621

5.23.1 INTRODUCTION

Although eating disorders have been recognized as part of psychiatric nomenclature for many years, it has only been in since the mid-1980s that they have commanded interest in mainstream psychology. Undoubtedly, among the factors responsible for the growing attention shown by researchers and clinicians alike is that eating disorders provide a prototypic model of the complex interaction between cultural, somatic, and psychological factors in the expression of disturbed functioning. The varied psychological features associated with eating disorders contribute to their heterogeneity on presentation. Finally, potentially serious physical and psychological sequelae must be addressed in the understanding and the treatment of the eating disorders. These factors, as well as the apparent increasing prevalence of eating disorders among adolescent and young adult women, have led to a rapid acceleration in theoretical formulations and research bearing on pathogenesis and treatment. Since the mid-1970s there has been divergence in etiological formulations as well as convergence of opinion regarding the utility of certain practical intervention principles. In spite of these advancements, current knowledge has yet to yield conclusive support for any one theoretical viewpoint or treatment modality.

The aims of this chapter are to review diagnostic features, some of the more popular models for understanding eating disorders, and to provide an overview of the major approaches to treatment that derive from different theoretical models.

5.23.2 DIAGNOSTIC AND DEFINING FEATURES

5.23.2.1 Phenomonology

A key feature of anorexia nervosa and bulimia nervosa is a persistent overconcern with body size and shape indicated by behavior such as prolonged fasting, strenuous exercise, and self-induced vomiting aimed at decreasing body weight and fat. In anorexia nervosa, the over-concern has been described variously as a relentless pursuit of thinness (Bruch, 1973), a morbid fear of fatness (Russell, 1970), and a weight phobia (Crisp, 1970). The term “bulimia nervosa” was proposed by Russell (1979) to define a syndrome in which the patient suffers from powerful and intractable urges to overeat leading to bouts of binge eating followed by extreme efforts to control weight through self-induced vomiting, laxative abuse, fasting, vigorous exercise, and other behaviors with similar intent. Bulimia nervosa shares many features in common with anorexia nervosa, however, the patient is not emaciated.

There is general agreement that eating disorder symptoms have psychological and developmental significance. Both individual and family theorists have emphasized that eating disorders often represent a developmental struggle for autonomy, independence, and individuality. These normal development hurdles become flashpoints in adolescence, when the vulnerable individual, parents, or entire family are forced to deal with emergent developmental realities. However, the phenomonology of eating disorders cannot be fully appreciated outside the context of cultural values. There has been intense pressure on women to diet in order to conform to ultra-slender role models for feminine beauty (Garner, 1997; Garner & Garfinkel, 1980). Unfortunately, very few women will ever achieve the admired physical form through restrictive dieting, largely owing to biological limits to achieving permanent weight loss. Nevertheless, constructs such as competence, control, attractiveness, self-worth, and self-discipline have continued to be associated with dieting and weight control in our culture. The consequence of the conflict between cultural imperatives and biological realities has been widespread dissatisfaction with body shape and weight that has even affected young girls yet to cross the pubertal threshold (Edlund, Halvarsson, & Sjödén, 1995).

Another aspect of the phenomonology of anorexia nervosa is the "ego-syntonic" nature of certain symptoms such as pathological dieting and other extreme weight controlling behaviors. In contrast to patients with other psychological disorders such as depression or anxiety, most anorexia nervosa patients actively embrace their eating disorder symptoms. It is not uncommon for patients to actually strive for and then cling to an "anorexic identity" due to the disorder's associations with celebrity status and socially desirable traits. The resistance to change seen in anorexia nervosa has obvious implications for treatment and has been a major focus of therapeutic strategies recommended for this disorder (Garner, Vitousek, & Pike, 1997).

5.23.2.2 Diagnosis

The prevailing diagnostic systems, the *Diagnostic and statistical manual of mental disorders, fourth edition* (*DSM-IV*; American Psychiatric Association [APA], 1994) and the *International Classification of Diseases, 10th revision* (World Health Organization, 1992), formally define as well as distinguish the two major eating disorders, anorexia nervosa and bulimia nervosa. By drawing the boundaries for these eating disorders, the current *DSM-IV* diagnostic criteria have substantial implications for clinical care and research.

5.23.2.2.1 Anorexia nervosa

The requirements for diagnosis of anorexia nervosa according to the *DSM-IV* are summarized as follows: (i) refusal to maintain a body weight over a minimally normal weight for age and height (e.g., weight loss leading to maintenance of a body weight less than 85% of that expected or failure to make expected weight gain during a period of growth); (ii) intense fear of gaining weight or becoming fat, even though underweight; (iii) disturbance in the way that body weight, size, or shape is experienced; (iv) amenorrhea in females (absence of at least three consecutive menstrual cycles).

The *DSM-IV* divides anorexia nervosa into two diagnostic subtypes: (i) restricting type and (ii) binge eating/purging type. The restricting type is defined by rigid restriction of food intake without bingeing or purging, whereas the binge eating/purging type is defined by stringent attempts to limit intake punctuated by episodes of binge eating as well as self-induced vomiting and/or laxative abuse. This diverges from earlier conventions in which anorexia nervosa was subdivided simply on the basis of the presence or absence of binge eating (Casper, Eckert, Halmi, Goldberg, & Davis, 1980; Garfinkel, Moldofsky, & Garner, 1980). Patients who regularly engage in bulimic episodes report greater impulsivity, social/sexual dysfunction, substance abuse, general impulse control problems, family dysfunction, and depression as part of a general picture of more conspicuous emotional disturbance compared to patients with the restricting subtype of anorexia nervosa (Garner, Garfinkel, & O'Shaughnessy, 1985; Garner, Garner, & Rosen, 1993; Herzog, Keller, Sacks, Yeh, & Lavori, 1992; Laessle, Wittchen, Fichter, & Pirke, 1989; Rosen, Murkofsky, Steckler, & Skolnick, 1989). In contrast, restricting anorexia nervosa patients have been described as overly compliant but at the same time obstinate, perfectionistic, obsessive-compulsive, shy, introverted, interpersonally sensitive, and stoical (Casper, 1990, Hsu, Kaye, & Weltzin, 1993; Strober, 1983).

5.23.2.2.2 Bulimia nervosa

The criteria for the diagnosis of bulimia nervosa according to the *DSM-IV* are summarized as follows: (i) recurrent episodes of binge eating (a sense of lack of control over eating a large amount of food in a discrete period of time); (ii) recurrent inappropriate compensatory behavior in order to prevent weight gain (i.e., vomiting, abuse of laxatives, diuretics or other medications, fasting or excessive exercise); (iii) a minimum average of two episodes of binge eating and inappropriate compensatory behaviors per week for the past three months; (iv) self-evaluation unduly influenced by body shape and weight; and (v) the disturbance does not occur exclusively during episodes of anorexia nervosa. Bulimia nervosa patients are further subdivided into purging and nonpurging types based on the regular use of self-induced vomiting, laxatives, or diuretics (APA, 1994).

5.23.2.2.3 *Eating disorders not otherwise specified*

The *DSM-IV* delineates a large and heterogeneous diagnostic category, "eating disorder, not otherwise specified" (EDNOS), for individuals with clinically significant eating disorders but who fail to meet all of the diagnostic criteria for anorexia nervosa or bulimia nervosa. Unfortunately, the term "not otherwise specified" could be interpreted as connoting eating problems of minor clinical significance. This assumption is incorrect since the clinical picture for many individuals with EDNOS can be every bit as complicated and serious as the two main eating disorders (Walsh & Garner, 1997). The binge eating disorder is a specific example of an EDNOS included in the current *DSM-IV*. This term applies to individuals who share a number of characteristics of individuals with bulimia nervosa, but who do not regularly engage in the inappropriate compensatory behaviors, such as self-induced vomiting, required for the diagnosis of bulimia nervosa. In order to qualify for this diagnosis, binge eating must occur, on average, at least twice a week for a six-month period. Although there is merit in adopting the binge eating disorder into the diagnostic nomenclature, it is critical to remain aware of the fact that binge eating and associated psychological symptoms, particularly in the obese, may be attributed to standard weight loss treatments (cf. Garner & Wooley, 1991). Other examples of EDNOS have received less attention but are of clear clinical importance. Terms such as "bulimia," "compulsive eating," "normal weight bulimia," often used to describe eating disorders, do not appear in the *DSM-IV*.

5.23.2.2.4 *Relationship between different eating disorders*

There is extraordinary variability within each of the diagnostic subgroups in terms of demographic, clinical, and psychological variables (Welch, Hall, & Renner, 1990). Patients with anorexia nervosa and bulimia nervosa have been observed to move between diagnostic categories at different points in time (Russell, 1979). For example, patients move between the two subtypes of anorexia nervosa (restricting and binge eating/purging types); however, the tendency is for restricters to move toward bulimia (and purging) more often than bulimic anorexics move to an exclusively abstaining mode (Garner et al., 1993; Hsu, 1988; Kreipe, Churchill, & Strauss, 1989).

Even though distinctions between diagnostic subgroups have been emphasized in the research literature, it is important to recognize that the different diagnostic subgroups tend to share many features in common. For example, even though anorexia nervosa patients are differentiated into "restricter" and "binge eating/purging" subtypes, it must be remembered that virtually all eating disorder patients "restrict their food intake," "diet" and probably "fast" for abnormally long periods of time. Some do this in association with binge eating, some with vomiting, and/or purgation, and some with neither of these symptoms. For some this occurs at a statistically "normal" body weight (bulimia nervosa) and for others it occurs well over the body weight norms (e.g., binge-eating disorder). Figure 1 characterizes the relationship between bingeing, purging, and restricting in subgroups of eating disorder patients at different body weights.

5.23.2.2.5 *Differential diagnosis*

Patients without anorexia nervosa, bulimia nervosa, or some variant can superficially resemble patients with an eating disorder diagnosis. Patients with a severe affective disorder can display marked weight loss (due to loss of appetite) or hyperphagia. Schizophrenia can present with an aversion to eating and occasionally binge eating or purging. Vomiting and weight loss also can be associated with what has been described as a conversion disorder (Garfinkel, Kaplan, Garner, & Darby, 1983). A range of physical illnesses producing weight loss (e.g., inflammatory bowel disease, chronic hepatitis, Addison's disease, Crohn's disease, undiagnosis cystic fibrosis, diabetes mellitus, hyperthyroidism, tuberculosis, malignancies, malabsorption diseases, or other wasting diseases) should be ruled out as the primary diagnosis (Powers, 1997).

5.23.3 EPIDEMIOLOGY

5.23.3.1 Incidence and Prevalence

Incidence rates are defined as the number of new cases in the population per year, whereas prevalence rates refers to the actual number of cases in the population at a certain point in time. Prevalence studies have been the most common and have often been conducted on high-risk populations such as college students and athletes. There are serious limitations to estimates of the incidence and prevalence of eating disorders, because most have been derived exclusively from self-report instruments and on samples that may not reflect important

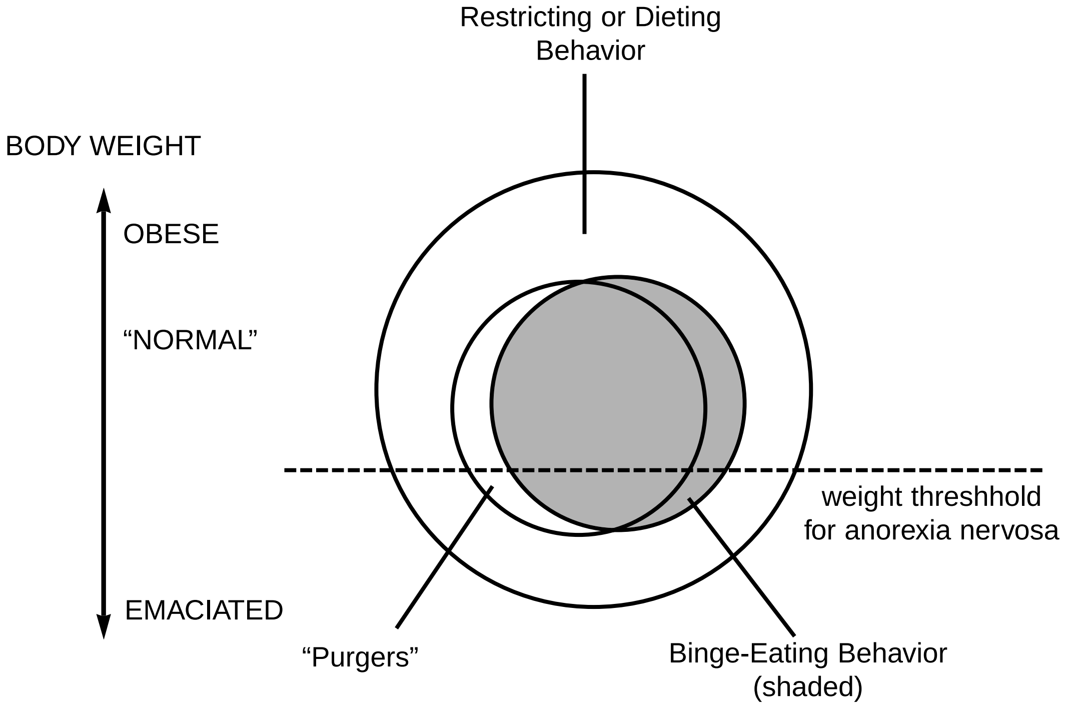
Restricting or Dieting Behavior
BODY WEIGHT
OBESE
"NORMAL"
EMACIATED
weight threshhold for anorexia nervosa
"Purgers"
Binge-Eating Behavior (shaded)

Figure 1 Venn diagram illustrating the relationship between bingeing, purging, and restricting symptoms in eating disorder patients at different body weights. Purging and bingeing are overlaping sets of symptoms that characterize subgroups of patients at different points along the weight spectrum. Anorexia nervosa patients fall below a critical weight threshold (adapted with permission from Garner, Garner, & Rosen, 1993).

demographic differences in base rates. In general, estimates based exclusively on questionnaires yield much higher rates of eating disorders.

Estimates of incidence based on detected cases in primary care practices yielded rates of 8.1 per 100 000 persons per year for anorexia nervosa and 11.5 for bulimia nervosa (Hoek et al., 1995). The most sophisticated prevalence studies using strict diagnostic criteria report rates of about 0.3% for anorexia nervosa and about 1% for bulimia nervosa among young females in the community (Fairburn & Beglin, 1990; Hoek, 1993). This compares to surveys using questionnaires that find that as many as 19% of female students report bulimic symptoms. Prevalence studies of higher risk samples indicate that serious eating disorders occur in as many as 4% of female high school and college students (cf. Fairburn & Beglin, 1990). Suspected cases of clinical eating disorders or subclinical variants are even more common among groups exposed to heightened pressures to diet or maintain a thin shape. Research has confirmed a high incidence of actual or suspected cases among samples of ballet students and professional dancers, wrestlers, swimmers, skaters, and gymnasts (cf. Garner & Rosen, 1991).

5.23.3.2 Gender Considerations

Epidemiological studies have consistently found that eating disorders are more common among women than men; in clinical samples, only 5–10% of patients are men (Hoek et al., 1995). Studies in the USA suggest the incidence of anorexia nervosa among males may be as little as 0.02% per year and the prevalence of current bulimia nervosa in men is between 0.1% and 0.5%. The differences in gender rates for the eating disorders are generally attributed to cultural factors (Garner & Garfinkel, 1980); however, biological and psychodynamic factors may also play a role (Scott, 1986). Eating disorders are less common in men, but when they occur, they have a clinical picture very similar to that observed in women (Andersen, 1990; Olivardia, Pope, Mangweth, & Hudson, 1995).

5.23.3.3 Developmental Considerations

From the earliest descriptions, there has been agreement that anorexia nervosa is a disorder primarily of adolescence and that the core psychopathology often relates to conflicts that emerge with pubertal changes in body shape.

Theander (1996) reviewed the topic of age at onset and concluded that the average age of onset is usually found to be about 17 years of age with about 30% of cases with an onset at 14 years or younger. Selection bias yields younger age at onset in pediatric treatment facilities. Theander (1996) summarizes the distribution of age at onset as follows: (i) there are occasional patients with an early age of onset (7–10 years of age); (ii) the number of cases rises steeply at ages 11 and 12; (iii) most patients develop in the teenage, 13–19; (iv) there is a gradual decline in onset from 20 to 30 years; and (v) there are fewer cases with an onset after 30 years of age. Some have found that age of onset tends to be a bit younger for males than females; however, this point remains controversial (Theander, 1996).

The age of onset of anorexia nervosa is paralleled by data on body dissatisfaction and dieting in nonclinical populations, again supporting developmental vulnerability. Feeling fat and restricting food intake are common among schoolchildren 7–13 years of age (Davies & Furnham, 1986; Edlund et al., 1995; Maloney, McGuire, & Daniels, 1988; Wardle & Beales, 1986). Dieting to lose weight and fear of fatness are common in girls as young as seven years old; these attitudes and behaviors escalate significantly during adolescence, particularly among those at the heavier end of the weight spectrum (Edlund et al.). In a study of 1410 students in grades 9 through 12 (mean age 16.5 years) Timmerman, Wells, and Chen (1990) reported that 2% of the girls and 0.14% of the boys met all of the diagnostic criteria for bulimia nervosa ascertained through an anonymous self-report instrument. Severe dieting because of a fear of fatness has been shown to lead to short-stature syndrome and delayed puberty in a subgroup of young children (Puglifse, Lifshitz, Grad, Fort, & Marks-Katz, 1983). In 1990, 44% of female high school students were trying to lose weight compared to 15% of male students (Serdula et al., 1993). An additional 26% of female students and 15% of male students were trying to keep from gaining more weight. It has been shown that the risk of developing an eating disorder is eight times higher in dieting compared to nondieting 15-year-old girls (Patton, Johnson-Sabine, Wood, Mann, & Wakeling, 1990). Although there are case reports of anorexia nervosa in young children (Fosson, Knibbs, Bryant-Waugh, & Lask; 1987), the consensus is that it is very rare in this age group (Jaffe & Singer, 1989; Lask & Bryant-Waugh, 1997). Similarly, though probably rarely, cases of anorexia nervosa also have been observed at the other end of the age spectrum (Kellett, Trimble, & Thorley, 1976).

5.23.3.4 Have Eating Disorders Changed Over Time?

Anorexia nervosa was once believed to be restricted to the higher socioeconomic classes; however, there is evidence that it has become more common in middle and lower socioeconomic groups (Gowers & McMahon, 1989). According to Russell (1985, 1997), the most dramatic evidence for a transformation in the psychopathology of anorexia nervosa is the increased appearance of binge eating, both in anorexia nervosa and eating disorder patients who are not emaciated. However, the actual extent to which binge eating has become more common in anorexia nervosa is unclear owing to the fact that the symptom may have been identified less reliably in earlier reports (Garner, 1993).

Casper (1983) suggested another shift in the psychopathology of eating disorders over the past several decades whereby "ascetic motives" for weight loss have become less prominent, replaced by the "drive for thinness" as the most common motivational theme. The theme of asceticism was common in early writings on anorexia nervosa (cf. Bell, 1985; Rampling, 1985) and is expressed in the conceptions of dieting as purification, thinness as virtue, and fasting as an act of penitence. Nevertheless, there is a subgroup of eating disordered patients who seem to be motivated by belief in the virtue of oral self-restraint (Garner & Bemis, 1982, 1985; Garner et al., 1997). Evidence suggests that "oral self-restraint" may be part of a more general theme of renunciation of physical gratification (Haimes & Katz, 1988). This is supported by findings that anorexia nervosa patients have elevated scores on the Asceticism subscale of the Eating Disorder Inventory (Garner, 1991), a subscale designed to identify the tendency to seek virtue through the pursuit of spiritual ideals, such as self-discipline, self-denial, self-restraint, self-sacrifice, and control of bodily urges.

5.23.4 COMORBIDITY

Since the earliest descriptions of anorexia nervosa there has been considerable controversy as to whether eating disorders represent discrete psychological entities or are simply manifestations of other illnesses (cf. Beumont, Al-Alami, & Touyz, 1987). Anorexia nervosa and bulimia nervosa each have been considered variants of affective disorder, obsessive-compulsive disorder, or borderline personality disorder. Traits such as hostility, somatization, social maladjustment, physical anhedonia, affective rigidity, interpersonal sensitivity, anxiety, poor self-esteem, external locus of control, and confused sex-role identity have been observed repeatedly in eating disorder patients (Beumont, George, & Smart, 1976; Garfinkel & Garner, 1982; Garner et al., 1993; Johnson & Connors, 1987; Rastam, Gillberg, & Gillberg, 1995; Strober, 1981; Swift & Wonderlich, 1988; Williamson, 1990). Although there is general acceptance that eating disorders are best considered as discrete clinical entities there is growing interest in psychological features that may define meaningful subgroups of eating-disordered patients.

5.23.4.1 Depression

Depression has been described as a common theme in eating disorders, but its precise role is controversial. Pope and Hudson (1988) originally argued that bulimia nervosa is a variant of depression based on evidence that eating disorder patients exhibit: (i) a high prevalence of depression, (ii) a family history of depression, (iii) biological markers of depression, and (iv) a positive response to antidepressant medications. Detractors from this view provide alternative interpretations for these associations and contend that the nature of bulimia nervosa as a variant of depression is "simplistic and theoretically limiting" (Strober & Katz, 1988). Although depression may play an important predisposing role in eating disorders, current evidence fails to support the proposition that eating disorders are simply depressive equivalents (Cooper & Fairburn, 1986; Hinz & Williamson, 1987; Laessle, Kittl, Fichter, Wittchen, & Pirke, 1987; Wilson & Lindholm, 1987). This said, the evidence is overwhelming that depression is one of the most common experiences of those with eating disorders, and to minimize the significance of this experience would be misleading. Moreover, bulimia nervosa patients with a history of affective disorder or substance abuse disorder report significantly more suicide attempts, social impairment, and previous treatment, both before and after the onset of the eating disorder (Hatsukami, Mitchell, Eckert, & Pyle, 1986). Hatsukami et al. (1986) found no suicide attempts among the group of patients with bulimia alone; however, 26.5% of those with concurrent affective disorder and 32.4% with substance abuse had attempted suicide. Bulimia nervosa patients with a history of substance abuse also report a high incidence of stealing, perhaps suggesting a more general problem with impulsivity.

Some studies indicate that depressive states of varying intensities foreshadow the development of anorexia nervosa with lifetime prevalence rates of major depressive illness ranging between 25% and 80% across different samples (Herzog, Keller, Sacks, Yeh, & Lavori, 1992). Halmi et al. (1991) found a lifetime prevalence of 68% for major depression in a sample of severely ill anorexia nervosa patients. Depression can be secondary to starvation and coexisting complications since improved mood often follows nutritional rehabilitation (Eckert, Halmi, Marchi, & Cohen, 1987). One population study found that depression typically did not precede the eating disorder, although it did correlate with onset (Rastam, 1992). In a six-year follow-up, Rastam et al. (1995) found that affective disorders were common throughout the follow-up period, but tended to follow the course of the disorder rather than to precede or postdate it. Thus, a history of affective disorder may be a foreboding complication for a subgroup of eating-disordered patients who may be particularly resistant to treatment efforts. Unfortunately, this subgroup of patients does not seem to respond more favorably to antidepressant therapy than those without a history of mood disorder (Brotman, Herzog, & Woods, 1984).

5.23.4.2 Anxiety

A lifetime prevalence of social phobia has been identified in more than 25% of anorexia nervosa patients (Halmi et al., 1991). The most common anxiety symptoms documented in anorexia nervosa are obsessive-compulsive (OC) in nature (Halmi et al., 1991; Rastam, 1992; Rastam et al., 1995; Rottenberg, 1988; Thiel, Broocks, Ohlmeier, Jacoby, & Schuβler, 1995). In reviewing comorbidity in anorexia nervosa, Rothenberg (1988) found OC symptoms reported in anywhere between 11% and 83% of patient samples, either during the active phase of the disorder or after weight restoration. In a sample of 93 women with either anorexia or bulimia nervosa, Thiel et al. (1995) found that 37% met diagnostic criteria for OC disorder. These patients with concomitant OC disorder also had higher mean scores on five of eight Eating Disorder Inventory subscale scores. Rastam et al. followed a sample of anorexia nervosa patients six years after onset and found a high rate of OC disorder and that underlying personality disorders tended to predict poor outcome. A common diathesis between eating disorders and OC disorder is suggested by the finding that 11% of 151 women presenting with OC disorder had a history of anorexia nervosa (Kasvikis, Isakiris, Marks, Basoglu, & Noshirvani, 1986).

5.23.4.3 Personality Disorders

Since the mid-1980s, there has been intense interest in the relationship between eating disorders and personality disorders. A number of reports have indicated that almost two-thirds of eating-disordered patients receive a concurrent diagnosis of a personality disorder, with borderline personality disorder being reported as particularly common (Bulik, Sullivan, Joyce, & Carter, 1995; Cooper et al., 1988; Gillberg, Rastam, & Gillberg, 1995; Johnson, Tobin, & Dennis, 1990; Levin & Hyler, 1986; Piran, Lerner, Garfinkel, Kennedy, & Brouilette, 1988). Levin and Hyler (1986) assessed 24 bulimia nervosa patients and found that 15 (63%) met diagnostic criteria for personality disorder, with six (25%) fulfilling the diagnosis for borderline personality disorder. Similarly, Bulik et al. (1995) found at least one personality disorder in 63% of a sample of 76 women with bulimia nervosa. Fifty-one percent of the personality disorders were in cluster C (specifically, avoidant, OC, or dependent personality disorders), 41% in cluster B (particularly, borderline or histrionic), and 33% in cluster A (paranoid, schizoid, or schizotypal). In an earlier evaluation of 35 patients with eating disorders, Gartner, Marcus, Halmi, and Loranger (1989) found that 57% met the *DSM-III-R* diagnostic criteria for at least one form of personality disorder, with borderline, self-defeating, and avoidant being the most common. Two or more Axis II diagnostic criteria were met by 40% of the patients, and 17% fulfilled all of the criteria for 5–7 personality disorder diagnoses. Wonderlich, Swift, Slotnick, and Goodman (1990) interviewed 46 eating-disordered patients and reported that 72% met criteria for at least one personality disorder. OC personality disorder was common among restricting anorexic patients. Histrionic and borderline personality disorder diagnoses were common among bulimic groups. Johnson et al. (1990) followed patients one year after an initial assessment and found that those who initially scored above a threshold on the self-report Borderline Syndrome Index had a worse prognosis in terms of eating behavior and general psychiatric symptoms. Gillberg et al. (1995) found that OC and avoidant personality disorders were particularly common in a study comparing 51 anorexia nervosa patients with an age-matched community sample.

Arguing that borderline assessment measures are confounded by certain eating symptoms, Pope and Hudson (1989) have challenged the interpretation that borderline personality disorder is over-represented among eating disorders. For example, bulimic eating patterns may be used to satisfy the *DSM-III-R* poor impulse control criterion for borderline personality disorder, making the association between disorders tautological. Nevertheless, the tendency toward poor impulse regulation has been identified as a negative prognostic sign in eating disorders (e.g., Hatsukami et al., 1986; Sohlberg, Norring, Holmgren, & Rosmark, 1989). Results from research on the incidence and prevalence of personality disorders in anorexia nervosa are inconsistent. Some studies indicate remarkably high rates, with avoidant personality disorder occurring in as many as 33% of anorexic restricters and borderline personality disorder occurring in almost 40% of anorexic bulimic patients (Piran et al., 1988). Other studies suggest that personality disorders are relatively uncommon in anorexia nervosa (Herzog et al., 1992; Pope & Hudson, 1989). Impulse control problems, such as self-mutilation, suicide attempts, and stealing, are reported in a subgroup of anorexia nervosa patients, particularly those with purging and/or bulimic symptoms (Fahy & Eisler, 1993; Garner et al., 1993). While personality disturbances are not uniform in eating disorders, their presence suggests meaningful subtypes that may be relevant to treatment planning and prognosis.

5.23.4.4 Sexual Abuse

There has been considerable interest and controversy regarding the role of sexual abuse as a risk factor for the development of eating disorders. Clinical accounts and the observation in some studies of a high incidence of sexual abuse in eating disorder patient samples (Oppenheimer, Howells, Palmer, & Chaloner, 1985) were followed by further clinical reports and numerous empirical studies yielding conflicting findings (Fallon & Wonderlich, 1997). Fallon and Wonderlich summarized the literature and concluded that: (i) childhood sexual abuse appears to be positively associated with bulimia nervosa; (ii) there is less evidence for this association in anorexia nervosa; (iii) childhood sexual abuse does not appear to be a specific risk factor for eating disorders (i.e., it is no higher for eating disorders than in psychiatric controls); (iv) childhood sexual abuse does appear to be associated with greater levels of comorbidity among those with eating disorders, however, there is not strong evidence that it predicts a more severe eating disorder; and (v) a more complex approach to the definition of sexual abuse has led to better prediction of later disturbances in eating. It is indisputable that a significant subgroup of women from some clinical eating disorder samples have a history of sexual abuse and that careful assessment and treatment is important in the process of dealing with resulting feelings of shame, distrust, and anger (Fallon & Wonderlich, 1997).

5.23.5 ETIOLOGY

A complete understanding of anorexia nervosa and bulimia nervosa must account for factors that predispose individuals to each of the eating disorders. Theory should be able to delineate the variety of developmental experiences that interact with these factors to initiate symptom expression as well as account for the maintaining variables (biological, psychological, and interpersonal) and the key variations in the symptom picture. Although current models are unable to specify all of these elements in precise detail, research and clinical observations since the mid-1970s have improved the understanding of eating disorders, which in turn, has led to more sophisticated treatment recommendations.

In the latter half of the twentieth century single-factor causal theories have been replaced by the view that eating disorders are "multidetermined" (Garfinkel & Garner, 1982; Garner, 1993; Garner & Garfinkel, 1980). The symptom patterns represent final common pathways resulting from the interplay of three broad classes of predisposing factors shown in Figure 2. The role of cultural, individual (psychological and biological), and familial causal factors are presumed to combine with each other in different ways leading to the development of eating disorders. The precipitants are less clearly understood except that dieting is invariably an early element. Perhaps the most practical advancements in treatment have come from increased awareness of the perpetuating effects of starvation with its psychological, emotional, and physical consequences (Figure 2).

5.23.5.1 Genetic Contributions

The evidence for genetic vulnerability to anorexia nervosa comes from approximately 100 twin pairs culled from selected twin case report summaries and from twin studies (Garfinkel & Garner, 1982; Holland, Sicotte, & Treasure, 1988). These reports indicate concordance rates of more than 50% for monozygotic twin pairs compared to less than 10% for dizygotic twins. This is in contrast to one study of 11 twin pairs (5 monozygotic) where none of the co-twins was found to be concordant for anorexia nervosa (Waters, Beaumont, Touyz, & Kennedy, 1990). Strober, Lampert, Morrell, Burroughs, and Jacobs (1990) found first degree relatives of eating disorder patients have a greater risk of having anorexia nervosa. Monozygotic twins have a concordance rate of about 50% compared to about 10% for dizygotic twin pairs (Holland et al., 1988). These studies suggest that there may be a genetic component to the transmission of anorexia nervosa; however, it is not at all clear what is inherited. Is it the specific disorder, a particular personality trait associated with the disorder, or a general vulnerability to psychiatric disturbance? Moreover, the concordance data on twins reared together does not conclusively distinguish between genetic and environmental transmission (Holland et al.).

5.23.5.2 Developmental Theory

Defined in broad terms, many theories have offered developmental explanations of eating disorders. Some formulations have emphasized intrapsychic mediation and others have highlighted interactional factors. Together they have provided a rich understanding of a range of psychological themes that may account for both anorexia nervosa and bulimia nervosa.

Many of the early psychodynamic writings on eating disorders emphasized anorexia nervosa as a rejection of adult femininity. This theme was refined and extended by Crisp (1965) who contends that the central psychopathology of both anorexia nervosa and bulimia nervosa is related to fears of the psychological and biological experiences associated with adult weight. According to this view, starvation becomes the mechanism for avoiding psychobiological maturity because it results in a return to prepubertal appearance and hormonal status. This regression is thought to provide relief from adolescent turmoil and conflicts within the family. Moreover, patients report experiencing themselves as younger following extreme weight loss, and this may be due to the reversal of mature adolescent hormonal profiles that form the biological substrata for psychological experience (Crisp, 1980).

Other developmental theorists have attributed eating disorders to various types of parenting failure. Bruch (1973) and Selvini-Palazzoli (1974) provided a developmental paradigm in which the mother superimposes her own inaccurate perceptions of the child's needs on the child. Such invalidation of the child's experiences results in an arrest of cognitive development and is manifest in debilitating feelings of ineffectiveness that are later evident in adolescent struggles for autonomy and control of the body. Bruch (1973) postulated

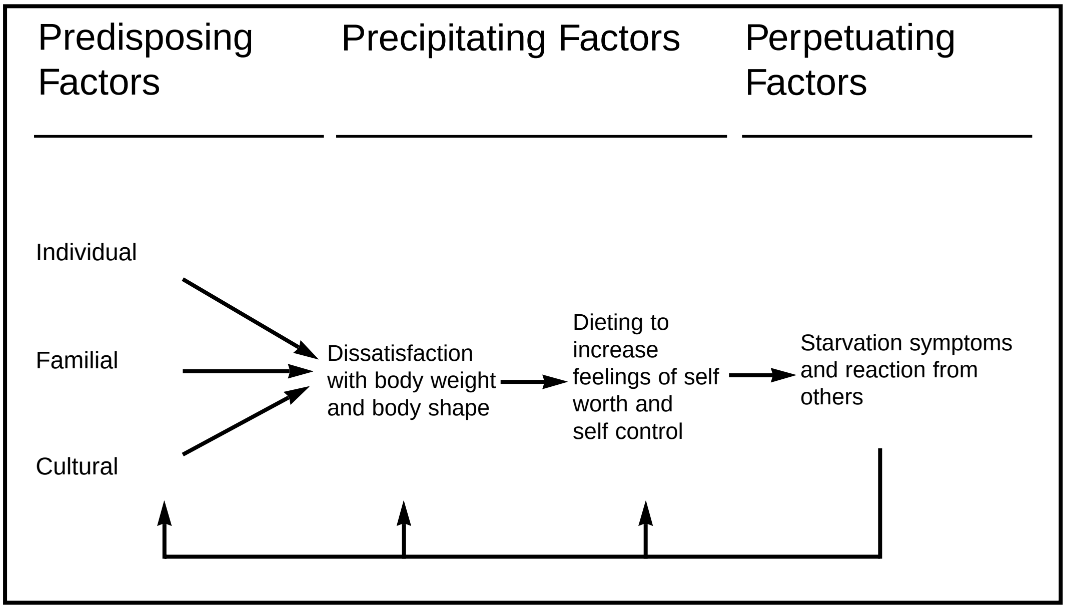
Predisposing Factors
Precipitating Factors
Perpetuating Factors
Individual
Familial
Cultural
Dissatisfaction with body weight and body shape
Dieting to increase feelings of self worth and self control
Starvation symptoms and reaction from others

Figure 2 Eating disorders as multidetermined disorders. ("Pathogenesis of anorexia nervosa," by D. M. Garner, 1993, *The Lancet, 341,* p. 1632. Copyright 1993 by *The Lancet*. Reprinted with permission.)

that these early parenting failures lead to fundamental deficits in self-awareness, including the way that the body is perceived and experienced.

A similar view is expressed by theorists working from an object relations model to account for eating disorders (Johnson & Connors, 1987; Masterson, 1977; Stern, 1986; Sugarman, Quinlan, & Devenis, 1982). According to this view, normal development is characterized by the process of separation–individuation, a state achieved by consistent interactions with a primary caretaker who is highly responsive to the needs of the child. Success is reflected by the child's progressive internalization of the ability to accurately recognize, respond, and regulate internal needs and impulses without contact with the immediate caregiver. Accordingly, the different personality structures of eating disordered patients reflect developmental arrest at particular phases of the process of separation–individuation. The eating disorder symptoms are an attempt to cope with needs stemming from incomplete self-development or an interruption of the separation–individuation process.

The self-psychology perspective of Kohut (1971) has also been used to account for eating disorders as a reflection of developmental arrest in the separation–individuation process (Casper, 1983; Goodsitt, 1985; Swift & Letven, 1984). According to this conceptual scheme, the eating disordered patient's lack of a cohesive sense of self is the direct result of the primary caregiver's failure to provide essential functions during development. These include mirroring, tension regulation, and integration. The obsessive concern about eating and repeated bouts of bingeing and vomiting serve organizing and tension-regulation functions in the attempt to modulate basic deficits in "self-structures." These symptoms become an organizing event in one's life and provide intense stimulation that numb the anguish and emptiness that pervade the eating disordered patient's experience. There is growing empirical evidence to support the clinical view that intense separation distress is common among eating disorders and that many patients demonstrate marked separation anxiety in response to severely as well as mildly stressful situations (Armstrong & Roth, 1989).

Others have emphasized the motivation for binge eating as an attempt to escape negative aspects of self-awareness (cf. Heatherton & Baumeister, 1991). According to this view, the binge eater experiences negative self-evaluations and concerns about perceived negative evaluations from others that are accompanied by anxiety and depression. Binge eating provides a means of avoiding this unpleasant state of negative self-awareness through the narrowing of attention to one facet of the immediate environment.

Strober (1991) has drawn together developmental theory, psychobiology, and personality genetics in offering a valuable understanding of the adaptive mechanisms behind the range of symptoms seen in eating disorder patients. He integrates modern psychoanalytic concepts of development with constructs indicating that: (i) individual differences may be expected in the internal regulators of arousability or temperament that organize self-experiences, and (ii) heritable personality traits, and their presumed biological substrata, set limits within which behavior patterns are expressed. Strober accounts for the differences in traits seen in subgroups of eating disorder patients by incorporating Cloninger's (1989) concept of three bipolar, and primarily genetically determined, dimensions of personality that appear to account for remarkable consistency over time in both normal and psychopathological behavior.

5.23.5.3 Familial Factors

Some of the earliest descriptions of anorexia nervosa emphasized the role of the family in the development of the disorder (cf. Garfinkel & Garner, 1982). The possible role of the family in the development of anorexia nervosa and bulimia nervosa has been described by writers from a range of theoretical orientations (Bruch, 1973; Dare & Eisler, 1997; Garfinkel & Garner, 1982; Humphrey, 1989; Minuchin, Rosman, & Baker, 1978; Root, Fallon, & Friedrich, 1985; Schwartz, Barrett, & Saba, 1985; Selvini-Palazzoli, 1974; Vandereycken, Kog, & Vanderlinden, 1989). A major advancement in family therapy with eating disorders was the formulation of the structural approach of Minuchin et al. (1978). They identified a number of characteristics of the interactions encountered in eating disordered families, including enmeshment, overprotectiveness, rigidity, and poor conflict resolution. The systemic model was applied to eating disorders by Selvini-Palazzoli (1974) and overlaps considerably with the structural model. The systemic approach holds that the identified patient serves an essential homeostatic or stabilizing role within the family. A second generation of family theorists have integrated and greatly elaborated earlier approaches providing very specific advice regarding the treatment of bulimia nervosa (Root et al., 1985; Schwartz et al., 1985; Vandereycken et al., 1989).

Regardless of their orientation, most individual and family theorists have emphasized that anorexia nervosa often represents a developmental struggle for autonomy, independence, and individuality. These issues are very likely to surface in adolescence, when the vulnerable individual, parents, or entire family are forced to deal with emergent developmental realities. In the only controlled trial examining the efficacy of family therapy for anorexia nervosa, Russell, Szmukler, Dare, and Eisler (1987) found that family therapy was superior to individual therapy for younger patients. This study, as well as clinical experience, suggests that family therapy should be routinely employed as the treatment of choice for young eating disorder patients.

5.23.5.4 The Addiction Model

The term compulsive overeating is somewhat problematical since the *DSM-III-R* diagnostic criteria for OC disorder specifically excludes those who meet an eating disorder diagnosis. Nevertheless, the suggestion that binge eating is compulsive behavior and a form of substance abuse has intuitive appeal on a number of grounds. Studies have shown both that substance abuse and poor impulse regulation are common features in bulimia nervosa (Garner, Rockert, et al., 1985; Garner et al., 1993; Hatsukami, Owen, Pyle, & Mitchell, 1982). Moreover, bulimia nervosa shares several features with addictive disorders that have been summarized by Bemis (1985) as: (i) "loss of control, (ii) preoccupation with the abused substance, (iii) use of the abused substance to cope with stress and negative feelings, (iv) secrecy about the behavior, and maintenance of the addictive behavior despite the aversive consequences" (p. 415).

The addiction model is the framework for Overeaters Anonymous (OA), one of the most popular self-help programs for bulimia nervosa and for other commercial eating disorder programs patterned after the OA or "12-step" approach to chemical dependency. Malenbaum, Herzog, Eisenthal, and Wyshak (1988) have described many of the benefits of OA, including group support, "sponsorship" by someone who has achieved some measure of success, and the clear guiding philosophy originally applied to alcoholism in Alcoholics Anonymous. OA programs are aimed at achieving abstinence from compulsive overeating which is achieved, in part, by adhering to several behavioral axioms. Participants are usually advised to eat nothing beyond three meals a day, on the basis of the premise that certain foods are "addicting" and will trigger bouts of compulsive overeating (Malenbaum et al., 1988).

Although there are parallels between eating disorders and chemical dependency, there are serious theoretical and practical limitations to the comparison. Bemis (1985), and more recently Wilson (1991, 1993), have provided a thorough analysis of the topic, and only the salient points will be reviewed. The most obvious flaw in the application of the addiction model to overeating is that abstinence from food (unlike abstinence from alcohol) is obviously unrealistic, and recommending "controlled eating" is analogous to recommending controlled drinking, which is antithetical to the abstinence model applied to substance abuse (Bemis, 1985; Garner, 1985; Malenbaum et al., 1988; Wilson, 1991, 1993). The parallel breaks down further in trying to understand the nature of the compulsive overeater's addiction. Empirical evidence is lacking for the assertion that specific foods are addicting. Moreover, the OA explanation of compulsive overeating is completely inconsistent with the far more parsimonious understanding of binge eating, food preoccupations and food cravings as biological adaptations to chronic undernutrition (Garner, 1997; Garner, Rockert, et al., 1985; Wardle, 1987). Thus, to the degree that OA programs encourage "control" by avoiding certain foods while not deviating from three small meals a day (Garner, 1985), they may inadvertently collude with the bulimia nervosa patient's restrictive eating patterns. Unfortunately, these programs also reinforce the fear that normal eating will result in inordinate weight gain. Moreover, encouraging participants to avoid "forbidden foods" is inconsistent with the evidence from the treatment literature on the effectiveness of incorporating these foods into a daily eating plan (Rosen & Leitenberg, 1985).

While strongly encouraging abstinence from bingeing or vomiting is consistent with recommendations of other treatment approaches, the relative intolerance for "slips" in some abstinence oriented eating disorder programs may exacerbate the harsh, self-critical, and dichotomous thinking patterns that are already prominent among these patients (Bemis, 1985). The OA model also has been criticized for presenting compulsive overeating as an "incurable illness," a characterization that contradicts the growing literature on treatment success. Finally, whereas some patients find the spiritual orientation of the OA approach appealing, others may find it incompatible with their religious or moral beliefs. In Wilson's (1991) review of the addiction model of eating disorders, he decisively concludes that: (i) it does not address the core clinical features of eating disorders, (ii) it fails to account for the effects of dieting on behavior, and (iii) it may seriously undermine the effectiveness of treatment. Despite these criticisms, it is important to acknowledge that many find the OA model useful and effective. Clearly what is needed is research extending the one uncontrolled report in the literature (Malenbaum et al., 1988).

5.23.5.5 Cultural Factors

A complete understanding of eating disorders requires attention to the cultural forces selectively impinging on young girls and women in modern times. One of the most pernicious has been the intense pressure to diet and to engage in strict weight control in order to meet unrealistic standards for thinness (Garner & Garfinkel, 1980; Garner, Garfinkel, Schwartz, & Thompson, 1980). Young women today are totally immersed in the cultural admiration of a physical form for women that has little to do with the actual shape of most women in our society. The disconcerting result is a norm in which women report being dissatisfied with their shapes and to feeling guilty about eating even reasonable amounts of food. It has been increasingly recognized that dieting can play a direct role in causing a range of symptoms such as binge eating and mood disturbances (Garner, 1997; Garner, Rockert, et al., 1985; Polivy & Herman, 1985; Striegel-Moore, Silberstein, & Rodin, 1986). The values surrounding slenderness have become sufficiently embedded in our cultural value system that many of the symptoms required for eating disorders are not viewed as unusual or abnormal by members of the general public (Huon, Brown, & Morris, 1988). There is even some evidence that eating disorders may have developed a positive social stereotype and, in some instances, may be spread by social contagion (Bruch, 1985; Chiodo & Latimer, 1983). Recognition of the impact of cultural factors on norms related to dieting and weight control has led to the conclusion that eating disorders may develop in those without underlying personality disturbances or family dysfunction, although secondary disruption in both of these areas may be present by the time the person presents for an assessment (Garner, 1997). Much of current psychological theorizing related to eating disorders may be criticized for not accounting for these cultural factors or, when they are mentioned, not specifying the details of how they must be integrated into the understanding of the psychology of the disorder.

5.23.6 MULTIMODAL ASSESSMENT

Descriptive refinements along with advances in the areas of diagnosis, pathogenesis, and treatment have resulted in tremendous improvements in the technology of assessment (Crowther & Sherwood, 1997). While there is no universally accepted assessment protocol for eating disorders, there is a consensus on the value of a multimodal assessment approach based on the conceptualization of eating disorders as multidetermined and heterogeneous in nature.

The targets for assessment of the eating disorder patient can be divided into two main areas. First is specific psychopathology and behavioral patterns that define the core features of the disorders. Symptoms such as binge eating, extreme weight control behaviors, and stereotypic attitudes toward weight or shape fall into this category. Second is psychopathology not necessarily specific to eating disorder patients, but that has particular theoretical or clinical relevance. Examples include psychological features such as low self-esteem, perfectionism, fears of psychobiological maturity, poor impulse control, and reactions to sexual abuse. There are also more general associated features such as depression, anxiety and poor social functioning which are important aspects of the psychopathology.

5.23.6.1 Targets of Assessment

5.23.6.1.1 Body weight history

Athough body weight is not "psychopathology" *per se*, it is obviously a fundamental aspect of eating disorders. Determining the patient's current body weight and weight history provide essential diagnostic information. This information serves as a basis for exploring the central psychopathology of eating disorders which is the meaning body weight and shape has to the patient. Since amenorrhea is required for anorexia nervosa, menstrual history should be determined within the context of the weight history.

5.23.6.1.2 Binge eating

Binge eating is a key symptom in bulimia nervosa and binge eating disorder. It occurs in about 50% of cases presenting with anorexia nervosa. Thus, determining the presence or absence of binge eating is critical to diagnosis and treatment planning. The *DSM-IV* defines binge eating as having two main characteristics: (i) consumption of a large amount of food, and (ii) the experience of loss of control. Despite the definitional requirements, research has indicated a significant minority of eating disorder patients describe binges involving relatively small amounts of food. Moreover, in some cases there is no real loss of control. Thus, until there is agreement on the significance of size of binge episodes, it is recommended that assessments follow the system proposed in the eating disorder examination (below) for dividing episodes of overeating into four types based on amount of food eaten (large or small) and loss of control (present or absent). It is important to determine age of onset, frequency, and duration of binge eating episodes. Likewise, circumstances surrounding binge episodes (settings, times of day, social context, thoughts, and emotions) should be ascertained.

5.23.6.1.3 Extreme weight losing behaviors

The intensity of the dieting efforts and the types of weight losing behaviors employed can reflect psychopathology. A dieting history should pinpoint when dieting first began and its course over time. Information should be gathered regarding extreme weight control behaviors including self-induced vomiting, laxative and diuretic abuse, use of diet pills or other drugs to control appetite, use of emetics, chewing and spitting food out before swallowing, prolonged fasting, and excessive exercise. Diabetic patients may manipulate insulin levels and patients taking thyroid replacement hormone may alter their dosage to control body weight. Establishing the frequency (as well as the number of symptom free days) for vomiting, laxative abuse, and other extreme weight control behaviors is essential in determining the severity of the disorder and the need for medical consultation.

5.23.6.1.4 Psychopathology related to weight or shape

The psychopathology related to weight or shape has been described in various ways over the years including a drive for thinness, fear of fatness, shape and weight dissatisfaction, body size misperception, body image disturbance, and fears associated with physical maturity. Dissatisfaction with overall body shape and disparagement directed toward specific bodily regions are common in eating disorders and should be a focus of assessment. Some patients may actually overestimate their body size. Although research has shown that this is not unique to eating disorder patients, it may have clinical importance, particularly for emaciated patients. A critical psychopathological feature in anorexia nervosa and bulimia nervosa is that patients must be more than merely dissatisfied with their body; they rely on weight or shape as the predominant or even the sole criterion for judging their self-worth. This criterion has been made mandatory for a diagnosis of the two main eating disorders according to the *DSM-IV*. In anorexia nervosa, there is often a denial of the seriousness of a current low weight as well as intense fear of weight gain.

5.23.6.1.5 General psychopathology relevant to eating disorders

Eating disorder patients have been described as suffering from ineffectiveness feelings, low self-esteem, lack of autonomy, obsessionality, interpersonal sensitivity, introversion, poor relationship skills, social anxiety, dependence, perfectionism, fears of psychobiological maturity, poor impulse control, external locus of control, conflict avoidance, developmental pathology, failure in separation–individuation, vulnerability to substance abuse, interoceptive deficits, and idiosyncratic or dysfunctional thinking patterns. Complete psychological assessment should include these areas as well as evaluation of stable personality features, overall psychological distress, depression, anxiety, family functioning, history of sexual abuse, social and vocational adaptation, all of which may be relevant to the development and maintenance of these syndromes.

5.23.6.1.6 Psychopathology secondary to starvation

It may not be apparent from the initial assessment whether psychological distress, cognitive impairment, and behavioral symptoms signal fundamental emotional disturbance or are secondary elaborations resulting from weight loss and chaotic dietary patterns. Symptoms such as poor concentration, lability of mood, depressive features, obsessional thinking, irritability, difficulties with decision making, impulsivity, and social withdrawal have been identified in normal subjects undergoing semistarvation (Garner, 1997). Thus, these symptoms may reflect secondary effects of weight suppression rather than primary psychopathogy in eating disorder patients.

5.23.6.2 Assessment Methods

Various approaches to information gathering have been developed for eating disorders including clinical interviews, self-report measures, self-monitoring, direct behavioral observation, symptom checklists, clinical rating scales, the stroop test, and standardized test meals. The most commonly used methods are semistructured clinical interviews and self-report measures.

5.23.6.2.1 Clinical interview

Clinical interviews have been the primary method for gathering information on eating disorders. The Eating Disorder Examination (EDE, Cooper & Fairburn, 1987) is the most well validated and has generated a large body of research. It is an investigator based semi-structured interview for assessing psychopathology specific to eating disorders and is the current interview method of choice. Responses are organized on four subscales (restraint, eating concern, shape concern, and weight concern). The EDE can be used to arrive at a diagnosis, it has proven sensitive to treatment effects, and it defines different forms of overeating based upon amount of food eaten (large vs. small) and presence or absence of loss of control. The EDE has the advantages of allowing a more fine-grained appraisal of the specific psychopathology of eating disorders and permitting investigator probes to clarify the meaning behind responses to questions. Disadvantages of the interview include the fact that it takes an hour or more to administer, it requires a trained interviewer, and it is not suitable when anonymity or group administration are required.

5.23.6.2.2 Self-report measures

Various self-report instruments have been introduced to measure eating disorder symptoms. The two most widely used in clinical and research settings are the Eating Attitudes Test (EAT) and the Eating Disorder Inventory (EDI). The EAT is a widely used, standardized, self-report measure of eating disorder symptoms (Garner, 1997). A factor analysis of the original 40-item version resulted in a brief, 26-item measure of global eating disorder symptoms. The EDI is a standardized, multiscale instrument with a much broader focus (Garner, 1991). It comprises three subscales tapping attitudes and behaviors concerning eating, weight, and shape (drive for thinness, bulimia, body dissatisfaction) plus subscales assessing more general psychological traits or organizing constructs clinically relevant to eating disorders (ineffectiveness, perfection, interpersonal distrust, interoceptive awareness, maturity fears). The EDI-2 (Garner, 1991) added three new subscales to the original instrument (asceticism, impulse regulation, and social insecurity). The EAT and EDI have good psychometric properties and are sensitive to treatment effects. The EDI provides a psychological profile that can be useful in clinical situations.

Self-report measures have the advantages of being relatively economical, brief, easily administered and objectively scored. They are not susceptible to bias from interviewer–subject interactions and can be administered anonymously. The major disadvantage of self-report measures is that they are less accurate than interview methods, particularly when assessing ambiguous behaviors such as binge eating. They need to be supplemented by symptom frequency data derived by interview or a symptom checklist. Under some circumstances, symptom data gathered using self-report symptom checklist (e.g., Eating Disorder Symptom Checklist; Garner, 1991) can provide information pertinent to a diagnostic formulation. The different methods for assessing psychopathology in eating disorders have different aims, strengths, and weaknesses. The strategy adopted should be guided by the aims of the assessment and, whenever possible, convergent methods should be employed.

5.23.7 CLINICAL DECISION MAKING IN TREATMENT

The number of psychotherapeutic options available in treating eating disorders has expanded remarkably since the mid-1980s. The major approaches to psychotherapy have been well articulated along with a growing list of alternative forms of treatment (Garner & Garfinkel, 1997). The aim of the following section is to briefly review the major treatment options for eating disorders within the context of a decision-making paradigm for integration and sequencing of these treatments.

5.23.7.1 Integration and Sequencing of Treatments

The wisdom of considering integration of different psychotherapeutic procedures is increasingly evident with the demonstrated effectiveness of different forms of treatment (Garner & Needleman, 1997). The notion of applying different treatments to different eating disorder patients is not new and has formed the basis for multidimensional approaches to psychotherapy (Garner, Garfinkel, & Bemis, 1982). However, there has been more recent interest in "stepped-care," "decision-tree," or "integration" models which rely on set rules for the delivery of the various treatment options (Agras, 1993; Fairburn, Agras, & Wilson, 1992; Fairburn & Peveler, 1990; Garner, Garfinkel, & Irvine, 1986; Garner & Needleman, 1997; Tiller, Schmidt, & Treasure, 1993). These overlapping concepts of treatment delivery share the value system of nonallegiance to a single theoretical orientation; however, they have somewhat different points of emphasis.

Figure 3 illustrates a tentative decision-making paradigm for sequencing and integration of treatments for people with eating disorders. The various treatment options are represented by the boxes with sharp corners (hospitalization, family therapy, self-help, individual therapy, etc.) as well as questions that need to be addressed (boxes with curved corners) related to symptom areas, patient characteristics, or response to previous treatments in determining the treatment to be offered. It does not include all treatment alternatives but concentrates on those for which there is good clinical and/or empirical support. With this schematic in mind, the major approaches to treatment will be reviewed along with the clinical criteria for their integration and sequencing.

5.23.7.2 Hospitalization and Day Treatment

Hospitalization may be considered in the treatment of eating disorders in order to meet at least two different objectives: (i) simple medical stabilization, and (ii) a strategic step aimed at the recovery process (Garner & Sackeyfio, 1993). When aimed at simple medical stabilization in treating physical complications, hospitalization is a medical priority that does not require the commitment by the patient to recover from the eating disorder. This contrasts with hospitalization intended as an important step in the process of recovery from an eating disorder (Andersen, Bowers, & Evans, 1997). As part of recovery, hospitalization not only addresses the physical aspects of the disorder but also targets the psychological problems that maintain it. This requires every effort to enlist the patient's commitment to actively participate in every stage of the treatment process.

There are several general guidelines for hospitalization including: (i) weight restoration or interruption of steady weight loss in patients who are emaciated, (ii) interruption of bingeing, vomiting, and/or laxative abuse that pose medical risks or complications, (iii) evaluation and treatment of other potentially serious physical complications, and (iv) management of associated conditions such as severe depression, risk of self-harm, or substance abuse. On rare occasions, hospitalization may be required to "disengage" a patient from a social system that both contributes to the maintenance of the disorder and disrupts outpatient treatment. Thus, as indicated above (guideline (ii)), the first major treatment consideration is whether or not the eating disorder patient is in sufficient medical danger to require hospitalization. If a careful evaluation of the patient's prior treatment and current motivation for treatment leads to the conclusion that they are clearly unresponsive to psychological treatment, then it is appropriate to consider hospitalization for medical stabilization only and referral to out-

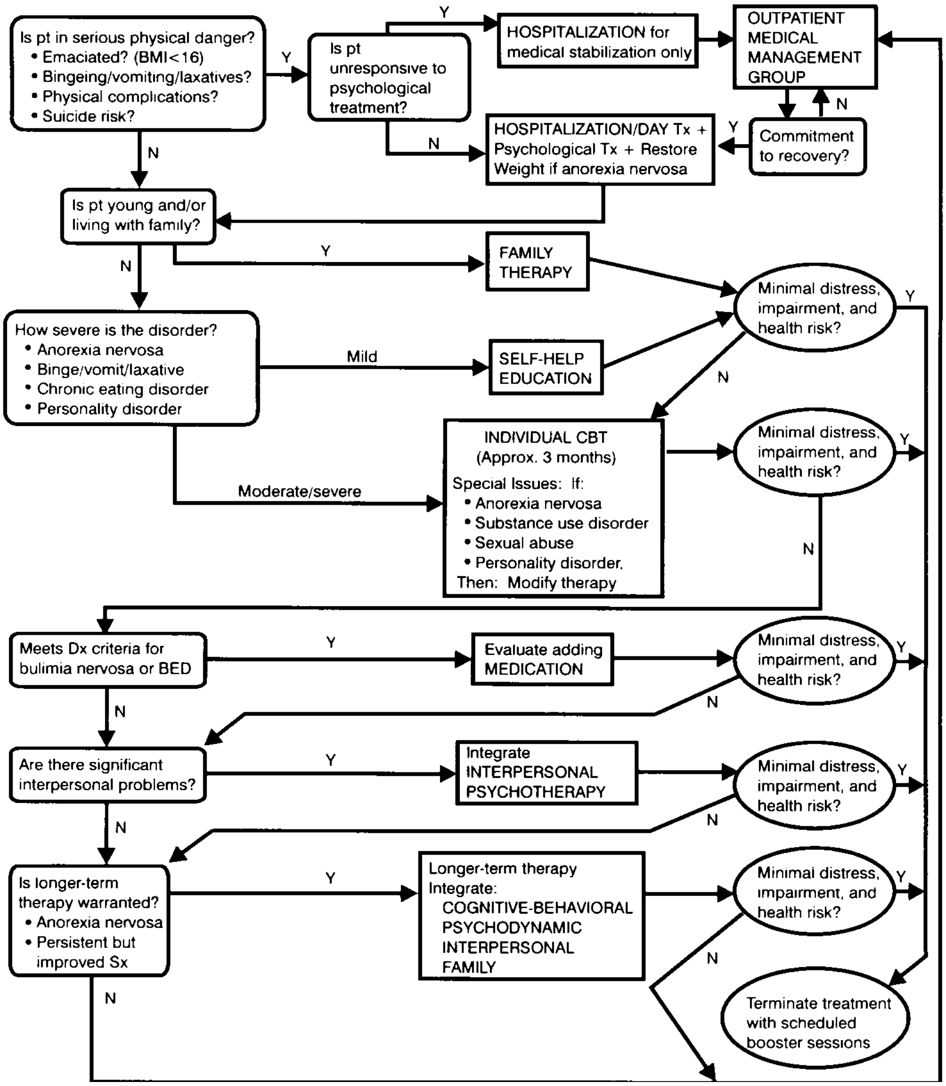
Is pt in serious physical danger?
• Emaciated? (BMI<16)
• Bingeing/vomiting/laxatives?
• Physical complications?
• Suicide risk?
Y
Is pt unresponsive to psychological treatment?
Y
HOSPITALIZATION for medical stabilization only
OUTPATIENT MEDICAL MANAGEMENT GROUP
N
HOSPITALIZATION/DAY Tx + Psychological Tx + Restore Weight if anorexia nervosa
Y
Commitment to recovery?
N
N
Is pt young and/or living with family?
Y
FAMILY THERAPY
N
Minimal distress, impairment, and health risk?
Y
How severe is the disorder?
• Anorexia nervosa
• Binge/vomit/laxative
• Chronic eating disorder
• Personality disorder
Mild
SELF-HELP EDUCATION
N
INDIVIDUAL CBT (Approx. 3 months)
Special Issues: If:
• Anorexia nervosa
• Substance use disorder
• Sexual abuse
• Personality disorder,
Then: Modify therapy
Moderate/severe
Minimal distress, impairment, and health risk?
Y
N
Meets Dx criteria for bulimia nervosa or BED
Y
Evaluate adding MEDICATION
Minimal distress, impairment, and health risk?
Y
N
N
Are there significant interpersonal problems?
Y
Integrate INTERPERSONAL PSYCHOTHERAPY
Minimal distress, impairment, and health risk?
Y
N
N
Is longer-term therapy warranted?
• Anorexia nervosa
• Persistent but improved Sx
Y
Longer-term therapy Integrate: COGNITIVE-BEHAVIORAL PSYCHODYNAMIC INTERPERSONAL FAMILY
Minimal distress, impairment, and health risk?
Y
N
Terminate treatment with scheduled booster sessions
N

Figure 3 General guidelines for integration and sequencing major treatment options for eating disorder patients. Begin in the upper left corner with the question "is the patient in serious physical danger?" Further criteria for decision making, including symptom areas, patient characteristics, or response to previous treatments, are indicated by boxes with rounded corners (mostly on the left side of the figure). The treatment options are represented by boxes with sharp corners (hospitalization, family therapy, self-help, individual therapy, etc.). The figure does not delineate all treatment alternatives or all considerations for decision making. However, it does include the main interventions for which there is good clinical and/or empirical support and key variables to consider in determining the most appropriate type and intensity of treatment. (*Handbook of Treatment of Eating Disorders* (pp. 50–66), by D. M. Garner and L. Needleman, 1997, New York: Guilford Press. Copyright 1997 by Guilford Press. Reprinted with permission.)

patient medical maintenance discussed below. More typically there is reason for optimism about recovery from the eating disorder. In these cases, hospitalization can usually be brief, again aimed at medical stabilization, but with the addition of psychological counseling and followed by referral to outpatient psychotherapy. The most common exception to brief hospitalization is the emaciated anorexia nervosa patient for whom longer hospitalization is usually required to make steady headway in the renourishment process. It is generally pointless to negotiate with patients or insurance carriers around the duration of hospitalization required for weight restoration, since the time needed is relatively straightforward and easy to calculate. It is the number of weeks or months required to reach at least 90% of expected weight, gaining at a rate of about three pounds a week and assuming optimal compliance with the treatment program. Even though this is a time-consuming and expensive process, it is an economical alternative if it leads to recovery since a chronic eating disorder inflicts heavy price, both in monetary and emotional terms. At the same time, it is legitimate to question the validity or timing of a lengthy hospitalization aimed at renutrition if steady weight gain does not occur.

Another situation in which hospitalization may be necessary is when a patient has had a protracted period of outpatient treatment and is simply unable to make improvement without the structure and containment offered by the inpatient environment. Indefinite outpatient treatment may be fruitless or even deleterious in the sense that it inadvertently colludes with the patient's ambivalence about recovery. It is like years of insight-oriented psychotherapy with an elevator phobic without ever requiring a ride on an elevator. Without some exposure, the therapist cannot ever assist the patient in confronting fears associated with weight restoration. The patient may perceive the recommendation for hospitalization as a threat or as abandonment, but they should be reassured that it is a humane alternative to the tremendous emotional and financial expense of prolonged and unproductive outpatient therapy. Psychotherapy can only achieve modest goals in the presence of the severe limits imposed by dietary chaos or chronic starvation.

Day treatment or partial hospitalization is the preferred alternative to inpatient care for most patients (Kaplan & Olmsted, 1997). This level of care provides structure around meal times plus the possibility for intensive therapy, without requiring the patient to become totally disengaged from the supports and the therapeutic challenges outside of the hospital. Day treatment programs offer the distinct advantage of being more economical than full hospitalization. They can also provide a useful bridge between inpatient and outpatient care. There are various models for day treatment programs which generally share many features with inpatient programs. The major difference is that patients receive the therapeutic services but do not stay overnight. Inpatient treatment is still the preferred modality for patients who are seriously emaciated, require close medical monitoring, fail to progress in partial care, or are at serious risk of self-harm.

5.23.7.3 Outpatient Medical Management

Concluding that a patient is unresponsive to psychological treatment and then referring to an outpatient medical management (Figure 3) requires careful analysis and considerable reflection. However, there are some chronic or recalcitrant patients who have participated in various forms of treatment over the course of years and finally reach the point where they no longer want treatment for their eating disorder or they can agree that further psychotherapy has a very low probability of success. In these circumstances, psychotherapy aimed at recovery can be highly frustrating for the therapist and the patient alike, often leading to termination with inadequate follow-up plans. Patients are then reluctant to seek further assistance when their medical or psychological condition deteriorates. The key is determining if there really has been sufficient exposure to truly adequate treatment since patients can have extensive exposure to poor treatment or therapists who are not knowledgeable about eating disorders. If careful evaluation leads to the conclusion that the patient has had an adequate course of psychotherapy and (i) is not in imminent psychological or medical danger, (ii) is sufficiently symptomatic to suggest ongoing potential medical risk, and (iii) there is no clear rationale for further psychotherapy, then termination from psychotherapy may be very appropriate as long as it is made conditional on the patient agreeing to participate in ongoing "medical management" (Figure 3). Again, the goal of medical management is very different from psychotherapy. Rather than focusing on overcoming the eating disorder, it is aimed at maintaining medical and psychological stability. It can be conducted in an individual or a group format, meeting weekly with medical supervision. Body weight, electrolytes, and vital signs should be checked with appropriate referral to medical specialists, as needed. Patients derive the added benefits of group support and sharing with similarly afflicted patients. In some cases, these groups set the stage for renewed efforts to actively address the eating disorder symptoms.

5.23.7.4 Self-help and Guided Self-help

There are eating disorder patients at the other end of the severity spectrum who have relatively mild symptoms and who may experience marked improvement or complete recovery with self-help or education-based interventions. Self-help organizations for eating disorders have been very popular since the late 1970s (Baker-Enright, Butterfield, & Berkowitz, 1985). From the mid-1980s, there has been a remarkable increase in number of self-help books available for eating disorder patients and their families (Santrock, Minnett, & Campbell, 1994). Interest in testing the efficacy of self-help has been slower to develop, but there are now a number of studies indicating that some forms of self-help or guided self-help may be quite effective with some eating disorders (see Fairburn & Carter, 1997). As indicated in Figure 3 self-help is recommended for patients with less severe eating disorders symptoms. While self-help is not sufficient as the sole treatment for most patients, there is a significant subgroup who benefit from this extremely economical intervention.

5.23.7.5 Psychoeducation

Figure 3 indicates that psychoeducation should be considered for bulimia nervosa patients with milder symptoms. As will be reviewed in a later section, there have been a number of research studies providing empirical support for the utility of educational approaches in the treatment of eating disorders. Education was originally recommended as part of early descriptions of cognitive behavioral treatment for eating disorders based on the premise that certain faulty assumptions evinced by patients were maintained, at least in part, by misinformation (e.g., Garner & Bemis, 1982). One early psychoeducational primer (Garner, Rockert et al., 1985) emphasized the conflict between the cultural pressures on women to diet and biological compensatory mechanisms which tend to defend a "set point" for body weight. The primer has been updated (Garner, 1997) and provides background for challenging society's harsh standards for women to be thin, education regarding the biology of weight regulation, recommendations for establishing regular eating patterns, and methods for interrupting bingeing and purging. There seems to be general agreement on the utility of incorporating psychoeducation into most treatment approaches for eating disorders. For people with mild eating disorder symptoms (Figure 3), guided self-help or educational groups should be considered as an economical and potentially effective option before proceeding to more costly interventions.

5.23.7.6 Family Therapy

Family therapy is the initial treatment choice for patients who are young and/or living at home (Figure 3). Family therapy should also be regarded as a desirable adjunct to individual therapy with older patients living outside of the home, particularly when family conflicts predominate. For young patients, there are practical as well as theoretical reasons for recommending family therapy. From a strictly practical point of view, the patient shares her home with her parents or guardians who are responsible for her well-being. Parents have the potential to provide powerful directives in support of therapeutic goals. Regardless of theoretical orientation of treatment, family members need assistance in dealing with a young eating disordered patient.

Family therapists have argued convincingly that eating disorders may reflect certain conflicts, dysfunctional roles, or problematic interactional patterns within the family (Dare & Eisler, 1997; Minuchin et al., 1978; Selvini-Palazzoli, 1974). The following are some of the more common problematic themes.

(i) Denial regarding the seriousness of the eating disorder may be manifest by the identified patient and/or the parents. Both the patient and parents may need assistance in accurately labeling the eating disorder which involves neither minimizing or exaggerating the significance of certain symptomatic behaviors.

(ii) Particularly with younger patients, parents may need help in developing an effective parenting style. Guilt and fear may have prohibited them from being firm and effective in establishing guidelines for behavior consistent with recovery. Parents should be encouraged to maintain usual expectations in areas unrelated to food and eating (bedtime, chores, language, treatment of siblings) except when unrealistic parental expectations are directly contributing to the problem. Treatment recommendations should be consistent with the family's value system and appropriate to the developmental stage of the child.

(iii) Parents may need assistance in accepting the patient's need for autonomy and self-expression through traits, interests, and values that may deviate from expectations within the family.

(iv) Under certain circumstances, parental attitudes about weight, shape, thinness, or "fitness" may be inappropriate. Inappropriate family eating patterns, beliefs about food or eating rituals should be identified, and practical interventions need to be designed that address areas of potential conflicts.

(v) An eating disorder can deflect members of the family away from potentially threatening developmental expectations emergent in the transition to puberty. It can function as a maladaptive solution to the adolescent's struggle to achieve autonomy in a family where any move toward independence is perceived as a threat to family unity. It can also become a powerful diversion enabling the parents and the child to avoid major sources of conflict.

(vi) The identified patient's symptoms may be functional within a disturbed family context and the meaning systems that underlie the resulting interactional patterns need to be identified and corrected. Problematic family interactional patterns need to be addressed such as "enmeshment," overprotectiveness, inadequate mechanisms for resolving conflicts, and inappropriate parent–child allegiances that undermine the marital relationship (Minuchin et al., 1978).

Even when family therapy is not the primary mode of treatment, it may be appropriate to integrate family therapy into the longer-term treatment plan. It can be a valuable adjunct to individual therapy in addressing trauma, such as sexual abuse, within the family (Andersen, 1985; Crisp, 1970; Garner et al., 1982; Wooley & Lewis, 1989).

5.23.7.7 Cognitive Behavioral Therapy

Cognitive behavioral therapy (CBT) has become the standard treatment for bulimia nervosa and forms the theoretical base for much of the treatment of anorexia nervosa. The empirical support for CBT in the treatment of eating disorders will be summarized in a later section. As indicated in Figure 3, CBT should be considered the treatment of choice for patients whose age does not mandate family therapy, and whose symptoms are moderate to severe.

The CBT for bulimia nervosa developed by Fairburn and colleagues (Fairburn, 1985; Fairburn, Marcus, & Wilson, 1993) has the following major points of emphasis: (i) self-monitoring of food intake, bingeing, and purging episodes as well as thoughts and feelings which trigger these episodes; (ii) regular weighing; (iii) specific recommendations such as the introduction of avoided foods and meal planning designed to normalize eating behavior and curb restrictive dieting; (iv) cognitive restructuring directed at habitual reasoning errors and underlying assumptions that are relevant to the development and the maintenance of the eating disorder; and (v) relapse prevention.

CBT is also effective in stabilizing the eating patterns of obese individuals who experience recurrent episodes of binge eating (Agras et al., 1992; Telch Agras, Rossiter, Wilfley, & Kenarely, 1990; Wilfley et al., 1993). The results with binge-eating disorder patients are promising; however, the long-term effectiveness of CBT with this group remains unclear.

In the case of anorexia nervosa, CBT has been recommended largely on clinical grounds (Garner, 1986, 1988; Garner & Bemis, 1982; 1985; Garner & Rosen, 1990; Garner et al., 1997; Vitousek & Orimoto, 1993). Case studies and preliminary research provide some grounds for optimism (Channon, DeSilva, Hemsley, & Perkins, 1989; Cooper & Fairburn, 1984); however, current data are insufficient to warrant meaningful conclusions regarding effectiveness.

There are many areas of overlap between the versions of CBT offered for anorexia and bulimia nervosa; however, there are also important differences that have implications for clinical care (Garner et al., 1997). A major focus for both disorders is the patient's underlying assumption that "weight, shape, or thinness can serve as the sole or predominant referent for inferring personal value or self-worth" (Garner & Bemis, 1982, p. 142). Fear of body-weight gain is a central theme for both anorexia and bulimia nervosa; however, most bulimia nervosa patients can be reassured that treatment will probably result in little weight gain. In contrast, in anorexia nervosa, therapeutic strategies must be aimed at actual weight gain in the face of the implacable wish to maintain a low weight. Establishing a sound and collaborative therapeutic relationship is particularly important in anorexia nervosa since it becomes a fulcrum for gradually helping the patient to relinquish the myriad of ego-syntonic symptoms. As indicated on Figure 3, CBT is modified for anorexia nervosa. It is typically longer in duration and the targets for cognitive interventions are broad, encompassing a wider range of personal and interpersonal subject domains than typical for bulimia nervosa (Garner et al., 1997). This is to address the marked deficits impeding normal psychosocial functioning in many of these patients. Thus, cognitive interventions are aimed not only at beliefs that maintain extreme dieting and chronic weight suppression but also at fundamental assumptions associated with interpersonal conflicts, feelings of ineffectiveness, struggles with autonomy, and fears associated with psychosocial development (Garner & Bemis, 1982; Garner & Rosen, 1990; Garner et al., 1997; Vitousek & Orimoto, 1993; Vitousek & Ewald, 1993). Components of interpersonal therapy and family therapy are often integrated into longer-term CBT for anorexia nervosa (see Figure 3). In the subset of anorexia nervosa patients who do not engage in binge eating, and those who show no obvious serious physical complications, there may be extraordinary resistance to complying with the therapeutic objectives of weight gain. However, as long as there is gradual improvement in symptoms (Figure 3), outpatient therapy is recommended; however, if the patient's condition deteriorates, hospitalization must be considered. Outpatient medical management (groups) should be considered if there is a protracted therapeutic impasse. This often takes the form of apparent psychological insights that invariably fail to translate into symptomatic change.

5.23.7.8 Interpersonal Therapy

Since the mid-1980s, the prevailing view was that CBT's effectiveness with eating disorders was tied to cognitive and behavioral methods aimed specifically at overconcern about weight and shape which are responsible for restrictive dieting and extreme weight controlling behaviors. Changing these attitudes presumably relaxed restrictive dieting and relieved the biological tension created by chronic attempts to suppress body weight. However, a series of studies using interpersonal psychotherapy (IPT), adapted to bulimia nervosa, has prompted re-examination of earlier speculations regarding the specific mechanisms of action in the treatment of binge eating (Fairburn et al., 1995). This is because IPT does not directly focus on eating problems. IPT was originally proposed by Klerman, Weissman, Rounsaville, and Chevron (1984) as a short-term treatment for depression. The IPT treatment process is divided into three stages (Fairburn, 1997). The first stage involves the identification of the interpersonal problems that led to the development and maintenance of the eating problems. The second stage consists of a therapeutic contract for working on these interpersonal problems. The final stage addresses issues related to termination. Fairburn et al. (1991) found IPT somewhat less effective than CBT at the end of treatment; however, patients who received IPT gradually improved during the follow-up period so that after one year, both treatments were equally effective (Fairburn et al., 1993). These findings are maintained over the longer term with patients recieving CBT or interpersonal therapy doing significantly better than those receiving behavior therapy (Fairburn et al., 1995). This pattern of improvement during follow-up was not found in a study of another very different form of interpersonally oriented therapy, supportive-expressive therapy (Garner et al., 1993). These findings suggest that the IPT used by Fairburn et al. contains specific therapeutic ingredients that facilitate change. Further support for the effectiveness of IPT comes from Wilfley et al. (1993) in a study of nonpurging bulimic patients, many of whom presented with obesity. They found both CBT and IPT equally effective in reducing binge eating assessed both at the end of treatment and at one year follow-up.

The evidence that certain interpersonal therapies are as effective with binge eating as CBT has implications for the decision process illustrated in Figure 3. It could be argued that there should be no priority for either IPT or CBT as the initial treatment of choice for bulimia nervosa since both treatments are equally effective in the long-term. Certainly, a therapist well trained in one of the interpersonal therapies described by Fairburn and colleagues should not be encouraged to abandon this form of treatment in favor of CBT. However, we are still inclined to recommend CBT as the preferred initial treatment at this time since it has been shown to have a more rapid effect on symptoms. Moreover, the efficacy of IPT for bulimia nervosa has been demonstrated in just one center compared to many studies in support of CBT. If the findings from the Oxford trials are replicated in other centers, then IPT may become another "standard" initial treatment for bulimia nervosa. At this time, there is no empirical basis for suggesting that IPT should be differentially applied to patients on the basis of premorbid features such as interpersonal conflicts. However, integrating IPT into treatment should be considered for bulimia nervosa patients who fail to respond favorably to an initial course of CBT when interpersonal conflicts predominate according to Figure 3. The IPT orientation should also be a leading candidate for integration into longer term psychotherapy for anorexia nervosa patients or for others with persistent symptoms.

5.23.7.9 Psychodynamic Therapy

There have been no controlled comparisons between longer-term psychodynamic psychotherapy and other forms of treatment; however, dynamically oriented treatments have been well articulated for eating disorders and may be recommended on clinical grounds when other short-term treatments are ineffective (Figure 3). Psychodynamic treatment for eating disorders may be roughly divided into two schools of thought. The first presumes eating disorders do not require fundamental modifications to orthodox dynamic interventions since neither the symptoms nor the disorders represent a unique underlying process. The second conceptualization implies that eating disorders are distinctive in that they require major modifications to traditional dynamic therapy in order to meet the special psychological needs of the patient. There has been considerable movement among dynamically oriented writers toward integrating psychodynamic therapy with active symptom management principles in the treatment of both anorexia nervosa and bulimia nervosa (Bruch, 1973; Casper, 1982; Crisp, 1980, 1997; Goodsitt, 1985, 1997; Stern, 1986; Strober, 1997; Strober & Yager, 1985). However, there are still some who espouse a traditional interpretive framework (Lerner, 1993; Sands, 1992; Sugarman, 1991). At this time, we recommend the integration of longer-term dynamically oriented therapy be considered as an alternative to long-term CBT for patients who fail less expensive approaches.

A new generation of feminist therapists deviate from the traditional views by incorporating the growing literature on the psychology of women into dynamic formulations regarding the etiology and treatment of eating disorders (Fallon, Katzman, & Wooley, 1994; Fallon & Wonderlick, 1997; Katzman, 1994; Kearney-Cooke & Striegel-Moore, 1997; Perlick & Silverstein, 1994; Steiner-Adair, 1994; Striegel-Moore, 1994). They have highlighted the importance of addressing issues such as sexual abuse and other forms of victimization in the development, maintenance, and treatment of eating disorders. Even though controlled studies are lacking at this time, the highly evolved conceptual framework presented by these dynamic-feminist treatment approaches warrants their consideration when abuse is identified during the course of assessment or treatment (Figure 3). At this time, the therapist's orientation, rather than the special needs of the patient, probably determine the decision to apply feminist therapy for eating disorders. However, many of the views articulated by feminist therapists have influenced the applications of other forms of therapy. There are some eating disorder patients, because of the nature of their predisposing and perpetuating factors, who should receive feminist therapy as the treatment of choice.

5.23.7.10 Pharmacotherapy

Figure 3 indicates medication should be considered for patients with bulimia nervosa or binge-eating disorder who fail an initial trial of cognitive behavioral therapy but is generally not indicated for anorexia nervosa. There have been many well-controlled trials since the mid-1980s indicating the effectiveness of some antidepressant medications for bulimia nervosa and the research evidence will be reviewed in detail in a later section. However, in reviewing the research in the field, Raymond et al. (1994) have suggested that medication should not be "the primary mode of therapy with patients with bulimia nervosa" (p. 241). This conclusion is based on the following observations: (i) psychological interventions have been shown to be very effective, (ii) there are high drop-out rates reported in most medication studies, (iii) there are risks of drug side effects, and (iv) data suggest high relapse rates with drug discontinuation. Based on a multisite collaborative study (Fluoxetine Bulimia Nervosa Collaborative Study Group [FBNC], 1992), fluoxetine hydrochloride (Prozac) is currently the first choice for the treatment of bulimia nervosa (daily dosages of 60 mg were generally superior to 20 mg) and probably should be used at least as an adjunct to psychotherapy in many cases failing in a course of adequate psychological treatment. Tricyclic antidepressants may be an alternative for some patients failing a course of fluoxetine, but side effects, high drop-out rates, and greater lethality with overdose may be a source of concern when treating some patients (Leitenberg et al., 1994; Mitchell et al., 1990). Similarly, monoamine oxidase inhibitors (MAOIs) may be useful in a small minority of patients who fail using fluoxetine and tricyclics. There is some evidence that patients who fail on one tricyclic may benefit from changing to an alternative medication (Mitchell, Raymond, & Specker, 1993). Alternatively, the specific decision to not prescribe medication may be therapeutic in some cases (Raymond et al., 1994).

There is still little evidence for change in the early recommendations that pharmacotherapy has a very limited value with emaciated anorexia nervosa patients and should never be the sole treatment modality (Garfinkel & Garner, 1982). Occasionally patients may benefit from medication to deal with overwhelming anxiety, severe depression or intolerable gastric discomfort after meals, but this only applies to a small minority of patients (Andersen, 1985; Garfinkel & Garner, 1982; Garfinkel & Walsh, 1997).

Thus, in the decision-tree approach for treating eating disorders, antidepressant medication should be considered as an adjunct or possibly an alternative for bulimia nervosa patients who fail to respond to psychosocial therapies or for those whose affective symptoms are clearly impeding progress in other forms of treatment. Drug treatments have generally not proven effective for anorexia nervosa.

5.23.8 PSYCHOLOGICAL TREATMENT RESEARCH

5.23.8.1 Treatment Research on Anorexia Nervosa

Evaluation of treatment response would be aided by an understanding of the natural history of the disorder. While there are no studies of the actual natural history of anorexia nervosa, Theander (1985) has provided an estimate from a long-term follow-up study of 94 patients seen in the early 1960s and thus, probably received treatment that was quite crude compared to modern methods. The patients were evaluated in 1966 (after a mean observation time of 15 years since the onset of the disorder) and again in 1984 (after a mean observation time of 33 years). Table 1 presents outcome status for the sample five, 15, and 33 years after the onset of the disorder and indicates an aggregation over time of patients into the recovered and deceased categories. Of the patients who died, one-third committed suicide and two-thirds died of complications from their disorder. The data in Table 1 dramatically illustrate that the evaluation of outcome in anorexia nervosa varies depending on when the assessment is performed and that, after many years of illness, patients tend either to improve or to die.

Since 1953 there have been more than 60 follow-up studies of anorexia nervosa (Eckert, Halmi, Marchi, Grove, & Cosby, 1995; Fichter & Quadflieg, 1995; Steinhausen, Rauss-Mason, & Seidel, 1991); however, there are few controlled trials comparing different treatment approaches. There is considerable variation in the outcomes reported across these follow-up reports which could be expected given differences among methodologies and treatment settings.

The major outcome criteria used in these follow-up studies have been diagnosis, weight, menses, eating behavior, general psychological functioning, and psychosexual status. Usually a global outcome (good, moderate, poor, death) is derived from some composite of these variables. The mortality rates have varied from 0% to 22%, with over half of the studies reporting deaths in 4% or fewer patients (Garfinkel & Garner, 1982; Herzog, Keller, Lavoni, & Ott, 1987; Sullivan, 1995). The mortality rates for anorexia nervosa are two times higher than for any other psychiatric disorder and more than 12 times higher than all-cause mortality for females. Although the definition of "recovery" has been ambiguous, between 12% and 87% of patients appear to be free of eating disorder symptoms at follow-up (Hsu, 1996). Consistent with the report by Theander (1985), studies of relatively brief follow-up find fewer patients who are recovered or deceased, with most patients congregating in the moderate to poor outcome. More patients recover and the number of deaths increase as the follow-up is extended.

Some patients can be considered fully recovered at follow-up while others who recover in terms of weight and eating remain disturbed in other areas of functioning. Casper (1990) reported that about 25% of patients classified as having a good outcome in an 8–10-year follow-up were still weight preoccupied and depressed. In a study of 60 anorexia nervosa patients followed 5–14 years after presenting for treatment, Toner, Garfinkel, and Garner (1986) found that 38% were asymptomatic, 27% improved, 27% unchanged, and 8% deceased. While a significant proportion of the sample met *DSM-III* criteria for affective (34%) or anxiety disorders (47%) in the year prior to follow-up, those in the asymptomatic group had scores equivalent to noneating disorder controls on most measures of psychosocial functioning. Rastam et al. (1995) followed a sample of anorexia nervosa patients six years after onset and found that most cases no longer met criteria for the original eating disorder, but many developed bulimia nervosa or atypical eating disorders. In a 12-year follow-up study of 69 hospitalized anorexia nervosa patients, Herzog, Schellberg, and Deter (1997) found that 50% of patients did not improve until more than six years after the first inpatient treatment.

One of the most striking observations regarding outcome is the relatively good prognosis for pediatric compared with older samples. In a summary of pediatric studies, Garfinkel and Garner (1982) indicated that the outcome status of patients in follow-up studies conducted before 1974 were quite similar to that reported in the literature on older adolescents and adults. However, later pediatric outcome studies indicate a very positive prognosis with recovery rates of between 76% and 87% (Cantwell, Sturzenburger, Burroughs, Salkin, & Green, 1977; Goetz, Succop, Reinhart, & Miller, 1977; Martin, 1985; Minuchin et al., 1978; Silverman, 1974). Steinhausen and Glanville (1983) present somewhat less optimistic findings from a 9.4-year follow-up of patients presenting between the ages of 11 and 17. They report that only 30% of their sample could be considered symptom free at follow-up. Similarly, only 50% of the early onset cases ($N = 18$) followed by Hawley (1985) after a mean of 8.7 years were classified as having a good outcome. Theander (1996) reviewed the literature on early onset anorexia nervosa and

Table 1 Long-term outcome in anorexia nervosa.

	Years after onset of the disorder		
Outcome status	*5 years* (%)	*15 years* (%)	*33 years* (%)
Good	55	63	76
Intermediate	19	17	1
Poor	18	7	6
Death	8	13	18

Source: Theander (1985), adapted by permission.

found that there are relatively more boys among patients with extremely early onset. He also reported that early onset predicted better outcome, but long-term outcome is unfavorable for many early onset cases.

Follow-up reports indicating that early age of onset confers a positive prognosis (cf. Garfinkel & Garner, 1982; Herzog, Keller, & Lavori, 1988) as well as the variable, but generally positive, outcome in a number of pediatric samples has led to the interest in family therapy in research studies on anorexia nervosa. In the first controlled trial examining two forms of treatment, Russell et al. (1987) compared family and individual supportive therapy for both anorexia nervosa and bulimia nervosa. It was concluded that family therapy was superior for patients whose disorder was not chronic and that had begun before the age of 19. A somewhat more tentative finding was that individual therapy yielded better results with older patients.

In another study of family treatment, Robin, Siegel, and Moye (1995) randomly assigned 22 adolescent anorexia nervosa patients to 16 months of either behavioral family systems therapy or ego-oriented individual therapy and found that both treatments had a positive effect on body mass index, menstruation, and family interactional patterns. There was no consistent pattern of differential treatment effects; however, the relatively small sample size used in the comparison makes it difficult to detect treatment differences.

Although family intervention was not the object of the study, Crisp et al. (1991) provided collateral evidence for the utility of family intervention in a controlled study in which 90 anorexia nervosa patients were randomly assigned to one of three types of treatment (or referral back to the community). The three treatment conditions were: (i) intensive inpatient treatment followed by 12 sessions of outpatient therapy, (ii) 12 outpatient individual and family psychotherapy sessions, and (iii) 10 outpatient group psychotherapy sessions (both for patients and parents separately). All three active treatments were accompanied by dietary counseling aimed at gradual weight restoration. Outcome was assessed at the end of treatment and at one-year follow-up. Of the 20 control group patients referred back to the community, 14 received some form of treatment and six were untreated. At one year, there were no significant differences between the three active treatment conditions in weight gain, return of menses, or important aspects of psychosocial functioning. Surprisingly, although the inpatient treatment was more intensive (several months of inpatient care followed by 12 outpatient sessions), presumably much more costly, and delivered as the treatment of choice by Crisp and colleagues for years, it was no more successful than the strictly outpatient approaches. In terms of weight gain at follow-up, inpatient treatment was not superior to the controls. Moreover, the inpatient treatment had a much higher rate of noncompliance (40%) compared to the outpatient conditions (10–15%). A two-year follow-up of this same sample (Gowers, Norton, Halek, & Crisp, 1994) indicated that the positive outcome with the outpatient treatment package was sustained. Thus, the results of this major study challenge the assumption that "more treatment is better" and that inpatient treatment is the treatment of choice for anorexia nervosa.

More evidence for the efficacy of less intensive treatment comes from a study by Treasure et al. (1995) in which 30 anorexia nervosa patients were randomly assigned to either cognitive analytical therapy (CAT) or educational behavioral therapy. At one year, both of these outpatient treatments led to significant weight gain. Patients assigned to the CAT reported significantly greater subjective improvement; however, there were no treatment group differences in other outcome parameters. All these findings are summarized in Table 2.

The relatively small number of carefully controlled trials of treatments for anorexia nervosa may relate, in part, to the protracted, complicated, and variable course in the treatment of the disorder. Many cases require one or more periods of hospitalization at various points during a lengthy and fluctuating recovery process. Since this may require numerous transitions between inpatient and outpatient treatment staff over an extended period of time in response to the patient's changing nutritional status, it is difficult to design a rigorously controlled study protocol suitable for the range of patients presenting for treatment.

5.23.8.2 Treatment Research on Bulimia Nervosa

In contrast to the paucity of controlled trials with anorexia nervosa, there have been more than 25 controlled studies evaluating the efficacy of different treatment approaches to bulimia nervosa (for reviews see Garfinkel & Walsh, 1997; Garner, Fairburn, & Davis, 1987; Mitchell, Raymond, & Specker, 1993). The majority of these studies have involved comparing different forms of treatment or identifying active ingredients within the more successful treatment packages (primarily cognitive-behavioral) using a dismantling approach. Most

Table 2 Comparative treatment studies for anorexia nervosa.

Investigator	*Orientation/groups*	*N*	*Format/duration*[a] (weeks)	*Dropouts* (%)	*Pretreatment BMI*	*Follow-up BMI*	*Mean weight gain post-treatment* (kg)	*Mean weight gain at follow-up* (kg)	*Duration of follow-up*
Growers et al. (1994)	Individual and family behavioral/dynamic	20	I & F (40)	10	15.5 (1.4)	19.0 (2.0)–12 mos.	?	?	12 months
						20.0 (2.8)–24 mos.	?	12.0 (8.4)	24 months
	"No treatment" control[b]	20		?	15.8 (1.7)	17.0 (2.8)–12 mos.	?	?	
						17.8 (3.2)–24 mos.	?	5.2 (6.3)	
Robin et al. (1995)	Behavioral family systems	11	F (54–80)	?	15.0 (1.4)	20.1 (1.1)	5.1 (1.6)	5.0 (1.6)	12 months
	Ego-oriented therapy	11	I (54–80)	?	16.3 (2.8)	19.0 (1.4)	2.7 (2.2)	2.6 (2.2)	
Treasure et al. (1995)	Education behavioral	16	I (20)	38	15.0 (1.0)	17.2 (3.2)	?	5.6 (4.5)	12 months
	Cognitive analytic	14	I (20)	29	15.6 (2.1)	18.2 (2.1)	?	6.9 (4.6)	

? = not reported

[a] F = Family; I = Indivudual. [b] Some participants may have been in treatment elsewhere.

treatment packages have been between 10 and 18 weeks in duration, delivered in either a group or an individual format (cf. Garner et al., 1987; McKisack & Waller, 1997). It is evident from these and other studies that treatment can lead to statistically and clinically significant reductions in bingeing, vomiting, and other associated symptoms.

A number of comparative treatment studies of bulimia nervosa that provide very encouraging findings are summarized in Table 3 (see Mitchell et al., 1993 for earlier studies). It is evident from this that CBT, as well as certain other forms of treatment with which it has been compared, lead to marked reductions in bingeing and vomiting. In a long-term prospective follow-up of 91 eating disorder patients from two randomized controlled trials involving CBT, behavior therapy (BT), and focal interpersonal therapy (FIT), Fairburn et al. (1995) found that those received BT did poorly while bulimia nervosa patients who received CBT or FIT did markedly better. Compared to subjects in the BT condition, those in the FIT were twice as likely to be in remission (defined as no longer meeting a *DSM-IV* diagnosis for an eating disorder) and those in the CBT condition were more than three times as likely to be in remission. Similarly, about half of the patients who received CBT or FIT were completely abstinent from all key behavioral symptoms such as binge eating, vomiting, and laxative and diuretic abuse compared to 18% of those receiving BT.

Agras and colleagues (1992, 1994) compared desipramine, CBT, and their combination (with the drug withdrawn at either 16 or 24 weeks) and found that the combined drug and CBT treatment was superior to the other conditions as long as the medication was continued for at least 24 weeks. The finding that the combined treatment was more effective than 16 and 24 weeks of medication alone suggests that the CBT may be important in preventing relapse once medication is withdrawn. In the 1994 follow-up study, both the CBT and the combined 24-week treatment were superior in reducing binge eating to desipramine given for 16 weeks. This study supports earlier research indicating that medication alone is not the most effective treatment for bulimia nervosa but, contrary to an earlier study (Mitchell et al., 1990), it provides important data suggesting that antidepressant therapy may add significantly to the effects of CBT.

There have been a number of empirical studies of the effectiveness of educational and self-help treatments in bulimia nervosa. Educationally-oriented treatment seems to have remarkable impact on behavior for some patients, although the overall treatment results are not as potent as some of the more intensive CBT interventions (Connors, Johnson, & Stuckey, 1984; Huon & Brown, 1985; Wolchik, Weiss, & Katzman, 1986). Ordman and Kirschenbaum (1985) compared CBT to a "brief intervention wait-list" that consisted of just three assessment sessions and minimal intervention. A study comparing the presentation of this educational material in a group "classroom" format (five 90-minute sessions during a one-month period) with individual CBT indicated, for the least symptomatic 25–45% of patients in the sample, both treatments were equally effective on important measures of outcome (Olmsted, Davis, Garner, Eagle, & Rockert, 1991). Nutritional counselling or nutritional management (Beaumont, O'Connor, Touyz, & Williams, 1987) overlaps with the psychoeducational approach described above. Laessle et al. (1991) compared nutritional management (NM) with stress management and found both treatments led to marked improvement in eating symptoms and general psychopathology at the end of treatment and at a 12-month follow-up. However, NM produced a significantly more rapid reduction in binge eating, greater abstinence rates, and greater improvements on some measures of psychopathology.

Huon (1985) was the first to test the effectiveness of a self-help program. Binge eating subjects were recruited from a woman's magazine article and 32% who followed a self-help manual were symptom free at six-month follow-up compared to only 7% of a comparison group. Two later studies have shown short-term benefits of following a self-help manual but neither presented follow-up data (Schmidt et al., 1993; Treasure et al., 1994). In contrast to "pure" self-help, Cooper, Coker, and Fleming (1995) have shown that self-help, supervised by a professional, can be sufficient for some patients. At a four- to six-month follow-up of 82 patients, 16 (20%) had dropped out of the program, and 22 (33%) of those remaining had not binged or vomited over the past month (Cooper et al., 1995). Thus, while self-help is not sufficient as the sole treatment for most patients, there is a significant subgroup who benefit from this extremely economical intervention.

While there has been overall optimism regarding the efficacy of CBT, it again must be emphasized that improvement also has been observed with other forms of treatment. Significant reductions in bingeing and vomiting, as well as associated symptoms, have been found with nondirective (Kirkley, Schneider, Agras, & Bachman, 1985), short-term focal (Fairburn,

Table 3 Comparative treatment studies for bulimia nervosa

Investigator	*Orientation/groups*	*N*	*Format/duration*[a] (weeks)	*Dropouts* (%)	*Reduction pre- to post-treatment* %		*Abstinent at last week of treatment* %		*Abstinent at follow-up* %		*Duration of follow-up*
					Binge	*Vomit*	*Binge*	*Vomit*	*Binge*	*Vomit*	
Agras et al. (1992)	Cognitive-behavioral	23	I (24)	4	82	83	50	48	55	?	32 weeks
	Desipramine	12	(16)	?	39	43	35	33	?	?	
	Desipramine	12	(24)	?	?	?	?	?	42	?	
	Cognitive-behavioral and desipramine	12	I (16 meds, 24 CBT)	17	?	83	81	65	64	?	
	Cognitive-behavioral and desipramine	12	I (24 meds, 24 CBT)	?	?	?	?	?	70	?	
Wolf and Crowther (1992)	Behaviour therapy	15	G (8)	0	(only report biweekly mean frequencies)						1 & 3 mos.
	Cognitive-behavioral	15	G (8)	0							
	Wait list control	11	NA	9							
Fairburn et al. (1993)	Cognitive-behavioral	25	I (18)	20	?	?	44 (binge & vomit)		36 (binge & vomit)		12 months
	Interpersonal therapy	25	I (18)	32	?	?	35 (b & v)		44 (b & v)		
	Behavior therapy	25	I (18)	48	?	?	26 (b & v)		20 (b & v)		
Gamer et al. (1993)	Cognitive-behavioral	30	I (18)	17	73	82	?	36	–	–	NA
	Supportive-expressive	30	I (18)	69	62	?	12	–	–		
Thackewray et al. (1993)	Cognitive-behavioral	Total *N* – 47	I(8)	Total – 17	?	?	92 (b & v)		69 (b & v)		6 months
	Behavior therapy		I (8)		?	?	100 (b & v)		38 (b & v)		
	Attention placebo		I (8)		?	?	69 (b & v)		15 (b & v)		
Agras et al. (1994)[b]	Cognitive-behavioral	23	I (24)	1					54 (b & v)		12 months
	Desipramine	12	(16)	1					18 (b & v)		
	Desipramine	12	(24)	3					67 (b & v)		
	Cognitive-behavioral and desipramine	12	I (16 meds, 24 CBT)	2					40 (b & v)		
	Cognitive-behavioral and desipramine	12	I (24 meds, 24 CBT)	3					78 (b & v)		
Leitenberg et al. (1994)	Cognitive-behavioral	7	I (20)	14	?	?	?	71	?	57	6 months
	Desipramine	7	(20)	57	?	?	?	0	?	0	
	Cognitive-behavioral and desipramine	7	I (20)	29	?	?	?	29	?	29	

Table 3 (continued)

					Reduction pre- to post-treatment %		*Abstinent at last week of treatment %*		*Abstinent at follow-up %*		*Duration of follow-up*
Investigator	*Orientation/groups*	*N*	*Format/duration*[a] (weeks)	*Dropouts* (%)	*Binge*	*Vomit*	*Binge*	*Vomit*	*Binge*	*Vomit*	
Treasure et al. (1994)	Cognitive-behavioral	21	I (8)	33	?	?	35	29			
	Self-care manual	41	S (8)	34	?	?	31	24			
	Wait list control	19		42	?	?	17	15			
Cooper and Steere (1995)	Cognitive-behavioral	15	I (18)	13	78	83	46	54	50	50	12 months
	Exposure w/response prevention	16	I (18)	13	79	91	50	43	15	8	
Fairburn et al. (1995)[c]	Cognitive-behavioral	35	I (18)	?					50 (b & v)		M = 5.8 years
	Behavioral	22	I (18)	?					18 (b & v)		
	Focal interpersonal	32	I (18)	?					52 (b & v)		
Jager et al. (1996)	Analytic inpatient	37	IP	14	?	?	34 (b & v)		66 (b & v)		38 months
	Systemic outpatient	46	I (?)	21	?	?	26 (b & v)		44 (b & v)		
Treasure et al. (1996)	Cognitive-behavioral	55	I (16)	27	?	?	30 (b & v)		41 (b & v)		18 months
	Self-care manual and cognitive-behavioral (if needed following self-care)	55	S (8) & I (8)[d]	16	?	?	30 (b & v)		40 (b & v)		

? = not reported

[a] G = group; I = individuals; S = self-care manual; IP = inpatient. [b] Reports on follow-up of Agras et al. (1992). [c] Report of follow-up from Fairburn et al. (1986, 1991). [d] Additional cognitive-behavioral therapy needed for 39 of 55 participants, median sessions = 3.

Kirk, O'Connor, & Cooper 1986; Fairburn et al., 1995), and supportive–expressive therapy (Garner et al., 1993). Understanding the active ingredients of some of these treatments is impeded by the fact that some included techniques such as self-monitoring, meal planning, education regarding the untoward effects of dieting, and behavioral contracting, which are generally considered as more conceptually aligned with CBT. Thus, it remains unclear whether the observed changes in behavior are attributed to the "psychological" components of the treatment or to the dietary management. In a comparison of bulimia nervosa treatment studies using dietary management with those that did not, Laessle, Zoett, and Pirke (1987) concluded that dietary management was associated with superior results. Further support for this hypothesis comes from a study indicating very poor outcome with a 12-week psychodynamically oriented group program that did not incorporate dietary management (Frommer, Ames, Gibson, & Davis, 1987). Nevertheless, there have been two studies of treatments for bulimia nervosa that did not emphasize dietary management (Fairburn et al., 1995; Garner et al., 1993) indicating that dietary management may not be essential for change in eating disorder symptoms.

In summary, a number of carefully controlled studies demonstrated that many patients derive benefit from either individual or group treatment and that CBT tends to perform better than drug and other psychological treatments. Dietary management may enhance the effectiveness of treatments of varying orientations. Since it is evident that not all patients improve with CBT, it is important to consider other treatment options, perhaps sequenced in a systematic manner, as part of a multidimensional model of treatment.

5.23.9 PSYCHOPHARMACOLOGICAL TREATMENT RESEARCH

A wide variety of medications have been tried with eating disorders and while they can play a role in helping some people, they should rarely become the exclusive mode of treatment (Garfinkel & Garner, 1987; Garfinkel & Walsh, 1997; Raymond et al., 1994). According to the decision-tree approach for treating eating disorders (Figure 3), antidepressant medication should be considered as an adjunct or possibly an alternative for bulimia nervosa patients who fail to respond to psychosocial therapies or for those whose affective symptoms are clearly impeding progress in other forms of treatment. Antidepressants have an even more limited role in the treatment of anorexia nervosa. With this general recommendation in mind, an overview of research on drug treatments will presented.

5.23.9.1 Treatment Research on Anorexia Nervosa

Historically, a variety of medications have been studied for their efficacy in promoting weight gain in women with anorexia nervosa. The rationale for trying many of the medications was based more on clinical observations than on a thorough understanding of the physiological mechanisms involved with this disorder. Unfortunately, we are still at a novice stage in our understanding of the physiological mechanisms that may contribute the development and maintenance of anorexia nervosa. Thus, although many medications have been tried in treating anorexia nervosa, few offer promise.

One class of medications that have been studied in the treatment of anorexia nervosa is antipsychotics. The rationale behind their use was the belief that the disturbed thinking regarding weight and shape, a cardinal feature of this disorder, may be a form of delusional thinking. Also, it was noted that dopamine mediates feeding and often people taking dopamine antagonists will experience weight gain. However, two double blind, placebo-controlled trials utilizing antipsychotic (pimozide and sulpiride) medications showed no statistically significant effects on weight (Vandereycken, 1984; Vandereycken & Pierloot, 1982).

A depressed mood is often associated with anorexia nervosa, and, as mentioned earlier, this is due most often to the effects of starvation. Given this, and that people treated with tricyclic antidepressants often experience an increase in appetite and weight gain, antidepressants have been utilized in treating anorexia nervosa. The results of open trials with antidepressants were promising; however, double-blind placebo-controlled studies yielded mixed results. Lacey and Crisp (1980) and Biederman et al. (1985) did not find significant effects on weight gain with either clomipramine or amitriptyline, respectively. Halmi, Eckert, LaDa, and Cohen (1986), however, did find that patients taking amitriptyline gained weight faster than those on a placebo. Substantial side effects from the amitriptyline were also noted.

One study reports on the effectiveness of lithium in treating anorexia nervosa (Gross et al., 1981). Again, the rationale for trying lithium was based on clinical observations of mood disturbances with anorexia nervosa and the weight gain experienced by many people who are prescribed lithium. The results of this study showed that the group of women with anorexia nervosa prescribed lithium did gain weight at a slightly faster rate, but that this was not statistically significant. Additionally, use of lithium may be problematic with this population due to frequent fluid imbalances if the person is purging.

Cyproheptadine, a serotonin and histamine antagonist, often used to treat allergies, has also been studied for its efficacy in treating anorexia nervosa. Cyproheptadine is a medication that is relatively free of side effects and promotes weight gain in some people who are prescribed it for allergies. Additionally, inhibiting serotonin transmission is believed to stimulate appetite. Thus, three controlled trials using cyproheptadine in treating anorexia nervosa were undertaken in the 1970s and 1980s. Two of the studies showed no significant effect on weight gain over placebo (Goldberg, Halmi, Eckert, Casper, & Davis, 1979; Vigersky & Loriaux, 1977). Halmi et al. (1986), utilizing a higher dosage, found that women with anorexia nervosa treated with cyproheptadine reached their goal weights an average of 10 days faster than those on placebo. This result was true only for the participants with the restricting subtype of anorexia nervosa, however, and the cyproheptadine was actually associated with delayed weight gain in those participants with the binge-purging subtype of anorexia nervosa. Moreover, there is little evidence to suggest that speed of weight gain is positively related to long-term recovery from anorexia nervosa.

Thus, the controlled trials of various medications designed to increase weight gain in women with anorexia nervosa have been generally disappointing. The practice guidelines for treating anorexia nervosa published by the APA (1993), indicate that medications should not be used routinely with this population. If an individual continues to experience mood disturbances following weight restoration, it is recommended that antidepressants be considered. Given that this is a volume that addresses clinical issues in children and adolescents, it should be noted that to date, no controlled trials of any medications reviewed here have been conducted with younger populations.

A more promising lead has been found when the focus is shifted from weight gain to relapse prevention in anorexia nervosa. Kaye, Guirtsman, George, and Ebert (1991) found that 29 of 31 women recovering from anorexia nervosa maintained their weight an average of 11 months when utilizing doses of fluoxetine between 20 and 80 mg daily. In a controlled study, Kaye et al. (1996) found that 63% of weight restored women with anorexia nervosa maintained a healthy body weight and a reduction in core eating disorder pathology for one year while taking fluoxetine. Only 16% of the participants taking a placebo maintained a healthy body weight. Thus, fluoxetine may be a promising agent in helping to prevent relapse in anorexia nervosa. More controlled studies are needed however, before firm conclusions can be drawn about the efficacy of fluoxetine with this population.

In sum, there is still little evidence for change in the early recommendations that pharmacotherapy has a very limited value with emaciated anorexia nervosa patients and should never be the sole treatment approach (Garfinkel & Garner, 1982). Occasionally patients may benefit from medication to deal with overwhelming anxiety, severe depression or intolerable gastric discomfort after meals, but this only applies to a small minority of patients (Andersen, 1985; Garfinkel & Garner, 1982; Garfinkel & Walsh, 1997).

5.23.9.2 Treatment Research on Bulimia Nervosa

In contrast to the lack of efficacy of medications in treating anorexia nervosa, there is some reason for optimism for the role of certain antidepressants in the treatment of bulimia nervosa (Garfinkel & Walsh, 1997; Raymond et al., 1994). The rationale for using antidepressants with bulimia nervosa was derived from several sources. First, as indicated earlier, Hudson and Pope (1988) argued that the efficacy of antidepressants in the treatment of bulimia nervosa suggests that it could be a variation of an affective illness. However, the failure to find any relationship between pretreatment levels of depression and response to medication does not support this view (Horne et al., 1988; Hughes, Wells, Cunningham, & Ilstruys, 1986; Walsh et al., 1988). Additionally, the dosage needed to be effective with bulimia nervosa is often substantially higher than the dosage used to treat mood disturbances (i.e., fluoxetine; FBNC, 1992).

Another more recent argument for pursuing antidepressants in treating bulimia nervosa is based on the hypothesis that dietary restraint may lead to lower levels of serotonin in the central nervous system. Binge eating may be driven by a need to increase brain serotonin (Walsh & Devlin, 1995). However, the merits of boosting serotonin with medication rather than with food is open to debate.

A review of the empirical literature on antidepressant treatment suggests that almost all antidepressants studied have been found to be superior to placebo in reducing binge frequency (Mitchell et al., 1993). Antidepressant therapy also leads to improvement in mood disturbances and reduces preoccupations with weight and shape. Only one controlled study, utilizing the antidepressant amitriptyline, failed to show a significant effect of the drug over placebo (Mitchell & Groat, 1984). Seven studies have been published that examine the efficacy of tricyclics such as imipramine and desipramine (Mitchell et al., 1993). Across these studies mean reduction in binge frequency was 69%. However, mean rate of abstinence from bulimic symptoms was only 24%. Moreover, the relatively high percentage of participants who relapsed while taking antidepressants is a concern (Pyle et al., 1990; Walsh, Hadigan, Devlin, Gladis, & Roose, 1991). For example, Walsh et al. (1991) found that 29% of participants who responded to desipramine during an eight-week trial relapsed during a four-month maintenance phase. Likewise, Pyle et al. (1990) assigned nine responders to a three-month trial of imipramine to either continue the drug treatment ($n = 3$) or to a placebo ($n = 6$). Five of the six participants in the placebo group relapsed and two of the three participants in the imipramine group relapsed. Of the tricyclics, desipramine is recommended as the first medication of choice given its suggested efficacy and minimal anticholineric effects.

There is growing evidence that third-generation antidepressants, selective serotonin reuptake inhibitors (SSRIs), can be effective in treating bulimia nervosa. A multisite collaborative study (FBNC, 1992) indicated that fluoxetine hydrocholride (Prozac) in daily dosages of 60 mg was superior to 20 mg dosages and probably should be used at least as an adjunct to psychotherapy in cases failing a course of adequate psychological treatment. Open studies have been conducted examining the efficacy of other SSRIs. Geretsegger, Greimel, Roed, and Hesselink (1995) conducted an open label study with ipsapirone and reported that 93% of the participants showed greater than a 50% reduction in weekly binges after four weeks of treatment. Mizes, Sloan, Ruderich, Freiheit, and Silverman (1996) found that sertraline yielded a significant reduction in bingeing and purging episodes among seven participants.

Studies that have compared the efficacy of antidepressant treatments to psychotherapeutic interventions have shown that psychotherapy is more effective than drug treatment alone and that participants who received combined drug and psychotherapy interventions may have slightly better outcomes (Agras et al., 1992; Leitenberg et al, 1994; Mitchell et al., 1990; Walsh et al., 1997). Agras et al. (1992) showed that a combined treatment of CBT and 24 weeks of treatment with desipramine yielded the highest rate of abstinence from bingeing at 32 weeks compared to CBT alone, desipramine alone, and CBT and desipramine treatment for 16 weeks. Likewise, Leitenberg et al. (1994) compared CBT alone, 20 weeks of desipramine treatment, and 20 weeks of combined treatments. There was a high (57%) dropout rate in the desipramine group, thus, the interpretation of the data is limited. However, the percentage of participants abstinent from vomiting at the six-month follow-up was substantially higher for those in the CBT group (57%) than for those in the desipramine group (0%) or in the combined treatment group (29%). Walsh et al. (1997) used a two-stage medication intervention for bulimia nervosa in which fluoxetine was employed if desipramine was either ineffective or poorly tolerated. This approach was compared to a combination of medication and two psychological treatments (cognitive-behavioral and psychodynamic-supportive). The results indicated that the two-stage medication intervention added modestly to the effects of the cognitive-behavioral treatment in reducing binge eating and depression. In addition, cognitive behavioral therapy plus medication was superior to medication alone; however, supportive psychotherapy plus medication was not more effective than medication alone.

MAOIs have been shown to be effective in reducing binge eating. Walsh et al. (1988) reported a 59% reduction in binge frequency in a controlled trial with an MAOI and Kennedy et al. (1988) reported a 22% reduction in binge frequency. However, side effects of MAOIs such as orthostatic hypotension, insomnia, and weight gain may be problematic. Additionally, dietary restrictions involving foods with a high concentration of tyramine or dopamine (aged cheese, beer) are necessary when taking MAOIs because high concentrations of tyramine or dopamine may cause a life threatening rapid increase in blood pressure. Placing people with bulimia nervosa on dietary restrictions is often counter to the therapeutic goals. Thus, it is generally recommended that MAOIs not be utilized as the first trial drug for this population (Kennedy et al., 1988; Walsh et al., 1988).

There have been a number of studies examining the utility of drugs other than antidepressants in the treatment of bulimia nervosa. Several studies of naltrexone, an opioid antagonist, have yielded mixed results. Mitchell, Davis, Goff, and Pyle (1986) reported a 23% decrease in binge size with naloxone. Alger, Schwalberg, Bigaouette, Michaleck, and Howard (1991) conducted a controlled trial with naltrexone and found that it significantly reduced the duration of bingeing episodes but was no better than placebo in reducing frequency of binges. In another controlled trial, naltrexone was no better than placebo in reducing symptoms of bulimia nervosa (Mitchell et al., 1989). Marrazzi, Bacon, Kinzie, and Luby (1995) reported on the results of a double-blind randomized crossover-design trial of naltrexone in treating participants with bulimia nervosa and anorexia nervosa. Out of the 19 participants, 18 showed a greater reduction in bingeing and purging while taking the naltrexone when compared to placebo. Given the mixed results of the above studies, the efficacy of treating bulimia nervosa and perhaps, anorexia nervosa with naltrexone warrants further investigation.

Mood stabilizing drugs such as lithium carbonate have been used in several trials, however, Garfinkel and Walsh (1997) concluded that "there is little evidence that most patients will benefit from their use" (p. 374). Fenfluramine has been studied with bulimia nervosa because of its effects in promoting satiety. An early study by Blouin et al. (1988) produced encouraging findings but a subsequent controlled trial showed no substantial clinical benefit of fenfluramine over placebo (Fahy, Eisler, & Russell, 1993).

5.23.10 FUTURE DIRECTIONS

A major premise in this chapter has been that eating disorders are multidetermined and heterogeneous syndromes resulting from the interplay of biological, psychological, familial, and sociocultural predisposing factors. In this sense, they are probably best understood as final common pathways that appear to have different psychological points of entry. Within this overall context, this chapter has selectively reviewed diagnostic issues, major etiological formulations, and associated psychopathology applied to eating disorders.

From research to date it may be concluded that there has been as much resistance to dispelling the "uniformity myth" applied to eating disorders as has been the case with disorders with a longer or more visible history in the psychological nomenclature. However, research since the mid-1980s has yielded notable advancements in refining the criteria for eating-disordered subgroups and in providing more rigorous research examining psychopathology or psychological topologies associated with diagnostic subtypes. An important direction for future research will be the clearer specification of the relative contribution of particular biological, psychological, and interpersonal predisposing features considered relevant to eating disorders. It is to be hoped that the extraordinary clinical and research interest in personality disorders, depression, sexual abuse, and addictive behaviors as related to eating disorders will continue to lead to improved understanding. Further research is clearly needed to determine the precise nature and the significance of the observed associations. If past research to date is any indication of future findings, the within-diagnostic-group variability will continue to be as noteworthy as the between-group differences.

It is now evident that certain patients respond relatively quickly to brief interventions, in contrast to others who require more intensive and protracted treatments. Perhaps the most significant goal for future research will be the identification of traits, personality features, or background factors that predict differential response to treatment. Even better would be a taxonomy yielding an accurate match between patient characteristics and treatment type.

Advances in research on psychopathology and treatment efficacy warrant genuine optimism with regard to bulimia nervosa. Less is known about personality and response to treatment for anorexia nervosa because of the relative absence of controlled treatment research with this eating disorder. It is hoped that controlled treatment research will assume a high priority, and that the results will generate the same level of progress as is now evident with bulimia nervosa.

5.23.11 REFERENCES

Agras, W. S. (1993). Short-term psychological treatments for binge eating. In C. G. Fairburn & G. T. Wilson (Eds.), *Binge eating: Nature, assessment, and treatment* (pp. 50–76). New York: Guilford Press.

Agras, W. S., Rossiter E. M., Arnow, B., Schneider, J. A., Telch, C. F., Raeburn, S. D., Bruce, B., Perl, M., & Koran, L. M. (1992) Pharmacologic and cognitive-behavioral treatment for bulimia nervosa: A controlled comparison. *American Journal of Psychiatry, 149,* 82–87.

Agras, W. S., Rossiter, E. M., Arnow, B., Telch, C., Raeburn, S., Bruce, B., & Koran, L. (1994). One year follow up of psychosocial and pharmacologic treatments for bulimia nervosa. *Journal of Clinical Psychiatry, 55,* 179–83.

Alger, S. A., Schwalberg, M. D., Bigaouette, J. M., Michaleck, A. V., & Howard, L. J. (1991). Effect of a tricyclic antidepressant and opiate antagonist on binge-eating in normal weight bulimic and obese, binge-eating subjects. *American Journal of Clinical Nutrition, 53,* 865–871.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.) Washington, DC: Author.

Andersen, A. E. (1985). *Practical comprehensive treatment of anorexia nervosa and bulimia.* Baltimore: Johns Hopkins University Press.

Andersen, A. E. (1990). *Males with eating disorders*. New York: Brunner/Mazel.

Andersen, A. E., Bowers, W., & Evans, K. (1997). Inpatient treatment of anorexia nervosa. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 327–353). New York: Guilford Press.

Armstrong, J. G., & Roth, D. M. (1989). Attachment and separation difficulties in eating disorders: A preliminary investigation. *International Journal of Eating Disorders, 8,* 141–155.

Baker Enright, A., Butterfield, P., & Berkowitz (1985). Self-help and support groups in the management of eating disorders. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of psychotherapy for anorexia nervosa and bulimia* (pp. 491–512), New York: Guilford Press.

Beaumont, P. J. V., Al-Alami, M. S., & Touyz, S. W. (1987). The evolution of the concept of anorexia nervosa. In P. J. V. Beaumont, G. D. Burrows, & R. C. Casper (Eds.), *Handbook of eating disorders: Part 1, Anorexia and bulimia nervosa* (pp. 105–116). New York: Elsevier.

Beaumont, P. J. V., George, G. G. W., & Smart, D. E. (1976). "Dieters" and "vomiters and purgers" in anorexia nervosa. *Psychological Medicine, 6,* 617–622.

Beaumont, P. J. V., O'Connor, M., Touyz, S. W., & Williams, H. (1987). Nutritional counseling in the treatment of anorexia and bulimia nervosa. In P. J. V. Beaumont, G. D. Burrows, & R. C. Casper (Eds.), *Handbook of eating disorders: Part 1: Anorexia and bulimia nervosa.* New York: Elsevier.

Bell, R. (1985). *Holy anorexia.* Chicago: University of Chicago Press.

Bemis, K. M. (1985). "Abstinence" and "nonabstinence" models for the treatment of bulimia. *International Journal of Eating Disorders, 4,* 389–406.

Biederman, J., Herzog, D., Rivinus, T., Harper, G., Ferber, R., Rosenbaum, J., Harmatz, J., Tondorf, R., Orsulack, P., & Schildkraut, J. (1985). Amitriptyline in the treatment of anorexia nervosa. *Journal of Clinical Psychopharmacology, 5,* 10–16.

Blouin, A. G., Blouin, J., Perez, E., Bushnik, T., Zuro, C., & Mulder, E. (1988). Treatment of bulimia with fenfluramine and desipramine. *Journal of Clinical Psychopharmacology, 8,* 261–269.

Brotman, A. W., Herzog, D. B., & Woods, S. W. (1984). Antidepressant treatment of bulimia nervosa: The relationship between bingeing and depressive symptomatology. *Journal of Clinical Psychiatry, 45,* 7–9.

Bruch, H. (1973). *Eating disorders: Obesity, anorexia nervosa and the person within.* New York: Basic Books.

Bruch, H. (1985). Four decades of eating disorders. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of psychotherapy for anorexia nervosa and bulimia* (pp. 7–18). New York: Guilford Press.

Bulik, C., Sullivan, P., Joyce, P., & Carter, F. (1995). Temperament character, and personality disorder in bulimia nervosa. *The Journal of Nervous and Mental Disease, 183,* 593–598.

Cantwell, D. P., Sturzenburger, S., Burroughs, J., Salkin, B., & Green, J. K. (1977) Anorexia nervosa: An affective disorder? *Archives of General Psychiatry, 34,* 1087–1093.

Casper, R. C. (1982). Treatment principles in anorexia nervosa. *Adolescent Psychiatry, 10,* 86–100.

Casper, R. C. (1983). On the emergence of bulimia nervosa as a syndrome: A historical view. *International Journal of Eating Disorders, 2,* 3–16.

Casper, R. C. (1990). Personality features of women with good outcome from restricting anorexia nervosa. *Psychosomatic Medicine, 52,* 156–170.

Casper, R. C., Eckert, E. D., Halmi, K. A., Goldberg, S. C., & Davis, J. M. (1980). Bulimia: Its incidence and clinical importance in patients with anorexia nervosa. *Archives of General Psychiatry, 37,* 1030–1034.

Channon, S., DeSilva, P., Hemsley, D., & Perkins, R. (1989). A controlled trial of cognitive-behavioral and behavioral treatment of anorexia nervosa. *Behavior Research and Therapy, 27,* 529–535.

Chiodo, J., & Latimer, P. R. (1983). Vomiting as a learned weight-control technique in bulimia. *Journal of Behavior Therapy and Experimental Psychiatry, 14,* 131–135.

Cloninger, C. R. (1989). A systematic method for clinical description and classification of personality variants. *Archives of General Psychiatry, 44,* 573–588.

Connors, M. E., Johnson, C. L., & Stuckley, M. K. (1984). Treatment of bulimia with brief psychoeducational group therapy. *American Journal of Psychiatry, 141,* 1512–1516.

Cooper, J. L., Morrison, T. L., Bigman, O. L., Abramowitz, S. I., Levin, S., & Krener, P. (1988). Mood changes and affective disorder in the bulimic binge-purge cycle. *International Journal of Eating Disorders, 7,* 469–474.

Cooper, P. J., & Fairburn, C. G. (1984). Cognitive behavioral treatment for anorexia nervosa: Some preliminary findings. *Journal of Psychosomatic Research, 28,* 493–499.

Cooper, Z., & Fairburn, C. G. (1987). The eating disorder examination: A semistructured interview for the assessment of the specific psychopathology of eating disorders. *International Journal of Eating Disorders, 6,* 1–8.

Cooper, Z., Coker, S., & Fleming, C. (1995). An evaluation of the efficacy of cognitive behavioural self-help for bulimia nervosa. *Journal of Psychosomatic Research, 40,* 281–287.

Cooper, P. & Steere, J. (1995). A comparison of two psychological treatments for bulimia nervosa: Implications for models of maintenance. *Behavior Research & Therapy, 33,* 875–885.

Crisp, A. H. (1965). Clinical and therapeutic aspects of anorexia nervosa: Study of 30 cases. *Journal of Psychosomatic Research, 9,* 67–78.

Crisp, A. H. (1970). Anorexia nervosa. "Feeding disorder," nervous malnutrition or weight phobia? *World Review of Nutrition, 12,* 452–504.

Crisp, A. H. (1980). *Anorexia nervosa.* New York: Grune & Stratton.

Crisp, A. H. (1997). Anorexia nervosa as flight from growth: Assessment and treatment based on the model. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 248–277). New York: Guilford Press.

Crisp, A. H., Norton, K., Gowers, S., Halek, C., Bowyer, C., Yeldham, D., Levett, G., & Bhat, A. (1991). A controlled study of the effect of therapies aimed at adolescent and family psychopathology in anorexia nervosa. *British Journal of Psychiatry, 159,* 325–333.

Crowther J., & Sherwood N. (1997). Assessment. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 34–49). New York: Guilford Press.

Dare, C., & Eisler, I. (1997). Family therapy for anorexia nervosa. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 307–326). New York: Guilford Press.

Davies, E., & Furnham, A. (1986). The dieting and body shape concerns of adolescent females. *Journal of Child Psychology and Psychiatry, 3,* 417–428.

Eckert, E. D., Halmi, K. A., Marchi, P., & Cohen, J. (1987). Comparison of bulimic and non-bulimic anorexia nervosa patients during treatment. *Psychological Medicine, 17,* 891–898.

Eckert, E. D., Halmi, K. A., Marchi, P., Grove, W., & Crosby, R. (1995). Ten-year follow-up of anorexia nervosa: Clinical course and outcome. *Psychological Medicine, 25,* 143–156.

Edlund, B., Halvarsson, K., & Sjödén, P. (1995). Eating behaviours, and attitudes to eating, dieting, and body image in a 7-year-old Swedish girl. *European Eating Disorders Review, 3,* 111/1–111/14.

Fahy, T., & Eisler, I. (1993). Impulsivity and eating disorders. *British Journal of Psychiatry, 162,* 193–97.

Fahy, T., Eisler, I., & Russell, G. (1993). A placebo-controlled trial of d-Fenfluramine in bulimia nervosa. *British Journal of Psychiatry, 162,* 597–603.

Fairburn, C. G. (1985). Cognitive-behavioral treatment for bulimia. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of psychotherapy for anorexia nervosa and bulimia* (pp. 160–192). New York: Guilford Press.

Fairburn, C. G., (1997). Interpersonal psychotherapy for bulimia nervosa. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 278–294). New York: Guilford Press.

Fairburn, C. G., Agras, W. S., & Wilson, G. T. (1992). The research on the treatment of bulimia nervosa: Practical and theoretical implications. *Biology of Feast and Famine,* 317–340.

Fairburn, C. G., & Beglin, S. J. (1990). Studies of the epidemiology of bulimia nervosa. *American Journal of Psychiatry, 147,* 401–408.

Fairburn, C. G., & Carter, J. (1997). Self-help and guided self-help for binge eating problems. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 494–500). New York: Guilford Press.

Fairburn, C. G., Jones, R., Peveler, R. C., Carr, S. J., Solomon, R. A., O'Connor, M. E., Burton, J., & Hope, R. A. (1991). Three psychological treatments for bulimia nervosa: A comparative trial. *Archives of General Psychiatry, 48,* 463–469.

Fairburn, C. G., Jones, R., Peveler, R. C., Hope, R. A., & O'Connor, M. E. (1993). Psychotherapy and bulimia nervosa: the longer-term effects of interpersonal psychotherapy, behavior therapy and cognitive behavior therapy. *Archives of General Psychiatry, 50,* 419–428.

Fairburn, C. G., Kirk, J., O'Connor, M. E., & Cooper, P. J. (1986). A comparison of two psychological treatments for bulimia nervosa. *Behaviour Research and Therapy, 24,* 629–643.

Fairburn, C. G., Marcus, M. D., & Wilson, G. T. (1993). Cognitive-behavioral therapy for binge eating and bulimia nervosa. In C. G. Fairburn & G. T. Wilson (Eds.), *Binge eating nature: Assessment and treatment* (361–404). New York: Guilford Press.

Fairburn, C. G., Norman, P. A., Welch, S. L., O'Connor, M. E., Doll, H. A., & Peveler, R. C. (1995). A prospective study of outcome in bulimia nervosa and the long-term effects of three psychological treatments, *Archives of General Psychiatry, 52,* 304–312.

Fairburn, C. G., & Peveler, R. C. (1990). Bulimia nervosa and a stepped care approach to management, *Gut, 31,* 1220–1222.

Fallon, P., Katzman, M. A., & Wooley, S. C. (1994). *Feminist perspectives on eating disorders.* New York: Guilford Press.

Fallon, P., & Wonderlich, S. (1997). Sexual abuse and other forms of trauma. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 394–414). New York: Guilford Press.

Fichter, M., & Quadflieg, N. (1995). Comparative studies on the course of eating disorders in adolescence and adults. Is age at onset a predictor of outcome? In H. C. Steinhausen (Ed.), *Eating disorders in adolescence. Anorexia and bulimia nervosa.* Berlin: de Gruyter.

Fluoxetine Bulimia Nervosa Collaborative Study Group (FBNC) (1992). Fluoxetine in the treatment of bulimia nervosa. *Archives of General Psychiatry, 49,* 139–147.

Fosson, A., Knibbs, J., Bryant-Waugh, R., & Lask, B. (1987). Early onset anorexia nervosa. *Archives of Disease in Childhood, 62,* 114–118.

Frommer, M. S., Ames, J. R., Gibson, J. W., & Davis, W. N. (1987). Patterns of symptom change in the short-term treatment of bulimia. *International Journal of Eating Disorders, 6,* 469–476.

Garfinkel, P. E., & Garner, D. M. (1982). *Anorexia nervosa: A multidimensional perspective.* New York: Brunner/Mazel.

Garfinkel, P. E., & Garner, D. M. (1987). *Psychotropic drug therapies for eating disorders.* Monograph Series. New York: Brunner/Mazel.

Garfinkel, P. E., Moldofsky, H., & Garner, D. M. (1980). The heterogeneity of anorexia nervosa. *Archives of General Psychiatry, 37,* 1036–1040.

Garfinkel, P. E., Kaplan, A. S., Garner, D. M., & Darby, P. L. (1983). The differentiation of vomiting/weight loss as a conversion disorder from anorexia nervosa. *American Journal of Psychiatry, 140,* 1019–1022.

Garfinkel, P. E., & Walsh B. (1997). Drug therapies. In D. M., Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 229–247). New York: Guilford Press.

Garner, D. M. (1985). Iatrogenesis in anorexia nervosa and bulimia nervosa. *International Journal of Eating Disorders, 4,* 701–726.

Garner, D. M. (1986) Cognitive therapy for anorexia nervosa. In K. D. Brownell & J. P. Foreyt (Eds.), *Handbook of eating disorders* (pp. 301–327). New York: Basic Books.

Garner, D. M. (1988) Anorexia nervosa. In M. Hersen & C. G. Last (Eds.), *Child behavior therapy casebook* (pp. 263–276). New York: Plenum.

Garner, D. M. (1991). *Eating disorder inventory–2: Professional manual.* Odessa, FL: Psychological Assessment Resources.

Garner, D. M. (1993). Pathogenesis of anorexia nervosa. *The Lancet, 341,* 1631–1635.

Garner, D. M. (1997). Psychoeducational principles in treatment. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 145–177). New York: Guilford Press.

Garner, D. M., & Bemis, K. M. (1982). A cognitive-behavioral approach to anorexia nervosa. *Cognitive Therapy and Research, 6,* 123–150.

Garner, D. M., & Bemis, K. M. (1985). Cognitive therapy for anorexia nervosa. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of psychotherapy for anorexia nervosa and bulimia* (pp. 107–146). New York: Guilford Press.

Garner, D. M., Fairburn, C. G., & Davis, R. (1987). Cognitive-behavioral treatment of bulimia nervosa: A critical appraisal. *Behavior Modification, 11,* 398–431.

Garner, D. M., & Garfinkel, P. E. (1980). Socio-cultural factors in the development of anorexia nervosa. *Psychological Medicine, 10,* 647–656.

Garner, D. M., & Garfinkel, P. E. (1997). *Handbook of treatment for eating disorders.* New York: Guilford Press.

Garner, D. M., Garfinkel, P. E., & Bemis, K. M. (1982). A multidimensional psychotherapy for anorexia nervosa. *International Journal of Eating Disorders, 1,* 3–46.

Garner, D. M., Garfinkel, P. E., & Irvine, M. J. (1986). Integration and sequencing of treatment approaches for eating disorders. *Psychotherapy and Psychosomatics, 46,* 67–75.

Garner, D. M., Garfinkel, P. E., & O'Shaughnessy, M. (1985). The validity of the distinction between bulimia with and without anorexia nervosa. *American Journal of Psychiatry, 142,* 581–587.

Garner, D. M., Garfinkel, P. E., Schwartz, D. M., & Thompson, M. M. (1980). Cultural expectations of thinness in women. *Psychological Reports, 47,* 483–491.

Garner, D. M., Garner, M. V., & Rosen, L. W. (1993). Anorexia nervosa "restricters" who purge: Implications for subtyping Anorexia nervosa. *International Journal of Eating Disorders, 13,* 171–185.

Garner, D. M., & Needleman, L. (1997). Sequencing and integration of treatments. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment of eating disorders* (pp. 50–66). New York: Guilford Press.

Garner, D. M., Rockert, W., Olmsted, M. P., Johnson, C. L., & Coscina, D. V. (1985). Psychoeducational principles in the treatment of bulimia and anorexia nervosa. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of psychotherapy for anorexia nervosa and bulimia* (pp. 513–572). New York: Guilford Press.

Garner, D. M., & Rosen, L. W. (1990). Anorexia nervosa and bulimia nervosa. In A. S. Bellack, M. Hersen, & A. E. Kazdin (Eds.), *International handbook of behavior modification and therapy* (pp. 805–817). New York: Plenum.

Garner, D. M., & Rosen, L. W. (1991). Eating disorders in athletes: Research and recommendations. *Journal of Applied Sports Research, 5,* 100–107.

Garner, D. M., & Sackeyfio, A. H. (1993). Eating disorders. In A. S. Bellack & M. Hersen (Eds.), *Handbook of behavior therapy in the psychiatric setting* (pp. 477–497). New York: Plenum.

Garner, D. M., Vitousek, K. & Pike, K. (1997). Cognitive-behavioral therapy for anorexia nervosa. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 94–144). New York: Guilford Press.

Garner, D. M., & Wooley, S. C. (1991). Confronting the failure of behavioral and dietary treatments for obesity. *Clinical Psychology Review, 11,* 729–780.

Gartner, A. F., Marcus, R. N., Halmi, K., & Loranger, A. W. (1989). DSM-III-R personality disorders in patients with eating disorders. *American Journal of Psychiatry, 14,* 1585–1591.

Geretsegger, C., Greimel, K., Roed, I., & Hesselink, J. (1995). Ipsapirone in the treatment of bulimia nervosa: An open pilot study. *International Journal of Eating Disorders, 17,* 359–363.

Gillberg, G., Rastam, M., & Gillberg, C. (1995). Anorexia nervosa 6 years after onset: Part I. Personality disorders. *Comprehensive Psychiatry, 36,* 61–69.

Goetz, P. L., Succop, R. A., Reinhart, J. B., & Miller, A. (1977). Anorexia nervosa in children: A follow-up study. *American Journal of Orthopsychiatry, 47,* 597–603.

Goldberg, S. C., Halmi, K. A., Eckert, E. D., Casper, R. C., & Davis, J. M. (1979). Cyproheptadine in anorexia nervosa. *British Journal of Psychiatry, 134,* 67–70.

Goodsitt, A. (1985). Self psychology and the treatment of anorexia nervosa. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of psychotherapy for anorexia nervosa and bulimia* (pp. 55–84). New York: Guilford Press.

Goodsitt, A. (1997). Eating disorders a self psychological perspective. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 205–228). New York: Guilford.

Gowers, S., & McMahon, J. B. (1989). Social class and prognosis in anorexia nervosa. *International Journal of Eating Disorders, 8,* 105–110.

Gowers, S., Norton, K., Halek, C., & Crisp, A. (1994). Outcome of outpatient psychotherapy in a random allocation treatment study of anorexia nervosa. *International Journal of Eating Disorders, 15,* 165–177.

Gross, H., Ebert, M., Faden, V., Goldberg, S., Nee, L., & Kaye, W. (1981). A double-blind controlled trial of lithium carbonate in primary anorexia nervosa. *Journal of Clinical Psychopharmacology, 1,* 376–381.

Haimes, A. L., & Katz, J. L. (1988). Sexual and social maturity versus social conformity in restricting anorectic bulimic, and borderline women. *International Journal of Eating Disorders, 7,* 331–341.

Halmi, K. A., Eckert, E., LaDu, T. J., & Cohen, J. (1986). Anorexia nervosa: Treatment efficacy of cyproheptadine and amitriptyline. *Archives of General Psychiatry, 43,* 177–181.

Halmi, K. A., Eckert, E., Marchi, P., Sampugnaro, V., Apple, R., & Cohen, J. (1991). Comorbidity of psychiatric diagnoses in anorexia nervosa. *Archives of General Psychiatry, 48,* 712–718.

Hatsukami, D., Mitchell, J. E., Eckert, E. D., & Pyle, R. (1986). Characteristics of patients with bulimia only, bulimia with affective disorder, and bulimia with substance abuse problems. *Addictive Behaviors, 11,* 399–406.

Hatsukami, D., Owen, P., Pyle, R., & Mitchell, J. (1982). Similarities and differences on the MMPI between women with bulimia and women with alcohol and drug abuse problems. *Addictive Behaviors, 7,* 435–439.

Hawley, R. M. (1985) The outcome of anorexia nervosa in younger subjects. *British Journal of Psychiatry, 146,* 657–660.

Heatherton, T. F., & Baumeister, R. F. (1991). Binge-eating as escape from self-awareness. *Psychological Bulletin, 110,* 86–108.

Herzog, D. B., Keller, M. B., Lavori, P. W., & Ott, I. L. (1987). Social impairment in bulimia. *International Journal of Eating Disorders, 6,* 741–747.

Herzog, D. B., Keller, M. B., & Lavori, P. W. (1988). Outcome in anorexia nervosa and bulimia nervosa: A review of the literature. *Journal of Nervous and Mental Disease, 176,* 131–143.

Herzog, D. B., Keller, M. B., Sacks, N. R., Yeh, C. J., & Lavori, P. W. (1992). Psychiatric morbidity in treatment-seeking anorexics and bulimics. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 810–818.

Herzog, W., Schellberg, D., & Deter, H. (1997). First recovery in anorexia nervosa patients in the long-term course: A discrete-time survival analysis. *Journal of Consulting and Clinical Psychology, 65,* 169–177.

Hinz, L. D., & Williamson, D. A. (1987). Bulimia and depression: A review of the affective variant hypothesis. *Psychological Bulletin, 102,* 150–158.

Hoek, H. W. (1993). Review of the epidemiological studies of eating disorders. *International Review of Psychiatry, 5,* 61–74.

Hoek, H. W., Bartelds, A., Bosveld, J., Graff, Y., Limpens, V., Maiwald, M., & Spaij, C. (1995). Impact of urbanization on detection rates of eating disorders. *American Journal of Psychiatry, 152,* 1272–1278.

Holland A. J., Sicotte, N., & Treasure, J. (1988). Anorexia nervosa: Evidence for a genetic basis. *Journal of Psychosomatic Research, 32,* 561–571.

Horne, R. L., Ferguson, J. M., Pope, H. G., Hudson, J. I., Lineberry, C. G., Ascher, J., & Cato, A. (1988). Treatment of bulimia with bupropion: A multicenter controlled trial. *Journal of Clinical Psychiatry, 49,* 262–266.

Hsu, L. K. G. (1988). The outcome of anorexia nervosa: A reappraisal. *Psychological Medicine, 18,* 807–812.

Hsu, L. K. G. (1996). Outcome of early onset anorexia nervosa: What do we know? *Journal of Youth and Adolescence, 25,* 563–568.

Hsu, L. K. G., Kaye, W., & Weltzin (1993). Are the eating disorders related to obsessive compulsive disorder? *International Journal of Eating Disorders, 14,* 305–318.

Hudson, J. I., & Pope, H. G. (1988). Depression and eating disorders. In O. G. Cameron (Ed.), *Presentations of depression*. New York: Wiley.

Hughes, P. L., Wells, L. A., Cunningham, C. J., & Ilstrup, D. M. (1986). Treating bulimia with desipramine: A double-blind placebo-controlled study. *Archives of General Psychiatry, 43,* 182–186.

Humphrey, L. L. (1989). Is there a causal link between disturbed family processes and eating disorders? In W. G. Johnson (Ed.), *Bulimia nervosa: Perspectives on clinical research and therapy*. New York: JAI Press.

Huon, G. F. (1985). An initial validation of a self-help program for bulimia. *International Journal for Eating Disorders, 4,* 573–588.

Huon, G. F., & Brown, L. B. (1985). Evaluating a group treatment for bulimia. *Journal of Psychiatric Research, 19,* 479–483.

Huon, G. F., & Brown, L., & Morris, S. (1988). Lay beliefs about disordered eating. *International Journal of Eating Disorders, 7,* 239–252.

Jager, B., Liedtke, R., Kunsebeck, H. W., & Lempa, W. (1996). Psychotherapy and bulimia nervosa: Evaluation and long-term follow-up of two conflict-oriented treatment conditions. *Acta Psychiatrica Scandinavica, 93, 268–278.*

Jaffe, A. C., & Singer, L. T. (1989). Atypical eating disorders in young children. *International Journal of Eating Disorders, 8,* 575–582.

Johnson, C. L., & Connors, M. E. (1987). *The etiology and treatment of bulimia nervosa: A biopsychosocial perspective.* New York: Basic Books.

Johnson, C., Tobin, D. L., & Dennis, A. (1990). Differences in treatment outcome between borderline and nonborderline bulimics at one-year follow-up. *International Journal of Eating Disorders, 9,* 617–627.

Kaplan, A., & Olmsted, M. (1997). Partial hospitalization. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 354–360). New York: Guilford Press.

Kasvikis, Y. G., Isakiris, F, Marks, I. M., Basoglu, M., & Noshirvani, H. F. (1986). Past history of anorexia nervosa in women with obsessive-compulsive disorder. *International Journal of Eating Disorders, 5,* 1069–1075.

Katzman, M. A. (1994). When reproductive and productive worlds meet: collusion or growth? In P. Fallon, M. A. Katzman, & S. C. Wooley, (Eds.), *Feminist perspectives on eating disorders* (pp. 132–151). New York: Guilford Press.

Kaye, W. H., Gwirtsman, H. E., George, D. T., & Ebert, M. H. (1991). Altered serotonin activity in anorexia nervosa after long-term weight restoration. *Archives of General Psychiatry, 48,* 556–562.

Kaye, W. H., McConaha, C., Nagata, T., Plotnicov, K. H., Sokol, M. S., Weltzin, T. E., Hsu, L. K. G., & La Via, M. C. (1996). Fluoxetine prevents relapse in a majority of patients with anorexia nervosa: A double-blind placebo-controlled study. Presented at the Eating Disorders Research Society Meeting, Pittsburgh, PA.

Kearney-Cooke, A., & Striegel-Moore, R. (1997). The etiology and treatment of body image disturbance. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 295–306). New York: Guilford Press.

Kellett, J., Trimble, M., & Thorley, A. (1976). Anorexia nervosa after the menopause. *British Journal of Psychiatry, 128,* 555–558.

Kennedy, S. H., Piran, N., Warsh, J. J., Pendergast, P., Mainprize, E., Whynot, C., & Garfinkel, P. E. (1988). A trial of isocaroxide in the treatment of bulimia nervosa. *Journal of Clinical Psychopharmacology, 8,* 391–396.

Kirkley, B. G., Schneider, J. A., Agras, W. S., & Bachman, J. A. (1985). Comparison of two group treatments for bulimia. *Journal of Consulting and Clinical Psychology, 53,* 43–48.

Klerman, G. L., Weissman, M. M., Rounsaville, B. J., & Chevron, E. (1984). *Interpersonal psychotherapy for depression.* New York: Basic Books.

Kohut, H. (1971). *The restoration of the self.* New York: International Universities Press.

Kreipe, R. E., Churchill, B. H., & Strauss, J. (1989). Long-term outcome of adolescents with anorexia nervosa. *American Journal of Diseases of Children, 143,* 1322–1327.

Lacey, J. H., & Crisp, A. H. (1980). Hunger, food intake and weight: The impact of clomipramine on a refeeding anorexia nervosa population. *Postgraduate Medical Journal, 56,* 79–85.

Laessle, R. G., Beumont, P. J. V., Butow, P., Lenneris, W., O'Connor, M., Pirke, K. M., Touyz, S. W., & Waadi, S. (1991). A comparison of nutritional management and stress management in the treatment of bulimia nervosa. *British Journal of Psychiatry, 159,* 250–261.

Laessle, R. G., Kittl, S., Fichter, M. M., Wittchen, H. U., & Pirke, K. M. (1987). Major affective disorder in anorexia nervosa and bulimia—a descriptive diagnostic study. *British Journal of Psychiatry, 151,* 785–789.

Laessle, R. G., Waadt, S., & Pirke, K. M. (1987). A structured behaviorally oriented group treatment for bulimia nervosa. *Psychotherapy and Psychosomatics, 48,* 141–145.

Laessle, R. G., Wittchen, H. U., Fichter, M. M., & Pirke, K. M. (1989). The significance of subgroups of bulimia and anorexia nervosa: Lifetime frequency of psychiatric disorders. *International Journal of Eating Disorders, 8,* 569–574.

Lask B., & Bryant-Waugh, R. (1997). Pre-pubertal eating disorders. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 476–483). New York: Guilford Press.

Leitenberg, H., Rosen, J. C., Wolf, J., Vara, L. S., Detzer, M. J., & Srebnik, D. (1994). Comparison of cognitive-behavior therapy and desipramine in the treatment of bulimia nervosa. *Behavior Research and Therapy, 32,* 37–45.

Lerner, H. D. (1993). Self-representation in eating disorders: A psychodynamic perspective. In Z. Segal & S. Blatt (Eds.), *The self-in emotional disorders: Cognitive and psychodynamic perspectives* (pp. 267–287). New York: Guilford Press.

Levin, A. P., & Hyler, S. E. (1986). DSM-III personality diagnosis in bulimia. *Comprehensive Psychiatry, 27,* 47–53.

Malenbaum, R., Herzog, D., Eisenthal, S., & Wyshak, G. (1988). Overeaters anonymous: Impact on bulimia. *International Journal of Eating Disorders, 7,* 139–143.

Maloney, M. J., McGuire, J. B., & Daniels, S. R. (1988). Reliability testing of a children's version of the Eating Attitudes Test. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 541–543.

Marrazzi, M., Bacon, J., Kinzie, J., & Luby, E. (1995). Naltrexone use in the treatment of anorexia nervosa and bulimia nervosa. *International Clinical Psychopharmacology, 10,* 163–172.

Martin, F. (1985). Treatment and outcome of anorexia nervosa in adolescence: Prospective study and 5 year follow-up. *Journal of Psychiatric Research, 19,* 509–514.

Masterson, J. F. (1977). Primary anorexia nervosa in the borderline adolescent: An object relations view. In P. Harticollis (Ed.), *Borderline personality disorders* (pp. 475–494). New York: International Universities Press.

McKisack, C., & Waller, G. (1997). Factors influencing the outcome of group psychotherapy for bulimia nervosa. *International Journal of Eating Disorders, 22,* 1–13.

Minuchin, S., Rosman, B. L., & Baker, L. (1978). *Psychosomatic families: Anorexia nervosa in context.* Cambridge, MA: Harvard University Press.

Mitchell, J. E., Davis, L., Goff, G., & Pyle, R. (1986). A follow-up study of patients with bulimia. *International Journal of Eating Disorders, 5,* 441–450.

Mitchell, J. E., & Groat, R. (1984). A placebo-controlled, double blind trial of amitriptyline in bulimia. *American Journal of Psychiatry, 140,* 554–558.

Mitchell, J. E., Pyle, R. L., Hatsukami, D., Goff, G., Glotter, D., & Harper, J. (1989). A 2–5 year follow-up study of patients treated for bulimia. *International Journal of Eating Disorders, 8,* 157–165.

Mitchell, J. E., Pyle, R. L., Eckert, E. D., Hatsukami, D., Pomeroy, C., & Zimmerman, R. (1990). A comparison study of antidepressants and structured intensive group psychotherapy in the treatment of bulimia nervosa. *Archives of General Psychiatry, 47,* 149–157.

Mitchell, J. E., Raymond, N., & Specker, S. (1993). A review of the controlled trials of pharmachotherapy and psychotherapy in the treatment of bulimia nervosa. *International Journal of Eating Disorders, 14,* 229–247.

Mitchell, J. S., Sloan, D. M., Ruderich, S., Freiheit, S., & Silverman, E. (1996). The efficacy of sertraline in bulimia nervosa patients: A preliminary report. Presented at the Eating Disorders Research Society Meeting, Pittsburgh, PA.

Mizes, J. S., Sloan, D., Ruderich, S., Freiheit, S., & Silverman, E. (1996, November). *The efficacy of sertraline in bulimia nervosa patients: A preliminary report.* Paper presented at the Eating Disorders Research Society Meeting, Pittsburgh, PA.

Olivardia, R., Pope, H., Mangweth, B., & Hudson, J. (1995). Eating disorders in college men. *American Journal of Psychiatry, 152,* 1279–1285.

Olmsted, M. P., Davis, R., Garner, D. M., Eagle, M. & Rockert, W. (1991). Efficacy of brief psychoeducational intervention for bulimia nervosa. *Behavior Research and Therapy, 29,* 71–83.

Oppenheimer, R., Howells, K., Palmer, R. L., & Chaloner, D. A. (1985). Adverse sexual experience in childhood and clinical eating disorders: A preliminary description. *Journal of Psychiatric Research, 19,* 357–361.

Ordman, A. M., & Kirschenbaum, D. S. (1985). Cognitive-behavioral therapy for bulimia: An initial outcome study. *Journal of Consulting and Clinical Psychology, 53,* 305–313.

Patton, G. C., Johnson-Sabine, E., Wood, K., Mann, A. H., & Wakeling, A. (1990). Abnormal eating attitudes in London schoolgirls—A prospective epidemiological study: outcome at 12 month follow-up. *Psychological Medicine, 20,* 383–394.

Perlick, D., & Silverstein, B. (1994). Faces of female discontent: Depression, disordered eating, and changing gender roles. In P. Fallon, M. A. Katzman, & S. C. Wooley (Eds.), *Feminist perspectives on eating disorders* (pp. 77–93). New York: Guilford Press.

Piran, N., Lerner, P., Garfinkel, P. E., Kennedy, S. H., & Brouilette, C. (1988). Personality disorders in anorexia patients. *International Journal of Eating Disorders, 5,* 589–599.

Polivy, J., & Herman, C. P. (1985). Dieting and bingeing: A causal analysis. *American Psychologist, 40,* 193–201.

Pope, H. G., & Hudson, J. I. (1989). Are eating disorders associated with borderline personality disorder? A critical review. *International Journal of Eating Disorders, 8,* 1–9.

Powers, P. (1997). Management of patients with comorbid medical conditions. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 424–436). New York: Guilford Press.

Puglifse, M. T., Lifshitz, F., Grad, G., Fort, P., & Marks-Katz, M. (1983). Fear of obesity: A cause of short stature and delayed puberty. *The New England Journal of Medicine, 309,* 513–518.

Pyle, R. L., Mitchell, J. E., Eckert, E. D., Hastsukami, D., Pomeroy, C., & Zimmerman, R. (1990). Maintenance treatment and 6 month outcome for bulimic patients who respond to initial treatment. *American Journal of Psychiatry, 147,* 871–875.

Rampling, D. (1985). Ascetic ideals and anorexia nervosa. *Journal of Psychiatric Research, 19,* 89–94.

Rastam, M. (1992). Anorexia nervosa in 51 Swedish adolescents: Premorbid problems and comorbidity. *Journal of the American Academy of Child & Adolescent Psychiatry, 31,* 819–829.

Rastam, M., Gillberg, C., & Gillberg, C. (1995). Anorexia nervosa 6 years after onset: Part II. comorbid psychiatric problems. *Comprehensive Psychiatry, 36,* 70–76.

Raymond, N. C., Mitchell, J. E., Fallon, P., & Katzman, M. A. (1994). A collaborative approach to the use of medication. In P. Fallon, M. Katzman, & S. C. Wooley (Eds.), *Feminist perspectives on eating disorders* (pp. 231–250). New York: Guilford Press.

Robin, A. L., Siegel, P., & Moye, A. (1995). Family versus therapy for anorexia: Impact on family conflict. *International Journal of Eating Disorders, 17,* 313–322.

Root, M. M. P., Fallon, P., & Friedrich, W. N. (1985). *Bulimia: A systems approach to treatment.* New York: Norton.

Rosen, J. C., & Leitenberg, H. (1985). Exposure plus response prevention treatment of bulimia. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of psychotherapy for anorexia nervosa and bulimia* (pp. 193–209). New York: Guilford Press.

Rosen, A. M., Murkofsky, C. A., Steckler, N. M., & Skolnick, N. J. (1989). A comparison of psychological and depressive symptoms among restricting anorexic, bulimic anorexic, and normal-weight bulimic patients. *International Journal of Eating Disorders, 8,* 657–663.

Rothenberg, A. (1988). Differential diagnosis of anorexia nervosa and depressive illness. A review of 11 studies. *Comprehensive Psychiatry, 29,* 427–432.

Russell, G. F. M. (1970). Anorexia nervosa: Its identity as an illness and its treatment. In J. H Price (Ed.), *Modern trends in psychological medicine* (Vol. 2, pp. 131–164). London: Butterworths.

Russell, G. F. M. (1979). Bulimia nervosa: An ominous variant of anorexia nervosa. *Psychological Medicine, 9,* 429–448.

Russell, G. F. M. (1985). The changing nature of anorexia nervosa: An introduction to the conference. *Journal of Psychiatric Research, 19,* 101–109.

Russell, G. F. M. (1997). The history of bulimia nervosa. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 11–24). New York: Guilford.

Russell, G. F. M., Szmukler, G. I., Dare, C., & Eisler, I. (1987). An evaluation of family therapy in anorexia nervosa and bulimia nervosa. *Archives of General Psychiatry, 44,* 1047–1056.

Sands, S. (1992). Bulimia dissociation, and empathy: A self-psychological view. In C. Johnson (Ed.), *Psychodynamic treatment of anorexia nervosa and bulimia* (pp. 34–49). New York: Guilford Press.

Santrock, J. W., Minnett, A. M., & Campbell, B. D. (1994). *The authoritative guide to self-help books.* New York: Guildford Press.

Schmidt, U., Jiwany, A., & Treasure, J. (1993). A controlled study of alexithymia in eating disorders. *Comprehensive Psychiatry,* 1, 54–58.

Schwartz, R. C., Barrett, M. J., & Saba, G. (1985). Family therapy for bulimia. In D. M. Garner & P. E. Garfinkel (Eds.) *Handbook of psychotherapy for anorexia nervosa and bulimia,* (pp. 280–307). New York: Guilford.

Scott, D. W. (1986). Anorexia nervosa: A review of possible genetic factors. *International Journal of Eating Disorders, 5,* 1–20.

Selvini-Palazzoli, M. P. (1974). *Self-starvation.* London: Chaucer.

Serdula, M. K., Collins, M. E., Williamson, D. F., Anda, R. F., Pamuk, E. R., & Byers, T. E. (1993). Weight control practices of US adolescents and adults. *Annals of Internal Medicine, 119,* 667–671.

Silverman, J. A. (1974). Anorexia nervosa: Clinical observations in a successful treatment plan. *Journal of Pediatrics, 84,* 68–73.

Sohlberg, S., Norring, C., Holmgren, S., & Rosmark, B. (1989). Impulsivity and long-term prognosis of psychiatric patients with anorexia nervosa/bulimia nervosa.

Journal of Nervous and Mental Disease, 177, 249–258.

Steiner-Adair, C. (1994). The politics of prevention. In P. Fallon, M. A. Katzman, & S. C. Wooley (Eds.), *Feminist perspectives on eating disorders* (pp. 381–394). New York: Guilford Press.

Steinhausen, H. C., & Glanville, K. (1983). A long-term follow-up of adolescent anorexia nervosa. *Acta Psychiatrica Scandinavica, 68,* 1–10.

Steinhausen, C. H., Rauss-Mason, C., & Seidel, R. (1991). Follow-up studies of anorexia nervosa: A review of four decades of outcome research. *Psychological Medicine, 21,* 447–454.

Stern, S. (1986). The dynamics of clinical management in the treatment of anorexia nervosa and bulimia: An organizing theory. *International Journal of Eating Disorders, 5,* 233–254.

Striegel-Moore, R. H. (1994). A feminist agenda for psychological research on eating disorders. In P. Fallon, M. A. Katzman, & S. C. Wooley (Eds.), *Feminist perspective on eating disorders* (pp. 438–454). New York: Guilford Press.

Striegel-Moore, R. H., Silberstein, L. R., & Rodin, J. (1986). Toward an understanding of risk factors in bulimia. *American Pychologist, 41,* 246–263.

Strober, M. (1981). A comparative analysis of personality organization in juvenile anorexia nervosa. *Journal of Youth and Adolescence, 10,* 285–295.

Strober, M. (1983). An empirically derived typology of anorexia nervosa. In P. L. Darby, P. E. Garfinkel, D. M. Garner, & D. V. Coscina (Eds.), *Anorexia nervosa: Recent developments in research* (pp. 185–198). New York: Liss.

Strober, M. (1991). Disorders of the self in anorexia nervosa: An organismic-developmental paradigm. In C. Johnson (Ed.), *Psychodynamic theory and treatment for eating disorders.* New York: Guilford Press.

Strober, M. (1997). Consultation and therapeutic engagement in severe anorexia nervosa. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 229–247). New York: Guilford Press.

Strober, M., & Katz, J. L. (1988). Depression in the eating disorders: A review and analysis of descriptive, family, and biological findings. In D. M. Garner & P. E. Garfinkel (Eds.), *Diagnostic issues in anorexia nervosa and bulimia nervosa* (pp. 80–111). New York: Brunner/Mazel.

Strober, M., Lampert, C., Morrell, W., Burroughs, J., & Jacobs, C. (1990). A controlled family study of anorexia nervosa: Evidence of familial aggregation and lack of shared transmission with affective disorders. *International Journal of Eating Disorders, 9,* 239–53.

Strober, M., & Yager, J. (1985). A developmental perspective on the treatment of anorexia nervosa in adolescents. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of psychotherapy for anorexia nervosa and bulimia* (pp. 363–390). New York: Guilford Press.

Sugarman, A. (1991). Bulimia: A displacement from psychological self to body self. In C. Johnson (Ed.), *Psychodynamic treatment of anorexia nervosa and bulimia* (pp. 3–33). New York: Guilford Press.

Sugarman, A., Quinlan, D., & Devenis, L. (1982). Ego boundary disturbance in anorexia: Preliminary findings. *Journal of Personality Assessment, 46,* 455–461.

Sullivan, P. F. (1995). Mortality in anorexia nervosa. *American Journal of Psychiatry, 152,* 1073–1074.

Swift, W. J., & Letven, R. (1984). Bulimia and the basic fault. *Journal of the American Academy of Child and Adolescent Psychiatry, 23,* 489–497.

Swift, W. J., & Wonderlich, S. A. (1988). Personality factors and diagnosis in eating disorders: Traits, disorders, and structures. In D. M. Garner & P. E. Garfinkel (Eds.), *Diagnostic issues in anorexia nervosa and bulimia nervosa* (pp. 112–165). New York: Brunner/Mazel.

Telch, C. F., Agras, W. S., Rossiter, E. M., Wilfley, D., & Kenardy, J. (1990). Group cognitive-behavioral treatment for the non-purging bulimic: An initial evaluation. *Journal of Consulting and Clinical Psychology, 58,* 629–635.

Thackwray, D., Smith, M., Bodfish, J., & Meyers, A. (1993). A comparison of behavioral and cognitive-behavioral interventions for bulimia nervosa. *Journal of Consulting & Clinical Psychology, 61,* 639–645.

Theander, S. (1985). Outcome and prognosis in anorexia nervosa and bulimia: Some results of previous investigations, compared with those of a Swedish long-term study. *Journal of Psychiatric Research, 19,* 493–508.

Theander, S. (1996). Anorexia nervosa with an early onset: Selection, gender, outcome, and results of a long-term follow-up study. *Journal of Youth and Adolescence, 25,* 419–429.

Thiel, A., Broocks, A., Ohlmeier, M., Jocoby, G., & Schuβler, G. (1995). Obsessive-compulsive disorder among patients with anorexia nervosa and bulimia nervosa. *American Journal of Psychiatry, 152,* 72–75.

Tiller, J., Schmidt, U., & Treasure, J. (1993). Compulsory treatment for anorexia nervosa: Compassion or coercion? *British Journal of Psychiatry, 162,* 679–680.

Timmerman, M. G., Wells, L. A., & Chen, S. (1990). Bulimia nervosa and associated alcohol abuse among secondary school students. *American Academy of Child and Adolescent Psychiatry, 29,* 118–122.

Toner, B. B., Garfinkel, P. E., & Garner, D. M. (1986) Long-term follow-up of anorexia nervosa. *Psychosomatic Medicine, 48,* 520–529.

Treasure, J., Schmidt, U., Troop, N., Tiller, J., Todd, G., Keilen, M., & Dodge, E. (1994). First step in managing bulimia nervosa: Controlled trial of therapeutic manual. *British Medical Journal, 308,* 686–687.

Treasure, J., Todd, G., Brolly, M., Tiller, J., Nehmed, A., & Denman, F. (1995). A pilot study of a randomized trial of cognitive analytical therapy vs. educational behavioral therapy for adult anorexia nervosa. *Behavioural Research and Therapy, 33,* 363–367.

Vandereycken, W. (1984). Neuroleptics in the short-term treatment of anorexia nervosa: A double-blind placebo-controlled study with sulpiride. *British Journal of Psychiatry, 144,* 288–292.

Vandereyecken, W., & Pierloot, R. (1982). Pimozide combined with behavior therapy in the short-term treatment of anorexia nervosa. *Acta Psychiatry Scandinavia, 66,* 445–450

Vandereycken, W., Kog, E., & Vanderlinden, J. (1989). *The family approach to eating disorders.* New York: PMA.

Vigersky, R. A., & Loriaux, D. L. (1977). Anorexia nervosa as a model of hypothalamic dysfunction. In R. Vigersky (Ed.), *Anorexia nervosa* (pp. 109–122). New York: Raven Press.

Vitousek, K. B., & Ewald, L. S. (1993). Self-representation in eating disorders: A cognitive perspective. In: Z. Segal & S. Blatt (Eds.), *The self-in emotional disorders: Cognitive and psychodynamic perspectives* (pp. 221–257). New York: Guilford Press.

Vitousek, K. B., & Rimoto, L. (1993). Cognitive-behavioral models of anorexia nervosa, bulimia nervosa, and obesity. In P. Kendal & K. Dobson (Eds.), *Psychopathology and cognition* (pp. 191–142). New York: Academic Press.

Walsh, B. T., & Devlin, M. J. (1995). Psychopharmacology of anorexia nervosa, bulimia nervosa, and binge eating. In F. Bloom & D. Kupfer (Eds.), *Psychopharmacology: The fourth generation of progress.* New York: Raven Press.

Walsh, B. T., & Garner D.M. (1997). Diagnostic issues. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of* *treatment for eating disorders* (pp. 25–33), New York/London: Guilford Press.

Walsh, B. T., Gladis, M., Roose, S. P., Stewart, J. W., Stetner, F., & Glassman, A. H. (1988). Phenelzine vs. placebo in 50 patients with bulimia. *Archives of General Psychiatry, 45,* 471–475.

Walsh, B. T., Hadigan, C. M., Devlin, M. J. Gladis, M., & Roose, S. P. (1991). Long-term outcome of antidepressant treatment for bulimia nervosa. *American Journal of Psychiatry, 148,* 1206–1212.

Walsh, B. T., Wilson, G. T., Loeb, K. L., Devlin, M. J., Pike, K. M., Roose, S. P., Fleiss, J., & Waternaux, C. (1997). Medication and psychotherapy in the treatment of bulimia nervosa. *American Journal of Psychiatry, 154,* 523–531.

Wardle, J. (1987). Compulsive eating and dietary restraint. *British Journal of Clinical Psychology, 26,* 47–55.

Wardle, J., & Beales, S. (1986). Restraint, body image and food attitudes in children from 12 to 18 years. *Appetite, 7,* 209–217.

Waters, B. G. H., Beaumont, P. J. V., Touyz, S., & Kennedy, M. (1990). Behavioural differences between twin and non-twin female sibling pairs discordant for anorexia nervosa. *International Journal of Eating Disorders, 9,* 265–273.

Welch, G. W., Hall, A., & Renner, R. (1990). Patient subgrouping in anorexia nervosa using psychologically-based classification. *International Journal of Eating Disorders, 9,* 311–322.

Wilfley, D. E., Agras, W. S., Telch, C. F., Rossiter, E. M., Schneider, J. A., Cole, A. G., Stifford, L., & Raeburn, S. D. (1993). Group cognitive-behavioral therapy and group interpersonal psychotherapy for the nonpurging bulimic individual: A controlled comparison. *Journal* of Consulting and Clinical Psychology, 2, 296–305.

Williamson, D.A. (1990). *Assessment of eating disorders: besity, anorexia, and bulimia nervosa.* New York: Pergamon.

Wilson, G.T. (1991). The addiction model of eating disorders: A critical analysis. *Advances in Behaviour Research and Therapy, 13,* 27–72.

Wilson, G. T., & Lindholm, L. (1987). Bulimia nervosa and depression. *International Journal of Eating Disorders, 6,* 725–732.

Wilson, G.T. (1993). Assessment of binge eating. In C. G. Fairburn & G. T. Wilson (Eds.), *Binge eating: Nature assessment, and treatment* (pp. 227–249). New York: Guilford Press.

Wolchik, S. A., Weiss, L., & Katzman, M.A., (1986). An empirically validated, short-termed psychoeducational group treatment program for bulimia. *International Journal of Eating Disorders, 5,* 21–34.

Wolf, E. & Crowther, J. (1992). An evaluation of behavioral and cognitive-behavioral group interventions for the treatment of bulimia nervosa in women. *International Journal of Eating Disorders, 11,* 3–15.

Wonderlich, S. A., Swift, W. J., Slotnick, H. B., & Goodman, S. (1990). DSM-III-R personality disorders in eating-disorder subtypes. *International Journal of Eating Disorders, 9,* 607–616.

Wooley, S. C., & Lewis, K.G. (1989). The missing woman: Intensive family-oriented treatment of bulimia. *Journal of Feminist Family Therapy, 1,* 61–83.

World Health Organization (1992). *The ICD-10 classification of mental and behavioral disorders: Clinical descriptions and diagnostic guidelines* (pp. 176–181). Geneva, Switzerland: Author.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.24
Sleep Disorders

AVI SADEH and REUT GRUBER
Tel Aviv University, Israel

5.24.1 INTRODUCTION 630

5.24.2 SLEEP DISORDERS IN CHILDREN 631

- *5.24.2.1 Phenomenology* 631
  - *5.24.2.1.1 Primary insomnia* 631
  - *5.24.2.1.2 Primary hypersomnia* 632
  - *5.24.2.1.3 Narcolepsy* 632
  - *5.24.2.1.4 Breathing-related sleep disorder* 632
  - *5.24.2.1.5 Circadian rhythm sleep disorder* 633
  - *5.24.2.1.6 Nightmare disorder* 633
  - *5.24.2.1.7 Sleep terror disorder* 633
  - *5.24.2.1.8 Sleepwalking disorder* 634
  - *5.24.2.1.9 Other parasomnias* 634
- *5.24.2.2 Diagnostic Features* 634
  - *5.24.2.2.1 Primary insomnia* 634
  - *5.24.2.2.2 Primary hypersomnia* 634
  - *5.24.2.2.3 Narcolepsy* 635
  - *5.24.2.2.4 Breathing-related sleep disorder* 635
  - *5.24.2.2.5 Circadian rhythm sleep disorder* 635
  - *5.24.2.2.6 The parasomnias* 635
- *5.24.2.3 Prevalence* 636
  - *5.24.2.3.1 Primary insomnia* 636
  - *5.24.2.3.2 Primary hypersomnia and narcolepsy* 636
  - *5.24.2.3.3 Breathing-related sleep disorder* 637
  - *5.24.2.3.4 Circadian rhythm sleep disorder* 637
  - *5.24.2.3.5 The parasomnias* 637
- *5.24.2.4 Etiology* 637
  - *5.24.2.4.1 Primary insomnia* 637
  - *5.24.2.4.2 Primary hypersomnia and narcolepsy* 638
  - *5.24.2.4.3 Breathing-related sleep disorder* 638
  - *5.24.2.4.4 Circadian rhythm sleep disorder* 639
  - *5.24.2.4.5 The parasomnias* 639
- *5.24.2.5 Developmental Course* 639

5.24.3 CONCEPTUALIZATION AND CLINICAL FORMULATION 639

- *5.24.3.1 Multimodal Assessment* 640
  - *5.24.3.1.1 Initial information* 640
  - *5.24.3.1.2 Assessment of sleep* 640
  - *5.24.3.1.3 Assessment of excessive sleepiness* 642
- *5.24.3.2 Treatment Formulation* 642

5.24.4 TREATMENT 642

- *5.24.4.1 Psychosocial Treatments* 642
  - *5.24.4.1.1 Primary insomnia* 642
  - *5.24.4.1.2 Primary hypersomnia and narcolepsy* 644
  - *5.24.4.1.3 Breathing-related sleep disorder* 644
  - *5.24.4.1.4 Circadian rhythm sleep disorder* 644
  - *5.24.4.1.5 Nightmares, sleep terror, sleepwalking, and other parasomnias* 645

5.24.4.2 Pharmacological and Other Medical Treatments 646
5.24.4.2.1 Primary insomnia 646
5.24.4.2.2 Primary hypersomnia and narcolepsy 646
5.24.4.2.3 Breathing-related sleep disorder 647
5.24.4.2.4 Circadian rhythm sleep disorder 647
5.24.4.2.5 Nightmares, sleep terror, sleepwalking, and other parasomnias 648
5.24.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 648
5.24.6 SUMMARY 648
5.24.7 REFERENCES 649

5.24.1 INTRODUCTION

Sleep or rest appears to be a basic need of every organism. Despite many years of sleep research, many of the fundamental questions such as the function of sleep have only been partially resolved or replaced by more specific questions, such as the function of rapid eye movement (REM) sleep or the influence of daytime naps on functioning.

In the attempt to unveil the function of sleep, research efforts have focused mainly on the effects of sleep deprivation on the individual's mental and physiological functioning. It has been shown that significant and prolonged sleep deprivation can lead to serious cognitive and behavioral deficits and can eventually cause medical complications and even death in extreme cases. The effects of partial sleep deprivation, which is closer to the phenomenology of sleep disorders, have been less striking, but in general the literature suggests that the partially sleep-deprived individual may pay a significant price in behavioral, cognitive, and medical terms (e.g., Babkoff, Caspy, Mikulincer, & Sing, 1991; Bonnet, 1985). Despite the fact that extensive research has been conducted on the effects of sleep deprivation in animals and in adults, only limited research efforts have been focused on children. These studies have demonstrated relationships between sleep disorders and learning and behavior problems in children and adolescents (see Alvarez, Dahlitz, Vignau, & Parkes, 1992; Dahl, 1996; Kahn, Mozin, Rebuffat, Sottiaux, & Muller, 1989; Quine, 1992; Thorpy, Korman, Spielman, & Glovinsky, 1988).

From a clinical perspective, sleep is a very sensitive barometer to the medical and mental well-being of the individual. The most prominent manifestation of the close relationships between sleep and psychopathology and medical disorders is the presence of sleep difficulties or sleep-related complaints as symptoms or even as major diagnostic features of many psychiatric and medical disorders.

The close ties between sleep disorders and psychopathology are best manifested in professional diagnostic systems such as the *DSM-IV* (American Psychiatric Association, 1994). In addition to more than 50 manual pages dedicated to sleep disorders, sleep-related problems are often included in the symptoms or diagnostic criteria of other major affective and anxiety disorders (see Table 1). In addition, it has been demonstrated repeatedly that sleep is very sensitive to psychosocial stressors. Extensive research has documented the vulnerability of the sleep–wake system to various experimental and natural stressors, ranging from sleeping in an unfamiliar setting (i.e., the sleep laboratory) to serious traumatic life events (see Sadeh, 1996, for a recent review).

Table 1 Sleep-related difficulties associated with *DSM-IV* diagnoses of the American Psychiatric Association, 1994.

Mood disorders

Major depressive episode: insomnia or hypersomnia, fatigue and loss of energy, irritable mood in children

Manic episode: decreased need for sleep, and distractibility

These symptoms are manifested in less severe forms of affective disorders.

Anxiety disorders

Generalized anxiety disorder: sleep disturbances (difficulty in falling or staying asleep, or restless unsatisfying sleep), being easily fatigued, irritability and restlessness

Post-traumatic stress disorder: difficulty in falling or staying asleep, difficulty concentrating, hypervigilance, recurrent distressing dreams (nightmares)

Separation anxiety disorder: persistent reluctance or refusal to go to sleep without being near a major attachment object, repeated nightmares involving the theme of separation

Source: American Psychiatric Association (1994).

Disrupted sleep has also been associated with difficult temperament in early childhood (e.g., Carey, 1974; Sadeh, Lavie, & Scher, 1994; Zuckerman, Stevenson, & Baily, 1987). It has also been shown that older children who suffer from sleep problems are more likely to exhibit other behavior problems (Gass & Strauch, 1984; Jenkins, Bax, & Hart, 1980; Richman, Stevenson, & Graham, 1982).

Most of the studies demonstrating strong links between sleep and behavior problems or psychopathology are correlative in nature and therefore preclude a better understanding of the cause-and-effect nature of these relationships (see Dahl, 1996; Sadeh, Hayden, McGuire, Sachs, & Civita, 1994; Sadeh et al., 1995c). This important issue is evident in the case of attention deficit and hyperactivity disorder (ADHD). A number 882044of studies have documented higher rates of sleep problems in children diagnosed with ADHD compared to normal children (see Dahl, 1996). These studies included subjective and objective measures that supported the notion that the sleep of ADHD children is significantly more disturbed or shorter (e.g., Tirosh, Sadeh, Munvez, & Lavie, 1993). Traditionally, clinicians and parents tended to interpret this finding intuitively by attributing the sleep problem to the fact that the child is running in "turbo" mode or is hyperalert, and therefore, cannot stop and initiate sleep or is still hyperactive during sleep. Research supports the competing hypothesis that sleep deprivation in children may lead to behaviors or symptoms associated with ADHD such as inattention, difficulties in concentration and sustaining attention, impulsiveness, and increased activity level (Anders, Carskadon, Dement, & Harvey, 1978; Navelet, Anders, & Guilleminault, 1976). There have been some descriptions of clinical cases of children diagnosed with ADHD that suffered from a major sleep disorder, where treating the underlying sleep disorder led to disappearance of the ADHD symptoms (Dahl, 1996; Dahl, Pelham, & Wierson, 1991).

Sleep disruptions have also been associated with various medical disorders. For example, asthma is considered to be one of the most common illnesses during childhood, affecting 5–15% of children (e.g., Strachan, Anderson, Limb, O'Neill, & Wells, 1994). Asthma is a disorder of the respiratory system which tends to become more dysfunctional during sleep. Indeed, sleep difficulties are very often associated with asthma during childhood, with a survey in 1995 reporting that as many as 61% of asthmatic children suffer from sleep difficulties (Madge et al., 1995).

5.24.2 SLEEP DISORDERS IN CHILDREN

Sleep disorders cover a wide range of phenomena associated with the multiple sleep–wake systems related to the specific mechanisms responsible for consolidated and refreshing sleep, inhibition of behavior during sleep or sleep–wake transitions, control of breathing during sleep, and the timing of sleep ("the biological clock").

A number of diagnostic systems exist for sleep disorders (e.g., American Psychiatric Association, 1994; Diagnostic Classification Steering Committee, 1990; World Health Organization, 1992). These systems vary in their emphasis on phenomenology, etiology, and medical or psychosocial factors. The following description will be based on the *DSM-IV* (American Psychiatric Association, 1994). Due to space limitations, many issues will be only partially addressed or not addressed at all. For additional information, the reader is referred to professional books on children's sleep disorders (Ferber & Kryger, 1995; Guileminault, 1987; Sheldon, Spire, & Levy, 1992).

The first diagnostic distinction used by *DSM-IV* with regard to sleep disorders is between the dyssomnias and the parasomnias. Dysomnias refers to the disorders of initiating or maintaining sleep and to those related to excessive sleepiness. These disorders include: primary insomnia, primary hypersomnia, narcolepsy, breathing-related sleep disorder, and circadian rhythm 1651828sleep disorder. 1472630Parasomnias are disorders manifested in unique behaviors or physiologic phenomena occurring during sleep or sleep–wake transitions.

5.24.2.1 Phenomenology31890000

The clinical presentation of various sleep disorders often changes as the child matures, and the involvement of parents as primary observers and reporters changes accordingly.

5.24.2.1.1 Primary insomnia

Primary insomnia is related to disorders in initiating and maintaining sleep (DIMS). This diagnosis covers the most common sleep complaints during childhood and include problems such as difficulty going to bed, difficulty falling asleep, and multiple and/or prolonged night-wakings. This category is equivalent to the *ICD-10* classification of "nonorganic insomnia" (World Health Organization, 1992). The clinical picture is related to the developmental status of the child, as are the source of the referral, the etiology, and the clinical interventions.

During infancy and early childhood, the most common complaints (usually of the parents) are related to the child's incapacity to "sleep through the night" which is manifested in difficulties such as falling asleep without parental involvement, multiple and prolonged night-wakings during which the child usually signals (by crying or calling), and requires parental intervention to resume sleep. In most cases, the child does not exhibit strong manifestations of distress during the day and compensates during daytime napping for the sleepless nights. The parents are often exhausted by their child's sleep difficulties, and in many cases they show signs of distress and tension that may lead to difficulties in their daily functioning and their relationship with the child.

As children mature, during the preschool and school ages, the clinical features of insomnia undergo significant changes. The dominance of night-wakings and difficulties in resuming sleep following nocturnal awakening during early childhood is replaced by sleep onset difficulties, bed refusal, fears, and bedtime struggles. When fears are involved the child often seeks the proximity of the parents during the night and may end up cosleeping with the parents. The child is usually able to address these fears and the subsequent reluctance to go to sleep or difficulties to resume sleep after night-waking. Older children and adolescents may lie in bed for prolonged periods unable to fall asleep or resume sleep. In many cases, the parents are unaware, or are only partially aware of the problem, because the child no longer seeks their attention.

Parental and cultural attitudes, and the different lifestyle of the family, may have a significant influence on whether a specific set of problems or behaviors are considered a sleep problem. For instance, parent–child cosleeping is considered a normal and acceptable practice in many cultures and by many parents. In these families, cosleeping is a habitual sleeping arrangement that meet the needs of the parents and their culture. However, in western societies, cosleeping is usually an unfavored practice that generally represents an attempt to solve a child's sleep problem.

5.24.2.1.2 *Primary hypersomnia*

Excessive sleepiness is usually manifested by very extended sleep periods and an unusual pressure to sleep during the daytime. This relatively rare cluster of disorders is not fully understood and may be the result of different neurologic or physiologic origins. One recurrent form of the disorder, the Kleine–Levin syndrome has, in addition to the excessive sleep episodes, other clinical manifestations that may be confused with psychiatric disorders. These behaviors include indiscriminate and inappropriate sexual behavior, compulsive overeating, and in some individuals other symptoms such as depression, confusion, irritability, and hallucinations. The Kleine–Levin syndrome is relatively rare and more prevalent in males (approximately 3:1 ratio). The age at onset is usually middle or late puberty but no prevalence estimates exist for children or adolescents.

ICD-10 specifies nonorganic hypersomnia as a separate classification that is usually associated with other mental disorders such as affective disorders (World Health Organization, 1992).

5.24.2.1.3 *Narcolepsy*

Unintentional sleep episodes are common phenomena in sleepy or sleep-deprived individuals. However, in narcolepsy, sleep attacks unexpectedly, regardless of the situation and in combination with cataplexy (sudden loss of muscular tone in response to emotional stimuli), hypnogogic hallucination, and loss of muscle tone (paralysis). The disorder is very serious and highly disruptive to normal life. Narcolepsy is associated with excessive daytime sleepiness, difficulty waking in the morning, and in many cases with hyperactivity (Guilleminault, 1987).

5.24.2.1.4 *Breathing-related sleep disorder*

The maintenance of continuous breathing during wakefulness as well as during sleep is a crucial function of the central nervous system (CNS) and other peripheral systems. During sleep, the breathing system is more vulnerable and more likely to fail than during wakefulness. Failures in breathing during sleep lead to arousals, fragmented, and nonrefreshing sleep. Clinical complaints in children may include excessive sleepiness and low level of arousal, or other symptoms associated with fatigue in children such as irritability and hyperactivity. Observers (usually the parents) may report loud snoring and long episodes of breathing pauses during sleep terminated by a loud and active breathing effort. Most affected children (as well as adults) are unaware of their snoring, breathing difficulties, and sleep fragmentation. Sometimes the child may even be considered a "good sleeper," since they readily go to bed and sleep through the night and for an extended duration. In addition, children with obstructive sleep apnea may experience breathing problems during wakefulness, may be more susceptible to complications in diseases associated with upper airways blockage, and present features such as swallowing difficulties and poor speech articulation.

Breathing-related sleep disorders are further subdivided into obstructive, central, and mixed sleep apnea and other nonapnea disorders. A discussion of these subcategories is beyond the scope of this chapter.

5.24.2.1.5 Circadian rhythm sleep disorder

The evolution of sleep–wake rhythms that favors the dark hours for nocturnal sleep occurs during the first few months of life. “Sleeping through the night” is the expression used to describe this maturational milestone. However, this association between the light–dark and the sleep–wake cycles cannot be taken for granted. Circadian rhythm sleep disorders (CRSD) relate to all the deviations or the incongruity between these basic rhythms. Children with CRSD may experience difficulties falling asleep at normal, age-appropriate evening hours and waking up in the morning, but would be “good sleepers” given the opportunity to go to sleep at very late hours and wake-up later in the morning (delayed sleep phase type, DSPT; Weitzman et al., 1981). Other children may experience the opposite problem, the advanced sleep phase, manifested in difficulty staying awake until normal evening bedtime hours with a tendency to wake up for the day at a very early morning hour. Other possible manifestations include the irregular sleep–wake patterns and the non-24-hour sleep–wake patterns that describe deviations from a regular sleep–wake schedule. These disorders, if not attributed to organic factors, fall under the *ICD-10* category of nonorganic disorder of the sleep–wake schedule (World Health Organization, 1992).

Children and adolescents with CRSD may experience serious problems when forced to adapt to normal social schedules such as early school hours. When they encounter serious difficulties waking up in the morning for the school day they are often considered unmotivated and face disciplinary actions from their parents and school authorities who fail to appreciate the potential biological origin of their problem.

Irregular and non-24-hour sleep–wake patterns are often reported in blind children and in children with pervasive developmental disorders or other neurobehavioral disorders (Lapierre & Dumont, 1995; Okawa et al., 1987; Sadeh, Klitzke, Anders, & Acebo, 1995; Tzischinsky, Skene, Epstein, & Lavie, 1991)

In addition to the above disorders of the biological clock, some externally induced schedule disorders exist. These disorders include the rhythm disturbances associated with the jet-lag phenomenon or with other forced shifts of the sleep–wake schedule. These disorders are usually transient and disappear with time after the normal sleep-wake schedule is reinstated.

5.24.2.1.6 Nightmare disorder

This is the first representative of the heterogeneous group of the parasomnias. The main feature of the nightmare disorder is the repeated occurrence of terrifying dreams that result in the child’s awakening. Following the nightmare, the child often exhibits high levels of anxiety that may lead to refusal to go to bed, or to stay alone in the dark. Young children usually respond to their parents’ attempts to soothe them but may insist on parental proximity for prolonged periods.

Verbal children may report the content of the dream and it may relate to their daily lives or their normal anxieties such as getting lost or being separated from their parents. Children may also report scary dreams about monsters, evil animals, or other creations of their imagination, or those they were exposed to while watching TV or movies. Other prevalent nightmares are those related to loss of control (e.g., free fall from different places) and a variety of threats of self-injury and death (e.g., drowning). It has been reported that many children refrain from sharing these private negative experiences with their parents and therefore parents may be unaware what their child is experiencing (Terr, 1987). A similar category exists in *ICD-10* (World Health Organization, 1992).

5.24.2.1.7 Sleep terror disorder

The sleep terror disorder is characterized by the repeated occurrence of abrupt awakenings from sleep with a scream or cry that conveys the sense that the child is terrorized. Usually the child continues to yell, cry, and fight during these episodes and is difficult to awaken or comfort. In many cases the child, even with open eyes, does not appear to recognize their parents and may struggle with them as they try to calm the child who also presents signs of autonomic arousal such as tachicardia, rapid breathing, sweating, pupils dilation, increased muscle tone, and flushing of the skin. The child does not remember these events the following morning. If awakened during the episode, the child looks confused and disoriented. Usually the child does not reach full awareness and returns to sleep following the episode. The night terror events commonly last between one and 10 minutes but can be longer if the child is not awakened by their parents. A similar category exists in *ICD-10* (World Health Organization, 1992).

5.24.2.1.8 Sleepwalking disorder

Sleepwalking episodes are common in childhood. The phenomenon may have different manifestations, but typically, the child gets out of bed in the middle of the night and starts wandering around the house. The child may look confused and disoriented but is capable of finding their way around and sometimes performs quite complex activities. Parents may find the child asleep on a different bed in a different room the next morning or may notice the episode while it is occurring and redirect the child to bed to resume normal sleep. Many different behaviors are seen during nightwalking episodes. These behaviors include eating, using the bathroom (sometimes urinating in inappropriate places), talking, or fearful running around. Usually the child has no recollection (or only vague memory) of the sleepwalking episode the following morning. Since these behaviors are highly counterintuitive many parents tend to be alarmed by these events and seek professional help. A similar category exists in *ICD-10* (World Health Organization, 1992).

5.24.2.1.9 Other parasomnias

There are some other unusual behavior patterns that have been associated with sleep and are considered parasomnias. These parasomnias include: REM sleep behavior disorder (activities and behaviors occurring during REM sleep and are associated with vivid dream recall), sleep paralysis (inability to move during sleep–wake transitions), and rhythmical movement disorders, periodic leg movements in sleep, and bruxism (tooth grinding).

5.24.2.2 Diagnostic Features

The diagnostic process of sleep disorders in children should include the following considerations:

(i) the disordered aspect of sleep (e.g., sleep schedule, sleep quality, or related behaviors);

(ii) the primary origin of the disorder (physiologic and organic vs. psychosocial);

(iii) the psychophysiological effects of the disorder (e.g., deteriorated daytime functioning);

(iv) the psychosocial effects of the disorder (e.g., disrupts parent–child relationships);

(v) Other potential external causes (e.g., substance abuse, another medical or psychiatric disorder).

All of the diagnostic systems are based on the principle that a medical origin for the disorder should be excluded (including comorbidity with another major mental disorder) before a psychosocial diagnosis is made. The origins of a number of specific disorders such as circadian rhythm sleep disorders and the parasomnias are not fully understood and therefore such a distinction is not always justified.

5.24.2.2.1 Primary insomnia

The *DSM-IV* diagnostic criteria for primary insomnia are as follows: (i) the predominant complaint is difficulty initiating or maintaining sleep, or nonrestorative sleep, for at least one month; (ii) the sleep disturbance (or associated daytime fatigue) causes clinically significant distress or impairment in social, occupational, or other important areas of functioning; (iii) the sleep disturbance does not occur exclusively during the course of another primary sleep disorder; (iv) the disturbance does not occur exclusively during the course of another mental disorder; and (v) the disturbance is not due to the direct physiological effects of a substance or a general medical condition.

In addition to these standard diagnostic criteria, a differential diagnosis of primary insomnia should exclude the diagnosis of CRSD. Furthermore, difficulties falling asleep at a certain age-appropriate hour may result from the child's anxieties, fears, and other psychosocial factors that would justify the diagnosis of primary insomnia. However, if these difficulties result from the child's biological tendency to fall asleep at a late hour (assuming the child is not experiencing any difficulties), then the child is probably suffering from sleep-phase delay and therefore a diagnosis of CRSD would be warranted.

5.24.2.2.2 Primary hypersomnia

Excessive sleepiness during the day is a normal experience explained by many factors including chronic insufficient sleep or sleep deprivation, transient effects such as jet-lag, and identified sleep disorders such as insomnia or breathing-related disturbances. All these factors must be excluded before considering primary hypersomnia as a diagnosis.

The major *DSM-IV* diagnostic criteria for primary hypersomnia is the predominant complaint of excessive sleepiness for at least one month (or less if recurrent) as evidenced by either prolonged sleep episodes or daytime sleep episodes that occur almost daily. All of the other criteria are similar to those of primary insomnia, with the exception that primary insomnia as well as other primary sleep, mental, and medical disorders must be excluded before reaching a primary hypersomnia diagnosis.

Primary hypersomnia as defined by *DSM-IV* tends to be very rare in general, and particularly so, in children. Most of these disturbances have a late age of onset, usually appearing during adolescence or adulthood. The Kleine–Levin syndrome is distinguished by prolonged sleep episodes and by the clinical manifestation of inappropriate sexual and aggressive behaviors described earlier.

5.24.2.2.3 Narcolepsy

The *DSM-IV* criteria for narcolepsy are as follows: (i) irresistible attacks of sleep that occur daily over at least three months; (ii) the presence of cataplexy (sudden loss of muscle tone, most often in association with intense emotion) and/or recurrent intrusions of elements of REM sleep into the transition between sleep and wakefulness, as manifested by either hypnopompic or hypnagogic hallucinations or sleep paralysis at the beginning or end of sleep episodes; and (iii) the disturbance is not due to the direct physiologic effects or a substance or another general medical condition.

5.24.2.2.4 Breathing-related sleep disorder

The *DSM-IV* uses only two criteria for breathing-related sleep disorder (BRSD): (i) sleep disruption due to a sleep-related breathing problem; and (ii) the disturbance is not better explained by another mental disorder or by direct physiological effects of a substance. Since the child is usually unaware of having disturbed sleep in BRSD and the parents may also be unaware of the disorder, the first diagnostic signs are very often signs of fatigue, sleepiness and other manifestations of difficulties in arousal level. The full diagnosis of BRSD requires a complete laboratory assessment including monitoring of the breathing system during sleep and a medical examination that includes assessment of the airway (nasal airflow, tonsillar size, mandibular, and facial abnormalities). These assessment procedures can verify the existence of a BRSD, its severity, its subtype (e.g., central vs. obstructive apnea), and the appropriate intervention.

5.24.2.2.5 Circadian rhythm sleep disorder

Like many other species in nature, the human mind is programmed to be alert during daytime hours and to rest or sleep during night time hours. The circadian clock is a very powerful mechanism underlying our behavior and functioning. It includes a nonspecific general clock that controls our rest–activity cycle and multiple clocks that affect various distinct domains of human behavior such as different cognitive functions (Monk, 1994; Montagner, De Roquefeuil, & Djakovic, 1992). The initial step in diagnosing CRSD is the identification of a persistent sleep problem that results from a mismatch between the individual's own circadian sleep-wake pattern and the environmentally appropriate sleep–wake schedule. In addition, *DSM-IV* requires the following requirements: clinically significant stress or a functional impairment and the exclusion of another sleep or mental disorder and any substance effects.

As indicated earlier the major diagnostic issue is differentiating between primary insomnia and CRSD (see Section 5.24.5).

5.24.2.2.6 The parasomnias

Isolated and infrequent episodes of the parasomnias are very common in children. Most forms of these episodes are prevalent during early and middle childhood and become uncommon during adolescence. Therefore, it is difficult to draw the line when these episodes should be considered a disorder. The basic guideline is that a diagnosis of a disorder should be made whenever the phenomenon becomes very frequent, involves violent or dangerous behaviors, or results in significant functional impairment. In addition, if the phenomenon persists from childhood to adolescence it is more likely to warrant a diagnosis.

The main diagnostic criteria for nightmare disorder is repeated awakenings from sleep with recall of terrifying dreams that usually involve threats to survival, security, and self-esteem. Very young children may not be able to produce verbal recall but usually appear oriented and communicative. On awakening from these dreams the child rapidly becomes oriented and alert and usually responds well to parental comfort.

Differential diagnosis from sleep terror disorder is based on these features of dream recall, orientation, and interactive responsiveness which are quite distinguishable from the disorientation and relative unresponsiveness that characterized the sleep terror. In addition, sleep terror disorder is associated with severe autonomic arousal response. Differentiation between sleep terror disorder and epilepsy or REM sleep behavior disorder is based on electroencephalogram (EEG) records and specific stereotyped behaviors.

Sleepwalking disorder is characterized by repeated episodes in which the child rises from bed during sleep and walks about relatively unresponsive to others. It is difficult to awaken the child, and once awake the child is usually amnesic regarding the episode. Another feature is that the child's normal behavior is usually restored shortly after a full awakening from the episode.

Another parasomnia that does not fall under specific *DSM-IV* diagnostic category is that of rhythmical movement disorder. This disorder refers to the very common phenomena of self-soothing behaviors performed by the child during sleep-wake transitions. These behaviors, such as body rocking, head rolling, and banging, are common in young children (more than 50%) and are usually considered normal behavior. But when the phenomenon becomes disruptive or is perceived to be dangerous (e.g., an infant that bangs its head fiercely against the wall), parents often become alarmed and seek professional help.

5.24.2.3 Prevalence

The review of the literature on issues pertaining to prevalence and incidence of sleep problems in children is quite disappointing. The number of solid and systematic studies in this field is very small. In most fields, these studies do not usually involve consensus clinical guidelines or diagnostic criteria, and therefore it is difficult to assess the prevalence rates of these disorders using any systematic diagnostic system. However, the existing surveys provide some sense as to which phenomena are more prevalent than the others and their peak age of onset and time course.

5.24.2.3.1 Primary insomnia

The prevalence of primary insomnia and its etiology appears to change systematically during development and therefore close attention should be given to developmental issues when assessing children presenting with insomnia.

Moore and Ucko (1957) defined a night-waking problem when the child awakened and cried at least once between midnight and 5 a.m. on at least four of seven nights, for at least four consecutive weeks. Using this definition and relying on maternal reports, they reported that at three months of age 70% of the infants "slept" through the night, whereas 83% of the six month olds and 90% of the nine month olds "slept" through the night. Richman (1981) defined sleep problem in one-year-old infants as either regular sleep onset fussing periods lasting more than 30 min, or night-waking episodes that required parental intervention occurring at least four nights a week. Other surveys using parental reports indicated that 20–30% of the children suffer from these problems during the first three years (e.g., Adair, Bauchner, Philipp, Levenson, & Zuckerman, 1991; Beltramini & Herzig, 1983; Jenkins, et al. 1980; Johnson, 1991; Richman, 1981; Scher et al., 1995; Zuckerman et al., 1987). Research based on objective sleep measures derived from methods like time-lapse video and actigraphy (activity based monitoring) suggests that night wakings are more prevalent than reported by parents (Anders; 1979; Paret, 1983; Sadeh, Lavie, Scher, Tirosh, & Epstein, 1991). For instance, it was found that nonreferred normal young children (9–24 months of age) woke-up twice a night on average in comparison to an average of more than four night-wakings in referred "sleep-disturbed" infants (Sadeh et al., 1991). Most of these infants resumed their sleep without fussing or requiring help and therefore their parents were not even aware of their night-wakings.

In school-age children the prevalence of primary insomnia appears to drop to 1–5% (Gass & Strauch, 1984; Richman et al., 1982). However, there are indications that these problems are underestimated because older children are less likely to alert their parents to their sleep problems (Anders et al., 1978). Kahn et al. (1989) surveyed predadolescents (aged 8–10 years) using parental reports. A prevalence of 43% for a variety of sleep difficulties was reported by the parents. "Poor sleep," a rough equivalent of primary insomnia, was reported in 14% of the children.

In adolescence, chronic or severe sleep difficulties have been self-reported by 11–33.4% of the youth (Bearpark & Michie, 1987; Kirmil-Gray, Eagleston, Gibson, & Thorensen, 1984; Morrison, McGee, & Stanton, 1992; Price, Coates, Thoresen, & Grinstead, 1978). Since these surveys did not use detailed criteria to identify the exact nature of the problems (e.g., difficulty falling asleep due to primary insomnia vs. delayed sleep phase), it is difficult to provide good estimates of primary insomnia in this age group.

5.24.2.3.2 Primary hypersomnia and narcolepsy

Primary hypersomnia and narcolepsy are relatively scarce in young children and tend to appear in adolescence or adulthood. It is estimated that the prevalence of narcolepsy in the general population ranges between 0.04% and 0.09% (Dement, Carskadon, & Ley, 1973).

Narcolepsy is usually diagnosed in late adolescence or in adulthood. In a study of 400 narcoleptic patients, Yoss and Daly (1960) reported that only 4% were younger than 15 years when their diagnosis was made. Retrospective studies suggest that early signs of the disorder may appear earlier in many children. Challamel et al. (1994) reported on the basis of aggregated studies of 235 adult patients that 34% of the patients presented their first narcoleptic symptoms before they were 15 years old and 16% before they were 10 years old. There are, however, claims that professionals often fail to diagnose narclopsy and therefore the disorder is either underdiagnosed or its diagnosis is seriously delayed (Baird, 1987).

5.24.2.3.3 Breathing-related sleep disorder

The prevalence of BRSD in children is unknown. It is certainly not a rare disorder, but data on prevalence in the general population is lacking.

5.24.2.3.4 Circadian rhythm sleep disorder

The prevalence of CRSD in children has not been established but the awareness of these disorders is growing rapidly. During adolescence when there is a normal tendency to delay the sleep phase the prevalence is estimated as high as 7% for delayed sleep phase type.

5.24.2.3.5 The parasomnias

It is estimated that 10–50% of children aged three to five years have nightmares severe enough to alarm their parents. Since nightmares are such a prevalent experience (at least on a periodic or sporadic basis), it is difficult to estimate how many of these children warrant a diagnosis of nightmare disorder. The prevalence of nightmares in older children has not been established although there are indications that the phenomenon decreases significantly as children grow beyond the preschool years.

The prevalence of night terror disorder is estimated between 1% and 6% in children (Beltramini & Hertzig, 1983; Gass & Strauch, 1984; Kales et al., 1980a; Klackenberg, 1987). It is also estimated that the prevalence of this disorder decreases with age.

Sleepwalking appears to be a prevalent phenomenon during childhood. One study reported that 15% of 5–12-year-olds have had at least one episode. An incidence of 10% for at least two episodes during a year period prior to the study was reported for 3–11-year-old Swiss children (Gass & Strauch, 1984). The peak age for sleepwalking in this study was 10. In a longitudinal study of 212 children studied between six and 16 years of age, the incidence of sleepwalking was 40%. But when the frequency of the episodes was examined, only 2–3% of the children had experienced at least one episode a month. The prevalence rates peaked at ages 11–12 years (Klackenberg, 1987).

5.24.2.4 Etiology

Sleep as a biobehavioral phenomenon is affected by multiple maturational, physiological, environmental and psychosocial factors. Different sleep disorders vary in their loading on medical and psychosocial factors. Primary insomnia is by definition the disorder that is most likely to originate or be maintained by psychosocial and environmental factors. Primary hypersomia and BRSD are more heavily influenced by medical factors. On the other hand, CRSD and the parasomnias are loaded by a complex matrix of physiological and psychosocial factors that are difficult to untangle.

5.24.2.4.1 Primary insomnia

The timing, duration, and structure of sleep change dramatically during the first year of life. On average, newborns spend close to 18 h in sleep each day. Their sleep is multiphasic and consists of five to six sleep episodes spread throughout the daytime and night-time hours. Significant individual differences in sleep duration exist already during the first few days of life (Sadeh, Dark, & Vohr, 1996). While some newborns sleep close to 22 h a day, others spend only 7–8 h in sleep. In a rapid maturational process during the first year of life, most infants significantly reduce sleep duration and their sleep progressively becomes more consolidated and concentrated during the night. These developmental changes are crucial for the understanding of early childhood insomnia because most of the problems seen in infants and young children could be viewed as a developmental delay. Most infants develop consolidated sleep, sometimes described as "sleeping through the night," toward the end of the first year. This significant shift is very sensitive and can be complicated and compromised by many factors.

Another significant process that occurs during the first year of life is the transition from the parents as the major soothing agents of the infant to a growing self-regulatory self-soothing capacity. This transition is manifested in the parents decreased involvement around bedtime and the child's growing ability to fall asleep without parental presence. Close relationships exist between parental involvement in the soothing processes and the child's difficulty falling asleep and night-waking problems (Adair et al., 1991; Johnson, 1991; Van Tassel, 1985). Many intervention studies show significant improvement or resolution of children's night-waking and difficulty falling asleep when parents are encouraged to withdraw from their active involvement around bedtime (see Section 5.24.4). These findings suggest that the etiology of many of these problems is heavily influenced by parent–child bedtime interactions that either reward the child for activities that are incompatible with self-soothing and consolidated sleep patterns or prevent the pressure to develop more appropriate ones.

Other psychosocial factors such as parental and in particular, maternal psychopathology have also been associated with increased likelihood of sleep problems in the child (e.g., Benoit, Zeanah, Boucher, & Minde, 1992; Guedeney & Kreisler, 1987; Richman, 1981; Scott & Richards, 1990; Seifer, Sameroff, Dickstein, Hayden & Schiller, 1996; Van Tassel, 1985; Zuckerman et al., 1987). In general this literature suggests that maternal depression, sense of incompetence, separation difficulties, and insecure attachment may contribute to or exacerbate sleep problems in early childhood. It has been suggested that these parental attitudes, personality and psychodynamic factors affect sleep through direct bedtime interactions (Sadeh & Anders, 1993).

Trauma, stress, and anxiety are the underlying cause of many forms of primary insomnia (Sadeh, 1996). It has been documented in numerous studies on stress and trauma in children that sleep difficulties are among the most common nonspecific responses to such conditions (see Sadeh, 1996, for a review). In many cases it is possible to link the onset of the disorder, as an immediate or delayed response, to stressful or anxiety provoking events. Often it is difficult to trace back the association and a disorder that originated from stress continues to be maintained by other intervening factors.

In older children, anxieties and fears are probably the most common cause of primary insomnia but the theme of these anxieties may change according to phase-specific issues. For instance, a 3-year-old child may develop nighttime fears and sleep onset difficulties in response to a transition to a new daycare program, while an 11-year-old child may develop anxieties and sleep onset difficulties in response to a growing sexual awareness and anxieties related to masturbation that are likely to rise when the child is alone in bed.

From a different angle, primary insomnia in children has been associated with the child's personality and psychopathology. Close links between sleep problems and "difficult temperament" and other behavior problems have been documented (see Section 5.24.1). The causal links are difficult to interpret but undoubtedly children with behavior or temperament problems, whatever the source of these difficulties, are more likely to develop sleep disorders. A similar picture exists in adolescence. Sleep problems have been associated with depression, anxiety, and related negative self-perceptions (Dahl, 1996; Kirmil-Gray et al., 1984; Marks & Monroe, 1976; Price et al., 1978).

Beyond the psychosocial and developmental factor in the etiology of primary insomnia, it is important to note that some common physiological factors may underlie the clinical picture. For instance, a number of medical and physiological factors may be involved in the etiology of night-waking problems in infants and young children. Among those identified factors are milk intolerance (Kahn et al., 1989), esophageal reflux (Kahn et al., 1991), atopic eczema (Reid & Lewis-Jones, 1995), and colic.

5.24.2.4.2 Primary hypersomnia and narcolepsy

Little is known about this cluster of relatively rare disturbances. The etiology of these disorders is usually attributed to neurological and hormonal processes. Genetic and hormonal influences have been identified in disorders such as Kleine–Levin and narcolepsy. Some of these processes have been associated with pubertal changes, thus explaining their onset and other phenomena such as the links, in some cases, between the clinical manifestations and the menstrual cycle. The origin of these disorders is organic in nature although most of the underlying neurophysiological mechanisms are still unknown. Insufficient or dysregulated sleep tends to exacerbate the clinical manifestations of these disorders.

5.24.2.4.3 Breathing-related sleep disorder

The most common cause for BRSD in children is airway obstruction due to anatomical and neurological complications. Enlarged tonsils and adenoids are among the prevalent reasons for upper airway blockage. Central sleep apnea is caused by CNS failure to control the breathing process. Full understanding of the various mechanisms involved in this family of disorders is still lacking, however.

5.24.2.4.4 *Circadian rhythm sleep disorder*

The understanding of the etiology of CRSD has significantly increased since the mid-1970s. Some of the mechanisms involved in mediating between the light–dark cycle and the sleep–wake rhythm have been identified and systematically studied. This new knowledge includes the specific role of the hormone melatonin, which is secreted from the pineal body during darkness, in transferring the external light–dark information into other brain centers controlling the sleep–wake cycle. It is now more clearly understood that the 24-h sleep–wake rhythm requires constant light–dark information and when such information is blocked (e.g., by experimental measures or blindness), the sleep–wake cycle loses its stable organization. This explains why reports on specific CRSD disorders, such as free-running or disorganized sleep–wake rhythms, have mostly been associated with blind adults and children, and more specifically with their different melatonin secretion patterns (e.g., Lewy & Newsome, 1983; Tzischinsky et al., 1991). Another group that has been associated with CRSD is children who suffer from "brain-damage" or psychoneurological, developmental disorders. It is assumed that these children either have dysfunctional neurological systems that fail to control the circadian clock or have difficulties interpreting the environmental cues that are essential for controlling the circadian system.

Delayed sleep phase has mostly been associated with adolescence as adolescents seem to have a natural tendency to delay sleep onset and insufficient sleep (Carskadon & Dement, 1987). It has been suggested that this delay tendency is not only due to the psychosocial influences of adolescence but also associated with pubertal changes (Carskadon, Viera, & Acebo, 1993).

5.24.2.4.5 *The parasomnias*

Little is known about the etiology of the parasomnias. A number of constitutional or predisposing genetic, developmental, and psychological factors have been suggested as well as factors associated with sleep structure. Anxiety, stress, and insufficient sleep are among the factors that may contribute to or exacerbate these phenomena.

5.24.2.5 Developmental Course

The issue of the persistency and continuity of sleep problems in early childhood is important since the common belief shared by many professionals and parents is that children outgrow these problems and therefore early interventions are not critical. Despite the fact that maturational processes seem to resolve the sleep difficulties in many young children, a number of studies have shown that these problems tend to be very persistent (Kateria, Swanson, & Trevarthin, 1987; Richman et al., 1982; Salzarulo & Chevalier, 1983; Zukerman et al., 1987). For instance, when infants identified as sleep-disturbed at 8 months of age were revisited when they were three years old, 41% of them still suffered from sleep problems (Zuckerman et al.). These studies indicate that many sleep-disturbed children continue to suffer from sleep problems for prolonged periods although the characteristics of their problem may change significantly. Some retrospective studies with adult insomnia patients report that a large proportion of these patients suffered from sleep problems from early childhood, suggesting that these problems are likely to turn into chronic disorders if not treated (Hauri & Olmstead, 1980).

The course of hypersomnia, BRSD, and the parasomnias is very specific to the source of the disorder and the interventions involved. As described earlier, some disorders (the Kleine–Levin syndrome, narcolepsy) tend to appear (or to be diagnosed) in adolescence. In contrast, the parasomnias tend to appear in early childhood and disappear as the child matures, usually prior to or during adolescence.

5.24.3 CONCEPTUALIZATION AND CLINICAL FORMULATION

The assessment and diagnosis of sleep disorders in children should be based on two components: (i) the assessment of sleep schedule, sleep duration, and sleep quality; and (ii) the assessment of daytime functioning and signs of fatigue, sleepiness, and other arousal disorders.

One unique feature of sleep disorders that separates it from many other psychosocial and biobehavioral disorders is that in addition to subjective or parental reports, sleep can be directly measured and quantified by objective measures that have been established by extensive research efforts. However, in clinical practice and clinical research the assessment of common sleep disorders of childhood is still mostly based on subjective reports. This section reviews some of the available objective measurement tools for clinical assessment of sleep disorders. Some of these assessment tools (such as polysomnographic (PSG) studies) are often necessary to identify organic or medical sources for the disorders and are therefore crucial for correct diagnosis and treatment. In many cases, it is also crucial to exclude major psychiatric or developmental disorders when diagnosing sleep disorders.

5.24.3.1 Multimodal Assessment

Since sleep–wake regulation plays such a major role in the arena of emotional regulation and psychopathology (Dahl, 1996), it is imperative for any clinical diagnostic procedure within the context of medical or behavioral disorders to include screening questions regarding the possible existence of sleep or alertness problems. This screening investigation should include questions regarding sleep schedule, sleep quality, known or suspected sleep disorders, and questions regarding daytime sleepiness and fatigue (Dahl & Carskadon, 1995; Ferber, 1995; Sadeh & Anders, 1993).

The assessment of sleep disorders involves a systematic intake interview and, in many cases, medical examination and laboratory or ambulatory monitoring of sleep–wake patterns. In addition, when excessive sleepiness is suspected, it is important to objectively evaluate the level of sleepiness of the child.

5.24.3.1.1 Initial information

Parents or other caregivers are generally the best source of information regarding the child's present sleep problem, its history and etiology, and the associated psychosocial factors. If the child is old enough they may be able to describe the difficulties, particularly when emotional and psychosocial factors play an important role. The preliminary information for screening and evaluating sleep disorders includes the child's sleep patterns (schedule, quantity, and quality), sleep history, bedtime rituals, and other psychosocial factors. In many cases, it is recommended that the parents and/or the child be asked to complete a daily sleep log for one week to provide more detailed information regarding specific issues that are of interest to the clinician and to serve as a baseline assessment for later treatment progress evaluation. It is also important to understand what kind of interventions or efforts have been made already to solve the problem (i.e., self- or professionally-driven attempts) (Dahl & Carskadon, 1995; Ferber, 1995; Sadeh & Anders, 1993).

In assessing primary insomnia, primary hypersomnia, and the parasomnias in infants and children it is particularly important to evaluate the psychosocial, individual, and family contexts. The assessment should include efforts to identify specific stressors in the child and family's life, given the important role that stress plays in these disorders (Sadeh, 1996).

When there is a suspected or potential physical or medical source for the disorder, a medical examination is required as part of the initial assessment.

5.24.3.1.2 Assessment of sleep

Clinicians interested in sleep disorders have the advantage of being able to measure and quantify the clinical problem objectively in addition to receiving subjective reports of the child or the parents. With infants and children the parents are usually the reporters; however, their ability to report is limited. Parents are mostly aware of their child's sleep problem if the child signals (cries and requires intervention) during the night or, in the case of older children, if they communicate their problem directly to their parents (Sadeh, 1996; Sadeh & Anders, 1993). It has been shown that parents are relatively good reporters of their child sleep–wake schedule, but when it comes to sleep quality (i.e., number or duration of night-wakings), their reports are less accurate and might be adversely affected by various motivational and situational factors (Sadeh, 1996; Sadeh et al., 1991).

The traditional objective and comprehensive method to assess sleep is PSG (e.g. Anders, Emde, & Parmelee, 1971). This method is usually based on a laboratory study (although ambulatory studies have started to gain in popularity in recent years) for a single night. PSG provides information regarding the EEG, the electro-oculogram (eye movements), the electromyogram (muscle tone), and respiratory patterns during sleep. These data enable assessment of sleep structure, sleep stages, and specific disturbances during sleep such as those related to disordered breathing.

A number of ambulatory methods have been developed to assess sleep in the child's natural home environment. These methods include direct observations (Thoman, 1975), time-lapse video recording (Anders & Sostek, 1976), pressure-sensitive mattress recording of motility and respiration (Thoman & Whitney, 1989), and actigraphy–activity monitoring with a wrist-watch like device that has already been adopted for clinical practice (Sadeh, 1994; Sadeh et al., 1991, 1995a). See Table 2 for a description of the advantages and shortcomings of each method.

Table 2 Assessment tools for diagnosis and treatment follow-up in sleep disorders.

Method	*Description*	*Advantages*	*Limitations*
Polysomnography	Laboratory or ambulatory multiple-channel physiologic recording that requires wiring electrodes, post-testing interpretation (Anders et al., 1971)	Provides the most elaborate data. Crucial for diagnosing organic causes for sleep disorders	Expensive, disrupts natural sleep settings, and causes stress. Usually limited to one night
Videosomnography	Time-lapse video enables the monitoring of child's sleep patterns at home (Anders & Sostek, 1976)	Provides information about sleep onset, sleep–wake states, night-wakings, and parental interventions	Requires transportation to the home, and a postmonitoring observer scoring
Sensitive mattress	A system based on a pressure sensitive mattress installed in the child's bed (e.g., Thoman & Whitney, 1989)	Provides useful information about sleep–wake states and respiratory patterns	Requires home installation and a complex postmonitoring analysis
Direct observation	This method is based on trained observers who observe the infant and record his or her sleep–wake states on an ongoing basis. In addition a sensor pad and a strip chart recorder provide respiration data that facilitate state differentiation (e.g., Thoman, 1975)	Provides valuable information on sleep–wake and to some extent on sleep states in infants ("quiet" vs. "active" sleep). Provides information on respiration patterns	May disrupt normal home-setting. Very demanding task for the observers and requires complex postmonitoring analysis. Has only been used with infants
Actigraphy	Actigraphy is based on a small computerized activity monitor attached to the child's wrist or ankle. The activity data is automatically translated to sleep–wake measures that have been validated against polysomnography (Sadeh et al., 1995a)	Enables 24 h continuous nonintrusive measurement of sleep–wake and activity patterns. Cost-effective. Provides information on sleep schedule, night-wakings, and sleep quality	Some artifact may result from the fact that the only channel measured is activity. Sleep disturbances seen in the activity data are difficult to interpret without complementary information
Sleep diary	Parents or child complete a sleep diary that includes bedtimes, sleep onset time, number and timing of night-wakings and other sleep-related events and disruptions (e.g., Douglas & Richman, 1984)	Easy and inexpensive to administer. Provides additional information regarding relevant events associated with sleep (e.g., how the parents intervene)	Limited by the subjective knowledge of the parents or the child. May be biased by a variety of motivational issues. Parents and children may lose interest with prolonged administration
Sleep questionnaires	Parents or child complete a general sleep questionnaire characterizing their sleep problems (e.g., "child wakes up regularly every night")	Easy to administer. Provides general information on sleep difficulties	The subjective components may play a major role since the questions are not anchored to specific nights as in the daily logs

5.24.3.1.3 Assessment of excessive sleepiness

As described earlier, fatigue and excessive sleepiness may have different manifestations in children. The primary sign of sleepiness is a tendency or a growing pressure to sleep. Other signs in infants and children include irritability, fussiness, hyperactivity, affective dysregulation, difficulties in attention, and concentration (Dahl, 1996). Many of these signs may be observed directly in the school or the clinical setting or reported by the parents.

The multiple sleep latency test (MSLT) was designed to measure and quantify the pressure to fall asleep in children and adults (Carskadon & Dement, 1982). The basic procedure is to ask the subject to lie in bed with eyes closed and to try to fall asleep. Individuals who fall asleep rapidly (e.g., within 5 minutes) are considered very sleepy. Higher sleep latencies are interpreted as increased alertness. This procedure is repeated a few times during daytime hours. Norms and clinical criteria have been established.

5.24.3.2 Treatment Formulation

Treatment strategy is heavily based on the diagnosis and the clinical interpretation of applicable intervention avenues. For most clinical sleep disorders, pharmacological treatments are available, although the pervasive belief is that whenever possible the use of sleep-related drugs should be avoided, particularly with children. Potent psychosocial treatments are available for many sleep disturbances and they should be considered the treatment of choice.

Although Section 5.24.5 presents psychosocial and medical–pharmacological treatments separately, it is important to remember that this is not the recommended approach, rather, an integrated approach is usually called for. Even when medical treatment is required, motivational and psychological factors of the child and the parents play a major role in their response to treatment and therefore should be evaluated and addressed in the course of treatment.

5.24.4 TREATMENT

This section discusses the different approaches to treatment in children's and adolescents' sleep disorders. Due to the heterogeneous nature and the wide variety of sleep disorders, there are multiple psychosocial and medical therapeutic approaches. The following section reflects an attempt to review this wide range of interventions with a special focus on the psychosocial approaches.

5.24.4.1 Psychosocial Treatments

The psychosocial treatments of sleep disorders in infants and children constitute a relatively well-documented domain of multiple behavioral and psychodynamic approaches. This is particularly true for infants and young children and less so for older children and adolescents. Most of the emphasis is on primary insomnia which is the disorder that responds best to these interventions.

5.24.4.1.1 Primary insomnia

The recommended psychological treatment should include both behavioral components aimed at changing the sleep habits of the child and a psychodynamic component enabling the working-through of psychodynamic issues related to sleep difficulties (whenever such issues exist). The main issues in such a brief and focused psychological treatment are usually related to separation, soothing capacities, and limit setting. Therefore, in addition to the benefits of solving the specific sleep disorder, the brief psychological treatment might serve as a vehicle for the parents and child to acquire skills and competencies that may be useful in broader contexts (e.g., self-soothing, limit-setting).

The central hypothesis that underlies most behavioral interventions in early childhood is that sleep problems are the result of habitual, learned parent–child interactional patterns (Sadeh & Anders, 1993). Sadeh & Anders (1993) suggest that one of the key factors in the evolution of sleep problems in childhood is related to soothing skills that are acquired by children during development. The lack of these self-soothing skills is manifested in the case of a 18-month-old boy who cannot fall asleep without being rocked or breast-fed by his mother or by the 10-year-old girl who refuses to sleep in her own bed separated from her parents because of fears associated with sleep, dreaming, or darkness. An additional assumption is that sleep difficulties may be associated with sleep-competitive activities, that have been maintained by a variety of positive and negative reinforcers (Bootzin, 1977; Bootzin & Nicassio, 1978), such as parental reactions that perpetuate the sleep difficulties (Blampied & France, 1993). Hence, the target of the different modes of the behavioral treatment is to break these associations by teaching the child to fall asleep and to maintain sleep with minimal interaction with the parents.

Since the capacity for "sleeping through the night" is acquired by most infants by six months of age, most parents and clinicians prefer to wait for the maturational process to occur during this period. Despite this general attitude, it has been shown that an improvement in sleep patterns may be achieved by behavioral intervention with breast-fed infants as young as eight weeks old (Pinilla & Birch, 1993), or even by prevention programs based on educating expecting parents (in delivery classes) how to help their children develop healthy and mature sleep habits (Wolfson, Lacks, & Futterman, 1992).

According to the "behavioral extinction" approach, parents are instructed to stop attending to the child during the night. They are instructed to ignore the child's crying and protests until the child falls asleep due to exhaustion and lack of reinforcers (Sadeh & Anders, 1993). However, the extinction approach is difficult to implement for most parents since there is usually a postextinction response burst, an increase in the frequency and intensity of the behavior that is likely to accompany the introduction of extinction (Lawton, France, & Blampied, 1991; Rickert & Johnson, 1988). This treatment is difficult for many parents who experience such behavior as abandonment on their part. Therefore, more gradual and less aversive variants of the extinction technique have been recommended.

In the gradual extinction approach, the parental response to the child during falling asleep or night-waking are suspended for constant (Douglas & Richman, 1984) or increasing intervals of time (Ferber, 1995), and parents are advised to decrease their proximity to and involvement with the child during the night. For example, a parent may initially stop rocking the child to sleep, and instead remain close to the child's bed for a while until the child falls asleep. In the next step, the parent would leave the child's bedroom once the child has been put to bed (Sadeh & Anders, 1993).

Another alternative aimed at to make the extinction less aversive for both child and parents is using a combination treatment of sedative medication for the child and extinction (France, Blampied, & Wilkinson, 1991).

The stimulus control therapy attempts to break the association between the sleep environment and wakefulness by teaching the child not to engage in sleep "incompatible" behavior in the bedroom, and to establish the bed and bedroom as cues for sleep (e.g., to be in the bedroom only when drowsy or asleep) (Bootzin & Perlis, 1992).

Another way to break the rewarding association among night-waking, crying, and parental intervention is that of "scheduled awakenings." According to this approach the parents awaken the child prior to the time of the expected spontaneous awakening (McGarr & Hovel, 1980).

Parental cosleeping with the child has been used as an intervention method. According to this method, one parent sleeps in the child's bedroom for a one-week period without having any other interactions with the child throughout the night. This method is based on the assumption that in many cases infant sleep problems are related to the child's separation anxiety. The presence of the parent in the child's room is expected to reassure the child and enable consolidated sleep (Sadeh, 1994).

The methods described above are usually implemented within a standard clinical setup including clinical sessions with a therapist. However, some programs have used only written information and guidelines to the parents as the sole intervention. Written manuals containing instructions for parents without therapist support were found as highly effective in reducing waking during the night as the "conventional" format. The improvement was more gradual than the improvement in the sleep program with therapist involvement, but there was no difference in the final achievements of both groups (Seymour, Brock, During, & Pole, 1989). Adoption of behavioral principles into a group therapy format was not found to be effective (Messer, Lauder, & Humphrey, 1994).

Many studies demonstrate the efficacy of behavioral interventions in treating sleep problems, particularly during early childhood. Generally, the reported success rate range has been between 73% and 90% (e.g., Adair, Zuckerman, Bauchner, Philipp, & Levenson, 1992; Jones & Verduyn, 1983; Richman, Douglas, Hunt, Landsdown, & Levere, 1985; Roberts, 1993; Sadeh, 1994). Moreover, it has been found that a brief behavioral intervention program focused on helping families manage children with sleep disturbances could be generalized to improving daytime mother–child interactions (Minde, Faucon, & Falkner, 1994).

High success rates have been reported for a variety of extinction strategies (e.g., Adair et al., 1992; Adams & Rickert, 1989; Durand & Mindell, 1990; France & Hudson, 1990; Jones & Verduyn, 1983; Pritchard & Appelton, 1988; Richman et al., 1985; Rickert & Johnson, 1988; Sadeh, 1994).

Most of the reports have documented dramatic improvement within a one-week period. The use of extinction is reportedly accompanied by reduction in parental anxiety, and improvement in parental self-esteem, depression, marital satisfaction, and sense of control following extinction-based treatments of sleep disturbance. In addition, an improvement in infant sense of security has also been documented (France et al., 1991). These findings could be used to reassure parents who fear that such a strategy may harm their child. The stimulus control therapy was found to be effective, though its efficacy was measured in combination with other therapeutic components (Weissbluth, 1982). The scheduled awakening strategy proved difficult for parents to follow though it was found to be effective in reducing the number of night-wakings (e.g., McGarr & Hovel, 1980; Ricket & Johnson, 1988; Weissbluth, 1982).

All of the above reported efficacy studies are based on parental reports with their obvious limitations with one exception that has used both parental reports and actigraphic monitoring (Sadeh, 1994). This study documented two processes that occur during these interventions: (i) the infant's sleep becomes more consolidated and less fragmented; and (ii) the children acquire the capacity to soothe themselves back to sleep without parental interventions.

Psychodynamic approaches have also been implemented in treating early childhood primary insomnia. Psychodynamically oriented approaches have been used in order to understand and interpret the meaning of the child's sleep behavior in the context of the child–parent relationship (Daws, 1989). Since many sleep problems in children reflect parent–child interactions, parental anxieties, and psychopathology, these issues become the focus of psychodynamic treatment. The efficacy of the psychodynamic approach is yet to be established empirically, however, we believe that family and psychodynamic issues should be addressed in most cases to facilitate the efficacy of behavioral interventions.

As the child matures, the characteristics of the insomnia progressively resemble those seen at older ages and the insomniac adolescent gradually resembles the insomniac young adult and similar approaches found to be effective with adults can be adopted (Ferber, 1995). These methods include improving sleep hygiene, normalizing sleep schedules, stimulus control, sleep restriction training, biofeedback, cognitive therapy, and dynamic psychotherapy (Ferber, 1995; Morin et al., 1994).

Since little research has been conducted with adolescent subjects, efficacy of the different approaches is based on studies with adult subjects. The findings from different studies and reviews indicate that nonpharmacological interventions produce reliable and durable changes in sleep patterns of patients with chronic insomnia (e.g., Bootzin & Perlis, 1992; Morin et al., 1994). For instance, Morin et al. conducted a meta-analysis of nonpharmacological interventions for insomnia to examine the efficacy of the different treatment approaches. They found that psychological intervention produced a reliable change in sleep latency and in time awake after sleep onset.

5.24.4.1.2 *Primary hypersomnia and narcolepsy*

In addition to the pharmacological treatments, psychosocial interventions in hypersomnia and narcolepsy focus on two main issues: (i) regulating the sleep–wake schedule and extending sleep to reduce sleepiness (e.g., scheduling daily naps, maintaining strict and regular sleep–wake schedule; and (ii) consulting the child, the parents, and the school system in an effort to optimize the child's social and school adjustment.

Daytime sleep schedules have been found to be effective in reducing somnolence during the day (Billiard, 1976; Roehrs et al., 1986; Rogers & Aldrich, 1993) and in improving cognitive functioning (Mullington & Broughton, 1993). In addition to the difficulties to function due to the excessive sleepiness, narcoleptic children may suffer adjustment problems due to the unique form of their syndrome. Adult narcoleptic patients have been described as depressed and prone to adjustment disorders, anxiety, personality disorders, and alcohol dependency. There are also some indications that most adult narcoleptic patients benefit from psychotherapy and support groups (Alaia, 1992; Zarcone, 1973). The efficacy of such interventions with children and adolescents has not been established, however.

5.24.4.1.3 *Breathing-related sleep disorder*

The treatment of these disorders is primarily medical. Psychological work with the child and family is often focused on compliance with some of the medical interventions.

5.24.4.1.4 *Circadian rhythm sleep disorder*

Readjustment of the circadian system requires small, gradual, and consistent changes in bedtime and wake-up time, and requires one to two weeks of stability to realign all the components of the system (Dahl, 1992). This gradual change could be achieved by the combined effect of chronotherapy and light therapy.

The traditionally accepted treatment for delayed sleep phase is chronotherapy (Czeisler et al., 1981). Chronotherapy is a behavioral technique in which bedtime is systematically delayed or advanced, following the child's natural biological tendency (Ferber, 1995). Czeisler et al. reported that chronotherapy was effective in advancing sleep onset and wake-times with no reduction in sleep efficiency in five adult patients. Research focused on treatment in infants and adolescents has yet to be conducted.

In dealing with the adolescent age group, treatment must include behavioral contracts with specifically delineated rules, rewards, consequences, and a measure of monitoring components (Dahl, 1992). The behavioral contract should specify the time of going to sleep and awakening, avoidance of napping during the day, and the amount and timing of exposure to bright light. Parents and child should agree upon the reinforcements for keeping the program. The reinforcements should be given on a weekly basis during follow-up meetings with the therapist. The importance of keeping to the new schedule as strictly as possible, including holidays and weekends, must be emphasized and explained to the child and the parents. It is only after the schedule has been stabilized that there may be a place to negotiate "exceptions" for special occasions (e.g., trips, parties).

The treatment of delayed sleep phase syndrome was found to be effective in adolescents who were cooperative and were not suffering from depression, but was not effective in adolescents who expressed a refusal to go to school (Ferber & Boyle, 1983). In other studies, it was demonstrated that patients who had delayed phase syndrome and significant psychopathology (e.g., depression) responded poorly to treatment (Billiard, Verge, Touchon, Carlander, & Besset, 1993; Lahmeyer & Lilie, 1987). Therefore, it seems that whenever the circadian rhythm disorder is complicated by additional psychopathology, the success of treatment depends on the success of additional psychotherapy and psychiatric treatments.

When there are motivational factors that help perpetuate the problem (e.g., social or academic problems that the child is trying to avoid by staying away from school when tired, or special incentives for the child to stay up late), individual and family psychotherapy is recommended in addition to chronotherapy and light therapy.

5.24.4.1.5 Nightmares, sleep terror, sleepwalking, and other parasomnias

Two components should be considered in the approach to treating specific parasomnias. The first component is that of management aimed at providing protection and safety for the child. This is the most important consideration in coping with sleepwalking or night-terror episodes (Kales et al., 1980b; Kales, Soldatos, & Kales, 1987). The second component is additional symptomatic treatment in those cases where the arousals are dangerous or disruptive to the child or the family and no obvious precipitant can be identified (Rosen, Mahowald, & Ferber, 1995). When attacks are very frequent, if they persist beyond the age they would normally cease, if they start beyond that age (Stores, 1990), or if they follow a traumatic event (Rosen et al., 1995; Stores, 1990), protection is not enough and there might be a need for additional treatment.

Regarding the management of night terrors and sleepwalking, the recommendations are very practical. It is advisable to remove everything from the bedroom (and house) that could be hazardous. If possible, it is recommended that the bedroom be located on the ground floor of the house. It is important to lock the windows and cover the panes with heavy drapes (Keefauver & Guilleminault, 1995). In addition, it is recommended that sleep deprivation or other circumstances that might cause the child excessive sleep be avoided (Driver & Shapiro, 1993) and that the sleep–wake schedule be made regular (Rosen et al., 1995).

It is important to assure both the parents and the child that the problem is temporary, and that it is neither "abnormal" nor bearing significant consequences for the future (Mahowald & Rosen, 1990). It is also important to recommend that the parents do not discuss the episodes with the child the following day, since such discussion might be stressful for the child (who usually does not remember anything) and it may inadvertently reinforce the behavior. Furthermore, during the night, parents should resist trying to awaken and comfort the child. Waking the child may increase disturbance (Stores, 1990). The most appropriate response is to help the child to settle down when the acute episode subsides or if the child awakens when it ends. In cases in which night terrors occur at a regular and predicted time, it is possible to wake the child before each episode, keep them awake for few minutes, and thus break the cycle of night terror (Stores, 1990).

Another treatment approach is based on relaxation mental imagery. This form of treatment allows the child to utilize self-regulation techniques to control the previously uncontrolled night-time behaviors (Mahowald & Rosen, 1990). This method has been described as effective in treating parasomnias (Driver & Shapiro, 1993; Pesikoff & Davis, 1971).

When treating nightmares there is a special need to address potential sources for the fears (e.g., horror films) and to remove them (Dahl, 1992), simultaneously providing support and comfort to the child (Driver & Shapiro, 1993).

Psychotherapy is indicated with children and in families where significant psychopathology is present.

5.24.4.2 Pharmacological and Other Medical Treatments

Medical interventions play a significant role in treating many sleep disorders. While in some disorders, drugs and other medical interventions are the only available treatment, in other disorders they are too readily used when they should not be considered the treatment of choice.

5.24.4.2.1 Primary insomnia

Most experts in the field believe that medication is not a recommended treatment for insomnia in children. However, in practice high rates of medication use have been reported for infants and children.

The most common hypnotics that have been used for treating insomniac children are niaprazine (Nopron R) and trimeprazine tartrate (Vallergan Forte R) (Navelet, 1996). Sedatives such as chloral hydrate (Adler, 1990), benzodiazepines, and sedatives containing cough medicine have also been used by clinicians (Chavin & Tinson, 1980; Dahl, 1992; Werry & Carlielle, 1983).

Generally, it has been found that in the short-run sleep problems improve with the use of sedatives. However, most of the studies have not differentiated between different types of sleep disorders or ages of the children, and possess methodological deficiencies. For example, Russo, Vymut, and Gurauraj-Allen (1976) found dipheramine HCl to be effective in decreasing sleep latency time and reducing the number of awakenings in children aged 2–12 years who had a variety of sleep difficulties. Besana, Fiocchi, DeBartolomies, Magno, and Donati (1984) found niaprazine to be effective in the treatment of sleep disorders in children aged 2–13 years who had difficulties with onset and maintaining of sleep due to different emotional and medical conditions. However, both studies did not differentiate between different age groups or different sleep disorders. In a different clinical setting, Simonoff and Stores (1987) found tromerazine to be an effective treatment in reducing night-waking. However, their sample consisted of children with severe psychopathology and therefore does not represent the majority of infants with sleep difficulties.

Despite short-term improvement in insomnia as the result of sedative use, studies have shown little or no long-term benefit (Chavin & Tinson, 1980; Jones & Verduyn, 1983; Richman, 1986; Richman et al., 1985; Seymour, 1987). In addition, side effects of using sedatives in children include tolerance effects, withdrawal effects, and the fact that children who are not able to self-comfort behaviorally will often fight sleep despite the sedatives. This could result in increased levels of activity and awareness of the child when he or she is medicated (Dahl, 1992). Moreover, the problem may recur when the medication is stopped.

The use of medication is usually not indicated for insomnia in adolescents unless it is to treat a specific disorder that is related to insomnia. Since the insomnias that seem most likely to occur in adolescents are primarily associated with either psychiatric problems or with a circadian rhythm disorder, it seems important to make a careful diagnosis, and, in cases of psychiatric disorders such as anxiety or depression, the medication should be appropriate in treating both the sleep problem and the affective problem (i.e., by sedatives that provide an anxiolytic or depressant effect the next day; see Nicholson, 1994). When choosing to use sedatives, it is advisable to prescribe relatively small doses that enable the preservation of normal sleep architecture both during ingestion and after withdrawal.

Sleep-inducing medications may have adverse effects such as impaired performance the following day and anterograde amnesia, rebound insomnia when treatment is stopped, dependence, and developed tolerance.

5.24.4.2.2 Primary hypersomnia and narcolepsy

There is little empirical data in the younger narcoleptic population to guide treatment decisions (Dahl, 1992). Hence, the treatment of narcolepsy and cataplexy in children reflects clinical experience with older adult patients (Lawrence & Billiard, 1995).

The goal of the pharmacological treatment is to control the narcoleptic symptoms and to allow the patient full participation in daily activities. Pharmacological treatment usually includes: REM-suppressant medication such as trycyclic antidepressants and clomipramine, desmethylimipramine, or other 5HT2 reuptake inhibitors aimed at dealing with cataplexy and hypnogogic hallucinations, and stimulant medication such as pemoline and methylphenidate, or amphetamines aimed at dealing with daytime sleepiness (Parkes, 1994). The effectiveness of these pharmacological treatments in increasing alertness of narcoleptic patients has been demonstrated in a number of studies (Guillerminault, 1993; Rogers & Aldrich, 1993). However, since narcolepsy is a long-lasting illness, it is important to take into account issues related to tolerance, addiction, and adverse side effects.

It is important to remember that despite medication, most adult narcoleptic patients experience narcoleptic symptoms, especially somnolence, which is very disruptive to functioning during the day (Scumacher, Merritt, & Cohen, 1993). Further research should be conducted to evaluate the efficacy of pharmacological treatments of narcolepsy in children.

5.24.4.2.3 Breathing-related sleep disorder

Optional medical treatments for obstructive sleep apnea syndrome (OSAS) include surgical, mechanical therapy, and medical–pharmacological treatment. The surgical approach includes adenotonsillectomy and tracheostomy that are often used in treating sleep apnea in children and adults. Adenotonsillectom, removing tonsils and adenoids, is a solution for enlarged tonsils and adenoids that have been found to cause obstruction (Stores, 1990; Stradling, Thomas, Warely, Williams, & Feeland, 1990). Tracheostomy is a surgical intervention aiming at enlarging the breathing passway itself by introducing a canula into it. Tracheostomy is indicated for children with severe and/or complicated OSAS in whom other treatment approaches are either not possible or unsuccessful.

Since there are potentially serious complications involved in these surgical procedures, they are recommended only in serious cases (Carrol & Louglin, 1995; Powell; Guilleminault, & Riley, 1994). There is little scientific information and no consensus or guidelines for surgical treatment of childhood OSAS (Carrol & Louglin, 1995). Hence, further research concerning efficacy of this treatment strategy is required.

Nasally applied continuous airway pressure (CPAP) has become the major nonsurgical, long-term form of treatment used in obstructive sleep apnea and it is currently the treatment of choice. The method is based on providing intermittent positive pressure through a nose mask to overcome obstructions and hypoventilation syndromes (Sullivan & Grunstein, 1994). When CPAP treatment is initiated there is an immediate reduction in daytime somnolence. This has been demonstrated using the multiple sleep latency test (Lamphere, Roehrs, & Wittig, 1989). Serious complications of CPAP appear to be rare. Nasal congestion is a side effect of CPAP therapy that can be pharmacologically or, in very rare cases, surgically treated.

CPAP is effective in reducing the severity of obstructive sleep apnea, but it does not lead to a permanent cure (Marcus et al., 1995). CPAP treatment for OSAS was found to be effective in more than 85% of the pediatric patients. Objective measures such as MSLT, before and after one year of treatment, have confirmed its efficacy in adults, indicating that after a year of CPAP treatment sleepiness as measured by the MSLT was significantly raised (Sforza & Lugaresi, 1995).

In contrast with the reported efficacy of the CPAP treatment, the use of medication to treat sleep apnea has been disappointing (Strollo & Rogers, 1996). Nasal decongestants and topical steroids can be used as a short-term solution in cases of OSAS in children.

5.24.4.2.4 Circadian rhythm sleep disorder

Since the architecture of sleep in CRSD has no abnormalities, and because the patients usually have no difficulties with sleep onset given the opportunity to sleep at a less commonly accepted time of day, there is no indication for drug therapy in this disorder.

A relatively new approach to the treatment of sleep problems is the use of melatonin for regulating sleep. Melatonin plays an important role in the induction of sleep in the human organism. The secretion of melatonin is controlled by an endogenous rhythm-generating system in the brain which is synchronized by the light-dark cycle, so that melatonin levels are high during dark hours and low during light hours.

Exogenous melatonin has been found useful in the treatment of sleep disorders in adults (Dahlitz et al., 1991; McFarlane, Cleghorn, Brown, & Streiner, 1991); however, some studies have shown that the results are not maintained when treatment is terminated (Alvarez et al., 1992; Dahlitz et al.). There are relatively few studies examining melatonin therapy for sleep disorder in children.

Melatonin treatment has been found to be effective in treating highly disabled children with severe sleep disorders that have resisted conventional treatment (Jan, Espezel, & Appelton, 1994). In addition, melatonin has been found to be effective in treating blind retarded children with a problematic sleep–wake cycle (Lapierre & Dumont, 1995). However, concerns regarding the safety of using melatonin in prepubertal children have been raised due to the links between melatonin and sexual maturation.

Since the circadian cycle, and hence the sleep–wake cycle, may be shortened or lengthened by exposure to bright light (Regestein & Monk, 1995), early morning exposure to bright light and activity can be an important adjunct to maintaining an early schedule (Dahl, 1992). Exposure to bright light has been shown to be effective in treating circadian rhythm disorders in adult patients (e.g., Lack & Wright, 1993; Lewy, Singer, & Sack, 1985). Guilleminault, McCann, Quera-Salva, and Cetel (1993) reported a positive response in five of 14 mentally retarded and neurologically damaged children with severe sleep disorders.

5.24.4.2.5 *Nightmares, sleep terror, sleepwalking, and other parasomnias*

Drug treatment may be considered if night terrors or sleepwalking occur very frequently, or if the problem interferes with the child participating in normal social interactions. The common drugs for treating parasomnias are diazepam, imipramine, benzodiazepines, and tricyclic antidepressants. For night terrors, both diazepam and imipramine have been shown to be effective. Diazepam is considered the drug of choice because it is associated with fewer side effects (McDaniel, 1987). Benzodiazepines and tricyclic antidepressants have been shown to be useful (Fisher, Kahn, Edwards, & Davis, 1973; Guilleminault, 1994; Kavey, Whyte, Resor, & Gidro-Frank, 1987). Low doses of a benzodiazepine may also be effective (Fisher et al., 1973; Pesikoff & Davis, 1971), but there is a high rate of relapse particularly during times of stress (Driver & Shapiro, 1993).

Combining pharmacological and psychological treatment may be effective: drugs could be used for treating acute current difficulties, and different therapeutic tools could be used for long-term treatment, thus avoiding tolerance and adverse side effects.

5.24.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

This chapter raises many areas of insufficient knowledge and research efforts. Information on diagnosis, prevalence, time course, and interventions is lacking for many specific sleep disorders. Many of the definitions of the disorders are too vague and rely heavily on subjective interpretations of the child, parents, and clinicians. Despite the fact that a number of objective assessment tools have been established, they have yet to be adopted for clinical research and practice.

Clinical sleep research has mostly focused on two age periods: infancy and adulthood. There is a significant gap in clinical knowledge on children during school years and the adolescent period. For instance, there are clinical assumptions that school children and adolescents can benefit from relaxation training and other behavioral methods that have demonstrated efficacy with adults; however, this efficacy has not been established for children.

Another area with limited information pertains to the maturational-developmental processes of the sleep–wake system beyond early childhood. Longitudinal research is needed to assess the continuity of sleep problems across the life span effects of early disturbances of the sleep–wake system on other behavioral and neuropsychological systems.

There is a growing body of research indicating that: (i) sleep disturbances are very common in early childhood; (ii) sleep problems in early childhood are persistent; and (iii) brief and focused interventions are effective in treating most of the common sleep problems. This knowledge should be translated to clinical practice by shifting the focus to prevention and early detection and intervention. The study of Wolfson et al. (1992) is a good example of a prevention program and its potential efficacy.

The strong links between sleep problems and psychopathology require special attention. Future experimental and clinical studies should focus on assessing the role of sleep disorders and sleep deprivation in the evolution of psychopathology in children.

We believe that in the near future the current trend of adopting a multidimensional approach in diagnosing and treating sleep disorders will become the prevailing standard of practice. This multidimensional approach will include: (i) objective and subjective assessment of sleep; (ii) assessment of psychosocial factors including psychopathology and stress; (iii) combining behavioral or medical treatment with psychosocial interventions aimed at psychodynamic and family issues associated with the disorder; and (iv) assessing the efficacy of interventions by subjective and objective measures.

5.24.6 SUMMARY

This chapter raises a number of clinical and empirical issues that are fundamental to the diagnosis and treatment of children's sleep disorders.

(i) Sleep disorders are very prevalent during infancy, childhood, and adolescence. Since many studies have not used standard criteria to define specific sleep disorders, and some disorders have been studied mostly in clinical settings, the prevalence of many specific disorders in the general population has not been well established.

(ii) Despite the fact that objective measures of specific disorders like night-waking problems have been established, they have not been incorporated into clinical research and practice which is still heavily based on subjective reports.

(iii) The diagnostic criteria for many specific disorders are still relatively vague with no empirical consensus. Therefore, in clinical research it is difficult to compare studies and in many cases to even identify the specific disorders under investigation.

(iv) Sleep disorders, daytime behavior, and psychopathology are highly inter-related. Screening for sleep problems should be included in any clinical assessment with children. Specific attention should be given to daytime sleepiness and paradoxical signs of fatigue.

(v) When signs of sleepiness and fatigue exist with no "simple" explanation, such as insufficient sleep, a full evaluation is needed, including identification of medical disorders such as breathing-related sleep disorders by means of a full polysomnographic study.

(vi) When a sleep disorder exists in the context of known psychopathology one should not assume that the psychopathology is the cause of the sleep problem. The sleep problem should be thoroughly assessed and treated appropriately. It may be the cause of the associated psychopathology.

(vii) Primary insomnia should be primarily treated by psychosocial interventions. The efficacy of short-term focused interventions has been strongly supported in young children and additional research is needed in the case of older children and adolescents.

(viii) Since stress and other psychosocial factors play an important role in most of the sleep disorders in children, it is important to assess these factors and intervene accordingly, even in those disorders which primarily require medical treatments.

(ix) Psychosocial and psychodynamic issues should always be addressed in the evaluation and treatment phases as underlying powerful factors in the evolution and maintenance of a child's sleep problem.

5.24.7 REFERENCES

Adair, R., Bauchner, H., Philipp, B., Levenson, S., & Zuckerman, B. (1991). Night waking during infancy: Role of parental presence at bedtime. *Pediatrics, 84,* 500–504.

Adair, R., Zuckerman, B., Bauchner, H., Philipp, B., & Levenson, S. (1992). Reducing night waking in infancy: A primary care intervention. *Pediatrics, 89,* 585–588.

Adams, L. A., & Rickert, V. I. (1989). Reducing bedtime tantrums: comparison between positive routines and graduated extinction. *Pediatrics, 84,* 756–761.

Adler, R. (1990). Common behavioral disturbances. In M. J. Robinson (Ed.), *Practical paediatrics* (2nd ed., pp. 543–549). Singapore: Longman.

Alaia, S. L. (1992). Life effects of narcolepsy: Measures of negative impact, social support and psychological well-being. In M. Goswami, C. P. Pollak, F. L. Cohen, M. J. Thropy, & N. B. Kavey (Ed.), *Psychosocial aspects of narcolepsy* (pp. 1–22). New York: Haworth Press

Alvarez, B., Dahlitz, M. J., Vignau, J., & Parkes, J. D. (1992). The delayed sleep phase syndrome: clinical and investigative findings in 14 subjects. *Journal of Neurology, Neurosurgery and Psychiatry, 55,* 665–670.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Anders, T. (1979). Night waking in infants during the first year of life. *Pediatrics, 63,* 860–864.

Anders, T. F., Carskadon, M. A., Dement, W. C., & Harvey, K. (1978). Sleep habits of children and the identification of pathologically sleepy children. *Child Psychiatry and Human Development, 9,* 56–62.

Anders, T. F., Emde, R., & Parmelee, A. A. (1971). *A manual of standardized terminology, techniques and criteria for the scoring of states of sleep and wakefulness in newborn infants.* Los Angeles: UCLA Brain Information Service.

Anders, T. F., & Sostek, A. M. (1976). The use of time-lapse video recording of sleep–wake behavior in human infants. *Psychophysiology, 13,* 155–158.

Babkoff, H., Caspy, T., Mikulinncer, M., & Sing, H. C. (1991). Monotonic and rhythmic influences: a challenge for sleep deprivation research. *Psychological Bulletin, 109,* 411–428.

Baird, W. P. (1987). *Narcolepsy: A non-medical presentation.* San Carlos, CA: American Narcolepsy Association.

Bearpark, H. M., & Michie, P. T. (1987). Prevalence of sleep/wake disturbances in Sydney adolescents. *Sleep Research, 16,* 304.

Beltramini, A., & Hertzig, M. (1983). Sleep and bedtime behavior in preschool-aged children. *Pediatrics, 71,* 153–158.

Benoit, D., Zeanah, C., Boucher, C., & Minde, K. (1992). Sleep disorders in early childhood: Association with insecure maternal attachment. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 86–93.

Besana, R., Fiocchi, A., DeBartolomies, L., Magno, F., & Donati, C. (1984). Comparison of Niaprazine and placebo in pediatric behaviour and sleep disorders: double-bind clinical trial. *Current Therapeutic Research, 36,* 58–66.

Billiard, M. (1976). Competition between the two types of sleep, and the recuperative function of REM sleep versus NREM sleep in narcoleptics. In C. Guilleminault, W. C. Dement, & P. Passouant (Eds.), *Narcolepsy* (pp. 77–96). New York: Spectrum.

Billiard, M., Verge, M., Touchon, J., Carlander, B., & Besset, A. (1993). Delayed sleep phase syndrome: subjective and objective data, chronotherapy and follow-up. *Sleep Research, 22,* 172.

Blampied, N. M., & France, K. G. (1993). A behavioral model of infant sleep disturbance. *Journal of Applied Behavior Analysis, 26,* 477–492.

Bonnet, M. H. (1985). Effect of sleep disruption on sleep, performance, and mood. *Sleep, 8,* 11–19.

Bootzin, R. R. (1977). Effects of self-control procedures for insomnia. In R. B. Stuart (Ed.), *Behavioral self-management: strategies, techniques and outcomes* (pp. 176–195). New York: Brunner/Mazel.

Bootzin R. R., & Nicassio, P. M. (1978). Behavioral treatment for insomnia. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification* (Vol. 6, pp. 1–45) New York: Academic Press.

Bootzin R. R., & Perlis M. E. (1992). Nonpharmacologic treatments of insomnia. *Journal of Clinical Psychiatry, 53,* 37–42.

Carey, W. (1974). Night waking and temperament in infancy. *Journal of Pediatrics, 84,* 756–758.

Carrol, J. L., & Louglin, G. M. (1995) Obstructive sleep apnea syndrome in infants and children: diagnosis and management. In R. Ferber & Kryger (Eds.), *Principles and practice of sleep medicine in the child* (pp. 193–216). Philadelphia: W. B. Saunders.

Carskadon, M. A., & Dement, W. C. (1982). The Multiple Sleep Latency Test: What does it measure? *Sleep, 5,* s67–s72.

Carskadon, M. A., & Dement, W. C. (1982). Sleepiness in the normal adolescent. In C. Guilleminault (Ed.), *Sleep and its disorders in children* (pp. 53–66). New York: Raven Press.

Carskadon, M. A., Viera, C., & Acebo, C. (1993). Association between puberty and delayed phase preference. *Sleep, 16,* 258–262.

Challamel, M. J., Mazzola, M. E., Nevsimalova, Cannard, C., Louis, J., & Revol, M. (1994). Narcolepsy in children. *Sleep, 17,* s17–s20.

Chavin, W., & Tinson, S. (1980). The developing child: Children with sleep difficulties. *Health Visitor, 53,* 477–480.

Czeisler, C. A., Richardson, G. S., Coleman, R. M., Zimmerman, J. C., Moore-Ede, M. C., Dement, W. C., & Weitzman, E. D., (1981). Chronotherapy: resetting the circadian clocks of patients with delayed sleep phase insomnia. *Sleep, 4,* 1–21.

Dahl, R. E. (1992). The pharmacologic treatment of sleep disorder. *Psychiatric Clinics of North America, 1,* 161–178.

Dahl, R. E. (1996). The regulation of sleep and arousal: development and psychopathology. *Development and Psychopathology, 8,* 3–27.

Dahl, R. E., & Carskadon, M. A. (1995). Sleep and its disorders in adolescence. In R. Ferber & M. Kryger (Eds.), *Principles and practice of sleep medicine in the child* (pp. 19–27). Philadelphia: W. B. Saunders.

Dahl, R. E., Pelham, W., & Wierson, M. (1991). The role of sleep disturbances in attention deficit disorder symptoms: a case study. *Journal of Pediatric Psychology, 16,* 229–239.

Dahlitz, M. J., Alvarez, B., Vignau, J., English, J., Arendt, J., & Parkes, J. D. (1991). Delayed sleep phase syndrome response to melatonin. *Lancet, 337,* 1121–1124.

Daws, D. (1989), *Through the night; Helping parents and sleepless infants,* London: Free Association Books.

Dement, W. C., Carskadon, M. A., & Ley, R. (1973). The prevalence of narcolepsy II. *Sleep Research, 2,* 147.

Diagnostic Classification Steering Committee (M. J. Thorpy, Chairman) (1990). *International classification of sleep disorders: diagnostic and coding manual* Rochester, MN: American Sleep Disorders Association.

Douglas, J., & Richman, N. (1984). *My child won't sleep.* Harmondsworth, UK: Penguin.

Driver, H. S., & Shapiro, C. M., (1993). ABC of sleep disorders: Parasomnias. *British Medical Journal, 306,* 921–924.

Durand, V. M., & Mindell, J. A. (1990). Behavioral treatment of multiple childhood sleep disorders. *Behavior Modification, 14,* 37–49.

Ferber, R. (1995). Sleeplessness in children. In R. Ferber & M. Kryger (Eds.), *Principles and practice of sleep medicine in the child* (pp. 79–89). Philadelphia: W. B. Saunders.

Ferber, R., & Boyle M. D. (1983). Delayed sleep phase syndrome versus motivated sleep phase delay in adolescents. *Sleep Research, 12,* 239.

Ferber, R., & Kryger, M. (Eds) (1995). *Principles and practice of sleep medicine in the child.* Philadelphia: W.B. Saunders.

Fisher, C., Kahn, E., Edwards, A., & Davis, D. M. (1973). The psychophysiological study of nightmares and night terrors. *Archives of General Psychiatry, 28,* 252–259.

France, K. G., & Hudson, S. M. (1990). Behavior management of infant sleep disturbance. *Journal of Applied Behavioral Analysis, 23,* 91–98.

France, K. G., Blampied, N. M., & Wilkinson, P. W. (1991). Treatment of infant sleep disturbance by trimeprazine in combination with extinction. *Developmental and Behavioral Pediatrics, 5,* 308–314.

Gass, E., & Strauch, I. (1984). The development of sleep behavior between 3 and 11 years. In *Proceedings of the 7th European Sleep Congress.* Munich.

Guedeney, A., & Kreisler, L. (1987). Sleep disorders in the first 18 months of life: Hypothesis on the role of mother–child emotional exchanges. *Infant Mental Health Journal, 8,* 307–318.

Guilleminault, C. (1987). Narcolepsy and its differential diagnosis. In C. Guilleminault (Ed.), *Sleep and its disorders in children* (pp. 181–193). New York: Raven Press.

Guilleminault, C. (1993). *Amphetamines and narcolepsy: use of the Stanford database.*

Guilleminault, C. (1994). *Narcolepsy syndrome.* In M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and practice of sleep medicine* (2nd ed., pp. 549–561). Philadelphia: W. B. Saunders.

Guilleminault, C., McCann, C. C., Quera-Salva, M., & Cetel, M. (1993). Light therapy as treatment of dyschronosis in brain impaired children. *European Journal of Pediatrics, 152,* 754–759.

Hauri, P., & Olmstead, E. (1980). Childhood-onset insomnia. *Sleep, 3,* 59–65.

Jan, G. E., Espezel, H., & Appleton, R. E. (1994). The treatment of sleep disorders with melatonin. *Developmental Medicine and Child Neurology, 36,* 97–107.

Jenkins, S., Bax, M. & Hart, H. (1980). Behavior problems in preschool children. *Journal of Child Psychology and Psychiatry, 21,* 5–17.

Johnson, M. (1991). Infant and toddler sleep: A telephone survey of parents in one community. *Journal of Developmental and Behavioral Pediatrics, 12,* 108–114.

Jones, D. P. H., & Verduyn, C. M. (1983). Behavioural management of sleep problems. *Archives of Diseases in Childhood, 58,* 442–444.

Kahn, A., Mozin, M., Rebuffat, E., Sottiaux, M., & Muller, M. F. (1989). Milk intolerance in children with persistent sleeplessness: A prospective double-blind crossover evaluation. *Pediatrics, 84,* 595–603.

Kahn, A., Rebuffat, E., Sottiaux, M., Dufour, D., Cadranel, S., & Reiterer, F. (1991). Arousals induced by proximal esophageal reflux in infants. *Sleep, 14,* 39–42.

Kales A., Soldatos C. R., Bixler, E. O., Ladda, R. L., Charney, D. S., Weber, G., & Schweitzer, P. K., (1980a). Hereditary factors in sleep walking and night terrors. *British Journal of Psychiatry, 137,* 111–118.

Kales A., Soldatos C. R., Caldwell, A. B., Kales, J. D., Humphrey, F. J., Charney I. L., & Schweitzer, P. K. (1980b). Somnambulism: clinical characteristics and personality patterns. *Archives of General Psychiatry, 37,* 1406–1410.

Kales, A., Soldatos, C. R., & Kales, J. D. (1987). Sleep disorders: insomnia, sleepwalking, night terrors, nightmares, and enuresis. *Annals of Internal Medicine, 106,* 582–592.

Kateria, S., Swanson, M., & Trevarthin, G. (1987).

Persistence of sleep disturbances in preschool children. *Journal of Pediatrics, 110,* 642–646.

Kavey, N. B., Whyte, J., Resor, S., & Gidro-Frank, S. (1987). Classification and treatment of somnambulism. *Sleep Research, 16,* 368.

Keefauver, S. P., & Guilleminault, C. (1995). Sleep terrors and sleepwalking. In R. Ferber, M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and practice of sleep medicine in the child* (pp. 567–573). Philadelphia: W. B. Saunders.

Kirmil-Gray, K., Eagleston, J. R., Gibson, E., & Thorensen, C. E. (1984). Sleep disturbances in adolescents: sleep quality, sleep habits, beliefs about sleep, and daytime functioning. *Journal of Youth and Adolescence, 13,* 375–384.

Klackenberg, G. (1987). Incidence of parasomnias in children in a general population. In C. Guilleminault (Ed.), *Sleep and its disorders in children* (pp. 99–113). New York: Raven Press.

Lack, L., & Wright, H. (1993). The effect of evening bright light in delaying the circadian rhythms and lengthening the sleep of early morning awakening insomniacs. *Sleep, 16,* 436–443.

Lahmeyer, H. W., & Lilie, J. K. (1987). Personality affects treatment outcome in phase delay sleep syndrome. *Sleep Research, 16,* 282.

Lamphere, J., Roehrs, T., & Wittig R., (1989). Recovery of alertness after CPAP in apnea, *Chest, 96,* 1364–1367.

Lapierre, O., & Dumont, M. (1995). Melatonin treatment of a non-24-hour sleep–wake cycle in a blind retarded child. *Biological Psychiatry, 38,* 119–122.

Lawrence, W. B., & Billiard, M. (1995). Narcolepsy, Kleine–Levin syndrome, and other causes of sleepiness in children. In R. Ferber & M. Kryger (Eds.), *Principles and practice of sleep medicine in the child* (pp. 125–134). Philadelphia: W. B. Saunders.

Lawton, C., France, K. G., & Blampied, N. M. (1991). Treatment of infant sleep disturbance by graduated extinction. *Child and Family Behavior Therapy, 13,* 39–56.

Lewy, A. J., Singer, C. M., & Sack, R. L. (1985). Treatment of appropriately phase typed sleep disorders using properly timed bright light. *Sleep Research, 14,* 304.

Lewy, A. J., & Newsome, D. A. (1983). Different types of melatonin circadian secretory rhythms in some blind subjects. *Journal of Clinical and Endocrinological Metabolism, 56,* 1103–1107.

Madge, P. J., Nisbet, L., McColl, J. H., Vallance, A., Paton, J. Y., & Beattie, J. O. (1995). Home nebuliser use in children with asthma in two Scottish Health Board Areas. *Scottish Medical Journal, 40,* 141–143.

Mahowald, M. W., & Rosen, G. M. (1990). Parasomnias in children. *Pediatrician, 17,* 21–31.

Marcus, C. L., Ward, S. L., Mallory, G. B., Rosen, C. L., Beckerman, R. C., Weese-Meyer, D. E., Brouillette, R. T., Trang, H. T., & Brooks, L. J. (1995). Use of continuous positive airway pressure as treatment of childhood constructive sleep apnea. *Journal of Pediatrics, 172,* 88–94.

Marks, P. A., & Monroe, J. L. (1976). Correlates of adolescent poor sleepers. *Journal of Abnormal Psychology,* 85, 243–246.

McDaniel, K. D. (1987). Pharmacologic treatment of psychiatric and neurodevelopmental disorders in children and adolescents: I, II, and III. *Annual Progress in Child Psychiatry and Child Development,* 462–493.

McFarlane, J. G., Cleghorn, J. M., Brown, G. M., & Streiner, D. L. (1991). The effects of exogenous melatonin on the total sleep time and daytime alertness of chronic insomniacs: a preliminary study. *Biological Psychiatry, 30,* 371–376.

McGarr, R. J., & Hovel, M. F. (1980). In search of the sandman: Shaping an infant to sleep. *Education and Treatment of Children, 3,* 173–182.

Messer, D. J., Lauder, L., & Humphery, S. (1994). The effectiveness of group therapy in treating children's sleeping problems. *Child: Care, Health and Development, 20,* 267–277.

Minde, K., Faucon, A., & Falkner, S. (1994), Sleep problems in toddlers: effects of treatment on their daytime behavior. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 1114–1121.

Monk, T. H. (1994). Circadian rhythms in subjective activation, mood and performance efficiency. In M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and practice of sleep medicine* (2nd ed., pp. 321–330). Philadelphia: W. B. Saunders.

Montangner, H., De Roquefeuil, G., & Djakovic, M. (1992). Biological, behavioral and intellectual activity rhythms of the child during its development in different education environments. In Y. Touitou & E. Haus (Eds.), *Biological rhythms in clinical and laboratory medicine* (pp. 214–229). Berlin: Springer-Verlag.

Moore, T., & Ucko, L. (1957). Night waking in early infancy, Part 1. *Archives of Disease in Childhood, 32,* 333–342.

Morin, C. M., Culbert, J. P., & Schwartz, M. (1994), Nonpharmacological interventions for insomnia: a meta-analysis of treatment efficacy. *American Journal of Psychiatry, 151,* 1172–1180.

Morrison, D. N., McGee, R., & Stanton, W. R. (1992). Sleep problems in adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 94–99.

Mullington, J., & Broughton, R. (1993). Scheduled naps in the management of daytime sleepiness in narcolepsy-cataplexy. *Sleep, 16,* 444–456.

Navelet, Y. (1996). Insomnia in the child and adolescent. *Sleep, 19,* S23–S28.

Navelet, Y., Anders, T. F., & Guilleminault, C. (1976). Narcolepsy in children. In C. Guilleminault, W. Dement, & P. Passouant (Eds.), *Narcolepsy* (pp. 171–177). New York: Spectrum.

Nicholson A. N., (1994). Hypnotics: clinical pharmacology and therapeutics. In M. H. Kryger., T. Roth., & W. C. Dement (Eds.), *Principles and practice of sleep medicine* (2nd ed., pp 355–363). Philadelphia: Saunders.

Okawa, M., Nanami, T., Wada, S., Shimizu, T., Hishikawa, Y., Sasaki, H., Nagamine, H., & Takahashi, K. (1987). Four congenitally blind children with circadian sleep–wake rhythm disorder. *Sleep, 10,* 101–110.

Paret, I. (1983). Night waking and its relationship to mother-infant interaction in nine-month-old infants. In J. Call, E. Galenson, & R. Tyson (Eds.), *Frontiers of infant psychiatry* (pp. 171–177). New York: Basic Books.

Parkes, D. (1994). Introduction to the mechanism of action of different treatments of narcolepsy. *Sleep, 17,* s93–s966.

Pesikoff, R. D., & Davis, P. C. (1971), Treatment of pavor nocturnus and somnambulism in children. *American Journal of Psychiatry, 128,* 134–137.

Pinilla, T., & Birch, L. L. (1993). Help me make it through the night: behavioral entrainment of breast-fed infants' sleep pattern. *Pediatrics, 91,* 436–444.

Powell, N. B., Guilleminault C., & Riley R. W. (1994). Surgical therapy for obstructive sleep apnea. In M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and practice of sleep medicine* (2nd ed., pp. 706–721). Philadelphia: W. B. Saunders.

Price, V. A., Coates, T. J., Thoresen, C. E., & Grindstead, O. A. (1978). Prevalence and correlates of poor sleep among adolescents. *American Journal of Disease in Children, 132,* 583–586.

Pritchard, A., & Appleton, P. (1988). Management of sleep problems in pre-school children. *Early Child Development and Care, 34,* 227–240.

Quine, L. (1992). Severity of sleep problems in children with severe learning difficulties: description and correlates. *Journal of Community & Applied Social Psychology, 2,* 247–268.

Regestein, Q. R., & Monk, T. H. (1995). Delayed sleep phase syndrome: a review of its clinical aspects. *American Journal of Psychiatry, 152,* 602–608.

Reid, P., & Lewis-Jones, M. S. (1995). Sleep difficulties and their management in preschoolers with atopic eczema.-*Clinical and Experimental Dermatology, 20,* 38–41.

Richman, N. (1981). A community survey of characteristics of one- to two-year-olds with sleep disruptions. *Journal of the American Academy of Child Psychiatry, 20,* 281–291.

Richman, N. (1986). Recent progress in understanding and treating sleep disorders. *Advances in Developmental and Behavioral Pediatrics, 7,* 45–63.

Richman, N., Douglas, J., Hunt, H., Landsdown, R., & Levere, R (1985). Behavioral methods in the treatment of sleep disorders—a pilot study. *Journal of Child Psychology and Psychiatry, 26,* 581–590.

Richman, N., Stevenson, J., & Graham, P. (1982). *Preschool to school: A behavioral study*. London: Academic Press.

Rickert, V. I., & Johnson, C. M. (1988). Reducing nocturnal awakening and crying episodes in infants and young children: A comparison between scheduled awakening and systematic ignoring. *Pediatrics, 81,* 203–212.

Roberts, S. (1993). Tackling sleep problems through clinically-based approaches. *Health Visitor, 66,* 173–174.

Roehrs, T., Zorick, F., Witting, R., Paxton, C., Sicklesteel, J., & Roth, T. (1986). Alerting effects of naps in patients with narcolepsy. *Sleep, 9,* 194–199.

Rogers, A. E., & Aldrich, M. (1993). The effects of regularly scheduled naps on sleep attacks and excessive daytime sleepiness associated with narcolepsy. *Nursing Research, 42,* 111–117.

Rosen, G., Mahowald, M. W., & Ferber, R. (1995). Sleepwalking, confusional arousals, and sleep terrors in the child. In R. Ferber & M. Kryger (Eds.), *Principles and practice of sleep medicine in the child* (pp. 99–106). Philadelphia: W. B. Saunders.

Russo, R. M., Vymut, J., & Gurauraj-Allen, J. E. A. (1976). The effectiveness of diphenhydramine HCl in pediatric sleep disorders. *The Journal of Clinical Pharmacology, 16,* 284–288.

Sadeh, A. (1994). Assessment of intervention for infant night waking: parental reports and activity-based home monitoring. *Journal of Consulting and Clinical Psychology, 62,* 63–98.

Sadeh, A. (1996). Stress, trauma and sleep in children. *Child and Adolescent Psychiatric Clinics of North America, 5,* 685–700.

Sadeh, A. (1996). Evaluating night-wakings in sleep-disturbed infants: Methodological study of parental reports and actigraphy. *Sleep, 19,* 757–762.

Sadeh, A., & Anders, T. F. (1993). Infant sleep problems: origins, assessment, intervention. *Infant Mental Health Journal, 14,* 17–34.

Sadeh, A., Dark, I., & Vohr, B. R. (1996). Newborns' sleep–wake patterns: The role of maternal, delivery and infant factors. *Early Human Development, 44,* 113–126.

Sadeh, A., Hauri, P., Kripke, D., & Lavie, P. (1995a). The role of actigraphy in sleep medicine. *Sleep, 18,* 288–302.

Sadeh, A., Hayden, R. M., McGuire, J., Sachs, H., & Civita, R. (1994). Somatic, cognitive and emotional characteristics of abused children in a psychiatric hospital. *Child Psychiatry and Human Development, 24,* 191–200.

Sadeh, A., Klitzke, M., Anders, T. F., & Acebo, C. (1995b). Sleep and aggressive behavior in a blind retarded adolescent: A concomitant schedule disorder? *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 820–824.

Sadeh, A., Lavie, P., & Scher, A. (1994). Maternal perceptions of temperament of sleep-disturbed toddlers. *Early Education and Development, 5,* 311–322.

Sadeh, A., Lavie, P., Scher, A., Tirosh, E., & Epstein, R. (1991). Actigraphic home monitoring of sleep-disturbed and control infants and young children: A new method for pediatric assessment of sleep–wake patterns. *Pediatrics, 87,* 494–499.

Sadeh, A., McGuire, J. P. D., Sachs, H., Seifer, R., Trembley, A., Civita, R., Hayden, R. M. (1995c). Sleep and psychological characteristics of children on a psychiatric inpatient unit. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 813–819.

Salzarulo, P., & Chevalier, A. (1983). Sleep problems in children and their relationships with early disturbances of the waking–sleeping rhythms. *Sleep, 6,* 47–51.

Scher, A., Tirosh, E., Jaffe, M., Rubin, L., Sadeh, A., & Lavie, P. (1995). Sleep patterns of infants and young children in Israel. *International Journal of Behavioral Development, 18,* 701–711.

Scott, G., & Richards, M. (1990). Night waking in infants: Effects of providing advice and support for parents. *Journal of Child Psychology and Psychiatry, 31,* 551–567.

Scumacher, A., Merritt, S., & Cohen, F. L. (1993). The relationship between symptom severity and drug effectiveness as perceived by persons with narcolepsy. *Sleep Research, 22,* 263.

Seifer, R., Sameroff, A. J., Dickstein, S., Hayden, L. C., & Schiller, M. (1996). Parental psychopathology and sleep variation in children. *Child and Adolescent Psychiatric Clinics of North America, 5,* 715–727.

Seymour, F. W. (1987). Sleep problems in young children. *New Zealand Medical Journal, 100,* 347–349.

Seymour, F. W., Brock, P., During, M., & Pole, G. (1989), Reducing sleep disruptions in young children: evaluation of therapist-guided and written information approaches: a brief report. *Journal of Child Psychiatry, 30,* 913–918.

Sforza, E., & Lugaresi, E. (1995). Daytime sleepiness and nasal continuous positive airway pressure therapy in obstructive sleep apnea syndrome patients: effects of chronic treatment and 1-night therapy withdrawal. *Sleep, 18,* 195–201.

Sheldon, H. S., Spire, J. P., & Levy, H. B. (1992). *Pediatrics sleep medicine*. Philadelphia: W. B. Saunders.

Simonoff, E. A., & Stores, G., (1987). Controlled trial of trimeprazine tartrate for night waking. *Archives of Disease in Childhood, 62,* 253–257.

Stores, G. (1990). Sleep disorders in children. *British Medical Journal, 301,* 351–352.

Strachan, D. P., Anderson, H. R., Limb, E. S., O'Neill, A., & Wells, N. (1994). A national survey of asthma prevalence, severity, and treatment in Great Britain. *Archives of Diseases in Childhood, 70,* 174–178.

Stradling, J. R., Thomas, G., Warely, A. R. H., Williams, P., & Feeland, A. (1990). Effect of tonsillectomy on nocturnal hypoxaemia, sleep disturbance, and symptoms in snoring children. *Lancet, 335,* 249–253.

Strollo, P. J., & Rogers, R. M. (1996). Current concepts: obstructive sleep apnea. *The New England Journal of Medicine, 11,* 99–104.

Sullivan, C. E., & Grunstein, R. (1994). Continuous positive airway pressure in sleep disordered breathing. In M. H. Kryger, T. Roth & W. C. Dement (Eds.), *Principles and practice of sleep medicine* (2nd ed., pp. 694–705). Philadelphia: Saunders.

Terr, L. C. (1987). Nightmares in children. In C. Guilleminault (Ed.), *Sleep and its disorders in children* (pp. 231–242). New York: Raven Press.

Thoman, E. B. (1975). Sleep and wake behaviors in the neonates: consistencies and consequences. *Merrill-Palmer Quarterly, 21,* 295–314.

Thoman E. B. & Whitney, M. P. (1989). Sleep states in infants monitored at home: Individual differences, development trends and origins of diurnal activity. *Infant Behavior and Development, 12,* 59–75

Thorpy, M. J., Korman, E., Spielman, A. J., & Glovinsky, P. B. (1988). Delayed sleep phase syndrome in adolescents. *Journal of Adolescents Health Care, 9,* 222–227.

Tirosh, E., Sadeh, A., Munvez, R., & Lavie, P. (1993). Effects of Methylphenidate on sleep in children with attention-deficit hyperactivity disorder. *American Journal of Diseases of Childhood, 147,* 1313–1315.

Tzischinsky, O., Skene, D., Epstein, R., & Lavie, P. (1991). Circadian rhythms in 6-sulfatoxymelatonin and nocturnal sleep in blind children. *Chronobiology International, 8,* 168–175.

Van Tassel, E. B. (1985). The relative influence of child and environmental characteristics on sleep disturbances in the first and second years of life. *Developmental and Behavioral Pediatrics, 6,* 81–86.

Weissbluth, M. (1982). Modification of sleep schedule with reduction of night waking: A case report. *Sleep, 5,* 262–266.

Weitzman, E. D., Czeisler, C. A., Coleman, R. M., Spielman, A. J., Zimmerman, J. C., & Dement, W. C. (1981). Delayed sleep phase syndrome, a chronobiological disorder with sleep-onset insomnia. *Archives of General Psychiatry, 38,* 737–746.

World Health Organization (1992). *International statistical classification of diseases and related health problems (ICD-10)* (10th ed.). Geneva, Switzerland: Author.

Wolfson, A., Lacks, P., & Futterman, A. (1992). Effects of parent training on infant sleeping patterns, parents' stress, and perceived parental competence. *Journal of Consulting and Clinical Psychology, 60,* 41–48.

Werry J. S., & Carlielle, J., (1983). The nuclear family, suburban neurosis, and iatrogenesis in Aukland mothers of young children. *Journal of the American Academy of Child Psychiatry, 22,* 172–179.

Yoss, R., & Daly, D. (1960). Narcolepsy in children. *Pediatrics, 25,* 1025–1033.

Zarcone, V. (1973). Narcolepsy. *New England Journal of Medicine, 188,* 156–166.

Zuckerman, B., Stevenson, J., & Baily, V. (1987). Sleep problems in early childhood: Predictive factors and behavioral correlates. *Pediatrics, 80,* 664–671.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.25

Somatoform Disorders in Childhood and Adolescence

JUDY GARBER
Vanderbilt University, Nashville, TN, USA

5.25.1 INTRODUCTION 656
5.25.2 HISTORY OF SOMATOFORM DISORDERS 656
5.25.3 PHENOMENOLOGY 657
5.25.4 DIAGNOSTIC FEATURES 658
  5.25.4.1 Somatization Disorder 658
  5.25.4.2 Undifferentiated Somatoform Disorder 658
  5.25.4.3 Conversion Disorder 658
  5.25.4.4 Pain Disorder 658
  5.25.4.5 Hypochondriasis 659
  5.25.4.6 Body Dysmorphic Disorder 659
  5.25.4.7 Somatoform Disorder Not Otherwise Specified 659
5.25.5 DIFFERENTIAL DIAGNOSIS 659
5.25.6 COMORBIDITY 659
5.25.7 PREVALENCE 662
  5.25.7.1 Age Effects 663
  5.25.7.2 Sex Differences 663
  5.25.7.3 Sociocultural Factors 664
5.25.8 ETIOLOGY 665
  5.25.8.1 Genetics and Personality 665
  5.25.8.2 Somatic and Perceptual Sensitivity 665
  5.25.8.3 Family Processes 666
  5.25.8.4 Stress and Trauma 667
  5.25.8.5 Emotional Communication 668
5.25.9 COURSE AND PROGNOSIS 669
5.25.10 CONCEPTUALIZATION AND CLINICAL FORMULATION 669
  5.25.10.1 Assessment 669
  5.25.10.2 Treatment Formulation 671
5.25.11 TREATMENT 672
  5.25.11.1 Psychosocial 672
  5.25.11.2 Medical Interventions 674
5.25.12 CONCLUSION AND FUTURE DIRECTIONS 674
5.25.13 REFERENCES 675

5.25.1 INTRODUCTION

Medically unexplained somatic symptoms are a common problem in adults and there is increasing evidence that these physical symptoms begin during childhood and adolescence (Kellner, 1986; Pilowsky, Bassett, Begg, & Thomas, 1982). Excessive somatic complaints and associated illness behavior can lead to serious developmental problems when accompanied by school absences, academic failure, and withdrawal from normal social activities. Moreover, the search for diagnoses and cures of these persistent physical complaints can place children at risk of unnecessary medical tests and treatments, which can be very expensive for the family, medical profession, and society. During this time of increasing attention to the cost of health care, there is a growing concern about the expenditure of both human and economic resources to deal with pediatric somatization. Several reviews of somatization in children (Campo & Fritsch, 1994; Campo & Garber, in press; Garralda, 1992, 1996; Nemzer, 1996) have highlighted the importance of this area to researchers and clinicians in both the psychiatric and medical communities. This chapter builds on these earlier reviews and further synthesizes the literature on somatoform disorders in children and adolescents.

5.25.2 HISTORY OF SOMATOFORM DISORDERS

The term *somat* is derived from the Greek word meaning body. As early as ancient Greece, the clinical entity of conversion symptoms was recognized. The word hysteria referred to physical symptoms that could not be explained medically. In the past, hysteria was thought to be a disorder in women due to abnormal movement of the uterus, and the term "somatization" was used to describe the belief that some neuroses could be manifested as bodily disorders (Steckel, 1943). Steckel defined somatization as a bodily disorder that arises as the expression of a deep-seated neurosis, "especially as a disease of the conscious" (p. 580). More recently, Lipowski (1988) defined somatization as "a tendency to experience and communicate somatic distress and symptoms unaccounted for by pathological findings, to attribute them to physical illness, and to seek medical help for them" (p. 1359).

Briquet (1859) was one of the first in modern times to observe a condition in which patients complained of multiple chronic, unexplainable physical symptoms. He also was one of the first to describe scientifically the syndrome of "hysteria." In the later 1800s, several physicians including Charcot, Bernheim, and Janet began treating hysterical symptoms with hypnosis. Freud (1896/1962) proposed that painful memories and affect from a traumatic event, typically of a sexual nature, can be "dissociated" from conscious awareness and then "converted" into somatic symptoms that symbolize an aspect of the trauma. Such dissociation and conversion were hypothesized to be the processes underlying hysterical symptoms.

Psychoanalytic theory changed the term hysteria to conversion hysteria, and in 1952 the *Diagnostic and statistical manual* (*DSM-I*) (American Psychiatric Association (APA), 1952) modified the term conversion hysteria to conversion reaction. Dissociative reaction referred to the mental manifestations of hysteria including amnesia, fugue states, and multiple personality. In 1968, *DSM-II* (APA, 1968) relabeled the categories to be "hysterical neurosis, conversion type" and "hysterical neurosis, dissociative type." *DSM-II* also created a category for the chronic multisystem somatoform disorder, and referred to it as "Briquet's syndrome." Separate categories were designated for "conversion disorder" and "dissociative disorder" and only parenthetically retained the term "hysterical neurosis" in *DSM-III* (APA, 1978) and *DSM-III-R* (APA, 1987). Some have questioned, however, whether conversion and dissociative disorders should be categorized separately because it is not unusual for both to occur in the same individual, and both disorders are characterized by psychological dissociation (Nemiah, 1985). *DSM-III* expanded the number of diagnoses under the rubric of somatoform disorders to include Body Dysmorphic Disorder, Hypochondriasis, and Somatoform Pain Disorder in addition to Conversion Disorder and Somatization.

Some of the specific criteria defining the various Somatoform Disorders have changed in *DSM-IV*. For example, whereas in *DSM-III* and *DSM-III-R* Conversion Disorder was broader and included autonomic symptoms (e.g., vomiting, endocrine dysfunction), *DSM-IV* has returned to a narrower set of criteria involving neurological symptoms. Autonomic symptoms that used to be part of the Conversion Disorder criteria in *DSM-III* now fall under the category of "undifferentiated somatoform disorder" or "somatoform disorder not otherwise specified." The separation of autonomic and sensory-motor symptoms in the criteria for defining conversion disorders has been questioned, particularly with regard to children in whom conversion-like disorders are particularly likely to be characterized by a mixture of motor, sensory, and autonomic symptoms (Looff, 1970).

The symptoms defining Somatization Disorder also have undergone several changes. In *DSM-III-R* any 13 symptoms from a list of 35 were required for the diagnosis. In *DSM-IV* a certain number of symptoms are required from each of the four physiological categories. This is more similar to the criteria outlined in *DSM-III*, which required a certain number of symptoms from each of six categories.

Thus, the specific names of the diagnoses that comprise somatoform disorders as well as the criteria that define them have undergone several changes with each new version of the *DSM*. This chapter focuses on somatoform disorders as defined by *DSM-III-R* and *DSM-IV*.

5.25.3 PHENOMENOLOGY

According to the fourth edition of the *Diagnostic and statistical manual of mental disorders* (*DSM-IV*; APA, 1994), "the common feature of Somatoform Disorders is the presence of physical symptoms that suggest a general medical condition ... and are not fully explained by a general medical condition, by the direct effects of a substance, or by another mental disorder" (p. 445). To receive a diagnosis of a somatoform disorder, the symptoms must cause clinically significant distress or impairment in social, occupational (academic), or other areas of functioning. Marital discord and diminished work performance are common. The group of conditions categorized as Somatoform Disorders do not necessarily share a common etiology or mechanism. Rather, they are classified together due to the clinical utility of needing to exclude medical conditions or substances as the cause of the bodily symptoms that comprise the disorders. Although each of the seven diagnoses included under the rubric of somatoform disorders has its own unique set of symptoms and criteria, they share several characteristics in common. Associated features often include depression, anxiety, suicidality, and impulsivity. It is not unusual for substance abuse to develop as a result of medications taken to alleviate the physical discomfort of the medical symptoms. Personality disorders, particularly histrionic, dependent, borderline, and antisocial are not uncommon. There also may be significant academic/occupational and social impairment due to frequent school or work absences and lack of involvement in normal social or extracurricular activities.

The course of somatoform disorders tends to be chronic with exacerbations of symptoms waxing and waning. Prognosis tends to be best when onset is acute, not complicated by personality disorders, associated with an identifiable stressor, and there is an absence of secondary gain as a result of symptom complaining (*DSM-IV;* APA, 1994).

Individuals with somatoform disorders typically take on the "sick role" and engage in illness behavior. Illness behavior is the manner in which individuals perceive and respond to somatic sensations that may or may not signify disease (Mechanic, 1972). Whereas most people pay little attention to such sensations, persons with somatoform disorders tend to interpret even normal sensations as indicating disease, and they respond by excessive complaining of symptoms, seeking medical treatment, and limiting their activities so that it interferes with their occupational/academic or social functioning.

"Exaggerated," "abnormal," or "medically unexplained" illness behavior has been examined extensively in adults (e.g., Pilowsky, 1987; Reesor & Craig, 1988) and is a problem not only for the individual who suffers from distress and impaired functioning, but also for society, which bears the costs associated with lost work days and overutilization of health services. Approximately 50% of physician visits are for complaints for which medical evaluation yields no evidence of disease (Barsky, Wyshak, & Klerman, 1986). Because individuals with somatoform disorders maintain the belief that there is something physically wrong with them, they often undergo many painful, expensive, and unnecessary medical tests and procedures in search of the presumed disease. They are generally not reassured by negative findings and therefore seek out multiple physicians ("doctor shop"), looking for one who will provide them with a medical explanation for their symptoms. They tend to be resistant to suggestions that their symptoms might be associated with psychosocial factors and they often ignore referrals to mental health professionals.

Despite speculation that symptoms of somatoform disorders and associated illness behavior begin in childhood, little is known about childhood forms and precursors of these diagnoses (Parmelee, 1986). Because children typically do not take themselves to the doctor, children with somatoform disorders must first convince their parents of the need to seek medical treatment. It also is not unusual for parents to share the belief that the child's symptoms require medical attention, thereby reinforcing or actually helping to generate the illness behaviors and beliefs in the child.

5.25.4 DIAGNOSTIC FEATURES

In *DSM-IV* (pp. 445–469), the Somatoform Disorders comprise seven distinct diagnoses: Somatization Disorder, Undifferentiated Somatoform Disorder, Conversion Disorder, Pain Disorder, Hypochondriasis, Body Dysmorphic Disorder, and Somatoform Disorder Not Otherwise Specified. Table 1 presents the *ICD-10* disorders that correspond to the somatoform disorders in *DSM-IV*. The criteria that define each of the *DSM-IV* somatoform disorders are the same for children, adolescents, and adults. These criteria are described briefly below.

5.25.4.1 Somatization Disorder

Somatization Disorder is characterized by a pattern of recurrent, multiple somatic complaints that cannot be explained fully by any known general medical condition or the direct effects of a substance. If a medical condition is present, the symptoms or associated dysfunction are in excess of what would be expected given the history or medical findings. The physical complaints must have begun before the age of 30 and have occurred over several years. Somatization disorder rarely is diagnosed in children; it is not uncommon, however, for symptoms to begin in adolescence with menstrual difficulties being one of the earliest symptoms in women. For a diagnosis of somatization disorder, a history of a certain number of symptoms are required in each of four areas: pain, gastrointestinal, sexual/reproductive, and pseudoneurological.

5.25.4.2 Undifferentiated Somatoform Disorder

Undifferentiated Somatoform Disorder is very similar to somatization disorder but fewer symptoms are required. It is defined by one or more medically unexplained, distressing, or impairing somatic complaints that persist for at least six months. Chronic fatigue or weakness of unknown medical origin, sometimes referred to as "Neurasthenia," would be given this diagnosis. Because a shorter duration and fewer symptoms are required, this may be a more appropriate diagnosis for children than somatization disorder.

5.25.4.3 Conversion Disorder

Conversion Disorder involves symptoms or deficits in voluntary motor or sensory functions that suggest a neurological or general medical condition. Conversion symptoms are referred to as "pseudoneurological" and include such problems as paralysis, blindness, and seizures. Because the onset or exacerbation of the symptom or deficit is preceded by conflict or other stressors, conversion disorder is believed to be associated with psychological factors.

5.25.4.4 Pain Disorder

Pain disorder is defined specifically by pain, which causes significant distress or impairment, and is of sufficient severity to warrant clinical attention. According to *DSM-IV*, two types of pain disorder can be identified: Pain Disorder Associated with Psychological Factors, and Pain Disorder Associated with Both Psycholo-

Table 1 Somatoform disorders in *DSM-IV* and *ICD-10*.

DSM-IV somatoform disorders		*ICD-10 somatoform disorders*	
300.81	Somatization disorder	F45.0	Somatization disorder
300.81	Undifferentiated somatoform disorder	F45.1	Undifferentiated somatoform disorder
300.81	Somatoform disorder not otherwise specified	F45.9	Other somatoform disorder
300.11	Conversion disorder	F44	Dissociative (conversion)
		F44.4–44.7	Dissociative disorder of movement and sensation
307.80	Pain disorder associated with psychological factors	F45.4	Persistent somatoform pain disorder
307.89	Pain disorder associated with psychological factors and a general medical condition		
307.7	Hypochondriasis	F45.2	Hypochondriacal disorder
300.7	Body dysmorphic		
		F45.3	Somatoform autonomic dysfunction
		F48	Other neurotic disorders
		F48.0	Neurasthenia

gical Factors and a General Medical Condition. In both subtypes, psychological factors are judged to play an important role in the onset, maintenance, and severity of the disorder. The primary difference is that in the latter diagnosis, both psychological factors and medical factors contribute to the onset and exacerbation of the pain. Several recurrent pain conditions in children would receive this diagnosis (e.g., recurrent abdominal pain, headaches, chest pain).

5.25.4.5 Hypochondriasis

In Hypochondriasis, the individual is preoccupied with fears of having or the belief of already having a serious disease despite medical reassurance to the contrary. The person misinterprets one or more bodily signs or symptoms such as heartburn, shortness of breath, or a minor cough as being indicative of a serious illness. The level of concern is not delusional, however, and the person can acknowledge that the fears may be unwarranted.

5.25.4.6 Body Dysmorphic Disorder

Body Dysmorphic Disorder is the preoccupation with some imagined or slight physical defect in appearance, particularly about flaws of the face or head (e.g., hair, nose, skin, eyes). Individuals often have difficulty controlling their preoccupation with the supposed deformity and may spend hours a day thinking about their "defect." In this respect, Body Dysmorphic Disorder (dysmorphophobia) is similar to obsessive-compulsive disorder (Phillips, McElroy, Keck, Pope, & Hudson, 1993).

5.25.4.7 Somatoform Disorder Not Otherwise Specified

Somatoform Disorder Not Otherwise Specified includes those cases with significant somatoform symptoms that do not meet all of the criteria for one of the somatoform disorders. Examples are when symptoms have been of shorter than the necessary duration (e.g., hypochondriacal symptoms for less than six months) or when the extent of functional impairment is less than clinically significant.

Table 2 summarizes information from *DSM-IV* concerning other characteristics of the five main somatoform disorders including age of onset, sex differences, cultural influences, prevalence, course, and associated features.

5.25.5 DIFFERENTIAL DIAGNOSIS

An important feature of somatoform disorders is that the symptoms are not produced voluntarily. This is in contrast to Factitious Disorder and Malingering, in which somatic symptoms are intentional and controllable. In the case of Factitious Disorder, symptoms are feigned in order to assume the sick role, whereas in Malingering symptoms are produced for some external incentive such as financial gain or avoidance of responsibilities. Factitious Disorder by proxy refers to situations in which a parent or caretaker produces, feigns, or stimulates disease in a child, with the incentive appearing to be an internal one for the caretaker.

Somatoform disorders also need to be distinguished from Psychological Factors Affecting Physical Conditions (PFAPC). In PFAPC there is a general medical condition that completely accounts for the physical symptoms but that is adversely affected by specific psychological or behavioral factors. These psychological factors tend to negatively influence the course or treatment of the medical condition. A common example of this is when the patient is noncompliant with a treatment regime (e.g., taking insulin for diabetes) due to emotional or behavioral problems.

5.25.6 COMORBIDITY

Comorbidity is defined as the co-occurrence of two or more disorders in the same individual at the same point in time. It is quite common for individuals with somatoform disorders to have other psychiatric problems and diagnoses including anxiety, depression, suicidality, impulsivity, substance abuse, and personality disorders. In an epidemiological study of psychiatric disorders in adults, Swartz, Landerman, and George (1991) reported that all persons with somatization disorder received another diagnosis, most commonly affective disorder, panic disorder, or simple phobia. In a community sample of adolescents, Taylor, Szatmari, Boyle, and Offord (1996) reported that about 42% of the youth who self-reported high rates of somatic complaints had at least one psychiatric diagnosis, most commonly an emotional disorder.

Several studies have found high rates of anxiety and mood symptoms and disorders among children with recurrent abdominal pain (RAP) (Garber, Zeman, & Walker, 1990; Hodges, Kline, Barbero, & Flanery, 1985a; Hodges, Kline, Barbero, & Woodruff, 1985b; Walker, Garber, & Greene, 1993; Wasserman, Whitington, & Rivera, 1988). Garber et al. (1990) conducted careful psychiatric interviews of children with RAP and found that 11 of 13 had an anxiety disorder and five of 13 had a

Table 2 Somatoform disorders (*DSM-IV*).

Disorder	*Somatization*	*Conversion*	*Pain*	*Hypochondriasis*	*Body dysmorphic*
Age	Onset must begin before age 30	Onset usually in late childhood to early adulthood	Can occur at any age	It can begin at any age, most commonly during early adulthood	Usually begins during adolescence
	Symptoms typically begin by adolescence	Onset rarely before age 10 year or after age 35 years			
	Rarely diagnosed before puberty	In children under 10 years old; gait problems or seizures are most common			
Gender	More common in women in US	More common in females than males (2:1 to 10:1)	Both males and females experience pain disorders	Equally common in males and females	About equal in males and females
	Higher rates in men in some cultures (e.g., Greece, Puerto Rico)	In women, conversion symptoms may develop into somatization disorder	Women experience more headaches than men		
		In men, associated with antisocial personality	Adolescent girls report more abdominal pain than boys		
Culture	Content of complaints varies with cultures	More common in lower SES, less educated, developing countries	Expressions of and responses to pain varies across cultures	Must consider cultural definitions of illness and what “healers” in their culture tell them about their illness	Culture can influence the extent of emphasis on appearance
	Sex ratio varies across cultures	Specific symptoms reflect local cultural ideas			
Prevalence	0.2–2% in women	1–3% of adult outpatients	Relatively common about 10–15% of adults have significant back pain	4–9% of patients in general medical practices	Little is known about its prevalence
	0.2% in men	Ranges from 11/100 000 to 300/100 000	About 10–15% of school children have abdominal pain		May be more common than previously thought

Table 2 (continued)

Disorder	*Somatization*	*Conversion*	*Pain*	*Hypochondriasis*	*Body dysmorphic*
Course	Chronic, but fluctuating Rarely remits	Onset is generally acute, but may have gradually increasing symptoms Episodes of acute symptoms are brief, but recurrent	Acute pain episodes tend to resolve after a brief duration Chronic pain can persist for years	Usually a chronic course, although symptoms may wax and wane	Onset may be gradual or abrupt Chronic course
Associated features	Functional disorders (e.g., irritable bowel syndrome)	*La belle indifference* Dissociation	With: psychological factors and/or general medical conditions Insomnia	May become homebound May have poor insight about excessiveness of complaints	Excessive grooming behavior and "mirror checking" Seek plastic surgery "Ideas of reference" about the presumed defect Eating disorders

mood disorder. Similarly, Wasserman et al. (1988) reported that 26 of 31 RAP patients had psychiatric disorders, with anxiety being the most common. Walker et al. (1993) reported that children with RAP had significantly higher levels of emotional symptoms than healthy controls, but less than children being evaluated in a psychiatric clinic specifically for emotional problems. Thus, children with RAP clearly have comorbid emotional problems, although their symptoms tend to be less severe than found among children receiving psychiatric treatment for anxiety and depression.

In their review, Bass and Murphy (1995) concluded that somatoform disorders are associated with very high rates of personality disorders and that the illness behavior of persons with somatoform disorders is "a lifestyle in which the sick role is a model of relating to self and others" (p. 424). They argued, therefore, that it might make more sense to think of the somatoform disorders as Axis II personality disorders rather than as Axis I disorders with comorbid personality disorders.

5.25.7 PREVALENCE

Epidemiological studies indicate that somatization disorder occurs in only 0.05–0.38% of the adult population (Swartz et al., 1991). The diagnosis of somatoform disorders is extremely rare in prepubertal children (Offord et al., 1987), although it has been diagnosed among adolescents. In a study of the six-month prevalence of child psychiatric disorders in a community sample, Offord et al. reported that somatization disorder was so rare among children 4–11-years old ($n = 1442$) that prevalence estimates for that age group could not be obtained reliably. Among adolescents aged 12–16 years, polysymptomatic somatization disorder was identified in 4.5% of boys ($n = 608$) and 10.7% of girls ($n = 624$). Thus, somatization disorder appears to increase with age and is about twice as common in adolescent girls as boys.

Other studies examining prevalence of child psychiatric disorders in pediatric (Costello et al., 1988) and community samples (Bird et al., 1988), however, have failed to report the rates of any of the somatoform disorders. This might be because somatoform disorders are considered to be so rare during childhood that researchers have decided that it was not worth the extra time to interview everyone about these syndromes. In addition, because information about actual medical status is needed to make the diagnoses, it is difficult to obtain this in large epidemiological studies. It is likely that it is for these reasons that somatoform disorders have been excluded from the various versions of the most frequently used interview in epidemiological studies (i.e., the Diagnostic Interview Schedule for Children, DISC; Costello; Edelbrock, Kalas, Kessler, & Klaric; 1982; Shaffer et al., 1996). Questions about somatoform disorders were not included in either the completed National Institute of Mental Health (NIMH)-funded Methods for the Epidemiology of Child and Adolescent Mental Disorders (MECA) study (Lahey et al., 1996) or the planned NIMH-funded Multi-Site Study of Mental Health Service Use, Need, Outcomes, and Costs in Child and Adolescent Populations (UNOCCAP) study. Therefore, the lifetime incidence and one-year prevalence rates of any of the Somatoform Disorders will remain unknown. A separate study in which both medical and psychiatric symptoms can be assessed in a large epidemiological sample is needed to determine more precisely the incidence and prevalence of somatoform disorders in children and adolescents.

Some children meeting the diagnostic criteria for somatization disorder have been identified in clinical samples (Kriechman, 1987; Livingston & Martin-Connici, 1985), but this is quite unusual. Several investigators have suggested that the *DSM* criteria are inappropriate for children, particularly given the fact that sexual and reproductive symptoms are required (Garber, Walker, & Zeman, 1991; Livingston & Martin-Connici, 1985). Other diagnoses such as undifferentiated somatoform disorder or pain disorder are probably more applicable to children.

In contrast to the finding that diagnosed somatoform disorders are rare in childhood, physical complaints have been shown to be quite common in children and adolescents (Garber et al., 1991; Goodman & McGrath, 1991). In a sample of 540 children in grades three through 12, Garber et al. assessed how frequently children had had each of 36 physical complaints in the preceding two weeks. About half the sample indicated on the Children's Somatization Inventory (CSI; Walker, Garber, & Greene, 1991; Walker & Greene, 1989, 1991a) at least one physical symptom during the prior two weeks; the most commonly reported symptoms were headaches, low energy, sore muscles, and abdominal discomfort. Factor analysis of the CSI revealed four symptom clusters that corresponded to four of the categories that defined somatization disorder in *DSM-III-R*: cardiovascular, gastrointestinal, pain/weakness, and pseudoneurological. About 1% of the sample endorsed at least 13 symptoms, and approximately 15% reported four or more symptoms. This was an important study because it provided information about the rates of self-reported somatic complaints across a wide age range. However, it is limited because independent medical assessments were not conducted to determine whether the reported physical symptoms were the result of actual disease. Thus, the rates of medically unexplained symptoms in children might be lower than that reported in this study.

Taylor et al. (1996) similarly conducted a community study with over 1000 Canadian adolescents (aged 12–16 years) and found the rates of some somatic symptoms and concerns to be quite high. About 50% (44.3–54.4%) of these teens reported that "their health should be better," headaches, worries about health, stomach aches or cramps, and aches and pains. Loss of functions (e.g., ability to swallow, voice, consciousness, hearing, sight) were reported much less frequently (2.9–7.5%). Although the rates of some individual symptoms were high, there did not appear to be a group of teens with extreme scores who might meet the criteria for somatization disorder. The average number of symptoms for those youth who had visited any medical service in the previous six months was 4.9, which is far below the 13 required for a diagnosis of somatization disorder.

Several other studies have found relatively high rates of somatic symptoms and concerns in children. The most common symptom is headaches, which has been reported in 10–30% of children and adolescents (Aro, Paronen, & Aro, 1987; Belmaker, Espinoza, & Pogrund, 1985; Larson, 1991; Oster, 1972; Rutter, Tizard, & Whitmore, 1970). Recurrent abdominal pain also is very common, reported in 10–24% of school-aged children and adolescents (Apley, 1975; Apley & Naish, 1958; Belmaker et al., 1985; Garber et al., 1991; Hyams, Burke, Davis, Rzepski, & Andrulonis, 1996). Other common somatic symptoms in children and adolescents are limb pains or so-called "growing pains" found in 5–20% of children (Apley, 1958; Belmaker et al., 1985; Garber et al., 1991; Larson, 1991; Oster, 1972), chest pain in about 7–15% (Belmaker et al.; Garber et al.), daily fatigue in about 15%, and 33.3–50% of adolescents complaining of fatigue weekly (Belmaker et al.; Larson, 1991), dizziness in about 15% of children (Garber et al.; Larson, 1991), and other gastrointestinal symptoms such as nausea and vomiting (Apley, 1975; Garber et al.; Larson, 1991). Worries about health also are common in children. Orr (1986) found that 40% of children and adolescents endorsed worrying about their health often, and 20% indicated they worried "all the time" about their health.

Pseudoneurological symptoms tend to be reported much less frequently (Garber et al., 1991; Rutter et al., 1970; Stefansson, Messina, & Meyerowitz, 1976). The most commonly reported conversion or pseudoneurological symptoms in children and adolescents are pseudoseizures, fainting, gait abnormalities, and sensory symptoms (Goodyer, 1981; Goodyer & Mitchell, 1989; Grattan-Smith, Fairley, & Procopis, 1988; Lehmkuhl, Blanz, Lehmkuhl, & Braun-Scharm, 1989; Leslie, 1988; Maloney, 1980; Spierings, Poels, Sijben, Gabreels, & Renier, 1990; Steinhausen, Aster, Pfeiffer, & Gobel, 1989; Volkmar, Poll, & Lewis; 1984).

5.25.7.1 Age Effects

Somatoform disorders are rarely diagnosed before puberty, although specific somatic complaints are not uncommon during childhood. For example, the rate of recurrent abdominal pain peaks at about age nine and the rate of headaches peaks about age 12 (Oster, 1972), and even preschoolers have been reported to have somatic symptoms (e.g., stomach aches, vomiting) associated with psychological distress (Prugh, 1983). Pseudoneurological symptoms and diagnosed conversion disorder are extremely rare before age six, but increase with age particularly during adolescence (e.g., Lehmkuhl et al., 1989; Stefansson et al., 1976; Volkmar et al., 1984). Similarly, polysymptomatic complaints become increasingly more common during adolescence (Achenbach, Conners, Quay, Verhulst, & Howell, 1989; Offord et al., 1987). Although all the criteria for somatization disorder are rarely met in childhood, the majority (55%) of adults receiving a diagnosis of somatization disorder report experiencing their first somatic symptoms before age 15, and their symptoms tended not to remit (Swartz et al., 1991). Thus, not all symptoms that comprise somatoform disorders may be present across development, although it is likely that some symptoms and precursors of these disorders can be observed by early adolescence.

5.25.7.2 Sex Differences

Table 2 shows the rates of each of the somatoform disorders in men and women as reported in *DSM-IV*. There are some differences between children and adults in the rates of somatoform disorders in girls vs. boys. Prior to puberty the prevalence of somatization disorders is about equally rare in girls and boys. With the onset of puberty and menarche in girls, there tends to be an increase in the rates of somatic complaints in girls compared to boys (Aro & Taipale, 1987; Belmaker, 1984). There appears to be a decline in symptom reporting among adolescent boys (Garber et al., 1991). Moreover, girls tend to be more consistent than boys in reporting somatic symptoms over time (Rauste-von Wright & von Wright, 1981; Walker & Greene, 1991b).

Pseudoneurological or conversion symptoms have also been found to be more prevalent in females than males, and this begins to become apparent during adolescence (Lehmkuhl et al., 1989; Turgay, 1980; Volkmar et al., 1984). Among prepubertal children, some studies have found the rates of conversion disorder to be about equal for boys and girls (Stevens, 1986), whereas other studies have reported the rates to be higher among girls than boys (Goodyer, 1981; Volkmar et al., 1984).

The greater rate of somatic complaints among adolescent girls than boys could be due to several factors. There may be an actual rise in the number and types of somatic symptoms girls experience during adolescence, particularly due to symptoms associated with their menstrual cycle. It also is possible, however, that the discrepant rates are due to sex differences in reporting tendencies. Girls may be more willing to express their distress to others and might view it as a legitimate means of obtaining social support. Indeed, girls generally seek support from others more than boys (Belle, Burr, & Cooney, 1987), and in particular girls seek care for their physical complaints more than boys (Lewis & Lewis, 1989). Moreover, in a study of children's use of emotion display rules, Zeman and Garber (1996) found that girls reported they would be significantly more likely than boys to express pain. In addition, compared to boys, girls reported that they thought other people would be more accepting and understanding of their expressions of emotions. Thus, boys may be less willing to endorse somatic symptoms because they think others will be less tolerant of such complaints or see it as unmasculine.

These sex differences could have their roots in children's socialization history. For example, Walker and Zeman (1992) found that parents encouraged illness behavior more in girls than in boys. Fuchs and Thelen (1988) found that children reported that their parents reacted differently to the emotional displays of girls vs. boys. In general, parental socialization pressures tend to be perceived by children to be directed toward the masking of distress in boys vs. girls.

Thus, there appear to be sex differences in the rates of somatic complaints and illness behavior in children, and this becomes particularly apparent during adolescence. These differences could be the result of girls actually experiencing more somatic distress than boys, as well as to girls' greater willingness to express this distress to others.

5.25.7.3 Sociocultural Factors

Although it is likely that social, cultural, and ethnic factors might influence the development and expression of somatic complaints (Kirmayer, 1984), there has been little research on this issue in children. Whereas some studies have found an association between somatoform disorders in children and low socioeconomic status (SES), low levels of parental education, or living in a rural area (Aro et al., 1987; Proctor, 1958; Steinhausen et al., 1989), other studies have not found such a relation (Stevenson, Simpson, & Bailey, 1988; Walker & Greene, 1991a). In developing countries, 1.3–5% of child psychiatry outpatients have been found to have conversion disorders (Herman & Simonds, 1975; Proctor, 1958).

It is possible that cultural factors influence the presentation of somatic symptoms and may contribute to the higher incidence of conversion disorder in less-developed countries. Prazar (1987) asserted that less sophisticated individuals are more likely to present with symptoms that are not easily explained physiologically, although he argued that there is no relation between conversion symptoms and SES. In contrast, Taylor et al. (1996) found in a community sample of Canadian adolescents a significant association between poverty and somatic complaints. It is possible, however, in the Taylor et al. study that these somatic symptoms reflected real illnesses due to their less favorable living conditions.

Studies in Turkey (Turgay, 1980) and India (Srinath, Bharat, Girimaji, & Seshadri, 1993) have found conversion symptoms to be a common presentation of psychiatric disorders. For example, Srinath et al. (1993) reported that conversion disorder (mostly pseudoseizures) was diagnosed in 30% of inpatients and 15% of outpatients.

Weisz and colleagues (Weisz et al., 1987; Weisz et al., 1993) have compared the behavioral and emotional problems of children and adolescents in the US vs. Thailand. Weiss et al. (1987) found that among children aged 6–11 years, five of the six differences that were considered to be medium to large effects involved somatic difficulties. Thai children were rated by their parents to be higher on "doesn't eat well," "constipated," "feels dizzy," and "underactive, lacks energy." Thai adolescents also were rated higher on the latter three symptoms (Weisz et al., 1993). However, Thai children were not more likely than American children to have the most common somatic complaints such as headaches and stomach aches. Thai children were rated higher than American children on overcontrolled behaviors. Weisz et al. (1987, 1993) suggested that there probably are more cultural pressures in this Asian society for children to have self-control, emotional restraint, and social inhibition. It could be that such societal demands increase the likelihood of children expressing their distress somatically rather than behaviorally.

5.25.8 ETIOLOGY

Multiple factors have been proposed to explain the etiology of somatoform disorders including genetics, personality, somatic and perceptual sensitivity, family processes, stress and trauma, and emotional communication. Because of the heterogeneity of somatoform disorders, it is likely that different models will explain the specific disorders that comprise this category. It is also likely that a more complex biopsychosocial model that incorporates different aspects of various perspectives will ultimately explain the processes leading to the development of somatoform disorders.

5.25.8.1 Genetics and Personality

Somatization is clearly familial (Edwards, Zeichner, Kuczmierczyk, Boczkowski, 1985; Garber et al., 1990; Kriechman, 1987; Routh & Ernst, 1984; Walker et al., 1991; Wasserman et al., 1988), which is likely due to a combination of genetic and environmental factors. Adoption studies have found somatization disorder to be over five times more common in female first-degree relatives of probands with somatization than in the general population (Cloninger, 1986). Twin studies, on the other hand, have not provided much evidence of a strong genetic contribution to the development of somatization (Gottesman, 1962; Inouye, 1972; Torgensen, 1986), although there might be a genetic factor for somatic anxiety (Kendler, Heath, Martin, & Eaves, 1987).

Cloninger and colleagues (Bohman, Cloninger, von Knorring, & Sigvardsson, 1984; Cloninger, Sigvardsson, von Knorring, & Bohman, 1984; Guze, Cloninger, Martin, & Clayton, 1986) have also found genetic links between somatization disorder and antisocial personality and alcoholism. The biological fathers of adopted-away women who were "high-frequency somatizers" tended to have a history of violent crime (Cloninger, 1986). Cadoret (1978) similarly found that adopted-away daughters of antisocial patients were more likely to suffer medically unexplained somatic symptoms compared to controls. Studies of somatizing children have also found high rates of externalizing disorders among their relatives (Kriechman, 1987; Routh & Ernst, 1984).

Shields (1982) proposed that the personality traits that predispose to somatization may be what is inherited rather than somatization *per se*. Cloninger (1986) suggested that persons with somatization disorder are characterized by distractibility, impulsivity, and failure to habituate to repetitive stimuli. It is possible that these traits contribute to the link between somatization and antisocial personality disorder.

In a study specifically designed to assess personality, Stern, Murphy, and Bass (1993) reported that 72% of somatization patients compared to only 36% of controls had personality disorders, particularly passive-dependent, histrionic, and sensitive-aggressive. Tyler, Fowler-Dixon, and Ferguson (1990) have described a hypochondriacal personality disorder that presumably underlies somatization disorder and includes excessive preoccupation with health, distorted perceptions of minor symptoms, and frequent medical consultations. In children, excessive somatic complaints have been associated with obsessiveness, conscientiousness, hypersensitivity, insecurity, lack of perceived competence, anxiety, and possibly behavioral inhibition (Cunningham et al., 1987; Garralda, 1992; Kagan, Reznick, & Snidman, 1988; Walker et al., 1993). The extent to which these temperamental features during childhood become personality disorders that then predispose individuals to develop adult somatoform disorders needs to be explored further.

5.25.8.2 Somatic and Perceptual Sensitivity

Individuals with somatoform disorders have also been hypothesized to have greater somatic and/or perceptual sensitivity (Barsky & Klerman, 1983; Kellner, 1990). Various forms of normal physiological arousal and bodily changes are associated with somatizing symptoms. For example, contractions of the scalp muscles are observed with headaches (e.g., Fujii, Kachi, & Sobue, 1981; Sainsbury & Gibson, 1954); elevated electromyograph levels in paravertebral muscles can be found in patients with lower back pain (Dolce & Raczynski, 1985; Flor, Turk, & Birbaumer, 1985); contractions of smooth muscle in the alimentary tract are common in abdominal pain patients (Whitehead & Schuster, 1985), and esophageal motility is found in persons with chest pain (Clouse & Lustman, 1983). These normal physiological reactions can develop into more persistent or severe symptoms in several ways. First, somatizers might have a more sensitive physiological system such that they respond more intensely than normals to the same level of physical stimulation. Second, even if their initial response is within the normal range, they might have a more sustained physiological reaction and a slower return to baseline. Third, it is possible that somatizers respond normally physiologically, but simply attend to their symptoms more than others. Finally, somatizers also might misinterpret their normal bodily sensations as indicating a more serious condition than is warranted.

It is likely that individuals differ in their degree of somatic sensitivity, that is, the intensity with which they experience a particular level of somatic stimulation. Moreover, under conditions of stress, high somatic sensitivity will increase the likelihood that symptoms of affective arousal will be translated into somatic symptoms. The existing evidence in children relevant to this hypothesis is mixed. Using a painful cold pressor stimulus, Feuerstein, Barr, Francoeur, Houle, and Rafman (1982) did not find any meaningful differences in the autonomic, subjective, or behavioral responses of children with RAP compared with healthy controls. In contrast, using lower intensity stimulation with an algometer, Alfven (1993) reported that RAP children had significantly lower pressure pain thresholds than well children. Thus, if somatizers are hypersensitive to physical stimulation, this might be particularly apparent at lower levels of stimulation and for some forms of stimulation but not others.

Even if somatizers do not respond more sensitively initially to somatic stimulation, it also might be that they have a more prolonged somatic reaction once exposed. The reasons for this sustained reaction could be either physical or psychological; that is, once aroused it simply might take more time for their physiology to return to baseline, and therefore they experience the symptoms longer. It also is possible, however, that individuals who somatize are more likely to continue to attend to their symptoms, thereby sustaining a higher level of arousal for a longer period of time.

The extent to which individuals attend to normal bodily sensations can be learned. For example, Pennebaker and Skelton (1978) found that the instructions given to subjects could influence whether or not they perceived a bodily sensation. Moreover, habitual attention to somatic symptoms in certain parts of the body is likely to increase one's perceptions of sensations in that area, (Kellner, 1986). Such attentional variation can lead to either an amplification or minimization of bodily sensations (Barsky, Goodman, Lane, & Cleary, 1988; Mayou, 1976).

Finally, normal physiological reactions are sometimes misinterpreted as reflecting a more serious somatic problem, that is, whereas most individuals normalize common symptoms and attribute them to minor physical variation or environmental factors, individuals who somatize tend to attribute such normal bodily symptoms to much more pathological causes (Robbins & Kirmayer, 1991). This misinterpretation itself could increase the amount of stress experienced by the individual, and thereby lead to a more severe or sustained somatic reaction.

Thus, individuals with somatoform disorders are hypothesized to have an increased somatic and/or perceptual sensitivity. Somatic sensitivity is likely the result of a more severe initial physiological response to stimulation or a more sustained reaction with a slower return to baseline. Perceptual sensitivity is reflected by a normal physiological reaction that receives greater attention or that is misinterpreted as reflecting a more serious somatic problem.

5.25.8.3 Family Processes

The family environment can influence the development of somatization symptoms in children in several ways. Although the specific amount of genetic vs. environmental factors which contribute to the high rates of somatization in relatives of children with somatoform disorders is unclear, the presence of "ill" persons in the family environment can influence the child's psychological development through such processes as modeling and reinforcement (Craig, 1978; Fordyce, 1978; Turkat, 1982). Walker et al. (1993) suggested that the high numbers of potential models of illness behavior in families of children with recurrent abdominal pain provide more opportunities for vicarious learning of such behavior and for the observation of these models receiving reinforcement of their illness behavior (Whitehead, Busch, Heller, & Costa, 1986).

Operant learning can also play a role in the development and maintenance of recurrent pain in children (McGrath, 1987) and in the reduction of symptoms and disability (Sanders et al., 1989; Sank & Biglan, 1974). Positive reinforcement of expressions of pain (Osborne, Hatcher, & Richtsmeier, 1989) and discouragement of adaptive coping responses have been found to be associated with higher rates of illness behavior in children (Dunn-Geier, McGrath, Rourke, Latter, & D'Astous, 1986). Walker et al. (1993) found that children with abdominal pain reported that their parents responded more frequently to their symptom complaints with increased attention and special privileges compared to well children.

Whereas some have reported that parental lack of care and childhood illness significantly predicted adult somatization (Craig, Boardman, Mills, Daly-Jones, & Drake, 1993), others have suggested that parents of somatizing children are overprotective and promote separation fears (Grattan-Smith et al., 1988; Lehmkuhl et al., 1989; Robinson, Alverez, & Dodge, 1990) which can increase children's sense of vulnerability and fragility (Green & Solnit, 1964). The children's somatic symptoms can serve to maintain proximity to important attachment figures (Henderson, 1974; Wooley, Blackwell, & Winget, 1978) and may be the consequence of children's anxiety about being separated from such persons (Garber et al., 1990; Hodges et al., 1985b).

Family system theorists have suggested that psychosomatic symptoms can serve a function within the family system to avoid conflict, particularly between the parents (Minuchin et al., 1975; Minuchin, Rosman, & Baker, 1978; Mullins & Olson, 1990). Disturbed patterns of communication, cohesion, and support have been reported among families of children with somatoform disorders (Maloney, 1980; Walker, McLaughlin, & Greene, 1988; Wasserman et al., 1988), although not necessarily worse than families of children with emotional disorders (Walker et al., 1993). In somatizing families, physical symptoms often become the means of communication of distress, and are the focus of attention rather than the relationship problems that exist among family members (Lask & Fosson, 1989; Maisami & Freeman, 1987).

Wood (1993) proposed a biobehavioral family model that focuses on the mutual influence of individual factors, such as psychophysiologic reactivity, and family patterns, such as parental discord, triangulation, and negative affectivity. She suggested that high levels of individual psychophysiologic reactivity and stressful family patterns will potentiate one another and interact to influence the outcome of disease processes. This model needs to be tested further with regard to children with somatoform disorders in particular.

Thus, socialization processes can contribute to the development of excessive somatic complaints in children (Parmelee, 1986) and the eventual onset of somatoform disorders. Further research is needed in which the actual behaviors of parents in response to their children's somatic complaints is observed. Longitudinal studies are also needed to untangle the direction of causality between parenting behavior and child symptoms. Such studies can serve as a basis for interventions aimed at preventing the development and maintenance of somatization symptoms and disorders.

5.25.8.4 Stress and Trauma

Higher levels of negative life events have been found to be associated with greater frequency of illness episodes and health service utilization (Beautrais, Fergusson, & Shannon, 1982; Boyce et al., 1977; Roghmann & Haggerty, 1973), symptom maintenance, and illness exacerbations (Chase & Jackson, 1981; Perrin, MacLean, & Janco, 1988). Among children with somatization disorders in particular, some studies have found an association with stressful life events (Aro, 1987; Greene, Walker, Hickson, & Thompson, 1985; Hodges et al., 1985a; Pantell & Goodman, 1983; Robinson et al., 1990; Scaloubaca, Slade, & Creed, 1988; Walker & Greene, 1987), whereas other studies have failed to find such a relation (McGrath, Goodman, Firestone, Shipman, & Peters, 1983; Turgay, 1980; Walker & Greene, 1991a; Wasserman et al., 1988). The life-event profiles of children with somatizing disorders have been found to be similar to those of children with psychiatric disorders and different from healthy controls (Hodges et al. ; Scaloubaca et al., 1988; Turgay, 1980), although other studies have reported that children with recurrent pain experience fewer negative life events than children with emotional disorders (Walker et al., 1993). The different results across these studies may be due to differences in the types of psychiatric comparison groups studied and in the types of life events assessed. For example, children with somatizing problems tend to report a greater number of events involving family illness in particular compared to healthy and psychiatric children (Hodges et al.; Walker et al.; Wasserman et al.). There also is some evidence that early loss is linked to somatization and recurrent pain (Mallouh, Abbey, & Gillies, 1995; Wasserman et al.).

It also is possible that stress is a general risk factor that contributes to the onset and maintenance of many types of disorders. The more interesting and important question is what factors in combination with stress make some people vulnerable to one disorder rather than another. That is, what is the particular model in which stress is a part that predicts specific disorders (Garber & Hollon, 1991)? In a prospective study of 197 pediatric patients with chronic abdominal pain, Walker, Garber, and Greene (1994) found that among children who had high levels of negative life events, those who had been reported to be low in social competence had higher levels of somatic complaints one year later compared to children who were high in social competence. In addition, Walker et al. found that among boys who experienced high levels of negative life events, those whose mothers were characterized by high levels of somatic symptoms had higher levels of somatic symptoms a year later. Thus, both social competence and maternal somatic complaints moderated the relation between stress and somatization in pediatric patients. This indicates that a more complex model which includes stress in combination with other psychosocial factors may be needed to explain the onset and maintenance of somatization in children. The extent to which such a model generalizes to normative populations needs to be studied further.

Traumatic life events, particularly physical and sexual abuse, have also been linked to somatization. Early theories suggested that childhood trauma was a factor in most cases of hysteria (Briquet, 1859; Freud, 1896/1962). Despite the general rejection of Freud's "seduction theory," there continues to be an interest in the association between childhood physical and sexual abuse and somatic and conversion symptoms (Green, 1993; Loewenstein, 1990; Morrison, 1989; Rimza, Berg, & Locke, 1988). Pseudoneurological symptoms (e.g., pseudoseizures), genitourinary and gastrointestinal symptoms, in particular, have been observed among victims of sexual abuse (Bowman, 1993; Goodwin, Simms, & Bergman, 1979; Hunter, Kilstrom, & Loda, 1985; Klevan & DeJong, 1990; LaBarbera & Dozier, 1980). Prospective studies of abused children are needed to establish the predictive link between early trauma and subsequent somatoform symptoms and disorders.

5.25.8.5 Emotional Communication

Traditional psychodynamic theory has been particularly prominent in the explanation of conversion disorder. According to this view, an unconscious intrapsychic conflict, wish, or need is "converted" into a somatic symptom, which presumably expresses symbolically some aspect of the conflict while at the same time protecting the person from conscious awareness (Engel, 1962). Critics of this perspective argue that the evidence is limited concerning the symbolic meaning of the symptom and its link to the unconscious.

More modern approaches assert that somatizing is a psychological defense whereby individuals experience and express emotional distress physically rather than acknowledge unpleasant affects, memories, or conflicts (Shapiro & Rosenfeld, 1987; Simon, 1991). Constructs such as alexithymia (Nemiah, 1977; Sifneos, 1973) and the defensive style of "repression" (Weinberger, Schwartz, & Davidson, 1979) have been proposed to explain the association between somatic symptoms and reduced emotional awareness and expression.

Alexithymia is characterized by difficulties in identifying and describing feelings, difficulty distinguishing between feelings and the bodily sensations of emotional arousal, limited imagination and fantasy life, and a concrete and reality-based cognitive style (Taylor, Baglay, Ryan, & Parker, 1990). Adult patients with somatization disorder have been found to have higher scores on a self-report measure of alexithymia compared to patients with other psychiatric disorders (Bagby, Taylor, & Parker, 1994; Taylor et al., 1990), although alexithymia has also been found to be associated with depression, neuroticism, and eating disorders (Bourke, Taylor, Parker, & Bagby, 1992; Cohen, Auld, & Brooker, 1994; Wise & Mann, 1994). More research needs to be conducted regarding the measurement and construct validity of alexithymia in adults, and its particular relation to somatoform disorders. There also has been almost no research on alexithymia in children.

Because young children have limited cognitive and linguistic abilities, it is not unusual to hear them express their emotional distress with physical symptoms. This might be particularly apparent in those families or cultures in which the expression of emotional distress is discouraged, whereas somatic complaints receive acceptance and attention (Escobar, Burnam, & Karno, 1987). Although there is a large developmental literature concerning the emergence of emotional competence in children (e.g., Saarni, 1990), little is known about the emotional development of children with somatoform disorders. Children who somatize might be especially likely to show delays or deviation in their emotional competence, particularly regarding emotional identification and expression; that is, they might be less competent at linking affects with situations, or expressing emotions in certain contexts, particularly those involving negative affect. Thus, although there has been considerable theoretical speculation about the role of affect expression and emotional competence in the development of somatization, the relevant empirical investigations have not yet been conducted.

5.25.9 COURSE AND PROGNOSIS

There have been few systematic longitudinal studies of children and adolescents with somatoform disorders and therefore little is known about their course and prognosis. Studies have varied with regard to the types of patients and disorders studied, the length of the follow-ups, and the specificity vs. generality of the outcomes assessed. In general, acute onset of fewer and transient symptoms tends to have a better prognosis than insidious onset of multiple symptoms that are recurrent or chronic (Ernst, Routh, & Harper, 1984; Grattan-Smith et al., 1988; Robins & O'Neal, 1953), although the findings have not always been consistent with this perspective (Goodyer & Mitchell, 1989; Lehmkuhl et al., 1989).

Several follow-up studies of children and adolescents with RAP have revealed that many (25–50%) of these patients continue to experience somatic complaints as adults (Apley & Hale, 1972; Christensen & Mortensen, 1975; Liebman, 1978; Stickler & Murphy, 1979; Stone & Barbero, 1970; Walker, Garber, Van Slyke, & Greene, 1995). In a five-year follow-up study, Walker et al. (1995) found that children and adolescents originally diagnosed with RAP reported significantly higher levels of abdominal discomfort, other somatic symptoms, and functional disability than formerly healthy controls. Moreover, the former RAP patients had higher levels of internalizing emotional symptoms (e.g., anxiety and depression) and significantly higher levels of mental health service utilization than the healthy comparison group. With regard to pseudoneurologic or conversion symptoms, studies have found that over 50% of patients with such symptoms tend to show significant clinical improvement or recovery, although a subset of children with conversion disorders go on to develop somatization disorder or other psychiatric problems including anxiety and personality disorders (Bangash, Worley, & Kandt, 1988; Goodyer, 1981; Goodyer & Mitchell, 1989; Grattan-Smith et al., 1988; Kotsopolous & Snow, 1986; Lehmkuhl et al., 1989; Leslie, 1988; Maisami & Freeman, 1987; Proctor, 1958; Schneider & Rice, 1979; Spierings et al., 1990; Turgay, 1990; Wyllie et al., 1991). Kotsopoulos and Snow (1986) reported that 40% of 20 children with conversion disorder continued to have somatic or psychiatric symptoms 1–3 years after the initial assessment.

There is also evidence that some children originally diagnosed as having a somatoform disorder (e.g., conversion disorder) were later discovered to have had or to have developed a medical condition that explained their earlier symptoms (Robins & O'Neil, 1953; Woodbury, De Maso, & Goldman, 1992). Some researchers have argued, however, that early subclinical impairments in medical disorders, such as multiple sclerosis or tumors of the nervous system, might trigger a vulnerable person to develop conversion symptoms (Caplan & Nadelson, 1980; Nemzer, 1996; Rivinus, Jamison, & Graham, 1975). Thus, careful follow-up studies are needed to examine the continuity of somatic symptoms over time and to explore the extent to which diagnoses of somatoform symptoms and disorders were originally or subsequently develop into actual medical conditions.

5.25.10 CONCEPTUALIZATION AND CLINICAL FORMULATION

5.25.10.1 Assessment

There are few standard methods and measures for assessing somatoform disorders in children and adolescents. Comparisons across studies are difficult because different criteria and different instruments containing different lists of symptoms have been used. Assessment and diagnosis of somatoform disorders is especially challenging because independent medical assessments are needed to determine whether the reported physical symptoms are the result of identifiable organic disease and, if there is organic disease, whether there is functional impairment beyond what would be expected. Furthermore, the diagnosis of somatoform disorders is not simply a matter of the absence of medical findings. Rather, clinicians need to conduct a careful assessment to determine whether the patient meets the various symptom criteria for a somatoform disorder as well as for any other psychiatric diagnoses.

Another difficulty in assessing and diagnosing childhood somatoform disorders is that, in contrast to children with most other kinds of psychiatric problems, children with somatoform disorders are typically brought to pediatricians and general practitioners rather than to psychiatrists or psychologists. It is critical that physicians lay the groundwork for possible psychosocial intervention by explaining to the child and family that the routine comprehensive evaluation involves assessing both physical and psychological factors. Presented in this way, physicians can reduce resistance by the patient and family to inquiries about emotional issues that they might view to be private and irrelevant to the medical condition for which they are seeking help.

In addition to conducting a physical examination and obtaining a complete medical history, the pediatrician or family practitioner also should get a psychosocial history, an assessment of any recent stressors, and conduct a thorough mental status examination. Physicians are obviously quite adept at performing the necessary medical evaluations to detect physical illness, however, they often miss the psychiatric symptoms in children (Costello & Edelbrock, 1985). This is partially due to their lack of experience with the methods for evaluating psychiatric symptoms in children. Therefore, it is important for physicians to have screening measures that they can use to identify those patients for whom referral for further psychiatric evaluation might be appropriate.

Several commonly used psychological instruments are available for assessing somatic complaints as well as other emotional and behavioral problems in children. The Child Behavior Checklist (CBCL; Achenbach & Edelbrock, 1983) has versions for parents, teachers, and children to complete. A somatic complaints subscale can be derived from the CBCL. The Personality Inventory for Children (PIC; Wirt, Lachar, Klinedinst, & Seat, 1984), which is a personality questionnaire completed by parents about their children, also contains a somatic complaints subscale. Neither the CBCL nor the PIC, however, can be used to diagnose somatoform disorders (Prichard, Ball, Culbert, & Faust, 1988). This is largely because the extent to which somatic symptoms are the result of an actual organic problem requires an independent medical evaluation. There are also measures that assess somatic complaints in particular. The CSI (Garber et al., 1991; Walker et al., 1991; Walker & Greene, 1989) was developed explicitly to assess symptoms of somatization disorder in children and adolescents. The CSI includes symptoms from the *DSM-III-R* (APA, 1987) criteria for somatization disorder and from the somatization factor of the Hopkins Symptom Checklist (Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974). Children rate the extent to which they have experienced each of 35 symptoms in the last two weeks using a five-point scale ranging from 0 = not at all to 4 = a whole lot. A total score, obtained by summing the ratings, can range from zero to 140. Both test–retest reliability and internal consistency for the CSI have been found to be quite good (Walker et al.). There also is a parent form of the Children's Somatization Inventory (P-CSI) that includes the same symptoms as the CSI and is completed by parents with reference to their children.

A similar questionnaire was developed by Benjamin and Eminson (1992) that measured lifetime prevalence of symptoms in 11–16-year-old British children. They found that somatic and emotional symptoms were correlated, particularly in girls. About 20% of the children reported worries about health that affected their school attendance.

The Ontario Child Health Study used a 17-item list containing questions about the experience of bodily distress, worries and concerns about health, and the experience of having lost faculties for no apparent reason. Items were checked either 0 = never or not true, 1 = sometimes true, or 2 = often or very true. A total "somasick" score can be calculated by adding items marked one or two, yielding a possible total score of 34. Psychometric properties of this scale have not been reported (Taylor et al., 1996).

Thus, when physicians suspect the possibility of a somatoform disorder in children, their assessment should include a thorough physical exam to rule out medical explanations for the child's symptoms, general measures of child behavioral and emotional symptoms (e.g., CBCL), more specific measures of the most likely emotional problems including anxiety and depression (e.g., Children's Depression Inventory, Kovacs, 1980/1981; Revised Children's Manifest Anxiety Scale, RCMAS, Reynolds & Richmond, 1985), specific measures of somatic complaints (e.g., CSI), and an assessment of the extent of impairment and functional disability due to the somatic symptoms (e.g., Functional Disability Inventory, FDI; Walker & Greene, 1991b). Although the physician will still not be able to make a diagnosis of a somatoform disorder on the basis of these instruments, these measures will provide important information to guide decisions about the possible need for further medical and psychological evaluation.

It also is generally considered to be a good idea to obtain information from multiple sources, although there are known problems with lack of agreement across informants (Achenbach, McConaughy, & Howell, 1987). Who is the best source of information regarding children's somatic problems? Because somatic symptoms are primarily subjective, it is crucial to ask the children themselves. Parents and teachers can be useful informants, however, about the onset and duration of children's physical complaints and the extent of impairment associated with these complaints, such as not playing with friends or not doing school work.

Results of studies examining the relation between parent and child report about somatic complaints have differed as a function of the sample used. Taylor et al. (1996) explored in a community sample the extent of agreement between adolescents and their parents' reports about the teens' somatic symptoms using the 17-item questionnaire from the Ontario Child Health Study. They found a high correlation between parents and youth ($r = 0.80$), but little agreement on the endorsement of specific symptoms (all kappas were less than 0.30).

Garber, Van Slyke, and Walker (1997) studied mother–child concordance regarding children's somatic and emotional symptoms in children diagnosed with either a physical, emotional, or no disorder. Whereas well children and their mothers showed a high level of concordance, mothers of children with recurrent abdominal pain reported more child somatic and emotional symptoms than their children. Discordance was in the direction of mothers reporting more symptoms in their children than the children themselves, and this was associated with higher levels of maternal distress. Thus, obtaining reports from multiple sources can provide useful information from different perspectives, although the level of concordance is likely to be low.

Finally, given the presumably unconscious nature of conversion symptoms, they are likely to be particularly difficult to assess. The Dissociative Experiences Scale is a self-rating scale that has been used to measure dissociative symptoms in adults (Bernstein & Putnam, 1986) and children (Putnam, Helmers, & Trickett, in press). When other diagnostic procedures have been exhausted, sodium amytal interviews are sometimes tried. Weller, Weller, and Fristad (1985) described several case studies in which they used amytal interviews with children to remove conversion symptoms and to differentiate between feigned and actual symptoms. Weller et al. (1985) concluded that a properly administered amytal interview can be a safe and effective diagnostic and therapeutic tool with children.

An interesting technique for diagnosing pseudoseizures is video-electroencephalographic monitoring. During a seizure, the patient is simultaneously videotaped and the EEG recorded. If there is no electrical brain activity, the seizure is assumed to be a conversion symptom (Nemzer, 1996).

Thus, the assessment of somatoform disorders in children and adolescents should involve multiple methods including child self-report inventories, parent questionnaires, psychiatric interviews, medical evaluations, family histories, and amytal interviews when appropriate. More research needs to be conducted to determine the best way to integrate the information obtained from these various methods and sources.

5.25.10.2 Treatment Formulation

Because patients with somatoform disorders tend to think of their symptoms as resulting from medical rather than psychological causes, they can be particularly reluctant to seek or accept psychological explanations for their symptoms, and they may be resistant to psychosocial interventions. Therefore, it is important for health-care workers to acknowledge the reality of the patient's physical distress and also to provide reassurance that the symptoms are not severe or life-threatening (Kellner, 1991; Maisami & Freeman, 1987), although too much reassurance should be avoided (Warwick & Salkovskis, 1985).

Routine use of an interdisciplinary team comprised of physicians, nurses, psychologists, and social workers can be especially helpful because it communicates to the patient and family that all aspects of the child's health are being considered, and it can provide a coordinated management approach (Mayou, Bass, & Sharpe, 1995). The treatment team can explain to patients that their physical problems can also have emotional consequences. This is sometimes a good way of introducing the need for psychological assessment and possible psychosocial intervention. At least at first, it is probably not necessary for patients with somatoform disorders to believe that psychological factors played a role in the development or maintenance of their symptoms. After some rapport has been developed with the patient, however, the possible contribution of psychological factors to their symptoms can and should be discussed. Moreover, the diagnosis of a somatoform disorder can be presented as "good news" if it means the absence of a serious medical condition. When a referral is made for mental health services, the child patient and family should be reassured that the pediatrician or family practitioner will continue to be involved in their treatment. Such close liaison can help minimize any sense of abandonment or rejection they might experience.

With regard to the treatment of conversion symptoms, in particular, there is an old controversy among clinicians as to whether it is first necessary to remove the physical symptom and then conduct psychotherapy to deal with the underlying psychopathology, or whether the psychotherapeutic intervention will simultaneously address the symptom and its causes. Various procedures used for removing the symptom include suggestion (Proctor, 1958; Rock, 1971), encouragement (Gold, 1965), and chemical abreaction through the use of medications such as amobarbital (Laybourne & Churchill, 1972). More recent perspectives do not recommend withholding psychotherapy until the symptom is removed. Rather, clinicians now suggest helping the patient and family recognize that the symptom is not life-threatening in order to reduce their anxiety, and thereby make them more open to other treatments (Kellner, 1986; Maisami & Freeman, 1987; Schulman, 1988).

An important goal of any intervention with children and adolescents experiencing somatoform symptoms and disorders is to encourage their return to normal activities and to discourage "sick role" behaviors such as missing school, staying in bed, and avoiding friends. This rehabilitative approach involves a reframing of the child's symptoms and has two important features. First, it emphasizes that patients do not need to wait until their symptom(s) have improved for them to engage in their normal academic and social activities. This teaches them that they can keep going even when they are experiencing symptoms and discomfort. Patients with pseudoneurologic symptoms, particularly motor problems, are sometimes encouraged to do prescribed exercises or to see a physical therapist who can assist them with gradually returning to normal functioning (Dubowitz & Hersov, 1976) and with giving up their symptom while also "saving face" (Bolton & Cohen, 1986).

Second, this approach shifts the responsibility for "cure" from the physician to the patient; that is, the child changes from being a passive recipient of care to becoming an active collaborator in his or her recovery. Rather than looking for attention and rewards for sick role behaviors, the child seeks reinforcement and praise for his or her efforts at overcoming the symptoms and persevering in spite of them.

Another method sometimes used to reduce sick role behaviors is to arrange regular appointments with the physician independent of the child's physical condition. This provides children and their parents with opportunities to meet with their doctor without the child having to be symptomatic. In addition, the physician can use these appointments to reassess whether any medical conditions have developed that could account for their somatoform symptoms.

One additional complication in the treatment of children and adolescents with somatoform disorders is that not only does the child patient need to be convinced of the role of psychological factors, but typically the parents also need to be persuaded. There is often a shared belief among family members of the medical origins of the child's symptoms. Parents often are the primary reinforcers of the child's sick role behaviors. Therefore, interventions need to target the beliefs and behaviors of parents as well as the children.

Family perspectives about psychosomatic conditions (Minuchin et al., 1975, 1978; Mullins & Olson, 1990) suggest that the child's somatic symptoms serve a function within the family system. Family therapists argue that interventions should not focus on the identified child's symptoms because that distracts the family from the more important family processes that are maintaining the symptom. Minuchin and colleagues recommend that family therapy targets the enmeshment, parental overprotection, and conflict avoidance that often characterize these families. Thus, the basic principles for treating somatoform disorders include acknowledging the patient's physical symptoms while at the same time explaining to both the child and parents the possible contribution of psychological factors to the onset and maintenance of the child's symptoms. A multidisciplinary approach that involves collaboration between the physician and mental health workers is recommended. A primary goal of any intervention is the minimization of sick role behaviors and encouragement of a return to normal activities as soon as possible, although for some children this may need to be a gradual process.

5.25.11 TREATMENT

5.25.11.1 Psychosocial

Psychosocial approaches to the treatment of somatoform disorders in children and adolescents have generally included behavioral and cognitive techniques, psychodynamic psychotherapy, and group and family therapies. The particular technique used depends on the specific symptoms and disorder. Relaxation training (Larsson, 1992; Larsson & Mellin, 1988), hypnosis (Elkins & Carter, 1986), and biofeedback (Klonoff & Moore, 1986; Mizes, 1985) have been used to reduce tension associated with headaches and other pains. Cognitive interventions teach self-monitoring and coping strategies (Sanders, Shepherd, Cleghorn, & Woolford, 1994). Behavioral methods are used to reinforce mastery over symptoms and to reduce secondary gain associated with sick role behaviors (Delameter, Rosenbloom, Conners, & Hertweck, 1983; Mizes, 1985). Psychodynamic approaches aim to help children with conversion disorders to gain insight into unconscious conflicts and to identify how psychological factors can influence the maintenance of symptoms (Goodyer & Taylor, 1985). Some form of parent training, family education, or family therapy can be used to apprise parents of their possible role in the onset and maintenance of their child's somatic symptoms and to improve communication among family members (Mullins & Olson, 1990). For those children with somatoform disorders who are shy, unassertive, and have difficulty expressing negative emotions, interventions that also improve self-esteem, promote assertiveness, and train in emotion expression and regulation techniques can be helpful (Nemzer, 1996).

Reattribution techniques that involve simply making the link for the child and parents between the occurrence of a stressor and the onset of symptoms can sometimes lead to symptom improvement. When such insight is followed by targeted problem-solving such as making appropriate environmental and behavioral changes, symptom reduction may happen rapidly (Wilkinson & Mynors-Wallis, 1994). For example, interventions such as changing a child's classroom or having a conference with the teacher to address the child's specific concerns can eliminate somatic symptoms being used to avoid school.

Although there are many interesting case studies in the literature using a variety of psychosocial approaches for the treatment of specific patients (e.g., Andrasik, Burke, Attanasio, & Rosenblum, 1985; Delameter et al., 1983; Elkins & Carter, 1986; Linton, 1986; Miller & Kratochwill, 1979), there are relatively few controlled trials for the treatment of somatoform disorders in children. These studies have targeted specific somatic complaints, particularly recurrent abdominal pain (Finney, Lemanek, Cataldo, Katz, & Fuqua, 1989; Sanders et al., 1989, 1994) and headaches (Larsson & Mellin, 1988; McGrath et al., 1988; Osterhaus et al., 1993).

Finney and colleagues (Edwards, Finney, & Bonner, 1991; Finney et al., 1989) have examined the efficacy of using several different techniques including self-monitoring, limited parent attention, relaxation training, increased dietary fiber, and required school attendance. Finney et al. found that a group of 16 children who received a brief multicomponent targeted intervention comprising all of these methods had improved pain symptoms and decreased school absences compared to the untreated comparison group ($n = 16$). In addition, the treated group had significantly decreased medical care utilization.

Sanders and colleagues (Sanders et al., 1989, 1994) have conducted two studies examining the efficacy of cognitive-behavioral treatment of recurrent abdominal pain in children. In the first study (Sanders et al., 1989), the treatment program consisted of differential reinforcement of well behavior, cognitive coping skills training, and various generalization enhancement procedures. Although both the treatment and control groups showed a reduction in pain symptoms, the cognitive-behavioral treatment group had a more rapid improvement that generalized to the school setting, and a larger proportion of the treated children were completely pain-free at the three month follow-up.

Their second study (Sanders et al., 1994) compared the efficacy of a cognitive-behavioral family intervention (CBFI) with standard pediatric care (SPC) in 44 children with recurrent abdominal pain aged 7–14 years. Again, children in both treatment conditions showed improvements on pain intensity and pain behavior. The children in the CBFI group, however, also had a higher rate of complete elimination of pain, lower levels of relapse, and lower levels of interference with their activities. Thus, these studies show that some form of intervention typically leads to symptom improvement, although the more extensive treatments have broader and possibly longer-lasting effects on symptoms and functioning.

Efficacy of specific treatments for headaches in children and adolescents has been mixed. Relaxation training has been found to be superior to attention placebo control for the treatment of tension headaches (Larsson & Mellin, 1988) and migraines in adolescents (Larsson, 1992). Osterhaus et al. (1993) found that a combined treatment consisting of relaxation training, temperature biofeedback, and cognitive techniques reduced headache frequency and duration, although not intensity, compared with a waiting list control. Osterhaus et al. suggested that an increase in the ability to relax and a decrease in anxiety contributed to improvement. Helm-Hylkema, Orlebek, Enting, Thijssen, and Ree (1990) used the same training program as that of Osterhaus et al. but in a clinical setting and found that 90% of their young adolescent patients showed significant improvement. McGrath et al. (1988), however, reported that among adolescents attending a pediatric neurology clinic for migraines, relaxation was not found to be more effective than a "placebo" discussion of psychological topics. It could be that their control condition, which involved receiving reassurance and assistance, and attending to psychological factors, was sufficiently effective in reducing migraines among these adolescents. Thus, relaxation procedures show some promise in reducing headaches in children and adolescents, although further research needs to be conducted exploring what might be the effective ingredients of "control" interventions. Although much has been written about how to deal therapeutically with conversion disorders (e.g., Hodgman, 1995; Maisami & Freeman, 1987), there have been few controlled treatment studies in either adults or children. Adult patients who received psychotherapy for pseudoseizures have been found to be significantly improved compared to those who did not undergo treatment (Lesser, Lueders, & Dinner, 1983; Meierkord, Will, Fish, & Shorvon, 1991).

5.25.11.2 Medical Interventions

Treating somatoform disorders with medical interventions such as pharmacology and hospitalization is a complex issue because of the message it sends to the child and family. One concern about using medication is that this might perpetuate the patients' beliefs about the organic etiology of their symptoms. Indeed, some physicians resort to prescribing placebos to avoid the patient's anger about not receiving medical treatment. This could have the negative effect, however, of perpetuating the child's view of him or herself as sickly and in need of medical intervention.

On the other hand, providing some medical intervention helps to validate the reality of patients' physical distress and thereby might increase their willingness to accept other forms of treatment as well; that is, if the physician does not simply dismiss patients' symptoms as being "all in their head," then patients might be more open to a multimodal approach. Moreover, some medical interventions, even if they are over-the-counter remedies, can at least alleviate some of the physical discomfort they are experiencing.

In addition, if a child with a somatoform disorder is experiencing comorbid symptoms such as anxiety or depression, then psychopharmacological treatment aimed at reducing these symptoms might be warranted. If children's somatic symptoms are reduced as a function of psychotropic medication, then this might help to convince them and their parents that psychological factors may be contributing to their physical problems. In adults, antidepressants have been found to reduce somatic symptoms in depressed patients, and anxiolytic medications appear to alleviate somatic symptoms in patients diagnosed with anxiety disorders, such as panic (Kellner, 1991). There have been no controlled treatment studies of the use of antidepressants or antianxiety medications for children or adolescents with somatoform disorders with comorbid emotional disorders. There is some evidence of the efficacy of selective serotonin reuptake inhibitors (SSRIs) for the treatment of adolescent depression (Emslie, Rush, & Weinberg, in press) and pediatric obsessive compulsive disorder (Riddle et al., 1992), and for the use of medications in the treatment of childhood anxiety disorders (Allen, Leonard, & Swedo, 1995). Given the high rates of anxiety and depressive disorders among children with some forms of recurrent somatic complaints (e.g., Garber et al., 1990; Hodges et al., 1985a; 1985b), such a treatment study might be informative.

Sometimes the patient's lack of improvement with medical treatment can assist in the diagnosis of a somatoform disorder; that is, if the symptoms do not get better with a medication known to be effective for a particular medical condition, then it is less likely that the patient has that condition. Thus, prescribing medication can serve as a diagnostic test as well as a form of treatment.

Are there any conditions under which hospitalization of children with somatoform disorders would be warranted? If the use of outpatient psychotherapy and pharmacological interventions have been found to be ineffective, then pediatric or psychiatric hospitalization may be appropriate. Inpatient hospitalization has been suggested when there is diagnostic uncertainty or marked functional impairment, such as persistent school refusal (Goodyer, 1981; Leslie, 1988). Such treatment can be accomplished best with a multidisciplinary team on a pediatric medical psychiatry inpatient unit (Campo & Raney, 1995). Hospitalization provides an opportunity for careful observation and control of environmental contingencies. Moreover, removal from the environmental conditions that might be producing or exacerbating the symptoms can be informative and therapeutic.

5.25.12 CONCLUSION AND FUTURE DIRECTIONS

The diagnostic criteria that define somatoform disorders have changed with each new edition of the *DSM*. The validity of these criteria for use with children still needs to be investigated. Epidemiological studies are needed that assess the full range of symptoms that comprise these disorders and that test the validity of different symptom cut-offs and combinations.

One of the real challenges for diagnosing somatoform disorders is the necessity of conducting comprehensive medical as well as psychiatric evaluations. Multidisciplinary teams that include medical and psychiatric staff are required for the assessment and treatment of patients with somatoform disorders. These evaluations should include information from multiple sources (i.e., children, parents, teachers) and should use multiple assessment methods including questionnaires, psychiatric interviews, observation, and amytal interviews when relevant. Similarly, interventions typically will require the cooperation of physicians and mental health specialists. Successful treatment is likely to involve individual therapy with the child as well as family therapy that addresses the system maintaining the child's symptoms.

Somatoform disorders are a heterogeneous category of loosely associated syndromes. They share the feature of physical symptoms that are not explained simply by a medical condition. It is likely that multiple factors comprising a biopsychosocial model contribute to the etiology of somatoform disorders. Moreover, it also is likely that a unique, although somewhat overlapping, set of processes underlie each disorder.

There also are age, sex, and cultural differences in somatic complaints and somatoform disorders that need to be studied further. How might the observed sex differences in the rates of somatoform disorder inform us about its etiology? Similarly, how are the different rates and patterns of somatic complaints in different cultures and subcultures explained?

Additional issues for future research concern the comorbidity of somatoform disorders with other internalizing problems. To what extent do other emotional disorders such as anxiety and depression underlie the somatic symptoms that comprise somatoform disorders? What is the extent of somatic symptoms in children with other emotional disorders (McCauley, Carlson, & Calderon, 1991)? How do children diagnosed with emotional disorders with and without somatic symptoms differ with regard to their prognosis, etiology, and treatment response?

Somatoform disorders are important because they drain society's economic and medical resources, strain the family environment, and hinder children's social and academic development. Further research is needed that addresses the assessment, etiology, prognosis, and treatment of each of the different types of somatoform disorders across the lifespan. Early interventions that target high-risk populations can help reduce the cost of these disorders to both individuals and society.

ACKNOWLEDGMENT

This work was supported in part by a grant from the National Institute of Child Health & Human Development (RO 1 HD 23264).

5.25.13 REFERENCES

Achenbach, T. M., Conners, C. K., Quay, H. C., Verhulst, F. C., & Howell, C. T. (1989). Replication of empirically derived syndromes as a basis for taxonomy of child/adolescent psychopathology. *Journal of Abnormal Child Psychology*, *17*, 299–323.

Achenbach, T. M., & Edelbrock, C. S. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, VT: University of Vermont.

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213–232.

Alfven, G. (1993). The pressure pain threshold (PPT) of certain children suffering from recurrent abdominal pain of non-organic origin. *Acta Paediatrica*, *82*, 481–483.

Allen, A. J., Leonard, H., & Swedo, S. E. (1995). Current knowledge of medications for the treatment of childhood anxiety disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *34*, 976–986.

American Psychiatric Association (1952). *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Press.

American Psychiatric Association (1968). *Diagnostic and statistical manual of mental disorders* (2nd ed.). Washington, DC: American Psychiatric Press.

American Psychiatric Association (1978). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: American Psychiatric Press.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. Rev.). Washington, DC: American Psychiatric Press.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Press.

Andrasik, F., Burke, E. J., Attanasio, V., Rosenblum, E. L. (1985). Child, parent and physician reports of a child's headache pain: Relationships prior to and following treatment. *Headache*, *25*, 421–425.

Apley, J. (1958). A common denominator in the recurrent pains of childhood. *Proceedings of the Royal Society of Medicine*, *51*, 1023–1024.

Apley, J. (1975). *The child with abdominal pain*. Oxford: Blackwell.

Apley, J., & Hale, B. (1973). Children with recurrent abdominal pain: How do they grow up? *British Medical Journal*, *3*, 7–9.

Apley, J., & Naish, N. (1958). Recurrent abdominal pains: A field survey of 1000 school children. *Archives of Diseases of Children*, *33*, 165–170.

Aro, H. (1987). Life stress and psychosomatic symptoms among 14 to 16-year-old Finnish adolescents. *Psychological Medicine*, *17*, 191–201.

Aro, H., Paronen, O., & Aro, S. (1987). Psychosomatic symptoms among 14–16-year old Finnish adolescents. *Social Psychiatry*, *22*, 171–176.

Aro, H., & Taipale, V. (1987). The impact of timing of puberty on psychosomatic symptoms among fourteen to sixteen year-old Finnish girls. *Child Development*, *58*, 261–268.

Bagby, R. M., Taylor, G. J., & Parker, D. A. (1994). The twenty-item Toronto Alexithymia Scale. *Journal of Psychosomatic Research*, *38*, 23–40.

Bangash, I. H., Worley, G., & Kandt, R. S. (1988). Hysterical conversion reactions mimicking neurological disease. *American Journal of Diseases of Childhood*, *142*, 1203–1206.

Barsky, A. J., Goodson, J. D., Lane, R. S., & Cleary, P. D. (1988). The amplification of somatic symptoms. *Psychosomatic Medicine*, *50*, 510–519.

Barsky, A. J., & Klerman, G. (1983). Overview: Hypochondriasis, bodily complaints, and somatic styles. *American Journal of Psychiatry*, *140*, 273–283.

Barsky, A. J., Wyshak, G., & Klerman, G. L., (1986). Psychiatric comorbidity in *DSM-III-R* hypochondriasis. *Archives of Senses Psychiatry*, *4*, 101–108.

Bass, C., & Murphy, M. (1995). Somatoform and personality disorders: Syndromal comorbidity and overlapping developmental pathways. *Journal of Psychosomatic Research*, *39*, 403–427.

Beautrais, A., Fergusson, D., & Shannon, F. (1982). Life events and childhood morbidity: A prospective study. *Pediatrics*, *70*, 935–940.

Belle, D., Burr, R., & Cooney, J. (1987). Boys and girls as social support theorists. *Sex Roles*, *17*, 657–665.

Belmaker, E. (1984). Nonspecific somatic symptoms in early adolescent girls. *Journal of Adolescent Health*, *5*, 30–33.

Belmaker, E., Espinoza, R., & Pogrund, R. (1985). Use of medical services by adolescents with non-specific somatic symptoms. *International Journal of Adolescent Medicine and Health*, *1*, 150–156.

Benjamin, S., & Eminson, D. M. (1992). Abnormal illness behavior: Childhood experiences and long term consequences. *International Review of Psychiatry*, *4*, 55–70.

Bernstein, E., & Putnam, F. W. (1986). Development, reliability and validity of a dissociation scale. *Journal of Nervous and Mental Disorders*, *174*, 727–735.

Bird, H. R., Canino, G., Rubio-Stipec, M., Gould, M. S., Ribera, J., Sesman, M., Woodbury, M., Huertas-Goldman, S., Pagan, A., Sanchez-Lacy, A., & Moscoso, M. (1988). Estimates of the prevalence of childhood maladjustment in a community survey in Puerto Rico. *Archives of General Psychiatry*, *45*, 1120–1126.

Bohman, M., Cloninger, C. R., von Knorring, A. L., Sigvardsson, S. (1984). An adoption study of somatoform disorders: III. Cross-fostering analysis and genetic relationship to alcoholism and criminality. *Archives of General Psychiatry*, *41*, 872–878.

Bolton, J., & Cohen, P. (1986). "Escape with honour": The need for face-saving. *Bulletin of Anna Freud Centre*, *9*, 19–33.

Bourke, M. P., Taylor, G. J., Parker, J. D. A., & Bagby, R. M. (1992). Alexithymia in women with anorexia nervosa: A preliminary investigation. *British Journal of Psychiatry*, *161*, 240–243.

Bowman, E. S. (1993). Etiology and clinical course of pseudoseizures: Relationship to trauma, depression and dissociation. *Psychosomatics*, *34*, 333–342.

Boyce, T. W., Jensen, E. W., Cassel, J. C., Collier, A. M., Smith, A. H., & Ramsey, C. T. (1977). Influence of life events and family routines on childhood respiratory-tract illness. *Pediatrics*, *60*, 609–615.

Briquet, P. (1859). *Traite de Hysterie*. Paris: J. B. Balliaere and Fils.

Cadoret, R. J. (1978). Psychopathology in adopted-away offspring of biologic parents with antisocial behavior. *Archives of General Psychiatry*, *35*, 176–184.

Campo, J. V., & Fritsch, S. L. (1994). Somatization in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *33*, 1223–1235.

Campo, J. V., & Garber, J. (in press). Somatization. In R. T. Ammerman & J. V. Campo (Eds.), *Handbook of pediatric psychology and psychiatry*. Boston: Allyn & Bacon.

Campo, J. V., & Raney, D. (1995). The pediatric medical-psychiatric unit in a psychiatric hospital. *Psychosomatics*, *36*, 438–444.

Caplan, L. R., & Nadelson, T. (1980). Multiple sclerosis and hysteria: Lessons learned from their association. *Journal of the American Medical Association*, *243*, 2418–2421.

Chase, H. P., & Jackson, G. G. (1981). Stress and sugar control in children with insulin-dependent diabetes mellitus. *Journal of Pediatrics*, *98*, 1011–1013.

Christensen, M. F., & Mortensen, O. (1975). Long-term prognosis in children with recurrent abdominal pain. *Archives of Diseases of Childhood*, *50*, 110.

Cloninger, C. R. (1986). *Somatoform and dissociative disorders*. Philadelphia: Saunders.

Cloninger, C. R., Sigvardsson, S., von Knorring, A. L., Bohman, M. (1984). An adoption study of somatoform disorders. *Archives of General Psychiatry*, *41*, 863–871.

Clouse, R. E., & Lustman, P. J. (1983). Psychiatric illness and contraction abnormalities of the esophagus. *New England Journal of Medicine*, *309*, 1337–1342.

Cohen, K., Auld, F., & Brooker, H. (1994). Is alexithymia related to psychosomatic disorder and somatising? *Journal of Psychosomatic Research*, *38*, 119–127.

Costello, A., Edelbrock, C. S., Kalas, R., Kessler, M. K., & Klaric, S. (1982). *National Institute of Mental Health Diagnostic Interview Schedule of Children*. Bethesda, MD: National Institutes of Health.

Costello, E. J., & Edelbrock, C. S. (1985). Detection of psychiatric disorders in pediatric primary care: A preliminary report. *Journal of the American Academy of Child Psychiatry*, *24*, 771–774.

Costello, E. J., Edelbrock, C., Costello, A. J., Dulcan, M. K., Burns, B. J., & Brent, D. (1988). Psychopathology in pediatric primary care: The hidden morbidity. *Pediatrics*, *82*, 415–424.

Craig, K. D. (1978). Social modeling influences on pain. In R. A. Sternbach (Ed.), *The psychology of pain* (pp. 73–109). New York: Raven Press.

Craig, T. K. J., Boardman, A. P., Mills, K., Daly-Jones, O., & Drake, H. (1993). The South London Somatisation Study I: Longitudinal course and the influence of early life experiences. *British Journal of Psychiatry*, *163*, 579–588.

Cunningham, S. J., McGrath, P. J., Ferguson, H. B., Humphreys, P., D'Astrous, J., Latter, J., Goodman, J. T., & Firestone, P. (1987). Personality and behavioral characteristics in pediatric migraine. *Headache*, *27*, 16–20.

Delameter, A. M., Rosenbloom, N., Conners, K., & Hertweck, L. (1983). The behavioral treatment of hysterical paralysis in a ten-year-old boy: A case study. *Journal of the American Academy of Child Psychiatry*, *1*, 73–79.

Derogatis, L. R., Lipman, R. S., Rickels, K. Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins Symptom Checklist. *Behavioral Science*, *19*, 1–15.

Dolce, J. J., & Raczynski, J. M. (1985). Neuromuscular activity and electromyography in painful backs: Psychological and biomechanical models in assessment and treatment. *Psychological Bulletin*, *97*, 502–520.

Dubowitz, V., & Hersov, L. (1976). Management of children with nonorganic (hysterical) disorders of motor function. *Developmental Medicine and Child Neurology*, *18*, 358–368.

Dunn-Geier, J., McGrath, P. J., Rourke, B. P., Latter, J., & D'Astous, D. (1986). Adolescent chronic pain: The ability to cope. *Pain*, *26*, 23–32.

Edwards, M. C., Finney, J. W., & Bonner, M. (1991). Matching treatment with recurrent abdominal pain symptoms: An evaluation of dietary fiber and relaxation treatments. *Behavior Therapy*, *22*, 257–267.

Edwards, P. W., Zeichner, A., Kuczmierczyk, A. R., & Boczkowski, J. (1985). Familial pain models: The relationship between family history of pain and current pain experience. *Pain*, *21*, 379–384.

Elkins, G. R., & Carter, B. D. (1986). Hypnotherapy in the treatment of childhood psychogenic coughing: A case report. *American Journal of Clinical Hypnosis*, *29*, 59–63.

Emslie, G., Rush, A. J., & Weinberg, A. W. (in press). A double-blind randomized placebo-controlled trial of fluoxetine in depressed children. *Archives of General Psychiatry*.

Engel, G. L. (1962). *Psychological development in health and disease*. Philadelphia: Saunders.

Ernst, A. R., Routh, D. K., & Harper, D. C. (1984). Abdominal pain in children and symptoms of somatization disorder. *Journal of Pediatric Psychology*, *9*, 77–86.

Escobar, J. I., Burnam, A., & Karno, M. (1987). Somatization in the community. *Archives of General Psychiatry*, *44*, 713–720.

Feuerstein, M., Barr, R., Francoeur, E., Houle, M., & Rafman, S. (1982). Potential biobehavioral mechanisms of recurrent abdominal pain in children. *Pain*, *13*, 287–298.

Finney, J. W., Lemanek, K. L., Cataldo, M. F., Katz, H. P., & Fuqua, R. W. (1989). Pediatric psychology in primary healthcare: Brief targeted therapy for recurrent abdominal pain. *Behavior Therapy*, *20*, 283–291.

Flor, H., Turk, D. C., & Birbaumer, N. (1985). Assessment of stress-related psychophysiological reactions in chronic back pain patients. *Journal of Consulting and Clinical Psychology*, *53*, 354–364.

Fordyce, W. E. (1978). Learning processes in pain. In R. A. Sternbach (Ed.), *The psychology of pain* (pp. 49–72). New York: Raven Press.

Freud, S. (1896/1962). The ethiology of hysteria. In *Complete psychological works of Sigmund Freud* (pp. 191–221). London: Hogarth.

Fuchs, D., & Thelen, M. H. (1988). Children's expected interpersonal consequences of communicating their affective state and reported likelihood of expression. *Child Development*, *59*, 1314–1322.

Fujii, S., Kachi, T., & Sobue, I. (1981). Chronic headache: Its psychosomatic aspect. *Japanese Journal of Psychosomatic Medicine*, *21*, 411–419.

Garber, J., & Hollon, S. D. (1991). What can specificity designs say about causality in psychopathology research? *Psychological Bulletin*, *110*, 129–136.

Garber, J., Van Slyke, D., & Walker, L. S. (1997). Concordance between mothers' and children's reports of somatic and emotional symptoms in patients with recurrent abdominal pain or psychiatric disorders. Manuscript under review.

Garber, J., Walker, L. S., & Zeman, J. L. (1991). Somatization symptoms in a community sample of children and adolescents: Further validation of the Children's Somatization Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *3*, 588–595.

Garber, J., Zeman, J. L., Walker, L. S. (1990). Recurrent abdominal pain in children and adolescents: Psychiatric diagnoses and parental psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry*, *29*, 648–656.

Garralda, M. E. (1992). A selective review of child psychiatric syndromes with a somatic presentation. *British Journal of Psychiatry*, *161*, 759–773.

Garralda, M. E. (1996). Somatization in children. *British Journal of Psychiatry*, 13–33.

Gold, S. (1965). Diagnosis and management of hysterical contracture in children. *British Medical Journal*, *1*, 21–23.

Goodman, J. E., & McGrath, P. J. (1991). The epidemiology of pain in children and adolescents: A review. *Pain*, *46*, 247–264.

Goodwin, J., Simms, M., & Bergman, R. (1979). Hysterical seizures in four adolescent girls. *American Journal of Orthopsychiatry*, *49*, 698–703.

Goodyer, I. M. (1981). Hysterical conversion reactions in childhood. *Journal of Child Psychology and Psychiatry*, *22*, 179–188.

Goodyer, I. M., & Mitchell, C. (1989). Somatic and emotional disorders in childhood and adolescence. *Journal of Psychosomatic Research*, *33*, 681–688.

Goodyer, I. M., & Taylor, D. C. (1985). Hysteria. *Archives of Diseases of Childhood*, *63*, 680–681.

Gottesman, I. I. (1962). Differential inheritance of the psychoneuroses. *Eugenics Quarterly*, *9*, 223–267.

Grattan-Smith, P., Fairley, M., & Procopis, P. (1988). Clinical features of conversion disorder. *Archives of Diseases of Children*, *63*, 408–414.

Green, A. (1993). Child sexual abuse: Immediate and long-term effects and intervention. *Journal of the American Academy of Child and Adolescent Psychiatry*, *32*, 890–902.

Green, M., & Solnit, A. J. (1964). Reactions to the threatened loss of a child: A vulnerable child syndrome. *Pediatrics*, *34*, 58–66.

Greene, J. W., Walker, L. S., Hickman, G., Thompson, J. (1985). Stressful life events and somatic complaints in adolescents. *Pediatrics*, *75*, 19–22.

Guze, S. B., Cloninger, C. R., Martin, R. L., Clayton, P. J. (1986). A follow-up and family study of Briquet's syndrome. *British Journal of Psychiatry*, *149*, 17–23.

Helm-Hylkema, J. van der, Orlebek, J. F., Enting, L. A., Thijssen, J. H. H., & Ree, J. van (1990). Effects of behavior therapy and plasma B-endorphin in young migraine patients. *Psychoneuroendocrinology*, *15*, 39–45.

Henderson, S. (1974). Care eliciting behavior in man. *Journal of Nervous and Mental Diseases*, *159*, 172–181.

Herman, R. M., & Simonds, I. F. (1975). Incidence of conversion symptoms in children evaluated psychiatrically. *Modern Medicine*, *72*, 597–604.

Hodges, K., Kline, J. J., Barbero, G., & Flanery, R. (1985a). Depressive symptoms in children with recurrent abdominal pain and in their families. *Journal of Pediatrics*, *107*, 622–626.

Hodges, K., Kline, J. J., Barbero, G., & Woodruff, C. (1985b). Anxiety in children with recurrent abdominal pain and their parents. *Psychosomatics*, *26*, 859–866.

Hodgman, C. H. (1995). Conversion and somatization in pediatrics. *Pediatrics in Review*, *16*, 29–34.

Hunter, R. S., Kilstrom, N., & Loda, F. (1985). Sexually abused children: Identifying masked presentations in a medical setting. *Child Abuse and Neglect*, *9*, 17–25.

Hyams, J. S., Burke, G., Davis, P. M., Rzepski, B., & Andrulonis, P. A. (1996). Abdominal pain and irritable bowel syndrome in adolescents: A community-based study. *Journal of Pediatrics*, *129*, 220–226.

Inouye, E. (1972). Genetic aspects of neurosis. *International Journal of Mental Health*, *1*, 176–189.

Kagan, J. Reznick, J. S., & Snidman, N. (1988). Biological bases of childhood shyness. *Science*, *240*, 167–171.

Kendler, K. S., Heath, A. C., Martin, N. G., & Eaves, L. J. (1987). Symptoms of anxiety and symptoms of depression. *Archives of General Psychiatry*, *44*, 451–457.

Kellner, R. (1986). *Somatization and hypochondriasis*. New York: Praeger.

Kellner, R. (1990). Somatization. *The Journal of Nervous and Mental Disease*, *178*, 150–160.

Kellner, R. (1991). *Psychosomatic syndromes and somatic symptoms*. Washington, DC: American Psychiatric Press.

Kirmayer, L. (1984). Culture, affect and somatization. Part I. *Transcultural Psychiatric Research Review*, *21*, 159–188.

Klevan, J. L., & DeJong, A. R. (1990). Urinary tract symptoms and urinary tract infection following sexual abuse. *American Journal of Diseases of Childhood*, *144*, 242–244.

Klonoff, E. A., & Moore, D. J. (1986). "Conversion reactions" in adolescents: A biofeedback-based operant approach. *Journal of Behavioral Therapy and Experimental Psychiatry*, *17*, 179–184.

Kotsopoulos, S., & Snow, B. (1986). Conversion disorders in children: A study of clinical outcome. *Psychiatric Journal of the University of Ottawa*, *11*, 134–139.

Kovacs, M. (1980/1981). Rating scales to assess depression in school-aged children. *Acta Paedopsychiatric*, *46*, 305–315.

Kriechman, A. M. (1987). Siblings with somatoform disorders in childhood and adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry*, *26*, 226–231.

LaBarbera, J. D., & Dozier, J. E. (1980). Hysterical seizures: The role of sexual exploitation. *Psychosomatics*, *21*, 897–903.

Lahey, B. B., Flagg, E. W., Bird, H. R., Schwab-Stone, M. E., Canino, G., Dulcan, M. K., Leaf, P. J., Davies, M., Brogan, D., Bourdon, K., Horwitz, S. M., Rubio-Stipec, M., Freeman, D. H., Lichtman, J. H., Shaffer, D., Goodman, S. H., Narrow, W. E., Weissman, M. M., Kandel, D. B., Jensen, P. S., Richters, J. E., & Regier, D. A. (1996). The NIMH methods for the Epidemiology of Child and Adolescent Mental Disorders (MECA) Study: Background and Methodology. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*, 855–864.

Larson, B. S. (1991). Somatic complaints and their relationship to depressive symptoms in Swedish adolescents. *Journal of Child Psychology and Psychiatry*, *32*, 821–832.

Larsson, B. (1992). Behavioral treatment of somatic disorders in children and adolescents. *European Child and Adolescent Psychiatry*, *1*, 68–81.

Larsson, B., & Mellin, L. (1988). The psychological treatment of recurrent headache in adolescents—short term outcome and its prediction. *Headache*, *28*, 187–195.

Lask, B., & Fosson, A. (1989). *Childhood illness: The psychosomatic approach.* New York: Wiley.

Laybourne, P. C., & Churchill, S. W. (1972). Symptom discouragement in treating hysterical reactions of childhood. *International Journal of Child Psychotherapy*, *1*, 111–123.

Lehmkuhl, G., Blanz, G., Lehmkuhl, U., & Braun-Scharm, H. (1989). Conversion disorder: Symptomatology and course in childhood and adolescence. *European Archives of Psychiatry and Neurological Sciences*, *238*, 155–160.

Leslie, S. A. (1988). Diagnosis and treatment of hysterical conversion reactions. *Archives of Diseases of Children*, *63*, 506–511.

Lesser, R. P., Lueders, H., Dinner, D. S. (1983). Evidence for epilepsy is rare in patients with psychogenic seizures. *Neurology*, *33*, 502–504.

Lewis, C. E., & Lewis, M. A. (1989). Educational outcomes and illness behaviors in participants in a child-initiated care system: A 12-year follow-up study. *Pediatrics*, *84*, 845–850.

Liebman, W. H. (1978). Recurrent abdominal pain in children: A retrospective survey of 119 patients. *Clinical Pediatrics (Philadelphia)*, *17*, 149–153.

Linton, S. J. (1986). A case study of the behavioral treatment of chronic stomach pain in a child. *Behavior Change*, *3*, 70–73.

Lipowski, Z. J. (1988). Somatization: The concept and its clinical application. *American Journal of Psychiatry*, *145*, 1358–1368.

Livingston, R., & Martin-Connici, C. (1985). Multiple somatic complaints and possible somatization disorder in prepubertal children. *Journal of the American Academy of Child Psychiatry*, *24*, 603–607.

Loewenstein, R. J. (1990). Somatoform disorders in victims of incest and child abuse. In R. P. Kluft (Ed.). *Incest related syndromes of adult psychopathology* (pp. 76–107). Washington, DC: American Psychiatric Press.

Looff, D. H. (1970). Psychophysiological and conversion reactions in children. *Journal of the American Academy of Child Psychiatry*, *9*, 318–331.

Maisami, M., & Freeman, J. M. (1987). Conversion reactions in children as body language: A combined child psychiatry/neurology team approach to the management of functional neurologic disorders in children. *Pediatrics*, *80*, 46–52.

Mallouh, S. K., Abbey, S. E., & Gillies, L. A. (1995). The role of loss in treatment outcome of persistent somatization. *General Hospital Psychiatry*, *17*, 187–191.

Maloney, M. J. (1980). Diagnosing hysterical conversion reactions in children. *Journal of Pediatrics*, *97*, 1016–1020.

Mayou, R. (1976). The nature of bodily symptoms. *British Journal of Psychiatry*, *129*, 55–60.

Mayou, R., Bass, C., & Sharpe, M. (1995). *Treatment of functional somatic symptoms.* Oxford: Oxford University Press.

McCauley, E., Carlson, G. A., & Calderon, R. (1991). The role of somatic complaints in the diagnosis of depression in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *30*, 631–635.

McGrath, P. J. (1987). The multidimensional assessment and management of recurrent pain syndromes in children. *Behavior Research and Therapy*, *25*, 251–262.

McGrath, P. J., Goodman, J. T., Firestone, P., Shipman, R., & Peters, S. (1983). Recurrent abdominal pain: A psychogenic disorder? *Archives of Diseases of Childhood*, *58*, 888–890.

McGrath, P. J., Humphreys, P., Goodman, J. T., Keene, D., Firestone, P., Jacob, P., & Cunningham, S. J. (1988). Relaxation prophylaxis for childhood migraine: A randomized placebo-controlled trial. *Developmental Medicine and Child Neurology*, *30*, 626–631.

Mechanic, D. (1972). Social psychologic factors affecting the presentation of bodily complaints. *The New England Journal of Medicine*, *286*, 1132–1139.

Meierkord, H., Will, B., Fish, D., & Shorvon, S. (1991). The clinical features and prognosis of pseudoseizures diagnosed using video-EEG telemetry. *Neurology*, *41*, 1643–1646.

Miller, A. J., & Kratochwill, T. R. (1979). Reduction of frequent stomach complaints by time out. *Behavior Therapy*, *10*, 211–218.

Minuchin, S., Baker, L., Rosman, B. L., Liebman, R., Milman, L., & Todd, T. C. (1975). A conceptual model of psychosomatic illness in children. *Archives of General Psychiatry*, *32*, 1031–1038.

Minuchin, S., Rosman, B. L., & Baker, L. (1978). *Psychosomatic families: Anorexia nervosa in context.* Cambridge, MA: Harvard University Press.

Mizes, J. S. (1985). The use of contingent reinforcement in the treatment of a conversion disorder: A multiple baseline study. *Journal of Behavior Therapy and Experimental Psychiatry*, *16*, 341–345.

Morrison, J. (1989). Childhood sexual histories of women with somatization disorder. *American Journal of Psychiatry*, *146*, 239–241.

Mullins, L. L., & Olson, R. A. (1990). Familial factors in the etiology, maintenance, and treatment of somatoform disorders in children. *Family Systems Medicine*, *8*, 159–175.

Nemiah, J. C. (1977). Alexithymia: Theoretical considerations. *Psychotherapy and Psychosomatics*, *8*, 199–206.

Nemiah, J. C. (1985). Somatoform disorders. In H. I. Kaplan & B. I. Sadock (Eds.), *Comprehensive textbook of psychiatry* (4th ed., pp. 924–942). Baltimore, MD: Williams & Wilkins.

Nemzer, E. D. (1996). Somatoform disorders. In M. Lewis (Ed.). *Child and adolescent psychiatry: A comprehensive textbook* (2nd ed., pp. 693–702). Baltimore, MD: Williams & Wilkins.

Offord, D. R., Boyle, M. H., Szatmari, P., Rae-Grant, N. I., Links, P. S., Cadman, D. T., Byles, J. A., Crawford, J. W., Blum, H. M., Byrne, C., Thomas, H., & Woodward, C. A. (1987). Ontario Child Health Study II.: Six-month prevalence of disorder and rates of service utilization. *Archives of General Psychiatry*, *44*, 832–836.

Orr, D. (1986). Adolescence, stress and psychosomatic issues. *Journal of Adolescent Health Care*, *7*(Suppl.), 975–1085.

Osborne, R. B., Hatcher, J. W., & Richtsmeier, A. J. (1989). The role of social modeling in unexplained pediatric pain. *Journal of Pediatric Psychology*, *14*, 43–61.

Oster, J. (1972). Recurrent abdominal pain, headache, and limb pains in children and adolescents. *Pediatrics*, *50*, 429–436.

Osterhaus, S. O. L., Passchier, J., Helm-Hylkema, H. van der, de Jong, K. T., Orlebeke, J. F., de Grauw, A. J. C., Dekker, P. H. (1993). Effects of behavioral psychophysiological treatment of schoolchildren with migraine in a nonclinical setting: Predictors and process variables. *Journal of Pediatric Psychology*, *18*, 697–715.

Pantell, R. H., & Goodman, B. W. (1983). Adolescent chest pain: A prospective study. *Pediatrics*, *71*, 881–887.

Parmelee, A. H. (1986). Children's illnesses: Their beneficial effects on behavioral development. *Child Development*, *57*, 1–10.

Pennebaker, J. W., & Skelton, J. A. (1978). Psychological parameters of physical symptoms. *Personality and Social Psychological Bulletin*, *4*, 524–530.

Perrin, J. M., MacLean, W. E., & Janco, R. L. (1988). Does stress affect bleeding in hemophilia? *The American Journal of Pediatric Hematology/Oncology*, *10*, 230–235.

Phillips, K. A., McElroy, S. L., Keck, P. E., & Hudson, J. (1993). Body Dysmorphic disorder: 30 cases of imagined ugliness. *American Journal of Psychiatry*, *150*, 302–308.

Pilowsky, I. (1987). Abnormal illness behavior and chronic pain. In G. D. Burrows, D. Elton, & G. V. Stanley (Eds.), *Handbook of chronic pain management* (pp. 131–136). New York: Elsevier.

Pilowsky, I., Bassett, D. L., Begg, M. W., Thomas, P. G. (1982). Childhood hospitalization and chronic intractable pain in adults: A controlled retrospective study. *International Journal of Psychiatry and Medicine*, *12,* 75–184.

Prazar, G. (1987). Conversion reactions in adolescents. *Pediatric Review*, *8*, 279–286.

Prichard, C. T., Ball, I. D., Culbert, I., & Faust, D. (1988). Using the personality inventory for children to identify children with somatoform disorders: MMPI findings revisited. *Journal of Pediatric Psychology*, *3*, 237–245.

Proctor, J. T. (1958). Hysteria in childhood. *American Journal of Orthopsychiatry*, *28*, 394–407.

Prugh, D. G. (1983). *The psychosocial aspects of pediatrics*. Philadelphia: Lea & Febiger.

Putnam, F. W., Helmers, K., & Trickett, P. K. (in press). Development, reliability and validity of a child dissociative scale. *Child Abuse and Neglect*.

Rauste-von Wright, M., & von Wright, J. (1981). A longitudinal study of psychosomatic symptoms in healthy 11–18 year old girls and boys. *Journal of Psychosomatic Research*, *25*, 525–534.

Reesor, K. A., & Craig, K. D. (1988). Medically incongruent chronic back pain: Physical limitations, suffering, and ineffective coping. *Pain*, *32*, 35–45.

Reynolds, C. R., & Richmond, B. O. (1985). *The Revised Children's Manifest Anxiety Scale Manual*. New York: Western Psychological Services.

Riddle, M. A., Scahill, L., King, R. A., Hardin, M. T., Anderson, G. M., Ort, S. I., Smith, J. C., Leckman, J. F., & Cohen, D. J. (1992). Double-blind, crossover trial of fluoxetine and placebo in children and adolescents with obsessive compulsive disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*, 1062–1069.

Rimza, M. E., Berg, R. A., Locke, C. (1988). Sexual abuse: Somatic and emotional reactions. *Child Abuse and Neglect*, *12*, 201–208.

Rivinus, T. M., Jamison, D. L., & Graham, P. J. (1975). Childhood organic neurological disease presenting as psychiatric disorder. *Archives of Diseases of Childhood*, *40*, 115–119.

Robbins, J. M., & Kirmayer, L. J. (1991). Attributions of common somatic symptoms. *Psychological Medicine*, *21*, 1029–1045.

Robins, E., & O'Neal, P. (1953). Clinical features of hysteria in children—with a note on prognosis: A two to seventeen year follow-up study of 41 patients. *Nervous Child*, *10*, 246–271.

Robinson, J. O., Alverez, J. H., Dodge, J. A. (1990). Life events and family history in children with recurrent abdominal pain. *Journal of Psychosomatic Research*, *34*, 171–181.

Rock, N. (1971). Conversion reactions in childhood: A clinical study on childhood neuroses. *Journal of the American Academy of Child Psychiatry*, *10*, 65–3.

Roghmann, K. J., & Haggerty, R. J. (1973). Daily stress, illness, and use of health services in young families. *Pediatric Research*, *7*, 520–526.

Routh, D. K., & Ernst, A. R. (1984). Somatization disorder in relatives of children and adolescents with functional abdominal pain. *Journal of Pediatric Psychology*, *9*, 427–437.

Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, health, and behavior*. London: Longman Group.

Saarni, C. (1990). Emotional competence: How emotions and relationships become integrated. In R. Thompson (Ed.), *Socioemotional Development: Nebraska Symposium on Motivation* (pp. 115–181). Lincoln, NE: University of Nebraska Press.

Sainsbury, P., & Gibson, J. G. (1954). Symptoms of anxiety and tension and the accompanying physiological changes in the muscular system. *Journal of Neurological and Neurosurgical Psychiatry*, *17*, 216–224.

Sanders, M. R., Rebgetz, M., Morrison, M., Bor, W., Gordon, A., Dadds, M., & Shepherd, R. (1989). Cognitive-behavioral treatment of recurrent nonspecific abdominal pain in children: An analysis of generalization, maintenance, and side effects. *Journal of Consulting and Clinical Psychology*, *57*, 294–300.

Sanders, M. R., Shepherd, R. W., Cleghorn, G., & Woolford, H. (1994). The treatment of recurrent abdominal pain in children: A controlled comparison of cognitive-behavioral family intervention and standard pediatric care. *Journal of Consulting and Clinical Psychology*, *62*, 306–314.

Sank, L. I., & Biglan, A. (1974). Operant treatment of a case of recurrent abdominal pain in a 10-year old boy. *Behavior Therapy*, *5*, 677–681.

Scaloubaca, D., Slade, P., & Creed, R. (1988). Life events and somatization among students. *Journal of Psychosomatic Research, 32*, 221–229.

Schneider, S., & Rice, D. R. (1979). Neurologic manifestations of childhood hysteria. *Journal of Pediatrics, 94*, 153–156.

Schulman, J. L. (1988). Use of a coping approach in the management of children with conversion reactions. *Journal American Academy of Child and Adolescent Psychiatry, 27,* 785–788.

Shaffer, D., Fisher, P., Dulcan, M. K., Davies, M., Piacentini, J., Schwab- Stone, M. E., Lahey, B. B., Bourdon, K., Jensen, P. S., Bird, H. R., Canino, G., & Regier, D. A. (1996). The NIMH Diagnostic Interview Schedule for Children version 2.3 (DISC-2.3): Description, acceptability, prevalence rates, and performance in the MECA study. *Journal of the Academy of Child and Adolescent Psychiatry, 35*, 865–877.

Shapiro, E. G., & Rosenfeld, A. A. (1987). *The somatizing child*. New York: Springer.

Shields, J. (1982). Genetic studies of hysterical disorders. In A. Roy (Ed.), *Hysteria*. New York: Wiley.

Sifneos, P. E. (1973). The prevalence of "alexithymic" characteristics in psychosomatic patients. *Psychotherapy and Psychosomatics, 22*, 255–262.

Simon, G. E. (1991). Somatization and psychiatric disorder. In L. J. Kirmayer & J. M. Robbins (eds.), *Current concepts of somatization research and clinical perspectives*. Washington, DC: American Psychiatric Press.

Spierings, C., Poels, P. J. E., Sijben, N., Gabreels, F. J. M., & Renier, W. O. (1990). Conversion disorders in childhood: A retrospective follow-up study of 84 patients. *Developmental Medicine and Child Neurology*, *32*, 865–871.

Srinath, S., Bharat, S., Girimaji, S., Seshadri, S. (1993). Characteristics of a child inpatient population with hysteria in India. *Journal of the American Academy of Child and Adolescent Psychiatry*, *32*, 822–825.

Steckel, W. (1943). *The interpretation of dreams*. New York: Liveright.

Stefansson, J. G., Messina, J. S., & Meyerowitz, S. (1976). Hysterical neurosis, conversion type: Clinical and epidemiological considerations. *Acta Psychiatra Scandinavian*, *53*, 119–138.

Steinhausen, H. C., Aster, M. V., Pfeiffer, E., & Gobel, D. (1989). Comparative studies of conversion disorders in childhood and adolescence. *Journal of Child Psychology and Psychiatry*, *30*, 615–621.

Stern, K., Murphy, M., & Bass, C. (1993). Personality disorders in patients with somatisation disorder: A controlled study. *British Journal of Psychiatry*, *163*, 785–789.

Stevens, H. (1986). Is it organic or is it functional, is it hysteria or malingering? *Neuropsychiatry*, *9*, 241–254.

Stevenson, J., Simpson, J., & Bailey, V. (1988). Research note: Recurrent headaches and stomachaches in preschool children. *Journal of Child Psychology and Psychiatry*, *29*, 897–900.

Stickler, G. B., & Murphy, D. B. (1979). Recurrent abdominal pain. *American Journal of Diseases of Childhood*, *133*, 486–489.

Stone, R., & Barbero, G. (1970). Recurrent abdominal pain in childhood. *Pediatrics*, *45*, 732–738.

Swartz, M., Landerman, R., & George, L. K. (1991). The diagnosis of somatization disorder. In L. N. Robins & D. A. Regier (Eds.), *Psychiatric disorders in America* (pp. 220–257). New York: Free Press.

Taylor, G. J., Baglay, M., Ryan, D.P., Parker, J. D. A. (1990). Validation of the alexithymia construct: A measurement-based approach. *Canadian Journal of Psychiatry*, *35*, 290–297.

Taylor, D. C., Szatmari, P., Boyle, M. H., & Offord, D. R. (1996). Somatization and the vocabulary of everyday bodily experiences and concerns: A communicate study of adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*, 491–499.

Torgensen, S. (1986). Genetics of somatoform disorders. *Archives of General Psychiatry*, *43*, 502–505.

Turgay, A. (1980). Conversion reactions in children. *Psychiatric Journal of the University of Ottawa*, *5*, 287–294.

Turgay, A. (1990). Treatment outcome for children and adolescents with conversion disorder. *Canadian Journal of Psychiatry*, *35*, 585–589.

Turkat, I. D. (1982). An investigation of parental modeling in the etiology of diabetic illness behavior. *Behavior Research and Therapy*, *2*, 547–552.

Tyler, P., Fowler-Dixon, R., & Ferguson, B. (1990). The justification for the diagnosis of hypochondriacal personality disorder. *Journal of Psychosomatic Research*, *34*, 637–642.

Volkmar, R. R., Poll, J., & Lewis, M. (1984). Conversion reactions in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *23*, 424–430.

Walker, L. S., Garber, J., & Greene, J. W. (1991). Somatization symptoms in pediatric abdominal pain patients: Relation to chronicity of abdominal pain and parent somatization. *Journal of Abnormal Child Psychology*, *19*, 379–394.

Walker, L. S., Garber, J., & Greene, J. W. (1993). Psychosocial correlates of recurrent childhood pain: A comparison of pediatric patients with recurrent abdominal pain, organic illness, and psychiatric disorders. *Journal of Abnormal Psychology*, *102*, 248–258.

Walker, L. S., Garber, J., & Greene, J. W. (1994). Somatic complaints in pediatric patients: A prospective study of the role of negative life events, child social and academic competence, and parental somatic symptoms. *Journal of Consulting and Clinical Psychology*, *62*, 1213–1221.

Walker, L. S., Garber, J., Van Slyke, D., & Greene, J. W. (1995). Long term health outcomes in patients with recurrent abdominal pain. *Journal of Pediatric Psychology*, *20*, 233–245.

Walker, L. S., & Greene, J. W. (1987). Negative life events, psychosocial resources, and psychophysiological symptoms in adolescents. *Journal of Clinical Child Psychology*, *16*, 29–36.

Walker, L. S., & Greene, J. W. (1989). Children with recurrent abdominal pain and their parents: More somatic complaints, anxiety, and depression than other patient families? *Journal of Pediatric Psychology*, *14*, 231–243.

Walker, L. S., & Greene, J. W. (1991a). Negative life events and symptom resolution in pediatric abdominal patients. *Journal of Pediatric Psychology*, *16*, 341–360.

Walker, L. S., & Greene, J. W. (1991b). The Functional Disability Inventory: Measuring a neglected dimension of child health status. *Journal of Pediatric Psychology*, *16*, 39–58.

Walker, L. S., McLaughlin, F. J., & Greene, J. W. (1988). Functional illness and family functioning: A comparison of healthy and somaticizing adolescents. *Family Process*, *27*, 317–325.

Walker, L. S., & Zeman, J. L. (1992). Parental response to child illness behavior. *Journal of Pediatric Psychology*, *17*, 49–71.

Warwick, H. M., & Salkovskis, P. M. (1985). Reassurance. *British Medical Journal*, *290*, 1028.

Wasserman, A. L., Whitington, P. F., & Rivera, F. P. (1988). Psychogenic basis for abdominal pain in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 179–184.

Weinberger, D. A., Schwartz, G. E., & Davidson, R. G. (1979). Low-anxious, high-anxious, and repressive coping styles: Psychometric patterns and behavioral physiological responses to stress. *Journal of Abnormal Psychology*, *88*, 369–380.

Weisz, J. R., Suwanlert, S., Chaiyasit, W., Weiss, B., Achenbach, T. M., & Walter, B. (1987). Epidemiology of behavioral and emotional problems among Thai and American children: Parent reports for ages 6 to 11. *Journal of the American Academy of Child and Adolescent Psychiatry*, *26*, 890–897.

Weisz, J. R., Suwanlert, S., Chaiyasit, W., Weiss, B., Achenbach, T. M., & Eastman, K. L. (1993). Behavioral and emotional problems among Thai and American adolescents: Parent reports for ages 12 to 16. *Journal of Abnormal Psychology*, *102*, 395–403.

Weller, E. B., Weller, R. A., & Fristad, M. A. (1985). Use of sodium amytal interviews in pre-pubertal children: Indications, procedure and clinical utility. *Journal of the American Academy of Child Psychiatry*, *24*, 747–749.

Whitehead, W. E., Busch, C. M., Heller, B. R., & Costa, P. T. (1986). Social learning influences on menstrual symptoms and illness behavior. *Health Psychology*, *5*, 13–23.

Whitehead, W. E., & Schuster, M. M. (1985). *Gastrointestinal disorders: Behavioral and physiological basis for treatment*. Orlando, FL: Academic.

Wilkinson, P., & Mynors-Wallis, L. (1994). Problem solving therapy in the treatment of unexplained physical symptoms in primary care: A preliminary study. *Journal of Psychosomatic Research*, *38*, 591–598.

Wirt, R. D., Lachar, D., Klinedinst, J. K., & Seat, R. D. (1984). *Multidimensional description of child personality: A manual for the Personality Inventory for Children*. Los Angeles: Western Psychological Services.

Wise, T. N., & Mann, L. S. (1994). The relationship between somatosensory amplification, alexithymia, and neuroticism. *Journal of Psychosomatic Research*, *38*, 515–521.

Wood, B. L. (1993). Beyond the "psychosomatic family": A biobehavioral family model of pediatric illness. *Family Process*, *32*, 261–278.

Woodbury, M. M., De Maso, D. R., & Goldman, S. J. (1992). An integrated medical and psychiatric approach to conversion symptoms in a four-year-old. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*, 1095–1097.

Wooley, S. C., Blackwell, B., & Windet, C. (1978). A learning theory model of chronic illness behavior: Theory, treatment, and research. *Psychosomatic Medicine*, *40*, 379–401.

Wyllie, E., Friedman, D., Luders, H., Morris, H., Rothner, D., Turnbull, J. (1991). Outcome of psychogenic seizures in children and adolescents compared with adults. *Neurology*, *41*, 742–744.

Zeman, J., & Garber, J. (1996). Display rules for anger, sadness, and pain: It depends on who is watching. *Child Development*, *67*, 957–973.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.26
Factitious Disorders

JUDITH A. LIBOW
Children's Hospital Oakland, CA, USA

5.26.1 INTRODUCTION 683
5.26.2 FACTITIOUS DISORDERS 684
5.26.2.1 Phenomenology 684
5.26.2.1.1 Munchausen syndrome and factitious disorder 684
5.26.2.1.2 Factitious disorder by proxy 685
5.26.2.2 Diagnostic Features/Differential Diagnosis 686
5.26.2.2.1 Diagnosis of childhood factitious disorder 686
5.26.2.2.2 Diagnosis of factitious disorder by proxy 688
5.26.2.3 Prevalence 689
5.26.2.3.1 Factitious disorder prevalence 689
5.26.2.3.2 Factitious disorder by proxy prevalence 690
5.26.2.4 Etiology 690
5.26.2.4.1 Origins of factitious disorders 690
5.26.2.4.2 Origins of factitious disorder by proxy 692
5.26.2.5 Developmental Course of Factitious Disorders 693
5.26.3 CONCEPTUALIZATION AND CLINICAL FORMULATION 694
5.26.3.1 Multimodal Assessment 694
5.26.3.1.1 The medical puzzle 694
5.26.3.1.2 Psychological assessment 695
5.26.3.2 Treatment Formulation 695
5.26.4 TREATMENT 696
5.26.4.1 Psychosocial Treatments 696
5.26.4.1.1 Psychotherapy for factitious disorder 696
5.26.4.1.2 Psychotherapy for factitious disorder by proxy 697
5.26.4.1.3 Efficacy of psychosocial treatment 698
5.26.4.2 Pharmacological Treatments 698
5.26.4.2.1 Uses of pharmacotherapy 698
5.26.4.2.2 Efficacy of pharmacotherapy 698
5.26.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 698
5.26.6 SUMMARY 699
5.26.7 REFERENCES 699

5.26.1 INTRODUCTION

Factitious disorder, simply put, is the intentional simulation of illness through a variety of false pretenses. This includes reporting exaggerated symptoms, fabricating laboratory results, and actually inducing symptoms through physical means such as ingestion of substances, injection, or suffocation. Factitious illnesses often result in numerous expensive and painful, not to mention unnecessary, laboratory tests, hospitalizations, and surgical procedures for the young "patient" before the charade is finally uncovered.

What is not so simple is the identification and understanding of this disorder, which poses serious challenges in differential diagnosis and effective treatment. Much better represented in the adult literature, often as "Munchausen syndrome," factitious disorder in youngsters is certainly known to occur, but has yet to be studied in any systematic manner. There is a very limited literature on factitious disorders in children *per se*, while there is an extensive literature of related disorders, diagnoses, and conditions which may provide intriguing clues to the etiology of this disorder. While this chapter will address factitious illness which involves a child or adolescent as the object of unnecessary medical attention, we will find that the perpetrator can actually be the child him or herself, the parent, the family unit, or a complex pathological collusion between parent and child. The relationships between these different yet similar presentations is still little understood, but poses intriguing clues for fiuture understanding of the development of the many forms of somatization and illness exaggeration throughout the lifespan.

5.26.2 FACTITIOUS DISORDERS

5.26.2.1 Phenomenology

Factitious illnesses have no doubt been observed for as long as there have been healers and physicians. Reports in the medical literature dating back to the early nineteenth century (Gavin, 1838; Jacobi, 1895) indicated a variety of falsely presented medical conditions such as factitious fevers. However, the most systematic and seminal discussion of fabricated illness dates back to the work of Asher (1951) writing in *Lancet*, who first coined the colorful name of "Munchausen's syndrome." This term was used to describe a perplexing type of patient who wandered far and wide seeking medical treatments and hospitalizations for nonexistent or self-induced illnesses. These patients were notorious for dramatic presentations, brief emergency room or surgical stays often ending in early detection and the flight of the patient to the next hospital or city. At the time, Asher reported that the three major organ system presentations involved abdominal, hemorrhagic, and neurologic symptoms. Since Asher's original paper, there have been hundreds of Munchausen syndrome papers in the adult medical literature elaborating on individual or sometimes a series of factitious illness cases involving all manner of factitious presentations, including fevers (Aduan, Fauci, Dale, Herzberg, & Wolff, 1979), rashes (Reich & Gottfried, 1983), and even psychiatric presentations (Hay, 1983; Phillips, Ward, & Ries, 1983; Pope, Jonas, & Jones, 1982).

5.26.2.1.1 Munchausen syndrome and factitious disorder

The disorder described by Asher was named after a fascinating man and equally colorful character, the Baron Karl Friedrich Hieronymus von Munchausen, an eighteenth century cavalry officer and mercenary who was renowned for telling elaborate tales of his imaginary adventures, chronicled by his friend Rudolf Erich Raspe (Sakula, 1978). Both men and women from all walks of life have been described as exhibiting the "Munchausen" disorder. The mean age of onset of the disorder is said to be most commonly in the patient's early twenties (Raymond, 1987) and perhaps up to half the cases may begin in adolescence (Plassmann, 1994a). The literature reports that men tend to predominate in the most chronic, "wandering" form of Munchausen syndrome (at a ratio of 2:1) in which multiple physicians, towns, and hospitals are caught up in the patient's relentless charade. Women have been found to predominate (at a 3:1 ratio) in the less dramatic forms of factitious disorder, more likely in episodic or less elaborate impostures (Eisendrath, 1996). However, the methods of self-abuse and the range of false symptoms and organ systems involved have been remarkably extensive and dramatic, confined neither to a single gender, ethnicity, socioeconomic, or age group. Many of these patients are involved in the medical, nursing, or health care professions (Plassmann, 1994b).

Within the factitious disorder (FD) literature are a handful of reports on children or adolescents discovered to have manufactured their own symptoms and illnesses. The reports of young children include a four year old feigning seizures (Croft & Jervis, 1989), a nine, ten, and fourteen year old creating skin lesions (Lovejoy, Marcuse, & Landrigan, 1971; White, Pearson, & Coddington, 1966), several series of children between 11 and 17 years with factitious fevers (Aduan et al., 1979; Herzberg & Wolff, 1972; Wedel, 1971), preadolescents adulterating laboratory specimens (Abe, Shinozima, Okuno, Abe, & Ochi, 1984; Abrol, Heck, Gleckel, & Rosner, 1990), and factitious charges of sexual and physical abuse by a 12 year old (Feldman & Smith, 1996).

Factitious illnesses show increasing sophistication with age, and the adolescent literature includes a patient deliberately ingesting steroids (Witt & Ginsberg-Fellner, 1981), a patient scraping her skin and pouring nutritional supplements down the drain (Sanders, 1995), self-induced chronic joint inflammation and pain in a 15 year old (Paperny, Hicks, & Hammar, 1980), use of a wrist tourniquet for self-induced edema by a 15 year old (Rodriguez-Moreno, Ruiz-Martin, Mateo-Soria, Rozadilla, & Roig-Escofet, 1990), the injection of exogenous egg protein into the bladder (Tojo et al., 1990), and 15 hospitalizations for a wide variety of nonexistent illnesses by an adolescent (Raymond, 1987).

The vast majority of these reports have appeared in the medical rather than the psychiatric literature, with the resultant emphasis more on the unraveling of the medical puzzle than an explication of the etiology or family dynamics. As in the case of adult FDs, it is clear that these cases, particularly with more clever adolescent imposters, can stymie physicians for several years before being discovered as factitious (Abe et al., 1984; Paperny et al., 1980). They generally result in an expensive, destructive process of multiple hospitalizations, laboratory procedures, and surgeries before the fabrication is finally unmasked. The physical and psychological toll on the patient, and the professional/emotional toll on the physicians and the health care system can be enormous (Schreier & Libow, 1993a).

5.26.2.1.2 Factitious disorder by proxy

That the great majority of FDs in children is seen in adolescents is no doubt due to the fact that a certain basic level of medical sophistication is required to create and maintain believable symptoms capable of entrapping physicians successfully. Unfortunately, a substantial number of very young children are also, in fact, victims of fabricated illness—victims of Munchausen by proxy syndrome (MBPS), or factitious disorder by proxy (FDbP). First systematically described and named by British pediatrician Meadow (1977), MBPS involves the use of the child (or less commonly, another adult) as the proxy "patient." The parent or caregiver exaggerates symptoms, or actually induces illness in her child in order to maintain an ongoing relationship with doctors and hospitals (Schreier & Libow, 1993a). The most common methods of inducing illness include poisoning, suffocation, tampering with specimens or intravenous lines, skin adulturation, witholding medication, and false reporting of symptoms. The child manifests the factitious illness, but it is the parent/mother who carries the FDbP diagnosis.

There are currently several hundred cases of FDbP reported in the medical literature, with the vast majority of perpetrators being mothers (Levin & Sheridan, 1995; Schreier & Libow, 1993a) with complex societal and systemic reasons for this gender imbalance. These mothers have been classically described as highly devoted, seemingly very concerned, and medically sophisticated parents, although a subset of these parents have also been noted to be more hostile and demanding (Libow & Schreier, 1986). A review in 1987 by Rosenberg reported on 117 cases in the literature, and found that 24% of the mothers themselves had Munchausen syndrome or at least some elements of FD. Classically, the fathers or partners of the MBPS mother are described as distant, uninvolved, or emotionally disengaged from the family. The child victims can be of any age, gender, or birth order, and sometimes several children in one family are victimized serially or simultaneously. Child victims can be biological children, adoptive, or foster children, and often do not show any obvious signs of attachment problems or physical child abuse. The age of the child victims have ranged from as young as fetuses subjected to prenatal fabrication (Goodlin; 1985; Goss & McDougall, 1992), to infants (Light & Sheridan, 1990; Meadow, 1990), preschoolers (Guandolo, 1985) and latency age (Libow & Schreier, 1986; Palmer & Yoshimura, 1984) as well as adolescent victims (Katz, Mazer, & Litt, 1985). Generally it is the infants and toddlers who are subjected to the most active forms of physical symptom inductions such as suffocation and poisoning (McClure, Davis, Meadow, & Sibert, 1996), while the older children tend to be victims of what have been called the "milder" form of MBPS (Roth, 1990; Warner & Hathaway, 1984). "Doctor shopping" and symptom exaggeration, in which the parent does not actively induce symptoms herself, can also subject the child to unnecessary procedures and medications through exaggerated or false reporting of symptoms (Masterson, Dunworth, & Williams, 1988). Often older children have been noted to engage in either passive or active collusion in such fabrications, confirming their mother's reports to the physicians (Woolcott, Aceto, Rutt, Bloom, & Glick, 1982), failing to report falsehoods (Sanders, 1995), or even participating in inducing symptoms.

Factitious disorder by proxy differs in a number of ways from FD, although there are obvious areas of overlap. Often the "by proxy" deception takes longer to be revealed, because the mothers are more convincing imposters and the fabrications are more sophisticated (Schreier & Libow, 1993a). Some percentage of the proxy victims actually die from the abuse. Rosenberg (1987) calculated that number to be approximately 9% of cases, although that likely represents an overestimation due to the preponderance of more severe cases in the literature she reviewed. Because the proxy cases involve an innocent victim, the morbidity and mortality of the patients is even more distressing, as is the devastating impact of this abuse, once it is revealed, on the treating physicians who are drawn into unintentional participation in the abuse (Schreier & Libow, 1993a).

5.26.2.2 Diagnostic Features/Differential Diagnosis

5.26.2.2.1 Diagnosis of childhood factitious disorder

As in all forms of factitious illness, the first and most important issue in differential diagnosis is the ruling out of bona fide physical illness. No matter how obscure or elusive, the presence of a true medical illness which can account for the presenting symptoms negates the diagnosis of FD, unless the symptoms are exaggerated way out of proportion to what should reasonably be expected. Thus, physicians generally expend considerable effort in attempting to provide the most exhaustive medical workup possible for a patient with mysterious symptoms, including extensive laboratory tests, radiological procedures, biopsies, genetic studies, and hospital observations, and often institute treatments in the hope that responses to treatment might provide clues to diagnosis. In most cases the diagnosis of FD is the last possibility considered, when all else has been ruled out, and when the "illness" fails to fit any usual pattern of symptoms, is physiologically impossible, fails to respond in an expected way to common treatments, and provides other clues as to its inconsistency and its tendency to reoccur only when the patient is not being observed closely. The identification of factitious illness in younger children tends to be easier due to the lack of sophistication of the falsifications, but as with adults, can be greatly complicated when the patient has coexisting, bona fide medical problems such as asthma, or serious iatrogenic illness induced by efforts to treat the original "problem." Livingston (1992) reminds us that several organic disease states have

> vague, strange, and transient symptoms and are easily confused with somatization: systemic lupus, multiple sclerosis, juvenile rheumatoid arthritis, acute intermittent porphyria, partial complex seizures, and inflammatory bowel disease ... it may be that only the longitudinal picture will establish the diagnosis firmly but a reasonable clinical judgment about working diagnosis can often be reached by complete history from multiple sources, thorough mental status exam, a careful probe for the family history, personality patterns and comorbid difficulties known to be associated with somatoform disorders, and ancillary psychological exams for characteristic profiles and findings. (pp. 14–15)

DSM-IV (American Psychiatric Association, 1994) includes "Factitious Disorder" as an Axis I disorder involving three elements, including: (i) the intentional production of physical/psychological signs and symptoms; (ii) the motivation to assume the sick role; and (iii) the absence of external incentives for the behavior. It includes three subcodes based on whether the signs and symptoms are predominantly psychological, predominantly physical, or a combination of physical and psychological in nature. *ICD-9* (Stegman & Aaron, 1996) lists factitious disorders as subcategories of "Hysteria" under "Neurotic Disorders."

There are separate categories for factitious illness with psychological symptoms (300.16) and other and unspecified factitious illness (300.19) which includes factitious physical symptoms. There is a different category for "Munchausen syndrome" or chronic factitious illness with physical symptoms (301.51) in *ICD-9*, coded separately under the "Histrionic personality disorders."

Seemingly quite straightforward, the diagnosis of FD in adults as well as children and adolescents is complicated by its considerable overlap with a number of related disorders which also involve the presentation of physical or psychological symptoms which are unobservable, do not appear to have an organic basis, or have an element of conscious production, unconscious motivation, or secondary gain for the patient. Somatization disorder, malingering, conversion disorder, anxiety disorders, and mood disorders can all be confused with FD, and differential diagnosis is often a challenge. The three most critical variables in differentiating the phenomenon of FD are the following.

(i) *Motivation.* Patients who are motivated to produce factitious symptoms for obvious external gain, such as disability payments, avoidance of military service, or evasion of criminal prosecution, rather than less conscious needs as in FD, are identified as malingerers, a V-Code in *DSM-IV*. Identification of the possible material gain/motivation of the patient requires careful identification of the tangible impact of the patient role on the person's lifestyle, finances, and responsibilities. In the case of children, the differential diagnosis of malingering from somatization and factitious disorder is complicated by the fact that in all cases, the child's creation of illness symptoms is likely to result in reduced school attendance and considerably increased parental involvement. In these cases, a careful psychosocial and family systems evaluation will be critical in sorting out the likely motivation for assuming the sick role, and determining whether fully conscious as opposed to less conscious needs are being gratified. In the case of the FDs, it is proposed that assuming a dependent yet controlling role in relationship to the powerful and caregiving physician is the primary motivation. However, in reality there are often tangible gains for young FD patients—such as missing school or engendering parental sympathy—which can be identified as possible secondary incentives, but not the primary motivators of the behavior.

(ii) *Intentionality*. Probably the most difficult variable to assess in FD is the degree to which the exaggeration or falsification of symptoms is conscious and intentional. Elaborately orchestrated deceptions are much easier to identify as intentional than exaggerated symptom complaints. By definition, FD is intentional, although believed to be compulsively driven, and therefore not entirely under the control of the patient. Patients who present repeatedly with a range of nonexistent medical complaints involving multiple symptoms and organ systems, but who are *not* consciously exaggerating or falsifying symptoms, are diagnosed with somatization disorder. As in FD, somatization can be episodic or chronic, and can pose major difficulties for the physician attempting to treat questionable symptoms with no apparent organic basis. Conversion disorder is diagnosed in the case of more restricted symptoms related to identified stressors (e.g., abdominal pain or "paralysis" following an incest experience). As in factitious disorder, somatization can be episodic or chronic, and can pose major difficulties for the physician atempting to treat questionable symptoms with no apparent organic basis.

(iii) *Chronicity*. There is some confusion about the use of the term "Munchausen syndrome," as it is often used generically to refer to all forms of fabricated illness. In general, presentation of the patient with chronic false symptoms, and a lifestyle centered around repeated wandering from hospital to hospital seeking medical care is classically known as "Munchausen syndrome," while FD is not as chronic, and can involve even a single episode (Folks & Freeman, 1985; Nadelson, 1979).

Munchausen patients have also been described as having poorer employment histories and more antisocial traits (Carney, 1980; Folks & Freeman, 1985; Spiro, 1968). Clearly, the term FD is more appropriate for use with child patients by these criteria. While many authors use both terms interchangably to describe patients who produce their own symptoms, others have attempted to identify differences in the groups in terms of gender distribution (Feldman & Eisendrath, 1996; Folks & Freeman, 1985), age, and degree of compulsivity (Spiro, 1968). What remains unclear is whether more episodic FD is qualitatively different, or whether the episodic patient eventually becomes the chronic Munchausen patient, given sufficient success over time in the impostering of illness.

What becomes clear is that the differentiation of patient intentionality, motivation, and even frequency of fabrication are not always easy to gauge by the clinician making the diagnosis, suggesting there may be considerable lack of specificity in the actual use of these terms or in the disorders themselves. A number of clinicians have proposed the concept of a continuum or spectrum of related disorders including Aduan et al. (1979), Barsky (1992), Fink (1992), Folks and Freeman (1985), and Nadelson (1979). Cramer, Gershberg, and Stern (1971) remind us that "A volitional factor in the genesis of hysterical symptoms has long been recognized ... " as well as " ... unconscious factors in the genesis of malingering" (p. 576). A number of authors have pointed to the multiple difficulties in differentiating the various "dissimulating disorders" including Jonas and Pope (1985) who remind us that malingering, conversion disorders and FDs have several striking similarities. All typically begin in late adolescence or early adulthood, all appear to have a relatively chronic course, therapeutic approaches appear to be relatively ineffective for all, and the patients all appear to display a similar pattern of psychiatric diagnoses which include histrionic, antisocial, and borderline personality disorders. They even suggest there may be a gender bias in diagnosis, with women tending to receive somatization and conversion disorder diagnoses, while men may tend to be perceived as Munchausen or malingering patients.

Spiro (1968) agrees that hysterical and factitious symptomatology are not always mutually exclusive categories, and he describes a classic Munchausen patient who also suffered occasional conversion abdominal pain. Fink (1992) also points out that while somatization and FD appear to constitute separate entities at any one point in time, the categories may actually merge or overlap over time. That is, his study of 113 patients found that 18% of his "persistent somatizers" were discovered to have intentionally self-induced their diseases, lesions, or illnesses at least once or to have manipulated test results. "The number of patients who had produced factitious objective symptoms were probably underestimated as they were only registered if clear evidence was present" (p. 134). He suggests that we may tend to underdiagnose the number of patients with FD because clear evidence of deliberate fabrication is not always present, and it is so difficult to know whether symptoms are actually consciously produced.

It is possible that this clinician bias towards diagnosing somatization is even stronger when it comes to child patients, as it may be even more psychologically difficult for medical providers to suspect children of deliberate falsifications, and there is less extensive psychosocial history available for assessing possible motivation, intent, etc. Some authors have indicated awareness that some supposed child "somatizers" may in fact be intentionally producing their symptoms (Garralda, 1996; Sanders, 1995). There is often a broad definition of child somatization used in the literature, such as Campo and Fritsch's (1994) use of the term for "medically unexplained physical symptoms" (p. 1223). In the absence of any attention to the issues of motivation or conscious awareness, this broad use of the term "somatization" could easily encompass an unknown percentage of deliberately "ill" children. This remains an issue for which our definitions and diagnostic tools require further refinement.

5.26.2.2.2 Diagnosis of factitious disorder by proxy

DSM-IV (American Psychiatric Association, 1994) includes a second category of factitious disorder, called "Factitious Disorder Not Otherwise Specified." This is supposed to incorporate cases of FDbP, although FDbP itself is included in the *DSM-IV* Appendix as needing additional study and refinement. It is differentiated from the category of "Factitious Disorder" by its definition as "Intentional production or feigning of physical or psychological symptoms *in another person who is under the individual's care*." Thus, their definition would include the abuse of an elderly person, hospital patient, or another adult. As in FD, the diagnosis requires the motivation to assume the sick role by proxy, the absence of external incentives, and the condition that the behavior is not better accounted for by another mental disorder. Perpetrators may, and often do, also receive additional diagnoses of various personality, somatoform, or mood disorders where appropriate. There is no discrete psychiatric diagnostic category within *ICD-10* for FDbP.

As in FD, the parental behaviors seen in MBPS or FDbP first and foremost must be differentiated from those seen in response to bona fide physical illness. If the medical problem in question is, in fact, due to real physical illness, the diagnosis of FDbP is inaccurate. Krener and Adelman (1988) point out that there are some superficial similarities between the "pathogenic parents with Munchausen's syndrome by proxy and the parents of chronically ill children" (p. 946). They demonstrate that chronic illness itself may create some features of parental behavior commonly seen in FDbP parents, including extensive medical knowledge, an unusual focus on the child's bodily products, and "inseparable maternal devotion." While there are very few known cases of misdiagnosed FDbP, the potential dangers of wrongly accusing a parent of this form of child abuse and mental disorder are very sobering. One of the few known cases of a misdiagnosis of FDbP involved a mother who was jailed for the murder of her infant son, believed to have poisoned him with ethylene glycol (Chin & Breu, 1991). It was only the illness of a second child, while his mother was in jail and without access to this child, that alerted authorities to the misdiagnosis and eventually allowed for the discovery of the rare genetic disorder that had killed one child and made the other so ill. In the majority of cases, physicians are generally extremely cautious about making the diagnosis, often spending months or years, and sometimes hundreds of thousands of dollars in repeated tests and procedures before acknowledging that FDbP is the only possible explanation of a persistant medical puzzle (Schreier & Libow, 1993a).

What is often confusing to diagnosticians and physicians is that the behavior of the parent with FDbP seems to fall on a continuum that includes a range of both normal and abnormal parent behavior in relation to the health care system. Eminson and Postlewaite (1992) suggest there exists a range of "parents' desire to consult for their child's symptoms" which includes nine points along the spectrum with classical neglect at one extreme and classical MBPS (or FDbP) at the other. They propose that other points along the spectrum include parental carelessness, noncompliance, and "normal" responses to child's symptoms, progressing up the scale to anxious behaviors, exaggeration, and finally, invention of symptoms. It clearly requires patience, expertise, considerable time, and investigation of the family system and the presenting symptoms in order to differentiate the anxious parent, the somatizing parent, or the thought-disordered parent, from the parent with FDbP. As in FD, the central elements of the FDbP diagnosis once again revolve around a persistent search for treatment of an induced or exaggerated illness, the parent's conscious intent to deceive, and an unconscious motivation for a relationship with the medical system rather than more material secondary gain. As Meadow (1984) reminds us, "Unfortunately there is no sharp dividing line between deliberate falsification (malingering for conscious gain) and abnormal illness behaviour in which there is unconscious gain (hysterical behaviour). The end result for the child is the same, and can be both cruel and dangerous, regardless of the origin of the mother's behaviour" (p. 156).

Even more complex is the difficulty of diagnostically sorting out the puzzle of FDs in the case of older children. Beyond the age of toddlerhood, many if not most child victims of FDbP are believed to have at least some level of awareness of the exaggeration or induction of their symptoms by their parents, and some are known to actively participate in verifying "the story" in the case of "doctor shopping" (Woollcott et al. 1982) or actually inducing their symptoms. In fact, the one study of adults victimized by FDbP as children (Libow, 1995) found that many of the victims were fully aware of the deception, but too intimidated or unsure of how to free themselves from this abuse to speak up. Sanders (1995) presents what she calls a "continuum of collusion with Munchausen by proxy syndrome" which includes the categories of the naive child, passive acceptance, active participation, and active harm.

Many studies indicate that some child and adolescent victims of FDbP actively collude in the deceptions with their parents in an effort to protect or maintain a relationship with their parents or continue the "only life they know" (Schreier & Libow, 1993a). What becomes very difficult at this point is to sort out the appropriateness of using the FD diagnosis for the child actively participating in the fabrication as opposed to the status of FDbP victim. It becomes particularly difficult to differentiate the child's possible beliefs and motivations regarding the illness—for example, should the colluding child be considered a somatizer if a parent is responsible for directing the deliberate falsifications?

Other concerns have been raised about the diagnosis of FDbP including those of Morley (1995) and Baldwin (1996) who are concerned about the nonspecificity of many of the criteria for this diagnosis, and the inclusion of other forms of fabrication beyond active induction of illness including false symptom reporting and exaggeration of genuine illness; some feel this may result in overinclusion of too many anxious or overdramatizing parents. Fisher and Mitchell (1995) take issue with the concept of FDbP as a "syndrome" as they feel that perpetrators manifest a wide range of psychopathology and presentations, with many "psychological, psychiatric, and environmental pathways leading to a behaviour of fabricating illness in a child" (p. 533).

5.26.2.3 Prevalence

5.26.2.3.1 Factitious disorder prevalence

Solid, systematic prevalence data on FD does not really exist, due to the significant challenges in assessment of intentionality and motivation, and the surreptitious nature of this disorder (Taylor & Hyler, 1993). Furthermore, since diagnostic criteria for factitious and somatoform conditions have been under revision over the years, epidemiological studies are difficult to conduct (Krener, 1994).

There is actually more prevalence-related literature available on FD in adults, with factitious fever one of the most studied single symptoms. Knockaert, Vanneste, Vanneste, and Bobbaers (1992) found that 3.5% of 199 cases of fever of unknown origin seen at one hospital over a 10 year period were factitious. Their review of four other studies of unexplained fever revealed rates of factitious fever ranging from 2.5% to 4.5%. Herzberg and Wolff (1972) reported that 13 patients in a sample of 200 (or 6.5%) at the National Institutes of Health (NIH) were considered to have factitious fever. A study of fever at Stanford University (Rumans & Vosti, 1978) found that 2.2% of some 500 cases were factitious in origin. Aduan et al. (1979) reported on another study of fever of unknown origin at NIH, involving 343 patients, which identified 9% as factitious (this group included 10 children under 18 years old).

A study that spanned a broader spectrum (Reich & Gottfried, 1983) surveyed 10 years of records on the medical and surgical wards of a Boston hospital, and identified 41 patients with a variety of documented FDs (they do not report prevalence figures) including some adolescents. The most common and serious disorders were infections produced by injection or insertion of contaminants through or into the skin surfaces. They concluded that "factitious disorders are far from rare on hospital wards, at least in tertiary hospitals" (p. 245). Pope et al. (1982) reported on a study of 219 psychiatric patients, which found that 6.4% probably had factitious psychotic symptoms. They suggest that clinicians should be especially suspicious of histrionic and borderline patients who experience "brief dips into psychosis."

A study of 282 patients with heavy utilization of in-patient admissions in Denmark (Fink, 1992) indicated that approximately 20% of the patients had no physical illness or disorder that could adequately explain their multiple admissions. They conducted several comparisons of what they called "persistant somatizers" and "nonsomatizers" and found that 18% of the "persistant somatizers" (defined as having more than six unexplained admissions) had at least one known admission for a factitious illness. It is at least clear that FD in adolescents and adults is not a trivial problem. Unfortunately, similar data does not exist for FD in children. But there are indications that it may not be rare in children, given observations of the high rates of somatization in children (Garralda, 1996) and of conversion symptoms in adolescents (Friedman, 1973) as well as Livingston's (1992) observation that "Mendacious presentations of physical disorder, currently classified either as factitious disorder or as malingering, are also far from rare among young people" (p. 13).

5.26.2.3.2 Factitious disorder by proxy prevalence

The collection of relevant prevalence data on MBPS or FDbP is also confounded by the absence of any national registry on this form of child abuse, and the likelihood that significant numbers of these cases are never identified or are misdiagnosed (Holborow, 1985; Schreier & Libow, 1993b). However, as with FD in adults, there are some intriguing studies which provide suggestive data that can be extrapolated to the larger population.

Some studies have been done on specific types of specialized medical problems in an effort to determine the possible rates of factitious presentations. Light and Sheridan (1990) surveyed 125 infant apnea programs across the United States and reported that 54 cases (2.7 per 1000) were believed to involve parents inducing apnea. Warner and Hathaway (1984) found that about 5% of some 300 parents seen in their allergy clinic appeared to have a "parental obsession" with food allergies that appeared to constitute FDbP. McClure, Davis, Meadow, and Sibert (1996) studied all cases in Ireland in a two year period involving a child protection conference, and identified 55 cases of MBPS out of 128 cases. They reported a combined incidence of MBPS, poisoning, and suffocation of 0.5 per 100 000 children, but admitted that these were likely an underestimate because cases were only identifed if they were reported and a child protection conference was called. They also found large regional differences in patterns of reporting, which also suggested differences in the ability of reporters to effectively recognize or pursue child abuse cases. Similar differences in rates of identified cases were reported by Schreier and Libow (1993b) who surveyed all 880 pediatric neurologists and 388 gastroenterologists in the United States in 1991 to collate data on their MBPS cases. With return rates of 21.8% and 32.4% respectively, these subspecialists reported a total of 273 confirmed cases of FDbP. As in the McClure study, the significant variability in the number of cases seen by different reporters suggests that there are great differences in the familiarity with and ability to identify these cases and that perhaps a significant number are missed entirely.

5.26.2.4 Etiology

5.26.2.4.1 Origins of factitious disorders

Since there are so few papers on FD in children and adolescents, and even fewer discussions of etiology in childhood cases, the adult literature provides the most useful directions for theory building. However, it must be acknowledged from the outset that all discussions of etiology and treatment are handicapped by the fact that these patients are notoriously resistant to psychotherapeutic intervention, and often flee from the hospital before mental health professionals even have an opportunity to obtain an adequate psychosocial history. However, case studies and theoretical discussions of this disorder provide intriguing ideas about the origins of fabricated illness.

One of the earliest discussions of etiology of chronic FD is offered by Spiro (1968) who points out the parallels between FD and imposture, including the patients' histories of early neglect and deprivation, a desire for mastery over early traumatic experience, and a discrepancy between the patient's self-image and the ego-ideal. He suggests that the patient repeatedly seeks mastery over his/her early experiences of pain and fear by his/her demanding behavior in the guise of being a medical patient, which is coupled with revenge and hostility.

Cramer, Gershberg, and Stern (1971) present four adult Munchausen cases and offer similar observations on their object relationships. In particular, they point to their patients' early experiences with physicians as authority figures who likely began as idealized figures but gradually generated disappointment and eventually, feelings of revenge. They point to their patients' intense wish for closeness and a state of dependency on the physician, paradoxically mixed with feelings of rivalry and desire for control. They describe serious ego disturbances, which are manifested in unstable identifications, primitive defenses and a surge of sadistic wishes against their abandoning parent(s), as represented by the physician. Plassmann (1994a) makes a distinction between the (chronic) Munchausen patient and the (episodic) FD patient, suggesting that the Munchausen patients have a history of infantile-acquired attachment to physicians who are omnipotent, idealized figures, while FD patients have a less idealized relationship with physicians and more of an early fear/interest in their own bodies. He found that many of his FD patients (54%) had been subjected to physical abuse in childhood, which is eventually transformed, often in adolescence, into the patient's destructiveness towards his own body.

Eisendrath (1996) also suggests that childhood experiences of physical or sexual abuse may be significant causes of later FD, generating feelings of guilt in response to sexual excitement or anger. The FD is hypothesized as a possible vehicle for the abuse to be masochistically re-enacted, with the physician serving as the abusing or depriving parent. The use of deception may create a sense of power and superiority in relation to the caregiver. He also suggests that FD may serve as a means for the patient to achieve a greater sense of control and mastery, needed perhaps as a result of childhood traumatic illness. Alternatively, he also proposes a more behavioral theory that FD may have its origins in childhood reinforcement for illness behaviors by eliciting parental attention, support and caring, and an escape from childhood responsibilities. Spiro (1968) describes just such a case in which the earliest memories of a chronic FD patient involved his happy recall of winning his parents' attention during a hospitalization at age four years and a series of later medical complications.

Several authors also highlight the crucial issue of how children learn illness behaviors within the family milieu. Barsky (1992) offers the concept of "amplification" to explain the process by which people learn to interpret their bodily sensations as intense and noxious. While sidestepping the issue of etiology, he suggests the three key elements in this process as bodily hypervigilence to unpleasant bodily sensations, a tendency to focus on weak or infrequent sensations, and the tendency to appraise these sensations as abnormal. The possible contribution of parents in this process is unspecified.

Krener (1994) points more directly to the ways within the family context that children may learn how to experience and communicate their emotional and physical states. "Parents in psychosomatic families may adduce physical symptoms as an explanation of their child's state, behavior, or response, just as they use them to explain their own feelings. Children are more vulnerable than adults to this influence and to developing somatic representations of emotional states, which, like other coping and mastery skills acquired in childhood, involves learning, suggestion, and psychological compliance" (p. 419). It is not difficult to imagine that a child living with a somatizing or even a fabricating parent might come to rely increasingly on somatized explanations and a patient role, eventually coming to participate more actively in convincing physicians to believe in a nonexistent illness.

Unfortunately, the rather limited literature on childhood and adolescent FD is even more restricted in the number of papers which present any detailed early history, parental or family psychosocial background that could shed light on the various theories of the genesis of FD. Many focus almost entirely on the medical presentation (Abe et al., 1984; Abrol et al., 1990; Lovejoy et al., 1971; Paperny et al., 1980; Wedel, 1971). However, a few papers on child FD point to an intriguing link between child and parent preoccupation with illness.

A careful reading of the child FD literature describing family background suggests that some of the mothers play a pivotal role in encouraging or even teaching illness fabrication to the child (Stankler, 1977). Sneed and Bell (1976) describe a 10-year-old boy who presented with abdominal pain, bloody urine, and "kidney stones" which were later found to be ordinary pebbles. The article described the mother as actively involved in the fabrication, reporting "blood" on the stones and describing her own passage of kidney stones years earlier; she "tended to exaggerate physical problems and illnesses for herself, the siblings, and the patient" (p. 127). Her son was described as passive and compliant, and almost completely dependent upon his mother.

Herzberg and Wolff's (1972) report on six children presenting with factitious fever of unknown origin at the NIH also included some fascinating data about the mothers' presentations. The children ranged from 11 to 17 years of age, and were all described as "good, compliant, serious, but lonely" children. The authors felt that four of the mothers had "very troubling chronic illnesses of their own" which appeared highly inconsistent with observations or history. Several of the mothers were actually noted to be involved in the fraudulent reporting of the childrens' temperature elevations; one mother persistently pictured her daughter as suffering from leukemia despite negative findings! The authors observed that the children appeared to be recruited to help the parents with their own inner conflicts. The "sick" children were noted to enjoy increased attention from their parents that might not have otherwise been available as well as a break from demanding schoolwork. Interestingly, two of the six children aspired to be medical doctors!

One of the most elaborate factitious illness presentations (Gilbert, Pierse, & Mitchell, 1987) involved a 13-year-old boy seen over six months time at an otolaryngology clinic with recurrent otalgia, bloody discharge from the ear, and chronic mastoiditis with a cerebrospinal fluid leak. Just as the child was facing a temporal craniotomy, forensic analysis identified the "ear fluid" as beetroot juice, and the boy admitted to placing the material in his ear. A retrospective review of this boy's three-year medical history revealed a long series of nonspecific complaints resulting in 28 outpatient visits, 15 specialist consultations, 4 surgical procedures, 2 endoscopies, 52 blood tests, and 20 x rays. The medical staff wondered if the boy's mother had been coaching him in his fabrications since "It seemed unlikely that a 13-year-old would have the medical sophistication to simulate a cerebrospinal fluid leak or comprehend the significance such a sign would have for the attending surgeons" (p. 232). Also suggestive of FDbP involvement was the mother's domineering style and devoted attendance at bedside, previous health care employment, and efforts to have her son immediately discharged when confronted with the fabrication. Many years earlier, Steinbeck (1961) reported on a 19-year-old girl simulating massive vaginal hemorrhage and hematemesis through concealing blood; her mother was also later found to be an accomplice in her deceptions.

Similar reports of maternal involvement are found in a case described by Croft and Jervis (1989) of a four-year-old, seen over a year's time, with 35 emergency room visits for repeated "fits," trembling, and a "paralyzed" arm which were all being feigned by the child. The boy had a history since age two which included "cyanosis" and episodes of shaking and unconsciousness at age three. The child eventually described how his mother had taught him to simulate his epileptic "fits" by shaking, falling out of his chair, and flickering his eyes, and he reported being rewarded by his mother with food and drink when he performed for her. His "fits" stopped after five months in foster care. Unfortunately, there is no psychosocial history provided for the mother herself.

Perhaps one of the most unusual descriptions of FD collaboration between parent and "child"—in this case a 25-year-old daughter and her 59-year-old mother—is offered by Janofsky (1986). Both mother and daughter were hospitalized together with similar complaints of fevers, nausea, headaches, weakness, etc., and both had long medical histories engaging over 100 physicians. When psychiatric consultations were recommended, both mother and daughter fled the hospital against medical advice. The author suggests that this case of chronic FD in the mother, and a budding FD daughter, would meet the classic criteria of "folie a deux," or even "folie a famille," given the husband's willing participation as well as the physician son's facilitation of the process!

So it is apparent that FD behavior can involve overt modeling, direct coaching of the child, active collaboration on the part of the parent, and behavioral reinforcement for the child's participation. Many of these parents would no doubt fit the definition of MBPS parents. What becomes increasingly difficult to differentiate diagnostically, and an area of increasingly blurred boundaries, is the distinction between the FDbP parent with a "colluding child," and a childhood FD with a cooperative parent. These cases suggest that at least some percentage of childhood FD cases have their origins in the child's direct exposure to the coaching of a parent with factitious disorder or factitious disorder by proxy.

5.26.2.4.2 Origins of factitious disorder by proxy

As in adult FD, theories of etiology for FDbP are limited by the difficulties in obtaining accurate and detailed psychosocial histories from these reluctant and often court-ordered patients. Spiro's (1968) and Cramer et al.'s (1971) theories of the etiology of FD have many parallels to the origins of FDbP, particularly in terms of histories of early neglect and deprivation, and the mother's enactment of an idealizing/hostile attachment to the physician. The description of a MBPS mother's early background by Wood, Fowlkes, Holden and Castro (1989) is fairly typical. They describe a childhood of early deprivation and humiliation, resulting in marked insecurity, deep resentment of authority figures, and inability to meet the patient's needs through the usual means. In the case of FDbP, the infant or child is hypothesized to serve as a type of "fetish object" in the mother's "perverse" relationship which is focused on reducing anxiety by controlling the physician/authority figure and warding off abandonment (Schreier & Libow, 1993a). While many FDbP mothers appear to have some history of paternal loss or abandonment, others have different, less dramatic types of loss, as in a loss of self- or parental esteem through neglect or traumatic illness, or loss of a father's interest. Interestingly, many mothers with FDbP seem to re-enact these unsatisfying relationships with men in later life by marrying aloof and distant husbands who are often absent and uninvolved in their wives' lives and the children's "illnesses." Somehow, these fathers are allowed to remain aloof and uninvolved by our physicians and health care institutions, who seem to accept their absence or fail to insist on their participation.

In fact, social factors which allow for distortions and inequities in gender roles also appear to play their part in the etiology of FDbP, in that many women grow up in a societal context in which they feel desperately unrecognized, and have few outlets for a sense of power and importance outside of their caretaking roles as mothers. Ironically, the role of "devoted mother" of a chronically ill child is one of the few traditional, socially sanctioned roles for women which also allows them access to an exciting and powerful world of hospitals and health care (Schreier & Libow, 1993a). Others place some responsibility on our medical systems (Donald & Jureidini, 1996; Fialkov, 1984) for the development of MBPS, or on the "medicalization of childhood, clinical and social iatrogenesis and the creation of a society obsessed with health and its concomitant illness" (Baldwin, 1996, p. 161). But while there are interesting theories about these different influences, we know little about the relative contributions of childhood neglect, character disorders, sociocultural influences, and family systems dysfunctions in their eventual evolution into the destructive and often life-threatening behavior of medical abuse.

5.26.2.5 Developmental Course of Factitious Disorders

While addressing a somewhat different population, the literature on the course of somatization disorder in children offers useful directions for our thinking about the course of FD in children. Livingston's (1992) excellent review of somatization in childhood and adolescence found that it tends to have a progressive or chronic course similar to that found in adults. Furthermore, he reported that somatoform syndromes in children that do not yet constitute full somatization disorder may do so over time, and childhood somatizers are unlikely to reach adulthood without developing other kinds of serious psychopathology. His general conclusions were that abdominal and other pains develop first in the early school-age population, followed by neurologic or conversion symptoms during later elementary or adolescent years. The number of symptoms then appears to peak in the female patient's early 20s and decline gradually, although for men the peak appears to be several decades later. Livingston (1993) has also found that children of adults with somatization disorder almost universally develop psychiatric disorders.

Factitious disorder is not ordinarily seen until at least early adolescence in most children due to lack of medical sophistication (Aduan et al., 1979). Children much younger than 12 are very likely to have an active parental ally either engineering or at least encouraging the deception, while adolescent patients might be expected to have some earlier history of medical problems or somatization. Reich and Gottfried's (1983) study of FD provides some data on the early histories of their patients. All of their patients who self-induced their infections began in adolescence although many were not identified until later in adulthood. All of them apparently showed an early interest in medical matters and hypochondriacal preoccupations, although none had histories of serious childhood diseases. Paperny et al. (1980) described the case of a late adolescent whose history of factitious illness spanned the five years between 15 and 20 years of age. He endured 23 hospitalizations and 13 surgical procedures, while functioning as an "excellent student" who worked for a year in an orthopedist's office. An even longer history was available on a 28-year-old FD patient described by Sale and Kalucy (1980) who found that this patient lived with a mother who was ill during most of her childhood, and had infrequent and probably genuine medical visits until age 12. From age 12 through 17 she made frequent use of the gynecology clinic with abdominal pain complaints and engaged in many struggles with her physicians; she manifested features of adult FD from age 17 onwards. Reich, Lazarus, Kelly, & Rogers (1977) provide detailed early history on a 15-year-old boy who simulated an enterovesicular fistula by contaminating his urine with feces and food by retrograde injection of substances into his bladder. They found that his early history included moderately severe asthma, multiple allergies, surgery at age two for hypospadias, and a prolonged episode of unexplained fever at age 12. The authors concluded that the patient's factitious behavior was due to a hypochondriacal anxiety state about his urogenital system, but did not address the question of familial or maternal contributions to his medical history or to his surprising level of medical sophistication at a young age.

Little can be said about the long-term developmental outcome of FD beyond the fact that the illness behavior is likely to continue into adulthood, and may or may not become chronic. All that is available is retrospective data, as neither the young patients actually diagnosed with this disorder nor those whose diagnoses are missed can be studied prospectively.

There is somewhat better developmental data available on child victims of FDbP. It is clear that some percentage of infants subjected to the most severe forms of this abuse will die as a result of suffocation and poisoning (Meadow, 1990) and some will have long-term physical damage (Rosenberg, 1987). Certainly there is known psychological damage to children who endure medical fabrications at the hands of their parents (Bools, Neale, & Meadow, 1993; Roth, 1990). McGuire and Feldman (1989) not only found many severe adverse psychological effects of FDbP on their child victims, but also found that some of their young patients developed conversion symptoms or were starting to manifest FDs themselves.

A fascinating case of a 9 year old being contaminated by human fecal matter by her mother was presented by Palmer and Yoshimura (1984). Their examination of the mother-perpetrator found that the mother herself had an extensive history of FD in adolescence, adult FD later in life (as well as a job as a nurse practitioner), and eventually FDbP using her own daughter! Investigations even earlier into this mother's history revealed that she herself had had an extensive childhood medical history which included an unexplained fracture as an infant, failure to thrive, episodic vomiting, frequent injections for minor complaints, and several questionable procedures done at an early age. The possibility in some families of a multigenerational cycle of MBPS alternating with FD is a fascinating and disturbing one.

A retrospective study of 10 adults who reported themselves as victims of childhood FDbP (Libow, 1995), some as much as 40 years earlier, noted that they experienced serious developmental and emotional problems throughout the lifespan including growth failure, school problems, poor self-esteem, difficulty in relationships, depression, and difficulty separating fantasy from reality. Remarkably, some even reported that their mothers sporadically attempted to continue to abuse them even into adulthood! Thus, unless it is actively stopped by physicians who discover the deception, FDbP may run a chronic course, although the most active forms of induction in the child's early years tend to give way to more indirect forms of illness exaggeration, often gradually drawing the child or adolescent into passive or active cooperation, with increasing risk of physical harm and psychological morbidity.

5.26.3 CONCEPTUALIZATION AND CLINICAL FORMULATION

5.26.3.1 Multimodal Assessment

5.26.3.1.1 The medical puzzle

There are many obstacles to the accurate identification of FDs, starting with the deception on the part of the patient and the lack of an accurate history, the patient's cooperation, or a meaningful medical picture. The episodic and emergency nature of many of the illnesses, the involvement of multiple physicians and institutions, and the difficulty for the physician of even contemplating deception on the part of a youngster or a seemingly exemplary parent adds to the challenge of identifying medical deception in a timely way before great harm is done (Guandolo, 1985). Schreier and Libow (1993a) also detail many of the ways in which the susceptible physician is drawn into the patient's imposture and is sometimes blinded to what otherwise appears to be rather obvious and blatant falsification. Often it is the close observation by nursing or ward staff of peculiar family interactions and inappropriate affect on the part of the patient or parent that finally alerts medical staff to take a very different perspective on the intractable medical problem.

In any case, the first step in the assessment of either FD or FDbP is a multidisciplinary team meeting of all involved physicians, social workers, psychologists, specialists, consultants, nurses, and other team members, in order to pool data, share observations, identify discrepancies, and develop a plan to further pursue investigation of possible deception. While it is often not possible to find direct evidence of fabrication, this is most desirable as it allows for a clear and unambiguous response from the health care team, and possible legal action in the case of child abuse. In the case of childhood factitious disorder, direct evidence can be gathered through such ruses as checking serial numbers on thermometers to identify planted ones (Herzberg & Wolff, 1972), and comparing urine specimens collected under observed and unobserved conditions (Meadow, 1977). In many pediatric hospitals surreptitious video surveillance is now being used under carefully controlled conditions for collecting evidence of medical abuse. Parents have been caught on camera in the act of suffocating as well as poisoning their young children (Samuels, McClaughlin, Jacobson, Poets, & Southall, 1992).

However, in many cases the deception can only be inferred through a careful process of eliminating all other possible medical explanations for the child's symptoms. The peculiar, sporadic, physiologically impossible, or unresponsive nature of the symptoms may be the major clues available to the alert physician. There are a number of cases of young children in which a direct confrontation has resulted in the child's confession of the fabrication (Abe et al., 1984; Croft & Jervis, 1989; Lovejoy et al., 1971), but in the case of older FD patients or

FDbP parents, there is often flight, denial, and lack of cooperation with treatment. In the case of child victims of serious medical abuse, the separation of the child from the parent often offers the most important diagnostic verification of the fabrication, as the child's condition rapidly improves and symptoms disappear in the absence of the suspected perpetrator. However, there are often legal and ethical obstacles to the separation of suspected victims from their mothers, and even the evidence of child's improvement in the mother's absence is often challenged in court.

5.26.3.1.2 Psychological assessment

If the patient is amenable to a psychological interview, the psychologist can make a significant contribution to the understanding of the patient's dynamics and treatment. Particularly important for these patients is the gathering of a detailed and thorough history, both psychosocial and medical, on the child patient as well as siblings and parents. Nothing can be taken at face value, and all history, particularly of previous illness and dramatic events (such as deaths, fires, car accidents), should be verified through other sources including detailed medical records from other institutions. Many therapists and examiners make the mistake of failing to question and verify *all* important background information provided by FD patients, which only further complicates the understanding of the etiology, dynamics, and personal issues of the patient.

Questions about the patient's motivation, chronicity, and conscious intent to deceive can be fruitfully addressed in psychological interviews and family observations. In particular, an exploration of the patient's relationship to his/her physicians, and the day-to-day impact of the "illness" role on the patient's life may help clarify the function of the deception and the needs being met. An exploration of the patient's current family life, relationships with parents, siblings, and significant others can be useful in clarifying important influences, reinforcers, and models for the sick role. Children are likely to be more candid and less deceptive in their presentation, as they may be less aware of the implications of what they reveal.

Psychological testing instruments can contribute to a comprehensive picture of the patient's functioning, particularly when the examiner uses instruments that are less susceptible to conscious deception. The dissimulation scales of the Minnesota Multiphasic Personality Inventory (MMPI-II) and projective tests such as the Rorschach and the Thematic Apperception Test have been particularly useful in this regard. For example, Aduan et al., (1979) reported psychological test results for five adult FD cases and elicited themes of inhibited expression of anger, affective deprivation, faulty concepts of identity, and difficulty separating self from others. They described FD patients as generally suffering from underlying severe personality disorder " ... characterized by hostility, dependency, imposture and poor impulse control with self-destructive acting-out" (p. 238). A review of psychological test data available on adult FD patients (Schreier & Libow, 1993a) found that identity confusion, and sexual and aggressive impulse control issues were significant; defenses included acting out, projection, somatization, and denial. Only limited and mostly case study test data has been available on child FD patients.

There has also been limited psychological test data presented for child victims of FDbP. Ojeda-Castro (1995) recommends thorough developmental and psychological assessments of child victims to assess for developmental delays due to chronic illness, multiple hospitalizations, or deprived environments. Other important concerns for these children are dependency issues, separation anxiety, issues of bodily integrity, and issues of trust and attachment to caregivers. However, it must be noted that the diagnosis of FD or FDbP cannot be made on the basis of a psychological evaluation alone, as there is no uniform profile of such patients, and many manage to present themselves as fairly "normal."

5.26.3.2 Treatment Formulation

The treatment process begins with the open acknowledgment directly to the patient that his/her caregivers suspect deception. The first (and often the last) phase of treatment is confrontation with the patient and family, when the physicians and health care team reveal their findings and directly address the factitious illness. This confrontation is essentially unavoidable in the case of fabricated illness being actively treated by medical staff, although milder forms of symptom exaggeration can sometimes be addressed more indirectly to help the patient "save face" while referring him/her on to treatment. In the case of illness exaggeration, Richtsmeier and Waters (1984) suggest it is not very useful to directly challenge a "family myth" or explain what is "really going on" if the family is unable to hear it. Instead they recommend developing trust, with the physician continuing to meet with the patient and gradually exploring psychological issues, hopefully eventually resulting in the acceptance of psychological referrals.

The approach selected for the confrontation of the blatantly fabricating patient is critical to the possibility of forming an alliance with the family and maintaining their accessibility for psychotherapeutic treatment. Without a successful confrontation, there is no possibility of ongoing follow-up. A successful confrontation does not necessarily require an admission by the patient, but should at least conclude with an agreement to seek psychological services. Unfortunately, it may not be possible, even with the most carefully chosen team and therapeutic approach, to successfully join with all FD or FDbP patients. In fact, it is very common for cases described in elaborate medical detail in the literature to conclude with the unfortunate fact that the patient was confronted, fled the facility, and was never heard from again (e.g., Rodriguez-Moreno et al., 1990; Witt & Ginsberg-Fellner, 1981). Many of the younger FD patients described in the literature apparently do confess fairly readily to their fabrications when confronted (Abe et al., 1984; Abrol et al., 1990; Croft & Jervis, 1989; Lovejoy et al., 1971; White et al., 1966), which is much less likely in older patients.

Given that denial and flight are common responses to confrontation of adolescents and FDbP perpetrators, practitioners have outlined a number of suggestions for approaching the confrontation to maximize success. Wedel (1971) recommends a "therapeutic confrontation" in which a trusted, supportive social worker is included in the team meeting, which should convey concern and acceptance to decrease the patient's sense of shame and anxiety. Many authors address the need to avoid a punitive, rejecting stance on the part of the medical team, as these feelings are common reactions of health care providers duped into active collusion with unnecessary treatment. Hollender and Hersh (1970) developed an approach which includes the pairing of a confronting physician and a supportive psychiatrist to maximize their success. It is important for the team to include at least one psychiatrist or psychologist, and for an emergency plan to be ready in case the patient becomes actively suicidal (Schreier & Libow, 1993a). If the patient agrees to psychotherapeutic treatment, it is recommended that the contract include the patient's agreement to confine all his/her medical treatment to a single facility or group of physicians familiar with the disorder, to help contain the behavior. In the case of FDbP, this restriction is often built directly into court orders for parent–child reunification.

If the confrontation is successful and the patient remains accessible, the task of treatment remains formidable. Some authors recommend initial psychiatric hospitalization for long-term in-patient care (Plassmann, 1994b) if that is feasible. However, it is much more common for out-patient psychotherapy to be the primary mode of treatment. A combination of behavioral and psychodynamic approaches have been recommended, with an emphasis on exploring the nature of the patient's historical relationships with physicians and authority figures and developing insight into the use of illness in daily life. There is little consensus on successful treatment approaches to the FD patient, and the full range of psychotherapeutic approaches, from behavioral to psychodynamic to family systems, have been utilized, with limited success.

5.26.4 TREATMENT

5.26.4.1 Psychosocial Treatments

5.26.4.1.1 Psychotherapy for factitious disorder

There is virtually no treatment literature available describing psychotherapy for children or adolescents who fabricate illness, and this is especially surprising compared with the considerably more extensive literature on child somatization disorder and its treatment (Campo & Fritsch, 1994; Garralda, 1996). This is partially due to the fact that many of the children fabricating illness make early admissions when confronted. Some of the reports on FD report that younger children, in particular, do not seem to repeat this behavior once confronted, although it is unlikely that many of these patients have been followed prospectively in any systematic way. Suggestions about therapy for these children that have been made, not surprisingly, include the need for involvement of the family system, including Herzberg and Wolff's (1972) recommendation for parental couples therapy to avoid sabotaging the child's treatment. White et al. (1966) suggest that the clinician needs to allay parental concerns about illness, and help the parent(s) minimize secondary gains by insisting the child resume full activity and school responsibilities.

Once again the adult FD literature is more informative on treatment and includes several descriptions of individual therapies. A German psychiatrist, Plassmann (1994c), reviewed approximately 350 papers on the treatment of FD and found that the "vast majority of psychotherapies undertaken do not go beyond an initial phase consisting of several psychiatric interviews" (p. 96). He found that the major issues in successful treatment involved the problem of how to confront the patient, selection of the appropriate setting (in- or out-patient), the establishment of a positive working relationship, and countertransference issues. His review found that the most frequently used therapeutic methods were analytically oriented therapy, hypnosis, psychoanalysis, and group therapy. His paper also described his own psychodynamic treatment of 12 patients, mostly in in-patient or long-term out-patient therapy, which utilized an "approaching phase," an "introjection phase," and a "separation phase."

Most other authors report on only single cases or very small samples. For example, Schoenfeld, Margolin, and Baum (1987) described the long-term treatment of a young adult patient with a combination of dynamic and supportive therapy. They described the value of working with a nurturing female therapist as a "corrective experience," and they used the patient's children to generate treatment motivation. Stone (1977) treated a young woman in an in-patient psychiatric unit, although she checked herself out against medical advice Two male adult FD patients were treated on an in-patient unit in Ireland with a variety of "communication techniques" and the fostering of a highly dependent relationship with the therapist (O'Shea & McGennis, 1982). Fras (1978) described a very supportive, cautious approach to psychotherapy which avoided challenging the patient's denial and resistance too vigorously. A supportive four-month treatment of a patient was described by Jamieson, McKee, and Roback (1979) focused on reinforcing reality-testing, although the patient was apparently still producing symptoms on follow-up. Schreier and Libow (1993a) and Eisendrath and Feder (1996) also report a variety of other approaches that have been used with FD patients including aversive therapy, biofeedback, paradoxical approaches, systems interventions, and behavioral contingencies.

It should be noted that countertransference issues for the therapist working with FD patients can be a significant obstacle to joining with the patient and establishing a positive working relationship. Schreier and Libow (1993a) present two detailed cases narrated directly by the therapists who treated the MBPS mothers which highlighted the therapists' own difficulties even recognizing their patients' manipulations. The patients' efforts at impostering the role of "cooperative patient" apparently sabotaged the therapeutic efforts of even these motivated and well-intentioned therapists.

5.26.4.1.2 Psychotherapy for factitious disorder by proxy

As in the treatment of FD, there is no single specific treatment for the parent or adult with FDbP, but instead a variety of psychotherapeutic techniques and approaches, assuming the patient is still amenable for treatment after the initial confrontation. Since few mothers with this disorder immediately acknowledge the problem, most enter therapy guardedly, or under court order. One of the few successful treatment cases was described by Nicol and Eccles (1985) who detailed a one-year treatment of a mother, focused on uncovering and interpreting her fantasies related to her ambivalent relationship with her father and her family culture of exaggerating illness. Lansky and Erickson (1974) described couples work in a family in which the child was being poisoned by the mother with caustic agents and medication.

Palmer and Yoshimura (1984) provided unusually detailed psychosocial histories for both a mother and her child victim of FDbP, and suggested that treatment of such children should be intensive and reality-focused, emphasizing the "gradual removal of pathologic defense mechanisms and on expanding the child's understanding of his or her state of physical health, of the parent's role in the creation of the illness, and of the parent's psychological state" (pp. 507–508). They described the issues of bodily integrity, feelings of helplessness and overly close mother–child interactions that emerged during a play therapy evaluation of their nine-year-old patient.

Suggestions for treatment of very young child victims of factitious illness are offered by Ojeda-Castro (1995), who described a play therapy approach for helping victims resolve issues such as the violation of basic trust, attachment problems, the "incongruous mother image," and the distorted self-image. The sessions begin with thorough cognitive and developmental evaluations of the child, move into regressive play therapy, and gradually work towards joint sessions with the mother, if reunification is anticipated.

One of the more unique, specific approaches to treating children involved in FDbP is offered by Sanders (1995). She proposes a narrative approach to working with families in which children collude in the exaggeration of illness. The goal of the therapist is to challenge the family's "story of illness" and help them develop an "alternative story" and means of coping. Involvement of the parents is considered very important to the success of treatment, and this approach seems mainly useful for more cooperative families.

5.26.4.1.3 Efficacy of psychosocial treatment

Unfortunately the difficulty of assessing successful outcome is probably greater in FDs than in most other psychiatric disorders, due to the ever-present risk of deception and the difficulty of accurately verifying the cessation of chronic or compulsive repetition of the behavior. Furthermore, the simple absence of the originally falsified medical symptom(s) may not represent the best outcome measure. As Campo and Fritsch (1994) remind us, "Outcome for patients with troubling somatization early in life may not be as positive when functional and psychiatric status is evaluated at follow-up, as opposed to the presence or absence of the original physical symptom" (p. 1231). That is, a child who stops fabricating an illness after being confronted or even treated in therapy is not necessarily psychologically healthy upon termination of this behavior. Furthermore, one has to questions some "positive" outcomes that have been reported for factitious disorder patients, such as the patient described by Fras (1978) who stopped inducing diabetic problems and started a nursing career, and another (Schoenfeld et al., 1987) who moved from factitious illness to work as a paramedic!

Certain factors do seem to point to better prognosis. The less chronic, more episodic FD patients seem to have a better response to both confrontation and therapy than Munchausen patients. Folks and Freeman (1985) remind us that patients with underlying mood or anxiety disorders have a better prognosis for treatment than those with coexisting personality disorders. We must treat any psychopathology such as depression, conversion symptoms or psychosis, often before the FD itself can be addressed. At the same time, factitious psychological disorders can coexist with factitious physical disorders and must also be ruled out before treatment is derailed.

Above and beyond all of these cautions, it must be noted that the efficacy of treatment approaches for both FD and FDbP is questionable. A handful of papers describe limited success on a case study basis. Even Plassmann's (1994c) report of an intensive treatment study involving 24 patients found that only half the patients even accepted the offer of treatment, and only three of the patients showed a disappearance of factitious symptomatology and marked changes in social competence as a result of treatment.

Reports on the effectiveness of treating mothers with MBPS are equally cautionary. Schreier and Libow's (1993a) review of the literature reports very limited success in the treatment of this disorder, as do McGuire and Feldman (1989), who note little effect of treatment on mothers' insight or medical abuse behavior. They also point out that even if the illness fabrications cease, the child victim may continue to manifest significant psychological problems and may eventually begin to seek inappropriate medical care him or herself. However, the increase in awareness, identification, and treatment of FD cases is beginning to offer some hope for the development of more effective treatments.

5.26.4.2 Pharmacological Treatments

5.26.4.2.1 Uses of pharmacotherapy

There are only a handful of reports on the use of pharmacotherapy in the treatment of FDs, and they generally involve treatment of the patient's concurrent mood or anxiety disorders. It has been suggested by Earle and Folks (1986) that antipsychotic medication can be useful for the treatment of brief psychotic episodes in FD patients, just as antidepressants may be appropriate for depressed patients. One apparently successful example in the literature is Schoenfeld et al.'s (1987) use of antianxiety and antidepressant medication in conjunction with dynamic therapy for an FD patient. One of the only known uses of pharmacological agents directly for treatment of FD is described by McDonald et al. (1979) who used sodium amytal ("truth serum") interviews to gather more accurate information from their patients. Unfortunately, they found that it had no real value for treatment itself.

5.26.4.2.2 Efficacy of pharmacotherapy

There is no known pharmacotherapy treatment designed specifically for FD or FDbp, although the use of psychoactive medication may be useful in establishing and maintaining the treatment relationship and treating coexisting psychiatric disorders in these patients.

5.26.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

For a variety of reasons already outlined, the study of FD in children and adolescents is in its very early stages. To begin with, untold numbers of cases of fabricating children and adolescents are likely being misidentified under the broader category of somatizers when their deliberate deceptions are not detected. Thus, there has been a limited pool of actual cases for study and it is likely that only the most blatant or persistant fabrications have been identified. An area ripe for more study is the clarification of the relationship between somatization, malingering, and FD in children, to determine if these represent qualitatively different diagnostic entities or whether there is considerable overlap. Further research may clarify whether there is even a developmental progression from one to the next. Furthermore, we need to better understand the similarities and differences between young people who consciously exaggerate their symptoms and those who actively induce illness; these distinctions are equally important to understand in the mothers who use illness exaggeration and those who engage in the active induction of illness in their children. The establishment of better diagnostic tools for assessing intent, motivation, and conscious deception, and a better program of education for physicians in recognizing the problem would be extremely helpful in collecting better data about these interesting and troubling cases of children who deliberately induce their own illnesses.

Beyond diagnostic considerations, it is clear that our understanding of the etiology and treatment of these disorders is also in its infancy. Earlier and more intensive involvement of pediatric psychologists in the assessment of FD patients, perhaps even before some of the families are confronted and lost to follow-up, would allow for the collection of more detailed psychosocial histories that should include not only the child's detailed medical, developmental, and social history, but the medical and psychosocial histories of mother, father, and all siblings. The intriguing hints in the literature of possible multigenerational replication of FD and FDbP in families can only be fruitfully explored through examination of the family systems and backgrounds of as many family members as possible. In particular, the possible role of mothers with FDbP raising children who themselves later become active fabricators of illness is a fascinating one in our understanding of FD and its genesis in young adults. Further, it suggests that research into the ways that older children may learn to either collude with, or resist, the efforts of a fabricating parent can point directions to effective treatment approaches for victims of MBPS.

One of the suggestions that has been discussed in child abuse and child protection circles is that of developing some type of national registry of FDbP cases, so that better data can be collected on the prevalence, psychosocial histories, and successful and unsuccessful treatment approaches to these cases. Psychological test data is beginning to be collected by individual clinicans working with FDbP parents and by-proxy patients, which, if systematically collected and shared, might help develop profiles useful for clinicians in identifying FD cases. Nation-wide registries might also save our health care systems and child protective services considerable resources, and prevent continued harm to vulnerable children by allowing earlier identification and follow-up of chronic, itinerant FD patients.

5.26.6 SUMMARY

Physicians and clinicians have been sharing their experiences and observations about FDs in the literature for several decades. These simulated illnesses, intentionally exaggerated or induced by the youngster and/or the parent, are often medically treated for long periods of time before finally being discovered as factitious. Physicians are often drawn into believing in these deceptions for many complex reasons, and can unintentionally cause considerable harm to the child through unnecessary treatments, procedures, hospitalizations and the fostering of the child's career as an invalid. While the discovery of fabrications by preadolescent children is generally fairly rapid and tends to result in confession and cessation of the deception, the treatment of factitious illness in adolescents and adults and of factitious illness by proxy is often a difficult, frustrating process complicated by widespread confusion about the diagnosis, and lack of systematic data on etiology and treatment of these disorders. There may be considerably more cases of FD in our health care systems than are commonly recognized, and there appear to be family systems which perpetuate the process of misusing medical resources to meet complex psychological needs, resulting in multiple generations of family somatizers or fabricators. The development of a better research database on family histories, psychological test findings and responses to different treatment approaches could help reduce the serious physical and psychological damage resulting from these forms of self-induced illness, and perhaps prevent some children from progressing to lifetime careers of medical abuse.

5.26.7 REFERENCES

Abe, K., Shinozima, K., Okuno, A., Abe, T., & Ochi, H. (1984). Munchausen's syndrome in children: Bizarre clinical and laboratory features. *Acta Paediatrica Japan, 26,* 539–543.

Abrol, R. P., Heck, A., Gleckel, L., & Rosner, F. (1990). Self-induced hematuria. *Journal of the National Medical Association, 82*(2), 127–128.

Aduan, R. P., Fauci, A. S., Dale, D. C., Herzberg, J. H., & Wolff, S. M. (1979). *Annals of Internal Medicine, 90,* 230–242.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Asher, R. (1951). Munchausen's syndrome. *Lancet, 1,* 339–341.

Baldwin, C. (1996). Munchausen syndrome by proxy: problems of definition, diagnosis and treatment. *Health and Social Care in the Community, 4*(3), 159–165.

Barsky, A. J. (1992). Amplification, somatization and the somatoform disorders. *Psychosomatics, 33*(1), 28–34.

Bools, C. N., Neale, B. A., & Meadow, S. R. (1993). Follow up of victims of fabricated illness (Munchausen syndrome by proxy). *Archives of Diseases of Childhood, 69,* 625–630.

Campo, J. V., & Fritsch, S. L. (1994). Somatization in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 33*(9), 1223–1235.

Carney, M. W. P. (1980). Artifactual illness to attract medical attention. *British Journal of Psychiatry, 136,* 542–547.

Chin, P., & Breu, G. (1991). The murder that never was. *People Magazine, Dec. 16,* 111–116.

Cramer, B., Gershberg, M. R., & Stern, M. (1971). Munchausen syndrome: Its relationship to malingering, hysteria, and the physician–patient relationship. *Archives of General Psychiatry, 24,* 573–578.

Croft, R. D., & Jervis, M. (1989). Munchausen's syndrome in a 4 year old. *Archives of Disease in Childhood, 64,* 740–741.

Donald, T., & Jureidini, J. (1996). Munchausen syndrome by proxy: Child abuse in the medical system. *Archives of Pediatric and Adolescent Medicine, 150,* 753–758.

Earle, J. R., & Folks, D. G. (1986). Factitious disorder and coexisting depression: a report of successful psychiatric consultation and case management. *General Hospital Psychiatry, 8,* 448–450.

Eisendrath, S. J. (1996). Current overview of factitious physical disorders. In M. D. Feldman & S. J. Eisendrath (Eds.), *The spectrum of factitious disorders* (pp. 21–36). Washington, DC: American Psychiatric Press.

Eisendrath, S. J., & Feder, A. (1996). Management of factitious disorders. In M. D. Feldman & S. J. Eisendrath (Eds.), *The spectrum of factitious disorders* (pp. 195–213). Washington, DC: American Psychiatric Press.

Eminson, D. M., & Postlewaite, R. J. (1992). Factitious illness: recognition and management. *Archives of Disease in Childhood, 67,* 1510–1516.

Feldman, M. D., & Eisendrath, S. J. (Eds.) (1996). *The spectrum of factitious disorders.* Washington, DC: American Psychiatric Press.

Feldman, M. D., & Smith, R. (1996). Personal and interpersonal toll of factitious disorders. In M. D. Feldman & S. J. Eisendrath (Eds.), *The spectrum of factitious disorders,* pp. 175–194. Washington, DC: American Psychiatric Press.

Fialkov, M. J. (1984). Peregrination in the problem pediatric patient. *Clinical Pediatrics, 23,* 571–575.

Fink, P. (1992). Physical complaints and symptoms of somatizing patients. *Journal of Psychosomatic Research, 36*(2), 125–136.

Fisher, G. C., & Mitchell, I. (1995). Is Munchausen syndrome by proxy really a syndrome? *Archives of Disease in Childhood, 72*(6), 530–534.

Folks, D. G, & Freeman, A. M. (1985), Munchausen's syndrome and other factitious illness. *Psychiatric Clinics of North America, 8*(2), 263–278.

Fras, J. (1978). Factitial disease: an update. *Psychosomatics, 19*(2), 119–122.

Friedman, S. B. (1973). Conversion symptoms in adolescents. *Pediatric Clinics of North America, 20*(4), 873–882.

Garralda, M. E. (1996) Somatisation in children. *Journal of Child Psychology and Psychiatry, 37*(1), 13–33.

Gavin, H. (1838). *On feigned and factitious diseases.* Edinburgh: Edinburgh University Press.

Goodlin, R. C. (1985). Pregnant females with Munchausen syndrome. *American Journal of Obstetrics and Gynecology, 153,* 207–210.

Goss, P. W., & McDougall, P. N. (1992). Munchausen syndrome by proxy: a cause of preterm delivery. *Medical Journal of Australia, 157,* 814–817.

Gilbert, R. W., Pierse, P. M., & Mitchell, D. P. (1987). Cryptic otalgia: A case of Munchausen syndrome in a pediatric patient. *Journal of Otolaryngology, 16*(4), 231–233.

Guandolo, V. L. (1985). Munchausen syndrome by proxy: an out-patient challenge. *Pediatrics, 75,* 526–530.

Hay, G. G. (1983). Feigned psychosis: A review of the simulation of mental illness. *British Journal of Psychiatry, 43,* 8–10.

Herzberg, J. H. & Wolff, S. M. (1972) Chronic factitious fever in puberty and adolescence: A diagnostic challenge to the family physician. *Psychiatry in Medicine, 3,* 205–212.

Holborow, P. L. (1985). A variant of Munchausen syndrome by proxy. *Journal of the American Academy of Child Psychiatry, 24*(2), 238.

Hollender, M. D., & Hersh, S. R. (1970). Impossible consultation made possible. *Archives of General Psychiatry, 23,* 343–345.

Jacobi, A. (1895). Hyperthermy in a man up to 148 degrees F. *Transactions of the Association of American Physicians, 10,* 159–191.

Jamieson, R., McKee, E., & Roback, H. (1979). Munchausen's syndrome: an unusual case. *American Journal of Psychotherapy, 33*(4), 616–621.

Janofsky, J. S. (1986). Munchausen syndrome in a mother and daughter: an unusual presentation of folie deux. *Journal of Nervous and Mental Disease, 174*(6), 368–370.

Jonas, J. M., & Pope, H. G., Jr. (1985). The dissimulating disorders: a single diagnostic entity? *Comprehensive Psychiatry, 26,* 58–62.

Katz, R. L., Mazer, C., & Litt, I. F. (1985). Anorexia nervosa by proxy. *The Journal of Pediatrics, 107*(2), 247–248.

Knockaert, D. C., Vanneste, L. J., Vanneste, S. B., & Bobbaers, H. J. (1992). Fever of unknown origin in the 1980s. An update of the diagnostic spectrum. *Archives of Internal Medicine, 152,* 51–55.

Krener, P. (1994). Factitious disorders and the psychosomatic continuum in children. *Current Opinion in Pediatrics, 6,* 418–422.

Krener, P., & Adelman, R. (1988). Parent salvage and parent sabotage in the care of chronically ill children. *American Journal of Diseases of Children, 142,* 945–951.

Lansky, S. B., & Erickson, H. M. (1974). Prevention of child murder. *Journal of the American Academy of Child Psychiatry, 13,* 691–698.

Levin, A. V., & Sheridan, M. S. (1995). *Munchausen syndrome by proxy: Issues in diagnosis and treatment.* New York: Lexington Books.

Libow, J. A. (1995). Munchausen by proxy victims in adulthood: A first look. *Child Abuse and Neglect, 19*(9), 1131–1142.

Libow, J. A., & Schreier, H. A. (1986). Three forms of factitious illness in children: When is it Munchausen syndrome by proxy? *American Journal of Orthopsychiatry, 56,* 602–611.

Light, M. J., & Sheridan, M. S. (1990). Munchausen syndrome by proxy and apnea—A survey of apnea programs. *Clinical Pediatrics, 29*(3), 162–168.

Livingston, R. (1992). Somatization in child, adolescent, and family psychiatry. *Psychiatric Medicine, 10*(3), 13–23.

Livingston, R. (1993). Children of people with somatization disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 32*(3), 536–544.

Lovejoy, F. H., Marcuse, E. K., & Landrigan, P. J. (1971). Two examples of purpura facititia. *Clinical Pediatrics, 10*(3), 183–184.

Masterson, J., Dunworth, R., & Williams, N. (1988). Extreme illness exaggeration in pediatric patients: A variant of Munchausen by proxy? *American Journal of Orthopsychiatry, 58*(2), 188–195.

McClure, R. J., Davis, P. M., Meadow, S. R., & Sibert, J. R. (1996). Epidemiology of Munchausen syndrome by proxy, non-accidental poisoning, and non-accidental suffocation. *Archives of Disease in Childhood, 75,* 57–61.

McDonald, A., Kline, S. A., & Billings, R. F. (1979). The limits of Munchausen's syndrome. *Canadian Journal of Psychiatry, 24,* 323–328.

McGuire, L. T., & Feldman, K. W. (1989). Psychologic morbidity of children subjected to Munchausen syndrome by proxy. *Pediatrics, 83*(2), 289–292.

Meadow, R. (1977). Munchausen syndrome by proxy. The hinterland of child abuse. *Lancet, 2,* 343–345.

Meadow, R. (1990). Suffocation, recurrent apnea and sudden infant death. *The Journal of Pediatrics, 117,* 351–357.

Meadow, S. R. (1984). Commentary. *Archives of Disease in Childhood, 59,* 156.

Morley, C. J. (1995). Practical concerns about the diagnosis of Munchausen syndrome by proxy. *Archives of Disease in Childhood, 72*(6), 528–530.

Nadelson, T. (1979). The Munchausen spectrum: borderline character features. *General Hospital Psychiatry, 1,* 11–17.

Nichol, A. R., & Eccles, M. (1985). Psychotherapy for Munchausen syndrome by proxy. *Archives of Disease in Childhood, 60,* 344–348.

Ojeda-Castro, M. (1995). Issues in play therapy with victims of Munchausen by proxy syndrome. *Presented at the 103rd Annual Convention of the American Psychological Association,* New York, August, 1995.

O'Shea, B., & McGennis, A. (1982). The psychotherapy of Munchausen's syndrome. *Irish Journal of Psychotherapy, 1,* 17–19.

Palmer, A. J., & Yoshimura, G. J. (1984), Munchausen syndrome by proxy. *Journal of the American Academy of Child Psychiatry, 23*(4), 503–508.

Paperny, D., Hicks, R., & Hammar, S. L. (1980). Munchausen's syndrome. *American Journal of Diseases of Childhood, 134,* 794–795.

Phillips, M. R., Ward, N. G., & Ries, R. K. (1983). Factitious mourning: Painless patienthood. *American Journal of Psychiatry, 140,* 420–425.

Plassmann, R. (1994a). The biography of the factitious-disorder patient. *Psychotherapy Psychosomatics, 62,* 123–128.

Plassmann, R. (1994b). Munchhausen syndromes and factitious diseases. *Psychotherapy Psychosomatics, 62,* 7–26.

Plassmann, R. (1994c). Impatient and outpatient long-term psychotherapy of patients suffering from factitious disorders. *Psychotherapy Psychosomatics, 62,* 96–107.

Pope, H. G., Jonas, J. M., & Jones, B. (1982). Factitious psychosis: phenomenology, family history and long-term outcome of 9 patients. *American Journal of Psychiatry, 139*(11), 1480–1486.

Raymond, C. A. (1987). Munchausen's may occur in younger persons. *Journal of the American Medical Association, 257*(24), 3332.

Reich, P., & Gottfried, L. A. (1983). Factitious disorders in a teaching hospital. *Annals of Internal Medicine, 99,* 240–247.

Reich, P., Lazarus, J. M., Kelly, M. J., & Rogers, M. P. (1977). Factitious feculent urine in an adolescent boy. *Journal of the American Medical Association, 238*(5), 420–421.

Richtsmeier, A. J., Jr., & Waters, D. B. (1984). Somatic symptoms as family myth. *American Journal Diseases of Childhood, 138,* 855–857.

Rodriguez-Moreno, J., Ruiz-Martin, J. M., Mateo-Soria, L., Rozadilla, A., & Roig-Escofet, D. (1990). Munchausen's syndrome simulating reflex sympathetic dystrophy. *Annals of Rheumatic Diseases, 49,* 1010–1012.

Rosenberg, D. (1987). Web of deceit: a literature review of Munchausen syndrome by proxy. *Child Abuse and Neglect, 11,* 547–563.

Roth, D. (1990). How "mild" is mild Munchausen syndrome by proxy? *Israel Journal of Psychiatry and Related Sciences, 27*(3), 160–167.

Rumans, L. W., & Vosti, K. L. (1978). Factitious and fraudulent fever. *American Journal of Medicine, 65,* 745–755.

Sakula, A. (1978). Munchausen: fact and fiction. *Journal of Royal College of Physicians, 12*(3), 286–292.

Sale, I., & Kalucy, R. (1980). An observation on the genesis of Munchausen syndrome. A case report. *Australian and New Zealand Journal of Psychiatry, 14,* 61-64.

Samuels, M. P., McClaughlin, W., Jacobson, R. R., Poets, C. F., & Southall, D. P. (1992). Fourteen cases of imposed upper airway obstruction. *Archives of Disease in Childhood, 67,* 162–170.

Sanders, M. J. (1995). Symptom coaching: Factitious disorder by proxy with older children. *Clinical Psychology Review, 15*(5), 423–442.

Schoenfeld, H., Margolin, J., & Baum, S. (1987). Munchausen syndrome as a suicide equivalent: abolition of syndrome by psychotherapy. *American Journal of Psychotherapy, 41*(4), 604–612.

Schreier, H. A., & Libow, J. A. (1993a). *Hurting for love: Munchausen by proxy syndrome.* New York: Guilford Press.

Schreier, H. A., & Libow, J. A. (1993b). Munchausen by proxy syndrome: diagnosis and prevalence. *American Journal of Orthopsychiatry, 63*(2), 318–321.

Sneed, R. C., & Bell, R. F. (1976). The dauphin of Munchausen: Factitious passage of renal stones in a child. *Pediatrics, 58,* 127–130.

Spiro, H. R. (1968). Chronic factitious illness. Munchausen's syndrome. *Archives of General Psychiatry, 18,* 569–579.

Stankler, L. (1977). Factitious skins lesions in a mother and two sons. *British Journal of Dermatology, 97,* 217–219.

Stegman, M. S., & Aaron, W. S. (Eds.) (1996). *ICD-9-CM Codebook.* Reston, VA: St. Anthony Publishing.

Steinbeck, A. W. (1961). Haemorrhagica histrionica—the bleeding Munchausen syndrome. *Medical Journal of Australia, 48,* 451–456.

Stone, M. H. (1977). Factitious illness. *Bulletin of the Meninger Clinic, 41*(3), 239–254.

Taylor, S., & Hyler, S. E. (1993). Update on factitious disorders. *International Journal of Psychiatric Medicine, 23,* 81–94.

Tojo, A., Nanba, S., Kimura, K., Hirata, Y., Matsuoka, H., Sugimoto, T., Watanabe, N., & Ohkubo, A. (1990). Factitious proteinuria in a young girl. *Clinical Nephrology, 33*(6), 299–302.

Warner, J., & Hathaway, M. J. (1984). Allergic form of Meadow's syndrome (Munchausen by proxy). *Archives of Disease in Childhood, 59,* 151–156.

Wedel, K. R. (1971). A therapeutic confrontation approach to treating patients with factitious illness. *Social Work, 16*(2), 69–73.

White, J. G., Pearson, H. A., & Coddington, R. D. (1966). Purpura factitia: Attracting attention by self-inflicted lesions. *Clinical Pediatrics, 5*(3), 157–160.

Witt, M. E., & Ginsberg-Fellner, F. (1981). Prednisone-induced Munchausen syndrome. *American Journal of Diseases of Childhood, 135,* 852–853.

Wood, P. R., Fowlkes, J., Holden, P., & Castro, D. (1989). Fever of unknown origin for six years: Munchausen syndrome by proxy. *Journal of Family Practice, 28*(4), 391–395.

Woolcott, P., Aceto, T., Rutt, C., Bloom, M., & Glick, R. (1982). Doctor shopping with the child as proxy patient: A variant of child abuse. *Journal of Pediatrics, 101,* 297–301.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

5.27
Schizophrenia in Children and Adolescents

CINDY M. YEE and MARIAN D. SIGMAN
University of California, Los Angeles, CA, USA

5.27.1 INTRODUCTION 703
5.27.2 SCHIZOPHRENIA AND OTHER PSYCHOTIC DISORDERS 704
  5.27.2.1 Phenomenology 704
    5.27.2.1.1 A case illustration 707
  5.27.2.2 Diagnostic Features and Differential Diagnosis 709
  5.27.2.3 Prevalence 712
  5.27.2.4 Etiology 712
  5.27.2.5 Developmental Course 714
5.27.3 CONCEPTUALIZATION AND CLINICAL FORMULATION 715
  5.27.3.1 Multimodal Assessment 715
  5.27.3.2 Treatment Formulation 716
5.27.4 TREATMENT 717
  5.27.4.1 Psychosocial Treatments 717
  5.27.4.2 Pharmacological Treatments 718
  5.27.4.3 Efficacy of Treatments 718
5.27.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 719
5.27.6 SUMMARY 720
5.27.7 REFERENCES 721

5.27.1 INTRODUCTION

Childhood-onset schizophrenia and adolescent-onset schizophrenia are among the most devastating of mental disorders to afflict the young. Fortunately, the incidence of schizophrenia in childhood and early adolescence is quite rare with the onset of schizophrenia typically occurring during late adolescence and early adulthood. In current clinical research and practice, the diagnostic criteria for schizophrenia are the same for children, adolescents, and adults, thereby allowing for developmental effects on the presentation and expression of symptoms. Schizophrenic illness has been identified in children as young as three years of age (Russell, Bott, & Sammons, 1989), although psychotic symptoms do not usually appear or are not detected reliably until a child is at least six or seven years of age (Russell, 1994; Werry, 1996). Schizophrenia in children and adolescents, sometimes referred to as very-early-onset and early-onset schizophrenia, respectively (Werry, 1996), are often treated as distinct disorders although there are conflicting reports about the nature of the differences between the two forms of the illness (R. F. Asarnow & J. R. Asarnow, 1994).

Some of the earliest observations of childhood- and adolescent-onset schizophrenia date back to Kraepelin's (1883/1915) description of dementia praecox, a form of mental deterioration that involves disturbances of thought and that can occur even in children. Upon observing such symptoms in a group of children, De Sanctis (1906/1973) described the disorder as *dementia praecocissima*. Bleuler (1911/1950) subsequently introduced the term *schizophrenia* to refer to the same condition in children and adults, as identical diagnostic criteria were being used to categorize both groups. By the 1930s, however, all children with profound mental disturbances, including autism and other psychoses, were placed in a general category that was labeled "childhood schizophrenia" (Fish & Ritvo, 1979). This tradition persisted with *DSM-II* (American Psychiatric Association, 1968) and *ICD-8* (World Health Organization, 1967–1969). However, with the introduction of *DSM-III* (American Psychiatric Association, 1980) and *ICD-9* (World Health Organization, 1978), the field returned to the practice of using the same diagnostic criteria for identifying schizophrenia regardless of age. Given the considerable diagnostic heterogeneity that characterized the broad category of childhood schizophrenia from the 1930s until the 1970s, it can be difficult, if not impossible at times, to meaningfully interpret the earlier literature.

The material covered in the present chapter relies on *DSM-III* through *DSM-IV* (American Psychiatric Association, 1994), *ICD-9* (World Health Organization, 1978), and *ICD-10* (World Health Organization, 1992) diagnostic criteria. By necessity, the primary emphasis of this chapter will be on schizophrenia that has been diagnosed before the age of 12 as early adolescent-onset schizophrenia has received relatively less attention in the clinical research literature to date. It should be noted that some variability exists between clinical research sites with regard to the maximum age allowed for receiving a diagnosis of childhood-onset schizophrenia. The National Institute of Mental Health project on childhood-onset schizophrenia, for instance, does not accept patients with an onset of psychosis beyond the age of 12 while the UCLA Childhood-Onset Schizophrenia Research Program has raised the maximum age from 11 to 14 years. Across sites, schizophrenia in childhood is defined by age and not by physical development because there are considerable individual differences in the age at onset of puberty (Werry, McClellan, Andrews, & Ham, 1994).

5.27.2 SCHIZOPHRENIA AND OTHER PSYCHOTIC DISORDERS

5.27.2.1 Phenomenology

Relying upon more recent diagnostic criteria, there have been reasonably consistent findings across studies on the clinical presentation of schizophrenia in children. Moreover, many of the same qualitative symptoms which are found in children can be seen in adult patients diagnosed with schizophrenia. This is not a complete surprise as the same diagnostic criteria are being applied to children and adults, although the possibility remains that there are distinct developmental differences in the expression of schizophrenia (J. R. Asarnow, 1994). It is noteworthy that any similarity in symptomatology has been obtained despite the inherent difficulty in ascertaining complex, internal symptoms in young children and with recognition of the need to distinguish between clinical symptoms, such as delusions, and normal experiences in childhood, such as vivid fantasies (J. R. Asarnow, 1994; Russell, 1994). In the following discussion, some examples of psychotic symptoms in children are provided; additional examples of symptom expression in preadolescence can be found in Russell (1994).

The most common presenting symptom for children, as with adults, is auditory hallucinations. As shown in Table 1, auditory hallucinations have been observed in approximately 80–100% of cases. Command hallucinations, which are hallucinations that involve telling an individual what to do, are perhaps the most frequent and have been noted in over 60% of children with a schizophrenic disorder (McKenna et al., 1994; Russell et al., 1989). Oftentimes, children report that the commands are negative or violent in nature (e.g., "kick the furniture," "hit your mother," "kill the goldfish"). Other auditory hallucinations involve voices that are conversing with one another (e.g., two stuffed animals speaking with one another), are maintaining a running commentary on the child's thoughts or behaviors (e.g., "It was nice that you helped your brother today"), are making persecutory statements (e.g., calling the child names), or are of a religious nature (e.g., the voice of God or the devil calling the child by name).

Visual hallucinations also appear to occur with considerable frequency and have been reported in approximately 40% or more of children diagnosed with schizophrenia, as shown in Table 1. Some visual perceptions may be reported as accessible only to the child (e.g., "The magician does not want anyone else to see him"), whereas other images are assumed

Table 1 Comparison of studies reporting on clinical symptoms in children with schizophrenia.

			Hallucinations						
Study	*N*	*Age* (years)	*Auditory* (%)	*Visual* (%)	*Tactile* (%)	*Olfactory* (%)	*Delusions* (%)	*Thought disorder* (%)	*Inappropriate or flat affect* (%)
Green, Padron-Gayol, Hardesty, & Bassiri (1992)	38	5–11	84	47	8		55	100	84
Kolvin (1971); Kolvin, Ounsted Humphrey, & McNay (1971)	33	5–15	81	30			57	60	63
McKenna et al. (1994)	19	8–11[a]	100	79	37	21	95	84	95
Russell, Bott, & Sammons (1989); Russell (1994)	35	4–13[b]	80	37	17	6	63	40	74
Spencer & Campbell (1994)	16	5–11	100	69	25		100	81	81
Volkmar, Cohen, Hoshino, Rende, & Paul (1988); Volkmar (1996)	42[c]	4–11[d]	79	28			86	93	71

[a] Subjects ranged in age from 12 to 16 years when screened for participation in the research study but the onset of psychosis was before age 12. [b] Schizophrenia was diagnosed before the age of 12. [c] The total sample includes children with clinical diagnoses of childhood schizophrenia, psychotic illness with schizophrenic features, and schizophrenia of childhood onset. Upon reclassification with *DSM-III* criteria, 14 of the 42 children were diagnosed with schizophrenia; the remainder generally appeared to exhibit schizoid disorder. [d] Approximate estimate of age, derived from the mean and standard deviation provided in the report.

to be visible to anyone present ("Can't you see? The branches are turning into snakes and they are going to get us").

Occurring far less frequently, by contrast, are tactile hallucinations (e.g., "I can feel bugs crawling on my legs") and olfactory hallucinations (e.g., "There's a stinky smell that makes me want to throw up"). As can be seen in Table 1, these hallucinations are typically observed in less than one-third of schizophrenic children. There is some evidence to suggest that non-auditory hallucinations may only occur in those children who experience auditory hallucinations (Kolvin et al., 1971; Russell et al., 1989).

As with adults, delusions are also quite common and have been found to occur in 55–100% of children with a schizophrenic disorder (see Table 1). The content of their delusions is varied and might include themes that are persecutory (e.g., someone is trying to harm the child), somatic (e.g., "My arms grow longer every time I go in the swimming pool"), grandiose (e.g., "I have the powers to decide who lives and who dies"), referential (e.g., the rotating emblem at a gas station is sending secret messages to the child), religious (e.g., "I'm not really a boy, I'm an angel that God sent to be on earth"), or bizarre (e.g., a child believes that he is a soccer ball) in nature. With increased age, hallucinations and delusions become more complex and elaborate, and the content reflects the developmental level of the patient. Themes of animals and monsters, for instance, are more prevalent among younger children, whereas sexual themes appear to be less common (Russell et al., 1989).

Another manifestation of schizophrenia is disorganized thinking or formal thought disorder. As shown in Table 1, thought disorder appears to be present in the majority of children diagnosed with schizophrenia and typically has been observed in 80–100% of cases. Signs of thought disorder include loose associations (e.g., "I am 10-years-old. My pen ran out of ink this morning"), illogical thinking ("I hate peanut butter but I love to eat it all the time"), incoherence ("The day to raindrops go"), and poverty of content of speech ("It does. Uh huh. Yes, does"). Obtaining reliable estimates of formal thought disorder in children can be very difficult, however, in view of developmental constraints on cognitive, linguistic, and practical social skills (e.g., Caplan, 1994; Russell, 1994; Werry et al., 1994), which may account for differences in reported rates between studies. Russell et al. (1989), for example, found that only 40% of their cases demonstrated formal thought disorder.

Inappropriate or flat affect are other common symptoms of schizophrenia and typically have been found in over 70% of children with a schizophrenic disorder (see Table 1). Schizophrenic patients may exhibit inappropriate or uncontrollable emotions, such as laughing at a sad story, responding irritably for no apparent reason, or fluctuating rapidly between emotions. Alternatively, children diagnosed with schizophrenia may show no facial expression and speak in a monotone voice.

Grossly disorganized behavior in schizophrenia can include unpredictable agitation (e.g., unprovoked shouting), extreme silliness, and difficulties with such goal-directed behaviors as maintaining personal hygiene. In children diagnosed with schizophrenia, grossly disorganized behavior has been found to varying degrees. McKenna et al. (1994) noted disorganized behavior in 84% of their sample, whereas Russell et al. (1989) detected it in only 40% of their cases. Green, Padron-Gayol, Hardesty, and Bassiri (1992) observed that either grossly disorganized or catatonic behavior was apparent in 32% of their childhood-onset patients. Catatonic motor behaviors involve psychomotor disturbances, ranging from motoric immobility to excited, excessive activity. In general, catatonic behaviors appear to be rare among children. McKenna et al. (1994) reported the highest rate at 37% and Russell (1994) did not find any cases of catatonic behaviors; other studies made no mention of catatonia, suggesting that it was rare or nonexistent in their samples.

Although suicidal ideation or suicide attempts are often present in adults with schizophrenia, rates of their occurrence in children have not been routinely reported in the literature. In a hospitalized sample of 21 childhood-onset schizophrenics, J. R. Asarnow, Tompson, and Goldstein (1994) observed evidence of suicidality in 76% of cases. The authors note that the rate of suicidality in their inpatient sample may be higher than in other samples as suicidal behavior can often precipitate hospitalization.

As Russell (1994) notes, some young children may not experience their psychotic symptoms as being particularly distressing or unusual. Instead, these symptoms may be perceived as indistinguishable from normal experience given their insidious and early onset. As the psychotic symptoms become more intrusive and interfere with attention, however, children may be more likely to become disturbed or frightened and will then describe their symptoms to parents or others. Parents and others who are close to the child initially may not recognize that a child is experiencing hallucinations and may, instead, attribute the child's behavior to daydreaming.

Documentation of the phenomenology of schizophrenic disorder in early adolescence has been relatively sparse in terms of the current number of studies available. Nonetheless, as shown in Table 2, the studies that do report on symptom patterns in adolescent-onset schizophrenia find considerable similarity in the relative proportion of symptom characteristics as compared with those observed in childhood-onset cases. The rates of symptoms noted by Remschmidt, Martin, Schulz, Gutenbrunner, and Fleischhaker (1991) and by Werry, McClellan, and Chard (1991) are somewhat lower than those obtained in children diagnosed with schizophrenia but this may be due in part to methodological differences in ascertaining symptom patterns. In the studies by McKenna et al. (1994), Russell et al. (1989) and Spencer and Campbell (1994), rates for psychotic symptoms were determined through interviews with the children and their families. Green et al. (1992) relied on clinical records, but noted that their charts received extensive documentation as it was understood that they would be reviewed for research purposes. In both studies of adolescents, symptoms were ascertained from reviews of hospital records although Werry et al. also attempted to subsequently interview patients and their families. It is possible, therefore, that reliance upon archival data could account for lower rates of psychotic symptomatology in adolescents than that observed in studies of childhood-onset schizophrenia.

While the constellation of symptoms appears to be similar across children, adolescents, and adults with schizophrenia, considerable heterogeneity exists in symptomatology across patients. To help delineate particular symptom patterns, the concept of positive and negative symptoms has been introduced in the clinical research literature on adults with schizophrenia. The more florid or bizarre symptoms, such as hallucinations and delusions, constitute positive symptoms, whereas deficient or absent behaviors, such as inappropriate affect, unchanging or flat facial expression and minimal speech, are considered negative symptoms. More recently, the concept of "deficit state" has been introduced to focus upon primary and enduring negative symptoms (Carpenter, Heinrichs, & Wagman, 1988). In adult schizophrenic patients, negative symptoms tend to be present early in the illness, are less responsive to treatment than positive symptoms, and are relatively stable over time (Arndt, Andreasen, Flaum, Miller, & Nopoulus, 1995). A predominant pattern of negative symptoms has also been related to poor outcome (e.g., Breier, Schreiber, Dyer, & Pickar, 1991), although Pogue-Geile and Harrow (1985) point out that the association may be due to the persistence of poor social and occupational functioning.

The distinction between positive and negative symptoms has been applied less frequently to children and adolescents with schizophrenia. Studies in the 1990s have determined that the positive and negative symptom patterns are present in children and adolescents, much as they are in adults (e.g., Maziade et al., 1996; Remschmidt et al., 1991; Remschmidt, Schulz, Martin, Warnke, & Trott, 1994). In examining the course of these two symptom clusters in childhood- and adolescent-onset schizophrenia, Remschmidt et al. (1994) found some suggestion that negative symptoms may occur over a longer period of time than positive symptoms in the period immediately prior to hospitalization, particularly in children.

5.27.2.1.1 A case illustration

Jeremy is an eight-year old whose family moved to the US from England as he was approaching his sixth birthday. His father is a corporate executive and his mother works part-time as a literary agent. He resides with his parents and a 14-year old half-sister from his mother's previous marriage. The pregnancy and delivery were unremarkable and developmental milestones were reached within normal limits. His mother described Jeremy as a bright, active, and inquisitive toddler who needed to be monitored more closely than his sister at this stage to prevent accidents. Family history of psychiatric illness is largely negative with the exception of a paternal aunt who has a fear of heights. Jeremy's father believes that he may have been hyperactive as a child but he was never diagnosed or treated for the disorder; he attributed his behavior to being bored and unchallenged during grade school.

Jeremy started the first grade at a private school shortly after arriving in the US and began to exhibit noncompliant and aggressive behaviors. His teachers and parents initially attributed his behavioral problems to the family's recent move and the transition to a new school. Rather than subside over time, his negative behaviors began to escalate. A behavior management program was implemented by the school and it appeared to be modestly successful. By the end of second grade, however, Jeremy had become increasingly disruptive in the classroom, had begun to destroy other children's property, and was unable to sit still for any length of time. His school performance had declined substantially and he was performing in the low average range. Peer relations were similarly poor. He was referred for a psychiatric

Table 2 Comparison of studies reporting on clinical symptoms in adolescents with schizophrenia.

			Hallucinations						
Study	*N*	*Age* (years)	*Auditory* (%)	*Visual* (%)	*Tactile* (%)	*Olfactory* (%)	*Delusions* (%)	*Thought disorder* (%)	*Inappropriate or flat affect* (%)
Remschmidt, Martin, Schulz, Gutenbrunner, and Fleischhaker (1991)	113	16–20[a]	23[b]				55	38	37
Werry, McClellan, and Chard (1991)	30[c]	7–17	53	13	13		47		57

[a] Approximate estimate of age, derived from the mean and standard deviation provided in the report. [b] Not reported by subtype. [c] Three of the cases were below the age of 12.

evaluation, at which time he was diagnosed with attention-deficit hyperactivity disorder and started on a trial of methylphenidate. The stimulant trial resulted in a further deterioration of behavior whereby Jeremy's behavior became more agitated and unpredictable. This course of medication was discontinued, Jeremy was rediagnosed with chronic adjustment and conduct disorders, and he was referred for psychotherapy.

A psychiatric hospitalization was precipitated when Jeremy attempted to repeatedly stab a classmate with a sharpened pencil and, two days later, tried to choke his sister claiming that she was not his "real sister." Jeremy also had begun to pick at his skin and scabs until he drew blood, discussing the possibility of bleeding to death. On admission, Jeremy was neatly groomed and presented as a friendly but fidgety child with a short attention span. He reported hearing his name whispered but denied any other hallucinatory experiences. When questioned about the violent acts towards his classmate and sister, Jeremy stated that they were not who they seemed to be and that he had to kill the space aliens that had taken over their bodies in order for his real classmate and his real sister to come back to life. Jeremy refused to discuss his own self-injurious behavior and denied any suicidal ideation.

Upon admission to the hospital, physical and neurological exams proved to be normal. The patient made a good adjustment to the hospital environment and initially responded well to the behavior management program that was implemented. After a few days, however, Jeremy began to exhibit increasingly aggressive, chaotic, and disruptive behaviors. In addition, he resumed the practice of picking at his skin. The staff also observed occasional staring spells. An EEG was obtained and although a mild irregularity was observed, the follow-up magnetic resonance image was normal and did not support the possibility of a structural lesion. Jeremy slowly began to disclose psychotic symptoms that included auditory and visual hallucinations. For instance, Jeremy described a wizard appearing and although he could not always see him, the wizard would whisper his name and instruct him to hurt himself because he had disobeyed the rules. By making himself bleed, Jeremy reported that he could make the wizard go away. Jeremy also described seeing daggers that would drop down from the sky. At night, Jeremy was unable to sleep for the entire duration in his own room; he would awaken and attempt to bring a blanket to sleep on the floor next to the nursing station. Jeremy acknowledged that he was afraid but refused to reveal his specific fear, stating that disclosure would endanger other people. Over time, Jeremy became increasingly disorganized, his social relatedness declined considerably, and he would speak incoherently. He responded extremely well to a trial of risperidone, an atypical antipsychotic medication, that was offered along with supportive psychotherapy and exhibited no psychotic symptoms upon discharge. The total duration of hospitalization was 45 days.

5.27.2.2 Diagnostic Features and Differential Diagnosis

The *DSM-IV* and *ICD-10* definitions of schizophrenia are similar and, as noted, both use the same criteria for children, adolescents, and adults. The primary differences between the two systems is the degree to which diagnostic criteria are specified and the required duration for certain symptomatology, such that the *ICD-10* definition might be more liberal (Werry, 1996). This chapter focuses on *DSM-IV* criteria as this has been the predominant system used in recent studies on childhood-onset and adolescent schizophrenia. Extensive discussions of the similarities and differences between these two diagnostic systems, as well as comparisons with past systems, can be found in J. R. Asarnow (1994) and Werry (1996).

The *DSM-IV* criteria for schizophrenia (see Table 3) requires the presence of psychotic symptoms, accompanied by a deterioration in social and occupational functioning. In children and adolescents, this might be indicated by a failure to achieve the expected level in social development or academic achievement. Psychotic symptoms must be apparent for at least one month while signs of the disorder must persist for at least six months. Duration of symptoms is a particularly critical factor when considering a diagnosis of schizophrenia in children who might otherwise exhibit transient delusions or hallucinations (Rapoport & Ismond, 1996). Finally, symptoms cannot be attributed to another psychiatric disorder, medication, substance use, or a general medical condition. Subtypes of schizophrenia (i.e., paranoid, disorganized, catatonic, undifferentiated, and residual types) can be defined by the pattern of presenting symptoms but this typology appears to be less clearly differentiated in children than it is in adults (Rapoport & Ismond, 1996).

Because certain types of delusions and auditory hallucinations are seen as hallmarks of schizophrenia, they are weighted more heavily when making a diagnosis of schizophrenia. The presence of bizarre delusions or

Table 3 *DSM-IV* diagnostic criteria for schizophrenia.

A. *Characteristic symptoms*. Two (or more) of the following, each present for a significant portion of time during a one-month period (or less if successfully treated):
  (i) delusions
  (ii) hallucinations
  (iii) disorganized speech (e.g., frequent derailment or incoherence)
  (iv) grossly disorganized or catatonic behavior
  (v) negative symptoms, that is, affective flattening, alogia, or avolition
  Note: Only one criterion A symptom is required if delusions are bizarre or hallucinations consist of a voice keeping up a running commentary on the person's behavior or thoughts, or two or more voices conversing with each other

B. *Social/occupational dysfunction*. For a significant portion of the time since the onset of the disturbance, one or more major areas of functioning such as work, interpersonal relations, or self-care are markedly below the level achieved prior to the onset (or when the onset is in childhood or adolescence, failure to achieve expected level of interpersonal, academic, or occupational achievement)

C. *Duration*. Continuous signs of the disturbance persist for at least six months. This six-month period must include at least one month of symptoms (or less if successfully treated) that meet criterion A (i.e., active-phase symptoms) and may include periods of prodromal or residual symptoms. During these prodromal or residual periods, the signs of the disturbance may be manifested by only negative symptoms or two or more symptoms listed in criterion A present in an attenuated form (e.g., odd beliefs, unusual perceptual experiences)

D. *Schizoaffective and Mood Disorder exclusion*. Schizoaffective Disorder and Mood Disorder with Psychotic Features have been ruled out because either (i) no Major Depressive, Manic, or Mixed Episodes have occurred concurrently with the active-phase symptoms; or (ii) if mood episodes have occurred during active-phase symptoms, their total duration has been brief relative to the duration of the active and residual periods

E. *Substance/general medical condition exclusion*. The disturbance is not due to the direct physiological effects of a substance (e.g., a drug of abuse, a medication) or a general medical condition

F. *Relationship to a Pervasive Developmental Disorder*. If there is a history of Autistic Disorder or another Pervasive Developmental Disorder, the additional diagnosis of Schizophrenia is made only if prominent delusions or hallucinations are also present for at least a month (or less if successfully treated)

Required with permission from the *Diagnostic and Statistical Manual of Mental Disorders*, Fourth edition. Copyright 1994 American Psychiatric Association.

auditory hallucinations involving commenting or conversing voices, therefore, is sufficient to satisfy criterion A, as shown in Table 3. As McKenna et al. (1994) indicate, thought processes may be developmentally immature in the younger patient but are not necessarily indicative of thought disorder or delusions. An adolescent, for instance, might engage in magical thinking which, while age-inappropriate, would not necessarily signify a psychotic disorder. Moreover, hallucinations are not uncommon in young children, and when they are accompanied by disorganized behavior they could lead to a misdiagnosis of schizophrenia. A key consideration, therefore, in making a diagnosis of schizophrenia in children and adolescents is the extent to which psychotic symptoms are pervasive and enduring rather than brief and transient. Disorganized speech is also likely to occur with increased frequency in children and may be indicative of a language deficit or communication disorder rather than schizophrenia (McKenna et al., 1994; Rapoport & Ismond, 1996). As noted by Volkmar and Schwab-Stone (1996), disorganized speech and grossly disorganized behavior in children may be nonpsychotic, and reliance upon these two characteristic symptoms will likely lead to a misdiagnosis of schizophrenia. Negative symptoms, such as affective flattening, can be difficult to distinguish from the psychomotor retardation that can accompany a mood disorder. Thus, the differential diagnosis of schizophrenia is not only critical but can prove to be a particular challenge in children and adolescents.

A misdiagnosis of childhood-onset schizophrenia, unfortunately, is not an uncommon event. McKenna et al. (1994) at the National Institute of Mental Health (NIMH) received more than 260 referrals from academic centers for children and adolescents with a presumptive diagnosis of childhood-onset schizophrenia. After screening the medical records, 71 children were evaluated at NIMH. Of this sample, fewer than 28% were found to meet diagnostic criteria for schizophrenia while the majority of children were best described by another diagnostic category. Misdiagnosis of schizophrenia is obviously a serious problem as the illness is associated with a relatively poor prognosis and neuroleptic medications that can have grave and potentially irreversible side effects. In addition to false-positives, false-negatives undoubtedly occur with some, as yet unknown, frequency.

One area that can prove to be confusing is the differential diagnosis of schizophrenia from other schizophrenia-spectrum disorders such as schizophreniform disorder, schizoaffective disorder, and schizotypal personality disorder. The primary distinction between schizophrenia and schizophreniform disorder is duration, with the latter requiring the presence of symptoms for less than six months. Schizophreniform disorder, however, is rarely diagnosed in childhood (Rapoport & Ismond, 1996). Schizoaffective disorder involves a concurrent manic or depressive episode during an uninterrupted and active phase of schizophrenic illness. As with adults, differentiating between schizoaffective disorder, mood disorder with psychotic features, and schizophrenia that is accompanied by brief disturbances in mood requires careful documentation of the time course and duration of symptoms. In general, the rate of psychotic symptoms accompanying depression is much higher than in adults, with estimates ranging from 48% in prepubertal depressed children (Chambers, Puig-Antich, Tabrizi, & Davies, 1982) to 31% in an inpatient sample of depressed adolescents (Strober, Lampert, Schmidt, & Morrell, 1993).

Schizotypal personality disorder also presents problems in the differential diagnosis of schizophrenia in children and adolescents. Because the onset of a schizophrenic disorder in children is generally insidious or subacute (J. R. Asarnow & Ben-Meir, 1988; McKenna et al., 1994), it may be difficult to distinguish schizotypal personality disorder from the premorbid phase of schizophrenia. In addition, premorbid schizoid traits have been found to occur with increased frequency among adolescents who subsequently develop schizophrenia (Dalkin, Murphy, Glazebrook, Medley, & Harrison, 1994).

Diagnostic confusion also can arise frequently between schizophrenia and mood disorders. McKenna et al. (1994) found that the majority of children in their sample were misdiagnosed with schizophrenia as a sizable group presented with major depression or bipolar disorder. Rapoport and Ismond (1996) caution that diagnoses such as mood disorder with psychotic features are often overlooked in children and that other depressive disorders also must be considered as well. Of course, many of the same issues must be considered in discriminating between a mood disorder and schizophrenia as when making a differential diagnosis with schizoaffective disorder. Discrimination between the irritability associated with depression and the inappropriate outbursts that can occur with schizophrenia, for instance, is one example of the type of difficulty that may be encountered when evaluating children and adolescents.

With certain anxiety disorders, there may be some similarities in symptom presentation. For instance, the recurrent ruminations and bizarre stereotyped behaviors that are associated with obsessive-compulsive disorder (OCD) could be mistaken for symptoms of schizophrenia. In adult patients, a key diagnostic feature of OCD is that obsessions and compulsions are seen as ego-dystonic or alien and not under the control of the patient. In children with a schizophrenic disorder, psychotic symptoms may seem relatively ego-syntonic and even "normal" given their longstanding presence in the child's life experience (Russell, 1994). Post-traumatic stress disorder also must be ruled out as flashbacks can be confused with the hallucinations that accompany schizophrenia.

The diagnosis of schizophrenia is made separately from the diagnosis of autistic and pervasive developmental disorders (see Table 3). This is specified explicitly in the *DSM-IV* criteria for schizophrenia, probably because of the previous practice of lumping autism and schizophrenia into a single diagnostic group labeled either "childhood schizophrenia" or "childhood psychosis." In fact, the diagnostic criteria for the two disorders are quite different. Autism involves qualitative impairments in both social interaction and communication as well as restricted and stereotyped patterns of behavior and interests. Moreover, delays or abnormal functioning have to occur before the child reaches three years of age, whereas schizophrenia is rarely manifested or detected before six to seven years. Finally, children diagnosed with schizophrenia also tend to have low to normal intellectual functioning, whereas 70–80% of children with autism are also mentally retarded. Diagnostic confusions sometime arise in older, nonretarded children with autism. In some of these children, their preoccupations with stereotyped areas of interest can seem somewhat delusional and some children with autism have communication deficits that include loose associations and illogical thinking. A small number of autistic children develop true hallucinations and delusions (Petty, Ornitz, Michelman, & Zimmerman, 1984) but this tends to be rather infrequent (Volkmar & Cohen, 1991). Asperger's syndrome is a rare form of autism in which the children do not exhibit any communication disorder but do show problems in social interactions and preoccupations with stereotyped interests; these children do not manifest the psychotic symptoms associated with children diagnosed with schizophrenia.

The alternative diagnoses considered thus far are not intended to represent an exhaustive list of candidates. Certainly, numerous other possibilities (e.g., delusional disorder, substance-induced psychotic disorder, psychotic disorder due to a medical condition, degenerative neurological disorders) must be considered along with a diagnosis of schizophrenia (see Rapoport & Ismond, 1996). Even granting that tremendous caution is taken before conferring a diagnosis of schizophrenia on a child, adolescent, or adult, misdiagnoses still can occur with considerable frequency.

Several processes may contribute to diagnostic misclassification of schizophrenia in children and adolescents. One possibility, mentioned earlier, is that it can be difficult to establish the presence of some symptoms, particularly in younger children; symptoms may be overlooked, misinterpreted, or remain uncorroborated. As McKenna et al. (1994) found, schizophrenia also may be applied overinclusively to severely disturbed children who exhibit psychotic symptoms. In their sample, almost one-third of the children did not meet diagnostic criteria for schizophrenia but were best-described as "multidimensionally impaired" (MDI). Such a group of children highlight the continued need to improve upon our current diagnostic classification schemes (Volkmar & Schwab-Stone, 1996). Another factor that may lead to a failure to diagnose schizophrenia is that the characteristic symptoms of the illness may only emerge over time. Children later found to have a schizophrenic disorder, for example, may initially present with significant behavioral problems and receive an attention-deficit or disruptive behavior disorder diagnosis (e.g., Russell, 1994). Still another subset of children may initially present with a disorder, such as psychotic mood disorder (Werry et al., 1994), and only later develop schizophrenia. Finally, as illustrated by Werry et al. (1991), a diagnosis of schizophrenia may precede the onset of a psychiatric disorder such as bipolar disorder.

The diagnostic process is further complicated by the high rate of co-occurrence of other psychiatric disorders with childhood-onset schizophrenia. Russell et al. (1989) determined that 69% of their sample of children with schizophrenia met diagnostic criteria for another psychiatric disorder. Associated diagnoses occurring with the greatest frequency were conduct disorder (29%), atypical depression (26%), dysthymia (11%), and enuresis or encopresis (14%). Among adolescent patients, schizophrenia may coexist with a substance abuse disorder. Werry et al. (1994), for instance, reported a 54% comorbidity rate in their US sample. Although a small number of autistic children have been found to later develop schizophrenia (Petty et al., 1984), there does not appear to be a comorbid association between autism and schizophrenia (Volkmar & Cohen, 1991).

5.27.2.3 Prevalence

The potentially high rate of misdiagnosis of schizophrenia in children and adults is due in part to its rarity. Estimates of the incidence of schizophrenia in childhood range from 0.19 to 1 in 10 000 (Beitchman, 1985; Burd & Kerbeshian, 1987; Gillberg, Wahlstrom, Forsman, Hellgren, & Gillberg, 1986; Häfner & Nowotny, 1995). In adolescence, schizophrenia continues to be rare although it shows a steady increase with each successive year (Gillberg et al., 1986; Häfner & Nowotny, 1995). With the majority of cases of schizophrenia developing during late adolescence and early adulthood, schizophrenia occurs in about 1% of the population across the entire lifespan.

Among childhood-onset schizophrenia cases, there is a consistent sex difference with males exceeding females by a ratio of about 2:1 (Green et al., 1992; Hollis, 1995; Kolvin et al., 1971; Russell et al., 1989; Volkmar et al. 1988). These differences, however, disappear by adolescence where the distribution appears to be equal for both sexes (Hollis, 1995; Remschmidt et al., 1991; Schmidt, Blanz, Dippe, Koppe, & Lay, 1995; Werry et al., 1991). Various suggestions have been offered to explain the apparent influence of age on the sex distribution of schizophrenia in young patients, including the possibility of etiological differences with males showing greater biological vulnerability (J. R. Asarnow, 1994).

5.27.2.4 Etiology

Research examining the relationship between schizophrenia in children and adults has indicated a strong likelihood that the two forms of the disorder are continuous. In addition to the investigations into the phenomenology of childhood-onset schizophrenia described earlier, studies on brain morphology (Frazier et al., 1996), event-related potentials (Strandburg, Marsh, Brown, R. F. Asarnow, & Guthrie, 1994), autonomic responsivity (Zahn et al., 1997), smooth pursuit eye tracking (Gordon, Frazier et al., 1994), attention and information processing (R. F. Asarnow, Brown, & Strandburg, 1995), and reaction time (Zahn et al., in press) in children have found similar patterns of abnormalities to those observed in adults with schizophrenia. Evidence also exists to suggest that childhood-onset schizophrenia may represent a more homogenous and severe form of the disorder as the genetic loading for schizophrenia, degree of central nervous system dysfunction, and level of environmental insult or stress may be higher in children than in adults with schizophrenia (J. R. Asarnow, 1994; McKenna, Gordon, & Rapoport, 1994).

Additional efforts to understand the etiology of childhood- and adolescent-onset schizophrenias, thus far, are represented largely by the application and extension of theories developed to explain schizophrenia in adults. This research has consistently failed to support the notion that there is a single cause of schizophrenia, and instead, accumulating evidence suggests that schizophrenia is a heterogenous disorder with multiple causes and presentations. Several of the prevailing theories of schizophrenia will be considered.

Considerable evidence indicates genetic contributions to schizophrenia. Results of family and twin studies suggest concordance rates of about 6% in a parent, 9% in a sibling, 13% when one parent is affected, 17% in dizygotic twins, 46% when both parents are affected, and 48% in monozygotic twins (Gottesman, 1991). These findings have been bolstered by the results of adoption studies that demonstrate the genetic influence on schizophrenia although familial environment also appears to play a role (e.g., Kety, Rosenthal, Wender, & Schulsinger, 1976; Tienari et al., 1987). Preliminary findings from the NIMH study on childhood-onset schizophrenia suggests that 13.6% of the children have first-degree relatives with non-affective psychotic disorders. It is likely, however, that this is an underestimate of a possible genetic influence due to a small sample size and selection bias (Gordon, Frazier et al., 1994). In one child diagnosed with a schizophrenic disorder, there was an apparently balanced translocation involving chromosomes 1 and 7 which, when combined with evidence from other research, suggests a potential gene locus for developmental brain disorders (Gordon, Krasnewich, White, Lenane, & Rapoport, 1994).

Although the evidence for a genetic contribution to schizophrenia is strong, genetic heritability cannot account for all cases of schizophrenia nor does it explain instances whereby the monozygotic twin of a schizophrenic patient fails to develop the illness. Several lines of evidence suggest that other personal and environmental factors may lead to schizophrenia or may influence the outcome of an underlying genetic predisposition. Weinberger (1987) has proposed that schizophrenia is a neurodevelopmental disorder in which a fixed brain lesion or some pathologic condition of the brain occurs early in life. This lesion or brain pathology is likely to involve the dopaminergic neural systems and the prefrontal cortex. Over the normal course of maturation of the neural system, the brain lesion becomes evident with the onset of schizophrenic symptoms. Weinberger suggests that schizophrenia may be rare in children but more common in older adolescents and young adults as maximum dopaminergic activity in the brain does not occur until early adulthood. As for the cause of the lesion, numerous possibilities exist. Mednick, Machon, Huttunen, and Bonett (1988), for instance, have found that maternal exposure to a severe influenza during the second trimester is associated with an elevated frequency of adult schizophrenia in the offspring.

Psychosocial factors also may have a significant impact on the course of schizophrenia. Communication deviance (CD), which refers to an unclear, confusing, and fragmented communication style (Singer & Wynne, 1963), has been found to occur more frequently in the parents of adult schizophrenic patients (e.g., Singer & Wynne, 1965), adolescents at risk of schizophrenia and related disorders (Goldstein, 1987), and children diagnosed with schizophrenia or schizotypal personality disorder (J. R. Asarnow, Goldstein, & Ben-Meir, 1988). The etiological relationship, if any, between CD and schizophrenia is unclear, although it has been suggested that CD may represent an environmental stressor or a biological predisposition to schizophrenia in parents that is reflected in subtle levels of disordered thought or communication (Singer & Wynne, 1965; Miklowitz et al., 1991).

Another psychosocial factor that may contribute to the course of schizophrenia is expressed emotion (EE) or the degree to which critical and emotionally overinvolved attitudes are conveyed to the schizophrenic patient. As reviewed by Goldstein, Strachan, and Wynne (1997), numerous studies have shown that high levels of EE predict relapse among schizophrenic adults. In contrast, the association between EE and schizophrenia in children appears to be less robust although parents of children with schizophrenia or a related condition have been found to be more likely to express harsh criticisms toward their child compared with parents of normal children (J. R. Asarnow et al., 1994; Hamilton, 1991). On the basis of these data, it is not clear if harsh criticisms arise in response to dealing with a child with a severe mental illness, if some family members of schizophrenic patients tend to express critical comments that contribute to symptoms, or perhaps most likely, if the emotional environment of the family reflects a combination of these factors.

5.27.2.5 Developmental Course

The developmental course of schizophrenia in children and adolescents entails consideration of a broad range of issues, beginning with premorbid adjustment and extending to the prodromal phase, the course of the schizophrenic episode, and subsequent outcome of the disorder that may include relapse. Premorbid adjustment, or the child's level of functioning prior to the onset of schizophrenia, has been found to range from normal to uneven development (Gordon, Frazier et al., 1994). In reviewing the premorbid history of children diagnosed with schizophrenia, several studies have found delays on one or more developmental milestones in approximately one-half of children examined, with the largest and most frequent lags typically occurring in the areas of language and gross motor development (Alaghband-Rad et al., 1995; Hollis, 1995; Watkins, R. F. Asarnow, & Tanguay, 1988). It is clear that a number of childhood-onset schizophrenic patients also experience difficulty with academic performance and school adaptation (J. R. Asarnow et al., 1994; Hollis, 1995). Alaghband-Rad et al. (1995) observed that over half of their sample repeated at least one grade; a substantial proportion had received placements in special education classes and estimates of premorbid IQ tended to be in the low to normal range.

In terms of social development, approximately one-half of patients diagnosed with childhood-onset schizophrenia have been noted to exhibit significant impulsivity and approximately half were viewed as relatively withdrawn and socially impaired (Alaghband-Rad et al., 1995; Hollis, 1995). It is not surprising that when assessed, peer relationships also were quite poor with childhood-onset schizophrenic patients showing significantly greater levels of impairment than children diagnosed with major depressive disorder (J. R. Asarnow et al., 1994).

For some children, the presence of significant developmental delays or disruptive behaviors has been associated with other psychiatric disorders. Approximately one-third of schizophrenic children have a premorbid history marked by at least some feature of pervasive developmental disorder while about one-third have a premorbid history of attentional and hyperactivity problems (Alaghband-Rad et al., 1995; Russell et al., 1994). Overall, Russell (1994) concluded that 86% of their sample had a history of some behavioral or psychiatric disturbance prior to the emergence of psychotic symptoms. There is some evidence to suggest that childhood-onset patients with schizoaffective disorder may tend to have a better premorbid course than children diagnosed with schizophrenia (Eggers, 1989).

The premorbid course for adolescent-onset schizophrenia also appears to be generally poor. In samples obtained in New Zealand and the United States, Werry et al. (1994) observed that 83% of participants exhibited odd, anxious, or disruptive personality styles and that the typical patient had a history of poor adaptive functioning. In comparing the premorbid course between childhood- and adolescent-onset cases, Hollis (1995) found that children with schizophrenia were more likely to have a history of impaired language or speech development. The onset of schizophrenia in early adolescence was more likely to be accompanied by a history of restless and fidgety behavior although there was no significant difference between the two groups in motor development.

The onset of schizophrenia in children is typically gradual and insidious, with an extensive prodromal phase of behavioral disturbance that eventually leads to the emergence of psychotic symptoms. The prodromal period might be characterized by attentional problems, hyperactivity, declines in school performance, poor peer relationships, deterioration in daily functioning (e.g., performing personal hygiene), and social withdrawal (e.g., Green et al., 1992; McKenna et al., 1994; Russell et al., 1989; Watkins et al., 1988). An acute onset of psychotic symptoms appears to be be more characteristic of schizophrenia in adolescents and some adults (e.g., Hollis, 1995; Werry, 1996).

While research into the course of schizophrenia as it occurs in children and early adolescents is quite limited, the available data suggests that many patients will experience recurrent episodes of schizophrenia. J. R. Asarnow et al. (1994) contacted childhood-onset patients one year after discharge from the hospital and found that 89% of their sample continued to meet diagnostic criteria for schizophrenia or schizoaffective disorder rather than some other psychiatric disorder. When the same patients were followed up after 2–7 years, 61% of the sample was diagnosed with schizophrenia while 33% was in remission. Among remitted patients, only one child was free of symptoms two years after their initial hospitalization. Similar rates have been obtained by Eggers (1989) who observed complete remission in 27% of schizophrenic patients included in his study. In comparing children diagnosed with schizophrenia to those with schizoaffective disorder, the rate of remission appears to be comparable although the overall prognosis appears to be better for schizoaffective patients. In both cases, there is evidence to suggest an unremarkable premorbid history is associated with a more favorable outcome (Eggers, 1989).

Among patients diagnosed with adolescent-onset schizophrenia, Werry et al. (1994) similarly found that over 80% continued to experience psychotic episodes two years after follow-up. Schmidt et al. (1995) contacted adolescent-onset patients 1–15 years after the index hospitalization and found that 27% had no further schizophrenic episodes. As with childhood-onset cases, normal premorbid adjustment has been found to be predictive of a good outcome (Werry et al., 1994).

5.27.3 CONCEPTUALIZATION AND CLINICAL FORMULATION

5.27.3.1 Multimodal Assessment

Formulating a diagnosis of schizophrenia in a child or adolescent is obviously a serious proposition given the social stigma, relatively poor prognosis, and potentially irreversible side effects that can accompany antipsychotic medications. As noted earlier, a number of factors may contribute to the misdiagnosis and overdiagnosis of schizophrenia in children and adolescents. Notwithstanding these considerations, care also must be taken to avoid underutilization of this diagnostic category; failure to make a diagnosis of schizophrenia could perpetuate suffering and preclude patients from receiving necessary treatment and appropriate services (McClellan & Werry, 1994). Given the multitude of issues that need to be confronted in considering a diagnosis of schizophrenia in a child or adolescent, a thorough and multimodal assessment is clearly indicated.

Perhaps the first step involved in the diagnostic assessment of schizophrenia in children and adolescents is to obtain a thorough and detailed premorbid history as well as a comprehensive evaluation of the prodromal period and current level of functioning. Development, intellectual functioning, academic achievement, cultural considerations, and family psychiatric history and background are all crucial factors that need to enter into any consideration of possible psychotic symptoms. A challenge inherent to any assessment of schizophrenia in children and adolescents is distinguishing true psychotic manifestations from idiosyncratic thinking and perceptions associated with developmental delays, speech and language disturbances, vivid imaginations, exposure to an extreme traumatic stressor, as well as religious ideation (e.g., McClellan & Werry, 1994; McKenna et al., 1994). A thorough pediatric and neurological evaluation is also indicated to rule out the possibility of an organic psychosis. Results of an investigation by Adams, Kutcher, Antoniw, and Bird (1996) in adolescent-onset schizophrenic patients suggest that current, routine endocrine, and neuroimaging screening tests are of no diagnostic utility unless symptoms suggestive of neurological or endocrinological disorders are present.

Reliance on multiple informants that include the child or adolescent, parents, teachers, caretakers, and other significant individuals is likely to yield the most information and improve diagnostic accuracy. As described above, psychotic symptoms may be initially associated with other psychiatric disorders or they may not be obvious, particularly in younger children. A child, for instance, may initially present with significant behavioral disturbance and receive a diagnosis of conduct disorder or attention-deficit hyperactivity disorder. Months may elapse before the clinical presentation of psychotic symptoms becomes apparent and even then, the disturbances may not present across situations or be evident to all potential informants. Alternatively, some children and adolescents with schizophrenia will require hospitalization because of suicidal, aggressive, or dangerous behavior problems. As illustrated in the case study described earlier, a psychiatric hospitalization may be beneficial when considering a diagnosis of schizophrenia in younger patients as it can provide an excellent opportunity to observe and document a range of behaviors in a controlled setting.

Regardless of whether the evaluation is conducted on an inpatient or outpatient basis, it is critical that the range of symptoms be thoroughly assessed. Reliable and well-validated structured diagnostic interview schedules, such as the Schedule for Affective Disorders and Schizophrenia for School-Age Children-Epidemiologic Version (K-SADS-E; Orvaschel, 1995) and the Diagnostic Interview for Children and Adolescents (DICA; Herjanic & Reich, 1982), are currently available to guide clinicians and clinical researchers through a detailed examination of various diagnoses in children and adolescents. Werry (1996) cautions that structured interviews may be tedious but are clearly warranted given the seriousness of the diagnoses under consideration. Systematic coverage of symptoms is particularly useful when attempting to resolve some of the complexities in distinguishing between disorders. (See McKenna et al. (1994) and McClellan and Werry (1994) for excellent discussions of other key diagnostic issues.)

In addition to diagnostic interviews, a number of rating scales exist for assessing specific schizophrenic symptoms such as thought disorder, negative symptoms, and positive symptoms. These measures include the Kiddie Formal Thought Disorder Rating Scale (K-FTDS) and Story Game (Caplan, Guthrie, Fish, Tanguay, & David-Lando, 1989), Thought Disorder Index (TDI; Arboleda & Holzman, 1985), Scale for the Assessment of Negative Symptoms in Schizophrenia (SANS; Andreasen, 1983), and Scale for the Assessment of Positive Symptoms in Schizophrenia (SAPS; Andreasen, 1984). Although some of these measures have traditionally been used in clinical research settings, they may be appropriate for clinical practice as well. It is important to note that while psychotic symptoms are the hallmark of schizophrenia, most children who report hallucinations are not schizophrenic as they do not present with all of the requisite diagnostic symptoms (McClellan & Werry, 1994). The use of reliable and valid measures of thought disorder that have been found to discriminate between children with schizophrenia, other psychiatric disorders, or no history of psychiatric illness, therefore, can be helpful in ascertaining perceptual, language, and communication disturbances. Moreover, the importance of developmental effects on the K-FTDS and TDI have been demonstrated (Arboleda & Holzman, 1985; Caplan et al., 1989; Caplan, Perdue, Tanguay, & Fish, 1990), so both measures can potentially serve as useful guides in differentiating between age-appropriate and abnormal behaviors.

Finally, children or adolescents who have received an initial diagnosis of schizophrenia will need to be closely monitored and reassessed periodically to determine the course of illness, the impact of treatment, and level of functioning. As demonstrated by McKenna et al. (1994) and Werry et al. (1991), a significant proportion of children and adolescents who are initially diagnosed with schizophrenia are later found to exhibit a different psychotic disturbance, such as bipolar disorder or schizoaffective disorder. Diagnostic reassessment, in some instances, may be complicated by the effects of pharmacological treatments which can lead to a partial, or even, complete remission in symptoms. Over time and with treatment, symptom patterns also are likely to change and influence the patient's level of functioning accordingly. In some instances, it may be difficult to distinguish between maturational effects and change due to clinical improvement. There also may be marked discrepancies in degree of improvement across domains, such that premorbid levels of functioning may be reached in some areas but not in others. J. R. Asarnow et al. (1994) have demonstrated that there is considerable variability in the course of schizophrenic illness with some children exhibiting deterioration while others show minimal or even substantial improvement over time. Psychoeducational assessment of multiple domains of functioning at multiple points in time, therefore, provides opportunities to determine current strengths and weaknesses for a child and any modifications that might be needed in service provision.

5.27.3.2 Treatment Formulation

A diagnosis of schizophrenia is naturally a devastating event for a child or adolescent and their family members. As with any psychiatric illness, the initial diagnosis may be met with denial or even relief as an explanation has been found for unusual and perhaps frightening behavior. Treatment, therefore, ideally should be oriented not only towards the child but towards addressing the needs of their families as well.

Neuroleptic medication is generally regarded as the cornerstone of treatment for schizophrenia in children, adolescents, and adults. While pharmacological treatment may be considered necessary, it is not sufficient. Various clinical researchers (e.g., Kane, 1987) have recommended that schizophrenia in adults be approached in phases, each with different strategies and goals. The first phase involves the acute presentation of florid psychotic symptoms, and therefore, the necessary emphasis is on the use of antipsychotic drugs to decrease the severity of symptoms. Oftentimes, the child or adolescent will have been admitted to the hospital at this point. Once the acute phase is controlled, there is a stabilization period in which some psychotic symptoms will have improved while others may continue to persist. The third phase is reached once most psychotic symptoms have resolved. During these latter two periods, psychosocial and school-based interventions become critical factors in improving the quality of life, decreasing the likelihood of relapse, and meeting educational needs. When appropriate, rehabilitation programs may help to optimize the social and occupational outcome for many children and adolescents with schizophrenia. Depending upon symptom course and the effectiveness of pharmacologic interventions, residential or day treatment programs may need to be considered.

As with any child who has a debilitating and potentially chronic condition, case management services may be essential to obtaining necessary social services. Access to services can have far-reaching effects. Compliance with neuroleptics, for example, may depend upon the availability of funds for purchasing medication. In general, the burden experienced by family members, who are required to care for a child or adolescent with schizophrenia, can be alleviated to some extent if caregivers can receive guidance on how to access relevant resources. Family psychoeducational programs, which are discussed below, and family advocacy groups may be of substantial benefit to family members of children and adolescents diagnosed with schizophrenia as well as to the patients themselves.

The American Academy of Child and Adolescent Psychiatry (AACAP) has developed a detailed set of practice parameters for the diagnosis and treatment of children and adolescents with schizophrenia (McClellan & Werry, 1994). Although these practice parameters were developed to assist physicians, other clinicians may find the guidelines useful. A number of diagnostic and treatment recommendations are offered that reflect recent research knowledge and advancements in childhood and adolescent-onset schizophrenia.

5.27.4 TREATMENT

5.27.4.1 Psychosocial Treatments

The schizophrenic child or adolescent is likely to face interpersonal, behavioral, and intellectual challenges. Intervention programs, therefore, need to be designed to not only treat the schizophrenic illness but also meet the unique needs of each child.

In light of the earlier discussion on premorbid course, it would be reasonable to expect that many children and adolescents diagnosed with schizophrenia will have exhibited delays in certain areas of functioning, such as language, gross motor, or social development. With the onset of the illness, normal developmental progress may be impeded further. Individualized school programs may need to be developed, therefore, to accommodate the educational needs of a child diagnosed with schizophrenia. Reintegration into the classroom or school may require academic remediation as well as assistance with modifying inappropriate social behaviors and developing suitable social skills. In the classroom, for instance, behavioral interventions may be necessary for minimizing disruptive behaviors and increasing appropriate behaviors.

Social skills training may be helpful in eliminating or reducing the boundary that may exist between a child or adolescent with schizophrenia and his or her peers. With adult patients diagnosed with schizophrenia, skills-focused training leads to more effective interpersonal functioning and improved social adjustment, particularly among patients with a relatively early onset of illness (Marder et al., 1996). The authors speculate that adults with an earlier onset of schizophrenia may have been more socially deficient as they had less time to develop interpersonal skills prior to the onset of their illness, and therefore, may have benefited more from the training program. Such a scenario would argue in favor of implementing social skills training with children and adolescents. To the best of our knowledge, a comprehensive social skills training program has yet to be adapted for children and adolescents with a severe psychiatric disturbance, although an effective program has been developed for adult patients with schizophrenia (Liberman, DeRisi, & Mueser, 1989).

Individual psychotherapy is a traditional mode of treatment for children and adolescents with any psychiatric disturbance. Such approaches typically range from supportive therapies and behavioral or cognitive-behavioral strategies to insight-oriented therapies. AACAP (McClellan & Werry, 1994), however, cautions against the use of insight-oriented therapies, particularly during the acute phases of illness. More recently, psychoeducational family intervention programs have generated considerable interest.

As discussed earlier, there may be an association between the emotional environment of the family and a propensity towards relapse. Current psychoeducational family treatment programs, in contrast to previous techniques, focus on the present and encourage the patient to gradually re-enter the family environment. Family treatment recognizes that intrafamilial processes arise from the complex transactions between family members and patients, who can pose genuine difficulties for their relatives and the resources available to those relatives. Rather than attempt to establish a causal link between families and the patient's illness, treatment is oriented towards assisting patients and their families in the effective management of the illness (Goldstein & Miklowitz, 1995). Although there are various forms of psychoeducationally-oriented family treatment programs, they share certain features that likely contribute to their effectiveness. These include engaging family members early in treatment;

encouraging an open atmosphere rather than attempting to assign fault; educating families about the phenomenology, course, prognosis, etiologic models, and rationale behind different treatments; offering recommendations for coping with the disorder; providing training on improving communication and problem-solving skills; and performing crisis intervention (Goldstein & Miklowitz, 1995). Family intervention can be based in the home or clinic, may involve individual or multiple families, and may include or exclude the patient. As with social skills training, systematic application of psychoeducational family intervention programs appears thus far to have been limited to adult populations.

For some children or adolescents with schizophrenia, a residential treatment center or other out-of-home placement may need to be considered. In a follow-up study of child psychiatric inpatients, J. R. Asarnow, Goldstein, Carlson et al. (1988) found that 57% of children with a schizophrenia spectrum disorder were placed out of the home within 13 months of being discharged from the hospital; the majority were placed in residential treatment centers. The decision to place a child or adolescent out of the home is obviously a complicated issue, around which multiple determining factors may operate. The extent to which long-term therapeutic placements are utilized is likely to depend, in part, on how a patient responds to pharmacological treatment and medication compliance, a topic to which we now turn.

5.27.4.2 Pharmacological Treatments

Children and adolescents diagnosed with schizophrenia have largely received the same medication treatments as adult patients. Pharmacological agents have included conventional neuroleptics such as haloperidol (Haldol), loxapine (Loxitane), thiothixene (Navane), thioridazine (Mellaril), and pimozide (Orap) (Spencer & Campbell, 1994). Treatment with these antipsychotic medications has often been limited by accompanying side effects, such as sedation, weight gain, and extrapyramidal symptoms (EPS) which include dystonia, akinesia, akathisia, and tremor. With conventional neuroleptic treatment of schizophrenia in children and adolescents, there also is the possibility of withdrawal dyskinesia and a potentially irreversible movement disorder, tardive dyskinesia (McClellan & Werry, 1994). More recently, atypical neuroleptic medications such as clozapine (Clozaril) and risperidone (Risperdal) have been introduced. Both of these newer atypical antipsychotic agents appear to be associated with less EPS (Findling, Grcevich, Lopez, & Schulz, 1996; Kumra et al., 1996). Clozapine, however, has been reserved for treatment-refractory patients as patients treated with this medication must be closely monitored for agranulocytosis and seizures (Kumra et al., 1996).

5.27.4.3 Efficacy of Treatments

As mentioned earlier, it does not appear that a formal social skills training program exists for children and adolescents. In Germany, Remschmidt et al. (1994) have developed a rehabilitation program that was deemed necessary for about 40% of the schizophrenic children and adolescents at their hospital. Results of their study suggest that it was possible to gradually reintegrate most patients back into their families, schools, or professional training programs.

Despite their longstanding tradition, the effectiveness of individual psychotherapies with children and adolescents with schizophrenia has yet to be determined. In schizophrenic adults, results of several controlled trials failed to indicate that individual psychotherapy, either alone or combined with pharmacotherapy, contributed significantly to remission after a psychotic episode or prevention of relapse once the patient had returned to the community (Goldstein, 1989). Psychoeducational family education programs, in contrast, when combined with medication management, have consistently been shown to reduce the likelihood of relapse beyond that associated with medication alone in adult patients. Moreover, there is evidence to suggest that when combined with pharmacologic treatment, family interventions are more effective over a two-year period than individual therapies in delaying relapses and improving social functioning (Goldstein & Miklowitz, 1995). Although there is little reason to doubt that similar effects would not be observed in children and adolescents with schizophrenia, it remains an empirical question as to whether younger patients would obtain similar benefits from family education.

There have been few studies on the effectiveness of neuroleptic medication with children and adolescents; the efficacy of such treatments has been based primarily on extrapolations from double-blind comparison studies conducted with adult schizophrenic patients and case studies involving children and adolescents. In one of the earliest studies utilizing a double-blind and placebo-controlled design, Pool, Bloom, Wielke, Roninger, and Gallant (1976) demonstrated that loxapine and haloperidol were superior to placebo in treating a large sample of adolescents diagnosed with schizophrenia. Spencer and Campbell (1994) similarly found haloperidol to have superior effects over placebo in children with schizophrenia. In a double-blind study comparing an atypical agent with a typical neuroleptic, clozapine was shown to be a more effective form of treatment than haloperidol on all administered measures of psychosis (Kumra et al., 1996). These data are particularly striking as only children who were found to be nonresponsive to at least two typical neuroleptics were included in the study. Results of a case study (Quintana & Keshavan, 1995) and a retrospective chart review study (Findling et al., 1996) suggest that another atypical antipsychotic medication, risperidone, also may be efficacious in the treatment of children and adolescents with schizophrenia.

5.27.5 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

With the shift in the late 1970s and 1980s towards greater diagnostic specificity of schizophrenia in childhood and adolescence, a nucleus of clinical research findings began to emerge, supporting the possibility of early-onset forms of schizophrenia. As reviewed in this chapter, a number of similarities were found to exist between child, adolescent, and adult forms of schizophrenia when the same diagnostic criteria were applied. Seeking to determine the degree of continuity with adult schizophrenia was a critical and natural first step. As the fields of childhood- and adolescent-onset schizophrenia mature, however, it will be important to not only continue to extrapolate from the adult research literature and seek similarities and differences between the various forms of schizophrenia but to move towards elucidating a developmental perspective on schizophrenia in children and adolescents. There has been notable movement in this direction, such as the work by Caplan (e.g., Caplan, 1994) on the influence of maturation on communication skills and the assessment of thought disorder.

Because of the many developmental issues that can arise, further progress is needed to assist with diagnostic ambiguities. Overdiagnosis of schizophrenia in children and adolescents is occurring at an alarming rate (McKenna, Gordon, Lenane et al., 1994). Misdiagnoses, such as with attention-deficit hyperactivity disorder, also are not uncommon. As more standardized measures of various aspects of schizophrenia become implemented, diagnostic decisions can become more reliable and valid. Identification of more homogenous samples will also greatly facilitate comparisons of studies conducted by different research groups. In addition, attention needs to be directed to understanding children with psychopathological conditions that appear to fall on the borderline of schizophrenia, such as the MDI group that was identified by McKenna et al. (1994).

The relative youth of the area is underscored by the meager number of longitudinal and long-term follow-up studies. Information from such investigations is crucial to our understanding of schizophrenia in children and adolescents and to its treatment. For instance, it will be important to determine the extent to which course of illness is influenced by stressors in the child's environment as well as protective factors. Information on the long-term course of schizophrenia as child and adolescent patients enter into adulthood will be of considerable utility as patients and their families prepare for the future.

For clinical researchers, childhood-onset schizophrenia provides a unique opportunity to trace the etiologic underpinnings of a rare, but presumably more homogenous form of schizophrenia. The available evidence suggests that childhood- and adolescent-onset schizophrenia are distinct forms of the illness. Accordingly, continued comparisons between these groups will help shed light on the underlying neuropathology as well as the influence of age of onset on neurocognitive and behavioral disturbances in schizophrenia.

Perhaps the most striking deficiency in the literature at present is the relative paucity of research on the efficacy of psychosocial treatments of schizophrenia when diagnosed in children and adolescents. Drawing from the literature on adult schizophrenic patients, it would appear that the development of social skills training programs for children and adolescents, and psychoeducational family intervention programs for their families, are clearly warranted. With such programs in place, multimodal treatment studies could then be conducted to determine the most effective form of treatment for these young patients. As Campbell and colleagues (Campbell & Cueva, 1995; Spencer & Campbell, 1994) have noted, additional research is also needed to assess the use of atypical neuroleptics with younger children and the long-term effects of all antipsychotic medications in children and adolescents with schizophrenia.

Considerable progress has been made in understanding childhood- and adolescent-onset schizophrenias. This is particularly noteworthy in light of the fact that this is a rare and often difficult to diagnose disorder. Further research will assist not only with improving our understanding of schizophrenia but with alleviating the distress and disruptions experienced by children and adolescents diagnosed with schizophrenia and their families.

5.27.6 SUMMARY

Schizophrenia is uncommon in childhood and adolescence, with prevalence estimates ranging from 0.19 to 1 in 10 000. By a ratio of about 2:1, male children are more likely than female children to be diagnosed with schizophrenia. Male and female adolescents, in contrast, are about equally likely to receive this diagnosis. Diagnosis of childhood- and adolescent-onset forms of schizophrenia involves the same criteria as adult forms of the disorder, although developmental variations can complicate diagnostic decision-making. Distinguishing psychotic thought disturbances from idiosyncratic thinking associated with developmental delays, language disturbances, vivid imagination, and stress responses may be difficult but, nonetheless, is essential.

Auditory hallucinations have been noted in approximately 80% of children diagnosed with schizophrenia, along with high rates of visual hallucinations, delusions, formal thought disorder, and inappropriate or flat affect. The presence of pervasive and enduring psychotic symptoms is necessary for making a diagnosis of schizophrenia, and careful documentation of the time course and duration of symptoms is essential for accurate classification. A high rate of misdiagnosis may be quite common (McKenna et al., 1994), with a complex presentation of symptoms and the presence of comorbid conditions often complicating the diagnostic picture. In particular, schizophrenia in children and adolescents is often confused with other schizophrenia-spectrum disorders, major depression, bipolar disorder, and mood disorder with psychotic features. Given the complexity of this disorder and its diagnosis, the use of multiple informants and structured interviews is strongly recommended. Moreover, because change in symptom patterns is common, frequent reassessment is necessary.

With regard to etiology, there appears to be marked continuity between childhood-onset schizophrenia and adult forms of the disorder. Research suggests that childhood-onset schizophrenia is a more homogeneous and severe form of the disorder, in terms of genetic loading, central nervous system dysfunction, and level of environmental insult or stress. As with adult forms of the disorder, theoretical models relating to children and adolescents emphasize the interplay of genetic and environmental factors. However, the adult models may need to be altered to account for biological changes that occur with time. For instance, key aspects of dopaminergic functioning may be low in childhood and increase into early adulthood or, to account for the increased salience of environmental variables, the quality of family communication may prove to be more etiologically significant at different points along the lifespan.

The onset of schizophrenia in children is typically gradual, with an extensive prodromal phase that is marked by attentional problems, hyperactivity, and deterioration in school and social functioning. Acute onset appears to be more common among adolescents and adults. Most children who are eventually diagnosed with schizophrenia have some history of developmental delays, behavioral difficulties, and social impairment; the premorbid course for individuals diagnosed with adolescent-onset schizophrenia is also poor. Recurrence of symptoms is common among childhood- and adolescent-onset cases, although an unremarkable premorbid history may be related to a more favorable outcome.

Treatment of schizophrenia in children and adolescents should be multimodal with pharmacological interventions forming the cornerstone. Conventional neuroleptics as well as atypical antipsychotic agents have been found to be successful in treating children and adolescents with schizophrenia. In adult patients with schizophrenia, their effectiveness in delaying relapses and improving social functioning is significantly improved when combined with a psychoeducational family intervention program. Comparable effects with childhood- and adolescent-onset cases are likely but have not yet been documented. School-based and social skills training interventions may also be beneficial.

In sum, considerable progress has been made in understanding and treating children and adolescents who are diagnosed with schizophrenia. Continued research, particularly involving prospective longitudinal designs, is needed to obtain a better understanding of the presentation, diagnosis, causes, course, and treatment of schizophrenia in children and adolescents.

ACKNOWLEDGMENT

This chapter is dedicated to the memory of Dr. Michael J. Goldstein, colleague, mentor, and friend, who devoted his career to improving our understanding and treatment of schizophrenia through his research and encouragement of young investigators.

5.27.7 REFERENCES

Adams, M., Kutcher, S., Antoniw, E., & Bird, D. (1996). Diagnostic utility of endocrine and neuroimaging screening tests in first-onset adolescent psychosis. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 67–73.

Alaghband-Rad, J., McKenna, K., Gordon, C. T., Albus, K. E., Hamburger, S. D., Rumsey, J. M., Frazier, J. A., Lenane, M. C., & Rapoport, J. L. (1995). Childhood-onset schizophrenia: The severity of the premorbid course. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*, 1273–1283.

American Psychiatric Association. (1968). *Diagnostic and statistical manual of mental disorders* (2nd ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Andreasen, N. C. (1983). *The Scale for the Assessment of Negative Symptoms (SANS)*. Iowa City, IA: University of Iowa.

Andreasen, N. C. (1984). *The Scale for the Assessment of Positive Symptoms (SAPS)*. Iowa City, IA: University of Iowa.

Arboleda, C., & Holzman, P. S. (1985). Thought disorder in children at risk for psychosis. *Archives of General Psychiatry, 42*, 1004–1013.

Arndt, S., Andreasen, N.C., Flaum, M., Miller, D., Nopoulus, P. (1995). A longitudinal study of symptom dimensions in schizophrenia: Prediction and patterns of change. *Archives of General Psychiatry, 52*, 352–360.

Asarnow, J. R. (1994). Annotation: Childhood-onset schizophrenia. *Journal of Child Psychology and Psychiatry, 35*, 1345–1371.

Asarnow, J. R., & Ben-Meir, S. (1988). Children with schizophrenia spectrum and depressive disorders: A comparative study of onset patterns, premorbid adjustment, and severity of dysfunction. *Journal of Child Psychology and Psychiatry, 29*, 477–488.

Asarnow, J. R., Goldstein, M. J., & Ben-Meir, S. (1988). Parental communication deviance in childhood onset schizophrenia spectrum and depressive disorders. *Journal of Child Psychology and Psychiatry, 29*, 825–838.

Asarnow, J. R., Goldstein, M. J., Carlson, G. A., Perdue, S., Bates, S., & Keller, J. (1988). Childhood-onset depressive disorders: A follow-up study of rates of rehospitalization and out-of-home placement among child psychiatric inpatients. *Journal of Affective Disorders, 15*, 245–253.

Asarnow, J. R., Tompson, M. C., & Goldstein, M. J. (1994). Childhood-onset schizophrenia: A follow-up study. *Schizophrenia Bulletin, 20*, 599–617.

Asarnow, R. F., & Asarnow, J. R. (1994). Childhood-onset schizophrenia: Editors' Introduction. *Schizophrenia Bulletin, 20*, 591–597.

Asarnow, R. F., Brown, W., & Strandburg, R. (1995). Children with a schizophrenic disorder: Neurobehavioral studies. *European Archives of Psychiatry and Clinical Neuroscience, 245*, 70–79.

Beitchman, J. H. (1985). Childhood schizophrenia: A review and comparison with adult-onset schizophrenia. *Psychiatric Clinics of North America, 8*, 793–814.

Bleuler, E. (1950). *Dementia Praecox or the group of schizophrenias.* (J. Zinkin, Trans.). New York: International Universities Press (Original work published 1911).

Breier, A., Schreiber, J. L., Dyer, J., & Pickar, D. (1991). National Institute of Mental Health Longitudinal Study of Chronic Schizophrenia: Prognosis and predictors of outcome. *Archives of General Psychiatry, 48*, 239–246.

Burd, L., & Kerbeshian, J. (1987). A North Dakota prevalence study of schizophrenia presenting in childhood. *Journal of the American Academy of Child and Adolescent Psychiatry, 26*, 347–350.

Campbell, M., & Cueva, J. E. (1995). Psychopharmacology in child and adolescent psychiatry: A review of the past seven years, part II. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*, 1262–1272.

Caplan, R. (1994). Thought disorder in childhood. *Journal of the American Academy of Child and Adolescent Psychiatry, 33*, 605–615.

Caplan, R., Guthrie, D., Fish, B., Tanguay, P. E., & David-Lando, G. (1989). The Kiddie Formal Thought Disorder Rating Scale (K-FTDS): Clinical assessment, reliability, and validity. *Journal of the American Academy of Child and Adolescent Psychiatry, 28*, 208–216.

Caplan, R., Perdue, S., Tanguay, P. E., & Fish, B. (1990). Formal thought disorder in childhood-onset schizophrenia and schizotypal personality disorder. *Journal of Child Psychology and Psychiatry, 31*, 169–177.

Carpenter, W. T., Jr., Heinrichs, D. W., & Wagman A. M. I. (1988). Deficit and nondeficit forms of schizophrenia: The concept. *American Journal of Psychiatry, 145*, 578–583.

Chambers, W., Puig-Antich, J., Tabrizi, M., & Davies, M. (1982). Psychotic symptoms in prepubertal major depressive disorder. *Archives of General Psychiatry, 39*, 921–927.

Dalkin, T., Murphy, P., Glazebrook, C., Medley, I., & Harrison, G. (1994). Premorbid personality in first-onset psychosis. *British Journal of Psychiatry, 164*, 202–207.

De Sanctis, S. (1973). On some varieties of dementia praecox (A. J. Gianascol & A. M. Corbiascio, Trans.). In S. A. Szurek & I. N. Perlin (Eds.), *Clinical studies in childhood psychoses: 25 years in collaborative treatment and research* (pp. 31–47). New York: Brunner/Mazel (Original work published 1906).

Eggers, C. (1989). Schizo-affective psychoses in childhood: A follow-up study. *Journal of Autism and Developmental Disorders, 19*, 327–342.

Findling, R. L., Grcevich, S. J., Lopez, I., & Schulz, S. C. (1996). Antipsychotic medications in children and adolescents. *Journal of Clinical Psychiatry, 57* (Suppl. 9), 19–23.

Fish, B., & Ritvo, E. R. (1979). Psychoses of childhood. In J. D. Noshpitz (Ed.), *Basic handbook of child psychiatry: Vol. 2. Disturbances in development* (pp. 249–304). New York: Basic Books.

Frazier, J., Giedd, J. N., Hamburger, S. D., Albus, K. E., Kaysen, D., Vaituzis, A. C., Rajapakse, J. C., Lenane, M. C., McKenna, K., Jacobsen, L. K., Gordon, C. T., Breier, A., & Rapoport, J. L. (1996). Brain anatomic magnetic resonance imaging in childhood-onset schizophrenia. *Archives of General Psychiatry, 53*, 617–624.

Gillberg, C., Wahlstrom, J., Forsman, A., Hellgren, L., & Gillberg, I. C. (1986). Teenage psychoses: Epidemiology, classification and reduced optimality in the pre-, peri- and neonatal periods. *Journal of Child Psychology and Psychiatry, 27*, 87–98.

Goldstein, M. J. (1987). Psychosocial issues. *Schizophrenia Bulletin, 13*, 157–171.

Goldstein, M. J. (1989). Psychosocial treatment of schizophrenia. In S. C. Schulz & C. A. Tamminga (Eds.), *Schizophrenia: Scientific progress* (pp. 318–324). New York: Oxford University Press.

Goldstein, M. J., & Miklowitz, D. J. (1995). The effectiveness of psychoeducational family therapy in the treatment of schizophrenic disorders. *Journal of Marital and Family Therapy, 21*, 361–376.

Goldstein, M. J., Strachan, A. M., & Wynne, L. C. (1997). Relational problems related to a mental disorder or general medical condition. In T. A. Widiger, H. A. Pincus, R. Ross, M. B. First, & W. Davis (Eds.), *DSM-*

IV Sourcebook: Volume 3 (pp. 531–567). Washington, DC: American Psychiatric Press.

Gordon, C. T., Frazier, J. A., McKenna, K., Giedd, J., Zametkin, A., Zahn, T., Hommer, D., Hong, W., Kaysen, D., Albus, K. E., & Rapoport, J. L. (1994). Childhood-onset schizophrenia: An NIMH study in progress. *Schizophrenia Bulletin, 20*, 697–712.

Gordon, C. T., Krasnewich, D., White, B., Lenane, M., & Rapoport, J. L. (1994). Brief report: Translocation involving chromosomes 1 and 7 in a boy with childhood-onset schizophrenia. *Journal of Autism and Developmental Disorders, 24*, 537–545.

Gottesman, I. I. (1991). *Schizophrenia genesis: The origins of madness*. New York: Freeman.

Green, W. H., Padron-Gayol, M., Hardesty, A. S., & Bassiri, M. (1992). Schizophrenia with childhood-onset: A phenomenological study of 38 cases. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*, 968–976.

Häfner, H., & Nowotny, B. (1995). Epidemiology of early-onset schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience, 245*, 80–92.

Hamilton, E. B. (1991). *Interactional styles in families of children with depressive disorders, schizophrenia spectrum disorders, and normal controls*. Unpublished doctoral dissertation, University of California, Los Angeles, CA.

Herjanic, B., & Reich, W. (1982). Development of a structured psychiatric interview for children: Agreement between child and parent on individual symptoms. *Journal of Abnormal Child Psychology, 10*, 307–324.

Hollis, C. (1995). Child and adolescent (juvenile onset) schizophrenia: A case control study of premorbid developmental impairments. *British Journal of Psychiatry, 166*, 489–495.

Kane, J. (1987). Treatment of schizophrenia. *Schizophrenia Bulletin, 13*, 171–186.

Kety, S. S., Rosenthal, D., Wender, P. H., & Schulsinger, F. (1976). Studies based on a total sample of adopted individuals and their relatives: Why they were necessary, what they demonstrated and failed to demonstrate. *Schizophrenia Bulletin, 2*, 414–428.

Kolvin, I. (1971). Studies in the childhood psychoses: I. Diagnostic criteria and classification. *British Journal of Psychiatry, 118*, 381–384.

Kolvin, I., Ounsted, C., Humphrey, M., & McNay, A. (1971). Studies in the childhood psychoses: II. The phenomenology of childhood psychoses. *British Journal of Psychiatry, 118*, 385–395.

Kraepelin, E. (1915). *Clinical psychiatry* (A. R. Eiefendorf, Trans.). Delmar, NY: Scholars' Facsimiles and Reprints (Original work published 1883).

Kumra, S., Frazier, J. A., Jacobsen, L. K., McKenna, K., Gordon, C. T., Lenane, M. C., Hamburger, S. D., Smith, A. K., Albus, K. E., Alaghband-Rad, J., & Rapoport, J. (1996). Childhood-onset schizophrenia: A double-blind clozapine-haloperidol comparison. *Archives of General Psychiatry, 53*, 1090–1097.

Liberman, R. P., DeRisi, W. J., & Mueser, K. T. (1989). *Social skills training for psychiatric patients*. New York: Pergamon.

Marder, S. R., Wirshing, W. C., Mintz, J., McKenzie, J., Johnston, K., Eckman, T. A., Lebell, M., Zimmerman, K., & Liberman, R. P. (1996). Two-year outcome of social skills training and group psychotherapy for outpatients with schizophrenia. *American Journal of Psychiatry, 153*, 1585–1597.

Maziade, M., Bouchard, S., Gingras, N., Charron, L., Cardinal, A., Roy, M., Gauthier, B., Tremblay, G., Cote, S., Fournier, C., Boutin, P., Hamel, M., Merette, C., & Martinez, M. (1996). Long-term stability of diagnosis and symptom dimensions in a systematic sample of patients with onset of schizophrenia in childhood and early adolescence. II: Positive/negative distinction and childhood predictors of adult outcome. *British Journal of Psychiatry, 169*, 371–378.

McClellan, J., & Werry, J. (1994). Practice parameters for the assessment and treatment of children and adolescents with schizophrenia. *Journal of the American Academy of Child and Adolescent Psychiatry, 33*, 616–635.

McKenna, K., Gordon, C. T., Lenane, M., Kaysen, D., Fahey, K., & Rapoport, J. L. (1994). Looking for childhood-onset schizophrenia: The first 71 cases screened. *Journal of the American Academy of Child and Adolescent Psychiatry, 33*, 636–644.

McKenna, K., Gordon, C. T., & Rapoport, J. L. (1994). Childhood-onset schizophrenia: Timely neurobiological research. *Journal of the American Academy of Child and Adolescent Psychiatry, 33*, 771–781.

Mednick, S. A., Machon, R. A., Huttunen, M. O., & Bonett, D. (1988). Adult schizophrenia following prenatal exposure to an influenza epidemic. *Archives of General Psychiatry, 45*, 189–192.

Miklowitz, D. J., Velligan, D. I., Goldstein, M. J., Nuechterlein, K. H., Gitlin, M. J., Ranlett, G., & Doane, J. A. (1991). Communication deviance in families of schizophrenic and manic patients. *Journal of Abnormal Psychology, 100*, 163–173.

Orvaschel, H. (1995). *K-SADS-E* (Fifth revision, *DSM-IV*). Center for Psychological Studies, Nova Southeastern University, Ft. Lauderdale, FL.

Petty, L. K., Ornitz, E., Michelman, J. D., & Zimmerman, E. G. (1984). Autistic children who become schizophrenic. *Archives of General Psychiatry, 41*, 129–135.

Pogue-Geile, M. F., & Harrow, M. (1985). Negative symptoms in schizophrenia: Their longitudinal course and prognostic significance. *Schizophrenia Bulletin, 11*, 427–439.

Pool, D., Bloom, W., Wielke, D. H, Roninger, J. J., & Gallant, D. M. (1976). A controlled evaluation of loxitane in seventy-five adolescent schizophrenic patients. *Current Therapeutic Research, 19*, 99–104.

Quintana, H., & Keshavan, M. (1995). Case study: Risperidone in children and adolescents with schizophrenia. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*, 1292–1296.

Rapoport, J. L., & Ismond, D. R. (1996). *DSM-IV training guide for diagnosis of childhood disorders*. New York: Brunner/Mazel.

Remschmidt, H., Martin, M., Schulz, E., Gutenbrunner, C., & Fleischhaker, C. (1991). The concept of positive and negative schizophrenia in child and adolescent psychiatry. In A. Marneros, N. C. Andreasen, & M. T. Tsuang (Eds.), *Negative versus positive schizophrenia* (pp. 219–242). Berlin: Springer.

Remschmidt, H. E., Schulz, E., Martin, M., Warnke, A., & Trott, G.-E. (1994). Childhood-onset schizophrenia: History of the concept and recent studies. *Schizophrenia Bulletin, 20*, 727–745.

Russell, A. T. (1994). The clinical presentation of childhood-onset schizophrenia. *Schizophrenia Bulletin, 20*, 631–646.

Russell, A. T., Bott, L., & Sammons, C. (1989). The phenomenology of schizophrenia occurring in childhood. *Journal of the American Academy of Child and Adolescent Psychiatry, 28*, 399–407.

Schmidt, M., Blanz, B., Dippe, A., Koppe, T., & Lay, B. (1995). Course of patients diagnosed as having schizophrenia during first episode occurring under age 18 years. *European Archives of Psychiatry and Clinical Neuroscience, 245*, 93–100.

Singer, M. T., & Wynne, L. C. (1963). Differentiating characteristics of parents of childhood schizophrenics, childhood neurotics, and young adult schizophrenics. *American Journal of Psychiatry, 120*, 234–243.

Singer, M. T., & Wynne, L. C. (1965). Thought disorder and family relations of schizophrenics: IV. Results and implications. *Archives of General Psychiatry, 12*, 201–212.

Spencer, E. K., & Campbell, M. (1994). Children with schizophrenia: Diagnosis, phenomenology, and pharmacotherapy. *Schizophrenia Bulletin, 20*, 713–725.

Strandburg, R. J., Marsh, J. T., Brown, W. S., Asarnow, R. F., & Guthrie, D. (1994). Information-processing deficits across childhood- and adult-onset schizophrenia. *Schizophrenia Bulletin, 20*, 685–695.

Strober, M., Lampert, C., Schmidt, S., & Morrell, W. (1993). The course of major depressive disorder in adolescents: Recovery and risk of manic switching in a 24-month prospective, naturalistic follow-up of psychotic and non-psychotic subtypes. *Journal of the American Academy of Child and Adolescent Psychiatry, 32*, 34–42.

Tienari, P., Sorri, A., Lahti, I., Naarala, M., Wahlberg, K., Moring, J., Pohjola, J., & Wynne, L. C. (1987). Genetic and psychosocial factors in schizophrenia: The Finnish Adoptive Family Study. *Schizophrenia Bulletin, 13*, 477–484.

Volkmar, F. R. (1996). Childhood and adolescent psychosis: A review of the past 10 years. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 843–851.

Volkmar, F. R., & Cohen, D. J. (1991). Comorbid association of autism and schizophrenia. *American Journal of Psychiatry, 148*, 1705–1707.

Volkmar, F. R., Cohen, D. J., Hoshino, Y., Rende, R. D., & Paul, R. (1988). Phenomenology and classification of the childhood psychoses. *Psychological Medicine, 18*, 191–201.

Volkmar, F. R., & Schwab-Stone, M. (1996). Annotation: Childhood disorders in *DSM-IV*. *Journal of Child Psychology and Psychiatry, 37*, 779–784.

Watkins, J. M., Asarnow, R. F., & Tanguay, P. E. (1988). Symptom development in childhood onset schizophrenia. *Journal of Child Psychology and Psychiatry, 29*, 865–878.

Weinberger, D. R. (1987). Implications of normal brain development for the pathogenesis of schizophrenia. *Archives of General Psychiatry, 44*, 660–669.

Werry, J. S. (1996). Childhood schizophrenia. In F. R. Volkmar (Ed.), *Psychoses and pervasive developmental disorders in childhood and adolescence* (pp. 1–48). Washington, DC: American Psychiatric Press.

Werry, J. S., McClellan, J. M., Andrews, L. K., & Ham, M. (1994). Clinical features and outcome of child and adolescent schizophrenia. *Schizophrenia Bulletin, 20*, 619–630.

Werry, J. S., McClellan, J. M., & Chard, L. (1991). Childhood and adolescent schizophrenic, bipolar, and schizoaffective disorders: A clinical and outcome study. *Journal of the American Academy of Child and Adolescent Psychiatry, 30*, 457–465.

World Health Organization (1967–69). *Manual of the international statistical classification of diseases, injuries, and causes of death: Based on the recommendation of the World Health Organization Eighth Revision Conference*. Geneva, Switzerland: World Health Organization.

World Health Organization (1978). *Mental disorders: Glossary and guide to their classification in accordance with the ninth revision of the international classification of diseases*. Geneva: World Health Organization.

World Health Organization (1992). *International statistical classification of disease and related health problems* (Vol. 1, 10th revision). Geneva, Switzerland: World Health Organization.

Zahn, T. P., Jacobsen, L. K., Gordon, C. T., McKenna, K., Frazier, J. A., & Rapoport, J. L. (1997). Autonomic nervous system markers of psychopathology in childhood-onset schizophrenia. *Archives of General Psychiatry, 54*, 904–912.

Zahn, T. P., Jacobsen, L. K., Gordon, C. T., McKenna, K., Frazier, J. A., & Rapoport, J. L. (in press). Attention deficits in childhood-onset schizophrenia: Reaction time studies. *Journal of Abnormal Psychology*.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.01
Clinical Formulation

GILLIAN BUTLER
University of Oxford, Warneford Hospital, UK

6.01.1 INTRODUCTION 1

6.01.2 DEFINITIONS: WHAT IS A FORMULATION? 2

6.01.2.1 Main Principles 2
6.01.2.2 Formulation and Diagnosis: Assumptions 4
6.01.2.3 Formulation and Diagnosis: Controversial Issues 5
6.01.2.4 The Difference Between a Formulation and a Model 6
6.01.2.5 Types of Formulation 7
6.01.2.6 Levels of Formulation 7

6.01.3 PURPOSES: WHAT A FORMULATION IS FOR 8

6.01.3.1 Understanding: The Overall Picture or Map 9
6.01.3.2 Prioritizing Issues and Problems 9
6.01.3.3 Planning and Selecting Intervention Strategies 10
6.01.3.4 Predicting Responses and Difficulties 10
6.01.3.5 Determining Criteria for Successful Outcome 11
6.01.3.6 Thinking About Lack of Progress 11

6.01.4 METHODS: HOW TO CONSTRUCT A FORMULATION 12

6.01.4.1 Sources of Information 12
6.01.4.2 Putting the Information Together 14
6.01.4.3 Key Factors and Basic Elements 17
6.01.4.4 Issue of Completeness 18
6.01.4.5 Conceptualizing Processes of Change 19

6.01.5 ACCURACY: HOW TO TELL IF A FORMULATION IS RIGHT 20

6.01.5.1 Criteria of Accuracy 20
6.01.5.2 Questions for Research 20

6.01.6 USING THE FORMULATION: PRACTICAL ISSUES 21

6.01.6.1 The Value of Organizing and Clarifying 21
6.01.6.2 Developing an Internal Supervisor 21
6.01.6.3 Communicating a Formulation 22

6.01.7 CONCLUDING DISCUSSION 22

6.01.8 REFERENCES 23

6.01.1 INTRODUCTION

> Patients come to psychotherapy because they are demoralized by the menacing meanings of their symptoms. The psychotherapist collaborates with the patient in formulating a plausible story that makes the meanings of the symptoms more benign and provides procedures for combatting them, thereby enabling the patient to regain his morale. (Frank, 1986)

Although not all therapists would be happy with the idea that they are "formulating a plausible story," the process of clinical formulation remains the lynch pin that holds theory and practice together. This is agreed by proponents of most major therapeutic traditions: for example, behavior therapy (Turkat & Maisto, 1985; Wolpe & Turkat, 1985), psychodynamic therapy (Barber & Crits-Christoph, 1993; Perry, Cooper, & Michels, 1987; Silberschatz, Fretter, & Curtis, 1986), family therapy (Minuchin, 1974), cognitive therapy (Freeman, 1992; Persons, 1989, 1993), cognitive analytic therapy (Ryle, 1978, 1990), and interpersonal therapy (Klerman, Weissman, Rounsaville, & Chevron, 1984). The attempt to construct and use a clinical formulation is central to the work of therapy. Various methods for systematizing the processes involved have recently been proposed (Horowitz, 1989; Luborsky & Crits-Christoph, 1990) and, thinking specifically about the issues involved in psychotherapy integration, Goldfried (1995) has put forward a case for developing a common language for case formulation that is independent of theoretical orientation. Personal discussions of many kinds may be more or less valued and helpful to someone experiencing a difficulty, including the informal advice traded between friends, but one of the major differences between informal discussions and responsible clinical practice is that they do not make use of the process of formulation. The attempt to formulate a case, so as to apply an appropriately chosen method of intervention in the light of a particular theory, is one of the activities that makes therapists, as opposed to friends, accountable for their practice.

This chapter discusses issues concerning clinical formulation that are relevant to therapists from different theoretical backgrounds. However, the illustrations of the general points made will largely be drawn from the author's own experience and will therefore reflect the author's original cognitive-behavioral training, together with a more recent interest in exploring possibilities for integration between different kinds of psychotherapy.

6.01.2 DEFINITIONS: WHAT IS A FORMULATION?

6.01.2.1 Main Principles

A formulation is the tool used by clinicians to relate theory to practice. Clinicians use theoretical as well as practical knowledge to guide their thinking about the problems and difficulties presented by the people who come to them for help, and this combination of ideas helps them decide how best to help those people. However, although the theories are relatively simple and clear—admittedly to varying degrees—the information brought to treatment, and gathered during the process of assessment, is always complex and often unclear. The process of marrying theory and practice is therefore fraught with difficulty. As well as having different reasons for requesting psychological help, people vary in their ability to describe or name their difficulties, in their histories and relationships with their families, friends and colleagues, in their ability to relate to a therapist, degree of psychological-mindedness, and emotional expressiveness. As well as having different theories, training, and clinical experience, therapists vary in the ways in which they understand, communicate with, and relate to their patients. Therapists bring with them to therapy specific skills, expertise, and information, and also their individual personalities and inclinations. The process of formulation is influenced by all these disparate factors, and this makes learning how to formulate a case with the necessary objectivity, clarity, and attention to the individual to guide a successful treatment one of the most fascinating, rewarding, and difficult tasks faced by clinicians.

The assumption that many clinicians of different orientations probably share about the psychological difficulties of others is this: at some level it all makes sense. Even though our understanding of the processes involved, and particularly of their inter-relationships, is incomplete, this assumption was given a simple, and relatively uncontroversial, diagrammatic form by Padesky and Mooney (1990). The difficulties that people describe to their therapists have four inter-related aspects (cognitive, affective, behavioral, and physiological), and change in any one of these variables affects all of the others, as shown by the bidirectional arrows in Figure 1. So, taking anxiolytic medication can make one feel calmer, think about problems more constructively, and do some of the things that previously seemed too difficult or overwhelming. Feeling more cheerful can lighten ones step, help one to feel more optimistic, and relate more productively to others. Changing ones perspective—or way of thinking—can provide the sort of new outlook that helps to dissipate distress, reduce tension, and encourage constructive activity, and so on. The four ways in which aspects of psychological life are conventionally categorized reflect the internal workings and psychological state of a person at a particular point in time. This person is at the same time relating to the external world through a personal social, political, and historical context. The factors that determine this context, and fashioned it to be the way that it now is, are not easy for psychological therapists to know about: hypotheses for explaining and understanding the way they interact with each of the four types of phenomena have been made. The overall configuration is the source of the narrative, or story, that a person brings to therapy. If we understood the rules governing the relationships between all these factors we would, no doubt, be better able to help our patients.

The business of therapy, to a large extent, involves intervening to facilitate change in (at least) one of the four main aspects of psychological life shown in Figure 1, and different kinds of therapy attend differently to these different aspects, entering the process of change through different gateways. The intention, however, is much the same—to help people solve the problem or problems that they bring to therapy. Pharmacological and traditional behavioral therapies provide perhaps the clearest examples as the methods that they use, and the formulations upon which these methods are based, can be isolated relatively easily. Cognitive therapies, which adopt both cognitive and behavioral methods, operate on at least two levels. They may concentrate on identifying and reexamining particular thoughts, thereby changing feelings and behavior, and/or they may focus on underlying meanings and beliefs and adopt more sophisticated and complex methods of intervention, often related to those used in more dynamic and experiential traditions. Experiential therapies make specialized use of the medium provided by the feelings and thoughts arising in the present context of therapy, and work with these to facilitate a dynamic process of change. In order to do this, it becomes essential to think about, and to formulate, what happens in the relationship between the two people involved in therapy—methods which were originally described and understood by proponents of the various psychodynamic schools of therapy. Interpersonal therapy and systems therapy also formulate problems in terms of relationships between the person requesting help and others around them, and use this understanding to help people change as they wish. All of these methods initiate the process of change in different ways, determined by the way in which they understand, or formulate, the problem presented, and it is this understanding that determines what therapists do—what steps they take to alleviate the problem.

The point is that the way in which a formulation is constructed will be influenced by the point at which a therapist enters, and attempts to influence, this dynamic relationship between these main aspects of psychological life. Some general points are important:

(i) each aspect influences all of the others, so none of the therapies has the exclusive aim of changing one factor. Rather, by focusing the process of change in one place, the aim is to bring about the change that the patient desires—usually to "feel better," in all the relevant respects.

(ii) The main medium of therapy is language—what one person says or suggests, to another. To this extent, the cognitive, implicational context within which therapies take place provides the basis for the way in which the presenting problems will be formulated.

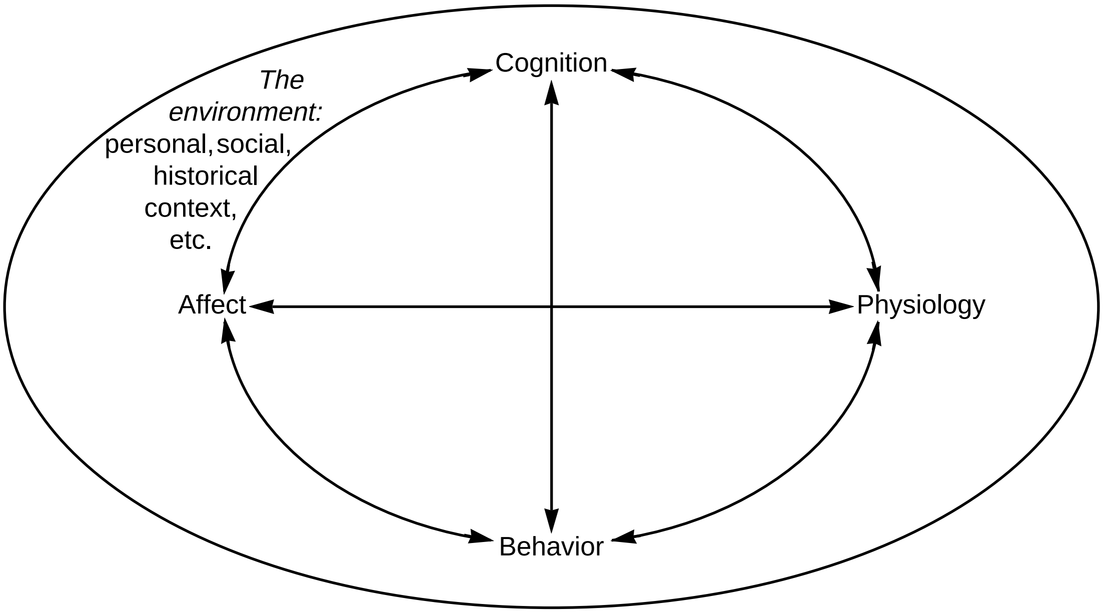

Figure 1 Inter-relationships between aspects of functioning (Padesky & Mooney, 1990).

(iii) Understanding of other people, and hence the ability accurately to formulate their problems, develops within the context of the relationship between them, mediated by factors such as trust and acceptance as well as by language.

(iv) Our understanding of the ways in which the aspects of psychological life are integrated is partial. At this point in time, psychology is an imperfect but developing science. The implication of this is that formulations for the purpose of therapy have to be speculative.

Formulations can best be understood as hypotheses to be tested, and the most obvious, if not the most logical, test of a formulation will be the response to the selected interventions. This is not to say that an expected change following a specific intervention proves that the formulation on which it was based is accurate. Unfortunately, the reasons why change occurs are far more complex and difficult to discern than this. However, the formulation used in this way is perhaps the main tool that the therapist has from which to draw such conclusions in the individual case. Thus, working in an open-minded way with a formulation provides a means of contributing as a therapist to the scientific endeavor involved in finding out which are the best, most effective, and most efficient, methods of treatment.

Although a formulation provides the link between theory and practice, it does so at a different level of generality. A theory is the source of general explanations and general hypotheses, whereas a formulation is specific to the person to whom it applies, and therefore is the source of more specific explanations and hypotheses. The specificity of the formulation is the source of ideas about the selection of specific interventions and about how to adapt them for use with a particular person. It is for these reasons that Wolpe and Turkat (1985) describe a formulation as a theoretically guided way of structuring the information concerning a patient's problem. It reflects the product of taking an individual approach to clinical phenomena and combining this with knowledge of relevant theories, scientific principles, and research findings. It involves imposing an explanatory system upon the material presented, and raises questions concerning the degree to which this explanatory system should reflect every aspect of a problem. One view is that it should reflect everything, including a patient's past development, characteristic ways of behaving and forming relationships, emotions, beliefs, assumptions, attitudes, self-evaluations, expectations, attributions, appraisals, and so on. In practice, the degree of elaboration required depends upon the purpose for which the formulation is made. At this point it is probably sufficient to enunciate one of the principles that will run through this chapter—that of parsimony. In principle, it is always better, and more useful, to keep the formulation as simple as possible. The temptation to elaborate a formulation is strong, especially when dealing with complex cases. However, the simpler and clearer it is the more readily will its implications be seen and the easier it will be to use.

Theoretically speaking, the principles that guide the practice of formulation are derived from the way in which the concept is defined. The three main ones to be proposed here are:

(i) A formulation should be based on a theory, reflecting an attempt to put the theory into practice.

(ii) A formulation should be hypothetical in nature, so that it can be modified by information gained during the course of treatment.

(iii) A formulation should be as parsimonious as possible.

6.01.2.2 Formulation and Diagnosis: Assumptions

In psychological practice there appears to be a common assumption that only those patients who participate in research trials have simple diagnoses, for example, of the kinds defined in the various versions of the *DSM*. Diagnostic systems are useful for ensuring that the populations studied in different places are similar in the relevant respects, and they are useful for insurance purposes, but from the point of view of the therapist they have limitations in that they rarely provide specific implications for treatment. Besides, unselected samples of patients often do not have single, clear problems—indeed informally they are commonly said to "fulfill criteria for an average of 2.3 diagnoses." A formulation, however, is designed precisely to fit the individual and is intended to help therapists to derive theoretically-based hypotheses about factors that contribute to causing and maintaining their specific problems—to explain as well as to describe. Therefore, the argument runs, diagnoses are less useful than formulations, from which specific treatment implications can be derived, and they may be less necessary than formulations. For example, one depressed person's sense of failure may be triggered by an inability to live up to exacting standards and another person's by an inability to form close relationships (for any number of reasons, which may be discovered during therapy and included in the formulation). Only having the diagnosis tells the therapist nothing about this difference, and ignoring the difference will reduce the chances of achieving a satisfactory outcome.

This argument has much to recommend it to the therapist, especially as diagnoses are largely atheoretical descriptions, and therapists can use their theoretical knowledge to construct formulations that are clinically useful. This does not mean thinking anew with each patient, but keeping in close touch with theoretical and clinical research so that, for example, empirically validated treatments can be selected when the diagnosis suggests they would be appropriate, and individual formulations then used to specify details of their application. Learning to work with a formulation instead of relying on a diagnosis also has advantages when the problems presented are unusually rare or complex and do not fit readily into a diagnostic system, or when the system does not succeed in "carving nature at the joints," and the demarcation between one diagnosis and another is difficult to establish. Of course there are difficulties with this point of view. Seen from the patient's perspective, over-reliance on the process of formulation may involve a degree of risk.

What if the theory is wrong? Or if the therapist is unclear about it? Or susceptible to bias? Or unable to come up with an adequate formulation? Or attempts to combine one theory with another without understanding sufficiently well the implications of doing so—as when borrowing from experiential or dynamic ideas when doing cognitive therapy for instance? This risk can be reduced by formalizing the requirements of responsible clinical practice—by providing adequate training and supervision, by clarifying ethical guidelines, and by defining criteria for professional accountability, including the expectation that practitioners will keep in touch with the literature relevant to their practice. Ultimately though, the mysterious faculty of clinical judgment has also to be brought into play. Without this—whatever it is—clinicians may well run into difficulties, both making and using formulations.

The implication of this argument is primarily that, much of the time, formulations are more useful than diagnoses, provided that therapists are well versed in the theories they are using, and that diagnoses, which can after all convey a large amount of information in a few words, may help to streamline the process of assessment, and may guide decisions about treatment in relatively straightforward cases. For example, knowing someone is socially phobic directs attention towards a fear of being humiliated or embarrassed, and knowing the diagnosis is of bulimia nervosa focuses attention on over-concern with shape and weight (among other things). Underlying problems of self-consciousness and poor self-esteem may be relevant in both cases, so the assessment which provides an adequate basis for a formulation, and for a specific treatment plan, must cover more than the criteria for inclusion and exclusion that determine whether or not someone qualifies for a diagnosis.

6.01.2.3 Formulation and Diagnosis: Controversial Issues

The assumption behind the argument presented above is that a treatment plan based on a formulation will have a better chance of success than one based on a diagnosis. However, there is considerable debate about this issue, and some recent research suggests that the assumption could be false. Schulte, Künzel, Pepping, and Schulte-Bahrenberg (1992) and Schulte (1997) found that patients with phobias, assigned to a standardized treatment (exposure *in vivo*) on the basis of their diagnoses, responded at least as well as, and possibly better than, patients whose treatments had been selected on the basis of individual problem analyses. With this finding in mind, Wilson (1996, 1997) summarized the arguments for using manual-based, empirically-validated treatments, also selected on the basis of diagnoses, and argued that there are inherent limitations involved in basing treatment on idiographic case formulation. As he points out, making formulations involves making judgments and judgments are fallible. They are demonstrably susceptible to bias and using them introduces an additional source of error. It would be better, he argues, to adopt an actuarial approach to assessment and treatment as this is more likely to result in a superior outcome than using clinical judgment, at least when treatment manuals are available.

The issue is complex (Beutler, Williams, Wakefield & Entwistle 1995; Hayes, Follette, Risley, Dawes & Grady, 1995; Norcross, Alford, & DeMichele, 1992; Seligman, 1995; Stricker & Trierweiler, 1995), and differences will not be settled here. Nevertheless, it is useful to clarify the basis of the disagreement, as two issues are frequently confounded. The first concerns the failure of practicing clinicians to adopt standardized practices and the second concerns the dangers of over-reliance on individual formulations. Those who argue against the use of formulations seem to forget that it is the job of practicing clinicians to bridge the gap between science and practice, and in doing so to balance the requirements of recommended procedures with clinical flexibility. A formulation, as defined above, is intended to facilitate this process—to assist the clinician in adapting the procedure to the particular circumstances. When treatments so adapted are reported to be less effective than expected, then many factors in addition to formulation could contribute to this finding. These include the quality, integrity, structure, and delivery of the treatment, the accuracy with which the effects of treatment can be measured, and the relevance of the measures used to the outcome desired by the patient. Proponents of the view that treatments can be selected on the basis of diagnoses alone seem to assume that case formulation is idiographic, in the sense that making one is unconstrained by theoretical ideas and using it to select interventions is independent of the findings of clinical research.

Neither of these points is accepted here. Instead it is argued that individual case formulation is always relevant, even when applying a manual-based treatment (examples will be found below). It is also argued that formulations have to be rooted in theory to be useful, and that using clinical judgment is not providing a licence for subjectivity, but recognizing that at least some of the time clinicians will not be able to follow the rule book, even when there is one. Then they have to use their judgment. In doing so, they can appeal to many sources of understanding, including theories about psychological dysfunction, and their knowledge of the relevant literature. As Stricker and Trierweiler (1995, p. 997) put it "it is likely that the practitioner always will be required to go beyond firm and available scientific knowledge"—less so when treating phobias than when treating a complex of depression and anxiety in someone with a dependent personality type, and not without keeping in touch with scientific advance—but individual judgment and case formulation remain indispensable clinical tools. Using these tools does not exempt the practitioner from being aware of the pitfalls of basing decisions about treatment on anecdotal case material, intuition, or subjective impression. On the contrary, working with a formulation that can be explained to others provides a check on the use of too much speculation and too many far-fetched inferences.

Therapists need to speak about their patients' problems in many settings and contexts, and to do so can make use of any of the available systems—labels, diagnoses, descriptions, and formulations. Labels (e.g., manipulative, hysterical, narcissistic, personality disordered) are efficient but can bring assumptions with them (and in these examples, assumptions that may not be to the advantage of the person being labeled). Diagnoses reflect agreed systems of categorization and for the most part are based on particular kinds of descriptions rather than on theories. They may or may not be subject to the same disadvantages as labels. Formulations differ in that they bring together the products of theoretical knowledge and clinical judgment. Their theoretical basis reflects ideas about the factors that cause and maintain problems, and that precipitate or prolong particular episodes of distress. This theoretical basis provides a framework for the type of personal, individual formulation on which precise decisions about treatment can be based. Their advantages and disadvantages are discussed further below.

6.01.2.4 The Difference Between a Formulation and a Model

Models are ways of conceptualizing particular disorders (e.g., the cognitive hypotheses of obsessive-compulsive disorder and of health anxiety described by Salkovskis (1996), or of formulating particular patterns of functioning (e.g., the role–relationship models developed by Horowitz, Eells, Singer and Salovey (1995) or the functional analytic causal model of Haynes, Uchigakiuchi, Meyer, Orimoto, and Blaine (1993). Models, as understood here, are constructed from a particular perspective, so there are separate cognitive models of panic disorder (Clark, 1988) and social phobia (Clark & Wells, 1995), and the psychopharmacological or interpersonal psychotherapy models of panic disorder differ from the cognitive model. These differences are valuable in that they stimulate useful research, as well as the development of sets of coherent treatment strategies. Using the cognitive model of panic disorder as an example, this would suggest that catastrophic misinterpretation of bodily symptoms plays a crucial role in triggering panic attacks, and that understanding this will help people who suffer from panic disorder to identify the symptoms that trigger their panics. They will then be in a position to think again about the meaning of these symptoms, and to reinterpret them in terms of (harmless but distressing) panic rather than of real, impending catastrophe. In order to facilitate the therapeutic process, the model has to be translated into a conceptualization (or formulation), and structured systems for doing this can be developed, as in this case has been done by Dattilio (1994). So the model provides guidelines for an individual formulation which encourages a new explanation—the leap in my heart could be a response to the coffee I have just drunk, or a normal arrhythmia that I notice more readily than I used to because it frightens me, and not a sign of imminent cardiac crisis.

Although a model has implications for treatment, it differs from a formulation in that it operates at a different level of generality, and has a different content. So, the way in which a formulation applies to particular people will depend upon their personal history and circumstances. One person's panic may be triggered by leaps in the heart and another's by losing concentration when being spoken to (and a third may find that memories of traumatic incidents, flashbacks, or nightmares precipitate panic, possibly because they trigger associated sensations that then trigger the panic attacks). There will in practice always be exceptions to the rule, cases in which, for example, no sensational trigger can be identified. Then the clinician may be best advised to base the formulation on a higher level theory rather than on the specific model—in this case on the general theory that cognitions, including meanings, are closely related to feelings and behavior, and that changing one is likely to change the others. Thus the formulation illustrates, in ways that are clinically relevant, how the model applies, and does not apply, to the case. It assists the therapist in looking for particular theoretical constructs or processes (catastrophic misinterpretations in this example), and also in making a judgment about the degree to which the case is typical.

Atypical cases arise when patients have more than one difficulty—social anxiety as well as panic disorder for example—or when they have especially complex or rare problems such as panic attacks in the context of avoidant or borderline personality disorder. Then, conceptually speaking, it may be more useful to draw on more than one model to construct a single formulation, or to look for models with a higher order of generality. Writing about psychodynamic formulation, Perry et al. (1987) point out that overlapping models of mental functioning may emphasize different aspects of development and psychopathology. They distinguish ego-psychological, self-psychological, and object relations models, and make the important point that a certain amount of trial and error may be needed in constructing a formulation that explains the presenting data: "the absence of a meta-model to explain all data makes this trial and error unavoidable" (p. 546). What clinicians are looking for in a formulation is a way of explaining and understanding the relationship between a patients' inner lives and their outer lives that is the product of their personal history, explains present difficulties, and guides future therapy. Their sources in this search include knowledge of diagnostic systems, of relevant theoretical models, and of outcome research, as well as information about the individual case—otherwise they would have to reinvent the wheel each time.

6.01.2.5 Types of Formulation

Typically, different therapeutic schools are thought to use different types of formulation. In general, behavioral and cognitive therapies make use of more mechanistic formulations, based on theories about learning and detailed functional analysis (Hayes & Follette, 1992), or on theories about processes such as the supposedly circular relationships between thoughts and feelings, and more dynamic therapies employ more narrative-based formulations, placing current problems in the context of a developmental history. Some systemic and experiential approaches to therapy adopt a third, essentially dynamic, approach, claiming that formulations have constantly to be re-formed in the present, as therapy focuses on moment-to-moment events (Goldman & Greenberg, 1997). They also point out that the process of formulation can be dangerous and limiting when it makes use of preset categories and ideas. A constantly changing situation then appears to be fixed, and opportunities for change may be obscured (Eells, 1996; Rosenbaum, 1996). However, the process of formulation is still thought to be essential, and its main purpose is still to look for patterns and links that assist in understanding, and to provide ideas about how to bring about change. So, distinctions can be applied too rigidly. The developmental history of a problem or a person, or the narrative, is always relevant (Nicholson, 1995; White, 1989), although it may be understood in different ways, and so are ideas about the mechanisms that precipitate an episode of distress or perpetuate a problem. Overt differences between types of formulation are therefore relatively unimportant to an understanding of the term, and of the functions that the activity of formulating a case performs for the therapist.

6.01.2.6 Levels of Formulation

When making a formulation, it is necessary to think at many different levels, and the number of levels postulated obviously varies with the theory being applied. Taking an example from cognitive-behavioural therapy (CBT) to illustrate the point, at the most superficial level, or the level of "overt difficulties" (Persons, 1989), the main task is to define the problems and the ways in which they are maintained, usually in terms of vicious cycles. Someone who feels depressed may withdraw from company, think about being all alone, and become increasingly depressed. Even such a simple formulation suggests a focus for intervention—working to reduce the withdrawal. Thinking about the factors that precipitated the depression adds another level to the formulation. The person might have become depressed when their job required them to move to a new place, when they got divorced, or when their children left home—factors that would demand different types of adaptation, to be promoted by the therapist in different ways. Stressors are additive, so many factors may be involved, and an apparently minor stressor may be the straw that broke the camel's back (and relatively irrelevant to the formulation), or it may reflect a particular personal vulnerability. Factors that predispose someone to become depressed, biological as well as psychological factors, add a further level, and the way in which these are understood, and formulated, will again influence the selection of interventions. At the most profound level of all, assuming that "at some level it all makes sense," the formulation is supposedly capable of reflecting the meaning of structures through which people interpret and think about, remember and recount, their experiences, and theoretical assumptions about the origin of these things.

This is the standard way in which psychiatric formulations have traditionally been made—in terms mainly of predisposing, precipitating, and perpetuating factors. However, there are yet other levels to consider, reflecting social, cultural, and historical factors. Social assumptions ("men should not show their feelings" or "women are bad organizers") influence the views of therapists as well as patients, and cultural assumptions may or may not be shared between therapist and patient. Some cultures, for example, do not share the common Western therapeutic goal of autonomy, especially for women. Others assume that a relationship between a professional person and their client is one involving activity—or authoritative pronouncements—on the one side, and passivity—or receptivity—on the other. In addition, different hierarchies of values can interfere profoundly with the therapeutic process. An example in our culture is when someone thinks it more important to avoid giving offence than it is to tell the truth. Although it is never possible to stand outside all of these factors, making a formulation helps therapists to think about them, to identify them clearly, and to become aware of their potential influence on the interpretation of other people's circumstances. It can help therapists to ensure that the ways in which they understand problems and select interventions are not influenced by unwanted biases. Seen in this way, a formulation assists therapists in achieving a relatively objective stance.

Formulations are always made from a particular perspective—in the author's case made (usually) from a cognitive-behavioral perspective, and from that of a White woman of a certain age, living in Britain now, whose ways of thinking have been formed by her own learning and experience. A formulation is neither about fitting information about a patient to a predetermined formula, whether that formula be derived from a general theory or from a more specific model, nor is it a personal judgment, though both things are relevant. It is about developing the kind of understanding of another person, their circumstances and their difficulties, that enables a therapist to apply the theoretical knowledge acquired during training to help that person. There is no single right way of making a formulation. The general aim is to map the territory so that one can then explore the possibilities for change, and not to let these be influenced by factors that are irrelevant to, or unwanted by, the person who is receiving help.

6.01.3 PURPOSES: WHAT A FORMULATION IS FOR

One common view of the purpose of formulation is that it is for explaining the past, making sense of the present, and suggesting what to modify in order to influence the future. It can also be an important means of communicating understanding, either to the patient or to another professional, whether in the role of supervisor or colleague. However, its prime purpose is to help therapists to apply the theory they have learned to their practice (a comprehensive account of different approaches to formulation is given by Eells (1997).

In practice, there are many answers to the question "What is a formulation for?" The main functions of formulation are listed in Table 1. The main point is that making formulations is an essential, and not an optional, element of the therapeutic process. Formulations do not have to be 100% accurate or complete in order to be useful precisely because they provide a source of testable hypotheses. They can be changed when they turn out to be wrong—and nothing is lost by using a partial or partially mistaken formulation which can be improved and corrected as the process of therapy continues, and reveals the initial mistake. They guide questioning, and open the therapist's mind to the kind of understanding from which effective treatment strategies can be derived, applied, and evaluated. Therefore, the author would argue, that therapists should work with a formulation in mind right from the start. Ideas about people and their problems cannot be kept at bay or excluded, even when first meeting them or reading a referral letter about them. One way of trying to ensure that this information is openly received and accurately assessed is to engage immediately in the process of formulation—in applying both theoretical and clinical knowledge to the particular case. Just as when first visiting a new place, a rough sketch map may set one on the right road at first, but will need expanding and revising if it is to guide more detailed exploration.

6.01.3.1 Understanding: The Overall Picture or Map

A formulation "provides the map of the territory and once you have that you can use whatever vehicle you are most comfortable with" (Beck, 1991). Formulations, just like maps, provide an overall view (often in diagrammatic, conventional form) of something that it is not possible to see directly all at once—the wood as well as the trees. They indicate which are the important features, their size and shape, and the way in which they relate to each other. Mapping the territory is clearly the product of accurate assessment (see Section 6.01.4), and formulating enables therapists to make and to justify such statements as "this lack of energy is part of the depression," or "in this case the anxiety seems to be primary and the depression secondary." Similarly, formulations can indicate where information is missing and prompt appropriate questions: where did this low self-esteem come from? Why does it become apparent in the context of close personal relationships but not at work?

6.01.3.2 Prioritizing Issues and Problems

An overall formulation helps to differentiate what is essential from what is secondary in a general sense. It also helps in a more particular way to decide which issues or problems should be prioritized. Someone who believes that they cannot change is unlikely to remain engaged in therapy unless they can see the point of it. Creating hope, or the context for a developing relationship—something with a future, in which change is inherent and undeniable—then becomes a priority. Likewise, an initial assessment may indicate primarily that inability to trust people will make it hard to disclose distressing material, and building trust within the therapeutic relationship is necessary before a more detailed and accurate formulation can be made.

It is probably not unusual for patients and therapists to start the process of therapy with somewhat different priorities. Usually this problem can be overcome during assessment and those early stages of therapy during which goals become clear or are specifically agreed. But sometimes different priorities persist, and then the process of re-formulating can help to solve the problem. For example, an anxious and hypochondriacal patient who was worried, among other things, about seeing "floaters" in his visual field, started to respond well to treatment that was formulated in terms of his underlying sense of vulnerability. The formulation reflected the way in which his various concerns made him feel threatened, and think that he was at risk for being unable to handle a number of initially rather vaguely specified distressing eventualities. However, although his confidence increased, his distress about the floaters did not diminish. If anything it increased, in tune, it must be acknowledged, with the therapist's frustration when discussing this issue became his main priority. Focusing the work of one session on the meaning or understanding of this problem revealed (for reasons which later became clear) that visual anomalies for this person felt, in his words, "like a bereavement." Formulating this aspect of the problem in terms of loss rather than in terms of vulnerability changed the focus of treatment, which then became more productive. This example also illustrates how characteristics of the process of therapy can contribute to ideas about the formulation, especially in those cases in which change is not proceeding as well as might otherwise be expected.

Table 1 Summary of the purposes of formulation.

Clarifying hypotheses and questions
Understanding; providing an overall picture or map
Prioritizing issues and problems
Planning treatment strategies
Selecting specific interventions
Predicting responses to strategies and interventions; predicting difficulties
Determining criteria for successful outcome
Thinking about lack of progress; trouble shooting
Overcoming bias

6.01.3.3 Planning and Selecting Intervention Strategies

Once a hypothesis about how the presenting problem can be understood has been formulated, the most important functions of a formulation are in planning a treatment strategy and selecting appropriate methods of intervention. Persons (1989) provides some lucid examples: someone who avoids exercise because they are bad at time management, scheduling, or self-organization has a different problem to overcome, and needs to acquire different skills from the person who avoids exercise because they are embarrassed about their appearance; insomnia that is associated with the fear of letting go may require different interventions from insomnia that results from overcommitment.

The way in which a problem is formulated thus determines what should be done about it (Blackburn & Twaddle, 1996; Butler & Low, 1994; Eells, 1997). If avoidance maintains the problem, then facing the fear is likely to reduce it, and in individual cases the formulation helps to specify idiosyncratic aspects of the avoidance (the spider phobic who will not walk under trees; the social phobic who is more fearful of silence than of conversation). The general vicious cycle model is common to both—and indeed, a standardized method of treatment of proven effectiveness, exposure *in vivo*, is readily available. The individual formulation is still necessary because it specifies exactly what steps to encourage the person to take.

Planning overall strategies is just as important a product of formulation as the selection of specific methods of intervention, but is a more complex task, and requires of the therapist more than one level of understanding. The way in which depression or anxiety is understood may suggest, for instance, that it would be helpful to increase levels of activity before discussing thoughts associated with depression; or to build up a repertoire of coping skills before facing fears. Many such imprecations are based on clinical judgment (or clinical intuition) as much as on theoretical or experimental work, and in these cases it is especially important that they should be made clear by means of a formulation. For example, it is often said that when working with people who have suffered abusive experiences in childhood, one should help them to develop a variety of support systems, ways of dealing with intense feelings or suicidal impulses and of creating around themselves a sense of safety, before exploring memories of early traumatic experiences, and the meanings of such events, in depth.

Clearly, this overall strategy reveals assumptions about how the effects of these events can be understood, about the effects of talking about them, and the interventions used—assumptions which formulations clarify, and which are potentially amenable to research, but which will differ according to the therapist's theoretical orientation. A secondary purpose of clarifying the formulation and its function in selecting strategies and interventions is to facilitate evaluation of interventions.

6.01.3.4 Predicting Responses and Difficulties

Because a formulation reflects theoretical assumptions, it helps therapists make two kinds of predictions that are essential in therapy: to predict the effect of the intervention, assuming it is successfully applied, and to predict the stumbling blocks and difficulties that will be encountered during therapy. An anxious person treated during a clinical research trial (Butler, Fennell, Robson, & Gelder, 1991) held the belief that "all my ideas are bound to be wrong." She became more confident as she learned to identify her ideas, to act upon them, and consciously to evaluate the consequences of doing so. Her formulation enabled us to predict first that she would feel especially vulnerable and be likely to overgeneralize and catastrophize the consequences when she made mistakes, and second, that she was likely to find it especially difficult to apply the new strategy when relating to her partner, but easier to build up the necessary skills (and courage), and to increase her confidence, in the context of other relationships (including ours). Treatment in this case was guided by the requirements of a treatment manual, and the example illustrates the important role played by clinical formulation in the application of standardized treatments.

It is probably true to say that interpersonal difficulties are one of the most common sources both of patients' problems and of problems encountered during psychological therapy; for example, an ability to form superficial relationships without being able to sustain deeper friendships, or veering between passivity and aggression when interacting closely with others. Such difficulties also play their part within the therapeutic relationship, and they are much more easily dealt with if the processes involved have been understood in terms of the theory being used, and problems predicted in advance. Formulating helps people to recognize such patterns, to develop hypotheses about their origins, functions and effects, and to think about whether and how to engage in a process of change.

6.01.3.5 Determining Criteria for Successful Outcome

Theoretically a formulation provides the basis for hypotheses about what needs to change for someone to feel better, or the goals of therapy in the broad sense of the term. This is obvious when a theoretical model for the condition being treated is available, but the point applies more generally as well. The present version of the cognitive model of social phobia (Clark & Wells, 1995), for example, suggests that self-awareness, or self-focused attention, plays a central role in the disorder. In outline, when in a socially frightening situation a social phobic feels self-conscious, notices symptoms of anxiety and tries to keep safe. An individual formulation based on this model would specify the way in which this actually happens. For example, when speaking to others (e.g., colleagues during a lunch break), Marie became aware of the sound of her own voice, felt anxious, flustered, hot, and shaky, and found it hard to listen to what was being said. She thought other people must be able to see how nervous she felt and tried to fade into the background as quickly as possible (keeping herself safe by avoiding eye contact, saying little, speaking in a quiet voice). Both general and specific goals for change can be derived from thinking along these lines. In simple terms, if Marie can focus her attention outside herself, and listen without self-criticism to those around her, if she can reverse the safety behaviors (make appropriate eye contact, speak more audibly, move around freely), she will break the cycle and start to feel less anxious. The general criteria for change are reflected in the three elements of the model specified here, the self-awareness, safety behaviors, and symptoms of anxiety, and specific ones reflect the individual ways in which these factors are manifested in the case of Marie.

Of course this might not be the whole story. Marie's social anxiety may be based on a belief in her own unworthiness relative to others, and reflect an unhappy history of family relationships. Such formulations again indicate criteria for change—a sense of worthiness or the ability to form more satisfying relationships in the present. The difficulty here is that more abstract and general phenomena are harder to identify, define with any precision, and measure than more superficial and specific ones. Criteria for change are therefore more easily derived from formulations at lower than at higher levels of abstraction, and indeed the more specific the formulation the easier it will be to be clear about what exactly needs to change.

6.01.3.6 Thinking About Lack of Progress

There are many possible reasons for lack of progress in therapy, including working without making a formulation. The first line of defence when this happens must be to formulate or to reformulate the problem. The way in which this is done will have specific implications for the next steps in therapy. For example, if the problem is a long-standing, chronic one, it may be that much practice is needed and that it is unrealistic to expect faster change, in which case it may be important to think about how to keep the momentum of change going—about how to maintain hope and create the energy for change when doing so is difficult. If the original formulation was inaccurate or incomplete, the failure to change may suggest that different strategies and interventions are needed. When lack of progress leads to frustration, and the reactions of both the patient and the therapist interfere with subsequent progress, including these factors in the reformulation can reveal ways of overcoming them. Blocks in treatment are nearly always informative and formulation skills should be used to identify their specific nature.

Often this is complex and involves making hypotheses about past events, the exact nature of which can never be known. Possible formulations in these circumstances, often derived from a combination of observation and understanding of the apparent effects of the past on the present, can suggest which avenues to explore so as to make further progress. For example, a patient who provided a cold and dispassionate account of a childhood in which she was neglected, often frightened and sometimes threatened with physical abuse, appeared to have developed a variety of ways of controlling both the experience and the expression of her emotions. Many, but by no means all, of these ways were dysfunctional. A possible formulation of this case suggests that improvement will remain blocked unless or until she becomes able to experience and express the relevant feelings. Doing this is likely in the first instance, to precipitate periods of distress, and the precise implications for therapy to be derived from it will depend on both the skill and the emotional sensitivity of the therapist as well as on a willingness to adapt the formulation according to what happens.

Drawing these points together, it is clear that formulations cannot be treated as a matter of last resort, only to be constructed and worked on when the going gets difficult, when dealing with chronic problems, when treatment has apparently gone on too long, or when preparing to report to someone else. Formulations do not provide the answers to questions, but a rich source of questions and ideas of potential therapeutic value. They should not become the tool for applying a preconceived theoretical plan to someone for whom the plan does not fit, nor should they focus exclusively on someone's problems and difficulties. Accurate formulation takes account of a person's strengths as well as failures, talents and potential, as well as shortcomings and failures.

If formulations can be so useful it is surprising that so little attention has been devoted to them both within training programs and in the literature. One reason for this may be that formulations were supposed to follow logically from the processes of assessment and functional analysis, and additional skills were not often specified. A more important one is probably that formulating is difficult. As already indicated, in practice it involves exercising clinical judgment as well as the ability to relate theory to practice. Also, until recently, there was less communication between people with different theoretical backgrounds, and fewer challenges to think about alternative methods of formulating specific cases. So, the next important question is "How do you construct a formulation?"

6.01.4 METHODS: HOW TO CONSTRUCT A FORMULATION

The main reason for considering the purposes of formulation before thinking about how it should be done is that there is no single correct method—how you do it is in general determined by understanding the purposes that it serves, and in particular by the theoretical orientation of the therapist. The end product should enable the therapist to relate theory to practice in a way that can direct and inform the process of therapy, and the methods used vary enormously. For the student this is both confusing and liberating, as it demands creativity and the ability to deal with abstractions as well as the more mundane skills primarily involved in assessment. Assessment is a necessary step in the development of a formulation, but it is not a sufficient condition for it. Unfortunately, it is possible to assess, in the data collection sense, without developing a formulation.

6.01.4.1 Sources of Information

An account of presenting problems, informed by knowledge of psychological processes and diagnostic systems, provides a common starting point, and assessment covers all of the four aspects of functioning illustrated in Figure 1 and their determinants: cognition (thoughts, assumptions, attitudes, beliefs, images, etc.); affect, behavior, and physiological sensation; the present context for the ways in which these things are manifested; and an account of their background and associated developmental history. It also draws on information gathered during the process of referral, such as a summary of the problems as understood by the referrer, of the reasons for requesting help and of responses to treatment received so far, and on the impressions and observations made during the first encounter with the therapist when the processes of mutual interaction are set in motion.

Therapists use many skills in helping them to understand this material: theoretical knowledge; products of academic learning and professional training; and clinical judgment. The process of encapsulating this understanding in a formulation, which at first takes time and becomes quicker with practice, is facilitated by adopting a questioning stance. The aim would be to be able adequately to answer three of the key questions that patients ask: Why me? Why now? What keeps it going? and in doing so it helps to draw on a further set of questions central to the process of formulation, which therapists can pose either to themselves or to their patients: How do you understand that (or make sense of it)? What do you think is going on? How does this all fit together? What might be the missing links? What does that mean about you now? Is there a pattern here?

Formulations are useful in helping people to think again about their difficulties, and see them in a new (e.g., clearer, more realistic, or more illuminating) light, and the process of assessment potentially reveals the patient's present point of view. In order to develop an understanding of such personal and unique phenomena, it is particularly useful to pay attention to the ways in which people react to their experiences. Their comments provide a rich source of such information—"I have to keep control" or "I need to know I am succeeding" are remarks that suggest hypotheses about the self and about underlying processes and mechanisms. Ideas expressed about others, such as "she'll be miserable alone" may fit with assumptions that precipitate or maintain presenting problems. General comments of the kind "you have to conform or you can't get on" reveal attitudes that may (or may not) dominate within the real world in which the person lives. Expectations about the future, including those about the process and outcome of therapy, are also revealing: "I won't be able to do what is needed," "There are some things I would rather not talk about," "I'm relying on you to make me better." In order to formulate, it is important to understand the personal significance of experiences as well as their phenomenology. These comments illustrate well how the processes of formulation and assessment meet, and indeed may overlap. Therapists assess to find out about problems and their context, and they formulate differently according to what they think their findings mean.

Patients' comments may need clarifying during assessment if they are to inform the process of formulation maximally. Statements that are apparently clear to the person expressing them may not be clear to the therapist, or may reveal ambiguities and contradictions, as when angrily saying "I'm not capable as a parent" (when sadness sounds more likely and, superficially, more appropriate), or when commenting wryly that "I felt sorry for myself," without elaborating on what that means. One of the most useful sources of information for formulation comes from the mutual reactions of the patient and therapist to each other—information that is used differently in different types of therapy, and which is understood using different theoretical systems, of varying degrees of sophistication, but which is always relevant.

The processes of assessment and formulation therefore go hand in hand, and inform each other, but they remain different processes. Ideas about how to understand (conceptualize or formulate) what is being said, about its personal meaning and implications for theorized psychological structures and processes, guide questions and observations. When formulating as well as when assessing, the information gathered changes and shapes these ideas as hypotheses are formed, revised, and (theoretically) refined. So, making a formulation is not a one-off activity that defines a fixed state, but the reflection of a dynamic process, and the resulting system of understanding develops and changes over time. This is why the process of formulation should start at the same time as the process of assessment—just as the process of finding ones way around a new place starts with the first encounter with it—and may be on paper rather than in person.

Two points that follow from this line of argument help to determine how a formulation is made. First, if therapists are always formulating as well as assessing, then their questions and statements should be guided by conceptual hypotheses. They should always be able to answer the question "Why did you ask that then?" The answer should not just be phrased in terms of curiosity or information gathering, but should relate to a hypothesis about how to understand the minutiae of the case. The patient's response to the therapist's comment or question is then maximally informative. This may sound unrealistically demanding, as if every sentence the therapist utters should be shaped by the developing formulation—indeed, it is intended as a rule of thumb rather than as a categorical imperative. However, it is less unrealistic than it might seem. The initial question in the therapist's mind could be quite a simple one, for example: Is the withdrawal described by this person associated with feelings of depression and sadness or is it a kind of avoidance motivated by fear? Will attentive listening help this person feel sufficiently comfortable to disclose significant material? Are my questions too specific and intrusive at this stage? Answers to these questions could of course lead to more complex ones: Is this person's reticence a product of experiences that have destroyed trust? Does it reflect a preference for an autonomous style of relating to others? Is it a product of inexperience and lack of practice or opportunity in talking about intimate and personal matters? Is this way of interacting culturally unfamiliar to them?

The second point is that the process of therapy should not be artificially separated into discrete stages of assessment, formulation, and treatment (or intervention). It is not that these processes cannot be distinguished, or that one or other of them may not predominate at a particular time, but that they cannot in practice be wholly separated from each other. Thus, one of the hardest tasks therapists have to learn is how to bear all three of them in mind at once—how to gather information, think about it in theoretical/structural terms, and remain aware of the various ways in which they are likely to exert an influence, so as to enhance the potential for productive change, rather than limit or delay it.

The many sources of information available to therapists when starting to develop a formulation, assuming an adequate process of assessment has been set in motion, are summarized in Table 2. This list includes both direct and indirect sources of information, information from standardized questionnaires, and from initial interventions such as self-monitoring and homework assignments (when these are used). The purpose of this summary, in the context of the preceding discussion, is not to overwhelm therapists with long and exhaustive lists of material to be gathered, items to consider, processes to complete, and so on, but to illustrate that there is an enormously rich source of relevant material potentially available, and the process of formulation can draw on any of it, beginning anywhere. The process of formulation is essentially one of abstraction and it works by relating observable phenomena to hypothetical underlying processes and mechanisms. It is not necessary to observe everything before making a guess at what lies underneath. An (informed) guess may either indicate the need for more assessment or it can short circuit the process. Because formulating is a dynamic process, and depends on the ability of the therapist to retain an open mind, the process can productively start to serve the functions listed in the previous section straight away. Therapists can focus their minds on the process of formulation by asking more formal questions: How can I understand the information I have been given in terms that make theoretical sense? What implications does that understanding have for what to do next? What difficulties will I have, working with this person? What difficulties will they have (working in this way) with me? What use will this person be able to make of treatment? Answers help to determine how to intervene and to predict what will or will not happen as a consequence.

6.01.4.2 Putting the Information Together

Given that a formulation provides connecting links between theory and practice, the precise form that it takes will be partly determined by the theoretical approach of the person making it. Nevertheless, some general points apply, and these are illustrated here using the cognitive-behavioral approach.

First, initial formulations can provide cross-sectional understanding of an aspect of the presenting problem. The most obvious example is probably that of a vicious cycle which summarizes the way in which a particular, readily accessible, symptom pattern is thought to be maintained. It is used here to illustrate the way in which a formulation helps to specify processes, links, and mechanisms. In this case the focus is on certain kinds of links. Other cross-sectional formulations might focus on other patterns, for instance in interpersonal functioning, sequences of behaviors and their consequences; thoughts, feelings, attitudes, and beliefs; dilemmas and traps. In this example (Figure 2), a woman living through a stressful period described feeling tired much of the time and being unable to relax. Asked about what goes through her mind when trying to switch off, she described a stream of worries, most of which were rather vague and hard to specify in detail. The worry disturbed her sleep pattern, which exacerbated the tiredness. A cycle, which symbolizes how one thing leads to another, can easily be illustrated diagrammatically, and it has obvious implications. Breaking the links will help to solve the problem, and this can be done in various ways, such as learning to relax, identifying and dealing with the worries, or taking hypnotic medication. The assumption behind the formulation so far is that the problem will subside if the process that maintains it is interrupted, and the intervention selected could be determined by the preferences, understanding, or skill of either of the parties involved.

However, a formulation essentially relates theory to practice. Applying the cognitive model to this case would suggest that a close relationship between thoughts and feelings is likely to be of central importance. There are at least three ways in which this initial formulation, in its hypothetical and simplified form, can help the cognitive therapist to focus on factors that theoretically are likely to be relevant. It identifies worry as an important cognitive-maintaining factor, it reflects an overall understanding of the problem, suggesting that the symptom pattern is recognizable, understandable, and changeable—attitudes which may differ strikingly from those the patient starts with—and it poses questions about the context of the problem. Nothing has been specified

Table 2 The main sources of information for use in formulation.

Examples of direct information

- Reports of present phenomena: cognitive, affective, behavioral, and physiological
- The context: historical background and development, real life problems
- Reactions, comments, and expectations, about the self, others, therapy, events, etc.
- Interactions within therapy: ability to relate, tenor of relationships
- Observations of body position, movement, facial expression, eye contact, etc.
- The outcome of interventions such as self-monitoring, homework assignments, behavioral experiments, etc.
- Products of questionnaires, tests, standardized interviews, systematic observation, etc.

Examples of indirect information

- Knowledge about diagnosis: *DSM*
- Referral information: summaries, previous treatment, opinions
- Knowledge of cultural norms (of the therapist and of the patient)
- The socioeconomic and political context

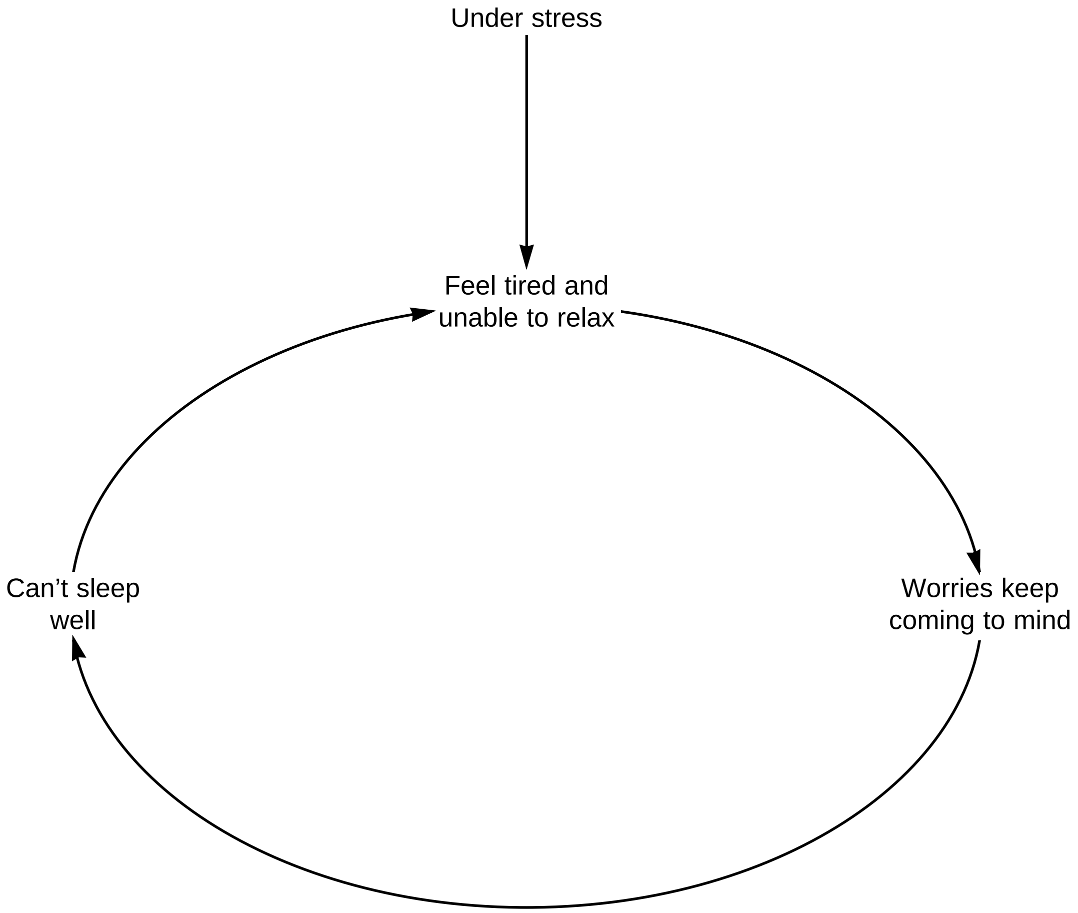

Figure 2 Example of a simple cross-sectional formulation: basis for a more complex formulation.

about why this is, for this person, a stressful period. Theory-driven questions help to develop more hypotheses: What does it mean about her habitual response to stress? How does she construe her present situation? What does her reaction to it mean to her—about herself, about other people, and/or about the world in which she finds herself? So, the initial formulation triggers further inquiry, and starting from a simple cross-sectional map can lead to more sophisticated levels of understanding, and to more complex formulations, as well as being practically useful. The precise way in which this happens will be determined by the theory being used.

Cross-sectional formulations can also provide an outline summary of the way in which complex underlying factors are understood, or of the way in which aspects of a problem are linked. Three statements made by an unemployed, unconfident young man with a wide range of social, interpersonal, and affective problems were used as the starting point for the initial formulation illustrated in Figure 3: "If I always please others they'll never find out about me," "I'll be OK if I stick to doing easy things," and "People will reject you if you don't toe the line." In this diagram, three aspects of his problem are represented in different ways. First there is a rather shapeless "thought bubble" at the top in which hypotheses about underlying cognitive structures, beliefs, attitudes, and rules about himself have been put into words: "I'm incompetent"; "I have to do what others ask"; "I'm thick (stupid) ... " These actual words were his responses to specific (theory-driven) questions, and they illustrate how the process of formulation interacts with that of assessment, and depends on the ability to abstract and to generalize. The broken line is labeled a "protective wall" because it represents the idea that the three statements listed—starting points for a more detailed formulation—reflect behaviors that serve a function. Reacting in these ways protects him from having to confront (the hypothetical) underlying beliefs and attitudes, and prevents others from discovering them, both of which would be painful experiences for him. However, these protective reactions cause problems, not specified here but referred to in the box in Figure 3. This formulation contributed to the process of developing a shared understanding of some complex problems, and it was used to explain how change would probably involve working at all three levels. It also has implications for decisions about general aspects of therapy. For example, it suggests that at times this will be a distressing process that will demand sensitivity and a good sense of timing from the therapist.

Cross-sectional formulations potentially reflect ideas about psychological processes and mechanisms as well as about the relevance and relative importance of different facets of a problem. Longitudinal formulations reflect assumptions about etiology as well. They are used in most kinds of therapy, and are readily illustrated in the case of CBT. The basis for using this theoretical model in clinical practice has been summarized in the form of a template (Table 3) which can be used to illustrate how theoretical understanding can be translated into practice. This shows that, theoretically, experience, both early in life and subsequently, gives rise to a set of beliefs and assumptions about the world, about other people, and about the self.

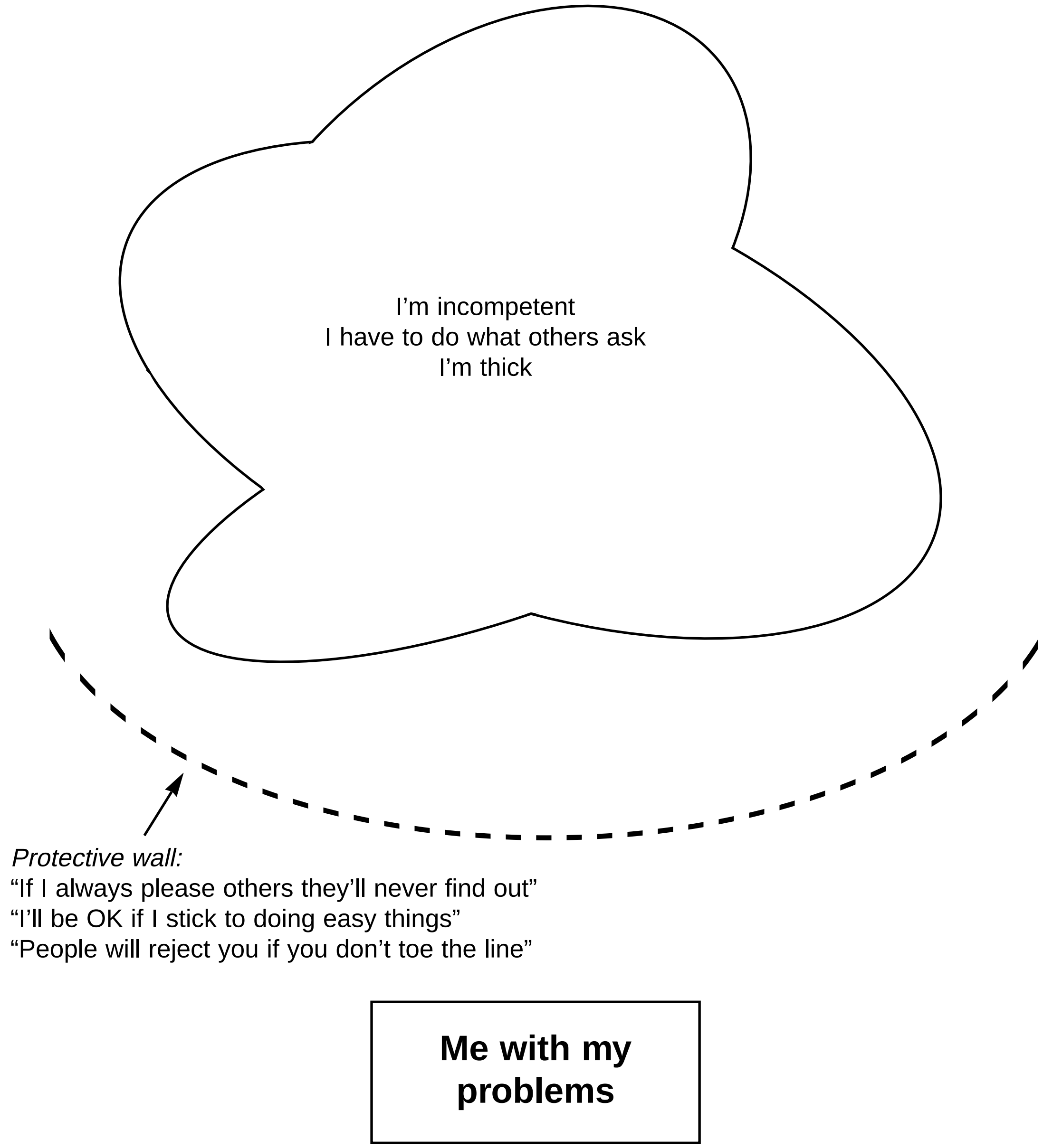
I'm incompetent
I have to do what others ask
I'm thick
Protective wall:
"If I always please others they'll never find out"
"I'll be OK if I stick to doing easy things"
"People will reject you if you don't toe the line"
Me with my problems

Figure 3 Example of a cross-sectional formulation.

These beliefs are seen as a product of the ways in which earlier events have been perceived, understood, and remembered. They can be functional or dysfunctional, actively influential or latent at any particular time, and relatively easy or hard to identify and to recognize. A critical incident (see also below) is an event that fits with a belief—being rejected for someone who believes they are not socially acceptable, or being let down for someone who believes that other people are unreliable or untrustworthy. Critical incidents activate the relevant beliefs and assumptions, and thus produce negative automatic thoughts (NATs). Then a variety of interacting cognitive, affective, behavioral, and physiological reactions follows. At this level the problem is theoretically maintained by cyclical processes of the kind summarized in the cross-sectional vicious cycle described above.

Clearly a template such as this can be used to structure information about a patient, and this will have implications for what the therapist does. For instance, if it appears that dysfunctional beliefs play a small part in the presenting problem, or are well balanced by a set of positive beliefs, the theory (and the formulation derived from it) suggest that the work should focus predominantly on the level of maintaining factors. Another type of implication might reflect the degree of verification available for the theory. For example, psychologists do not yet know which are the most effective ways of changing beliefs (the cognitive frameworks with which people approach the world). One common strategy is therefore to begin working at the level of the NATs and to evaluate the degree of belief change that follows. The processes of change may, or may not, be set in motion by work at this level. If not, then another hypothesis might be that one of the many processes now available for changing beliefs should be adopted as well as or instead of. This example is not meant to explain how to do CBT, but to illustrate how the internal map provided by a theoretical understanding relates to a specific formulation, and how therapists can use such maps as guides even when there is incomplete evidence for the theories upon which they are based. Doing so enables them to explain what they have been doing, and it enables others to decide whether their actions were skilful, appropriate, and so on.

6.01.4.3 Key Factors and Basic Elements

This example also illustrates that when learning how to construct a formulation, it can be helpful to think in terms of key factors. Critical incidents provide a good example of these as they reflect the way in which hypothetical underlying mechanisms are manifested, and link these with observations about present phenomena. Critical incidents are "critical" because they provoke a high degree of affect, often in excess of what might otherwise be expected (an over-reaction, such as becoming enraged if kept waiting for 10 minutes); they are easy to notice and remember, and are of special significance for the person who experiences them. Examining them potentially reveals other elements of the CBT template: underlying

Table 3 Template for a longitudinal formulation using cognitive-behavior therapy.

Experience (early or otherwise)

↓

Beliefs, about the self, the world, and others, which are expressed in categorical statements: I am . . . ; the world is . . . ; others are . . .

↓

Assumptions derived from beliefs, which can be expressed in conditional statements: If I . . . then . . . ; One should . . . otherwise . . .

↓

Critical incidents

↓

Activated beliefs and assumptions

↓

Negative automatic thoughts (NATs)

↓ ↑

Cognitive, behavioral, affective, and physiological reactions beliefs, preferred coping mechanisms, maintenance cycles, and so on. Focusing on critical incidents is thus theoretically helpful when stuck in constructing a cognitive formulation. This is not to say that all cognitive therapists think about them, or base their formulations upon them. An alternative method might involve working from a problem list, weighting the problems for importance, and going on to abstract and understand the connecting themes and links in ways that fit with the theory. The point is that within a particular method of working there are many ways of constructing a formulation, but it can be helpful to keep those factors in mind which play a central part in the theory, or in revealing the manifestations of important theoretical constructs whether these are core beliefs, core interpersonal schemata, or core conflicts. To repeat, there is no single correct method.

Use of the word "core" suggests that formulations may be thought to have certain basic elements, and that unless these are identified the formulation will, in Perry et al.'s words, "lack an integrative coherence." When writing about psychodynamic formulation and about central conflicts, Perry et al. (1987, p. 546) say "The aim is to find a small number of pervasive issues that run through the course of the patient's illness and can be traced back through his or her personal history, and then to explain how the patient's attempts to resolve these central conflicts have been both maladaptive ... and adaptive." The overall intention is clearly closely similar across different therapeutic orientations, as is the general approach: first, apply a particular, theory-driven model; if that does not in practice fit the particular case, explore further using questions and trial and error in the (scientific) search for a formulation that fits better.

This process might be facilitated if there was agreement over which were the basic elements of a formulation and an atheoretical way of linking them together. One way of doing this has been developed by Goldfried and his collaborators. This transtheoretical coding system "was developed as a common language for use in conducting comparative process research across orientations" (Goldfried, 1995, p. 222). It specifies which are the relevant components of functioning (e.g., self-observation, self-evaluation, intention, emotion, and action) and the types of links that can be made between them (vicious cycles, patterns, contradictions). These can be manifested both in intra- and interpersonal contexts, involving other people or not, over a particular time frame. One advantage of this type of formulation, the coding system of therapeutic focus (CSTF), is that it indicates what the problem is, and where to intervene, but (being atheoretical) cannot indicate how to do so. It cannot therefore provide specific implications for treatment, but it does provide a common language, and using this it is potentially easier to find out precisely how theories differ when put into practice.

6.01.4.4 Issue of Completeness

The formulations illustrated so far have been kept simple for the sake of clarity, to emphasize the point made at the beginning about the principle of parsimony, and because they demonstrate the point that it is never too soon to start formulating. They are examples of initial hypotheses. As treatment progresses they would be likely to become more complex and also to take more account of a person's developmental history and the supposed underlying mechanisms.

This raises an important issue for discussion. Many people assume that formulating is a difficult and lengthy process, the aim of which is to encompass, systematize, and explain all relevant factors about a particular case. This view can lead therapists either to bypass the process of formulation and start treatment straight away, or to delay the start of treatment until they have got the picture right. Both of these reactions cause problems: bypassing the process makes it hard to move beyond the stage of trial and error; interventions are selected in the absence of a coherent underlying strategy. This seems to be successful when the patient responds well (as many patients do initially), but it leaves both parties feeling confused and unable to understand what has happened when half a dozen sessions later progress is halted and setbacks are encountered. It is rather like trying to stop a car rattling by cleaning and adjusting those parts of the engine that are most accessible. Delaying the start of treatment is another false economy, for many reasons: the initial momentum provided by a fresh start and a new encounter may be lost; the impact of being listened to, heard, and understood by someone new may be dissipated; and the goodwill, advice, and new ideas derived from interacting with a trained therapist may not be harnessed in a way that is either helpful or informative (or both). So, opportunities to test hypotheses may be lost.

In an ideal world therapists, believing that "at some level it all makes sense", would be able to use their formulations to make sense of the material presented in a particular case. But at present complete formulations, like complete theories, are not possible. A person cannot be summarized in a diagram. But some of their problems and patterns of behaving can be understood in theoretical terms and this understanding can be represented in a way that helps to guide treatment. The complexity and accuracy with which this is done varies according to the stage of treatment.

The emphasis on completeness that is often found in discussions about conceptualization may be a consequence of the historical associations between medical practice, psychiatry, and psychology, and the common use of the word "treatment" to refer to the actions of people trained in those professions when they are trying to help others. The assumption is that it could be dangerous to miss something serious or to apply the wrong treatment. So, a complete understanding is supposedly an essential (or important) prerequisite for deciding how to intervene.

The situation is different in psychotherapy (or psychological therapy), first because the psychological influence of one person on another cannot be withheld (as can a medical treatment), and then applied when ready, in a self-contained package. Various (partially unspecified) factors are always operating, and in psychotherapy the ways in which these function will to some degree be influenced by the theoretical views and assumptions of the therapist (as well as by their personal characteristics). The business of formulating can direct this process, clarify what is intended, and make the way in which theories are being applied accessible. Formulations do not have to be complete to perform this function but the method of working with them does have to be in place. Second, when dealing with psychological matters, the process of formulation is overtly interactive. Patients' comments and reactions contribute to the process; their opinions are relevant, and these may change over the course of treatment.

6.01.4.5 Conceptualizing Processes of Change

Therapists seek understanding of the way in which change takes place as well as of the way in which problems arise and persist, and they may also formulate this understanding in theoretical terms. Conceptualizing the processes of change is thus another way of relating theory to practice, and formulations may be technical, phrased in terms that are derived from the particular theory being used, or metaphorical. Technical formulations might explain how changing reinforcement patterns would change behaviors, how change in one person will prompt the system around them to adjust, or how changing patterns of defensiveness might change opinions of the self. Examples of metaphorical formulations are provided in this section to illustrate how metaphors can encapsulate information about complex processes that may be hard to specify otherwise. Some examples are well known and their use has become quite conventional, such as "a journey of a thousand miles begins with the first step." Others are created in a particular therapy context. For example, a manager of an engineering company, whose habitual rigidity was exacerbated by various (personal and industrial) crises, saw himself as "hanging on for dear life," and being unable to contemplate change. He was asked to think about how to build a building to withstand an earthquake. Thinking about this enabled him to reconceptualize change as a way of developing the combination of flexibility and rigidity needed to provide stability in difficult times. Another relatively simple way of representing and summarizing a complex process of change was spontaneously developed by a woman with longstanding problems involving low self-esteem and lack of confidence. She saw herself as "wobbly" and at risk of falling, as if trying to sit on a two-legged stool. The process of change for her was like "putting down the third leg." In practice this meant many things that contributed to a sense of stability: developing new skills and abilities, thinking about herself in new ways, and making more respectful and open relationships with those around her.

People often use metaphorical language to communicate their experience of distress. Indeed, it might be more accurate to say that it is difficult to describe such experiences without using metaphor—people explain to clinicians how they feel broken, trapped, fenced in, cast adrift, close to the edge, messed up, out of reach, cut off, high, low, and so on. Perhaps the most common methaphors describe life as a journey and ourselves as traveling through different kinds of emotional weather. Patients' understanding (or personal formulation) of the processes involved is also reflected in the words used to describe their experiences: "I've hidden myself away ... built a protective wall around me ... had to harden my shell ... can't see my way out of the tunnel ... waited to be rescued." It is hardly surprising that the processes involved in therapy are similarly described. Someone who came to understand the stultifying and self-destructive effects of overt compliance with the wishes of those around her, despite her own inclinations, and the relationship of this pattern of behavior to the fear and anger for which she was requesting help, said that she felt as if she had spent her whole life trying to grow flowers in her garden and cutting off the buds before they could flower. She saw therapy as a process that would help her to allow the flowers in her garden to bloom. Undoubtedly, the process of developing a shared understanding is a complex one, and the more abstract the material considered, the more difficult this process will be. Although a metaphor is not a formulation, and it may reflect only part of what is involved, using one can help to fulfill some of the purposes of formulation that were described above, and it can do so with a startling degree of economy and emotional sensitivity because it operates at more than one level.

These examples have been chosen because they illustrate a point not so far emphasized about formulation, that it is a way of summarizing meanings, and of negotiating for shared ways of understanding them and communicating about them. When these are complex it can be helpful to use metaphor, and of course this applies generally, not just when formulating processes of change. A formulation provides a source of common language, and when this is available it can then be used to relate a theoretical framework, at a high level of abstraction, to practice, so as to facilitate the process of change.

6.01.5 ACCURACY: HOW TO TELL IF A FORMULATION IS RIGHT

Formulations can never be shown to be right as they are hypotheses not statements of facts. The evidence may support them or it may not, and they should be judged according to probabilities rather than on an absolute scale of rightness. Like other scientific hypotheses, formulations can only be shown, conclusively, to be wrong. Nevertheless, practical guidelines are useful, and a number of attempts have recently been made to evaluate their inter-rater reliability and predictive validity (Barber & Crits-Christoph, 1993; Horowitz & Eells, 1993; Persons, Mooney, & Padesky, 1995).

6.01.5.1 Criteria of Accuracy

A summary of the kinds of practical guidelines that might provide clinicians with criteria of accuracy is given in Table 4. Unfortunately, the fact that a formulation makes good internal sense (provides a plausible narrative for instance) is not a guarantee of its accuracy, which should therefore be tested out in practice. It goes without saying perhaps that a formulation which is simple, clear, and easy to understand, and therefore easy to explain, is more readily testable than one which is overly complex. One which is more specific and low level will have clearer implications than one which is phrased in more general, abstract, and high-level terms. Presenting the formulation to someone else, or putting it onto paper, is therefore a useful and revealing exercise.

6.01.5.2 Questions for Research

It would probably be fair to say that, of the many questions that could be asked, few have been studied and none have been conclusively answered. Persons, Padesky, and Mooney (1996) found only moderately good inter-rater reliability of cognitive-behavioral formulations when tapes of initial therapy sessions were rated by a large group of therapists who had been trained in CBT, and who varied in their level of experience. Surprisingly perhaps, agreement was better with respect to underlying mechanisms than in listing patients' overt problems. Barber and Crits-Christoph (1993) found, when reviewing the psychodynamic literature, that when clinicians based their formulations on preset categories, formulations were more reliable, and in addition the predictions of the psychotherapy process and outcome were better. Both these findings fit with the view that the more clearly specified the activity (as in CBT and interpersonal psychotherapy, or when using clearly defined conceptual categories), the less room there is for wide-ranging, speculative inferences, and the more agreement there is both about particular case formulations and about their utility.

As discussed above, there has been some suggestion that making overall decisions about treatment purely on the basis of a diagnosis may be at least as useful as basing them on an idiographic formulation. However, a diagnosis only enables therapists to make general decisions about which set of interventions to employ; for example, to use exposure *in vivo* to help someone with a simple phobia, or those techniques that will assist in resolving a role dispute in a case of depression treated with interpersonal psychotherapy. In both cases the actual steps used will still depend on the way in which the individual case is formulated (Markowitz & Swartz, 1997). The question as to whether treatment that is based on a formulation is more successful than treatment that is not is more complex than at first appears. Most clinicians bring their theoretical knowledge to bear in the way that they understand, and communicate understanding about, a case. They use covert formulations, which may not be made overtly communicable even though they inform and direct the process of treatment. This happens because, once therapists are

Table 4 Ten tests of a formulation.

1. Does it make theoretical sense?
2. Does it fit with the evidence? (symptoms, problems, reactions to experiences)
3. Does it account for predisposing, precipitating, and perpetuating factors? (both overall and with respect to episodes of difficulty)
4. Do others think it fits? (the patient, supervisors, colleagues)
5. Can it be used to make predictions? (about difficulties, aspects of the therapeutic relationship, etc.)
6. Can you work out how to test these predictions? (to select interventions, to anticipate responses and reactions to therapy)
7. Does the past history fit (with respect to the person's strengths as well as weaknesses)
8. Does treatment based on the formulation progress as would be expected, theoretically?
9. Can it be used to identify future sources of risk or difficulties for this person?
10. Are there important factors that are left unexplained?

thoroughly familiar with the theoretical background to their work, and with the process of map-making, the activity of formulation cannot be wholly suspended. Once able to recognize signs of core beliefs or core conflicts, for example, such theoretically meaningful constructs cannot suddenly be rendered invisible again. Formulation skills may still need sharpening, and there is certainly a need for more and better training (Sperry, Gudeman, Blackwell, & Faulkner, 1992), especially now that clinicians appear increasingly likely to incorporate ideas from theoretical orientations other than their main one into their work (Messer, 1996b). The effects of working with (or without) a formulation will remain hard to evaluate. The more important question, in practical terms, is whether or not a particular way of seeing things is put to good use, successfully to do the things that a formulation is for. The struggle is to find a way of seeing things that helps. Although the assumption that "at some level it all makes sense" still underpins much clinical work, it is not necessary to believe that there is such a thing as a "correct" formulation. As Messer (1996a, p. 136) says, "An alternative outlook is that there is no one version of truth possible because we largely construct our realities, which inevitably leads to multiple perspectives on that reality. Wearing different glasses provides different views of the world."

6.01.6 USING THE FORMULATION: PRACTICAL ISSUES

A formulation does not have to be correct, but it does have to be useful. The purposes of formulation are discussed in Section 6.01.3. Here, three practical factors that influence whether a particular formulation succeeds in fulfilling its purposes are mentioned briefly.

6.01.6.1 The Value of Organizing and Clarifying

Formulating is a way of classifying information, putting it into (conceptual) boxes, and drawing links between them. It organizes information, treatment strategies, and the choice of interventions, and it also clarifies understanding of a case, and therefore the meaning of what is observed. This process has some less obvious advantages as well as the obvious ones. In particular, it helps therapists to see problems and difficulties as understandable, and this influences their attitudes and expectations. For example, hostile or passive–aggressive behaviors frequently create frustrations and difficulties for therapists, especially when they persist despite all their best efforts. Organizing and formulating the information helps therapists to see these as characteristic and predictable difficulties for which they can plan appropriate strategies.

6.01.6.2 Developing an Internal Supervisor

The process of formulation provides therapists with an opportunity to achieve on their own many things that otherwise they would achieve through supervision. It prompts them to reflect about their work with individual cases, and to rethink when progress seems blocked. It helps them to become aware of their own assumptions and beliefs, and to look out for ways in which these may cause problems, such as making it hard for them to notice, understand, or work with particular issues. It helps them to work well with unusual cases or with types of problems that they have not previously encountered. In doing so it helps to build confidence. Formulation is no substitute for supervision but, used well, it complements and extends it—provided that the formulation does not become a fixed way of seeing things that obscures the significance of information that does not fit.

6.01.6.3 Communicating a Formulation

Some obvious principles can be derived from the preceding arguments: the simpler the formulation, the easier it will be to communicate; it should be presented as a hypothesis, not as fact; and initial guesses are worth checking out as they can indicate whether a particular way of seeing things is likely to be productive.

To some degree a formulation is a matter of judgment. It is based on clinical judgment as well as on knowledge and facts. As judgments about people are bound to reflect some of the attitudes and assumptions of the person who makes them, the question arises as to what should be done with those judgments. Who should be told about them? Are there people who should not be told, or circumstances in which they should not be disclosed?

Answers to these questions are partly determined by practitioners' ethical guidelines and procedures for professional accountability. They also depend partly on the theoretical orientation of the therapist. In cognitive analytic therapy, interpersonal psychotherapy, CBT, and in some forms of short-term psychodynamic psychotherapy, therapists make their formulations explicit, and have therefore considered carefully how and when this should be done (Beck, 1995; Beck, Freeman, & associates, 1990; Butler & Booth, 1991; Markowitz & Swartz, 1997; Ryle, 1995). The method used is immensely variable, using imagery, metaphor, diagram, or verbal explanation, presented in person or in a letter. There is room here for creative thinking, and sensitive adaptation of communication skills, though it may help to specify some general principles.

Being on the receiving end of a formulation can feel like being weighed up, evaluated, or judged—like being "seen through" or "rumbled" rather than understood. This is less likely if the formulation is presented questioningly and collaboratively, at a time when therapists are clear that patients are able honestly to give feedback, and while thinking about how to facilitate the process of feeding back reactions in a way that is not just superficial or polite. It is important to focus on strengths as well as weaknesses, and to draw out implications for change, otherwise patients with chronic problems may conclude that "this is the way that they are," and become hopeless about change. The language used should be simple and jargon free. It may help to give a small amount of information at a time and to be ready to repeat explanations, or introduce technical terms, as necessary. Therapists often underestimate how much patients can themselves contribute to the process of formulation, for instance, by elaborating details, filling in missing links, or providing contradictory information that shows how the formulation can usefully be adjusted.

Formulation thus goes hand-in-hand with reformulation, and it is this, as Rosenbaum (1996) points out, that stops it becoming a way of "fitting something to a known formula."

6.01.7 CONCLUDING DISCUSSION

Formulations reflect the way in which therapists make sense of someone else's predicament. They reflect the assumptions brought to bear when thinking about it, the theories learned, and the meaning made of it. However, making sense is not the only thing that they do. All therapists are aware that sometimes (albeit rarely) providing a formulation can be sufficient to bring about change. Such cases show that formulations do more than supply understanding—they enable someone to see things differently, to reformulate, or to find a new meaning. A business executive whose whole career was threatened by an episode of severe stress and anxiety was suddenly able to see himself as engaged in a genuine struggle. It was then legitimate, in his view, to experience reactions indicative of both fight and flight. His symptoms became acceptable, diminished immediately, and he remained well over the following six months. Of course this could be understood in many ways: as a healthy consequence of a reformulation, as a miracle cure, or as a flight into health. So therapists are also in a predicament. Most of the time only some of the facts are available to them, whether these are about someone's past life, their internal experience or their present relationships, and the facts that are available are consistent with a wide range of plausible interpretations. Different mechanisms can be inferred from the same event, as in the example above, or from the same overt problems—the bather's hand movements could signify waving or drowning. Equally, the same mechanisms could be inferred from different problems—a fear of abandonment could underlie both hostile and dependent behavior. The skills of functional analysis may help to advance the process of formulation here. To end where we began, Frank (1986, p. 343) said that "the best hope of bringing conceptual order into the field of psychotherapy may lie in thinking of all psychotherapeutic enterprises as lying in the realm of meanings" ... thinking, feeling, and behavior are ... "responses to the meanings of events as much as to the events themselves." Our assumptions and knowledge about the ways in which these meanings are stored, represented, and recalled, and about the degree to which they can be brought into awareness, will therefore greatly influence the meaning we give to our formulations and the uses we make of them.

Therapy can be understood in many ways—as managing anticipated transferences, countertransferences, and resistances; as seeking new perspectives and using these to restructure a belief system; as a process of constantly meeting and adjusting to what is happening each moment; or as a way of influencing the contingencies that relate behaviors to their antecedents and consequences. In all of them, the process of formulation serves similar functions. It is useful because it helps to determine what we, as therapists, do and enables us to understand and to explain that better.

6.01.8 REFERENCES

Barber, J. P., & Crits-Christoph, P. (1993). Advances in measures of psychodynamic formulations. *Journal of Consulting and Clinical Psychology, 61,* 574–585.

Beck, A. T. (1991). Workshop on cognitive therapy of personality disorders. Brighton, UK: Royal College of Psychiatrists.

Beck, A. T., Freeman, A., & associates (1990). *Cognitive therapy of personality disorders* (chap. 4). New York: Guilford Press.

Beck, J. S. (1995). *Cognitive therapy: basics and beyond.* New York: Guilford Press.

Beutler, L. E., Williams, R. E., Wakefield, P. J., & Entwistle, S. R. (1995). Bridging scientist and practitioner perspectives in clinical psychology. *American Psychologist, 50,* 984–994.

Blackburn, I-M., & Twaddle, V. (1996). *Cognitive therapy in action: A practitioner's casebook.* London: Souvenir Press.

Butler, G., & Booth, R. (1991). Developing psychological treatments for generalized anxiety disorder. In R. M. Rapee & D. H. Barlow (Eds.), *Chronic anxiety and generalized anxiety disorder* (pp. 187–209). New York: Guilford Press.

Butler, G., Fennell, M., Robson, P., & Gelder, M. (1991). A comparison of behavior therapy and cognitive behavior therapy in the treatment of generalised anxiety disorder. *Journal of Consulting and Clinical Psychology 59*, 167–175.

Butler, G., & Low, J. (1994). Brief psychotherapy. In M. Pokorny & P. Clarkson (Eds.), *A handbook of psychotherapy* (pp. 208–224). London: Routledge.

Clark, D. M. (1988). A cognitive model of panic attacks. In S. Rachman & J. D. Maser (Eds.), *Panic: Psychological perspectives* (pp. 71–90). Hillsdale, NJ: Erlbaum.

Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg, M. K. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment and treatment* (pp. 69–93). New York: Guilford Press.

Dattilio, F. M. (1994). SAEB: A method of conceptualisation in the treatment of panic attacks. *Cognitive and Behavioral Practice, 1,* 179–191.

Eells, T. D. (1996). Commentary on three case formulations of Jim. *Journal of Psychotherapy Integration, 6,* 119–126.

Eells, T. D. (1997). *Handbook of psychotherapy case formulation.* New York: Guilford Press.

Frank, J. (1986). Psychotherapy—the transformation of meanings: Discussion paper. *Journal of the Royal Society of Medicine, 79,* 341–346.

Freeman, A. (1992). The development of treatment conceptualisations in cognitive therapy. In A. Freeman & F. Dattilio (Eds.), *Comprehensive caseboook of cognitive therapy* (pp. 13–23). New York: Plenum.

Goldfried, M. R. (1995). Toward a common language for case formulation. *Journal of Psychotherapy Integration, 5,* 221–244.

Goldman, R., & Greenberg, L. (1997). Case formulation in process–experiential therapy. In T. Eells (Ed.), *Handbook of psychotherapy case formulation.* New York: Guilford Press.

Hayes, S. C., & Follette, W. C. (1992). Can functional analysis provide a substitute for syndromal classification? *Behavioral Assessment, 14,* 345–365.

Hayes, S. C., Follette, V. M., Risley, T., Dawes, R. D., & Grady, K. (1995). *Scientific standards of psychological practice.* Reno, NV: Context Press.

Haynes, S. N., Uchigakiuchi, P., Meyer, K., Orimoto, L., & Blaine, D. (1993). Functional analytic causal models and the design of treatment programs: Concepts and clinical applications with childhood behavior problems. *European Journal of Psychological Assessment, 9,* 189–205.

Horowitz, M. J. (1989). Relationship schema formulation: Role-relationship models and intrapsychic conflict. *Psychiatry, 5,* 260–274.

Horowitz, M. J., & Eells, T. D. (1993). Case formulations using role-relationship model configurations: A reliablity study. *Psychotherapy Research, 3,* 57–68.

Horowitz, M. J., Eells, T., Singer, J., & Salovey P. (1995). Role-relationship models for case formulation. *Archives of General Psychiatry, 52,* 625–632.

Klerman, G. L., Weissman, M. M., Rounsaville, B. J., & Chevron, E. S. (1984). *Interpersonal psychotherapy of depression.* New York: Basic Books.

Luborsky, L., & Crits-Christoph, P. (1990). *Understanding transference: The CCRT method.* New York: Basic Books.

Markowitz, J. C., & Swartz, H. A. (1997). Case formulation in interpersonal psychotherapy of depression. In T. Eells (Ed.), *Handbook of psychotherapy case formulation.* New York: Guilford Press.

Messer, S. B. (1996a). Concluding comments: Special section: Case formulation. *Journal of Psychotherapy Integration, 6,* 135–137.

Messer, S. B. (1996b). Introduction to special section: Case formulation. *Journal of Psychotherapy Integration, 6,* 81–83.

Minuchin, S. (1974). *Families and family therapy.* Cambridge, MA: Harvard University Press.

Nicholson, S. (1995). The narrative dance—a practice map for White's therapy. *Australian and New Zealand Journal of Family Therapy, 16,* 23–28.

Norcross, C. A., Alford, B. A., & DeMichele, J. T. (1992). The future of psychotherapy: Delphi data and concluding observations. *Psychotherapy, 29,* 150–158.

Padesky, C. A., & Mooney, K. A. (1990). Presenting the cognitive model to clients. *International Cognitive Therapy Newsletter, 6,* 13–14.

Perry, S., Cooper, A. M., & Michels, R. (1987). The psychodynamic formulation: Its purpose, structure and clinical application. *American Journal of Psychiatry, 144,* 543–550.

Persons, J. B. (1989). *Cognitive therapy in practice: A case formulation approach*. New York: Norton.

Persons, J. B. (1993). Case conceptualisation in cognitive behavior therapy. In K. T. Kuelwein & H. Rosen (Eds.), *Cognitive therapy in action*. San Francisco: Jossey-Bass.

Persons, J. B., Mooney, K. A., & Padesky, C. A. (1995). Interrater reliability of cognitive-behavioral case formulation. *Cognitive Therapy and Research, 19,* 21–34.

Rosenbaum, R. (1996). Form, formlessness and formulation. *Journal of Psychotherapy Integration, 6,* 107–118.

Ryle, A. (1978). A common language for the psychotherapies. *British Journal of Psychotherapy, 132,* 585–594.

Ryle, A. (1990). *Cognitive-analytic therapy: Active participation in change*. Chichester, UK: Wiley.

Ryle, A. (Ed.) (1995). *Cognitive analytic therapy: Developments in theory and practice*. Chichester, UK: Wiley.

Salkovskis, P. M. (1996). The cognitive approach to anxiety: threat beliefs, safety-seeking behavior, and the special case of health anxiety and obsessions. In P. M. Salkovskis (Ed.), *Frontiers of cognitive therapy*. New York: Guilford Press.

Schulte, D. (1997). Behavioural analysis: Does it matter? *Behavioural and Cognitive Psychotherapy, 25,* 231–249.

Schulte, D., Künzel, R., Pepping, G. & Schulte-Bahrenberg, T. (1992). Tailor made versus standardised therapy of phobic patients. *Advances in Behaviour Research and Therapy, 14,* 67–92.

Seligman, M. E. P. (1995). The effectiveness of psychotherapy. *American Psychologist, 50,* 965–974.

Silbertschatz, G., Fretter, P. B., & Curtis, J. T. (1986). How do interpretations influence the process of psychotherapy? *Journal of Consulting and Clinical Psychology, 54,* 646–652.

Sperry, L., Gudeman, J. E., Blackwell, B., & Faulkner, L. R. (1992). *Psychiatric case formulations*. Washington, DC: American Psychiatric Press.

Stricker, G., & Trierweiler, S. J. (1995). The local scientist; A bridge between science and practice. *American Psychologist, 50,* 995–1002.

Turkat, I. D., & Maisto, S. A. (1985). Personality disorders: Application of the experimental method to the formulation and modification of personality disorders. In D. H. Barlow (Ed.), *Clinical handbook of psychosocial disorders: A step by step treatment manual.* New York: Guilford Press.

White, M. (1989). The externalizing of the problem and the re-authoring of lives and relationships. Adelaide, Australia: *Dulwich Centre Newsletter, Summer,* 5–28.

Wilson, G. T. (1996). Manual-based treatments: The clinical application of research findings. *Behaviour Research and Therapy, 34,* 295–314.

Wilson, G. T. (1997). Treatment manuals in clinical practice. *Behaviour Research and Therapy, 35,* 205–210.

Wolpe, J., & Turkat, I. D. (1985). Behavioral formulation of clinical cases. In I. D. Turkat (Ed.), *Behavioral case formulation* (pp. 5–36). New York: Plenum.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.02
Behavioral Approaches

JÜRGEN MARGRAF
Technische Universität Dresden, Germany

6.02.1 INTRODUCTION 26
  6.02.1.1 One Behavior Therapy or Many Behavior Therapies? 26
6.02.2 WHAT IS MODERN BEHAVIOR THERAPY? 27
  6.02.2.1 Definition 27
  6.02.2.2 Basic Principles 27
    6.02.2.2.1 Principle 1: behavior therapy is based upon empirical psychology 27
    6.02.2.2.2 Principle 2: behavior is problem-oriented 27
    6.02.2.2.3 Principle 3: behavior therapy addresses predisposing, triggering, and maintaining problem conditions 28
    6.02.2.2.4 Principle 4: behavior therapy is goal-oriented 28
    6.02.2.2.5 Principle 5: behavior therapy is action-oriented 28
    6.02.2.2.6 Principle 6: behavior therapy is not limited to the therapeutic setting 28
    6.02.2.2.7 Principle 7: behavior therapy is transparent 28
    6.02.2.2.8 Principle 8: behavior therapy helps the patients to help themselves 28
    6.02.2.2.9 Principle 9: behavior therapy strives for continuous development 28
  6.02.2.3 Hollywood or Realistic Expectations: Treatment Goals in Behavior Therapy 28
  6.02.2.4 Underlying Methodology 29
    6.02.2.4.1 Different types of behaviorism 29
    6.02.2.4.2 Guideline 1: the search for lawfulness 29
    6.02.2.4.3 Guideline 2: observability 29
    6.02.2.4.4 Guideline 3: operationalization 29
    6.02.2.4.5 Guideline 4: empirical testability 30
    6.02.2.4.6 Guideline 5: experimentation 30
  6.02.2.5 The Etiological Approach of Behavior Therapy 30
  6.02.2.6 Typical Therapeutic Methods 30
  6.02.2.7 Indications 31
6.02.3 HISTORICAL DEVELOPMENT OF BEHAVIOR THERAPY 32
  6.02.3.1 Situation at Onset and Precursors 32
  6.02.3.2 The "Founding Period" 33
  6.02.3.3 The Influence of Operant Approaches 35
  6.02.3.4 Consolidation and Broadening of Scope 35
  6.02.3.5 Integration of Behavioral and Cognitive Approaches 37
  6.02.3.6 Continuous Development and the Future 38
6.02.4 EMPIRICAL STATUS OF BEHAVIOR THERAPY 39
  6.02.4.1 Problems in Empirically Approaching Efficacy 39
  6.02.4.2 Results of Outcome Research 40
  6.02.4.3 Consequences of Research Findings 43
6.02.5 PROBLEMS AND CRITICISMS OF BEHAVIOR THERAPY 45
  6.02.5.1 Misunderstandings About Behavior Therapy 45
  6.02.5.2 Scientifically Well Established, Practically Neglected? 46
  6.02.5.3 Current Criticisms of Behavior Therapy 46
6.02.6 CONCLUSIONS 47
6.02.7 REFERENCES 47

6.02.1 INTRODUCTION

More than most other psychotherapeutic approaches, behavior therapy has been the subject of misunderstandings. While the self-image of behavior therapists and research results consistently yield the positive image of a pragmatic, problem-solving, empathic, and frequently successful approach, the outside perception is often quite negative. The scientifically oriented language of behavioral publications seems to have led some observers to the conclusion that behavior therapy is a highly technical endeavor that neglects human relationship factors. Indeed, surveys of the acceptance of behavior therapy (Eschenröder, 1994; Heekerenz, 1991; Lutz, Bezold, Bloem, Dietrich, & Wittmann, 1992; Woolfolk, Woolfolk, & Wilson, 1977) frequently show negative judgments which often are more pronounced in persons that have little information on the approach. A remarkable example on attitudes toward behavioral approaches is given below.

"A rose by any other name . . . : Labeling bias and attitudes toward behavior modification" Using this title, Woolfolk et al. (1977) published two studies, in which beginning and advanced students in educational sciences had been shown a film excerpt of a teacher who applied reinforcement methods. Half of the subjects had been told that the film would show the application of "humanistic methods." For the other half, the method was labeled "behavior modification." This simple labeling massively influenced the ratings of the teacher and the session. With the "humanistic" label, subjects in both studies rated the teacher significantly more positive, competent, flexible, and personally attractive. Moreover, they expected clearly better academic results and emotional growth from the "humanistic" method. The authors attribute their results to the negative effects of a frequently technical and mechanistic presentation of behavioral methods.

In addition to external misunderstandings, there are also "self-misunderstandings" and discrepancies that result from the broad and frequently stormy development of the behavioral approach. Today, there are many different types and opinions about behavior therapy. For example, the classical approach of Joseph Wolpe is as different from modern cognitive-behavior therapy as is traditional progressive muscle relaxation from Öst's applied relaxation or the early operant treatment of depression from Beck's cognitive-behavioral approach. The reasons for the misunderstandings on behavior therapy have been sought in the polarizing form of its early self-presentation, in a lack of information of the public, in feelings of threat in the face of superior proofs of efficacy, in hostility toward rational science, etc. But regardless of their origins, the misunderstandings clearly show how important an explicit discussion of the nature of behavior therapy is.

In the "founding phase" there was a relative consensus on the definition of the term "behavior therapy." It was generally agreed that this was the clinical application of the learning principles that had been established by psychological research (cf. Eysenck, 1959). However, behavior therapy was even in its beginnings a heterogeneous movement of considerable broadness. Feedback from growing clinical practice and vivid research quickly softened the classical learning theory understanding of behavior therapy. Together with the progress in the underlying psychological science, its claim to be theoretically founded and empirically tested generated an ongoing discussion. It is therefore hardly surprising that a large number of definitions has been proposed.

6.02.1.1 One Behavior Therapy or Many Behavior Therapies?

Definitions of behavior therapy typically have varied with respect to the broadness of theoretical orientation and underlying methodology. Wolpe (1976) saw behavior therapy solely as methods that have been derived from experimentally established principles and paradigms of learning. Similarly, Eysenck (1959) understood behavior therapy as an attempt to change human behaviors and emotions based upon the laws of modern learning theory. According to Agras, Kazdin, and Wilson (1979), behavior therapy included already in the 1970s behavioral and cognitive approaches. Even broader, Hollandsworth (1986) saw behavior therapy generally as the application of scientific methods to clinical problems. In the same vein, Yates' (1970) influential definition emphasized that behavior therapy utilizes all the scientific knowledge accumulated in psychology and its neighboring disciplines. Among other important definitions were those by the Association for the Advancement of Behavior Therapy (see Franks & Wilson, 1975) and by Rachman (1988).

The considerable variability of these definitions makes it even more important to ask what constitutes modern behavior therapy. An "eternal" answer to this question is not possible. The mere attempt to answer this question "once and for all" would counterproductively institutionalize the status quo and impair future developments. The continuous evolution of behavior therapy implies that the question about its nature has to be discussed at more or less regular intervals. The next section is therefore devoted to a characterization of modern behavior therapy at the end of its first half century of existence. This is followed by a brief sketch of its historical development in order to better understand its present position. Section 6.02.4 is devoted to the important question of the empirical status of behavior therapy and the practical consequences that this should imply. The chapter ends with a discussion of some of the problems and criticisms of behavior therapy in today's practice and research.

6.02.2 WHAT IS MODERN BEHAVIOR THERAPY?

A mere definition is not sufficient to adequately describe the nature of behavior therapy (Margraf, 1996; Margraf & Lieb, 1994). Any meaningful characterization that is open to the future needs to discuss the basic principles, methodological position, and etiological approach that underly the behavioral approach. Moreover, the characterization could easily become a listing of abstract statements if it does not include typical treatment methods and their indications.

6.02.2.1 Definition

Behavior therapy is a genuine psychological treatment approach that includes a large variety of specific techniques and interventions. In clinical applications, these different methods are utilized either alone or in combination depending on the nature of the problem to be treated. Behavior therapy can therefore not be understood as a single, circumscribed treatment method that relies on one single theoretical model. On the contrary, its theoretical background is composed of a multitude of general and disorder-specific etiological theories and psychological models of change. The common link is the orientation toward empirical psychology. A future-oriented characterization cannot limit itself to a list of present methods. The definition of behavior therapy therefore has to fulfill the following requirements:

(i) the theoretical and methodological breadth of the behavioral approach needs to be included;

(ii) in spite of broad borders the specific aspects of behavior therapy have to be named explicitely; and

(iii) openess for future developments.

This has led me to the proposal to consider the behavioral approach as a broad basic orientation rather than as a single therapy school or group of treatment methods (Margraf, 1996; Margraf & Lieb, 1994). Modern behavior therapy can then be defined as follows: Behavior therapy is a broad psychotherapeutic orientation that is based upon empirical psychology. It includes disorder-specific and general treatment methods that aim at a systematic improvement of target problems on the basis of as much as possible tested knowledge of disorders and psychological principles of change. Interventions have concrete and operationalized goals on the different levels of behavior and experience. They are derived from the diagnosis of disorders and individual problem analysis and target the predisposing, triggering and maintaining conditions of the problem. Behavior therapy is continuously evolving and explicitly asserts to test its statements empirically. The levels of behavior and experience can be conceptualized in different ways. In behavior therapy, they are typically defined according to Lang's three-systems model (1971). This includes a behavioral, a subjective, and a physiological level of response. Although this model can be criticized in some respects (e.g., Fahrenberg, 1987), a multimodal approach has become the standard in most of behavior therapy research (Seidenstücker & Baumann, 1987).

6.02.2.2 Basic Principles

More important than the abstract definition is a concrete description of the basic principles that underly all of behavior therapy. These can be conceptualized as follows (Margraf, 1996).

6.02.2.2.1 Principle 1: behavior therapy is based upon empirical psychology

Empirical psychology is the scientific foundation of the behavioral approach. Behavior therapy therefore endeavors to operationalize its theoretical concepts and therapeutic methods and to test them empirically. Testing should be comprehensive using objective, reliable, and valid measurements. In addition to psychological knowledge on change principles and methods, the findings of nonpsychological neighbor disciplines such as biology or medicine are taken into account.

6.02.2.2.2 Principle 2: behavior is problem-oriented

Treatment as a rule aims at present problems. Therapeutic proceedings are as much as possible tailored to the respective disorders and individual patients. Thus, different disorders are typically treated in individualized form with different methods that are based upon empirical knowledge of the disorder. Beyond the solution of the actual problem treatment typically strives for a general increase in problem-solving capacities. This can be achieved indirectly by making the therapeutic interventions transparent or by promoting new experiences and directly by problem-solving trainings.

6.02.2.2.3 Principle 3: behavior therapy addresses predisposing, triggering, and maintaining problem conditions

Behavior therapy distinguishes between predisposing, triggering, and maintaining factors. Interventions target those conditions whose alterations are regarded as necessary for a durable solution of the problem. Often these are the maintaining conditions, because they are especially important for future well-being. With respect to predisposing and triggering conditions, most often their present effects are at the center of attention, because these types of problem conditions typically cannot be changed *post hoc*.

6.02.2.2.4 Principle 4: behavior therapy is goal-oriented

Identification of the problem and the joint definition of treatment goals are integrative parts of behavior therapy. The problem is the target of the treatment, its solution means that the goal of treatment is attained and the intervention can be terminated. Ideally, the explicit agreement on treatment goals should prevent the pursuit of different goals by therapist and patient or the subsistence of unrealistic expectations.

6.02.2.2.5 Principle 5: behavior therapy is action-oriented

The success of most behavioral treatments presupposes an active participation by the patient. Mere insight is not a sufficient condition for the alteration of fixed problems. Behavior therapy therefore does not limit itself to the discussion and reflexion of problems, but motivates the patient to actively try new behaviors, experiences, and problem-solving strategies.

6.02.2.2.6 Principle 6: behavior therapy is not limited to the therapeutic setting

Behavior therapy strives for a generalization of therapeutic effects on everyday life. The therapeutic setting and a good therapeutic relationship offer the framework for learning and testing new behaviors and experiences. They do not, however, ensure transfer into the patient's individual environment. For this, the patient has to practice newly acquired strategies between sessions. Although behavior therapists frequently accompany their patients for exercises outside of their office, the ultimate goal is always to master problems without therapeutic assistance.

6.02.2.2.7 Principle 7: behavior therapy is transparent

Behavior therapy wants its patients to be informed and active. A plausible model of the disorder, an explicit treatment rationale, and the intelligible explanation of all aspects of the therapeutic interventions are parts of behavior therapy that fulfill the legitimate need of the patients for an understanding of their condition. They elevate treatment acceptance and help to prevent relapse. Transparence thus increases compliance, comprehension of the treatment process, and indirectly problem-solving capacity. In this way, the acquired skills are better available for application to future problems without requiring renewed therapeutic assistance.

6.02.2.2.8 Principle 8: behavior therapy helps the patients to help themselves

Beyond the increase in general problem-solving capacity and the transparent derivation of therapeutic actions from an explanatory model of the disorder, the treatment aims at giving the patient skills for the independent analysis and mastery of future problems. Behavior therapy thus increases the patient's self-help potential and prevents relapses and the development of new problems.

6.02.2.2.9 Principle 9: behavior therapy strives for continuous development

Behavior therapy's orientation on empirical psychology leads to a permanent process of evaluation and further differentiation of its theoretical concepts and practical procedures. Behavior therapy therefore continuously evolves.

6.02.2.3 Hollywood or Realistic Expectations: Treatment Goals in Behavior Therapy

What claims should or may psychotherapy make? Some patients and therapists pursue a "hollywood perspective," in which the end of treatment should be like the happy end of a movie. After successful cure the patient disappears from the therapist's office into his life like the victorious cowboy from the movie screen. Although it has become an acknowledged triviality that for instance love stories end regularly when the relationships and thus new challenges begin, the image of permanent happiness shows an astonishing persistence in the "therapy market." Yet, far-reaching explicit or implicit promises of entire restructuring of personalities, complete freedom from problems, "imploding" symptoms, everlasting happiness, and perfectly painless mastery of life are not only unrealistic but also harmful. Disappointed hopes are especially embittering. Compared to the hollywood standard, one's own achievements and experiences may look like failures. They pursuit of chimeras wastes energy and diverts from realistic coping. The more one believes in "healing promises," the more dependent one gets. Psychotherapy cannot mean lifelong guidance. Realistic treatment goals therefore have to be coping and help to self-help. Even in complex problem constellations, therapy can at best attempt to teach new ways of coping and to identify cardinal points that can be used to break old habits. Psychotherapy can teach swimming, but the swimming has to be done by oneself.

6.02.2.4 Underlying Methodology

Behavior therapy sees itself as an applied science whose understanding of science is strongly influenced by methodological questions. Being that part of logics that deals with the logics of research and methods in general, methodology is a central constituent of the theory of science. It makes statements about questions such as "what is a hypothesis" or "what is the goal of scientific research." The methodology of behavior therapy is typically called "methodological behaviorism." This may not be confounded with other types of behaviorism.

6.02.2.4.1 Different types of behaviorism

The term behaviorism was from its early beginnings part of the struggle of paradigms in psychology. Coined by Watson in order to push his view of scientific psychology, the term was later used primarily by the opponents of Watson, Skinner, and their followers. The intense debate frequently overlooked that there is not a single type of behaviorism; instead there are several competing positions. The most common classification differentiates three types of behaviorism.

(i) *Methaphysical behaviorism* denies the existence of consciousness or mental events. The object of psychological science is exclusively seen in observable (overt) behavior (prominent advocate: Watson).

(ii) *Radical behaviorism* (also called analytical behaviorism) is a version of radical materialism which asserts that the world consists only of one matter. Mental phenomena are regarded as mere linguistic illusions that can therefore not be the object of scientific psychology (prominent advocate: Skinner).

(iii) Opposed to these two types is *methodological behaviorism* that defines itself not through statements about the existence of mental phenomena but through methodological guidelines that distinguish scientific from nonscientific procedures. Even though many representatives do not use the term, methodological behaviorism represents the majority in today's empirical psychology, including, for instance, cognitive psychology.

The basic guidelines of methodological behaviorism as it is applied today in clinical psychology and psychotherapy can be summarized as follows (Margraf, 1996; Reinecker, 1994; Westmeyer, 1984).

6.02.2.4.2 Guideline 1: the search for lawfulness

The goal of scientific work is to find lawful relationships that permit description and explanation of the subject of investigation. "Laws" do not have to be deterministic, probabilistic statements are also recognized. As a rule, different classes of causes are distinguished, although functional relationships traditionally have attracted most attention (see Section 6.02.2.5).

6.02.2.4.3 Guideline 2: observability

Only observable events or phenomena that are regularly related to observable signs can be subject of scientific analyses. This does not imply the restriction to observable motor behavior as the exclusive subject of psychology. Today behavior and experience are the widely recognized subjects of the discipline. Interestingly, even Skinner did not refuse introspection as a method if it fulfilled the above requirement.

6.02.2.4.4 Guideline 3: operationalization

Assessment of the subject of research has to be guided by explicit measurement instructions. Theoretical constructs have to be operationalized, that is, it must be stated in what way they are represented in variables that can be assessed empirically.

6.02.2.4.5 Guideline 4: empirical testability

Hypotheses principally have to be accessible to empirical testing, and they must be sensitive to experience. Immunizing strategies that attempt to make theoretical statements irrefutable are unacceptable because they prevent possible progress in scientific knowledge. Testing of hypotheses can operate through confirmation or refutation. Under the influence of Popper the possibility to falsify general hypotheses ("for all X is true ... ," e.g. "all mental disorders derive from learning") has gained great importance. Also relevant is the verification of hypotheses of existence ("There are some Y, for whom is true ... ," e.g., "some phobias are acquired by classical conditioning").

6.02.2.4.6 Guideline 5: experimentation

The best method to test statements is offered by controlled experiments (which do not necessarily have to be conducted in the laboratory). For ethical and pragmatic reasons, the experimental method has clear limits in clinical research. Often important variables cannot be varied arbitrarily as would be implied in a true experiment. Thus, it is ethically unacceptable to induce mental disorders for experimental reasons. At the most, weak experimental analogues for pathological states may be induced transitorily (e.g., hallucinations, sensorical deprivation, anxiety states, false feedback of performance). However, the question for the ethically acceptable limit has to be answered for each individual case. The opposite approach to reduce pathological states or to treat mental disorders also bears ethical problems when conducted in an experimental frame. For instance, randomization of patients to treatments presupposes informed consent. Therefore, research often has to rely on quasi-experimental designs.

6.02.2.5 The Etiological Approach of Behavior Therapy

The time of the "great" monistic theories to explain all mental disorders by one cause or constellation of causes is over. It has become obvious that such complex phenomena cannot be explained by simplistic or reductionistic "solutions." Today, slogans such as "biopsychosocial approach" or "vulnerability–stress model" dominate the debate. These, however, can be criticized for being overly unspecific or lacking concrete content. The behavioral approach attempts to identify specific causal constellations for individual disorders. Mental disorders are regarded as clinically significant behavioral or psychological syndromes or patterns that are accompanied by distress, disability, or reduced functioning on the behavioral, subjective, or social level.

It is important to distinguish between different classes of "causal" factors and to investigate their relevance to therapeutic change. The basic approach of behavior therapy differentiates between three major classes of etiological factors:

(i) *Predisposing factors:* These are also called vulnerability factors or diatheses. Pre-existing genetic, somatic, psychological, or social characteristics make the appearance of a disorder or problem possible or more probable.

(ii) *Triggering factors:* Psychological, somatic, or social conditions (e.g., experiences, burdens, demands, events, "stress") elicit the first occurrence of a disorder or problem under the possible influence of an individual vulnerability.

(iii) *Maintaining factors:* False responses (by the subject or its environment) or lasting demands prevent the quick reduction of the complaints and make the problem chronic.

This "three-factor model" (Margraf, 1996) is not meant to be an all-inclusive explanation of mental disorders. Instead, it is seen as a heuristic to help etiological research, the judgment of possible starting points for therapeutic change, and the formulation of individual models of pathogenesis.

The three classes of causal factors can coincide or differ completely, they can be more or less accessible to change, etc. For instance, predispositions frequently cannot be changed or problem-evoking traumata typically cannot be undone. The modification of maintaining factors in contrast is often the central mechanism of change for future well-being. Behavior therapy therefore addresses exactly this point (e.g., reduction of avoidance behavior in phobias, training of social skills in schizophrenic or depressed patients).

6.02.2.6 Typical Therapeutic Methods

One reason to conceptualize behavior therapy as a basic therapeutic orientation rather than as a single therapeutic school is the large number of often remarkably different methods that characterize the behavioral approach. Three classes of methods have to be distinguished.

(i) Basic skills that apply to the therapeutic dialogue, the therapeutic relationship or treatment motivation. Although behavior therapists in clinical practice typically display high levels of relationship skills, they neglected to present these basic aspects in their early writings. This may have contributed to the overly technical image of behavior therapy. Today it is clear that a good therapeutic relationship is typically a necessary although not sufficient condition for therapeutic success.

(ii) General therapeutic methods that do address specific disorders. These are methods that every behavior therapist has to know and that need to be incorporated flexibly into individual treatment plans. A list of these methods is given in Table 1 (left column). Several of these methods have been the backbone of behavior therapy in its early beginnings (e.g., systematic desensitization, self-assertiveness training, relaxation training).

(iii) Disorder-specific treatment programs that are tailored as much as possible to the specific characteristics of the different disorders. Such programs came later in the development of behavior therapy. Today, they have been developed and tested for most important mental disorders. Ideally, they rely upon psychological knowledge of the disorders and general change principles. Most widely disseminated are programs for various anxiety disorders, depression, schizophrenia relapse prevention, eating disorders, sexual dysfunctions or marital problems in adults, or enuresis, hyperactivity, and aggressiveness in children. A list of exemplary programs is also given in Table 1 (right column).

Most of the general and disorder-specific methods are supported by detailed treatment manuals and a sufficient body of empirical research on efficacy. The mechanisms of change are often less clear. This question forms an important focus of process research in behavior therapy.

6.02.2.7 Indications

The last important aspect that is needed for a sufficient characterization of behavior therapy are the indications or prescriptions for its application. Decisions about indication are made every day and in every type of therapeutic activity. This begins with the question whether treatment is at all indicated, continues with the choice between different treatment methods and their adaptation to the individual, and stops with the decision about termination. Indication decisions thus deal with the optimal fit between patients and treatments (and other conditions such as therapists, settings, etc.). Unfortunately, only a small proportion of these decisions are made by psychoherapists. Typically, it is the patient, his relatives, his physician, his health assurance, or other lay institutions that decide whether and what kind of psychotherapy is sought. Moreover, even in the case of psychotherapists these decions are often made in implicit or irrational ways. For example, one does what one has learned or what fits into one's therapeutic ideology. Empirically-based decisions are the exception rather than the rule in much of psychotherapy.

This situation is partly due the research problems that make a fully satifactory scientific solution of the question of differential indication impossible. The most popular version of this question was formulated by Paul (1966, 1967) as early as 1966: "which is the most effectice treatment for this individual with this specific problem, by whom and under which circumstances?" A complete answer to this question would involve large experiments with factorial designs that combine all factors. Considering the number of combinations, this is illusory. On the other hand, indication decisions cannot be avoided in clinical practice. Behavior therapy therefore pragmatically considers soluble partial questions. First, it is asked whether psychotherapy is at all indicated, then what type of treatment should be applied, and how it should be adapted to the individual case. For practical reasons, clinicians need to consider not only specific scientific knowledge but also sometimes untested assumptions, individual practical experience, expert or colleague opinions, and everyday practical knowledge.

In this context, a very important achievement of behavior therapy is the development of disorder-specific treatment programs (Task Force, 1995). In its beginnings, behavior therapy meant to a large degree the application of general psychological (learning) principles to the individual case. Therapeutic procedures were therefore typically described in the rather abstract terms of general principles. Over the course of time, procedures were described in more concrete details. This led to the development of treatment manuals which were written for groups of patients rather than for individual cases. After its publication in 1980, the *Diagnostic and statistical manual of mental disorders* (3rd ed., *DSM-III*) and its succesors rapidly became the the basis for grouping patients. Standard programs for diagnostic groups often proved very efficacious in empirical testing. Thus, modern behavior therapy has two bases for its decisions: the general characteristics of the disorder and the individual characteristics of the patient.

The pragmatic approach of behavior therapy to the difficult issue of differential indication can thus be summarized as follows: for specific disorders or problem constellations those methods should be preferred whose efficacy

Table 1 Overview of typical cognitive-behavioral treatment methods.

General methods	*Disorder-specific programs*
Exposure-based methods	Agoraphobia
Systematic desensitization	Social phobia
Flooding	Specific phobia
Response prevention	Panic disorder
Operant methods	Generalized anxiety disorder
Positive reinforcement	Obsessive-compulsive disorder
Extinction	Post-traumatic stress disorder
Response cost	Depression
Time out	Anorexia nervosa
Token economies	Bulimia nervosa
Cognitive methods	Obesity
Self-instruction training	Hypochondriasis and health anxiety
Problem-solving training	Somatization disorder
Modification of dysfunctional cognitions	Somatoform pain disorder and chronic pain syndromes
Reattribution	Schizophrenia
Analysis of erroneous logics	Sexual dysfunctions
Decatastrophizing	Marital problems
Modeling	Hyperactivity and attention deficit disorder
Social skills training	Aggressiveness
Self-control methods	Autism
Self-observation	
Self-reinforcement	
Stimulus control	
Multimodal therapy or broad-spectrum therapy	

has been validated empirically. Based on a detailed problem analysis and the course of treatment, standard interventions should be adapted to the individual case. This should take individual strengths and weaknesses, personality, life situation, and setting variables as well as possible interactions of these factors into account. The decision to apply a specific treatment program presupposes competence in the diagnosis of disorders and the analysis of problem constellations. The mutual completion of classificatory diagnosis and problem analysis is therefore the basis for rational indication decisions in behavior therapy. The explicit question for the optimal therapeutic method for a given disorder as well as the offer of concrete alternatives for disorder or problem-oriented indication decisions are specific characteristics of modern behavior therapy. Examples of the various mental disorders for which treatment programs have been developed can be taken from Table 1 and Table 3.

6.02.3 HISTORICAL DEVELOPMENT OF BEHAVIOR THERAPY

6.02.3.1 Situation at Onset and Precursors

Knowledge of the past may help understanding of the present. Behavior therapy evolved from the application of the principles of experimental psychology to clinical problems (Kazdin, 1978; Schorr, 1984, 1995). Its growth was closely related the development of clinical psychology as an applied science. Although there had been sporadic early clinical applications of psychology (e.g., the case of "little Peter" [Jones, 1924a, 1924b] or the early treament of enuresis [Mower & Mower, 1938]), a broad movement started only in the middle of the century when two conditions came together:

(i) The enormous productivity of basic research on learning theory explications for clinical phenomena had become obvious (e.g., Mowrer's two-factor theory of phobias, research on experimental neurosis, Solomon and Wynne's work on traumatic conditioning, Dollard and Miller's experiments on originally psychoanalytic concepts). These findings demanded for clinical application and testing.

(ii) There were strong criticisms of the low efficacy of the then available psychotherapeutic (i.e., psychoanalytic) methods and their poor empirical basis (e.g., Eysenck's criticism of psychoanalysis [Eysenck, 1952]). Such a fundamental critique immediately posed the question of alternatives. These of course were regarded especially critically by those who had been attacked before and thus received particular attention.

It was at this time that research groups in South Africa, England, and the USA at first independently reported great successes with learning theory-based methods in the treatment of anxiety and other clinical problems. Together with the two other named conditions these startling outcomes of the new, at first experimental methods gave the impulse for a development whose breadth and dynamic until today have no parallel in psychotherapy.

6.02.3.2 The "Founding Period"

While England and the USA are widely regarded as the origin of behavior therapy, the contribution of South Africa is still underestimated although this is where many of the founding personalities of behavior therapy started their careers. The first publication of the term "behaviour therapy" in a scientific journal took place in the *South African Medical Journal* (Lazarus, 1958). Since the end of the 1940s and during the 1950s Joseph Wolpe tried to bring together learning theory and neurophysiology at the University of Witwatersrand. This was consistent with his training as a physician with central interest in learning psychology. Wolpe had been influenced by the American work of Masserman on experimental neurosis and by Salter on "self-assertiveness training." In South Africa he worked with psychologists such as Stanley Rachman and Arnold Lazarus who together with their compatriotes G. Terence Wilson and Isaac Marks belong to the most prominent founders of behavior therapy. In such groups experimental research was discussed and therapy sessions were "supervised" and observed through one-way mirrors.

In his research on "experimental neurosis" in cats, Wolpe developed new techniques to eliminate experimentally induced fear and avoidance. Based upon the notion that conditioned fear and food intake should be antagonistic and thus inhibit each other reciprocally, he assumed that feeding could be used to reduce fears resulting from specific situations. Wolpe demostrated this successfully in his animal subjects by systematically decreasing the distance between the feeding place and the place where fear had been conditioned using electric shocks. In an article entitled without modesty *Reciprocal inhibition as the central basis of psychotherapy*, Wolpe (1954) postulated reciprocal inhibition as a universal principle: "Fear reduction is achieved if fear inducing stimuli are presented together with stimuli that produce a dominant antagonistic response to fear (i.e., reciprocal inhibition)." To make certain that inhibition was stronger, he presented fear-inducing stimuli with increasing intensity, that is, in hierachical order ("fear hierarchy").

In extending his results to humans, Wolpe primarily considered three response domains for reciprocal inhibition: sexual, assertive, and relaxation responses. Most widely employable seemed a modified version of Jacobson's (1938) progressive muscle relaxation. Wolpe believed relaxation and eating to lead to similar neurophysiological effects. In order to reduce fear responses by reciprocal inhibition, Wolpe first taught his patients progressive muscle relaxation and then encouraged them to go through their feared situations step by step while staying in the relaxed state. Originally Wolpe used exposure *in vivo* (i.e., in real life situations), but then he employed imagined situations (*in sensu*) because these were easier to realize and were better controllable. In addition, patients completed comprehensive homework *in vivo* (i.e., practiced in their natural environment) between therapeutic sessions. He called this approach "systematic desensitization" and described it in his classical book *Psychotherapy by reciprocal inhibition*, (Wolpe, 1958) that was published by Stanford University Press in the USA after a recommendation by Albert Bandura. Systematic desensitization probably still is the most famous treatment method in behavior therapy, although there are now more effective methods for many clinical problems and reciprocal inhibition theory has been shaken by contradicting evidence. Wolpe's formulation of a theory on the basis of testable hypotheses with the goal of clearly defined treatment strategies for minutely described clinical applications had a tremendous impact on the development of behavior therapy.

Many of the important actors of the South African behavioral scene pursued their work in the United States and England. Wolpe's publications came at a time where the efficacy of the psychoanalytic approach was severely attacked. Especially Eysenck's (1959) controversial argument that the success rates of psychotherapy (then largely identical to psychoanalytic therapy) were not better than spontaneous remission, that is, the percentage of improved patients was not higher with psychotherapy than without it, yielded vehement debates. It is not surpsing then that Eysenck's department was the European cradle of behavior therapy. Head of the famous Institute of Psychiatry at the Maudsley Hospital in London was Aubrey Lewis, a strong advocate of the relevance of psychological research for psychiatry. In 1950 he made Eysenck the first head of a psychological department of this leading institution. Soon collaborators such as Gwynne Jones, Victor Meyer, Aubrey Yates, or M. B. Shapiro were interested in the application of conditioning theories to psychological problems. This group knew Wolpe's publications more than the operant work of Lindsley in the USA. Single case experiments and theoretical seminars that were at first purely diagnostical were soon extended to therapeutic topics. As early as 1957 Meyer emphasized the importance of a good relationship between therapist and patient for exercises *in vivo*.

The first clinical application of a behavioral approach at the Maudsley Hospital occurred more or less by chance (Schorr, 1995). When drinking coffee with a medical student, Gwynne Jones and M. B. Shapiro discussed a patient that had been treated psychotherapeutically without success. The young dancer was unable to work because she had to urinate very frequently which had in the meantime led to secondary anxiety responses and a lack of self-confidence. The conversation led to the idea to attempt a new treatment approach using conditioning techniques. A combination of systematic desensitization *in vivo* for the main complaint and a stepwise training *in vivo* for the other anxiety responses outside of the hospital brought a treatment success that proved durable at five-year follow-up (Jones, 1956, 1960).

In the 1960s the application of learning-based treatments was advanced at the Maudsley Hospital by Rachman, who had before worked with Wolpe and now had good contact with the psychiatric side of the Maudsley (personal communication by H. J. Eysenck, September 1995). Rachman played a central role in the development of aversion therapy (Rachman & Teasdale, 1969) (which was soon to be given up again), behavioral medicine, and the treatment of obsessive-compulsive disorder (Rachman & Hodgson, 1980). Other colleagues in hospitals in London and Oxford (Warneford Hospital) such as Gelder, Marks, and Mathews developed and tested exposure treatments for phobias. At the same time American researchers such as Davison (1968) investigated the process of desensitization and other anxiety reduction techniques in detail. They came to the conclusion that exposure *in vivo* was the most important and effective component of treatment (see also Kazdin & Wilcoxon, 1976). An important reason why behavioral approaches gained influence in the treatment of anxiety-related problems so fast was the fact that their efficacy was tested systematically in controlled studies. One such study that set standards for psychotherapy research was presented by Paul in 1966.

In addition to the treatment of anxiety, such diverse problems as writer's cramp, tics, and stuttering were addressed. An important element were the publications by Shapiro (1961) on experimental single-case methodology. Clinical single-case experiments typically involve a series of measurements of a clinically relevant variable in regular intervals (time series). At a predetemined point in this series an intervention is made and the effect of this intervention is then assessed by looking at the changes in the measured variable. In this way, the effects of most intervention strategies can be determined. Later, complex experimental designs were developed (see Barlow & Hersen, 1984) that made it possible to apply single-case experiments to a large number of clinical and scientific problems as a part of everyday work. Although this method is not limited to the behavioral approach, it has a close tie to it and plays an important role in its ongoing development.

At the end of the 1950s and the beginning of the 1960s behavior therapy offered already a broad scope of therapeutic methods on the basis of experimental psychology. These became fast known outside the inner circle of active researchers. Important for the dissemination were publications and the international appointments of "founding personalities" to universities and clinical institutions. The volume *Behaviour therapy and the neuroses* edited by Eysenck in 1960 already included contributions from the USA, the UK, South Africa, and Czechoslovakia. In the early 1960s, Wolpe, Lazarus, and Cyril Franks accepted professorships in the USA. In 1963 Eysenck founded the journal *Behaviour Research and Therapy*, whose editor later became Rachman. At this time the new movement became known under the name "behavior therapy," although alternative terms (e.g., "behavior modification" which was preferred by the promoters of operant methods) existed and some prominent participants simply wanted to add explanatory elements to the traditional "psychotherapy" (e.g., Wolpe's "psychotherapy by reciprocal inhibition").

It is hardly possible to determine who really first coined the term behavior therapy. This is primarily due to the fact that there was no single "founding father." Instead, the behavioral approach developed as a relatively broad movement simultaneously in several places in South Africa, England, and the USA. For the same reason, only a starting period but not a single starting date can be given. The development as a broad movement based upon empirical psychology differentiates behavior therapy from all other forms of psychotherapy. These were regularly "invented" by single charismatic personalities with more or less distance to scientific psychology. It may be argued that behavior therapy's type of origin already contains the nucleus for the development of a broad "basic orientation" rather than a narrow "therapy school." At the same time this origin can also be seen as an important protection against dogmatic immobility and as an impetus for continuous development. Better than the invention of the term, its dissemination can be attributed. Its widespread usage goes back primarily to Hans J. Eysenck and Arnold Lazarus. While Wolpe rejected such a "brand name," they postulated that the large differences to the traditional approaches should be underlined with a new name. The first printed appearance of *Behaviour therapy* was authored by Lazarus in 1958. The first public usage of the new term, however, was made by Eysenck in a talk entitled *Learning theory and behaviour therapy* (published in 1959). Here, the highly confrontative marketing strategy of Eysenck is clearly visible: he combined a fundamental and partly polemical critique of traditional methods with a partly overly optimistic view of the new approach.

6.02.3.3 The Influence of Operant Approaches

The mainstream of behavior therapy at first developed outside of America where at that time the operant approach was very popular in psychology. This is perhaps one reason why the potential clinical applications of operant methods were hardly acknowledged. Another reason was the fact that operant researchers did not come out of the clinical sector. Skinner himself never worked clinically. But even those of his followers that moved into the clinical world typically did not limit themselves to it. Instead, they always saw other fields such as education, economy, or administration as important areas of applications for their methods.

In the late 1950s, Skinner and Lindsley described the potential applications of operant methods. Corresponding therapies, however, were only conducted in the 1960s and at first with children and mentally handicapped adults. The first clinical applications are related to the names of Charles Ferster, Ivar Lovaas, Donald Baer, Sidney Bijou, Leonard Krasner, Leonard Ullman, Nathan Azrin, and T. Ayllon. The "operant group" had constructed their own network of relationships and publication outlets. They utilized their proper, highly technical terminology and restricted themselves to the narrow approach of Skinnerian psychology whose clinical application they called applied behavioral analysis or behavior modification. The term behavior therapy was rejected as much as the term "patient" for the addressee of their interventions. Eysenck, Rachman, Wolpe, and their colleagues became aware of the clinical work of the American operant school only in the early and mid-1960s. Even after that communication between the two groups was rather reluctant. Integration of the operant methods into the behavior therapy movement was not unproblematic. It was only since the 1970s that the terms behavior therapy and behavior modification became regarded as more or less equivalent. Until today, a small group of strictly operant researchers have kept their own tradition outside of the more clinically oriented behavior therapy ("applied behavior analysis").

All in all, the relevance of operant methods is frequently overestimated, especially from outside behavior therapy, although they are only rarely applied as sole treatment methods. They do, however, have a firm place in parts of behavior therapy, for instance in the treatment of childhood behavioral problems, mental retardation, or chronically institutionalized patients. Studies such as those on "token economies" in long-term institutionalized mental patients underlined the importance of social reinforcement for change in general, for long-term generalization and the maintenance of desired or acceptable behaviors (Ayllon & Azrin, 1968). More recent research, however, has yielded doubts about the theoretical basis of token economies and other methods that were originally conceptualized as purely operant techniques. Thus, it has been shown that social feedback and specific guidelines for action were the most important factors in such programs, more important than the "tokens" themselves or the reinforcers they represented. In spite of this, the development of token economy programs was important beyond their immediate applications because they underlined the relevance of a comprehensive approach in rehabilitation. The usage of structured social reinforcement (e.g., praise) has been accepted more widely in clinical practice than the usage of tokens or symbols for reinforcement. Emphasizing the role of changing and structuring social interactions has for instance strong importance in the treatment of schizophrenic patients (Fallon, Boyd, & McGill, 1984; Hahlweg, Dürr, & Müller, 1994). Another very important influence of the operant approach was the acceptance of functional analysis (behavior or problem analysis). Behavior analysis in behavior therapy is based upon the Skinnerian concept of explaining behavior by studying the conditions under which it emerged. The relevance of living conditions, environmental factors, and social relationships had before been underestimated or even overlooked.

6.02.3.4 Consolidation and Broadening of Scope

With the founding of clinical/scientific societies devoted to behavior therapy a first period of consolidation of the stormy development began.

The American AABT was founded in 1966 as the Association for the Advancement of Behavioral Therapies. Later, the plural was given up and the name was changed to Association for the Advancement of Behavior Therapy. The first members of AABT came largely from two groups: first, a primarily academic group of scientist-clinicians interested in "clinical psychology as an experimental science," who had been organized in a subgroup of the American Psychological Association (Division 12, Section 3); second, a primarily clinically active group that had been influenced directly by founding personalities of behavior therapy such as Wolpe, Franks, Salter, or Reyna. Cyril Franks was elected as the first president of AABT, the vice-president was Wolpe. The council of the society consisted of equal numbers of scientists and practitioners. While AABT experienced an enormous increase in membership, the European sibling society EABT (European Association of Behaviour Therapy) was founded in 1971. The intiative for this was taken by Johannes C. Brengelmann, who had spent long years with Eysenck in England and in the USA. He returned in 1967 to Germany and became Director at the Max-Planck-Institute of Psychiatry in Munich and head of its psychological department. Brengelmann, who was the decisive personality for the development of behavior therapy in Germany and several other European countries, became the first president of EABT. Before this, individual national societies for behavior therapy had been founded in several European countries (e.g., the German society was founded in 1968 and one year later had already 450 members). The members of EABT at first came from similar sources as those of AABT, namely academic researchers and practioners with a close contact to founding centers such as London or Munich. The "official" founding of EABT took place during a congress in Munich, which was attended by 1200 participants from 14 countries.

Consolidation was also furthered by a rapidly increasing flood of research works and publications in journals or books and by the implementation of the first teaching or training centers. In the 1970s behavior therapy therefore came to a first stage of maturity while there were still many new techniques being developed, tested and refined. Toward the end of the decade, most of these treatment approaches were internationally accepted. Behavioral methods were the treatment of choice for such diverse problems as phobias, obsessions and compulsions, sexual dysfunctions, or the rehabilitation of chronic mental patients. At the same time existing techniques were further developed (e.g., shortening of the duration of exposure or relaxation treatments) and new methods were introduced or disseminated (e.g., social skills trainings in groups, behavioral treatment of depression). The treatment of sexual dysfunctions developed less from behavioral research on these disorders, but rather from the work of Masters and Johnson (1970) on the physiology of sexual responses. Even though this approach had an independent origin, it shared the emphasis on concrete operationalization of treatment strategies and on empirical testing with behavior therapy. This was the main reason for its rapid integration into behavior therapy which gives further testimony to the concept of behavior therapy as a broad basic orientation.

A significant broadening of the behavioral approach was introduced by the development of behavioral medicine. The term had originally been used by Birk (1973) to describe the application of biofeedback to medical problems. In biofeedback, patients are given immediate information about relevant changes in their physiological system (e.g., by auditory or visual feedback) in order to acquire control even of involuntary physiological responses. Behavioral medicine soon extended to a substantially larger field. Today, it encompasses all applications of psychological knowledge and techniques to purely somatic problems (e.g., burnings, tumor pain), disorders of possible partial psychological origin (e.g., irritable bowel syndrome), or risk factors (e.g., smoking, nutrition, physical exercise). Behavioral medicine has become a largely interdisciplinary movement that has developed a multitude of applications.

Another important development consisted of overcoming the narrow borders of the behavioristic heritage of the early learning theory-based behavior therapy. Strictly speaking, Mowrer's (1947) two-factor theory which was frequently used to explain phobias already contained a departure from behaviorism. The notion of negative reinforcement of phobic avoidance by anxiety reduction implied the concept of anxiety as an internal, "mental" state that was not directly observable. In addition, Eysenck and others had always accepted an influence of biological and genetic factors. Especially important for the advancement of behavior therapy, however, was the acceptance of the so-called "three-systems approach." Peter Lang, Rachman, and others (Hodgson & Rachman, 1974; Lang, 1993; Rachman & Hodgson, 1974) argued that psychological reactions and problems related to them needed to be seen as multidimensional, loosely connected responses systems acting on different "levels." The most popular classification differentiates between a subjective/cognitive, a behavioral, and a physiological level of response. Although these response systems are interconnected, they do not necessarily have to change in the same direction, at the same time, to the same degree etc., a phenomenon that has been termed "desynchrony" by behavioral writers (Hodgson & Rachman, 1974; Rachman & Hodgson, 1974).

This alternative to a uniform view of psychological problems was important because it allowed explanation of the wide variation in the symptom patterns reported by the patients. It also permitted a more systematic and precise assessment of treatment outcomes. In addition, differential outcomes could now be observed. Thus, relaxation exercises might influence the physiological aspects of a problem more readily than the behavioral or subjective problems. Moreover, the relevance of experience was more strongly acknowledged although a largely behavioristic, response-oriented language continued to be used. The threefold classification by Lang or Rachman has frequently been criticized. There is indeed no *a priori* reason to assume three levels rather than four or more response systems. Some authors argue that it would be better to distinguish between a cognitive and an affective system which would result in a four-systems approach. Other authors such as Fahrenberg (1987) pointed to the large body of findings that question the assumption of a unitarian, homogenous physiological level. Nevertheless, it was important to overcome the early monistic and one-dimensional view of human responses. Today, a "many-systems approach" with its emphasis on desynchrony is generally accepted in spite of the ongoing discussion about the best classification of response systems. In clinical practice, Lang's (1993) three-systems approach is the established (although simplified) standard.

At the end of the 1970s the usefulness of behavior therapy was widely accepted. Because researchers now no longer had to prove the general efficacy of their approach, some of them directed their attention to those patients that were not helped by behavior therapy, even when it was applied correctly. These studies culminated in the first scientific psychotherapeutic book on the failures of treatment by Foa and Emmelkamp (1983). Already in the late 1960s and early 1970s a beginning discontentment with the strict behavioristic guidelines of the early years was apparent. Especially Lazarus (1971) criticized the "mechanistic" terminology of early behavior therapy. He argued that the multitude of behavior therapy's interventions could not be explained by learning theory alone. Instead, he proposed the introduction of "broad-spectrum behavior therapy" that should include all techniques with empirically demonstrated efficacy regardless of their theoretical background. In clinical practice, this approach was accepted by an increasing number of behavior therapists. A further important development at this time was the attempt to apply behavioral methods to other mental disorders, especially depression. After Ferster (1965) and Lewinsohn (1974) had assumed that depression resulted from a lack of positive reinforcement, first treatment approaches consisted of elevating the rate of positive reinforcement. The success of this method stayed limited, perhaps because patients typically judged efforts and outcomes negatively even if they attempted to comply with the treatment regimen. This made the relevance of cognitive factors increasingly clear. The lack of satisfaction with the strictly behavioral techniques led to the attempt to add cognitive methods to them. This opened the way for a more systematic integration of these methods and for the development of new cognitive-behavioral approaches.

6.02.3.5 Integration of Behavioral and Cognitive Approaches

Lang's theory of three relatively independent response systems had strengthened the basis for the acceptance of cognitive ideas within the behavioral approach. Empirical psychology had already acknowledged the relevance of cognitive variables to such a degree that the "cognitive turning-point" had been passed by most actors. The delayed integration of the cognitive dimension into behavior therapy probably had its roots in the continuing influence of Watson's rejection of introspection and in the experiences with its abuse potential in "talking cures" of limited success. In addition, the polemical position toward other psychotherapeutic approaches that behavior therapists typically took in the founding phase made it difficult to abandon identity-producing boundaries. Bandura's work on vicarious learning, however, had directed the attention even of behavior therapists toward cognitive factors. The fact that people were able to learn new behaviors by observing the behavior of others even without being reinforced themselves transgressed traditional learning theory. Later on, Bandura developed a model of self-regulation that he termed "self-efficacy." This approach postulates that any conscious behavior change is based on the subject's conviction of its capacity to perform this concrete behavior. The rigorous empirical methodology of Bandura's work facilitated its acceptance even by researchers and practitioners who had originally been behavioristically oriented. In addition, the increasing interest in the concept of self-control had an important influence on the development of behavior therapy. People observe themselves, judge their own performance, and reinforce (or punish) themselves accordingly. On each of these levels, problems may arise and lead to clinical states such as depressed mood. Research into these phenomena led to the investigation of cognitive constructs including attribution and self-instruction.

Perhaps the first cognitive therapeutic method that gained acceptance in behavior therapy was Meichenbaum's (1975) self-instruction training. The popularity of this approach was bolstered by its simple, consistent theoretical basis and its resemblance with the operant concept of "coverants." Meichenbaum argued that behavior changes can be induced by changing the instructions that patients give themselves. Inadequate and negative thoughts should be transformed into more adequate statements. It is not by chance, that self-instruction training was first implemented in the treatent of impulsive children. The inner monologue and the thinking of these children were termed self-verbalizations and seen as equal to the coverants in operant terminology. The artifical term "coverant" was derived from "covert" and "operant" and thus denoted covert operant behavior. With the help of this terminology, mental contents were "smuggled" into the operant movement.

In contrast, independent cognitive approaches that had developed outside of behavior therapy encountered much more hesitation. Although Beck's (1967) cognitive therapy and Ellis' (1962) rational-emotive therapy originated in the 1960s, it took many years before they grew together with the behavioral movement. Until late in the 1970s cognitive and behavioral approaches were seen as separate "schools." This led to sometimes inadequate differentiations or to attempts to prove that the other school was "in reality" using methods of one's own school. It was only in the 1980s that the two movements converged and only in 1995 the first joint "World Congress of Behavioural and Cognitive Therapies" was held in Copenhagen. Integration also became obvious in the renaming of EABT into EABCT (European Association for Behavioural and Cognitive Therapies) in 1992 on the occasion of the 22nd Congress of the "old" EABT in Coimbra, Portugal. A common basis between the cognitive and the traditional behavioral movement were the methodological standards emphasizing operationalization, experimental research, and systematic testing of efficacy on the one hand and the rational, pragmatic orientation on the other. Today, cognitive, behavioral, and cognitive-behavioral treatments together form the broad basic therapeutic orientation alluded to in Section 6.02.2. There are, however, also critics of this convergence (for instance, Krasner's "paradigm lost," that eloquently mourns for the loss of the clear operant paradigm).

6.02.3.6 Continuous Development and the Future

Modern cognitive-behavioral therapy attempts a true integration of its components. It aims to help patients recognize their individual structure of distorted thinking and inadequate behavior. Systematic cognitive restructuring and carefully structured behavioral tasks are designed to help the patients in dealing with problems in both domains. In the meantime, cognitive-behavioral treatment programs have been developed for most mental disorders that occur in everyday practice. In some instances it is the behavioral part, in others the cognitive part that dominates. Moreover, a host of findings from process research points out that even in classical behavioral treatments such as exposure *in vivo* cognitive processes may contribute as central mechanisms of change.

The behavioral approach is still characterized by rapid change. One of the more important new achievements is the development of specific treatment programs for a rapidly increasing number of disorders or specific problem constellations. These are frequently represented in the form of concrete therapy manuals in order to facilitate practical implementation. Therapy manuals were originally used in treatment research as a mean to ensure treatment integrity in clinical studies. They were quickly disseminated in the practice sector, because they made therapeutic methods more concrete and thus accessible to the broad public of practitioners. Texts on psychotherapy had for a long time been overly abstract and often lacked any concrete descriptions or guidelines of what needed to be done in practice. Manuals with concrete guidelines therefore fulfilled an important need. The positive view of treatment manuals is reflected among others by the criteria for empirically validated treatments that were established by the Task Force on Promotion and Dissemination of Psychological Procedures of the American Psychological Association (APA). Their criteria see the existence of a manual as one precondition for any therapy method to be accepted as scientifically validated. Another important line of development is ongoing professionalization of behavior therapy in routine patient care and therapist training. Training issues are especially relevant in most European countries where a number of different private and public training models coexist. Several countries have passed laws regulating the practical application of psychotherapy including behavior therapy. These laws have important implications for research and practice as do managed care and health maintenance organizations in the USA.

The historical perspective shows what developments have led to acknowledgment of the benefit that the behavioral approach has for most mental disorders. The vast amount of change in the past leads to the prediction that the future of behavior therapy will also be characterized by major modifications of present methods and concepts. This is facilitated by the fact that ongoing research and development are major attributes of behavior therapy. But even in its present form the behavioral approach offers effective help for patients as well as a valuable basis for the understanding of mental disorders. The empirical basis for these assertions is the topic of the following section.

6.02.4 EMPIRICAL STATUS OF BEHAVIOR THERAPY

6.02.4.1 Problems in Empirically Approaching Efficacy

Psychotherapy research has moved from one difficulty to another: after a long period with a massive lack of empirical findings we now have the opposite problem. Largely due to the behavioral approach there are now so many studies on the outcome and process of psychotherapy that it has become hard for anyone to keep track. Aggregating findings across a vast number of different studies is a problem in itself (see below). Without any doubt behavior therapy is by far the most intensely researched type of psychotherapy (Grawe, 1992; Grawe, Donnati, & Bernauer, 1994). For behavioral and cognitive methods there have been over 10 times more published studies than for all other types of psychotherapy together (Grawe et al., 1994). Even some individual behavioral methods have been put to test more frequently than the whole group of psychoanalytic or humanistic approaches, respectively. In addition, behavioral methods have been applied to the broadest spectrum of mental disorders with typically positive results. However, such a general statement needs to be specified because "automatic" generalization across all disorders is senseless. Instead, individual disorders have to be regarded separately. Of course a larger number of studies does not necessarily imply better efficacy compared to other treatments. A more detailed assessment is thus needed although the whole body of psychotherapy research cannot be summarized here. I will therefore only briefly summarize the findings on behavior therapy.

How should treatment studies be aggregated? The most important approaches to the problem of secondary analysis used so far can roughly be classified into three groups:

(i) *Narrative summaries.* Here, the collected studies on a topic are simply interpreted more or less intuitively and a personal judgment of their outcomes is given. A severe disadvantage of this approach is its lack of objectivity. It has been shown that most people distort their judgment in the light of pre-existing biases. Moreover, the rapidly increasing number of publications makes it very difficult for any individual to have a complete overview.

(ii) *Box-score method.* Based on the justified criticisms of narrative summaries, this method attempts a more systematic and quantitative approach. It simply consists of counting the number of statistically significant effects that were obtained in the different studies. This, however, is subject to so many limitations that the approach is of dubious utility. Especially important is the problem of low statistical power in most psychotherapy studies. The difficulties in recruiting patients and conducting the studies typically result in small sample sizes that do not allow the detection of even clinically relevant effects with sufficient certainty. In addition, simply looking at statistical significance overlooks information about the strength of the effects ("clinical significance"). Moreover, it is unacceptable to give studies with widely varying methodological quality or very different outcome criteria equal weight.

(iii) *Meta-analysis.* This last problem also applies to meta-analysis, which otherwise avoids the power problem of the box-score method. Meta-analysis is at present the most objective method for aggregating findings from different studies. It starts by calculating "effect sizes" for each measure taken in each study. Effect sizes express by how many standard deviations the mean of two groups (e.g., treatment and control group) or of one group at two times (e.g., pre- and post-therapy) differ. This standardization makes different measures directly comparable, even across separate studies. In a second step, all calculated effect sizes are then averaged to yield one "integrated" effect size as a global index of therapy outcome. As stated above, a criticism is that the methodological quality of the studies or even clinical relevance (e.g., clinical vs. analogue sample) typically are not taken into account. Of course, one could introduce weights for methodological quality or set methodological criteria for inclusion into the meta-analysis. However, this introduces, a subjective element and thereby affects objectivity that together with its quantitative nature is a great advantage of the method. The utility of meta-analytic results has been challenged by referring to the fact that frequently very different studies are thrown together ("comparing apples and pears"). In spite of these criticisms, meta-analysis has established itself as an objective statistical method for the aggregation of heterogenous findings across separate studies.

6.02.4.2 Results of Outcome Research

The most comprehensive and differentiated analysis of outcome research so far has been published by Grawe and co-workers (1994). Shortly after this work appeared in print, it became a standard for any discussion of the topic. In contrast to earlier meta-analyses, Grawe et al. analyzed the methodological quality of the studies, took the different types of outcome measures into account, and included all controlled studies published until 1983 regardless of language, country of origin, or publication outlet in their thorough analysis. Table 2 shows the number of controlled studies that Grawe found for the different classes of therapy methods and for the individual cognitive-behavioral techniques. Table 3 shows for what types of disorders behavioral methods have been tested.

In these studies, behavioral methods have proved effective in achieving the immediate goals of treatment as well as generalized benefits that most of the individual techniques have to be given the status of empirically validated therapy methods (Grawe, 1992). Especially well researched are exposure techniques that were frequently applied to agoraphobia, obsessive-compulsive disorder, and specific phobias. They led almost always to massive improvement of the target symptoms, while the effects on general well-being and other masures were somewhat smaller. Well established is also the effect of systematic desensitization in treating anxiety disorders. A very broad range of effects was found for social skills training that almost regularly achieved significant positive changes beyond the target symptoms. Broad efficacy was also found for cognitive methods such as Beck's treatment for depression, rational-emotive therapy, Meichenbaum's self-instruction, and stress-inocculation trainings or problem-solving treatments. Studies revealed not only a very broad range of clinical applications but also some degree of specific relations between certain treatments and individual disorders. Thus, anxiety disorders were most frequently treated with exposure-based methods followed by systematic desensitization and cognitive methods. "The application of behavior therapy in clinical practice is thus based on a broad spectrum of treatment

Table 2 The number of controlled studies that were found by Grawe et al. in their meta-analysis of published therapy outcome studies.

Different types of therapy	*Number of studies*	*Only cognitive-behavioral methods*	*Number of studies*
Cognitive-behavioral methods	567	Social skills training	74
Progressive muscle relaxation[a]	66	Exposure	62
Autogenic training	14	Biofeedback	62
Hypnosis	19	Systematic desensitization	56
Meditation	15	Cognitive coping methods	38
Psychoanalytic short-term treatments	27	Aversion therapies	31
Psychoanalytic therapy	12	Problem-solving therapies	25
Long-term psychoanalysis (Freud) and analytic therapy (Jung)	each 0	Sexual therapy	22
Client-centered therapy	35	Lewinsohn's therapy of depression	17
Gestalt therapy	7	Rational-emotive therapy	17
Psychodrama and transactional analysis	each 6	Beck's therapy of depression	16
Marital therapy	35	Alcoholism programs	14
Family therapy	18	Paradoxical intention	10
Interpersonal therapy	10	Broad-spectrum behavior therapy	8

[a] The classifiaction of progressive muscle relaxation is controversial. It can be subsumed under behavioral methods (of which it typically was part) or counted as a separate entity. We chose the later approach in order to give more detailed information. Other methods not mentioned in the table were either not specfied clearly enough (unspecified humanistic therapies $N = 11$, encounter-groups $N = 9$, unspecified psychodynamic therapies $N = 8$, psychodynamic therapies with drugs N = 13, eclecticistic therapies $N = 22$) or counted only 1–3 studies (music therapy, body therapy, dance/art therapy, bioenergetics, Ich-Analyse, Individual therapy (Adler), Katathymes Bilderleben, Daseinsanalyse). Because many behavioral studies used combinations of methods, the numbers in the right column do not add to the total given in the left column for cognitive-behavioral methods.

Table 3 Types of mental disorders for which behavioral methods have been tested empirically in controlled treatment studies based on Grawe's 1994 meta-analysis. For each disorder, the number of studies published until 1983 is given.

Therapeutic method	*Childhood disorders*	*Organic mental disorders*	*Substance abuse disorders*	*Schizophrenia*	*Other psychoses*	*Affective disorders*	*Anxiety disorders*	*Sexual dysfunctions*	*Disorders of impulse cntrol*	*Disorders of adaptation*	*Somatic states with psychological factors*	*V-codes for states without mental disorders*	*Additional codes*
Cognitive			10	1		18	19	1	1	5	15	12	20
Various	10		11	1		16	5	2	2		3	28	9
Biofeedback	7		2	1		1	9	1		2	57		2
Systematic desensitization	2		2	3		1	37	12		4	4		6
Operant	4	1	6	11	3	4	4	3		1	4	11	19
Social skills training	1		7	4	4	7	9	4		3	6	3	22
Exposure			1		1	1	57			1		1	3
Aversion	3	1	24		1	2	910			1		1	
Modeling	1		3	5	1		11	2		2	3	3	18
Self-control	10		12	1		4	4		1		4	1	7
Covert conditioning			7	1		2	7	1			2		11
Speech therapy	7						1				1		

Table 4 Results of meta-analyses comparing the efficacy of different types of psychotherapy. Because of the low number of studies, psychodynamic and humanistic methods were typically grouped together.

	Psychodynamic/ humanistic methods	*Cognitive-behavioral methods*	*Direct comparisons within one study*
Meta-analysis	*Mean effect sizes*	*Mean effect sizes*	*Difference between effect sizes*
Smith et al. (1980)	0.64	1.03	−0.49
Shapiro & Shapiro (1982)	0.40	1.08	−0.53
Nicholson & Berman (1983)	0.29	0.75	−0.44
Wittmann & Matt (1986)	0.30/0.25	0.50	
Grawe et al. (1994) (psychoanalytic methods)	0.83	1.21	only direct comparisons
Grawe et al. (1994) (client-centered methods)[a]	0.87	1.13	only direct comparisons

[a] Without the study of Angulo (1983; quoted in Grawe et al., 1994), whose results represent clear outliers (effect sizes for behavior therapy about 8, for client-centered therapy about 4, patients with obsessive-compulsive symptoms).

methods with proven efficacy. With a clear distance to other types of treatment, behavior therapy can therefore claim to have proven its effectiveness sufficiently for playing a prominent role in psychotherapeutic mental health care" (Grawe, 1992, p. 139).

In addition to the general statements on the efficacy of behavior therapy, Grawe et al. (1994) also analyzed all studies directly comparing behavior therapy with psychoanalytic, client-centered, or systemic treatments that were published until 1991. Further inclusion criteria were at least three outcome measures, at least six hours of treatment duration, similar duration ("dosage") of the different treatment conditions (unless differences were explicitly justified on a theoretical basis), and the publication of means and standard deviations of outcome measures. In this way, they found 41 studies with a total of 1401 effect sizes. The large number of effect sizes results from the great number of treatment conditions and outcome measures. The direct comparisons underlined impressively the results reported above for the behavior therapy studies: The comparison with psychoanalytic (19 studies, 215 outcome measures, 487 individual comparisons), client-centered (10 studies, 133 outcome measures, 723 individual comparisons), and family-therapy methods (3 studies, 18 outcome measures, 40 individual comparisons) always yielded a clear-cut superiority of behavioral approaches. These results are in line with those of earlier meta-analyses as can be seen in Table 4. This table summarizes the effect sizes that have been calculated in the different meta-analyses. Because the calculatory basis differed considerably across meta-analyses, the comparison within each meta-analytic study should primarily be regarded (thus Wittmann and Matt (1986) included only German-language publications that were in addition analyzed with a different algorithm; Grawe et al. (1994) analyzed only direct comparisons, etc.). In every meta-analysis cognitive-behavioral methods fared better than other treatments. This result holds up after methodological problems (such as restriction to direct comparisons, consideration of different types of outcome measures, limitation to certain types of disorders, liberal or conservative inclusion criteria, etc.) have been addressed.

An important alternative to the meta-analytic approach was proposed by the Task Force on Promotion and Dissemination of Psychological Procedures of the APA (1995). Their approach takes into consideration one major criticism of most meta-analytic studies published so far, namely the neglect of the differences between the various mental disorders. It is inadequate to exclude the type of disorder treated in outcome research. The assumption that disorders as different as alcoholism, sexual dysfunctions, agoraphobia, or schizophrenia can be treated by the same method regardless of the specific characteristics of the disorder is unrealistic. The question for efficacy therefore always has to address clearly defined categories of problems or disorders. In addition, the APA Task Force considered treatment integrity. In contrast to, for instance, pharmacological treatments, psychotherapeutic approaches need to clarify what concrete actions are included under such labels as "behavior therapy" or "client-centered therapy." For this purpose the treatment manuals mentioned above are pivotal. Moreover, the criteria reflect the notion that controlled treatment studies may be the best but not the only valid source of empirical information on treatment outcomes. For long-term treatments these studies have traditionally been regarded as being hard or even impossible to perform. Single-case experimental studies offer an alternative if they fulfill basic methodological standards. Another requirement was that positive reports on a treatment's efficacy should come from more than one independent center, thus avoiding situations where the "founder" of a therapy keeps finding his treatment perfectly useful. On the basis of these and additional considerations, minimal criteria for empirically validated treatments were established. In addition, the Task Force made a difference between "well-established" and "probably efficacious" methods. The criteria are listed in Table 5. A first, still provisorial listing of treatments that fulfill the two sets of criteria is given in Table 6.

The results of the APA Task Force resemble those discussed above for the meta-analytic approach. The vast majority of treatments listed in Table 6 come from the cognitive-behavioral approach. This is even more important, because the proposal for the APA criteria which also accepts single-case studies as valid data originated from an psychoanalytically oriented member of the Task Force (1995). Using a quite different approach, this Task Force composed of members with various therapeutic orientations came to the same result as Grawe et al. (1994) or the other quoted meta-analyses.

6.02.4.3 Consequences of Research Findings

The behavioral approach therefore very clearly deserves the certificate of being the best validated psychotherapeutic orientation. In addition to the large number of studies on the efficacy of the individual treatment methods, the vast body of knowledge collected by clinical psychology and its neighboring disciplines on the etiology, diagnosis, and epidemiology of mental disorders contributes to this success. For most disorders, specific treatment programs tailored to their individual characteristics have been developed and tested. Behavioral treatments can therefore have very different contents depending on the nature of the problem to be treated. In principle, the superior empirical basis for behavioral treatments should imply that these methods should play a central role in psychotherapeutic patient care. However, one cannot derive from the research findings that behavioral methods are the only ones that should be applied. There are still no "100% methods" for most mental disorders. Even the best treatments yield a certain amount of failures. As long as this is the case, we need a broad spectrum of treatment options that address different facets of the problematic structures. Yet this does not mean that every

Table 5 Criteria for empirically validated treatments established by the APA Task Force on Promotion and Dissemination of Psychological Procedures (1995, 1998).

Well-established treatments

I. At least two good between group design experiments demonstrating efficacy in one or more of the following ways:
  A. Superior to pill or psychological placebo or to another treatment
  B. Equivalent to an already established treatment in experiments with adequate statistical power (about 30 per group)

OR

II. A large series of single case design experiments demonstrating efficacy
These experiments must have:
  A. Used good experimental designs and
  B. Compared the intervention to another treatment as in I.A

FURTHER CRITERIA FOR BOTH I AND II:

III. Experiments must be conducted with treatment manuals

IV. Characteristics of the client samples must be clearly specified
V. Effects must have been demonstrated by at least two different investigators or investigatory teams

Probably efficacious treatments

I. Two experiments showing that the treatment is more effective than a waiting-list control group

OR

II. One or more experiments meeting the well-established treatment citeria I, III, and IV, but not V

OR

III. A small series of single case design experiments otherwise meeting well-established criteria II, III, and IV

Table 6 Treatments named by the APA Task Force on Promotion and Dissemination of Psychological Procedures as examples of their criteria for empirical validation (1995).

- Well-established treatments
  - Cognitive therapy for depression
  - Behavior modification for developmental disorders
  - Behavior modification for enuresis and encopresis
  - Behavior therapy for headache and irritable bowel syndrome
  - Behavior therapy for orgasm and erectile dysfunctions
  - Behavioral marital therapy
  - Cognitive-behavioral therapy for chronic pain
  - Cognitive-behavioral therapy for panic disorder and agoraphobia
  - Cognitive-behavioral therapy for generalized anxiety disorder
  - Exposure therapy for phobias and post-traumatic stress disorder
  - Exposure and response prevention for obsessive-compulsive disorder
  - Psychoeducative family intervention for schizophrenia
  - Cognitive-behavioral group therapy for social phobia
  - Interpersonal therapy for bulimia
  - Interpersonal therapy for depression
  - Training programs for parents with behavior problems
  - Systematic desensitization for specific phobias
  - Token economy programs
- Probably efficacious treatments
  - Applied relaxation for panic disorder
  - Psychodynamic short-term therapy
  - Behavior modification for sexual offenders
  - Dialectical behavior therapy for borderline personality disorder
  - Emotion centered couples therapy
  - Habit reversal training and habit control training
  - Psychoeducational therapy of depression

therapist can do what they want regardless of the research findings. We need to ask ourselves how to choose the most promising method for each individual case or for the rank order of the different approaches that could possibly be applied to the case. These decisions need to take into account other variables such as the motivation and personality of the patient, the possible existence of comorbidity, or the availability of treatments or therapists. In addition, it has to be clarified whether and when the patient should be transferred to another specialist and what role primary care physicians and self-help measures could play. As a general rule, measures with the best cost-benefit ratio should be tried first. These considerations lead to the proposal of more or less rational "sequential models" for clinical practice (Margraf, 1996) such as the one detailed in Table 7. Of course, the model delineated below represents an abridgement, that in clinical practice needs to be specified according to disorders and other factors.

The ideas behind this sequential model can be summarized as follows.

(i) It is generally better if the patient is able to solve his problems himself. Therefore, given an adequate motivation the person should be encouraged to try self-help measures. The input of lay personnel such as relatives or friends also belongs here.

(ii) If self-help fails or is insufficient, however, professional help should be sought. Here we can distinguish between general, nonspecific counseling or supportive measures on the one hand and a specfic therapy for the mental disorder on the other. As long as there are no acute crises, suicidal ideation, active psychoses, or a long history of failed treatments, generalists such as the family physician can try to achieve improvement using nonspecific measures.

(iii) If this does not lead to clear-cut improvement or the listed conditions are not met, a specific treatment of the mental disorder should be started as fast as possible in most cases, beacuse most mental disorders otherwise have a high risk of chronification. The decision between the competing psychological or pharmacological treatments should not simply be based on the personal preference of therapist or patient. Instead, it should consider probability of success, cost-benefit ratio, durability of treatment outcomes, and the other conditions discussed above (e.g., availability).

(iv) Among psychological treatments, cognitive-behavioral methods should typically (for most disorders) form the first line of

Table 7 A sequential model for choosing treatment interventions in clinical practice.

1. Self-help, lay interventions
2. Advice, counseling, supportive therapy
3. Specific treatment of the mental disorder
    3.1. Behavior therapy
    3.2. Other psychotherapeutic and pharmacological interventions
    3.3. Long-term interventions after treatment failures

Source: Margraf (1996).

intervention because their efficacy is best validated and they have a favorable cost-benefit ratio. Only if this does not lead to sufficient success should other types of psychotherapy be applied because these typically are less well established empirically and often involve greater expense (e.g., longer duration).

(v) Similar considerations apply to the choice of pharmacological methods. Here, the questions of durability of outcomes or relapse after drug withdrawal and of side effects (cost-benefit ratio) may lead to critical reflections.

(vi) In some cases, the necessity of long-term support may arise. This could for instance be the case if intensive attempts have not led to clinically significant improvement. Here its is important to convey to the patient that he or she is not alone and to avoid making the patient "guilty" for failure. Of course, it needs to be clarified whether all promising treatment methods have been attempted competently. In clinical practice, one frequently sees cases where due to their training or theoretical orientation, care providers used only a narrow part of the total spectrum of treatment methods. In these cases, transfer to another specialist is often helpful. In other cases, however, all that is left is to convey to the patient a realistic notion of his chances for recovery and to support him in dealing with the chronic aspects of his condition.

6.02.5 PROBLEMS AND CRITICISMS OF BEHAVIOR THERAPY

Like any other human endeavor, behavior therapy has its share of problems. These are for instance the treatment failures that may occur even with good motivation on the patient's side and optimal application on the therapist's side. Here, research is called for in order to ameliorate success rates, attrition, etc. Other important problems of behavior therapy lie in limiting setting factors such as lack of information of patients and professionals or obstinate misunderstandings about the behavioral approach. In addition, the availability of well-trained behavior therapists is a major problem in most parts of the world, even in those countries where the behavioral movement started. Moreover, it is very difficult for patients or other physicians or psychologists to judge the competence of any given colleague. Many national health insurance schemes (or the lack of them) are responsible for limitations to the availability of competent behavior therapy. Another aspect are the wishes and prejudices of the patients that often come with strong feelings about their treatment. These are frequently conveyed by media, popular literature, friends, or relatives. Some decline any notion of pharmacological help ("how can pills help with psychological distress?"), others definitely do not want to be treated by psychotherapy ("I am not crazy!," "They only dig in the past!"). Some people, especially with an academic background, have powerful convictions even with respect to the specific kind of psychotherapy ("in depth," "thorough," "explain everything"). All of these points together may be responsible for the fact that behavior therapy is applied far less frequently than one should assume on the basis of the research findings and the productive publication activities of its proponents.

6.02.5.1 Misunderstandings About Behavior Therapy

An important obstacle to the dissemination of behavioral methods have been several obstinate misunderstandings about its nature or techniques (Baer & Minichiello, 1990; Yates, 1977). Some of these misconceptions will be briefly addressed here.

(i) Behavior therapy does not lead to symptom substitution.

(ii) The experience of strong emotions during exposure (e.g., in phobias, grief reactions, post-traumatic disorders, or eating disorders) is not dangerous for the patients.

(iii) The thoughts and feelings of the patients are not ignored. Instead, they are addressed directly by a multitude of therapeutic interventions.

(iv) Modern behavior therapy does not assume that all mental disorders have been acquired by simple conditioning processes.

(v) The use of medication is not generally incompatible with behavioral interventions.

Inasmuch as these misunderstandings are being clarified and together with the mounting reception of the results of psychotherapy research, a growing interest in behavioral methods can be observed. But even today, the lack of well-trained specialists is a major problem.

6.02.5.2 Scientifically Well Established, Practically Neglected?

A remarkable example of the underutilization of behavioral methods is given by the treatment of anxiety disorders. If anything, these disorders should be a primary area of application for behavioral methods. However, procedures such as exposure *in vivo* are apparently given only to a small minority of afflicted persons. Taylor et al. (1989) studied a large sample of patients with anxiety disorders that had applied for treatment to the anxiety disorders clinic at Stanford University School of Medicine. Although the region surrounding the clinic was densely populated with care providers, they found that less than 10% of agoraphobic patients had been given a trial of exposure.

One criticism of this study can be seen in the fact that the nonrepresentative sample consisted of self-selected patients searching treatment. However, we recently found the same effect in a sample 3000 people representative of the German adult population (Margraf, 1996; Margraf & Poldrack, in press). The study investigated the prevalence of clinical anxiety in the general population using the Beck Anxiety Inventory using a cut-off score from clinical anxiety research. In face-to-face interviews, subjects who reported anxiety were asked whether (and if so, how) they had been treated for their anxiety problem. Results showed that even with the broadest definition of treatment only 40% of all afflicted persons had received some kind of treatment and that primary care physicians were by far the most frequent care providers (82% of all treated patients). Genuine specialists such as clinical psychologists (3%) or psychiatrists (6%) together performed fewer treatments than other somatic physicians (17%, multiple responses possible). With respect to the treatments used, there was a remarkable discrepancy between drugs (roughly 90% of all patients, multiple responses possible) and nonspecific advice (roughly 75%) on the one hand and psychotherapy (17%) on the other. In addition, an elevated rate of inpatient treatment (10%) became apparent.

Yet the most impressive result to us was that psychotherapy consisted with roughly equal numbers of relaxation methods (including hypnosis and autogenic training) and "talking cures" (i.e., primarily psychoanalytic or client-centered methods), while cognitive-behavioral methods accounted for only 1% of all treatment cases. In view of this fact, it may seem less surprising that patients on the whole were not very satisfied with their treatment successes. For any type of treatment, the proportion reporting durable improvement was not attained even in one-third of respondents. The lowest success rate was given for psychotherapy with only 14% of respondents reporting lasting success (compared to 29% for drugs!). Of course, these results are limited by the fact that they are based on the patients' classification of the treatment received. However, assessment of treatments was not simply limited to giving labels such as "behavior therapy" or "client-centered therapy" that may be unknown to many patients. Instead we gave brief descriptions of various treatment methods (e.g., "exposing yourself systematically to feared situations or objects"). In any case, this study makes it possible to generalize from the findings from patients seeking treatment to the general public.

6.02.5.3 Current Criticisms of Behavior Therapy

In addition to the problems caused by misunderstandings, misapplications, or the lack of application, there are several criticisms that apply to behavior therapy in its current form. In my opinion, behavior therapy at least partly has not addressed some of these criticisms well enough, which should make its proponents more modest.

(i) Theory and practice frequently do not overlap very well (Eysenck & Martin, 1987; O'Donoghue & Krasner, 1995). For instance, exposure methods have a well-established efficacy in the treatment of phobias, but the theory upon which they are based (two-factor theory, habituation) are outdated in their classical form. In the same vein, functional behavior analysis classifies problematic into operant and respondent types although it is quite clear that most clinical problems are neither one nor the other.

(ii) Because its treatment methods are most often disorder-specific and rely strongly on the empirical knowledge about these disorders, behavior therapy has neglected unspecific or diffuse complaints or problem constellations. The guidelines for dealing with the problem of comorbidity (common occurrence of more than one disorder, which is the rule rather than the exception) are sparse and often unsatisfactory.

(iii) There are hardly any approaches to the problems of counseling, although this forms a large part of the work not only in couseling centers but (to a lesser degree) also in psychotherapy. How should one work through inadequate feelings of guilt after separation, deal with leaving-home situations, educational problems, existential problems, etc.?

(iv) The strong emphasis on change together with the explicit induction of hope for improvement has not only positive consequences because there are limits for change and some situations should not not be changed at all (Wilson, 1996). An example of this problem is behavioral marital therapy. Not all partnerships can be optimally restructured, even in well-functioning partnerships there may be instances where the partners do not agree and that should not be changed. There is a danger of promising too much or things that cannot be attained (see the discussion of therapeutic goals in Section 6.02.2.3).

(v) The "classical" criticism of early behavior therapy, namely the neglect of experience compared to an overemphasis on behavior, does not apply to modern cognitive-behavioral approaches. It is, however, justified for those forms of behavior therapy that have not passed the "cognitive turning-point."

6.02.6 CONCLUSIONS

This representation of the behavioral approach has attempted to give a satisfactory answer to the nature of this therapeutic orientation by describing its past and present and simultaneously trying to stay open to future developments. The close relationship between research and practice has always been a special advantage of the behavioral approach. It is to be hoped that this will not be lost as behavior therapy becomes a more and more professionalized endeavor. In psychotherapy we stand in front of the remarkable fact that we have a far from satisfactory state in practical patient care and at the same time highly effective methods to ameliorate this situation. We should not let ourselves be kept from bridging this gap. Problems such as the lack of application of successful methods, the widespread misallocation of resources, and the strong need for high-quality psychotherapy point to urgent tasks for the health care system:

(i) disseminate knowledge about already available methods;

(ii) ameliorate the conditions for their implementation;

(iii) ameliorate the assessment of mental disorders;

(iv) develop simple interventions for routine care by nonspecialists;

(v) develop prevention programs; and

(vi) routine evaluation and quality assurance.

In dealing with these challenges we need to avoid the two contradictory dangers of trivializing or medicalizing psychological problems. Many psychological problems of a transitory or clearly subclinical nature should not be medicalized, that is, they should not be declared "illnesses" requiring treatment ("for every bad feeling a therapy"). On the other hand, full-blown mental disorders may not be trivialized and thereby be withdrawn from adequate and necessary treatments ("depressed people simply need to get their act together"). As a rule, all concerned parties will only profit from a movement away from ideology and toward a more empirical approach in psychotherapy. Hopefully, the current debates on psychotherapy research will contribute to that process.

6.02.7 REFERENCES

Ayllon, T., & Azrin, N. H. (1968). *The token economy: A motivational system for therapy and rehabilitation*. New York: Appleton-Century-Crofts.

Agras, S., Kazdin, A. E., & Wilson, G. T. (1979). *Behavior therapy: Toward an applied clinical science*. San Francisco: Freeman.

Baer, L., & Minichiello, W. E. (1990). Behavioral treatment for obsessive-compulsive disorder. In R. Noyes, M. Roth, & G. D. Burrows (Eds.), *Handbook of anxiety* (Vol. 4). *The treatment of anxiety*. Amsterdam: Elsevier.

Barlow, D. H., & Hersen, M. (1984). *Single case experimental designs: Strategies for studying behavior change* (2nd ed.). New York: Pergamon.

Beck, A. T. (1967). *Depression: Clinical, experimental and theoretical aspects*. New York: Harper & Row.

Birk, L. (1973). *Biofeedback: Behavioral medicine*. New York: Grune & Stratton.

Davison, G. C. (1968). Systematic desensitization as a counterconditioning process. *Journal of Abnormal Psychology, 73,* 91–99.

Ellis, A. (1962). *Reason and emotion in psychotherapy*. New York: Lyle Stuart.

Eschenröder, C. T. (1994). Über die schriftliche Darstellung der verhaltenstherapeutischen Praxis. *Verhaltenstherapie, 4,* 112–115.

Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology, 16,* 319–324.

Eysenck, H. J. (1959). Learning theory and behaviour therapy. *Journal of Mental Science, 195,* 61–75.

Eysenck, H. J. (Ed.) (1960). *Behaviour therapy and the neuroses*. Oxford, UK: Pergamon.

Eysenck, H. J., & Martin, I. (Eds.) (1987). *Theoretical foundations of behaviour therapy*. New York: Plenum.

Fahrenberg, J. (1987). Zur psychophysiologischen Methodik: Konvergenz, Fraktionerung oder Synergismen? Aktivierungsforschung. *Diagnostica, 33,* 272–287.

Falloon, I. R., Boyd, J. L., & McGill, C. W. (1984). *Family care of schizophrenia*. New York: Guilford Press.

Ferster, C. A. (1965). Classification of behavioral pathology. In L. Krasner & L. Ullman (Eds.), *Research in behavior modification*. New York: Holt, Rinehart and Winston.

Foa, E. B., & Emmelkamp, P. M. G. (1983). *Failures in behavior therapy*. New York: Wiley.

Franks, C. M., & Wilson, G. T. (1975). *Annual review of behavior therapy: Theory and practice* (Vol. 3). New York: Brunner/Mazel.

Grawe, K. (1992). Psychotherapieforschung zu Beginn der neunziger Jahre. *Psychologische Rundschau, 43,* 132–162.

Grawe, K., Donati, R., & Bernauer, F. (1994). *Psychotherapie im Wandel. Von der Konfession zur Profession.* Göttingen: Hogrefe.

Hahlweg, K., Dürr, H., & Müller, U. (1994). *Familienbetreuung schizophrener Patienten.* Weinheim: Beltz, Psychologie-Verlags-Union.

Heekerenz, H. P. (1991). *Zur sozialen Akzeptanz der Verhaltenstherapie. Verhaltensmodifikation und Verhaltensmedizin, 12,* 221–236.

Hodgson, R., & Rachman, S. (1974). Desynchrony in measures of fear. *Behaviour Research and Therapy, 12,* 319–326.

Hollandsworth, J. G. (1986). *Physiology and behavior therapy. Conceptual guidelines for the clinican.* New York: Plenum.

Jacobson, E. (1938). *Progressive relaxation.* Chicago: University of Chicago Press.

Jones, H. G. (1956). The application conditioning and learning techniques to the treatment of a psychiatric patient. *Journal of Abnormal and Social Psychology, 52,* 414–420.

Jones, H. G. (1960). Continuation of Yates' treatment of a tiqueur. In H. J. Eysenck (Ed.), *Behaviour therapy and the neuroses.* Oxford, UK: Pergamon.

Jones, M. C. (1924a). The elimination of children's fears. *Journal of Experimental Psychology, 7,* 382–390.

Jones, M. C. (1924b). A laboratory study of fear: The case of Peter. *Pediatric Seminars, 31,* 308–315.

Kazdin, A. E. (1978). *History of behavior modification.* Baltimore: University Park Press.

Kazdin, A. E., & Wilcoxon, L. A. (1976). Systematic desensitization and non-specific treatment effects: A methodological evaluation. *Psychological Bulletin, 83,* 729–758.

Lang, P. (1971). The application of psychophysiological methods to the study of psychotherapy and behavior change. In A. E. Bergin, & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change. An empirical analysis.* New York: Wiley.

Lang, P. (1993). The three-system approach to emotion. In N. Birbaumer & A. Öhman (Eds.), *The structure of emotion.* Seattle, WA: Hogrefe & Huber.

Lazarus, A. A. (1958). New methods in psychotherapy: A case study. *South African Medical Journal, 32,* 660–664.

Lazarus, A. A. (1971). *Behavior therapy and beyond.* New York: McGraw-Hill.

Lewinsohn, P. H. (1974). A behavioral approach to depression. In R. J. Friedman & M. M. Katz (Eds.), *The psychology of depression: Contemporary theory and research.* Washington, DC: Winston-Wiley.

Lutz, R., Bezold, G., Bloem, R., Dietrich, M., & Wittmann, L. (1992). Sehen und gesehen werden. Neue Psychologien mit den alten Philosophien? Gespräche mit erfahrenen VerhaltenstherapeutInnen über Therapieziele und Prognosen. In H. Lieb & R. Lutz (Eds.), *Verhaltenstherapie. Ihre Entwicklung—ihr Menschenbild.* Göttingen: Verlag für Angewandte Psychologie.

Margraf, J. (1996). Grundprinzipien und historische Entwicklung. In J. Margraf (Ed.), *Lehrbuch der Verhaltenstherapie* (Vol. 1). Berlin: Springer.

Margraf, J., & Lieb, R. (1994). Was ist Verhaltenstherapie? Versuch einer zukunftsoffenen Neucharakterisierung. Editorial. *Zeitschrift für Klinische Psychologie, 24,* 1–7.

Margraf, J., & Poldrack, A. (in press). Angstsyndrome in Ost- und Westdeutschland. Eine repräsentative Bevölkerungserhebung. Submitted for publication.

Masters, W. H., & Johnson, V. E. (1970). *Human sexual inadequacy.* Boston: Little, Brown.

Meichenbaum, D. (1975). Self-instructional methods. In F. H. Kanfer & A. P. Goldstein (Eds.), *Helping people change: A textbook of methods.* New York: Pergamon Press.

Meyer, V. (1957). The treatment of two phobic patients on the basis of learning principles. *Journal of Abnormal and Social Psychology, 55,* 261–166.

Mowrer, O. H. (1947). On the dual nature of learning—a reinterpretation of "conditioning" and "problem-solving." *Harvard Educational Review, 17,* 102–148.

Mowrer, O. H., & Mowrer, W. M. (1938). Enuresis—a method for its study and treatment. *American Journal of Orthopsychiatry, 8,* 436–459.

Nicholson & Berman (1983). Quoted in Shapiro, D. (1985). Recent applications of meta-analysis in clinical research. *Clinical Psychology Review, 3,* 13–34.

O'Donahue, W., & Krasner, L. (Eds.) (1995). *Theories of behavior therapy. Exploring behavior change.* Washington, DC: American Psychological Association.

Paul, G. L. (1966). *Insight versus desensitization in psychotherapy.* Stanford, CA: Stanford University Press.

Paul, G. L. (1967). Strategy of outcome research in psychotherapy. *Journal of Consulting Psychology, 31,* 109–118.

Rachman, S. (1988). Verhaltenstherapie. In W. Arnold, H. J. Eysenck, & R. Meili (Eds.), *Lexikon der Psychologie* (6 Vols, pp. 2460–2469). Freiburg, Germany: Herder.

Rachman, S., & Hodgson, R. (1974). Synchrony and desynchrony in fear and avoidance. *Behaviour Research and Therapy, 12,* 311–318.

Rachman, S.J., & Hodgson, R.J. (1980). *Obsessions and compulsions.* Englewood Cliffs, NJ: Prentice-Hall.

Rachman, S., & Teasdale, J. (1969). *Aversion therapy and behavior disorders: An analysis.* Coral Gables, FL: University of Miami Press.

Reinecker, H. (1994). *Grundlagen der Verhaltenstherapie. 2.* Auflage. Weinheim: Psychologie-Verlags-Union.

Schorr, A. (1984). Die Verhaltenstherapie. *Ihre Geschichte von den Anfängen bis zur Gegenwart.* Weinheim: Beltz.

Schorr, A. (1995). Behaviour therapy in Europe. A brief excursion into its history. In K. G. Götestam & E. Ö. Arnarson (Eds.), *Twenty-fifth anniversary of the European Association for Behavioural and Cognitive Therapies (EABCT)—History.* EABCT.

Seidenstücker, G., & Baumann, U. (1987). Multimodale Diagnostik als Standard in der Klinischen Psychologie. *Diagnostica, 33, 243–258.*

Shapiro, M. B. (1961). The single case in fundamental psychological research. *British Journal of Medical Psychology, 34,* 255–262.

Shapiro, D., & Shapiro, D. (1980). Meta-analysis of comparative therapy outcome research: A critical appraisal. *Behavioural Psychotherapy, 10,* 4–25.

Smith, M. L., Glass, G. V., & Miller, T. I. (1980). *The benefits of psychotherapy.* Baltimore: Johns Hopkins University Press.

Task Force on Promotion and Dissemination of Psychological Procedures (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist, 48,* 3–23.

Task Force on Promotion and Dissemination of Psychological Procedures (1998). An update on empirically validated treatments. Washington, DC: Division of Clinical Psychology, American Psychological Association.

Taylor, C. B., King, R., Margraf, J., Ehlers, A., Telch, M., Roth, W. T., & Agras, W. S. (1989). Use of medication and *in vivo* exposure in volunteers with panic disorder. *American Journal of Psychiatry, 146, 1423–1426.*

Westmeyer, H. (1984). Von der Schwierigkeit, ein Behaviorist zu sein oder Auf der Suche nach einer behavioristischen Identität. In H. Lenk (Ed.), *Handlungstheorien – interdiszilinär* (Band 3). Munich, Germany: Fink.

Wilson, G. T. (1996). Acceptance and change in the treatment of eating disorders and obesity. *Behaviour Therapy, 27,* 417–439.

Wittmann, W. W., & Matt, G. E. (1986). Meta-Analyse als Integration von Forschungsergebnissen am Beispiel deutschsprachifer Arbeiten zur Effektivitat von Psychotherapie. *Psychologische Rundschau, 37,* 20–40.

Wolpe, J. (1954). Reciprocal inhibition as the main basis of psychotherapeutic effects. *Archives of Neurological Psychiatry, 72,* 205–226.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition.* Stanford, CA: Stanford University Press.

Wolpe, J. (1976). Behavior therapy and its malcontents. 1. Denials of its bases and psychodynamic fusionism. *Journal of Behavior Therapy and Experimental Psychiatry, 7,* 1–5.

Woolfolk, A. E., Woolfolk, R. L., & Wilson, G. T. (1977). A rose by any other name . . . : Labeling bias and attitudes toward behavior modification. *Journal of Consulting and Clinical Psychology, 45,* 184–191.

Yates, A. J. (1970). *Behavior therapy.* New York: Wiley.

Yates, A. J. (1977). Falsche Auffassungen über die Verhaltenstherapie: Ein Standpunkt. In H. Westmeyer & N. Hoffmann (Eds.), *Verhaltenstherapie: Grundlegende Texte.* Hamburg: Hoffmann & Campe.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.03
Cognitive Therapy

IVY-MARIE BLACKBURN
Cognitive Therapy Centre, Saint Nicholas Hospital, Newcastle upon Tyne, UK and University of Durham, UK

6.03.1 INTRODUCTION 51
6.03.2 PRINCIPLES OF COGNITIVE THEORY OF EMOTION 52
6.03.2.1 A.T. Beck (Early Model) 52
6.03.2.2 The Evolution of Beck's Model 53
6.03.2.2.1 Personality variables 54
6.03.2.3 Interpersonal Processes 56
6.03.2.4 Constructivism 56
6.03.2.5 Information Processing and Cognitive Science 57
6.03.2.6 Other Cognitive Approaches 60
6.03.2.6.1 Rational-emotive therapy 60
6.03.2.6.2 Self-instructional training 62
6.03.2.6.3 Hopelessness theory of depression 62
6.03.3 APPLICATION OF COGNITIVE THERAPY 63
6.03.3.1 General Principles of Cognitive Therapy 63
6.03.3.2 Treatment Methods 67
6.03.3.2.1 Behavioral methods 67
6.03.3.2.2 Cognitive methods 68
6.03.3.3 Developments 75
6.03.4 PROCESS OF COGNITIVE THERAPY 76
6.03.4.1 Mediators of Change in Cognitive Therapy 76
6.03.4.2 Moderators of Change in Cognitive Therapy 78
6.03.5 EFFICACY 78
6.03.6 CONCLUSION 79
6.03.7 REFERENCES 79

6.03.1 INTRODUCTION

Nearly 40 years have elapsed since the early work of Beck and Ellis in the late 1950s and early 1960s in cognitive approaches to the treatment of emotional disorders. Both were reacting against classical psychoanalysis in which they had been trained and thus started what has been termed "The cognitive revolution" (Dember, 1974) which, however, is probably best conceptualized as an evolution from the two embattled camps of psychoanalysis and behaviorism (Blackburn, 1986; Salkovskis, 1986).

As with all "new" ideas, the cognitive movement did not arise *de novo*. Several influences within the general field of psychology and from clinical practice have been acknowledged, in particular Piaget (1972) who described the hierarchical structuring of knowledge; the ego analysts (e.g., Adler, 1936; Horney, 1950) who stressed patterns of thinking regarding the self and the world as central to understanding feelings and behavior; Kelly (1955) who was a real ancestor in his descriptions of personal constructs; Arnold (1960) and Lazarus (1966) who stressed the role of cognition in their theories of emotion.

Be it revolution or evolution, it can certainly be asserted that cognitive therapy has been a most successful movement, in view of the highly prolific research literature, the number of specialized journals, the demands for training courses in cognitive therapy, and continuous refinement of the underlying theoretical principles. The strength of cognitive therapy from its inception has been its emphasis on empirical backing. This tradition continues, both in experimental studies and in clinical outcome studies, and, consequently, the theory has not remained static, being also informed by advances in related areas of study, namely cognitive science, social science, and biology.

This chapter will consider the key principles of the cognitive theory of emotion and how they apply to cognitive treatments. The range of approaches and recent developments will be discussed.

6.03.2 PRINCIPLES OF COGNITIVE THEORY OF EMOTION

6.03.2.1 A.T. Beck (Early Model)

Beck's early work (1963, 1967), related to thinking style in depression and later expanded to other emotional disorders (1976). He described the typical negative content of thought in depression as a pervasive negative view of the self, of the environment, and of the future (the negative cognitive triad), expressed in automatic, habitual thoughts in reaction to trigger stimuli. The negative automatic thoughts are maintained by various processing errors with a negative bias (arbitrary inferences, selective abstractions, personalizations, overgeneralizations, minimizations, and magnifications). These processing errors do not differ necessarily from the type of processing errors made by nondepressed individuals (labeled heuristics by Kahnneman, Slovic, & Tversky, 1982), the difference being in the direction of the bias, which is usually positive in the nondepressed, expressed as a self-serving bias (Taylor & Brown, 1988). The consequence of processing information with a negative bias is a congruent negative emotion. Thus, the information processing model of emotional disorders was set in a straightforward vertical or unidirectional model, which has since been reviewed. To explain why some individuals process information with such unhelpful biases, Beck evoked the concept of basic structures or schemata derived from cognitive science (Bartlett, 1932; Neisser, 1976; Piaget, 1950) These represent the sum of previous experiences, serving as templates that direct attention, influence encoding and interpretation of stimuli, and facilitate recall. In depression, the schemata reflect themes of loss and of self deficiency, for example, "unless I do everything perfectly, I am worthless" or "unless I do everything to please others, they will reject me." It is noteworthy that at this stage of the theory, no differentiation was made between conditional schemata or basic assumptions, rules, and attitudes on the one hand, and unconditional schemata or core beliefs on the other, for example, "I am inadequate" or "I am unlovable." An example linking the three elements described above would be:

Schema: (conditional)
"If people do not like me, I cannot be happy"

Stimulus:
"A friend does not telephone as promised"

↓

Interpretation: (Content of thought expressed in negative automatic thoughts)

"She does not like me"	(*arbitrary inference personalization*)
"Nobody likes me"	(*magnification*)
"I shall always be alone and miserable"	(*overgeneralization*)

↓

Feeling:
"Down in the dumps"

Once the negative feeling is triggered, it is likely to feed forward and lead to other stimuli being interpreted negatively in accordance with the schema which has been activated and which becomes progressively more widely applied to inappropriate stimuli. Similarly, maladaptive behaviors, such as ruminations and inactivity, become preponderant and increase the incidence of negative automatic thoughts and the corresponding dysthymic moods. The typical schemata, beliefs, rules, and attitudes reflect themes of love, approval, entitlement, omnipotence, perfectionism, autonomy, and achievement, as measured by the Dysfunctional Attitude Scale (DAS) which was developed specially to assess this aspect of Beck's theory (Weissman & Beck, 1978)

The cognitive model of the anxiety disorders (Beck, 1976; Beck & Emery, 1985) followed the same principles, emphasizing specific patterns of thinking which differentiate these disorders from depression. Beck, Laude, and Bohnert (1974), in two studies designed to elicit thoughts and visual imagery associated with anxiety, found that anxious patients experience threatening thoughts or images which often precede attacks of anxiety. These thoughts relate to anticipated or visualized danger and extreme vulnerability. It was apparent that imagery was important in anxiety, an aspect that had not been stressed in depression, but which is emphasized in cognitive therapy (Edwards, 1989) The content of thought in anxiety was found to relate to an anxiogenic triad; seeing the world as threatening, the self as vulnerable, and the future as uncontrollable. This approach was similar to that of Lazarus (1966) who distinguished between two cognitive processes in anxiety, namely primary and secondary appraisal. Primary appraisal relates to an individual's evaluation of a situation as dangerous and threatening, whereas secondary appraisal relates to the evaluation of the self as not having the internal and/or external resources to deal with that situation.

The processing errors in anxiety do not differ essentially from those described in depression, but the schemata are likely to be different, referring to themes of personal vulnerability, of unpredictability, and of threat, for example, "If I feel anxious, this means I have no control of myself" or "I must always be on my guard, if not something awful will happen."

As in depression, the dysfunctional schemata are self- or world-referent; they tend to be rigid and undifferentiated. These characteristics have been taken as implying that they have been learnt in early childhood and continue to survive through the process of assimilation, rather than accommodation (Piaget, 1977).

6.03.2.2 The Evolution of Beck's Model

The implications of the original model were of a causative link between cognitions and emotions, assigning primacy to cognitions. This was challenged by Zajonc (1980) and Rachman (1981, 1984). Beck (1987) restated his approach more clearly, emphasizing that negative cognitions are inherent to depression, as delusions are inherent to psychosis, and that, therefore, they cannot be conceptualized as causing depression, as delusions cannot be said to cause psychosis. Negative cognitions are one side of the coin and biological changes the other side. While research findings have given ample evidence for a negative content of thought in depression (Haaga, Dyck, & Ernst, 1991) and thoughts relating to personal danger and inability to cope in anxiety (Hibbert, 1984; Butler & Mathews, 1983), the causative role of cognitions, and the specificity of cognitive therapy, relative to antidepressants medication, in changing cognitions have not been demonstrated (Blackburn & Bishop, 1983; Simons, Garfield, & Murphy, 1984).

Regarding particular biases in information processing, in general, studies support an attentional bias in anxiety and a memory bias in depression (Dalgleish & Watts, 1990; Mogg et al., 1991; Wells & Matthews, 1994; Williams, Watts, MacLeod, & Mathews, 1988) Williams et al. (1988) proposed that "anxiety involves biased allocation of attention at the preattentive stage, and depression involves biased use of mnemonic cueing at the elaborative stage" (p. 181). Indeed, several studies have found no bias in memory for threat words in anxious subjects, *inter alia* (MacLeod, 1991; Mineka, 1992; Mogg, Mathews, & Weinman, 1987), while there is substantial evidence of a negative bias in retrieval process in depression (Blaney, 1986; Breslow, Kocsis, & Belkin 1991; Teasdale & Russell, 1983). The extensive research in information processing in depression and the anxiety disorders (reviewed by Wells & Matthews, 1994; Williams et al., 1988) directly inform cognitive therapy as applied to these disorders to date.

The structural aspect of the cognitive model of the emotional disorders has been the most elusive of its components to put to experimental validation. As alluded to above, the term "schema" has been used differently by different authors and a reliable and valid methodology to measure and assess schemata is not well established.

In the conceptualization of cases for the development of cognitive treatment strategies (Persons, 1989), it is usual to differentiate between conditional schemata or basic assumptions and core or unconditional schemata. Thus, in the example given in the previous section, the core schema may be "I am unlovable" leading to the conditional schema "If people do not like me, I cannot be happy" and the rule "I must do everything people ask of me." The same basic or core schema may lead to the conditional schema "If people get close to me and get to know me, they will reject me" leading to the rule "I must avoid close relationships". This differentiation developed from the extension of cognitive therapy to personality disorders (Beck, Freeman, et al., 1990; Young, 1990); from greater emphasis on the role of conceptualization in treatment (Blackburn, Twaddle, et al., 1996); and from the clinical observation that although there seems to be a limited range of themes in basic schemata, there is a wide variety of conditional schemata leading to different behaviors, different ways of constructing the self and the environment, different mood states, and different disorders. The typical themes of core schemata relate to personal worth (e.g., I am worthless/useless), to moral worth (I am bad/evil), to abnormality (I am a freak/abnormal/different), and to personal ability (I am incapable/inadequate). Examples of how these different levels of the structural aspect of cognitive theory are used in case conceptualizations will be given in Section 6.04.3.

The standard assessment tool of underlying cognitive structures remains the DAS, Form A, (Weissman & Beck, 1978), a 40-item, seven-point scales questionnaire which primarily is aimed at measuring the presence and intensity of belief in conditional schemata in depression. A questionnaire by Young (1990) attempts to measure core schemata (Early Maladaptive Schemata in Young's terminology), but the scale has not been validated adequately. Several studies, using the DAS, have shown that level of dysfunctional attitudes discriminates between depressed subjects and controls, *inter alia* (Blackburn, Jones, & Lewin, 1986; Hamilton & Abramson, 1983) and at post-treatment, predicts future levels of depression (Rush, Weissenberger, & Eaves, 1986; Simons, Murphy, Levine, & Wetzel, 1986). However, not surprisingly according to Beck's theory, the level of dysfunctional attitudes decreases markedly with remission and becomes indistinguishable from normal level (Blackburn, Roxborough, Muir, Glabus, & Blackwood, 1990; Simons et al., 1984).

The endorsement of self-referent adjectives is another method of assessing the self-schema (Segal, 1988), which is more in line with the definition of schemata in cognitive psychology. Using this methodology, Teasdale and colleagues (Teasdale, 1988; Teasdale & Dent, 1987) found that subjects vulnerable to depression recall more global self-referent adjectives (e.g., "pathetic," "worthless," "stupid") than never depressed subjects who recall more specific and milder negative self-referent adjectives (e.g., "rude," "thoughtless," "inconsiderate"). The number of global negative adjectives endorsed also predicted level of depression five months later in mildly depressed women. Teasdale (1988) put forward his differential activation hypothesis on the basis of such findings; this states that the type of thinking activated in mild-dysphoric moods, in combination with high neuroticism, a personality variable (Eysenck, 1970) which also affects selective recall of depressed memories, is likely to activate higher levels of depression. These globally negative self-referent adjectives are likely to reflect core schemata. Figure 1 describes the generic model of the emotional disorders, following Beck's approach as developed from the original model, differentiating core schemata and basic assumptions. In Figure 2, the application of the generic model to a specific case is shown in a case formulation, where good as well as depressed periods can be understood through the individual's core schema.

6.03.2.2.1 Personality variables

A further development in the area of vulnerability to depression has been the delineation of higher-order, more stable personality characteristics which may determine the content of dysfunctional schemata. These have been defined as sociotropy and autonomy by Beck and his colleagues (Beck, 1983; Beck, Epstein, & Harrison, 1983). An overly sociotropic individuals is described as socially dependent, highly invested in positive interchange with others, valuing above all acceptance, intimacy, support, and guidance. In contrast, an overly autonomous individual is highly invested in independent functioning, mobility, freedom, choice, achievement, and the integrity of the personal domain.

A specially constructed scale, the sociotropy-autonomy scale (SAS; Beck et al., 1983) consists of 60 items rated on five-point scales. A large number of studies have been conducted (Blackburn, 1996; Clark & Beck, 1991; Gilbert & Reynolds, 1990; Hammen, Ellicott, Gitlin, & Jamison, 1989; Moore & Blackburn, 1994; Robins, 1990) indicating that sociotropy is related to vulnerability to interpersonal life events, to level of depressive symptoms, to level of neuroticism, and to the subscale of the DAS which measures social approval. Moreover, Moore and Blackburn (1996) have shown, in a large clinical sample of depressed patients assessed pre- and post-treatment, that sociotropy, as assessed by the SAS, is relatively stable over time in contrast with level of dysfunctional attitudes as assessed by the DAS. On the other hand, the conceptual validity of autonomy has not been demonstrated. It has been shown to be only minimally related to level of depression (Gilbert & Reynolds, 1994; Moore & Blackburn, 1994, 1996), to correlate poorly with other measures of independence and not to predict consistently an interaction with negative achievement events (Hammen et al., 1989; Robins & Block, 1988). The autonomy scale of the SAS is in the process of being revised for new validation studies (Clark, Steer, Beck, & Ross, 1995).

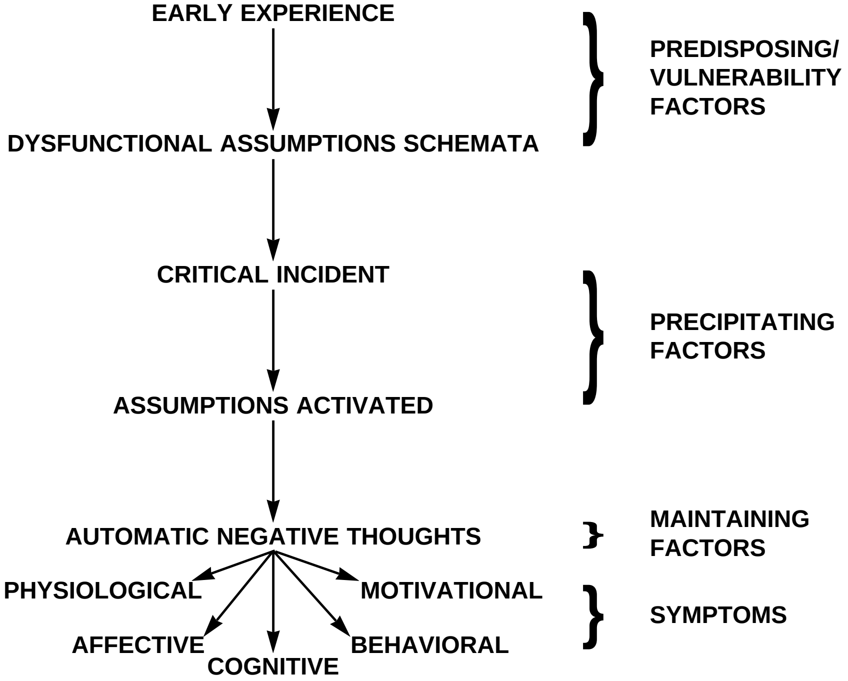
EARLY EXPERIENCE
DYSFUNCTIONAL ASSUMPTIONS SCHEMATA
PREDISPOSING/ VULNERABILITY FACTORS
CRITICAL INCIDENT
ASSUMPTIONS ACTIVATED
PRECIPITATING FACTORS
AUTOMATIC NEGATIVE THOUGHTS
MAINTAINING FACTORS
PHYSIOLOGICAL
MOTIVATIONAL
AFFECTIVE
COGNITIVE
BEHAVIORAL
SYMPTOMS

Figure 1 Cognitive model of emotional disorders.

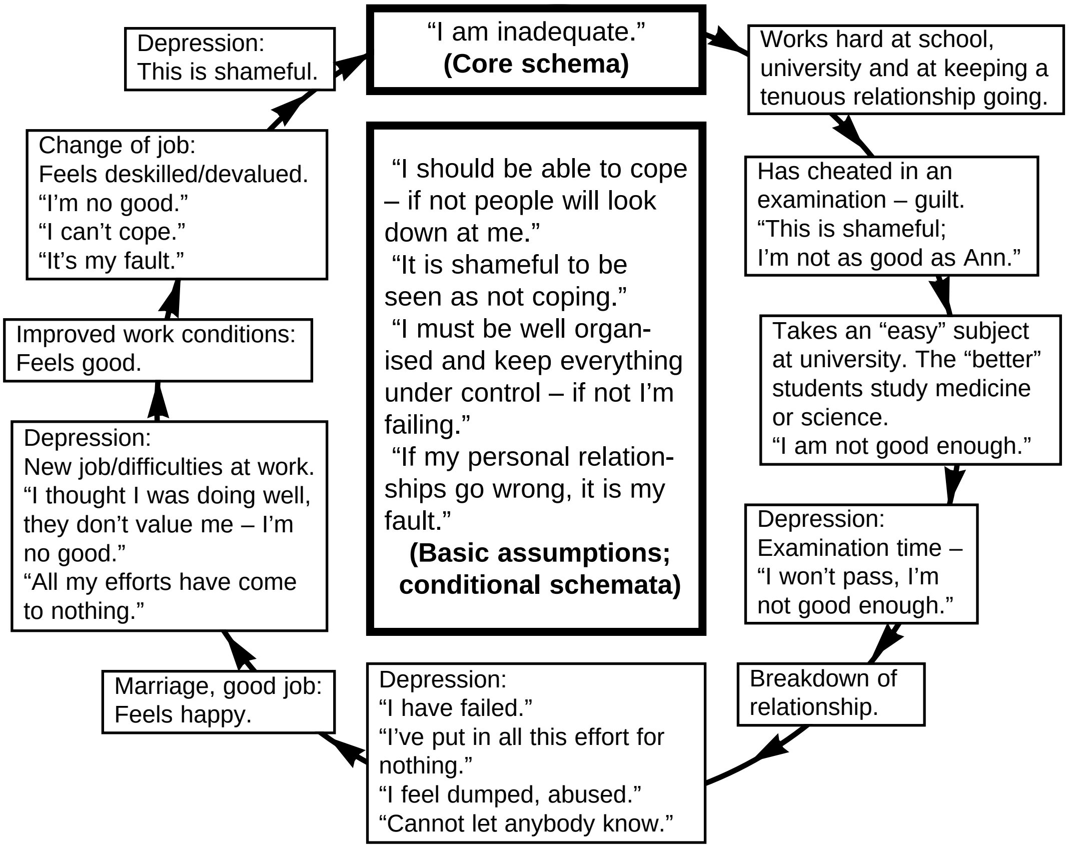
"I am inadequate."
(Core schema)
Works hard at school, university and at keeping a tenuous relationship going.
Has cheated in an examination – guilt.
"This is shameful;
I'm not as good as Ann."
Takes an "easy" subject at university. The "better" students study medicine or science.
"I am not good enough."
Depression:
Examination time –
"I won't pass, I'm not good enough."
Breakdown of relationship.
Depression:
"I have failed."
"I've put in all this effort for nothing."
"I feel dumped, abused."
"Cannot let anybody know."
Marriage, good job:
Feels happy.
Depression:
New job/difficulties at work.
"I thought I was doing well, they don't value me – I'm no good."
"All my efforts have come to nothing."
Improved work conditions:
Feels good.
Change of job:
Feels deskilled/devalued.
"I'm no good."
"I can't cope."
"It's my fault."
Depression:
This is shameful.
"I should be able to cope – if not people will look down at me."
"It is shameful to be seen as not coping."
"I must be well organised and keep everything under control – if not I'm failing."
"If my personal relationships go wrong, it is my fault."
(Basic assumptions; conditional schemata)

Figure 2 Depression conceptualization (reproduced by permission of Souvenir Press from *Cognitive therapy in action. A practitioner's case-book*. Blackburn, Twaddle et al., 1996).

6.03.2.3 Interpersonal Processes

It may be that the relationship of sociotropy, and not autonomy, to depression simply reflects that interpersonal processes are of particular importance in depression, and possibly in other emotional disorders. Early descriptions of cognitive therapy were criticized for ignoring interpersonal processes (Coyne & Gotlib, 1983). Safran and Segal (1990) have made a welcome contribution emphasizing the role of interpersonal factors. They state "the individual must always be understood as part of the interpersonal systems in which he or she is participating. Thus, one cannot understand the patients in therapy independently of the therapist" (p. 5). They view the therapeutic relationship as central to cognitive therapy, in that it is through the therapeutic relationship that core dysfunctional interpersonal schemata can be revealed. The interpersonal schema is defined as a generic cognitive representation of interpersonal events which derives from interactions with attachment figures of the past and serves to predict future interactions with these figures and with others. The stress on the interpersonal perspective has exercised a strong influence on the practice of cognitive therapy, in that more emphasis is put on the therapeutic relationship and more attention is paid to developmental factors, especially in the treatment of personality disorders. Issues which might have been considered as stumbling blocks in therapy, for example, lack of trust, overdependence, avoidance, and aggression, can now be directly employed to guide formulation and questioning. For example, if the patient appears to be aggressive in therapy, the therapist may comment: "It seems to me that you feel angry with me. Is this how you feel at the moment?" "Is it something that I have said or something about me?" "Is there some other situation like this outside of therapy that made you feel the same?" "Did this happen in your family?" Thus, the typical socratic questioning or guided discovery style of cognitive therapy is used to uncover possible core interpersonal schemata, rather than, perhaps more traditionally, just to elicit the current automatic thoughts so as to revaluate them and to look for disconfirming evidence.

6.03.2.4 Constructivism

Mahoney (1993) suggests that one of the "major conceptual developments in the cognitive psychotherapies over the past three decades" has been the differentiation of rationalist and constructivist approaches to cognition. Constructivism or constructive metatheory emphasizes the active and proactive nature of all knowing, the operation of tacit or unconscious processes, the complexity of human experiences, and the need for a developmental, process-focused approach to knowing. Some of the proponents of this approach, apart from Mahoney, are Neimeyer (1992, 1993), Guidano and Liotti (1983), and Anderson (1990) who are all indebted to Kelly's (1955) personal construct theory. Constructivism is contrasted by these authors to logical positivism and rationalism which are described as characterizing Beck, Rush, Shaw, and Emery's (1979) information processing model and Ellis' (1962) rational-emotive approach. Neimeyer (1993) considers constructivism to be predicated on a postmodern epistemology and gives a quote from Kelly (1977, p. 6) "what we think we know is anchored only in our assumptions, not in the bedrock of truth itself, and that world we seek to understand remains always on the horizons of our thoughts," as representing a quintessentially postmodern conclusion.

Neimeyer (1993) contrasts traditional cognitive theory and therapy with constructivist theory and therapy. In so doing, he gives a distorted and narrow view of the Beckian approach, at least in its modern form. He interprets Beck et al. (1979) and Ellis (1973) in particular, as equating emotional adjustment with rationality (logical empiricism), so that when humans deviate from rationality and the scientific method, they form distorted views of themselves, their world, and their future, and experience dysphoric emotions and emotional disorders. The word "rational" may have been emphasized in earlier cognitive theory, but it is not now; the term "dysfunctional" being preferred, indicating that thoughts or attitudes are dysfunctional when they have negative consequences, not because they are irrational. These changes in the theory have followed research findings that indicate, for example, that depressed subjects do not have exclusive negative thoughts (Clifford & Hemsley, 1987), that depressed subjects choose as many positive as negative adjectives as self-descriptors (Derry & Kuiper, 1981), and that normal controls show a ratio of 1:17 of negative to positive thoughts. (Schwartz, 1986). Ellis (1993) has argued that even rational-emotive therapy (RET) has never been truly rationalistic, although he accepts that he too has moved from a more rationalistic to a more constructivist position. Cognitive therapy (Beck et al., 1979) has, in any case, always used questioning that stresses the disadvantages of thinking in certain ways and the advantages of thinking in alternative ways, rather than logical vs. illogical thinking.

However, it must be acknowledged that a number of the changes in cognitive therapy, as it has evolved, have been influenced by the philosophical stance of constructivist theories. The examination of developmental factors to understand the cognitive structures through which an individual construes their world, the stress on interpersonal processes, and developments in the understanding of information-processing to explain individual emotional reaction, as described in the next section, probably reflect constructivist influences.

6.03.2.5 Information Processing and Cognitive Science

Cross-fertilization from cognitive science has inspired a vast number of ingenious experiments to test various aspects of information processing in depression and anxiety disorders, and has prompted more complex models of information processing (Ingram, 1988; Teasdale, 1993; Teasdale & Barnard, 1993; Wells & Matthews, 1994; Williams et al., 1988). Ingram and Kendall (1988, p. 13) present a complex model which is adapted in Figure 3.

These authors attempt to delineate the "pattern of relationships among the various components of the cognitive taxonomic system." The diagram indicates that the different components which may be involved in information processing operate in a reciprocal, interactional, and multidirectional fashion. Cognitive structure is defined as the architecture of the system in which information is stored and organized. This would consist of long- and short-term memories, sensory memories, and associative networks. Cognitive content represents the information that is actually stored, that is, the content of the cognitive structure, for example, semantic and episodic memory. The term schema, as used in cognitive therapy, is represented by both the cognitive structure and the cognitive proposition. Cognitive operations represent the various procedures by which information is processed, for example, attention, selection, encoding, interpretation, and retrieval. These represent the various cognitive processes which have been found to show particular biases in the emotional disorders (attentional biases in anxiety and retrieval biases in depression). Cognitive products represent thoughts, both automatic and controlled, and images. Arrows labeled as operations indicate cognitive processes leading to cognitive products, whereas arrows not labeled as operations indicate "procedures that may not be primarily cognitive," for example, the interactions between affective, biochemical, and cognitive factors.

The model suggests that behavior can be affected by a variety of sources, by cognitions, by cognitive structures directly (e.g., unconscious routines such as riding a bike), and by affective and biochemical factors. The acknowledgement of other cognitive structures, which remain undefined, recognizes that understanding of the cognitive system is still incomplete, particularly as regards unconscious structures. The model also suggests a generic model for the emotional disorders which can take into account similarities and differences across different disorders. For example, in depression and anxiety, the operation through which information is processed may be largely similar, but the content of the cognitive structures and the cognitive products differ as described in Section 6.03.2.2. Such a model would evidently be far too complex to use in therapy. Instead, as shown in Figure 4, a simplified generic model giving the same information can be given to patients (Greenberger & Padesky, 1995). The diagram indicates the interactions among all elements; biology (which includes physiological and biochemical changes), thoughts (automatic thoughts, beliefs, attitudes, schemata), emotions, and behavior and stresses that the individual exists within an environment which triggers their reactions and which reacts to their reactions.

Bower (1981) presented his influential associative network theory of mood and memory which derived experimentally to explain the relationships between cognition and mood. According to this theory, emotions, as well as cognitions and events, are represented in the brain/mind by discrete nodes or units which are linked in an associative network. Thus, an emotion can be activated by an appropriate external stimulus or by the activation of other linked nodes in the associative network, for example, a sad memory or some physical sensation. Once activated, a depressed mood or an anxious mood will influence future information processing by the spreading of activation through the associative network, so that events or situations are more likely to be interpreted negatively or as denoting threat and danger. The prediction from Bower's (1981) paper was of a recriprocal relationship between mood and thinking, so that there would be mood state-dependant retrieval (superior recall of material retrieved in the same mood state as was present during learning) and mood-congruent retrieval (increased ease of recall of sad memories when in a depressed mood state). These predictions have been confirmed in several studies using both naturally occurring depressed moods (Clark & Teasdale, 1982; Miranda & Persons, 1988; Miranda, Persons, &

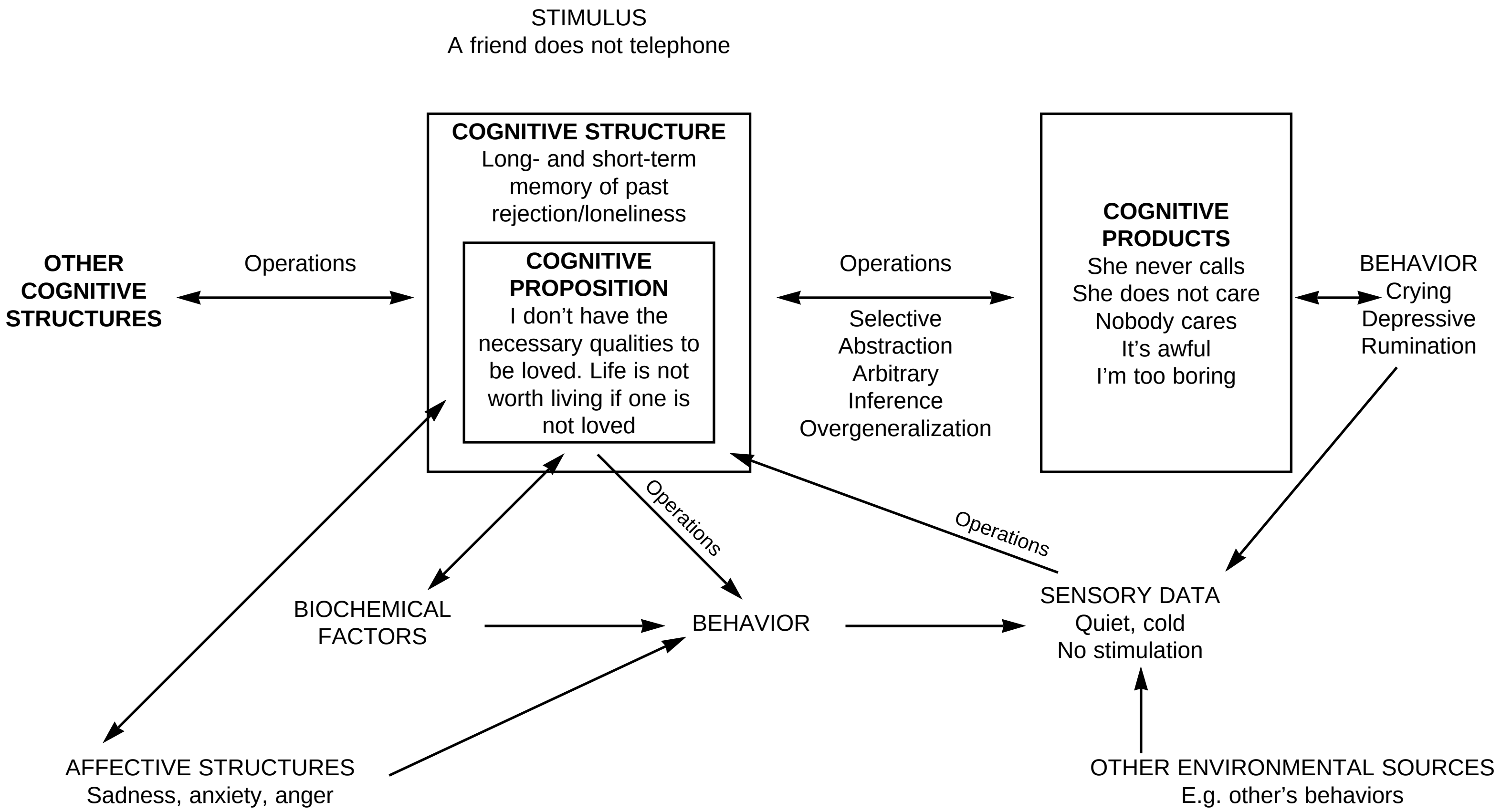
STIMULUS
A friend does not telephone
COGNITIVE STRUCTURE
Long- and short-term memory of past rejection/loneliness
COGNITIVE PROPOSITION
I don't have the necessary qualities to be loved. Life is not worth living if one is not loved
OTHER COGNITIVE STRUCTURES
Operations
Operations
Selective Abstraction
Arbitrary Inference
Overgeneralization
COGNITIVE PRODUCTS
She never calls
She does not care
Nobody cares
It's awful
I'm too boring
BEHAVIOR
Crying
Depressive Rumination
Operations
Operations
BIOCHEMICAL FACTORS
BEHAVIOR
SENSORY DATA
Quiet, cold
No stimulation
AFFECTIVE STRUCTURES
Sadness, anxiety, anger
OTHER ENVIRONMENTAL SOURCES
E.g. other's behaviors

Figure 3 Information processing model and depression (Source: Ingram and Kendall, 1988).

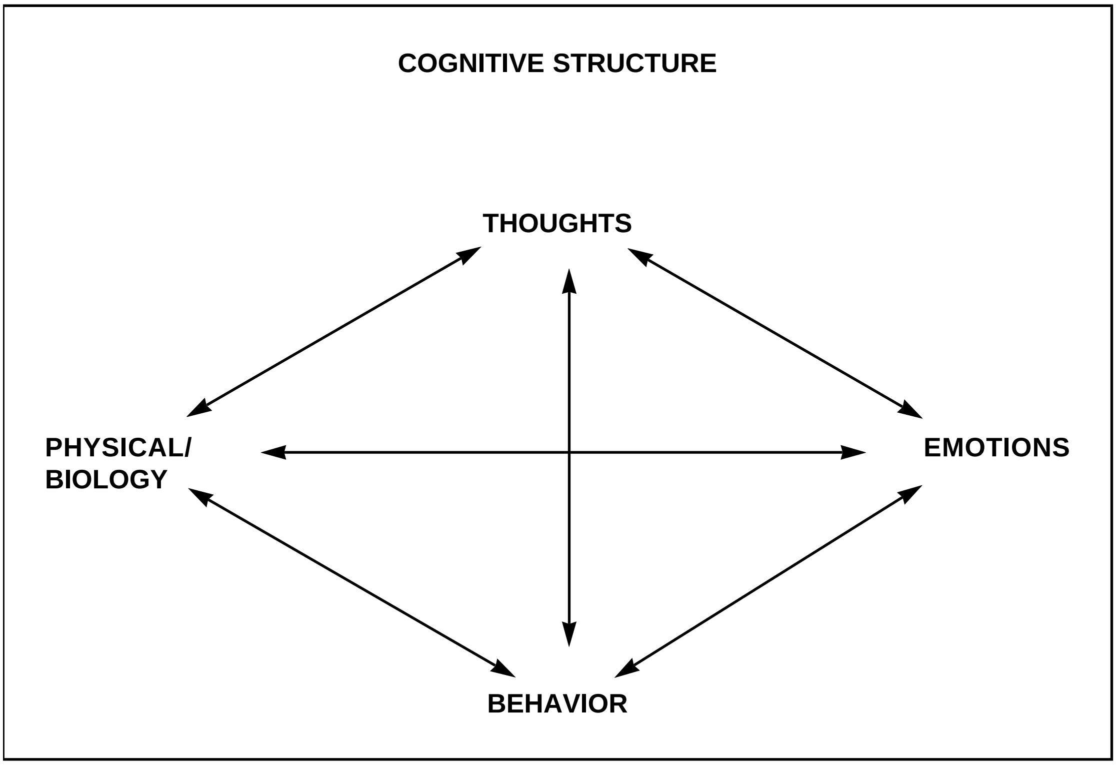

Figure 4 Generic cognitive model of the emotional disorder (Source: Greenberger & Padesky, 1995).

Buyers, 1990) and in induced moods (Teasdale & Fogarty, 1979; Teasdale & Taylor, 1981). However, a number of disconfirming findings have also been reported, as seen in meta-analyses and Ucros (1989) and Bower (1987) concluded that his theory needed revising. Bower (1992) has expanded his theory by proposing that emotions may activate not only isolated semantic concepts, but wider rule-based action plans.

Teasdale (1993) finds Bower's theory an improvement on Beck's theory, for several reasons: its clearly testable predictions; its acknowledgement that negative thinking can be an antecedent, as well as a consequence of depressed affect; its view of depression as an extreme form of the normal effect of mood on information processing, which thus provides an alternative view of cognitive vulnerability without having to evoke the matching of trigger events with dysfunctinal assumptions and schemata; and its ability to explain the maintenance of depression. In short, he considers Bower's theory more scientifically sound than Beck's. However, Teasdale (1993) and Teasdale and Barnard (1993) criticize Bower's theory on several counts and have put forward an alternative theory, the Interacting Cognitive Subsystems (ICS) framework which will be described briefly because of its potential influence on cognitive therapy.

ICS aims to bring to clinical psychology insights derived from cognitive science. Teasdale (1993) considers that Beck's model depends on lay concepts of cognition where cognitions refer only to consciously experienced thoughts and images and schemata to beliefs and assumptions. He considers that research studies and clinical observations have higlighted certain problems with cognitive therapy: cognitive therapy is not uniquely more effective than other psychological treatments (e.g., behavior therapy, Gallagher & Thompson, 1982; Interpersonal Psychotherapy, Elkin et al., 1989); antidepressant medication reduces negative thinking to the same extent as cognitive therapy (Simons et al., 1984); there is no evidence of dysfunctional assumptions in remitted patients (Blackburn et al., 1990); in therapy, it sometimes appears that patients can experience emotional reactions without identifiable negative automatic thoughts; rational argument in therapy often appears ineffective in changing emotional responses; and for more refractory problems, for example, personality disorders, noncognitive methods of treatment have had to be borrowed from other schools of therapy, for example, Gestalt therapy (Beck, Freeman et al., 1990).

On the other hand, Bower's (1981) model, according to Teasdale (1993), fails because of the lack of differentiation in the semantic networks which are described. These appear to represent knowledge in a single format, so that specific concepts and the relationship between them is not differentiated. Higher order of meaning is not explained and, therefore, no distinction can be made between "hot" cognitions which are associated with a high level of emotion and "cold" cognitions which are reported at an intellectual level, without accompanying emotions.

Teasdale and Barnard (1993) describe ICS as an information processing model which attempts to take into account *all* aspects of information processing, and thus, probably succeed in developing in more detail Ingram and Kendall's (1988) model described earlier. They postulate nine types of information, encoded in separate mental codes, representing different aspects of experience; for example, sensory experience (sensory and proprioceptive codes); regularities in the sensory code patterns (intermediate structural description codes); specific and generic levels of meaning (meaning codes); and information required for effector action (effector codes). Each type of information is processed by its own specialized subsystem and encoded in separate memory stores. Information processing involves the transfer of information between subsystems and its transformation from one mental code to another.

Importantly for cognitive therapy, in the "meaning codes" two different codes are described relating to two highly different levels of meaning. The propositional code encodes specific meanings, discrete concepts, and the relationship between them. This is akin to Bower's semantic network, for example, "my neighbour is gray haired." The implicational code encodes generic, holistic levels of meaning which are difficult to convey in language: it relates to recurring very high order regularities across all information codes. Only this level of meaning is linked directly to emotion, with implicit meaning content, for example, "something wrong." Implicational meaning has a high level of abstraction; is influenced by contextual sensory features, for example, facial expression, tone of voice, bodily arousal; and represents generic features of experience which cannot be expressed in single sentences. The implicational code represents schematic models of experience with which mood biases are associated. Thus, depression is associated with the regeneration of depression-related schematic models when there is a shift in the prevailing high-order mental models of self and world which dominate information processing. This shift in the interrelationship between high order mental codes, not the activation of a particular cognitive node, leads to a negative content of thought in the form of negative attributions, evaluations, and memories.

Although this model appears at times overly complex and abstract, the clear differentiation between propositional and implicational levels of meaning is useful to understand the process of therapy and to make sense of failures and successes in therapy. The central goal of therapy is evidently to change implicational meaning related to depressive or anxious schematic models by substituting alternative patterns of meaning related to more adaptive schematic models. Since this abstract, generic level of meaning is made up of various elements, such as specific meanings, bodily state, and sensory information, such as tone and volume of voice and visual information, changes in any of the elements may bring therapeutic change. Therapy would thus involve different elements: change in meanings and/or change in sensory input, adding new elements to an experience through guided imagery, and paying greater attention to feelings through the use of emotional methods of treatment as used in Gestalt therapy. Interestingly, such therapeutic approaches have been adopted in cognitive therapy, especially for personality disorders. It seems that clinicians inspired by clinical insight have reached the same point as cognitive theorists, although through a different route, and perhaps without a proper rationale to explain their methods of treatment.

6.03.2.6 Other Cognitive Approaches

Beck's theories in the original form, and as further developed through research and clinical experience, have been the main emphasis in the previous sections because it is the approach which has led most clearly to specific methods of treatment which have been tested widely in controlled outcome studies. However, some other approaches have also been influential, in particular Ellis' RET, Meichenbaum's self-instructional training, and clinical developments of Seligman's learned helplessness theory by Abramson and her colleagues in the hopelessness theory of depression. These will be described briefly in this section.

6.03.2.6.1 Rational-emotive therapy

Ellis (1989) describes himself as the first major cognitive-behavioral therapist in view of his writings and practice dating back to 1957

(Ellis, 1957, 1962). RET, now relabeled REBT (rational-emotive-behavioral therapy), is based on an "ABC" theory of psychopathology, stating that activating events (A) do not directly cause emotional and behavioral consequences (C), but that the mediating variable of beliefs (B) about these events exercises the major influence on feelings and behavior. As such, the theory is quite close to Beck's, but the emphasis is clearly on beliefs only, as opposed to different aspects of cognition, for example, the content, form, and structure of thought. The therapeutic thrust of RET has, therefore, been on irrational beliefs as the cause of emotional problems. Rational beliefs are defined as those promoting survival and happiness and as likely to find empirical support in the environment, in contrast with irrational beliefs which are unlikely to find empirical support and reflect "musts" and "should." Rational beliefs will lead to appropriate negative emotions (sorrow, annoyance, regret) in the face of losses or difficulties, whereas irrational beliefs will lead to inappropriate negative emotion (depression, extreme anger, guilt). Haaga and Davison (1993) argue that the definition of rationality and irrationality is problematic in that irrationality is not the prerogative of emotionally disturbed individuals. Extensive research has shown that much of the thinking of nondisturbed people is irrational (Alloy & Abramson, 1979; Taylor & Brown, 1988). Irrational beliefs in RET are often defined by lists of typical beliefs which are often encountered in clinical practice (Ellis, 1962), for example:

(i) there is a dire necessity to be loved for everything one does;

(ii) certain acts are awful and wicked and people should be punished who perform such acts;

(iii) it is horrible when things are not the way one would like them;

(iv) human misery is externally caused and is forced on one by outside people and events;

(v) if something is or may be dangerous, one should be terribly upset about it;

(vi) it is easier to avoid than face life's difficulties and self-responsibilities;

(vii) one needs something other or stronger or greater than oneself on which to rely;

(viii) one should be thoroughly competent, intelligent, and achieving in all possible aspects;

(ix) because something once strongly affected one's life, it should indefinitely affect it;

(x) one must have certain and perfect control over things;

(xi) human happiness can be achieved by inertia and inaction; and

(xii) one has no control over one's emotions and one cannot help feeling certain things.

Such beliefs have been described in lists of a dozen to lists of several hundreds specific ones, to a few general categories with many exemplars (Ellis & Bernard, 1985). Specific questionnaires to measure irrational beliefs, such as the Rational Belief Inventory (Shorkey & Whiteman, 1977) and the Irrational Beliefs Test (Jones, 1968) have been criticized for confounding irrational beliefs with negative emotions.

Ellis and Whiteley (1970) make a difference between "elegant" and "inelegant" RET techniques. Elegant or preferential RET involves forceful disputation of the patient's irrational beliefs with the aim of making profound philosophical changes in the patient. However, this approach is not suitable for all patients, for example, for those with limited intellectual resources or with poor motivation. In such cases other methods, as used in cognitive therapy, are acceptable. This is described as inelegant or general RET, involving manipulation of environmental circumstances, teaching coping strategies, and "persuading clients that their perceptions of events are incorrect" (Raitt, 1988, p. 202). Persuasion is, however, definitely not a Beckian method of cognitive therapy which advocates instead the use of inductive methods to guide patients to evaluate their cognitions and make their own discoveries. Ellis (1980), in his comparison of RET and cognitive behavior therapy, emphasizes the philosophical stance of RET which is that: (i) people create their own emotional disturbances through irrational, absolutistic beliefs; (ii) as people have free-will, they can choose to disturb themselves or not; (iii) to bring about change, active work at modifying thoughts, feelings, and behaviors is necessary; (iv) profound philosophical change will bring about modification in emotional and behavioral reactions; (v) long-range hedonism is more healthy than short-term hedonism; and (vi) a scientific outlook, rather than an unscientific religious or mystical outlook, is likely to bring greater emotional health and satisfaction.

He rightly says that such a philosophical stance is absent in cognitive behavioral therapy which emphasizes evaluation and modification of cognitive processes to bring about long-term symptomatic changes, instead of deep philosophical changes.

The efficacy of RET has been tested in a number of studies (Haaga & Davison, 1989; Kendall, 1984; Zettle & Hayes, 1980), with encouraging results. However, the studies have been highly criticized for their poor methodology, involving nonclinical samples, no follow-up evaluation, poor outcome measures, no attempt at measuring treatment adherence, and the lack of differentiation between the two versions of RET (Haaga & Davison, 1993).

6.03.2.6.2 Self-instructional training

Self-instructional training (SIT) is the therapeutic method described by Meichenbaum (1977) in cognitive-behavior modification (CBM). As the term indicates, CBM was a development of behavior therapy and behavior modification at the time when efforts were being made to incorporate cognition as a valid variable in behavior therapy (Mahoney, 1974), because of the growing dissatisfaction with the empirical and theoretical basis of radical behaviorism. Cognitive-behavior therapists proposed that cognitions could be viewed as covert behaviors, subject to the same laws of learning as overt behaviors. Homme (1965) coined the term "coverants" to denote cognitions as covert operants, in Skinner's language, which are responsive to both external and internal contingencies and altered by contiguous pairings, through covert sensitization (Cautela, 1973). Meichenbaum (1993) describes the process as "client's self-statements and images were viewed as discriminative stimuli and as conditioned responses that come to guide and control overt behavior" (p. 202). Following from that, SIT aimed to decondition maladaptive patterns of behavior by training in self-talk which would establish and strengthen new patterns of behavior and by rehearsing adaptive coping skills. "The technology of behavior therapy, such as modelling, mental rehearsal, and contingency manipulations, was used to alter not only clients' overt behaviors, but also their thoughts and feelings" (p. 202) (Meichenbaum, 1993).

Meichenbaum (1977) describes the application of SIT and successful outcomes in hyperactive and impulsive children, socially withdrawn children, adult schizophrenics, and in creativity training. An example of SIT with impulsive children involves the following steps:

(i) an adult performs the task while talking to himself out loud (cognitive modeling);

(ii) the child performs the same task under the direction of the model's instructions (overt, external guidance);

(iii) the child performs the task while instructing himself or herself aloud (overt self-guidance);

(iv) the child whispers to himself while he or she goes through the task (faded, overt self-guidance);

(v) finally, the child performs the task while guiding his or her performance via private speech (covert self-instruction).

Over a number of training sessions, the package of self-statements modeled by the experimenter and rehearsed by the child is enlarged by the means of response chaining and successive approximation procedures. The skills which are taught through modeling are: problem definition, focusing attention and response guidance, self-reinforcement, self-evaluative coping skills, and error correcting options (I'm doing okay ... if I make a mistake, I can correct it and go on more slowly). The same operant conditioning principles of chaining and shaping were used over a number of training sessions to teach schizophrenic patients to use complex sets of self-statements.

The modification of self-talk or of automatic thoughts is central to cognitive therapy of the emotional disorders in adults, but the specific techniques of SIT appear particularly useful for patients with less well-developed verbal skills, for example, children, adults with learning difficulties, and severely impaired individuals such as schizophrenic patients. It has also been shown to be effective as a version of cognitive-behavioral therapy with phobic and anxious patients (Chambless & Gillis, 1993).

6.03.2.6.3 Hopelessness theory of depression

Seligman's theory of learned helplessness (Seligman, 1975; Seligman & Maier, 1967), as a model of depression derived from animal experiments where uncontrollable shocks were administered, was criticized for not being an adequate model for depression in humans. The model did not appear to reflect the complexity of human depression and the varied presentation of depression. As a result, a reformulated learned helplessness model was proposed by Abramson, Seligman, and Teasdale (1978), based on attributional theory, which described a particular type of depression, namely hopelessness depression. In brief, the theory stated that when an individual makes causal attributions which are internal, global, and stable for negative events, but which are external, specific, and unstable for positive events, expectations are created that highly desired outcomes are unlikely to occur or that highly aversive outcomes are likely to occur and that the individual has no available response to change the likelihood of occurrence of these outcomes. These expectations and the occurrence of a negative event were seen as a sufficient proximal cause of depression. However, Abramson et al. (1978) were aware of problems in this analysis, in that not all highly improbable, but highly desired outcomes, will be a sufficient cause of depression. For example, it may be desirable to win several million dollars on the weekly lottery, but people do not generally become depressed if they fail to win. There are other characteristics of the desired outcomes, for example, how much concern exists about them and how possible they are, which may determine whether their nonoccurrence is a sufficient proximal cause of depression.

Abramson, Alloy, and Metalsky (1988) have clarified and revised the helplessness theory of depression. They describe it as a diathesis stress model, relevant to one hypothesized subtype of depression. The hopelessness theory specifies not only a proximal sufficient cause, but also potential distal causes. A negative life event is interpreted by an individual who has a depressogenic attributional style (the diathesis) as due to stable, internal, global factors and high importance is attached to the event. These attributions may be modulated by situational cues (consequences, consistency of occurrence, and distinctiveness). The depressogenic attributions lead to lowered self-esteem and to feelings of hopelessness (which can be increased by social factors, for example, lack of social support) which are seen as sufficient to lead to hopelessness depression. Hopelessness depression is characterized by retarded initiation of voluntary responses (motivation symptom), difficulty in seeing that similar or related outcomes can be controlled (cognitive symptom), and sad affect (emotion symptom). Although the specification of a subtype of depression as hopelessness depression in reaction to negative situations has not led to specific methods of treatment, the theory helps to give a cognitive rationale for the well established finding of the role of life events in depression (Brown & Harris, 1978). It also pinpoints the thinking style which may be of particular importance in situational depression, and which should, therefore, be targeted in cognitive therapy.

6.03.3 APPLICATION OF COGNITIVE THERAPY

6.03.3.1 General Principles of Cognitive Therapy

There are several textbooks which describe cognitive therapy methods in detail for various psychiatric disorders: for depression (Beck et al., 1979; Blackburn & Davidson, 1995; Williams, 1992); for anxiety (Beck, Emery, & Greenberg, 1985; Blackburn & Davidson, 1995); for personality disorders (Beck, Freeman et al., 1990; Layden, Newman, Freeman, & Morse, 1993; Linehan, 1993; Young, 1990); for groups of disorders (Beck, 1995; Blackburn, Twaddle et al., 1996; Dobson, 1995; Hawton, Salkovskis, Kirk, & Clark, 1989; Padesky & Greenberger, 1995); for psychosis (Birchwood & Tarrier, 1994; Fowler, Garrety, & Kuiper, 1995; Kingdon & Turkington, 1994); for substance abuse (Beck, Wright, Newman, & Liese, 1993); for bipolar disorder (Basco & Rush, 1996); and for inpatients (Wright, Thase, Beck, & Gudgate, 1993). In this chapter, the general characteristics of cognitive therapy are described and attention is focused on developments which have evolved following the theoretical developments described above.

Cognitive therapy was first developed as a short-term treatment of 12–16 weeks, for unipolar depressed outpatients (Beck et al., 1979). The extension of its application to an ever increasing range of conditions and of disorders may be a cause for concern for some, but need not be surprising in view of the fact that cognitive theory is meant to be a comprehensive theory of psychopathology. For each disorder, the underlying generic model is applied, with additional components specific to each disorder (see later chapters on specific disorders). Newer applications have not been tested in controlled outcome studies, but the empirical basis of cognitive therapy remains very much the tradition, with models tested in experimental studies, followed by case studies, and then by controlled trials.

For most disorders cognitive therapy remains short-term, except in the case of personality disorders when therapy can extend to one or two years. Whereas in short-term cognitive therapy, weekly one-hour sessions are the norm, in personality disorders therapy sessions are often more infrequent, allowing the patient more time for homework assignments and more time to get over the high levels of emotion which may be stirred up during therapy. Typically, short-term therapy may last 10–20 sessions, although in some cases (especially in panic disorder) fewer sessions may suffice. Sessions are structured by the use of an agenda which is set collaboratively by the patient and the therapist. This ensures the active participation of the patient and the appropriate use of time. Generally, a session agenda will include a review of previous assigned homework; setting up one or two relevant areas for discussion, these being often related to previous homework; session feedback and deciding on appropriate homework for the coming week. Throughout therapy, the therapist must ensure that the process is collaborative, by giving and asking for feedback, by rehearsing, and by questioning. The therapist remains active and directive, but never prescriptive. The style is gentle and questioning, guiding the patient to elucidate their emotions and thoughts, to evaluate them, and to arrive at alternative interpretations and solutions. This style of questioning has been described as socratic questioning (Overholser, 1993a, 1993b). Questions are used for gathering information, for discovering new perspectives, for finding commonalities between situations, patterns of thinking and emotions (synthesis), and for achieving change. Overholser (1993a) differentiates between seven types of socratic questions: (i) memory questions (When did the problem first begin? When did it last happen? What did you do when it happened?); (ii) translation questions (What does it mean to you? How can we make sense of this?); (iii) interpretation questions (How are these situations similar? Do your problems at work seem similar in any way to your problems at home?); (iv) application questions (What have you tried in order to solve this problem? How will you go about making these changes?); (v) analysis questions (What evidence do you have for this? How could you tell if you are right or wrong?); (vi) synthesis questions (In what other ways could you look at this situation? What does all this say to you? What does it mean to you to be a mother?); and (vii) evaluation questions (What does it mean to you to be a success? How do you rate yourself as a person?).

Another general characteristic of cognitive therapy is openness, that is, the therapist is explicit about the model of therapy, about the rationale for the procedures which are used, about the formulation of problems, about their own reaction in the therapeutic relationship, and about the rationale for homework assignments. Feedback is elicited not only regarding the content of therapy, but also about the patient's reaction to the therapist. This explicit and open style not only fosters collaboration, the model being of two scientists working together to solve problems, but also ensures that therapist and patient share a common understanding of what is going on in therapy. It is particularly important at the onset of therapy to explain the model of therapy by sharing the generic model of cognitive therapy with the patient, using the diagram shown in Figure 4, with real examples from the information gathered in the first assessment.

Therapy is problem oriented: therapist and patient engage in a functional analysis of problems in the assessment stage, so that therapy remains focused. The functional analysis includes symptoms (panic attacks, mood, hopelessness, suicidal wishes or behaviors, situational avoidance, etc.), problem situations (work, marital relationship, interpersonal relationship, etc.), life condition (employment, social support, social network, etc.), proximal trigger situations (e.g., loss or threat situations), maintaining factors (current hassles and problems, coping strategies), distal factors (short developmental history), typical automatic thoughts, and typical behaviors. This analysis leads to an agreed list of problems which can be prioritized and targeted for therapy. An example of assessment is given in Figure 5.

Figure 5 indicates that although a detailed conceptualization, as shown in Figure 2, cannot be made at the assessment stage of therapy, nonetheless the therapist may already have some clues about the basic assumptions and the core schemata of the patient derived from the themes of the negative automatic thoughts and from the pattern of hypothesized predisposing factors from early childhood experiences. In the example, a basic assumption may be "If people do not treat me properly, this means I am worthless" and a core schema may be "I am worthless." Through questioning (e.g., What do you think is the most pressing problem for us to work on? What would help you most at the moment?), problems are prioritized to bring about effective change quickly. This is a two-way process, so that the therapist and the patient reach conclusions collaboratively about what priorities suggest themselves. In the case described in Figure 5, the suicidal ideas and wishes of the patient would take priority and it is usual to work then at a symptom level, using behavioral methods of treatment to increase activity, in particular pleasurable activities. However, concurrently, during the therapy sessions, the therapist would start socializing the patient into cognitive methods of treatment to increase awareness of automatic thoughts, and to begin to evaluate these thoughts. An initial conceptualization can then follow at about the fifth or sixth session of therapy, which is shared with the patient.

Conceptualization is an aspect of cognitive therapy which is more emphasized now than it used to be in the early days of cognitive therapy (Blackburn et al., 1996; Persons, 1989; see Chapter 6.02, this volume for a detailed discussion). An example of a conceptualization was given in Figure 2 above. The cognitive conceptualization of cases will vary from disorder to disorder and is done within the specific cognitive theory framework of that disorder. For example, a conceptualization of a case of panic disorder is given in Figure 6 to highlight the difference from the case of depression in Figure 2.

As Persons (1989) points out, conceptualization plays a crucial role in cognitive therapy:

(i) It helps the therapist to choose an intervention strategy, for example, it might indicate that environmental changes are possible and indicated or that the patient's view of themself as vulnerable and weak is more central than their avoidance behaviors.

(ii) It helps to choose the treatment modality, couple therapy instead of individual therapy.

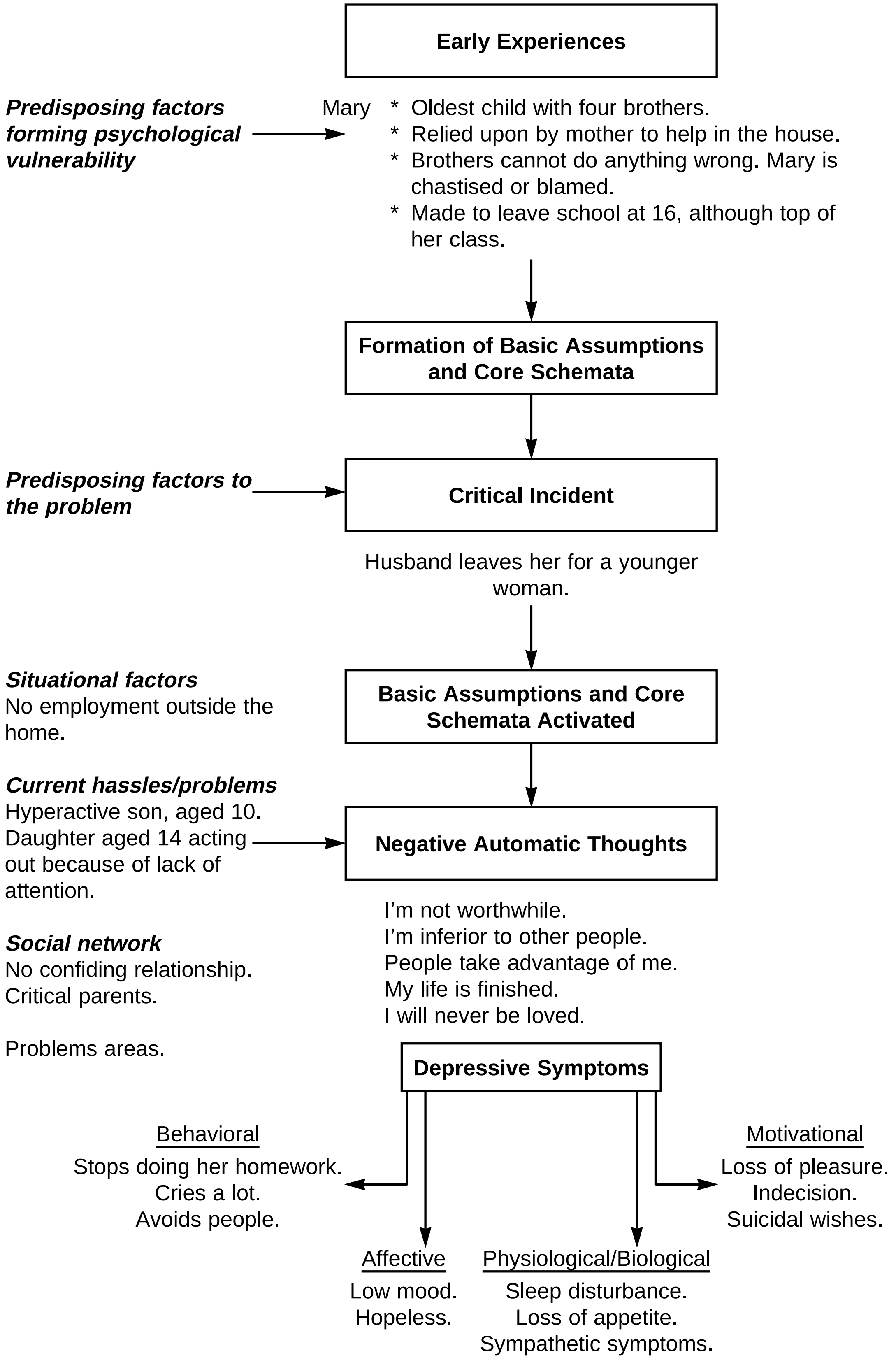
Early Experiences
Predisposing factors forming psychological vulnerability
Mary
* Oldest child with four brothers.
* Relied upon by mother to help in the house.
* Brothers cannot do anything wrong. Mary is chastised or blamed.
* Made to leave school at 16, although top of her class.
Formation of Basic Assumptions and Core Schemata
Predisposing factors to the problem
Critical Incident
Husband leaves her for a younger woman.
Situational factors
No employment outside the home.
Current hassles/problems
Hyperactive son, aged 10.
Daughter aged 14 acting out because of lack of attention.
Social network
No confiding relationship.
Critical parents.
Problems areas.
Basic Assumptions and Core Schemata Activated
Negative Automatic Thoughts
I'm not worthwhile.
I'm inferior to other people.
People take advantage of me.
My life is finished.
I will never be loved.
Depressive Symptoms
Behavioral
Stops doing her homework.
Cries a lot.
Avoids people.
Motivational
Loss of pleasure.
Indecision.
Suicidal wishes.
Affective
Low mood.
Hopeless.
Physiological/Biological
Sleep disturbance.
Loss of appetite.
Sympathetic symptoms.

Figure 5 Examples of a functional analysis of a problem at assessment (Mary, age 42).

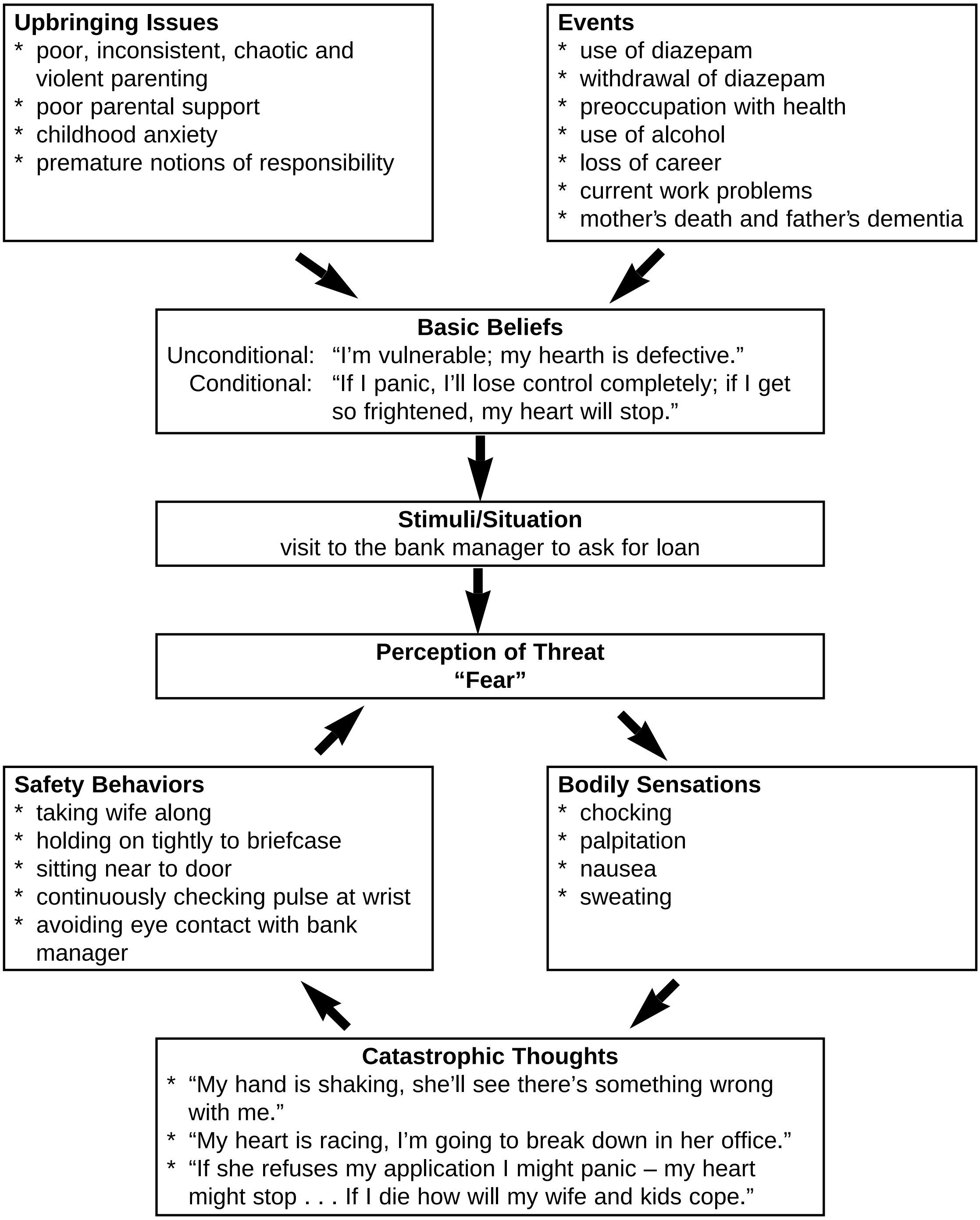
Upbringing Issues
* poor, inconsistent, chaotic and violent parenting
* poor parental support
* childhood anxiety
* premature notions of responsibility
Events
* use of diazepam
* withdrawal of diazepam
* preoccupation with health
* use of alcohol
* loss of career
* current work problems
* mother's death and father's dementia
Basic Beliefs
Unconditional: "I'm vulnerable; my hearth is defective."
Conditional: "If I panic, I'll lose control completely; if I get so frightened, my heart will stop."
Stimuli/Situation
visit to the bank manager to ask for loan
Perception of Threat
"Fear"
Safety Behaviors
* taking wife along
* holding on tightly to briefcase
* sitting near to door
* continuously checking pulse at wrist
* avoiding eye contact with bank manager
Bodily Sensations
* chocking
* palpitation
* nausea
* sweating
Catastrophic Thoughts
* "My hand is shaking, she'll see there's something wrong with me."
* "My heart is racing, I'm going to break down in her office."
* "If she refuses my application I might panic – my heart might stop . . . If I die how will my wife and kids cope."

Figure 6 Conceptualization of a panic disorder case (reproduced by permission of Souvenir Press from *Cognitive therapy in action. A practitioner's case-book*. Blackburn, Twaddle et al., 1996).

(iii) It guides the therapist in the choice of an intervention point, for example, what changes are likely to bring about the most gain and are likely to have the biggest impact on the hypothesized basic assumptions and schemata.

(iv) It helps the therapist to predict the patient's behavior, for example, is the patient likely to become over-dependent on the therapist or likely to be noncompliant in terms of attendance and homework assignments?

(v) It helps to make sense of the patient's relationship difficulties both in and out of therapy and this guides the therapist's behaviors.

(vi) It may help to understand why therapy is not working. The formulation may be wrong and the therapist needs to consider new information that might have transpired in the course of therapy and revise the original formulation of the patient's problems.

(vii) Because a good formulation should lead to appropriate strategies, methods of treatment, and intervention points, length of treatment can be much reduced as time is used effectively and economically.

6.03.3.2 Treatment Methods

Treatment strategies in cognitive therapy are driven by the case conceptualization which is itself firmly based in the cognitive theory of the emotions, as described in the earlier part of this chapter. Different targets of therapy may require a different balance of cognitive, behavioral, and emotional therapeutic methods, which emphasizes that cognitive therapy is not a series of techniques to be mastered and applied technically. It is a system of psychotherapy requiring a thorough understanding of the cognitive theory of the emotional disorders, familiarity with the experimental literature on which it is based, and knowledge and experience of psychiatric syndromes in general. In addition, the competent cognitive therapist needs to display the general qualities of a good psychotherapist; namely, genuineness, warmth, empathy, and understanding.

6.03.3.2.1 Behavioral methods

Behavioral methods of treatment typically are applied early on in therapy as they can often provide relief for distressing symptoms which may prevent therapy from progressing. It is also perhaps easier for patients to master behavioral techniques, while concurrently getting socialized in cognitive techniques. Behavioral techniques are not used in cognitive therapy without examining concomitant effects on cognitions and emotions.

Typical behavioral methods of treatment and the problems for which they are suitable are listed below:

(i) *Inactivity*. Scheduling of graded activities with reasonable goals to approximate in the long run the patient's normal routine.

(ii) *Indecivesiveness*. Planning daily activities and rescheduling activities for another day if the original plan has not been adhered to because of unforeseen circumstances. For life decisions, a more cognitive approach may be indicated (illustrated in Table 1).

(iii) *Procrastination*. Establishing detailed daily plans of activities to reduce decision making. Again, it is necessary to ensure that the plans are feasible and attainable.

(iv) *Lack of pleasure*. Scheduling of potentially pleasurable activities and rating activities for pleasure. The therapist needs to ensure that the patient is not showing an all or none response, that is not rating activities as pleasurable or not pleasurable and not confusing ratings of pleasure with ratings of achievement

(v) *Lack of concentration*. Practice engaging in activities (e.g., reading) in small bursts which are reinforcing, instead of persevering when concentration has gone or giving up altogether. The therapist must watch out for the patient's self-talk or automatic thoughts.

(vi) *Low mood*. Engaging in distracting and pleasurable activities. Mental imagery of pleasant memories.

(vii) *Anxious mood*. Distraction, relaxation techniques. Distraction techniques may involve focusing on an object and describing it in minute details; developing sensory awareness by concentration on visual, auditory, olfactory, and bodily sensations; mental exercises, for example, reading, watching a film, or doing the crosswords.

Relaxation is an important technique for patients who feel they have no control over how they feel. Several relaxation methods have been described (Bernstein & Borkovec, 1973; Goldfried & Davison, 1976), however, the applied relaxation method described by Ost (1987) has been shown to be particularly effective. It involves eight stages which can be taught sequentially or in subsets: giving a rationale for relaxation training; recognizing the early signs of anxiety; progressive relaxation; relaxation exercises without tensing the muscles first; cue-controlled relaxation, that is, inhaling and relaxing to the cue word relax; differential relaxation, that is, learning to relax while engaged in everyday activities; rapid relaxation by using several cues in the patient's environment; and, finally, application training which involves applying the learnt relaxation in anxiety provoking situations.

(viii) *Lack of motivation*. Schedule activities and grade for mastery and pleasure. The underlying rationale here is that the main problem lies in starting an activity and that, once started, motivation increases through the desire to finish the task (Zeigarnik, 1927) and the unexpected pleasure experienced, much as appetite may actually be stimulated by the act of eating, although there might not have been much desire to eat at the outset.

(ix) *Panic attacks*. Respiratory control exercises which reduce panic symptoms due to hyperventilation (Clark, 1989). The exercises consist in pacing breathing to 12 breaths per minute or eight breaths per minute by breathing for two seconds "in" and for two seconds "out"

(12 breaths/min) or three seconds "in" and three seconds "out" (eight breaths/min)

(x) *Avoidance*. Exposure to avoided situations, in a graded and repeated fashion, for example, social, phobic, or agoraphobic situations.

(xi) *Safety behaviors*. These are behaviors that anxious patients engage in when anxiety or panic feelings have already started, for example, holding on to solid objects, sitting down, pressing their head, checking their pulse rate. These behaviors can be very subtle and need to be elicited carefully or observed during a behavioral test. They need to be eliminated gradually as they maintain the patient's beliefs that symptoms are dangerous.

(xii) *Rituals*. Deliberate exposure to previously avoided situations or feared stimuli (including thoughts) and prevention of compulsive rituals and neutralizing behaviors, including neutralizing thoughts.

(xiii) *Problem situations*. Rehearsal in role-plays, with the therapist, of coping techniques and strategies, for example, assertiveness training, negotiations with a spouse or at work, and anger control.

It is already evident from the list of behavioral methods described above that they are applied within a general cognitive strategy to help in identifying the cognitions which are related to the problems, to test the patient's predictions and to modify expectations about the self, the world, and the future.

6.03.3.2.2 Cognitive methods

(i) Identifying automatic thoughts

The main emphasis at the beginning of therapy, even while applying behavioral methods of treatment, is the patient's automatic thoughts, which are considered as the basic data of cognitive therapy. The patient needs to understand the nature of automatic thoughts and how they affect emotions and behaviors and this is done from the time of assessment, through questioning, noting spontaneous reports of examples of automatic thoughts, and including them in the model of therapy which is presented to the patient. Because the very nature of automatic thoughts indicates that they are habitual and involuntary, some patients have difficulty in getting access to their "hot" automatic thoughts for several reasons: they have been used to taking them for granted; or they report a commentary about their automatic thoughts (thoughts about thoughts) through excessive intellectualization; or they feel embarrassed about their automatic thoughts; or they consider them as silly. A great deal of socialization is, therefore, necessary to guide the patient early on in therapy. A good way for the therapist to decide whether they have accessed the right automatic thought is to ask themself: "Is this thought consistent with the emotion that is reported?", "Is this thought consistent with the degree of emotion that is reported?", and "If I had the same thought, would I feel as bad?"

There are several methods to guide the therapist about how to access the patient's automatic thoughts:

(a) Direct questions. "What was going through your mind when you felt panicky?"; "What went through your mind when your husband criticized your housework?" This is evidently the easiest method and it is surprising how often this is enough to trigger the appropriate negative automatic thoughts. This questioning also serves as a model for the patient to tap into their automatic thoughts at the time of strong emotions or of self-defeating behaviors.

(b) If direct questions do not prove useful, the therapist poses a series of questions to guide the patient to access the automatic thoughts. This approach has been labeled "guided discovery" or "socratic questioning." It is the main technique which cognitive therapists need to master. It demands good listening skills, an understanding of the patient's style of thinking and a formulation of the problems which acts as a guide or a map.

(c) Sometimes the patient describes situations or events which may have caused problems in a detached way, the way they may describe a film or a book or a road accident that one might have witnessed. These are probably examples of propositional and not implicational meaning, as described by Teasdale (1993). In such cases, the therapist can use mental imagery to help the patient relive the actual situation as vividly as possible, with all the sensory input which might be involved. This method helps in triggering strong emotions and the accompanying automatic processing which might have taken place.

(d) For interpersonal situations, role-plays serve the same purpose. Having elicited the details of what actually took place, the therapist reenacts the situation by playing the role of the other person involved as realistically as possible, using posture, tone, and volume of voice to trigger the emotions and, hence, the automatic thoughts.

(e) Moments of strong emotions are, therefore, of crucial importance to access automatic thoughts. If emotion is not shown or expressed in therapy, the therapist tries to elicit emotions through imagery or role plays. If it is present during the session, the therapist uses the

Table 1 A method for dealing with indecisiveness. A married woman, with two school age children, having recovered from a depressive illness, cannot decide about her future plans.

Alternatives	*Pros*	*Importance*	*Cons*	*Importance*	*Balance*
1. Go back to the same job on a full-time basis	I enjoyed the job, because of its responsibility.	80	It was not a 9am-5pm job, so I never knew when I would get home and sometimes I had to take work home with me.	100	
	I knew the job and could do it well.	80	The pressure might make me ill again.	100	
	If I got depressed again I could have time off, as the employers were very understanding the last time.	100	It was not fair on the children not to spend more time with them.	100	
	It gave me a good income of my own and made me feel independent.	20			
		Σ + 280		Σ + 300	Σ − 20
2. Go back to the same job on a part-time basis.	It would give me more time with the children.	100	I would no longer have the responsibility which I enjoyed.	80	
	The part-time income would still make me feel independent.	20	It is a difficult job to do part-time. I might find myself doing as much as before for less money.	80	
	It would be less pressure and I might feel better.	100			
	I would be able to move to full-time again when the children grow up.	100			
		Σ + 320		Σ − 160	Σ + 160
3. Look for another part-time job which is less demanding	It would involve regular hours—more time at home.	100	Better the devil you know. A new job might turn out inappropriate.	80	
	May be less interesting, but would get out of the house nonetheless.	60	I get bored by routine jobs.	80	
	My husband would be happier with this.	50	Less money.	20	
	Less pressure would be better for my health.	100			
		Σ + 310		Σ − 180	Σ + 130
4. Not go back to work at all.	A lot of free time to do things with the children.	100	I would miss having an interesting job outside the home.	100	
	I could use the spare time at home to study a language or something else.	30	I would be bored.	50	
	I could spend more time on my hobbies: gardening, cooking, going out.	20	I would be resentful.	50	
	Both sets of parents and my husband would approve.	50	In nine years time, by the time children have grown up, I will be older and it would be more difficult to get into the job market.	20	
	May be better for my health.	100			
		Σ + 300		Σ − 220	Σ + 80

Final decision: There are advantages and disadvantages for all alternatives, but going back to the same job on a part-time basis is preferable.

Source: Blackburn, 1993, pp. 94–97.

moment to help the patient access the automatic thought. It is particularly important to be aware of emotions which may not appear to be related to the current subject of discussion: the patient may become fidgety or tense, blush or start crying and the therapist then asks "what"s going through your mind just now?"

(f) Behavioral tasks have already been described in the previous section as an invaluable method to train the patient to become aware of their thoughts as they feel the negative emotions aroused while encouraging in behavioral tests.

(g) At the beginning of therapy, patients can increase their awareness of negative thoughts by simply counting them, using some sort of counter which they keep with them at all times or keeping a tally on an index card.

(h) The classical method of increasing awareness of negative automatic thoughts is the use of Dysfunctional thoughts records (DTRs) where the patient records the situation (what was going on) when they felt a strong negative emotion, what the emotion was and at what intensity and what were the automatic thoughts and what was their degree of belief in their automatic thoughts. Only these three columns are used at the beginning of therapy and they should be used in the session for training, before being given as a homework assignment. Several examples of DTRs have been given in cognitive therapy textbooks (Blackburn & Davidson, 1995; Hawton et al., 1989).

(ii) Evaluating automatic thoughts

In the therapy sessions and as homework assignments, the patient learns to evaluate their automatic thoughts. This can probably be considered the core of cognitive therapy, involving the acquisition of the necessary skills to distance oneself from one's automatic processing of information and to treat the products of thoughts as interpretations of reality, rather than reality itself. These skills are of crucial importance as they are generalisable to the evaluation of others and to basic assumptions and core beliefs as well.

The evaluation of automatic thoughts is done through careful questioning to guide the patient to consider various alternatives to their original, automatic processing and to adopt more functional alternatives that they can test and come to believe in. The therapist's questioning becomes a model for the patient to apply to their thinking outside of therapy.

Questioning may take the form of:

(a) What is the evidence for this interpretation?

(b) Are there alternative interpretations which may be more realistic?

(c) What is the effect of thinking that on me, on others?

(d) What thinking errors am I making?

(e) Am I jumping to conclusions? (arbitrary inference),

(f) Have I considered all aspects of the situation (selective abstraction),

(g) Am I using one instance to draw general conclusions, when there is no basis to do this? (overgenalization),

(h) Am I catastrophizing? (minimization and maximization),

(i) Am I taking this personally when it may have nothing to do with me? (personalization),

(j) Am I thinking in black and white? (dichotomous thinking),

(k) Am I applying arbitrary rules, "shoulds" and "musts" when, in fact, there are no such rules?

Table 2 gives a list of 20 questions which can be given as a handout to patients to help them evaluate their automatic thoughts.

When the degree of belief in the original automatic thought is put in question, alternative interpretations can be considered, and the probability of these alternatives can be assessed. The alternative interpretation, after careful collaborative evaluation, can sometimes be seen clearly as more likely and more realistic, but even if this is not the case, when the alternatives are not more probable than the original interpretation, it is beneficial to see that one interpretation—is not necessarily correct, so that the degree of belief in the distressing interpretation becomes reduced.

In the process of revaluation, a negative event or outcome may be reattributed to external instead of internal, less global, and less stable causes (Abramson et al., 1978).

It is particularly important that the alternative interpretations or conclusions that are arrived at be owned by the patient, rather than dictated by the therapist, as the degree of belief in the alternatives will depend on that. Degree of belief in the alternatives are rated and the resultant emotions rated. Thus, it is not taken for granted that because alternatives have been considered, the patient would automatically believe in them and feel better. The DTR, with the five columns, Situation, Emotion, Automatic Thought, Alternative Response, and Outcome is used for that purpose. An example is given in Table 3.

To consolidate the patient's revaluation of their automatic thoughts and the degree of belief in them, several methods can be used, for example, rehearsal, role-plays, and behavioral tests. Rehearsal simply restates the same situation or equivalent situations and the typical automatic thoughts and the patient is invited to

Table 2 Twenty questions to help challenge negative thinking.

Question	*Response*
Am I confusing a thought with a fact?	The fact that you believe something to be true does not necessarily mean that it is. Would your thought be accepted as correct by other people? Would it stand up in court, or be dismissed as circumstantial? What objective evidence do you have to back it up, and to contradict it?
Am I jumping to conclusions?	This is the result of basing what you think on poor evidence. For instance, depressed people often believe that others are thinking critically about them. But none of us are mind-readers. How do you *know* what someone else is thinking? You may be right, but don't jump to conclusions—stick to what you know, and if you don't know, see if you can find out.
What alternatives are there?	Are you assuming your view of things is the only one possible? How would you have looked at this situation before you got depressed? How would another person look at it? How would you look at it if someone else described it to you?
What is the effect of thinking the way I do?	What do you want? What are your goals? Do you want to be happy and get the most out of life? Is the way you are thinking now helping you to achieve this? Or is it standing in the way of what you want?
What are the advantages and disadvantages of thinking this way?	Many distorted thought patterns do have some pay-off—that is what keeps them going. But do the disadvantages outweigh the advantages? If so, you can think out a new way of looking at things which will give you the advantages, but avoid the disadvantages of the old way.
Am I asking questions that have no answer?	Questions like "How can I undo the past?" "Why am I not different?" "What is the meaning of life?" "Why does this always happen to me?" "Why is life so unfair?" Brooding over questions like these is a guaranteed way to depress yourself. If you can turn them into answerable questions, so much the better. If not, don't waste time on them.
Am I thinking in black-and-white, all-or-nothing terms?	Nearly everything is relative. People, for instance, are not usually all good or all bad. They are a mixture of the two. Are you applying this kind of black-and-white thinking to yourself?
Am I using global words in my thinking?	Watch out for words like always/never, everyone/no-one, everything/nothing. The chances are that the situation is actually less clear-cut than that. Mostly it's the case of sometimes, some people, and some things.
Am I condemning myself as a total person on the basis of a single event?	Depressed people often take difficulties to mean that they have no value at all as a person. Are you making this kind of a blanket judgement?
Am I concentrating on my weaknesses and forgetting my strengths?	When people become depressed, they often overlook problems they handled successfully in the past and resources which would help them overcome current difficulties. Once they can change their thinking, they are often amazed at their ability to deal with problems. How have you coped with similar difficulties in the past?

Table 2 (continued)

Question	*Response*
Am I blaming myself for something which is not really my fault?	Depressed people, for instance blame themselves for being depressed. They put it down to lack of willpower, or weakness, and criticise themselves for not "pulling themselves together." In fact, scientists have been studying depression for many years and they are still not certain what causes it. Depression is a difficult problem to solve and blaming yourself for it will only make you more depressed.
Am I taking something personally which has little or nothing to do with me?	When things go wrong, depressed people often believe that in some way this is directed at them personally, or caused by them. In fact, it may have nothing to do with them.
Am I expecting myself to be perfect?	It is simply not possible to get everything right all the time. Depressed people often set unrealistically high standards for themselves. Then they condemn themselves for making mistakes, or acting in ways they would rather not have done. Accepting that you can't be perfect does not mean you have to give up trying to do things well. It means that you can learn from your difficulties and mistakes, instead of being upset and paralysed by them.
Am I using a double standard?	You may be expecting more of yourself than you would of another person. How would you react to someone else in your situation? Would you be so hard on them? You can afford to be as kind to yourself as you would be to someone else. It won't lead to collapse.
Am I paying attention only to the black side of things?	Are you, for instance, focusing on everything that has gone wrong during the day and forgetting or discounting things you have enjoyed or achieved?
Am I overestimating the chances of disaster?	Depressed people often believe that if things go at all wrong, disaster is sure to follow. If the day starts badly, it can only get worse. How likely is it that what you expect will really happen? Is there really nothing you can do to change the course of events?
Am I exaggerating the importance of events?	What difference does a particular event really make to your life? What will you make of it in a week, a year, 10 years? Will anyone else remember what you now see as a terrible thing to do? Will you? If you do, will you feel the same way about it? Probably not.
Am I worrying about the way things ought to be, instead of accepting and dealing with them as they are?	Are you allowing events in the world at large to feed your depression? Telling yourself life is unjust and people awful? It is sad that there is so much suffering in the world and you may decide to do what you can to change things, but getting depressed about it does nothing to help.
Am I assuming I can do nothing to change my situation?	Pessimism about the chances of changing things is central to depression. It makes you give up before you even start. You can't know that there is no solution to your problems until you try. Is the way you are thinking helping you to find answers, or is it making you turn down possible solutions without even giving them a go?
Am I predicting the future instead of experimenting with it?	The fact that you have acted in a certain way in the past does not mean to say that you have to do so in the future. If you predict the future, instead of trying something different, you are cutting yourself off from the chance of change. Change may be difficult, but it is not impossible.

Table 3 Example of automatic thoughts and how to answer them.

Situation	*Emotion* (rate degree, 0–100%)	*Automatic thoughts* (rate belief, 0–100%)	*Alternative responses* (rate belief, 0–100%)	*Outcome* (re rate belief automatic thought and emotion)
Woke up and kept thinking of office situation	Disturbed, anxious, low (70%)	What mess am I going to have to go to? (70%)	I'm crystal ball gazing. It may not be a mess. However, if it is less organized than before, it will be the responsibility of the supervisor. I can only do my work as well as I can and let other people worry about their work (100%)	Automatic thought (10%) Anxious (20%)

Source: Blackburn, 1993, p. 66.

evaluate these thoughts along the lines of the preceding discussion. In role plays, the therapist may reenact a typical distressful scenario and elicit thoughts and emotions or the therapist may play devil's advocate and play back the patient's automatic thoughts which they then have to modify using the skills which have been practiced beforehand. This is a particularly impactful method which usually triggers high emotional responses. Several attempts using role reversals are usually necessary. Finally, behavioral tests of the new interpretation can be devised collaboratively to check whether the new interpretation is indeed more probable. These may involve polling friends and relatives to check their opinions in reality rather than mind reading; facing rather than avoiding a confrontational or self-assertive situation to assess others' reactions; engaging in, say, panic provoking situations, making predictions, and reviewing the evidence after the behavioral test and drawing conclusions.

When dealing with automatic thoughts, several caveats need to be drawn to the attention of the therapist: working on peripheral rather than key automatic thoughts; dealing with questions instead of the underlying thought which prompts the question; buying into the patient's system instead of keeping to a questioning style; not engaging in a thorough evaluation of the thoughts; and not engaging in consolidation methods.

(iii) Basic assumptions and core schemata

These have been described in previous sections and are an integral part of the information processing model of the emotional disorders. It is considered that therapy would be incomplete without dealing with this aspect of the patient's cognitive system which underlies the way information is processed and which is considered to represent psychological vulnerabilities. Therapy at that level is therefore considered to be preventative. However, the relative emphasis put on eliciting and modifying core structures may vary from disorder to disorder, being most central in the personality disorders and in depression and general anxiety and possibly less so in obsessive-compulsive, panic, health anxiety, and psychotic disorders.

(iv) Determining basic assumptions and core schemata

As therapy progresses, general themes in problem situations and in automatic thoughts become evident, as are personal rules, and these are discussed with the patient within the conceptualization of problems as shown in Figures 2, 5, and 6.

The therapist listens attentively for the implicit meanings of the patient's statements and makes them explicit, as usual in question form, for example, "Does this mean that if people do not like you, you are no good?"; "Does this apply to everybody or only to some people?"; and "Does this mean that you are a failure?"

It is interesting to note that often, quite early on in therapy, the patient may express an automatic thought such as "I am a failure" which, in fact, is their core schema. At this stage, the patient may be totally unaware that this is their general view of themself and the therapist may have only an inkling that it is the patient's core schema. It is therefore advisable to treat the statement as an automatic thought at this stage, as the patient has not yet acquired the skills to evaluate automatic thoughts and an overall understanding of the common thread in all the problems has not yet been reached by the patient and/or the therapist.

To bring out the common themes in the automatic thoughts, the therapist can use several examples to extract the general implications of the automatic thoughts. This method has been labeled the "downward arrow" technique, where the automatic thoughts are accepted as possibly true and their ulterior catastrophic implications arrived at through questioning. An example is given in Table 4.

It is to be noted in the example illustrated in Table 4 that the patient is not encouraged to evaluate automatic thoughts at this stage, but instead is guided to consider the meaning implicit in the automatic thoughts which causes distress. At the end of this exercise, the therapist would help put the conclusion explicitly: "It seems from this that you are saying that people will not respect you if you do not do things perfectly? Is this right?" (basic assumption); "You also say that it would simply confirm the fact that you are worthless?" Other examples are then used to construct a consistent formulation which gives the patient an overall understanding.

It may be an aid to therapy to use standard questionnaires to elicit basic assumptions and core schema, for example, the DAS (Weissman & Beck, 1978) or Young's schema questionnaire (1990). However, questionnaires should never be used by themselves, as discovery of the underlying structure needs to be a joint endeavor between therapist and patient. Questionnaires are also open to response biases, and, by definition, cannot capture individual meanings and idiosyncratic terminology.

(v) Modifying basic assumptions and core schemata

Methods for modifying core structures are very similar to those used for the evaluation of automatic thoughts, using the same style of socratic questioning and of behavioral tests. They might be classified as cognitive, behavioral, emotional, and inerpersonal methods. A summary is presented below.

(a) *Cognitive methods.* These involve weighing the advantages and disadvantages of thinking in certain ways, for example, "I must do everything perfectly" (see Table 5); examining the evidence for and against core beliefs, such as, "I am worthless"; establishing continua of zero to 100 to rate personal qualities; evaluating the validity of personal rules; contrasting the short- and long-term utility of personal rules; examining the validity of the conclusions drawn at the time when the basic schemata were arrived at in childhood; collecting evidence contrary to the basic schema. Padesky (1990) describes the self-schema as self-prejudice to explain to the patient how the basic schema acts to filter information, so that only negative confirming information

is accepted and processsed, while disconfirming information is either transformed into confirming information or ignored or disqualified. Padesky (1994) describes in detail several cog-

Table 4 Example of "downward arrow" technique to arrive at basic assumptions and core schemata.

Situation: This is Helen's first day at the office	
Emotions: Depressed (60%), anxious (70%)	
Automatic thoughts	
Helen	I'm really dumb. I should have organised myself better.
Therapist ↓	Supposing that were true, what would it mean about you?
Helen	It means that I can never get things right or do things properly.
Therapist ↓	Supposing that were true, what would it mean about you?
Helen	It would mean that people cannot trust me to do things perfectly.
Therapist ↓	Supposing that were true, what would it mean to you?
Helen	It would mean that I am second rate, useless.
Therapist ↓	Supposing that were true, what would it mean?
Helen	It would mean that nobody would respect me.
Therapist ↓	And if nobody respected you, what would it mean?
Helen	It would mean what I've always known—that I am worthless.

nitive change methods and methods for consolidating the new schema (e.g., keeping a positive data log).

(b) *Behavioral methods*. These invlove engaging in activities which disconfirm the basic belief (e.g., do a piece of work in less than the usual time and testing whether not attaining perfection entails social or professional catastrophies); or stopping behaviors which maintain the belief (e.g., stop avoiding meeting people).

(c) *Emotional methods*. These may involve some procedures from Gestalt therapy, for example, role plays of key painful experiences of the past, when the therapist, appropriately briefed by the patient, can play the role of a punishing parent or of an unfair teacher, and the patient plays the role of the child, but with an adult understanding. These role plays are rehearsed in role reversals several times, usually bringing strong emotional responses. Other emotional methods may involve examining key experiences of the past and using reattribution methods to decrease self-blame and guilt which might have been reinforced by critical parents in childhood. Figure 7 below describes the conclusions reached after discussing a patient's sense of responsibility for the unhappiness of her parents.

(d) *Interpersonal methods*. These involve group, marital, or family therapy if indicated; but more usually; the therapist uses the interpersonal relationship in the therapeutic situation to develop ideas about problems in interpersonal style which are then discussed openly, as is usual in cognitive therapy.

6.03.3.3 Developments

The area of cognitive therapy which has developed the most in the 1990s is the methodology for evaluating and modifying schemata. These changes have evolved gradually through theoretical developments (e.g., Teasdale & Barnard, 1993) and through the expansion of cognitive therapy to new areas of psychopathology, in particular to the personality disorders (Beck et al., 1990; Layden et al., 1993; Linehan, 1993; Young, 1990). Similarly, through, the influence of the constructivist approach (Guidano & Liotti, 1983; Mahoney, 1993) and because of the wider application of cognitive therapy, more emphasis is probably put on developmental and interpersonal issues in the formulation of cases. Another area of development is the application of relapse prevention methods at the end of treatment (Wilson, 1992). The importance of this stage of therapy is reflected in the dedication of a whole chapter to this topic in this volume.

Table 5 Weighing the advantages and disadvantages of a dysfunctional basic assumption.

Basic assumptions: I must do everything perfectly, if not, people will not respect me and I am worthless.

Advantages of this belief	*Disadvantages of this belief*
It makes me try hard to do well	It increases my anxiety, so that my performance suffers.
It makes me produce good work and be successful.	It stops me from doing many of the things I would like to do, because I may not succeed.
When something goes well, I feel really good.	It makes me very critical of myself, so that I cannot take pleasure in what I do.
	I cannot afford to let my mistakes be noticed by anyone, and therefore I probably miss out on valuable constructive comments.
	When I get criticized, I become defensive and angry.
	My successes are undermined, because any subsequent failure wipes out their significance.
	I become very intolerant of others. I find so many faults in others, that I cannot be warm and friendly. I will end up without any friends.
	I can never think well of myself because it is impossible to get it right all the time.
	Because I get so upset by failures, I cannot use them as valuable experiences to learn how to do things better next time.

Source: Blackburn, 1993, pp. 113–1140

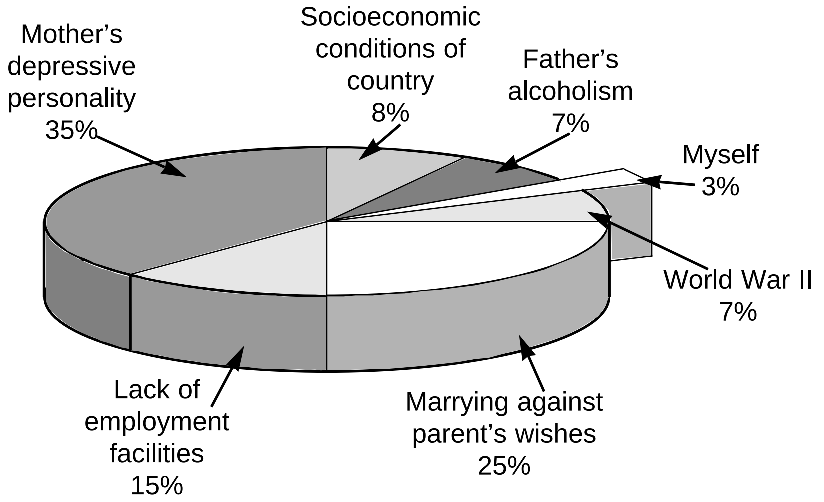

Figure 7 Responsibility pie chart. "I was responsible for my mother's unhappiness."

6.03.4 PROCESS OF COGNITIVE THERAPY

A large number of controlled outcome studies, reviewed by Blackburn, Twaddle et al. (1996), attest to the efficacy of cognitive therapy in depression (Blackburn, Bishop, Glen, Whalley, & Christie, 1981; Elkin et al., 1989; Hollon et al., 1992; Murphy, Simons, Wetzel, & Lustman, 1984; Rush, Beck, Kovacs, & Hollon, 1977; Teasdale, Fennell, Hibbert, & Amies, 1984); in general anxiety disorders (Butler, Fennell, Robson, & Gelder, 1991; Durham et al., 1994; Durham & Turvey, 1987; Power et al., 1990); in panic disorder (Barlow, Craske, Cerney, & Klosko, 1989; Beck et al., 1992, 1994; Clark et al., 1994) and to a lesser extent in obsessive-compulsive disorder (Emmelkamp & Beens, 1991; Emmelkamp, Van der Helm, Van Zainten, & Ploghg, 1980; Van Oppen et al., 1995). However, there is little understanding of what are the actual critical components of cognitive therapy in effecting change and most of the work to date is in depression. Relevant studies have been reviewed by Whisman (1993) and by Robins and Hayes (1993). A summary of the main findings is provided here. Research has attempted to differentiate between mediators and moderators of change in cognitive therapy. Mediators are the mechanisms or the patients' characteristics that are changed by treatment and which precede change in the dependent variable of interest (level of depression). In contrast, moderators are the variables that predict treatment outcome, for example therapists' competence level and patients' characteristics (gender, age, diagnosis).

6.03.4.1 Mediators of Change in Cognitive Therapy

Cognitive therapy aims to bring about improvement by changing the hypothesized cognitive components which maintain the disorder. Cognitive changes after treatment have been found to correlate with level of improvement (Blackburn & Bishop, 1983; DeRubeis et al., 1990; Rush, Beck, Kovacs, Weissenburger, & Hollon, 1982; Seligman et al., 1988). Garamoni, Reynolds, Thase, Frank, and Fasiezka (1992) reported that the balance of positive to negative cognitions and related affect changed to an optimal range in responders to cognitive therapy, but not in nonrespondents. Persons and Burns (1985; 1986) also found that changes in negative-automatic thoughts within session were highly correlated with within-session changes in mood. The covariation between change in cognition and change in depression has been found in studies using self-rating measures (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), as well as observer ratings of depression (Hamilton, 1960).

However, the effect of cognitive therapy on cognitive mediational variables is not unique to cognitive therapy, pharmacotherapy having the same effect (Simons et al., 1984). Blackburn and Bishop (1983) attributed the larger effect of cognitive therapy relative to pharmacotherapy on views of self, the world, and the future to the larger effect on level of depression at the end of treatment in patients treated with cognitive therapy relative to patients treated with pharmacotherapy. Hollon, DeRubeis, and Evans (1987) argued that covariation between change in cognitive variables and change in depression is not sufficient to prove that change in cognitive variables plays a causal mediational role in the recovery process with cognitive therapy. In addition, there must be a primary effect, that is, cognitive change must precede change in depression and experimental manipulation of the degree of change in cognition must correspond to the degree of change in depression. Few studies have succeeded in demonstrating the primary change in cognition. Rush, Kovacs, Beck, Weissenburger, and Hollon (1981), using cross-lagged correlations, found that during the first four of 11 weeks of cognitive therapy, improvement in hopelessness, in view of self, and in mood preceded changes in vegetative and motivational symptoms of depression. No specific pattern of change was found in patients treated with pharmacotherapy. DeRubeis et al. (1990) reported that changes in cognitive variables (attributional style, dysfunctional attitudes, and hopelessness) at midtreatment predicted overall improvement at the end of treatment with cognitive therapy, but not with pharmacotherapy. If this effect is replicated in future studies, these authors' conclusion, that cognitive constructs play a mediational role in cognitive therapy, but that this effect is not sufficient as it was not found in pharmacotherapy which was equally effective, will be an important one in the understanding of the mode of action of cognitive therapy and of medication.

Experimentally, a number of studies have indicated that the manipulation of cognitive content, using the Velten procedure (1968) of reading self-referent sad statements, induces sad mood (Coleman, 1975; Hale & Strickland, 1976; Teasdale & Fogarty, 1979). These studies have been much criticized because of the experimental bias introduced by the inherent demand characteristics of the task. Perhaps more convincingly, Teasdale and Fennell (1982) compared effects on mood produced by 30 minutes of exploration compared with 30 minutes of active modification of depressive thoughts. They found, in a within-subjects design, that active modification led to a greater reduction of self-rated depression.

When cognitive therapy has been broken down into components to analyze the different aspects of cognitive therapy, mixed results have been obtained. McNamara and Horan (1986) found that cognitive procedures reduced depressive cognitions and improved social skills more than behavioral procedures. Jarrett and Nelson (1987) divided the cognitive therapy package into three components: self-monitoring, logical analyses, and hypothesis testing. All subjects received each treatment component in one of two sequences. The results indicated that self-monitoring did not bring about a change in symptoms, but logical analysis and hypothesis testing brought about a decrease in depressive symptoms and in negative automatic thoughts, a better level of interpersonal relationships, and a greater frequency of pleasant activities. The two active components, logical analysis and hypothesis testing, were more effective in combination than singly. Jacobson et al. (1996) randomly allocated 150 outpatients with major depression to partial cognitive therapy (behavioral component and modifying automatic thoughts) or to full cognitive therapy (behavioral component, modifying automatic thoughts, and modifying core schemata). They found partial cognitive therapy as effective as full cognitive therapy at the end of treatment and at six months follow-up. Both the behavioral component and the automatic component were as effective as full cognitive therapy in altering negative thinking and dysfunctional attributional style. These results are in contrast with McNamara and Horan's (1986) and Jarrett and Nelson's (1987) findings, indicating that more work is required in this area.

The debate about what type or level of cognitive change actually is achieved in cognitive therapy is unresolved. Persons (1993), in a theoretical paper, considers whether cognitive therapy changes the basic schemata which are presumed vulnerability factors or only teaches compensatory skills. The two models generate different predictions regarding the timing of change (which would occur earlier in the schema change model) and the generalizability of what is learned in therapy (the compensatory skills model providing more general skills which might provide more protection against future episodes of illness). The two models are empirically testable and answers may be provided in future studies.

6.03.4.2 Moderators of Change in Cognitive Therapy

The Cognitive Therapy Scale (CTS; Young & Beck, 1988) is a 13-item rating scale which evaluates several aspects of cognitive therapy, the therapeutic alliance, adherence to the procedures of cognitive therapy (including the structure and content of therapy), and competence or skill in application of the methods and techniques. The CTS has been found to have adequate psychometric properties (Dobson, Shaw, & Vallis, 1985; Vallis, Shaw, & Dobson, 1986), but the relationship of ratings on the CTS and outcome has not been demonstrated.

The quality of the therapeutic relationship has been found by some researchers to be positively related to outcome (DeRubeis et al., 1990; Persons & Burns, 1985). In contrast, Bercham (1989) found no relationship between therapeutic alliance ratings during early sessions of cognitive therapy and outcome in the treatment of depression. DeRubeis and Feeley (1990) found that the "facilitative" aspects of cognitive therapy (empathy, warmth, and understanding) and a measure of the helping alliance did not predict change in depression level after the sessions in which they were rated. However, ratings of the helping alliance made during later sessions of therapy were related to prior change in depression symptoms.

Several studies have demonstrated that cognitive therapy can be discriminated reliably from other psychotherapies, for example, interpersonal psychotherapy (DeRubeis et al., 1990; DeRubeis, Hollon, Evans, & Bemis, 1982; Hill, O'Grady, & Elkin, 1992), indicating that the procedures of cognitive therapy are specific. Only one study, by DeRubeis and Feeley (1990), has examined the relationship between adherence to different aspects of the treatment protocol and outcome. These authors found that "concrete" and "symptom-focused" methods of cognitive therapy predicted outcome, whereas "abstract" discussions did not.

Level of competence has been reported to be related to outcome (Beckham, 1990; Burns & Nolen-Hoeksema, 1992; Hollon, Shelton, & Davis, 1993), but contrary findings have been reported in psychotherapy (Shapiro et al., 1994). The issue of level of competence of therapists is of evident importance for outcome studies and for training courses.

Since therapy involves a diadic relationship, patient's characteristics may be of equal importance as therapists' characteristics for outcome. Sociodemographic variables have usually been found not to be important in determining response to cognitive therapy, except that married status appears to be a positive predictor (Jarrett, Eaves, Grannemann, & Rush, 1991; Sotsky et al., 1991).

In terms of illness characteristics, several studies have shown that the endogenous subtype of depression responds as well to cognitive therapy as to antidepressant medication (Blackburn et al., 1981; Imber et al., 1990; Kovacs, Rush, Beck, & Hollon, 1981). Severity of depression was found to be a negative predictor of response by Elkin et al. (1989), but other studies have failed to replicate this finding (Hollon et al., 1992; McLean & Taylor, 1992; Thase, Simon, Cahalance, McGreary, & Harden, 1991). Since cognitive therapy targets cognitive dysfunction, it might be hypothesized that cognitive therapy would be relatively more effective than control treatments (e.g., pharmacotherapy) for patients with higher levels of cognitive dysfunction. In a review paper, Rude and Rehm (1991) conclude that research findings do not support this prediction.

In terms of personal characteristics, level of intelligence has not been found to be related to outcome (Haaga et al., 1991) and the presence of a concurrent Axis II disorder (personality disorder) has been found to be related to more residual symptoms of depression at outcome and poorer social functioning (Shea et al., 1990). Simons, Lustman, Wetzel, and Murphy (1985) examined a number of predictor variables in an outcome study of depression and found that a cognitive variable, learned resourcefulness, was the only reliable predictor. High learned resourcefulness (as measured by the Self-Control Schedule; Rosenbaum, 1980) predicted better response to cognitive therapy and low learned resourcefulness predicted better outcome with pharmacotherapy. Learned resourcefulness refers to a set of coping and problem-solving skills that facilitate the monitoring, control, and change of dysfunctional or unpleasant events. However, the finding by Simons et al. (1985) has not been replicated (Beckham, 1989; Jarrett, Giles, Gullion, & Rush, 1991; Kavanagh & Wilson, 1989).

The short preceding review of the process of action of cognitive therapy indicates that findings are limited as they relate primarily to depression. Both mediating and moderating variables remain relatively unclear or unreplicated.

6.03.5 EFFICACY

The efficacy of cognitive therapy relative to other treatment methods has been well demonstrated, as can be seen in the various chapters of this volume relating to specific disorders. Of particular interest is the long-term or prophylactic effect of cognitive therapy. Cognitive therapy aims to decrease psychological vulnerabilities by targeting the underling cognitive structures which are hypothesized to make up these vulnerabilities. Whether it succeeds in doing so by modifying the underlying structures (schemata) or by teaching coping skill (Persons & Miranda, 1992), it would be expected to reduce the likelihood of relapse or recurrence of illness after successful treatment. In general, this expectation has been fulfilled in follow-up studies of depression (Beck et al., 1985; Blackburn et al., 1986; Evans et al., 1991; Kovacs et al., 1981; Shea et al., 1992; Simons et al., 1986). Fewer long-term studies have found the same prophylactic effect in general anxiety (Borkovec & Costello, 1993; Butler et al., 1991; Chambless & Gillis, 1993) and in panic disorder (Clark et al., 1994; Craske, Brown, & Barlow, 1991; Shea, Pilkonis, Cloitre, & Leon, 1994).

Unfortunately, all follow-up studies have been primarily naturalistic, so that no control was exercised on what actually happened to the patients after the end of treatment of the acute episode. However, the data from various studies are sufficiently congruent to indicate that cognitive therapy does have a prophylactic effect. Better controlled studies are awaited.

6.03.6 CONCLUSION

In this chapter, a general overview of the theories and research underlying cognitive therapy has been given. The application of cognitive therapy indicates an on-going cross-fertilization between evolving theories, research, and practice. The fact that cognitive therapy has not remained static since its inception is a healthy sign of a growing and living theory. It is evident that future research is needed in information processing, in the bidirectional interaction between emotion and cognition and in the process of therapy to elucidate many of the issues raised in this chapter which still remain open. Moreover, as the methodology of outcome research improves and assessment measures get more sophisticated, clinicians will be able to make better predictions about which patient responds best to cognitive therapy. The expanding application of cognitive therapy remains to be tested.

6.03.7 REFERENCES

Abramson, L. Y., Alloy, L. B., & Metalsky, G. I. (1988). The cognitive diathesis-stress theories of depression: Toward an adequate evaluation of the theories' validities. In L. E. Alloy (Ed.), *Cognitive processes in depression* (pp. 3–30). New York: Guilford Press.

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87,* 49–74.

Adler, A. (1936). The neurotic's picture of the world. *International Journal of Individual Psychology, 2,* 3–10.

Alloy, L. B., & Abramson, L. Y. (1979). Judgement of contingency in depressed and non-depressed students: Sadder but wiser? *Journal of Experimental Psychology, 108,* 441–485.

Anderson, K. J. (1990). Arousal and the inverted U-hypothesis: A critique of Neiss's "Reconceptualising arousal." *Psychological Bulletin, 107,* 96–100.

Arnold, M. (1960). *Emotions and personality* (Vol. 1). New York: Columbia University Press.

Bartlett, F. E. (1932). *Remembering: A study in experimental and social psychology*. London: Cambridge University Press.

Barlow, D. H., Craske, M. G., Cerny, J. A., & Klosko, J. S. (1989) Behavioural treatment for panic disorder. *Behaviour Therapy, 20,* 261–282.

Basco, M. R., & Rush, A. J. (1996). *Cognitive behavioral therapy for bipolar disorder*. New York: Guilford Press.

Beck, A. T. (1963). Thinking and depression. I: Idiosyncratic content and cognitive distortions. *Archives of General Psychiatry, 9,* 324–333.

Beck, A. T. (1967). *Depression: Clinical, experimental and theoretical aspects*. New York: Harper & Row.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.

Beck, A. T. (1983). Cognitive therapy and depression: New perspectives. In P. J. Clayton & J. E. Barrett (Eds.), *Treatment of depression: Old controversies and new approaches* (pp. 265–284). New York: Raven Press.

Beck, A. T. (1987). Cognitive models of depression. *Journal of Cognitive Psychotherapy, 1,* 5–37.

Beck, A. T., Emery, G., & Greenberg, R. L. (1985). *Anxiety disorders and phobias: A cognitive perspective*. New York: Basic Books.

Beck, A. T., Epstein, H., & Harrison, R. (1983). Cognitions, attitudes and personality dimensions in depression. *British Journal of Cognitive Psychotherapy, 1,* 1–16.

Beck, A. T., & Freeman, A. et al. (1990). *Cognitive therapy of personality disorders*. New York: Guilford Press.

Beck, A. T., Hollon, S. D., Young, J. E., Bedrosian, R. C., & Budenz, D. (1985). Treatment of depression with cognitive therapy and amitriptylene. *Archives of General Psychiatry, 42,* 144–148.

Beck, A. T., Laude, R., & Bohnert, M. (1974). Ideational components or anxiety neurosis. *Archives of General Psychiatry, 31,* 319–326.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression: A treatment manual.* New York: Guilford Press.

Beck, A. T., Sokol, L., Clark, D. A., Berchick, B. L., & Wight, F. (1992). Focused cognitive therapy of panic disorder: A crossover design and one-year follow-up. *American Journal of Psychiatry, 147,* 778–783.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. E., & Erbaugh, J. K. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561–571.

Beck, A. T., Wright, F. W., Newman, C. F., & Liese, B. (1993). *Cognitive therapy of substance abuse*. New York: Guilford.

Beck, J. G., Stanley, M. A., Baldwin, L. E., Deagle, E. A., & Averill, P. M. (1994). Comparison of cognitive therapy and relaxation training for panic disorder. *Journal of Consulting and Clinical Psychology, 62,* 818–826.

Beck, J. S. (1995). *Cognitive therapy. Basics and beyond.* New York: Guilford Press.

Beckham, E. E. (1989). Improvement after evaluation in psychotherapy of depression: Evidence of a placebo effect? *Journal of Clinical Psychology Review, 45,* 945–950.

Beckham, E. E. (1990). Psychotherapy of depression.

Research at a cross-roads: Directions for the 1990s. *Clinical Psychology Review, 10,* 207–228.

Bernstein, B. A., & Borkovec, T. D. (1973). *Progressive relaxation training: A manual for the helping professions.* Champaign, IL: Research Press.

Birchwood, M., & Tarrier, N. (1994). *Psychological management of Schizophrenia.* Chichester, UK: Wiley.

Blackburn, I.-M. (1986). The Cognitive Revolution: An ongoing Evolution. *Behavioural Psychotherapy, 14,* 274–277.

Blackburn, I.-M. (1993). *Coping with depression* (5th ed.). Edinburgh, UK: Chambers.

Blackburn, I.-M. (1996). Cognitive vulnerability to depression. In P. M. Salkovskis (Ed.), *Frontiers of cognitive therapy* (pp. 250–265). New York: Guilford Press.

Blackburn, I.-M., & Bishop, S. (1983). Changes in cognition with pharmacotherapy and cognitive therapy. *British Journal of Psychiatry, 143,* 609–617.

Blackburn, I.-M., Bishop, S., Glen, A. I. M., Whalley, L. J., & Christie, J. E. (1981). The efficacy of cognitive therapy in depression: A treatment trial using cognitive therapy and pharmacotherapy, each alone and in combination. *British Journal of Psychiatry, 139,* 181–189.

Blackburn, I.-M., Eunson, K. M., & Bishop, S. (1986). A two-year naturalistic follow-up of depressed patients treated with cognitive therapy, pharmacotherapy and a combination of both. *Journal of Affective Disorders, 10,* 67–75.

Blackburn, I.-M., Jones, S., & Lewin, R. J. P. (1986). Cognitive style in depression. *British Journal of Clinical Psychology, 25,* 241–251.

Blackburn, I.-M., Roxborough, H. M., Muir, W. J., Glabus, M., & Blackwood, J. D. R. (1990). Perceptual and physiological dysfunction in depression. *Psychological Medicine, 20,* 95–103.

Blackburn, I.-M., & Twaddle, V. et al. (1996). *Cognitive therapy in action. A practitioner's casebook.* London: Souvenir Press.

Blaney, P. H. (1986). Affect and memory: A review. *Psychological Bulletin, 99,* 229–246.

Borkovec, T. D., & Costello, E. (1993). Efficacy of applied relaxation and cognitive behavioural therapy in the treatment of generalized anxiety disorder. *Journal of Consulting and Clinical Psychology, 61,* 611–619.

Bower, G. H. (1981). Mood and memory. *American Psychologist, 36,* 129–148.

Bower, G. H. (1987). Commentary on mood and memory. *Behaviour Research and Therapy, 25,* 443–455.

Bower, G. H. (1992). How might emotions effect learning? In S. A. Christianson (Ed.), *The handbook of emotion and memory: Research and theory.* Hillsdale, NJ: Erlbaum.

Breslow, R., Kocsis, J., & Belkin, B. (1991). Contribution of the depressive perspective to memory function in depression. *American Journal of Psychiatry, 138,* 227–230.

Brown, G. W., & Harris, T. (1978). *Social origins of depression.* London: Tavistock.

Burns, D. D., & Nolen-Hoeksema, S. (1992). Therapeutic empathy and recovery from depression in cognitive behavioural therapy: A structural equation model. *Journal of Consulting and Clinical Psychology, 60,* 441–449.

Butler, G., Fennell, M., Robson, P., & Gelder, M. (1991). Comparison of behaviour therapy and cognitive behaviour in the treatment of generalised anxiety disorders. *Journal of Consulting and Clinical Psychology, 59,* 167–175.

Butler, G., & Mathews, A. (1983). Cognitive processes in anxiety. *Advances in Behaviour Research and Therapy, 5,* 51–62.

Cautela, J. (1973). Covert processes and behavior modification. *Journal of Nervous and Mental Disease, 157,* 27–35.

Chambless, D. L., & Gillis, M. M. (1993). Cognitive therapy of anxiety disorders. *Journal of Consulting and Clinical Psychology, 61,* 248–260.

Clark, D. A., & Beck, A. T. (1991). Personality factors in dysphoria: A psychometric refinement of Beck's Sociotropy-Autonomy Scale. *Journal of Psychopathology and Behavioral Assessment, 13,* 369–388.

Clark, D. A., Steer, R. A., Beck, A. T., & Ross, L. (1995). Psychometric characteristics of revised sociotropy and autonomy scales in college students. *Behaviour Research and Therapy, 33,* 325–334.

Clark, D. M. (1989) Anxiety states: Panic and generalised anxiety. In K. Hawton, P. Salkovskis, J. Kirk, D. Clark (Eds.), *Cognitive behaviour therapy for psychiatric problems: A practical guide* (pp. 52–96). Oxford, UK: Oxford University Press.

Clark, D. M., Salkovskis, P. M., Hackmann, A., Middleton, H., Anastasiades, P., & Gelder, M. (1994). A comparison of cognitive therapy, applied relaxation and imipramine in the treatment of panic disorder. *British Journal of Psychiatry, 164,* 759–769.

Clark, D. M., & Teasdale, J. D. (1982). Diurnal variation in clinical depression and accessibility of memories of positive and negative experiences. *Journal of Abnormal Psychology, 91,* 87–95.

Clifford, P. I., & Hemsley, D. R. (1987). The influence of depression on the processing of personal attributes. *British Journal of Psychiatry, 150,* 95–103.

Coleman, R. E. (1975). Manipulation of self-esteem as a determinant of mood of elated and depressed women. *Journal of Abnormal Psychology, 84,* 693–700.

Coyne, J., & Gotlib, I. (1983). The role of cognition in depression: A critical appraisal. *Psychological Bulletin, 94,* 472–505.

Craske, M. G., Brown, T. A., & Barlow, D. H. (1991). Behavioural treatment of panic disorder: A two-year follow-up study. *Behaviour Therapy, 22,* 289–304.

Dalgleish, T., & Watts, F. N. (1990). Biases of attention and memory in disorder of anxiety and depression. *Clinical Psychology Review, 10,* 589–604.

Dember, W. (1974). Motivation and the cognitive revolution. *American Psychologist, 29,* 161–168.

Derry, P. A., & Kuiper, N. A. (1981). Schematic processing and self reference in clinical depression. *Journal of Abnormal Psychology, 90,* 286–297.

DeRubeis, R. J., Evans, M. D., Hollon, S. D., Garvey, M. J., Grove, W. M., & Tuason, V. B. (1990). How does cognitive therapy work. Cognitive change and symptom change in cognitive therapy and pharmacotherapy for depression. *Journal of Consulting and Clinical Psychology, 58,* 862–869.

DeRubeis, R. J., & Feeley, M. (1990). Determinants of change in cognitive therapy for depression. *Cognitive Therapy and Research, 14,* 469–482.

DeRubeis, R. J., Hollon, S. D., Evans, M. D., & Bemis, K. M. (1982). Can psychotherapies for depression be discriminated? A systematic investigation of cognitive, therapy and interpersonal therapy. *Journal of Consulting and Clinical Psychology, 50,* 744–756.

Dobson, K. S. (1995). *Handbook of cognitive-behavioral therapies.* New York: Guilford Press.

Dobson, K. S., Shaw, B. F., & Vallis, T. M. (1985). Reliability of a measure of the quality of cognitive therapy. *British Journal of Clinical Psychology, 24,* 295–300.

Durham, R. C., Murphy, I. J. C., Allan, T., Richard, K., Treliving, L. R., & Fenton, G. W. (1994). A comparison of cognitive therapy, analytic psychotherapy and anxiety management training in the treatment of generalised anxiety disorder. *British Journal of Psychiatry, 165,* 315–323.

Durham, R. C., & Turvey, A. A. (1987). Cognitive therapy vs. behaviour therapy in the treatment of chronic general anxiety. *Behaviour Research and Therapy, 25,* 229–235.

Edwards, D. J. A. (1989), Restructuring through guided imagery: Lessons from gestalt therapy. In A. Freeman, K. M. Simon, L. E. Beutler, & H. Arkowitz (Eds.), *Comprehensive handbook of cognitive therapy* (pp. 283–297). New York: Plenum.

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., Glass, D. R., Pilkoniz, P. A., Leber, W. R., Docherty, J. P., Fiester, S. J., & Parloff, M. B. (1989). NIMH treatment of depression collaborative research program: General effectiveness of treatments. *Archives of General Psychiatry, 46, 971–982.*

Ellis, A. (1957). Outcome of employing three techniques of psychotherapy. *Journal of Clinical Psychology,* 13, 344–350.

Ellis, A. (1962). *Reason and emotion in psychotherapy*. New York: Lyle Stuart.

Ellis, A. (1973). Rational-emotive therapy. In R. Jurjevitch (Ed.), *Direct psychotherapy* (pp. 295–331). Coral Gables, FL: University of Miami Press.

Ellis, A. (1980). Rational-emotive therapy and cognitive behavior therapy: Similarities and differences. *Cognitive Therapy and Research, 4,* 325–340.

Ellis, A. (1989). The history of cognition in psychotherapy. In A. Freeman, K. M. Simon, L. E. Beutler, & H. Arkowtiz (Eds.), *Comprehensive handbook of cognitive therapy* (pp. 5–19). New York: Plenum.

Ellis, A. (1993). Reflections on rational-emotive therapy. *Journal of Consulting and Clinical Psychology, 61,* 199–201.

Ellis, A., & Bernard, M. E. (1985). What is rational-emotive therapy (RET)? In A. Ellis & M. E. Bernard (Eds.), *Clinical applications of rationale-emotive therapy* (pp. 1–30). New York: Plenum.

Ellis, A., & Whiteley, J. M. (Eds.) (1970). *Theoretical and empirical foundations of rational-emotive therapy*. Monterey, CA: Brooks Cole.

Emmelkamp, P. M., & Beens, H. (1991). Cognitive therapy with obsessive compulsive disorder: A comparative evaluation. *Behaviour Research and Therapy, 29,* 293–300.

Emmelkamp, P. M., Van der Helm, M., Van Zainten, B. L., & Ploghg, I. (1980) Treatment of obsessive-compulsive patients: The contribution of self-instructional training to the effectiveness of exposure. *Behaviour Research and Therapy, 18,* 61–66.

Evans, M. D., Hollon, S. D., DeRubeis, R. J., Piasecki, J. M., Grove, W. M., Garvey, M. J., & Tuason, V. B. (1992). Differential relapse following cognitive therapy and pharmacotherapy for depression. *Archives of General Psychiatry, 49,* 802–808.

Eysenck, H. J. (1970). *The structure of human personality* (3rd ed.). London: Methuen.

Fowler, D., Garety, P., & Kuiper, E. (1995). *Cognitive behaviour therapy for psychosis. Theory and practice*. Chichester, UK: Wiley.

Gallagher, D. E., & Thompson, L. W. (1982). Treatment of major affective disorder in older adult outpatients with brief psychotherapies. *Psychotherapy, Theory, Research and Practice, 19,* 82–490.

Garamoni, G. L., Reynolds, C. F., Thase, M. E., Frank, E., & Fasiezka, A. L. (1992). Shifts in affective balance during cognitive therapy of major depression. *Journal of Consulting and Clinical Psychology, 60,* 260–266.

Gilbert, P., & Reynolds, S. (1990). The relationship between the Eysenck Personality Questionnaire and Beck's concepts of sociotropy and autonomy. *British Journal of Clinical Psychology, 29,* 319–325.

Goldfried, M. R., & Davison, G. C. (1976). Clinical behavior therapy. New York: Holt, Rinehart and Winston.

Greenberger, D., & Padesky, C. (1995). *Mind over mood. A cognitive therapy treatment manual for clients*. New York: Guilford Press.

Guidano, V. F., & Liotti, G. (1983). *Cognitive processes and emotional disorders: A structural approach to psychotherapy*. New York: Guilford Press.

Haaga, D. A. F., & Davison, G. C. (1989) Outcome studies of rational-emotive therapy. In M. E. Bernard & R. DiGuiseppe (Eds.), *Inside rationale-emotive therapy: A critical appraisal of the theory and therapy of Albert Ellis* (pp. 155–197). San Diego, CA: Academic Press.

Haaga, D. A. F., & Davison, G. C. (1993) An appraisal of rational-emotive therapy. *Journal of Consulting and Clinical Psychology, 61,* 215–220.

Haaga, D. A., Dyck, M. J., & Ernst, D. (1991). Empirical status of cognitive theory of depression. *Psychological Bulletin, 110,* 215–236.

Hale, W. D., & Strickland, B. R. (1976). Induction of mood states and their effect on cognitive and social behaviors. *Journal of Clinical and Consulting Psychology, 44,* 55.

Hamilton, M. A. (1960). A rating scale for depression. *Journal of Neurological and Neurosurgical Psychiatry, 23,* 56–62.

Hamilton, E. W., & Abramson, L. Y. (1983). Cognitive patterns and major depressive disorder: A longitudinal study in a hospital setting. *Journal of Abnormal Psychology, 92,* 173–184.

Hammen, C., Ellicott, A., Gitlin, M., & Jamison, K. R. (1989). Sociotropy/autonomy and vulnerability to specific life events in patients with unipolar depression and bipolar disorders. *Journal of Abnormal Psychology, 98,* 154–160.

Hawton, K., Salkovskis, P. M., Kirk, J., & Clark, D. M. (1989). *Cognitive behaviour therapy for psychiatric problems. A practical guide*. Oxford, UK: Oxford University Press.

Hibbert, G. A. (1984). Ideational components of anxiety: Their origin and content. *British Journal of Psychiatry, 144,* 618–624.

Hill, G. E., O'Grady, K. E., & Elkin, I. (1992). Applying the collaborative study Psychotherapy Rating Scale to rate therapist adherence in cognitive-behavior therapy, interpersonal therapy, and clinical management. *Journal of Consulting and Clinical Psychology, 60,* 73–79.

Hollon, S. D., DeRubeis, R. J., & Evans, M. D. (1987). Causal medication of change in treatment of depression: Discrimination between non-specificity and non-causality. *Psychological Bulletin, 102,* 139–149.

Hollon, S. D., DeRubeis, R. J., Evans, M. D., Wiemer, M. J., Garvey, M. J., Grove, W. M., & Tuason, V. B. (1992). Cognitive therapy and pharmacotherapy for Depression singly and in combination. *Archives of General Psychiatry, 49,* 774–781.

Hollon, S. D., Shelton, R. C., & Davis, D. D. (1993). Cognitive therapy for depression: Conceptual issues and clinical efficiency. *Journal of Consulting and Clinical Psychology, 61,* 270–275.

Homme, L. (1965). Perspectives in psychology: Control of coverants, the operants of the mind. *Psychological Record, 15,* 501–511.

Horney, K. (1950). *Neurosis and human growth: The struggle toward self-realization*. New York: Norton.

Imber, S. D., Pilkonis, P. A., Sotsky, S. M., Elkin, I., Watkins, J. T., Collins, J. F., Shea, M. T., Leber, W. R., & Glass, D. R. (1990). Mode-specific effect effects among three treatments of depression. *Journal of Consulting and Clinical Psychology, 58,* 352–359.

Ingram, R. E. (Ed.) (1988). *Information processing approaches to clinical psychology*. Orlando, FL: Academic Press.

Ingram, R. E., & Kendall, P. C. (1988). Cognitive clinical psychology. Implications of an information processing perspective. In R. E. Ingram (Ed.), *Information processing approaches to clinical psychology* (pp. 3–21). Orlando, FL: Academic Press.

Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., Gortner, E., & Prince, S. E. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 64,* 295–304.

Jarrett, R. B., Eaves, G. G., Grannemann, B. D., & Rush, A. J. (1991). Clinical, cognitive and demographic predictors of response to cognitive therapy for depression: A preliminary report. *Psychiatry Research, 37,* 245–260.

Jarrett, R. B., Giles, D. A., Gullion, C. M., & Rush, A. J. (1991). Does learned resourcefulness predict response to cognitive therapy in depressed outpatients? *Journal of Affective Disorders, 23,* 223–229.

Jarrett, R. B., & Nelson, R. D. (1987). Mechanisms of change in cognitive therapy of depression. *Behavior Therapy, 18,* 227–241.

Jones, R. (1968). *A factored measure of Ellis's irrational belief system.* Unpublished doctoral dissertation. Texas Technological College.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.) (1982). *Judgement under uncertainty: Heuristics and biases.* Cambridge, UK: Cambridge University Press.

Kavanagh, D. J., & Wilson, P. H. (1989). Prediction of outcome with group cognitive therapy for depression. *Behaviour Research and Therapy, 27,* 333–43.

Kelly, G. (1955). *The psychology of personal constructs.* New York: Norton.

Kelly, G. A. (1977). The psychology of the unknown. In D. Bannister (Ed.), *New perspectives in personal construct theories* (pp. 1–19). San Diego, CA: Academic Press.

Kendall, K. D. (1984). Cognitive processes and procedures in behavior therapy. In C. M. Franks, G. T. Wilson, P. C. Kendall, & K. D. Brownell (Eds.), *Annual review of behavior therapy; theory and practice* (Vol. 10, pp. 123–163). New York: Guilford Press.

Kingdon, D. G., & Turkington, D. (1994). *Cognitive behaviour therapy for Schizophrenia.* Hove, UK: Erlbaum.

Kovacs, M., Rush, A. J., Beck, A. T., & Hollon, S. D. (1981). Depressed outpatients treated with cognitive therapy and pharmacotherapy. A one-year follow-up. *Archives of General Psychiatry, 38,* 33–39.

Lazarus, R. S. (1966). *Psychological stress and the coping process.* New York: McGraw-Hill.

Layden, M. A., Newman, C. F., Freeman A., & Morse, S. B (1993). *Cognitive therapy of borderline personality disorder.* Boston: Allyn & Bacon.

Linehan, M. M. (1993). *Skills training manual for treating borderline personality disorder.* New York: Guilford Press.

MacLeod, C. (1991). Clinical anxiety and the selective encoding of threatening information. *International Review of Psychiatry, 3,* 279–292.

Mahoney, M. J. (1974). *Cognition and behaviour modification.* Cambridge, UK: Ballinger.

Mahoney, M. J. (1993). Introduction to special section: Theoretical development in the cognitive psychotherapies. *Journal of Consulting and Clinical Psychology, 61,* 187–193.

McLean, P., & Taylor, S. (1992) Severity of unipolar depression and choice of treatment. *Behaviour Research and Therapy, 30,* 443–451.

McNamara, K., & Horan, J. J. (1986). Experimental construct validity in the evaluation of cognitive and behavioral treatments for depression. *Journal of Counselling Psychology, 33,* 23–30.

Meichenbaum, D. (1977). *Cognitive-behavior modification. An integrative approach.* New York: Plenum.

Meichenbaum, D. (1993). Changing conceptions of cognitive behavior modification: Retrospect and prospect. *Journal of Consulting and Clinical Psychology, 61,* 202–204.

Mineka, S. (1992). Evolutionary memories, emotional processing, and the emotional disorders. In D. Medlin (Ed.), *The psychology of learning and motivation* (Vol. 28, pp. 161–206). New York: Academic.

Miranda, J., & Persons, J. B. (1988). Dysfunctional attitudes are mood state dependent. *Journal of Abnormal Psychology, 97,* 76–79.

Miranda, J., Persons, J. B., & Buyers, C. (1990). Endorsement of dysfunctional beliefs depends on current mood state. *Journal of Abnormal Psychology, 92,* 237–241.

Mogg, K., Mathews, A. M., May, J., Grove, M., Eysenck, M., & Weinman, J. (1991). Assessment of cognitive bias in anxiety and depression using a colour perception task. *Cognition and Emotion, 5,* 221–238.

Mogg, K., Mathews, A. M., & Weinman, J. (1987). Memory bias in clinical anxiety. *Journal of Abnormal Psychology, 96,* 94–98.

Moore, R. G., & Blackburn, I.-M. (1994). The relationship of sociotropy and autonomy to symptoms, cognition and personality in depressed patients. *Journal of Affective Disorders, 32,* 239–245.

Moore, R. G., & Blackburn, I.-M. (1996). The stability of sociotropy and autonomy in depressed patients undergoing treatment. *Cognitive Therapy and Research, 20,* 69–80.

Murphy, G. E., Simons, A. D., Wetzel, R. D., & Lustman, P. J. (1984) Cognitive therapy and pharmacotherapy, singly and together in the treatment of depression. *Archives of General Psychiatry, 41,* 33–41.

Neimeyer, R. A. (1992). Constructivist approaches to the measurement of meaning. In G. J. Neimeyer (Ed.), *Casebook of constructivist assessment* (pp. 58–103). Newbury Park, CA: Sage.

Neimeyer, R. A. (1993). An appraisal of constructivist psychotherapies. *Journal of Consulting and Clinical Psychotherapy, 61,* 221–234.

Neisser, U. (1976). *Cognition and reality.* San Francisco: Freeman.

Overholser, J. C. (1993a). Elements of the Socratic method: I. Systematic questioning. *Psychotherapy, 30,* 67–74.

Overholser, J. C. (1993b). Elements of the Socratic method: II. Inductive reasoning, *Psychotherapy, 30,* 75–85.

Ost, L. B. (1987). Applied relaxation: description of a coping technique and review of controlled studies. *Behaviour Research and Therapy, 25,* 397–410.

Padesky, C. A. (1990). Schema as self-prejudice. *International Cognitive Therapy Newsletter, 6,* 22–23.

Padesky, C. A. (1994). Schema change processes in cognitive therapy. *Clinical Psychology and Psychotherapy, 1,* 267–278.

Padesky, C. A., & Greenberger, W. (1995). *Clinician's guide to mind over mood.* New York: Guilford Press.

Persons, J. B. (1989). *Cognitive therapy in practice. A case formulation approach.* New York: Norton.

Persons, J. B. (1993). The process of change in cognitive therapy: Schema change or acquisition of compensatory skills? *Cognitive Therapy and Research, 17,* 123–137.

Persons, J. B., & Burns, D. D. (1985). Mechanisms of action of cognitive therapy: The relative contributions of technical and interpersonal interventions. *Cognitive Therapy and Research, 9,* 539–551.

Persons, J. B., & Burns, D. D. (1986). The process of cognitive therapy: The first dysfunctional thought changes less than the last one. *Behavior Research Therapy, 24,* 619–624.

Persons, J. B., & Miranda, J. (1992). Cognitive theories of vulnerability to depression: Reconciling negative evidence. *Cognitive Therapy and Research, 16,* 485–502.

Piaget, J. (1950). *The psychology of intelligence.* New York: Harcourt Brace.

Piaget, J. (1972). *The principles of genetic epistemology.* London: Routledge & Kegan Paul.

Piaget, J. (1977). *The development of thought.* New York: Viking Penguin.

Power, K. G., Simpson, R. J., Swanson, V., Wallace, L. A., Feistner, A. C. T., & Sharp, D. (1990). A controlled comparison of cognitive behaviour therapy, diazepam and placebo, alone and in combination for the treatment of generalised anxiety disorder. *Journal of Anxiety Disorders, 4,* 267–292.

Rachman, S. (1981). The primacy of affect: Some theoretical implications. *Behaviour Research and Therapy, 19,* 279–296.

Rachman, S. (1984). A reassessment of the "primacy of affect." *Cognitive Therapy and Research, 8,* 579–584.

Raitt, A. (1988). Weight Control: A rational-emotive approach. In W. Dryden & P. Trower (Eds.), *Developments in rational-emotive therapy* (pp. 197–209). Milton Keynes, UK: Open University Press.

Robins, C. J. (1990). Crongruence of personality and life events in depression. *Journal of Abnormal Psychology, 99,* 393–397.

Robins, C. J., & Block, P. (1988). Personal vulnerability, life events, and depressive symptoms: A test of a specific interactional model. *Journal of Personality and Social Psychology, 54,* 847–852.

Robins, C. J., & Hayes, A. M., (1993). An appraisal of cognitive therapy. *Journal of Consulting and Clinical Psychology, 61,* 205–214.

Rosenbaum, M. (1980). A schedule for assessing self-control behaviors: Preliminary findings. *Behavior Therapy, 11,* 109–121.

Rude, S. S., & Rehm, L. P. (1991). Response to treatment for depression: The role of initial status on targeted cognitive and behavioral skills. *Clinical Psychology Review, 11,* 493–514.

Rush, A. J., Beck, A. T., Kovacs, M, & Hollon S. D. (1977). Comparative efficacy of cognitive therapy versus pharmacotherapy in outpatient depression. *Cognitive Therapy and Research, 1,* 17–37.

Rush, A. J., Beck, A. T., Kovacs, M. Weissenburger, J., & Hollon, S. D. (1982). Comparison of the effects of cognitive therapy and pharmacotherapy on hopelessness and self-concept. *American Journal of Psychiatry, 139,* 862–866.

Rush, A. J., Kovacs, M., Beck, A. T., Weissenburger, J., & Hollon, S. D. (1981). Differential effects of cognitive therapy and pharmacotherapy on depressive symptoms. *Journal of Affective Disorders, 3,* 221–229.

Rush, A. J., Weissenberger, A. J., & Eaves, G. (1986). Do thinking patterns predict depression symptoms? *Cognitive Therapy and Research, 10,* 225–236.

Safran, J. D., & Segal, Z. V. (1990). *Interpersonal processes in cognitive therapy.* New York: Basic Books.

Salkovskis, P. M. (1986). The cognitive revolution: New way forward, backward somersault or full circle? *Behavioural Psychotherapy, 14,* 278–282

Schwartz, R. M. (1986). The internal dialogue: On asymmetry between positive and negative coping thoughts. *Cognitive Therapy and Research, 6,* 591–605.

Segal, Z. V. (1988). Appraisal of the self-schema construct in major depression. *Psychological Bulletin, 103,* 147–162.

Seligman, M. E. P. (1975). *Helplessness: On depression, development and death.* San Francisco: Freeman.

Seligman, M. E. P., Castellon, C., Cacciola, J., Schulman, P., Luborsky, L., Ollove, M., & Downing, R. (1988). Explanatory style change during cognitive therapy for unipolar depression. *Journal of Abnormal Psychology, 97,* 13–18.

Seligman, M. E. P., & Maier, S. F. (1967). Failure to escape traumatic shock. *Journal of Experimental Psychology, 74,* 1–9.

Shapiro, D. A., Harper, H., Startup, M., Reynolds, S., Bird, D., & Suskas, A. (1994). The high-water work of the drug metaphor: A meta-analytic critique of process-outcome research. In R. L. Russell (Ed.), *Reassessing psychotherapy research* (pp. 1–35). New York: Guilford Press.

Shea, M. T., Elkin, I., Imber, S. D., Sotsky, S. M., Watkins, J. T., Collins, J. F., Pilkonis, P. A., Beckham, E., Glass, D. R., Dolan, R. T., & Parloff, M. B. (1992). Course of depressive symptoms over follow-up. *Archives of General Psychiatry, 49,* 782–787.

Shea, M. T., Pilkonis, P. A., Beckham, E., Collins, J. F., Elkin, I., Solsky, S. M., & Docherty, J. P. (1990). Personality disorders and treatment outcome in the NIMHT Treatment of depression Collaborative Research program. *American Journal of Psychiatry, 147,* 711–718.

Shear, M. K., Pilkonis, P. A., Cloitre, M., & Leon, A. C. (1994). Cognitive behavioural treatment compared with non prescriptive treatment of panic disorders. *Archives of General Psychiatry, 51,* 395–401.

Shorkey, C. T., & Whiteman, V. L. (1977). Development of the Rational Behavior Inventory: Initial validity and reliability. *Educational and Psychological Measurement, 37,* 527–534.

Simons, A. D., Garfield, S. L., & Murphy, G. E. (1984). The process of change in cognitive therapy and pharmacotherapy for depression: Changes in mood and cognition. *Archives of General Psychiatry, 41,* 45–51.

Simons, A. D., Lustman, P. J., Wetzel, R. D., & Murphy, G. E. (1985). Predicting response to cognitive therapy of depression: The role of learned resourcefulness. *Cognitive Therapy and Research, 9,* 79–89.

Simons, J. D., Murphy, G. E., Levine, J. L., & Wetzel, R. D. (1986). Cognitive therapy and pharmacotherapy for depression: Sustained improvement over one year. *Archives of General Psychiatry, 43,* 43–48.

Sotsky, S. M., Glass, D. R., Shea, M. T., Pilkonis, P. A., Collins, J. F., Elkin, I., Watkins, J. T., Imber, S. D., Leber, W. R., Moyer, J., & Oliveri, M. E. (1991). Patient predictors of response to psychotherapy and pharmacotherapy: Findings in the NIMH Treatment of Depression Collaborative Research Program. *American Journal of Psychiatry, 148,* 997–1009.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being. *Psychological Bulletin, 103,* 193–210.

Teasdale, J. D. (1988). Cognitive vulnerability to persistent depression. *Cognition and Emotion, 2,* 247–274.

Teasdale, J. D. (1993). Emotion and two kinds of meaning: Cognitive therapy and applied cognitive science. *Behaviour Research and Therapy, 31,* 339–354.

Teasdale, J. D., & Barnard, P. J. (1993). *Affect, cognition and change, Re-modelling depressive thought.* Hove, UK: Erlbaum.

Teasdale, J. D., & Dent, J. (1987). Cognitive vulnerability and depression: An investigation of two hypotheses. *British Journal of Clinical Psychology, 26,* 113–126.

Teasdale, J. D., & Fennell, M. J. V. (1982). Immediate effects on depression of cognitive therapy intervention. *Cognitive Therapy and Research, 6,* 343–351.

Teasdale, J. D., Fennell, M. J. V., Hibbert, G. A., & Amies, P. L. (1984) Cognitive therapy for major depressive disorders in primary care. *British Journal of Psychiatry, 144,* 400–406.

Teasdale, J. D., & Fogarty, S. J. (1979). Differential effects of induced mood on retrieval of pleasant and unpleasant events from episodic memory. *Journal of Abnormal Psychology, 88,* 248–257.

Teasdale, J. D., & Russell, M. L. (1983). Differential effects of induced mood on the recall of positive, negative and neutral words. *British Journal of Clinical Psychology, 22,* 163–171.

Teasdale, J. D., & Taylor, R. (1981). Induced mood and accessibility of memories: An effect of mood state or induction procedure? *British Journal of Clinical Psychology, 20,* 39–48.

Thase, M. E., Simons, A. D., Cahalance, J., McGreary, J., & Harden, T. (1991). Severity of depression and response to cognitive behavior therapy. *American Journal of Psychiatry, 148,* 784–789.

Ucros, C. G. (1989). Mood state-dependent memory: A meta-analysis. *Cognition and Emotion, 3,* 139–167.

Vallis, T. M., Shaw, B. F., & Dobson, K. S. (1986). The Cognitive Therapy Scale: Psychometric properties. *Journal of Consulting and Clinical Psychology, 54,* 381–385.

Van Oppen, P., de Haan, E., Van Balkom, A. T., Spinhoven, P., Hoogduin, K., & Van Dyck, R. (1995). Cognitive therapy and exposure *in vivo* in the treatment of obsessive compulsive disorder. *Behaviour Research and Therapy, 33* 379–390.

Velten, E. (1968). A laboratory task for induction of mood states. *Behavior Research and Therapy, 6,* 473–482.

Weissman, A. N., & Beck, A. T. (1978). *Development and validation of the dysfunctional attitude scale*. Paper presented at the annual meeting of the Association for the Advancement of Behaviour Therapy, Chicago.

Wells, A., & Matthews, G. (1994). *Attention and emotion. A clinical perspective*. Hove, UK: Erlbaum.

Whisman, M. A. (1993). Mediators and moderators of change in cognitive therapy of depression. *Psychological Bulletin, 114,* 248–265.

Wilson, P. H. (1992). *Principles and practice of relapse prevention*. New York: Guilford.

Williams, J. M. G. (1992). *The psychological treatment of depression* (2nd ed.). London: Routledge.

Williams, J. M. G., Watts, F. N., MacLeod, C., & Mathews, A. M. (1988). *Cognitive psychology and emotional disorders*. Chichester, UK: Wiley.

Wright, J. H., Thase, M. E., Beck, A. T., & Ludgate, J. W. (1993). *Cognitive therapy with inpatients: Developing a cognitive milieu*. New York: Guilford Press.

Young, J. E. (1990). *Cognitive therapy for personality disorders: A schema-focused approach*. Sarasoto, FL: Professional Resource Press.

Young, J. E., & Beck, A. T. (1988). *Cognitive Therapy Scale*. Unpublished manuscript, University of Pennsylvania, Philadelphia, PA.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35,* 15–175.

Zeigarnik, B. (1927). Über das Behalten von erledigten und unerledigten Handlungen. *Psychologische Forschung, 9,* 1–85.

Zettle, R. D., & Hayes, S. C. (1980). Conceptual and empirical status of rational-emotive therapy. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.) *Progress in behavior modification* (Vol. 9, pp. 126–166). San Diego, CA: Academic Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.04

Family Therapy and Systemic Approaches

ARLENE L. VETERE
University of Reading, UK

6.04.1 INTRODUCTION 85

6.04.2 FAMILY SYSTEMS THEORY 86

6.04.3 FAMILY LIFE CYCLE THEORY 87

6.04.4 SOME MAJOR SCHOOLS OF FAMILY THERAPY AND SYSTEMIC PRACTICE 88

6.04.4.1 Structural Family Therapy 88
6.04.4.2 Milan Family Therapy 89
6.04.4.3 Range of Therapeutic Techniques 89

6.04.5 DEVELOPMENTS IN FAMILY THERAPY THEORY AND PRACTICE 90

6.04.5.1 Feminist-led Critiques 90
6.04.5.2 Constructivism and Social Constructionism 91
6.04.5.3 User-friendly Approaches 92

6.04.6 RAPPROCHEMENT WITH OTHER THEORETICAL APPROACHES 93

6.04.6.1 Psychodynamic Influences 93
6.04.6.2 Cognitive Behavior Therapy Influences 94

6.04.7 ASSESSMENT IN FAMILY THERAPY AND SYSTEMIC PRACTICE 95

6.04.8 THE PROCESS OF CHANGE IN FAMILY THERAPY 97

6.04.9 CLINICAL PRACTICE PATTERNS OF FAMILY THERAPISTS 99

6.04.10 FAMILY THERAPY AND SYSTEMIC PRACTICE OUTCOME RESEARCH 100

6.04.11 TRAINING AND SUPERVISION 101

6.04.12 SYSTEMIC CONSULTATION 102

6.04.13 CONCLUSION AND THE WAY FORWARD 103

6.04.14 REFERENCES 103

6.04.1 INTRODUCTION

Gurman, Kniskern, and Pinsof (1986) defined family therapy as "any psychotherapeutic endeavor that explicitly focuses on altering the interactions between or among family members; and seeks to improve the functioning of the family as a unit, or its subsystems, and/or the functioning of individual members of the family" (p. 565) The term "systemic psychotherapy" is a broader definition that includes, in addition, intervention in other groups and organizations using systems ideas and keeping a relational focus, for example, networking (Dimmock & Dungworth, 1985) and systemic consultation (Boscolo & Bertrando, 1993).

Systemic psychotherapy recognizes the recent important developments in family therapy theory and practice, building on the earlier pioneering work of the family therapists.

6.04.2 FAMILY SYSTEMS THEORY

Family therapy developed during the 1950s and 1960s as practitioners experimented with involving family members in the treatment process. Early observations, such as "see-saw" effects in marriage, where improvements in the psychological well-being of one partner as a result of individual psychotherapy seemed to herald a worsening for the other partner's well-being; or "status quo" effects where improvements in individually treated psychological symptoms were not sustained; or the inability of individual approaches to address relationship difficulties *per se*, led early family therapists to speculate as to the importance of the social and relational context for the genesis and/or maintenance of individual symptomatology. Thus, the family system or kinship network was successfully placed at the center of therapeutic thinking. However, the early practitioners found their practice had developed ahead of their conceptual thinking. They turned to general system theory (von Bertalanffy, 1968) for its potential for application to the study of family process and family therapy.

Systems research is concerned with the concept of self-organized complexity in living systems, a central tenet of which was that a whole functions as a whole through the interdependence of its parts. General system theory attempted to explain how this obtained in the widest variety of systems, spanning the social, physical, and biological world. Thus, the theory attempted to classify systems according to the way the parts were interrelated, and to describe typical patterns of behavior for the different classes of systems as defined.

Within the model of organized complexity, there exists a hierarchy of levels of organization, for example, organelle, cell, tissue, organ, organism, such that each level of organization is more complex than the one below, with each level characterized by emergent properties that do not exist at lower levels and as such are not reducible to previous events. Our example of a biological system is said to have emergent properties, such as reproduction, self-organization, and self-reflective behavior. Systems are described as open if they exchange materials, energies, and information with their environment. Thus, open systems can be considered subsystems of higher order systems, for example, the individual as a member of the family, the family as part of the kinship network, the kinship network as part of a culture, and so on.

In this way, general system theory was adapted to the study of complex organization and interaction in family household groups and kinship networks. A major contribution has been in the study of system adaptability, the balancing of the homeostatic tendency, and the capacity for transformation. The family system is defined in terms of its structure; its structure is defined as the network of relationships amongst its component members; and its relationships are defined in terms of interactions that are mediated by communication, information exchange, and the development and transmission of meaning. Thus, the family system is said to function to develop networks of operations suitable for coping with the varied and changeable environmental inputs and internal stresses to which it is subject.

Family system theorists are concerned with the description of family rules, identifying hierarchies of feedback and control. Feedback or recursive processes are believed to be characteristic of social systems, such as family groups, where family members' output is recognized as input at some later stage. Thus, explanations of behavior embrace circular notions of causality. Boundaries are said to determine system and subsystem membership in family groups, with family rules operating to define who belongs to the subsystem and their roles and tasks within the subsystems. Boundaries are described as more or less open according to the degree of exchange with the system's environment.

General system theory provided a conceptual framework for early attempts to describe complex, time-related interactional behavior amongst family members, for which psychological and sociological theories were not suited. It focused attention on the role of each family member in the maintenance of psychological symptoms and took account of the various social and cultural variables impinging on family groups. The development of individual distress was seen within a contextual, social matrix and the treatment process addressed both individual needs and the needs of other family members. For example, Prince and Jacobson (1995) examined the hypothesis that the treatment for depression in married women needed to address the interpersonal environment. They examined three recent studies which compared behavioral marital therapy for distressed couples with individual cognitive therapy where the women were diagnosed with unipolar depression. They found that both therapeutic approaches were successful in reducing depression at six and twelve month follow-ups, but that the couples therapy was effective in reducing the marital distress.

The application of general system theory to the study of the family has been critiqued extensively (Pam, 1993; Vetere, 1987). Difficulties include a lack of clear conceptual definitions, necessary for operational definitions, and overlap between the *description* and the *explanation* of family system behavior. The conceptual and practical focus on relational processes led to a perceived diminution of the importance of individual emotional experiences and motivations, and the emphasis on circular causality led to a perceived diffusion of responsibility within family groups for unacceptable behaviors, such as violence and abuse. It is debatable to what extent family therapy practitioners took circular causality to mean that victims of violence played a part in the development of violent patterns of behavior for which they were accountable or blamed, or used the concept of circularity to understand how such processes arose. The current position separates the issue of responsibility for violence from the explanation of how violent behaviors occur. A linear moral stance is taken with respect to the responsibility of the perpetrator for the abuse, and then only working therapeutically, using explanations and notions of patterning and circularity, when the perpetrator has admitted responsibility and engaged in a nonviolence contract (Bentovim & Davenport, 1992).

6.04.3 FAMILY LIFE CYCLE THEORY

It has been proposed that family groups pass through phases, whose characteristics are determined by changes within family members, the impact of external events on family members, and the influence of sociocultural norms and requirements (Carter & McGoldrick, 1989). Thus, the family system is seen to move through time, expanding and contracting its membership as individuals join and leave the system. The family is usually a multigenerational system at any time, so that each generation can be influenced by previous and later generations. Special challenges to the family's organization, membership, and belief systems, which may be experienced as stressful, are said to arise at transitional points in family development, such as the birth of children, or young adults establishing their own households, where there is a confluence of external and internal demands for changes, understood and interpreted by family members within the context of transgenerational influences, that is, the handing down of family cultural beliefs, expectations, and practices. So, for example, grandparents might have a similar or different set of expectations to the parents about their involvement in rearing their grandchild, or a young couple might experience some difficulty in establishing what is their province of decision making relative to that of their extended family.

Carter and McGoldrick (1989) delineate a typical set of stages in development as follows: courting couples, couples without children, childbearing families with children in the preschool years, families with school-age children, families with adolescents at home, families with adolescents beginning to evolve separate lives, families with adult children, and families in retirement. Family members are said to face different challenges at different phases of the life cycle, with different expectations of self and others according to external demands and maturational and social demands for change and adaptation.

The life cycle stage model has been critiqued as applying to Western, middle-class, nuclear families with less relevance for different family forms, some ethnic groupings, poor families, and so on (Vetere & Gale, 1987). Certainly, if we accept the concept that a family is made up of different individuals at different stages of growth, the concept of family is difficult to grasp within this model. However, as a model it should purport to show how transitions affect adaptation, which demands a longitudinal perspective within research. The model, while not easy to put to the test, has construct validity and its implications for other theories of family functioning is considerable.

Definition of the family presents considerable difficulties for family researchers and therapists (Trost, 1990). Trost's survey of general public definitions of the family revealed little agreement, for example, with opinion literally divided over whether a heterosexual marital couple without children constituted a family or whether a lone parent and child constituted a family. The criteria for definition of the family used by respondents to the survey were legal ties, biological relatedness, common residence, and psychological significance, another concept which is difficult to define. In addition these criteria were weighted differently by the respondents. Thus, family researchers and therapists differentiate between definition of an individual's family, where an individual's perceptions of family life are accorded prominence, and definition of the family unit, which is believed necessary for the understanding of the development of family groups over time, family-wide difficulties, impact of traumatic effects on family processes, and so on. For many purposes therapists and researchers use the family household as the defining criterion.

A fully comprehensive theory of family functioning is not available (Burr, Hill, Nye, & Reiss, 1979a, 1979b). Family therapy approaches involve a range of schools and theories. The rationale for family therapy assumes that most individuals are born into and develop in the context of family groups. The family is seen as the cradle and web of emotional development and the early source of our attitudes and beliefs about ourselves, our relationships, our past, present, and future circumstances. The family is the origin of our basic patterns of social interaction and interpersonal adjustment. Thus, the system that is seen to create and support patterns of behavior may be the means of describing, assessing, and changing interpersonal behavior.

Family therapy is said to be indicated:

(i) when a child or adolescent is the referred person;

(ii) when family members define the problem as a family issue, such as relationship and communication difficulties;

(iii) when relationship difficulties threaten the future of the couple relationship or the adequate care of the children;

(iv) when the family has experienced recent stress, such as death, serious illness or injury, loss of employment, "leaving home" issues;

(v) when psychological symptoms have secondary gain effects; and

(vi) when family members become organized into "helping" with the problem in such a way that the attempted solutions become problematic themselves.

Family therapy is contra-indicated:

(i) by practical limitations, such as the unavailability of key members or the unavailability of an experienced therapist;

(ii) when family members are "sentenced" to therapy by the courts as an alternative to legal proceedings and motivation to participate in therapy is highly ambivalent;

(iii) when a family presents "too late";

(iv) when mediation might be more appropriate;

(v) in circumstances of precarious emotional equilibrium or emotional deprivation, when family therapy is more appropriately considered part of a larger intervention programme, that might include individual therapy, social supports, and practical aid.

6.04.4 SOME MAJOR SCHOOLS OF FAMILY THERAPY AND SYSTEMIC PRACTICE

The family systems tradition and the associated discipline of family psychology have given rise to a number of different schools of family therapy, such as structural (Minuchin & Fishman, 1981), strategic/problem solving (Haley, 1976; de Shazer et al., 1986; Weakland, Fisch, Watzlawick, & Bodin, 1974), Milan (Selvini-Palazolli, Boscolo, Cecchin, & Prata, 1980a), transgenerational (Lieberman, 1979), constructivist and social constructionist (Andersen, 1987; Hoffman, 1993), and narrative (Freedman & Combs, 1996).

6.04.4.1 Structural Family Therapy

Structural family therapy (Minuchin, 1974; Minuchin & Fishman, 1981) is a body of theory and techniques that approaches individuals in their social contexts and conceptualizes family interactions as habitual and sequential. Structure refers to the family's organizational characteristics, the subsystems it contains, and the rules which govern interactional patterns among family members. An aim of therapy is to alter organizational patterns, particularly where modes of communication among family members are seen to be dysfunctional. When the structure of the family group is transformed, the positions of members in that group are altered accordingly. As a result, each individual's experiences change. Thus, individual change is predicated upon system change.

Structural theory has five principal features:

(i) The family is a system which operates through transactional patterns, that is, repeated transactions establish patterns of how, when, and to whom to relate, and these patterns underpin the system.

(ii) The functions of the family system are carried out by bounded subsystems.

(iii) Such subsystems are made up of individuals on a temporary or permanent basis, and members can be part of one or more subsystems, within which their roles will differ.

(iv) Subsystems are hierarchically organized in a way which regulates power structure within and between subsystems.

(v) Cohesiveness and adaptability are key characteristics of the family.

Structural family therapy is described in three stages: joining, middle therapy, and termination. Therapy is believed to be effective when the therapist forms a new system with the family. The therapist relies heavily on techniques of joining and accommodation, such as planned support for the existing family structure, tracking the process and content of family members' communication, and accommodating to the range and style of family affect. Middle therapy interventions confront and challenge family members at the three levels of symptomatic behavior, family structure, and family members' beliefs, in the attempt to create therapeutic change. Interventions can include the enactment and re-enactment of interactional patterns, the development of negotiation and problem-solving skills, reinforcing parental authority, reframing interpersonal dilemmas, creating opportunities for empathic appreciation of others' perspectives and needs, providing support for change, and negotiating tasks within and between sessions. The therapy contract is terminated when the family members have rehearsed their ability to solve new problems and have had the opportunity to solve earlier ones.

6.04.4.2 Milan Family Therapy

Milan family therapy approaches the family as a history-containing system with entrenched meanings (Selvini-Palazolli et al., 1980a). It is based on Bateson's (1972) circular epistemology and focuses on information and difference. The Milan family therapist searches for differences—in behavior, in relationships and in the way family members perceive and construe events—and connections and links between ideas, behaviors, relationships and events. The approach assumes that these connections hold the system in balance.

Milan family therapy has been described as "long brief therapy" (Tomm, 1984) with sessions usually held monthly over a one-year period. The therapists work as a team, with two therapists usually interviewing the family and observed by two colleagues from behind the one-way screen. The session format often consists of an interview, followed by a break when the two interviewing therapists retire behind the one-way screen to consult with their colleagues, and finishing with the delivery of a message to the family. The Milan interviewing techniques consist of hypothesizing, circular questioning, neutrality, positive connotation, and the use of rituals. Hypothesizing refers to the presession formulation when the team hypothesizes about what might be maintaining the family problems. Hypothesizing helps organize incoming information during the family session, and is constructed continuously as a result of feedback received from the family to the questioning developed from the original hypotheses. Circularity throughout the system is stressed; problems and events are depicted as interpersonal, such that a problem can be seen as an event between two people.

The use of circular questioning in the family session reflects and is guided by systemic hypothesizing. Every question explores differences in family members' perceptions about events, relationships, beliefs, and the connections between them, illustrating Bateson's ideas about difference, information, and circularity. For example, different family members might be asked about their attitudes to a particular event, allowing the therapist to focus on difference without being too confrontational, exploring differences in meaning afforded to the same event. Triadic questioning, asking one person for their view of the relationship between two other family members, often produces changes in relationships as well as providing information for the therapists. Family members learn to think in circular fashion and to become observers of family process in a way that allows different views to be enlightening and helpful. Neutrality refers to the therapist's attempts to be even-handed in the session, allied to all without getting involved in family coalitions or alliances. It does not mean that the therapist is indifferent or ethically neutral, especially when working with abuse and violence in families. The therapist is more concerned with understanding the meanings in the system than with attempts to change the system. This is believed to be the process through which change occurs, that is, by not intervening to direct family members, the therapist encourages the family members' ability to generate their own solutions.

Positive connotation is a form of reframing, whereby symptomatic behavior is seen as positive or helpful because it keeps the system in balance and furthers family cohesiveness. Thus, the intention behind symptomatic behavior is seen in a good light, preparing the family for a paradoxical injunction around the need for family cohesion apparently to require the presence of symptomatic behavior. Family rituals often mark and facilitate developmental transitions in family life. They may be developed as part of therapy to help clarify family dilemmas and promote new ways of doing things which may help family members change their views and attitudes and see different options for behavior.

6.04.4.3 Range of Therapeutic Techniques

All family therapists use a range of therapeutic techniques, which vary somewhat according to school, such as problem-solving techniques, solution-focused techniques, directive techniques (e.g., enactment, restructuring boundaries), neutral techniques (e.g., emphasizing patterns in relationships), cognitive intervention (e.g., reframing of relationships), between-session activity (e.g., tasks, rituals), narrative and restorying techniques.

The hallmark of the family systems approaches has been the development of methods of live supervision and consultation (Campbell, Draper, & Huffington, 1989), such as the use of video and live observation teams located behind one-way screens (Madanes, 1984; Whiffen & Byng-Hall, 1982), in-room supervision (Kingston & Smith, 1983), reflecting teams and reflective practice (Andersen, 1987), and so on.

6.04.5 DEVELOPMENTS IN FAMILY THERAPY THEORY AND PRACTICE

Contemporary systemic thinking and practice represents a knitting together of different theoretical positions and concepts, such as: positive and negative feedback processes (recursiveness); hierarchy of system levels; boundaries (open and closed); pattern in relationships and meaning; family life cycle; communication theory; stability and change processes; symptomatic and system change; the position of the observer in the system; feminist critiques of gender, power, and inequality; narrative approaches; and solution focused approaches. Increased emphasis on the conceptual integration of different family therapy approaches (Burnham, 1992; Liddle, 1991a), the fit between therapist style and preferred models, and the needs of the family members and collaboration in the therapy endeavor (Hoffman, 1993) can be traced to recent theoretical developments in the field. For example, Goldner, Penn, Sheinberg, and Walker (1990) describe the integration of social learning, psychodynamic, sociopolitical, and systemic ideas in their therapeutic work with male violence in couple relationships.

Current systemic thinking and practices have been strongly influenced by the feminist-led critiques of power and inequality in family life (Perelberg & Miller, 1990), the postmodern developments of constructivism and social constructionism, and the emergence of user-friendly approaches to family therapy (Reimers & Treacher, 1995).

6.04.5.1 Feminist-led Critiques

The feminist-led critiques of family systems theory and practice heralded extensive rethinking and revision of the assumptions and models of working with families. Generation had long been recognized in the field as an important organizing feature of family life. Feminist family therapists pointed out the dimension of gender as another organizing feature of family life and proceeded to analyze and deconstruct the concept of power and gender-based access to sources of power (Hare-Mustin, 1986; Goldner, 1985, 1988). The sociological structural-functional theories of Parsons, which posited that men held instrumental roles in family life and women held expressive roles, no longer went unchallenged. Thus, family therapy was seen as a political process, in that it dealt with the allocation and distribution of power among family members, the therapist and family members, and so on.

Some significant contributions from feminist theorizing which influenced the development of family therapy thinking and practice included:

(i) recognition that men and women have different experiences of self, of others, and of life;

(ii) recognition that men's experiences had been more widely articulated than women's, which was underrepresented;

(iii) recognition that women do not have equality of opportunity, despite recent legislation in both the UK and USA (Central Statistical Office, 1990);

(iv) placing the family in a historical context and challenging accepted views of the family; and

(v) calling for a re-examination of family life, such as redistributing household and "mothering" responsibilities, validating nontraditional sexual and living arrangements, campaigning for reproductive rights, and calling for an end to women's economic dependence on men.

Such theorizing highlighted how slow family therapists had been to identify the links between social inequalities and psychological distress.

Williams and Watson (1987) have argued that the issue of asserting authority and power and having it acknowledged and respected is particularly problematic for women in a variety of situations. In addition to the relative lack of experience with leadership roles and the overt use of power in public life and social institutions, most women are reared in a multitude of nonverbal behaviors that communicate submission and indecision. They suggest that men and women traditionally have had differential access to types of interpersonal power, and these means of access are determined largely by sex-role stereotypes and expectations. Women are said to influence others more indirectly and to rely on their own personal resources including attractiveness, kindness, empathy, warmth, and close interpersonal relationships for the exercise of referent power. Men are said to utilize more direct means of influence, relying on a different set of resources, such as strength, skill, and competence and tend to be more indifferent to the interpersonal/intimate dimension for the exercise of instrumental/expert power.

The feminist critiques of differential gender-based access to power, and the new focus on gender as a once hidden dimension of family life, had enormous implications for the practice of family therapy. A gender-sensitive approach to family therapy developed, not as a set of specific skills or techniques, but as a process between therapist and family members which provided an opportunity for all family members to negotiate both their individual and system needs. Thus family therapists strived to be more aware of their own values regarding gender, as they are expressed in training, therapy and supervision; and began looking at the extent to which our ideas about differences between men and women are based on sexist stereotypes (Wheeler, Avis, Miller, & Chaney, 1989).

Family therapists are curious about the ways in which gender roles and stereotyping affect each individual in the family, the relationships between family members, the relationships between family members and other social institutions, and relationships between family members and the therapist (Gregory & Leslie, 1996). They ask questions that make explicit the issues, expectations, decisions, and behaviors that demonstrate the degree to which equality of opportunity and reciprocity exist between men and women in the family. Using analyses of interpersonal power, family therapists formulate questions about how the presence or withdrawal of different sources of power affect everyday processes, such as decision making, negotiation, and conflict resolution. Family therapists can use positive reframing and relabeling to shift the conceptual and emotional perspective on an individual or a relationship. For example, what may have been seen as personal inadequacy may be reinterpreted as socially prescribed, by exploring with men and women what they have been taught about their own gender roles and each other's, such as passivity as a model for female behavior or emotional impassivity as a model for male behavior. Gender-sensitive family therapists can facilitate consideration of a wider range of perspectives, behavior, and solutions that are less constrained by more traditional definitions of roles and personal identity. For example, when discussing parental teamwork and shared responsibilities, attention can be paid to the implications for both partners by checking that the woman is willing to share parental responsibility and has other ways of expressing her competence, and that the man is willing to bear the cost in the workplace of being more involved in the family. Family therapists can use their "gendered selves" in therapy in a therapeutic manner, for example, by modeling alternatives to traditional roles.

Recently, there has been a shift from a woman-centred focus to gender-sensitive family therapy to an incorporation of analyses of male psychological experience (Frosch, 1992; Mason & Mason, 1990). The debate on how to engage men/fathers in the therapy process has been informed by research which seems to indicate that men are less likely to seek help with emotional concerns than women (Verbrugge, 1985) and that when they do seek help, their problems are at a more serious stage of development (O'Brien, 1990). The risk of men not being engaged is heightened: (i) if family therapy is seen to take place in "women's time" and in a "woman's world," so to speak; (ii) if men are slower on the whole than women to perceive indicators of relationship difficulties, and (iii) if expressive differences between men and women leave them believing that fathers get involved in therapy only when family distress is very high. The greater difficulty in engaging men in the therapy process leaves women shouldering the burden of the responsibility for change, with fathers, albeit inadvertently, made peripheral.

Issues of culture and race have also been critically examined in the practice of systemic family therapists (Lau, 1987). Similar distinctions have been drawn between cultural awareness and culturally sensitive practices and the role of training in promoting sensitive practice (Hardy & Laszloffy, 1995). Reflexivity in the training process has led to examination of how therapists' cultural identities (which includes consideration of ethnicity, gender, social class, age, and so on) influence their understanding and acceptance of those who are both culturally similar and dissimilar. Falicov (1988) suggests that one way to help family therapists in training to think culturally is for them to interview a nonclinic family of a distinct ethnic or socioeconomic group. To date, the literature consists of theoretical critiques with clinical case descriptions of culturally sensitive practice.

6.04.5.2 Constructivism and Social Constructionism

Constructivism and social constructionism have had a significant impact on contemporary systemic thinking and practice (Hayward, 1996). A postmodern perspective embraces issues of meaning and language, narrative, politics, and practices of power. Family therapy theory increasingly attends to the interpretive meaning-making dimension of experience and the multiple contexts in which it occurs and evolves (Cronen & Pearce, 1985). The usefulness of systemic metaphors of family life has been well-documented; their further development has been prompted by the interest in family members' beliefs and stories embedded in language, with family members' ability to reconstrue their worlds in accordance with values and aspirations seen as the central impetus to change (White & Epston, 1990). This contrasts with earlier notions of change in the field, where change at the level of a relationship was seen to lead to change in an individual's felt experience. Anderson and Goolishian (1988) argue that we are dialogical beings who evolve knowledge and meaning and attribute meaning to action through conversation and other forms of social interaction. These ideas suggest that we construct stories or accounts about ourselves, others, and our relationships through social interaction. The major implication of these ideas is to suggest that the therapist's task is to help family members construct more useful stories about themselves and their relationships, for example, by restorying the past and altering the definition of the problem to change its meaning and change its perceived effect on the present.

The critical evaluation of systemic ideas and practice, alongside the recognition of "observer-created reality" has led to the development of reflective practice and reflecting teams (Friedman, 1995). Deriving from the work of Andersen (1987) and colleagues, reflecting team practices illustrate the notion of observing systems and the position of the observer in the system. Reflecting team practice assumes that the therapist and family form a new system during the process of therapy and that the supervising reflecting team can observe and reflect on the therapist/family system, family dilemmas and problems, creating an enriched image of the family and fresh perspectives on their problems and potential solutions. Since family members are invited to listen to the conversation of the reflecting team one or two times during each therapy session, family members themselves assume the position of the observer whilst listening, and when subsequently invited to comment on the reflecting team's conversation, we see an iterative process of both therapist's and family members' views and perspectives folding back on each other, much as in the way one might knead bread dough, as pointed out by Lax (1989).

Andersen first outlined the reflecting team in his 1987 article. He suggested that the working assumptions of the reflecting team included the following:

(i) The observer generates many of the views and distinctions we call "reality," with many possible interpretations present in those distinctions.

(ii) When people share their views, each person hears many different versions of this "reality," and these shared views constitute an "ecology of ideas."

(iii) At any point in time, people can only respond or take part in interactions that are within or partially within their known repertory and experience.

(iv) Sharing different versions of the same world or reality creates the conditions for "stuck" family systems to change, as family members realize that they have more than one perspective, option, or solution available to them.

A fundamental tenet of this approach is that information needs to be shared rather than withheld. This view is held in common with "user-friendly" approaches to family therapy, discussed below.

There are several different guidelines available to the practicalities of reflecting team practice, although they all agree that discussion should positively connote family members' motivations where ethically feasible, and contributors to the reflecting team discussion should build their comments on previous comments so a coherent account emerges during the reflection. The theory and research underpinning this approach is still in the early stages of development. Jenkins (1996) has both critiqued the approach and begun research into the theoretical assumptions, techniques, model of change, indications and contra-indications, using the Delphi technique.

6.04.5.3 User-friendly Approaches

User-friendly approaches to family therapy (Reimers & Treacher, 1995) developed in the UK in direct response to both a perceived diminution in significance of individual subjectivity within some areas of family systems theorizing and a lack of research attention to the experiences of family members in therapy. A recent book by Howe (1989) put forward some devastating criticisms of family therapy. Howe interviewed a small sample of families offered therapy by a team of social workers in social services in the UK. The therapy was a mix of brief strategic and Milan approaches undertaken by a single therapist, connected to a supervising team by closed circuit television and an ear bug. The majority of family members interviewed found the hi-tech approach very alienating and objected to "their" therapist apparently being controlled by some all-seeing yet unknown team.

Howe's book was helpful in that it mobilized a response to alienating practices, yet was flawed in that he failed to acknowledge the many developments in family systems theory and practice, not least of which is a longstanding debate on the ethics of family therapy thinking (Walrond-Skinner & Watson, 1987) and the necessity for therapy to be both ethically and politically defensible (Waldegrave, 1990).

User-friendly approaches recognize that many family members find it very distressing and problematic to come to therapy. Thus, it places a premium on providing adequate information about therapy and pays attention to convening and engaging family members (Treacher & Carpenter, 1983). The relationship of the therapeutic alliance to treatment effectiveness has not been researched in the field of family systems therapy anywhere near as extensively as in the individual psychotherapies, where the therapeutic alliance is held to be a common factor underlying much of effective helping. Research conducted by Bennun (1989) found that fathers were more likely to engage in therapy if they perceived the therapist to be competent, to show a positive liking for the family and to have a problem-solving orientation, whereas mothers preferred therapists to provide an opportunity for the airing of common concerns amongst the family members. Interestingly there seems to be a crossover effect in middle therapy, where mothers preferred a problem-solving approach and fathers had come to appreciate "just talking."

The effect of the therapist's race and gender on the engagement process has received little empirical attention in the family therapy literature, despite its prominence in family therapy discourse over the 1990s. Preliminary research by Gregory and Leslie (1996) with 63 adult heterosexual couples suggests that black females rate their initial sessions more negatively than white females when seeing a white therapist, and that black men had a more positive response than white men to the initial sessions, regardless of the race of the therapist. These differences decreased over time. They found no significant effects for gender of the therapist, which is consistent with the very limited research to date. Gregory and Leslie speculate as to the role of perceived racial difference in engaging family members in therapy and call for the identification of procedures which enhance the probability of family members staying in therapy.

6.04.6 RAPPROCHEMENT WITH OTHER THEORETICAL APPROACHES

Family systems models and practice differ from individual psychotherapeutic approaches because they focus primarily on relationships. They emphasize relationships between people, events, beliefs, behaviors, and so on, as opposed to emphasizing the internal world of individuals, as in individual psychotherapeutic models. Thus, the emphasis in family systems models of the creation of meaning through interaction is different from a primary focus on the cognition of the individual. The understanding of psychological symptoms is rooted in explanations of interactional process. For example, individual development is conceptualized as a complex series of relationships which extend beyond the individual, so a child's development is influenced by the child's relationship with the mother, the mother's relationship with her partner, extended family relationships, prevailing economic and political conditions, and so on, in a rich contextual interplay of different system levels. It is assumed that individual development and maturation are important and the systemic approach focuses on how development is understood and how these meanings affect relationships. For example, a family systems approach recognizes the importance of the different meanings attached to a child's genetic condition and how those meanings might influence relationships.

6.04.6.1 Psychodynamic Influences

Psychodynamic ideas have long informed family systems theory and practice as exemplified in the work of Ackerman (1958), Framo (1982), and Skynner (1976). The psychodynamic view of personal motivation and sources of anxiety and psychological discomfort has been generalized to the level of family group functioning. The psychodynamic systems approach seeks to understand how intrapersonal conflicts and motivations interlock and are expressed at the interpersonal level, with a specific focus on how such interlocking affects the development and expression of psychological symptoms in family members. Proponents of this approach would be interested in how couples bring to their relationship separate psychological heritages rooted in their past parent–child relationships and the extent to which past patterns of relating and introjection influence the current relationship. Applications in practice would seek to understand the extent to which such past attachments are problematic for family members in their current relationships and, by using processes of insight, attempt to bring about change in family members' relationships. Goldenberg and Goldenberg (1991) summarize the main differences between psychodynamic approaches to family therapy and systemic approaches in the differing emphasis on the role of past experiences and unconscious processes; the emphasis on insight- vs. action-oriented techniques; the role of the therapist in making interpretations of individual and family behavior patterns; and the major focus on the individual, in which the way family members feel about each other is emphasized.

Wachtel and Wachtel (1986), writing from the perspective of individual psychodynamic psychotherapists, suggest a number of ways in which systemic ideas can inform the thinking and practice of individual psychotherapists. They recommend that individual therapists explore the systemic meaning of psychological symptoms by asking such questions as:

(i) What is the functional significance of the individual's symptom(s) for the family as a whole?

(ii) How would the family be stressed if the individual were to change?

(iii) Does the symptom serve to restabilize a family whose stability has been threatened?

(iv) Is the symptom a result of an attempted solution to the problem that has in some sense "backfired" and itself become the problem?

(v) Is the symptomatic individual acting out someone else's distress?

Such questions serve to contextualize our understanding of the meaning of symptoms within the individual's significant emotional relationships and puts the individual's problems in a larger framework.

Wachtel and Wachtel also recommend the use of genograms (see Section 6.04.7) in individual psychotherapy. Genograms are useful tools for exploring the transmission of multigenerational patterns and influences, and provide a way of gathering information about an individual's implicit and explicit assumptions, wishes, fears, and values. Wachtel and Wachtel use the genogram like a projective test, a map of the unconscious. In addition they advocate the use of active interventions in individual therapy as developed in family therapy, for example, setting tasks (both individual and systems-oriented); devising rituals of celebration, mourning, and healing; using role play and role reversal interventions systemically; using systemic reframing and paradoxical tasks that embody collaboration; and predicting relapses. Finally, they recommend meeting "the cast of characters," such as adult siblings and partners, rather than engaging them in the therapeutic process itself, with the following possible benefits: correcting the tendency to blame, sampling the individual's interactional style, helping the system become more receptive to individual change, and helping to reconstruct the past.

6.04.6.2 Cognitive Behavior Therapy Influences

Behavior therapists and cognitive behavior therapists are interested increasingly in the utility of family systems ideas and practices. Emmelkamp and Foa (1983) have written about three sources of treatment failure, one of which is the neglect of, and/or the incomplete assessment of, the social contingencies and factors affecting clients' problems, in particular the extent to which a client's behavior is entrenched in habitual and longstanding patterns of family interactions and expectations. The recognition that some treatment failures might be the result of incomplete assessment of powerful social contingencies has led some cognitive therapists to address the question of what additional help clients might need in order to benefit from their therapies. Bandura and Goldman (1995) describe how they developed a family systems and cognitive behavioral analysis for use during assessment to address the above question. They recognize that family systems models and cognitive behavioral models have different theoretical underpinnings and do not attempt any theoretical integration. Instead they point to the areas of overlap and how the overlap can be useful in aspects such as the joint emphasis on beliefs and rule systems governing behavior, the importance of attributions and expectations in perceptions of self and others and the recognition of options for change, and the emphasis on problem-solving patterns and the significance of interpersonal contingencies to the understanding of symptomatic behavior. Thus in therapy, the focus on training in cognitive and behavioral skills to improve and enhance adaptive coping can involve the modification of environmental contingencies that influence clients' problems. This is an interesting and useful area of overlap, they would argue, because the family systems approaches also focus on interrupting interactional sequences that are thought to influence symptomatic behavior. Some of these interactional sequences are thought not to be reducible to the level of individual behavior as they are unique to systems functioning, examples are communication processes, power and hierarchy, patterns of disengagement and overinvolvement. Family systems ideas can thus help cognitive behavior therapists expand their functional analyses to include extended interactional analyses underpinned by models of circular causality rather than the linear models of reinforcement contingencies. Systemic analyses and practices can be used to help understand and overcome difficulties experienced by clients during the therapeutic change process, such as the broader costs associated with change, the client's involvement in familywide dilemmas, and the influence of family process on the content and function of schemata.

6.04.7 ASSESSMENT IN FAMILY THERAPY AND SYSTEMIC PRACTICE

The current and most widely used diagnostic system, the *Diagnostic and statistical manual of mental disorders* (4th ed., *DSM-IV*) (American Psychiatric Association, 1994), assesses and diagnoses individual psychopathology, and largely ignores the significance of interpersonal context to the development and maintenance of psychological problems. The marital and family therapies are not covered by third party insurance payments in the UK and the USA, despite the overwhelming evidence of their efficacy. Assessing systems of individuals demands more complex methodology than individual assessment and differs in the following ways. Assessment in family therapy is an ongoing process, occurring simultaneously with treatment. It is based on an understanding of multiple levels of systems hierarchy and feedback processes within and between levels. Assessment guides intervention. It can be conducted either as a clinical process through interviewing and observation or as formal psychometric procedure using both "insider" and "outsider" reports.

Minuchin (1984) writes that for therapy to be effective and for assessment to take place, the therapist needs to form a new system with the family: the therapist plus family system. In this, the therapist relies on techniques of accommodation and joining. Accommodation is the adjustment of the therapist to the family system. Aspects of accommodation are: *maintenance*, or planned support of the existing family structure; *tracking*, following the content and process of the family communication; and *accommodating* to the family's style and range of affect. Joining is the therapist's own method and style of helping to form the new therapist plus family system, so that the therapist does not lose the position of facilitator.

The areas for assessment in systemic work with families cover the following:

(i) the family's structure of subsystems, roles and boundaries, preferred transactional patterns, and available alternatives;

(ii) family members' strengths, flexibility, and capacity for change as revealed by responses to changing circumstances, past and present;

(iii) family members' perceptions of and sensitivity to each other's needs, behaviors, attitudes, and so on;

(iv) the context of family life, with specific reference to sources of support and sources of stress;

(v) both family and family members' life cycle stages and perceived expectations and tasks; and

(vi) the meaning and significance of symptomatic behavior for family members.

The emphasis on areas of assessment will vary according to the school of therapy. Common pitfalls of assessment include ignoring the developmental process, ignoring some family subsystem, and joining and supporting some family members at the expense of others. Gurman and Kniskern (1981) suggest that all family assessments should make clear at what organizational and psychological level the assessment is conducted, how soon therapy will follow assessment, and the relationship between the different methods for collecting data and the means by which the data should be collated.

Selvini-Palazolli, Boscolo, Cecchin, and Prata (1980b) paved the way for including the referrer in the assessment process, partly to prevent the waste of family and staff time with inappropriate or ill-prepared referrals. Referrers are often invited to preliminary consultation meetings with family members and the therapy team or wider network meetings to establish who in the system wants the referral and any differing views and expectations of the referral. The importance of thinking of referrers systemically as part of the process of convening, engaging, and assessing families has tended to lead to a decline in nonattendance rates (Reimers & Treacher, 1995).

Lieberman (1995) and his colleagues at the Prudence Skynner Family Therapy Clinic have developed an assessment and observation form that is used by both observing team members and the therapist to collate their views after the first family interview. They adapted the Current State Family Assessment (Loader, Burck, Kinston, & Bentovim, 1980) into five separate areas of observation:

(i) process:

(a) communication patterns, such as interruptions, listening, speaking for self and others, contradictory verbal and nonverbal messages;

(b) family atmosphere, such as the predominant mood, whether it is shared by all, what makes it change;

(c) family alliances, such as rigid vs. flexible alliances, presence of conflict and problem solving, scapegoating;

(d) feelings, such as the range, intensity and expression of feelings, presence of empathy;

(e) family boundaries, such as rigid or diffuse generational boundaries, cross-generational alliances, parental role responsibilities, differentiation, and connectedness;

(ii) family seating arrangements;
(iii) content of the interview;
(iv) formulation and themes; and
(v) feedback and tasks given to the family.

A popular atheoretical assessment tool used by family therapists is the genogram, derived from the work of the transgenerational family therapists (McGoldrick & Gerson, 1985). As many families are made up of many people experiencing a multitude of events across many generations, recording this information in a concise and coherent way can be difficult. The genogram provides a vehicle for recording social, emotional, and demographic data across the generations, such as births, deaths, marriages, divorces, life events, emotional connectedness, and so on. It provides a means of exploring the meaning and impact of events across the generations, engaging all family members in the telling. Thus, the process of constructing a genogram allows the therapist and team to contextualize such assessment questions as "what is the current family problem?," "what factors maintain the problem?," "why do family members come for therapy now?," "what are their expectations of therapy and the therapist?," "how have they tried to solve the problem previously?" within an understanding of background data on the characteristics of both the household family group and extended family network.

It has been noted by many beginning family therapists that constructing a genogram is a way of putting family members at their ease in the early stages of therapy, not to mention helping to reduce the anxiety of the therapist. It is a visual and active method that concentrates the attention of all participants, and provides a rationale for negotiating change. The construction of a genogram also has strong therapeutic influence as it can be used to reframe behaviors, events, relationships, and time connections; to make links across the generations and across emotional cut-offs; and to normalize some perceptions. It may facilitate alternate interpretations of a family's experience and point the way to new possibilities in the future.

Other more formal, psychometric assessment methods are available, such as the Beavers Systems Model scales (Beavers & Hampson, 1990). These include the Beavers Interactional Competence and Style scales and the Self-Report Family Inventory. The Competence and Style scales are completed by observers, using videotaped family behavior as the basis of their ratings and the Self-Report Inventory is completed by all family members above the age of 12. The Competence scale is a 13-item structured observational rating scale, with ratings derived from trained observers' evaluations of family interaction over a 10 minute period. The scale evolved from general system theory and studies of well-functioning vs. dysfunctional families. Competence is defined broadly as how well the family performs its necessary tasks, such as providing support and nurturance, establishing effective generational boundaries and leadership, promoting developmental separation and autonomy of offspring, negotiating conflict, and communicating effectively.

The Style scale is a nine-item observational rating scale. The Style dimension represents elements of enmeshment and disengagement at its extreme points, such as extremes in binding and expelling patterns; and affective patterns, such as subdued conflict where anger is threatening vs. open conflicts and hostility. Families at similar "competence" levels may show different functional "styles" of relating. The model assumes that competence in small tasks is related to competence in larger domains of living. The Self-Report Family Inventory is a 36-item self-report scale and measures the major elements of family competence and style, the cornerstones of the Beavers Model. The scales have been developed during 30 years of research into normal family functioning. The internal consistency of the scales, their construct validity and inter-rater reliability have been extensively researched and demonstrate acceptable levels of reliability. Such rigorous attention to issues of reliability and validity requires considerable investment in the training of raters. Most family therapists and supervising teams are unable to invest such resources in training, so the more formal psychometric scales are rarely used in day-to-day practice, mainly being used in research trials. Family therapists rely on clinical interviewing and observation, teamwork, knowledge and experience of family patterns for the assessment and formulation of family difficulties, sometimes supplementing their assessment with more formal procedures.

Termination of therapy will be discussed here as an assessment issue because the iterative and interactive nature of assessment and formulation in part determines the decision to end therapy. Theorizing and empirical research on ending therapy is under-represented in the family therapy field compared to the other psychotherapies (Treacher, 1989). Since family therapy was developed by and is often practiced by clinicians trained in other modes of therapy, notably individual therapies, there is a tendency in the family therapy literature to assume knowledge of both convening and engaging issues and termination issues in family therapy. Readiness for termination is assessed when family members resolve or learn to cope with the presenting difficulties, when they demonstrate increased independence and/or cooperation, when they display more "open" styles of communication, when they report an increased sense of security and show greater flexibility in performing family roles. According to the orientation of the family therapist and team, they will typically review whether family members' expectations of the therapy process have been met, summarize the treatment process and enquire about helpful and unhelpful aspects of the therapy, predict and rehearse future coping in similar situations, and offer a follow-up meeting far enough ahead for the family members to have tested out their newfound confidence.

6.04.8 THE PROCESS OF CHANGE IN FAMILY THERAPY

The main change mechanisms in family therapy are considered to be at the symptomatic level, the level of family structure, and the level of beliefs held individually and collectively. The different schools vary in their emphases. Structural therapy posits change in individual experience as a result of change at the level of the family's organizational structure. For example, a child's experience of parenting is said to change if the two parents learn how to work as a team rather than undermining each other's decisions. Strategic and brief therapies focus on symptomatic change, and use an understanding of interpersonal dilemmas and difficulties as ineffective solutions to problems that are maintained because people are unable to generate alternative solutions. Milan therapy uses intervention questioning to identify and highlight connections between beliefs, behaviors, and relationships, with change occurring as a result of the development of new perspectives within relationships. These changes are achieved using a range of techniques, described earlier in the chapter. The following discussion will consider how the family therapy process studies have investigated the processes hypothesized to bring about change and actual change.

Family therapy process studies investigate change mechanisms common to the various types of family therapy. To date, most research has focused on family therapy outcomes, demonstrating the efficacy of family therapy for many different psychological disorders as classified. The question of how interpersonal change is facilitated during family therapy is of most interest to the practicing family therapist, who is concerned with effective interventions, yet this is the area where there is a relative paucity of research. Thus family therapists tend to seek information on interpersonal change processes from theoretical articles, technical articles, and clinical workshops. This material tends on the whole to focus on what therapists should do, rather than considering what family members themselves need to do. The exception is the work of Kuehl and colleagues, who used ethnographic interviewing procedures in an iterative manner to elucidate accounts of interpersonal change from family members who had undertaken strategic/structural therapy for help with adolescent drug misuse (Kuehl, Newfield, & Joanning, 1990).

It is important to note when reviewing family process studies, that family therapists are challenged by complex demands and intrapersonal and interpersonal tasks during the therapy session that do not occur during individual therapy. For example, the therapist may create interpersonal alliances with family members who may well be in conflict with one another, manages the multiperson conversation in a collaborative and facilitative manner to prevent it becoming destructive, uses family members as "cotherapists," assesses and intervenes in live enactments of problematic family interactions and problem-solving attempts, and so on.

Friedlander, Wildman, Heatherington, and Skowron (1994) reviewed family process research, including naturalistic studies of conjoint therapy in which the focus of the study was verbal behavior of the therapy participants during therapy or their self-reported perceptions of actual interactions during therapy. Thirty-six studies, dating from 1963, met their inclusion criteria. Following Greenberg (1986), they organized their review hierarchically and focused on three levels of in-session behavioral processes: (i) speech acts during therapy, (ii) important incidents or change episodes during therapy, and (iii) the therapeutic relationship. Their review builds on the previous review of family process studies conducted by Gurman, Kniskern, and Pinsof (1986).

The bulk of family process research has been conducted at the speech act level of analysis, where all behaviors occurring during a specified segment of interaction are observed. Frequencies or proportions of these observed behaviors, such as rates of participation, types of response, and modes of expression are compared or otherwise used to predict successful therapy outcome. The following questions have been addressed using speech act analyses: What factors predict premature termination of family therapy? What are the common and distinctive features across different modalities of therapy? What is the relationship between participants' gender and therapy process? What interpersonal changes take place during the course of therapy, and so on. For example, Alexander, Barton, Schiavo, and Parsons (1976) found that the proportion of defensive to supportive speech acts was significantly higher in families who terminated their therapy prematurely. Shields, Sprenkle, and Constantine (1991) compared families who ended therapy prematurely with those who completed therapy in agreement with the therapist. They found that families who ended therapy early were characterized by more in-session disagreements and more attempts to structure the therapist, compared to more family problem-solving conversations and more therapist structuring in response to family disagreements in families who completed therapy. The results suggest that defensiveness amongst family members may predict early termination, although we do not have other information about the families and their contexts, so that the correlations may be misleading. A study by Dowling (1979) examined the consistency of therapist verbal behavior according to their role, therapist versus cotherapist. She found that therapists behaved similarly with different cotherapists and with different families, supporting the notion that therapists had a consistent cotherapy style.

Postner, Guttman, Sigal, Epstein, and Rakoff (1971) investigated family members' verbal behaviors as predictors of outcome. Interactional segments were analyzed at four points in therapy for 11 families. Speech acts were coded into emergency, welfare, or neutral emotional states. Outcomes were coded as good or bad by three independent judges. Results from the good outcomes showed that family members tended to speak more to each other during the course of therapy, that welfare statements increased during the course of the therapy, and that significant changes in emotional expression occurred between the second and sixth sessions. Unfortunately the speech act studies reviewed by Friedlander and her colleagues provide little accumulated knowledge towards answering how change occurs over time during successful therapy. Different investigators have used different coding systems within different theoretical frameworks, thus the generalizability of the findings is limited.

There have been a few attempts to research therapeutic episodes in the hope that by identifying significant moments in therapy, interpersonal change processes will be more readily elucidated. An example is provided in the work of Patterson and Forgatch (1985), who identified and coded instances of maternal "noncompliance." They found that noncompliance responses were more likely to follow therapists' attempts to "teach or confront" mothers, as coded, whereas a decrease in noncompliance was coded following therapists' "support or facilitate" interventions, as coded. Using a larger sample from the same child behavior management project, Patterson and Chamberlain (1988) identified instances of "mother–father within-session conflict" and using path analyses suggested that extraneous forces, such as parental stressors, marital conflict, and depression, appear to increase the within-session conflict among family members. This study is important in elucidating the role of factors outside therapy sessions that have a bearing on behavior in therapy. In addition they found high correlations between in-session conflict and antisocial behavior scores for the children. Patterson and Chamberlain conclude that therapists should plan interventions that help to reduce external parental and marital stressors outside the sessions. The Patterson studies rely on small samples but do provide some interesting support for family systems ideas.

The third level of analysis in the family process studies is the therapeutic relationship itself. Family therapists have not researched the ingredients believed to be important by individual therapists in establishing a therapeutic relationship, such as warmth, empathy, acceptance, and unconditional regard. Instead they have attended to the more strategic and systemic aspects of the therapeutic relationship, such as engaging and joining the family, and the development of therapeutic coalitions. Gurman and Kniskern (1978) concluded that the family therapist's ability to establish a positive therapeutic relationship with family members was most predictive of successful outcome. Since that time very few studies have been conducted and only six met the inclusion criteria of the Friedlander et al. review discussed here. Shapiro (1974) concluded that greater therapist emotional responsiveness to family members, as measured by questionnaire, predicted client continuance in therapy beyond the initial assessment phase. Families who were not seen in such a positive light by their therapists tended not to continue in therapy beyond assessment. These ratings are global with little psychometric support. Later research by Pinsof and Catherall (1986), using a more robust psychometric measure of therapeutic alliances, explored the development of therapeutic alliances across family therapy sessions, and found that such alliances often develop variably rather than uniformly. Their research showed a tendency for positive correlations between alliance ratings and therapist-rated outcomes, with most family members rating their therapist positively, and interestingly, that alliances are best understood as both multidimensional and occurring on multiple levels. Examples are:

whole system alliances, subsystem alliances, and individual alliances.

The study by Kuehl et al. (1990), referred to earlier in this section, will be discussed here because it is an interesting exploration of family members' views of therapy and the change process. Twelve families (37 individuals) who had completed family therapy with successful outcomes for adolescent substance misuse were interviewed. Family members identified stages in the therapy process, collectively described as the introductory meeting, assessment, getting down to basics and generating suggestions, putting suggestions into practice, sharing successes with the counselor, and troubleshooting and follow-up. Satisfaction with the therapy seemed to depend on family members' perceptions of the therapist as caring and understanding and able to generate relevant suggestions, whereas dissatisfaction was reported if the therapist was thought to be on "too strict a program." The flexible use of theory was valued by family members. Where parents reported that they considered their marital problems to be contributing to their child's problem, there was a willingness to explore this connection. Where parents believed their own problems to be separate from those of the adolescent, attempts by the therapist to explore a possible connection was the point at which therapy was reported to "stall."

Friedlander et al. (1994) conclude that most family therapy process studies are descriptive and at the speech act level of therapeutic process. Compared to the wealth of family therapy outcome research, there are few process studies, but a few conclusions are possible. These include: individual symptoms as observed in therapy occur in the context of predictable interpersonal events; there are more commonalities across family therapy approaches than differences, with therapists behaving consistently across family work, central in their position with families and skillful in indirect communications; changes over time in the course of therapy with positive outcomes as affective, cognitive, and behavioral, observed both interpersonally and intrapersonally; affective changes among family members appear to be crucial to effective therapy; family members' responses to their therapists and their therapists' responsiveness to them appear to be important; and finally, family members' motivations to engage in therapeutic activities are predictive of good outcome. More research is needed to address how family members themselves construe therapy and its effective ingredients. Observer ratings represent one possible viewpoint; the covert feelings and thoughts of family members have still to be explored alongside observer ratings. Interpersonal process recall as developed by Elliott (1984) in the context of exploring individual therapy process has promise for exploring family members' perspectives during family therapy. The bulk of the process studies have researched the more established styles and schools of family therapy, with the newer constructivist approaches remaining to be explored. But perhaps the real challenge to family process research lies in developing methods that both avoid isolating behavior from its social context and isolating individual behaviors from the stream of behaviors.

6.04.9 CLINICAL PRACTICE PATTERNS OF FAMILY THERAPISTS

There is little published data on the clinical practice patterns of family therapists, whereas in both the UK and the USA there is a growing body of data on the practitioner demographics and clinical practice patterns of clinical psychologists, psychiatrists, and social workers (VandenBos & Stapp, 1983). Because of the increased interest in family therapy, its effectiveness as a mode of therapy (Pinsof & Wynne, 1995), and the expanding role of family therapists in larger health care teams, Simmons and Doherty (1995) surveyed the clinical practice patterns of marital and family therapists in Minnesota, USA. Marriage and family therapists are currently regulated in 31 states through State and Federal legislation. In the UK registration of family therapists is through the United Kingdom Council for Psychotherapy, and in Europe many countries require the registration of family therapists as specialist psychotherapists.

Simmons and Doherty surveyed American Association for Marriage and Family Therapy (AAMFT) members from Minnesota ($N = 76$) for educational qualifications, demographic characteristics, and practice-related issues. In addition they sought data on 199 treatment cases involving a total of 351 clients, which they claim as a first in the field of marital and family therapy.

Their major findings were that: (i) short-term therapy is practiced, with an average case involving 11 sessions over a four-month period; (ii) therapy with families (average of eight sessions) and couples (average of 10 sessions) is briefer than that with individuals (14 sessions average); and (iii) a wide range of serious problems were treated by the surveyed therapists, including marital problems, depression, anxiety, child problems, and parent–child problems. Although academic training in marital and family therapy seemed to result in a systemic orientation as exemplified by the tendency of respondents to identify family and larger system problems, there was consistency between the respondents and their psychology and social work trained counterparts in the types of client problems treated, the utilization of *DSM* diagnoses, and the lengths of treatment. Simmons and Doherty's overall conclusion was that the clinical practice patterns of marital and family therapists were similar to those of the other established mental health professions.

6.04.10 FAMILY THERAPY AND SYSTEMIC PRACTICE OUTCOME RESEARCH

In both the UK and the USA there is an increased emphasis on the evaluation and audit of clinical practices and outcomes, aimed at the limitation of escalating health care costs whilst improving the quality of care. Family therapists have been collecting outcome data for many years. According to Bergin and Garfield (1994), marital and family therapy approaches have been subjected to rigorous scrutiny, with only a few other forms of psychotherapy studied as often. Different outcome studies have reported the use of single case designs and controlled and uncontrolled group comparison designs. These outcome studies have been reviewed by Hazelrigg, Cooper, and Borduin (1987) and Pinsof and Wynne (1995), and using meta-analytic techniques by Markus, Lange, and Pettigrew (1990) and Shadish, Ragsdale, Glaser, and Montgomery (1995). The overwhelming findings from all these reviews is that family therapy works compared to untreated control groups, with some demonstrated superiority to standard and individual treatments for certain disorders and populations. Meta-analysis demonstrates moderate, statistically significant and often clinically significant effects. The research literature supporting this conclusion is at least as robust as it is for other modes of psychotherapy.

Shadish et al. (1993) conducted a meta-analysis of 163 randomized experimental comparisons of the effects of marital and family therapy (marital = 62; family = 101) with distressed clients, published up to 1988. Seventy-one studies compared marital and family therapy to an untreated control group, and 105 compared it to another kind of marital and family therapy or to another model of psychotherapy. The number is higher than 163 because some studies contained multiple comparisons. The 71 studies showed that therapy clients were better off than the untreated comparison group. The size of this effect was approximately half a standard deviation, which means that the odds of a treated client doing better at post-test than a randomly chosen control client are two out of three. The effect sizes for marital and family therapy were both significant and roughly similar. Comparisons are hard to make because marital and family therapies are offered for the treatment of different presenting problems. The review considered specific presenting problems for the family therapy studies and found that family therapy clients were significantly better off than control clients for general child conduct disorders, child aggression, global family problems, and communication/problem-solving difficulties. Their review included 23 studies that compared marital and family therapies to individual psychotherapies. The differences in outcomes were small and nonsignificant across a range of presenting problems.

Chamberlain and Rosicky (1995), in their review of seven family intervention studies for adolescent conduct disorder and delinquency, published since the Shadish et al. (1993) meta-analysis which included 18 such studies, found that family therapy approaches appeared to decrease adolescent conduct problems and delinquent behavior when compared to individual therapy, treatment as usual, and no therapy, with similar effect sizes of 0.53. Treatment failure in the studies reviewed by Chamberlain and Rosicky correlated highly with poverty and/or social isolation for the family. The Florida Network Study (Nugent, Carpenter, & Parks, 1993) with high risk families found that families who received family therapy were four times as likely to stay together as families who did not, and families who received more than five treatment sessions were twice as likely to stay together as families who did not. So, for these high risk families, family therapy may be a necessary treatment component, but is not sufficient in itself.

Despite the improvements to methodology in the outcome literature, we are still a long way from answering the specific family therapy outcome question posed by Gurman, Kniskern, and Pinsof (1986): "What are the specific effects of specific interventions by specified therapists at specific points in time with particular types of clients with particular presenting problems?" (p. 569).

Gale (1980) provided a series of useful questions to ask when evaluating the quality and character of extant family therapy outcome research, covering issues such as theoretical rational and therapeutic schools, methods of training for therapists, pretreatment characteristics of clients, techniques of treatment, and evaluation of treatment outcome. For example, we might ask:

(i) Who has conducted the research, clinicians or researchers?

(ii) Are the studies well designed and free from bias?

(iii) Are control groups used for different types of family therapy, for family therapy and other forms of psychotherapy, for family therapy and no treatment control?

(iv) Are the effects of the therapist shown independent of the type of treatment?

(v) Are the outcome criteria for success and failure clearly specified?

(vi) Are multiple outcome criteria used and do they intercorrelate?

(vii) Is there a follow-up period and are the effects of treatment persistent over time?

(viii) What characteristics of families/therapists are associated with success and failure?

(ix) Are there nonspecific treatment effects?

(x) What is the comparative cost of family therapy? Is there a cost to the community for failing to provide financial support for family therapy?

Family therapists have a variety of therapeutic approaches to choose from; they are concerned increasingly to provide the best "fit" for clients with their particular circumstances and presenting problems, often without clear evidence supporting one systemic approach over another (Orlinsky, Grawe, & Parks, 1994). When working therapeutically with severe problems, such as adolescent conduct disorder, there is increasing evidence of the value of treatment packages, of which family therapy is a part (Pinsof & Wynne, 1995).

6.04.11 TRAINING AND SUPERVISION

Family therapy training courses at both introductory and qualifying level have been established in the UK since the 1970s, provided by family therapy training institutes. All courses are accredited by the Association for Family Therapy. The criteria for qualifying level training courses include completed training at introductory level, 320 academic hours, and 320 practice hours of which 40 are live supervised, and a personal development component which focuses on family-of-origin experiences and their effects on current thinking and practice.

Many of the qualifying level courses are program-centered (Street, 1988) and concerned to identify theories and skills for systemic practice within their curricula. Thus, the focus of training is very much on the process of teaching. Reviewers of research on family therapy training in both the UK (Street, 1988) and the USA (Liddle, 1991b) have expressed concern at the program-centered nature of the training and questioned to what extent courses address the dilemmas of adult education and adult centered learning, such as exploring the relevance of client change processes in family therapy to family therapy trainees, and learning from the experiences of other psychotherapy training courses. In addition, they point out that courses might do more to close the gap between clinicians and researchers said to exist in the family therapy field, by promoting reflective practice models and qualitative research methodology.

Research into the effectiveness of family therapy training has been limited compared to the proliferation of regulated training courses. Avis and Sprenkle (1990) suggest there are many reasons for this, including the increased complexity of the training issues and their relationship to practitioner outcomes, problems of sampling and replicability, and a lack of reliable and valid measures of training effectiveness. Despite these limitations, Avis and Sprenkle conclude from their review of the training research that conceptual and executive skills may develop at different rates, that different training approaches can bring about similar conceptual and intervention skills developments, that introductory level assessment skills and concepts can be taught using lectures and textbook methods, and that "in-house" training courses for agency staff have been shown to be effective in promoting systemic practices and broadening the remit of mental health services.

Breunlin and his co-workers (1989) attempted to predict which factors might contribute to improved performance in family therapy trainees before and after training on measures of conceptual, executive, and observational skills. Interestingly, previous family therapy experience did not appear significantly to affect change scores, although prior life experience and maturity did seem to be important, but only in changes on measures of executive skills, rather than conceptual and observational skills. Previous experience as an individual psychotherapist predicted improvement on conceptual skills change scores, perhaps suggesting that a previous training prepares trainee family therapists for thinking about therapy processes. Pulleybank and Shapiro (1986) were also interested in the acquisition of cognitive and therapeutic behavior skills during family therapy training. They suggest that their results indicate a developmental progression in learning, with cognitive skills acquired earlier in training, followed by development in planning and then intervention skills. Research by Anderson (1992) raised the knotty problem of how to reconcile differing observer opinions when assessing family therapy trainee performance. Anderson compared the assessments of placement supervisors, academic supervisors, and so-called neutral observers and found more change reported by placement supervisors than the "neutral" observers. The role of evaluation in supervisors' judgments has yet to be systematically researched in the field.

Researchers of family therapy training effectiveness have focused primarily on issues of skills acquisition, theory–practice linking, and preparing trainees for systemic practice outside of the somewhat protected environment of the training course. These issues continue to be of interest and when we map them on to Gurman and Kniskern's (1992) predictions for the future of the field, such as increased diversity of intervention formats and methods, increased recognition of the importance of the relationship between the family therapist and the family members, and increased interest in postmodern developments, the research agenda for training effectiveness into the first decade of the twenty-first century becomes broader.

6.04.12 SYSTEMIC CONSULTATION

Brunning and Huffington (1990) define consultancy as a direct or indirect process enabling individuals, groups, or organizations to fulfill their role, function, or tasks better. It is a process by which the person or persons seeking the consultation ask for help in identifying or clarifying concerns and in considering the options available for problem resolution. Thus, the consultees have legal, ethical, and administrative responsibility for initiative and action, as opposed to supervision where there may well be a hierarchical relationship between the supervisor and the supervisee, or therapy where there is pressure to accept suggestions or directives in order for the therapy to be effective. In the case of family consultation, it may well be a prologue or alternative to family therapy (Street, Downey, & Brazier, 1991). Wynne, McDaniel, and Weber (1986) discuss consultation with families as a process whereby the family members' responsible decision making is assumed and their resources and competence are directly tapped. The decision as to whom to invite to the consultation meetings will depend largely on where the impasse in the system is located. In addition, the systemic view of consultation would see the role of the consultant as a necessary participant in the system requiring consultation.

The stages in the process of systems consultation have been described by Wynne, McDaniel, and Weber (1986) as exploring, contracting, connecting, assessing, implementing, evaluating, and leaving. The stages are not mutually exclusive and represent a process that guides the activity of the consultant. Exploring involves clarifying both the request for consultation and who is requesting the consultation, by considering how the request came about, who approves the request and the role of the consultee in the team or organization requesting the consultation. In particular any political ramifications of the request are explored. Contracting is the process whereby agreement is reached on the goals of the consultation, the services provided by the consultant, consideration of the risks and consequences of the consultation, procedures for sharing information, and other practical arrangements. Connecting is the process of engaging key members of the consultation system in the consultation process and deciding how they will be involved in goal setting. Assessing refers to the methods used for gathering information and the systemic concepts used to understand the organization's structure, function, and dynamics. The assessment process particularly focuses on organizational life cycle issues, recent events triggering the consultation request, previous attempted solutions to the problem, and the belief systems of the consultee. Implementing involves specifying the systemic interventions chosen to meet the consultation request, decisions about whether education is provided, the means of collaborating with the consultee over implementation and procedures to ensure the maintenance of change. Evaluating is the process of deciding how the consultation goals and organizational changes have been met, who will take part in the evaluation process, plans for follow-up evaluation, and reflection on the process of consultation. Leaving describes the means by which a consultation is ended or the consultant renegotiates a different role within the organization.

Campbell (1985) describes an alternative model of systemic consultation specifically designed to help other family therapists who are "stuck" in their clinical work with a family. He adapts the Milan systemic model of family therapy to this task, using theoretical concepts of meaning, pattern, recursiveness, and difference. The consultation interview follows a similar procedure to a Milan-style family therapy interview with presession hypothesizing based on preliminary information, the use of circular questioning and reframing, characterized by the neutral stance of the therapist, and midsession breaks for team discussion and final formulation. The main aim of this style of consultation is to provide a space to comment on and clarify some of the confusions that arise when working in systems where different meanings for behavior arise from the differing perspectives from which that behavior is viewed.

In summary, systems consultation has the advantages of not prejudging the nature of the problem requiring consultation; the consultant can take a meta position from which to develop a perspective on systemic relationships and patterns; the consultation process facilitates the reframing of problems and the search for competence and strengths; a more collaborative relationship between consultant and consultee can be established because the consultee is "free" to take the advice offered or not; and the role of consultant does not preclude a shift to an alternative role, such as therapist.

6.04.13 CONCLUSION AND THE WAY FORWARD

Looking ahead to the next millennium is both exhilarating and daunting. The development of family therapy and systemic approaches to practice has been rapid, encapsulating and condensing the natural history of the development of systems of thought and professional disciplines. A time for consolidation of practice and ideas would be welcome. Rapprochement with other psychotherapies proceeds apace. The emphasis on context as an infinitely layered and dynamic process has enormous implications beyond the practice of family and systemic therapies. An ecosystemic vision and practice awaits us all.

6.04.14 REFERENCES

Ackerman, N. W. (1958). *The psychodynamics of family life*. New York: Basic Books.

Alexander, J. F., Barton, C., Schiavo, R. S., & Parsons, B. V. (1976). Systems-behavioral intervention with families of delinquents: Therapist characteristics, family behavior, and outcome. *Journal of Consulting and Clinical Psychology, 44,* 656–664.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Andersen, T. (1987). The reflecting team: Dialogue and meta-dialogue in clinical work. *Family Process, 26,* 415–428.

Anderson, H., & Goolishian, H. (1988). Human systems as linguistic systems: Preliminary and evolving ideas about the implications for clinical theory. *Family Process, 27,* 371–393.

Anderson, S. A. (1992). Evaluation of an academic family therapy training programme: Changes in trainees relationship and intervention skills. *Journal of Marital and Family Therapy, 18,* 365–376.

Avis, J. M., & Sprenkle, D. H. (1990). Outcome research on family therapy training: A substantive and methodological review. *Journal of Marital and Family Therapy, 16,* 241–264.

Bandura, M. M., & Goldman, C. (1995). Expanding the contextual analysis of clinical problems. *Cognitive and Behavioral Practice, 2,* 119–141.

Bateson, G. (1972). *Steps to an ecology of mind*. New York: E. P. Dutton.

Beavers, W. R., & Hampson, R. B. (1990). *Successful families: Assessment and intervention*. New York: Norton.

Bennun, I. (1989). Perceptions of the therapist in family therapy. *Journal of Family Therapy, 11,* 243–255.

Bentovim, A., & Davenport, M. (1992). Resolving the trauma organised system of sexual abuse by confronting the abuser. *Journal of Family Therapy, 14,* 29–50.

Bergin, A. E., & Garfield, S. L. (Eds.) (1994). *Handbook of psychotherapy and behavior change* (4th ed.). New York: Wiley.

Bertalanffy, L. von (1968). *General system theory*. Harmondsworth, UK: Penguin.

Boscolo, L., & Bertrando, P. (1993). *The times of time: A new perspective in systemic therapy and consultation*. New York: Norton.

Breunlin, D. C., Schwartz, R. C., Krause, M. S., Kochalka, J., Puetz, R. A., & Van Dyke, J. U. (1989). The prediction of learning in family therapy training programs. *Journal of Marital and Family Therapy, 15,* 387–395.

Brunning, H., & Huffington, C. (1990). Jumping off the fence—developing the consultancy model. *Clinical Psychology Forum, 29,* 31–33.

Burnham, J. (1992). Approach—method—technique: Making distinctions and creating connections. *Human Systems, 3,* 3–26.

Burr, W. R., Hill, R., Nye F. I., & Reiss, I. L. (Eds.) (1979a). *Contemporary theories about the family; Vol. 1. Research based theories*. New York: Free Press.

Burr, W. R., Hill, R., Nye, F. I., & Reiss, I. L. (Eds.) (1979b). *Contemporary theories about the family; Vol. 2. General theories/theoretical orientations*. New York: Free Press.

Campbell, D. (1985). The consultation interview. In D. Campbell & R. Draper (Eds.), *Applications of systemic family therapy* (pp. 193–202). New York: Grune and Stratton.

Campbell, D., Draper, R., & Huffington, C. (1989). *A systemic approach to consultation*. London: Karnac Books.

Carter, B., & McGoldrick, M. (1989). *The changing family lifecycle* (2nd ed.). New York: Allyn and Bacon.

Central Statistical Office (1990). *Social Trends 20*. London: HMSO.

Chamberlain, P., & Rosicky, J. G. (1995). The effectiveness of family therapy in the treatment of adolescents with conduct disorders and delinquency. *Journal of Marital and Family Therapy, 21,* 441–459.

Cronen, V. E., & Pearce, W. B. (1985). Towards an explanation of how the "Milan method" works: An invitation to systemic epistemology and evolution of family systems. In D. Campbell & R. Draper (Eds.), *Applications of systemic family therapy* (pp. 69–84). New York: Grune and Stratton.

Dimmock, B., & Dungworth, D. (1985). Beyond the family: Using network meetings with statutory child care cases. *Journal of Family Therapy, 7,* 45–68.

Dowling, E. (1979). Co-therapy: A clinical researcher's view. In S. Walrond-Skinner (Ed.), *Family and marital psychotherapy: A critical approach* (pp. 173–199). London: Routledge and Kegan Paul.

Elliott, R. (1984). A discovery-oriented approach to significant events in psychotherapy: Interpersonal process recall and comprehensive process analysis. In L. N.

Rice & L. S. Greenberg (Eds.), *Patterns of change*. New York: Guilford Press.

Emmelkamp, P. M. G., & Foa, E. B. (1983). Failures are a challenge. In E. B. Foa & P. M. G. Emmelkamp (Eds.), *Failures in behavior therapy*. New York: Wiley.

Falicov, C. J. (1988). Learning to think culturally. In H. Liddle, D. Breunlin, & D. Schwartz (Eds.), *Handbook of family therapy training and supervision*. New York: Guilford Press.

Framo, J. L. (1982). *Explorations in marital and family therapy: Selected papers of James L. Framo*. New York: Springer.

Freedman, J., & Combs, G. (1996). *Narrative therapy: The social construction of preferred realities*. New York: Norton.

Friedlander, M. L., Wildman, J., Heatherington, L., & Skowron, E. A. (1994). What we do and don't know about the process of family therapy. *Journal of Family Psychology, 8,* 390–416.

Friedman, S. (Ed.) (1995). *The reflecting team in action: Collaborative practice in family therapy*. New York: Guilford Press.

Frosch, S. (1992). Masculine ideology and psychological therapy. In J. Ussher & P. Nicolson (Eds.), *Gender issues in clinical psychology* (pp. 153–170). London: Routledge.

Gale, A. (1979). Problems of outcome research in family therapy. In S. Walrond-Skinner (Ed.), *Family and marital psychotherapy: A critical approach* (pp. 225–244). London: Routledge and Kegan Paul.

Goldenberg, I., & Goldenberg, H. (1991). *Family therapy: An overview*. Pacific Grove, CA: Brooks/Cole.

Goldner, V. (1985). Feminism and family therapy. *Family Process, 24,* 31–37.

Goldner, V. (1988). Generation and gender: Normative and covert hierarchies. *Family Process, 27,* 17–31.

Goldner, V., Penn, P., Sheinberg, M., & Walker, G. (1990). Love and violence: Gender paradoxes in volatile attachments. *Family Process, 29,* 343–364.

Greenberg, L. S. (1986). Change process research. *Journal of Consulting and Clinical Psychology, 54,* 4–9.

Gregory, M. A., & Leslie, L. A. (1996). Different lenses: Variations in clients' perception of family therapy by race and gender. *Journal of Marital and Family Therapy, 22,* 239–51.

Gurman, A. S., & Kniskern, D. P. (1978). Research on marital and family therapy: Progress, perspective and prospect. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis* (2nd ed., pp. 817–902). New York: Wiley.

Gurman, A. S., & Kniskern, D. P. (1991). Family therapy outcome research: Knowns and unknowns. In A. S. Gurman & D. P. Kniskern (Eds.), *Handbook of family therapy* (pp. xv–xxi). New York: Brunner/Mazel.

Gurman, A. S., & Kniskern, D. P. (1992). The future of marital and family therapy. *Psychotherapy, 29,* 65–71.

Gurman, A. S., Kniskern, D. P., & Pinsof, W. (1986). Research on the process and outcome of marital and family therapy. In S. Garfield & A. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed., pp. 565–624). New York: Wiley.

Haley, J. (1976). *Problem-solving therapy: New strategies for effective family therapy*. San Francisco: Jossey-Bass.

Hardy, K. V., & Laszloffy, T. A. (1995). The cultural genogram: Key to training culturally competent family therapists. *Journal of Marital and Family Therapy, 21,* 227–37.

Hare-Mustin, R. (1986). The problem of gender in family therapy theory. *Family Process, 26,* 15–27.

Hayward, M. (1996). Is second order practice possible? *Journal of Family Therapy, 18,* 219–42.

Hazelrigg, M. D., Cooper, H. M., & Borduin, C. M. (1987). Evaluating the effectiveness of family therapies: An integrative review and analysis. *Psychological Bulletin, 101,* 428–442.

Hoffman, L. (1993). *Exchanging voices*. London: Karnac.

Howe, D. (1989). *The consumer's view of family therapy*. London: Gower.

Jenkins, D. (1996). A reflecting team approach to family therapy: A Delphi study. *Journal of Marital and Family Therapy, 22,* 219–238.

Kingston, P., & Smith, D. (1983). Preparation for live consultation and live supervision when working without a one-way screen. *Journal of Family Therapy, 5,* 219-233.

Kuehl B. P., Newfield, N. A., & Joanning, H. (1990). A client-based description of family therapy. *Journal of Family Psychology, 3,* 310–321.

Lau, A. (1987). Family therapy and ethnic minorities. In E. Street & W. Dryden (Eds.), *Family therapy in Britain* (pp. 270–290). Milton Keynes, UK: Open University Press.

Lax, W. (1989). Systemic family therapy with young children in the family: Use of the reflecting team. *Journal of Psychotherapy and the Family, 5,* 55–74.

Liddle, H. A. (1991a). Empirical values and the culture of family therapy. *Journal of Marital and Family Therapy, 17,* 327–348.

Liddle, H. A. (1991b). Training and supervision in family therapy: A comprehensive and critical analysis. In A. S. Gurman & D. P. Kniskern (Eds.), *Handbook of family therapy* (Vol. II, 2nd ed., pp. 638–697). New York: Brunner/Mazel

Lieberman, S. (1979). *Transgenerational family therapy*. London: Croom Helm.

Lieberman, S. (1995). How I assess for family therapy. In C. Mace (Ed.), *Assessment in psychotherapy* (pp. 61–77). London: Routledge.

Loader, P., Burck, C., Kinston, W., & Bentovim, A. (1980). Method for organizing the clinical description of family interaction: The family interaction summary format. *Australian Journal of Family Therapy, 2,* 131–141.

Madanes, C. (1984). *Behind the one-way mirror*. London: Jossey-Bass.

Markus, E., Lange, A., & Pettigrew, T. (1990). Effectiveness of family therapy: A meta-analysis. *Journal of Family Therapy, 12,* 205–221.

Mason, B., & Mason, E. (1990). Masculinity and family work. In R. J. Perelberg & A. C. Miller (Eds.), *Gender and power in families* (pp. 209–220). London: Routledge.

McGoldrick, M., & Gerson, R. (1985). *Genograms in family assessment*. New York: Norton.

Minuchin, S. (1974). *Families and family therapy*. London: Tavistock.

Minuchin, S. (1984). *Family kaleidoscope*. London: Harvard University Press.

Minuchin, S., & Fishman, H. C. (1981). *Family therapy techniques*. Cambridge, MA: Harvard University Press.

Urgent W. R., Carpenter, D., & Parks, J. (1993). A statewide evaluation of family preservation and family reunification services. *Research on Social Work Practice, 3,* 40–65.

O'Brien, M. (1990). The place of men in a gender-sensitive therapy. In R. J. Perelberg & A. C. Miller (Eds.), *Gender and power in families* (pp. 195–208). London: Routledge.

Orlinsky, D. E., Grawe, K., & Parks, B. K. (1994). Process and outcome in psychotherapy. In A. Bergin & S. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 311–384). New York: Wiley.

Pam, A. (1993). Family systems theory: A critical view. *New Ideas in Psychology, 11,* 77–94.

Patterson, G. R., & Chamberlain, P. (1988). Treatment process: A problem at three levels. In L. C. Wynne (Ed.), *The state of the art in family therapy research: Controversies and recommendations* (pp. 189–226). New York: Family Process Press.

Patterson, G. R., & Forgatch, M. S. (1985). Therapist behavior as a determinant for client non-compliance: A paradox for the behavior modifier. *Journal of Consulting and Clinical Psychology, 53,* 846–851.

Perelberg, R. J., & Miller, A. C. (Eds.) (1990). *Gender and power in families.* London: Routledge.

Pinsof, W. M., & Catherall, D. R. (1986). The integrative psychotherapy alliance: Family, couple and individual therapy scales. *Journal of Marital and Family Therapy, 12,* 137–151.

Pinsof, W. M., & Wynne, L. C. (1995). The efficacy of marital and family therapy: An empirical overview, conclusions, and recommendations. *Journal of Marital and Family Therapy, 21,* 585–613.

Postner, R. S., Guttman, H. A., Sigal, J. J., Epstein, N. B., & Rakoff, V. M. (1971). Process and outcome in conjoint family therapy. *Family Process, 10,* 451–474.

Prince, S. E., & Jacobson, N. S. (1995). A review and evaluation of marital and family therapies for affective disorders. *Journal of Marital and Family Therapy, 21,* 377–401.

Pulleybank, E., & Shapiro, R. J. (1986). Evaluation of family therapy trainees: Acquisition of cognitive and therapeutic behavior skills. *Family Process, 25,* 591–598.

Reimers, S., & Treacher, A. (1995). *Introducing user-friendly family therapy.* London: Routledge.

Selvini-Palazolli, M., Boscolo, L., Cecchin, G., & Prata, G. (1980a). Hypothesizing—circularity—neutrality: Three guidelines for the conductor of the session. *Family Process, 19,* 3–12.

Selvini-Palazolli, M., Boscolo, L., Cecchin, G., & Prata, G. (1980b). The problem of the referring person. *Journal of Marital and Family Therapy, 6,* 3–9.

Shadish, W. R., Montgomery, L. M., Wilson, P., Wilson, M. R., Bright, I., & Okwumabua, T. (1993). Effects of family and marital psychotherapies: A meta-analysis. *Journal of Consulting and Clinical Psychology, 61,* 992–1002.

Shadish, W. R., Ragsdale, K., Glaser, R. R., & Montgomery, L. M. (1995). The efficacy and effectiveness of marital and family therapy: A perspective from meta-analysis. *Journal of Marital and Family Therapy, 21,* 345–360.

Shapiro, R. J. (1974). Therapist attitudes and premature termination in family and individual therapy. *Journal of Nervous and Mental Disease, 159,* 101–107.

Shazer, S. de, Kimberg, I., Lipchik, E., Nunnally, E., Molnar, A., Gingerich, W., & Weiner-Davis, M. (1986). Brief therapy: Focused solution development. *Family Process, 25,* 207–222.

Shields, C. G., Sprenkle, D. H., & Constantine, J. A. (1991). Anatomy of an initial interview: The importance of joining and structuring skills. *American Journal of Family Therapy, 19,* 3–18.

Simmons, D. S., & Doherty, W. J. (1995). Defining who we are and what we do: Clinical practice patterns of marriage and family therapists in Minnesota. *Journal of Marital and Family Therapy, 21,* 3–16.

Skynner, A. C. R. (1976). *Systems of family and marital psychotherapy.* New York: Brunner/Mazel.

Street, E. (1988). Family therapy training research: Systems model and review. *Journal of Family Therapy, 10,* 383–402.

Street, E., Downey, J., & Brazier, A. (1991). The development of therapeutic consultations in child-focused family work. *Journal of Family Therapy, 13,* 311–334.

Tomm, K. M. (1984). One perspective on the Milan approach: Part 1. Overview of development, theory and practice. *Journal of Marital and Family Therapy, 10,* 113–125.

Treacher, A. (1989). Termination in family therapy—developing a structural approach. *Journal of Family Therapy, 11,* 135–148.

Treacher, A., & Carpenter, J. (1983). On the failure to take convening strategies seriously—a reply to Campion. *Journal of Family Therapy, 5,* 259–262.

Trost, J. (1990). Do we mean the same by the concept of family? *Communication Research, 17,* 431–443.

VandenBos, G. R., & Stapp, J. (1983). Service providers in psychology: Results of the 1982 APA human resources survey. *American Psychologist, 38,* 1330–1352.

Verbrugge, L. (1985). Gender and health: An update on hypothesis and evidence. *Journal of Health and Social behavior, 26,* 156–182.

Vetere, A. L. (1987). General system theory and the family: A critical evaluation. In A. Vetere & A. Gale (Eds.), *Ecological studies of family life* (pp. 18–33). Chichester, UK: Wiley.

Vetere, A. L., & Gale, A. (1987). *Ecological studies of family life.* Chichester, UK: Wiley.

Wachtel, E. F., & Wachtel, P. L. (1986). *Family dynamics in individual psychotherapy: A guide to clinical strategies.* New York: Guilford.

Waldegrave, C. (1990). Social justice and family therapy. *Dulwich Centre Newsletter*, No. 1. Adelaide, Australia: Dulwich Centre Publications.

Walrond-Skinner, S., & Watson, D. (1987). *Ethical issues in family therapy.* London: Routledge and Kegan Paul.

Weakland, J. H., Fisch, R., Watzlawick, P., & Bodin, A. M. (1974). Brief therapy: Focused problem resolution. *Family Process, 13,* 141–168.

Wheeler, D., Avis, J. M., Miller, L. A., & Chaney, S. (1989). Rethinking family therapy training and supervision: A feminist model. In M. McGoldrick, C. M. Anderson, & F. Walsh (Eds.), *Women in families: A framework for family therapy* (pp. 135–151). New York: Norton.

Whiffen, R., & Byng-Hall, J. (Eds.) (1982). *Family therapy supervision.* London: Academic Press.

White, M., & Epston, D. (1990). *Narrative means to therapeutic ends.* New York: Norton.

Williams, J., & Watson, G. (1987). Sexual inequality, family life and family therapy. In E. Street & W. Dryden (Eds.), *Family therapy in Britain* (pp. 291–311). Milton Keynes, UK: Open University Press.

Wynne, L., McDaniel, S., & Weber, T. (1986). *Systems consultation: A new perspective for family therapy.* New York: Guilford Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.05

Psychodynamic Approaches

PETER FONAGY

University College London, UK

6.05.1 THE ORIGINS OF PSYCHODYNAMIC THERAPY 108
  6.05.1.1 Definition of Psychodynamic Therapy 108
  6.05.1.2 The Short-term Emphasis in Psychodynamic Therapy 108
6.05.2 TECHNICAL ISSUES IN PSYCHODYNAMIC THERAPY 109
  6.05.2.1 Suitability for Psychodynamic Therapy 109
  6.05.2.2 Formulation of Patients' Problems 109
  6.05.2.3 Mechanisms of Defense 110
    6.05.2.3.1 The concept of defense in various theoretical frameworks 110
    6.05.2.3.2 Primitive defenses 110
    6.05.2.3.3 Neurotic and mature defenses 110
  6.05.2.4 The Context of Psychotherapy 111
    6.05.2.4.1 The therapeutic contract 111
    6.05.2.4.2 Abstinence and neutrality 111
    6.05.2.4.3 Regression 111
    6.05.2.4.4 Resistance 111
  6.05.2.5 Transference 112
    6.05.2.5.1 The history of the concept 112
    6.05.2.5.2 The origin of the transference experience 112
    6.05.2.5.3 The limits of transference interpretations 112
    6.05.2.5.4 Special forms of transference 113
  6.05.2.6 Counter-transference 113
    6.05.2.6.1 History of the concept 113
    6.05.2.6.2 Types of counter-transference 113
    6.05.2.6.3 Counter-transference and interpersonal approaches 114
  6.05.2.7 Therapeutic Interventions 114
    6.05.2.7.1 Historical overview 114
    6.05.2.7.2 Supportive and directive aspects of psychodynamic interventions 114
    6.05.2.7.3 Interpretation in psychodynamic therapy 115
  6.05.2.8 The Result of Psychodynamic Interventions 116
    6.05.2.8.1 The role of insight 116
    6.05.2.8.2 The role of working through 116
  6.05.2.9 Ending Treatment 116
    6.05.2.9.1 Indications for ending treatment 116
    6.05.2.9.2 The process of ending treatment 116
6.05.3 ILLUSTRATION OF LONG-TERM PSYCHODYNAMIC THERAPY: PSYCHOTHERAPY FOR BORDERLINE PERSONALITY DISORDER 117
  6.05.3.1 Theoretical Approaches to Borderline Conditions 117
  6.05.3.2 Treatment Strategies 117
    6.05.3.2.1 Limits and boundaries 117
    6.05.3.2.2 Interpretive focus 118
    6.05.3.2.3 Counter-transference 118
  6.05.3.3 Handling Crises 118
    6.05.3.3.1 Desirable outcomes 118
  6.05.3.4 Outcome Research 118
6.05.4 FORMS OF BRIEF PSYCHODYNAMIC THERAPY 119

6.05.4.1 The Historical Roots of Brief Psychodynamic Therapy 119
6.05.4.2 Indications for Brief Psychodynamic Therapy 119
6.05.4.3 Techniques of Brief Psychodynamic Therapy 120
6.05.4.3.1 Malan's brief intensive psychotherapy (BIP) 120
6.05.4.3.2 Sifneos' short-term anxiety-provoking psychotherapy (STAPT) 121
6.05.4.3.3 Davanloo's intensive short-term dynamic psychotherapy (ISTDP) 121
6.05.4.3.4 Luborsky's supportive-expressive time-limited therapy (SETLT) 122
6.05.4.3.5 Strupp's time-limited dynamic therapy (TLDP) 123
6.05.4.3.6 Weiss and Sampson's plan formulation method 123
6.05.4.3.7 Horowitz's person schema theory 124
6.05.4.3.8 Mann's time-limited psychodynamic therapy (TLPT) 125
6.05.4.3.9 Ryle's cognitive analytic therapy (CAT) 126
6.05.4.3.10 Hobson's conversational model 128
6.05.5 CONCLUSIONS 128
6.05.6 REFERENCES 129

6.05.1 THE ORIGINS OF PSYCHODYNAMIC THERAPY

6.05.1.1 Definition of Psychodynamic Therapy

The term "psychodynamic psychotherapy" covers a somewhat heterogeneous range of psychological interventions which draw their inspiration from psychoanalytic theory. Various implementations of this form of treatment (see Chapter 14, Volume 1) emphasize different aspects of this rich body of ideas including: (i) notions of psychic conflict as an all-pervasive aspect of human experience; (ii) the internal organization of the mind to avoid unpleasure arising out of conflict and maximize a sense of the experience of safety; (iii) the use of defensive strategies for the adaptive manipulation of ideas and experience to minimize unpleasure; (iv) a developmental view of psychopathology as arising out of the long-term adverse consequences of adaptations at earlier phases of development; (v) the organization of experience in terms of the internal representations of relationships between self and other throughout the life span; and (vi) the expectable re-emergence of these experiences in the relationship with the therapist. Psychodynamic therapies are predominantly verbal and interpretive aimed at the modifications of restructuring of the representations of relationships primarily, but not exclusively, through the use of insight.

6.05.1.2 The Short-term Emphasis in Psychodynamic Therapy

Psychodynamic psychotherapies are rooted in long-term, time-unlimited approaches, and many of the most important advances in theory and technique originate from this context. The emphasis in the present chapter will, however, be on short-term psychodynamic approaches. The rationale for this emphasis may be stated as follows: The demand for psychotherapeutic treatment as either an adjunct to pharmacotherapy or as a unique treatment is considerable and has increased substantially over recent years as a function of cultural, social, and economic changes (e.g., lessened stigma attached to seeking assistance with psychological problems, increase in some risk factors for mental illness such as drug and alcohol use) and the greater availability of providers (Howard et al., 1996).

Surveys of the delivery of psychotherapy in both public and private health care settings demonstrate that the majority of clients receive relatively short-term treatment. The majority of patients receive less than 10 sessions (DeLeon, VandenBos, & Bulatao, 1991; Olfson & Pincus, 1994), although therapists spend the majority of their time with the relatively small proportion (15–20%) who attend more than 26 sessions (Howard, Davidson, O'Mahoney, Orlinsky, & Brown, 1991; Taube, Kessler, & Feuerberg, 1984) and thus may come to believe that long-term treatment is more common than it actually is. It should be remembered however that estimates of the mean length of treatment for psychotherapy are reduced by the contribution that general (primary care) physicians make to this type of mental health care provision (approximately 30%), the majority of whom provide extremely brief forms of treatment (one to three sessions) (Olfson & Pincus, 1994; Olfson, Pincus, & Dial, 1994).

Managed care, with its declared intention of cost containment for health care provision, requires a more structured, focused, treatment-goal-oriented psychotherapeutic approach (Brosowski, 1995; Richardson & Austad, 1991). There has been a shift in the formulation of the nature of both psychological disorder and interventions aimed to address it from an acute to a chronic illness model: certain individuals are seen as requiring therapeutic input, briefly, but repeatedly, throughout the life cycle, especially in relation to life crises (Cummings, 1988; Kazdin, 1988; Roth & Fonagy, 1996).

6.05.2 TECHNICAL ISSUES IN PSYCHODYNAMIC THERAPY

6.05.2.1 Suitability for Psychodynamic Therapy

Whereas medical treatments tend to have clear indications and contraindications for specific interventions, the emphasis of assessment for psychodynamic therapy tends to emphasize the looser notion of general "suitability" (Tyson & Sandler, 1971). Nevertheless, some writers have made specific recommendations concerning suitability for long-term (e.g., Coltart, 1988) and short-term therapy (e.g., Malan, 1980). While some authors have suggested relatively systematic assessment, which yields both diagnostic and prognostic features (e.g., Kernberg, 1981), in terms of the developmental level of the patient's personality organization most psychodynamic clinicians rely on clinical judgments based on interpersonal aspects of their first meeting with the patient (Etchegoyen, 1991). The three areas of greatest relevance to suitability are: patients' personal history, content of the interview, and style of presentation.

In terms of the patient's history, good evidence of personal achievement and at least one good relationship has been traditionally regarded as a good prognostic indicator (Malan, 1980). A history of psychotic breakdown, severe obsessional states, somatization, and a lack of frustration tolerance are normally regarded as contraindications. There are many psychodynamic clinicians who report working successfully with patients who in the past had been regarded as unsuitable because of their histories: for example, psychotic (Rosenfeld, 1952), learning disabled (Sinason, 1992), and chronically poorly controlled diabetic individuals (Moran, Fonagy, Kurtz, Bolton, & Brook, 1991). Psychodynamic clinicians treating clients in long-term psychotherapy, particularly those working intensively, are more likely to make such risky selections.

There is even less agreement about suitability concerning the content of assessment interviews. In general, clinicians tend to make their judgment on the basis of the presence of a "mutuality" between them and the patient and observed responses to "trial interpretations" in which they attempt to summarize their initial impressions concerning the patients' presenting problem in the context of their current and past life situation (see below). In addition, therapists may try to identify if the patient has the capacity to respond emotionally within the sessions to allow feelings of fear, sadness, or anger to come to the surface (Orlinsky, Grawe, & Parks, 1994). Some clinicians, though by no means all, attempt to assess the patient's motivation for treatment. This is, however, hard to do as superficial expression of commitment may or may not predict the willingness to confront unpleasant aspects of oneself. For this reason it is probably impossible to obtain genuinely informed consent in psychodynamic psychotherapy research.

More recently, psychodynamic therapists have paid more attention to the nature of the patient's discourse rather than just its content. Holmes (1995), for example, observes if the patient's narrative styles are markedly avoidant and dismissing of interpersonal issues or appear entangled and enmeshed with them with much current anger about past hurts and insults. There is at least one study that shows that the former type of patient has a somewhat better prognosis in psychodynamic therapy (Fonagy et al., 1996). Patients' narrative style may also be a clue to the extent of their reflectiveness: the capacity to see oneself from the outside (Sandler, Dare, & Holder, 1992), autobiographical competence (Holmes, 1992), the ability to reflect on one's inner world (Coltart, 1988), and fluidity of thought (Limentani, 1972) are all considered indicators of suitability.

6.05.2.2 Formulation of Patients' Problems

Psychodynamic theory is too diverse to permit definitive formulations. Formulations identify central unconscious conflicts, maladaptive defenses, unhelpful unconscious fantasies and expectations, deficits in personality development and the like. Formulation depends on the theoretical orientation of the psychodynamic clinician. In Chapter 14, Volume 1 of this work we have reviewed the range of currently popular orientations. Agreements, however, are hard to reach even when clinicians follow the same orientation (Horowitz, Rosenberg, Ureño, Kalehzan, & O'Halloran, 1989). Some standardized approaches have, however, been developed (Perry, Cooper, & Michels, 1987; Perry, Luborsky, Silberschatz, & Popp, 1989).

While there is no generally accepted schema for formulations, there are several key parameters that clinicians generally consider. These are: (i) the extent to which representations of relationships are mature, that is, involve three or more persons rather than just a two-person, self–other dimension (Karasu, 1990); (ii) the quality of psychic defenses, particularly the predominance of primitive defenses rather than more mature ones (Vaillant, 1992, and see below); (iii) the extent of whole, as opposed to part object relations (where individuals are represented as whole persons rather than just an aspect or a function of a person, e.g., feeding or nurturance, sexual gratification, a container for evacuation) (Kernberg, 1984).

Considerations such as these usually serve two functions. The first of these is to suggest the likely effectiveness of the type of treatment: short vs. long-term, intensive vs. nonintensive psychodynamic therapy. On the whole, patients seen as more severe on parameters such as the three suggested above are less likely to do well according to most studies of psychodynamic treatment (e.g., Wallerstein, 1986). The second function of formulations is to give an initial focus to the clinical work, which in the case of brief therapy may be the sole focus of the treatment. In long-term therapy these formulations tend to change, sometimes radically, on the basis of information emerging in the course of treatment. Winnicott (1965) referred to psychodynamic treatment as "an extended form of history taking."

6.05.2.3 Mechanisms of Defense

6.05.2.3.1 The concept of defense in various theoretical frameworks

Within classical theory, which sees conflict as the core of mental function (e.g., Brenner, 1982), defenses are seen as adaptations to intrapsychic conflict. Within object relations theories, defenses are seen as assisting the individual to maintain an authentic self-representation, a true (Winnicott, 1965) or nuclear (Kohut, 1984) self. Within attachment theory, defenses are seen as maintaining desirable relationships (Holmes, 1993). Within a Klein–Bion frame of reference, defenses are often conceived of in terms of complex structures or systems called organizations. The term underscores the relative inflexibility of some defensive structures. Personality types appear to be characterized by specific types of defensive organizations. For example, in narcissistic disorders idealization and destructiveness, the devaluation of genuine love and truth, may have been protective at one developmental stage but came to acquire a stability which seems to be based on the emotional pay-off from this form of adaptation (Rosenfeld, 1964, 1971; Steiner, 1982).

6.05.2.3.2 Primitive defenses

Regardless of theoretical orientation, there is general agreement about common forms of defense and their relative sophistication from a developmental standpoint (Vaillant, 1992). The primitive defenses tend to be found together in certain individuals. Borderline individuals, who idealize and then derogate their therapists in order to maintain their self-esteem, are using splitting and projection. Projective identification (Klein, 1946) is also common in this group. Projection, the simple ascribing of an undesirable mental state to the other, becomes a more powerful mechanism when the other can be unconsciously forced to accept the projection and experience its impact, thus the "identification" is in the recipient of the projection. This current interpretation of the term is particularly well described by Ogden (1979). It clearly goes beyond Klein's original meaning: the patient fantasizing that the recipient of the projection acquired a part of the patient's self. Spillius (1994) suggests a helpful clarification; she calls projective identification which invites the therapist to actualize the projection, "evocative projective identifications."

Either in fantasy or in actualized form, projective identification offers a primitive method of experiencing control over the other within a relationship (therapeutic or care-giving) (Bion, 1955). When parts of the self are experienced as being within another person, the individual frequently makes attempts to control these split-off parts by exerting total control over the recipient or container of the projection. Bion (1962) also stresses that good, as well as undesirable, aspects of the self may be externalized via this route, making projective identification one of the principle avenues for communication in infancy. Other functions of projective identification, beyond communication and control, include: the acquisition of the object's attributes in fantasy, the protection of a good quality from internal persecution by evacuating it into an object, and the avoiding or denying of separateness. It is thus a fundamental aspect of an interpersonal relationship focused on fantasy, and its appreciation is critical for adequate psychodynamic psychotherapeutic practice.

6.05.2.3.3 Neurotic and mature defenses

Anna Freud (1936) is to be credited for the delineation of most of the mechanisms of defense commonly used in clinical formulations today. It is not possible to give detailed consideration to each of these and most are, by now, part of common parlance (and exist as terms of mild rebuke, at least between mental health professionals). Defenses involving access to mental representations (e.g., repression) or the attribution of emotional significance to these (e.g., denial or disavowal) are perhaps most commonly encountered as part of individual coping strategies. For example, jokes containing emotionally threatening, sexual, and aggressive material are notoriously hard to remember. Denial of the emotional significance of incurable disease may be helpful in increasing the patient's chance of survival (Greer, Morris, & Pettingale, 1979). Turning a response into its opposite—terror into aggression (reaction formation)—is particularly common in children in response to abuse, and identification with the aggressor (becoming the tormentor instead of the victim) is common amongst maltreated children who become abusers in adulthood (Oliver, 1993). Other commonly noted neurotic defenses include: undoing (performing a magical reparative act), intellectualization and rationalization, and humor.

The diagnostic significance of mechanisms of defense is controversial. Some workers claim predictive specificity for such assessments beyond that of psychiatric diagnosis (Perry et al., 1989; Vaillant, 1992). The theoretical ambiguity which surrounds the concept makes its widespread use as a diagnostic device unlikely, at least in the short term. Its inclusion in this chapter is justified by its heuristic value: conceptualizing patients' reactions in terms of defenses makes psychodynamic work practicable. All of these defenses are common reactions in psychodynamic therapy, particularly at times of emotionally challenging work. Recognizing them has to become part of the working routine of the psychotherapist. For example, the forgetting of material from last week's session is more common than remembering it, even amongst people whose memory for other, far more trivial, aspects of their daily life is nothing short of admirable. The psychotherapist recognizes the patient's defenses and mostly uses them as an indication of the presence of underlying anxiety rather than as occasion for confrontation (but see some brief therapeutic strategies below).

6.05.2.4 The Context of Psychotherapy

6.05.2.4.1 The therapeutic contract

Although contracts are far more relevant for short-term than for long-term therapy, most psychodynamic therapists explicitly or implicitly convey objectives and expectations to their clients. The details of the agreement usually include time and place, the length and frequency of sessions and an initial idea concerning the likely duration of therapy, the expected behavior of the patient and the therapist and so on. It is the emotional context of this agreement that is often more important than the items which it includes. It implies a mutual recognition by both parties of the importance placed upon protecting the process of therapy.

6.05.2.4.2 Abstinence and neutrality

Freud (1915) originally wrote about the analyst resisting the temptation of gratifying the patient's sexual desires. Although this is clearly an ethical issue, it also pertains to analysts forgoing gratifying patients' curiosity, or using patients to gratify their own personal needs. Equally, the patient undertakes to forgo major life changes where these are currently the subject of psychotherapeutic attention. Whereas this is particularly important in short-term treatment, even long-term psychodynamic treatment can flounder if the emotional experiences of the therapy are obscured by the upheavals from major life events.

Abstinence ensures therapeutic neutrality. Psychodynamic therapists go to great lengths not to direct their patients' associations and remain neutral no matter what the content of the patients' past experiences or fantasies may be. Although this is frequently caricatured, it is important for the psychodynamic therapist to retain a certain emotional distance from the client, sufficient for the latter's fantasies and hidden fears to emerge. Needless to say, neutrality must be balanced by sensitivity. The recent literature on the working alliance (see below) underscores that the therapist's genuine concern for the client must come through if significant progress is to be made.

6.05.2.4.3 Regression

One aim of psychodynamic psychotherapy is to activate and explore aspects of personality functioning which are normally obscured behind the patient's need to adapt to the demands of everyday reality. Psychodynamic therapy, to a degree, encourages access to these representations through the process of regression. This is less of an active encouragement than a passive permission. Sandler and Sandler (1994) suggested that an "antiregressive function" is normally active in all of us, and psychotherapy can function as a way of disinhibiting this function in much the same way as intimate interpersonal and certain types of large group situations and alcohol appear to do. Some therapists consider regression to be an essential part of successful treatment (Balint, 1949; Winnicott, 1971). Certainly the fear of regression can become an important source of resistance (Sandler & Sandler, 1994).

6.05.2.4.4 Resistance

Resistance is an essential component of psychodynamic treatment. The word dynamic implies the presence of forces rallied against change. Resistance, like regression, fluctuates in intensity across treatment and may serve different functions at different times and in different individuals. In narcissistic and borderline states, resistance may protect the patient's self-esteem (Kernberg, 1988; Kohut, 1984); in more neurotic cases it may focus around preventing the integration of experience (Thomä & Kächele, 1987).

It may also take different forms. Clients access to their mental life (e.g., remembering dreams) may be experienced as restricted (repression resistance). Others may wish to keep their relationship with their therapist at an extremely superficial level (transference resistance). Some patients show a paradoxical reaction to their treatment: The better it progresses the worse they feel (e.g., new symptoms emerge). Freud (1923) attributed this to unconscious guilt. In some patients, at least, it is more likely to be linked to their pervasive predisposition (envy) to destroy all things in their life which they experience as good but beyond their immediate control (Kernberg, 1975; Rosenfeld, 1975).

6.05.2.5 Transference

6.05.2.5.1 The history of the concept

The concept of transference originated with resistance. In his work with the cathartic method (Breuer & Freud, 1895) Freud initially understood the patient's intense emotional reaction to the therapist as an interference with what was an essentially verbal method. He quickly realized that the patient's emotional experiences could be better understood as representations of earlier relationship experiences (particularly Oedipal strivings and disappointments) which could be understood in terms of past experience and in fact made that experience emotionally more meaningful for the patient (Freud, 1914).

6.05.2.5.2 The origin of the transference experience

The patient's intense feelings of anger, suspicion, and disappointment, or love, admiration, and excitement about the person of the therapist are unlikely to be a realistic response, since two patients may experience such opposite feelings towards the same therapist at similar times. There are currently a range of views about the origin of transference experiences. Some believe that transference is based on actual past experience: the expression of expectations based on past experiences of relationships (Bowlby, 1980). In this view, transference distorts the actual relationship; patients see what they expect to see and resist understanding by repeating past relationships in a rigid manner (Levenson, 1983). There is an implication in this view of a "correspondence" between present and past. Many therapists are reluctant to accept the idea of such a direct relationship. For them transference is an aspect of the patient's narrative; it gives "coherence" to the patient's experience of the therapeutic relationship, but it is not an expression of an underlying truth about the patient's history (Gill, 1982; Spence, 1982). Transference is constructed in the present and cannot give us clues to the past. By contrast, some who work within a Klein–Bion perspective see transference as a unique window to the patient's current internal reality (e.g., Joseph, 1986). For example, confronted with an idealized transference they might see the patient's true state as organized around psychotic anxieties related to the death instinct. The idealization is a desperate defense against the destruction which they fantasize may engulf both them and the therapist. Cavell (1994) demonstrated that the dialectic between these conceptions of the transference has important philosophical roots in the debate between "correspondence" and "coherence" models of truth.

Additional suggestions by object relations theorists have in some ways complicated the issue. Here transference is based neither on expectations, defenses against expectations, nor constructions aimed to achieve coherence, but rather on a process of externalization of internal objects (Kernberg, 1984). Such representations cannot be seen as true or false; they are distorted by fantasy and defensive processes.

6.05.2.5.3 The limits of transference interpretations

There is a further important debate with regard to the relevance of a transference focus in psychodynamic therapy. Some therapists are inclined to see transference as covering all aspects of the analytic situation. Joseph, (1986) for example, sees the total situation in the therapy as reflecting an internal state of affairs in the patient's mind. Thus other aspects of the therapeutic relationship, for example, the so-called alliance and the so-called real relationship, are all subsumed under the transference. By contrast, Strachey (1934) conceived of transference as externalization of the patient's superego. The therapist, unlike other people in the patient's life, does not accept this externalization, be it idealized, denigrated, or judgmental, and conveys his or her understanding of it by so-called "mutative interpretations." This view implies that the only truly therapeutic interpretations are those which involve the transference, but clearly admits of other aspects of the relationship. Other therapists, particularly in the Freudian school (e.g., Anna Freud, 1936), see transference interpretations as just one category of interpretive intervention and give them relatively little particular priority.

6.05.2.5.4 Special forms of transference

There are several psychodynamic observations concerning special forms of transference which are helpful to particular groups of patients. Kohut (1984) described some common transference patterns with narcissistic patients. Individuals whose self-objects (parents) failed in their mirroring function may continue to crave for approbation and admiration and manifest this pattern in therapy. This is termed the mirroring transference, which Kohut recommends the therapist should not puncture by premature interpretations. Idealizing transference similarly aims to meet defective self-esteem by vicariously identifying with the perfect therapist. It is a highly controversial recommendation of Kohut's for therapists to accept such transferences. On the one hand, behind such exaggerated accolades may lie deep frustration and even rage; on the other, if Kohut is correct, destroying this image of the therapist is tantamount to a destruction of the patient's self-regard.

Erotic transference is also thought to have infantile roots and occurs relatively commonly in therapeutic relationships, whatever the gender of the therapist and the patient. It tends to be highly embarrassing for the patient and sometimes also for the therapist. Most agree that it is a distraction and some suggest that it is a manifestation of a defense against either recognizing damage done to the object or the fragmented state of the self (e.g., Steiner, 1982). Alternatively, from an attachment theory perspective, erotic transference may be a way of forcing the unresponsive object to pay attention. At the extreme end of this dimension is erotized transference where the demand for sexual gratification is not experienced by the patient as unrealistic (Etchegoyen, 1991). Such reactions are particularly common in severely traumatized individuals.

6.05.2.6 Counter-transference

6.05.2.6.1 History of the concept

Counter-transference in its broadest sense refers to the thoughts and feelings of the therapist during a treatment session which are in some ways relevant to the patient's current experience and thus may be involved in illuminating the patient's reaction or, indeed, obscuring it. Langs (1976) usefully distinguished between counter-transferences which may be attributed to the patient and those which result from the analyst's neurotic reaction to some aspect of the therapeutic situation.

The concept of counter-transference, like that of transference, is rooted in the notion of resistance. Freud, throughout his career, never recognized the value of counter-transference as an indicator of the patient's subjective state (Freud, 1912). Heimann (1950) expanded the usefulness of the concept by pointing out that the therapist's feelings and thoughts about the patient's communications, if reflected upon, could provide important clues about the patient's current state of mind. The idea was controversial (e.g., Fliess, 1953) yet it gradually gained acceptance. Those psychotherapists who were committed to an interpersonalist tradition (e.g., Sullivan, 1953), and saw the relationship aspect of psychotherapy as its most important facet, welcomed this expansion of the counter-transference concept. From their point of view the omniscient neutral analyst was an anachronistic anathema; accepting the analyst's human reaction was a welcome democratic humanistic development (Abend, 1989).

6.05.2.6.2 Types of counter-transference

Racker (1968) distinguished between complementary and concordant counter-transference. Concordant counter-transferences are based upon primitive empathic processes within the therapist. The latter resonates with as yet unverbalized experiences of the patient. Complementary counter-transferences arise when the patient treats the analyst in a manner congruent with an earlier relationship pattern. King's (1978) notion of "reverse transference," where patients treat the therapist as they felt treated when children, is a special instance of this category.

The case of concordant counter-transference particularly raises the issue of the mechanisms by which analysts may become aware of the patient's unconscious experience. Sandler (1993) offers the notion of "primary identification" as a rapid process of automatically mirroring one's partner in a communication and only becoming aware of this upon reflection. There is good evidence for such a process, particularly in the infant development literature (Osofsky, 1995).

Only when such primary identifications touch on unconscious conflicts within the analyst, with a consequent mobilization of defensive processes, will counter-transference start to be a distorting process. Grinberg (1962) pointed out that such experiences could lead therapists to withdraw from the therapeutic relationship. Therapists sensitive to, and resonating with, feelings of inadequacy in patients may be made anxious by such feelings and become defensively angry or hypermotivated to show their effectiveness. To some degree this may be an inevitable part of the process with the therapists only gradually understanding the reasons behind their unusually defensive style of relating. Bion (1962) stressed the importance of the recipient of the projection (the container) being able to "metabolize" and feed back the products of such primitive communications. Sandler (1976b) pointed out that therapists needed to allow themselves sufficient freedom of action to be able first to enact the patient's projection and then respond to the enactment in appropriate ways. Along similar lines, Brenman-Pick (1985) cautioned that the therapists' psychopathology frequently required that they worked through counter-transference identifications so that their wish not to know was turned into a potential for knowing.

Not all counter-transference reactions are provoked by patients' projections or reactions to the anxieties that these create. Winnicott (1949) was perhaps the first to point out that the outrageous provocative behavior of certain borderline or psychotic patients produced a normal reaction of "objective hate." Kernberg (1984) formalized this, suggesting that such behaviors activated primitive aspects of the therapist's personality. Understanding these reactions naturally helps the therapist empathize but it is neither a complementary nor concordant reaction, nor is it a defense against affective resonance. It is simply one further indication that the therapist too is human.

6.05.2.6.3 Counter-transference and interpersonal approaches

Modern psychodynamic theory considers counter-transference as firmly located in the interpersonal field. The patient and psychotherapist are seen as two mutually influencing psychological systems (Langs, 1978). Both transference and counter-transference are seen as the product of a subtle interplay between conscious and unconscious systems of both patient and analyst. Influences occur in both directions at both conscious and unconscious levels (Arlow, 1993). Some writers go so far as to suggest that patient and therapist share an unconscious fantasy of creating an intersubjective field between them (Baranger, 1993). Whether correct or not, these modern approaches emphasize that it is no longer possible to consider the therapist's role as "neutral" or "mirror-like" and that the psychotherapeutic process is a highly subjective admixture of a range of complementary processes that establish themselves in a unique configuration for each therapy.

6.05.2.7 Therapeutic Interventions

6.05.2.7.1 Historical overview

There has been a dialectic throughout the history of psychodynamic approaches between orientations that emphasize insight and interpretation and those that stress the unique emotional relationship with the therapist as the primary vehicle of change. The controversy dates back to the earliest days of psychoanalysis, to Freud and Ferenczi, but re-emerged powerfully with first the work of Balint and Winnicott set against classical Freudian and Kleinian theorists, and later Kohut and self psychology opposing the classical ego psychology tradition. This dichotomy is clearly artificial, as it is hard to envision effective psychodynamic therapy without both components (see Wallerstein, 1992). It is nonetheless helpful in presenting the range of techniques used by most psychodynamic therapists which span the spectrum between emotional relationship-oriented techniques and insight-oriented ones. Winnicott (1971) referred to this dimension as "being with" versus "doing to." As supportive interventions tend to suggest particular lines of association, an overlapping dichotomy exists between the use of directive and nondirective techniques.

6.05.2.7.2 Supportive and directive aspects of psychodynamic interventions

There are a whole range of supportive techniques used more or less deliberately by psychodynamic psychotherapists. These include: explicit support and affirmation; concern and sympathy (e.g., in response to bereavement or major setback); reassurance, commonly concerning irrational anxieties about the therapeutic arrangements; empathy with the patients' painful internal struggles, and so on. The complexity of such interventions was well illustrated by Feldman (1993). He showed how patients may experience the therapist's submission to a demand for reassurance as anxiety provoking rather than calming, since they may be unconsciously aware that the therapist's genuine position is incompatible with offering such palliatives and are thus faced with the deep fears concerning their own omnipotence and the therapist's weakness. Nevertheless, Kohut's (1984) emphasis on empathy was undoubtedly a welcome counterbalance to the rigid interpretive techniques of many classical ego psychologists and seems appropriate with patients who have experienced little genuine concern in their history.

Psychodynamic therapists are most likely to use supportive and directive techniques which encourage the therapeutic process. Elaborative techniques, such as "Tell me more about that," undoubtedly direct the patient's attention to specific issues and focus the treatment but may be essential precursors to interpretive work proper. Clarification is also a commonly used technique which involves a restatement of the patient's utterance or just offering a label (a symbol) for an internal state which the patient has limited capacity to represent. Clarifications shade into interpretations and the distinction is hard to draw in actual practice. Confrontation is a subtype of elaboration and clarification. The therapist, having identified inconsistency, brings this to the attention of the patient, usually in a supportive but firm manner (e.g., drawing the patient's attention to the absence of affect appropriate to a specific situation).

6.05.2.7.3 Interpretation in psychodynamic therapy

Interpretation is the paradigmatic intervention. Perhaps not surprisingly therefore it is often idealized as the sole or uniquely effective method for bringing about psychic change. Menninger (1958) offered a useful classification for psychodynamic interpretation. He suggested that interpretations addressed principally one of three aspects of a conflict: the defense, the anxiety, or the underlying wish or feeling. The content of the interpretation could be further subdivided as concerning current external reality, the transference, and childhood relationships.

The phase of the treatment is most likely to determine the therapist's choice. Commonly interpretations move from current events through the transference to the distant past. Similarly, generally interpretations start with the anxiety, through the identification of the defense to accessing the repudiated affects. In practice, neither of these patterns can be considered more than a guideline to be loosely applied. For example, very long-term therapy often ends up focusing on supportive exploration of current events (Blum, 1989). This is not surprising since the distant past can only be worked over a limited number of times. Similarly, defenses, if interpreted without regard to the anxiety, may well come across as implicit criticisms of the patient. Yet early interpretations of an unconscious wish can be puzzling and confusing even for relatively intact patients. Furthermore, severely personality disordered individuals respond poorly to interpretations of the distant past. Their reality testing may be too tenuous and they may overattribute their current experience to past events. It is clearly preferable to work with such patients uniquely in the here and now if gross distortions of history are to be avoided.

(i) Transference interpretations

Strachey (1934) placed transference interpretation at the center of psychodynamic work. As discussed above, his rationale for this was constructed as a route out of a vicious cycle of the external confirmation of the patient's repudiation of certain fantasies and feelings. Others, however, saw different strengths in "working in the transference." A focus on the therapeutic relationship offers the patient the opportunity to internalize the thinking function of the therapist (Hoffer, 1950), the relationship with the analyst as a new object (Klauber, 1972), the therapist as a self-object (Kohut, 1977; Kohut & Wolf, 1978), or adopt a pretend stance facilitative of the development of an awareness of other minds (Fonagy, 1991, 1995).

Steiner (1993) pointed to this latter aspect of interpretations when he distinguished "analyst-centered" and "patient-centered" aspects of interpretations. When therapists interpret what they imagine to be going on in the patient's mind they make patient-centered interpretations; when they interpret the patient's reactions in terms of what they imagine the patient is thinking about what the therapist is thinking or feeling then this is an analyst-centered interpretation. Here the patient is directly learning about how minds interact within social relationships (Fonagy, Moran, & Target, 1993). Whereas the former type of interpretations, if used to excess, may appear to be blaming the patient, the latter type make the therapists appear as if their concerns were purely narcissistic and not sympathetic to the patients' real difficulties in the external world.

(ii) Extratransference interpretations

Most psychodynamic clinicians now agree that a balance needs to be struck between transference and extratransference interpretations (Stewart, 1989). O'Shaughnessy (1992) evocatively described how a treatment focused too much within the transference can become "an over-close enclave" and extratransference interpretations had the power to break the claustrophobic atmosphere. Sometimes, the spontaneous and direct communication of the analyst's experience of frustration ("objective hate") may help the patient see the therapist as a real person and break a repetitive unproductive pattern in the therapy (Coltart, 1986; Symington, 1983). Kernberg (1995) is probably accurate in saying that patients in the borderline spectrum benefit specifically from well-structured, here-and-now transference interpretations.

6.05.2.8 The Result of Psychodynamic Interventions

6.05.2.8.1 The role of insight

Just as interpretation is paradigmatic but by no means the sole effective component of psychodynamic interventions, so insight may be an oft idealized but, in reality, relatively rare outcome of therapeutic work. Insight is the conscious recognition of the role of unconscious factors (feelings, experiences, fantasies) on current experience and behavior. "True insight" and mere and intellectual knowledge should be carefully distinguished (Zilboorg, 1952). Thomä and Kächele (1987) identify insight as equidistant between the poles of pure intellect and simple emotional experience. Etchegoyen (1991) helpfully distinguished between descriptive or verbal insights on the one hand and ostensive or shown insights on the other. The latter is a more direct form of knowing and pertains to the common experience when one is in emotional contact with an event one has known before.

There is general agreement that insight is an integrative experience (Thomä & Kächele, 1987). Those who follow a Klein–Bion model would describe this as a healing of defensively instituted splits in the patient's mental representations of others and their relationships to them (e.g., Segal, 1962). In more general terms, it may be seen as an instance of a more general predisposition to think in terms of the mental states of one's objects and understand one's relation to them in mental state terms (Fonagy & Target, 1996). This tendency may also be described in terms of a willingness to see the interpersonal world from a "third-person perspective" (Britton, 1989, 1992).

6.05.2.8.2 The role of working through

Neither intellectual nor emotional insight is sufficient for progress (Freud, 1914). The function of working through is to help the patient to practice a newly arrived-at integration. This has two components: first, to unlearn the implications of prior misconceptions, and second to assimilate, and practice working with, new constructions. It is interesting to note that the literature on long-term therapy pays far less attention to this aspect of the outcome of interventions than the process of interpretation which merely is the starting point of the change process. Short-term therapies necessarily pay far more attention to the need to present and re-present psychotherapeutic understanding. The relative effectiveness of short-term interventions may owe much of their potency to the systematic way in which the task of working through is tackled (see below).

6.05.2.9 Ending Treatment

6.05.2.9.1 Indications for ending treatment

Like much of psychoanalytic psychotherapy, the ending of treatment is often idealized. The desirable final outcome is often stated in terms of the process of treatment. Kennedy and Moran (1991), following Anna Freud , helpfully separate the process aims from the outcome aims of psychodynamic treatment. The former is likely to be stated in theoretical terms (e.g., a move from paranoid to depressive anxieties, an increase in the coherence of the patient's narrative, an increased awareness of impulses and fantasies, a manifestation of genuine concern for others and so on). All these are observed in the context of the treatment and are at best loosely coupled with the goals the patient might have for ending the treatment.

The latter are often external changes such as the decline of symptoms, improvement of relationships, decrease of unpleasant affect, an increased capacity for assertiveness, and so on. These external criteria are sadly regarded by many psychodynamic clinicians as superficial as they can be achieved without fulfilling the process aims of the treatment (Grünbaum, 1984). Evidence will have to be gathered which clearly demonstrates that external change associated with process change is more extensive or longer lasting than external changes achieved in isolation.

6.05.2.9.2 The process of ending treatment

There is general agreement that ending treatment is "a process." Different authors identify different processes: a new beginning (Balint, 1949), weaning (Meltzer, 1967), mourning (Klein, 1950), detachment (Etchegoyen, 1991), and maturation (Payne, 1950). An inevitable part of ending is disillusionment with not having achieved the ideal (Pedder, 1988), and the loss of the object who has been the receptacle for projections (Steiner, 1993). As part of this process, symptoms might return, if briefly, and problems already worked through may appear to resurface. Most clinicians agree that other than acknowledging the unconscious issues around ending, no specific technical maneuvers are indicated.

6.05.3 ILLUSTRATION OF LONG-TERM PSYCHODYNAMIC THERAPY: PSYCHOTHERAPY FOR BORDERLINE PERSONALITY DISORDER

6.05.3.1 Theoretical Approaches to Borderline Conditions

Psychodynamic theories of borderline personality disorder (BPD) are discussed extensively in Chapter 14, Volume 1 and only a brief summary of these views will be presented here. Broadly there are two approaches to understanding this relatively rare but troublesome disorder characterized by intense but unstable personal relationships, self-destructiveness, impulsivity, poor social adaptation, self-damaging behavior or suicidality, chronic dysphoria, transient psychotic episodes, and so on. There are those who suggest conflict as the central theme and those who favor an explanation in terms of deficit.

The conflict model is best represented by the Klein–Bion tradition (e.g., Steiner, 1993) as well as by the work of Kernberg (1984). A key feature of these formulations is the unsuccessful integration of good and bad part-objects, and the use of primitive defenses (particularly projective identification and idealization and derogation). It is assumed that children cannot deal with excessive aggression arising from abandonment experiences and are forced to split again and again their internal representations to protect their internal sense of good. By contrast, deficit models ascribe the internal fragmentation of borderline patients to deprivation experiences which leave the ego weak and unable to self-soothe. Consequently such individuals draw on exogenous stimuli such as drugs, binge eating, or self-harm to induce mood states (Adler, 1985). Narcissistic features which accompany borderline states are seen by Kernberg as representing the conflict between the individual's need for an object and the rage felt towards that object. They deal with the conflict by self-absorbedly becoming their own ideal self. For Kohut (1977) and Winnicott (1965) absent, insensitive, or abusive primary figures force children to retreat into themselves and become the missing ideal object. In both cases infantile omnipotence is maintained into adulthood.

A critical feature from the point of view of treatment is the high prevalence of trauma in this group. A large proportion of hospitalized borderline individuals meet diagnostic criteria for post-traumatic stress disorder (PTSD) (Gunderson & Sabo, 1993) and 70–80% have histories of severe physical or sexual abuse (Herman, Perry, & van der Kolk, 1989). The almost ubiquitous presence of trauma may help bring together conflict or deficit theories, particularly since the trauma is frequently one of relatively late childhood or even adolescence. It is conceivable that certain individuals whose early attachment relationships were insecure, and who are consequently somewhat poorer at understanding mental states in others and in themselves, when confronted with an abusing relationship respond to it by selectively obliterating detailed representations of the mental states of other attachment figures (Fonagy et al., 1995). A number of common features of borderline personality disorder could be explained in terms of the defensively deactivated mentalizing capacity of such individuals in the context of attachment relationships (e.g., confusion in interpersonal relationships, apparent callousness towards others, poor capacity to communicate).

6.05.3.2 Treatment Strategies

It is to be expected that two sets of theoretical approaches (conflict and deficit) lead to categorically different treatment recommendations. The former identifies the early interpretation of the negative transference as critical, whilst the proponents of the deficit view stress the importance of the holding environment and empathic responding. There is no shortage of forceful recommendations and warnings of the dire consequences should these not be followed (e.g., Kernberg, 1984; Ryle, 1994), but evidence supporting either position is at the moment sparse. A number of reviews have, however, brought together a generic framework for the psychodynamic treatment of borderline patients based on the assumption that both conflict and deficit models are of relevance and both would probably cause harm if inexperienced therapists were to follow the recommendations without qualification (see Gabbard, 1994; Higgitt & Fonagy, 1992; Waldinger, 1987).

6.05.3.2.1 Limits and boundaries

The therapist needs to recognize the importance of both setting definite limits for patients and the patients' likely failure to keep to traditional boundaries. The setting must be protected from the overwhelming demands which such difficult patients may produce, yet must have the flexibility to contain the patients at times when self-regulatory capacities are no longer available to them. As these remarks suggest, there are major advantages to seeing the more severe cases in institutional (although not necessarily inpatient) rather than private settings.

6.05.3.2.2 Interpretive focus

Psychological space must be created in the mind of the patient for interpretive work which assumes mentalizing (thinking about mental states in self and other). It is essential that moment-to-moment changes in affective states are noted and clarified by the therapist, even if the reason for these rapid changes of affect often remains obscure. Failures of understanding are often reacted to rather dramatically (with complete withdrawal, paranoia, physical violence, self-harming). The interpretation (more properly clarification) of these emotional reactions at their earliest stages may avoid a vicious cycle of ever-increasing anger and ever-decreasing possibility for genuine understanding. Therapists should aim to make "microinterpretations"—simple but frequent verbalizations which address states-of-mind—using words to make room for thinking and feeling without making assumptions that at times of intense affect the patient can understand complex causal relations between mental states. It should be remembered that the destruction of thought might be the strategy closest to hand for many of these patients and thus, unless cautious and nimble, the therapists may well find themselves squeezed out of the therapeutic space, where there is now no longer room for any kind of understanding.

6.05.3.2.3 Counter-transference

Feelings with borderline patients are intense. From one meeting to the next, "therapist the savior" may turn into "therapist the tormentor." At these moments several things are hard to remember. First and most obvious, that there is no truth in either attitude and that all that is certain is that nothing is permanent. Second, that there is some painful element of truth to even the wildest projection and that the therapist is a real cause as well as an imaginary one. Third, that any pretense of being unaffected by the close proximity of such intense emotion creates an atmosphere of unauthenticity, not just between therapist and patient but also within the therapist. Unfortunately, the patients' intolerance of other minds causes them to attack the space which therapists usually reserve for themselves for thinking. Mindless anxiety or a bewildered state of numbness can fill not just the session itself, but times for reflection before and after.

6.05.3.3 Handling Crises

Crises in treating borderlines are inevitable but their timing may well be a surprise. The general advice of those who regularly deal with emergencies across a wide range of contexts is that disaster planning is best when it precedes rather than follows the event. A well-rehearsed contract with a patient may go some way to limit extremes of acting out, but having a plan for dealing with suicidality, self-harm, drunkenness, intrusiveness, and sometimes violence is the key. Knowing what to do will avoid acting out on the part of the therapist which mostly takes the form of aggressive, overpunitive, rejecting actions arising mainly out of damaged self-esteem: "How can you do this to me when I have tried so hard with you?" Good collaborative relationships with colleagues also working with the patient is a *sine qua non* of psychotherapeutic treatment.

The management of crises, however, cannot be mechanistic. Careful scrutiny with the patient of thoughts and feelings that led up to the event must take place even if not at the time and not as part of the emergency procedures. The overarching goal is replacing action with mental work. Despite there being no foolproof way of preventing crises, certain patterns of therapeutic intervention (e.g., excessive passivity, complete nondirectiveness, or other expressions of rigidity and unthinking conduct) will inevitably cause these patients to incubate anxiety and "blow."

6.05.3.3.1 Desirable outcomes

The goals of long-term psychodynamic therapy are ambitious in some contexts but have to be modest with borderline patients. The process goals are largely clustered around the notion of tolerance: tolerance for affect, for fantasy, for a variety of therapeutic interventions (particularly therapist errors). The external goals are also limited. Although anxiety may abate, depression rarely responds significantly. Interpersonal issues may improve but never become normal. The behavioral aspects are most likely to respond to treatment.

6.05.3.4 Outcome Research

Evidence for the efficacy of psychodynamic therapy for borderline individuals is limited. A long-term naturalistic follow-up (Stone, 1993) of 500 patients demonstrated some degree of spontaneous remission in middle life but also high suicide rates and selective response to treatment. Clearly, the *DSM* diagnosis describes several subgroups of patients probably with somewhat different etiologies and significant differences in expected outcomes. Work at Cornell Medical Center by John Clarkin, Otto Kernberg, and colleagues may provide us with a questionnaire instrument which will assist in the assessment of suitability. Treatment drop-out rates are very high, in one sample up to 50% (Aronson, 1989). Patients who stay in treatment appear to do relatively well with success rates of 60–70% reported from open trials (Rosser, Birch, Bond, Denford, & Schachter, 1987; Stevenson & Meares, 1992; Stone, 1993).

6.05.4 FORMS OF BRIEF PSYCHODYNAMIC THERAPY

6.05.4.1 The Historical Roots of Brief Psychodynamic Therapy

Paradoxically, the prototypical long-term psychodynamic therapy, psychoanalysis, shares its roots with brief therapy. The psychoanalytic treatment of many of the pioneers of psychoanalysis would by present standards be considered brief, intensive psychotherapies. The early cathartic method advocated in Breuer and Freud (1895) may be construed as a focused hypnotic brief intervention. Freud's case load, even after his discovery of the method of free association, contained brief therapy cases including some very distinguished clients, Bruno Walter and Gustav Mahler amongst them.

Three of the early generations of psychoanalysts sharing Hungarian origins contributed most directly to brief psychodynamic interventions. These are Otto Rank, Sandor Ferenczi, and Franz Alexander. Both Rank and Ferenczi were concerned that long-term therapy could reinforce regressed overdependent aspects of the client's personality and that the goal of psychoanalytic research through the in-depth exploration of the patient's psyche may at times conflict with the immediate aim of addressing the abnormal dynamics of the patient's mental life (Ferenczi & Rank, 1925, p. 52). Ferenczi was particularly keen to discover ways in which the process of symptomatic cure could be accelerated. In the course of this research he discovered many of the central principles of behavior therapy. Because of the conservative nature of the psychoanalytic establishment, these remained outside the psychodynamic frame of reference and became a major challenge rather than part of its armamentarium. There is no difficulty in incorporating these methods into modern psychoanalysis (Fonagy, 1989; Wachtel, 1977), just as Ferenczi had little difficulty in so doing. His discoveries included the principle of exposure to combat phobic avoidance, the principle of response prevention for obsessional rituals, and the focused elaboration of key ideas akin to some strategies of cognitive therapy. Ferenczi 's aim in these and other elaborations was to accelerate the process of change. Neither he nor Otto Rank shied away from the possibility of using techniques from other therapeutic modalities (e.g., hypnosis) if these were going to advance their underlying aim of enhancing the curative emotional experience of psychodynamic therapy (Ferenczi & Rank, 1925, pp. 63–64).

Franz Alexander was influenced by Ferenczi in his training. Together with Thomas French, he elaborated Ferenczi's suggestion of provoking specific affective experiences in the therapeutic relationship. They advocated that the therapist should purposefully counteract the pathogenic influence of particular significant figures from the past. For example, the accepting attitude of the therapist may contrast with that of an excessively harsh and authoritarian parental figure. The choice of the therapist's attitude should be dictated by the history of the specific patient. Even more important than these controversial suggestions is the general framework advanced by Alexander and French which included a structured approach initiated by a detailed assessment followed by a comprehensive formulation, the setting of treatment goals, and the systematic anticipation of problems that may be encountered in the course of treatment. The total treatment package included homework assignments, a focus on current relationships, and an open acceptance of educational as well as insight-related goals of therapy.

6.05.4.2 Indications for Brief Psychodynamic Therapy

The diagnostic groups that are likely to be considered for brief psychodynamic therapy are the less severe anxiety and depressive disorders, adjustment disorders, and some of the milder personality disorders. By contrast, individuals with a history of suicidal threats, alcohol and substance abuse, poor impulse control, incapacitating depression or anxiety, or dramatic cluster personality disorders are normally deemed unsuitable (Messer & Warren, 1995). Proponents of brief therapy normally add psychodynamic criteria to the phenomenological. Sifneos (1987) for example, suggested that, in addition to a circumscribed central complaint, suitable clients would have a history of at least one good childhood relationship, a capacity to relate flexibly to the interviewer, evident psychological mindedness and a motivation for change beyond symptom relief. By contrast, Malan (1976a) lists six dynamic exclusion criteria which are the mirror image of Sifneos' selection criteria: for example, the inability to make contact, lack of motivation for treatment, rigid defenses and severe dependence. Davanloo (1980) emphasizes the importance of "trial interpretations." If clients respond to firmly put, but necessarily tentative hypotheses by a "deepening involvement," they are more likely to be regarded as suitable than individuals whose response is decompensatory (e.g., anxiety, confusion, paranoia).

Strupp and Binder (1984) also list current emotional discomfort. They suggest that patients have to be sufficiently uncomfortable with their feelings and/or behavior to seek help from psychotherapy. On the whole, brief psychotherapists recommend expressive techniques for healthier patients and supportive techniques for sicker ones (Luborsky & Mark, 1991).

No formal, structured interviews and assessments have been developed to aid clinicians in these assessments. There is, however, significant empirical data available which suggests that some of these psychodynamic parameters are pertinent to the likely success of psychotherapy. Piper and his colleagues have, however, carried out a number of excellent studies validating some of the underlying concepts. Piper, de Carufel, and Szkrumelak (1985) demonstrated that the quality of object relations (QOR) and defensive style together predicted good outcome in time-limited psychodynamic therapy. The clinical judgment of the quality of object relations is based on an evaluation of the quality of object relation patterns throughout the life span, the capacity to regulate affect and self-esteem, and historical antecedents of these (Piper et al., 1985).

Quasi-experimental studies demonstrated that clients with high QOR were more likely to benefit from therapy than low QOR ones to a clinically significant extent (Piper, Azim, McCallum, & Joyce, 1990a). QOR may be a better predictor of therapeutic alliance than measures of interpersonal functioning but may account for the latter's association with outcome (Piper et al., 1990). Norwegian studies have independently demonstrated that a clinical assessment of the patient's quality of interpersonal relations was a good predictor of long-term change following dynamic psychotherapy (Høglend, 1993a, 1993b; Høglend, Sørlie, Heyerdahl, Sørbye, & Amlo, 1993b).

6.05.4.3 Techniques of Brief Psychodynamic Therapy

6.05.4.3.1 Malan's brief intensive psychotherapy (BIP)

David Malan, a British psychoanalyst working at the Tavistock Clinic alongside one of the pioneers of object relations theory, Michael Balint, was one of the first to adapt standard psychodynamic therapy as practiced in this psychoanalytically oriented outpatient public mental health facility for brief interventions (Malan, 1963, 1976a, 1976b). The normal length of BIP is approximately 20 sessions. In the initial session(s) central conflicts for the patient are identified and the therapist focuses on these selectively, ignoring other conflicts and interpreting only those aspects of the patient's material which pertain to these concerns. There is no particular type of conflict to which Malan's approach gives preference, although Malan offers an overriding structure, similar to that of Karl Menninger, which he regards as pertinent to all conflicts. The "triangle of conflict" includes: the impulse or affect, the defense erected against it, and the symptom or anxiety which ensues after the failure of defence. An example (Malan, 1980, pp. 178–184) of a focal conflict may be someone who is angry about being imposed upon but defends against this and manifests usually intense anxiety about asserting herself and becomes depressed as a consequence. Malan recommends addressing conflict in at least three contexts: in the patient's current life, with the therapist, and in relation to past caregiving figures.

Perhaps because of its proximity to standard psychotherapeutic practice, Malan's Brief Intensive Therapy has been extensively validated empirically. Malan's own studies (Malan, 1976a, 1976b) are methodologically too weak to warrant review. However, the Canadian studies by Piper have provided evidence that (i) Malan's therapy is as effective as long-term individual or group psychotherapy (Piper, Debbane, Bienvenu, & Garant, 1984), (ii) patients undergoing this form of therapy are significantly better off in 78% of cases than untreated control patients (Piper et al. 1990), and (iii) that for highly object-related patients the accuracy of transference interpretations corresponded to outcome at six-month follow-up (Piper, Joyce, McCallum, & Azim, 1993). Norwegian studies demonstrated that level of insight gained (as assessed independently) in 43 outpatients correlated with overall dynamic change at four-year follow-up (Høglend, Engelstad, Sørbye, Heyerdahl, & Amlo, 1994).

There are controversial findings concerning the role of transference interpretations. Malan reported that the frequency of therapist, parent, current figures triangular interpretations correlated with therapeutic outcome (Malan, 1963, 1976b). This finding corresponds to Strachey 's classical assumptions concerning "mutative" aspects of interpretation (Strachey, 1934). Unfortunately, although preliminary replications confirmed Malan's observations (Marziali, 1984; Silberschatz, Fretter, & Curtis, 1986), more careful analyses by the Canadian group indicated that, for low QOR patients at least, the frequency of transference interpretation was associated with less rather than more symptom change (Høglend et al., 1993a; Piper, Azim, Joyce, & McCallum, 1991).

An interesting study by another Norwegian research group has demonstrated the effectiveness of Malan's BIP in reducing the relapse rate following clomipramine treatment of panic disorder (Wiborg & Dahl, 1996). Thus Malan's therapy is relatively well validated although the clinical groups on which the treatment was assessed are relatively heterogeneous. Studies also offer some indication that the treatment process corresponds to those hypothesized by the originator of the therapy, although evidence on this point remains equivocal.

6.05.4.3.2 Sifneos' short-term anxiety-provoking psychotherapy (STAPT)

In the USA, short-term psychodynamic psychotherapy retained close links with the classical ego psychology tradition. Sifneos developed a psychodynamic treatment focused on the oedipal concern of individuals whose psychological problems could be relatively readily linked to this common type of unconscious conflict (Sifneos, 1979, 1987, 1992). Sifneos, probably accurately, pointed out that psychodynamic therapists were frequently more ready to acknowledge issues of dependency and frustration in relation to caregiving figures than concerns about childhood sexual fantasies about parents of the opposite gender. The recommended strategy is for the therapist to listen carefully for material pertaining to oedipal issues and to address this directly without regard to the defenses which individuals may have erected to protect themselves from the anxieties these thoughts might engender.

Sifneos adopts a somewhat didactic stance and does not shrink from explaining his reasoning in identifying material as relating to oedipal anxieties. He also confronts patients' defenses, being particularly sensitive to instances of intellectualization. A strong point of his approach is the availability of a relatively comprehensive manual for short-term anxiety-provoking psychotherapy (Sifneos, 1992).

There is only very limited evidence available to support the usefulness of this approach. Sifneos and colleagues reported a comparison of 22 treated patients and eight waiting list controls (Sifneos, Apfel, Bassuk, Fishman, & Gill, 1980). While 18 out of 22 were reported to have recovered in the treated group or to be much better, none of the waiting list group reported a similar degree of change. In a somewhat larger study reported in 1987, 30 of 36 patients were rated as having recovered or being much improved whereas 80% of the 14 waiting list patients were unchanged. Unfortunately in neither study were the measures sufficiently clearly described to permit generalization, nor were raters blind as to treatment group. Independent studies examining the relationship of therapeutic process to outcome found no evidence that therapists' competence at practicing STAPT predicted good outcome. In fact, competence was inversely related to improvement (Svartberg & Stiles, 1992, 1994).

6.05.4.3.3 Davanloo's intensive short-term dynamic psychotherapy (ISTDP)

Davanloo's approach is also quite confrontational, aiming to create a degree of emotional arousal and even discomfort while trying to address presumed feelings in clients which they might have consistently avoided (Davanloo, 1978, 1980). Davanloo's aim is to intensify the emotional charge of the therapeutic situation so that within it important past emotionally charged experiences will once again come to life. The therapeutic strategy entails offering empathic support in relation to the past hardships suffered by the patient. Nevertheless, the pattern of interaction as revealed by Davanloo's detailed account is one of tenacious and unremitting confrontation of the patient, focused on the patient's problems and defenses against them, identified on an initial inquiry.

There is a single trial which supports the effectiveness of this form of therapy (Winston et al., 1991). Thirty-two patients were assigned to one of two brief psychodynamic therapies. Patients assigned to the other therapy (brief adaptive psychotherapy) did comparably to those receiving ISTDP and both did substantially better than the waiting list controls. In a follow-up and extension of this investigation (Winston et al., 1991, 1994) a larger sample of patients were seen to maintain their improvement. An interesting feature of this investigation was the inclusion of some nondramatic Axis II patients. The study, however, only used self-report measures and did not succeed in differentiating the two forms of psychotherapy.

6.05.4.3.4 Luborsky's supportive-expressive time-limited therapy (SETLT)

Luborsky's brief psychodynamic approach is fundamentally an adaptation of psychodynamic therapy defined quite broadly (Luborsky, 1984). The technical principles are outlined by Luborsky and Mark (1991). They include recommendations for the therapist to be sensitive to allow the patient to form a "helping alliance," to identify and respond about central relationship patterns, and identify where the client's symptoms fit into these. They explicitly recognize the patient's needs to test the relationship in transference terms and recommend that the patient's symptoms should be identified as problem-solving coping patterns.

Thus, Luborsky's approach has two critical focuses. The first is the relationship patterns which they label as "core conflictual relationship themes" (CCRTs) (Luborsky & Crits-Christoph, 1990). These themes consist of three components: the wish (or need), the anticipated response from others to this wish, and the response from the self to the other's response. These themes are conflictual because the response of the other to the wish is anticipated to be negative. The CCRT is derived from clinical material. "Relationship episodes" are identified from transcripts of sessions where the patient narrates an episode of interaction. The CCRT represents a summary of the most frequent types of components. Luborsky and Schaffler (1990) offer an illustration of an individual whose wish is to be assertive, dominant yet reassured, but anticipates disapproval and is left feeling annoyed, angry, and upset about not feeling in control. It is important to note that relationship episodes may be drawn from narratives which apparently do not involve the self, and in this respect the method is analogous to psychodynamic psychotherapists' approach to the understanding of the transference.

Considerable research supports the value of CCRTs (see Luborsky & Luborsky, 1995). The reliability of both the overall formulation and its components is relatively high when performed by trained judges (Crits-Christoph, Luborsky, Popp, Mellon, & Mark, 1990). The measure is also replicable across settings. The CCRT with the therapist parallels the CCRT for others in the patient's life (Fried, Crits-Christoph, & Luborsky, 1990) and even the CCRT derived from patients' dreams (Popp, Luborsky, & Crits-Christoph, 1990). The extent of these observed congruencies indicates that these patterns of representations of relationships are stable structures of the personality. It is an important validation of Luborsky's approach that the frequency of conflict themes negatively correlated with measures of change (Crits-Christoph & Luborsky, 1990).

The second component of Luborsky's approach is the helping alliance. In addition to attending to current, past, and transference relationship fears, therapists are required, through their timing and restraint in responding interpretively to clients' material, to convey respect, acceptance, realistic optimism, to encourage self-expression and thus create a collaborative atmosphere (Luborsky, 1984, pp. 81–89). The helping alliance thus generated has two dimensions according to Luborsky: the patient perceives the therapist as a provider of help that is needed, and the patient perceives the therapy as a collaborative exercise (Luborsky, Crits-Christoph, Mintz, & Auerbach, 1988).

Helping alliance has been shown to be positively associated with outcome. The duality of Luborsky's approach (CCRT and helping alliance) is supported by critical findings that the quality of the helping alliance and therapist accuracy (as defined in terms of the degree of correspondence with the CCRT) independently predict outcome (Crits-Christoph, Barber, & Kurcias, 1993; Crits-Christoph, Cooper, & Luborsky, 1988). In the more recent study, there was an indication that accuracy of early intervention had a positive impact later on in therapy. Thus it is possible that whilst the quality of the helping alliance may be the immediate cause of change, this aspect of the relationship may be enhanced by accurate formulations by the therapist.

Although no randomized controlled trials compare Luborsky's supportive-expressive therapy with no treatment controls for milder neurotic disorders, the Penn project yielded numerous important findings concerning the outcome of this treatment. Most impressively, three-quarters of the large group of patients treated showed moderate or much improvement with an average effect size of over one standard deviation, and much of this gain was maintained at one-year follow-up (Luborsky et al., 1988). In addition, in a major study comparing supportive-expressive therapy, cognitive therapy, and drug counseling for opiate-dependent patients, Woody and his colleagues (Woody, Luborsky, McLellan, & O'Brien, 1990; Woody et al., 1983; Woody, McLellan, Luborsky, & O'Brien, 1987) found supportive-expressive therapy to be effective in reducing psychiatric symptoms, opiate-positive urine specimens, employment, and legal problems. In a replication study the same group demonstrated that supportive-expressive therapy was a useful adjunct to drug counseling as part of a typical community-based drug program (Woody, McLellan, Luborsky, & O'Brien, 1995). It is notable that supportive-expressive therapy was particularly helpful in maintaining improvement at follow-up. Although studies on depressed or anxious individuals would be helpful, it is clear that supportive-expressive therapy is a valuable intervention strategy.

6.05.4.3.5 Strupp's time-limited dynamic therapy (TLDP)

The hallmark of Strupp's model of therapy lies in its interpersonal emphasis and the persistent use of the transference relationship in the here and now. The therapy has its intellectual roots in Sullivan's (1953) interpersonal psychoanalytic tradition. Binder and Strupp (1991) understand psychopathology as arising out of cycling maladaptive patterns whereby patients perceive themselves in maladaptive roles with self-defeating expectations, negative self-appraisals and adverse affects consequent on these. The cyclical nature of the process arises out of the client's unconscious tendency to induce others to behave in ways that reinforce the patient's negative and painful expectations, thus further reinforcing these expectations and the interpersonal behaviors which arise from these. This formulation is close to that proposed by Joseph Sandler in the 1970s and early 1980s (Sandler, 1976b, 1990, 1992).

Similarly to Luborsky, Binder and Strupp (1991) distinguish actions of the self, expectations about others' actions, acts of others towards the self, and acts of the self towards the self. Although these terms sound somewhat behavioral, from their description it is clear that the authors are concerned with mental representation of these interpersonal behaviors rather than the behaviors themselves. The cyclical maladaptive patterns are identified by the therapist on the basis of the client's characteristic patterns of relating which involve the patient's perception of self, others, and their interactions. In therapy, maladaptive patterns are identified, their meaning interpreted and the client is helped to articulate and modify entrenched and limiting views of the self, others, and their interaction.

TLDP creates individualized theories for each client, making use of these four headings. The theories connect together behavioral and experiential phenomena which otherwise would appear discontinuous (Strupp & Binder, 1984). The patient's relationship with the therapist is seen as a key component of the change process. This new relationship may disconfirm maladaptive expectations of others and provides a chance to examine the way that patients' acts towards the self may ensnare them in maladaptive patterns (Binder & Strupp, 1991, p. 142). The Vanderbilt University Project, of which Strupp's TLDP is a product, was key in identifying the therapeutic relationship and the nature of the patient's experience with this new figure as a key aspect of therapeutic change. An important finding of the project was that successful treatments could be predicted on the basis of the patient feeling accepted, understood, and liked by the therapist as early as the third session of treatment (Hartley & Strupp, 1983; O'Malley, Suh, & Strupp, 1983; Windholz & Silberschatz, 1988). The importance of the relationship is further underscored by the absence of an observed significant difference between trained psychotherapists and sensitive but untrained college professors in their ability to administer time-limited dynamic therapy (Strupp & Hadley, 1979).

6.05.4.3.6 Weiss and Sampson's plan formulation method

Weiss and Sampson (1986) proposed a further useful framework for the articulation of clinical focus. At the core of the theory is the assumption that patients enter therapy with an unconscious plan about how they may overcome their problems with therapeutic help. Pathogenic beliefs are "obstacles" to this. Weiss and Sampson assume that patients will inevitably "test," in the context of their therapy, whether their pathogenic beliefs are true. If the therapist "passes" the test, the patient's experience will contribute to the enlargement of understanding ("insight") to counteract the pathogenic belief. An example of a pathogenic belief may be an individual with a mother unhappily married throughout the patient's childhood, unconsciously believing that her happiness means that her mother will feel hurt and abandoned. The patient might then test the therapist to see if the therapist also expects the patient to be self-sacrificing and feigns hurt if the patient attends "selfishly" to her own needs. Such an unconscious belief arises out of actual historical experience, rather than an unconscious wish to harm the mother (Sampson, 1992, p. 515).

It follows that the principal therapeutic task of short-term dynamic psychotherapy is that the therapist should recognize the patients' conscious or unconscious attempts to replicate with the therapist their pathogenic situations or past life experiences. Therapists may be deemed to pass the patient's test by giving recognition of the situation through appropriate interpretation or nonverbal behavior, such as maintaining their therapeutic stance despite the patient's determination to traumatize and unsettle them.

Thus the nonoccurrence of certain strongly anticipated outcomes may in and of itself be therapeutic.

Research, building on this approach, has demonstrated that plan formulation may be reliably assessed using recorded clinical material (Curtis, Silberschatz, Sampson, & Weiss, 1994; Curtis, Silberschatz, Sampson, Weiss, & Rosenberg, 1988; Silberschatz & Curtis, 1993) and that such plans are stable over time (Collins & Messer, 1991). The compatibility of the therapist's interventions with independently assessed plans of the patient was found to predict progress in the early and middle phase of therapy (Messer, Tishby, & Spillman, 1992). One of the strengths of Weiss and Sampson's approach is that it may be assumed to apply to patient's behavior, regardless of the orientation of the therapist. Indeed, the Plan Compatibility of Interventions Scale predicts response to therapist intervention in cognitive-dynamic, as well as object-relations theory-based approaches (Tishby & Messer, 1995). It should be noted that in these studies outcome refers to relative improvements within the therapeutic process rather than overall improvements at the end of therapy. In fact, no large-scale study of Weiss and Sampson's method has as yet been undertaken.

6.05.4.3.7 Horowitz's person schema theory

Horowitz (1988a, 1988b, 1991a) has offered a general systems theory reformulation of object relations constructs strongly influenced by Bowlby's (1973, 1980) notion of internal working models, Sandler's (1976a, 1976b) notion of role responsiveness, and Kernberg's (1975, 1984) model of self–object dyadic units, as well as current cognitive science. He proposes that through development the individual evolves multiple schemas of self and other which exist either as person schemas or as role-relationship models (RRMs). He defined the former as nonexperiential (codifications and) meaning structures with the potential to influence the formation of the self-concept. They are seen as combining in more complex schemata of the self-in-relationship-with-the-other (see also Stern, 1994). The self-schemas integrate the individual's prior experiences and, ideally, present a stable image of the self as invulnerable. These RRMs are templates of relationships which can affect the formation of the concept of relationships as well as actual patterns of interpersonal transactions. RRMs are assumed to specify interaction patterns as sequences, much like scripts of plays, but in terms of expectations, wishes, and appraisals of one person toward the other.

RRMs are organized into affectively coherent configurations (RRMCs) each of which are made up of RRMs with a set of wishes, fears, and defenses in relation to a specific theme. Desired RRMCs contain strong wishes, and dreaded RRMCs are made up of feared RRMs. The derivatives of defensive operations are compromise RRMs which can be either adaptive (if successful) or problematic (if not), but in either case the affective valence of the RRM has been attenuated. A problematic RRMC will contain either negative affects or maladaptive traits, at a more manageable level than in dreaded RRMs. If the enactment of a desired RRM is blocked by the threat of entry of a dreaded RRM, an attenuated solution to the desired RRM is found which provides a partial gratification of the wish. An RRMC may represent a firm linkage between RRMs so that a mental state organized by a desired RRM can trigger a mental state organized by a dreaded RRM. Such mental states may be represented by patterns of activation, such as those envisaged in parallel distributed computer models of neural activity.

Horowitz (Horowitz, 1991a; Horowitz, Fridhandler, & Stinson, 1991) views anxiety as a mismatch between schemas and incoming information. If information is interpreted as suggesting a dreaded schema, anxiety will result. If a wished-for RRM is to be in relation with a powerful guiding figure but this brings with it a dreaded RRM of a state of exploitation, the individual will experience anxiety at the moment of being approached by an actually benign but powerful figure because this person's presence brings with it the threat of exploitation. This may put control processes in place which will reduce the distance of this figure. The anticipation of the dreaded RRM is experienced as anxiety without the dreaded RRM ever being fully activated (i.e., coming into awareness). The control processes can, in the extreme, become severe enough to imply personality disorder, in this case perhaps a somewhat schizoid state.

In certain cases the dreaded RRM may be partly experienced and this is also expected to lead to anxiety. A woman who lost her husband, whom she felt was dependable but whom she did not love, may develop anxiety when she starts a relationship with another man for whom she has more intense feelings. The dreaded RRM which is briefly activated is the experience of seeing herself as the unfaithful wife humiliating her dependable but unexciting husband (see Horowitz, 1991a).

In post-traumatic stress disorder an experience is vividly encoded in memory. Because it is not integrated into the individual's prior integrated self-schema, it is liable to be activated as incoming information, and misinterpreted to imply the reoccurrence of the trauma. The trauma may also threaten to actualize a dreaded RRM, for example, of the self as weak and overwhelmed. The compromise state may be denial, depersonalization, restricted affect, and hypervigilance (see Horowitz, 1986, 1988a). In generalized anxiety disorder the dreaded RRM is seen as inescapable, either because a compromise cannot be reached or because the desired RRM contains some dreaded components (e.g., RRM of the self as blundering and stupid in face of punishing mentors in a situation which is experienced as one of constant evaluation).

Horowitz's model is most extensively elaborated for the 12-session treatment of post-traumatic stress disorder (Horowitz, 1986, 1991b). The therapy is aimed at the realignment of RRMs. In the "intrusive repetitive phase of the stress response" the recommended strategy is largely supportive and ameliorative. The therapist takes on the auxiliary ego function of self-regulation and the reduction of overwhelming affect states. In the "denial numbing phase" of the stress response, Horowitz emphasizes the reduction of controls over self-expression and emotional exploration. The stress event remains the therapeutic focus, serving to organize the therapist's activities in relation to the transference and other feared topics. Time limit and termination are also explicitly identified as themes to be addressed from as early as session six.

Horowitz's model stands out among psychodynamic formulations for brief therapy in offering a comprehensive framework specific for a range of psychiatric disorders at the same time as remaining amenable to empirical examination. Clinical judges appear to be able reliably to assess roles, characteristics, and traits of self and others, link them in wish-fear dilemmas, and assemble them into presumed RRMCs of particular patients observed in psychotherapy sessions (Horowitz, Milbrath, Reidbord, & Stinson, 1993; Horowitz, 1995; Horowitz & Eells, 1993). There appears to be considerable convergence between RRMC formulations and Luborsky's core conflictual relationship theme approach (Horowitz, Luborsky, & Popp, 1991). Unfortunately, randomly controlled outcome studies on Horowitz's therapeutic strategies are lacking.

6.05.4.3.8 *Mann's time-limited psychodynamic therapy (TLPT)*

Of the brief psychodynamic therapies it is perhaps Mann's approach that takes the time-limited nature of the intervention most seriously. He believes that the time limitations inevitably brings into the foreground the difficulty most of us experience with separation (Mann, 1973). Time-limited psychotherapy is focused on the patient overcoming separation–individuation issues through the mastery of separation anxiety. The conflicts encountered in this context concern independence, activity, self-esteem, and delayed grief. Mann suggests that the beginning of therapy recreates a symbiotic unity, the middle phase recreates ambivalence, whilst the end phase introduces the necessity to give up the object, but this time without hatred, anger, despair, or guilt.

Mann's formulation focuses on clients' feelings about themselves in relation to the painful events described. He acknowledges the patients' active coping efforts and their experience of failure with regard to being able to adapt adequately to this distress. Patients' symptoms are not directly addressed but rather the underlying emotional state, particularly their injured self-esteem, is highlighted. In contrast to Sifneos and Davanloo, Mann is not at all confrontative and uses confirming and mirroring to bypass defenses rather than tackling them head on (Mann & Goldman, 1982). The therapy lasts 12 sessions excluding sessions for history taking. The initial phase of four sessions sets up an alliance generating hope and frequently a remission of symptoms. In the middle phase ambivalence sets in as patients come to recognize that their unconscious expectations will not be fulfilled. In this phase the therapist is no longer simply mirroring and affirming but is introducing clarifications, mild confrontations, and interpretations especially regarding the current situation of the person's life and thus directly reinforcing the patient's separateness. In the final four sessions the patient's reactions to termination are the focus. The therapist attempts to build a sense of mastery and competence at the same time as addressing the patient's disappointment and ambivalence concerning separation, often linked to past experiences of inadequate resolutions of separation and loss.

In this therapy the time-limited nature of the intervention is used directly to elicit a set of conflicts associated with earlier separations and losses where maladaptive emotions were generated (anger, disappointment, sadness, guilt) which ultimately resulted in the disorder which led to the referral. The therapy creates a situation where separation can occur with a degree of resolution which is less contaminated by negative emotions. The internalizations thus lead to a less angry, more benign introject.

Although Mann's therapy is well-established, it has not yet generated a great deal of research.

A notable exception is a study reported by Shefler and colleagues (Shefler, Dasberg, & Ben-Shakhar, 1995). This was a randomized controlled trial with a waiting list control design. Only nine out of the 33 patients did not have a *DSM* diagnosis. The effect size was only fractionally below one on a range of measures. However, the patients accepted for the trial were highly selected and only 11% of those seen were included in the trial. In an uncontrolled trial, Joyce and Piper (1990) found that 14 patients diagnosed with separation-individuation problems were highly successfully treated in 12 sessions using Mann's technique and the results were maintained at six-months follow-up. There is further evidence that TLPT reduces drop-out rate (Sledge, Moras, Hartley, & Levine, 1990), and patients who do well show a trend towards increasingly appreciating their sessions as the treatment progresses (Joyce & Piper, 1990).

Mann's approach has been criticized for attempting to provide a generic model on the basis of a specific model of pathogenesis (Grand, Rechetnick, Podrug, & Schwager, 1985; Westen, 1986). Some fundamentally disagree that termination is an inevitable crisis in therapy (Quintana, 1993). It is also unlikely that Mann's model of cure is accurate. A process such as internalization is more likely in long-term than in short-term therapy (Westen, 1986). However, although the treatment is only appropriate for a relatively limited group of patients, its very specificity may be its strongest feature. If a method was found to identify this group using reliable operational criteria which also matched concerns of purchasers of mental health care (e.g., generalized anxiety disorder), this form of therapy could be a valuable component of the repertoire of psychodynamic approaches.

6.05.4.3.9 Ryle's cognitive analytic therapy (CAT)

Relatively unknown in the USA, but increasingly influential in the UK, is CAT, a time-limited integrative psychotherapy (Ryle, 1982, 1990). The procedural sequence model (PSM) is the framework used by Ryle to restate psychoanalytic ideas using cognitive language. The model conceptualizes intentional acts as procedures entailing a series of steps including appraisal of plans and predicted consequences, the evaluation of the consequences of the enactment, and the confirmation or revision of aims and means following this evaluation. The therapeutic method is centered on the process of reformulation. Over the course of the first month of the treatment, patients normally engage in monitoring their symptoms, undesirable behaviors, and mood shifts. Neurotic patterns are described in terms of three categories: dilemmas, traps, and snags. Dilemmas, traps, and snags are described in the "psychotherapy file" which is given to patients at the end of the first session. They rate items within it to indicate how characteristic they are of them. These ratings are discussed at subsequent sessions when main target problems (TPs) are also established, on the basis of the self-monitoring, together with the underlying dilemmas, traps, and snags. Traps are things we cannot escape from, such as "a fear of hurting others" trap, "trying to please" trap, or "social isolation" trap. Dilemmas are false choices about oneself or about one's relation to others, for example, "either I feel I spoil myself and am greedy or I deny myself things and punish myself and feel miserable" or "either I am a brute or a martyr." Snags are ways we stop ourselves from changing, for example, "for fear of the response of others." The dilemmas, traps, and snags characteristic of an individual are the target problem procedures (TPPs) which are thought to underlie that person's central problems. "TPs" and "TPPs" form the agenda of the therapy. The remaining sessions (usually once a week over three months) are devoted to recognition of the TPPs using diaries and other self-monitoring devices as well as close monitoring of the client's behavior in the therapeutic situation. Modification of TPPs is principally achieved through behavioral techniques such as role play, as well as enhanced self-reflection. The explicit noncollusive relationship with the therapist is also thought to facilitate the development of new procedures.

Ryle (1985) incorporated object-relations theory into CAT introducing the notion of reciprocal role procedures. These are thought to develop on the basis of early object relationships. It is assumed that a relationship teaches children both the behaviors expected of them and the behaviors they expect of others. Self-management is learned through incorporating into the child's behavioral repertoire the caretaker's behaviors.

The emphasis in CAT is on early and profound deprivation as the cause of primitive defenses such as splitting, which characterizes individuals who fail to integrate their self-structure and elicit confirmations from others for each of their split-off self states. Whereas neurotic clients restrict or distort their procedures, borderline personality disorder patients manifest dissociated self-states containing different procedures in each.

Therapists summarize their assessments and present it to the client in writing after the fourth session. These summaries represent client history and present circumstances and trace how current problematic procedures could be repetitions of early harmful patterns or are solutions to early situations. The accounts conclude with a list of problem procedures identified and attempt to anticipate how these may influence the course of therapy. A further tool used in CAT is the sequential diagramatic reformulation (SDR) which is a flow diagram representation of how TPPs maintain neurotic patterns.

Thus CAT is a genuine integration of cognitive therapy (Beck, 1976) and object-relations theory-oriented psychodynamic therapy (Ogden, 1986). The approach to psychodynamic diagnosis is similar to many considered above, particularly Luborsky and Horowitz. The therapeutic techniques suggested are innovative and share much with schema-oriented cognitive therapy where emotional problems are seen as the reactivation of schemas which have been dormant for many years (Beck & Freeman, 1990; Bricker & Young, 1993; Young, 1990). Ryle's integration also has much in common with other integrative models such as Gold and Wachtel's "cyclical psychodynamics," which also emphasizes self-maintaining vicious cycles and intra- and inter-psychological processes and structured intervention techniques (Gold & Wachtel, 1993). Safran (1990a, 1990b) also links concepts of interpersonal schema and the cognitive interpersonal cycle, and the therapy program outlined has as its target the disconfirmation of dysfunctional interpersonal schemas. CAT, however, is far more coherently integrated with the traditional psychodynamic formulations than any of these alternatives (e.g., 1992; Leiman, 1994b; Ryle, 1994).

The procedural sequence object relations model (PSORM) illustrates the thoughtfulness of this integration. The PSORM identifies procedural patterns which explain the persistence of neurotic behavior. For example, self-destructive acts may be attributed to a dilemma ("as if I must harm myself or harm others") or to a snag ("as if guilty and therefore self-punishing"). Within SDR the self-maintaining nature could be clearly demonstrated through the connection between procedures. For example, the expectation of abandonment may generate a dilemma between being involved and thus risking abandonment and avoiding closeness. Being involved thus necessitates procedures for controlling emotionally significant others by compensatory procedures such as bulimia, that are seen as a substitute for emotional emptiness. The PSORM postulates reciprocal role patterns which constitute a central core and are stated in terms of inner parent–inner child (IP–IC) relations. An example may be a powerfully rejecting inner parent relating to a submissive and needy inner child. The reciprocal nature of the role patterns encompass psychodynamic concepts such as identification, introjection, and projection, internal objects and part-objects. Roles which are experienced as untenable are projected, that is induced in the other, and can be replaced by symptomatic procedures or defensive ones. Procedures acquire their stability from confirming reciprocations which are generally readily elicited from others, thus leaving the central core repertoire unchanged. While Ryle and others writing from a CAT perspective are keen to acknowledge Soviet theoreticians such as Vygotsky, Bakhtin, and Leonjew (e.g., Leiman, 1994a), their views are consistent with psychoanalysts writing in the interpersonalists' tradition (e.g., Mitchell, 1988).

The key difference between CAT and traditional psychodynamic therapy is the shift from interpretive work to description. The CAT therapist describes the state of affairs, often in writing, which is then subject to discussion and modification in direct therapeutic conversations. Ryle repudiates interpretive techniques as potentially regression-inducing, reflecting an unbalanced power relationship between client and therapist and feeding on the omnipotent fantasies of the therapist (Ryle, 1992, 1993). Ryle's approach lends more weight to conscious processes and his technique is based on insight coupled with the activation of self-corrective mechanisms. It is striking that, notwithstanding the emphasis on such mature mental processes, Ryle and his colleagues have reported significant successes in the brief psychotherapeutic treatment of borderline personality disorder. In an ongoing study, BPD patients are offered up to 24 sessions of CAT and follow-up sessions at one, two, three and six months. There is a three-month and one-year follow-up. Initial results are promising (Ryle, 1995). Eight out of 13 patients no longer meet BPD criteria four months after termination but seven were rereferred for a variety of other treatments. Five patients assessed at one year all showed continuing reductions in symptomatology and only one has remained in treatment.

A number of other outcome studies support the usefulness of CAT. A study of 48 outpatients randomly assigned to 12 sessions of CAT or Mann-type brief therapy demonstrated the superiority of CAT on a grid measure of change of construing problems (Brockman, Poynton, Ryle, & Watson, 1987). Unfortunately the measure was neither standardized nor sufficiently independent from the treatment to justify firm conclusions. A study of poorly controlled diabetics randomized 32 patients to intensive education or CAT (Fosbury, 1994). At nine-months follow-up CAT-treated patients had better diabetic control in terms of $HbAl^c$ levels. Other studies were either uncontrolled clinical reports (e.g., Cowmeadow, 1994; Duignan & Mitzman, 1994; Pollock & Kear-Colwell, 1994) or yielded insignificant differences between CAT and control treatments (e.g., the outpatient treatment of anorexia; Ryle, 1995). Thus the empirical basis of CAT cannot yet be considered well-established (although relative to many other psychodynamic treatments its empirical status is highly favorable).

6.05.4.3.10 Hobson's conversational model

An approach which integrates many of the characteristics of the brief psychodynamic approaches considered above is the conversational model outlined by Hobson and his colleagues (Goldberg et al., 1984; Hobson, 1985). The approach combines psychodynamic, interpersonal, and experiential concepts, and emphasizes the therapist–client relationship as the main vehicle for revealing and resolving interpersonal difficulties. In contrast to a number of other short-term therapies, therapists are encouraged to present their views as tentative statements rather than assertions and to make clear that these are open to correction and modification. The therapy is conceived of as an interpersonal negotiation, with therapists inviting elaboration of their ideas by the patient, as well as feedback. The language of the treatment is one of mutuality, with the therapist putting forward hypotheses concerning the client's experiences and possible relationships between these. The therapy has been manualized as part of the Sheffield Psychotherapy Project under the direction of David Shapiro (Shapiro & Firth, 1985).

A unique outcome study by Shapiro and colleagues (Shapiro et al., 1994) assessed the effectiveness of this mode of intervention, contrasting it with cognitive-behavioral therapy in a sample of 117 patients. Both therapies were administered for either eight or 16 weeks. Overall, both therapies were found to be effective and to have comparable results. There was an interaction between initial symptom level and duration of therapy. Patients with severe depression showed significantly better outcomes when they received 16 weeks of therapy. Eighty-eight percent of the sample were followed up one year after the end of treatment (Shapiro et al., 1995). Of the 103 patients, 52% were defined as treatment responders (remained asymptomatic for four months). Of these, 57% maintained their gains, 32% partially maintained their gains, and 11% relapsed. No differences were found between cognitive-behavioral and psychodynamic therapy, although patients who only received eight sessions of psychodynamic therapy did less well than those who received eight or 16 sessions of CBT.

While this particular form of therapy is neither widely known nor widely practiced, it is unique in terms of having been subjected to a rigorous randomized, controlled trial. As the methods used in Hobson's conversational model are fairly generic and consistent with most psychodynamic approaches, the Sheffield psychotherapy trial provides encouraging evidence for the value of brief psychodynamic approaches for the treatment of one of the most common disorders, major depression.

6.05.5 CONCLUSIONS

This chapter has reviewed psychodynamic approaches to adult mental health problems. Psychodynamic therapy is most appropriate for individuals with psychiatric disorder who are relatively well-functioning and have a capacity to understand and respond to interpretive work. The chapter reviewed key clinical concepts in psychodynamic work and demonstrated an increasing concern among psychodynamic clinicians with feelings and ideas provoked by the therapeutic situation itself. Long-term psychodynamic therapy was illustrated in the review of therapeutic strategies with individuals with a borderline personality disorder diagnosis. Several forms of brief psychodynamic therapy were discussed, together with evidence for their effectiveness.

No single approach to such brief treatments emerged as clearly superior to others. While there is a surprising amount of empirical work which has been performed over recent years, few of the studies appear to be conclusive. But, taken together, they underscore the merits of the psychodynamic approach. There is good evidence that psychodynamic therapy is effective with depressed, substance abusing, and some mixed groups of neurotic patients. There are indications of its appropriateness for individuals with PTSD, physical illnesses such as diabetes, and even BPD. Clearly much empirical work remains to be done to identify which of these treatments has the greatest potential value for which patient group.

Work over the last few decades, however, goes a considerable way towards overcoming many of the weaknesses frequently noted in connection with this approach. A number of psychodynamic clinicians have done extensive work on operationalizing their interventions, specifying modifications to generic approaches to make them more appropriate with particular clinical groups, to provide information concerning the way theories of pathology may relate to aspects of technique, to focus and shorten interventions in order to enhance their cost-effectiveness, to provide empirical evidence concerning suitability, and develop specific measures of therapeutic outcome to relate the goals of interventions to the outcomes observed.

Psychodynamic approaches have been around for over 100 years. Considering the mature nature of this category of clinical interventions, the amount of empirical work available in the literature is lamentably small. The underlying trend is for an improvement in this state of affairs. The integration of cognitive and psychodynamic approaches has certainly facilitated this process. Changes to the systems of health care delivery certainly represent a threat to psychodynamic approaches. However, the priority for psychodynamic researchers must be clearly to establish which clinical context is uniquely suited to this approach. There is no doubt that, in the past, psychodynamic clinicians have made exaggerated claims concerning the value of this mode of therapy.

The first decade of the twenty-first century should be a period of readjustment where empirical work could establish the true value of psychodynamic therapy in specific clinical contexts. Purchasers and users of these services have been undoubtedly patient in waiting for such data. Their faith in this approach has probably been fueled by the exceptional intellectual contributions made by many of those who have occupied leadership positions in the psychodynamic field as well as the value of psychodynamic theoretical contributions. Equally there is no doubt that their patience is wearing thin and that compelling empirical data of efficacy is urgently required if this form of therapy is to retain a legitimate place in handbooks such as the present one.

6.05.6 REFERENCES

Abend, S. M. (1989). Countertransference and and psychoanalytic technique. *Psychoanalytic Quarterly, 58,* 374–395.

Adler, G. (1985). *Borderline psychopathology and its treatment.* New York: Aronson.

Arlow, J. A. (1993). Discussion of Baranger's paper on "The mind of the analyst: From listening to interpretation." *International Journal of Psycho-Analysis, 74,* 1147-1154.

Aronson, T. (1989). A critical review of psychotherapeutic treatments of the borderline personality: Historical trends and future directions. *Journal of Nervous and Mental Disease, 177,* 511–528.

Balint, M. (1949). On the termination of analysis. *International Journal of Psycho-Analysis, 31,* 196–199.

Baranger, M. (1993). The mind of the analyst: From listening to interpretation. *International Journal of Psycho-Analysis, 74,* 15–24.

Beck, A. (1976). *Cognitive therapy and the emotional disorders.* New York: International Universities Press/Meriden.

Beck, A. T., & Freeman, A. (1990). *Cognitive therapy of personality disorders.* New York: Guilford

Binder, J. L., & Strupp, H. H. (1991). The Vanderbilt approach to time limited dynamic psychotherapy. In P. Crits-Christoph & J. P. Barber (Eds.), *Handbook of short-term dynamic psychotherapy* (pp. 137–165). New York: Basic Books.

Bion, W. (1955). Language and the schizophrenic. In M. Klein, P. Heimann, & R. Money-Kyrle (Eds.), *New directions of psychoanalysis* (pp. 220–239). London: Tavistock.

Bion, W. R. (1962). *Learning from experience.* London: Heinemann.

Blum, H. (1989). The concept of termination and the evolution of psychoanalytic thought. Annual meeting of the American Psychoanalytic Association (1987, Montreal, Canada). *Journal of the American Psychoanalytic Association, 37,* 275–295.

Bowlby, J. (1973). *Attachment and loss: Vol. 2: Separation: Anxiety and anger.* London: Hogarth Press and Institute of Psycho-Analysis.

Bowlby, J. (1980). *Attachment and loss: Vol. 3: Loss: Sadness and depression.* London: Hogarth Press and Institute of Psycho-Analysis.

Brenman-Pick, I. (1985). Working through in the countertransference. *International Journal of Psycho-Analysis, 66,* 157–166.

Brenner, C. (1982). *The mind in conflict.* New York: International Universities Press.

Breuer, J., & Freud, S. (1895). *Studies on hysteria.* London: Hogarth.

Bricker, D. C., & Young, J. E. (1993). *A client's guide to schema-focussed cognitive therapy.* New York: Cognitive Therapy Center of New York.

Britton, R. (1989). The missing link: Parental sexuality in the Oedipus complex. In J. Steiner (Ed.), *The Oedipus complex today* (pp. 83–102). London: Karnac.

Britton, R. (1992). The Oedipus situation and the depressive position. In R. Anderson (Ed.), *Clinical lectures on Klein and Bion* (pp. 34–45). London: Routledge.

Brockman, B., Poynton, A., Ryle, A., & Watson, J. P. (1987). Effectiveness of time-limited therapy carried out by trainees: Comparison of two methods. *British Journal of Psychiatry, 151,* 602–609.

Broskowski, A. T. (1995). The evolution of health care: Implications for the training and carreer of psychologists. *Professional Psychology: Research and Practice, 26,* 156–162.

Cavell, M. (1994). *The psychoanalytic mind.* Cambridge, MA: Harvard University Press.

Collins, W. D., & Messer, S. B. (1991). Extending the plan formulation method to an object relations perspective: Reliability, stability and adaptability. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 75–81.

Coltart, N. (1986). Slouching towards Bethlehem—or thinking the unthinkable in psychoanalysis. In G. Kohon (Ed.), *The British school of psychoanalysis* (pp. 185–199). London: Free Association Books.

Coltart, N. (1988). Diagnosis and assessment for suitability for psycho-analytic psychotherapy. *British Journal of Psychotherapy, 4,* 127–134.

Cowmeadow, P. (1994). Deliberate self-harm and cognitive analytic therapy. *International Journal of Short-Term Therapy, 9,* 135–150.

Crits-Christoph, P., Barber, J. P., & Kurcias, J. S. (1993). The accuracy of therapists' interpretations and the development of the therapeutic alliance. *Psychotherapy Research, 3,* 25–35.

Crits-Christoph, P., Cooper, A., & Luborsky, L. (1988). The accuracy of therapists' interpretations and the outcome of dynamic psychotherapy. *Journal of Consulting and Clinical Psychology, 56,* 490–495.

Crits-Christoph, C., & Luborsky, L. (1990). Changes in CCRT pervasiveness during psychotherapy. In L. Luborsky & P. Crits-Christoph (Eds.), *Understanding transference: The CCRT method* (pp. 133–146). New York: Basic Books.

Crits-Christoph, P., Luborsky, L., Popp, C., Mellon, J., & Mark, D. (1990). The reliability of choice of narratives and of the CCRT measure. In L. Luborsky & P. Crits-Christoph (Eds.), *Understanding transference: The CCRT method* (pp. 93–101). New York: Basic Books.

Cummings, N. A. (1988). Emergence of the mental health complex: Adaptive and maladaptive responses. *Professional Psychology: Research and Practice, 19,* 308–315.

Curtis, J. T., Silberschatz, G., Sampson, H., & Weiss, J. (1994). The plan formulation method. *Psychotherapy Research, 4,* 197–207.

Curtis, J. T., Silberschatz, G., Sampson, H., Weiss, J., & Rosenberg, S. E. (1988). Developing reliable psychodynamic case formulation: An illustration of the plan diagnosis method. *Psychotherapy, 25,* 256–265.

Davanloo, H. (Ed.) (1978). *Basic principles and techniques in short-term dynamic psychotherapy.* New York: Spectrum.

Davanloo, H. (Ed.) (1980). *Short-term dynamic psychotherapy.* New York: Aronson.

DeLeon, P. H., VandenBos, G. R., & Bulatao, E. Q. (1991). Managed mental health care: A history of the federal policy initiative. *Professional Psychology: Research and Practice, 22,* 15–25.

Duignan, I., & Mitzman, S. F. (1994). Measuring individual change in patients receiving time-limited cognitive analytic therapy. *International Journal of Time-Limited Psychotherapy, 9,* 151–160.

Etchegoyen, H. (1991). *The fundamentals of psychoanalytic technique.* London: Karnac.

Feldman, M. (1993). The dynamics of reassurance. *International Journal of Psycho-Analysis, 74,* 275–285.

Ferenczi, S., & Rank, O. (1925/1986). *The development of psychoanalysis.* Madison, CT: International Universities Press.

Fliess, R. (1953). Countertransference and counter-identification. *Journal of the American Psychoanalytic Association, 1,* 268–284.

Fonagy, P. (1989). On the integration of psychoanalysis and cognitive behaviour theory with psychoanalysis. *British Journal of Psychotherapy, 1015,* 557–563.

Fonagy, P. (1991). Thinking about thinking: Some clinical and theoretical considerations in the treatment of a borderline patient. *International Journal of Psycho-Analysis, 72,* 1–18.

Fonagy, P. (1995). Playing with reality: The development of psychic reality and its malfunction (in borderline patients). *International Journal of Psycho-Analysis, 76,* 39–44.

Fonagy, P., Leigh, T., Steele, M., Steele, H., Kennedy, R., Mattoon, G., & Target, M. (1996). The relationship of attachment status, psychiatric classification, and response to psychotherapy. *Journal of Consulting and Clinical Psychology, 64,* 22–31.

Fonagy, P., Moran, G. S., & Target, M. (1993). Aggression and the psychological self. *International Journal of Psycho-Analysis, 74,* 471–485.

Fonagy, P., Steele, M., Steele, H., Leigh, T., Kennedy, R., Mattoon, G., & Target, M. (1995). Attachment, the reflective self, and borderline states: The predictive specificity of the Adult Attachment Interview and pathological emotional development. In S. Goldberg, R. Muir, & J. Kerr (Eds.), *Attachment theory: Social, developmental and clinical perspectives* (pp. 233–278). New York: Analytic Press.

Fonagy, P., & Target, M. (1996). Playing with reality: I. Theory of mind and the normal develpment of psychic reality. *International Journal of Psycho-Analysis, 77,* 217–233

Fosbury, J. A. (1994, 27 September—1 October). *Cognitive analytic therapy with poorly controlled Type I diabetic patients.* Paper presented at the European Association for the study of diabetes, Dusselforf, Germany.

Freud, A. (1936/1946). *The ego and the mechanisms of defense.* New York: International Universities Press.

Freud, S. (1912). Recommendations to physicians practising psychoanalysis. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 12, pp. 109–120). London: Hogarth.

Freud, S. (1914). Remembering, repeating, and working through. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 12, pp. 145–156). London: Hogarth.

Freud, S. (1915). Observations on transference love. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 12, pp. 157–171). London: Hogarth.

Freud, S. (1923). The ego and the id. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 19, pp. 1–59). London: Hogarth.

Fried, D., Crits-Christoph, P., & Luborsky, L. (1990). The parallel of the CCRT for the therapist with the CCRT for other people. In L. Luborsky & P. Crits-Christoph (Eds.), *Understanding transference: The CCRT method* (pp. 147–157). New York: Basic Books.

Gabbard, G. O. (Ed.). (1994). *Psychodynamic psychiatry in clinical practice: The DSM-IV edition.* Washington, DC: American Psychiatric Press.

Gill, M. M. (1982). *Analysis of transference.* New York: International Universities Press.

Gold, J. R., & Wachtel, P. L. (1993). Cyclical psychodynamics. In J. Stricker & J. R. Gold (Eds.), *Comprehensive handbook of psychotherapy integration* (pp. 59–72). New York: Plenum.

Goldberg, D. P., Hobson, R. F., Maguire, G. P., Margison, F. R., O'Dowd, T., Osborn, M. S., & Moss, S. (1984). The clarification and assessment of a method of psychotherapy. *British Journal of Psychiatry, 114,* 567–575.

Grand, S., Rechetnick, J., Podrug, D., & Schwager, E. (1985). *Transference in brief psychotherapy.* Hillsdale, NJ: Analytic Press.

Greer, S., Morris, T., & Pettingale, K. W. (1979). Psychological response to breast cancer: Effect on outcome. *Lancet, II,* 785–787.

Grinberg, L. (1962). On a specific aspect of countertransference due to the patient's projective identification. *International Journal of Psycho-Analysis, 43,* 436–440.

Grünbaum, A. (1984). *The foundations of psychoanalysis: A philosophical critique.* Berkeley, CA: University of California Press.

Gunderson, J., & Sabo, A. (1993). The phenomenal and conceptual interface between borderline personality disorder and PTSD. *American Journal of Psychiatry, 150,* 19–27.

Hartley, D. E., & Strupp, H. H. (1983). The therapeutic alliance: Its relationship to outcome in brief psychotherapy. In J. Masling (Ed.), *Empirical studies of psychoanalytical theories* (Vol. 1, pp. 1–37). Hillsdale, NJ: Analytic Press.

Heimann, P. (1950). On countertransference. *International Journal of Psycho-Analysis, 31,* 81–84.

Herman, J. L., Perry, C., & van der Kolk, B. A. (1989).

Childhood trauma in borderline personality disorder. *American Journal of Psychiatry, 146,* 490–495.

Higgitt, A., & Fonagy, P. (1992). The psychotherapeutic treatment of borderline and narcissistic personality disorder. *British Journal of Psychiatry, 161,* 23–43.

Hobson, R. F. (1985). *Forms of feeling: The heart of psychotherapy*. New York: Basic Books.

Hoffer, W. (1950). Three psychological criteria for the termination of treatment. *International Journal of Psycho-Analysis, 31,* 194–195.

Høglend, P. (1993a). Personality disorders and long-term outcome after brief psychodynamic psychotherapy. *Journal of Personality Disorders, 7,* 168–181.

Høglend, P. (1993b). Suitability for brief dynamic psychotherapy: Psychodynamic variables as predictors of outcome. *Acta Psychiatrica Scandinavica, 88,* 104–110.

Høglend, P., Engelstad, V., Sørbye, Ø., Heyerdahl, O., & Amlo, S. (1994). The role of insight in exploratory psychodynamic therapy. *British Journal of Medical Psychology, 67,* 305–317.

Høglend, P., Heyerdahl, O., Amlo, S., Engelstad, V., Fossum, A., Sørbye, Ø., & Sørlie, T. (1993a). Interpretations of the patient–therapist relationship in brief dynamic psychotherapy: Effects on long-term mode specific change. *Journal of Psychotherapy Practice and Research, 2,* 296–306.

Høglend, P., Sørlie, T., Heyerdahl, O., Sørbye, Ø., & Amlo, S. (1993b). Brief dynamic psychotherapy: Patient suitability, treatment length, and outcome. *Journal of Psychotherapy Practice and Research, 2,* 230–241.

Holmes, J. (1992). *Between art and science: Essays in psychotherapy and psychiatry*. London: Routledge.

Holmes, J. (1993). *John Bowlby and attachment theory*. London: Routledge.

Holmes, J. (1995). How I assess for psychoanalytic psychotherapy. In C. Mace (Ed.), *The art and science of assessment in psychotherapy* (pp. 25–48). London: Routledge.

Horowitz, L. M., Rosenberg, S. E., Ureño, G., Kalehzan, M., & O'Halloran, P. (1989). Psychodynamic formulation, consensual response method, and interpersonal problems. *Journal of Consulting and Clinical Psychology, 57,* 599–606.

Horowitz, M. J. (1986). *Stress response syndromes*. Northvale, NJ: Aronson.

Horowitz, M. J. (1988a). *Introduction to psychodynamics: A new synthesis*. New York: Basic Books.

Horowitz, M. J. (1988b). Psychodynamic phenomena and their explanation. In M. J. Horowitz (Ed.), *Psychodynamics and cognition* (pp. 3–20). Chicago: University of Chicago Press.

Horowitz, M. J. (1991a). Emotionality and schematic control processes. In M. J. Horowitz (Ed.), *Person schemas and maladaptive interpersonal patterns* (pp. 413–423). Chicago: University of Chicago Press.

Horowitz, M. J. (1991b). Short-term dynamic therapy of stress response syndromes. In P. Crits-Christoph & J. P. Barber (Eds.), *Handbook of short-term dynamic psychotherapy* (pp. 166-198). New York: Basic Books.

Horowitz, M. J. (1995). Defensive control states and person schemas. In T. Shapiro & R. N. Emde (Eds.), *Research in psychoanalysis: Process, development, outcome* (pp. 67–89). Madison, CT: International Universities Press.

Horowitz, M. J., & Eells, T. D. (1993). Case formulations using role-relationship model configurations: A reliability study. *Psychotherapy Research, 3,* 57–68.

Horowitz, M. J., Fridhandler, B., & Stinson, C. D. (1991). Person schemas and emotion. *Journal of the American Psychoanalytic Association, 39,* 173–208.

Horowitz, M. J., Luborsky, L., & Popp, C. (1991). A comparison of the role relationship models configuration and the core conflictual relationship theme. In M. J. Horowitz (Ed.), *Person schemas and maladaptive interpersonal patterns* (pp. 213–219). Chicago: University of Chicago Press.

Horowitz, M., J. Milbrath, C., Reidbord, S., & Stinson, C. (1993). Elaboration and dyselaboration: Measures of expression and defence in discourse. *Psychotherapy Research, 3,* 278–293.

Howard, K. I., Cornille, T. A., Lyons, J. S., Vessey, J. T., Lueger, R. J., & Saunders, S. M. (1996). Patterns of mental health service utilization. *Archives of General Psychiatry, 53,* 696–703.

Howard, K. I., Davidson, C. V., O'Mahoney, M. T., Orlinsky, D. E., & Brown, K. P. (1991). Patterns of psychotherapy utilization. *American Journal of Psychiatry, 146,* 775–778.

Joseph, B. (1986). Transference: The total situation. *International Journal of Psycho-Analysis, 66,* 447–454.

Joyce, A. S., & Piper, W. E. (1990). An examination of Mann's model of time-limited psychotherapy. *Canadian Journal of Psychiatry, 35,* 41–49.

Karasu, T. (1990). Towards a clinical model of psychotherapy for depression. *American Journal of Psychiatry, 147,* 269–278.

Kazdin, A. E. (1988). *Child psychotherapy: Developing and identifying effective treatment*. New York: Pergamon.

Kennedy, H., & Moran, G. (1991). Reflections on the aims of child psychoanalysis. *The Psychoanalytic Study of the Child, 46,* 181–198.

Kernberg, O. F. (1975). *Borderline conditions and pathological narcissism*. New York: Aronson.

Kernberg, O. F. (1981). Structural interviewing. *Psychiatric Clinics of North America, 4,* 169–195.

Kernberg, O. F. (1984). *Severe personality disorders: Psychotherapeutic strategies*. New Haven, CT: Yale University Press.

Kernberg, O. F. (1988). Object relations theory in clincial practice. *Psychoanalytic Quarterly, LVII,* 481–504.

Kernberg, O. F. (1995). An ego psychology-object relations theory approach to the transference. *Psychoanalytic Quarterly, 51,* 197–221.

King, P. (1978). Affective response of the analyst to the patient's communications. *International Journal of Psycho-Analysis, 59,* 329–334.

Klauber, J. (1972). On the relationship of transference and interpretation in psychoanalytic therapy. *International Journal of Psycho-Analysis, 53,* 385-391.

Klein, M. (1946). Notes on some schizoid mechanisms. In M. Klein, P. Heimann, S. Issacs, & J. Riviere (Eds.), *Developments in psychoanalysis* (pp. 292–320). London: Hogarth.

Klein, M. (1950). On the criteria for the termination of a psychoanalysis. *International Journal of Psycho-Analysis, 31,* 78–80.

Kohut, H. (1977). *The restoration of the self*. New York: International Universities Press.

Kohut, H. (1984). *How does analysis cure?* Chicago: University of Chicago Press.

Kohut, H., & Wolf, E. S. (1978). The disorders of the self and their treatment: An outline. *International Journal of Psycho-Analysis, 59,* 413–425.

Langs, R. (1976). *The bipersonal field*. New York: Aronson.

Langs, R. (1978). Some communicative properties of the bipersonal field. *International Journal of Psychoanalytic Psychotherapy, 7,* 87–135.

Leiman, M. (1994a). The development of cognitive analytic therapy. *International Journal of Short Term Psychotherapy, 9,* 67–82.

Leiman, M. (1994b). Projective identification as early joint action sequences: A Vygotskian addendum to the procedural sequence object relations model. *British Journal of Medical Psychology, 67,* 97–106.

Levenson, E. (1983). *The ambiguity of change*. New York: Basic Books.

Limentani, A. (1972). The assessment of analysability: A major hazard in selection for psychoanalysis. *International Journal of Psycho-Analysis, 53,* 351–361.

Luborsky, L. (1984). *Principles of psychoanalytic psychotherapy: A manual for supportive-expressive (SE) treatment*. New York: Basic Books.

Luborsky, L., & Crits-Christoph, P. (1990). *Understanding transference: The CCRT method*. New York: Basic Books.

Luborsky, L., Crits-Christoph, P., Mintz, J., & Auerbach, A. (1988). *Who will benefit from psychotherapy? Predicting therapeutic outcomes*. New York: Basic Books.

Luborsky, L., & Luborsky, E. (1995). The era of measures of transference: The CCRT and other measures. In T. Shapiro & R. N. Emde (Eds.), *Research in psychoanalysis* (pp. 329–351). Madison, CT: International Universities Press.

Luborsky, L., & Mark, D. (1991). Short-term supportive-expressive psychoanalytic psychotherapy. In P. Crits-Christoph & J. P. Barber (Eds.), *Handbook of short-term dynamic psychotherapy* (pp. 110–136). New York: Basic Books.

Luborsky, L., & Schaffler, P. (1990). Illustrations of the CCRT Scoring Guide. In L. Luborsky & P. Crits-Christoph (Eds.), *Understanding transference: The CCRT method* (pp. 51–81). New York: Basic Books.

Malan, D. H. (1963). *A study of brief psychotherapy*. New York: Plenum.

Malan, D. H. (1976a). *The frontier of brief psychotherapy*. New York: Plenum.

Malan, D. H. (1976b). *Toward the validation of dynamic psychotherapy: A replication*. New York: Plenum.

Malan, D. H. (1980). Criteria for selection. In H. Davanloo (Ed.), *Short-term dynamic psychotherapy* (pp. 169–189). New York: Aronson.

Mann, J. (1973). *Time-limited psychotherapy*. Cambridge, MA: Harvard University Press.

Mann, J., & Goldman, R. (1982). *A casebook in time-limited psychotherapy*. New York: McGraw-Hill.

Marziali, E. (1984). Prediction of outcome of brief psychotherapy from therapist interpretive interventions. *Archives of General Psychiatry, 41,* 301–304.

Meltzer, D. (1967). *The psychoanalytical process*. London: Heinemann.

Menninger, K. (1958). *Theory of psychoanalytic technique*. New York: Basic Books.

Messer, S. B., Tishby, O., & Spillman, A. (1992). Taking context seriously in psychotherapy research: Relating therapist interventions to patient progress in brief psychodynamic therapy. *Journal of Consulting and Clinical Psychology, 60,* 678–688.

Messer, S. B., & Warren, C. S. (1995). *Models of brief psychodynamic therapy: A comparative approach*. New York: Guilford.

Mitchell, S. A. (1988). *Relational concepts in psychoanalysis: An integration*. Cambridge, MA: Harvard University Press.

Moran, G. S., Fonagy, P., Kurtz, A., Bolton, A. M., & Brook, C. (1991). A controlled study of the psychoanalytic treatment of brittle diabetes. *Journal of the American Academy of Child Psychiatry, 30,* 241–257.

Ogden, T. (1986). *The matrix of the mind: Object relations and the psychoanalytic dialogue*. New York: Aronson.

Ogden, T. H. (1979). On projective identification. *International Journal of Psycho-Analysis, 60,* 357–373.

Olfson, M., & Pincus, H. A. (1994). Outpatient psychotherapy in the United States: II. Patterns of utilization. *American Journal of Psychiatry, 151,* 1289–1294.

Olfson, M., Pincus, H. A., & Dial, T. H. (1994). Professional practice patterns of US psychiatrists. *American Journal of Psychiatry, 151,* 89–95.

Oliver, J. E. (1993). Intergenerational transmission of child abuse: Rates, research, and clinical implications. *American Journal of Psychiatry, 150,* 1315–1324.

O'Malley, S. S., Suh, C. S., & Strupp, H. H. (1983). The Vanderbilt Psychotherapy Process Scale: A report on the scale development and a process-outcome study. *Journal of Consulting and Clinical Psychology, 51,* 581–586.

Orlinsky, D. E., Grawe, K., & Parks, B. K. (1994). Process and outcome in psychotherapy. In A. E. Bergin & S. L. Garfield, *Handbook of psychotherapy and behaviour change* (pp. 270–378). New York: Wiley.

O'Shaughnessy, E. (1992). Enclaves and excursions. *International Journal of Psycho-Analysis, 73,* 603–611.

Osofsky, J. (1995). Applied psychoanalysis: How research with infants and adolescents at high psychosocial risk informs psychoanalysis. In T. Shapiro & R. N. Emde (Eds.), *Research in psychoanalysis* (pp. 193–207). Madison, CT: International Universities Press.

Payne, S. (1950). Short communication on criteria for terminating analysis. *International Journal of Psycho-Analysis, 31,* 205.

Pedder, J. (1988). Termination reconsidered. *International Journal of Psycho-Analysis, 69,* 495–505.

Perry, J., Cooper, A., & Michels, R. (1987). The psychodynamic formulation. *American Journal of Psychiatry, 144,* 543–550.

Perry, J., Luborsky, L., Silberschatz, G., & Popp, C. (1989). An examination of three methods of psychodynamic formulation based on the same videotaped interview. *Psychiatry, 52,* 302–323.

Piper, W. E., Azim, H. F. A., Joyce, A. S., & McCallum, M. (1991). Transference interpretations, therapeutic alliance, and outcome in short-term individual psychotherapy. *Archives of General Psychiatry, 48,* 946–953.

Piper, W. E., Azim, H. F. A., McCallum, M., & Joyce, A. S. (1990). Patient suitability and outcome in short-term individual psychotherapy. *Journal of Consulting and Clinical Psychology, 58,* 475–481.

Piper, W. E., de Carufel, F. L., & Szkrumelak, N. (1985). Patient predictors of process and outcome in short term individual psychotherapy. *Journal of Nervous and Mental Disease, 173,* 726–733.

Piper, W. E., Debbane, E. G., Bienvenu, J. P., & Garant, J. (1984). A comparative study of four forms of psychotherapy. *Journal of Clinical and Consulting Psychology, 52,* 268–279.

Piper, W. E., Joyce, A. S., McCallum, M., & Azim, H. F. A. (1993). Concentration and correspondence of transference interpretation in short-term psychotherapy. *Journal of Consulting and Clinical Psychology, 61,* 586–610.

Pollock, P. H., & Kear-Colwell, J. J. (1994). Women who stab: A personal construct analysis of sexual victimisation and offending behaviour. *British Journal of Medical Psychology, 67,* 13–22.

Popp, C., Luborsky, L., & Crits-Christoph, P. (1990). The parallel of the CCRT from therapy narratives with the CCRT from dreams. In L. Luborsky & P. Crits-Christoph (Eds.), *Understanding transference: The CCRT method* (pp. 157–172). New York: Basic Books.

Quintana, S. M. (1993). Toward an expanded and updated conceptualization of termination: Implications for short-term, individual psychotherapy. *Professional Psychology: Research and Practice, 24,* 426–432.

Racker, H. (1968). *Transference and countertransference*. New York: International Universities Press.

Richardson, L. M., & Austad, C. S. (1991). Realities of mental health practice in managed-care settings. *Professional Psychology: Research and Practice, 22,* 53–57.

Rosenfeld, H. (1952). Transference phenomena and transference analysis in an acute catatonic schizophrenic patient. *International Journal of Psycho-Analysis, 33,* 457–464.

Rosenfeld, H. (1964). On the psychopathology of narcissism: A clinical approach. *International Journal of Psycho-Analysis, 45,* 332–337.

Rosenfeld, H. (1971). A clinical approach to the psychoanalytic theory of the life and death instincts: An investigation into to the aggressive aspects of narcissism. *International Journal of Psycho-Analysis, 52,* 169–178.

Rosenfeld, H. (1975). Negative therapeutic reaction. In P. L. Giovacchini (Ed.), *Tactics and techniques in psychoanalytic therapy* (Vol. 2, pp. 217–228). London: Hogarth.

Rosser, R., Birch, S., Bond, H., Denford, J., & Schachter, J. (1987). Five year follow-up of patients treated with inpatient psychotherapy at the Cassel Hospital for Nervous Diseases. *Journal of the Royal Society of Medicine, 80,* 549–555.

Roth, A., & Fonagy, P. (1996). *What works for whom: Limitations and implications of the research literature*. New York: Guilford.

Ryle, A. (1982). *Psychotherapy: A cognitive integration of theory and practice*. London: Academic Press.

Ryle, A. (1985). Cognitive theory, object relations and the self. *British Journal of Medical Psychology, 58* 1–7.

Ryle, A. (1990). *Cognitive analytic therapy: Active participation in change*. Chichester, UK: Wiley.

Ryle, A. (1992). Critique of a Kleinian case presentation. *British Journal of Medical Psychology, 65,* 309–317.

Ryle, A. (1993). Addiction to the death instinct? A critical review of Joseph's paper "Addiction to near death". *British Journal of Psychotherapy, 10,* 88–92 (with response by A. Scott, 93–96).

Ryle, A. (1994). Psychoanalysis and cognitive analytic therapy. *British Journal of Psychotherapy, 10,* 402–405.

Ryle, A. (1995). Research relating to CAT. In A. Ryle (Ed.), *Cognitive analytic therapy: Developments in theory and practice* (pp. 175–189). Chichester, UK: Wiley.

Safran, J. D. (1990a). Towards a refinement of cognitive analytic therapy in the light of interpersonal theory: Practice. *Clinical Psychology Review, 10,* 107–121.

Safran, J. D. (1990b). Towards a refinement of cognitive analytic therapy in the light of interpersonal theory: Theory. *Clinical Psychology Review, 10,* 87–105.

Sampson, H. (1992). The role of "real" experience in psychopathology and treatment. *Psychoanalytic Dialogues, 2,* 509–528.

Sandler, J. (1976a). Actualization and object relationships. *Journal of the Philadelphia Association of Psychoanalysis, 3,* 59–70.

Sandler, J. (1976b). Countertransference and role-responsiveness. *International Review of Psycho-Analysis, 3,* 43–47.

Sandler, J. (1990). Internal objects and internal object relationships. *Psychoanalytic Inquiry, 10,* 163–181.

Sandler, J. (1992). Reflections on developments in the theory of psychoanalytic technique. *International Journal of Psycho-Analysis, 73,* 189–198.

Sandler, J. (1993). Communication from patient to analyst: Not everything is projective identification. *British Psycho-Analytical Society Bulletin, 29,* 8–16.

Sandler, J., Dare, C., & Holder, A. (1992). *The patient and the analyst.* (2nd ed.). London: Karnac.

Sandler, J., & Sandler, A.-M. (1994). Theoretical and technical comments on regression and anti-regression. *International Journal of Psycho-Analysis, 75,* 431–439.

Segal, H. (1962). The curative factors in psycho-analysis. *International Journal of Psycho-Analysis, 43,* 212–227.

Shapiro, D., Rees, A., Barkham, M., Hardy, G., Reynolds, S., & Startup, M. (1995). Effects of treatment duration and severity of depression on the maintenance of gains after cognitive-behavioral and psychodynamic-interpersonal psychotherapy. *Journal of Consulting and Clinical Psychology, 63,* 378–387.

Shapiro, D. A., Barkham, M., Rees, A., Hardy, G. E., Reynolds, S., & Startup, M. (1994). Special Feature. Effects of treatment duration and severity of depression on the effectiveness of cognitive-behavioral and psychodynamic-interpersonal psychotherapy. *Journal of Consulting and Clinical Psychology, 62,* 522–534.

Shapiro, D. A., & Firth, J. A. (1985). *Exploratory therapy manual for the Sheffield Psychotherapy Project (SAPU Memo 733)*. Sheffield, UK: University of Sheffield.

Shefler, G., Dasberg, H., & Ben-Shakhar, G. (1995). A randomized controlled outcome and follow-up study of Mann's time-limited psychotherapy. *Journal of Consulting and Clinical Psychology, 63,* 585–593.

Sifneos, P. E. (1979). *Short-term dynamic psychotherapy: Evaluation and technique*. New York: Plenum.

Sifneos, P. E. (1987). *Short-term dynamic psychotherapy: Evaluation and technique* (2nd ed.). New York: Plenum.

Sifneos, P. E. (1992). *Short-term anxiety-provoking psychotherapy*. New York: Basic Books.

Sifneos, P. E., Apfel, R. J., Bassuk, E., Fishman, G., & Gill, A. (1980). Ongoing outcome research on short-term dynamic psychotherapy. *Psychotherapy and Psychosomatics, 33,* 233–241.

Silberschatz, G., & Curtis, J. T. (1993). Measuring the therapist's impact on the patient's therapeutic progress. *Journal of Consulting and Clinical Psychology, 61,* 403–411.

Silberschatz, G., Fretter, P. B., & Curtis, J. T. (1986). How do interpretations influence the process of psychotherapy. *Journal of Consulting and Clinical Psychology, 54,* 646–652.

Sinason, V. (1992). *Mental handicap and the human condition*. London: Free Association Books.

Sledge, W. H., Moras, K., Hartley, D., & Levine, M. (1990). Effect of time-limited psychotherapy on patient dropout rates. *American Journal of Psychiatry, 147,* 1341–1347.

Spence, D. P. (1982). *Narrative truth and historical truth. Meaning and interpretation in psychoanalysis*. New York: Norton.

Spillius, E. B. (1994). Developments in Kleinian thought: Overview and personal view. *Psychoanalytic Inquiry, 14,* 324–364.

Steiner, J. (1982). Perverse relationships between parts of the self: A clinical illustration. *International Journal of Psycho-Analysis, 63,* 241–251.

Steiner, J. (1993). *Psychic retreats: Pathological organisations in psychotic, neurotic and borderline patients*. London: Routledge.

Stern, D. J. (1994). One way to build a clinically relevant baby. *Infant Mental Health Journal, 15,* 36–54.

Stevenson, J., & Meares, R. (1992). An outcome study of psychotherapy for patients with borderline personality disorder. *American Journal of Psychiatry, 149,* 358–362.

Stewart, H. (1989). Technique at the basic fault: Regression. *International Journal of Psycho-Analysis, 70,* 221–230.

Stone, M. (1993). Long-term outcome in personality disorders. *British Journal of Psychiatry, 162,* 299–313.

Strachey, J. (1934). The nature of the therapeutic action of psychoanalysis. *International Journal of Psycho-Analysis, 50,* 275–292.

Strupp, H. H., & Binder, J. L. (1984). *Psychotherapy in a new key: A guide to time-limited dynamic psychotherapy*. New York: Basic Books.

Strupp, H. H., & Hadley, S. W. (1979). Specific versus nonspecific factors in psychotherapy. *Archives of General Psychiatry, 36,* 1125–1136.

Sullivan, H. S. (1953). *The interpersonal theory of psychiatry*. New York: Norton.

Svartberg, M., & Stiles, T. C. (1992). Predicting patient change from therapist competence and patient–therapist complementarity in short-term anxiety-provoking psychotherapy: A pilot study. *Journal of Consulting and Clinical Psychology, 60,* 304–307.

Svartberg, M., & Stiles, T. C. (1994). Therapeutic alliance, therapist competence, and client change in short-term anxiety-provoking psychotherapy. *Psychotherapy Research, 4,* 20–33.

Symington, N. (1983). The analyst's act of freedom as agent of therapeutic change. *International Review of Psycho-Analysis, 10,* 783–792.

Taube, C. A., Kessler, L., & Feuerberg, M. (1984). *Utilization and expenditures for ambulatory mental health care during 1980 (Data Report 5).* Washington, DC: US Department of Health and Human Services.

Thomä, H., & Kächele, H. (1987). *Psychoanalytic practice. I: Principles.* New York: Springer-Verlag.

Tishby, O., & Messer, S. B. (1995). The relationship between plan compatibility of therapist interventions and patient progress: A comparison of two plan formulations. *Psychotherapy Research, 5,* 76–88.

Tyson, A., & Sandler, J. (1971). Problems in the selection of patients for psychoanalysis: comments on the application of concepts of "indications", "suitability", and "analyzability". *British Journal of Medical Psychology, 44,* 211–228.

Vaillant, G. E. (1992). *Ego mechanisms of defense: A guide for clinicians and researchers.* Washington, DC: American Psychiatric Press.

Wachtel, P. (1977). *Psychoanalysis and behaviour therapy: Toward an integration.* New York: Basic Books.

Waldinger, R. J. (1987). Intensive psychodynamic therapy with borderline patients: An overview. *American Journal of Psychiatry, 144,* 267–274.

Wallerstein, R. S. (Ed.) (1992). *The common ground of psychoanalysis.* Northvale, NJ: Aronson.

Wallerstein, R. S. (1986). *Forty-two lives in treatment: A study of psychoanalysis and psychotherapy.* New York: Guilford.

Weiss, J., & Sampson, H. (1986). *The psychoanalytic process: Theory, clinical observations and empirical research.* New York: Guilford.

Westen, D. (1986). What changes in short-term psychodynamic psychotherapy. *Psychotherapy, 23,* 501–512.

Wiborg, I. M., & Dahl, A. A. (1996). Does brief dynamic psychotherapy reduce the relapse rate of panic disorder? *Archives of General Psychiatry, 53,* 689–694.

Windholz, M. J., & Silberschatz, G. (1988). Vanderbilt Psychotherapy Process Scale: A replication with adult outpatients. *Journal of Consulting and Clinical Psychology, 56,* 56–60.

Winnicott, D. W. (1949). Hate in the countertransference. *International Journal of Psycho-Analysis, 30,* 69–75.

Winnicott, D. W. (1965). *The maturational process and the facilitating environment.* London: Hogarth.

Winnicott, D. W. (1971). *Playing and reality.* London: Tavistock.

Winston, A., Laikin, M., Pollack, J., Samstag, L. W., McCullough, L., & Muran, J. C. (1994). Short-term dynamic psychotherapy of personality disorders. *American Journal of Psychiatry, 15,* 190–194.

Winston, A., Pollack, J., McCullough, L., Flegenheimer, W., Kestenbaum, R., & Trujillo, M. (1991). Brief psychotherapy of personality disorders. *Journal of Nervous and Mental Disease, 179,* 188–193.

Woody, G. E., Luborsky, L., McLellan, A. T., & O'Brien, C. P. (1990). Corrections and revised analyses for psychotherapy in methadone maintenance programs. *Archives of General Psychiatry, 47,* 788–789.

Woody, G. E., Luborsky, L., McLellan, A. T., O'Brien, C. P., Beck, A. T., Blaine, J., Herman, I., & Hole, A. (1983). Psychotherapy for opiate addicts: Does it help? *Archives of General Psychiatry, 40,* 639–645.

Woody, G. E., McLellan, A. T., Luborsky, L., & O'Brien, C. P. (1987). Twelve-month follow-up of psychotherapy for opiate dependence. *American Journal of Psychiatry, 144,* 591–596.

Woody, G. E., McLellan, A. T., Luborsky, L., & O'Brien, C. P. (1995). Psychotherapy in community methadone programs: A validation study. *American Journal of Psychiatry, 192,* 1302–1308.

Young, J. E. (1990). *Cognitive therapy for personality disorders: A schema-focused approach.* Sarasota, FL: Professional Resource Exchange.

Zilboorg, G. (1952). The emotional problem and the therapeutic role of insight. *Psychoanalytic Quarterly, 21,* 1–24.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.06
Psychopharmacology

PHILIP J. COWEN
University of Oxford, Warneford Hospital, UK

6.06.1 INTRODUCTION 136
  6.06.1.1 History of Drug Treatment in Psychiatry 136
  6.06.1.2 Mode of Action of Psychotropic Drugs 137
  6.06.1.3 Pharmacokinetics of Psychotropic Drugs 138
  6.06.1.4 Prescribing Psychotropic Drugs 138
6.06.2 CLASSIFICATION OF DRUGS USED IN PSYCHIATRY 139
6.06.3 ANXIOLYTIC DRUGS 139
  6.06.3.1 Benzodiazepines 139
    6.06.3.1.1 Pharmacology 139
    6.06.3.1.2 Compounds available 139
    6.06.3.1.3 Adverse effects 140
  6.06.3.2 Azapirones 140
  6.06.3.3 Other Drugs Used to Treat Anxiety 141
    6.06.3.3.1 Antidepressant drugs 141
    6.06.3.3.2 Antipsychotic drugs 141
    6.06.3.3.3 β-Adrenoceptor antagonists 141
6.06.4 HYPNOTIC DRUGS 141
  6.06.4.1 Compounds Available 141
6.06.5 ANTIPSYCHOTIC DRUGS 141
  6.06.5.1 Pharmacology 142
  6.06.5.2 Compounds Available 142
    6.06.5.2.1 Typical antipsychotics 142
    6.06.5.2.2 Atypical antipsychotic drugs 142
  6.06.5.3 Adverse Effects 143
    6.06.5.3.1 Movement disorders 143
    6.06.5.3.2 Autonomic and endocrine effects 143
    6.06.5.3.3 Neuroleptic malignant syndrome 144
    6.06.5.3.4 Other adverse effects 144
    6.06.5.3.5 Adverse effects of clozapine 144
  6.06.5.4 Dosage of Antipsychotic Drugs 145
6.06.6 ANTI-PARKINSONIAN DRUGS 145
  6.06.6.1 Preparations Available 145
  6.06.6.2 Adverse Effects 145
6.06.7 ANTIDEPRESSANTS 145
  6.06.7.1 Mechanism of Action 146
  6.06.7.2 Tricyclic Antidepressants 146
    6.06.7.2.1 Pharmacological properties 146
    6.06.7.2.2 Adverse effects of tricyclic antidepressants 146
    6.06.7.2.3 Amoxapine 147
    6.06.7.2.4 Clomipramine 147
    6.06.7.2.5 Lofepramine 147
    6.06.7.2.6 Maprotiline 148
  6.06.7.3 Selective Serotonin Reuptake Inhibitors 148

6.06.7.3.1 Pharmacological properties 148
6.06.7.3.2 Efficacy in depression 148
6.06.7.3.3 Unwanted effects of SSRIs 148
6.06.7.4 Monoamine Oxidase Inhibitors 149
6.06.7.4.1 Pharmacology 149
6.06.7.4.2 Compounds available 149
6.06.7.4.3 Efficacy of MAOIs in depression 149
6.06.7.4.4 Unwanted effects 149
6.06.7.4.5 Interactions with foodstuffs and drugs 149
6.06.7.4.6 Moclobemide 150
6.06.7.5 Other Antidepressant Drugs 150
6.06.7.5.1 Mianserin 150
6.06.7.5.2 Mirtazapine 151
6.06.7.5.3 Trazodone 151
6.06.7.5.4 Nefazodone 151
6.06.7.5.5 Venlafaxine 151
6.06.7.5.6 Bupropion 152
6.06.7.5.7 L-Tryptophan 152
6.06.8 MOOD-STABILIZING DRUGS 152
6.06.8.1 Lithium 153
6.06.8.1.1 Pharmacology 153
6.06.8.1.2 Efficacy 153
6.06.8.1.3 Adverse effects 153
6.06.8.1.4 Toxic effects 154
6.06.8.1.5 Lithium and pregnancy 154
6.06.8.2 Carbamazepine 155
6.06.8.2.1 Pharmacology 155
6.06.8.2.2 Efficacy 155
6.06.8.2.3 Adverse effects 155
6.06.8.3 Sodium Valproate 155
6.06.8.3.1 Pharmacology 155
6.06.8.3.2 Efficacy 155
6.06.8.3.3 Adverse effects 155
6.06.9 CLINICAL USE OF PSYCHOTROPIC DRUGS 156
6.06.9.1 Anxiety Disorders 156
6.06.9.1.1 Generalized anxiety disorder 156
6.06.9.1.2 Panic disorder and agoraphobia 156
6.06.9.1.3 Obsessive-compulsive disorder 157
6.06.9.1.4 Social phobia 157
6.06.9.2 Insomnia 157
6.06.9.3 Depression 157
6.06.9.3.1 Choice of antidepressant 157
6.06.9.3.2 Prophylaxis of recurrent major depression 158
6.06.9.3.3 Psychological therapies and antidepressant drug treatment 158
6.06.9.4 Mania 158
6.06.9.5 Prophylaxis of Bipolar Illness 158
6.06.9.6 Schizophrenia 159
6.06.10 REFERENCES 159

6.06.1 INTRODUCTION

It is only since the late 1940s that drug treatment has been able to play a useful role in the management of psychiatric disorders. Little is known about the pathophysiology of psychiatric illness, and the drugs that are currently in use were discovered by chance or by modification of compounds known to be efficacious. Indeed, such neurochemical theories as there are of psychiatric disorder have, in the main, been derived from a knowledge of the pharmacological effects of psychotropic drugs in animal experimental studies.

Drug treatment now plays in indisputable role in management of severe psychiatric disorder, such as schizophrenia and bipolar disorder. Drugs are also commonly used to treat less severe illnesses, such as depressive and anxiety disorders. Wherever drug treatments are used, they should form part of an overall management plan that takes into account psychological and social needs of the individual patient and their family.

6.06.1.1 History of Drug Treatment in Psychiatry

Drugs that produce changes in the function of the central nervous system, such as opiates and anticholinergic agents, have been used in the treatment of mental disorders for hundreds of years. Although some of these drugs may have had calming effects, they were of no specific value in the treatment of psychiatric disorders. Particular drug treatments tended to be used because their cause was espoused by vigorous and eminent physicians rather than on the basis of proven efficacy. In any case, assessment of efficacy depended on uncontrolled clinical observation.

The first drug that was discovered to have a specific effect on a prticular psychiatric disorder was lithium (Cade, 1949; Table 1). Lithium is a toxic agent and Cade's important clinical observations did not make a significant impact on clinical practice until the following decades when controlled trials showed that lithium was effective in both the acute treatment of mania and the prophylaxis of recurrent mood disorders. In addition, monitoring of plasma lithium levels enabled safe dosing regimes to be established.

Since the late 1960s there has been a period of consolidation in psychopharmacology during which clinical trials have been extensively employed to refine the indications of particular drug treatments and to maximize their benefit risk ratios. New compounds have continuously become available but because, in the main, these agents have been derived from previously described agents, their range of activity is not strikingly different from their predecessors. In general, however, the newer agents are better tolerated and are sometimes safer, and both these developments are important for clinical practice.

It is, perhaps, possible now to be more optimistic about the prospects for advances in psychopharmacology. There is, for example, rapidly increasing knowledge about the nature of chemical signaling in the brain. A multiplicity of neurotransmitters and neuromodulators interact with specific families of receptors, many of which exist in several different subtypes. Several of these receptors have been cloned and selective agents for them are becoming available. We are also learning much more about how these chemical messengers may modify behavior through their interactions with specific brain regions and distributed neuronal circuits.

Table 1 Introduction of some drug treatments in psychiatry.

Year	*Drug treatment*
1949	Lithium
1952	Chlorpromazine
1954	Benzodiazepines
1957	Iproniazid (MAOI)
1957	Imipramine
1971	Carbamazepine
1980	Selective serotonin reuptake inhibitors
1988	Clozapine for treatment-resistant schizophrenia

The availability of novel compounds, likely to have a quite different range of behavioral effects compared to currently available drugs, will lead to some exciting developments in psychopharmacology. It seems probable, given the complexity and multifactorial origin of psychiatric disorders, that detailed knowledge of etiology and pathophysiology, may lag behind advances in therapeutics. This, of course, is not uncommon in general medicine but serves to reinforce the importance of controlled clinical trials in the assessment of new psychopharmacological treatments.

6.06.1.2 Mode of Action of Psychotropic Drugs

Psychotropic drugs act in one way or another on the process of chemical signaling, a mechanism employed by the brain for the purposes of communication between neurones. Neurones make connection with each other at specialized regions of the cell membrane called synapses. Chemical transmission at synapses occurs when a chemical messenger or neurotransmitter is released from one neurone and interacts with a specific binding site or receptor on an adjacent neurone. In general, psychotropic drugs act to increase or decrease the impact of a neurotransmitter on its receptor. This can be achieved in a number of ways (Table 2).

Much is known about the pharmacological effects of psychotropic drugs in experimental animal studies. However, it is much less clear how such pharmacological changes are translated into therapeutic benefit in patients with psychiatric disorders. Studies in healthy volunteers may not be helpful in elucidating this issue. For example, a therapeutic course of an antidepressant drug has no effect on the mood of an individual who is not depressed. A further complication is that most psychotropic drugs take a number of weeks before their full clinical effect becomes apparent (Table 2). From this it has been deduced that the therapeutic effect of drug treatment is likely to be due to slowly evolving adaptive changes in neurotransmitter mechanisms and the neuronal circuits that depend on them. (A review of the mechanisms of drug action is given by Stahl 1996.)

When a patient has responded to psychotropic drug treatment it is usual for drug therapy to be continued for some time. Where a disorder is believed to run a self-limiting course, for example a single uncomplicated episode of major depression, it is customary to withdraw

Table 2 Mechanism of action of some psychotropic drugs and time to onset of therapeutic effect.

Drug	*Action*	*Time to onset of action*
Antidepressant	Increase action of noradrenaline and serotonin	2–4 weeks
Antipsychotic	Block dopamine receptors	2–6 weeks
Anxiolytic (benzodiazepine)	Increase action of γ-aminobutyric acid	immediate
Lithium	Decrease activity of receptor-linked enzyme systems	2–4 weeks (mania)

treatment after about six months. However, many psychiatric disorders are persistent or highly recurrent, and here drug treatment may need to be continued in the longer term.

Where circumstances are favorable it is often worthwhile trying to discontinue drug treatment, or at least lower the dose, while the patient is carefully monitored. In general, psychotropic drugs should not be discontinued suddenly because withdrawal symptoms may result. These are seen particularly with anxiolytic and antidepressant drugs. In addition, abrupt discontinuation of medication can sometimes result in "rebound" illness. For example, sudden withdrawal of lithium in patients with bipolar disorder confers a risk of mania of about 50% over the next six weeks (Goodwin, 1994). This is substantially greater than that which would be expected from the natural history of the illness.

6.06.1.3 Pharmacokinetics of Psychotropic Drugs

Before psychotropic drugs can produce their effects they need to reach the brain in adequate amounts. This depends on how well they are absorbed from the gastrointestinal tract into the blood stream and their ability to cross the blood–brain barrier. Subsequently, drugs are broken down or metabolized in the liver and then eliminated from the body in the urine by the kidney. Patients with liver or kidney disease can have exaggerated effects from small doses of drugs.

Plasma concentrations of drugs throughout the day vary, rising immediately after each dose and falling at a rate that differs between individual drugs and to some extent between individual people. However, this rate of decline influences how long the drug persists in the body.

The concept of plasma half-life is useful here. The half-life of a drug in plasma is the time taken for its concentration to fall by a half, once dosing has ceased. With most psychotropic drugs, the amount eliminated over time is proportional to plasma concentration and in this case it will take approximately five times the half-life for the drug to be eliminated from plasma. Equally, when dosing with a drug begins, it will take five times the half-life for the concentration in plasma to reach a steady state.

This can be important when planning treatment. For example, monoamine oxidase inhibitors (MAOIs) should not be given with selective serotonin reuptake inhibitors (SSRIs) because of the danger of drug interaction (see Section 6.06.7.4.5). If, for example, a patient is taking sertraline which has an elimination half-life of about 26 hours, it will be important to leave at least five times the half-life (a week is recommended) before starting MAOI treatment. When sertraline treatment begins, the plasma concentrations will continue to rise for about a week before reaching steady state.

Most psychotropic drugs have fairly long half-lives and once or twice daily dosing is sufficient. This aids compliance. Some antipsychotic preparations are made in the form of long-acting intramuscular preparations. These depot injections may need only to be given only once or twice monthly, which again has advantages in terms of compliance for some patients.

6.06.1.4 Prescribing Psychotropic Drugs

It is good practice to use well-tried drugs with therapeutic actions and side effects that are clearly understood. When a drug is prescribed it is necessary to determine the dose, the interval between doses and the likely duration of treatment. Until a clinician is thoroughly familiar with a drug it is important to consult the manufacturer's literature or other appropriate reference to decide the dosing schedule. Particular care should be taken to assess the possibility of drug interaction because many patients will be taking more than one kind of medicine.

Before providing a prescription the clinician should explain what effects are likely to be expected on first taking the drug, for example, drowsiness or dry mouth with a tricyclic antidepressant. They should also explain how long it will be before therapeutic effects will appear and what signs a patient should look for.

Many patients do not take their prescribed medication either because they do not perceive the need to take it or have fears about the possible consequences if they do. Time spent in discussing patient's concerns is well spent and is likely to improve compliance with medication. Written instructions (which are now often provided as inserts to drug packaging) can be a valuable adjunct.

There are special problems about prescribing in pregnancy because of the risk that drugs might produce adverse effects on fetal development (teratogenesis). Withdrawal symptoms in the new-born may also occur. Where possible, therefore, psychotropic drugs should not be prescribed to women who are pregnant or who are considering becoming pregnant. Where drug treatment is strongly indicated, for example, in a psychotic illness or with severe depression, a careful clinical risk benefit assessment should be made and discussed with the patient. It is usually possible to select preparations that appear on current evidence to be unlikely to produce terratogenic effects.

6.06.2 CLASSIFICATION OF DRUGS USED IN PSYCHIATRY

Psychotropic drugs are those whose main clinical effect is to produce a change in the psychological state. Psychotropic drugs used in psychiatry are conventionally divided into different classes, but the therapeutic actions of particular compounds are not confined to one diagnostic category. For example, SSRIs are classified as antidepressants and are effective in the treatment of major depression, but they also produce useful therapeutic effects in panic disorder, obsessive compulsive disorder and social phobia (Cowen, 1997). This breadth of effect does not mean that the latter syndromes are forms of depression. It merely emphasizes that the neuropsychological consequences of facilitating brain serotonin function may provide beneficial effects in a variety of psychiatric disorders.

Although there is considerable understanding of the pharmacological actions of psychotropic drugs, little is known about the neuropsychological consequences of these pharmacological actions and about the ways in which neuropsychological changes are translated into clinical benefit in different diagnostic syndromes. At present, therefore, the best plan is to classify drugs according to their major therapeutic use but to bear in mind therapeutic effects of different classes of drugs may overlap (Table 3).

These groups of drugs will be discussed in turn. Subsequently, general advice will be given about the use of psychotropic drugs in different psychiatric disorders and how they may be combined with psychological methods of treatment.

6.06.3 ANXIOLYTIC DRUGS

Anxiolytic drugs are indicated in the treatment of anxiety disorders. It must be remembered, however, that the classification of anxiety disorders in the *Diagnostic and statistical manual of mental disorders* (4th ed., *DSM-IV*) and the *International classification of diseases* (10th ed., *ICD-10*) subsumes a number of different disorders, some of which have a distinct drug response. For example, although benzodiazepines are effective in generalized anxiety disorder and to some extent in panic disorder, they are not useful in the treatment of obsessive compulsive disorder (Jenike, 1992).

Anxiolytic drugs are prescribed widely and often inappropriately. There is evidence, however, that this trend may be subsiding (Tyrer, 1997). To some extent this may reflect substitution of antidepressant drugs for benzodiazepines, since the former are effective in the treatment of anxiety disorders and may be less likely to cause dependence.

6.06.3.1 Benzodiazepines

6.06.3.1.1 Pharmacology

Benzodiazepines are anxiolytic, sedative, and in larger doses hypnotic. They also have muscle relaxant and anticonvulsant properties. Their pharmacological actions are mediated through specific receptor sites, located in a supramolecular complex with γ-aminobutyric acid (GABA) receptors. Benzodiazepines enhance GABA neurotransmission, thereby altering indirectly the activity of many other neurotransmitters, for example, noradrenaline and serotonin (Stahl, 1996).

6.06.3.1.2 Compounds available

Many different benzodiazepines are available. They differ both in the potency with which they interact with a benzodiazepine receptor, and in their plasma half-life. In general, high potency benzodiazepines and those with short half-lives are more likely to be associated with dependence and withdrawal. Benzodiazepines with short half-lives (less than 12 hours) include lorazepam and temazepam. Because of problems with dependence, long-acting compounds are preferable for the management of anxiety, even if such treatment is to be given intermittently and on an as-required basis. The long-acting benzodiazepines include drugs such as diazepam and chlordiazepoxide.

Table 3 Classification of clinical psychotropic drugs.

Class of drug	*Examples of classes*	*Indications*
Antipsychotic	Phenothiazines, Butyrophenones, Dibenzazepine	Acute treatment of schizophrenia and mania; prophylaxis of schizophrenia
Antidepressant	Tricyclic antidepressants MAOIs SSRIs SNRIs	Major depression (acute treatment and prophylaxis); anxiety disorders; obsessive-compulsive disorder (SSRIs)
Mood stabilizer	Lithium, Carbamazepine, Valproate	Acute treatment of mania; prophylaxis of recurrent mood disorder
Anxiolytic	Benzodiazepines, azapirones (buspirone)	Generalized anxiety disorder
Hypnotic	Benzodiazepines, Cyclopyrrolones (zopiclone), imidazopyridine (zolpidem)	Insomnia

Diazepam is rapidly absorbed, and can be used both for continuous treatment of anxiety and for treatment as required. Alprazolam, a high-potency and long-acting benzodiazepine, is used widely outside the UK for the treatment of panic disorder. This therapeutic efficacy is not confined to alprazolam, however, because equivalent doses of other high-potency agents, such as clonazepam, are also effective (Nutt & Bell, 1997). Flumazenil is a benzodiazepine receptor antagonist. This drug produces little pharmacological effect by itself but blocks the action of other benzodiazepines (Nutt, Cowen, & Little, 1982). It is therefore used to reverse acute toxicity of benzodiazepines, but carries a risk of provoking withdrawal symptoms in chronic users.

6.06.3.1.3 Adverse effects

Benzodiazepines are generally well tolerated. When they are given as anxiolytics, their main side effects are due to sedative properties that can lead to ataxia and drowsiness. A degree of cognitive impairment may be detectable, which is obviously a matter of concern when people are driving or operating machinery. The adverse effects of benzodiazepines are potentiated by alcohol. Although in some circumstances benzodiazepines may lower tension and aggression, in some people they can increase aggressive behavior, probably through disinhibition (Cowdrey & Gardner, 1988). They should therefore be prescribed only with great caution to those with a previous history of impulsive aggressive behavior.

It is now generally agreed that physical dependence develops after prolonged use of benzodiazepines. The frequency depends on the drug and the dosage, and has been estimated at between 5% and 50% among patients taking the drugs for more than six months. Although escalation during treatment is unusual, problems can appear when patients try to discontinue their medication, whereupon a withdrawal syndrome becomes apparent. This is characterized by anxiety, insomnia, nausea, and tremor, together with perceptual disturbances (Pertursson & Lader, 1984; Tyrer, 1997).

Withdrawal symptoms generally begin within 2–3 days of stopping a short-acting benzodiazepine, or within about seven days of stopping a longer-acting one. The symptoms generally last for 3–10 days. If benzodiazepines have been taken for a long time, it is best to withdraw them gradually over several weeks under supervision (Tyrer, Rutherford, & Huggett, 1981). Despite this, a few patients either cannot discontinue their benzodiazepines satisfactorily or are troubled by persistent long-standing withdrawal symptoms for long periods of time after the drugs have been discontinued.

6.06.3.2 Azapirones

The only drug in this class currently marketed for the treatment of anxiety is buspirone. This drug has no affinity for benzodiazepine receptors but stimulates a subtype of serotonin receptor called the serotonin-1A receptor. This receptor is found in high concentration in the raphe nuclei in the brain stem, where it regulates the firing of serotonin cell bodies. Administration of buspirone lowers the firing rate of serotonin neurones and thereby decreases serotonin neurotransmission in certain brain regions. This action may be the basis of its anxiolytic effects (Yocca, 1990).

Buspirone is different to benzodiazepines in that its anxiolytic effects take several days to develop, whereas those of benzodiazepines are apparent very quickly. Its side effect profile also differs; for example, it is associated with lightheadedness, nervousness, and headache early in treatment. There is little evidence that tolerance and dependence occur during buspirone use, although such judgment must always be made with circumspection. There is some evidence that patients who have previously responded to treatment with benzodiazepines do not respond well to buspirone. Buspirone cannot be used to treat benzodiazepine withdrawal. Although buspirone appears to be effective in the treatment of generalized anxiety disorder, current evidence does not suggest that it confers benefit in the treatment of panic disorder (Cowen, 1992).

6.06.3.3 Other Drugs Used to Treat Anxiety

6.06.3.3.1 Antidepressant drugs

Antidepressants usually ameliorate the anxiety that accompanies depressive disorders. In addition, tricyclic antidepressants and trazodone have been shown to be as effective as benzodiazepines in the management of generalized anxiety and panic disorder (Rickels, Downing, Schweizer, & Hassman, 1993). SSRIs and MAOIs are also effective in the treatment of panic disorder, but the selective noradrenaline reuptake inhibitor, maprotiline, is not (Den Boer & Westenberg, 1988).

6.06.3.3.2 Antipsychotic drugs

These drugs are sometimes prescribed for their anxiolytic effect. They are not more effective than benzodiazepines, but may have a place in the group of patients who have become irritable and disinhibited with benzodiazepines.

6.06.3.3.3 β-Adrenoceptor antagonists

These include drugs such as propanolol which are used to treat hypertension. Such drugs relieve some of the autonomic (peripheral) symptoms of anxiety, such as tachycardia and tremor, by blocking peripheral β-adrenoceptors. In general they are not particularly helpful in the treatment of patients with anxiety disorders, but can be helpful in otherwise healthy subjects who develop marked autonomic symptoms coupled to performance anxiety.

6.06.4 HYPNOTIC DRUGS

Hypnotic drugs are used to improve sleep. The benzodiazepine drugs described previously are widely used as hypnotics. Most hypnotics in common use act at the benzodiazepine GABA receptor complex; this includes more recently introduced compounds such as zolpidem and zopiclone. Occasionally, low-dose tricyclic antidepressants are used as hypnotics because of their sedating effects. Similarly, sedating antihistamines such as chlopheniramine have been employed to promote sleep.

Hypnotic drugs have two major problems; first, the development of tolerance, with rebound insomnia when medication is discontinued; and second, hangover effects that can compromise psychological performance the next day.

6.06.4.1 Compounds Available

The most commonly used hypnotics are benzodiazepines with short half-lives, such as temazepam and lormetazepam. However, there is increasing use of nonbenzodiazepine drugs such as zopiclone, a cyclopyrrolone, and zolpidem, an imidazopyridine. With these shorter-acting compounds daytime hangover is less common but is still experienced by some patients.

Both zopiclone and zolpidem bind to a site close to the benzodiazepine receptor, thereby facilitating brain GABA function. Zopiclone and zolpidem produce fewer changes in sleep architecture than benzodiazepines, and are also claimed to be less liable to produce tolerance and dependence (Langtry & Benfield, 1990). This has not yet been fully substantiated. The most common side effect of zopiclone is a bitter after-taste following injection, but behavioral disturbances including confusion, amnesia, and depressed mood have been reported. Zolpidem has also been associated with behavioral disturbances but more commonly causes nausea and dizziness.

Other hypnotic drugs include chloral hydrate and chlormethiazole. The latter has a short half-life and is commonly used to facilitate sleep in the elderly. In addition, in some countries, chlormethiazole is used to prevent withdrawal symptoms in patients dependent on alcohol. For this reason it is sometimes thought, mistakenly, to be a suitable hypnotic for alcoholic patients. In fact, chlormethiazole has barbiturate-like actions and can cause respiratory depression when combined with alcohol or in overdose.

6.06.5 ANTIPSYCHOTIC DRUGS

This term is applied to drugs that reduce psychomotor overactivity and diminish symptoms of psychosis. Alternative terms for these drugs are neuroleptic and major tranquilizer. None of these names is wholly satisfactory. Neuroleptic refers to the side effects rather than to the therapeutic effects of the drugs, and major tranquilizer does not refer to the most important clinical action, that of ameliorating the symptoms of psychosis. Therefore, the term antipsychotic drug is preferred here.

The main therapeutic uses of antipsychotic drugs are to reduce hallucinations, delusions, agitation, and psychomotor excitement in schizophrenia, mania, or psychosis secondary to a medical condition. The drugs are also used prophylactically to prevent relapse of schizophrenia and occasionally mania (Gelder, Gath, Mayou, & Cowen, 1996).

6.06.5.1 Pharmacology

Antipsychotic drugs share the property of blocking brain dopamine receptors. Dopamine receptors are of several biochemical subtypes, but most antipsychotic drugs bind strongly to dopamine-D_2 receptors, and this action appears to account both for their antipsychotic activity and their propensity to cause movement disorders.

Actions at other neurotransmitter receptors may offset the liability of D_2 receptor antagonists to produce movement disorders. For example, thioridazine is a potent antagonist at muscarinic cholinergic receptors, and anticholinergic drugs are known to possess anti-Parkinsonian effects. This might account for the diminished liability of thioridazine to cause movement disorders. Similarly, the lack of movement disorders associated with risperidone and olanzapine has been attributed to the ability of these drugs to block serotonin_2 receptors as well as D_2 receptors (see Stahl, 1996).

6.06.5.2 Compounds Available

A large number of antipsychotic compounds have been developed. The main distinction of clinical utility is into typical and atypical antipsychotics. Atypical antipsychotics are so called because they have a decreased likelihood to cause extrapyramidal side effects or movement disorders.

6.06.5.2.1 Typical antipsychotics

These drugs have comparable efficacy in the treatment of psychosis but different side effect profiles because of their other pharmacological properties. For example, phenothiazines such as chlorpromazine and thioridazine are sedating and tend to lower blood pressure. They are, however, less likely to cause movement disorders than nonsedating phenothiazines, such as trifluoperazine and fluphenazine. The pharmacological profile of the thioxanthenes, such as flupenthixol and clopenthixol, resembles the nonsedating phenothiazines and the same is true of the butyrophenone, haloperidol.

A number of typical antipsychotic drugs (the decanoates of haloperidol, fluphenazine, flupenthixol and zuclopenthixol) are available as long-acting intramuscular depot preparations, given at intervals of 2–4 weeks. For many patients continued medication is the only way to prevent psychotic relapse. In some subjects compliance with medication is better if they will accept a long-acting intramuscular preparation.

Another intramuscular preparation is zuclopenthixol acetate whose action lasts for 1–2 days. This preparation can be useful in an acutely psychotic patient where antipsychotic medication is needed but where oral drug administration is ineffective or not possible. The disadvantage of this approach is that it is not easily possible to titrate the dosage of medication for an individual patient (Royal College of Psychiatrists, 1993).

6.06.5.2.2 Atypical antipsychotic drugs

These include a number of compounds that differ strikingly in their structure and pharmacological properties. Sulpiride is a D_2 receptor blocker, but it appears to act more selectively on D_2 receptors on mesolimbic and mesocortical regions than on those in the basal ganglia. This might account for its reduced liability to cause extrapyramidal side effects. In contrast, risperidone is a potent D_2 receptor antagonist but is even more potent at blocking serotonin_2 receptors. Concomitant serotonin_2 receptor blockade is believed to attenuate the movement disorders caused by unnopposed D_2 receptor antagonism (Livingston, 1994).

Olanzapine and sertindole are also serotonin_2 receptor antagonists but have less effect than risperidone on D_2 receptors. Both these factors may account for their decreased liability to cause movement disorders (Gerlach & Peacock, 1995; Reus, 1997). In addition, it has been claimed that risperidone, olanzepine, and sertindole are more effective than typical antipsychotics, particularly having some activity against so-called negative symptoms of schizophrenia, which are difficult to treat pharmacologically (Marder & Meibach, 1994; Tollefson & Sanger, 1997). These claims are currently being tested but clinically risperidone and olanzepine are being used increasingly in preference to typical antipsychotic drugs.

Clozapine is an important drug because it is the only drug with established efficacy in patients who are resistant to other antipsychotic treatments (Kane, Honigfeld, Singer, & Meltzer, 1988). Clozapine has complex pharmacology with weak binding to dopamine-D_2 receptors but strong antagonist properties at a variety of serotonin and noradrenergic receptors. Because of a rare but serious adverse effect on white blood cells, clozapine can be used only with special monitoring (see below).

6.06.5.3 Adverse Effects

6.06.5.3.1 Movement disorders

Probably the most troublesome side effect of antipsychotic drugs, particularly typical antipsychotics, are movement disorders, which are a consequence of D_2 receptor blockade in the basal ganglia. At least four different kinds of movement disorder are recognized: acute dystonia, akathisia (a Parkinsonian syndrome), and tardive dyskinesia (Table 4).

Acute dystonia occurs soon after treatment begins, especially in young men. It is observed most often with drugs such as haloperidol and trifluoperazine. The main features are caused by acute contraction of muscle groups, resulting in torticollis (neck twisting), tongue protrusion, grimacing and oculogyric crisis (rolling upward of the eyes). These symptoms are distressing and alarming for patients, and can be controlled by anticholinergic drugs, such as benztropine or a benzodiazepine.

Akathisia is an unpleasant feeling of physical restlessness, with a compelling need to move. Patients are usually very distressed by akathisia and can present in a state of severe agitation. It usually occurs in the first few weeks of treatment with antipsychotic drugs but may only begin after several months. Akathisia is not reliably treated by anticholinergic drugs, but may disappear when the dose of the antipsychotic agent is reduced. Some cases are helped by treatment with β-adrenoceptor antagonists, such as propranolol, and short-term benzodiazepine administration has also been employed.

The Parkinsonian syndrome caused by antipsychotic drugs presents with the usual clinical triad of akinesia, tremor, and muscular rigidity. These symptoms can be controlled by lowering the dose of the antipsychotic drug or with anticholinergic agents (see below).

The last syndrome, tardive dyskinesia, is particularly serious because, unlike the other extrapyramidal effects, it does not always recover when antipsychotic drugs are stopped. It is usually characterized by chewing and sucking movements of the lips and jaw, but can involve limbs and occasionally the whole body. Although this syndrome is seen occasionally among patients who have not taken antipsychotic drugs, clinical observations suggest it is much more common in those who have received antipsychotic agents for extended periods of time. Estimates of the frequency of the syndrome vary in different series, but it seems to develop in 20–30% of patients with schizophrenia treated with long-term antipsychotic drugs (Jeste & Caligiuri, 1993). Whatever the exact incidence, the risk of this syndrome should be a deterrent to the long-term prescription of antipsychotic drugs unless clearly indicated.

Many treatments for tardive dyskinesia have been tried but none is universally effective. The antipsychotic drug should be stopped if the state of the psychiatric illness allows this. About 50% of cases may then remit. Where this in not possible, increasing the dose of antipsychotic drug may produce some suppression of movement symptoms but this relief is usually temporary. Sulpiride, however, may provide rather more sustained benefit with few adverse effects. Interestingly, treatment with vitamin E has been found useful in some studies (Jeste & Caligiuri, 1993).

6.06.5.3.2 Autonomic and endocrine effects

Several of the antipsychotic drugs, particularly phenothiazines, such as chlorpromazine

Table 4 Some unwanted effects of antipsychotic drugs.

Effect	*Symptoms*
Movement disorders	Acute dystonia, akathisia, parkinsonism, tardive dyskinesia
Autonomic and endocrine effects	Dry mouth, constipation, urinary hesitancy, blurred vision, sedation, postural hypotension, hypothermia, amenorrhoea, galactorrhoea, decreased libido
Other	Cardiac arrhythmias, weight gain, agranulocytosis (clozapine)

and thioridazine, are antagonists at muscarinic cholinergic receptors. This leads to dry mouth, urinary hesitancy and retention, constipation, reduced sweating, blurred vision, and rarely the precipitation of glaucoma.

Phenothiazines, and risperidone also have strong antagonist activity at α_1-adrenoceptors. This leads to sedation, hypotension and sexual dysfunction. Antipsychotic drugs that block histaminic and some subtypes of serotonin receptors can cause troublesome weight gain. This can be a particular problem with chlopromazine, thioridazine, clozapine, and olanzepine.

Blockade of D_2 receptors elevates plasma prolactin levels which can cause amenorrhoea and galactorrhoea in women and loss of libido in both sexes. The atypical antipsychotic drugs, with the exception of risperidone and sulpiride appear to have less effect on plasma prolactin.

6.06.5.3.3 Neuroleptic malignant syndrome

This rare but serious disorder occurs in a small minority of patients taking antipsychotic drugs, particularly high-potency compounds. Most reported cases have followed the use of antipsychotic drugs for schizophrenia, but in some cases the drugs were used for mania, depressive disorder, and psychosis secondary to a medical condition. The clinical picture includes the rapid onset, over 24–72 hours, of severe motor, cognitive, and autonomic disorders. The prominent motor symptom is generalized rigidity. The psychological symptoms include mutism, stupor, or fluctuating levels of consciousness. Hyperpyrexia develops with evidence of autonomic instability in the form of rapidly changing blood pressure, tachycardia, excessive sweating and urinary incontinence. Plasma levels of the enzyme, creatinine phosphokinase, are increased to very high levels. The neuroleptic malignant syndrome has a significant mortality, which may be declining but is probably still about 10%. The syndrome lasts for 1–2 weeks after stopping an oral neuroleptic, but may last 2–3 times longer after stopping long-acting preparations. Patients who survive usually make a complete recovery but residual movement disorders are sometimes seen (Addonizio & Susman, 1991).

The mainstay of treatment in neuroleptic malignant syndrome is symptomatic, with support in an intensive care unit if needed. No drug treatment is of definite utility. The dopamine receptor agonist, bromocriptine, is often tried, but there is no definite evidence of effectiveness. Patients who have developed neuroleptic malignant syndrome may need in the future to be treated with antipsychotic drugs. At least two weeks should elapse before antipsychotic treatment is reinstated, and it is prudent to start treatment cautiously with an oral low potency drug such as thioridazine.

6.06.5.3.4 Other adverse effects

A rare adverse effect, but one which is of particular concern, is the development of cardiac arrhythmias. Subclinical electrocardiogram (ECG) changes are not uncommon and usually take the form of prolongation of the QT interval. This appears to be more problematic with some drugs than others, for example, pimozide and sertindole, where the use of ECG monitoring has been recommended. It should also be noted that most antipsychotic drugs appear to lower the seizure threshold to some extent, so must be used with caution in patients with a tendency to seizure disorder.

Typical antipsychotic drugs have not been shown to be teratogenic, but nevertheless should be used cautiously in early pregnancy. There is presently insufficient data on the possible teratogenic effects of the atypical antipsychotic drugs.

6.06.5.3.5 Adverse effects of clozapine

As noted above, the use of clozapine requires special monitoring with regard to blood white cell count. For this reason both the hematological and other adverse effects of clozapine will be discussed here.

The use of clozapine is associated with a significant risk of leucopenia (about 2–3%) which can progress to agranulocytosis (Krupp & Barnes, 1992). Weekly blood counts for the first 18 weeks of treatment and at two week intervals thereafter are mandatory. With this intensive monitoring the early detection of leucopenia can be followed by immediate withdrawal of clozapine and reversal of the low white cell count. This procedure greatly reduces, but does not eliminate, the risk of progression to agranulocytosis. It is usually recommended that clozapine be used as the sole antipsychotic agent in a treatment regimen. Clearly, it is wise to avoid concomitant use of drugs such as carbamazepine, which may also lower the white cell count.

Because of its relatively weak blockade of dopamine D_2 receptors, clozapine is less likely than other antipsychotic drugs to cause extrapyramidal movement disorders. It does not increase plasma prolactin, hence galactorrhoea does not occur. However, its use is associated with hypersalivation, drowsiness, postural hypotension, weight gain, and hyperthermia. Seizures may occur at higher doses. Rarely, myocarditis has been reported.

6.06.5.4 Dosage of Antipsychotic Drugs

Doses of antipsychotic drugs need to be adjusted for the individual patient and changes should be made gradually. There is a growing trend for lower doses to be recommended. This is based in part on recent studies with positron emission tomography which have demonstrated that adequate dopamine D_2 receptor blockade can be obtained with low doses of conventional antipsychotic drugs, for example, 5 mg a day of haloperidol (Farde, Wiesel, & Nordstrom, 1989) (Table 5). Such doses produce an adequate antipsychotic effect in the majority of patients. Higher doses may cause further calming but are also likely to be associated with significant adverse effects, some of which may be serious, for example, cardiac arrhythmias. A view of growing influence is that the combination of modest doses of antipsychotic drugs with a benzodiazepine is a safer and more effective means of producing rapid sedation than high doses of antipsychotic drugs (Pilowsky, Ring, Shine, Battersby, & Lader, 1992).

The association of sudden unexplained death with antipsychotic drug treatment is a matter of continuing debate. For example, it is not established whether the rate of such deaths is greater in patients receiving antipsychotic drugs than in those receiving other treatments, or whether the rate in psychiatric patients is higher than in the general population. However, antipsychotic drugs are known to alter cardiac conduction (see above), and some drugs, such as chlorpromazine, also produce hypotension. Although the relationship between high doses of antipsychotic drug treatment and sudden death is not established, it is clearly prudent to use as low a dose of an antipsychotic drug as the clinical circumstances permit (Royal College of Psychiatrists, 1993).

6.06.6 ANTI-PARKINSONIAN DRUGS

Although these drugs have no direct therapeutic use in psychiatry, they are often required to control the extrapyramidal side effects of antipsychotic drugs, particulary acute dystonia and symptoms of Parkinsonism. They are of modest benefit in akathisia and actually worsen tardive dyskinesia. The pharmacological effect of these drugs is to block peripheral and central muscarinic cholinergic receptors.

6.06.6.1 Preparations Available

Many anticholinergic drugs are available and there is no reason for choosing one over the others for the treatment of neuroleptic induced movement disorders. Those most often used in psychiatric practice are benzhexol, benztropine, procyclidine, and orphenadrine. An injectable preparation of biperiden is useful for the treatment of acute dystonias.

Table 5 Daily dose of different antipsychotic drugs required to produce greater than 70% blockade of D_2 receptors in basal ganglia.

Drug	*Dose* (mg)
Chlorpromazine	200
Thioridazine	300
Trifluoperazine	10
Haloperidol	4
Flupenthixol	10
Sulpiride	800

Source: Farde et al. (1989).

6.06.6.2 Adverse Effects

In large doses these drugs may cause an acute organic syndrome, particularly in the elderly. Their anticholinergic activity can summate with those of antipsychotic drugs so that glaucoma or retention of urine in men with enlarged prostates may be precipitated. Drowsiness, dry mouth and constipation also occur. These effects tend to diminish as the drug is continued. Anticholinergic drugs can also exacerbate tardive dyskinesia but are probably not a predisposing factor in its development. Anticholinergic drugs can also be abused for their euphoriant and psychomimetic effects at high doses.

6.06.7 ANTIDEPRESSANTS

Antidepressant drugs are indicated in the treatment of major depression and dysthymia. Certain antidepressants are also effective in the treatment of anxiety disorders, and the eating disorder bulimia nervosa. Currently used antidepressant drugs can be divided into three main classes, depending on their acute pharmacological properties. The first class consists of compounds that inhibit the reuptake of noradrenaline and/or serotonin. This class includes the tricyclic antidepressants and the SSRIs. The second class consists of drugs that inhibit the enzyme MAO. The third class consists of drugs with complex effects on monoamine mechanisms, for example, mianserin and nefazodone, which cannot be easily categorized under the first two headings.

In the broad range of major depression, these drugs are of equivalent efficacy. The main distinctions between them are in their adverse effects, toxicity, and cost (Table 6). These three classes of drugs will be considered in turn after some comments on the possible mechanism of action of antidepressants.

6.06.7.1 Mechanism of Action

The primary pharmacological action of reuptake inhibitors and MAOIs can be detected within hours of the start of treatment, and yet the antidepressant effects of drug therapy can be delayed for a number of weeks. For example, it has been suggested that 4–6 weeks should elapse before an assessment of the effect of an antidepressant drug can be made in an individual patient (Depression Guideline Panel, 1993; Quitkin et al., 1996).

In animal experimental studies, the acute effect of antidepressants to facilitate noradrenaline and serotonin neurotransmission is followed by numerous secondary adaptive changes in noradrenaline and serotonin pathways. It is thought that it is these neuroadaptive changes that lead to the clinical antidepressant effect, probably by further enhancing the acute potentiation that antidepressant drugs produce on noradrenaline and serotonin neurotransmission (Blier & Montigny, 1994; Svensson & Usdin, 1978).

An important action observed in animal experimental studies following repeated administration of antidepressant drugs is a gradual desensitization of inhibitory autoreceptors on serotonin and noradrenaline cell bodies. These receptors normally have the effect of decreasing the firing of these cells, and hence their desensitization would be expected to add to the facilitation of neurotransmission produced by the antidepressant. Thus the clinical effects of antidepressant treatment may result from an increasing potentiation of noradrenaline and serotonin neurotransmission over time.

6.06.7.2 Tricyclic Antidepressants

6.06.7.2.1 Pharmacological properties

Tricyclic antidepressants have a three-ringed structure with an attached side chain. The first tricyclic to be introduced was imipramine, and there have been many modifications, so that a range of tricyclic compounds is now available. These can broadly be divided into tertiary and secondary amines, the distinction being that tertiary have a terminal methyl group on the side chain, whereas the secondary amines do not. In general, compared with the secondary amines, tertiary amines have a higher potency to block the serotonin reuptake site and are stronger antagonists of noradrenaline α_1-adrenoceptors and muscarinic cholinergic receptors. Therefore, in clinical use tertiary amines are more sedating and cause more anticholinergic side effects than secondary amines. In the UK, tertiary amines such as amitriptyline, dothiepin, and imipramine are popular in the treatment of depression, whereas in the USA secondary amines such as desipramine and nortriptyline are more commonly used.

Antidepressant drugs also antagonize histamine H_1 receptors, which can cause drowsiness and weight gain. Tricyclics also have quinidine-like membrane stabilizing effects. This may explain why they impair cardiac conduction and cause high toxicity in overdose.

6.06.7.2.2 Adverse effects of tricyclic antidepressants

The adverse effects of tricyclic antidepressants can, in general, be derived from their receptor blocking properties (Table 7). As mentioned above, anticholinergic and antiadrenergic effects are common. In addition, the drugs can cause tiredness and drowsiness, although secondary amines, such as desipramine, can cause insomnia. Fine tremor and muscle twitching have been observed and like

Table 6 Side effect profiles of some antidepressant drugs.

Drug	*Anticholinergic*	*Sedation*	*Insomnia*	*Cardiotoxic*	*Nausea/vomiting*
Amitriptyline	+ + +	+ + +	0	+ + +	0
Desipramine	+ +	+	+	+ + +	0
Lofepramine	+ +	+	+	0	0
Trazodone	0	+ + +	0	+	+ +
Nefazodone	0	+	0	0	+ +
Mirtazapine	0	+ + +	0	0	0
SSRIs	0	0/+	+ +	0	+ + +
SNRI	0	0/+	+ +	0	+ + +

0 = not present; + + + = strong.

antipsychotic drugs, tricyclic antidepressants lower the seizure threshold, which means they must be used with caution in people predisposed to seizure disorders. Allergic skin rashes, cholestatic jaundice, and agranulocytosis, are seen rarely. Weight gain and sexual dysfunction are more common.

Sudden withdrawal of tricyclics can produce an abstinence syndrome, characterized by nausea, anxiety, sweating, and insomnia. Current evidence does not suggest that tricyclics such as amitriptyline and imipramine are important human teratogens. There is less data about other tricyclic antidepressants.

In overdose, tricyclic antidepressants produce a large number of effects, of which some are extremely serious. Therefore urgent expert treatment in a general hospital is required. The main danger comes from cardiovascular effects, which include cardiac arrhythmias with ventricular fibrillation. In addition, respiratory depression can occur, and the resulting hypoxia increases the likelihood of cardiac complications. Tricyclics delay gastric emptying, and so gastric lavage is valuable for several hours after the overdose.

6.06.7.2.3 Amoxapine

Some tricyclics have pharmacological properties sufficiently distinct to be worth separate mention. Amoxapine is a fairly selective inhibitor of noradrenaline uptake but, unusually for a tricyclic antidepressant, produces significant blockade of dopamine D_2 receptors. The combined effect of amoxapine to increase noradrenaline neurotransmission and antagonize D_2 receptors has led to suggestions that this compound may be particularly useful in the treatment of depressive psychosis when combined treatment with antidepressant and antipsychotic drugs is often required. However, the use of a single preparation to produce combined pharmacological effect limits prescribing flexibility. Furthermore, as might be expected, the D_2 receptor blocking properties of amoxapine may result in extrapyramidal disorders (Rudorfer & Potter, 1980).

6.06.7.2.4 Clomipramine

Clomipramine is the most potent of the tricyclic antidepressants in inhibiting the reuptake of serotonin. Probably because of this, unlike other tricyclic antidepressants, clomipramine is useful in treating the symptoms of obsessive compulsive disorder whether or not there is a coexisting major depression (Jenike, 1992). Clomipramine is also available as an intravenous infusion, but in general this form of administration does not appear to produce better therapeutic effects than the oral route.

6.06.7.2.5 Lofepramine

Lofepramine is a fairly selective inhibitor of noradrenaline reuptake, and has fewer anticholinergic and antihistaminic properties than amitriptyline. It has been widely compared with other tricyclic antidepressants, and in general its antidepressant efficacy is equivalent. Lofepramine is not sedating; early in treatment it can be experienced as activating, an effect which some depressed patients find unpleasant. Similarly, impaired sleep does not usually improve until the underlying depression remits. The most important feature of lofepramine is that in overdose it is not cardiotoxic and it is therefore much safer than conventional tricyclic antidepressants. Therefore lofepramine is likely to be safer than other tricyclics for patients with cardiovascular disease, although caution is still required (Lancaster & Gonzalez, 1990).

Table 7 Some adverse effects of tricyclic antidepressants.

Pharmacological action	*Adverse effect*
Anticholinergic	Dry mouth, tachycardia, blurred vision, glaucoma, constipation, urinary retention, cognitive impairment
Antiadrenergic	Drowsiness, postural hypotension, sexual dysfunction
Histamine H_1 receptor blockade	drowsiness, weight gain
Other	Cardiac conduction defects, cardiac arrhythmias, epileptic seizures, (all common in overdose), rash, oedema, sweating, low white cell count (rare)

6.06.7.2.6 Maprotiline

Maprotiline is often referred to as a quadricyclic antidepressant because the tricyclic nucleus is supplemented by an ethylene bridge across the middle ring. It is the most selective noradrenaline uptake inhibitor of the tricyclic antidepressants currently available, but has moderate antihistaminic properties and rather less anticholinergic effects than imipramine. It appears as effective as reference tricyclics. The use of maprotiline at doses above 200 mg daily have been associated with a higher incidence of seizures than is usual during tricyclic treatment. Therefore a dose range of 75–150 mg daily has been recommended, and the coprescription of other drugs that may lower the seizure threshold, such as phenothiazines, should be approached with caution. Maprotiline has effects on the heart that are similar to those of conventional tricyclics, and in overdose it is at least as toxic (Rudorfer & Potter, 1989).

6.06.7.3 Selective Serotonin Reuptake Inhibitors

6.06.7.3.1 Pharmacological properties

Five SSRIs, citalopram, fluoxetine, fluvoxamine, paroxetine, and sertraline, are available at present for clinical use in the UK. SSRIs are a structurally diverse group, but they all inhibit the reuptake of serotonin with a high potency and selectivity. None of them has an appreciable affinity for the noradrenaline uptake site, and present data suggest that they have a low affinity for other monoamine neurotransmitter receptors (Stahl, 1996).

6.06.7.3.2 Efficacy in depression

The SSRIs have been extensively compared with placebo and with reference tricyclic antidepressants. The SSRIs are all clearly superior to placebo and are generally as effective as tricyclics antidepressants in the treatment of major depression. Most comparative studies have been of moderately depressed outpatients, and there has been concern that SSRIs may be less effective than conventional tricyclic antidepressants for more severely depressed patients. This is based to some extent on the work of the Danish University Antidepressant Group (1990), who found that the tricyclic clomipramine is significantly more effective than either paroxetine or citalopram in depressed inpatients. In addition, more recent meta-analyses have suggested that SSRIs may be slightly less effective than drugs that potently block the reuptake of both serotonin and noradrenaline. This categorization includes clomipramine and the new antidepressant, venlafaxine (Anderson, 1997) (see below).

6.06.7.3.3 Unwanted effects of SSRIs

The SSRIs have a different side effect profile to tricyclic antidepressants and are somewhat better tolerated at therapeutic dosage (Table 8). In meta-analyses against conventional tricyclic antidepressants the drop-out rate due to adverse effects with SSRIs is about 25% less. Whether this confers significant cost benefit is controversial.

The SSRIs are less cardiotoxic than tricyclic antidepressants and are safer in overdose. They also lack anticholinergic effects and are less sedating. Their adverse effect profile consists mainly of gastrointestinal symptoms and central nervous system effects. Sexual dysfunction is also common, occurring in up to about 25% of people. A rare but problematic side effect, more frequent in the elderly, is a low sodium state. They can rarely cause movement disorders. Discontinuation of paroxetine and sertraline

Table 8 Some side effects of SSRIs.

Effect	*Symptoms*
Gastrointestinal	Common: nausea, appetite loss, dry mouth, diarrhea, constipation, dyspepsia Uncommon: vomiting, weight loss
Central nervous system	Common: headache, insomnia, dizziness, anxiety, fatigue, tremor, somnolence Uncommon: extrapyramidal reaction, seizures
Other	Common: sweating, delayed orgasm, anorgasmia Uncommon: rash, pharyngitis, dyspnoea, serum sickness, hyponatremia, alopecia

has been associated with a withdrawal syndrome consisting of nausea, irritability, impaired sleep and ataxia (Cowen, 1996).

It is not clear whether SSRIs are human teratogens. Fluoxetine may increase rates of premature delivery and perhaps minor fetal abnormalties, but this is controversial (Chambers, Johnson, & Dick, 1996; Robert, 1996).

6.06.7.4 Monoamine Oxidase Inhibitors

MAOIs were introduced just before the tricyclic antidepressants, but their use has been less widespread because of both troublesome interactions with foods and drugs and uncertainty about their therapeutic efficacy. Recent controlled studies have shown that in adequate doses MAOIs are useful antidepressants, often producing clinical benefit in depressed patients who have not responded to other medications or electroconvulsive therapy (ECT). In addition, MAOIs can be useful in refractory anxiety states (Nutt & Glue, 1989; Paykel, 1990). These beneficial effects have to be weighed against the need to adhere to strict dietary and drug restrictions in order to avoid reactions with tyramine and other sympathomimetic agents. In practice this means that MAOIs are very rarely used as first-line treatment. It remains to be seen whether this approach will be altered by the recent availability of MAOIs, such as moclobemide, that do not potentiate tyramine.

6.06.7.4.1 Pharmacology

The MAOIs inactivate enzymes that oxidize noradrenaline, serotonin, tyramine, and other amines that are widely distributed in the body as transmitters, or are taken in food and drink or as drugs. Monoamine oxidase (MAO) exists in a number of forms that differ in their substrate and inhibitor specificities. From the point of view of psychotropic drug treatment it is important to recognize that there are two forms of MAO (type A and type B), encoded by separate genes. In general, MAO-A metabolizes intraneuronal noradrenaline and serotonin, whereas both MAO-A and MAO-B metabolize dopamine and tyramine.

6.06.7.4.2 Compounds available

Phenelzine is the most widely used and widely studied compound. Isocarboxazid is reported to have fewer side effects than phenelzine and can be useful for patients who respond to the latter drug but suffer from its side effects of hypotension or sleep disorder. Tranylcypromine differs from the other compounds in combining the ability to inhibit MAO with an amphetamine-like stimulating effect. There are more reports of adverse drug and food reactions with tranylcypromine than other MAOIs, so it should be prescribed with particular caution.

Moclobemide is the most recently developed MAOI to be marketed. It differs from the other compounds in selectively binding to MAO-A, which it inhibits in a reversible way. This results in a lack of significant interactions with foodstuffs and a quick offset of action (see below).

6.06.7.4.3 Efficacy of MAOIs in depression

For many years MAOIs were in relative disuse because several studies, in particular a large controlled trial by the Medical Research Council, found them no better than placebo in the treatment of depressive disorders (Clinical Psychiatry Committee, 1965). It seems likely that the doses of MAOIs were too low in these early investigations; in the Medical Research Council study the maximum dose of phenelzine was 45 mg daily as against the current practice of doses up to 90 mg daily if side effects permit. In this wider dose range MAOIs are superior to placebo and generally equivalent to tricyclic antidepressants in their therapeutic activity (Paykel, 1990).

6.06.7.4.4 Unwanted effects

MAOIs have numerous unwanted effects (Table 9). In clinical practice the main problems are insomnia, weight gain, and postural hypotension. Ankle edema is also not uncommon. Phenelzine and isocarboxazid have been associated with hepatocellular jaundice. There are little data on the possible teratogenicity of MAOIs in humans.

6.06.7.4.5 Interactions with foodstuffs and drugs

Some foods contain tyramine, a substance that is normally inactivated by MAO in the liver and gut wall. When MAO is inhibited, tyramine is not broken down and is free to exert its hypertensive effects. These effects are due to release of noradrenaline with a consequent elevation in blood pressure. This may reach dangerous levels and may occasionally result in cerebrovascular accident. Important early symptoms of such a crisis include a severe and usually throbbing headache.

There have been reports of many foods being implicated in hypertensive reactions with MAOIs, but many of these have cited single cases and hence are of uncertain validity. Another complication is that the tyramine content of a particular food item may vary, as

Table 9 Adverse effects of MAOIs.

Effect	*Symptoms*
Central nervous system	Insomnia, drowsiness, agitation, headache, fatigue, weakness, tremor, mania, confusion
Autonomic	Blurred vision, difficulty passing urine, sweating, dry mouth, postural hypotension, constipation
Other	Sexual dysfunction, weight gain, peripheral neuropathy (rare), edema, rashes, hepatocellular toxicity (rare), leucopenia (rare)

may the susceptibility of an individual patient to a hypertensive reaction. If a forbidden food has been consumed on one occasion without adverse effects, this does not preclude a future reaction. It has been concluded that the following foods and drinks should be avoided (Davidson, 1992):

(i) all cheeses except cream, cottage, and ricotta cheeses;

(ii) red wine, sherry, beer, and liquors;

(iii) pickled or smoked fish;

(iv) brewer's yeast products, for example, Marmite, Bovril, and some packet soups;

(v) broad bean pods;

(vi) beef or chicken liver;

(vii) fermented sausage, for example, pepperoni, salami; and

(viii) unfresh, overripe, or aged food, for example, pheasant, venison, unfresh dairy products.

Despite this list, case reports suggest that cheese is the food most often incriminated in serious adverse reactions. Hypertensive reactions should be treated with parenteral administration of an α_1-adrenoceptor antagonist, such as phentolamine. Chlorpromazine can be used if the latter is not available. Oral nifedipine may also be useful. Whatever treatment is given, blood pressure must be monitored carefully.

A number of drugs cause serious interactions with MAOIs. In particular, drugs that increase brain serotonin function such as the SSRIs and clomipramine can cause a fatal neurotoxicity syndrome (Sternbach, 1991). In general medicine coadministration of opiate analgesics, particularly pethidine, can produce a similar effect. Finally, drugs that potentiate noradrenaline can cause serious hypertensive reactions. The importance of this is that such compounds are often present in "over-the-counter" cold and flu remedies. Patients receiving MAOI treatment must be warned not to take any other medication until its safety with MAOIs has been specifically checked.

6.06.7.4.6 Moclobemide

In its freedom from tyramine reactions, the reversible type A MAOI, moclobemide, has a clear advantage over conventional MAOIs (Simpson & De Leon, 1989). As with all new antidepressants, however, the therapeutic efficacy of moclobemide, particularly in severely ill patients is not as well established as that of phenelzine or tranylcypromine. Moreover, a recent meta-analysis suggested that it might be somewhat less effective than clomipramine and imipramine in depressed inpatients (Angst, Amrein, & Stabl, 1995).

Moclobemide is better tolerated than conventional MAOIs but can cause insomnia and nausea. Caution still has to be exercised when moclobemide is coprescribed with other drugs. It is recommended that it should not be prescribed with SSRIs or clomipramine because of the risk of serotonin syndrome (see above). As for the irreversible MAOIs, moclobemide may react adversely with opiates and noradrenaline potentiating drugs.

6.06.7.5 Other Antidepressant Drugs

Under this heading are discussed drugs whose mechanism of action cannot be easily grouped with tricyclic antidepressants, SSRIs, or MAOIs. These drugs also have differing adverse-effect profiles. They are therefore discussed individually below.

6.06.7.5.1 Mianserin

Mianserin is a quadricyclic compound with complex pharmacological actions. It has weak noradrenaline reuptake inhibiting effects, and is a fairly potent antagonist at a number of neurotransmitter receptors, including serotonin$_2$, serotonin$_3$, and noradrenergic α_1- and α_2-adrenoceptors. It is not a muscarinic cholinergic antagonist and is not cardiotoxic.

Controlled trials have shown that mianserin is superior to placebo in the management of depression, and comparative studies against imipramine and clomipramine have shown no difference in effect. These studies are difficult to assess because of the wide range of doses that have been used. Many early studies of mianserin used doses of 30–60 mg daily, whereas much higher doses of up to 200 mg daily have sometimes been advocated for inpatients (Montgomery, Bullock, & Pinder, 1991).

The main adverse effects of mianserin are drowsiness and dizziness. Significant cognitive impairment is more likely with mianserin than with SSRIs, and weight gain is a common problem. The most serious adverse effect of mianserin is a lowering of the white cell count, and fatal agranulocytosis has been reported. It is recommended that a blood count be obtained before starting mianserin treatment, and that the white cell count be monitored monthly for three months after treatment has started.

6.06.7.5.2 Mirtazapine

Mirtazapine is a new antidepressant which is structurally related to mianserin. Its pharmacological properties are similar but its noradrenergic α_1-adrenoceptor blockade is less potent which means that in practice it is a little less sedating. Placebo-controlled trials have shown an efficacy in the treatment of major depression, although the relative efficacy of mirtazapine against other antidepressants is not fully clear (Bruijn et al., 1996; Davies & Wilde, 1996).

Mirtazepine is generally well tolerated, with the most common side effects being drowsiness, dizziness, and weight gain.

6.06.7.5.3 Trazodone

Trazodone is a triazolopyridine derivative with complex actions on serotonin pathways. Studies *in vitro* suggest that trazodone has some weak serotonin reuptake inhibiting properties but these are probably not manifest during clinical use. Trazodone is an antagonist at serotonin$_2$ receptors but its active metabolite, *m*-chlorophenylpiperazine (*m*CPP), is a serotonin receptor agonist. Trazodone also blocks postsynaptic α_1-adrenoceptors, which gives it a distinct sedating profile.

Several placebo-controlled studies have shown that trazodone in doses of 150–600 mg is superior to placebo in the treatment of depressed patients. Trazodone also appears to have equivalent antidepressant activity to compounds such as imipramine. Many of these studies were carried out in moderately depressed outpatients, and the efficacy of trazodone in depressed inpatients is not as well established (Rudorfer & Potter, 1998)

The major unwanted effect of trazodone is excessive sedation, which can result in significant cognitive impairment. Nausea and dizziness are also reported, particularly if the drug is taken on an empty stomach. The α_1-adrenoceptor antagonist properties of trazodone may lower blood pressure to some extent, and postural hypotension has been reported. Trazodone is less cardiotoxic than conventional tricyclics, but there are reports that cardiac arrhythmias may be worsened in patients with cardiac disease. Nevertheless, trazodone is much less toxic in overdose than tricyclic antidepressants. A serious side effect of trazodone is priapism. This reaction is seen rarely (about 1 in 6000 male patients). It can cause considerable problems, requiring the local injection of noradrenaline agonists or even surgical treatment (Rudorfer & Potter, 1989).

6.06.7.5.4 Nefazodone

Nefazodone is related to trazodone but lacks α_1-adrenoceptor antagonist properties and is therefore less sedating. Like trazodone it has mild serotonin reuptake blocking properties, and is metabolized to the serotonin receptor agonist *m*CPP.

Controlled trials in patients with major depression have shown that in doses of 400 mg and greater, nefazodone is more effective than placebo and generally equal in therapeutic activity to comparator drugs (Rickels, Schweizer, Clary, Fox, & Weise, 1994). As with trazodone, these studies have focused on outpatients with moderate depressive disorders. Nefazodone is usually given in two divided doses starting at 200 mg daily with titration to 400 mg daily after about a week. The maximum dose is 600 mg.

Nefazodone is generally well tolerated with the most common side effects being headache, loss of energy, dizziness, dry mouth, nausea, and somnolence. It appears less cardiotoxic than tricyclic antidepressants and is less likely than SSRIs to cause insomnia and sexual dysfunction (Robinson et al., 1996).

6.06.7.5.5 Venlafaxine

Venlafaxine is a phenylethylamine derivative which produces a potent blockade of both serotonin and noradrenaline reuptake. In this respect the pharmacological properties of venlafaxine resemble those of clomipramine. However, unlike clomipramine and other tricyclic antidepressants, venlafaxine has a negligible affinity for other neurotransmitter receptor sites and so lacks sedative and anticholinergic effects. Venlafaxine is classified as a selective serotonin and noradrenaline reuptake inhibitor (SNRI). Venlafaxine has been studied in both inpatients and outpatients with major depression and compared with placebo and active comparators. Current studies suggest that it is more effective than placebo and at least of equal efficacy to other available antidepressant drugs. Venlafaxine also appears to be effective in depressed inpatients, perhaps more so than fluoxetine (Clerc, Ruimy, & Verdeau-Pailles, 1994; Feighner, 1994).

Venlafaxine has a wider dosage range than SSRIs, from 75 mg to 375 mg daily, in two divided doses. Higher doses are associated with a greater incidence of adverse effects. The usual starting dose of venlafaxine is 75 mg daily which may be sufficient for many patients. Upward titration can be considered where there is insufficient response, or if a faster onset of therapeutic activity is needed.

The adverse effect profile of venlafaxine resembles that of SSRIs, with the most common adverse effects being nausea, headache, insomnia, and sexual dysfunction. Venlafaxine occasionally causes postural hypotension but, in addition, dose-related increases in blood pressure can occur. Blood pressure monitoring may be advisable in patients receiving more than 200 mg of venlafaxine daily. Sudden discontinuation of venlafaxine has been associated with symptoms of fatigue, nausea, and dizziness. It is recommended that the dose should be reduced gradually over at least a one week period. Preliminary evidence suggests that venlafaxine is less toxic in overdose than tricyclic antidepressants (Feighner, 1994).

6.06.7.5.6 Bupropion

Bupropion is marketed for the treatment of depression in the USA but not in Europe. It is a unicyclic compound whose pharmacological properties are not well characterized. It may, however, have some activity as a dopamine and noradrenaline reuptake inhibitor (Ascher et al., 1995). Bupropion has activating properties, and early in treatment can cause restlessness and insomnia. However, it does not cause significant sexual dysfunction. Bupropion is associated with an increased risk of seizures, particularly where the dose exceeds 450 mg daily (Rudorfer & Potter, 1989).

6.06.7.5.7 L-Tryptophan

L-Tryptophan is a naturally occurring amino acid, present in the normal diet. About 500 mg of tryptophan is consumed daily in the typical Western diet. Most ingested tryptophan is used for protein synthesis and the formation of nicotinamide nucleotides; only a small proportion (about 1%) is synthesized to serotonin via 5-hydroxytryptophan. There is only weak evidence that L-tryptophan has antidepressant activity, although it may be superior to placebo in moderately depressed outpatients. There is rather better evidence that L-tryptophan combined with MAOI treatment can enhance the antidepressant effects of MAOIs. Similar synergistic effects have been reported in some studies of L-tryptophan combined with tricyclics, although overall the therapeutic effect of this combination seems inconsistent (Chalmers & Cowen, 1990).

L-Tryptophan is generally well tolerated, although nausea and drowsiness soon after dosing are not unusual. In the early 1990s, the prescription of L-tryptophan began to be associated with the development of a severe scleroderma-like illness, the eosinophilia-myalgic syndrome, in which there is a very high circulating eosinophil count (about 20% of peripheral leucocytes). Associated symptoms were severe muscle pain, edema, skin sclerosis and peripheral neuropathy. Some fatalities occurred. It is now reasonably well established that EMS is not caused by L-tryptophan itself but rather by a contaminant formed in the manufacturing process used by a particular manufacturer (Kilbourne, Philen, Kamb, & Falk, 1996). L-Tryptophan remains available for the treatment of severe refractory depression when it is used as an adjunct to other antidepressant medication. Patients receiving L-tryptophan require close supervision, including monitoring for possible symptoms of EMS and regular blood eosinophil counts. L-Tryptophan should be withdrawn if there is any evidence that EMS may be developing.

6.06.8 MOOD-STABILIZING DRUGS

Under this heading are grouped three agents, lithium, carbamazepine, and sodium valproate. These three drugs have efficacy in the prevention of recurrent mood disorders and also in the acute treatment of mania. Lithium also has useful antidepressant effects in some circumstances (Price, 1989), but the antidepressant activity of carbamazepine and sodium valproate is less well established. Lithium has also been shown to lower the frequency of aggressive behavior in patients with learning difficulties (Nilsson, 1993).

6.06.8.1 Lithium

6.06.8.1.1 Pharmacology

The mode of action of lithium is uncertain. Lithium does not affect neurotransmitters or their receptors directly, but appears to have important effects on intracellular signaling molecules, or second messengers, that are activated when a neurotransmitter or agonist binds to a specific receptor. Through these actions lithium could exert profound effects on a wide range of neurotransmitter pathways. It has been proposed that the effects of lithium may be particularly apparent when the turnover and recycling of second messengers is increased, and accordingly lithium may act preferentially to inhibit overactive neurotransmitter systems (Lithium Mechanisms Study Group 1993).

6.06.8.1.2 Efficacy

It is estimated that about 50% of patients with bipolar disorder will have a good prophylactic response to lithium. Some patients who respond well have a complete cessation of mood swings, whereas others experience markedly dimished symptomatology in which subclinical mood swings can still be discerned (Goodwin & Jamison, 1990). Lithium is also effective in the acute treatment of mania but may need to be supplemented with antipsychotic drugs if psychotic symptoms are present (Chou, 1991). About 50% of patients with a depressive syndrome unresponsive to antidepressant drugs will show a clinical response if lithium is added to their drug treatment (Price, 1989). Lithium alone has some acute antidepressant activity, most apparent in patients with an underlying bipolar disorder (Goodwin & Jamison, 1990).

6.06.8.1.3 Adverse effects

Common side effects include tremor of the hands, dry mouth, a metallic taste, feelings of muscular weakness, and fatigue (Table 10). Thirst and increased urine volume are also common. Most patients taking lithium have some minimal defect of renal tubular concentrating ability, but this rarely leads to clinical problems. A few patients, however, develop a diabetes insipidus-like syndrome with frequent passage of large volumes of water.

Weight gain during lithium treatment is quite common, and partial hair loss has sometimes been reported. Thyroid gland enlargement occurs in about 5% of patients taking lithium. Lithium also interferes with thyroid production, and hypothyroidism occurs in up to 20% of patients. Tests of thyroid function should be performed every six months to help detect these changes, but patients should also be monitored clinically for signs of hypothyroidism, particularly lethargy and substantial weight gain. If hypothyroidism develops and the reasons for lithium treatment are still strong, thyroxine treatment should be added. Lithium has rarely been associated with elevated serum calcium levels in the context of hyperparathyroidism.

Lithium is also associated with reversible ECG changes that do not seem to be of particular consequence for cardiac conduction. Other changes include an elevated white count and occasional rashes.

Most concern around the adverse effects of lithium is centred on possible long-term renal damage. As mentioned above, most patients taking lithium have some mild impairment of renal tubular concentrating ability. However, this usually recovers when the drug is stopped, although there are some reports of persisting cases. There have also been reports of tubular damage in patients taking prolonged lithium treatment.

Several follow-up studies have examined the effect of longer-term lithium maintenance treatment on glomerular function. It has been concluded that long-term lithium treatment, in the absence of toxic blood levels, does not result in a lowering of renal filtering ability (glomerular filtration rate) (Gelder et al., 1996). However, although lithium may not significantly lower the mean glomerular filtration rate in groups of patients with bipolar illness, there are case reports of increases in plasma creatinine in lithium-treated subjects when other causes of nephrotoxicity appear to be absent. Whether lithium treatment, in the absence of toxic blood levels, can cause frank renal failure is unclear (Gelder et al., 1996). With the current trends towards long-term prophylaxis of mood disorders, it is clearly wise to monitor biochemical measures of renal function regularly. It seems likely that the risk of nephrotoxicity will be minimized by avoiding toxic blood levels and maintaining plasma lithium levels at the lower end of the therapeutic range.

Effects on memory are sometimes reported by patients who complain of every-day lapses of memory, such as forgetting well-known names. Although this impairment may be associated with the mood disorder rather than lithium, there is evidence that lithium can impair certain cognitive tasks in healthy volunteers (Glue, Nutt, Cowen, & Broadbent, 1987).

Sudden discontinuation of lithium in patients with bipolar illness can result in the rapid development of mania in up to 50% of subjects. This is thought to represent a rebound phenomenon (Goodwin, 1994). Rates of relapse

Table 10 Some adverse effects of lithium and carbamazepine.

Effect	*Lithium*	*Carbamazepine*
Neurological	Tremor, weakness, dysarthria, ataxia, impaired memory, seizures (rare)	Dizziness, weakness, drowsiness, ataxia, headache, visual disturbance
Renal/fluid balance	Increased urine output with decreased urine-concentrating ability; thirst, diabetes insipidus (rare), edema	Low sodium states, edema
Gastrointestinal/hepatic	Altered taste, anorexia, nausea, vomiting, diarrhea, weight gain	Anorexia, nausea, constipation, hepatocellular damage
Endocrine	Decreased thyroxine with increase TSH,[a] goitre, hyperparathyroidism (rare)	Decreased thyroxine with normal TSH[a]
Hematological	Leucocytosis	Leucopenia, agranulocytosis (rare)
Dermatological	Acne, exacerbation of psoriasis	Rashes
Cardiovascular	ECG changes (usually clinically benign)	Cardiac conduction disturbances

[a] TSH = thyroid-stimulating hormone.

are significantly less when lithium is stopped gradually over several weeks (Baldessarini, Tondo, Floris, & Rudas, 1997).

6.06.8.1.4 *Toxic effects*

Toxic effects of lithium are related to dose. Because therapeutic blood levels (0.5–1.0 mmol/1) are close to levels at which toxicity may be experienced (> 1.5 mmol/l) it is important for both patient and clinician to be aware of symptoms of toxicity. They include ataxia, poor coordination of limb movements, muscle twitching, slurred speech, and confusion. Such symptoms constitute a serious medical emergency for they can progress through coma and fits to death. If these symptoms appear, lithium must be stopped at once and a high intake of fluid provided. In severe cases renal dialysis may be needed. There have been reports of permanent neurological damage in patients who have suffered from lithium toxicity. It is important to note that certain commonly used medical drugs, such as thiazide diuretics and nonsteroidal antiinflammatory drugs, can elevate lithium levels and cause toxic effects (Gelder et al., 1996).

6.06.8.1.5 *Lithium and pregnancy*

Lithium crosses the placenta, and retrospective studies have found increased rates of abnormalities in the babies of mothers receiving lithium in pregnancy. For example, a rate of 7% has been reported, with most abnormalities affecting the baby's heart. However, a prospective study of 148 women found no increase in congenital malformation in patients exposed to lithium in the first trimester of pregnancy compared with matched controls. These authors concluded that lithium did not appear to be an important human teratogen (Jacobson et al., 1992). However, this conclusion was based on relatively few patients.

Clearly, it is desirable for patients to be medication-free during the first trimester of pregnancy, and the decision whether or not to continue with lithium treatment must be carefully weighed. Important factors include the likelihood of affective relapse if lithium is withheld. If pregnant patients continue with lithium, plasma levels should be monitored closely. Ultrasound examination and fetal echocardiography are valuable screening tests as the pregnancy progresses.

Patients with a history of bipolar disorder have a substantially increased risk of psychotic relapse in the postpartum period. In such patients it may be worth considering the introduction of lithium shortly after delivery to provide a prophylactic effect. However, it should be noted that lithium is secreted into breast milk and that significant concentrations of lithium can be measured in the plasma of breast-fed infants.

6.06.8.2 Carbamazepine

Carbamazepine was originally introduced as an anticonvulsant and was found to have useful effects on mood in certain patients. Subsequently it was found to be beneficial in many bipolar patients, including those who had proved refractory to lithium. The acute antidepressant effect of carbamazepine is not established (Post, 1991).

6.06.8.2.1 Pharmacology

Like certain other anticonvulsants, carbamazepine blocks neuronal sodium channels. It is unclear whether this action plays a role in its mood-stabilizing effects. Like lithium, carbamazepine facilitates some aspects of brain serotonin function.

6.06.8.2.2 Efficacy

In the treatment of acute mania, carbamazepine is of about equal efficay to lithium. In the prophylaxis of bipolar illness, carbamazepine is also about as effective as lithium, although the quality of the trials has been criticized (Dardennes, Even, Bange, & Heim, 1995). Based on a survey of controlled and uncontrolled trials (Post, Denicoff, Frye, & Leverich, 1997) it is estimated that about 65% of patients with bipolar illness show a clinically significant prophylactic response to carbamazepine either given alone or added to lithium. However, there is some evidence that in some patients an initial response can diminish with time suggesting that tolerance to the mood-stabilizing effects of lithium can occur (Post, Leverich, Rosoff, & Altschuler, 1990).

6.06.8.2.3 Adverse effects

Adverse effects with carbamazepine are common at the beginning of treatment (Table 10). They include drowsiness, dizziness, ataxia, diplopia, and nausea. Tolerance to these effects usually develops quickly. A potentially serious side effect of carbamazepine is agranulocytosis, although this complication is very rare (variously estimated from 1 in 10 000 to 1 in 125 000 patients). A relative leucopenia is more common, with the white cell count falling in the first few weeks of treatment, though usually remaining within normal levels. Rashes occur in about 5% of patients. Elevations in liver enzymes may occur and, rarely, hepatitis has been reported. Carbamazepine can also cause significant disturbances of cardiac conduction, and therefore is contraindicated in patients with preexisting abnormalities of cardiac rhythm. Carbamazepine has also been associated with low sodium states.

The use of carbamazepine in pregnancy has been associated with neural tube defects in the fetus. The risk is diminished by adequate folate intake.

6.06.8.3 Sodium Valproate

Like carbamazepine, sodium valproate was first introduced as an anticonvulsant. In recent years there has been increasing interest in using the drug in the management of mood disorders, particularly acute mania and the prophylaxis of bipolar disorder in patients unresponsive to lithium and carbamazepine. There is presently little evidence that valproate has acute antidepressant effects (McElroy, Kerk, & Pope, 1987; Post, 1991).

6.06.8.3.1 Pharmacology

Valproate is a simple branch-chain fatty acid with a mode of action that is unclear. However, there is some evidence that it can slow the breakdown of the inhibitory neurotransmitter GABA. This action could account for the anticonvulsant properties of valproate, but whether it also underlies the psychotropic effects is unclear.

6.06.8.3.2 Efficacy

There have been several controlled studies suggesting that valproate is effective in the acute management of mania, but there are no controlled investigations of its efficacy in the prophylaxis of bipolar disorder (McElroy, Kerk, Pope, & Hudson, 1992). There have, however, been numerous case series and open studies that have reported useful prophylactic effects of valproate in patients unresponsive to lithium and carbamazepine, including those with rapid cycling mood disorders (McElroy et al., 1987, 1992; Post, 1991).

6.06.8.3.3 Adverse effects

Common side effects of valproate include gastrointestinal disturbances, tremor, sedation, and tiredness. Other troublesome side effects include weight gain and transient hair loss with changes in texture on regrowth. Patients taking valproate may have some elevation in hepatic transaminase enzymes; provided this increase is not associated with hepatic dysfunction the drug can be continued while enzyme levels and liver function are carefully monitored. However, there have been several reports of fatal hepatic toxicity associated with valproate; thus far these reports have been confined to children taking multiple anticonvulsant drugs. Valproate must be withdrawn immediately if vomiting, anorexia, jaundice or sudden drowsiness occur. Valproate may also cause thrombocytopenia and may inhibit platelet aggregation. Acute pancreatitis is another rare but serious side effect, and increases in plasma ammonia have been reported. Other possible side effects include edema, amenorrhoea and rashes. The use of valproate in pregnancy has been associated with neural tube defects and bleeding in the neonate.

6.06.9 CLINICAL USE OF PSYCHOTROPIC DRUGS

6.06.9.1 Anxiety Disorders

6.06.9.1.1 Generalized anxiety disorder

In generalized anxiety disorder (GAD) drugs are generally used as an adjunct to psychological methods of treatment (Table 11). Benzodiazepines are now prescribed only for short-term use (2–4 weeks) in patients where anxiety is causing severe distress and functional impairment. Benzodiazepines have a number of advantages as short-term treatment in that they are rapidly effective and have a wide safety margin. It is usually best to give benzodiazepine treatment on an as-required basis in doses of 2.5–5 mg. Intermittent use of this nature is less likely to result in tolerance (Tyrer, 1997).

Tricyclic antidepressant drugs and trazodone have also been shown to be effective in GAD in the same doses that are effective in treating major depression (Rickels et al., 1993). SSRIs have not been formally tested in this disorder but clinical impression suggests that they are likely to be useful. The azapirone buspirone is also effective in GAD, particularly in patients who have not received significant prior benzodiazepine treatment. All these treatments take longer to work than benzodiazepines. Their ultimate effect over 6–8 weeks, however, is at least as great (Cowen, 1997).

Where psychological treatments have not helped or are not available, buspirone or antidepressant drugs rather than benzodiazepines should now be used for longer-term treatment of GAD. The choice will lie between a sedating compound, such as a tricyclic antidepressant, or nonsedating treatments such as buspirone or an SSRI. Although, sedation has some advantages for patients with sleep disturbance and agitation, in the longer-term, buspirone or SSRIs are likely to be as effective and have less risk of cognitive impairment. There is little information about the combined use of medication and psychological treatment in GAD. There has been concern, however, that the cognitive impairment produced by benzodiazepines may decrease the ability of the patient to carry out psychological treatments.

6.06.9.1.2 Panic disorder and agoraphobia

A number of drug treatments are effective in panic disorder. These include high potency benzodiazepines such as alprazolam and clonazepam, as well as tricyclic antidepressants such as imipramine and clomipramine (Lydiard & Ballenger, 1987; Modigh, Westberg, & Eriksson, 1992). SSRIs are also effective and paroxetine is licensed for the treatment of panic disorder in the UK (Oehrberg et al., 1995). MAOIs are good antipanic drugs but are little used because of their adverse food and drug interactions (Nutt & Bell, 1997; Nutt & Glue, 1989).

Overall, all these drugs have efficacy in both preventing panic attacks and lessening phobic avoidance. However, antidepressants are difficult to use in panic disorder because initial treatment often produces symptomatic worsening and jitteriness. For this reason a low starting dose and careful titration are required. However, to obtain the best clinical response it is necessary eventually to build the doses up so that they are similar to those required in major depression (Mavissakalian & Perel, 1995). The

Table 11 Spectrum of activity for anxiolytic drugs.

Drug	*Generalized anxiety disorder*[a]	*Panic disorder*[b]	*Social phobia*	*Obsessive-compulsive disorder*
Benzodiazepines	+	+[c]	?[c]	0
Buspirone	+	–	?	0
TCAs	+	+	–	0[d]
SSRIs	?	+	+	+
MAOIs	?	+	+	0

[a] + = effective; – = not effective, ? = uncertain. [b] With or without agoraphobia. [c] High potency compound (alprazolam, clonazepam). [d] Clomipramine effective.

high potency benzodiazepines, alprazolam, and clonazepam, have the advantage of not causing increased anxiety early in treatment and this makes them easier to use (Schweizer, Rickels, Weiss, & Zavodnick, 1993). However, withdrawal of these drugs in patients with panic disorder can be very difficult (Fyer et al., 1987).

Cognitive behavior therapy is effective in panic disorder and has the advantage that relapse is less common after the end of treatment than it is after drug treatment is withdrawn (Clark et al., 1994). Antidepressant treatment retains a useful place in patients who are not able to benefit from cognitive therapy. It is posssible that the combination of behavior therapy and antidepressant treatment might produce increased efficacy (de Beurs, Vanbalkom, Lange, Koele, & Van Dyke, 1995). A small potentiation of treatment efficacy was also apparent when alprazolam was added to exposure therapy for patients agarophobia and panic. However, alprazolam-treated patients showed less improvement after the end of the study (Marks et al., 1993). Alprazolam also produced significant impairments on word recall tasks (Curran et al., 1994).

6.06.9.1.3 Obsessive-compulsive disorder

Drugs that produce marked potentiation of brain serotonin function, such as SSRIs and clomipramine, are effective in obsessive-compulsive disorder (OCD). About 50% of patients are much improved, although recovery is rarely complete. Similar doses or somewhat higher are used in the treatment of OCD as in major depression. Both obsessional ruminations as well as rituals respond to drug treatment. The time course of response is rather longer than in major depression with a linear rate of improvement beginning at about 4 weeks and continuing for 12 weeks and more (Jenike, 1992).

There is little evidence that drug treatment increases the effect of behavior therapy in patients able to comply with this treatment (Cobb, 1992). However, drug treatment has a role in patients who cannot undertake behavior therapy or who are unresponsive to it. Relapse rates are much less after successful behavior therapy than after withdrawal of antidepressant drug treatment. It has therefore been suggested that, where patients have improved with drug treatment, behavior therapy may be added with the aim of facilitating drug discontinuation (Cobb, 1992).

6.06.9.1.4 Social phobia

There is less evidence about the utility of drug treatment in the management of social phobia, although there is evidence from controlled trials that treatment with SSRIs and MAOIs may confer benefit (van Vliet, den Boer, & Westenberg, 1994; Versioni, Nardi, & Mundim, 1992). Interactions of drug treatment with psychological treatment have not been systematically studied.

6.06.9.2 Insomnia

Insomnia is a common health problem in community samples (Lasagna, 1995). Although hypnotic drugs such as temazepam and zopiclone are effective in the short term, continued use may result in tolerance and rebound insomnia often occurs upon drug discontinuation. It is recommended therefore that hypnotic drugs be employed only for short-term treatment. Many people with insomnia use other pharmacological remedies such as sedating antihistamines or alcohol (Lasagna, 1995). These measures are generally of limited utility. The role of the pineal hormone melatonin is arousing increasing interest but controlled longitudinal studies are rare (Gafinkel, Laudon, Nof, & Zisapel, 1995).

There are a number of psychological methods of helping insomnia (Morin, Culbert, & Schwartz, 1994). These should be preferred to drug treatment in the first instance because they have fewer adverse effects and their benefits, if obtained, are likely to persist for longer.

6.06.9.3 Depression

Antidepressant drugs are used in the treatment of major depression and dysthymia (Depression Guideline Panel, 1993). Antidepressant drugs appear to be of definite value in more clinically severe depressions, particularly those that meet criteria for melancholic features (*DSM-IV*) or somatic symptoms (*ICD-10*) (Depression Guideline Panel, 1993). Although some improvement in depressive symptoms may be seen in the first week of treatment, generally antidepressant drugs can take 4–6 weeks to exert clinically important effects (Depression Guideline Panel, 1993; Quitkin et al., 1996).

6.06.9.3.1 Choice of antidepressant

As noted above several kinds of antidepressant treatment are available which are of generally equivalent efficacy in the broad range of depressed patients. For the more severely depressed subjects, however, treatment with a tertiary tricyclic antidepressant, such as amitriptyline or clomipramine or the SNRI venlafaxine, should be considered (Anderson, 1997).

For other patients the choice can be made according to the symptom profile of the antidepressant and the needs of the patient. For example, in subjects striving to carry on with their usual work and social activities, relatively nonsedating compounds such as lofepramine, nefazodone, SSRIs or venlafaxine would be suitable. Tertiary tricyclics may be helpful when sleep disturbance is severe or when rapid sedation is needed. In some patients, however, tricyclic antidepressants will be contraindicated because of their anticholinergic and cardiovascular side effects. Sedating compounds suitable in this situation include trazodone and mirtazepine. With the exception of lofepramine, tricyclic antidepressants should not be prescribed where there is a risk of deliberate overdose. MAOIs will generally be used as second- or third-line drugs because of their food and drug interactions.

6.06.9.3.2 Prophylaxis of recurrent major depression

Once a patient has responded to antidepressant drug therapy, drug treatment is usually continued for at least 6 months to prevent relapse of syptoms; this is called continuation therapy (NIMH Consensus Development Conference Statement, 1985). In patients with recurrent major depression longer-term prophylactic drug therapy is often required to prevent frequent disabling recurrences. Generally, if an antidepressant is effective in acute phase treatment, it will provide a useful prophylactic effect during long-term treatment (Quitkin et al., 1996). However, some patients respond better to lithium prophylaxis or to the combination of antidepressant and lithium.

6.06.9.3.3 Psychological therapies and antidepressant drug treatment

Specific psychotherapies such as interpersonal therapy (IPT) and cognitive behavior therapy are effective in the treatment of major depression. In general, patients with more severe depression appear to do less well with psychotherapy. For example, both IPT and cognitive behavior therapy appear less effective in patients with disturbances of sleep architecture (Thase et al., 1997; Thase, Simons, & Reynolds, 1996), a common feature of depression with melancholic features. It has also been reported that patients who fail respond to cognitive therapy can subsequently show benefit with antidepressant medication (Stewart, Mercier, Agosti, Guardino, & Quitkin, 1993).

There is a strong clinical impression that combinations of psychotherapy and antidepressant medication may be more efficacious than either alone in patients whose depression is complicated by interpersonal and social difficulties, but few controlled trials exist to sustain this opinion (Depression Guideline Panel, 1993).

There have been studies of drug and psychotherapy interactions in the longer-term treatment of depression. For example, Frank et al. (1990) found that interpersonal therapy given once monthly delayed, but did not prevent, depressive recurrence compared to placebo medication and clinical management. In the same study interpersonal therapy did not augment the effect of impiramine to prevent depressive recurrence. There is some evidence that cognitive therapy can decrease the risk of subsequent relapse (Blackburn, Eunson, & Bishop, 1986). If this is the case it could be a very useful treatment for the many patients who wish to discontinue medication but are unable to do so because of symptomatic recurrence.

6.06.9.4 Mania

Mania is generally treated with mood-stabilizing drugs and antipsychotic medication. In the USA antipsychotic drugs are avoided as far as possible because of the risk of movement disorders and treatment with lithium or valproate preferred (Chou, 1991). It is important to note that the plasma level of lithium effective in mania (0.8–1.2 mmol/l) is somewhat higher than the range recommended for prophylaxis (0.5–0.8 mmol/l) (Prien, Caffey, & Glett, 1992).

In the UK, because of the risk of lithium rebound upon sudden disontinuation (Goodwin, 1994), antipsychotic medication tends to be used more in the treatment of acute mania. In general, conventional antipsychotic drugs such as chlopromazine and haloperidol are employed; there is little information on the efficacy of newer antipsychotic drugs such as risperidone and olanzepine. Benzodiazepines may be employed for sedation (Chou, 1991). The treatment of mania requires skilled nursing and psychological management but specific interpersonal or cognitive therapies are not generally used.

6.06.9.5 Prophylaxis of Bipolar Illness

Bipolar illness is a recurrent disorder and prophylaxis with mood-stabilizing drugs is a mainstay of treatment. A major problem in the management of bipolar disorder is lack of compliance with treatment. This has a variety of causes, ranging from adverse effects of drugs to the difficulty that people experience in coming to terms with life-long serious illness which requires continued drug treatment. In addition, patients need to learn to recognize minor mood changes that may be the prodrome for more serious mood disturbances unless action is taken. It is possible to address these issues in a number of ways, ranging from education to formal cognitive behaviour therapy. Such measures improve the prognosis in bipolar disorder (Jamison, 1994).

6.06.9.6 Schizophrenia

Antipsychotic drugs are necessary for the acute treatment of psychosis and are also required for long-term prophylaxis in many patients. First-line drug treatment has usually been with chlorpromazine or haloperidol. However, such drugs are likely to be gradually replaced with newer atypical agents such as risperidone and olanzepine because of the decreased incidence of movement disorders and somewhat greater efficacy of the latter agents (Marder & Meibach, 1994; Tollefson & Sanger, 1997).

Atypical agents are not available yet in parenteral form and, where oral medication is not possible, intramuscular administration of conventional antipsychotic drugs has to be employed. As noted above, current trends are to use low doses of a high potency agent such as haloperidol with additional benzodiazepine medication if required (Pilowsky et al., 1992).

For longer-term prophylactic treatment, again atypical agents may be preferred. If depot preparations are needed, however, conventional drugs such as flupenthixol, fluphenazine or haloperidol have to be used.

Most patients with schizophrenia continue to experience a level of positive and negative symptoms despite drug treatment. There is growing interest in the use of cognitive behavior therapy in addition to medication in schizophrenia. There is evidence that combined drug and psychotherapy in selected patients can improve the resolution of delusions and hallucinations (Drury, Birchwood, Cochrane, & Macmillan, 1996).

As with the long-term management of bipolar illness, lack of compliance with medication is a major problem in patients with schizophrenia. There is evidence that cognitive therapy designed to help the patient evaluate their attitudes to their illness and its treatment may improve the take up of drug therapy (Kemp, Hayward, Applewhaite, Everitt, & David, 1996).

6.06.10 REFERENCES

Addonizio, G., & Susman, V. L. (1991). *Neuroleptic malignant syndrome*. St Louise, MO: Mosby-Year Book.

Anderson, I. M. (1997). Lessons to be learnt from meta-analyses of newer versus older antidepressants. *Advances in Psychiatric Treatment*, *3*, 57–62.

Angst, J., Amrein, R., & Stabl, M. (1995). Moclobemide and tricyclic antidepressants in severe depression: A meta-analysis and prospective studies. *Journal of Clinical Psychopharmacology*, (Suppl. 15), 16S–23S.

Ascher, J. A., Cole, J. O., Colin, J. N., Feighner, J. P., Ferris, R. M., Fibiger, H. C., Golden, R. N., Martin, P., Potter, W. Z., Richelson, E., & Saulser, F. (1995). Bupropion: A review of its mechanism of antidepressant activity. *Journal of Clinical Psychiatry, 56,* 395–401.

Baldessarini, R. J., Tondo, L., Floris, G., & Rudas, N. (1997). Reduced morbidity after gradual discontinuation of lithium treatment for bipolar I and II disorders: A replication study. *American Journal of Psychiatry, 154,* 551–553.

Blackburn, I. M., Eunson, K. M., & Bishop, S. (1986). A two-year naturalistic follow-up of depressed patients treated with cognitive therapy, pharmacotherapy and a combination of both. *Journal of Affective Disorders, 10,* 67–75.

Blier, P., & de Montigny, C. (1994). Current advances and trends in the treatment of depression. *Trends in Pharmacological Sciences*, *15*, 220–226.

Bruijn, J. A., Moleman, P., Mulder, P. G. H., van den Broek, W. W., Van Hulst, A. M., Van der Mast, R. C., & Van de Wetering, B. J. M. (1996). A double-blind fixed blood-level study comparing mirtazapine with imipramine in depressed inpatients. *Psychopharmacology*, *127*, 231–237.

Cade, J. F. (1949). Lithium salts in the treatment of psychotic excitement. *Medical Journal of Australia, 2,* 349–352.

Chalmers, J. S., & Cowen, P. J. (1990). Drug treatment of tricyclic resistant depression. *International Review of Psychiatry, 2,* 239–248.

Chambers, C. D., Johnson, K. A., & Dick, L. M. (1996). Birth outcomes in pregnant women taking fluoxetine. *New England Journal of Medicine*, *335*, 1010–1056.

Chou, J. C. (1991). Recent advances in the treatment of mania. *Journal of Clinical Psychopharmacology, 11,* 3–21.

Clark, D. M., Salkovskis, P. M., Hackmann, A., Middleton, H., Anastasiades, P., & Gelder, M. G. (1994). A comparison of cognitive therapy, applied relaxation and imipramine in the treatment of panic disorder. *British Journal of Psychiatry, 164,* 759–769.

Clerc, G. E., Ruimy, P., & Verdeau-Pailles, J. (1994). A double-blind comparison of venlafaxine and fluoxetine in patients hospitalised for major depression and melancholia. *International Clinical Psychopharmacology, 9,* 139–143.

Clinical Psychiatry Committee (1965). Clinical trials of the treatment of depressive illness: report to the Medical Research Council. *British Medical Journal i*, 881–886.

Cobb, J. (1992). Serotonin reuptake inhibitors in obsessive-compulsive disorder: What is their therapeutic role? In K. Hawton & P. Cowen (Eds.), *Practical problems in clinical psychiatry* (pp. 63–76). Oxford, UK: Oxford University Press.

Cowdrey, R. W., & Gardner, D. L. (1988). Pharmacotherapy of borderline personality disorder. *Archives of General Psychiatry, 45*, 111–119.

Cowen, P. J. (1997). Pharmacotherapy for anxiety disorders: Drugs available. *Advances in Psychiatric Treatment, 3*, 66–71.

Cowen, P. J. (1996). Antidepressant drugs. In J. K. Aronson & C. J. van Boxte (Eds.), *side effects of drugs* *annual 19* (pp. 7–13) Amsterdam: Elsevier.

Cowen, P. J. (1992). Drugs in focus: Buspirone. *Prescribers' Journal, 32*, 200–204.

Curran, H. V., Bond, A., O'Sullivan, G., Bruce, M., Marks, I., Lelliot, P., Shine, P., & Lader, M. (1994). Memory functions, alprazolam and exposure therapy: A controlled longitudinal study of agoraphobia with panic disorder. *Psychological Medicine, 24,* 969–976.

Danish University Antidepressant Group. (1990). Paroxetine: A selective serotonin reuptake inhibitor showing better tolerance but weaker antidepressant effect than clomipramine in a controlled multicentre study. *Journal of Affective Disorders*, *18*, 289–299.

Dardennes, R., Even, C., Bange, F., & Heim, A. (1995). Comparison of carbamazepine and lithium in the prophylaxis of bipolar disorders. A meta-analysis. *British Journal of Psychiatry, 166,* 378–381.

Davidson, J. R. T. (1992). Monoamine oxidase inhibitors. In E. S. Paykel (Ed.), *Monoamine oxidase inhibitors* (pp. 345–358). Edinburgh, UK: Churchill Livingstone.

Davies, R., & Wilde, M. I. (1996). Mirtazapine. A review of its pharmacology and therapeutic potential in the management of major depression. *CNS Drugs*, *5*, 389–402.

de Beurs, E., Vanbalkom, J. L. M., Lange, A., Koele, P., & van Dyke, R. (1995). Treatment of panic disorder with agoraphobia: Comparison of fluvoxamine, placebo, and psychological panic management combined with exposure and of exposure in vivo alone. *American Journal of Psychiatry, 152,* 683–691.

Den Boer, J. A., & Westenberg, H. G. M. (1988). Effect of a serotonin and noradrenaline uptake inhibitor in panic disorder: A double-blind comparative study with fluvoxamine and maprotiline. *International Clinical Psychopharmacology, 3*, 59–74.

Depression Guideline Panel (1993). *Depression in primary care* (Vol. 2). Clinical Practice Guidelines. No. 5. Rockville: MD: US Department of Health and Human Services.

Drury, V., Birchwood, M., Cochrane, R., & Macmillan, F. (1996). Cognitive therapy and recovery from acute psychosis: A controlled trial. *British Journal of Psychiatry, 169*, 593–601.

Farde, L., Wiesel, F. A., Nordstrom, A. L., & G. S. (1989). D_1- and D_2-dopamine receptor occupancy during treatment with conventional and atypical neuroleptics. *Psychopharmacology, 99,* S28–S31.

Feighner, J. P. (1994). The role of venlafaxine in rational antidepressant therapy. *Journal of Clinical Psychiatry, 59*(9 Suppl. A), 62–68.

Frank, E., Kupfer, D. J., Perel, J. M., Cornes, C., Jarrett, D. B., Mallinger, A. G., Thase, M. E., McEachran, A. B., & Grochocinski, V. J. (1990). Three year outcomes of maintenance therapies in recurrent depression. *Archives of General Psychiatry, 47*, 1093–1099.

Fyer, A. J., Liebowitz, M. R., Gorman, J. M., Campeas, R., Levin, A., Davies, S. O., Goetz, D., & Klein, D. F. (1987). Discontinuation of alprazolam treatment in panic patients. *American Journal of Psychiatry, 144,* 303–308.

Garfinkel, D., Laudon, M., Nof, D., & Zisapel, N. (1995). Improvement of sleep quality in elderly people by controlled-released melatonin. *Lancet, 346,* 541–543.

Gelder, M., Gath, D., Mayou, R., & Cowen, P. (1996). *Oxford Textbook of Psychiatry* (pp. 548–558). Oxford, UK: Oxford University Press.

Gerlach, J., & Peacock, L. (1995). New antipsychotics: The present status. *International Clinical Psychopharmacology, 10* (Suppl. 3), 39–48.

Glue, P. W., Nutt, D. J., Cowen, P. J., & Broadbent, D. (1987). Selective effects of lithium on cognitive performance in man. *Psychopharmacology, 91,* 109–111.

Goodwin, G. M. (1994). Recurrence of mania after lithium withdrawal. *British Journal of Psychiatry, 164,* 149–152.

Goodwin, F. K., & Jamison, K. R. (1990). Medical treatment of acute bipolar depression. In F. K. Goodwin & K. R. Jamison (Eds.), *Manic depressive illness* (pp. 630–664). Oxford, UK: Oxford University Press.

Jacobson, S. J., Jones, K., Johnson, K., Ceolin, L., Kaur, P., Sahn, D., Donnenfeld, A. E., Rieder, N., Santelli, R., Smyth, J., Pastuszak, A., Einarson, T., & Koren, G. (1992). Prospective multicentre study of pregnancy outcome after lithium exposure during the first trimester. *Lancet, 339,* 530–533.

Jamison, K. R. (1992). Manic-depressive illness. In K. Hawton & P. Cowen (Eds.), *Practical problems in clinical psychiatry* (pp. 33–50). Oxford, UK: Oxford University Press.

Jenike, M. A. (1992). Pharmacologic treatment of obsessive compulsive disorders. *Psychiatric Clinics of North America, 15,* 895–919.

Jeste, D. V., & Caligiuri, M. P. (1993). Tardive dyskinesia. *Schizophrenia Bulletin, 19,* 303–315.

Kane, J., Honigfeld, G., Singer, J., & Meltzer, H. Y. (1988). Clozapine for the treatment-resistant schizophrenic: A double blind comparison with chlorpromazine. *Archives of General Psychiatry, 45,* 789–796.

Kemp, R., Hayward, P., Applewhaite, G., Everitt, B., & David, A. (1996). Compliance therapy in psychotic patients: randomised controlled trial. *British Medical Journal, 312*, 345–349.

Kilbourne, E. M., Phiten, R. M., Kamb, M. L., & Falk, H. (1996). Tryptophan produced by Showa Denko and epidemic eosinophilia-myalgia syndrome. *Journal of Rheumatology, 23*(Suppl. 46), 81–88.

Krupp, P., & Barnes, P. (1992). Clozapine-associated agranulocytosis: Risk in aetiology. *British Journal of Psychiatry, 160*(Suppl. 17), 38–40.

Lancaster, S. G., & Gonzalez, J. P. (1990). Lofepramine: A review of its pharmacodynamic and pharmacokinetic properties and therapeutic efficacy in depressive illness. *Drugs*, *37*, 123–140.

Langtry, H. D., & Benfield, P. (1990). Zolpidem: A review of its pharmacodynamic and pharmacokinetic properties and therapeutic potential. *Drugs, 40*, 291–313.

Lasagna, L. (1995). Over-the-counter hypnotics and chronic insomnia in the elderly. *Journal of Clinical Psychopharmacology, 15,* 383–386.

Lithium Mechanisms Study Group. (1993). Mechanisms of lithium action. *Reviews in Contemporary Pharmacotherapy, 4*, 287–317.

Livingston, M. G. (1994). Risperidone. *Lancet, 343*, 457–460.

Lydiard, R. B., & Ballenger, J. C. (1987). Antidepressants in panic disorder and agoraphobia. *Journal of Affective Disorders, 13,* 153–168.

Marder, S. R., & Meibach, R. C. (1994). Risperidone in the treatment of schizophrenia. *American Journal of Psychiatry, 151,* 825–835.

Marks, I. M., Swinson, R. P., Basoglu, M., Kuch, K., Noshirvani, H., O'Sullivan, G., Lelliot, P. T., Kirby, M., McNamee, G., Sengun, S., & Wickwire, K. (1993). Alprazolam and exposure alone and combined in panic disorder with agoraphobia. *British Journal of Psychiatry, 162,* 776–787.

Mavissakalian, M. R., & Perel, J. M. (1995). Imipramine treatment of panic disorder with agoraphobia: Dose ranging and plasma level–response relationships. *American Journal of Psychiatry, 152,* 673–682.

McElroy, S. L., Keck, P. E., & Pope, H. G. (1987). Sodium valproate: Its use in primary psychiatric disorders. *Journal of Clinical Psychopharmacology, 7,* 16–24.

McElroy, S. L., Keck, P. E., Pope, H. G. & Hudson, J. I. (1992). Valproate in the treatment of biopolar disorder: Literature review and clinical guideline. *Journal of Clinical Psychopharmacology, 12,* 42S–52S.

Modigh, K., Westberg, P., & Eriksson, E. (1992). Superiority of clomipramine over imipramine in the treatment of panic disorder: A placebo-controlled trial. *Journal of Clinical Psychopharmacology, 12,* 251–261.

Montgomery, S. A., Bullock, T., & Pinder, R. M. (1991). The clinical profile of mianserin. *Nord Psykiatric Tidsskr*, *45* (Suppl. 24), 27–35.

Morin, C. M., Culbert, J. P., & Schwartz, S. M. (1994). Non-pharmacological interventions for insomnia: A meta-analysis of treatment efficacy. *American Journal of Psychiatry, 151,* 1172–1180.

Nilsson, A. (1993). The anti-aggressive actions of lithium. *Reviews in Contemporary Pharmacotherapy, 4,* 269–285.

NIMH Consensus Development Conference Statement (1985). Mood disorders: Pharmacological prevention of recurrence. *American Journal of Psychiatry, 142,* 469–476.

Nutt, D., & Bell, C. (1997). Practical pharmacotherapy for anxiety. *Advances in Psychiatric Treatment, 3,* 79–85.

Nutt, D. J., Cowen, P. J., & Little, H. J. (1982). Unusual interactions of benzodiazepine receptor antagonists. *Nature, 295*, 436–438.

Nutt, D. J., & Glue, P. W. (1989). Monoamine oxidase inhibitors: Rehabilitation from recent research. *British Journal of Psychiatry*, *154*, 287–291.

Oehrberg, S., Christiansen, P. E., Behnke, K., Borup, A. L., Severin, B., Soegaard, J., Calberg, H., Judge, R., Ohrstrom, J. K., & Manniche, P. N. (1995). Paroxetine plus cognitive behaviour therapy reduces the number of panic attacks in panic disorder. *British Journal of Psychiatry, 167,* 374–379.

Paykel, E. S. (1990). Monoamine oxidase inhibitors: When should they be used. In K. Hawton & P. Cowen (Eds.), *Dilemmas and difficulties in the management of psychiatric patients* (pp. 17–30). Oxford, UK: Oxford University Press.

Pertursson, H., & Lader, M. H. (1984). *Dependence on tranquillisers*. Oxford, UK: Oxford University Press.

Pilowsky, L. S., Ring, H., Shine, P. J., Battersby, M., & Lader, M. (1992). Rapid tranquillisation: A survey of emergency prescribing in a general psychiatric hospital. *British Journal of Psychiatry, 160,* 831–835.

Post, R. M. (1991). Anticonvulsants as adjuncts or alternatives to lithium in refractory bipolar illness. In J. D. Amsterdam (Ed.), *Advances in neuropsychiatry and psychopharmacology. Vol 2: Refractory Depression* (pp. 155–165). New York: Raven Press.

Post, R. M., Denicoff, K. D., Frye, M. A., & Leverich, G. S. (1997). Re-evaluating carbamazepine prophylaxis in bipolar disorder. *British Journal of Psychiatry, 170,* 202–204.

Post, R. M., Leverich, G. S., Rosoff, A. S., & Altschuler, L. L. (1990). Carbamazepine prophylaxis in refractory affective disorders: A focus on long-term follow-up. *Journal of Clinical Psychopharmacology, 10,* 318–322.

Price, L. H. (1989). Lithium augmentation in tricyclic resistant depression. In I. L. Extein (Ed.), *Treatment of tricyclic-resistant depression* (pp. 51–79). Washington, DC: American Psychiatric Press.

Prien, R. F., Caffey, E. M., & Glett, C. J. (1972). Comparison of lithium carbonate and chlorpromazine in the treatment of mania. *Archives of General Psychiatry, 26,* 146–153.

Quitkin, F. M., McGrath, P. J., Stewart, J. W., Ocepek-Welikson, K., Taylor, B. P., Nunes, E., Deliyannides, D., Agosti, V., Donovan, S. J., Petkova, E., & Klein, D. F. (1996). Chronological milestones to guide drug change: When should clinicians switch antidepressants? *Archives of General Psychiatry*, *53*, 785–792.

Reus, V. I. (1997). Olanzapine: A novel atypical neuroleptic agent. *Lancet, 349*, 1264–1265.

Rickels, K., Downing, R., Schweizer, E., & Hassman, H. (1993). Antidepressants for the treatment of generalised anxiety disorder: A placebo-controlled comparison of imipramine, trazadone and diazepam. *Archives of General Psychiatry, 50,* 884–895.

Rickels, K., Schweizer, E., Clary, C., Fox, I., & Weise, C. (1994). Nefazodone and imipramine in major depression—a placebo controlled trial. *British Journal of Psychiatry, 164,* 802–805.

Robert, E. (1996). Treating depression in pregnancy. *New England Journal of Medicine*, *335*, 1056–1058.

Robinson, D. S., Roberts, D. L., Smith, J. M., Stringfellow, J. C., Kaplita, S. B., Seminara, J. A., & Marcus, R. N. (1996). The safety profile of nefazodone. *Journal of Clinical Psychiatry, 57* (Suppl. 2), 31–38.

Royal College of Psychiatrists (1993). *Consensus statement on the use of high dose antipsychotic medication. Council Report CR26*. London: Royal College of Psychiatrists.

Rudorfer, M. V., & Potter, W. Z. (1989). Antidepressants: A comparative review of the clinical pharmacology and therapeutic use of the "newer" versus the older drugs. *Drugs*, *37*, 713–738.

Schweizer, E., Rickels, K., Weiss, S., & Zavodnick, S. (1993). Maintenance drug treatment of panic disorder. 1. Results of a prospective, placebo-controlled comparison of alprazolam and imipramine. *Archives of General Psychiatry, 50,* 51–60.

Simpson, G. M., & De Leon, J. (1989). Tyramine and new monoamine oxidase inhibitor drugs. *British Journal of Psychiatry*, 155 (Suppl. 6), 32–37.

Stahl, S. M. (1996). *Essential Psychopharmacology*. Cambridge, UK: Cambridge University Press.

Sternbach, H. (1991). The serotonin syndrome. *American Journal of Psychiatry*, *148*, 705–713.

Stewart, J. W., Mercier, M. A., Agosti, V., Guardino, M., & Quitkin, F. M. (1993). Imipramine is effective after unsuccessful cognitive therapy: Sequential use of cognitive therapy and imipramine in depressed outpatients. *Journal of Clinical Psychopharmacology, 13*, 114–119.

Svensson, T. H., & Usdin, T. (1978). Feedback inhibition of brain noradrenaline neurones by tricyclic antidepressants: Alpha-receptor mediation. *Science*, *202*, 1089–1091.

Thase, M. E., Buysse, D. J., Frank, E., Cherry, C. R., Cornes, C. L., Mallinger, A. G., & Kupfer, D. J. (1997). Which depressed patients will respond to interpersonal psychotherapy? The role of abnormal EEG sleep profiles. *American Journal of Psychiatry, 154*, 502–509.

Thase, M. E., Simons, A. D., & Reynolds, C. F. (1996). Abnormal electroencephalographic sleep profiles in major depression: Association with response to cognitive behaviour therapy. *Archives of General Psychiatry, 53,* 99–108.

Tollefson, G. D., & Sanger, T. M. (1997). Negative symptoms: A path analytic approach to double-blind, placebo—and haloperidol—controlled clinical trial with olanzapine. *American Journal of Psychiatry, 154,* 466–474.

Tyrer, P. (1997). Pharmacotherapy for anxiety disorders: Using the available drugs. *Advances in Psychiatric Treatment, 3,* 72–78.

Tyrer, P., Rutherford, D., & Huggett, T. (1981). Benzodiazepine withdrawal symptoms and propranolol. *Lancet, i*, 520–522.

van Vliet, I. M., den Boer, J. A., & Westenberg, H. G. M. (1994). Psychopharmacological treatment of social phobia: a double blind placebo controlled study with fluvoxamine. *Psychopharmacology, 115,* 128–134.

Versiani, M., Nardi, A. E., & Mundim, F. D. (1992). The pharmacology of social phobia. A controlled study of moclobemide and phenelzine. *British Journal of Psychiatry, 161,* 353–360.

Yocca, F. D. (1990). Neurochemistry and neurophysiology of buspirone and gepirone: Interactions of pre-synaptic and post-synaptic 5-HT$_{1A}$ receptors. *Journal of Clinical Psychopharmacology, 10*(Suppl. 3), 6S–12S.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.07

Experiential Treatments: Humanistic, Client-centered, and Gestalt Approaches

LARRY E. BEUTLER, KEVIN BOOKER, and STACEY PEERSON
University of California, Santa Barbara, CA, USA

6.07.1 INTRODUCTION 163

6.07.1.1 The Diversity of Experiential Therapies 165
6.07.1.1.1 Experiential theory as a method of study 165
6.07.1.1.2 Experiential theory as a set of assumptions 165
6.07.1.1.3 Experiential theories view the source of behavior 167
6.07.1.2 Chapter Overview 167

6.07.2 HISTORICAL DEVELOPMENT 167

6.07.2.1 Existential Models: The Rise of Logotherapy 167
6.07.2.2 Phenomenological Approaches: The Rise of Client-centered Therapy 169
6.07.2.3 Humanistic Approaches: The Rise of Gestalt Therapy 172

6.07.3 CONTEMPORARY DEVELOPMENTS 173

6.07.3.1 Theoretical Developments and Applications 174
6.07.3.1.1 Existential theories 174
6.07.3.1.2 Phenomenological theories 175
6.07.3.2 Humanistic Theories 176
6.07.3.3 Research and the Status of Experiential Therapies 178

6.07.4 CONCLUSIONS 179

6.07.5 REFERENCES 180

6.07.1 INTRODUCTION

The theoretical development of psychotherapy has not been smooth; theories have evolved largely through conflict and revolution rather than through an orderly progression of evidence and discovery. From the beginning, Freud demanded unswerving loyalty to his viewpoints, an unrealistic expectation given that the topic of discourse was, at that time, so poorly understood, and that the concepts were so complex and subjective. Freud's lack of sympathy toward those whose perspectives came to diverge from his own extended to a point that even minor disagreements with former disciples resulted frequently in the development of different "schools"; old allegiances and friendships were lost, and theoretical constructs became rigidly reified by the vigorous defenses erected against annihilation by their proponents. This history of dispute and division left a legacy of fragmentation in psychotherapy theory; hundreds of schools of thought have evolved with the very nature of evidence being in hot dispute. As clients and prospective clients, as well as practitioners and scientists, this is the legacy that remains (Bergin & Garfield, 1994; Freedheim, 1992).

In order to both capture the breadth of the field and to bring some order to it, it is useful to conceptualize the field as being composed of various overlapping levels of specificity. At least three such levels, systems, models, and theories, are necessary to capture the color and diversity of the field of psychotherapy.

Most theorists generally identify psychotherapy as evolving from three major systems or schools (Rice & Greenberg, 1992). Each of these intellectual systems continue in somewhat modified form from their initial beginnings and in contemporary psychotherapy, are composed of models and theories, all of which share a common but distinctive view of what motivates behavior. The first system in this evolution is the collection of viewpoints that are often referred to as "psychoanalytic" or "psychodynamic." This system is distinguished from the other two by its reliance on the concept of intrapsychic conflict as the basis of motivation and change. It arose with the ideas of Freud in the first two decades of the twentieth century, and expanded through the contributions both of Freud's (former) disciples and from the application of his ideas to the theories that are known as ego-psychology, self-psychology, and object relations.

The second system in the evolution of psychotherapy came to prominence in the 1940s and 1950s. The behavioral school replaced the concept of intrapsychic conflict with "Reinforcement" as the basic motivating force of behavior. Conditioning, along with its variants of learning by association and consequences, became the primary explanatory construct.

Experiential schools, the "third force" in psychotherapy (Rice & Greenberg, 1992), in contrast to the first two, view motivation as an inherent struggle for integration and growth. Thus, rather than being either driven by base instincts, that were inherently destructive and negative, or by external consequences, that were inherently neutral in social value, self-actualization assumed a decidedly positive valence, moving the individual toward increasingly social and enhancing ends.

Within each of these three systems there are a variety of more specific models that describe how the general system or movement is translated to the specific example of psychotherapy. Each of these models share the view of the larger group, the school, regarding the basis of motivation, but differ in how they think these motives are manifest in psychopathology and how they are addressed in psychotherapy. Specifically, for example, within the larger movement of "Psychoanalytic Theory," as has already been mentioned, there are models based on ego-psychology, self-psychology, and objects-relations theory. "Experiential Theory," likewise, is comprised of existential, humanistic, and person-centered models of psychotherapy; within the behavioral school are radical behaviorism, social learning theory, and cognitive models. Each of these models poses a slightly different view of how behavior is best changed and what aspects and patterns of motives are likely to be manifest in this process.

A third and more specific ordering of theory identifies specific theories of psychotherapy. These theories typically are identified by certain strategies and techniques that distinguish them from others, even within the same model and system. Beck's theory of cognitive therapy is different from that of Ellis, for example, though they both are representative of a cognitive model of psychotherapy.

Again, using the example of Experiential Theory, within the model of existentialism, Daseinsanalysis and Logotherapy represent different, specific, theoretical contributions; within the humanistic tradition, Gestalt Therapy and Redecision Therapy may be recognizable and distinct, and within a phenomenological model, client-centered and person-centered therapies represent a line of evolution within a single theory.

It bears mentioning that some believe that two other systems or schools are sufficiently distinct from their roots and are vying for status as major, independent "forces" in psychotherapy. Cognitive models of behavior are certainly strong contenders for this status. Given that they pose the same motivational system that characterizes behavioral theories, however, it is unlikely that they qualify as a distinctive, major system of the order of psychoanalytic, behavioral, and experiential theories. Similarly, while integrative models are widespread and influential, it is difficult to find a characteristic, common, and distinguishing set of theoretical assumptions. These issues will likely prevent both cognitive and integrative models of intervention from being elevated to the level of a "fourth force" in psychotherapy.

The very presence of these emerging and competing systems, however, and the many variations of theory that they represent, testify to the fragmentation of the field. Certainly, the diversity of psychotherapeutic approaches is substantially greater than might be judged if the view were only at the level of the three broad systems to which theorists often refer. Though it seems that no one has taken the time to count them, there are clearly more than 400 specific theories and an indefinite number of intermediate level models representing these.

6.07.1.1 The Diversity of Experiential Therapies

"Experiential Therapy" is a broad designation within which there is a rich array of diversity. Humanistic, existential, and phenomenological models are well accepted as being subdivisions of this general system. The identity of the more specific theories that represent these models, however, is less consistent. Some theories are not easily classified within a single model while others are blends of two or more. Thus, some authors identify constructivist approaches such as that of Kelly (1955) as an experiential theory (e.g., Bugental & McBeath, 1995; Feshbach, Weiner, & Bohart, 1996), while others identify it as a variant of behavioral or psychodynamic schools (e.g., Mahoney, 1993). The basis of this disparity is not obvious. While this variability suggests that there is a lack of reliability in assigning at least some theories of psychotherapy to general models or schools, it also illustrates the cross-theory blending that is a part of the experiential movement itself.

Culturally, experiential philosophies are ingrained deeply in the values of hard work and sacrifice, the values of the industrial and Protestant revolutions. The evolution of these theories was forever colored by the experiences of holocaust survivors and others who suffered the ravages of war and prolonged catastrophe (e.g., Frankl, 1961). While their theoretical roots date to the European philosophers of the eighteenth and nineteenth centuries, as a domain of practice and service, experiential therapies are largely a post-WWII development. They represent the "transplanting of existential thinking from their chiefly European intellectual roots to a broader US audience of nonmedical, unphilosophically sophisticated, practicing counselors and therapists" (Bugental & McBeath, 1995, p. 112).

Rice and Greenberg (1992; modified from Tageson, 1982) point out that the various theoretical variations within experiential theory share: (i) a reliance on phenomenology as a method of study, (ii) an assumption that in-born actualizing tendencies motivate behavioral development and change, and (iii) a belief in individual choice as the causal locus of behavior.

6.07.1.1.1 Experiential theory as a method of study

The methods that are used by experiential psychotherapists are phenomenological and individualistic, contrasting with the objective, external, and pluralistic methods of psychoanalytic and behavioral models. It is assumed that healing of the most complex problem will occur naturally by acquiring a deep understanding of each individual's unique personal experience of the world in which they live. An assessment of either the truth or value of this personal experience is not reliant on external criteria, but on the person's own subjective standard, their capacity to assess that standard through a process of reflective awareness, and the degree to which the resulting understanding furthers the self-actualization of the individual. Knowledge, as defined through the conventional scientific standards of empirical observation and externally measured but collective responses, is thought to be inadequate unless it deepens an understanding of the unique phenomenology of the person.

The assumption that "Truth" is in the patient's experience contrasts both with psychoanalytic and behavioral views. These latter views identify the therapist, rather than an external body, as the nucleus of knowledge. In psychoanalysis, for example, insight and knowledge come from the therapist's or analyst's interpretation, not from the patient, through the mechanism of interpretation. Thus, it is the analyst, not the analysand, who defines the veracity and validity of experience. Likewise, in behavioral models, the therapist is considered to be the expert or authority whose critical understanding of behavior serves as the guide for developing a technical correction. From the therapist, not the patient, come the suggestions and guidance that are then further supported by homework assignments from the therapist. In both of these views, the value of evidence is proportional to the degree to which the therapist-observer can remain objectively distant from the observed, the patient.

In contrast, in existential/humanistic analysis, information flows from patient to therapist, rather than vice versa. The veracity of an experience is defined, for the experiential-humanist, by self-reflection and personal authenticity. The process of exploring and identifying the nature and content of self-appraisal is not controlled by the therapist; it is only facilitated and allowed to expand by a permissive clinician. It is this self-initiated growth that is the basis of healing.

6.07.1.1.2 Experiential theory as a set of assumptions

Beutler, Bongar, and Shurkin (in press) emphasize that three assumptions both distinguish experiential therapy from psychoanalytic and behavior therapy and represent the bases that bind the specific theories within the domain of experiential models together. These binding assumptions include the beliefs that: (i) there is an innate process that directs emotional growth;

(ii) constraining this inborn tendency from its normal process of unfolding produces distress and psychopathology; and (iii) removal of these constraints in a permissive environment releases the healing process. Among the various models and theories that constitute the experiential school, however, the priority and nature of these beliefs change.

Within the mainstream of the tradition, self-actualization is the fundamental motivational force. It is an innate process that moves the organism to become increasingly complex, balanced, and integrated. Thus, self-actualization requires emotional and sensory as well as cognitive input. Sensory and emotional experiences are thought to be necessary in order to insure the adaptive incorporation of new cognitive information. As a person adapts, accommodates, and incorporates new information, they naturally become more differentiated and complex, both intellectually and interpersonally. Cognition becomes differentiated, emotions become more varied, and behavior is increasingly discriminating.

Put in this way, self-actualization is not tied to a particular set of social values. However, some theories within the phenomenology and humanistic traditions extend the definition to include the view that this growth process includes a movement toward greater productivity, self-sufficiency, creativity, and social adaptability. These definitions add a dimension of social "goodness" and tie the theory more closely to democratic social systems.

The assumption that self-actualization is an inherent drive, drawing one toward growth and differentiation is probably less easily accepted by existential theorists than it is by humanistic and phenomenological theorists (Rice & Greenberg, 1992). In existential models, self-actualization assumes a teleological definition in which growth derives from a pull from individualized goals and aspirations rather than from a force that pushes one toward more abstract, societal goals.

The second assumption underlying experiential models identifies the basis for psychopathology. This assumption points to social and familial constraint and restriction of emotional experience as a pathological force that constrains the natural process of growth and differentiation. Beutler et al. (in press) observe that most Western societies seek to restrain intense emotions and discourage the development and expression of emotions. Experiential theorists argue that this social constraint separates an individual from some of the experience that is necessary for effective information processing. This emotional constraint is based on three beliefs that are considered by experiential theorists to be fallacious and mistaken: (i) cognitive knowledge can be separated from its emotional concomitants, (ii) destructive acts are the products of intense emotions, and (iii) constraint of emotions will protect the society from these destructive acts.

Experiential theories assert that the very act of trying to protect a society by restricting the emotional experiences of its citizens has the paradoxical effect of increasing, rather than reducing, the likelihood of social discord and violence. In this view, human problems are caused, not by excessive expression, but by restraining and fragmenting the normal and necessary acts of thinking, feeling, sensing, and intuiting. Emotional constraint during the course of emotional development, in other words, prevents incorporation of the very information that promotes integration and emotional growth.

Thus, rather than reducing the likelihood of destructive acts, the efforts of society and families to protect themselves from strong emotions is thought to interrupt and fragment the experiences that allow people to develop social conscience and constraint. By becoming separated from their feelings, individuals may then engage in destructive behavior without feeling and sensing the consequences of these experiences. Both the benign failure to actualize and grow from new experience and the malignant social destructiveness of impulsive behavior, therefore, are products of these misdirected societal efforts.

The third assumption that characterizes the experiential movement identifies the basis of correcting the pathological effects of emotional constraint. This assumption defines the nature and objectives of psychotherapy. It assumes that self-actualizing tendencies and powers that have been aborted and fragmented can be reactivated in a therapeutic environment that is characterized by unconventional acceptance of emotional experience. Such an environment is thought to unencumber, allow, and encourage the reconnection of emotions and their expression. Behaviors become reattached to their emotional and sensory concomitants and consequences. By reconnecting these components of experience, self-actualization motives are reactivated, and the normal processes of growing is allowed to continue.

Experiential therapies provide an environment that is free from those arbitrary constraints that are imposed against emotion and sensing by society and by families. The psychotherapist endeavors to focus on, facilitate, sometimes even to create experiences that will magnify emotional and sensory reactions so that it can penetrate defenses that have been deadened by societal rules. The idea that human emotional needs and drives are sources of prosocial and proindividual behavior stands in contrast both with the negative nature of humankind espoused by the id psychology of psychoanalysis and with the *tabula rasa* view of personal development that characterizes most modern reinforcement psychologies.

6.07.1.1.3 Experiential theories view the source of behavior

Experiential philosophies view individuals as having ultimate control in their own lives. They are thought to be active, not only in selecting their own behavior, but in constructing their own realities. It is out of this human capacity to make their own meanings of experience that people construct goals and aspirations. It is also on these meanings that people rely when faced with oppression, physical restriction, and intellectual constraint in order to survive even the most atrocious of events.

The internal construction of meaning and worth, in these eventualities, provides armor against external, destructive forces. Self-determinism, in other words, is the prevailing principle of the experiential movement, and applies as much to thoughts and feelings as it does to behaviors.

Experiential models attempt to counter the victim-based mentality that characterizes both psychoanalytic and behavioral positions. These latter systems assign people to a position of being reactive either to instinctive urges, social disruption, or reinforcements. Unlike these alternative systems, experiential models assume that people have a proactive posture. They retain personal choice over their own perceptions and meanings. This choice is optimized when using and integrating information from the full range of sensory, perceptual, and cognitive experience. Both psychoanalytic and behavioral models are viewed as excluding sensory and emotional experience and of exaggerating the significance of societal standards, insight, and behavior.

6.07.1.2 Chapter Overview

In the rest of this chapter, attention will turn to a variety of specific psychotherapy theories that represent three models of behavior within the experiential tradition: existential, phenomenological, and humanistic psychotherapies. The discussion will be restricted largely to post-WWII developments. Using representative approaches of each of the models that comprise the experiential school as examples, theories will be traced from their post-WWII roots to their status in the 1990s.

The theoretical approaches that have been chosen to be representative of the development of the experiential system include the person-centered (or "client-centered") therapy of Carl Rogers as an example of phenomenological models; the Gestalt therapy of Fritz Perls as a representative example of humanistic models; and the Logotherapy of Victor Frankl as an example of existential models. These early theories will be traced through an evolution that has produced a large array of contemporary modifications, variations, and approaches. An exhaustive description of the offspring and research associated with the theories developed by Rogers, Perls, and Frankl is not possible within the framework of this chapter, but some representative descriptions of the progeny of each will be provided. The contemporary Gestalt therapies of Greenberg and Daldrup and their colleagues will be described; the contemporary experiential therapies of Gendlin and Mahrer and the contemporary existential therapies of May, Maslow, and Bugental.

Following a description of the historical roots of the three general theories, a more general perspective of the experiential system will describe developments in theory and practice, with a review of the research.

6.07.2 HISTORICAL DEVELOPMENT

6.07.2.1 Existential Models: The Rise of Logotherapy

Existential philosophy was transported to the field of psychotherapy from Europe via the immigration of victims of the holocaust following WWII. This philosophy found voice in the US through Frankl's (1963) *Man's search for meaning*, a treatise on survival in the concentration camps. Existential conceptualization of the individual sought to qualify and validate the experience of immediate existence, independent of the theoretical projections about human psychological functioning that were implicated historically by other theories of psychotherapy.

Frankl observed that interpersonal alienation, related either to economic status or political posturing, in conjunction with the precarious stability of social (family) structures, had the capacity to invoke debilitative anxiety. Under a war-time period of constant life-threat, existentialism congealed as a response that restored the dignity of human life and resurrected the human spirit. Theories that explain motivation by reference to homeostasis alone were incapable of capturing the complexity of the human struggle. To Frankl, therefore, it was the act of striving for meaning that formed the basis of motivation and the impetus for change.

Frankl was born in 1905 in Vienna, the seat of Freud's psychoanalysis. He founded the Youth Advisement Centers in 1928, in response to a long-standing interest in emotional well-being, shortly before completing the work for his M.D. degree (1930). As a psychiatrist and neurologist, he joined the faculty of the Department of Neurology at the Rothschild Hospital in Vienna in 1936 and rose to Department Head by 1942. He published papers in prominent journals on the topic of psychiatric and neurological treatment of neurosis, beginning in 1939, and became quite prominent in Vienna psychiatric circles even before Hitler's 1939 invasion of Poland (Patterson & Watkins, 1996).

Frankl was interned both in Auschwitz and Dachau between the years of 1942 and 1945, where he lost virtually all of his family to the gas chambers. *Man's search for meaning*, which reported his experiences there, was published in Vienna shortly after the end of the war. Subsequently he lectured widely in North America, capturing the imagination both of postwar Europe and of the American public. In spite of his great influence on American psychotherapy, Frankl never immigrated to the United States. Following the war, he assumed a full-time faculty position at the University of Vienna where he rose to the rank of Professor in 1955.

Frankl's experiences in Auschwitz and Dachau, and his intimate relationships with despair, hopelessness, depression, and imminent death, qualify him as an effective authority on suffrage and the human condition. Frankl found himself stripped to his naked existence by the atrocities of the Nazis; his father, mother, brother, and even his wife had been sent to the gas ovens. With the exception of his sister, Frankl lost his entire family to the Nazis. He faced the question of how a person who had lost virtually every possession, had every value destroyed, who was hungry and suffering with the constant thought of imminent extermination, could find life worth preserving.

Responsibility and meaning are the hallmarks of Frankl's theoretical construct of "being." These concepts speak directly to the condition of the human sufferer who has made the concession that they have nothing to lose except life, and that life itself is irrelevant without value and choice. Reflecting on his own senseless suffering and misery, Frankl concluded that to live is to suffer, and to survive is to find meaning in the suffering. Frankl believed that each individual was responsible for finding a purpose in life and that each must accept the responsibility that is prescribed by this meaning. Thus, Frankl believed, with Nietzsche, that he who has a "why" to live can bear with almost any "how."

"Being" to Frankl was trivariate, comprising spirituality, freedom, and responsibility. Spiritual awareness was cardinal in the process of being, taking precedence over freedom and responsibility, what he called the "psychophysicum." To reflect on one's self, to project meaning on experience, and to relate with forces that were bigger than one's self were the products of spiritual awareness. Freedom was second only to spirituality in the course of being. Freedom was defined as the capacity to rise above instincts, to reject or accept experience, and to do or not to do. Thus, a person had the freedom to pursue the objectives set by spiritual awareness, and this freedom represented a cardinal distinction between people and other animals. Yet the motivation to achieve meaning through spiritual consciousness and freedom were not driving but pulling forces. That is, to Frankl, freedom and striving for meaning were achieved because they drew people to action, not because it caused or compelled them to act.

In his own struggle, Frankl was forced to deal with the conspiracy of the Nazis to force the oppressed to give up on life as they saw all of their familiar benchmarks, attachments, and goals ripped away. He discovered that when life is thus controlled, and external objects and relationships are lost, some basic "human freedoms" remain; the God-given ability to choose the meanings that one assigns to experience and to choose the attitude one holds in and about the extant conditions and circumstances. It is this ultimate freedom to choose attitudes that both informed Frankl's experiences in the camps of Auschwitz and Dachau and that continue to have the capacity to guide modern people's resolution of the existential dilemma.

The third factor in Frankl's conception of being was responsibility. He came to believe that spiritual awareness and choice were matters of responsibility. One is responsible both to find meaning in one's life and to act on that meaning. Logotherapy (Frankl, 1961), therefore, placed much emphasis on helping people assess the implied responsibilities associated with the meanings they gave to their lives, and to act on their own behalf. This process of developing meaning, claiming freedom, and taking on responsibility, however, was changing constantly with one's position and place in life. The emphasis given by existential theory to the ontological predisposition of the patient is the most fundamental and salient marker for human functioning, and is the lens through which one is viewed.

In Logotherapy, human problems are seen as arising in the form of neurotic anxiety that becomes attached either to the somatic, psychologic, or spiritual aspects of one's life. Frankl gave special attention to the spiritual neuroses, those that arise from existential concerns as opposed to those arising from intrapsychic conflicts. Frankl had great faith in and hope for the capacity of humans to rise above these anxieties. His accounts of life as a prisoner resonate with examples of tremendous human resilience and surprising capacity, of the common man or woman, to transcend the spiritual emaciation and weakness that so often is associated with human suffering.

Frankl bemoaned that spirituality, the experience which is so central to human experience, was ignored by Freud and his followers. He believed that it was from the spiritual sense of humankind that consciousness arose, and in a corresponding fashion, Frankl identified not only an instinctive unconscious, but a spiritual one. By the process of uncovering the spiritual unconscious, esthetic awareness, love, and life values arose.

The will to live in the face of death, the will to hope in the presence of hopelessness, and ultimately the "will-to-meaning" in the presence of meaninglessness were the basis for psychotherapeutic applications of existentialism. Thus, Frankl's goal in psychotherapy was to help patients find meaning and responsibility in life independent of their life's circumstance. In an attempt to help "ground" his patients, he would often ask, "Why do you not commit suicide?" (Frankl, 1963). From their responses, he would seek the emotional basis for their existence.

True to the phenomenological assumption that characterized all of the experiential system, Frankl placed trust in the struggles of lost or searching patients to find a meaning from which to extract the basis for existential preservation; personal elements and experiences with which to illuminate the meaning in their lives. For some, meaning may be found in love for family; for others it may be in a relationship with a church or organization. Whatever the scenario, the intricate and delicate processes of a therapeutic search for meaning involved in the integration or weaving together of these threads of broken lives into a supportive pattern.

6.07.2.2 Phenomenological Approaches: The Rise of Client-centered Therapy

Like existentialism, client-centered therapy had its roots in religious tradition and practice. However, whereas existentialism arose from religious persecution and was supportive of contemporary religious views, the Client-centered movement began as a reaction against the attitudes and philosophies that characterized much of American Protestantism.

Carl R. Rogers (1902–1987) was raised in a close-knit, Methodist family that was committed to the Protestant traditions of work and obedience. He struggled with the religious dogmas of his family, most of which seemed arbitrarily to dictate how he should feel and what he should do. He viewed these dictates as constraining both of his intellectual growth and of his enjoyment, and found little comfort in the admonition that they should be accepted on faith alone. When, as an adolescent, his family moved to a farm, Rogers became enthralled with the science of agriculture. He devoured articles and books describing agricultural research and, deriving support from the dictums of science, he began to question his family's religious traditions. He found in science the hope that abstract principles and assumptions could be put to a test. He also found hope for an escape from unbending religious doctrines. Eventually, the development of a nondirective, nondoctrinaire approach to helping people was his escape from the rigid views of a conservative "(almost fundamentalist) Protestant Christianity" (Rogers, 1959, p. 184).

The route to the development of client-centered therapy passed through a number of stages as Rogers grappled with how to reconcile his scientific need for structure, his personal needs for freedom, and the dogmatic and controlling religion of his youth. Rogers originally called his approach to psychotherapy "Non-directive," a reflection of his de-emphasis on therapist authority. Later, he adopted the term "Client-centered" (Rogers, 1951) to place the focus on the process rather than the techniques. In later years, Rogers came to prefer the term "Person-centered" to reflect his evolving interest in the application of his ideas to environments beyond that of psychotherapy.

In adolescence, Rogers felt distant from the family religion, a distance that was exacerbated by the liberal views to which he was exposed at the University of Wisconsin, where he graduated with a B.A. degree in 1924. Still trying to maintain a foothold in religion, upon entering young adulthood, he sought out a liberal Christian denomination and undertook studies as a minister in the Union Theological Seminary. He found there a religious environment that gave him both the opportunity and the encouragement to engage in free, philosophical thought. This freedom of thought, however, ultimately moved him even further from his religious roots and he left the seminary after two years in order to enter the graduate program in clinical psychology at Columbia University's Teacher's College.

Rogers graduated with an M.A. (in 1928) and a Ph.D. (in 1931), but during these years his break with formal religion became complete and he began to seek other avenues for expressing his humanistic values. Client-centered Therapy may well have been his long-sought alternative to religion in providing meaning and structure in his life.

Rogers initially was encouraged to work with children and families, probably because in these prewar years this domain still offered the only significant opportunity for a nonphysician to offer treatment. In 1928, when he was nearing the completion of his Ph.D. work, Rogers was accepted as a Fellow at the Child Guidance Center in Rochester, New York. Here, he was exposed to a deeply ingrained Freudian viewpoint. He found the rigidity of this viewpoint at contrast to the free-thinking environment that had first attracted him to Union Theological Seminary, and reminiscent of the religion of his youth. Like religion, he found the speculative and nonempirical nature of Freudian methods to be remarkably at odds with the statistical and methodological rigor that characterized both the agricultural science of his youth and the curriculum at Teacher's College. Ultimately, he rejected the rigidity of Freudian dogma as he had rejected fundamental Methodism. In what was a compromise between the contrasting values that he held, at once favoring structure and needing freedom, he adopted the structure of the scientific method as a means of demonstrating the value of his freedom-giving psychotherapy procedures.

After completing a Ph.D. degree from Columbia Teachers College, Rogers continued as a staff member and Director of the Child Study Center as it merged with the University of Rochester. He spent nine years in this position, ending his tenure there in 1940. During these years, America was in heated debate about entering the European war. Scant attention was given to other social ills. Rogers found that his interest in and work as a psychotherapist were not given great value either by his academic colleagues or by the courts and schools that purported to rely on psychological services. Yet, in the midst of an eclectic and supportive group of colleagues in Rochester, he began to formulate his ideas and to conduct experiments on the effectiveness of his nondirective methods.

Rogers became familiar with the controversial works of Otto Rank, and with that of Rank's devotees at the Philadelphia Child Study Center, and this work left its mark on Rogers. On the one hand, he found these approaches to be challenging because they forced him to search for and to find the order that he suspected to exist in clinical work. On the other hand, the stranglehold that psychoanalytic views held on clinical practice prevented his receiving the encouragement and assistance he needed to conduct research on his evolving theories. This led him to seek an academic environment in which he hoped to find more support for his research interests.

Rogers took a position in the Department of Psychology and at the University Counseling Services at Ohio State University in 1940. His hopes that his evolving theory and research program were allied closely enough with experimentally based views of behavior to earn him support from colleagues in this environment proved to be frustrated. These were the war years, a time when the concepts of personal freedom were highly valued, but the encouragement of constraint, control, and obedience were also of signal concern. Devotion to personal freedom, a concept that represented a national mantra during that time, synergistically attached itself to Rogers' own history of constraint and structure. The product was a unique form of psychotherapy that valued freedom, that saw an environment (including a therapy environment) that reduced structure and authoritative demands as the avenue for its realization, and that, paradoxically, sought evidence of its value through the structure and rigor of the scientific method.

At Ohio State University, Rogers discovered that his notions of the subjective were at odds with the rigid laboratory science of psychology that held sway in academia, however. It was through this experience that Rogers became convinced that his theoretical perspectives, indeed, represented a new direction in psychology. His ideas flowered while he was in Ohio, but the fruit emerged later, while he was at the University of Chicago (1945–1957).

WWII opened the door for psychologists to provide services to adults. The needs, both for mental health treatment and for vocational guidance, that were exacerbated as veterans reentered society at the end of the war, became the means for counseling and clinical psychology to separate from school and child psychology, and to enter the world of adult mental health. Correspondingly, Rogers found his niche at the University of Chicago Counseling Center, where he served as the Executive Secretary and Director and enjoyed the richness of working with individuals of varied backgrounds. He taught students of psychology, theology, education, human development, and sociology, and carried on a practice that included work with veterans, families, and young adults. His teaching was supplemented by the responsibilities and excitement of a patient/client load that consumed 15–20 hours per week. It was during these years that client-centered Therapy caught the imagination of the field and became firmly established within clinical and counseling psychology.

In 1957 Rogers returned to his alma mater, the University of Wisconsin, where he was to serve on the faculties of both psychology and psychiatry. His research continued to earn widespread recognition there. His ideas caught on and widely dispersed research groups began developing and refining his ideas. Research on therapist-offered conditions promised both to establish psychotherapy's empirical base and provide a bridge across various theoretical points of view.

In 1962, Rogers became a Fellow at the Center for Advanced Study in the Behavioral Sciences at Stanford, California. By this time, his interests in applying the client-centered approach to group therapy were well entrenched and his ideas were expanding to include community relations, education, and other nontherapy environments. To foster these "person-centered" ideas, following his year at Stanford, Rogers became a Resident Fellow at the Western Behavioral Sciences Institute in La Jolla, California. He went on to found and direct the Center for the Studies of the Person there, and remained in that position, as Founding Fellow and Resident Fellow, until his death in 1987 (Patterson & Watkins, 1996).

As might be guessed from his history, Rogers was preoccupied with establishing the science and the practice of psychology as an integrated discipline. Writing in 1959, he summarized the relationship between science and practice in very contemporary terms. He asserted that while research and theory share the common goal of bringing order to experience, their methods frequently are incompatible. However, he maintained that scientific knowledge can come from clinical observations as well as through controlled research; in fact, naturalistic observation is frequently the first step in empirical research. He considered traditional controlled research paradigms as being too simplistic to understand the complex variables that affect human growth in psychotherapy. Rogers observed that a progressive and helpful science capitalizes on serendipity and chance; it captures in unexpected moments kernels of truth that cannot be preordered on demand within the laboratory.

Finally, he urged that the book should never be closed on any theory. Since every theory contains an unknown element of error, no theory is ever completely valid or invalid. The role of research is to see beyond the dogma of theory and search for the nuclear truths that were germane to and common among all theories and that transcend the clinic, the laboratory, and even the discipline of study. While phenomenological understanding was necessary on the personal, therapeutic level, understanding the nature of this phenomenological truth would require the objectivity of science. The structure of science provided the best available protection against self-deception.

The structure of personality and the nature of adaptation outlined by client-centered therapy is found in the concept of the "self" (Rogers, 1951; 1961). The emergence of Self as a viable explanatory concept provided a base from which the therapist could justify a personal identity in therapy and could directly encourage and foster personal choice on the part of patients and clients.

In its focus on choice and freedom, client-centered therapy revealed its roots, not to be in existential crises, but in a reaction against controlling and oppressive religions that threatened moral choice and self-governance. Unlike Frankl, to Rogers the threat was not to physical existence and humanness, but to mental and moral autonomy. The correction was not to move closer to the mysteries of spiritualism and subjectivity, but further away, clearly reflecting the different religious experiences of the two authors. While Frankl found solace in religious philosophy, Rogers found it constricting. But client-centered therapy, no less than the whole of existential approaches, offered both a positive view of growth and hope for freedom from mental domination. Client-centered therapy added psychotherapy research methodologies to the other important and respected methods used by other humanists to understand the processes of change.

The essence of client-centered therapy was summarized in Rogers (1993/1997) classic paper, "The necessary and sufficient conditions of therapeutic personality change." Though subsequently criticized for assuming an all-important position as both necessary and sufficient (Lambert, 1991), variations of these qualities have been adopted by most therapeutic approaches as, at least, "highly desirable." The conditions emphasized the importance of an emotionally congruent therapist who is able to convey empathic understanding and positive regard.

While frequently interpreted as being a philosophy only about therapist qualities, in truth, Rogers offered, in these principles, a view that included the patient variables and perceptions in creating necessary conditions for change. Specifically, he identified patient distress or incongruence as a motivational construct. This anxiety complemented the natural growth processes by providing the impetus to work and to engage the therapist.

There were other patient qualities hidden in Rogers' treatise as well. For example, he asserted that the patient must perceive the therapist's empathic understanding and regard, implying that a degree of interpersonal facility and competence, as well as a capacity for realistic interpersonal appraisal, are necessary for effective work. In later explanations, it became clear that Rogers was not suggesting that this therapy would be ineffective with those who lacked these skills, but that all individuals had the capacity for the level and type of sensitivity that would allow them to benefit from a therapeutic environment. In his later book, *On becoming a person*, Rogers (1961) revealed his ties to existentialism and humanism as a method of understanding, as a set of assumptions about the nature of human problems, and as a philosophy about the causal locus of change (Rice & Greenberg, 1992).

The famous Rogers–Skinner debates (Kirshenbaum & Henderson, 1989) highlighted Rogers' beliefs in personal choice and emphasized the value he placed on the patient's responsibility for their own change experiences. While considering the locus of change to be within the person, he asserted that free choice could only be expressed when certain environmental events and conditions are present. The therapeutic environment was designed to consist of those contextual conditions. The therapist's sensitive support and the client's willingness to explore work synergistically to foster the conditions for growth.

6.07.2.3 Humanistic Approaches: The Rise of Gestalt Therapy

Gestalt psychotherapy took root as a counter-response to the negativism of psychoanalysis. It was a logical extension of the movement toward personal direction and freedom that characterized Logotherapy and client-centered therapy. It did so, however, without relying on the abstract concepts of spirituality of these former approaches. Although the emergence of Gestalt therapy is generally thought to have been in the postwar years of the 1940s, its genesis actually was some 20 years earlier.

Friedrich (Fritz) Perls (1893–1970) was born in Berlin (Patterson & Watkins, 1996). He obtained an M.D. degree from the Frederich Wilhelm University in 1920. Following his medical training, Perls worked under the direction of Professor Kurt Goldstein at the Frankfurt Neurological Institute for Brain-Damaged Soldiers. With the advent of WWII, Perls took his family to South Africa and in 1935, he established the South African Institute for Psychoanalysis. In 1946, he emmigrated to the US having become disillusioned with the arise of apartheid in South Africa.

By training, Perls was a psychoanalyst, but he was heatedly dissatisfied with the dogma and structure of psychoanalysis. He was also influenced by the experimental work of Kohler, Wertheimer, and Lewin, with whose work he gained familiarity during his early years in Germany. Other existential philosophers with whose work he also became familiar also eventually affected his work, but initially he was too preoccupied with orthodox psychoanalysis to assimilate their work (Perls, 1947). Perhaps this is why he came to identify more closely with psychology than with psychiatry.

Fritz Perls met Laura Posner in 1926 while she was working on her Ph.D. in psychology. She was to become his wife and cofounder of Gestalt therapy. Laura Posner Perls' family was very affluent and culturally enriched, while Fritz was from a lower middle-class Jewish family. These class distinctions were to influence their relationship for many years.

Laura Posner Perls was heavily influenced by Martin Buber and Paul Tillich who were prominent contemporary existentialists. Laura and Fritz Perls worked closely together for nearly 25 years until they separated in the 1950s. Fritz Perls wrote his first book *Ego, hunger and aggression* subtitled "A Revision of Freud's theory and Method" in 1941 and 1942 while serving as a captain in the South African Medical Corps. Although it was not for many years that the name and character of "Gestalt therapy" was succinctly expressed, this first book introduced many of the Gestalt concepts which would later become central ideas in Gestalt therapy.

During the years that Fritz and Laura Posner Perls were together, Gestalt therapy was developing and maturing, although the particular contributions of Laura Perls often are obscured in the available writings. Although Gestalt therapy was first introduced in the US by Fritz and Laura Perls, it was not until Fritz found a home at Esalen Institute at Big Sur, California, in the 1970s that Gestalt therapy was recognized nominally as an independent theory. It was Fritz, not Laura, who came to be recognized as the discoverer, the father, and developer of Gestalt therapy. His estranged wife, Laura Perls, was residing in New York City, outside the mainstream of the human potential movement which was beginning in

California. She published few papers, and her contributions to the theory and methods of Gestalt work were known only to a handful of people (Corsini & Wedding, 1989).

Those familiar with the work of Laura, and Fritz Perls note interesting differences in how they implemented Gestalt therapy. One of the most noted of these differences pertained to the dimension of control and permissiveness that they applied in treatment. Laura employed procedures that were characterized by apparent permissiveness, while Fritz's work emphasized therapist authority and control (Hatcher & Himmelstein, 1976).

The 1930s and 1940s were a dynamic time for Fritz Perls. He was heavily influenced by a number of prominent people including Wilheim Reich, who was Perls' analyst in the 1930s, Karen Horney, and Otto Rank. Horney was said to have directed the young and rebellious Perls to a very eccentric and rebellious Reich. Most notable in his influence was Reich, who introduced Perls to a theory of psychosomatic medicine that considered physical movement and symptoms as the body's armor against threat (Perls, 1947). The body work and physical techniques of Gestalt therapy was a product of this earlier association.

However, it was Freud's theory of psychoanalysis that provided Perls with a theoretical framework for all of his future thinking; despite their differences, psychoanalytic theory was the major foundation upon which Perls built his understanding of human behavior, and it was psychoanalytic theory that he used as a standard against which to evaluate his own emerging theory. In Fritz Perls' autobiography, *In and out of the garbage pail* (Perls, 1969), he makes the following comment on Freud, "Rest in peace, Freud, you stubborn saint-devil-genius," reflecting his own ambivalent attitude toward both Freud and psychoanalysis.

Gestalt therapy differs from other systems and models in a number of important ways. For example, Perls accepted psychoanalysis as a general theory from which his own view derived. However, he chose to omit certain aspects of Freud's theory from his own view, such as psychosexuality, the tripartite anatomy of the personality (id, ego, and superego), and the nature of the unconscious. As applied to technique, these omissions led Perls to emphasize *how* rather than *why*, and explored experience within the "here and the now" rather than the "there and then."

Another difference is in the value assigned to various bodies of scientific research. Gestalt therapy draws from a broader scientific literature than most systems of psychotherapy. It placed greatest value on research that describes the nature of perception and information processing, as well as from literature on defense and psychopathology. At the same time, traditionally it has eschewed psychotherapy outcome research. This priority of values contrasts with client-centered therapy, for example, that has always valued outcome research, but has given little acknowledgment to research on psychopathology and personality development.

Still another distinguishing aspect of Gestalt therapy is its adoption of a holistic view of behavior. It regards individuals as being inherently integrated; their behavior reflects an integrated system whose collective activity cannot be understood by simply viewing isolated acts or structures. This humanistic view is borrowed from Rank, whose concept of the "total organism" contrasts with the psychoanalytic view that separated mind and body and divided the psyche into discrete elements, for example, id, ego, superego, that engaged in a struggle for power over one another. Instead, Gestalt theory asserts that people struggle and experience conflict because of the difficulty of incorporating new information into perceptions based on old knowledge.

Finally, Gestalt therapy defined self-actualization in a manner that contrasted with other experiential approaches. Self-actualization was reflected in balance, differentiation, and integration of cognitive, sensory, and emotional systems, rather than a motive toward social goodness. The ability of conceptual systems to communicate was manifest in the concept of self-response-ability, that is, the ability to choose to be active and to overcome apathy. Rather than encouraging social compliance, as might psychoanalytic therapy, Gestalt therapy encourages social rebellion and individualism.

6.07.3 CONTEMPORARY DEVELOPMENTS

This section will summarize some of the most significant developments and extensions of existential, humanistic, and phenomenological traditions. Both in the interest of space and because time has resulted in an inevitable blurring of theoretical boundaries, no attempt will be made to maintain clear distinctions among these models. Existential theories have become more phenomenological and humanistic ones have become more existential. While some of the more visible theorists within these different systems will be mentioned no attempt will be made to be inclusive or exhaustive and detailed consideration to the fine distinctions among their points of view will not be given.

This will allow a summary of the more important findings from the research.

6.07.3.1 Theoretical Developments and Applications

6.07.3.1.1 Existential theories

Experiential theories have continued to expand, but probably at a lower rate than in their heyday of the 1970s. Theories have become more integrative of phenomonological, humanistic, and existential perspectives, and have been applied to an increasing array of problems, formats, and environments.

Because they did not arise from a single strand of philosophy, it is difficult to find a coherent direction of development that has characterized the evolution of existential therapy. Several major authors, both European and American (e.g., Boss, Binswanger, Bugental, May, Maslow, Yalom), have offered contrasting directions. The works of May, Bugental, and Maslow have been particularly forward-looking and serve as a representative view of the field.

Rollo May, a psychoanalyst by training and practice, probably is most responsible for introducing existential therapy to the United States (Rice & Greenberg, 1992). Along with Irvin Yalom, one of his major contributions (May & Yalom, 1984) was the blending of existential and phenomenological viewpoints. Adopting the phenomenological view, he believed that the personal experience of an individual person was the most important vehicle to knowledge and understanding, and emphasized that a person could be understood through their own experiences more adequately than through a therapist's theory about people (May, 1961; May & Yalom, 1984). Likewise, May thought that access to a patient's motivation for change was best obtained through an exploration of the meanings, ideals, and goals that directed their life.

May's perspective began with a fascination with anxiety and its meaning. This topic served as the basis for his doctoral dissertation in 1950 as well as for a major book, *The meaning of anxiety* (May, 1977) that was revised and reprinted several times. He saw anxiety, defense, and abortion of developmental progress as the inevitable products of conflict between goals and aspirations and the demands and constraints of reality. However, in a viewpoint that was reminiscent of Rogers and others, he maintained that anxiety was growth enhancing.

Pushing this point further, May launched an insightful and provocative attack against psychotherapy theory. He noted that psychotherapists, in the interest of effecting good therapy, observe and attribute meanings to patterns or mechanisms of behavior, using language and concepts that are constructed by their theoretical orientation. He pointed out that both the experience of and resolution of anxiety was often aborted by therapists and other observers when they construed the patient's experience in ways that forced it into line with their own conceptual frameworks.

The tendency of therapists to filter perceptions of the patient through a rigid lens of theoretical views clouded their ability to accurately perceive and relate to the patient. More importantly, it frustrated the objective of bringing the patient into contact with the anxiety, preventing self-knowledge and resolution. He pointed out that a rigid or dogmatic theory prevented the therapist from adopting a phenomonological perspective. The therapists' task is to separate themselves sufficiently from their own histories and dispositions to accurately perceive the meanings offered by a patient's own developmental history.

Perhaps May's major contribution was his piercing questions of fellow therapists. He provocatively asserted the possibility that what therapists view as an understanding of the patient is, in reality, a projection of their own self-reinforcing theory.

The foregoing is not to diminish May's contributions to experiential theory. May's vision was to understand how people become aware of their own growth and potential, that is one's "beingness" (May, Angel, & Ellenberger, 1958). It is this phenomenon of becoming aware of one's beingness that characterizes May's most frequently noted contribution to theory, the "I-Am" experience.

The definition of the "I-Am" includes the perception of what is real, but adds an awareness of some emerging or existent potential. Thus, a person exists in the present but with the potential to *become* (May, 1961). This dynamic aspect of self experience is central to the capacity to change; it is a precondition for their solution (May et al., 1958). The success of this solution, however, depends upon a therapist being able to help clients recognize and experience their own existence.

May's "I-Am" experience was so ontologically sound that it resonated through much of the existential movement. The value of his teachings seemed to help organize and guide the search for what Bugental (1976) calls the "lost sense of being." May's "I-Am" experience was incorporated by Bugental into what he called the "existential sense." Like May, Bugental identified existing or being alive as the fundamental concern of each human being.

Bugental, however, emphasizes that the striving for existence supersedes the mere act of being alive. It is a motivator that drives a person to become more vibrant and sensitive to life. This quest for more life was inextricably related to what he called the tragedy of the human condition, the inability both to recognize and seize opportunities for fuller living.

Bugental's concept of the "lost being" is a person who was invisibly crippled, blind, and deaf to their own state of needing (Bugental, 1976, 1987). This blindness to one's own state of emotional impairment represented a loss of the inward vision that made it possible to assess how well outer and inner experiences match.

Abraham Maslow was another major figure in blending existential and humanistic theory. His most noted contribution was his assertion of a jurisdiction for psychology which dealt explicitly with issues of growth, motivation, and creativeness (Maslow, 1968). His hierarchy of needs expressed his motivational theory and replaced the unimotivational concept of self-actualization, embodied in experiential theory, with a conception that arranged this drive within an array of more basic motives. The ultimate questions of, "Who am I?" and "What am I?" according to Maslow could only be answered by the individual asking the questions and proceeding through the hierarchically arranged answers.

Another contribution of Maslow's was his conceptualization of inner awareness as a process of "listening to the impulse voices" (Maslow, 1971). He proposed that failures in personal growth, such as neuroses, were conditions in which these "impulse voices" or "inner signal" became weak or disappeared altogether. With the inability to hear these signals, the person was incapacitated and became detached from their own essence.

Maslow envisioned the result of detachment to resemble a zombie, the experientially empty person, rather than an anxious person as proposed by others. Anxiety represented a level of being aware of disconnection, but complete disconnection included a lack of awareness of this disconnection. Maslow viewed the disconnected person as one who was empty, and believed that recovering the self must include the recovery of the ability to have and recognize these inner signals (Bugental, 1987).

One major implication of this theory was in the nature of obsessive and compulsive behaviors. The experientially empty person, lacking direction from within, turned to outer cues for guidance and reassurance of existence. This person relied on external cues, such as clocks, rules, calendars, schedules, agenda, and other people as a substitute for personal resources (Maslow, 1971).

6.07.3.1.2 Phenomenological theories

The persuasive power of Rogers' ideas, bolstered by ample research support (see Beutler, Machado, & Neufeldt, 1994; Orlinsky, Grawe, & Parks, 1994), ensured that the concepts of therapy process (therapist empathy, acceptance, regard, and congruence) have been absorbed into the body of psychotherapy. Virtually all schools of psychotherapy find common ground in the views that a helpful therapist is kind, respectful, caring, and understanding, the essence of the Rogerian "necessary" conditions. The differences between the views of these therapist contributions and Rogers' earlier views are twofold: (i) these factors are considered to be minimal rather than optimal conditions for change, and (ii) none are considered to be necessary or sufficient of themselves. While there are a few "pure" client-centered theorists remaining, most within this tradition have accepted these modifications to Rogers' original tenets.

Out of this perspective, the translation of Rogers' views over time are best embodied in extensions of his work to the broader domain of experiential therapy. The most prominent of these are represented by the writings of Laura Rice, Eugene Gendlin, and Alvin Mahrer. These authors have added several specific procedures and philosophies to the general framework of client-centered therapy. The major contributions of each of these writers, however, has been more often remembered for the technical procedures introduced than in the philosophical truths that they have added to understanding (Hart & Tomlinson, 1970; Wexler & Rice, 1974). Indeed, they all represent an emerging view that the therapist-offered conditions (empathy, warmth, positive regard, and congruence) are insufficient to assure change. Thus, they offer the beginnings of a technology to encourage and even to direct change (Rice & Greenberg, 1992).

By the mid-1970s, client-centered therapy had become widely used as both a group and an individual therapy. It was also becoming more cognitive. Information-processing theory was popular and there was a concerted effort to integrate these views with the workings of client-centered therapy (Patterson & Watkins, 1996). Laura Rice proved to be one of the leaders in this movement. She formulated the therapist's role as an assistant to the patient; together patient and therapist embarked on a process of learning to symbolize experience. Rogers had discussed at length the idea that one of the goals of psychotherapy was for patients to begin to symbolize experience that was outside of their awareness. Symbolization represents the process of bringing information into focus and consciousness by giving it mental representation. This occurred largely, though not exclusively, through words.

Rice (1974) proposed that symbolization was necessary in order to accurately process new information and suggested that this process could be facilitated by "Evocative Reflection." This technique involved efforts to reactivate emotional experiences by drawing the client's attention to key experiences. Then the client was asked systematically to process or talk through the experience in a heightened state of arousal. The emotional information was thought to be more easily recoded and integrated by maintaining a heightened state of arousal and this was, in turn, enhanced by emotion-focused reflections and focus.

Gendlin (1981) took the concept of focus a further step, increasing the attention given to client-centered therapy's affect on sensory experiencing. Gendlin's method of focusing (Gendlin, 1981) included the use of nonverbal exercises, such as directed imagery and magnification of multisensory systems, designed to heighten awareness of sensory cues that signaled the presence of emotional experience.

Gendlin (1969) defined focusing as the process in which people make contact with a special kind of bodily awareness known as the "felt sense." This felt sense was neither an emotion nor an intellectual emotional derivative (cognition). Rather, it was a bodily sense or intelligence by which the body directs the person to answers for personal problems. Moving through the six stages of focusing, emotional and cognitive awareness were thought to emerge as by-products (Gendlin, 1969). While Rice's conceptualizations moved client-centered therapy ever closer to cognitive conceptualization, Gendlin's moved it closer to existential and Gestalt perspectives.

Mahrer (1983, 1986, 1996) furthered this movement of client-centered therapy toward humanistic models, increasing its resemblence to Gestalt therapy. In Mahrer's Experiential therapy, the therapist and client both develop images, the therapist's being an effort to replicate that of the patient. Laying side by side, the patient's experience is recreated in the therapist's imagery and both attempt to focus and move closer to the source of intense sensory and emotional experience. Mahrer (1989) has attempted to manualize the key steps that a therapist and patient make in effecting his "Experiential Therapy," making this approach both more researchable and more easily applied than conventional applications.

The steps of Mahrer's Experiential/Humanistic therapy are constructed to provide an increasingly intense therapeutic experience (Mahrer, 1983). In this process, cognition and emotions are salient to the extent that they occur within this genre of client *experiencing*. Thoughts and feelings of the client are both accompanied and facilitated by focusing attention on bodily sensations. The key to unlocking the power of the therapy resides in the client's ability to allow these physical experiences to occur and to then give them meaning. In this process, feelings such as fear, anger, anxiety, excitement, gloom, as well as various cognitive postures such as confusion are identified as signals of the client's experience. Associations with these signals are then used to provide a landscape of meaning to the experiences.

6.07.3.2 Humanistic Theories

Integration has characterized the development of humanistic models of psychotherapy. It is no longer easy to identify a psychotherapy that is only humanistic. The infusion and cross-fertilization of cognitive therapy, experiential/humanistic therapy, and client-centered therapy, with traditional Gestalt therapy, for example, have produced a rich array of interventions and novel directions to the conceptualization and application of psychotherapy. When cross-bred with the active orientation of Gestalt therapy, the experiential approaches of Gendlin and Mahrer have emphasized phenomenological interventions within a provocative and therapist-guided medium. Some developments have highlighted the bridging of cognitive and experiential theories while others have highlighted differences between two somewhat different strands of experiential therapy, the nondirective, person-centered approach of Rogers and the process-directive Gestalt approach (Greenberg, Elliott, & Lietaer, 1994).

Drawing on classical client-centered theory, the person-centered approach regards the "relationship" as necessary and sufficient for therapeutic change; humanistic experiential approaches advocate the importance of an active, process-directive intervention and cognitive therapies focus on the role of conceptual structures in emotional dysfunction. The result is a number of approaches that stress active interventions that utilize emotional access to facilitate a change of perspective and the power of which is deepened within the context of a person-centered relationship.

These evolving experiential/humanistic models of therapy are applied in an increasing variety of formats. While most often used in an individual format (Greenberg et al., 1994), group therapy formats are practiced widely (Daldrup, Beutler, Greenberg, & Engle 1988), and there are guidelines for their application to marital/couples therapy (Greenberg & Johnson, 1988).

Focused Expressive Psychotherapy (FEP) is an integrated experiential method for resolving blocked affect that is based on Gestalt therapy concepts. FEP works with a full range of inhibited emotions such as anger, fear, joy, and sadness with a particular emphasis on the range of emotions that surround the blocked expression of anger (Daldrup et al., 1988). FEP, first attempts to access the client's schematic memory by directing attention to potential relationships in which these memories developed. That is, the client is directed to painful memories involving interactions with parents, spouses, children, and other significant members. Therapy is then designed to heighten the intensity and expression of emotion associated with the memories so that new schematic structures can be realized (Engle, Beutler & Daldrup, 1991).

Although FEP attends to a full range of emotions which may be blocked or inhibited in an individual, particular attention is given to the emotion of anger. FEP assumes a five-step process that identifies emotions, specifies targets, determines markers of their presence, implements experiments in change, and evaluates progress. The patient is actively engaged in all of these processes, but the therapist is responsible for leading them through the processes. The experiments are designed to facilitate full awareness of the emotion itself as well as its implications and genesis.

FEP assumes that there are no "good" or "bad" emotions, per force, but rather that emotions exist in all human beings. Why an individual chooses to react in a dysfunctional or enhancing way to a stressful event is reliant on how accessible they are to emotional and sensory experience and the internalized rules that govern their behavior. The influence of early experiences in childhood cannot be overlooked in the formation of emotional development and the way in which an individual will react in emotionally laden situations. If a child is unable to complete an emotional cycle, then the expression of this emotion (most notably, anger) will become frozen and part of the individual's "unfinished business."

In contrast to most forms of psychotherapy, FEP was designed to exert specific effects for individuals who tend to constrain and inhibit emotional expression, regardless of the specific symptoms manifest. Individuals who can identify the hurtful relationships and experiences that may have caused the inhibition of emotion and who can identify past and present relationships in which they are unable to express important desires are considered to be good candidates for this type of therapy.

Process-experiential therapy (Greenberg, 1994) represents a similar model to FEP, but places more emphasis on the integration of broadly-based nondirective (client-centered) relationship strategies with Gestalt therapy methods, and less emphasis on the role of enduring patient traits. In this approach, there is also an effort to integrate the interventions within a theory of cognitive–emotion relationships. This theory blends cognitive and experiential theory (Greenberg & Safran, 1987). It assumes that the barriers to healthy functioning result from (i) difficulties in finding words or images to symbolize experiences, and (ii) dysfunctional emotional schemes through which to interpret experiences. Accordingly, the goal of therapy is to enable client's to access dysfunctional schemas within a therapeutic environment in order to facilitate relevant schematic change (Greenberg et al., 1993).

In process-experiential therapy, the client and therapist interact on two different levels. The first level concerns the moment-by-moment effects that each of the therapist responses has following a client response. Accurate and empathetic understanding of the client's words and messages provide the client with a feeling of being understood and received by the therapist. It is anticipated that these therapeutic responses will allow the client to initiate work on increasingly more difficult and painful issues.

The second level on which the client and therapist interact is viewed as more "molar" than the first. In this level, the therapist attempts to enable the client to resolve the larger affective problems that present themselves as in-session therapeutic tasks. This requires emotional experiencing, which is facilitated by attending to the emotions and senses in-the-moment, under the assumption that so doing activates emotional meanings that are then subject to inspection and change.

Process-experiential therapy is considered to be most appropriate for use among ambulatory outpatients who are experiencing moderate clinical distress and symptomatology. It is not suited for clients with major thought disorder or schizophrenia, impulse control or antisocial personality patterns, or for those who may be in need of immediate crisis intervention or case management (e.g., acutely suicidal persons) (Elliott & Greenberg, 1995). It is not recommended for those who are unable to be self-reflective and those who find the therapist's nondirective stance of not advising or interpreting to be unacceptable.

6.07.3.3 Research and the Status of Experiential Therapies

The suspicion with which experiential therapies have traditionally regarded both diagnosis and empirical research has given way to produce an emerging body of scientific evidence. Led by the efforts to blend Gestalt therapy with client-centered and cognitive therapy, research has included both naturalistic studies, randomized clinical trials of patients in clearly defined diagnostic groups, and meta-analyses of contributors to efficacy. Findings from all of these research areas have generally supported the value of these approaches.

A prominent characteristic in the evolution of this research is the increasing focus on differentiating the client characteristics and problems for which experiential treatments are likely to be effective. Specific focus has been on clinical disorders such as anxiety disorders, major depression, and personality disorders. Some research has looked at broader dimensions than those identified by diagnosis, particularly focusing on the differential response of patients with different coping styles (internalizing and externalizing), levels of resistance, and those with physical symptoms.

Greenberg et al. (1994) summarized the results of studies conducted since 1978, using meta-analytic techniques. They compared pre- to postchanges on symptom measures and, where possible, compared treatment effect sizes to those observed among patients who were assigned to control or alternative treatments. They found moderate to strong effect sizes for experiential therapies, which compared favorably both with those observed among other therapies and demonstrated that these treatments were more effective than no-treatment and placebo-treatment groups.

Naturalistic and quasi-experimental studies have been used both to assess the relative efficacy of experiential therapy relative to alternative treatments (e.g., Beutler & Mitchell, 1981) and to assess the contributors to effective outcome within different types of psychotherapy (e.g., Burgoon et al., 1993; Greenberg & Foerster, 1996; Hill, Beutler, & Daldrup, 1989; Watson, 1996). Findings from these studies have suggested both that experiential therapies may be at least as effective as alternative treatments and that the level of arousal and experiencing induced by experiential methods facilitate the resolution of interpersonal conflict and reduce emotional distress.

Randomized clinical trial research designs have also incorporated various experiential therapies among the treatment comparisons. Especially promising results have been found in the treatment of those with depression using these controlled comparison procedures (e.g., Paivio & Greenberg, 1995). In a well-controlled investigation, Watson and Greenberg (1996) compared a Gestalt-based experiential therapy with a client-centered intervention, observing that while initial results were equivalent, long-term effects favored the more active experiential approach.

FEP (Beutler et al., 1987) has also been used successfully in the treatment of individuals who present with chronic organic pain. This domain of study is based on the theoretical connection between chronic pain and depression. Beutler, et al. (1987) undertook the task of testing the hypothesis that the intensity of pain correlates with the degree to which a person has persistent difficulty expressing anger and controlling intense emotions (e.g., Beutler, Engle, Oro'-Beutler, Daldrup, & Meredith, 1986). Using a multiple baseline design with six patients who had rheumatoid arthritis, Beutler et al. (1987) found that FEP substantially activated beta-endorphin discharge and correspondingly reduced depressive symptoms. This was particularly true in the early and late phases of treatment. However, the beta-endorphin response was not correlated with alterations of pain, suggesting that they functioned as stress markers more than as anesthetics of pain.

The evidence that certain "types" of patients are particularly responsive to experiential therapy has been the area in which the largest body of research has accumulated. Identifying clients by coping styles is another way of identifying a "type" of client who may benefit from experiential therapy. Whether a client is an externalizer vs. an internalizer may interact with the type of treatment that the client will receive. Beutler et al. (1991) found that those who coped with internalizing styles may profit more than externally-oriented clients in client-centered or nondirective therapies than from either cognitive therapy or Gestalt-based therapy (FEP). This has also been supported in several other studies that have used different approaches to identify or measure internalization. For example, Tscheulin (1990) reported that self-oriented clients (internalizers) did better with a nonconfrontational therapist than with a confrontational therapist.

Several studies (see Greenberg et al., 1994) have also concluded that high levels of client resistance is a contraindicator to the directiveness that is inherent in many experiential treatments, particularly those based on Gestalt/humanistic (e.g., FEP, process-experiential therapy) models. Specifically, highly resistant (high dominance, low submissiveness) patients do better in client-centered or nondirective therapies (Beutler et al., 1991, 1993) than in various directive, experiential therapies. In contrast, clients with low levels of resistance do better in these directive alternatives than they do in nondirective ones (Beutler et al., 1991). Although more research is needed on this general theme, the findings suggest that clients with good interpersonal interests and skills may be the ones who are most receptive to experiential therapies.

6.07.4 CONCLUSIONS

Experiential therapies comprise a heterogeneous collection of theoretically diverse interventions. They share certain assumptions about the nature of existence, the most reliable avenue to knowledge, and the mechanisms of symptom development and change. Humanistic, existential, and phenomenological traditions have had a long history within the fields of psychotherapy. Existential views emphasize the importance of personal meaning and choice in facilitating adaptation and survival in times of crisis; phenomenological approaches have emphasized the role of subjective experience, safety, and therapist-offered conditions as the avenues to understanding and change; and humanistic models have emphasized the role of active therapist interventions and the provision of structure in facilitating change.

Though deriving from different traditions, these models of behavior have merged over time, both in their theoretical development and in their methods of intervention. Central to these emerging, amalgamated theories is a common reliance on the concept of self as an organizing principle of personality, and on self-actualization as a foundation motivational principle.

Experiential models of behavior have historically observed that the "self" has been defined in Western cultures indirectly. Thus, one identifies him or herself through either occupational role (e.g., "I am a bus driver") or through some other external attribute or possession (e.g., "I am rich"). Experiential therapies pose the thesis that such definitions relegate self-definitions to the status of superficial attributes that can be threatened or destroyed by external change. The result is an identity that lacks stability and that is changeable and uncertain. The resulting construct of "self" comes to serve only the most tenuous of functions, and in this process projects a social image that is not authentic. Such reliance on superficial structures provides some limited protection from external threat but does not provide protection from the most basic of fears, the fear of nonexistence. Indeed, its presence makes one vulnerable to related fears of being irrelevant, ignored, and nonessential when these defining roles and possessions are no longer present.

Experiential therapies concentrate on redefining self-views, a process that is generally thought to be natural and orderly when environments are conducive. Thus, the specific therapy models that constitute the experiential tradition all seek to remove obstructions to these normal processes. By providing a permissive atmosphere, they attempt to facilitate and enhance normal movement toward differentiation of self, foster creativity, and stabilize personal identity. Specific theories of psychotherapy vary in the degree of directiveness exercised by the therapist (Perls vs. Rogers), in the roles they assign to spiritual and teleological drives (e.g., Frankl vs. Rogers), and to the acknowledgment of other motives (e.g., Maslow vs. Perls). However, they have in common efforts to focus patient attention on the moment, to encourage attendance to present sensory and perceptual impressions, and to expand therapist awareness of patient experience. While there is not now nor has there ever been absolute concordance among experiential models and therapists' theories, most have embraced this foundation principle.

Modern developments of experiential therapies have blended theories, often including principles and practices of cognitive and behavior therapies, as well as various components from humanistic, existential, and phenomenological viewpoints. There has also been a significant movement to reconceptualize human experience within a developmental perspective. In this conceptualization, emphasis is placed on growth as an immediately ontological experience, the person is a consciously and developing "being." These modern experiential therapies urge patients to focus both on immediate experience and on the progressive change and evolution of that experience as a way of obtaining self-knowledge. This view contrasts with both a static description of personal meanings and with models of behavior that emphasize either the effects of historical experience in their own right, the roles of unconscious drives, or the influence of biologically determined behaviors.

Traditionally, research has not been widely accepted in the experiential traditions. This is with the exception of Carl Rogers, whose client-centered therapy introduced the scientific method to the study of psychotherapy process and outcome. Led by active movements from client-centered and Gestalt traditions, however, there has been a proliferation of research. This research has sought both to specify and operationalize the application of experiential therapy so that they can studied objectively. The resulting manualized therapies have then applied these methodologies to the study of various clinical conditions. The resulting studies have found that experiential therapies produce modest and reasonably strong effects when compared either to patients' pretreatment status, control conditions, or comparison treatments.

Research also confirms that experiential treatments can be adapted to treat specific client disorders (e.g., depression, chronic pain, anxiety) as well as to a variety of client parameters. Research has been particularly fruitful in applying different aspects or tasks of treatments to different clients and situations. Therapists should be sensitive to the possibility that clients who are high in autonomy or reactance may react negatively to the effects of the more process-directive experiential therapies (e.g., Gestalt), whereas more dependent or externally oriented clients may react negatively to nondirective therapies. Likewise, experiential and nonexperiential traditions may be differentially effective for patients who vary in coping style. Experiential therapies may be most appropriate to those who are self-reflective, have organized cognitive processes, and who are capable of establishing stable interpersonal relationships.

Clearly more research is needed to study the effectiveness of experiential treatments. Research is especially needed that assesses treatment outcomes among specific populations and disorders. However, it is also important to direct research attentions to studies of the transfer of treatment to alternative settings, including research on methods of training, procedures to ensure the retention of therapeutic effects, and applications to populations that are generally unresponsive to other treatments.

6.07.5 REFERENCES

Bergin, A. E., & Garfield, S. L. (Eds) (1994). *Handbook of psychotherapy and behavior change* (4th ed.). New York: Wiley.

Beutler, L. E., Bongar, B., & Shurkin, J. (in press). *Am I crazy or is my shrink?* New York: Oxford University Press.

Beutler, L. E., Daldrup, R. J., Engle, D., Oro'-Beutler, M. E., Meredith, K., & Boyer, J. T. (1987). Effects of therapeutically induced affect arousal on depressive symptoms, pain and beta-endorphins among rheumatoid arthritis patients. *Pain, 29,* 325–334.

Beutler, L. E., Engle, D., Mohr, D., Daldrup, R. J., Bergan, J., Meredith, K., & Merry, W. (1991). Predictors of differential responses to cognitive, experiential, and self-directed psychotherapeutic procedures. *Journal of Counsulting and Clinical Psychology, 59,* 333–340.

Beutler, L. E., Engle, D., Oro'-Beutler, M. E., Daldrup, R. J., & Meredith, K. (1986). Inability to express intense affect: A common link between depression and pain? *Journal of Counsulting and Clinical Psychology, 54*(6), 752–759.

Beutler, L. E., Machado, P.P.P., & Neufeldt, S. A. (1994). Therapist variables. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 229–269). New York: Wiley.

Beutler, L. E., & Mitchell, R. (1981). Psychotherapy outcome in depressed and impulsive patients as a function of analytic and experiential treatment procedures. *Psychiatry, 44,* 297–306.

Bugental, J. F. T. (1976). *The search for existential identity.* San Francisco: Jossey-Bass.

Bugental, J. F. T. (1987). *The art of the psychotherapist.* New York: Norton.

Bugental, J. F. T., & McBeath, B. (1995). Depth existential therapy: Evolution since World War II. In B. Bongar & L. E. Beutler (Eds.), *Comprehensive textbook of psychotherapy: Theory and practice* (pp. 111–122). New York: Oxford University Press.

Burgoon, J. K., Beutler, L. E., LePoire, B. A., Engle, D., Bergan, J., Salvio, M., & Mohr, D. (1993). Nonverbal indicies of arousal in group psychotherapy. *Psychotherapy, 30*(4), 635–645.

Corsini, R. J., & Wedding. D. (Eds.) (1989). *Current psychotherapies* (4th ed.). Itaska, IL: Peacock Publishers.

Daldrup, R., Beutler L., Greenberg, L., & Engle, D. (1988). *Focused expressive therapy: A treatment for constricted affect.* New York: Guildford Press.

Elliott, R., & Greenberg, L. S. (1995). Experiential therapy in practice: The process experiential approach. In B. Bongar & L. E. Beutler (Eds.), *Comprehensive textbook of psychotherapy* (pp. 123–139). New York: Oxford University Press.

Engle, D., Beutler, L. E., & Daldrup, R. J. (1991). Focused expressive psychotherapy: Treating blocked emotions. In J. D. Safran & L. S. Greenberg (Eds.), *Emotion psychotherapy & change.* New York: Guilford Press.

Feshbach, S., Weiner, B., & Bohart, A. (1996). *Personality* (4th ed.). Lexington, MS: D. C. Heath.

Freedheim, D. K. (Ed.) (1992). *History of psychotherapy: A century of change.* Washington, DC: American Psychological Association.

Frankl, V. E. (1961). Logotherapy and the challenge of suffering. *Review of Existential Psychology and Psychiatry, 1,* 3–7.

Frankl, V. E. (1963). *Man's search for meaning* (Rev. ed.). New York: Washington Square Press.

Gendlin, E. T. (1969). Focusing. *Psychotherapy: Theory, Research, and Practice, 6,* 4–15

Gendlin, E. (1981). *Focusing* (2nd ed.). New York: Bantam Books.

Greenberg, L. S., Elliott, R. K., & Lietaer, G. (1994). Research on experiential psychotherapies. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 509–539). New York: Wiley.

Greenberg, L. S., & Foerster, F. S. (1996). Task analysis exemplified: The process of resolving unfinished business. *Journal of Consulting and Clinical Psychology, 64,* 438–446.

Greenberg, L. S., & Johnson, S. M. (1988). *Emotionally focused therapy for couples.* New York: Guilford Press.

Greenberg, L. S., & Safran, J. D. (1987). *Emotion in psychotherapy, affect, cognition, and the process of change.* New York: Guilford Press.

Hart, J. T., & Tomlinson, T. M. (Eds.) (1970) *New directions in client-centered therapy.* Boston: Houghton Mifflin.

Hatcher, C., & Himmelstein, P. (Eds.) (1976). *The handbook of Gestalt therapy.* New York: Aronson.

Hill, D., Beutler, L. E., & Daldrup, R. J. (1989). The relationship of process to outcome in brief experiential psychotherapy for chronic pain. *Journal of Clinical Psychology, 45*(6), 951–957.

Kelly G. A. (1955). *The psychology of personal constructs* (Vols. 1–2). New York: Norton.

Kirschenbaum, H., & Henderson, V. L. ((Ed.) (1989). *Carl Rogers: dialogues*. Boston: Houghton Mifflin.

Lambert, M. J. (1991). Introduction to psychotherapy research. In L. E. Beutler & M. Crago (Eds.), *Psychotherapy research: An international review of programmatic studies* (pp. 1–11). Washington, DC: American Psychological Association.

Mahoney, M. J. (1993). Introduction to special section: Theoretical developments in the cognitive psychotherapies. *Journal of Consulting and Clinical Psychology, 61,* 187–193.

Mahrer, A. R. (1983). *Experiential psychotherapy: Basic practices*. New York: Brunner/Mazel.

Mahrer, A. R. (1986). *Therapeutic experiencing: The process of change*. New York: Norton.

Mahrer, A. R. (1989). *How to do experiential psychotherapy: A manual for practitioners*. Ottawa: University of Ottawa Press.

Mahrer, A. R. (1996). *The complete guide to experiential psychotherapy*. New York: Wiley.

Maslow, A. H. (1968). *Toward a psychology of being* (2nd ed.). New York: Van Nostrand Reinhold.

Maslow, A. H. (1971). *The further reaches of human nature*. New York: Viking Press.

May, R. (Ed.) (1961). *Existential psychology*. New York: Random House.

May, R. (1977). *The meaning of anxiety* (2nd ed.). New York: Norton.

May, R., Angel, E., & Ellenberger, H. F. (Eds.) (1958). *Existence: A new dimension in psychiatry and psychology*. New York: Basic Books.

May, R., & Yalom, I. (1984). Existential therapy. In R. J. Corsini (Ed.), *Current psychotherapies* (3rd ed., pp. 354–391). Itaska, IL: Peacock.

Orlinsky, D. E., Grawe, K., & Parks, B. K. (1994). Process and outcome in psychotherapy: Noch Einmal. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 270–376). New York: Wiley.

Paivio, S. C., & Greenberg, L. S. (1995). Resolving unfinished business efficacy of experiential therapy using empty-chair dialogue. *Journal of Consulting and Clinical Psychology, 63*(3), 419–425.

Patterson, C. H., & Watkins, C. E., Jr. (1996). *Theories of psychotherapy* (5th ed.). New York: Harper Collins.

Perls, F. S. (1947). *Ego, hunger and aggression: The beginning of Gestalt therapy*. New York: Random House.

Perls, F. S. (1969). *In and out of the garbage pail*. Lafayette, CA: Real People Press.

Rice, L. N. (1974). The evocative function of the therapist. In D. A. Wexler & L. N. Rice (Eds.), *Innovations in client-centered therapy* (pp. 282–302). New York: Wiley.

Rice, L. N., & Greenberg, L. S. (1992). Humanistic approaches to psychotherapy. In D. K. Freedheim (Ed.), *History of psychotherapy* (pp. 197–224). Washington, DC: American Psychological Association Press.

Rogers, C. R. (1951). *Client-centered thereapy*. Boston: Houghton-Mifflin.

Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology, 21,* 95–103.

Rogers, C. R. (1959). A theory of therapy, personality, and interpersonal relationships as developed in the client-centered framework. In S. Koch (Ed.), *Psychology: A study of science* (Vol. 3, pp. 184–256). New York: McGraw-Hill.

Rogers, C. R. (1961). *On becoming a person*. Boston: Houghton-Miflin.

Tageson, W. C. (1982). *Humanistic psychology: A synthesis*. Homewood, IL: Dorsey Press.

Tscheulin, D. (1990). Confrontation and non-confrontation as differential techniques in differential client-centered therapy. In G. Lietaer, J. Rombauts, & R. Van Balen (Eds.), *Client-centered and experiential psychotherapy in the nineties*. Leuven, Belgium: Leuven University Press.

Watson, J. C. (1996). The relationship between vivid description, emotional arousal, an in-session resolution of problematic reactions. *Journal of Counsulting and clincial Psychology, 64* (3), 1–6.

Watson, J. C., & Greenberg, L. S. (1996). Pathways to change in the psychotherapy of depression: Relating process to session change and outcome: *Psychotherapy, 33,* 262–274.

Wexler, D. A., & Rice, L. N. (1974). *Innovations in client-centered therapy*. New York: Wiley.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.08

Social Skills Training and Problem Solving

KIM T. MUESER
New Hampshire–Dartmouth Psychiatric Research Center, Concord, NH, USA

6.08.1 INTRODUCTION 183

6.08.2 THEORETICAL MODELS 184
  6.08.2.1 The Stress–Vulnerability–Coping Skills Model 185
  6.08.2.2 The Social Skills Model 185

6.08.3 ASSESSMENT 186
  6.08.3.1 Components of Social and Problem Solving Skills 186
  6.08.3.2 Assessment Strategies 187
  6.08.3.3 Nonskill Factors which can Affect Social Competence 188
  6.08.3.4 Integration of Assessment and Treatment 189

6.08.4 TRAINING TECHNIQUES 189
  6.08.4.1 Motor Skills Model 189
  6.08.4.2 The Problem Solving Model 189
  6.08.4.3 Common Learning Principles 190
  6.08.4.4 Basic Skills Training Techniques 190
  6.08.4.5 Problem Solving Training 192

6.08.5 FORMAT OF SOCIAL SKILLS AND PROBLEM SOLVING TRAINING 194
  6.08.5.1 Group Skills Training 194
  6.08.5.2 Individual Format 195
  6.08.5.3 Family Format 195

6.08.6 CLINICAL APPLICATIONS AND RESEARCH 195
  6.08.6.1 Utility of Social and Problem Solving Skills Training 196

6.08.7 SUMMARY AND CONCLUSIONS 198

6.08.8 REFERENCES 198

6.08.1 INTRODUCTION

Over the past several decades, social skills training and problem solving training have become some of the most widely practiced techniques in clinical psychology, with applications spanning a broad range of adult disorders, and numerous books published on their theoretical foundations, clinical applications, and research (e.g., Bellack & Hersen, 1979; Hollin & Trower, 1986; Kleinmuntz, 1966; L'Abate & Milan, 1985; Nezu & Nezu, 1989; O'Donohue & Krasner, 1995b). Although various definitions of social skills have been given by different clinical researchers, Bellack and Hersen offer one of the most comprehensive descriptions of social skills as the:

> ability to express both positive and negative feelings in the interpersonal context without

> suffering consequent loss of reinforcement. Such skill is demonstrated in a large variety of interpersonal contexts and involves the coordinated delivery of appropriate verbal and nonverbal responses. In addition, the socially skilled individual is attuned to the realities of the situation and is aware when he is likely to be reinforced for his efforts. (1979, p. 512)

Problem solving can be defined as the ability to recognize problems or formulate goals, and to develop strategies to successfully reduce or eliminate the problem or to make progress towards achieving a goal. Training in social skills and problem solving generally refers to the systematic application of principles of social learning to the teaching of requisite skills hypothesized to underlie dysfunctions in interpersonal relationships, mood, and functional capacity (Davis & Butcher, 1985; Goldstein, 1982; Larson, 1984).

Skills training has a long history, and no single individual can be given sole credit for developing the method. However, several contributors to the development of social skills training should be noted. In the 1940s, Salter (1949) engaged individuals in role-plays in order to facilitate self-expression and to help them overcome symptoms such as depression and anxiety. In the 1950s, Wolpe incorporated role-playing into his approach to psychotherapy based on reciprocal inhibition. Wolpe (1958) theorized that assertive interpersonal behavior would be experienced as incompatible with feelings of anxiety, leading to the extinction of anxiety in certain social situations. Role-playing was used to help clients develop more assertive interpersonal skills. In the 1960s, Bandura's (1969) work on observational learning led to the formal inclusion of role modeling as a critical ingredient in social skills training. Also around this time, Lazarus (1966) introduced the use of repeated role-plays paired with instructions to facilitate the behavior-shaping approach employed in social skills training.

In their review of the history of psychological skills training, O'Donohue and Krasner (1995a) note several other key influences on the development of skills training which have played a prominent role in the emergence of behavior therapy as a system of psychotherapy. These influences include the concepts of behaviorism, instrumental and classical conditioning, social role learning, and finally social learning theory, which posited that deviant or maladaptive behavior could be learned, rather than being the product of a disease, leading to interventions designed to unlearn such behaviors or, as in the case of social skills and problem solving training, to teach more adaptive behavior. By the 1970s, the critical ingredients of social skills and problem solving training had begun to be packaged, and research was underway evaluating the effects of skills training on clinical populations.

Since the "packaging" of skills training procedures into standardized approaches, social skills training and training in problem solving have become some of the most widely practiced clinical techniques in individual, group, and family psychotherapy. Indeed, aside from clinical applications, the methods employed in skills training are widely applied with nonclinical populations as well, such as in work with maritally distressed couples (Gottman & Rushe, 1995), training parenting skills (Forehand & McMahon, 1981), and teaching employment skills (Berg, Wacker, & Flynn, 1990). Thus, although the focus of this chapter is on the application of skills training procedures with clinical populations, the principles of training can be used to teach skills to any desired audience.

This chapter begins with a review of theoretical models which serve as heuristics in understanding the impact of social and problem solving skills on interpersonal competence, and the effect of skills on the course of psychiatric disorders. Next, the fundamentals of assessing social and problem solving skills are described, followed by a review of the basic procedures for training skills. Different formats for training skills are then considered, including group, individual, and family-based approaches. Following this, different clinical applications of skills training are discussed, and research supporting the effects of training social and problem solving skills is reviewed. The chapter concludes with a brief summary of social skills and problem solving training for clinical populations.

6.08.2 THEORETICAL MODELS

Two different models have been proposed for understanding the impact of social and problem solving skills on adaptive functioning and the course of psychiatric disorders: the stress–vulnerability–coping skills model and the social skills model. These two models are compatible with one another, but each addresses a somewhat different domain of functioning. The stress–vulnerability–coping model addresses the relationship between social and problem solving skills, and the severity and the course of severe psychiatric disorders such as schizophrenia and bipolar disorder. The social skills model, on the other hand, addresses the relationship between social and problem solving skills, and functional capacity, including the ability to achieve desired goals. Both of these models are described below.

6.08.2.1 The Stress–Vulnerability–Coping Skills Model

This model provides a general framework for understanding the interactions between psychobiological vulnerability, psychosocial stress, and coping skills in determining the severity and course of psychiatric illnesses (Liberman et al., 1986; Nuechterlein & Dawson, 1984; Zubin & Spring, 1977). According to the model, psychobiological vulnerability is necessary for the development of a psychiatric illness. It is determined relatively early in life by factors such as genetic loading and early environmental contributions (e.g., insults to the fetus such as *in utero* exposure to the influenza virus or obstetric complications). Stress is defined as any environmental change or set of contingencies that requires adaptation to minimize noxious effects (e.g., the loss of support from a significant other, living in a stressful home environment, stressful life events). Stress interacts with psychobiological vulnerability, increasing the chances of either developing a psychiatric disorder, precipitating the occurrence of previously dormant symptoms in an individual who already has a psychiatric illness, or worsening symptoms in a currently symptomatic individual. Coping skills are those abilities that enable an individual to buffer the negative effects of stress on psychobiological vulnerability. Coping skills operate by either eliminating the source of stress (e.g., solving a pressing problem) or by decreasing the unpleasant effects of stress (e.g., talking about feelings with friends after the death of a loved one) (Lazarus & Folkman, 1984). Social skills and problem solving skills are examples of coping skills.

The stress–vulnerability model of psychiatric disorders has important implications for treatment. Although psychobiological vulnerability is assumed to be present at a relatively early age, it can be decreased by encouraging adherence to prescribed psychotropic medications and minimizing the abuse of substances which may exacerbate psychiatric conditions (e.g., alcohol, cannabis, cocaine; Drake & Brunette, in press). The negative effects of stress on vulnerability can be minimized in three ways. First, clients' exposure to stress can be reduced by modifying the environment in which they live (e.g., decreasing negative affect in the family environment; Dixon & Lehman, 1995). Second, clients can be taught skills for minimizing the noxious effects of stress (e.g., stress management skills; Meichenbaum, 1985). Third, clients can be taught social and problem solving skills to enable them to decrease stressors that impinge upon them (Liberman, DeRisi, & Mueser, 1989). Thus, social skills and problem solving training offer the promise of decreasing symptom severity and the course of psychiatric illness by enhancing the ability of individuals to manage stress effectively.

6.08.2.2 The Social Skills Model

This model addresses the relationship between social and problem solving skills on the one hand and social competence and social adjustment on the other (Bellack, Mueser, Gingerich, & Agresta, 1997). Social competence is defined as the ability to achieve desired goals. Social adjustment refers to an individual's actual attainment of those goals, including the ability to function in different social roles (e.g., worker, parent, spouse, student), to enjoy leisure and recreational activities, and to care for oneself (Mueser, Bellack, Morrison, & Wixted, 1990). The social skills model postulates four assumptions about the relationships between social skills and problem solving skills and social functioning:

(i) social competence requires the integration of a set of component behaviors;

(ii) impairments in component skills contribute to poor social competence;

(iii) social skills are learned or are learnable;

(iv) deficits in social and problem solving skills can be rectified by skills training.

As noted by Bellack et al. (1997) and discussed further in Section 6.8.3, social skills are not the only determinent of social functioning; a wide range of other factors may also influence social adjustment, such as psychotic symptoms, environmental conditions, and mood. However, skills are postulated to be critical ingredients for interpersonal success, and deficits in these skills can be rectified through skills training techniques. While the stress–vulnerability model addresses the interface between skills and psychiatric illness, the social skills model is aimed at explaining the relationships between skills, social functioning, and goal attainment.

It is beyond the scope of this chapter to review the wealth of evidence supporting the stress–vulnerability–coping skills model and the social skills model (e.g., Trower, 1995; Yank, Bentley, & Hargrove, 1993). The practical importance of these models is that they have served as heuristics for clinicians and researchers in understanding the roles of social and problem solving skills in social functioning, and the interaction between these skills and the course of psychiatric illnesses. Furthermore, these models have resulted in testable predictions about the effects of social skills and problem solving training on social adjustment in psychiatric disorders, predictions which have been empirically supported by research reviewed later in this chapter.

6.08.3 ASSESSMENT

This section discusses: (i) the components of social and problem solving skills; (ii) the range of strategies for assessing skills; (iii) nonskill factors which can influence social competence; and (iv) strategies for integrating assessment into treatment.

6.08.3.1 Components of Social and Problem Solving Skills

Prior to the late 1970s, social skills were broadly conceptualized strictly in terms of specific behaviors such as eye contact, voice tone, and the verbal content of what is said. However, in the late 1970s and 1980s, a number of different clinical research teams independently proposed a tripartite typology of social skills, which included social perception, cognitive (or problem solving) skills, and behavioral skills (McFall, 1982; Morrison & Bellack, 1981; Trower, Bryant, & Argyle, 1978; Wallace et al., 1980). Based on this broader conceptualization of social skills, social competence is conceptualized as requiring three different types of skills.

First, in order to be effective in a social situation, the person must be able to accurately perceive relevant situational parameters, such as their relationship to the other person, whether the setting is public or private, and the other person's affective response (i.e., social perception skills). Recognition of these situational features is crucial, as they may constrain the appropriateness of social behaviors. There is ample evidence showing that social perception skills, such as the inability to accurately perceive the facial expressions of others or recognize interpersonal problems, are correlated with poorer role performance in clinical populations (e.g., Bellack et al., 1994; Hellewell & Whittaker, 1998; Mueser et al., 1996).

Second, after the relevant social information has been extracted from a situation, the individual must be able to formulate a goal, generate possible response alternatives for achieving the goal, weigh the benefits and disadvantages of each possible solution, and choose the best solution. Although these cognitive skills may occur implicitly, they have an important bearing on the success of any plan, and hence are of direct relevance to an individual's social competence. In addition to the cognitive skills generally subsumed under the rubric of "problem solving skills," other cognitive skills can also influence social competence. Abstract thinking can be crucial for an individual to grasp a concept related to solving a particular problem or achieving a desired goal. Memory impairment may interfere with social competence by rendering it more difficult for individuals to learn from past mistakes or recall critical features of situations. Cognitive functioning has been found to be a robust predictor of psychosocial adjustment, with research suggesting that the ability to cognitively process socially-oriented information (i.e., social cognition) is especially critical to success in interpersonal spheres (Penn, Corrigan, Bentall, Racenstein, & Newman, 1997).

Third, after social perception and cognitive skills have been used to appraise a situation and formulate a plan of action, behavioral skills are required to carry out the plan. Behavioral skills refer to the actual behaviors emitted in interpersonal situations that are necessary to achieve a particular goal. Broadly speaking, behavioral skills can be divided into four different areas, including nonverbal components, paralinguistic skills, verbal content, and interactive balance. Nonverbal components are behaviors such as eye contact and use of gestures that convey meaning or affect during an interaction. Paralinguistic skills correspond to the vocal characteristics of speech, such as voice tone, loudness, and inflection. Like nonverbal components, paralinguistic skills often communicate vital information during an interaction, such as the speaker's mood or underlying motives.

Verbal content refers to what is actually said, regardless of the manner in which it is said. If the verbal content is difficult to understand, bizarre, or socially offensive, it will interfere with social competence. However, in most day-to-day interactions, nonverbal and paralinguistic skills are as important or more important than verbal content in determining a person's social effectiveness. Interactive balance concerns the reciprocity between two individuals in a social interaction. For example, the amount of time each person spends speaking, and the amount of reinforcement each speaker provides to the other when the other person is speaking, can influence whether a person experiences an interaction as rewarding. Thus, the interactive balance between two persons may partly determine whether either person desires similar interactions with the other in the future. Numerous studies have documented that the social behaviors emitted in interpersonal encounters, including paralinguistic features, nonverbal skills, verbal content, and interactive balance, are less effective in clinical populations such as schizophrenia than in nonclinical populations, and are correlated with social functioning (e.g., Bellack, Morrison, Mueser, Wade, & Sayers, 1990; Bellack, Morrison, Wixted, & Mueser, 1990). Table 1 provides a summary of the different components of social skills.

6.08.3.2 Assessment Strategies

A number of different strategies can be used to assess social and problem solving skills, including self-report, reports by significant others, naturalistic observations, and role-play tests (Liberman, 1982). Each of these strategies has advantages and disadvantages, and a combination of strategies is preferable (Curran, 1979; Wallace, 1986). It is usually best in clinical practice to combine at least two or more assessment strategies for identifying specific deficits in social and problem solving skills that will be the focus of subsequent training. Different assessment strategies are briefly described below, including the advantages and disadvantages of each approach.

Self-report measures of social and problem solving skills (e.g., Connor, Dann, & Twentyman, 1982; Rathus, 1973), based on either questionnaires or interviews, have the advantage of being easy to administer and score. In addition, assessments are helpful because they provide therapists with valuable insights into clients' perceptions of their own needs. Because of these distinct advantages, the assessment of skills usually includes at least some information from the client's point of view. Although self-report information is easily obtained, it often has limited validity, especially in clients with severe mental illnesses such as schizophrenia and bipolar disorder. An additional problem is that self-reports often lack the behavioral specificity necessary to target specific situations and behaviors for skills training. For these reasons, self-reports are of limited utility in the assessment of social and problem solving skills, although they are frequently obtained as an adjunctive measure.

Table 1 Components of social skills.

- Nonverbal behaviors
  - Eye contact
  - Facial expression
  - Posture
  - Use of gestures
  - Body orientation
  - Interpersonal distance
- Paralinguistic skills
  - Loudness
  - Tone
  - Pitch
  - Affect
  - Rate of speech
  - Clarity of speech
  - Duration of utterance
- Verbal content
  - Verbal message
  - Choice of wording
  - Appropriateness of self disclosure
- Interactive balance
  - Smoothness of turn-taking
  - Use of social reinforcers (e.g., reflective listening skills)
  - Balance of time-talking

Reports by significant others offer a number of advantages over self-reports of social and problem solving skills. Significant others often directly experience the consequences of poor skills in clients, and may be able to pinpoint specific situations in which these skills are most prominent. For individuals with psychotic disorders, reports by significant others also have the advantage of not being susceptible to the cognitive distortions and denials often present in persons with these illnesses. Despite these advantages, there are a number of limitations of these reports. Significant others are privy to only certain types of social interactions, and are thus unable to inform about clients' skills in many other situations. In addition, although significant others can often identify situations in which clients experience difficulty, their reports typically lack the level of behavioral specificity necessary to target skills for training. Therefore, reports by significant others are often useful for identifying problem areas and specific situations in which clients experience difficulties, which can then be the focus of more fine-grained behavioral analysis.

After self-reports and reports by significant others have identified problematic social situations, detailed assessments of specific situations can be conducted through the use of naturalistic observations and role-play tests. Naturalistic observations can be important sources of information for several reasons. First, by their very nature such observations have a high generalizability to the social environment which clients face. Second, naturalistic observations can be easily performed in certain settings in which clinical staff have regular contact with clients. Third, direct behavioral observations provide information at a level of specificity necessary for targeting behaviors in skills training.

Of course, there are also limits to the value of naturalistic observations (Foster, Bell-Dolan, & Burge, 1988). Only certain social situations can be readily observed by others, and some important situations are almost always unobservable (e.g., intimate communications, offers to use drugs or alcohol). A related problem is that the behavior of some clients may be reactive to the presence of an observer, rendering the observation less naturalistic than intended. Despite these limitations, naturalistic observations provide important information about the social behavior of clients in "real world" settings.

A final assessment method is the use of role-play tests for the assessment of social skills. Role-play tests involve the engagement of clients in a simulated interaction which may be recorded and later rated on different dimensions of social and problem solving skills. There are several advantages of role-play tests over other assessment methods. Because role-plays can be scripted to address specific problem situations, they can be used to assess behavior across a wide range of situations, including situations in which naturalistic behavior cannot ordinarily be observed. Because of the contrived nature of role-play tests, the responses of confederates can be scripted, permitting the comparison of a client's performance both with other clients as well as over time. A final advantage of role-play tests is that there is extensive research supporting both their reliability and validity. For example, role-play tests tend to be stable over time in the absence of social skills training, and are correlated with independent measures of more naturalistically observed social behavior as well as social role performance (e.g., Bellack et al., 1990; Mueser, Bellack, Douglas, & Morrison, 1991).

Probably the most significant limitation of role-play tests is the time and effort required to administer them properly. Additional persons need to be trained and on hand to participate as confederates in role-play tests or the therapist must be able to serve as the partner. Depending on the rigor necessary for the rating of social skills, extensive training may be necessary to obtain reliable ratings of skill in role-play tests. Aside from the effort involved in conducting role-play tests, another limitation is the fact that role-plays tend to provide information about whether a client is capable of performing a requisite skill, but not necessarily whether he or she will perform that skill when an appropriate situation arises. These limitations notwithstanding, role-play tests have been shown to be useful for assessing social skills across numerous studies, and are sufficiently sensitive to the effects of social skills training to show improvements consistent with predictions (Bellack, 1979, 1983).

In summary, each of the different strategies for assessing social skills has both advantages and disadvantages, and a combination of methods is optimal in most cases. Assessment usually begins with identifying general areas of dysfunction based on interviews with clients and significant others. Information gleaned from these interviews can then be used to specify situations for more specialized assessment, employing naturalistic observations and/or role-play tests to better characterize the nature of any behavioral deficits or excesses.

6.08.3.3 Nonskill Factors which can Affect Social Competence

Although social and problem solving skills are hypothesized to be important determinants of social competence and social functioning, not all impairments in social adjustment are the result of deficits in these skills. The recognition of other factors which can also affect social functioning is critical, since successful intervention may require attention to these factors in addition to, or instead of, social skills. These factors fall under the general categories of medication side effects, mood, other psychiatric symptoms, environmental factors, and cultural mores.

A number of medication side effects can interfere with social functioning (Kane & Lieberman, 1992). For example, akinesia is a side effect of antipsychotic medications, characterized by a diminution of facial expressiveness and use of gestures. Another common side effect of antipsychotic medications is akathisia, reflected by an inability to sit still and a need to pace. Both of these medication side effects can interfere with social functioning because they may either make it difficult for the client to be sufficiently expressive (akinesia) or have a distracting influence on social interactions (akathisia).

Problems with mood, such as depression, may result in clients not using social and problem solving skills that are in their behavioral repertoire. For example, positive affect has been experimentally demonstrated to facilitate creative problem solving (Isen, Daubman, & Nowicki, 1987). Depression, reflected by a sense of futility and hopelessness, may cause a person not to use critical social or problem solving skills during an interaction, or to give up pursuing interpersonal goals altogether (Marx, Williams, & Claridge, 1992). Anxiety can interfere with a person's ability to use skills that he or she is ordinarily capable of, or may result in the avoidance of situations relevant to achieving interpersonal goals. Similarly, intense feelings of anger or hostility may impede the ability of a client to use effective social and problem solving skills.

Other psychiatric problems may also contribute to poor social competence independent of social skill. One group of symptoms that can have a negative effect on social functioning is the negative symptoms of schizophrenia, such as blunted affect, anhedonia, asociality, and paucity of speech (Andreasen, 1982). For example, clients with a diminished capacity to experience pleasure (anhedonia) may fail to initiate interactions or use relevant skills because they lack the motivation to pursue personal goals (Blanchard, Bellack, & Mueser, 1994). Another group of symptoms that can interfere with social functioning are positive symptoms such as hallucinations, delusions, and bizarre behavior. Clients with prominent positive symptoms often experience difficulties in their interpersonal relationships because their tenuous contact with reality interferes with establishing a common ground of understanding, a necessary precondition for much human communication (Chadwick, Birchwood, & Trower, 1996; Fowler, Garety, & Kuipers, 1995).

Environmental factors can have a profound impact on the likelihood that clients will use skills that are in their behavioral repertoires. Similarly, environmental factors can also interfere with the acquisition of skills during social skills training, by either limiting the opportunity clients have to use particular skills, or by not providing sufficient reinforcement for using the skills in appropriate situations. For example, in some state hospital settings, clients are reinforced by staff for assuming the "sick role" (i.e., extremely passive behavior), and attempts to break out of this mold by more goal-directed behavior may be actively discouraged (Goffman, 1961; Wing & Brown, 1970). For another example, a depressed client who lives with a domineering spouse may be actively discouraged from becoming more assertive unless that spouse is involved in, understands, and accepts the treatment plan.

Finally, cultural mores can influence both social skill and social competence. Cultures may vary in the established norms for behavior based on factors such as gender, age, and relationship to others. Behavior deemed to be "unassertive" in one culture may be viewed as "normal" and desirable in another (Sue & Sue, 1990). Awareness of the cultural norms of the groups to which clients belong is critical in order to understand cultural factors contributing to what appear to be problems in social functioning.

6.08.3.4 Integration of Assessment and Treatment

Although the assessment of social and problem solving skills serves as the cornerstone for selecting target behaviors for intervention, ongoing assessment over the course of treatment is critical for evaluating the success of intervention. Improvements in the specific social and problem solving skills targeted for treatment can be evaluated through the use of role-play tests and naturalistic observations in some settings. The larger question of whether improved social and problem solving skills translate into better role functioning is best addressed through interviews with clients and significant others. More objective information may be obtained through the selected use of standardized instruments for evaluating social functioning.

6.08.4 TRAINING TECHNIQUES

There are two basic models which are commonly used to train social skills: the motor skills model and the problem solving model (Bellack, Morrison, & Mueser, 1989).

6.08.4.1 Motor Skills Model

The primary focus of the motor skills model is on training the specific component skills necessary for successful interactions through repetition and programming the generalization of skills to a variety of real-life situations (e.g., Bellack et al., 1997). The fundamental assumption underlying the motor skills model is that overlearning specific behavioral skills through practice in both simulated (role play) and real social situations will result in these skills becoming automatic in relevant situations. Improved social competence, according to the motor skills model, develops in a fashion similar to the way in which expert performance develops over the course of extensive practice, resulting in complex skills that can be performed without contemplating the necessary steps in advance (Ericsson & Charness, 1994).

6.08.4.2 The Problem Solving Model

The problem solving model also places emphasis on the importance of repetition for acquiring necessary social skills. However, according to this model, the generalization of social skills to novel situations requires cognitive or problem solving skills to be maximally effective (e.g., Liberman et al., 1989). Real-life situations provide a multitude of different challenges, not all of which can be anticipated and prepared for in advance. The basic thesis of the problem solving model is that if clients are able to systematically apply problem solving skills in order to formulate goals and deal with obstacles or problems in social situations, their social competence will be improved. Thus, in addition to training motor skills, this model incorporates problem solving training in order to improve the generalization of social skills to novel situations, to enable clients to overcome potential obstacles to achieving goals, and to develop alternative strategies when initial ones fail.

6.08.4.3 Common Learning Principles

Both the motor skills and problem solving models employ a common set of learning principles, based mainly on instrumental (or operant) and observational (or social) learning theories (Bandura, 1969; Skinner, 1938), to train new social skills. Modeling (demonstrating a skill in a role-play) is frequently employed to familiarize clients with the basic steps of targeted skills. Verbal reinforcement is generously used to encourage effort and to draw attention to particular component skills that were performed well in a role-play situation.

Shaping refers to the reinforcement of successive approximations to a goal. Social skills require the complex integration of a number of component skills. Typically, these skills are learned gradually over many role-plays and with much practice outside of the sessions. Therefore, in order to encourage clients to keep trying, and to recognize their progress in acquiring targeted component skills, behavior needs to be shaped gradually over time by providing ample reinforcement along the way.

Generalization is the ability to transfer a skill learned in one setting to another situation. In order for social skills training to improve social functioning, clients must be able to use the skills acquired in training sessions in real-life settings. Therefore, programming the generalization of skills to client's natural living environments is an integral part of social skills training. Some of the strategies employed to facilitate generalization include community trips for clients to practice skills on their own, homework assignments, and teaching significant others (e.g., family members, staff members) to prompt clients to use skills in appropriate situations.

6.08.4.4 Basic Skills Training Techniques

The basic techniques of social skills training are outlined in Table 2 and are summarized briefly here. More information about techniques for training social skills are available in a variety of books, including McFall (1976), Trower et al. (1978), Goldstein (1982), Kelly (1982), Hargie and McCartan (1986), Liberman et al. (1989), and Bellack et al. (1997). The table describes social skills training in a group format, although the basic techniques are the same when working with individuals.

In order to teach a new social skill, a rationale must first be established for the importance of learning this skill. A combination of strategies can be used to develop the rationale, including asking questions in the Socratic style (e.g., "Why might it be helpful to express a positive feeling to someone who has just done something for you?"), providing additional reasons for the importance of a skill, and exploring the relevance of the skill to clients' personal goals and circumstances. The therapist's most immediate goal is to harness clients' motivation to learn the new skill.

After the importance of a skill has been established, the therapist discusses the specific component steps of the skill. For example, the skill of "expressing negative feelings" can be broken down into the following five component behaviors: (i) look at the person; (ii) speak in a firm voice tone; (iii) tell the person what they did to upset you; (iv) tell them how it made you feel; (v) suggest how this can be prevented from happening again in the future. The importance of each component step of the skill is discussed (e.g., it is important to look at the person so that you can be sure that you have their attention when you speak to them). After discussing the different steps of the skill, the therapist demonstrates the skill in a role-play. Role-plays are planned in advance, are usually quite brief, and are based on situations that are both highly plausible and likely to be encountered by clients. Immediately following the role-play, the therapist obtains feedback from clients about which component steps of the skill were observed and the overall effectiveness of the therapist in the role-play.

When clients have had an opportunity to observe the therapist model the skill, one client is engaged in a role-play of the same skill, usually based on the same situation. The advantage of using the same role-play situation at this point of the training is that it minimizes the amount of work the client must do in order to achieve a successful performance. Immediately following the role-play, the therapist provides positive feedback to the client about which specific steps of the skill were performed well. A critical feature of social skill training is that the therapist always provides immediate, positive, and specific feedback following each role-play. This feedback serves to encourage clients' efforts for trying to perform the skill as well as specific reinforcement for behaviors that were done especially well.

After positive feedback has been provided, the therapist provides the client with corrective feedback, conveyed in a helpful, upbeat manner.

Table 2 Steps of social skills training.

1. Establish rationale for the skill
    - Elicit reasons for learning the skill from group participants
    - Acknowledge all contributions
    - Provide additional reasons not mentioned by group members
2. Discuss the steps of the skill
    - Break the skill down into three or four steps
    - Write the steps on a board or poster
    - Discuss the reason for each step
    - Check for understanding of each step
3. Model the skill in a role-play
    - Explain that you will demonstrate the skill in a role-play
    - Plan out the role-play in advance
    - Use two leaders to model the skill
    - Keep the role-play simple
4. Review the role-play with the participants
    - Discuss whether each step of the skill was used in the role-play
    - Ask group members to evaluate the effectiveness of the role model
    - Keep the review brief and to the point
5. Engage a client in a role-play of the same situation
    - Request the client to try the skill in a role-play with one of the leaders
    - Ask the client questions to make sure he or she understands their goal
    - Instruct members to observe the client
    - Start with a client who is more skilled or is likely to be compliant
6. Provide positive feedback
    - Elicit positive feedback from group members about the client's skills
    - Encourage feedback that is specific
    - Cut off any negative feedback
    - Praise effort and provide hints to group members about good performance
7. Provide corrective feedback
    - Elicit suggestions for how client could do the skill better next time
    - Limit the feedback to one or two suggestions
    - Strive to communicate the suggestions in a positive, upbeat manner
8. Engage the client in another role-play of the same situation
    - Request that the client change one behavior in the role play
    - Check by asking questions to make sure the client understands the suggestion
    - Try to work on behaviors that are salient and changeable
9. Provide additional feedback
    - Focus first on the behavior that the client was requested to change
    - Engage client in two to four role-plays with feedback after each one
    - Use other behavior shaping strategies to improve skills such as coaching, prompting, and supplemental modeling
    - Be generous but specific when providing positive feedback
10. Assign homework
    - Give an assignment to practice the skill
    - Ask group members to identify situations in which they could use the skill
    - When possible, tailor the assignment to each client's level of skill

Rather than providing negative feedback about component skills that were performed poorly, one or two suggestions are made for how the client could improve his or her performance in another role-play. In addition to giving verbal suggestions for how to improve performance in the next role-play, the therapist can also model the skill again, drawing the client's attention to specific component behaviors that are targeted for change.

When corrective feedback has been provided, the client is engaged in a second role-play with specific instructions to modify particular component behaviors. The same role-play situation is used as in the first role-play. The second role-play is followed by the same sequence of positive specific reinforcement, with initial emphasis on the component skills targeted for change, followed by corrective feedback. Typically, a client can be engaged in two to five role-plays of a skill, depending on his or her motivation and improvement over the role-plays. If verbal instructions and praise alone are insufficient to bring about behavior change in the role-plays, the therapist may use a variety of other teaching techniques such as supplemental modeling by the therapist, coaching (i.e., whispering verbal prompts to the client during a role-play), prompting (i.e., providing the client with nonverbal cues, such as hand signals, to modify his or her behavior during a role-play) (Bellack et al., 1997), or, for severely impaired clients, attention focused training (i.e., combining verbal cues for response with primary reinforcers such as food) (Massel et al., 1991). The most critical concern when engaging a client in a series of role-plays is that he or she demonstrates some improvement in the targeted skill from the first to the last behavioral rehearsal. This is the essence of the shaping process, in which role-plays provide learning opportunities to improve performance over multiple trials.

After sufficient progress has occurred over the role-plays (and other clients have had the opportunity to practice the skill in similar role-plays), the therapist develops a homework assignment for clients to practice the skill on their own. The rationale for practicing the skill outside of the session may need to be reviewed with the client. Homework assignments are most effective when a specific situation to practice the skill can be identified by the client and therapist in advance. Possible obstacles to completing the assignment should be anticipated. When possible, significant others should be informed of the homework assignment so they may remind or prompt the client to practice the skill in appropriate situations.

The preceding sequence describes the introduction of new skills in social skills training. Usually several sessions are spent teaching one specific skill before moving onto another skill. Following the introduction to a skill, subsequent sessions begin with a review of homework, including the identification of situations where the client has unsuccessfully tried to use the skill or could have used the skill but did not. Instead of getting a description from the client of what happened, the therapist engages him or her in setting up the role-play of that situation. Following the role-play, positive and corrective feedback are provided using the principles previously described. Role-plays from more than one situation can be practiced, as well as situations that the client expects to encounter or hypothetical situations. Practicing the targeted skill across a variety of role-play situations, as well as trying the skill in real-life situations, serves as a form of generalization training. When the client has demonstrated an ability to perform the skill spontaneously in real-life situations, additional skills are introduced and trained.

6.08.4.5 Problem Solving Training

As in social skills training, training in problem solving skills involves following a specific sequence of steps (D'Zurilla & Goldfried, 1971). These steps are designed to allow consideration of as many solutions as possible and to consider what is needed in order to put a chosen plan into action. While social skills training focuses on teaching specific component behaviors necessary for effective interactions, problem solving training aims to teach a process for approaching problems and achieving goals. The process of problem solving has been likened to the scientific method (Kuhn, 1970; Popper, 1979), in which the essential task is to educate clients how to think and approach psychological problems in a more systematic, rational, and empirically based manner (Beck, 1976; Ellis, 1962). Thus, while social skills training tends to be content-oriented, training in problem solving is oriented towards teaching a set of processing skills designed to maximize goal attainment.

The basic steps of problem solving are outlined in Table 3 and are briefly described here. Further information about problem solving training can be found in a variety of books and book chapters, including Falloon, Boyd, and McGill (1984), D'Zurilla (1986), Hawton and Kirk (1989), Nezu, Nezu, and Perri (1989), and Mueser and Glynn (1995). The same steps of problem solving listed in the table are followed when teaching in an individual, group, couple, or family format.

Prior to the initiation of problem solving training, the therapist endeavors to develop in clients a problem solving "orientation" in which problems are viewed as obstacles which can be overcome or improved upon by systematically exploring and trying different response options. Similarly, achieving short- and long-term goals is construed as requiring individuals to overcome a series of obstacles which are ordered in logical sequence. The development of a problem solving orientation can be facilitated by providing examples of problems and their solutions, and reviewing the steps of problem solving. However, in the long-run, clients learn to adopt a problem solving orientation through repeated prompting, practicing the steps of problem solving on personally relevant problems, and experiencing the natural consequences of implementing effective solutions to their problems.

Problem solving begins with the identification of a problem or goal that the client wishes to resolve or achieve. As the definition of the

Table 3 Steps of problem solving.

Step 1: Define the problem
- Get different opinions about the nature of the problem
- Define the problem or goal in behaviorally specific terms
- If the problem is shared by more than one person, make sure each one agrees on the definition

Step 2: Generate possible solutions to the problem
- Brainstorm as many different solutions as possible
- Do not evaluate any solutions at this time
- Be creative and include "wild and crazy" ideas

Step 3: Evaluate the solutions
- Consider the advantages and disadvantages of each solution for solving the problem
- Systematically evaluate one solution at a time
- Avoid settling on one "best" solution before reviewing all solutions

Step 4: Select the best solution(s)
- Select the solution that seems most likely to be effective
- Consider how difficult the solution will be to implement
- Choose more than one solution if they can be easily combined

Step 5: Plan on how to implement the solution(s)
- Consider what resources are needed to implement the solution (e.g., money, skills, information)
- Anticipate possible obstacles to implementing the solution
- Establish a time frame for implementing the solution(s)
- If more than one person is involved, establish specific tasks for each person

Step 6: Review problem solving plan at a later time
- Plan a time to evaluate whether the problem solving plan was successful
- Do additional problem solving and modify the plan if the desired goal has not been obtained
- Praise all efforts at solving the problem, even if the problem is not entirely resolved

problem is crucial to the solutions that will be generated and likely success of solving it, the problem should be discussed in detail, and questions should be posed such as "Why is it a problem?," "For whom is it a problem?," and "How have you tried to solve this problem in the past and what happened?" If problem solving is conducted with others who may be invested in the problem and its solution (e.g., family members), then multiple perspectives on the problem are sought to ensure each person's involvement. When the nature of the problem or goal has been considered, the client must arrive at a specific definition of the problem. In general, the more behaviorally specific the definition, and the more circumscribed the problem, the greater the chance of solving the problem. Large, complex problems and ambitious goals are best approached by breaking them into smaller, more manageable chunks, each of which is the focus of problem solving. For couple- and family-based problem solving, all involved persons must agree on the definition of the problem in order for them to be involved in its resolution (Falloon et al., 1984).

When a problem or goal has been articulated, multiple solutions are identified for solving the problem or achieving the goal. Clients are urged to be as creative as possible at this point of problem solving, and to avoid editing, censuring, or evaluating any of the solutions that come to mind. Instead, all solutions are acknowledged, with the expectation that suggesting even bad or inappropriate solutions may lead to the identification of novel and innovative strategies. Clients who tend to be overly self-critical and punitive may require extra practice at not immediately rejecting solutions as soon as they are generated.

After a variety of possible solutions have been identified, each one is systematically evaluated in terms of its perceived effectiveness for solving the problem. This evaluation can be standardized by routinely assessing the advantages and disadvantages of each solution. Following the evaluation of solutions, the best solution or combination of solutions is selected. Often, the best solution is quite evident after the advantages and disadvantages of each solution have been considered. Sometimes no one solution is obviously best, more than one solution appears equally effective, or, when more than one person has a stake in the problem, there is disagreement as to the "best" solution. In such cases, the best solution is determined by combining different solutions, selecting more than one solution to implement, and determining which one should be tried first, or creating a new solution drawn from the previously discussed ones.

Solutions to problems can only be effective if they are implemented. Furthermore, a variety of obstacles can conspire to interfere with the implementation of a potentially effective solution to a problem. Therefore, planning on how to implement the solution is critical to successful resolution. Several factors are useful to consider when determining a plan for solving the problem. First, if more than one person is involved in solving the problem, roles for implementing the solution need to be agreed upon. Second, the resources needed to implement the solution must be evaluated, such as money, expertise, information, or skills. Role-plays may be useful at this stage to help clients develop or practice the requisite skills for enacting a solution. Third, possible obstacles to effective implementation should be explored and, if realistic obstacles are identified, tentative plans for dealing with them should be determined. Finally, a time-frame should be established for putting the different steps of the solution into action. This time-frame should include a follow-up time during which the success (or lack thereof) of the problem solving plan can be reviewed.

Although some problems are solved after a single attempt at problem solving, many are not, and it is common for repeated efforts to be required to make sufficient progress at resolving a problem or achieving a goal. An important part of developing in clients a problem solving orientation is conveying the idea that problem solving is often an iterative process that requires multiple efforts in order to secure success. In order to ensure that problem solving continues to be applied to problems that remain unsolved, it is helpful to establish follow-up times to evaluate the status of the problem or goal.

If the problem has been successfully resolved, then a new problem or goal can be targeted. On the other hand, if the problem remains, the therapist teaches the client how to systematically "debug" the problem solving plan. Finding the problem in a problem solving plan is accomplished by going through the steps of the plan in reverse order until a problem with the plan is identified, at which point the problem in the plan is corrected, and any necessary changes in the following steps are determined (Mueser & Glynn, 1995). Thus, the first step in debugging a problem solving plan is to determine if the solution was implemented as intended. If it was not, then the implementation plan needs to be altered in order to determine whether the selected solution will work. If the solution was implemented, but it did not work and the problem remains, then the second step is to re-evaluate the other possible solutions, and choose a different solution (or combination of solutions) for solving the problem. When a different solution has been selected, a new implementation plan must be made. If every solution has been tried and implemented, but the problem remains, the third step is to generate additional solutions for resolving the problem or achieving the goal, to then evaluate these new solutions, select the best one, and plan on how to implement it. Finally, if repeated attempts at solving a problem are unsuccessful and all viable solutions have been exhausted, a fourth and last step is to define the problem differently in order to increase the probability that the new problem will be more solvable than the old one.

6.08.5 FORMAT OF SOCIAL SKILLS AND PROBLEM SOLVING TRAINING

Social and problem solving skills can be taught in a variety of formats, ranging from individual psychotherapy, to group therapy, to couples or family therapy.

6.08.5.1 Group Skills Training

The group format is one of the most common formats for teaching social and problem solving skills. Skills training in groups generally follows a preplanned curriculum, which is provided over a limited time period, ranging from several months to over a year. Clients usually begin participating in the group at its initiation, and continue until the targeted skills have been acquired. Groups are often led by two therapists and comprise clients with deficits in similar areas.

There are several advantages to teaching social skills and problem solving in a group format. First, there is the obvious advantage of economy, considering that one or two therapists can work with five to eight clients simultaneously in a group setting. Second, group-based skills training provides clients with a variety of different role models, which may facilitate their acquisition of targeted skills. Third, feedback from other clients can be obtained in a group format, providing additional reinforcement for clients to practice the requisite skills. Fourth, group-based skills training can provide the opportunity for role-playing with a variety of different partners (i.e., different clients), a task which is much more difficult to accomplish when conducting individual psychotherapy. Fifth, clients often appreciate the opportunity of working with others who share similar difficulties and goals. For example, group-based social skills training can be conducted with individuals with severe mental illnesses (e.g., schizophrenia), social phobia, or poor assertiveness skills. Problem solving training in a group format offers the same advantages. Clients often enjoy being able to help each other make progress towards desired goals and appreciate knowing that they are not alone in the problems or obstacles they face. In addition, since most problems or goals that clients have are not unique, more than one client may benefit from progress made on problem solving about a particular goal.

6.08.5.2 Individual Format

Although the group format for skills training has a number of advantages, it is often not practical, especially for clinicians working in private practice. Social skills and problem solving training in an individual format follows the same basic principles as when working with groups. There are several advantages to individual-based work. By its very nature, it is easier to tailor treatment sessions to the specific needs presented by the client. There is also more time for training skills because of the exclusive focus on one client. Finally, skills training can be conducted on an *ad hoc* basis, rather than following a structured curriculum, which may be more beneficial for some clients participating in other types of psychotherapy, such as cognitive restructuring or relaxation training.

6.08.5.3 Family Format

Social skills and problem solving training can be conducted in the context of family or couples intervention. For example, behavioral marital therapy typically incorporates both training in basic communication skills for couples and problem solving skills as a strategy for minimizing conflict and increasing positive interactions (e.g., Bornstein & Bornstein, 1986; Jacobson & Margolin, 1979; Liberman, Wheeler, deVisser, Kuehnel, & Kuehnel, 1980). These interventions are based on the assumption that most of the difficulties maritally distressed couples experience are due to deficits in communication rather than fundamental differences in values and preferences.

Similarly, behavioral family therapy involves teaching basic communication skills and problem solving skills to family members (Mueser & Glynn, 1995). For example, Falloon, Boyd, and McGill (1984) developed a model in which families are taught four basic communication skills (active listening, expressing positive feelings, making positive requests, expressing negative feelings), followed by the six steps of problem solving. Families are taught communication skills following the same basic procedures used in the group format for teaching social skills, and are then given homework assignments to practice these skills on their own. Role-plays are used to assess family members' acquisition of targeted skills. When family members have demonstrated improvements in basic communication skills, they are then taught problem solving skills. Initially, the therapist leads the family to demonstrate the steps of problem solving. Subsequent to this, family members elect their own "chairman" to lead the problem solving discussion and "secretary" to record the family's problem solving efforts. The members are encouraged to have weekly meetings to practice their problem solving skills. Over time, as the family members' skills improve, increasingly more difficult problems are tackled, including problems that may be the source of major conflict between family members.

It is also possible to combine the group and family formats into a multiple family group format. For example, McFarlane (1990) has developed a model of multiple family group intervention for persons with schizophrenia and their relatives. In groups that are held every two weeks, problem solving is taught and regularly practiced in order to identify effective solutions to common problems, thereby reducing the risk of relapse and rehospitalization (McFarlane et al., 1995).

6.08.6 CLINICAL APPLICATIONS AND RESEARCH

Social skills and problem solving training have been applied to very wide range of psychiatric disorders and interpersonal difficulties. In fact, there are few areas of psychopathology or interpersonal dysfunction for which skills training approaches have not been developed. For example, skills training approaches have been developed to help people with problems such as anger management, interpersonal shyness and dating anxiety, poor assertiveness, conflict resolution, difficulties with interpersonal relationships on the job, marital discord, and dealing with social situations involving substance abuse. Although social and problem solving skill interventions generally focus on improving interpersonal competence in specific domains of functioning, programs are often targeted at groups of clients with specific diagnoses, including major depression, schizophrenia, borderline personality disorder, and social phobia, as well as individuals with developmental disabilities.

Table 4 summarizes clinical applications of social and problem solving skills training to different problem areas and populations. Many of the programs described in this table have been empirically validated in controlled studies. The research literature supporting the efficacy of social and problem solving training has grown rapidly over the last two decades, and these approaches now enjoy some of the best empirical support of all psychological interventions. It is beyond the scope of this chapter to critically review research on the effects of social skills training and problem solving training. However, numerous books and review articles are available that critically evaluate these different training programs, including Bedell and Lennox (1996), O'Donahue and Krasner (1995a), Dilk and Bond (1996), Benton and Schroeder (1990), Nezu and Nezu (1989), and Smith, Bellack, and Liberman (1996).

6.08.6.1 Utility of Social and Problem Solving Skills Training

From a treatment planning perspective, it is important to consider when skills training is a treatment on its own and when it best serves as an adjunct to other approaches to therapy. To some extent, the answer to this question depends on whether skills training is used to treat a complex constellation of behaviors and symptoms (e.g., a psychiatric disorder) or a more specific and narrower class of situations and behaviors. There have been numerous controlled studies of skills training for schizophrenia, with some evidence documenting benefits in the areas of social functioning, symptom severity, and relapses and rehospitalizations (e.g., Bellack, Turner, Hersen & Luber, 1984; Hogarty et al., 1991; Marder et al., 1996). Schizophrenia is a complex illness involving impairments across multiple domains, including social adjustment, cognitive functioning, and symptomatology. When skills training is conducted with persons with schizophrenia, it is provided in the context of a comprehensive treatment program (Bellack & Mueser, 1993), which may include a veriety of other interventions such as medication and symptom monitoring, case management, family psychoeducation, and vocational rehabilitation.

With respect to major depression, both social skills training and problem solving training have been shown to be effective in reducing severity of depression without the use of other interventions (Bellack, Hersen, & Himmelhoch, 1983; Nezu & Perri, 1989). However, in contrast to schizophrenia, where relatively few psychological interventions have been established to be clinically effective, a wide variety of therapeutic approaches appear to result in improvements in depression comparable to those produced by social and problem solving skills training, such as cognitive therapy and scheduling pleasant events (Dobson, 1989; Zeiss, et al., 1979). Therefore, skills training can be used as a primary treatment approach for depression, although its effects can be expected to be similar to those of other psychological interventions. In clincal practice, skills training is most often provided as an adjunctive intervention for some clients, in combination with cognitive or interpersonal therapy, rather than as the sole intervention.

Similar in some respects to its use in major depression, social skills training is frequently employed as an adjunctive strategy in the treatment of social phobia. For example, in a program for social phobia developed by Heimberg et al. (1990), cognitive therapy and exposure are combined with social skills training (role playing) in order to provide clients with feedback about their behavior, to challenge their distorted perceptions about their own social behavior, and to encourage them not to avoid feared social situations. Although some research suggests that cognitive therapy may not reduce social anxiety above and beyond that provided by the combination of social skills training and exposure (Hope, Heimberg, & Bruch, 1996; Stravynski & Shahar, 1983), other research indicates that cognitive therapy alone may be effective (Emmelkamp, Mersch, Vissia, & van der Helm 1985). However, in practice no single treatment approach is usually provided for social phobia; rather, most treatment programs involve a combination of strategies, including social skills training, exposure, and cognitive therapy (e.g., Heimberg et al., 1990; Turner, Beidel, & Cooley, 1994).

Social and problem solving skills training can be effective as treatments on their own (or in combination with training in other self-regulatory skills, such as stress inoculation, relaxation, cognitive restructuring) when they are provided to address a specific problem (or class of problems), often within a specific population (e.g., clients with developmental disabilities, psychiatric disorders). There are numerous studies in the research literature (some included in Table 4) of skills training interventions demonstrating superior outcomes compared with other treatment or waitlist (no treatment) comparison groups. For example, skills training has been found to be useful in decreasing problems related to anger (Benson, Rice, & Miranti, 1986; Deffenbacher, 1988), improving social anxiety, loneliness, and dating skills (e.g., Fox, McMorrow, Storey, & Rogers, 1984; MacDonald, Lindquist, Kramer, McGrath, & Rhyne, 1975), improving assertiveness (Gambrill, 1995), enhancing job-related skills, such as

Table 4 Examples of clinical applications of social and problem solving skills training.

Program developers	*Target population*	*Training format*	*Focus of training*[a]
Alberti and Emmons (1990)	Unassertive persons	Group	SST for skills to help people "stand up for their rights" and express feelings directly
Becker, Heimberg, and Bellack (1987)	Persons with depression	Individual	SST for interpersonal skills
Benson (1991)	Mentally retarded adults	Group	SST and PS to teach strategies for managing anger and interpersonal conflict
Falloon et al. (1984) Miklowitz and Goldstein (1997) Mueser and Glynn (1995)	Families of persons with severe psychiatric disorders	Family	SST for basic communication skills and PS to help families solve problems and achieve goals
Fisher and Carstensen (1990)	Elderly nursing home residents	Group	SST to improve conversational and other social skills in elderly persons
Heimberg et al. (1990)	Socially anxious clients	Group	Exposure to social situations, cognitive restructuring, and SST for peer intentions
Jacobson and Margolin (1979)	Maritally distressed couples	Couples	SST for communication skills and PS to address problem areas
Kelly (1995)	Persons with high HIV risk behavior	Group	SST to reduce HIV risk behaviors
Kleiner, Marshall, and Spevack (1987)	Agoraphobics	Individual	PS and exposure to feared situations
Liberman et al. (1989) Bellack et al. (1997)	Clients with schizophrenia	Group (can be adapted for individuals)	SST for communication skills, resolving conflicts, making friends, dealing with care providers, recreational and leisure activities
Linehan (1993)	Borderline personality disorder clients	Group and individual	SST and PS to address interpersonal skills, emotion regulation, and distress tolerance
Maiuro (1991)	Domestically violent men	Group	PS to reduce domestic violence
Monti et al. (1989)	Alcoholics	Group	SST for conversational skills, expressing feelings, dealing with alcohol-related situations
Mynors-Wallis, Davies, Gray, Barbour, and Gatz (1997)	Primary care clients	Individual	PS to address anxiety and depression problems interfering with functioning
Nezu, Nezu, and Perri (1989)	Persons with major depression	Individual	PS to address interpersonal problems and goals related to depression
Novaco(1975)	Persons with anger problems	Individual	Stress management skills, PS to identify suitable alternatives in conflict situations, SST for dealing with provoking situations
Salkovskis, Atha, and Storer (1990)	Suicidal persons	Individual	PS to identify problems and arrive at solutions that are alternatives to self-injurious behaviors
Valenti-Hein and Mueser (1990)	Mentally retarded adults	Group	SST and PS to develop dating skills (e.g., conversational skills, asking for a date, resisting persuasion)
Waldo, Roath, Levine, and Freedman (1987)	Mothers with schizophrenia	Group	SST to teach parenting skills
Wong and Woolsey (1989)	Chronic psychotic inpatients	Individual	SST to re-establish conversational skills in severely impaired clients

[a] SST = social skills training; PS = Problem solving training.

interviewing, managing interactions with supervisors, and peer relationships (e.g., Hughes & Rusch, 1989; Mueser, Foy, & Carter, 1986), and dealing with social situations involving alcohol and drug abuse (e.g., Hawkins, Catalano, Gillmore, & Wells, 1989). Thus, social and problem solving skills training are most often combined with other therapeutic modalities when used in the treatment of psychiatric syndromes or disorders, whereas skills training can be used on its own when the goals are to improve specific areas of functioning related to specific types of social situations.

6.08.7 SUMMARY AND CONCLUSIONS

Over the past several decades, social skills and problem solving training have become some of the most widely practiced interventions for the treatment of psychological disorders in adults. Therapies designed to improve social and problem solving skills are based on the assumptions that clients are capable of learning more adaptive interpersonal and self-management skills, and that these skills are most effectively taught in a systematic fashion, employing the principles of social learning theory. Training in social and problem solving skills can be conducted in a variety of different formats, such as with individuals, groups, couples, or families. Skills training approaches have enjoyed success across a broad range of clinical problems, including depression, poor social functioning in schizophrenia, social anxiety, anger, marital distress, and families coping with severe psychiatric disorders. Expertise in teaching social and problem solving skills is an important tool for clinical psychologists' armamentarium of treatment techniques.

6.08.8 REFERENCES

Alberti, R. E., & Emmons, M. L. (1970). *Your perfect right*. San Luis Obispo, CA: Impact Press.

Andreasen, N. C. (1982). Negative symptoms in schizophrenia: Definition and reliability. *Archives of General Psychiatry, 39,* 784–788.

Bandura, A. (1969). *Principles of behavior modification*. New York: Holt, Rinehart and Winston.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.

Becker, R. E., Heimberg, R. G., & Bellack, A. S. (1987). *Social skills training treatment for depression*. New York: Pergamon.

Bedell, J. R., & Lennox, S. S. (1996). *Handbook for communication and problem-solving skills training*. New York: Wiley.

Bellack, A. S. (1979). A critical appraisal of strategies for assessing social skill. *Behavior Assessment, 1,* 157–176.

Bellack, A. S. (1983). Recurrent problems in the behavioral assessment of social skill. *Behaviour Research and Therapy, 21,* 29–42.

Bellack, A. S., & Hersen, M. (Eds.) (1979). *Research and practice in social skills training*. New York: Plenum.

Bellack, A. S., Hersen, M., & Himmelhoch, J. M. (1983). A comparison of social skills training, pharmacotherapy, and psychotherapy for depression. *Behaviour Research and Therapy, 21,* 101–107.

Bellack, A. S., Morrison, R. L., & Mueser, K. T. (1989). Social problem solving in schizophrenia. *Schizophrenia Bulletin, 15,* 101–116.

Bellack, A. S., Morrison, R. L., Mueser, K. T., Wade, J. H., & Sayers, S. L. (1990). Role play for assessing the social competence of psychiatric patients. *Psychological Assessment, 2,* 248–255.

Bellack, A. S., Morrison, R. L., Wixted, J. T., & Mueser, K. T. (1990). An analysis of social competence in schizophrenia. *British Journal of Psychiatry, 156,* 809–818.

Bellack, A. S., & Mueser, K. T. (1993). Psychosocial treatment for schizophrenia. *Schizophrenia Bulletin, 19,* 317–336.

Bellack, A. S., Mueser, K. T., Gingerich, S., & Agresta, J. (1997). *Social skills training for schizophrenia: A step-by-step guide*. New York: Guilford Press.

Bellack, A. S., Sayers, M., Mueser, K. T., & Bennett, M. (1994). An evaluation of social problem solving in schizophrenia. *Journal of Abnormal Psychology, 103,* 371–378.

Bellack, A. S., Turner, S. M., Hersen, M., & Luber, R. F. (1984). An examination of the efficacy of social skills training for chronic schizophrenic patients. *Hospital and Community Psychiatry, 35,* 1023–1028.

Benson, B. A. (1991). *Teaching anger management to persons with MR/DD*. Worthington, OH: IDS Publishing.

Benson, B. A., Rice, C. J., & Miranti, S. V. (1986). Effects of anger management training with mentally retarded adults in group treatment. *Journal of Consulting and Clinical Psychology, 54,* 728–729.

Benton, M. K., & Schroeder, H. E. (1990). Social skills training with schizophrenics: A meta-analytic evaluation. *Journal of Consulting and Clinical Psychology, 58,* 741–747.

Berg, W. K., Wacker, D. P., & Flynn, T. H. (1990). Teaching generalization and maintenance of work behavior. In F. R. Rusch (Ed.), *Supported employment: Models, methods, and issues* (pp. 145–160). Sycamore, IL: Sycamore Publishing Company.

Blanchard, J. J., Bellack, A. S., & Mueser, K. T. (1994). Affective and social-behavioral correlates of physical and social anhedonia in schizophrenia. *Journal of Abnormal Psychology, 103,* 719–728.

Bornstein, P. H., & Bornstein, M. T. (1986). *Marital therapy: A behavioral communications approach*. New York: Pergamon.

Chadwick, P., Birchwood, M., & Trower, P. (1996). *Cognitive therapy for delusions, voices and paranoia*. Chichester, UK: Wiley.

Connor, J., Dann, L., & Twentyman, C. (1982). A self-report measure of assertiveness in young adolescents. *Journal of Clinical Psychology, 38,* 101–106.

Curran, J. P. (1979). Pandora's Box reopened? The assessment of social skill. *Journal of Behavioral Assessment, 1,* 55–72.

Davis, H. (1985). Training professionals in behavior modification. *British Journal of Medical Psychology, 58,* 241–248.

Davis, H., & Butcher, P. (1985). Sharing psychological skills. *British Journal of Medical Psychology, 58,* 207–216.

Deffenbacher, J. L. (1988). Cognitive-relaxation and social skills treatments of anger: A year later. *Journal of Counseling Psychology, 35,* 414–418.

Dilk, M. N., & Bond, G. R. (1996). Meta-analytic evaluation of skills training research for persons with severe mental illness. *Journal of Consulting and Clinical Psychology, 64,* 1337–1346.

Dixon, L. B., & Lehman, A. F. (1995). Family interventions for schizophrenia. *Schizophrenia Bulletin, 21,* 631–643.

Dobson, K. S. (1989). A meta-analysis of the efficacy of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 57,* 414–419.

Drake, R. E., & Brunette, M. F. (in press). Complications of severe mental illness related to alcohol and drug use disorders. In M. Galanter (Ed.), *Recent developments in alcoholism* (Vol. 14): *The consequences of alcohol.* New York: Plenum.

D'Zurilla, T. J. (1986). *Problem-solving therapy: A social competence approach to clinical intervention.* New York: Springer.

D'Zurilla, T. J., & Goldfried, M. R. (1971). Problem solving and behavior modification. *Journal of Abnormal Psychology, 78,* 107–126.

Ellis, A. (1962). *Reason and emotion in psychotherapy.* New York: Lyle Stuart.

Emmelkamp, P. M. G., Mersch, P. P., Vissia, E., & van der Helm, M. (1985). Social phobia: A comparative evaluation of cognitive and behavioral interventions. *Behaviour Research and Therapy, 23,* 365–369.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49,* 725–747.

Falloon, I. R. H., Boyd, J. L., & McGill, C. W. (1984). *Family care of schizophrenia: A problem-solving approach to the treatment of mental illness.* New York: Guilford Press.

Fisher, J. E., & Carstensen, L. L. (1990). Generalized effects of skills training among older adults. *Mental Health in the Nursing Home, 9,* 91–107.

Forehand, R., & McMahon, R. J. (1981). *Helping the noncompliant child: A clinician's guide to effective parent training.* New York: Guilford Press.

Foster, S. L., Bell-Dolan, D. J., & Burge, D. A. (1988). Behavioral observation. In A. S. Bellack & M. Hersen (Eds.), *Behavioral assessment: A practical handbook* (3rd ed., pp. 119–160). New York: Oxford University Press.

Fowler, D., Garety, P., & Kuipers, E. (1995). *Cognitive behaviour therapy for psychosis: Theory and practice.* Chichester, UK: Wiley.

Fox, R. M., McMorrow, M. J., Storey, K., & Rogers, B. M. (1984). Teaching social/sexual skills to mentally retarded adults. *American Journal of Mental Deficiency, 89,* 9–14.

Gambrill, E. (1995). Assertion training. In W. O'Donohue & L. Krasner (Eds.), *Handbook of psychological skills training: Clinical techniques and applications* (pp. 81–118). Boston: Allyn & Bacon.

Goffman, E. (1961). *Asylums: Essays on the social situation of mental patients and other inmates.* Garden City, NY: Anchor.

Goldstein, A. P. (1982). *Psychological skill training: The structured learning technique.* New York: Pergamon.

Gottman, J., Rushe, R. (1995). Communication and social skills approaches to treating ailing marriages: A recommendation for a new marital therapy called "minimal marital therapy". In W. O'Donohue & L. Krasner (Eds.), *Handbook of psychological skills training: Clinical techniques and applications* (pp. 287–305). Boston: Allyn & Bacon.

Hargie, O., & McCartan, P. J. (1986). *Social skills and psychiatric nursing.* London: Croom Helm.

Hawkins, J. D., Catalano, R. F., Jr., Gillmore, M. R., & Wells, E. A. (1989). Skills training for drug abusers: Generalization, maintenance, and effects on drug use. *Journal of Consulting and Clinical Psychology, 57,* 559–563.

Hawton, K., & Kirk, J. (1989). Problem-solving. In K. Hawton, P. M. Salkovskis, & J. W. Kirk (Eds.), *Cognitive-behavioural treatment of adult psychological disorders: A practical guide* (pp. 406–426). Oxford: Oxford University Press.

Heimberg, R. G., Dodge, C. S., Hope, D. A., Kennedy, C. R., Zollo, L. J., & Becker R. E. (1990). Cognitive behavioral group treatment for social phobia: Comparison with a credible placebo control. *Cognitive Therapy and Research, 14,* 1–23.

Hellewell, J. S. E., & Whittaker, J. F. (1998). Affect perception and social knowledge in schizophrenia. In K. T. Mueser & N. Tarrier (Eds.), *Handbook of social functioning in schizophrenia* (pp. 197–212). Needham Heights, MA: Allyn & Bacon.

Hogarty, G. E., Anderson, C., Reiss, D., Kornblith, S., Greenwald, D., Ulrich, R., & Carter, M. (1991). Family psychoeducation, social skills training, and maintenance chemotherapy in the aftercare treatment of schizophrenia, II: Two year effects of a controlled study on relapse and adjustment. *Archives of General Psychiatry, 48,* 340–347.

Hollin, C. R., & Trower, P. (Eds.) (1986). *Handbook of social skills training.* Oxford: Pergamon.

Hope, D. A., Heimberg, R. G., & Bruch, M. A. (1995). Dismantling cognitive-behavioral therapy for social phobia. *Behaviour Research and Therapy, 33,* 637–650.

Hughes, C., & Rusch, F. R. (1989). Teaching supported employees with severe mental retardation to solve problems. *Journal of Applied Behavior Analysis, 22,* 365–372.

Isen, A. M., Daubman, K. A., & Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology, 52,* 1122–1131.

Jacobson, N. S., & Margolin, G. (1979). *Marital therapy: Strategies based on social learning and behavior exchange principles.* New York: Brunner/Mazel.

Kane, J. M., & Lieberman, J. A. (Eds.) (1992). *Adverse effects of psychotropic drugs.* New York: Guilford Press.

Kelly, J. A. (1982). *Social skills training: A practical guide for interventions.* New York: Springer.

Kelly, J. A. (1995). *Changing HIV risk behavior.* New York: Guilford.

Kleiner, L., Marshall, W. L., & Spevack, M. (1987). Training in problem-solving and exposure treatment for agoraphobics with panic attacks. *Journal of Anxiety Disorders, 1,* 219–238.

Kleinmuntz, B. (Ed.) (1966). *Problem solving: Research, method, and theory.* New York: Wiley.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed). Chicago: University of Chicago Press.

L'Abate, L., & Milan, M. (Eds.) (1985). *Handbook of social skills training and research.* New York: Wiley.

Larson, D. (1984). *Teaching psychological skills: Models for giving psychology away.* Monterey, CA: Brooks/Cole.

Lazarus, A. A. (1966). Behavior rehearsal vs. nondirective therapy vs. advice in effecting behavior change. *Behaviour Research and Therapy, 4,* 209–212.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping.* New York: Springer.

Liberman, R. P. (1982). Assessment of social skills. *Schizophrenia Bulletin, 8,* 62–82.

Liberman, R. P., DeRisi, W. J., & Mueser, K. T. (1989). *Social skills training for psychiatric patients.* Needham Heights, MA: Allyn & Bacon.

Liberman, R. P., Mueser, K. T., Wallace, C. J., Jacobs, H. E., Eckman, T., & Massel, H. K. (1986). Training skills in the psychiatrically disabled: Learning coping and competence. *Schizophrenia Bulletin, 12,* 631–647.

Liberman, R. P., Wheeler, E. G., deVisser, L. A. J. M., Kuehnel, J., & Kuehnel, T. (1980). *Handbook of marital therapy: A positive approach to helping troubled relationships.* New York: Plenum.

Linehan, M. M. (1993). *Skills training manual for treating borderline personality disorder*. New York: Guilford Press.

MacDonald, M. L., Lindquist, C. U., Kramer, J. A., McGrath, R. A., & Rhyne, L. D. (1975). Social skills training: Behavioral rehearsal in groups and dating skills. *Journal of Counseling Psychology, 22,* 224–230.

Maiuro, R. (1991). The evolution and treatment of anger and hostility in domestically violent men. *Revista Intercontinental de Psicologia y Educacion, 4,* 165–189.

Marder, S. R., Wirshing, W. C., Mintz, J., McKenzie, J., Johnston, K., Eckman, T. A., & Johnston-Cronk, K. (1996). Two-year outcome of social skills training and group psychotherapy for outpatients with schizophrenia. *American Journal of Psychiatry, 153,* 1585–1592.

Marx, E. M., Williams, J. M. G., & Claridge, G. C. (1992). Depression and social problem solving. *Journal of Abnormal Psychology, 101,* 78–86.

Massel, H. K., Corrigan, P. W., Liberman, R. P., & Milan, M. A. (1991). Conversational skills training of thought-disordered schizophrenic patients through attention focusing. *Psychiatry Research, 38,* 51–61.

McFall, R. M. (1976). *Behavioral training: A skill-acquisition approach to clinical problems*. Morristown, NJ: General Learning Press.

McFall, R. M. (1982). A review and reformulation of the concept of social skills. *Behavioral Assessment, 4,* 1–33.

McFarlane, W. R. (1990). Multiple family groups and the treatment of schizophrenia. In M. I. Herz, S. J. Keith, & J. P. Docherty (Eds.), *Handbook of schizophrenia* (Vol. 4): *Psychosocial treatment of schizophrenia* (pp. 167–189). Amsterdam: Elsevier.

McFarlane, W. R., Lukens, E., Link, B., Dushay, R., Deakins, S. A., Newmark, M., Dunne, E. J., Horen, B., & Toran, J. (1995). Multiple-family groups and psychoeducation in the treatment of schizophrenia. *Archives of General Psychiatry, 52,* 679–687.

Meichenbaum, D. (1985). *Stress innoculation training*. New York: Pergamon.

Miklowitz, D. J., & Goldstein, M. J. (1997). *Bipolar disorder: A family focused treatment approach*. New York: Guilford Press.

Monti, P. M., Abrams, D. B., Kadden, R. M., & Cooney, N. L. (1989). *Treating alcohol dependence*. New York: Guilford Press.

Morrison, R. L., & Bellack, A. S. (1981). The role of social perception in social skill. *Behavior Therapy, 12,* 69–79.

Mueser, K. T., Bellack, A. S., Douglas, M. S., & Morrison, R. L. (1991). Prevalence and stability of social skill deficits in schizophrenia. *Schizophrenia Research, 5,* 167–176.

Mueser, K. T., Bellack, A. S., Morrison, R. L., & Wixted, J. T. (1990). Social competence in schizophrenia: Premorbid adjustment, social skill, and domains of functioning. *Journal of Psychiatric Research, 24,* 51–63.

Mueser, K. T., Doonan, R., Penn, D. L., Blanchard, J. J., Bellack, A. S., Nishith, P. & deLeon, J. (1996). Emotion recognition and social competence in chronic schizophrenia. *Journal of Abnormal Psychology, 105,* 271–275.

Mueser, K. T., Foy, D. W., & Carter, M. J. (1986). Social skills training for job maintenance in a psychiatric patient. *Journal of Counseling Psychology, 33,* 360–362.

Mueser, K. T., & Glynn, S. M. (1995). *Behavioral family therapy for psychiatric disorders*. Needham Heights, MA: Allyn & Bacon.

Mynors-Wallis, L., Davies, I., Gray, A., Barbour, F., & Gath, D. (1997). A randomized controlled trial and cost analysis of problem-solving treatment for emotional disorders given by community nurses in primary care. *British Journal of Psychiatry, 170,* 113–119.

Nezu, A. M. & Nezu, L. M. (Eds.) (1989). *Clinical decision making in behavior therapy: A problem solving perspective*. Champaign, IL: Research Press.

Nezu, A. M., Nezu, C. M., & Perri, M. G. (1989). *Problem-solving therapy for depression: Theory, research, and clinical guidelines*. New York: Wiley.

Nezu, A. M., & Perri, M. G. (1989). Social problem-solving therapy for unipolar depression: An initial dismantling investigation. *Journal of Consulting and Clinical Psychology, 57,* 408–413.

Nuechterlein, K. H., & Dawson, M. E. (1984). A heuristic vulnerability/stress model of schizophrenic episodes. *Schizophrenia Bulletin, 10,* 300–312.

Novaco, R. W. (1975). *Anger control: The development and evaluation of an experimental treatment*. Lexington, MA: D. C. Heath.

O'Donohue, W., & Krasner, L. (1995a). Psychological skills training. In W. O'Donohue & L. Krasner (Eds.), *Handbook of psychological skills training: Clinical techniques and applications* (pp. 1–19). Boston: Allyn & Bacon.

O'Donohue, W., & Krasner, L. (Eds.) (1995b). *Handbook of psychological skills training: Clinical techniques and applications*. Boston: Allyn & Bacon.

Penn, D. L., Corrigan, P. W., Bentall, R. P., Racenstein, J. M., & Newman, L. (1997). Social cognition in schizophrenia. *Psychological Bulletin, 121,* 114–132.

Popper, K. R. (1979). *Objective knowledge*. Oxford, UK: Oxford University Press.

Rathus, S. A. (1973). A 30-item schedule for assessing assertive behavior. *Behavior Therapy, 4,* 398–406.

Salkovskis, P. M., Atha, C., & Storer, D. (1990). Cognitive-behavioural problem solving in the treatment of patients who repeatedly attempt suicide: A controlled trial. *British Journal of Psychiatry, 157,* 871–876.

Salter, A. (1949). *Conditioned reflex therapy*. New York: Farrar, Strauss.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century-Crofts.

Smith, T. E., Bellack, A. S., & Liberman, R. P. (1996). Social skills training for schizophrenia: Review and future directions. *Clinical Psychology Review, 16,* 599–617.

Stravynski, A., & Shahar, A. (1983). The treatment of social dysfunction in nonpsychotic psychiatric outpatients: A review. *Journal of Nervous and Mental Disease, 171,* 721–728.

Sue, D. W., & Sue, D. (1990). *Counseling the culturally different: Theory and practice* (2nd ed.). New York: Wiley.

Trower, P. (1995). Adult social skills: State of the art and future directions. In W. O'Donohue & L. Krasner (Eds.), *Handbook of psychological skills training: Clinical techniques and applications* (pp. 54–80). Boston: Allyn & Bacon.

Trower, P., Bryant, B. M., & Argyle, M. (1978). *Social skills and mental health*. London: Methuen.

Turner, S. M., Beidel, D. C., & Cooley, M. R. (1994). *Social effectiveness therapy: A program for overcoming social anxiety and phobia*. Mount Pleasant, SC: Turndel.

Valenti-Hein, D., & Mueser, K. T. (1990). *The dating skills program: Teaching social–sexual skills to adults with mental retardation*. Orland Park, IL: International Diagnostic Systems.

Waldo, M. C., Roath, M., Levine, W., & Freedman, R. (1987). A model program to teach parenting skills to schizophrenic mothers. *Hospital and Community Psychiatry, 38,* 1110–1112.

Wallace, C. J. (1986). Functional assessment in rehabilitation. *Schizophrenia Bulletin, 12,* 604–630.

Wallace, C. J., Nelson, C. J., Liberman, R. P., Aitchison, R. A., Lukoff, D., Elder, J. P., & Ferris, C. (1980). A review and critique of social skills training with schizophrenic patients. *Schizophrenia Bulletin, 6,* 42–63.

Wing, J. K., & Brown, G. W. (1970). *Institutionalism and* *schizophrenia*. Cambridge, UK: Cambridge University Press.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press.

Wong, S. E., & Woolsey, J. E. (1989). Re-establishing conversational skills in overtly psychotic, chronic schizophrenic patients: Discrete trials training on the psychiatric ward. *Behavior Modification, 13,* 415–430.

Yank, G. R., Bentley, K. J., & Hargrove, D. S. (1993). The vulnerability–stress model of schizophrenia: Advances in psychosocial treatment. *American Journal of Orthopsychiatry, 63,* 55–69.

Zeiss, A. M., Lewinsohn, P. M., & Munoz, R. F. (1979). Nonspecific improvement effects in depression using interpersonal skills training, pleasant activity schedule, or cognitive training. *Journal of Consulting and Clinical Psychology, 45,* 543–551.

Zubin, J., & Spring, B. (1977). Vulnerability: A new view of schizophrenia. *Journal of Abnormal Psychology, 86,* 103–123.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.09

Arousal Reduction Methods: Relaxation, Biofeedback, Meditation, and Hypnosis

GRAHAM C. H. TURPIN and MICHAEL HEAP
University of Sheffield, UK

6.09.1 INTRODUCTION 203
6.09.2 A THEORETICAL OVERVIEW OF AROUSAL DYSFUNCTION AND AROUSAL MODIFICATION INTERVENTIONS 204
6.09.2.1 Arousal as a Hypothetical Construct 204
6.09.2.2 Arousal Dysfunction and Health 206
6.09.2.3 Theoretical Basis of Arousal Modification Methods 208
6.09.2.4 Summary 210
6.09.3 AROUSAL REDUCTION METHODS: GENERAL PROCEDURES AND SPECIFIC TECHNIQUES 210
6.09.3.1 General Procedural Issues 211
6.09.3.1.1 Assessment 211
6.09.3.1.2 Therapeutic rationale 211
6.09.3.1.3 Practical considerations 212
6.09.3.1.4 Presentation of the technique 212
6.09.3.1.5 Evaluation of procedural factors 212
6.09.3.2 Specific Techniques 213
6.09.3.2.1 Relaxation-based methods 213
6.09.3.2.2 Biofeedback 214
6.09.3.2.3 Meditation 215
6.09.3.2.4 Hypnosis 215
6.09.3.3 Indications, Side Effects, and Contraindications 216
6.09.3.3.1 Client selection 216
6.09.3.3.2 Side effects and adverse reactions 217
6.09.3.3.3 Contraindications and procedural modifications 217
6.09.4 COMPARATIVE OUTCOMES AND THERAPEUTIC MECHANISMS 219
6.09.4.1 Overview of Outcomes 219
6.09.4.2 Therapeutic Mechanisms 221
6.09.5 CONCLUSION 223
6.09.6 REFERENCES 223

6.09.1 INTRODUCTION

Heightened arousal, elevated somatic tension, increases in anxiety and worry, are common consequences of many everyday life experiences, and are also frequently associated with more serious mental health problems such as panic, obsessional disorders, and psychosis. It is not surprising, therefore, that some of the earliest proposed psychological treatments (e.g., Jacobson, 1938), have targeted arousal reduction. Indeed, many of these techniques (e.g., meditation, hypnosis) have their origins in nonscientific or religious practices and customs. Because of this, relaxation, biofeedback, meditation, and hypnosis are frequently regarded, particularly within the USA, as alternative or complementary therapies within medicine and are provided by specialist practitioners such as stress counsellors, hypnotherapists, and biofeedback practitioners. A recent review of "unconventional medicine" practised within the USA (Eisenberg et al., 1993) revealed, by means of a telephone survey of 1539 adults, that one in three respondents reported using at least one unconventional therapy in the past year, and a third of these had engaged in an average of 19 visits to a practitioner. Extrapolating from this pattern of contact to the entire population, the authors concluded that the extent of usage of unconventional therapies exceeded the number of primary care consultations and the economic cost ($13.7 billion) was comparable to out-of-pocket expenditure for all hospitalizations within the USA. Relaxation (13%), chiropracty (10%), and massage (7%) accounted for the most frequently used therapies, and back problems (36%), anxiety (28%), headaches (27%), chronic pain (26%), and cancer (24%) represented the most common health problems for which people had sought unconventional therapies. Although reliance on a telephone survey may have overestimated the prevalence of these techniques, Eisenberg et al. (1993, p. 251) conclude that these therapies have "an enormous presence in the US health care system." They further concluded that they represent a sizeable expenditure in healthcare, are used largely as adjuncts to conventional medical interventions, and are commonly undertaken independently of consultation with the medical practitioner responsible for conventional treatment of the presenting problem.

The above review included a wide range of treatments unrelated to the focus of this particular chapter. Nevertheless, relaxation (13%), imagery (4%), biofeedback (1%), and hypnosis (1%) were reported as being widely used within a 12-month period for a range of psychological problems (e.g. anxiety, depression, insomnia, and headache) which are frequently referred to a clinical or health psychologist, and this raises a number of intriguing questions. For example, what is the efficacy of these techniques, how severe are the problems for which these therapies are commonly employed, how effectively are these therapies delivered, what training and levels of competence are associated with the practitioners responsible for delivery of these therapies, and what proportion of therapists identified within Eisenberg et al.'s survey are clinical psychologists? Many of these questions go far beyond the scope of this chapter, since it is our intention to focus only on those specific therapies that are delivered as part of a formally constructed and scientifically evaluated package of psychological treatment. Nevertheless, clinical psychologists have to be aware of a wider context, whereby clients and their referral sources already have access to and are familiar with these treatments, and which goes far beyond that formally delivered within a clinic setting. It is likely, therefore, that such experiences may well influence how the efficacy and appropriateness of many of these therapies are judged, irrespective of any formal scientific evaluation of their effectiveness.

The major aims of this chapter, therefore, are to overview therapeutic techniques directed at modifying arousal and inducing relaxation. The therapeutic rationales underlying these techniques will be critically examined, together with an evaluation of their efficacy. The review will restrict itself primarily to adults and to problems broadly conceived as being associated with mental health. Some reference, however, will be made to physical health and behavioral medicine when reviewing more contemporary applications. We will conclude by attempting to appraise whether any common process or mechanism might underlie change brought about by the application of these diverse psychological treatments.

6.09.2 A THEORETICAL OVERVIEW OF AROUSAL DYSFUNCTION AND AROUSAL MODIFICATION INTERVENTIONS

A fundamental assumption that underlies this chapter is that elevated "arousal" is associated with a variety of mental and physical health problems, and that treatment methods targeted at "arousal" reduction will alleviate the severity of these conditions. Before we review these proposed treatments, it is important that this assumption is examined further. The critical questions are: what is arousal, what associations exist between elevated arousal and physical and mental health problems, what techniques exist to modulate arousal, and what are the mechanisms that mediate therapeutic change?

6.09.2.1 Arousal as a Hypothetical Construct

The construct of arousal has been revisited extensively within the literature, with respect to psychophysiology (Andreassi, 1989; Cacioppo

& Tassinary, 1990; Gale & Eysenck, 1992), emotion (Tucker, Vannatta, & Rothlind, 1990; Wagner, 1988), and clinical applications (Cacioppo, Berntson, & Anderson, 1991; Lader, 1975; Turpin, 1989, 1990; Wieisse, Davidson, & Baum, 1989). A variety of different arousal constructs can be identified, including arousal as central nervous system activation underlying an "arousal–sleep" continuum (e.g. Duffy, 1951), arousal as a general drive system affecting behavior as characterized by the inverted U model (e.g. Andrew, 1974; Claridge, 1987), and arousal as a psychophysiological construct identified by peripheral physiological response patterning (e.g., Lacey, 1967).

Despite the almost ubiquitous presence of the "arousal construct" within psychophysiology, its current usage has markedly declined. Its demise can be traced to Lacey's classic critique (Lacey, 1967) wherein the unitary nature of a physiologically mediated arousal drive, as advocated by Duffy (1951), was discarded because different autonomic response measures, behaviors, and situational factors were found to be dissociable. More recent psychophysiological studies have also failed to demonstrate a unitary arousal construct and largely support Lacey's original position (Venables, 1984). Peripheral autonomic measures of arousal, therefore, fail to demonstrate high intercorrelations, and instead display response patterns specifically determined by either individual differences or situational factors. More recent arousal theories have also discarded the notion of "arousal" as an unitary construct, but instead have stressed the complexities of stimulus—response relationships, different arousal systems, motivational systems, and neurophysiological or neurochemical substrates (Turpin, 1989).

Given the paucity of evidence to support a unitary arousal construct, it is perhaps paradoxical that the construct of arousal continues to receive widespread support as an explanatory construct within clinical psychology. This is, perhaps, best characterized by Gale and Eysenck's (1992) quote from Claridge (1987, p. 134): "I have often felt that as an explanatory concept in psychology "arousal" has many of the qualities of a difficult but persuasive lover, whom reason tells one to abandon yet who continues to satisfy an inescapable need." The question arises, therefore, as to why, particularly within the clinical arena, this irrational infatuation should continue. One answer is that "arousal" is best viewed as a hypothetical construct and, as such, it is difficult to define precisely or to operationalize. However, clinical psychology deals with many such similar constructs including, for example, "affect," "emotion," and "anxiety," and would be severely limited without the explanatory power that such constructs bestow on clinical formulations and conditions. Indeed, we suggest that arousal has important clinical utility and that its use can be progressed by treating it in a similar fashion to other clinical constructs such as anxiety. Just as anxiety is best conceptualized according to Lang's three-system model (Lang, 1968; Turpin, 1991), arousal might be broken down into three component systems: physiological, cognitive, and behavioral (Schilling & Poppen, 1983). Physiological arousal components will include both central activation and autonomic responding; cognitive components might refer to both subjective experience and verbal report, together with attentional consequences in terms of vigilance and selective attention; and behavioral components would account for levels of activity, together with the integration and coordination of motor responding. It is suggested that these three systems are loosely coupled in a fashion similar to that described by Öhman (1987) for emotion (Figure 1). The failure to observe perfect coupling by measures either within or between component systems is essentially similar to the *de synchrony* concept commonly entertained for anxiety (Rachman & Hodgson, 1974; Turpin, 1991) and frequently encountered elsewhere within psychophysiology (Cacioppo & Tassinary, 1990; Öhman, 1992).

Several implications for therapy emerge if arousal is treated according to the systems identified in Figure 1. First, a comprehensive assessment of arousal will involve attempts to measure all three systems. Indeed, Poppen (1988) has written extensively about the multimodal assessment of arousal and the outcome of relaxation training. He suggests the use of self-report relaxation and symptom scales, specific physiological measures (e.g., electrodermal, cardiovascular, electromyogram [EMG], skin temperature, and electroencephalogram [EEG] changes), and direct behavioral observations (e.g., Behavioral Relaxation Scale). Second, it becomes evident how overarousal might lead to dysfunction and the expression of a variety of symptoms including subjective distress, hypervigiliance and heightened attention, selective appraisal, hyperactivity, disorganized behavior, and autonomic and somatic overreactivity. The presence of such symptoms within a fear situation is consistent with the constructs of anxiety or panic. Finally, the three-system approach suggests a diversity of potential interventions that might moderate the expression of hyperarousal. The choice between a unitary model of arousal reduction and relaxation, on the one hand, and a multicomponent

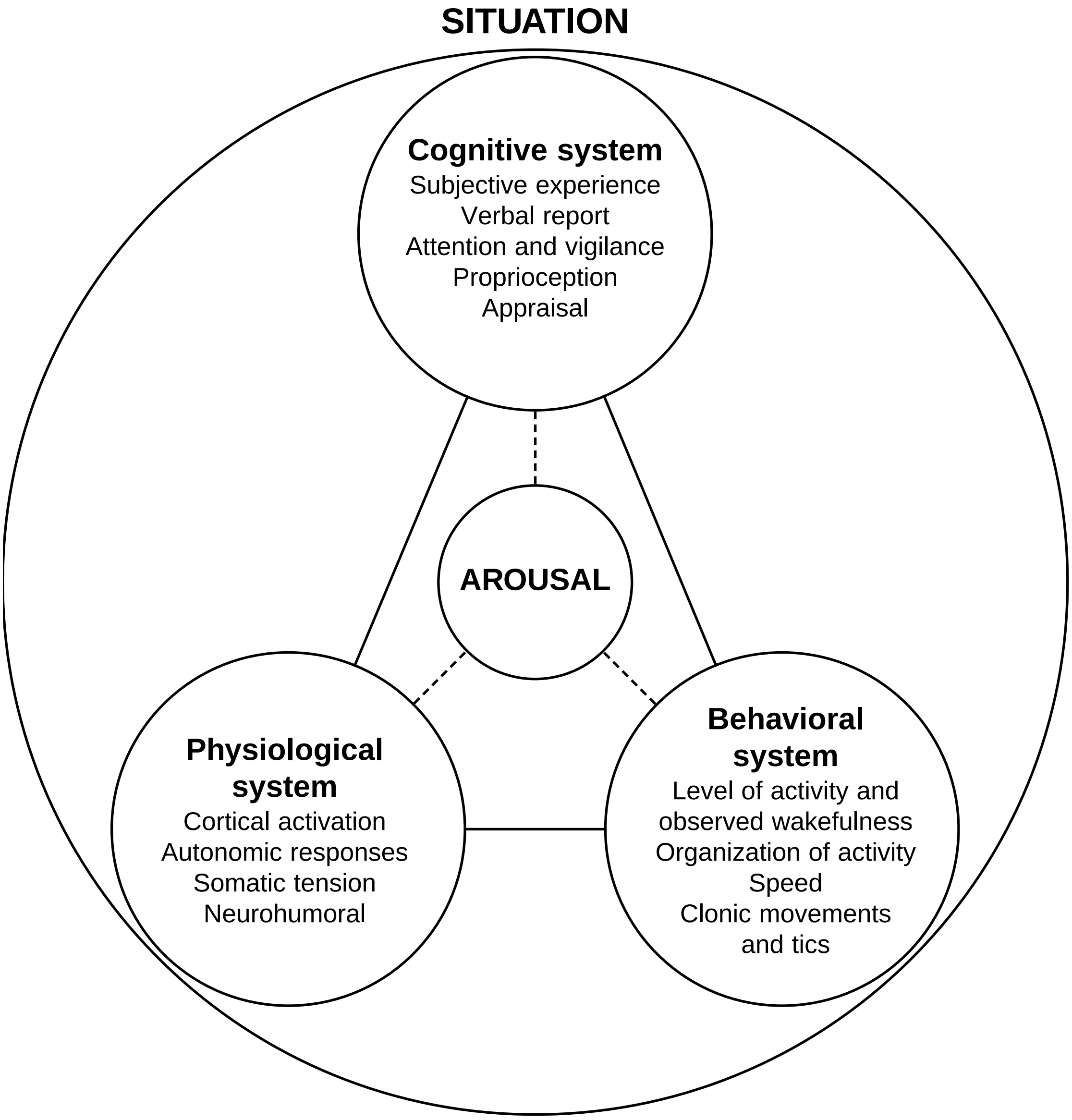
SITUATION
Cognitive system
Subjective experience
Verbal report
Attention and vigilance
Proprioception
Appraisal
AROUSAL
Physiological system
Cortical activation
Autonomic responses
Somatic tension
Neurohumoral
Behavioral system
Level of activity and observed wakefulness
Organization of activity
Speed
Clonic movements and tics

Figure 1 The inferential basis of arousal (based on Öhman's (1992) construct of emotion). (—) Observables and relationships between observables; (– – – –) inferred constructs.

model with a variety of specific treatment effects, on the other, is discussed later in this section.

6.09.2.2 Arousal Dysfunction and Health

Despite the difficulties described earlier in defining and measuring arousal, the construct has been used widely in association with etiologic theories of both physical and mental illness. Perhaps the most ubiquitous approach concerns diathesis–stress models of psychophysiological disorders (Gatchel, 1993). In particular, Sternbach (1966) postulated that situationally specific patterns of psychophysiological responding, if not adequately modulated by individual homeostatic mechanisms, would lead to dysfunctional patterns of physiological activation characteristic of so-called "psychosomatic disorders." Indeed, a diversity of theoretical models has been postulated in which elevated physiological arousal is included as a component of "stress" and is linked to both predisposition and the subsequent expression of a variety of physical health problems including headache, gastrointestinal disorders, hypertension, and asthma (McEwen & Stellar,1993; Steptoe, 1991; Weiner, 1977; Wiebe & Williams, 1992). Such models either incorporate physiological arousal as a component of an undifferentiated stress response (Seyle, 1950), or identify specific pathways associated with autonomic hyperreactivity and end-organ dysfunction, as exemplified by Freedman's hypothesized adrenergic mechanism implicated in Raynaud's disorder (Freedman, 1989). Although psychosomatic disorders are covered elsewhere, they represent a major area of study that relies on "stress–arousal" explanations of disease, and hence provide explicit therapeutic rationales for arousal reduction interventions (Lehrer, Carr, Sargunaraj, & Woolfolk, 1993). Nevertheless, as with arousal, many have argued that the concept of "stress" is too ill-defined and lacks specificity when defining individual mechanisms underlying disease or dysfunction (Steptoe, 1980).

A similar approach has been taken with regard to mental health, whereby stress is seen as a potential threat to mental well-being, and physiological stress or arousal is seen as a frequent consequence of psychosocial stressors and is manifested in a variety of psychological problems (Turpin & Lader, 1986). Such models underpin stress management approaches to mental health and the promotion of more adaptive life-styles, together with the development of more appropriate coping and adaptive responses to challenging stresses and life strain. Generic stress management approaches (Meichenbaum, 1993) are widely used for both physical and mental health problems, and are frequently employed within occupational and vocational settings (Reynolds & Shapiro, 1991).

Within the area of mental health, several specific rationales can be identified that link elevated arousal manifested across physiological, cognitive, and behavioral response modes with psychological dysfunction. Several psychological disorders are defined with particular reference to elevated levels of physiological arousal, which are situationally inappropriate and are frequently present in association with other aspects of overarousal, such as heightened attention and vigilance, somatic tension, and behavioral hyperactivity. Generally, arousal is treated nonspecifically and, for example, can be considered as a component of anxiety or even sleep dysfunction in the form of insomnia. With reference to anxiety, the three-system conceptualization of anxiety (Lang, 1968; Kozak & Miller, 1982; Turpin, 1991) is still the prevalent conceptual model, particularly when used as the basis for the evaluation of cognitive behavioral therapies. Physiological, behavioral, and cognitive components of anxiety and their manifestations can easily incorporate the arousal construct, and may even be reformulated whereby arousal is seen as an energetic or drive component of situational fear or anxiety. Similarly, insomnia is frequently (Borkovec, 1979; Nicassio & Buchanan, 1981) conceptualized as arising from both elevated physiological arousal and cognitive activation such as preoccupation. We will return in more detail to the three-system analysis of anxiety and arousal when specific therapeutic processes are discussed in later sections of this chapter.

Specific arousal components can also be identified for other psychological problems. For example, many sexual disorders are associated with specific patterns of arousal responding in the form of nonspecific physiological arousal, together with abnormal sexual dysfunction. Conditions such as vagismus, erectile dysfunction, and paraphelia might all be considered as involving a specific sexual arousal dysfunction. Similarly, heightened arousal is an important component of the expression of aggression and hostility, and may be specifically associated with particular patterns of cardiovascular and neuroendocrine response. Accordingly, it is frequently the focus of treatments aimed at anger management. Another disorder that might also be considered to include a specific arousal component concerns post-traumatic stress disorder, in which heightened anxiety, tension, hypervigiliance, and exaggerated startle responses are symptomatic features.

Finally, the construct of arousal has been used as a mediational variable for several other psychological disorders. Although physiological arousal may not necessarily be an essential component of the expression of the disorder, it is hypothesized that an underlying arousal dysfunction is instrumental in the expression of the disorder. For example, the consequences of arousal on the habituation and conditioning of fear responses has been implicated both in theories of phobic fear (Lader & Mathews, 1970) and in treatment through desensitization (Wolpe, 1982). Habituation is also suggested as an essential component of anxiety reduction within *in vivo* exposure treatments of anxiety disorders (Foa & Kozak, 1986) and also in the response prevention treatment of obsessive-compulsive disorder (Mills & Salkovskis, 1988). Elevated physiological arousal has also been suggested as a substrate underlying suggested psychomotor disorders such as torticollis (Meares, 1973), tics (Corbett & Turpin, 1985), and writer's cramp (Cottraux, Juenet, & Collet, 1983). More recently, arousal dysfunction has been implicated in relapse for psychosis, whereby arousal is seen as a mediating pathway in the stress–vulnerability model of schizophrenia (Nuetcherlein & Dawson, 1984; Tarrier and Turpin, 1992). Therapeutic approaches that have sought to modify the stress–vulnerability relationship, such as coping enhancement strategies (Barrowclough & Tarrier, 1992), frequently include arousal modulation techniques and rationales.

In summary, a variety of theoretical models exist that relate arousal dysfunction to the expression of physical and mental health problems. These frameworks, therefore, provide therapeutic rationales for arousal modification methods, assuming that such methods exist and can be shown to be efficacious. It is worth noting the wide range of problems that have been associated with an arousal dysfunction, and this perhaps lends greater credence to Eisenberg et al.'s survey that indicated widespread use of "unconventional" therapies for a wide range of conditions. In the next section the theoretical rationale underlying the treatment methods claimed to reduce arousal is reviewed.

6.09.2.3 Theoretical Basis of Arousal Modification Methods

Various psychological interventions have been suggested to reduce the level of elevated arousal (Lehrer & Woolfolk, 1993). The rationale for their use depends on how elevated arousal and the mechanisms underlying it are conceptualized. As seen in the previous section, arousal can be defined in a variety of ways—physiological, behavioral, and cognitive—and each mode can be measured and assessed independently. Moreover, it is likely that elevated arousal is the product of several distinct processes that may or may not be interrelated (see Figure 1). Accordingly, different arousal modification methods might be targeted at different components and rely on different mediational processes.

For example, a cognitive arousal component consisting of hypervigilance and appraisal of threat might be ameliorated either by cognitive therapy directed at the beliefs maintaining the negative appraisal or by the reduction and substitution of the prevalence of negative thoughts or worries *per se* using a cognitive distraction technique, perhaps involving meditation or imagery. A further reduction in a cognitive arousal component might also arise as a consequence of applied relaxation training directed at reducing physiological arousal, thereby eliminating a cue associated with interoceptive sensations of overarousal. Conversely, elevated somatic arousal, as induced by, for example, increased muscle tension and accompanying headache, might be reduced specifically by EMG biofeedback and training directed at the muscle group concerned. An emphasis on different arousal components, therefore, gives rise to what Poppen (1988) describes as multimodal theories of relaxation induction. Such theories originated from dualistic accounts of arousal reduction, which emphasized both cognitive and somatic relaxation (Davidson & Schwartz, 1976; Heide & Borkovec, 1983; Lehrer, Woolfolk, Rooney, McCann, & Carrington, 1983). The implications of these approaches are that the arousal requires assessing across a range of different domains, and that different aspects of arousal reduction methods might target different sources of arousal.

The above approach can be contrasted with unitary theories of arousal and relaxation induction, as exemplified by Benson's (1975) relaxation response. Many arousal reduction techniques are designed to induce a state of relaxation, which is conceived of as being antithetical to either arousal or stress and is typified as subjective calmness, lowered muscle tension, and lowered levels of physiological arousal. Indeed, many authors refer to the relaxation response in a similar, but opposite, fashion, to Selye's (1950) stress response. Moreover, several theoretical models have even implicated specific physiological mechanisms, particularly those related to autonomic balance, as mediating the relaxation state (Gellhorn, 1958). The rationale for some arousal reduction methods, therefore, is to lower arousal by the induction of an opposite and incompatible state of relaxation. This distinction between theories of arousal modulation, which rely on specific pathways, as opposed to a unitary model of relaxation induction is important when evaluating the differential efficacy of these techniques. If relaxation provides a common pathway, it may be that all arousal reduction interventions are equivalent in terms of process, and lead to identical outcomes, albeit with differing degrees of effectiveness in how easily achievable these outcomes are. Alternatively, if arousal reduction or relaxation induction is conceived of as a variety of specific components, different techniques might have differing degrees of efficacy expressed in different modalities for different individuals and disorders.

Indeed, the issue of specificity concerning a variety of relaxation methods has been emphasized by Lehrer and colleagues (Lehrer, 1996; Lehrer, Carr, Sargunaraj, & Woolfolk, 1994; Lehrer & Woolfolk, 1993). In contrast to Benson's (1975) unitary relaxation response, they propose an extension of Davidson and Schwartz's (1976) model, in which relaxation is viewed as a multicomponent process (of up to 12 separate components), and that specific techniques give rise to specific effects. For example, they suggest that cognitive changes might be expected from cognitive oriented methods such as meditation; somatic tension changes would be brought about by muscle-oriented methods, such as Jacobson's (1938) progressive muscle relaxation (PMR) or EMG biofeedback; and autonomic changes would be brought about by either electrodermal/thermal biofeedback or autogenic training. It should also be acknowledged, however, that specific effects may actually overlie a more generalized induction of a relaxation response, and hence these two approaches are not necessarily seen as mutually incompatible. Moreover, it is by no means conclusive as to how different interventions should be classified using multicomponent models. Autogenic training, for example, can be said to be directed at both autonomic and cognitive components. Similarly, hypnosis can also be considered as multimodal.

Lehrer and Woolfolk have tried to resolve the question of unitary vs. multicomponent models of relaxation by comparing the effects of different relaxation procedures on a range of outcome measures chosen to reflect the component processes described by Davidson and Schwartz (1976). This has been attempted either using a comparative treatment approach reviewing the results from different outcome studies (e.g. Lehrer et al., 1994) or by performing essentially analogue laboratory studies in which the effects of these different relaxation techniques on a battery of psychological and physiological measures are contrasted (Lehrer et al., 1983).

Generally, these studies are consistent with the view that the effects of different relaxation techniques are mediated by specific changes in somatic and cognitive arousal. Reviews, principally by Lehrer and Woolfolk (Lehrer 1996; Lehrer et al., 1994; Lehrer & Woolfolk, 1993; Lehrer et al., 1983), concerning the specific effects of relaxation techniques on physiological activity support the contention that muscle relaxation training, either in its original form as PMR (Jacobson, 1938) or in its modified form as abbreviated progressive relaxation (APR) (Bernstein & Carlson, 1993), produces significant decreases in muscle tension as measured by EMG. Similarly, EMG biofeedback, either delivered alone or in conjunction with modified PMR, has been shown to produce significant decreases in target EMG levels. However, whether the effects produced by EMG biofeedback delivered to a single muscle site can be generalized to other EMG sites is debatable (Alexander, 1975). Similarly, the data remain equivocal as to the comparative effectiveness of PMR, APR and EMG biofeedback on reducing EMG levels (Lehrer et al., 1994; Lehrer & Woolfolk, 1993) although, EMG biofeedback is probably superior to APR with regard to EMG reduction (Lehrer et al., 1994).

Nevertheless, the therapeutic rationale for muscle relaxation training does not restrict itself solely to reductions in muscle tension and EMG. Jacobson and his followers (see McGuigan, 1993) stressed that proprioceptive feedback from the musculature directly affects peripheral autonomic activity via the sympathetic and parasympathetic nervous system. This mechanism is based on Gellhorn's theories linking emotional experience to the skeletal musculature (Gellhorn, 1958; Gellhorn & Kiely, 1972). Accordingly, studies of relaxation training that include autonomic measures have generally demonstrated additional effects of, for example, lowered heart rate and electrodermal levels (Lehrer et al., 1994). Indeed, several studies have attempted to assess relaxation effects on the adrenergic system (see Freedman, 1994), and have included measures of peripheral muscle sympathetic nerve activity, blood norepinephrine levels (Hjemdahl et al., 1989), and platelet monoamine oxidase activity (Mathew et al., 1981). Other studies have also suggested effects on the immune system in the form of, for example, changes in killer cell activity and lymphocyte density (Gruber et al., 1993). Other arousal reduction methods have also been demonstrated to yield specific physiological effects. For example, breathing regulation techniques have been shown to lower physiological arousal and to normalize the EEG (Fried, 1993; Lehrer & Woolfolk, 1993).

The physiological effects of more cognitively mediated procedures such as autogenics and meditation have also been reviewed by Lehrer and colleagues (Lehrer et al., 1994; Lehrer & Woolfolk, 1993). Although the specific effects model of Davidson and Schwartz would predict greater cognitive than physiological effects for interventions such as meditation, Lehrer et al. conclude that the differential effects of meditation vs. relaxation training are equivocal. With respect to autogenic training, Linden (1990) cites a collection of studies demonstrating physiological effects of autogenic training, which support the relationship between subjective somatic sensations and physiological change, particularly in relation to changes in breathing and peripheral skin temperature.

Although the above studies generally substantiate the fact that the intended effects of relaxation on the levels of physiological activity are observed within treatment sessions, findings regarding the longer term effects of relaxation on physiological functioning are few. Indeed, a recent appraisal of the area which has been most thoroughly examined, that is, stress management and hypertension, reveals that a cautious approach to extrapolating from short-term effects to long-term outcomes might be warranted. Although relaxation-based stress management protocols revealed significant falls in blood pressure within hypertensive subjects (Patel et al., 1985), recent meta-analyses of studies within this area have produced less-promising findings (Jacob, Chesney, William, Ding, & Shapiro, 1991; Johnston, 1994; but see also Linden & Chambers, 1994; Linden, Stossel, & Maurice, 1996). Some of the factors responsible for these disparate findings probably include pre-existing levels of hypertension, subject selection and entry to the study, habituation and initial testing effects on blood pressure measurement, and failure of blood pressure reductions to generalize to everyday life situations (Jacob et al., 1991). It remains to be seen, therefore, whether the long-term effects of relaxation on elevated physiological arousal can be demonstrated, particularly if the methodological rigors adopted within hypertension research were to be more generally applied.

Most of the evidence reviewed in this section was derived from short-term treatment or laboratory analogue studies, and as such emphasizes process-related changes in relaxation measures. Whether these changes are also directly related to stable therapeutic outcomes with regard to reduction in symptoms and increased functioning remains to be demonstrated. Moreover, this assumes that the fundamental process underlying therapeutic change is physiological arousal reduction. However, this assumption has recently been challenged by Smith (1988, 1989), who has proposed that the therapeutic changes brought about by relaxation techniques are best explained by cognitive-behavioral formulations, as opposed to changes in physiological functioning. Hence changes in attitudes towards relaxation and stress, together with specific changes in personal cognitive schema, underlie therapeutic improvement, which may or may not be accompanied by a lowering of physiological arousal. The issue of specificity of effect will be examined in greater detail in the final section of this chapter.

6.09.2.4 Summary

In summary, arousal reduction methods are based on a variety of different explanatory therapeutic models. Indeed, the construct of arousal has been difficult to define and cannot be satisfactorily operationalized. Nevertheless, despite the construct being less frequently employed within experimental psychology, its use as an explanatory factor in clinical psychology continues to be prevalent. Many physical and mental disorders can be considered to involve an element of arousal dysfunction, either resulting directly in the expression of symptoms or indirectly as a result of some mediational role of arousal in triggering specific symptoms. Accordingly, a variety of therapeutic techniques targeted at arousal reduction have been devised in order to alleviate the effects of overarousal.

A variety of mechanisms have been proposed to underlie these techniques, ranging from the unitary induction of a relaxation response through to specific processes responsible for therapeutic change across different arousal components. It would appear that the majority of relaxation techniques produce some measurable effects on a variety of cognitive, somatic, autonomic, and central measures. However, the effects produced by different techniques do not appear to be equivalent, as might be predicted by a unitary theory such as Benson's. Nevertheless, there would appear to be evidence of some more generalized physiological effects of relaxation on autonomic, neurohumoral, and immune system functioning that underlie specific treatment-induced changes. Moreover, it should be acknowledged that some specific predictions of the multicomponent models have not been supported.

6.09.3 AROUSAL REDUCTION METHODS: GENERAL PROCEDURES AND SPECIFIC TECHNIQUES

In this section we systematically review clinical interventions derived in order to reduce arousal and facilitate positive functioning and therapeutic change. Only those techniques that are supported by a clear therapeutic rationale and that have been scientifically evaluated in terms of their efficacy are included. In addition, the review focuses primarily on mental health problems prevalent in adults. However, where appropriate, procedural modifications for child and elderly populations are noted. Physical conditions such as hypertension, headache, or pain control will not be systematically covered as these are reviewed elsewhere, but will be referred to if the studies are able to expand on some general procedural issues surrounding the implementation of these therapeutic techniques. The main techniques reviewed are relaxation methods, autogenic training, breathing modification, biofeedback, meditation, and hypnosis. However, many of these techniques have been advocated for use in a variety of different clinical problems based on a diverse range of therapeutic rationales. Hypnosis, for example, is claimed to be therapeutically beneficial in inducing a relaxation state, accessing "inaccessible memories and feelings," and promoting cognitive restructuring. Accordingly, we will restrict ourselves to covering only specific applications of these techniques where the therapeutic rationale upon which their use is based concerns the reduction or modification of arousal.

Before reviewing each technique in turn, we briefly describe and evaluate some general procedural factors. Contraindications and modification of the technique for specific problems or client groups will also be reviewed. The comparative efficacy of different techniques and their use for specific conditions are dealt with in the final section of this chapter, together with a discussion of the therapeutic mechanisms that might underlie their efficacy.

6.09.3.1 General Procedural Issues

The limited scope of this chapter prevents a detailed account of the clinical protocols used to deliver these methods. Instead, we aim to deal with some general procedural issues common to most relaxation or arousal reduction techniques, followed by a brief description of the essential features of each technique. Several excellent practical guides and training manuals have been published (Howell & Whitehead, 1989; Payne, 1995; Whitehead & Adams, 1991) and these should be consulted by the novice practitioner alongside this chapter.

6.09.3.1.1 Assessment

The successful application of relaxation techniques will depend on sound assessment and formulation of the clinical problem, leading to an appropriate rationale for their use. Assessment should focus on identifying the clinical problems amenable to relaxation techniques (Lehrer & Woolfolk, 1993; Poppen, 1988), together with establishing the presence of elevated physiological arousal, and self-reports or inferred observations of tension, anxiety, or worry (Crist, Rickard, Prentice-Dunn, & Barker, 1989; Poppen, 1988; Smith, 1989).

As discussed in Section 6.09.2.2, specific health problems said to be associated with elevations in arousal are generally identified as anxiety disorders (e.g., generalized anxiety disorder, panic disorder, and specific phobias), psychosomatic complaints (e.g., hypertension, tension headache, migraine, and asthma), elevated tension or somatic arousal (e.g., chronic pain, insomnia, and skeletomuscular disorders), and palliative problems and applications (e.g., side effects of chemotherapy). The purpose of assessment, therefore, is to substantiate the presence of any of these clinical indications, and to provide sufficient information to arrive at an individual formulation of the presenting problems. Such a formulation should address the rationale for the adoption of a relaxation-based intervention and should also rule out other more appropriate treatment approaches. The latter might concern the use of physically based medical treatments or more psychotherapeutically based interventions. Not that these treatments are necessarily mutually exclusive; in the case of psychological interventions in particular, relaxation techniques may well be used in conjunction with other cognitive and behavioral approaches. A final purpose of assessment should be to identify contraindications of applying relaxation-based therapies (see Section 6.09.3.2.1).

In order to formulate the presence of an arousal-related health problem that might be amenable to relaxation, it is important to verify the presence of elevated arousal either directly through physiological measurement or by inference from observation or self-report. Indeed, given the multidimensional nature of the arousal construct (see Figure 1), it is likely that individual measures will not be interchangeable and that a multicomponent approach to assessment will be required. Physiological measures of arousal will require sophisticated equipment and expertise, which is unlikely to be available in most clinical settings, but may well be available if biofeedback techniques are also on offer within the clinic. Simple assessments of elevated EMG or heart rate may also be accomplished using inexpensive commercially available monitors. Further advice and information on psychophysiological assessment procedures are available elsewhere (e.g., Cacioppo & Tassinary, 1990; Turpin, 1989).

An alternative to direct measurement is the use of either therapist observational checklists or client self-report instruments. The most widely used observational scale is the behavioral relaxation scale (BRS) (Schilling & Poppen, 1983), which is discussed in great detail in Poppen's instructional text (Poppen, 1988). A wide range of self-report scales is available to assess, for example, anxiety and worry, somatic arousal, stress and tension, state of relaxation and calmness. The use of these scales has been reviewed by Smith (1989).

6.09.3.1.2 Therapeutic rationale

Assuming that assessment has revealed a need for relaxation training, and the most appropriate method has been chosen for the individual, a common feature of most methods is to present a clear therapeutic rationale to the client (Bernstein & Carlson, 1993). This will usually involve an explanation or reformulation of the client's problem in terms of anxiety or an inability to relax, the provision of basic information on stress and anxiety, and an explanation of how relaxation can be induced and its potential benefits to the client. The latter will usually stress the importance of regular practice and emphasize that the client must take responsibility for his or her own therapy. It is important that at the beginning of therapy the client has an accurate expectation of the nature of the therapeutic intervention and his or her own commitment to the therapy. The adoption of a clear therapeutic rationale, therefore, together with agreed therapeutic goals between therapist and client cannot be overstated. At the very least it provides the therapeutic framework within which relaxation techniques can be learnt and practised. Indeed, some would argue (e.g., Smith, 1989) that the therapeutic benefits of relaxation are dependent on the client's reframing of the clinical problem cognitively, and arise as much from changes in cognitive schema and appraisal processes as they do from any accompanying changes in physiological arousal.

6.09.3.1.3 Practical considerations

In addition to introducing to the client an appropriate rationale, it is also important that the client is provided with instructions concerning environmental considerations for relaxation training. Important requirements are a comfortable chair and the adoption of a reclining position, avoiding extraneous noises or interruptions, subdued lighting and a comfortable ambient room temperature, and the wearing of loose clothing that does not restrict body movement. These factors apply both to clinic sessions and homework practice. Finally, the adoption by the therapist of a smooth, quiet, and almost monotonous voice for the presentation of instructions is also an important consideration.

6.09.3.1.4 Presentation of the technique

Different techniques clearly have to be delivered to the client by the therapist in many different ways. Nevertheless, certain choices regarding the mode of training have to be made. For example, what amount of information is provided prior to training (booklets, audiocassettes, videos, etc.)? The role of the therapist in providing within-session modeling of the techniques for the client, and the opportunity for the client to rehearse techniques before the therapist and to receive feedback are examples of procedural questions that have to be resolved. Similarly, the therapist has to decide between the use of individual sequences of exercises, whereby individual progression is based on mastering each stage and the strict adherence to a standardized and, possibly, inflexible protocol. The majority of published studies of relaxation training relate to individualized, one-to-one therapy as opposed to group delivery methods. However, in practice, many practical relaxation training programs are directed at group approaches (Howell & Whitehead, 1989; Whitehead & Adams, 1991). Other considerations include the number of sessions, homework practice, compliance, and follow-up and booster sessions. The overall impact of these factors on treatment efficacy are now reviewed.

6.09.3.1.5 Evaluation of procedural factors

Several of the above general procedural factors commonly found in most forms of relaxation training have been investigated in order to determine their influence on therapeutic outcomes. These have included taped vs. live instruction, the use of homework practice sessions, and brief vs. extended practice, and they have been the subject of several reviews (Bernstein & Carlson, 1993; Borkovec & Sides, 1979; Blanchard et al., 1991; Hillenberg & Collins, 1982; Lehrer & Woolfolk, 1993). Although none of these reviews yield unequivocal conclusions, several consistent recommendations emerge. First, live instructions, whereby the client is able to control the progress of training by moving systematically from one muscle group to the next, with access to appropriate observation of and feedback from the therapist, appears superior to the sole use of standard taped instructions. Second, most clinical trials have included home practice sessions, perhaps supplemented by taped instructions and measures of homework compliance. Some data exist to support the additional benefits of prolonged and regular homework practice, although some studies (Borkovec et al., 1987; Nelson & Borkovec, 1989) have failed to support such an effect. Compliance and overreporting of homework practice sessions might explain some of the inconsistencies in these studies (Bernstein & Carlson, 1993). Third, it would appear that the greater exposure to relaxation training, either through extended clinic sessions or through continued application outside of the clinic, is more likely to be associated with maintenance of therapeutic change. The continued application of brief relaxation procedures, as opposed to more extended periods of relaxation training, has been argued by Lake and Pingel (1988) to be an important determinant of outcome for headache patients. It should be emphasized, however, that the above recommendations are largely based on clinical guidelines rather than extensive outcome studies. Moreover, specific procedural recommendations for different techniques or subject groups have yet to be evolved.

6.09.3.2 Specific Techniques

The purpose of this section is to provide the reader with a brief account of the essential components of each technique. More detailed clinical protocols are available elsewhere (Lehrer & Woolfolk, 1993). As relaxation-based techniques have been the focus of this chapter so far, greater emphasis will be placed in this section on detailing the background to other techniques, such as biofeedback, meditation, and hypnosis, which have also been commonly employed as arousal reduction methods.

6.09.3.2.1 Relaxation-based methods

(i) Jacobson progressive muscle relaxation

The development and application of Jacobson's ideas have been reviewed extensively by McGuigan (1993). Jacobson's method was devised in 1938 and seeks to train subjects in the ability to perceive muscle tension accurately, to control and hence reduce tension, and to relax their muscles differentially. The training involves detailed familiarization with the majority of muscle groups and an emphasis on enhancing awareness of muscle tension. Discrimination training is brought about by the method of "diminishing tensions," whereby the client learns to effect and detect the smallest possible muscle contraction in a variety of different muscle groups. Once the skill of discrimination in muscle tension has been acquired, the client can then bring about "differential relaxation" of his or her muscles, which entails only contracting those specific muscles that are functionally required for action, and relaxing all other muscles unassociated with the ongoing physically active group. The distinctive feature of Jacobson's approach is its specific emphasis on muscle tension and control, and the absence of any cognitive or suggestive components directed at the induction of a subjective feeling of relaxation.

(ii) Abbreviated progressive muscle relaxation

Jacobson's techniques were adapted, notably by Wolpe (1958) and Paul (1966), and modified in the 1960s to form part of the basis for several behavioral treatments, including systematic desensitization. The major changes were a shift from actual to subjective tension, an abbreviated and standardized protocol, use of the tension-release cycle, and greater emphasis on subjective and cognitive approaches. Many of these changes were incorporated into a highly influential treatment manual produced by Bernstein and Borkovec (1973). The specific instructions underlying the technique involve the client performing a series of tension-release cycles in a systematic manner across 16 muscle groups, beginning with the hand and forearm, moving through the forehead, face and neck, down to the abdomen, and finishing with the feet. Each tension-release cycle involves the client contracting the specified muscle according to a cue "tense" or "now," holding and sensing the contraction for 5–7 seconds, and relaxing the muscle on the command "relax," followed by a period of 30–40 seconds during which the client focuses on the sensations of relaxation. A detailed account of these procedures is provided by Bernstein and Carlson (1993). Various modifications have also been proposed to the standard Abbreviated Progressive Relaxation Training (ABRT) protocol, including the establishment of relaxation cues in the form of cue-controlled relaxation (Grimm, 1980) and attempts to generalize the relaxation response into everyday settings, such as in applied relaxation training (Ost, 1987).

(iii) Breathing training

A variety of techniques exist to induce relaxation through the modification of breathing. Specific instructions to entrain diaphragmatic breathing have commonly been used as adjuncts to APR (Bacon & Poppen, 1985; Poppen, 1988) and have also been the focus of cue-controlled relaxation procedures (Grimm, 1980). For disorders such as panic, whereby sensations arising from hyperventilation have been implicated, controlled breathing exercises have been used effectively (Clark, Salkovskis, & Chalkley, 1985). More specific techniques involving respiratory feedback and slowly paced breathing have also been developed (Fried, 1993; Montgomery, 1994). Finally, breathing regulation is also a common component of yoga (Patel, 1993) and other meditation-based techniques (Carrington, 1993), and these are discussed later.

(iv) Autogenic training

Instead of focusing directly on muscle tension as a means to induce relaxation, autogenic training (AT) involves a set of exercises targeted at various physiological end-organs whereby a state of relaxation might be induced. These techniques have been described in detail by Linden (1990, 1993) and are based on the work of Schultz and further developed by Luther. "Autogenics", as a term, essentially means "self-exercise" or "self-induction" therapy and relies on a theoretical rationale whereby the client gains psychophysiological self-control of autonomic functioning and is able to achieve relaxation. Control is said to be gained by the client focusing on body sensations in a *passive* manner. This is achieved by the client *subvocally* repeating six standard formulas that are designed to induce physiological change: these include the "heaviness formula" directed at the muscles, the "warmth formula" aimed at peripheral blood flow, the "heart formula" associated with the heart rate, and the "breathing formula" targeted at a natural breathing rhythm. Like many other relaxation techniques, AT has been modified and used in combination with other methods, particularly thermal biofeedback (Norris & Fahrion, 1993). Indeed, AT techniques and imagery are commonly employed to assist relaxation induction within a range of different procedures. However, it should be emphasized that the classical use of AT also implies a therapeutic rationale whereby the techniques facilitate cathartic experiences in the form of "autogenic discharges."

6.09.3.2.2 Biofeedback

Biofeedback is the regulation of autonomic nervous system functions, such as blood pressure, heart rate, and sweating, by the subject continually monitoring that function and being rewarded (usually simply by the knowledge of his or her success) for changing that activity in a desired direction. Feedback is usually conveyed either auditorily by the pitch of a continuous tone or by using a visual display.

The idea that a person, with training, can selectively influence the activity of a particular autonomic function, notably in the direction of diminished arousal, has been in existence for a long time and is traditionally associated with Eastern practices such as meditation and yoga. It has been claimed that practitioners of these forms of meditation have the ability to attenuate their vital functions (e.g. heart rate) to an extraordinary degree. In contrast, Western scientific interest in autonomic control has been a recent development and was stimulated by the work of Miller and his colleagues on the operant conditioning of autonomic responses in animals, particularly the curarized rat. This research was significant from a theoretical standpoint because the findings suggested that autonomic and visceral responses, hitherto held only to be amenable to alteration by classical conditioning, could be modified by positive reinforcement within an instrumental or operant-conditioning paradigm. Around that time also, interest had been shown (Kamiya, 1969) in the operant conditioning of the alpha EEG rhythm (as a means of achieving altered states of consciousness) and of changes in skin conductance (Shapiro, Crider, & Tursky, 1964) and heart rate (Engel & Hansen, 1966). However, these studies on humans were criticized by Katkin and Murray (1968) for their lack of controls for mediating responses such as breathing.

Despite the many problems encountered in replicating some of the above animal laboratory work, it proved very influential and a wide range of clinical applications have been investigated since then and have focused mainly on two types of problem. First, anxiety and stress disorders, where the aim of biofeedback is to reduce general arousal levels; and, second, disorders presenting as somatic problems (e.g., pain, tension headaches, and irritable bowel syndrome), which may be triggered or exacerbated by overarousal and where again the biofeedback is either aimed at general relaxation or is targeted more specifically to the affected organ or function in order to achieve self-regulation.

Biofeedback techniques associated with generalized arousal reduction have included from the outset alpha rhythm EEG feedback (Budzynski & Stoyva, 1973) and other methods such as electrodermal (skin conductance) feedback (usually from the hand) and surface EMG feedback, for example, from the frontalis muscle, or in combination. The latter techniques are now more commonly used and may also be employed to augment systematic *in vitro* or *in vivo* desensitization by enabling closer monitoring of the return-to-relaxation phase, and thereby assisting patients to achieve this (Stoyva & Budzynski, 1993).

Procedurally, treatment by biofeedback begins with an assessment of the patient's presenting problem and an explanation of the rationale of biofeedback. Initial training sessions with the therapist enable the patient to become more attuned to cognitions and bodily events that are associated with increase in arousal. This training has been reported to be productive in establishing anxiety-evoking cognitions in the case of generalized anxiety (Budzynski & Stoyva, 1973). This stage is followed by daily practice at home (multiple short sessions) until mastery of the target response has been achieved and the patient can be weaned off the biofeedback device. Again, it should be noted that biofeedback is often combined in treatment with other self-relaxation methods, such as autogenic training or progressive relaxation (Lehrer et al., 1993).

6.09.3.2.3 Meditation

Shapiro (1982) defines meditation as referring to "a family of techniques which have in common a conscious attempt to focus attention in a non-analytical way and an attempt not to dwell on discursive, ruminating thought" (p. 268). However, he points out that this definition covers a wide variety of methods of meditation. Moreover, the methods may involve either a state of mental relaxation, or excitement and arousal; either physical restfulness or, as in the case of the whirling dervish or tai chi, physical activity; and different prescriptions as regards the distribution of attention (focus on awareness generally, focus on a specific thought, or a shifting form one to the other). In general, however, it is not the contents of awareness that are significant but the process of attending to the objects of one's awareness.

Such a definition of meditation presupposes no adherence to any cultic philosophy or beliefs, a point made by Carrington (1993) in her overview, citing transcendental meditation as the most widely known and extensively studied of the Western methods. Carrington goes on to distinguish two methods that have been used in clinical settings, namely clinically standardized meditation and the respiratory one method. In the former, the meditator selects a mantra (usually not a real word) which he or she initially repeats aloud with the instructor in the training phase. Later, the mantra becomes the focus for silent meditation, twice-daily for 20 minutes being recommended, although many meditators reduce the frequency of sessions to one per day or less after training (Carrington, 1993). Training is facilitated by written instructions and a cassette tape. The respiratory one method (Benson, 1975) requires the meditator to repeat silently the mantra (which may simply be the number one), pacing this with the outward breath. It is thus more structured, and is presumed to be more "effortful" than the above method (Carrington, 1993).

Meditation is recommended as an activity for everyday living, one which reduces the stress of daily life, encourages a calmer outlook, promotes greater self-awareness and inner contentment, and so on. However, it has also been used clinically in the case of anxiety (Kabat-Zinn et al., 1992; Kirsch & Henry, 1979; Thomas & Abbas, 1978), insomnia (Woolfolk, Carr-Kaffeshan, & McNulty, 1976), hypertension (Surut, Shapiro & Good, 1978), psoriasis (Gaston, 1988–1989), asthma (Honsberger & Wilson, 1973), and alcohol and drug abuse (Benson & Wallace, 1971; Parker, Gilbert, & Thoreson, 1978; Taub, Steiner, Weingarten, & Walton, 1994).

6.09.3.2.4 Hypnosis

The origins of hypnosis may be traced back to the ideas and practices of the Austrian physician Franz Anton Mesmer (1734–1815). However, so great has been the transformation since that time that mesmerism, on the one hand, and modern hypnosis, on the other, bear little resemblance to one another. Hypnosis as it is now practised may be operationally defined as

> an interaction between two people (or one person and a group) in which one of them, the hypnotist, by means of verbal communication, encourages the other, the subject or subjects, to disattend to their immediate realities and concerns and to focus their awareness on inner experiences such as thoughts, feelings and imagery. The hypnotist further attempts to create changes in the way subjects are feeling, thinking and behaving by directing them to imagine various events or situations which, were they to occur in reality, would evoke the intended changes in the subjects. (Heap, 1995, p. 649)

The state of inner focus (which is almost invariably, though not always, associated with feelings of calmness and physical relaxation) is what nowadays most practitioners appear to mean by a "trance." As such it is continuous with everyday experiences that are popularly given that label (e.g., absorption in a daydream). It is generally achieved by the hypnotist's use of suggestions and imagery, which denote feelings of calmness, heaviness, tiredness, sleepiness, and so on, a procedure known as "hypnotic induction." Whichever method is used, it may be learned by the patient and practised regularly at home, either by himself or herself or with the aid of a recorded tape. At its simplest then, "self-hypnosis" resembles meditation and other self-directed relaxation procedures such as progressive muscular relaxation and autogenic relaxation, and the general aims are similar (Sanders, 1993).

In the "heterohypnotic" context, having performed an induction procedure, the hypnotist proceeds to give suggestions to the patient that are intended to elicit the desired therapeutic changes. It is useful, however, to bear in mind that rather than a therapy in itself, hypnosis is best regarded as an adjunct to therapy (Gibson & Heap, 1991), and the range of problems to which it is applied almost matches that seen in any mental health context. Thus it may be used to augment a cognitive-behavioral approach to, say, anxiety management by the use of suggestions of calmness and symptom control in mentally rehearsed, anxiety-provoking situations, and posthypnotic suggestions (i.e. suggestions to take effect after the hypnotic session) of the immediate and effective use of self-control techniques in such situations. These procedures, which may be targeted at any of the autonomic, cognitive, or behavioral levels, may be augmented by the use of cues or "anchors." For example, a positive feeling may be coupled with a simple gesture such as the clenching of a fist (Stein, 1963); this technique may be rehearsed repeatedly in imagination, and the posthypnotic suggestion be given that this feeling may be re-evoked by the same action as and when required.

Another very common routine is that of "ego strengthening"; this consists of suggestions, usually delivered towards the end of a session of hypnosis, which are intended to promote self-confidence and more positive self-esteem. Use is often made in this procedure of symbolic and metaphorical imagery and guided fantasy (Stanton, 1990, 1991).

Hypnotic procedures are often used as an adjunct to dynamic psychotherapy, notably by the use of what are termed "exploratory" and "uncovering" techniques. These include age regression and various guided imagery techniques (Gibson & Heap, 1991; Karle, 1988), and are generally based on the notion that unconscious processes lie at the root of the patient's presenting problem. Nevertheless, these techniques may be put to good use by the eclectically minded cognitive or behavior therapist (Gibson & Heap, 1991; Heap, 1991).

One difference between hypnosis and other methods of arousal control is that the capacity to respond to hypnotic suggestions varies (around a central tendency) among the population, and is a relatively stable trait (Hilgard, 1991). Standard scales exist for measuring hypnotic susceptibility (or, as some would prefer, suggestibility) and these have good psychometric properties (for a useful overview see Fellows, 1988). However, because so many factors contribute to the outcome of therapy and because hypnosis tends to be used in an adjunctive manner, the relationship of hypnotic susceptibility to therapeutic outcome is highly variable and tends to be of modest proportions (Spinhoven, 1987; Wadden & Anderton, 1982).

6.09.3.3 Indications, Side Effects, and Contraindications

6.09.3.3.1 Client selection

The self-regulation methods presented here have the advantage that their immediate aim is to create pleasant experiences purely from something the patient does for him- or herself, and ultimately, at any time, control of treatment is in the hands of the patient, and may be instigated, directed and terminated at his or her own initiative. Consequently, patients require a certain level of autonomy and ability to manage their own daily activities, and perhaps a more "internal" locus of control (see the review by Lehrer & Woolfolk, 1993).

Adherence to a prescribed program of self-directed relaxation has been noted to be a problem by a number of authors (Bernstein & Carlson, 1993; Carrington 1977; Linden, 1993) and some studies have indicated overreporting of home-based practice by patients (Hoelscher, Lichtenstein, & Rosenthal, 1986). A problem that is more serious than simple lack of discipline on the patient's part (or inadequate instruction or direction on the therapist's part) may be that training in self-regulation does not satisfactorily address the patient's problems. There may also be issues relating to secondary gain or marital and family relationships, or unacknowledged or undisclosed areas of difficulty and anxiety that need to be the focus of a more extensive cognitive behavioral or psychodynamic intervention.

Another characteristic of potential relevance to patient selection is absorption, the tendency of an individual to become deeply immersed in the object of his or her attention (e.g., a book, some music, or a daydream). This quality may be measured using a standardized questionnaire (Tellegen & Atkinson, 1974). There is a modest relationship between absorption and measured hypnotic susceptibility (Council, Kirsch, & Hafner, 1986), and there is some evidence of a tendency for good meditators to be high on absorption (Davidson & Goleman, 1977).

Hypnotic susceptibility is itself related to outcome in several of these self-regulation methods, although only modestly so, and not at all in the case of biofeedback (Lehrer & Woolfolk, 1993; Sigman & Phillips, 1985). Whereas, as was stated earlier, practitioners may make use of self-regulation procedures in combination, biofeedback and hypnosis do not appear to complement one another in this way, and the net effect may be less positive than either technique used alone (Edmonds, 1979; Sigman, 1988). Dumas (1980) has suggested that different attentional processes may be involved in these two methods.

There is a case for favoring high susceptibility subjects as candidates for hypnotic self-regulation in, for example, an anxiety-management training program that relies heavily on suggestions of automatic responding (Spinhoven, 1987). There is also a case for offering hypnosis, and even simply explicitly calling the procedures "hypnosis," when it is the treatment requested by the patient, and not doing so when the patient has a negative attitude to hypnosis (Kirsch, 1996). Perhaps equivalent recommendations may be made in the case of the other methods.

6.09.3.3.2 Side effects and adverse reactions

Notwithstanding the relatively benign nature of the procedures under discussion, adverse effects have been noted in the literature, and precautions and contraindications have been detailed by various authors (e.g. Shapiro, 1982, for meditation; Bernstein & Borkovec, 1973, for progressive relaxation; Linden, 1993, for autogenic training; and Crawford, Hilgard, & Macdonald, 1982, for hypnosis). Reported effects, which are generally transient, include unwelcome and intrusive imagery, anxiety, panic, dizziness, confusion, restlessness, and headache. For example, Edinger and Jacobsen (1982) surveyed 116 clinicians who reported on the effects of relaxation training on 17 542 clients. The following prevalence of problems was reported: intrusive thoughts (15%), fear of losing control (9%), disturbed sensory experiences (4%), muscle cramps and spasms (4%), inappropriate sexual arousal (2%), and dissociation (0.4%). Other problems less frequently reported included laughing, coughing, and falling asleep.

Indeed, the presence of such side effects has led to the awareness of a particular problematic response known as "relaxation-induced anxiety (or panic)," which appears to represent a cumulative build-up of arousal during the relaxation procedure itself. Explanations have included fear of losing control, fear of the sensations of relaxation itself, "interoceptive conditioning" and brief bursts of hyperventilation (Adler, Craske, & Barlow, 1987; Bernstein & Borkovec, 1973; Borkovec, 1987; Braith, McCullough, & Bush, 1988; Ley, 1980).

These effects do not in themselves contraindicate the use of relaxation procedures in the individual so affected; exploration of the reasons for their occurrence, and adjustments to the procedure may be required. For example, within the pretreatment rationale it may be useful to acknowledge the existence of side effects, while emphasizing that their occurrence is usually infrequent. When side effects do occur, this is confirmation of the prior information provided and less likely to be interpreted in a catastrophic manner. The client should be assured that side effects are probably of a transitory nature and that, from a positive and constructive perspective, they indicate that the training is clearly having *some* effect. In this way the client is also engaged with the treatment. Adaptive strategies may include focusing on breathing in order to reduce the likelihood of hyperventilation, and the use of mantras such as "calm" and "relax." In the case of fear of loss of control, the use of a more individualized approach, self-paced by the client, should be considered. If problems persist, the therapist should consider adopting a different technique, preferably based on some formulation as to why the side effects are being experienced with the current relaxation method employed.

Finally, it should be recognized that the possible presence of unreported side effects might account for treatment noncompliance and rates of uptake and dropout within therapy. As will be discussed in greater detail in Section 6.09.4.2, relaxation-based treatments have compared favorably with more recently evolved cognitive-behavioral techniques, sometimes rather unexpectedly since they have frequently been included in therapy outcome studies as a control condition. However, there are strong indications that they may be less acceptable to clients, resulting in lower uptake and higher dropout rates in therapy. Such effects may be due to a variety of factors, such as treatment credibility and client expectations; however, negative side effects might also account for possible reduced acceptability.

6.09.3.3.3 Contraindications and procedural modifications

There is a limited literature regarding the suitability of arousal reduction techniques for specific clinical problems or client groups. Where contraindications have been recommended they tend to have been founded on clinical observation rather than on outcome data, and it is often uncertain whether specific warnings of adverse effects can be generalized across all methods. A brief overview of these issues is provided here, together with some possible procedural modifications for specific client groups that might result in making these techniques more appropriate.

The presence of clinical depression is an important factor to consider. Shapiro (1982), for example, suggests that meditation may not be appropriate for chronically depressed individuals, for whom a low level of arousal is more characteristic. Neither, the same author argues, is it pertinent to the needs of the socially withdrawn individual. Depression has also traditionally been said to contraindicate hypnosis (Burrows, 1980), one argument being that it can have mood-elevating properties that may provide the suicidally depressed patient with the impetus to take his or her own life. There is not, however, a great deal of evidence for this, and in recent years hypnosis has been used to augment cognitive-behavioral therapy of depression (Alladin & Heap, 1991). Lehrer et al. (1993)

have also reviewed whether stress management techniques in general, but particularly those involving relaxation training, should be directed at depression, since received clinical wisdom would counsel against such an approach. Again, they report that stress management is associated with improvements in depression, particularly when it is combined with other therapeutic approaches. Nevertheless, they only cautiously suggest the possible inclusion of stress management in programs for the treatment of depression in view of the problems of motivating clients to comply with an active homework-based therapeutic intervention, and the need to ensure that the most effective treatments are employed in the case of potentially suicidal clients. We may conclude that self-regulation may be used with the clinically depressed patient with due care, and as one component of a broader therapeutic program.

One adverse consequence of the methods in question seems to be, as in the case of depression, the amplification of an existing negative internal state. Individuals who lack well-developed reality-testing processes, such as psychotic and borderline patients, and indeed who may hallucinate, may be unsuitable candidates for self-regulation methods, at least without close supervision (Carrington, 1977; Gibson & Heap, 1991), and there is some evidence that prolonged meditation may precipitate psychotic episodes in some highly vulnerable individuals (Lazarus, 1976). Similarly, Adler and Morrisey-Adler (1983) have reported the possibility of increased hallucinations with biofeedback training. Nevertheless, Lehrer et al. (1993) have reviewed several studies purporting to report beneficial effects of stress management within psychosis. Given the contemporary interest in psychological management techniques for psychotic symptoms, it is perhaps time to reassess the application of these techniques to schizophrenia. Moreover, many of these recent approaches to the psychological management of psychotic symptoms are based on the stress–vulnerability model (Clements & Turpin, 1992; Tarrier & Turpin, 1992), which identifies increases in arousal as an underlying mechanism in the production of psychotic states. Further support for such an approach may also be derived from studies of clients' own self-control strategies for psychotic phenomena, which commonly include arousal reduction methods. Indeed, these have been capitalized on by Tarrier and Barrowclough's therapeutic approach known as self-control enhancement therapy. It remains to be seen, however, if these techniques are useful in both relieving psychotic symptoms as well as reducing excessive levels of anxiety in these clients.

There are some other conditions the presence of which warrants special consideration by the therapist. For example, Carrington (1993) advises careful monitoring of patients' medication requirements (antihypertensive and anxiolytic drugs), as these may alter during treatment. Stoyva and Budzynski (1993) give a similar warning for patients with insulin-dependent diabetes. Clearly, patients should not deny themselves their essential medication in the belief that otherwise they are failing in their self-regulation procedure (Gibson & Heap, 1991).

In addition to specific clinical problems that might mitigate against the use of arousal reduction techniques, the applicability of these techniques in general to certain client groups has been questioned. Essentially, there are two major considerations: do clients possess sufficient cognitive functions to benefit from such an approach, and are there any physical barriers that might prevent the implementation of relaxation or other related exercises? Client groups that might be considered unresponsive to a stress reduction approach because of cognitive or developmental limitations might include children, people with a learning disability, or individuals with acquired brain damage. Poppen (1988) specifically reviews the application of behavioral relaxation training with these populations, and provides examples of successful interventions. Similarly, Lehrer et al. (1993) have reviewed applications of stress management techniques specifically for children, and report positive findings in relation to hyperactivity and academically related anxiety problems.

Physical restrictions such as arthritis or neuromuscular dysfunction can give rise to serious obstacles to using muscle relaxation based methods with either brain-injured or elderly populations. Two different approaches might be used to overcome this problem. First, a variant of PMR has been designed that is based not on muscle contraction but on muscle stretch. It is argued by Kay and Carlson (1992) that this approach might be more acceptable to older adults. Second, it could be argued that more cognitively focused interventions might be of greater benefit to older adults who experience either restricted movement or arthritic pain. For example, Scogin, Richard, Keith, Wilson, and McElreath (1992) have demonstrated that imagination-based relaxation procedures were as effective as APR for an elderly population. Deberry, Davies, and Reinhard (1989) have also reported that a meditation based relaxation approach was more effective than a purely cognitive approach in dealing with emotional problems in an elderly group of clients.

6.09.4 COMPARATIVE OUTCOMES AND THERAPEUTIC MECHANISMS

In this final section we briefly address the issue of comparative outcomes for a variety of adult mental health problems, focusing particularly on anxiety disorders. In doing so, we draw attention to relationships between various techniques of arousal reduction and differential therapeutic responses across a range of outcome measures. We also speculate about possible therapeutic processes that might underlie these different relationships.

6.09.4.1 Overview of Outcomes

In Section 6.09.2.3 we discussed the evidence supporting the notion of treatment specificity across a range of arousal reduction methods. However, the majority of these studies focused on within-session changes and relied on either experimental or analog designs as opposed to clinical outcome studies employing patient samples. Moreover, the question that was addressed concerned the effectiveness of individual techniques in modifying various measures of arousal or stress. For example, if EMG biofeedback is compared to PMR, which method gives rise to the greatest change in arousal as measured by muscle tension, autonomic activity, and self-reported calmness? Although such studies lend credence to the validity of the proposed therapeutic mechanisms said to underlie these techniques, and although they support, to some degree, the treatment specificity hypothesis (see Lehrer & Woolfolk, 1984), they do not address the question of clinical efficacy for individual disorders. To achieve this, it is essential that studies include clinical samples that fulfil specified diagnostic criteria, that therapeutic change is assessed with respect to specific measures of psychological distress and functioning and not just arousal dysfunction, that suitable designs are employed to assess the contribution of nonspecific treatment effects, and that long-term follow-up assessments are also included.

Many reviews exist that focus on clinically relevant studies as described above. Some of these were discussed in Section 6.09.3.1 in relation to the effectiveness of different procedural components of relaxation techniques, such as homework compliance and taped instructions. Others have compared the general effectiveness of arousal reduction techniques in modifying psychological distress in patient samples. An early review by Glaister (1982) concluded that relaxation training leads to decreases in autonomic arousal both within and across sessions, and that these changes were associated with reductions in behavioral avoidance and distress. However, he suggested that the use of only relaxation is probably inferior to exposure treatments. Using a meta-analysis approach, Eppley, Abrams, and Shear (1989) compared the effects of different relaxation techniques on measures of trait anxiety. They concluded that transcendental meditation was superior to other techniques, and suggested that this might be related to the emphasis given to "effortless relaxation," the concentration of training sessions within the first few weeks, and the regularity and duration of practice. Expectancies and demand characteristics were said not to influence effect size. However, these conclusions should be treated cautiously, especially when generalizing to adult mental health populations, since studies involving psychiatric patients were excluded from these analyses, mainly because of the lack of application of transcendental meditation to this subject population. Similar narrative reviews have been published by Lehrer et al. (1994) and provide useful overviews of the comparative efficacy of relaxation techniques in relation to, for example, psychosomatic disorders, insomnia, anger management, and substance abuse.

Instead of focusing solely on arousal reduction techniques, more recent outcome studies have tended to address the comparative efficacy of a range of cognitive-behavioral interventions for specific disorders. This reflects the growth of psychotherapy outcome research for anxiety disorders and the continued inclusion of relaxation-based therapy components within these studies. However, it should be recognized that the rationale for the inclusion of arousal reduction methods has differed across individual studies. Some researchers, probably influenced by Mark's (1976) negative critique of the importance of relaxation in exposure treatment (but also see Wolpe, 1989), have included relaxation based techniques as control conditions in order to assess the additional beneficial effects of behavioral or cognitive treatment modalities. Other researchers (Borkovec & Costello, 1993; Ost, 1987) have employed relaxation protocols as treatments in their own right.

We will focus mainly on panic disorder, agoraphobia, and generalized anxiety disorder. Before reviewing work in these areas, however, the issue of whether relaxation training leads to any additional beneficial effects on fear reduction in simple phobias when compared to exposure alone needs to be addressed. McGlynn, Moore, Rose, and Lazarte (1995), in a study of clinically anxious snake-phobic students, reported that group abbreviated PMR training, in association with *in vivo* exposure, led to less behavioral avoidance, reduced autonomic activity, and lower levels of subjective distress compared to exposure alone. It would appear, therefore, that Marks had prematurely dismissed the potency of relaxation techniques in fear reduction.

The relationship between relaxation and exposure has been a focus of outcome research in both agoraphobia and panic disorder. Michelson, Mavissakalian, and Marchione (1988) compared paradoxical intention, therapist-assisted exposure, and relaxation on a range of outcome measures. Overall, the three treatments were equally efficacious and demonstrated significant improvements post-treatment and at a three-month follow-up. It should be noted that all groups underwent a general induction program that stressed the importance of self-directed exposure. In a later analysis, Michelson et al. (1990) examined differences in the three treatment modalities on psychophysiological processes and outcomes. Relaxation produced greater reductions in physiological arousal than did graded exposure, whereas paradoxical intention had no effect. However, despite earlier research indicating that patterns of synchrony–desynchrony might predict differential outcomes for these treatments (Turpin, 1990), no reliable relationships were obtained.

In a review of the treatment of panic disorder, Craske (1991) reports the superiority of a cognitive exposure group to relaxation alone on decreasing panic-attack frequency. However, relaxation was more effective at reducing general somatic symptoms and daily anxiety ratings during the course of therapy. Unfortunately, at 6- and 24-month follow-up, the relaxation group showed deterioration, while the exposure and combined exposure and relaxation groups demonstrated maintenance of therapeutic gains. Moreover, the relaxation alone group experienced higher attrition rates from therapy. In contrast, more positive results have been obtained by Ost and colleagues (Ost & Westling, 1995; Ost, Westling, & Hellstrom, 1993) using applied relaxation training (Ost, 1987). The first of these studies suggested that relaxation, exposure, and cognitive therapy were equally efficacious in the treatment of agoraphobia, if used in combination with self-exposure instructions. The more recent study attempted to replicate Clark et al.'s (1994) findings that cognitive-behavior therapy was superior to applied relaxation. Instead, Ost and Westling (1995) again reported generally equivalent outcomes for the two therapies. However, in agreement with the cognitive model of panic proposed by Clark (1986), the present study demonstrated that measures of cognitive beliefs predicted therapy outcomes, and this was equally true for both treatments. We discuss this finding further in the final section of this chapter.

Given the relative success of cognitive-behavioral treatments for agoraphobia and panic, psychosocial interventions have also been advocated for generalized anxiety disorder. Barlow, Rapee, and Brown (1992) compared relaxation, cognitive therapy, and their combination to a waiting-list control group. All three treatments were superior to the control condition on measures of worry and anxiolytic medication usage, and these gains were maintained at two-year follow-up. However, no differences emerged between treatment groups and following treatment most patients were left with significant residual symptoms. Moreover, the dropout rate was high among the active treatment groups. Similar results were obtained by Borkovec and Costello (1993), and indicated that cognitive-behavioral therapy (CBT) and applied relaxation were superior to a nondirective therapy control group. Some between-treatment differences did, however, suggest themselves. Applied relaxation appeared to favor greater reductions in daily anxiety ratings and diary measures, whereas CBT seemed more effective for worry and depression. At a 12-month follow-up, CBT showed greater maintenance of therapeutic gain (58% vs. 38% high end-state functioning) than applied relaxation. Following a review of existing outcome studies for generalized anxiety disorder, Borkovec and Whisman (1996) come to similar conclusions, and recommend a combination package of CBT and relaxation training, targeting both somatic and cognitive features of anxiety.

Finally, some mention should be made of outcome studies that have used nonrelaxation-based arousal reduction methods. Rice, Blanchard, and Purcell (1993) compared different biofeedback treatments (EMG and alpha EEG) with two control treatment conditions (EEG and pseudomeditation), and a waiting-list group. All four treatment groups, including the control treatment conditions groups, showed some improvement compared to the waiting-list control group. However, small sample sizes and a restricted number of treatment sessions limit the conclusions that can be drawn from this study. The utility of "mindfulness meditation" used in conjunction with a stress-management program for patients diagnosed with a variety of anxiety disorders (panic, agoraphobia, generalized anxiety disorder, etc.) is reported by Kabat-Zinn et al. (1992). In discussing the success of their program, the authors reflect on the potential benefits of "mindfulness meditation," and draw attention to the ability of clients to experience their thoughts directly as "just thoughts" in a nonreactive way, which leads to a decatastrophizing of the situation. It may be that this is a similar process identified by Borkovec and Roemer (1994) as "letting go" within applied relaxation, or the "effortless relaxation" discussed by Eppley et al. (1989) in relation to transcendental meditation.

6.09.4.2 Therapeutic Mechanisms

We wish to conclude this chapter by addressing the question of whether different therapeutic pathways can be identified, albeit rather speculatively, that share some commonality both within different arousal reduction methods and across cognitive-behavioral approaches in general. Starting with Lehrer and Woolfolk's treatment specificity position, which was derived from Davidson and Schwartz's specific effects model, we can classify a range of arousal reduction methods according to their putative therapeutic effects. Hence, progressive relaxation and biofeedback are essentially considered as somatic or physiologically oriented techniques, whereas meditation and cognitive stress management approaches are regarded as cognitive techniques. The logic of this model, therefore, is to match the therapeutic method to the presenting problem. Hence, it is suggested that somatically based treatments might benefit problems characterized by excessive autonomic and muscular arousal. Implicit in this approach is a rather simplistic opponent process model of therapy, whereby presenting problems are ameliorated by the addition of an opposing therapeutic component. However, recent outcome studies appear not to support such a simple relationship.

If panic disorder is considered, the specific effects model would specify that relaxation ought to target more somatic change, whereas cognitive therapy should give rise to greater attitudinal shifts, etc. Although there has been some support for this suggestion (Craske, 1991), in as much as relaxation was reported to bring about greater reduction in daily symptomatic ratings whereas cognitive approaches brought about a greater decrease in overall panic frequency, the majority of outcome studies have yielded more equivocal results. For example, Ost and Westling have reported that both applied relaxation and cognitive therapy yield roughly equivalent positive outcomes for agoraphobia. However, the major predictor of overall therapeutic outcome across both treatments were measures of cognitive beliefs (Ost & Westling, 1995; Westling & Ost, 1995). This would suggest that a common pathway involving cognitive beliefs mediated therapeutic change for both these interventions, as predicted by Clark's (1986) cognitive model of panic.

The question arises, therefore, of how the specific effects model can accommodate findings such as the above. The answer is probably to be found in more recent models of therapeutic change, which involve concepts such as emotional processing (Borkovec, 1994; Foa & Kozak, 1986; Rachman, 1980) and give rise to more dynamic explanations of change that involve an interplay between a variety of more discrete psychological processes such as physiological arousal, preattentive biases, affective response, associative networks, imagery, cognitive avoidance, worry, and extinction. The exact nature of the interplay depends on the therapeutic technique considered, together with the nature of the clinical problem. Within this context, we would argue that arousal reduction techniques might yield more subtle therapeutic gains through a variety of possible pathways. To illustrate this, we provide here some speculative examples of how relaxation techniques might yield the suitable conditions that bring about enhanced emotional processing and therapeutic change.

Before we can discuss the link between arousal reduction and emotional processing, we need to outline what is meant by the latter. According to Foa and Kozak (1986), emotional processing occurs when maladaptive fear structures are accessed in memory repeatedly, and fully processed, resulting in affective and physiological responding. In addition, this processing results in changes to the structure of the associative network responsible for representing the meaning of the fear object. With repeated exposure, together with full processing, extinction of the fear cues and their associations is brought about. Factors that are said to influence this process are level of arousal, early detection of threat cues, cognitive engagement and avoidance, depth of affective processing and emotional responding, worry, and distraction, together with overarching changes in attitude associated with mastery, coping, safety, and other metacognitive structures.

We would argue, therefore, that in the case of emotional disorder, arousal reduction techniques yield a variety of therapeutic effects through their ability to enhance emotional processing. We further speculate that these techniques provide several quite discrete mechanisms through which such facilitation might be achieved. These influences of arousal reduction techniques on emotional processing and consequent therapeutic change are outlined in Table 1. First, we suggest that the consequences of gaining competence in any arousal reduction/relaxation induction technique can be considered with respect to three general areas of functioning. These are direct changes in physiological activity, changes in ongoing attention and other associative processes (e.g., access to associative networks), and longer term cognitive, attitudinal, and behavioral changes.

Clearly, many relaxation techniques are said to result in altered physiological functioning, such as greater cortical deactivation, increased parasympathetic activity, decreased sympathetic activity, and lowered muscle tension. These changes may well be intrinsically therapeutic, since they represent direct therapeutic targets in their own right (i.e., reduced muscle tension, lowered blood pressure, etc.) and may also be associated with longer term physiological benefits with respect to neuroendocrine and immune functioning. We would argue, moreover, that these changes in physiological arousal may also lead to therapeutic changes within a second area of functioning known as "emotional processing." Several authors (Borkovec, Lyonfields, Wiser, & Deihl, 1993; Foa & Kozak, 1986; Shapiro, 1995; Turpin, 1991) have suggested that the depth of processing might be enhanced by relaxation. An associated observation from Borkovec's research is the observation that generalized anxiety disorder and worry are associated with autonomic inflexibility, and that therapeutic improvement is associated with an increase in parasympathetic activity. It is too early, however, to speculate about the direction of this relationship. Finally, the association with decreased arousal and extinction of fear has been fundamental to many accounts of desensitization (Lader & Mathews, 1970) and is still relevant to more contemporary models of fear conditioning (Davey & Matchett, 1996) and other therapeutic procedures such as eye-movement desensitization and post-traumatic stress disorder (MacCulloch & Feldman, 1996; Shapiro, 1995).

Table 1 Targets for potential therapeutic change arising from the implementation of arousal reduction techniques.

Physiological

Alteration of autonomic and endocrine functioning:
Central deactivation
Increased parasympathetic activity
Decreased sympathetic activity
Reduced muscle tension
Altered breathing
Endocrine changes
Immune function changes

Emotional processing

Changes in fear structures and memory:
Preattentive biases and early cue-detection of threat
Associative networks consisting of stimulus and responses propositions Accessibility to network and imagery
Changes in associative strengths and meaning
Cognitive avoidance and worry

Cognitive change

Changes in sensation, attributions, and beliefs:
Altered subjective experience of tension and calmness
Redistribution of focused attention
Reattributions of anxiety and panic cues/situations
Reduction in worry and metacognitions
Enhanced mastery and coping
Commitment to approach and direct exposure of feared object/situation

Changes in physiological activity may also be represented indirectly as altered interoreceptive feedback within a third general area involving cognitive change. Since it has been argued (Clark, 1986) that nonspecific bodily sensations might provide a cue for catastrophic attributions within panic disorder, any modification of these sensations, either due to relaxation or even medication, might have beneficial effects. This approach has been specifically developed further by Borkovec and Roemer (1994), whereby applied relaxation is combined with early cue detection and self-control desensitization. It would be expected that decreases in subjective bodily sensations and tension would also be accompanied by increases in subjective calmness, etc. Finally, a word of caution is warranted, since it has also been argued that a potential therapeutic strategy is to train clients to disattend to interoreceptive sensations (Wells, 1990). Somatically based relaxation techniques might contradict such an approach, and this may be apposite when clients occasionally experience relaxation-induced panic.

The extent to which interoceptive feedback or any other anxiety-related cues, such as distressing imagery or worrisome thoughts, are detected and attended to might also be affected by arousal reduction techniques. In particular, techniques that focus on cognitive strategies as a means of controlling, for example, worry, troublesome imagery, and unpleasant thoughts, such as guided imagery, autogenics, meditation, and hypnotic relaxation procedures, might also have an important role in emotional processing. Although these techniques may not necessarily have a direct facilitative role, they may be expected to have therapeutic benefits by counteracting other maladaptive cognitive phenomena such as worry and disturbing imagery which, according to Borkovec (1994), inhibit emotional processing. Worry is said to lead to cognitive avoidance, suppression of affective responding to imagery, delayed habituation, and the preservation of intrusive thoughts and images. The degree to which various cognitively directed relaxation strategies can result in a client gaining control of these negative cognitive strategies requires further confirmation and research. Interestingly, as discussed in Section 6.09.4.1, there are several different observations that have independently stressed the importance of decontextualizing negative thoughts or images by such techniques as "letting go," "effortless relaxation," "just thoughts," or "mindfulness meditation." Moreover, the therapeutic rationale provided with various relaxation techniques, together with the nature of the therapeutic relationship and the experience of therapy, should contribute to positive attitudinal change in relation to self-coping and efficacy, re-attributions of anxiety and panic, and overall changes in life-style and outlook. Indeed, Smith (1988) has emphasized the importance generally of the cognitive components of relaxation training and cognitive restructuring.

6.09.5 CONCLUSION

In this chapter we have attempted to provide a theoretical context within which different arousal reduction techniques might be evaluated. We have briefly reviewed the practical applications of these techniques and attempted some discussion of their relative efficacy, together with a review of the relevant outcome data. From this review, we would suggest that relaxation techniques have been shown to be beneficial for a variety of physical and mental health problems, particularly in the short term and during the course of ongoing therapy. Although other techniques involving more cognitively oriented approaches have tended to eclipse relaxation-based therapies, we would argue that therapists continue to give these methods serious consideration, particularly when used in conjunction with other cognitive-behavioral methods. This recommendation is based on evidence from outcome studies that have demonstrated a degree of efficacy, even for relaxation applied alone, and also from consideration of potential factors that might, albeit rather speculatively, affect the depth and efficiency of emotional processing.

6.09.6 REFERENCES

Adler, C. M., Craske, M. G., & Barlow, D. H. (1987). Relaxation-induced panic (RIP): When resting isn't peaceful. *Integrative Psychiatry, 5,* 94–112.

Adler, C. S., & Morrisey-Adler, S. (1983). Stategies in general psychiatry. In J. Basmajian (Ed.), *Biofeedback: Principles and practice for clinicians.* Baltimore: Williams & Wilkins.

Alexander, A. B. (1975). An experimental test of assumptions relating to the use of electromyographic biofeedback as a general relaxation training technique. *Psychophysiology, 12,* 656–662.

Alladin, A., & Heap, M. (1991). Hypnosis and depression. In M. Heap & W. Dryden (Eds.), *Hypnotherapy: A handbook* (pp. 49–67). Buckingham, UK: Open University Press.

Andreassi, J. L. (1989). *Psychophysiology: human behaviour and physiological response.* New York: Oxford University Press.

Andrew, R. J. (1974). Arousal and the causation of behaviour. *Behaviour, 51,* 135–165.

Bacon, M., & Poppen, R. (1985). A behavioral analysis of diaphragmatic breathing and its effects on peripheral temperature. *Journal of Behavior Therapy and Experimental Psychiatry, 16,* 15–21.

Barlow, D. H., Rapee, R. M., & Brown, T. A. (1992). Behavioural treatment of generalized anxiety disorder. *Behaviour Therapy, 23,* 551–570.

Barrowclough, C., & Tarrier, N. (1992). *Families of schizophrenic patients.* London: Chapman & Hall.

Benson, H. (1975) *The relaxation response.* New York: William Morrow.

Benson, H., & Wallace, R. K. (1971). Decreased drug abuse with transcendental meditation. *Congressional Record, 92nd Congress, First Session*, Serial No. 92–1. Washington, DC: US Congress.

Bernstein, D. A., & Borkovec, T. D. (1973). *Progressive relaxation training: A manual for the helping professions.* Champaign, IL: Research Press.

Bernstein, D. A., & Carlson, C. R. (1993) Progressive relaxation: Abbreviated methods. In P. M. Lehrer & R. L. Woolfolk (Eds.), *Principles and practice of stress management* (2nd ed., pp. 53–88). New York: Guilford Press.

Blanchard, E. B., Nicholson, N. L., Taylor, A. E., Steffek, B. D., Radnitz, C. L., & Appelbaum, K. A. (1991). The role of regular home practice in the relaxation treatment of tension headache. *Journal of Consulting and Clinical Psychology, 59,* 467–470.

Borkovec, T. D. (1979). Pseudo-(experimental) insomnia and idiopathic (objective) insomnia: Theoretical and therapeutic issues. *Advances in Behaviour Research and Therapy, 2,* 27–55.

Borkovec, T. D. (1987) Commentary. *Integrative Psychiatry, 5,* 104–106.

Borkovec, T. D. (1994). The nature, functions, and origins of worry. In G. C. L. Davey & F. Tallis (Eds.), *Worry: Perspectives on theory, assessment, and treatment* (pp. 5–34). Chichester, UK: Wiley.

Borkovec, T. D., & Costello, E. (1993). Efficacy of applied relaxation and cognitive–behavioural therapy in the treatment of generalized disorder. *Journal of Consulting and Clinical Psychology, 61,* 611–619.

Borkovec, T. D., Lyonfields, J. D., Wiser, S. L., & Deihl, L. (1993). Case histories and shorter communications. *Behaviour Research and Therapy, 31,* 321–324.

Borkovec, T. D., Mathews, A. M., Chambers, A., Ebrahimi, S., Lytle, R., & Nelson, R. (1987). The effects of relaxation training with cognitive therapy or non-directive therapy and the role of relaxation-induced anxiety in the treatment of generalized anxiety. *Journal of Consulting and Clinical Psychology, 55,* 883–888.

Borkovec, T. D., & Roemer, L. (1994). Generalized anxiety disorder. In M. Herson & R. T. Ammerman (Eds.), *Handbook of prescriptive treatments for adults* (pp. 261–281). New York: Plenum.

Borkovec, T. D., & Sides, J. K. (1979). Critical procedural variables related to the physiological effects of progressive relaxation: A review. *Behaviour Research and Therapy, 17,* 119–126.

Borkovec, T. D., & Whisman, M. A. (1996). Psychosocial treatment for generalized anxiety disorder. In M. R. Mavissakalian & R. Prien (Eds.), *Long-term treatments of anxiety disorders* (pp. 171–199). Washington, DC: American Psychiatric Association Press.

Braith, J. A., McCullough, J. P., & Bush, J. P. (1988). Relaxation-induced anxiety in a subclinical sample of chronically anxious subjects. *Journal of Behaviour, Therapy and Experimental Psychiatry, 19,* 193–198.

Budzynski, T. H., & Stoyva, J. M. (1973). Biofeedback techniques in behaviour therapy. In N. Birbaumer (Ed.), *Reihe Fortschritte der klinischen psychologie: Vol. 3. Neuropsychologie der Angst* (pp. 248–270). Munich, Germany: Urban & Schwarzenberg.

Burrows, G. D. (1980). Affective disorders and hypnosis. In G. D. Burrows & L. Dennerstein (Eds.), *Handbook of hypnosis and psychosomatic medicine* (pp. 149–170). New York: Elsevier.

Cacioppo, J. T., Berntson, G. G., & Andersen, B. L. (1991). Psychophysiological approaches to the evaluation of psychotherapeutic process and outcome: Contributions from social psychophysiology. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 321–336.

Cacioppo, J. T., & Tassinary, L. G. (1990). *Principles of psychophysiology. Physical, social, and inferential elements.* New York: Cambridge University Press.

Carrington, P. (1977). *Freedom in meditation.* Garden City, NY: Doubleday/Anchor.

Carrington, P. (1993). Modern forms of meditation. In P. M. Lehrer & R. L. Woolfolk (Eds.), *Principles and practice of stress management* (2nd ed., pp. 139–168). New York: Guilford Press.

Claridge, G. (1987). Psychoticism and arousal. In J. Strelau & H. J. Eysenck (Eds.), *Personality dimensions and arousal.* New York: Plenum.

Clark, D. (1986). A cognitive approach to panic. *Behaviour Research and Therapy, 24,* 461–470.

Clark, D. M., Salkovskis, P. M., & Chalkley, A. J. (1985). Respiratory control as a treatment for panic attacks. *Journal of Behaviour Therapy and Experimental Psychology, 16,* 23–30.

Clark, D. M., Salkovskis, P. M., Hackman, A., Middleton, H., Anastasiades, P., & Gelder, M. (1994). A comparison of cognitive therapy, applied relaxation, and imipramine in the treatment of panic disorder. *British Journal of Psychiatry, 164,* 759–769.

Clements, K., & Turpin, G. (1992). Vulnerability models and schizophrenia: The assessment and prediction of relapse. In M. Birchwood & N. Tarrier (Eds.), *Innovations in the psychological management of schizophrenia* (pp. 21–48). Chichester, UK: Wiley.

Corbett, J. A., & Turpin, G. (1985). Tics and Tourette's syndrome. In M. Rutter & L. Hersov (Eds.), *Child and adolescent psychiatry* (pp. 516–525). Oxford, UK: Blackwell Scientific.

Cottraux, J. A., Jeunet, C., & Collet, L. (1983). The treatment of writer's cramp with multimodal behaviour therapy and biofeedback: A study of 15 cases. *British Journal of Psychiatry, 142,* 180–183.

Council, J. R., Kirsch, I., & Hafner, L. P. (1986). Expectancy vs. absorption in the prediction of hypnotic responding. *Journal of Consulting and Clinical Psychology, 50,* 182–189.

Craske, M. G. (1991). Models and treatment of panic: Behavioural therapy of panic. *Journal of Cognitive Psychotherapy: An International Quarterly, 5,* 199–214.

Crawford, H. J., Hilgard, J. R., & Macdonald, H. (1982). Transient experiences following hypnotic testing and special termination procedures. *International Journal of Clinical and Experimental Hypnosis, 30,* 117–126.

Crist, D. A., Rickard, H. C., Prentice-Dunn, S., & Baker, H. R. (1989). The relaxation inventory: Self-report scales of relaxation training effects. *Journal of Personality Assessment, 53,* 716–726.

Davey, G. L., & Matchett, G. (1996). The effects of response feedback on conditioned responding during extinction: implications for the role of interoception in anxiety based disorders. *Journal of Psychophysiology, 10,* 291–302.

Davidson, R., & Goleman, D. (1977). The role of attention in meditation and hypnosis: A psychobiological perspective on transformation of consciousness. *International Journal of Clinical and Experimental Hypnosis, 25,* 291–308.

Davidson, R. J., & Schwartz, G. E. (1976). The psychobiology of relaxation and related states: A multiprocess theory. In *E. Mostofsky* (Ed.), *Behavior control and modification of physiological activity.* New York: Prentice-Hall.

Deberry, S., Davis, S., & Reinhard, K. E. (1989). A comparison of meditation-relaxation and cognitive/behavioural techniques for reducing anxiety and depression in a geriatric population. *Journal of Geriatric Psychiatry, 22,* 231–247.

Duffy, E. (1951). The concept of energy mobilization. *Psychological Review, 58,* 30–40.

Dumas, R. A. (1980). Cognitive control in hypnosis and biofeedback. *International Journal of Clinical and Experimental Hypnosis, 28,* 53–62.

Edinger, J. D., & Jacobsen, R. (1982). Incidence and significance of relaxation treatment side effects. *Behavior Therapy, 5,* 137–138.

Edmonds, G. M. (1979). Hypnotic facilitation of temperature biofeedback. *Dissertation Abstracts International, 39*(9-B), 4611.

Eisenberg, D., Kessler, R. C., Foster, C., Norlock, F. E., Calkins, D. R., & Delbanco, T. L. (1993). Unconventional medicine in the United States. Prevalence, costs and patterns of use. *New England Journal of Medicine, 328,* 246–252.

Engel, B. T., & Hansen, S. P. (1966) Operant conditioning of heart rate slowing. *Psychophysiology, 3,* 176–187.

Eppley, K. R., Abrams, A. I., & Shear, J. (1989). Differential effects of relaxation techniques on trait anxiety: A meta-analysis. *Journal of Clinical Psychology, 45,* 957–974.

Fellows, B. J. (1988). Hypnotic susceptibility scales. In M. Heap (Ed.), *Hypnosis: Current clinical, experimental and forensic practices* (pp. 40–50). London: Croom Helm.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99,* 20–35.

Freedman, R. R. (1989). Raynaud's disease. In G. Turpin (Ed.), *Handbook of clinical psychophysiology* (pp. 469–494). Chichester, UK: Wiley.

Freedman, R. R. (1994). Adrenergic mechanisms in behavioral interventions. In O. G. Cameron (Ed.), *Adrenergic dysfunction and psychobiology* (pp. 299–301). Washington, DC: American Psychiatric Press.

Fried, R. (1993). The role of respiration in stress control: Toward a theory of stress as a hypoxic phenomenon. In P. M. Lehrer & R. L. Woolfolk (Eds.), *Principles and practice of stress management* (pp. 301–332). New York: Guilford Press.

Gale, A., & Eysenck, M. W. (1992). *Handbook of individual differences: Biological perspectives.* Chichester, UK: Wiley.

Gaston, L. (1988–1989). Efficacy of imagery and meditation techniques in treating psoriasis. *Imagination, Cognition and Personality, 13,* 155–173.

Gatchel, R. J. (1993). Psychophysiological disorders: Past and present perspectives. In R. J. Gatchel & E. B. Blanchard (Eds.), *Psychophysiological disorders* (pp. 1–22). Washington, DC: American Psychological Association.

Gellhorn, E. (1958). The physiological basis of neuromuscular relaxation. *Archives of Internal Medicine, 102,* 392–399.

Gellhorn, E., & Kiely, W. F. (1972). Mystical states of consciousness: Neurophysiological and clinical aspects. *Journal of Nervous and Mental Disease, 154,* 399–405.

Gibson, H. B., & Heap, M. (1991). *Hypnosis in therapy.* Hove, UK: Erlbaum.

Glaister, B. (1982). Muscle relaxation training for fear reduction of patients with psychological problems: A review of controlled studies. *Behaviour Research and Therapy, 20,* 493–504.

Grimm, L.G. (1980). The evidence for cue-controlled relaxation. *Behaviour Research and Therapy, 11,* 283–293.

Gruber, B. L., Hersh, S. P., Hall, N. R. S., Waletzky, L. R., Kunz, J. F., Carpenter, J. K., Kverno, K. S., & Weiss, S. M. (1993). Immunological responses of breast cancer patients to behavioural interventions. *Biofeedback and Self-regulation, 18,* 1–21.

Heap, M. (1991) Role and uses of hypnosis in psychotherapy. In M. Heap & W. Dryden (Eds.), *Hypnotherapy: A handbook* (pp. 15–30). Buckingham, UK: Open University Press.

Heap, M. (1995). Hypnosis: Applications in general medicine. *British Journal of Therapy and Rehabilitation, 12,* 649–654.

Heide, F., & Borkovec, T. D. (1983). Relaxation-induced anxiety: Paradoxical anxiety enhancement due to relaxation training. *Journal of Consulting and Clinical Psychology, 51,* 171–182.

Hilgard, E. R. (1991). A neo-dissociation interpretation of hypnosis. In S. J. Lynn & J. W. Rhue (Eds.), *Theories of hypnosis: Current models and perspectives* (pp. 83–104). New York: Guilford Press.

Hillenberg, J. B., & Collins, F. L. (1982). A procedural analysis and review of relaxation training research. *Behaviour Research and Therapy, 20,* 251–260.

Hjemdahl, P., Fagius, J., Freyschuss, U., Wallin, B. G., Daleskog, M., Bohlin, G., & Perski, A. (1989). Muscle sympathetic activity and norepinephrine release during mental challenge in humans. *American Journal of Physiology, 257,* E654–E664.

Hoelscher, T. J., Lichtenstein, K. L., & Rosenthal, T. L. (1986) Home relaxation practice in hypertension treatment: Objective assessment and compliance induction. *Journal of Consulting and Clinical Psychology, 54,* 217–221.

Honsberger R. W., & Wilson, A. F. (1973). Transcendental meditation in treating asthma. *Respiratory Therapy: The Journal of Inhalation Technology, 3,* 79–80.

Howell, M., & Whitehead, J. (1989). *Survive stress: A training programme.* Cambridge, UK: East Anglian Regional Health Authority/Winslow Press.

Jacob, R. G., Chesney, M. A., William, D. M., Ding, Y., & Shapiro, A. P. (1991). Relaxation therapy for hypertension: Design effects and treatment effects. *Annals of Behavioral Medicine, 13,* 5–15.

Jacobson, E. (1938). *Progressive relaxation.* Chicago: University of Chicago Press.

Johnston, D. W. (1994). Hypertension: Psychological factors in aetiology and management. *The Irish Journal of Psychology, 15,* 27–42.

Kabat-Zinn, J., Massion, A. O., Kristeller, J. Peterson, L. G., Fletcher, K. E., Pbert, L., Lenderking, W. R., & Santorelli, S. F. (1992). Effectiveness of a meditation-based stress reduction program in the treatment of anxiety disorders. *American Journal of Psychiatry, 149,* 936–943.

Kamiya, J. (1969). Operant control of the EEG alpha rhythm and some of its reported effects on consciousness. In C. T. Tart (Ed.), *Altered states of consciousness* (pp. 507–517). New York: Wiley.

Karle, H. (1988). Hypnosis in analytical psychotherapy. In M. Heap (Ed.), *Hypnosis: Current clinical, experimental and forensic practices* (pp. 208–220). London: Croom Helm.

Katkin, E. S., & Murray, E. N. (1968). Instrumental conditioning of autonomically mediated behaviour. *Psychological Bulletin, 70,* 52–68.

Kay, J. A., & Carlson, C. R. (1992). The role of stretch-based relaxation in the treatment of chronic neck tension. *Behaviour Therapy, 23,* 423–431.

Kirsch, I. (1996) Hypnosis in psychotherapy: Efficacy and mechanisms. *Contemporary Hypnosis, 13,* 109–114.

Kirsch, I., & Henry, D. (1979). Self-desensitization and meditation in the reduction of public speaking anxiety. *Journal of Consulting and Clinical Psychology, 47,* 536–541.

Kozak, M. J., & Miller, G. A. (1982). Hypothetical constructs vs. intervening variables: A reappraisal of the three-system model of anxiety assessment. *Behavioral Assessment, 4,* 347–358.

Lacey, J. I. (1967). Somatic response patterning and stress: Some revisions of activation theory. In M. H. Appley & R. Trumbull (Eds.), *Psychological stress* (pp. 14–38). New York: Appleton-Century-Crofts.

Lader, M. H. (1975). *The psychophysiology of mental illness.* London: Routledge and Kegan Paul.

Lader, M., & Mathews, A. (1970). Physiological changes during spontaneous panic attacks. *Journal of Psychosomatic Research, 14,* 377–382.

Lake, A. E., & Pingel, J. D. (1988). Brief vs. extended relaxation: Relationship to improvement at follow-up in mixed headache patients. *Medical Psychotherapy, 1,* 119–129.

Lang, P. J. (1968). Fear reduction and fear behaviour: Problems in treating a construct. In J. M. Shlien (Ed.), *Research in psychotherapy.* Washington, DC: American Psychological Society.

Lazarus, A. A. (1976). Psychiatric problems precipitated by transcendental meditation. *Psychological Reports, 10,* 39–74.

Lehrer, P. M. (1996). Varieties of relaxation methods and their unique effects. *International Journal of Stress Management, 3,* 1–15.

Lehrer, P. M., Carr, R., Sargunaraj, D., & Woolfolk, R. L. (1993). Differential effects of stress management therapies in behavioral medicine. In P. M. Lehrer & R. L. Woolfolk (Eds.), *Principles and practice of stress management* (2nd ed., pp. 571–605). New York: Guilford Press.

Lehrer, P. M., Carr, R., Sarunaraj, D., & Woolfolk, R. L. (1994). Stress management techniques: Are they all equivalent, or do they have specific effects? *Biofeedback and Self-Regulation, 19,* 353–401.

Lehrer, P. M., & Woolfolk, R. L. (1984). Are stress reduction techniques interchangeable or do they have specific effects? A review of the comparative empirical literature. In R. L. Woolfolk & P. M. Lehrer (Eds.), *Principles and practice of stress management* (1st ed., pp. 404–477). New York: Guilford Press.

Lehrer, P. M., & Woolfolk, R. L. (1993). Research on clinical issues in stress management. In P. M. Lehrer & R. L. Woolfolk (Eds.), *Principles and practice of stress management* (2nd ed., pp. 521–538). New York: Guilford Press.

Lehrer, P. M., Woolfolk, R. L., Rooney, A. J., McCann, B., & Carrington, P. (1983). Progressive relaxation and meditation. A study of psychophysiological and therapeutic differences between two techniques. *Behaviour Research and Therapy, 21,* 651–662.

Ley, R. (1980). Panic attacks during relaxation and relaxation-induced anxiety: A hyperventilation interpretation. *Journal of Behaviour Therapy and Experimental Psychiatry, 19,* 253–259.

Linden, W. (1990). *Autogenic training: A clinical guide.* New York: Guilford Press.

Linden, W. (1993). The autogenic training method of J. H.

Schultz. In P. M. Lehrer & R. L. Woolfolk (Eds.), *Principles and practice of stress management* (2nd ed., pp. 205–230). New York: Guilford Press.

Linden, W., & Chambers, L. (1994). Clinical effectiveness of non-drug treatment for hypertension: A meta-analysis. *Annals of Behavioral Medicine, 16*, 35–45.

Linden, W., Stossel, C., & Maurice, J. (1996). Psychosocial interventions for patients with coronary artery disease. *Archives of Internal Medicine, 156*, 745–752.

MacCulloch, M. J., & Feldman, P. (1996). Eye-movement desensitization-treatment utilizes the positive visceral element of the investigatory reflex to inhibit the memories of post-traumatic-stress-disorder—a theoretical analysis. *British Journal of Psychiatry, 168*, 571–579.

Marks, I. M. (1976). Current status of behavioral psychotherapy: Theory and practice. *American Journal of Psychiatry, 133*, 253–261.

Mathew, R. J., Ho, B. T., Kralick, P., Weinman, M., & Claghorn, J. L. (1981). Anxiety and platelet MAO levels after relaxation training. *American Journal of Psychiatry, 138*, 371–373.

McEwen, B. S., & Stellar, E. (1993). Stress and the individual. *Archives of International Medicine, 153*, 2093–2101.

McGlynn, F. D., Moore, P. M., Rose, M. P., & Lazarte, A. (1995). Effects of relaxation training on fear and arousal during *in vivo* exposure to a caged snake among *DSM-III-R* simple (snake) phobics. *Journal of Behavioural Therapy and Experimental Psychiatry, 26*, 1–8.

McGuigan, F. J. (1993). Progressive relaxation: Origins, principles, and clinical applications. In P. M. Lehrer & R. L. Woolfolk (Eds.), *Principles and practice of stress management* (2nd ed., pp. 17–52). New York: Guilford Press.

Meares. R. (1973). Electromyographic studies in patients with spasmodic torticollis. *Journal of Psychosomatic Research, 15*, 13–18.

Meichenbaum, D. (1993). Stress innoculation training: A 20-year update. In P. M. Lehrer & R. L. Woolfolk (Eds.), *Principles and practice of stress management* (2nd ed., pp. 373–406). New York: Guilford Press.

Michelson, L., Mavissakalian, M., & Marchione, K. (1988). Cognitive, behavioral and psychophysiological treatments of agoraphobia: A comparative outcome investigation. *Behavior Therapy, 19*, 97–120.

Michelson, L., Mavissakalian, M., Marchione, K., Ulrich, R. F., Marchione, N., & Testa, S. (1990). Psychophysiological outcome of cognitive, behavioural and psychophysiologically-based treatments of agoraphobia. *Behaviour Research and Therapy, 28*, 127–139.

Mills, I., & Salkovskis, P. M. (1988). Mood and habituation to phobic stimuli. *Behaviour Research and Therapy, 26*, 435–439.

Montgomery, G. T. (1994). Slowed respiration training. *Biofeedback and Self-Regulation, 19,* 211–225.

Nelson. R. A., & Borkovec, T. S. (1989). Relationship of client participation to psychotherapy. *Journal of Behavioural Therapy and Experimental Psychiatry, 20,* 155–162.

Nicassio, P. M., & Buchanan, D. C. (1981). Clinical application of behaviour therapy for insomnia. *Comprehensive Psychiatry, 22,* 512–521.

Norris, P. A., & Fahrion, S. L. (1993). Autogenic biofeedback in psychophysiological therapy and stress management. In P. M. Lehrer & R. L. Woolfolk (Eds.), *Principles and practice of stress management* (2nd ed., pp. 231–262). New York: Guilford.

Nuechterlein, K. H., & Dawson, M. E. (1984). A heuristic vulnerability/stress model of schizophrenic episodes. *Schizophrenia Bulletin, 10,* 300–312.

Öhman, A. (1987). The psychophysiology of emotion: An evolutionary-cognitive perspective. *Advances in Psychophysiology, 2,* 79–127.

Öhman, A. (1992). Orienting and attention: Preferred preattentive processing of potentially phobic stimuli. In B. A. Campbell, H. Haynes, & (Eds.), *Attention and information processing in infants and adults: perspectives from human and animal research*. Hillsdale. NY: Erlbaum.

Ost, L. (1987). Applied relaxation: Description of a coping technique and review of controller studies. *Behaviour Research and Therapy, 25,* 379–397.

Ost, L. G., & Westling, B. E. (1995). Applied relaxation vs. cognitive behaviour therapy in the treatment of panic disorder. *Behaviour Research and Therapy, 33,* 145–158.

Ost, L.-G., Westling, B. E., & Hellstrom, K. (1993). Applied relaxation exposure in vivo and cognitive methods in the treatment of panic disorder with agoraphobia. *Behaviour Research and Therapy, 31,* 383–394.

Parker, J. C., Gilbert, A. S., & Thoreson, R. W. (1978). Reduction of autonomic arousal in alcoholics. *Journal of Consulting and Clinical Psychology, 46,* 879–886.

Patel, C. (1993). Yoga-based therapy. In P. M. Lehrer & R. L. Woolfolk (Eds.), *Principles and practice of stress management* (2nd ed., pp. 89–138). New York: Guilford Press

Patel, C., Marmot, M. G., Terry, D. J., Carruthers, M., Hunt, B., & Patel, M. (1985). Trial of relaxation in reducing coronary risk: Four year follow up. *British Medical Journal, 290,* 1103–1106.

Paul, G. L. (1966). *Insight vs. desensitization in psychotherapy: An experiment in anxiety reduction*. Stanford, CA: Stanford University Press.

Payne, R. A. (1995). *Relaxation techniques: A practical handbook for the health care professional*. Edinburgh, UK: Churchill Livingstone.

Poppen, R. (1988). *Behavioral relaxation training and assessment*. New York: Pergamon.

Rachman, S. J. (1980). Emotional processing. *Behaviour Research and Therapy, 18,* 51–60.

Rachman, S. J., & Hodgson, R. (1974). Synchrony and desynchrony in fear and avoidance. *Behaviour Research and Therapy, 12,* 311–318.

Reynolds, S., & Shapiro, D. A. (1991). Stress reduction in transition: Conceptual problems in the design, implementation and evaluation of work-site stress management interventions (SMIs). *Human Relations, 44,* 717–733.

Rice, K. M., Blanchard, E. B., & Purcell, M. (1993). Biofeedback treatments of generalized anxiety disorder: Preliminary results. *Biofeedback and Self-Regulation, 18,* 93–105.

Sanders, S. (1993). Clinical self-hypnosis: Transformation and subjectivity. In J. W. Rhue, S. J. Lynn, & I. Kirsch (Eds.), *Handbook of clinical hypnosis* (pp. 251–270). Washington, DC: American Psychological Association.

Shilling, D. J., & Poppen, R. (1983). Behavioral relaxation training and assessment. *Journal of Behaviour Therapy and Experimental Psychiatry, 14,* 99–107.

Scogin, F., Rickard, H. C., Keith, S., Wilson, J., & McElreath, L. (1992). Progressive and imaginal relaxation training for elderly persons with subjective anxiety. *Psychology and Ageing, 7,* 419–424.

Selye, H. (1950). *Stress.* Montreal, Canada: ACTA.

Shapiro, D. H. (1982) Clinical and physiological comparison of meditation and other self-control strategies. *American Journal of Psychiatry, 139,* 267–274.

Shapiro. D., Crider, A. B., & Tursky, B. (1964). Differentiation of an autonomic response through operant reinforcement. *Psychonomic Science, 1,* 147–148.

Shapiro, F. (1995). *Eye movement desensitization and reprocessing: Basic principles, protocols and procedures.* New York: Guilford Press

Sigman, A. (1988). Hypnotic and biofeedback procedures for self-regulation of autonomic nervous system functions. In M. Heap (Ed.), *Hypnosis: current clinical, experimental and forensic practices* (pp. 126–138). London: Croom Helm.

Sigman, A., & Phillips, K. R. (1985). Biofeedback and hypnosis. A review of the recent literature. *British Journal of Experimental and Clinical Hypnosis, 3,* 13–24.

Smith, J. C. (1988). Steps toward a cognitive-behavioural model of relaxation. *Biofeedback and Self-Regulation, 13,* 307–329.

Smith, J. C. (1989). *Relaxation dynamics: A cognitive–behavioral approach to relaxation*. Champaign, IL: Research Press.

Spinhoven, P. (1987). Hypnosis and behavior therapy: A review. *International Journal of Clinical and Experimental Hypnosis, 35,* 8–31.

Stanton, H. E. (1990). Use of ego enhancement to increase assertiveness. *British Journal of Experimental and Clinical Hypnosis, 7,* 133–137.

Stanton, H. E. (1991). The reduction in secretarial stress. *Contemporary Hypnosis, 1,* 45–50.

Stein, C. (1963). The clenched fist technique as a hypnotic procedure in clinical psychotherapy. *American Journal of Clinical Hypnosis, 6,* 113–119.

Steptoe, A. (1980). Stress and medical disorders. In S. J. Rachman (Ed.), *Contributions to medical psychology*. Oxford, UK: Pergamon.

Steptoe, A

. (1991). The links between stress and illness. *Journal of Psychosomatic Research, 35,* 633–644.

Sternbach, R. (1966). *Principles of psychophysiology*. New York: Academic Press.

Stoyva, J. M., & Budzynski, T. H. (1993). Biofeedback methods in the treatment of anxiety and stress disorders. In P. M. Lehrer & R. L. Woolfolk (Eds.), *Principles and practice of stress management* (2nd ed., pp. 263–300). New York: Guilford Press

Surut, R. S., Shapiro, D., & Good, M. I. (1978). Comparison of cardiovascular biofeedback, neuromuscular feedback, and meditation in the treatment of borderline hypertension. *Journal of Consulting and Clinical Psychology, 46,* 252–263.

Tarrier, N., & Turpin, G. (1992). Psychosocial factors, arousal and schizophrenic relapse—the psychophysiological data. *British Journal of Psychiatry, 161,* 3–11.

Taub, E., Steiner, S. S., Weingarten, E., & Walton, K. G. (1994). Effectiveness of broad spectrum approaches to relapse prevention in severe alcoholism: A long-term, randomized, controlled trial of transcendental meditation, EMG biofeedback and electronic neurotherapy. *Alcoholism Treatment Quarterly, 11,* 187–220.

Tellegen, A., & Atkinson, G. (1974). Openness to absorbing and self-altering experiences (absorption), a trait related to hypnotic susceptibility. *Journal of Abnormal Psychology, 83,* 268–277.

Thomas, D., & Abbas, K. A. (1978). Comparison of transcendental meditation and progressive relaxation in reducing anxiety. *British Medical Journal, 2,* 1749.

Tucker, D. M., Vannatta, K., & Rothlind, J. (1990) Arousal and activation systems and primitive adaptive controls and cognitive priming. In N. L. Stein, B. Leventhal, & T. Trabasso (Eds.), *Psychological and biological approaches to emotion* (pp. 145–166). Hillsdale, NJ: Erlbaum.

Turpin, G. (1989). An overview of clinical psychophysiological techniques: Tools or theories. In G. Turpin (Ed.), *Handbook of clinical psychophysiology*. Chichester, UK: Wiley.

Turpin, G. (1990). Psychophysiology and behavioral assessment: Is there scope for theoretical frameworks? In P. Martin (Ed.), *Handbook of behavior therapy and psychological science: An integrative approach* (pp. 348–382). New York: Pergamon.

Turpin, G. (1991). The psychophysiological assessment of anxiety disorders: Three-systems measurement and beyond. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 366–375.

Turpin, G., & Lader, M. (1986). Life events and mental disorder: Biological theories of their mode of action. In H. Katschnig (Ed.), *Life events and psychiatric disorders: Controversial issues*. Cambridge, UK: Cambridge University Press.

Venables, P. H. (1984). Arousal: An examination of its status as a concept. In G. H. Coles,, J. R. Jennings, & J. A. Stern (Eds.), *Psychophysiological perspectives: Festschrift for Beatrice and John Lacey*. New York: Van Nostrand Reinhold.

Wadden, T. A., & Anderton, C. H. (1982). The clinical use of hypnosis. *Psychological Bulletin, 91,* 215–43.

Wagner, H. L. (1988). *Social psychophysiology and emotion theory and clinical applications*. Chichester, UK: Wiley.

Weiner, H. (1977). *Psychobiology and human disease*. New York: Elsevier.

Wells, A. (1990). Panic disorder in association with relaxation induced anxiety: An attentional training approach to treatment. *Behavior Therapy, 21*, 273–80.

Westling, B. E. & Ost, L. G. (1995). Cognitive bias in panic disorder patients and changes after cognitive-behavioral treatments. *Behavior Research and Therapy, 33*, 585–588.

Whitehead, C., & Adams, L. (1991). *Relaxation techniques*. Cambridge UK: Daniels Publishing Winslow Press.

Wiebe, D. J., & Williams, P. G. (1992). Hardiness and health: A social psychophysiological perspective on stress and adaptation. *Journal of Social and Clinical Psychology, 11,* 238–262.

Wiesse, C. S., Davidson, L. M., & Baum, A. (1989). Arousal theory and stress. In H. Wagner & A. Manstead (Eds.), *Handbook of social psychophysiology* (pp. 283–302). Chichester, UK: Wiley.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press.

Wolpe, J. (1982). *The practice of behavior therapy*. New York: Pergamon.

Wolpe, J. (1989). The derailment of behavior therapy: A tale of conceptual misdirection. *Journal of Behavior Therapy and Experimental Psychiatry, 20,* 3–15.

Woolfolk, R., Carr-Kaffeshan, L., & McNulty, T. F. (1976). Meditation training as a treatment for insomnia. *Behavior Therapy, 7,* 359–365.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.10
The Therapeutic Relationship

FRANK M. DATTILIO
University of Pennsylvania School of Medicine, Philadelphia, PA, USA
and
ARTHUR FREEMAN and JOHN BLUE
Philadelphia College of Osteopathic Medicine, PA, USA

6.10.1 INTRODUCTION 230

6.10.2 THE EXISTENTIAL PERSPECTIVE 231

6.10.3 THE PSYCHOANALYTIC PERSPECTIVE 232

6.10.4 THE HUMANISTIC PERSPECTIVE 236

6.10.5 THE BEHAVIORAL APPROACH 237

6.10.6 THE COGNITIVE-BEHAVIORAL PERSPECTIVE 237

6.10.7 IMPEDIMENTS TO THERAPY 238
- *6.10.7.1 Definitions of Resistance* 239
- *6.10.7.2 Client Factors* 240
- *6.10.7.3 Therapist Factors* 240
- *6.10.7.4 Problem/Pathology Factors* 240
- *6.10.7.5 Client Factors* 240
    - *6.10.7.5.1 Lack of client skill* 240
    - *6.10.7.5.2 Client cognitions regarding previous therapy failure* 240
    - *6.10.7.5.3 Client cognitions regarding consequences to others change* 241
    - *6.10.7.5.4 Secondary gain* 241
    - *6.10.7.5.5 Fear of changing* 241
    - *6.10.7.5.6 Lack of client motivation* 241
    - *6.10.7.5.7 Negative set* 241
    - *6.10.7.5.8 Lack of limited motivation* 241
    - *6.10.7.5.9 Limited or poor self-monitoring* 241
    - *6.10.7.5.10 Limited or poor monitoring of others* 241
    - *6.10.7.5.11 Narcissistic style* 242
    - *6.10.7.5.12 Client frustrated with lack of therapy progress* 242
    - *6.10.7.5.13 Client perception of lowered status in therapy* 242
- *6.10.7.6 Therapist Factors* 242
    - *6.10.7.6.1 Lack of therapist skill* 242
    - *6.10.7.6.2 Client and therapist distortions are congruent* 242
    - *6.10.7.6.3 Poor socialization to the treatment model* 242
    - *6.10.7.6.4 Lack of collaboration/alliance* 242
    - *6.10.7.6.5 Lack of data* 243
    - *6.10.7.6.6 Therapeutic narcissism* 243
    - *6.10.7.6.7 Poor timing of interventions* 243
    - *6.10.7.6.8 Lack of experience* 243
    - *6.10.7.6.9 Therapy goals are unstated, unrealistic, or vague* 243
    - *6.10.7.6.10 Lack of agreement with therapy goals* 243

6.10.7.7 Problem/Pathology Factors 244
  6.10.7.7.1 Client rigidity foils compliance 244
  6.10.7.7.2 Medical/physiological problems 244
  6.10.7.7.3 Difficulty in establishing trust and cooperation 244
  6.10.7.7.4 Press of autonomy 244
  6.10.7.7.5 Impulsivity 244
  6.10.7.7.6 Confusion 244
  6.10.7.7.7 Limited cognitive ability 244
  6.10.7.7.8 Symptom profusion 244
  6.10.7.7.9 Dependence 244
  6.10.7.7.10 Self-devaluation 244
  6.10.7.7.11 Limited energy 245
  6.10.7.7.12 Substance use 245
6.10.8 ABUSES OF THE THERAPEUTIC RELATIONSHIP 245
6.10.9 GENERAL GUIDELINES 246
6.10.10 APA CODE 246
6.10.11 REFERENCES 248

> The meeting of two personalities is like the contact of two chemical substances: if there is a reaction, both are transformed. (C.G. Jung, 1933)

6.10.1 INTRODUCTION

Few would argue that the therapeutic relationship between client and therapist is the cornerstone of any psychotherapeutic modality. Whether one practices classical psychoanalysis or radical behaviorism, the practice of psychotherapy involves a relationship, an interaction, and an emotional exchange between two or more individuals. Regardless of whether or not the goal is in sight or is directed behavioral change, one cannot ignore the essential ingredient of the bidirectional influence of one party to the other within the therapeutic collaboration. With each client, the therapist must start from the beginning and build a relationship from the ground up. While the process of relationship building can become well developed over years of clinical practice, the therapist learns that each and every client requires the same regime. That is, to identify the needs of the client, discern the pathways to best relate to the client, and then use those pathways to establish and maintain the therapeutic relationship.

The therapeutic relationship has been addressed in the professional literature in a number of different manners (Alexander & Luborsky, 1986; Horvath & Greenberg, 1986; Safran & Wallner, 1991; Suh, Strupp, & O'Malley, 1986; Sullivan, 1954; Wright & Davis, 1994).

The therapeutic relationship typically involves factors of confession, atonement and absolution, encouragement, positive and negative reinforcement, aversive procedures, behavioral modeling, promotion of values, and cheer leading (Frank, 1985). What makes the therapeutic relationship different from a long talk with a close friend can be summed up in Frank's definition of psychotherapy. He states:

> Psychotherapy is a planned, emotionally charged, confiding interaction between a trained, socially sanctioned healer and a sufferer. During this interaction the healer seeks to relieve the sufferer's distress and disability through symbolic communications, primarily words but also sometimes bodily activities. The healer may or may not involve the client's relatives and others in the healing rituals. Psychotherapy also often includes helping the client to accept and endure suffering as an inevitable aspect of life that can be used as an opportunity for personal growth. (Frank, 1985, p. 50)

The relationship is not the reason that therapy can have a helping effect for an individual, but rather the vehicle and context that makes the probability of relief more likely.

This chapter will address four issues of the therapeutic relationship. The first focus will be to define and compare the view of the therapeutic relationship from different theoretical perspectives, including the Psychodynamic, Existential, Humanistic, and Cognitive-behavioral domains. The second area of focus will be to relate some of the empirical data on the therapeutic relationship and highlight the correlation between relationship factors and therapeutic efficacy. The third focus will be to identify and describe issues and circumstances that negatively impact on the therapeutic relationship, broadly termed resistance. Finally, we will address abuses of the therapeutic relationship and the impact that these abuses have on the therapeutic process. A final summary will attempt to integrate the factors discussed as well as review some general guidelines for the therapeutic relationship and the American Psychological Association's code for ethical behavior.

Early psychoanalytic work identified three relational components of the therapy process that had to be addressed in the therapy. The meaning, value, and importance of these factors has been the subject of debate for more than 100 years. Nevertheless, we can use them as a broad framework upon which to stretch the fabric of our discussion of the relationship in psychotherapy. These issues were termed transference, countertransference, and resistance. Broadly, they capsule the client's expectations, beliefs, reactions, and responses to the therapist and to the therapeutic environment (transference), the therapist's reactions and responses to the client's transference or, alternatively, the therapist's reaction to the client generally (countertransference), and the client's difficulty in therapy or avoidance of issues and areas of concern (resistance). These factors may be overtly expressed or more covertly expressed either subtly or symbolically depending on the particular client and the situation.

6.10.2 THE EXISTENTIAL PERSPECTIVE

Within the context of Existential Therapy the therapeutic relationship is one characterized by mutual respect, individual uniqueness, authenticity, and pursuit of meaning. According to Seguin (1965), 'It is through total acceptance, that the client comes to value his or her own uniqueness, becomes free to exert choice, to make commitments, and to find meaning in life." According to Bugental and Sterling (1995), it is necessary to develop what they term a 'meaningful realistic therapeutic contract." This suggests an agreement or arrangement (contract) between the parties and that the individuals involved both contribute and hope to gain from the arrangement. The therapeutic contract facilitates understanding between the client and therapist, rendering it meaningful and realistic. Bugental and Sterling (1995) further identify the existence of other therapeutic functions characterizing the therapeutic course. Besides developing a meaningful, realistic therapeutic contract, they propose the importance of 'Fostering the growth of a resilient alliance, working through the situational resistances, working through some of the character resistances, exploring client–therapist collusion (transference–countertransference issues), working with residual transferential elements and preparing for termination" (p. 240). The therapeutic contract is viewed as providing a foundation for therapy with the therapeutic alliance being conceptualized as a container that will hold the struggles, emotions, and relationships necessary to a major life undertaking. According to May and Yalom (1989), existentially oriented therapists strive toward honest, mutually open relationships with their clients. The therapist and client address one another equally, generally both on a first-name basis, with the therapist striving toward demystification of the therapy process, answering questions fully and openly, as opposed to remaining impassive in an effort to evoke transferential distortions.

The canon of humanistic existentialism is that the human element must be the focus of the therapeutic endeavor. No theory of the person, of therapy, or of pathogenesis can take precedence over the personhood and human elements. Further, those subscribing to a "scientific materialistic modality" as thinking about client's productions in atomistic terms such as units of energy, drives, impulses, needs, instincts, stimulus response links, or in transformation-of-energy terms lose the essence of the human element. For example, Ofman (1985) suggests that proponents of psychoanalysis will be guided by theory rather than by an attention to the person. He explains that the client should be seen in that (theoretical) light; taught that way of thinking (by the analyst); and the client's predictions will be traded in on the "quotidian" view because it is difficult for behavioral scientists to proceed otherwise. In Ofman's view, "Humanistic existentialism translated that position to dealings with clients, but not in counter transference terms. In such a translation there is not subject–object split, but there is a mutual process" (p. 12).

Within the context of the therapeutic contract/alliance, the existential therapeutic task is the fostering of explicit awareness. The goals are for the person to utter, clarify, and identify projects, and their personal myth; to assume ownership and responsibility for one's projects; to validate and embrace both the inevitable positive and negative aspects of relationships and of acts, so that the divided, unattended parts of the personality become integrated. The attainment of such goals leads to greater autonomy, responsibility, and an attendant equality in relationships. Progress toward these goals is sought by means of acceptance, affirmation, and authentic relating between therapist and client (Ofman, 1985). The authenticity can occur only when the therapist sees that the client has freely chosen to do and to be what they are, and to make choices for the best of all possible reasons, personal growth.

May (1969) states, "Intentionality presupposes an intimate relationship with the world that we would not be able to go on existing in except if we could block the world out at times. This should not be called simply by the condemnatory term 'resistance.' I do not doubt the reality of resistance, as Freud and other's elucidated it, but I am emphasizing here a broader, structural phenomenon. That is 'every intention is an attention, and attention is I-can,' as Merleau-Ponty puts it. We are, therefore, unable to give attention to something until we are able in some way to experience an 'I-can' with regard to it"(p. 36).

In essence, the Existential view places a major importance on the development of the relationship as central to the therapy. The relationship is famed in terms such as "real" or "authentic," with the realness of the therapist serving as both a model and guide for the client to emulate.

6.10.3 THE PSYCHOANALYTIC PERSPECTIVE

Munroe (1955) capsules Karen Horney's position by conveying that she believed that man was born a potentially harmonious organism, capable of expanding his own capacities normally, in happy relationships within his surroundings. Conflict, Horney believed, occurred only as a consequence of social mishandling which instituted exaggerated impulses and fears in contradictory directions. Normally, a child is confident of help from his parents, because he has experienced it regularly. The neurotic child believes he must fight for everything he wants in a hostile world, because he has experienced deprivation. Yet open aggression in a creature as helpless as the human infant brings swift and overwhelming counteraggression. Thus, aggression and the fear of being aggressive both develop to an exaggerated degree.

One important area addressed by Horney in the therapeutic relationship was her new perspective on the psychology of women. Tasman, Kay, and Lieberman (1997) identify very clearly that it was actually Horney who "argued that Freud's model of feminine development was phallocentric, a masculine fantasy that protected men from their unconscious fear of women. Horney's arguments have been of central importance in feminist critiques of Freud and ultimately in the reworking of views of feminine psychology within the mainstream of psychoanalysis" (p. 16).

Horney's therapeutic and theoretical approach was holistic and placed an emphasis upon the notion of self-realization as well as phenomenological and existential influences noted in her earlier works. Such a position is characterized by Horney (1945) in the following excerpt: "Freud's pessimism as regards neuroses and their treatment arose from the depths of his disbelief in human goodness and human growth. Man, he postulated, is doomed to suffer or to destroy. The instincts which drive him can only be controlled, or at best 'sublimated.' My own belief is that man has the capacity as well as the desire to develop his potentialities and become a decent human being, and that these deteriorate if his relationship to others and hence to himself is, and continues to be, disturbed. I believe that man can change and go on changing as long as he lives" (p. 72). Rejecting Freud's instinctual approach Horney stressed a stronger contextual stance as influential on the individual's development.

Tasman et al. (1997) make several cogent points that are important in understanding the therapeutic relationship:

(i) "Horney believed that humans were not the tragic figures that Freud depicted, driven by biological endowment to seek sexual and aggressive satisfaction and inevitably meeting conflict."

(ii) "Normal development was not inherently conflictual. The neurotic child sacrificed the pursuit of pleasure and the fulfillment of his or her unique potential which Horney termed the real self, to ensure safety in an environment that the child perceived as hostile and dangerous."

(iii) "Ultimately, neuroses was the result of the environment; the neurotic child responded to the parents' failure to love her or him as she or he was; and neurotic symptoms were importantly determined by cultural norms."

(iv) "Her emphasis on the current context of analysis has been absorbed into mainstream psychoanalysis; her concepts of the real self and human potential have contributed to the broader expanse of dynamic psychotherapies" (p. 96).

In the matter of issues directly involving the role of the analyst in the context of the therapeutic relationship, Horney (1939) writes: "When I feel uncertain about a suggestion made to the client I point out its tentative character. If then my suggestion is not to the point, the fact that the client feels that I too am searching for a solution may elicit his active collaboration in correcting or qualifying my suggestion." In effect, Horney takes the active position that the analyst should exercise a more deliberate influence not only on the direction of the client's associations but also on those psychic forces which cause the neuroses but also on those forces that may help him eventually overcome the neuroses. She further states that the work a client has to accomplish is most strenuous and most painful. It implies no less than relinquishing or greatly modifying all the strivings for safety and satisfaction which have hitherto prevailed. It implies relinquishing illusions about himself which in his eyes have made him significant. It also advocates placing his entire relations to others and to himself on a different basis. What drives the client to do this hard work? Clients come for analytical help because of different motivations and with different expectations. Most frequently they want to get rid of manifest neurotic disturbances. Sometimes they wish to be better able to cope with certain situations. Sometimes they feel arrested in their development and wish to overcome a dead point. Very rarely do they come with the outright hope for more happiness. The strength and constructive value of these motivations vary in each client, but all of them can be actively used in effecting a cure. Tasman et al. (1997) state, "The analyst should focus on the here and now, interpreting current wishes and defensive structures so the client's unconflicted potential could emerge" (p. 36).

Concerning Horney's views on the therapist's and client's direction in therapy, the following excerpt best illuminates Horney's position, "The aim of therapy is then not to help the client to gain mastery over his instincts but to lessen his anxiety to such an extent that he can dispense with his 'neurotic trends.' Beyond this aim there looms an entirely new therapeutic goal, which is to restore the individual to himself, to help him regain his spontaneity and find his center of gravity in himself" (Horney, 1939).

In an examination of issues of transference, Tasman et al. (1997) explain, "Although childhood influences shaped character, current defensive needs rather than infantile wishes were the chief determinants of current behavior and of the client's transference." Horney (1937) states, "In relationships in which one person becomes dependent on the other there is invariably a great deal of resentment. The dependent person resents being enslaved; he resents having to comply, but continues to do so out of fear of losing the other. When the fear is very great they may seek to protect themselves against this dependence by not attaching themselves to anyone. These processes are evident also in a client's attitude during analysis" (p. 24).

In a more traditional sense, "Resistance often uses the mechanism of transference as its weapon to interfere with the progress of analysis. In this, the client unconsciously strives to avoid the insight which the cure demands. The transfer of unconscious infantile wishes to the person of the doctor facilitates their admission into consciousness" (Eidelberg, 1968). According to Ferenczi and Rank (1925), the transference is an immediate goal of analysis; the replacement of the manifest neurosis by a transference neurosis. According to Horney (1937), "Transference (in the traditional sense) should refer to the sum total of all the client's irrational reactions toward the analyst, not only the emotional dependence. The problem here is not so much why this dependence takes place in analysis, because persons in need of such protection will cling to any physician, social worker, friend, member of the family, but why it is particularly strong and why it occurs with such frequency. The answer is comparatively simple: analyzing means, among other things, tackling defenses built up against anxiety, and thereby stirring up the anxiety lurking behind the protecting walls. It is this increase of anxiety that causes the client to cling to the analyst in one way or another." This view is emphasized in the following comment by Horney (1937). "Sexual desires concerning the analyst are usually interpreted as repetitions of a sexual fixation on the father or the mother, but often they are not genuine sexual wishes at all, but a reaching out for some reassuring contact to allay anxiety. The client, to be sure, often relates associations or dreams—expressing, for example, a wish to lie at the mother's breast, or to return to the womb—which suggest a father or mother transference. We must not forget, however, that such an apparent transference may be only the form in which a present wish for affection or shelter is expressed" (p. 48).

In a discussion of issues of transference infatuation, Horney (1937) points out that the analyst becomes cognizant of the impersonal character of the infatuation and attributes this indiscrimination to the client's tendency to repeat old patterns. The client feels relieved because he/she recognizes that there is something compulsory, something not genuine in his/her feelings of love. As a result, the actual infatuation diminishes, dependent on the analyst's strengths. The client's view may hold that his security and satisfaction may depend upon fastening himself to others or, more accurately, merging with them. Consequently, giving and hopefully obtaining affection is for him, a means of reassurance and thereby safety.

It would seem inconceivable that a therapy could begin or proceed without some level of anxiety being raised for the client. Anxiety may reflect the need to discuss personal issues and problems. It may also involve the need, expectation, or requirement that the client relate to the therapist. Whenever anxiety emerges, the client's need to hang on to the analyst intensifies. Therefore, whenever the client displays a more than usual attachment or transference reaction to the analyst, the analyst should connect and reflect the attachment back to the client as an indication of anxiety or insecurity.

The effect of this interpretation would be to open the gate toward a greater recognition on the client's part, and eventually to lead to an understanding of the underlying structure responsible for his/her anxiety reaction. As it is mainly the client's anxiety, according to the Horneyan view, which makes him dependent on the analyst, interpretations of this kind counteract from the beginning the danger of dependency.

On matters of resistance, Horney (1939) contends that when the client is analyzed, he/she will realize the futility of their efforts, though with some reluctance. He may politely and intellectually follow the analyst's suggestions that the irritations are only "bubbles" surfacing. But as soon as the analyst identifies one of the deeper disturbances, the client will react with a mixture of concealed irritation and diffuse anxiety, and soon will argue most cleverly that the analyst is wrong, that at least he is exaggerating his reaction. Resistance for Horney (1939) involves the energy with which an individual protects repressed feelings or thoughts against their integration into conscious awareness. This concept is based on our knowledge that the client has good reasons for not acknowledging certain drives and will seek to avoid them.

Regarding notions of countertransference and transference, Horney (1939) believes that when a therapist reacts with inner irritation to a client's tendency to defeat his/her efforts, he/she may be identifying the client with his/her own father, and thus repeating an infantile situation in which they felt defeated by the father. If, however, the therapist's emotional reactions are understood in the light of his/her own character structure as it is affected by the client's actual behavior, then it will be viewed as the therapist's need to cure every case. Failure to do so may be viewed as a personal humiliation.

In summary, Horney's view states that neuroses are ultimately the expression of disturbances in human relationships. The analytical relationship is a very unique form of human interaction. The therapy then becomes a microcosm of the client's relational life. Relationships and existing disturbances are bound to appear in the therapy as they appear elsewhere. The particular controlled and contained conditions under which an analysis is conducted render it possible to study these relationship disturbances more accurately than elsewhere. Finally, the relationship problems in the analysis can serve to convince the client of their existence and of the role they play in forming and maintaining the problems.

At the bottom is the question which is essential to the goals of therapy. Do we intend to put out a finished product with all problems solved for good and all? In fact, do we believe it to be desirable? Or do we think of life as a process of development which does not end and should not end until the very last day of existence?

The work of Harry Stack Sullivan is, by definition, based on the interaction between people. His work termed "Interpersonal Psychiatry" is representative of the Neo-Freudian factions of his time that espoused a collection of ideas that could be found operative within the context of the therapeutic relationship. Sullivan held the belief that human beings possess an actualizing tendency, that is to say, an innate predisposition for fulfillment of one's human potential. According to Mullahy (1970), Sullivan believed that there was a built-in tendency of the self to maintain its current organization and functional activity. Furthermore, by our very nature we would rather be healthy than to suffer from the pains of mental illness. Sullivan clearly recognized the importance of creativity, purposefulness, and in stressing a holistic approach in the treatment of his clients.

May and Yalom (1989) state that in the interpersonal model of personality an individual is not instinctually guided and preprogrammed, but is instead almost entirely shaped by his cultural and interpersonal environment. Hall and Lindzey (1970) regard personality as a hypothetical entity which is unable to be isolated from interpersonal situations and interpersonal behavior. It is all that can be observed as personality. In a sense, it is vacuous according to Sullivan to speak of the individual as the object of study because the individual does not and cannot exist apart from his relations with other people. Although Sullivan does not deny the importance of heredity and maturation in forming and shaping the organism, he contended that, that which is distinctly human is the product of social interactions.

Some believe that Sullivan's interpersonal psychiatry emphasized psychopathology as a personal method for coping with terrifying relationships. He believed that the therapist as a participant-observer offered a new mutative quality of relatedness to the therapeutic relationship (Greenberg & Cheselka, 1995). Sullivan believed in experience and cognitive processes as represented by three modalities: the prototaxic, parataxic, and syntaxic modes. According to Kaplan, Sadock, and Grebb (1994), the prototaxic mode is undifferentiated thought that is unable to separate the whole into the parts or to use symbols. This is something Kaplan, Sadock, and Grebb believe occurs normally in infancy and is also seen in schizophrenia.

Within the context of the therapeutic relationship we find that the process is a collaborative and active one. According to Kaplan et al. (1994), "The therapy process requires the active participation of the therapist, who is known as a 'participant-observer' " (p. 236). Modes of experience, particularly the parataxic, need to be clarified and new patterns of behavior need to be implemented. In his critical analysis of Sullivan's works, Mullahy (1970) states that the therapeutic relationship is a performance of two people, a transaction, in which the client's behavior and what he says and does are adjusted, in accordance with the best of his information and ability, to what he guesses or surmises about the therapist. In addition, the rationale of the interview is based on a progressively unfolding expert client (therapist–client) relationship for the purpose of discovering the client's characteristic patterns of living. The therapist seeks to discover those (frequently obscure) patterns of living of the client that he experiences as particularly troublesome as well as those he experiences as particularly valuable.

Within the scope of the therapeutic techniques, interrogation and interpretation warrant brief definition at this point. It was Sullivan's position to avoid direct interrogation. Sullivan viewed a direct question in a psychiatric interview as not likely to provoke the most informative answer. Rather, Sullivan would have the client test out hypotheses that would either support or disconfirm his/her existing thoughts on a subject which is more reflective of an indirect approach. Direct interrogative questioning was only to be used if there was a very justifiable reason and purpose. Mullahy (1970) writes, "The interrogation proceeds from a given point in a direction that is easy for a client to follow" (p. 16). Regarding the use of interpretations, they must be timely and accurate as well as and constructive. Further interpretation as a therapeutic strategy if used inappropriately might serve to fuel anxiety. According to Mullahy (1970), one must usually avoid questions and interpretations that arouse anxiety since, among other things, it provokes aspects that interfere with growth of the client's information and insight. Clearly if an untimely inappropriate interpretation is made and results in an unnecessary escalation of the client's anxiety, the therapeutic relationship could fall into serious jeopardy. Sullivan was aware that at times anxiety could serve a constructive function when it helped to mobilize a client. In general, however, he viewed it as a "formidable enemy" (Mullahy, 1970).

In reading Sullivan's works, one can see that his methodology represents an honest attempt at scientific inquiry. It is clear that his psychiatric interview is an attempt at the collection of relevant data. The results or accomplishments of the interview are to ensue partly from interpretation of clearly documented facts, from the building of inferential bridges that carry one from particular concrete instances to a generalized formulation, and partly from considering alternative hypotheses regarding situations in which the client has preferred misleading formulations. It was Sullivan's contentions that psychiatrists, particularly some psychoanalysts, were prone to interpretation of the material expressed by their clients and that the abundance of interpretations was greatly in excess of the need for them.

The therapist, within the context of the therapeutic relationship, seeks to facilitate identification and awareness of parataxic distortions and the development of insight into that process.

The notion of resistances were renamed "security operations" by Sullivan (1953). He felt that the discovery of a parataxic distortion brought about a sharp fall in one's level of security, causing one to become intensely anxious. This was viewed as injurious to the relationship, and ultimately to the therapy because, "the client's self-esteem, his interpersonal security, is often fragile or at least infirm, then, the disturbing consciousness of the 'mistake' tends to arouse essentially disintegrative, distance-producing security operations" (p. 34). The therapist must first be aware of the tendency for this to occur, be skilled in dealing with it, and to then intervene and counter the disintegrative operation.

For Sullivan, counter transference (in the classical Freudian sense) was the result of inadequate skill or some personal inadequacy (Mullahy, 1970). Sullivan did not believe a therapist should become seriously ego-involved with a client, and would, in fact, consider any personal involvement with his client therapeutically dangerous and part of what he termed "social hokum." For Sullivan, the key to the effectiveness of therapy was the interpersonal relationship. He saw the therapist as an expert in the problems that "ail" the client but did not recommend that the therapist become intimate with any of them.

Sullivan expanded upon and reformulated the notion of transference in terms of parataxic "me–you" patterns. He interpreted transference as a one-way process that showed the direct interaction of personalities. "By transference the client manifests interpersonal processes that open the gates of memory sealed by dissociations, reorients his experience, and facilitates the development of arrested or distorted systems of motives so that he moves forward toward the conditions of adult personality organization" (Sullivan, 1962, p. 87).

In the initial stages of the therapeutic relationship the therapist has the responsibility when the client presents a problem, especially in the early stages of therapy, to not arouse so much anxiety that the client feels threatened. This would occur before a client learned how to withstand a good deal of anxiety when experiencing any perceived attack on their self-esteem. In the initial stages of the therapeutic relationship, the therapist must ensure within the structure of the therapeutic situations that both the therapist and client can arrive at an agreement on the overall goal, end, or inclusive "purpose" of the therapy, so that there is collaboration.

Mullahy (1970) identifies that "Sullivan held that in the course of identifying all the more parataxically 'surviving,' unresolved situations of the client's past and their consequent dissolutions, there progressively occurs an expanding of the self to the extent that the client as known to himself is much of the same person as the client behaving with others. This was, for Sullivan, the therapeutic cure. Ideally, the client would achieve social cure (though not always the outcome) that might bring about a more abundant life in the community."

6.10.4 THE HUMANISTIC PERSPECTIVE

A discussion of the therapeutic relationship would not be complete without addressing Carl Rogers and his person-centered approach. Rogers was a major spokesman for the humanistic movement in psychology and led a personal life that reflected the ideas that he developed and used in his psychotherapeutic approach.

He displayed a questioning stance and a deep openness to change which spiraled person-centered therapy to a height of becoming a major force in psychology.

The goals of person-centered therapy differ from those of traditional approaches. The person-centered approach aims toward a greater degree of independence and integration of the individual in treatment. Its focus is on the person, not on the person's presenting problem. Rogers (1977) contends that his aim of therapy is not merely to solve problems, but to assist clients in their growth process so that they can better cope with problems they are currently facing as well as deal with future issues.

The specific role of the person-centered therapist is rooted in their ways of being and attitudes, not techniques designed to get the client to do something to change. Research on person-centered therapy indicates that the attitudes of therapists, rather than their knowledge, theories, or techniques, facilitate personality change in a client. Basically, the therapist uses himself/herself within the relationship as an instrument of change. Hence, encountering the client on a person-to-person level, their role as a therapist is actually to be "without a role." The main function then is to establish a therapeutic climate that helps the client grow in an individual way.

The therapist's role becomes essential in the person-centered modality in that it creates a helping relationship in which clients experience the necessary freedom to explore areas of their life that are now either denied awareness or distorted. A genuine aspect of the therapist's role involves the willingness to be real in the relationship wih a client instead of viewing him/her in a preconceived diagnostic fashion. In essence, the relationship then meets the client on a moment-to-moment experiential basis in an attempt to help them divulge their world. It is through the therapist's relationship of genuine caring, respect, acceptance, and understanding that they are able to help the client loosen their defense mechanisms and rigid perceptions in order to evolve into a level of personal functioning.

In his early work, Rogers (1961) wrote "If I can provide a certain type of relationship, the other person will discover within himself or herself the capacity to use that relationship for growth and change, and personal development will occur" (p. 33). It was Rogers' contention that the significant positive personality change of an individual occurs only as a result of the relationship between therapist and client.

The characteristics of the therapeutic relationship that are conducive to creating a suitable psychological climate involve creating an atmosphere in which the client will experience the necessity to initiate personalty change.

Rogers believed that the following six conditions were necessary and sufficient for personality changes to occur:

(i) Two persons are in psychological contact.

(ii) The first, whom we shall term the client, is experiencing incongruency.

(iii) The second person, whom we shall term the therapist, is congruent or integrated in the relationship

(iv) The therapist experiences unconditional positive regard or real caring and acceptance for the client.

(v) The therapist experiences an empathic understanding of the client's internal frame of reference and endeavors to communicate this experience to the client.

(vi) The communication to the client or the therapist's empathic understanding and unconditional positive regard is to a minimal degree achieved (Rogers, 1987, pp. 39–41).

Rogers hypothesized that no other conditions were necessary to facilitate change in the therapeutic relationship. He believed that if the aforementioned six conditions occurred over a given period of time, constructive personality change would occur.

In addition, three personal characteristics, or attitudes, of the therapist form the central part of the therapeutic relationship. These include congruence, or genuineness, unconditional positive regard and acceptance, and accurate empathic understanding.

One final important aspect of Rogers' work was that he believed that the therapist and the person of the therapist were the same individual. He believed that therapists should live their lives in accordance with theory which is used in the course of psychotherapy. Rogers believed this to be paramount in the therapeutic relationship.

6.10.5 THE BEHAVIORAL APPROACH

Clinical and research evidence suggests that a therapeutic relationship, even in the context of a behavioral orientation, can contribute significantly to the process of behavior change (Granvold & Wodarski, 1994). A good therapeutic relationship increases the chances that the client will be receptive to therapy. Not only is it important for the client to cooperate with the therapeutic procedures, but the client's positive expectations about the effectiveness of therapy may often contribute to successful outcomes as well. The behavior therapist is one who can conceptualize problems behaviorally and make use of the client/therapist relationship in facilitating change.

As opposed to some of the other modalities of treatment, behavioral practitioners do not assign an all important role to relationship variables. Instead, they typically contend that factors such as warmth, empathy, authenticity, permissiveness, and acceptance are considered necessary, however, not sufficient for behavioral change to occur (Bandura, 1977). It is not a matter of the importance of the relationship *per se*, but rather the role of the relationship as a foundation on which therapeutic strategies are built to help clients change in the direction that they desire. Lazarus (1989) maintains that unless clients respect their therapist it will be difficult to develop the trust necessary for them to engage in significant self-disclosure. At the same time therapists need an array of clinical skills and techniques to employ once an effective client/therapist relationship has been established. Since behavior therapy demands such a high level of skills and sensitivity as well as the ability to form a working relationship with clients, the therapeutic relationship is important. Behavior therapists tend to be more active and directive and can function as consultants in helping the client solve problems as opposed to allowing the relationship in and of itself to facilitate the change. Since behavior therapists use a coping model in initiating behavioral change in the clients natural environment, it is important that they be personally supportive to them. However, once again the actual process of change comes from the implementation of techniques as opposed to a byproduct of the actual relationship. At the very least, behavior tharapists view the relationship as being an important supportive aspect, yet not necessarily essential to change.

6.10.6 THE COGNITIVE-BEHAVIORAL PERSPECTIVE

Despite criticism by some in the field, cognitive-behavior therapy places an important emphasis on the therapeutic relationship as an integral part of the effectiveness in the treatment process. Constructs from cognitive-behavioral therapy can be used to understand variability in client behavior, and cognitive-behavioral formulations can be used to guide individuals in dealing with difficult problems in the therapeutic relationship. What is more, vicissitudes in the therapeutic relationship can provide individuals with an opportunity to work directly with clients' most significant maladaptive schemas which are an essential part of the therapeutic relationship. Schemas are defined as cognitive structures that may be at the core of an individual's particular dysfunction. It is a maladaptive, cognitive–interpersonal cycle that clients' perceptions of the therapist's behavior provide a phenomonological link to the dysfunctional interpersonal schemas and associated patterns of behavior (Safran, 1990; Safran & Segal, 1990).

The notion of an individual's maladaptive beliefs or schema was first proported by Beck, Rush, Shaw, and Emery (1979). Schemas, which are cognitive distortions at the individual's deepest level, consist of concepts that the client habitually uses in viewing reality. These biases direct their focus in the retrieval of information. Typical negative schemas include rejection, abandonment, control, uniqueness, and unrelenting standards—prime aspects that are found in any therapeutic relationship.

Therefore, a client who may maintain a negative schema of rejection will tend to focus on any sign that he or she is being rejected by the therapist. The same goes for their concept of abandonment and consists of what may be defined by the specific bias of their content. These schemas are usually formed during early childhood and selectively focus on information processing.

In his cognitive model, Beck, Freeman and associates (1990) used personality as a combination of core beliefs and a characteristic set of basic strategies for interacting with others. Inter-related beliefs and action plans are organized within a generic knowledge structure referred to in the cognitive literature as schema. These schemas guide and actively participate in the processing of information and actions so that we see recurring patterns of thoughts, feelings, and behaviors.

When deeply held beliefs and well-ingrained behavioral patterns are repeated in the therapeutic relationship and interfere with the effective use of what cognitive therapists refer to as collaborative empiricism, this creates a threat in the therapeutic process. The clients' expectations and perceptions of the therapist are typically distorted to remain consistent with his/her underlying beliefs. Behavior also remains consistent with dysfunctional beliefs. Cognitive-behavior therapists state that "cognitive transference develops when the aforementioned patterns are repeated in the therapeutic process (Wright & Davis, 1994). Cognitive-behavior therapists typically use certain procedures to identify and modify maladaptive beliefs and dysfunctional behavior when cognitive transference occurs. When an individual has a serious personality disorder, this process may require longer-term therapy than is usually the case with a major depressive or anxiety disorder.

The examination of beliefs, related behavioral strategies, and patterns in the therapeutic relationship can help therapists recognize and manage each client's unique expectations within the therapeutic process. It is extremely important that this be done and addressed forthright since any schema level analysis that is ignored may contribute to the clients' resistance and noncompliance and thwart the progress in therapy. In addition, excessive dependency or stagnation in the treatment process may also occur.

It is therefore that cognitive-behavior therapists emphasize the intense supervision on relationship issues in training programs (Michenbaum & Turk, 1987). In fact, it has been suggested that this element of the therapeutic process should require as much attention as learning how to implement specific cognitive and behavioral techniques (Wright & Davis, 1994).

The degree of importance of the therapeutic relationship depends not only on the approach, but in accordance with the complexity or the problem itself. The more specific and simple the problem, the more the therapeutic relation shall be of less importance. Results from a series of investigations give evidence to the importance that the therapeutic relation has gained as an instrument of social change. Sloan (1975) found that consultants who solved their problems satisfactorily through behavioral therapy considered the relation with their therapist as the most important factor in their recovery. Alexander (1976) reports that the therapeutic relationship contributed significantly to the behavioral treatment of delinquents and their families. Mathews et al. (1976) found that a group of agoraphobics considered that the support, enthusiasm, and sympathy received from their therapist helped them to overcome their fears more than the actual practice of exposure exercises. Researchers Orlinsky and Howard carried out a revision in 1986 about research undertaken in order to study the predictive capacity for diverse factors in psychotherapy with regard to their efficiency. They were surprised by the great predictive capacity shown by patients' feelings about the empathy shown by the therapist. This prediction went far beyond what could be done based on the empathy displayed by the therapist. In general, these results suggest that the interpretation which the client has on the actions of the therapist during treatment is a significant factor in the effectiveness of the respective treatment. The therapeutic relationship may therefore be used as a process facilitator of social influence of different forms.

In summary, the therapeutic relationship is an essential, interactive component of the process of cognitive-behavior therapy. Cognitive-behavior therapists need to remain sensitive to both the general and idiosyncratic expectations of their clients, without compromising the necessary limits or boundaries of the relationship. It is attention to these principles that will aid in building productive therapeutic relationships.

6.10.7 IMPEDIMENTS TO THERAPY

It is tempting (and simplistic) to blame treatment difficulty or lack of progress in therapy on the client's noncompliance or "resistance." Therapists may assume that when progress lags it is because the client does not want to change or "get well," for either conscious or unconscious reasons. However, there are many different problems which can slow or block progress in therapy and few of them are due to the client's desire to retain his or her problems. To illustrate this point, one might consider a "self-help" or "self-improvement" task such as losing weight, exercising regularly, or completing unfinished paperwork which one has been slow to complete. Is the lack of follow-through due to a secret desire to remain overweight, out of shape, or behind in paperwork, and to what extent might it be due to a number of more mundane factors such as personal comfort, time limitations, and so on. Might it be viable to assume that sometimes it is a combination of the two?

6.10.7.1 Definitions of Resistance

Common themes of client resistance in therapy involves distrust of the therapist, personal shame, grievances against others, depreciation, or fear of rejection. Typically, resistance may be manifested directly (e.g., tardiness or missing of appointments) or more subtly through omissions in the material reported in the sessions. A number of publications have addressed this important issue (Ellis, 1985; Shelton & Levy, 1981; Stark, 1994; Wachtel, 1982). There are, however, many reasons for noncompliance other than the client not wanting to change or the lack of compliance indicating a pitched battle between intrapsychic structures.

Clinically, we can identify several reasons for noncompliance. They can appear in any combination or permutation, and the relative strength of any noncompliant action may change with the client's life circumstance, progress in therapy, relationship with the therapist, and so on. We can divide these impediments to therapy into four broad categories. The first are problems emanating from the client. The second category include those problems stemming directly from the therapist. The third category pertains to those attributes involving the type, severity, and nature of the client's problem(s) or diagnosis. The final factors are those related to the client's life situation, personal context, and significant others. For each problem discussed, we will describe therapeutic interventions for reducing or ameliorating these obstacles and impediments to therapeutic change.

Resistance or noncompliance may take many forms. Some would appear more directly negative, for example, verbal evasion, verbal or physical aggression, or threats against the therapist or agency. Other forms of resistance may be more subtle, for example, lateness for appointments, missing appointments, extended silence in sessions, forgetting homework, or failing to pay the therapist's bill. And yet, others may appear positive, but still have the effect of resistance, for example, trying to win the therapist over with praise, gifts, or devotion. Resistance may be conscious, for example, withholding information and details to appear healthier, smarter, and so on. On the other hand, the same material might not be reported because of "forgetting," a process that might be more unconscious.

A starting point for understanding resistance is to conceptualize it as a normal and adaptive response to any type of perceived threat. Whether the resistant behavior is viewed as an ego defense or as a response to unfamiliar demands, it is unreasonable to accept that an individual will move directly to change without some level of discomfort. In point of fact, an individual who moves too quickly to change may be seen as impulsive or labile. An extreme to this would be the client with a borderline personality disorder who moves quickly from overidealization to complete devaluation. Normally, there is initially some hesitation with this process. Such resistance needs to be identified and addressed directly within the therapy process. Resistance is a predictable and therefore expected and at times even welcome part of the therapy. It can become an arena for the practice of a range of therapeutic interventions and in some cases adds to the change process.

Many of our clients have been involved in therapy prior to submitting to their present course of treatment. If the previous therapeutic experience has not been successful, it might be essential for the therapist to first assess what occurred in the previous therapy in terms of the focus, direction, content, timing, cadence, and style of treatment. With this information in hand, the therapist can then work to alter any and all of the factors emblematic of the previous therapy so that the present therapy is experienced as different, and not more of the same old pattern. For example, if the previous therapy was unstructured, structure would offer a different therapeutic experience and possibly draw forth new or different material. If the content revolved around a particular issue shifting to a variant of that issue, or even a different issue may be in order. The use of different examples would be useful in this regard. The short-term treatment approach is often enough of a different focus to make the therapy a new experience and opportunity for change. If the previous therapy was unstructured, structure would be helpful. If the previous therapy was a short-term treatment approach, the therapist must be even more creative in finding new directions or foci.

Use of the broad umbrella term of resistance may do more to limit therapy than to encourage or support it. Several reasons for noncompliance with a given therapeutic regimen may be isolated. It would be impossible to totally isolate each one as a separate and completely distinct entity from all others. There will, of necessity, be similarity and overlap. To make it easier to identify and deal with each, we have divided them into three categories, client factors, therapist factors, and problem or pathology factors.

6.10.7.2 Client Factors

(i) Lack of client skill
(ii) Client cognitions regarding previous therapy failure
(iii) Client cognitions regarding consequences to others of change
(iv) Secondary gain
(v) Fear of changing
(vi) Lack of client motivation
(vii) Negative set
(viii) Limited or poor self-monitoring
(ix) Limited or poor monitoring of others
(x) Narcissistic style
(xi) Client frustrated with lack of therapy progress
(xii) Patent perception of lowered status in therapy.

6.10.7.3 Therapist Factors

(i) Lack of therapist skill
(ii) Client and therapist distortions are congruent
(iii) Poor socialization to the model
(iv) Lack of collaboration/alliance
(v) Lack of data
(vi) Therapeutic narcissism
(vii) Poor timing of interventions
(viii) Lack of experience
(ix) Therapy goals are unstated, unrealistic, or vague
(x) Lack of agreement with therapy goals.

6.10.7.4 Problem/Pathology Factors

(i) Client rigidity foils compliance
(ii) Medical/physiological problems
(iii) Difficulty in establishing trust
(iv) Autonomy press
(v) Impulsivity
(vi) Confusion
(vii) Limited cognitive ability
(viii) Symptom profusion
(ix) Dependence
(x) Self-devaluation
(xi) Limited energy
(xii) Dissociation.

6.10.7.5 Client Factors

6.10.7.5.1 Lack of client skill

Therapists cannot make the assumption that every client has developed the skills to effectively perform a particular behavior or sequence of behaviors. For many clients their difficulty in therapy will parallel their inability to cope with life stressors. Both may be based on inadequately developed skills. These skills may be broad, that is, social skills, or more discrete, that is, making eye contact while speaking to another person. For many individuals, their skills may be adequate for "getting by" in familiar and highly structured areas of life experience, when their skills are tested in novel situations they have far more difficulty and may withdraw, or fail. If however, they are overtly successful at coping, they experience such a high level of discomfort that they will avoid future encounters. Given that the client may never have developed skills, or not developed them to the level necessary for adequate functioning, the therapist may need to teach particular skills to help the client move along in therapy and thereby in life.

6.10.7.5.2 Client cognitions regarding previous therapy failure

When the client has cognitions of failing to be able to successfully make changes in thought or behavior, the therapist needs to help the client to carefully examine their cognitions. Examining the cognitions, the underlying assumptions/schema and learning to respond in an adaptive manner to these negative and self-deprecatory thoughts, is a major goal of the therapy work. One aspect of "failure" which inhibits many clients is their anticipation regarding the therapist's reaction if homework assignments are not done "right." If the client anticipates receiving harsh criticism, anger, expressions of disappointment, or other aversive responses from the therapist when the homework is discussed, this can easily result in his or her avoiding the homework and coming up with excuses for not having done it. Obviously, it is important for the therapist to respond to noncompliance without being punitive or authoritarian and instead to work with the client to understand what blocked compliance. However, it is also important for the therapist to be alert for negative anticipation based on the client's previous experience with parents and teachers and to address these explicitly if they impede therapy. In particular, perfectionistic clients often anticipate extreme reactions if the homework is not done perfectly, and it can be quite useful to address these anticipations early in therapy.

Usually it is possible to honestly present the client's task as a "no-lose" situation by pointing out to the client that incidents of noncompliance or unexpected results provide opportunities for making valuable discoveries. For example, the therapist might follow the first homework assignment with, "One of the nice things about this sort of approach is that whatever happens, we come out ahead. If you go ahead and do [the assignment] and it goes the way we expect, great! We're making progress towards your goals. If you unexpectedly cannot get yourself to do it or if it does not work out the way we expect, then we have an opportunity to look at what happened and at your thoughts and feelings to discover more about what blocks you from your goals. If it goes smoothly we're making progress and if it doesn't, we're making a discovery." For many clients this greatly reduces the fear of failure.

6.10.7.5.3 Client cognitions regarding consequences to others change

Another set of cognitions involves the client having catastrophic ideas relative to the result of their attempting to change on others. The client often catastrophizes the result or consequences of their changing and needs to not only decatastrophize the potential, but to examine whether there are still advantages to changing.

6.10.7.5.4 Secondary gain

There may be situations where the client may not change because of the gain that accrues from continuing their dysfunctional suicidal thinking and/or behavior. In the case of suicidal behavior or ideation, this may force family members to treat the client with "kid gloves," not put any pressure on the client, avoid confrontation, and generally allow the client to do whatever they wished, rather than increase the suicidal potential. This gain may be obtained from family, friends, employers, or other individuals with whom the client has interaction. This client needs to look at the "primary loss" that goes into achieving their secondary gain. The client needs to be helped to achieve their gain in other ways.

6.10.7.5.5 Fear of changing

For some clients, changing means relinquishing ideas, beliefs, or behaviors that they see as inimical to their survival. While this may appear paradoxical in that their thinking makes them suicidal, these clients fear change as an unknown. They often choose the familiarity of their pain to the uncertainty of a new mode of thinking or behaving.

6.10.7.5.6 Lack of client motivation

Clients may arrive for therapy under protest. Therapy may be mandated as part of a legal penalty, that is, "go to therapy or go to jail." For other clients, therapy is coerced by family members, that is, "If you don't get help, I'm leaving you." For child and adolescent clients, the referral will come from the school or parents, that is, "You are in trouble and must change what you are doing, now."

In these circumstances, clients come to therapy with the message that the therapist is an agent of the referring (or coercive) individual, whether that individual is parent, spouse, judge, probation officer, guidance counselor, or teacher. This may facilitate an adversarial situation prior to the first therapy session, setting the stage for possible failure.

6.10.7.5.7 Negative set

Often a client is seen to have a "bad attitude," or "a very negative view." What is labeled as negative or attitudinal is often an issue of negative set. The negative set might be manifested directly as "Yes-but" behavior quickly disqualifying whatever the therapist says, or as directly arguing with the therapist on issues both large and small.

6.10.7.5.8 Lack of limited motivation

The reason for referral, the reasons offered for the client to change, the client's comfort or limited discomfort with behaving and feeling the way that they do, the level of dysfunction, and the press of demands of significant others all may work to reduce motivation. Often, the client may see small parts of the problems, but perceives them as "part of them," or as too small to bother with in therapy. The demands of therapy, cost, time, money and effort all further contribute to the limiting of motivation.

6.10.7.5.9 Limited or poor self-monitoring

Individuals may see the flaws and foibles of others, but remain blind to their own. Difficulty or inability to self-monitor will often be a major stumbling block to therapy. Being either unaware or unable to self-monitor and to then self-evaluate will lead to depressive affect and behavior. The client will often not self-monitor but will self-devaluate by developing and maintaining negative ideas about themselves.

6.10.7.5.10 Limited or poor monitoring of others

For some clients, the monitoring of others is a problem. They tend to look at others, but see very little. Their response to others would more likely be based on the client's images and distortions of others rather than a databased assessment. They will often see others as brighter, more attractive, more skilled, and so on. They will view all other relationships as more rewarding than any they have or might have, and do not seem to open their eyes to collect the needed data regarding others.

6.10.7.5.11 Narcissistic style

A narcissistic style needs to be differentiated from a diagnosis of clinical narcissism. The narcissistic style causes the client to be so self-involved that any attempt to have them look at others or at themselves is met with resistance. The reaction is typically framed as, "It can't be me," or "Why would you (or anyone) expect so much from me."

6.10.7.5.12 Client frustrated with lack of therapy progress

Clients may have unrealistic expectations of therapy and possible therapy progress. When the expectations are not met, the client may blame self or therapist for the lack of therapy progress and withdraw or withhold from the therapeutic collaboration.

6.10.7.5.13 Client perception of lowered status in therapy

For some individuals, being a client is a mark of lowered status. For many, being in therapy is the mark of being "sick," "disturbed," "weird," or "crazy." Given that belief, it would follow that leaving or avoiding therapy makes one less of all of the above.

6.10.7.6 Therapist Factors

6.10.7.6.1 Lack of therapist skill

Just as clients come into therapy with a particular set of skills, so too do therapists. Because of limited experience with a particular client problem or population, the therapist may not be best equipped to work with a particular client. The therapist working within the context of an agency or hospital setting may be able to call in colleagues for consultation on the case or to seek supervision on the particular case/problem. It would be incumbent upon therapists to constantly develop, enhance, and upgrade their skills through additional training. Post-graduate courses, continuing education programs, seminars, workshops, or institutes would be part of the professional growth of the therapist, no matter what their training. If the therapist's skills are poorly developed to effectively cope with a problem, transfer to another therapist is the ethical requirement.

6.10.7.6.2 Client and therapist distortions are congruent

This therapist blind-spot may be very destructive to the therapeutic process in that it would generally incline the therapist to accept the client's dysfunctional beliefs. If client and therapist share a particular dysfunctional idea, for example, "everything is hopeless and cannot change," it will bode poorly for the therapy. This sharing of an idea or belief can result in the therapist "buying into" the client's hopeless ideas and beliefs, not testing these beliefs, or even encouraging them.

6.10.7.6.3 Poor socialization to the treatment model

The client who does not understand what is expected of them will have difficulty complying with the therapeutic regimen. It is essential that the therapist assess the level of understanding of the model throughout the therapy work, especially with the suicidal client. Often their ability to listen and understand may be impaired by their hopelessness. The therapist cannot assume that having read books about therapy guarantees adequate socialization to therapy. Further, there may be proactive interference because of previous therapy. Clients who have been in therapy have, ideally, been socialized to that previous therapy model. They will continue to use the same strategies and approach to therapy and to life in general unless and until they are taught differently.

6.10.7.6.4 Lack of collaboration/alliance

Collaboration is an essential ingredient for all psychotherapy. This is crucial in working with the suicidal client. If the client and therapist do not have a good working alliance, it would seem to follow that the client may be less motivated to work with the therapist, do homework, follow the therapist's direction, or generally work towards making changes. The lack of collaboration, if not based on socialization difficulty or the skill of the therapist, may be due to the client's cognitions relative to cooperation or collaboration. Certain clients may actively work to thwart the therapist. This type of passive–aggressive behavior may be motivated by any of a variety of client cognitions, that is, issues of control, fear, competition, or displaced anger may all serve to cause difficulty in the therapy.

This client may be directly challenging or more covertly avoidant as in the classic "yes, BUT" response.

For example, a therapist in training asked for advice in handling a client's noncompliance with behavioral experiments designed to reduce the client's perfectionism. The client's goals for therapy were to resolve some relatively minor marital problems, but the therapist saw the client's perfectionism and the stress and job dissatisfaction which resulted from it as more significant problems. Rather than discussing this issue with the client and reaching an agreement on the goals of therapy, the therapist had unilaterally begun working on perfectionism and this led to the noncompliance.

Collaboration involves both the therapist and the client and either of them can disrupt it. If the client feels that he or she has no voice in how therapy proceeds, either because this is indeed the case or because of his or her beliefs and expectations, this is likely to interfere with collaboration and produce problems with compliance. It is important for the therapist to actively solicit and value the client's input in setting agendas, determining the focus of therapy, and developing homework assignments, particularly with clients who tend to be unassertive. It is also important to be alert for any cognitions on the part of the client which could block collaboration.

6.10.7.6.5 Lack of data

The basis for treatment is the assessment and general collection of data. The therapeutic conceptualization and the resultant treatment plan is then databased. If the therapy is focused on theory without data, the therapy will suffer. The therapist may then make major conceptual leaps without a solid footing upon which to base the therapy.

6.10.7.6.6 Therapeutic narcissism

An issue that can be a major impediment to therapy is what we term "therapeutic narcissism." This results from the therapist being so taken with themselves that they are blinded by the need for greater humanity and empathy. The therapeutic narcissism may take the form of telling rather than asking the client how they feel. It may take the form of deciding what the client needs without consulting the client.

6.10.7.6.7 Poor timing of interventions

Interventions that are untimely can have the effect of the client not seeing the importance or relevance of the therapeutic work, and thereby appearing to be noncompliant. If the therapist, because of his or her anxiety, tries to push or rush the client, the result may be the loss of collaboration, the missing of sessions, a misunderstanding of the therapeutic issues, or a premature termination of therapy, a possibly fatal issue with a hopeless client.

The timing and pacing of interventions can be quite important. If the therapist tries to push or rush the client, the result may be the loss of collaboration, poor compliance, poor attendance, or premature termination of therapy.

6.10.7.6.8 Lack of experience

Inexperience is something that all therapists face at the advent of their careers. This impediment is unintentional and a standard part of the mental health training system. Frontline therapists working with the most disturbed and problematic clients may be therapists in practicum, internship, or residency settings.

6.10.7.6.9 Therapy goals are unstated, unrealistic, or vague

When the goals of therapy are unstated, unrealistic, or vague, the client may be in the position of unknowingly resisting the treatment. This issue also raises problems with informed consent. The client must be part of the treatment planning process and informed as to the goals, strategies, and interventions of the therapy so that they can best comply rather than being noncompliant out of ignorance.

6.10.7.6.10 Lack of agreement with therapy goals

Obviously, the client who does not understand and agree to what is expected of him or her will have difficulty complying with the therapeutic regimen. However, it is easy for therapists to overlook the possibility that their instructions and explanations may not be understood and accepted by the client. It is important for the therapist to repeatedly solicit feedback from the client and to encourage the client to raise any concerns and objections, so that therapist and client can develop a shared understanding of the client's problems which forms a basis for collaboration and so that it is clear that the client understands and accepts the homework assignments. Generally this proves to be sufficient, but when the client holds strong preconceptions about therapy, the therapist may need to compromise to some extent in order to facilitate collaboration.

6.10.7.7 Problem/Pathology Factors

6.10.7.7.1 Client rigidity foils compliance

With some clients, their personality rigidity foils their ability to actively comply with therapy. This is particularly true with clients who are obsessive-compulsive, paranoid, among others, in which their disorder may preclude their compliance. They may question the therapist's motives or goals. They may be unable to break out of the rigid position that they see themselves as having to maintain.

6.10.7.7.2 Medical/physiological problems

It is essential for every client coming for therapy to have a complete medical evaluation, with blood work as part of a comprehensive assessment and treatment plan. It is unethical and dangerous for the therapist to be treating what may appear to be psychological disorders but have a medical etiology, for example, a client with hypothyroidism may appear depressed because of the slowed action and thinking. Conversely, hyperthyroidism might be confused with anxiety disorders.

6.10.7.7.3 Difficulty in establishing trust and cooperation

Trust is a central issue in therapy. The trust must be bidirectional where the client trusts the therapist and the therapist can trust the client. For clients where problems of trust are a diagnostic part of the disorder will have problems in therapy.

6.10.7.7.4 Press of autonomy

The autonomous individual will be reluctant to come for therapy. Their view is that if they cannot help themselves, how can anyone else help them. The idea of coming to someone's office at a time set by someone else's schedule, to talk about themselves is, at best, uncomfortable. Their avoidance of therapy is seen as one way of maintaining their autonomy.

6.10.7.7.5 Impulsivity

Clients who are impulsive, and this includes most children and adolescents, see therapy as restrictive and limiting. At best it is out of line with how they generally respond, at worst the therapy is seen as an onerous and problematic requirement that they cannot meet. Their standard and accustomed manner of response is to act without thinking and self-monitoring rather than to self-monitor and to think about actions.

6.10.7.7.6 Confusion

Clients who are confused because of schizophrenia, bipolar illness, or neurological injury or deficit will have difficulty making use of therapy. They may have memory problems, difficulty in follow-through, difficulty with homework, and problems dealing with any abstractions.

6.10.7.7.7 Limited cognitive ability

Clients may have limited cognitive ability that is a result of limited intellectual ability or neurological deficit. Their processing will be limited by the lowered level of cognitive integration.

6.10.7.7.8 Symptom profusion

Anxious clients will often overwhelm the therapist with graphic, elaborate, and detailed descriptions of their symptoms. Their idea is that if anything is left out they run the risk that the omitted piece will be the essential piece that makes it impossible for them to be helped. If part of their symptom picture is gastrointestinal distress, they will regale the therapist with images of their distress. The therapist will quickly learn more about the clients gastrointestinal tract than is necessary. They will often avoid therapy by speaking quickly to get it all in.

6.10.7.7.9 Dependence

The client who is dependent will often work to insure that the therapist is totally and completely on their side. They may overwhelm the therapist with data, bring the therapist gifts, or praise the therapist for the wit, insight, sensitivity, and perspicacity. The goal is closeness. Without the closeness, they believe that they will be injured or even destroyed. They are often frightened by short term therapy in that it means that in a relatively short time they will be without their helper.

6.10.7.7.10 Self-devaluation

Often termed low self-esteem or poor self-image, this involves devaluing everything that one does, or the concomitant overvaluing of what everyone else does (and therefore by comparison devaluing self). This often leads to "yes-but" behavior and to devaluing both the therapy and the therapist. This devaluation can be summed up by paraphrasing Groucho Marx's comment, "I would never join any club that would have me as a member." The therapist and the therapy are, by extension of working with the client, tainted.

6.10.7.7.11 Limited energy

Depression is a major contributor to this impediment. Individuals who are depressed will often have vegetative signs that include lowered energy. It then becomes difficult to cooperate in therapy given that the major goal is to avoid any activity that requires action or energy. Levels of energy both within the session, for homework, or for interpersonal relationships is minimal.

6.10.7.7.12 Substance use

If substance abusing clients come to therapy drunk or stoned, they cannot make use of the therapy. What may occur is state-dependent learning whereby they can only act in certain ways when under the influence of the substance. Clients who are semiaware or semirelated to the therapist will make little use of the therapy. When sober they either will not remember what was said or done or will have distorted it due to the filter of the drugs.

6.10.8 ABUSES OF THE THERAPEUTIC RELATIONSHIP

Smith (1988) writes, "The practice of psychotherapy involves much more than the mere application of psychological theory and techniques" (p. 59). The therapist–client relationship is a unique establishment of a working partnership with respect, trust, and confidentiality as its fundamental principles. This interactional process is the mechanism by which the effects of treatment are created and realized (Smith, 1988). As with most interpersonal relationships, the psychotherapeutic process has a frame or structure that delineates and identifies the purpose and meaning of the relationship. This frame consists of socially dictated components which serve as guidelines to how the therapeutic process should occur (Epstein, 1994).

Trust is an element of psychotherapy which is of primary importance. Clients expect that within reason, their needs are the focal concern of the therapeutic relationship, and that the therapist's responses and interactions have the intention of promoting effective treatment gains (Galletly, 1993). The ethical dictum is that the therapist must act so as to avoid injury or harm to the client.

The client has been encouraged to suspend their usual defenses, reveal deeply personal thoughts and feelings, and ultimately become somewhat dependent upon the clinician. The client enters treatment with the expectation of being helped. In addition, the client experiences the development of intense feelings toward the therapist as a manifestation of the transference. In this way, the therapist is in the position of power. The client becomes dependent, and depending on the presenting problems, possibly needy. The client is vulnerable to abuse. This abuse may take many forms from the more subtle (i.e., the therapist starting therapy sessions later than scheduled) through more moderate abuse (canceling therapy sessions) to the severe abuse of physical or sexual aggression. Sexual contact between a client and therapist will invariably cause situations that manifest themselves as the client experiencing greater difficulties with trust, self-esteem, and problems expressing anger (Simon, 1989). The clients usually feel exploited, used, and confused (Blackshaw & Patterson, 1992). Despite the documented harmful effects of sexual involvement with clients, 7–12% of therapists admit to having had sexual contact with a client (Gartrell, Herman, Olarte, Feldstein, & Localio, 1986). The gender breakdown of the therapists was approximately 10% of male therapists and 2–3% of female therapists having engaged in sexual activity with a client. Pope, Sonne, and Holroyd (1993) report that although the sexual exploitation of a client is the most explicitly defined ethical principle, it continues to be a problem within the mental health discipline.

Other areas of potential problems occur in regard to fees and financial arrangements. The Ethics Code for Psychology states that the psychologist as early as possible define and make an agreement as to how the fee schedule is arranged. Fees are consistent with laws, and are not misrepresented. Psychologists are permitted to use collection agencies (Canter, Bennett, Jones, & Nagy, 1994).

With regard to therapy itself, psychologists are required to advise the client as early as possible the structure and outline of therapy, confidentiality and its limits, and the proposed treatment plan. The psychologist is to provide coverage in the event of their absence, and terminate the professional relationship in a collaborative stance when it becomes evident that the client no longer requires the service or is not benefiting from continued treatment (Canter et al., 1994).

In conclusion, therapists are bound by clinical, ethical, and legal standards to maintain the integrity of the psychotherapeutic process. Epstein (1994) writes, "Ideally, the therapist will be able to fine-tune the frame into an empathic dynamic structure that is sensitive to the client's changing needs" (p. 17). The therapist is required to be not only competent in their chosen profession, but respectful, conscious of, and adhere to ethical standards espoused by their colleagues, licensing boards, legal mandates, and society as a whole in order to protect those vulnerable to potential harm. Gorton and Samuel (1996) propose that comprehensive training and mandatory education on the topic of sexual and ethical issues be included as a national requirement for graduate program accreditation.

The therapeutic work must take place within a working partnership with respect, trust, and confidentiality as its fundamental principles. This interactional process is the mechanism by which the effects of treatment are created and realized (Smith, 1988). Like every interpersonal relationship, the psychotherapeutic process has a frame or structure that delineates and identifies the purpose and meaning of the relationship. This frame consists of socially dictated components which serve as guidelines to how the therapeutic process should occur (Epstein, 1994).

Trust is an element of psychotherapy which is of primary importance. Clients expect that within reason, their needs are the focal concern of the therapeutic relationship, and that the therapist's responses and interactions have the intention of promoting effective treatment gains (Galletly, 1993). The ethical dictum is that the therapist must act so as to avoid injury or harm to the client.

The client has been encouraged to suspend their usual defenses, reveal deeply personal thoughts and feelings, and ultimately become somewhat dependent upon the clinician. The client enters treatment with the expectation of being helped. In addition, the development of intense feelings toward the therapist as a manifestation of transference occurs. In this way, the therapist, doctor, or psychologist is in the position of power, while the client assumes a more vulnerable position (Blackshaw & Patterson, 1992).

In addition, legal duties mandated by courts and statutes have further defined the psychotherapy relationship and the relevant and appropriate boundaries between the therapist and client. In sum, therapists are therefore accountable to professional, ethical, and legal standards (Simon, 1992).

The process of transference provides an avenue for either beneficial treatment to become actualized or noxious to the client by the therapist exploiting the established trust. In addition, countertransference, or feelings toward a client a therapist may have also become a mode of either positive or negative influence to the therapeutic process (Smith, 1988). The therapeutic relationship is described as fiduciary in nature (Simon, 1992). A fiduciary relationship exists when the client's compliance requires trust and vulnerability while the therapist's position renders a clearly influential role upon the relationship. The client relies on the therapist to provide treatment resulting in an asymmetrical or fiduciary relationship (Epstein, 1994).

6.10.9 GENERAL GUIDELINES

Sigmund Freud provided specific guidelines defining the ideal psychotherapeutic structure. Cited in Epstein (1994), Freud stated that his guidelines were imperative for creating and maintaining trust between therapist and client, and provided self-protection for the therapist. Freud advised that the client should receive informed consent explaining the nature of treatment. The therapist should practice abstinence and the nonexploitation of a client which included not confusing transference for real or true feelings or expressions toward the therapist (Lakin, 1991). The therapist should maintain a neutral stance including not advocating any one position or stance over another. The therapist should avoid dual agency or becoming involved with a client already personally known to the therapist. The therapist should practice relative anonymity and keep personal disclosures to a minimum. Finally the therapist should collect coherent and rational fees at regular intervals (cited in Epstein, 1994). Recently, Simon (1992) conceptualized treatment boundaries or frames necessary to maintain the integrity of treatment. These guidelines and principles mirror the ethical standards espoused by the American Psychological Association (APA). The most recent published set of APA guidelines are the 1992 version (Canter et al., 1994). Simon advises that the therapist must follow the guidelines below in order to create an effective ethical treatment process (Simon, 1992).

(i) Maintain therapist neutrality

(ii) Encourage psychological separateness between client and therapist

(iii) Maintain confidentiality

(iv) Secure informed consent for treatment

(v) Encourage verbal interaction with client

(vi) Ensure that no current or future personal relationship occurs with the client

(vii) Minimize or avoid physical contact

(viii) Preserve anonymity of the therapist

(ix) Establish and maintain a consistent fee policy

(x) Provide a safe, consistent, and private treatment setting

(xi) Provide clear definition of time and length of treatment sessions.

6.10.10 APA CODE

The 1992 Ethics Code of the APA clearly identifies general principles and ethical standards for all areas of psychological treatment.

The creation of the first formal APA Ethics Code was the result of the increasing professional activity of psychologists, both industrial and in mental health settings. As activity increased, regulations and guidelines were needed to ensure the integrity of treatment and the protection of all clients. The first Code was formulated from a critical incident method in 1948 which involved psychologists describing what they felt was an ethical situation or dilemma. The Ethics Committee of the APA then created a draft and final version of guidelines and standards in response to the previously identified situations from the empirical data obtained from the critical incidents listed and defined by psychologists. The first version was published in 1953, with subsequent revisions in 1958, 1962, 1965, 1972, 1977, 1979, 1981, 1989, and 1992 (Canter et al., 1994).

Following the outline of General Standards espoused by the APA Ethics Code, psychologists are expected to provide services for which they are competent based on their education and training. Continued training and education is required for the psychologist to remain in touch with current and updated treatment methods and research outcomes. The psychologist should speak with the client in a language understood by the client to ensure that informed consent has occurred. In this way, the client is aware and cognizant of the proposed treatment plan, procedures, or research (Canter et al., 1994).

Psychologists need to be aware of how human differences such as age, gender, race, sexual orientation, and disabilities can affect their work and recognize when they may or may not be qualified to treat any one particular individual. If this occurs, the expectation exists that the therapist will provide an appropriate referral to a competent therapist with the relevant expertise. Psychologists should always respect the rights of others to hold values, beliefs, or opinions different from their own. Psychologists do not degrade those of differing religious values, cultural or ethnic backgrounds, or politics. In this way, they are nondiscriminative (Canter et al., 1994).

Psychologists are prohibited from engaging in any form of sexual harassment, defined by sexual solicitation, physical advances, or verbal or nonverbal conduct that is sexual in nature that occurs in the therapeutic encounter. If a client has complained of being sexually harassed, the psychologist does not make prejudicial decisions in response to the client being a complainant, for example, denying admission to a graduate program. Additionally, psychologists are prohibited from harassing or demeaning clients in any form (Canter et al., 1994).

Psychologists must take reasonable steps and precautions to avoid harming their clients, and to minimize harm when it is foreseeable and unavoidable. In this way, psychologists must be able to recognize when their own personal problems or conflicts may interfere with their effectiveness as a treating clinician. If a psychologist is aware that their personal problems may be interfering with treatment, they are obliged to seek assistance and determine if they need to suspend or terminate treatment. A psychologist must also not misuse their influence upon a client, especially as the relationship between the therapist and client is asymmetrical with the position of power lying with the therapist (Canter et al., 1994).

The 1992 Ethics Code specifically prohibits the psychologist and client engaging in any type of multiple relationship. A psychologist must refrain from beginning a personal, scientific, professional, financial, or other relationship with a client as it may involve future harm or exploitation of the client. The Code does, however, recognize that in some rural areas of the country, situations in which clinicians who know the client before treatment begins may not always be avoided due to lack of resources. Psychologists are also advised not to barter with a client which means refraining from accepting goods or services as a way of paying for psychological services. The psychologist may, however, participate in bartering providing it does not cause the client harm, and does not involve any exploitation of the client (Canter et al., 1994).

Exploitation of the client is strictly forbidden in the Ethics Code. This includes not engaging in any sexual or personal relationship with students, supervisees, and clients (Canter et al., 1994). Clear evidence exists that clients are harmed.

In conclusion, therapists are bound by clinical, ethical, and legal standards to maintain the integrity of the psychotherapeutic process. Epstein (1994) writes, “Ideally, the therapist will be able to fine-tune the frame into an empathic dynamic structure that is sensitive to the client’s changing needs” (p. 17). This is essential, particularly since the therapeutic relationship plays such an important part as an agent in the change process.

ACKNOWLEDGMENT

The authors wish to acknowledge Gina Fusco, a doctoral student in clinical psychology at the Philadelphia College of Osteopathic Medicine, for her assistance with this chapter.

6.10.11 REFERENCES

Alexander, J. F. (1976). Social reinforcement in the modification of agoraphobia. *Archives of General Psychiatry, 19,* 423–427.

Alexander, L. B., & Luborsky, L. (1986). The Penn helping alliance scales. In L. S. Greenberg & W. M. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 35–366). New York: Guilford Press.

Bandura, A. (1977). *Social learning theory*. Hillsdale, NJ: Prentice-Hall.

Beck, A. T., Freeman, A. and associates (1990). *Cognitive therapy of personality disorders*. New York: Guilford Press.

Beck, A. T., Rush, J. A., Shaw, B., & Emery, G. (1979). *Cognitive therapy for depression*. New York: Guilford Press.

Blackshaw, R., & Patterson, P. (1992). The prevention of sexual exploitation of clients: Educational issues. *Canadian Journal of Psychiatry, 37,* 350–353.

Bugental, J. F. T., & Sterling, M. M. (1995). Existential–humanistic psychotherapy: new perspectives. In A. S. Gurman & S. B. Messer (Eds.), *Essential psychotherapies: Theories and practice* (pp. 226–260). New York: Guilford Press.

Canter, M., Bennett, B., Jones, S., & Nagy, T (1994). *Ethics for psychologists a commentary on the APA ethics code*. Washington, DC: American Psychological Association.

Eidelberg, L. (1968). *Encyclopedia of psychoanalysis*. New York: The Free Press.

Ellis, A. (1985). *Overcoming resistance: Rational-emotive therapy with difficult clients*. New York: Springer.

Epstein, R. (1994). *Keeping boundaries maintaining safety and integrity in the psychotherapeutic process*. Washington, DC: American Psychiatric Press.

Ferenczi, S., & Rank, O. (1925). *The development of psychoanalysis*. NMDP.

Frank, J. (1985). Therapeutic components shared by all psychotherapies. In M. Mahoney & A. Freeman (Eds.), *Cognition and psychotherapy* (pp. 49–79). New York: Plenum.

Galletly, C. (1993). Psychiatrist–client sexual relationships: The ethical dilemmas. *Australian and New Zealand Journal of Psychiatry, 27,* 133–139.

Gartrell, N., Herman, J., Olarte, S., Feldstein, M., & Localio, R. (1986). Psychiatrist–client sexual contact: Results of a national survey. *American Journal of Psychiatry, 143,* 112–131.

Greenberg, J., & Cheselka, O. (1995). Relational approaches to psychoanalytic psychotherapy. In A. S. Gurman & S. B. Messer (Eds.), *Essential psychotherapies theory and practice* (p. 56). New York: Guilford Press.

Gorton, G., & Samuel, S. (1996). A national survey of training directors about education for prevention of psychiatrist-client sexual exploitation. *Academic Psychiatry, 20*(2), 92–98.

Granvold, D. K., & Wodarski, J. S. (1994). Cognitive and behavioral treatment: Clinical issues, transfer of trauma, and relapse prevention. In D. K. Granvold (Ed.), *Cognition and behavioral treatment: Method and applications* (pp. 353–375). Pacific Grove, CA: Brooks/Cole.

Hall, C. S., & Lindzey, G. (1970). *Theories of personality*. New York: Wiley.

Horney, K. (1937). *The neurotic personality of our time*. New York: W. W. Norton.

Horney K. (1939). *New ways in psychoanalysis*. New York: W. W. Norton.

Horney, K. (1945). *Our inner conflicts*. New York: W. W. Norton.

Horvath, A., & Greenberg, L. S. (1986). The development of the working alliance inventory. In L. S. Greenberg & W. S. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 529–556). New York: Guilford Press.

Kaplan, H. I., Sadock, B. J., & Grebb, J. A. (1994). *Kaplan and Sadock's' synopsis of psychiatry: Behavioral sciences clinical psychiatry* (7th ed.). Baltimore: Williams & Wilkins.

Lakin, M. (1991). *Coping with ethical dilemmas in psychotherapy*. New York: Pergamon.

Lazarus, A. (1989). *The practice for multimodal therapy*. Baltimore: Johns Hopkins University Press.

Mathews, A. M., Johnston, D. W., Lancashire, M., Munby, M., Shaw, P. N., & Gelder, M. G. (1976). Imaginal floofing and exposure to read phobic situations: Treatment outcome with agoraphobic patients. *British Journal of Psychiatry, 129,* 362–371.

May, R. (1969). *Love and will*. New York: W. W. Norton.

May, R., & Yalom, 1. (1989). Existential psychotherapy. In R. J. Corsini & D. Wedding (Eds.), *Current psychotherapies*, (4th ed., pp. 363–402). Ithaca, IL: F. E. Peacock Publishers Inc.

Meichenbaum, D., & Turk, D. C. (1987). *Facilitating treatment: A practitioners guidebook*. New York: Plenum.

Mullahy, P. (1970). *Psychoanalysis and interpersonal psychiatry*. New York: Science House, Inc.

Munroe, R. (1955). *Schools of psychoanalytic thought*. New York: Dryden.

Ofman, W. V. (1985). Existential psychotherapy. In H. Kaplan & B. J. Sadock (Eds.), *Comprehensive textbook of psychiatry/IV* (4th ed., Vol. 2). Baltimore: Williams & Wilkins.

Orlinsky, D. E., & Howard, R. I. (1986). The relation of process to outcome in psychotherapy. In S. L. Garfield & A. E. Berguin (Eds.), *Handbook of psychotherapy and behavior change*. New York: Wiley.

Pope, K., Sonne, J., & Holroyd, J. (1993). *Sexual feelings in psychotherapy*. Washington, DC: American Psychological Association.

Rogers, C. (1961). *On becoming a person*. Boston: Houghton Mifflin.

Rogers, C. (1977). *Carl Rogers on personal power: Inner strength an its revolutionary impact*. New York: Delacorte Press.

Rogers, C. R. (1987). The underlying theory: Drawn from experiences with individuals and groups. *Counseling and Values, 32*(1), 38–45.

Safran, J. D. (1990). Towards a refinement of cognitive therapy in light of interpersonal theory. I: Theory. *Clinical Psychology Review, 10,* 87–105.

Safran, J. D., & Segal, Z. V. (1990). *Interpersonal process in cognitive therapy*. New York: Basic Books.

Safran, J. D., & Wallner, L. (1991). The relative predictive validity of two therapeutic alliance measures in cognitive therapy. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 188–195.

Seguin, C. A. (1965). *Love and psychotherapy*. New York: Libra Publishers.

Shelton, J. L., & Levy, R. I. (1981). *Behavioral assignments and treatment compliance: A handbook of clinical strategies*. Champaign, IL: Research Press.

Simon, R. (1989). Sexual exploitation of clients: How it begins before it happens. *Psychiatric Annals, 19*(2), 104–112.

Simon, R. (1992). Treatment boundary violations: Clinical, ethical, and legal considerations. *Bulletin of the American Academy of Psychiatry & Law, 20*(3), 269–288.

Sloan, R. B. (1975). *Psychotherapy versus behavior therapy*. Cambridge, MA: Harvard University Press.

Smith, J. (1988). Therapist–client sex: Exploitation of the therapeutic process. *Psychiatric Annals, 18*(1), 59-63.

Stark, M. (1994). *Working with resistance*. Northvale, NJ: Jason Aronson, Inc.

Suh, C., Strupp, H., & O'Malley, S. (1986). The Vanderbilt process measures: The Psychotherapy Process Scale

(VPPS) and the Negative Indicators Scale (VNIS). In L. Greenberg & W. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 285–323). New York: Guilford Press.

Sullivan, H. S. (1953). *The interpersonal theory of psychiatry*. New York: W. W. Norton.

Sullivan, H. S. (1954). *The psychiatric interview*. New York: W. W. Norton.

Sullivan, H. S. (1962). *Schizophrenia as a human process*. New York: W. W. Norton.

Tasman, A., Kay, J., & Lieberman, J. A. (1997). *Psychiatry* (Vol. 1). Philadelphia: W. B. Saunders.

Wachtel, P. L. (1982). *Resistance: Psychodynamic and behavioral approaches*. New York: Plenum.

Wright, J. H., & Davis, D. (1994). The therapeutic relationship in cognitive-behavior therapy: Patient perceptions and therapists responses. *Cognitive and Behavioral Practice, 1,* 25–45.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.11
Treatment Maintenance and Relapse Prevention

JOHN W. LUDGATE
Bristol Regional Medical Center, Bristol, TN, USA

6.11.1 OVERVIEW 251
6.11.2 DEFINITIONS OF RELAPSE AND RECURRENCE 252
6.11.3 REVIEW OF RELAPSE RATES AND PREDICTORS OF RELAPSE 252
  6.11.3.1 Depression 252
  6.11.3.2 Anxiety Disorders 253
  6.11.3.3 Substance Abuse 255
  6.11.3.4 Eating Disorders 255
  6.11.3.5 Overall Review of the Extent of the Problem of Relapse 256
  6.11.3.6 Review of the Predictors of Relapse 256
  6.11.3.7 Treatment Implications that Follow from Research on Predictors of Relapse 256
6.11.4 MODELS OF RELAPSE 257
  6.11.4.1 The Medical Biological Model 257
  6.11.4.2 Cognitive Behavioral Models 257
    6.11.4.2.1 Marlatt and Gordon's model 257
    6.11.4.2.2 Shiffman's model 258
    6.11.4.2.3 Teasdale et al.'s model 258
6.11.5 CLINICAL STRATEGIES TO REDUCE RELAPSE RISK 258
  6.11.5.1 General Overview of Relapse Prevention Strategies 258
  6.11.5.2 Application of Specific Relapse Prevention Procedures at Different Points in Therapy 259
    6.11.5.2.1 Assessing the risk of relapse 259
    6.11.5.2.2 Early therapy activities 259
    6.11.5.2.3 Throughout therapy strategies 259
    6.11.5.2.4 Pretermination procedures 260
    6.11.5.2.5 Follow-up and aftercare activities 260
  6.11.5.3 Working with the Patient who has Relapsed 261
6.11.6 CURRENT STATUS AND FUTURE DIRECTIONS FOR RELAPSE PREVENTION 261
6.11.7 REFERENCES 262

6.11.1 OVERVIEW

In the psychotherapy literature the main source of attrition in psychotherapy has been seen as dropouts during the acute phase of therapy. Rather less attention has been given to the problem of relapse and recurrence following sucessful treatment. However, since the late 1970s there has been an upsurge of interest in the long-term effects of psychological treatments and how treatment effects can be maximized and maintained. As will be demonstrated in this chapter there is still a considerable relapse rate associated with all treatments of psychological disorders including psychotherapy and pharmacotherapy. Cognitive therapy (Beck, Rush, Shaw, & Emery, 1979), behavior therapy (Wolpe & Lazurus, 1966) and interpersonal psycotherapy (Klerman, Weissman, Rounsaville, & Chevron, 1984) have all demonstrated some degree of efficacy in the acute treatment phase and also seem to have some preliminary promise in reducing relapse in certain clinical populations. However, the task of further reducing the still high relapse rates found in several disorders such as depression (Belsher & Costello, 1988) and substance abuse (Hunt, Barnett, & Branch, 1971) needs to be given serious attention by clinicians and researchers. In this chapter the relapse rates and predictors of relapse in a number of disorders will be reviewed, and the implications of this research for clinical practice will be outlined. A number of models of relapse that can guide reseach and clinical practice will be described. Attention will be given to the assessment of relapse risk and to planning treatment in a manner designed to reduce the risk of relapse. Relapse prevention and maintenance facilitating strategies will be outlined. Some concluding comments on the present status of maintenance and relapse prevention therapy procedures will be offered and future directions for research and clinical practice in this area will be suggested.

6.11.2 DEFINITIONS OF RELAPSE AND RECURRENCE

It is important to clearly define the terms "relapse" and "recurrence" as these are often used interchangeably. In the field of depression, Klerman (1978) has suggested that clinicians and researchers adopt the convention that the term "relapse" be used to refer to a return of symptoms within six to nine months after the onset of the index episode and that the term "recurrence" be used to denote a return of symptoms after a period of six or twelve months of remaining symptom-free. Klerman suggests that relapse may refer to continuity or re-emergence of symptoms of the original episode, while recurrence would be considered a new episode. Despite these guidelines there is still considerable confusion regarding this distinction in the literature.

Distinguishing between a chronic continuity of symptoms or behavior and a relapse following recovery or remission has been problematic in this field and made it difficult to establish accurate figures on relapse and recurrence. Designating patients as relapsed presupposes that they have experienced a recovery or remission from the original index episode. Recovery can be defined in a number of different ways. In the field of depression, for example, recovery is said to occur when the number and severity of symptoms falls below the threshold used for defining onset for a specified period of time (Keller, Shapiro, Lavori, & Wolfe, 1982).

Defining relapse can also be problematic. Brownell, Marlatt, Lichenstein, and Wilson (1986) have noted two different dictionary definitions of the term "relapse." The first refers to an outcome and defines relapse as "the recurrence of symptoms of disease after a period of improvement," and the second refers to a process, namely, "the act of backsliding, worsening, or subsiding." The outcome definition, often associated with biological or medical models, has very different implications from the process definition, which suggests that relapse involves a number of stages and is not a single event (Marlatt & Gordon, 1985).

Studies on long-term outcome in treatment often do not clearly specify the criteria used to define relapse. Also, since different studies do not utilize standardized criteria for recovery, relapse, and recurrence, comparing studies is difficult. Wilson (1992a) points out that there are many problems inherent in attempting to draw conclusions about relapse in psychological disorders due to the difficulty in drawing distinctions between lapse, relapse, and recurrence and clarifying the usage and measurement of these concepts in different disorders. Other methodological problems with the studies conducted in this field include inadequate cell size, differential attrition, failure to include control or comparison groups, and insufficient duration of follow-up.

6.11.3 REVIEW OF RELAPSE RATES AND PREDICTORS OF RELAPSE

6.11.3.1 Depression

A great deal of research has been conducted into the long-term outcome of depressive disorders (see also Chapters 3 and 16, this volume). The NIMH consensus panel (NIMH, 1985) concluded that as many as 50% of patients with recurrent unipolar depressive disorder who recover from an episode of depression will have a recurrence within two years. Other reviewers in this field (Belsher & Costello, 1988; Lavori, Keller, & Klerman, 1984) confirm these findings and additionally report a 20% relapse rate within two months and a 40% rate within a year. Approximately 70% of unipolar depressed patients respond to antidepressant medication in the short term but there is a considerable relapse rate with short-term pharmacotherapy (Ludgate, 1991). There is now an emerging body of evidence demonstrating that the use of maintenance medication (antidepressants or lithium) reduces the probability of relapse. Prien and Kupfer (1986) concluded in a review of this field that the risk of relapse was approximately 20% in patients continuing to take lithium or antidepressant medication compared to 50% who were switched to a placebo. Thase (1990) suggests that maintenance pharmacotherapy reduces the risk of recurrence over a two to three year period from 80% to 50%.

Behavioral approaches have shown some promise in producing positive long-term outcome and reducing relapse in depressed patients (Gonzales, Lewinsohn, & Clarke, 1985; Thompson & Gallagher, 1984). Social skills training has been found to be superior to both psychotherapy and tricyclic maintenance medication at six month follow-up (Hersen, Bellack, Himmelhoch, & Thase, 1984). Behavioral therapy was found to have a better course than pharmacotherapy, relaxation therapy, and nondirective therapy in a two and a half year follow-up of depressed outpatients (McLean & Hakistian, 1990).

Encouraging results have also been found using interpersonal psychotherapy with depressed patients (Weissman, 1994). Weisman, Klerman, Pruscoff, Shalomskas, and Padian (1981) found no difference between interpersonal psychotherapy and antidepressant medication at one year follow-up. Frank et al. (1990) used continuation interpersonal psychotherapy (on a monthly basis for 20 weeks) and antidepressant medication as a continuation treatment for unipolar depressed patients initially successfully treated with either interpersonal psychotherapy or medication. They found that maintenance medication had significant prophylactic effects and interpersonal psychotherapy had modest effects in reducing relapse risk.

A number of reviews (Hollon, Shelton, & Loosen, 1991; Ludgate, 1995; Wilson, 1992) have concluded that cognitive behavioral therapy has a prophylactic effect in the treatment of depression. It appears that patients treated with short-term antidepressant medication have roughly twice the relapse risk compared to patients receiving cognitive behavioral therapy (22–38% vs. 65–78%). In the NIMH study (Shea et al., 1992), cognitive therapy, interpersonal psychotherapy, and short-term Imipramine were compared. Relapse rates at 18-month follow-up were: 36% for cognitive therapy, 33% for interpersonal psychotherapy, 50% for Imipramine, and 33% for placebo. Return to treatment was significantly lower in the cognitive therapy group than in the other groups. In the only studies that compared cognitive therapy to maintenance antidepressant medication (Blackburn, Eunson, & Bishop, 1986; Evans et al, 1992), cognitive therapy still fared better in terms of relapse rates although the use of maintenance medication also resulted in relatively low rates of relapse. Wilson (1992b) concludes that cognitive therapy is at least as effective as maintenance medication properly maintained for one year. In all of the studies carried out on the long-term outcome of cognitive therapy for depression, no explicit relapse prevention or maintenance procedures are described. The therapy administered follows that described by Beck et al. (1979) and focuses on providing cognitive and behavioral strategies that aim to alleviate depressive symptomatology with the addition in some studies of booster sessions and some additional therapy (Blackburn et al., 1986). Only one study (Berlin, 1985) used an explicit relapse prevention focus and attempted to determine its effects. Self-critical patients were randomly assigned to standard cognitive behavior therapy or to a specifically designed relapse prevention program. Both treatments were equivalent in their effects on maintenance of treatment effects (reduced depression and increased self-esteem) at six-month follow-up.

Despite lower relapse rates in depressed patients treated with maintenance pharmacotherapy, cognitive therapy and interpersonal psychotherapy compared to other treatment methods, relapse is still a significant problem in this area.

Wilson (1992b) suggests that the following variables appear to put depressed individuals at greatest risk for relapse: chronicity, number of previous episodes, being in an environment of high expressed emotion or where a high number of critical comments are made by a spouse (Hooley & Teasdale, 1989), adverse life events, residual depression after treatment, posttherapy dysfunctional cognitive style (including the presence of dysfunctional attitudes and a negative attributional style), low self-efficacy concerning control over negative thoughts, dissatisfaction with major life areas, and coexisting medical problems. Other predictors for which there is at least some empirical support include endogenous subtype, duration of index episode, the presence of an Axis II disorder, higher levels of hopelessness at discharge, absence of social support, and persistent neuroendocrine dysregulation after recover (Belsher & Costello, 1988; Lavori et al., 1984; Ludgate, 1994).

6.11.3.2 Anxiety Disorders

While there has been a good deal of research carried out on the short-term effects of a number of therapies for anxiety disorders (see Chapters 17–22, this volume), there are few studies that provide outcome data on patients with anxiety disorders followed up for any significant period of time. Rappee (1991), reviewing research on generalized anxiety disorders, concludes that the positive effects of psychological treatments tend to endure over 6–12 month follow-up periods. In a three-year follow-up of generalized anxiety disorder patients receiving short-term stress innoculation training (Meichenbaum, 1977), Holcomb (1986) found that these patients required fewer hospital readmissions compared to patients receiving pharmacological treatment.

In agoraphobia, Telch, Tearnan, and Taylor (1983) found relapse rates of 27–50% in patients treated with Imipramine alone. Zitrin, Klein, and Woerner (1980) found that active psychotherapy methods such as exposure, when used in conjunction with Imipramine in the treatment of agoraphobia, reduces the relapse rate to 19% vs. 31% with Impipramine alone over a two-year follow-up. Encouragingly, Emmelkamp and Kuipers (1979) found that 60–70% of agoraphobic patients who responded to cognitive behavior therapy maintained their improvement at four-year follow-up. In treating panic disorder, medications such as Alprozolam and Imipramine have been found to be very effective at eliminating panic attacks in the short term but relapse rates of 70–90% following discontinuation have been found with Alprozolam and 35–40% with Imipramine (Brown & Barlow, 1992). A series of studies on cognitive therapy with panic disorder patients (Beck, Sokol, Clark, Berchick, & Wright, 1992; Clark et al., 1994; Craske, Brown, & Barlow, 1991; Ost & Westling, 1995) show impressively low relapse rates ranging from 0 to 19% in patients treated with cognitive behavior therapy whose panic attacks resumed in the follow-up period. In a review of behavioral treatments of agoraphobia, simple phobia and obsessive compulsive disorder with a follow-up period of at least one year, Ost (1989) reports relapse rates of 24% for agoraphobia, 4% for simple phobia, and 15% for obsessive compulsive disorder. Ost concludes that: further improvement after treatment is moderate in agoraphobia and obsessive compulsive disorder and nonexistent in simple phobia, overall about one-fifth of these patients relapse, and a large proportion seek further treatment. Ost (1989) describes a cost-effective maintenance program for anxiety disorders following initial behavioral treatment which is introduced in the last session of treatment when a contract to implement maintenance activities is signed by the patient. Over a six-week period following termination, the patient and therapist have 1.5–2 hours of telephone contact. Maintenance activities include: reviewing changes achieved and skills acquired during therapy, continued practicing of skills, distinguishing a setback from a relapse, identifying high-risk situations for relapse, and coping with setbacks. Results from three studies where agoraphobic panic and simple phobia patients completing this maintenance program were evaluated at 6–15 months, follow-up showed that overall, only 7% of patients relapsed and 10% needed further treatment, which compares favorably with the means of previously reported studies where 14% relapsed and 25% needed further treatment.

A number of longer-term follow-up studies have been conducted in obsessive compulsive disorder treated with cognitive behavior therapy. Rates of relapse reported are consistently in the 20–25% range in follow-up periods ranging from seven months to three years (Foa et al., 1983). Similarly, relapse rates of 20–30% have been reported over follow-up periods of up to six years following treatment by exposure therapy with this population (Emmelkamp & Kuipers, 1979).

While follow-up studies examining relapse and maintenance of treatment effects in patients with anxiety disorders are somewhat limited, the literature suggests that patients treated with cognitive behavior therapy have a lower risk of relapse compared to other treatment methods. However, as in the case of depression, a significant proportion of patients do not maintain treatment gains and experience a recurrence. There are few well-controlled studies that investigate predictors of maintenance and relapse following treatment for anxiety disorders. However, certain clinical characteristics have been implicated in poorer longer-term outcome. These include a comorbid diagnosis of a personality disorder (Turner, 1987), concurrent depression (Brown & Barlow, 1992), negative life events (Munroe & Wade, 1988), and interpersonal stressors and health problems (Tearnan, Telch, & Keefe, 1984) Rappee (1991) has hypothesized that information processing deficits are implicated in the long-term maintenance of generalized anxiety disorder. In the case of panic disorder, a cognitive factor, that is, misinterpretation of body sensations, is hypothesized to be involved in the etiology and maintenance of panic attacks (Clark et al., 1988), and this hypothesis has been given support in two studies which showed that higher scores on the Body Sensations Interpretation Questionnaire (Clark et al. 1988), an instrument which measures the tendency to misattribute sensations, were predictive of relapse at 15-month follow-up in panic patients who were panic-free at termination (Clark et al., 1988; Westling & Ost, 1995). In the area of obsessive compulsive disorder, Emmelkamp and colleagues (Emmelkamp, Kloek, & Blaauw 1992) have empirically demonstrated that life events, inadequate coping style, and high expressed emotion environments are predictive of relapse. In conclusion, though little systematic research has been done in this area, the following factors appear to increase the risk of relapse in anxiety disorders: comorbid depression or personality disorder, adverse life events, dysfunctional cognitive style after treatment, deficient coping skills, and a high expressed emotion environment.

6.11.3.3 Substance Abuse

The treatment of substance abuse has become increasingly more research-driven in recent years (see Chapter 25, this volume). Stimulated by the seminal work of Marlatt and colleague (Marlatt & Gordon, 1985), a good deal of research on relapse in addictive behaviors has been carried out. In this field, rates of relapse are disturbingly high across the spectrum of addictive behaviors (Marlatt & Gordon, 1985). Additionally, the temporal course of the relapse process is remarkably consistent across the various disorders. In a comprehensive review of a large number of treatment studies that followed successfully treated addictive behavior patients over a 12-month follow-up, Hunt et al. (1971) plotted survival rates for a number of addictive disorders and concluded that: the vast majority (65–75%) of treatment successes relapse within 12 months, two out of three relapses occur within three months, relapse rates seem to stabilize and plateau over time (after six months), and relapse rates and time to relapse are simular for heroin, nicotine, and alcohol addictions. Marlatt and Gordon (1985) confirm the finding that approximately two-thirds of all relapses across addictive disorders occur within 90 days of initiation of a treatment program or attempts at cessation. The high rate of relapse in addictive disorders is exemplified by research on smoking where as many as 88% of treated patients will relapse over a two-year period with the majority of relapses occurring within seven months with some stabilization of rates after 15 months (Brandon, Tiffany, Obremski, & Baker, 1990).

Marlatt and his colleagues (Marlatt & Gordon, 1985) have carried out extensive research into the predictors of relapse in addictive disorders. Based on an analysis of 311 relapse episodes obtained from patients with problems in the areas of smoking, drinking, gambling, and overeating, it was found that the following three factors accounted for over 70% of relapses: negative emotional states (e.g. frustration, anger, anxiety, depression, and boredom): social pressure; and interpersonal conflict (Cummings, Gordon, & Marlatt, 1980). In reviewing this area, Marlatt and Gordon (1985) note that there has been condiderable consistency found in studies examining predictors of relapse across various addictive behaviors. They further state that these three factors (negative emotional states, social pressure, and interpersonal conflict) appear to be the most frequently found high-risk situations for relapse in addictive disorders with urges and temptations, negative physical states, testing personal control, and positive emotional states being implicated in a somewhat smaller number of relapse situations.

6.11.3.4 Eating Disorders

Psychological treatments for eating disorders have grown in range and sophistication in recent years (see Chapter 29, this volume). A good deal of attention is now being given to long-term outcome in eating disorders although, as Orimoto and Vitousek (1992) note, these studies often track the status of diagnosed individuals over various time periods following diverse and often poorly specified forms of treatment, which makes comparison of findings across studies difficult. A review of the literature on anorexia nervosa suggests that a good long-term outcome is attained by roughly 40%, with 11–33% in an intermediate category and 14–42% judged to be doing poorly (Orimoto & Vitousek, 1992). Relapse rates following treatment are hard to estimate since a great number of anorexic patients do not attain recovery or remission after treatment and cannot thus be candidates for relapse by definition. Orimoto and Vitousek conclude that only roughly 50% of anorexics will eventually attain normal weight and menstruation and a significant subgroup will experience a relapse in the future, with the exact percentage of actual relapses being hard to ascertain.

Bulimia nervosa is a relatively recent diagnosis and the long-term course of the disorder is still being investigated. However, some tentative conclusions can be drawn from the literature. Approximately 53% of patients improve over 12–15 months without formal therapy (Mitchell, Davis, & Goff, 1985). However, only one-third of a large sample of treated and untreated patients remained abstinent from bulimic behavior over a one-year period (Mitchell, Davis, Goff, & Pyle, 1986). In addition Mitchell et al. (1986) found that as many as one-third of relapses take place within one month of treatment. Keller, Herzog, Lavori, and Bradburn (1992) found that one-third of bulimics remained in the index episode three years later and there was a cumulative probability of relapse of 63% over a follow-up of 35–42 months. Orimoto and Vitousek (1992) summarize the cognitive behavior therapy outcome studies carried out in this area and show a median figure for abstinence from binging and from purging of 61% and 45.5%, respectively, at follow-ups of three months and longer. They also suggest, based on some preliminary findings, that interpersonal psychotherapy may show some promise in preventing relapse and maintaining treatment gains. Few studies using cognitive behavioral methods have used treatment formats that use explicit relapse prevention procedures, and this is an area for further investigation following the promise of initial studies in this area (Fairburn, Kirk, O'Connor, & Cooper, 1986; Wilson, Rossiter, Kleifield, & Lindholm, 1986).

In the field of eating disorder, there are few systematic studies of predictors of long-term outcome following treatment. Keller et al. (1992) found that maintenance of recovery in bulimia nervosa could be predicted by less disturbed eating behaviors, more positive self-body image at intake, and having good friends. Mitchell et al. (1985) found that the antecedent conditions for relapse included coping with stressful situations (80%), feeling anxious or nervous (23%), and feeling depressed (23%). Root (1990) showed that triggers to relapse in bulimic patients were found to include inability to cope with feelings, attitudes that give rise to negative affect, and stressful events. Freeman, Beach, Davis, and Solyom (1985) found that body image dissatisfaction at the end of therapy was found to be the most powerful predictor of subsequent relapse. Cognitive states such as guilt and the tendency to make internal, global, and uncontrollable causal attributions for an initial binging have been found to be predictors of subsequent binge behavior (Grillo & Shiffman, 1994).

6.11.3.5 Overall Review of the Extent of the Problem of Relapse

Relapse rates following treatment have been well-established for depression. The incidence of relapse in depression has been shown to be reduced by the use of maintenance medication and psychosocial treatments such as cognitive behavior therapy and interpersonal therapy. Nevertheless, relapse still remains a significant problem (up to 50% over two years) in depression. High rates of relapse (of the order of 60–80%) and a short survival time (the majority of relapses occurring within three months) are consistently found in addictive disorders. Rather less is known about long-term outcome in the fields of anxiety and eating disorders, and so no firm conclusions can be drawn. However, clinical experience would suggest that relapse and recurrence are significant problems in these disorders too. There is a considerable economic cost involved in the still high rate of recurrence in psychological disorders. The human cost is also very high in terms of distress and demoralization for the relapsing patients, their families, and for the clinicians who treat them.

6.11.3.6 Review of the Predictors of Relapse

The prediction of relapse has a number of important consequences. First, it assists in our understanding of the fundamental processes responsible for change during and after treatment and on the nature of relapse. At a more practical level this work can help to identify individuals likely to show a poorer long-term response to treatment. This can help in treatment planning. Insofar as the predictors established are potentially modifiable (e.g., personal variables, such as dysfunctional beliefs or a poor sense of self-efficacy), these can become the focus of therapy in an attempt to prevent relapse.

Across the disorders reviewed there appear to be some common factors that put individuals at risk for relapse. These include negative life events, inability to cope with stress or negative affect, persistence of certain cognitive patterns or dysfunctional beliefs, and a low sense of self-efficacy. It could be postulated that the above risk factors may also predispose patients to relapse in other psychological disorders that are beyond the scope of this review.

6.11.3.7 Treatment Implications that Follow from Research on Predictors of Relapse

Research and clinical experience on relapse and its prediction suggest a number of rules of clinical practice, which if observed may lower the risk of relapse in patients with psychological disorders (Ludgate, 1995). The following strategies may be expected to facilitate maintenance and prevent relapse:

(i) cognitive distortions, dysfunctional assumptions, and other personal variables that can create a vulnerability to relapse should be dealt with prior to termination of therapy;

(ii) residual symptoms of emotional disorder, even those at mild or subclinical levels, should be targeted therapeutically prior to discharge;

(iii) patients who evidence high symptom levels and a maladaptive cognitive style post-treatment should be kept in treatment longer and maintenance therapy provided;

(iv) comorbid Axis I or Axis II disorders should also be addressed in treatment;

(v) general problem-solving and self-management skills with wide application should be practiced regularly during therapy, and a sense of self-efficacy should be fostered;

(vi) future life stressors should be anticipated and planned for where possible;

(vii) significant others should be involved in treatment, and therapy should focus on lifestyle modification, both of which may help to reinforce changes made in treatment.

6.11.4 MODELS OF RELAPSE

The importance of deriving treatment and relapse prevention methods from a coherent model has been stressed in several reviews in this field (Ludgate, 1995, Teasdale, Segal & Williams, 1995; Wilson, 1992a). A conceptualization of relapse benefits not only theoretical and research developments in this field but, most importantly, can guide clinical practice and assist in the development of rationally derived relapse prevention procedures. Finally, a coherent explanation of relapse can assist relapsed patients in making sense of this often baffling and mysterious process and, as a result, help generate both hope and a sense of control.

6.11.4.1 The Medical Biological Model

Until recently the predominant model of relapse was the medical biological view which stressed physiological or biochemical events that led to a return of symptoms or behaviors characteristic of the syndrome. This is seen in the definition of relapse as "the recurrence of a disease after a period of improvement." In the extreme version of this model (the disease model), a relapse into addictive behavior or into depression is an event over which the individual has little control as it is assumed that it is precipitated by physiological or biochemical changes that lead invariably to symptom re-emergence or behavioral loss of control.

This model has been criticized on a number of grounds including the charge that it ignores known psychological factors in relapse and induces a sense of hopelessness and helplessness in relapsed patients, stemming from the belief that they have no control over the relapse process, which may result in a further downward spiral of negative cognition, helplessness, and symptom recurrence (Marlatt & Gordon, 1985; Wilson, 1992).

6.11.4.2 Cognitive Behavioral Models

6.11.4.2.1 Marlatt and Gordon's model

Marlatt and Gordon's (1985) model of the relapse process in addictive disorders has had a major impact in the field of relapse prevention since the late 1980s. Marlatt and Gordon postulate that newly abstinent patients experience a sense of perceived control up to the point at which they encounter a high-risk situation, which most commonly entails a negative emotional state, an interpersonal conflict, or an experience of social pressure. If individuals cope effectively in the high-risk situation, perceived control and self-efficacy increase, which in turn makes the probability of relapse decrease. Conversely, the hypothesized result of a failure to cope with a high-risk situation is a decrease in a sense of self-efficacy, which in turn increases the probability of relapse. Each experience of successful or unsuccessful coping with a high-risk situation builds up a greater or lesser sense of self-efficacy, which determines the future risk of relapse in similar circumstances. Marlatt and Gordon (1985) contend that individuals' reactions to the initial slip and their attributions regarding the cause of the slip are the determining factors in the escalation of a lapse or setback into a full-blown relapse. The transition from slip or lapse to relapse involves the "abstinence violation effect," which results from a state of cognitive dissonance regarding the nonabstinent behavior and the individual's image of being abstinent. This dissonance can be reduced by either changing the behavior or changing the image, and characteristically in this population is resolved by the latter. Internal and stable attributes for the slip also lead to further lapse behavior. This model has received a good deal of empirical support and has the merit of dismantling the process of relapse and exploring subjective and cognitive variables in a manner that has important treatment implications.

Marlatt and Gordon (1985) argue that relapse prevention and maintenance should focus on teaching individuals who are trying to change their behavior how to anticipate and cope with the problem of relapse. Relapse prevention therapy should include: identification and preparation for high-risk situations; coping with negative emotional states, interpersonal conflict, and social pressure; training in general coping or self-management skills; slip recovery and relapse-crisis debriefing; programmed relapse and relapse rehearsal; lifestyle interventions; and education about the disorder and relapse. The use of Marlatt and Gordon's model in disorders outside the area of addictions has been advocated. Wilson (1992b) examines the possible role of this model in efforts to deal with depressive relapse. In particular he stresses the need to enhance depressed patients' sense of self-efficacy, and suggests strategies to foster this. Emmelkamp et al. (1992) have adapted this model in arriving at a cognitive behavioral model of relapse in obsessive compulsive disorder. Ludgate (1994) describes how many of the strategies described by Marlatt and Gordon are also applicable at various stages in the therapy of emotionally distressed patients.

6.11.4.2.2 Shiffman's model

Shiffman (1989) argues that the risk of relapse is related to the operation of three mechanisms: enduring personal characteristics (e.g., coping style and individual vulnerability factors such as low self-esteem); background variables (e.g., life events, relationship issues, mood, motivation); and precipitants (e.g., thoughts, feelings, and events that can lead to a relapse). Background and precipitating factors can both be viewed as high-risk situations. Background factors serve to "set the stage" for relapse and may reach a threshold level where relapse becomes more likely. How individuals react to these high-risk situations will influence where their level of relapse proneness crosses the threshold for relapse. The individuals' reactions to these situations will be determined by their coping skills, their sense of self-efficacy, and their expectancies regarding outcome and the effect of using their problem-solving skills. This model is useful in introducing the valuable concept of relapse proneness and the notion of factors from different domains mutually interacting, with a threshold for relapse being established by background variables. The treatment implications of this model are: first, background factors need to be identified, monitored, and modified in order to keep relapse proneness below a threshold level; second, enduring personal characteristics need to be worked on to reduce vulnerability to relapse; and third, precipitants need to be anticipated and planned for where possible.

6.11.4.2.3 Teasdale et al.'s model

Teasdale and colleagues (1995) have proposed a model of depressive relapse which attempts to explain the process of relapse in depression and also the mechanisms by which cognitive therapy achieves its prophylactic effects in the treatment of depression. This model involves an information-processing analysis of depressive relapse. It hypothesizes that following recovery, mild states of depression can reactivate depressogenic cycles of cognitive processing similar to those found during a major depressive episode. Teasdale et al. suggest that preventive interventions such as cognitive therapy operate by changing the patterns of cognitive processing that become active in states of mild negative affect preceding a full relapse into major depression. They suggest that the redeployment of attention utilized in stress-reduction procedures based on the techniques of mindfulness meditation (Kabat-Zinn, 1990) can be integrated with cognitive therapy procedures into a system of attentional control training. This approach would be applicable to recovered depressed patients and would serve as a means of preventing relapse. Teasdale and colleagues provide a description of this training which teaches generic psychological, self-control skills and can be used on a continuing basis to maintain skills after initial training. While no data on the effectiveness of this approach in preventing relapse exist to date, this appears to be a useful and stimulating conceptualization of relapse and relapse prevention that deserves further attention.

6.11.5 CLINICAL STRATEGIES TO REDUCE RELAPSE RISK

6.11.5.1 General Overview of Relapse Prevention Strategies

Maintenance and relapse prevention depend on the nature and natural course of the specific disorder involved, and the specific procedures selected to target relapse prevention and maintenance in therapy also need to be linked to a detailed case conceptualization for each patient (Persons, 1989). However, there are a number of strategies which might be expected to facilitate maintenance and prevent relapse across the various psychological disorders (Greenwald, 1988; Krantz, Hill, Foster-Rawlings, & Zeeve 1984; Ludgate, 1995; Marlatt & Gordon 1985; Miller, 1984; Shiffman, 1989). These strategies include:

(i) increasing patient responsibility within sessions and ensuring more between-session activities as therapy proceeds;

(ii) promoting internal attributions of change;

(iii) getting the patient to practice beyond a criterion (overlearning);

(iv) working on a variety of targets and providing training in general problem-solving or self-management skills;

(v) developing a self-therapy program for use after formal therapy ends;

(vi) helping patients anticipate high-risk situations and develop an emergency plan to deal with setbacks;

(vii) education regarding relapse, developing realistic expectations, and working on cognitions regarding possible lapses or setbacks;

(viii) assisting patients in recognizing early warning signs of possible relapse and intervening quickly and appropriately when these occur;

(ix) fading the frequency of sessions later in therapy;

(x) using booster or refresher sessions after acute therapy ends;

(xi) modifying the patient's lifestyle or environment to ensure reinforcement of changes made;

(xii) involving significant others in treatment.

6.11.5.2 Application of Specific Relapse Prevention Procedures at Different Points in Therapy

6.11.5.2.1 Assessing the risk of relapse

During the assessment phase and in the early stages of therapy, data should be collected to allow a comprehensive individual conceptualization which guides treatment planning (Persons, 1989). As a part of an overall comprehensive case conceptualization, a risk analysis can be carried out for each patient. This will include: an analysis of enduring personal characteristics (e.g., coping style, dysfunctional beliefs); background variables (e.g., life events, stressors); and possible precipitants to relapse based on the patient's history (i.e., events, thoughts, and feelings that could lead to a recurrence). In addition, the patient's sense of self-efficacy and their perception of their resources or skills to deal with life problems can be assessed at different points in therapy as new information becomes available. In the case of depression, a risk-analysis might be based on the following factors: number of prior episodes, duration of episode, comorbidity (especially Axis II disorders), coexisting medical conditions, post-treatment symptom levels, vulnerability factors (poor self-esteem, dependency), dissatisfaction with life areas, post-treatment cognitive characteristics (dysfunctional attitudes, attributional style), sense of self-efficacy after treatment, social support, and life stressors. Individual patients can then be characterized as a result of this case analysis as high risk (if a significant number of these factors exist), moderate risk (where several of these factors are in evidence), or low risk (where none of these factors are in evidence). The focus of therapy and the choice of treatment strategies can then be devised in the light of the assessment of relapse risk. This risk analysis also allows the therapist to make decisions regarding aftercare. Patients at high risk can be offered continuation therapy and, if possible, engage in active therapy even after symptom remission to work on relapse prevention and maintenance. Patients at moderate risk may be offered booster sessions while patients at low risk may have their last several therapy sessions spaced out before discharge and be invited back for occasional refresher sessions (Ludgate, 1992).

6.11.5.2.2 Early therapy activities

In the early part of therapy the "mental set" necessary for eventual maintenance and relapse prevention activities can be facilitated by stressing that therapy will be time-limited and that the goal is to provide methods that patients can subsequently use to "become their own therapist." Eventual termination should be discussed and the therapist should emphasize the importance of the maintenance and underscore the notion that therapy involves learning self-control methods which patients can use on their own after therapy ends. The therapist should also stress the importance of patients being active in the therapy process, practicing new skills, and setting agendas. This will help facilitate the shift later in therapy to patients working on maintenance and relapse prevention on their own.

6.11.5.2.3 Throughout therapy strategies

As an attempt to maximize the gains made in therapy, it is recommended (Ludgate, 1995) that throughout therapy attention be paid to the following:

(i) reviews of skills at regular intervals;

(ii) self-monitoring of progress;

(iii) overpractice of skills (overlearning);

(iv) teaching skills with wide application;

(v) generalizing skills or tools;

(vi) working at the schema or belief level and dealing with Axis II issues and comorbid diagnoses, where necessary;

(vii) dealing with vulnerabilty issues;

(viii) ensuring that the patient become more and more active in therapy;

(ix) increasing the emphasis on between-session behavior;

(x) fostering internal attributions for change; and

(xi) working on reactions to setbacks.

6.11.5.2.4 Pretermination procedures

In the later stages of therapy it is often beneficial to gradually taper off or space therapy sessions. There are three important therapeutic tasks to be accomplished prior to termination which may help reduce the risk of relapse. These are: first, developing a self-therapy or maintenance plan; second, reviewing skills learned in therapy and rehearsing the application of these to future high-risk situations; and third, learning to identify early warning signals of lapses or slips and developing an "emergency plan" to cope with these setbacks. A good rationale for maintenance activities needs to be given and thoughts and feelings related to continued self-therapy should be identified and restructured if necessary at this time. A list of possible helpful maintenance activities can then be drawn up collaboratively with the patient. To introduce the procedure of self-monitoring of early warning signals of relapse, the patient first needs to be socialized to the concept of relapse or recurrence as a process, not an event, and hence to the importance of early detection and intervention in the sequence. Using the patient's past history and the pattern of events in the most recent episode, it may be possible to come up with some key prodromal symptoms or behaviors that signal the beginning of the relapse process. A regular monitoring system carried out by patients themselves or by the therapist can then be set up to determine the presence of any of these early warning signs. Finally, a specific "emergency plan" can be generated before therapy ends, which should contain steps that can be taken in the event of symptoms or maladaptive behaviors reaching certain levels. It is also helpful at this point to discuss the criteria that will be used in making a decision to return to therapy or seek hospitalization.

6.11.5.2.5 Follow-up and aftercare activities

There are a number of options for the therapist in terms of aftercare activities. These include formal maintenance or continuation therapy, booster sessions, and refresher sessions. These decisions should follow from an analysis of the potential for relapse in each particular case. Maintenance or continuation therapy (Thase, 1990) is clearly the most intensive in terms of therapist time and effort and would usually be reserved for patients who are symptomatically recovered but at high risk for relapse.

The essential difference between continuation therapy and booster sessions lies in the frequency and duration of sessions and also in the content of these sessions. Continuation therapy involves working actively on areas of vulnerability, anticipating high-risk situations, and continuing to practice and overlearn skills. Booster sessions are focused more on reviewing progress and dealing with any residual problems or current difficulties experienced by the patient. Although reviews of the effectiveness of booster sessions have shown that the addition of these follow-up sessions are only moderately successful in maintaining behavior change (Whisman, 1990), there are some obvious advantages clinically to using booster sessions after the initial treatment has ended. Booster sessions encourage patients to take more responsibility in solving problems or managing their dysfunctional behaviors or mood states with the therapist's role becoming more that of a consultant or adviser. They also allow continuing monitoring of symptoms and appropriate intervention if necessary. Patients are more likely to be accountable in terms of maintenance activities if they know they will be reporting back on their progress in booster sessions. Whisman (1990) recommends that the natural relapse rate for the problem being treated be first determined, and maintenance sessions be scheduled during this high-risk period rather than arbitrarily scheduled. He suggests that the effectiveness of booster sessions can be improved by including the following components: a review of skills covered in therapy and problem-solving difficulties encountered in their implementation; teaching new or familiar skills to deal with problems arising during the follow-up period; reinforcing the individual's accomplishments and positive achievements as well as anticipating and planning for future stressors; encouraging individuals to involve themselves in self-help or community social support systems; and promoting lifestyle changes that are incompatible with relapse. Group refresher sessions may be an alternative to individual booster therapy sessions for patients who are doing well and have a somewhat lower risk of relapse. As described by Ludgate (1992), this is a cost-effective way to promote maintenance in former patients who are invited to attend group refresher workshops that review general strategies for solving problems, managing moods and stress, and changing behaviors. The emphasis is on skills rather than specific patient issues and these workshops focus on education and skill consolidation rather than therapy *per se*. In addition to being cost-effective in terms of the therapist's time, these workshops allow monitoring of former patients' progress, which can facilitate decisionmaking regarding the need for further aftercare interventions.

6.11.5.3 Working with the Patient who has Relapsed

The tasks facing the therapist in this situation are: to conceptualize and help the patient understand the relapse process, to re-instill a sense of self-efficacy, and to collaborate with the patient in generating a plan to recover from the lapse or setback. The extent and duration of the recurrence should be carefully explored in order to help distinguish a lapse from a full relapse. The task of the therapist and patient is to interupt the relapse process. It is important to explore patients' expectations regarding recovery, relapse, and the course of their disorder. If these are unrealistic, the therapist may help patients to substitute more realistic expectations. The therapist should help the patient "make sense" of what happened in this relapse episode, which will help to demystify the recurrence, give a greater sense of control, and encourage hope that the situation can be improved. The therapist should help foster external, unstable, and specific attributions regarding the relapse. The therapist can review with the patient strategies used during this episode and what impact they had. By a process of guided discovery the patient can identify skills learned in therapy that could be applied to this situation. A list of options can be drawn up to deal with the setback. It is important that the patient be as active as possible in the task of generating the remedial plan as this offsets hopelessness and increases the sense of self-control. Finally, patients can be encouraged to consider what has been learned from this experience and what they would do differently if faced with a similar situation in the future.

6.11.6 CURRENT STATUS AND FUTURE DIRECTIONS FOR RELAPSE PREVENTION

Since the late 1980s considerable progress has been made in the area of predicting relapse and in relapse prevention. While there have been appreciable advances in producing models of relapse and therapeutic methods to prevent relapse in addictions (Gossup, 1989; Marlatt & Gordon, 1985; Shiffman, 1989) and sexual deviancy (Laws, 1989), there is a need for clinical guidelines for relapse prevention in other fields such as depression, anxiety, and eating disorders. Some practical suggestions have been offered in the literature (Greenwald, 1988; Ludgate, 1994; Miller, 1984), but there are few specific guidelines for working with individuals with these disorders to promote maintenance and reduce the risk of relapse.

Studies are needed across all areas of psychological disorders that will employ more standardized utilization of relapse prevention strategies during and after therapy.

The development of treatment protocols for the continuation phase of therapy for patients at high risk is also a priority in this field. Research into factors involved in the maintenance of treatment effects and in the prevention of relapse is an important goal for both researchers and clinicians. There are a number of research questions that need to be answered by well-controlled studies. Wilson (1992a) argues that studies with long-term follow-up periods (e.g., three years or more) are needed, as few currently exist and most studies are confined to periods of 12 months or less. Research is also needed into the mechanisms responsible for relapse, or conversely for maintenance of initial remission. Clear predictions about conditions for relapse arising out of well-formulated theoretical models need to be made and empirically tested. Increasing our ability to predict the occurrence of relapse in particular disorders can be considered a high-priority task as it has obvious clinical implications in terms of treatment planning for at-risk individuals. Research into the development of valid and reliable measures of relapse risk is also an important research and clinical task at this time. Long-term outcome needs to be investigated in well-controlled studies which include clearly defined relapse prevention strategies applied in either the acute phase of treatment or during a continuation or maintenance phase of therapy following initial recovery. More research into the efficacy of booster sessions is needed and, in particular, into the optimal timing of booster sessions after treatment. Knowledge of survival time for a particular disorder may help to identify particular points in time for certain disorders when relapse is most likely and when booster sessions could usefully be scheduled. Wilson (1992a) advocates that several different types of relapse prevention procedures such as maintenance strategies integrated into the initial treatment program, boosters sessions and post-therapy, and minimal-contact procedures (self-help groups, telephone contact, bibliotherapy) be compared in terms of efficacy. To date, the shotgun or "more is better" approach appears to be the norm and research is needed on the contribution made by the different components of relapse prevention programs (e.g., anticipation and planning for high-risk situations, early detection of warning signals of relapse, programmed relapse, relapse rehearsal, self-efficacy training, lifestyle modification, and interventions for vulnerability issues). Studies that investigate predictors of relapse and sustained improvement, particularly therapy and patient factors that are amenable to change, are needed to inform clinical practice.

Pursuing some of the suggestions outlined above would serve to further expand some of the recent exciting developments in this crucially important field of maintenance of treatment effects and prevention of relapse.

6.11.7 REFERENCES

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy for depression*. New York: Guilford Press.

Beck, A. T., Sokol, L., Clark, D. A., Berchick, R., & Wright, F. (1992). Focused cognitive therapy of panic disorder: A cross-over design and one year follow-up. *American Journal of Psychiatry, 147,* 778–783.

Belsher, G., & Costello, C. G. (1988). Relapse after recovery from unipolar depression: A critical review. *Psychological Bulletin, 104,* 84–96.

Berlin, S. (1985). Maintaining reduced levels of self-criticism through relapse prevention treatment. *Social Work Research and Abstracts, 21*(1), 21–33.

Blackburn, I. M., Eunson, K. M., & Bishop, S. (1986). A two year naturalistic follow up of depressed patients treated with cognitive therapy, pharmacotherapy and a combination of both. *Journal of Affective Disorders, 10,* 67–75.

Brandon, T. H., Tiffany, S. T., Obremski, K. M., & Baker, T. B. (1990). Postcessation cigarette use: The process of relapse. *Addictive Behaviors, 15,* 105–114.

Brown, T. A., & Barlow, D. H. (1992). Panic disorder with agoraphobia. In P. H. Wilson (Ed.), *Principles and practice of relapse prevention* (pp. 191–213). New York: Guilford Press.

Brownell, K. D., Marlatt, G. A., Lichenstein, E., & Wilson, G. T. (1986). Understanding and preventing relapse. *American Psychologist, 4,* 765–782.

Clark, D. M., Salkovskis, P. M., Gelder, M., Koehler, K., Martin, M., Anastasiades, P., Hackman, A., Middleton, H., & Jeavons, A. (1988). Tests of a cognitive theory of panic. In I. Hand & H. U. Wittchen (Eds.), *Panic and phobias II* (pp. 149–158). Berlin: Springer.

Clark, D. M., Salkovskis, P. M., Hackman, A., Middleton, H., Anastasiades, P., & Gelder, M. (1994). A comparison of cognitive therapy, applied relaxation and Imipramine in the treatment of panic disorder. *British Journal of Psychiatry, 164,* 759–769.

Craske, M. G., Brown, T. A., & Barlow, D. H. (1991). Behavioral treatment of panic disorder: A two year follow-up. *Behavior Therapy, 22,* 289–304.

Cummings, C., Gordon, J. R., & Marlatt, G. A. (1980). Relapse: Strategies of prevention and prediction. In W. R. Miller (Ed.), *The addictive behaviors: treatment of alcoholism, drug abuse, smoking and obesity* (pp. 291–321). Oxford, UK: Pergamon.

Emmelkamp, P. M., Kloek, J., & Blaauw, E. (1992). Obsessive-compulsive disorders. In P. H. Wilson (Ed.), *Principles and practice of relapse prevention* (pp. 213–235). New York: Guilford Press.

Emmelkamp, P. M. J., & Kuipers, A. C. M. (1979). Agoraphobia: A follow-up study four years after treatment. *British Journal of Psychiatry, 128,* 86–89.

Evans, M. D., Hollon, S. D., De Rubeis, R. J., Piasecki, J. M., Grove, W. M., Garvey, M. J., & Tuason, V. B. (1992). Differential relapse following cognitive therapy and pharmacotherapy for depression. *Archives of General Psychiatry, 49,* 802–808.

Fairburn, C. G., Kirk, J., O'Connor, M., & Cooper, P. J. (1986). A comparison of two psychological treatments for bulimia nervosa. *Behaviour Research and Therapy, 24,* 629–643.

Foa, E. B., Grayson, J. B., Steketee, G., Doppelt, H. C., Turner, R. M., & Lattimer, P. L. (1983). Success and failure in the behavioral treatment of obsessive-compulsives. *Journal of Consulting and Clinical Psychology, 15,* 287–297.

Frank, E., Kupfer, D. J., Perrel, J. M., Cornes, C., Jarrett, D. B., Mallinger, A. G., Thase, M., McEachran, A. B., & Grochocinski, V. J. (1990). Three year outcomes for maintenance therapies in recurrent depression. *Archives of General Psychiatry, 47,* 1093–1099.

Freeman, R. J., Beach, B., Davis, R., & Solyom, L. (1985). The prediction of relapse in bulimia nervosa. *Journal of Psychiatric Research, 19,* 349–353.

Gonzales, L. R., Lewinsohn, P. M., & Clarke, G. N. (1985). A longitudinal follow up of unipolar depressives: An investigation of predictors of relapse. *Journal of Consulting and Clinical Psychology, 53,* 461–469.

Gossup, M. (Ed.) (1989). *Relapse and addictive behavior*. New York: Tavistock/Routledge.

Greenwald, M. A. (1988). Programming treatment generalization. In L. Michelson & M. Asher (Eds.), *Handbook of anxiety and stress disorders* (pp. 583–616). New York: Plenum.

Grillo, C. M., & Shiffman, S. (1994). Longitudinal investigation of the abstinence violation effect in binge eaters. *Journal of Consulting and Clinical Psychology, 62*(3), 611–619.

Hersen, M., Bellack, A. S., Himmelhoch, J. M., & Thase, M. E. (1984). Effects of social skills training, amitriptyline and psychotherapy in unipolar depressed women. *Behavior Therapy, 15,* 21–40.

Holcomb, W. R. (1986). Stress innoculation therapy with anxiety and stress disorders of acute psychiatric inpatients. *Journal of Clinical Psychology, 42,* 864–872.

Hollon, S. D., Shelton, R. C., & Loosen, P. T. (1991). Cognitive therapy and pharmacotherapy for depression. *Journal of Consulting and Clinical Psychology, 59,* 88–89.

Hooley, J. M., & Teasdale, J. D. (1989). Predictors of relapse in unipolar depression: Expressed emotion, marital distress and perceived criticism. *Journal of Abnormal Psychology, 1,* 14–26.

Hunt, W. A., Barnett, L. W., & Branch, L. G. (1971). Relapse rates in addiction programs. *Journal of Clinical Psychology, 27,* 455–456.

Kabat-Zinn, J. (1990). *Full catastrophe living: The program of the stress reduction clinic at the University of Massachusetts Medical Center*. New York: Dell.

Keller, M. B., Herzog, D. B., Lavori, P. H., & Bradburn, I. S. (1992). The naturalistic history of bulimia nervosa: Extraordinarily high rates of chronicity, relapse, recurrence and psychosocial morbidity. *International Journal of Eating Disorders, 12,* 1–9.

Keller, M. B., Shapiro, R. W., Lavori, P. M., & Wolfe, N. (1982). Relapse in major depressive disorder: Analysis of the life table. *Archives of General Psychiatry, 39,* 911–915.

Klerman, G. L. (1978). Long-term maintenance of affective disorders. In C. Lipton, A. Dismascio, & K. Killam (Eds.), *Psychopharmacology: A generation of progress* (pp. 1303–1311). New York: Raven Press.

Klerman, G. L., Weismann, M. M., Rounsaville, B. J., & Chevron, E. (1984). *Interpersonal psychotherapy of depression*. New York: Basic Books.

Krantz, S. E., Hill, R. D., Foster-Rawlings S., & Zeeve, C. (1984). Therapist's use of and perceptions of strategies for maintenance and generalization. *The Cognitive Behaviorist, 6,* 19–22.

Lavori, P. W., Keller, M. B., & Klerman, G. L. (1984). Relapse in affective disorders: A re-analysis of the literature using life table methods. *Journal of Psychiatric Research, 18*(1), 13–25.

Laws, D. R. (Ed.) (1989). *Relapse prevention with sex offenders*. New York: Guilford Press.

Ludgate, J. W. (1991). *The long-term effectiveness of cognitive therapy in the treatment of depression: A five year follow-up study of treated patients*. Unpublished doctoral dissertation, University of Dublin, Ireland.

Ludgate, J. W. (1992). Relapse prevention groups for former cognitive therapy patients. *International Cognitive Therapy Newsletter, 6,* 10–12.

Ludgate, J. W. (1994). Cognitive behavior therapy and depressive relapse. Justified optimism or unwarranted complacency? *Behavioural and Cognitive Psychotherapy, 22*(1), 1–12.

Ludgate, J. W. (1995). *Maximizing psychotherapeutic gains and preventing relapse in emotionally distressed clients*. Sarasota, FL: Professional Resource Press.

Marlatt, G. A., & Gordon, J. R. (Eds.) (1985). *Relapse prevention: maintenance strategies in the treatment of addictive behaviors*. New York: Guilford Press.

McLean, P. D., & Hakistian, A. R. (1990). Relative endurance of unipolar depression treatment effects: Longitudinal follow-up. *Journal of Consulting and Clinical Psychology, 58,* 482–488.

Meichenbaum, D. (1977). *Stress innoculation*. New York: Pergamon.

Miller, I. W. (1984). Strategies for maintenance of treatment gains for depressed patients. *The Cognitive Behaviorist, 6,* 10–13.

Mitchell, J. E., Davis, L., Goff, G. (1985). Relapse in patients with bulimia. *International Journal of Eating Disorders, 4,* 457–463.

Mitchell, J. E., Davis, L., Goff, G., & Pyle, R. L. (1986). A follow up study of patients with bulimia. *International Journal of Eating Disorders, 5,* 441–450.

Munroe, S. M., & Wade, S. L. (1988). Life events. In C. G. Last & M. Hersen (Eds.), *Handbook of anxiety disorders* (pp. 293–305). New York: Pergamon.

National Institute of Mental Health National Institute of Health, Consensus Development Conference Statement (1985). Mood disorders: Pharmacological prevention of recurrences. *American Journal of Psychiatry, 142,* 469–476.

Orimoto, L., & Vitousek, K. B. (1992). Anorexia nervosa and bulimia nervosa. In P. H. Wilson (Ed.), *Principles and practice of relapse prevention* (pp. 85–128). New York: Guilford Press.

Ost, L. G. (1989). A maintenance group for behavioral treatment of anxiety disorders. *Behaviour Research and Therapy, 27,* 123–130.

Ost, L. G., & Westling, B. E. (1995). Applied relaxation versus cognitive behavior therapy in the treatment of panic disorder. *Behaviour Research and Therapy, 33*(2), 145–158.

Persons, J. B. (1989). *Cognitive therapy in practice: A case formulation approach*. New York: W. W. Norton.

Prien, R. F., & Kupfer, D. J. (1986). Continuation drug therapy for major depressive episodes: How long should it be maintained? *American Journal of Psychiatry, 143,* 18–23.

Rappee, R. M. (1991). Generalized anxiety disorder: A review of clinical features and theoretical concepts. *Clinical Psychology Review, 111,* 4019–440.

Root, M. P. P. (1990). Recovery and relapse in former bulimics. *Psychotherapy, 27,* 397–403.

Shea, T. M., Elkin, I., Imber, S. D., Sotsky, S. M., Watkins, J. T., Collins, J. F., Pilkonis, P. A., Leber, W. R., Krupnick, J., Dolan, R. T., & Parloff, M. B. (1992). Course of depressive symptoms over follow-up. Findings from the National Institute of Mental Health Treatment of depression collaborative research program. *Archives of General Psychiatry, 49,* 782–789.

Shiffman, S. (1989). Conceptual issues in the study of relapse. In M. Gossup (Ed.), *Relapse and addictive behavior* (pp. 149–179). New York: Tavistock/Routledge.

Tearnan, B. H., Telch, M. J., & Keefe, P. (1984). Etiology and onset of agoraphobia: a critical review. *Comprehensive Psychiatry, 25,* 511–562.

Teasdale, J. D., Segal, Z., & Williams, J. M. G. (1995). How does cognitive therapy prevent depressive relapse and why should attentional control (mindfullness) training help? *Behaviour Research and Therapy, 33,* 25–39.

Telch, M. J., Tearnan, B. H., & Taylor, C. B. (1983). Antidepressant medication in the treatment of agoraphobia: A critical review. *Behavior Research and Therapy, 21,* 505–527.

Thase, M. E. (1990). Relapse and recovery in unipolar depression: short-term and long-term approaches. *Journal of Clinical Psychiatry, 6,* 51–57.

Thompson, L. W., & Gallagher, D. (1984). Efficacy of psychotherapy in the treatment of late life depression. *Behaviour Research and Therapy, 6,* 127–139.

Turner, R. M. (1987). The effects of personality disorder on the outcome of social anxiety reduction. *Journal of Personality Disorders, 1,* 136–146.

Weissman, M. M. (1994). Psychotherapy in the maintenance treatment of depression. *British Journal of Psychiatry, 165*(26), 42–50.

Weissman, M. M., Klerman, G. L., Pruscoff, B. A., Shalomskas, D., & Padian, N., (1981). Depressed outpatients: Results one year after treatment with drugs and or interpersonal psychotherapy. *Archives of General Psychiatry, 38,* 51–55.

Westling, B. E., & Ost, L. G. (1995). Cognitive bias in panic disorder patients and changes after cognitive behavioral treatments. *Behaviour Research and Therapy, 33,* 585–588.

Whisman, M. A. (1990). The efficacy of booster maintenance sessions in behavior therapy: Review and methodological critique. *Clinical Psychology Review, 10,* 155–170.

Wilson, P. H. (1992a). Relapse prevention: conceptual and methodological issues. In P. H. Wilson (Ed.), *Principles and practice of relapse prevention* (pp. 1–22). New York: Guilford Press.

Wilson, P. H. (1992b). Depression. In P. H. Wilson (Ed.), *Principles and practice of relapse prevention* (pp. 128–156). New York: Guilford Press.

Wilson, G. T., Rossiter, E., Kleifield, E. I., & Lindholm, L. (1986). Cognitive behavioral treatment of bulimia nervosa: A controlled evaluation. *Behaviour Research and Therapy, 24,* 277–288.

Wolpe, J., & Lazurus, A. A. (1966). *Behavior therapy techniques*. New York: Pergamon.

Zitrin, C. M., Klein, D. F., & Woerner, M. G. (1980). Treatment of agoraphobia with group exposure in vivo and Imipramine. *Archives of General Psychiatry, 37,* 63–72.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.12

Use of Self-help Books in the Practice of Clinical Psychology

MICHAEL V. PANTALON
Yale University School of Medicine, New Haven, CT, USA

6.12.1 INTRODUCTION 265

6.12.2 CATEGORIZATION OF SELF-HELP BOOKS 266

- *6.12.2.1 Standardized Therapy Manuals* 266
- *6.12.2.2 General Self-help Books* 267
- *6.12.2.3 Problem-focused Self-help Books* 267
- *6.12.2.4 Technique-focused Self-help Books* 267
- *6.12.2.5 Internet Self-help Books* 267
- *6.12.2.6 Educational Self-help Books* 267
- *6.12.2.7 Other Self-help Books* 268

6.12.3 OVERALL EFFECTIVENESS OF SELF-HELP BOOKS 268

6.12.4 EFFECTIVENESS OF EACH TYPE OF SELF-HELP BOOK 269

- *6.12.4.1 General Self-help Books* 269
- *6.12.4.2 Problem-focused Self-help Books* 269
- *6.12.4.3 Technique-focused Self-help Books* 270
- *6.12.4.4 Internet Self-help Books* 270

6.12.5 RECOMMENDATIONS FOR FUTURE RESEARCH 271

6.12.6 INTEGRATING SELF-HELP BOOKS INTO THE PROCESS OF PSYCHOTHERAPY 271

- *6.12.6.1 Selection* 272
- *6.12.6.2 When to Use Self-help Books* 272
- *6.12.6.3 Introducing and Assigning Self-help Books* 273
- *6.12.6.4 Addressing Nonadherence* 273

6.12.7 HOW SELF-HELP BOOKS CAN IMPROVE PSYCHOTHERAPY OUTCOMES 274

6.12.8 POTENTIAL RISKS AND LIMITATIONS OF SELF-HELP BOOKS 274

6.12.9 SUMMARY AND CONCLUSIONS 275

6.12.10 REFERENCES 275

6.12.1 INTRODUCTION

Review articles and meta-analytic studies have suggested that the use of self-help books (SHBs) is an effective adjunctive treatment in the practice of clinical psychology (Gould & Clum, 1993; Marrs, 1995; Pantalon, Lubetkin, & Fishman, 1995; Pardeck, 1990; Scogin, Bynum, Stephens, & Calhoun, 1990; Starker, 1986; 1988; 1990). In this chapter, SHBs will be broadly defined as the use of bibliotherapy and other written, psychotherapeutic materials that are usually read independent of a health professional. It seems likely that this effectiveness is due to the great appeal and availability of SHBs to the general public, including those who read them as study participants—more than 2000 SHBs are published annually (American Psychological Association, 1989)—as well as the study of SHBs which are based on empirically supported psychological interventions (see Gould & Clum, 1993). Whether based on their great appeal, effectiveness, or other reasons, SHBs are widely prescribed by psychotherapists to their patients. One estimate of the number of therapists who have ever assigned a SHB to a client is 60% (Starker, 1988), and the number of therapists who prescribe SHBs on a regular basis ranges from 55% to 88% (Marx, Gyorky, Royalty, & Stern, 1991; Starker, 1986, respectively). In addition, favorable therapist and client ratings of the helpfulness of a variety of SHBs has been documented (Starker, 1986, 1988).

Though the popularity of SHBs may have given rise to the empirical evaluations of their effectiveness, relatively few studies have been published on the critical factors which bear on their selection and appropriate use. Because information on these factors is unavailable, it is unclear whether the general public or the relatively large number of therapists who are assigning SHBs to their clients are selecting the most appropriate and/or effective SHBs. Another question regarding SHBs is whether their use is closely monitored by the prescribing therapist or whether they are merely suggested to clients, and not integrated into the course of therapy. Despite these and other unanswered questions, this author believes that SHBs have the potential to provide effective treatments, as well as to increase client involvement in the process of therapy and perhaps, improve its outcomes. These ideas are timely given the continued popularity of SHBs and the increase in managed care which sometimes severely limits the availability of psychotherapy sessions to its consumers. This author believes research on SHBs could be undertaken in a more systematic manner if the various types of SHBs available to the public were better defined and categorized, and if the degree to which readers follow the recommendations given in SHBs is assessed. Such improvements would allow empirical investigations of each type of book separately and comparisons among different SHBs, as well as information regarding the best way to use SHBs. This chapter will address these and other issues regarding the use of SHBs in the practice of clinical psychology.

There has been a long-standing controversy surrounding the use of SHBs. Rosen (1987) stated that SHBs "commercialize" (p. 46) psychotherapy by making overgeneralized statements about psychological problems and their causes, and by promising unrealistic results. He conjectured that this commercialization is mainly due to the market pressures to sell such books. Rosen believes that an SHB simply cannot present psychotherapeutic techniques in an adequate manner, and therefore unlikely to be sufficiently effective. Further, because SHBs may not accurately diagnose and because readers usually do not adequately comply with the recommendations therein, Rosen states that SHBs should not be used as a substitute for psychotherapy (Rosen, 1987). Although there have not been calls to have SHBs replace psychotherapy, this author agrees that the high degree of enthusiasm over some SHBs may have led to an overstatement of their effectiveness. Other authors disagree with Rosen's claims and suggest that an individual's ability to profit from a wide variety of SHBs is great (Mahoney, 1988; Starker, 1988). This author agrees with the statement that SHBs can be very helpful, but identifying the various conditions under which this can occur, and for whom, is crucial to advancing our knowledge in this area. This chapter will also attempt to address such issues.

In the following sections, this author will first categorize the various types of SHBs available today. Second, the available empirical studies on the effectiveness of SHBs, when used alone and with therapist contact, will be reviewed. Third, the great appeal of SHBs will be discussed along with the manner in which this appeal can be used to increase the understanding of and adherence to psychotherapeutic interventions, and can therefore enhance clinical results. Finally, this chapter will include a description of an approach for integrating SHBs into the course of psychotherapy, which could be applied to various forms of psychotherapy. Specific recommendations for selecting, assigning, and using SHBs in the context of a clinical practice will be offered, along with identifying both the client and therapist variables that appear to be crucial in the effective use of SHBs.

6.12.2 CATEGORIZATION OF SELF-HELP BOOKS

6.12.2.1 Standardized Therapy Manuals

Standard therapy manuals (STMs) describe a standardized treatment protocol to be used in conjunction with psychotherapy. They are therapist-directed and usually in an outline format. Both therapist and client versions are usually available. In essence, they are used as therapy guides. They are more technical than easy-reading, which differs from what one might generally find in a local bookstore in the self-help section. Marsha Linehan's (1993) treatment manual for borderline personality disorder is an example of an STM. Each chapter of Linehan's manual for dialectical behavior therapy illustrates in great detail the goals, methods, and objectives which will be used in treatment, similarly to what one would find in therapy sessions. It is considered a type of SHB because the client manual specifies tasks that clients can do on their own (e.g., homework). Other examples of STMs are *Mastery of your anxiety and panic* (Barlow & Craske, 1989) and *Problem drinkers: Guided self-change treatment* (Sobell & Sobell, 1993).

6.12.2.2 General Self-help Books

These do not address specific disorders, but rather general emotional, relationship, or life development issues without delineating specific techniques and homework exercises within a structured protocol. They are usually self-administered, but can also be therapist-assisted. Examples of such SHBs are *The road less traveled* (Peck, 1978) and *Notes to myself* (Prather, 1970). General SHBs can be used by therapists from a wide variety of orientations because they do not espouse any particular brand of therapy over another. Instead, they can be used to motivate clients to address problems which they may have avoided for some time or they can help clarify the goals and objectives of psychotherapy.

6.12.2.3 Problem-focused Self-help Books

These address a particular disorder (some actually assist the reader in diagnosing him- or herself) and delineate specific techniques and homework exercises within a structured protocol. They are usually self-administered, but can also be therapist-assisted. Examples of this type of SHB are *Stop obsessing* (Foa & Wilson, 1991), *Control your depression* (Lewinsohn, Munoz, Youngren, & Zeiss, 1992), *Coping with depression* (Blackburn, 1987), and *The feeling good handbook* (Burns, 1989). Although these are all examples of cognitive-behavior therapy (CBT) books, problem-focused SHBs are also written from a psychodynamic perspective (Bass & Davis, 1997). Problem-focused SHBs can be used by therapists with a wide variety of therapeutic orientations (e.g., behavioral, psychodynamic, interpersonal) for a wide variety of problems. However, it is advisable that they be used by therapists and clients who are prepared to systematically apply problem-solving and other techniques to a specific problem or symptom.

6.12.2.4 Technique-focused Self-help Books

These are similar to problem-focused SHBs, but instead focus on a specific technique, rather than a disorder, that could be used for several or many problems. Unlike the previous SHBs, these are usually therapist-assisted, because therapists are usually the ones to recommend these books to their clients. However, they can also be self-administered. Examples of SHBs are *The anxiety and phobia workbook* (Bourne, 1990) and *The relaxation response* (Benson, 1975). As above, these books can be used by therapists from a wide variety of orientations.

6.12.2.5 Internet Self-help Books

These are defined as any of the preceding SHBs that are found on the Internet. This category challenges the traditional notion of the genre that self-help is an individual helping him- or herself with the use of materials written by another or several individuals. Once this material is formatted for the Internet, however, the opinions, advice, and counsel of potentially thousands of other, mostly lay, but some professional, people can be accessed via chatrooms, message boards, and e-mail. The advice obtained from such sources can be quite personalized, because of the Internet user's ability to e-mail or post very specific questions about their condition and/or treatment. However, for the purposes of this chapter, it shall be considered self-help as long as individuals are not seeking direct consultation with a psychotherapist. Some of the examples of SHBs listed above are actually on the Internet (albeit in severely condensed form), but many Internet SHBs are not in the form of a book *per se*, but are rather in the form of a homepage with a myriad of options for further information or directions to other resources depending upon a reader's specific interests. These "books" are usually self-administered or with the assistance of "online therapists." The reader is referred to Dow, Kearns, & Thornton (1996) for a list self-help sites.

6.12.2.6 Educational Self-help Books

These attempt to help the patient, spouse, family member, or significant other learn more about the problems the identified patient is experiencing, and their prevalence, treatment, and course, as well as ways in which to effectively cope with the patient. This type of book is often helpful because it can shorten the educational process, allow patients and their families to become collaborators rather than passive recipients of treatment, and enhance adherence to treatment recommendations (Mueser, 1994). A number of such books have been written for many of the psychotic, affective, and anxiety disorders, such as *Surviving mental illness: Stress, coping, and adaptation* (Hatfield & Lefley, 1993) which pertains largely to schizophrenia, *When once is not enough: Help for obsessive-compulsives* (Steketee & White, 1990), or *Panic disorder: The facts* (Rachman & de Silva, 1996). A lengthier list of other organizations, SHBs, and resources for educating both patients and their families about psychological disorders has been prepared by Mueser (1994).

6.12.2.7 Other Self-help Books

Finally, there are some SHBs which do not fall into one of the above categories. These are books that are not explicitly intended as self-help, but rather are books (e.g., novels, biographies) which, for example, chronicle a well-known individual's life or attempts to cope with a particular problem which is in some way instructive, such as *Darkness visible* by William Styron (1990), Pulitzer prize-winning fiction writer. Most of these SHBs are self-administered, but when they are used with the assistance of a therapist, it is usually within the framework of psychodynamic psychotherapy (Pardeck, 1991). However, their use in other types of psychotherapy, as with general SHBs, is not contraindicated, because they usually serve to motivate efforts to change and give clients a lexicon with which to better describe their complaints. Unfortunately, there are no controlled studies, known to this author, testing the efficacy of this type of SHB. In the next section, the effectiveness of SHBs in general is reviewed, followed by a summary of the research on each of the above SHB categories.

6.12.3 OVERALL EFFECTIVENESS OF SELF-HELP BOOKS

Although the treatments found in many SHBs are usually not empirically supported, there are some significant exceptions and the number is increasing (Blackburn, 1987; Burns, 1989; Clum, 1990; Ellis & Tafrate, 1997; Foa & Wilson, 1991; Lewinsohn et al., 1992). However, even though there are some books which are based on empirically validated treatments, many of the SHBs themselves are not evaluated directly, though once again there are notable exceptions (e.g., Burns, 1989; Clum, 1990; Lewinsohn et al., 1992). For example, Foa & Wilson's (1991) book is based on a good deal of methodologically sound research study demonstrating the effectiveness of the techniques therein (e.g., graduated exposure with response prevention). However, the book itself (i.e., how the treatment works when a client reads about it and decides on his or her own to what degree he or she adheres to the recommendations) has not been empirically tested in controlled trials. However, others have been empirically tested in this manner, such as Burns (1989). Because an SHB presents psychotherapeutic techniques in a unique manner, it can have a very different impact on a client as compared with the impact of seeing a therapist, even if the therapist and the author of the SHB are presenting the very same techniques. This is true, in part, because rationales for the use of therapeutic interventions may differ significantly from those presented in an SHB, and because different rationales can lead to very different perceptions about treatment and outcome (Newman, 1994). For some individuals, the therapist could be more motivating because he or she represents another human being expressing understanding of the client's problems, but for others an SHB could be more effective because it is less personally threatening.

Gould and Clum's (1993) meta-analysis of 40 self-help studies resulted in a measurable overall effect size for self-help interventions of 0.76 at post-treatment and 0.53 at follow-up, which is comparable to those reported for psychotherapy (Stiles, Shapiro, & Elliot, 1986). This effect was strongest when SHBs were compared to no treatment conditions (e.g., wait-list and delayed-treatment controls) and when behavioral observations were used as the dependent measure (e.g., behavioral avoidance test). The lowest effect sizes were found when SHBs were compared to placebos (e.g., attention control conditions) and when physiological instruments (e.g., heart rate) were used as the dependent measure. Another meta-analysis, conducted by Marrs (1995), found a positive, though somewhat smaller effect size for SHBs (0.57), but no significant differences between SHBs and therapist-assisted treatments. In both of these more recent meta-analyses, SHBs for assertion training, anxiety, and sexual dysfunction were more effective than those for weight loss, impulse control, and studying problems.

Additionally, Gould and Clum found that the mean effect size for pure self-help conditions (with no therapist contact at all) was greater than that of the minimal therapist contact conditions (where therapists periodically helped patients with the materials). However, the mean effect size for self-help conditions was less than that of the therapist-assisted conditions. The mean effect size for self-help conditions (SHB assigned with no therapist contact) was also less than that of the combination of self-help and therapy (assignment of an SHB plus periodic therapist supervision). An important difference is that SHB plus therapy was a more intensive integration of therapy and self-help regimens than the minimal contact conditions. Obviously there is a need for more research in this area so that a definitive statement about usefulness of SHBs and their combination with psychotherapy can be made. Interestingly, in the Marrs (1995) study, the amount of contact the therapist made with participants in SHB-only conditions was not significantly correlated with outcome. However, it appeared that participants whose primary goal was weight loss or anxiety reduction improved significantly more with increased therapist contact.

Thus, it appears that there are mixed results when comparing self-help vs. therapist-assisted conditions. One hypothesis, put forward by Gould and Clum (1993) is that the therapist-assisted conditions may not adequately resemble how a therapist would actually assist a client with an SHB in therapy and therefore may not have enhanced the SHBs effectiveness; it may have even diminished it to some degree. One factor that may have influenced these results is the fact that only 15 of the 40 studies analyzed by Gould and Clum utilized a clinical sample and it is unclear whether SHBs would be as effective with patients who consistently met specific diagnostic criteria.

Although Marrs (1995), based on a sample of nine studies, found that there was no significant difference between SHBs and therapist-directed treatments, a definitive statement about whether SHBs are more or less effective than therapy cannot be made because of the small number of studies that compare active treatment without an SHB to (i) self-help in general (only nine were included in Marrs' study), and (ii) pure self-help conditions (no therapist contact).

6.12.4 EFFECTIVENESS OF EACH TYPE OF SELF-HELP BOOK

6.12.4.1 General Self-help Books

Although general SHBs sometimes focus on a particular area of an individual's life, the exact nature of the problem is ill-defined (e.g., low self-esteem, failed relationship), and they do not assist a reader in ascertaining a proper diagnosis. Such books also do not offer specific techniques and homework exercises within a structured protocol. For these reasons, these books are simply more difficult to assess in terms of effectiveness. They are often based on the clinical intuition and nonempirically based theories of the author. Another reason why they are difficult to investigate in research is that the advice described within them often appears to be overgeneralized and superficial. Despite the above shortcomings of general SHBs, their popularity is staggering. However, this popularity has not prompted much empirical investigation, and thus, very little data are available regarding the impact of these books (Starker, 1990). Therefore, it would be difficult at this stage to even recommend how they could be integrated into the context of psychotherapy or for what problems. However, it appears that such books (e.g., Peck, 1978) could be effective in helping a patient clarify exactly what thoughts, feelings, and/or behaviors are most distressing. They could also assist the therapist and client in setting goals and objectives at the outset of therapy, especially if the client is unclear about how he or she hopes to benefit from the experience. Such books have been described as being helpful in treating physically and/or sexually abused or neglected children (Pardeck, 1991), though only one controlled trial has been conducted (Ogles, Lambert, & Craig, 1991), which demonstrated that a general SHB worked equally well as behaviorally and stage theory-based SHB for coping with divorce. General SHBs seem to be utilized mostly by therapists who are experientially, Gestalt, or psychodynamically oriented therapists (Pardeck, 1991). This may be because these SHBs focus on events (current and/or historical) which are thought to be the cause of the problem behaviors, rather than on specific procedures to ameliorate them.

6.12.4.2 Problem-focused Self-help Books

There are a number of other studies that demonstrate the effectiveness of problem-focused SHBs for particular disorders. Regarding depression, it appears that reading a cognitive or behavioral SHB is more effective than delayed-treatment groups (Scogin et al., 1990; Scogin, Jamison, & Gochneaur, 1989; Smith, Floyd, Scogin, & Jamison, 1997; Wollersheim & Wilson, 1991), equally effective as group CBT (Wollersheim & Wilson, 1991), and that cognitive and behavioral books were also equally effective (Scogin et al., 1989). Treatment gains have reported to be maintained or enhanced as long as three years beyond the end of self-help treatment (Smith et al., 1997). Additionally, Mahalik & Kivlighan (1988) reported that participants who are high on measures of independence, need for structure, internal locus of control, and self-efficacy, had responded better to a self-help approach to depression, than those who were low on these attributes.

Regarding panic disorder, Ghosh and Marks (1987) found that self-exposure instructions worked equally well to reduce anxiety when given by a psychiatrist, SHB (*Coping with panic* [Clum, 1990]), or computer program, and that each resulted in substantial improvement up to six-month follow-up. Gould, Clum, and Shapiro (1993) similarly found that the effectiveness of an SHB was not significantly different from that of individual guided imaginal coping. Both treatments were significantly more effective than a wait-list control group. In a follow-up to the above study, Gould and Clum (1995) concluded that the SHB *Coping with panic* (Clum, 1990) was significantly more effective than a wait-list control group at post-treatment and at two-month follow-up. Lidren et al. (1994) also report that *Coping with panic* is as effective as group therapy based on the same book, which was required reading, and more effective than a wait-list control.

Ogles et al. (1991) report findings from one of the few studies to compare problem-focused books based on different therapeutic orientations, namely behavior therapy and stage theory for the treatment of reactions to divorce. These books were also compared to a general SHB, which focused broadly on coping with life events (not only divorce). It was found that all three were effective in reducing a variety of symptoms associated with loss, but that they all worked equally well. In addition, those who reported high expectations of being helped fared better than those who did not, as evidenced by greater reductions in symptom scores, and attributed the change largely to their reading of the books.

Other behaviors, such as problematic drinking (Heather, Robertson, MacPherson, Allsop, & Fulton, 1987; Hester & Delaney, 1997; Miller & Taylor, 1980; Sobell & Sobell, 1993) and smoking cessation (Glasgow, Schafer, & O'Neill, 1981) have also been shown to be modifiable with the use of SHBs. Heather et al. (1987), in particular, demonstrated that a problem-focused book on controlled drinking was more effective than an educational book regarding alcohol problems. In addition, many of these studies report maintenance of gains at six months, and one, two, and three years. Hester and Delaney (1997) and Gould and Clum (1995) have also demonstrated that computerized SHBs are also effective.

In summary, the SHBs in this category which have some level of empirical support have been based predominantly on CBT. Thus, more research is required in other areas, as meta-analytic studies inform us that overall, most psychotherapies are equally effective (Stiles et al., 1986). SHBs, based on other models of psychotherapy, could thus prove as helpful as those above. However, the books that take up the most space in the self-improvement section of a local bookstore (e.g., SHBs on dieting, relationships, parenting, hypnosis), or have been found to sell the best (Starker, 1990) are not the ones that have been empirically validated directly, nor are they based on treatments shown to be effective in research studies (Marrs, 1995). Conversely, some SHBs that are based on empirical data do not sell nearly as well as those that are not based on such information, perhaps because claims regarding change and its maintenance are more realistic in the former (Mahoney, 1988). However, as consumers of health care become more sophisticated in their knowledge of available treatments, which this author believes to be a by-product of managed care, the demands to publish SHBs with some empirical support vs. the need to make such books appealing to the general public may become less competing.

6.12.4.3 Technique-focused Self-help Books

A number of very clearly written workbooks on specific techniques for coping with various problems (e.g. tension headaches, depression, anxiety) have been written. Some popular examples are *The anxiety and phobia workbook* (Bourne, 1990) and *The relaxation response* (Benson, 1975), and *Mind over mood* (Greenberger & Padesky, 1995). Although some of the techniques therein have been empirically demonstrated to be effective, especially those in Benson (1975), the use of the books by clients on their own or with minimal therapist contact has not. Perhaps this is partially due to the fact that such books seem most appropriate for use as a supplement to the training of clients in various techniques (e.g., progressive muscle relaxation). Oftentimes, they do not offer an entire treatment plan for addressing all the various aspects of a problem (e.g., behavioral, cognitive, interpersonal), as does a problem-focused book, and may therefore be difficult to assign and test for a particular disorder. However, it is conceivable that their effectiveness in reducing some circumscribed aspect of a disorder (general anxiety or feelings of hopelessness), or in treating subclinical samples, could be empirically tested.

6.12.4.4 Internet Self-help Books

At the time of writing, there are no controlled trials on Internet SHBs. As described earlier, Internet SHBs could greatly expand the number and type of SHBs to be tested. The Internet also adds a new medium to test. For example, the difference between the effectiveness of an SHB that is read in the traditional manner vs. one that is read on the World Wide Web could be illuminating. Internet books have the potential to make SHBs significantly more personalized via their potential to cross-reference a greater amount of information than a traditional SHB (see Dow et al., 1996).

6.12.5 RECOMMENDATIONS FOR FUTURE RESEARCH

It is recommended that statements regarding the effectiveness of SHBs should be made based on a careful investigation of specific types of books separately with more attention paid to the characteristics of individuals who can benefit from them and the experimenters/therapists who utilize them. More controlled trials comparing the effectiveness of the various types of SHBs, especially those in the general, technique, Internet, and other categories would be helpful toward this end. Research on the characteristics of participants, such as reading and grade level, expectations of change, and motivation, should be addressed (Marrs, 1995). The issue of subject selection is also important. Perhaps individuals who typically read SHBs differ substantially from those who do so merely to be in a study for which they will receive some evaluation and/or financial remuneration. Investigations of the difference in outcomes between individuals who are permitted to read and apply the precepts of the book on their own vs. those who have periodic reviews with the experimenter, could be illuminating. Perhaps the use of diary recordings of progress and issues during reading would obviate the need for such frequent contacts with the experimenter. Such frequent contacts may contaminate the findings of self-help research in two important ways: (i) the contact with the therapist, rather than with the SHB, may account for some, if not all, of the improvement noted; and (ii) similarly demand characteristics, rather than the SHB, may elicit some portion of the improvement noted. However, research which compares the effectiveness of psychotherapy, with and without SHBs is not available at this time.

Therapist characteristics which bear on the effectiveness of SHBs, which pertain mostly to the manner in which they are utilized and monitored in treatment (e.g., prescribed vs. suggested, monitored closely or informally), should also be addressed in future research so that practitioners may have guidance on how to make good use of such books. In addition, investigating what aspects of SHBs (e.g., instillation of hope via case examples vs. techniques) account for the most change in clients' behavior, would also be helpful. Both of these issues will be discussed in a later section of this chapter.

In summary, although a carefully selected and closely read SHB, without any assistance from a therapist, may be good enough for some, the available data do not speak to the characteristics of such individuals. Thus, a more conservative approach, one consisting of an SHB plus therapy, may be warranted at this stage, given that the data are not conclusive and the fact that only a small number of SHBs have been empirically validated for use alone. Another reason why this approach might be more advantageous is that readers/clients would have guidance in how to select an SHB and how to adequately adhere to its recommendations. Poor selection and nonadherence are among the most deleterious results of an unstructured approach to self-help (Rosen, 1987). Therefore, because of the tentative nature of the research on SHBs, this author believes that psychotherapy, supplemented by judicious use of SHBs, is a safe and useful alternative to the pure self-help approach. Perhaps future research will explicate further under what conditions this may be the case and under what conditions a pure self-help approach may be indicated. Despite the fact that the prescription of self-help books and other materials is widespread, little research regarding various methods for assigning SHBs to clients or incorporating them into the course of therapy has not been undertaken. Future research should address this and the aforementioned questions as it suggests the possibility of using SHBs to enhance psychotherapy outcomes. It is hoped that the following guidelines for integrating SHBs in the practice of clinical psychology, which have been helpful in this author's work with clients, could facilitate such research.

6.12.6 INTEGRATING SELF-HELP BOOKS INTO THE PROCESS OF PSYCHOTHERAPY

Many studies testing the effectiveness of SHBs do not clearly state how the books are to be used. Some authors, however, have addressed related issues, such as the prescription of SHB (Pardeck, 1990; Starker, 1990); which clients benefit most from a self-help approach (Mahalik & Kivlighan, 1988); and bibliographies of SHBs (Quackenbush, 1991). However, only Pantalon et al. (1995) and Pardeck (1991) have detailed a model of fully integrating SHBs into the course of psychotherapy. A brief overview of the former approach is given below.

6.12.6.1 Selection

Before a therapist can adequately help select the SHB that will be best for a client, it is helpful for him or her to be aware of the most frequently cited positive characteristics of such books. These are: (i) optimistic frame, (ii) encouragement, (iii) advice to seek professional help, (iv) general self-understanding, and (v) understanding about specific problems (Halliday, 1991). This author believes that, in addition to the above, one of the most valued aspects of SHBs is the social validation offered through their suggestion that problems are understandable given life experiences, and that the readers are not to blame. Hence, fear of embarrassment or evaluation by a therapist are nonexistent. Also, instead of a complex diagnosis, the reader is given simple, easy-to-understand explanations for why they are the way they are, via accounts of cases which often mirror their own personal experiences. They also raise motivation to change through examples of others who have succeeded. They typically offer an optimistic outlook on problems and a multitude of suggestions for better coping. In addition, they are perceived by readers to be more efficient than therapy, especially in terms of time, cost, and effort (Starker, 1986).

This author believes that SHBs will be able to offer all of the above, if they are goal-directed and written in a clear and user-friendly format (e.g., tables, figures, charts). They should also offer the reader a multitude of therapeutic exercises with rationales and encouragement, and descriptions of methods to measure noticeable changes (e.g., charts, checklists). In an effort to promote self-acceptance, it should prepare readers for setbacks or failure. In this author's opinion, however, the two most important things an SHB should offer a reader are research-based information and a plethora of case examples of those treated successfully, so that readers can identify with the case examples in a book and be convinced that they have selected an SHB that is appropriate to their needs. Both have also been demonstrated in research to significantly impact on expectancies for therapeutic change (Kazdin & Kranse, 1983). The reader is referred to Pantalon et al. (1995) and Santrock, Minnett, and Campbell (1994) for bibliographies of SHBs recommended for use in the practice of clinical psychology. Seligman's (1994) *What you can change and what you can't* is strongly recommended, as it can help clients, and therapists of any orientation develop realistic expectations for behavior change.

6.12.6.2 When to Use Self-help Books

The second step in integrating SHBs into the course of therapy involves deciding when and why to use them. When clients are experiencing difficulty understanding the nature of the problems they are facing, when the direction therapy should take is somewhat unclear, when clients either resist or have difficulty prioritizing goals, are only some examples of when SHBs can be used. A general rule of thumb, however, is that the use of SHBs could be effective at a point when either or all of these three situations arise:

(i) the client is asking many questions about his or her symptoms or the process of therapy ("How is this going to help me?" or "How do you know this is right for me?");

(ii) the client is resistant in some way and/or does not adhere to therapeutic exercises (e.g., "I can't understand why you want me to do this homework?"); or

(iii) the client appears to be of the belief that he or she is the only person that is attempting to cope with such problems, which may accompany a sense of helplessness (e.g., "No one else has such major problems; I just can't do anything about it").

In these situations, an SHB could bolster the therapist's explanation of a problem, clarify the process of treatment, offer motivation and rationales for doing therapeutic exercises, normalize symptoms, and give hope.

How ready a client is to engage in actual strategies to change behavior may affect how effectively he or she utilizes the advice in a particular type of SHB. For example, if a client is still thinking about whether or not he or she has a problem in the first place or whether he or she believes anything can be done about it, or whether the costs of change are too great, he or she may not be motivated to read a book which offers specific advice regarding how to change behavior directly. Instead they may profit from a book or a section of a book that discusses the pros and cons of their behavior first (e.g., general SHB or the "diagnosis" section of a problem-focused book). This then impacts on what type of book the theorist should assign to the client. One way to assess this is by using one of the various stages of change assessments available which have been shown to correlate with actual attempts to change and which are useful in planning clinical strategies (Prochaska, DiClemente, & Norcross, 1992).

6.12.6.3 Introducing and Assigning Self-help Books

The rationale for asking a client to use an SHB in conjunction with therapy should be planned after the therapist has (i) decided on the target problem; (ii) conceptualized his or her reason for seeking additional materials (e.g., client is asking many questions about the process of therapy); (iii) formulated his or her criteria for selecting an SHB; and (iv) selected an SHB, as well as, the introductory selections he would like the client to read. It may be best not to suggest exercises other than reading for the first two sessions after introducing an SHB, even if the client has already practiced such exercises prior to adding the book. Once all four conditions have been met the therapist can introduce the SHB he or she has selected following the suggestions found in the next several sections.

When introducing the idea of using an SHB to a client, begin with a review of what gains the client has made thus far and how. It is then explained that if these efforts were extended, during nontherapy days and with new exercises, such as with the help of an SHB, greater improvements could be achieved. The client is then asked for his or her thoughts on this. Subsequently, the author shares some of his book selection criteria with the clients and points out the ways that the SHBs generally used by this author are different from other titles (e.g., research-based), many of which their friends may adore. If a client appears motivated to integrate an SHB into the work of therapy, then this author informs the client about the appropriate and reasonable expectations for what can be gained from reading an SHB. The degrees of behavior change and change as a process is emphasized. The client is told that the learning does not necessarily end by the time he or she finishes reading the book, but that it continues and the SHB becomes a reference, which the reader can go back to time and time again. This author then goes through the table of contents of the book with the client, remarking on how each of the sections (preselected by the therapist) is relevant to the client's particular problems and treatment. This exercise is sometimes quite illuminating because it makes it clear how therapist and client conceptualizations of the primary problems or methods of addressing them differ. Therefore, the therapist must be flexible at this early stage and willing to compromise on certain selections or exercises that the client may or may not see as relevant at first. This process of introducing and assigning an SHB also usually helps clients develop their own criteria for future self-help selections made on their own.

This author believes that it is almost always better to prescribe an SHB rather than give the patient a choice of reading it or not. The more confident, directive, and positive a therapist is when assigning an SHB, the better the chance the SHB has of making a therapeutic impact on the client. Therapists should also display high regard for the selected SHB. Therefore, SHBs, and the exercises therein, are almost always assigned as homework, which is monitored by the therapist on a regular basis.

6.12.6.4 Addressing Nonadherence

We have found that it is usually more effective to integrate SHBs into the context of therapy rather than to simply assign them. If they are not an integral part of the therapy, the patients may devalue such books, behavioral homework assignments, and even the therapy itself. If they are not good or relevant enough to use systematically, their use should be seriously questioned. This is true even if a therapist is only using one section of a particular book. However, some patients simply will not comply with such prescriptions as described above and will require more in the way of motivation to actively change problematic behaviors (e.g., a change in type of SHB). Readers of SHBs can mistakenly believe that reading a book, but failing to follow through on specific assignments, constitutes being actively engaged in the changing behavior.

Due to the above, the following suggestions are made for dealing with nonadherence when using an SHB:

(i) self-monitoring, which allows clients to become more aware of their resistant behavior, because a lack of such awareness may be stopping a client from acknowledging the severity of their problem;

(ii) shaping of behavior, which involves positively reinforcing (e.g., giving praise, feedback, or acknowledgment for) each incremental effort toward more compliant behavior;

(iii) agreeing on relevant sections. Frequent reevaluation of therapy goals and objectives facilitates this, and allows for a pace that is comfortable for both;

(iv) assessing therapist demands regarding change. If a therapist is vigilant about monitoring his or her irrational expectations, then the chances of making unrealistic demands of the client, and hence of noncompliance, can be minimized;

(v) anticipating nonadherence, which gives the therapist an early opportunity to plan strategies for addressing this issue. However, it also allows the therapist to view such resistant behavior as part of the process of change, rather than as an annoyance or interference. With this attitude, the therapist can respond to nonadherence in a nondefensive, reflective manner, which can be critical in motivating resistant clients; and

(vi) traditional CBT techniques (e.g., desensitization, cognitive restructuring).

6.12.7 HOW SELF-HELP BOOKS CAN IMPROVE PSYCHOTHERAPY OUTCOMES

Therapists may be able to enhance their effectiveness by adding SHBs to their armamentarium. In general, the manner in which SHBs could do this is by offering to therapists a new medium through which they can motivate behavior change. Specifically, SHBs can:

(i) increase the number and variety of rationales given for treatment interventions, via case examples and interesting anecdotes in understandable language;

(ii) increase opportunities to self-monitor problem behaviors as well as to rehearse adaptive ones, through the many therapeutic exercises they offer;

(iii) enhance adherence to homework by offering the client motivational accounts of other clients and breaking down complex tasks into their component parts;

(iv) increase self-efficacy via a heightened sense that they themselves (rather than the therapist) are solving their problems; and

(v) improve generalization to other settings, because their suggestions are read and implemented in a wide variety of situations outside of the therapy office.

6.12.8 POTENTIAL RISKS AND LIMITATIONS OF SELF-HELP BOOKS

Rosen (1987) suggests that the potential risks associated with the use of SHBs are associated mainly with the misdiagnosis of problems, lack of adherence with procedures, overgeneralized advice, and overstatement of the effects of SHBs. However, this author believes that these risks are associated with the *misuse* of SHBs, rather than by the books themselves. Specifically, SHBs may not help produce positive behavior change when the following conditions exist:

(i) when expectations of change are unrealistic;

(ii) when specific directions are not followed;

(iii) when an irrelevant or inappropriate book is selected (e.g., due to a misunderstanding of the various types of books available);

(iv) when the book selected is not written in terms the reader can understand; or

(v) when it does not match the goals of the reader.

Since the psychotherapy plus SHB approach, described above, addresses each of these conditions, it should help the therapist avoid the potential risks associated with SHBs. Unfortunately, diligent attempts on the part of the therapist to ensure that each of these conditions is met does not always guarantee good outcome with an SHB. One of the most important aspects of using an SHB, as with any problem-focused therapy, is to achieve a good balance between efforts toward change and efforts toward acceptance (Hayes, Jacobson, Follette, & Dougher, 1994; Linehan, 1993). Thus, therapists attempting to integrate an SHB into the course of therapy should remember that at certain points he or she should either put the SHB (e.g., change effort) aside (with the mutual agreement that this is only a temporary change) or change the type of SHB that is being used (e.g., change from a problem-focused to a general SHB, which promotes more self-understanding or acceptance).

One of the fundamental limitations of psychotherapy manuals in general, and SHBs in particular, is the assumption that the same disorder manifested by different people, with different learning histories and contingencies that maintain the symptoms, are to be treated in the same manner. Fishman (1981) refers to this as mistaking "topographical equivalence" for "functional equivalence" (p. 244). In order to treat an individual effectively, the therapist may need to functionally assess each of the client's problematic behaviors so as to fully understand the reinforcement contingencies and how they relate to each problem area. Assuming that the same behavior exhibited by two individuals, compulsions related to unwanted, intrusive thoughts (i.e., topographical equivalence) serves the same function in both (i.e., functional equivalence), such as avoidance of intimate relationships vs. a desire to always be in control of one's thoughts, may lead the therapist to implement minimally helpful interventions. Thus, when using an SHB in therapy, it would be advisable for the therapist to individualize its procedures, either by timing or choosing the interventions or target behaviors in such a way that they are reasonable given the client's entire clinical picture (e.g., symptoms, motivation, social support, family life, occupational functioning, strengths and weaknesses). For example, an empirically validated SHB for generalized anxiety disorder may begin the course of treatment with progressive muscle relaxation, which for some may be anxiety-provoking, rather than relaxing. For such cases, reframing, or more assessment may be indicated. Without a functional analysis of each of this client's problem areas, the therapist might have incorrectly assumed relaxation training to be the best intervention with which to start.

In fact, many of the issues that impinge on the use and effectiveness of manualized treatments in therapy also affect the effectiveness of SHBs. For example, the degree to which therapists individualize such manuals affects clinical outcome, and there is great controversy regarding whether modifications based on clinical intuition or actuarial judgment leads to greater efficacy of the treatment (Wilson, 1997). How a therapist individualizes an SHB for use by a particular client with a unique learning history may impact on its effectiveness in the same way that it does with treatment manuals. However, if future research attempts to experimentally control for differences in administration or individualization, more could be learned about the most productive way to do this. One of the more basic questions would involve a comparison of individualization vs. no individualization, or individualization with and without formal rules, as has been done with standardized treatment manuals. The results of such studies are mixed, but appear to favor actuarial judgment, which would suggest less individualization of standardized books. However, it appears that even those who favor the actuarial side of this debate, suggest that tailoring a treatment manual to reduce disruptive behaviors, and enhance adherence and regard for the interventions (much in the way described above), is both appropriate and improves the effectiveness of the treatment (Eifert, Schulte, Zvolensky, Lejuez, & Lau, 1997; Linehan 1993).

A final purpose SHBs could serve is to further researchers efforts at wide dissemination of research findings regarding empirically supported psychotherapeutiic techniques (Persons, 1997). If SHBs based on empirically validated treatments became as popular to the general public as other books, clinicians who would otherwise be unaware of such treatments might have an incentive to learn and use them (i.e., if individuals in therapy were aware of such approaches, it is conceivable that they would seek such treatment from their therapist).

6.12.9 SUMMARY AND CONCLUSIONS

This chapter reviewed (i) the various types of SHBs available today; (ii) the available empirical studies on the effectiveness of SHBs; (iii) the possible reasons for the appeal of SHBs to the general public and, more recently, to researchers; (iv) an approach for integrating SHBs into the course of psychotherapy, with specific recommendations for selecting, assigning, and using SHBs in the context of a clinical practice; (v) the manner in which SHBs could enhance (a) clinical results, especially via a better understanding of and adherence to psychotherapeutic interventions, generalization, and increased efficiency, as well as (b) dissemination of empirically supported treatments to practitioners and clients alike; and finally (vi) the potential hazards associated with the misuse of SHBs. It was proposed that if therapists used SHBs in the manner outlined, SHBs could be an effective adjunctive treatment in psychotherapy. It is the belief of this author that, with the above issues clarified, future research into the effectiveness of SHBs, including its potential positive impact on psychotherapy outcome, may be more systematically addressed.

ACKNOWLEDGMENT

The author wishes to express his gratitude to Wendy Bobadilla and Beatrice Martineau for their assistance in the preparation of this chapter.

6.12.10 REFERENCES

American Psychological Association (1989). *First annual golden fleece awards for do-it-yourself therapies*. Presentation at the annual meeting of the American Psychological Association, New Orleans, LA.

Barlow, D. H., & Craske, M. G. (1989). *Mastery of your anxiety and panic*. Albany, NY: Graywind.

Bass, E., & Davis, L. (1997). *Courage to heal: A guide for women survivors of childhood sexual abuse*. New York: Harper Perennial.

Benson, H. (1975). *The relaxation response*. New York: Avon.

Blackburn, I. M. (1987). *Coping with depression*. Edinburgh, UK: Chambers.

Bourne, E. J. (1990). *The anxiety and phobia workbook*. Oakland, CA: New Harbinger.

Burns, D. D. (1989). *The feeling good handbook*. New York: William Morrow.

Clum, G. A. (1990). *Coping with panic*. Pacific Grove, CA: Brooks/Cole.

Dow, M. G., Kearns, W., & Thornton, D. H. (1996). The Internet II: Future effects on cognitive behavioral practice. *Cognitive & Behavioral Practice, 3,* 137–157.

Eifert, G. H., Schulte, D., Zvolensky, M. J., Lejuez, C. W., & Lau, A. W. (1997). Manualized behavior therapy: Merits and challenges. *Behavior Therapy, 28,* 499–509.

Ellis, A. E., & Tafrate, R. C. (1997). *How to control anger before it controls you*. Secaucus, NJ: Carol Publishing Group.

Fishman, S. T. (1981). Narrowing the generalization gap in clinical research. *Behavioral Assessment, 3,* 243–248.

Foa, E., & Wilson, R. (1991). *Stop obsessing: How to overcome your obsessions and compulsions*. New York: Bantam.

Ghosh, A., & Marks, I. M. (1987). Self-treatment of agoraphobia by exposure. *Behavior Therapy, 18,* 3–16.

Glasgow, R. E., Schafer, L., & O'Neill, H. K. (1981). Self-help books and amount of therapist contact in smoking cessation programs. *Journal of Consulting & Clinical Psychology, 49,* 659–667.

Gould, R. A., & Clum, G. A. (1993). A meta-analysis of self-help treatment approaches. *Clinical Psychology Review, 13,* 169–186.

Gould, R. A., & Clum, G. A. (1995). Self-help plus minimal therapist contact in the treatment of panic disorder: A replication and extension. *Behavior Therapy, 26,* 533–546.

Gould, R. A., Clum, G. A., & Shapiro, D. (1993). The use of bibliotherapy in the treatment of panic: A preliminary investigation. *Behavior Therapy, 24,* 241–252.

Greenberger, D., & Padesky, C. A. (1995). *Mind over mood: Change how you feel by changing the way you think.* New York: Guilford.

Halliday, G. (1991). Psychological self-help books: How dangerous are they? *Psychotherapy, 28,* 678–680.

Hatfield, A. B., & Lefley, H. P. (1993). *Surviving mental illness: Stress, coping, and adaptation.* New York: Guilford.

Hayes, S. C., Jacobson, N. S., Follette, V. M., & Dougher, M. J. (Eds.) (1994). Acceptance and change: Content and context in psychotherapy. Reno, NV: Context Press.

Heather, N., Robertson, I., MacPherson, B., Allsop, S., & Fulton, A. (1987). Effectiveness of a controlled drinking self-help manual: One-year follow-up results. *British Journal of Clinical Psychology, 26,* 279–287.

Hester, R. K., & Delaney, H. D. (1997). Behavioral self-control program for windows: Results of a controlled clinical trial. *Journal of Consulting & Clinical Psychology, 65,* 686–693.

Kazdin, A. E., & Krause, R. (1983). The impact of variations in treatment rationales on expectancies for therapeutic change. *Behavior Therapy, 14,* 657–671.

Lewinsohn, P. M., Munoz, R. F., Youngren, M. A., & Zeiss, A. (1992). *Control your depression.* New York: Simon & Schuster.

Lidren, D. M., Watkins, P. L., Gould, R. A., Clum, G. A., Asterino, M. & Tulloch, H. L. (1994). A comparison of bibliotherapy and group therapy in the treatment of panic disorder. *Journal of Consulting & Clinical Psychology, 62,* 865–869.

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder.* New York: Guilford.

Mahalik, J. R., & Kivlighan, Jr., D. M. (1988). Self-help treatment for depression: Who succeeeds? *Journal of Counseling Psychology, 35,* 237–242.

Mahoney, M. (1988). Beyond the self-help polemics. *American Psychologist, 43,* 598–599.

Marrs, R. W. (1995). A meta-analysis of bibliotherapy studies. *American Journal of Community Psychology, 23,* 843–870.

Marx, J. A., Gyorky, Z. K., Royalty, G. M., & Stern, T. E. (1991). Use of self-help books in psychotherapy. *Professional Psychology: Research and Practice, 23,* 300–305.

Miller, W. R., & Taylor, C. A. (1980). Relative effectiveness of bibliotherapy, individual and group self-control training in the treatment of problem drinkers. *Addictive Behaviors, 5,* 13–24.

Mueser, K. T., & Glynn, S. M. (1995). *Behavioral family therapy for psychiatric disorders.* Boston: Allyn & Bacon.

Newman, F. (1994). Understanding client resistance: Methods for enhancing motivation to change. *Cognitive & Behavioral Practice, 1,* 47–70.

Ogles, B. M., Lambert, M. J., & Craig, D. E. (1991). Comparison of self-help books for coping with loss: Expectations and attributions. *Journal of Counseling Psychology, 38,* 387–393.

Pantalon, M. V., Lubetkin, B. S., & Fishman, S. T. (1995). Use and effectiveness of self-help books in the practice of cognitive and behavioral therapy. *Cognitive & Behavioral Practice, 2,* 213–228.

Pardeck, J. T. (1991). Using books in clinical practice. *Psychotherapy in Private Practice, 9,* 105–115.

Peck, M. S. (1978). *The road less traveled.* New York: Simon & Schuster.

Persons, J. B. (1997). Dissemination of effective methods: Behavior therapy's next challenge. *Behavior Therapy, 28,* 465–471.

Prather, H. (1970). *Notes to myself: My struggle to become a person.* Moab, UT: Real People Press.

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist, 47,* 1102–1114.

Quakenbush, R. L. (1991). The prescription of self-help books by psychologists: A bibliography of selected bibliotherapy resources. *Psychotherapy, 28,* 671–677.

Rachman, S., & de Silva, P. (1996). *Panic disorder: The facts.* Oxford, UK: Oxford University Press.

Rosen, G. M. (1987). Self-help treatment books and the commercialization of psychotherapy. *American Psychologist, 42,* 46–51.

Santrock, J. W., Minnett, A. M., & Campbell, B. D. (1994). *The authoritative guide to self-help books.* New York: Guilford.

Scogin, F., Bynum, J., Stephens, G., & Calhoun, S. (1990). Efficacy of self-administered treatment programs: Meta-analytic review. *Professional Psychology, 21,* 42–47.

Scogin, F., Jamison, C., & Davis, N. (1990). A two-year follow-up of the effects of bibliotherapy for depressed older adults: *Journal of Consulting & Clinical Psychology, 58,* 665–667.

Scogin, F., Jamison, C., & Gochneaur, K. (1989). Comparative efficacy of cognitive and behavioral bibliotherapy for mildly and moderately depressed older adults. *Journal of Consulting & Clinical Psychology, 57,* 403–407.

Seligman, M. E. P. (1994). *What you can change and what you can't: The complete guide to successful self-improvement.* New York: Alfred A. Knopf.

Smith, N. M., Floyd, M. R., Scogin, F., & Jamison, C. S. (1997). Three-year followup of bibliotherapy for depression. *Journal of Consulting & Clinical Psychology, 65,* 324–327.

Sobell, M. B., & Sobell, L. C. (1993). *Problem drinkers: Guided self-change treatment.* New York: Guilford.

Starker, S. (1986). Promises and prescriptions: Self-help books in mental health and medicine. *American Journal of Health Promotion, 1,* 19–24, 68.

Starker, S. (1988). Do-it-yourself therapy: The prescription of self-help books by psychologists. *Psychotherapy, 25,* 142–146.

Starker, S. (1990). Self-help books: Ubiquitous agents of health care. *Medical Psychotherapy, 3,* 187–194.

Steketee, G., & White, K. (1990). *When once is not enough: Help for obsessive-compulsives.* Oakland, CA: New Harbinger.

Stiles, W. B., Shapiro, D. A., & Elliot, R. (1986). Are all psychotherapies equivalent? *American Psychologist, 41,* 165–180.

Styron, W. (1990). *Darkness visible: A memoir of madness.* New York: Random House.

Wilson, G. T. (1997). Treatment manuals in clinical practice. *Behavior Research & Therapy, 35,* 205–210.

Wollersheim, J. P., & Wilson, G. L. (1991). Group treatment of unipolar depression: A comparison of coping, supportive, bibliotherapy, and delayed treatment groups. *Professional Psychology: Research and Practice, 22,* 496–502.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.13
Preventive Goals and Indirect/Consultation Strategies: Meeting Current Needs Through a Recommitment to Underused Means and Ends

RAYMOND P. LORION
Ohio University, Athens, OH, USA

6.13.1 INTRODUCTION 277
6.13.2 ANTECEDENTS TO PREVENTIVE CONSULTATION/INDIRECT STRATEGIES 280
6.13.3 BEFORE THE BEGINNING; RETRACING PAST STEPS 282
6.13.4 THE COMMUNITY MENTAL HEALTH MOVEMENT: A LOST OPPORTUNITY 283
6.13.4.1 Shifting One's Conceptual Frame 284
6.13.4.2 Shifting Emphasis from Inside to Outside 287
6.13.5 EXTERNALIZING THE STUDY OF BEHAVIOR 288
6.13.6 ENVIRONMENTAL MODELS OF BEHAVIOR 289
6.13.7 TRANSACTIONAL MODELS OF BEHAVIOR 290
6.13.7.1 Linking Inside and Outside 291
6.13.8 SHIFTING TO A PREVENTION FOCUS 293
6.13.9 MERGING INDIRECT/CONSULTATION STRATEGIES WITH PREVENTION GOALS 294
6.13.10 FINAL COMMENT 297
6.13.11 REFERENCES 297

6.13.1 INTRODUCTION

Chapters encouraging the adoption of alternative approaches to service delivery typically combine an overview of intervention strategies with a critical review of evidence for or against each strategy. Such an approach reflects several related assumptions. First, information will increase the acceptance and delivery of the proposed alternative. Second, existing resistance to application of these services reflects a lack of information. Third, information about these has been developed in spite of resistance. Fourth, information developed for strategies which have typically been limited to highly controlled use will be accepted as relevant to the decision to incorporate these approaches into existing service delivery strategies. Finally, with sufficient evidence, the existing service delivery strategies may be replaced by the alternative approaches.

An alternative perspective is that understanding resistance to changing current services and professional definitions is an essential first step to achieving such changes. The transition of mental health from direct to indirect services and from a clinical treatment model to a public health prevention model has been understandably slow and difficult. This chapter reflects the belief that such transitions are necessarily and appropriately slow and that professional resistance is best understood as natural and well-intended rather than as selfish and mean-spirited. Rather than look for enemies to progress, one must understand its impediments and plan accordingly. Such is the focus of this chapter which reviews the history, conceptual, and practical obstacles to the widespread acceptance, development, and implementation of indirect/consultation and preventive approaches by the mental health professions generally and clinical psychology specifically. Rather than attempt to present a compendium of strategies of questionable scientific merit and pragmatically uncertain dissemination, I will examine the circumstances under which progress had developed thus far and propose avenues for accelerating the pace at which further development occurs.

Admittedly, the chapter focuses on the development of these strategies in the US. Systematic descriptions of how these strategies have evolved internationally (e.g., in South America, Australia, and Europe) are being prepared for publication in 1999 and 2000 in the *Journal of Community Psychology*.

In his book, *The creation of settings and the future societies*, Sarason (1972) posits that to understand a human service one must examine the circumstances under which it is developed and delivered. One must particularly appreciate those pre-existing conditions which Sarason (1972, p. 24) labeled as occurring during the period "before the beginning." Change as reflected in an "innovative" service, for example, does not arise without antecedents! Rather, it represents a sequence of events (some of which can be identified and some of which cannot) leading to recognition of the need for change and of the form(s) which that change must take to be accepted and incorporated within the surrounding service structure. As explained in this chapter, such is the case for the growing recognition of preventive interventions as necessary components of a comprehensive mental health and health care strategy. Seemingly, such was also the case for the unfortunate de-emphasis of indirect/consultation strategies following their most recent introduction. At the same time, currently existing and expected conditions are likely to alter the perceived importance of these very strategies.

Originally focused solely on "indirect/consultation strategies," the chapter took an unexpected and fortuitous turn. The invitation to prepare the chapter for this set of volumes came as somewhat of a surprise to the author. Although involved with the development, implementation, and evaluation of community-based mental health interventions for nearly a quarter century, I had not associated my work on the prevention of emotional and behavioral disorders with "indirect/consultative approaches." On reflection, the link is as obvious as it is appropriate. Preventive interventions, by definition, focus on antecedents and concomitants of dysfunction rather than on dysfunction itself. As such, they both represent and rely heavily on indirect means of responding to mental health needs. Conceptually, their influence depends on understanding the nature and dynamic processes among individual and situational factors which are salient to pathogenesis (Lorion, Price, & Eaton, 1989). Pragmatically, the intervention's target is not the pathology but its antecedents, that is, enabling circumstances and related individual diatheses.

A second reason for the surprise was that contemporary interest in consultation and indirect strategies was quite unexpected given the nation's seeming abandonment of both the goals and means of the community mental health movement born in the 1960s. The few remaining vestiges of what was to be a "bold new approach" for responding to the nation's mental health needs seemed doomed in the current environment of managed care, welfare reform, and the search for the origins of emotional and behavioral impairment through neuroscience, genetic mapping, and research on biopsychosocial models of development (e.g., Bronfenbrenner, 1979). Four decades after beginning with lofty goals and expectations, the community mental health center system appears but a skeleton of its former self. Whether even the skeleton will remain in the next decade seems uncertain.

Ironically, what may remain of the "bold new approach" is the seemingly least emphasized and least valued of the five original mandatory services, that is, consultation and education. Again, on reflection, the timeliness of these approaches is understandable. If emotional and behavioral problems are to be economically and efficiently addressed, they must be confronted early and on as many fronts as possible. Consultation and indirect strategies provide potentially cost-effective avenues for encircling situational and systemic elements of dysfunction and thereby the potential for accelerating its amelioration. Less recognized, however, is that these strategies might also contribute to a most timely goal, that is, the actual avoidance of dysfunction. In effect, these strategies represent routes to the prevention of mental disorder and to the promotion of optimal development and improved quality of life.

The latter point highlights a link which has heretofore been largely ignored but which is particularly salient to this chapter. The importance of indirect/consultation approaches was acknowledged nearly four decades ago in President Kennedy's landmark message to the nation calling for a "bold new approach" for responding to mental disorders:

> ... We must seek out the causes of mental illness and mental retardation and eradicate them. Here, more than in any other area, "an ounce of prevention is worth more than a pound or cure." For prevention is far more desirable for all concerned. It is far more economical and far more likely to be successful. Prevention will require both selected specific programs directed especially at known causes, and the general strengthening of our fundamental community, social welfare, and educational programs which can do much to eliminate or correct the harsh environmental conditions which often are associated with mental retardation and mental illness. (Kennedy, 1963, reported in Bloom, 1977, p. 264)

To pursue this goal, "consultation services to other community agencies and mental health information and education" (Bloom, 1977, p. 267) were included as mandatory components of community mental health services. The link between indirect/consultation strategies and preventive interventions was reaffirmed when the editors of this work extended the chapter's focus to include all three strategies.

Confronting the responsibility for designing and delivering a new array of services (i.e., partial hospitalization, emergency, consultation, and education) to heretofore underserved segments (e.g., rural, low-income, or minority) of the population in unfamiliar settings (e.g., in sparsely populated villages, inner-city ghettoes, or working-class neighborhoods), community mental health service providers initially focused on established and evident need rather than incipient or potential disorder (Bloom, 1975; Cowen, 1973). As clinicians they were trained to recognize and treat disorder. Interest in doing these very activities, as professionals, determined, in part at least, their original career choice. Understandably, therefore, their initial forays into consultation and education efforts reflected contemporary assumptions about the nature of disorder and established intervention practices arising from such assumptions.

The implementation of the remaining and predominant mandatory elements of the community mental health movement (i.e., outpatient and inpatient services) focused primarily on pathology as figure and the expert/therapist/provider-serving-the-learner/client/recipient-exchange as ground. Within this framework, consultation and indirect/educational services were designed substantively and procedurally to maximize the recognition of disorder, the efficiency and timeliness of referral for treatment, and the appropriate utilization of community-based resources for supporting the effects of intervention. As explained below, social legislation of the 1960s and 1970s presented significant opportunities for the mental health system and mental health professions to change their conceptions of problems and solutions. Few (e.g., Levine & Perkins, 1997; Mrazek & Haggerty, 1994; Sarason, 1981) believe that the opportunities were taken or their potential fulfilled. Current and anticipated revisions of health and mental health care policies present another chance for substantive change. Whether the outcome will differ remains, for now, uncertain. What is certain, however, is that prior efforts at change were sincere but insufficiently cognizant of how **established** practices, perspectives, and assumptions about the nature of problems and solutions shaped the likelihood of true innovation. As we enter yet another era in which change is seemingly pursued, we must reflect on how current stances will necessarily shape our future.

What seems quite certain is that the potential for change which arose in the 1960s remains unfulfilled! "Community" in the community mental health movement, for example, referred primarily to the recipient and the location of services rather than to the rationale for and procedures of those services (Bloom, 1984). For this and other reasons, the heuristic and programmatic importance of the link between prevention as goal and consultation/indirect strategies as means lost its immediacy. Also clouded was appreciation of the vast array of alternatives which would have become available had the focus been on health and its maintenance rather than on pathology and its treatment or control. Ideally, this chapter will focus attention on the potential of those links and result in a broadening of perspectives as contemporary health and mental health service providers design intervention systems for the future. Toward that goal, it is important to note that the community mental health movement represented a return to rather than discovery of the value of indirect/consultation strategies. That recognition occurred decades before! In fact, the link between indirect strategies and prevention was unquestioned in the early twentieth century!

6.13.2 ANTECEDENTS TO PREVENTIVE CONSULTATION/INDIRECT STRATEGIES

Legislative initiatives have understandably been assigned as much if not more of the credit than scientific findings and professional interest for moving to center stage the mental health disciplines' recognition of and involvement with prevention and indirect/consultation strategies. Attention to indirect/consultation strategies, for example, is assumed to have resulted from passage of the Community Mental Health Center Act (Public Law 88-164) of 1963. Legislative antecedents to that Act included the National Mental Health Act (Public Law 79-487) of 1946. This law was written to facilitate the incorporation of public health principles into the design and delivery of mental health services. It also authorized the establishment of the National Institute of Mental Health which included the prevention of mental disorder among its stated purposes. More than a decade would pass, however, before that goal was publically acknowledged. Nearly two more decades would pass before substantive action was taken to achieve this goal (Goldston, 1986). Nearly 50 years after passage of Public Law 79-487, preventive intervention research remains in the early stages of development and the products of its efforts, although increasing in number and documented value, account for a very small portion of the nation's mental health service effort (Durlak & Wells, 1997; Mrazek & Haggerty, 1994).

Further legislative actions fueled the potential for momentum in establishing and disseminating indirect/consultation strategies. The Mental Health Study Act (Public Law 84-182) of 1955 created the Joint Commission on Mental Illness and Health charged with conducting a thorough and objective study of the human and economic costs of the nation's mental health needs, services, and service delivery systems. The Joint Commission took this assignment seriously. Over five years, it conducted comprehensive studies of topics ranging from the adequacy of staffing resources (Albee, 1959) to public attitudes toward mental health needs and services (Gurin, Veroff, & Feld, 1960) to epidemiological analyses of mental illnesses. In addition to publication of numerous monographs reporting the results of individual studies, the Joint Commission (1961) submitted an integrative summary of its findings and recommendations in its final report entitled *Action for mental health*. That report provided the substantive foundation for the Community Mental Health Centers Act of 1963.

The Joint Commission's report was supported by two significant epidemiological studies (Hollingshead & Redlich, 1958; Srole, Langner, Michael, Opler, & Rennie, 1962) published during that period. Whether measured in terms of need or demand for services, a serious imbalance was evident between the availability of mental health services and the prevalence of mental disorder in adults. A subsequent epidemiological study confirmed the existence of at least an equivalent imbalance between children's needs and available services (Glidewell & Swallow, 1969). These and related studies provided what appeared to be convincing support of "Albee's (1959) pessimistic assessment that sufficient mental health staffing resources would *never* be available to treat those in need within the dominant individual service delivery models.

Other sources of personnel (e.g., paraprofessionals, Durlak, 1979, and "natural care givers," Cowen, 1982) and other models of service (e.g., empowerment strategies, Rappaport, 1977) if developed and supported offered routes to redressing these imbalances. The founders of the emerging field of community psychology and their followers devoted considerable effort to pursuing such alternatives (Cowen, Gardner, & Zax, 1967; Iscoe, Bloom, & Spielberger, 1977). Although overlapping in terms of their goals and objectives, community psychology (e.g., Heller, Price, Reinharz, Riger, & Wandersman, 1984; Rappaport, 1977) and community mental health (Bloom, 1984) took somewhat differing routes to finding alternative solutions to resolving the needs/services imbalance. Each, however, emphasized the development and dissemination of preventive interventions and indirect/consultation strategies as outcomes to be achieved. Each spoke of the value of expanding definitions of providers of service, of the timing of services, and of the nature of services (Riessman, 1965). Both community psychology and community mental health, however, retained a strong identification with clinical psychology. With notably few exceptions (Iscoe et al., 1977; Phares & Trull, 1997), doctoral programs in these emerging fields devoted far more time to training in traditional theories of pathology and traditional methods of diagnosis and treatment than in those uniquely distinct theories and methods focused on communities. Understandably, professional identification was more likely to default to clinical than community psychology.

Simultaneously, prevention's advocates reminded policymakers of the public health maxim that no disease or disorder had ever been controlled by treatment but only by prevention. As early as 1962, the American Public Health Association identified six categories of mental disorder which, because their etiology was known, were preventable (Bloom, 1984). These categories included diseases caused by: (i) poisons (e.g., brain syndromes following exposure to solvents); (ii) infections (e.g., rubella); (iii) genetic disorders (e.g., TAY-SACHS disease); (iv) nutritional deficiencies (e.g., pellagra); (v) injuries (e.g., traumatic head injuries); and (vi) general systemic disease (e.g., prematurity). It is important, however, to note that early in the development of the community mental health movement, prevention represented not simply an idea but also an attainable goal. Moreover, guidelines for shifting to alternative approaches were provided by Caplan in his influential works, *Principles of preventive psychiatry* (1964) and *The theory and practice of mental health consultation* (1970). Riessman (1965) among others (e.g., Fairweather, Sanders, Maynard, & Cressler, 1969) described innovative ways in which self-help and environmental change, alone or in combination, offered relief which could precede and potentially avoid dysfunction and exacerbation of existing problems. These efforts did not depend on professional resources to operate and, under appropriate circumstances, may have become self-sustaining.

Yet the solution to a problem is determined in large part by how it is defined and how it is presumed to have arisen. As noted, the movement toward community-based service delivery occurred at a time that conceptions of health and especially pathology were focused primarily on *internal individual* states (Sarason, 1981). Changes in the individual's affective response to stress, behavioral response to external stimuli, and the capacity to alter those responses were assumed to occur through interventions targeted to individuals. From that perspective, a reasonable response to Albee's (1959) prediction of chronic shortages in mental health professionals capable of delivering individual services was to resolve that shortage by increasing the nation's capacity to train such professionals. Concurrent with legislation to incorporate public health concepts into mental health theory and practice and to establish the community mental health movement, Congress funded programs to train substantial numbers of mental health professionals to provide traditional diagnostic and therapeutic services (Phares & Trull, 1997). The inclusion of reimbursement for mental health services in health care policies and the growing public acceptance of outpatient mental health services led, in psychology, to a proliferation of university and nonuniversity-based programs for health service providers (Phares & Trull, 1997). The shortages in mental health professionals capable of serving individual needs predicted by Albee (1959) appears to have been substantially resolved, if not replaced by a surplus (Routh, 1998).

The surplus becomes undeniable if subdoctoral providers in psychology, counselor education, marital and family practice, counselor education, nursing, and social work are included in service delivery resources.

Ignored in this solution were the alternatives suggested by Albee (1959) and echoed in the Community Mental Centers Act, that is, to expand the range of personnel available to provide services and to expand the range and nature of those services. Indirect/consultation efforts targeted at other than the detection of those in need of referral for treatment or the enhancement of nonmental health service providers (e.g., teachers, social service case workers) to support mental health interventions could not be justified within the prevailing understandings of pathogenesis, dysfunction, and treatment. Similarly, efforts to establish the feasibility of responding to mental health needs using providers who were indigenous to the communities in which the problems arose (e.g., Sobey, 1970) may have demonstrated their effectiveness (e.g., Dawes, 1994; Durlak, 1979) but the weight of evidence was insufficient to topple the individual focus and internal emphases dominant at the time.

Being from the neighborhood, for example, could only be of advantage to a service provider if characteristics of the neighborhood were assumed to contribute to the etiology, maintenance, and resolution of emotional and behavioral dysfunction. If one assumed otherwise, one would reasonably question the therapeutic benefit of attracting and retaining heretofore untreated segments of the population if the insights needed for symptom removal could not be achieved in the presence of an "unqualified therapist." Those who challenged the innovativeness of the community movement pointed to its seeming lack of theory and focus. Levine and Perkins (1997) describe reactions to the community mental health movement as follows:

> The new thrust went off in all directions at once, with little coherence and little conceptual clarity.

> Critics committed to traditional medical model practice looked askance at social activism. Those committed to "intrapsychic supremacy " (Levine, 1969)—the belief that problems in living result from people's internal psychological structures, which in turn dictate perceptions, feelings, and actions in everyday situations—viewed the activists as misguided romantics who had foolishly strayed from proper professional roles and activities. Community-oriented critics of traditional practice were equally firm in their convictions but had little to offer by way of alternate conceptualizations. (pp. 59–60)

Alternative perspectives had yet another chance to establish themselves as a viable element of federal health care policies. Shortly after his election, President Carter created the President's Commission on Mental Health to review the state of the nation's needs and response capacity. In essence, the Commission's numerous task forces and panels revisited the issues raised by President Kennedy's group and recommended continued and vigorous pursuit of many of its most innovative elements. The Commission's findings were translated into the Mental Health Systems Act (Public Law 96-398). Included within this legislation was the requirement that the National Institute of Mental Health establish an administrative unit focused on funding, evaluating, and disseminating preventive interventions. The prevention of emotional and behavioral disorders was to become a central and adequately supported element of the nation's mental health policy (Bloom, 1984). Significant changes in service and training priorities were expected along with consequent shifts in the balance of direct and indirect approaches to serving the nation's mental health needs. Just as suddenly, the Mental Health Systems Act was repealed! Newly elected President Reagan's initial budget shifted to the states responsibility for deciding which services to provide with the federal government's newly designed block-grant funding approach to underwriting the costs of such services. If they were to be funded, the preventive and indirect services needed to compete for limited funds which were primarily used to support the provision of direct services to those with established diagnosed needs. Unfortunately, the promises of benefits to be gained from support for preventive interventions significantly exceeded the evidence of such benefits (Lorion et al., 1989; Mrazek & Haggerty, 1994) and left the field open to significant challenge (Albee, 1986). Ironically, an appreciation of the history of mental health strategies would have revealed that justification for these alternative services could be found in the efforts of mental health providers decades earlier.

6.13.3 BEFORE THE BEGINNING; RETRACING PAST STEPS

Examination of the scientific and legislative accomplishments reviewed thus far might lead one to conclude erroneously that prevention and indirect/consultation strategies were neither recognized nor attempted before the latter half of the twentieth century. The legislative initiatives and epidemiological studies just described, however, were not mental health's initial foray into indirect/consultation and preventive strategies. Human services generally and the mental health services specifically have not only applied such strategies throughout the twentieth century but periodically such strategies represented a dominant and explicit objective. Considered in their entirety, it appears quite possible that indirect/consultative approaches and even preventive interventions may have both a longer and a more successful history than interventions which provide direct services to individuals.

In their analysis of the history of consultation, for example, Mannino and Shore (1986) report that clinical psychology has been associated with such interventions since its inception. Long identified as a defining milestone of clinical psychology's formation as a profession, the establishment of Lightner Witmer's psychological services clinic at the University of Pennsylvania in 1896 is also quite relevant to this chapter. Reportedly, this clinic's services documented early professional appreciation of the value of indirect services and of both the desirability and feasibility of balancing attention to referred needs with creation of ways to avoid such needs (Mannino & Shore, 1986). Little known is the fact that Witmer's clinic relied on consultative interventions to deliver a substantial portion of its services. Comprehensive assessments of the nature and origins of children's cognitive, behavioral, and impulse control problems were, for example, as likely to result in interventions with the child's family or school as with the child. Frequently, analysis of the needs of a referred child included diagnostic assessments of instructional procedures. In turn, the intervention often involved prescriptive changes in school curriculum and detailed suggestions to teachers for handling disruptive behavior within the classroom setting.

Established during the initial decade of the twentieth century to respond to the needs of immigrants and low-income families, the Settlement House Movement hired teachers to visit the schools enrolling the children of families served by the Movement. These itinerant teachers would observe classroom procedures and recommend changes in curriculum, classroom organization, and instructional methods to improve children's educational performance. Such visits were intended both to reduce the needs of identified children and to enhance the overall effectiveness of the classroom. Without being labeled as such, these visits represent very early attempts to prevent emotional and behavioral problems and to promote optimal development and adaptive functioning.

In the 1920s, similar goals were pursued by the Child Guidance Movement. Healy's Juvenile Psychopathic Institute in Chicago, for example, developed and implemented programs to assist troubled youth by enhancing the capacity of relevant community agencies, including the schools, to serve the needs of referred and nonreferred youth. Healy's staff also worked to increase public understanding of mental health principles and public acceptance of emotional problems. The recipients of these educational programs were the communities in which the youth lived and the public agencies which served them. Child Guidance staff in other cities established service priorities for a community based on systematic assessments of local needs and resources. Findings from this planning step identified gaps in services and achievable objectives. Problem solutions often included programs to educate staff from relevant agencies about mental health principles and strategies.

Child Guidance staff also worked to improve a community's readiness to develop its own child guidance capacity. As described by Mannino and Shore (1986), such capacity development activities included the aforementioned local assessments as well as field services which carried out site visits, assisted existing service organizations to recognize the contributions of other human services, and identified ways for these organizations to network and collaborate. Through their activities, child guidance workers sought to make explicit the close relationship between the child and family as case and the community in which services were provided. Aware of the futility of attempting to respond individually to each case, child guidance workers often chose instead to alter the settings (e.g., home, school) in which the child lived and learned by influencing the decisions made by teachers, parents, and other caretakers. To effect such changes, caretakers were provided information about human development and child care methods. Child guidance workers explained these principles and provided direct models of alternative parenting or teaching behaviors. Child guidance workers recognized that their consultative and educational services should not be designed to make their recipients into mental health workers but rather to improve their functioning as parents or teachers. When the recipient was a service organization, the goal was not to change it into a mental health agency but to improve the developmental sensitivity, efficiency, and effectiveness of its intended services.

Reflecting on these approaches, it becomes apparent that attention to indirect/consultation strategies preceded by several decades the community mental health movement. Lessons reflected in these services included appreciation of the need to understand, value, and respect the perspectives and skills of a variety of service organizations. Movements developed at the turn of the twentieth century (e.g., Child Guidance, Settlement Houses, Mental Hygiene) appreciated the significant problems in adjustment confronting families attempting to cope with immigration, emigration, and poverty. Neither the families nor the organizations associated with these movements had adequate resources; both confronted challenging and unforgiving circumstances. Early on, the movements recognized that attempting to solve problems individually could not succeed. With their limited financial and staff resources, they focused on reducing the demanding quality of the environment, on enhancing the responsiveness of existing organizations, and the benefits of interorganizational collaboration and (long before the term was popular) on empowering parents, teachers, and citizens to respond to needs confronting them.

6.13.4 THE COMMUNITY MENTAL HEALTH MOVEMENT: A LOST OPPORTUNITY

As noted, following an intensive study of the nation's mental health needs and resources, the Joint Commission on Mental Health and Illness concluded that a "bold new approach" was needed if services were to respond to need (Joint Commission on Mental Health and Illness, 1961). The Joint Commission's report urged the nation to pursue alternatives to a mental health delivery system which was recognized as unavailable, inaccessible, and unacceptable to significant portions of the population (Joint Commission on Mental Health and Illness, 1961). Then as now, the need for responsiveness and relevance mandated the development of strategies which fit the actual circumstances within which people lived, learned, worked, and recreated. Then as now, the need to respond early rather than later was perceived as both economically responsible and humane. Then as now, responsibility for the emotional and behavioral health of individuals extended beyond the mental health professions. As reported by Gurin et al. (1960), physicians and the clergy were considerably more likely than psychologists, psychiatrists, or social workers to be approached for help with emotional problems. Subsequently, Cowen (1982) noted that hairdressers, bartenders, retired persons, and many others in the community served or could serve as valuable resources for those experiencing emotional discomfort.

What in the past may have been defined as "simply being a good neighbor" has the potential of being organized into a pervasive network of support for those suffering distress which can be tapped naturally without risking the personal and public stigma of mental health diagnosis and treatment. Many forms of indirect/consultation approaches are, in fact, designed to sensitize others to the responsibility and ease of "simply being a good neighbor" and to facilitate their capacity to fulfill it. Regrettably, these approaches received far too little attention to flourish during and since the 1960s. In fact, few "bold new approaches" were developed which redefined the timing and nature of mental health interventions. The question presented in this chapter is: "will the potential of indirect/consultation be recognized and exploited now and in the decades to come?"

One by one, whether labeled client, patient, or consultee, those presenting emotional and behavioral dysfunction have been approached by growing legions of mental health service providers. The sincere and somewhat effective efforts of these providers have not, however, stemmed the tide. As reported in Rouse (1995), the proportion of the nation's citizens suffering from emotional and behavioral (including substance abuse) disorders remains unacceptably high. Nor is there any basis for anticipating this tide to abate in the foreseeable future. Whatever their modality, direct approaches are limited in this regard by their focus on the individual and by the requirement that pathology must be present or imminent before intervention can take place.

In his pointedly titled book, *The absurd healer*, Dumont (1968) foresaw this outcome as the community mental health movement was just beginning. A decade later, I experienced Dumont's insightful moment as I entered the waiting room of a public clinic. I was then part of a team of mental health professionals excited by the opportunity to develop a short-term therapy program for indigent clients. We seemed to be a success! Clients were referred in substantial numbers; most cooperated with treatment protocols and participated in the 8–10 session program. We gloated to colleagues when an attempted comparison with long-term protocols could not be completed because few of the long-term clients attended the minimum of four sessions required for inclusion in the study. For a moment after I entered on this day, the waiting room full of therapy clients confirmed my certainty that our protocol represented a solution to the unacceptable levels of measured pathology in low-income populations (Lorion, 1973, 1974).

Suddenly, I realized that however successful we were, the room would never empty! Clients would come and go; their improvement ranging from considerable to none. Yet, without a substantive change in the system of care, others would always take the place of those who left. This prediction appeared valid whether long- or short-term interventions were provided and whether the recipients were individuals, groups, or families. To empty the waiting room, I and my colleagues would have to change our sense of when, where, and how we intervened. We would also have to shift our definition of the problem and of our responsibility for its solution from the waiting room to the circumstances and conditions which led to its need. We would, in effect, have to shift our existing theories of how problems arose and our theories of their solutions. Such shifts presented enormous challenges to a system seemingly committed to change. As noted, our profession has yet to commit wholeheartedly to confronting those challenges!

6.13.4.1 Shifting One's Conceptual Frame

With hindsight, it seems difficult to understand the mental health fields' reluctance to adopt wholeheartedly the flexibility and potential breadth of effect associated with applying indirect/consultation strategies in the pursuit of preventing mental illness and responding to emotional disorders. Passage of the Community Mental Health Center's Act provided the political support and financial backing seemingly necessary for shifting from an emphasis on treatment to understanding and ultimately controlling factors relevant to the development and maintenance of disorder. Bloom's (1977) analysis of this "bold new approach" makes evident that talented, dedicated, and hard working individuals were committed to its development. Soon after passage of this landmark legislation, the Swampscott Conference laid the foundation for the field of community psychology (Bennett et al., 1966). Subsequent organizational meetings of this group (e.g., Iscoe et al., 1977) gave evidence that this emerging field had momentum as reflected in budding programs of research, the development of programs at the master's and doctoral levels, and a growing corps of adherents anxious to enlist in this movement. Voicing a commitment to the integration of science and practice, this emerging field's established leaders led the way for successive generations to establish a critical mass of providers of indirect/consultation and preventive interventions (Iscoe et al., 1977).

Within a decade of its inception, community psychology appeared to have the commitment, momentum, and tools to move mental health toward acceptance of conceptions of etiology that extended beyond the individual and conceptions of intervention that redefined service delivery in partnership (Tyler, Pargament, & Gatz, 1983) rather than caretaker roles (e.g., Heller & Monahan, 1977; Levine & Perkins, 1997; Rappaport, 1977). After a decade of building a foundation for a national program of prevention work (Klein & Goldston, 1977), this perspective shift appeared to have taken root and become ready to blossom. For the next decade, leading advocates of prevention science, policy, and practice met regularly at the Vermont Conference on the Primary Prevention of Psychopathology (Kessler & Goldston, 1986). Funding for research on the prevention of emotional and addictive disorders began to rise with passage of the Community Mental Health Centers Act in 1963, a trend which has generally continued to the present (Mrazek & Haggerty, 1994).

Yet, three decades after passage of the Community Mental Health Centers Act and five decades after formation of the National Institute of Mental Health with its mandate to bring public health principles to mental health care, the mental health fields remain at best ambivalent toward and distant from prevention as a goal and indirect/consultation strategies as means to that goal. Proposing recommitment to that goal or attempting to convince readers that recent scientific findings argue convincingly that it lies within reach seem quite unlikely to have other than limited success. Rather, it seems important to examine the mental health field's seeming resistance to what all agree is a desirable outcome, that is, the avoidance of disorder and the promotion of positive mental health. Levine and Perkins (1997) raise a plausible explanation, that is, those proposing the shift from treatment to prevention and from direct to indirect/consultation services "had little to offer by way of alternate theoretical conceptualizations" (p. 60).

Other explanations seem rather less easily supported. Resistance to changing mental health's goals and methods must be considered historically. As reported in Mannino and Shore's (1986) review and echoed in Sarason's autobiography (1988), we have been there before! Attempts to develop and disseminate indirect strategies early in the twentieth century make evident that their limited survival cannot be attributed to a lack of talent in its advocates, a lack of social support for the resulting programs, or even a lack of evidence of their contribution to the welfare of those served. It is seemingly not about the creativity of those who design interventions, the importance of the research questions raised, the social problems addressed, or the explanatory power of the methods used to find answers (Sarason, 1996).

Having devoted much of his career to understanding the intractability of educational methods to genuine change, Sarason (1990, 1993) warns against attempts to understand the intransigence of existing approaches by finding and blaming those who oppose change. He notes:

> Our task would be incomparably easier if such villains existed. Being imprisoned in tradition, being resistant to and fearful of anything other than superficial change and window dressing, puzzled by the failure of past efforts, allergic to fads and fashions, disenchanted with quick fixes—such attitudes and reactions are not those of villains. (Sarason, 1993, p.13)

Subsequent comments about impediments to educational change seem equally applicable to understanding the pace at which indirect/consultation and preventive approaches have been adopted by the mental health fields:

> If everyone is in agreement that we are faced with a set of problems that we have to do something about, and if in the post-World War II era serious people have committed themselves to valiant efforts at reform, why have the results been either so modest, minuscule, or nonexistent? That was the question the book (Sarason, 1990) tried to illuminate. It was not about what to do; it was about how we have been thinking, what axioms we have unreflectively accepted as right, natural, and proper, and alternative ways we should begin think. (Sarason, 1993, p. 13)

Sarason explains that existing conceptions of how emotional and behavior disorder developed and needed to be remedied were reflected in priorities for research and interventions. Assuming that the seeds of pathology lay within the individual leads to the conclusion that changes in the individual are necessary to correct disorder. That perspective was dominant before, during, and after earlier attempts to shift from direct to indirect approaches and was reinforced by clinical psychology's acceptance of psychiatry and medicine as the definers of the nature of problems and solutions (Phares & Trull, 1997; Sarason, 1981, 1985).

Reluctance to shift to the public health perspective are understandable. Public health practitioners traditionally have viewed disease and disorder as the consequence of complex relationships between individual vulnerabilities (i.e., diatheses) and situational factors (i.e., stress) and of interventions (especially those seeking to prevent disorder) as requiring an understanding and alteration of aspects of the host (i.e., the individual presenting the disorder), the agent (i.e., the cause of the disorder), or of the environment (i.e., of the circumstances which bring the host in contact with the agent) (Lilienfeld & Lilienfeld, 1980). Ideally, at least a portion of the contributions of each element of this classic public health triad (i.e., host, agent, environment) will be understood in planning the preventive intervention. That understanding will identify those points along the etiological chain vulnerable to intervention and thus opportunities for avoiding or interrupting the pathogenic process.

The public health perspective challenges many core elements of clinical psychology's most influential theories, research methods, and concerns (e.g., to understand, predict, and control individual differences). Not surprisingly, clinicians tend to focus their research and services at the individual level. By contrast, public health interventions target populations or segments thereof and measure their impacts at the population rather than individual level. Of concern to proponents of indirect/consultation and preventive approaches is the seeming reluctance of psychology's community-oriented fields to emphasize public health rather than traditional clinical orientations. This individual orientation is confirmed in examinations of the content and methods of published community psychology research in its initial decades (Lounsbury, Cook, Leader, & Mears, 1985) and thereafter (Heller, 1989; Levine & Perkins, 1997). Community psychology's endorsement of the prevention of disorder and the development and adoption of interventions to change social policies and resolve social inequities as defining goals are slowly translating into substantial research and programmatic activity beyond the individual level of analysis or impact.

Echoing Sarason (1993), I would assert that the villains who are obstructing our pursuit of innovation and appreciation of indirect/consultation and preventive approaches are not "them"; contrary to Pogo, nor are they us! The impediment is the difficulty of leaving the traditions and methods of clinical training when such training is valued both within an existing professional culture and within the larger society. Sarason (1993) argues that "if we have learned anything with near certainty, both from the organizational and psychotherapeutic literature, it is that those who seek to change find themselves at some point resisting change" (p. 28). His perceptions about the difficulty of creating new forms of service settings seem to apply equally to the creation of alternative services. Resistance to change must be recognized and expected as part of the process of change which must be thoughtfully and sensitively countered rather than simply regretted or condemned. Regarding the introduction of innovative settings, Sarason appreciated that "if there was anything about which we were truly secure on the basis of past experience, it was that we were embarking on a venture that was very complicated, guaranteed to arouse anxiety and doubt, and, in all respects a venture that could bring out the best and worst in us" (Sarason, Zitnay, & Grossman, 1970, p. 71). Reflecting on the fate of indirect/consultation and preventive approaches thus far, one might apply the same lesson.

Moving from an individual-clinical to a public health perspective required abandoning what was known and valued by the profession and, admittedly, reimbursed by society. As noted, the design, development, and delivery of public health interventions ran (and continues to run) counter to many aspects of clinical training. Many indirect/consultation and preventive interventions are intended to preclude either the need for clinical intervention or for professional intervention. In either case, clinical involvement in mental health service delivery would be reduced. The fact that, once established, the majority of public health interventions would not even belong to us raises yet another impediment to shifting from clinical to public health perspectives. Few professionals readily accede to such generosity! More likely we are influenced by "professional preciousness," that is,

> the tendency of ... professionals to view their technical skills in very precious kinds of ways, that is, to overestimate the differences in skills among the professions and to underestimate the communalities ... . Each field does have a core of distinctiveness, and thus should be both recognized and treasured. It is one thing, however, to say that each field has a distinctive core of skills and it is quite another thing to say that *everything* a particular profession does is either distinctive or not, in part at least, learnable by the other professions in the settings in which they work together. (Sarason, Levine, Goldenberg, Cherlin, & Bennett, 1966, p. 587)

If this applies to our dealings with other professions, one can only assume it is not diminished when those with whom we are to share, if not turn over entirely, our clinical and intervention skills are nonprofessionals.

Consultants, after all, teach elements of their expertise to others so that those others can apply the elements. Health promotion and disease/disorder prevention involves us with segments of the population to which we cannot lay claim based on our expertise in the diagnosis and treatment of psychopathology. Without an accepted taxonomy of health, we cannot use a specialized lexicon to stake out that claim. Indirect strategies which require entry into the realm of public policy, system analysis, and institutional change frequently requires that we do so without a recognized credential which asserts *prima facie* that such involvement is both our right and responsibility. Indirect/consultation and prevention strategies demand that we operate within alien territory in which generic knowledge and skills are the currency for goal attainment. Encouraging people in need to take advantage of self-help groups, enabling those without such experience to recognize and assert their right to influence the allocation of resources, and applying our science and profession to make teachers, parents and social service providers better at what **they** do seem quite discordant with established definitions of professional functioning.

In the sole course on consultation I encountered in graduate school nearly three decades ago (during community psychology's presumed take-off phase), I was both impressed and frightened to learn that effective consultation is documented by evidence that the service is no longer needed by its recipient. It seemed that the task required providing recipients of the services with possession of critical elements of what I offered to be shaped in whatever form **they** deemed useful. My job, in effect, was to make my job unnecessary! Providing prevention and indirect/consultation services seemed to require literal acceptance of Miller's (1969) mandate that we "give psychology away." It simultaneously challenged the content of most of my didactic and clinical courses! It turned around basic assumptions about professional responsibility, and about the patient's capacity to change their problem behaviors and feelings. Most importantly, perhaps, it required significant revision in the very understanding I was developing of why people behave and feel as they do.

6.13.4.2 Shifting Emphasis from Inside to Outside

Suppose that the emotions and behaviors which concern mental health professionals reflect not only internal processes but also the circumstances under which they are displayed? Were this the case, to what should one attend . . . the people or the circumstances? Three decades ago, as the Community Mental Health Movement was getting underway, Barker (1968) and his colleagues proposed that **behavior settings** have greater influence on the expression of certain behaviors than the people who inhabit those settings. In other words, one might predict what will happen if one knows **where** it is expected to happen more so than if one knows **who** will be present. If generalizable, this assumption implies that understanding the conditions under which social services are provided, classroom instruction is given, play is attempted, or employment opportunities arise may predict expressed emotion and behavior better than (or if combined with) information about the individual's early childhood experiences, measured intelligence, or personality profile. It also implies that structuring those conditions might influence predictably the experience and expression of emotions and behaviors. Such a perspective would provide a theoretical rationale for program, setting, or system consultation focusing on the conditions under which people lived, worked, played, and learned. Although given little attention, a substantial theoretical and empirical basis for responding to individual emotional and behavioral needs through indirect/consultation approaches has been developing for decades.

It should be noted that attributions to factors external to the individual have a long and respected history in public health. The classic "host–agent–environment" model for understanding the occurrence of disease has successfully informed efforts to prevent and control threats to physical health. It worked because it looked both **within and outside** the individual to understand vulnerability and resistance to disease. Consideration of individual and environmental factors allows for the design and implementation of interventions whose health promotion or disease/disorder prevention effects can be delivered unobtrusively. To be effective, it is not necessary that the recipients of these interventions be either identified or participate actively. Rather, not unlike fluoride in the water or guard rails along the highway, the environment can be changed in ways which are preventive and health promotive. Many of these interventions operate without professional involvement and independent of either the cooperation or knowledge of the vulnerable individual. By contrast, the application and ultimate success of treatment generally depend on multiple patient variables including acknowledging the need for, having access to, being able to pay for (in terms of money and time), and complying with a treatment protocol. Interventions which have their effects independent of such factors and, ideally, which preclude the need for treatment would clearly be logistically and humanly advantageous. To consider those as viable, however, one must believe that factors **external to the individual** are salient to health and illness.

In the early eighteenth century before germ theory accounted for contagion, the spread of disease among the poor was attributed to "miasma" (i.e., bad air). As Bloom explained (1977):

> Miasma theory held that soil polluted with waste products of any kind gave off a "miasma" into the air, which caused many major infectious diseases of the day. People living near swamps, and thus particularly vulnerable to marsh gases, were thought to develop fever from these gases—a fever that came to be known as *malaria* (bad air). The miasmist felt that the way to prevent disease was to modify the environment—to remove the sources of the miasma. (p. 71)

Given confirmation that a different external factor,"germs" (i.e., viruses and bacteria) explained most infectious disorders, miasma theory had limited influence on public health policy and practice. Its major tenet, however, informs concepts of vulnerability and resistance to emotional and behavioral disorders.

Miasma theory posits that nonspecific characteristics of an environment are pathogenic. This position echoes widespread beliefs among mental health epidemiologists (e.g., Dohrenwend, B. P. & Dohrenwend, B.S., 1981; Dohrenwend, B. S. & Dohrenwend, B. P., 1981; Dohrenwend et al., 1987) that elements of the environment linked to experienced stress are significant pathogens. Analogously, it echoes explanations of delinquency based on social learning and modeling processes (Berkowitz, 1993; Goldstein & Keller, 1987) and links observed among neighborhood disorganization, deterioration (e.g., the presence of broken windows and littered vacant lots), and antisocial acts (Gold, 1987; Levine & Rosich, 1996). Continuing the analogy, indirect/consultation strategies could be thought of as avenues to "modify the environment—to remove the sources of the miasma" (Bloom, 1977, p. 71).

6.13.5 EXTERNALIZING THE STUDY OF BEHAVIOR

Albeit outdated and dismissed, miasma theory is offered to focus readers' attention on nonspecific characteristics of settings and situations in which disease and dysfunction disproportionately occur. Alternatives to miasma were sought because of the variability of its effects, that is, not all exposed became ill. Not understood at the time were how individual characteristics enhanced resistance or raised vulnerability. Such distinctions are central to contemporary diathesis-stress models of illness (Lilienfeld & Lilienfeld, 1980). The complexity of understanding how nonspecific setting factors mix with individual variability raises questions about the processes by which each influences the other. More than six decades ago, Lewin (1935) asserted the importance of contextual factors in his widely cited formula:

$$B = f(P, E)$$

This formula reflects Lewin's (1944) concept of "psychological ecology," that is, consideration of setting characteristics and opportunities *and* individual characteristics and constraints into predictions of human behavior. Lewin's perspective has been studied along the lines of research emphasizing, respectively, **transactional** (i.e., the person in environment) and **environmental** contributions to behavioral and developmental processes. The distinction between the two lines of inquiry is one of relative rather than absolute emphases of the contributions of individual and situational factors in determining behavior. Both lines provide important scientific and clinical bases for the design and evaluation of indirect/consultation strategies and offer avenues for altering pathogenic processes toward preventive goals.

The qualifier,"transactional," represents an important element of Lewin's conceptualization of behavioral dynamics. A transaction refers to the inter-relational processes between individual and situational factors in which each reciprocally influences the other. The distinction between a "transaction" and an "interaction" lies in the former's bidirectional nature, that is, each component in the transaction influences the **successive** states of the other. Dewey and Bentley (1949) described transactional processes as a basis for epistemological analysis. Based on the models of thought articulated in that work, the study of human behavior requires consideration of the actor (i.e., the knower), that which is acted upon (i.e., the known), and the circumstances under which the action occurred (i.e., the situation in which the knower confronts the known). It is within that entirety that the meaning of an action or event must be understood. As explained by Allport (1955):

> Perceptual order and stability can be regarded not merely as a matter of the appraisal of cues, but also as a more dynamic relationship between the organism and the environment. When percepts

> are attained their attainment is likely to involve direct participation in some process of overt adjustment. *Perception and action can be seen to have a close functional relationship* (emphasis added). (pp. 271–272)

From a transactional perspective, violent acts occurring in a school, for example, would be understood in terms of when, how, and where they occur, who is present at the time, and how each person present interprets the events and the options available to respond to those events. The coalescence of those elements gives meaning to the violence and determines how its occurrence influences subsequent events. Among those events are the responses of others to the violence, the likelihood of its continuation by its perpetrator(s), and of its adoption by others (Lorion, Brodsky, & Cooley-Quille, 1998).

As used in this chapter,"environmental" research refers to studies which examine the contributions of setting characteristics to behavior. As explained by Schoggen (1989), environmental studies are exemplified by analyses of the physical or built environment on behavior (e.g., the design of playgrounds on student interactions or of classrooms on reading) and of physical qualities of a setting on functioning (e.g., noise and lighting levels on classroom on-task behavior). Such studies may, however, focus on the "the ecological environment (which) includes also the objectively observable *standing patterns of behavior* of people—that is, specific sequences of people's behavior that regularly occur within particular settings" (Schoggen, 1989, p.2). This subset of environmental studies must, themselves, be distinguished from work on the "psychological environment," that is,"the subjective representation of the objective environment by a given person at a particular time" (Schoggen, 1989, p. 3).

6.13.6 ENVIRONMENTAL MODELS OF BEHAVIOR

As noted, Lewin's (1935, 1944) assertion that behavior is a function of person and environment stimulated inquiry into the contribution of setting characteristics in shaping behavior. Barker, a colleague of Lewin, developed this line of research. In Barker's (1968; updated in Schoggen, 1989) view, the distinction between Lewin's concept of "life-space" or the psychological environment (i.e., the world as an individual *perceives* and is affected by it) and the ecological environment (i.e., the objective characteristics of real-life settings) makes evident the need to understand how the physical properties of settings, including the objects within them and the shared rules for relating to them, influence behavior.

Support for the importance of attending to the ecological environment was provided in the findings of a series of early studies (summarized in Barker, 1968) on children's behavior across school settings. Barker and his colleagues determined that knowledge of situational characteristics of such settings resulted in better prediction of children's behavior than knowledge of the characteristics of individual children. Across a number of settings, children acted more like each other within a given setting than like themselves across settings. To the researchers, measurable properties of settings appeared to constrain the range of alternative responses and opportunities for behavior within each setting.

Barker and his colleagues used the concept of "behavior settings" to describe such person-setting linkages. As explained by Barker:

> A behavior setting has both structural and dynamic attributes. On the structural side, a behavior setting consists of one or more standing patterns of behavior-and-milieu, with the milieu circumjacent and synomorphic to the behavior. On the dynamic side, the behavior-milieu parts of a behavior setting, the synomorphs, have a specified degree of interdependence among themselves that is greater than their interdependence with parts of other behavior settings. (1968, p. 18)

In effect, a behavior setting refers to a pattern of behavior which appears regularly under specific setting-defined circumstances. As explained by Schoggen (1989): "From this viewpoint, the environment is seen to consist of *highly structured, improbable arrangements of objects and events that coerce behavior in accordance with their own dynamic patterning*" (p. 4).

Such patterns of behavior appear regularly and predictably; their occurrence is independent of the specific individuals present. These patterns of behavior are linked both in place (i.e., are circumjacent) and form (i.e., are synomorphic) to characteristics of the setting but *not* to the presence of specific individuals as much as to the presence of categories (e.g., students and teachers; players and coaches) of individuals. Within a behavior setting, the behavior and setting are "interdependent" to the extent that the former is unlikely to occur outside of the latter and that qualitative aspects of the behavior are shaped by characteristics of the setting.

Examples of behavior settings include the fit between the organization and structure of a classroom and the behaviors expected of, and displayed by, students and teachers; the distinct activities of students within a gym, in a library, or on a playground's basketball court or hopscotch grid. The physical setting and its associated rules make some behaviors highly likely and other behaviors highly unlikely. Schoggen (1989) explains that neither the physical nor social aspects of settings preclude all but a single behavior; rather, these settings aspects make "actions of some kinds easier than others" (p. 43). Wicker (1979) explains this process as follows:

> *Behavior settings are self-regulating, active systems.* They impose their program of activities on the persons and objects within them. Essential persons and materials are drawn into settings, and disruptive components are modified or ejected. It's as if behavior settings were living systems intent on remaining alive and healthy, even at the expense of their individual components. . . .
>
> . . . to summarize some of the essential features of behavior settings. Most of them can be presented in a single sentence: *A behavior setting is a bounded, self-regulated and ordered system composed of replaceable human and nonhuman components that interact in a synchronized fashion to carry out an ordered sequence of events called the setting program.* (p. 12)

Although the heuristic potential of Barker's work on setting characteristics has yet to be aggressively mined (readers are encouraged to review Levine & Perkins, 1987; Moos & Insel, 1974; and Wandersman & Hess, 1985 for illustrations of the potential richness of this vein), Barker and other environmental psychologists offer a heuristic avenue for understanding and responding to external contributors to health and pathology. That the concept of behavior settings may help unravel interpersonal processes within educational settings is strongly suggested in findings reported by Barker and his colleagues (e.g., Barker, 1968; Barker & Gump, 1964; Barker & Schoggen, 1973; Barker & Wright, 1951; Schoggen, 1989). Extrapolating from that work, it seems reasonable to assume that aspects of the environments in which people live, learn, play, and work can be shaped such that they can contribute to optimal adjustment to developmental demands and reduced potential for emotional and behavioral dysfunction.

6.13.7 TRANSACTIONAL MODELS OF BEHAVIOR

In his biopsychosocial theory of behavior, Bronfenbrenner (1977, 1979) focused attention on the synergistic relationship between the individual (e.g., IQ, temperament) and setting (e.g., parental responsiveness) in the processes of human development. Extending that theory, Sameroff (Sameroff & Chandler, 1975; Sameroff & Fiese, 1989) proposed a transactional-ecological model for understanding variation in the outcomes associated with presumably established risk factors for developmental disabilities. Unique to this model is a hypothesized "transaction" between individual and contextual factors. For Sameroff, emotional and behavioral development is the product of an ongoing synergistic series of: (i) responses by an individual to situational circumstances and demands; (ii) alterations of those situational circumstances and demands as a function of the individual's responses; and (iii) responses by an individual to that altered situation. Individual and setting characteristics, however, differ in their malleability and range of variation. Consequently, it should not be assumed that each transaction represents equivalent changes in individual and setting characteristics.

Nevertheless, within Sameroff's model, individuals and settings evolve continually (albeit generally gradually rather than dramatically) over time. This element of the theory augments substantially the complexity of understanding and controlling behavioral development. It also suggests that behavior can be shaped through controlled changes in the individual, the context, or both. Sameroff's model mirrors the continuing occurrence and adaptation to events which characterize the ongoing flow of daily life. It also offers a basis for contextualizing variations in behavior across situations. In effect, unyielding settings would constrain behaviors within circumscribed parameters. In the most recent version of the model, Sameroff and Fiese (1989) make explicit that societal (and, presumably, subgroup) expectations have regulatory influence on definitions of behavioral roles and thereby set contextual limitations on developmental outcomes. Insofar as such expectations and setting demands are fixed, the behavioral responses of those within a setting may be more restricted than assumed or desired.

Kellam's (Kellam, Branch, Agrawal, & Ensminger, 1975) Life Course-Social Field model offers further enhancement of Lewin's and Bronfenbrenner's theories. Kellam's approach integrates elements of the biopsychosocial perspective with Erikson's (1963) recognition that development involves confronting and resolving a series of social tasks across the life span. Thus, Life Course-Social Field theory defines development in terms of movement over the life span across diverse settings (i.e., social fields), each presenting its own demands and adaptive challenges.

The success with which the unique and common tasks associated within and across social fields are resolved is assessed both by oneself and by significant others in those settings. Kellam labels the substance of assessments of one's own psychological state as "psychological well-being." Examples include one's sense of self-efficacy, anxiety, self-esteem, depression, disappointment, inadequacy, etc. By contrast,"social-adaptational status" refers to the assessments which others make of the adequacy of an individual's response to these developmental tasks across the multiple contexts (labeled "Social Fields" in Kellam's model) in which we live, play, or work, for example.

Unique to Kellam's model is his emphasis on the fact that the individuals (i.e.,"natural raters") who make such assessments vary across tasks and social fields. In that sense, social fields are *social*, that is, they involve exchanges among the players in settings who have setting-related roles and expectations. Depending on their responses, these raters influence for better or worse the individual's resolution of task demands. The salience of particular social fields and of their respective natural raters changes over time and across developmental demands. Thus, the school for example, has particular salience during childhood and adolescence but much less so during the adult years. Similarly, as peers grow in importance, the family becomes less salient during adolescence and early adulthood and more so thereafter.

6.13.7.1 Linking Inside and Outside

As noted, Levine and Perkins (1997) suggested that an important obstacle to the widespread adoption of alternative perspectives of emotional and behavioral problems (i.e., as adaptations to setting conditions rather than as reflections of endogenous psychopathology) and their resolution (e.g., indirect/consultation and preventive approaches rather than psychotherapy) was the absence of "alternate theoretical conceptualizations" (p. 60). Such an alternative is now available in the growing body of work linking internal processes and external conditions. Environmental analyses reviewed thus far (e.g., Barker, 1968) provide evidence supporting the impact of external factors on emotional and behavioral functioning. Developmental theorists such as Sameroff (e.g., Sameroff & Fiese, 1989), Kellam (e.g., Kellam et al., 1975), Bronfenbrenner (1979), and Rutter (e.g., Rutter et al., 1997) extend our understanding of the programmatic significance of this impact by clarifying the transactional processes by which external and internal events influence each other.

Acceptance of such insights is heuristically important for multiple reasons. Orford (1992) notes that "one of the consequences of complex, transactional influence is that outcome is difficult to predict" (p. 33). Rather than a limitation for intervention design and application, this consequence argues for use of indirect/consultation approaches, particularly in the pursuit of preventive goals. As explained by Sroufe (1997) and Rutter et al. (1997), conceptualizing psychopathology as a developmental phenomenon shifts the focus from the individual to the individual–environment dynamic. Particularly important is recognition of such phenomena as reflecting adaptations to environmental demands rather than as the inevitable result of individual failings, unresolved early experience, biological determinism, or other endogenous factors. Within this framework, adaptive states represent points along developmental pathways. Rather than reflecting linear and deterministic processes whose outcomes are fixed from the outset (e.g., based on inherited characteristics) or soon thereafter (e.g., based on early traumatic experiences), such pathways evolve over time branching this way and that depending on specific mixes of adaptive resources, coping demands, and as yet unknown determinants of salience at the moment (e.g., Bell, 1986).

Outcome becomes difficult to predict if one requires that knowledge of beginnings implies knowledge of endings or vice versa. By contrast, emotional and behavioral states become manageable if understood in terms of antecedent processes and likely subsequent presentations (some of which may be mutually exclusive but most of which may overlap or co-occur). Such a view of pathogenesis has multiple implications (Sroufe, 1997):

(i) adaptation and maladaptation represent evolving rather than established conditions determined by what has occurred in the past, what is currently happening and the range of options allowed by cultural, setting and individual circumstances. This temporal quality of functional states links what has been with what is and *the likelihood*, ***but not certainty*** of subsequent states. Kandell's ( ) "gateway theory" of drug involvement exemplifies this principle. Use of alcohol or tobacco, for instance, is associated with an increased probability of marijuana use just as the latter is associated with an increased probability of use of cocaine or heroin. Any individual may, however, not go beyond a particular gate; hence the discontinuity between antecedents and outcomes (Kandel, Single, & Kessler, 1976);

(ii) because many characteristics of experience and environment are shared across individuals (e.g., who live in a given neighborhood or are raised within a common culture), diverse pathways may lead to a common outcome.

General systems theorists and developmental psychologists refer to this pattern as "equifinality" (Cicchetti & Rogosh, 1996; Orford, 1992). Highly diverse economic, familial, educational and occupational antecedents, for example, have been associated with drug involvement;

(iii) similarly, although individuals may share common antecedents (e.g., poverty, being reared in a single parent household; a history of physical or sexual abuse) differences in circumstances subsequent to those antecedents give rise to a variety of outcomes (e.g., depression, alcoholism, conduct disorder). General systems theorists and developmental psychologists refer to this pattern as "multifinality" (Cicchetti & Rogosh, 1996; Orford, 1992);

(iv) by combining a, b and c, one appreciates that developmental pathways offer multiple opportunities for changing direction and hence influencing outcome. Branching off may reflect the occurrence of positive or negative events (e.g., an increase in family income, parental death or divorce), the addition of individual or circumstantial resources which enhance or reduce adaptive capacity (e.g., acquisition of interpersonal problem-solving skills or access to a caring teacher) or a change in the environment (e.g., moving to a more or less violent neighborhood). Whatever the case, a seemingly normative or pathogenic process may be interrupted and its direction changed for good or ill;

(v) recognition that developmental trajectories can be altered must be tempered with appreciation of the limitations of that assertion. It appears that the further along a path one has gone and the longer the pathogenic process has been underway, the less likely it is that the path's direction can be altered. This seems to be the case because the capacity to recognize, access and use additional resources in the environment to adapt to health compromising situations depend on prior experiences, verbal abilities, interpersonal skills etc. In effect, the further along a path one has traveled, the more likely one is to reach its end point. "More likely," however, refers to probability rather than inevitability!

The conceptual option presented links exogenous and endogenous processes in a way that enriches comprehension of etiological processes and expands the avenues for interventions to influence those processes. Developmentally, emotional and behavioral patterns reflect adaptations to situational demands based on individual and environmental resources and the adaptive history of both the individual and the setting. (Mal)adaptation may proceed linearly or circuitously, winding like a river or branching out like the roots of a tree. Depending on the resistance encountered as well as what has come before, one among many alternative paths will be followed. Understanding that path will likely require understanding the history of the resistance encountered. Some of it may reflect the prior experiences of a specific individual; some may instead represent a community's, a neighborhood's, or a family's *post-hoc* response to the behavior or needs of others. Some paths will proceed in a relatively linear fashion; antecedents are followed by clinically defined emotional or behavioral problems. Removing, neutralizing, or interrupting antecedents prevents the undesired outcome.

Unfortunately, few mental health conditions can be linearly modeled. More typical is Sroufe's (1997) portrayal of emotional and behavioral development as the branches of a tree or the meanderings of its roots. Within this perspective, growth and survival requires maximizing one's access to necessary resources regardless of the impediments encountered in the environment. At any point in time, emotional and behavioral status reflects the cumulative effects of its antecedents. The further back one traces a branch the more completely one can understand what has been and will be encountered. With such information, the better one can predict the **probability** of the branch's likely direction. Where the branch connects to the trunk is but one element to be considered in predicting its various termini. That beginning is likely to be common across many outcomes (i.e., the principle of multifinality). Equally likely is that termini in close proximity may have begun at different points along the trunk (i.e., the principle of equifinality). Adding to the complexity is that some branches will end close to where they began, whereas others continue on toward other terminating points. Presumed antecedents (e.g., alcohol or tobacco use) for some conditions (e.g., substance dependence or addiction) may, for many, represent an endpoint (i.e., substance use is limited, controlled, or even ended). In other cases, however, the process continues toward dependence and addiction.

Common antecedents leading to diverse outcomes; diverse antecedents leading to common outcomes; the co-occurrence of multiple outcomes—may be perceived as too complex and misunderstood an etiological model to inform intervention design, especially preventive interventions (Lamb & Zusman, 1981). Presumably, one should wait until the die is cast, the nature of the pathology is clear, and treatment can be selected and applied. Public health practitioners might accept this approach were the costs of delay (health related and financial) minimal, the effectiveness of treatment certain, and the secondary consequences of the disorder acceptable. The common cold and the 24 hour flu fit within these parameters.

Most emotional and behavioral disorders, however, do not.

Alternatively, the aforementioned complexity may be recognized as allowing for, indeed requiring, nonspecific interventions focusing on common etiological factors, be they individual, environmental, or transactional. Targeting common risks factors or pathways which are epidemiologically linked with undesirable outcomes has the potential for impacting **simultaneously** multiple outcomes. Simply stated, interventions targeting common antecedents may impact on multiple pathways and thereby serve as a stone which kills (or at least weakens) multiple birds! By aggregating the measured preventive effects across those outcomes sharing common risks, the true impact of such interventions may finally be documented (Lorion et al., 1989). Applying the principle of multifinality to intervention design leads to targeting a limited number of shared risk factors and assessing their impact across alternative pathogenic expressions. Similarly, the principle of equifinality requires that interventions intended to reduce the occurrence of a specific outcome must target the various pathways and risks associated with that outcome. In both instances, basic elements of developmental theory supported by epidemiological findings argue against the likelihood of documenting substantial evidence of preventive outcomes if only a limited number of risks factors are targeted with the intent of impacting a single form of disorder or dysfunction. As noted below, neither prevention theory nor the traditional intervention taxonomy made that point clearly.

6.13.8 SHIFTING TO A PREVENTION FOCUS

Initial forays by the mental health sciences and professions into preventive efforts were led by Caplan (1964) who recognized that the mental health disciplines had to move beyond sole reliance on treatment if they were to serve public needs. To organize initial efforts in this direction, Caplan urged mental health to adopt the goals of and the classification system for prevention extant in public health at the time (Commission on Chronic Illness, 1957). This system was centered around the epidemiological concept of "caseness," that is, confirmation that the diagnostic criteria defining a syndrome were met. Intervention categories were differentiated within this system in terms of their proximity to the targeted condition's fulfillment of diagnostic criteria, that is, meeting the symptomatic definition of a syndrome. Preventive interventions were designed to reduce the prevalence of a targeted disease or disorder.

Within this framework, **primary** prevention interventions are implemented prior to caseness. Successful interventions reduce prevalence (i.e., overall presence of cases of a target disorder in the population) by reducing incidence (i.e., the occurrence of new cases). Potential recipients of such interventions range from the population at large, to asymptomatic subgroups targeted on the basis of epidemiologically-defined risk, to subgroups presenting prodromal signs antecedent to diagnostic status. By contrast, **secondary** interventions seek to lower prevalence by reducing the duration of caseness, that is, through the application of effective treatment. Combinations of sensitive screening procedures which identify cases early and involve them in treatment exemplify this category. Secondary approaches seek both a return to premorbid status and avoidance of subsequent episodes. Finally, **tertiary** interventions reduce the long-term disabilities and sequella consequent to caseness. Rehabilitation strategies and support services exemplify tertiary approaches. Insofar as they enable a person with schizophrenia to live independently, tertiary goals are achieved. If recurrence of disabling symptoms is avoided, secondary goals are achieved.

Ideally, primary prevention interventions are preferable. This assumes, of course, the availability of necessary etiological information, access to the population at risk, and effective intervention strategies. It also assumes that the intervention's effects are, at worse, neutral. This is a particularly important point given that the diseases and disorders of most interest to mental health are relatively low-frequency events. Hence, unless substantial risk information is available and targeting is very precise, it is most likely that the majority of those receiving the intervention are unlikely to manifest the disease or disorder in the absence of the intervention. If the intervention has iatrogenic consequences, then these individuals may be at greater risk for those consequences than to the etiological outcome to be avoided. Dismissal of this concern notwithstanding (Albee, 1986), evidence for caution has been reported (e.g., Lorion, 1987; Sameroff & Fiese, 1989) although the frequency of negative consequences appears to be low (Durlak & Wells, 1997). Secondary and tertiary prevention efforts, by contrast, are targeted only to those meeting diagnostic criteria and hence having confirmed need for the intervention.

As noted earlier (Lorion, 1983), the classic public health triad of approaches applied to the disorders (e.g., infectious diseases) and etiological processes of primary concern to public health practitioners. The applicability of its underlying causal assumptions, however, to emotional and behavioral disorders has been challenged from the outset (e.g., Albee, 1982; Lamb & Zusman, 1979). As reflected in the work of Brofenbrenner (1979), Rutter (1989), and Sameroff and Fiese (1989), there is an emerging appreciation of the etiological complexity of emotional and behavioral problems involving biological, psychological, social, and environmental parameters. This complexity obscures determination of the onset of pathogenic processes, the presence of disorder, and thereby designation of "caseness" (Lorion et al., 1989). Unresolved, these challenges produce disagreements in the categorization of primary and secondary interventions and obstacles to empirical verification of their efficacy. In either case, they allow little room for consideration of the principles of equifinality and multifinality in intervention design or evaluation.

Gordon (1983, 1987) has proposed an alternative taxonomy for organizing preventive efforts which allows for consideration of these most important developmental principles. Given the low base rate at which emotional and behavioral disorders occur in the general population, Gordon's system also allows for consideration of both the financial costs and the iatrogenic potential of the interventions. Weighting expected intervention risks with target selection, Gordon (1983) proposed that "universal," "selective," and "indicated" interventions should be designed for and targeted to "persons not motivated by current suffering" (p. 108) and neither currently seeking nor in need of treatment. **Universal** interventions are applied to the population at large and combine low costs per contact with limited likelihood of iatrogenic consequences. Examples of such approaches include public service announcements advocating seat-belt use and physical exercise or discouraging tobacco use and the consumption of alcohol during pregnancy. In the mental health realm, universal interventions include the "Friends can be good medicine" program implemented in California to enhance use of social support resources and emerging strategies to teach parenting skills to adolescents. Perhaps the designator "universal" may also be applied to the generic nature of the risk factors targeted by such interventions. If so, the difficulty of measuring their impacts becomes understandable given the multiplicity of outcomes which may be affected at undefined points in the future.

Selective interventions are targeted to segments of the population for which there is an epidemiologically established risk. Selection may be based on identified links between gender, ethnicity, economic status, or family history, for example, and the presence of emotional or behavioral disorders. Targeted subgroups have a higher likelihood of disorder and hence a more focused (and potentially more iatrogenic) intervention can be justified. Examples include programs targeted to families considering marital separation or divorce (Wolchik et al., 1993) or the recently widowed (Silverman, 1988).

Indicated interventions target individuals (rather than subgroups) who present risk factors or prodromal signs indicating substantial individual risk for subsequent disorder. Examples of these approaches include programs for women presenting early indices of depression (e.g., Vega, Valle, Kolody, & Hough, 1987) and adolescents referred to juvenile court for status offenses (Lochman, 1992). It also includes Olds' substantially validated approach to assisting high-risk adolescent mothers (Olds & Korfmacher, 1997, 1998a, 1998b).

Gordon's proposal was intended to stimulate development of preventive efforts and to make explicit how they were clearly distinct from treatment of symptoms and syndromes. In an important sense, Gordon (1983) proposal offers a conceptual expansion of the original "Primary prevention" category. His approach highlights the breadth of opportunity available to influence prevalence by intervening along that functional continuum ranging from health through various prodromal stages prior to meeting diagnostic criteria. The richness of this alternative system lies in its clarification of concepts of "risk" and "vulnerability" as applied to individuals, situations, and their combinations. It also lies in its clear distinction between the end-state conditions to be avoided (e.g., substance abuse or depression) and susceptible individuals. Sameroff (1977) made a similar plea in his early warning to prevention scientists that their efforts must not compromise respect for the humanity of the recipients of those efforts. In his advocacy for health promotion efforts, Cowen (1996) similarly urged program developers to recognize that removal or reduction of risks and vulnerabilities could leave individuals both empowered and likely to reach their potential. In their meta-analytic study of primary prevention programs, Durlak and Wells (1997) report clear evidence of the value of skill-building and competency enhancement as a means of avoiding dysfunction as well as improving quality of life for at-risk youth.

6.13.9 MERGING INDIRECT/ CONSULTATION STRATEGIES WITH PREVENTION GOALS

Whatever classificatory scheme is adopted, epidemiology and developmental psychopathology provide important conceptual and methodological foundations for meaningful advances in prevention (Mrazek & Haggerty, 1994). Application of case–control methodology, for example, enables epidemiologists to discern potential risk and protective factors associated with symptomatic and syndromal expression. Developmental psychopathology's longitudinal methods offer insights into the pathogenic processes defining vulnerability and allow for modeling the sequence of states and events which define the trajectory from health to dysfunction, to distress, and potentially (but not necessarily) to pathology. As noted, examination of developmental processes from a system's perspective brings to the forefront the importance of the principles of equifinality and multifinality. In turn, the application of these principles argue for the design and implementation of preventive interventions which target multiple risk factors linked to one or more negative emotional and behavioral outcomes. This view is supported by Mrazek and Hall (1997) who question Durlak and Well's (1997) decision to exclude programs which sought to improve academic achievement and reduce substance use from their meta-analysis of preventive interventions:

> Even though this is a legitimate way to narrow the huge task they faced, it also represents one of the main problems in prevention science today, that is, the categorical approach to mental, social, educational, behavioral and legal problems. Many of these problems have common risk factors that interact in complex causal chains. Addressing clusters of risk and protective factors increases the chances of preventing multiple problems in many areas of functioning. It is the accumulation of multiple proximal outcomes across various domains of functioning, such as educational status, social adjustment, and behavioral and emotional well-being that will be the most convincing arguments for the effectiveness of prevention. (Mrazek & Hall, 1997, p. 223)

Impacting on such a diversity of targets requires in my view the application of indirect and consultation strategies to maximize achievement of desired outcomes. As noted earlier in this chapter, such approaches characterized early twentieth century attempts to respond to the needs of children and families at risk for emotional and behavioral problems. Across multiple settings and a diversity of goals relating to improving the status of those in need, such early efforts at consultation applied expertise in developmental processes and mental health to enhancing the effectiveness with which parents, teachers, and other members of a community met their responsibilities. Consultants observed behavior, examined policies, and procedures, identified and examined the assumptive bases of definitions of problems and solutions, and thereby opened new perspectives and alternatives to those whom they served. The recipients of their services, in turn, impacted on the lives of the children and families deemed to be at risk.

Reviews of mental health consultation (e.g., Bloom, 1984; Gallessich, 1986; Mannino & Shore, 1986) reveal both the promise and limitations of attempts to extend these early efforts to meeting the mandates of the Community Mental Health Centers Act. Whether focused on the needs of a client being provided mental health services, of the provider delivering those services or of the agency in which those services, were delivered, consultation was generally targeted to responding to the needs of those with identified dysfunction or incipient forms of disorder. The delivery system itself was not altered nor were the conditions for providing its services.

What seems necessary at this time is the application of indirect/consultation strategies which redefine service delivery systems and thereby impact on the common risk factors seemingly essential to the etiology of diverse emotional and behavioral problems. Durlak and Wells' (1997) meta-analysis of prevention programs for children and adolescents makes evident the promise of such interventions. Similar optimism is supported by the findings of the recent Institute of Medicine review of the state of preventive science (Mrazek & Haggerty, 1994). What both reviews underline is the importance of focusing on building skills in the individual at risk or in those who surround the individual. Preventive interventions with the most promise appear to strengthen basic academic skills, interpersonal resources, and the responsiveness of the social systems which can influence developmental outcomes (Durlak & Wells, 1997; Lorion, 1990; Mrazek & Haggerty, 1994). The best of such work is theory-based, systematic and, most importantly, **incremental** (Price, 1997; Price & Lorion, 1989). Like the pieces of a puzzle whose image gradually emerges, program development proceeds through the sequence depicted by Price (1983) as the "four domains of prevention science" (p. 291). These domains include: (i) problem analysis (i.e., epidemiological and etiological studies); (ii) innovation design (i.e., program development); (iii) field trial (i.e., efficacy research); and (iv) innovation diffusion (i.e., effectiveness research).

A most important element of Price's model for the development of preventive intervention was the interconnectedness of these domains. In effect, Price recognized how progress through these stages modifies one's initial understanding of the originating problem and hence catalyzes the refinement of each subsequent stage. Earlier, Cowen (1980) had differentiated the initial two domains as "generative" and the latter two as "executive" components of prevention science. Generative studies have scientific importance beyond their capacity to inform the design and implementation of interventions for they inform us "about the very processes whereby functional and dysfunctional states develop and are maintained. Preventive research can increase our understanding of the environmental parameters which engender, maintain, and modify human behavior" (Lorion, 1983, p. 264). The continuing heuristic value of Cowen's distinction between generative and executive studies and Price's four domains is reflected in the Institute of Medicine's analysis of the state of prevention science. To catalyze the field, the report advocates that the field adopt the "preventive intervention research cycle" as its map for the systematic accumulation of information (Mrazek & Haggerty, 1994). Optimistically, they conclude:

> If the research standards and methodology outlined here are systematically and rigorously applied within the preventive intervention research cycle and the guidelines on cultural, ethical, and economic issues are carefully considered at each step, prevention research will yield progressively more powerful results over the next decade. The ensuing development of prevention into a science will provide a firm base of knowledge for policymakers. This knowledge will inform their decisions on the allocation of available resources toward the ultimate goal of realizing the opportunities presented by the science for the alleviation of the personal and societal suffering and burdens associated with mental disorders. (pp. 408–409)

The validity of this assertion is clearly represented by the Home Visitation Program developed by Olds and colleagues and described in detail in two special issues of the *Journal of Community Psychology* (Olds & Korfmacher, 1997, 1998a, 1998b). Built slowly and methodically over more than two decades, this program exemplifies the challenges and potential payoffs of the systematic pursuit of preventive goals. Of direct relevance to this chapter is that the program relies on a mixture of direct and indirect intervention components to accomplish its goals. These components derive from recognizing the complementary aspects of self-efficacy, human attachment, and human ecology theory. Recipients of the program's direct intervention are primiparous (< 26 weeks gestation), young (< 19 years of age) women who are deemed at risk because they are single and/or low-income. Program goals for these women included the reduction of pre- and perinatal risks to decrease epidemiologically established risks for health problems in late pregnancy, delivery complications, and associated developmental sequella for the infant. The program also sought to alter parenting practices associated with negative child health and developmental outcomes. Finally, the program sought to improve maternal life course by extending the interval between the initial and subsequent pregnancies and thereby increase the likelihood of mothers returning to school, obtaining work, and becoming independent of welfare. Women were randomly assigned to intervention or control conditions and the intervention was provided initially by registered nurses (a comparison of the differential effectiveness of nurses and paraprofessionals is currently underway).

The program began in Elmira, NY with a predominantly White, rural sample and was replicated and extended with an urban, African-American sample in Memphis, TN and a diverse sample in Denver, CO. The initial cohorts have now been followed for nearly two decades. In considering the findings, the complexity of this program becomes evident. Through contacts with the home visitor during the final stages of the pregnancy, mother's knowledge, attitudes, and behaviors concerning health compromising risks (e.g., diet, exercise, alcohol, and tobacco use) are to be influenced with the hope of reducing birth complications and optimizing fetal development. Following the birth, home visitation continues its focus on maternal health as well as child care, developmental processes, and, importantly, the implications of maternal choices on the life course. The latter is important insofar as the mother assumes control of the timing of subsequent pregnancies as she recognizes its implications for her educational and employment options and their implications for her immediate and long-term economic situation.

From this perspective, the program delivers a direct service to the mother. The lack of a mental health diagnosis notwithstanding, the service appears similar to many forms of treatment. The visitor meets regularly with mother to discuss issues, provide information, and offer counsel and support. The notable difference, however, is that at least a portion of the changes in the mother are intended to alter permanently **her** behavior toward her child. Not unlike the Child Guidance and Settlement House social programs of the early twentieth century, the Home Visitation program represents a form of indirect/consultation service. Toward this goal, the program strategy is designed to enhance the mother's capacity to parent and to assert herself in employment and social service situations. If this element of the program is successful, significant risks for the child (e.g., developmental lags, physical or sexual abuse, early academic failure) can be reduced which have been linked epidemiologically to school failure, conduct disorder, substance involvement and, later in life, early pregnancy, and cross-generational child abuse.

Changes in a mother participating in the Home Visitation program, therefore, can ripple across the family, home, neighborhood, and other levels. Assuming that the program's effects on the mother are lasting, developmental benefits observed in the developmental course for the firstborn should hold for subsequent children as well. To the extent that mother's overall heath behaviors and especially prenatal health behaviors are improved, further reductions in neonatal risk should occur for subsequent children. If the intervention represents a viable approach to assist women to move from welfare to employment, program effects support the achievement of indirect/consultative goals at the public policy level.

Results reported thus far confirm the benefits of merging direct and indirect strategies in the pursuit of preventive goals. Summarizing findings from the Elmira and Memphis samples, Olds et al. (1998) report that compared to mothers randomly assigned to the control condition, those assigned to the Home Visitation intervention smoked less and improved their diets during pregnancy, had fewer hypertensive disorders and preterm deliveries. Postnatal differences favoring the intervention group included fewer health care encounters for the infants relating to injuries and ingestions and, in the Elmira sample, lower rates of child abuse and neglect. These results suggests that intended improvements in maternal caregiving were achieved. During the four-year period following program participation, Home Visited mothers had fewer pregnancies; more program mothers completed their education, became employed, and independent of welfare than control mothers. Cost-effectiveness analyses confirm the long-term economic benefits of this program; program costs are recovered within a few years and continue to accrue over the life of the mother and her offspring. These investigators acknowledge that the nature and extent of program effects vary across individual (Olds & Korfmacher, 1998b), familial (Cole, Kitzman, Olds, & Sidaro, 1998), and program process (Korfmacher, Kitzman, & Olds, 1998) characteristics.

It is important to note that similarly broad and significant effects have not been reported for approaches to home visitation (Olds & Kitzman, 1993). Olds (personal communication, April, 1998) supports my impression that what differentiates his approach from its counterparts is not only reliance on licensed nurses as the providers (the Denver replication specifically examines the differential impact of nurse and paraprofessional provider) but its clear intent to produce lasting changes in the mother. In effect, an essential component of the Home Visitation program is for the visitor to make herself unnecessary, that is, to achieve the consultant's goal of putting oneself out of business. Echoing the commitment of the Child Guidance Movement (Mannino & Shore, 1986), the program enabled the mothers both to solve immediate problems and to anticipate and thereby avoid or solve future problems. My impression is that Olds' approach empowers the mothers and embodies the principle reflected in the adage "give me a fish and I eat for a day. Teach me to fish and I eat for a lifetime."

6.13.10 FINAL COMMENT

The position presented in this chapter is that significant health and mental health goals can be achieved through the development, evaluation, and dissemination of theory-based, scientifically validated, and politically and economically supported preventive interventions delivered through the application of indirect/consultation strategies. A review of the history of human service strategies over the past century and particularly since passage of the Community Mental Health Centers Act documents repeated attempt to pursue this goal. Good intentions and sincere commitment, however, have not been sufficient to place this effort into the fabric of the human and social service network. For that to occur, I am arguing, requires a combination of theoretical justification, flexibility in professional attitudes and roles, and viable examples of effective applications. All of those conditions appear to be presently available! Ideally, this chapter will contribute in some small way to the discipline's recognition and seizure of this opportunity.

6.13.11 REFERENCES

Albee, G. W. (1959). *Mental health manpower trends*. New York: Basic Books.

Albee, G. W. (1982). Preventing psychopathology and promoting human potential. *American Psychologist, 37,* 1043–1050.

Albee, G. W. (1986). Advocates and adversaries of prevention. In M. Kessler & S. E. Goldston (Eds.), *A decade of progress in primary prevention* (pp. 309–332). Hanover, NH: University Press of New England.

Allport, F. H. (1955). *Theories of perception and the concept of structure*. New York: Wiley.

Barker, R. G. (1968). *Ecological psychology*. Stanford, CA: Stanford University Press.

Barker, R.G., & Gump, P. V. (1964). *Big school, small school: High school size and student behavior*. Stanford CA: Stanford University Press.

Barker, R.G., & Schoggen, P. (1973). *Qualities of community life: Methods of measuring environment and behavior applied to an American and an English town*. San Francisco: Jossey-Bass.

Barker, R.G., & Wright, H.F. (1951). *One boy's day*. New York: Harper & Row.

Bell, R. Q. (1986). Age specific manifestations in changing psychosocial risk. In D. C. Farran & J. D. McKinney (Eds.), *The concept of risk in intellectual and psychosocial development* (pp. 169–185). New York: Academic Press.

Bennett, C. C., Anderson, L. S., Cooper, S., Hassol, L. Klein, D.C., & Rosenblum, G. (1966). *Community psychology: A report on the Boston conference on the education of psychologists for community mental health*. Boston: Department of Psychology, Boston University.

Berkowitz, L. (1993). *Aggression: Its causes, consequences and control*. New York: McGraw-Hill.

Bloom, B. L. (1975). *Changing patterns of psychiatric care*. New York: Human Sciences Press.

Bloom, B. L. (1977). *Community mental health: A general introduction*. Monterey, CA: Brooks/Cole.

Bloom, B. L. (1984) *Community mental health: A general introduction* (2nd ed.). Monterey, CA: Brooks/Cole.

Bronfenbrenner, U. (1977). Toward and experimental ecology of human development. *American Psychologist, 32,* 513–531.

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.

Caplan, G. (1964). *Principles of preventive psychiatry*. New York: Basic Books.

Caplan, G. (1970). *The theory and practice of mental health consultation*. New York: Basic Books.

Cicchetti, D., & Rogosh, F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology, 8,* 597–600.

Cole, R., Kitzman, H., Olds, D. L., & Sidaro, S. (1998). Family context as a moderator of program effects in prenatal and early childhood home visitiation. *Journal of Community Psychology, 26,* 37–48.

Commission on Chronic Illness (1957). *Chronic illness in the United States* (Vol. I). Published for the Commonwealth Fund. Cambridge, MA: Harvard University Press.

Cowen, E. L. (1973). Social and community interventions. *Annual Review of Psychology, 24,* 423–472.

Cowen, E. L. (1980). The wooing of primary prevention. *American Journal of Community Psychology, 8,* 258–284.

Cowen, E. L. (1982). Help is where you find it: Four informal helping groups. *American Psychologist, 37,* 385–395.

Cowen, E. L. (1996). The ontogenesis of primary prevention: Lengthy strides and stubbed toes. *American Journal of Community Psychology, 24,* 235–250.

Cowen, E. L., Gardner, E. A., & Zax, M. (Eds.) (1967). *Emergent approaches to mental health problems*. New York: Appleton-Century-Crofts.

Dawes, R. (1994). *House of cards: Psychology and psychotherapy built on myth*. New York: Free Press.

Dewey, J., & Bentley, A.F. (1949). *Knowing and the known*. Boston: The Beacon Press.

Dohrenwend, B. P., & Dohrenwend, B.S. (1981). Socioenvironmental factors, stress, and psychopathology. Part 1: Quasi-experimental evidence on the social causation-social selection issue posed by class differences. *American Journal of Community Psychology, 9,* 129–146.

Dohrenwend, B. S., & Dohrenwend, B. P. (1981). Socioenvironmental factors, stress, and psychopathology. Part 2: Hypotheses about stress processes linking social class to various types of psychopathology. *American Journal of Community Psychology, 9,* 146–159.

Dohrenwend, B. P., Levav, I., Shrout, P. E., Link, B. G., Skodol, A. E., & Martin, J. L. (1987). Life stress and psychopathology: Progress on research begun with Barbara Snell Dohrenwend. *American Journal of Community Psychology, 15,* 677–716.

Dumont, M. P. (1968). *The absurd healer: Perspectives of a community psychiatrist*. New York: Science House.

Durlak, J. E. (1979). Comparative effectiveness of paraprofessional and professional helpers. *Psychological Bulletin, 86,* 80–92.

Durlak, J. E., & Wells, A. M. (1997). Primary prevention mental health programs for children and adolescents: A meta-analytic review. *American Journal of Community Psychology, 25,* 115–152.

Erikson, E. (1963). *Childhood and society*. New York: Norton.

Fairweather, G. W., Sanders, D. H., Maynard, H., & Cressler, D. L. (1969). *Community life for the mentally ill*. Chicago: Aldine.

Gallesich, J. (1986). Training of mental health consultants. In M. F. V. Mannino, E. J. Trickett, M. F. Shore, M. G. Kidder, & G. Levin (Eds.), *Handbook of mental health consultation* (pp. 279–317). US Department of Health and Human Services (DHHS Publication No. ADM 86-1446). Washington, DC: US Government Printing Office.

Glidewell, J. C., & Swallow, C. S. (1969). *The prevalence of maladjustment in elementary schools: A report prepared for the Joint Commission on the Mental Health of Children*. Chicago: University of Chicago Press.

Gold, M. (1987). Social ecology. In H. C. Quay (Ed.), *Handbook of juvenile delinquency* (pp. 62–105). New York: Wiley.

Goldstein A. P., & Keller, H. (1987). *Aggressive behavior: Assessment and intervention*. New York: Pergamon.

Goldston, S. E. (1986). The federal scene: Ten years later. In M. Kessler & S. E. Goldston (Eds.), *A decade of progress in primary prevention* (pp. 363–388). Hanover, NH: University Press of New England.

Gordon, R. (1983). An operational classification of disease prevention. *Public Health Reports, 98,* 107–109.

Gordon, R. (1987). An operational classification of disease prevention. In J. A. Steinberg & M. M. Silverman (Eds.), *Preventing mental disorders: A research perspective* (pp. 20–26). Rockville, MD: Department of Health and Human Services.

Gurin, G., Veroff, J., & Feld, S. (1960). *Americans view their mental health*. New York: Basic Books.

Heller, K. (1989). Ethical dilemmas in community intervention. *American Journal of Community Psychology, 17,* 367–378.

Heller, K., & Monahan, J. (1977). *Psychology and community change*. Homewood, IL: The Dorsey Press.

Heller, K., Price, R. H., Reinharz, S., Riger, S., & Wandersman, A. (1984). *Psychology and community change: Challenges for the future*. Homewood, IL: The Dorsey Press.

Hollingshead, A. B., & Redlich, F. C. (1958). *Social class and mental illness: A community study*. New York: Wiley.

Iscoe, I., Bloom, B. L., & Spielberger, C. D. (1977). *Community psychology in transition: Proceedings of the National Conference on Training in Community Psychology*. New York: Hemisphere.

Joint Commission on Mental Illness and Health (1961). *Action for mental health*. New York: Basic Books.

Kandel, D. B., Single, E., & Kessler, R. (1976). The epidemiology of drug use among New York state high school students: Distribution, trends and changes in rates of use. *American Journal of Public Health, 66,* 43–53.

Kellam, S. G., Branch, J.D., Agrawal, K. C., & Ensminger, M.E. (1975). *Mental health and going to school: The*

Woodlawn program of assessment, early intervention, and evaluation. Chicago: University of Chicago Press.

Kennedy, J. F. (1963). *Message from the president of the United States relative to mental illness and mental retardation*. (88th Congress, First Session, U.S. House of Representatives Document No. 58). Washington, DC: US Printing Office.

Kessler, M., & Goldston, S. E. (Eds.) (1986). *A decade of progress in primary prevention*. Hanover, NH: University Press of New England.

Klein, D. C., & Goldston, S. E. (Eds.) (1977). *Primary prevention: An idea whose time has come*. Department of Health, Education and Welfare Publication No. (ADM) 77–447. Washington, DC: US Government Printing Office.

Korfmacher, J., Kitzman, H., & Olds, D. L. (1998). Intervention process as predictors of outcomes in a preventive home-visitation program. *Journal of Community Psychology, 26,* 49–64.

Lamb, H. R., & Zusman, J. (1979). Primary prevention in perspective. *American Journal of Psychiatry, 136,* 12–17.

Lamb, H. R., & Zusman, J. (1981). A new look at primary prevention. *Hospital and Community Psychiatry, 32,* 843–848.

Levine, M. (1969). Some postulates of community psychology practice. In F. Kaplan & S. B. Sarason (Eds.), *The psycho-educational Clinic papers and research studies*. Springfield, MA: Department of Mental Health.

Levine, M., & Perkins, D. V. (1987). *Principles of community psychology: Perspectives and applications*. New York: Oxford University Press.

Levine, M., & Perkins, D. V. (1997). *Principles of community psychology: Perspectives and applications* (2nd ed.). New York: Oxford University Press.

Levine, F. J., & Rosich, K. J. (1996). *Social causes of violence: Crafting a science agenda*. Washington, DC: American Sociological Association.

Lewin, K. (1935). *A dynamic theory of personality: Selected papers* (D. K. Adams trans.). New York: McGraw-Hill.

Lewin, K. (1944). Constraints in psychology and psychological ecology. *University of Iowa Studies in Child Welfare, 20,* 17–20.

Lilienfeld, A. M., & Lilienfeld, D. E. (1980). *Foundations of epidemiology*. New York: Oxford University Press.

Lochman, J. E. (1992). Cognitive-behavioral intervention with aggressive boys: Three-year follow-up and preventive effects. *Journal of Consulting and Clinical Psychology, 60,* 426–432.

Lorion, R. P. (1973). Socioeconomic status and traditional treatment approaches reconsidered. *Psychological Bulletin, 79,* 263–270.

Lorion, R. P. (1974). Patient and therapist variables in the treatment of low-income patients. *Psychological Bulletin, 81,* 344–354.

Lorion, R. P. (1983). Evaluating preventive interventions: Guidelines for the serious social change agent. In R. D. Felner, L. A. Jason, J. N. Moritsugu, & S. S. Farber (Eds.), *Preventive psychology: Theory, research, and practice in community interventions* (pp. 251–272). New York: Pergamon.

Lorion, R. P. (1987). The other side of the coin: The potential for negative consequences of preventive interventions. In J. A. Steinberg & M. M. Silverman (Eds.), *Preventing mental disorders: A research perspective* (pp. 243–250). (DHHS Publication No. ADM 87-1492). Washington, DC: US Government Printing Office.

Lorion, R. P. (1990). Basing preventive interventions on theory: Stimulating a field's momentum. In R. P. Lorion (Ed.), *Protecting the children: Strategies for optimizing emotional and behavioral development* (pp. 7–32). New York: Haworth Press.

Lorion, R. P., Brodsky, A., & Cooley-Quille, M. (1998). Exposure to urban violence: Responding to the contaminating effects of risk settings. In D. E. Biegel & A. Blum (Eds.), *Innovation in practice and service delivery across the lifespan* (Vol. 1). New York: Oxford University Press.

Lorion, R. P., Price, R. H., & Eaton, W. W. (1989). The prevention of child and adolescent disorders: From theory to research. In D. Shafer, I. Phillips, & N. B. Enzer (Eds.), *Prevention of mental disorders, alcohol and other drug use in children and adolescents* (pp. 55–96). OSAP Prevention Monograph - 2 (DHHS Publication No. ADM 89-1646). Washington, DC: US Government Printing Office.

Lounsbury, J., Cook, M., Leader, D., & Mears, E. (1985). A critical analysis of community psychology research. In E. Susskind & D. Klein (Eds.), *Community research: Methods, paradigms and applications*. New York: Praeger.

Mannino, F. V., & Shore, M. F. (1986). History and development of mental health consultation. In F. V. Mannino, E. J. Trickett, M. F. Shore, M. G. Kidder, & G. Levin (Eds.), *Handbook of mental health consultation* (pp. 3–28). US Department of Health and Human Services (DHHS Publication No. ADM 86-1446). Washington, DC: US Government Printing Office.

Miller, G. A. (1969). Psychology as a means of promoting human welfare. *American Psychologist, 24,* 1063–1075.

Moos, R. H., & Insel, P. M. (Eds.) (1974). *Issues in social ecology*. Palo Alto, CA: National Press Books.

Mrazek, P. J., & Haggerty, R. J. (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington, DC: National Academy Press.

Mrazek, P. J., & Hall, M. (1997). A policy perspective on prevention. *American Journal of Community Psychology, 25,* 221–226.

Olds, D. L., Henderson, C., Kitzman, H., Eckenrode, J., Cole, R., & Tatelbaum, R. (1998). The promise of home visitation: Results of two randomized trials. *Journal of Community Psychology, 26,* 5–22.

Olds, D. L., & Kitzman, H. (1993). Review of research on home visiting. *The Future of Children, 3,* 51–92.

Olds, D. L., & Korfmacher, J. (1997). The evolution of a program of research on prenatal and early childhood home visitation: Special issue introduction. *Journal of Community Psychology, 25,* 1–8.

Olds, D. L., & Korfmacher, J. (1998a). Findings from a program of research on prenatal and early choldhood home visitation: Special issue introduction. *Journal of Community Psychology, 26,* 1–4.

Olds, D. L., & Korfmacher, J. (1998b). Maternal psychological characteristics as influences on home visitation contact. *Journal of Community Psychology, 26,* 23–36.

Orford, J. (1992). *Community psychology: Theory and practice*. Chichester, UK: Wiley.

Phares, E. J., & Trull, T. J. (1997). *Clinical psychology* (5th ed.). Pacific Grove, CA: Brooks/Cole.

Price, R. H. (1983). The education of a prevention psychologist. In R. D. Felner, L. A. Jason, J. N. Moritsugu, & S. S. Farber (Eds.), *Preventive psychology: Theory, research, and practice in community interventions* (pp. 290–296). New York: Pergamon.

Price, R. H. (1997). In praise of a cumulative prevention science. *American Journal of Community Psychology, 25,* 169–176.

Price, R. H., & Lorion, R. P. (1989). Prevention programming as organizational reinvention: From research to implementation. In D. Shafer, I. Phillips, & N. B. Enzer (Eds.), *Prevention of mental disorders, alcohol and other drug use in children and adolescents* (pp. 97–123). OSAP Prevention Monograph - 2 (DHHS Publication No. ADM 89-1646). Washington, DC: US Government Printing Office.

Rappaport, J. (1977). *Community psychology: Values, research and action*. New York: Holt, Rinehart &

Winston.

Riessman, F. (1965). The "helper-therapy" principle. *Social Work, 10,* 27–32.

Riessman, F. (1967). Strategies and suggestions for training nonprofessionals. *Community Mental Health Journal, 3,* 103–110.

Rouse, B. A. (1995). *Substance abuse and mental health statistics sourcebook*. SAMHSA DHHS Publication No. (SMA) 95–3064. Washington, DC: US Government Printing Office.

Routh, D. (1998). The internship crisis. *The Clinical Psychologist, 51,* 1–2.

Rutter, M. (1989). Pathways from childhood to adult life. *Journal of Child Psychology and Psychiatry, 30,* 23–51.

Rutter, M., Dunn, J., Plomin, R., Siminoff, E., Pickles, A., Maughan, Ormel, J., Meyer, J. & Eaves, L. (1997). *Development and Psychopathology, 9,* 335–364.

Sameroff, A. J. (1977). Concepts of humanity in primary prevention. In G. W. Albee & J. M. Joffe (Eds.), *The issue: An overview of primary prevention* (pp. 42–64). Hanover, NH: University Press of New England.

Sameroff, A. J., & Chandler, M. J. (1975). Reproductive risk and the continuum of caretaking casualty. In F. D. Horowitz, M. Hetherington, S. Scarr-Salapatek, & G. Siegel (Eds.), *Review of child development research, 4,* (pp. 187–244) Chicago: University of Chicago Press.

Sameroff, A. J., & Fiese, B. H. (1989). Conceptual issues in prevention. In D. Shaffer, I. Phillips, & N. B. Enzer (Eds.), *Prevention of mental disorders, alcohol and other drug use in children and adolescents* (pp. 23–54). Rockville, MD: Office for Substance Abuse Prevention. DHHS Publication No. (ADM) 89-1646.

Sarason, S. B. (1972). *The creation of settings and the future societies*. San Francisco: Jossey-Bass.

Sarason, S. B. (1981). *Psychology misdirected*. New York: Free Press.

Sarason, S. B. (1985). *Caring and compassion in clinical practice*. San Francisco: Jossey-Bass.

Sarason, S. B. (1988). *The making of an American psychologist: An autobiography*. San Francisco: Jossey-Bass.

Sarason, S. B. (1990). *The predictable failure of educational reform: Can we change course before it's too late?* San Francisco: Jossey-Bass.

Sarason, S. B. (1993). *The case for change: Rethinking the preparation of educators*. San Francisco: Jossey-Bass.

Sarason, S. B. (1996). *Revisiting the culture of the school and the problem of change*. New York: Teachers College Press.

Sarason, S. B., Levine, M., Goldenberg, I. I., Cherlin, D. L., & Bennett, E. M. (1966). *Psychology in community settings: Clinical, educational, vocational, social aspects*. New York: Wiley.

Sarason, S. B., Zitnay, G., & Grossman, F. (1972). *The creation of a community setting*. Syracuse: Division of Special Education and Rehabilitation, Syracuse University Press.

Shoggen, P. (1989). *Behavior settings: A revision and extension of Roger G. Barker's Ecological Psychology*. Stanford, CA: Stanford University Press.

Silverman, P. R. (1988). Widow-to-widow: A mutual help program for the widowed. In R. H. Price, E. L. Cowen, R. P. Lorion, & J. Ramos-McKay (Eds.), *Fourteen ounces of prevention: A casebook for practitioners* (pp. 175–186). Washington, DC: American Psychological Association.

Sobey, F. (1970). *The nonprofessional revolution in mental health*. New York: Columbia University Press.

Srole, L., Langner, T. S., Michael, S. T., Opler, M. K., & Rennie, T. A. C. (1962). *Mental health in the metropolis: The midtown Manhattan study*. New York: McGraw-Hill.

Sroufe, L. A. (1997). Psychopathology as an outcome of development. *Development and Psychopathology, 9,* 251–268.

Tyler, F. T., Pargament, K. I., & Gatz, M. (1983). The resource-collaborator role. *American Psychologist, 38,* 488–498.

Vega, W. A., Valle, R., Kolody, B., & Hough, R. (1987). The Hispanic social network prevention intervention study: A community-based randomized trial. In R. F. Munoz (Ed.), *Depression prevention. Research directions* (pp. 217–231). Washington, DC: Hemisphere.

Wandersman, A., & Hess, R. (1985). *Beyond the individual: Environmental approaches and prevention*. New York: Haworth.

Wicker, A. W. (1979). *An introduction to ecological psychology*. Monterey, CA: Brooks/Cole.

Wolchik, S. A., West, S. G., Westover, S., Sandler, I. N., Martin, A., Lustig, J., Jenn-Yun Tien, & Fisher, J. (1993). The Children of Divorce parenting intervention: Outcome evaluation of an empirically based program. *American Journal of Community Psychology, 21,* 293–332.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.14
Working with Images in Clinical Psychology

ANN HACKMANN
University of Oxford, Warneford Hospital, UK

6.14.1 INTRODUCTION 301
6.14.2 AN OVERVIEW OF THE USE OF IMAGERY IN THERAPY 302
6.14.3 TYPES OF IMAGES IN PSYCHOLOGICAL DISORDERS 304
6.14.4 WORKING WITH IMAGES IN BEHAVIOR THERAPY 307
6.14.4.1 Desensitization 307
6.14.4.2 Covert Conditioning 307
6.14.4.3 Implosion and Imaginal Flooding 308
6.14.4.4 Reliving Treatment in Post-traumatic Stress Disorder 309
6.14.4.5 Eye Movement Desensitization and Reliving 309
6.14.4.6 Thought Stopping 310
6.14.5 WORKING WITH IMAGES IN COGNITIVE THERAPY 310
6.14.5.1 Images and Memories in Cognitive Therapy 310
6.14.5.2 Eliciting Images and their Meanings 311
6.14.5.3 Transforming Images and their Meanings 312
6.14.5.4 Working with Beliefs about Having Images, Intrusive Memories, and Nightmares 313
6.14.5.5 Working with Core Beliefs Using Imagery Techniques 313
6.14.6 CONCLUSION 315
6.14.7 REFERENCES 316

6.14.1 INTRODUCTION

Images are defined as contents of consciousness that possess sensory qualities, as opposed to those which are purely verbal or abstract. This definition embraces dreams and waking images of various kinds. Whilst images can have qualities associated with any of the sensory modalities, it seems that visual imagery is the most common (Horowitz, 1970). Memories, too, can come to awareness as mental contents with sensory qualities, most vividly in the case of "flash bulb" memories, and also in intrusive memories and reliving experiences following trauma. Memories themselves are not reexcitations of fixed traces, but are imaginative reconstructions (Bartlett, 1932) that can be altered or even partly created by varying retrieval procedures (Loftus & Palmer, 1974). These observations give weight to the idea that some memories are in fact more like imaginative images, although these may be as vivid and convincing as "real" memories. However, this distinction may be a difficult one to make, since it is also true (as discussed later in this chapter) that imaginative images themselves contain much input from memory, even though the person having the image may not be aware of the connections with their past.

Among the contents of consciousness are verbal thoughts and images in various modalities, some of which are considered to be memories. One might wonder why one would choose to work on imagery rather than purely verbal thought. Two possible reasons are that images condense a great deal of information, and they reveal layers of meaning which are idiosyncratic. It has also been suggested (e.g., Paivio, 1986; Singer, 1974) that imagery via sensory modalities is one of several major coding systems the brain has for organizing and storing experience, and that images are likely to afford greater emotional arousal than the verbal encoding system. Buzan and Buzan (1997) suggest that images are often more evocative than words, and are more precise and potent in triggering a wide range of associations. Thus, changing or transforming an image might be hypothesized to be more likely to bring about a significant emotional shift than challenging a verbal thought.

When dramatic, life altering events occur people tend to remember a great deal of detail about the circumstances in which the event occurred. Such memories are called "flashbulb memories," and it almost appears as if the mind takes a kind of photograph of the situation at such times. An example would be the way in which everyone can remember where they were when the death of John F. Kennedy occurred (Ornstein, 1992). Work in post-traumatic stress disorder (PTSD; Ehlers, 1997) suggests that during traumatic incidents processing may be data-driven rather than semantic and memories may be in a fragmented, sensory form, with associated high affect when they are recalled. Such memories are often deliberately avoided by the client, although they reappear as intrusive memories, nightmares, and spontaneous experiences of reliving. Brewin, Dalgleish, and Joseph (1996) have suggested that there are two types of memory, namely verbally accessible memories and situationally accessible memories (SAMS). The latter are not verbally accessible, but may be triggered by sensory environmental cues or emotions which match those present during the trauma. Several other writers have suggested that there may be a number of different representation systems that are involved in the processing of information (cf the Interacting Cognitive Subsytems approach, Barnard & Teasdale, 1991; and the SPAARS model, Power & Dalgleish, 1997).

Following resolution of trauma, memories seem to be stored in a more orderly, narrative form which is less fragmented (Foa, Molnar, & Cashman, 1995) and they are less likely to be upsetting to recall, or to be evoked involuntarily by sensory cues. Such observations support the idea that there may be several ways of encoding material, and there may be a place in therapy for bringing more fragmented imaginal material into awareness in order to integrate it with other material available in a verbal store. There are parallels here with what have been referred to as primary and secondary process thinking (Freud, 1962).

There are other properties of images which affect their usefulness in therapy. They may occur spontaneously, but they can also be deliberately evoked, transformed, or suppressed. They can reflect perspectives on the past, the present, or the future, and can be literal or symbolic. They can appear to be veridical memories, or at the other end of the scale they can be seen as imaginative images. However, their content is neccessarily determined by experience.

6.14.2 AN OVERVIEW OF THE USE OF IMAGERY IN THERAPY

In the history of psychology "mental contents" have at times been seen an appropriate area of study (e.g., by the early introspectionists, such as William James, Galton, and Titchener), and at times relegated to the essentially unknowable "black box" (e.g., by Skinner and Watson). Indeed, Brown (1958) expressed a prevailing attitude by writing that "In 1913 John Watson mercifully closed the bloodshot inner eye of American psychology. With great relief the profession trained its exteroceptors on the laboratory animal." However, early in the history of behavior therapy imaginal material crept back into the arena, and was used instead of *in vivo* material, for example, in various forms of exposure therapy, such as Wolpe's desensitization (Wolpe, 1958) or Stampfl's implosion therapy (Stampfl & Levis, 1967) and imaginal flooding (e.g., Foa, Skeketee, & Grayson, 1985). Images have also been used in other behavioral treatments, such as various types of covert conditioning (Cautela, 1977). Wolpe (1958) legitimized the use of covert processes within a behavioral framework through the assumption that if someone is led to imagine particular events this can lead to predictable changes in behavior. Jacobsen (1942) had also paved the way for this work, by demonstrating that simply imagining a situation like running produced responses in the body such as those produced by real running. It had also been shown that imagining a feared situation produced evidence of the "fight or flight" reflex, with changes, for example, in the galvanic skin response (GSR). In behavior therapy, images have been used to provide attenuated forms of both stimuli and responses.

Despite the use of imaginal material the main therapeutic techniques used in behavior therapy were thought to have been derived from learning theory, and to involve classical and operant conditioning, deconditioning, and extinction. The cognitive therapies took a fresh approach by emphasizing the fact that between stimulus and response there appear to be appraisals, which can be worked on directly, with profound effects on behavior and emotion. These appraisals include "meanings, significances, and imagery" (Beck, 1970), and we are advised that in cognitive therapy we should attend to thoughts and images (Beck, Laude, & Bonhert, 1974; Ottaviani & Beck, 1987). More specifically, Beck (1976) held that meanings were the things we needed to change, and that these are accessible through thoughts and images. The picture is further complicated by the fact that not all appraisals are necessarily in full awareness or reported spontaneously, although they are potentially accessible to consciousness.

Also within a cognitive therapy framework, Freeman (1981) and Freeman and Boyll (1992) discuss the way in which the images in dreams and fantasies provide access to the individual's idiosyncratic interpretations of reality, and distortions therein. This is also true where visual imagery accompanies strong emotion, and such images can form a basis for the reexamination and restructuring of beliefs.

In an interesting paper on the use of Gestalt techniques in cognitive therapy, Edwards (1989) briefly reviews the use of directed visualization techniques in Tibetan Buddhism, Jung's active imagination (Watkins, 1976), Assagioli's (1965) guided fantasy in psychosynthesis, Leuner's (1984) guided affective imagery, and the imagery psychodrama of Perls and other Gestalt therapists (Perls, 1971; Stevens, 1971). He also describes its application to the treatment of depression (Shultz, 1984) and phobias (Habeck & Sheik, 1984), and in controlling pain (Bresler, 1984) and increasing resistance to cancer (Hall, 1984). Much fuller reviews of many imagery techniques and their theoretical backgrounds have been given by Singer (1974) and Samuels and Samuels (1975). It is beyond the scope of this chapter to review all previous attempts to treat physical and psychological disorders using imagery techniques, so the main focus will be on the major work done in this area by clinical psychologists, working in the cognitive-behavioral framework, with the stance of the scientist practitioner. However, as Edwards points out, work in different theoretical frameworks is still worth studying, since despite its dissimilar surface characteristics the cognitive restructuring that takes place through guided imagery has close affinity to that which takes place through guided discovery in cognitive therapy, and techniques such as those used by Gestalt therapists, for example, can offer useful methods of identifying key cognitions (Arnkoff, 1981).

For this reason it may be useful to consider some of the earlier uses of visualization in psychiatry and clinical psychology. The first clinician credited with eliciting images to treat patients was Breuer (see Freud & Breuer, 1955). In 1881, when Freud was his student, he treated Anna O. by hypnotizing her, and encouraging her to produce what he called fancies, or daydreams. Freud describes the way in which her symptoms of hysterical paralysis would disappear, when in hypnosis she was made to remember the situation and associative connections under which they first appeared, provided free expression was given to the emotions these evoked. Later Freud applied similar techniques, but he abandoned the use of hypnosis when he realized that patients who could not be hypnotized could still produce similar types of early traumatic memory images and affect. Initially these occurred in response to the suggestion that a memory would come to mind when Freud put his hand on the patient's head, but this was later altered to an emphasis on free association and the analysis of dreams and symptomatic acts. Freud considered images to be part of primary process thinking, more primitive than ordinary thinking, developed earlier in childhood, and more direct, immediate, and less restrained by logic. He also felt that memory images were often not lost to the person, but stored in the unconscious.

In hypnosis, both imaginative and memory images may be elicited, again whilst the patient is in a deeply relaxed state. Age regression techniques are used to evoke memory images, from earlier in the patient's life, or even in some cases from supposed past lives. These can be extremely vivid and realistic, and this capacity of people in relaxed states to be able to virtually relive certain experiences has been noted by many therapists. Grof (1976) has described such experiences in his patients who explored altered states of consciousness under the effects of LSD, or during holotropic breathwork sessions, following hyperventilation and listening to music whilst physically relaxed. Where memory images are evoked using such methods they may encompass sensory material in any or all modalities, and associated affects. The same is true with images that can be recognized as imaginative images, leaving a gray area where it is impossible to say to what extent a veridical memory has been evoked, and thus making it important to consider the idea that some memories may be "false" ones, a problem which Freud himself encountered.

In the section on working with images in cognitive therapy some attention will be paid to the restructuring of memories, a topic which also has relevance in the treatment of PTSD, where treatments based on the repeated reliving of traumatic memories seem to hold the most promise (e.g., Foa, Rothbaum, Riggs, & Murdock, 1991). Edwards (1990) has written a masterly review of work on early traumatic memories in various theoretical frameworks, including the Freudian approach. Freud was initially very interested in early memories, but later moved away from his work on this topic, partly because he came to believe that some of the memories reported were in fact fantasies or imaginative images, and partly because in any case he believed that simply uncovering and analyzing traumatic memories would only be useful if accompanied by a wish to use these discoveries as a basis for changing things in the present. This second notion would be agreeable to cognitive therapists working in the 1990s. Bowlby (1985) criticised analysts for steering away from work on early memories, recognizing that recovery of a traumatic memory can allow distorted schemas to be recognized and restructured. Arieti (1985) reformulated the Freudian approach, and based his new approach on a cognitive developmental model. The techniques in use in the 1990s in cognitive therapy will be described below.

Jung gave great importance to the use of symbolic or imaginative images in therapy. He believed that images are often archetypal in nature, and that there is a universal language of symbols, common across different cultures and periods of history. Jung wrote that he attempted to translate emotions into images, to find the images which were concealed in the emotions. He felt that if emotion-laden images were left in the unconscious they could give rise to neurotic symptoms, but the more they could be brought into consciousness the calmer the person would become. Jung developed a technique called active imagination, during which the person is asked to meditate, remaining free of any goal. They then invite images to appear, and watch them without intervening. It is also possible to talk to the images, or to question them. Subsequently these visualizations are discussed with the therapist. This technique is more likely to evoke imaginative rather than memory images. Many of the images symbolize aspects of personal growth. Jung felt that simply bringing images to awareness and reflecting on them could be therapeutic, as Freud originally believed would be the case with memory images.

Imaginative images have been used in a variety of other therapies. Many of these have been more directive in their approach than Jung. The content of the image may be specified to some extent, and suggestions made for ways in which the image may be manipulated or controlled. An example would be Leuner's Guided Affective Imagery (Leuner, 1969). The therapist suggests a theme that the patient then works with. They may be asked to visualize a scene such as a mountain, a meadow, a house, or a stream, or they may evoke an image of a relative, a sexual partner, or someone of their own sex. Within the imaginary scene they could be prompted to make a journey, and then deal with the symbolic material which appears. This may involve confrontation, reconciliation, or other ways of dealing with what comes up in the imagery. The theory behind the technique is that the material elicited will symbolize important areas of conflict in the person's life. Once again the imaginary material is evoked following a period of relaxation in a darkened room.

Assagioli (1965) has used symbolic visualization as part of the therapy in psychosynthesis. The patient closes their eyes and relaxes, and is then asked to dwell on symbols of integration and balance, of harmony in relationships, of the masculine and the feminine, and of affective states. Focusing on particular symbols in this way is supposed to bring about increased control in the inner and outer life. Sometimes it is suggested that the patient should transform the image, so that, for example, a seed becomes a tree, or a worm becomes a butterfly.

Other therapists have emphasized the importance of interacting with the elements of the imaginative image. For example, Gestalt therapists (Perls, 1971; Stevens 1971), working with an image or daydream of an imaginary scene, would suggest to the patient that they imagine themselves as each of the other characters in the image, and even as an animal, or an inanimate object, such as the stick which is beating a dog. In this way different perspectives are developed which may be useful in therapy.

Before the sections on working with images in behavioral and cognitive therapies a brief outline of phenomenology in the realm of imagery is provided. This is biased towards a study of imagery in neurotic disorders, but is worth noting that a great deal of work has been done on imagery in other fields, notably in physical medicine and in sports psychology.

6.14.3 TYPES OF IMAGES IN PSYCHOLOGICAL DISORDERS

The literature suggests that images abound in psychological states, perhaps particularly in anxiety disorders. Ottaviani and Beck (1987) revealed that images are common in people with panic disorder, and frequent themes are those of physical and mental catastrophe, as is typical of their anxious thoughts. Beck et al. (1974) looked at ideational components in anxiety neurosis, and found themes to do with danger, both physical and psychosocial, in the thoughts and images collected.

Images in phobias have been studied by several authors. Watts, Sharrock, and Trezise (1986) hypothesized that phobics would have imagery connected with their phobias that was relatively lacking in detail and elaboration. They discovered in a study of visual imagery of coping with a spider that the accounts of phobics were briefer and less elaborated, suggesting a lack of coping imagery. However, in another study of stimulus imagery immediately following exposure, phobic subjects reported being more aware of the image, but did not report less detail (as predicted). This was surprising in view of a previous study by Watts, Trezise, and Sharrock (1986) in which poor recognition memory for phobic stimuli was demonstrated and taken as evidence for poorly elaborated stimulus processing.

Beck (1976) reported on his studies of the content of dreams and other ideational material in people suffering with depression. These dreams reflect themes of defeat, defectiveness, thwarting, and deprivation, rather than the themes predicted by psychoanalysis, to do with hostility turned inwards. It was this work, an unsuccessful attempt to provide empirical support for psychodynamic views on depression, which led to the development of the cognitive model of depression. There are several studies on the themes in depressive dreams (Beck & Hurvich, 1959; Beck & Ward, 1961), but less has been written about depressive images. Layden (personal communication) also suggests that depression is characterized by a lack of positive images, particularly of the future, which might, if they existed, have the power to motivate the depressed individual to work towards realizing this hoped for vision.

De Silva (1986) has studied images in obsessive-compulsive disorder. He found, like the authors above, that in general the images were not different in content from other obsessional cognitions. The most common themes were those of death and decay, illness and injury, violence and disaster, and sex and blasphemy. He distinguishes between four different types of image: (i) obsessional ones, which are unwanted and intrusive; (ii) compulsive images, which are formed deliberately to neutralize an obsessional image, (which can be corrective versions of the original, or independent images bearing no relation to the obsessional one they are meant to neutralize); (iii) disaster images, of what might happen if the obsessions are not acted upon; and (iv) disruptive images, which can interrupt compulsive rituals, and make it seem imperative to start them again. The most interesting distinction here is probably the observation that images have different functions, some constituting obsessions (i.e., unwanted, intrusive images), and some being covert compulsions or rituals (i.e., images which are deliberately formed; see also Salkovskis & Westbrook, 1989). This highlights the observation that some images are seen as threatening or likely to lead to harmful consequences, whilst some are seen as capable of warding off harm.

Wells and Hackmann (1993) looked at images in health anxiety, and found that they contained a wealth of information. As with automatic thoughts (Warwick & Salkovskis, 1990) the themes expressed were those of misinterpretation of bodily sensations and symptoms, and overestimation of the likelihood of illness or death. However, the images also began to answer the question of what is so particularly bad about such calamities for people with health anxiety. For many the real calamity lay in the interpersonal consequences of illness or death (see Section 6.14.5). Like obsessional patients these individuals also had many superstitious or metaphysical beliefs, often involving fears that thinking in a particular way, or having certain images, could affect the future, or tempt providence. For example, some individuals fear that imagining death or illness could make it happen, or thinking of how lucky they are to be well, or picturing a happy future, could invite disaster. Sometimes the belief is like the popular view that negative thinking is bad for the immune system, but more frequently the mechanism is thought to be God or Fate. These ideas lead to attempts to control or neutralise images, in ways which may well exacerbate the problem, since suppressing thoughts and images often leads to an increase in intrusions.

In generalized anxiety disorder patients complain about worrying. Worry has been defined as consisting of long chains of negative thought that are predominantly verbal in form, and aimed at problem solving (Borkovec, Robinson, Pruzinsky, & DePree, 1983). Borkovec and Inz (1990) have suggested that the predominantly verbal activity of worrying diverts attention away from imaginal activity which is more closely linked to the activation of physiological arousal. Furthermore, they also suggest this may block exposure to it and lead to symptoms of failed emotional processing, including more intrusions of imaginal material. Where there is imaginal material typical themes are of psychosocial or physical danger (see Beck et al., 1974).

Wells, Ahmed, and Clark (1995) have also studied images in social phobics. Here, the finding was that the images of social phobics when deliberately generated depicted the self, seen in a distorted way, and as if seen from an external perspective. Thus, a social phobic tends to see themself as they imagine others might, with a bright red or sweaty face, or with shaking hands or a childish appearance. This also happens in spontaneous images social phobics get in naturalistic situations (Hackmann, Surawy, & Clark, 1998). Here the images are usually visual ones, but are sometimes sensory impressions rather than visual images. However, they are still perceived from an external view-point, in contrast with the images less socially anxious people get when anxious in a social situation. These images are often recurrent, and depict the individual in an embarrassing or humiliating predicament, with some awareness of an audience. When asked when in their lives they first remember feeling that way a specific incident is often recalled, in which they were criticised or bullied (Hackmann & Clark, in press). Images in social phobia have an important part to play in maintaining the disorder, in that like the images in other disorders they tend to be regarded as reflecting reality, and therefore they are often acted upon, giving rise to avoidant behaviours. They also form part of the postmortem, in which a negative, ruminative going over of past imagined and predicted future failures leads to a set of exaggerated and distorted images and memories and images, with a consequent enhancement of anxiety (Clark & Wells, 1995).

Agoraphobics often have recurrent images (Hackmann & Surawy, in press). In these the patients image themselves from an external viewpoint, experiencing the sort of physical or mental calamities feared by panic disorder patients. However, in their images they also see an interpersonal context in which the events are occurring. Typically they are either ignored and experience themselves as far from home or help, or they attract unwelcome attention, which is embarrassing, and may also have the effect of making it harder for them to reach the safety of home. For example, someone might picture themself being taken to hospital and kept there against their will, whist another might see themself lost and alone in an unreal world, where it has become impossible to reach home.

In therapy with patients suffering from a variety of anxiety disorders it has been noted that there is often close similarity between the sensory and interpersonal content of recurrent images and upsetting childhood memories (Hackmann, 1995; Hackmann & Clark, in press; Hackmann & Surawy, in press; Wells & Hackmann, 1993). It has also been reported that in depression individuals often have intrusive memories, linked to their current mood, of childhood traumas (Brewin, Hunter, Carroll, & Tata, 1996; Kuyken & Brewin, 1994; Spenceley & Jerrom, 1997). This suggests that some images may have more than a passing resemblance to reliving experiences in PTSD.

Images in PTSD tend to appear like flashbacks, often to scenes which have been especially traumatic. Some of them are so vivid that it is as if the patient is reliving the trauma, in rich sensory detail, as if it were happening again. Often the intrusions people get spontaneously are always of the same fragment (or fragments) of memory. One patient always saw the headlights of a car approaching her head-on, on the wrong side of the road. It was the headlights which she always saw when waking up in terror in the night, or when lying on the sofa and looking at the corner into which she had gazed as she spent weeks lying down, recovering from the accident. Often it appears that the fragments which intrude are those which have not been resolved in the mind of the patient. For example, the patient who described the headlights on the wrong side of the road linked to the idea that she might have been able to prevent the accident. Revisiting the scene of the accident helped her to realize that there was nothing she could have done, upon which the intrusions decreased markedly. Nightmares in PTSD often involve a total and accurate replay of the traumatic events. For example, someone who was attacked whilst recovering from surgery repeatedly dreamed that she saw the hands of the hospital clock pointing to 10 minutes to three, then experienced the attack again, and woke at exactly three a.m., the time at which the attack ended and she was safe again.

As mentioned above, images in PTSD can involve any of the senses. The first patient described above was able to actually hear the crunch of metal, smell the petrol and taste it, see the steam rising, and feel the actual impact. Imagining the crash gave her pain in her chest and neck, which had been hurt in the crash a year before. Sometimes these flashbacks are not completely veridical memories. For example, the same patient had been told that since the man who crashed into her was foreign he could not be prevented from coming back in the future and hiring another car. At one point this was represented in an image of him crashing into her again in the future, and walking up to talk to her in a way that had not happened in the real accident. She took this to mean that this was what would really happen in the future. Similar features are noted after sexual abuse—there is reliving, not only in the visual modality, but also of sounds, tastes, smells, and sensations, and there can be mixture of images and memories. There is also often fear in both disorders that letting these mental contents into full awareness could have serious mental or physical consequences. Once again this leads to attempts to suppress such intrusions, which can then perpetuate the problem.

6.14.4 WORKING WITH IMAGES IN BEHAVIOR THERAPY

6.14.4.1 Desensitization

In 1958, Wolpe published his famous book about desensitization. This was based on experiments on the counter-conditioning of fear responses in laboratory animals through the pairing of stimuli which had been paired previously with noxious stimuli with positive stimuli (e.g., Masserman's experiments with cats). Wolpe called the process "reciprocal inhibition," and paired the feared stimuli with relaxation, or with eating, or sexual stimuli. This process, when repeated a number of times, appeared to decrease the amount of fear evoked by the negative stimulus, and increase the amount of approach behavior. The feared stimuli were approached in a graded way, via a carefully arranged hierarchy of feared situations. Wolpe discovered that imaginal presentation of stimuli was easier to organize than *in vivo* presentation, and also that the most practical activity to pair with the stimulus was relaxation. Results with this method were very impressive, and in 1977 Smith and Glass concluded that desensitization-type behavior therapies produced larger average therapeutic effects than any other type of psychotherapy, at that time, particularly in the treatment of phobias.

Yet the important ingredients in the procedure remain uncertain. The use of a systematic hierarchy and progressive relaxation of the muscles have been shown not to be essential (Dawson & McMurray, 1978). Wilkins (1971) found that the research literature seemed to indicate that the significant ingredient of systematic desensitization and similar treatments is the use of imagery by the client. Weitzman (1967) interviewed patients after every session of desensitization. He found that during the period of silence in which the client held the image the experience was not a static one. Other images and associations tended to occur, and the image at the end of the period had often undergone some changes. Also, those who observed Wolpe conducting the therapy underscored the process of cognitive exploration and discussion which occurred. The relationship of this approach to learning theory has been questioned, by Breger and McGaugh (1965). Singer and Pope (1978) conclude that therapy seems to work by changing the client's anticipations, self-communications, and images with respect to the feared situations. The mechanism remains to be fully investigated.

6.14.4.2 Covert Conditioning

Cautela (1977) produced a useful review of the assumptions and procedures in covert conditioning, in which it was conceptualized in operant-conditioning terms. The main assumption was that behavior patterns could be strengthened or weakened by imagining them to be followed by either positive or negative reinforcement, or by punishment, or by an absence of the previous reinforcers. This was not the original way in which covert sensitization was conceptualized (Cautela, 1967), when it was seen more as a classical conditioning procedure, in which the objective to was the opposite from that of desensitization, in that the aim was to render a previously positive stimulus noxious, by pairing it with imaginal stimuli which are aversive, and thus decrease the chance that presenting the original stimulus will lead to undesired behaviors. By 1977 the procedure was considered to work best if the maladaptive behavior was imagined after the triggering stimulus, and then followed by negative consequences. Self-control or relief scenes were also added, and alternated with aversive scenes. A relief scene might involve having the temptation to carry out an undesirable behavior, but starting to feel ill, then deciding to resist the temptation, and then feeling immediately better, and proud to have resisted it.

The most extensive use of such techniques has been by Cautela, and in a series of papers he has reported its use in the treatment of many maladaptive approach behaviors, such as smoking, drinking, obesity, compulsive stealing, and deviant sexual behaviors (Cautela, 1966, 1967, 1971). Cautela has developed a set of graphic and disgusting images, which can be used as they are, or elaborated with the patient. These are paired with mental images of the desired stimulus or maladaptive behavior. For example, a patient might be asked to imagine a desired sexual partner in a disgustingly filthy cellar, which causes the patient to vomit (in imagination) all over himself and the desirable partner. The disgusting scene is terminated by his escape from the cellar, and the other individual. It is interesting to note that in many studies of covert sensitization progressive relaxation is used, not for the purpose of reciprocal inhibition, but to enhance the generation of vigorous imagery, or stronger emotional reactions. Covert response cost has also been described by Cautela (1970a). In this case the removal of an imagined positive reinforcer is contingent on a real or (more usually) imagined response, which can be either a maladaptive approach or avoidance response. This technique is relatively untested.

Covert positive reinforcement has also been used by Cautela (1970b). In this procedure the person rewards himself in imagination for the desired behavior in real life or in imagination. Covert-positive reinforcement can be used to reinforce both approach and avoidance behaviors. It is considered to be an operant procedure in which a covert reinforcer is used to increase probability. It need not involve relaxation or hierarchical presentation of material, so it differs in a number of respects from desensitization.

Cautela also describes a process he calls covert extinction, in which the client is asked to imagine himself carrying out the behavior he wishes to curtail, without it bringing any of the usual satisfactions. For example, a smoker might imagine himself smoking without being able to taste or smell the cigarette, and without any decrease in his level of stress. Cautela (1971) reports a few studies where these techniques were used successfully, and also warns of several aspects of treatment which the therapist and patient need to be aware of. Finally, Cautela describes covert modeling, in which the learning or alteration of new behaviors is effected by imagining others carrying out the behavior, and observing the imagined consequences. A shaping process can also be used, where the patient first imagines observing someone else carry out the behavior, then watches themself do it, then imagines actually doing it, and finally carries it out. There is experimental evidence to support the idea that this is a useful technique (Cautela, 1977; Kazdin 1974).

6.14.4.3 Implosion and Imaginal Flooding

This technique attracted considerable attention in the 1960s, following a theoretical paper by Stampfl and Levis (1967). Stampfl viewed the defense mechanisms of the neurotic as avoidance responses which had been learned and perpetuated because they reduced anxiety. Originally the patient, often when very young, has traumatic experiences, in which they has been rejected, humiliated, deprived, or punished. Subsequently everything associated with these events, such as the sights, sounds, and smells, elicit anxiety, and these cues motivate behavior. This behavior can be considered defensive if it reduces anxiety, and is consequently reinforced.

Stampfl theorized that these neurotic defensive behaviors could be unlearned if something very similar to the original trauma could be strongly recreated, in the absence of any real punishment, deprivation, and so on. The use of imaginal material proved a convenient vehicle for this, as the patient was often able to evoke the visual cues, and the therapist could provide suitable sound effects. It was hypothesized that this process of play-acting the fears, and reliving the trauma could lead to the extinction of anxiety. Thus, the theory spanned psychoanalytic and learning theory ideas about anxiety.

The first research study was by Hogan (1966). In it psychotics were given either implosive or more traditional psychotherapy. There was a significant difference between the two groups, with more of the implosive therapy patients being discharged after one year. Levis and Carrera (1967) reported on the effects of 10 hours of implosive therapy in neurotic patients, with a significant difference in two of the 16 MMPI scales compared with control groups. In a second study, Hogan and Kirchner (1967) studied rat phobics. Those in the implosive group were significantly more likely to pick up a rat in the behavior test than those who had been given only pleasant imagery. Hogan and Kirchner (1968) also reported a successful one-session intervention with snake phobics. In another study, Kirchner and Hogan (1966) administered the implosive therapy standardized tape simultaneously to a group of rat phobics in a single session, and played music on a standardized tape to the control group. The implosive group were superior to the control group in their willingness to pick up a rat. It is interesting that these results were obtained with such standardized material, in view of the original theory behind implosion, which aimed to extinguish the anxiety reaction to highly idiosyncratic sensory cues present at the time of the supposed original trauma. This suggests that the psychodynamic aspects of the theory may be redundant.

However, it is also worth noting that Levis (1980) remarks that one of the most interesting phenomena to occur with the use of implosion therapy is the apparent reactivation or redintegration of early memories during or following the presentation of imaginal scenes depicting the most feared outcomes possible in an individual's feared situations. The emergence of this new material helps shape the course of new imaginal sessions, and is thought to provide a closer approximation of the avoided conditioned stimulus complex from which one is attempting to extinguish anxiety. Several examples are provided in this chapter of the reemergence of traumatic early memories during therapy, and an example of the same phenomenon is described with reference to implosion therapy for a recuurent nightmare.

In the 1970s there were several studies comparing imaginal flooding with other behavioral treatments. For fearful volunteers and simple phobics *in vivo* exposure treatment generated better results than imaginal procedures (Bandura, Blanchard, & Ritter, 1969; Barlow, Leitenberg, Agras, & Wincze, 1969; Dyckman & Cowan, 1978; Mathews, 1978; Sherman, 1972). With obsessive compulsives both media produce similar outcomes (Foa et al., 1985; Rabavilas, Boulougouris, & Stelanis, 1976). With agoraphobics two studies indicate the superiority of *in vivo* exposure (Emmelkamp & Wessells, 1975; Stern & Marks, 1973) and two studies indicate equivalent gains (Chambless & Goldstein, 1982; Mathews et al., 1976).

Foa, Steketee, Turner, and Fischer (1980) looked more closely at this with a group of obsessive compulsives with checking rituals. In addition to their feared situations which they were exposed to *in vivo*, they also had disaster fears for the future which could only be presented in imagination. Half the subjects received two hours of *in vivo* exposure only. The other half received only half an hour of exposure, with 90 minutes of imaginal exposure to the feared disasters. Both groups improved significantly, to the same degree. However, the group that received the combined treatment retained their gains, whilst some relapse was evident in the *in vivo* only group, suggesting that more of the elements of the fear structures (Foa & Kozac, 1986) were addressed in the imaginal exposure than in the *in vivo* exposure group, leading to more successful emotional processing.

6.14.4.4 Reliving Treatment in Post-traumatic Stress Disorder

A technique which has proved relatively successful in the treatment of PTSD is that of repeated imaginal reliving of the trauma (Foa et al, 1991). After preparation and socialization as to the treatment rationale clients are asked to repeatedly relive their traumatic experiences, describing them aloud, in the first person, and attempting to reexperience exactly what happened to them. Particular attention is paid to the hot spots, that is, the most emotionally laden parts of the memory. Clients are asked to take home tape-recordings of these sessions, and listen to them daily if possible. This procedure frequently results in a decrease in the associated affect, particularly the emotion of fear. It also results in the accounts becoming less fragmented, more organized (i.e., indicating more attempts to understand the traumatic events), more complete, longer, and richer in thoughts and emotions, but with less detail about the events themselves (Foa et al., 1995). Overall this procedure also produces good symptomatic relief in real life. Also, there was greatest change in indices of fragmentation and organization in the narratives of those who improved the most.

These results are in line with Foa's theoretical position, that PTSD may occur in cases where there has been inadequate emotional processing, possibly because the terrible memories have somehow been held out of awareness (Foa & Kozac, 1986). Jaycox and Foa (1996) discuss possible obstacles in implementing exposure therapy in certain cases of PTSD. These include extreme anger, emotional numbing, and overwhelming anxiety. Practical suggestions are made, such as projecting the imagery on to an imaginary wall, and "zooming out" from the images if the distress of the client becomes extreme, to the point where the sense of current reality is about to be lost.

It remains to be seen whether the results of reliving treatment can be improved upon by giving the intervention a more complex rationale, and tackling some of the issues raised by using techniques such as those derived from cognitive therapy to challenge distorted ideas about the trauma itself, or its psychological sequelae.

6.14.4.5 Eye Movement Desensitization and Reliving

Shapiro (1989) published details of a new treatment for PTSD, which she called eye movement desensitization and reliving (EMDR). Subsequently she has also published a book (Shapiro, 1995) on the subject. The therapeutic technique was described as involving eliciting from the client sequences of large-magnitude, rhythmic saccadic eye movements, while holding in the mind the most salient elements of a traumatic memory. The client is asked to come up with a prototypical image which symbolizes the area of trauma. After being subjected to a set number of eye movements they are encouraged to talk about what remains of their concerns, and what has come up for them during this brief period of time. When they come to the end of discussing this a fresh image is evoked, symbolizing the imaginal or verbal material which has emerged. The process is then repeated until the client reports a significant reduction in the subjective units of distress (SUDS). This appears to result in rapid reduction in anxiety, changes in cognitive assessment of the memory, and cessation of flashbacks, intrusive thoughts, and sleep disturbance. In her book Shapiro writes about what she calls accelerated emotional processing, and there is some resonance with the work of Foa's group, described in the previous section. It is also interesting to recall that in Wolpe's work (described in the section on desensitization) merely bringing upsetting images to mind, and holding them in awareness, tended to alter both their content and their meaning, with a subsequent drop in affect.

However, the literature suggests that EMDR is a far from uncontroversial area. There are a number of review articles pointing out that many of the studies reported are characterized by serious methodological flaws, including failure to use repeated objective measures, and the use of other interventions in combination with EMDR (Acierno et al., 1994; Lohr, Kleinkrecht, Tolin, & Barrett, 1995). Nevertheless, there are many reports of rapid improvement with this technique, particularly where there are disturbing memories, and a recent meta-analysis found as much evidence to support this technique as there is for other efficacious approaches to PTSD (Van Etten & Taylor, 1997). Research questions which remain to be answered decisively include whether or not the rapid eye movements are a necessary component, whether or not relaxation facilitates the process, and whether any additional benefit is gained above merely repeating the presentation of the image.

6.14.4.6 Thought Stopping

It might appear that therapeutic strategies in which the patients are directed to immerse themselves in fearful imaginings, or relive distressing memories, would be in direct contradiction of an approach like thought stopping, which aims to suppress thoughts or images. However, this approach was used for a number of years, and was sometimes considered useful in the treatment of obsessions and other unwanted thoughts and impulses.

Thought stopping has been another controversial technique, with some review articles sounding a cautionary note. Where there have been positive results the emphasis has not been on blocking the obsessional cognitions as such, but on stopping before attempts to covertly neutralize the upsetting thoughts or images can occur. Thus, as a behavioral strategy the best use of thought stopping would be in the area of response prevention, that is, the prevention of covert rituals (Salkovskis & Westbrook, 1989). This would mean that in order to decide whether an imaginal exposure approach or a thought stopping approach was relevant one would need a functional analysis of the cognitions concerned to decide which strategies were appropriate for which parts of the problem.

6.14.5 WORKING WITH IMAGES IN COGNITIVE THERAPY

6.14.5.1 Images and Memories in Cognitive Therapy

Internal reality is said by Beck (1976) to comprise meanings, significances, and imagery, and we are advised to work both with images and thoughts. Sometimes in therapy an image seems to resonate more fully with an expressed emotion than the thought expressed in words. For example, the anxiety of an agoraphobic who thinks that they might fall down makes more sense to us when they describe an image of this fall, in which they see a crowd gather, summon an ambulance and take the patient to hospital, where they are kept against their will. This type of difference between the first negative automatic thoughts reported in a situation, and the spontaneous images which are described, has been exemplified in a paper by Wells and Hackmann (1993). In this paper the negative automatic thoughts and images about illness and death were studied in a group of hypochondriacs, and material presented highlights the extra, idiosyncratic layers of meaning depicted in the images. It is clear that for many hypochondriacs the feared catastrophe is not just illness or death, but the imagined interpersonal cost to them of these events (such as being neglected in hospital by relatives, or dying and ending up in Hell, or alone or trapped in a coffin with full awareness, or being able to witness one's family, carrying on quite happily without one).

Thus, attending to the content of images can help people get a "handle" on why they are getting upset (see Gendlin, 1988) and can also provide a quicker route in to what Barnard and Teasdale (1991) would call the implicational level of meaning which tends to be more poetic and more idiosyncratic than the less colorful, shared level of propositional meaning. It is also true that, as Lang (1977) points out, images contain both stimulus and response elements, and thus give access to more of any particular fear structure. Of course, when working with verbal material in cognitive therapy we can access more meaning of a more idiosyncratic nature with judicious use of questions such as "What did that mean to you?," "What was the worst thing about that?," and "Can you tell me a bit more about that?" These are the type of questions used in the "downward arrow" technique (e.g., Greenberger & Padesky, 1995), which is a standard technique in cognitive therapy to elicit the full meaning of the appraisal leading to the strong emotions reported in therapy. What is being suggested here is that spontaneous images accompanied by high affect can be a rapid way into the system.

In cognitive therapy there are a number of aspects of imagery one might choose to work on. These include unpacking the meaning encapsulated in the image, attempting to transform it, and examining and working on beliefs about the meaning of actually having the image. It is also possible to work on memories, which are imaged by the patient when asked to relive the remembered experience. Once again the meaning given to the memory can be explored, and sometimes transformed using imagery techniques.

6.14.5.2 Eliciting Images and their Meanings

Greenberger and Padesky (1995) advocate staying with questioning and careful reflection until cognitions arise which seem to be an adequate match for the level of emotion experienced. They pose a series of useful questions such as "What does this mean about me, my life, other people, the future?," "What am I afraid of?," "What might happen?," "What images and memories do I have in this situation?" Whilst doing this visual images or sensory impressions can occur in any modality, reflecting the present, the past, or the future, and these can be literal or metaphorical images. Any of these images can be used in cognitive therapy, in the sorts of ways described in subsequent sections.

It is also possible to enquire directly about the presence of an image. For example, in agoraphobia clients are often rather vague about the feared catastrophic consequences of their bodily symptoms of anxiety, whilst their avoidant behavior suggests that they are quite frightened. Asking them whether they had a picture of what might happen in a given situation often provides graphic information about the exact nature of the feared catastrophes. For example, one person might picture themselves as having collapsed in a crowded place, and being totally ignored, whilst another might have an image of a large crowd gathering and an ambulance arriving with a wailing siren.

Sometimes it is possible to spot when a client has a spontaneous image whilst describing something from the way they point, gesture, or direct their gaze. One woman with hypochondriasis was describing what she thought would happen when she was dead, and indicated with her hands how her Elvis flag would be draped across the coffin, whilst her family were grouped around it. This was a cue for her therapist to enquire whether she was picturing this at the time, which she was. This then led her say that she was, but that she habitually tried to suppress this image, as she feared that if she pictured it this would make it more likely to happen. This type of information is also useful for the therapist, as it is possible that her voluntary efforts to suppress the image are likely to make it recur more frequently, thus maintaining her distress, as in thought suppression experiments (Trinder & Salkovskis, 1993).

It is also important in some disorders, such as obsessions, to enquire about the existence of neutralising images. Sometimes these have the function of compulsions, in that they are evoked voluntarily in order to neutralize unwanted obsessional intrusive thoughts. For example, someone who had a spontaneous image of her husband being ill might deliberately try to form a positive image of him looking healthy and well. Such different functions and origins of the images have to be taken into account when planning a strategy, whether the treatment is a cognitive or a behavioral one.

Images are common in PTSD. In this case they are frequently fragments of disturbing memories, which press into conscious awareness at frequent intervals, to the distress of the individual. Whilst they are usually closely related to the trauma they are not always identical to it. For example, one man was traumatized by having to witness the body of a woman who had had her head chopped off. He had recurrent images of her husband swinging the chopper towards her, which he had not actually seen. However, the image seemed to have input from an earlier, linked memory of seeing his father attack his mother with chopping movements when he was a boy. The image of the murder was accompanied by the sense that he should have been able to prevent it, an inappropriate response since he was not there, nor did he know the people. However, the response was identical to the way he had felt as a boy. Images which appear to be fragments of memory, or which appear to have much input from memory in sensory modalities are often triggered by sensory cues, which incidentally can also sometimes trigger the distress without the patient consciously remembering the original event. This brings us back to the starting point of allowing enough time for the patient to stay with the strong emotion evoked, for any associated thoughts, images, and memories to come into awareness. Needless to say this can temporarily increase affect as meaning is accessed, and may be resisted at times by patients.

It is also important to note that images which seem in fact fragments of upsetting memories, or amalgamations of several of them, are also present in other Axis I disorders (Hackmann, 1995; Hackmann & Clark, in press; Hackmann & Surawy, in press; Wells & Hackmann, 1993). Other authors have also noted the intrusion of traumatic memories during depressive episodes (Brewin, Hunter, Carroll, & Tata, 1996; Kuyken & Brewin, 1994).

Once images have been elicited the therapist needs to help the patient understand the meaning attached to them. Images can be considered at the level of automatic thoughts, in which case the therapist can ask what the image means to the person, what has led up to the events in the image, and what is the worst thing about it. Typically this sort of question will give rise to a set of negative automatic thoughts, which can be worked on using verbal techniques, or imagery techniques as described below. It is also important to find out what sort of beliefs the patient has about actually having the image. For example, do they see it as a prediction of what will happen in the future, or do they think that such images can affect reality in some way? If so, such beliefs may also need to be worked on in therapy.

Further questions about what the image means about the individual, other people, and the world in a more general way can lead on to the elicitation of conditional and unconditional core beliefs, which can also be worked upon at a later stage.

Once all the meaning and sensory data in the image has been unpacked a useful bridge can be made back to early experience by asking when in life the client can first remember having the types of thoughts and feelings and sensory experiences reflected in the image under consideration. Often this takes the patient right back to traumatic memories from childhood which appear to have set the scene for later maladaptive schemas (e.g., Wells & Hackmann, 1993). For example, a patient with an obsessional fear of getting AIDS repeatedly visualized himself in hospital, in a white, cast-iron bed, all alone because he is blamed by others for contracting the disease. In the patient's mind there was a direct link between this recurrent image and a memory of lying in hospital with a broken arm, alone and not aware of the pain so much as the fear of his father blaming and rejecting him for having had an accident. To his amazement he realized that the bed in the hospital in the image and in the memory was in fact the identical white, cast-iron bed with the same white sheets and blankets. Having located significant memories in this way (or if they come up spontaneously in therapy) it is then possible to work on early memories using imagery techniques such as those described below, often with profound results.

As mentioned above, some of the images clients bring are symbolic rather than literal images. For example, someone feeling exhausted and resentful about the way in which their family depends on them may picture themself with their family in a canoe, having to row them all upstream in the Amazon. Similar questions can be used to explore the meaning of this type of image, and its possible history in the life of the individual, before proceeding to consider ways of transforming its meaning (Edwards, 1989). This work considers thematic rather than symbolic content of the images or dreams being presented, in the sense that the client provides their own hunches about the history and meanings of the images, rather than the therapist trying to interpret them in the light of what the therapist might consider to be the usual symbolic meanings (Freeman, 1981).

6.14.5.3 Transforming Images and their Meanings

The meanings elicited from consideration of the image can be worked on using the full range of verbal techniques and behavioral experiments familiar to the cognitive therapist. The encapsulated meanings can also be changed by altering the image in the modality or modalities in which it was experienced. A short summary of possible interventions to transform images and their meanings at the level of negative automatic thoughts will now be given.

In many anxious images a scenario is pictured, in which the worst possible outcome has been visualized, and the image is effectively frozen in time at this worst point. In such a case it may be sufficient to simply let the image run on in time. For example, a client who was afraid of being sick on arrival at a party, and thus totally ruining the evening for everyone, was helped by running his recurrent image on in time past the point where it usually stopped (i.e., when he had just vomited) to discover that as it unfolded people appeared in the image to help him clear up, were sympathetic, and encouraged him to stay on until he felt better. Thus, he could appreciate the fact that the evening might not in fact be ruined. This was easier to believe once he had also tried picturing what might happen if his girlfriend vomited, rather than him (a decentring technique which is often useful in social anxiety).

Another client was terribly distressed by recurrent images of her own gravestone, and the happy faces of her husband and child when they visited it. Verbal discussion had helped her to see that this image was only an image, and did not necessarily depict what would happen in the future, but without lessening her distress. What did help was to imagine herself alive and well, and looking at this image on a television. She was then able to picture herself switching off the television, and driving away from the house. This changed her beliefs about the meaning of having the image, in the sense that she no longer felt that it reflected her future, but was only a product of her fearful mind. This simple intervention got rid of both images and nightmares on this theme, which had been tormenting her. Putting the image on the screen gave her the message that after all it was only an image, and one which she could control. This maneuver is also useful in dealing with any very frightening imaginal material, since the image can then be switched on or off, or made smaller or dimmer, as a prelude to dealing with it further. In these cases the meaning that is changed is at a meta-level: what is changed is the meaning of having the image, rather than the meaning of the image. Similar techniques can be used in the treatment of distressing images following trauma. Putting an image or vivid memory of a screen and altering it can help keep affect at a managable level, and reinforce the idea that what happens in that working space is not what is happening in the present: it is only an image or memory, and belongs in the past. Further work can then also be done to change the content if appropriate.

In some cases when one enquires about images those that are reported are not accompanied by much affect, and are rather small, or distant, blurred, or foggy. In such a situation the patient can be asked to zoom in on the image, and make it bigger or clearer, in order to examine and work on the content.

Where there is an absence of positive images, for example in the hopeless depressive attempting to picture the future, it may be helpful to prompt the patient to generate detailed, positive images. For example, they can be asked to picture themselves in the future when the depression is over, and they can be asked to picture where they are living, with whom, and what work they are doing, and what their hobbies are. They are to imagine who is there and what is actually happening at given points in time. This type of work can subsequently have a motivating effect on the patient (Layden, personal communication).

In a similar way where the patient is about to try some new behavior in real life it can be very helpful to imagine carrying out the desired coping behavior. This gives the opportunity for a useful rehearsal of the skill, and can then either go into memory as a success experience, or if snags arise in the imaginary scenario more verbal discussion and imagery work can be done to deal with these.

Another useful technique is to abandon the spectator position. In social phobia it has already been mentioned that individuals are likely to have very distorted images of themselves, as seen from an external point of view (i.e., as an observer might see them). Since these images can be shown to be inaccurate, using feedback from others, or video feedback, patients can be advised to ignore them, and focus instead on the external environment, and the true reactions of others. In a similar way people with health anxiety can have images of what they imagine is happening inside their bodies, and can learn to ignore these or replace them with more realistic mental pictures.

6.14.5.4 Working with Beliefs about Having Images, Intrusive Memories, and Nightmares

Often patients who experience affectively laden intrusions are further distressed about the meaning of actually having the intrusions. They may believe that these intrusions reflect reality, for example, that they foretell the future. Or they may believe that they are a sign that they are losing control or going mad, particularly if attempts to banish them lead to increased occurrence of the intrusions. Sometimes they believe that having particular thoughts or images may mean something bad about them as a person, or may directly affect what will happen (e.g., thinking about illness can make you become ill, or thinking about being well or happy could tempt providence). Such beliefs can be tackled in therapy using verbal techniques and behavioral experiments.

6.14.5.5 Working with Core Beliefs Using Imagery Techniques

A variety of verbal and behavioral techniques are also available for use in changing core beliefs (Beck, Emery, & Greenberg, 1985; Beck, Rush, Shaw, & Emery, 1979). These can be used to examine both conditional and unconditional beliefs about the self, the world, and other people. However, there may be ways of supporting this verbal work with imaginal material in ways which enhance the emotional shift, and help to affirm and strengthen more adaptive ways of viewing life.

The cognitive model of emotional disorder suggests that our negative automatic contents of consciousness reflect our core beliefs, as activated by present events. These beliefs were formed as a result of early experience. Very often there are particular early memories which seem to reflect core beliefs as they emerged in a person's life. These memories can be accessed and transformed using imagery techniques, as described by Edwards (1990), Layden, Newman, Freeman, and Byers Morse (1993), and Smucker, Dancu, Foa, and Niederee (1995).

Layden et al. (1993) describe the sequence in which one can work when attempting to "repair" traumatic early memories. The first step is to notice when a negative core belief is activated, and instruct the client to dwell on how they feel emotionally, in terms of bodily sensations, and what sorts of thoughts and images they are having. Once they have dwelled on these contents of consciousness they are asked whether there is an early memory which is linked to such emotions, sensations, and thoughts. Very often one or more traumatic memories come to mind. Once a memory has been selected as seemingly important in the formation of the negative core belief the patient is instructed to close their eyes, and relive the memory, frame by frame, reporting their experience aloud, and the meanings they are giving to events. Once all the painful meanings have been discussed and challenged where appropriate using verbal techniques it is often extremely helpful for the patient to mentally replay the memory, bringing in any corrective information which has been pinpointed during discussion. This can be done in many ways. For example, the adult self can visit the unhappy child, and present an adult perspective on their predicament, and the conclusions they have drawn from it. Also, other coping and rescue factors can be highlighted.

To give an example, a man suffering from agoraphobia had great anxiety on going far from home, accompanied by fears of becoming ill in some way and images of being taken to hospital, kept there indefinitely against his will, and never being visited. Whilst attempting to transform one such image of being trapped in hospital he retrieved a traumatic memory of being taken to hospital when he was three, left there by his parents for three days without explanation of where they had gone. At the time he had become extremely afraid that, as the unwanted fourth son of a poor family with a sick mother, he had been gladly abandoned by his parents. The transformation of the meaning of this memory image was effected by having the adult self visit his parents in the hospital, and warn them not to leave the child there without explaining why they were doing it, and reassuring him that they loved him. This was accomplished after reviewing through discussion in therapy the evidence he had that in fact his parents did love him, and also that back in the 1950s people did not realize how traumatic a lonely hospital stay could be for a small child. This intervention had the effect of abolishing his panic attacks for several months, during which additional work was done on his core beliefs.

Quite fanciful scenarios can be utilized to good effect. One patient who had concluded at an early age that she was unlovable after certain messages were given to her by her mother repeatedly saw her child self as unattractive, with the word "precocious" over her head. In imagination she visited an "inner guide," whom she visualized as a wizard, and asked him to show her in a crystal ball what sort of a child she actually was. In the crystal ball many pleasant images of herself appeared, including one of herself aged six putting a hot water bottle in her parents' bed. This image had the word "kind" over her head, and was accompanied by a deep and lasting shifting of the affect associated with memories of her childhood, and her feelings about herself.

More orthodox schema work can also be reinforced by the use of symbolic material. For example, one patient had many unconditional negative core beliefs about herself and others, to the effect that she was an emotionally deprived person, whom others should look after unconditionally, and never leave her to fend for herself. If they did it meant that she had been rejected and abandoned in a harsh world, and could trust no-one. In therapy she was working towards being able to stand criticism, and take care of herself without a catastrophic emotional response. She was greatly supported in this by reading the story of Vasalisa, whose mother died when she was small, but left her a doll who lived in her pocket, and helped her decide what to do when the going got tough. The patient found this a most useful metaphor for her new sense of herself as supported, but able to be independent, with growing intuition and skills of her own.

Finally, it may be worth noting that some of the techniques described above which can be used to transform the meaning of early memories can also be used when working with intrusive traumatic fragments of memory. The client can be asked to bring the traumatic memory vividly to mind, and then unpack its meaning, as described above. Once the meaning has been discussed and verbally challenged the memory can be replayed in imagery, incorporating material depicting the new perspective. For example, a man whose father had died in an explosion discovered that the meaning behind his intrusive image of his father's body was the idea that his father's spirit had not been able to leave the scene of the accident. Imagining his father's spirit as now being free, and able to travel to his favorite place and then be released provided relief, and the images of the body disappeared.

6.14.6 CONCLUSION

Clinical psychologists who work within the scientist–practitioner tradition would probably agree on a model of human functioning which proposes that in a given situation conscious and unconscious appraisals will be made, leading to a set of automatic reactions (such as autonomic changes and probably some automatic adjustments in information processing, together with behavioral reactions (such as escape, avoidance, etc). Both automatic and behavioral reactions may change the perceived stimulus situation to some extent, so fresh appraiasals may be made. For example, in a panic attack bodily changes may be perceived, and catastrophically misinterpreted (perhaps in an image of an imagined heart attack) resulting in an escalation of anxiety. Also, in social anxiety a person might talk quickly, and attempt to fill silences for fear of being thought boring, but might then get a distorted image of themselves as babbling or talking terribly quickly, from which they might conclude that that is how they are being seen by others at that moment.

The therapist can attempt to break in to apparently closed systems of situations, appraisals, and automatic and behavioral reactions in a number of ways. In life people are continually appraising stimuli for their meanings, and then reacting accordingly. If we agree that this takes place partly consciously and partly unconsciously then there need be no disagreement between various schools of thought, but instead they can be seen to be focusing on different aspects of functioning.

Images and memories with sensory components form a sizable proportion of the conscious appraisals of many people, and it is also possible to bring them under voluntary control, at least to some extent. Since they can contain both stimulus and response elements (Lang, 1977, 1979) they can provide a useful working space in therapy in which to attempt to change appraisals of stimuli, or of automatic or behavioral reactions to them.

There is a certain amount of literature to support the view that emotion is more closely linked to mental contents with sensory components than those in the verbal modality. In most individuals conscious experience is a rapidly changing mix of these two ingredients. It has been hypothesized by Borkovec and Inz (1990) that one of the possible functions of verbal worrying is to keep more frightening or upsetting images or memories at bay, or in the unconscious, to use the terminology of the analysts.

It is certainly true that after a traumatic experience people often try to keep memories at bay, and the fragmented memories appear to be stored in sensory modalities, and can be readily accessed by situational cues, to the extent that in PTSD people can find themselves involuntarily reliving various aspects of the trauma. As time goes by (in cases where the problem is successfully resolved) a narrative version emerges, which incorporates details previously forgotten, and which is no longer fragmented or disjointed. This narrative can be run through in imagination or conversation without enormous distress, and the more sensory type of memory is less likely to be triggered by sensory cues. This interesting process has been described by Jaycox and Foa (1996), and is also consistent with theoretical models such as those of Brewin, Dalgleish, and Joseph (1996), Dalgleish (1997), Ehlers (1997), Ehlers and Clark (in press), and others discussed by Power (1997).

It may be that such theoretical accounts of PTSD have something to tell us about other emotional problems. It has been noted above that in various anxiety disorders there are spontaneous images in anxious situations, which not only depict possible catastrophes which might occur now or in the future, but also are often rich with sensory and interpersonal material which appears to be linked to past experiences in which there were similar themes and stimuli. There is evidence that bringing recurrent upsetting images into full awareness can often spontaneously evoke such traumatic early memories. Both images and memories echo current concerns and important themes for the person, encapsulating idiosyncratic meanings, linked to core beliefs.

A wide range of techniques are available for use in this "working space," and they seem to share many characteristics. They allow rich, multifaceted material to emerge into full consciousness, and provide opportunities for the incorporation of corrective information, and the possibility of moving towards storing material in a semantic, orderly form, which does not intrude, and is not so affectively laden as when it was stored in a more sensory mode. It is interesting to note that in a number of the therapeutic approaches described merely bringing upsetting images or memories to mind repeatedly seems often to result in spontaneous changes over time, with the imagined material becoming less emotionally charged, less vivid, and less threatening in meaning. In other cases imagery techniques can be used to access meanings, which can then be challenged using verbal discussion techniques, behavioral experiments, and direct restructuring of images and memories. There also appear to be direct effects on behavior from the imaginal rehearsal of possible future scenarios. However, this whole area has not been extensively researched, and much interesting work remains to be done.

6.14.7 REFERENCES

Acierno, R., Hersen, M., Van Hasselt, V. B., & Tremont, G. (1994). Review of the validation and dissemination of eye-movement desensitization and reprocessing: A scientific and ethical dilemma. *Clinical Psychology Review, 14,* 287–299.

Arieti, S. (1985). Cognition in psychoanalysis. In M. J. Mahoney & A. Freeman (Eds.), *Cognitive and psychotherapy*. New York: Plenum.

Arnkoff, D. B. (1981). Flexibility in practicing cognitive therapy. In G. Emery, S. D. Hollon, & R. C. Bedrosian (Eds.), *New directions in cognitive therapy: A casebook*. New York: Guilford.

Assaglioli, R. (1965). *Psychosynthesis*. New York: Hobbs, Dorman, & Co.

Bandura, A., Blanchard, E. B., & Ritter, B. (1969). Relative efficacy of desensitization and modeling approaches for inducing behavioral, affective, and attitudinal changes. *Journal of Personality and Social Psychology, 13,* 173–199.

Barlow, D. H., Leitenberg, H., Agras, W. S., & Winche, J. P. (1969). The transfer gap in systematic desensitization: An analogue study. *Behaviour Research and Therapy, 7,* 191–196.

Barnard, P. J., & Teasdale, J. D. (1991). Interactive cognitive subsystems: A systemic appraoch to cognitive-affective interaction and change. *Cognition and Emotion, 5,* 1–39.

Bartlett, F. C. (1932). *Remembering*. Cambridge, UK: Cambridge University Press.

Beck, A. T. (1970). The role of fantasies in psychotherapy and psychopathology. *Journal of Nervous and Mental Disease, 150,* 3–17.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: New American Library.

Beck, A. T., Emery, G., & Greenberg, R. L. (1985). *Anxiety disorders and phobias*. New York: Basic Books.

Beck, A. T., & Hurvich, M. (1959). Psychological correlates of depression. *Psychosomatic Medicine, 21,* 50–55.

Beck, A. T., Laude, R., & Bohnert, M. (1974). Ideational components of anxiety neurosis. *Archives of General Psychiatry, 31,* 319–325.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford.

Beck, A. T., & Ward, C. H. (1961). Dreams of depressed patients: Characteristic themes in manifest content. *Archives of General Psychiatry, 5,* 462–467.

Borkovec, T. D., & Inz, J. (1990). The nature of worry in generalised anxiety disorder: A predominance of thought activity. *Behaviour Research and Therapy, 28,* 153–158.

Borkovec, T. D., Robinson, E., Pruzinsky, T., & DePree, J. A. (1983). Preliminary exploration of worry: Some characteristics and processes. *Behaviour Research and Therapy, 21,* 9–16.

Bowlby, J. (1985). The role of childhood experience in cognitive disturbance. In M. J. Mahoney & A. Freeman (Eds.), *Cognition and psychotherapy*. New York: Plenum.

Breger, L., & McGaugh, J. L. (1965). Critique and reformulation of "learning-theory" approaches to psychotherapy and neurosis. *Psychological Bulletin, 63,* 338–358.

Bresler, B. (1984). Mind-controlled analgesia: The inner way to pain control. In A. A. Sheikh (Ed.), *Imagination and healing*. Farmingdale, NY: Baywood.

Brewin, C. R., Dalgleish, T., & Joseph, S. (1996). A dual representation theory of post-traumatic stress disorder. *Psychological Review, 103,* 670–686.

Brewin, C. R., Hunter, E., Carroll, F., & Tata, P. (1996). Intrusive memories in depression: An index of schema activation? *Psychological Medicine, 26,* 1271–1276.

Brown, R. W. (1958). *Words and things*. Glencoe, USA: Free Press.

Buzan, T., & Buzan, B. (1997). *The mindmap book*. London: BBC Books.

Cautela, J. R. (1966). Treatment of compulsive behavior by covert sensitization. *Psychological Record, 16,* 33–41.

Cautela, J. R. (1967). Covert sensitization. *Psychological Reports, 20,* 459–468.

Cautela, J. R. (1970a). Covert reinforcement. *Behavior Therapy, 1,* 33–50.

Cautela, J. R. (1970b). Covert negative reinforcement. *Journal of Behavior Therapy and Experimental Pyschiatry, 1,* 273–378.

Cautela, J. R. (1971). Covert extinction. *Behavior Therapy, 2,* 192–200.

Cautela, J. R. (1977). Covert conditioning: Assumptions and procedures. *Journal of Mental Imagery, 1,* 53–64.

Chambless, D. L., & Goldstein, A. J. (1982). *Agoraphobia: Multiple perspectives on theory and treatment*. New York: Wiley.

Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. Heimberg, M. Liebowitz, D. A. Hope, & F. R. Scheier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 69–93). New York: Guilford.

Dalgleish, T. (1997). *Phenomenology and the clinical course of PTSD from the perspective of a multi-level theory of emotion*. Paper presented at the 25th Anniversary Conference of the British Association for Behavioural and Cognitive Psychotherapy, Canterbury, UK.

Dawson, R. W., & McMurray, N. E. (1978). Desensitization without hierarchical presentation and concomitant relaxation. *Australian Journal of Psychology, 30,* 119–132.

de Silva, P. (1986). Obsessional-compulsive imagery. *Behaviour Research and Therapy, 24,* 333–350.

Dyckman, J. M., & Cowan, P. A. (1978). Imaging vividness and the outcome of *in vivo* and imagined scene desensitization. *Journal of Consulting and Clinical Psychology, 46,* 1155–1156.

Edwards, D. J. (1989). Cognitive restructuring through guided imagery: Lessons from gestalt therapy. In A. Freeman, K. M. Simon, L. E. Beutler, & H. Arkowitz (Eds.), *Comprehensive handbook of cognitive therapy*. New York: Plenum.

Edwards, D. J. (1990). Cognitive therapy and the restructuring of early memories through guided imagery. *Journal of Cognitive Psychotherapy: An International Quarterly, 4*(1), 33–49.

Ehlers, A. (1997). *Data-driven versus semantic processing driven trauma*. Paper presented at the 25th annual conference of the British Association for Behavioural and Cognitive Psychotherapy, Canterbury, UK.

Ehlers, A., & Clark, D. M. (in press). A cognitive model of persistent posttraumatic stress disorder. *Behaviour Research and Therapy*.

Emmelkamp, M. G., & Wessels, H. (1975). Flooding in imagination vs. flooding *in vivo*: A comparison with agoraphobics. *Behaviour Research and Therapy, 13,* 7–15.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99,* 20–35.

Foa, E. B., Molnar, C., & Cashman, L. (1995). Change in rape narratives during exposure thearpy for posttraumatic stress disorder. *Journal of Traumatic Stress, 8,* 675–690.

Foa, E. B., Rothbaum, B. O., Riggs, D. S., & Murdock, R. B. (1991). Treatment of posttraumatic stress disorder in rape victims: A comparison between cognitive-behavioral procedures and counselling. *Journal of Consulting and Clinical Psychology, 59,* 715–723.

Foa, E. B., Steketee, G., & Grayson, J. B. (1985). Imaginal and *in vivo* exposure: A comparison with obsessive-compulsive checkers. *Behavior Therapy, 16,* 292–302.

Foa, E. B., Steketee, G., Turner, R. M., & Fischer, S. C. (1980). Effects of imaginal exposure to feared disasters in obsessive-compulsive checkers. *Behaviour Research and Therapy, 18,* 449–455.

Freeman, A. (1981). Dreams and images in cognitive therapy. In G. Emery, S. D. Hollon, & R. C. Bedrosian (Eds.), *New directions in cognitive therapy: A case book*. New York: Guilford.

Freeman, A., & Boyll, S. (1992). The use of dreams and the dream metaphor in cognitive-behavior therapy. Special issue: Psychotherapy in independent practice: Current issues for clinicians. *Psychotherapy in Private Practice, 10,* 173–192.

Freud, S. (1962). Freud's psychoanalytic procedure. In J. Strachey (Ed.), *The standard edition*. London: Hogarth.

Freud, S., & Breuer, J. (1955). Studies of hysteria. In J. Strachey (Ed.), *The standard edition*. London: Hogarth.

Gendlin, E. T. (1978). *Focussing*. New York: Bantam Books.

Greenberger, D., & Padesky, C. A. (1995). *Mind over mood: A cognitive therapy treatment manual for clients*. New York: Guilford.

Grof, S. (1976). *Realms of human unconscious: Observations from LSD research*. New York: Dutton.

Habeck, B. K., & Sheikh, A. A. (1984). Imagery and the treatment of phobic disorders. In A. A. Sheikh (Ed.), *Imagination and healing*. Farmingdale, NY: Baywood.

Hackmann, A. (1995). *The "felt sense" of things in anxiety disorders: Looking at ways in which images and metaphors echo early traumatic memories*. Paper presented at the World Congress of Behavioural and Cognitive Therpaies, Copenhagen, Denmark.

Hackmann, A., & Clark, D. M. (in press). *Recurrent images and memories in social phobia*. Manuscript in preparation.

Hackmann, A., & Surawy, C. (in press). *Recurrent images and memories in agoraphobia*. Manuscript in preparation.

Hackmann, A., Surawy, C., & Clark, D. M. (1998). Seeing yourself through others' eyes: A study of spontaneously occurring images in social phobia. *Behavioural and Cognitive Psychotherapy, 26,* 3–12.

Hall, H. (1984). Imagery and cancer. In A. A. Sheikh (Ed.), *Imagination and healing*. Farmingdale, NY: Baywood.

Hogan, R. A. (1966). Implosive therapy in the short term treatment of psychotics. *Psychotherapy: Theory Research and Practice, 3,* 25–32.

Hogan, R. A., & Kirchner, J. H. (1967). A preliminary report of the extinction of learned fears via short term implosive therapy. *Journal of Abnormal Psychology, 2,* 106–109.

Hogan, R. A., & Kirchner, J. H. (1968). Implosive, eclectic, verbal and bibliotherapy in the treatment of fears of snakes. *Behaviour Research and Therapy, 6,* 167–171.

Horowitz, M. (1970). *Image formation and cognition*. New York: Appleton-Century-Crofts.

Jacobson, E. (1942). *Progressive relaxation*. Chicago: University of Chicago Press.

Jaycox, L. H., & Foa, E. B. (1996). Obstacles in implementing exposure therapy for PTSD: Case discussions and practical solutions. *Clinical Psychology and Psychotherapy, 3,* 176–184.

Kazdin, A. E. (1974). Effects of covert modeling and model reinforcement on assertive behavior. *Journal of Abnormal Psychology, 83,* 240–252.

Kirchner, J. H., & Hogan, R. A. (1966). The therapist variable in the implosion of phobias. *Psychotherapy, 3,* 102–104.

Kuyken, W., & Brewin, C. R. (1994). Intrusive memories of childhood abuse during depressive episodes. *Behaviour Research and Therapy, 32,* 525–528.

Lang, P. J. (1977). Imagery in therapy: An information processing analysis of fear. *Behavior Therapy, 8,* 862–886.

Lang, P. J. (1979). A bio-informational theory of emotional imagery. *Psychophysiology, 16,* 495–512.

Layden, M. A., Newman, C. F., Freeman, A., & Byers Morse, S. (1993). *Cognitive therapy of borderline personality disorder*. Boston: Allyn and Bacon.

Leuner, H. (1969). Guided affective imagery (GAI): A method of intensive psychotherapy. *American Journal of Psychotherapy, 23,* 4–22.

Levis, D. J. (1980). Implementing the technique of implosive therapy. In A. Goldstein & E. B. Foa (Eds.), *Handbook of behavioral interventions*. New York: Wiley.

Levis, D. J., & Carrera, R. M. (1967). The effects of 10 hours of implosive therapy in the treatment of outpatients. *Journal of Abnormal Psychology, 6,* 504–508.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behaviour, 13,* 585–589.

Lohr, J. M., Kleinknecht, R. A., Tolin, D. F., & Barrett, R. H. (1995). The empirical status of the clinical application of eye movement desensitization and reprocessing. *Journal of Behaviour Therapy and Experimental Psychiatry, 26,* 285–302.

Mathews, A. M. (1978). Fear-reduction reseach and clinical phobias. *Psychological Bulletin, 85,* 390–404.

Mathews, A. M., Johnston, D. W., Lancashire, M., Munby, M., Shaw, P. M., & Gelder, M. G. (1976). Imaginal flooding and exposure to real phobic situations: Treatment outcome with agoraphobic patients. *British Journal of Psychiatry, 129,* 361–371.

Ornstein, R. (1992). *Evolution of consciousness*. New York: Touchstone.

Ottaviani, R., & Beck, A. T. (1987). Cognitive aspects of panic disorders. *Journal of Anxiety Disorders, 1,* 15–28.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, UK: Oxford University Press.

Perls, F. S. (1971). *Gestalt therapy verbatim*. New York: Bantam.

Power, M. J. (1997). Conscious and unconscious representations of meaning. In M. Power & C. R. Brewin (Eds.), *The transformation of meaning in psychological therapies*. Chichester, UK: Wiley.

Power, M. J., & Dalgleish, T. (1997). *Cognition and emotion: From order to disorder*. Hove, UK: Erlbaum.

Rabavilas, A. D., Boulougouris, J. C., & Stefanis, C. (1976). Duration of flooding sessions in the treatment of obsessive-compulsive patients. *Behaviour Research and Therapy, 14,* 349–355.

Salkovskis, P. M., & Westbrook, D. (1989). Behaviour therapy and obsessional ruminations can failure be turned into success? *Behaviour Research and Therapy, 27,* 149–160.

Samuels, M. & Samuels, N. (1975). *Seeing with the mind's eye: The history, techniques and uses of visualisation*. New York: Random House.

Shapiro, F. (1989). Eye movement desensitization: A new treatment for post-traumatic stress disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 20,* 211–217.

Shapiro, F. (1995). *Eye movement desensitization and reprocessing: Basic principles, protocols, and procedures*. New York: Guilford.

Sherman, A. R. (1972). Real-life exposure as a primary therapeutic factor in the desensitization treatment of fear. *Journal of Abnormal Psychology, 79,* 19–28.

Shultz, K. D. (1984). The use of imagery in alleviating depression. In A. A. Sheikh (Ed.), *Imagination and* *healing*. Farmingdale, NY: Baywood.

Singer, J. L. (1974). *Imagery and daydream methods in psychotherapy and behavior modification*. New York: Academic Press.

Singer, J. L., & Pope, K. S. (1978). The use of imagery and fantasy techniques in psychotherapy. In J. L. Singer & K. S. Pope (Eds.), *The power of human imagination: New methods in psychotherapy*. New York: Plenum.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32,* 752–760.

Smucker, M. R., Dancu, C., Foa, E. B., & Niederee, J. L. (1995). Imagery rescripting: A new treatment for survivors of childhood sexual abuse suffering from posttraumatic stress. *Journal of Cognitive Psychotherapy, 9,* 3–17.

Spenceley, A., & Jerrom, B. (1997). Intrusive traumatic childhood memories in depression: A comparison between depressed, recovered and never depressed women. *Behavioural and Cognitive Psychotherapy, 25,* 309–318.

Stampfl, T., & Levis, D. J. (1967). Essentials of implosive therapy: A learning theory based psychodynamic behavioural therapy. *Journal of Abnormal Psychology, 72,* 496–503.

Stern, R., & Marks, I. (1973). Brief and prolonged flooding: A comparison in agoraphobic patients. *Archives of General Psychiatry, 28,* 270–276.

Stevens, J. O. (1971). *Awareness: Exploring, experimenting, experiencing*. Moab, UT: Real People Press.

Trinder, H. & Salkovskis, P. M. (1993). Personally relevant intrusions outside the laboratory: Long-term suppression increases intrusion. *Behavior Research and Therapy, 32*(8), 833–842.

Van Etten, M. L. & Taylor, S. (1997). *Comparative efficacy of treatments for posttraumatic stress disorder: A meta-analysis*. Paper presented at the XXVII Congress of the European Association for Behavioural and Cognitive Therapies, Venice, Italy.

Warwick, H. M., & Salkovskis, P. M. (1990). Hypochondriasis. *Behaviour Research and Therapy, 28,* 105–117.

Watkins, M. (1976). *Waking dreams*. New York: Harper.

Watts, F. N., Sharrock, R., & Trezise, L. (1986). Detail and elaboration in phobic imagery. *Behavioural Psychotherapy, 14,* 115–123.

Watts, F. N., Trezise, L., & Sharrock, R. (1986). Processing of phobic stimuli. *British Journal of Clinical Psychology, 35,* 253–259.

Weitzman, B. (1967). Behavior therapy and psychotherapy. *Psychological Review, 74,* 300–317.

Wells, A., Ahmed, S., & Clark, D. M. (in press). Perspective taking in social phobic imagery. *Behaviour Research and Therapy*.

Wells, A., & Hackmann, A. (1993). Imagery and core beliefs in health anxiety: Content and origins. *Behavioural and Cognitive Psychotherapy, 21,* 265–273.

Wilkins, W. (1971). Desensitization: Social and cognitive factors underlying the effectiveness of Wolpe's procedure. *Psychological Bulletin, 76,* 311–317.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Palo Alto, CA: Stanford University Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.15

Group Therapy: A Cognitive-behavioral Interactive Approach

SHELDON D. ROSE
University of Wisconsin–Madison, WI, USA

6.15.1 INTRODUCTION 319
6.15.2 THE IMPLICATION OF THE GROUP IN CBIGT 321
6.15.3 THERAPIST ACTIVITIES IN CBIGT 324
  6.15.3.1 Pregroup Planning 324
  6.15.3.2 Orientation to Group Therapy 325
  6.15.3.3 Assessment 325
  6.15.3.4 Intervention 328
  6.15.3.5 Generalization 328
6.15.4 INTERVENTION STRATEGIES USED IN CBIGT 329
  6.15.4.1 Systematic Problem-solving 329
  6.15.4.2 The Modeling Sequence 329
  6.15.4.3 Cognitive Change Methods 331
    6.15.4.3.1 Cognitive restructuring 331
    6.15.4.3.2 Self-instructional training 331
  6.15.4.4 Relaxation Methods 332
  6.15.4.5 Reinforcement and Stimulus Control 332
  6.15.4.6 Small-group Procedures 333
    6.15.4.6.1 Role-playing 333
    6.15.4.6.2 Subgrouping 333
    6.15.4.6.3 Buddy system 333
    6.15.4.6.4 Group exercises 333
    6.15.4.6.5 Group feedback 334
  6.15.4.7 Identifying and Resolving Group Problems 334
  6.15.4.8 Phases of Group Development 335
  6.15.4.9 Integrating Diverse Intervention Strategies in CBIGT 335
6.15.5 SUMMARY AND CONCLUSION 336
6.15.6 REFERENCES 336

6.15.1 INTRODUCTION

Group therapy is widely used as the context of treatment for patients with a variety of problems and concerns. In a recent survey of mental health providers in a middle-sized American city, it was noted that 72% of the providers made the option of group therapy available to their patients (Rose, 1998a). There are many types of group therapy, one of which makes use primarily, though not exclusively, of cognitive and behavioral procedures. The purpose of this chapter is to describe the rationale for, review research concerning, and present one general approach and the techniques commonly applied in this form of group therapy primarily for adults. Unfortunately, in the existing literature, the role of the group is rarely explicated (see Rose, Tolman, & Tallant, 1985, for a review of the cognitive-behavioral group therapy literature until 1984).

Because of this neglect the potential use of the group as a set of interventions in itself in administrating cognitive and behavioral procedures is described in this chapter. Because the approach lends itself to evaluation and specific description of its primary intervention, assessment, and generalization strategies, cognitive-behavioral group therapy appears to be growing in popularity among practitioners and patients. Its relevance in particular for women has been described by Wolfe (1987) and for minorities in studies by La Framboise and Rowe (1983) and Comas-Diaz and Duncan (1985).

Cognitive-behavioral group therapy (CBGT) refers to an approach that occurs in groups and makes use of behavioral (e.g., modeling and reinforcement), cognitive (e.g., cognitive restructuring, self-instructional training, problem-solving), relational change methods, and group procedures to enhance the coping skills of the participants and to resolve or ameliorate relational problems they may be experiencing. Coping skills refer to that set of behaviors and cognitions which facilitate adaptation to stressful or problematic day-to-day situations. Some coping skills which participants learn in CBGT are systematic problem-solving, relaxation, meditation, imagery, techniques for "decatastrophizing events," correcting other cognitive distortions, being appropriately assertive, distancing or objectifying one's pain or strong emotions, communication skills in dealing with others, negotiation, and making more effective use of social networks and increasing social-recreational skills. Many of these skills are both target behaviors and a means of intervention for learning more specific target behaviors.

Group treatment approaches in general when compared with individual approaches seem to fare favorably. A meta-analysis (Toseland & Siporin, 1986) reviewing all group vs. individual therapy comparisons (many of which were cognitive and behavioral) showed evidence that groups are at least as effective and sometimes more effective than individual treatment and more efficient in that more patients can be dealt with effectively for the same cost. This finding was confirmed by Tilliski (1990) who used only those studies from the Toseland and Siporin review that contrasted group, individual, and a control group. In all cases, regardless of theory or problem focus, there was no difference between treatments and significant differences between treatment and the control group. Since that time several studies involving CBGT have appeared which also support the equivalency of individual and group therapy and their greater effectiveness than control groups in the treatment of a variety of presenting problems. For example, Spence (1991) demonstrated that CBGT was equivalent to individual cognitive-behavioral therapy and both were more effective than a wait-list control in the treatment of chronic occupational pain. Oneytt and Turpin (1988) found that a six-week cognitive-behavioral group treatment was more effective than a control group but equivalent to individual general practitioner appointments in the reduction of benzodiazepine use and anxiety. All changes were maintained at 15-week follow-up. The results are confounded somewhat by the fact that the general practitioner may not have used cognitive-behavioral methods. Teri and Lewinsohn (1986) also found that both individual and group behavioral treatment were equivalent and both were more effective than a wait-list control in the treatment of patients with moderate depression.

Support for the effectiveness of CBGT has also been demonstrated in studies in which CBGT was compared to wait-listed and other approaches. When Wolf and Cowther (1992) compared CBGT with behavioral treatment for 41 bulimic women, both were more effective than an extreme weight control method in terms of reduced bulimic tendencies. CBGT was also more effective than the behavioral treatment measures in the treatment reduction of severity of psychopathological symptoms and preoccupation with dieting. Lee and Rush (1986) had similar results with 30 bulimic patients. Those receiving five weeks of CBGT reduced their binging and purging behavior significantly more than those in a wait-list control group. However, only four of the CBGT patients showed a full remission of the binging and purging behavior.

When Bottomley, Hunton, Roberts, and Jones (1996) compared nine patients in CBGT to a eight patients in social support group and to 14 "decliners" for newly diagnosed and psychologically stressed cancer patients, they found that the CBGT patients significantly improved their coping styles in comparison with the other two conditions. At three-months follow-up, the differences washed out possibly because two CBGT patients had died.

Ehlers, Stangier, and Gieler (1995) compared cognitive and behavioral methods carried out in groups with intensive or standard dermatological treatment. They found that the behavioral treatments for atopical dermatitis resulted in significant improvement in skin condition and in topical steroids used when compared to dermatological condition alone, at one-year follow-up. Once again, the group factor and the behavioral methods were confounded in the experimental condition.

Fals-Stewart, Marks, and Schafer (1993) compared a behavioral group therapy condition ($n=30$) consisting of group exposure and response prevention with an individual therapy condition ($n=31$) using the same methods and a relaxation control condition ($n=32$). The subjects were randomly assigned to the conditions. Patients in all three conditions showed stress reduction at the end of treatment, but only in the behavioral conditions were the gains significantly maintained at the six-month follow-up.

Lutgendorf, Antoni, Ironson, and Klimas (1997) tested the effects of a 10-week cognitive-behavioral group stress management program on mood and immunologic parameters in HIV-seropositive gay men whose disease had progressed to a symptomatic stage. The patients were randomly assigned to the CBGT ($n=14$) or a wait-list control group ($n=12$). At the end of the 10 weeks patients in the CBGT condition significantly decreased self-rated dysphoria, anxiety, and total distress. The intervention also decreased Herpes simplex virus-type 2 (Hsv-2) immunoglobulin G antibody titers. No such changes were found in the control group. There was no follow-up measurement.

Subramanian (1991) found that 39 chronic pain patients in eight weekly sessions of CBGT improved significantly more than a wait-list control ($n = 20$) in the areas of physical and psychosocial dysfunction, but showed no difference in the experiencing of pain. Group strategies as well as cognitive-behavioral strategies were described in this study. However, in a follow-up study, Subramanian (1994) found that the gains of a subsample that could be located were not maintained. However, no specific steps in and after the therapy program had been taken to achieve maintenance.

Patients in CBGT in a number of before and after multiple cases without control groups have consistently shown significant change over time and for the most part these changes were maintained 2–3 months following treatment. These studies include coping with generalized anxiety (Lindsay, Gamsu, McLaughlin, Hood, & Espie, 1987; Power, Simpson, Swanson, & Wallace, 1990), adaptation problems of recently divorced women (Graff, Whitehead, & LeCompte, 1986), managing one's depression (Abraham, Neundorfer, & Currie, 1992), social phobia (Enright, 1991), obsessive-compulsive disorders (Krone, Himle, & Nesse, 1991; Van Hoppen, Rassmussen, & Eisen, 1991), insomnia (Kupych-Woloshyn, MacFarlane, & Shapiro, 1993), and a heterogeneous population of 531 patients suffering from various psychiatric problems in intensive CBGT at a private psychiatric clinic (Manning, Hooke, Tannenbaum, & Blythe,1994).

Although each of the above researchers treat the given problem area in groups using cognitive-behavioral and/or behavioral intervention strategies, most use different combinations of these techniques and, as already mentioned, a few describe specific uses of groups in the treatment process such as member modeling and building cohesion of the group. In summary, modest support for CBGT can be identified. Usually in the control group and databased case studies, the group factor and the cognitive-behavioral techniques are confounded, thus making interpretation of the findings unclear as to whether group or the cognitive-behavioral strategies are the major causes of change.

Because of the variations mentioned above, different kinds of cognitive-behavioral approaches to CBGT exist. Some primarily rely on cognitive, relaxation, relational, and didactic procedures, while others integrate with the cognitive procedures extensive behavioral methods such as modeling and reinforcement into the treatment program. Only a few make specific uses of the group. Because of our clinical and research experience with a model of CBGT that incorporates group techniques into treatment, the purpose of this chapter is to present a cognitive-behavioral model that indeed makes use of the group. In this chapter the specific model that includes the extensive use of group methods and a focus on group variables that impinge on outcome along with the use of cognitive-behavioral procedures is referred to as cognitive-behavioral interactive group therapy (CBIGT), while CBGT refers to all group models using cognitive-behavioral methods, whether explicit use of the group has been determined or not.

6.15.2 THE IMPLICATION OF THE GROUP IN CBIGT

Originally the group was used in therapy because of its convenience and efficiency in treating more than one or two people at the same time. As many scholars of group therapy have noted, it has a number of other advantages (e.g., Yalom, 1985, p. 3). Improved opportunities for assessment is a major contribution of therapy in groups. The group therapist has the opportunity to observe live interactions as opposed solely to second-hand accounts of interactions and make solid conclusions as to social skill strengths and deficiencies. He or she can hear the cognitive distortions typical of many patients being played out in the group. This is possible only insofar as there is broad participation among the patients which is the focus of many interventions in CBIGT.

The group provides the patient with a source of feedback about those behaviors which are irritating or acceptable to others and about those cognitions which can be viewed as distorted or dysfunctional. In CBIGT, participants are trained to help each other be concrete in the formulation of problem situations and the cognitive, emotional, and behavioral responses of the patient to those situations and to identify specific goals towards which intervention is aimed.

Another advantage of using the group in CBIGT is the frequent and varied opportunity for mutual reinforcement which for patients is often far more powerful than reinforcement by the therapist only. Each patient is given the chance to learn or to improve his or her ability to mediate rewards for others in social interactive situations (with acquaintances, friends, family members, acquaintances in other groups, fellow therapists, or employer). The group therapist in CBIGT creates situations in which each patient has frequent opportunities, instructions, and rewards for reinforcing others in the group. In CBIGT, special exercises have been designed to train patients specifically in mutual reinforcement, and often extragroup tasks (homework) are used to encourage them to practice reinforcement skills in the real world.

High levels of reinforcement, among other techniques, contribute to the cohesion of the group. As Yalom (1985, p. 49) argues, group cohesion is the equivalent in groups to relationships in individual treatment. At least in the early phase of therapy, cohesion must be high if other methods are to be maximally effective.

If cohesion is high, the group creates the opportunity for the group therapist to mobilize the members to use of a wide variety of coping skill training procedures that are either unavailable or less efficient in the therapeutic dyad. For example, the most effective way of systematic problem-solving is in the group. In the "brainstorming" phase of problem-solving, the multiplicity of patients and patients' experiences results in numerous ideas and suggestions. The group also provides a large number of models, role-players for behavioral rehearsal, manpower for monitoring, and partners for use in a "buddy system."

In CBIGT, as well as most other CBGT models, members are expected to carry out homework or extragroup tasks which are cognitive, behavioral, or both, to practice what they have learned in the group. Much of the program in CBIGT is designed to prepare the members for completing those tasks. If appropriate norms are established, the group serves to enhance the design through group feedback and mutual supervision in subgroups, and to reinforce the completion of extragroup task completion if the members report publicly at the beginning of the following session what they did for their task during the week.

Negotiating rather than assigning homework tasks in CBIGT and the use of extensive feedback from members and brainstorming increases the power of the members and reduces the excessive didactic quality of many other models of CBGT. Specific questions on weekly postsession questionnaires (see Table 1) keeps the therapist in touch with the group's evaluation of their own participation and power in the group.

Finally, Yalom (1985, p. 14) points out the importance of helping others (altruism) in facilitating therapy in groups. Although he refers to insight-oriented group therapy, the same appears to be true for CBIGT provided that the therapist creates conditions which permit patients to help each other. In this way the participant is not only a patient, he or she is also a co-therapist–teacher within the given structure. This appears to increase the participant's sense of self-efficacy or the belief that he or she can perform in such a way as to achieve a desired outcome (Bandura, 1977). Yalom (1985, p. 8) also reminds the reader of the principle of universality. Patients experience that they are not the only person with the given problem, as serious as it may be.

Of course, groups are not without disadvantages. The time allotted to each individual is drastically reduced in most groups compared to individual treatment. Thus, individualization of patients may suffer. If the group process is utilized in treatment, the amount of material covered is less than when the group process is ignored. In addition, it is difficult to assure the patients of the confidentiality of their comments, even though the therapist emphatically points out its importance. For this reason, some individuals in group therapy are less likely to self-disclose relevant material than they would in individual therapy. However, many others are encouraged by the modeling of other members to self-disclose more readily. Usually as the group progresses, interpersonal trust increases, and the patients become more comfortable with self-disclosing significant concerns. Another danger of CBGT is its tendency to become excessively didactic because of the amount of material to be disseminated. This didactic quality may reduce the cohesion of the group and reduce the benefits which might otherwise ensue. Also, the opportunity for altruistic activities is diminished. For this reason, in CBIGT group therapists are encouraged as early as possible to create conditions that maximize patient participation and self-determination. The use of small group

Table 1 Example of a postsession questionnaire.

code name——group name——date——

1. How useful was this session?
1——2——3——4——5——6——7——8——9
not at all very little somewhat quite a bit extremely
2. How actively involved were the members in today's session?
1——2——3——4——5——6——7——8——9
not at all very little somewhat quite a bit extremely
3. How helpful were members to each other during this session?
1——2——3——4——5——6——7——8——9
not at all very little somewhat quite a bit extremely
4. How much did the members reveal about themselves (their real thoughts, feelings, motivations, and or concerns) during this session?
1——2——3——4——5——6——7——8——9
not at all very little somewhat quite a bit extremely
5. How bored or tired looking were the members during this session?
1——2——3——4——5——6——7——8——9
not at all very little somewhat quite a bit extremely
6. How close did the members feel to each other during this session?
1——2——3——4——5——6——7——8——9
not at all very little somewhat quite a bit extremely
7. How upset or angry were the members during this session?
1——2——3——4——5——6——7——8——9
not at all very little somewhat quite a bit extremely
8. How task oriented were the members during this session?
1——2——3——4——5——6——7——8——9
not at all very little somewhat quite a bit extremely
9. How important were the problems the group worked on in this session?
1——2——3——4——5——6——7——8——9
not at all very little somewhat quite a bit extremely
10. How much did the members control the content and direction of this session?
1——2——3——4——5——6——7——8——9
not at all very little somewhat quite a bit extremely
11. How much conflict was there during this session?
1——2——3——4——5——6——7——8——9
not at all very little somewhat quite a bit extremely
12. To what degree were the goals of the session acheived?
1——2——3——4——5——6——7——8——9
not at all very little somewhat quite a bit extremely
13. How anxious were you during this session?
1——2——3——4——5——6——7——8——9
not at all very little somewhat quite a bit extremely
14. How prepared were you for this session?
1——2——3——4——5——6——7——8——9
not at all very little somewhat quite a bit extremely
15. What specifically did you find useful in todays session?
16. What specifically did you find unhelpful in todays session?

exercises especially in subgroups (see Rose, 1997, for a description of most of these) also enhances broad participation.

6.15.3 THERAPIST ACTIVITIES IN CBIGT

One cannot merely assume that if a problem and the appropriate coping skills are identified, one can then intervene at any time in the therapeutic process with the best techniques available and help the patient to solve it. In CBIGT just as in other group therapies, one can identify a number of sets of activities each of which is linked to unique and overlapping therapist functions. They are somewhat overlapping in time and in content. Differentiating these sets is useful insofaras each set of activities provides a guide for the group therapist as to when emphasis should be shifted. These include a pregroup planning, an orientation, an assessment, an intervention, a generalization and termination, and a postgroup set of activities.

6.15.3.1 Pregroup Planning

In planning for therapy, the CBIGT therapist must establish the group's purposes, assess potential membership, recruit members, decide on the group social environment or structure, and create the group's physical environment. In determining the group's purposes, the group therapist can draw on several sources. Based on experience with patients in the agency or community, the agency may have identified a need for a certain type of program. Patients may have been requesting help for certain types of problems for which help is not generally or readily available. To the degree that these needs can be translated into behavioral or cognitive responses to identifiable problematic situations, the CBIGT approach can be considered. Decisions must be made as to the theme of the group, group size, number of therapists, frequency and length of sessions, a sufficient number of candidates for the group, and group composition.

In citing the literature above, it appears that most cognitive-behavioral groups have a theme, that is, all patients in the group have similar problems. A few practitioner–researchers (e.g., Flowers & Schwartz, 1985; Manning et al., 1994) also use CBIGT to deal with more heterogeneous groups (i.e., groups consisting of persons with diverse presenting problems) with positive results. Many of the homogeneous groups, although comprising members who have similar presenting problems, are diverse in terms of race, gender, and ethnic background, although in most of the reports from the literature the groups are predominantly White middle class. Some groups are homogeneous in terms of gender and race. Some single case research and descriptions of clinical experience point to the efficacy of CBGT (although not necessarily CBIGT) approach with Hispanics (see Comas-Diaz & Duncan, 1985), Native Americans (La Framboise & Rowe, 1983; Schinke & Singer, 1994), and economically disadvantaged depressed women (Azocar, Miranda, & Dwyer, 1991). Wolfe (1987) makes use of all-women groups because consciousness raising is a component of the CBGT model she employs. Wolfe (1987) asserts that CBGT and in particular, rational emotive therapy in groups, seem to come the closest to meeting the criteria for feminist therapy. Sometimes all-women groups exist by default rather than by plan as in the case of many parents groups or groups of depressed persons where only women sign up. All-men groups in CBGT can be found for male abusers, anger management groups in prisons, and men who are HIV positive.

At least two kinds of basic organization for therapy groups can be identified. The first is open-ended groups in which new group participants come at any time in the history of the group and leave at any time. These are usually found in institutions although some community groups are organized in this way as well. It is a more common model with support groups than CBGT groups. The model duration of most open-ended groups is approximately a year, although some groups may go on for much longer. However, participants may stay for any length of time, and they often come and go sporadically. The session length varies from one to two hours.

The second type of organization is closed groups which have a fixed beginning and fixed end-date for all participants. Most of the available research is on this type of group and most cognitive-behavioral community groups are closed. Although there are many exceptions, most closed groups last from 6–16 sessions once a week for one and a half to two hours. The modal number is eight. Experience indicates that 12–16 sessions are required to achieve complex or multiple goals. There is usually only one group therapist except in training situations. The size of the groups varies from as few as four to as many as 20 participants. However, with adults the modal number is eight, which permits participation by everyone at any given session and provides a wide variety of ideas and experiences. Most CBGT groups are sponsored by social agencies, clinics providing mental health services, schools and colleges, social welfare agencies, and private practitioners.

6.15.3.2 Orientation to Group Therapy

As part of recruitment and later during the first group sessions, members are oriented to the purposes of the group, the methods to be used, the potential goals that can be achieved, and the importance of keeping what goes on in the group confidential. An overview of the group activities and expectations are presented and discussed. As part of orientation, group contracts are often developed; these contracts establish what the patients can expect from the group therapist and the agency and what the members can be expected to do. These contracts are often in writing. Patients are also oriented to the basic assumptions underlying each of the treatment techniques used. In this way not only do patients know what is happening to them, expectations of positive outcomes can be stimulated.

In orientation, the therapist in CBIGT also introduces the patients to the use of the group as a vehicle of treatment. The use of subgroup exercises (Rose, 1997) are also initiated. The therapist draws on patient experience to orient members to the assumptions of the approach. Broad group discussion is encouraged by means of assigned tasks in subgroups and/or in exercises, which is the first step taken to enhance group cohesion.

Other means of building group cohesion are by using introductory exercises in which members interview each other, keeping the tone and interaction of the group positive, encouraging others to provide frequent mutual reinforcement, noting similarities as well as differences among members in terms of background and presenting concerns, keeping the group small, providing occasions for the patients to help each other, permitting and encouraging physical movement during the session, using humor, using role-playing, providing variation in program and program media, and providing or having the members provide refreshments (see Rose, 1989, pp. 250–251, for additional detail.).

One can also develop cohesion by enhancing the motivation of the members. Groups can serve to develop motivation primarily if the group therapist takes specific actions (see Miller & Rollnick, 1991, for a discussion of these principles in more detail as they apply to the individual therapy of people who abuse drugs and alcohol). In the first place the therapist should accept and encourage the patients to treat ambivalence or reluctance to participate as normal phenomena. No pressure is placed on the participants initially to change or to get out of "denial" either by the therapist or the other group members. The members are encouraged to review the advantages of the problem state as well as the disadvantages and to weigh the relative merits of changing or not changing. The patients are taught to identify motivational statements made by one another and to reinforce each other as they occur. Throughout the process the therapist makes ample use of empathic statements which are often emulated by the members with each other. The level of motivation is an ever changing process and requires constant attention by both therapist and group members.

6.15.3.3 Assessment

Assessment is a concept central to all empirical approaches. The purpose of assessment is to determine the specific targets of interventions and the specific coping skills to be learned, in such a way as to make them amenable to intervention, the situations in which these coping skills and other target behaviors should be applied, the social and material resources of each patient that might impinge on treatment, and the potential barriers to effective treatment. It has the additional purpose of determining whether the given group or another type of therapy might be the most appropriate setting for each potential patient. Finally, it forms the basis for establishing treatment goals.

The goals of CBIGT may be changes in the level of intensity of specific behaviors or an increase of more general adaptive coping behaviors. Among those specific target behaviors that patients have worked on in CBIGT and on which research exists have been mentioned earlier. These include reducing the extent or intensity of agoraphobia, social phobia, mood swings, obsessive-compulsive behaviors, smoking, alcohol and drug abuse, binging and purging, stress reduction, anger responses, and pain responses. Positive targets involve sleeping regularly, specific ways of making and keeping friends, increasing social activities, making more effective use of the patient's social network, improving relationships and communication skills. Another set of target phenomena are the development of coping skills which mediate resolution of the target behavior. Two kinds of coping skills can be identified, cognitive and behavioral.

Cognitions refer to thoughts, images, thinking patterns, self-statements, expectations, or other private or covert events which may be inferred from verbal or other overt behavior. Cognitive coping skills are those cognitions which facilitate coping with internal and social phenomena. Examples are skill in analyzing one's own cognitions, in labeling appropriately one's self-defeating self-statements, in observing and rehearsing new, more appropriate self-statements, and in reinforcing oneself covertly. Though important skills in their own right, some cognitive coping skills also mediate the attainment of the more observable social skills and other behaviors mentioned above or other coping behaviors. Thus, the goal of increasing cognitive coping skills is important as a means of reducing the frequency of anxiety-inducing and behavior-inhibiting cognitions, in diminishing the intensity of anxiety, and in improving social behavior (Beck & Emery, 1985; Meichenbaum, 1977). Self-statements, such as "everyone thinks I'm strange," not only may produce anxiety, but promote inaction or self-defeating action. Changing such self-statements to something like "Sure, I'm different than others in many ways, some things I like and some I'll change" may reflect a more accurate appraisal, may suggest avenues of change, is more self-respecting, may reduce anxiety, and ultimately should improve social behavior. In order to teach these coping skills, assessment involves having the patients examine each others cognitions in the situational analysis of stressful or anger-inducing situations. Checklists (e.g., see Piacentini, 1993, for a description of most of these) are also used: keeping track of one's cognitions with such procedures as Beck's (1976) three-column technique in which the patient records in the first column an anxiety- or anger-producing situation; in the second, his or her automatic thoughts or thinking errors; and in the third, types of errors found in these thoughts. These are then shared with other members of the group.

Another set of cognitive coping skills is the problem-solving sequence described earlier (D'Zurilla, 1986). In assessment, the therapist determines the patterns of problem-solving of each of the patients by means of the social problem-solving inventory (D'Zurilla & Nezu, 1988). Simulated problems can also be presented to the group for the members' solutions. The therapist can observe the process. Those who are either impulsive or chaotic problem-solvers would profit from being able to approach problem situations in a careful, step-by-step, analytic, and planful way.

Self-management refers to those cognitive coping skills by which patients control their own environment as a means of controlling their own behavior. Procedures such as the use of environmental cues, self-monitoring, self-instruction, self-modeling, self-evaluation, and self-reinforcement fall under this rubric. These are used in changing patterns of smoking, drug and alcohol abuse, studying, and other problems of self-control, which are targets of change or the themes of many groups in CBGT. As part of assessment the therapist ascertains the degree to which these self-management skills are present and the target behaviors in which the self-management skills can be directed primarily by means of patient self-report and situational analysis. These self-management skills, if learned, represent a set of strategies that the patient can utilize even when only limited external support is available.

Among the most common behavioral coping skills are social skills, a set of learned performance behaviors that relate an individual to others. These include such behaviors as responding to criticism, techniques of dealing with other people (family, friends, colleagues, strangers) in stressful situations, asking for help when needed, disagreeing constructively. They may even be more specific such as giving appropriate eye contact in conversations or smiling occasionally. These social behaviors are often essential for effective social fuctioning and can be readily taught in the group. Social skill inventories (e.g., Gambrill & Richey, 1973) and role-play tests (Magen & Rose, 1998) are commonly used to assess patient skills in this area.

Recreational and leisure time skills may also be regarded as behavioral coping skills. The extent of these interests are explored in assessment to determine whether they should be amplified or modified. Although some are social in nature, they may be regarded as a separate category for coping with general life stress. Information about these skills and interests are obtained in a group exercise in which the members interview each other in pairs about their interests and leisure time activities and then each reports his or her partner's interests and skills to the group.

Because therapy does not go on forever, patients will have to learn to make use of their social network more adequately. A number of social network surveys have been developed which the patients fill in and ascertain the relative helpfulness and limitations of the various social units of which they are a part (see e.g., Rose, 1998b, p. 329, for an example). In the group they share the results of the survey with other members.

Numerous other noninteractive coping skills can be explored as part of assessment such as the patients's ability to manage his or her time, or to relax or mediate during stressful situations. Skills in relaxation can be readily observed as they are taught and tried out in the group. Relaxation is frequently used as part of most cognitive-behavioral programs. In most cases, coping skills and specific behavioral targets are intertwined.

One cannot learn coping skills in a vacuum. One must identify the specific situations with which the individual must learn to cope. To this end the group members are taught to identify and define their unique problematic situations. The therapist first provides models of such situations, which the members discuss in terms of the criteria for formulating troublesome situations they have thus far had trouble dealing with. For example, in a group of men working on reducing the frequency of drinking behavior, an example of a situation in which a coping response was required was as follows:

> A friend of yours stops you after work. He urges you to have drink with him. He notes that several other old friends will be there. You have refused and then he says, "what kind of friend are you, if you can't have a drink with your old buddies."

A group of patients who have problems managing their stress are presented with the following example of a situation that might increase stress and which they will analyze as a group:

> Your car has just been ticketed. You are only eight minutes late for your meter. You really rushed to be back on time. You tell yourself it isn't fair. You know people who are hours late and don't get parking tickets. You feel the tension rising in your stomach.

An example used for a group of men who batter is the following:

> You get home right on time. You've had a couple of drinks, nothing much, you feel, and your wife says she smells alcohol on your breath, you feel anger rising, your head hurts, you say to yourself she can't talk to me that way. If she loves you, you think, she won't nag me. It's my life, I'll drink what I want.

The criteria for a situation that readily lends itself to examination are that the situation is important to the patient and likely to occur again, that it is specific as to time and place and the people involved, and that it represents a situation the patient might have difficulty in dealing with. After discussion of the application of these principles to the model situations, as an extragroup task it is suggested that each person develop one or more such situations which would be presented at the following session in the group in terms of how well each situation meets the criteria. After working with a group for a while, the therapist can put together a role-play test consisting of a wide variety of such situations (see e.g., Magen & Rose, 1998) to ascertain the social skills needed to deal with social situations the patients with a common problem area encounter. The group therapist develops a set of predetermined or "canned" situations for those members who cannot develop one for themselves. An exercise is used to train the members in defining relevant and useful situations. In this exercise a number of models of situations are presented and the members analyze whether the criteria for "useful" situations are met.

As part of assessment and to determine the progress of treatment, data on patient behavior and the resolution of problems are usually systematically collected before, throughout, and following therapy, and several weeks or months after therapy. In order to understand patients' reponse session by session to the program as a whole, a postsession questionnaire is filled out by all the patients and the therapist at the end of each session (see Table 1). In summary, some of the methods of collecting data mentioned above include diaries, personality inventories and checklists, role-play tests, sociometric tests, self-observation, direct observations of the group or of individuals when not in the group, postsession questionnaires, and interviews. (Each of these are discussed in more detail in Rose, 1989, pp. 109–136 for adults, and Rose, 1998, pp. 130–151 for adolescents.) The more structured methods permit evaluation of outcome, one of the purposes of assessment.

Based on the initial data collected in the first part of assessment, goals are eventually established together with the patient. These goals are usually in terms of the specific target behaviors and cognitions that the patients need to achieve by the end of treatment and the coping behaviors required to deal with the problem situations. Knowing each other well because of their intense interaction, the group members provide each other with ideas as to goals each might pursue and feedback as to the formulation of the goals, the importance of the goals, and the degree to which these goals are realistic. Goals may be either behavioral, affective, and/or cognitive change.

Some examples of common treatment goals that people have worked towards in these groups are:

1. (For the depressed patient) By the end of treatment, I will participate in at least two social activities every week with friends or family members. I will also identify any self-defeating thoughts I make and will change each into a coping thought. On my self-monitoring card, I will score an average of 5 or lower on the 10 point depression scale.
2. (For the angry and abusive male) By the end of treatment when I begin to feel angry, I will take a deep breath and let it out slowly and remind myself that I can destroy my marriage if I give way to it. When I experience the first signs of anger with my

> wife, I will excuse myself from the situation and walk away. I will also increase the number of compliments I make to my wife to a minimum of once a day.

The criteria for effective goal formulation are that the goal is important to the individual, that a time frame be provided (e.g., by the end of treatment or by next month), that the goal be sufficiently specific that the patient knows when he or she has achieved it, and that the goal is realistic. If necessary, subgoals may also be formulated which are even more concrete.

The patients are trained in goal formulation by means of an exercise in which the model goals such as those above are provided, the above criteria for effective goals listed, simulated goals are presented and these are corrected in group discussion. Then the members formulate their own goals which are evaluated by the group in terms of how well the goals meet the above criteria. Once goals are formulated, an additional means of measurement, goal attainment scaling, can be used to ascertain progress towards goals (see Cardillo, 1994). The goal attainment scales are developed in the group with the members helping each other. Following the formulation of goals, group programs and interventions can be planned which facilitate the achievement of these goals.

6.15.3.4 Intervention

Based on the goals, intervention activities change or maintenance or decisionmaking strategies are selected. The patients are oriented to these procedures, and with their concurrence, the strategies are applied. Specific strategies are selected which have empirical support as well as a relationship to the goals.

In order to develop these and specific behavioral and coping skills, usually no one intervention technique is sufficient. A number of methods of teaching patients necessary specific skills include problem-solving, modeling (overt and covert), rehearsal, coaching, cognitive-restructuring, rational-emotive techniques, self-instructional training, reinforcement and stimulus control, sociorecreational, relaxation training, and small-group techniques. In CBIGT most of these methods with special emphasis on the group are combined into one integrated approach. A method is selected for inclusion preferably if it has some independent empirical foundation and some relationship to the above mentioned targets. In the following sections, most of these methods will be reviewed in terms of their contribution to the total approach.

6.15.3.5 Generalization

Generalization, a concern of all therapies, refers to the process of transferring what the patient has learned in the group to the outside world and maintaining what he or she has learned beyond the end of therapy. The most fundamental principle of generalization is that it rarely occurs without taking steps to see that it occurs (Stokes & Baer, 1977). One of the major strategies in CBIGT for transferring learning from the group to other situations involves the use of extragroup tasks, described above, which are used in almost every session.

The patients are prepared by modeling and rehearsal on how to deal with persons unsympathetic to their changes. Former patients are sometimes brought in to discuss the possibilities of setbacks and how they might be handled. Possible self-referral sources are discussed, such as a counselor or a local clinic or health service. Finally, as part of this phase, a follow-up or "booster" session would be held two or three months following therapy.

The members are taught the general principles of what they practice specifically. For example, they learn the general steps for problem-solving. They learn the principles of how anxiety and stress is in part a function of how we evaluate situations. They learn the principle that if one practices a given behavior it is more likely that they can learn that behavior. In CBIGT one goes from the specific exercise and experience to learning the general principles behind them.

Members are finally prepared for termination by getting them to develop a plan for how they intend to apply what they have learned in the group when the group ends, and by designing activities appropriate to practicing their newly learned skills. These plans often contain such actions as joining a nontherapeutic group, reading a self-help book, practicing relaxation on a regular basis, meeting with a group member to talk about the principles learned in the group, and meeting again for a booster session.

In order to diminish the intensity of the relationships of the members with each other and with the group therapist as the group approaches termination, the therapist encourages members to establish relationships outside of the group and to become involved in extragroup activities such as family activities, bowling leagues, bridge clubs, dancing lessons, and social organizations. Furthermore, relationships with nongroup members are encouraged. These new activities and social relationship become the focus of the later sessions in CBIGT. Extragroup tasks become less structured but more extensive. Preparation is largely in the hands of the patient. Monitoring is less strict. Social, recreational, and other cohesion-building activities are kept to a minimum in the group. Many of the leadership functions are performed by group members (see Table 2).

6.15.4 INTERVENTION STRATEGIES USED IN CBIGT

The most commn intervention strategies used in CBIGT are systematic problem-solving, the modeling sequence, cognitive change procedures, relaxation, reinforcement and stimulus control, and small group procedures.

6.15.4.1 Systematic Problem-solving

Systematic problem-solving is a central method to CBIGT insofaras patients bring problematic situations of concern to the group and the group under the guidance of the therapist attempts to help them find solutions to those problems. It is systematic insofaras the members follow (or deviate by plan from) specified steps. The steps characteristic of the problem-solving process include orienting the members to the basic assumptions of problem-solving, defining the problem, generating alternative solutions, selecting the best set of solutions, planning and preparing for implementation, implementing the solution, and evaluating the outcome (except for "preparation for implementation" which has been added, these steps are taken from D'Zurilla, 1986). Many of the steps of problem-solving can be classified other than as intervention. For example, orientation to problem-solving is an orientation activity, defining the problem is clearly a part of assessment, implementation through extragroup tasks is part of the generalization of change, and the other steps are intervention activities.

Problem-solving is not only a general paradigm. It can be identified as a set of coping skills to be learned in their own right to replace impetuous or uninformed problem-solving. In using the problem-solving method to find more effective ways of copying with the many problems of concern to group members, patients learn the skills involved in carrying out the general problem-solving paradigm. Such a set of procedures can be effectively taught in groups since the general paradigm is repeated frequently in the course of therapy not only dealing with individual problems but also in dealing with group problems should they arise. Moreover, once learned, it can be used by patients to deal with problems long after the group has ended.

Within the framework of problem-solving, a number of other cognitive and behavioral procedures are used to prepare the group members to impliment the solutions they have agreed to. Each of these have their own empirical foundation. One of the most important is the modeling sequence.

6.15.4.2 The Modeling Sequence

This sequence is designed to teach specific interactive behaviors for coping with various problems situations, and includes such techniques as overt modeling, behavior rehearsal, coaching, and group feedback. Modeling refers to learning that occurs through the observation of a model who might be the group therapist, another member of the group, someone in the patient's environment, or an admired person on stage, on screen, in novels, or in public life. Modeling may be role-played in the group or it may be observed directly in real life. Modeling is specifically used to demonstrate how a situation problematic to one or more patients in the group may be handled effectively. Behavioral rehearsal is a role-play technique in which a patient with a given problem situation practices new, more effective ways of handling that situation. Coaching refers to instructions, verbal or physical cues given to the patient when she or he is modeling or rehearsing a set of behaviors in a given situation.

Group feedback is the verbal evaluation from others as to how effectively the patient role-played or modeled. Following the modeling sequence, the patient prepares for and carries out extragroup tasks to practice the newly learned coping skill in the real world.

> Following the situational analysis of the situation calling for a refusal of drink when offered (see example above), the therapist asks the group members for possible responses to the one of agreeing with the friend and going with him which is what the patient, Pete, had done in the original event. Based on the suggestions of his fellow patients, Pete decided that he would like to refuse clearly and in a matter-of-fact tone of voice, reminding his friend that he (Pete) was on the wagon and was working hard to stay there and he would appreciate whatever help his friend could give him. Pete would repeat the statement if pressured further. The therapist set the scene outside a bar, and then modeled the situation with one of the other group members acting as the friend. Then Pete, who felt that one demonstration was enough, played himself and his friend was played by the same person as in the first role play. When it was finished the therapist asked the other patients to tell Pete what he did well and then what he might consider doing differently. The members responded that Pete made an impressive statement and didn't argue. They suggested that he might consider giving better eye contact and speaking in

Table 2 Group development in a structured group.

Group phase	*Therapist behavior*	*Group processes*
Orientation phase	Orients members to theory and approach Ascertains levels of motivation Encourages broad participation Orients members to each other Introduces rules Promotes protherapeutic norms	Factual communication, limited self-disclosure High mutual anxiety Feedback is limited and polite Norms loosely established Leadership functions primarily controlled by therapist
Preliminary work phase	Stimulates moderate self-disclosure Examines problem situations Extragroup tasks introduced Trains in effective feedback Teaches basic concepts Initiates brief role-plays	Weak subgroups begin to form Increase somewhat in cohesion Members focus attention on therapist Feedback is descriptive and positive Distribution of participation broadens
Deterioration or conflict phase	Examines nonproductive activities Group processes becomes more flexible and empathic Begins to look at consequences of present behavior Initiates discussion of group problems Introduces concept of cognitive distortions Introduces group problem-solving	Homework completion rate drops Feedback is more negative Anger and withdrawal begin to intensify Cohesion begins to weaken Roles and norms are challenged by some members
Resolution phase	Assists members to assume responsibility for therapy Encourages leadership from members Deals with complex situations	Feedback to each other more constructive Cohesiveness decreases slightly Group engages in systematic problem-solving Distribution of participation increases broadly Interaction more task oriented New protherapeutic norms Higher levels of cohesion
Secondary work	Therapist reduces own activity Reduces frequency of reinforcement Encourages members to work on complex problems Interaction highly task oriented Rate of homework completion high	Members assume major phase leadership functions Significant self-disclosure to each other Cohesion decreases slightly Rate of homework completion high
Termination phase	Points to principles of generalization Prepares for termination Increases own activity slightly Summarizes progress Helps individuals to plan post-therapy activity	Group is more spontaneous Cohesion diminishes, focuses on extragroup social units Group focuses on future actions

> a more matter-of-fact tone of voice. Pete rehearsed one more time, trying to incorporate the suggestions with the "friend" in the role play being much more insistent.

The therapist is the director of the role-play. He or she coaches significant others and the target person when necessary. He or she stops the role-play before it goes on too long or if the person cannot handle the situation. He or she instructs the observers as to what they should observe. He or she makes sure the tempo is reasonable. He or she makes sure that the role-play remains task oriented. Modeling is an excellent group procedure because of the presence of many potential models and rich source of feedback, and the opportunity for client leadership if these characteristics are taken advantage of.

6.15.4.3 Cognitive Change Methods

Cognitive change methods refer to the steps taken to train the patient in more effective ways of thinking about or evaluating him or herself as the patient responds in specific problematic situations. In groups many different cognitive procedures are used, often in combination with each other and with other types of procedures such as the modeling sequence. It is assumed that in a given set of circumstances cognitions in part mediate overt behavioral and affective responses. These cognitions include how one values oneself and one's action and how one specifically thinks in or evaluates a given situation. In CBIGT the most commonly used cognitive procedures are cognitive restructuring (Beck, 1976), which includes rational emotive techniques (Ellis, 1973), and self-instructional training (Meichenbaum, 1977).

6.15.4.3.1 Cognitive restructuring

Cognitive restructuring is characterized by two sets of procedures. First, one set of procedures is to identify the distorted patterns of thinking and/or dysfunctional schemata (Beck, 1976) which interfere with social functioning or create intensive emotions. Second, a set of methods are used to replace such distorted or dysfunctional thinking with self-statements that facilitate effective coping with day-to-day life events and reduce anxiety and stress. In the first set of procedures the patients are trained to identify cognitive distortions in case examples or group exercises and to label them into such categories as absolutizing, catastrophizing, mind reading, selective perception, or prophesizing (Beck, 1976). In one exercise the group members are given a list of self-statements which represent both cognitive distortions and coping statements. They are asked to identify independently to which category each statement belongs, and if a distortion, to identify the nature of that distortion and how it might interfere with social functioning and the eliciting of strong emotions. In the group the members exchange their ideas until each person has a clear picture of what a cognitive distortion is and why a given statement is a cognitive distortion.

Later using the theoretical framework they have learned, they are better able to identify their own unique patterns of cognitive distortions, dysfunctional thinking, or illogical expectations and those of their peers. Continuing with the above exercise the clients are asked to identify one statement they are likely to make which they might be identified as a cognitive distortion. To help each other in the process of identification of each person's distorted thinking, the other group members are taught to "rub in" the errors in each person's thinking through logical disputation and analysis. Supportive evidence is also sought through interviewing by the other group members either to invalidate or validate the assumptions of each client. In addition, as members relate ongoing problem situations to the group, they are asked to note their thinking in these situations and to identify and ask the peers to identify any cognitive distortions.

Once the distorted thinking is identified and labeled, each patient, assisted by the other group members, is asked to replace the distortions with coping statements. Group exercises are used to facilitate this replacement in which the patients first correct the list of "canned" statements of distorted or dysfunctional statements. To facilitate this process, a list of coping statements is presented as points-of-departure. For example, "if I take one step at time I can handle this," or "I should remind myself to take a deep breath and relax," or "if I make a mistake it's not so terrible; nobody is perfect."

Once the members become skilled in replacing cognitive distortions with coping statements on simulated examples, they can begin with the help of the group to replace distorted thoughts of their own with coping statements. If they have difficulty the group can provide each other with corrective information or they can "brainstorm" alternate ways of thinking about a given situation which might help them to cope with the situation.

Finally, the members assign each other tasks to perform outside of the group to try out and self-monitor their use of coping self-statements to replace cognitive distortions. At a subsequent session each patient reports back his or her observations to the group.

6.15.4.3.2 Self-instructional training

In self-instructional training, the members are taught first in the face of a problematic or stress-inducing series of situations to utilize functional self-statements at each step of the series. This process consists of step-by-step verbalizations concerning the problem definition ("What's wrong with the way I'm thinking about this?"), problem focus ("What can I do about it?"), focusing of attention ("I should think about how that will get me in trouble."), coping statements ("If I keep relaxing I won't blow it!"), and self-reinforcement ("Wow! I did it! See, I can do it!"). In groups, the members are trained in the method first by providing the members with a model. Examples of situations in which self-instructional training would be appropriate are the following: (i) a situation in which a person is unfairly critizing the patient and escalates the criticism over time; (ii) a situation in which the patient is being persistently hassled to do something she does not want to do; and (iii) a situation in which a person must make a presentation to colleagues and others of which he has little experience. If a given member identifies such a situation, the members brainstorm coping self-statements the given patient might make in the several steps involved in learning to cope with the situation. A group member models the statements (cognitive modeling) for the given patient who then practices the same statement (cognitive rehearsal). The given patient is assigned the task of self-monitoring how he or she handles the given situation should it occur, and reports his or her experience back to the group. Often this method is combined with social skill training.

Other methods with cognitive as well as behavioral elements used in groups are thought stopping in the face of persistent recurring thoughts, systematic desensitization or exposure methods in the case of phobias, and guided imagery. Gradual exposure methods have been used in the treatment of agoraphobia (Hand, Lamontagne, & Marks, 1974) and obsessive-compulsive disorder (Fals-Stewart et al., 1993) with positive results. In the treatment of agoraphobia the group members accompanied each other to face the feared setting together. The group served as support and encouragement for each other. Later the members accompany one person at a distance and eventually each patient faces the feared situation alone.

In the group treatment of obsessive-compulsive disorder (OCD), Fals-Stewart et al. (1993) used both exposure to the feared object (e.g., dirty materials, garbage cans, etc.) and response prevention (e.g., by preventing the individual from washing their hands for at least an hour after exposure to a contaminant). Where *in vivo* exposure was not feasible, imaginal flooding was used. This combination seemed to be effective in alleviating OCD symptoms of patients in groups as well as in individual therapy. The groups are used effectively because of the multiple modeling effect. Also members can help plan each other's hierarchy of anxiety-provoking activities with which they would expect to be confronted in the near future.

6.15.4.4 Relaxation Methods

These are strategies for helping patients to cope with strong emotional responses such as anger, stress, or depression. Thus, relaxation training is an intervention and the skill in relaxation is a coping target in its own right. The training involves teaching patients a modified version of the alternate tension and relaxation technique (developed by Jacobsen, 1978, and adapted by Bernstein & Borkovec, 1973) and then later fading the tension phases. Various alternatives uniquely suited to various populations are also taught. Modest research support for the use of relaxation procedures in reducing anxiety and stress is to be found in studies by Stovya (1977) and Lyles, Burish, Korzely, and Oldham (1982). However, Heide and Borkovec (1983) warn that for a few persons relaxation may increase anxiety. In groups the members are taught the procedures by the therapist modeling the procedure then by having them practice the procedures a few muscle groups at a time. Once the therapist has relaxed the entire group together, in order to make use of the group, the members then go through the same procedures with each other in pairs. The therapist prompts the "teachers" and monitors the process and is available if someone gets stuck. This affords each member the opportunity to be both teacher as well as learner. The members are given tapes and encouraged to practice the relaxation out of the group, if possible with a "buddy" so that they can monitor each other's practice. In addition, meditation (Carington, 1978) and diaphragmatic breathing in some groups are taught as alternatives to deep muscle relaxation.

6.15.4.5 Reinforcement and Stimulus Control

These are not central methods in most applications of CBIGT except in parent training groups where the parents are taught to modify antecedent conditions and consequences with their children. Reinforcement is derived from operant theory and involves the target person performing a target behavior followed by a reinforcing event (positive reinforcement) or the withdrawal of an aversive stimulus (negative reinforcement). Both forms of reinforcement increase the probability that the target behavior will increase under similar conditions in the future. Stimulus control involves procedures in which the immediate conditions which lead to or are parallel with a given behavior are changed to create conditions more amenable to the performance of a desired behavior. Creating conditions in the group through subgrouping as a means of getting broad participation is an example of modifying the stimulus conditions. In groups patients receive many kinds of reinforcement for the performance of prosocial group behavior and the completion of extra-group tasks. With adults, this reinforcement takes the form of praise by the group therapist or other group members. Occasionally it takes the form of smiles, applause, approving nods, and delighted laughter. Also reinforcement is an effective communication skill in its own right if carried out appropriately. Finally, high levels of mutual reinforcement have been shown to to increase the cohesion of therapy groups (Goldstein, Heller, & Sechrest, 1966). For these reasons every effort is made to include as much reinforcement as possible in the early sessions. Reinforcers can also be withheld in response to undesirable behaviors. This is referred to as extinction and is an occasional response in groups when someone is frequently off-task or complains a great deal. Extinction is difficult to perform in groups since it requires the cooperation and prior agreement of all group members along with the therapist.

6.15.4.6 Small-group Procedures

Because CBIGT is also a small-group approach, it is possible to take advantage of the multiple interaction to enhance the effectiveness of the intervention and assessment procedures described earlier and to provide entirely new ones. Since most group therapies make use of some of these techniques, only a few are illustrated in terms of their unique application in CBIGT. These include role-playing, the buddy system, subgrouping, leadership delegation, and group exercises. (These and other group procedures commonly used in CBIGT are described in detail by Rose, 1989, as they apply to adults and Rose, 1998a, as they apply to children and adolescents.)

6.15.4.6.1 Role-playing

Role-playing is a set of procedures that has several purposes in CBIGT. First it is used to demonstrate or model behaviors targeted in problem-solving and for the patient to practice (behaviorally rehearse) them before trying them out in the real world as part of the modeling sequence discussed earlier. It is highly focused and followed by highly specific feedback. Role-playing may also used for assessment by having a patient role-play a problem situation as much as possible as it originally occurred. A set of situations may also be presented to the patients who role-play their responses. It is also used to teach specific therapy skills such as providing constructive feedback and receiving critical feedback, and in generalization training to practice preparing for an unsympathetic environment or the eventuality of setbacks. When using role-playing, practitioners have found it useful to encourage physical movement (knocking at and opening imagined doors, sitting and standing when appropriate), separate the players clearly from the observers, give the observers a clear observational task, and to keep the role-plays brief at least in the beginning. Where patients are reluctant to participate, the therapist will model role-playing with colleagues, write out scripts, and let members express their thoughts with "I would say ... " before having them actually act it out. Usually in groups there are some members who are willing to try it out first and who consequently serve as models for the others.

6.15.4.6.2 Subgrouping

Subgrouping is a simple procedure of working in pairs, triads, or other sized subgroups as a means of increasing interaction among the members and providing them an opportunity to work without the constant oversight of the therapist. Many of the group exercises are performed in subgroups. Subgrouping also creates an opportunity to practice leadership skills. Care is taken to assign subgroups with various compositions in order to avoid "clique" formation which might work against the pursuit of treatment goals. In working with subgroups, the therapist must give them clear assignments, the results of which they report back to the larger groups. He or she must also carefully monitor the activities of the subgroups to make sure they are working on the given tasks.

6.15.4.6.3 Buddy system

The buddy system is a special subgrouping procedure for patients to work together outside of the group (see O'Donnell, Lydgate, & Fo, 1979, for a review of the research with children). In CBIGT the buddy system is especially useful in facilitating completion of the group tasks. Buddies check with each other on their respective progress and reinforce each other for small successes. In addition to the advantages mentioned above, it contributes to the transfer of learning within the group to situations outside of the group.

6.15.4.6.4 Group exercises

Group exercises refer to the use of structured interactive activities as ways of teaching patients the skills which mediate the achievement of therapeutic goals. For example, an introduction exercise is used in which patients interview at least two patients in the group and introduce them to the others. Another exercise is one in which the patients study a case and discuss how each of them is different from the person in that case. In stress groups an exercise is used in which members identify their own unique stress response (see Rose, 1998, for a list and description of commonly used exercises, most of which can be used with adults as well as children). Often the exercises are performed in part in subgroups of varying sizes. Each exercise has a rationale, a stated purpose, and a set of instructions for individuals and for the group as whole. In one particular application of CBIGT, exercises are used in every session. The exercises serve to provide variation in program, enhance interest and cohesion, and create opportunities for broad participation.

6.15.4.6.5 Group feedback

Group feedback is another group technique that plays an important role in group therapy. Members are constantly giving each other feedback as to what they do well and what they might consider doing differently. In CBIGT the patients are trained in how to give feedback in such a way as to maximize its positive reception. The criteria which the participants are trained (in exercises) to use are to give positive feedback first, describe what was done in specific terms, and use an I statement. In giving critical feedback the members are asked to affix the word, "what I might do differently is ..." or "one thing you might consider doing differently is ..." In order to protect the recipient of feedback, he or she can stop feedback any time he or she chooses. Even with these protections the therapist must be careful to monitor feedback to make sure that no one is being devastated by what he or she hears. Usually the norms of giving helpful feedback are quickly established and the group members correct inappropriate feedback whenever it should occur. The group members in contrast with the individual therapist provide a wide range of diverse observations and feedback. The danger of multiple comments being too destructive or aversive can be limited by training the members receiving and giving feedback and by careful monitoring by the group therapist.

6.15.4.7 Identifying and Resolving Group Problems

In all groups one can observe a number of group problems. These can be defined as an intragroup interactive event (or series of events) or a product of interactive events which interfere with effective member task performance or goal attainment. The responsibility for achieving a change in group process cannot usually be linked to a change in behavior of any one member or the group therapist, but to interactive changes among all or most of the members and the therapist. Several problems in particular stand out such as when group cohesion is too low, when the communication pattern primarily directly from members to the therapist and not to each other, when self-disclosure is too low, when some members dominate the interaction while others are virtually excluded, when antitherapeutic norms exist such as a pattern of lateness, and where members are frequently off-task. In general, most of these group problems are avoided by careful treatment programming. Occasionally when a group problem persists, it is necessary to use systematic problem-solving.

The therapist points out the problem as he or she experiences it and checks with the group if they perceive it. Often the therapist can back up his or her perception by the results on the postsession questionnaire or other ongoing measurement procedure. If in discussing the problem, blame is laid at the feet of the therapist, one member, or one subgroup, the therapist points out that the problem cannot exist without the shared contribution of all members and the therapist. It requires a change in behavior or attitude of everyone. If the group agrees, the members are asked to brainstorm what the group can do, what each member him- or herself can do, and what the therapist can do to ameliorate the problem. Then the group uses the ideas to negotiate a group plan and a therapist plan and each person commits him- or herself to an individual plan all of which are found in the following example.

> In a hospital group of patients with borderline personality disorders, the therapist noted that everyone seemed to be addressing all their questions to her and not to each other, in spite of the fact that many of them knew the answers better than the therapist. The co-therapist had been noting the direction of all interaction at the past session and provided the group this information. The group members agreed, stating that it was easier, somehow, to do that, and besides she had the most power. The therapist pointed out that eventually they would have to depend on their own resources and those of their peers, and this was a good opportunity to practice solving a problem. Since the group members agreed with her assessment, she suggested that everyone write down one idea as to what the group could do differently, what the therapist could do differently, and what each of them could do differently. An extensive list for each category was generated, and after an evaluation, the group selected for the group that whenever a question is asked, everyone would write down what

> they each thought was the answer. Then the person asking the question would ask the others what they had written. They agreed that the therapist would throw back to the group any question directed to her and remind them of the agreement. The individual plans included relaxing before the session and before answering a question, looking at other members instead of the therapists, and listening more carefully to each question asked by writing down the question.

6.15.4.8 Phases of Group Development

Many of these group problems are a function of the phase of group development in which the group finds itself. Group development refers to the ways in which norms, roles, cohesion, communication patterns, subgroups, and leadership shift over time. In that sense group development is a kind of metaprocess. Common elements seem to run through most paradigms of group development (see e.g., Corey & Corey, 1997; Forsyth, 1990; Garvin, 1987, pp. 110–111; Sarri & Galinsky, 1985; Tuckman, 1965), which suggest that some phenomena are at least in part a function of time. It is useful for group therapists to be aware of group process insofaras the phase of therapy is one of many conditions which may contribute to the appearance of certain behaviors viewed as obstructionist or constructive such as one might experience in the conflict phase.

In Table 2 these often overlapping phases are differentiated into the phase of development, the therapist activity, and the observable group phenomena.

6.15.4.9 Integrating Diverse Intervention Strategies in CBIGT

In the following example the way in which various intervention strategies are developed into an integrated treatment package in CBIGT are examined.

> In the fourth session of a stress management group the therapist reviewed the postsession data from the previous week. He noted that satisfaction and a feeling of self-control of the activities of the group had showed a lot of improvement over the previous session and the members agreed that that was a good sign. Then the therapist asked each member to review what he or she had done on the extragroup tasks during the week. Applause and cheers greeted those who completely or partially succeeded.
>
> A group exercise was used to train all the members in correcting the cognitive distortions that they had identified the previous week in the first part of the exercise. In this exercise each person eventually identified one distortion that each usually made. Following the exercise, each member brought in a stressful situation he or she had encountered the previous week and had recorded in their diaries. For example, Laura told about a commonly occurring situation in which she felt upset by the half joking way in which her boss implies criticism of her in front of others but never states it. She always felt anxious at work and constantly afraid he would criticize her at any time. The other members by asking questions helped to spell out the details of the situation and to relate it to other situations in which she is also often quite anxious. They also examined her thoughts in these situations. In each situation, the group suggested how her self-defeating statements, "I can't stand being criticised," "If I'm criticized, I must be an awful person" seem to keep her in a state of constant anxiety and she agreed that something must be done about these cognitions and some action with respect to her boss was called for. Based on suggestions generated through "brainstorming" from the group, she decided to do two things differently. First, she would let her boss know that she was dissatisfied (while reminding herself that she was a highly regarded employee and she was willing to risk any potential fall-out) with his public suggestion of criticism; if he had any criticism she would like to hear it explicitly stated in private. Whenever she began to feel negative about herself, she would take a deep breathe and relax. In order to prepare Laura to carry out the solution, several members of the group demonstrated in a role play how she might act with her boss, and then she rehearsed what she would actually say first to herself and then aloud to her boss. Then several members demonstrated how, when confronted with implied criticism, what she might think to herself and say aloud to her boss. Laura also rehearsed her thoughts and her actions until she indicated she was comfortable with the way she handled them. Most of the other members handled a similar stress inducing situation with emphasis on the cognitions that prevented them from dealing with them.
>
> Towards the close of the session, all the members designed with a partner their new extragroup task and announced it to the group. For example, Laura reported that she would make an appointment with her boss and then carry out the plan developed earlier in the session. She would also monitor her self put-downs and her self-praise statements throughout the following week (which she has been working on for several weeks) and practice relaxation at least four times for 15 minutes each time. After the others also reported a plan for their extragroup task, they practiced relaxation (all had problems with handling strong emotions) in pairs. The session was closed by filling in the postsession questionnaire.

Although sessions vary considerably, the above session is typical insofaras previous extragroup tasks are monitored, the results of the postsession questionnaire are reviewed and evaluated, a group exercise is carried out, problematic situations are presented by most of the members of the group, alternative responses to those situations are generated, cognitive and overt modeling of the alternative responses takes place, the members practice cognitive and behavioral alternatives for which they receive group feedback, cognitive distortions the client may have had in relation to the situation are considered, relaxation is practiced, and extra-group tasks are planned. The group was used in many ways: the use of role-playing, subgroups, the generation of ideas, group exercise, group feedback, and the structuring of broad participation. All were integrated into a coherent and fast-moving session.

6.15.5 SUMMARY AND CONCLUSION

Some of the research supporting the use of CBGT with various populations has been reviewed. It appears that group therapy is as effective as individual therapy and more efficient. It is also more effective than doing nothing and several other alternative therapies. Follow-up information is either limited or reveals little maintenance. Most of the research on CBGT does not report on any use of the group beyond it being the context of treatment. In this chapter the reader has been presented with one type of cognitive-behavioral group therapy in which use is explicitly made of the group. The various assessment, intervention, and generalization strategies employed in CBGT have been presented specifically as they could be uniquely applied in a group and which take advantage of the group context. Specific group interventions are identified. The cornerstone of using the group is maximum involvement of the members in the treatment process both as patient and as cotherapist. Because of the focus on interaction, this particular approach to CBGT has been referred to throughout as cognitive-behavioral interactive group therapy or CBIGT.

It is clear that intense involvement of members in the therapeutic process is costly in time. Much less can be covered in a given number of sessions. According to our clinical observations, failure to become involved often results in a highly didactic quality of therapy, high drop-out rates, and lower levels of learning. More research on these very issues is required before one can be firmly committed to this conclusion.

6.15.6 REFERENCES

Abraham, I. L., Neundorfer, M. M., & Currie, L. J. (1992). Effects of group interventions on cognition and depression in nursing home residents. *Nursing Research, 41*(4), 196–202.

Azocar, F., Miranda, J., & Dwyer, E. V. (1991). Treatment of depression in disadvantaged women. *Women and Therapy, 18*(3–4), 91–105.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84,* 191–215.

Beck, A. T. (1976). *Cognitive therapy and emotional disorders.* New York: International Universities Press.

Beck, A. T., & Emery, G. (1985). *Anxiety disorders and phobias.* New York: Basic Books.

Bernstein, D. A., & Borkovec, T. D. (1973). *Progressive relaxation training: A manual for the helping professions.* Champaign, IL: Research Press.

Bottomley, A., Hunton, S., Roberts, G., & Jones, L. (1996). A pilot study of cognitive behavioral therapy and social support group interventions with newly diagnosed cancer patients. *Journal of Psychosocial Oncology, 14*(4), 65–83.

Cardillo, J. E. (1994). Goal setting, follow-up, and goal monitoring. In T. J. Kiresuk, A. Smith, & J. E. Cardillo (Eds.), *Goal attainment scaling: Applications, theory, and measurement* (pp. 39–59). Hillsdale, NJ: Erlbaum.

Carrington, P. (1978). *Learning to meditate: Clinically standardized meditation (CSM). Course workbook.* Kendall Park, NJ: Pace Educational Systems.

Comas-Diaz, L., & Duncan, J. W. (1985). The cultural context: A factor in assertiveness training with mainland Puerto Rican women. *Psychology of Women Quarterly, 9,* 463–475.

Corey, M. S., & Corey, G. (1997). *Groups process and practice* (5th ed.). Pacific Grove, CA: Brooks/Cole.

D'Zurilla, T. J. (1986). *Problem-solving therapy: Social competence approach to clinical intervention.* New York: Springer.

D'Zurilla, T. J., & Nezu, A. M. (1988). Development and preliminary evaluation of the social problem-solving inventory. Paper presented at AABT, New York.

Ehlers, A., Stangier, U., & Gieler, U. (1995). Treatment of atopic dermatitis: A comparison of psychological and dermatological approaches to relapse prevention. *Journal of Consulting and Clinical Psychology, 63*(4), 624–635.

Ellis, A. (1973). *Humanistic psychotherapy.* New York: McGraw-Hill.

Enright, S. J. (1991). Group treatment for obsessive-compulsive disorder: An evaluation. *Behavioural Psychotherapy, 19*(2), 183–192.

Fals-Stewart, W., Marks, A. P., & Schafer, J. (1993). A comparison of behavioral group therapy and individual behavior therapy in treating obsessive-compulsive disorder. *The Journal of Nervous and Mental Disease, 181*(3), 189–193.

Flowers, J. V., & Schwartz, B. (1985). Behavioral group therapy with clients with homogeneous problems. In S. Ross & D. Upper (Eds.), *Handbook of behavioral group therapy.* New York: Plenum.

Forsyth, D. R. (1990). *Group dynamics* (2nd ed.). Pacific Grove, CA: Brooks/Cole.

Gambrill, E. D., & Richey, C. A. (1973). An assertion inventory for use in assessment and research. *Behavior Therapy, 6,* 550–561.

Garvin, C. (1987). *Contemporary group work.* Englewood Cliffs, NJ: Prentice-Hall.

Goldstein, A. P., Heller K., & Sechrest, L. B. (1966). *Psychotherapy and the psychology of behavior change.* New York: Wiley.

Graff, R. W., Whitehead, G. I., & LeCompte, M. (1986). Group treatment with divorced women using cognitive-behavioral and supportive-insight methods. *Journal of Counseling psychology, 33*(3), 276–281.

Hand, I., Lamontagne, Y., & Marks, I. M. (1974). Group exposure (flooding) *in vivo* for agoraphobics. *British Journal of Psychiatry, 124,* 588–602.

Heide, F. J., & Borkovec, T. D. (1983). Relaxation-induced anxiety: Paradoxical anxiety enhancement due to relaxation training. *Journal of Consulting and Clinical Psychology, 51,* 171–182.

Krone, K. P., Himle, J. A., & Nesse, R. M. (1991). A standardized behavioral group treatment program for obsessive-compulsive disorder: Preliminary outcomes. *Behaviour Research and Therapy, 29,* 627–631.

Kupych-Woloshyn, N., MacFarlane, J. G., & Shapiro, C. M. (1993). A group approach for the management of insomnia. *Journal of Psychosomatic Research, 37* (Suppl. 1), 39–44.

La Framboise, T., & Rowe, W. (1983). Skills training for bicultural competence: Rationale and application. *Journal of Counseling Psychology, 30,* 589–595.

Lee, N. F., & Rush, A. J. (1986). Cognitive-behavioral group therapy for bulimia. *International Journal of Eating Disorders, 5*(4), 599–615.

Lindsay, W. R., Gamsu, C. V., McLaughlin, E., Hood, E. M., & Espie, C. A. (1987). A controlled trial of treatments for generalized anxiety. *British Journal of Clinical Psychology, 26,* 3–15.

Lutgendorf, S. K., Antoni, M. H., Ironson, G., & Klimas, N. (1997). *Journal of Consulting and Clinical Psychology, 65*(1), 31–43.

Lyles, J. N., Burish, T. G., Korzely, M. G., & Oldham, R. K. (1982). Efficacy of relaxation training and guided imagery in reducing the aversiveness of cancer chemotherapy. *Journal of Consulting and Clinical Psychology, 50,* 509–524.

Magen, R., & Rose, S. D. (1998). Assessing parenting skills through role play: Development and reliability. *Research on Social Work Practice, 8.*

Manning, J. J., Hooke, G. R., Tannenbaum, D.A., & Blythe, T. H. (1994). Intensive cognitive behaviour group therapy for diagnostically heterogeneous groups of patients with psychiatric disorder. *Australian and New Zealand Journal of Psychiatry, 28*(4), 667–674.

Meichenbaum, D. (1977). *Cognitive-behavior modification.* New York: Plenum.

Miller, W. R., & Rollnick, S. (1991). *Motivational interviewing.* New York: Guilford.

O'Donnell, C. R., Lydgate, T., & Fo, W. S. O. (1979). The buddy system: review and follow-up. *Child Behavior Therapy, 1*(2), 161–170.

Onyett, S. R., & Turpin, G. (1988). Benzodiazepine withdrawal in primary care: A comparison of behavioural group training and individual sessions. *Behavioural Psychotherapy, 16*(4), 297–312.

Piacentini, J. (1993). Checklists and rating scales. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of child and adolescent assessment.* Boston: Allyn and Bacon.

Power, K. G., Simpson, R. J., Swanson, V., & Wallace, L. A. (1990). A controlled comparison of cognitive-behaviour therapy, diazepam, and placebo, alone and in combination, for the treatment of generalised anxiety disorder. *Journal of Anxiety Disorder, 4*(4), 267–292.

Rose, S. D. (1989). *Working with adults in groups: A multimethod approach.* San Francisco: Jossey-Bass.

Rose, S. D. (1997). *A training program for leaders of coping skill training groups. A book of exercises.* Madison, WI: School of Social Work, University of Wisconsin.

Rose, S. D. (1998a). *A survey of organizations serving mental health needs in groups in Madison, Wisconsin.* University of Wisconsin, Madison.

Rose, S. D. (1998b). *Group therapy with troubled youth.* Thousand Oaks, CA: Sage.

Rose, S. D., Tolman, R., & Tallant, S. (1985). Group process in cognitive-behavioral therapy. *The Behavior Therapist, 8,* 71–75.

Sarri, R., & Galinsky, M. (1985). A conceptual framework for group development. In M. Sundel, P. Glasser, R. Sarri, & R. Vinter (Eds.), *Individual change through small groups* (pp. 70–86). New York: Free Press.

Schinke, S. P., & Singer, B. R. (1994) Prevention of health care problems. In D. K. Grand (Ed.), *Cognitive and behavioral treatment* (pp. 285–298). Belmont, CA: Brooks/Cole.

Spence, S. H. (1991). Cognitive-behaviour therapy in the treatment of chronic, occupational pain of the upper limbs: A 2 year follow-up. *Behaviour Research and Therapy, 29*(5), 503–509.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10,* 349–367.

Stoyva, J. (1977) Why should muscular relaxation be useful? In I. J. Beatty & H. Legewie (Eds.), *Biofeedback and behavior.* New York: Plenum.

Subramanian, K. (1991). Structured groupwork for the management of chronic pain: An experimental investigation. *Research on Social Work Practice, 1,* 32–45.

Subramanian, K. (1994). Long-term follow-up of a structured group treatment for the management of chronic pain. *Research on Social Work Practice, 4,* 208–223.

Teri, L., & Lewinsohn, P. M. (1986). Individual and group treatment of unipolar depression: Comparison of treatment outcome and identification of predictors of successful treatment outcome. *Behavior Therapy, 17*(3), 215–228.

Tilliski, C. J. (1990). A meta-analysis of estimated effect sizes for group versus individual versus control treatments. *The International Journal of Group Psychotherapy, 40*(1), 215–224.

Toseland, R. W., & Siporin, M. (1986). When to recommend group treatment: A review of the clinical and the research literature. *International Journal of Group Psychotherapy, 36,* 171–206.

Tuckman, B. (1965). Developmental sequence in small groups. *Psychological Bulletin, 63,* 384–399.

Van Hoppen, B. V., Rasmussen, S. A., & Eisen, J. L. (1991, May). *A time limited behavioral group treatment for OCD.* Paper presented at the American Psychiatric Association Annual Meeting, New Orleans, LA.

Wolf, E. M., & Crowther, J. H. (1992). An evaluation of behavioral and cognitive-behavioral group interventions for the treatment of bulimia nervosa in women. *International Journal of Eating Disorders, 11*(1), 3–15.

Wolfe, J. L. (1987). Cognitive behavioral group therapy for women. In G. M. Brody (Ed.), *Women's therapy groups: Paradigms of feminist treatment* (pp. 163–173). New York: Springer.

Yalom, I. D. (1985). *The theory and practice of group psychotherapy* (3rd ed.). New York: Basic Books.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.16
Affective Disorders

ROBERT J. DeRUBEIS, PAULA R. YOUNG,
and KATHERINE K. DAHLSGAARD
University of Pennsylvania, Philadelphia, PA, USA

6.16.1 INTRODUCTION 340

6.16.2 PHENOMENOLOGY 340

- *6.16.2.1 Phenomenology of Depressive Episodes* 340
- *6.16.2.2 Phenomenology of Manic Episodes* 341

6.16.3 CLASSIFICATION 342

- *6.16.3.1 Unipolar–Bipolar Distinction* 342
- *6.16.3.2 Melancholic–Nonmelancholic Distinction* 342
- *6.16.3.3 Seasonal–Nonseasonal Distinction* 343

6.16.4 ETIOLOGICAL FACTORS 344

- *6.16.4.1 Epidemiology* 344
  - *6.16.4.1.1 Epidemiology of unipolar disorder* 344
  - *6.16.4.1.2 Epidemiology of bipolar disorder* 345
- *6.16.4.2 Genetics* 345
  - *6.16.4.2.1 Genetics of unipolar disorder* 345
  - *6.16.4.2.2 Genetics of bipolar disorder* 346
- *6.16.4.3 Life Events* 347
  - *6.16.4.3.1 Life events in unipolar disorder* 347
  - *6.16.4.3.2 Life events in bipolar disorder* 348
- *6.16.4.4 Effects of Early Environment on Unipolar Disorder* 348

6.16.5 COMORBIDITY 349

- *6.16.5.1 Comorbidity with Anxiety Disorders* 349
- *6.16.5.2 Comorbidity with Substance Use Disorders* 350

6.16.6 SUICIDE 351

6.16.7 THEORIES AND TREATMENTS OF UNIPOLAR DISORDER 352

- *6.16.7.1 Biological Theories and Treatments of Unipolar Disorder* 352
- *6.16.7.2 Cognitive Theories and Treatments of Unipolar Disorder* 354
- *6.16.7.3 Behavioral Theories and Treatments of Unipolar Disorder* 356
- *6.16.7.4 Psychodynamic Theories and Treatments of Unipolar Disorder* 357

6.16.8 THEORIES AND TREATMENTS OF BIPOLAR DISORDER 358

- *6.16.8.1 Biological Theories and Treatments of Bipolar Disorder* 358
- *6.16.8.2 Psychosocial Theories and Treatments of Bipolar Disorder* 359

6.16.9 SUMMARY AND FUTURE DIRECTIONS 360

6.16.10 REFERENCES 361

6.16.1 INTRODUCTION

The array of human problems which has come to be known as the affective disorders offers students of psychopathology an engrossing course of study. The modern conception of this nosologic category encompasses both depression and mania, disorders in which the most basic human drives and desires are altered, in many cases to such a degree that life itself is threatened. Hunger is extinguished, or in some cases excessive. Sleep is severely disrupted. Sexual appetite evaporates or goes out of control. And the desire to live, taken for granted by most people most of the time, can be replaced by the urge to die. Secondary drives and motives suffer as well. Achievement strivings, pride, and a sense of belonging in a social network become distant memories if they are remembered at all, or they are perceived out of proportion to reality.

Nearly all of psychology's philosophical, methodological, and scientific battles have taken place in the field of the affective disorders. Although many of the secrets of the affective disorders have yielded to scientific inquiry, and there is little controversy about the nature and treatment of mania, uncertainties regarding many of the most basic issues in depression remain. Is the distinction between a sad mood and major depression one of degree or of kind? Is depression essentially psychological or biological in nature? To what degree is it caused by genetic factors, and to what degree by environment? Of the environmental influences, is the key factor childhood upbringing, or is it a difficult environment encountered in adulthood? And finally, which are the more effective treatments for depression: biological or psychological ones? The answers to each of these questions are complex, but with the enormous amount of research conducted since the 1960s, the field is moving closer to an integrated understanding of depression and its treatment.

In the last four decades the field has also witnessed several new discoveries and the development of potent treatments for the affective disorders. With the exception of electroconvulsive therapy, none of the treatments used 40 years ago are among those considered effective today. Specific antidepressant medications, as well as targeted psychosocial interventions, have been added to the list of effective treatments. A newly described variant of affective disorder, seasonal affective disorder, has led to the development of light therapy, a treatment that would have been the stuff of science fiction 25 years ago. Advances in genetics have not yet informed our understanding of the biological processes involved in the affective disorders, but we now understand that genes play a substantial part in the development of depression and mania.

Note that brand names given in this chapter refer to the most commonly prescribed brands in the USA. Brand names other than those mentioned here are also used in the USA as well as internationally.

6.16.2 PHENOMENOLOGY

The phenomena that make up affective disorders form classic syndromes in that a wide array of distinctive experiences (symptoms) and behaviors (signs) are found to appear together to form coherent clusters. Though disturbed mood is usually a central feature in a person who is diagnosed with an affective disorder (sadness in depression; euphoria in mania), these symptoms are not required for either diagnosis. Instead, the person must possess a collection of phenomenologically disparate signs and symptoms. The system embedded in the fourth edition of the *Diagnostic and statistical manual of mental disorders* (*DSM-IV*; American Psychiatric Association [APA], 1994) will be used to describe the syndromes of major depression and mania.

6.16.2.1 Phenomenology of Depressive Episodes

Modern diagnostic systems recognize a distinction between short-term emotional upset and a major depressive episode, in part by requiring the presence of the depressive syndrome over a period of time. Two weeks has been chosen as a reasonable if arbitrary cutoff. Moreover, the depressive syndrome is not considered to be present if the signs and symptoms can be accounted for by a normal bereavement reaction, a medical condition, or a reaction to a medication or a drug of abuse. In this way, the depressive syndrome designation is reserved for cases in which the appearance of depressive signs and symptoms is neither normative nor readily explained by gross somatic disturbance.

The *DSM-IV* requires that, for a major depressive episode, either sadness or lack of interest/pleasure be present most of the day, each day. Sadness is usually a reported symptom, but its presence can be inferred from crying spells or from facial expression. Diminished interest in activities can be obtained from the patient's subjective report, or it can be inferred from reduced participation. Markedly reduced interest in sexual activities is common.

The vegetative symptoms, especially sleep and appetite disturbance, are seen by many as the essential indicators of depression. The classic forms of these disturbances entail sleep loss (insomnia) and a decreased appetite (hypophagia), but it is now widely recognized that the reverse, hypersomnia and hyperphagia, are commonly found in persons with the other symptoms of the depressive syndrome. Thus, these reversed vegetative signs are included in the most recent diagnostic systems and assessment tools. Persons whose affective episodes follow a seasonal pattern (see Section 6.16.3.3) are particularly likely to experience these reversed vegetative symptoms.

The motor symptoms of agitation and retardation are especially observed in the more severe forms of the syndrome. Agitation is manifest in pacing or fidgeting, and the experience that it is difficult to sit still. Psychomotor retardation is observed in slowed speech and movements, and is reported as the feeling of moving in slow motion. Depressed patients also report fatigue and reduced energy that cannot be attributed entirely to either sleep deprivation or diminished interest.

Cognitive symptoms are also observed in depression: excessive self-blame or guilt, along with the belief that one is inadequate or unworthy, are nearly always present. Such beliefs can become so unconnected to reality as to be delusional. Extreme pessimism about the future, a state referred to as hopelessness, is also common.

Depressed persons often report great difficulty with concentration and decision-making. Simple acts such as reading the newspaper, following a conversation, or watching television, can become virtually impossible for the severely depressed person. Problems with concentration and decision-making result in part from the distraction caused by persistent worries, but they also reflect cognitive inefficiency, possibly associated with hormonal disturbance (Rubinow, Post, & Savard, 1984).

Depressive disorders also include milder forms, which are nonetheless bothersome enough to warrant attention and treatment, especially when they are experienced over long periods of time. Dysthymic disorder refers to long-lasting episodes of mood disturbance that are not severe enough to qualify as a major depressive episode. For a person to be deemed dysthymic, mild to moderate levels of depressive symptoms need to be experienced most of the day, more days than not, over a two year period (APA, 1994).

6.16.2.2 Phenomenology of Manic Episodes

Modern conceptions of manic syndromes have their roots in the descriptions of manic depressive illness articulated by Kraepelin (1921) in the early part of the twentieth century. These descriptions continue to be some of the most illustrative accounts of the phenomenology of mania.

The central feature of a manic episode is a distinct change in mood. This change can be reflected in either an elevation of mood (elation, euphoria), or in irritability. Elevated mood may be inferred by behaviors such as joking, laughing and singing. Euphoria is not necessarily perceived as problematic by the patient. *DSM-IV* requires that the change in mood lasts for at least one week, or less if the impairment is severe enough to warrant hospitalization. Dysphoria (sad mood) may occur within a manic episode, and there may be frequent changes between elation and depression. Persons with this pattern have been labeled dysphoric manics, or as those having a mixed episode (Post et al., 1989).

Inflated self-esteem and grandiosity are characteristics of mania. The manic person often has an exaggerated sense of achievements and abilities. Grandiosity may also be observed in the manic person's behavior, such as embarking impulsively on ambitious but unrealistic projects, or giving advice on matters about which he or she has little knowledge. The grandiosity may also manifest in delusions, such as a belief that one is a deity or a famous public figure.

Increases in energy are observed in mania. In earlier stages and milder forms of mania, this may manifest in increased goal-directed activities in social or occupational realms, some of which may be successful. In later stages, manic persons may be unable to sit still, and they may initiate many tasks that they are unable to complete because of severe distractibility. The increase in activities may include imprudent involvement in pursuits of pleasure which have high potential for painful or dangerous consequences. Increased sexual drive can lead to impulsive and reckless sexual behavior, such as infidelity or sexual encounters with strangers. Buying sprees that involve expensive, unneeded items, gambling binges, or foolish business investments create financial problems, which can exacerbate subsequent depressive episodes.

Decreased need for sleep is common. Fogarty, Russell, Newman, and Bland (1994) reported reduced sleep as the most frequent symptom in those with a history of mania. The sleep disturbance seen in mania differs from that observed in depression; in mania, the disturbance is related to a marked decrease in perceived need for sleep. Some manic individuals go for days without any sleep, yet they report no fatigue. By contrast, in depression the decreased sleep (insomnia) is unwanted, and the restless sleep that does occur is not experienced as restorative.

The disturbed thought processes of mania are distinct and can be readily observed in the speech of the manic patient. The rate of speech is rapid; interrupting the manic patient can be difficult. Speech is often colorful, and characterized by the use of rhymes, puns, and word associations based on sounds rather than meaning (clang associations). Persons in a manic episode may report that thoughts are racing through their mind. The thought disorder characteristic of mania, which is reflected in the content of speech, is known as "flight of ideas." Speech may be tangential, or there may be frequent, loosely associated shifts from one topic to another. Individuals in a manic episode are characterized by their inability to screen out irrelevant stimuli, both internal (thoughts) and external (sights, sounds). This distractibility can be observed in the quickly shifting speech content and activities of the manic individual. The symptom known as "pressured speech" refers to the forced quality of utterances. It is experienced by the manic person as an inability to slow down or inhibit speech.

Psychotic symptoms such as delusions or hallucinations are present in more than half of manic patients (Silverstone & Hunt, 1992). Delusions in mania tend to be grandiose, but paranoid delusions are not uncommon. Hallucinations are most frequently auditory, and are usually congruent with prevailing mood and themes of grandiosity.

6.16.3 CLASSIFICATION

The heterogeneity of the affective disorders has led to numerous attempts to discern meaningful subcategories. The usual nosologic aims of improving the search for causes and treatments have driven these efforts. The severity of depressive symptoms varies considerably, from mild dysphoria to psychotic forms. Moreover, the observed variation appears to be continuous. Meaningful distinctions have been made, however, with implications for etiology and treatment.

6.16.3.1 Unipolar–Bipolar Distinction

Of the many attempts to divide the affective disorders into meaningful subclassifications, no distinction has proven more fruitful than that between unipolar and bipolar forms. Through the middle part of the twentieth century, Kraepelin's (1921) diagnostic category of "manic-depressive insanity" included all patients who experienced a manic or a depressive episode, whether or not they had experienced both poles of affective disturbance. The use of the term "manic-depressive" is today more a source of confusion, since it implies that one suffering from it would have experienced (or could expect to experience) both mania and depression.

In more modern classification systems this confusion has been eliminated. Separate categories have been recognized for (i) patients who have experienced only depressive episodes—unipolar disorder; and (ii) patients who have experienced both manic and depressive episodes—bipolar disorder. There was some concern that a third category, for patients who experienced only manic episodes, would be needed, but research on the treatments, genetics and clinical courses of these patients has shown that they are best considered bipolar, rather than unipolar manic, despite the absence of a depressive episode. The organization of much of this chapter recognizes the importance of the unipolar–bipolar distinction. Differences between unipolar and bipolar disorders have been found in regard to (i) epidemiologic factors, including gender distributions; (ii) genetics; (iii) psychosocial causes; (iv) course of illness; (v) biologic variables; and (vi) treatments.

Perris (1992) reviewed findings from family studies, twin studies, and adoption studies that together support the claim that unipolar and bipolar forms of affective disorder are, at least to some degree, genetically distinct. In the family studies, unipolar (or bipolar) probands were identified, and their first-degree relatives were assessed for the presence of unipolar or bipolar disorder. The mean morbidity risk for bipolar disorder across seven family studies of unipolar probands was less than 2%, close to the population risk, whereas the risk for unipolar disorder was considerably higher, at 15%. These figures can be contrasted with those obtained in 14 studies of bipolar probands. The mean risk was 9% for bipolar disorder, which is much higher than the population rate. However, the risk for unipolar disorder was rather high also, at 13%. Although some of the unipolar relatives in this last group may have been misdiagnosed—early episodes in bipolar disorder tend to be depressions—clearly the relatives of bipolar patients are at increased risk not only for bipolar disorder, but for the unipolar form of affective pathology as well.

6.16.3.2 Melancholic–Nonmelancholic Distinction

The most enduring subtype within the depressive spectrum has gone by several names, but the use of "melancholic" to refer to the classic presentation of depression has endured through most of the twentieth century. The concept, which has obtained inconsistent support from research findings, is that melancholic (or endogenous) depressions are caused more by biological factors, and should respond better to biological treatments, than their opposite—the so-called exogenous (or nonmelancholic, reactive, or neurotic) depressions.

There is indeed a melancholic symptom pattern that coheres fairly well (Haslam & Beck, 1994), although the presence of endogenous symptoms is also indicative of, and may simply reflect, a more severe form of depression. Melancholic features include: early morning awakening, weight loss, loss of pleasure, and loss of reactivity to usually pleasurable stimuli. By contrast, a nonmelancholic pattern may involve no sleep or appetite disturbance, or a reversed pattern, characterized by excessive sleeping and weight gain. In addition, such individuals often experience a lifting of mood in response to rewarding or pleasurable events in their environment.

Attempts to validate the melancholic–nonmelancholic distinction have produced mixed results. The assumption that melancholic (endogenous) depressions "arise from within," and are therefore not likely to have been preceded by a negative life event, has not generally been confirmed with data (Grove & Andreasen, 1992). Most researchers have found that plausible precipitating events are common in both endogenous and exogenous depressions, although the association may be stronger in melancholia (Brown, Harris, & Hepworth, 1994).

However, the endogenous–exogenous distinction does predict response to biological treatments. Both electroconvulsive therapy and antidepressant medications are more effective for endogenous episodes than for episodes with a reactive symptom pattern (Grove & Andreasen, 1992). The "biological vs. non-biological" connotation of this distinction would receive even more support if the reverse pattern were found for psychosocial interventions. There has been little investigation of this question. Not one large-sample prospective comparison of biological and psychological treatments has been conducted, but results thus far have indicated that psychosocial treatments such as cognitive therapy may not be as effective with melancholic depressions as they are with nonmelancholic depressions (Thase et al., 1996). Finally, biological abnormalities associated with depression, such as the dexamethasone suppression test, are found more often in melancholic than in nonmelancholic depressed patients (APA Task Force on Laboratory Tests in Psychiatry, 1987).

6.16.3.3 Seasonal–Nonseasonal Distinction

Since the 1970s, increased attention has been given to a subtype of affective disturbance that follows an annual pattern—seasonal affective disorder (SAD). An appreciation of this pattern has led to important advances in its treatment, and to an increased understanding of the biological mechanisms that play a role in all depressions. Moreover, the fact that some depressions follow a seasonal pattern has fueled an interest in the evolutionary function that this depressive complex might have served for our ancestors (Whybrow & Bahr, 1988).

The typical annual course for SAD is an onset in mid to late autumn, with either remission of symptoms in early spring, or a period of hypomania or mania followed by a return to normal mood. The symptom pattern found in SAD differs from that found in nonseasonal depressions. The "reversed vegetative symptoms" of hypersomnia (increased sleep) and hyperphagia (increased appetite) are common. In this regard, SAD is similar to another subtype of depression—atypical depression (Lam & Stewart, 1996); however, it differs from atypical depression in important ways. Craving for carbohydrates, especially sweets, is pronounced in SAD, but not in atypical depression. Interpersonal sensitivity and avoidance of rejection are not common in SAD, but they are defining features of atypical depression (Tam et al., 1997). Lethargy and disinterest are especially common in SAD. The signs of SAD are thus akin to those observed in a hibernating animal (Whybrow & Bahr, 1988).

SAD is found predominantly in the higher latitudes, such as in Scandinavian countries or the southern parts of South America. It is nonexistent near the equator, because light and climate conditions vary seasonally in the higher latitudes, but not in the lower. Separate influences of light and temperature have been difficult to distinguish (Molin, Mellerup, Bolwig, Scheike, & Dam, 1996), because they tend to covary. Although it is possible that temperature plays an independent role, it is believed that some aspect of exposure to light is critical for the development of SAD. Whether the critical variable is length of day (i.e., time from light onset to light offset), or total amount of light, has been difficult to determine.

Treatment for SAD involves therapeutic exposure to high intensity light (see Tam, Lam, & Levitt, 1995). Although most research has been conducted on light presented before 8 a.m., the timing of the light treatment may not matter for most patients (Meesters et al., 1993). The spectral properties of the light seem also not to be critical—white, red, green, and blue light all appear to be equally effective. Intensity of light is important, however—2500 lux for one to two hours per day is a well-documented regimen, whereas less intense light (e.g., 500 lux) is not as effective (Tam et al., 1995). Briefer exposure (e.g., 30 minutes) with a higher intensity light (10 000 lux) appears to be as effective as the standard treatment, lending credence to the view that it is the total amount of light that falls on the retina which is critical in the development and treatment of SAD (Terman et al., 1990).

What physiological mechanisms explain how the absence of light can lead to depressive symptoms? How can treatment that involves only the presentation of extra light alleviate SAD? The answer lies in the fact that humans possess a tract of nerves that runs directly from the retina to the hypothalamus, without passing through the visual cortex. Presumably, this nonvisual tract plays a role in the regulation of, for example, mood, appetite, and sleep—functions controlled by the hypothalamus.

6.16.4 ETIOLOGICAL FACTORS

6.16.4.1 Epidemiology

6.16.4.1.1 Epidemiology of unipolar disorder

Depressive disorders occur on a continuum from mild (dysthymic disorder) to very severe (psychotic, melancholic forms), with no clear thresholds marking the transition points. This has made it difficult to study the prevalence of depression across time and across settings, as well as across cultures. Since the 1970s, epidemiological studies have benefited greatly from the development of standardized assessment methods, beginning with the criteria of Feighner et al. (1972) in the USA, and the Present State Examination (Wing, Cooper, & Sartorius, 1974) in the UK. In a review of studies that were conducted prior to the development of these instruments, Boyd and Weissman (1981) reported estimates of lifetime prevalence of unipolar depression that ranged from 3% to 18%. Most researchers now cite the newer Epidemiological Catchment Area study (ECA; Weissman et al., 1988), a large, five-site community survey that employed the Diagnostic Interview Schedule (Robins, Helzer, Ratcliff, & Seyfried, 1982) to determine *DSM-III* (APA, 1980) diagnoses. In the ECA, lifetime prevalence of unipolar (major) depression was estimated at 4.4%. A further 3.2% were found to suffer from the milder, chronic version of depression, dysthymia (Gwirtsman, 1994). A substantially higher estimate of major depression (8%) was obtained in the recent, large Edmonton study (Spaner, Bland, & Newman, 1994), with methods similar to those used in the ECA study. It appears, then, that findings from the earlier studies overestimated the prevalence of unipolar depression. At this point, it would be safe to say that lifetime prevalence of major depression in western societies is in the range of 4–8%. However, this may be changing rapidly.

Indeed, the most disturbing findings to emerge from depression research since the 1970s concern indications that depression rates have been increasing dramatically in industrialized cultures during the twentieth century (Klerman & Weissman, 1989). Klerman and Weissman estimate that the birth cohorts born after World War II incur more risk for depression than cohorts born early in the twentieth century, perhaps by a factor of five or more. The same data suggest that the mean age of onset has been decreasing sharply in the same time period. Whereas the older studies yielded mean age-of-onset figures in the 50s or 60s, recent estimates suggest a mean age of onset in modern industrialized nations in the late 20s. The most dramatic change in age of onset also appears to have occurred in the cohorts born around the time of World War II and later. It is impossible to rule out conclusively all of the potential artifacts in such findings, especially since most of the data rely on respondents' memories and their willingness to report potentially embarrassing feelings and behaviors. However, the phenomenon survives careful scrutiny. For example, the same epidemiologic methods have been applied in nonindustrialized cultures, and no increases in depression rates over time have been observed. Even though Klerman and Weissman's estimate of the increase in rates is almost certainly an overestimate (Fombonne, 1994), there is little doubt that depression is and will continue to be a more common disorder than it was in our great-grandparents' day, when it was a rare disorder.

Despite between-study and between-era differences in rates of depression, a consistent finding is that women experience unipolar depression at about twice the rate of men (Weissman & Klerman, 1985). Several explanations have been advanced to account for the gender difference; it appears that it cannot be explained by a reporting artifact, or by differential willingness to consult mental health professionals. Women have been found to engage in more ruminative, self-reflective thinking when they experience a low mood, whereas men engage more often in distracting activities (Nolen-Hoeksema, 1987). Hormonal differences and differences in the traditional roles adopted by men and women appear also to play a role in the gender discrepancy (Nolen-Hoeksema, 1987; Weissman & Klerman, 1977).

There has been a consistent and surprising lack of association between socioeconomic status and depression in epidemiological investigations (Weissman et al., 1988). Likewise, within Western societies, racial and ethic differences in rates of depression have not been observed. Marital status is associated with depression, however. Married persons who have never divorced have the lowest rates, followed by those who have never married. The highest rates tend to be found for divorced persons, especially those recently divorced. Although the nature of the causal relationships involved is unclear, there is good evidence that the presence of a supportive relationship, such as can be found in a good, stable marriage, acts as a buffer against environmental stressors (Brown & Harris, 1978). Alternatively, the interpersonal adjustment of depressed people is found to be impaired (Barnett & Gotlib, 1988). Thus, a portion of the relation of marital status to depression is undoubtedly due to the damaging effects of depression on intimate relationships, including marriages.

6.16.4.1.2 Epidemiology of bipolar disorder

Bipolar illness is much less common than unipolar depression. In the ECA study, the lifetime prevalence was of bipolar disorder was 1.2% (Weissman et al., 1988). Whereas depression is more common among females than males, gender does not appear to be a significant risk factor for bipolar illness. Results of the ECA study indicate that the prevalence of bipolar illness is nearly equivalent in the two genders, with a female:male ratio of 1.2:1 (Weissman et al., 1988).

The average age of onset of bipolar disorder is 20. About one in four bipolar patients experience their first affective episode prior to age 20 (Goodwin & Jamison, 1990). ECA data (Weissman et al., 1988) indicate that incidence of bipolar affective disorder was highest in the sample of 18–44 year olds (1.4%), and that there were virtually no cases reported in the group aged 65 and over (0.1% incidence).

It is estimated that in just over half of bipolar patients, the illness commences with a manic episode. On average, manic episodes develop within a few days, whereas depressive episodes develop less abruptly, usually within a few weeks. Manic episodes are also shorter in duration than depressive episodes (Keller et al., 1986; Goodwin & Jamison, 1990). Gender differences in the course of the illness have also been examined. Manic episodes were predominant in males with bipolar illness, whereas depressive episodes were more frequent in females (Coryell et al., 1989).

Prior to the advent of effective drug treatments, the average natural duration of a manic episode was about 17 weeks (Wertham, 1928). Since somatic treatments have been available, the average duration of an episode has decreased to 10 weeks (Coryell & Winokur, 1992). Bipolar disorder is a chronic, recurrent illness. Most individuals with this disorder will experience more than one episode during their lifetime (Goodwin & Jamison, 1990). There is evidence that the time from the onset of one episode to the onset of the next (cycle length) decreases for each subsequent episode (Goodwin & Jamison, 1990), and that each affective episode increases the probability of future relapse (Keller, Shapiro, Lavori, & Wolfe, 1982).

It is estimated that 13–20% of bipolar patients are rapid cyclers. Rapid cycling describes a subgroup of bipolar patients who experience more than four affective episodes in a year (Dunner & Fieve, 1974). Females are more likely than males to exhibit rapid cycling (Goodwin & Jamison, 1990). Although rapid cycling is present from the onset of illness in about 20% of cases, the majority (80%) develop rapid cycling later in the course of the illness. The fact that rapid cycling generally develops later in the course of the illness may reflect an underlying pathophysiological mechanism, such as progressive kindling or sensitization (Post, Rubinow, & Ballenger, 1986). Kindling refers to the process whereby alterations in neurotransmitter systems, such as occur in depression or mania, transform the brain so as to make it more vulnerable to dysregulation in the future.

6.16.4.2 Genetics

6.16.4.2.1 Genetics of unipolar disorder

Progress in the behavior genetics of depression has lagged behind that of similar work in schizophrenia. However, there have been enough replications of well-conducted family studies and twin studies to warrant the conclusion that risk for unipolar depression is, at least in part, heritable. From family studies it is estimated that the risk to first-degree relatives of depressive probands is 11–18%, compared to 0.7–7% for relatives of normal controls (Moldin, Reich, & Rice, 1991). The higher figures within each of these ranges more likely correspond to the thresholds employed in modern estimates of population risk cited in the earlier section on epidemiology in this chapter. The main advantage of family studies over other behavior genetic investigations is their large sample sizes (usually in the hundreds). The drawback is that, although findings from family studies can reveal general familial risk, they cannot distinguish genetic from environmental contributions.

Findings from twin studies can help to disentangle genetic from environmental factors. There have been five twin studies completed in which the concordance rates in monozygotic (MZ) twins have been compared to the concordance rates in dizygotic (DZ) twins. The greater the ratio of MZ concordance to DZ concordance, the stronger the evidence for a genetic contribution to depression. In each of the studies (Allen et al., 1974, Kallman, 1953, Slater, 1953 [pooled together in Allen, 1976]; Bertelsen et al., 1977; Torgersen, 1986; all cited in Nurnberger & Gershon, 1992), the MZ concordance was higher than the DZ rate. The MZ:DZ ratios ranged between 2.3:1 and 3.4:1, suggesting a substantial genetic contribution to unipolar depression.

In behavior genetic work, adoption studies are critical because they can help rule out confounds that remain in even the best family and twin studies. Only four adoption studies have been conducted, however, in which unipolar depression was a focus of investigation. Although three of the four studies (Cadoret, 1978; Mendlewicz & Rainer, 1977; Wender et al., 1986) yielded findings consistent with a genetic hypothesis, very small sample sizes (Cadoret) or very low rates of unipolar illness in the sample (Wender et al.) reduce confidence in two of them. Furthermore, Mendlewicz and Rainer did not distinguish unipolar from bipolar adoptees in their study, although they did ascertain the polarity of the biological and adoptive relatives. Finally, von Knorring, Cloninger, Bohman, and Sigvardsson (1983) found no differences in the rates of unipolar depression between biological and adoptive relatives of unipolar adoptees. Surprisingly, these null findings extended to comparisons between the relatives of unipolar adoptees and the relatives of control adoptees, despite a rather robust sample size.

To summarize, the evidence of a genetic contribution to unipolar depression has been inconsistent, and weaker than the evidence for several other mental disorders, including bipolar disorder. McGuffin and Katz (1989) point out that the inconsistencies in findings across genetic investigations of unipolar depression may arise from differences in the definition of depression used in the various investigations. They cite evidence that as one moves from the more severe forms of affective disorder (including bipolar disorder) to the moderate forms, and finally to the milder, or neurotic forms, genetic contribution diminishes, whereas the contributions of both familial (common to all members of a family) and non-shared (unique to the individual) environments increase.

In part due to the difficulty of ascertaining cases of unipolar depression, and because the environmental contribution to risk is substantial, efforts to identify the mode of genetic transmission for unipolar depression, and to identify the specific genes involved, have not been successful (Moldin et al., 1991). It appears that for the time being, biologically oriented researchers will need to focus on the biological markers and treatments for depression, knowing that genes play some role. An understanding of the ways in which genes play their role awaits refinements in the ascertainment of unipolar depression and its subtypes.

6.16.4.2.2 Genetics of bipolar disorder

The primary etiologic understanding of bipolar disorder is that it is transmitted genetically. Belief in the heritability of bipolar disorder is buttressed by findings from family, twin, and adoption studies. More recently, studies of genetic linkage and genetic association have been conducted. These are attempts to identify specific genes that play a part in the pathogenesis of bipolar illness.

Studies of families have consistently shown that the prevalence of affective disorders (both unipolar and bipolar) are higher in the families of patients with bipolar disorder than in families of nonpsychiatric controls. When relative risk factors are calculated (ratio of the risk in relatives of probands to risk in relatives of controls), family studies of bipolar disorder produce relative risk values upwards of 15 for bipolar disorder, and between two and five for unipolar illness (Alda, 1997). Findings from these studies have also suggested that the various forms of affective illness are related in a hierarchical way—relatives of bipolar probands have increased risk for both bipolar and unipolar disorder, whereas relatives of unipolar probands tend only to have increased risk for unipolar illness (Perris, 1992).

Studies of twins have also provided substantial evidence for the heritability of bipolar illness. Bertelsen, Harvald, and Hauge (1977) observed an 80% concordance for bipolar disorder in MZ twins. Overall concordance rates for bipolar disorder are estimated at 57% for MZ twins, as compared to 14% for DZ twins (Alda, 1997). Moreover, the concordance rates of MZ twins reared apart do not differ from those reared together (McGuffin & Katz, 1989).

Results of adoption studies are consistent with family study data (McGuffin & Katz, 1989) Mendlewicz and Rainer (1977) have conducted the largest adoption study of bipolar probands. For biological relatives of bipolar patients who had been adopted away from their biological parents, the risk was similar to that of relatives of bipolars who were not adopted (approximately 25%). Adoptive relatives did not show increased risk. Therefore, genetic factors were found to contribute to risk, but family environment was not.

The family, twin, and adoption studies indicate that bipolar disorder aggregates in families and is heritable. However, since the concordance rates in MZ twins do not approach 100%, genetics alone is not a sufficient explanation for the occurrence of bipolar disorder. There must be non-genetic causes as well.

Recently, genetic research has focused on the identification of the means of transmission through studies of linkage and association. Linkage refers to the identification of a chromosomal region containing an "illness" gene, whereas association refers to the identification of particular characteristics more frequent in patients than controls. Initial positive linkage results were found for the X chromosome as well as chromosome 11, but these have not been replicated. More promising results have been found to implicate chromosomes 18 and 21 (Alda, 1997).

6.16.4.3 Life Events

6.16.4.3.1 Life events in unipolar disorder

Since the 1960s, a voluminous literature has indicated a relationship between stressful life events and increased risk for unipolar depression. Paykel (1994) reviewed 28 studies in which more negative life events were reported by depressed people prior to episode onset than were reported in a similar period by people who did not become depressed. Similar results have also been obtained using community samples (Brown, Ahmed, Gary, & Milburn, 1995; Brown, Adler, & Bifulco, 1988), and samples of depressed children (Goodyer, Herbert, Tamplin, Secher, & Pearson, 1997). The association between stressful life events and depression appears to be stronger for nonmelancholic depressions; evidence for the link between melancholic depression and precipitating life events remains less well documented (Brown et al., 1994).

Despite the research directed into this area, the strength of the causative link between stressful life events and unipolar depression remains difficult to gauge. Depression, be it melancholic or nonmelancholic, is a heterogeneous illness with many causative, perpetuating, and modifying factors. Most individuals experience stressful life events and do not develop depression. Rather, life stress interacts, or is additive with, other predisposing factors to induce depression.

Researchers have investigated what types of events may be especially "meaningful" to individual persons and as such may particularly dispose them to development of depression. Beck and colleague's Sociotropy-Autonomy Scale (SAS; Beck, Epstein, & Harrison, 1983) distinguishes sociotropic people, those for whom affiliation with others is especially important, from autonomous people, those for whom issues of achievement and failure are paramount. Hammen, Ellicott, Gitlin, and Jamison (1989) found that patients who experienced stress consistent with their SAS type reported significantly more depressive symptoms than those who did not, despite no overall difference in the amount of stress experienced.

There has also been some study of the possibility that different types of events may differentially dispose people to depressive versus anxiety responses. In a community sample of 400 adult women, Brown (1993) found that loss provoked the onset of depression, danger provoked anxiety, and a combination of the two types of stressors provoked comorbid anxiety and depressive disorders.

Other researchers have examined how the association of stressful life events, maladaptive coping responses, and depression may be interrelated. Bifulco and Brown (1996) found that, in a person's reaction to severe events, the negative cognitive coping responses of inferred denial, self-blame, and pessimism were associated with increased depression risk, whereas a positive cognitive coping response, downplaying, was not.

Depressed persons may also create more negative life events than persons not prone to depression. Cui and Vaillant (1997), in a prospective study, found that depressed men experienced a greater amount of self-induced (dependent) negative life events after their first episode. The authors concluded that depression can generate these dependent negative events, and that this in turn may contribute to the chronic nature of the disorder. Similarly, Hammen (1991) found that women with unipolar depression, but not bipolar disorder, experienced more stressful life events and more dependent events than women with chronic medical illness, or with no physical or psychological impairment. She theorized that these women tended to generate stressful conditions for themselves, particularly in the interpersonal domain, which further contributed to their ongoing cycle of depressive symptoms and stress.

Research into the interaction of genetic factors and life events has also indicated that depressed persons play a role in bringing about their own negative stressors. Using a large sample of female twins, Kendler and Karkowski-Shuman (1997) found that genetic liability to depression was associated with a greater risk for assault, serious marital problems, divorce or breakup, job loss, serious illness, or financial problems. The authors suggested that genetic liability for depression is associated with greater life stress because genetic factors increase the likelihood that these individuals will place themselves in high risk environments.

Finally, some research has indicated the importance of positive events in the recovery from depression. Brown et al. (1988) and Brown (1993) found that, in women with chronic depression, life events that promise some hope for a better future, termed "fresh start events," were associated with recovery or improvement, even if these events were threatening.

6.16.4.3.2 Life events in bipolar disorder

The role of psychosocial stressors is not as well understood in the onset and course of bipolar disorder as it is in unipolar disorder. This is largely because bipolar disorder is more heavily determined by biological factors, in particular genetics. Even though bipolar illness is considered largely biological in nature, the impact of life events and psychosocial stressors have been shown to influence recovery and relapse from affective episodes in bipolar patients (Johnson & Roberts, 1995).

Bipolar patients have been found to have a higher number of severe stressors in the six months prior to the onset of a manic episode, compared to nonpsychiatric controls (Kennedy, Thompson, Stancer, Roy, & Persad, 1983), but no more so than unipolar depressive or schizophrenic patients (Johnson & Roberts, 1995).

Negative life events also affect relapse in patients with bipolar disorder. Ellicott, Hammen, Gitlin, Brown, and Jamison (1990) found that patients who had the highest total life event scores had a risk of relapse 4.5 times that of patients who did not experience stress. Similarly, Hunt, Bruce-Jones, and Silverstone (1992) found the rate of severe negative life events was significantly higher in the three months before relapse than in periods of time that did not precede relapse. Rates were especially elevated in the month just before relapse, such that severe negative events were four times more likely in the month prior to relapse than during other months. Other researchers have found that patients with bipolar disorder are more susceptible to the effects of severe stressors, such as a hurricane, than persons without bipolar disorder (Aronson & Shukla, 1987; Hammen & Gitlin, 1997).

Negative life events may also delay time to recovery for patients who are experiencing a manic episode. Using a prospective design, Johnson and Miller (1997) found that those who experienced a major stressor during a manic episode took three times as long to achieve a full recovery than those who did not experience a major stressor.

A hypothesized mechanism by which stressful life events might precipitate relapse is through alteration of circadian rhythms. Models employing this hypothesis have posited that stressful life events disrupt normal social–environmental cues and physiological rhythms, including sleep and neuroendocrine functioning. Such disturbances in physiological processes may precipitate manic and depressive episodes in vulnerable individuals (Johnson & Miller, 1997).

6.16.4.4 Effects of Early Environment on Unipolar Disorder

The onset of depressive episodes is associated with environmental circumstances that can serve as proximal causes, as described in Section 6.16.4.3.1. The effect of such life events could be to expose a genetic vulnerability, but they may also reveal a vulnerability that is in large part created by events in childhood, long before the emergence of depressive symptoms. Early empirical efforts to isolate predisposing events in childhood focused on parental loss, especially the child's bereavement experience, as a risk factor for later depression. More recently, the childhood home environment of depressed adults has been studied, primarily through their recollection of their parents' attitudes and behaviors.

With the availability of dozens of studies that have examined the association of parental loss to later depression, it has become possible to separate the effects of different types of loss, such as parental death, separation, or extended absences (Parker, 1992). When examined carefully, parental death during childhood, long believed to predispose an individual to adult depression, has not consistently been found to be related to the disorder. Rather, parental separations, especially those that reflect or result in family discord, are the types of "loss" that are most clearly and consistently associated with later depression. This analysis of the effects of parental loss focuses on the nature of the parenting environment, rather than the experience of bereavement in childhood *per se* (Brier, Kelsoe, Kirwin, Beller, & Wolkowitz, 1988;

Harris, Brown, & Bifulco, 1986; Tennant, 1988). Researchers have thus turned their attention to the effects of parenting, with or without parental loss, as they may predispose a child to depression in adulthood.

The most frequently studied parenting styles in this domain concern the two dimensions of (i) parental care (ranging from caring to unaffectionate style), and (ii) parental protection (from appropriate encouragement of autonomy to overprotective style). Lack of care is hypothesized to predispose the child to lasting self-esteem problems. Overprotection is thought to deprive the child of education in self-reliance and problem-solving behaviors, which may result in a sense of helplessness and hopelessness when stressors are encountered in later adult life (Perris, 1988).

Several studies of parental style have used the Parental Bonding Instrument (PBI; Parker, Tupling, & Brown, 1979), a retrospective measure of parents' behavior during the respondent's childhood. High scores on two dimensions of the PBI, low care and overprotection, have been associated with higher depression levels in clinical (Hickie et al., 1990) and community (Parker, Hadzi-Pavlovic, Greenwald, & Weissman, 1995) samples. This combination of styles has been termed "affectionless control."

These findings appear to be specific to nonmelancholic depression. Research has revealed no association between low care or overprotection and adult melancholic depression (Parker & Hadzi-Pavlovic, 1992; Parker, Kiloh, & Hayward, 1987). The association of parental affectionless control and later adult nonmelancholic depression (but not melancholic depression) has recently been replicated, using a non-Western sample of depressives and controls (Sato, Sakado, Uehara, Nishioka, & Kasahara, 1997).

Research shows that the repercussions of parents' use of affectionless control are additive. That is, if both parents exhibit the affectionless control style, the risk for later depression is greater than if only one parent does so; the risk is decreased if one parent exhibits optimal (high care–low control) parenting style (Parker & Hadzi-Pavlovic; 1992).

The research findings cited thus far do not prove that a parenting style of affectionless control is a cause of depression. Some noncausal explanations, however, have been ruled out. Depressed persons' reports of parental low care and overprotection do not appear to be the consequence of (i) response bias in the form of a blaming personality style (Hickie et al., 1990), (ii) retrospective bias resulting from a depressed mood state (Gerlsma, Das, & Emmelkamp, 1993), (iii) a propensity to help-seeking behavior (Parker et al., 1995).

However, other noncausal relationships have not been ruled out. Genetic factors common to parent and child could both predispose the parent to a style of affectionless control, and the child to the development of depression. Alternatively, premorbid child characteristics may themselves elicit pathogenic parenting. In this scenario, the child's behavior is both an influence on parental behavior and a premorbid (temperamental) characteristic that is associated with later depression.

Recent findings from a study of adult female twins (Neale et al., 1994) are consistent with a model in which poor parental style causes depression. Fathers' influence on their daughters' depression status appeared to be due to their parental style, as well as to a genetic effect. Mothers' influence appeared to be due only to their parenting style, and not genetics. Further use of behavior genetic methods (adoption studies, twin studies) should help to clarify the role of parental style in the development of depression in later life.

6.16.5 COMORBIDITY

6.16.5.1 Comorbidity with Anxiety Disorders

There is substantial symptom overlap in depression and anxiety, including negative affect, restless sleep, decreased energy, irritability, worry, indecisiveness, and negative self-evaluation (Alloy, Kelly, Mineka, & Clements, 1990). Both anxious and depressed individuals attach unrealistically high probabilities to negative outcomes (Beck, 1967, 1976; Beck & Emery, 1985). The similarity in symptomatology is exemplified in the assessment instruments used to measure the constructs of depression and anxiety. For example, there is a 70% item overlap between the Hamilton Rating Scale for Depression (Hamilton, 1967) and the Hamilton Rating Scale for Anxiety (Hamilton, 1959; DiNardo & Barlow, 1990).

Prior to the publication of the *DSM-III-R* (APA, 1987), comorbid anxiety disorders were excluded in the presence of affective disorders. It was presumed that affective disorders held a higher position in the classification system. This hierarchical classification was modified in the *DSM III-R* and *DSM-IV* to allow for additional diagnoses of anxiety disorder (i.e., panic disorder, agoraphobia, obsessive-compulsive disorder, generalized anxiety disorder) in the presence of a major depression.

When anxiety–depression comorbidity is recognized, as it is in the *DSM-IV*, persons with a major depressive disorder are found to be more than 10 times as likely to be diagnosed with an anxiety disorder than are nondepressed persons (Boyd et al., 1984). High rates of depression are commonly found in the course of anxiety disorders, especially in individuals with panic disorder, agoraphobia, and obsessive-compulsive disorder (Akiskal, 1990). Although cases of anxiety without depression are relatively common, cases of depression in the absence of anxiety are quite rare (Hamilton, 1983). There is some evidence that anxiety tends to precede the onset of depression, and may be a response to the debilitating effects of anxiety. Once established, anxiety disorders are not uncommonly followed by a depressive disorder, whereas depressive disorders are rarely succeeded by an anxiety disorder. Angst, Vollrath, Merikangas, and Ernst (1990) reported that 49% of their cases with pure anxiety disorders went on to develop depression, whereas only 33% of the cases with pure depression later developed an anxiety disorder.

Merikangas (1990) reviewed family and genetic studies of anxiety and affective disorders, which provide some evidence for a common genetic pathway for these disorders. Although the results of this review were inconsistent with respect to familial overlap, nearly all the studies reviewed found a high degree of comorbidity of anxiety and depression in the probands and to a lesser extent, their relatives. Families in which probands exhibited both anxiety and depressive symptoms resembled depressed families more so than families of probands with anxiety disorders alone. Likewise, families of probands with anxiety or panic disorder without depression did not tend to have increased familial loading for depression. In addition to finding support for the notion of pure forms of anxiety and depression breeding true, there was also substantial evidence for a mixed form of anxiety–depression in which anxiety and depression in families appear to have a common underlying genetic contribution. More recently, in a study of 133 depressed and 82 controls probands, Merikangas, Risch, and Weissman (1994) assessed patterns of comorbidity and cotransmission of anxiety disorders, major depression, and alcoholism. Their findings indicated that anxiety, depression and alcoholism are cotransmissible in families. This relationship was especially strong for anxiety disorders and major depression. Interestingly, their cotransmission analyses indicated that the observed comorbidity could be explained entirely by transmitted (i.e., genetic) rather than environmental factors. This finding supports the hypothesis that these disorders are manifestations of a shared genetic factor.

6.16.5.2 Comorbidity with Substance Use Disorders

Another type of Axis I disorder commonly found among individuals with affective disorders involves alcohol and substance abuse and dependence. Among individuals with major depression the lifetime prevalence rates for alcohol use and substance use disorders are 17% and 18%, respectively (Regier et al., 1990). When these figures are used to calculate odds ratios (the increased risk of developing a substance use disorder associated with depression), individuals with major depression are 1.5 times more likely to have a comorbid alcohol use disorder and 1.8 more likely to develop a substance use disorder.

In the ECA study, bipolar disorder was found to have the highest comorbidity with drug or alcohol abuse of all the Axis I disorders (Regier et al., 1990). The lifetime prevalence of alcohol-use disorders in individuals with bipolar disorder was 46%, which represents a risk ratio of 5.3. The percentage of substance abusers among individuals with bipolar illness was 34%, equivalent to an odds ratio of 8.3. Among patients in treatment for bipolar illness, cocaine and alcohol appear to be the most common agents of abuse (Fogarty et al., 1994; Sonne, Brady, & Morton, 1994). Initially, substance abuse during a manic episode may simply be part of the pattern of excessive and risky behavior. However, repeated substance use may cause changes in the mood-regulating centers of the brain, which then leave the brain more vulnerable to recurrences of affective episodes (Post et al., 1986). This is similar to the kindling process that may result from affective episodes themselves. In each case, it may be that there is a permanent degradation of brain regulatory processes.

One hypothesis regarding the increased risk of alcohol and substance abuse among individuals with affective disorders is that their use serves to alter negative mood states (Kantzian, 1985). Freud (1884, cited in Goodwin & Jamison, 1990, p. 223), recognizing its mood elevating properties, believed that cocaine could be used as a potent antidepressant. However, manics' increased cocaine use appears to be related to their attempts to enhance or prolong euphoric mood states, rather than to ameliorate depressive symptoms (Weiss & Mirin, 1987).

Evidence for a common genetic pathway of alcohol use disorders and affective disorders is inconclusive. Studies examining this issue have compared the prevalence of illness in relatives of patients against that found in controls. The concept of depression spectrum disease (Winokur, Cadoret, Dorzab, & Baker, 1971) came from the finding that unipolar depressed patients with alcoholic relatives differ in important ways from those without such relatives. Specifically, female relatives of patients with early-onset depression were more likely to be depressed, whereas male relatives were more likely to be alcoholic. Depressions could be divided into pure depressive disease (only depressions in the family) and depression spectrum disease (depressions with family histories of male alcoholism). Although many family studies have not shown an increase in alcoholism in the relatives of depressed patients (Merikangas & Gelernter, 1990), several adoption studies have demonstrated an increase in alcoholism in the biological relatives of depressed adoptees (Cadoret et al., 1996; von Knorring et al., 1983; Wender et al., 1986). Cadoret et al. concluded that the findings of adoption studies indicate the presence of a genetic factor for which alcoholism is a marker, and that a gene–environment interaction is an important etiological factor in depressive spectrum disease. Whereas the relationship between unipolar depression and alcoholism is mixed, it is clear that alcoholism does not belong in the genetic spectrum of bipolar manic-depressive illness (Gershon, 1990).

6.16.6 SUICIDE

Suicide is the ninth leading cause of death in the USA, accounting for more than 31 000 deaths in 1995 (Anderson, Kochanek, & Murphy, 1997). It is estimated that psychiatric illness is implicated in suicide 90% of the time (Goldstein, Black, Nasrallah, & Winokur, 1991). Affective illness is a well-established risk factor for suicide, and as such is a psychiatric disorder associated with serious risk of mortality. Indeed, the estimated lifetime probability for suicide among psychiatric patients with affective disorder is 15% (Guze & Robins, 1970).

Several risk factors for suicide in psychiatric and community samples have been identified, including demographic variables such as male gender, divorced or widowed civil status, unemployment, and older age (see Hawton, 1992, for a review). Psychiatric history variables, other than affective disorder, have also been investigated. Alcoholism (Martin, Cloninger, Guze, & Clayton, 1985), previous suicide attempts (Borg & Stahl, 1982), past psychiatric hospitalizations (Pokorny, 1983), and premature termination or unfavorable outcome of therapy (Dahlsgaard, Beck, & Brown, 1998; Goldstein et al., 1991) are associated with increased suicide risk.

Most recently, research efforts have been focused on what presenting symptoms, or psychological variables, may predict suicide in patients suffering from affective disorders. Unlike demographic and psychiatric history variables, psychological variables are considered "modifiable risk factors" (Fawcett et al., 1990), in that they may change as a result of treatment. One such factor is hopelessness. Hopelessness in the depressed patient manifests itself as a belief that there are no solutions to serious life problems, and that the future holds no respite from intolerable life circumstances. In inpatients hospitalized with suicide ideation, hopelessness has been found to be a better predictor of suicide than depression (Beck, Steer, Kovacs, & Garrison, 1985); it has also been found to predict suicide for up to one year after its assessment (Fawcett et al., 1990). In outpatients, hopelessness has also been demonstrated as a significant predictor of eventual suicide (Beck, Brown, Berchick, Stewart, & Steer, 1990; Beck, Brown, Steer, Dahlsgaard, & Grisham, in press).

The insidious nature of hopelessness in depression is illustrated in studies by Young et al. (1996) and Dahlsgaard et al. (1998). In both studies, suicide was not predicted by levels of hopelessness assessed at the beginning or middle of treatment. In Young et al., suicide attempts were predicted by high, stable levels of hopelessness assessed when these patients' depressive episodes had remitted. Similarly, Dahlsgaard et al. found that suicide was most likely to occur in patients whose levels of hopelessness did not decrease during treatment.

Anxiety symptoms have been related to suicide in work by Fawcett et al. (1990). Severe psychic anxiety, including the presence of panic attacks, and anxiety symptoms such as insomnia and diminished concentration, were strongly related to completed suicide occurring within one year of their assessment.

Suicide ideation has also been proposed as an obvious, and in principle modifiable, risk factor for suicide. However, communication of suicide ideation at initiation of clinical contact has not been found to predict eventual suicide (Beck et al., 1985). More powerful predictions can be made from retrospective reports of suicide ideation at its most severe point (Beck et al., in press).

Biological markers of suicide risk have recently been discovered. Indications of deficits in the neurotransmitter serotonin have been observed in depressed suicide attempters (Lester, 1995). This has been found to be especially true in those with more serious attempts (Mann & Malone, 1997), and those who went on to complete suicide within 1 year (Nordström et al., 1994).

6.16.7 THEORIES AND TREATMENTS OF UNIPOLAR DISORDER

6.16.7.1 Biological Theories and Treatments of Unipolar Disorder

There is no one biological theory of unipolar depression. Instead, the biological characterization of the state of depression is quite complex and incomplete, and it involves many biological systems and levels of analysis. One can begin at the level of the neurotransmitter, which is where modern theories of the biology of depression have focused. Abnormalities in hormonal functioning, as well as structural abnormalities in the brain, have also been examined.

In its simplest form, the amine theory of depression, originally proposed by Schildkraut (1965) and Glassman (1969), is that depression is caused by decreases in brain levels of the neurotransmitter norepinephrine (NE) or serotonin (5-HT), both of which belong to a class of neurotransmitters called amines. Evidence for this theory came in part from observing neuronal effects of drugs that reduce depressive symptoms. The tricyclic antidepressants (TCAs) inhibit the reuptake of NE and 5-HT into the presynaptic terminal. The inhibition of reuptake is presumed to make more amines available for neurotransmission, which in turn is presumed to return the amine systems to normal levels of functioning. Normal amine functioning, in turn, translates into normal mood functioning (but see Richelson [1991] who argues that the amine effects explain the side effects, not the therapeutic actions of the drugs).

The amine story cannot be as simple as the version described above. The effects of TCAs on the inhibition of reuptake are nearly immediate, yet these drugs take more than a week to exert their therapeutic effect. A partial explanation of this delay involves a regulatory system called the α_2 autoreceptor. This receptor, which sits on the presynaptic membrane, regulates the release of NE into the synapse by detecting how much NE is in the synapse already, and inhibiting the further release of NE. After days or weeks of treatment, during which time the α_2 autoreceptor is fighting to maintain the pretreatment synaptic level of NE, it downregulates. This means that it reduces its activity, and thereby allows more NE to be released into the synapse.

An even more sophisticated analysis of the role of amines in depression has recently been articulated by Heninger, Delgado, and Charney (1996), based on findings from biological challenge tests. Heninger et al. conclude that the role of amines is to regulate other biological mechanisms, likely to be found within nerve cells, that have a more direct effect on mood and other symptoms of depression.

The brain structure most consistently implicated in mood disorders is the hypothalamus. The effect of hypothalamic activity on other biologic systems has thus been of great interest (Musselman & Nemeroff, 1996). Hypothalamic activity affects the pituitary gland, which in turn affects activity in the thyroid and adrenal glands. Abnormalities of thyroid activity have been found in depressed patients, and medication treatments that include the use of synthetic thyroid hormone have been found to be effective. The adrenal hormone cortisol, commonly considered a stress hormone, tends to be elevated in depressed people. Moreover, the cortisol regulation system, as assessed by the dexamethasone suppression test (DST), is impaired in many depressed patients.

Individual tests of thyroid and adrenal malfunction have thus far not been sensitive and specific enough to guide diagnosis and treatment. Not surprisingly, the combined results of several tests yield better results. For example, when Banki, Arato, Papp, Rihmer, and Kovacs (1986) used four neuroendocrine tests, they found that 11 of 21 (52%) depressed patients obtained abnormal results on at least three of them. Only 3 of 44 (7%) of nondepressed psychiatric patients, and 0 of 15 nonpsychiatric patients, achieved a similar result.

In addition to neurotransmitter and neuroendocrine abnormalities, abnormalities of brain structure have been identified (Soares & Mann, 1997). Although global brain atrophy has not been found, the basal ganglia, the cerebellum, and the frontal lobe appear to be reduced in size in unipolar depression. Each of these areas has connections to brain structures known to participate in the regulation of mood. For example, the cerebellum projects to norepinephrine cell groups in the limbic system. Therefore, cerebellar atrophy is suspected to play a role in the inability of the limbic system to regulate processes that are responsible for normal mood functioning.

The major biological treatment modalities for unipolar depression are medication and electroconvulsive therapy. Antidepressant medications fall into three classes: tricyclic antidepressants, selective serotonin reuptake inhibitors, and monoamine oxidase inhibitors.

Treatment of depression with antidepressants is effective in cases of moderate to severe depression. The utility of medicines for milder forms of depression, however, is in doubt. Elkin et al. (1989) found no advantage of the antidepressant medication imipramine over a placebo condition for patients who, at intake, scored below 20 on the Hamilton rating scale for depression (Hamilton, 1967). The advantage of imipramine over placebo was clear, however, among the more severely depressed patients (see also Paykel, Hollyman, Freeling, & Sedgwick, 1988; Stewart et al., 1983). Typical symptom clusters that predict a favorable response to antidepressant medications are: decreased appetite, insomnia, acute onset, a family history of depression, and a personal or family history of positive response to antidepressant medications or electroconvulsive therapy (APA, 1980).

The efficacy of TCAs is demonstrated most thoroughly by research on imipramine. The oldest and most widely prescribed TCA, its most common brand name in the USA is Tofranil. Imipramine has been compared to placebo in nearly 50 double-blind studies. The combined findings reveal a 65% improvement rate with imipramine, compared to 30% with placebo (Goodwin, 1992). Other commonly used TCAs are amitriptyline (Elavil) and doxepin (Sinequan).

As with all medicines, side effects can occur with TCAs. Common cardiovascular side effects include tachycardia (increased heart rate) and postural hypotension (a sudden drop in blood pressure when arising from a bed or a chair). Neurological side effects are fairly common, especially sedation. Due to the neurological and cardiovascular toxicity of TCAs, the use of these medicines in suicide attempts is a serious risk. For this reason, special caution is exercised when these drugs are used to treat suicidally depressed patients.

Administration of selective serotonin reuptake inhibitors (SSRIs) is indicated in the same cases of depression that indicate administration of TCAs—moderate to severe depression of a fairly typical character. These newer antidepressants are often preferred to TCAs because they exhibit comparable effectiveness but have reduced side-effects and a much reduced overdose potential relative to TCAs (see Greenberg, Bornstein, Zborowski, Fisher, & Greenberg's [1994] meta-analysis of fluoxetine [Prozac] outcome). Besides fluoxetine, other commonly prescribed SSRIs are paroxetine (Paxil) and sertraline (Zoloft). Common side effects of the SSRIs tend to be gastrointestinal or sexual, including nausea, diarrhea, and sexual dysfunction. Decreased appetite, nervousness, and tremor are also reported (Guze & Gitlin, 1994). Because the cardiac and neurological side effects are less than those of TCAs, SSRIs are associated with a much lower risk for use in lethal overdose.

Monoamine oxidase inhibitors (MAOIs) are generally indicated in cases of atypical depression that would not respond adequately to TCAs or SSRIs (Quitkin et al., 1993). The category atypical depression includes nonmelancholic depression, or depression with anxiety, rejection sensitivity, chronic pain, or the reversed vegetative symptoms of appetite increase and hypersomnia (Lam & Stewart, 1996). The medicines phenelzine (Nardil) and tranylcypromine (Parnate) belong to the class of MAOIs.

Whereas the TCAs and SSRIs are presumed to alter neurotransmission functioning by inhibiting reuptake of amines into the presynaptic cleft, MAOIs act through the inhibition of monoamine oxidase (MAO). MAO is an enzyme that breaks down norepinephrine, either in the synapse or in the presynaptic terminal (McNeal & Cimbolic, 1986).

Frequently reported side effects of MAOIs include headaches, dry mouth, insomnia, constipation, sedation, blurred vision, and nausea. Of particular interest is the potential for unpleasant, or even fatal, interactions of MAOIs with foods containing high levels of tyramine. While taking MAOIs, one must avoid aged cheeses, red wines, soy sauce, smoked fish, and other foods containing tyramine (Davidson, 1992).

Electroconvulsive therapy (ECT) is generally considered a last resort treatment, to be used with patients whose depression severity requires hospitalization, or those who are not adequately responsive to antidepressant medications (Fink, 1992). There are cases, however, in which ECT is considered a primary treatment, such as when a rapid treatment response is needed, when risks of medications are greater than those of ECT, when a poor history of drug response and/or a positive history of response to ECT exists, or when the patient prefers ECT to other treatments. ECT is especially effective when depression is accompanied by psychosis or marked psychomotor retardation and depressive stupor (Enns & Reiss, 1992; Fink, 1992).

ECT has been shown to be more effective than sham ECT in double-blind controlled studies in the treatment of major depression (Gregory, Shawcross & Gill, 1985). Further, ECT has been shown to be more effective than moderate doses of TCAs (Ghangadhar, Kapur, & Kalyanasundaram, 1982).

Although the mechanisms through which ECT achieves its antidepressant effects are poorly understood, it is known that the induction of a *grand mal* seizure is essential for the success of ECT. One possibility is that the release of norepinephrine and dopamine during the seizure leads to overall improvement in the functioning of amine systems (Sackeim, 1989).

Patients with major depression generally receive seven to nine treatments, three times per week, with improvement generally evident after four or five treatments. The most common side effects associated with ECT as it is currently implemented are cognitive impairment and memory loss. Although memory loss is generally limited to the periods around the time of the treatments, some patients report loss of memory of events that occurred before ECT, or even before the depressive episode (Fink, 1992).

6.16.7.2 Cognitive Theories and Treatments of Unipolar Disorder

The cognitive theory has been the most influential psychological theory of depression since the 1970s. Its heuristic value as a way of describing the experience and behavior of the depressed person has been widely accepted. Its value as an etiological theory and as a basis for treatment, especially in severe forms of depression, has been more controversial.

Beck's (1967) early writings about depression came against the backdrop of the influence of psychoanalytic understandings of depressive phenomena. He argued that cognitive phenomena that are available to introspection are at the core of, and can explain the presence of, the other features of depression. Depressive mood is said to be the natural consequence of negative beliefs about the self, the future, and the world (the negative cognitive triad). Likewise, behavioral disturbances, such as response initiation deficits, are seen as resulting from a belief in the futility of action. For example, whereas in a psychoanalytic formulation of depression the failure of a depressed person to keep an appointment for a job interview might be interpreted as an instance of unconsciously motivated self-punishment, the cognitive model would focus on the negative expectations of the depressed person (e.g., "I'll make a fool of myself"; "I'll never get that job anyway"). Findings from literally hundreds of studies (Haaga, Dyck, & Ernst, 1991) have supported this characterization of the thinking of depressed people.

What makes a person vulnerable to this kind of thinking? Is depressive thinking, like sleep disturbance, merely a symptom of depression that arises during the episode, or are there premorbid cognitive phenomena that act as vulnerabilities? Beck has argued that depressed people possess "depressive schemas" even before the first onset of a depressive episode. However, because these schemas are latent, they are difficult to detect. Miranda and Persons (1988) have proposed that latent schemas can be detected, in principle, by assessments conducted when the vulnerable person is in a sad mood (Hedlund & Rude, 1995).

A research literature relevant to the question of cognitive vulnerability concerns the reformulated learned helplessness, or hopelessness, model of depression (Abramson, Metalsky, & Alloy, 1989). According to this model, vulnerability to depression is in the form of characteristic ways of explaining events, especially bad ones. This pattern is referred to as a depressogenic inferential style. The person with this style has a tendency to attribute bad events to causes that are internal (due to, or inherent in, the person), stable (likely to persist over time), and global (applicable to many domains of the person's life). Given this style, the occurrence of a bad event is especially likely to lead to specific depressogenic inferences, which will then lead to negative expectations, or a state of hopelessness. The state of hopelessness then generates the other symptoms of depression.

Research findings have yielded equivocal support for this model (Barnett & Gotlib, 1988). Abramson, Alloy, and Metalsky (1988) argue, however, that many investigators have used research designs that are not appropriate to test the set of hypotheses embedded in the hopelessness model. A large study designed specifically to test the hopelessness model is near completion; its findings strongly support the hypothesis that cognitive vulnerability interacts with stressful life events to produce the beliefs, and then the symptoms, associated with depression (Abramson, Alloy, Hogan, & Whitehouse, 1997; Alloy, Abramson, Murray, Whitehouse, & Hogan, 1997).

A critical aspect of any psychological model of depression is its ability to delineate distal causes. If there are persons who possess a thinking style that makes them vulnerable to depression in the face of negative life events, it is likely that both genetic (Schulman, Keefe, & Seligman, 1993) and environmental factors play a role in the development of that style. The recent findings reported by Garber and Flynn (in press) suggest that a relative absence of maternal acceptance and the presence of maternal psychological control contribute to the development of thinking styles characteristic of persons at risk for depression. These findings accord with the observation that depressed adults report having experienced a parental style of affectionless control (Parker, 1992). Garber and Flynn also found that whereas correlations between children's and parent's thinking styles were low, mothers' explanations of their children's behavior were related to their children's own explanations regarding their behavior. Hence, only on explanations of child-focused events, and not general thinking style, were mothers and children similar.

Beck's cognitive therapy for depression (Beck, Rush, Shaw, & Emery, 1979; DeRubeis & Beck, 1988) has for the last 20 years epitomized modern, brief, semistructured forms of psychosocial interventions for serious mental disorders. Two factors have been especially important in the popularity of this form of therapy for depression. First, beginning in 1977 (Rush, Beck, Kovacs, & Hollon, 1977), there has been a series of impressive demonstrations of the efficacy of cognitive therapy. Second, Beck et al.'s (1979) description of cognitive therapy, as detailed in their classic therapy manual, has been admired for its clarity, detail, and thoroughness. Professionals who find descriptions of psychoanalytic methods too unstructured, and behavior therapy too restrictive, find a middle ground in cognitive therapy. Similarly, it has captured a segment of the educated public, who find that the ideas embedded in cognitive therapy (Burns, 1980) are both practical and consistent with common sense.

The beginning assumption in cognitive therapy for depression is that depressed patients frequently engage in ill-considered negative thinking that, when held up to scrutiny, will be found in error. Although there has been an interest by some writers, beginning with Freud (1917/1957; Alloy & Abramson, 1982, 1988), in the idea that depressed people see the world more clearly than nondepressed people, there is no good evidence for this view (Ackermann & DeRubeis, 1991). Moreover, this depressive realism hypothesis runs counter to common sense; the most interesting version of this hypothesis implies that the most severely depressed person will also be the wisest, clearest thinker. This proposition is, *prima facie*, incorrect if not foolish.

The therapist's first task in cognitive therapy, then, is to teach the patient to attend to their thoughts, some of which will be distorted. These thoughts may be quite prominent in the patient's attentional focus, even before the beginning of therapy, in which case the patient merely needs to learn to remember them or, better yet, write them down for further consideration. It may take extensive instruction and practice before the patient learns to capture the stream of thoughts that accompanies almost any experience. Of course, the greatest efforts are directed at observing thoughts that occur before, or during, an unpleasant emotional state, such as sadness or despair. The term "automatic thoughts" has been used to refer to the spontaneous nature of many of these thoughts. Their automatic, effortless, tacit, and habitual nature can make them especially difficult to notice.

The kinds of negative thoughts most frequently reported by depressed patients involve beliefs about the self (e.g., "I'm worthless"), the future ("I'll never get another job"), or the world ("People are uncaring"), that is, Beck's (1967) negative cognitive triad. Hollon and Kriss (1984) classify automatic thoughts into (i) attributions, which are explanations about past events (e.g., "The breakup with my girlfriend was all my fault"); (ii) ascriptions, which are descriptions of present states (e.g., "I'm unlovable"); and (iii) expectations, which are predictions about the future (e.g., "I'll never marry"). However one classifies these thoughts, the assumption is that they are observable, either as (complete or partial) sentences, or as visual images (e.g., an image of oneself, old and alone, with no loved ones present). It is further assumed that many of the thoughts are incorrect or invalid, or that the implications depressed patients derive from the thoughts are exaggerated toward the negative (e.g., "Since my relationship is over, I will never find another good relationship").

For many patients, once the thoughts are noticed, they can be seen as exaggerations or distortions. Therapy is facilitated further when patients write the thoughts on paper. The concreteness of thoughts written down lends them to serious, dispassionate reconsideration.

The next element of therapy, then, entails teaching the patient the most effective means of reconsidering their thoughts. Four classes of questions address most thoughts reported in therapy. The first question is, "What is the evidence that the thought is true, or not true?" The first half of the question is aimed at getting the patient to explicate the basis for a belief; the second half will challenge the belief, if there is relevant evidence to counter it. The second question is, "Are there alternative explanations that have not been considered?" This question is especially relevant when patients assign self-blame inappropriately or excessively. The third question—"What are the implications if the belief is true?"—is directed at the tendency for depressed patients to imagine exaggerated consequences of real failures or shortcomings. The fourth question—"What can I do?"—follows from the others. Once a problem has been identified, relatively free of distortion and exaggeration, a plan of action can be developed rationally.

Cognitive therapy has also been called cognitive behavior therapy, in part because the fourth question focuses on action. Moreover, especially early in treatment, behavioral techniques are employed to counter the failure of depressed people to initiate action that could lead to pleasure or a sense of accomplishment. Cognitive therapists help patients break tasks into their smaller components, and to plan to engage in activities they have been avoiding. At each point, the cognitive therapist is interested both in increasing the patient's engagement in the world, and in discovering the distorted beliefs that have prevented them from doing so during the depression.

Cognitive therapy is a skills-based treatment (Barber & DeRubeis, 1989). The hope is that patients will not only benefit from the therapists' interventions, but that they will learn how to apply the interventions to themselves. But cognitive therapy, especially as recent writers have described it (Persons, 1989), is also a kind of insight therapy. Therapists help patients to discover idiosyncratic patterns of belief, or schemas, that they apply across a wide range of situations. Therapy can then focus on the developmental origin of the schemas. Many modern cognitive therapists believe that therapy is unsuccessful if it does not include the identification of schemas and their origins. However, all of the outcome research has been conducted with therapists who have not placed great emphasis on the schema aspect of cognitive therapy. It remains to be seen whether it is the acquisition of skills, or the understanding of themes and origins, that is most critical to the success of cognitive therapy (Barber & DeRubeis, 1989).

In 1977, Rush, Beck, Kovacs, and Hollon published the first findings indicating that a psychosocial treatment could intervene in a depressive syndrome as quickly and effectively as a pharmacological treatment. In fact, due in part to a design artifact, cognitive therapy was said to be significantly more effective than imipramine pharmacotherapy in that study. Findings from subsequent studies (Hollon et al., 1992; Murphy, Simons, Wetzel, & Lustman, 1984) converged on the conclusion that cognitive therapy is as effective and as fast-acting as antidepressant medications as an acute treatment of depressed outpatients. Moreover, follow-up results from these studies, especially the Evans et al. (1992) follow-up of the Hollon et al. study, suggested that short-term cognitive therapy prevents relapse relative to equally short-term pharmacotherapy (Hollon, Shelton, & Loosen, 1991).

These conclusions were widely accepted until the publication of findings from the large, multicenter Treatment of Depression Collaborative Research Program (TDCRP; Elkin et al., 1989). Whereas none of the previously mentioned studies included a control group in their design, the TDCRP included a pill-placebo condition, as well as cognitive therapy, imipramine pharmacotherapy, and interpersonal therapy conditions. In the TDCRP, cognitive therapy did not outperform placebo significantly, and some analyses of the outcome data indicated that imipramine was more effective than cognitive therapy for the more severely depressed patients (Elkin, Gibbons, Shea, & Sotsky, 1995). This pattern of findings from the TDCRP led the APA (1993) to publish treatment guidelines that questioned the effectiveness of cognitive therapy for depression, especially in its more severe forms. Recently, DeRubeis, Gelfand, Tang, and Simons (1998) have reanalyzed data from only the more severely depressed patients from the Rush et al. (1977), Murphy et al (1984), Hollon et al. (1992), and TDCRP studies. Considering the patients in all four studies combined, cognitive therapy was at least as effective as medications during the acute treatment phase. Thus, treatment guidelines that focus only on the failure of cognitive therapy in the TDCRP appear to underestimate the short-term benefits of cognitive therapy for severely depressed outpatients. Further research on cognitive therapy and pharmacotherapy for severely depressed outpatients is currently being conducted. This study will also examine the putative relapse prevention benefits of cognitive therapy.

The specificity of cognitive therapy for depression, once inferred primarily from its equivalence to pharmacotherapy (Dobson, 1989), is now being questioned. Not only have there not been many comparisons of cognitive therapy to other psychosocial interventions, there has also been virtually no research comparing cognitive therapy to control conditions. In the TDCRP, cognitive therapy did not surpass the interpersonal therapy condition or the placebo condition. Indeed, there were some indications of the superiority of interpersonal therapy in that study, especially in the more severe sample. A recent study by Jacobson et al. (1996) suggested that behavior therapy is as effective as cognitive therapy for depression. Jacobson is following up this finding in a larger design that will include a pharmacotherapy group, as well as a placebo control. Indications of the potency of cognitive therapy gleaned from the studies summarized in DeRubeis et al. (1998) cannot be ignored, but the set of investigations currently underway will refine the answers to questions about the short- and long-term benefits of cognitive therapy for depression.

6.16.7.3 Behavioral Theories and Treatments of Unipolar Disorder

Lewinsohn (1974) has written most extensively about the role of behavior theory and behavior therapy in depression. In the behavioral view, depression is seen as the result of a low rate of positive reinforcement. The depressed person is unable to obtain gratification, either because there are very few reinforcing stimuli in their environment, or because of a deficit in skills, especially social skills. Depressed mood and behavior become part of a vicious cycle, in which depressive behavior, because others find it unpleasant, is especially unlikely to lead to positive reinforcement from others. According to this view, there are two promising points of entry in therapy for depression: acquisition of skills, especially social skills; and enrichment of the environment, by identifying, planning, and engaging in potentially reinforcing activities.

In two studies behavior therapy has outperformed control conditions. Shaw (1977) found that patients in his behavior therapy group improved more than those in a wait-list condition. (However, they improved less than those in his cognitive therapy group.) McLean and Hakstian (1979) conducted a large study in which they compared behavior therapy to psychotherapy, relaxation therapy, and amitriptyline pharmacotherapy. Significantly greater symptom change was observed in their behavior therapy condition than in each of the other three conditions.

In a recent study of 150 outpatients by Jacobson et al. (1996), behavior therapy produced as much change in depressive symptoms as either of two cognitive therapy conditions. This result was not due to poor outcome in the cognitive therapy conditions—recovery rates and mean end-of-treatment scores on standard measures were impressive in all three groups. Jacobson and colleagues are following up these findings with an even larger scale study that will pit behavior therapy and cognitive therapy against each other, as well as against medication therapy and a placebo control condition.

Nezu (1986) has described a behavioral program he calls problem-solving therapy, which assumes that depressed persons have lost (or never possessed) the skills necessary to solve life problems. Problem solving therapy has performed well in two controlled comparisons with depressed outpatients (Nezu 1986; Nezu & Perri, 1989). Problem solving approaches that are directed at the suicidal behavior of depressed people also appear to be promising.

Linehan (1993) has developed a cognitive behavioral approach to the treatment of chronically suicidal patients with borderline personality disorder. Termed dialectical behavior therapy (DBT), the goals of this treatment are to remediate deficits in interpersonal, emotion regulation, self-regulation, and distress tolerance skills, and to encourage the use of these skills, especially during times of suicide ideation. In a randomized comparison with a treatment as usual condition, DBT patients exhibited better overall adjustment, as well as anger control, at treatment termination (Linehan, Tutek, Heard, & Armstrong, 1994). Because the chronically suicidal borderline patient is considered especially difficult to treat, these preliminary indications of the effectiveness of DBT have been greeted with enthusiasm by practitioners and treatment researchers.

6.16.7.4 Psychodynamic Theories and Treatments of Unipolar Disorder

The earliest modern psychological theory of depression was described by Freud (1917/1957) in his classic paper, *Mourning and melancholia*. Central to Freud's theory was the proposition that loss precedes all episodes of melancholia, as it does in all instances of mourning. Freud delineated the similarities between mourning and melancholia, such as sad mood and decreased interest in people and activities, all of which could be accounted for by loss.

Freud then focused on the unique, defining features of melancholia, most importantly self-reproach, and suicidal thoughts and actions. He posited that desires to harm or punish oneself are best understood as deriving from a wish to harm the lost object, or some aspect of it. So, for example, the depressed person who fails to show up for a job interview could be seen as manifestly punishing or harming him- or herself, but the unconscious motive for this may be a wish to punish his or her deceased father. Necessary for the development of this wish to punish another object is a history of ambivalent love for that object. Under conditions of ambivalent love, when the object is lost and is incorporated into the self (ego), both the loved and hated elements of the object become part of the ego. From that point on, when a wish to retaliate against or harm the lost object occurs unconsciously, it is manifest as a wish to harm the self, through the defense of retroflection, whereby anger towards another is transformed into anger turned inward. Many mental health professionals believe that retroflected anger is an important basis for depressive symptoms, especially those that are self-defeating. Belief in the phenomenon of retroflected anger has formed the basis for the popularity of efforts to help depressed people release their anger. Nonetheless, there have been no systematic investigations of the effectiveness of Freudian psychoanalysis for depression, or of any treatment based on the anger-turned-inward hypothesis of depression.

Whereas the psychoanalytic theory of depression focuses on a pathological motivational system, and cognitive theory focuses on pathology of belief, interpersonal–psychodynamic theory of depression focuses on the pathological relation of the person to their interpersonal environment. The interpersonal theory of depression is a specific application of a general interpersonal theory of psychopathology, whose best known advocates are Meyer (1957) and Sullivan (1953). In the interpersonal theory of depression (Klerman, Weissman Rounsaville, & Chevron, 1984), four types of interpersonal pathology are elaborated. A given depressed person is expected to exhibit one or more of these. Abnormal grief refers to either a prolonged grief reaction, or a distorted grief reaction in which normal grieving processes do not occur. Interpersonal role disputes involve relationships that are one-sided or nonreciprocal. Abnormal role transitions refer to a person's inability to cope with life changes. Finally, interpersonal deficits are seen in people who have had a poverty of social relations throughout their lives.

The interpersonal theory of depression is associated specifically with a brief, structured psychosocial intervention for depression, termed interpersonal psychotherapy (IPT; Klerman et al., 1984). The benefits of IPT in depressed outpatients were first examined by DiMascio et al. (1979). IPT achieved symptom reduction equivalent to that observed in a medication condition, although the medications achieved their effects more rapidly. In comparison to a minimal contact control condition, IPT effects were superior.

Interpersonal therapy also performed well in the TDCRP (Elkin et al., 1989). At post-treatment, the depressive symptoms of IPT patients reached levels almost as low as those achieved in the medication condition, and slightly lower than those of patients treated with cognitive therapy. (Neither difference was significant.) When the authors examined data from the patients who were more severely depressed prior to treatment, there were some indications that IPT was superior to placebo.

Another psychodynamically based intervention, psychodynamic interpersonal psychotherapy (PI; Shapiro et al., 1994), has been examined as a treatment for depression. The main difference between PI and IPT appears to be that PI focuses more on the relationship between the patient and the therapist, using it as a vehicle for understanding the patient's interpersonal relationships more generally. In psychodynamic terms, PI attends to the transference relationship between patient and therapist. Shapiro et al. found that on some measures, PI performed as well as a version of cognitive behavioral therapy in a large sample of depressed outpatients. However, there were also some indications that PI was not as effective as cognitive behavior therapy.

Thus, although there has been little effort to test Freud's psychoanalytic treatment for depression, active research programs exist that aim to evaluate more structured versions of therapies that derive from psychoanalytic thought. This is part of a modern trend for psychodynamic thinking to be instantiated in short-term, focused interventions for specific disorders (Barber & Crits-Christoph, 1995).

6.16.8 THEORIES AND TREATMENTS OF BIPOLAR DISORDER

6.16.8.1 Biological Theories and Treatments of Bipolar Disorder

Neurochemical theories of mania invoke the same neurotransmitter systems as are implicated in unipolar depression. In Schildkraut's (1965) classic formulation, whereas depressive disorders were associated with a functional deficit of catecholamines (particularly norepinephrine), manias were said to involve a functional excess of catecholamines. The major metabolite of the catecholamine norepinephrine (NE), namely 3-methoxy-4-hydroxyphenylglycol (MHPG), has been shown to be elevated in bipolar patients when they were manic, compared with levels when depressed (Halaris, 1978). Similarly, bipolar patients, when manic, demonstrate significantly higher plasma NE and epinephrine levels than when they are depressed or euthymic (Maj, Ariano, Arena, & Kemali, 1984). Swann et al. (1987) reported that MHPG and NE levels in cerebral spinal fluid were significantly higher in manic patients than in either depressed or control subjects.

The permissive hypothesis of serotonin (Prange, Wilson, Lynn, Alltop, & Stikeleather, 1974) posits that a deficit in serotonergic transmission permits the expression of manic depressive illness, although a deficit is not sufficient to cause it. According to this theory, both the manic and depressive phases of bipolar illness are characterized by low serotonin (5-HT) function, but differ in terms of NE activity (high in mania, low in depression). Serotonin systems inhibit a variety of functions served by other neurotransmitters. Defective (low) serotonergic function, and the corresponding lack of inhibition of other neurotransmitter systems, especially NE and dopamine (DA), permits wide variations in affective symptomatology from depression to mania.

Goodwin and Sack (1974) analyzed the effects of drugs on the DA or NE system in relation to the relative efficacy of the drugs against different symptoms of mania. Their analysis led them to suggest that DA abnormalities are primarily involved in the hyperactivity and psychosis associated with the more severe stages of mania, whereas excesses in NE underlie the euphoria and grandiosity more characteristic of less severe forms of mania, or hypomania.

Janowsky, El-Yousef, Davis, and Sekerke (1972) postulated that an increased ratio of cholinergic to adrenergic activity underlies depression, and that the reverse is true in mania. In this view, mania results when increases in NE (an adrenergic transmitter) effect an imbalance relative to cholinergic (i.e., acetylcholine) activity. This hypothesis derives from the fact that antidepressant medications have substantial anticholinergic properties. Moreover, drugs that stimulate acetylcholine decrease manic symptoms in manic patients, and they increase depressive symptoms in depressed patients. These same drugs also induce depressions in nondepressed subjects.

Clinical and pharmacological data suggest that γ-aminobutyric acid (GABA) metabolism is altered in bipolar disorder. GABA has inhibitory input into other neurotransmitter systems such as NE and DA, and regulates many central nervous system functions, including seizure threshold. The role of GABA in convulsive disorders has been extensively studied, and antiepileptic drugs appear to act directly or indirectly on GABA. The antiepileptic agents valproic acid and carbamazepine were serendipitously noted to have a mood-stabilizing effect, decreasing the intensity and frequency of both manic and depressive phases of bipolar disorder. In addition to valproic acid, other antimanic agents, such as lithium and carbamazepine, appear to stabilize mood by increasing GABA-ergic transmission, leading to the hypothesis that GABA deficiency plays a role in mania (Bernasconi, 1982).

The behavior of the manic person, especially when agitated or violent, can be so extreme as to warrant immediate, aggressive therapeutic action. Medical management is almost always the first line treatment for a manic episode.

Lithium is the oldest of the modern antimanic agents. Lithium's antimanic properties were first discovered by John Cade in 1949, but it was not used in the USA until the 1960s. Since then, it has remained the medication of choice for mania. Plasma lithium levels must be monitored to maintain a therapeutic level, and to avoid toxic effects. The plasma level of lithium required to produce clinical response varies from patient to patient, and can vary within the same patient; patients in an active manic episode typically require higher dosages than they do to remain in a normal mood. Lithium is very effective in the treatment of acute mania; Goodwin and Zis (1979) reported a 78% overall response rate in a review of four double-blind, placebo-controlled studies. In addition to acute treatment, lithium is also effective in reducing future manic and depressive episodes in bipolar patients when used prophylactically (Prien, Caffey, & Klett 1974; see Goodwin & Jamison, 1990, for a review).

The anticonvulsants carbamazepine and valproate are also commonly used. They are used alone or in conjunction with lithium, and are commonly used in the treatment of lithium nonresponders. In a recent meta-analysis, Emilien, Maloteaux, Seghers, and Charles (1996) found the therapeutic efficacies of carbamazepine and valproate to be comparable to lithium in the treatment of acute mania. In Strömgren and Boller's (1985) review of 15 studies, 55% of patients were reported to have had a marked or moderate response to carbamazepine. Carbamazepine and valproate have both been shown to have prophylactic effects as well (Coxhead, Silverstone, & Cookson, 1992; McElroy, Keck, Pope, & Hudson, 1992).

6.16.8.2 Psychosocial Theories and Treatments of Bipolar Disorder

Psychosocial theories and treatments of bipolar disorder follow from an understanding that, although the disorder has predominantly genetic causes, its course is responsive to environmental stressors (Johnson & Miller, 1997). As adjuncts to medication therapy, psychosocial treatments aim to enhance medication compliance, and to increase resilience to stress. Miklowitz and Goldstein (1990) have developed family-focused treatment (FFT) for bipolar disorder. The goal of FFT is to reduce family conflict, which is believed to precipitate relapse. In the initial phase of FFT, patients and their families are educated about the symptoms of the disorder, the importance of medication compliance, and the strategies to employ if the patient begins to relapse. In the later stages of FFT, the psychoeducational focus is on improving family interaction—all family members learn communication and problem-solving skills (Miklowitz, Frank, & George, 1996).

A similar psychoeducational approach is found in the treatment known as interpersonal and social rhythm therapy (IPSRT; Frank et al, 1997). IPSRT methods were developed in recognition that irregular sleep and circadian patterns can be detrimental to the bipolar patient. Patients are shown that interpersonal stressors and social role transitions may upset daily routines, which in turn may disrupt circadian regularity and lead to relapse. During the course of therapy, patients are taught how to negotiate interpersonal dilemmas and role transitions, and they are coached in how to regularize and maintain stable daily routines, even when they encounter stressors in their environments.

Both FFT and IPSRT therapies are currently under investigation in randomized clinical studies. Preliminary results show that both treatments produce high patient retention rates through the first year of treatment (Miklowitz et al., 1996). In addition, IPSRT patients have demonstrated increasing stability of daily routines over the course of treatment (Frank et al., 1997).

Newman and Beck (1992) have described a cognitive model in which mania is characterized by a positive cognitive triad: the self is seen as highly valued and powerful, experience is viewed as overly positive, and the future is replete with opportunity. The distortions associated with this triad are the focus of cognitive theory and treatment.

Cochran (1984) found that the addition of a six-week cognitive therapy component reduced medication noncompliance, as well as the frequency of hospitalizations, during a six-month follow-up. More recently, Basco and Rush (1996) and Newman and Beck (1992) have given extensive descriptions of adjunctive cognitive-behavioral programs for bipolar disorder. The primary goals of these programs are to educate the patient, to increase adherence to medication regimens, and to promote the use of cognitive-behavioral techniques when confronted with affective symptoms. Education about bipolar illness includes instruction in the self-monitoring of prodromal symptoms. Awareness of subtle changes in symptoms or behaviors can then serve as a warning system for relapse. This warning system can cue the patient to have their medications monitored or increased. Efforts to facilitate adherence to medication regimens include addressing patient concerns and beliefs about medication. Behavioral techniques, such as pairing medication intake with another daily activity, are used to ensure that medication is taken regularly.

Cognitive therapy also focuses on the cognitive distortions of manic patients. For example, feelings of grandiosity are perpetuated by magnification of the positive and minimization of, or obliviousness to, negative feedback. Paranoid thinking in mania is reinforced by selective attention to evidence that confirms paranoid beliefs, along with dismissal of disconfirming evidence. These tendencies, as well as the patient's desire to maintain a manic or hypomanic state, are addressed with standard cognitive therapy methods.

In cognitive therapy, behavioral interventions are used to limit self-stimulating behaviors. Activities that have a high potential for dangerous consequences or which might serve to exacerbate the episode are identified early on, and the therapist helps to plan strategies for avoiding or limiting these activities. These include, as in IPSRT, maintaining a regular sleep schedule, as well as limiting spending, alcohol and drug use, and other risky behavior. Problem-solving skills are employed to address stressful life events that might precipitate relapse, and to address difficulties arising in the aftermath of an affective episode, such as the financial difficulties that result from excessive gambling or the loss of employment. Thus far, there have been no controlled clinical studies of cognitive therapy of bipolar disorder.

6.16.9 SUMMARY AND FUTURE DIRECTIONS

As we have seen, the affective disorders comprise a multifaceted, heterogeneous group. Heterogeneity is seen in the variety of causal influences, and in the perspective that is achieved by examining mood disorders from diverse vantage points. Indicative of this diversity is the extraordinarily broad array of effective treatments, from discussions of erroneous thinking in cognitive therapy to the induction of seizures by the electrical stimulation of the brain.

The greatest progress since the 1950s has occurred in diagnostic practices, phenomenological and biological characterizations, and in the development of effective treatments. The next half-century should produce further advances in these areas, but we will also witness the emergence of sophisticated and detailed causal theories that draw from the successes of the last half-century, especially once clarification is obtained about the subtypes of affective disorder. Genetic research, in particular, will inform us about the mechanisms involved in the development of brain processes in a person who is vulnerable to affective disturbance. Studies of the early social factors that lead to vulnerability in adulthood will also be critical in this period, as will research that yields more detailed understandings of the psychological states that predispose a person to mood disorder.

Attempts to understand affective disorders will succeed only if knowledge is integrated from several disparate areas of research. Explorations of the mechanisms and influences of genetics, research on parenting styles, investigations of cognitive styles, and elucidation of the roles played by the interpersonal environment and by life's difficulties will come together to inform a synthetic view of affective pathology and its treatment.

6.16.10 REFERENCES

Abramson, L. Y., Alloy, L. B., Hogan, M. E., & Whitehouse, W. G. (1997, November). Depressogenic cognitive style and the prospective development of depression. In J. Riskind (Chair), *Cognitive styles and psychopathology*. Symposium conducted at the annual meeting of the Association for the Advancement of Behavior Therapy, Miami Beach, FL.

Abramson, L. Y., Alloy, L. B., & Metalsky, G. I. (1988). The cognitive diathesis-stress theories of depression: Toward an adequate evaluation of the theories' validities. In L. B. Alloy (Ed.), *Cognitive processes in depression* (pp. 3–30). New York: Guilford.

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review, 96,* 358–372.

Ackermann, R., & DeRubeis, R. J. (1991). Is depressive realism real? *Clinical Psychology Review, 11,* 565–584.

Akiskal, H. S. (1990). Toward a clinical understanding of the relationship of anxiety and depressive disorders. In J. D. Maser & C. R. Cloninger (Eds.), *Comorbidity of mood and anxiety disorders* (pp. 597–607). Washington, DC: American Psychiatric Press.

Alda, M. (1997). Bipolar disorder: From families to genes. *Canadian Journal of Psychiatry, 42,* 378–387.

Alloy, L. B., & Abramson, L. Y. (1982). Learned helplessness, depression, and the illusion of control. *Journal of Personality and Social Psychology, 42,* 1114–1126.

Alloy, L. B., & Abramson, L. Y. (1988). Depressive realism: Four theoretical perspectives. In L. B. Alloy (Ed.), *Cognitive processes in depression* (pp. 223–265). New York: Guilford.

Alloy, L. B., Abramson, L. Y., Murray, L. A., Whitehouse, W. G., & Hogan, M. E. (1997). Self-referent information-processing in individuals at high and low cognitive risk for depression. *Cognition and Emotion, 11,* 539–568.

Alloy, L. B., Kelly, K. A., Mineka, S., & Clements, C. M. (1990). Comorbidity of anxiety and depressive disorders: A helplessness–hopelessness perspective. In J. D. Maser & C. R. Cloninger (Eds.), *Comorbidity of mood and anxiety disorders* (pp. 499–543). Washington, DC: American Psychiatric Press.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd Rev. ed.). Washington, DC: Author.

American Psychiatric Association (1993). Practice guidelines for major depressive disorder in adults. *American Journal of Psychiatry, 150*(Suppl. 4).

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association Task Force on Laboratory Tests in Psychiatry (1987). *The dexamethasone suppression test: An overview of its current status in psychiatry, 144,* 1253–1262.

Anderson, R. N., Kochanek, K. D., & Murphy, S. L. (1997). Report of Final Mortality Statistics, 1995. *Monthly Vital Statistics Report, 45,* 1–80.

Angst, J., Vollrath, M., Merikangas, K. R., & Ernst, C. (1990). Comorbidity of anxiety and depression in the Zurich Cohort Study of Young Adults. In J. D. Maser & C. R. Cloninger (Eds.), *Comorbidity of mood and anxiety disorders* (pp. 123–137). Washington, DC: American Psychiatric Press.

Aronson, T. A., & Shukla, S. (1987). Life events and relapse in bipolar disorder: The impact of a catastrophic event. *Acta Psychiatrica Scandinavica, 75,* 571–576.

Banki, C. M., Arato, M., Papp, Z., Rihmer, Z., & Kovacs, Z. (1986). Associations among dexamethasone nonsuppression and TRH-induced hormonal responses: Increased specificity for melancholia? *Psychoneuroendocrinology, 11,* 205–211.

Barber, J. P., & Crits-Christoph, P. (Eds.) (1995). *Dynamic therapies for psychiatric disorders (Axis 1).* New York: Basic Books.

Barber, J. P., & DeRubeis, R. J. (1989). On second thought: Where the action is in cognitive therapy for depression. *Cognitive Therapy and Research, 13,* 441–457.

Barnett, P. A., & Gotlib, I. H. (1988). Psychosocial functioning and depression: Distinguishing among antecedents, concomitants, and consequences. *Psychological Bulletin, 104,* 97–126.

Basco, M. R., & Rush, A. J. (1996). *Cognitive-behavioral therapy for bipolar disorder*. New York: Guilford.

Beck, A. T. (1967). *Depression: Clinical, experimental and theoretical aspects*. New York: Harper & Row.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.

Beck, A. T., Brown, G., Berchick, R. J., Stewart, B. L., & Steer, R. A. (1990). Relationship between hopelessness and ultimate suicide: A replication with psychiatric outpatients. *American Journal of Psychiatry, 147,* 190–195.

Beck, A. T., Brown, G. K., Steer, R. A., Dahlsgaard, K. K., & Grisham, J. R. (in press). Suicide ideation at its worst point: A predictor of eventual suicide in psychiatric outpatients. *Suicide and Life-threatening Behavior*.

Beck, A. T., & Emery, G. (1985). *Anxiety disorders and phobias: A cognitive perspective*. New York: Basic Books.

Beck, A. T., Epstein, N., & Harrison, R. (1983). Cognitions, attitudes and personality dimensions in depression. *British Journal of Cognitive Psychotherapy, 1,* 1–16.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford.

Beck, A. T., Steer, R. A., Kovacs, M., & Garrison, B. (1985). Hopelessness and eventual suicide: A 10-year prospective study of patients hospitalized with suicidal ideation. *American Journal of Psychiatry, 142,* 559–563.

Bernasconi, R. (1982). The GABA hypothesis of affective illness: Influence of clinically effective antimanic drugs on GABA turnover. In H. M. Emrich, J. B. Aldenhoff, & H. D. Lux (Eds.), *Basic mechanisms in the action of lithium* (pp. 183–192). Amsterdam: Elsevier.

Bertelsen, A., Harvald, B., & Hauge, M. (1977). A Danish twin study of manic-depressive disorder. *British Journal of Psychiatry, 130,* 330–351.

Bifulco, A., & Brown, G. W. (1996). Cognitive coping response to crises and onset of depression. *Social Psychiatry & Psychiatric Epidemiology, 31,* 163–172.

Borg, S. E., & Stahl, M. (1982). Prediction of suicide: A prospective study of suicides and controls among psychiatric patients. *Acta Psychiatric Scandinavia, 65,* 221–232.

Boyd, J. H., Burke, J. D., Gruenberg, E., Holzer, C. E., III, Rae, D. S., George, L. K., Karno, M., Stoltzman, R., McEvoy, L., & Nestadt, G. (1984). Exclusion criteria of *DSM-III:* A study of co-occurrence of hierarchy-free syndromes. *Archives of General Psychiatry, 41,* 983–989.

Boyd, J. H., & Weissman, M. M. (1981). Epidemiology of affective disorders: A re-examination and future directions. *Archives of General Psychiatry, 38,* 1039–1046.

Brier, A., Kelsoe, J. R., Kirwin, P. D., Beller, S. A., & Wolkowitz, D. P. (1988). Early parental loss and development of adult psychopathology. *Archives of General Psychiatry, 45,* 987–993.

Brown, D. R., Ahmed, F., Gary, L. E., & Milburn, N. G. (1995). Major depression in a community sample of African Americans. *American Journal of Psychiatry, 152,* 373–378.

Brown, G. W. (1993). Life events and affective disorder: Replications and limitations. *Psychosomatic Medicine, 55,* 248–259.

Brown, G. W., Adler, Z., & Bifulco, A. (1988). Life events, difficulties and recovery from chronic depression. *British Journal of Psychiatry, 152,* 487–498.

Brown, G. W., & Harris, T. O. (1978). *Social origins of depression: A study of psychiatric disorder in women.* New York: Free Press.

Brown, G., Harris, T., & Hepworth, C. (1994). Life events and endogenous depression: A puzzle re-examined. *Archives of General Psychiatry, 51,* 525–534.

Burns, D. D. (1980). *Feeling good: The new mood therapy.* New York: Morrow.

Cadoret, R. J. (1978). Evidence for genetic inheritance of primary affective disorder in adoptees. *American Journal of Psychiatry, 135,* 463–466.

Cadoret, R. J., Winokur, G., Langbehn, D., Troughton, E., Yates, W. R., & Stewart, M. A. (1996). Depression spectrum disease. I: The role of gene-environment interaction. *American Journal of Psychiatry, 153,* 892–899.

Cochran, S. (1984). Preventing medical non-compliance in the outpatient treatment of bipolar affective disorder. *Journal of Consulting and Clinical Psychology, 52,* 873–878.

Coryell, W., Keller, M., Endicott, J., Andreasen, N., Clayton, P., & Hirschfeld, R. (1989). Bipolar II illness: Course and outcome over a five-year period. *Psychological Medicine, 19,* 129–141.

Coryell, W., & Winokur, G. (1992). Course and outcome. In E. S. Paykel (Ed.), *Handbook of affective disorders* (2nd ed., pp. 89–108). New York: Guilford.

Coxhead, N., Silverstone, T., & Cookson, J. (1992). Carbamazepine versus lithium in the prophylaxis of bipolar affective disorder. *Acta Psychiatrica Scandinavica, 85,* 114–118.

Cui, X. J., & Vaillant, G. E. (1997). Does depression generate negative life events? *Journal of Nervous and Mental Disease, 185,* 145–150.

Dahlsgaard, K. K., Beck, A. T., & Brown, G. K. (1998). Inadequate response to therapy as a predictor of suicide. *Suicide and Life-threatening Behavior, 28,* 197–204.

Davidson, J. R. T. (1992). Monoamine oxidase inhibitors. In E. S. Paykel (Ed.), *Handbook of affective disorders* (pp. 345–358). New York: Guilford.

DeRubeis, R. J., & Beck, A. T. (1988). Cognitive therapy. In K. S. Dobson (Ed.), *Handbook of cognitive-behavioral therapies* (pp. 273–306). New York: Guilford.

DeRubeis, R. J., Gelfand, L. A., Tang, T. Z., & Simons, A. D. (1998). *The equivalence of cognitive therapy and antidepressant medications for severely depressed outpatients.* Manuscript submitted for publication.

DiMascio, A., Weissman, M. M., Prusoff, B. A., Neu, C., Zwilling, M., & Klerman, G. L. (1979). Differential symptom reduction by drugs and psychotherapy in acute depression. *Archives of General Psychiatry, 36,* 1450–1456.

DiNardo, P. A., & Barlow, D. H. (1990). Syndrome and symptom co-occurrence in the anxiety disorders. In J. D. Maser & C. R. Cloninger (Eds.), *Comorbidity of mood and anxiety disorders* (pp. 205–230). Washington, DC: American Psychiatric Press.

Dobson, K. S. (1989). A meta-analysis of the efficacy of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 57,* 414–419.

Dunner, D. L., & Fieve, R. R. (1974). Clinical factors in lithium carbonate prophylaxis failure. *Archives of General Psychiatry, 30,* 229–233.

Elkin, I., Gibbons, R. D., Shea, M. T., & Sotsky, S. M. (1995). Initial severity and differential treatment outcome in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology, 63,* 841–847.

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., Glass, D. R., Pilkonis, P. A., Leber, W. R., Docherty, J. P., Fiester, S. J., & Parloff, M. B. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Archives of General Psychiatry, 46,* 971–982.

Ellicott, A., Hammen, C., Gitlin, M., Brown, G., & Jamison, K. (1990). Life events and the course of bipolar disorder. *American Journal of Psychiatry, 147,* 1194–1198.

Emilien, G., Maloteaux, J. M., Seghers, A., & Charles, G. (1996). Lithium compared to valproic acid and carbamazepine in the treatment of mania: A statistical meta-analysis. *European Neuropsychopharmacology, 6,* 245–252.

Enns, M., & Reiss, J. (1992). Electroconvulsive therapy. *The Canadian Journal of Psychiatry, 37,* 671–678.

Evans, M. D., Hollon, S. D., DeRubeis, R. J., Piasecki, J. M., Garvey, M. J., Grove, W. M., & Tuason, V. B. (1992). Differential relapse following cognitive therapy, pharmacotherapy, and combined cognitive-pharmacotherapy for depression. *Archives of General Psychiatry, 49,* 802–808.

Fawcett, J., Scheftner, W., Fogg, L., Clark, D. C., Young, M. A., Hedeker, D., & Gibbons, R. (1990). Time-related predictors of suicide in major affective disorder. *American Journal of Psychiatry, 147,* 1189–1194.

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G., & Muñoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry, 26,* 57–63.

Fink, M. (1992). Electroconvulsive therapy. In E. S. Paykel (Ed.), *Handbook of affective disorders* (pp. 359–367). New York: Guilford.

Fogarty, F., Russell, J., Newman, S., & Bland, R. (1994). Mania. *Acta Psychiatrica Scandinavica,* (Suppl. 376), 16–23.

Fombonne, E. (1994). Increased rates of depression: Update of epidemiological findings and analytical problems. *Acta Psychiatrica Scandinavica, 90,* 145–156.

Frank, E., Hlastala, S., Ritenour, A., Houck, P., Tu, X. M., Monk, T. H., Mallinger, A. G., & Kupfer, D. J. (1997). Inducing lifestyle regularity in recovering bipolar disorder patients: Results from the maintenance therapies in bipolar disorder protocol. *Biological Psychiatry, 41,* 1165–1173.

Freud, S. (1917/1957). *Mourning and melancholia.* In J. Strachey (Ed.), *The complete psychological works of Sigmund Freud.* London: Hogarth.

Garber, J., & Flynn, C. (in press). Origins of depressive cognitive style. In D. K. Routh & R. J. DeRubeis (Eds.), *The science of clinical psychology.* Washington, DC: American Psychological Association.

Gerlsma, C., Das, J., & Emmelkamp, P. M. G. (1993). Depressed patients' parental representations: Stability across changes in depressed mood and specificity across diagnoses. *Journal of Affective Disorders, 27,* 173–181.

Gershon, E. S. (1990). Genetics. In F. K. Goodwin & K. F.

Jamison, *Manic-depressive illness* (pp. 373–401). New York: Oxford University Press.

Ghangadhar, B. N., Kapur, R. L., & Kalyanasundaram S. (1982). Comparison of electroconvulsive therapy with imipramine in endogenous depression: A double blind study. *British Journal of Psychiatry, 141,* 367–371.

Glassman, A. (1969). Indoleamines and affective disorder. *Psychosomatic Medicine, 31,* 107–114.

Goldstein, R. B., Black, D. W., Nasrallah, A., & Winokur, G. (1991). The prediction of suicide: Sensitivity, specificity, and predictive value of a multivariate model applied to suicide among 1906 patients with affective disorders. *Archives of General Psychiatry, 48,* 418–422.

Goodwin, G. M. (1992). Tricyclic and newer antidepressants. In E. S. Paykel (Ed.), *Handbook of affective disorders* (pp. 327–343). New York: Guilford.

Goodwin, F. K., & Jamison, K. R. (1990). *Manic depressive illness*. New York: Oxford University Press.

Goodwin, F. K. & Sack, R. L. (1974). Central dopamine function in affective illness: Evidence from precursors, enzyme inhibitors, and studies of central dopamine turnover. In E. Usdin (Ed.), *The neuropsychopharmacology of monoamines and their regulatory enzymes* (pp. 261–279). New York: Raven Press.

Goodwin, F. K., & Zis, A. P. (1979). Lithium in the treatment of mania: Comparisons with neuroleptics. *Archives of General Psychiatry, 36,* 840–844.

Goodyer, I. M., Herbert, J., Tamplin, A., Secher, S. M., & Pearson, J. (1997). Short term outcome of major depression: II. Life events, family dysfunction, and friendship difficulties as predictors of persistent disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 474–480.

Greenberg, R. P., Bornstein, R. F., Zborowski, M. J., Fisher, S., & Greenberg, M. A. (1994). A meta-analysis of fluoxetine outcome in the treatment of depression. *Journal of Nervous and Mental Disease, 182,* 547–551.

Gregory, S., Shawcross, C. R., & Gill, D. (1985). The Nottingham ECT study: A double-blind comparison of bilateral, unilateral and simulated ECT in depressive illness. *British Journal of Psychiatry 146,* 520–524.

Grove, W. M., & Andreasen, N. C. (1992). Concepts, diagnosis, and classification. In E. S. Paykel (Ed.), *Handbook of affective disorders* (2nd ed., pp. 25–41). New York: Guilford.

Guze, B. H., & Gitlin, M. (1994). New antidepressants and the treatment of depression. *Journal of Family Practice, 38,* 49–57.

Guze, S. B., & Robins, E. (1970). Suicide and primary affective disorders. *British Journal of Psychiatry, 19,* 437–448.

Gwirtsman, H. E. (1994). Dysthymia and chronic depressive states: Diagnostic and pharmacotherapeutic considerations. *Psychopharmacology Bulletin, 30,* 45–51.

Haaga, D. A., Dyck, M., & Ernst, D. (1991). Empirical status of cognitive theory of depression. *Psychological Bulletin, 110,* 215–236.

Halaris, A. E. (1978). Plasma 3-methoxy-4-hydroxyphenylglycol in manic psychosis. *American Journal of Psychiatry, 135,* 493–494.

Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology, 32,* 50–55.

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology, 6,* 276–296.

Hamilton, M. S. (1983). The clinical distinction between anxiety and depression. *British Journal of Clinical Pharmacology, 15*(Suppl. 2), 165S–169S.

Hammen, C. (1991). Generation of stress in the course of unipolar depression. *Journal of Abnormal Psychology, 100,* 555–561.

Hammen, C., Ellicott, A., Gitlin, M., & Jamison, K. (1989). Sociotropy/autonomy and vulnerability to specific life events in patients with unipolar depression and bipolar disorders. *Journal of Abnormal Psychology, 98,* 154–160.

Hammen, C., & Gitlin, M. (1997). Stress reactivity in bipolar patients and its relation to prior history of disorder. *American Journal of Psychiatry, 154,* 856–857.

Harris, T., Brown, G. W., & Bifulco, A. (1986). Loss of parent in childhood and adult psychiatric disorder: The role of lack of adequate parental care. *Psychological Medicine, 16,* 641–659.

Haslam, N., & Beck, A. T. (1994). Subtyping major depression: A taxometric analysis. *Journal of Abnormal Psychology, 103,* 686–692.

Hawton, K. (1992). Suicide and attempted suicide. In E. S. Paykel (Ed.), *Handbook of affective disorders* (2nd ed., pp. 635–650). New York: Guilford.

Hedlund, S., & Rude, S. S. (1995). Evidence of latent depressive schemas in formerly depressed individuals. *Journal of Abnormal Psychology, 104,* 517–525.

Heninger, G., Delgado, P., & Charney, D. (1996). The revised monoamine theory of depression: A modulatory role for monoamines based on new findings from monoamine depletion experiments in humans. *Pharmacopsychiatry, 29,* 2–11.

Hickie, I., Wilhelm, K., Parker, G., Boyce, P., Hadzi-Pavlovic, D., Brodati, H., & Mitchell, P. (1990). Perceived dysfunctional intimate relationships: A specific association with the non-melancholic depressive type. *Journal of Affective Disorders, 19,* 99–107.

Hollon, S. D., DeRubeis, R. J., Evans, M. D., Wiemer, M. J., Garvey, M. J., Grove, W. M., & Tuason, V. B. (1992). Cognitive therapy and pharmacotherapy for depression: Singly and in combination. *Archives of General Psychiatry, 49,* 774–781.

Hollon, S., & Kriss, M. R. (1984). Cognitive factors in clinical research and practice. *Clinical Psychology Review, 4,* 35–76.

Hollon, S. D. Shelton, R. C., & Loosen, P. T. (1991). Cognitive therapy and pharmacotherapy for depression. *Journal of Consulting and Clinical Psychology, 59,* 88–99.

Hunt, N., Bruce-Jones, W., & Silverstone, T. (1992). Life events and relapse in bipolar affective disorder. *Journal of Affective Disorders, 25,* 13–20.

Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., Gortner, E., & Prince, S. E. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 64,* 295–304.

Janowsky, D. S., El-Yousef, M. K., Davis, J. M., & Sekerke, H. J. (1972). A cholinergic-adrenergic hypothesis of mania and depression. *Lancet, 2,* 632–635.

Johnson, S. L., & Miller, I. (1997). Negative life events and time to recovery from episodes of bipolar disorder. *Journal of Abnormal Psychology, 106,* 449–457.

Johnson, S. L., & Roberts, J. (1995). Life events and bipolar disorder: Implications from biological theories. *Psychological Bulletin, 117,* 434–449.

Kantzian, E. J. (1985). The self-medication hypothesis of addictive disorders: Focus on heroin and cocaine dependence. *American Journal of Psychiatry, 142,* 1259–1264.

Keller, M. B., Lavori, P. W., Coryell, W., Andreasen, N. C., Endicott, J., Clayton, P. J., Klerman, G. L., & Hirschfeld, R. M. A. (1986). Differential outcome of pure manic, mixed/cycling, and pure depressive episodes in patients with bipolar illness. *Journal of the American Medical Association, 255,* 3138–3142.

Keller, M. B., Shapiro, R. W., Lavori, P. W., & Wolfe, N. (1982). Relapse in major depressive disorder. Analysis with the life table. *Archives of General Psychiatry, 39,* 911–915.

Kendler, K. S., & Karkowski-Shuman, L. (1997). Stressful life events and genetic liability to major depression:

Genetic control of exposure to the environment? *Psychological Medicine, 27,* 539–547.

Kennedy, S., Thompson, R., Stancer, H. C., Roy, A., & Persad, E. (1983). Life events precipitating mania. *British Journal of Psychiatry, 142,* 398–403.

Klerman, G. L., & Weissman, M. M. (1989). Increasing rates of depression. *Journal of the American Medical Association, 261,* 2229–2235.

Klerman, G. L., Weissman, M. M., Rounsaville, B. J., & Chevron, E. S. (1984). *Interpersonal psychotherapy of depression.* New York: Basic Books.

Kraepelin, E. (1921). *Manic-depressive insanity and paranoia.* Edinburgh, UK: Livingstone.

Lam, R. W., & Stewart, J. N. (1996). The validity of atypical depression in *DSM-IV. Comprehensive Psychiatry, 37,* 375–383.

Lester, D. (1995). The concentration of neurotransmitter metabolites in the cerebrospinal fluid of suicidal individuals: A meta-analysis. *Pharmacopsychiatry, 28,* 45–50.

Lewinsohn, P. M. (1974). A behavioral approach to depression. In R. J. Friedman & M. M. Katz (Eds.). *The psychology of depression: Contemporary theory and research* (pp. 176–178). Washington, DC: Winston-Wiley.

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder.* New York: Guilford.

Linehan, M. M., Tutek, D. A., Heard, H. L., & Armstrong, H. E. (1994). Interpersonal outcome of cognitive behavioral treatment for chronically suicidal borderline patients. *American Journal of Psychiatry, 151,* 1771–1776.

Maj, M., Ariano, M. G., Arena, F., & Kemali, D. (1984). Plasma cortisol, catecholamine and cyclic AMP levels, response to dexamethasone suppression test and platelet MAO activity in manic-depressive patients: A longitudinal study. *Neuropsychobiology, 11,* 168–173.

Mann, J. J., & Malone, K. M. (1997). Cerebrospinal fluid amines and higher lethality suicide attempts in depressed inpatients. *Biological Psychiatry, 41,* 162–171.

Martin, R. L., Cloninger, C. R., Guze, S. B., & Clayton, P. J. (1985). Mortality in a follow-up of 500 psychiatric outpatients. *Archives of General Psychiatry, 42,* 58–66.

McElroy, S. L., Keck, P. E., Pope, H. G., & Hudson, J. I. (1992). Valproate in the treatment of bipolar disorder: Literature review and clinical guidelines. *Journal of Clinical Psychopharmacology, 12*(Suppl. 1), 42–52.

McGuffin, P., & Katz, R. (1989). The genetics of depression and manic-depressive disorder. *British Journal of Psychiatry, 155,* 294–304.

McLean, P. D., & Hakstian, A. R. (1979). Clinical depression: Comparative efficacy of outpatient treatments. *Journal of Consulting and Clinical Psychology, 47,* 818–836.

McNeal, E., & Cimbolic, P. (1986). Antidepressants and biochemical theories of depression. *Psychological Bulletin, 99,* 361–374.

Meesters, Y., Jansen, J. H. C., Lambers, P. A., Bouhuys, A. L., Beersma, D. G. M., & van den Hoofdakker, R. H. (1993). Morning and evening light treatment of seasonal affective disorder: Response, relapse and prediction. *Journal of Affective Disorders, 28,* 165–177.

Mendlewicz, J., & Rainer, J. D. (1977). Adoption study supporting genetic transmission in manic-depressive illness. *Nature, 368,* 327–329.

Merikangas, K. R. (1990). Comorbidity for anxiety and depression: Review of family and genetic studies. In J. D. Maser & C. R. Cloninger (Eds.), *Comorbidity of mood and anxiety disorders* (pp. 331–348). Washington, DC: American Psychiatric Press.

Merikangas, K. R., & Gelernter, C. S. (1990). Comorbidity for alcoholism and depression. *Psychiatric Clinics of North America, 13,* 613–632.

Merikangas, K. R., Risch, N. J., & Weissman, M. M. (1994). Comorbidity and co-transmission of alcoholism, anxiety, and depression. *Psychological Medicine, 24,* 69–80.

Meyer, A. (1957). *Psychobiology: A science of man.* Springfield, IL: Thomas.

Miklowitz, D. J., Frank, E., & George, E. L. (1996). New psychosocial treatments for the outpatient management of bipolar disorder. *Psychopharmacology Bulletin, 32,* 613–621.

Miklowitz, D. J., & Goldstein, M. J. (1990). Behavioral family treatments for patients with bipolar affective disorder. *Behavior Modification, 14,* 457–489.

Miranda, J., & Persons, J. B. (1988). Dysfunctional attitudes are mood-state dependent. *Journal of Abnormal Psychology, 97,* 76–79.

Moldin, S., Reich, T., & Rice, J. (1991). Current perspectives on the genetics of unipolar depression. *Behavior Genetics, 21,* 211–242.

Molin, J., Mellerup, E., Bolwig, T., Scheike, T., & Dam, H. (1996). The influence of climate on development of winter depression. *Journal of Affective Disorders, 37,* 151–155.

Murphy, G. E., Simons, A. D., Wetzel, R. D., & Lustman, P. J. (1984). Cognitive therapy and pharmacotherapy: Singly and together in the treatment of depression. *Archives of General Psychiatry, 41,* 33–41.

Musselman, D., & Nemeroff, C. (1996). Depression and endocrine disorders: Focus on the thyroid and adrenal system. *British Journal of Psychiatry, 168*(Suppl. 30), 123–128.

Neale, M. C., Walters, E., Heath, A. C., Kessler, R. C., Perusse, D., Eaves, L. J., & Kendler, K. S. (1994). Depression and parental bonding: Cause, consequence, or genetic covariance? *Genetic Epidemiology, 11,* 503–522.

Newman, C., & Beck, A. T. (1992). *Cognitive therapy in rapid cycling disorder: A treatment manual.* Unpublished manuscript, Center for Cognitive Therapy, University of Pennsylvania, Philadelphia, PA.

Nezu, A. M. (1986). Efficacy of a social problem-solving therapy approach for unipolar depression. *Journal of Consulting and Clinical Psychology, 54,* 196–202.

Nezu, A. M., & Perri, M. G. (1989). Social problem-solving therapy for unipolar depression: An initial dismantling investigation. *Journal of Consulting and Clinical Psychology, 57,* 408–413.

Nolen-Hoeksema, S. (1987). Sex differences in unipolar depression: Evidence and theory. *Psychological Bulletin, 101,* 259–282.

Nordström, P., Samuelsson, M., Åsberg, M., Träskman-Bendz, L., Åsberg-Wisted, A., Nordin, C., & Bertilsson, L. (1994). CSF 5-HIAA predicts suicide risk after attempted suicide. *Suicide and Life-threatening Behavior, 24,* 1–9.

Nurnberger, J. I., Jr., & Gershon, E. S. (1992). Genetics. In E. S. Paykel (Ed.), *Handbook of affective disorders* (2nd ed., pp. 131–148). New York: Guilford.

Parker, G. (1992). Early environment. In E. S. Paykel (Ed.), *Handbook of affective disorders* (2nd ed., pp. 171–183). New York: Guilford.

Parker, G., & Hadzi-Pavlovic, D. (1992). Parental representations of melancholic and non-melancholic depressives: Examining for specificity to depressive type and for evidence of additive effects. *Psychological Medicine, 22,* 657–665.

Parker, G., Hadzi-Pavlovic, D., Greenwald, S., & Weissman, M. (1995). Low parental care as a risk factor to lifetime depression in a community sample. *Journal of Affective Disorders, 33,* 173–180.

Parker, G., Kiloh, L., Hayward, L. (1987). Parental representations of neurotic and endogenous depressives. *Journal of Affective Disorders, 13,* 75–82.

Parker, G., Tupling, H., & Brown, L. B. (1979). A parental bonding instrument. *British Journal of Medical Psychology, 52,* 1–10.

Paykel, E. S. (1994). Life events, social support and depression. *Acta Psychiatrica Scandinavica*, (Suppl. 377), 50–58.

Paykel, E. S., Hollyman, J. A., Freeling, P., & Sedgwick, P. (1988). Predictors of therapeutic benefit from amitriptyline in mild depression: A general practice placebo-controlled trial. *Journal of Affective Disorders, 14,* 83–95.

Perris, C. (1988). A theoretical framework for linking the experience of dysfunctional parental rearing attitudes with manifest psychopathology. *Acta Psychiatrica Scandinavica*, (Suppl. 377), 93–100.

Perris, C. (1992). Bipolar–unipolar distinction. In E. S. Paykel (Ed.), *Handbook of affective disorders* (2nd ed., pp. 57–75). New York: Guilford.

Persons, J. B. (1989). *Cognitive therapy in practice: A case formulation approach*. New York: Norton.

Pokorny, A. D. (1983). Prediction of suicide in psychiatric patients. *Archives of General Psychiatry, 40,* 249–257.

Post, R. M., Rubinow, D. R., & Ballenger, J. C. (1986). Conditioning and sensitization in the longitudinal course of affective illness. *British Journal of Psychiatry, 149,* 191–201.

Post, R. M., Rubinow, D. R., Uhde, T. W., Roy-Byrne, P. P., Linnoila, M., Rosoff, A., & Cowdry, R. (1989). Dysphoric mania: Clinical and biological correlates. *Archives of General Psychiatry, 46,* 353–358.

Prange, A. J., Jr., Wilson, A. C., Lynn, C. W., Alltop, L. B., & Stikeleather, R. A. (1974). L-Tryptophan in mania: Contribution to a permissive hypothesis of affective disorders. *Archives of General Psychiatry, 30,* 56–62.

Prien, R. F., Caffey, E. M., Jr., & Klett, J. (1974). Factors associated with treatment prophylaxis: Report of the Veterans Administration and National Institute of Mental Health Collaborative Study Group. *Archives of General Psychiatry, 31,* 189–192.

Quitkin, F. M., Stewart, J. W., McGrath, P. J., Tricamo, E., Rabkin, J. G., Ocepek-Welikson, K., Nunes E., Harrison, W., & Klein, D. F. (1993). Columbia atypical depression: A subgroup of depressives with better response to MAOI than to tricyclic antidepressants. *British Journal of Psychiatry, 163*(Suppl. 21), 30–34.

Regier, D. A., Farmer, M. E., Rae, D. S., Locke, B. Z., Keith, S. J., Judd, L. L., & Goodwin, F. K. (1990). Comorbidity of mental disorders with alcohol and other drug abuse. *Journal of the American Medical Association, 264,* 2511–2518.

Richelson, E. (1991). Biological basis of depression and therapeutic relevance. *Journal of Clinical Psychiatry, 52,* 4–10.

Robins, L. N., Helzer, J. E., Ratcliff, K. S., & Seyfried, W. (1982). Validity of the Diagnostic Interview Schedule, Version II: *DSM-III* diagnosis. *Psychological Medicine, 12,* 855–870.

Rubinow, D., Post, R., & Savard, R. (1984). Cortisol hypersecretion and cognitive impairment in depression. *Archives of General Psychiatry, 41,* 279–285.

Rush, A. J., Beck, A. T., Kovacs, M., & Hollon, S. (1977). Comparative efficacy of cognitive therapy and pharmacotherapy in the treatment of depressed patients. *Cognitive Therapy and Research, 1,* 17–37.

Sackeim, H. A. (1989). Mechanism of action. *Convulsive Therapy, 6,* 207–310.

Sato, T., Sakado, K., Uehara, T., Nishioka, K., & Kasahara, Y. (1997). Perceived parental styles in a Japanese sample of depressive disorders: A replication outside Western culture. *British Journal of Psychiatry, 170,* 173–175.

Schildkraut, J. J. (1965). The catecholamine hypothesis of affective disorders: A review of supporting evidence. *American Journal of Psychiatry, 122,* 509–522.

Schulman, P., Keith, D., & Seligman, M. E. P. (1993). Is optimism heritable? A study of twins. *Behaviour Research and Therapy, 31,* 569–574.

Shapiro, D. A., Barkham, M., Rees, A., Hardy, G., Reynolds, S., & Startup, M. (1994). Effects of treatment duration and severity of depression on the effectiveness of cognitive-behavioral and psychodynamic-interpersonal psychotherapy. *Journal of Consulting and Clinical Psychology, 62,* 552–534.

Shaw, B. F. (1977). Comparison of cognitive therapy and behavior therapy in the treatment of depression. *Journal of Consulting and Clinical Psychology, 45,* 543–551.

Silverstone, T., & Hunt, N. (1992). Symptoms and assessment of mania. In E. S. Paykel (Ed.), *Handbook of affective disorders* (2nd ed., pp. 15–24). New York: Guilford.

Soares, J., & Mann, J. (1997). The anatomy of mood disorders: Review of structural neuroimaging studies. *Biological Psychiatry, 41,* 86–106.

Sonne, S. C., Brady, K. T., & Morton, W. A. (1994). Substance abuse and bipolar affective disorder. *Journal of Nervous and Mental Disease, 182,* 349–352.

Spaner, D., Bland, R., & Newman, S. (1994). Major depressive disorder. *Acta Psychiatrica Scandinavica*, (Suppl. 376), 7–15.

Stewart, J. W., Quitkin, F. M., Liebowitz, M. R., McGrath, P. J., Harrison, W. M., & Klein, D. F. (1983). Efficacy of desipramine in depressed outpatients: Response according to research diagnostic criteria and severity of illness. *Archives of General Psychiatry, 40,* 202–207.

Strömgren, L. S., & Boller, S. (1985). Carbamazepine in treatment and prophylaxis of manic-depressive disorder. *Psychiatric Developments 3,* 349–367.

Sullivan, H. S. (1953). *The interpersonal theory of psychiatry*. New York: Norton.

Swann, A. C., Koslow, S. H., Katz, M. M., Maas, J. W., Javaid, J., Secunda, S. K., & Robins, E. (1987). Lithium carbonate treatment of mania: Cerebrospinal fluid and urinary monoamine metabolites and treatment outcome. *Archives of General Psychiatry, 44,* 345–354.

Tam, E., Lam, R., & Levitt, A. (1995). Treatment of seasonal affective disorder: A review. *Canadian Journal of Psychiatry, 40,* 457–466.

Tam, E. M., Lam, R. W., Robertson, H. A., Stewart, J. N., Yatham, L. N., & Zis, A. P. (1997). Atypical depressive symptoms in seasonal and non-seasonal mood disorders. *Journal of Affective Disorders, 44,* 39–44.

Tennant, C. (1988). Parental loss in childhood: Its effect in adult life. *Archives of General Psychiatry, 45,* 1045–1050.

Terman, J. S., Terman, M., Schlager, D., Rafferty, B., Rosofsky, M., Link, M. J., Gallin, P. F., & Quitkin, F. M. (1990). Efficacy of brief, intense light exposure for treatment of winter depression. *Psychopharmacology Bulletin, 26,* 3–11.

Thase, M., Dube, S., Bowler, K., Howland, R., Myers, J., Friedman, E., & Jarrett, D. (1996). Hypothalamic–pituitary–adrenocortical activity and response to cognitive behavior therapy in unmedicated, hospitalized depressed patients. *American Journal of Psychiatry, 153,* 886–891.

von Knorring, A., Cloninger, R., Bohman, M., & Sigvardsson, S. (1983). An adoption study of depressive disorders and substance abuse. *Archives of General Psychiatry, 40,* 943–950.

Weiss, R. D., & Mirin, S. M. (1987). Substance abuse as an attempt at self medication. *Psychiatry Medicine, 3,* 357–367.

Weissman, M. M., & Klerman, G. L. (1977). Sex differences and the epidemiology of depression. *Archives of General Psychiatry, 34,* 98–111.

Weissman, M. M., & Klerman, G. L. (1985). Gender and depression. *Trends in the Neurosciences, 8,* 416–420.

Weissman, M. M., Leaf, P. J., Tischler, G. L., Blazer, D.

G., Karno, M., Bruce, M. L., & Florio, L. P. (1988). Affective disorders in five United States communities. *Psychological Medicine, 18,* 141–153.

Wender, P. H., Kety, S. S., Rosenthal, D., Schulsinger, F., Ortmann, J., & Lunde, I. (1986). Psychiatric disorders in the biological and adoptive families of adopted individuals with affective disorders. *Archives of General Psychiatry, 43,* 923–929.

Wertham, F. I. (1928). A group of benign psychoses: Prolonged manic excitements. *American Journal of Psychiatry, 9,* 17–78.

Whybrow, P., & Bahr, R. (1988). *The hibernation response.* New York: Arbor House/Morrow.

Wing, J. K., Cooper, J. E., & Sartorius, N. (1974). *Measurement and classification of psychiatric symptoms: An instructional manual for the PSE and CATEGO program.* New York: Cambridge University Press.

Winokur, G., Cadoret, R., Dorzab, J., & Baker, M. (1971). Depressive disease: A genetic study. *Archives of General Psychiatry, 24,* 135–144.

Young, M. A., Fogg, L. F., Scheftner, W., Fawcett, J., Akiskal, H., & Maser, J. (1996). Stable trait components of hopelessness: Baseline and sensitivity to depression. *Journal of Abnormal Psychology, 105,* 155–165.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.17
Obsessive-compulsive Disorder

GAIL S. STEKETEE
Boston University, MA, USA
and
RANDY O. FROST
Smith College, Northampton, MA, USA

6.17.1 INTRODUCTION 368
6.17.2 SYMPTOMS AND CHARACTERISTICS OF OCD 368
  6.17.2.1 Symptoms and Diagnosis of OCD 368
  6.17.2.2 Differential Diagnosis 369
6.17.3 EPIDEMIOLOGY 369
  6.17.3.1 Sex Ratio and Age of Onset 369
  6.17.3.2 Clinical Course 369
  6.17.3.3 Demographic and Cultural Variables 370
6.17.4 COMORBIDITY AND RELATED DISORDERS 370
  6.17.4.1 Axis I Disorders 370
  6.17.4.2 Axis II Disorders 372
6.17.5 OCD PATIENT CHARACTERISTICS 372
  6.17.5.1 Family Factors in OCD 372
  6.17.5.2 Personality and Cognitive Characteristics 373
    6.17.5.2.1 Overestimation of risk 374
    6.17.5.2.2 Responsibility 374
    6.17.5.2.3 Controllability of thoughts 374
    6.17.5.2.4 Overimportance of thoughts 374
    6.17.5.2.5 Tolerance for ambiguity 374
  6.17.5.3 Information/Cognitive Processing in OCD 375
6.17.6 THEORETICAL MODELS AND TREATMENTS FOR OCD 376
  6.17.6.1 Behavioral Models for OCD 376
  6.17.6.2 Behavioral Treatments for OCD 377
    6.17.6.2.1 Degree of improvement from ERP 377
    6.17.6.2.2 Specific effects of exposure and response prevention 378
    6.17.6.2.3 Processes during exposure 378
    6.17.6.2.4 Problems in the treatment of pure obsessive symptoms 378
    6.17.6.2.5 Variants in delivery of ERP 381
    6.17.6.2.6 Group treatment for OCD 381
    6.17.6.2.7 Family treatment for OCD 382
  6.17.6.3 Cognitive Theories of OCD 383
  6.17.6.4 Cognitive Treatments for OCD 384
  6.17.6.5 Biological Models of OCD 385
    6.17.6.5.1 The role of serotonin 385
    6.17.6.5.2 Neuroanatomy of OCD 386
    6.17.6.5.3 Genetics 387
  6.17.6.6 Biological Treatments for OCD: Medications and Psychosurgery 387

6.17.6.6.1 Medications 387
6.17.6.6.2 Psychosurgery 389
6.17.6.7 Combined Interventions for Treatment of OCD 389
6.17.7 FUTURE DIRECTIONS 390
6.17.8 REFERENCES 391

6.17.1 INTRODUCTION

Since the mid 1960s, many researchers have invested considerable effort in defining and characterizing obsessive-compulsive disorder (OCD) and its myriad symptoms, and, perhaps most importantly, in developing effective treatments. The results of these efforts by researchers, clinicians, patients, families, and their advocates, are that OCD is now a well-recognized condition, with effective and available pharmacological and behavioral treatments. In the late 1990s, newer forms of medications and of cognitive and behavioral interventions are under study. Consequently, it is likely that our knowledge of OCD and how to treat it effectively will continue to expand for some time. Below we present established and recent findings in the study of OCD.

6.17.2 SYMPTOMS AND CHARACTERISTICS OF OCD

6.17.2.1 Symptoms and Diagnosis of OCD

According to the *Diagnostic and statistical manual of mental disorders* (4th ed., *DSM-IV*) (American Psychiatric Association [APA], 1994), obsessions are recurring persistent thoughts, images, or impulses that are experienced at least at some point as intrusive, unreasonable, and distressing. The individual attempts to ignore or suppress these ideas or to "neutralize" them using some ritualistic thought or action. Compulsions are repeated behaviors or mental acts, rigidly or excessively applied to relieve anxiety provoked by the obsessions. Often, these rituals are designed to restore safety or to prevent a dreaded event (Rachman, 1976a), an intention that is often deliberately hidden from casual observers and even close family members. These symptoms interfere with the person's normal activities and consume at least an hour a day. As for many anxiety and affective disorders, in the general population, the symptoms of OCD are relatively common, but usually fail to meet the severity and interference criteria that characterize the disorder. It is interesting to note that the content of clinical obsessive intrusions differs very little from ordinary intrusive thoughts exhibited by most people, but the former provoke considerably more anxiety and are more difficult to dismiss than the latter (Dent & Salkovskis, 1986; Rachman & DeSilva, 1978; Salkovskis & Harrison, 1984).

Unlike most other Axis I disorders, the symptoms of OCD vary considerably from one patient to the next, sharing only the disturbing nature of the intrusions and the ritualistic efforts to neutralize the obsessions. Many patients report having more than one type of obsession and ritual, often apparently unrelated to each other, although on closer inspection they may share some meaningful thematic content. Most obsessions and compulsions can be grouped into categories, somewhat more easily by type of ritual than type of obsession because of the many variations in the latter category. Most common are checking rituals, which are usually designed to prevent a particular catastrophe like fire, burglary, causing someone harm, embarrassing oneself, or being rejected. Washing rituals serve to remove "contamination" or "dirt" caused by such sources as human or animal body waste, chemicals, or other substances. In many cases washing is thought to prevent a feared disaster such as disease or debilitation from occurring. Repeating compulsions are "magical" rituals in which ordinary actions (such as crossing a threshold or lifting an object) are repeated to prevent harm from occurring (for example, a loved one dying in an accident). Ordering rituals are less common and involve arranging objects to produce symmetry or balance, thereby relieving a feeling of discomfort. Individuals may hoard or collect seemingly useless objects like old magazines and receipts or empty containers to prevent the loss of potentially important information or to avoid waste.

Although subtypes based on obsessive content have been proposed, these have not been widely adopted, probably because they have not been particularly useful in predicting differential treatment response. Instead, obsessive phenomena have been merely described. Goodman et al. (1989) developed a useful Symptom Checklist for this purpose. Their categories include aggressive or harming obsessions, contamination, sexual content, hoarding/saving, religious obsessions, somatic concerns, and a variety of miscellaneous fears about mental images, numbers, and knowing or saying things, among others. From a clinical standpoint, it may be useful to determine the patient's complete list of obsessive and compulsive symptoms, since these can fluctuate over time and may be differentially affected by cognitive and behavioral treatments.

An added component of recent diagnostic criteria for OCD is determination of "poor insight" type, for those who do not recognize the unreasonableness of their behavior (APA, 1994). Historically, OCD was at least partly defined by patients' recognition of the irrationality of obsessions, with those who were convinced of their obsessions considered psychotic or at least "overvalued ideators" (Foa, 1979; Insel & Akiskal, 1986). More recent study indicates that patients range on a continuum from complete awareness (insight) to delusional, with most having at least some insight (e.g., Kozak & Foa, 1994).

6.17.2.2 Differential Diagnosis

According to the *DSM-IV*, OCD must be distinguished from several other conditions that share similar features. For example, obsessions are defined as mental experiences that are not merely excessive worries about real-life problems (e.g., finances, family well-being, health, etc.) that are characteristic of generalized anxiety disorder (GAD) or hypochondriasis. Likewise, obsessions must be different in character from the guilty or depressive ruminations of those with major depression. Preoccupations with food, alcohol, drugs, gambling, and buying are sometimes colloquially termed "obsession" or "compulsion," but are readily distinguishable from obsessive fears of contamination, causing harm, preventing danger, and other anxiety-provoking concerns and rituals. Such appetitive preoccupations are discussed further below, along with other impulsive symptoms and disorders sometimes characterized as "OC Spectrum" disorders.

6.17.3 EPIDEMIOLOGY

Prior to publication of the first Epidemiologic Catchment Area (ECA) study (Karno, Golding, Sorenson, & Burnam, 1988; Robins et al., 1984), the prevalence rates for OCD were believed to be quite low, and were estimated only from rates among psychiatric patient populations. The ECA study was the first to estimate prevalence rates for OCD from community-based samples and using *DSM-III* criteria. Annual prevalence rates ranged from 0.8 to 2.3%. Lifetime prevalence rates ranged from 1.9 to 3.3%, considerably higher than previously thought (Karno et al., 1988; Robins et al.). These rates correspond to similar annual and lifetime prevalence rates reported in Canada (Kolada, Bland, & Newman, 1994), Finland (Vaisaner, 1975), Taiwan (Hwu, Yeh, & Chang, 1989), and Africa (Orly & Wing, 1979). Furthermore, a recent cross-national epidemiological study of OCD (Weissman et al., 1994) reported remarkably similar findings across seven countries (USA, Canada, Puerto Rico, Germany, Taiwan, Korea, & New Zealand), where annual prevalence rates ranged from 1.1 to 1.8, and lifetime prevalence ranged from 1.9 to 2.5.

6.17.3.1 Sex Ratio and Age of Onset

Although early studies found no differences in the frequency of OCD in men and women (Black, 1974; Karno et al., 1988), later ones have shown a slight preponderance of women among OCD patients (Rasmussen & Eisen, 1990; Weissman et al., 1994). A number of studies have noted a preponderance of males with early onset OCD (Bellodi, Sciuto, Diaferia, Ronchi, & Smeraldi, 1992; Flament et al., 1988; Rapoport, 1989; Rasmussen & Eisen, 1990). The mean onset for males ranged from 14 to 19.5 and for females 21.7 to 22 (Bellodi et al., 1992; Rasmussen & Eisen, 1990, 1992a). Therefore during childhood, boys with OCD outnumber girls by up to 2 to 1.

Sixty-five percent of OCD patients developed symptoms prior to age 25 while only 15% developed symptoms after age 35 (Rasmussen & Tsuang, 1986). Thus, the average age of onset appears to be in the late teens or early twenties (Tallis, 1995). The latency from onset to help seeking is estimated at seven to eight years (Yaryura-Tobias & Neziroglu, 1983). Onset of symptoms can be either acute or insidious (Kolada et al., 1994). Some investigators have identified precipitants related to changes in life roles or demands (e.g., pregnancy, childbirth, etc.), with symptoms often focusing on content related to the precipitant (see Diaz, Grush, Sichel, & Cohen , in press). However, at least 40% of OCD patients showed no clear-cut precipitant (Kolada et al., 1994).

6.17.3.2 Clinical Course

In a review of early follow-up studies, Goodwin, Guze, and Robins (1969) concluded that there were three types of clinical course in OCD: a chronic, unremitting course in which 10–15% of patients show deterioration, an episodic course with periods of complete remission, and an episodic course with periods of incomplete remission. Eisen and Steketee's (1997) recent review confirmed some of their conclusions: episodic OCD with full remissions occurs rarely (10–15%), although this percentage increased with longer follow-up periods. Two recent prospective studies of clinical course assessed OCD symptoms using a Psychiatric Rating Scale that allowed determination of whether subjects met diagnostic criteria for OCD, had subclinical symptoms, or remitted entirely. Eisen, Rasmussen, Goodman, et al. (1995) observed that 57% of 51 subjects still met full criteria after two years, although some of these had improved substantially. Of the remainder, 12% had minimal or no symptoms and 31% were partially improved and did not meet diagnostic criteria. Once remitted, 48% later relapsed. A second study of 107 clinic patients followed up to five years after intake showed a slightly better outlook, probably because of a longer treatment follow-up (Steketee, Eisen, Dyck, Warshaw, & Rasmussen, 1996). These patients showed a 22% probability of full remission and 53% for partial remission. Still, as in retrospective studies, more than three-quarters remained symptomatic and at least partially impaired by their symptoms.

A number of unsystematic observations about the clinical course are worthy of note and further study. It has been observed that OCD symptoms can switch (for example, from cleaning to checking or vice versa), but no systematic research has examined this phenomenon (Rasmussen & Eisen, 1992a; Tallis, 1995). Certain features such as mild level of symptoms, atypical symptoms, short duration, and good premorbid personality indicate a favorable prognosis (Kolada et al., 1994). Features suggestive of poor prognosis include being male with early onset of symptoms, symmetry or exactness-related symptoms, symptoms of hopelessness, delusions or hallucinations, family history of OCD, and presence of tics (Eisen & Steketee, 1997; Kolada et al.).

6.17.3.3 Demographic and Cultural Variables

Most studies have not found differences in prevalence of OCD based on race (Hispanic, African-American, whites), religion, or socioeconomic status (Burnam et al., 1987; Karno et al., 1988; Valleni-Basile et al., 1994). However, Karno et al. found OCD to be more common in young, divorced, separated, or unemployed subjects and slightly less common among blacks than whites. Although findings are somewhat conflicting, most studies have concluded that patients with OCD had higher than average levels of intelligence.

Despite approximately equivalent prevalence rates, clients from minority groups have rarely presented for treatment in mental health settings, and in 1991 Neal and Turner noted that no studies had yet examined treatment for OCD in African-Americans. According to recent research, blacks with washing rituals seek assistance more often from medical settings such as dermatology clinics (Friedman, Hatch, Paradis, Popkin, & Shalita, 1995) and may have more severe symptom presentations (Chambless & Williams, 1995). Symptoms of OCD appear to be recognizable across many disparate cultures. The tendency for some people with OCD to show changing symptoms (e.g., checking to cleaning), the wide variety and cultural specificity of symptom expression (e.g., fears of penile shrinkage in China, fears of leprosy in Africa), and the shifts in the content of obsessional fears over time (e.g., from fears of cancer to AIDS-related fears) suggests that OCD is at least partly influenced by cultural factors (e.g., Weissman et al., 1994). Dulaney and Fiske (1994) noted the similarities between OCD rituals and religious or cultural rituals, while other investigators (Greenburg & Witztum, 1994) have found certain religious practices to be associated with symptom expression in OCD (Okasha, Saad, Khalil, Dawla, & Seif, 1994). Thus, there is little evidence that culture is important in the development of OCD, but it may well influence the way the disorder is expressed.

6.17.4 COMORBIDITY AND RELATED DISORDERS

6.17.4.1 Axis I Disorders

Rasmussen and Eisen (1990) reported that nearly 60% of their OCD clinic patients had at least one other disorder, and some estimates are even higher (e.g., Lucey, Butcher, Clare, & Dinan, 1994). A wide variety of Axis I disorders appear to be associated with OCD, including anxiety disorders, depression, alcohol abuse (dependence), eating disorders, Tourette's syndrome, body dysmorphic disorder (BDD), hypochondriasis, and schizophrenia. Comorbidity of OCD with other anxiety disorders is quite high, ranging from 25–60% (Weissman et al., 1994). The highest frequencies are evident for specific phobias (Karno et al., 1988), with social phobia and panic disorder co-occurring in 10–50% of samples (see Steketee, Henninger, & Pollard, in press). Interestingly, GAD, which shares a ruminative focus with OCD, occurred infrequently among those with primary OCD (Turner, Beidel, & Stanley, 1992). Family studies have also suggested a shared genetic vulnerability with other anxiety disorders (Black, Noyes, Pfohl, Goldstein, & Blum, 1993).

There is strong evidence for a shared vulnerability for depression as well, with various studies indicating concurrent prevalence rates of 12–80% (Steketee et al., in press). Most studies concur that at least one-third of patients with OCD report histories of clinical depression (e.g., Crino & Andrews, 1996; Weissman et al., 1994), although 25–33% do not report depression (Rassmussen & Eisen, 1992b; Rasmussen & Tsuang, 1986). Interestingly, recent evidence indicates that the frequency of bipolar disorder in OCD is similar to that of major depression in this group (Kruger, Cooke, Hasey, Jorna, & Persad, 1995). The frequency of affective disturbance in OCD, together with genetic and treatment response findings, have led some to question whether OCD might better be classified as a variant of depression. Crucial to this issue is whether the depression which accompanies OCD is primary or secondary to OCD. Most patients report that the onset of depressive symptoms occurred after the onset of OCD (e.g., Rasmussen & Eisen, 1992b; Wilner, Reich, Robins, Fishman, & van Doren, 1976), and several investigators have found that successful treatment of OCD resulted in the reduction of depressive symptoms (e.g., Foa, Steketee, Grayson, Turner, & Latimer, 1984; Yaryura-Tobias & Neziroglu, 1983). These findings, as well as the centrality of anxiety in obsessive thinking and the comorbidity with anxiety disorders, argue for classification of OCD as an anxiety rather than depressive disorder.

Despite a wide range in the degree of insight that accompanies obsessions, OCD is relatively rarely accompanied by schizophrenia (Jenike, 1990; Tallis, 1995), although some research suggests a link (Yaryura-Tobias, Campisi, McKay, & Neziroglu, 1995). Furthermore, Enright (1996) proposed that an information-processing view could account for both disorders. Schizotypy is a normal personality dimension which ranges from healthy to schizophrenic and is proposed as the continuum which links OCD to schizophrenia. Enright and Beech (1990) observed that OCD patients had more schizotypal traits than patients with other anxiety disorders (see also Baer & Jenike, 1992), and, like schizotypal and schizophrenic subjects, had less cognitive inhibition. They proposed that obsessional thoughts result from a failure in the cognitive inhibition of associations to ongoing stimulation. Further research is necessary in order to examine Enright's hypothesis more fully.

A number of impulse control disorders have been found to be comorbid with OCD or to share common features (Tynes, White, & Steketee, 1990) which has led to their labeling as "OC Spectrum Disorders." These include Tourette's syndrome, trichotillomania, kleptomania, and exhibitionism. Many studies have shown an association of OCD to Tourette's syndrome (e.g., Leckman & Chittenden, 1990; Pauls, 1990; Robertson, 1991), as well as to other tic disorders and Sydenham's chorea (Rapoport, Swedo, & Leonard, 1992). In view of observations that one form of Tourette's may be OCD-related and some OCD cases are Tourette's-related (Leckman & Chittenden, 1990), OCD and Tourette's may be different expressions of a similar underlying genetic abnormality (Pauls & Leckman, 1986; Pauls, Towbin, Leckman, et al., 1986), for example, in basal ganglia-frontal cortex pathways (Rapoport et al., 1992). There are also reports of associations between OCD and kleptomania (McElroy, Hudson, Pope, & Keck, 1991), and exhibitionism (Snaith, Baugh, Clayden, Husain, & Sipple, 1983). Although relatively little empirical research has been done examining these relationships, enough suggestive evidence exists to warrant a closer examination.

DSM-IV diagnostic criteria rule out a diagnosis of OCD if the content of the obsessions or compulsions focus on the primary symptoms of several other disorders (e.g., eating disorders, body dysmorphic disorder, hypochondriasis, even if all other inclusion criteria are met. However, the observed similarity of food-related thoughts in eating disorders to obsessions has resulted in a substantial body of research on the relationship between OCD and eating disorders. Beyond similarities in the nature of the symptoms, this research suggests that OCD symptoms are more prevalent in eating disorder patients than in controls, and that eating disorders are found at a higher rate in OCD (8–11%) than in other disorders (Formea & Burns, 1995; Hsu, Kaye, & Weltzin, 1993; Kaye et al., 1992). Also suggestive is the finding that dieting and weight loss can cause changes in serotonergic activity (Goodwin, Fairburn, & Cowen, 1987). Furthermore, there is evidence that behavioral treatments (exposure and response prevention) and serotonergic reuptake inhibitors (SRIs), both effective in treating OCD, are similarly useful for eating disorders (Hsu et al., 1993; Rosen & Leitenberg, 1982). Rothenberg (1986) has suggested that eating disorders are a modern day variant of OCD, and that cultural pressures for thinness in Western societies lead young women vulnerable to OCD to anorexia and bulimia. This reasoning is consistent with Weissman et al's. (1994) proposal that cultural influences partly determine symptom presentation.

Similarly, body dysmorphic disorder has been linked to OCD. BDD is characterized by excessive preoccupation with a presumed defect in appearance. The obsessional nature of thoughts about appearance and anxiety-reducing behaviors (e.g., mirror checking, asking for reassurance) resemble classical OCD symptoms. Simeon, Hollander, Stein, Cohen, and Aronowitz (1995) found that 12% of OCD patients were comorbid for BDD during their lifetime, and Neziroglu and Yaryura-Tobias (1993) have proposed that BDD may actually be a form of OCD. Consistent with this notion, Hollander, Liebowitz, Winchel, Klumker, and Klein (1989) found that all five BDD patients in a clinical series responded to SRIs. The most apparent difference was that BDD patients showed less insight into their illness and may suffer from overvalued ideas rather than obsessions (Vitiello & de Leon, 1990). Nonetheless, in reviewing the limited evidence, Hollander, Neville, Frenkel, Josephson, and Liebowitz (1992) suggested that considering BDD a subgroup of OCD may be useful.

Hypochondriasis is characterized by obsessive-like fears of having a serious disease. The similarities between hypochondriacal and obsessional symptoms has been noted (Fallon, Javitch, Hollander, & Liebowitz, 1991; Tynes et al., 1990), as has the frequency of somatic obsessions in OCD patients (Rasmussen & Tsuang, 1986). In one of the few empirical investigations on this question, Barsky, Wyshak, and Klerman (1986) found comorbid OCD in 9.5% of 42 hypochondriacal patients, while only 2.6% of a nonhypochondriacal control group had OCD. These findings indicate some association of the two disorders.

6.17.4.2 Axis II Disorders

The picture of comorbidity with personality disorders (PD) is somewhat more confused. While it is clear that personality disorders occur frequently in conjunction with OCD, the specific type of personality disorder most often associated with OCD has varied substantially from study to study. This is probably a result of relatively small sample sizes and personality assessment instruments with lower reliability than those used for Axis I disorders. Most studies of Axis II comorbidity have found that over half of patients with OCD also have at least one diagnosable personality disorder (Baer & Jenike, 1992; Cassano, Del Buono, & Catapano, 1993). Along with obsessive-compulsive personality disorder (OCPD), the most frequently diagnosed personality disorders among patients with OCD are avoidant, dependent, histrionic, passive-aggressive, and schizotypal personality disorder. Pfohl and Blum's (1991) conclusion regarding the frequency of PD occurrence in those with OCD may be the most accurate reflection of current knowledge: most patients with OCD do not have comorbid OCPD, and for those who have a personality disorder, it is just as likely to be one other than OCPD. Interesting, no studies have shown antisocial personality disorder to occur frequently in OCD, and Mavissakalian, Hamann, and Jones (1990a) found it to be the least frequent. This finding is consistent with observations that patients with OCD generally show excessive responsibility (see below), a trait quite incompatible with antisocial behavior.

The issue of whether PD is the primary or secondary disorder when comorbid with OCD has rarely been addressed. However, findings such as those by Mavissakalian, Hamann, and Jones (1990b) that the number of personality disorder diagnoses, and scores on personality disorder measures, declined following successful treatment for OCD indicate that such traits may not be intractable patterns of behavior. In particular, avoidant and dependent traits commonly associated with other anxiety disorders may merely reflect a general vulnerability to anxiety and possibly depression, both being responsive to focused treatments. However, some traits such as passive-aggressive behaviors, borderline and schizotypal features may be problematic for treatment outcome (e.g., Baer & Jenike, 1992; Steketee, 1990).

6.17.5 OCD PATIENT CHARACTERISTICS

6.17.5.1 Family Factors in OCD

Several studies report that the celibacy rate, especially among males with OCD who have early onset, is higher than that found in the general population (cf Eisen & Steketee, 1997). Among those who are married, marital dysfunction has been observed, but overall marital satisfaction among clinical patients does not differ from the general population (e.g., Riggs, Hiss, & Foa, 1992). Family members (spouses and parents) often become quite involved in patients' symptoms, especially in performance of rituals and taking over tasks that are difficult or avoided (Calvocoressi et al., 1995; Shafran, Ralph, & Tallis, 1995). They often modify routines to adapt to OC symptoms, sometimes to the extent of restructuring household arrangements altogether. Greater family involvement symptoms was significantly correlated with family dysfunction and negative attitudes toward the patient (Calvocoressi et al.). However, the marital relationships of patients with OCD are no worse than those of the general population (Riggs et al., 1992). Family responses to OCD have been observed to span the range from accommodating to antagonistic toward patients (Livingston-Van Noppen, Rasmussen, Eisen, & McCartney, 1990). Including family members in assessment is recommended, since they may provide important information not volunteered by the patient (Van Noppen, Steketee, McCorkle, & Pato, 1996).

With regard to etiological factors, a number of characteristics of parents have been hypothesized to play a role in the development of OCD. Rachman (1976a) suggested that among individuals with a personal vulnerability (i.e., oversensitivity to criticism), certain parental traits may lead to OCD symptoms. He hypothesized that parents of OCD cleaners are more overcontrolling and overprotective and may model excessive cleaning behaviors. Parents of checkers were hypothesized to set high standards for their children, be overcontrolling and overcritical, as well as modeling meticulousness. Other investigations have observed or speculated on similar themes in the parents of OCD patients, including demanding or critical attitudes (Clark & Bolton, 1985; Hoover & Insel, 1984; Kringlen, 1965; Steketee, Grayson, & Foa, 1985); overprotectiveness (Hafner, 1988; Honjo et al., 1989); perfectionism (Allsopp & Verduyn, 1990; Honjo et al., 1989; Hoover & Insel, 1984; Lo, 1967; Rasmussen & Tsuang, 1986), and rejecting attitudes (Emmelkamp, Hoekstra, & Visser, 1985). However, systematic research on hypothesized differences between the parents of OCD patients and controls has been limited.

Merkel, Pollard, Wiener, and Staebler (1993) found that OCD patients perceived their mothers to be more overprotective than did depressed patients. Contrary to expectation, however, OCD patients described their fathers as less demanding than did panic disorder patients. Chambless, Gillis, Steketee, and Tran's (1996) study of parental warmth and overprotection in OCD and agoraphobia indicated that parents' and patients' recollections of parental behavior did not agree. Patients reported that their parents most often raised them using affectionless control, whereas parents most often rated themselves as having provided optimal parenting. Reports of poorer parenting were associated with worse social adjustment and higher levels of anxious personality disorder characteristics. However, the lack of a normative sample limits conclusions about the relevance of parenting ratings to obsessive symptomatology. Studying volunteers, Frost, Steketee, Cohn, and Griess (1994) found that subjects with subclinical OCD described their parents as more overprotective than nonobsessive-compulsives. They also found some evidence that both parents of the subclinical OCD group were more risk-aversive, and fathers were more critical and perfectionistic than fathers of control subjects. In a related vein, Hibbs et al. (1993) and Hibbs, Hamburg, Kruesi, and Lenane (1993) observed that families of children with OCD had higher levels of expressed emotion (criticism and overinvolvement) than families of control subjects (82% vs. 41%).

While the nature and characteristics of the families of OCD patients are far from clear, there is enough clinical observation and preliminary evidence to hypothesize that the parents of people who develop OCD are more likely to show a constellation of behaviors, including perfectionism, overprotectiveness, overcritical and demanding attitudes, and high levels of risk avoidance. Once further research has established that these characteristics are more prominent in the families of OCD patients than in controls, the question of their role in the development of the disorder can be addressed.

6.17.5.2 Personality and Cognitive Characteristics

Much of the early description of the obsessive-compulsive personality was tied to psychodynamic assumptions of the role of the anal personality in both OCPD and OCD, with little distinction made between these two conditions. The orderliness, parsimony, and obstinacy thought to typify people with OCD were expanded into more specific personality traits such as those articulated by Pollack (1987): "procrastination, indecision, self-doubting, driven by obligations and responsibilities, attempting to attain order and perfection, bound by routines, uncomfortable and avoidant of novelty and change, interpersonally formal and even cold, authoritarian and controlling, autocratic, critical, and punitive" (pp. 249–250). Although some of these latter traits better describe OCPD, others have recently been viewed as characteristic cognitive aspects of OCD.

A host of personality/cognitive features are closely linked to OCD (see Reed, 1985; Tallis, 1995). Prominent among these are overestimation of risk, exaggerated responsibility and guilt, beliefs about the need to control thoughts, fusion of thought and action (thinking something makes it so), perfectionism, uncertainty and doubt, and intolerance for emotional discomfort. These domains were also identified by a group of international researchers attempting to develop a coherent approach to the assessment of cognitions in OCD (International OCD Cognitions Working Group, 1996). The group noted that many of these domains overlap considerably. Using an expert consensus strategy, they selected a subset of five areas and defined these in order to generate items for a self-report measure of these domains. In discussing the several cognitive domains listed below that may be common features of OCD, we include the expert consensus definitions of the working group where relevant.

6.17.5.2.1 Overestimation of risk

It has long been noted that patients with OCD exaggerate the probability and severity of bad outcomes (Carr, 1971, 1974). Clinical (Rasmussen & Eisen, 1989) as well as empirical (Steketee & Frost, 1994) evidence supports the hypothesis that people with OCD are risk-aversive. An expert working group defined threat overestimation as "exaggeration of the probability or severity of harm believed to result from or follow intrusive internal experiences" (Obsessive Compulsive Cognitions Working Group 1997).

6.17.5.2.2 Responsibility

Overimportance assigned to intrusive thoughts and an exaggerated sense of responsibility for harm coming to self or others are closely related cognitive features widely noted in the OCD literature (Freeston, Rhéaume, & Ladouceur, 1996; Rachman, 1993). Salkovskis (1985) argued that exaggerated responsibility is central to the symptoms of OCD, and empirical literature supports its importance in OCD (cf Lopatka & Rachman, 1995; Rachman, 1993; Salkovskis, 1985). Closely related to responsibility is the emotional experience of guilt. People with OCD have been found to score significantly higher than other groups on measures of guilt (Frost, Heimberg, Holt, Mattia, & Neubauer, 1993; Niler & Beck, 1989; Steketee, Quay, & White, 1991). Salkovskis' and others' definition of responsibility adopted by the working group was "the belief that one has power which is pivotal to bring about or prevent subjectively crucial negative outcomes. These outcomes are perceived as essential to prevent. They may be actual, that is, having consequences in the real world, and/or at a moral level" (Salkovskis, Rachman, Ladouceur, & Freeston, 1992).

6.17.5.2.3 Controllability of thoughts

Another important cognitive feature of OCD is disruption in the control of thoughts. Clark and Purdon (1993) suggested that in normal cognitive processing, irrelevant and intrusive thoughts are simply ignored, whereas individuals with OCD attend to them, consider them important, and attempt to suppress them. The failure of these efforts at suppression results in sensitization and vigilance to all thoughts similar to the intrusion. The cycle of attention to intrusions followed by attempts at suppression are hypothesized to escalate to an obsessional pattern (Salkovskis & Campbell, 1994; Wegner, 1989). The cognition working group identified a domain of "need for control" and defined this as "overvaluation of the importance of exerting complete control over intrusive thoughts, images, and impulses, and the belief that this is both possible and desirable." Dimensions of the need to control were considered to include tracking mental events and hypervigilance; moral consequences—control as a virtue; psychological and behavioral consequences—insanity, decreased ability to function; and immediate efficiency and future success of control efforts.

6.17.5.2.4 Overimportance of thoughts

The reasons for the overimportance of intrusive thoughts and the perceived necessity for controlling them may lie in a belief about the meaning of these thoughts. Rachman and others proposed that people with OCD tend to believe that simply having a thought, particularly an unwanted and objectionable thought, is morally equivalent to engaging in the objectionable act (Rachman, Thordarson, Shafran, & Woody, 1995). That is, thinking something morally reprehensible is just as bad as doing it. Evidence for such "thought-action fusion" (TAF) can be found in the perception of OCD patients that imagined negative events are more likely actually to occur (Rachman, Shafran, et al., 1995). Such beliefs that thoughts can influence the probability of occurrence of negative events may be a part of an exaggerated sense of responsibility (for negative thoughts and events) by people with OCD. Recent attempts to measure and study TAF support the importance of this construct (Rachman, Shafran, et al., 1995; Shafran, Thordarson, & Rachman, 1996). The Obsessive Compulsive Cognitions Working Group (OCCWG, 1997) identified "overimportance of thoughts" as a central feature of OCD and defined this as a "belief that the presence of an intrusive thought per se gives it significance or means that it is important. This includes concepts of thought/action fusion, Cartesian reasoning, and magical thinking."

6.17.5.2.5 Tolerance for ambiguity

Part of the construct of perfectionism (Frost, Marten, Lahart, & Rosenblate, 1990), one that is perhaps more uniquely tied to OCD, is the experience of doubt about the veracity of experience and the quality of one's actions. Although not part of the formal diagnostic criteria, doubting is a hallmark of OCD (Reed, 1985). Doubt about the validity of experience can leave one uncertain about whether actions are appropriate or adequate to ward off danger. The need for absolute certainty or the inability to tolerate uncertainty is also a commonly noted feature of OCD (Kozak, Foa, & McCarthy, 1987; Rasmussen & Eisen, 1990). Guidano and Liotti (1983) suggested that doubt stems from a belief that every situation has a perfect solution and the person with OCD will not feel comfortable unless that perfect solution is found. Others have suggested that the need for certainty maintains ritualistic behaviors which must be repeated because certainty has not been achieved (Beech & Liddell, 1974). Such doubt is believed to be the root of the indecision seen in people with obsessive-compulsive tendencies (Frost & Shows, 1993). The working group identified intolerance of ambiguity as encompassing "beliefs about the necessity of being certain, about the capacity to cope with unpredictable change, and about adequate functioning in situations which are inherently ambiguous."

6.17.5.3 Information/Cognitive Processing in OCD

Quite apart from personality traits and beliefs which characterize OCD, several information-processing capabilities have been hypothesized to distinguish OCD patients from nonpatients. These processes involve memory, categorization, and attention. Reed (1977) suggested that a "pathologically faulty memory" was a central feature of OCD. Patients with OCD may suffer from deficits in memory for actions (especially among those with checking rituals, general memory, confidence in memory functioning, and reality monitoring (distinguishing actual from imagined actions). For nonclinical compulsive checkers, several studies have established the presence of deficits in memory for discrete actions performed (e.g., Rubenstein, Peynircioglu, Chambless, & Pigott, 1993; Sher, Frost, Kushner, Crews, & Alexander, 1989; Sher, Frost, & Otto, 1983). However, among clinical patients, this finding is less consistent, with one study supportive and another failing to replicate (see Tallis, 1995). With respect to general memory deficits, similar findings have emerged from neuropsychological studies of OCD. Nonclinical subjects showed poorer overall memory functioning (e.g., Sher et al., 1989), whereas findings for clinical patients were more mixed, although most studies reported poorer memory functioning, especially visual memory (e.g., Boone, Ananth, Philpott, Kaur, & Djenderedjian, 1991; Christensen, Kim, Dysken, & Hoover, 1992; Martinot et al., 1990; for an exception see Constans, Foa, Franklin, & Matthews, 1995). Research has consistently shown lowered confidence in memory capacity but not worse reality monitoring (Brown, Kossyln, Breiter, Baer, & Jenike 1994; Constans et al., 1995; McNally & Kohlbeck, 1993; Sher et al., 1983).

Consistent with hypotheses about memory dysfunction in OCD, Reed (1985) also argued that such patients fail to categorize and integrate information, leading to underinclusive categorization in which many small categories are generated with narrow boundaries and overspecified rules for category membership. Persons and Foa (1984) found some empirical support for this notion, and Frost and Hartl (1996) have described this as a nettlesome problem for compulsive hoarders. Because of a general failure to recognize reasonable boundaries for categories and experiences, OCD patients create artificial structures to govern their behavior. Thus, rituals, particularly checking rituals, are attempts to create an artificial structure to determine the end of a sequence of behaviors.

Other theorists have emphasized deficits in attentional processing in OCD. Enright and Beech (1990, 1993a, 1993b) proposed an information-processing theory of OCD which focuses on deficits in selective attention. They hypothesized that OCD patients fail to adequately inhibit irrelevant stimuli during normal cognitive processing, interfering with their focus on relevant stimuli. Consequently, processing of ordinary information requires a more conscious and effortful strategy, and suppression of irrelevant or unwanted thoughts is more difficult. The same cognitive inhibition deficit is hypothesized to characterize schizophrenia and schizotypal personality disorder (Enright, 1996). Enright and Beech (1993a, 1993b) report several studies which support their hypotheses using a negative priming paradigm.

Another approach to studying attentional processes in OCD has been to examine the processing of fear-relevant information, mainly using a Stroop task in which subjects observe fearful and nonfearful words presented in different colors, and are asked to name the color of the word, thereby inhibiting processing of its meaning. Fearful word meanings should be harder to inhibit and therefore take longer. Consistent with this hypothesis, Foa, Ilai, McCarthy, Shoyer and Murdock (1993) found that OCD washers took longer to name contamination words than neutral words, and that patients with other types of fears did not show such an effect, supporting the specificity of the processing effect. Lavy, van Oppen, and van den Hout (1994) found that OCD subjects selectively attended to negative OC-related words (e.g., disease, disaster) but did not show an attentional bias to positive ones (e.g., clean, precise). Of considerable interest is Foa and McNally's (1986) observation that after successful treatment OCD patients no longer showed attentional bias toward contamination items in a dichotic listening task.

The above research findings provide some confirmation of OCD deficits in memory for actions and in general memory. Decreased confidence in memory functioning is evident in clinical samples, but not problems in distinguishing actual from imagined actions. Problems with overspecification in categorization have been observed, and OCD patients also show problems in inhibiting attention to irrelevant information and in attending to relevant ones. Specific causes for these difficulties will require further research.

6.17.6 THEORETICAL MODELS AND TREATMENTS FOR OCD

The review presented below includes three major theoretical models for the development and maintenance of OCD, followed by treatment methods that are generally derived from these models. We then present evidence for their effectiveness. We begin with behavioral theories and treatments, followed by a discussion of newer cognitive models and derived treatments, and biological models and pharmacological treatments for OCD. In juxtaposing models with treatment methods, some anomalies in the practical application of methods derived from theory will be apparent. In the case of biological theories, the discovery of effective treatments (serotonergic reuptake inhibitors, SRIs) have partly driven speculations about explanatory models such as the role of serotonin. At the end of this section, we examine combined behavioral and pharmacological treatments and follow this with an effort to determine the comparative efficacy of behavioral, cognitive, and pharmacological treatments and their combinations.

6.17.6.1 Behavioral Models for OCD

Early psychological conceptualizations of the development of OCD were based mainly on psychodynamic models. However, although treatments derived from these models lead to some improvement, more for outpatients than inpatients, average rates of approximately 60% improved were not considered impressive (see Cawley, 1974). Not surprisingly, OCD gained a reputation as intractable (e.g., Coryell, 1981). Early behavioral treatments (thought stopping, aversion therapy) based on contingent reinforcement and punishment models were also successful for only a limited group of patients, often fewer than 50% (Emmelkamp & Kwee, 1977; Kenny, Mowbray, & Lalani, 1978; Stern, 1978). Interestingly, early efforts to utilize conditioning models alone (systematic desensitization, paradoxical intention, repeatedly verbalizing obsessive thoughts) improved the situation only slightly (Beech & Vaughn, 1978; Emmelkamp & Kwee, 1977; Solyom, Garza-Perez, Ledwidge, & Solyom, 1972). In the late 1960s and early 1970s newly developed intensive behavioral treatments that incorporated both conditioning models and negative reinforcement notions showed considerable promise. Labeled exposure and response (ritual) prevention (ERP), these strategies were derived from learning theory models discussed below.

The first part of a two-stage theory for acquiring and maintaining fear and avoidance behavior posits that an otherwise neutral event acquires the capacity to provoke fear because of its pairing with an aversive experience, much as a dog phobia might develop from being bitten (Dollard & Miller, 1950; Mowrer, 1960). Evidence supporting this fear acquisition process is inadequate, since many patients cannot recall conditioning experiences associated with symptom onset, and although onset often follows very stressful life events, it rarely does so immediately, as postulated by the traumatic onset theory (e.g., Rachman & Wilson, 1980). A more likely explanation is that stressful events sensitize some individuals to cues that have an innate tendency to elicit fear, were learned during early traumatic experiences, or have special cultural significance (Rachman, 1971; Teasdale, 1974; Watts, 1971). Observational or informational learning may also account for some OCD onsets (Foa & Kozak, 1985; Rachman, 1977), particularly when clients report that their symptoms resemble their parents' behavior. For example, hoarding rituals may mimic a cluttered home resulting from excessive saving by a parent of unneeded but "potentially useful" objects.

In the second stage of the two-factor model, any actions (escape or avoidance behavior, compulsions) that relieve obsessive anxiety/discomfort are negatively reinforced because they result in reduction of discomfort. Their demonstrated ability to terminate unpleasant experiences renders them very likely to be repeated in future situations. As Foa and Kozak (1985) have noted, both external cues (objects or situations) and internal triggers (thoughts, images, or impulses) can produce obsessive discomfort. Many of these triggers (e.g., locking the front door) cannot be avoided and therefore passive avoidance behaviors which enable phobics to manage their fears are often insufficient to control anxiety for those with OCD. More active compulsions are needed to prevent harm or restore a feeling of safety (Rachman, 1976a).

Substantial evidence supports a behavioral view of OCD as a functional anxiety-based disorder in which obsessions increase discomfort and compulsions reduce it. In the laboratory, production of obsessive thoughts increased heart rate and skin conductance more than neutral thoughts (e.g., Boulougouris, Rabavilas, & Stefanis, 1977), and contact with contaminants also increased subjective and physiological anxiety reactions (Hodgson & Rachman, 1972; Hornsveld, Kraaimaat, & van Dam-Baggen, 1979). Corresponding to the second stage of the behavioral theory, completing compulsions reduced anxiety in most instances (e.g., Hodgson & Rachman, 1972; Hornsveld et al., 1979). Occasionally reported increases in discomfort following compulsions appear to be frustration reactions to having to complete time-consuming unwanted behaviors. Thus, evidence supports a two-stage behavioral model characterizing the maintenance of OCD symptoms, but it does not account for many instances of onset unless other explanatory variables are added.

6.17.6.2 Behavioral Treatments for OCD

Treatment strategies that follow logically from the model attempt to disconnect obsessions from their associated discomfort (deconditioning), rendering them more like intrusions in normal samples, and to eliminate rituals that reinforce obsessive fears. In an elaboration on this two-factor model, Foa and Kozak (1985) concurred with Peter Lang's theories that fear or anxiety could be viewed as a memory network containing representations of concrete stimuli, associated physiological responses, and meanings assigned to these (cognitive elements). Correction of all three components would be necessary to eliminate fear, although exposure to at least one part of the network might evoke other elements. Following from both models, behavioral intervention for obsessions and chronic rituals involves exposure to overt and covert cues that provoke obsessions and blocking of rituals (response prevention) to prevent them from prematurely reducing obsessive fears and thus interfering with habituation of the negative mood generated by obsessions. This ERP treatment has been widely studied and found highly effective, although as for medication treatment, residual symptoms and treatment refractory cases plague the field. Salient research on ERP and its components are described below.

Considerable improvement in effectiveness of treatments came with the systematic application of prolonged exposure to obsessions and blocking or prevention of compulsions. Meyer and his colleagues are credited with development of effective ERP treatment with hospital patients who feared contamination and hand washing rituals (Meyer & Levy, 1973; Meyer, Levy, & Schnurer, 1974). Using a strict and intensive program of daily direct contact with contaminants and complete prevention of rituals (turning off plumbing in patients' rooms and having nurses observe all washing), these investigators reported that 10 of their 15 patients were much improved or symptom-free, and the remaining five were moderately improved. Only two relapsed after five to six years. Not surprisingly, such remarkable results led to many investigations of ERP in inpatient and outpatient settings in several countries. To date, approximately 30 open trials and controlled studies representing over 600 clients with OCD have reported good outcomes with variants of ERP. Tables detailing the outcome of these studies are available in a review by Steketee and Shapiro (1993). Highlights of findings from these studies along with their clinical implications, particularly for variations of ERP, are discussed below.

6.17.6.2.1 Degree of improvement from ERP

A series of early studies of inpatients by Rachman, Marks, and colleagues at the Maudsley Hospital in London indicated that 15 sessions of exposure and blocking over a threec-week period was considerably more effective than relaxation training, and overall effectiveness was excellent even two years later (Marks, Hodgson, & Rachman, 1975). Seventy-five percent of patients were responsive to ERP, and the majority showed substantial change. This improvement rate has been replicated across numerous studies using 10–20 sessions of treatment. On average, 85% of clients were at least "improved" immediately after treatment, and about 55% fell into the "much improved" or "very much improved" categories, implying that target symptoms improved by more than half (Steketee, 1993). At follow-up, improvement rates remained high, averaging about 75%, and 50% in the "much improved" or "very much improved" groups. Thus, results were largely maintained with some relapse evident and some clients needing additional therapy (Emmelkamp, Visser, & Hoekstra, 1988; Foa & Goldstein, 1978).

With regard to symptom improvement, the most positive results were evident on target ratings of treated obsessions and compulsions, where the average degree of benefit immediately after ERP ranged from 40%–75% (Steketee & Shapiro, 1993). On other standardized self-report measures, the level of improvement was slightly lower. At follow-ups ranging from three months to six years, treatment gains were only slightly different across studies, falling in the 45–70% range for target symptoms, and averaging 50% improvement for questionnaire measures. The consistency of ERP results and the degree of benefit across multiple treatment sites and countries is quite impressive, clearly indicating that exposure and response prevention is an effective treatment for OCD symptoms. However, the above summary does not address the question of how many of these treated patients no longer meet criteria for the diagnosis of OCD or no longer require intervention for their symptoms. Some more recent investigations have assessed improvement using the standardized Yale–Brown Obsessive Compulsive Scale (YBOCS) in which scores correspond closely to diagnostic categorizations (16 and above usually meets criteria for OCD). Outcomes of studies using the YBOCS to assess ERP treatment are presented in Table 1.

6.17.6.2.2 Specific effects of exposure and response prevention

Examining the separate effects of exposure and response prevention on obsessions and compulsions, Foa and colleagues, as well as other investigators, found support for the behavioral model from which ERP is derived. In their first trial with 21 patients (Foa & Goldstein, 1978), 86% had stopped ritualizing after treatment, but only 57% improved substantially on obsessions, suggesting that compulsions responded more to ERP than obsessions. According to behavioral theory, exposure should particularly reduce anxiety associated with obsessions, whereas prevention of ritualistic behavior should influence rituals more than obsessions. This dual hypothesis was supported in multiple case studies of OCD patients (Mills, Agras, Barlow, & Mills, 1973; Turner, Hersen, Bellack, Andrasik, & Capparell, 1980), as well as in group trials. Both interventions delivered separately reduced obsessions and rituals, but exposure affected subjective anxiety more than rituals, and response prevention reduced rituals more than obsessions. The addition of the missing component led to maximum improvement in both symptoms (Foa et al., 1984; Foa, Steketee, & Milby, 1980). These findings support the suggestion that exposure and prevention of rituals operate by separate mechanisms.

6.17.6.2.3 Processes during exposure

Foa and colleagues have suggested that effective ERP relies on subjective and physiological processes activated by exposure and blocking, that reflect the processing of emotional information pertinent to the obsessive symptoms. These processes include activation of obsessive anxiety (reactivity) and gradual decrease (habituation) of these reactions during exposure, as well as across exposure sessions (Foa & Kozak, 1986). Habituation during direct imagined exposure, and habituation across sessions, have been demonstrated (Grayson, Foa, & Steketee, 1982; Shahar & Marks, 1980) and associated with more change in therapy (Foa, Steketee, Grayson, & Doppelt, 1983; Kozak, Foa, & Steketee, 1988). Habituation rates may vary as a function of the intensity of arousal (Foa et al., 1983; Kozak et al., 1988). Therefore clinicians can use intensity of arousal and habituation rates to judge the course of therapy and make adjustments in the type and length of exposure necessary. Findings by Grayson and colleagues regarding the benefits of attention-focusing versus distraction from exposure (e.g., Grayson, Steketee, & Foa, 1986), suggest that in standard clinical practice therapists are advised to ask clients to concentrate on their fears during exposure, and avoid distracting conversation that might interfere with the processing of fearful information.

6.17.6.2.4 Problems in the treatment of pure obsessive symptoms

Early evidence indicated that patients with obsessions but no overt compulsions tended to have poorer outcome than those with clear rituals (e.g., Emmelkamp & Kwee, 1977; Rachman, 1976b). A possible reason for their failure to respond consistently to exposure treatment is that obsessions alternated with undetected and untreated mental rituals (see Salkovskis & Westbrook, 1989; Steketee & Foa, 1985). In such cases, prolonged exposure might actually have strengthened mental rituals by repeatedly provoking them and thereby preventing habituation of obsessions, consistent with Salkovskis and Westbrook's (1989) observation that following an obsessive thought with a "neutralizing" one (mental ritual) actually *increased* discomfort. According to behavioral theories of

Table 1 Studies of treatment outcome for OCD using YBOCS assessment and Clinical Improvement Scales.

Study (Year)	*N*	*Treatment*	*Type*	*# of sessions*	*# of weeks*	*YBOCS pretest*	*YBOCS post-test*	*YBOCS follow-up*	*Time of follow-up*	*Percentage of subjects clinically improved post-test follow-up*	*Comments*
Jenike et al. (1990)	18	SRI	Fluv		10	22.6 (3.5)	18.8 (4.0)				
Greist et al. (1995)	160	SRI	Fluv		10	23.3 (4.8)	18.4			38[b]	
Freeman et al. (1994)	34	SRI	Fluv		10	26.2	17.7			59[b]	
	32	SRI	Cmi		10	25.5	17.6			53[b]	
Clomipramine Collaborative Study Group (1991)	118	SRI	Cmi	12	10	26.3 (5.5)	16.2 (0.7)			60[b]*	* combined % from both sides
	134	SRI	Cmi	12	10	26.2 (4.9)	14.7 (0.7)				
Jenike et al. (1989)	13	SRI	Cmi		10	24.1 (5.8)	15.2 (5.2)				
	61	SRI	Fluox		12	22.2 (6.2)	13.9 (7.0)				depressed pts did not differ from nondepressed
Tollefson et al. (1994)	90	SRI	Fluox (60mg/d)	8	13	24.4 (5.1)	16.8 (7.8)			38[b]*	* at all doses, a greater proportion of Ss who took Fluox showed more clinical improvement than control (clinical improvement: 35% > baseline)
Goodman et al. (1989)	21	SRI	Fluv		8	25.0 (6.0)	14.3 (7.0)				
Chouinard et al. (1990)	43	SRI	Ser	5	8	23.4 (4.9)	19.6			56 or 25?[b]	
Greist et al. (1995)	80	SRI	Ser 200mg/d	∼7	12	23.6*	16.1*	17.3	48 weeks	39[b]	* estimated from graph
Jenike et al. (1990)	10	SRI	Ser	6	10	22.8 (6.0)	20.6 (9.2)				

Table 1 (continued)

Study (Year)	*N*	*Treatment*	*Type*	*# of sessions*	*# of weeks*	*YBOCS pretest*	*YBOCS post-test*	*YBOCS follow-up*	*Time of follow-up*	*Percentage of subjects clinically improved post-test follow-up*	*Comments*
Orloff et al. (1994)	85	SRI	Mixed		10 +	23.7 (6.7)	–	10.1 (7.0)	1–3.5 year		Chart review method, 29% also had 10 + hours of ERP, most pts still on meds.
Krone et al. (1991)	17	ERP-without meds*	Indiv	7	7	20.4 (6.3)	17.2 (6.9)				* 55% were already on SRIs
	18	ERP-with meds	Indiv	7	7	21.9 (6.6)	15.1 (6.1)	11.6* (6.4)	3 months		* Follow-up is for all subjects (*n* = 24) on or off meds.
Steketee et al. (1995)	45	ERP	Indiv*	22	16	24.4 (5.4)	15.1 (7.2)	14.5 (7.9)	3 months		* Some subjects on SRIs; at 3 month follow-up *n* = 34
Hiss et al. (1994)	8	ERP + RP*	Indiv	19 + 9-phone	4 12	23.1 (5.4)	8.8 (4.5)	8.1 (7.1)	6 months		* RP + Relapse prevention
	10	ERP + AT*	Indiv	19	4	23.9 (7.3)	9.8 (7.2)	13.4 (8.8)	6 months		* AT = Associative therapy (placebo)
Fals-Stewart et al. (1993)	30	ERP	Indiv	24	12	20.2 (4.0)	12.1 (3.7)	12.9 (3.5)	6 months		
	31	ERP	Group	24	12	22.1 (4.1)	12.0 (3.4)	14.0 (3.6)	6 months		
Van Noppen et al. (1996)	17	ERP	Group	10	10	23.9 (7.2)	16.6 (7.2)	14.5 (7.3)	1 year	31[a]	43[a]
	19	ERP	Multi-family	10–12	10–12	23.9 (5.4)	15.6 (6.5)	14.9 (7.8)	1 year	47[a]	58[a]
Van Oppen et al. (1995)	29	ERP	Indiv	16	16	25.4 (7.0)	17.3 (8.3)			28[a]	
	28	COG	Beckian	16	16	24.1 (5.5)	13.3 (8.5)			50[a]	

ERP = exposure and response prevention, Fluv = fluvoxamine, Cmi = clomipramine, Fluox = fluoxetine, Ser = sertraline, SRI = serotonin reuptake inhibitor, COG = cognitive therapy, YBOCS = Yale–Brown Obsessive Compulsive Scale. [a] Jacobson and Truax method. [b] Clinical Global Improvement rating.

OCD, effective treatment would require that exposure be applied only to obsessive thoughts, and mental compulsions should be blocked (e.g., by thought stopping, distraction, or substitution). Several researchers have demonstrated the success of such a strategy especially using uninterrupted audiotaped exposures to narrated obsessions (e.g., Rachman, 1976c; Salkovskis & Westbrook, 1989).

6.17.6.2.5 Variants in delivery of ERP

Several variations on ERP procedures have been studied. Emmelkamp and colleagues in Holland reported positive effects for a home-based therapy using ERP without the careful monitoring of *in vivo* exposure and rituals carried out with hospitalized patients (Boersma, Den Hengst, Dekker, & Emmelkamp, 1976). In a controlled study, therapist-directed exposure did not improve on the outcome of self-controlled treatment (Emmelkamp, van den Heuvell, Ruphan, & Sanderman, 1989) and the latter may have produced more durable gains (Emmelkamp & Kraanen, 1977). In Kirk's (1983) efforts to apply ERP to consecutive patients in an outpatient clinic setting, therapists relied more on homework assignments than on accompanied direct practice to achieve gains, unless progress required the therapist's presence. Correspondingly, in a third of cases, relatives and friends played a significant role in assisting in treatment. The deliberate involvement of relatives in behavioral treatment is described in more detail below.

Dutch investigators also reported that 10 sessions of treatment was too short, since on average clients required an additional 15 treatment sessions during follow-up (Emmelkamp, van der Helm, van Zanten, & Plochg, 1980). Foa and colleagues concurred, as evident from the expansion of their program from 10 sessions to 15 sessions plus home visits (see Foa & Goldstein, 1978; Steketee & Foa, 1985). Rachman and colleagues observed that 30 sessions led to somewhat more improvement than 15 sessions, but the gains were not substantial (Rachman et al., 1979). However, in a clinical study, Kirk (1983) obtained good outcome despite the fact that half of the patients attended 10 or fewer sessions directed specifically at the OCD symptoms. However, it may be that Kirk's clinic patients would have benefited from more sessions but could not afford the time or money to do so. In research trials, frequency of sessions has not made a difference in outcome (Emmelkamp et al., 1989), and examination of the outcome of studies that used different frequencies of treatment produced the same findings (Steketee & Shapiro, 1993).

Direct (*in vivo*) exposure has been delivered in nearly all studies of ERP, but some studies have also included imagined exposure or have compared these variants. Findings indicate that direct exposure is needed (Rabavilas, Boulougouris, & Stefanis, 1976) and that imagined exposure may confer added benefits, particularly at follow-up, for those patients for whom catastrophic outcomes are salient components of their obsessions (e.g., Foa, Steketee, & Grayson, 1985). From the above findings it appears that 10–20 sessions of self- or therapist-controlled ERP applied one to five times per week provides an adequate trial of ERP for most patients with OCD. Longer direct exposure sessions or assignments (generally about an hour) are more advisable than short exposures (Rabavilas et al., 1976) to permit reduction in their obsessive anxiety. Imagined exposure may be added for patients with problematic catastrophic fears.

6.17.6.2.6 Group treatment for OCD

Group interventions for delivering ERP have been examined in several studies in an effort to identify cost-effective strategies. Hand and Tichatzky (1979) treated 17 OCD patients in three groups targeting OC symptoms, social interaction problems, and problem-solving ability in twice-weekly group sessions containing *in vivo* exercises, homework assignments, communication skills training, and separate support group meetings for spouses. Some decrease was observed in OCD symptoms and anxiety, but this varied considerably across groups. Two smaller case series employed 10–14 group sessions of education, goal setting, behavioral skills training, exposure, response prevention, homework logs, and cognitive restructuring (Epsie, 1986; Taylor & Sholomoskas, 1993). Benefits at the end of treatment and at follow-up were substantial and generally comparable to individual treatment trials.

Several large trials have been reported. Enright (1991) treated 24 patients (approximately six per group) for nine weekly 90-minute sessions using a similar multidimensional behavioral treatment that included assertiveness training. The author reported significant improvement in OCD symptoms and functioning at post-test and six-month follow-up, although only 17% were clinically significantly improved. In a second uncontrolled trial, Krone, Himle, and Nesse (1991) treated 36 OCD patients in seven-week groups that included instruction in self-treatment and therapist-directed ERP during treatment sessions with an optional family psychoeducational session. YBOCS scores reduced significantly to below clinical levels at three-month follow-up, and improvement occurred regardless of medication use. In two studies by Van Noppen and colleagues, 10 sessions of group ERP were used to treat a total of 90 patients with OCD. YBOCS scores again reduced significantly (an average of 5 to 10 points) and general functioning also improved (Van Noppen, Pato, Marsland, & Rasmussen, 1995; Van Noppen et al., 1996). See Table 1 for details of outcome on YBOCS scores for group ERP treatment.

In the only controlled trial of group ERP, Fals-Stewart, Marks, and Schafer (1993) compared group ERP with or without imagined exposure ($N=30$) to comparable individual treatment ($N=31$) and to an individual relaxation control treatment ($N=32$). Following 24 sessions over a 12-week period, subjects in both exposure treatment conditions improved significantly in OCD symptoms, whereas the relaxation control group changed only on anxiety. Post-test YBOCS scores fell below the clinical range. Individual treatment led to comparable but somewhat more rapid change (see Table 1). Because this study excluded patients with major depression and Axis II diagnoses, both common in OCD as noted above, a replication is needed to determine the generalizability of these findings to a more representative sample. Nonetheless, the benefits of a group modality for delivery of ERP are quite clear.

6.17.6.2.7 Family treatment for OCD

As noted earlier, researchers have commented on the adverse effects of OCD symptoms on family functioning and the often extensive family involvement in patients' OCD symptoms (Marks et al., 1975; Calvocoressie et al., 1995). Several researchers have suggested advantages for support group involvement of both patients and family members. In an early study, Marks et al. (1975) employed an open-ended monthly group for family members and patients to discuss the effect of OCD on the family and to plan coping strategies and rehearse behavioral exercises. More recently, several reports have outlined psychoeducational focuses for time-limited family support groups, including sessions on diagnosis, assessment, theories of OCD, behavioral treatment, medications, and prevention of relapse (e.g., Black & Blum, 1992; Tynes et al., 1992). Psychoeducational goals included improving self-esteem, sharing feelings and experiences, accepting patients' realistic limitations, and learning strategies for coping with OCD symptoms. These reports noted high participant satisfaction with psychoeducational groups, but provided no outcome data.

Hafner (1982) and others (e.g., Cobb, McDonald, Marks, & Stern, 1980; Hoover & Insel, 1984) described cases in which patients improved when spouses or family members participated in the behavioral treatment process. However, group trials that examined inclusion of family members in ERP have produced mixed results. Although early results by Emmelkamp and DeLange (1983) indicated that spouse assistance in exposure treatment might improve outcome, a second larger study indicated no difference in immediate or follow-up outcome, despite improvement in marital satisfaction (Emmelkamp, de Haan, & Hoogduin, 1990). However, in this trial, treatment may have been too brief (eight sessions in five weeks) to permit differences to be detected, and the lack of communication training regarding OCD symptoms (see Emmelkamp, Kloek, & Blaauw, 1992) may have reduced family-assisted treatment efficacy. These negative findings are contradicted by Mehta (1990) who reported that involving spouses and other family members in ERP for 30 patients in India led to significantly greater gains in OCD symptoms, mood state, and social and occupational functioning, compared with unassisted treatment. At follow-up, family-treated patients continued to improve, whereas individually treated patients lost some gains, suggesting that family involvement may be especially useful in forestalling relapse. An analysis of family processes affecting outcome indicated that nonanxious, consistent family members were more successful in providing support and supervision than anxious and inconsistent ones, especially those who engaged in argument and ridicule. The discrepant findings from these studies could reflect study differences in treatment length, family role, and cultural differences in the populations of Western Europe and India (see Hafner, 1992).

Thornicroft, Colson, and Marks's (1991) impatient treatment program in the UK emphasized self-treatment and teaching relatives to reduce their involvement in rituals and to encourage self-exposure in a noncritical manner. Relatives practiced these skills under the therapist's supervision on the ward. Average symptom decreases of 45% at discharge and 60% at a six-month follow-up, with clear improvement in functioning, indicated good success for this very disabled inpatient population. Similarly, Van Noppen et al. (1996) treated 19 patients in three co-therapy groups that included spouses/partners, parents, and other relatives in 10–12 two-hour sessions. YBOCS scores were reduced substantially to below clinical levels and the reductions were still significant at one-year follow-up. Improvement was also evident on disability scores. In this era of managed care, efforts to reduce costs while maintaining efficacy argue for the potential value of group and family behavioral interventions for OCD.

6.17.6.3 Cognitive Theories of OCD

A number of cognitive theories of OCD grow out of an appraisal model (Lazarus, 1966) in which appraisals of threat (primary appraisal) and appraisals of coping resources (secondary appraisal) are considered central to the development of OCD. As noted earlier, Carr (1974) observed that obsessive compulsives make abnormally high estimates of the probability of unfavorable outcomes. That is, they more often anticipate danger or negative outcomes than do nonobsessionals. Carr (1974) suggested that compulsive behaviors are intended to reduce the threat created by this erroneous judgment of risk. McFall and Wollersheim (1979) expanded on this idea and theorized that the overestimate of threat by patients with OCD is part of an inaccurate primary appraisal process (cf Lazarus, 1966) which contributes to the development of OCD. They proposed that four types of beliefs or assumptions contribute to inaccurate appraisal and result in the overestimation of negative outcomes. These include beliefs about the need for perfect competence to feel worthwhile and avoid criticism; that making mistakes or failing to meet one's ideals should result in punishment or condemnation; that one has power to prevent disastrous outcomes by magical rituals; and that certain thoughts and feelings are unacceptable, potentially catastrophic, and worthy of punishment. McFall and Wollersheim also suggested that patients with OCD are deficient in secondary appraisal of their abilities to cope with threat.

In a very similar vein, Guidano and Liotti (1983) outlined a cognitive model of OCD that emphasized two primary features: the need for perfection and for certainty to avoid criticism and risk. These theorists further suggested that three of the irrational beliefs described by Ellis (1962) are especially relevant in OCD. These are that one should be thoroughly competent, adequate, and achieving in all possible respects to consider oneself worthwhile; that one should be very concerned about potential danger and dwell on the possibility of its occurring; and that perfect solutions to all problems are available and should be sought. A similar model was also proposed by Warren and Zgourides (1991) whose rational emotive therapy model derives from Ellis's concepts. Like other appraisal models, it also emphasizes the role of irrational beliefs. They propose that unpleasant or stressful events produce intrusive thoughts. If irrational beliefs are related to these intrusive thoughts, they become important and develop into obsessions. The above theories are clearly similar with regard to the specific irrational beliefs hypothesized to be most central to OCD. As evident below, these proposals have been largely subsumed by more recent cognitive theorizing.

The best articulated and experimentally supported cognitive theory of OCD has been proposed by Salkovskis (1985, 1989) who has drawn heavily from Beck's (1976) cognitive theory of emotional disorders. In this model, Salkovskis relies on extensive research showing that intrusive cognitions are normal phenomena experienced by over 90% of the population (e.g., Freeston, Ladouceur, Thibodeau, & Gagnon, 1991; Rachman & deSilva, 1978; Salkovskis & Harrison, 1984). Thus, what distinguishes people with OCD is not the experience of intrusive thoughts, but the way in which their occurrence and content are interpreted. While most people simply are surprised by and merely ignore these intrusions, people with OCD ascribe special importance to them. According to Salkovskis, a set of underlying beliefs determines the reaction to negative intrusive experiences in people with OCD. These underlying beliefs are characterized by an exaggerated sense that one is responsible for harm to oneself or others and must act to prevent it. When the intrusive thought is interpreted in this light, the person seeks to reduce the discomfort by engaging in some form of neutralization (i.e., compulsion, suppression, or avoidance). Although the role of responsibility in OCD has been described by a number of previous theorists (Rachman & Hodgson, 1980; McFall & Wollersheim, 1979), Salkovskis refined the concept and elevated it to an explanatory level. This theory suggests that two things happen in OCD. First, patients appraise the intrusive thought as an indication that they are in some way responsible for harm or its prevention. Second, this appraisal elicits neutralizing behavior (overt or covert). If some other form of negative appraisal occurs without the appraisal of responsibility, Salkovskis suggests that neutralization will not take place and the result will be anxiety and/or depression, but not OCD. Thus, the core hypothesis/assumption of this model is that OCD symptoms are neutralizing efforts to address the appraisal of responsibility for harm.

Responsibility has been discussed already in the context of cognitive characteristics of OCD. Neutralization has been defined as "voluntarily initiated and conducted activity which is intended to have the effect of reducing the perceived level of responsibility" (Salkovskis, Richards, & Forester, 1995, p. 285), linking such efforts to responsibility. Included in neutralization are overt behaviors (washing, checking, etc.), as well as mental events, attempts to put things right, thought suppression, and reassurance seeking. Consistent with earlier anxiety-reduction models, Salkovskis et al. (1995) suggested that neutralization effectively reduces the level of perceived responsibility in the short run, but increases the probability of intrusive thoughts in the longer term.

A great deal of research has been generated by this model, much of it supporting the basic contention that responsibility is a primary feature of OCD. Experimental manipulations of responsibility (Lopatka & Rachman, 1995) and refinements in the nature and measurement of obsessional responsibility (Rhéaume, Freeston, Dugas, Letarte, & Ladouceur, 1995; Salkovskis et al., 1995) tend to support the theory. Some revisions and critiques have been proposed. Clark and Purdon (1993) propose that more emphasis should be given to dysfunctional beliefs about controlling thoughts and the costs of not doing so and to the role of depression in actually impairing patients' ability to control their thoughts. They also suggest that neutralization is not so much an essential component in the formation and maintenance of OCD, but develops only when other thought control efforts fail. O'Connor and Robillard (1995) criticize Salkovskis' theory, concluding that people with OCD do not react to an actual feared stimulus, nor to the perceived consequences, but instead to what they imagine might be there, despite a lack of sensory evidence to support the belief. This concept is similar to the thought-action fusion in that what is imagined seems a real possibility. They argue that a faulty inference process is responsible for OCD beliefs, and posit several types of inference errors. For example, rather than forming a hypothesis about an obsessional fear (the table is dirty) and testing it (feeling the table), people with OCD revise the evidence to fit the hypothesis (it must be dirty because I can imagine it dirty). Thus, OCD rituals are attempts to change the "fictional narrative" by altering reality. Further reality testing difficulties are evident in explanations for obsessional beliefs that make reference to superstitions, magical or pseudoscientific explanations, or rely on information that is irrelevant to the current situation. Further, O'Connor and Robillard propose that OCD patients make irrelevant or incidental associations between objects or events and use these to establish links in a way quite similar to sympathetic magic. Several other theories are cognitive in nature, but have been described earlier in the sections on information processing (Enright, 1996) and biological models (Pitman, 1987).

It is clear from the above review that observation of patients with OCD have led to a rich variety of theoretical ideas that involve cognitive conceptualizations of the origins and maintenance of this disorder. Although early cognitive treatments were not derived from theories that were specific to OCD symptoms, more recent studies have examined rational-emotive therapy (RET) and Beckian-derived cognitive therapy based somewhat more directly on some of the above ideas. The effects of these methods are detailed below.

6.17.6.4 Cognitive Treatments for OCD

Several writers have observed the ability of exposure-based therapies to modify thoughts and beliefs, particularly those about risk and danger, merely by requiring clients to remain in anxiety-provoking situations until their fear has subsided. Foa and Steketee (1979) suggested that exposure treatment may not actually correct cognitive deficits, but results in reclassification of some contexts as nondangerous. Whether exposure also modifies other types of cognitions common in those with OCD is an important question only partly addressed in studies by Emmelkamp and Van Oppen. These investigators noted minimal change in the Irrational Beliefs Inventory (IBI) following ERP (Emmelkamp & Beens, 1991; Emmelkamp et al., 1988; Van Oppen et al., 1995), but this measure may not adequately capture the cognitive domains of concern in OCD. The effects of ERP on cognitions requires further study.

Early studies of the effects of cognitive therapies on OCD examined self-instructional training (SIT) and rational-emotive therapy (RET), neither of which was designed especially for OCD. SIT consisted mainly of efforts to modify self-statements. When combined with ERP, SIT did not improve significantly on the outcome of ERP alone (Emmelkamp et al., 1980). Questioning whether SIT was an appropriate cognitive intervention for OCD, Emmelkamp et al (1988) went on to study the effects of RET, consisting of disputing irrational beliefs and making rational analyses of obsessional situations, compared with self-controlled ERP in 18 patients. Both treatments produced significant changes in most measures of OCD symptoms (average improvement 78%), with continued gains six months later (average improvement 94%) when several clients received additional therapy. RET demonstrated some superiority over exposure in improving depressed mood. In a second study of RET, Emmelkamp and Beens (1991) replicated their earlier findings in 21 clients who were randomly assigned to either cognitive therapy or exposure treatment, followed by the addition of exposure treatment for both groups. Significant improvement was evident in both groups, with no differences between them. RET was just as effective as exposure, and interestingly, combined treatment did not have an additive effect. Irrational beliefs (again measured by the IBI) declined somewhat more after RET, and interestingly, required longer to improve (four weeks after treatment) than OC symptoms.

Cognitive treatments using Beck's model (see Beck, Emery, & Greenberg, 1985), in which faulty beliefs and underlying assumptions are examined and challenged, have been examined in case studies of OCD patients. Examples of Beckian cognitive methods are explicated by Van Oppen and Arntz (1994) and by Freeston, Rhéaume, and Ladouceur (1996), particularly for obsessional patients without overt rituals. Kearney and Silverman (1990) reported good outcome when they alternated cognitive therapy and response prevention for a suicidal adolescent who refused exposure. When exposure therapy was ineffective for contamination fears related to skin cancer, Salkovskis and Warwick (1986) found that cognitive therapy altered unrealistic beliefs and produced positive gains, implying possible benefits for OCD patients with overvalued ideas. In a recent study of four patients with checking rituals, cognitive intervention focused on excessive responsibility proved successful without ERP (Ladouceur, Leger, & Rhéaume, 1995).

Two controlled trials have been conducted using Beck's cognitive methods. Cognitive therapy using exposure only in the context of behavioral experiments was compared to exposure alone by Van Oppen et al. (1995). These authors found both treatments equally effective after therapy and at follow-up, but a trend was evident for slightly better outcome with cognitive therapy (see Table 1). This trend can be questioned since the effects of ERP in this study were somewhat lower than observed in some other centers. A second study by Ladouceur, Freeston, Gagnon, Thibodeau, and Dumont (1995) treated 29 ruminators with combined exposure and cognitive therapy in a controlled waiting-list study. The combined treatment produced an 84% success rate and therapeutic gains were maintained at one-year follow-up.

From the above reports, it is evident that treatments designed to correct cognitive distortions common in OCD clients enhance immediate outcome with or without ERP and may also reduce relapse. However, some preliminary research establishing which cognitions are important in the development or maintenance of OCD symptoms, and how to assess these, has not yet been completed. Only with such information can researchers begin to determine whether desirable cognitive changes are best achieved via cognitive treatments or by other (e.g., behavioral, biochemical) methods. In the late 1990s, given the recency of study of cognition and cognitive therapy for OCD, these methods are not routinely included in standard behavioral treatment programs. Informally, however, many clinicians engage in dialogue with clients regarding their beliefs and assumptions as these surface during exposure sessions and interfere with reduction of anxiety or discomfort.

6.17.6.5 Biological Models of OCD

The discovery of effective pharmacological agents for treating OCD produced a surge of interest in biological explanations for the disorder. With the recent development of sophisticated instrumentation, methods of testing and refining hypotheses about neural substrates have produced dramatic advances in our knowledge about the biology of OCD. Current biological models can roughly be organized into two types: neuropharmacological and anatomical. The neuropharmacological explanations involve a disregulation of the serotonergic transmitter system. The anatomical explanations involve the basal ganglia, orbiotfrontal cortex, and their connecting loops. These models are not mutually exclusive.

6.17.6.5.1 The role of serotonin

The most consistent neuropharmacological evidence comes from the results of the administration of serotonin reuptake inhibitors (SRIs) as a treatment for OCD (Barr, Goodman, & Price, 1993). A plethora of evidence now indicates that the SRIs are especially effective in reducing symptoms of OCD (see Section 6.17.6.6). These findings suggest that the serotonergic system may be particularly important in the biology of OCD. Unfortunately, drug response data are not the best source of evidence for testing underlying neurochemical processes.

More direct evidence for serotonergic involvement require studies of peripheral markers of serotonin (i.e., concentrations of known metabolites of serotonin in the blood or cerebrospinal fluid) or pharmacological challenges (i.e., administering a serotonin agonist which provokes a serotonergic response in people whose serotonergic system is dysregulated). Unfortunately, studies of blood platelet measures of serotonin and cerebrospinal fluid concentrations of serotonin metabolites (5-hydroxyindoleacetic acid: 5-HIAA) have been inconsistent in establishing the role of serotonin function in OCD (Barr et al., 1993; Rauch & Jenike, 1993). It is possible that peripheral markers of serotonin in a resting state do not show functional differences between OCD and normal subjects because the body may compensate for serotonin dysregulation (Barr et al.).

The results of pharmacologic challenge studies have produced findings opposite to what would be expected based on the evidence from drug treatment studies. The effectiveness of SRIs suggest OCD patients suffer from a functional deficit in serotonin (Yaryura-Tobias & Bhagavan, 1977). However, administration of serotonin agonists have produced exacerbated OCD symptoms, suggesting that OCD sufferers have a heightened sensitivity to serotonergic stimulation (Tallis, 1995). Strengthening this finding is the reversal of this effect by administration of serotonin receptor antagonists (Benkelfat et al., 1990). Challenge study findings have not been entirely consistent, however. In reviewing this evidence, Barr et al. (1993) conclude that dysregulation of the serotonergic system is probable, but the nature of the dysregulation is unclear. While neuroendocrine responses to challenge suggest hyposensitivity, behavioral findings suggest a hypersensitivity. In light of these inconsistencies, a number of investigators suggest a more complex model of OCD taking into account multiple transmitter systems (Leckman et al., 1995; Rauch & Jenike, 1993).

6.17.6.5.2 Neuroanatomy of OCD

Anatomical models of OCD have focused on selected parts of the brain, specifically the basal ganglia, the frontal cortex, and the limbic structures connecting them. The basal ganglia are a subcortical group of nuclei involved in the regulation of movement, as well as cognitive functioning (Crutcher, 1985). Three structures in the basal ganglia make up the corpus striatum: the caudate nucleus, putamen, and globus pallidus. It is the structure and activity of the corpus striatum which has been most closely associated with OCD. Computerized tomography (CT) studies have found caudate volume to be smaller in OCD subjects than in controls (Luxenburg et al., 1988). Positron emissions tomography (PET) scan studies have found increased metabolic activity in the caudate nucleus of OCD patients during a resting state (Baxter et al., 1987, 1988). Other studies have found increased metabolic activity in the striatum during exposure among OCD cleaners (McGuire, Bench, Marks, Frackowiak, & Dolan, 1994; Zohar, Fux, Taub, & Lusel, 1989), but reduced metabolic activity among compulsive checkers (Cottraux et al., 1996). Cottraux et al. (1996) suggest that the latter converse finding may be due to the ability of compulsive checkers to engage in covert neutralization during the PET scan procedure.

In reviewing the evidence on basal ganglia involvement in OCD, Rauch and Jenike (1993) conclude that there is compelling evidence for a striatal abnormality among OCD patients, but that it is not clear whether this abnormality is primary or secondary. These studies have led to speculation about the role of basal ganglia dysfunction in the development of OCD. Wise and Rapoport (1989) have postulated that OCD is a basal ganglia disorder involving the dysfunctional release of normally inhibited "fixed-action patterns" (innate species-specific responses). In this model it is assumed that the basal ganglia stores motor programs for action which are innate, and that it forms a gating mechanism for their release. The basal ganglia detects stimuli which activate fixed-action patterns, and allows the release of these behaviors which take the form of grooming and safety rituals.

A variety of studies suggest that frontal lobe activity in conjunction with the basal ganglia are implicated in OCD. Neuropsychological assessments (see Tallis, 1995) and EEG studies are suggestive of frontal lobe deficits (Rauch & Jenike, 1993). PET studies have also suggested some type of frontal-striatal abnormality. PET studies and SPECT (single-photon-emission computed tomography) studies have consistently shown increased metabolic activity in the frontal cortex of OCD patients compared to controls (Baxter et al., 1990; Cottraux et al., 1996; Rauch & Jenike, 1993). More recent studies have shown increased metabolic activity to provocative stimulation among OCD patients (Cottraux et al., 1996; Rauch et al., 1994) as well as normal controls (Cottraux et al., 1996). Other PET studies have found that elevated metabolic activity is no longer present following successful treatment with medication or behavior therapy (Baxter et al., 1987, 1992). Evidence of hypermetabolic activity in the orbitofrontal gyri and the basal ganglia has led to theorizing that a fronto-striatal dysfunction is at the core of OCD. Insel (1992) proposed that these regions make up a cortical-striatal-thalamic-cortical circuit that is hyperactive. In this circuit the caudate nucleus sends inhibitory signals to the globus pallidus which then fails to inhibit the thalamus from sending signals to the cortex. The cortex thus receives signals that it would not receive if the circuit were operating normally. Further inhibition by the caudate nucleus prevents the circuit from being interrupted, and therefore the messages continue to reverberate.

Baxter et al. (1990, 1992) have proposed a slightly different way in which this circuit may be dysfunctional. The hyperactivity may result from faulty screening of sensory information. The gating and screening activities of the striatum may not work properly and thereby fail to inhibit impulses/thoughts which "leak" into consciousness and are experienced as obsessions. Compulsions are fixed-action patterns that occur automatically in response to this material and require effortful and conscious suppression and/or neutralization. This effort produces the hyperactivity in the basal ganglia and frontal lobes observed in OCD patients. Other models have postulated a problem with the "comparator mechanism" in the limbic system (Gray, 1984) or in the basal ganglia (Pitman, 1989) which matches sensory information with what is anticipated. When a mismatch occurs, general arousal and corrective procedures are undertaken. However, a defective comparator mechanism might result in faulty evaluation of sensory input (e.g., my hands feel dirty when they are not), and ineffective corrective action (washing my hands to remove dirt that is not there). Because of a faulty comparator, the dirt can only be felt and not removed.

While evidence is strong for some form of frontal-striatal circuit involvement as a neurological substrate of OCD, the etiological significance of this circuit is unclear. It may simply be the biological expression of the disorder and not the cause of it. There is some evidence that hyperactivity in the frontal lobes can be produced in normal subjects by processing obsessional material (Cottraux et al., 1996). Nonetheless, the substantial evidence for biological substrates associated with OCD will undoubtedly foster considerable further research to clarify the mechanisms and roles of the various biochemical substances and neuroanatomical structures.

6.17.6.5.3 *Genetics*

The findings from twin and family studies have suggested some degree of genetic influence in OCD (for reviews see Kolada et al., 1994; Pauls, 1990, 1992; Rasmussen, 1993), but these studies have not always been consistent or without problems. When diagnosis of OCD is used to define concordance, the evidence for genetic transmission is weak or inconsistent (Kolada et al., 1994). However, when subclinical OCD symptoms are used to establish concordance, the evidence is strong for a genetic factor in OCD (Black, Noyes, Goldstein, & Blum, 1992; Carey & Gottesman, 1981). The evidence of concordance for other disorders, especially anxiety disorders, has led some to conclude that what is transmitted in families with OCD is a diathesis for anxiety disorders rather than specifically for OCD (Black et al., 1992). Recent interest has been generated by the hypothesis that at least a subtype of OCD and Tourette's syndrome (TS) are different phenotypic expressions of a common underlying genotype (Pauls, 1990, 1992). As noted earlier, OCD occurs more often in TS patients and their family members, and TS occurs more often in OCD patients and their family members (see Pauls, 1990). The evidence has led some to suggest that these disorders share a common genetic underpinning (Leckman & Chittenden, 1990; Pauls, 1990).

6.17.6.6 Biological Treatments for OCD: Medications and Psychosurgery

6.17.6.6.1 *Medications*

Nearly three decades have passed since the late 1960s when the first reports of the potential efficacy of clomipramine (Anafranil) in the treatment of OCD (Stein, Spadaccini, & Hollander, 1995). Since that time a number of more selective SRIs have been developed and investigated as treatments for OCD, including fluvoxamine (Luvox), fluoxetine (Prozac), and sertraline (Zoloft). A substantial number of clinical and placebo-controlled trials have now been conducted with these medications. Qualitative reviews of the treatment outcome literature (Jenike, 1990; Rasmussen, Eisen, & Pato, 1993; Steketee, 1993), as well as quantitative (meta-analytic) studies (Greist, Jefferson, Kobak, Katzelnick, & Serlin, 1995; Piccinelli, Pini, Bellantuono, & Wilkinson, 1995; Stein et al., 1995), provide a useful overview of the effectiveness of the serotonin transport inhibitors. Based on double-blind, placebo-controlled studies, these reviews conclude that all of these SRIs are effective in the treatment of OCD with at least a moderate improvement occurring in from 55–70% of previously untreated patients (Rasmussen, Eisen, & Pato, 1993; Steketee, 1993). The average effect sizes from meta-analytic studies range as high as 1.48–1.64 (Greist et al., 1995; Stein et al., 1995). Table 1 presents findings for medication studies using the YBOCS as an outcome measure.

Although early studies with clomipramine suggested that its antidepressant effect was responsible for at least some benefit in OCD symptoms (Marks, Stern, Mawson, Cobb, &

McDonald, 1980), more recent studies indicate that SRIs work for OCD patients whether or not they are depressed. Comparisons of SRIs with other non-SRI drugs have all favored the SRIs in the treatment of OCD (Jenike, 1993). A newer SRI, paroxetine (Paxil) has also shown promise in recent studies (Kaye & Dancu, 1994).

Very few studies directly compare the effectiveness of SRIs to each other. Comparisons of the efficacy of CMI versus fluoxetine have failed to find differences (Freeman, Trimble, Deakin, Stokes, & Ashford, 1994; Pigott et al., 1990). However, a meta-analytic comparison found CMI effect sizes (1.53 to 1.84) to be larger than those for fluoxetine (1.34) (Jenike, Baer, & Greist, 1990). Other meta-analytic studies have tended to show a similar preference for CMI (Greist et al, 1995; Stein et al., 1995), although at least one such study has been equivocal on this issue (Piccinelli et al., 1995). (See also Table 1.) Despite these indications that CMI is more effective than the other SRIs, more large-scale direct comparisons of treatment efficacy are needed before drawing definitive conclusions regarding the differential effectiveness of these drugs. There are several reports which suggest the method of administration of SRIs may influence their efficacy. In several clinical series, intravenously administered CMI worked significantly faster than oral methods (Sallee, Pollock, Perel, Ryan, & Stiller, 1989; Koran, Faravelli, & Pallanti, 1994) and, of particular importance, IV administration was effective for treatment of refractory patients (Fallon et al., 1992).

While the effectiveness of SRIs in the treatment of OCD has been established, it should be noted that the response to these medications is not usually a complete cessation of symptoms. In fact, Jenike (1990) estimates that patients treated with SRIs show only a 30–60% reduction in their symptoms, and 30% or more show no improvement at all (Rasmussen et al., 1993). Furthermore, most patients remain "chronically symptomatic." There is no question that the medications improve the symptom picture for patients with OCD, but the SRIs do not represent an ideal solution in the treatment of OCD.

Although the tolerableness of the side effects of SRIs is generally better than that of tricyclic antidepressants (Montgomery & Kasper, 1995), it is still of some concern in OCD. Major side effects include sedation, sexual dysfunction, and weight changes (Rasmussen et al., 1993). Other side effects vary somewhat as a function of the specific medication. Rasmussen et al. conclude that CMI has the most problematic side effects profile. These side effects may compromise patient compliance or necessitate dosage reductions which may make the treatment less effective (Clomipramine Collaborative Study Group, 1991). For CMI the side effects are mainly anticholinergic. There is a slight chance of seizures, especially at higher dosages (DeVeaugh-Geiss, Landau, & Katz, 1989) and partial or total anorgasmia is common (Monteiro, Noshirvani, Marks, & Lelliott, 1987). Although many patients may be willing to tolerate the anorgasmia for the symptom relief provided by CMI, it is a major cause of treatment dropout (Rasmussen et al.). For fluoxetine, up to 50% of patients reported side effects including fatigue, nausea, insomnia, tremor, and sexual dysfunction (Jenike, 1990). Similar side effects have been reported for fluvoxamine (Jenike, 1990), although in direct comparison, CMI had more anticholinergic side effects and more sexual dysfunction while fluvoxamine-treated patients reported more headaches and insomnia (Freeman, Trimble, Deakin, Stokes, & Ashford, 1994). In a comparison of the side effects of CMI and fluoxetine, Jenike et al. (1991) found that only 3% of patients had no side effects from CMI while 43% had no side effects on fluoxetine.

Other problems associated with use of medications include dropout, relapse on or off medication, and inadequate response. Dropout or noncompliance rates for SRIs have not been well studied. Recent reports range from 10 (Tollefson, et al. 1994) to 26% (Freeman et al., 1994). Most studies provide insufficient information to determine dropout rates. In a comparison of four large clinical trials, Greist et al. (1995) reported significantly fewer dropouts for CMI (12.5%) than for other SRIs (23–27%). Little information exists on relapse rates following cessation of SRI treatment. Pato, Zohar-Kadouch, Zohar, and Murphy (1988) found a nearly 90% relapse rate within seven weeks of discontinuation of CMI. Furthermore, the length of treatment was not related to relapse. Likewise Leonard et al. (1991) found an 89% relapse rate when desipramine was substituted for CMI. In two reports on relapse during continued medication treatment, Leonard et al. found an 18% relapse rate two months after successful CMI treatment, and Fontaine and Chouinard (1989) found a 23% relapse rate after one year on fluoxetine.

As noted earlier, a substantial number of OCD patients do not respond or respond marginally to SRI treatment. Although it is not clear who will not respond favorably to SRIs, there are several suggestions in the literature. In Ackerman, Greenland, Bystritsky, Morgensterns, and Katz's (1994) review of predictors from a multicenter clinical trial of over 500 patients, the only predictor of treatment response was age of onset.

Patients who develop OCD later in life showed a significantly better response to SRIs, a finding contrary to that for behavioral treatment where early onset is associated with better outcome (Foa et al., 1983). Additionally, patients with comorbid schizotypal personality disorder do not appear to respond well to SRI treatment (Jenike, 1990). Augmenting SRI treatment for patients who show an inadequate response has been the subject of some research. Evidence suggests that patients with comorbid tic disorders or schizotypal personality who are treatment resistant to SRIs may respond to the addition of neuroleptics (e.g., Goodman, McDougle, Barr, Aranson, & Price, 1993; McDougle, Goodman, Leckman, & Price, 1993).

6.17.6.6.2 Psychosurgery

For patients who do not respond to multiple medication trials or behavior therapy, psychosurgery may be an option. The development of stereotactic techniques appears to have resulted in more precise lesion placement, increasing the safety of psychosurgery techniques and reducing their side effects. Four types of psychosurgery are currently used for treatment of OCD: cingulotomy, capsulotomy, limbic leucotomy, and subcaudate tractotomy (Mindus & Jenike, 1992). No controlled trials of the effectiveness of psychosurgery have been conducted on the uncontrolled studies; improvement rates between 25 and 100% have been reported (Chiocca & Martuza, 1990; Jenike et al., 1991; Baer et al., 1995). These improvement rates should be viewed with caution since most are based on early studies before current behavioral and medication treatments were adequately developed. Whether such high rates of improvement will be observed among patients who have failed to respond to multiple medication trials and behavior therapy remains to be seen. At least one recent study (Baer et al.) reported a 45% improvement rate (substantially improved plus partial treatment responders) following cingulotomy for patients who had adequate but unsuccessful presurgery treatments. Following a thorough review of the available evidence on effectiveness and side effects, Chiocca and Martuza (1990) recommend cingulotomy as the first procedure, based primarily on the low rate of complications following this surgery.

6.17.6.7 Combined Interventions for Treatment of OCD

We have already discussed the effects of combined cognitive and behavioral treatments for OCD: combinations did not usually improve upon either treatment administered alone. Likewise, behavioral treatments have also been combined with pharmacological ones, although to date surprisingly few studies have undertaken controlled comparisons of these combinations. Two studies by Marks and colleagues compared clomipramine to placebo, and behavioral treatment to relaxation (Marks et al., 1980) and to antiexposure instructions (Marks et al., 1988). In the first study, CMI improved mood and OCD symptoms and behavior therapy led to additional improvement on most OCD symptoms, but not on mood. Combined clomipramine plus exposure therapy was only slightly better than drug alone at 10 weeks. After two years (Mawson, Marks, & Ramm, 1982) and six years (O'Sullivan, Noshirvani, Marks, Monteiro, & Lelliott, 1991), improvement in OCD symptoms generally continued. At the latter follow-up, most patients (61%) reported taking medication, but only 25% had continued on clomipramine. After six years, the original CMI-treated group improved more than the placebo group on only one of 16 measures. In Marks et al.'s second study (1988), CMI initially enhanced the effect of exposure but this disappeared with continued exposure treatment and was not evident at follow-up. In both Marks et al. studies (1980, 1988) greater average improvement occurred with exposure compared to CMI, and in the later study 75–80% of patients improved with ERP compared to 22–33% for CMI. Combined treatment led to more subjects who were "much improved" after one year (73% vs. 50% for exposure alone), but not to more average improvement.

A third study compared fluvoxamine to this medication plus exposure and to placebo plus exposure (Cottraux et al., 1989). Behavior therapy had a poorer showing compared to other studies (see above). All three groups improved, with only a slight advantage for combined drug and behavioral treatment at six months. In contrast to the studies by Marks et al., fluvoxamine led to more average improvement than exposure and produced more successes (54% vs. 40%) at post-test (Cottraux et al.). However, one year later, mean improvement was nearly equivalent across treatments and slightly more exposure patients were improved (50% vs. 45%). The combination treatment showed some potentiating effect at six months, but this was minimal after a year. With only three studies conducting a head-to-head comparison of ERP and SRI or SSRI medications, there is clearly inadequate evidence to judge whether combined treatment offers advantages or whether either treatment alone is the best choice for particular patient groups. Recent research by Foa, Liebowitz and colleagues (Foa et al., 1993) suggests that ERP is more effective than clomipramine and more effective than the combination. However, conclusions are premature since this study has not yet been completed and study inclusion criteria were stringent, reducing generalizability of findings for clinical OCD patients.

Meta-analysis of ERP and SRI-medication effects provide some information pertinent to this issue. In 1987, Christensen and colleagues concluded that behavioral and pharmacological interventions were equivalent in efficacy, though they acknowledged that their methodology may have overestimated medication effects (Christensen, Hadzi-Pavlovic, Andrews, & Mattick, 1987). A more recent meta-analysis by van Balkom et al. (1994) concluded that both ERP alone and ERP plus medications was more effective than SRIs on self-rated measures. However, they found no differences when assessor-rated measures were used. A recent review of findings concurred with van Balkom and others that the lack of equivalent measurement instruments makes the task of comparing these treatments and their combinations very difficult and equivocal at best (Stanley & Turner, 1995). With the general adoption of the Yale–Brown Obsessive Compulsive Scale (YBOCS, Goodman et al., 1989), it appears that this problem is being addressed. Table 1 includes several pharmacological, behavioral, and cognitive studies that employed the YBOCS to assess outcome. A visual comparison of findings shows substantial reductions using medications, with post-test scores generally ranging from 14 to 19. Unfortunately, few studies included follow-up results, but these were positive in two studies so long as medications were continued. Post-test scores for ERP treatments (and the single cognitive therapy) ranged from 9 to 17 with consistently stable gains at follow-up. These studies show a range of outcomes, with somewhat more average benefits for the behavioral methods.

6.17.7 FUTURE DIRECTIONS

In the above review, we have discussed a variety of areas of clinical research regarding the psychopathology and treatment of OCD. Although much is now understood about this disorder, there remain significant areas in need of further investigation. For example, the level of insight into obsessions varies across subclinical and clinical populations, and even within subjects depending on the mood state or context, but it is not clear whether some treatments (e.g., cognitive therapies) can change insight or even whether insight influences treatment response. The similarities of some diagnoses to OCD (e.g., bulimia, hypochondriasis) argue for possible underlying pathology, but to date research has only hinted at some shared characteristics (e.g., biological traits, treatment response). In view of the increasingly common use of the term "OC Spectrum" conditions that include impulsive disorders (e.g., gambling, trichotillomania), determination of the exact relationship, if any, of these conditions to OCD is needed before the field becomes attached to conceptual views and treatment tactics that may be inaccurate.

Some findings suggest that there may be subtypes of OCD. For example, one subgroup appears to be early-onset males with tics and/or schizotypal traits and poor insight, who are refractory to common treatment packages such as SRIs and ERP. It is important to determine whether patients with particular combinations of OCD characteristics or comorbidity benefit more from specialized combinations of treatments (e.g., addition of neuroleptics, social skills training, or change in living situation). Correspondingly, in view of cultural similarities and differences in OCD symptom pictures, should modifications be made to psychosocial treatments to accommodate cultural context, and if so how? For example, Williams, Chambless, and Steketee (1996) have suggested some modifications to ERP for African-American patients who showed specialized treatment response patterns. Along the same vein, before we proceed too far with inclusion of significant others in the treatment process, it is important to understand how these individuals influence patients' symptoms before and during therapy. It may be that for some patients, including family members in therapy is detrimental rather than helpful, whereas for others, family involvement may be essential to recovery.

Two areas engendering intensive research in recent years are, first, biological underpinnings of OCD and corresponding pharmacological treatments; and second, cognitive traits and theoretical models, and cognitive psychotherapy. With the availability of more precise capabilities to physically examine chemical and structural mechanisms, researchers have intensified their search for essential factors involved in the pathogenesis of this disorder. Efforts to identify effective biological treatments have also progressed rapidly, although these two areas of research are not always closely connected. Baxter and colleagues' (1992) case reports of neurochemical changes following effective drug and behavioral therapy are extremely interesting but require replications in controlled trials from other centers. Likewise, research on cognitive features and information processing has burgeoned recently with studies ongoing across various centers.

In the search for both biological and cognitive substrates and treatments, there is some tendency to test treatments for which theoretical models and assessment strategies are not fully conceptualized. As a result, treatment components are sometimes underdeveloped. This cart-before-the-horse phenomenon is understandable given the severity and debilitating costs associated with OCD. Nonetheless, thoughtful progression from conceptual models to treatment strategies is needed to ensure that newly developed interventions will work. In the current climate of health care cost containment, research on cost-effective treatments for OCD has, not surprisingly, focused on medication regimens and group interventions. However, it is important to bear in mind that cost-effectiveness may not mean briefer interventions, particularly for long-standing or complex symptom pictures. Further research in this area will certainly be needed in the immediate future.

6.17.8 REFERENCES

Ackerman, D., Greenland, S., Bystritsky, A., Morgenstern, H., & Katz, R. (1994). Predictors of treatment response in obsessive compulsive disorder: Multivariate analyses from a multicenter trial of clomipramine. *Journal of Clinical Psychopharmacology, 14,* 247–254.

Allsopp, M., & Verduyn, C. (1990). Adolescents with obsessive compulsive disorder: A case note review of consecutive patients referred to a provincial regional adolescent psychiatry unit. *Journal of Adolescence, 13,* 157–169.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Baer, L., & Jenike, M. (1992). Personality disorders in obsessive compulsive disorder. *Psychiatric Clinics of North America, 15,* 803–812.

Baer, L., Rauch, S., Ballantine, T., Martuza, R., Cosgrove, R., Cassem, E., Giriunas, I., Mauzo, P., Dimino, C., & Denike, M. (1995). Cingulotomy for intractable obsessive-compulsive disorder: Prospective long-term follow-up of 18 patients. *Archives of General Psychiatry, 52,* 384–392.

Barr, L., Goodman, W., & Price, L. (1993). The serotonin hypothesis of obsessive compulsive disorder. *International Clinical Psychopharmacology, 8,* 79–82.

Barsky, A., Wyshak, G., & Klerman, G. (1986). Hypochondriasis: An evaluation of the *DSM-III* criteria in medical outpatients. *Archives of General Psychiatry, 43,* 493–500.

Baxter, L., Phelps, M., Mazziotta, J., Guze, B., Schwartz, J., & Selin, C. (1987). Local cerebral glucose metabolic rates in obsessive compulsive disorder. A comparison with rates in unipolar depression and in normal controls. *Archives of General Psychiatry, 44,* 211–218.

Baxter, L., Schwartz, J., Bergman, K., Szuba, M., Guze, B., Mazziota, J., Alazraki, A., Selin, C., Feng, H., Munford, P., & Phelps, M. (1992). Caudate glucose metabolic rate changes with both drug and behavior therapy for obsessive compulsive disorder. *Archives of General Psychiatry, 49,* 681–689.

Baxter, L., Schwartz, J., Guze, B., Bergman, K., & Saba, M. (1990). PET imaging in obsessive compulsive disorder with and without depression. *Journal of Clinical Psychiatry, 51,* 61–69.

Baxter, L. R., Schwartz, J. M., Mazziotta, J. C., Phelps, M. E., Pahl, J. J., Guze, B. H., & Fairbanks, L. (1988). Cerebral glucose metabolic rates in nondepressed obsessive-compulsives. *American Journal of Psychiatry, 145,* 1560–1563.

Beck, A. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.

Beck, A., Emery, G., & Greenberg, R. (1985). *Anxiety disorders and phobias: A cognitive perspective*. New York: Basic Books.

Beech, H. R., & Liddell, A. (1974). Decision-making, mood states and ritualistic behaviour among obsessional patients. In H. R. Beech (Ed.), *Obsessional states* (pp. 143–160). London: Methuen.

Beech, H. R., & Vaughn, M. (1978). *Behavioral treatment of obsessional states*. New York: Wiley.

Bellodi, L., Sciuto, G., Diaferia, G., Ronchi, P., & Smeraldi, E. (1992). Psychiatric disorders in the families of patients with obsessive-compulsive disorder. *Psychiatry Research, 42,* 111–120.

Benkelfat, C., Nordahl, T., Semple, W., King, C., Murphy, D., & Cohen, R. (1990). Local cerebral glucose metabolic rates in obsessive compulsive disorder. *Archives of General Psychiatry, 47,* 840–848.

Black, A. (1974). That natural history of obsessional neurosis. In H. R. Beech (Ed.), *Obsessional states* (pp. 19–54). London: Methuen.

Black, D. W., & Blum, N. S. (1992). Obsessive-compulsive disorder support groups: The Iowa model. *Comprehensive Psychiatry, 33,* 65–71.

Black, D., Noyes, R., Goldstein, R., & Blum, N. (1992). A family study of obsessive-compulsive disorder. *Archives of General Psychiatry, 49,* 362–368.

Black, D., Noyes, R., Pfohl, B., Goldstein, R., & Blum, N. (1993). Personality disorder in obsessive compulsive volunteers, well comparison subjects, and their first degree relatives. *American Journal of Psychiatry, 150,* 1226–1232.

Boersma, K., Den Hengst, S., Dekker, J., & Emmelkamp, P. M. G. (1976). Exposure and response prevention: A comparison with obsessive compulsive patients. *Behaviour Research and Therapy, 14,* 19–24.

Boone, K., Ananth, J., Philpott, L., Kaur, A., & Djenderedjian, A. (1991). Neuropsychological characteristics of nondepressed adults with obsessive compulsive disorder. *Neuropsychiatry, Neuropsychology and Behavioral Neurology, 4,* 96–109.

Boulougouris, J. C., Rabavalis, A. D., & Stefanis, C. (1977). Psychophysiological responses in obsessive compulsive patients. *Behaviour Research and Therapy, 15,* 221–230.

Brown, H., Kosslyn, S., Breiter, H., Baer, L., & Jenike, M. (1994). Can patients with obsessive compulsive disorder discriminate between percepts and mental images? A signal detection analysis. *Journal of Abnormal Psychology, 103,* 445–454.

Burnam, M. A., Hough, R. L., Escobar, J. I., Karno, M., Timbers, D. M., Telles, C. A., & Locke, B. Z. (1987). Six-month prevalence of specific disorders among Mexican Americans and non-Hispanic whites in Los Angeles. *Archives of General Psychiatry, 44,* 687–694.

Calvocoressi, L., Lewis, B., Harris, M., Trufan, S., Goodman, W., McDougle, C., & Price, L. (1995). Family accommodation in obsessive-compulsive disorder. *American Journal of Psychiatry, 152,* 441–443.

Carey, G., & Gottesman, I. (1981). Twin and family studies of anxiety, phobic, and obsessive disorders. In D. Klein & J. Rabkin (Eds.), *Anxiety: New research and changing concepts* (pp. 117–136). New York: Raven Press.

Carr, A. (1971). Compulsive neurosis: Two psychophysiological studies. *Bulletin of the British Psychological Society, 24,* 256–257.

Carr, A. (1974). Compulsive neurosis: A review of the literature. *Psychological Bulletin, 81,* 311–318.

Cassano, D., Del Buono, G., & Catapano, F. (1993). The relationship between obsessive compulsive personality and obsessive compulsive disorder: Data obtained by the personality disorder examination. *European Psychiatry, 8,* 219–221.

Cawley, R. (1974). Psychotherapy and obsessional disorders. In H. R. Beech (Ed.), *Obsessional states* (pp. 259–290). London: Methuen.

Chambless, D. L., Gillis, M. M., Steketee G., & Tran, G. Q. (1996). Parental bonding reports of clients with obsessive compulsive disorder and agoraphobia. *Clinical Psychology and Psychotherapy, 3,* 77–85.

Chambless, D. L., & Williams, K. E. (1995). A preliminary study of African Americans with agoraphobia: Symptom severity and outcome of treatment with *in vivo* exposure. *Behavior Therapy, 26,* 501–515.

Chiocca, E., & Martuza, R. (1990). Neurosurgical therapy of the obsessive-compulsive disorder. In M. Jenike, L. Baer, & W. Minichiello (Eds.), *Obsessive-compulsive disorders: Theory and management* (pp. 283–294). St. Louis, MO: Mosby-Year Book.

Choninard, G., Goodman, W., Girest, J., Jenike, M., Rasmussen, S., White, K., Hackett, E., Gaffney, M., & Bick, P. (1990). Results of a double-blind placebo-controlled trial of new serotonin re-uptake inhibitor, sertralline, in the treatment of obssessive-compulsive disorder. *Psychopharmacology Bulletin, 26,* 279–284.

Christensen, H., Hadzi-Pavlovic, D., Andrews, G., & Mattick, R. (1987). Behavior therapy and tricyclic medication in the treatment of obsessive-compulsive disorder: A quantitative review. *Journal of Consulting and Clinical Psychology, 55,* 701–711.

Christensen, K., Kim, S., Dysken, M., & Hoover, K. (1992). Neuropsychological performance in obsessive compulsive disorder. *Biological Psychiatry, 31,* 4–18.

Clark, D., & Bolton, D. (1985). Obsessive compulsive adolescents and their parents: A psychometric study. *Journal of Clinical Psychiatry, 26,* 267–276.

Clark, D., & Purdon, C. (1993). New perspectives for a cognitive theory of obsessions. *Australian Psychologist, 28,* 161–167.

Clomipramine Collaborative Study Group. (1991). Clomipramine in the treatment of patients with obsessive compulsive disorder. *Archives of General Psychiatry, 48,* 730–738.

Cobb, J., McDonald, R., Marks, I. M., & Stern, R. (1980). Marital versus exposure therapy: Psychological treatments of co-existing marital and phobic obsessive problems. *Behavioural Analysis and Modification, 4,* 3–16.

Constans, J., Foa, E., Franklin, M., & Mathews, A. (1995). Memory for actual and imagined events in OC checkers. *Behaviour Research and Therapy, 33,* 665–671.

Coryell, W. (1981). Obsessive-compulsive disorder and primary unipolar depression. *Journal of Nervous and Mental Disease, 169,* 220–224.

Cottraux, J., Gérard, D., Cinotti, L., Froment, J., Deiber, M., Le Bars, D., Galy, G., Millet, P., Labbé, C., Lavenne, F., Bouvard, M., & Mauguière, F. (1996). A controlled positron emission tomography study of obsessive and neutral auditory stimulation in obsessive compulsive disorder with checking rituals. *Psychiatry Research, 60,* 101–112.

Cottraux, J., Mollard, E., Bouvard, M., Marks, I., Sluys, M., Nury, A. M., Douge, R., & Cialdella, P. (1989). A controlled study of fluvoxamine and exposure in obsessive compulsive disorder. *International Clinical Psychopharmacology, 5,* 1–14.

Crino, R., & Andrews, G. (1996). Obsessive compulsive disorder and Axis I comorbidity. *Journal of Anxiety Disorders, 10,* 37–46.

Crutcher, C. (1985). Motor functions of the basal ganglia and diseases of transmitter metabolism. In E. R. Kandel & J. H. Schwartz (Eds.), *Principles of neural science* (2nd ed., pp. 523–535). New York: North-Holland.

Dent, H. R., & Salkovskis, P. M. (1986). Clinical measures of depression, anxiety and obsessionality in nonclinical populations. *Behaviour Research and Therapy, 24,* 689–691.

DeVeaugh-Geiss, J., Landau, P., & Katz, R. (1989). Treatment of obsessive-compulsive disorder with clomipramine. *Psychiatric Annals, 19,* 97–101.

Diaz, S., Grush, L., Sichel, D., & Cohen, L. (1997). Obsessive compulsive disorder in pregnancy and the puerperium. In M. Pato & G. Steketee (Eds.), *Annual review of psychiatry* (Vol. 16, III 97–III 112). Washington, DC: American Psychiatric Press.

Dollard, J., & Miller, N. E. (1950). *Personality and psychotherapy: An analysis in terms of learning, thinking and culture*. New York: McGraw-Hill.

Dulaney, S., & Fiske, A. P. (1994). Cultural rituals and obsessive-compulsive disorder: Is there a common psychological mechanism? *Ethos, 22,* 243–283.

Eisen, J., Rasmussen, S., Goodman, W. et al. (1995, May). *Patterns of remission and relapse: A 2-year prospective study of OCD*. Poster presented at the annual meeting of the American Psychiatric Association, Miami, FL.

Eisen, J., & Steketee, G. (1997). Course of illness in obsessive compulsive disorder. In M. Pato & G. Steketee (Eds.), *Annual review of psychiatry* (Vol. 16, III 73–III 95). Washington, DC: American Psychiatric Press.

Ellis, A. (1962). *Reason and emotion in psychotherapy*. New York: Lyle Stuart.

Emmelkamp, P. M. G., & Beens, H. (1991). Cognitive therapy with obsessive-compulsive disorder: A comparative evaluation. *Behaviour Research and Therapy, 29,* 293–300.

Emmelkamp, P. M. G., de Haan, E., & Hoogduin, C. A. L. (1990). Marital adjustment and obsessive-compulsive disorder. *British Journal of Psychiatry, 156,* 55–60.

Emmelkamp, P. M. G., & DeLange, I. (1983). Spouse involvement in the treatment of obsessive-compulsive patients. *Behaviour Research and Therapy, 21,* 341–346.

Emmelkamp, P. M. G., Hoekstra, R., & Visser, S. (1985). The behavioral treatment of obsessive compulsive disorder: Prediction of outcome at 3.5 years follow-up. In A. Brenner (Ed.), *Psychiatry: The state of the art* (Vol. 4, pp. 265–270). New York: Plenum.

Emmelkamp, P. M. G., Kloek, J., & Blaauw, E. (1992). In P. H. Wilson (Ed.), *Obsessive-compulsive disorders in principles and practice of relapse prevention* (pp. 213–234). New York: Guilford Press.

Emmelkamp, P. M. G., & Kraanen, J. (1977). Therapist-controlled exposure *in vivo* versus self-controlled exposure *in vivo*: A comparison with obsessive-compulsive patients. *Behaviour Research and Therapy, 15,* 491–495.

Emmelkamp, P. M. G., & Kwee, K. G. (1977). Obsessional ruminations: A comparison between thought stopping and prolonged exposure in imagination. *Behaviour Research and Therapy, 15,* 441–444.

Emmelkamp, P. M. G., van den Heuvell, C. V. L., Ruphan, M., & Sanderman, R. (1989). Home-based treatment of obsessive-compulsive patients: Intersession interval and therapist involvement. *Behaviour Research and Therapy, 27,* 89–93.

Emmelkamp, P. M. G., van der Helm, M., van Zanten, B. L., & Plochg, I. (1980). Contributions of self-instructional training to the effectiveness of exposure *in vivo*: A comparison with obsessive compulsive patients. *Behaviour Research and Therapy, 18,* 61–66.

Emmelkamp, P. M. G., Visser, S., & Hoekstra, R. J. (1988). Cognitive therapy vs. exposure *in vivo* in the treatment of obsessive-compulsives. *Cognitive Therapy and Research, 12,* 103–114.

Enright, S. (1991). Group treatment for obsessive-compulsive disorder: An evaluation. *Behavioural Psychotherapy, 19,* 183–192.

Enright, S. (1996). Obsessive compulsive disorder: Anxiety disorders or schizotype? In R. Rapee (Ed.), *Current controversies in the anxiety disorders* (pp. 161–190). New York: Guilford Press.

Enright, S., & Beech, A. (1990). Obsessional states: Anxiety disorders or schizotypes? An information processing and personality assessment. *Psychological Medicine, 20,* 621–627.

Enright, S., & Beech, A. (1993a). Reduced cognitive inhibition in obsessive compulsive disorder. *British Journal of Clinical Psychology, 32,* 67–74.

Enright, S., & Beech, A. (1993b). Further evidence of reduced cognitive inhibition in obsessive-compulsive disorder. *Personality and Individual Differences, 14,* 387–395.

Epsie, C. A. (1986). The group treatment of obsessive-compulsive ritualizers: Behavioral management of identified patterns of relapse. *Behavioural Psychotherapy, 14,* 21–33.

Fallon, B., Campeas, R., Schneier, F., Hollander, E., Feerick, J., Hatterer, J., Goetz, D., Davies, S., & Liebowitz, M. (1992). Open trial of intravenous clomipramine in five treatment-refractory patients with obsessive compulsive disorder. *Journal of Neuropsychiatry and Clinical Neurosciences*, 4, 70–75.

Fallon, B., Javitch, J., Hollander, E., & Liebowitz, M. (1991). Hypochondriasis and obsessive compulsive disorder: Overlaps in diagnosis and treatment. *Journal of Clinical Psychiatry, 52,* 457–460.

Fals-Stewart, W., Marks, A. R., & Schafer, J. (1993). A comparison of behavioral group therapy and individual behavior therapy in treating obsessive-compulsive disorder. *Journal of Nervous and Mental Disease, 181,* 189–193.

Flament, M., Whitaker, A., Rapoport, J., Davies, M., Berg, C., Kalikow, K., Sceery, W., & Shaffer, D. (1988). Obsessive compulsive disorder in adolescence: An epidemiological study. *Annual Progress in Child Psychiatry and Child Development, 27,* 499–515.

Foa, E. B. (1979). Failure in treating obsessive compulsives. *Behaviour Research and Therapy, 17,* 169–176.

Foa, E. B., & Goldstein, A. (1978). Continuous exposure and complete response prevention of obsessive-compulsive disorder. *Behavior Therapy, 9,* 821–829.

Foa, E. B., Ilai, D., McCarthy, P., Shoyer, B., & Murdock, T. (1993). Information processing in obsessive-compulsive disorder. *Cognitive Therapy and Research, 17,* 173–189.

Foa, E. B., & Kozak, M. (1985). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99,* 20–35.

Foa, E. B., Kozak, M., Liebowitz, M., Gorfinkk, K., Campeas, R., Stehk, S., Street, L., Riggs, D., Franklin, M., Davies, S., Hearst, D., Del Benc, D., & Nixon, W. (1993, November). *Treatment of obsessive-compulsive disorder by behavior therapy, clomipramine, and their combination: Preliminary results of a multicenter double-blind controlled trial.* Paper presented at the meeting of the Association for Advancement of Behavior Therapy, Atlanta.

Foa, E. B., & McNally, R. (1986). Sensitivity to feared stimuli in obsessive-compulsives: A dichotic listening analysis. *Cognitive Therapy and Research, 10,* 477–485.

Foa, E. B., & Steketee, G. S. (1979). Obsessive-compulsives: Conceptual issues and treatment interventions. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification* (Vol. VIII). New York: Academic Press.

Foa, E. B., Steketee, G. S., & Grayson, J. B. (1985). Imaginal and *in vivo* exposure: A comparison with obsessive-compulsive checkers. *Behavior Therapy, 16,* 292–302.

Foa, E. B., Steketee, G. S., Grayson, J. B., & Doppelt, H. G. (1983). Treatment of obsessive-compulsives: When do we fail? In E. B. Foa & P. M. G. Emmelkamp (Eds.), *Failures in behavior therapy*. New York: Wiley.

Foa, E. B., Steketee, G. S., Grayson, J. B., Turner, R. M., & Latimer, P. R. (1984). Deliberate exposure and blocking of obsessive-compulsive rituals: Immediate and long term effects. *Behavior Therapy, 15,* 450–472.

Foa, E. B., Steketee, G. S., & Milby, J. B. (1980). Differential effects of exposure and response prevention in obsessive compulsive washers. *Journal of Consulting and Clinical Psychology, 48,* 71–79.

Fontaine, R., & Chouinard, G. (1989). Flouxotine in the long-term treatment of obsessive compulsive disorder. *Psychiatric Annals, 19,* 88–91.

Formea, G., & Burns, L. (1995). Relation between the syndromes of bulimia nervosa and obsessive compulsive disorder. *Journal of Psychopathology and Behavioral Assessment, 17,* 167–176.

Freeman, C., Trimble, M., Deakin, J., Stokes, T., & Ashford, J. (1994). Fluvoxamine versus clomipramine in the treatment of obsessive-compulsive disorder. *Journal of Clinical Psychiatry, 55,* 301–305.

Freeston, M., Ladouceur, R., Thibodeau, N., & Gagnon, F. (1991). Cognitive intrusions in a nonclinical population: I. *Behaviour Research and Therapy, 29,* 585–597.

Freeston, M., Rhéaume, J., & Ladouceur, R. (1996). Correcting faulty appraisals of obsessional thoughts. *Behavior Research and Therapy, 34,* 433–446.

Friedman, S., Hatch, M., Paradis, C. M., Popkin, M., & Shalita, A. R. (1995). Obsessive compulsive disorders in two black ethnic groups: Incidence in an urban dermatology clinic. *Journal of Anxiety Disorders, 7,* 343–348.

Frost, R., & Hartl, T. (1996). A cognitive behavioral model of compulsive hoarding. *Behavior Research and Therapy, 34,* 341–350.

Frost, R., Heimberg, R., Holt, C., Mattia, J., & Neubauer, A. (1993). A comparison of two measures of perfectionism. *Personality and Individual Differences, 14,* 119–126.

Frost, R., Marten, P., Lahart, C., & Rosenblate, R. (1990). The dimensions of perfectionism. *Cognitive Therapy and Research, 14,* 449–468.

Frost, R., & Shows, D. (1993). The nature and measurement of compulsive indecisiveness. *Behavior Research and Therapy, 31,* 683–692.

Frost, R., Steketee, G., Cohn, L., & Griess, K. (1994). Personality traits in subclinical and non-obsessive compulsive volunteers and their parents. *Behavior Research and Therapy, 32,* 47–56.

Goodman, W., McDougle, C., Barr, L., Aronson, S., & Price, L. (1993). Biological approaches to treatment resistant obsessive compulsive disorder. *Journal of Clinical Psychiatry, 54*(Suppl. 6), 16–26.

Goodman, W., Price, L., Rasmussen, S., Mazure, C., Fleischman, R., Hill, C., Heninger, G., & Charney, D. (1989). The Yale–Brown Obsessive Compulsive Scale: I. Development, use, and reliability. *Archives of General Psychiatry, 46,* 1006–1011.

Goodwin, G., Fairburn, C., & Cowen, P. (1987). Dietary changes serotonergic function in women, not men: Implications for the aetiology of anorexia nervosa? *Psychological Medicine, 17,* 839–842.

Goodwin, D., Guze, S., & Robins, E. (1969). Follow up studies in obsessional neurosis. *Archives of General Psychiatry, 20,* 182–187.

Gray, J. (1984). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system.* Oxford, UK: Oxford University Press.

Grayson, J. B., Foa, E. B., & Steketee, G. S. (1982). Habituation during exposure treatment: Distraction versus attention focusing. *Behaviour Therapy and Research, 20,* 323–328.

Grayson, J. B., Steketee, G. S., & Foa, E. B. (1986). Exposure *in vivo* of obsessive-compulsives under distracting and attention-focusing conditions: Replication and extension. *Behaviour Research and Therapy, 24,* 475–479.

Greenburg, D., & Witztum, E. (1994). The influence of cultural factors on obsessive compulsive disorder: Religious symptoms in a religious society. Special Section: Psychiatry and religion. *Israel Journal of Psychiatry and Related Sciences, 31,* 211–220.

Greist, J., Jefferson, J., Kobak, K., Katzelnick, D., & Serlin, R. (1995). Efficacy and tolerability of serotonin transport inhibitors in obsessive compulsive disorder: A meta analysis. *Archives of General Psychiatry, 52*, 53–60.

Guidano, V., & Liotti, G. (1983). *Cognitive processes and emotional disorder*. New York: Guilford Press.

Hafner, R. J. (1982). Marital interaction in persisting obsessive-compulsive disorders. *Australian and New Zealand Journal of Psychiatry, 16,* 171–178.

Hafner, R. (1988). Obsessive compulsive disorder. A questionnaire survey of a self-help group. *International Journal of Social Psychiatry, 34,* 310–315.

Hafner, R. (1992). Anxiety disorders and family therapy. *Australian and New Zealand Journal of Family Therapy, 13,* 99–104.

Hand, I., & Tichatzky, M. (1979). Behavioral group therapy for obsessions and compulsions: First results of a pilot study in trends in behavior therapy. In P.O Sjoden, D. Bates, & W. S. Dockens (Eds.), *Trends in behavior therapy* (pp. 269–297). New York: Academic Press.

Hibbs, E., Hamburger, S., Kruesi, M., & Lenane, M. (1993). Factors affecting expressed emotion in parents of ill and normal children. *American Journal of Orthopsychiatry, 63,* 103–110.

Hibbs, E., Hamburger, S., Lenane, M., Rapoport, J., Kruesi, M., Keysor, C., & Goldstein, M. (1991). Determinants of expressed emotion in families of disturbed and normal children. *Journal of Child Psychology and Psychiatry, 32,* 757–770.

Hiss, H., Foa, E. B., & Kozak, M. J. (1994). Relapse prevention program for treatment of obsessive-compulsive disorder. *Journal of Consulting and Clinical Psychology, 62,* 801–808.

Hodgson, R. J., & Rachman, S. (1972). The effects of contamination and washing in obsessional patients. *Behaviour Research and Therapy, 10,* 111–117.

Hollander, E., Liebowitz, M., Winchel, R., Klumker, A., & Klein, D. (1989). Treatment of body dysmorphic disorder with serotonin reuptake blockers. *American Journal of Psychiatry, 146,* 768–770.

Hollander, E., Neville, D., Frenkel, M., Josephson, S., & Liebowitz. (1992). Body dysmorphic disorder: Diagnostic issues and related disorders. *Psychosomatics, 33,* 156–165.

Honjo, J., Hirano, C., Murase, S., Kaneko, T., Sugiyama, T., Ohtaka, K., Aoyama, T., Takel, Y., Inoko, K., & Wakabayashi, S. (1989). Obsessive compulsive symptoms in childhood and adolescence. *Acta Psychiatrica Scandinavica, 80,* 83–91.

Hoover, C., & Insel, T. (1984). Families of origin in obsessive-compulsive disorder. *The Journal of Nervous and Mental Disease, 172,* 207–215.

Hornsveld, R. H. J., Kraaimaat, F. W., & van Dam-Baggen, R. M. J. (1979). Anxiety/discomfort and handwashing in obsessive compulsive and psychiatric control patients. *Behaviour Research and Therapy, 17,* 223–228.

Hsu, G., Kaye, W., & Weltzin, T. (1993). Are the eating disorders related to obsessive compulsive disorder? *International Journal of Eating Disorders, 14,* 305–318.

Hwu, Y., Yeh, E., & Chang, L. (1989). Prevalence of psychiatric disorders in Taiwan defined by the Chinese Diagnostic Interview Schedule. *Acta Psychiatrica Scandinavia, 79,* 136–147.

Insel, T. (1992). Toward a neuroanatomy of obsessive compulsive disorder. *Archives of General Psychiatry, 49,* 739–744.

Insel, T. R., & Akiskal, H. S. (1986). Obsessive-compulsive disorder with psychotic features: A phenomenological analysis. *American Journal of Psychiatry, 143,* 1527–1533.

Jenike, M. A. (1990). Predictors of treatment failure. In M. A. Jenike, L. Baer, & W. E. Minichiello (Eds.), *Obsessive-compulsive disorders: theory and management* (pp. 306–311). Chicago: Year Book Medical Publishers.

Jenike, M. A. (1993). Obsessive compulsive disorder: Efficacy of specific treatments as assessed by controlled trials. *Pharmacology Bulletin, 29,* 487–499.

Jenike, M. A., Baer, L., Ballantine, T., Martuza, R., Tynes, S., Girivnas, I, Buttolph, M., & Cassen, N. (1991). Cingulotomy for refractory obsessive-compulsive disorder. *Archives of General Psychiatry, 48,* 548–555.

Jenike, M. A., Baer, L., & Greist, J. (1990). Clomipramine versus fluoxetine in obsessive compulsive disorder: A retrospective comparison of side effects and efficacy. *Journal of Clinical Psychopharmacology, 10,* 122–124.

Jenike, M. A., Buttolph, L. Baer, L., Ricciardi, J., & Holland, A. (1989). Open trial of fluoxetine in obsessive-compulsive disorder. *American Journal of Psychiatry, 146,* 909–911.

Jenike, M. A., Hyman, S., Baer, L., Holland, A., Minichiello, W. E., Buttolph, L., Summergrad, P., Seymour, R., & Ricciardi, J. (1990). A controlled trial of fluvoxamine in obsessive-compulsive disorder: Implications for a serotonergic theory. *American Journal of Psychiatry, 147,* 1209–1215.

Karno, M., Golding, J., Sorenson, S., & Burnam, M. (1988). The epidemiology of obsessive-compulsive disorder in five US communities. *Archives of General Psychiatry, 45,* 1094–1099.

Kaye, N., & Dancu, C. (1994). Paroxetine and obsessive compulsive disorder. *American Journal of Psychiatry, 151*(10), 1523.

Kaye, W., Weltzin, T., Hsu, G., Bulik, G., McConaha, C., & Sobkiewicz, T. (1992). Patients with anorexia nervosa have elevated scores on the Yale–Brown Obsessive Compulsive scale. *International Journal of Eating Disorders, 12,* 57–62.

Kearney, C. A., & Silverman, W. K. (1990). Treatment of an adolescent with obsessive-compulsive disorder by alternating response prevention and cognitive therapy: An empirical analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 21,* 39–47.

Kenny, F. T., Mowbray, R. M., & Lalani, S. (1978). Faradic disruption of obsessive ideation in the treatment of obsessive neurosis: A controlled study. *Behavior Therapy 9,* 209–221.

Kirk, J. W. (1983). Behavioural treatment of obsessional-compulsive patients in routine clinical practice. *Behaviour Research and Therapy, 21,* 57–62.

Kolada, J., Bland, R., & Newman, S. (1994). Obsessive compulsive disorder. *Acta Psychiatrica Scandinavica, 89,* 24–35.

Koran, L., Faravelli, C., & Pallanti, S. (1994). Intravenous clomipramine for obsessive-compulsive disorder. *Journal of Clinical Psychopharmacology, 14*(3), 216–218.

Kozak, M., & Foa, E. (1994). Obsessions, overvalued ideas, and delusions in obsessive-compulsive disorder. *Behavior Research Therapy, 32,* 343–353.

Kozak, M., Foa, E., & McCarthy, P. (1987). Assessment of obsessive–compulsive disorder. In C. Last & M. Hersen

(Eds.), *Handbook of anxiety disorders*. Elmsford, NY: Pergamon.

Kozak, M. J., Foa, E. B., & Steketee, G. (1988). Process and outcome of exposure treatment with obsessive–compulsives: Psychophysiological indicators of emotional processing. *Behavior Therapy, 19,* 157–169.

Kringlen, E. (1965). Obsessional neurotics: A long-term follow-up. *British Journal of Psychiatry, 111,* 709–722.

Krone, K. P., Himle, J. A., & Nesse, R. M. (1991). A standardized behavioral group treatment program for obsessive-compulsive disorder: Preliminary outcomes. *Behaviour Research and Therapy, 29,* 627–632.

Kruger, S., Cooke, R., Hasey, G., Jorna, T., & Persad, E. (1995). Comorbidity of obsessive compulsive disorder with bipolar disorder. *Journal of Affective Disorders, 34,* 117–120.

Ladouceur, R., Freeston, M., Gagnon, F., Thibodeau, N., & Dumont, J. (1995). Cognitive behavioral treatment of obsessions and ruminations. *Behavior Modification, 19,* 247–257.

Ladouceur, R., Leger, E., & Rheaume, J. (1995). *Cognitive treatment of compulsive checkers*. Manuscript submitted for publication.

Lavy, E., van Oppen, P., & van den Hout, M. (1994). Selective processing of emotional information in obsessive compulsive disorder. *Behaviour Research and Therapy, 32,* 243–246.

Lazarus, R. (1966). *Psychological stress and the coping process*. New York: McGraw-Hill.

Leckman, J., & Chittenden, E. (1990). Gilles de la Tourette's syndrome and some forms of obsessive-compulsive disorder may share a common genetic diathesis. *Encephale, 16,* 321–323.

Leckman, J., Goodman, W., Anderson, G., Riddle, M., Chappell, P., McSwiggan-Hudin, M., McDougle, C., Scahill, L., Ort, S., Pauls, D., Cohen, D., & Price, L. (1995). Cerebrospinal fluid biogenic amines in obsessive compulsive disorder, Tourette's syndrome, and healthy controls. *Neuropsychopharmacology, 12,* 73–86.

Leonard, H., Swedo, S., Lenane, M., Rettew, D., Cheslow, D., Hamburger, S., & Rapoport, J. (1991). A double-blind desipramine substitution during long-term clomipramine treatment in children and adolescents with obsessive-compulsive disorder. *Archives of General Psychiatry, 48,* 922–927.

Livingston-Van Noppen, B., Rasmussen, S., Eisen, J., & McCartney, L. (1990). Family function and treatment in obsessive-compulsive disorder. In M. Jenike, L. Baer, & W. Minichiello (Eds.), *Obsessive-compulsive disorders: Theory and management* (pp. 325–340). Chicago: Year Book Medical Publishers.

Lo, W. (1967). A follow-up study of obsessional neurotics in Hong Kong Chinese. *British Journal of Psychiatry, 113,* 823–832.

Lopatka, C., & Rachman, S. (1995). Perceived responsibility and compulsive checking: An experimental analysis. *Behaviour Research and Therapy, 6,* 673–684.

Lucey, J., Butcher, G., Clare, A., & Dinan, T. (1994). The clinical characteristics of patients with obsessive compulsive disorder: A descriptive study of an Irish sample. *Irish Journal of Psychiatry, 11,* 11–14.

Luxenberg, J. S., Swedo, S. E., Flament, M. F., Friedland, R. P., Rapoport, J., & Rapoport, S. I. (1988). Neuroanatomical abnormalities in obsessive-compulsive disorder detected with quantitative x-ray computed tomography. *American Journal of Psychiatry, 145,* 1089–1093.

Marks, I. M., Hodgson, R., & Rachman, S. (1975). Treatment of chronic obsessive-compulsive neurosis *in vivo* exposure, a 2 year follow-up and issues in treatment. *British Journal of Psychiatry, 127,* 349–364.

Marks, I. M., Lelliott, P., Basoglu, M., Noshirvani, H., Monteiro, W., Cohen, D., & Kasvikis, Y. (1988). Clomipramine, self exposure and therapist-aided exposure for obsessive compulsive rituals. *British Journal of Psychiatry, 152,* 522–534.

Marks, I. M., Stern, R. S., Mawson, D., Cobb, J., & McDonald, R. (1980). Clomipramine and exposure for obsessive-compulsive rituals. *British Journal of Psychiatry, 136,* 1–25.

Martinot, J., Allilaire, J., Mazoyer, B., Hantouche, E., Huret, J., Legant-Demare, F., Deslauriers, A., Hardy, P., Pappata, S., Baron, H., & Syrota, A. (1990). Obsessive compulsive disorder: A clinical neuropsychological and positron emission tomography study. *Acta Psychiatrica Scandinavica, 82,* 233–242.

Mavissakalian, M., Hamann, M., & Jones, B. (1990a). Correlates of *DSM-III* personality disorder in obsessive compulsive disorder. *Comprehensive Psychiatry, 31,* 481–489.

Mavissakalian, M., Hamann, M., & Jones, B. (1990b). *DSM-III* personality disorders in obsessive compulsive disorder: Changes with treatment. *Comprehensive Psychiatry, 31,* 432–437.

Mawson, D., Marks, I. M., & Ramm, L. (1982). Clomipramine and exposure for chronic obsessive-compulsive rituals: Two year follow-up and further findings. *British Journal of Psychiatry, 140,* 11–18.

McDougle, C., Goodman, W., Leckman, J., & Price, L. (1993). The psychopharmacology of obsessive compulsive disorder: Implications for treatment and pathogenesis. *Psychiatric Clinics of North America, 16,* 749–766.

McElroy, S., Hudson, J., Pope, H., & Keck, P. (1991). Kleptomania: Clinical characteristics and associated psychopathology. *Psychological Medicine, 21,* 93–108.

McFall, M., & Wollersheim, J. (1979). Obsessive compulsive neurosis: A cognitive behavioral formulation and approach to treatment. *Cognitive Therapy and Research, 3,* 333–348.

McGuire, P., Bench, C., Marks, I., Frackowiak, R., & Dolan, R. (1994). Graded activation of symptoms in obsessive compulsive disorder. *British Journal of Psychiatry, 164,* 459–468.

McNally, R., & Kohlbeck, P. (1993). Reality monitoring in obsessive compulsive disorder. *Behavior Research and Therapy, 31,* 249–253.

Mehta, M. (1990). A comparative study of family-based and patient-based behavioral management in obsessive-compulsive disorder. *British Journal of Psychiatry, 157,* 133–135.

Merkel, W., Pollard, C., Wiener, R., & Staebler, C. (1993). Perceived parental characteristics of patients with obsessive compulsive disorder, depression, and panic disorder. *Child Psychiatry and Human Development, 24,* 49–56.

Meyer, V., & Levy, R. (1973). Modification of behavior in obsessive-compulsive disorders. In H. E. Adams & P. Unikel (Eds.), *Issues and trends in behavior therapy*. Springfield, IL: Charles C. Thomas.

Meyer, V., Levy, R., & Schnurer, A. (1974). A behavioral treatment of obsessive-compulsive disorders. In H. R. Beech (Ed.), *Obsessional states*. London: Methuen.

Mills, H. L., Agras, W. S., Barlow, D. H., & Mills, J. R. (1973). Compulsive rituals treated by response prevention. *Archives of General Psychiatry, 28,* 524–527.

Mindus, P., & Jenike, M. A. (1992). Neurosurgical treatment of malignant obsessive compulsive disorder. *Psychiatric Clinics of North America, 15,* 921–938.

Monteiro, W., Noshirvani, H., Marks, I., & Lelliott, P. (1987). Anorgasmia from clomipramine in obsessive compulsive disorder. *British Journal of Psychiatry, 151,* 107–112.

Montgomery, S., & Kasper, S. (1995). Comparison of compliance between serotonin reuptake inhibitors and tricyclic antidepressants: A meta-analysis. *International Clinical Psychopharmacology, 9*(Suppl. 4), 33–40.

Mowrer, O. H. (1960). *Learning theory and behavior*. New York: Wiley.

Neal, A. M., & Turner, S. M. (1991). Anxiety disorders research with African-Americans: Current status. *Psychological Bulletin, 109,* 400–410.

Neziroglu, F., & Yaryura-Tobias, J. (1993). Exposure, response prevention, and cognitive therapy in the treatment of body dysmorphic disorder. *Behavior Therapy, 24,* 431–438.

Niler, E., & Beck, S. (1989). The relationship among guilt, dysphoria, anxiety and obsessions in a normal population. *Behaviour Research and Therapy, 24,* 213–220.

Obsessive Compulsive Cognitions Working Group (1997). Cognitive assessment of obsessive-compulsive disorder. *Behavior Research and Therapy, 335,* 667–681.

O'Connor, K., & Robillard, S. (1995). Inference processes in obsessive compulsive disorder: Some clinical observations. *Behavior Research and Therapy, 33,* 887–896.

Okasha, A., Saad, A., Khalil, A., Dawla, A., & Seif, E. (1994). Phenomenology of obsessive compulsive disorder: A transcultural study. *Comprehensive Psychiatry, 35,* 191–197.

Orloff, L. M., Battle, M. A., Baer, L., Ivanjack, L., Pettit, A. R., Buttolph, M. L., & Jenike, M. A. (1994). Long term follow-up of 85 patients with obsessive-compulsive disorder. *American Journal of Psychiatry, 15,* 441–442.

Orly, J., & Wing, J. (1979). Psychiatric disorders in two African villages. *Archives of General Psychiatry, 36,* 513.

O'Sullivan, G., Noshirvani, H., Marks, I., Monteiro, W., & Lelliott, P. (1991). Six-year follow-up after exposure and clomipramine therapy for obsessive compulsive disorder. *Journal of Clinical Psychiatry, 52,* 150–155.

Pato, M., Zohar-Kadouch, R., Zohar, J., & Murphy, D. (1988). Return of symptoms after discontinuation of clomipramine in patients with obsessive compulsive disorder. *American Journal of Psychiatry, 145,* 1521–1525.

Pauls, D. (1990). Emerging genetic markers and their role in potential preventive intervention strategies. In P. Muehrer (Ed.), *Conceptual research models for preventing mental disorders* (pp. 184–195). Rockville, MD: National Institute of Mental Health.

Pauls, D. (1992). The genetics of obsessive compulsive disorder and Gilles de la Tourette's syndrome. *Psychiatric Clinics of North America, 15,* 759–766.

Pauls, D., & Leckman, J. (1986). The inheritance of Gilles de Tourette's syndrome and associated behaviors. *New England Journal of Medicine, 315,* 993–997.

Pauls, D., Towbin, K., Leckman, J., Zahner, G., & Cohen, D. (1986). Gilles de Tourette's syndrome and obsessive-compulsive disorder: Evidence supporting a genetic relationship. *Archives of General Psychiatry, 43,* 1180–1182.

Persons, J., & Foa, E. (1984). Processing of fearful and neutral information by obsessive compulsives. *Behavior Research and Therapy, 22,* 259–265.

Pfohl, B., & Blum, N. (1991). Obsessive compulsive personality disorder: A review of available data and recommendations for *DSM-IV*. *Journal of Personality Disorders, 5,* 363–375.

Piccinelli, M., Pini, S. Bellantuono, C., & Wilkinson, G. (1995). Efficacy of drug treatment in obsessive-compulsive disorder: A meta-analytic review. *British Journal of Psychiatry, 166,* 424–443.

Pigott, T., Pato, M., Bernstein, S., Grover, G., Hill, J., Tolliver, T., & Murphy, D. (1990) Controlled comparisons of clomipramine and fluoxetine in the treatment of obsessive compulsive disorder: Behavioral and biological results. *Archives of General Psychiatry, 47,* 926–932.

Pitman, R. (1987). Pierre Janet on obsessive compulsive disorder (1903). *Archives of General Psychiatry, 44,* 226–232.

Pitman, R. (1989). Animal models of compulsive behavior. *Biological Psychiatry, 26,* 189–198.

Pollack, J. (1987). Obsessive-compulsive personality: Theoretical and clinical perspectives and recent research findings. *Journal of Personality Disorders, 1,* 248–262.

Rabavilas, A., Boulougouris, J., & Stefanis, C. (1976). Duration of flooding sessions in the treatment of obsessive-compulsive patients. *Behaviour, Research and Therapy, 14,* 349–355.

Rachman, S. (1971). Obsessional ruminations. *Behaviour Research and Therapy, 9,* 229–235.

Rachman, S. (1976a). Obsessive compulsive checking. *Behaviour Research and Therapy, 14,* 269–277.

Rachman, S. (1976b). Obsessional ruminations. *Behaviour Research and Therapy, 9,* 229–235.

Rachman, S. (1976c). The modification of obsessions: A new formulation. *Behaviour Research and Therapy, 14,* 437–443.

Rachman, S. J. (1977). The conditioning theory of fear acquisition: A critical examination. *Behaviour Research and Therapy, 15,* 375–387.

Rachman, S. (1993). Obsessions, responsibility and guilt. *Behaviour Research and Therapy, 39,* 149–154.

Rachman, S., Cobb, J., Grey, S., McDonald, B., Mawson, D., Sartory, G., & Stern, R. (1979). The behavioral treatment of obsessional-compulsive disorders with and without clomipramine. *Behaviour Research and Therapy, 17,* 467–468.

Rachman, S., & DeSilva, P. (1978). Abnormal and normal obsessions. *Behaviour Research and Therapy, 16,* 233–248.

Rachman, S., & Hodgson, R. (1980). *Obsessions and Compulsions*. Englewood Cliffs, NJ: Prentice-Hall.

Rachman, S., Shafran, R., Thordarson, D., Neufeld, H., Sawchuk, C., McIver, J., & Fairbrother, N. (1995). Thought-action fusion in obsessive compulsive disorder. Poster session at the World Congress of Behavioural and Cognitive Therapies, Copenhagen, Denmark.

Rachman, S. Thordarson, D., Shafran, R., & Woody, S. (1995). Perceived responsibility: Structure and significance. *Behaviour Research and Therapy, 33,* 779–784.

Rachman, S., & Wilson, G. T. (1980). *The effects of psychological therapy*. Oxford, UK: Pergamon.

Rapoport, J. (1989, March). The biology of obsessions and compulsions. *Scientific American,* 83–89.

Rapoport, J., Swedo, S., & Leonard, H. (1992). Childhood obsessive compulsive disorder. *Journal of Clinical Psychiatry, 53,* 11–16.

Rasmussen, S. (1993). Genetic studies of obsessive compulsive disorder. *Annals of Clinical Psychiatry, 5,* 241–247.

Rasmussen, S., & Eisen, J. (1989). Clinical features and phenomenology of obsessive compulsive disorder. *Psychiatric Annals, 19,* 67–73.

Rasmussen, S., & Eisen, J. (1990). Epidemiology of obsessive compulsive disorder. *Journal of Clinical Psychiatry, 51,* 10–13.

Rasmussen, S., & Eisen, J. (1992a). The epidemiology and clinical features of obsessive compulsive disorder. *Psychiatric Clinics of North America, 15,* 743–758.

Rasmussen, S., & Eisen, J. (1992b). The epidemiology and differential diagnosis of obsessive compulsive disorder. *Journal of Clinical Psychiatry, 53,* 4–10.

Rasmussen, S., Eisen, J., & Pato, M. (1993). Current issues in the pharmacologic management of obsessive compulsive disorder. *Journal of Clinical Psychiatry, 54*(Suppl. 6), 4–9.

Rasmussen, S., & Tsuang, M. (1986). Clinical characteristics and family history in *DSM-III* obsessive-compulsive disorder. *American Journal of Psychiatry, 143,* 317–322.

Rauch, S., & Jenike, M. A. (1993). Neurobiological models of obsessive compulsive disorder. *Psychosomatics, 34,* 20–32.

Rauch, S., Jenike, M. A., Alpert, N., Baer, L., Breiter, H., Savage, C., & Fischman, A. (1994). Regional cerebral blood flow measured during symptom provocation in obsessive-compulsive disorder using oxygen 15-labeled carbon dioxide and positron emission tomography. *Archives of General Psychiatry, 51,* 62–70.

Reed, G. (1977). Obsessional personality disorder and remembering. *British Journal of Psychiatry, 130,* 177–183.

Reed, G. (1985). *Obsessional experience and compulsive behavior: A cognitive-structural approach.* Orlando, FL: Academic Press.

Rhéaume, J., Freeston, M., Dugas, M., Letarte, H., & Ladouceur, R. (1995). Perfectionism, responsibility and obsessive compulsive symptoms. *Behaviour Research and Therapy, 33,* 785–794.

Riggs, D. S., Hiss, H., & Foa, E. B., (1992). Marital distress and the treatment of obsessive compulsive disorder. *Behavior Therapy, 23,* 585–597.

Robertson, M. (1991). The Gilles de la Tourette's syndrome and obsessional disorder. *International Clinical Psychopharmacology, 6,* 69–84.

Robins, L. Helzer, J., Weissman, M., Ovasehel, H., Gruenberg, E., Burke, J., & Reigner, D. (1984). Lifetime prevalence of specific psychiatric disorders in three sites. *Archives of General Psychiatry, 41,* 949–958.

Rosen, J. C., & Leitenberg, H. (1982). Bulimia nervosa: Treatment with exposure and response prevention. *Behavior Therapy, 13,* 117–124.

Rothenberg, A. (1986). Eating disorder as a modern obsessive compulsive syndrome. *Psychiatry, 49,* 45–53.

Rubenstein, C., Peynircioglu, Z., Chambless, D., & Pigott, T. (1993). Memory in sub-clinical obsessive compulsive checkers. *Behaviour Research and Therapy, 36,* 759–765.

Salkovskis, P. M. (1985). Obsessional compulsive problems: A cognitive behavioral analysis. *Behavioral Research and Therapy, 23,* 571–583.

Salkovskis, P. M. (1989). Cognitive-behavioral factors and the persistence of intrusive thoughts in obsessional problems. *Behaviour Research and Therapy, 27,* 677–682.

Salkovskis, P. M., & Campbell, P. (1994). Thought suppression induces intrusion in naturally occurring negative intrusive thoughts. *Behavioral Research and Therapy, 32,* 1–8.

Salkovskis, P. M., & Harrison, J. (1984). Abnormal and normal obsessions: A replication. *Behavior Research and Therapy, 22,* 549–552.

Salkovskis, P. M., Rachman, S., Ladouceur, R., & Freeston, M. (1992). *Proceedings of the Toronto Cafeteria .*

Salkovskis, P. M., Richards, C., & Forrester, E. (1995). The relationship between obsessional problems and intrusive thoughts. *Behavioral and Cognitive Psychotherapy, 23,* 281–299.

Salkovskis, P. M., & Warwick, H. M. C. (1986). Morbid preoccupations, health anxiety and reassurance: A cognitive-behavioural approach to hypochondriasis. *Behaviour Research and Therapy, 24,* 597–602.

Salkovskis, P. M., & Westbrook, D. (1989). Behavior therapy and obsessional ruminations: Can failure be turned into success? *Behavior Research and Therapy, 27,* 149–160.

Sallee, F., Pollock, B., Perel, J., Ryan, N., & Stiller, R. (1989). Intravenous pulse loading of clomipramine in adolescents with depression. *Psychopharmacology Bulletin, 28,* 114–118.

Shafran, R., Ralph, J., & Tallis, F. (1995). Obsessive-compulsive symptoms and the family. *Bulletin of the Menninger Clinic, 59,* 472–479.

Shafran, R., Thordarson, D., & Rachman, S. (1996). Thought action fusion in obsessive compulsive disorder. *Journal of Anxiety Disorders, 5,* 379–391.

Shahar, A., & Marks, I. M. (1980). Habituation during exposure treatment of compulsive rituals. *Behavior Therapy, 11,* 397–401.

Sher, K., Frost, R., Kushner, M., Crews, T., & Alexander, J. (1989). Memory deficits in compulsive checkers: Replication and extension in a clinical sample. *Behavior Research and Therapy, 27,* 65–69.

Sher, K., Frost, R., & Otto, R. (1983). Cognitive deficits in compulsive checkers: An exploratory study. *Behaviour Research and Therapy, 21,* 357–363.

Simeon, D., Hollander, E., Stein, D., Cohen, L., & Aronowitz, B. (1995). Body dysmorphic disorder in the *DSM-IV* field trial for obsessive compulsive disorder. *American Journal of Psychiatry, 152,* 1207–1209.

Snaith, R., Baugh, S., Clayden, A., Husain, A., & Sipple, M. (1983). The clinical anxiety scale: An instrument derived from the Hamilton anxiety scale. *British Journal of Psychiatry, 141,* 518–523.

Solyom, L., Garza-Perez, J., Ledwidge, & Solyom, C. (1972). Paradoxical intention in the treatment of obsessive thoughts: A pilot study. *Comprehensive Psychiatry, 13,* 291–297.

Stanley, M., & Turner, S. (1995). Current status of pharmacological and behavioral treatment of obsessive-compulsive disorder. *Behavior Therapy, 26,* 163–186.

Stein, D., Spadaccini, E., & Hollander, E. (1995). Meta-analysis of pharmacotherapy trials for obsessive compulsive disorder. *International Clinical Psychopharmacology, 10,* 11–18.

Steketee, G. (1990). Personality traits and disorders in obsessive compulsives. *Journal of Anxiety Disorders, 4,* 351–364.

Steketee, G. (1993). Social support and treatment outcomes of obsessive compulsive disorder at 9-month follow up. *Behavioural Psychotherapy, 21,* 81–95.

Steketee, G., Chambless, D., & Tran, G. (1995, July). *Comorbidity and outcome for behaviorally treated clients with OCD and agoraphobia.* Paper presented at the World Congress of Behavioral and Cognitive Therapy. Copenhagen, Denmark.

Steketee, G., Eisen, J., Dyck, I., Warshaw, M., & Rassmussen, S. (1996). *Long-term course of illness of OCD in an anxiety-disordered clinical population.* Unpublished manuscript.

Steketee, G., & Foa, E. B. (1985). Obsessive-compulsive disorder. In D. H. Barlow (Ed.), *Psychological treatment of adult disorders: A clinical guide* (pp. 68–142). New York: Guilford Press.

Steketee, G., & Frost, R. (1994). Measurement of risk-taking in obsessive-compulsive disorder. *Behavioural and Cognitive Psychotherapy, 22,* 287–298.

Steketee, G., Grayson, J., & Foa, E. (1985). Obsessive compulsive disorder: Differences between washers and checkers. *Behaviour Research and Therapy, 23,* 197–201.

Steketee, G., Henninger, N., & Pollard, C. A. (in press). Predicting treatment outcome for OCD: Effects of comorbidity. In W. K. Goodman, M. Rudorfer, & J. D. Maser (Eds.), *Treatment challenges in obsessive compulsive disorder.* Rahway, NJ: Earlbaum.

Steketee, G., Quay, S., & White, K. (1991). Religion and guilt in obsessive compulsive disorder patients. *Journal of Anxiety Disorders, 5,* 359–367.

Steketee, G., & Shapiro, L. (1993). Obsessive compulsive disorder. In A. S. Bellack & M. Hersen (Eds.), *Handbook of behavior therapy in the psychiatric setting* (pp. 99–127). New York: Plenum.

Steketee, G., & Shapiro, L. J. (1995). Predicting behavioral treatment outcome for agoraphobia and obsessive compulsive disorder. *Clinical Psychology Review, 15,* 317–346.

Stern, R. S. (1978). Obsessive thoughts: The problem of therapy. *British Journal of Psychiatry, 133,* 200–205.

Tallis, F. (1995). The characteristics of obsessional thinking: Difficulty demonstrating the obvious? *Clinical Psychology and Psychotherapy, 2,* 24–39.

Taylor, C. J., & Sholomskas, D. E. (1993, March). *Group exposure and response prevention for OCD.* The annual meeting of the Anxiety Disorders Association of America, Santa Monica, CA.

Teasdale, J. D. (1974). Learning models of obsessional compulsive disorder. In H. R. Beech (Ed.), *Obsessional states*. London: Methuen.

Thornicroft, G., Colson, L., & Marks, I. (1991). An in-patient behavioural psychotherapy unit description and audit. *British Journal of Psychiatry, 158,* 362–367.

Tollefson, G., Rampey, A., Potvin, J., Jenike, M., Rush, J., Dominguez, R., Koran, L., Shear, M., Goodman, W., & Genduso, L. (1994). A multicenter investigation of fixed dose fluoxetine in the treatment of obsessive compulsive disorder. *Archives of General Psychiatry, 51,* 559–567.

Turner, S., Beidel, D., & Stanley, M. (1992). Are obsessional thoughts and worry different cognitive phenomena? *Clinical Psychology Review, 12,* 257–270.

Turner, S. M., Hersen, M., Bellack, A. S., Andrasik, F., & Capparell, H. V. (1980). Behavioral and pharmacological treatment of obsessive-compulsive disorders. *Journal of Nervous and Mental Disease, 168,* 651–657.

Tynes, L., Salins, C., Skiba, W., & Winstead, D. K. (1992). A psychoeducational and support group for obsessive-compulsive disorder patients and their significant others. *Comprehensive Psychiatry, 33,* 197–201.

Tynes, L., White, K., & Steketee, G. (1990). Toward a new nosology of OCD. *Comprehensive Psychiatry, 31,* 465–480.

Vaisaner, E. (1975). Psychiatric disorders in Finland. *Acta Psychiatrica Scandinavica, 62,* (Suppl. 263), 27.

Valleni-Basile, L., Garrison, C., Jackson, K., Waller, J., Makeowa, R., Addy, C., & Cuffe, S. (1994). Frequency of obsessive compulsive disorder in a community sample of young adolescents. *Journal of American Academy of Child and Adolescent Psychiatry, 6,* 782–791.

Van Balkom, A., van Oppen, P., Vermeulen, A., van Dyck, R., Nauta, M., & Vorst, H. (1994). A meta-analysis on the treatment of obsessive compulsive disorder: A comparison of antidepressants, behavior, and cognitive therapy. *Clinical Psychology Review, 14,* 359–381.

Van Noppen, B., Pato, M. T., Marsland, R., & Rasmussen, S. (1995). *A time-limited behavioral group for treatment of OCD*. Unpublished manuscript.

Van Noppen, B. L., Steketee, G. S., McCorkle, B., & Pato, M. T. (1996). *Group and family treatment of obsessive compulsive disorder*. Manuscript submitted for publication.

Van Oppen, P., & Arntz, A. (1994). Cognitive therapy for obsessive-compulsive disorder. *Behaviour Research and Therapy, 32,* 79–87.

Van Oppen, P., deHaan, E., van Balkom, A., Spinhoven, P., Hoogduin, K., & van Dyck R. (1995). Cognitive therapy and exposure *in vivo* in the treatment of obsessive compulsive disorder. *Research and Therapy, 4,* 379–390.

Vitiello, B., & de Leon, J. (1990). Dysmorphophobia misdiagnosed as obsessive compulsive disorder. *Psychosomatics, 31,* 220–222.

Warren, R., & Zgourides, G. (1991). *Anxiety disorders: A rational emotive perspective*. New York: Pergamon.

Watts, F. N. (1971). Habituation model of systematic desensitization, *Psychological Bulletin, 86,* 627–637.

Wegner, D. (1989). *White bears and other unwanted thoughts*. New York: Viking Penguin.

Weissmann, M., Bland, R., Canino, G., Greenwald, S., Hwo, H., Lee, C., Newman, S., Oakley-Browne, M., Rubio-Stipek, M., Wickramaratne, P., Wittchen, H., & Eng-Kung, Y. (1994). The cross national epidemiology of obsessive compulsive disorder. *Journal of Clinical Psychiatry, 55,* 5–10.

Williams, K. E., Chambless, D. L., & Steketee, G. S. (1996). *Behavioral treatment of OCD in African Americans: Clinical issues*. Manuscript submitted for publications.

Wilner, A., Reich, T., Robins, I., Fishman, R., & van Doren, T. (1976). Obsessive-compulsive neurosis. *Comprehensive Psychiatry, 17,* 527–539.

Wise, P., & Rapoport, J. (1989). Obsessive compulsive disorder: Is it basal ganglia dysfunction? In J. Rapoport (Ed.), *Obsessive-compulsive disorder in children and adolescents*. Washington, DC: American Psychiatric Press.

Yaryura-Tobias, J., & Bhagavan, H. (1977). L-Typtophan in obsessive compulsive disorders. *American Journal of Psychiatry, 134,* 1298–1299.

Yaryura-Tobias, J., Campisi, T., McKay, D., & Neziroglu, F. (1995). Schizophrenia and obsessive compulsive disorder: Shared aspect of pathology. *Neurology, Psychiatry and Brain Research, 3,* 143–148.

Yaryura-Tobias, J., & Neziroglu, F. (1983). *Obsessive-compulsive disorders: Pathogenesis-diagnosis-treatment*. New York: Marcel Dekker.

Zohar, J., Fux, M., Taub, M., & Insel, T. (1989). *Specificity of clomipramine response in obsessive compulsive disorder*. Presented at the 28th annual meeting of the American College of Neuropsychopharmacology, December 15, Maui, HI.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.18

Panic Disorder and Agoraphobia

PAUL M. SALKOVSKIS

University of Oxford, Warneford Hospital, UK

6.18.1 INTRODUCTION 400

6.18.2 DEVELOPMENT OF THE PANIC CONCEPT 401
- *6.18.2.1 Lay and Psychiatric Conceptualizations* 401
- *6.18.2.2 The Development of Psychological Conceptualizations of Panic* 402

6.18.3 PREVALENCE OF PANIC 403
- *6.18.3.1 Panic Attacks in Nonclinical Populations ("Nonclinical Panic")* 403
- *6.18.3.2 Panic Disorder* 404
- *6.18.3.3 Correlates of Panic Disorder* 404
- *6.18.3.4 Comorbidity* 404
- *6.18.3.5 Natural History of Panic* 405
- *6.18.3.6 Other Issues* 405

6.18.4 THEORIES OF PANIC 406
- *6.18.4.1 The Original Biological Hypothesis: Klein's View* 406
- *6.18.4.2 The Neuroanatomical Hypothesis* 408
- *6.18.4.3 Neurochemical Hypotheses: Noradrenaline and the Locus Coeruleus* 410
- *6.18.4.4 Physiological Accounts of Panic: Panic as a "Hyperventilation Syndrome"* 410
  - *6.18.4.4.1 Evidence for an association between panic and hyperventilation* 411
  - *6.18.4.4.2 Hyperventilation as a CAUSE or EFFECT of panic* 412

6.18.5 PSYCHOLOGICAL THEORIES OF PANIC 413
- *6.18.5.1 Barlow's Model: False Alarms and Learned Alarms* 413
- *6.18.5.2 The Cognitive Hypothesis of Panic* 414
- *6.18.5.3 The Importance of the Idiosyncratic Nature of Catastrophic Misinterpretations* 415
  - *6.18.5.3.1 Panic, cognitions, and avoidance* 416
- *6.18.5.4 Agoraphobia* 418
  - *6.18.5.4.1 Correlates of agoraphobia* 419
- *6.18.5.5 The Cognitive Account and Research into Panic* 420
- *6.18.5.6 The Relevance of Psychobiological Factors to the Cognitive Hypothesis of Panic* 421
  - *6.18.5.6.1 The cognitive link between panic and hyperventilation-induced symptoms* 421

6.18.6 EVALUATING THE COGNITIVE HYPOTHESIS 423

6.18.7 TREATMENT 425
- *6.18.7.1 The Theoretical Roots of Treatment* 425
  - *6.18.7.1.1 Cognitive therapy for panic: some general issues* 427
  - *6.18.7.1.2 Clinical implementation of cognitive-behavioral therapy* 427
  - *6.18.7.1.3 Effectiveness of cognitive-behavioral therapy* 430

6.18.8 CONCLUSION 432

6.18.9 REFERENCES 432

6.18.1 INTRODUCTION

Panic attacks are characterized by sudden and apparently inexplicable increases in anxiety accompanied by a range of intense bodily sensations. Over the last two decades, clinical problems characterized by recurrent panic attacks have become a major focus of biological and psychological research into anxiety. Probably the main reason for this upsurge in interest in panic was the nearly simultaneous development of effective psychological and biological treatments for a range of anxiety disorders, coupled with the growing awareness of the relevance of panic attacks to treatment outcome (Rachman, 1990).

The strong association between panic attacks and avoidance behavior, particularly in agoraphobia, has led many writers to the conclusion that panic attacks are the principal *cause* of agoraphobia (Barlow, 1988; Klein, Ross, & Cohen, 1987; Rachman, 1990; Thyer, 1986; Thyer & Himle, 1985). Even when not accompanied by agoraphobic avoidance, panic attacks are associated with considerable distress as well as social and occupational handicap (i.e., in Panic Disorder). The importance of the phenomenon of panic goes beyond these conditions, however; it has been demonstrated that panic attacks are prevalent in all other anxiety disorders (Barlow et al., 1985). A number of studies have also shown that panic also occurs relatively frequently in nonclinical subjects (Margraf & Ehlers, 1988; Norton, Dorward, & Cox, 1986; Norton, Harrison, Hauch, & Rhodes, 1985), and that subjects with infrequent (nonclinical) panic share many of the characteristics of patients diagnosed as suffering from Panic Disorders. Such observations have led to the suggestion that infrequent panic represents the nonclinical end of an agoraphobic continuum (Rachman, 1990), challenging earlier suggestions that agoraphobia is discontinuous from normal fear (Arrindell, 1980; Roth, 1959).

Current psychological thinking tends to assume that panic and agoraphobic avoidance may have a common basis in psychological factors which are normally distributed in the general population, and that anxious patients are people who are especially vulnerable to the development of specific anxiety disorders as a result of being at the extreme end of such a continuum or several continua. The basis of the putative underlying dimensions seems likely to be cognitive (McNally, 1990, 1994; Rachman, 1990; the psychological basis of dimensional approaches to panic is discussed in greater detail in Salkovskis, 1988a).

Although panic attacks have been recognised since the nineteenth century (albeit in various forms and with a range of descriptive labels), the importance of panic attacks was only formally recognised on the introduction of the third edition of the *Diagnostic and statistical manual of mental disorders* (*DSM-III*) of the American Psychiatric Association (American Psychiatric Association, 1980). In *DSM-III*, the presence of relatively frequent panic attacks, some of which were spontaneous (i.e., uncued and unexpected) was defined as pathognomonic of two conditions in the anxiety disorders grouping; Panic Disorder and Agoraphobia with Panic.

The *DSM* description of panic (currently in the form of *DSM-IV*) has become the standard way of operationally defining panic attacks and Panic Disorder, with and without agoraphobia. Table 1 gives the *DSM-IV* criteria for panic. There are a number of shortcomings of this definition. Perhaps the most important of these concerns the definition of panic frequency; it is possible to obtain a diagnosis of Panic Disorder as a result of a single panic attack followed by a period of at least one month's apprehension of having another. This definition could have the result that some control subjects in investigations into panic may experience more frequent panic attacks than some patients in the Panic Disorder groups. The obvious solution (and the one employed by many research groups) is to supplement *DSM-IV* definitions of Panic Disorder with the requirement that panic subjects should be currently experiencing panic attacks at a predefined frequency.

As is clear from the criteria in Table 1 that the strong association between panic attacks and avoidance behavior, particularly in agoraphobia, has led many writers to the conclusion that panic attacks are the principal cause of agoraphobia (Barlow, 1988; Klein et al., 1987; Rachman, 1990; Thyer, 1986; Thyer & Himle, 1985). Even when the diagnosis is of agoraphobia without a history of panic attacks this depends on the "presence of agoraphobia related to fear of developing panic-like symptoms (e.g., dizziness or diarrhea)." Although it is recognized here that panic and panic-like symptoms are important for (and often central to) the experience of a substantial proportion of agoraphobics, the attribution of all agoraphobia to panic and panic-related symptoms may be premature (Salkovskis & Hackmann, 1997). Nevertheless, the last two decades has seen the concept of agoraphobia fade and become subsidiary to panic.

Table 1 Criteria for panic disorder with and without agoraphobia.

For a diagnosis of Panic Disorder, at least two unexpected attacks must occur, characterized by at least four of the symptoms listed below

Also requires:

Recurrent and unexpected panic attacks and at least one of the attacks is followed by at least one month of persistent concern about having additional attacks, or worry about the implications of the attack or its consequences (such as losing control or having a heart attack) or a significant change in behavior related to the attacks

Panic must also be unexpected, i.e., did not occur immediately before or on exposure to a situation that almost always caused anxiety, or produced by another disorder (social phobia, obsessive-compulsive disorder, post-traumatic stress disorder, or separation anxiety disorder)

Symptoms (criteria specify an increase in at least four in at least one attack)

Physical
- Shortness of breath
- Choking
- Palpitations or tachycardia
- Chest pain or discomfort
- Sweating
- Nausea or abdominal distress
- numbness or tingling sensations
- depersonalization or derealization
- flushes or chills
- trembling or shaking
- feeling dizzy, lightheaded, unsteady, faint

Mental
- Fear of dying
- fear of losing control or going crazy

Other characteristics

At least four or more symptoms developed abruptly and reached a peak within 10 minutes of the beginning of the first symptom noticed

Can be associated with avoidance behavior (especially agoraphobia and anticipatory anxiety)

Onset can be associated with physical disorder or factors; diagnosis depends on the persistence beyond the cessation of precipitating organic factors (e.g., Hyperthyroidism, caffeine intoxication, etc.)

Painc Disorder with Agoraphobia: as above, but in addition

Anxiety about being in places or situations from which escape might be difficult (or embarassing) or in which help may not be available in the event of having an unexpected or situationally predisposed panic attack or panic-like symptoms. Agoraphobic fears typically involve characteristic clusters of situations that include being outside the home alone; being in a crowd or standing in a line; being on a bridge; and traveling in a bus or train. The situations are avoided or endured with marked distress or with anziety about having a panic attack or panic-like symptoms, or require the presence of a companion

6.18.2 DEVELOPMENT OF THE PANIC CONCEPT

6.18.2.1 Lay and Psychiatric Conceptualizations

The term "panic" is said to be derived from the Greek *to panikon deima*, which was a description of the effects on lone travelers of sudden noises made by the God Pan (Baker, 1989). There are a number of literary allusions to panic from Classical Greek times onwards. Most descriptions seem to concern identifiable responses to extraneous stimuli and situations, such as the effects of apparent reverses during battle or of sudden noises. More "scientific" descriptions of anxiety began to appear during the nineteenth century (notably those of Charles Darwin and William James), although these do not appear to have been related to panic as such. Modern writers have suggested that early descriptions of a range of syndromes (such as neurasthenia, effort syndrome, etc.) may have been descriptions of Panic Disorder (Skerrit, 1983). While this is possible, such attempts to pinpoint early descriptions of Panic Disorder tend to be overinclusive. For example, careful examination of early and late descriptions of "Neurasthenia" (e.g. Beard, 1869; Cohen & White, 1951) indicate a quite different pattern of symptoms, probably including characteristic subgroups of anxious and depressed patients whose main complaint was weakness on exercise, often more closely resembling modern accounts of Chronic Fatigue Syndrome. That there may have been overlaps with modern diagnoses is not impossible, but the presence of these type of accounts cannot be taken as evidence of the ubiquity of the diagnosis of Panic Disorder.

More relevant are explicitly psychiatric accounts of anxiety problems such as that of Westphal (1872), who identified a specific constellation of symptoms which were characteristic of what we now recognize as panic associated with Agoraphobia. It fell to Freud (1894, 1895) to specifically identify episodic and "anxiety attacks" (*Angstanfaelle*) which were not triggered by specific situations and tended to be unexpected; he described such episodes as being important symptoms in defining "anxiety neurosis." Anxiety neurosis was said to be characterized by acute dread and fear of death, night terrors, phobias, vertigo, parasthesias, hypervigilance, excessive arousal, and anxious expectation (Lewis, 1967). Freud defined anxiety attacks, which were often preceded by periods of generally heightened anxiety, as suddenly breaking through and being accompanied by a range of physical sensations similar to those recognized today by the *DSM-III-R* definition. Anxiety attacks continued to be intermittently described by psychiatric writers such as Diethelm and Fenichel, but invariably as a symptom of underlying anxiety (Baker, 1989) rather than as defining features.

The 1960s were marked by an increased prominence afforded to panic attacks, with the importance of the concept of panic being championed by Klein (1964), who described what he believed to be the specific responsiveness of anxiety characterized by panic to imipramine but not to phenothiazines (this work is often wrongly described as being a comparison of imipramine compared with benzodiazepines, based on the use of the term "tranquilizers"). Subsequently, this "pharmacological dissection" was extended to show that benzodiazepines benefited those with generalized and phobic anticipatory anxiety but not those with panic (Zitrin, 1983; Zitrin, Woerner, & Klein, 1981). Others enthusiastically took up the view that panic attacks were a distinct marker of a different, possibly endogenous type of anxiety (e.g., Snaith 1968). At about the same time, Pitts and McClure (1967) demonstrated that anxiety attacks could be precipitated by intravenous infusions of the alkaline substance, sodium lactate. Later work showed that this effect was specific to patients with Panic Disorder as opposed to other anxiety disorders (Klein, 1981), and that prior administration of imipramine blocked the response to lactate (Rifkin, Klein, Dillon, & Levitt, 1981). Taken together, this data was believed to indicate that panic attacks were the core manifestation of a specific biological dysfunction and that psychological aspects such as agoraphobic avoidance were entirely secondary (Klein, 1980; Klein et al., 1987; Sheehan, 1982). This view on the primacy of panic attacks was in sharp contrast to Freudian conceptualizations current in the 1970s, in which panic was regarded as a secondary phenomenon, with free-floating anxiety (unfocused anxiety arising from subconscious conflicts) being the supposed origin of panic.

An important corollary of Klein's biological hypothesis was that panic anxiety unlike anticipatory or phobic anxiety was, by virtue of its endogenous nature, *unaffected by psychological influences*. This assertion extended to statements concerning the ineffectiveness of psychological treatments, including exposure (Zitrin, Klein, & Woerner, 1980). Klein's view was the basis for the definition adopted by *DSM-III* (American Psychiatric Association, 1980) and its modification, *DSM-III-R* (American Psychiatric Association, 1987). Not surprisingly, psychological theorists and practitioners began to challenge the primacy of endogenous and biological factors in panic and its treatment.

6.18.2.2 The Development of Psychological Conceptualizations of Panic

With the early interest in panic arising from biological findings, those involved in psychological conceptualization were slow to acknowledge the diagnostic importance of panic, and some still dispute the value of the label (Marks, 1987a, 1987b). Perhaps the most influential early *psychological* account of the nature of panic attacks was that of Lader and Mathews (1968,1970). Drawing on their previous theorizing concerning the tendency of arousal to inhibit habituation, and, in more extreme instances, to result in sensitization, they postulated that panic attacks were the consequence of an inhibitory function which occurred if the arousal/sensitization process proceeded too far. Thus, the function of panic attacks was said to be to dampen down physiological responses in an "over excited" system. Although this hypothesis did not receive support from subsequent research, it was influential because it introduced the idea of the interaction between psychological and physiological factors in panic. Aspects of this view can be found in Barlow's hypothesis concerning the role of background stress (Barlow, 1988).

The later emphasis on cognitive conceptualizations resulted in a number of psychological hypotheses intended to account for panic attacks occurring in association with agoraphobia, based on the notion that panic was the result of "fear of fear." The most important advocates of this position are Goldstein and Chambless (1978), Chambless and Goldstein (1980), Chambless and Gracely (1989), and Weekes (1977). The fear of fear hypothesis is based on the idea of an ascending spiral in which the patient who enters or anticipates confronting an anxiety provoking situation becomes afraid of the possibility of experiencing fear (e.g., Foa, Steketee, & Young, 1984). Clark (1988) has pointed out that earlier versions of the fear of fear hypothesis tend to over predict the occurrence of the panic spiral; in panic prone subjects, any occurrence of anxiety should result in panic. Later versions of this hypothesis propose that fear of the consequences of anxiety and panic may also play an important part (e.g., Chambless & Gracely, 1989). With later modifications, the fear of fear view closely resembles the cognitive hypothesis of panic discussed below. These modifications concern reconceptualizing fear of fear to fear of the *consequences* of fear (e.g., patients may fear that anxiety will put a dangerous strain on their heart).

6.18.3 PREVALENCE OF PANIC

The importance of the phenomenon of panic is highlighted by epidemiological findings that not only is Panic Disorder *per se* very common, but also that a considerable proportion of nonclinical subjects surveyed experience panic attacks, and that panic attacks frequently coexist with a range of other disorders, especially (but not exclusively) other anxiety disorders. In considering the prevalence of panic, it is important to distinguish between the prevalence of occasional *panic attacks* (often described as nonclinical panic when noted in the general population) as opposed to *Panic Disorder*, which is an operationally defined diagnosis in *DSM-IV*; see below). A number of large-scale studies have been carried out on the prevalence of Panic Disorder with and without agoraphobia (Bourdon et al., 1988, Myers, Weissman, Tischler, et al., 1984; von Korff, Eaton, 1989; von Korff, Eaton, & Keyl, 1985; Wittchen, 1986). Other investigations have been carried out into the prevalence of panic attacks *per se* (Margraf & Ehlers, 1988; Norton et al., 1986; Norton et al., 1985). Taken together, these studies are comprehensive and representative enough to allow some confidence in current estimates of prevalence and demographic factors.

Bebbington (personal communication) described epidemiological studies as being an exercise in "counting swans on the lawn." There is some value in doing so, but the exercise does not necessarily give accurate information as to how the swans got there and what might happen to them subsequently. Instead, the preoccupation tends to be with deciding whether swans are truly present or whether they may be particularly attractive ducks. In panic, as in most areas, there are few longitudinal studies which might help provide more definitive clarifications of issues of mechanism. Cross-sectional studies cannot provide unambiguous information in this respect, as they are vulnerable to cohort effects, in which changes in prevalence and comorbidity in different age groups reflect samples subject to variations in critical influences and naturally occurring selection processes.

6.18.3.1 Panic Attacks in Nonclinical Populations ("Nonclinical Panic")

Norton et al. (1985) found that, as measured by questionnaire, 34.5% of 186 students screened had experienced a panic attack in the preceding year, and that 2.2% had experienced at least three such attacks in the preceding two weeks. In a second questionnaire study with some methodological refinements, Norton et al. (1986) obtained figures of 35.9% and 3.1% for the same criteria as in the earlier study, this time among a sample of 256 students. Margraf and Ehlers (1988) conducted two studies intended to replicate the findings from Norton's group in a German sample, to determine the validity and reliability of the questionnaire method as compared to structured clinical interview methodology, and to evaluate the response of nonclinical panickers to laboratory procedures. Margraf and Ehlers were able to replicate the questionnaire results in both of their studies, obtaining higher rates for the year prevalence of panic (46% and 59%) and similar rates for three panic attacks in three weeks (2% and 1%). However, analysis of agreement between questionnaire and interview data indicated that the questionnaire tended to be systematically overinclusive, and that many (approximately 50%) of those identified by the questionnaire as panickers had a milder, subthreshold variant of the problem as assessed by interview. The differentiation of nonclinical panic vs. Panic Disorder presents a number of problems of definition and comparability; nevertheless, as a paradigm for the investigation of the processes involved in panic attacks, it seems particularly promising, especially if there is a psychological dimension underpinning both infrequent panic in normal subjects and frequent panic in Panic Disorder as proposed above. That is, the former may be a milder version of the processes involved in the latter. In studies carried out by both the German group and the Norton group, it has been found that the characteristics of the nonclinical panickers were similar to those identified in clinical samples with respect to the response to procedures such as lactate infusions and scores on questionnaire measures of cognition and avoidance, and so on.

6.18.3.2 Panic Disorder

The largest and most influential studies are those carried out as part of the Epidemiologic Catchment Area (ECA) surveys, carried out in five centers in the USA. Area probability samples (i.e., samples intended to be representative of the characteristics of the populations as a whole) of 3000 people were selected at each site (i.e., a total of 15 000 people). Response rates of 75–80% were obtained; subjects were interviewed using an instrument called the Diagnostic Interview Schedule (Eaton, Holtzer, von Korff et al., 1984). Different levels of definition were employed; the simplest, six-month prevalence of panic, provided figures consistent with Margraf and Ehlers's (1988) nonclinical panic findings (28.5–31.1%; von Korff & Eaton, 1989). Panic Disorder prevalence rates were higher than those previously (6.3–10.0% in the ECA study compared to Weissman's figures of 3.3–6.8%) (Weissman, Leaf, Blazer, Boyd, & Florio, 1986). It is possible that the discrepancy may arise from the use of lay interviewers in the ECA study. Nevertheless, epidemiological studies are consistent in placing the prevalence of Panic Disorder in the general population at around 5%.

6.18.3.3 Correlates of Panic Disorder

As noted above, the factors found to be associated with Panic Disorder also tend to be found in nonclinical panickers. Panic Disorder is more common among women (75% of sufferers). Panic attacks are most likely to develop in the age range 18–24 years, while Panic Disorder is most likely to peak in the age range 24–44 years (von Korff & Eaton, 1989); this finding might indicate that panic attacks represent a precursor to the full disorder. Onset of Panic Disorder is associated with negative life events such as marital conflict and bereavement (Pollard, Pollard, & Corn, 1989).

Boyd (1986) analyzed ECA data across a range of diagnoses, and found that people with Panic Disorder were more likely to seek help for their problem than those with any other psychiatric diagnoses (including schizophrenia). The high prevalence of panic, both in other anxiety disorders and in affective disorders as noted by Barlow (1988), Barlow and Craske, (1988), and Barlow et al. (1985), has been proposed as a factor influencing the decision of these other groups to seek professional help (Barlow, 1988). Boyd (1986) suggests that the strong element of somatic sensations and fears of associated catastrophes might account for consultation. Work by Katon (Katon, 1984; Katon, Vitaliano, Russo, Jones, & Anderson, 1987) provides some support for this view. He noted that early presentation of panic to primary care physicians focused almost exclusively on somatic symptoms (Katon et al., 1987), and that physical causes were initially attributed in 89% of a sample of panic cases; misdiagnoses persisted "for months or even years" (although exact figures are not given).

The association between somatic diagnoses and panic is well known, particularly in the cardiac clinic (Skerrit, 1983; Waxler, Kimbiris, & Dreifus, 1971). The involvement of chest physicians in this area has led to a search for pathophysiological correlates of chest pain and other cardiac symptoms; most commonly, this has focused on the role of hyperventilation, with a diagnosis of "hyperventilation syndrome" frequently being advocated (e.g. Lum, 1976; Magarian, 1982). This is discussed in greater detail below.

6.18.3.4 Comorbidity

The coexistence of panic with a wide range of other psychiatric diagnoses has been noted, particularly by Barlow (Barlow & Craske, 1988; Barlow et al., 1985). Panic attacks occur in between 83% and 100% of patients with anxiety disorders and in those with depressive disorders; the lowest figure tending to be for Generalized Anxiety Disorder (Barlow & Craske, 1988). Strict *DSM-III* diagnoses of Panic Disorder range from rates of 30–50% in major depression (Leckman, Weissman, Merikangas, Pauls, & Prusoff, 1983; Vermilyea, 1987, cited in Barlow, 1988) to 83% in anxiety disorders (Barlow et al., 1985). There were few differences between the characteristics of panic in Panic Disorder groups and the other diagnoses. The report of dizziness and fear of going crazy or losing control were significantly higher in the Panic Disorder patients, a finding which may be of specific relevance to psychological hypotheses which emphasize bodily sensations and their misinterpretation signs of catastrophe. Boyd (1986) has reported high rates of panic in patients diagnosed as part of the ECA study; especially striking in this study are the panic comorbidity figures of 28–62% in schizophrenia.

From a different perspective, panic has usually been described in association with a number of clinical phenomena, particularly avoidance behavior and agoraphobia. It has been suggested elsewhere (Salkovskis, 1988a) that the definition of avoidance behavior associated with panic attacks needs to be broadened to include more subtle aspects of avoidance behavior, particularly the immediate reactions to panic such as holding on to supermarket trolleys, using distraction and so on.

A further likely association is with health anxiety or hypochondriasis, dealt with in greater detail below. The structure of the interview for *DSM-III* is such that hypochondriasis is not classified as an Anxiety Disorder, but instead is part of the Somatoform Disorders Group. For this reason, Barlow et al. (1985) did not include screening for hypochondriasis in their study of the prevalence of panic in other disorders, and this has not yet been evaluated in any other study. From a somewhat different perspective, Noyes, Reich, Clancy, and O'Gorman (1986) used a measure of Hypochondriasis (Pilowsky's Illness Behaviour Questionnaire (IBQ)) and found that scores of 60 patients with Panic Disorder and Agoraphobia were similar to norms obtained from groups of Hypochondriacal patients. In the same study, good treatment outcome in respect of panic was associated with a significant improvement in IBQ scores. One of the main reservations of this study concerns the use of the IBQ, which is a poor measure of hypochondriasis.

6.18.3.5 Natural History of Panic

The absence of high-quality longitudinal studies means that little is known definitively about the natural history of panic and Panic Disorder. In general, it appears that, in some patients, an anxiety diathesis is present in childhood and that chronicity is common (Degonda & Angst, 1993; Pollack et al., 1990; Pollack & Otto, 1997). Most information comes from cross-sectional studies, and is subject to a range of problems such as cohort effects. As outlined above, there is some evidence that panic attacks precede the development of Panic Disorder. Von Korff and Eaton (1989) make an interesting theoretical case for Panic Disorder as a sequenced and staged problem. The sequence proposed is based on that suggested by Klein et al. (1987), and runs as follows:

(i) occurrence of panic attacks;
(ii) cognitive appraisal of symptoms as harmful;
(iii) sensitization to symptoms; and
(iv) development of avoidance.

These stages are said to represent different phases in the development of a disease. Agoraphobia is thus said to represent a later stage of Panic Disorder. The evidence supporting this epidemiologically-based view is scanty and, in some instances may be contradictory. For example, Thyer, Nesse, Cameron, and Curtis (1985) demonstrated that, in their sample, Panic Disorder patients have a (nonsignificantly) *longer* duration of disorder compared to agoraphobics. The disease-stage model would predict a significantly shorter duration. The link between Panic Disorder and Agoraphobia is discussed in detail in Salkovskis (1988a); it is concluded that the data is consistent with two possibilities; either that important idiosyncratic individual differences in the reaction to anxiety and panic may be involved, or that there may be a difference in severity, with agoraphobia with panic representing the more severe variant (Pollack & Smoller, 1995).

A further shortcoming of the disease-stage model is the implication of a passive relationship between panic and avoidance such that avoidance is *caused by* panic. There is a considerable amount of data supporting Marks (1987a, 1987b) assertion that a functional relationship exists between agoraphobia and panic and vice versa; not least is the demonstration that exposure treatments not only reduce avoidance but also ameliorate panic attacks (Marks et al., 1983; see also review in Michelson, 1988). Again, all concerned conclude that the requirements for longitudinal studies are paramount (Klein et al., 1987; Marks, 1987a).

6.18.3.6 Other Issues

There are a range of other issues relevant to the phenomenology and assessment of panic attacks and Panic Disorder; these are reviewed in Barlow (1988), Rachman (1990), and Salkovskis (1988a). The main questions posed can be summarized as follows:

(i) Is panic qualitatively distinct from other forms of anxiety?

(ii) Is there a dimensional basis for panic, with the presence of Panic Disorder reflecting an extreme score on an underlying psychological dimension; if so, what is the nature such a continuum?

(iii) Is there a difference between cued and uncued ("spontaneous") panic attacks?

These issues have not been fully resolved (Salkovskis, 1988a); however, it now seems most likely that the answers to such questions will depend less on epidemiological and phenomenological studies and more on experimental investigations based on the evaluation of particular theories of the nature of panic.

6.18.4 THEORIES OF PANIC

There are two main approaches to the conceptualization of panic attacks. These are:

(i) Biomedical hypotheses, which consider that Panic Disorder is an illness or disease with a specific pathophysiology (caused by a lesion, a biochemical deficiency, or some other abnormality of structure or function). These hypotheses include Klein's original position, his suffocation-alarm hypothesis, Charney's neuropharmacological hypotheses, Leibowitz et al.'s detailed neuroanatomical hypothesis, and the physiological/hyperventilation syndrome theories of Lum and Ley.

(ii) Psychological hypotheses, which propose that panic is the outcome of normal processes which may be quantitatively but not qualitatively different from everyday experience. These include van den Hout's interoceptive conditioning hypothesis and the cognitive hypotheses advocated by Beck, Clark, Ehlers, and Salkovskis. Barlow's "false alarm" hypothesis is described here as psychological because many of the processes invoked are psychological in nature. However, it also shares a key characteristic with the biomedical hypothesis in that it starts from the supposition of dysfunction in a biological "alarm" mechanism which determines the uniqueness of panic.

As Seligman (1988) and Rachman (1990) comment, the debate between the two positions concerning panic has considerable implications in other areas, being the "centerpiece" of the debate between two competing models of psychopathology.

Despite the major redefinition which was involved in *DSM-III*, earlier behavioral conceptualizations of phobic anxiety and Agoraphobia remain of vital importance (cf. Marks, 1987a; Salkovskis & Hackmann, 1997) and have much to say about the implications of phobic avoidance behavior (particularly agoraphobic avoidance and its treatment). However, at this point behavioral theories have not directly contributed to the debate on the nature of panic, and for a variety of reasons tend to become aligned with cognitive views in the form of cognitive-behavioral approaches (Hawton, Salkovskis, Kirk, & Clark, 1989; Salkovskis, 1986). For detailed accounts of older behavioral approaches to the understanding of agoraphobic behavior, the interested reader is referred to Foa et al. (1984), Hallam, (1978, 1985), Marks (1981, 1987b), Mathews, Gelder, and Johnston (1981), and Rachman (1990).

6.18.4.1 The Original Biological Hypothesis: Klein's View

The importance of Klein's hypothesis goes beyond its intrinsic value, both because it resulted in the adoption of pharmacotherapy as the dominant North American treatment approach to panic and because other biological theories which developed later have tended to take his views and evidence as their own starting points. For this reason, the foundations of this theory will be considered particularly carefully. Klein (1980, 1981; Klein & Klein, 1988, 1989a, 1989b) argues that several findings indicate a qualitative distinction between panic anxiety and anticipatory or phobic anxiety. These are: (i) the apparently spontaneous nature of some panic attacks; (ii) the observation that panic but not anticipatory anxiety responds to antidepressant medication (Sheehan, 1982), and that panic does not respond to phenothiazines or (as shown later) benzodiazepines; (iii) the finding that patients have experienced more separation anxiety and school "phobia" (50%) than mixed psychiatric controls who were not diagnosed as suffering from Panic Disorder (27%) (Gittleman-Klein & Klein, 1984; Klein, 1964); (iv) the finding that sodium lactate frequently precipitates panic attacks in Panic Disorder patients but does not do so as frequently in nonclinical control subjects or socially anxious patients (Liebowitz et al., 1984; Liebowitz, Fyer, et al., 1985; Pitts & McClure, 1967); (v) evidence of a familial (possibly genetic) component in panic; particularly the higher concordance rate for panic in monozygotic as compared to dizygotic twins (Crowe, Pauls, Slymen, & Noyes, 1980; Torgerson, 1983).

Klein regards panic as similar to epilepsy, with uncontrolled discharges of activity in the biobehavioral anxiety system producing semi-random bursts of activity which manifest as panic attacks. This is said to happen because a normally stabilizing negative feedback loop has become converted to a pathological positive feedback loop, possibly as a result of sensitizing separation experiences during childhood. All other panic-related phenomena follow on from these bursts of pathological activity, including Agoraphobia (Klein et al., 1987).

Klein's position has been criticised as being too vague, especially in terms of the failure to specify the exact nature of the disordered mechanism (Rachman, 1990). There have been a variety of critiques of the evidential basis of Klein's theory; these are fully discussed in Clark (1986a, 1986b, 1988), Margraf, Ehlers, and Roth (1986), Rachman (1990), and Barlow (1988). In the order specified above, these can be summarized as follows.

(i) The degree of spontaneity of panic, even early attacks, is debatable; in the longer term, work such as that reported by Faravelli (1985), Finlay-Jones and Brown (1981), Pollard, Pollard, and Corn (1989) and reviewed in Roy-Byrne and Uhde (1988) indicates that threat-related life events tend to precede the onset of panic. In more immediate terms, there is evidence that acute panic is often preceded by identifiable stimuli such as bodily sensations and frightening thoughts (Beck, Laude, & Bonhert, 1974; Hibbert, 1984). These interview studies have received support both from studies 24 hour ambulatory psychophysiological monitoring (reviewed in Freedman, 1989) and from work using telemetry and computer based *in situ* questionnaires (Taylor, Fried, & Kenardy, 1990).

(ii) The initial supposition (Klein & Fink, 1962) that phenothiazines were specifically anxiolytic in generalized anxiety has not been supported by subsequent research. In later work, the emphasis of the "pharmacological dissection" shifted to the demonstrably anxiolytic benzodiazepine class of drugs. Early studies suggested that the benzodiazepine class of drugs were not effective in the treatment of panic attacks (McNair & Kahn, 1981), whilst a range of outcome studies continued to provide evidence for the specific efficacy of tricyclic antidepressants (Garakani, Zitrin, & Klein, 1984; McNair & Kahn, 1981; Zitrin, Klein, Woerner, & Ross, 1983). However, there is now a considerable amount of evidence available for short-term (i.e., over a period of six weeks) beneficial effects of benzodiazepine administration in panic, particularly the triazolobenzodiazepine alprazolam (Ballenger et al., 1988), but also to higher doses of more conventional benzodiazepines such as diazepam (Noyes et al., 1984). However, the efficacy of imipramine as an effective treatment for Panic Disorder is now well established (Ballenger, 1986; Sheehan, Ballenger, & Jacobsen, 1980). Despite Marks' assertion that the antipanic effects of imipramine are mediated by its specific antidepressant action (based on the results of the study carried out by Marks et al., 1983; see also Marks, 1987a), this conclusion is not supported by a large number of other well-conducted studies (Clum & Pendrey, 1987; Klein & Klein, 1989a, pp. 151–153). The more convincing reservation (Telch, Tearnan, & Taylor, 1983) focuses on whether or not the effectiveness of imipramine is *independent* of the self-exposure instructions which invariably accompany pharmacological treatment. The effectiveness of such instructions on their own and combined with pharmacological treatments has been the subject of much investigation. At present, there is some evidence that the effects of imipramine and self-exposure instructions might be synergistic (Margraf & Ehlers, 1989; Telch et al., 1983). Finally, a number of authors have pointed out that the effectiveness of a particular treatment modality (i.e., biological or psychological) is not convincing evidence favoring the basis of a disorder in that modality (Clark, 1988; Rachman, 1990).

(iii) A number of carefully controlled studies (including prospective ones) have failed to find differences between panic patients and others in terms of separation anxiety and school phobia (Berg, Butler, & Pritchard, 1974; Coyrell, Noyes, & Clancy, 1983; see also reviews by Margraf, Ehlers, & Roth 1986a, 1986b, and Tearnan, Telch, & Keefe, 1984). Earlier assertions in this respect were largely based on retrospective reporting.

(iv) 40% or more of panic patients do not panic when sodium lactate is administered. This observation is not consistent with the hypothesis that response to lactate is a specific "biological marker" for the mechanisms involved in naturally occurring panic attacks. Data from the Oxford group indicates that, far from being impervious to psychological influences as previously argued (Klein, 1981), the anxiety-provoking effects of lactate in panic patients can be blocked by instruction (Clark et al., in preparation). A further problem for Klein's original position is the fact that sodium lactate infusions are only one of a range of "biological challenges" which have the effect of provoking panic attacks in Panic Disorder patients; the range of such panic-provoking challenges is remarkably wide, including adrenaline, caffeine, increased carbon dioxide, hyperventilation, isoproterenol, *m*-CPP, and yohimbine (see reviews by van den Hout & Griez, 1986; Margraf, Ehlers, & Roth, 1986a, 1986b); Uhde & Tancer, 1989). The importance of the finding that all of these diverse provocations produce a panic response in panic patients is the fact that these challenges have not been identified as having any commonality in terms of either their biochemical or physiological effect. Thus, it is very unlikely that they are all working on the same mechanism, and no view has yet been proposed concerning possible multiple inter-related influences involved in the biochemical triggering of panic. It is therefore not clear whether some biochemical inductions are valid models of naturally occurring panic and others are not, whether none are, and so on. As will be described below, a psychological account best encompasses the full range of inductions used, and does so in a particularly parsimonious way.

(v) The evidence for a familial component is relatively stronger than other evidence for the biological hypothesis (Kendler, Heath, Martin, & Eaves, 1987). However, there are some distinct oddities in the concordance data which are difficult to make sense of from the standpoint of a genetic diathesis for the occurrence of panic attacks. For example, groups of relatives who should show similar genetic risk have been shown to have widely differing rates of panic and other anxiety disorders. Dizygotic twins were shown by Torgersen (1983) to have lower panic concordance rates than nontwin siblings, and concordance between siblings is less than between parents and children (Moran & Andrews, 1985). The extent to which familial risk for Panic Disorder differs from that for Generalized Anxiety Disorder is a subject of debate (Margraf et al., 1986a, 1986b). Margraf and others have suggested that concordance rates are only high for panic when liberal diagnostic criteria are used (i.e., when the presence of panic is recorded because of what are described as limited symptom attacks [increases in anxiety involving less than the four symptoms required by *DSM-III*]) as opposed to the full diagnostic criteria for Panic Disorder. That is, it is not possible to rule out the suggestion that relatively mild symptoms of general anxiety are being inappropriately identified as panic attacks. This view is consistent with the findings that those studies finding the highest concordance for panic also find the lowest concordance for Generalized Anxiety Disorder. Barlow (1988, pp. 172–174) summarizes this data, and proposes that the evidence on the heritability of anxiety is most consistent with the hypothesis of the *trait* of anxiety (often described as emotionality, nervousness, or neuroticism). He argues that studies reporting a difference between heritability for Panic Disorder and other anxiety disorders have not controlled for the fact that all Panic Disorder subjects are generally anxious with different levels of severity which are inherent in the structure of the *DSM-III* diagnostic rules. When severity of disorder is taken into account, Barlow points out that the different anxiety disorders appear to be equally heritable.

In later theoretical writings, Klein and Klein (1989a, 1989b) suggest that the specific mechanism of panic may involve an evolutionarily evolved central alarm mechanism which functions in relation to separation anxiety or impending asphyxia, and that catastrophic cognitions are incidental to this biological mechanism. He argues that "Many spontaneous panics occur when the brain's suffocation monitor erroneously signals a lack of useful air, thereby maladaptively triggering an evolved suffocation alarm system" (Klein, 1993; see also Klein, 1996). Again the suggestion is that panic is relatively impervious to psychological influences. The evidence against the "irrelevance" of psychological influences is now overwhelming. There have been a number of more direct tests of the suffocation/alarm system which suggest that it is not correct (e.g., Taylor, Woody, Koch, McLean, et al., 1996).

6.18.4.2 The Neuroanatomical Hypothesis

A group of researchers from Klein's center proposed a "neuroanatomical hypothesis," which they describe as being inspired by his work (Gorman, Liebowitz, Fyer, & Stein, 1989). This complex hypothesis takes Klein's basic assertions (p. 26) as its starting point, but attempts both to make the position more explicit and to identify the neuroanatomical basis of the components of the hypothesis. The paper also attempts to incorporate cognitive-behavioral findings and circumvent the psychological/biological debate described above (p. 25). They argue that it is not useful to attempt to specify whether panic is "a behavioural cognitive disease" (sic) or "a biological disease," and that it should be regarded as having elements of both. By describing the basis of their hypothesis in this way, they seem to misconstrue the basis of the debate, which is not biological vs. psychological, but disease vs. nondisease (Rachman, 1990). The neuroanatomical hypothesis proposes that:

> there are three distinct components of the illness Panic Disorder—the acute panic attack, anticipatory anxiety and phobic avoidance. These three clinical phenomena arise from excitation of three distinct neuroanatomical locations, respectively: the brainstem, limbic lobe, and prefrontal cortex. Reciprocal innervation among nuclei in these three centres explains the genesis of the disease and its clinical fluctuations over time. (Gorman et al., 1989, p. 149)

Like Klein's view, the model proposed is a disease-stage one, with acute panic being the necessary first stage, followed by the development of anticipatory anxiety and agoraphobic avoidance. The problem of the development of agoraphobia without preceding panic attacks is acknowledged (Gorman et al., p. 149) but not dealt with except by suggesting that "phobic avoidance and anticipatory anxiety could conceivably be maintained without the induction of panic if only these reciprocal pathways between prefrontal cortex and limbic areas were involved" (Gorman et al., p. 156). Gorman *et al.* specify that the complexity of the hypothesis means that a great many experiments will be required to validate or disprove the model; they also argue in their conclusion that, because the disorder is complex and involves so many facets, it requires a complex theory to account for it (p. 158). Perhaps the principal problem with this position is the fact that they do *not* specify any differential predictions vis-à-vis alternative hypotheses, relying instead on a proposed synthesis which incorporates alternative accounts ("combined rather than opposed," p. 158). However, they do insist on the basic assumption of "the neural circuits of the brain as the basis for pathology."

The argument for a brainstem basis of acute panic is based on the physical character of panic symptoms and the biochemical panic provocation studies discussed above in the context of Klein's theory. The suggestion is that the various panic-provoking agents produce paroxysmal brainstem activity as a result of activity in one or more specific brainstem loci. However, the central basis of at least one panic-provoking agent, isoproterenol (which provokes panic at the same rate as lactate infusions; Rainey et al., 1984), is particularly hard to explain in these terms as it does not cross the blood–brain barrier (Conway, Tejani-Butt, & Brunswick, 1987) and therefore cannot have its panic-provoking effects as a result of direct action on the brainstem. Gorman et al. discuss yohimbine as an exemplar of a panic-provoking agent because its action is relatively well understood; it has the effect of increasing CNS noradrenergic activity, particularly in the locus coeruleus and pontine nucleus. The effect of yohimbine in activating the brainstem noradrenergic activity which is hypothesized to be implicated in panic is described in some detail (Gorman et al., p. 156). The role of clonidine, which has the opposite effect to yohimbine on noradrenergic activity, is more briefly alluded to; two studies of the effects of clonidine in panic patients are cited as additional evidence of the proposed noradrenergic hypersensitivity in panic patients (Charney & Heninger, 1986; Nutt, 1986). However, as is described below (p. 39), although the results of these clonidine studies appear to show an abnormality, it is *not* consistent with a simple hypersensitivity as proposed by the neuroanatomical model.

The evidence for a link between the limbic system ("limbic lobe") and anticipatory anxiety is somewhat more consistent, being mainly derived from the work of animal researchers such as Gray (1982). The specific involvement of this system in Panic Disorder in man is more difficult to justify. Little evidence supporting such a link is presented, other than general allusions to the vulnerability of the limbic system to the generalization of spasmodic CNS electrical activity ("kindling effects") coupled with speculation that general anxiety associated with panic might result from such kindling effects. The importance of the limbic system to the neuroanatomical model is related in large part to the proposition that it is through this system that nonpharmacological interventions (particularly exposure) have their effects. It is difficult to see how the propositions made concerning the role of limbic system can be evaluated directly, and whether this component of the hypothesis in any way advances the understanding of mechanisms involved in panic attacks and their treatment.

The prefrontal cortex is linked to panic and anticipatory anxiety by its presumed cognitive function; it is here that interpretations are linked to brainstem discharge, according to the model. The extensive innervation (afferent and efferent) of this area is said to make it particularly suitable for such a function. Again, no specific evidence in support of a link with panic is offered other than previous general studies into the possible function of the prefrontal cortex. If this aspect of the hypothesis is correct, then it may be possible to evaluate this directly by experiments involving patients who have previously been subject to psychosurgery in this region.

Perhaps the most valuable aspect of Gorman et al.'s paper is that it offers specific predictions which are said to allow experimental testing of the hypothesis. These predictions generally specify the likely locus of brain activity which should accompany various forms of stimulation and treatment. The prediction which is particularly relevant to psychological hypotheses concerns treatment, specifically that cognitive-behavioral treatments will leave a core of occasional panic attacks (whilst drug treatments will not) because cognitive-behavior therapies "exert ameliorative effects at a different 'site' of disease progression than do antipanic drugs" (Gorman et al., p. 157). This proposition already seems to have been invalidated (Clark et al., 1994; Clark, Salkovskis, & Chalkley, 1985; see also below).

Altogether, the neuroanatomical hypothesis is disappointing in that it does not utilize new data, and seems to explain less than Klein's original hypothesis. The attempted integration is unhelpful because it retains the basic proposition of a disease process, and because psychological factors are only included in the account as secondary phenomena to the unmodified "fact" of biologically determined panic. That is, all psychological factors are considered to operate "downstream" of fundamental biological mechanisms modifying the secondary response to panic itself, with the most minimal of interactions even at this secondary level. This position does not take account of the wide range of psychological findings described below. The problems associated with the empirical foundations of Klein's position (described above) remain unresolved by this newer theorizing.

6.18.4.3 Neurochemical Hypotheses: Noradrenaline and the Locus Coeruleus

Some of the most scientifically interesting work on biological approaches to panic has emanated from Charney's group (Charney & Heninger, 1986), who have concentrated on a neurochemical hypothesis in which it is proposed that panic attacks are a result of dysregulation in the noradrenergic system. This work formed an elaboration of earlier basic research in primates implicating the locus coeruleus (LC) in fear situations and behaviors. Redmond and colleagues found that direct stimulation of the LC in stump-tailed monkeys results in a behavioral response said to mimick a human panic attack; the monkeys displayed behaviors characteristic of fear in the wild (e.g., scratching and grimacing) (Redmond, 1985; Redmond & Huang, 1979). Ablation of the LC in monkeys was associated with less fear responses when in the presence of anxiety-provoking stimuli, as was the administration of many antianxiety drugs (Redmond, 1979; 1985; Redmond & Huang, 1979). Charney and Heninger (1986) argue that the importance of the noradrenergic LC system lies in its responsivity to sensory input and the ramifications of this system throughout the cortex. They propose that "abnormally high responsivity of brain noradrenergic systems is etiologically related to the development of panic disorders." In particular they focus on the possible importance of alpha2-adrenergic presynaptic receptors (autoreceptors), which regulate the release of noradrenaline within their own neurones through a negative feedback mechanism. Charney proposed that impairment of this inhibitory mechanism leads to inappropriate paroxysmal discharges in the arousal system in response to low levels of stimulus input. Evidence for this position was derived from the fact that some panic-inducing substances, such as yohimbine and caffeine, are known to stimulate the LC, while drugs which have an antipanic effect are known to depress LC activity (e.g., tricyclic antidepressants; but also diazepam).

There are a number of problems with the LC hypothesis. There are doubts about the generalizability of much of the animal research upon which it is based (Mason, 1981), and a variety of findings from well-conducted studies contradict the earlier positive results. For example, biochemically specific noradrenergic LC lesions in rats do not impair the response to threatening stimuli (File, Deakin, Longden, & Crowe, 1979) and do not impair the acquisition of one-way, two-way, and Sidman avoidance responses, nor modify conditioned suppression. In humans, the effect of LC and central noradrenergic disturbances are impaired perceptual organization and memory, disorientation, hallucinations and delusions, with anxiety not being a prominent feature (van Dongen, 1981). Drugs such as buspirone and carbamazepine which increase LC activity and mianserin which blocks the activity of alpha2-autoreceptors do not induce panic, and have been shown to be therapeutic in some instances. The evidence for Charney's LC hypersensitivity hypothesis is not particularly convincing. However, this line of research has generated a series of findings in panic patients which are not readily explained by any other position, and which suggest some biological differences between panic patients and nonclinical subjects. Charney and Heninger (1986), Nutt (1986), and Uhde, Siever, and Post (1984) evaluated the response to the alpha2-receptor agonist clonidine in panic patients and nonclinical controls. They found that panic patients show a modified hormonal response (significantly greater decreases in the noradrenergic metabolite MHPG (3-methoxy-4-hydroxyphenylglycol) and smaller increase in growth hormone) and greater decreases in blood pressure (but not heart rate). While these differences cannot satisfactorily be explained in any detail by the LC hypothesis as originally stated, they might indicate what Charney describes as a "dysregulation," involving both over- and under-responsiveness of different subsystems, perhaps even the same subsystems at different times. This line of research may be of interest if it can be shown that the key results can be replicated when panic patients are compared to anxious controls.

6.18.4.4 Physiological Accounts of Panic: Panic as a "Hyperventilation Syndrome"

The notion that panic attacks may occur as a result of episodic hyperventilation is an old one which has recently been enthusiastically revived

(Hibbert, 1984; Ley, 1985a, 1985b). It has periodically been suggested since the 1930s that the physiological effects of breathing in excess of metabolic requirements could account for sudden severe attacks of anxiety accompanied by a wide range of bodily sensations (e.g., Gibson, 1978; Kerr, Dalton, & Gliebe, 1937; see also review by Magarian, 1982). Briefly, the hyperventilation hypothesis is that, under conditions of stress, patients breathe in an excess of metabolic requirements, that is, they breathe out more CO_2 than they produce. When this happens, it results in a "blowing off" of CO_2. That is, because of the more rapid movement of air in and out of the patient's lungs, more CO_2 than usual diffuses out of the pulmonary blood vessels and into expired air, resulting in a drop in the partial pressure of arterial CO_2 (p_aCO_2). Because CO_2 is dissolved in blood as carbonic acid (as part of the pH buffering system), the effect of decreased p_aCO_2 is an increased alkalinity of blood, usually described as a "respiratory alkalosis." In physically healthy subjects, respiratory alkalosis produces a range of physical sensations, potentially including most of those regarded as symptoms of panic. Thus, it has been argued that panic attacks are caused by or are even synonymous with episodes of acute hyperventilation.

Normally, respiratory alkalosis is self-correcting, in that respiration self-regulates; lowered p_aCO_2 results in the activation of a range of negative feedback mechanisms, such as increased tone in the chest wall (intercostal) muscles. Note that increased breathing rate or depth are not necessarily indicative of hyperventilation; only blood CO_2 levels (as measured directly from arterial blood [p_aCO_2] or the mixed venous [$p_{mv}CO_2$] systems) give true indications of whether hyperventilation is occurring or not, and the degree to which it is present. It also appears that it is not the absolute level of p_aCO_2 which is important, but is instead the extent to which this variable shows a rapid decrease. This has been shown in an elegant way by Hout van den and Griez (1985), who carried out an experiment on breath-holding after CO_2 inhalations. They demonstrated that the symptoms induced by inhalations of 35% CO_2 in 65% O_2 did not arise from the increased carbon dioxide associated with the inhalation, but by the subsequent rapid *fall* which occurs when an inhalation is released and followed by a reflex hyperventilation and unusually large decrease in p_aCO_2. (See Hardonk & Beumer, 1979, and Lum, 1976, for more detailed accounts of physiological factors involved in hyperventilation; Hibbert (1984) provides a concise account of the simplest form of the hyperventilation hypothesis of panic attacks.) The role of change in p_aCO_2 levels in the production of symptoms has also led to a specific proposal concerning the role of hyperventilation in the vulnerability to acute episodes of panic symptoms. It has been proposed that panic patients are vulnerable to such acute episodes of panic arising from hyperventilation because of a chronic, asymptomatic hyperventilation (Lum, 1976). Thus, it is argued that frequent hyperventilation leads to the blood buffering system "resetting," so that a constantly lowered p_aCO_2 is present at all times, but only becomes symptomatic when a further acute drop occurs, which will occur relatively frequently because the reset buffering system is more reactive to small p_aCO_2 changes.

Hyperventilation therefore provides an apparently simple and clinically appealing explanation of an otherwise complex and perplexing phenomenon, and this type of hypothesis has enjoyed intermittent popularity and "rediscovery" every few years. The recently renewed attention to the proposed link between panic anxiety and respiration probably arises from the emphasis currently placed on the diagnostic importance of panic attacks and the search for biological explanations and treatment already described above. This is because explanations based on the idea of "hyperventilation syndrome" invariably have an implicit basis in biomedical ideas of underlying dysfunction or pathology of respiratory control.

6.18.4.4.1 Evidence for an association between panic and hyperventilation

The evidence for an association between panic and hyperventilation is considerable. However, this falls far short of evidence for causation. First, it has frequently been observed that voluntary hyperventilation reproduces symptoms similar to those experienced during panic attacks, including increased anxiety (e.g., Lum, 1976). Second, substantial drops in p_aCO_2 levels have been observed as a result of contrived and naturally occurring stressors in panic patients (Hibbert, 1986; Salkovskis, Clark, & Jones, 1986; Salkovskis, Warwick, Clark, & Wessels, 1986) and in normal subjects. Third, a number of studies have reported that panic patients have significantly lower resting levels of p_aCO_2 than normal controls (Gorman et al., 1987; Gorman et al., 1986; Rapee, 1986; Salkovskis, Jones, & Clark, 1986), possibly indicating the occurrence of renal compensation for repeated acute hyperventilation ("resetting" of the buffer system; Gledhill et al., 1975). That is, if hyperventilation (and therefore the respiratory alkalosis) is prolonged, the metabolic response involves excretion of bicarbonate, which restores the acid/base balance at a lower level of p_aCO_2. Fourth, respiratory symptoms and fears are particularly prominent among panic patients (Pollard, 1986) and symptoms associated with hyperventilation are significantly more likely to occur in panic patients than in patients suffering from general anxiety disorder (Rapee, 1985). Finally, there have been a number of reports of the success of treatments which contain an element of respiratory control (Clark et al., 1985; Lum, 1976; Salkovskis, Jones, & Clark, 1986), and that successful psychological and physical treatment results in a significant decrease in resting p_aCO_2 (Gorman et al., 1987; Salkovskis & Clark, 1986).

6.18.4.4.2 Hyperventilation as a CAUSE or EFFECT of panic

Ley (1985a, 1985b) strongly argues that hyperventilation causes panic. In a formulation very similar to previous "hyperventilation syndrome" hypotheses (Garssen, van Veenendaal, & Bloemink, 1983; Hardonk & Beumer, 1979; Lum, 1976), he proposed that "the panic attack is unique, that it consists of a synergistic interaction between hyperventilation and fear" (Ley, 1985b, p. 272). The basis of the relationship between fear and hyperventilation is said to be found in the way in which fear activates the sympathetic nervous system and thence drives the intercostal muscles of the thorax, thus increasing respiration. Ley regards the occurrence of hyperventilation as the crucial (necessary) factor for panic to occur, and states that "the crucial assumption is that *sensitive* individuals who have recently experienced stressful events are constantly overbreathing and, therefore, on the brink of respiratory alkalosis" (Ley, 1985b, p. 281).

The foundations of this approach do not substantially differ from other attempts to account for panic attacks as part of a centrally determined defect of respiratory control. As with previous statements of the hyperventilation hypothesis, the way the role of fear is dealt with presents some problems, although it is at least considered in some detail. Fear is said to arise either as a direct physiological effect on the autonomic nervous system, or the intensity of the symptoms in some direct (and unspecified) way produces fear and alarm. An interaction between fear and overall levels of stress and sensitivity to fear is proposed, as well as individual differences in the respiratory system and its reactivity.

Most problematic for the view that hyperventilation is a necessary cause of panic attacks are the results of work on ambulatory monitoring of CO_2. This technique measures the concentration of CO_2 passing across the skin (transcutaneous CO_2) and record the information on a miniaturized belt-mounted recorder while the patient goes about their normal business. This measure of transcutaneous CO_2 has been shown to correlate highly with arterial CO_2 (Pilsbury & Hibbert, 1987). Studies using this measure have shown that some patients do not hyperventilate during panic attacks (Hibbert, 1986; Hibbert & Pilsbury, 1989). The effectiveness of treatment involving elements of respiratory control has been taken as evidence for the hyperventilation hypothesis. Early cognitive treatments emphasized respiratory control as part of a broader program of treatment directed at reattribution. Hibbert and Chan (1989) attempted to replicate the Clark et al. (1985) and Salkovskis, Jones, and Clark (1986) findings for respiratory-based treatment but without the inclusion of a cognitive element; the results of this relatively large study suggest that the effectiveness of previous treatments may have been due to the cognitive element, and is certainly not likely to be accounted for by the physiological basis of respiratory control alone.

Nevertheless, while it is difficult to support the view that hyperventilation causes panic, it is also difficult to maintain the view that acute hyperventilation, when it occurs, is irrelevant to the experience of panic. The differences reported between panic patients and control subjects in resting CO_2 seem certain to result in reduced blood buffering capacity and should have the effect of making panic patients more prone to experiencing substantial respiratory changes in response to relatively minor challenges. In a consecutive series of single-case experiments with panic patients, very substantial drops in p_aCO_2 occurred as a response to mild specific psychological stressors (Salkovskis, Clark, & Jones, 1986, discussed together with other related data in Salkovskis & Clark, 1986). It was concluded in this paper that the rapidity and magnitude of the observed changes in CO_2 were sufficiently large to infer that a symptomatic respiratory alkalosis must at least have contributed to the symptoms experienced. That similar profound changes can be observed in a proportion of naturally occurring panic attacks has been demonstrated by a single case in which it was possible to directly measure arterial p_aCO_2 and pH (Salkovskis et al., 1986) and by a study using ambulatory measures of transcutaneous CO_2 (Hibbert & Pilsbury, 1989). There is evidence from a study on the clinical presentation of anxiety disorders that the main symptoms which trigger thoughts of physical and mental harm in panic patients (but not in generally anxious patients) are the symptoms commonly produced by hyperventilation (Rapee, 1985).

It thus appears that it is not possible to account fully for panic by invoking concepts such as "hyperventilation syndrome," but neither does it seem possible to totally discount a role for hyperventilation in the production of symptoms experienced during panic attacks. The way in which the hyperventilation hypothesis enjoys periodic popularity has been based on the fact that there is indeed compelling evidence of an association with anxiety, and on enthusiastic clinical reports of the efficacy of treatments involving breathing exercises (Lum, 1975, 1976). Any hypothesis must therefore reconcile the strong evidence for an association between hyperventilation and panic with the equally strong evidence that hyperventilation is neither necessary nor sufficient for the occurrence of panic. After reviewing psychological hypotheses of panic, I will return to consider a cognitive explanation of the relationship between panic and hyperventilation.

6.18.5 PSYCHOLOGICAL THEORIES OF PANIC

Although psychological hypotheses of panic are a relatively new development, behavioral theories of agoraphobia have been generating research and treatment studies directly relevant to panic since the early 1960s (see Hallam, 1985; Marks, 1987b; Mathews et al., 1981). Such work continues to generate valuable results, and has become substantially integrated with the more recent research and theory emphasizing panic attacks. Some of the most substantial and important work has approached panic from the perspective of panic attacks associated with phobic behavior (Barlow, 1988; Rachman, 1990).

There are two main psychological hypotheses of panic, both of which draw on a variety of influences. These are interoceptive conditioning models (Barlow, 1988; Hout van den, 1988) which are most closely modeled on theories of classical conditioning particularly Russian work described by Razran (1961) and cognitive theories (Beck, 1988; Clark, 1986a, 1986b, 1988; Salkovskis, 1988a) which are based on Beck's (1976; Beck, Emery, & Greenberg, 1985) cognitive approach to emotional disorders. The discussion of Clark's cognitive model subsumes similar models, particularly the psychophysiological models described by Margraf and Ehlers; as pointed out in Ehlers, Margraf, and Roth (1988), all such theories (with the exception of Barlow's) share common assumptions and differ only in emphasis. No differential predictions have, as yet, been identified for any of these other psychophysiological models which have been proposed to account for cognitive–physiological interactions (Ehlers & Margraf, 1989); Clark's (1986a, 1986b) hypothesis therefore remains the most parsimonious and influential of these.

6.18.5.1 Barlow's Model: False Alarms and Learned Alarms

Barlow's model (Barlow, 1988, 1991) is particularly interesting because it is presented in the context of a major monograph in which he attempts to develop an overarching theory relevant to all anxiety disorders and to synthesise conditioning, neurobiological and psychosocial factors. He acknowledges that the result lacks specificity at times, although he also expresses the hope that it represents a testable set of propositions. Interestingly, the model bears some resemblance to the neuroanatomical hypothesis (p. 32), in that it is a disease-stage model with an emphasis on concepts of malfunctioning subsystems. Unlike other psychological models of panic, Barlow's specifically proposes a qualitative distinction between panic and other forms of anxiety. The basis for the model concerns a "complex biopsychosocial process," involving the interaction between "an ancient alarm system, crucial for survival, with inappropriate and maladaptive learning and subsequent cognitive and affective complications" (Barlow, 1988, p. 209). Thus, a biological (evolutionarily relevant) alarm system is postulated; this primitive system is designed to react to actual danger with a set of energizing responses, termed a "true alarm"; such responses prepare the organism to react in appropriate ways to threat, and are therefore associated with "action tendencies," which are either escape/flight, freezing/immobility, or aggression/attack. However, under conditions of nonfocal threat and consequent generalized hyperarousal (e.g., during a period following a major threat-related life event), biologically (genetically) vulnerable individuals become prone to spontaneous discharges of the alarm system. Specifically, "false alarms" (panic attacks experienced at the outset of development of Panic Disorder) are said to "spike off" from a stress response in such vulnerable individuals. That is, generalized anxiety increases under stress and forms a kind of platform which puts the alarm system on a hair-trigger and makes the sparking off of an alarm-type response particularly likely. This initial occurrence of spontaneous panic as "false alarms," which are said to be characteristic of the panic attacks experienced by nonclinical panickers may, in turn, become conditioned to interoceptive stimuli if a false alarm was associated with such stimuli. The formation of such conditioned associations are particularly likely if the association occurs in a situation where the ability to carry through "action tendencies" is blocked. This conditioning process then results in the occurrence of "learned alarms." For Panic Disorder to result, there is a further crucial step, however. Anxious expectation over the occurrence of negative events or subsequent "alarms" must develop, and is said to be "a crucial psychological vulnerability," possibly arising as a result of a "pre-existing disposition to focus anxiety specifically on somatic events." This tendency would make such individuals "exquisitely sensitive to false alarms." Avoidance and cognitive symptoms seem to follow on in a relatively passive way, although this is not made more explicit than already described except as a diagram in which these are depicted as a noninteracting end-point of the sequence. The lack of emphasis on cognitive factors in the theory contrasts markedly with Barlow and Cerny's (1988) and Rapee and Barlow's (1989) description of treatment, which includes a major component of cognitive intervention.

Barlow does not himself suggest any specific predictions from this model, nor does he differentiate it in this way from other hypotheses. However, he does state that "there should be enough coherence within the theory to enable investigators to design research protocols that result in confirmation or disconfirmation" (Barlow, 1988, p. 209). Perhaps the most striking aspect of his theory for the present thesis is the way in which cognitive factors are regarded as a passive "downstream" consequence of a response to the experience of physiological sensations previously associated with excessive arousal. Cognition is thus seen as a complication of a biologically initiated sequence.

6.18.5.2 The Cognitive Hypothesis of Panic

A cognitive hypothesis of panic (Beck, 1988; Clark, 1986a, 1986b, 1988, 1996, 1997; Salkovskis, 1988a; Salkovskis & Clark, 1986) has been developed, based on initial work by Clark (1979) and Beck, Emery, and Greenberg (1985). This fundamentally simple hypothesis states that patients with recurring panic attacks have an *enduring tendency* to misinterpret certain bodily sensations as a sign of imminent disaster. The way in which this accounts for acute panic attacks is summarized in Figure 1.

According to the hypothesis, acute panic results from the misinterpretation of bodily or mental sensations as signs of imminent personal disaster. The anxiety engendered by the misinterpretation of sensations produces an increase in sensations, which can in turn feed back to the misinterpretation, rapidly culminating in a full-blown panic attack. Thus, according to the cognitive hypothesis, panic occurs only if sensations are misinterpreted; patients who experience repeated panic attacks are particularly vulnerable to making such misinterpretations by virtue of a pre-existing set of beliefs. A variety of things may trigger off the panic vicious circle in the first instance; it may be triggered by anxiety from general stress, by the anxiety of entering a situation where panic previously occurred, by a sensation from an unrelated source such as a hangover or too much coffee, by a frightening thought, and so on. It is important to note that the whole process of panic generation can occur within a few seconds, as the person responds physiologically and psychologically to what seems, at that time, to be a profound threat to his or her safety.

For example, a patient may interpret palpitations and a racing heart as a sign that he is having a heart attack. This misinterpretation in turn results in intense anxiety and thus increases the sensations associated with autonomic/physiological arousal, including palpitations, tachycardia, dizziness, and precordial pain. This increase in intensity and range of symptoms appeared to confirm the original misinterpretation, further increasing anxiety and symptoms and so on, rapidly culminating in an acute panic attack.

A wide range of symptoms can give rise to misinterpretations, but the cognitive hypothesis proposes that those involved in panic attacks are particularly likely to be sensations which can result from autonomic or central nervous system responses to anxiety itself, given the immediate nature of the feedback loop. A further factor in the readiness to misinterpret bodily sensations should be the extent to which the person has available any innocuous alternative explanations of the intense and wide-ranging sensations associated with panic attacks. For example, if someone were to notice their heart racing, being short of breath, feeling dizzy, and found they were sweating a great deal immediately after running to catch a train, misinterpretation is unlikely. The same sensations experienced without any obvious cause would, at best, be bewildering and at worst be taken as a sign of serious illness.

The catastrophic beliefs specified by the cognitive hypothesis account for the intensity of anxiety and the associated physical symptoms, insofar as the anxiety reaction experienced would not be considered abnormal if occurring

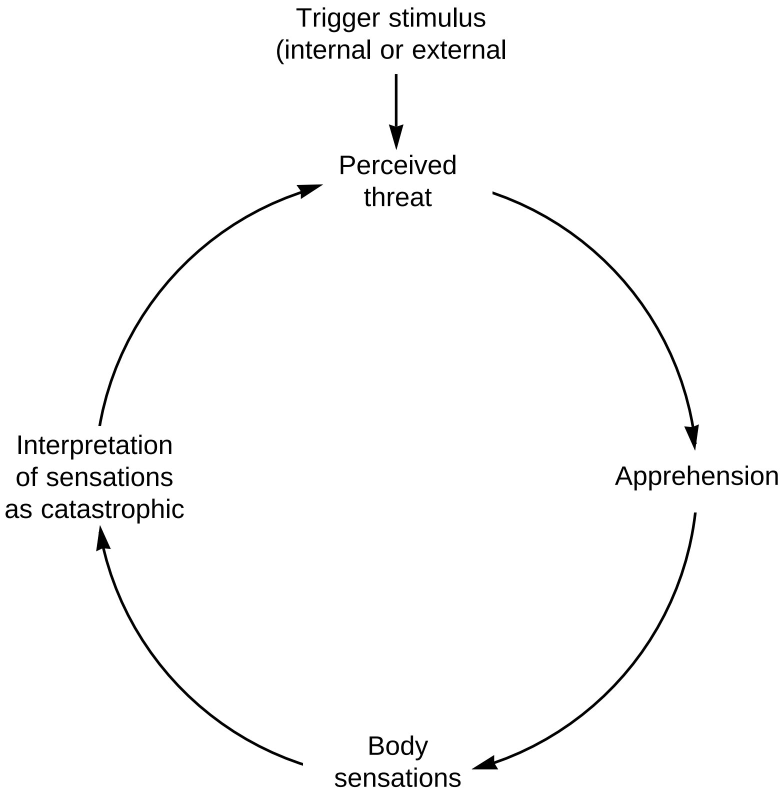

Figure 1 The suggested sequence of events in a panic attack. (Reprinted with permission from Clark, 1986, p. 463).

in the context of a truly life-threatening situation. Typically, patients express very strong beliefs in the possibility of feared disaster during the attacks themselves, although in the clinician's office (when the symptoms are not present and a trusted individual is) patients will report that they do not, at that time, feel convinced that the disasters would happen. The apparent unexpectedness or "spontaneousness" of some panic attacks can also be accounted for; although the panic may start with a mild (and often normal) physical sensation, if such a sensation triggers off the thoughts of disaster and hence a panic, the originating sensation may then be lost in the surge of much more intense, anxiety-generated symptoms which follow. Often, there is no specific explanation for the triggering symptom; it was simply one of the many ordinary physical fluctuations which most people experience every day.

6.18.5.3 The Importance of the Idiosyncratic Nature of Catastrophic Misinterpretations

The misinterpretations made by individual patients are based on aspects of their experience, such as the person's previous experience (direct or vicarious) of symptoms and their meaning. For each patient, there tends to be an internally consistent and logical link between the sensations experienced during panic and the particular interpretations made. Some of the most commonly occurring sensations and related interpretations are shown in Table 2.

The link between particular sensations and interpretations are frequently based on conventional wisdom and "common sense" so that the links are immediately obvious (e.g., shortness of breath interpreted as a sign that the patient is about to stop breathing), although the links may not always be so clear to the observer unaware of the patient's specific pattern of beliefs.

Although it has not as yet been established why some people begin to misinterpret bodily sensations in the first place, there are two main possibilities (which may operate in separate ways or, more likely, in combination). Some individuals may be in a transiently vulnerable state because of biochemical factors (e.g., drugs, hormonal disturbances, etc.; see Roy-Byrne & Uhde, 1988) or recent stressful life events (Finlay-Jones & Brown, 1981, Pollard, Pollard, & Corn, 1989), particularly when those events

Table 2 Examples of the association between symptoms and catastrophic misinterpretations occurring during panic attacks.

Sensation	*Typical misinterpretations*
Heart racing, pounding, palpitations	I'm having a heart attack, my heart will stop, I'm dying
Breathlessness	I'm going to stop breathing, suffocate
Feeling unreal and distant	I'm going to go crazy, lose my mind
Loss of sensation and tingling in arms and legs	I'm having a stroke
Feeling dizzy, faint, weak legs	I'm going to faint, fall over, pass out
Feeling distant, tense, sweaty, and confused	I'm about to lose control of my behavior
Feeling dizzy, heart pounding, chest tight, palpitations, flushed, tingling	I'm dying

concern serious illness in the patient him- or herself or in friends or relatives. Critical incidents of an apparently less serious type may serve the same triggering function (e.g., hearing of a similarly aged colleague at work who had a heart attack). Under such circumstances, patients may more readily misinterpret bodily sensations as harmful when they are unable to generate any convincing innocuous explanation. Second, particular attitudes and assumptions may predispose patients to the misinterpretation of particular circumstances, especially those concerning the effects of anxiety. An individual who held the belief that anxiety can kill may become alarmed as they go through a period of intense anxiety (due, for example, to excessive job stress or illness) at a level which they had not previously experienced. A substantial subgroup of such attitudes have been subsumed under the heading of "Anxiety sensitivity" (Reiss, Peterson, Gursky, & McNally, 1986), which has previously been shown to be high in panic patients (e.g., McNally & Lorenz, 1987) and nonclinical panickers (Donnell & McNally, 1990) and to predict the response of nonclinical subjects to panic-related tasks such as hyperventilation (Holloway & McNally, 1987).

6.18.5.3.1 Panic, cognitions, and avoidance

A variety of researchers have attempted to conceptualize the relationship between panic and avoidance. Rachman has carried out an influential series of experiments in which he examines the relationship between predicted fear and actual fear in subjects repeatedly confronting aversive events (summarized in Rachman & Bichard, 1988). The basic paradigm used involved asking student volunteers, previously identified as claustrophobic, to repeatedly enter a very small enclosed space; in the course of experimental sessions they are asked to rate anxiety and a range of other variables, particularly their expectation of anxiety prior to entering the enclosed space. Using this paradigm, it was found that anxious subjects were consistently more likely to overpredict than underpredict the fear associated with phobic situations, and overpredictions were particularly likely to follow underpredictions. Reductions in fear were most likely after overpredictions, but there was evidence of an inherent bias towards overprediction of aversive events, including both anxiety provocation and pain (Rachman & Bichard, 1988). Rachman (1990, chap. 17) proposes that if agoraphobics show the same type of tendency as people with other types of fear to overpredict, then avoidance behavior may be a product of overprediction. Rachman goes on to suggest that, in agoraphobia as in claustrophobia, the occurrence and intensity of panic attacks will form the basis of predictions of fear. He also suggests that the anticipation of fear may itself provoke situational panic attacks, and that the occurrence of such panic in turn affects subsequent avoidance. Although this view is consistent with clinical observations of agoraphobics' pessimism regarding the impact of exposure despite repeated positive experiences of anxiety reduction (and reluctance to continue or repeat exposure sessions), more direct evaluation is required if overpredictions are to be regarded as a key factor in the maintenance of avoidance behavior (Rachman, 1990).

Other direct attempts to conceptualize and predict the relationship between avoidance related to naturally occurring panic outside the laboratory setting, or as a factor related to individual differences in anxiety responding among panic patients, have met with less success. A range of factors are weak predictors of avoidance behavior (Craske, Rapee, & Barlow, 1988; Mavissakalian, 1988; Warren, Zgourides, & Jones, 1989). Craske and Barlow (1988) review the data, and summarize the range of factors which have been found to relate to avoidance. They note that most important identified to date are expectancy of anxiety, the extent to which patients have a predominantly somatic focus of anxiety, the extent to which patients perceive panic to be related to negative consequences beyond anxiety itself (as in anxiety sensitivity), and access to safety signals. They attempt to synthesize this work within the context of Barlow's biobehavioral hypothesis. However, Craske and Barlow's discussion has a major limitation in the present context, because the approach applies only to situational and expected panic attacks. The importance of spontaneous and/or unexpected panic attacks in the origin and maintenance of avoidance is stressed by most theories including Barlow's own; Craske and Barlow propose no specific basis for this in their discussion, presumably because they assume that spontaneous panic attacks are biologically determined.

The cognitive hypothesis specifies that situations which generate anxiety might become the focus of avoidance, but this will only become debilitating when accompanied by fears that the anxiety might lead to other, more serious consequences. Furthermore, the scope of avoidance could be regarded as broader than generally encompassed by behavioral descriptions (e.g., Marks, 1987b), with both subtle and more generalized avoidance playing a major role in the maintenance of panic attacks (Salkovskis, 1988a). For example, a patient who interprets weak legs as a sign that he may collapse attempts to prevent collapse by holding on to nearby objects, tensing his legs and seeking a seat at the earliest possible opportunity. By doing so, he has a further effect; he prevents disconfirmation of his fears of collapse. The patient is unaware of the anxiety maintaining effects of his avoidant behavior (described here as "safety seeking" behavior). The effect of his behavior on his perception of what happened is to transform an incident which could have provided a disconfirmation of his fears into evidence bolstering his negative interpretation of symptoms, in the form that he has just experienced a "near miss." The situation would be interpreted as "If I had not prevented it by tensing my legs, then I would certainly have collapsed." The cognitive hypothesis predicts an internally logical match between such beliefs and behaviors during panic; for example, fears of loss of control should be associated with attempts to control oneself. Similar considerations can be applied to more generalized forms of actual avoidance and escape behavior; an agoraphobic who decides not to go shopping on a particular day may conclude "It's lucky I did not go, otherwise I would have had an enormous panic; if I had panicked today, I would certainly have collapsed." Once again, avoidance has "prevented" collapse. However, these types of behaviors involving rapid "escape" from the immediate situation are only logically possible when the feared catastrophe has external correlates; that is, simply leaving a situation has little value as a strategy for dealing with an impending heart attack. The main "usefulness" of such an escape strategy would be in situations where social evaluative concerns predominate; removing oneself from a place where there are other people is likely to be perceived as helpful both as direct catastrophe avoidance and as a means of removing the person from the scrutiny of others.

As panic and avoidance become more chronic, the behaviors involved probably become habitual and awareness of the specific cognitive component of avoidance will tend to diminish except when directly provoked by deliberate of accidental exposure. Often patients express this as the desire to escape; for example, when asked what went through his mind in a phobic situation, patients typically respond "I just had to get out." The cognition motivating this wish to escape is readily revealed by asking what, at that time, would he have thought the worst thing that could have happened should he have been unable to get out. Thus, the cognitive analysis of avoidance involves an analysis of what the patient is avoiding rather than just of the anxiety arousal/relief associated with the specific situation. The importance of the minutiae of anxiety related and safety seeking behavior also takes on a new importance, particularly for therapy.

The cognitive hypothesis proposes a specific set of links, both between panic and avoidance and vice versa. If panic patients fear imminent disasters because of their misinterpretation of bodily sensations, it follows that their behavior should be logically related to the avoidance of the feared catastrophes rather than of anxiety *per se*. This means that avoidance behavior would be meaningful but idiosyncratic both to the individual and to attacks involving specific types of misinterpretations and the individual experiencing them.

Support for this view comes from a study by Salkovskis, Clark, and Gelder (1996). In a preliminary investigation of this hypothesis, data were collected from a questionnaire devised as an assessment of safety seeking behaviors in panic together with assessments of some of the more common panic-related cognitions in a sample of 147 panic patients. The pattern of associations was consistent with predictions from the safety seeking perspective. Holding on to an object was associated with fears of fainting and being paralyzed with fear. Holding on to or leaning on another person was associated with fear of fainting. Sitting down was associated with fear of heart attack, acting foolishly, and being paralyzed with fear. As expected, keeping still was exclusively associated with fear of having a heart attack. Doing more exercise was only associated with being paralyzed with fear. Focusing attention on one's body was associated with thoughts of having a tumor, choking to death, and having a stroke. The specific association with controlling one's behavior were with thoughts of acting foolishly and not being able to control oneself. These cognitions were also associated with trying to move more slowly, reflecting the attempt to exercise control over one's behavior. The beliefs associated with looking for an escape route are exclusively with loss of control type concerns (acting foolish, losing control, and being paralyzed with fear), and not with thoughts concerning the more somatically-focused catastrophes. This finding was expected; leaving a situation is a relatively ineffective remedy for internally located problems, where the catastrophe would be carried with the patient out of the situation. By contrast, the negative consequences of loss of control type events are almost exclusively social evaluative. Asking people around for help was associated with thoughts of having a heart attack and choking to death, both likely to be perceived as extremely serious and uncontrollable somatic catastrophes, likely to require immediate external aid. The pattern of associations between specific behaviors and cognitions found was therefore consistent with the hypothesis that cognitions may "drive" the anxiety related behaviors in a meaningful and idiosyncratic way. Particular avoidant behaviors may be occurring because they are intended by the patients as logical reactions designed to avert feared catastrophes. These findings are consistent with the cognitive hypothesis as a "normalizing" account of panic and associated phenomena (particularly avoidant behaviors), highlighting the internal logic of the problem.

6.18.5.4 Agoraphobia

Within the cognitive framework, agoraphobia is therefore regarded as motivated avoidance. In agoraphobia, the avoidance is of situations which involve being away from safety, most commonly embodied by the person's home (or other trusted helper's homes) and its extension by other "safe" objects (e.g., the person's car). This does not necessarily mean that the person is free of anxiety when at home, but that they feel relatively safer there than in other situations. The fear in agoraphobic situations revolves around not being able to leave the situation for a safer one (or not being able to do so without embarrassment). This distinguishes agoraphobia from problems in which other types of phobic avoidance predominate because the person does not exclusively avoid (or endure with considerable anxiety) situations where they are likely to be negatively evaluated (as in social phobia) or be in a confined space (as in claustrophobia) or avoid because of some other type of anxiety (such as obsessive-compulsive disorder or post-traumatic stress disorder). However, fear of these other types of situations may be present in many people suffering from agoraphobia, as is it perfectly possible to experience multiple fears of this type.

Such a view has the merit of considering agoraphobic avoidance as a distinct psychological phenomenon (with much in common with other types of avoidance) rather than as a symptom of some other diagnosis. According to this view, people develop a pattern of avoidance of particular situations because these situations have become associated with unpleasant experiences (or the expectation of unpleasant experiences). These associations may be historical (and may therefore be relatively inaccessible to introspection), or may be more recent and vivid. This helps make sense of why panic attacks are neither necessary nor sufficient for the experience of agoraphobia, but why there is such a strong association. Other types of anticipated unpleasant experience which can lead to an agoraphobic pattern of behavior include epileptic attacks, incontinence, and irritable bowel (i.e., episodes where the person believes they are in danger of losing control over their bowels). Agoraphobia can develop in the elderly who fear falling and breaking their bones. This way of viewing agoraphobic avoidance has the further advantage of normalizing the behavior. Avoidance as a reaction to the perceived possibility of physical or social threat is, of course, an understandable reaction.

Within this framework, panic is the single most common factor associated with agoraphobia. This then raises the important question of why only some people develop agoraphobia when they experience panic attacks. (The same question clearly applies to irritable bowel problems, epilepsy, and other associated features; for understandable reasons research has tended to focus on the association with panic.) The second and more difficult question concerns those people who have no panic or "panic like symptoms" associated with their agoraphobia. There are several ways in which a tendency to avoid particular situations may result in actual avoidance behavior. The person may be particularly strongly motivated because their fear of agoraphobic situations is especially intense (as is usually the case when panic is present or anticipated) or they may believe they are particularly vulnerable to even relatively small amounts of fear, perhaps because they underestimate their ability to cope either with the agoraphobic situation or their own emotional reactions. Before a more detailed analysis of theories of agoraphobia, clinical factors thought to be meaningfully correlated with agoraphobia will be considered.

6.18.5.4.1 Correlates of agoraphobia

One view about agoraphobics which has been fairly prevalent is that they may show two basic dispositional qualities: dependence on others, and a tendency to cope with difficulties by using avoidance as a coping method (e.g., Andrews, 1966; McGennis, Nolan, & Hartman, 1977; Shafer, 1976). Andrews also suggests that this may be due to maternal overprotection. Mathews et al. (1981) attempted to survey the evidence which could be brought to bear with regard to this theory. The seven studies reviewed present somewhat equivocal results, although several did find evidence of overprotection, dependency, and somewhat unstable and more "anomalous" families (with more adopted children, etc.). Mathews et al. conclude that few properly conducted studies have been done, but that there is some suggestive evidence of dependence as a pre-existing trait. Reich, Noyes, and Troughton 1987) compared subjects who had panic disorder with varying amounts of agoraphobic avoidance, and found that about 40% of subjects with some (as opposed to no) phobic avoidance met criteria for dependent personality disorder, although it was not possible to determine whether the dependent behavior was primary or secondary to the agoraphobia. Kleiner and Marshall (1987) reported a high degree of dependent behavior in agoraphobics, prior to the onset of the phobia, and Borden and Geller (1981) in their review suggested that dependent personality traits are a key feature of agoraphobia.

There is some evidence that people who go on to develop agoraphobia may have been more fearful of being out and about in the world premorbidly. A history of school phobia is more common in agoraphobics than in panic patients, and the relatives of agoraphobics are more likely to suffer from school phobia or agoraphobia (Deltito, Perugi, Maremman, Mignani, & Cassano, 1986; Gittelman-Klein & Klein, 1984). There is also evidence of behavioral inhibition (defined as a tendency to have a constricted range of exploratory behavior) in agoraphobics premorbidly (e.g., see Rosenbaum et al., 1988). Renneberg, Chambless, and Gracely (1992) studied the incidence of *DSM-III-R* Axis II personality disorders in a large group of agoraphobics, and found that 56% of the sample had at least one personality disorder, of which avoidant personality was the single most frequent disorder. They also note that personality disorders seem more common in agoraphobics than in individuals with panic disorder (see also Friedman, Shear, & Frances, 1987, and Green & Curtis, 1988)

Chambless and Goldstein (1982) report that agoraphobic patients show low scores on the Bernreuter Inventory of self-sufficiency (although they also note that it incorporates items to do with activities to be done alone). Buglass, Clarke, Henderson, Kreitman, and Fredey (1977) reported no extra premorbid dependency in agoraphobics than in a matched group of normals, but found more conflict over dependency on the mother in the agoraphobic group. They may have contaminated their results by excluding those agoraphobics with marked dependency. Torgesen (1979) studied a small group of monozygotic twins of patients with agoraphobia. He found that the agoraphobic twin was more likely to have been dependent, self-doubting, and suggestible on a personality inventory, and to report having been like that since childhood. They were also likely to have been the second twin, and the smaller one. Chambless and Goldstein (1982) report two other studies on assertiveness in agoraphobics. In one they found that the agoraphobics were less assertive than simple phobics on therapist ratings, and in the other they found that they were less assertive on assertion inventories at intake than a group of college students, and more anxious about assertion. Rapee and Murrell (1988) also found that very avoidant individuals scored higher on the Eysenck Personality Questionnaire and lower on a measure of assertiveness than minimally avoidant subjects.

Most studies report a higher percentage of women in groups of agoraphobic patients than in groups with panic disorder without avoidance (Chambless, 1985; Wittchen, 1988), this has led to the stereotyped notion of the typical agoraphobic as a housebound housewife. Chambless (1985) does, however, point out that of her sample of 378 agoraphobics, 64 were men and only 52% were currently married, so that this description fitted less than half of the patients. There has, however, been considerable interest in the relationship between gender and avoidance in patients with agoraphobia. There has been speculation that, where the roles adopted are more "traditional," it is easier and more acceptable for women to become housebound than men, and for men to use alcohol as an alternative coping strategy. The stereotypical female gender role is also associated with traits of dependency and unassertiveness, both found to be prominent in agoraphobics. This literature is reviewed in detail by Chambless (1985).

A number of studies indicate that greater degrees of depression are associated with more severe avoidance in agoraphobia (Sanderson & Barlow, 1986; Klosko, Shadick, Heimberg, & Barlow, 1986; Telch et al., 1989; Vitaliano et al., 1987). This finding was not replicated in several other studies (e.g., Rapee & Murrell, 1988; Thyer, Himte, Curtis, Cameron, & Neese, 1985). Aronson and Logue (1987) reported a reliable difference between PD and PDA such that depression was more likely in individuals who later developed avoidance, and McNally and Lorenz (1987) found that in a group of PDA patients depression predicted a resurgence of avoidance. The association between depression and avoidance may be conceptualized within the theoretical framework of learned helplessness and attributional style, if it is found that PD patients with avoidance have weaker beliefs in their own ability to cope with stress. Klosko et al. (1986), who showed that panic disorder patients with avoidance were more likely to attribute negative events to stable and global factors, and Telch et al. (1989) showed that panic disorder patients with avoidance were likely to rate their ability to cope with panic as lower than patients with panic disorder without avoidance.

Several studies indicate that there is very little difference between the thoughts of possible physical consequences of bodily sensations (or in the extent to which they fear them) in people suffering from panic with agoraphobia as opposed to panic patients with minimal avoidance. The main difference observed is a greater fear of dizziness, and a higher frequency and belief in thoughts of fainting among the agoraphobics (Salkovskis, 1990; Telch et al., 1989). In the study by Telch et al., the more avoidant group showed much higher scores on scales reflecting fears of loss of mental control, or social ridicule. These results echo those of other studies which have found that agoraphobics tend to have more social evaluative concerns than patients with panic disorder (de Ruiter & Garssen, 1989). In the study by Telch et al., agoraphobics were found to have less perceived self-efficacy relating to dealing with panic than other panic patients. Hoffart (1995) has also shown that low scores on a measure of perceived self-efficacy predict situational avoidance more strongly than other measures, including a measure of catastrophic beliefs about the symptoms of panic.

6.18.5.5 The Cognitive Account and Research into Panic

Before examining previous work specifically designed to evaluate predictions derived from the cognitive hypothesis, it is first important to consider how one of the key observations thought to support a more biological view is encompassed. The provocation of panic by challenge tests such as lactate infusion and carbon dioxide inhalation has been described as a "biological marker" for panic, in that such provocations do not have the same effect on nonpanic control subjects. Clark (1986a,1986b, 1988) suggests that the research into "biological" panic provocation potentially represents some of the strongest support for the cognitive hypothesis. As noted above, a wide range of procedures provoke panic attacks in Panic Disorder patients but not in control subjects, although the procedures concerned have no identifiable common biochemical or physiological basis, making it difficult to envisage any specific biological mechanism. They do, however, have a common psychophysiological feature relevant to the cognitive hypothesis, which is that all of these substances induce a range of intense physical sensations in all subjects, panic or otherwise. Thus, the tendency to panic under such diverse provocations is consistent with the hypothesis that panic arises from the misinterpretation of bodily sensations. For example, it may be that panic patients interpret the effects of panic-inducing procedures as a sign of some kind of adverse reaction in their body to the inductions, all of which involve some kind of unusual physical procedure. To test this possibility, a number of researchers have carried out experiments into psychological factors affecting biological inductions.

In a particularly elegant experiment, Sanderson, Rapee, and Barlow (1989) demonstrated that the effects of carbon dioxide inhalation are modified if panic patients believe that they can control the inhalation by switching it off. In the experiment, the button which was supposed to turn off the carbon dioxide was not, in fact, connected, so the fact that only two patients out of the 10 in the perceived control condition panicked as compared to eight of 10 patients in the comparison group is attributable to the illusion of control. In an experiment in which hyperventilation was used as a provocation, the effect of hyperventilation was compared in panic patients and patients having a diagnosis as general anxiety disorder (Rapee, 1986). In this study, panic patients recognized the effects of hyperventilation as very similar to their experience during panic attacks, while patients suffering from generalized anxiety disorder did not. However, none of the 20 panic patients actually experienced a full attack, and the degree of similarity between hyperventilation and panic was highly variable. This result can be readily accounted for by the cognitive hypothesis, because in this experiment the patients will have a clear explanation of why symptoms are occurring, so that they are unlikely to misinterpret their sensations catastrophically, particularly as the provoking "agent" is a readily controllable behavior (breathing quickly and deeply). The absence of anxiety-provoking thoughts noted during the experiment would also have had the effect of minimizing the autonomic arousal which normally occurs at the same time as symptoms of hyperventilation during normally occurring panic, reducing the perception of similarity.

Another experiment from the same laboratory (Rapee, Mattick, & Murrel, 1986) manipulated cognitive factors prior to 50% CO_2/50% O_2 inhalations, a procedure which has been shown to induce panic attacks in panic patients but not in normal controls. In this experiment, subjects were either given a full account of the effects of the inhalation or given only minimal instructions. Panic patients who were not given the full explanation perceived the effects of the inhalation as more similar to panic and experienced more catastrophic cognitions than patients who were given an account of the effects. Groups of social phobics compared on the same basis showed no effect of instructions, demonstrating the specificity of this finding to panic patients as opposed to other anxious patients.

Recently, an experiment has been carried out with the most important of these substances, sodium lactate; Clark et al. (in preparation) demonstrated that, if patients are given detailed information which prevents misinterpretation of the symptoms of the infusion, panic is blocked. Thus, the cognitive hypothesis provides a testable account of the effects of so-called biological panic provocation tests, and this account has already received some support from experimental investigations of the mechanisms involved.

6.18.5.6 The Relevance of Psychobiological Factors to the Cognitive Hypothesis of Panic

The cognitive model does not rule out the possibility that some Panic Disorder patients may be physiologically more vulnerable to episodes of panic, for instance because of biological differences in autonomic lability associated with an increased perception of bodily or "mental" sensations. For example, it seems likely that there may be a genetic component to emotionality, which may make particular individuals prone to experiencing relatively intense reactions to threat. Indeed, such an observation would be quite consistent with the cognitive hypothesis. For example, if there was an identifiable subgroup of subjects who were particularly likely to experience intense and apparently inexplicable physiological sensations, it is quite likely that such subjects would be predisposed to panic. However, without the element of catastrophic misinterpretations, such individual differences would be insufficient to account for panic attacks. Indeed, many people actively seek out circumstances in which they are likely to experience heightened levels of arousal (e.g., horror films, exciting sports). This point is illustrated by the example of hyperventilation, which presents some problems of classification briefly discussed above, where it was concluded that although hyperventilation was neither necessary nor sufficient to account for panic, the data concerning the relationship between hyperventilation and panic strongly suggested some kind of link.

6.18.5.6.1 The cognitive link between panic and hyperventilation-induced symptoms

As described above there are a range of findings which strongly suggest that hyperventilation may be involved in some way in panic attacks. However, the hyperventilation hypothesis of Ley and others is not well supported by the experimental evidence. The association between panic and hyperventilation can best be accounted for by the cognitive hypothesis of panic, the early development of which was intimately connected with interest in the role of hyperventilation in the physical symptoms experienced by some patients (Clark & Helmsley, 1982; Salkovskis, 1988a, 1988b).

The way in which the cognitive hypothesis can account for the apparently conflicting data concerning the link between panic and hyperventilation is illustrated in Figure 2, which is a specific development of Figure 1. Although in his schema for the link between hyperventilation and panic, triggering sensations may sometimes arise from hyperventilation, this is not in any sense required in the way demanded by Ley (1985a, 1985b) (cf. Hibbert & Pilsbury, 1989). Upsetting thoughts, images, and situations which have previously been associated with panic are probably more likely to serve as triggers (Clark, 1988). The perception of threat results in apprehension, which in turn results in arousal and, in some instances, hyperventilation. If (and only if) the bodily sensations which then occur are interpreted as a sign of imminent personal harm of internal origin, then anxiety and bodily sensations are likely to show a further and sudden increase which will in turn be catastrophically misinterpreted, and so on round in a vicious circle culminating in a panic attack.

There are several additional points which emerge from this analysis. It is proposed that stress induced hyperventilation can be a source of intense and frightening bodily sensations, although such sensations are not intrinsically either unpleasant or frightening (Clark & Helmsley, 1982). Thus, hyperventilation can be an effect of anxiety and stress; if and only if the sensations are misinterpreted in a catastrophic fashion, then hyperventilation-induced sensations can then also become a "cause" of stress as part of the cognitive mechanism described above (Salkovskis & Clark, 1986). Increases in respiration are a normal part of the stress response and can occur without awareness, probably making misattribution more likely in vulnerable individuals.

Even in instances where hyperventilation does occur during a panic attack, note that the anxiety arising from the catastrophic misinterpretation of hyperventilation-induced bodily sensations is very likely to be accompanied by autonomic arousal in addition to the effects of hyperventilation. This means that, even in panic attacks in which hyperventilation

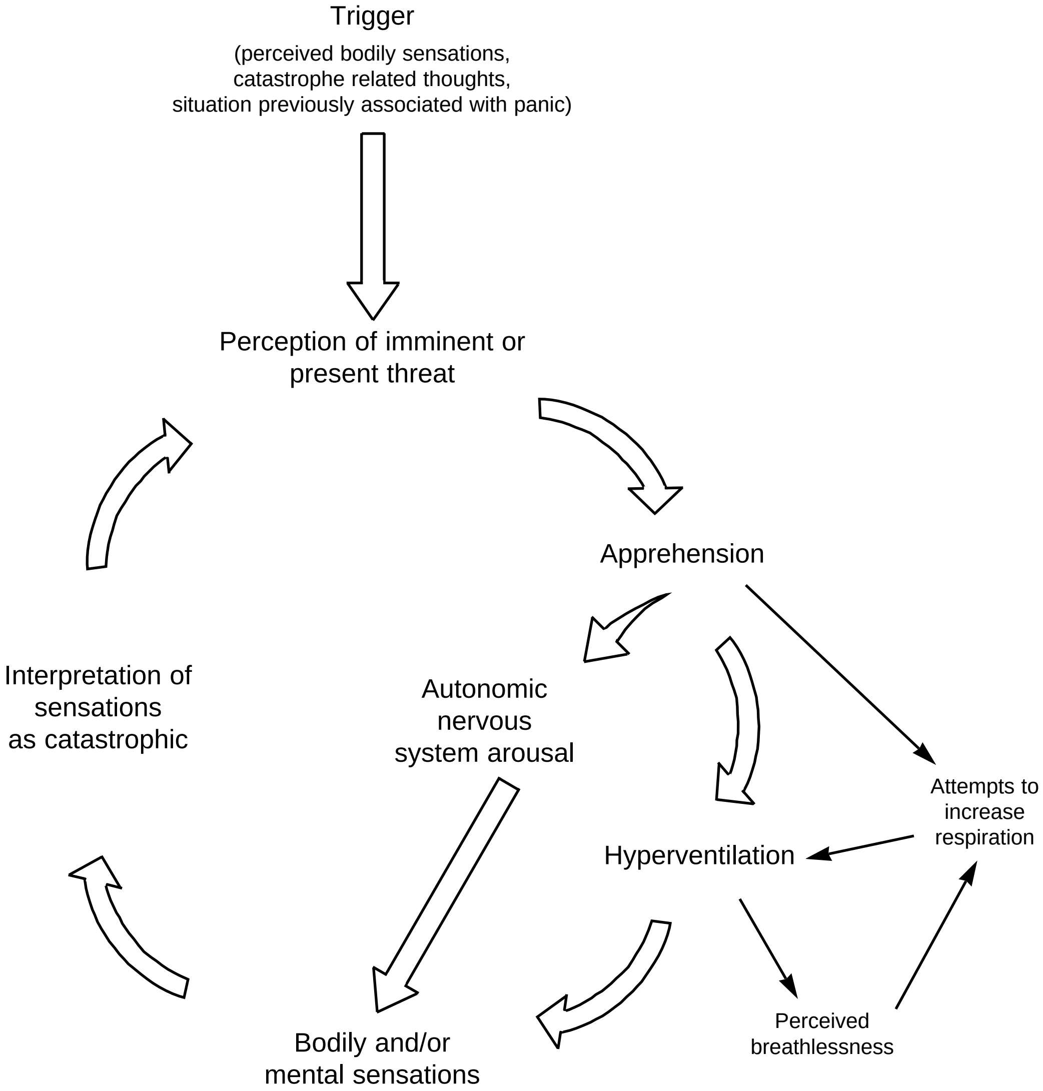
Trigger
(perceived bodily sensations,
catastrophe related thoughts,
situation previously associated with panic)
Perception of imminent or
present threat
Apprehension
Interpretation of
sensations
as catastrophic
Autonomic
nervous
system arousal
Attempts to
increase
respiration
Hyperventilation
Perceived
breathlessness
Bodily and/or
mental sensations

Figure 2 The principal components of the cognitive model applying to panic attacks in which hyperventilation is involved in the production of symptoms.

is involved, many of the sensations which occur may not be due to the effects of the respiratory alkalosis alone, making the importance of the effects of hyperventilation more difficult to determine.

There are several reasons why hyperventilation is a particularly powerful mechanism for the production of sensations which are subsequently misinterpreted. Hyperventilation can produce a wide range of sensations very quickly. Some of the sensations produced by hyperventilation are often unfamiliar even to patients who have previously experienced intense anxiety in specific situations (e.g., parasthesias, dyspnea, tetany), as is the sudden onset and rapid increase of symptoms. When dyspnea is associated with panic, as is often the case (Pollard, 1986), this may be misinterpreted as a sign that breathing is about to stop. This specific misinterpretation increases the likelihood that patients will respond to sensations by a deliberate increase in their breathing (Bass & Gardener, 1985), which will in turn have the effect of further intensifying sensations. Patients often report taking "deep breaths" in an attempt to calm themselves down when anxious. Such behavior is particularly associated with misinterpretation of breathlessness as a sign of imminent cessation of breathing. This means that, for some patients, panic associated hyperventilation might be entirely determined by voluntary effort directed at the prevention of suffocation.

Finally, there is evidence that repeated episodes of hyperventilation over a longer period can result in a change in the buffer set point as renal excretion of bicarbonate compensates for respiratory alkalosis of longer durations, as happens in altitude adaptation (Okel & Hurst, 1961). This adaptation has the effect of lowering resting p_aCO_2 levels and increasing the sensitivity of the buffer system to challenge. That is, low resting p_aCO_2 does not itself produce symptoms, but small changes in CO_2 (such as are associated with exercise or mild stress) produce relatively greater increases in respiration. This phenomenon will also have cognitive effects, as patients notice their increased sensitivity, and interpret this as further evidence for their belief that panic attacks are signs of a serious illness or harm. Although several studies have demonstrated that the resting p_aCO_2 levels of panic patients are significantly lower than those of normal controls, the only well-conducted study to examine the study of central CO_2 regulation failed to find any evidence of increased sensitivity (Woods et al., 1986). (This negative finding might be accounted for by the fact that panic patients did not differ from controls in resting level of p_aCO_2.) Thus, it can be concluded that the evidence is consistent with the view that hyperventilation can play an important role in the maintenance of panic in some individuals, but that it is neither necessary or sufficient. The symptoms of hyperventilation will only be involved in the mechanism of panic to the extent that they subject to specific catastrophic misinterpretations.

Consistent with such a cognitive view is the data from a study by Salkovskis and Clark (1990). Neurobiological hypotheses regard localized (brainstem) effects of pH changes (particularly secondary to respiratory alkalosis; Liebowitz et al., 1986) as being the basis of acute panic attacks. Hyperventilation-based theories (Fried, 1987; Ley, 1985a, 1985b; Lum, 1976) also regard the aversive effects of respiratory alkalosis to be central to the experience of panic. In both instances, hyperventilation is regarded as a biologically-based panic induction procedure. On the other hand, the cognitive hypothesis specifies that hyperventilation is not intrinsically aversive, and that the affect experienced will be a function of the way symptoms are interpreted. If this is so, this would also be further evidence that hyperventilation is neither necessary nor sufficient for the experience of panic, and highlights the intrinsic weakness of unmodified breathing control exercises as a means of treating panic. Salkovskis and Clark (1990) experimentally manipulated the way in which symptoms of hyperventilation were interpreted. Nonclinical subjects were provided with a relatively positive or relatively negative interpretation of the effects of hyperventilation, and the results evaluated both in terms of subjective and physiological responses. Those given a positive interpretation responded to the sensations as pleasant, while those given a negative interpretation found them aversive. Furthermore, affect was strongly correlated with the intensity of bodily sensations experienced, but only the affect congruent with the interpretation given.

6.18.6 EVALUATING THE COGNITIVE HYPOTHESIS

Clark et al. (1988) enumerated predictions directly arising from the cognitive hypothesis of panic disorder. Experimental studies have evaluated the cognitive account of laboratory-based biochemical panic inductions (Clark, 1993) and the presence of catastrophic misinterpretations of bodily sensations, both as an enduring characteristic of patients suffering from Panic Disorder and as a more acute phenomenon occurring during panic attacks themselves (reviewed in Salkovskis, 1989).

Chambless and Gracely (1989) evaluated the cognitions of patients with diagnoses of Panic Disorder, Agoraphobia with Panic, Generalized Anxiety Disorder, Social Phobia, Obsessive-Compulsive Disorder, and Depression. They found that Agoraphobics scored higher than all other groups on fear of physical sensations; both Agoraphobics and Panic Disorder patients scored higher than all other groups on thoughts that physical harm would result from their anxiety. They also found that, in Agoraphobics, fear of bodily sensations and agoraphobic thoughts (i.e, catastrophic cognitions) predicted self-reported avoidance behavior even when the effects of trait anxiety were statistically removed. Similar results were found by Hout van den, van der Molen, Griez, and Lousberg (1987) for fear of bodily sensations in Panic Disorder patients, who were significantly different from nonpanic "neurotic" patients and normal controls.

Rachman and Levitt (1985) tested claustrophobic subjects, who were asked to enter a very small confined space (a "closet"). They demonstrated that, in subjects experiencing a panic attack in this situation, breathlessness was particularly associated with thoughts of suffocation. Furthermore, they found an association between failure to habituate to repeated exposure and the specific presence of thoughts of suffocation, as they had predicted from a cognitive analysis based on the misinterpretation hypothesis.

McNally and Foa (1987) and Dattilio and Foa (1987) (described in Foa, 1988) used variants of questionnaires employed by Butler and Mathews (1983). The questionnaires were intended to tap two main phenomena:

(i) The subjective probability and perceived aversiveness of bodily sensations, in which subjects rate the likelihood of occurrence of interoceptive sensations and externally-based fear situations, together with judgments of how aversive such occurrences would be. Although the results show elevated subjective probabilities and cost in panic groups compared to anxious controls for interoceptive stimuli and not external fear situations, the results for this questionnaire are difficult to interpret because the *actual* probabilities for the experience of interoceptive (i.e., panic) symptoms are, by definition, likely to be considerably higher for the panic patients.

(ii) The interpretations which might be made of vignettes concerning interoceptive or external stimuli (McNally & Foa, 1987). This questionnaire provides the subject with a range of imaginary scenarios and asks the subject what they think is happening in the specified situation. Subjects respond by endorsing one of three options, which include two neutral interpretations and one negative or catastrophic option. For example, an item might be

> You notice your heart is beating quickly and pounding.
> Why?

and offer the options:

> Because you have been exercising
> Because there is something wrong with your heart
> Because you are excited

Panic patients were found to score higher than controls not only on internal (body sensation) items, but also on external threat situations. However, direct examination of Foa's version of the questionnaire reveals a specific problem, which is that there is confusion between negative interpretations *per se* and panic-related responses. Thus, many of the "negative" choices offered are "Because you are about to panic" or "Because you are very anxious." In the same paper, Foa presents laboratory data from an otherwise unpublished experiment by Mansueto, Grayson, and Foa on the subjective and physiological responses to imagery of internal and external fear stimuli. The results show that agoraphobic patients show significantly elevated subjective anxiety and heart rate responses to imagery of external fear situations, imagery of autonomic fear responses alone, and the two combined compared with neutral imagery. Specific phobics showed a similar pattern, except that imagery of autonomic fear responses alone was not different from neutral imagery. However, the analysis presented is solely within-group; Foa does not specify whether the groups differed on this item (i.e., whether there was a significant group X condition interaction). Inspection of means provided suggests that there may not have been.

In a study based on the McNally and Foa (1987) experiment described above, but designed more specifically to evaluate the cognitive hypothesis, Clark et al. (1997) used a further variant of the Butler and Mathews (1983) questionnaire in order to evaluate the tendency to misinterpret bodily sensations. This study found that panic patients were significantly more likely to misinterpret physical sensations associated with panic than other anxious patients who did not have panic attacks, but were no more likely than other anxious patients to misinterpret general and social situations.

The results of questionnaire studies of this kind are often criticized because it may be that patients are aware of how they are expected to respond. In a study designed to evaluate the interpretation question without possible subjective biases, a technique called Contextual Priming was used (Clark et al., 1988). This technique, derived from psycholinguistics, is intended to determine subject's expectations of particular meanings as provided by specific contextual cues. An incomplete sentence is projected on a computer screen; for example, "The cat drank the . . .". The sentence fades, and a word replaces it; the person watching simply has to say the word as quickly as possible, and the subject's reaction time is measured using a voice key triggered by the subject's verbal response. Subjects are faster at naming the final word when the word makes a relatively expected completion compared to relatively unexpected completions. In the above example, subjects would be appreciably quicker with the word "milk" than the word "vodka," revealing something about the person's belief about cats' drinking habits. In the experiment with panic patients, the incomplete sentences included items such as "My heart is beating quickly and pounding because I am . . . ". Two possible completions were used for this sentence: "dying" or "excited." Consistent with the cognitive theory, panic patients were significantly faster when catastrophic completions were presented (e.g., the word "dying" in the above example), while controls did not show this difference. This suggests that panic patients have specific negative beliefs, and that such misinterpretations can be rapid and automatic.

Although the demonstration that panic patients specifically misinterpret bodily sensations, it could still be argued that these patients show such misinterpretations because they are experiencing panic attacks rather than vice versa. More convincing and less ambiguous evidence for the theory comes from the prediction that activating misinterpretations should make panic patients (but not anxious controls) experience panic-like states. The task which was devised to do this was a simple word-reading task, where subjects were asked to read words such as

breathlessness-suffocate
palpitations-heart attack
unreality-insane

slowly, as pairs, pausing to dwell on each pair. Seventy-five percent of the panic patients asked to do this experienced a panic attack (as defined by *DSM-III*, viz., experiencing an increase of at least four physical symptoms and a sudden increase in anxiety), as opposed to 17% of anxious controls and no nonclinical controls. Another stressful task (backwards serial sevens) did not differentiate panic patients from anxious controls; although anxiety increased substantially in both groups, none of the subjects experienced a panic attack. Interestingly, no panic attacks occurred when the paired associates task was carried out with a group of panic patients who had been successfully treated by cognitive therapy.

A range of other experimental investigations has now been completed by the Oxford group and others. The cognitive model has been able to explain findings from the biological field (e.g., how biological challenges such sodium lactate provoke panic) and experimentally test these findings (Clark, 1993). On the other hand, biological theorists have been unable to account for findings in the cognitive field. Many specific predictions arising from biological research have been falsified; for example, Gorman et al. (1989) predict that psychological treatments will not reduce anxiety, but will only have effects on avoidance.

6.18.7 TREATMENT

An important clinical prediction from the cognitive hypothesis is that interventions which result in an enduring modification of catastrophic misinterpretations made by panic patients should result in clinically significant improvements in panic. It is important to note that the effectiveness of treatment does not provide any conclusive evidence for the hypothesis upon which it is based. However, ineffectiveness of treatment based on the hypothesis would present considerable problems (Gorman et al., 1989; Teasdale, 1988). Cognitive treatment now emphasizes two major components: (i) strategies designed to help patients to change misinterpretations of bodily sensations as they occur during panic attacks, and (ii) modifying the beliefs upon which such catastrophic misinterpretations are based.

6.18.7.1 The Theoretical Roots of Treatment

Both testable theoretical refinements (such as the importance of safety-seeking behavior as described above) and experimental studies can make contributions to the refinement of treatment techniques and strategies. The way that these examples have had an impact on treatment will be described next.

Some of the most important implications of the cognition-behavior links described above concern treatment (Salkovskis, 1991). In general, cognitive-behavioral treatment emphasizes the need to deal with idiosyncratic factors which bolster the continued misinterpretation of bodily sensations. For example, a patient who became confused during panic describes mentally "holding on to my sanity." In each successive panic he becomes *more* convinced that he would have gone mad were it not for this effort. Another example is the patient who, hundreds of very severe panic attacks later, still believes in each new attack that she is about to go crazy, pass out, or die. Clinically, questioning of such patients reveals that no disconfirmation has occurred because the patients believe that they have, in every instance, been able to take successful preventative action. Thus, each panic is a near miss and further confirmation of the risk. Once such behavioral responses are identified, the patient can be helped to begin the process of reappraisal by withholding such protective responses and learning the true extent of risk. Typically, this may involve suggesting to the patient that they challenge their worries by actively trying to cause the feared disaster, for example, going into the supermarket and trying to faint or try to go mad. This helps the patient to discover that their efforts to prevent these disasters have been misdirected; this can also help the patient to then reinterpret their usually considerable past experiences of such anxiety-provoking situations as true disconfirmations instead of "near misses." This particular type of association might be a fruitful one for investigation, as controlling one's mind is a common response in panic patients as well as other groups, including nonclinical subjects. Given the readiness with which patients can reproduce this type of behavior as either safety-seeking or coping when requested in the laboratory, experiments concerning the use of the same degree of effort directed at different targets (i.e., reduction of anxiety vs. preservation of sanity) could be used to assess the validity of the cognitive basis of such behavior and its putative anxiety-preserving effects when used as a safety-seeking behavior. A further implication is that for the most complete and enduring treatment effects, controlling panic should not be emphasized. Helping people to change their beliefs about the meaning of their symptoms in a more enduring way is the target, and should reduce subsequent vigilance and preoccupation.

The present analysis also suggests ways of combining cognitive procedures and brief exposure in a way which should be particularly effective in bringing about belief change without repeated and prolonged exposure being necessary. Thus, exposure sessions are devised in the manner of behavioral experiments, intended as an information gathering exercise directed towards invalidation of threat-related interpretations. Most previous studies in which cognitive and behavioral treatments have been combined (such as those reviewed in Marks, 1987a, 1987b) have used cognitive procedures as a way of dealing with general and "background" life stresses, most of which tend not to be directly relevant to the specific experience of anxiety. The particular strategy of using exposure as an exercise in testing alternative nonthreatening interpretations of experience would be predicted to succeed better than brief exposure with "supplementary" (threat-irrelevant) cognitive change procedures. That is, such a study should show that "general" cognitive therapy combined with exposure has an additive effect, while "anxiety-focused" cognitive therapy would be expected to multiply the effect of exposure, resulting in maximal cognitive change through behavioral experiments. Thus, according to the cognitive hypothesis, the value of behavioral experiments transcends mere exposure; such experiments allow patient and therapist to collaborate in the gathering of new information assessing the validity of nonthreatening explanations of anxiety and associated symptoms. (Findings consistent with this hypothesis have been found in Social Phobia by Wells, Clark, Salkovskis, Ludgate et al. (1995) and in Panic Disorder with Agoraphobia by Salkovskis et al. (1998).)

Since devising the paired associates task described above as an experimental task for use in the laboratory, we have found that it is a particularly effective strategy as part of treatment. Given that cognitive therapy involves a process of reattribution in which the panic sufferer is helped to understand that the symptoms which they experience do not arise from the feared catastrophes, the most effective way of doing this is to provide the person with an alternative, noncatastrophic explanation (the personalized vicious circle model). The paired associates can be a particularly powerful way of convincing the sufferer of the validity of the vicious circle model. A personalized version of the paired associates is used (derived from a careful assessment of the symptoms experienced and the misinterpretations made by that person). If the person begins to panic, they are asked what they make of the experience. If reading these particular pairs of words can induce panic, how does that fit with the idea that your symptoms mean that you are having a heart attack? How does this fit with the vicious circle explanation?

We have also discovered, in the context of therapy, that quite simple strategies will readily modify panic patients' response to the task. For example, a patient undergoing therapy was concerned that her initial panic when reading the paired associates meant that she was so out of control that even reading words on a card could make her have a panic. She was asked to read the same card, but where the hyphen appeared in the pair (e.g., breathless-suffocate) she was to say "does not mean I will." This procedure completely abolished her anxious response to the card. When asked what she had learned, she replied that she now understood that changing the meaning of panic symptoms was the key to dealing with panic attacks, and that she did not need to exercise control over panic. This clinical illustration shows both the potential strengths of the technique and the dangers of its uncritical use.

6.18.7.1.1 Cognitive therapy for panic: some general issues

Cognitive therapy has similar prerequisites to other forms of psychological treatment; primary among these is the formation of a good, collaborative therapeutic relationship. The style of therapy emphasizes guided discovery, so that the patient is helped to understand and try to modify their problems by using careful questioning. Therapy also makes extensive use of diaries and other aids; a particularly helpful adjunct to the most efficient use of treatment time is to tape record therapy sessions so the patient can listen to the tape at home. These elements are considered to be a minimum baseline against which any effective therapy needs to be judged. The general cognitive theory predicts that specific techniques will be of minimal effectiveness in the absence of a good therapeutic relationship as described above. Part of this relationship involves a proper assessment of the idiosyncratic problems experienced by the person in their own unique individual context. Therapy is based on the view that if a person holds a "distorted" belief, this is because they have good present, contextual, and historical reasons for doing so. This is equally true not only in the obvious case of general feelings of low self-worth in a person whose main difficulty is low self-esteem, but also for specific misinterpretations of bodily sensations in someone suffering from frequent panic attacks. For example, it is often important not only to pinpoint just what the person means by the idea of losing control, but also loss of control is a major issue for them. Most commonly, these types of concern reflect "developmental" issues.

Evaluation of specific techniques or strategies must therefore take place against the necessary background of good and consistent general therapeutic skills. People (clients and therapists) seldom work well with those who they tend to distrust. Many of the negative findings in the field can be explained not only by the failure of the specific intended manipulation (which is, of course, a fatal error) but also by inappropriate therapy styles. One of the most common errors in this respect is the use of cognitive-behavioral treatments in a prescriptive fashion. Another example is to abstract a single element (e.g., the emphasis by some on "correct breathing," by others on "positive thinking"). In order to ensure therapy integrity, therapy evaluations need to combine assessment of specific techniques with the evaluation of therapy style and quality of relationship. Other factors which are of key importance (and which need to be equated across any more refined comparisons of different therapies) are therapist time, the initial credibility and expectation of improvement engendered by the different treatments prior to any differential treatment effects, any assessments and self-monitoring, and so on.

6.18.7.1.2 Clinical implementation of cognitive-behavioral therapy

Once it has been established that panic is a major problem for the person and a more general assessment of the patient's situation has been carried out, a specific panic-oriented assessment is undertaken. This assessment focuses on eliciting an account of recent panic attacks and devising an individualized vicious circle explanation of the sequence of events during panic (loosely based on Figure 1, but always tailored to the idiosyncratic description of the patient). The links which comprise the vicious circle should be elicited by systematic questioning, thus ensuring both that the assessment is accurate and that the patient is clear about the way in which the formulation has been reached. The patient has to recall the last major panic they experienced, and then to describe, step by step, how it developed.

Questioning is guided by use of the sequence in the upper part of Figure 3, leading to the construction of a vicious circle as shown the lower part of the figure. Thus, if the first thing the patient noticed was a physical sensation, the next question should concern the way the patient interpreted it (e.g., "When you noticed your heart pounding, what went through your mind at that time?"). Another question which may be useful at this crucial stage in the assessment is "Right then, at that time, what did you think was the worst thing that could happen?" Once a specific catastrophic thought (or image) is elicited, questioning shifts to the anxiety experienced (e.g., "When you had the thought, "I'm having a heart attack,' how did that make you feel?"). The next question concerns physical reactions to anxiety (e.g., "when you became anxious about having a heart attack, how did that affect you physically?"). Once the vicious circle has been elicited in this way, it is summarized verbally and outlined on paper for the patient. Thus, in the above example, the summary would go: "Let's see if I have understood what you are telling me.

During the attack you experienced on Wednesday night, you first noticed your heart pounding. You then had the thought 'I'm having a heart attack,' and this, not surprisingly, frightened you. The anxiety made your heart pound more, which seemed to confirm the idea that you were having a heart attack, and so on, round in a vicious circle as I've drawn here. Does that seem to you to be right?" In this way, a positive description of the factors involved in the production of panic is built up from information provided by the patient. Note that the emphasis is on providing a positive, noncatastrophic account of panic rather than on disproving feared catastrophes. The internal logic of panic is emphasized as a way of normalizing the intensity of the anxiety experienced (e.g., "How do you think that anyone would feel if they noticed that they were feeling dizzy and believed that this meant they were about to faint? ... When someone becomes anxious because they have a very frightening thought, such as about having a heart attack, how do you think they would feel physically?").

A helpful pointer for the assessment, and a further illustration of the logic of panic, is the way in which catastrophic interpretations tend to be meaningfully associated with particular symptoms. Examples of the way in which this association works are illustrated in Table 2. Assessment as described above is the first component of treatment, because it already involves a major element of decatastrophizing, helping patients to make sense of the panics which had previously been especially frightening because they appeared to be inexplicable. The procedures used for further decatastrophizing are divided into verbal techniques, in which discussion of the patients' past experience is used to change beliefs and behavioral experiments, in which the patient and therapist devise experiments to test the way in which particular processes may be involved in panic attacks.

(i) *Discussion.* The vicious circle model is again the starting point for discussion aimed at helping the patient to attribute the symptoms they experience during panic away from catastrophic explanations and on to less threatening interpretations. The therapist starts from the assumption that, if the patient makes catastrophic interpretations, they do so because they have evidence which they find convincing. This evidence may be past experiences (e.g., a close relative or friend who experienced similar symptoms associated with serious illness), in generalized attitudes (e.g., "If I can't control my thoughts I will go mad"), in the intensity or nature of the physical sensations themselves (e.g., "I feel so faint that I must be about to pass out"), and so on.

Discussion involves helping the patient to evaluate such evidence and consider alternatives. For example, assessment of a patient revealed that, during panic, she interpreted dizziness, unsteadiness, and feelings of unreality as a sign that she was about to faint or collapse. Attacks invariably started with an episode of dizziness, often triggered by postural change or feeling angry and frustrated. The evidence for her belief that she would faint was that the symptoms, especially faintness, became very intense when she was at the height of an attack, and that she had, on a previous occasion, fainted. During the first stage of this discussion, the patient and therapist went over the factors involved in a true faint, and agreed that a drop in blood pressure was required in order to faint. The patient, when asked what she thought might happen to her blood pressure during panic, said that she knew that it went up a little and this therefore must mean that she could not faint during panic. The discussion then moved on to her previous fainting episode, which had occurred after giving blood. The reasons for blood pressure drop in this specific situation were discussed. The patient was also asked to compare the experience of fainting with the experience of feeling faint in panic; she identified these as quite different. In particular, she noted that when she fainted she was not anxious, but instead became rather sleepy and distant. When asked whether that distant feeling resembled the unreality she experienced during panic, she said that it was quite different, and summarized by saying that she would easily recognize a true faint if it were to occur, and that fainting was particularly unlikely during panic attacks. This was an especially useful manoeever, because it meant that previously anxiety-provoking situations had now become safety signals. Following on from this, the discussion moved on to consideration of the nature of "feeling faint," and whether such feelings were necessarily associated with the fact of fainting. The comparison with having a splitting headache was used; did a splitting headache mean that one's head could split? The patient agreed that this was unlikely and her ratings of belief in the idea that she might faint during a panic attack dropped substantially, but not completely. When questioned about the basis of her beliefs, she said that the symptoms were so intense at the time she believed that there had to be something wrong that could lead her to faint. The therapist then asked her if she had been excited recently. She said that she had. Three weeks previously, she had thought that she had won a very large sum of money (on the football pools). She described her bodily sensations at that time, which were almost identical in type and intensity

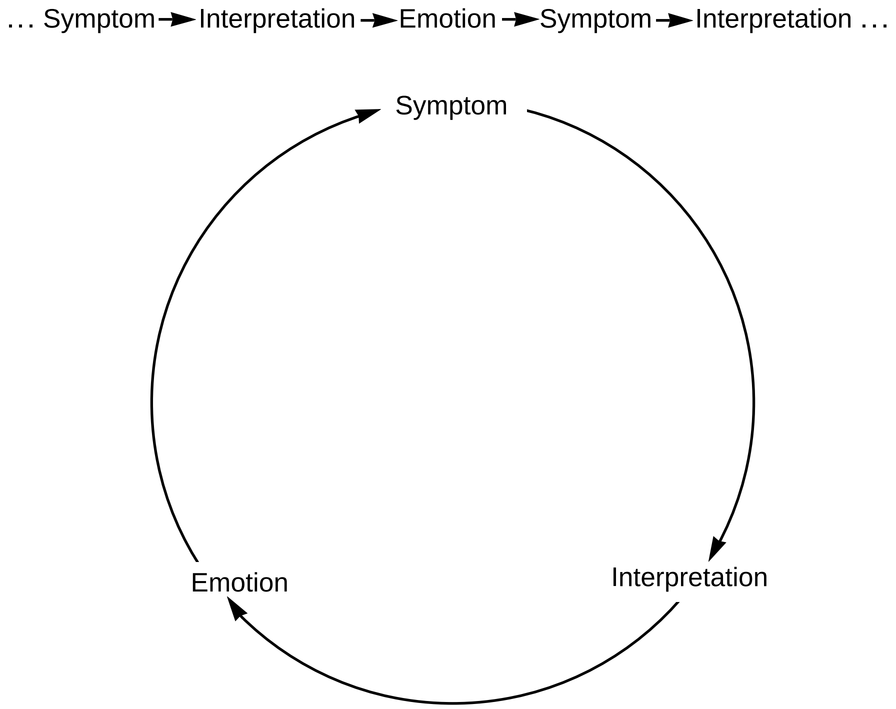

Figure 3 Construction of a vicious circle.

to those she experienced during a panic, although she had not panicked and found the experience pleasurable and exciting. Careful questioning elicited the main difference to be the way in which she interpreted the sensations of excitement; she had not had any catastrophic thoughts. The conclusion she drew from this discussion was that the symptoms she experienced during panic attacks were the normal effects of adrenaline, and not a sign of something physically amiss. Finally, ways in which the patient could apply the reasoning outlined above during panic attacks were devised and discussed.

(ii) *Behavioral experiments*. Discussion in which the evidence upon which patients base their catastrophic misinterpretations is reviewed continues throughout therapy. However, patients sometimes do not have ready access to information which would help them to challenge their misinterpretations, or understand the alternative explanations which have been generated in the course of discussion but require further evidence to support the alternative. This is where behavioral experiments are particularly valuable as part of treatment. A wide range of behavioral experiments are useful. One of the most commonly used involves the use of hyperventilation, which often occurs in association with panic when the associated symptoms are misinterpreted. The metabolic alkalosis which accompanies hyperventilation produces a wide range of physical symptoms, such as tachycardia, dizziness, paresthesia, breathlessness, and so on; these symptoms are a particularly potent basis of misinterpretation because they are often relatively unfamiliar to the patient in the context of his or her previous experience of anxiety (e.g., paresthesia and dizziness).

Particularly common is the experience of breathlessness, which, if misinterpreted as a sign of impending suffocation, can have the effect of making the patient try to breathe more, worsening the breathlessness, and so on. An especially helpful way of dealing with this is to ask the patient to show how he or she breathes (or tries to breathe) at the height of a panic attack. If the pattern the patient demonstrates appears to involve an element of hyperventilation, the therapist asks the patient to continue to breathe in the same way (or, if it is only a slight degree of hyperventilation, to exaggerate this) for a few minutes, then to describe the physical feelings they experience. This type of procedure reproduces the bodily sensations experienced during a panic attack by about 50%; the fact that the patient is simply doing what they would normally be doing during a panic makes this demonstration of the role of hyperventilation particularly convincing to him or her. Discussion of the extent to which the experience of deliberate overbreathing was similar or different from the patient's naturally occurring panic will usually show that the physical sensations were similar, but the experience differed in the extent to which anxiety was present. This, of course, is because the patient in the therapist's office has an unambiguous and nonthreatening explanation of their symptoms, and because the patient is in control of the onset (and therefore the offset) of the symptoms themselves. This difference can itself help pinpoint the crucial role of misinterpretations; thus, the patient is taught that hyperventilation does not cause panic, but is a likely source of symptoms which can be misinterpreted. Subsequently, patients may find it helpful to learn simple breathing control techniques which they can use during panic attacks as part of the general cognitive strategy. It is important to stress that breathing control is a helpful way of dealing with the nuisance value of symptoms, but only at times when the patient notices the physical symptoms. It emphasizes that the control of breathing is in no way essential; indeed, in the later stages of therapy the patient deliberately hyperventilates in situations which tend to make them anxious; this exercise provides further confirmation of the noncatastrophic nature of bodily sensations. Throughout, the emphasis is on the idea that the key factor is misinterpretation; breathing control is most useful as evidence for a noncatastrophic interpretation of the symptoms. During an acute attack, the patient reminds him- or herself (a) that the physical symptoms of panic were reproduced by overbreathing in the therapist's office, suggesting that the symptoms are both a sign of anxiety induced by catastrophic misinterpretations and also a result of stress-induced hyperventilation, and (b) that strategies which help them to control their thoughts and breathing, although unnecessary for their safety, have the effect of reducing the symptoms, thus providing further confirmation of the noncatastrophic interpretation. (Note that, despite encouraging early clinical reports, breathing control without the crucial cognitive element has been shown to be ineffective; the cognitive element appears to be necessary for successful treatment using respiratory control.)

The therapist who wishes to deal effectively with the problems of patients suffering from panic attacks should also deal with those safety-seeking behaviors which the patient uses during panic attacks. Typically, three main types of behavior are seen: (a) avoidance of situations which the patient believes might provoke panic (e.g., avoiding supermarkets); (b) escape from a situation when a panic attack occurs (e.g., leaving a shop once the symptoms of panic begin); and (c) safety-seeking behaviors carried out during panic with the intention of actively preventing the feared catastrophe (e.g., when dizziness leads to the thought "I'll faint," holding on to another person, shopping trolley, or sitting down). Each of these behaviors can have the effect of maintaining the panic-related beliefs; the patients logically infer that they have prevented the occurrence of feared catastrophes by their behavior (e.g., "If I had gone to the supermarket yesterday, then I would have passed out; If I had not left immediately I would have fainted; If I had not sat down then I would have fainted." Thus, behavior of the type described above prevents disconfirmation of the feared catastrophes, and transforms potential disconfirmations into "near misses." In cognitive therapy for panic, patients are helped to understand how their behavior can maintain anxiety (in the same way as outlined above), and the therapist helps the patient to use this information to construct an explanation of the persistence of their beliefs. For example, a patient summarized this as "Although I believe that my behavior has prevented disaster, maybe this is only because I have never tried to check this out, and have just assumed that my *not* fainting was because of the precautions I took. By continuing to behave *as if* I might have a heart attack might have kept my anxiety going, because it was a constant reminder that I have to be careful, that I am vulnerable. I have never given myself the opportunity to test the truth of my fears." The therapist uses the patient's understanding of how behaviors can prevent disconfirmation of feared disasters to devise direct tests; for example, a patient who felt weak during attacks and believed that she was going to fall would tense her muscles to prevent collapse. In the office therapy session, patient and therapist tried out how weak muscles would have to be to result in collapse. The patient decided to test the possibility of falling at times when she felt panicky and weak by trying to relax her muscles in the same way she had in the therapy session. This exercise had the effect of completely changing her belief in the possibility of collapsing, integrating cognitive techniques with short behavioral experiments. These behavioral experiments, in which the patient experiences a clear disconfirmation of the beliefs which had previously been the focus of safety-seeking behavior, allow rapid and complete modification of both avoidance behavior and catastrophic interpretations.

6.18.7.1.3 *Effectiveness of cognitive-behavioral therapy*

Early clinical studies on both cognitive and behavioral treatments used consecutive single case series (see Salkovskis & Clark, 1986, and Barlow, 1988, for a review of these). These early studies used quasiexperimental designs which required the measurement of a baseline, followed by the demonstration of change associated with treatment that was greater in magnitude than any variation seen during the baseline period. Not only was it possible to show this, but it was also shown in these relatively unselected samples that the changes associated with treatment (most patients being panic free at the end of treatment) persisted at two-year follow-up. These studies were criticized for not being sufficiently selective about the patients included and not adhering exactly to a preset treatment protocol. For example, several of the patients in Clark, Salkovskis, and Chalkley's (1985) study required and were given additional treatment for nonpanic related problems. As these people were treated as part of routine clinical practice, a certain (small) amount of time was devoted to dealing with problems such as assertiveness difficulties, marital problems, and so on. This was fully documented in the published report, but was nevertheless a source of some criticism, and highlighted the need to conduct better controlled studies. A later single case series (Salkovskis, Clark, & Hackmann, 1991) was used in a quite different way. In this series, it was demonstrated that it was possible to make a substantial impact on panic attack by using brief (two hour) and purely cognitive (verbal) procedures. This study was consistent with the hypothesis that cognitive change (focusing on the modification of the misinterpretation of bodily sensations) could be an element in the effectiveness of cognitive therapy for panic in reducing panic attack frequency. This study also gathered preliminary evidence that cognitive procedures which do not target misinterpretations may not be effective.

Following these clinically based studies which demonstrated that the effectiveness of cognitive therapy was unlikely to be due spontaneous remission or to nonspecific factors (see also Beck, 1988), it was clearly important to assess the effectiveness of cognitive therapy compared with the best available alternative psychological and pharmacological treatments. Clark et al. (1994) compared cognitive therapy to (i) a waiting list control group; (ii) imipramine delivered in the optimum fashion by an experienced biologically-oriented psychiatrist (gradually increased to high doses maintained to six months then gradually tapered); (iii) applied relaxation, a specially modified form of relaxation for use with panic (Ost, 1988); and (iv) cognitive therapy. All active treatments were delivered by experienced therapists in the fashion currently considered optimum for each therapy; thus, psychological treatments were 12 sessions over 12 weeks, pharmacotherapy was continued for much longer.

Important features of this study were (i) the selection of patients who were particularly unlikely to show spontaneous remission (i.e., those patients experiencing frequent panic attacks with a minimum duration of six months); (ii) delivering each treatment in an optimum fashion, with considerable care being taken with treatment integrity (e.g., sessions audiotaped and checked against the competency scale for cognitive therapy adapted for panic treatment); (iii) ensuring that each treatment had an equally credible rationale and expectance of good outcome at the outset (i.e., before any differential effectiveness was apparent); (iv) the inclusion of as many common elements as possible (e.g., daily recording/self-monitoring, printed therapy handouts early in treatment, self-exposure instructions in the later stages of treatment, and so on); (v) use of blind assessor ratings as well as self-ratings; (vi) inclusion of high end-state functioning criteria in outcome assessment; (vii) maximum attention to follow-up completeness; and (viii) inclusion of theoretically relevant variables to predict relapse.

The outcome of this study clearly indicated (i) that all three treatments were active (superior to waitlist); (ii) that both in the short term (12 weeks) and long term (one-year post-treatment follow-up), cognitive therapy was significantly superior to both applied relaxation and imipramine; (iii) that the degree to which catastrophic misinterpretations of bodily sensations remained at the end of treatment predicted relapse at one-year follow-up, and that this remained the case when any residual symptomatology at the end of treatment was partialled out; and (iv) that all treatments resulted in good levels of high end-state functioning at the end of treatment and at one-year follow-up (CT 80% and 70%; AR 25% and 32%; imipramine 40% and 45%).

The particular effectiveness of cognitive therapy has been replicated in a number of centers; a recent feature of this work has been the coordination of outcome measures allowing detailed comparisons across centers (Clark, 1994). Treatment outcome studies are also described particularly well in Clark (1994).

The next stages in this line of research is to (i) improve cost-effectiveness, that is, to evaluate whether it is possible to conduct treatment in a smaller number of sessions without sacrificing the outstanding results found by our own and other groups, and, if a briefer treatment is found to be as effective. Work carried out by the Oxford group on the treatment of panic disorder suggests that it is possible, by the use of self-help manuals integrated with therapist contact, to reduce therapy time to around five sessions without loss of efficacy (Clark et al., in preparation); (ii) to return to the initial strategy of applying the therapy to a less selected clinical sample of people suffering from panic attacks; (iii) to evaluate the generalizability of cognitive methods and theory developed in panic by adapting these to the needs of other patient groups such as hypochondriasis (Salkovskis, 1989; Salkovskis & Clark, 1993; Salkovskis & Warwick, 1986; Warwick & Salkovskis, 1990).

It is important to note that a range of other psychological treatments have been shown to be more or less effective in the amelioration of panic attacks, including applied relaxation (Ost, 1988), exposure to interoceptive sensations (Hout van den et al. 1987), paradoxical intention (Mitchelson, Mavissakalian, & Marchione, 1988) and a treatment which combines elements of traditional cognitive therapy, respiratory control, and relaxation, termed Panic Control Treatment (Barlow, 1991; Barlow & Cerny, 1988). Barlow (1997) suggests appropriately that some caution is needed before concluding that psychological treatments have been fully developed to be effective in the long term.

6.18.8 CONCLUSION

Both biological and psychological theories have generated a great deal of specific research into Panic Disorder and its treatment. Over the last decade, the evidential basis for psychological approaches has strengthened, in large part as a result of the specific predictions made both for experimental and clinical studies. By contrast, biological theories have been been gradually weakening. Nevertheless, there is a specific need to consider further the nature of the links between catastrophic interpretations and acute panic and physiological and behavioral correlates. Such clarification should improve the extent to which the misinterpretations hypothesis can be matched against findings from the biological and behavioral fields, and allow yet more specific predictions to be made in these areas. The clinical benefits of a more coherent approach to the conceptualization of panic attacks, Panic Disorder, and associated phenomena would be considerable.

6.18.9 REFERENCES

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. Rev.). Washington, DC: Author.

Andrews, J. D. W. (1966). Psychotherapy of phobias. *Psychological Bulletin, 66,* 455–480.

Aronson, T. A., & Logue, C. M. (1987). On the longitudinal course of panic disorder: Development history and predictors of phobic complications. *Comprehensive Psychiatry, 28,* 344–355.

Arrindell, W. A. (1980). Dimensional structure and psychopathology correlates of the Fear Survey Schedule (FSSIII) in a phobic population: a factorial definition of agoraphobia. *Behaviour Research and Therapy, 18,* 229–242.

Baker, R. (1989). Introduction: Where does "Panic Disorder" come from? In R. Baker (Ed.), *Panic disorder: Theory research and therapy*. Chichester, UK: Wiley

Ballenger, J. C. (1986). Pharmacotherapy of the panic disorders. *The Journal of Clinical Psychiatry, 47*(Suppl.), 27–32.

Ballenger, J. C., Burrows, G. D., DuPont, R. L., Lesser, I. M., Noyes, R., Pecknold, J. C., Rifkin, A., & Swinson, R. P. (1988) Alprazolam in panic disorder and agoraphobia: results from a multicenter trial. *Archives of General Psychiatry, 45,* 413–423.

Barlow, D. H. (1988). *Anxiety and its disorders*. New York: Guilford Press.

Barlow, D. H. (1991). Disorders of emotion. *Psychological Inquiry, 2*(1), 58–71.

Barlow, D. H. (1997). Cognitive-behavioral therapy for panic disorder: Current status. *Journal of Clinical-Psychiatry, 58*(Suppl. 2), 32–37.

Barlow, D. H., & Cerny, J. (1988). *Psychological treatment of panic*. New York: Guilford Press.

Barlow, D. H., & Craske, M. G. (1988). The phenomenology of panic. In S. J. Rachman & J. Maser (Eds.), *Panic: Psychological perspectives*. Hillsdale, NJ: Erlbaum.

Barlow, D. H., Vermilyea, J., Blanchard, E. B., Vermilyea, B. B., Di Nardo, P. A., & Cerny, J. A. (1985). The phenomenon of panic. *Journal of Abnormal Psychology, 94,* 320–328.

Bass, C., & Gardner, W. (1985). Emotional influences on breathing and breathlessness. *Journal of Psychosomatic Research, 29,* 559–609.

Beard, G. (1869). Neurasthenia, or nervous exhaustion. *Boston Medical and Surgical Journal, 3,* 217–221.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International University Press.

Beck, A. T. (1988). Cognitive approaches to panic disorder: theory and therapy. In S. J. Rachman & J. Maser (Eds.), *Panic: Psychological perspectives*. Hillsdale, NJ: Erlbaum.

Beck, A. T., Emery, G., & Greenberg, R. L. (1985). *Anxiety disorders and phobias*. New York: Basic Books.

Beck, A. T., Laude, R., & Bohnert, M. (1974). Ideational components of anxiety neurosis. *Archives of General Psychiatry, 31,* 319–325.

Berg, I., Butler, A., & Pritchard, J. (1974). Psychiatric illnesses in the mothers of school phobic adolescents. *British Journal of Psychiatry, 128,* 466–467.

Borden, J. W., & Geller, E. S. (1981). Agoraphobia: Appraisal of research and proposal for an integrative model. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behaviour modification, 12,* 1–66. New York: Academic Press.

Bourdon, K. H., Boyd, J. H., Rae, D. S., Burns, B. J. et al. (1988). Gender differences in phobias: Results of the ECA community survey. *Journal of Anxiety Disorders, 2,* 227–241.

Boyd, J. H. (1986). Use of mental health services for the treatment of panic disorder. *American Journal of Psychiatry, 143,* 1569–1574.

Buglass, D., Clarke, J., Henderson, A. S., Kreitman, N., & Presley, A. S. (1977). A study of agoraphobic housewives. *Psychological Medicine, 7,* 73–86.

Butler, G., & Mathews, A. M. (1983). Cognitive processes in anxiety. *Advances in Behaviour Research and Therapy, 5,* 51–63.

Chambless, D. (1985). The relationship of severity of agoraphobia to associated psychopathology. *Behavior Research and Therapy, 23*(3), 305–310.

Chambless, D. L., Caputo, G. C., Bright, P., & Gallagher,

R. (1984). Assessment of fear of fear in agoraphobics: the Bodily Sensations Questionnaire and the Agoraphobic Cognitions Questionnaire. *Journal of Consulting and Clinical Psychology, 52,* 1090–1097.

Chambless, D. L., & Goldstein, A. (1980). The treatment of agoraphobia. In A. Goldstein & E. B. Foa (Eds.) *Handbook of behavioral interventions*. New York: Wiley.

Chambless, D. L., & Goldstein, A. J. (1982). *Agoraphobia: Multiple perspectives on theory and treatment*. New York: Wiley.

Chambless, D. L., & Gracely, E. J. (1989). Fear of fear and the anxiety disorders. *Cognitive Therapy and Research, 13,* 9–20.

Charney, D. S., & Heninger, G. R. (1986). Abnormal regulation of noradrenergic function in panic disorders. *Archives of General Psychiatry, 43,* 1042–1054.

Clark, D. M. (1979). *Therapeutic aspects of increasing* pCO_2 *by behavioral means,* Unpublished M. Phil. thesis, University of London.

Clark, D. M. (1986a). A cognitive approach to panic. *Behaviour Research and Therapy, 24,* 461–470.

Clark, D. M. (1986b). Cognitive therapy for anxiety. *Behavioural Psychotherapy, 14,* 283–294.

Clark, D. M. (1988). A cognitive model of panic. In S. J. Rachman & J. Maser (Eds.), *Panic: Psychological perspectives*. Hillsdale, NJ: Erlbaum.

Clark, D. M. (1993). Cognitive mediation of panic attacks induced by biological challenge tests. *Behaviour Research and Therapy, 15,* 75–84.

Clark, D. M. (1994). Cognitive therapy for panic disorder. In B. E. Wolfe & J. D. Maser (Eds.), *Treatment of panic disorder: A consensus development conference*. Washington, DC: American Psychiatric Press

Clark, D. M. (1996). Panic disorder: from theory to therapy. In P. M. Salkovskis (Ed.), *Frontiers of cognitive therapy*. New York: Guilford Press.

Clark, D. M. (1997). Panic disorder and social phobia. In D. M. Clark & C. G. Fairburn (Ed.), *The science and practice of cognitive-behaviour therapy*. Oxford, UK: Oxford University Press.

Clark, D. M., & Hemsley, D. R. (1982). The effects of hyperventilation: individual variability and its relation to personality. *Journal of Behaviour Therapy and Experimental Psychiatry, 13,* 41–47.

Clark, D. M., Salkovskis, P. M., Anastasiades, P., Middleton, H., Hackmann, A., & Gelder, M. G. (in preparation). *Cognitive mediation of lactate induced panic.*

Clark, D. M., Salkovskis, P. M., & Chalkley, A. J. (1985). Respiratory control as a treatment for panic attacks. *Journal of Behaviour Therapy and Experimental Psychiatry, 16,* 23–30.

Clark, D. M., Salkovskis, P. M., Gelder, M. G., Koehler, C., Martin, M., Anastasiades, P., Hackmann, A., Middleton, H., & Jeavons, A. (1988). Tests of a cognitive theory of panic. In I. Hand & H. U. Wittchen (Eds.), *Panic and phobias II*. New York: Springer.

Clark, D. M., Salkovskis, P. M., Hackmann, A., Middleton, H., Anastasiades, P., & Gelder, M. (1994). A comparison of cognitive therapy, applied relaxation and imipramine in the treatment of panic disorder. *British Journal of Psychiatry, 164*(6), 759–769.

Clark, D. M., Salkovskis, P. M., Ost, L. G., Breitholtz, E., Koehler, K. A., Westling, B. E., Jeavons, A., & Gelder, M. (1997). Misinterpretation of body sensations in panic disorder. *Journal of Consulting and Clinical Psychology, 65*(2), 203–213.

Clum, G. A., & Pendrey, D. (1987). Depression symptomatology as a non-requisite for successful treatment of panic with anti-depressant medications. *Journal of Anxiety Disorders, 1,* 337–344.

Cohen, M. E., & White, P. D. (1951). Life situations, emotions and neurocirculatory asthenia (anxiety neurosis, neurasthenia, effort syndrome). *Psychosomatic Medicine, 13,* 335–357.

Conway, P. G., Tejani-Butt, S., & Brunswick, J. (1987). Interaction of *Beta* adrenergic agonists and antagonists with brain *Beta* adrenergic receptors *in vivo*. *The Journal of Pharmacology and Experimental Therapeutics, 241,* 755–762.

Coryell, W., Noyes, R., & Clancy, J. (1983). Panic disorder and primary unipolar depression: a comparison of background and outcome. *Journal of Affective Disorders, 5,* 311–317.

Crowe, R. R., Pauls, D. L., Slymen, D. J., & Noyes, R. (1980). A family study of anxiety neurosis: morbidity risk in families of patients with and without mitral valve prolapse. *Archives of General Psychiatry, 37,* 77–79.

Craske, M. G., & Barlow, D. H. (1988). A review of the relationship between panic and avoidance. *Clinical Psychology Review, 8,* 667–686.

Craske, M. G., Rapee, R. M., & Barlow, D. H. (1988). The significance of panic expectancy for individual patterns of avoidance. *Behavior Therapy, 19,* 577–592.

Degonda, M., & Angst, J. (1993). The Zurich study. XX. Social phobia and agoraphobia. *European Archives of Psychiatry and Clinical Neuroscience, 243*(2), 95–102.

Deltito, J. A., Perugi, G., Maremmani, I., Mignani, V., & Cassano, G. B. (1986). The importance of separation anxiety in the differentiation of panic disorder from agoraphobia. *Psychiatric Developments, 3,* 227–236.

Donnell, C. D., & McNally, R. J. (1990). Anxiety sensitivity and panic attacks in a nonclinical population. *Behaviour Research and Therapy, 28,* 83–86.

Eaton, W. W., Holzer, C. E., von Korff, M., et al. (1984). The design of the ECA surveys: The control and measurement of error. *Archives of General Psychiatry, 41,* 942–948.

Ehlers, A., & Margraf, J. (1989). The psychophysiological model of panic attacks. In P. M. G. Emmelkamp, W. T. A. M. Everaerd, F. W. Kraaimaat, & M. J. M. van Son (Eds.), *Fresh perspectives on anxiety disorders: Annual series of European research in behaviour therapy: 4*. Amsterdam: Swets and Zeitlinger

Ehlers, A., Margraf, J., & Roth, W. T. (1988). Selective information processing, interoception and panic attacks. In I. Hand & H. U. Wittchen (Eds.), *Panic and phobias II*. New York: Springer.

Faravelli, C. (1985). Life events preceding the onset of panic disorders. *Journal of Affective Disorders, 9,* 103–105.

File, S. E., Deakin, J. F. W., Longden, A., & Crowe, T. J. (1979). An investigation of the role of the locus coeruleus in anxiety and agonistic behaviour. *Brain Research, 169,* 411–420.

Finlay-Jones, R., & Brown, G. W. (1981). Types of stressful life event and the onset of anxiety and depressive disorders. *Psychological Medicine, 11,* 801–815.

Foa, E. B. (1988). What cognitions differentiate panic disorder from other anxiety disorders? In I. Hand & H. Wittchen (Eds.), *Panic and phobias II*. New York: Springer.

Foa, E. B., Steketee, G., & Young, M. C. (1984). Agoraphobia: Phenomenological aspects, associated characteristics and theoretical considerations. *Clinical Psychology Review, 4,* 431–457.

Freedman, R. R. (1989). Ambulatory monitoring findings on panic. In R. Baker (Ed.), *Panic disorder: Theory research and therapy*. Chichester, UK: Wiley

Freud, S. (1894). The justification for detaching from neurasthenia a particular syndrome: the anxiety-neurosis. Reprinted in Freud, S. (1940). *Collected papers* (Vol. 1). London: Hogarth Press.

Freud, S. (1895). Obsessions and phobias: their physical mechanisms and their aetiology. Reprinted in S. Freud, (1940). *Collected papers* (Vol. 1). London: Hogarth Press.

Fried, R. (1987). *The Hyperventilation syndrome*. Baltimore: Johns Hopkins University Press.

Friedman, C. J., Shear, M. K. & Frances, A. (1987). *DSM-III* personality disorders in panic patients. *Journal of Personality Disorders, 1,* 132–135.

Garakani, H., Zitrin, C. M., & Klein, D. F. (1984). Treatment of panic disorder with imipramine alone. *American Journal of Psychiatry, 141,* 446–448.

Garssen, B., van Veenendaal, W. V., & Bloemink, R. (1983). Agoraphobia and the hyperventilation syndrome. *Behaviour Research and Therapy, 21,* 643–649.

Gibson, H. B. (1978). A form of behaviour therapy for some states diagnosed as "affective disorder." *Behavioural Research Therapy, 16,* 191–195.

Gittleman-Klein, R., & Klein, D. F. (1984). Relationship between separation anxiety and panic and agoraphobic disorders. *Psychopathology,* Suppl. 1, 56–65.

Gledhill, N., Beirne, G. J., & Dempsey, J. A. (1975). Renal response to short-term hypocapnia in man. *Kidney International, 8,* 376.

Goldstein, A. J., & Chambless, D. L. (1978). A reanalysis of agoraphobia. *Behavior Therapy, 9,* 47–59.

Gorman, J. M., Cohen, B. S., Liebowitz, M. R., Fyer, A. J., Ross, D., Davies, S. O., & Klein, D. F. (1986). Blood gas changes and hypophosphatemia in lactate-induced panic. *Archives of General Psychiatry, 43,* 1067–1071.

Gorman, J. M., Fyer, A. J., Ross, D. C., Cohen, B. S., Marinez, J. M., Liebowitz, M. R., & Klein, D. F. (1987). Normalization of venous pH, $p\mathrm{CO}_2$, and bicarbonate levels after blockade of panic attacks. *Psychiatry Research, 14,* 57–65.

Gorman, J. M., Liebowitz, M. R., Fyer, A. J., & Stein, J. (1989). A neuroanatomical hypothesis for panic disorder. *American Journal of Psychiatry, 146,* 148–161.

Gray, J. A. (1982). *The neuropsychology of anxiety*. New York: Oxford University Press.

Green, M., & Curtis, G. (1988). Personality disorders in panic patients. *Journal of Personality Disorders, 2,* 303–314.

Hallam, R. S. (1978). Agoraphobia: A critical review of the concept. *British Journal of Psychiatry, 133,* 314–319.

Hallam, R. S. (1985). *Anxiety: Psychological perspectives on panic and agoraphobia*. London: Academic Press.

Hardonk, H. J., & Beumer, H. M. (1979). Hyperventilation syndrome. In P. J. Vincken & G. W. Bruyn (Eds.), *Handbook of neurology, 38,* 618–624.

Hawton, K., Salkovskis, P., Kirk, J., & Clark, D. M. (1989). *Cognitive behaviour therapy for psychiatric problems: A practical guide*. Oxford: Oxford University Press.

Hibbert, G. A. (1984). Ideational components of anxiety: Their origin and content. *British Journal of Psychiatry, 144,* 618–624.

Hibbert, G. A. (1984). Hyperventilation as a cause of panic attacks. *British Medical Journal, 288,* 263–264.

Hibbert, G. A. (1986). Ambulatory monitoring of transcutaneous $p\mathrm{CO}_2$. In L. Lacey & J. Sturgeon (Eds.), *Proceedings of the 15th European Conference on Psychosomatic Research*. London: Libby.

Hibbert, G. A., & Chan, M. (1989). Respiratory control: its contribution to the treatment of panic attacks: A controlled study. *British Journal of Psychiatry, 154,* 232–236.

Hibbert, G. A., & Pilsbury, D. (1989). Hyperventilation: Is it a cause of panic attacks? *British Journal of Psychiatry, 155,* 805–809.

Hoffart, A. (1995). Cognitive mediators of situational fear in agoraphobia. *Journal of Behaviour Therapy and Experimental Psychiatry, 26*(4), 313–320.

Holloway, W., & McNally, R. (1987). Effects of anxiety sensitivity on the response to hyperventilation. *Journal of Abnormal Psychology, 96,* 330–334..

Hout, van den, M. R. (1988). Metabolism, reflexes and cognitions in experimental panic. In S. J. Rachman & J. Maser (Eds.), *Panic: Psychological perspectives*. Hillsdale, NJ: Erlbaum.

Hout, van den, M. A., & Griez, E. (1982). Cognitive factors in carbon dioxide therapy. *Journal of Psychosomatic Research, 26,* 209–214.

Hout, van den, M. A., & Griez, E. (1986). Experimental panic. Biobehavioural notes on empirical findings. In I. Hand & H. U. Wittchen (Eds.), *Panic and phobias*. New York: Springer.

Hout, van den, M. A., & Griez, E. (1984). Panic symptoms after inhalation of carbon dioxide. *British Journal of Psychiatry, 144,* 503–507.

Hout, van den, M. A., & Griez, E. (1985). Peripheral panic symptoms occur during changes in alveolar carbon dioxide. *Comprehensive Psychiatry, 26,* 381–387.

Hout, van den, M. A., van der Molen, G. M., Griez, E., & Lousberg, H. (1987). Specificty of interoceptive fear panic disorders. *Journal of Psychopathology and Behavioural Assessment, 9,* 99–109.

Katon, W. (1984). Panic disorder and somatization. *The American Journal of Medicine, 77,* 101–106.

Katon, W., Vitaliano, P. P., Russo, J., Jones, M., & Anderson, K. (1987). Panic disorder: Spectrum of severity and somatization. *Journal of Nervous and Mental Disease, 175,* 12–19.

Kendler, K. S., Heath, A. C., Martin, N. G., & Eaves, L. J. (1987). Symptoms of anxiety and symptoms of depression. Same genes, different environment? *Archives of General Psychiatry, 44,* 451.

Kerr, W. J., Dalton, J. W., & Gleibe, P. A. (1937). Some physical phenomena associated with anxiety states and their relation to hyperventilation. *Annals of Internal Medicine, 11,* 961–992.

Klein, D. F. (1964). Delineation of two-drug responsive anxiety syndromes. *Psychopharmacologia, 5,* 397–408.

Klein, D. F. (1980). Anxiety reconceptualised. *Comprehensive Psychiatry, 21,* 411–427.

Klein, D. F. (1981). Anxiety reconceptualised. In D. F. Klein & J. Rabkin (Eds.), *Anxiety: New research and changing concepts*. New York: Raven Press.

Klein, D. F. (1993). False suffocation alarms, spontaneous panics, and related conditions. An integrative hypothesis. *Archives of General Psychiatry, 50,* 306–317.

Klein, D. F. (1996). Panic disorder and agoraphobia: Hypothesis hothouse. *Journal of Clinical Psychiatry, 57*(Suppl. 6), 21–27.

Klein, D. F., & Fink, M. (1962). Psychiatric reaction pattern to imipramine. *American Journal of Psychiatry, 119,* 438.

Klein, D. F., & Klein, H. M. (1988). The status of panic disorder. *Current Opinion in Psychiatry, 1,* 177–183.

Klein, D. F., & Klein, H. M. (1989a). The nosology, genetics and theory of spontaneous panic and phobia. In P. Tyrer (Ed.), *Psychopharmacology of anxiety*. Oxford, UK: Oxford University Press.

Klein, D. F., & Klein, H. M. (1989b). The definition and psychopharmacology of spontaneous panic and phobia. In P. Tyrer (Ed.), *Psychopharmacology of anxiety*. Oxford, UK: Oxford University Press.

Klein, D. F., Ross, D. C., & Cohen, P. (1987). Panic and avoidance in agoraphobia. Application of path analysis to treatment studies. *Archives of General Psychiatry, 44,* 377-385.

Kleiner, L., & Marshall, W. (1987). The role of interpersonal problems in the development of agoraphobic with panic attacks. *Journal of Anxiety Disorders, 1,* 313-324.

Klosko, J. S., Shadick, R., Heimberg, R., & Barlow, D. H. (1986). *The attributional styles of patients diagnosed with agoraphobia with panic attacks versus patients diagnosed with panic disorder*. Paper presented at the annual meeting of the Association for the Advancement of Behaviour Therapy. Chicago, November.

Lader, M., & Mathews, A. M. (1968). A physiological model of phobia anxiety and desensitisation. *Behaviour Research and Therapy, 6,* 411–418.

Lader, M., & Mathews, A. M. (1970). Physiological changes during spontaneous panic attacks. *Journal of Psychosomatic Research, 14,* 377–380.

Leckman, J. F., Weissman, M. M., Merinkangas, K. R., Pauls, D. L., & Prusoff, B. A. (1983). Panic disorder and major depression. *Archives of General Psychiatry, 40,* 1055–1060.

Lewis, A. (1967). Problems presented by the ambiguous word "anxiety" as used in psychopathology. *Israel Annals of Psychiatry, 5,* 72–88.

Ley, R. (1985a). Agoraphobia, the panic attack and the hyperventilation syndrome. *Behaviour Research and Therapy, 23,* 79–82.

Ley, R. (1985b). Blood, breath and fears: A hyperventilation theory of panic attacks and agoraphobia. *Clinical Psychology Review, 5,* 271–285.

Liebowitz, M. R., Fyer, A. J., Gorman, J. M., Dillon, D., Appleby, I. L., Levy, G., Anderson, S., Levitt, M., Palij, M., Davies, S. O., & Klein, D. F. (1984). Lactate provocation of panic attacks: I. Clinical and behavioural findings. *Archives of General Psychiatry, 41,* 764–770.

Liebowitz, M. R., Fyer, A. J., Gorman, J. M., Dillon, D., Davies, S., Stein, J. M., Cohen, B. S., & Klein, D. F. (1985). Specificity of lactate infusions in social phobia versus panic disorders. *American Journal of Psychiatry, 142,* 947–950.

Liebowitz, M. R., Gorman, J. M., Fyer, A., Dillon, D., Levitt, M., & Klein, D. F. (1986). Possible mechanisms for lactate's induction of panic. *American Journal of Psychiatry, 143,* 495–502.

Liebowitz, M. R., Gorman, J. M. Fyer, A. J., Levitt, M., Dillon, D., Gail, L., Appleby, I. L., Anderson, S., Palij, M., Davies, S. O., & Klein, D. F. (1985). Lacate provocation of panic attacks: II. Biochemical and physiological findings. *Archives of General Psychiatry, 42,* 709–719.

Lum, L. C. (1975). Hyperventilation: the tip and the iceberg. *Journal of Psychosomatic Research, 19,* 375–383.

Lum, L. C. (1976). The syndrome of habitual chronic hyperventilation. In O. W. Hill (Ed.), *Modern trends in psychosomatic medicine* (Vol. 3). London: Butterworths.

McGennis, A., Nolan, G., & Hartman, M. (1977). The role of self-help associations in agoraphobia: One years experience with "out and about." *Irish Medical Journal, 70,* 10–13.

McNair, D. M., & Kahn R. J. (1981). Imipramine compared with a benzodiazepine for agoraphobia. In D. F. Klein & J. Rabkin (Eds.), *Anxiety: New research and changing concepts.* New York: Raven Press.

McNally, R. J. (1990). Psychological approaches to panic disorder: a review. *Psychological Bulletin, 108,* 403–419.

McNally, R. J. (1994). *Panic disorder: A critical analysis.* New York: Guilford Press.

McNally, R. J., & Foa, E. B. (1987). Cognition and agoraphobia: Bias in the interpretation of threat. *Cognitive Therapy and Research, 11,* 567–581.

McNally, R. J., & Lorenz, M. (1987). Anxiety sensitivity in agoraphobics. *Journal of Behavior Therapy and Experimental Psychiatry, 18,* 3–11.

Magarian, G. J. (1982). Hyperventilation syndromes: infrequently recognised common expressions of anxiety and stress. *Medicine, 61,* 219–236.

Margraf, J., & Ehlers, A. (1988). Panic attacks in nonclinical subjects. In I. Hand & H. U. Wittchen (Eds.), *Panic and phobias II.* Berlin: Springer.

Margraf, J., Ehlers, A., & Roth, W. T. (1986a). Biological models of panic disorder and agoraphobia: A review. *Behaviour Research and Therapy, 24,* 553–567.

Margraf, J., Ehlers, A., & Roth, W. T. (1986b). Sodium lactate infusions and panic attacks: A review and critique. *Psychosomatic Medicine, 48,* 23–51.

Marks, I. M. (1981). *The cure and care of neurosis.* Chichester, UK: Wiley.

Marks, I. M. (1987a). Behavioral aspects of panic disorder. *American Journal of Psychiatry, 144,* 1160–1165.

Marks, I. M. (1987b). *Fears, phobias and rituals.* New York: Oxford University Press.

Marks, I. M., Grey, S., Cohen, S. D., Hill, R., Mawson, D., Ramm, E. M., & Stern, R. S. (1983). Imipramine and brief therapist-aided exposure in agoraphobics having self exposure homework: a controlled trial. *Archives of General Psychiatry, 40,* 153–162.

Mason, S. T. (1981). Norepinephrine in the brain: progress in theories of behavioural function. *Progress in Neurobiology, 16,* 263–303.

Mathews, A. M., Gelder, M. G., & Johnston, D. W. (1981). *Agoraphobia: Nature and treatment.* New York: Guildford Press.

Mavissakalian, M. (1988). The relationship between panic, phobic and anticipatory anxiety in agoraphobia. *Behaviour Research and Therapy, 26,* 235–240.

Michelson, L. (1988). Cognitive, behavioural and psychophysiological treatments and correlates of panic. In S. J. Rachman & J. Maser (Eds.), *Panic: Psychological perspectives.* Hillsdale, NJ: Erlbaum.

Michelson, L., Mavissakalian, M., & Marchione, K. (1988). Cognitive, behavioural and psychophysiological treatments of agoraphobia: a comparative outcome trial. *Behavior Therapy, 19,* 97–120.

Moran, C., & Andrews, G. (1985). The familial occurrence of agoraphobia. *British Journal of Psychiatry, 146,* 262–267.

Myers, J. K., Weissman, M. M., Tischler, C. E., et al. (1984). Six-month prevalence of psychiatric disorders in three communities. *Archives of General Psychiatry, 41,* 959–967.

Norton, G. R., Dorward, J., & Cox, B. J. (1986). Factors associated with panic attacks in nonclinical subjects. *Behavior Therapy, 17,* 239–252.

Norton, G. R., Harrison, B., Hauch, J., & Rhodes, L. (1985). Characteristics of people with infrequent panic attacks. *Journal of Abnormal Psychology, 94,* 216–221.

Noyes, R., Anderson, D. J., Clancy, J., Crowe, R. R., Slymen, D. J., Ghoneim, M. M., & Hinrichs, J. V. (1984). Diazepam and propranolol in panic disorder and agoraphobia. *Archives of General Psychiatry, 41,* 287–292.

Noyes, R., Reich, J., Clancy, J., & O'Gorman, T. W. (1986). Reduction in hypochondriasis with treatment of panic disorder. *British Journal of Psychiatrty, 149,* 631–635.

Nutt, D. J. (1986). Increased central alpha2-adrenoceptor sensitivity in panic disorder. *Psychopharmacology, 90,* 268–269.

Okel, B. B., & Hurst, J. W. (1961). Prolonged hyperventilation in man: associated electrolyte changes and subjective symptoms. *Archives of Internal Medicine, 108.*

Ost, L. G. (1988) Applied relaxation vs. progressive relaxation in the treatment of panic disorder. *Behaviour Research and Therapy, 26,* 13–22.

Pilsbury, D., & Hibbert, G. (1987). An ambulatory system for long term continuous monitoring of transcutaneous $p\text{CO}_2$. *Bulletin of European Physiopathololgical Respiration, 23,* 9–13.

Pitts, F. N., & McClure, J. N. (1967). Lactate metabolism in anxiety neurosis. *New England Journal of Medicine, 277,* 1329–1336.

Pollack, M. H., & Otto, M. W. (1997). Long-term course and outcome of panic disorder. *Journal of Clinical Psychiatry, 58*(Suppl. 2), 57–60.

Pollack, M. H., Otto, M. W., Rosenbaum, J. F., Sachs, G.

S., C, O. N., Asher, R., & Meltzer Brody, S. (1990). Longitudinal course of panic disorder: Findings from the Massachusetts General Hospital Naturalistic Study. *Journal of Clinical Psychiatry, 51* (Suppl. A), 12–16.

Pollack, M. H., & Smoller, J. W. (1995). "The longitudinal course and outcome of panic disorder." *Anxiety Disorders: Longitudinal Course and Treatment, 18,* 785–801.

Pollard, C. A. (1986). Respiratory distress during panic attacks associated with agoraphobia. *Psychological Reports, 58,* 61–62.

Pollard, C. A., Pollard, H. J., & Corn, K. J. (1989). Panic onset and major events in the lives of agoraphobics: A test of contiguity. *Journal of Abnormal Psychology, 98,* 318–321.

Rachman, S. J. (1990). *Fear and courage* (2nd ed.). New York: Freeman.

Rachman, S. J., & Bichard, S. (1988). The overprediction of fear. *Clinical Psychology Review, 8,* 303–312.

Rachman, S. J., & Levitt, K. (1985). Panics and their consequences. *Behaviour Research and Therapy, 23,* 585–600.

Rainey, J. M., Pohl, R. B., Williams, M., Kritter, E., Freedman, R. R., & Ettedugi, E. (1984). A comparison of lactate and isoproterenol anxiety states. *Psychopathology, 17,* 74–82.

Rapee, R. M. (1985). Distinction between panic disorder and generalized anxiety disorder. Clinical presentation. *Australian and New Zealand Journal of Psychiatry, 19,* 227–232.

Rapee, R. M. (1986). Differential response to hyperventilation in panic disorder and generalized anxiety disorder. *Journal of Abnormal Psychology, 95,* 24–28.

Rapee, R. M., & Barlow, D. H. (1989). Psychological treatment of unexpected panic attacks: cognitive/behavioural components. In R. Baker (Ed.), *Panic disorder: Theory, research and therapy*. Chichester, UK: Wiley.

Rapee, R., Mattick, R., & Murrell, E. (1986). Cognitive mediation in the affective component of spontaneous panic attacks. *Journal of Behaviour Therapy and Experimental Psychiatry, 17,* 245–253.

Rapee, R., & Murrell, E. (1988). Predictors of agoraphobic avoidance: An explanatory analysis. *Journal of Anxiety Disorders, 2,* 203–217.

Razran, G. (1961). The observable unconscious and the inferable conscious in current Soviet psychophysiology: interoceptive conditioning, semantic conditioning and the orienting reflex. *Psychological Review, 68,* 81–150.

Redmond, D. E. (1979). New and old evidence for the involvement of a brain norepinephrine system in anxiety. In W. E. Fann, I. Karacan, A. D. Pokorny, et al. (Eds.), *Phenomenology and treatment of anxiety*. New York: SP Scientific and Medical Books.

Redmond, D. E. (1985). Neurochemical basis for anxiety and anxiety disorders: evidence from drugs which decrease human fear and anxiety. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders*. Hillsdale, NJ: Erlbaum.

Redmond, D. E., & Huang, Y. (1979). Current concepts II: New evidence for a locus coeruleus-norepinephrine connection with anxiety. *Life Sciences, 25,* 2149–2162.

Reich, J., Noyes, R., & Troughton, E. (1987). Dependent personality disorder associated with phobic avoidance in patients with panic disorder. *American Journal of Psychiatry, 144,* 323–326.

Reiss, S., Peterson, R. A., Gursky, G. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behaviour Research and Therapy, 24,* 1–8.

Renneberg, B., Chambless, D., & Gracely, E. (1992). Prevalence of SCID-diagnosed personality disorders in agoraphobic outpatients. *Journal of Anxiety Disorders, 6,* 111–118.

Rifkin, A., Klein, D. F., Dillon, D., & Levitt, M. (1981). Blockade by imipramine or desimpramine of panic induced by sodium lactate. *American Journal of Psychiatry, 138,* 676–677.

Rosenbaum, J. F., Biederman, J., Gersten, M., Hirschfield, D., et al. (1988). Behavioural inhibition in children of parents with panic disorder and agoraphobia: A controlled study. *Archives of General Psychiatry, 45*(5), 463–470.

Roth, M. (1959). The phobic-anxiety-depersonalisation syndrome. *Proceedings of the Royal Society of Medicine, 52,* 587–594.

Roy-Byrne, P. P., & Uhde, T. W. (1988). Exogenous factors in panic disorder: Clinical and research implications. *Journal of Clinical Psychiatry, 49,* 56–61.

de Ruiter, C., & Garssen, B. (1989). Social anxiety and fear of bodily sensations in panic disorder and agoraphobia: a matched comparison. *Journal of Psychopathology and Behavioral Assessment, 11,* 175–184.

Salkovskis, P. M. (1986).The cognitive revolution: New way forward, backward somersault or full circle? *Behavioural Psychotherapy, 14,* 278–282.

Salkovskis, P. M. (1988a). Phenomenology, assessment and the cognitive model of panic. In S. J. Rachman & J. Maser (Eds.), *Panic: psychological perspectives*. Hillsdale, NJ: Erlbaum.

Salkovskis, P. M. (1988b). Hyperventilation and anxiety. *Current Opinion in Psychiatry, 1,* 76–82.

Salkovskis, P. M. (1989). Cognitive models and interventions in anxiety. *Current Opinion in Psychiatry, 2,* 795–800.

Salkovskis, P. M. (1990). *The nature of and interaction between cognitive and physiological factors in panic attacks and their treatment,* Unpublished Ph.D. thesis, University of Reading.

Salkovskis, P. M. (1991). The importance of behaviour in the maintenance of anxiety and panic: A cognitive account. *Behavioural Psychotherapy, 19,* 6–19.

Salkovskis, P. M., & Clark, D. M. (1986). Cognitive and physiological processes in the maintenance and treatment of panic attacks. In I. Hand & U. H. Wittchen (Eds.), *Panic and phobias*. New York: Springer.

Salkovskis, P. M., & Clark, D. M. (1990). Affective response to hyperventilation; a test of the cognitive model of panic. *Behaviour Research and Therapy, 28,* 51–61.

Salkovskis, P. M., & Clark, D. M. (1993). Panic and hypochondriasis. *Advances in Behaviour Research and Therapy, 15,* 23–48.

Salkovskis, P. M., Clark, D. M., & Gelder, M. G. (1996). Cognition–behaviour links in the persistence of panic. *Behaviour Research and Therapy, 34,* (5–6), 453–458.

Salkovskis, P. M., Clark, D. M., & Hackmann, A. (1991). Treatment of panic attacks using cognitive therapy without exposure to feared situations or bodily sensations. *Behaviour Research and Therapy, 29,* 161–166.

Salkovskis, P. M., Clark, D. M., & Jones, D. R. O. (1986). A psychosomatic mechanism in anxiety attacks: the role of hyperventilation in social anxiety and cardiac neurosis. In H. Lacey & J. Sturgeon (Eds.), *Proceedings of the 15th European Conference on Psychosomatic Medicine*. London: Libby.

Salkovskis, P. M., & Hackmann, A. (1997). Agoraphobia. In G Davey (Ed.), *Phobias: A handbook*. Chichester, UK: Wiley.

Salkovskis, P. M., Jones, D. R. O., & Clark, D. M. (1986). Respiratory control in the treatment of panic attacks: replication and extension with concurrent measurement of behaviour and $p\mathrm{CO}_2$. *British Journal of Psychiatry, 148,* 526–532.

Salkovskis, P. M., & Warwick, H. M. C. (1986). Morbid preoccupations, health anxiety and reassurance: A cognitive-behavioral approach to hypochondriasis. *Be-*

haviour Research and Therapy, 24(5), 597–602.

Salkovskis, P. M., Warwick, H. M. C., Clark, D. M., & Wessels, D. J. (1986). A demonstration of acute hyperventilation during naturally occurring panic attacks. *Behaviour Research and Therapy, 24,* 91–94.

Sanderson, W. C., & Barlow, D. (1986). *A description of the sub-categories of the DSM-III revised category of panic disorder: Characteristics of 100 patients with different levels of avoidance*. Paper presented at the annual meeting of the Association for the Advancement of Behavior Therapy, Chicago, November.

Sanderson, W. C., Rapee, R. M., & Barlow, D. H. (1989). The influence of an illusion of control on panic attacks induced via inhalation of 5.5% carbon dioxide enriched air. *Archives of General Psychiatry, 46,* 157–164.

Seligman, M. E. P. (1988). Competing theories of panic. In S. J. Rachman & J. Maser (Eds.), *Panic: Psychological perspectives*. Hillsdale, NJ: Erlbaum.

Shafer, S. (1976). Aspects of phobic illness: A study of 90 personal cases. *British Journal of Medical Psychology, 49,* 221–236.

Sheehan, D. V. (1982). Panic attacks and phobias. *New England Journal of Medicine, 307,* 156–158.

Sheehan, D. V., Ballenger, J. C., & Jacobsen, G. (1980). Treatment of endogenous anxiety. *Archives of General Psychiatry, 37,* 51–59.

Skerrit, P. W. (1983). Anxiety and the heart—a historical review. *Psychological Medicine, 13,* 17–25.

Snaith, R. P. (1968). A clinical investigation of phobias. *British Journal of Psychiatry, 114,* 673–697.

Taylor, C. B., Fried, L., & Kenardy, J. (1990). The use of a real-time computer diary for data acquisition and processing. *Behaviour Research and Therapy, 28,* 93–97.

Taylor, S., Woody, S., Koch, W. J., McLean, P. D., et al. (1996). Suffocation false alarms and efficacy of cognitive behavioral therapy for panic disorder. *Behavior Therapy, 27*(1), 115–126.

Tearnan, B. H., Telch, M. J., & Keefe, P. (1984). Etiology and onset of agoraphobia: a critical review. *Comprehensive Psychiatry, 25,* 51–62.

Teasdale, J. D. (1988). Cognitive models and treatments for panic: a critical evaluation. In S. J. Rachman & J. Maser (Eds.), *Panic: Psychological perspectives*. Hillsdale, NJ: Erlbaum.

Telch, M. J., Brouillard, M., Telch, C. F., Agras, W. S., & Taylor, C. B. (1989). Role of cognitive appraisal in panic-related avoidance. *Behavior Research and Therapy, 27,* 373–383.

Telch, M. J., Tearnan, B. H., & Taylor, C. B. (1983). Antidepressant medication in the treatment of agoraphobia: a critical review. *Behaviour Research and Therapy, 21,* 505–527.

Thyer, B. (1986). Agoraphobia: A superstitious conditioning perspective. *Psychological Reports, 58,* 95–100.

Thyer, B. A., & Himle, J. (1985). Temporal relationship between panic attack onset and phobic avoidance in agoraphobia. *Behaviour Research and Therapy, 23,* 607.

Thyer, B. A., Himle, J., Curtis, G. C., Cameron, O. G., & Neese, R. M. (1985). A comparison of panic disorder and agoraphobia with panic attacks. *Comprehensive Psychiatry, 26,* 208–214.

Thyer, B. A., Nesse, R. M., Cameron, O. G., & Curtis, G. C. (1985). Agoraphobia: A test of the separation anxiety hypothesis. *Behaviour Research and Therapy, 23,* 75–78.

Torgersen, S. (1979). The nature and origin of common phobic fears. *British Journal of Psychiatry, 134,* 343–351.

Torgerson, S. (1983). Genetic factors in anxiety disorders. *Archives of General Psychiatry, 40,* 1085–1089.

Uhde, T. W., Siever, L., & Post, R. M. (1984). Clonidine: Acute challenge and clinical trial paradigms for the investigation and treatment of anxiety disorders, affective illness and pain syndromes. In R. M. Post & J. C. Ballenger (Eds.), *Neurobiology of mood disorders*. Baltimore: Williams and Wilkins.

Uhde, T. W., & Tancer, M. E. (1989). Chemical models of panic: a review and critique. In P. Tyrer (Ed.), *Psychopharmacology of anxiety*. Oxford, UK: Oxford University Press.

van Dongen, P. A. (1981). The human locus coeruleus in neurology and psychiatry. *Progress in Neurobiology, 17,* 97–139.

Vitaliano, P., Katon, W., Russo, J., Maiuro, R., Anderson, K., & Jones, M. (1987). Coping as an index of illness behaviour in panic disorder. *The Journal of Nervous and Mental Disease, 175,* 78–83.

von Korff, M., & Eaton, W. W. (1989). Epidemiologic findings on panic. In R. Baker (Ed.), *Panic disorder: Theory research and therapy*. Chichester, UK: Wiley

von Korff, M., Eaton, W. W., & Keyl, P. M. (1985). The epidemiology of panic attacks and panic disorder: the results of three community surveys. *American Journal of Epidemiology, 122,* 970–981.

Warren, R., Zgourides, G., & Jones, A. (1989). The relationship between panic, phobic and anticipatory anxiety in agoraphobia. *Behaviour Research and Therapy, 27,* 181–189.

Warwick, H. M. C., & Salkovskis, P. M. (1990). Hypochondriasis. *Behaviour Research and Therapy, 28,* 105–118.

Waxler, E. B., Kimbiris, D., & Dreifus, L. S. (1971). The fate of women with normal coronary arteriograms and chest pain resembling angina pectoris. *American Journal of Cardiology, 28,* 25–32.

Weekes, C. (1977). *Simple effective treatment of agoraphobia*. London: Angus and Robertson

Weissman, M. M., Leaf, P. J., Blazer, D. G., Boyd, J. H., & Florio, L. (1986). The relationship between panic disorder and agoraphobia: An epidemiologic perspective. *Psychopharmacology Bulletin, 43,* 787–791.

Wells, A., Clark, D. M., Salkovskis, P., Ludgate, J. et al. (1995). Social phobia: The role of in-situation safety behaviors in maintaining anxiety and negative beliefs. *Behavior Therapy, 26,* 153–161.

Westphal, C. (1872). Die Agoraphobie: Eine neuropathische Erscheinung. *Archives of Psychiatric Nerv. Krankh, 3,* 138–171.

Wittchen, H.-U. (1986). Epidemiology of panic attacks and panic disorders. In I. Hand & H. U. Wittchen (Eds.), *Panic and phobias*. Berlin: Springer.

Wittchen, H.-U. (1988). Epidemiology of panic attacks and panic disorder. In I. Hand & H. U. Wittchen (Eds.), *Panic and phobias*. Berlin: Springer.

Woods, S. W., Charney, D. S., Loke, J., Goodman, W. K., Redmond, D. E., & Heninger, G. R. (1986). Carbon dioxide sensitivity in panic anxiety. *Archives of General Psychiatry, 43,* 900–910.

Zitrin, C. M. (1983). Differential treatment of phobias: use of imipramine for panic attacks. *Journal of Behavioral and Therapeutic Experimental Psychiatry, 14,* 11–18.

Zitrin, C. M., Klein, D. F., & Woerner, M. G. (1980). Treatment of agoraphobia with group exposure *in vivo* and imipramine. *Archives of General Psychiatry, 37,* 63–72.

Zitrin, C. M., Klein, D. F., Woerner, M. G., & Ross, D. C. (1983). Treatment of phobias. I. Comparison of imipramine hydrochloride and placebo. *Archives of General Psychiatry, 40,* 125–133.

Zitrin, C. M., Woerner, M. G., & Klein, D. F. (1981). Differentiation of panic anxiety from anticipatory anxiety and avoidance behaviour. In D. F. Klein & J. Rabkin (Eds.), *Anxiety: New research and changing concepts*. New York: Raven Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.19

Worry and Generalized Anxiety Disorder

THOMAS D. BORKOVEC and MICHELLE G. NEWMAN
Pennsylvania State University, University Park, PA, USA

6.19.1 NORMAL AND PATHOLOGICAL WORRY 439
6.19.2 DIAGNOSTIC DESCRIPTION OF GENERALIZED ANXIETY DISORDER 440
6.19.3 IMPORTANCE OF GAD 440
6.19.4 BASIC RESEARCH ON THE NATURE, FUNCTIONS, AND ORIGINS OF WORRY AND GAD 441
  6.19.4.1 GAD and Information Processing 441
    6.19.4.1.1 Worrisome thinking and emotional processing 441
    6.19.4.1.2 Worrisome thinking and other information processing functions 442
  6.19.4.2 The Psychophysiology of GAD 443
  6.19.4.3 GAD and Interpersonal Factors 444
    6.19.4.3.1 GAD and early childhood interpersonal factors 445
    6.19.4.3.2 GAD and adult interpersonal factors 445
6.19.5 THERAPY OUTCOME INVESTIGATIONS OF GAD 445
  6.19.5.1 Review of Past Controlled Therapy Outcome Studies of GAD 445
6.19.6 CLINICAL DESCRIPTION OF COGNITIVE-BEHAVIORAL THERAPY FOR GAD 447
  6.19.6.1 Self-monitoring and Early Cue Detection 447
  6.19.6.2 Stimulus Control Methods 448
  6.19.6.3 Relaxation Methods 449
  6.19.6.4 Applied Relaxation 451
  6.19.6.5 Self-control Desensitization 452
  6.19.6.6 Cognitive Therapy 453
6.19.7 SUMMARY AND IMPLICATIONS FOR FUTURE THERAPY DEVELOPMENT 456
6.19.8 REFERENCES 457

6.19.1 NORMAL AND PATHOLOGICAL WORRY

Research on worry began in the test anxiety area in the 1970s, where the distinction was made between worry as a cognitive aspect of anxiety and emotionality reflecting the physiological features of anxious experience. An early definition of worry suggested that it involved "a chain of thoughts and images, negatively affect laden and relatively uncontrollable. The worry process represents an attempt to engage in mental problem solving on an issue whose outcome is uncertain but contains the possibility of one or more negative outcomes. Consequently, worry relates closely to fear process" (Borkovec, Robinson, Pruzinsky, & DePree, 1983a). So the early theoretical view of worry was that it was a cognitive avoidance response to detected threat (Borkovec, Metzger, & Pruzinsky, 1986). Several studies indicated that worry was indeed functionally separate from somatic anxiety and that it was primarily responsible for the interfering effects of anxiety on test and grade performance (Deffenbacher, 1980). During the 1980s, experimental research devoted to exploration of the nature of the worry process began. Inductions of worry were found to produce several important effects: incubation of negative thought intrusions (Borkovec et al., 1986; Pruzinsky & Borkovec, 1990; York, Borkovec, Vasey, & Stern, 1987), lengthening of decision-making times (Metzger, Miller, Cohen, Sofka, & Borkovec, 1990) due to elevated evidence requirements (Tallis, Eysenck, & Mathews, 1991), absence of change in heart rate (Borkovec et al., 1983a; York et al., 1987), elimination of the incubating effects of unconditional stimulus rehearsal in the unconditioned stimulus revaluation phenomenon (Davey & Matchett, 1994), and the generation of predominantly anxious but also depressive affect (Andrews & Borkovec, 1988). Remarkably, nonanxious groups displayed the very same effects during or after brief elicitation of the worry process. Related research indicated that anxious thoughts were harder to dismiss than neutral or depressing thoughts or intrusive negative images. Such research also showed that both the degree of worry about anxious thoughts and the presence of depressed states could further decrease control over anxious thoughts (Clark, 1986; Clark & DeSilva, 1985; Edwards & Dickerson, 1987a, 1987b; Sutherland, Newman, & Rachman, 1982). What generally emerged from this literature was that normal negative intrusive thoughts (including those of worry) were distinguishable from the pathological worry found in anxiety disorders by neither their process nor effects but rather by their frequency and intensity, which in turn appeared to be due to their uncontrollability (Clark, 1986; Clark & DeSilva, 1985; England & Dickerson, 1988; Kent & Jambunathan, 1989; Parkinson & Rachman, 1981; Salkovskis & Harrision, 1984). Despite growing interest in the phenomenon of worry during this period, the greatest boost to its investigation occurred when worry became the central defining feature of generalized anxiety disorder.

6.19.2 DIAGNOSTIC DESCRIPTION OF GENERALIZED ANXIETY DISORDER

The diagnostic description of generalized anxiety disorder (GAD) has undergone several changes in diagnostic definition. The label first occurred in the third edition of the *Diagnostic and statistical manual of mental disorders* (*DSM*, American Psychiatric Association, 1980); the older category of anxiety neurosis was divided into panic disorder and GAD. The definition of GAD focused on chronic, diffuse anxiety and its symptoms of apprehensive expectation (worry), vigilance/scanning, motor tension, and autonomic hyperactivity. If other Axis I conditions were present, however, GAD was not to be diagnosed. The arrival of *DSM-III-Revised* (American Psychiatric Association, 1987) was an important turning point. GAD was allowed as a principal diagnosis even in the presence of most other disorders, and it was centrally defined as apprehensive expectation (worry). Its remaining symptoms were organized into an 18-item checklist and designated as associated characteristics.

Further revision occurred in *DSM-IV* (American Psychiatric Association, 1994) based on comprehensive reviews of the extant empirical literature on GAD (Borkovec, Shadick, & Hopkins, 1991). The disorder is now defined by excessive and uncontrollable worry about a number of events or activities not confined to other Axis I disorders, occurring more days than not for at least six months, causing clinically significant distress or impaired functioning, and not due to the direct physiological effects of a substance or a general medical condition. The associated symptoms checklist was reduced in size, based on empirical identification of the most frequent and reliable symptoms (Marten et al., 1993). Diagnosis requires the presence of only three of six symptom groups (restless/keyed-up/on-edge, fatigue, difficulty concentrating, irritability, muscle tension, and sleep disturbance). Interestingly, all six symptoms reflect central nervous system (CNS) rather than autonomic nervous system (ANS) activity, matching what has been discovered in the psychophysiology of GAD. Also of importance for the scientific generalizability and continuity of GAD research, clients who meet GAD criteria by *DSM-III-R* routinely meet *DSM-IV* criteria as well (Abel & Borkovec, 1995).

6.19.3 IMPORTANCE OF GAD

GAD may be a particularly important area of experimental outcome and basic research. In terms of psychopathology: (i) estimates in the 1990s (Kendler, Neale, Kessler, Health, & Eaves, 1992; Wittchen, Zhao, Kessler, & Eaton, 1994) suggest 3.6–5.1% lifetime and 3.1% one-year prevalence rates, and it is one of the most common comorbid conditions in the anxiety and mood disorders (Brown & Barlow, 1992); (ii) it may be the basic anxiety disorder (based on its early onset, chronicity, and resistance to change) out of which other anxiety disorders often emerge (Brown, Barlow, & Liebowitz, 1994); (iii) worry, the central feature of GAD, is pervasive in the anxiety and mood disorders (Barlow, 1988); (iv) worry has been experimentally demonstrated to prevent emotional processing and thus may maintain any disorder wherein such processing is important for therapeutic change (Borkovec, 1994); and (v) successful treatment of GAD results in dramatic reduction in comorbid conditions (Borkovec, Abel, & Newman, 1995). So understanding the mechanisms of GAD and developing effective therapies based on that understanding may not only contribute to the amelioration of a very common disorder which is difficult to treat but may also have significant implications for the treatment and/or prevention of other disorders. Investigation of GAD is also important in the more general pursuit of understanding human nature and its adaptive and maladaptive processes. Research on GAD and worry has implicated the mutually interactive influences of several response systems (e.g., attention, interpretation of the environment, implicit and explicit memory, abstract thought, imagery, affect, central and peripheral physiology, and behavior) in their processes and maintenance. Thus, the study of GAD may yield knowledge in general, about the nonlinear dynamical nature of human functioning, viewed as a constantly changing process in response to a changing environment. Finally, recent basic research on GAD, especially on the role of developmental and interpersonal factors, holds considerable promise for the development of more effective therapy.

6.19.4 BASIC RESEARCH ON THE NATURE, FUNCTIONS, AND ORIGINS OF WORRY AND GAD

Clinical experience and empirical results suggest a straightforward description of the inner life of GAD clients: it is necessary to search constantly for cues about possible future threat ("What if ...?") in order to avoid catastrophes or to prepare ways to cope with their occurrence (Beck & Emery, 1985). Because the dangers exist only in their minds and only in the future, behavioral avoidance is not an available response. This leaves conceptual activity (i.e., worry) as one of the few coping resources. Human thought may well have evolved specifically to anticipate the future, to avoid negative events, and to increase positive events (McGuire & McGuire, 1991). The central questions for GAD research, however, focus on why so much threat is perceived, what the specific nature, functions, and consequences of the pathological worry response to threat might be, what factors are involved in its origins and maintenance, and how it is best treated. Results suggest that the answers will likely be found in dysfunctional processes within and interactively between three domains: information processing, physiology, and interpersonal functioning.

6.19.4.1 GAD and Information Processing

Evidence indicates that GAD and its central worry feature are associated with impairments in several types of information processing, the most important of which is the processing of emotional material.

6.19.4.1.1 Worrisome thinking and emotional processing

When people worry, they are mostly talking to themselves. Worry involves a prevalence of abstract, verbal-linguistic, thought as opposed to imaginal process (Borkovec, 1994). This is true for GAD clients and nonanxious controls when they are worrying. During relaxation states, GAD clients continue to experience the predominance of thoughts, whereas nonanxious persons shift to a predominance of imagery. Therapy for GAD clients leads to the normalization of thought and image frequencies (Borkovec & Inz, 1990). The extensive findings on the role of such cognitive factors and many others in GAD (Mathews; 1990; Mathews & MacLeod, 1994) has long suggested the usefulness of cognitive therapy (CT) for the disorder.

The role of such cognitive factors is also fundamental to a specific theoretical view of the nature and functions of worry. The distinction between thought and imagery is crucial. Imagery is closely tied to efferent commands into the affective, physiological, and behavioral systems; thought is not. Imagination of fear material elicits strong cardiovascular response; verbally articulating the same material does not (Vrana, Cuthbert, & Lang, 1986). It is adaptive that thought evolved in this way. Its relative isolation from efferent systems increases choice and flexibility.

Worrisome thought, however, has the disadvantage that it precludes emotional processing. Phobic persons who think worrisome thoughts just prior to repeated phobic image presentations show no cardiovascular response at all to those images; those who think neutral or relaxing thoughts display significant response, with relaxation producing habituation over repeated images (Borkovec & Hu, 1990; Borkovec, Lyonfields, Wiser, & Diehl, 1993a). Moreover, worry after exposure to a stressful event also results in a failure to process the emotional material adaptively (Butler, Wells, &

Dewick, 1995; Wells & Papageorgiou, 1995). From a neobehavioristic perspective, absence of cardiovascular response to phobic information reflects a failure to access the complete fear structure stored in memory (especially its efferent emotional meanings) and therefore an absence of the kind of emotional processing necessary for extinction (Foa & Kozak, 1986). One of the immediate functions of worry, then, is the avoidance of aversive images, somatic anxiety, and other negative emotions, and the worrying is thereby strengthened via negative reinforcement. A significant consequence of this avoidance function is the maintenance of anxious meanings despite daily exposures to threat-related material and the potentially corrective information that accompanies such exposures. These findings and their theoretical context have long suggested the usefulness of imagery exposure (combined with relaxation) for accessing and changing emotional meaning in therapy for GAD.

The above research supports the hypothesis that worry functions as a cognitive avoidance response to threatening material. Analogue GAD and control groups actually view worry in a similar way (Borkovec & Roemer, 1995). When these groups rated possible reasons for worry, only one scale discriminated the two groups: "Worrying about most of the things I worry about is a way to distract myself from worrying about even more emotional things, things that I don't want to think about." Although this could merely represent attempts of the GAD participants to rationalize why they worry, it may also reflect a further avoidance function of worry: avoidance of more painful emotional material. Rather than focus on primary affect, the superficial content of worry may be focused on secondary or instrumental affects.

What underlying content might be avoided is speculative, but three candidates exist, and all relate to interpersonal issues. Two (negative attachment experiences and current interpersonal problems) are discussed later. The third possibility resides in the fact that both GAD clients and analogue GAD groups report more frequent past traumas than controls, and 65% of these are interpersonal in nature (Roemer, Borkovec, Posa, & Lyonfields, 1991). Of course, when considering this evidence, we must be mindful that recall is subject to distortion, the traumas could have occurred after the onset of GAD, and many psychiatric groups report frequent trauma. However, content analyses of the trauma topics and worry topics reveal indirect evidence for avoidance of trauma memories (Shadick, Roemer, Hopkins, & Borkovec, 1991). Even though the traumatic events involved death, illness, or injury related to self or others (as required by *DSM* trauma definition), this is the very category about which GAD clients worried the least. The avoidance of trauma content can in and of itself lead to the maintenance of anxious meanings, but a traumatic experience may also provide actual evidence that the world (especially involving other people) can indeed be a threatening place and further contribute to the origins and/or maintenance of GAD.

The implication of worry's function as cognitive avoidance and its prevention of emotional processing is far-reaching. Worry contains a process that maintains GAD or indeed any emotional disorder. Emotional change due to daily life experiences or to deliberate exposures in therapy to emotional material is less likely to take place to the degree that worry is present before, during, or after such exposures.

The mechanisms for worry's inhibitory effect on emotional processing are unsubstantiated, but four possibilities have been suggested (Borkovec, 1994): (i) lessened attentional resource for processing other information (Mathews, 1990); (ii) difficulty of shifting attention from excessive, habitual (especially negatively valanced) thought activity (Parkinson & Rachman, 1981); (iii) semantic satiation effects of repetitious worrisome thought, thus insulating worry content from the rest of its associative network, especially affective associations (Smith, 1984); and (iv) less mismatch between information expected and received (a central condition for anxiety in Gray's neuropsychological theory of anxiety (Gray, 1982)). Future basic research might aim usefully at isolating the specific mechanisms involved.

6.19.4.1.2 Worrisome thinking and other information processing functions

Research evidence indicates that GAD and worry are associated with various information processing characteristics that interfere with learning from experience and thereby maintain anxious meaning. GAD clients display (i) preattentive bias to threat cues, even outside of awareness and often with task-interfering effects, (ii) rapid cognitive avoidance of detected threats such that explicit memory for the material is reduced but implicit memory is increased, and (iii) negative interpretations of and predictions from ambiguous and even neutral information. Several of these effects are accentuated for stimulus material directly associated with the client's worrisome concerns (Mathews, 1990; Mathews & MacLeod, 1994).

Three further instances of worry effects on information processing have emerged. First, in the current therapy trial at the Pennsylvania State GAD Project, clients monitor their daily worry predictions and rate actual outcomes once they occur. In this study, the outcomes turned out better than expected 84% of the time, and in 78% of the remaining situations, clients still coped better than they predicted. Thus, in only 3% of all worries did the core feared event ("The predicted bad event will occur, and I won't be able to cope with it") actually happen. It is possible that GAD clients continue to worry despite consistent real world evidence that feared outcomes do not occur because they frequently fail to process this corrective evidence. Second, people who worry about a topic prior to generating alternative predictions of what might happen generate exclusively negative predictions, whereas people who relax before generating the alternatives list only positive predictions (Borkovec, 1995). This suggests that worrying increases cognitive rigidity and that this rigidity precludes the retrieval of realistic corrective information. Third, several studies have demonstrated poor recall (reflecting cognitive avoidance) of threat words in GAD (cf. Mathews, 1990). However, a study by Thayer and Borkovec (1995) attempted to replicate these findings using a higher-order classical conditioning task (i.e., emotional words as the unconditional stimuli). In this study, GAD clients were exposed to 10 threat words (preceded by a colored block, the conditional stimulus) and 10 nonthreat words (preceded by a differently colored block) in random order. Clients merely read the words silently as they appeared, but they were given a surprise recall task at the end. Contrary to results of prior research, clients recalled more threat than nonthreat words and to a greater degree than nonanxious participants. Whereas previous research was based on brief exposures and/or instructions to actively respond to the words, Thayer and Borkovec used a longer exposure time. This suggests that lengthier exposures during a passive task will enhance threat recall from explicit memory, perhaps by overriding the defensive process of cognitive avoidance.

Taken as a whole, the above research indicates that GAD is characterized by habitual, inflexible cognitive and affective processes. A parallel rigidity has also been found in the tonic physiology of GAD and the phasic physiology of worry.

6.19.4.2 The Psychophysiology of GAD

Unlike the results from other anxiety disorders, physiological activation is not always characteristic of GAD, and inductions of worry do not produce differential activation compared to relaxed states (Borkovec et al., 1983b; Elliott, 1990). One study did find greater CNS-mediated muscle tension in GAD but no other activation differences compared to controls (Hoehn-Saric, McLeod, & Zimmerli, 1989). However, GAD was characterized by reduced variability in autonomic measures, leading to the hypothesis that GAD involves sympathetic inhibition and an autonomic inflexibility rather than sympathetic activation (Hoehn-Saric & McLeod, 1988).

Rigid autonomic functioning is implicated in several biological disorders (Thayer, Friedman, & Borkovec, 1996). Models of nonlinear dynamical systems view flexible responding as a marker of system integrity, and these models interpret reduced variability to be pathological (Goldberger, 1992). The hypothesis that GAD is characterized by a distinctive physiology involving restricted range of variability thus takes on added importance. However, the speculation that this is due to sympathetic inhibition may be an oversimplification of complex autonomic activity which involves both sympathetic and parasympathetic branches.

Evidence supports the specific importance of chronic parasympathetic (vagal) deficiency in GAD. Analogue GAD participants display tonically lowered vagal tone and little variation in either vagal tone or heart rate during rest, aversive imagery, and worrisome thinking (Lyonfields, Borkovec, & Thayer, 1995). Nonanxious participants show decreases in vagal tone from rest to emotional tasks and the greatest reduction during worry. Thus, GAD is characterized by autonomic inflexibility, such rigidity is due to a chronic deficiency in parasympathetic tone, and worrisome thinking can cause such reductions. Replicated findings emerged from a comparison of GAD clients to nonanxious participants during periods of self-relaxation and worry (Thayer et al., 1996). GAD was characterized by a tonic deficiency in vagal tone, and a main effect of tasks indicated that worry phasically caused that deficiency in GAD and nonanxious participants. These results make dramatic sense if we combine what is known empirically about GAD (e.g., chronic vigilance to threat, frequent aversive cognitions with a predominance of worrisome thought, excessive muscle tension, and difficulty concentrating) with results from experimental physiology: tasks that phasically cause reduced vagal tone among normals include threat of shock, recall of past aversive events, mental problem-solving, and isometric hand-grip (Grossman, Stemmler, & Meinhardt, 1990; Grossman & Svebak, 1987). Moreover, vagal deficiency provides a physiological basis for the habitual, GAD-characteristic attentional processes mentioned previously, given that poor attentional concentration and distractibility are related to lower vagal tone in both infants (Richards, 1987) and adults (Porges, 1992). The above research demonstrating vagal deficiency and autonomic inflexibility in GAD reinforces the long-held view of the usefulness of relaxation methods in the treatment of GAD. These results also provide a clear physiological basis for the finding that CNS but not ANS symptoms characterize the self-reports of GAD clients (Marten et al., 1993).

This GAD-distinctive inflexibility also makes sense when considering the psychological situation that GAD clients face. Their threats are illusory, exist only in their minds, refer to a nonexistent future, and are based on fears of events having a low probability of occurrence. The clients thus face an insoluble problem wherein fight-or-flight activation is inhibited because there is no one to fight and nowhere to flee, analogous to the freezing response of rats when motoric avoidance to feared stimuli is prevented.

Electroencephalogram (EEG) differences have also been discovered between nonanxious groups and both GAD clients and chronic worriers (e.g., greater frontal beta and less parietal alpha for the latter) as well as between relaxation and induced worry tasks (e.g., greater frontal and parietal beta during worry) (Carter, Johnson, & Borkovec, 1986; Inz, 1990). Importantly, inflexibility was also revealed in the cortical activity of GAD clients; lowered variability was found in alpha and beta over several laboratory tasks (Inz, 1990).

Although restricted, tonic cardiovascular functioning appears to be characteristic of GAD, a close look at phasic heart rate reactions to threatening material reveals that brief initial activation can occur. Results from the threat/nonthreat word conditioning task described earlier have shown that: (i) nonanxious participants showed larger orienting responses to threat and nonthreat words than GAD clients; (ii) their magnitude was related to resting vagal tone (described earlier to be deficient in GAD and intimately related to attentional mechanisms); (iii) the control group, but not GAD group, showed habituation of physiological response over both threat and nonthreat trials; (iv) GAD clients, but not the control participants, displayed heart rate acceleration (the classic "defensive response") to threat words; and (v) the GAD group, but not the control group, developed an anticipatory heart rate deceleration to the colored block preceding threat words but not to the block preceding nonthreat words (Thayer & Borkovec, 1995). The latter result indicates that classically conditioned orienting responses (indicating increased attention) emerge in GAD over repeated exposures to verbal threat material, and provides a basis both for GAD hypervigilence and for the generalization of threatening meaning via associative learning processes which depend solely on words as the unconditional stimulus. These results may have significant implications for understanding the interrelationships between the processing of verbal material, attention, associative emotional learning, and physiology in GAD.

Thus far, we have reviewed information on the intrapsychic patterns of GAD showing inflexible and repetitious cognition, affect, and physiology. Several lines of research have also explored the possibility of habitual patterns in overt behavior, particularly those in the domain of interpersonal functioning. Such research has found evidence for recurrent, problematic, interpersonal behavior as well as for possible connections between adult interpersonal styles and early attachment experiences.

6.19.4.3 GAD and Interpersonal Factors

For many years, researchers have known of a significant connection between GAD and interpersonal factors. Several lines of evidence have contributed to this knowledge. First, patterns of familial history in GAD (Noyes, Clarkson, Crowe, Yates, & McChesney, 1987; Noyes et al., 1992) along with a low degree of heritability (Kendler et al., 1992; Torgersen, 1986) imply that interpersonal developmental experiences may provide some etiologic foundation for GAD. Second, although GAD clients worry about many topics, they are particularly afraid of interpersonal situations: social phobia is the most frequent comorbid diagnosis for GAD (Barlow, 1988; Borkovec et al., 1995); trait-worry correlates more highly with social fears than with nonsocial fears (Borkovec et al., 1983b); content analysis of GAD worry topics reveals interpersonal concerns to be more frequent than any other domain (Shadick et al., 1991). Finally, worry involves a predominance of thought; thought is based upon language; language is based on social communication. Thus, some features of thought (talking to ourselves) are inherently interpersonal (Borkovec et al., 1991). Together, these observations suggest that a history of negative interpersonal experiences may contribute to the origins or maintenance of GAD, to current interpersonal functioning, to the maintenance of an excessive perception of threat, and/or to learned worrisome thought styles to cope with threat.

6.19.4.3.1 GAD and early childhood interpersonal factors

Bowlby (1982) was quite specific about hypothesized relationships between attachment and anxiety. If a child has a primary attachment figure who is repeatedly unavailable, the resulting insecure base may manifest itself in the child's developing mental models of the world as a dangerous, frightening place, with an overestimation of the probability and severity of feared events and an underestimation of coping resources. These mental representations are strikingly similar to those present in adult GAD. Indeed, diffuse anxiety was felt by Bowlby to be the typical consequence of some forms of insecure attachment.

The Inventory of Adult Attachment (IAA, Lichtenstein & Cassidy, 1991) is based on Main and Goldwyn's (in press) system for scoring their Adult Attachment Interview (George, Kaplan, & Main, 1985) and on Bowlby's attachment theory. Results using the IAA indicate that the GAD groups score higher than nonanxious groups on role-reversed/enmeshed relationships but not on childhood rejection (Borkovec, 1995). The child had to take care of the parent rather than the parent taking care of the child. This suggests an understandable basis for the adult GAD view of the world as a potentially dangerous place, of oneself as unable to cope, and of the need to constantly anticipate threats as the necessary means to obtain love, approval, and nurturance. The clients also reported greater unresolved feelings of anger and vulnerability toward their primary caregiver. Both the enmeshment and anger/vulnerable findings are discussed more fully below in relation to findings on adult interpersonal styles. Finally, the GAD group reported greater difficulty remembering their childhood. Whether the latter reflects motivated cognitive avoidance of negative memories or the fact that GAD clients have a harder time remembering anything due to the sheer amount of conceptual activity in which they are constantly engaging is not yet known.

6.19.4.3.2 GAD and adult interpersonal factors

Clinicians often discover that many of the concerns raised by their GAD clients relate to interpersonal relationships. Research at the Pennsylvania State Project (Pincus & Borkovec, 1994) has made use of the Inventory of Interpersonal Problems Circumplex Scales (IIP-C, Alden, Wiggins, & Pincus, 1990) to assess the interpersonal functioning of GAD clients in a psychometrically sound way. Results indicate that GAD clients show not only greater interpersonal distress (not surprisingly) but also a greater rigidity in their interpersonal styles across differing situations than nonanxious people (Pincus & Borkovec, 1994). Thus, GAD has now been shown to reflect inflexibility in cognitive, affective, central physiological, peripheral physiological, and interpersonal behavioral domains. More importantly, cluster analysis has revealed three subtypes of GAD, replicated in both initial and cross-validation samples. The first of the two smaller clusters (24.3% of the clients) involved social avoidance and nonassertiveness, whereas the second cluster (13.5%) was characterized by elevated domineering and vindictive scales. The primary interpersonal problems for GAD (62.1% of the clients) were found on the overly nurturant and intrusive scales. Relating this to the attachment results, the majority of GAD clients may have learned in childhood to care for others in order to get love, approval, and nurturance, and they may continue to repeat this pattern in adulthood for the same reason. Worry may thus function, as it did in childhood, as a means of anticipating the needs of, and threats to, significant others in the pursuit of satisfying interpersonal needs. The discovery of this majority cluster, along with the attachment results, also fits well with research (Peasley, Molina, & Borkovec, 1994) showing that GAD groups report higher degrees of empathy on the Interpersonal Reactivity Index (Davis, 1980). This empathy was most prevalent in the realm of feeling other people's pain. Thus, attempts by GAD clients in the overly nurturant cluster to take care of others may also be motivated by attempts to reduce one's own pain caused by intensely feeling the pain of others.

6.19.5 THERAPY OUTCOME INVESTIGATIONS OF GAD

Although the findings on the possible role of interpersonal factors in GAD suggest that such factors may become a very important feature in the future development of more effective therapies, all controlled therapy outcome investigations to date have involved evaluations of behavioral and cognitive-behavioral therapy (CBT) treatments that target intrapersonal anxiety process.

6.19.5.1 Review of Past Controlled Therapy Outcome Studies of GAD

Systematic development and evaluation of psychotherapies for GAD began only recently for two reasons. First, GAD has historically been an ambiguous disorder with frequent changes in diagnostic definition. Second, the etiology of GAD was not well explained by the discrete conditioning events that have been used to elucidate the origins of other anxiety disorders. Moreover, because conditioned environmental triggers for anxiety were not obvious, and internal cues were more functionally relevant, the utility of exposure methods was less clear for the diffuse forms of anxiety found in GAD. Thus, early treatments for GAD emphasized the somatic aspects of the client's anxiety and trained clients to use relaxation techniques as a general coping response whenever they felt anxious. Later developments involved the addition of elements from behavioral self-control and cognitive treatments. These elements included an emphasis on thorough applied relaxation training (Öst, 1987), rehearsal of relaxation and cognitive coping responses during imaginal exposures to environmental and internal fear cues (e.g., self-control desensitization (Goldfried, 1971), anxiety management training (Suinn & Richardson, 1971), and CT aimed at modifying the GAD client's constant perception of threat (Beck & Emery, 1985). The emerging view of GAD, based on growing evidence concerning its nature and the specific functions of worry, now sees it as a loosely structured cognitive-affective state (Barlow, 1988) whose process is characterized by spiraling, habitual, inflexible interactions of multiple systems responding to constantly perceived threat (Borkovec & Inz, 1990). Thus, intrapersonally focused CBT was designed to provide multiple coping responses for the targeting of each of the maladaptive reactions and their interactions and to replace anxiety-maintaining spirals among these systems with alternative responses. Such a treatment should be able to help an individual to develop a flexible, adaptive lifestyle conducive to reduced anxious experience.

Only 12 controlled outcome studies on *DSM*-defined GAD have been published. Review of 11 of these studies indicated that indeed CBT produces both statistically and clinically significant change, with maintained or increased improvement up to a year later (Borkovec & Whisman, 1996). This therapy has also been associated with low drop-out rates, declining use of psychotropic medication, and the largest degrees of change among contrasted therapies and control conditions on both anxiety and depression measures. Fifteen earlier outcome studies of "general anxiety" are consistent with the GAD trials: anxiety management has produced long-lasting gains that were sometimes superior to component conditions, and CT techniques may add to post-therapy and especially follow-up improvement (Durham & Turvey, 1987; Lindsay, Gamsu, McLaughlin, Hood, & Espie, 1987). Finally, a twelfth GAD study found CBT to be superior to anxiety management at six-month follow-up and superior to brief psychodynamic therapy at post-therapy and follow-up (Durham et al., 1994). However, several methodological flaws render the conclusions questionable (brief duration of GAD, therapist-by-treatment confound, lower credibility for the psychodynamic condition, no therapy manuals, large drop-out rate, higher severity for the psychodynamic condition, and no therapy integrity checks).

Despite these promising effects, three areas of concern exist (Borkovec & Whisman, 1996). First, the role of nonspecific factors or individual CBT components remained unclear until the late 1990s. Although placebo and individual elements have yielded less change than CBT, between-group effects have not always emerged. Second, clinically significant change for CBT has been moderate at best (50% high endstate functioning on average across studies). An earlier review of nine of these studies came to similar conclusions (Chambless & Gillis, 1993). Finally, although their methodology has been fairly good (e.g., frequent use of manuals, diagnostic interviews, integrity checks, and expectancy/credibility assessments), significant deficiencies exist in the majority of the studies. Most importantly, only three studies required two independent diagnostic interviews. Principal GAD has the lowest interassessor kappa among anxiety disorders (Barlow & DiNardo, 1991). Absence of independent diagnosis raises the possibility of false-positive cases (as many as 25–30%). Increased error variance due to false-positive cases may explain why CBT has not always produced significantly greater improvement than nonspecific or component conditions and why prior basic research on GAD has often yielded ambiguous or conflicting results.

Somewhat greater clarity on the issue of nonspecific effects and the contribution of individual CBT elements was achieved in a comparison of applied relaxation, CBT (applied relaxation, self-control desensitization, and brief CT), and a reflective listening control condition (Borkovec & Costello, 1993). Unlike many prior studies, this study required two independent diagnostic interviews, as well as quality and adherence checks of the therapy provided. At post-therapy, applied relaxation and CBT were superior to reflective listening, despite the fact that reflective listening was superior to the other two conditions on depth of emotional processing as objectively measured by the Experiencing Scale (Klein, Mathieu-Coughlan, & Kiesler, 1986). However, at 12-month follow-up, clients in the reflective listening condition had lost most of their gains and had more frequent subsequent therapy, clients in applied relaxation had maintained their gains, and those in the CBT condition showed further gains and had higher endstate functioning (50% of the clients) compared to the other two conditions. These findings support the view that a cognitive behavioral treatment for a multidimensional disorder like GAD will be most effective if it targets each (cognitive, physiological, and imaginal) system found in basic research to play a significant role in GAD.

The fact that reflective listening produced greater levels of emotional processing but failed to achieve a better outcome in the above trial suggests that the type of emotional processing that occurred was not therapeutically relevant to this disorder. From an experiential perspective, the primary affects most centrally involved in emotional disturbance are not easily accessed in GAD when reflective listening alone is provided; the avoidant function of worry may yield a client focus on secondary or instrumental affect (Greenberg & Safran, 1987). For emotional processing to occur, more potent techniques for accessing and deeply experiencing underlying, primary emotions in the present moment may be necessary. But the development of such a therapeutic focus, either for exposure treatment or for experiential techniques, must await future research which identifies what those primary fears might be in GAD.

The current therapy study in the Pennsylvania State Project compares applied relaxation/self-control desensitization, CT, and their combination. Therapy time was increased by 50% to allow for more complete CT than in the previous trial. Preliminary results based on a partial sample have indicated that one-third of the clients in each of the three conditions reached high endstate functioning by the end of therapy. At both follow-ups, however, the combined CBT group had increased to a 50% rate, whereas the two single-element conditions lost endstate gains. Should these outcomes hold up once all clients have completed therapy, they will support the already existing trend in the outcome literature suggesting that long-term maintenance is indeed best produced by a combined CBT therapy that targets each of the intrapersonal anxiety response systems. However, it is also clear that the attempt to increase the efficacy of CBT by doubling therapy time was not successful. CBT outcome at follow-up was not significantly different from previous trials. This strongly suggests the wisdom of looking elsewhere to develop more powerful therapeutic approaches for GAD. Further preliminary results from this ongoing trial strongly suggest that this may well involve the therapeutic targeting of interpersonal problems (Borkovec, 1995). Three of the IIP-C scales that significantly contributed to the clustering results (domineering, overly nurturant, and intrusive) failed to change significantly after CBT. Most importantly, these three scales (plus the vindictive scale) negatively predicted outcome, and these relationships were stronger at follow-up than at post-test. Thus, interpersonal problems, if left unattended in therapy, may be especially associated with failure to maintain therapeutic gains obtained by intrapersonally oriented CBT.

6.19.6 CLINICAL DESCRIPTION OF COGNITIVE-BEHAVIORAL THERAPY FOR GAD

Despite limitations in CBT, both in degree of improvement and failure to address potentially important interpersonal issues, it is the only therapy method for GAD which has been empirically validated (Chambless et al., 1996). This section will provide a clinical description of the CBT elements which are most frequently used in controlled trials with the disorder. The protocol treatments in these trials have depended on such basic resources as Beck and Emery's (1985) text on CT with anxiety disorders and Bernstein and Borkovec's (1973) manual on progressive relaxation training. The principles and techniques of those approaches will not be reviewed here in detail. Instead, the present section will describe clinical adaptations of basic behavioral and cognitive therapy methods specifically for GAD, adaptations which have often been guided by the basic empirical research on the disorder summarized earlier in this chapter.

6.19.6.1 Self-monitoring and Early Cue Detection

Self-monitoring of anxiety cues on a moment-to-moment basis is foundational to CBT with GAD. Although the identification of environmental cues is important, greater emphasis is placed on determining relevant internal cues that trigger anxiety and worry. The latter include attention, thoughts, images, bodily sensations (especially muscle tension), emotions, and behaviors. Human experience is described to the client as an active process involving the interaction of these internal cues as they respond to a constantly changing environment. Anxiety is the result of an incremental process that develops as each internal cue is triggered by other cues in a combined effort to respond to perceived threat.

For example, a person has a thought, which triggers an emotion, and that emotion triggers another thought which triggers muscle tension, which leads to the spiraling up of anxiety. It is not our reactions to events that are the problem; the initial reaction is quite natural and understandable, given our past history of learning. The problem resides in our reactions to our reactions. These are the habits which we have developed and that are further strengthened each time we allow their entire sequence to occur. In therapy, the deliberate changing of reactions to our reactions and the strengthening of new coping responses in each relevant system are targeted. Emphasis is therefore placed on increasingly early detection of internal changes occurring in an anxious direction. The earlier we can identify incipient cues for anxiety, the earlier we can intervene. Any new intervention will be more effective in reducing anxiety if it is applied earlier in the spiral.

Asking clients to describe typical anxiety experiences in terms of what they attend to, do, think, imagine, and somatically and emotionally feel is a good starting point. However, the use of imagery of past or upcoming anxiety-provoking situations can help clients more accurately identify the internal contents and sequences of actual anxious experience. The use of a repeated image can help clients attend to each relevant system and to practice the early detection of relevant cues. This in-session exercise also provides a concrete rehearsal of the very self-monitoring that clients are asked to practice in daily life. Of particular importance is the actual worry process. Asking clients to worry silently about a typical topic of concern allows them to observe and describe this process in a more objective way. Asking them to subsequently worry out loud is often useful as well. This exercise gives the therapist some hints about what the nature and content of worry are for a specific client.

The therapist also observes clients for any detectable shift in their affective state and immediately asks them to describe what is occurring. In this way, the therapist serves the role of early detector of incipient anxious spirals and facilitates client recognition of early cues. Clients are encouraged to tell the therapist when they notice the beginning of anxiety cues. Such encouragement allows the therapist to fade out prompting as clients rehearse their own early detection, the very goal of self-monitoring in daily life.

Clients are encouraged to practice self-monitoring on a daily basis between sessions. Remembering to do so is the most difficult challenge, so the therapist provides a variety of suggestions for establishing reminders in the client's environment (e.g., checking every half-hour, at every change in activity, and in response to noticing post-it notes placed in noticeable locations). Over time, increasingly specific reminder cues which are particularly suited to the anxiety process of individual clients can be identified and used (e.g., whenever the phone rings, whenever a particular co-worker or family member is first seen during the day, whenever the client notices foot-tapping or hair-twisting behaviors). The important thing is that a sufficient number of effective cues are established to frequently remind clients to observe their inner processes. The ultimate goals are to establish a habit of recognizing incipient anxiety and to intervene at that moment with whatever coping responses have been learned up to that point.

Self-monitoring need not be focused only on anxiety cues. Both worry in general and GAD in particular are associated with other important emotional experiences (e.g., anger, frustration, depression) as well. Learning to detect the incipient cues for each of these emotional experiences and to distinguish them from anxiety will facilitate the application and mastery of interventions better suited to them and/or may lead to the identification of their functional role as the beginning point or consequence of an anxiety spiral.

6.19.6.2 Stimulus Control Methods

Because it occurs in numerous situations, the worry process is usually under poor discriminative control. However, once clients have learned to detect the occurrence of worry, the therapist can introduce a very simple stimulus control method for reducing the amount of time spent worrying. A degree of stimulus control can be re-established by providing clients with five instructions.

(i) Establish a 30-minute worry period to occur at the same time each day and in the same place (a place not associated with other daily activities).

(ii) Learn to detect the worrying at earlier and earlier moments, with the goal to ultimately identify its onset.

(iii) As soon as it is detected, postpone it to your worry period. Remember that you will have plenty of time to focus on your worry at that time and there is no need to make yourself miserable by worrying now when you have better things to do. Also keep in mind that short-circuiting worry will help reduce its habit strength.

(iv) After postponing the worry, focus your attention on the present environment or the task at hand.

(v) Use your worry period to worry intensely or (later in therapy) to employ some of the CT strategies to eliminate the source of the worrying.

Although empirical support for the efficacy of this technique is available on college samples of chronic worriers (Borkovec, Wilkinson, Folensbee, & Lerman, 1983c), a full clinical trial to evaluate its usefulness in GAD samples has not yet been conducted. Clinical experience suggests, however, that this simple method can often provide some beginning relief fairly quickly.

6.19.6.3 Relaxation Methods

Because autonomic inflexibility is characteristic of GAD physiological process, training in multiple relaxation methods for flexible application as well as to strengthen parasympathetic tone is desirable. Such relaxation methods can include paced diaphragmatic breathing, progressive muscle relaxation, guided imagery, meditation, and, most importantly, daily application of applied relaxation. Clients are encouraged to experiment with the different techniques in response to different response system reactions and to identity which ones are most helpful under which circumstances. With several coping responses available, the likelihood of finding an effective one in any given situation increases, and clients develop an increased sense of choice and flexibility as well.

When new cognitive or behavioral techniques are introduced in therapy, it is helpful to do demonstrations wherein the clients' own deployment of strategies produces an immediate, noticeable, and positive effect, however slight. In this way, clients obtain a sense that the technique can make a difference for them with further practice and application. Slowed, paced, diaphragmatic breathing produces a rapid relaxation response with little training and can be immediately used in daily living. The therapist models and has clients imitate both rapid, shallow, thoracic breathing (a sympathetic elicitor) and slowed diaphragm breathing (a parasympathetic elicitor). In this way, clients learn that the way that they breathe affects how they feel and that they can exert some immediate control on physiological and psychological states merely by controlling respiration. Typical instructions for diaphragmatic breathing would be: "Shift to breathing from your stomach rather than from your chest. Allow your diaphragm to rise and fall without expanding your chest. Also, slow your breathing down to a rate slower than usual but not so slow that it is unpleasant or uncomfortable. You might do this by counting from one to three as you breathe in evenly and then again as you evenly exhale." A few minutes of diaphragmatic breathing practice helps clients learn what relaxation feels like and how to elicit it quickly. They are encouraged to shift to this type of breathing whenever they notice thoracic breathing in order to develop diaphragm breathing as a habit. This coping response can also be used upon early cue detection during daily self-monitoring as an initial method for learning to control anxiety spirals. Moreover, when the therapist cues the client to become aware of increased anxiety during the session, the therapist can ask them to apply this relaxation response for a brief moment to reduce those cues, thus giving clients immediate opportunities for rehearsing applied relaxation.

Training in progressive muscle relaxation (PMR) involves systematically tensing and releasing various muscle groups throughout the body. This exercise causes a lengthening of muscle fibers which results directly in reduced muscular tension and indirectly in reduced sympathetic activity. Clients formally practice the technique twice a day to deepen their relaxation response and to strengthen their ability to generate it in daily applications. As therapy progresses, muscle groups are combined, and tension production is eventually replaced by relaxation-by-recall, wherein clients merely identify extant muscle tension and let go of that tension by recalling how those muscles felt when tension was produced and released.

The notion of "letting go" of muscular tension is inherent in PMR. Each time clients let go of tension in a particular muscle group, they become a little more relaxed. Once clients are accustomed to producing relaxation by concretely releasing muscular tension, the therapist can introduce the analogous notion of letting go of other processes involved in their anxiety spirals. For example, clients can also practice the gentle letting-go of distressing thoughts, worries, images, bodily sensations, and negative emotional experiences. They can be reminded that it is their reactions to their reactions that make a difference.

Meditational techniques are often useful because they specifically emphasize attention to a single focusing device. Strengthening one's ability to shift attention away from anxiety-provoking cognitions to one pleasant, internal stimulus is particularly helpful for GAD clients whose mental life is often consumed with attention to worrisome thoughts and catastrophic images. The therapist can ask clients to use diaphragmatic breathing for this purpose, or they can work with the client to identify a specific image or word which taps into previous senses of safety, comfort, security, love, and/or tranquillity. The word, "home," for example, for some clients may be associated with special, positive meaning. For other clients, interpersonal connections (e.g., an image of one's infant daughter) might serve this purpose. Clients are asked to practice meditational methods by focusing on the pleasant internal stimulus for a few moments at the end of formal PMR practice and to incorporate the use of this technique into their daily relaxation applications.

Guided imagery can also be introduced by selecting scenes which are associated with tranquillity and pleasantness. Scenes can be constructed such that they sequentially produce deeper states of relaxation and varying types of additional positive affect. Although guided imagery is not useful for the rapid elicitation of a relaxation response during early cue detection, it is a helpful device for deepening relaxation during formal practice sessions or at other moments in the day when a relaxation break is taken in order to reduce built-up anxiety and tension.

Many GAD clients are perfectionistic, with high performance standards for themselves and fear of the judgments of others. Consequently, attempting to achieve a completely relaxed state may, in itself, generate anxiety which may interfere with that very goal. The therapist can describe relaxation training in terms of "movement toward relaxation." Applied relaxation is merely a matter of shifting the direction of the internal process away from a spiral of tension and anxiety; the movement toward relaxation process, no matter how slight, is one of the most useful directions because of its direct incompatibility with anxiety. Frequency of application is more important than depth of relaxation. Other than movement toward relaxation, there is no end-state being sought either during training and practice or in application.

Perfectionism is only one example of a basis for relaxation-induced anxiety (RIA), a phenomenon common among diffusely anxious individuals. Other possible reasons include the potentially frightening nature of unfamiliar cognitive or physiological activity elicited by relaxation, fear of losing control, and increased awareness of feelings of anxiety or other emotions (Heide & Borkovec, 1984). If the very method for reducing anxiety is itself anxiety-provoking, the likelihood of a therapeutic effect in its application is minimal. Indeed, in the three published GAD therapy investigations which measured the occurrence of RIA during training, the degree of its occurrence negatively predicted outcome (Borkovec & Costello, 1993; Borkovec & Mathews, 1988; Borkovec et al., 1987). Fortunately, methods exist for circumventing this problem. Stressing the notion of movement toward relaxation is one. Another is to teach an alternate relaxation technique; clients who experience RIA with one type of relaxation tend not experience it with another (Heide & Borkovec, 1983). Moreover, when the emphasis is on frequent applications of brief relaxation responses to detected anxiety cues rather than on twice-a-day formal practice of deep relaxation, there is less likelihood that clients will experience RIA. Finally, treatment of RIA in more difficult cases is analogous to exposure treatment of any feared event: training can be systematically conducted in a graduated fashion, with exposures to increasingly deeper levels of relaxation over sessions. In our current therapy trial using these methods, the correlation between in-session RIA and therapeutic improvement is no longer significant.

One of the overarching goals of relaxation training and application with GAD clients is to increase the amount of time spent focusing on the present moment rather than on anxiety-provoking thoughts and images about the future or the past. Early sessions devoted to the identification of internal triggers to the anxiety spiral have already demonstrated to clients that what they are thinking and imagining contributes significantly to their anxiety; they are partly creating their realities. The next step is to discuss the fact that, when we are thinking about the future or the past, we are creating an illusion. These future and past events do not exist now. But we still react emotionally as if they are actually occurring. Images are particularly powerful elicitors of this effect; they can cause nearly the same degree of physiological reactivity which the occurrence of the real events they reflect would cause. So we produce much of our anxious experience merely by thinking worrisome thoughts and imagining catastrophic scenes. A second disadvantage exists when clients pay excessive attention to the future or the past. As long as they are not paying attention to the immediate environment, they are not processing further information from the actual world; there can be little learning and development in their knowledge and adaptive behavior. The therapeutic implications of this are clear: the more time clients spend focused on what is real and in front of them, the less time they spend creating negative emotional experience and the more time they spend processing adaptive information. Consequently, the therapist works with clients during sessions to practice focusing attention on what is actually present. Initially, with eyes closed, clients are asked merely to attend to sounds occurring in the environment, the tactile feelings of the chair, and the odors in the air. They are instructed to return their attention to these sensations whenever they notice their attention drifting to other thoughts or images and to sense them without making judgments or elaborating on their further meaning or associations (i.e., no reactions to reactions). The same exercise is then repeated with eyes open, and later as they walk around the office. They are finally encouraged to increasingly practice shifting to this kind of focus throughout the day and to let the therapist know in the next session what new sensations they have discovered across multiple situations.

As clients report that relaxation applications are beginning to have some noticeable effects, they are further encouraged to begin noticing what emerges from states of greater tranquillity. We know from decades of research that anxiety interferes with many adaptive processes and that relaxation facilitates many forms of performance. The inverted-U nature of arousal effects on performance can be described to clients (i.e., anxiety at moderate levels is optimal but can interfere with performance at very low and high levels), and they are asked to begin observing what adaptive behaviors become disinhibited or facilitated when they are creating greater states of relaxation for themselves. These may include more effective interpersonal behaviors, greater clarity or logic in thought, or better efficiency and productivity in work. They are thus invited to explore and discover on their own what benefits relaxation may yield. Behavioral experiments are often useful devices for demonstrating such effects. For example, clients can conduct an ABAB experiment at work, where "A" represents their usual state during work and "B" represents work done in a more relaxed state. Emotional state as well as productivity, accuracy, or efficiency can be dependent measures for them to assess the impact of relaxation on their work.

6.19.6.4 Applied Relaxation

Merely training clients in a relaxation technique and asking them to practice it twice a day are unlikely to yield significant benefits. It is surprising that before the mid-1980s, many therapy trials which included relaxation methods provided only these elements. Twice-a-day formal practicing is designed to provide momentary breaks to return oneself to a more tranquil state and to further build the skill of eliciting relaxation responses and of identifying what relaxation feels like so that the clients know what it is that they can usefully move toward. But the real potency of relaxation resides in its deployment on a moment-to-moment basis whenever incipient anxiety spirals are detected. Although clinical applications of relaxation had certainly been grounded more thoroughly in the importance of daily applications, it was not until Öst's article on applied relaxation (AR) that clinically thorough relaxation training was incorporated into therapy research (Öst, 1987). Evidence has now accrued to indicate that this technique by itself can be a potent coping strategy for several anxiety disorders.

The essential feature of AR is the frequent elicitation of relaxation in response to anxiety and tension triggers. With GAD clients, this would include worrying. The therapist has already laid the essential foundation of AR in first-session work with clients on early cue detection and the use of diaphragmatic breathing as a rapid coping response to detected cues. Rehearsals of AR take place in the session whenever the therapist suggests a brief application of relaxation in response to observed signs of increased anxiety. Differential relaxation training is also conducted to facilitate the generalization of relaxation skills to daily living. Clients are taught to relax themselves not merely when reclining in a comfortable chair but even when sitting in upright chairs, walking, and engaging in various common activities. The therapist tells clients that unneeded muscular tension comes in two forms: excessive tension in muscle groups required for an activity (e.g., neck muscles needed for holding the head erect) and any tension in muscles that are unnecessary for an activity. The goal of differential relaxation is to identify when either excessive or unnecessary tension exists and to let go of that tension while continuing to carry out the activity. In-session practice involves engaging in commonly performed activities while identifying and relaxing away excessive tension, and between-session practice is encouraged to adopt a more habitually relaxed lifestyle. Cue-controlled relaxation is often instituted wherein a particular word (e.g., "relax," "calm," or a word taken from meditational focusing devices) is paired with the release of muscle tension during PMR and then used as an initial response to initiate applied relaxation in daily coping.

Clients can also be encouraged to use the relaxation response as a method of opening up to experience. From a more tranquil center established via relaxation, they are asked to shift their attention more completely and frequently to the world around them, taking in the information which it provides and allowing stressful stimuli merely to "pass through" them without reacting to such events.

Finally, clients are told that when stressful events are upcoming, they can think in terms of three phases during which they can apply their relaxation coping responses: in anticipation of the event, during the event, and in recovery after the event. It is of course useful to deploy their coping strategies throughout these phases. But often, especially in the early stages of therapy, the most noticeable effect of relaxation applications occurs during recovery. It is helpful for clients to realize that more quickly recovering from an anxious experience is a desirable event for increasing their sense of self-control and as a starting point for more effectively reducing anxiety earlier and earlier in its sequence. As skill in applying relaxation to daily events increases, the therapist and clients can begin to create behavioral homework assignments that gradually increase their exposure to stressful, anxiety-provoking situations in which to apply the methods during these three phases.

6.19.6.5 Self-control Desensitization

Because worrying prevents emotional processing and imagery can be an effective method for engaging such processing, imagery exposure techniques could be very useful in the treatment of GAD. The problem is that we do not yet know what the core fears of GAD clients are. They are anxious about many situations, therefore either *in vivo* or imagery exposure to relevant fear-evoking stimuli would be a daunting, potentially excessive therapeutic task. Until the central fears and the images that effectively access those fears can be identified by future research, self-control desensitization (SCD) is a helpful intervention tool. SCD was originally designed to deal with more diffuse anxiety problems for which discrete, graduated hierarchies of circumscribed phobic stimuli were not possible (Goldfried, 1971). In SCD, clients first focus on relaxation process until they have achieved a relaxed state. Next, the therapist asks the clients to imagine themselves in a situation where internal and/or external anxiety cues occur. Clients signal as soon as actual anxious experience is elicited by the imagery, and the therapist then has them imagine coping effectively with the imagined situation by using the relaxation response. Clients signal again as soon as they no longer notice anxious responding, and they continue to imagine themselves in the original situation but effectively coping with it. Finally, images are terminated altogether, and clients return to focusing only on the relaxation process. The goal of SCD is to have the client rehearse relaxation applications to the kinds of situations that they often find problematic. Some degree of extinction process may take place, but the usefulness of the method probably resides more in the strengthening of adaptive coping responses to typical fear cues. Thus, clients are more likely to remember to employ coping responses during daily living given associative priming effects from previous imaginal rehearsals in therapy.

"Hierarchy" construction for SCD is less detailed than it is in systematic desensitization. The therapist helps clients identify both the internal cues of their anxiety spirals (somatic, cognitive (especially worry and catastrophic images), and behavioral) and representative types of external situations where anxiety and worry often occur. Images that include both internal and external cues are constructed for use during SCD and only roughly graded from mild to moderate to severe in anxiety-provoking value. The paragraph below exemplifies one possible method for conducting a formal SCD procedure with such images once the client is in a deeply relaxed state. It includes both therapist patter and procedural instructions.

> I will now begin to present the images which we discussed earlier. Remember to visualize yourself as if you are actually in the situation as realistically as you can. Also, be sure to raise your index finger immediately when you first notice the beginnings of any tension, anxiety, or shift away from the feelings of deep relaxation. Once you have signaled, be sure to become fully aware of what the feelings are like, continue to imagine the scene that I am describing, and hold your finger raised until you notice that the anxiety feelings are disappearing ... Visualize yourself at work at your desk with the boss nearby ... Imagine that your mind is thinking, "what if I make a mistake?" ... What if I should do something foolish? ... You can feel the anxiety beginning to increase as your mind is thinking these things ... The tension and anxiety are rising as your mind continues to worry about acting foolishly ... Imagine that your heart is beginning to pound ... More and more quickly and intensely ...

When the client signals anxiety, or after 60 seconds of visualization without a signal, introduce the coping statements:

> Just continue to imagine yourself at work at your desk with the boss nearby but now imagine yourself relaxing ... Just visualize yourself relaxing away the tension and anxiety, letting go of the worries ... and as you do, imagine that the relaxation is reducing the speed of your heart, the tension in your muscles ... Calm and peaceful ... quiet and relaxed ... Muscles becoming more and more deeply and more completely relaxed ... Heart-racing just melting away, dissipating ... Breathing becoming more and more smooth and regular ... As you imagine yourself at work with the boss nearby and relaxing more and more ... Just letting go of the tension and anxiety and worrisome thoughts and focusing on your work ...

Continue such patter for another 20 seconds if there is no client signal, then terminate the image and use periodic relaxation patter to deepen the relaxed state for 20 seconds. If there is a client signal, continue with the above patter (eliminating references to the anxiety cues) for another 20 seconds beyond the point where the anxiety signal is dropped:

> Just continue imagining yourself at work with the boss nearby ... Relaxing more and more deeply ... Muscles continuing to smooth out and unwind ... Loosening up and becoming more and more comfortably and deeply relaxed ... Nothing to do but to enjoy the pleasant sensations of relaxation as the relaxation process continues to take place as you imagine yourself peacefully at the office focusing on your work.

Then instruct the client to stop visualizing that scene and to continue focusing on the relaxation for another 20 seconds:

> Stop visualizing that scene and go on relaxing ... Allowing the relaxation process to continue to take place now ... Calm and quiet ... Peaceful and relaxed ... Simply enjoy the pleasant feelings of relaxation and notice how very good it feels to be so deeply and completely relaxed.

Repeat each image a minimum of three times, or, if anxiety was signaled, until the client twice consecutively no longer signals anxiety within 60 seconds or until the client is able to terminate the anxiety fairly quickly (e.g., drops signal within seven seconds).

Over time, the therapist introduces more anxiety cues into the images. Ultimately, the client should be able to imagine all cues typical of their anxiety experience, either without reacting with anxiety or being able to rapidly terminate the anxiety. Varied environmental contexts which incorporate the client's typical anxiety-provoking situations can be used in the imagery scenes. Not every cue will have a chance to be paired with every environmental context. It is ideal to avoid over-repetition of any one context and to cover a representative sampling of contexts and at the same time provide a complete coverage of the kinds of internal anxiety cues typically experienced by the client.

Once formal SCD has been initiated in therapy, clients are asked to use the technique at the end of their own daily relaxation practice sessions. The client begins by repeatedly employing the same image that was introduced in the last therapy session. This image should be used until it no longer elicits anxiety, at which time another image that represents the kinds of anxious experience that they have during the week can be substituted.

As therapy progresses, the therapist can also employ an informal variant of SCD: whenever clients begin discussing upcoming stressful events or current worries, the therapist can, with or without first eliciting a relaxation response, ask them to close their eyes and engage in an SCD presentation involving these situations or worries. For worries in particular, the therapist can merely ask the client to begin worrying, and, once it is established, to let it go and relax. Finally, toward the end of therapy, formal SCD can be done in two additional ways that mimic real life events to a closer degree. One involves sequentially presenting several different images representing a variety of different stressful events or circumstances without the client planning these images beforehand and with only a single presentation of each image. Thus, an element of surprise typical of naturally occurring events can be incorporated into imaginal rehearsals. The other method is to present unplanned images that form a seamless story that is both typical of the client's daily experiences with anxiety and represents a building accumulation of stressful events. During both methods, the client's task is merely to engage in ongoing coping in the usual imaginal way as previously done, but without breaks in the scene presentations.

Once CT (discussed below) has been initiated, self-statements that are anxiety-reducing for a particular client can be added to their repertoire of coping responses during SCD imaginal exposures. Thus, the client is encouraged to use both relaxation responses and perspective shifts to more accurate ways of thinking about the represented situation as means of coping with the anxiety and worry elicited by the scenes.

6.19.6.6 Cognitive Therapy

Basic CT for anxiety disorders (Beck & Emery, 1985) involves a fairly straightforward series of steps which can be summarized by a set of questions that the therapist can ask the client:

(i) How are you perceiving the situation (what are you thinking, imagining, predicting, saying to yourself, interpreting, believing)?

(ii) Is this way of perceiving accurate (what is the evidence)?

(iii) If this way may not be true, what alternative ways are there to see the situation that have evidence for being more or equally accurate?

(iv) Whenever you notice yourself using the old way, why not try viewing the situation with one of the alternatives which may be more true, or try testing out the perspectives by conducting experiments during your daily life to get the evidence you need?

In initial sessions, the therapist already helped clients identify the variety of internal cues associated with their anxiety and worry, including thoughts and images, to notice what immediate effect these cognitions can have, and to notice that their minds and bodies are reacting to something that is not now present and does not exist in the present. Clients are encouraged to view cognitive events as hypotheses rather than as facts, as predictions which may come true in the future. They are asked to estimate the probability that their feared outcome will actually happen. This question is linked to the idea that it is wise for them to avoid only probable dangers rather than every conceivable outcome. If the danger does not have a high likelihood of occurrence, there may be more accurate predictions available. If the danger does have a high likelihood, the client is asked to rate how well he or she could cope with the feared outcome. This question is linked with the idea that we often underestimate our ability to cope with predicted negative events.

When analyzing the truth-value of particular perspectives, clients are asked to consider the advantages and disadvantages of this way of seeing things and to examine the evidence for it. Quite often, very few advantages can be identified, and the anxiety and worry that follow from it are obvious to the client. The probability of something happening in the future is often best estimated by how frequently it has happened in the past, and the therapist and client can search personal history to come up with a meaningful estimate of the chances for the occurrence of the feared event. In addition, statistics are sometimes available to specify the likelihood of some negative events.

Although such traditional methods of helping clients logically analyze the likelihood that their perspectives are accurate or inaccurate are useful, the most important therapeutic task involves learning to observe the environment on a daily basis, fully processing the new information that is available, and using this information to construct new models of the way things actually are. GAD clients have two characteristic tendencies which ordinarily counteract such a process: focusing on mental illusions of the past and the future and excessively attending to the negative features of environmental information. Therapists therefore encourage them to become more objective in their view of events, distinguishing what is actually and descriptively occurring from their interpretations of events or their guesses about what the events may mean. One of the best ways of helping them with these tendencies with regard to their worrying involves daily monitoring of worry content and the specific predictions they are making and then observing the actual outcomes which eventually occur. As mentioned earlier, when an outcome relevant to a worry does occur, they rate how bad the outcome was and how well they coped with it. They are asked to think carefully about each actual outcome and what it might mean for revising their models of the way things are. By using the worry outcome monitoring method, clients create their own new history of events, more objectively considered, upon which to base predictions about the future and ways of seeing themselves in the world.

Because so many anxieties and worries relate to social evaluative situations, distancing methods are often initially useful, wherein clients are asked to logically analyze the thoughts and predictions of another person rather than of themselves (e.g., a friend or acquaintance, a hypothetical stranger, or the therapist role-playing someone else). Such distancing reduces the likelihood of the clients' habitual fear and avoidance responses to such material and allows them to think more clearly and to reason more accurately.

As clients engage in logical analysis and generate more accurate perspectives, they commonly report that the alternative perspective just does not feel true in the way which the old view does, even though it is more logical. There is very good reason for this. Worry contains numerous brief images of negative events, and imagery can result in bodily reactions similar in type and intensity to the actual occurrence of an event. The more one imagines catastrophes, the more real they feel. Of course images of more likely positive outcomes have been occurring far less frequently, so the "feeling" that these are true is less strong. This is one of the reasons why the cognitive products from CT are incorporated into SCD procedure and why realistic images that indicate everything turning out well can be usefully added toward the end of coping image rehearsals.

Although CT begins with specific thoughts and images, the therapist and client eventually identify recurrent, underlying themes that reflect their core beliefs and CT proceeds in a similar fashion using the core beliefs as targets for intervention. Formal assessment of characteristic thought styles can facilitate the identification of such schemata. Among the factors from the Dysfunctional Attitude Scale (Beck, Brown, Steer, & Weissman, 1991), GAD clients in the current Pennsylvania State Project score highest on items referring to fear of disapproval, dependence on others for their happiness, and excessively rigid rules about the way things "should" be. A major goal for CT with GAD, therefore, involves the creation of greater freedom from others' opinions and more flexible rules by which to judge oneself and others. Central to these is the gradual shifting of perspective from an extrinsic to an intrinsic focus. If clients can learn to attend only to the process of their actions and experiences rather than to the outcomes, less time will be spent generating anxiety-provoking cognitions about what negative things might happen, and more time is devoted to attending to the form and quality of the process of living. For example, when people are at work, they can choose to pay attention to the external rewards which may come from the work (and thus create performance anxiety due to fear of failure), or they can choose to focus on the work itself, its intrinsic meaning, or how it can be done in a beautiful, skillful, or joyful manner. Therapists can help clients to identify the specific elements of any activity or role which can be pleasing or life-affirming. Rehearsals in SCD of adopting such alternative perspectives can strengthen the tendency to employ such views on a daily basis.

Clients can have a difficult time at the beginning of CT when they are asked to generate alternative ways of thinking or believing; old habits have often been present for a long time. To facilitate flexible thinking, it is useful first to practice generating multiple perspectives on neutral topics where avoidant tendencies are less strong. For example, how many ways are there to see this pencil? What shapes can we see in the cloud formations? How many different ways do characters in situation comedies on television view the same "objective" event? Such rehearsals of flexible thinking show clients that many ways exist to perceive something and that we can create what those ways are for any situation. When the therapist eventually addresses anxiety-provoking situations, the first step is to generate several possible perspectives (even humorous ones) without initially judging their reasonableness. The second step is to choose one of the perspectives for application, using such touchstones as evaluating the advantages and disadvantages of each perspective and judging its likely accuracy based on logic, probability, and evidence.

In defining their core beliefs, GAD clients also often use words which have a strong emotional impact and which are vague enough to apply to any situation. For example, a common core belief for someone with GAD might be, "I must be perfect at everything I do or I am a failure." A helpful intervention is to ask the client to define what they mean by each word in the above phrase and to determine how realistic it is. For example, how do they determine their own perfection? Do they really need to do everything perfectly, or are there only some things which are important? Is it possible for them to be perfect at everything? Could they still be successful at something if it was not perfect? The therapist can encourage clients to reword their core beliefs so that they are much more specific. For example, instead of using the above phrase, the client might state, "I choose to be the best (mother, teacher, writer) that I can." The latter phrase is much more specific, avoids social comparison by refocusing on self-competition, focuses on the process (trying one's best) instead of the outcome, and implies choice.

A goal of CT for GAD involves the creation of more balanced perspectives, that is, the addition of realistic views to older, habitual, negative views. It is not a matter of looking at the world with rose-colored as opposed to brown glasses, but rather of using clear glasses. With a greater number of perspectives available, clients have choices rather than being determined by habit. Once more adaptive cognitions are employed with success in their daily lives, clients can be introduced to one further step: expectancy-free living. Although planning and anticipation are necessary human processes under certain circumstances, much of the time expectancies are not necessary. In fact, expectancy can distort the processing of information, influence what is recalled from memory, create negative mood states during anticipatory periods if its content involves negative predictions, and can never fully anticipate the way things will turn out anyway. Living with a focus on the present, processing information from the world as it arrives, and trusting one's ability to cope with whatever comes down the road are alternative perspectives potentially useful to practice to help the client discover under what circumstances such a lifestyle can be adaptive and enjoyable. As an anonymous poet from Crete wrote, "I fear nothing; I hope for nothing; I am free."

Because so much of the inner life of the GAD client revolves around fear of potential bad things which might happen, a perspective shift which defines "bad" in a more relativistic way can be helpful. Certainly we need to evaluate and react to events as they occur partly on the basis of our immediate judgments about "good" and "bad." But events reflect only relative and not absolute qualities. For example, the first author as a teenager spent his life savings on a used car, whose engine died two months later. This was a bad thing. However, it forced him to find a new job reachable by public transportation. The new job location was where he met his wife, the best thing that has ever happened to him. We cannot adaptively seek out bad events just because they might lead to something good. But when either bad or good events happen, it is useful to realize that new choices are made available, new doors are opened, and nothing is absolutely bad. If this is the case, then anticipations of bad events, so frequently characteristic of GAD, take on a slightly less rigid meaning.

Decatastrophizing is a useful, traditional CBT method with GAD. Given that clients frequently fear the worst, identifying what exactly that would be, the various steps which it would take to get to that point, the probability of each step, what alternative perspectives could be adopted, and what coping resources exist for each step can often significantly reduce the fear associated with the area of concern. If the probability of all steps occurring together is established to be quite low and the likelihood of inability to cope is determined to be nearly nonexistent, clients can be encouraged to consider a horse-betting analogy. Would they bet their life savings on a horse which has come in last every time in the past or on a horse which has always won? They may wish to consider how they are betting their daily emotional life in a similar context.

Behavioral experiments can be constructed for clients to conduct between sessions to test old and alternative perspectives. Such experiments often provide opportunities for graduated exposure as well. Asking clients to approach feared situations to obtain evidence for beliefs can thus serve a dual purpose. Although circumscribed anxiety-provoking situations are not as salient in GAD, clients do often engage in a variety of subtle avoidances (Butler, Cullington, Hibbert, Klimes, & Gelder, 1987). One frequent area of avoidance inherent to GAD is cognitive avoidance. For true exposure to be enacted, cognition and behavior must both reflect an approach. What perspective clients adopt during an approach behavior can make a difference to the extinction process. Fear of bees is an instructive example. Fearful people often swat at bees flying around them, an irrational, nonapproach act that increases rather than decreases the likelihood of being stung. Eliminating the swatting and sitting motionless until the bee flies away still includes avoidant behavior (a freezing response). Opening up one's palm to the bee in an invitation to land on the hand is an example of a true approach response in both cognition and behavior and is likely to facilitate extinction and changes in threatening meaning.

GAD clients are excellent "yes, but ..." debaters. They have rich associative networks which construct an elaborate view of why the world is a dangerous place and why they may not be able to cope. Although CT can be seen as a very systematic, step-by-step system for modifying cognitions, there is no need to follow those steps in textbook fashion. Indeed, overly rigid therapist pursuits can lead to arguments, frustration on the part of both therapist and client, and ruptures in the therapeutic alliance which make therapy less likely to be beneficial (Safran & Segal, 1990). The alternative is to reduce the pressure on both parties, deemphasizing the need to reach a specific outcome in any particular session. The real goal is the creation of a process designed to loosen up the client's rigid, habitual modes of thinking and to generate more flexible thinking in a gradual fashion. There are many specific techniques in CBT, and the therapist can often move smoothly among topics and techniques from moment to moment in the session. As therapists do so, they watch for a softening of the client's reactions to any one of these and pursue any openings which emerge more deeply until signs of avoidance or resistance are encountered. Such an approach keeps the therapeutic process flowing without significant obstacles and no sense of failure.

6.19.7 SUMMARY AND IMPLICATIONS FOR FUTURE THERAPY DEVELOPMENT

Basic research has revealed several interrelated features for GAD which fall into three major domains: GAD (i) involves a process (worry) which functions as cognitive avoidance to anxiety-provoking material, prevents emotional processing, and thus contributes to the maintenance of anxious meanings; (ii) has a distinctive physiology typified by autonomic inflexibility and deficiency in parasympathetic tone; and (iii) is associated with specific current interpersonal problems which may have developed partly on the basis of insecure childhood attachment. Outcome findings suggest that combined CBT is likely to be the most effective intrapersonal treatment approach for producing long-lasting change, but absolute degree of improvement remains unsatisfactory. It seems likely that something else must be added to "best available" CBT techniques to yield greater and maintained clinical improvement.

Prior theorizing about GAD has emphasized the role of habitual, nonlinear, dynamical interaction of multiple intrapersonal processes (attention, thought, imagery, affect, physiology) over time in response to constantly perceived threat. The superiority of combined CBT relative to its elements does suggest that the therapeutic targeting of all relevant systems will yield the greatest change. However, attachment and interpersonal problem results now suggest an extension of this model to include another relevant system, that is, the habitual cycles of interpersonal behavior which may also contribute to maintenance of the maladaptive anxiety process. Perceived threats are often significantly related to interpersonal situations (e.g., fears of criticism for socially avoidant clients, fears of being attacked or controlled by others for dominant-hostile clients, fears of failure to take care of others and so to obtain love and approval for overly nurturant clients). The interpersonal behaviors learned to cope with these fears can be readily seen in a functional sense as negatively reinforced avoidance of such feared interpersonal consequences and of the developmentally based anxieties which underlie those interpersonal fears.

Taken together, results and theory extension provide a rich context for understanding the nature, functions, maintenance, and origins of GAD and thereby for guiding the development of more effective treatment. They suggest that future research and therapy development would benefit from a focus on (i) identification of the specific nature of the core fears in GAD and the development of more powerful methods for the emotional processing of those fears, (ii) the creation of more impactful relaxation methods for increasing parasympathetic tone, and (iii) the pursuit of the functional role of interpersonal problems in GAD process and how best to modify these problems.

ACKNOWLEDGMENT

This research was supported in part by National Institute of Mental Health Research Grant MH-39172 to T.D.B.

6.19.8 REFERENCES

Abel, J. L., & Borkovec, T. D. (1995). Generalizability of *DSM-III-R* generalized anxiety disorders to proposed DSM-IV criteria and cross-validation of proposed changes. *Journal of Anxiety Disorders, 9,* 303–315.

Alden, L. E., Wiggins, J. S., & Pincus, A. L. (1990). Construction of circumplex scales for the Inventory of Interpersonal Problems. *Journal of Personality Assessment, 55,* 521–536.

Andrews, V. H., & Borkovec, T. D. (1988). The differential effects of induction of worry, somatic anxiety, and depression on emotional experience. *Journal of Behavior Therapy and Experimental Psychiatry, 19,* 21–26.

American Psychiatric Association (1980). *Diagnostic and statistical manual* (3rd ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association (1987). *Diagnostic and statistical manual* (3rd ed. Rev.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (1994). *Diagnostic and statistical manual* (4th ed.) Washington, DC: American Psychiatric Association.

Barlow, D. H. (1988). *Anxiety and its disorders: The nature and treatment of anxiety and panic.* New York: Guilford Press.

Barlow, D. H., & DiNardo, P. A. (1991). The diagnosis of generalized anxiety disorder: Development, current status, and future direction. In R. M. Rapee & D. H. Barlow (Eds.), *Chronic anxiety: Generalized anxiety disorder and mixed anxiety-depression* (pp. 95–118). New York: Guilford Press.

Beck, A. T., Brown, G., Steer, R. A., & Weissman, A. N. (1991). Factor analysis of the Dysfunctional Attitude Scale in a clinical population. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 478–483.

Beck, A. T., & Emery, G. (1985). *Anxiety disorders and phobias: A cognitive perspective.* New York: Basic Books.

Bernstein. D. A., & Borkovec, T. D. (1973). *Progressive relaxation training.* Champaign, IL: Research Press.

Borkovec, T. D. (1994). The nature, functions, and origins of worry. In G. C. L. Davey & F. Tallis (Eds.), *Worrying: Perspectives in theory, assessment, and treatment* (pp. 5–34). New York: Wiley.

Borkovec, T. D. (1995, July). *Implication of basic research for treatment developments in generalized anxiety disorder.* Paper presented at the World Congress of Behavioural and Cognitive Therapies, Copenhagen.

Borkovec, T. D., Abel, J. L., & Newman, H. (1995). The effects of therapy on comorbid conditions in generalized anxiety disorder. *Journal of Consulting and Clinical Psychology, 63,* 479–483.

Borkovec, T. D., & Costello, E. (1993). Efficacy of applied relaxation and cognitive behavioral therapy in the treatment of generalized anxiety disorder. *Journal of Consulting and Clinical Psychology, 61,* 611–619.

Borkovec, T. D., & Hu, S. (1990). The effect of worry on cardiovascular response to phobic imagery. *Behaviour Research and Therapy, 28,* 69–73.

Borkovec, T. D., & Inz, J. (1990). The nature of worry in generalized anxiety disorder: A predominance of thought activity. *Behaviour Research and Therapy, 28,* 153–158.

Borkovec, T. D., Lyonfields, J. D., Wiser, S. L., & Diehl, L. (1993a). The role of worrisome thinking in the suppression of cardiovascular response to phobic imagery. *Behaviour Research and Therapy, 31,* 321–324.

Borkovec, T. D., & Mathews, A. M. (1988). Treatment of nonphobic anxiety disorders: A comparison of nondirective, cognitive, and coping desensitization therapy. *Journal of Consulting and Clinical Psychology, 56,* 877–884.

Borkovec, T. D., Mathews, A. M., Chambers, A., Ebrahimi, S., Lytle, R., & Nelson, R. (1987). The effects of relaxation training with cognitive therapy or nondirective therapy and the role of relaxation-induced anxiety in the treatment of generalized anxiety. *Journal of Consulting and Clinical Psychology, 25,* 883–888.

Borkovec, T. D., Metzger, R., & Pruzinsky, T. (1986). Anxiety, worry, and the self. In L. Hartman & K. R. Blankstein (Eds.), *Perception of self in emotional disorder and psychotherapy* (pp. 219–260). New York: Plenum.

Borkovec, T. D., Robinson, E., Pruzinsky, T., & DePree, J. A. (1983b). Preliminary exploration of worry: Some characteristics and processes. *Behaviour Research and Therapy, 21,* 9–16.

Borkovec, T. D., & Roemer, L. (1995). Perceived functions of worry among generalized anxiety disorder subjects: Distraction from more emotional topics? *Journal of Behavior Therapy and Experimental Psychiatry, 26,* 25–30.

Borkovec, T. D., Shadick, R. N., & Hopkins, M. (1991). The nature of normal versus pathological worry. In R. Rapee & D. H. Barlow (Eds.), *Chronic anxiety and generalized anxiety disorder* (pp. 29–51). New York: Guilford Press.

Borkovec, T. D., & Whisman, M. A. (1996). Psychological treatment for generalized anxiety disorder. In M. R. Mavissakalian & R. F. Prien (Eds.), *Long-term treatments of anxiety disorders* (pp. 171–199). Washington, DC: American Psychiatric Association.

Borkovec, T. D., Wilkinson, L., Folensbee, R., & Lerman, C. (1983c). Stimulus control applications to the treatment of worry. *Behaviour Research and Therapy, 21,* 247–251.

Bowlby, J. (1982). *Attachment and loss.* New York: Basic Books.

Brown, T. A., & Barlow, D. H. (1992). Comorbidity among anxiety disorders: Implications for treatment and *DSM-IV. Journal of Consulting and Clinical Psychology, 60,* 835–844.

Brown, T. A., Barlow, D. H., & Liebowitz, M. R. (1994). The empirical basis of generalized anxiety disorder. *American Journal of Psychiatry, 151,* 1272–1280.

Butler, B., Cullington, A., Hibbert, G., Klimes, I., & Gelder, M. (1987). Anxiety management for persistent generalized anxiety. *British Journal of Psychiatry, 151,* 535–542.

Butler, G., Wells, A., & Dewick, H. (1995). Differential effects of worry and imagery after exposure to a stressful stimulus: A pilot study. *Behavioural and Cognitive Psychotherapy, 23,* 45–56.

Carter, W. R., Johnson, M. C., & Borkovec, T. D. (1986). Worry: An electrocortical analysis. *Advances in Behaviour Research and Therapy, 8,* 193–204.

Chambless, D. L., & Gills, M. M. (1993). Cognitive therapy of anxiety disorders. *Journal of Consulting and Clinical Psychology, 61,* 248–260.

Chambless, D. L., Sanderson, W. C., Shoham, V., Johnson, S. B., Pope, K. S., Crits-Christoph, P., Baker, M., Johnson, B., Woody, S. R., Sue, S., Beutler, L., Williams, D. A., & McCurry, S. (1996). An update on empirically validated therapies. *The Clinical Psychologist, 49,* 5–18.

Clark, D. A. (1986). Factors influencing the retrieval and control of negative cognitions. *Behaviour Research and Therapy, 24,* 151–159.

Clark, D. A., & DeSilva, P. (1985). The nature of depressive and anxious, intrusive thoughts: Distinct or uniform phenomena? *Behaviour Research and Therapy, 23,* 383–393.

Davey, G. C. L., & Matchett, G. (1994). Unconditioned stimulus rehearsal and the retention and enhancement of differential "fear" conditioning: Effects of trait and state anxiety. *Journal of Abnormal Psychology, 103,* 708–718.

Davis, M. A. (1980). A multidimensional approach to individual differences in empathy. *Journal of Personality and Social Psychology, 53,* 397–410.

Deffenbacher, J. L. (1980). Worry and emotionality in test anxiety. In I. G. Sarason (Ed.), *Test anxiety: Theory, research and application.* Hillsdale, NJ: Erlbaum.

Durham, R. C., Murphy, T., Allan, T., Richard, K., Treliving, L. R., & Fenton, G. W. (1994). Cognitive therapy, analytic psychotherapy and anxiety management for generalised anxiety disorder. *British Journal of Psychiatry, 165,* 315–323.

Durham, R. C., & Turvey, A. A. (1987). Cognitive therapy vs. behaviour therapy in the treatment of chronic generalised anxiety disorder. *Behaviour Research and Therapy, 25,* 229–234.

Edwards, S., & Dickerson, M. (1987a). Intrusive unwanted thoughts: A two-stage model of control. *British Journal of Medical Psychology, 60,* 317–328.

Edwards, S., & Dickerson, M. (1987b). On the similarity of positive and negative intrusions. *Behaviour Research and Therapy, 25,* 207–211.

Elliott, T. (1990). *The role of autonomic arousal in generalized anxiety disorder.* Unpublished doctoral dissertation, Pennsylvania State University, University Park, PA.

England, S. L., & Dickerson, M. (1988). Intrusive thoughts: Unpleasantness not the major cause of uncontrollability. *Behaviour Research and Therapy, 26,* 279–282.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99,* 20–35.

George, C., Kaplan, N., & Main, M. (1985). *An adult attachment interview.* Unpublished manuscript, University of California, Berkeley, CA.

Goldberger, A. L. (1992). Applications of chaos to physiology and medicine. In J. H. Kim & J. Stringer (Eds.), *Applied chaos.* New York: Wiley.

Goldfried, M. R. (1971). Systematic desensitization as training in self-control. *Journal of Consulting and Clinical Psychology, 37,* 228–234.

Gray, J. A. (1982). Precis of "The neurophysiology of anxiety: An enquiry into the functions of the septo-hippocampal system." *Behavioral and Brain Sciences, 5,* 469–534.

Greenberg, L. S., & Safran, J. D. (1987). *Emotions in psychotherapy.* New York: Guilford Press.

Grossman, P., Stemmier, G., & Meinhardt, E. (1990). Paced respiratory sinus arrythmia as an index of cardiac parasympathetic tone during varying behavioral tasks. *Psychophysiology, 27,* 404–416.

Grossman, P., & Svebak, S. (1987). Respiratory sinus arrhythmia as an index of parasympathetic cardiac control during active coping. *Psychophysiology, 24,* 228–235.

Heide, F., & Borkovec, T. D. (1984). Relaxation-induced anxiety: Mechanisms and theoretical implications. *Behaviour Research and Therapy, 22,* 1–12.

Heide, F. J., & Borkovec, T. D. (1983). Relaxation-induced anxiety: Paradoxical anxiety enhancement due to relaxation training. *Journal of Consulting and Clinical Psychology, 51,* 171–182.

Hoehn-Saric, R., & McLeod, D. R. (1988). The peripheral sympathetic nervous system: Its role in normal and pathological anxiety. *Psychiatric Clinics of North America, 11,* 375–386.

Hoehn-Saric, R., McLeod, D. R., & Zimmerli, W. D. (1989). Somatic manifestations in women with generalized anxiety disorder: Physiological responses to psychological stress. *Archives of General Psychiatry, 46,* 1113–1119.

Inz, J. (1990). *EEG activity in generalized anxiety disorder.* Unpublished doctoral dissertation, Pennsylvania State University, University Park, PA.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992). Generalized anxiety disorder in women: A population-based twin study. *Archives of General Psychiatry, 49*, 267–272.

Kent, G., & Jambunathan, P. (1989). A longitudinal study of the intrusiveness of cognitions in text anxiety. *Behaviour Research and Therapy, 27,* 43–50.

Klein, M. H., Mathieu-Coughlan, P., & Kiesler, D. (1986). The experiencing scales. In L. S. Greenberg & W. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 21–71). New York: Guilford Press.

Lichtenstein, J., & Cassidy, J. (1991, March). *The inventory of adult attachment: Validation of a new measure.* Paper presented at the annual meeting of the Society for Research in Child Development, Seattle, WA.

Lindsay, W. R., Gamsu, C. V., McLaughlin, E., Hood, E., & Espie, C. A. (1987). A controlled trial of treatment for generalized anxiety disorder. *British Journal of Clinical Psychology, 26, 3–15.*

Lyonfields, J. D., Borkovec, T. D., & Thayer, J. F. (1995). Vagal tone in generalized anxiety disorder and the effects of aversive imagery and worrisome thinking. *Behavior Therapy, 26,* 457–466.

Main, M., & Goldwyn, R. (in press). Interview-based adult attachment classifications: Related to infant–mother and infant–father attachment. *Developmental Psychology*.

Marten, P. A., Brown, T. A., Barlow, D. H., Borkovec, T. D., Shear, M. K., & Lydiard, R. B. (1993). Evaluation of the ratings comprising the associated symptom criterion of DSM-III-R generalized anxiety disorder. *Journal of Nervous and Mental Disease, 181,* 676–682.

Mathews, A. (1990). Why worry? The cognitive function of anxiety. *Behaviour Research and Therapy, 28,* 455–468.

Mathews, A., & MacLeod, C. (1994). Cognitive approaches to emotion and emotional disorders. *Annual Review of Psychology, 45,* 25–50.

McGuire, W. J., & McGuire, C. V. (1991). The content, structure, and operation of thought systems. In R. S. Wyer & T. K. Srull (Eds.), *Advances in social cognition* (pp. 1–78). Hillsdale, NJ: Erlbaum.

Metzger, R. L., Miller, M. L., Cohen, M., Sofka, M., & Borkovec, T. D. (1990). Worry changes decision making: The effect of negative thoughts on cognitive processing. *Journal of Clinical Psychology, 46,* 78–88.

Noyes, R., Clarkson, C., Crowe, R. R., Yates, W. R., & McChesney, C. M. (1987). A family study of generalized anxiety disorder. *American Journal of Psychiatry, 144,* 1019–1024.

Noyes, R., Woodman, C., Garvey, M. J., Cook, B. L., Suelzer, M., Chancy, J., & Anderson, D. J. (1992). Generalized anxiety disorder vs. panic disorder: Distinguishing characteristics and patterns of comorbidity. *Journal of Nervous and Mental Disease, 180,* 369–379.

Öst, L. (1987). Applied relaxation: Description of a coping technique and review of controlled studies. *Behaviour Research and Therapy, 25,* 397–409.

Parkinson, L., & Rachman, S. (1981). The nature of intrusive thoughts. *Advances in Behaviour Research and Therapy, 3,* 101–110.

Peasley, C. E., Molina, S., & Borkovec, T. D. (1994, November). *Empathy in generalized anxiety disorder*. Paper presented at the annual meeting of the Association for the Advancement of Behavior Therapy, San Diego, CA.

Pincus, A. L., & Borkovec, T. D. (1994, June). *Interpersonal problems in generalized anxiety disorder: Preliminary clustering of patients' interpersonal dysfunction*. Paper presented at the annual meeting of the American Psychological Society, New York.

Porges, S. W. (1992). Autonomic regulation and attention. In B. A. Campbell, H. Hayne, & R. Richardson (Eds.), *Attention and information processing in infants and adults* (pp. 201–223). Hillsdale, NJ: Erlbaum.

Pruzinsky, T., & Borkovec, T. D. (1990). Cognitive and personality characteristics of worriers. *Behaviour Research and Therapy, 28,* 507–512.

Richards, J. E. (1987). Infant visual development and respiratory sinus arrhythmia. *Child Development, 58,* 488–496.

Roemer, L., Borkovec, M., Posa, S., & Lyonfields, J. D. (1991, November). *Generalized anxiety disorder in an analogue population: The role of past trauma*. Paper presented at the annual meeting of the Association for the Advancement of Behavior Therapy, New York.

Safran, J., & Segal, Z. V. (1990). *Interpersonal process in cognitive therapy*. New York: Basic Books.

Salkovskis, P. M., & Harrision, J. (1984). Abnormal and normal obsessions: A replication. *Behaviour Research and Therapy, 22,* 549–552.

Shadick, R. N., Roemer, L., Hopkins, M. B., & Borkovec, T. D. (1991, November). *The nature of worrisome thoughts*. Paper presented at the annual meeting of the Association for the Advancement of Behavior Therapy, New York.

Smith, L. C. (1984). Semantic satiation affects category membership decision time but not lexical priming. *Memory and Cognition, 12,* 483–488.

Suinn, R. M., & Richardson, F. (1971). Anxiety management training: A nonspecific behavior therapy program for anxiety control. *Behavior Therapy, 2,* 498–510.

Sutherland, G., Newman, B., & Rachman, S. (1982). Experimental investigations of the relations between mood and intrusive unwanted cognitions. *British Journal of Medical Psychology, 55,* 127–138.

Tallis, F., Eysenck, M., & Mathews, A. (1991). Elevated evidence requirements and worry. *Personality and Individual Differences, 12,* 21–27.

Thayer, J., & Borkovec, T. D. (1995, July). *Cardiovascular evidence for higher-order classical aversive conditioning in generalized anxiety disorder*. Paper presented at the World Congress of Behavioural and Cognitive Therapies, Copenhagen.

Thayer, J. F., Friedman, B. H., & Borkovec, T. D. (1996). Autonomic characteristics of generalized anxiety disorder and worry. *Biological Psychiatry, 39,* 255–266.

Torgersen, S. (1986). Childhood and family characteristics in panic and generalized anxiety disorders. *American Journal of Psychiatry, 143,* 630–632.

Vrana, S. R., Cuthbert, B. N., & Lang, P. J. (1986). Fear imagery and text processing. *Psychophysiology, 23,* 247–253.

Wells, A., & Papageorgiou, C. (1995). Worry and the incubation of intrusive images following stress. *Behaviour Research and Therapy, 33,* 579–583.

Wittchen, H. -U., Zhao, S., Kessler, R. C., & Eaton, W. W. (1994). DSM-III-R generalized anxiety disorder in the national comorbidity survey. *Archives of General Psychiatry, 51,* 355–364.

York, D., Borkovec, T. D., Vasey, M., & Stern, R. (1987). Effects of worry and somatic anxiety induction on thoughts, emotion, and physiological activity. *Behaviour Research and Therapy, 25,* 523–526.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.20
Specific Phobias

PETER MURIS and HARALD MERCKELBACH
University of Maastricht, The Netherlands

6.20.1 INTRODUCTION 461
6.20.2 SYMPTOMATOLOGY AND EPIDEMIOLOGY 462
6.20.2.1 Case Vignette 462
6.20.2.2 Symptomatology 462
6.20.2.3 Epidemiology 463
6.20.2.4 Genetics 464
6.20.3 DEVELOPMENTAL ASPECTS 464
6.20.4 PATHWAYS TO SPECIFIC PHOBIA 465
6.20.4.1 Classical Conditioning 465
6.20.4.2 Shortcomings of the Traditional Conditioning Approach 466
6.20.4.3 Neoconditioning Model 467
6.20.4.4 Indirect Pathways to Phobia: Modeling and Information 468
6.20.4.5 Disgust Sensitivity 468
6.20.4.6 Retrospective Clinical Studies 469
6.20.5 MAINTENANCE OF PHOBIC FEARS 469
6.20.5.1 Attentional Bias 469
6.20.5.2 Judgmental Bias 470
6.20.5.3 Thought Suppression 470
6.20.6 TREATMENT OF SPECIFIC PHOBIAS 470
6.20.6.1 Exposure-based Treatments 470
6.20.6.2 Eye Movement Desensitization and Reprocessing 471
6.20.6.3 Benzodiazepines 471
6.20.7 CONCLUSIONS 471
6.20.8 REFERENCES 472

6.20.1 INTRODUCTION

Specific phobias are irrational and persistent fears of certain objects or animals. Whereas specific phobia is a relatively recent concept, descriptions of the condition to which it refers have a long tradition in medical history (Errera, 1962). For example, in his writings, Hippocrates refers to a case of a man who displayed an irrational fear of bridges. Likewise, detailed discussions of phobic symtoms can be found in the work of seventeenth or eighteenth century authors such as Descartes and Le Camus.

Since the 1960s our understanding of specific phobias has steadily increased: it is now recognized that learning mechanisms and developmental processes play a crucial role in the etiology of specific phobias; there are now strong indications that certain cognitive processes contribute to the maintenance of phobic symptoms; and we have a rather clear picture of how we should treat patients who suffer from a specific phobia. The present chapter evaluates these aspects in more detail.

6.20.2 SYMPTOMATOLOGY AND EPIDEMIOLOGY

6.20.2.1 Case Vignette

Paula, aged 24, visits a psychologist because of her intense fear of spiders. She reports that she has been afraid of spiders since early childhood. In the last few months, the spider fear has become more intense. Until recently, she lived with her parents, but at the moment she has her own apartment. She is, however, seriously considering the possibility of returning to her parents' home. The reason for this is quite simple: at her parents', she was able to call her father when she suspected that there was a spider in her bedroom. On these occasions, her father inspected the room and removed a spider. Now that she lives in her own apartment, she is unable to call upon her father because her parents live in another town, some 40 miles away. Paula is especially afraid of spiders at night. Consequently, she does not sleep very well. During the day she feels tired and finds it difficult to concentrate on her work. At home, she is constantly alert and keeps scanning the environment for spiders. Returning to her parents' home would probably mean having to give up her job as a secretary.

6.20.2.2 Symptomatology

The case vignette presented above illustrates the criteria for specific phobia that are described in detail in the *Diagnostic and statistical manual of mental disorders*, fourth edition (*DSM-IV*; American Psychiatric Association, 1994). Briefly, the *DSM-IV* stresses that specific phobia contains three crucial elements: (i) the fear is directed at a clearly discernable and circumscribed stimulus (e.g., spiders), (ii) confrontation with this stimulus elicits intense fear and avoidance behavior, and (iii) the fear is excessive and unreasonable to such a degree that it interferes with daily life.

As for the first element of the *DSM-IV* definition, survey studies have shown that some fears (e.g., fear of spiders or snakes) are far more prevalent than others (e.g., fear of electricity). Factor analyses have shown that the most prevalent fears can be grouped together in a relatively small number of distinct categories (Arrindell, Pickersgill, Merckelbach, Ardon, & Cornet, 1991; Frederikson, Annas, Fischer, & Wik, 1996). Likewise, the *DSM-IV* recognizes four types of specific phobias: animal type (e.g., spiders), natural environment type (e.g., water), situational type (e.g., enclosed spaces), and blood–injection–injury type (e.g., dental phobia).

The observation that specific phobias tend to be focused on relatively narrow and circumscribed classes of stimuli has led several authors (e.g., Menzies & Clarke, 1995; Seligman, 1971) to interpret specific phobias in evolutionary terms. Briefly, these authors suggest that the stimuli feared by phobics reflect the dangers that our prehistoric ancestors faced in their (presumed) Pleistocene savannah environment. As a result of natural selection, a readiness to acquire fear of these evolutionary relevant dangers became genetically coded. Consequently, modern man still possesses a preparedness (Seligman, 1971) or an innate tendency (Menzies & Clarke, 1995) to develop fear of spiders, snakes, or blood, for example. Although these evolutionary accounts have attracted considerable attention, other interpretations of the selectivity of fears do exist. For example, the selectivity of specific fears may be a consequence of the negative connotations that certain stimuli have in our culture (e.g., spiders are commonly seen as dirty [Davey, 1995]).

The second crucial element of specific phobias is concerned with the nature of intense fear emotions. A good framework for discussing this point is the three-systems-model proposed by Lang (1968). According to this model, emotions such as fear consist of three relatively independent components: a physiological component, a subjective component, and a behavioral component. Thus, fear is reflected in autonomic symptoms (e.g., tachycardia, increased respiration, etc.), subjective feelings of apprehension, and avoidance or escape behavior. In diagnosis or therapy outcome evaluation, it is important to measure each of these components. It should be noted that each component can be measured: the physiological component can be monitored with psychophysiological techniques (recording of heart rate, respiration, etc. [Hugdahl, 1989]), the subjective component can be quantified with specially constructed self-report questionnaires (e.g., the Spider Phobia Questionnaire; Klorman, Weerts, Hastings, Melamed, & Lang, 1974), and the behavioral component can be measured with standardized approach tasks (i.e., tasks which require a stepwise approach towards a phobic object [Öst, Salkovskis, & Hellström, 1991b]).

In passing, it should be noted that not all types of specific phobias display a comparable profile with regard to the three components. For instance, both animal phobia and blood–injection–injury phobia are accompanied by subjective reports of distress. However, whereas in animal phobia distress usually takes the form of fear, blood–injection–injury phobia is associated with strong subjective feelings of disgust and repulsion (e.g., Page, 1994). Also, confrontation with the phobic stimulus elicits sympathetic activation (e.g., tachycardia) in animal phobia, but parasympathetic activation (e.g., bradycardia) in blood–injection–injury phobia. That is, although exposure to the phobic stimulus produces heightened arousal in animal phobics, lowered arousal that may result in fainting is seen in blood–injection–injury phobia (Thyer, Himle, & Curtis, 1985).

As far as the content of subjective fear is concerned, there are also major differences between situational phobia (especially claustrophobia) and animal phobia. Subjective fear in claustrophobia is not only focused on danger, but also on anxiety expectations (e.g., fear of going crazy) and bodily sensations. The latter components are less prominent in animal phobias (Craske, Mohlman, Yi, Glover, & Valeri, 1995).

The third characteristic of specific phobias that requires some elaboration is their irrational nature. The patient suffering from, say, an intense spider phobia, readily admits that his or her fear is excessive. Nevertheless, he or she is unable to inhibit fear responses when exposed to spiders. How can one account for this failure to control fear responses? A first possibility is that, although the patient is convinced about the irrationality of the phobia when questioned in the absence of the feared stimulus, the presence of this stimulus may activate unreasonable and frightening danger expectations. There is some empirical support for this line of reasoning. For example, De Jongh, Muris, Schoenmakers, and Ter Horst (1995a) asked dental phobics to rate how much they were inclined to believe highly unrealistic dentist-related propositions. Data showed that in anticipation of dental treatment, dental phobics report various catastrophic beliefs about the dentist (e.g., dentists are often impatient), the treatment (e.g., everything will go wrong), and their own reactions (e.g., I can't control myself). Such catastrophic expectations also have been described in other types of specific phobia (e.g., Arntz, Lavy, Van den Berg, & Van Rijsoort, 1993; Thorpe & Salkovskis, 1995).

A somewhat different stance is taken by Öhman (1993). According to Öhman, phobic stimuli are analyzed by fast and subcortical information processing routines. These information processing routines provide a rough analysis of the stimulus and then initiate an immediate fear response. Consequently, even before the phobic patient becomes fully aware of the phobic stimulus, a fear response is already on its way. This might explain why phobics experience their fears as automatic and uncontrollable. Support for Öhman's (1993) hypothesis comes from a study in which it was shown that subliminal presentations of phobic slides (i.e., spiders and snakes) are sufficient to elicit autonomic responding in phobic but not in control subjects (Öhman & Soares, 1994). This finding fits well with the animal studies of LeDoux (1992). This reseacher found evidence for a subcortical fear pathway linking the thalamus with the amygdala. The thalamo–amygdala pathway is a "quick and dirty" transmission route in which the thalamus carries out a crude analysis of the sensory input and then activates the amygdala, which, in turn, generates an emotional response. According to LeDoux (1992, p. 276), this pathway has adaptive value: "the thalamo–amygdala system might be especially useful as a processing channel under conditions where rapid responses are required to threatening stimuli. In such situations, it may be more important to respond quickly than to be certain that the stimulus merits a response." Thus, one may speculate that in specific phobia, learning experiences have affected the thalamo–amygdala pathway in such a way that it is oversensitive to certain phobic cues.

6.20.2.3 Epidemiology

Epidemiological studies indicate that the lifetime and the 12-month prevalence of specific phobias are as high as 11.3% and 8.8%, respectively (Kessler et al., 1994). These data suggest that specific phobia is one of the most frequent mental disorders.

In his community study among women, Costello (1982) found that animal fears (including fear of dogs, snakes, etc.) were the most prevalent, followed by nature (e.g., heights), and mutilation (e.g., injections) fears. In effect, animal fears occurred in nearly 43% of the women. Whether these fears represent clinical phobias is not clear. Employing the *DSM-IV* criteria for specific phobias, Frederikson et al. (1996) found that situational and environmental phobias had the highest point-prevalence rate (13.2%), followed by animal phobias and blood–injection–injury phobias (7.9% and 3.0%, respectively).

Most specific phobias are diagnosed more frequently in women than in men. This is especially true for animal phobias and situational and environmental phobias, but less so for blood–injection–injury phobias (Frederikson et al., 1996). There are strong indications that this sex difference in the prevalence of specific phobias cannot be explained fully by assuming that it is more socially permissible for women to report fear than for men (Cornelius & Averill, 1983). Probably, other predisposing factors (such as sex hormones, see discussed below) play a significant role in the skewed sex distribution of fears and phobias.

With regard to the natural course of specific phobia, Wittchen (1988, p. 14) summarized the results of his longitudinal community study as follows: "The natural course of simple phobias is in the majority of cases chronic and can be characterized by the persistence of mild rather than severe symptoms of anxiety over decades. Only 16% remitted completely over the follow-up period of 7 years: thus, only very few spontaneous remissions could be observed." In passing, it should be noted that specific phobia often occur as an additional diagnosis in panic disorder and generalized anxiety disorder (De Ruiter, Rijken, Garssen, Van Schaik, & Kraaimaat, 1989).

6.20.2.4 Genetics

Several large-scale twin studies have been carried out to explore to what extent genetic factors predispose to phobic fear (e.g., Kendler, Neale, Kessler, Heath, & Eaves, 1992). Briefly, these studies have found that the genetic contribution to specific phobias is significant, though small. Moreover, the genetic factors underlying specific fears are of a general nature. That is, there is no inheritance of a specific fear, but inheritance of a general trait that predisposes to neurotic complaints. In some cases these complaints amount to a specific phobia, but in other cases the complaints take the form of social phobia, agoraphobia, and so on (Andrews, Stewart, Morris-Yates, Holt, & Henderson, 1990).

6.20.3 DEVELOPMENTAL ASPECTS

In general, specific phobias are characterized by early onset ages. For example, Öst (1987) found mean onset ages of seven and nine years for animal and blood–injury–injection phobias, respectively. An early onset is also found in natural environment phobias. For example, a considerable percentage of height-fearful subjects report that their fear has always been present (Menzies & Clarke, 1993). It should be noted, however, that situational fears such as claustrophobia differ from this pattern in that they have a later onset. For instance, Öst (1987) reported a mean onset age of 20 years for his sample of claustrophobics. Nevertheless, a substantial proportion of specific phobias begin in childhood. This finding is underlined by surveys of subclinical fears among children. In an early study of MacFarlane, Allen, and Honzik (1954), it was found that only a small minority of a large sample of children ($N = 1096$) displayed no fear reactions. A similar finding was reported by Lapouse and Monk (1959). These authors found, for example, that in 6–8 year olds, nearly 44% had a mild fear of blood. Altogether, it appears that mild to moderate fears are quite common in childhood. In this connection, there are several points that deserve comment.

First, studies evaluating specific fears in children have consistently found that girls report more fears than boys (Ollendick, Yule, & Ollier, 1991). Second, the mild fears seen in children often represent transient developmental phenomena. That is to say, childhood fears follow a predictable course (Marks, 1987). For example, Bauer (1976) found that fear of ghosts and fear of animals are common in children aged 4–8. In contrast, fear of injury is more characteristic of children aged 10–12. According to Bauer, these fluctuations in childhood fear are closely tied to cognitive development. Younger children would rely on global perception and animistic concepts in interpreting cause–effect relationships. This would explain why their fears are directed at ghosts and animals. Bauer's emphasis on cognitive development makes sense. For example, the visual cliff phenomenon (fear of heights) that typically occurs in young babies between four and nine months has long been interpreted as a prototype of a genetically based fear. However, experiments show that this phenomenon critically depends on locomotor development in infants (Bertenthal & Campos, 1984). A (social) developmental interpretation of the different onset ages of fears and phobias is also presented by Öhman (1987), who points out that animal fears typically occur during a developmental phase in which young children become more independent from their parents. From an evolutionary perspective, it may be assumed that during this phase, children are vulnerable to predators. The sensitivity of young children to phobogenic animal stimuli would therefore have survival value.

Third, fear rank orders based on what children indicate as their most feared object have demonstrated that typical topics of adult specific phobias are also very common in children. For example, snakes and spiders feature high in fear rank orders of children (Muris, Merckelbach, Meesters, & Van Lier, 1997). This suggests that the content of childhood fears resembles that of specific phobias in adults.

Fourth, although it is clear that the majority of childhood fears disappear spontaneously, there are also indications that in a subgroup of children, the specific fears persist (Ollendick, 1979). Thus, the critical question to be asked is why in some cases, specific fears appear to continue from childhood into adulthood. Part of the answer probably lies in the heritability factor associated with fear. For example, Stevenson, Batten, and Cherner (1992; p. 984) found suggestive evidence that the influence of this heritability increases with age: "as children move through adolescence this would lead to avoidance behaviour which would reinforce the fearfulness of these individuals genetically prone to fearfulness. The net effect will be an increasing significance of genetic influences over individual differences with age." Furthermore, research suggests that apart from genetic factors, three types of learning might contribute to the persistence of childhood fears, namely classical conditioning, modelling, and negative information transmission.

6.20.4 PATHWAYS TO SPECIFIC PHOBIA

6.20.4.1 Classical Conditioning

Ever since Watson and Rayner's (1920) little Albert study, researchers have tried to explain the etiology of specific phobias in terms of aversive classical conditioning. Little Albert was an infant who initially showed no fear of live animals. However, he did react with fear whenever noise was produced by stricking a steel bar with a hammer. Watson and Rayner then exposed Albert to a white rat and produced a loud noise with the hammer and steel. Albert was given six trials of white rat and noise. Eventually, little Albert reacted with crying and avoidance as soon as he was confronted with the white rat. Watson and Rayner argued that the little Albert case illustrates that phobias can be understood as the outcome of a Pavlovian or classical conditioning process. In technical terms the phobic object (e.g., a rat) is a conditioned stimulus (CS) that initially elicits no fear. An aversive event (e.g., a loud startling noise) is an unconditioned stimulus (UCS) that, by virtue of its biological characteristics, reliably induces a fear response. This fear response is therefore termed the unconditioned response (UCR). If the CS is paired with the UCS, a fear response occurs and becomes associated with the CS. Thereafter, the CS is able to elicit a full blown fear response, even in the absence of the UCS. The fear response has then become a conditioned response (CR) to the CS and, consequently, the CS has turned into a phobic object.

The basic finding of Watson and Rayner, that fear can be conditioned by pairing a neutral CS with an aversive or traumatic UCS, has been replicated many times. For example, in a study by Campbell, Sanderson, and Laverty (1964), subjects listened to a tone (CS) that was followed by an injection of scoline. This injection produced a temporary respiratory paralysis (UCS) and, consequently, intense fear (UCR). Even weeks after this single conditioning trial, subjects reacted with strong autonomic responses (CR) to isolated presentations of the tone (i.e., the tone was not followed by paralysis).

Although laboratory findings such as these illustrate that conditioning processes can, in principle, produce phobia-like fears, they do not establish that in everyday life, classical conditioning is the cause of phobias. However, there is evidence to support the notion that classical conditioning processes do operate outside the laboratory chambers. For example, Kuch, Cox, Evans, and Shulman (1994) reported that of 55 survivors of road vehicle accidents (UCS), 21 (38%) developed a phobia of driving (CS). Furthermore, 80% of the subjects with a severe weather phobia indicated that their fear began after experiencing a severe storm (Westefeld, 1996). Likewise, choking phobia is nearly always the result of a traumatic incident in which the patient chokes on, for example, a fishbone (Greenberg, Stern, & Weilburg, 1988).

The classical conditioning model of specific phobias explains why persons react with subjective fear and physiological arousal when they are exposed to the phobic CS. It does not, however, account for the persistent avoidance behavior that phobics display. The influential two-stage theory of Mowrer (1960) attempts to incorporate this point. Briefly, Mowrer's theory assumes that the development of phobias involves two stages. During the first stage, a pairing of a neutral CS and an aversive UCS results in a conditioned fear response to the CS. During the second stage, the person learns that fear responses to the CS can be reduced by avoiding the CS. A reduction in fear levels that follows avoidance is experienced as a positive state. Consequently, avoidance behavior is reinforced and in time becomes an integral part of the phobic fear. Again, research shows that the scenario depicted by the two-stage theory works in the laboratory. Thus, when subjects are confronted with pairings of light (CS) and shock (UCS) and then are given the opportunity to react with an escape response every time the CS appears, persistent avoidance responding occurs (Malloy & Levis, 1988).

The two-stage theory of phobic fear has straightforward treatment implications. According to this theory, avoidance behavior maintains phobic fear because it prevents prolonged exposure to the phobic object. Consequently, phobic subjects have no opportunity to experience that the phobic CS is no longer accompanied by an aversive UCS. The treatment lesson to be learned from this is that one should confront phobic subjects with their phobic CS while ensuring that they do not react with avoidance to this CS. This strategy, known as exposure therapy, has proven to be very successful in the treatment of specific phobias (Marks, 1987; Öst, 1989).

6.20.4.2 Shortcomings of the Traditional Conditioning Approach

In the 1970s, several authors drew attention to the fact that the classical conditioning model of phobias and its extension, Mowrer's two-stage theory, suffered from a number of serious shortcomings (Eysenck, 1979; Rachman, 1977). First, the conditioning approach fails to explain why specific fears are nonrandomly distributed. Indeed, from a conditioning point of view, one would expect that each object can become a phobic stimulus (CS), provided that it is followed by an aversive event (UCS). As previously stated, this is obviously not the case. For instance, fear of snakes occurs more often than fear of electricity, although there is no obvious reason to suspect that snakes are more frequently associated with UCSs than electricity (Agras, Sylvester, & Oliveau, 1969).

Second, in the literature, excellent illustrations can be found of aversive situations (UCSs e.g., air raids) in which, contrary to all predictions that can be derived from the classical conditioning model, people did not acquire phobic fears (e.g., Saigh, 1984). In the aforementioned study of Kuch et al. (1994) not all accident survivors developed a fear of driving. Similarly, there is abundant evidence to show that aversive experiences (UCSs) with spiders, dentists, or dogs do not give rise to spider phobia, dental phobia or dog phobia, respectively (Merckelbach, Arrindell, Arntz, & De Jong, 1992; Lautch, 1971; DiNardo, Guzy, & Bak, 1988). Clearly, then, trauma (UCS) without subsequent phobia occurs.

Third, over the years, it has become clear that not all specific phobias can be traced back to a confrontation with a traumatic or aversive event (UCS). In other words, some people develop a fear response to a CS, although this CS has never been paired with a UCS (Wolpe, Lande, McNally, & Schotte, 1985). In other words, phobia without preceding trauma (UCS) occurs.

These and other shortcomings of the traditional conditioning approach to phobias paved the way for two types of revision. The first type of revision stressed that conditioning concepts should be connected with biological notions in order to understand the etiology of phobias. The aforementioned preparedness hypothesis (Seligman, 1971) is a good example of such a biologically oriented revision. This hypothesis assumes that conditioning interacts with evolutionary processes to produce phobic fear. Thus, aversive events (UCSs) in the context of evolutionary recent objects (e.g., electricity) would not produce a phobic fear, whereas such events in the context of evolutionary relevant items (e.g., snakes) would easily give rise to a phobic fear. Several experiments have tested this idea in the laboratory. By and large, the results of these experiments are disappointing in that they found no evidence for a connection between ease of conditioning and evolutionary relevance of the CS (Hugdahl & Johnsen, 1989).

Another attempt to revise the conditioning model along biological lines proposes that the phobogenic effects of CS–UCS pairings interact with the neurohormonal state of the organism (Kelley, 1987). In Kelley's words: "neuroses = conditioning × neurohormones" (p. 403). More specifically, Kelley argues that high levels of adrenocorticotropic hormone (ACTH) and vasopressin enhances the effects of CS–UCS pairings whereas high levels of endogenous opioids (i.e., endorphins) attenuate these effects. Accordingly, individual differences in neurohormonal levels might explain why, for example, an automobile accident produces a driving phobia in some people, but not in others (Kuch et al., 1994): some subjects are more vulnerable to CS–UCS pairings because they habitually have heightened levels of ACTH and vasopressin and/or lowered endorphin levels. Although it is evident that phobic fear is accompanied by elevated or reduced levels of certain (neuro)-hormones (Cameron & Nesse, 1988), Kelley's proposal has not been put to a direct test. All that one can conclude from the evidence at hand is that lowering the endorphin levels by applying naltrexone (i.e., an opioid antagonist) intensifies fear in subjects who are already phobic (Merluzzi, Taylor, Boltwood, & Götestam, 1991; Arntz, Merckelbach, & De Jong, 1993).

A related line of research has stressed the involvement of sex hormones in the acquisition of phobic fear. For example, in their review Cameron and Nesse (1988, p. 291) conclude that "sex hormone fluctuations in women during the menstrual cycle may be related to changes in anxiety levels, although fluctuations have not always been observed." Interestingly, Van der Molen, Merckelbach, and Van den Hout (1988) found that pre-menstrual women show an increased susceptibility to fear conditioning. Sex hormone fluctuations may also explain why specific phobias are more often found in women than in men.

6.20.4.3 Neoconditioning Model

The second type of revision of the classical conditioning model is more cognitively oriented. Advocates of this approach point out that conditioning is not the blind, reflexive process that it is often assumed to be (Van den Hout & Merckelbach, 1991). The critical question to be asked is: what exactly happens during conditioning? Take the example of little Albert. Did conditioning result in a reflexive connection between the white rat stimulus (CS) and the fear response (CR)? Or did little Albert learn that the white rat stimulus (CS) predicts the onset of the noise stimulus (UCS) and was it this knowledge that elicited a CR? The first scenario is called S-R learning; the second scenario is known as S-S learning. There are good reasons for believing that, in humans, S-S learning is the most common classical conditioning scenario. That is to say, conditioning is usually a cognitive process during which one learns that the CS predicts the UCS (Rescorla, 1988). It is important to note that an S-S approach (i.e., the neoconditioning model) has a number of interesting ramifications and is able to avoid some of the inadequacies of the traditional conditioning model (Rachman, 1991). Two specific aspects of the neoconditioning model deserve comment.

First, the neoconditioning model emphasizes the phenomenon of latent inhibition. Suppose a subject has extensive experience with a CS (e.g., light) that is not followed by an aversive UCS. What will happen when this subject is later exposed to a new situation in which the CS is suddenly accompanied by a painful UCS (e.g., electric shock)? Will the CS now elicit a conditioned fear response? The answer is no. Laboratory experiments show that it is difficult to condition fear responses to a familiar CS that was formerly never associated with an aversive UCS (Booth, Siddle, & Bond, 1989). Recall that, during conditioning, subjects learn that a CS is the predictor of a UCS. If the subject has experienced a number of times that the CS predicts nothing, an incidental co-occurrence of the CS with an aversive UCS is not powerful enough to make the subject believe that the CS has changed into a predictor of the UCS. In other words, prior experiences with the CS inhibit fear conditioning. A study by Doogan and Thomas (1992) nicely illustrates that the mechanism of latent inhibition is relevant when one tries to understand how trauma (UCS) without phobia can occur. These researchers found that traumatic encounters with dogs are reported by both subjects with fear of dogs and subjects without fear of dogs. However, the latter group reported more previous nonaversive contacts with dogs. In other words, the nonfearful subjects had experienced many occasions on which dogs (CS) did not attack. This suggests that in the nonfearful group, latent inhibition helped to minimize the effects of fear conditioning (Davey, 1989a; De Jongh, Muris, Ter Horst, & Duyx, 1995b).

A second phenomenon that must be considered is called UCS inflation (Davey, 1989b). Briefly, this phenomenon refers to a situation in which subjects are exposed to pairings of a CS and a mild UCS. As a result of this, subjects acquire an S–S association. However, given the low intensity of the UCS, this S–S association leads to a weak conditioned fear response (CR). Now, the following may happen: if subjects during a next phase learn that the UCS is nevertheless dangerous, this postconditioning information will lead to an inflation of the UCS value. Consequently, the conditioned fear response will grow in strength. A clinical example taken from White and Davey illustrates how this kind of scenario can account for cases in which conditioned fear emerges without pertinent trauma:

> an individual may witness an unknown person die of a heart attack on a bus or a train. On future occasions, riding on public transport may evoke memories of this incident but no anxiety. ... Subsequently, however, that individual may be present when a close friend or relative dies of a heart attack, thus inflating the aversive properties of heart attacks. This may then give rise to acute anxiety when riding on public transport. In this particular scenario, public transport has never been directly associated with anxiety-eliciting trauma, but the public transport phobia results from a prior learned association between public transport and heart attacks, and subsequent independent inflation of heart attacks as aversive events. (White & Davey, 1989, p. 165)

To recapitulate, neoconditioning concepts such as latent inhibition and UCS inflation can deal with some inadequacies of the traditional conditioning approach. More specifically, latent inhibition can explain why UCS trauma does not always lead to phobia, whereas UCS inflation can explain why phobia without pertinent UCS trauma is possible. Apart from that, these concepts have implications for the treatment of phobias (Davey, 1992). For example, one may use the principle of latent inhibition as a way of producing immunization against the effects of subsequent CS–UCS trials. Giving dental patients a series of painless sham treatments will probably reduce the likelihood that a subsequent painful treatment results in a dental phobia (De Jongh et al., 1995b). As another example, the principle of UCS inflation suggests that, during treatment, it might be worthwhile to systematically devaluate the UCS. In the case of dental phobia, this could mean that one stresses the point that aversive experiences result in healthy teeth.

6.20.4.4 Indirect Pathways to Phobia: Modeling and Information

The classical conditioning pathway to fear assumes that the subject has direct experience with the CS and the UCS. This is true for straightforward conditioning in which a CS is paired with a traumatic UCS, as well as for subtle conditioning scenarios such as latent inhibition and UCS inflation. However, there are cases in which subjects develop a specific phobia although they have no history of direct experience with the CS (i.e., phobic object) and/or UCS (Rachman, 1991). In these cases, indirect pathways of fear acquisition might play a pivotal role. Rachman (1991) provides a detailed description of two indirect pathways that are important in this context: modeling (imitation or vicarious transmission of fear) and negative information.

The modeling pathway assumes that phobic fears can be acquired by watching parents or peers reacting fearfully to a stimulus (e.g., a spider). This assumption is supported by animal and human laboratory studies. For example, Cook and Mineka (1989) found that rhesus monkeys can acquire a fear of snakes through watching videotapes of two model monkeys reacting fearfully to toy snakes. Likewise, Hygge and Öhman (1978) demonstrated that subjects respond with sympathetic activation to pictures of, for example snakes when they have previously watched others reacting fearfully to these pictures. The reverse is also true: the pioneers of behavior therapy already knew that it is possible to eliminate specific fears by presenting nonfearful models and enhancing social learning through modeling (Kornfeld, 1989).

It seems obvious that negative verbal information about certain stimuli such as that provided by books, television, or significant others may give rise to fear of these stimuli. If this was not the case, prevention programs concerned with smoking or AIDS would make no sense. Although there is a paucity of experimental studies on the fear-evoking characteristics of negative information, social psychological studies show that subjects assign more value to negative than to positive information. There might be good evolutionary reasons for this asymmetry (Pratto & John, 1991).

Assuming that modeling and negative information transmission are closely tied to familial interactions, one would expect that there is a familial aggregation of specific phobias. A study by Fyer et al. (1990) shows this to be the case. These researchers found, for example, that 31% of the relatives of specific phobia probands as compared to 11% of control relatives had a life-time diagnosis of specific phobia. Although this study did not distinguish between different etiological mechanisms (genetic inheritance, modeling, etc.), its results are consistent with the idea that modeling and information do operate in families. Likewise, a recent study by Muris, Steerneman, Merckelbach, and Meesters (1996) demonstrated that fears reported by children are a function of the extent to which mothers express their own fears in the presence of their children. This association was independent of the influence of trait anxiety levels in mothers.

6.20.4.5 Disgust Sensitivity

Disgust sensitivity is believed to play a role in the etiology of certain types of specific phobias, especially animal phobias (Matchett & Davey, 1991). For example, several studies have found that spider-phobic subjects have a higher disgust sensitivity than nonphobic subjects (Merckelbach, De Jong, Arntz, & Schouten, 1993). Furthermore, spider-phobic subjects consider spiders to be more disgusting than do nonphobic subjects (De Jong, Andrea, & Muris, 1997). Finally, behavioral treatment results in a parallel decline of fear of spiders and spiders' disgust-evoking status (De Jong et al., 1997). These data indicate that there is a clear association between disgust and disgust sensitivity on the one hand, and clinical manifestations of spider fear on the other hand.

It should be noted that most studies on disgust and phobia are correlational in nature, and that they do not rule out the possibility that fear enhances disgust and/or disgust sensitivity rather than vice versa. However, an experiment of Webb and Davey (1992) indicates that fear can be causally affected by disgust. In that study, it was found that exposure to revulsive material (such as video scenes of a medical operation) leads to an increase in fear of, for example, spiders or snakes.

6.20.4.6 Retrospective Clinical Studies

Several studies have examined to what extent the aforementioned pathways to fear play a role in the etiology of specific phobia (Kleinknecht, 1994; Merckelbach et al., 1992; Öst, 1987). Most of these studies have used the Phobic Origin Questionnaire (POQ; Öst & Hugdahl, 1981), a self-report instrument that asks phobic patients to indicate to what degree their phobia developed along a direct (i.e., conditioning) and/or an indirect (i.e., modeling, negative information) pathway. A potential problem in these studies is that phobics are invited to provide a retrospective judgement as to the cause of their complaints. With this type of research strategy, all kinds of biases (e.g., memory distortions) may occur. On the other hand, Brewin, Andrews, and Gotlib (1993) have summarized evidence to show that claims about the unreliability of retrospective reports are often exaggerated.

Several interesting findings have emerged from the POQ studies. First, although each of the three pathways is represented in the various types of specific phobia, there are large differences between the types. For example, animal phobics and blood–injection–injury phobics more often ascribe their fears to modeling and negative information than claustrophobics. In contrast, the conditioning pathway is more pronounced in claustrophobia than in animal, blood–injection–injury illness or height phobia (Menzies & Clarke, 1993; Öst, 1987). Interestingly, Öst (1987) found evidence to suggest that modeling and negative information are associated with an early onset age. Second, several studies (Merckelbach, Arntz, & De Jong, 1991; Ollendick & King, 1991) have found that fearful subjects often report more than one pathway. There are even indications that subjects with mixed pathways (e.g., conditioning and modeling) have higher levels of fear. Finally, the three pathways to fear are also found in fearful and phobic children aged 9–14 years (Merckelbach, Muris, & Schouten, 1996; Ollendick & King, 1991). This suggests that these etiological mechanisms do, indeed, operate during the early developmental phases.

6.20.5 MAINTENANCE OF PHOBIC FEARS

Specific phobias are acquired on the basis of direct or indirect learning pathways. But how are phobic fears maintained, once they are acquired? As mentioned earlier, Mowrer's (1960) two-stage model suggests that avoidance behavior is responsible for the conservation of phobic fear: avoidance would minimize direct and prolonged contact with the phobic object and, hence, phobics would not have the opportunity to learn that the CS is a neutral object or otherwise extinguish fear response.

Although the role of avoidance behavior in the conservation of fear seems self-evident, there are other mechanisms that are highly relevant in this context. These mechanisms are cognitive in nature and have been the object of extensive research in the past years (Dalgleish & Watts, 1990; Mineka & Sutton, 1992). There is agreement among most researchers in this field that pathological anxiety (e.g., specific phobia) is not accompanied by a general cognitive dysfunction. In other words, it is not the case that pathological anxiety is associated with deficits in memory, attention, motor function and so on. Instead, cognitive dysfunctions in specific phobias are restricted. First, phobic subjects show evidence of dysfunctions in attentional and inferential (i.e., judgmental) processes. In contrast, depressive patients exhibit dysfunctions in memory rather than attention or judgment (Mineka & Sutton, 1992). Second, the attentional and judgmental biases that characterize phobias are domain specific. That is to say, these biases become apparent if and only if subjects are confronted with fear-relevant stimuli (e.g., the word spider in case of a spider phobic subject).

6.20.5.1 Attentional Bias

Attentional bias refers to the phenomenon of hyperattention to threatening material. One frequently employed technique for demonstrating attentional bias is the modified Stroop color task. This task requires subjects to name the color of a word while ignoring the meaning of that word. A consistent finding in Stroop studies with anxious patients is that their color naming of threatening words is slower than that of neutral words (Eysenck, 1992). For example, spider phobics display retarded color-naming times when they are confronted with spider-related words (e.g., web), but not when they have to color name neutral words (e.g., car). This is because spider phobics automatically direct their attention to the content of the threatening words and this interferes with their main task (i.e., color naming). Note that the phenomenon of attentional bias accords well with Öhman's (1993) hypothesis of a "quick and dirty" subcortical transmission route of phobic information (discussed above). Interestingly, the attentional bias may disappear in successfully treated spider phobics (Watts, McKenna, Sharrock, & Trezise, 1986). As to the clinical consequence of attentional bias, it seems likely that an increased focus on danger and threat stimuli perpetuates phobic fear: attentional bias implicates an increased encoding of threatening material and this, in turn, will elevate fear levels (Mineka & Sutton, 1992).

6.20.5.2 Judgmental Bias

Two types of judgmental bias probably play a role in the maintenance of specific phobia. The first is covariation bias, the tendency to overestimate the association between phobic stimuli and aversive outcomes (Tomarken, Mineka, & Cook, 1989). The experimental demonstration of covariation bias in phobias is straightforward. Phobic and normal subjects are shown a series of slides consisting of fear-relevant (e.g., spiders) and fear-irrelevant pictures (e.g., flowers). Slide offset is followed by one of three outcomes: an aversive shock, a tone, or nothing. Fear-relevant and fear-irrelevant pictures are equally often followed by each of the outcomes. After the series of slides, subjects are asked to estimate the contingencies between slides and outcomes—if a spider picture appeared, how great was the chance that this picture was followed by a shock? Under these experimental conditions, phobic subjects systematically overestimate the contingency between phobic stimuli and aversive outcomes. Again, in successfully treated spider phobics, the covariation bias disappears (De Jong, Merckelbach, Arntz, & Nijman, 1992). De Jong, Van den Hout, and Merckelbach (1995) found a strong and positive correlation between residual covariation bias in treated spider phobics and relapse: the stronger the (post-treatment) overestimation of the contingency between spider picture and aversive shock, the higher the spider fear at two-year follow up. The implication of this covariation bias is clear: phobics have a tendency to attribute aversive experiences to the phobic object and this, in turn, will sustain phobic fear.

A second judgmental bias that may occur in specific phobias is a style of reasoning known as *ex consequentia* inference. Like other people, phobics believe that dangerous situations elicit anxiety. However, unlike many nonphobic subjects, spider phobics seem to believe that anxiety symptoms imply the presence of danger, such as dangerous spiders (Arntz, Rauner, & Van den Hout, 1995). The *ex consequentia* inference probably serves to legitimize the phobic fear and, thus, may maintain the phobia. However, more studies are needed to firmly establish the role of this reasoning style in the maintenance of fear.

6.20.5.3 Thought Suppression

As mentioned earlier, phobic patients are frequently disturbed by anxiety-arousing thoughts. To avoid feelings of discomfort, phobic patients may try to suppress such thoughts. However, research has shown that thought suppression has paradoxical effects. That is, the effort "not to think about" a particular thought may, in fact, lead to the continuation of that thought (Wegner, Schneider, Carter, & White, 1987).

Several authors (Salkovskis, 1989; Wegner, 1989) have put forward that thought suppression might play a role in the etiology and maintenance of anxiety disorders such as specific phobias. In a recent study, Muris, De Jongh, Merckelbach, Postema, and Vet (in press) found some evidence for this claim. In their study, effects of thought suppression in phobic and nonphobic dental patients were examined. The results were as follows. First, as expected, phobic patients reported higher levels of intrusive and negative thinking during dental treatment than nonphobic patients. Second, phobic patients tried harder to suppress their negative thoughts than nonphobic subjects. Finally, in the nonphobic group, suppression resulted in increased levels of anxiety and intrusive thinking. Altogether, the findings suggest that thought suppression is related to pathological manifestations of dental fear. That is to say, suppression intensifies fear in nonphobic subjects, whereas it appears to be part of the habitual "coping" style in phobic patients.

6.20.6 TREATMENT OF SPECIFIC PHOBIAS

6.20.6.1 Exposure-based Treatments

Exposure *in vivo* is the treatment of choice for specific phobias (Marks, 1987; Rachman, 1990). Exposure treatment involves graded and prolonged confrontation with the phobic object. Meanwhile, the therapist encourages the patient to approach the phobic object and to refrain from avoidance behavior. Exposure is often combined with other techniques: modeling by the therapist in the case of animal phobias (Öst, 1989), applied tension to prevent fainting in the case of blood–injection–injury phobias (Öst, Fellenius, & Sterner, 1991a), and cognitive interventions to correct catastrophic misinterpretations of bodily symptoms in the case of claustrophobia (Craske et al., 1995). Exposure treatments yield good results: success percentages of 90% are no exception. Furthermore, controlled studies show that the efficacy of exposure is maintained at long term follow-up (Öst, 1996; Öst et al., 1991b).

The precise mechanism underlying the beneficial effects of exposure are a matter of considerable debate. Some authors have favored an extinction or habituation interpretation of exposure effects (Marks, 1987; Watts, 1979). According to these authors, exposure is the extinction or habituation of a fear CR to a CS that is no longer followed by a UCS. Other authors have argued that exposure effects are linked to cognitive changes rather than response habituation (Foa & Kozak, 1986). They emphasize that, during exposure, the patient learns that the phobic object is not associated with catastrophic events. This will eventually lead to a correction of the phobic fear. This formulation is compatible with the S-S view of phobic fear (Eysenck, 1987), and also with cognitive theories about phobic fear (Last, 1987). There is preliminary evidence that the interpretations of exposure that emphasize cognitive change are closer to the truth than extinction/habituation accounts. For example, Shafran, Booth, and Rachman (1993) recently showed that treatment of claustrophobia is successful to the degree that it corrects certain key cognitions (e.g., "I will suffocate"). Similarly, in their study of one-session exposure treatment of spider phobia, Öst et al. (1991b, p. 421) remark that "the clinical impression from treating these patients is that the most important factor in the one-session treatment is making explicit the patients" catastrophic thoughts concerning the phobic situation and devising the exposure situation in such a way that these can be tested out."

6.20.6.2 Eye Movement Desensitization and Reprocessing

Eye movement desensitization and reprocessing (EMDR) is a relatively new technique that was originally proposed as a treatment method for post-traumatic stress disorder (Shapiro, 1989). During EMDR, patients imaginally expose themselves to a traumatic or aversive memory, while simultaneously engaging in lateral eye movements that are induced by the therapist. The idea is that through eye movements, negative memories are emotionally processed and assimilated. The therapist-induced eye movements would simulate the inhibitory function of rapid eye movement (REM) sleep. Some authors have claimed that EMDR is not only effective in post traumatic stress disorder, but also in specific phobias (Marquis, 1991; Shapiro, 1995). However, this claim rests entirely on case studies. Controlled outcome studies designed to compare the effectivity of EMDR and exposure *in vivo* in the treatment of specific phobias indicate that exposure yields better results, especially on the behavioural level of fear (Muris & Merckelbach, 1997; Muris, Merckelbach, van Haaften, & Mayer, 1997).

6.20.6.3 Benzodiazepines

There is abundant evidence showing that benzodiazepines have anxiolitic effects (Rickels, 1978). Furthermore, animal research indicates that benzodiazepines inhibit conditioned fear responses. Consequently, one would predict that benzodiazepines can be useful in the treatment of specific phobias. However, studies designed to evaluate the effect of benzodiazepines on phobic avoidance have generally yielded disappointing results (Bernadt, Silverstone, & Singleton, 1980; Sartory, MacDonald, & Gray, 1990). Although benzodiazepines might inhibit subjective fear during phobic confrontation, they do not increase approach behavior (Sartory et al., 1990). Gray (1987) has gone one step further and suggested that benzodiazepines might even be harmful in that they create state dependency effects: phobics learn to approach the phobic stimulus when drugged, but this learning is not transferred to a nondrugged state.

6.20.7 CONCLUSIONS

Specific phobias have been well researched. As a result, we now have a comprehensive picture of the important elements that constitute this disorder. To summarize the most important points, specific phobias form a heterogeneous class of disorders (Himle, Crystal, Curtis, & Fluent, 1991). The most radical difference is found between animal phobias and claustrophobia. Compared to animal phobias, claustrophobia is characterized by a more elaborated set of cognitions, a later age of onset, and a more pronounced role of conditioning experiences in the etiology (Öst, 1987). Accordingly, in claustrophobia, but not in animal phobias, effects of exposure treatment are enhanced when the treatment is combined with cognitive interventions (Craske et al., 1995).

Specific fears are common in childhood and there are reasons to believe that cognitive development is closely linked to the content of these fears. In most children these fears represent transitory phenomena. In a small subgroup of children and adolescents, the specific fears become chronic due to conditioning, modeling, negative information and/or a heightened disgust sensitivity. It is also possible that neuroticism may make a general genetic contribution to the acquisition of such phobias. Once a specific phobia exists, it is maintained by processes such as cognitive biases and thought suppression. Consequently, in most cases, the natural course of specific phobias is chronic. However, exposure-based treatments are very effective in reducing specific phobias.

In summary, research has yielded a clear picture of the etiology and maintenance of specific phobias, and of how this disorder should best be treated. Hopefully, in the future, clinicians will profit more from this information than they often do now (Andrews, Hadzi-Pavlovic, Christensen, & Mattick, 1987).

6.20.8 REFERENCES

Agras, S., Sylvester, D., & Oliveau, D. (1969). The epidemiology of common fears and phobias. *Comprehensive Psychiatry, 10,* 151–156.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Andrews, G., Hadzi-Pavlovic, D., Christensen, H., & Mattick, R. (1987). Views of practicing psychiatrists on the treatment of anxiety and somatoform disorders. *American Journal of Psychiatry, 144,* 1331–1334.

Andrews, G., Stewart, G., Morris-Yates, A., Holt, P., & Henderson, S. (1990). Evidence for a general neurotic syndrome. *British Journal of Psychiatry, 157,* 6–12.

Arntz, A., Lavy, E., Van den Berg, G., & Van Rijsoort, S. (1993). Negative beliefs of spider phobics: A psychometric evaluation of the spider phobia beliefs questionnaire. *Advances in Behaviour Research and Therapy, 15,* 257–277.

Arntz, A., Merckelbach, H., & De Jong, P. J. (1993). Opioid antagonist affects behavioral effects of exposure in vivo. *Journal of Consulting and Clinical Psychology, 61,* 865–870.

Arntz, A., Rauner, M., & Van den Hout, M. A. (1995). "If I feel anxious, there must be danger": The fallacy of ex consequentia reasoning in inferring danger in anxiety disorders. *Behaviour Research and Therapy, 33,* 917–925.

Arrindell, W. A., Pickersgill, M. J., Merckelbach, H., Ardon, A. M., & Cornet, F. C. (1991). Phobic dimensions: III. Factor analytic approaches to the study of common phobic fears: An updated review of findings with adult subjects. *Advances in Behaviour Research and Therapy, 13,* 73–130.

Bauer, D. H. (1976). An exploratory study of developmental changes in children's fears. *Journal of Child Psychology and Psychiatry, 17,* 69–74.

Bernadt, M. W., Silverstone, T., & Singleton, W. (1980). Behavioural and subjective effects of beta-adrenergic blockade in phobic subjects. *British Journal of Psychiatry, 137,* 452–457.

Bertenthal, B. L., & Campos, J. J. (1984). A reexamination of fear and its determinants on the visual cliff. *Psychophysiology, 21,* 413–417.

Booth, M. L, Siddle, D. A. T., & Bond, N. W. (1989). Effects of conditioned fear-relevance and preexposure on expectancy and electrodermal measures of human Pavlovian conditioning. *Psychophysiology, 26,* 281–291.

Brewin, C. R., Andrews, B., & Gotlib, I. H. (1993). Psychopathology and early experience: A reappraisal of retrospective reports. *Psychological Bulletin, 113,* 82–98.

Cameron, O. G., & Nesse, R. M. (1988). Systemic hormonal and physiological abnormalities in anxiety disorders. *Psychoneuroendocrinology, 13,* 287–307.

Campbell, D., Sanderson, R. E., & Laverty, S. G. (1964). Characteristics of a conditioned response in human subjects during extinction trials following a single traumatic conditioning trial. *Journal of Abnormal and Social Psychology, 68,* 627–639.

Cook, M., & Mineka, S. (1989). Observational conditioning of fear to fear-relevant versus fear-irrelevant stimuli in rhesus monkeys. *Journal of Abnormal Psychology, 98,* 448–459.

Cornelius, R. R., & Averill, J. R. (1983). Sex differences in fear of spiders. *Journal of Personality and Social Psychology, 45,* 377–383.

Costello, C. G. (1982). Fears and phobias in women: A community study. *Journal of Abnormal Psychology, 91,* 280–286.

Craske, M. G., Mohlman, J., Yi, J., Glover, D., & Valeri, S. (1995). Treatment of claustrophobia and snake/spider phobias: Fear of arousal and fear of context. *Behaviour Research and Therapy, 33,* 197–203.

Dalgleish T., & Watts, F. N. (1990). Biases of attention and memory in disorders of anxiety and depression. *Clinical Psychology Review, 10,* 589–604.

Davey, G. C. L. (1989a). Dental phobics and anxieties: Evidence for conditioning processes in the acquisition and modulation of a learned fear. *Behaviour Research and Therapy, 27,* 52–58.

Davey, G. C. L. (1989b). UCS revaluation and conditioning models of acquired fears. *Behaviour Research and Therapy, 27,* 521–528.

Davey, G. C. L. (1992). Classical conditioning and the acquisition of human fears and phobias: A review and synthesis of the literature. *Advances in Behaviour Research and Therapy, 14,* 29–66.

Davey, G. C. L. (1995). Preparedness and phobias: Specific evolved associations or a generalized expectancy bias. *Behavioral and Brain Sciences, 18,* 289–325.

De Jongh, A., Muris, P., Schoenmakers, N., & Ter Horst, G. (1995a). Negative cognitions of dental phobics: Reliability and validity of the Dental Cognitions Questionnaire. *Behaviour Research and Therapy, 33,* 507–515.

De Jongh, A., Muris, P., Ter Horst, G., & Duyx, M. (1995b). Acquisition and maintenance of dental anxiety: The role of aversive experiences and cognitive factors. *Behaviour Research and Therapy, 33,* 205–210.

De Jong, P. J., Andrea, H., & Muris, P. (1997). Spider phobia in children: Disgust and fear before and after treatment. *Behaviour Research and Therapy, 35,* 559–562

De Jong, P. J., Merckelbach, H., Arntz, A., & Nijman, H. (1992). Covariation detection in treated and untreated spider phobics. *Journal of Abnormal Psychology, 101,* 724–727.

De Jong, P. J., Van den Hout, M. A., & Merckelbach, H. (1995). Covariation bias and the return of fear. *Behaviour Research and Therapy, 33,* 211–213.

De Ruiter, C., Rijken, H., Garssen, B., Van Schaik, A., & Kraaimaat, F. (1989). Comorbidity among the anxiety disorders. *Journal of Anxiety Disorders, 3,* 57–68.

DiNardo, P. A., Guzy, L. T., & Bak, R. M. (1988). Anxiety response patterns and etiological factors in dog-fearful and non-fearful subjects. *Behaviour Research and Therapy, 26,* 245–252.

Doogan, S., & Thomas, G. V. (1992). Origins of fear of dogs in adults and children: The role of conditioning processes and prior familiarity with dogs. *Behaviour Research and Therapy, 30,* 387–394.

Errera, P. (1962). Some historical aspects of the concept phobia. *Psychiatric Quarterly, 36,* 325–336.

Eysenck, H. J. (1979). The conditioning model of neurosis. *Behavioral and Brain Sciences, 2,* 155–199.

Eysenck, H. J. (1987). Behavior therapy. In H. J. Eysenck & I. Martin (Eds.), *Theoretical foundations of behavior* *therapy* (pp. 3–35). New York: Plenum.

Eysenck, M. W. (1992). *Anxiety: The cognitive perspective*. Hillsdale, NJ: Erlbaum.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99,* 20–35.

Frederikson, M., Annas, P., Fischer, H., & Wik, G. (1996). Gender and age differences in the prevalence of specific fears and phobias. *Behaviour Research and Therapy, 34,* 33–39.

Fyer, A. J., Mannuzza, S., Gallops, M. S., Martin, L. Y., Aaronson, C., Gorman, J. M., Liebowitz, M. R., & Klein, D. F. (1990). Familial transmission of simple phobias and fears. *Archives of General Psychiatry, 47,* 252–256.

Gray, J. A. (1987). Interactions between drugs and behavior therapy. In H. J. Eysenck & I. Martin (Eds.), *Theoretical foundations of behavior therapy* (pp. 433–447). New York: Plenum.

Greenberg, D. B., Stern, T. A., & Weilburg, J. B. (1988). The fear of choking: Three successfully treated cases. *Psychosomatics, 29,* 126–129.

Himle, J. A., Crystal, D., Curtis, G. C., & Fluent, T. E. (1991). Mode of onset of simple phobia subtypes: Further evidence of heterogeneity. *Psychiatry Research, 36,* 37–43.

Hugdahl, K. (1989). Simple phobias. In G. Turpin (Ed.), *Handbook of clinical psychophysiology* (pp. 283–308). Chichester, UK: Wiley.

Hugdahl, K., & Johnsen, B. (1989). Preparedness and electrodermal fear conditioning: Ontogenetic vs. phylogenetic explanations. *Behaviour Research and Therapy, 27,* 269–278.

Hygge, S., & Öhman, A. (1978). Modelling processes in the acquisition of fears: Vicarious electrodermal conditioning to fear-relevant stimuli. *Journal of Personality and Individual Differences, 36,* 271–279.

Kelley, M. J. (1987). Hormones and clinical anxiety: An imbalanced neuromodulation of attention. In H. J. Eysenck & I. Martin (Eds.), *Theoretical foundations of behavior therapy* (pp. 403–432). New York: Plenum

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992). The genetic epidemiology of phobias in women: The interrelationship of agoraphobia, situational phobia, and simple phobia. *Archives of General Psychiatry, 49,* 273–281.

Kessler, R. C., McConagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H. U., & Kendler, K. S. (1994). Lifetime and 12-month prevalence of *DSM-III-R* psychiatric disorders in the United States. *Archives of General Psychiatry, 51,* 8–19.

Kleinknecht, R. A. (1994). Acquisition of blood, injury, and needle fears and phobias. *Behaviour Research and Therapy, 32,* 817–823.

Klorman, R., Weerts, T. C., Hastings, J. E., Melamed, B. G., & Lang, P. J. (1974). Psychometric description of some specific fear questionnaires. *Behavior Therapy, 5,* 401–409.

Kornfeld, A. D. (1989). Mary Cover Jones and the Peter case: Social learning versus conditioning. *Journal of Anxiety Disorders, 3,* 187–195.

Kuch, K., Cox, B. J., Evans, R. E., & Shulman, I. (1994). Phobias, panic, and pain in 55 survivors of road vehicle accidents. *Journal of Anxiety Disorders, 8,* 181–187.

Lang, P. J. (1968). Fear reduction and fear behavior: Problems in treating a construct. In J. M. Schlien (Ed.), *Research in psychotherapy* (Vol. III, pp. 90–103). Washington, DC: American Psychiatric Association.

Lapouse, R., & Monk, M. A. (1959). Fears and worries in a representative sample of children. *American Journal of Orthopsychiatry, 29,* 803.

Last, C. (1987). Simple phobias. In L. Michelson & M. Ascher (Eds.), *Anxiety and stress disorders* (pp. 179–189). New York: Guilford Press.

Lautch, H. (1971). Dental phobia. *British Journal of Psychiatry, 119,* 151–158.

LeDoux, J. (1992). Emotion as memory: Anatomical systems underlying indelible neural traces. In S. A. Christianson (Ed.), *The handbook of emotion and memory: Research and theory* (pp. 269–288). Hillsdale, NJ: Erlbaum.

MacFarlane, J. W., Allen, L., & Honzik, M. P. (1954). *A developmental study of the behavior problems of normal children between 21 months and 14 years*. Berkeley, CA: University of California Press.

Malloy, P., & Levis, D. J. (1988). A laboratory demonstration of persistent human avoidance. *Behavior Therapy, 19,* 229–241.

Marks, I. (1987). *Fears, phobias, and rituals*. Oxford, UK: Oxford University Press.

Marquis, J. N. (1991). A report of seventy-eight cases treated by eye movement desensitization. *Journal of Behavior Therapy and Experimental Psychiatry, 22,* 187–192.

Matchett, G., & Davey, G. C. L. (1991). A test of a disease avoidance model of animal phobias. *Behaviour Research and Therapy, 29,* 91–94.

Menzies, R. G., & Clarke, J. C. (1993). The etiology of fear of heights and its relationship to severity and individual response patterns. *Behaviour Research and Therapy, 31,* 355–366.

Menzies, R. G., & Clarke, J. C. (1995). The etiology of phobias: A nonassociative account. *Clinical Psychology Review, 15,* 23–48.

Merckelbach, H., Arntz, A., & De Jong, P. J. (1991). Conditioning experiences in spider phobics. *Behaviour Research and Therapy, 29,* 333–335.

Merckelbach, H., Arrindell, W. A., Arntz, A., & De Jong, P. J. (1992). Pathways to spider phobia. *Behaviour Research and Therapy, 30,* 543–546.

Merckelbach, H., De Jong, P. J., Arntz, A., & Schouten, E. (1993). The role of evaluative learning and disgust sensitivity in the etiology and treatment of spider phobia. *Advances in Behaviour Research and Therapy, 15,* 243–255.

Merckelbach, H., Muris, P., & Schouten, E. (1996). Pathways to fear in spider phobic children. *Behaviour Research and Therapy, 34,* 935–938.

Merluzzi, T. V., Taylor, C. B., Boltwood, M., & Götestam, K. G. (1991). Opioid antagonist impedes exposure. *Journal of Consulting and Clinical Psychology, 59,* 425–430.

Mineka, S., & Sutton, S. K. (1992). Cognitive biases and emotional disorders. *Psychological Science, 3,* 65–69.

Mowrer, O. (1960). *Learning theory and behaviour*. New York: Wiley.

Muris, P., De Jongh, A., Merckelbach, H., Postema, S., & Vet, M. (in press). Thought suppression in phobic and non-phobic dental patients. *Anxiety, Stress, and Coping*.

Muris, P., & Merckelbach, H. (in press). Treating spider phobics with eye movement desensitization and reprocessing: A controlled study. *Behavioural and Cognitive Psychotherapy, 25,* 39–50.

Muris, P., Merckelbach, H., Meesters, C., & Van Lier, P. (1997). What do children fear most often? *Journal of Behavior Therapy and Experimental Psychiatry, 28,* 263–267.

Muris, P., Merckelbach, H., Van Haaften, H., & Mayer, B. (1997). Eye movement desensitization and reprocessing versus exposure *in vivo*. A single-session crossover study of spider phobic children. *British Journal of Psychiatry, 171,* 82–86.

Muris, P., Steerneman, P., Merckelbach, H., & Meesters, C. (1996). Parental modelling and fearfulness in middle childhood. *Behaviour Research and Therapy, 34,* 265–268.

Öhman, A. (1987). Evolution, learning, and phobias. In D. Magnusson & A. Öhman (Eds.), *Psychopathology: An interactional perspective* (pp. 143–158). New York: Academic Press.

Öhman, A. (1993). Fear and anxiety as emotional phenomena: Clinical phenomenology, evolutionary perspectives, and information processing mechanisms. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 511–536). New York: Guilford Press.

Öhman, A., & Soares, J. J. F. (1994). On the automaticity of phobic fear: Conditioned electrodermal responses to masked fear-relevant stimuli. *Journal of Abnormal Psychology, 103,* 231–240.

Ollendick, T. H. (1979). Fear reduction techniques with children. In P. Miller & R. Eisler (Eds.), *Progress in behavior modification* (Vol. 8, pp. 127–168). New York: Academic Press.

Ollendick, T. H., & King, N. J. (1991). Origins of childhood fears: An evaluation of Rachman's theory of fear acquisition. *Behaviour Research and Therapy, 29,* 117–123.

Ollendick, T. H., Yule, W., & Ollier, K. (1991). Fears in British children and their relationship to manifest anxiety and depression. *Journal of Child Psychology and Psychiatry, 32,* 321–331.

Öst, L. G. (1987). Age of onset in different phobias. *Journal of Abnormal Psychology, 96,* 223–229.

Öst, L. G. (1989). One-session treatment for specific phobias. *Behaviour Research and Therapy, 27,* 1–8.

Öst, L. G. (1996). One-session group treatment of spider phobia. *Behaviour Research and Therapy, 34,* 707–715.

Öst, L. G., Fellenius, J., & Sterner, U. (1991a). Applied tension, exposure *in vivo*, and tension-only in the treatment of blood phobia. *Behaviour Research and Therapy, 29,* 561–575.

Öst, L. G., & Hugdahl, K. (1981). Acquisition of phobias and anxiety response patterns in clinical patients. *Behaviour Research and Therapy, 19,* 439–447.

Öst, L. G., Salkovskis, P. M., & Hellström, K. (1991b). One-session therapist directed exposure vs. self-exposure in the treatment of spider phobia. *Behavior Therapy, 22,* 407–422.

Page, A. C. (1994). Blood-injury phobia. *Clinical Psychology Review, 14,* 443–461.

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention grabbing power of negative social information. *Journal of Personality and Social Psychology, 61,* 380–391.

Rachman, S. (1977). The conditioning theory of fear acquisition: A critical examination. *Behaviour Research and Therapy, 15,* 375–387.

Rachman, S. (1990). The determinants and treatment of simple phobias. *Advances in Behaviour Research and Therapy, 12,* 1–30.

Rachman, S. (1991). Neoconditioning and the classical theory of fear acquistion. *Clinical Psychology Review, 11,* 155–173.

Rescorla, R. A. (1988). Pavlovian conditioning: It is not what you think it is. *American Psychologist, 43,* 151–160.

Rickels, K. (1978). Use of anti-anxiety agents in anxious outpatients. *Psychopharmacology, 58,* 1–17.

Saigh, P. A. (1984). Pre- and post-invasion anxiety in the Lebanon. *Behavior Therapy, 15,* 185–190.

Salkovskis, P. M. (1989). Cognitive-behavioural factors and the persistence of intrusive thoughts in obsessional problems. *Behaviour Research and Therapy, 27,* 677–682.

Sartory, G., MacDonald, R., & Gray, J. A. (1990). Effects of diazepam on approach, self-reported fear and psychophysiological responses in snake phobics. *Behaviour Research and Therapy, 28,* 273–282.

Seligman, M. E. P. (1971). Phobias and preparedness. *Behavior Therapy, 2,* 307–320.

Shafran, R., Booth, R., & Rachman, S. (1993). The reduction of claustrophobia: Cognitive analyses. *Behaviour Research and Therapy, 31,* 75–85.

Shapiro, F. (1989). Efficacy of the eye movement desensitization procedure in the treatment of traumatic memories. *Journal of Traumatic Stress, 2,* 199–223.

Shapiro, F. (1995). *Eye movement desensitization and reprocessing. Basic principles, protocols, and procedures.* New York: Guilford.

Stevenson, J., Batten, N., & Cherner, M. (1992). Fears and fearfulness in children and adolescents: A genetic analysis of twin data. *Journal of Child Psychology and Psychiatry, 33,* 977–985.

Sturgis, E. T., & Scott, R. (1984). Simple phobia. In S. M. Turner (Ed.), *Behavioral theories and treatment of anxiety* (pp. 91–141). New York: Plenum.

Thorpe, S. J., & Salkovskis, P. M. (1995). Phobic beliefs: Do cognitive factors play a role in specific phobias? *Behaviour Research and Therapy, 33,* 805–816.

Thyer, B. A., Himle, J., & Curtis, G. C. (1985). Blood-injury-illness phobia: A review. *Journal of Clinical Psychology, 41,* 451–459.

Tomarken, A. J., Mineka, S., & Cook, M. (1989). Fear relevant selective associations and covariation bias. *Journal of Abnormal Psychology, 98,* 381–394.

Van den Hout, M. A., & Merckelbach, H. (1991). Classical conditioning: Still going strong. *Behavioural Psychotherapy, 19,* 59–79.

Van der Molen, G. M., Merckelbach, H., & Van den Hout, M. A. (1988). The possible relation of the menstrual cycle to susceptibility to fear acquisition. *Journal of Behavior Therapy and Experimental Psychiatry, 19,* 127–133.

Watson, J., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3,* 1–22.

Watts, F. N. (1979). Habituation model of systematic desensitization. *Psychological Review, 86,* 627–637.

Watts, F. N., McKenna, F. P., Sharrock, R., & Trezise, L. (1986). Colour naming of phobia-related words. *British Journal of Psychology, 77,* 97–108.

Webb, K., & Davey, G. C. L. (1992). Disgust sensitivity and fear of animals: Effect of exposure to violent or revulsive material. *Anxiety Stress and Coping, 5,* 329–335.

Wegner, D. M. (1989). *White bears and other unwanted thoughts. Suppression, obsession, and the psychology of mental control.* New York: Viking.

Wegner, D. M., Schneider, D. J., Carter, S. R., & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology, 53,* 5–13.

Westefeld, J. S. (1996). Severe weather phobia: An exploratory study. *Journal of Clinical Psychology, 52,* 509–515.

White, K., & Davey, G. C. L. (1989). Sensory preconditioning and UCS inflation in human fear conditioning. *Behaviour Research and Therapy, 27,* 161–166.

Wittchen, H. (1988). Natural course and spontaneous remissions of untreated anxiety disorders. In I. Hand & H. Wittchen (Eds.), *Panic and phobias: Treatment and variables affecting course and outcome* (pp. 3–17). Berlin, Germany: Springer.

Wolpe, J., Lande, S. D., McNally, R. J., & Schotte, D. (1985). Differentiation between classically conditioned and cognitively based neurotic fears: Two pilot studies. *Journal of Behavior Therapy and Experimental Psychiatry, 16,* 287–293.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.21
Social Phobia

HARLAN R. JUSTER
Pine Bush Mental Health, Albany, NY, USA
and
RICHARD G. HEIMBERG
Temple University, Philadelphia, PA, USA

6.21.1 INTRODUCTION 475
6.21.1.1 Definition of Social Phobia 476
6.21.1.2 Description of Social Phobia 477
6.21.1.2.1 Prevalence 477
6.21.1.2.2 Gender ratio 477
6.21.1.2.3 Age of onset 478
6.21.1.2.4 Chronicity 478
6.21.1.3 Functional Impairment 479
6.21.1.4 Comorbidity 479
6.21.1.5 Social Phobia and Avoidant Personality Disorder 480
6.21.1.6 Summary 480
6.21.2 UNDERSTANDING THE ETIOLOGY AND MAINTENANCE OF SOCIAL PHOBIA 481
6.21.3 TREATMENT STRATEGIES FOR SOCIAL PHOBIA 481
6.21.3.1 Social Skills Training 481
6.21.3.1.1 Review of empirical studies of SST 482
6.21.3.2 Relaxation Techniques 484
6.21.3.2.1 Review of empirical studies of relaxation techniques 484
6.21.3.3 Exposure-based Techniques 485
6.21.3.3.1 Review of empirical studies of exposure-based techniques 485
6.21.3.4 Cognitive-behavioral Techniques 486
6.21.3.4.1 Review of empirical studies of cognitive-behavioral techniques 486
6.21.3.5 Psychodynamic Formulations 492
6.21.4 PREDICTORS OF OUTCOME 492
6.21.5 SUMMARY AND FUTURE DIRECTIONS 493
6.21.6 REFERENCES 495

6.21.1 INTRODUCTION

Social anxiety is a common experience, and most individuals can recall some event in their lives when they experienced concern about being evaluated by others. Anyone who has interviewed for a job, given a presentation, or gone on a first date with someone to whom they were attracted has experienced the concerns associated with coming under the scrutiny of others. When anxiety in these situations becomes severe and interferes with the adaptive functioning of the individual, the diagnosis of social phobia may be assigned. The following are examples:

> Mr S., a man in his mid-20s, reports being anxious when meeting new people. Nevertheless, he managed to graduate from high school and secure a job in sales which appears to cause him little difficulty. He has a small circle of friends from his adolescent years but is very uncomfortable when meeting anyone for the first time. He actively avoids socializing in situations which may involve "too many new people." He managed to meet and marry a woman two years ago, but she is frustrated by his reluctance to meet new people. Mr. S. copes, in part, by smoking marijuana with his small circle of friends. He has also recently begun to use marijuana on his own outside of social situations. He is not happy about this behavior but reports that it is the only thing that calms him down after a long day of interacting with others.
>
> Ms. B. is a single woman in her late 30s who currently works as a reference librarian. She has dated only occasionally, and her social life is a source of much distress. She dreams of having a family but doubts her ability to achieve this dream. She has become more depressed, beginning to feel hopeless about her future prospects. To complicate her situation, the library is beginning a new outreach program and she will be expected to speak with students in the local schools, something she dreads. She seeks treatment now for what she views as her last opportunity to turn her life around.

As these examples illustrate, social phobia can be a debilitating disorder, causing considerable impairment in occupational, educational, and social functioning. Individuals with social phobia often have difficulty making friends, finding partners, and remaining socially connected to others. They may make important career decisions which accommodate their social anxiety even if these decisions make it unlikely that they will achieve their personal goals. For example, they may opt for college courses which require no public speaking, choose vocations because they present few demands for social interaction, or simply opt out of employment altogether. Some may settle for an easily available romantic partner regardless of level of attraction. Understanding social phobia and its concomitants is a crucial task, especially since social phobia may be much more common than previously believed. The recognition of social phobia and knowledge of effective treatment options are important considerations for the clinical psychologist.

6.21.1.1 Definition of Social Phobia

Our understanding of social phobia has evolved considerably since the 1960s, and it is instructive to examine that evolution in some detail. One of the first descriptions of social phobia was provided by Marks and Gelder (1966). Social phobia was described as the experience of anxiety in situations where an individual might perform some task under the scrutiny of others. These "performance" tasks included public speaking, writing or eating in front of others, and using public restrooms while others were present. Fears expressed by the individual with social phobia in these situations focused primarily on one of two concerns: that the individual would do or say something embarrassing or that symptoms of anxiety would be visible to others. When these performance tasks could adequately be performed in the absence of others, social phobia was said to be present.

This early conceptualization provided the framework for the diagnostic criteria for social phobia in the third edition of the *Diagnostic and statistical manual of mental disorders* (*DSM-III*; American Psychiatric Association [APA], 1980). Social phobia was defined by the experience of fear in and a compelling desire to avoid one of several specific situations. Notably absent from this early view was anxiety associated with interpersonal interaction. Interpersonal anxiety was not considered to be part of the social phobia "picture," and individuals with this type of anxiety were instead given the Axis II diagnosis of avoidant personality disorder (APD). This, however, would change with the next revision of the diagnostic criteria.

As more was learned about individuals with social phobia, it became apparent that many individuals who experienced anxiety in performance situations also feared interacting with others. Thus, when the revision of *DSM-III* (*DSM-III-R*; APA, 1987) was published, the scope of the diagnosis was broadened to include social interactional fears. Several additional findings and concerns supported this decision, including research documenting that most patients with social phobia reported multiple fears (Turner, Beidel, Dancu, & Keys, 1986), the observation that clinicians often applied the diagnosis of social phobia to individuals with multiple social fears (Spitzer & Williams, 1985), and the concern that classifying interpersonal anxiety as an Axis II disorder might discourage research into its treatment (Liebowitz, Gorman, Fyer, & Klein, 1985) (see Heimberg, 1996, for a more detailed discussion of these issues).

With the broadening of the diagnostic criteria, a generalized subtype was added to describe those individuals with social phobia who feared most social interactions. Avoidance, while common, was not essential for the diagnosis in recognition of the fact that some individuals with social phobia often endure the personal distress associated with feared situations. Social phobia, previously viewed as rare and rarely debilitating, was described as much more common and associated with potentially severe impairment.

The definition of social phobia remains essentially unchanged in the *DSM-IV* (APA, 1994). Our knowledge of social phobia has matured as further research has been conducted and our knowledge base has expanded. In fact, two important volumes have been published surveying the current "state of our knowledge" regarding this disorder (Heimberg, Liebowitz, Hope, & Schneier, 1995; Stein, 1995). What was once described as a "neglected anxiety disorder" (Liebowitz et al., 1985) is now much better understood.

6.21.1.2 Description of Social Phobia

6.21.1.2.1 Prevalence

In the National Institute of Mental Health Epidemiological Catchment Area (ECA) studies, the six-month prevalence for *DSM-III* social phobia was reported to be 1.5–2.6% for women and 0.9–1.7% for men (Myers et al., 1984). The lifetime prevalence of social phobia at four ECA sites was 2.4% for men and women combined (Schneier, Johnson, Hornig, Liebowitz, & Weissman, 1992). More recently, the National Comorbidity Survey (NCS; Kessler et al., 1994) was conducted in the United States with a national probability sample of more than 8000 participants between the ages of 15 and 54 years. The 12-month prevalence of *DSM-III-R* social phobia was reported to be 7.9% (6.6% for men and 9.1% for women), and the lifetime prevalence of social phobia was reported to be 13.3% (11.1% for men and 15.5% for women). The high lifetime prevalence of social phobia makes it the third most common mental disorder examined in the NCS after major depression (17.1%) and alcohol dependence (14.1%).

Similar prevalence rates for *DSM-III-R* social phobia were reported in two other studies conducted outside the United States. A study of 470 Swiss adults revealed a lifetime prevalence rate of 16% for *DSM-III-R* social phobia, similar to the rate reported in the NCS (Wacker, Mullejans, Klein, & Battegay, 1992). In Winnepeg, Manitoba, a telephone survey examined the nature of social anxiety in the community as well as the prevalence of *DSM-III-R* social phobia (Stein, Walker, & Forde, 1994). Of 526 respondents, 7.1% met criteria for the diagnosis of social phobia at the time of assessment, similar to the 12-month NCS prevalence rate.

The higher prevalence rates reported in these studies compared with the much lower rates reported in the ECA studies may arise from multiple sources. For example, there were a number of methodological differences between the ECA and NCS studies that might account for differences in reported prevalence. Specifically, in the ECA studies, interviewers inquired about the individual's anxiety in only three situations (eating in front of others, public speaking, and speaking to strangers or meeting new people). The NCS interview inquired about fear in these three situations, but also queried three additional situations: using toilets away from home, writing while being watched, and sounding foolish in front of others. Since both studies required endorsement of fear and interference in only one situation in order to receive a diagnosis of social phobia, the odds of a positive diagnosis were higher in the NCS than the ECA. Furthermore, the probe regarding "sounding foolish in front of others" reflects the broadening of the social phobia diagnosis to include a greater focus on anxiety in social interaction, not reflected in *DSM-III.* Clearly, the increase in the number of situational probes, as well as the broadening of the scope of inquiry to include more social interaction situations, may identify more individuals with social phobia.

Methodological differences and criterion changes may not be the only reasons that prevalence rates differ between studies. Magee, Eaton, Wittchen, McGonagle, and Kessler (1996) report that the prevalence of social phobia may, in fact, be increasing. Analysis of the NCS data demonstrated the existence of cohort effects, such that younger cohorts of subjects (ages 15–24) showed higher prevalence rates than older cohorts (ages 25 and up). These findings were most pronounced for agoraphobia, but were also noted for simple and social phobia. This trend towards increasing prevalence rates in younger cohorts has also been noted for major depression in several studies around the world (Cross-National Collaborative Group, 1992; Klerman & Weissman, 1989; Wickramaratne, Weissman, Leaf, & Holford, 1989).

6.21.1.2.2 Gender ratio

Epidemiological data suggest that more women than men suffer from social phobia. For example, across four ECA sites, approximately 70% of individuals with social phobia were women (Schneier et al., 1992). In the NCS, women with social phobia outnumbered men with social phobia about 1.4 to 1. This gender difference in prevalence is not apparent in treatment settings, however, where men and women with social phobia present in approximately equal numbers (Mannuzza, Fyer, Liebowitz, & Klein, 1990). This may be related to different sociocultural expectations for men and women. For example, women perceived as shy may fit stereotypical images of femininity and experience less demand to overcome their social anxiety. Conversely, men may experience more pressure to alter their social anxiety and avoidance which are less consistent with stereotypical conceptions of masculinity.

The study of gender differences was extended to examine whether there were differences between men and women with social phobia in the severity of their anxiety or in the nature of the situations in which they reported clinically significant anxiety (Turk, Heimberg et al., in press). Women exhibited more severe social fears on several measures. They also reported significantly greater fear than men while talking to authority, acting/performing/giving a talk in front of an audience, working while being observed, entering a room when others are already seated, being the center of attention, speaking up at a meeting, expressing disagreement or disapproval to people they don't know very well, giving a report to a group, and giving a party. Men reported significantly more fear than women regarding urinating in public bathrooms and returning goods to a store. Not surprisingly, more men than women with social phobia reported full-time employment.

6.21.1.2.3 Age of onset

Hazen and Stein (1995) reviewed 15 epidemiological and clinical studies and found the mean age of onset of social phobia to range between 13 and 20 years. However, age of onset may differ by subtype of social phobia. For example, Holt, Heimberg, and Hope (1992) found that the age of onset for a sample of subjects with generalized social phobia was about 13 years of age, whereas the age of onset for subjects with nongeneralized social phobia was 22.6 years. This finding was later replicated by Brown, Heimberg, and Juster (1995). To further complicate this picture, data from the NIMH ECA studies reported by Schneier et al. (1992) suggest that the distribution of age of onset of social phobia may be bimodal with one peak at 13 years of age and an earlier peak at under five years of age. This pattern suggests that there may be more than one pathway to the development of social phobia.

6.21.1.2.4 Chronicity

Although no true "natural history" studies of social phobia have been conducted, several lines of research converge to support the perspective that social phobia is a chronic disorder. In one such study, Reich, Goldenberg, Vasile, Goisman, and Keller (1994) followed 140 subjects with social phobia for over one year. Seventy-four subjects were diagnosed with generalized social phobia and the remainder (66) with discrete social phobia (fear of one or more specific performance situations). After 65 weeks, only 11% of the sample achieved full remission despite the fact that 88% received some form of psychotherapy or pharmacotherapy during this period. Importantly, there were no differences in chronicity between subjects with generalized vs. discrete social phobia, suggesting a similar course for these two variants of the disorder.

Studies on shyness, a construct closely related to social phobia, may shed additional light on the issue of chronicity. Shyness, which is not a diagnostic entity, has never been defined consistently by researchers. However, every definition of shyness has included the concept of anxiety regarding social interaction, suggesting that shyness and social phobia are closely allied (Turner, Beidel, & Townsley, 1990). Research reviewed by Turner et al. (1990) examining social phobia and shyness further suggests that the two conditions, while not identical, broadly overlap on behavioral, cognitive, and physiological indices. Rapee (1995) has suggested that the two conditions differ more quantitatively than qualitatively, especially on indices of impairment and avoidance, with social phobia being the more extreme condition. Thus, research on shyness in children and its influence on later levels of functioning may reveal something about the chronicity of social anxiety. One prospective study found that boys identified as shy were likely to marry and achieve occupational stability later in life than nonshy boys (Caspi, Elder, & Bem, 1988). Childhood shyness also influenced girl's later development. Compared with nonshy girls, shy girls were more likely to become housewives or to remain at home after childbirth. Thus, early experiences with social anxiety, specifically shyness in children, were predictive of adult behavior which could be broadly construed as the demonstration of social impairment or avoidance.

Finally, "behavioral inhibition," or the tendency to withdraw in the face of novel people, places, and things, has been identified in very young children and may persist for years (Kagan, Reznick, & Snidman, 1988). Kagan et al. (1988) report that behavioral inhibition occurs in approximately 10–15% of children. A majority of children classified as behaviorally inhibited prior to age three remained inhibited for at least 10 years (Rosenbaum, Biederman, Pollock, & Hirshfeld, 1994). Furthermore, behaviorally inhibited children were more introverted and more physiologically reactive to stress than uninhibited children. More revealing, however, is that the presence of behavioral inhibition in children was associated with a greater occurrence of childhood anxiety disorders (Biederman et al., 1990) and a higher likelihood of anxiety disorders, especially social phobia, in the parents of these children (Rosenbaum et al., 1991).

6.21.1.3 Functional Impairment

Social phobia is often associated with significant impairment in occupational and educational attainment as well as social functioning (Amies, Gelder, & Shaw, 1983; Liebowitz et al., 1985; Magee et al., 1996; Schneier et al., 1992, 1994; Turner et al., 1986). For example, Schneier et al. (1992) re-evaluated the data of subjects with social phobia from four of the five ECA study sites. Subjects with social phobia were less educated, of lower socioeconomic status, and less likely to be married when compared to subjects with other disorders (but not social phobia) and to nondisordered subjects. Another study found that subjects with social phobia were less likely to be married than subjects with agoraphobia, although they also were likely to be younger at the time of referral than the agoraphobic subjects (Amies et al., 1983). Both of these studies examined social phobics diagnosed according to *DSM-III*. However, similar impairment is observed whether *DSM-III* or *DSM-III-R* criteria for social phobia are employed. Magee et al. (1996) reported that *DSM-III-R* social phobia was associated with less educational attainment, lower income, reduced likelihood of being married or employed, and a greater likelihood of living with one's parents. Schneier et al. (1994) compared the self-rated impairment of 32 *DSM-III-R* social phobic subjects who were not currently depressed and a group of normal control subjects. The social phobic subjects reported significantly higher levels of impairment in every category of social functioning (family relations, romantic relationships, and social network) and occupational functioning (education and employment), as well as alcohol and drug use, mood regulation, and activities of daily living. Finally, individuals with social phobia report less satisfaction with their lives than persons without social phobia (Safren, Heimberg, Brown, & Holle, 1997).

6.21.1.4 Comorbidity

Comorbidity is high among individuals with social phobia. Schneier et al. (1992) reported that among ECA subjects with *DSM-III* social phobia, 69% had at least one comorbid disorder (excluding dysthymia and simple phobia). When comparing rates of comorbidity between subjects with social phobia and subjects with other disorders but not social phobia and subjects with no disorders (the latter two groups being combined into a no social phobia group), elevated rates of comorbidity in the social phobia group were found for all disorders studied. For example, simple phobia and agoraphobia were the most commonly occurring comorbid diagnoses in the social phobia group with lifetime rates of 59% and 45%, respectively, compared to rates of 13.2% and 6.1% in the no social phobia group. Other disorders with elevated comorbidity rates were panic disorder (4.7% vs. 1.3%), major depression (16.6% vs. 4.0%), obsessive-compulsive disorder (11.1% vs. 2.5%), somatization disorder (1.9% vs. 0.2%), dysthymia (12.5% vs. 3.1%), and bipolar disorder (4.7% vs. 1.1). When a comorbid disorder was diagnosed, social phobia was the temporally primary disorder 77% of the time.

In the NCS study, 81% of persons with *DSM-III-R* social phobia had at least one other disorder (Magee et al., 1996). Again, social phobia was highly comorbid with simple phobia (37.6%), agoraphobia (23.3%), and major depression (37.2%). Substantial rates of comorbidity were evident for any affective disorder (41.4%) and any anxiety disorder (56.9%). Again, social phobia was the temporally primary disorder in the majority of cases.

Substance-related disorders also commonly co-occur with social phobia. In the ECA studies, 18.8% of social phobics were comorbid for alcohol abuse and 13% for drug abuse (Schneier et al., 1992). In a subset of subjects from the ECA studies, social phobia was highly comorbid with alcohol-related disorders (Himle & Hill, 1991). Nearly 28% of social phobics also had an alcohol-related disorder compared with 13–24% for nonsocial phobics. In the NCS, these figures are even higher. Alcohol abuse and alcohol dependence were diagnosed in over 30% of *DSM-III-R* social phobics, whereas drug abuse and drug dependence were diagnosed in over 20% of social phobics (Magee et al., 1996). Nearly 40% of *DSM-III-R* social phobics had a comorbid substance-related disorder.

While alcohol-use disorders are more common in social phobics than in the community at large, social phobia is also more common among alcoholics than in the community. Estimates of the prevalence of social phobia in subjects diagnosed with alcohol-related disorders run as high as 40% (Ross, Glaser, & Germanson, 1988; Smail, Stockwell, Canter, & Hodgson, 1984). Kushner, Sher, and Beitman (1990) reported that social phobia was nine times more common among patients with alcohol disorders than in a community sample.

Data provided by Magee et al. (1996) and Schneier et al. (1992) are remarkably consistent despite the differences between the ECA and NCS studies. Comorbidity of social phobia with other anxiety disorders, affective disorders, and substance-related disorders is common. In the majority of cases, social phobia is the temporally primary condition, its onset preceding the onset of comorbid conditions.

6.21.1.5 Social Phobia and Avoidant Personality Disorder

Early on, APD and social phobia were defined as relatively distinct diagnostic entities. Social phobia referred to fear in circumscribed performance situations, whereas APD referred to more broad-based fear of social interaction. With the publication of *DSM-III-R*, changes in the diagnostic criteria were introduced which blurred the distinction between these two disorders. For example, the generalized subtype of social phobia, added in *DSM-III-R*, referred to fear in most social situations, a profile which previously would have excluded a diagnosis of social phobia. Also, the criteria for APD were modified such that there was now a greater emphasis on fear of negative evaluation and social discomfort, a central feature of social phobia. Whereas in *DSM-III*, five of five criteria had to be met for the diagnosis of APD, only four of seven were required in *DSM-III-R*, making it easier for the diagnosis to be awarded. This situation remains essentially unchanged in *DSM-IV* (see Heimberg, 1996, for a fuller discussion of the issues related to generalized social phobia and APD).

A number of articles have addressed the degree of overlap between generalized social phobia with and without APD (Brown et al., 1995; Herbert, Hope, & Bellack, 1992; Holt et al., 1992; Schneier, Spitzer, Gibbon, Fyer, & Liebowitz, 1991; Tran & Chambless, 1995; Turner, Beidel, & Townsley, 1992). Despite differences in sampling procedures, subtype definition, and diagnostic instruments, the degree of overlap was high, ranging from 25 to 89% in these six studies. Widiger (1992) concluded that generalized social phobia and APD were more alike than not and that the combination of generalized social phobia and APD probably represents a more severe manifestation of social phobia.

Differential response to treatment by generalized social phobics with and without APD might highlight distinctions between these two disorders, and the effects of treatment designed specifically for social phobia on the diagnosis of APD might further clarify this relationship. Brown et al. (1995) examined these issues in the context of cognitive-behavioral group treatment for social phobia. On the majority of outcome measures, there were no post-treatment differences between generalized social phobics with or without APD. On the remaining measures, social phobics with APD remained more severely impaired. Importantly, 47% (8 of 17) of the patients diagnosed with APD prior to treatment no longer met APD criteria following treatment. Hope, Herbert, and White (1995b) reported that APD did not differentially predict outcome of cognitive-behavioral group treatment for social phobia. However, in this study, subtype of social phobia and APD diagnosis were not considered simultaneously. Feske, Perry, Chambless, Renneberg, and Goldstein (1996) treated generalized social phobics with and without APD with a combination of behavioral techniques (exposure, relaxation, social skills training, and systematic desensitization). Although both groups significantly improved, generalized social phobics with APD began and ended treatment more impaired than social phobics without APD. This pattern of results was also evident at three-month follow-up.

These results lend further support to Widiger's (1992) conclusion noted above, that generalized social phobia and APD differ more quantitatively than qualitatively. In his review, Heimberg (1996) takes this argument a step further by questioning the appropriateness of separating these two diagnoses on to Axis I and Axis II. Most importantly for the clinical psychologist who may be treating generalized social phobics with and without APD is that response to treatment for these subjects appears to be similar.

6.21.1.6 Summary

Social phobia, a chronic anxiety disorder, is highly prevalent in both men and women, and is associated with significant functional impairment and personal distress. Although the average age of onset is in the early teens, some individuals appear to suffer with the disorder from a much earlier age. Comorbidity with other conditions is the rule although social phobia usually begins first. We turn our attention now to possible explanations of the development of social phobia.

6.21.2 UNDERSTANDING THE ETIOLOGY AND MAINTENANCE OF SOCIAL PHOBIA

As knowledge of social phobia has increased, researchers have developed models of the causes and maintaining factors of social anxiety and social phobia (e.g., Clark & Wells, 1995; Heimberg, Juster, Hope, & Mattia, 1995; Leary & Kowalski, 1995; Rapee & Heimberg, 1997). An integrative model of the etiology of social phobia was put forth by Heimberg et al. (1995). Drawing from a broad range of empirical and theoretical literature, they argue for a diathesis-stress conceptualization in which the principal etiological mechanism is an interaction between biological predisposition (Fyer, Mannuzza, Chapman, Liebowitz, & Klein, 1993) and environmental factors. Results from twin studies and family studies combine to support the view that a tendency to interpret situations as potentially threatening may be partially inherited. This general tendency to become anxious and interpret situations as potentially threatening may represent a general response pattern that becomes more specifically focused on social stimuli as a result of environmental factors (Bruch & Heimberg, 1994; Bruch, Heimberg, Berger, & Collins, 1989). Such environmental factors may include a parent's anxiety about social interaction which is communicated to the child directly or indirectly. Parents may also emphasize the opinions of others as excessively important, withhold approval for anything less than perfect performance, and/or isolate their child from social interaction. Early negative peer group interactions may further sensitize individuals to the "potential danger" of social interaction.

Repeated exposure to early sensitizing experiences may contribute to the development of cognitive structures (i.e., schemata) which define other people as inherently critical and social situations as inherently dangerous and threatening to one's self-esteem. Furthermore, the individual may come to believe that only perfect performance would allow him/her to avoid the danger of criticism and negative evaluation by others but that performance up to that standard is highly unlikely (Juster, Heimberg, Frost, et al., 1996). These cognitive structures are postulated to be an underlying, maintaining mechanism for the experience of anxiety in social situations.

Rapee and Heimberg (1997) have delineated a theoretical model of the experience of anxiety when a person with social phobia anticipates or enters a social situation. As noted above, individuals with social phobia make assumptions about the nature of social situations which provide the foundation upon which the experience of social anxiety is built. These assumptions compel them to monitor their performance and others' evaluation of them at all times. Thus, when individuals with social phobia enter social situations, they form mental representations of themselves as they believe they appear to their audience. Simultaneously, they compare their own perception of their performance in this situation with the standard of performance they believe to be held for them by their audience. While attention is given to this internal representation, individuals also attend to the environment, searching for potential threat as well as internal cues (e.g., tachycardia) which may suggest inadequate performance. Negative signs from without or within may then serve as indicators that negative evaluation is likely, and anxiety ensues. The entire experience has further negative impact as it may serve to strengthen maladaptive beliefs and lower expectations for success in similar future situations.

6.21.3 TREATMENT STRATEGIES FOR SOCIAL PHOBIA

Several comprehensive reviews of treatment strategies for social phobia have been written (Heimberg, 1989; Heimberg & Barlow, 1988; Heimberg & Juster, 1995; Hope, Holt, & Heimberg, 1993; Juster, Heimberg, & Holt, 1996) including one which focused on long-term outcome (Juster & Heimberg, 1995). The paper by Heimberg (1989) reviewed just 17 studies. In contrast, the current chapter reviews 37 studies, five of which focus specifically on long-term outcome. The field of social phobia treatment is rapidly growing in size and sophistication.

6.21.3.1 Social Skills Training

Early attempts to treat individuals with social phobia were based on the premise that the person's social anxiety was a result of a deficit in social skills, that is, a lack of knowledge of how to perform or capacity to perform in social situations. Although a logical possibility, there is little empirical support for this notion (Juster & Heimberg, 1995). It is clear that individuals with social phobia exhibit deficits in their social performance; however, this performance deficit has never been linked to a singular cause (i.e., social skills deficit). Possible causes of performance deficits include inhibition of behavior by anxiety or maladaptive beliefs, avoidance of certain situations due to anticipatory anxiety, or a social skills deficit, among others. Our clinical experience suggests that the first two are more likely explanations for the performance deficits seen in social phobic individuals. In clinical settings, social phobics repeatedly perform feared activities as part of exposure exercises during treatment or as part of behavioral assessment tasks, and their performance during these tasks is typically "within normal limits." Despite usually adequate performance, the individual's self-report is almost always marked by negative self-bias (Rapee & Lim, 1992). Although social phobics may report that they do not know what to say or how to act in social situations, their actual behavior often suggests otherwise.

Despite the lack of support for social skills deficits as an overall explanatory mechanism for social phobia, it clearly remains a possibility for some individuals. Whether or not an individual lacks social skills, social skills training (SST) may provide some benefit by providing exposure to anxiety-evoking stimuli, providing information about one's ability to perform in social situations, etc. Social skills training is an amalgam of behavioral techniques designed to give patients experience with and exposure to various types of behavioral practice. Techniques such as therapist modeling, behavioral rehearsal, corrective feedback, social reinforcement, and homework assignments may all be components of social skills training interventions.

6.21.3.1.1 Review of empirical studies of SST

We now turn to the review of treatment strategies designed specifically for the remediation of social skills deficits. Marzillier, Lambert, and Kellet (1976) compared SST with systematic desensitization (SD; the combination of progressive muscle relaxation and the presentation of scenes of gradually increasing anxiety-evoking potential to imagination) in the treatment of 20 subjects who, by description, met criteria for generalized social phobia. This study predated the inclusion of social phobia in the official diagnostic nomenclature, but it appears that subjects would have met criteria for the current conceptualization of the disorder. Subjects were assigned randomly to one of the two treatment conditions (or a wait-list control) and received 15 weekly, 45-minute sessions of SST or SD.

Subjects treated with SST increased their range of social activities, while subjects in both treatment conditions increased social contacts in comparison to subjects on the waiting-list. Between-groups analysis during a six-month follow-up was precluded due to a high number of drop-outs, although it appears that SST-treated subjects maintained their gains. Neither treatment produced significant improvement in social skills or reduction in anxiety. Furthermore, with the exception of those differences noted above, improvements found in the treated subjects were generally not reflective of improvements over and above those achieved by the wait-list group.

Trower, Yardley, Bryant, and Shaw (1978) also compared SST and SD. They specifically identified subjects as having either a social-skills deficit or excessive anxiety as their principal difficulty. Half of the identified subjects in each group were then assigned randomly to one of the two treatments, so that half of the subjects were assigned to treatments matching their deficit while the other half were in a mismatched condition. This treatment-matching appears frequently in the social phobia treatment literature.

Patients in all treatment conditions evidenced reduction in the severity of their phobias, social inadequacy, and general anxiety following treatment regardless of the specific type of deficit they exhibited. There was also some specific benefit for skill-deficit subjects treated with SST. For example, skill-deficit subjects who received SST engaged in more social activities and experienced less difficulty in social situations compared with skills-deficit patients treated with SD. No such treatment-matching advantage occurred for the primarily anxious subjects who benefited equally from both treatments.

Shaw (1979) extended the results of Trower et al. (1978) by treating 10 additional anxious subjects with imaginal flooding (repeated presentation of highly anxiety-provoking scenes to imagination). These additional subjects showed reductions in phobic severity, social inadequacy, and general anxiety similar to those attained by subjects treated with SST and SD.

Although patients treated with SST, SD, and imaginal flooding were significantly improved, the lack of a control condition in Trower et al. (1978) and Shaw (1979) is problematic. This is especially the case given the results of the study by Marzillier et al. (1976) in which treated subjects were only marginally improved over a wait-list control condition. The lack of adequate experimental control plagues several of the studies reviewed in this paper.

Öst, Jerremalm, and Johansson (1981) also employed a treatment-matching design. Thirty-two patients with social phobia were classified as either "behavioral reactors" or "physiological reactors" based upon their responses during a role-played social interaction. Physiological reactors evidenced greater cardiac reactivity during the interaction, whereas behavioral reactors evidenced greater overt signs of anxiety during the conversation. Half the subjects in each group were then assigned randomly to either SST or applied relaxation (AR). AR combines multiple relaxation procedures (see below) and was hypothesized to be a better treatment match for the physiological responders. SST was presumed to be the best treatment match for patients with disrupted behavioral responses.

There was some limited support for treatment matching in this study. Behavioral reactors treated with SST (matched condition) increased social activity and social contacts, evidenced greater improvement on self-report measures of social fear, social difficulty, and avoidance of social situations, and reported lower subjective anxiety during a behavioral test when compared with behavioral reactors treated with AR (mismatched condition). Physiological reactors participated in more social activities, made more social contacts, and evidenced fewer signs of anxiety during the behavioral test when treated with AR (matched condition) than when treated with SST (mismatched condition). Most analyses showed no differences between the matched and mismatched groups.

Mersch, Emmelkamp, Bögels, and van der Sleen (1989) used a similar treatment-matching approach to identify behavioral reactors, as defined in Öst et al. (1981), and "cognitive reactors." This latter group comprised subjects with social phobia scoring relatively lower on a measure of rational thinking. SST and rational-emotive therapy (RET; Ellis, 1962) were administered in group format to half of the subjects in each group. RET is a complex cognitive-behavioral therapy that involves the analysis and disputation of clients' irrational beliefs (Ellis, 1962). Social anxiety was reduced significantly in every condition, and there were some findings which ran counter to the treatment-matching hypothesis. For example, patients receiving SST were rated as being less anxious and more skillful on a behavioral test but also reported more positive self-statements at post-treatment. Subjects receiving RET rated themselves as more skillful and less anxious at post-treatment and six-week follow-up. There were virtually no differences between subjects in the matched and mismatched conditions. Fourteen months after treatment ended, SST and RET treated subjects maintained their within-group gains (Mersch, Emmelkamp, & Lips, 1991). There were no differences between treatment conditions or between subjects treated in the matched vs. mismatched conditions.

One study compared SST with exposure to feared social situations for the treatment of social phobia. Wlazlo, Schroeder-Hartwig, Hand, Kaiser, and Münchau (1990) compared an individually-tailored but group-administered SST program (Liberman, King, De Risi, & McCann, 1975) with individual- and group-administered exposure. Seventy-eight patients were judged retrospectively to have either a "primary social skills deficit" or a "primary social phobia." Subjects were assigned to treatments based only upon the time they applied for treatment. Thus, there was no random assignment of subjects to treatment condition, and there was no systematic matching of subject type to treatment type.

Both groups of subjects receiving SST reported reduced social anxiety, avoidance, and interference of their symptoms with daily life activities. They also evidenced reductions in depression and obsessive rumination. Similar gains were made by subjects treated in the exposure conditions, and there were few differences at either the post-treatment or three-month follow-up assessment. Although patients treated with SST made significant gains, it is difficult to determine how much of these gains should be attributed to SST or the remediation of social skills deficits since all patients were given specific instructions for exposure as part of the homework component of all treatments.

(i) Addition of other treatment components

Two studies have examined the efficacy of adding other treatment components to SST for the treatment of social phobia. Stravynski, Marks, and Yule (1992) examined the addition of cognitive modification procedures to SST in the treatment of 22 subjects meeting *DSM-III* criteria for social phobia. Subjects received 12 weekly 1.5-hour individual or group sessions (assignment to individual or group modes of treatment was not random) following a control waiting period in which all subjects received only assessments. Patients in both conditions (SST with or without cognitive modification) reported increases in degree of social interaction and reductions in depression and irrational beliefs regarding social situations following treatment. Similar results were reported at post-treatment assessment in both individual and group treatment conditions. At six-month follow-up, gains were maintained with essentially no difference between subjects in the two treatment conditions. Cognitive procedures did not produce any benefit over and above SST alone. This issue will be discussed further in Section 6.21.3.4.1(iii).

Falloon, Lloyd, and Harpin (1981) examined whether the addition of propanolol, a beta-blocker, would enhance the effects of SST for 16 subjects meeting *DSM-III* criteria for social phobia. Treatment which began after a four-week waiting period, was targeted at two problematic situations for each subject, with each problem the focus of one-half of the total treatment effort. SST began with a six-hour marathon session in which each subject met with another subject and a lay therapist focusing on the first problem situation. This was followed by two weeks of *in vivo* practice with the aid of the therapist. The focus then shifted to the second problem situation and the procedures were repeated.

No changes occurred during the pretreatment waiting period. Analyses of within-group changes showed reductions in social anxiety, decreased difficulty in targeted situations, and increases in positive self-image. A six-month follow-up of 81% of the subjects indicated that gains were maintained. There were no between-groups differences, however, indicating that propanolol did not enhance the efficacy of SST.

(ii) Recent advances

Although interest in SST had declined in recent years, it has received renewed interest as part of multicomponent treatment packages for the treatment of more severely disturbed individuals. Two studies present preliminary data on the efficacy of treatments which combine several treatment components and which prominently feature social skills training. Feske et al. (1996) used a combination of group exposure, relaxation training, systematic desensitization, and SST in the treatment of generalized social phobics with and without avoidant personality disorder. Both groups of patients improved from pretest to post-test and achieved further gains by three-month follow-up. Differences between patients with and without APD have been previously described.

Social effectiveness therapy (SET) combines education, social skills training, imaginal exposure, and therapist-directed exposure (Turner, Beidel, Cooley, Woody, & Messer, 1994). Turner et al. (1994) employed a broad set of assessments including self-report measures, a standardized behavioral test, and clinician ratings as well as a composite index of outcome (Turner, Beidel, Long, Turner, & Townsley, 1993) in an uncontrolled study of 17 patients meeting criteria for generalized social phobia. A combination of individual and group-treatment sessions were administered for 29 sessions conducted over 16 weeks. At post-treatment, 84% of the 13 patients who completed treatment (59% of the full sample) were classified as having achieved either moderate or high end-state functioning.

Turner, Beidel, and Cooley-Quille (1995) contacted patients for follow-up assessment two years after they completed SET. Eight of the original 13 patients completed a self-report battery while the remaining five completed a telephone interview. These patients maintained their gains or made some additional gains 21–24 months after the conclusion of treatment. Although these studies produced promising results, they are clearly preliminary and require follow-up evaluation in controlled treatment trials. Furthermore, they tell us little about the efficacy or mechanism of SST for social phobia.

6.21.3.2 Relaxation Techniques

The use of relaxation techniques for social phobia is based on the straightforward notion that they should provide the person with a means of coping with the physiological manifestations of anxiety. Several studies have examined the effectiveness of relaxation strategies in the treatment of social phobia. Most commonly, studies of this sort employ one of only a few strategies. In this section, we review studies that have employed either progressive muscular relaxation (Jacobson, 1938), systematic desensitization (Wolpe, 1958), or applied relaxation (Öst, 1987).

6.21.3.2.1 Review of empirical studies of relaxation techniques

Three studies already reviewed (Marzillier et al., 1976; Shaw, 1979; Trower et al., 1978) examined SD in comparison with SST and imaginal flooding. As noted, those studies found within-group improvements for SD but few group differences between SD and the other treatments. Furthermore, just one study included a control condition, and that study showed little benefit for SD above and beyond the wait-list control group (Marzillier et al.).

Two other studies compared progressive muscular relaxation to exposure techniques (Al-Kubaisy et al., 1992; Alström, Nordlund, Persson, Hårding, & Ljungqvist, 1984). In both studies (reviewed in more detail in the section on exposure techniques), relaxation fared poorly when compared with exposure.

(i) Applied relaxation

Öst and colleagues have had more success with alternative strategies employing relaxation techniques by linking relaxation more closely to actual social situations. Applied relaxation (AR) involves a series of techniques in which subjects (i) increase awareness of the earliest signs of anxiety and (ii) apply a series of relaxation techniques during anxiety-provoking situations (Öst, 1987). These techniques begin with standard progressive muscular relaxation strategies but then aim to decrease the time required to achieve a relaxed state. With practice, subjects reach moderate levels of relaxation in just 20–30 seconds. They then apply their relaxation skills, first in practice and then in real-life feared situations. These procedures have been used in two studies of patients with social phobia. Both studies used AR and a treatment-matching approach. A study described earlier indicated that physiological reactors who received AR participated in more social activities, made more social contacts, and showed fewer behavioral signs of anxiety during a social interaction test than physiological reactors who received SST (Öst et al., 1981).

Jerremalm, Jansson, and Öst (1986) used a similar strategy for physiological reactors and cognitive reactors (defined by high rates of endorsement of negative cognitions after a behavioral test), this time matching cognitive reactors with self-instructional training, a highly structured cognitive-behavioral technique. AR was more effective than a waiting-list condition and as effective as stress-inoculation training (SIT) on a number of measures. There was no support for the treatment-matching hypothesis and on some outcome measures, SIT outperformed AR for the physiological reactors as well. This study will be reviewed in more detail in Section 6.21.3.4.1.

6.21.3.3 Exposure-based Techniques

Exposure to a phobic stimulus has been the cornerstone of behavioral treatments and is viewed as a central component for effective fear reduction (Barlow & Beck, 1984). Thus, exposure has been evaluated in several studies of the treatment of social phobia. Exposure takes many forms in these studies, varying on several dimensions including degree of therapist involvement, length of exposure sessions and total exposure time in therapy, and imaginal vs. role-played vs. actual (*in vivo*) exposure, to name just a few. There is virtually no research that systematically compares the differential effects of these dimensions of exposure in the treatment of social phobia.

6.21.3.3.1 Review of empirical studies of exposure-based tehniques

Fava, Grandi, and Canestrari (1989) provided instructions for self-directed exposure to 10 social phobics in an uncontrolled study. Eight treatment sessions produced significant reductions in self-report and observer-rated anxiety in the seven patients who completed treatment. Al-Kubaisy et al. (1992) compared self-directed *in vivo* exposure with and without therapist assistance to relaxation strategies in the treatment of 28 social phobic patients (this was part of a larger study which included agoraphobic and specific phobia patients). Patients in the exposure conditions experienced greater reductions in anxiety and greater increases in approach behavior than patients treated with relaxation. Therapist-guided exposure was only minimally more effective than self-directed exposure instructions.

Alström et al. (1984) compared therapist-directed exposure to dynamically oriented supportive therapy and relaxation therapy. Forty-two patients were randomly assigned to one of the three treatments or a control condition ("basal therapy") which included unspecified anxiolytic medication plus instructions for self-exposure. Patients in the three treatment conditions also received basal therapy. Exposure patients received therapist-directed, prolonged exposure to moderately difficult but common situations. Patients were instructed to remain in feared situations until habituation occurred. Situation difficulty was graduated and therapist assistance gradually removed. Imaginal exposure supplemented this procedure for some patients treated with exposure.

Patients treated with exposure evidenced greater reductions in anticipatory anxiety and anxiety in difficult situations, avoided less when faced with difficult situations, and showed greater improvement on a measure of global disturbance compared with patients treated with relaxation or basal therapy alone. Exposure patients appeared to lose some of their gains by nine-month follow-up, no longer maintaining an advantage over patients treated with relaxation or supportive therapy. They did, however, continue to outpace patients receiving basal therapy alone on measures of situational and anticipatory anxiety.

Interpretation of the results of Alström et al. (1984) is difficult for several reasons. Patients selected for participation in this study were first deemed to be "unsuitable for insight-oriented psychotherapy" and it is unclear how this exclusion effects the generalizability of the findings. Furthermore, differences between groups at pretreatment on gender, other demographic characteristics, and level of impairment, and the unsystematic application of treatment within and across groups further compromises the reliability of these findings.

Emmelkamp, Mersch, Vissia, and van der Helm (1985) compared exposure with RET and SIT in the treatment of 34 patients meeting

DSM-III criteria for social phobia. Exposure consisted of within-session role-plays (such as public speaking and maintaining eye contact) and homework instruction to attempt *in vivo* exposure (such as making inquiries in shops and speaking to strangers). The cognitive treatments will be described in Section 6.21.3.4.1.

Exposure resulted in reductions in social anxiety and general psychopathology at post-treatment assessment and additional improvements at one-month follow-up. At post-test, exposure patients evidenced significantly greater reductions in heart rate before and after a behavior test than patients treated with RET or SIT. No other differences favored exposure at post-test or follow-up.

In a previously described treatment-matching study, Wlazlo et al. (1990) compared individual- and group-administered exposure with SST for patients classified as having primary social skills deficits or primary social phobia. Within-group changes occurred in all treatments on most measures. Patients with primary social phobia who received group exposure reported less fear of social contact and greater ability to refuse requests than patients with social skills deficits treated with group exposure. As we noted previously, problems with this study (e.g., all groups received instruction for self-exposure) compromise our ability to determine the relative benefits of these treatments.

6.21.3.4 Cognitive-behavioral Techniques

Social phobia has been a focus of attention of cognitive theorists for a considerable time. For example, Butler (1985, p. 651) asserted that patients' "thoughts and attitudes seem to play a central role in the maintenance of social phobia." Butler went on to suggest that treatment of social phobia which does not address these cognitive factors may be less effective. Several researchers have proposed models of social phobia (see earlier discussion) which highlight cognitive factors as the likely maintaining mechanism of social phobia in the face of exposure to feared situations (e.g., Clark & Wells; 1995; Rapee & Heimberg, 1997).

Cognitive-behavior therapy (CBT) draws heavily from two sources—RET, first put forth by Ellis in the early 1960s, and Beck's cognitive therapy, first put forth a decade later. Although various labels exist for what is termed CBT, most cognitive techniques share their origins and many common features with these two approaches. RET is based on the principal that human beings are prone to fall victim to one or more irrational beliefs which are common in Western civilization (Ellis, 1962). Examples include the belief that it is essential to be loved and approved by all significant people in one's life and the belief that it is necessary to be completely competent, adequate, and achieving in all areas to be worthwhile. Social phobics may easily be understood to fall victim to these two (and other) irrational beliefs. RET is designed to confront and contradict irrational beliefs through frank discussion, encouragement, and persuasion, with the ultimate goal to change a person's basic philosophy and reliance on irrational beliefs (Ellis, 1972).

Unlike RET, Beck's cognitive therapy focuses less on generalizations about one's misfortunes and more on errors of cognitive processing which lead to distortion and misinterpretation (Beck, 1976). These so-called cognitive errors include selective abstraction (focusing on a detail out of context and missing the significance of the situation) and arbitrary inference (reaching conclusions without adequate evidence or contrary to available evidence). Cognitive therapy involves openly challenging these types of errors with logic and Socratic discussion. Ultimately, patients begin to view their situation in more realistic and adaptive terms. To be precise, however, there is really no such thing as a cognitive therapy which is not also behavioral, because these therapies rely heavily on exposure to feared situations and the conduct of behavioral experiments.

6.21.3.4.1 Review of empirical studies of cognitive-behavioral technique

In a study briefly described earlier, Emmelkamp et al. (1985) compared RET and SIT with role-played and *in vivo* exposure. RET focused on challenging irrational beliefs commonly held by social phobic persons. SIT is a modified version of Meichenbaum's (1985) stress-inoculation training without relaxation procedures. SIT patients identify and record negative thoughts and affective responses that occur in social situations. Patients then work with their therapist to develop more realistic appraisals designed to increase coping with anxiety-provoking situations. Neither SIT or RET as conducted in this study included an exposure component.

Patients in the cognitive treatments reported reductions in social anxiety, general psychopathology, and irrational beliefs. RET patients reported less social anxiety on one measure at post-test and maintained this difference one month later compared with SIT patients. Additional improvement occurred in both conditions during the brief follow-up period and patients in both cognitive treatments reported fewer irrational beliefs than patients treated with exposure.

Jerremalm et al. (1986) classified 38 social phobics as physiological reactors (greater heart rate reactivity during a role-played social interaction) or as cognitive reactors (relatively greater frequency of negative thoughts during the interaction). Using a treatment-matching paradigm, patients of each type were assigned randomly to either AR, SIT, or a wait-list condition. Physiological reactors receiving either treatment were more improved than waiting-list patients on measures of anxiety and heart rate during the behavior test. Also, physiological reactors assigned to SIT reported fewer negative thoughts and more positive thoughts than the control group. Cognitive reactors receiving either treatment were also significantly improved. When treated with SIT, cognitive reactors improved more than when treated with AR on three of seven self-report measures. No advantage was found for the physiological reactors treated with AR, and on two measures, SIT was actually more effective than AR.

In another study described earlier, Mersch et al. (1989) also used a treatment matching model, comparing RET with SST in the treatment of social phobic patients classified as either cognitive or behavioral reactors. Both groups of patients evidenced numerous improvements regardless of treatment including reductions in social anxiety. Contrary to the treatment-matching hypothesis, patients treated with SST improved on measures of cognitive functioning and patients treated with RET rated themselves as more skillful.

(i) Combined cognitive-exposure techniques

Several studies have combined cognitive techniques with exposure more fully. Kanter and Goldfried (1979) first evaluated a cognitive treatment with 68 volunteers who appear to have met diagnostic criteria for social phobia (this study preceded publication of *DSM-III*). In this variation of RET known as systematic rational restructuring (SRR; Goldfried, DeCenteceo, & Weinberg, 1974), patients use imagery of anxiety-provoking situations to identify unrealistic thoughts, logically challenge those thoughts, and substitute more adaptive ones in their place. Anxiety experienced during imagery cues the patient to begin cognitive procedures. SRR was compared to self-control desensitization (SCD; Goldfried, 1971) in which patients respond to imagery-induced anxiety with progressive relaxation, to the combination of SCD and SRR, and to a wait-list control. Patients were treated in groups of 8–10 for seven weekly 1.5-hour sessions.

Patients in the three active treatments improved significantly more than waiting-list patients. Patients receiving SRR (singularly or as part of the combined treatment) showed significant improvement on social anxiety, trait anxiety, and irrational beliefs; improvement was seen on 16 of 19 measures overall. SCD treated patients improved on 10 of 19 measures. SRR patients also reported lower trait anxiety, fewer irrational beliefs, and lower anxiety during a behavior test than SCD patients. Combined treatment patients reported less anxiety about the prospect of giving a speech and less anxiety during a behavior test than SCD patients. At a nine-week follow-up assessment, patients on active treatments generally maintained their gains. Patients treated with SRR had significantly greater reductions in trait anxiety, fear of negative evaluation, and irrational beliefs than SCD patients.

Stravynski (1983) treated a single patient with psychogenic vomiting which occurred whenever the patient visited his girlfriend's parent's home and when he went to theaters and bars. *In vivo* exposure was combined with social skills training and cognitive restructuring and all components were introduced simultaneously. Vomiting was eliminated after seven sessions, and anxiety in three of four target situations dramatically decreased. Seven months following treatment the patient continued to perform the target behaviors, and an informal follow-up after two years suggested that gains had been maintained. In an uncontrolled study, Mersch, Hildebrand, Lavy, Wessel, and van Hout (1992) administered a unique combination of RET and paradoxical interventions to three patients with fears that public embarrassment would result from the occurrence of their physiological symptoms of anxiety (e.g., blushing, sweating, trembling). Paradoxical techniques incorporated exposure. All patients improved on a broad array of measures and frequency of feared symptoms was reduced, although anxiety experienced on occurrence of these symptoms remained high. While these gains were maintained at an 18-month follow-up, methodological considerations limit conclusions about the efficacy of this treatment approach.

Heimberg, Becker, Goldfinger, and Vermilyea (1985) treated seven social phobic patients in small groups with a combination of cognitive restructuring, imaginal exposure, exposure to simulated social situations, and homework assignments for *in vivo* exposure. Cognitive restructuring was conducted immediately following exposure exercises and focused on negative cognitions related to the simulated or imagined situation. These procedures resulted in significant reductions in social anxiety, general anxiety, and fear of negative evaluation immediately following treatment. Patients rated their own performance during a behavior test as higher in quality, and they were rated by others as less anxious. Attributions for negative outcomes became less internal and stable. For six of the seven patients, gains were maintained at a six-month follow-up on both self-report and behavior test measures, whereas one patient's anxiety returned to baseline levels.

Based on this preliminary study, Heimberg et al. (1990) removed imaginal exposure and developed a more highly integrated treatment combining cognitive restructuring and exposure exercises. In this treatment package, referred to as cognitive behavioral group therapy for social phobia (CBGT), cognitive restructuring is employed before, during, and after simulated and *in vivo* exposures (the latter comprise homework assignments). CBGT is described in more detail by Heimberg et al. (1995).

Heimberg et al. (1990) compared CBGT with an attention-placebo group in the treatment of 49 social phobic patients. The comparison treatment combined educational presentations on topics relevant to social phobia and guided peer support. Patients were assigned randomly to either CBGT or the placebo group. At post-test, patients treated with CBGT were less anxious during a behavior test and were rated as less severely impaired by clinical assessors. CBGT patients maintained their gains at a six-month follow-up and reported more positive and fewer negative thoughts during the behavior test. Eighty-one percent of CBGT patients, compared with 47% of placebo patients, were rated as having made clinically significant improvement at follow-up.

Heimberg, Salzman, Holt, and Blendell (1993) reassessed 19 of the patients who participated in the Heimberg et al. (1990) study an average of 5.5 years following treatment (range 4.5–6.25 years). Although these patients were less impaired at pretest than patients who did not participate in this follow-up, there were no differences between participating CBGT and control patients. Nonetheless, generalizations based on these results should be limited to less impaired patient populations.

Within these limits, this study supports the long-term efficacy of CBGT. Specifically, CBGT patients were rated by independent assessors as less severely impaired and as experiencing less interference at work, in social activities, and in their family life compared with attention-placebo patients. Eighty-nine percent of CBGT patients and 44% of placebo patients were judged to be significantly improved. CBGT patients were also rated as evidencing more social skills and less anxiety during a behavioral test.

(ii) Cognitive-behavioral techniques and pharmacotherapy

Several pharmacological treatments of social phobia exist and have been studied to varying degrees. Although pharmacological approaches will not be reviewed here (see Liebowitz & Marshall, 1995; Potts & Davidson, 1995, for reviews of pharmacological treatments), three studies have compared cognitive restructuring plus exposure with pharmacological treatments. Two of those studies used Heimberg's CBGT. Gelernter et al. (1991) compared CBGT with the monoamine oxidase inhibitor phenelzine, the triazolobenzodiazepine alprazolam, and pill placebo. Patients assigned to receive medication or placebo were also provided with self-exposure instructions, a procedure that makes this study more clinically relevant but reduces the distinctiveness of treatments. Patients in all four groups experienced significant improvement on all measures, with few differential effects. Patients treated with phenelzine were more improved than other patients on a measure of trait anxiety at two-month follow-up, but similarities among outcomes far outweighed differences. At follow-up, patients receiving phenelzine or CBGT maintained their gains while alprazolam patients did not. Several methodological concerns complicate interpretation of these results. These include sole reliance on self-report measures, the inclusion of self-exposure instructions in the medication and placebo conditions, and an increase in the recommended number of patients in each CBGT group from six to 10.

A large-scale, multisite collaborative study compared CBGT, phenelzine, pill placebo, and the attention-placebo treatment employed by Heimberg et al. (1990). Heimberg et al. (1997) randomly assigned 133 patients to 12 weeks of one of the four treatments. Patients judged to be responders to CBGT or phenelzine continued for an additional six months of therapy with reduced (monthly) contact and six months of treatment-free follow-up. CBGT and phenelzine produced equivalent response rates, superior to both control conditions, after 12 weeks of treatment. Phenelzine led to faster response and was somewhat more effective than CBGT on some measures after 12 weeks. During the six-month untreated follow-up, CBGT patients were less likely to relapse than patients treated with phenelzine (Liebowitz et al., 1997). This was especially the case for patients with generalized social phobia. A study examining the efficacy of the combination of these two treatment modalities is underway.

The third study comparing pharmacological treatments with cognitive restructuring and exposure treated 34 musicians with performance anxiety (Clark & Agras, 1991). The combination of cognitive restructuring, exposure, and relaxation was evaluated with and without the anxiolytic medication buspirone. Musicians were randomly assigned to one of four conditions: cognitive-behavioral treatment with buspirone, cognitive-behavioral treatment with placebo, buspirone alone, and placebo alone. Generally, patients treated with cognitive-behavioral therapy experienced improvement on several important outcome variables, whereas buspirone was no more effective than placebo. Cognitive-behavioral therapy plus placebo was the most effective treatment combination, and these patients outperformed all others on a measure of confidence in performing at a one-month follow-up assessment.

(iii) Component analysis of cognitive-exposure techniques

From this review, it is clear that the combination of cognitive restructuring and exposure can be a highly effective treatment for social phobia. However, these studies do not tell us whether or not it is necessary to combine these two treatment approaches and whether the addition of cognitive restructuring to exposure techniques is in any way more effective than exposure alone. A series of studies has examined this issue and is reviewed below.

Butler, Cullington, Munby, Amies, and Gelder (1984) compared exposure only with exposure plus anxiety management training (Suinn & Richardson, 1971). The latter comprises relaxation, distraction, and rational self-talk. Patients in the exposure-only condition received a filler treatment which occupied the same amount of time as anxiety management training. After treatment, patients in both exposure conditions were superior to a wait-list group, having achieved reductions in phobic severity, anxiety during a behavior test, difficulty during social situations, general anxiety and depression. Combined treatment patients achieved lower scores on two measures of social anxiety compared with patients receiving exposure only.

Differences between combined treatment patients and exposure-only patients were maintained or increased at a six-month follow-up assessment. Fully 40% of patients receiving exposure only sought further treatment in the following 12 months, whereas no combined treatment patient did so. Although anxiety management training is not a solely cognitive intervention, *post hoc* analyses suggested that rational self-talk was the most frequently utilized coping technique. Problems with procedures in the exposure condition are elaborated elsewhere (Juster, Heimberg, & Holt, 1996) but include limitation on the total time spent on exposure and its procedures, and low credibility ratings for the exposure-only condition possibly due, in part, to the filler treatment.

Mattick and his colleagues conducted two studies examining exposure and cognitive techniques in various combinations (Mattick & Peters, 1988; Mattick, Peters, & Clarke, 1989). The first study compared therapist-assisted exposure with and without cognitive restructuring (Mattick & Peters, 1988). In contrast with the study by Butler et al. (1984), no filler treatment was used. Both treatments yielded similar post-treatment outcomes, including reductions in avoidance, phobic severity, and depression. At three-month follow-up, patients in the combined treatment group completed a greater percentage of tasks during a behavioral test and evidenced lower avoidance and higher scores on composite measures of improvement and end-state functioning than patients treated with exposure alone.

In their second study, Mattick et al. (1989) compared the same conditions as in the previous study, but added a cognitive restructuring only condition and a wait-list condition. Patients in the three active treatments completed more tasks during a behavior test and scored lower on self-report measures of phobia severity and avoidance than patients on the wait-list. Exposure only and the combined treatment were more effective than cognitive restructuring alone in terms of the tasks completed during the behavior test. At the three-month follow-up, however, patients in the cognitive restructuring only and combined treatment conditions continued to improve, whereas patients receiving exposure only appeared to deteriorate. Patients in the combined treatment condition experienced the greatest improvement on percentage of behavioral tasks completed.

Not all studies support the combination of exposure and cognitive techniques. In fact, the majority of studies on this issue do not support the need for cognitive procedures to be added to exposure techniques. Biran, Augusto, and Wilson (1981) examined the effect of adding cognitive therapy to *in vivo* exposure in a multiple-baseline-across-subjects design. Two of three women with fears of writing in public received five sessions of cognitive restructuring followed by five sessions of exposure. The third patient received exposure only. Exposure, conducted according to a standardized hierarchy of feared writing situations, produced significant increases in the number of writing tasks performed. Cognitive restructuring did not enhance the effectiveness of exposure. However, reductions in fear ratings during performance of the exposure situations were less clearly tied to specific treatment.

At nine-month follow-up, avoidance of feared situations remained low. However, fear in those situations returned to pretreatment levels. Although no further follow-up was conducted, one might speculate that this desynchrony between approach behavior and anxiety could predict poor long-term outcome. In fact, two of the three patients sought additional treatment during follow-up, one for unspecified "social issues," the other for further anxiety reduction. All three indicated some disappointment that their anxiety reduction had not been maintained. Although this study has been criticized for several reasons, probably the most critical concern involves the total segregation of treatment components in a manner inconsistent with clinical practice and which may reduce treatment effectiveness (Heimberg, 1989; Heimberg, Dodge, & Becker, 1987).

More recently, Taylor et al. (1997) randomly assigned 65 patients to eight sessions of cognitive restructuring without exposure or eight sessions of the filler treatment used by Butler et al. (1984). Cognitive restructuring was clearly more effective than the filler condition after this initial phase of treatment. Following an assessment, both groups received additional sessions of exposure to determine if treatment response would be facilitated. There were no differences between groups after exposure on a battery of self-report measures administered at the three-month follow-up. Thus, after both groups were treated with exposure, patients who received the filler treatment were as improved as patients who had previously received cognitive restructuring.

Hope, Heimberg, and Bruch (1995a) compared exposure alone with the complete CBGT protocol described earlier to determine the relative importance of the cognitive component. Both conditions were more effective than a wait-list condition after 12 weeks of treatment. Patients treated with CBGT reported less anxiety during an individualized behavior test at post-treatment, but this was the only measure in which the full protocol outperformed exposure alone. Differences between exposure and CBGT at post-treatment, most of which favored exposure, disappeared at the six-month follow-up, at which time the two treatments were equally effective. In this trial, however, CBGT was less effective than in other studies of this protocol.

Three studies employed RET in assessing the efficacy of cognitive therapy with and without exposure techniques. Scholing and Emmelkamp (1993a) treated 30 social phobics with primary fears of blushing, sweating, or trembling. Individual treatment was administered in two four-week blocks separated by a four-week period of no treatment. Patients were randomly assigned to either *in vivo* exposure followed by RET, RET followed by *in vivo* exposure, or integrated RET and *in vivo* exposure. It is important to note that in the integrated treatment, exposures were never conducted during the therapy session. Thus, the primary focus of integrated therapy sessions was on changing irrational beliefs relative to situations which would then be the focus of home-based exposure exercises. All treatments produced substantial and equal improvements on composite measures of avoidance of target situations, irrational cognitions, and somatic complaints. No differences were evident after three-month follow-up.

Scholing and Emmelkamp (1993b) conducted an essentially similar study with 73 generalized social phobics. Patients were randomly assigned to exposure only, exposure followed by RET, or integrated RET and exposure. Treatments were provided in blocks as in the previous study and included both individual and group modalities. Patients assigned to integrated treatment in the individual modality had no exposures during their sessions. Patients assigned to exposure or integrated RET and exposure in the group modality received exposure during sessions, although discussion of maladaptive cognitions was clearly the primary emphasis in the integrated treatment. Patients in the integrated treatment reported more somatic complaints than patients in the other two treatments after the first treatment block. No significant differences were found between treatments after the second block or at three-month follow-up. Importantly, fully 47% of the exposure-only patients and 28% of the RET-followed-by-exposure patients sought additional treatment in the 18 months following their participation (Scholing & Emmelkamp, 1996). Only 6% of the patients in the integrated treatment condition sought further treatment during this period. While the authors reported no statistical analysis of this finding, it is similar to that reported by Butler et al. (1984).

Mersch (1995) compared exposure alone with an integrated treatment combining *in vivo* exposure, RET, and social skills training. As in the individual treatments administered by Scholing and Emmelkamp (1993a, 1993b), no exposures were conducted during treatment sessions. Thirty-four patients were randomly assigned to either a wait-list control group or to one of the two treatment groups. Sixteen sessions of individual therapy were provided, and follow-up assessments were conducted at three months and 18 months. Both treatments were superior to the wait-list control at post-test, but there were no differences between exposure alone and the integrated treatment. At 18-month follow-up, patients in both active treatments achieved additional gains on self-report measures of social anxiety, but they still did not differ from each other.

(iv) Results of meta-analyses

An alternative method of examining the relative importance of treatment components is meta-analysis, a technique which compares averaged effect sizes for particular treatments or treatment components across several studies. Three meta-analyses examining cognitive-behavioral treatment of social phobia have been conducted (Feske & Chambless, 1995; Gould, Buckminister, Pollack, Otto, & Yap, 1997; Taylor, 1996). In the meta-analysis by Feske and Chambless (1995), exposure plus cognitive restructuring and exposure alone did not differ in dropout rates and yielded similar effect sizes on both post-treatment and follow-up measures of social anxiety, cognitive symptoms, and depression/anxiety. Gould et al. (1997) evaluated 24 studies which examined cognitive-behavioral and pharmacological therapies which included a control condition. CBT and pharmacological treatment produced similar effect sizes for reductions in social anxiety, and both effect sizes were significantly different from zero. Within CBT, exposure with cognitive restructuring and exposure alone yielded similar attrition rates and effect sizes. Lastly, Taylor (1996) examined 42 treatment studies and compared six conditions: waiting-list control, placebo (the combination of pill placebo and attention-placebo), exposure alone, cognitive restructuring without exposure, cognitive restructuring with exposure, and SST. The dropout rate did not differ across the six conditions. The four active therapies and placebo produced greater effect sizes than the waiting-list control; however, only cognitive restructuring with exposure had a significantly larger effect size than placebo. Effects remained stable during follow-up for all treatments.

(v) Summary

The study of the relative efficacy of exposure vs. combined cognitive restructuring plus exposure represents one of the critical issues in the treatment of social phobia. However, this collection of studies presents results that are contradictory, confusing, and difficult to interpret. Several studies did not support the need for the integration of exposure and cognitive restructuring, while others suggest that combined treatment is superior and that additional gains are more likely during the follow-up period. The meta-analyses produced contradictory results regarding this question as well. Thus, the question of the importance of the integration of exposure and cognitive restructuring in the treatment of social phobia remains unresolved. However, cognitive change may be central to good clinical outcomes, regardless of how it is produced. The importance of cognitive change is demonstrated by studies in which change in fear of negative evaluation from pretreatment to post-treatment predicts end-state functioning (Mattick & Peters, 1988; Mattick et al., 1989). However, cognitive interventions may not be necessary to engender cognitive change. Exposure alone may provide sufficient disconfirmatory information to alter dysfunctional beliefs (Hope et al., 1995a; Mattick & Peters, 1988; Newman, Hofmann, Trabert, Roth, & Taylor, 1994).

Turk, Fresco, and Heimberg (in press) outline several additional issues that should be considered before conclusions are drawn regarding the necessity of cognitive interventions and the role of cognitive change in successful treatment outcome. The first issue concerns the measures used to demonstrate cognitive change. Questionnaire measures commonly used to assess cognition correlate highly with social phobia symptom measures. When change is noted on these measures, the observed gains may be a product of change in the portion of variance shared with symptom measures. In other words, change on commonly used questionnaire measures of cognition may reflect changes in cognition, anxiety, or both (Heimberg, 1994). Second, many “purely behavioral” interventions may implement “exposure only” procedures in a manner which facilitates cognitive change (Newman et al., 1994). Specifically, “exposure only” treatments which include therapist and/or group member feedback regarding patient skill, visibility of anxiety symptoms, ability to cope with anxiety, etc., share many features with cognitive approaches, although these aspects of therapy may not be labeled as cognitive interventions by noncognitive therapists. Finally, while studies that find equivalent outcomes for exposure vs. exposure plus cognitive restructuring typically interpret this pattern to mean that cognitive restructuring does not add to the efficacy of exposure, alternative interpretations are available. It is also reasonable to suggest that cognitive restructuring activities reduce the amount of exposure necessary for effective treatment.

6.21.3.5 Psychodynamic Formulations

Little has been written about the psychodynamic perspective on social phobia, and even less research specifically testing psychodynamic assumptions has been conducted. Still, psychodynamic theory and therapy remain popular and merit some discussion here. Gabbard (1992) proposes a traditional psychodynamic conception of social phobia. In one formulation, social interaction characterized by anxiety and withdrawal are presumed to be manifestations of important early developmental events re-experienced as adult anxiety symptoms. Specifically, it is thought that some social phobics have an unconscious wish to be the center of attention and to receive positive affirmation by those in their environment. This wish produces an immediate shame reaction in the social phobic individual believed to be associated with early negative shame-producing experiences with disapproving parental figures. Social phobia is also thought by some to be related to separation anxiety. In this variation, social phobia is seen as a recapitulation of childhood struggles with individuation and attachment. Efforts to become increasingly autonomous are countered by fears of losing parental love. According to Gabbard (1992), certain characteristic internal representations are common to psychodynamic views of social anxiety including criticism, ridicule, humiliation, abandonment, and embarrassment associated with parents, caregivers, and siblings. These so-called introjects are projected onto others in the social phobic person's environment and become the object of his/her adult anxiety.

In contrast to the more traditional psychodynamic perspective of social phobia, Cloitre and Shear (1995) espouse a psychobiological model which draws heavily on ethological theory, especially the work of Trower and Gilbert (1989). This model posits that competitive and cooperative modes of relating which are important to many species are central to human social interaction as well. Specifically, a competitive or defensive system of interaction is characteristic in groups that develop dominance hierarchies. Submission and dominance are critical behaviors that maintain the organization's hierarchical structure. A cooperative or safety system allows sharing of resources and effort, minimizing differences between members of the group. Hierarchies are more horizontal in a cooperative mode and more vertical in a competitive mode. Applied to social phobia, it is hypothesized that the person with social phobia is overly reliant on the defensive/competitive system of relating, always treating social interaction situations as dominant–submissive in nature (with the social phobic individual usually taking on a submissive role but desperately wanting to be the dominant figure). This bias towards a defensive posture leads the social phobic individual to become hypervigilant for threat cues in others' behavior, to be overly competitive or overly submissive, and to become hypersensitive to criticism. Although both the competitive and cooperative systems may be adaptive in different situations, social interaction is typically cooperative. Similar to Gabbard's (1992) perspective, it is believed that early interaction with parental figures stimulates overuse of the defensive system. In particular, parents may focus the child's attention on "dangers" in the social environment or conversely overprotect them unnecessarily. Children may experience their parents as controlling, hostile, or rejecting, and parents may communicate that caring is contingent on certain behaviors such as a submissive attitude with parents. Although some research does support early experiences as potential progenitors of later social anxiety in adult social phobics, this does not necessarily provide support for other aspects of these psychodynamic formulations.

6.21.4 PREDICTORS OF OUTCOME

Although there are now effective treatments for social phobia, not everyone who receives treatment benefits from it. For example, Heimberg et al. (1997) reported that 75% of patients who completed a 12-week trial of CBGT were independently judged to be moderate or full responders to that treatment. These figures are reduced (58%) when drop-outs are considered in an intent-to-treat analysis. Clearly, many patients with social phobia do not benefit maximally from the treatments they receive. It is not known whether patients who fail to respond to one treatment might be likely to respond to a different treatment, nor is it known which patient characteristics might predispose toward better or worse response to one or another treatment. However, several studies have begun to examine this issue. The majority of these studies have attempted to identify predictors of response to Heimberg's CBGT (Brown et al., 1995; Chambless, Tran, & Glass, 1997; Holt, Heimberg, & Hope, 1989, 1990; Hope et al., 1995b; Juster, Heimberg, & Mattia, 1993), whereas one other study examined predictors of response to flooding (Turner, Beidel, Wolff, Spaulding, & Jacob, 1996). It is not known how these results might generalize to the prediction of outcome of other treatments.

Demographic factors have shown little consistency in the prediction of outcome of cognitive-behavioral therapy for social phobia. Factors such as age, gender, marital status, employment, and educational status have been unrelated to outcome (Holt et al., 1989, 1990; Juster et al., 1993). Holt et al. (1989) found age of onset, locus of fear, and lower scores on the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) to predict outcome. Specifically, later onset, public speaking vs. social interaction fears and lower BDI scores were associated with better outcomes. However, Holt et al. (1990) found that similar results all but disappeared when pretreatment social phobia severity was statistically controlled. Juster et al. (1993) found no relationship between demographic factors, pretreatment clinical measures of social anxiety and depression, locus of control or general anxiety and response to CBGT.

Several studies have examined social phobia subtype as a potential predictor of CBGT outcome. Two studies found that patients with generalized social phobia improved with treatment as much as patients with nongeneralized social phobics. However, because they came to treatment more impaired than nongeneralized social phobics, they remained more impaired after treatment (Brown et al., 1995; Hope et al., 1995b). Patients with generalized social phobia were less likely to achieve high levels of endstate functioning than patients with nongeneralized social phobia. Heimberg et al. (1997) reported essentially the same pattern of results.

Turner et al. (1996) found that generalized and specific social phobics experienced similar levels of distress and habituation in response to flooding techniques. Moreover, measures of endstate functioning showed a similar pattern of response to that reported above for CBGT. Specifically, generalized social phobics' response to flooding was similar to the response of specific social phobics, but because they began treatment more impaired, they finished treatment that way as well. Subtype differences in response to treatment appear to reflect, for the most part, differences in severity.

The effect of APD diagnosis on outcome of the treatment of social phobia was addressed in Section 6.21.1.5. To summarize, generalized social phobics with and without APD show similar effects from cognitive-behavioral therapy. The presence of APD before treatment predicts lower endstate functioning after treatment, but as for social phobia subtype, this is related to pretreatment severity.

Patient expectancy regarding the benefits to be obtained from treatment may contribute to the outcome of cognitive-behavioral treatment of social phobia, and two studies have examined the expectancies of patients treated with CBGT (Chambless et al., 1997; Safren, Heimberg, & Juster, 1997). Both studies found expectancy to be a unique predictor of outcome even when other important variables were controlled. For example, Chambless et al. found that higher expectancies at the beginning of treatment were associated with improvement on self-ratings of anxiety and skill even after removing the effect of depression. Similarly, Safren et al. found that expectancy was significantly associated with improvement on a variety of self-report and interviewer measures of psychopathology after controlling for pretreatment severity. Thus, patient expectancies regarding the likelihood of obtaining benefit are a unique predictor of outcome. As noted by Chambless et al., the mechanism by which expectancies are associated with outcome is unknown and not necessarily causal.

6.21.5 SUMMARY AND FUTURE DIRECTIONS

Social phobia, described as the neglected anxiety disorder just 13 years ago (Liebowitz et al., 1985), has become an important focus of the attention of clinical psychologists and other mental-health professionals. The focus of this research has been on developing effective treatments as well as trying to better understand the phenomenology of social phobia. In this chapter, an overview of that effort has been provided. It is known, for example, that social phobia is more common than once believed, affecting nearly 8% of the US population in a given year. Social phobia affects more women than men; however, men and women seek treatment for social phobia in approximately equal numbers. Onset of the disorder is usually early in life, beginning for some in childhood, while for many, it may not appear until their teens or early 20s. Evidence points to social phobia as a chronic condition when left untreated, and it is not uncommon for patients to wait many years before pursuing treatment. Impairment in functioning due to social phobia may be extensive, reducing occupational and educational attainment in addition to social functioning. Finally, social phobia is usually associated with one or more comorbid conditions, especially other anxiety and mood disorders. Alcohol, drug abuse, and Axis II disorders are common complications of social phobia.

The serious nature of social phobia underscores the critical need for effective treatments for this disorder. In this chapter, the principal psychosocial treatment approaches for social phobia have been reviewed. These included

SST, relaxation techniques, exposure, and various cognitive-behavioral techniques. Patients treated with social skills training appear to achieve and maintain gains following treatment. However, as noted elsewhere, the lack of well-controlled treatment trials of SST reduces our ability to make strong statements about its efficacy (Juster & Heimberg, 1995). Despite this, SST may be an important ingredient in treatment of the most difficult patients with social phobia, especially those who also meet criteria for APD. The efficacy of this approach awaits further investigation.

Relaxation techniques have had mixed success in alleviating social phobia. Some studies found relaxation to be better than a wait-list condition, while others have not. Öst has reported more consistently positive results for AR, a combination of relaxation and exposure procedures which appears to benefit some patients with social phobia. However, AR procedures did less well than cognitive treatments in another study.

Exposure has fared well as a treatment for social phobia. However, as with SST, methodological problems limit conclusions about the comparative efficacy of exposure and other techniques. The bulk of studies reviewed combined exposure with some form of cognitive therapy, and several of these were methodologically sophisticated studies which allowed strong inference. These studies demonstrated that exposure plus cognitive restructuring is an effective treatment for social phobia, resulting in significant within-group gains and surpassing the performance of attention-placebo control conditions. CBGT, in particular, has demonstrated long-term efficacy as well. Despite these positive results, other studies which examined the components of cognitive restructuring and exposure separately and in combination have yielded conflicting and contradictory findings. Some studies support the combination of exposure and cognitive restructuring as superior, whereas others fail to show that the combined approach is better. Three meta-analyses also provide equivocal evidence for the utility of combining exposure and cognitive restructuring. Future studies are needed to address this issue further, paying special attention to methodological and conceptual issues raised herein.

The distinction between efficacy and effectiveness has become a central point of discussion (Seligman, 1995; VandenBos, 1996). Efficacy studies determine treatment response under relatively strict laboratory conditions. These conditions include the use of control groups or comparative treatments, random assignment, manualized treatments, a fixed number of sessions, raters who remain uninformed as to treatment assignment, patients with few or no comorbid conditions, and a fixed follow-up period (Seligman, 1995). The studies reviewed in this chapter are all studies of efficacy. In contrast, effectiveness studies establish a treatment's utility under real-world conditions in which one and usually several of the conditions of efficacy studies are not met. Effectiveness studies make sense once efficacy studies establish the empirical validity of a treatment. Thus, the effectiveness of combined exposure and cognitive restructuring for social phobia appears to be a timely issue.

One difficulty with efficacy studies is that the sample of patients who enter a study represents a narrow sampling of the population of patients with the disorder. Patients are often excluded from efficacy studies (e.g., due to the presence of comorbid conditions) or they may decline to participate in a study because of concern over random assignment, assignment to a placebo condition, etc. The question remains whether or not patients who are treated in an efficacy study respond similarly or differently to the treatment than patients who decide not to participate or are excluded from participation. These latter groups may represent a substantial portion of the population of patients with the disorder (perhaps the majority) but are often not represented in samples of patients for whom a treatment may be validated. One study of social phobia has addressed this issue. Juster, Heimberg, and Engleberg (1995) examined subjects screened for participation in Heimberg et al.'s (1997) comparative treatment trial of CBGT and phenelzine described earlier. One group of subjects was excluded for medical or diagnostic reasons and another refused entry in to the study, citing concerns about the side effects of phenelzine as the primary reason. The remaining subjects were treated in the comparative outcome study. Subjects who refused and were excluded from the trial were treated with CBGT outside the controlled trial, partly in an effort to establish the effectiveness of CBGT for a broader set of patients with social phobia. The three groups differed somewhat on demographic variables including measures of socioeconomic status and social support. Despite these differences, the three groups of patients demonstrated comparable gains in CBGT, suggesting that results of the clinical trial may be extended to these other groups of patients.

The Juster et al. (1995) study only begins to examine the effectiveness of CBGT and similar treatments for social phobia. Further research is needed to establish the effectiveness of these techniques outside university research centers and under less well-controlled conditions. Also, patients currently served by these centers tend to be more homogeneous than the broader population of social phobics. Treatment trials in nontraditional settings serving a diversity of patients need to be conducted. As we have noted before, the development of more effective approaches for the subset of patients who do not benefit from current methods is crucial (Juster & Heimberg, 1995). Whether that be in longer trials of existing treatments, combined pharmacotherapy and psychotherapy, or as yet untested treatment methods (e.g., interpersonal psychotherapy), there remains a proportion of patients who obtain little or no benefit from current treatments. Creative adjuncts to current treatments should also be considered and studied further. For example, the use of videotaped feedback appears to be a particularly useful therapeutic intervention in socially anxious subjects (Rapee & Hayman, 1996) and in patients with social phobia (Flynn, Polland, Carmin, & Logan, 1994), yet these methods have not been systematically studied. The search for efficacious and effective treatments for social phobia continues to be an important research agenda.

6.21.6 REFERENCES

Al-Kubaisy, T., Marks, I. M., Logsdail, S., Marks, M. P., Lovell, K., Sungur, M., & Araya, R. (1992). Role of exposure homework in phobia reduction: A controlled study. *Behavior Therapy, 23*, 599–621.

Alström, J. E., Nordlund, C. L., Persson, G., Hårding, M., & Ljungqvist, C. (1984). Effects of four treatment methods on social phobic patients not suitable for insight-oriented psychotherapy. *Acta Psychiatrica Scandinavica, 70*, 97–110.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. Rev.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Amies, P. L., Gelder, M. G., & Shaw, P. M. (1983). Social phobia: A comparative clinical study. *British Journal of Psychiatry, 142*, 174–179.

Barlow, D. H., & Beck, J. G. (1984). The psychosocial treatment of anxiety disorders: Current status, future directions. In J. B. W. Williams & R. L. Spitzer (Eds.), *Psychotherapy research: Where are we and where should we go*? (pp. 29–66). New York: Guilford Press.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: Meridian.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561–571.

Biederman, J., Rosenbaum, J. F., Hirshfeld, D. R., Faraone, S. V., Bolduc, E. A., Gersten, M., Meminger, S. R., Kagan, J., Snidman, M., & Reznick, J. S. (1990). Psychiatric correlates of behavioral inhibition in young children of parents with and without psychiatric disorders. *Archives of General, Psychiatry, 47*, 21–26.

Biran, M., Augusto, F., & Wilson, G. T. (1981). *In vivo* exposure vs. cognitive restructuring in the treatment of scriptophobia. *Behaviour Research and Therapy, 19*, 525–532.

Brown, E. B., Heimberg, R. G., & Juster, H. R. (1995). Social phobia subtype and avoidant personality disorder: Effect on severity of social phobia, impairment, and outcome of cognitive-behavioral treatment. *Behavior Therapy, 26*, 467–486.

Bruch, M. A., & Heimberg, R. G. (1994). Differences in perceptions of parental and personal characteristics between generalized and nongeneralized social phobics. *Journal of Anxiety Disorders, 8*, 155–168.

Bruch, M. A., Heimberg, R. G., Berger, P., & Collins, T. M. (1989). Social phobia and perceptions of early parental and personal characteristics. *Anxiety Research, 2*, 57–65.

Butler, G. (1985). Exposure as a treatment for social phobia: Some instructive difficulties. *Behaviour Research and Therapy, 23*, 651–657.

Butler, G., Cullington, A., Munby, M., Amies, P., & Gelder, M. (1984). Exposure and anxiety management in the treatment of social phobia. *Journal of Consulting and Clinical Psychology, 52*, 642–650.

Caspi, A., Elder, G. H., Jr., & Bem, D. J. (1988). Moving away from the world: Life-course patterns of shy children. *Developmental Psychology, 24*, 824–831.

Chambless, D. L., Tran, G. Q., & Glass, C. R. (1997). Predictors of response to cognitive-behavioral group therapy for social phobia. *Journal of Anxiety Disorders, 11*, 221–240.

Clark, D. B., & Agras, W. S. (1991). The assessment and treatment of performance anxiety in musicians. *American Journal of Psychiatry, 148*, 598–605.

Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 69–93). New York: Guilford Press.

Cloitre, M., & Shear, M. K. (1995). Psychodynamic perspectives. In M. B. Stein (Ed.), *Social phobia: Clinical and research perspectives* (pp. 163–188). Washington, DC: American Psychiatric Press.

Cross-National Collaborative Group (1992). The changing rate of major depression: Cross-national comparisons. *Journal of the American Medical Association, 268*, 3098–3105.

Ellis, A. (1962). *Reason and emotion in psychotherapy*. New York: Lyle Stuart.

Ellis, A. (1972). Rational-emotive psychotherapy. In J. T. Huler, & H. L. Millman, (Eds.), *Goals and behavior in psychotherapy and counseling*. Columbus, OH: Charles E. Merrill.

Emmelkamp, P. M., Mersch, P. P., Vissia, E., & van der Helm, M. (1985). Social phobia: A comparative evaluation of cognitive and behavioral interventions. *Behaviour Research and Therapy, 23*, 365–369.

Falloon, I. R. H., Lloyd, G. G., & Harpin, R. E. (1981). The treatment of social phobia: Real-life rehearsal with nonprofessional therapists. *Journal of Nervous and Mental Disease, 169*, 180–184.

Fava, G. A., Grandi, S., & Canestrari, R. (1989). Treatment of social phobia by homework exposure. *Psychotherapy and Psychosomatics, 52*, 209–213.

Feske, U., & Chambless, D. L. (1995). Cognitive-behavioral versus exposure treatment for social phobia: A meta-analysis. *Behavior Therapy, 26*, 695–720.

Feske, U., Perry, K. J., Chambless, D. L., Renneberg, B., & Goldstein, A. J. (1996). Avoidant personality disorder as a predictor for severity and treatment outcome among generalized social phobics. *Journal of Personality Disorders, 10*, 174–184.

Flynn, T., Pollard, C. A., Carmin, C., & Logan, A. (1994, November). *Exposure to taped enactments of feared social catastrophe: Successful treatment of four previously refractory cases of social anxiety*. Poster presented at the annual meeting of the Association for Advancement of Behavior Therapy, San Diego, CA.

Fyer, A. J., Mannuzza, S., Chapman, T. F., Liebowitz, M. R., & Klein, D. F. (1993). A direct family interview study of social phobia. *Archives of General Psychiatry, 50,* 286–293.

Gabbard, G. O. (1992). Psychodynamics of panic disorder and social phobia. *Bulletin of the Menninger Clinic, 56,* A3–13.

Gelernter, C. S., Uhde, T. W., Cimbolic, P., Arnkoff, D. B., Vittone, B. J., Tancer, M. E., & Bartko, J. J. (1991). Cognitive-behavioral and pharmacological treatments for social phobia: A controlled study. *Archives of General Psychiatry, 48,* 938–945.

Goldfried, M. R. (1971). Systematic desensitization as training in self-control. *Journal of Consulting and Clinical Psychology, 37,* 228–235.

Goldfried, M. R., DeCenteceo, E. T., & Weinberg, L. (1974). Systematic rational restructuring as a self-control technique. *Behavior Therapy, 5,* 247–254.

Gould, R. A., Buckminster, S., Pollack, M. H., Otto, M. W., & Yap, L. (1997). Cognitive behavioral and pharmacological treatment for social phobia: A meta-analysis. *Clinical Psychology: Science and Practice, 4,* 291–306.

Hazen, A. L., & Stein, M. B. (1995). Clinical phenomenology and comorbidity. In M. B. Stein (Ed.), *Social phobia: Clinical and research perspectives* (pp. 3–42). Washington, DC: American Psychiatric Press.

Heimberg, R. G. (1989). Cognitive and behavioral treatments for social phobia: A critical analysis. *Clinical Psychology Review, 9,* 107–128.

Heimberg, R. G. (1994). Cognitive assessment strategies and the measurement of outcome of treatment for social phobia. *Behaviour Research and Therapy, 32,* 269–280.

Heimberg, R. G. (1996). Social phobia, avoidant personality disorder and multiaxial conceptualization of interpersonal anxiety. In P. M. Salkovski (Ed.), *Trends in cognitive and behavioural therapies* (pp. 43–61). Chichester, UK: Wiley.

Heimberg, R. G., & Barlow, D. H. (1988). Psychosocial treatments for social phobia. *Psychosomatics, 29,* 27–37.

Heimberg, R. G., Becker, R. E., Goldfinger, K., & Vermilyea, J. A. (1985). Treatment of social phobia by exposure, cognitive restructuring, and homework assignments. *Journal of Nervous and Mental Disease, 173,* 236–245.

Heimberg, R. G., Dodge, C. S., & Becker, R. E. (1987). Social phobia. In L. Michelson & M. Ascher (Eds.), *Cognitive behavioral assessment and treatment of anxiety disorders* (pp. 280–309). New York: Plenum.

Heimberg, R. G., Dodge, C. S., Hope, D. A., Kennedy, C. R., Zollo, L., & Becker, R. E. (1990). Cognitive-behavioral group treatment of social phobia: Comparison to a credible placebo control. *Cognitive Therapy and Research, 14,* 1–23.

Heimberg, R. G., Juster, H. R. (1995). Cognitive-behavioral treatments: Literature review. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 261–309). New York: Guilford Press.

Heimberg, R. G., Juster, H. R., Hope, D. A., & Mattia, J. I. (1995). Cognitive behavioral group treatment for social phobia: Description, case presentation and empirical support. In M. B. Stein (Ed.), *Social phobia: Clinical and research perspectives* (pp. 293–321). Washington, DC: American Psychiatric Press.

Heimberg, R. G., Liebowitz, M. R., Hope, D. A., & Schneier, F. R. (Eds.) (1995). *Social phobia: Diagnosis, assessment, and treatment.* New York: Guilford Press.

Heimberg, R. G., Liebowitz, M. R., Hope, D. A., Schneier, F. R., Holt, C. S., Welkowitz, L. A., Juster, H. R., Campeas, R., Bruch, M. A., Cloitre, M., Fallon, B., & Klein, D. F. (1997). *Cognitive-behavioral group therapy versus phenelzine in social phobia: 12-week outcome*. Manuscript submitted for publication.

Heimberg, R. G., Salzman, D. G., Holt, C. S., & Blendell, K. A. (1993). Cognitive-behavioral group treatment for social phobia: Effectiveness at five-year follow-up. *Cognitive Therapy and Research, 17,* 325–339.

Herbert, J. D., Hope, D. A., & Bellack, A. S. (1992). Validity of the distinction between generalized social phobia and avoidant personality disorder. *Journal of Abnormal Psychology, 101,* 332–339.

Himle, J. A., & Hill, E. M. (1991). Alcohol abuse and the anxiety disorders: Evidence from the Epidemiologic Catchment Area Survey. *Journal of Anxiety Disorders, 5,* 237–245.

Holt, C. S., Heimberg, R. G., & Hope, D. A. (1989, November). *Baseline predictors of group psychotherapy outcome for the treatment of social phobia*. Paper presented at the annual meeting of the Association for Advancement of Behavior Therapy, Washington, DC.

Holt, C. S., Heimberg, R. G., & Hope, D. A. (1990, November). *Success from the outset: Predictors of cognitive-behavioral therapy outcome among social phobics*. Paper presented at the annual meeting of the Association for Advancement of Behavior Therapy, San Francisco.

Holt, C. S., Heimberg, R. G., & Hope, D. A. (1992). Avoidant personality disorder and the generalized subtype of social phobia. *Journal of Abnormal Psychology, 101,* 318–325.

Hope, D. A., Heimberg, R. G., & Bruch, M. A. (1995a). Dismantling cognitive-behavioral group therapy for social phobia. *Behaviour Research and Therapy, 33,* 637–650.

Hope, D. A., Herbert, J. D., & White, C. (1995b). Diagnostic subtype, avoidant personality disorder, and efficacy of cognitive behavioral group therapy for social phobia. *Cognitive Therapy and Research, 19,* 285–303.

Hope, D. A., Holt, C. S., & Heimberg, R. G. (1993). Social phobia. In T. R. Giles (Ed.), *Handbook of effective psychotherapy* (pp. 227–251). New York: Plenum.

Jacobson, E. (1938). *Progressive relaxation*. Chicago: University of Chicago Press.

Jerremalm, A., Jansson, L., & Öst, L.-G. (1986). Cognitive and physiological reactivity and the effects of different behavioral methods in the treatment of social phobia. *Behaviour Research and Therapy, 24,* 171–180.

Juster, H. R., & Heimberg, R. G. (1995). Social phobia: Longitudinal course and long-term outcome of cognitive-behavioral treatment. In M. H. Pollack & M. Otto (Eds.), *Psychiatric Clinics of North America Vol. 18. Anxiety disorders: Longitudinal course and treatment* (pp. 821–842). New York: W. B. Saunders.

Juster, H. R., Heimberg, R. G., & Engelberg, B. (1995). Self selection and sample selection in a treatment study of social phobia. *Behaviour Research and Therapy, 33,* 321–324.

Juster, H. R., Heimberg, R. G., Frost, R. O., Holt, C. S., Mattia, J. I., & Faccenda, K. (1996). Social phobia and perfectionism. *Personality and Individual Differences, 21,* 403–410.

Juster, H. R., Heimberg, R. G., & Holt, C. S. (1996). Social phobia: Diagnostic issues and review of cognitive behavioral treatment strategies. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification* (Vol. 30, pp. 74–98). Pacific Grove, CA: Brooks/Cole.

Juster, H. R., Heimberg, R. G., & Mattia, J. I. (1993, November). *Judgment and prediction of response to cognitive-behavioral group therapy for social phobia.*

Poster presented at the annual meeting of the Association for the Advancement of Behavior Therapy, Atlanta, GA.

Kagan, J., Reznick, J. S., & Snidman, N. (1988). Biological bases of childhood shyness. *Science, 240,* 167–171.

Kanter, N. J., & Goldfried, M. R. (1979). Relative effectiveness of rational restructuring and self-control desensitization in the reduction of interpersonal anxiety. *Behavior Therapy, 10,* 472–490.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen. H.-U., & Kendler, K. S. (1994). Lifetime and 12-month prevalence of *DSM-III-R* psychiatric disorders in the United States: Results from the National Comorbidity Survey. *Archives of General Psychiatry, 51,* 8–19.

Klerman, G. L., & Weissman, M. M. (1989). Increasing rates of depression. *Journal of the American Medical Association, 261,* 2229–2235.

Kushner, M. G., Sher, K. J., & Beitman, B. D. (1990). The relation between alcohol problems and the anxiety disorders. *American Journal of Psychiatry, 147,* 685–695.

Leary, M. R., & Kowalski, R. M. (1995). The self-presentation model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 94–112). New York: Guilford Press.

Liberman, R. P., King, L. W., De Risi, W. J., & McCann, M. (1975). *Personal effectiveness: Guiding people to assert themselves and improve their social skills.* Champaign, IL: Research Press.

Liebowitz, M. R., Gorman, J. M., Fyer, A. J., & Klein, D. F. (1985). Social phobia: Review of a neglected anxiety disorder. *Archives of General Psychiatry, 42,* 729–736.

Liebowitz, M. R., Heimberg, R. G., Schneier, F. R., Hope, D. A., Davies, S., Holt, C. S., Goetz, D., Juster, H. R., Lin, S-H., Bruch, M. A., Marshall, R. D., & Klein, D. F. (1997). *Cognitive-behavioral group therapy versus phenelzine in social phobia: Long term outcome.* Manuscript submitted for publication.

Liebowitz, M. R., & Marshall R. D. (1995). Pharmacological treatments: Clinical applications. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 366–383). New York: Guilford Press.

Magee, W. J., Eaton, W. W., Wittchen, H.-U., McGonagle, K. A., & Kessler, R. C. (1996). Agoraphobia, simple phobia, and social phobia in the National Comorbidity Survey. *Archives of General Psychiatry, 53,* 159–168.

Mannuzza, S., Fyer, A. J., Liebowitz, M. R., & Klein, D. F. (1990). Delineating the boundaries of social phobia: Its relationship to panic disorder and agoraphobia. *Journal of Anxiety Disorders, 4,* 41–59.

Marks, I. M., & Gelder, M. G. (1966). Different ages of onset in varieties of phobia. *American Journal of Psychiatry, 123,* 218–221.

Marzillier, J. S., Lambert, C., & Kellet, J. (1976). A controlled evaluation of systematic desensitization and social skills training for socially inadequate psychiatric patients. *Behaviour Research and Therapy, 14,* 225–238.

Mattick, R. P., & Peters, L. (1988). Treatment of severe social phobia: Effects of guided exposure with and without cognitive restructuring. *Journal of Consulting and Clinical Psychology, 56,* 251–260.

Mattick, R. P., Peters, L., & Clarke, J. C. (1989). Exposure and cognitive restructuring for social phobia: A controlled study. *Behavior Therapy, 20,* 3–23.

Meichenbaum, D. (1985). *Stress inoculation training.* New York: Pergamon.

Mersch, P. P. A. (1995). The treatment of social phobia: The differential effectiveness of exposure *in vivo* and and integration of exposure *in vivo*, rational emotive therapy and social skills training. *Behaviour Research and Therapy, 33,* 259–269.

Mersch, P. P. A., Emmelkamp, P. M. G., Bögels, S. M., & van der Sleen, J. (1989). Social phobia: Individual response patterns and the effects of behavioral and cognitive interventions. *Behaviour Research and Therapy, 27,* 421–434.

Mersch, P. P. A., Emmelkamp, P. M. G., & Lips, C. (1991). Social phobia: Individual response patterns and the long-term effects of behavioral and cognitive interventions. A follow-up study. *Behaviour Research and Therapy, 29,* 357–362.

Mersch, P. P. A., Hildebrand, M., Lavy, E. H., Wessel, I., & van Hout, W. J. P. J. (1992). Somatic symptoms in social phobia: A treatment method based on rational emotive therapy and paradoxical interventions. *Journal of Behavior Therapy and Experimental Psychiatry, 23,* 199–211.

Myers, J. K., Weissman, M. M., Tischler, G. I., Holzer, C. E., III, Leaf, P. J., Orvaschel, H., Anthony, J. C., Boyd, J. H., Burke, J. D., Jr., Kramer, M., & Stolzman, R. (1984). Six-month prevalence of psychiatric disorders in three communities: 1980 to 1982. *Archives of General Psychiatry, 41,* 959–967.

Newman, M. G., Hofmann, S. G., Trabert, W., Roth, W. T., & Taylor, C. B. (1994). Does behavioral treatment of social phobia lead to cognitive changes? *Behavior Therapy, 25,* 503–517.

Öst, L.-G. (1987). Applied relaxation: Description of a coping technique and review of controlled studies. *Behaviour Research and Therapy, 25,* 397–409.

Öst, L.-G., Jerremalm, A., & Johansson, J. (1981). Individual response patterns and the effects of different behavioral methods in the treatment of social phobia. *Behaviour Research and Therapy, 19,* 1–16.

Potts, N. L. S., & Davidson, J. R. T. (1995). Pharmacological treatments: Literature review. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 334–365). New York: Guilford Press.

Rapee, R. M. (1995). Descriptive psychopathology of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 41–66). New York: Guilford Press.

Rapee, R. M., & Hayman, K. (1996). The effects of video feedback on the self-evaluation of performance in socially anxious subjects. *Behaviour Research and Therapy, 34,* 315–322.

Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy, 35,* 741–756.

Rapee, R. M., & Lim, L. (1992). Discrepancy between self- and observer ratings of performance in social phobics. *Journal of Abnormal Psychology, 101,* 727–731.

Reich, J., Goldenberg, I., Vasile, R., Goisman, R., & Keller, M. (1994). A prospective follow-along study of the course of social phobia. *Psychiatry Research, 54,* 249–258.

Rosenbaum, J. F., Biederman, J., Hirshfeld, D. R., Bolduc, E. A., Faraone, S. V., Kagan, J., Snidman, N., & Reznick, J. S. (1991). Further evidence of an association between behavioral inhibition and anxiety disorders: Results from a family study of children from a non-clinical sample. *Journal of Psychiatric Research, 25,* 49–65.

Rosenbaum, J. F., Biederman, J., Pollock, R. A., & Hirshfeld, D. R. (1994). The etiology of social phobia. *Journal of Clinical Psychiatry, 55*(Suppl. 6), 10–16.

Ross, H. E., Glaser, F. B., & Germanson, T. (1988). The prevalence of psychiatric disorders in patients with alcohol and other drug problems. *Archives of General Psychiatry, 45,* 1023–1031.

Safren, S. A., Heimberg, R. G., Brown, E. J., & Holle, C. (1997). Quality of life in social phobia. *Depression and Anxiety, 4,* 126–133.

Safren, S. A., Heimberg, R. G., & Juster, H. R. (1997).

Client expectancies and their relationship to pretreatment symptomatology and outcome of cognitive behavioral group treatment for social phobia. *Journal of Consulting and Clinical Psychology, 65,* 694–698.

Schneier, F. R., Heckelman, L. R., Garfinkel, R., Campeas, R., Fallon, B. A., Gitow, A., Street, L., Del Bene, D., & Liebowitz, M. R. (1994). Functional impairment in social phobia. *Journal of Clinical Psychiatry, 55,* 322–331.

Schneier, F. R., Johnson, J., Hornig, C. D., Liebowitz, M. R., & Weissman, M. W. (1992). Social phobia: Comorbidity and morbidity in an epidemiologic sample. *Archives of General Psychiatry, 49,* 282–288.

Schneier, F. R., Spitzer, R. L., Gibbon, M., Fyer, A. J., & Liebowitz, M. R. (1991). The relationship of social phobia subtypes and avoidant personality disorder. *Comprehensive Psychiatry, 32,* 496–502.

Scholing, A., & Emmelkamp, P. M. G. (1993a). Cognitive and behavioural treatments of fear of blushing, sweating, or trembling. *Behaviour Research and Therapy, 31,* 155–170.

Scholing, A., & Emmelkamp, P. M. G. (1993b). Exposure with and without cognitive therapy for generalized social phobia: Effects of individual and group treatment. *Behaviour Research and Therapy, 31,* 667–681.

Scholing, A., & Emmelkamp, P. M. G. (1996). Treatment of generalized social phobia: Results at long-term follow-up. *Behaviour Research and Therapy, 34,* 447–452.

Seligman, M. E. P. (1995). The effectiveness of psychotherapy: The Consumer Reports study. *American Psychologist, 50,* 965–974.

Shaw, P. M. (1979). A comparison of three behaviour therapies in the treatment of social phobia. *British Journal of Psychiatry, 134,* 620–623.

Smail, P., Stockwell, T., Canter, S., & Hodgson, R. (1984). Alcohol dependence and phobic anxiety states. *British Journal of Addiction, 82,* 971–979.

Spitzer, R. L., & Williams, J. B. W. (1985). Proposed revisions in the DSM-III classification of anxiety disorders based on research and clinical experience. In A. H. Tuma & J. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 759–773). Hillsdale. NJ: Erlbaum.

Stein, M. B. (Ed.) (1995). *Social phobia: Clinical and research perspectives*. Washington, DC: American Psychiatric Press.

Stein, M. B., Walker, J. R., & Forde, D. R. (1994). Setting diagnostic thresholds for social phobia: Considerations from a community survey of social anxiety. *American Journal of Psychiatry, 151,* 408–412.

Stravynski, A. (1983). Behavioral treatment of psychogenic vomiting in the context of social phobia. *British Journal of Psychiatry, 134,* 620–623.

Stravynski, A., Marks, I., & Yule, W. (1982). Social skills problems in neurotic outpatients: Social skills training with and without cognitive modification. *Archives of General Psychiatry, 39,* 1378–1385.

Suinn, R. M., & Richardson, F. (1971). Anxiety management training: A nonspecific behavior therapy program for anxiety control. *Behavior Therapy, 2,* 498–510.

Taylor, S. (1996). Meta-analysis of cognitive-behavioral treatments for social phobia. *Journal of Behavior Therapy and Experimental Psychiatry, 27,* 1–9.

Taylor, S., Woody, S., Koch, W., McLean, P., Paterson, R., & Anderson, K. W. (1997). Cognitive restructuring in the treatment of social phobia: Efficacy and mode of action. *Behavior Modification, 21,* 487–511.

Tran, G. Q., & Chambless, D. L. (1995). Subtypes of social phobia and avoidant personality disorder. *Journal of Anxiety Disorders, 9,* 489–501.

Trower, P., Yardley, K., Bryant, B., & Shaw, P. (1978). The treatment of social failure: A comparison of anxiety-reduction and skills acquisition procedures on two social problems. *Behavior Modification, 2,* 41–60.

Trower, P., & Gilbert, P. (1989). New theoretical conceptions of social anxiety and social phobia. *Clinical Psychology Review, 9,* 19–35.

Turk, C. L., Fresco, D. M., & Heimberg, R. G. (in press). Social phobia: Cognitive behavior therapy. In M. Hersen & A. S. Bellack (Eds.), *Handbook of comparative treatments of adult disorders* (2nd ed.). New York: Wiley.

Turk, C. L., Heimberg, R. G., Orsillo, S. M., Holt, C. S., Gitow, A., Street, L. L., Schneier, F. R., & Liebowitz, M. R. (in press). An investigation of gender differences in social phobia. *Journal of Anxiety Disorders.*

Turner, S. M., Beidel, D. C., Cooley, M. R., Woody, S., & Messer, S. C. (1994). A multi-component behavioral treatment for social phobia: Social Effectiveness Therapy. *Behaviour Research and Therapy, 32,* 381–390.

Turner, S. M., Beidel, D. C., & Cooley-Quille, M. R. (1995). Two-year follow-up of social phobics treated with Social Effectiveness Therapy. *Behaviour Research and Therapy, 33,* 553–555.

Turner, S. M., Beidel, D. C., Dancu, C. V., & Keys, D. J. (1986). Psychopathology of social phobia and comparison to avoidant personality disorder. *Journal of Abnormal Psychology, 95,* 389–394.

Turner, S. M., Beidel, D. C., Long, P. J., Turner, M. W., & Townsley, R. M. (1993). A composite measure to determine the functional status of treated social phobics: The Social Phobia Endstate Functioning Index. *Behavior Therapy, 24,* 265–275.

Turner, S. M., Beidel, D. C., & Townsley, R. M. (1990). Social phobia: Relationship to shyness. *Behaviour Research and Therapy, 28,* 497–505.

Turner, S. M., Beidel, D. C., & Townsley, R. M. (1992). Social phobia: A comparison of specific and generalized subtypes and avoidant personality disorder. *Journal of Abnormal Psychology, 101,* 326–331.

Turner, S. M., Beidel, D. C., Wolff, P. L., Spaulding, S. & Jacob, R. G. (1996). Clinical features affecting treatment outcome in social phobia. *Behaviour Research and Therapy, 34,* 795–804.

VandenBos, G. R. (Ed.) (1996). Outcome assessment of psychotherapy (Special issue). *American Psychologist, 51*(10).

Wacker, H. R., Mullejans, R., Klein, K. H., & Battegay, R. (1992). Identification of cases of anxiety disorders and affective disorders in the community according to *ICD-10* and *DSM-III-R* by using the Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research, 2,* 91–100.

Wickramaratne, P. J., Weissman, M. M., Leaf, D. J., & Holford, T. R. (1989). Age, period, and cohort effects on the risk of major depression: Results from five United States communities. *Journal of Clinical Epidemiology, 42,* 333–343.

Widiger, T. A. (1992). Generalized social phobia versus avoidant personality disorder: A commentary on three studies. *Journal of Abnormal Psychology, 101,* 340–343.

Wlazlo, Z., Schroeder-Hartwig, K., Hand, I., Kaiser, G., & Münchau, N. (1990). Exposure *in vivo* vs. social skills training for social phobia: Long-term outcome and differential effects. *Behaviour Research and Therapy, 28,* 181–193.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition.* Palo Alto, CA: Stanford University Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.22

Post-traumatic Stress Disorder

LISA H. JAYCOX and EDNA B. FOA

Allegheny University of the Health Sciences, Philadelphia, PA, USA

6.22.1 DEFINITIONS	499
6.22.2 SCOPE OF THE PROBLEM	500
6.22.3 RISK FACTORS FOR CHRONIC POST-TRAUMA PSYCHOPATHOLOGY	500
6.22.4 THEORIES ABOUT PTSD	501
6.22.4.1 Psychodynamic Theories	501
6.22.4.2 Cognitive-behavioral Theories	502
6.22.4.2.1 Learning theory	502
6.22.4.2.2 Cognitive theories	502
6.22.4.2.3 Emotional processing theory	502
6.22.5 EFFICACY OF PSYCHOSOCIAL TREATMENTS FOR PTSD	503
6.22.5.1 Hypnotherapy	503
6.22.5.2 Psychodynamic Treatments	503
6.22.5.3 Cognitive-behavioral Therapy	504
6.22.5.4 Studies of Imaginal and In Vivo *Exposure*	504
6.22.5.4.1 Studies of eye movement desensitization and reprocessing	505
6.22.5.4.2 Studies of cognitive therapy and anxiety management training	506
6.22.5.4.3 Studies of combination therapies	506
6.22.6 PREVENTION PROGRAMS	507
6.22.7 PROLONGED EXPOSURE THERAPY FOR PTSD: THEORY AND PRACTICE	508
6.22.7.1 Mechanisms Underlying Prolonged Exposure	508
6.22.7.2 How To Implement Exposure-based Therapy	509
6.22.7.2.1 Information gathering	509
6.22.7.2.2 Breathing retraining	509
6.22.7.2.3 Education about symptoms	510
6.22.7.2.4 In vivo *exposure*	510
6.22.7.2.5 Imaginal exposure	511
6.22.7.3 Indications and Contra-indications	512
6.22.7.3.1 Anger	513
6.22.7.3.2 Emotional numbing	513
6.22.7.3.3 Overwhelming anxiety	513
6.22.8 REFERENCES	514

6.22.1 DEFINITIONS

Post-traumatic stress disorder (PTSD) is an anxiety disorder that consists of a sustained and dysfunctional emotional reaction to an extreme stressor. The revised third edition of the *Diagnostic and statistical manual of mental disorders* (*DSM-III-R*; American Psychiatric Association [APA], 1987) defined a trauma as an event that is outside the range of normal human experience and that would be markedly distressing to almost anyone. However, a normative definition of a trauma proved inadequate as epidemiological studies found high prevalence rates of traumatic experiences in segments of the general population (e.g., Resnick, Kilpatrick, Dansky, Saunders, & Best, 1993). Therefore, in *DSM-IV* (APA, 1994) the criteria were modified to include two aspects: the objective characteristics of the stressor, and the subjective experiences of the victim. Specifically, the traumatized person must have experienced or witnessed a situation involving actual or threatened bodily injury or death, or threat to physical integrity, and must have felt helpless, horrified, or terrified during the experience.

According to *DSM-IV*, PTSD symptoms fall into three general clusters: re-experiencing, avoidance, and arousal. To be eligible for a diagnosis of PTSD, three criteria must be met. First, the individual must re-experience the trauma in one of the following ways: nightmares, flashbacks, intrusive thoughts, and emotional distress or physiological arousal in response to internal or external cues that serve as reminders of the trauma. Second, the individual must have at least three avoidance symptoms: avoidance of thinking about the trauma, avoidance of reminders of the event, emotional numbing, detachment or distance from other people, less interest in activities, psychogenic amnesia, or a sense of a foreshortened future. Finally, individuals must experience at least two of the following arousal symptoms: sleep disturbance, hypervigilance, exaggerated startle response, irritability or outbursts of anger, or difficulty concentrating. If these symptoms persist for at least one month and cause considerable impairment in daily living, a diagnosis of PTSD is appropriate. In addition to the anxiety symptoms of PTSD itself, there is also high comorbidity between PTSD and depressive symptoms, other anxiety disorders, and substance abuse (Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995), and these can exacerbate dysfunction in daily living.

Although PTSD cannot be formally diagnosed unless symptoms persist for at least one month, many individuals show marked symptoms and impairment beginning immediately after a trauma and may benefit from some psychological intervention. Interest in this initial reaction to trauma gave impetus to the introduction of a new disorder called "acute stress disorder" into *DSM-IV* (APA, 1994). This disorder includes symptoms of PTSD with a special emphasis on dissociative symptoms and lasts anywhere from two days to four weeks post-trauma.

6.22.2 SCOPE OF THE PROBLEM

Both the prevalence of traumatic experiences and the rate of PTSD are alarming. Approximately 39% of people in the USA have experienced at least one traumatic event in their lifetime. Lifetime prevalence of PTSD is estimated at 24% among trauma victims and at 9% in the general population, with up to a third (3.4%) of these cases having chronic PTSD (Breslau, Davis, Andreski, & Peterson, 1991). In addition to individuals who meet full criteria for PTSD, up to 15% of the general population suffer from subthreshold symptoms (Davidson, Hughes, Blazer, & George, 1991).

Women appear to be somewhat more likely than men to develop PTSD following trauma (10.4% and 5%, respectively; Kessler et al., 1995). In a national sample of women, 69% had experienced some type of trauma, and 36% had been victims of crime: sexual assault, aggravated assault, or the homicide of a close friend or relative. The lifetime prevalence of PTSD in this sample was 12.9%, with six month prevalence of 4.6% (Resnick et al., 1993).

The experience of trauma according to *DSM-IV* includes not only the direct experience of a trauma, but also witnessing or learning about a traumatic event (APA, 1994). Symptoms can therefore develop in those who have little or no direct exposure to the trauma, although the symptom severity is directly related to degree of exposure (March, 1992). For example, ferry workers who were *not* among the crew of the capsized *Herald of Free Enterprise* ferry manifested PTSD symptoms, presumably because they identified strongly with co-workers who were in the ferry disaster (Dixon, Rehling, & Shiwach, 1993). This phenomenon has been referred to as the "ripples outward" effect.

Certain professionals are at risk for PTSD by virtue of increased probability of repeated direct exposure to trauma. In one study, 26% of police officers reported symptoms severe enough to meet symptom criteria of PTSD (Martin, McKean, & Veltkamp, 1986). Firefighters, emergency medical technicians, emergency room staff, and disaster workers are also thought to be at risk. For example 30% of firefighters of an Australian bushfire had chronic PTSD 29 months after the event (McFarlane, 1989).

6.22.3 RISK FACTORS FOR CHRONIC POST-TRAUMA PSYCHOPATHOLOGY

Factors related to risk of developing chronic disturbance may be divided into three types: pretrauma factors, factors related to the trauma itself, and post-trauma factors. We will briefly review research relevant to each of these factors.

Research on demographic variables has not identified reliable predictors of who will develop chronic PTSD (e.g., Foa & Riggs, 1993; Green, 1994). In contrast, it appears that poor psychological and social functioning prior to the trauma renders the individual vulnerable to developing chronic disturbances (Ruch & Leon, 1983). Also, a history of prior traumatic events in childhood or adulthood augments the response to subsequent traumas (Ellis, Atkeson, & Calhoun, 1981; Riggs, Foa, Rothbaum, & Murdock, 1992; Roth, Waylan, & Woolsey, 1990; Ruch & Leon, 1983).

It is safe to state that traumas differ in their likelihood of producing PTSD. In a retrospective study mentioned earlier, Resnick et al. (1993) found that 17.8% of female victims of aggravated assault and 12.4% of rape victims had current PTSD. In contrast, only 3.4% of female victims of noncrime trauma had PTSD. Three prospective studies of female victims of rape and nonsexual assault revealed higher rates of PTSD among the former at each data point up to three month postassault (Jaycox et al., 1996; Riggs, Rothbaum, & Foa, 1995; Rothbaum, Foa, Riggs, Murdock, & Walsh, 1992). The severity of the trauma also appears to influence subsequent severity of PTSD: the degree of combat exposure among Vietnam veterans predicted pathology (Foy, Sipprelle, Rueger, & Carroll, 1984; Pitman, Altman, & Macklin, 1989), and the perception of imminent life threat appears to predict post-trauma pathology among rape victims (Girelli, Resick, Marhoefer-Dvorak, & Hutter, 1986; Kilpatrick et al., 1989; Resick, 1986).

Just as traumas differ in their likelihood to produce PTSD, so do individuals differ in their reaction to a traumatic experience. Several post-trauma factors have been found to exacerbate symptoms. Dissociative experiences during the traumatic encounter or immediately after have been found to predict later distress or PTSD (e.g., Bremner et al., 1992; Koopman, Classen, & Spiegel, 1994; Marmar et al., 1994, Nishith, Mechanic, Griffin, & Resick, 1995; Tichenor, Marmar, Weiss, Metzler, & Ronfeldt, 1996). Also, on average, assault victims who exhibit more severe initial reactions to the trauma evidence more symptoms later on (Perry, DiFede, Musgni, Frances, & Jacobsberg, 1992; Riggs et al., 1995; Rothbaum et al., 1992; Shalev, Peri, Canetti, & Schreiber, 1996) although delayed peak PTSD symptoms were associated with more severe pathology compared with early peak (Gilboa & Foa, 1997). Evidence on the role of social support as facilitating or hindering recovery is equivocal. Indeed, empirical studies indicate that negative social interactions (e.g., victim blame or disbelief) have strong negative effects on victim adjustment, whereas positive reactions from others have little impact on adjustment (Davis, Brickman, & Baker, 1991; Ullman, 1995). Excessive anger or guilt following the trauma also appear to block readjustment (Amir, Foa, & Cashman, 1997; Riggs, Dancu, Gershuny, Greenberg, & Foa, 1992).

Thus, a number of factors must be taken into account in evaluating a trauma victim, including the nature of the trauma, the victim's prior psychological adjustment and trauma history, and the level of distress in the aftermath of the trauma. As will be discussed further in the sections on theory, the meaning of the trauma and changes in beliefs about the self and world are also important factors to consider.

6.22.4 THEORIES ABOUT PTSD

For over 100 years, reactions to trauma have captured the interest of theorists of psychopathology. Janet (1889) was among the first to develop an integrated theory about trauma, suggesting extreme stressors beseige the individual with an excess of thoughts and feelings too numerous or intense to integrate. Consequently, some individuals selectively attend away from the trauma, focusing instead on trauma-irrelevant thoughts and feelings. In this way certain trauma-related ideas become split off (dissociated) from consciousness, where they cause fragmented reliving of the trauma in the form of visual images, somatic states, emotional disturbances, or behavioral re-enactment. As will be evident in the discussion below, Janet's ideas about reactions to trauma have influenced both early and contemporary theorists.

6.22.4.1 Psychodynamic Theories

Freud's thoughts about the effects of trauma on the psyche have shifted over time (van der Kolk, Weisaeth, & van der Hart, 1996). In his early writing Freud was highly influenced by Janet, elaborating the dissociation phenomenon (Breuer & Freud, 1955). He later abandoned the dissociation theory, proposing instead that trauma reactions persist because of an association between the traumatic event and childhood repressed conflicts (Freud, 1953), thus rejecting the possibility that a traumatic event alone can cause severe emotional impact. Confronted by the grim effects on people of World War I, Freud focused again on external reality and returned to the view that the emotional upheaval generated by trauma is the source of traumatic neurosis. He suggested that the intensity of the trauma, the inability to find conscious expressions for it, and the unpreparedness of the individual cause a breach to the stimulus barrier and overwhelm the defense mechanisms (Freud, 1955).

6.22.4.2 Cognitive-behavioral Theories

Current cognitive-behavioral theory about the development and maintenance of PTSD stems from two lines of thinking. First is learning theory, which includes both classical and operant conditioning principles. Second is cognitive or schema theories, which propose evaluative processes as mediators in the development and maintenance of pathological anxiety. We will discuss how post-trauma psychopathology is conceived of within these two theoretical approaches and then present a third approach, emotional processing theory, that incorporates elements of the two within an information processing framework.

6.22.4.2.1 Learning theory

Mowrer's two-factor theory (1960) has been the most influential in the explanation of the pathological anxiety seen in PTSD. In this theory, fear is acquired through classical conditioning, in which a neutral stimulus (CS) is paired with an aversive stimulus (UCS), so that the CS comes to elicit a conditioned fear response (CR). Several theorists invoked this theory to explain the symptoms of PTSD (e.g., Foa, Steketee, & Rothbaum, 1989; Keane, Fairbank, Caddell, Zimering, & Bender, 1985; Kilpatrick et al., 1985), such that previously neutral stimuli that were present during the trauma come to elicit anxiety themselves. Through generalization and second-order conditioning, additional stimuli that are associated with either the feared stimuli or neutral stimuli that were present during the trauma also come to evoke fear. For example, words, thoughts, and images acquire the capacity to cause anxiety.

Subsequently, avoidance behavior is established through the process of operant conditioning. In this process, an individual learns to reduce trauma-related anxiety though avoidance of, or escape from, the CS. Escape and avoidance behaviors become established through the process of negative reinforcement, via their predicted capacity to end the aversive fear state. Because avoidance obstructs the realization that the CS has ceased to be followed by the UCS, fear is maintained.

6.22.4.2.2 Cognitive theories

Even during the influence of Mowrer's theory, discontent with non mediational accounts of anxiety stimulated the development of cognitive approaches. However, the originators of cognitive theory for anxiety disorders gave little attention to PTSD itself. They did suggest, though, that people with traumatic neuroses do not discriminate between safe and unsafe signals, and that their "thinking is dominated by the concept of danger." They also proposed that traumatic fear can be maintained through a sense of incompetence to handle stressful events (Beck, Emery, & Greenberg, 1985).

More specific hypotheses about cognitive factors that mediate emotional responses to trauma were advanced by scholars who employed personality and social psychology theories (e.g., Epstein, 1991; Janoff-Bulman, 1992; McCann & Pearlman, 1990). Central to their theories is the concept of schemas: core assumptions and beliefs that guide the perception and interpretation of incoming information. These theories share the supposition that processing a traumatic experience requires modification of existing assumptions. Based on Piaget's (1971) model of cognitive development, it is thought that such modification is accomplished through two mechanisms: assimilation and accommodation.

Both Epstein (1991) and Janoff-Bulman (1992) hold the position that people in general assume that the world is benevolent, the world is meaningful, the self is worthy, and that these assumptions are incompatible with a traumatic experience. Therefore, the victim must struggle either to assimilate the traumatic experience into the old set of assumptions or, more often, to change the assumptions such that they can accommodate the traumatic experience. Expanding on this view, McCann and Pearlman (1990) proposed seven fundamental psychological need areas: frame of reference, safety, dependency/trust of self and others, power, esteem, intimacy, and independence. Trauma, they argue, may cause disruptions in any of these need areas and thereby lead to troublesome emotions and thoughts or images.

Drawing on both psychoanalytic and information processing concepts, Horowitz (1986) suggested that individuals have a basic need to match trauma-related information with their "inner models based on old information" (p. 92). Recovery from trauma is a process of repetitive "revision of both [sources of information] until they agree" (p. 92), called the completion tendency, which is reflected in the re-experiencing symptoms observed in individuals with PTSD.

6.22.4.2.3 Emotional processing theory

Foa and her colleagues (Foa & Jaycox, in press; Foa & Riggs, 1993; Foa et al., 1989; Foa, Zinbarg, & Rothbaum, 1992) advanced a theory of PTSD that integrates learning, cognitive, and personality theories within an information processing framework, to explain why some individuals recover from a traumatic experience while others develop chronic disturbances. The starting point is the observation that emotional experiences are often relived well after the original emotional events have occurred, although the frequency and the intensity of emotional reliving usually decreases over time (Foa & Kozak, 1991). When these natural recovery processes are impeded, psychopathology emerges. To explain this pathological anxiety, Foa and Kozak (1985, 1986) adopted Lang's (1979) conceptual framework in which fear is viewed as a cognitive structure that serves as a program for escaping danger. Accordingly, fear is represented as a cognitive structure that includes three kinds of information: information about the feared stimulus; information about verbal, physiological, and overt behavioral responses; and interpretive information about the meaning of the stimulus and response elements of the structure. A fear structure is distinguished by the information that stimuli and/or responses are dangerous. In a subsequent exposition of emotional processing theory, Foa and Kozak, (1991) took the position that it is essentially this meaning information that distinguishes the fear structure from other information structures. With regard to PTSD, Foa and her colleagues (Foa & Jaycox, in press; Foa & Riggs, 1993) suggested that two classes of pathological cognitions underlie this disorder are: the world is utterly dangerous and the self is completely incompetent; a corollary of these is the idea that persistent PTSD symptoms are dangerous and are indicators of self-incompetence.

6.22.5 EFFICACY OF PSYCHOSOCIAL TREATMENTS FOR PTSD

Several types of psychosocial treatment have been employed for post-trauma disturbances, including hypnotherapy, psychodynamic psychotherapy, and cognitive-behavioral therapies. However, most outcome studies have focused on examining the efficacy of the latter (see Foa & Meadows, 1997, for a critical review). This section will review these studies.

6.22.5.1 Hypnotherapy

Hypnosis has long been used in the treatment of post-trauma disturbances (Spiegel, 1989). Freud used hypnosis to facilitate abreaction and catharsis, which he felt were necessary for recovery. Spiegel advocates the use of hypnosis in treating PTSD because processes akin to hypnotic trance, such as dissociation, occur naturally and are commonly used during or after a trauma. Spiegel proposes that hypnosis may facilitate the remembering of traumatic experiences that were encoded in a dissociative state.

Several case reports have described the usefulness of hypnosis for individuals with PTSD stemming from a variety of traumas (e.g., Jiranek, 1993; Kingsbury, 1988; Leung, 1994; MacHovec, 1983; Peebles, 1989; Spiegel, 1988, 1989). However, most of these lack methodological rigor and therefore the conclusions that can be drawn from them about the efficacy of hypnosis are limited.

One large controlled study compared hypnosis, desensitization, and psychodynamic psychotherapy with a wait-list control group among individuals who had experienced stressors (Brom, Kleber, & Defres, 1989). The majority of the participants, however, did not directly experience the traumatic event, but rather had lost a loved one. All three conditions produced superior improvement to the wait-list condition on self-report measures, but no differences across the three treatments were observed. Specifically, improvement in post-trauma symptoms was 29% for psychodynamic therapy, 34% for hypnotherapy, and 41% for desensitization, compared with about 10% improvement in the wait-list condition. This study offers evidence for the efficacy of hypnosis and other psychosocial therapies in reducing bereavement-related symptoms.

6.22.5.2 Psychodynamic Treatments

A variety of psychodynamic concepts have been utilized in individual and group therapy for trauma survivors (e.g., Horowitz, 1976; Yalom, 1995). For example, Horowitz's brief dynamic therapy emphasizes the concepts of denial, abreaction, catharsis, and stages of recovery from trauma in the treatment for post-trauma difficulties. The target of this brief psychodynamic therapy is the resolution of the intrapsychic conflict arising from the traumatic experience.

Only a few studies have been conducted to evaluate psychodynamic interventions for PTSD, and many have methodological difficulties that make their results difficult to interpret. One study used brief psychodynamic therapy to address difficulties following the Beverly Hills Supper Club fire (Lindy, Green, Grace, & Titchener, 1983). In this study, survivors included those present at the scene, rescue workers, people who lost significant others, and individuals who identified the bodies. Only 9 of 30 met criteria for PTSD prior to treatment. Survivors who agreed to be in the research study but did not request psychotherapy served as the comparison group. Patients who completed treatment exhibited somewhat lower symptoms than did patients who did not complete treatment and the non-treated comparison group did not improve as much as treatment completers.

Another study examined the efficacy of a group therapy based on Horowitz's (1976) model of responses to trauma with 13 sexual assault victims (Roth, Dye, & Lebowitz, 1988). The control group comprised 13 women who agreed to undergo assessements, and random assignment was not used. There were no clear effects at post-treatment, but high attrition rates in both the treated and control groups may have obscured effects. Results over a longer time period showed greater improvement in the therapy. Again, methodological problems render these results inconclusive. As described above, a study comparing Horowitz's (1976) brief psychodynamic therapy did as well as hypnosis, and desensitization; all did better than the wait-list control group (Brom et al., 1989).

In a study of psychodynamic treatment for conjugal bereavement, Marmar, Horowitz, Weiss, Wilner, & Kaltreider (1988) randomized bereaved individuals (widows) to 12 weekly sessions of either brief dynamic therapy or mutual help group treatment. Although death of a husband does not necessarily qualify as a *DSM-IV* trauma, some of the women in this study were diagnosed as having PTSD. Results indicated that patients in both conditions improved slightly, but there were no differences between groups. On one of many measures (the total score of the Symptom checklist, SCL-90) the dynamic therapy conditions showed greater reduction in symptoms at follow-up.

Scarvalone, Cloitre, & Difede (1995) used a quasi-experimental design to compare interpersonal process group therapy (IPGT) with a naturally occurring wait-list control in a sample of female childhood sexual abuse survivors, many of whom were diagnosed with PTSD. Patients in the IPGT group improved on a number of measures and fewer evidenced PTSD at post-treatment than in the wait-list group. However, because the evaluators were not blind to treatment condition, the extent to which expectancy for improvement influenced the evaluation is unknown.

6.22.5.3 Cognitive-behavioral Therapy

As noted earlier, cognitive-behavioral treatments have been evaluated far more extensively than have other psychosocial treatments for PTSD, via case reports and well-controlled studies. In this section, we will focus on controlled studies only that evaluate the efficacy of cognitive-behavioral programs developed to ameliorate PTSD symptoms: exposure therapy, anxiety management training, cognitive therapy, and their combinations.

6.22.5.4 Studies of Imaginal and *In Vivo* Exposure

Exposure therapy is a set of procedures that involve confrontation with feared stimuli, either *in vivo* or in imagination (e.g., flooding, systematic desensitization), which will be described in detail in the next section. With PTSD, exposure therapy typically includes repeated reliving of the traumatic event (in imagination) and *in vivo* confrontation with trauma-related situations that evoke fear but are not realistically dangerous.

Three controlled studies of male Vietnam veterans, each examining 6–16 sessions of imaginal exposure to the traumatic experience, have been conducted. In one study, patients were given up to nine sessions of imaginal flooding exposure in addition to "standard" treatment, or standard treatment alone (weekly individual and group therapies) (Cooper & Clum, 1989). In the second study, imaginal exposure (45 minutes of exposure preceded by relaxation) was compared with a no-treatment wait-list condition (Keane, Fairbank, Caddell, & Zimering, 1989). In the third study, all patients received group treatment. In addition to group treatment, half the patients received weekly imaginal exposure and the other half received weekly individual traditional psychotherapy (Boudewyns & Hyer, 1990; Boudewyns, Hyer, Woods, Harrison, & McCranie, 1990).

These three studies found that imaginal exposure produced some improvement in PTSD compared with the control groups, but the effects were rather small. In the Cooper and Clum (1989) study, imaginal exposure improved the PTSD symptoms, but had little effect on depression or trait anxiety. In the Keane et al. (1989) study a mixed picture emerged: therapists rated patients who received exposure as more improved on PTSD symptoms than control subjects, but on self-report measures of PTSD symptoms, no group differences emerged. However, patients who received exposure did rate themselves as more improved on general psychopathology measures than did those in the wait-list control. Boudewyns and Hyer (1990) found no group differences on psychophysiological measures, but the exposure group had improved more than controls on the Veterans Adjustment Scale at the three month follow-up.

Two completed studies of female assault victims examined the efficacy of exposure treatments for PTSD in comparison to other treatments and to wait-list controls. The assaults targeted for treatment occurred in all cases at least three months prior to treatment, to control for the effect of natural decline in symptoms that occurs shortly following an assault (Rothbaum et al., 1992). Exposure treatment consisted of nine twice-weekly individual sessions, seven of which included active exposure exercises and two of which were devoted to information gathering and education.

In the first study (Foa, Rothbaum, Riggs, & Murdock, 1991), exposure treatment (PE) was compared with stress inoculation training (SIT, which will be described later), to supportive counseling (SC), and to a wait-list control. Immediately post-treatment, patients in the PE and the SIT groups improved on all three clusters of PTSD symptoms. Patients in the SC and wait-list conditions evidenced improvement on arousal symptoms, but not on avoidance or re-experiencing symptoms. A follow-up evaluation that was conducted approximately three months after treatment revealed no differences between the active treatments; however, on all measures of psychopathology the PE group had less pathology. At follow-up, 55% of those receiving PE, 55% of patients who received SIT, and 45% of those receiving SC did not meet criteria for PTSD.

The second study compared PE, SIT, and a group treated with the combination of SIT and PE with one another and with a wait-list control group. All three active treatments showed significant improvement in PTSD symptoms, state anxiety, and depressive symptoms at post-test, and the wait-list did not improve. These treatment effects were maintained at the 3, 6, and 12 month follow-ups. On an index of good end-state functioning (defined as 50% improvement in PTSD symptoms, Beck Depression Inventory less than 9, and State-Trait Anxiety Inventory less than 40), 50% of patients who received PE achieved this criterion, whereas 37% and 38% of patients who received SIT and SIT/PE achieved the criterion, respectively; none of the patients in the control group reached this status (Foa, Dancu, Hembree, Jaycox, & Meadows, 1997). Moreover, those that received PE (PE or SIT/PE) improved more on all symptoms measures as compared with to those who did not receive PE (SIT or wait-list).

Two additional studies with female assault victims are in progress. Resick and colleagues (Resick, Nishith, & Astin, 1996) are examining the relative efficacy of PE and cognitive processing therapy (CPT), a therapy that will be described later when we discuss combined treatments. At post-treatment and at follow-up, both interventions resulted in large reduction of PTSD symptoms, but no clear group differences have emerged from the preliminary data. The second ongoing study (Foa, Jaycox, Meadows, Hembree, & Dancu, 1996) compares 9–12 sessions of PE alone with a program that includes PE and cognitive restructuring. Again, preliminary analysis suggests that both programs were highly effective but, unexpectedly, PE alone appears slightly more effective than the combined treatment.

One study that utilized a cross-over design examined the efficacy of PE with victims of diverse traumas, comparing the efficacy of *in vivo* and imaginal exposure by randomly assigning 14 patients with PTSD to one of two conditions: four sessions of *in vivo* exposure followed by four sessions of imaginal exposure; or eight sessions in the reverse order (Richards, Lovell, & Marks, 1994). Considerable improvement was observed in both groups, with symptom reduction between 65% and 80% and all patients losing the diagnosis of PTSD at post-test and at one-year follow-up. Results indicated that phobic avoidance was more reduced by *in vivo* than by imaginal exposure, regardless of the order in which it was received.

6.22.5.4.1 Studies of eye movement desensitization and reprocessing

A recently developed exposure technique called "eye movement desensitization and reprocessing" (EMDR; Shapiro, 1995), which combines imaginal exposure with saccadic eye movements, is beginning to be examined in empirical studies. However, to date, many of the outcome studies conducted have been plagued with methodological problems, making them difficult to interpret.

Several studies have compared EMDR to standard treatments in veterans, but none used independent evaluations of symptoms (e.g., Boudewyns, Stwertka, Hyer, Albrecht, & Sperr, 1993; Jensen, 1994). Another study that utilized a mixed sample of traumatized individuals, only some of whom had PTSD, showed that victims who received EMDR reported decreases in presenting complaints and in anxiety at post-test, whereas the wait-list group reported no such changes (Wilson, Becker, & Tinker, 1995). In an attempt to evaluate the effect of the saccadic eye movements, Pitman et al. (1996) compared EMDR to eyes-fixed exposure in a cross-over design using male veterans. No differences were found between EMDR and the eyes-fixed exposure condition on self-reported distress within sessions, with both groups showing modest gains on self-reported symptoms. Because these studies all relied solely on self-report data, the results may reflect expectancies for improvement after treatment rather than specific effects of EMDR.

Two studies did use independent evaluations of PTSD symptoms. The first used victims of heterogeneous traumas to compare EMDR with imagery habituation training and with applied muscle relaxation training (Vaughan et al., 1994). All three groups improved equally on independent ratings of PTSD symptoms after treatment. The second controlled study utilized female assault victims who had PTSD, and found EMDR to be more effective compared with a wait-list control group (Rothbaum, 1995). After four weekly sessions of EMDR, assessor's ratings of PTSD indicated 57% symptom reduction, and a self-report measure indicated 74% symptom reduction, and these gains were maintained at a three month follow-up.

In summary, the two controlled studies that used blind assessment provide preliminary evidence that EMDR is effective in the treatment of PTSD, but its value over and above other more established treatments is as yet unknown. Further studies about the efficacy of EMDR are clearly needed.

6.22.5.4.2 Studies of cognitive therapy and anxiety management training

Cognitive techniques aim at teaching patients systematically to examine and challenge maladaptive cognitions which are thought to underlie negative emotions. Early work with rape victims used cognitive therapy to address post-rape anxiety and depression, but these studies did not evaluate PTSD symptoms *per se* (Frank et al., 1988; Frank & Stewart, 1984; Turner & Frank, 1981). The therapy produced positive change on ratings of fear, anxiety, depression, and social adjustment. However, some patients had been recently raped, and therefore their improvement could reflect natural recovery. Although the effects of cognitive therapy on PTSD has not yet been studied, this therapy is often included as a component of anxiety management programs.

Anxiety management training (AMT) provides patients with a repertoire of strategies to handle anxiety. These include relaxation training (Bernstein & Borkovec, 1973; Jacobson, 1938), self-instruction (Meichenbaum, 1974), breathing retraining (Clark, Salkovskis, & Chakley, 1985), biofeedback (e.g., Blanchard & Abel, 1976), social skills training (e.g., Becker, Heimberg, & Bellack, 1987), and distraction techniques (Wolpe, 1973). AMT does not target specific mechanisms presumed to underlie anxiety, as exposure and cognitive therapy do, but instead aims to provide the patient with ways to manage anxiety when it occurs (Suinn, 1974). Because of the multiple elements contained in most AMT programs, it is not possible to determine the contribution of any specific procedure to overall outcome.

SIT (Kilpatrick, Veronen, & Resick, 1982) was developed as a treatment for rape victims with chronic disturbances. Modifying Meichenbaum's (1974) stress inoculation training, it included deep muscle relaxation training, breathing control, role playing, covert modeling, thought-stopping, and guided self-dialogue. In a sample of rape victims, Resick, Jordan, Girelli, Hutter, & Marhoefer-Dvorak (1988) compared the efficacy of a modified SIT, assertion training, and supportive psychotherapy with a naturally occurring wait-list control group. All three treatments were equally but only moderately effective in reducing trauma-related symptoms, depression, self-esteem, and social fears; gains were maintained at six month follow-up on rape-related fear measures but not on other measures. No improvements were found in the wait-list control group.

Two studies, described in Section 6.22.5.4, examined the efficacy of SIT with female rape and nonsexual assault victims who met criteria for PTSD (Foa et al., 1997, 1991). The SIT program in both studies was adapted from the Kilpatrick et al.'s (1982) program described above. In the first study, both SIT and PE produced significantly more improvement on PTSD symptoms than the wait-list group and supportive counseling immediately following treatment. However, patients who received PE continued to improve after treatment termination, whereas patients in the SIT and supportive counseling conditions evidenced no change between post-treatment and follow-up (Foa et al., 1991). In the second study, SIT, PE, and the combination both showed immediate and long-term (up to one year) treatment effects as compared with the wait-list. However, when patients who received SIT were compared with patients who did not, no significant differences emerged on any of the measures. Also the effect size for PE relative to the control condition was larger than that for SIT on all measures (Foa et al., 1997).

6.22.5.4.3 Studies of combination therapies

The success of both PE and SIT for rape-related symptoms spurred the development of combination programs that include confrontations with feared stimuli in combination with cognitive therapy or anxiety management skills. Two such programs have been studied with female assault victims who manifested chronic PTSD and a third one is currently being evaluated. The first consisted of the combination of SIT and PE (SIT/PE), already described above (Foa et al., 1997). Contrary to the expectations that SIT/PE would enhance treatment benefit, all three active treatments (SIT, PE, and SIT/PE) were effective in reducing PTSD severity. In fact, it appeared that PE was superior to SIT and their combination (Foa et al., 1997). Foa et al. suggested that since the combined treatment was delivered in the same number of sessions, it is possible that patients did not receive as much imaginal exposure as the PE group. Also, the SIT program contained seven different techniques, some of which may not be effective. As noted above, Foa and her colleagues are currently examining the relative efficacy of PE combined with cognitive restructuring to PE alone and to a wait-list control group (see Jaycox, Zoellner, & Foa, 1997, for a case example and discussion).

The second combined program for rape victims with PTSD is CPT and which has been tested in a quasi-experimental design (Resick & Schnicke, 1992). CPT can be conducted in a group format. It includes education, an exposure component consisting of writing the trauma narratives and sharing it with the therapist and group members, and cognitive restructuring. The efficacy of CPT was compared with that of a naturally occurring wait-list, but no random assignment was instituted. Patients who received CPT showed significant and lasting improvement whereas those on the wait-list did not. As noted earlier in the section on exposure therapy, Resick (Resick et al., 1996) is conducting a study comparing CPT delivered individually with PE alone. Preliminary results suggest that both programs are highly effective and do not differ from one another.

6.22.6 PREVENTION PROGRAMS

Many trauma researchers and mental health professionals emphasize the need for intervention immediately after a trauma, with the aim of preventing chronic disturbances (e.g., Bell, 1995). While there is scant evidence in support of the efficacy of immediate intervention, many programs have followed the "critical incident debriefing" model developed by Mitchell and Bray (1990). This program typically includes seven phases conducted in small groups and delivered within three days of the traumatic event: establishing the purpose of the debriefing and the ground rules, "recreating" the event by having all participants give their perspective on what occurred, discussion of individuals' thoughts during the event, description of the worst part of the event for each individual, discussion of reactions to the event, a presentation of common reactions to trauma by group leaders to normalize participants' responses, and a wrap-up phase.

These types of interventions have been applied to survivors of a variety of traumatic situations, including bank employees held at gunpoint (Manton & Talbot, 1990), emergency workers (e.g., Armstrong, O'Callahan, & Marmar 1991), and military personnel (e.g., Fitzgerald et al., 1993). However, the efficacy of these programs has not been ascertained in well-controlled studies and some uncontrolled trials suggest that the debriefing can have a deleterious effect. As noted by Raphael, Meldrum, and McFarlane (1995), randomized well-controlled studies of such programs have not yet been conducted.

The recognition that victims who exhibit severe reactions shortly after the trauma are more likely to develop chronic dysfunction (Riggs et al., 1995; Rothbaum et al., 1992) has prompted researchers to implement interventions that aim to prevent chronic PTSD. Foa and colleagues conducted a study to compare PTSD severity of 10 female assault victims who received a brief prevention (BP) program (four individual therapy sessions) with that of 10 matched victims who underwent an assessment procedure (Foa, Hearst-Ikeda, & Perry, 1995). Victims who received the brief prevention program showed symptom reductions of 74%, compared with a mean reduction of 33% in the assessment control group. Although the small sample size and lack of random assignment precludes drawing definitive conclusions as to the efficacy of the BP, the results are encouraging, and Foa and her colleagues are now studying this intervention in a larger sample of female trauma victims.

In summary, there is no evidence to date that the commonly used crisis interventions are effective, but they have not yet been rigorously tested. With female assault victims, short-term behavioral interventions may help in preventing chronic post-trauma problems. Foa and Meadows (1997) noted that one difference between the crisis intervention programs and the Foa, Hearst-Ikeda, and Perry (1995) study is the time frame in which the intervention was delivered. A typical crisis intervention is employed within three to four days after the traumatic event, whereas the BP program was instituted about two weeks post-trauma. Clinical observations suggest that during the first few days after the trauma victims are still disoriented and preoccupied with immediate logistical concerns, such as interaction with law enforcement or medical systems. Perhaps then a program that is implemented during this period cannot be processed adequately to provide benefit and may even add to the already existing state of confusion. Studies should investigate the optimal time for the delivery of crisis interventions.

6.22.7 PROLONGED EXPOSURE THERAPY FOR PTSD: THEORY AND PRACTICE

Despite a wide array of treatments that are being employed with trauma victims, most outcome studies focus on cognitive-behavioral programs which, as note above, were proven to be quite effective in ameliorating PTSD and related disturbances. Of these, exposure therapy seems to occupy a favored position. First, variants of exposure used with a variety of trauma populations, including veterans, assault victims, and sample victims of mixed traumas, have all indicated at least some degree of success. Second, the evidence for the efficacy of exposure comes from controlled studies conducted in different research centers, thus reducing the possibility of researchers' bias. In contrast, the two controlled studies on SIT were conducted in a single center, as were the studies of CPT. Third, PE has the advantage that compared with SIT and CPT, it is relatively straightforward and thus therapists can easily be trained to deliver it. Because of the apparent advantage of PE, we will focus on PE in the next section. We will first discuss the mechanisms thought to underlie exposure techniques as a background for understanding the subtleties of how to implement it successfully.

6.22.7.1 Mechanisms Underlying Prolonged Exposure

Emotional processing theory discussed earlier (Foa & Kozak, 1985, 1986) posits that, regardless of the type of therapeutic intervention, two conditions are required for fear reduction to occur. First, the fear structure must be activated via introduction of fear-relevant information, because if the fear structure is not activated (fear is not evoked), the structure would not be available for modification. In other words, successful treatment requires that the patient be engaged emotionally with the feared or traumatic material. Second, new information must be provided that includes aspects that are incompatible with the pathological elements, so that these can be corrected. Accordingly, successful therapy can be understood as correcting erroneous associations in the pathological fear structure targeted in treatment. In the case of PTSD, Foa and colleagues (Foa & Jaycox, in press; Foa & Riggs, 1993) suggested that treatment should correct the mistaken beliefs that "the world is extremely dangerous" and "I [the victim] am extremely incompetent."

As noted earlier, exposure procedures for PTSD consist of confronting the patient with trauma-related information, thus activating the trauma memory (i.e., eliciting the trauma-related fear). This activation constitutes an opportunity for corrective information to be integrated, and thus for the pathological elements of the trauma memory to be modified. Foa and Kozak (1986) suggested that fear reduction (habituation) within and across exposure sessions is an indicator that meaning changes in the fear structure have taken place. Several studies lend support to the proposition that emotional engagement or fear activation enhances treatment efficacy (c.f. Foa & Kozak; Kozak, Foa, & Steketee, 1988). Of particular relevance to PTSD are two studies demonstrating that emotional engagement with the trauma memory during treatment predicts successful outcome (Foa, Riggs, Massie, & Yarczower, 1995; Jaycox, Foa, & Morral, in press).

Several mechanisms have been proposed to underlie observed improvement of PTSD (Foa & Jaycox, in press; Foa & Riggs, 1993). First, repeated, prolonged, imaginal reliving of the trauma is thought to promote habituation of the anxiety previously associated with the trauma memory, thus correcting the erroneous belief common to patients with PTSD that anxiety stays forever unless avoidance or escape is realized. Second, deliberate confrontation with the feared memory blocks negative reinforcement, the process whereby relief of anxiety following cognitive avoidance of trauma-related thoughts and feelings reinforces avoidance behavior. Third, exposure to the trauma memory in a safe and supportive therapeutic setting incorporates safety information into the trauma memory, thereby refuting the belief that remembering the trauma is dangerous. Fourth, focusing on the trauma memory for a prolonged period helps the patient to discriminate between the trauma and other nontraumatic events, thereby rendering the trauma as a specific occurrence rather than a representation of a dangerous world and of an incompetent self. Fifth, successful reliving changes the belief that PTSD symptoms are a sign of personal incompetence to a belief that they are a sign of courage and competence. That is, the patient realizes that the distress experienced during reliving (PTSD symptoms) does not lead to anticipated loss of control or "going crazy."

Sixth, prolonged, repeated reliving of the traumatic event allows the patient to focus on specific elements of the trauma that drive negative evaluations about themselves, and to modify those evaluations. For example, a 25-year-old patient who suffered from PTSD following a rape by an acquaintance was convinced at the beginning of treatment that she was weak because she had not been able to prevent the rape. Through the process of imaginal reliving, she spontaneously realized "He was a very big man, and I am small," and this realization helped to modify her negative self-image.

In addition to the six mechanisms suggested above, Foa and Riggs (1993) suggested that the fear structures of traumatic memories are more disorganized than nontraumatic memories. A disorganized memory, they proposed, is particularly resistant to modification. Repeated, prolonged trauma reliving generates a more organized memory record that can be more readily integrated with existing schemas. Support for this contention comes from a study analyzing the victims' narratives of the trauma during exposure (Foa, Molnar, & Cashman 1995). Indices of disorganization, such as unfinished thoughts and repetitions, decreased from the first to the last narrative and this decrease was correlated with improvement.

Most of the mechanisms discussed above are also thought to operate in *in vivo* exposure. However, the most salient mechanism during *in vivo* exposure is the correction of overestimates of the probability of danger. As will be described in the next section, the trauma victim with PTSD often comes to fear a certain time of day, a physical characteristic of an assailant, or a part of the city because the fear has generalized to situations that bear similarity to the original trauma even when he or she is objectively safe. Some situations are avoided because they are reminders of the trauma and thus elicit distress and PTSD symptoms. Repeated confrontation with such objectively harmless situations is thought to prompt a more realistic appraisal of danger.

6.22.7.2 How To Implement Exposure-based Therapy

Often, cognitive-behavioral therapy (CBT) researchers neglect to discuss the therapeutic relationship when reporting outcome data. However, in CBT, just as in any other therapeutic environment, the therapist-patient alliance is indispensable. This is usually established in the first two sessions when the therapist elicits information about the trauma and its aftermath. This discussion provides an opportunity to express empathy and personal concern with the well-being of the patient. The therapy is presented as a collaborative effort, where the therapist and patient work together to achieve specific goals. With trauma victims, the explicit goal is the reduction of PTSD symptoms. Often patients are preoccupied with a number of other life issues, such as relationship problems, work stress, or parenting issues. These issues, while important, can divert attention away from the work on trauma-related disturbances and hinder treatment efficacy. Therefore, the therapist and the patient should come to an understanding from the outset that such issues will be set aside to be addressed after the exposure therapy has successfully been implemented. If daily crises around such issues are too pressing, the therapist and the patient may decide to delay exposure therapy.

6.22.7.2.1 Information gathering

Therapy begins with a comprehensive evaluation of PTSD and related symptoms, as well as information about the trauma itself. Objective characteristics of the trauma (e.g., extent of injury, time of day, characteristics of assailant) are important for treatment planning, so that the therapist can plan exposures that effectively address trauma reminders. Subjective factors of the assault (e.g., fear of death, dissociation during the trauma) are also important, since they will help guide the therapist to include these details in the imaginal reliving. A detailed trauma history will help the therapist identify the trauma or traumas that generate PTSD symptoms and make it the focus of treatment. In some cases, it is possible to work on more than one traumatic event sequentially during the course of treatment, especially if the traumas are similar.

6.22.7.2.2 Breathing retraining

Since exposure techniques involve direct confrontation with feared stimuli, PTSD symptoms often increase in the beginning of therapy and this should be discussed with the patient before exposure sessions begin. Training patients to use a simple relaxation technique helps ease this initial increase in distress. We have chosen breathing retraining in our program because of its versatility: patients can use this technique in almost any life situation without disruption of their activity or attracting the attention of others. Teaching this technique at the end of the first session seems to help the patient calm down after discussing in detail the traumatic experiences.

In breathing retraining the patient is taught to inhale a "normal" amount of air, and then to exhale very slowly (four or five seconds). During exhalation, the patient is asked to repeat the word "calm" or "relax" mentally and to hold his or her breath for about four seconds before inhaling again. Ideally, the therapist would practice this technique in the session, and to make an audiotape of the instructions and the demonstration for the patient's use at home. Patients are instructed to practice the breathing when they are relatively calm at first, and then begin implementing the technique when they are anxious.

6.22.7.2.3 Education about symptoms

Patients with PTSD interpret PTSD symptoms as a sign that they are "losing their minds," or "going crazy." It is extremely important to educate the patient about common reactions to trauma in order to provide them with a framework for understanding their own symptoms and to help them comprehend the treatment rationale. Education about reactions to trauma includes information about PTSD symptoms: the nature of fear and anxiety in response to triggers and cues, re-experiencing the trauma, concentration problems, feeling overly alert and on guard, avoidance of thoughts and feelings as well as reminders of the trauma, emotional numbing, and anger. In this context some related difficulties are discussed, including sadness about losses related to the trauma, feeling out of control or crazy, negative self-image and ideas of incompetence, lack of trust in others and thoughts about the whole world or all people being dangerous, relationship problems, sexual dysfunction or fears (usually following sexual assault), and a rekindling of emotion related to prior traumatic experiences. The education component is best delivered in an interactive, collaborative style, providing the patient with the opportunity to discuss his or her own symptoms and disturbing thoughts.

6.22.7.2.4 In vivo *exposure*

A crucial component of exposure therapy is its rationale. In exposure therapy the therapist is asking the patients to abandon their avoidance strategies used to manage their symptoms and instead to engage in imagery and in activities that are likely to cause some emotional distress in the beginning. Understanding the rationale for this therapy that "asks them to suffer" is essential for enlisting the patient's motivation. *In vivo* exposure aims at reducing the anxiety that patients with PTSD experience when they confront trauma reminders, and thereby return the patient to their pretrauma level of functioning. The rationale is explained as follows:

> In this program, we are going to focus on the fears that you are experiencing, and your difficulty coping, both of which are directly related to your assault. We've talked about the feelings, thoughts, and memories that are connected with the assault. We've also talked about some of the ways in which you are coping with that distress, such as avoiding situations and memories that remind you of the assault.
>
> Although most of the symptoms that you and I have talked about gradually decline with time after the trauma, for many victims like yourself, some of these symptoms endure and continue to cause marked distress. It is possible to spend up your recovery process by understanding what causes your reactions.
>
> A major factor is *avoidance* of situations, memories, thoughts, and feelings. It is quite normal for people to want to escape or avoid memories, situations, thoughts, and feelings that are painful and distressing. However, while the strategy of avoiding painful experiences works in the short-run, it actually prolongs the post-trauma reactions and prevents you from getting over your trauma-related difficulties. (The therapist may want to elicit examples of the client's avoidance based on previous discussion of common reactions.)
>
> When you *confront* the painful experiences, rather than avoid them, you will have the opportunity to process the traumatic experience. For example, if you avoid assault-related situations that are objectively safe, you do not give yourself the opportunity to get used to being in these situations. Unless you confront the situations, you may continue to believe that they are dangerous and that your anxiety in these situations will remain indefinitely. However, if you confront these situations you will find out that they are not actually dangerous and that your anxiety will diminish with repeated, prolonged confrontations. As a result of this process your symptoms will decline. The same is true for painful memories. For this reason we will ask you to relive repeatedly in your imagination the assault and to confront relatively safe situations that you are now avoiding.
>
> Today we are going to concentrate on your tendency to avoid situations and people that are related to the trauma. In order to help you stop avoiding situations and people that were once enjoyable or important to you, we are going to work together to make a list of situations that you have been avoiding since the assault. I call this list "the hierarchy." I also want to find out from you how much distress or discomfort these situations would cause you if you weren't avoiding them. Therefore I will teach you a method to indicate your level of distress. Of course we will not ask you to confront unsafe situations. The goal is not to help you view dangerous situations as safe, but rather to help you stop avoiding situations that are realistically quite safe. (Foa & Rothbaum, 1998)

After explaining the rationale, the therapist and patient construct a hierarchy of feared situations, listing all the situations or objects that the patient either avoids or endures with intense anxiety. The hierarchy may include situations that are directly related to the trauma (e.g., being in the place where the trauma occurred), or situations that have come to evoke anxiety through generalization (e.g., sleeping with lights off). A list of about 10 to 12 items is constructed and the therapist describes the "subjective units of distress" (SUDs) scale. The SUDs scale is anchored individually for each patient, so that a 0 indicates the most calm and relaxed that the patient has ever felt (e.g., lying on the beach); and a 100 indicates the most fearful or distressed the patient has ever been (e.g., during the trauma). The therapist then elicits SUDs ratings for each of the situations on the hierarchy, indicating the level of distress expected when confronting that situation. For example, Bonnie is a 37-year-old African-American woman who was raped by a male acquaintance in a friend's home. During the information gathering, she reported fear that had generalized to being alone in many public situations. Her *in vivo* exposure hierarchy follows:

50	sleeping with the door closed and curtains closed
50	taking the bus alone in the early evening
60	going to sister's house (near assault)
60	wearing a skirt on the bus in the early evening
75	sitting on a bench in a mall or shopping center
75	taking a walk alone (daytime)
75	walking by house where assault occurred with friend
100	sitting in park alone (daytime)
100	looking at pictures of assailant
100	sitting on steps of building where assault occurred with friend

After the hierarchy is constructed, specific instructions about how to implement exposure exercises between sessions are given:

> When you are practicing in the mall, for example, you may initially experience anxiety symptoms, such as your heart beating rapidly, your palms sweating, feeling faint; you may want to leave the situation immediately. But in order to get over the fear it is important that you remain in the situation until your anxiety decreases. Once your anxiety has decreased at least 50%, then you can stop the exposure and resume other activities. (Foa & Rothbaum, 1998)

In each session, *in vivo* homework assignments are reviewed to gauge progress and to make modifications if the patient has difficulty. Avoidance of homework is gently confronted and the rationale reviewed if necessary. New assignments are explicitly planned with the patient each session, and some discussion of logistical constraints, anticipated problems, and alternate plans are helpful. Of course, some *in vivo* exposure may also be conducted in session (e.g., looking at picture of the assailant) with the therapist's guidance.

6.22.7.2.5 *Imaginal exposure*

As described earlier, the first step in exposure treatment is explaining the rationale for the procedure. The following is an example of the rationale being explained to a patient:

> Today we are going to spend most of the session having you relive the memory of your assault in your imagination. It is not easy to understand and make sense of traumatic experiences. When you think about the rape, or you are reminded of it, you may experience extreme anxiety and other negative feelings such as shame or anger. The assault was a very frightening and distressing experience, so you tend to push away or avoid the painful memories. You may tell yourself, "Don't think about it; time heals all wounds" or, "I just have to forget about it." Other people often advise you to use these same tactics. Also, your friends, family, and partner may feel uncomfortable hearing about the assault, and this may influence you not to talk about it. But, as you have already discovered, no matter how hard you try to push away thoughts about the assault the experience comes back to haunt you through nightmares, flashbacks, phobias, and distressing thoughts and feelings. These symptoms signal us that your assault is still "unfinished business." In this treatment our goal is to help you process the memories connected with the assault by having you remember them for an extended period of time. Staying with these memories, rather than running away from them, will help decrease the anxiety and fear that are associated with them. It is quite natural to want to avoid painful experiences such as memories, feelings, and situations that remind you of the assault. However, as we already discussed, the more you avoid dealing with the memories, the more they disturb your life.
>
> What we are going to do today is to begin helping you process the memories associated with your assault. The goal of imaginal exposure as well as *in vivo* exposure is to enable you to have thoughts about the rape, talk about it, or see triggers associated with it without experiencing intense anxiety that disrupts your life. This part of the program includes having you confront situations and memories that generate both anxiety and an urge to avoid. Gradually, the memories will become less painful. You will get used to them.

> Remember, we call this "habituation." Before we begin, do you have any questions about anything that I have said?
>
> Repeated reliving of the trauma will also help you in other ways. It will teach you that remembering is not re-experiencing the trauma. In other words, it will help you discriminate between remembering the trauma and being traumatized again. There is no danger in remembering and therefore there is no good reason to become overwhelmed with anxiety every time that you think about the rape. Also repeated reliving will teach you that you are not going to lose control or go crazy if you engage with the traumatic memory. On the contrary, through repeated reliving of the rape you are going to gain control over your memories instead of them having control over you.
>
> And finally, engaging with the traumatic memory repeatedly will allow you to differentiate between the traumatic event and other events that are similar but not dangerous. For example, now you are afraid of all bald men because you were raped by a bald man. By repeated reliving of the rape you will realize that "baldness" itself is not dangerous. (Foa & Rothbaum, 1998)

After explaining the rationale, the therapist explains the method for prolonged imaginal exposure or reliving of the traumatic experience. The patient is asked to relive the trauma vividly by describing it in the present tense, with eyes closed. Patients are instructed to include in the trauma recounting details of the physical environment, their actions; others' actions; their thoughts and feelings; and bodily sensations that had occurred during the trauma. Because of the findings that emotional engagement enhances treatment efficacy, the therapist emphasizes the importance of reliving the emotions that were experienced at the time of the trauma. An example of the instructions is as follows:

> I'm going to ask you to recall the memories of the assault. It is best for you to close your eyes so you won't be distracted and so that you can envision these events in your mind's eye. I will ask you to recall these painful memories as vividly as possible. We call this "reliving." I don't want you to tell a story about the assault in the past tense. What I would like you to do is describe the assault in the present tense, as if it were happening now, right here. I'd like you to close your eyes and tell me what happened during the assault in as much detail as you remember.
>
> We will work on this together. If you start to feel too uncomfortable and want to run away or avoid it by leaving the image, I will help you to stay with it. We will audiotape the narrative so you can take the tape home and listen to it for homework. From time to time, while you are reliving the assault, I will ask you for your anxiety level on the 0 to 100 SUDS scale, in which 0 indicates no anxiety or discomfort and 100 indicates panic-level anxiety. Please answer quickly and don't leave the image. Do you have any questions before we start? The rating should indicate how anxious you are at the time I ask, sitting here in my office, not how you felt during the assault. (Foa & Rothbaum, 1998)

The therapist may probe for more information if the patient does not provide enough detail. The reliving exercise is maintained continuously for 45–60 minutes, without discussion or interruptions, and is tape-recorded. If recounting the entire trauma takes less than the allotted time, the therapist asks the patient to repeat it over again, until the full time is completed. Thus, several repetitions of the trauma may occur in each session. Ideally, reliving should continue until some reduction in anxiety takes place, but it is sometimes necessary to end the exercise without reduction in SUDs. In the first few sessions, reliving should focus on the entire trauma, but in successive sessions, the therapist and patient select the most distressing parts, as indicated by high SUDs levels, as the focus of the reliving. We call these parts "hot spots" and they usually represent the most distressing moments during the trauma. The entire session can be focused on one or two hot spots. Once these hot spots are being processed and habituation occurs, the distress during the remaining parts of the trauma narrative usually diminishes. We advise that towards the end of treatment the entire trauma be relived again so that any newly emerging hot spots can be addressed before therapy is completed.

The patient takes home the audiotape of the reliving exercise, and is instructed to listen to it daily while imagining the trauma as vividly as possible. At the end of each reliving session, the therapist and patient discuss changes in the reliving experience within and between sessions, and the therapist reinforces the patient's ability to confront such painful information without losing control or going crazy.

6.22.7.3 Indications and Contra-indications

As noted earlier, often in the beginning of exposure treatment, the patient experiences an increase in anxiety owing to the confrontation with trauma reminders. Because of this initial increase in anxiety, patients should be relatively stable prior to the beginning of therapy. Thus, active psychosis, alcohol or drug dependence, or self-injurious behavior should be addressed prior to implementing exposure therapy. In addition, there is some indication that exposure to an event about which the patient feels intense guilt or shame, rather than fear, is contraindicated, since imaginal exposure in such cases can accentuate guilt and shame (Pitman et al., 1991).

Other factors that require attention may not become apparent until exposure is initiated. These include severe anger, overwhelming anxiety, and emotional numbing (see Jaycox & Foa, 1996; Jaycox, Zoellner, & Foa, 1997).

6.22.7.3.1 Anger

Anger is a natural response to a traumatic event. Often there is an identifiable party that can be blamed for the trauma (e.g., the manufacturer of faulty equipment, the assailant), and anger towards these parties appears to be justified. Even when the blame cannot be clearly ascribed to another, as in some cases of natural disaster, trauma victims often experience a basic sense of injustice (e.g., "Why me?"). Not surprisingly, then, anger is one of the defining symptoms of PTSD.

While some anger is a frequent feature of the PTSD clinical picture, studies suggest that anger may have a role in the formation or maintenance of chronic post-trauma disturbances. For example, Riggs et al. (1992) found that female victims of a recent crime who reported higher anger within two weeks of the assault, also reported higher PTSD symptom scores one month later. The authors hypothesized that intense anger impedes the natural recovery process by interfering with activation of the fear associated with the trauma, thereby obstructing emotional processing.

During treatment, intense anger appears to impede improvement in therapy by hindering emotional engagement. Facial fear expressions during the first reliving of rape memory during therapy predicted improvement following exposure therapy and were inversely related to self-reported anger prior to treatment (Foa, Riggs, Massie, & Yarczower, 1995). Moreover, anger at pretreatment was negatively correlated with improvement. The latter result was replicated in a larger sample (Jaycox, Perry, Freshman, Stafford, & Foa, 1995).

In contrast to severe and constant anger, clinical observations suggest that moderate episodic anger and irritability do not seem to impede improvement in therapy. That is, a certain degree of increase in anger is expected as anxiety increases during the initial stage of therapy, and this anger usually dissipates as arousal decreases. Thus, the therapist is advised to contract with the patients to delay addressing anger-related issues until other emotions such as fear have been processed.

6.22.7.3.2 Emotional Numbing

Numbing of emotions, such as anger, is part of the definition of PTSD. Numbing is usually defined as including several symptoms that constitute part of the avoidance cluster in PTSD: disconnection or detachment from others, blunted emotional responsiveness or expression, and lack of enjoyment in activities (e.g. Foa & Riggs, 1993; Foa, Riggs, & Gershuny, 1995; Litz, 1993). Foa and Riggs noted that the inclusion of numbing with effortful avoidance symptoms reflects the view that these symptoms are motivated and maintained by their capacity to reduce distress. It follows that direct confrontation with feared stimuli will exacerbate numbing in patients who use it as a coping strategy. However, increased numbing would be expected to obstruct emotional processing since it would hinder fear activation, thereby decreasing the effect of treatment. One way to overcome the obstacle presented by numbing during the exposure exercise is to remind the patient of the rationale for reliving. The therapist is also advised to probe for details, thoughts, and feelings during the imaginal exposure.

6.22.7.3.3 Overwhelming anxiety

Although a rare occurence, patients may become so engaged in the reliving exercise that they experience the traumatic memory as if the trauma is actually reoccuring. Consequently, the patient experiences the same overwhelming anxiety and lack of control as he or she did during the trauma. This flashback usually happens when the patient fails to incorporate into the traumatic memory the information that he or she is in a safe place, that is, in the therapist's office. The therapist can take measures to remind the patient that he or she is remembering the trauma, not re-experiencing it. For example, the therapist can instruct the patient to relive the trauma with eyes open rather than closed, or can insert additional encouraging and reassuring comments during the procedure.

In summary, the therapist should be attentive to these obstacles to exposure therapy, and should modify procedures, when needed, to overcome them (Jaycox & Foa, 1996). In addition to modification of the exposure techniques, the therapist may wish to consider alternative techniques, such as relaxation training to address overwhelming anxiety, or cognitive restructuring to address anger and guilt. These techniques are clearly helpful for trauma victims, as reviewed earlier, and can therefore be used to augment or replace exposure techniques in difficult cases.

6.22.8 REFERENCES

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. Rev.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Amir, N., Foa, E., & Cashman, L. (1997). Predictors of post-traumatic stress disorder: A prospective study of rape and non-sexual assault victims. Unpublished manuscript.

Armstrong, K., O'Callahan, W., & Marmar, C. R. (1991). Debriefing Red Cross disaster personnel: The multiple stressor debriefing model. *Journal of Traumatic Stress, 4,* 581–593.

Beck, A. T., Emery, G., & Greenberg, R. L. (1985). *Anxiety disorders and phobias*. New York: Basic Books.

Becker, R. E., Heimberg, R. G, & Bellack, A. S. (1987). *Social skills training treatment for depression*. New York: Pergamon.

Bell, J. L. (1995). Traumatic event debriefing: Service delivery designs and the role of social work. *Social Work, 40,* 36–43.

Bernstein, D. A., & Borkovec, T. D. (1973). *Progressive relaxation training*. Champaign, IL: Research Press.

Blanchard, E. B., & Abel, G. G. (1976). An experimental case study of the biofeedback treatment of a rape induced psychophysiological cardiovascular disorder. *Behavior Therapy, 7,* 113–119.

Boudewyns, P. A., & Hyer, L. (1990). Physiological response to combat memories and preliminary treatment outcome in Vietnam veteran PTSD patients treated with direct therapeutic exposure. *Behavior Therapy, 21,* 63–87.

Boudewyns, P. A., Hyer, L., Woods, M. G., Harrison, W. R., & McCranie, E. (1990). PTSD among Vietnam veterans: An early look at treatment outcome using direct therapeutic exposure. *Journal of Traumatic Stress, 3,* 359–368.

Boudewyns, P. A., Stwertka, S. A., Hyer, L. A., Albrecht, J. W., & Sperr, E. V. (1993). Eye movement desensitization for PTSD of combat: A treatment outcome pilot study. *Behavior Therapist, 16,* 29–33.

Bremner, J. D., Southwick, S., Brett, E., Fontana, A., Rosenheck, R. & Charney, D. S. (1992). Dissociation and posttraumatic stress disorder in Vietnam combat veterans. *American Journal of Psychiatry, 149*(3), 328–332.

Breslau, N., Davis, G. C., Andreski, P., Peterson, E. (1991). Traumatic events and posttraumatic stress disorder in an urban population of young adults. *Archives of General Psychiatry, 48,* 218–228.

Breuer, J., & Freud, S. (1955). Studies on hysteria. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 2, pp. 1–305). London: Hogarth Press. (Original work published 1893–1895).

Brom, D., Kleber, R. J., & Defres, P. B. (1989). Brief psychotherapy for posttraumatic stress disorders. *Journal of Consulting and Clinical Psychology, 57,* 607–612.

Clark, D. M., Salkovskis, P. M., & Chalkley, A. J. (1985). Respiratory control as a treatment for panic attacks. *Journal of Behavior Therapy and Experimental Psychiatry, 16,* 23–30.

Cooper, N. A., & Clum, G. A. (1989). Imaginal flooding as a supplementary treatment for PTSD in combat veterans: A controlled study. *Behavior Therapy, 20,* 381–391.

Davidson, J. R. T., Hughes, D., Blazer, D. G., & George, L. K. (1991). Post traumatic stress disorder in the community: An epidemiological study. *Psychological Medicine, 21,* 713–721.

Davis, R. C., Brickman, E. R., & Baker, T. (1991). Effects of supportive and unsupportive responses of others to rape victims: Effects on concurrent victim adjustment. *American Journal of Community Psychology, 12,* 443–451.

Dixon, P., Rehling, G., & Shiwach, R. (1993). Peripheral victims of the *Herald of Free Enterprise* disaster. *British Journal of Medical Psychology, 66,* 193–202.

Ellis, E. M., Atkeson, B. M., & Calhoun, K. S. (1981). An assessment of long-term reaction to rape. *Journal of Abnormal Psychology, 90*(3), 263–266.

Epstein, S. (1991). The self-concept, traumatic neurosis, and the structure of personality. In D. Ozer, J. M. Healy, Jr., & A. J. Stewart (Eds.), *Perspectives on personality,* (Vol. 3, pp. 63–98) London: Jessica Kingsley.

Fitzgerald, M. L., Braudaway, C. A., Leeks, D., Padgett, M. B., Swartz, A. L., Samter, J., Gary-Stephens, M., & Dellinger, N. F. (1993). Debriefing: A therapeutic intervention. *Military Medicine, 158,* 542–545.

Foa, E. B., Dancu, C. V., Hembree, E., Jaycox, L. H., & Meadows, E. A. (1997). The efficacy of prolonged exposure and stress inoculation training in the treatment of chronic PTSD. Submitted for publication.

Foa, E. B., Hearst-Ikeda, D. E., Perry, K. J. (1995). Evaluation of a brief cognitive-behavioral program for the prevention of chronic PTSD in recent assault victims. *Journal of Consulting and Clinical Psychology, 63,* 948–955.

Foa, E. B., & Jaycox, L. H. (in press). Cognitive-behavioral treatment of post-traumatic stress disorder. In D. Spiegel (Ed.), *Psychotherapeutic frontiers: New principles and practices*. Washington, DC: American Psychiatric Press.

Foa, E. B., Jaycox, L. H., Meadows, E. A., Hembree, E., & Dancu, C. V. (1996, November). Preliminary efficacy of prolonged exposure (PE) vs. PE and cognitive restructuring for PTSD in female assault victims. In P. Resick (Chair), *Treating Sexual Assault/Sexual Abuse Pathology: Recent Findings*. Symposium presented at the Association for Advancement of Behavior Therapy 30th Annual Convention, New York.

Foa, E. B., & Kozak, M. J. (1985). Treatment of anxiety disorders: Implications for psychopathology. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 421–452). Hillsdale, NY: Erlbaum Associates.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99,* 20–35.

Foa, E. B., & Kozak, M. J. (1991). Emotional processing: Theory, research, and clinical implications for anxiety disorders. In J. D. Safran & L. S. Greenberg (Eds.), *Emotion, psychotherapy, and change* (pp. 449–480). New York: Guilford Press.

Foa, E. B., & Meadows, E. A. (1997). Psychosocial treatments for post-traumatic stress disorder: A critical review. In J. Spence (Ed.), *Annual review of psychology*. Palo Alto, CA: Annual Reviews.

Foa, E. B., Molnar, C., & Cashman, L. (1995). Change in rape narratives during exposure therapy for PTSD. *Journal of Traumatic Stress, 8*(4), 675–690.

Foa, E. B., & Riggs, D. S. (1993). Post-traumatic stress disorder in rape victims. In J. Oldham, M. B. Riba, & A. Tasman (Eds.), *American psychiatric press review of psychiatry* (Vol. 12, pp. 273–303). Washington, DC: American Psychiatric Press.

Foa, E. B., Riggs, D. S., & Gershuny, B. S. (1995). Arousal, numbing and intrusion: Symptom structure of post traumatic stress disorder following assault. *American Journal of Psychiatry, 152,* 116–120.

Foa, E. B., Riggs, D. S., Massie, E. D., & Yarczower, M. (1995). The impact of fear activation and anger on the efficacy of exposure treatment for PTSD. *Behavior Therapy, 26,* 487–499.

Foa, E. B., & Rothbaum, B. O. (1998). *Treating the trauma* *of rape: Cognitive-behavioral therapy for PTSD*. New York: Guilford Press.

Foa, E. B., Rothbaum, B. O., Riggs, D. S., & Murdock, T. (1991). Treatment of post-traumatic stress disorder in rape victims: A comparison between cognitive-behavioral procedures and counseling. *Journal of Consulting and Clinical Psychology, 59,* 715–723.

Foa, E. B., Steketee, G., & Rothbaum, B. (1989). Behavioral/cognitive conceptualizations of post-traumatic stress disorder. *Behavior Therapy, 20,* 155–176.

Foa, E. B., Zinbarg, R., & Rothbaum, B. O. (1992). Uncontrollability and unpredictability in post-traumatic stress disorder: An animal model. *Psychological Bulletin, 112*(2), 218–238.

Foy, D., Sipprelle, R., Rueger, D., & Carroll, E. (1984). Etiology of posttraumatic stress disorder in Vietnam veterans: Analysis of premilitary, military and combat exposure influences. *Journal of Consulting and Clinical Psychology, 52,* 667–669.

Frank, E., Anderson, B., Stewart, B. D., Dancu, C., Hugues, C., & West, D. (1988). Efficacy of cognitive behavior therapy and systematic desensitization in the treatment of rape trauma. *Behavior Therapy, 19,* 403–420.

Frank, E., & Stewart, B. D. (1984). Depressive symptoms in rape victims. *Journal of Affective Disorders, 1,* 269–277.

Freud, S. (1953). The interpretation of dreams. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 4, pp. 1–388; Vol. 5, pp. 339–627). London: Hogarth Press. (Original work published 1905.)

Freud, S. (1955). Beyond the pleasure principle. In J. Strachey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 18, pp. 3–64). London: Hogarth Press. (Original work published 1920).

Gilboa, E. & Foa, E. B. (1997). Patterns of recovery after trauma: Individual differences and trauma characteristics. Unpublished manuscript.

Girelli, S. A., Resick, P. A., Marhoefer-Dvorak, S., & Hutter, C. K. (1986). Subjective distress and violence during rape: Their effects on long-term fear. *Victims and Violence, 1*(1), 3–46.

Green, B. L. (1994). Psychosocial research in traumatic stress: An update. *Journal of Traumatic Stress, 7*(3), 341–362.

Horowitz, M. J. (1976). *Stress-response syndromes*. Northvale, NJ: Jason Aronson.

Horowitz, M. J. (1986). *Stress-response syndromes* (2nd ed.). Northvale, NJ: Jason Aronson.

Jacobson, E. (1938). *Progressive relaxation*. Chicago: University of Chicago Press.

Janet, P. (1889). *L'Automisme psychologique*. Paris: Felix Alcan.

Janoff-Bulman, R. (1992). *Shattered assumptions: Towards a new psychology of trauma*. New York: Free Press.

Jaycox, L. H., Amir, N., Williams, M., Meadows, E., Cashman, L., Dancu, C., & Foa, E. B. (1996, November). A prospective study of post-traumatic stress disorder in female victims of sexual and non-sexual assault. Poster presented at the Association for Advancement of Behavior Therapy 30th Annual Convention, New York.

Jaycox, L. H., & Foa, E. B. (1996). Obstacles in implementing exposure therapy for PTSD: Case discussions and practical solutions. *Clinical Psychology and Psychotherapy: An International Journal of Theory and Practice, 3*(3), 176–184.

Jaycox, L. H., Foa, E. B., & Morral, A. R. (in press). The influence of emotional engagement and habituation on outcome of exposure therapy for PTSD. *Journal of Consulting and Clinical Psychology.*

Jaycox, L. H., Perry, K., Freshman, M., Stafford, J., & Foa, E. B. (1995). *Factors related to improvement in assault victims treated for PTSD*. Paper presented at the annual meeting of the International Society for Traumatic Stress Studies, Boston, MA.

Jaycox, L. H., Zoellner, L., & Foa, E. B. (1997). Cognitive-behavior therapy for PTSD in rape survivors. *In Session: Psychology in Practice, 3*(4), 43–58.

Jensen, J. A. (1994). An investigation of eye movement desensitization and reprocessing (EMD/R) as a treatment for posttraumatic stress disorder (PTSD) symptoms of Vietnam combat veterans. *Behavior Therapy, 25*(2), 311–325.

Jiranek D. (1993). Use of hypnosis in pain management and post-traumatic stress disorder. *Australian Journal of Clinical and Experimental/Hypnosis, 21,* 75–84.

Keane, T. M., Fairbank, J. A., Caddell, J. M., & Zimering, R. T. (1989). Implosive (flooding) therapy reduces symptoms of PTSD in Vietnam combat veterans. *Behavior Therapy, 20,* 245–260.

Keane, T. M., Fairbank, J. A., Caddell, J. M., Zimering, R. T., & Bender, M. E. (1985). A behavioral approach to assessing and treating post-traumatic stress disorder in Vietnam veterans. In C. R. Figley (Ed.), *Trauma and its wake* (pp. 257–294). New York: Brunner/Mazel.

Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C. B. (1995). Posttraumatic stress disorder in the National Comorbidity Survey. *Archives of General Psychiatry, 52,* 1048–1060.

Kilpatrick, D. G., Best, C. L., Veronen, L. J., Amick, A. E., Villeponteaux, L. A., & Ruff, G. A. (1985). Mental health correlates of criminal victimization: A random community survey. *Journal of Consulting and Clinical Psychology, 53,* 866–873.

Kilpatrick, D. G., Saunders, B. E., Amick-McMullan, A., Best, C. L., Veronen, L. J., & Resnick, H. S. (1989). Victim and crime factors associated with the development of crime-related post-traumatic stress disorder. *Behavior Therapy, 20,* 199–214.

Kilpatrick, D. G., Veronen, L. J., & Resick, P. A. (1982). Psychological sequelae to rape: Assessment and treatment strategies. In D. M. Dolays & R. L. Meredith (Eds.), *Behavioral medicine: Assessment and treatment strategies* (pp. 473–497). New York: Plenum.

Kingsbury, S. J. (1988). Hypnosis in the treatment of posttraumatic stress disorder: An isomorphic intervention. *American Journal of Clinical Hypnosis, 31,* 81–90.

Koopman, C., Classen, C., & Spiegel, D. A. (1994). Predictors of posttraumatic stress symptoms among survivors of the Oakland/Berkeley, Calif., firestorm. *American Journal of Psychiatry, 151*(6), 888–894.

Kozak, M. J., Foa, E. B., & Steketee, G. (1988). Process and outcome of exposure treatment with obsessive-compulsives: Psychophysiological indicators of emotional processing. *Behavior Therapy, 19,* 157–169.

Lang, P. J. (1979). A bio-informational theory of emotional imagery. *Psychophysiology, 6,* 495–511.

Leung, J. (1994). Treatment of post-traumatic stress disorder with hypnosis. *Australian Journal of Clinical and Experimental Hypnosis, 22,* 87–96.

Lindy, J. D., Green, B. L., Grace, M., & Titchener, J. (1983). Psychotherapy with survivors of the Beverly Hills Supper Club fire. *American Journal of Psychotherapy, 4,* 593–610.

Litz, B. T. (1993). Emotional numbing in combat-related post-traumatic stress disorder: A critical review and reformulation. *Clinical Psychology Review, 12,* 417–432.

MacHovec, F. J. (1983). Treatment variables and the use of hypnosis in the brief therapy of post-traumatic stress disorders. *International, Journal of Clinical and Experimental Hypnosis, 1,* 6–14.

Manton, M., & Talbot, A. (1990). Crisis intervention after an armed hold-up: Guidelines for counsellors. *Journal of Traumatic Stress, 3,* 507–522.

March, J. S. (1992). What constitutes a stressor: The "Criterion A" issue. In J. R. T. Davidson & E. B. Foa (Eds.), *Posttraumatic stress disorder: DSM-IV and beyond* (pp. 27–54). Washington, DC: American Psychiatric Press.

Marmar, C. R., Horowitz, M. J., Weiss, D. S., Wilner, N. R., & Kaltreider, N. B. (1988). A controlled trial of brief psychotherapy and mutual-help group treatment of conjugal bereavement. *American Journal of Psychiatry, 145,* 203–209.

Marmar, C. R., Weiss, D. S., Schlenger, W. E., Fairbank, J. A., Jordan, B. K., Kulka, R. A., & Hough, R. L. (1994). Peritraumatic dissociation and posttraumatic stress in male Vietnam theater veterans. *American Journal of Psychiatry, 151*(6), 902–907.

Martin, C. A., McKean, H. E., & Veltkamp, L. J. (1986). Post-traumatic stress disorder in police work and working with victims: A pilot study. *Journal of Police Science and Administration, 14*(2), 98–101.

McCann, I. L., & Pearlman, L. A. (1990). *Psychological trauma and the adult survivor: Theory, therapy and transformation*. New York: Brunner/Mazel.

McFarlane, A. C. (1989). The aetiology of post-traumatic morbidity: Predisposing, precipitating and perpetuating factors. *British Journal of Psychiatry, 154,* 221–22.

Meichenbaum, D. (1974). *Cognitive behavior modification*. Morristown, NJ: General Learning Press.

Mitchell, J. T., & Bray, G. (1990). *Emergency services stress: Guidelines for preserving the health and careers of emergency services personnel*. Englewood Cliffs, NJ: Prentice-Hall.

Mowrer, O. A. (1960). *Learning theory and behavior*. New York: Wiley.

Nishith, P., Mechanic, M. B., Griffin, M. G., & Resick, P. A. (1995, November). *Peritraumatic dissociation and chronic PTSD in rape victims*. Poster presented at the 11th Annual Meeting of the International Society for Traumatic Stress Studies, Boston, MA.

Peebles, M. J. (1989). Through a glass darkly: The psychoanalytic use of hypnosis with post-traumatic stress disorder. *International Journal of Clinical and Experimental Hypnosis, 37,* 192–206.

Perry, S., DiFede, J., Musgni, G., Frances, A. J., & Jacobsberg, L. (1992). Predictors of posttraumatic stress disorder after burn injury. *American Journal of Psychiatry, 149*(7), 931–935.

Piaget, J. (1971) *Psychology and epistemology: Towards a theory of knowledge*. New York: Viking Press.

Pitman, R. K., Altman, B., Greenwald, E., Longpre, R. E., Macklin, M. L., Poire, R. E. & Steketee, G. S. (1991). Psychiatric complications during flooding therapy for posttraumatic stress disorder. *Journal of Clinical Psychology, 52,* 17–20.

Pitman, R., Altman, B., & Macklin, M. (1989). Prevalence of posttraumatic stress disorder in wounded Vietnam veterans. *American Journal of Psychiatry, 146,* 667–669.

Pitman, R. K., Orr, S. P., Altman, B., Longpre, R. E., Poire, R. E., & Macklin, M. L. (1996). Emotional processing during eye-movement desensitization and reprocessing therapy of Vietnam veterans with chronic post-traumatic stress disorder. *Comprehensive Psychiatry, 37*(6), 409–418.

Raphael, B., Meldrum, L., & McFarlane, A. (1995). Does debriefing after psychological trauma work? *British Medical Journal, 310,* 1479–1480.

Resick, P. A. (1986). *Reactions of female and male victims of rape or robbery*. (Grant No. R01-MH37296, Final Report). Washington, DC: National Institute of Mental Health.

Resick, P. A., Jordan, C. G., Girelli, S. A., Hutter C. K., & Marhoefer-Dvorak, S. (1988). A comparative outcome study of behavioral group therapy for sexual assault victims. *Behavior Therapy, 19,* 385–401.

Resick, P. A., Nishith, P., & Astin, M. C. (1996, November). Preliminary results of an outcome study comparing cognitive processing therapy and prolonged exposure. In P. Resick (Chair), *Treating Sexual Assault/Sexual Abuse Pathology: Recent Findings*. Symposium presented at the Association for Advancement of Behavior Therapy 30th Annual Convention, New York.

Resick, P. A., & Schnicke, M. K. (1992). Cognitive processing therapy for sexual assault victims. *Journal of Consulting and Clinical Psychology, 60,* 748–756.

Resnick, H. S., Kilpatrick, D. G., Dansky, B. S., Saunders, B. E., & Best, C. L. (1993). Prevalence of civilian trauma and posttraumatic stress disorder in a representative national sample of women. *Journal of Consulting and Clinical Psychology, 61*(6), 984–991.

Richards, D. A., Lovell, K., & Marks, I. M. (1994). Post-traumatic stress disorder: Evaluation of a behavioral treatment program. *Journal of Traumatic Stress, 7*(4), 669–680.

Riggs, D. S., Dancu, C. V., Gershuny, B. S., Greenberg, D., & Foa, E. B. (1992). Anger and posttraumatic stress disorder in female crime victims. *Journal of Traumatic Stress, 5*(4), 613–625.

Riggs, D. S., Foa, E. B., Rothbaum, B. O., & Murdock, T. (1992). *Post-traumatic stress disorder following rape and non-sexual assault: A predictive model*. Unpublished manuscript.

Riggs, D. S., Rothbaum, B. O., & Foa, E. B. (1995). A prospective examination of symptoms of posttraumatic stress disorder in victims of nonsexual assault. *Journal of Interpersonal Violence, 10*(2), 201–214.

Roth, S., Dye, E., & Lebowitz, L. (1988). Group therapy for sexual-assault victims. *Psychotherapy, 25,* 82–93

Roth, S., Wayland, K., & Woolsey, M. (1990). Victimization history and victim–assailant relationships as factors in recovery from sexual assault. *Journal of Traumatic Stress, 3,* 169–180.

Rothbaum, B. O. (1995, November). *A controlled study of EMDR for PTSD*. Poster presented at the Association for the Advancement of Behavior Therapy, Washington, DC.

Rothbaum, B. O., Foa, E. B., Riggs, D. S., Murdock, T., & Walsh, W. (1992). A prospective examination of post-traumatic stress disorder in rape victims. *Journal of Traumatic Stress, 5,* 455–475.

Ruch, L. O., & Leon, J. J. (1983). Sexual assault trauma and trauma change. *Women and Health, 8,* 5–21.

Scarvalone, P., Cloitre, M., & Difede, J. (1995, June). *Interpersonal process therapy for incest survivors: Preliminary outcome data*. Presented at the Society for Psychotherapy Research, Vancouver, BC.

Shalev, A. Y., Peri, T., Canetti, L. & Schreiber, S. (1996). Predictors of PTSD in injured trauma survivors: A prospective study. *American Journal of Psychiatry, 153*(2), 219–225.

Shapiro, F. (1995). *Eye movement desensitization and reprocessing: Basic principles, protocols, and procedures*. New York: Guilford.

Spiegel, D. (1988). Dissociation and hypnosis in post-traumatic stress disorders. *Journal of Traumatic Stress, 1,* 17–33.

Spiegel, D. (1989). Hypnosis in the treatment of victims of sexual abuse. *Psychiatric Clinics of North America, 12,* 295–305.

Suinn, R. (1974). Anxiety management training for general anxiety. In R. Suinn & R. Weigel (Eds.), *The innovative therapy: Critical and creative contributions* (pp. 66–70). New York: Harper and Row.

Tichenor, V., Marmar, C. R., Weiss, D. S., Metzler, T. J., & Ronfeldt, H. M. (1996). The relationship of peritraumatic dissociation and posttraumatic stress: Findings in female Vietnam theater veterans. *Journal of Consulting and Clinical Psychology, 64*(5), 1054–1059.

Turner, S. M. & Frank, E. (1981). Behavior therapy in the treatment of rape victims. In L. Michelson, M. Hersen, & S. M. Turner (Eds.), *Future perspectives in behavior therapy* (pp. 269–291). New York: Plenum.

Ullman, S. E. (1995). *Impact of social relations on sexual assault victims*. Manuscript submitted for publication.

van der Kolk, B. A., Weisaeth, L., & van der Hart, O. (1996). History of trauma in psychiatry. In B. A. van der Kolk, A. C. McFarlane, & L. Weisaeth (Eds.), *Traumatic stress: The effect of overwhelming experience on mind, body, and society* (pp. 47–76). New York: Guilford Press.

Vaughan, K., Armstrong, M. S., Gold, R., O'Connor, N., Jenneke, W., & Tarrier, N. (1994). A trial of eye movement desensitization compared to image habituation training and applied muscle relaxation in post-traumatic stress disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(4), 283–291.

Wilson, S. A., Becker, L. A., & Tinker, R. H. (1995). Eye movement desensitization and reprocessing (EMDR) treament for psychologically traumatized individuals. *Journal of Consulting and Clinical Psychology, 63*(6), 928–937.

Wolpe, J. (1973). *The practice of behavior therapy*. New York: Pergamon.

Yalom, I. (1995). *The theory and practice of group psychotherapy* (4th ed.). New York: Basic Books.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.23

Psychoses: The Management of Severe and Enduring Mental Illness

GEOFF SHEPHERD
Health Advisory Service (HAS 2000), London, UK

6.23.1 INTRODUCTION 519
6.23.2 DEFINING SEVERE AND ENDURING MENTAL ILLNESS 520
6.23.3 A PSYCHOLOGICAL FORMULATION 520
6.23.4 CHANGES IN THE ORGANIZATION OF MENTAL HEALTH SERVICES 523
6.23.5 SOCIAL AND PSYCHOLOGICAL DEFICITS 526
6.23.6 MODELS OF SERVICE DELIVERY 527
6.23.7 EFFECTIVENESS AND EFFICIENCY 530
6.23.8 METHODOLOGICAL PROBLEMS 532
6.23.9 DEVELOPING A "USER-LED" SERVICE 533
6.23.10 STAFF TRAINING 536
6.23.11 CONCLUSIONS 538
6.23.12 REFERENCES 538

6.23.1 INTRODUCTION

The effective management of people with severe and enduring mental illness poses a number of complex psychological, social, and economic problems. It begins with the application of a range of therapeutic interventions—drugs, cognitive-behavioral therapy (CBT), "Expressed Emotion (EE) based" family therapy, etc.—all aimed at addressing treatable symptoms, but goes on to include "management" interventions aimed at helping the individual make a positive psychological adjustment to their difficulties and in facilitating their social reintegration. It is this dimension of managing chronic disability, with particular emphasis on the importance of social outcomes, which form the core of modern concepts of mental health rehabilitation and which mark out the care of severe and long-term mental illness from more straightforward, "treatment oriented" models. In this chapter we will focus specifically on the contribution of clinical psychology to this process.

Rehabilitation remains one of the most poorly understood areas of mental health. Traditionally, it has been synonymous with improving the functioning of long-stay patients in institutions so that they can move out of hospital and live independently in the community, but this is now no longer of much relevance. As the locus of care for most people with severe and enduring mental health problems has shifted from hospital to community, so our ideas about their long-term management has also changed to reflect a much more dynamic view of the interaction between the psychiatrically disabled person and his/her social environment. Rehabilitation has become much more than a set of techniques aimed at a single "end point": it is a "philosophy," a way of looking at the psychological and social problems of people with severe and long-term mental health difficulties in their social context. It also provides a framework in which the social/environmental dimension has as much place as the biological and in which the management of "chronic" disabilities is just as important as the treatment of symptoms. These ideas form the conceptual framework for this chapter.

But how can we provide operational definitions for "severe and long-term mental illness"? How can we produce a psychological formulation for the management of severe and long-term mental illness which links together biological, psychological, and social influences? What are the options in terms of delivering a service to people with these kinds of problems and can we be confident that they will meet their needs? What do we know about the effectiveness of different models of service delivery and what are the methodological limitations of this evidence? How can we go about effectively involving users in the planning and development of services and what implications does this have for staff and staff training? These are the specific questions we need to address. Let us begin with the question of definition.

6.23.2 DEFINING SEVERE AND ENDURING MENTAL ILLNESS

As community-based services for the care of people with severe and enduring mental illness have begun to develop, so it has become increasingly clear that we need to define the "severely mentally ill" in order to discriminate them from the majority of those using mainstream mental health services. These questions of definition have real practical significance as, without clear definitions, priorities cannot be set and scarce resources cannot be targeted to those in greatest need.

Attempts to define severe and persistent mental illness have generally relied on multidimensional criteria (Bachrach, 1988; Powell & Slade, 1996; Schinnar, Rothbard, Kanter, & Jung, 1990). These usually employ a combination of diagnosis (usually major psychosis); indices of service contact (frequency and duration of admissions); presence of disability (long-term difficulties with social functioning and adaptation); availability of formal and informal support; and degree of risk to self or others. The typical criteria used are set out in Table 1.

The criteria in Table 1 highlight the issue of "comorbidity" in defining "severe and enduring mental illness." It is not just that these individuals may be suffering from the active symptoms of major mental illness (schizophrenia, bipolar disorder), they also have a range of social and occupational problems (poor or unstable housing, poor work records, lack of social support) and may pose periodic risks to their own safety and/or those around them. It is this combination of problems and needs which makes this population particularly difficult to care for—and often particularly difficult to engage in conventional services. Thus, from a clinical psychology point of view, it may seem difficult to know where to "start." They have so many problems, and there are so many possibilities in terms of intervention, that the difficulty is in selecting the most potentially effective strategy (and, as we shall see later, this may not always be a strictly "psychological" intervention). Nevertheless, the skills of a clinical psychologist are highly relevant when it comes to assessing needs for care, formulating care packages, and delivering effective interventions.

Notwithstanding the difficulties, such a list of defining criteria also give those in a local service a useful basis for defining those patients in their locality who have the greatest needs for complex, long-term, management programs and who may therefore form the target for specialist community teams. These individuals are actually already likely to be known to a variety of local agencies (e.g., inpatient teams, primary health care, police, accident and emergency services, etc.) and it is often useful to begin by cross-checking the list of names with each of the agencies. A local "case register" can then be compiled which may be used for subsequent planning and monitoring of the service (for an example of this approach in action, see Audit Commission, 1994, p.55).

6.23.3 A PSYCHOLOGICAL FORMULATION

Having operationally defined severe and enduring mental illness we then need a psychological formulation which will help us organize our thinking and bring together information in a way that has both theoretical and practical implications for the management

Table 1 Criteria for defining people with severe and enduring mental illness.

A typical "severely mentally ill" person might be expected to have:
(i) diagnosis of schizophrenia or major affective psychosis
  and
(ii) at least two admissions in last year or six months continually in hospital in the last three years
  and
(iii) significant problems functioning in at least two of the following areas:
  basic literacy and numeracy
  self-care
  financial support (including money management).
  housing (poor quality or unstable).
  lack of social supports
  lack of occupation/employment
  difficulties with close personal relationships (family, etc.).

In addition, they are likely to show:
(i) poor compliance with prescribed medication
(ii) some degree of drug/alcohol abuse
(iii) difficulties in engaging in follow-up and aftercare
(iv) frequent crises and re-admissions
(v) significant history of self-harm, self-neglect, or harm to others

of care. The classic formulation, enshrined in the World Health Organization (WHO) classification of *impairment*, *disability,* and *handicap*, is derived from the pioneering work of the English social psychiatrist John Wing (Wing, 1978).

Impairment refers to disturbances in function at the level of organ or bodily system. For example, in schizophrenia brain functioning appears to be impaired in relation to neurotransmission in dopaminergic pathways and this gives rise to the primary symptoms such as hallucinations, thought disorder, etc. *Disability* refers to the effects of primary impairment on the person's ability to perform a range of necessary social tasks and roles, for example, to shop, to cook, to work, to keep oneself reasonably clean and tidy, to establish and maintain relationships, etc. A characteristic of disabilities is that they are long term and persist after symptoms have been treated and must therefore be regarded as semipermanent difficulties in functioning. Adverse personal reactions to illness (loss of confidence, denial, loss of motivation) are often also regarded as disabilities because their effects may be enduring. A person may thus be "disabled" by their illness without it necessarily having a direct or long-lasting effect on their cognitive or intellectual ability simply through its impact on their self-confidence and self-esteem. The importance of these adverse personal reactions and their implications for models of management will be addressed in more detail later. *Handicap* then refers to the *social* consequences of impairments and disabilities, for example, unstable housing, poor work record, loss of social supports, stigma, etc. Again, the degree of handicap may not be proportionate to either the degree of impairment or the level of disability. Thus, a person may be symptomatically quite well recovered, with little residual disability, but still very handicapped because of the reactions of those around him/her—friends, employers, neighbors, etc. Again, more will be said about the independence of symptoms, functioning, and social adjustment in the following paragraphs.

Although it is sometimes difficult to be sure in a given case how much a person's problems are related to the underlying impairment, how much is disability based on an adverse personal reaction, and how much is due to social handicaps, nevertheless this conceptualization provides a useful heuristic to help us think about the design and implementation of multifaceted interventions—biological, psychological, social—which will separately address these different areas of need.

The distinction between symptoms and functioning was first clearly demonstrated in the classic studies on the course and outcome of schizophrenia (Strauss & Carpenter, 1974, 1977; Strauss, Klorman, & Kokes, 1977). They showed that each outcome "domain"—symptoms, social functioning, work—was relatively independent of each other. This is analogous to the "desynchrony" observed between different outcomes in other kinds of mental health problems (e.g., fear and anxiety, see Rachman & Hodgson, 1974) and has important implications for both treatment and management. For example, it implies that changes in symptoms may not always produce generalized improvements in functioning. Thus, the person may show positive improvements in symptoms, while remaining socially quite disabled. Indeed, this is quite common (De Jong, Giel, Sloof, & Wiersma, 1986). Conversely, it may be possible for an individual to function quite well socially and vocationally while still experiencing severe symptoms. This is probably less common, but certainly not rare. Interventions therefore need to be targeted to producing specific changes in functioning and it cannot be assumed that if symptoms are treated (whether physically or psychologically) then this will automatically have a generalized effect on behavior.

By definition, "enduring mental illness" is long-term and a psychological model of rehabilitation therefore has one other cardinal feature: it is concerned with people who may not "get better." Rehabilitation asks, "How can we help people manage despite their disabilities?" It says, "By all means try and treat people, but if they don't respond then how can we help them make the best adaptation possible, given the fact that they have difficulties in functioning which are, in the current state of knowledge, essentially untreatable?"

This concept of psychological "disability" (i.e., untreatable symptomatology) still creates problems for many mental health professionals. They do not like to admit the existence of problems which they cannot do anything about and managing disabilities seems an altogether less glamorous—and lower status—activity than treatment and "cure." In dealing with people with severe and enduring problems, many professionals therefore find themselves in a state of psychological "dissonance," that is, there is a discrepancy between their expectations of themselves and what they find they can achieve with their patients. There is then a temptation to resolve this conflict either by devaluing the patient (i.e., "they are untreatable and therefore not worthy of further time and resources") or by feeling personally devalued ("I am ineffective and therefore don't have anything to offer this person"). Either way, staff may wish to avoid or withdraw from those with the most severe problems and seek out more therapeutically "promising" clients. This is compounded by a culture in which health services are very "outcome" and "throughput" oriented—often for simple, financial reasons. Long-term, intractable difficulties are therefore unwelcome news for both clinicians and managers alike and the challenge of rehabilitation services is how to think creatively about services which will promote social functioning and facilitate social reintegration while, at the same time, accepting the possibility of very small and slow changes over time. This can only be based on a sophisticated psychological understanding of the nature of the problems and the difficulties that they cause both for the patients and for the staff and other carers who are looking after them.

In stressing the importance of disability, we should also emphasize that psychiatric disabilities are not fixed and immutable in the same way as physical disabilities. In physical rehabilitation, the problems of adaptation may be solved by the provision of a range of physical supports (e.g., wheelchairs, ramps, large print, induction hearing loops, etc.). By contrast, the problems of the psychiatrically disabled person center around "social access" (e.g., to ordinary housing, ordinary jobs, nonsegregated leisure activities, etc.) and these depend on the provision of social supports—families, professionals, nonprofessional carers, etc.—and social attitudes (e.g., of friends, family, employers, etc.). Long-term social adaptation will be largely determined by the stability of these supports and, since they are very much dependent on individuals, they are inherently fragile. Social tolerances may also change: employment opportunities may diminish (or increase), people become tired, move on, get promoted, lose hope, die, etc. The pattern of psychiatric disabilities therefore can change—both for the better and the worse—and rehabilitation services have continually to monitor and adjust the level of their interventions, offering more, or less, support as the person's social circumstances change. Good rehabilitation is about the quality of this "continuing care" and how to achieve this through effective teamwork, training, and leadership are all topics to which we shall return to later.

This formulation of the dynamic nature of the relationship between the individual's functioning and the social environment lies at the heart of our current attempts to provide "community care." Psychiatrically disabled persons living in the community are struggling to maintain their social adaptation, while being constantly exposed to a changing (and therefore stressful) social environment. In these circumstances, it is not surprising that their adaptation breaks down from time to time and they need to be "rerehabilitated." This is not "failure": it is the inevitable outcome of change, life events, and limited coping capacities. But it does imply a very specific approach to the organization of services and to the training and support of staff and family carers.

How the organization of mental health services for people with severe and long-term mental health problems has changed over the last 40 years and what implications this has for the context in which professionals, patients, and their families now have to operate will now be discussed.

6.23.4 CHANGES IN THE ORGANIZATION OF MENTAL HEALTH SERVICES

Most countries have been struggling to move away from an organization of mental health services that depends heavily on residential provisions sited in large, old, psychiatric hospitals. This change is partly motivated by a belief—now supported by considerable evidence—that most people with serious and enduring mental health problems prefer to live and function better in various kinds of sheltered and supported settings in the community where they can have access to the same range of activities and social supports that are available to every other citizen.

The movement towards the community has proceeded at different rates in different countries. Some (e.g., The Netherlands) have managed to maintain a very good standard of care in their hospitals, while at the same time vigorously developing community alternatives. Others—and here the UK is a good example—have had more difficulty in ensuring adequate access to inpatient provision (both short- and long-term) as the overall number of long-stay beds has reduced. These changes have confirmed the warnings of scholars like Leona Bachrach who reminded us more than 20 years ago that the traditional mental hospital was a complex organization, fulfilling a number of medical, psychological, and social needs and that, if we had ambitions to manage without it, we would need to develop an equally complex range of services in the community (Bachrach, 1976).

All countries have experienced some difficulties in ensuring that there is an adequate range of housing, occupation, social, psychological, and medical support for all those who need it. Part of the reasons for this are clearly financial, but part are also related to the difficulties of interagency cooperation implied by such an agenda (Mechanic & Aiken, 1987). To recreate comprehensive community services for the most severely psychiatrically disabled demands a high level of cooperation between different providers—health, social services, housing agencies, work and occupational programs, primary care teams, etc. These agencies are all organized and managed differently, with different priorities, different planning cycles, and different funding streams. It is therefore not surprising that effective interagency cooperation is difficult to achieve and that there have been repeated calls for health and social care agencies to work more closely together. This has certainly been the thrust of successive planning guidance from the UK Department of Health (e.g., Department of Health, 1995). However, simply extolling the need for greater cooperation does not solve the bureaucratic and organizational problems of getting large agencies, with very different cultures, to understand one another's aims and priorities. These are essentially social and psychological problems and require psychological solutions.

The model of a single agency provider has actually been subject to experimental evaluation in the US through the Robert Wood Johnson Foundation program (Goldman, Morrisey, & Ridgely, 1994). This ambitious project aimed to create unified mental health authorities across eight US cities. In organizational terms it generally succeeded very well, however, it failed to deliver much in the way of improved health and social outcomes for patients, mainly because insufficient attention was paid to changing the practice of professionals through new training initiatives, staff development, and support (Lehman, Postrado, Roth, McNary, & Goldman, 1994; Morrisey et al., 1994). The lesson is clear: organizational change is not enough. If services are going to improve we therefore need to think about how to change the behavior of staff as well as how to change organizational structures.

However, in both the US and the UK there have been substantial reductions in long-stay beds in mental hospitals over the past 40–50 years (Bachrach, 1997). In the US there has been an 86% reduction in the number of beds provided in state hospitals over a period of 40 years (560 000 in 1955 to 77 000 in 1995). In the UK, there has been a similar dramatic reduction (from 155 000 in 1955 to less than 25 000 today). What effects have these changes had on the overall service "systems" and on the lives of individual patients and their families?

At the outset, it is important to distinguish between those "old" long-stay patients who have been most directly affected by the reductions in mental hospital beds and the next generation of potentially "new" long-term patients who would perhaps have gone into long-stay beds in the past, but for whom this is now no longer an option. Many of the "old" long-stay patients have simply swapped one form of institutional care for another (so-called "transinstitutionalization"). However, the quality of their lives may still have improved, providing the new accommodation is at least superior to a large, impersonal, and old-fashioned psychiatric hospital. But the "new" long-term patients: young, with major mental illness, severe symptoms, often complicated by drug and/or alcohol misuse, sometimes with a history of violence or self-harm, have fared less well. They may have ended up in acute inpatient units, often on general hospital sites, and become "new" long stay inpatients (Lelliot, Wing, & Clifford, 1994; Mann & Cree, 1976). Alternatively, they may be caught in the "revolving door" of acute admission, followed by premature discharge to inadequate support in the community, leading to rapid relapse, and re-admission. Most worrying of all, they may have dropped out of formal mental health services altogether and be existing on the margins of society, in poor or unstable housing, on the street, or in prison. They are the most obvious casualties of these changes in the patterns of service provision. Let us now examine the fate of these two groups in more detail.

As indicated, the quality of life for older patients who have been reprovided for in the community as part of a systematic reduction in beds in old, psychiatric institutions may or may not have been enhanced depending on the quality of their new accommodation. It is difficult to generalize because so much depends on local conditions, but the largest study of its kind in the UK (possibly in the world) of the outcomes for patients moved from two large psychiatric hospitals in North London—the so-called TAPS (Team for the Assessment of Psychiatric Services) study—suggested that those who move into reasonably good-quality residential care generally showed similar clinical and symptomatic outcomes compared with matched cohorts who remained in hospital, had better social outcomes, enhanced social networks, showed decreased behavioral problems and improved medication compliance, and were much happier with their living situation, compared with matched cohorts who remained in hospital (Leff, 1997). There was no evidence of increased death rates, suicide, crime, or vagrancy. The cost of care in the community was generally slightly less than of continuing care in hospital, but showed significant increases with increasing levels of disability (Beecham et al., 1997).

A more detailed study by Shepherd, Muijen, Dean, and Cooney (1996) attempted to examine some of the variability in the quality of care in residential provision for ex-long-stay patients in the community. They surveyed 20 residential homes in the London area, all of which had been established in connection with mental hospital closure programs. The homes were selected randomly to represent the range of current providers (housing associations, local Social Services departments, private-for-profit, etc.) and were compared with a small sample of "rehabilitation" wards from the local psychiatric hospitals. Residents and staff were assessed using a battery of standardized measures, the organization and quality of care was examined, and care was directly observed using a behavioral observation measure. Residents were also interviewed individually to determine their quality of life and satisfaction with their current living situation.

The results showed that the most disabled residents were concentrated in hospital, which generally offered the poorest quality of care and the most "institutionalised" staff. Not surprisingly, patients in hospital were much less satisfied and complained of lack of privacy, autonomy, and independence. In comparison with the large differences between hospital and community, the differences between community providers were small. Although, as in hospital, the personality and leadership style of the project manager appeared to be very important. Where project leaders encouraged a positive and respectful attitude towards residents, other aspects of the quality of care appeared to be good. Where they were more careless or neglectful, then this was also usually reflected in the overall running of the home. This study was consistent with the findings of the TAPS team in suggesting that for most of those directly affected by reductions in long-stay beds, the quality of their care has improved and, with it, their perceived quality of life. For them, community care has been a success and certainly has proved preferable to life in a traditional, psychiatric institution.

It is important to recognize that the reduction in hospital beds does not necessarily mean that there has been a reduction in the overall amount of residential provision. As the notion of transinstitutionalisation implies, the system of care may have changed in terms of who is providing residential accommodation, but the absolute number of places may not have altered greatly. This is illustrated by Martin Knapp and his colleagues, again in the TAPS study, who showed that there has been a significant growth in provision managed by "voluntary" and "independent sector" organizations. (These are a mixture of private "for profit" and "not-for-profit" providers). In the UK, this almost completely compensated for the loss of inpatient beds (Knapp, Beecham, & Hallam, 1997). However, the "mixed economy" of care does pose problems in terms of regulation and the monitoring of standards and it is worth remembering that the public mental hospital system was created precisely because of fears that a private system would be difficult to regulate and inspect (Jones, 1972). Such problems do not go away.

Probably the most important effect of the reduction in mental hospital beds in the UK has been the impact on the current acute services caused by the failure to develop adequate alternatives to long-stay inpatient provision.

This is illustrated in another recent study in the UK by Shepherd, Beardsmoore, Moore, Hardy, and Muijen (1997). They constructed a nationally representative sample of acute psychiatric units, stratified according to levels of local social deprivation. Information was then collected regarding the numbers of available beds, staffing levels, and "activity" (i.e., throughput). The range of local services, such as community mental health teams, crisis services, and supported housing was also described. A census was completed of all the patients resident on a given day and background data collected together with ratings of clinical and social functioning using the "Health of the Nation Outcome Scale" (Wing, Curtis, & Beevor, 1996). Staff were also asked to indicate whether or not they judged that the individual still required to remain on an acute inpatient ward and completed a standardized checklist covering the possible reasons preventing discharge or transfer to a more appropriate setting.

The results provided information on over 2000 patients admitted on nearly 40 sites covering approximately half the population of England and Wales. Fifty-two percent were female; the mean age was 41 (s.d. = 15.07), 88% were White; and 30% were currently detained under a Section of the Mental Health Act. Seventy-four percent had a diagnosis of schizophrenia or mood disorder. Diagnoses of substance abuse, neurosis, and personality disorder accounted for another 6% each. The median length of stay was 28.5 days, but 7% had current lengths of stay greater than six months (i.e., would be classified as "new" long stay). Their prevalence was much higher—almost double—in general hospital based acute units, compared with those sited in mental hospitals (9% vs. 5%). There was a negative relationship between the number of available adult acute beds and bed occupancy and there was also a relationship between levels of social deprivation, acute bed availability, and bed occupancy. Those services with low bed availability and high social deprivation therefore had the highest bed occupancy levels.

Despite these problems with overoccupancy, just over a one-quarter (27%) of current inpatients were judged not to require continuing admission and this proportion was significantly larger among those with admissions of more than six months. The most common reasons preventing discharge were the nonavailability of various community and residential options, in particular the lack of suitable supported housing. As might be expected, a lack of specialized rehabilitation places was seen as a more important reason preventing discharge for the new long-stay group. The higher prevalence of new long-stay patients in general hospital based acute units, compared with those based in mental hospitals, is consistent with the hypothesis that these patients are accumulating in acute units because of a lack of suitable alternatives. Those requiring higher levels of supervision in the form of secure accommodation (for reasons of violence) amounted to around 10%, although this proportion was higher (17%) among the new long-stay patients. The HoNOS severity scores were consistent with the staff judgments, thus validating the placement options selected by the staff.

These results therefore support the hypothesis that patients are now accumulating on acute admission wards who would previously have been accommodated on long-stay wards, in secure provisions, or in various community alternatives. The resultant shortage of acute beds seems to be related to both social deprivation and lack of bed availability. However, it is clear that a simple expansion of acute beds would not effectively address the problem. If all the patients who are judged to be inappropriately placed on acute admission wards could be relocated, then the problem of overoccupancy would effectively be solved. Acute units could then get on with the task of looking after the patients they are meant to be targeting (i.e., those whose needs are for immediate treatment and symptom stabilization). But this does mean the creation of truly comprehensive community services, with a spectrum of residential options geared to different levels of need and different kinds of problems. There are clearly considerable difficulties in achieving this, particularly in inner cities, with a background of poor housing, high unemployment, and other social problems.

Apart from new resources, the creation of comprehensive community services which can look after the most disturbed (and disturbing) individuals also implies a very considerable change in public attitudes towards the care of people with severe and enduring mental health problems. At the present time—at least in the UK—attitudes could hardly be less favorable. The inadequate development of community alternatives, combined with the reduction in long-stay beds, has led to some highly publicized failures of community care in which members of the public have been injured, or even killed, by people regarded as "mentally ill." The perception that community care has failed has been reinforced by sensationalist press and media coverage, although there is actually no evidence that the rate of homicides committed by people with a mental illness is actually increasing, nor that the care of those who committed these acts would have been improved if more psychiatric beds had been available. In reality, the problems seem to have been not so much due to a shortage of resources, but of a failure to coordinate resources effectively and to communicate between different agencies (Ritchie, Dick, & Lingham, 1994). It is thus the organization of services that has been deficient, not their basic structures (or even, in many cases, the overall level of resourcing).

As this new group of very disturbed and very difficult young people have begun to emerge, it has also become clear that traditional models of service delivery will simply not be effective with them. Many of these young people are very sensitive to stigma issues and are reluctant to engage with traditional services which they feel simply add to their sense of alienation and isolation. They want new forms of services—housing, work, etc.—which are not substantially different from those available to normal people. Furthermore, they want services delivered at the times and places most acceptable to them, rather than those which are most convenient for the professionals. Such services have been slow to develop and, as a result, some of these individuals have opted out of formal mental health care, sometimes with tragic consequences both for themselves and others. So, what can done to improve services for this group in the future?

6.23.5 SOCIAL AND PSYCHOLOGICAL DEFICITS

First, we must clearly understand the nature of their psychological and social deficits. In 1994, the UK Royal College of Psychiatrists' research unit reported on a national audit of new long-stay inpatients—those who had been in hospital continuously for more than six months, but less than three years (Lelliot & Wing, 1994; Lelliot et al., 1994). Almost two-thirds (62%) had a diagnosis of schizophrenia, half had more than five previous admissions to hospital, and just under one-third (29%) were detained under a Section of the Mental Health Act (mostly for compulsory treatment). They showed multiple disabilities, with a high prevalence of severe, treatment-resistant symptoms and a range of other behavioral problems (violence, self-harm, extreme antisocial behavior, etc.) which made them difficult to look after outside hospital. They tended to be socially unskilled, with poor work records, few family ties, and poor physical health. About one-third were judged to be capable of living in the community if highly supervised hostel placements were available, about one-third were deemed to require further treatment in hospital, and the remainder were so multiply handicapped that it was difficult to see where they might be placed.

A very similar pattern of results have been obtained from comparable populations in the US. For example, in a classic paper Sheets, Prevost, and Rehman (1982) hypothesized three subgroups among the new, "young chronics." The first group they described as characterized by "low energy/low demand" with long histories of contact with services, passivity, low motivation, poor personal hygiene, and treatment-resistant symptomatology. They appeared "burnt out," accepting of their status as "chronic" patients, despite their relative youth. The second group was termed "high energy/high demand" and these individuals were characterized by low frustration tolerance, high geographical mobility, rapidly fluctuating levels of functioning and symptoms, "acting out" behavior, aggression, contact with the law, etc. 5 > They were very demanding, uncooperative, and difficult to manage. Finally, there was a "high functioning/high aspiration" group, characterized by relatively good premorbid history, coming from the middle class, having high ambitions, but unable to realize them. They often also had significant problems with drug or alcohol abuse. This group were also rejecting of their status as chronic patients, but accepting of help from mental health services insofar as it would enable them to lead relatively normal, unstigmatized lives.

Gudeman and Shore (1984) also described the characteristics of patients in their service in Boston who had proved very difficult to look after in conventional community accommodation. They consisted of a mixture of elderly, demented people with behavioral disturbance (20%), mentally retarded persons with concomitant psychiatric illness and aggressive behavior (20%), people with acquired brain damage and loss of impulse control (10%), and two groups of predominantly schizophrenic patients, one with unremitting assaultive or suicidal behavior (17%) and the others who, while not being a frank danger to themselves or others, showed such unacceptable social behavior (e.g., disrobing in public, eating garbage, continually vomiting food at mealtimes) that they would not be tolerated in an ordinary community setting (33%). Similarly, Bigelow, Cutler, Moore, McComb, and Leung (1988) studied a group nominated as "hard to place" by staff in a state hospital. Almost three-quarters (70%) had a primary diagnosis of schizophrenia, often in association with drug and alcohol problems and almost two-thirds (62%) showed significant levels of violent and assaultive behavior. They also note the high prevalence of poor self-care, leisure, and social skills.

Thus, a consistent picture emerges, across very different settings and local conditions, of the psychological and social deficits found in people who are difficult to place in traditional services. First, there is a group of young men, almost all of whom have a primary diagnosis of schizophrenia, who show severe and intractable "positive" and "negative" symptoms and a variety of behavioral problems. They may also have concomitant drug or alcohol misuse. This is the largest group and, depending on local circumstances, is likely to account for between two-thirds and three-quarters of the total. Second, there is a group of older persons, primarily female, who are more likely to have treatment-resistant affective psychoses (depressed type or rapid cycling bipolar conditions) and who present a significant risk to themselves and may show extreme self-neglect when they are acutely depressed. They are likely to account for a further 20%. Third, there is an heterogenous group, usually having organic brain syndromes (acquired brain damage, presenile dementia, low IQ, alcoholism, etc.) again, often with other associated behavioural problems, sometimes also showing comorbidity with psychosis. They usually account for about 10%. What are the options for these groups in terms of models of service delivery?

6.23.6 MODELS OF SERVICE DELIVERY

The range of services necessary to meet this complex combination of needs, in the absence of large numbers of beds in mental hospitals, have been well described in a number of strategy documents published on both sides of the Atlantic (Department of Health, 1994, 1997; Stroul, 1986). At the center of these systems lie various forms of sheltered and supported housing.

It is something of a truism to state that without adequate accommodation effective management of those with the most severe and enduring mental health problems will not be achievable, but nevertheless it bears repeating. What is difficult is to specify the precise levels of the different kinds of accommodation that are required. Attempts have been made (e.g., Strathdee, Davies, Perry, & Thompson, 1996; Wing, 1992) but any general guidelines always require qualification depending on local circumstances. The greatest shortages are usually regarding high-support housing of various kinds, but the possibilities of substituting one kind of facility or service by another (e.g., intensive support teams instead of high-support housing) make "normative planning" difficult. Each location is different and each service therefore needs to evolve in order to fit the local conditions.

At the extreme end of the housing spectrum there are those individuals with severe and enduring mental health problems who are actually homeless, or in unstable, or very poor quality housing. The problem of severe mental illness among homeless populations has received considerable attention worldwide (Bachrach, 1992; Bhugra, 1996; Timms, 1993) yet there is still controversy concerning the link between reductions in hospital beds and apparently increasing numbers of mentally ill people among the homeless. Most of the evidence seems to point to an indirect link, with other factors—particularly general housing, benefit, and unemployment policies—also exerting a considerable influence. Thus, the typical mentally ill homeless person in the UK is actually young (i.e., under 30 years), with some degree of alcohol or other drug abuse, and a range of social and occupational disabilities in addition to their mental illness, but without a previous history of long admissions to psychiatric hospitals (Craig, Bayliss, Klein, Manning, & Reader, 1995). Often they may have had a series of short admissions (i.e., "revolving door") but generally do not have the characteristics of an old long-stay patient who has been turned out of a mental hospital. In terms of service provisions, apart from the obvious need for stable, good-quality housing, the most successful management approaches appear to be using specialist community teams, who practice an "assertive outreach" model (Dixon, Krauss, Kernan, Lehman, & DeForge, 1995). Breakey (1996) also notes four stages in providing psychiatric services for the homeless: (i) engagement; (ii) provision of basic services; (iii) transition towards mainstream services; and (iv) integration into ordinary housing and work programs. We will have more to say about the work of these kinds of specialist community teams later, suffice to say at this point that the contribution of clinical psychology to all these stages in terms of both assessment and management is very important.

As indicated, the goal of working with people with severe mental illness who are homeless is generally to try and get them into some kind of stable, sheltered, or supported housing. So, how can we judge the quality of care in such housing provisions? The principles underlying good-quality care in supported housing have recently been described by Shepherd (1998a). First, each person must be regarded as an individual with a unique set of needs and abilities. This is the antithesis of the "block treatment" which characterized traditional institutional settings (Goffman, 1961). Nobody should be given care that he or she does not need and, conversely, nobody should be denied help because of general rules or expectations regarding levels of functioning or performance. Staff should make use of their everyday observations of residents in a variety of community settings (shops, public transport, eating houses, etc.) to make direct assessments of their skills and deficits. Residents should be fully involved in the formulation of their own care plans and these should be reviewed—together with staff—at least weekly. The aim should be to create an active, information-rich, environment in which staff use their relationships with residents, in the context of collaboration over everyday activities, to shape behavior. In general terms, there should be an emphasis on practical tasks and daily living skills and a deemphasis of pathology. (This is not to say that symptoms or psychological distress should be ignored but, by definition, these residents will have exhausted most conventional therapeutic approaches.) The crucial question is, "To what extent do their symptoms interfere with their ability to function?" This emphasis on functioning rather than symptoms gives a clear "normalizing" message to residents and contributes to the maintenance of an ordinary household atmosphere.

Finally, staff should be encouraged to monitor the quality of their interactions with residents, as well as their quantity. Quality in this context means not only a respect for the individual, but also a style of interaction that conveys clear expectations about performance and tries to minimize criticism and censure. This may be characterized as a low EE style and follows from the recent evidence that high levels of criticism and hostility expressed by nonfamily carers may also be detrimental, just as within certain dysfunctional families (Ball, Moore, & Kuipers, 1992; Moore, Kuipers, & Ball, 1992). In terms of staff support, there needs to be an emphasis on consistency, continuity, teamwork, and communication. Maintaining a positive therapeutic environment is probably the most difficult task of all, staff need to meet regularly and continually to have the opportunity to review their work with each other and with senior colleagues. Teams are fragile and a lot of time and effort has to go into "feeding" them to keep them functioning at their highest level.

Next to accommodation, most service users place work and/or employment as their greatest priority (Rogers, Pilgrim, & Lacey, 1993; Shepherd, Murray, & Muijen, 1995). In terms of developing a comprehensive service, the aim must again be to provide a range of work opportunities—assessment and training, placement in open employment with support, sheltered social firms and cooperative businesses, etc.—to cover the range of users' needs (Grove, 1994; Pozner, Ng, Hammond, & Shepherd, 1996). Just as in housing, one should not expect everyone necessarily to "progress" through the system, acquiring more and more skills, and therefore requiring less and less support. Some will and some won't: and it is important to allow individuals to settle at their own "level." There must be opportunities to move on (e.g., toward open employment) but also opportunities to stay if this is more consistent with their abilities and aspirations.

Users particularly value services which they see as having a positive image and as being well regarded by others (Dick & Shepherd, 1995). It is thus insufficient that professionals regard a particular work program as representing good practice. The crucial question is, "Do users want to go?" If not, then it will fail to engage some of the most difficult, younger people with psychosis who actively reject conventional day and work programs. Of course, not everyone will want to engage in work programs, however attractive services try to make them—that's also fine. But we cannot afford to alienate users by offering services which essentially reflect the therapeutic interests and values of highly paid professionals, rather than what they say they want. In this connection it is worth remembering the classic study of Linn, Coffey, Klett, Hogarty, and Lamb (1979) which showed that the number of psychologists and other therapists was inversely correlated with a good outcome in day programs supporting people with schizophrenia in the community. Modern, relevant, high-status work activities—computing, silk screening, designing T-shirts, producing interesting looking booklets and newsletters—combined with loud music (and financial remuneration) may therefore be the necessary ingredients to attract many of the target group. If this is what is needed to give work programs a positive image, then this is what must be provided.

The Clubhouse Model (Beard, Propst, & Malamud, 1982) has also proved a very popular approach with some service users who are seeking a more positive image from mental health day care programs. The principles of this model are set out in Table 2. As can be seen, the clubhouse model emphasizes mutual cooperation and consensus between staff and members in the development of the program and the importance of meaningful occupation both within and outside the clubhouse. It has proved very popular, particularly in the US and, more recently, also in the UK and other parts of Europe (e.g., Sweden). However, it has not been without its detractors.

For example, Transitional Employment Placements (TEPs) have been criticized because of

Table 2 The principles of the Clubhouse Model.

It is a club and, as in all clubs, it belongs to those who participate in it
Members and staff work together to design and implement the clubhouse program
The clubhouse believes in the value and dignity of meaningful work activities
All members are made to feel on a daily basis that their presence is expected and that their attendance really makes a difference
There are never enough staff to run the program on their own: members' contribution is therefore necessary for the success of the program
Members have an opportunity to experience Transitional Employment Placements (TEPs). These offer fixed-term, part-time, entry-level work experience and may be used as stepping stone to paid employment
Membership is for life and the clubhouse is open seven days a week, 365 days per year

Source: Beard et al. (1982).

the restrictions placed on entry-level positions. This is necessary because the model demands that any member or staff person can, in principle, cover any position. As a consequence, it means that TEPs tend to be limited to very simple jobs and may therefore be unattractive to those seeking re-entry into highly skilled or professional careers. The style of shared decision-making between members and staff has also been criticized as not being genuinely participative and therefore not really "user led" (Chamberlain, 1988). Whether this is fair probably depends on the specific clubhouse concerned, but there is some evidence to suggest that users do feel involved and empowered in clubhouses and that this contributes significantly to their overall life satisfaction (Rosenfield & Neese-Todd, 1993).

Perhaps the most serious criticism of clubhouses is that the intentional community created by the clubhouse culture, while being very attractive to some individuals, may not be helpful to others in terms of their reintegration into mainstream social activities. There is clearly a dilemma here: on the one hand, the clubhouse provides a sheltered setting which protects vulnerable and disadvantaged people from the stresses of ordinary social contact—asylum in the best sense of the word. On the other, an asylum, however positive, may become difficult to leave and therefore runs the risk of creating an unhelpful dependency. There is no satisfactory answer to this dilemma (and, of course, it occurs in other models than the clubhouse). Services simply have to strive to strike a sensitive balance between shelter and independence and ensure that they have clearly permeable boundaries so that individuals do not become isolated and cut off from normal social life. This is a fundamental problem for all dedicated community services (housing, work programs, social programs, etc.) which provide specialized help and support in a segregated form.

Regarding social support, evidence has continued to accumulate that different individuals have very different needs in terms of social support and that these may vary significantly over time. For example, Cresswell, Kuipers, and Power (1992) demonstrated how people with severe negative symptoms may have less need for close emotional support, but greater need for more practical kinds of help. This is consistent with other evidence regarding the limited needs of some very disabled patients for close emotional support (e.g., Mitchell & Birley, 1983). In a classic paper, Brier and Strauss (1984) described systematic changes in social support needs as the people recovering from psychotic episodes progressed from convalescence in which reality testing, material help, social approval, etc. were the priority, through to rebuilding in which the attention shifted to motivation, reciprocal relations, and symptom monitoring. These considerations have led writers like Harris, Bergman, and Bachrach (1986) to stress the importance of "individualized network planning," that is, tailoring interventions much more closely to the individual's unique pattern of social needs.

In the social support literature, there has also been a change from a preoccupation with the size of social networks to an examination of the influence of other network variables, particularly their composition and density (i.e., interconnectedness). Thus, it has been known for some time that social isolation tends to be associated with a higher risk of admission, but it is now clear that repeated admission to hospital also affects the composition of the network. Over time, friends and relatives are gradually replaced by mental health professionals and other service users (Holmes-Eber & Riger, 1990). This is not only likely to lead to lowered levels of satisfaction with social supports (Goering et al., 1992) but may also create networks which are relatively poor at providing help in a crisis, since they are diffuse and lack the interconnectedness necessary to mobilize support quickly in an emergency (Dozier, Harris, & Bergman, 1987). Improving communication (interconnectedness) within support networks should therefore be one of the central goals for specialist community teams.

Specialist community teams are the glue that keeps services together around individuals and allows them to reach their potential while, at the same time, ensuring that they receive sufficient support. These teams were pioneered for the severely mentally ill by Leonard Stein and his colleagues in Madison, Wisconsin (Test & Stein, 1978) and have now been operating successfully for more than 25 years (Thompson, Griffiths, & Leaf, 1990). The model has been exported to Australia (Hambridge & Rosen, 1994) and several European countries (e.g., The Netherlands and Italy). Such teams are beginning to appear in the UK (Ford et al., 1995) but progress is slow. The model for these teams (including those that work with the homeless) is based on assertive outreach, with small caseloads (n = 10–15), extended hours of operation (i.e., outside Mon./Fri. 9–5), and the capacity to deliver intensive support (i.e., daily visits) where necessary.

It is important to distinguish between these "intensive support" teams and "crisis intervention" services. Although the two kinds of teams share some features in common (e.g., assertive outreach, extended hours, intensive visiting), the skills mix and methods of working are quite different. Thus, crisis teams tend to be focused on the short-term, with an emphasis on symptom resolution, whereas intensive support teams have a longer-term perspective, with more of an emphasis on support and the optimization of social functioning. Crisis teams are also much more reliant on the use of medication, and hospitalization, with the aim of rapid symptom stabilization, while intensive support teams give more emphasis to practical help, social support, and facilitating access to mainstream community activities.

The final element in the necessary range of services to support people with serious and enduring mental illness in the community are inpatient beds. No matter how good the specialist intensive support (or crisis intervention) teams are in the community, from time to time inpatient admission will continue to be required. Ideally, community and inpatient teams should work together as an integrated system, with the same medical input to both elements. However, there is often poor integration between hospital and community services and community teams may thus need to develop "assertive inreach" to the inpatient units, as well as "assertive outreach" to the community.

So, what is the evidence for the effectiveness of these different components of a comprehensive community service?

6.23.7 EFFECTIVENESS AND EFFICIENCY

There is a dearth of evidence regarding the effectiveness of different kinds of community accommodation. As indicated earlier, most long-stay patients show lower levels of symptoms (particularly negative symptoms), have better social integration, and report higher levels of satisfaction, compared with matched cohorts who remain in hospital (Leff, 1997). The higher levels of reported satisfaction among the community cohorts seem to be directly related to the less restrictive environments in community residential settings. This provides a link between quality of care measures and quality of life indices (Abrahamson, Swatton, & Wills, 1989; Shepherd et al., 1996). In order to improve the quality of life for people with severe and enduring mental health problems we therefore need to focus on changes which will increase their range of choices and opportunities in relation to simple, everyday life goals (e.g., meals and mealtimes, access to rooms and personal possessions, etc.). Even in some community settings this may be difficult to arrange: in hospital it is virtually impossible.

In terms of user preferences, it is quite clear that most patients prefer more independent living arrangements, combined with flexible staff support, rather than staff "on site" and having to live with a group of other people with severe mental illness (Tanzman, 1993). These aspirations are understandable, but may cause problems when it comes to the care of the most disturbed (and potentially "risky") individuals. It is clearly desirable to try to offer them placements in ordinary housing, but in practice this may heighten fears for their—and others'—safety. It may also provoke considerable resistance from the local community if they become aware that such developments are planned.

Because of these problems, the Department of Health in the UK has suggested the concept of "24 hour nursed beds" in an attempt to combine high levels of professional supervision with a relatively low degree of institutionalisation (NHS Executive, 1996). This idea is very similar to the concept of a "ward-in-a-house" or "hospital hostel" proposed some years ago as a solution for some of the most difficult new long-stay inpatients (Young, 1991). The evidence for the effectiveness of these models has been reviewed in Shepherd (1995) and more recently in Shepherd (1998a). It seems that they may be effective in improving the functioning of up to 40% of those referred sufficiently for them to be resettled into less highly supervised accommodation in the community after an average two to three years. Residents generally make more progress than controls regarding their social functioning, they show increased contact with the community, and higher levels of satisfaction. Again this seems to be associated with the increased privacy and choice associated with domestic scale living arrangements and the more "normal" atmosphere. For those who are not resettled, the units are more effective in maintaining their functioning than traditional long-stay wards or acute admission units. Costs are generally less than acute beds, but greater than long-stay wards in mental hospitals.

Despite these rather positive results, such options still remain unattractive to many of the people they are specifically designed to serve. For them, privacy, autonomy, and practical help with money, food, laundry, etc. may be much more important than professional support, medication, and considerations of public safety (Rose & Muijen, 1997). There is therefore a problem: on the one hand it is clearly desirable to go along as far as possible with users' expressed preferences, or we risk losing their trust and undermining their engagement with services. On the other, we need to ensure that the settings in which people are cared for are both cost-effective and safe. This is a real dilemma. The only way forward would seem to continue with the development of ordinary housing options, but combine this with intensive support from specialized teams (see below).

In relation to work and employment, the circumstantial evidence in favor of the effectiveness of work to favorable long-term outcomes for people with serious mental illness is very strong (Shepherd, 1997; Warner, 1985). However, the specific evidence in favor of the effectiveness of particular models of vocational rehabilitation—TEPs, social firms, cooperatives, "job coach" models, etc.—is rather weak (Grove, Frendenberg, Harding, & O'Flynn, 1997). Many of these projects are small scale and do not lend themselves easily to conventional random controlled trials. Hence, most of the available evidence consists of uncontrolled, descriptive studies (e.g., McCrum, Burnside, & Duffy, 1997; Nehring, Hill, & Poole, 1993). An interesting recent report concerns the employment of service users in various capacities within mental health services (Perkins, Buckfield, & Choy, 1997).

The best controlled evidence comes from the US and this has recently been reviewed by Bond, Drake, Mueser, and Becker (1997). They examined seven uncontrolled and six controlled studies of vocational rehabilitation programs and concluded that supported employment models were significantly more effective in achieving successful placements in competitive employment compared with conventional approaches (unweighted mean 58% for experimental programs vs. 21% for controls). They also suggest that "place-and-train" models seem much more effective than the more traditional "train-and-place" approach. Bond et al. emphasize the need to integrate clinical and vocational supports, as in the program described by Becker and Drake (1994).

The outcome evidence from controlled trials of specialist assertive community teams suggests that they can retain contact and improve levels of engagement with the most difficult and disabled people currently presenting to community services (Burns & Santos, 1995; Ford et al., 1996). The outcomes in terms of enhanced engagement and maintained contact are particularly important as loss of contact has been a common feature of the unfortunate incidents where community care has "failed." The use of inpatient admission can also be significantly reduced without any disadvantages in terms of symptom stabilization and control (Marks et al., 1994; Muijen, Marks, Connolly, Audini, & McNamee, 1992; Olfson, 1990; Taube, Morlock, Burns, & Santos, 1990; Wright, Heiman, Shupe, & Olvera, 1989). These effects on hospitalization rates are important, both in clinical and financial terms, and seem to be stronger with regard to reducing the number of days in hospital, rather than the number of admissions. Costs are generally slightly lower in the community treatment conditions compared with standard hospital care, mainly because of the reduction in hospital days (Muijen et al., 1992; Taube et al., 1990).

Satisfaction ratings on the part of both patients and families also tend to be much higher for these community-based interventions compared with standard hospital care and this persists even when other treatment gains may have been lost (Marks et al., 1994). This seems at least partly attributable to the reduced use of inpatient admission, underlining the unpopularity of acute hospital admissions among many service users. Increased satisfaction is also undoubtedly related to the greater input of time, better continuity of care, and more effective follow-up, which are central elements of successful intensive community support.

There is some variability in these results, but this is largely due to "fidelity" problems (i.e., failure to adhere to interventions of known effectiveness, see Section 6.23.8).

In relation to crisis intervention *per se*, Kluiter (1997) has reviewed the outcome evidence and, based on eight studies of home-based treatment, he estimated that the average number of days in hospital was less than a third (17:60) for the home-based treatment groups compared with the conventionally-treated inpatient groups. He concluded that various community alternatives (day treatment, home-based treatment, assertive case management, etc.) can produce at least equivalent outcomes to inpatient admission for the majority of those referred. Community-based services generally showed no difference with regard to symptomatic outcomes providing there was adequate access to accurate diagnosis and effective medication regimes, social recovery was usually quicker and, as with longer-term intensive support, user and carer satisfaction was enhanced. Not surprisingly costs were less where there were significant reductions in inpatient days (Knapp et al., 1994).

The evidence in favor of standard acute inpatient care is surprisingly limited. This is particularly striking given the fact that acute inpatient care generally accounts for about three-quarters of the total budget for most mental health services. As indicated, acute care has often been used as the "control condition" against which various community alternatives have been compared and it has seldom been found to be superior. Of course, it has to be acknowledged that most of these alternative community models (crisis teams, day hospitals, etc.) are quite selective in terms of who they will consider and there are some individuals who are simply too disturbed, too much of a danger to themselves or others, or too vulnerable in other ways, for admission to be avoided.

For example, Creed's evaluation of acute day hospital treatment suggested that inpatient admission was only feasible for approximately 60% of those referred (Creed et al., 1990). Similarly, in Hoult's famous study of community treatment in Sydney, 40% of the experimental group were actually hospitalized initially (compared with 98% of controls). However, the experimental group then went on to spend an average of only 8.4 days in hospital in the succeeding year, compared with 53.5 days for the controls (Hoult, Reynolds, Charboneau-Powis, Weekes, & Briggs, 1993). The aim of community-based models should therefore not be to avoid admission "at all costs": it should be to offer treatment in the least restrictive, and most appropriate, setting and in many instances this will be hospital. The more important question is then, "What kind of treatment do they receive while they are there?"

Geddes et al. (1996) studied all patients admitted to one ward over a one-month period ($n=43$) and found that less than two-thirds (65%) received medication regimes which could be said to conform to "evidence-based" practice. Regarding therapeutic interventions other than medication, researchers from the Sainsbury Centre for Mental Health (a research-oriented mental health charity in the UK) studied over 200 consecutive admissions across nine sites and noted a marked absence of evidence-based practice. The most common interventions for inpatients were "creative" therapies (art, drama, music–46%), social activities (33%), ward groups (23%), and relaxation (21%). Only 5% of the patients received any specific psychological therapies (Beardsmoore, personal communication) and there was essentially no sign of effective psychosocial interventions like CBT (Drury, Birchwood, Cochrane, & Macmillan, 1996a, 1996b) or EE-based family interventions (Linszen et al., 1996). The links between community and inpatient teams were also so weak that it was difficult for the ward-based staff to make valid assessments of the patients' needs and therefore to plan effective community management strategies. These rather depressing findings underline the gap which is still present in many acute settings between current practice and evidence-based regimes and highlight the important contribution that clinical psychologists might make to improving effectiveness in acute inpatient settings.

6.23.8 METHODOLOGICAL PROBLEMS

Regarding the outcome literature on models of community care, there are the usual methodological problems of poor sample definition, variable quality outcome measures, unexplained sample attrition, etc. (see Shepherd, 1998b). One of the more important issues concerns the paucity of long-term follow-up data. The problems of living with psychosis are, inherently, long term and it is therefore crucial that good data are available on long-term outcomes (even if this has to sacrifice some rigor in the design). Most complex psychosocial interventions tend to show a gradual diminution over time, particularly after the intervention stops. This is not surprising and simply demonstrates the importance of maintaining good continuity of care and an active state of treatment over long periods of time. Successfully managing psychosis is not like prescribing a course of antibiotics, it requires effective management strategies as well as effective treatment interventions (Shepherd, 1984).

Complex psychosocial interventions must also be given sufficient time to work. It is unreasonable to expect that significant improvements in housing status, occupation, or social networks can be achieved in three (or even six) months. Bond, Miller, Krumwied, and Ward (1988) note that such changes may take two to three years and that is important to give assertive case management programs a chance to settle down. Having said this, there are also examples in the literature of long-term effects of case management programs (e.g., McRae, Higgins, Lycan, & Sherman, 1990). It is certainly not the case that such programs cannot be sustained and that staff will inevitably burn out. As indicated earlier, the PACT program in Madison has been operating for more than 20 years and intensive support and crisis teams now cover most of the Australian states of New South Wales and Victoria. Sustainability is therefore not impossible but it depends on good training, support, and leadership. More will be said about this in Section 6.23.10.

A second major methodological problem concerns poor fidelity of the independent variables (i.e., the interventions themselves). Thus, whether it is a housing project, a work program, or a specialist treatment team, one often never quite knows to what extent the intervention being evaluated actually contains those elements which one would suppose, based on previous research, are likely to be associated with positive outcomes. Without a clear specification of the independent variable results, especially negative or "no-difference" results, it becomes very difficult to interpret and it is easy to conclude that a particular kind of intervention does not work when, in reality, it has never actually been tested. This problem is particularly important with regard to specialist community teams where some practitioners seem to believe that by simply calling a team intensive, or saying that it is practicing case management, means that this is what is actually happening.

In order to improve the "fidelity" of complex, community interventions, Teague, Drake, and Ackeson (1995) have set out some of the criteria that would need to be present in specialist, assertive, community teams if they are to adhere to models of known effectiveness. These are shown in Table 3. Such a list provides a convenient way for both purchasers and providers to check on the quality of the service being provided. If it does not match up to these criteria, then there is little reason to suppose that it will be effective.

The final methodological problem that it is worth highlighting is the magic ingredient without which these complex interventions do not seem to work. This seems to be different in different settings. In some cases it may be associated with the expertise or charisma of individual practitioners. In others it seems to be more a reflection of local conditions, "readiness to change," enthusiasm, staff morale, etc. In either event the literature is full of examples of apparently the same intervention which works in some places, but not in others (e.g., Bond et al., 1988; Creed et al., 1991). This phenomenon should not be dismissed as "Hawthorne effects" or "nonspecifics." It reflects the little understood processes of organizational change and development. If we could articulate and define it more precisely, then clearly we would be in a much better position to engineer more effective service change.

6.23.9 DEVELOPING A "USER-LED" SERVICE

The desire to involve users (and their carers) more fully in the process of service planning and delivery is now widely accepted, but most places are struggling to translate this laudable ambition into reality. How can we appropriately involve service users when their expressed wishes sometimes seem unrealistic, self-defeating, even frankly irrational? How can we appropriately involve service users when, under some conditions, we are prepared to deprive them of their liberty and force them to accept treatment against their will? How can we strike the right balance between the needs of users and those of carers when they often seem diametrically opposed?

These are not easy questions to answer. They touch on some of the basic moral and ethical dilemmas in psychiatry. They are certainly very difficult to answer in any general sense and are probably best addressed in relation to specific instances. However, historically, the voice of the user (and the carer) has been much neglected in psychiatry and there is therefore much to be done to redress the balance of power between user, carer, and professional opinions. There are signs that this is beginning to happen, but progress is slow. It is also clear that the involvement of users in making decisions about the content and structure of their care is now not just a matter of ethics and civil rights, it is a pragmatic necessity. Unless we can actively engage some of the most disturbed and disabled young people currently presenting to mental health services, then the consequences may be very serious for them and for others. We need to give them a stronger sense that services are for them and that they understand their needs and

Table 3 Ingredients of effective "high support" (assertive outreach) teams.

Small caseloads (n = 10–15)
Intensive contacts (1–2 hours/week)
Emphasis on "assertive outreach" (i.e., > 75% of contacts in community)
Team initiates all admissions to hospital (and discharges)
"Team" approach to sharing difficult clients (as opposed to strict "individual" case management)
Service not restricted to Mon.–Fri./9–5 pm
Regular (i.e., daily). team meetings
Clear operational management
Easy access to inpatient beds (preferably same RMO as community)

Source: Teague et al. (1995).

priorities. They need to feel some sense of control and "ownership" of their own care, even if—in extreme circumstances—this has to be taken away. So, what would mental health services look like if they were designed by service users, instead of clinicians and managers?

Finding out what users really think about services is not easy. First, there are problems with sampling and representativeness and care has to be taken to ensure that the group of users consulted does reflect the broad range of user opinion and is not simply constructed from those who are currently attending the service, or who are already active in local user groups. They are, by definition, biased groups and, although the views of "activists" are obviously important, they may not be representative of the "silent majority" of those who are not active (and possibly not very engaged) in the services they receive.

The method chosen to measure users' views may also significantly influence the results. For example, some qualitative approaches (e.g., focus groups) may give very different answers compared with paper-and-pencil questionnaires. Again, this is often attributable to the dynamics of small groups which favor those who are more assertive (and generally more "radical"). On the other hand, standardized, quantitative measures may be more easily replicable, but may lack the depth and "richness" of more qualitative approaches (and therefore have limited validity). A mixture of quantitative and qualitative methods is therefore preferable (see Mechanic, 1989). With careful sampling and multiple methods, it is therefore possible to survey users' and carers' views and then compare these with professionals' views. The results can then sometimes be quite surprising.

For example, Shepherd et al. (1995) asked over 400 patients, relatives, and professionals what they thought was most important for the care of people with schizophrenia in the community. The data were collected across six different districts, with a mixture of rural and urban settings. The patient and carer samples included some people who were contacted through mental health professionals (including a number who were not well engaged with current services) and others who were active in local user or carer groups. A mixture of quantitative (e.g., structured questionnaires) and qualitative methods (e.g., focus groups) were used. The results showed a reasonably good consensus between the three groups regarding what should be provided (housing, day care, symptomatic control, crisis planning, etc.) but significant differences between them regarding the relative importance attached to each element. Users emphasized the importance of practical help—housing, money, physical health care—while the professionals tended to rate professional interventions—treatment, support, symptom control, etc.—as being most important.

There were also some interesting differences between those users consulted through mental health professionals and those contacted through the local user groups. Both rated the provision of information about their individual disorders (diagnosis, treatment, outcome, etc.) and about services in general (availability of day care, accommodation options, etc.) as important; both also noted the difficulties of accessing professionals (or any other mental health services) outside normal working hours (Monday–Friday/9.00–5.00). However, as might be expected, the users who were active in user groups were much more radical in their views and more rejecting of services aimed at providing information about medication and "illness." They clearly did not wish to accept their status as "psychiatric patient" and therefore did not see the relevance of further information. In this respect, they were different from the majority of users and carers.

Findings such as these suggest that users do have different priorities from those of professionals or managers. If they were designing mental health services they would probably begin by addressing their basic needs—somewhere to live, something to do, and a decent level of financial support. They would then build clinical ("illness") services in around this framework. Given all the difficulties discussed earlier about the development of community services and the problems of interagency cooperation, even this seemingly simple prescription is not easy to meet.

The results also highlight the heterogenous nature of "user views." Just like nurses, doctors, or psychologists, there is no one, simple version of users' views. They differ depending on who is consulted and what methods are used. The involvement of users in service planning must therefore proceed on the assumption that there is no magic solution to the problem of user consultation. A variety of methods must be used, with a variety of groups, repeatedly over time. Perhaps most importantly, users must be given the opportunity to speak directly to professionals and be regularly involved in training and education programs. Only by regular, face-to-face contact, in a setting that allows users time to speak and professionals time to listen, can one be confident that the users' agenda will be kept in the foreground.

As indicated, one area in which the majority of users agree is regarding the importance of information about their disorders (causes, treatments, outcomes, etc.). By giving information the user is empowered to take a more active role in managing their own care just as in the management of a long-term physical health condition like heart disease, diabetes, or asthma. The aim is to provide information and support so that those affected can manage their own disorder as much as possible and prevent the accumulation of secondary disabilities. Of course, this approach will not work with everyone (it does not work with all those suffering from diabetes or coronary heart disease) but it does imply a fundamental shift in attitude (and power) away from the professional and towards the patient and his/her family or other carers. The role of the professional is then to be "on tap, not on top" providing expert knowledge and support, but not taking away from the individual their ultimate responsibility for their own health care and its management.

If professionals are going to be able to function in this new kind of role, then they will have to be prepared to share information and knowledge in a much more open way than has been common in mental health services up to now. For far too long mental health professionals have been overly secretive about what they do and why. They have felt uncomfortable about trying to answer straight questions like "Why did I get ill? Will I get better? What do I have to do to prevent another breakdown? Sometimes this is because they simply do not know, not surprisingly, since these are very difficult questions. Sometimes they may feel that their credibility will be undermined if they reveal a degree of ignorance (despite the fact that most patients would prefer an honest answer to defensive interpretation). But sometimes they may rationalize their silence by appeals to confidentiality and professional values. In effect, whether consciously or unconsciously, they are then using their access to information to maintain their position of power and authority. A psychoeducational approach fundamentally challenges this position. Except in very special circumstances, a psychoeducational model assumes that patients have a right to know what professionals think they know about them and why they intend to treat or manage them in particular ways—whether they agree with this or not. An educational model is thus very different from a therapeutic model: good teachers aim to empower their students by sharing their knowledge and skills, not disempower them by concealment and manipulation.

To date, the evidence in favor of the effectiveness of psychoeducational approaches is mixed. There do seem to be some measurable gains in knowledge resulting from the provision of simple, clear information to both users and carers, but the increases tend to be small and are heavily influenced by the receiver's preconceived ideas (Sidley, Smith, & Howells, 1991; Smith & Birchwood, 1987; Smith, Birchwood, & Haddrell, 1992). Psychoeducation thus seems to work best when integrated with a comprehensive psychosocial "package," including problem-solving, skills training, etc. (e.g., Hogarty et al., 1991). At the heart of psychoeducational packages is an attempt to help the person—and their immediate carers—recognize the early warning signs of possible relapse and intervene accordingly (Birchwood et al., 1989).

It has been recognized for some time that many patients do not relapse suddenly into a psychotic state. Most go through a "prodromal" (i.e., prepsychotic) period, often characterized by an increase in neurotic symptoms (agitation, sleeplessness, poor concentration) before clear psychotic symptoms appear. Around two-thirds of people with relapsing psychosis can accurately report these early warning signs and over three-quarters of relatives and other carers (Herz & Melville, 1980). Furthermore, the particular pattern of symptomatic changes is often highly specific to each individual patient: they might therefore be said to possess a kind of personalized relapse "signature" (Birchwood, Macmillan, & Smith, 1992). The potential importance of identifying such relapse signatures is obvious. Close monitoring of prepsychotic symptoms may enable patients and clinicians to work together much more effectively and thereby manage with significantly lower overall doses of medication. Since the prevalence of drug side effects is very much related to the size of the dose, and since for many patients it is the severity of side effects which most significantly influences their decision to continue (or not) with medication, these approaches hold out considerable promise in improving medication compliance and delivering the maximum positive benefits from pharmacological interventions while minimizing their adverse side effects.

Other psychological approaches to improving compliance by involving users more fully in their own treatment and management plans have now also been developed (e.g., Corrigan, Liberman, & Engel, 1990; Eckman, Liberman, Phipps, & Blair, 1990; Eckman et al., 1992). One of the most interesting is the application of a technique known as "motivational interviewing." This was originally formulated to assist people with tackling problems of drug or alcohol abuse and is based on the assumption that most important behavioral choices contain some element of conflict (e.g., "How do I weigh up the immediate benefits of alcohol or other drug consumption against its longer-term ill effects ?"). In the case of medication compliance in schizophrenia, the question is, "How do I weigh up the degree of symptom relief associated with taking my medication, as opposed to the clear disadvantages in terms of side effects—lethargy, tiredness, weight gain, loss of libido, etc.?" The motivational interviewing intervention is built around a structured exploration of these conflicted beliefs and an effort is made to encourage maximum personal commitment to whatever final decision is reached (Miller & Rollnick, 1991). Results from a random controlled trial using this approach suggest that such interventions cannot only improve compliance and enhance symptomatic outcomes, they can also increase community tenure up to 18 months postdischarge (Kemp, Hayward, Applewhaite, Everitt, & David, 1996). Such results are very encouraging and one would hope to see more of these new approaches aimed at addressing some of the psychological issues around medication compliance. The effectiveness of medication in the treatment and management of severe mental illness is, after all, as much determined by the psychological decisions surrounding the decisions to take it (or not) as it is about its pharmacological action.

A psychoeducational format therefore involve users in their own care in a way that is quite different from the traditional patient–therapist relationship. The therapist talks openly about symptoms, he/she acknowledges that these experiences are not uncommon, and that it isn't therefore just a personal nightmare. The individual is thus helped to develop a new "relationship" with their disorder. Instead of being, "a schizophrenic," they become, "a person who is trying to manage their schizophrenia," not a schizophrenic. There is a parallel here with people with anxiety or depressive symptoms who develop a different relationship with their symptoms as a result of successful treatment. They start to see the possibility of exerting some degree of control over these painful and distressing experiences, rather than being controlled by them. This can be the basis for a more positive adjustment to their condition.

6.23.10 STAFF TRAINING

All this implies some very fundamental changes in the way that mental health professionals are currently trained and supported. What do we need to make staff more effective in working with people with severe and enduring mental health problems?

First, we need to ensure that they are able to deliver the range of psychosocial treatments of known effectiveness. We have emphasized throughout this chapter that rehabilitation must proceed from a basis of effective treatment, yet it is still rare to find staff who are well enough trained to deliver the full range of psychosocial interventions that have been demonstrated to be effective with this client group. For example, despite the familiar problems of lack of generalization and maintenance, traditional skills training models clearly have some value in promoting the acquisition of social and other skills (Corrigan, 1991). The effectiveness of skills training models are also likely to be enhanced if they are combined with suitable medication regimes, behavioral family management programs, and psychoeducational approaches (Hogarty et al., 1991). Similarly, the evidence for the effectiveness of EE-based family interventions in preventing—or at least postponing—relapse in schizophrenia is now very strong (Barrowclough & Tarrier, 1998; Mari & Streiner, 1994), again particularly in combination with medication and psychological interventions. Finally, CBT has produced some promising evidence for its effectiveness in terms of reducing symptom severity and alleviating some of the subjective distress associated with residual psychotic symptoms in outpatient samples (Kuipers et al., 1997) and, as indicated earlier, there are some very interesting reports of CBT applied in inpatient settings resulting in significantly more rapid symptomatic recovery and reduced overall lengths of admission (Drury et al., 1996a, 1996b). There is thus a very strong evidence base for the effectiveness of a number of psychological approaches and a considerable task to be done in terms of training and supporting staff to implement such interventions.

Staff must also be prepared to deliver care which is sensitively tuned to the "user agenda." As indicated earlier, this means that services have to be delivered much more on the user's terms and much less according to a "professional view" of what is most important. This can be difficult for those staff with a strong desire to show off their therapeutic skills, but there is a kind of "higher professionalism" which recognizes that sometimes sorting out someone's housing benefits and finding them meaningful work is more important—and makes a better starting point—for a therapeutic relationship than delivering "state-of-the-art" cognitive-behavioral intervention for residual psychotic symptoms, or detailed assessments of cognitive functioning, self-care skills, etc.

Staff must also be credible. What makes staff credible will be different for different users: some may prefer younger workers, some may prefer older, some may prefer the same gender, same ethnic group, etc. Whatever their background, users must feel that the worker has a real understanding of their situation and is able to offer something constructive to help. Teams therefore require a mixture of ages, backgrounds, experience, etc. Some users may even prefer staff who themselves have direct experience of receiving mental health services and there are some interesting examples from the US of service users receiving training and then acting as case managers (e.g., Mowbray et al., 1996; Solomon & Draine, 1995).

But it is not sufficient that staff simply have an intellectual understanding of users' problems, they must be able to empathize with their situation and convey this effectively. Users should also feel that staff like them (at least most of the time!). Sensitivity, humor, honesty, warmth may all sound like cliches, but they are central to forming effective working relationships in this, as in any other, form of therapeutic activity. Users may need to be seen in their own homes, in the street, in shops, the park, the pub, and the cafe. Staff need to be prepared to work outside normal offices and formal facilities; they must be prepared to work alongside the user in a variety of apparently mundane activities and recognize that these are all vital opportunities for assessment and treatment. They must be creative in their ability to exploit these contacts and to extract the maximum therapeutic benefit from what may often be unconventional encounters. Staff must also be very clear about the limits of their own competence (whatever their specific areas of expertise) and be prepared to enlist the help of colleagues from the same or other agencies (e.g., in housing, employment agencies, etc.) who also have expert knowledge in their own fields. This openness to working with others and clear recognition of one's own limitations are key ingredients for effective staff working with this client group.

Can we say anything more about the kind of personality that we should be looking for? As indicated earlier, the concept of low EE has recently been extended to staff and other informal carers (Ball et al., 1992; Moore et al., 1992) and this forms a useful framework for thinking about staff characteristics. It is a common clinical observation that staff often make the same "blame attributions" for negative symptoms which seem to underlie the responses of high EE relatives (Brewin, MacCarthy, Duda, & Vaughn, 1991). It is therefore possible that the same kinds of educational and skills-based programs that have been developed for families (e.g., psychoeducation, problem-solving, communication training, etc.) may be equally relevant for staff (Ranz, Horen, McFarlane, & Zito, 1991). Low EE staff will not get too irritated or frustrated in the face of repeated failure and avoid blaming the patient for his or her difficulties. This does seem to be related to a particular kind of personality type, but may also be facilitated by having clear and realistic expectations for change, which in turn is based on good support and supervision.

Effective leadership, organization, and management of staff are therefore also crucial. Leaders can be from any discipline, but they need both good analytic skills (to break down problems, set priorities, assign tasks, etc.) and good socio-emotional skills (to involve and support others, develop a strong sense of common ownership, commit them to carrying out agreed actions, etc.). There is a strong consensus now regarding the importance of team case management systems, as opposed to individually-based case managers (Stein, 1992; Test, 1979). Obviously, there have to be clear individual accountabilities, but this does not remove the necessity for good teamwork and sharing of responsibilities—and, to a certain extent, clients—within the team. This is necessary to cover for staff when they are absent or unavailable. Effective teamwork is really the only way in which good continuity of care can be maintained.

In the UK, the management of community teams has been identified as an important area of weakness (Onyett, Pillinger, & Muijen, 1994). This has been partly attributed to a lack of managerial authority on the part of team leaders, related to the continuing tension between professional and operational management. The outcome of this confusion is a lack of clarity regarding the roles and accountabilities of individual staff and this is associated with feelings of dissatisfaction and emotional exhaustion (burnout). Onyett et al. argue that in order to counteract this, team leaders should have clear managerial authority over operational matters (i.e., day-to-day running of the team) even if this means weakening professional line management. This may be controversial in some countries, but seems necessary if teams are to function effectively.

In the end, teams are delicate and complex structures. They require good training, good leadership, good organization and management, and good support. None of these components can be ignored if they are going to function to their maximum effect.

6.23.11 CONCLUSIONS

This chapter has covered a considerable range of topics. The effective treatment and management of people with severe and enduring mental health problems raises issues at all levels of services—individual, team, facility, subculture—and these are played out differently in different countries with differing social conditions, welfare systems, levels of unemployment, etc. In my view, this is what makes rehabilitation such a rich and intellectually challenging field and there is much that clinical psychology has contributed to the analysis and solution of these problems. But clinical psychology will never be enough on its own. Psychologists have to learn to work with other professionals—doctors, nurses, social workers, occupational and vocational therapists—and a range of social scientists—medical researchers, sociologists, economists—to produce integrated strategies to address these complex problems.

At their heart remain the experiences of patients, staff, and families. People with severe and enduring mental health problems the world over are still waiting to see services emerge which will effectively address their very basic—and seemingly very simple—needs. Somewhere to live, something to do, and a decent level of financial support. These do not seem unreasonable, or unrealistic, demands, but it is still rare to find them completely met. Similarly, staff need to be trained in effective interventions and then organized and supported in such a way that these can be delivered in the most efficient way possible. Again such conditions are rarely met. Finally, there are families and other carers. Families provide more health care than any formal system of professional care ever does, or ever will do, and severe and enduring mental illness is no exception. Yet families are still often ignored, seldom consulted, and sometimes even blamed for being the cause of the problem. It is not surprising that many are frustrated to the point of bitterness, asking "How long will it be before we are treated as genuine *partners* in care?" As professionals we have a responsibility to respond to this question and to ensure that families are included in the process of care and not simply exploited as a convenient (and cheap) resource.

So, there is much to be done. But, nevertheless I am optimistic. I believe that systems of care for people with severe and enduring mental illness have generally improved over the last 40 years and there are some outstanding examples, in several countries, of truly comprehensive, community-based services which do address peoples' housing, vocational, social, and clinical needs. Our task is to make these the norm rather than the exception. To achieve this, we must combine what the Italian philosopher Gramsci referred to as, "The optimism of will, tempered with the pessimism of reason." This is surely the task for the next millenium.

6.23.12 REFERENCES

Abrahamson, D., Swatton, J., & Wills, W. (1989). Do long-stay patients want to leave hospital? *Health Trends, 21,* 16–21.

Audit Commission (1994). *Finding a place*. London: HMSO.

Bachrach, L. L. (1976). *Deinstitutionalisation: An analytical review and sociological perspective*. Maryland: US Department of Health, Education and Welfare.

Bachrach, L. L. (1988). Defining chronic mental ilness: A concept paper. *Hospital and Community Psychiatry, 39,* 383–388.

Bachrach, L. L. (1992). What we know about homelessness among mentally ill persons: An analytical review and commentary. *Hospital and Community Psychiatry, 45,* 453–464.

Bachrach, L. L. (1997). Lessons from the American experience in providing community-based services. In J. Leff (Ed.), *Care in the community—illusion or reality?* (pp. 21–36). Chichester, UK: Wiley.

Ball, R., Moore, E., & Kuipers, L. (1992). Expressed emotion in community care staff. *Social Psychiatry and Psychiatric Epidemiology, 27,* 35–39.

Barrowclough, C., & Tarrier, N. (1998). Social functioning and family interventions. In K. T. Meuser & N. Tarrier

(Eds.), *Social functioning and schizophrenia* (pp. 327–341). New York: Allyn & Bacon.

Beard, J., Propst, R. N., & Malamud, T. J. (1982). The fountain house model of psychiatric rehabilitation. *Psychosocial Rehabilitation Journal, 5,* 47–59.

Becker, D. R., & Drake, R. E. (1994). Individual placement and support: A community mental health centre approach to vocational rehabilitation. *Community Mental Health Journal, 30,* 193–206.

Beecham, J., Hallam, A., Knapp, M., Baines, B., Fenyo, A., & Asbury, M. (1997). Costing care in hospital and in the community. In J. Leff (Ed.), *Care in the Community—illusion or reality?* (pp. 93–109). Chichester, UK: Wiley.

Bhughra, D. (1996). *Homelessness and mental health* (pp. 187–229). Cambridge, UK: Cambridge University Press.

Bigelow, D.A., Cutler, D.J., Moore, L.J., McComb, P., & Leung, P. (1988). Characteristics of state hospital patients who are hard to place. *Hospital and Community Psychiatry, 39,* 181–185

Birchwood, M., MacMillan, F., & Smith, J. (1992). Early intervention. In M. Birchwood & N. Tarrier (Eds.), *Innovations in the psychological management of schizophrenia* (pp. 115–145). Chichester, UK: Wiley.

Birchwood, M., Smith, J., MacMillan, F., Hogg, B., Prasad, R., Harvey, C., & Bering, S. (1989). Predicting relapse in schizophrenia: the development and implementation of an early signs monitoring system using patients and families as observers. *Psychological Medicine, 19,* 649–656.

Bond, G. R., Drake, R. E., Mueser, K. T., & Becker, D. R. (1997). An update on supported employment for people with severe mental illness. *Psychiatric Services, 48,* 335–346.

Bond, G. R., Miller, L. D., Krumwied, R. D., & Ward, R. S. (1988). Assertive case management in three CMHCs: A controlled study. *Hospital & Community Psychiatry, 39,* 411– 418.

Breakey, W. R. (1996). Clinical work with homeless people in the U.S.A. In D. Bhugra (Ed.), *Homelessness and mental health* (pp. 110–132). Cambridge, UK: Cambridge University Press.

Breier, A., & Strauss, J.S. (1984). The role of social relationships in the recovery from psychotic disorders. *American Journal of Psychiatry, 141,* 949–955.

Brewin, C. R., MacCarthy, B., Duda, K., & Vaughn, C. E. (1991). Attribution and expressed emotion in the relatives of patients with schizophrenia. *Journal of Abnormal Psychology, 100,* 546–554.

Burns, B. J., & Santos, A. B. (1995). Assertive community treatment: An update of randomized trials. *Psychiatric Services, 46,* 669–675.

Chamberlain, J. (1988). *On our own: Patient controlled alternatives to the mental health system.* London: MIND.

Corrigan, P. W. (1991). Social skills training in adult psychiatric populations: A meta-analysis. *Journal of Behaviour Therapy and Experimental Psychiatry, 22,* 203–210.

Corrigan, P. W., Liberman, R. P., & Engel, J. D. (1990). From noncompliance to collaboration in the treatment of schizophrenia. *Hospital and Community Psychiatry, 41,* 1203–1211.

Craig, T. K. J., Bayliss, E., Klein, O., Manning, P., & Reader, L. (1995). *The homeless mentally ill initiative: An evaluation of four clinical teams.* London: Department of Health.

Creed, F., Black, D., Anthony, P., Osborn, M., Thomas, P. & Tomenson, B. (1990). Randomised controlled trial of day versus inpatient psychiatric treatment. *British Medical Journal, 300,* 1033–1037.

Creed, F., Black, D., Anthony, P., Osborn, M., Thomas, P., Franks, D., Polley, R., Lancashire, S., Saleem, P., & Tomenson, B. (1991). Randomised controlled trial of day and in-patient psychiatric treatment 2: Comparison of two hospitals. *British Journal of Psychiatry, 158,* 183–189.

Cresswell, C. M., Kuipers, L., & Power, M. J. (1992). Social networks and support in long-term psychiatric patients. *Psychological Medicine, 22,* 1010–1026.

De Jong, A., Giel, R., Sloof, C.J., & Wiersma, D. (1986). Relationship between symptomatology and social disability: empirical evidence from a follow-up study of schizophrenic patients. *Social Psychiatry, 21,* 200–205.

Department of Health (1994). *The health of the nation key area handbook* (2nd ed.). London: HMSO.

Department of Health (1995). *Building bridges.* London: HMSO.

Department of Health (1997). *The spectrum of care: Local services for people with mental health problems.* London: HMSO.

Dick, N., & Shepherd, G. (1995). Work and mental health: A preliminary test of Warr's model in sheltered workshops for the mentally ill. *Journal of Mental Health, 3,* 387–400.

Dixon, L. B., Krauss, N., Kernan, E., Lehman, A .F., & DeForge, B. R. (1995). Modifying the PACT model to serve homeless persons with severe mental illness. *Hospital and Community Psychiatry, 46,* 684–688.

Dozier, M., Harris, M., & Bergman, A. (1987). Social network density and rehospitalisation among young adult patients. *Hospital and Community Psychiatry, 38,* 61–65.

Drury, V., Birchwood, M., Cochrane, R., & Macmillan, F. (1996a). Cognitive therapy and recovery from acute psychosis: A controlled trial I. Impact on psychotic symptoms. *British Journal of Psychiatry, 169,* 593–601.

Drury, V., Birchwood, M., Cochrane, R., & Macmillan, F. (1996b). Cognitive therapy and recovery from acute psychosis: A controlled trial II. Impact on recovery time. *British Journal of Psychiatry, 169,* 602–607.

Eckman, T. A., Liberman, R. P., Phipps, C. C., & Blair, K. (1990). Teaching medication management skills to schizophrenic patients. *Journal of Clinical Psychopharmacology, 10,* 33–38.

Eckman, T. A., Wirshing, W. C., Marder, S. R., Liberman, R. P., Johnston-Cronk, K., Zimmerman, K., & Mintz, J. (1992). Technique for training patients in illness self-management: a controlled trial. *American Journal of Psychiatry, 149,* 1549–1555.

Ford, R., Beardsmoore, A., Ryan, P., Repper, J., Craig, T., & Muijen, M. (1995). Providing the safety net: Case management for people with serious mental illness. *Journal of Mental Health, 4,* 91–97.

Ford, R., Ryan, P., Norton, P., Beardsmoore, A., Craig, T., & Muijen, M. (1996). Does intensive case management work? Clinical, social and quality of life outcomes from a controlled study. *Journal of Mental Health, 5,* 361–368.

Geddes, J. R., Game, D., Jenkins, N. E., Peterson, L. A., Pottinger, G. R., & Sackett, D. L. (1996). What proportion of primary psychiatric interventions are based on evidence from randomised controlled trials? *Quality in Health Care, 5,* 215–217.

Goering, P., Durbin, J., Foster, R., Bayles, S., Babiak, T., & Lancee, B. (1992). Social networks of residents in supported housing. *Community Mental Health Journal, 28,* 199–213.

Goffman, E. (1961). *Asylums: Essays on the social situation of mental patients and other inmates.* New York: Doubleday.

Goldman, H. H., Morrisey, J. P., & Ridgely, M .S. (1994). Evaluation of the Robert Wood Johnson Foundation Program on chronic mental illness. *Millbank Quarterly, 72,* 37–47.

Grove, R. (1994). Supported employment and social firms.

In M. Floyd, M. Povall, & G. Watson (Eds.), *Mental health at work* (pp. 96–99). London: Kingsley.

Grove, R., Freudenberg, M., Harding, A., & O'Flynn, D. (1997). *The social firm handbook*. Brighton, UK: Pavilion.

Gudeman, J. E., & Shore, M. F. (1984). Beyond deinstitutionalisation: A new class of facilities for the mentally ill. *New England Journal of Medicine, 311,* 832–836.

Hambridge, J. A., & Rosen, A. (1994). Assertive community treatment for the seriously mentally ill in suburban Sydney: A programme description and evaluation. *Australian & New Zealand Journal of Psychiatry, 28,* 438–445.

Harris, M., Bergman, H. C., & Bachrach, L. (1986). Individualised network planning for chronic psychiatric patients. *Psychiatric Quarterly, 58,* 51–56.

Herz, M., & Melville, C. (1980). Relapse in schizophrenia. *American Journal of Psychiatry, 137,* 801–812.

Hogarty, G. E., Anderson, C. M., Reiss, D. J., Kornblith, S. J., Greenwald, D. P., Ulrich, R. F., & Carter, M. (1991). Family psychoeducation, social skills training, and maintainance chemotherapy in the aftercare treatment of schizophrenia. *Archives of General Psychiatry, 48,* 340–347.

Holmes-Eber, P., & Riger, S. (1990). Hospitalisation and the composition of mental patients' social networks. *Schizophrenia Bulletin, 16,* 157–164.

Hoult, J., Reynolds, I., Charboneau-Powis, M., Weekes, P., & Briggs, J. (1983). Psychiatric hospital versus community treatment: The results of a randomised trial. *Australian & New Zealand Journal of Psychiatry, 17,* 160–167.

Jones, K. (1972). *A history of the mental health services*. London: Routledge & Kegan Paul.

Kemp, R., Hayward, P., Appletwhaite, G., Everitt, B., & David, A. (1996). Compliance therapy in psychotic patients: a randomised controlled trial. *British Medical Journal, 312,* 345–349.

Kluiter, H. (1997). Inpatient treatment and care arrangements to replace or avoid it—searching for an evidence-based balance. *Current Opinion in Psychiatry, 10,* 160–167.

Knapp, M., Beecham, J., & Hallam, A. (1997). The mixed economy of psychiatric reprovision. In J. Leff (Ed.), *Care in the community—illusion or reality?* (pp. 37–49). Chichester, UK: Wiley.

Knapp, M., Beecham, J. K., Koutsogeorgopoulou, V., Hallam, A., Fenyo, A., Marks, I. M., Connolly, J., Audini, B., & Muijen, M. (1994). Service use and costs of home-based versus hospital-based care for people with serious mental illness. *British Journal of Psychiatry, 165,* 195–203.

Kuipers, E., Garety, P., Fowler, D., Dunn, G., Bebbington, P., Freeman, D., & Hadley, C. (1997). The London–East Anglia randomised controlled trial of cognitive-behavioural therapy for psychosis 1: Effects of the treatment phase. *British Journal of Psychiatry, 171,* 319–327.

Leff, J. (1997). *Care in the community—illusion or reality?* Chichester, UK: Wiley.

Lehman, A., Postrado, L., Roth, D., McNary, S. W., & Goldman, H. H. (1994). Continuity of care and client outcomes in the Robert Wood Johnson Foundation Program on chronic mental illness. *Millbank Quarterly, 72,* 105–123.

Lelliot, P., & Wing, J.K. (1994). A national audit of new long-stay psychiatric patients. II: Impact on services. *British Journal of Psychiatry, 165,* 170–178.

Lelliot, P., Wing, J. K., & Clifford (1994). A national audit of new long-stay psychiatric patients. I: Method and description of the cohort. *British Journal of Psychiatry, 165,* 160–169.

Linn, M. W., Caffey, E. M., Klett, J., Hogarty, G. E., & Lamb, H. R. (1979). Day treatment and psychotropic drugs in the aftercare of schizophrenic patients. *Archives of General Psychiatry, 36,* 1055–1066.

Linszen, D., Dingeman, P., Van Der Does, J. M., Nugter, A., Scholte, P., Lenoir, R., & Goldstein, M. (1996). Treatment, expressed emotion and relapse in recent onset schizophrenia disorders. *Psychological Medicine, 26,* 565–578.

Mann, S., & Cree, W. (1976). "New" long stay psychiatric patients: A national sample survey of fifteen mental hospitals in England and Wales 1972/73. *Psychological Medicine, 6,* 603–616.

Mari, J. J., & Streiner, D. (1994). An overview of family interventions and relapse in schizophrenia: Meta-analysis of research findings. *Psychological Medicine, 24,* 565–578.

Marks, I. M., Connolly, J., Muijen, M., Audini, B., McNamee, G., & Lawrence, R. E. (1994). Home-based versus hospital-based care for people with serious mental illness. *British Journal of Psychiatry, 165*, 179–194.

McRae, J., Higgins, M., Lycan, C. & Sherman, W. (1990). What happens to patients after five years of intensive case management stops? *Hospital and Community Psychiatry, 41,* 175–179.

McCrum, B. W., Burnside, L. K., & Duffy, T. L. (1997). Organising for work: A job clinic for people with mental health needs. *Journal of Mental Health, 6,* 503–513.

Mechanic, D. (1989). Medical sociology: Some tensions among theory, method, and substance. *Journal of Health and Social Behaviour, 30,* 147–160.

Mechanic, D., & Aiken, L. (1987). Improving the care of patients with chronic mental illness. *New England Journal of Medicine, 317,* 1634–1638.

Miller, W. R., & Rollnick, S. (1991). *Motivational interviewing: Preparing people to change*. New York: Guilford Press.

Mitchell, S., & Birley, J. L. T. (1983). The use of ward support by psychiatric patients in the community. *British Journal of Psychiatry, 142,* 9–15.

Moore, E., Kuipers, E., & Ball, R. (1992). Staff–patient relationships in the care of the long-term mentally ill. *Social Psychiatry and Psychiatric Epidemiology, 27,* 28–34.

Morrisey, J. P., Calloway, M., Bartico, W. T., Ridgely, M. S., Goldman, H. H., & Paulson, R. (1994). Local Mental Health Authorities and service system change: Evidence from the Robert Wood Johnson Foundation Program on chronic mental illness. *Millbank Quarterly, 72,* 49–80.

Mowbray, C. T., Moxley, D. P., Thrasher, S., Bybee, D., McCrohan, N., Harris, S., & Clover, G. (1996). Consumers as community support providers: Issues created by role innovation. *Community Mental Health Journal, 32,* 47–67.

Muijen, M., Marks, M., Connolly, B., Andini, B., & McNamee, G. (1992). The Daily Living Programme. Preliminary comparison of community versus hospital-based treatment for the seriously mentally ill facing emergency admission. *British Journal of Psychiatry, 160,* 379–384.

Nehring, J., Hill, R. G., & Poole, L. (1993). *Work, empowerment and community: Opportunities for people with long-term mental health problems*. London: RDP.

NHS Executive (1996). *24 hour nursed care for people with severe and enduring mental illness*. Leeds, UK: NHS Executive.

Perkins, R., Buckfield, R., & Choy, D. (1997). Access to employment: A supported employment project to enable mental health service users to obtain jobs within mental health teams. *Journal of Mental Health, 6,* 307–318.

Powell, R., & Slade, M. (1996). Defining severe mental illness. In G. Thornicroft & G. Strathdee (Eds.), *Commissioning mental health services* (pp. 13–27). London: HMSO.

Pozner, A., Ng, M. L., Hammond, J., & Shepherd, G. (1996). *Working it out—creating work opportunities for* *people with mental health problems*. Brighton, UK: SCMH/Pavilion Publishing.

Olfson, M. (1990). Assertive community treatment: An evaluation of the experimental evidence. *Hospital and Community Psychiatry, 41,* 634–641.

Onyett, S., Pillinger, T., & Muijen, M. (1994). *Making community mental health teams work*. London: The Sainsbury Centre for Mental Health.

Rachman, S.J., & Hodgson, R. (1974). I. Synchrony and desynchrony in fear and avoidance. *Behaviour Research & Therapy, 12,* 311–318.

Ranz, J. M., Horen, B. T., McFarlane, W. R., & Zito, J. M. (1991). Creating a supportive environment using staff psychoeducation in a supervised residence. *Hospital and Community Psychiatry, 42,* 1154–1159.

Ritchie, J. H., Dick, D., & Lingham, R. (1994). *The report of the inquiry into the care and treatment of Christopher Clunis*. London: HMSO

Rogers, A., Pilgrim, D., & Lacey, R. (1993). *Experiencing psychiatry—users' views of services*. London: MIND/Macmillan

Rose, D., & Muijen, M. (1997). Nursing doubts. *Health Services Journal, 107,* 34–35.

Rosenfield, S., & Neese-Todd, S. (1993). Elements of a psychosocial clubhouse program associated with a satisfying quality of life. *Hospital and Community Psychiatry, 44,* 76–78.

Schinnar, A. P., Rothbard, A. B., Kanter, K., & Jung, Y. S. (1990). An empirical literature review of definitions of severe and persistent mental illness. *American Journal of Psychiatry, 12,* 1602–1608.

Sheets, J., Prevost, J., & Rehman, J. (1982). Young adult chronic patients: three hypothesized subgroups. *Hospital and Community Psychiatry, 33,* 197–203.

Shepherd, G. (1984). Rehabilitation. In B. Bradley, & C. Thompson (Eds.), *Psychological applications in psychiatry* (pp. 201–222). Chichester, UK: Wiley.

Shepherd, G. (1995). The "ward-in-a-house"—Residential care for the severely disabled. *Journal of Mental Health, 31,* 53–69.

Shepherd, G. (1997). Vocational rehabilitation in psychiatry—An historical perspective. In R. Grove, M. Freudendberg, A. Harding, & D. O'Flynn (Eds.), *The Social firm handbook* (pp. 1–7). Brighton, UK: Pavilion Press.

Shepherd, G. (1998a). Social functioning and challenging behaviour. In K. T. Meuser & N. Tarrier (Eds.), *Social functioning and schizophrenia* (pp. 407–423). New York: Allyn & Bacon.

Shepherd, G. (1998b). Models of community care. *Journal of Mental Health, 7,* 165–177.

Shepherd, G., Beardsmoore, A., Moore, C., Hardy, P., & Muijen, M. (1997). Relation between bed use, social deprivation, and overall bed availability in acute adult psychiatric units, and alternative residential options: a cross sectional survey, one day census data and staff interviews. *British Medical Journal, 314,* 262–266.

Shepherd, G., Muijen, M., Dean, R., & Cooney, M. (1996). Residential care in hospital and in the community—quality of care and quality of life. *British Journal of Psychiatry, 168,* 448–456.

Shepherd, G., Murray, A., & Muijen, M. (1995). Perspectives on schizophrenia: The views of users, relatives and professionals. *Journal of Mental Health, 4,* 403–422.

Sidley, G., Smith, J., & Howells, K. (1991). Is it ever too late to learn? Information provision to relatives of long-term schizophrenia sufferers. *Behavioural Psychotherapy, 19,* 305–320.

Smith, J., & Birchwood, M. J. (1987). Specific and non-specific effects of educational intervention with families living with a schizophrenic relative. *British Journal of Psychiatry, 150,* 645–652.

Smith, J., Birchwood, M., & Haddrell, A. (1992). Informing people with schizophrenia about their illness: The effect of residual symptoms. *Journal of Mental Health, 1,* 61–70

Solomon, P., & Draine, J. (1995). The efficacy of a consumer case management team: 2 year outcomes of a randomised controlled trial. *Journal of Mental Health Administration, 22,* 126–134.

Stein, L. (1992). On the abolishment of the case manager. *Health Affairs, 11,* 172–177.

Strathdee, G., Davies, S., Perry, M., & Thompson, K. (1996). Commissioning and managing hospital and community beds. In G. Thornicroft & G. Strathdee (Eds.), *Commissioning mental health services* (pp. 147–165). London: HMSO.

Strauss, J. S., & Carpenter, W. T. (1974). The prediction of outcome in schizophrenia: II. Relationships between predictor and outcome variables. *Archives of General Psychiatry, 31,* 37–42.

Strauss, J. S., & Carpenter, W. T. (1977). The prediction of outcome in schizophrenia: III. Five-year outcome and its predictors. *Archives of General Psychiatry, 34,* 159–163.

Strauss, J. S., Klorman, R., & Kokes, R. F. (1977). Premorbid adjustment in schizophrenia: Part V. Implications of findings for understanding and research application. *Schizophrenia Bulletin, 3,* 240–244.

Stroul, B. A. (1986). *Models of community support services: Approaches to helping persons with long-term mental illness*. Boston: Centre for Psychiatric Rehabilitation.

Tanzman, B. (1993). An overview of surveys of mental health consumers' preferences for housing and support services. *Hospital and Community Psychiatry, 44,* 450–455.

Taube, C.A., Morlock, L., Burns, B. J., & Santos, A. B. (1990). New directions in research on assertive community treatment. *Hospital & Community Psychiatry, 41,* 642–651.

Teague, G. B., Drake, R. E., & Ackerson, T. H. (1995). Evaluating use of continuous treatment teams for persons with mental illness and substance abuse. *Psychiatric Services, 46,* 689–695.

Test, M. A. (1979). Continuity of care. In L. I. Stein (Ed.), *Community support systems for the long term patient* (pp. 31–47). San Francisco: Jossey-Bass.

Test, M. A., & Stein, L. I. (1978). Training in community living: Research design and results. In L. I. Stein & M. A. Test (Eds.), *Alternatives to mental hospital treatment* (pp. 57–74). New York: Plenum.

Thompson, K., Griffiths, E., & Leaf, P. (1990). A historical review of the Madison model of community care. *Hospital and Community Psychiatry, 41,* 625–633.

Timms, P. W. (1993). Origins of homelessness specific to the mentally ill. In K. Fisher (Ed.), *Homelessness, health care and welfare provision* (pp. 100–116). London: Routledge.

Warner, R. (1985). *Recovery from schizophrenia*. London: Routledge & Kegan Paul.

Wing, J. K. (1978). Clinical concepts of schizophrenia. In J. K. Wing (Ed.), *Schizophrenia: Towards a new synthesis* (pp. 17–32). London: Academic Press.

Wing, J. K. (1992). *Epidemiologically-based mental health needs assessments*. London: Royal College of Psychiatrists.

Wing, J. K., Curtis, R. H., & Beevor, A. S. (1996). *HoNOS: Health of the Nation Outcome Scales—Report on Research and Development*. London: Royal College of Psychiatrists, Research Unit.

Wright, R. G., Heiman, J. R., Shupe, J., & Olvera, G. (1989). Defining and measuring stabilization of patients during 4 years of intensive community support. *American Journal of Psychiatry, 146,* 1293–1298.

Young, R. (1991). *Residential needs of severely disabled psychiatric patients—the case for hospital hostels*. London: HMSO.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.24
Somatoform Disorders

GEORG H. EIFERT and CARL W. LEJUEZ
West Virginia University, Morgantown, WV, USA
and
THEO K. BOUMAN
University of Groningen, The Netherlands

6.24.1 CLINICAL PICTURE 544
  6.24.1.1 DSM-IV Classification 544
  6.24.1.2 Disease, Illness, and Somatization 545
6.24.2 HISTORICAL PERSPECTIVES AND DIAGNOSTIC CHANGES 546
  6.24.2.1 Somatization Disorder 546
  6.24.2.2 Hypochondriasis 546
  6.24.2.3 Pain Disorder 546
  6.24.2.4 Conversion Disorder 547
  6.24.2.5 Body Dysmorphic Disorder 547
6.24.3 PREVALENCE, COURSE, AND DEVELOPMENT 547
  6.24.3.1 Somatization Disorder 547
  6.24.3.2 Hypochondriasis 547
  6.24.3.3 Pain Disorder 548
  6.24.3.4 Conversion Disorder 548
  6.24.3.5 Body Dysmorphic Disorder 548
6.24.4 PROBLEMS OF DIFFERENTIAL DIAGNOSIS AND COMORBIDITY 549
  6.24.4.1 Somatoform Disorders vs. General Medical Conditions 549
  6.24.4.2 Differentiating Somatoform Disorders 550
    6.24.4.2.1 Somatization disorder vs. hypochondriasis 550
    6.24.4.2.2 Somatization disorder vs. pain disorder 550
    6.24.4.2.3 Somatization disorder vs. conversion disorder 550
    6.24.4.2.4 Hypochondriasis vs. body dysmorphic disorder 550
  6.24.4.3 Relation Between Somatoform Disorders and Depression 551
  6.24.4.4 From Differential Diagnosis to Dimensional Classification 551
    6.24.4.4.1 Hypochondriasis and disease fear: a dimensional perspective 551
    6.24.4.4.2 Relation between hypochondriasis and anxiety disorders 552
6.24.5 CONTEMPORARY THEORETICAL PERSPECTIVES 553
  6.24.5.1 Dysfunctional Processes in the Somatoform Disorders 553
    6.24.5.1.1 Abnormal illness behavior 553
    6.24.5.1.2 Perception and attention 553
    6.24.5.1.3 Deficits in emotion processing (alexithymia) 554
  6.24.5.2 Disorder-specific Theoretical Models 554
    6.24.5.2.1 Hypochondriasis 554
    6.24.5.2.2 Pain 555
  6.24.5.3 Future Challenge: Integrative Biobehavioral Theories 556
6.24.6 TREATMENT OF SOMATOFORM DISORDERS 557
  6.24.6.1 Somatization Disorder 558

6.24.6.2 Hypochondriasis 559
6.24.6.3 Pain Disorder 559
6.24.6.4 Conversion Disorder 560
6.24.6.5 Body Dysmorphic Disorder 560
6.24.7 CONCLUSIONS 561
6.24.8 REFERENCES 562

6.24.1 CLINICAL PICTURE

Many individuals present to medical practitioners with physical complaints for which no medical explanation can be found. For instance, a European study (van Hemert, Hengeveld, Bolk, Rooijmans, & Vandenbroucke, 1993) found that among 191 new referrals to a general medical outpatient clinic, 52% of patients had symptoms that ultimately remained unexplained. An earlier Australian study reported that of 95 patients visiting a general practitioner and presenting with somatic complaints, 41% had no demonstrable somatic pathology (Pilowsky, Smith, & Katsikitis, 1987). Comparable percentages have been reported in other studies from different countries investigating unexplained symptoms of patients with abdominal pain (Harvey, Salih, & Read, 1983) and chest pain (Eifert, 1992; Mayou, Bryant, Forbar, & Clark, 1994).

Although many of these patients will be satisfied with negative medical examination results, and some reassurance to that effect, a significant subgroup will anxiously continue to ruminate about the possibility of suffering from a yet-undiagnosed physical disease. They are likely to continue to seek help for the same or different physical symptoms, demand more physical examinations and specialist referrals, undergo costly laboratory tests, and, in some cases, even end up on an operating table (Warwick & Salkovskis, 1990). Despite the high prevalence of ambiguous and unexplained physical problems, somatoform disorders have received only minimal research attention and are poorly understood. A major reason for this relative lack of knowledge is that persons with such problems are typically reluctant to see a psychologist and prefer to remain within the medical service system. As a result, they have literally evaded the attention of clinical psychologists.

We have organized our discussion in this chapter around the categories and classification put forward in the *Diagnostic and statistical manual of mental disorders* (4th ed., *DSM-IV*; American Psychiatric Association [APA], 1994). There is considerable overlap between somatoform disorders and several of the anxiety disorders. We will therefore make an attempt to move beyond *DSM* categories toward a more function-based dimensional perspective of the problems of persons who present with either unexplained somatic symptoms or excessive concerns over physical symptoms. We believe that such an approach is clinically more beneficial because it provides information that is directly useful for the design of clinical interventions.

6.24.1.1 *DSM-IV* Classification

The present classification of somatoform disorders stems from the *DSM-III* taskforce decision to eliminate the concept of neurosis because of its etiological connotations. This change resulted in disorders previously classified in one category as neuroses (anxiety neurosis, neurotic depression, hypochondriasis, and hysteria) to be redefined and reclassified in four separate *DSM* categories: anxiety disorders, mood disorders, somatoform disorders, and dissociative disorders (Murphy, 1990). According to *DSM-IV* (APA, 1994), the common feature of the somatoform disorders is the presence of physical symptoms that suggest a general medical condition (hence, the term somatoform) but are not fully explained by a general medical condition, the direct effects of substance, or by another mental disorder. Physical symptoms result in substantial personal, social, and occupational impairment, and are not feigned or voluntarily produced, as in malingering or factitious disorder.

DSM-IV distinguishes between six somatoform disorders: somatization disorder, undifferentiated somatization disorder, conversion disorder, pain disorder, hypochondriasis, and body dysmorphic disorder. These disorders can be categorized under two larger headings of the classical "hysterical disorders" and "preoccupation disorders." Hysterical disorders involve actual loss or alteration in bodily functions as in the case of somatization disorder, conversion, and pain disorder. Reports of suffering often exceed symptom reports of patients with known medical illnesses. Patients with hysterical disorders typically experience no anxiety, whereas persons with preoccupation disorders are excessively concerned that there is something physically wrong with their body either in terms of disease (hypochondriasis) or in terms of the shape and size (body dysmorphic disorder). The general features of each of these disorders will be outlined below.

Somatization disorder is characterized by multiple physical complaints without clear or known physical causes. The condition may last for many years and, in some cases, extend over the entire adult life-span. To make the diagnosis, an individual needs to present with a history of pain related to at least four different sites or functions (e.g., head, back, abdomen, joints), two gastrointestinal symptoms (e.g., diarrhea, food intolerance), one sexual symptom (e.g., irregular menses, indifference to sex), and one pseudoneurological symptom (e.g., poor balance, numbness, paralysis). These symptoms lead to frequent and multiple medical consultations, complex medical history, and to alterations of the person's lifestyle. Physical and laboratory findings cannot detect a plausible medical condition to be the cause of the symptoms, and if a cause exists, the patient's reaction seems to be in excess of what would be expected.

Hypochondriasis is characterized by unjustified fears, suspicions, or convictions that one has a serious and often fatal illness such as heart disease, cancer, or acquired immune deficiency syndrome (AIDS). Patients frequently seek reassurance, check their bodies, and avoid illness-related situations. Merely informing patients of the absence of a disease process, or explaining the benign nature of the symptoms, only results in temporary reassurance which is followed by renewed worry over symptoms and continuing overuse of medical services (Salkovskis & Warwick, 1986).

Pain disorder is characterized by severe acute or chronic pain in one or more body parts. The pain is not easily understood, or cannot be fully accounted for, by a known medical condition. Pain is *ipso facto* a subjective phenomenon and psychological factors such as mood, anxiety, and attention may be involved in the onset, maintenance, or exacerbation of pain and complicate differential diagnosis. Pain is considered acute when it exists for less than six months and chronic when it persists beyond six months. Chronic pain, in particular, is often associated with major changes in behavior such as decreased activity and somatic preoccupation (Pilowsky, Chapman, & Bonica, 1977).

Conversion disorder is characterized by symptoms suggesting a neurological disorder, although appropriate medical investigations fail to identify a neurological or general medical disorder. At times, symptoms may even be inconsistent with general neurological knowledge. *DSM-IV* describes four subtypes with (i) motor symptoms or deficits (e.g., paralysis), (ii) seizures of convulsions, (iii) sensory symptoms or deficits (e.g, blindness, anesthesia, and aphonia), and (iv) a mixed presentation. Some patients show an indifference or lack of worry about their symptoms (*la belle indifference*). An important requirement for the diagnosis is the temporal relation between conversion symptoms and a psychological stressor such as acute grief or victimization. Patients are typically unaware of the psychological basis for their symptoms and report being unable to control them.

Body dysmorphic disorder (BDD) is characterized by a preoccupation with an imagined or exaggerated body disfigurement or an excessive concern that there is something wrong with the shape or appearance of one's own body parts. The perceived defect or abnormality is generally not or hardly noticeable to others. Objects of concern typically refer to the face (nose, mouth, eyes, teeth), head (hair thinning), sexual characteristics (size of penis, breasts), or general body appearance (too long or too short, ugliness). Other concerns may involve scars, wrinkles, or body odor. Cognitive features are excessive preoccupation, intrusive thoughts, and sometimes ideas of reference. On a behavioral level, features include avoidance (e.g., of body exposure, direct social contact, talking about the problem, looking in the mirror), camouflaging or concealing of imagined deformities (wearing a hat or glasses), excessive grooming and checking, and reassurance seeking.

6.24.1.2 Disease, Illness, and Somatization

Disease is often described as the presence of objective biological abnormalities in the structure and/or function of bodily organs and systems. Illness refers to the subjective perception of being unwell, and may be unrelated to the presence of an objectifiable disease. The sick role is a sociological construct initially put forward by the American sociologist Talcot Parsons pertaining to a role granted to an individual by society with accompanying privileges (e.g., staying home from work) and obligations (complying with medical regimens).

Pivotal to the somatoform disorders is the phenomenon of somatization, defined by Lipowsky (1988, p. 275) as "a tendency to experience and express psychological distress in the form of somatic symptoms which the individual misinterprets as significantly serious physical illness and seeks medical help for them." Rather than referring to a discrete group of disorders, this definition reflects the current opinion of somatization as a complex psychopathological process that may be involved in a variety of clinical problems (cf. Kellner, 1986).

Psychosomatic disorders have traditionally been regarded as medical conditions in which psychological factors play a role. The concept of psychosomatic disorders is misleading, however, because it presupposes a special class of somatic disorders of psychogenic etiology (Lipowsky, 1988). A conceptually more neutral concept is that of "functional somatic symptoms," which Kellner (1986) defined as somatic symptoms that are not attributed to organic pathology demonstrable by physical examination or routine laboratory tests—although transient physiological changes may be found in some of the disorders or can be detected by special techniques such as measurement of bowel contractions or of changes in striated muscle tension.

Conceptual proliferation and confusion dominate the field of somatization and somatoform disorders. Many concepts are used interchangeably or remain ill-defined in research and clinical practice. Despite considerable comorbidity, and even phenomenological overlap between psychological and somatic disorders, Cartesian dualism is still implicitly or explicitly influential in clinical practice. This may, in part, be due to the traditional division of labor between medical and psychological disciplines, each claiming (or hoping) to answer some of the questions in the mind–body realm. Such reductionist models, however, are a serious obstacle to a more comprehensive understanding of somatoform disorders.

6.24.2 HISTORICAL PERSPECTIVES AND DIAGNOSTIC CHANGES

Early views of somatoform disorders assumed that these disorders have common roots and are somehow related to the female reproductive system. The French neurologist Charcot (1825–1893) demonstrated that "hysterical conversions," involving symptoms such as convulsions and paralysis, could be induced by hypnotic techniques. He assumed that conversions originate from mental rather than physical processes. Physical symptoms were regarded as a defense mechanism against unacceptable unconscious conflicts where massive repression forces psychic energy to be transformed into bodily symptoms. Although specific psychodynamic notions have been eliminated from the current *DSM-IV*, conflict is still believed to be involved in some of the somatoform disorders. For instance, in the case of conversion disorders "psychological factors are judged to be associated with the symptom or deficit because the initiation or exacerbation of the symptom or deficit is preceded by conflicts or other stressors" (APA, 1994, p. 457).

6.24.2.1 Somatization Disorder

Somatization disorder is the current term for what is arguably the oldest mental health diagnosis, "hysteria." More than 2000 years ago, the ancient Greeks believed that multiple somatic symptoms were caused by the uterus wandering through the female body. In the middle of the nineteenth century, the French physician Briquet described a polysymptomatic somatic condition ("Briquet's syndrome"), which was redefined in the 1950s as somatization disorder (cf. Bass & Murphy, 1990) and became the basis for the current diagnostic entity.

6.24.2.2 Hypochondriasis

Hypochondriasis is a Greek word meaning "below the cartilage." The ancient Greeks derived the concept of hypochondriasis from humoral theories of disease and considered it a special form of melancholia resulting from an excess of black bile. In the seventeenth century, Thomas Sydenham, an English physician, argued that hypochondriasis occurred only in men and was equivalent to hysteria occurring in females. Also around this period, Descartes proposed that the mind and body were separate entities, and there could be no causal relation between the two. Subsequent psychodynamic views suggested that hypochondriacal patients direct their libido inwards, whereas healthy persons typically direct their libido at external objects. Eventually, internally directed libido would build up and result in physical symptoms (Freud, 1956).

6.24.2.3 Pain Disorder

Throughout history, there has been considerable speculation concerning the nature and cause of pain. For instance, Aristotle viewed pain as an emotion as opposed to a sensation, Descartes viewed it as a result of physical stimuli impinging upon the body, Epictetus viewed it as the result of cognitive activity, and religious leaders viewed it as a test of faith imposed by god or punishment for sins (Turk, Meichenbaum, & Genest, 1983). A commonality of these early theories was a unidimensional view of pain. Theories were based either upon organic or psychological causes, and few attempted to integrate the two. It was not until the twentieth century that Cartesian dualism was seriously questioned and multidimensional theories of pain began to develop.

6.24.2.4 Conversion Disorder

The origin of the contemporary concept of conversion disorder can be traced back to the Middle Ages when "conversio" referred to diseases caused by a "suffocation of the womb." Freud considered conversion a form of defense in which the resulting symptoms were fixations of physical patterns relating to events at the time of a patient's (sexual) trauma (Mace, 1992). Primary gain was seen as the warding off of these forbidden impulses, whereas secondary gain was the attention and privileges patients received. The loss of bodily functions in conversion disorder was thought either to be the direct effect of the trauma or to symbolize the unconscious conflict.

6.24.2.5 Body Dysmorphic Disorder

The term "dysmorphophobia" was coined in the late nineteenth century by the Italian psychiatrist Enrico Morselli (cf. Berrios & Kan, 1996). Based upon his clinical observations, Morselli described a condition in which there is a sudden onset and subsequent persistence of an idea that the body is (or might become) deformed. Morselli stressed the obsessive nature of this condition and the strong desire to check the perceived body abnormality. Although Morselli described his patients as being fearful about their deformities, more recent investigations (e.g., Fava, 1992a) could not substantiate the presence of phobic anxiety in this condition. As a result, the term dysmorphophobia was abandoned in *DSM-IV* and replaced with the diagnosis of body dysmorphic disorder (cf. Bass & Murphy, 1995).

6.24.3 PREVALENCE, COURSE, AND DEVELOPMENT

There is a preponderance of studies showing that the presentation of symptoms for which no demonstrable organic pathology can be found is a common occurrence in a variety of medical settings (Pilowsky et al., 1987; van Hemert et al., 1993). Yet the exact prevalence of somatoform disorder is still unclear. This is largely due to the fact that studies frequently did not adequately differentiate between somatoform and related disorders (e.g., anxiety, depression), and epidemiological studies have used various diagnostic criteria and different samples (cf. Kellner, 1986). Although systematic knowledge is lacking, factors such as parental rearing and childhood development, stressful life events, high negative affect, and aspects of the relationship and communication between patients and doctors have been related to the development of somatoform disorders (e.g., Bass & Murphy, 1990; Craig, Boardman, Mills, Daly-Jones, & Drake, 1993).

6.24.3.1 Somatization Disorder

Somatization disorder is a relatively rare phenomenon with recent community studies citing prevalence rates between 0.4% and 0.7% (cf. Bass & Murphy, 1990). Its onset is in early adulthood, course is often chronic, and prognosis is generally regarded as poor. Somatization is associated with frequent absences from work or school, overuse of medical care, and excessive use of drugs and alcohol (APA, 1994). In recent years sexual abuse and traumatization have been cited as a precursor of somatization disorder (Salmon & Calderbank, 1996). For instance, female psychiatric patients with somatization disorder have reported more sexual and physical abuse than patients with other disorders (Pribor, Yutzy, Dean, & Wetzel, 1993).

6.24.3.2 Hypochondriasis

Most prevalence research related to hypochondriasis has been conducted in some type of medical or psychiatric setting and results have been inconclusive (cf. Iezzi & Adams, 1993). Research has been impeded by the use of different definitions of hypochondriasis. For example, Kenyon (1976) used a strict definition and determined that 3–14% of patients in a medical setting were hypochondriacal. With regard to illness fears, Agras, Sylvester, and Oliveau (1969) interviewed 325 randomly selected subjects from the general population and found that 16% had fears of illness and 3.1% qualified as having an illness phobia. Moreover, individuals with frequent exposure to medical settings (e.g., medical students) appear to have increased health concerns (Hunter, Lohrenz, & Schwartzman, 1964; Jacob & Turner, 1984). The onset of hypochondriasis is frequently in early adulthood. Although symptoms may wax and wane, the course is typically chronic and the condition frequently takes a dominant role in the person's life and relationships.

In children, somatic complaints and attention to physical symptoms have been shown to be influenced by attention received from parents (Mechanic, 1964). In addition, adult patients who rate themselves as having high hypochondriacal concerns describe their parents as more caring and overprotective than patients with other psychiatric disorders (Parker & Libscombe, 1980). Studies also found that somatic symptoms of child and adult hypochondriacal patients resemble those of their parents (Apley, 1958). An individual's perception of and exposure to physical symptoms and disease in the family, in combination with parental attitudes towards illness influencing the development of hypochondriacal concerns in children (Eifert, 1992; Flor, Birbaumer, & Turk, 1990; Kellner, 1985). These concerns are likely to continue into adulthood, particularly when persons experience symptoms that they cannot easily explain or understand. For instance, Eifert and Forsyth (1996) found that exposure to heart disease in parents may make observing children more vulnerable to developing fears of heart disease. Furthermore, the death or terminal illness of a relative or friend has also been shown to precipitate hypochondriacal fears and behaviors with developing symptoms often resembling the deceased relative's symptoms (Eifert, Hodson, Tracey, Seville, & Gunawardane, 1996; Parkes, 1972). Finally, sexual trauma has also been linked to the development of excessive health anxiety. Barsky, Brener, Coeytaux, and Cleary (1995) found that hypochondriacal hospital outpatients recalled more childhood trauma (parental upheaval, sexual trauma, and beatings) before the age of 17 years than a patient control group.

6.24.3.3 Pain Disorder

The exact prevalence of pain disorder is unknown, but it appears to be relatively common and may occur in adults of all ages. For example, pain is the most common complaint of individuals presenting to a physician (Ford, 1995). The proliferation of special pain clinics could be seen as another indication of the high number of pain patients seeking professional help. According to *DSM-IV* (APA, 1994) estimates, in any given year, 10–15% of adults in the United States have some form of work disability due to back pain alone. A study of internal medicine private practice patients found that 13% of patients suffered from chronic pain (Margolis, Zimny, & Miller, 1984). Using a sample of HMO enrollees, von Korff, Dworkin, and LeResche (1990) observed that 45% of individuals reported persistent or recurrent pain with 8% reporting severe pain and 2.7% reporting severe pain that limited their normal activity.

Although research aimed at determining the prototypical pain patient continues (cf. Gamsa, 1994), there is no compelling evidence that shows pain patients as a homogeneous and separate group from individuals without pain (Love & Peck, 1987; Turk & Flor, 1984; Turk & Salovey, 1984). Consequently, in the place of uniformity, many researchers have begun to look for subgroups with different psychological profiles (Armentrout, Moore, Parker, Hewett, & Feltz, 1982; Jensen, Turner, Romano, & Karoly, 1991). Apart from a high frequency of prior physical abuse, sexual abuse, and other trauma, occupational factors play an important role. For instance, overuse of a body part has been shown to lead to specific pain syndromes (Newmark & Hochberg, 1987; Schuldt, Eckholm, Harms-Ringdahl, Aborelius, & Nemeth, 1987).

6.24.3.4 Conversion Disorder

Isolated conversion symptoms are believed to be fairly common and symptoms often disappear after a relatively brief period. In contrast, the diagnosis of conversion disorder is rare and difficult to establish with estimates ranging between 0.001% and 0.3% (APA, 1994). One reason is that symptoms seemingly indicative of conversion disorder are later discovered to be linked to a gradually developing, physical (neurological) disease such as a brain tumor or multiple sclerosis (cf. Fishbain & Goldberg, 1991). Although this condition may occur at any age, onset is typically in late childhood or early adulthood. Onset is often sudden and in response to conflicts or stressful situations such as unresolved grief and sexual trauma. Emotional stress was found to be present before the onset of conversion in 87% of the 53 outpatients in an Indian study by Sharma and Chaturvedi (1995). According to *DSM-IV* (APA, 1994), conversion is five times more common in women than in men and more common in persons of lower socioeconomic status with limited medical or psychological knowledge.

6.24.3.5 Body Dysmorphic Disorder

The prevalence of BDD is largely unknown, but preoccupation with body image and dissatisfaction with some aspect of one's appearance are believed to be widespread in the general population. In fact, according to *DSM-IV* (APA, 1994), BDD may be more common than was previously thought and under-recognized in settings where cosmetic procedures are performed. In terms of a diagnosable condition, Rosen (1995) found that about 1% of a community sample met criteria for BDD with virtually no gender differences in prevalence.

Onset of BDD may be gradual or sudden, and its course is generally continuous and chronic, though fluctuating in intensity. At present, we have only fragmentary and anecdotal reports on etiological factors, and no prospective longitudinal study of BDD has been undertaken (Rosen, 1995). Nevertheless, BDD is thought to start in adolescence when preoccupation with physical appearance is very common. Sociocultural factors influencing people's attitudes towards and dissatisfaction with their bodies seem to play a role in determining the extent to which a real or imagined physical abnormality becomes a cause for concern and preoccupation. Perfectionistic features seem to be related to BDD (Frost, Williams, & Jenter, 1995).

6.24.4 PROBLEMS OF DIFFERENTIAL DIAGNOSIS AND COMORBIDITY

The diagnostic validity of the somatoform disorders has been questioned repeatedly and for good reason. Even the authors of the *DSM-IV* concede that the grouping of these disorders in a single section is based on clinical utility rather than on assumptions regarding shared etiology or mechanisms (APA, 1994). Murphy (1990) indicated that the disorders are not qualitatively distinct but rather merge into each other, making distinctions between individual somatoform disorders hard to define. As an example, pain may occur in any of the somatoform disorders. Ambiguous normative criteria such as "the person's concern is markedly excessive," "grossly in excess of what would be expected," and "slight physical abnormality" create further problems. The distinction between BDD and normal concerns about appearances are as difficult to make as the distinction between BDD and delusional disorder (somatic subtype). The relation and distinction between somatoform disorders and personality disorders have also recently been questioned (Bass & Murphy, 1995; Tyrer, 1995). On the other hand, *DSM* criteria are very selective and will only allow diagnosing patients with a particular symptom profile (Fink, 1996).

In view of all these problems, Fava (1992b) suggests that the concept of abnormal illness behavior should probably replace the somatoform disorders rubric because it provides a more useful conceptual framework for disorders that would otherwise be scattered and unrelated in the *DSM* or that would not find a place at all. Fink (1996) makes a similar point arguing that researchers have been preoccupied with physical symptoms in search for reliable and valid diagnostic criteria for somatization, but have neglected the psychopathology, behavior, and other aspects of the problem. A simplistic reliance upon physical symptoms is questionable for diagnosing somatoform disorders and useless for the purpose of designing treatments. Symptom-focused diagnoses may be artifacts biased by patient suggestibility and a clinician's preoccupation for some disease. Rather than the presence of a specific set of physical symptoms, it is the way a patient interprets, experiences, and responds to a symptom that constitutes psychopathology (Fink). In particular, a somatic attributional style (Robbins & Kirmayer, 1991) may contribute to the translation of personal and social problems into physical symptoms, and prompt patients to present with somatic distress and request somatic treatments.

In sum, the diagnostic validity of somatoform disorders in relation to each other as well as to other clinical syndromes is problematic. The problem even extends to the distinction between the somatoform disorders and general medical conditions. Using a categorical classification system, diagnostic distinctions cannot be made to a satisfactory degree. What is needed is a multidimensional classification system incorporating psychological, somatic, and social dimensions (cf. Mayou, Bass, & Sharpe, 1995). Nonetheless, in the present diagnostic criteria the origin of somatic complaints assumes an important role for differential diagnosis. *DSM-IV* distinguishes symptoms as part of a "real disease" (general medical condition) from symptoms that are under a patient's voluntary control or intentionally produced. We will therefore briefly discuss attempts to distinguish somatoform disorders from each other and from a general medical condition. The interesting relation between somatoform and anxiety disorders will then be discussed to show how a classification system that focuses on the functions of maladaptive health-related behaviors, and defines problems in a dimensional rather than categorical manner, may be more beneficial for clinicians, researchers, and patients alike. Such an approach avoids the pitfalls of putting people into distinct categories and provides information that is directly useful for the design of clinical interventions.

6.24.4.1 Somatoform Disorders vs. General Medical Conditions

The presence of a general medical condition that could account for the presenting symptoms must be carefully examined and considered in every case where physical problems are the focus of a patient's complaints. Symptoms such as pain or fatigue may be related to a wide array of problems ranging from normal sensations to fatal diseases. Health professionals are very much aware of the danger of misdiagnosing a somatoform disorder and of missing the presence of actual physical problems—particularly diseases with a slow or diffuse onset such as multiple sclerosis, brain tumors, or systemic lupus. Some patients diagnosed with somatoform disorders or "functional problems" are ultimately diagnosed with a demonstrable medical condition and must be regarded as initial false-positives (for a particularly poignant example, see Fishbain & Goldberg, 1991). On the other hand, advances in medical diagnostic procedures (e.g., PET and magnetic resonance imaging (MRI) scans) have resulted in more accurate diagnostic decisions and reduced the number of false-positive diagnoses of somatization disorder (Kent, Tomasson, & Coryell, 1995).

Any diagnosis of somatoform disorders should be made with caution and only after careful physical examination. This recommendation is also supported by the fact that it is occasionally difficult for physicians and psychologists alike to determine the "true" nature of somatic complaints. In a study of 200 successive patients undergoing cardiac catheterization, it was found that neither standard medical tests before catheterization nor standard psychological questionnaires alone were able to discriminate reliably chest pain patients with coronary artery disease from patients without heart disease. Diagnostic accuracy only improved when cardiac catheterization was considered along with results from psychological assessments (Eifert, Edwards, Thompson, Haddad, & Frazer, 1997).

In addition, there is at times a reciprocal relationship between health anxiety and somatic symptoms. For instance, Salkovskis (1996) describes how the very safety-seeking behaviors designed to reduce anxiety frequently increase the symptoms that are the focus of anxiety. He cites the example of patients who palpate or rub lumps until they swell and cause pain. Several authors also point to potential pathophysiological mechanisms that may underlie unexplained physical symptoms. Sharpe and Bass (1992) describe various pathophysiological mechanisms in abdominal pain, chest pain, chronic fatigue, breathlessness, and irritable bowel syndrome that can be detected by routine or advanced medical evaluation. For instance, symptoms can be due to excessive physiological activity (e.g., smooth-muscle contraction, striated-muscle contraction, changes in endocrine secretion and in blood flow) that may be accentuated by stress and intense emotions. Both in fairness to the patient, and to design the most appropriate intervention, we should be careful in the use of such terms as unexplained physical symptoms and functional defects. Terms such as "nonorganic ..." should be avoided completely.

6.24.4.2 Differentiating Somatoform Disorders

Differentiating the various somatoform disorders using *DSM-IV* criteria is difficult because criteria are vague and there is considerable overlap of symptoms, Nonetheless, several distinctions can be made and are briefly discussed below.

6.24.4.2.1 Somatization disorder vs. hypochondriasis

In somatization disorder the patient's attention is directed at the somatic symptoms themselves, whereas in hypochondriasis symptoms are generally less elaborate and the patient is concerned about a possible underlying illness (Murphy, 1990). As a result, hypochondriacal patients experience higher levels of anxiety, whereas patients with somatization disorder experience lower levels of anxiety or no anxiety at all.

6.24.4.2.2 Somatization disorder vs. pain disorder

Although pain symptoms are included in the diagnostic criteria for somatization disorder, symptoms other than pain must be present before a diagnosis can be assigned. In contrast, pain is the predominant (and frequently exclusive) focus of the clinical presentation of a person with pain disorder.

6.24.4.2.3 Somatization disorder vs. conversion disorder

A key difference is the sheer number of symptoms. A person with somatization disorder, by definition, must report at least eight symptoms of four different types, whereas a person with conversion disorder typically manifests only one symptom.

6.24.4.2.4 Hypochondriasis vs. body dysmorphic disorder

Although hypochondriasis and BDD share some features (e.g., bodily preoccupation, repetitious body checking, reassurance seeking, and medical consultations), individuals engage in these behaviors for different reasons. Persons with hypochondriasis are afraid of serious illness, whereas persons with BDD are concerned about the physical appearance of their body.

6.24.4.3 Relation Between Somatoform Disorders and Depression

There is a high comorbidity of all somatoform disorders with depression and to a lesser extent with dysthymic disorder. For instance, in a prospective study of 30 psychosomatic clinic inpatients, Rief, Hiller, Geissner, and Fichter (1995) found high lifetime comorbidity between various somatoform disorders and both depression (86%) and anxiety (43%). Remission rates were higher when somatoform disorders were not accompanied by other psychiatric disorders (e.g., anxiety, depression, or addiction) than in cases with comorbidity. In addition, the presence of somatoform disorders increased the risk for other psychiatric conditions (Rief et al.). Several studies found a striking relation between pain and major depression or dysthymic disorder but the specifics of that relation remain elusive (Chaturvedi & Michael, 1986; France, Krishnan, Houpt, & Maltbie, 1984; Turk, Okifuji, & Scharff, 1995).

6.24.4.4 From Differential Diagnosis to Dimensional Classification

Despite considerable degree of overlap in the symptoms of persons with hypochondriasis, disease phobia, and panic disorder, also reflected in reports of high comorbidity rates (Warwick & Salkovskis, 1990), our understanding of the relation between somatoform and anxiety disorders is poor (Barlow, 1988; Forsyth & Eifert, 1998). In the following discussion, we will adopt a dimensional approach to understanding illness-related concerns that might be useful for both assessment and treatment purposes. Delineating points of overlap and differences is also an important first step toward designing more thorough empirical investigations of the subject.

6.24.4.4.1 Hypochondriasis and disease fear: a dimensional perspective

According to Pilowsky's (1967) classic study, hypochondriasis has three dimensions: disease phobia, disease conviction, and bodily preoccupation. These dimensions are still the basis for the current *DSM* criteria for hypochondriasis. Unfortunately, the *DSM* definition fails to distinguish hypochondriasis as a fear of disease from a conviction of having a disease. This failure could, at least in part, account for the ambiguity in the use of the diagnostic label of hypochondriasis (cf. Eifert, 1992). Fear of having a physical disease has quite different theoretical, diagnostic, and therapeutic implications from a conviction of being seriously ill, regardless of whether the conviction is accompanied by fear or not.

A clear separation of the different dimensions of hypochondriasis may contribute to a better understanding of such problems and result in more appropriate treatments. We view hypochondriasis as a syndrome where persons present with problems that fall on a continuum along four dimensions.

(i) *Preoccupation with the body and its functioning*. Excessive awareness of and interest in bodily sensations and functioning, with or without physical complaints, constitutes what Starcevic (1988) aptly called "the hypochondriacal core." Such bodily preoccupation, especially when coupled with somatic complaints, may produce a state of somatic uncertainty and form the basis for the other three dimensions of the disorder.

(ii) *Disease suspicion or conviction*. The person has the suspicion or is convinced of having a serious physical disease; suspicion and conviction are on a continuum of strength, and in rare cases the conviction may reach delusional intensity.

(iii) *Disease fear*. The person fears having a serious physical disease.

(iv) *Safety-seeking behaviors*. The function of behavior such as repeated requests for medical examinations and tests, bodily checking, verbal complaints, and seeking reassurance is to reduce worry and anxiety over physical illness (Eifert, 1992; Salkovskis, 1996; Warwick & Salkovskis, 1990).

Although a person could score highly on any one or all four dimensions, bodily preoccupation and disease suspicion/conviction are most central to hypochondriasis. As indicated, disease suspicion/conviction may or may not be accompanied by fear of the suspected disease. Clinically, this feature is most apparent in patient's resistance to medical reassurance. For instance, Fava and Grandi (1991) indicated that patients may continue to be extremely worried about their health even though their fear of suffering from a specific disease may be eliminated by a satisfactory medical examination. In other words, a patient's morbid preoccupation with disease may continue in the absence of any specific disease phobia maintaining and prolonging a person's suffering and keeping them from consulting a psychologist.

Disease phobia is a persistent unfounded fear of suffering from or contracting a disease (Bianchi, 1973, Eifert, 1992) and may occur in the absence of a conviction or suspicion of having a disease. For example, a person may be afraid of having a heart attack or contracting cancer without being convinced or suspicious of having heart disease or cancer. According to

Fava and Grandi (1991), there are two main clinical characteristics of disease phobia that distinguish it from more general hypochondriasis. One feature is the specificity and temporal stability of symptoms. For instance, it was found that most cardiophobic patients were only afraid of having a heart disease but not of other diseases (Eifert et al., 1996). Even over time, these patients are unlikely to develop fears focused on other organ systems. This specificity led Marks (1987) to view disease phobias as a subtype of hypochondriasis focused on a specific illness. The second feature of disease fears is that they manifest themselves in cue-controlled attacks of shorter durations. For example, a cardiophobic person's fear of having a heart attack becomes salient and particularly strong only in response to the acute experience of chest pain and heart palpitations (Eifert 1992; Eifert et al. 1996). This feature of disease phobias is in contrast to the morbid preoccupations of persons with primary hypochondriasis who constantly worry and ruminate about their health.

6.24.4.4.2 Relation between hypochondriasis and anxiety disorders

A discussion of the dimensions of excessive health anxiety raises several questions regarding the relation of this type of anxiety to other anxiety-related problems, which will be discussed next.

(i) Comorbidity of disease phobias and hypochondriasis with panic disorder

Warwick and Salkovskis (1990) report that 59% of hypochondriacal patients seen in their clinic also met *DSM-III-R* criteria for panic disorder. Other studies (Beitman et al., 1987; Eifert et al., 1996) found that between 25% and 50% of cardiology persons with heart-focused anxiety also suffer from panic disorder. Panic patients with agoraphobia have been shown to score as high as hypochondriacal patients in areas such as somatic preoccupation, disease phobia, and illness conviction (Noyes, Reich, Clancy, & O'Gorman, 1986). Following treatment, significant reductions in hypochondriasis scores occurred in those panic patients who improved with treatment. Fava, Kellner, Zielezny, and Gurand (1988) also found that before treatment, patients with panic disorder and agoraphobia scored significantly higher on several illness fear and behavior measures than a group of healthy controls. Following the successful treatment of panic, however, these group differences had disappeared.

(ii) Modulation and timing of anxiety

Fava and Grandi (1991) describe the relation of disease phobia to hypochondriasis as similar to that of panic disorder to generalized anxiety. This perspective is based on a difference in the modulation and timing of anxiety in disease phobia and panic compared to hypochondriasis and generalized anxiety disorder. As in panic disorder, which is characterized by a sudden surge of aversive sensations and fear that tends to subside within 30–45 minutes, persons with disease phobias tend to experience anxiety attacks of a relatively short duration when exposed to feared cues. In contrast, persons with primary hypochondriasis experience more constant and chronic levels of worry about body sensations and illness, similar to the chronic levels of worry found in generalized anxiety disorder—although in both disorders worries and fears may occasionally escalate into a "crescendo of panic" (Salkovskis & Clarke, 1993; Warwick & Salkovskis, 1990). In addition, these authors point out that persons with hypochondriasis typically judge the danger they fear to occur at some distant time in the future, whereas panic patients fear an imminent catastrophe.

(iii) Misinterpreted symptoms and focus of fear

In panic disorder as well as disease phobia, individuals misinterpret autonomic symptoms that are involved in acute anxiety attacks. Persons with disease fears, however, tend to report fewer and less severe panic symptoms (Beck, Berisford, Taegtmeyer, & Bennett, 1990) and their fear is more specific and focused on one organ or organ system (Eifert, 1992; Eifert et al., 1996). Such individuals do not fear the symptoms themselves but what they seem to indicate (i.e., serious disease). On the other hand, persons with panic disorder tend to fear generalized uncontrollable aversive autonomic body sensations, that is, they fear the actual bodily sensations (see Chambless & Graceley, 1989). They also have fewer and less pronounced illness behaviors and beliefs than persons with either disease phobias or hypochondriasis (Eifert, Seville, Antony, Brown, & Barlow, 1992). Hypochondriacal patients are more likely to misinterpret symptoms which are not subject to direct amplification such as lumps, skin rashes, and blemishes (Salkovskis & Clark, 1993), and their fear is focused on these health/disease issues.

(iv) Behavioral avoidance

Individuals with disease phobia typically cannot avoid the stimuli they fear (pain and palpitations) and therefore attempt to avoid activities which they believe bring on the dreaded physical symptoms (Eifert et al., 1996). They may also engage in a number of behaviors that are designed to protect their body or a specific organ that they are afraid of damaging. Warwick and Salkovskis (1990) have identified another important difference between hypochondriacal and panic patients in patterns of escape and avoidance related to illness. Hypochondriacal patients have more time to prevent the anticipated disaster by seeking medical attention because the anticipated harm is perceived as much less imminent. In contrast, panic patients are likely to be overwhelmed by the sudden surge of arousal and fear during a panic attack and perceive leaving the situation and their future avoidance as their only option (Salkovskis, 1996).

In conclusion, we believe that a dimensional classification system could help overcome the pitfalls of pigeon-holing individuals into existing diagnostic categories. Identifying dimensions that allow a classification of illness behavior based on the function that such behavior serves, rather than its topography, might lead to a better understanding and improved treatments of persons with somatoform problems. Apart from the number and type of physical complaints, some of these dimensions include the presence and extent of preoccupation with body and health, symptom misinterpretation, disease suspicion or conviction, disease fear, safety-seeking approach and avoidance behavior, focus and modulation of worry and fear, and pathophysiological processes. Barlow (1988) noted that classification of any disorder, whether dimensional or categorical, should reliably describe subgroups of symptoms or behaviors that are readily identifiable by independent observers on the basis of operational definitions. There should also be some clinical usefulness or value in identifying these subgroups or dimensions such as predicting specific response to treatment, course of the disorder, and tailoring treatment (Eifert, 1996; Eifert, Evans, & McKendrick, 1990). We hope that a dimensional analysis of somatoform disorders will move us closer toward reaching that goal.

6.24.5 CONTEMPORARY THEORETICAL PERSPECTIVES

Although some theoretical models have been advanced in relation to the process of somatization (cf. Kellner, 1991), somatoform disorders are understudied and poorly understood (Barlow, 1988; Hitchcock & Mathews, 1992). As cognitive-behavioral concepts predominate contemporary perspectives on somatoform disorders, we will briefly discuss some of these concepts. In terms of specific disorders, we will concentrate on hypochondriasis and pain since much of recent research has focused on these problems.

6.24.5.1 Dysfunctional Processes in the Somatoform Disorders

6.24.5.1.1 Abnormal illness behavior

Pilowsky (1993, p. 62) defined abnormal illness behavior as the "persistence of a maladaptive mode of experiencing, perceiving, evaluating, and responding to one's own health status, despite the fact that a doctor has provided a lucid and accurate appraisal of the situation and management to be followed (if any), with opportunities for discussion, negotiation and clarification, based on adequate assessment of all relevant biological, psychological, social and cultural factors."

Abnormal illness behavior is not a diagnosis as such, but refers to the disagreement between the doctor and patient about the sick role to which the patient feels entitled (Pilowsky, 1993). Sharpe, Mayou, and Bass (1995) argued that the concept of abnormal illness behavior is not only valuable for understanding patients with functional somatic symptoms but for understanding the behavioral aspects of all illness.

6.24.5.1.2 Perception and attention

Cognitive processing of medical information as well as attention to and attribution of symptoms are believed to be central features in all somatoform disorders. Pennebaker (1982) conducted various studies on the role of attention in symptom perception and concluded that the more salient a somatic symptom (and the lower the amount of external distraction), the more likely and more intense the original symptom is perceived. An implication of this finding is that deficient external stimulation may pave the way for an increased attention to somatic symptoms Pennebaker also demonstrated that simply directing a person's attention to bodily sensations increases reports of symptoms. Selective attention to internal stimuli could thus augment somatic concerns.

Although attentional bias has been well documented in individuals with anxiety disorders showing that attention is biased towards threat-related stimuli (cf. MacLeod & Mathews, 1991), empirical evidence for this type of bias in somatoform disorders is mixed. Most studies have examined hypochondriacal patients

(Barsky et al., 1995; Tyrer, Lee, & Alexander, 1980) or normal subjects with high scores on a hypochondriasis scale (Hanback & Revelle, 1978; Hitchcock & Mathews, 1992). In a series of experiments, Hitchcock and Mathews were able to demonstrate enhanced registration and availability of illness information. They also found that hypochondriacal subjects interpreted ambiguous information as threatening, although this bias was applied equally to illness and social ambiguity. By contrast, subjects did not show a bias for making rapid automatic inferences about bodily sensations implying illness. Hitchcock and Mathews rightly caution, however, that it is unclear whether enhanced perceptual sensitivity to illness cues is a cause or consequence of hypochondriacal anxiety.

A particular perceptual disturbance in body image has been demonstrated in persons with BDD although the specific processes involved have not been studied in detail. Rosen (1995) proposed an integrative formulation of BDD linking processes such as thinking about the presumed abnormality in appearance with behavioral features such as avoidance of social situations, camouflaging, checking, and undertaking beauty remedies. Preoccupation with the imagined defect is especially salient in social situations and may take on delusional or obsessional qualities.

6.24.5.1.3 Deficits in emotion processing (alexithymia)

Alexithymia literally means “no words for feelings.” The concept refers to a hypothesized communicative function of somatic symptoms. Its key features are a relative constriction of emotional functioning, poverty of fantasy life, and inability to find appropriate words to describe one’s emotions (Taylor, Bagby, & Parker, 1991). Alexithymic individuals seem to be vulnerable to mounting tension from undifferentiated states of unpleasant emotional arousal. This vulnerability is presumably caused by a disturbance in the processing of emotional awareness that is believed to interfere with the subject’s ability to experience and express emotions. For example, in a sample of normal volunteers, Vingerhoets, Van Heck, Grim, and Bermond (1995) found strong negative correlations between alexithymia and the expression of emotions, daydreams, and fantasies, and planful and rational actions. Bach and Bach (1995) found high alexithymia scores to be predictive of persistent somatization. More recently, the concept of alexithymia has also been discussed in relation to a variety of psychological traumata (Salminen, Saarijävi, & Äärelä, 1995).

6.24.5.2 Disorder-specific Theoretical Models

As indicated, cognitive-behavioral perspectives and research have been helpful in providing a fledgling basis for a better understanding and treatment of persons with somatic problems. Below we will focus on cognitive-behavioral perspectives of hypochondriasis and pain as most recent research has dealt with these problems.

6.24.5.2.1 Hypochondriasis

Behavioral theories of excessive health anxiety stipulate that internal cues which have been associated with threat and bodily harm (unconditioned stimuli) can serve as conditioned stimuli (Forsyth, Eifert, & Thompson, 1996; Miller, 1977). In the presence of these stimuli, individuals will begin to exhibit conditioned responses such as anxious behavior and physiological changes. Classically conditioned instances of hypochondriacal behavior can be maintained through operant reinforcement. Somatic complaints may lead to attention, sympathy, and escape from or avoidance of undesirable tasks or situations. Such consequences may reinforce symptom reporting and other hypochondriacal behavior (Kellner, 1985). If somatic complaints are ignored in the patient’s home, hypochondriacal behavior in the presence of family and friends may be extinguished, but it may also lead the patient to increase medical assistance-seeking in the hope of receiving attention from a physician instead.

These are examples of contingency-shaped behavior, but observers may also learn to exhibit hypochondriacal behaviors through the process of modeling (Craig, 1986) or rule-governed behavior (Hayes & Wilson, 1994). According to Kellner (1985), hypochondriacal behavior worsens as the conditioning process repeats.

As can be seen in Figure 1, cognitive theories of health anxiety emphasize the role of misinterpretations of innocuous bodily changes or of information provided by doctors, friends, or the media. Patients focus on essentially harmless physical sensations, which they consistently misrepresent and misinterpret as indications of physical illness (Barsky & Klerman, 1983; Warwick & Salkovskis, 1990). As a result, symptoms and medical information tend to be perceived as more dangerous than they really are and a particular illness is believed to be more probable than it really is (Salkovskis, 1996). According to Salkovskis, the onset and maintenance of health anxiety involves the complex interaction of several factors. Previous experiences and perception of illness in self and

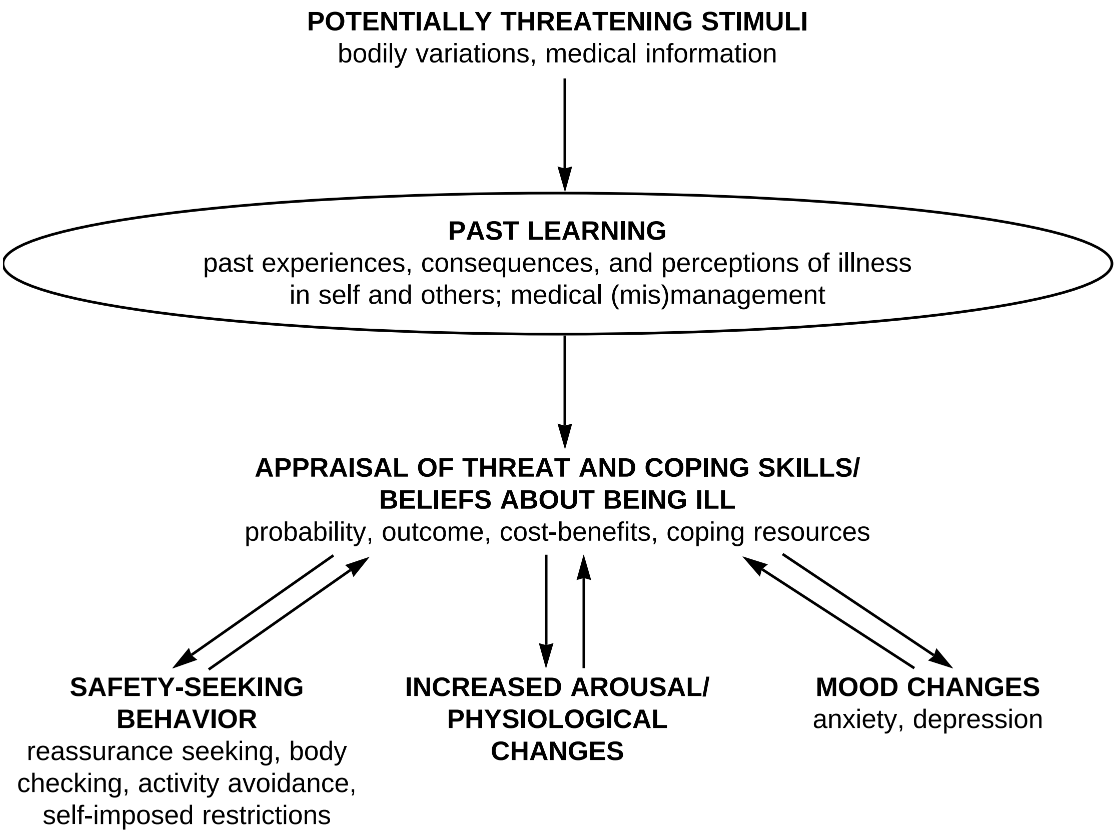
POTENTIALLY THREATENING STIMULI
bodily variations, medical information
PAST LEARNING
past experiences, consequences, and perceptions of illness in self and others; medical (mis)management
APPRAISAL OF THREAT AND COPING SKILLS/ BELIEFS ABOUT BEING ILL
probability, outcome, cost-benefits, coping resources
SAFETY-SEEKING BEHAVIOR
reassurance seeking, body checking, activity avoidance, self-imposed restrictions
INCREASED AROUSAL/ PHYSIOLOGICAL CHANGES
MOOD CHANGES
anxiety, depression

Figure 1 Cognitive-behavioral model of maladaptive illness-related behavior (adapted from Warwick & Salkovskis, 1990, and Salkovskis, 1996; cf. Eifert, 1992).

others ('whenever I had physical symptoms as a child I was taken to the doctor") are often the root of future hypochondriacal behavior and may result in the formation of dysfunctional beliefs ('bodily symptoms are always a sign that something is wrong"). As a result, harmless incidents or physical symptoms activate these negative assumptions and other catastrophic thoughts ('this is going to get worse"), which in turn lead to behavior changes such as increased safety-seeking and avoidance as well as heightened physiological arousal, anxiety, and depression. Repeated cycles of this kind may lead to an exacerbation of hypochondriacal symptoms and perceptions of oneself as a sick and incapable individual (Salkovskis, 1996).

6.24.5.2.2 Pain

According to the behavioral theory of pain developed by Fordyce (1976), pain behavior is often acquired through respondent conditioning ("sensory pain") and maintained through operant conditioning ("psychological pain"). Initially, pain is considered to be an overt reflexive response to an antecedent noxious stimulus such as an injury. Following onset, future pain behavior is subject to selection by consequences. Verbal reports and nonverbal expressions of pain are maintained because of the associated delivery of positive and negative reinforcers. As indicated, pain behavior may be positively reinforced through tangible benefits, attention, or stimulation (i.e., concern or sympathy). Pain behavior may also be negatively reinforced when it results in the avoidance of or escape from undesirable activities or stimulation such as work or unwanted sexual encounters. In the absence of direct exposure to the environmental contingencies, these behaviors may be acquired or taught through modeling (Craig, 1986). Respondent and operant conditioning may interact in the onset and maintenance of pain behavior (Rachlin, 1985). For instance, the avoidance of physical activity can lead to muscle fibers shortening and losing elasticity. This can be quite painful and, as a consequence, persons tend to avoid or escape from physical activity. Such avoidance and escape are negatively reinforced when they lead to a short-term decrease of pain, but avoidance and escape may also result in long-term cost by exacerbating the physical condition of the individual and lead to a vicious circle (Flor et al., 1990). A number of studies have shown that pain behavior is susceptible to social reinforcement. For example, Cairns and Pasino (1977) reported that exercise behavior increased when praise was delivered contingently and decreased soon after praise delivery was discontinued. Moreover, pain patients who perceived their wives as solicitous exhibited more pain behavior in the presence of their wives than patients who perceived their wives as nonsolicitous (Flor, Kerns, & Turk, 1987).

Although pain is a subjective phenomenon and difficult to quantify, behavioral researchers have attempted to quantify pain behavior. In a choice situation, the relative frequency of responding (either the number of responses or the amount of time responding) allocated to a particular alternative closely approximates the relative frequency of obtained reinforcement for that alternative. This relation is called the matching law (Baum, 1974; Herrnstein, 1970). Applied to pain, the matching law suggests that the frequency of an individual's adaptive behavior (e.g., pain coping without complaining), compared to maladaptive pain behavior (e.g., verbal complaints, facial contortions), will match the relative frequency of reinforcement obtained for each type of behavior. In other words, if maladaptive pain behavior is reinforced more than adaptive behavior, the likelihood of future maladaptive behavior will increase, whereas adaptive pain coping behavior becomes less likely. Conversely, if a patient's maladaptive behavior is reinforced less than adaptive pain behavior, preference will shift to adaptive pain behavior. The matching law is clinically useful because it allows us to quantify how much reinforcement will be necessary to sustain a particular level of a given behavior. Fernandez and McDowell (1995) found the matching law to provide accurate descriptions of pain behavior. Although there are no other studies applying the matching law to pain, the clinical utility of the matching law has been demonstrated in the study of self-injurious behavior (McDowell, 1981, 1982), social behavior (Conger & Killeen, 1974), and disruptive behavior of mentally retarded children (Martens & Houk, 1989). The matching law could provide a promising, more objective approach to quantifying pain behavior.

Cognitive-behavioral perspectives emphasize that maladaptive pain behavior is due to negative or unrealistically high expectations about experiencing pain as well as low levels of perceived competence to deal with pain (Schermelleh-Engel, Eifert, Moosbrugger, & Frank, 1997). Apart from the physical characteristics of the pain stimulus, or the amount of tissue damage, differences in coping style are particularly important for the perception of pain and associated mood states (Turk et al., 1995). Our own research specifically supports the importance of perceived competence for coping with chronic pain, indicating that perceived competence determines how intense persons experience pain, how much they are disturbed by it, and whether they engage in adaptive or maladaptive coping behavior (Schermelleh-Engel et al.). Although the coping process is not systematically affected by the characteristics of the pain stimulus, a person's type of coping behavior affects the impact of the pain stimulus (cf. Asmundson & Norton, 1995). Using path analytic structural equation modeling, we found that coping behavior is directly influenced by a person's perceived level of competence and only indirectly affected by pain intensity and pain-related anxiety and depression. Decreasing competence appears to lead to increasing pain intensity and increasing pain emotions, which, in turn, increase maladaptive behavior. Conversely, individuals who trust in their abilities to cope adequately with pain are more likely to engage in adaptive behavior, irrespective of how anxious they are; such persons also do not suffer as much as individuals with low perceived competence.

6.24.5.3 Future Challenge: Integrative Biobehavioral Theories

Throughout this chapter we have alluded to the significant degree of overlap between somatization, hypochondriasis, and pain. Comprehensive theoretical accounts are just beginning to emerge for these problems, but they are impeded by a relative lack of agreement on what precisely somatization refers to. For instance, do persons diagnosed with "unexplained physical symptoms" (e.g., Speckens et al., 1996) suffer from somatization disorder, hypochondriasis, or are they a mix between the two? Moreover, presenting with unexplained physical symptoms does not imply that there is no medical or other explanation for the symptoms. In some (although certainly not all) cases, an explanation might be found if more sophisticated tests such as MRI or PET scans were conducted. Moreover, there is a significant "gray area" that medical colleagues often refer to when they speak of persons who do have some degree of demonstrable pathophysiology. Yet patient symptom reports, or their response to symptoms, seem to be "exaggerated" in view of the degree of actual tissue damage or other physical changes. Since neither physicians nor psychologists can adequately pinpoint or understand the symptoms and behavior of such patients from their respective perspective alone, these patients should be approached from a multidisciplinary perspective. Unfortunately, theories that attempt to integrate biological/medical with psychological knowledge are rare. A notable exception is the psychobiological account of chronic pain, which we will briefly summarize as an example of what more adequate theories for persons with somatic problems might look like in the future.

The early physiological views of pain regarded pain simply as a sensation resulting from patterns of stimulation from free nerve endings; pain intensity was believed to be directly proportional to the amount of peripheral nociceptive input related to tissue damage (Dallenbach, 1939). Proposed by Melzack and Wall (1965), the gate control theory (GCT) was one of the first psychophysiological theories to challenge this assertion. According to the GCT, pain is due to the complex interplay of receptors, afferent and efferent processes, and spinal and supraspinal processes. More recent theories (cf. Cailliet, 1993) have expanded the basic concepts of the GCT to relate current information on physiological processes (e.g., the role of acid metabolites and opiate receptors) to psychological factors. An integrative psychobiological theory of chronic pain proposed by Flor et al. (1990) relates concepts and findings from areas such operant, respondent, observational, and cognitive learning and relates them to biological concepts and physiological findings.

Flor et al. (1990) propose that pain is a complex response that comprises subjective-psychological, motor-behavioral, and physiological-organic components. Stress may precipitate and facilitate a particular pain disorder to which the individual is already predisposed on an unlearned or learned basis. The learned aspect of this predisposition consists of a reduced threshold for nociceptive activation due to previous trauma or social learning experiences resulting in a physiological response stereotypy of a specific body system or group of muscles. Stress and pain episodes may trigger a host of autonomic and muscular reactions, particularly sympathetic activation and elevated muscle tension levels. If stress or pain-related muscular contractions occur repeatedly, a number of muscular and sympathetic reflexes lead to increases in muscle tension and to sympathetically mediated vasoconstriction. If the muscular contractions are of sufficient intensity, frequency, and duration, there will not be sufficient blood and oxygen in the affected muscles and algogenic (pain-inducing) substances will be released. The ensuing pain experience increases muscular and sympathetic hyperactivity and may thus lead to a vicious circle. As these processes also increase the sensitivity of chemo-nociceptors, the likelihood of future pain also increases. This feedback circle, in particular, nicely illustrates the interplay between physiological and psychological processes and helps us understand how both physiological changes and the person's response to such changes may interact to create the problems that persons with chronic pain struggle with.

6.24.6 TREATMENT OF SOMATOFORM DISORDERS

Psychologically distressed medical patients who present with unexplained somatic symptoms are high users of care and their doctors regard them as frustrating and difficult to treat (Mayou & Sharpe, 1995). There is often a mismatch between the expectations of these patients and their doctors' abilities and communication skills. For instance, Eifert (1992) pointed out that diagnoses such as functional heart problem, nervous heart, atypical chest pain, and pseudoangina can be easily misinterpreted by a cardiophobic patient who is determined to believe that some significant cardiac disease is being described. As a result, healthcare providers often feel frustrated and emotionally drained because these patients are obviously in need of psychological support but resent being referred to a psychologist or psychiatrist. Owing to the overlapping psychological processes involved in the various somatoform disorders, we will first outline treatment recommendations that are applicable to most individuals with any of the somatoform disorders. This discussion will be followed by some specific recommendations and outcome data for individual disorders.

Patients often perceive the use of diagnostic labels such as hypochondriasis as an insult because these labels are seen to imply that patient problems are not real and are "just in their head." Accepting and understanding the symptoms rather than refuting or arguing with the patient is therefore the most important condition for engaging the patient in a therapeutic working relationship (cf. Bass & Benjamin, 1993). In the engagement stage of treatment, patients are helped to see that there may be an alternative explanation for the difficulties they are experiencing (Salkovskis, 1996). For instance, although chest pain can be due to coronary artery disease, it can also be caused by hyperventilation-induced chest wall muscle tension (Eifert, 1992). The general treatment strategy is to test such alternative nonmedical or benign medical explanations for symptoms and to conduct therapy in the context of an experiment that provides an opportunity for testing alternative hypotheses (Salkovskis, 1996).

Rather than merely telling patients there is no "organic" reason for their chest pain, an explanation of symptoms that overcomes the nonorganic–organic dualism provides the patient with a more acceptable rationale and reassurance. For instance, to provide a patient with a credible explanation of how anxiety and chest wall muscle tension can result in chest pain, patients may be given a chest-focused relaxation with EMG feedback that literally shows them how they can change their chest tension levels (Eifert, 1996). Salkovskis and Warwick (1986) found that reassurance which only informs the patients that there is nothing organically wrong with their body will actually increase future reassurance-seeking rather than decrease it. On the other hand, appropriate reassurance and feedback that provides the patient with new and alternative explanations is the key to successful treatment, and in some cases, may help prevent the development of chronic somatization problems in the first place. It is therefore important that medical professionals are trained to deal with patients in this manner. For example, Goldberg, Gask, and O'Dowd (1989) developed a training program for primary care physicians that included making patients feel understood by carefully listening to their complaints and symptoms, changing the agenda of the actions to be undertaken from a somato-medical focus on bodily symptoms to a psychological perspective focusing on emotional distress. Physicians were also taught how to help patients reattribute bodily symptoms by linking them to psychological factors.

Another important treatment strategy is withholding unnecessary medication and medical examinations (response prevention for safety-seeking behaviors). A symptom diary is used to get a prospective, rather than retrospective, view of the symptoms' course, intensity, and relation to life stresses. Rief (1996) suggests starting off with a symptom-oriented treatment and, in the second phase of therapy, shifting towards more general goals, such as occupational problem-solving, social skills training, and quality of life enhancement.

6.24.6.1 Somatization Disorder

Most articles refer to the management rather than to the treatment of patients with persistent somatic problems (e.g., Bass & Benjamin, 1993). Early diagnosis and the prevention of unnecessary medical and surgical investigation are deemed to be of primary importance. Most somatization patients have distinct expectations regarding treatment goals and procedures and try to persuade their doctors to follow their wishes for further medical investigations and treatments. Bass and Murphy (1990) state that treatment often involves long-term supportive psychotherapy and must be directed toward controlling the demands on medical care as well as the treatment of symptoms and social disability. These authors recommend the following five steps: (i) establish a diagnosis by collecting appropriate medical and psychosocial information; (ii) encourage a long-term supportive relationship with only one understanding primary care physician to prevent doctor-shopping and to coordinate all actions; (iii) see patients on regular appointments rather than on demand to prevent reinforcement of illness behavior; (iv) regard the patient's physical complaints as a form of communication rather than as evidence of disease; and (v) minimize the use of psychotropic drugs and/or analgesic medication. In general, adaptive behavior is encouraged and promoted, whereas sick role behavior is ignored as much as possible. A study by Smith, Monson, and Ray (1986) found that such a treatment program did not lead to improvements on any of the measures of mental, physical, or social health, but patient healthcare was less costly. In contrast, Krasner, Rost, Cohen, Anderson, and Smith (1995) demonstrated that eight sessions of brief group therapy improved physical and mental health at one year follow-up. Treatment focused on coping with the nature and consequences of the physical symptoms, general problem-solving, and helping patients take more control of their lives.

A randomized controlled trial examining a comprehensive cognitive-behavioral approach for medically unexplained physical symptoms was conducted by Speckens et al. (1996). This team compared a cognitive-behavioral intervention group of 39 general medical outpatients with a control group of 40 patients receiving optimized medical care. The 6–16 sessions of cognitive-behavioral therapy included (i) imaginary exposure and distraction techniques to break the vicious circles of cognitive avoidance and preoccupation; (ii) activity scheduling, exposure *in vivo*, and response prevention to decrease avoidance behavior; (iii) relaxation training, breathing exercises, and physical exercises; and (iv) problem-solving or social skills training to overcome any problems in interpersonal relationships. At both 6 and 12 month follow-ups, compared to the control group, the intervention group reported lower intensity and frequency of symptoms, reduced illness behavior, less sleep impairment, and fewer limitations in social and leisure activities.

6.24.6.2 Hypochondriasis

Traditionally, individuals with hypochondriasis were considered difficult to treat and their prognosis was regarded as poor. For instance, early studies assessing the effectiveness of a variety of psychological interventions available at that time (e.g., Kenyon, 1964) found 40% of patients with primary hypochondriasis to be unchanged or worse following treatment. Although patients with unexplained physical symptoms and/or health anxiety still pose a considerable challenge for therapists, cognitive-behavioral interventions have yielded encouraging results.

The first comprehensive cognitive-behavioral treatment formulations were provided by Warwick and Salkovskis (1990). Patients are instructed to self-monitor during hypochondriacal episodes, paying careful attention to environmental events, physical symptoms, associated cognitions, and their resulting hypochondriacal behavior. As indicated, their treatment is directed at evaluating alternative, nonthreatening explanations of body-related observations that patients misinterpret as signs of serious disease. Two possible explanations for the patient's problem are considered alongside each other rather than as mutually exclusive alternatives (Salkovskis, 1996). Patients are then asked to engage in a variety of behavioral experiments to test these new explanations and therapy proceeds as an evaluation of the relative merits of the alternative views. Salkovskis emphasizes that inappropriate safety-seeking prevents individuals from discovering that their fears are groundless. Hence, the key function of exposure exercises and response prevention is to allow patients to repeatedly experience feared bodily sensations to disconfirm fears. Thereby patients learn that the things they are afraid of do not actually happen or that consequences are not as bad as they expected.

A system of differential reinforcement may need to be added for individuals who gain sympathy or interest for physical complaints but may receive little attention otherwise. In those cases, family members and other individuals the patient has contact with (including the therapist) should reinforce patient initiation of conversation on any topic other than symptoms (healthy conversation). This reinforcement could consist of praise or increased attention. As the patient engages in more healthy conversation, the natural contingencies (e.g., reduced medical bills, more pleasant interactions with others) should begin to maintain behavior and the frequency and magnitude of artificial reinforcement can be slowly reduced.

Salkovskis and Warwick (1986) reported two successfully treated single cases of primary hypochondriasis using cognitive-behavioral methods. In another study, Visser and Bouman (1992) found that four of six hypochondriacal patients improved significantly following *in vivo* exposure, response suppression, and cognitive therapy. Interestingly, in this small study, exposure and response suppression (behavioral components) appeared to have accounted for more of the improvement than the cognitive component they used. The positive results of these series of single case studies were supported in a first controlled trial that showed cognitive-behavioral therapy to be superior to a wait-list control condition (Warwick, Clark, Cobb, & Salkovskis, 1996).

For patients whose main hypochondriacal concerns involve disease or illness phobia without significant conviction of having the illness, a more focused program of tension reduction (e.g., through relaxation), exposure to the feared stimuli, and prevention of checking and safety-seeking can be sufficient (Eifert, 1992, 1996). A representative example of this type of program was implemented by Warwick and Marks (1988). Their treatment employed exposure to feared stimuli, paradox (deliberate attempts to induce a panic attack), and response prevention such as banning reassurance-seeking and physician visits. There was a significant decrease in illness fear and increase in work and social adjustment for the group of 17 patients who were treated.

Overall, cognitive-behavioral treatments are very promising and prognosis for patients treated with such interventions seems more favorable than previously thought. More extensive and controlled studies are needed to examine what the most crucial treatment components are and to determine the mechanisms for their success.

6.24.6.3 Pain Disorder

Although there have been suggestions that treatment of pain based on psychodynamic principles may be effective (Adler, Zlot, Hurny, & Minder, 1989), empirical evidence has been scarce and not very compelling (Gamsa, 1994; Turk et al., 1983). Biofeedback (e.g., Budzynski, Stoyva, Adler, & Mullaney, 1973) and hypnosis (Sacerdote, 1970) used to be popular treatments for pain, in the 1970s in particular. These approaches, however, have been harshly criticized on a number of grounds (for a review, see

Turk et al.). Most importantly, critics claim that these techniques are ineffective when used alone, and any treatment effects are due to the unwitting inclusion of behavioral and cognitive treatment components (Turk et al.) We will therefore focus on cognitive-behavioral approaches to the management of pain.

According to operant principles, pain behavior is not an expression or side effect of a pain problem, instead it is the problem (Rachlin, 1985). Therefore, the goal of treatment is to extinguish maladaptive pain behavior and to teach and encourage the use of more adaptive pain coping strategies (Nicholas, Wilson, & Goyen, 1991). Treatment begins with a thorough functional analysis of current pain behavior and the environmental contingencies associated with that behavior. Once the consequences maintaining the maladaptive pain behavior are determined, attempts are made to reduce reinforcement for that behavior and to increase reinforcement for adaptive pain-coping behavior (cf. Fordyce, Roberts, & Sternbach, 1985).

In addition to changing a patient's maladaptive pain behavior, cognitive-behavioral therapy also attempts to alter associated thought patterns and the patient's perceived competence to deal with pain (Schermelleh-Engel et al., 1997). Pain management programs typically include: (i) decreasing demoralization through problem reconceptualization; (ii) enhancing outcome efficacy by motivating active patient participation; (iii) fostering a patient's real and perceived competence to deal with pain; (iv) breaking maladaptive behavior and thought patterns; (v) teaching skills and adaptive responses to problems; (vi) teaching self-attribution for change and success; and (vi) planning maintenance.

Cognitive methods involving imagery and distraction have often been added to pain management programs (Turk et al., 1983). Based on the assumption that pain is mediated by thoughts and images, patients are encouraged to create and focus on those thoughts and images that alleviate pain, while ignoring those that exacerbate pain. Particularly for patients who engage in few activities, behavioral distraction may be beneficial in more than one way. This technique involves the introduction of overt activities that require attention in order to be performed (i.e., hobbies). As clients continue with these activities, they frequently recognize the value of the activities beyond their role as distractors.

Although the cognitive-behavioral approach to pain management has been identified as an empirically validated treatment by the Task Force on Promotion and Dissemination of Psychological Procedures (Chambless, 1995), it is unclear which specific methods are most effective. In a review of the experimental literature, Turk et al. (1983) concluded that patients successfully use cognitive strategies, but the data "do not convincingly establish the efficacy of any cognitive coping strategy relative to the strategies that subjects bring to experiments, nor is there sufficient evidence to support the use of any one strategy compared to any one other" (p. 96). Others have raised more fundamental criticisms and argued that the effects attributed to direct cognitive manipulations could be due to associated environmental manipulations (Rachlin, 1985).

6.24.6.4 Conversion Disorder

An important first step in the treatment of conversion symptoms is their early recognition in which a physical examination plays a crucial role. In many cases, a positive diagnosis can be made on the basis of the rather atypical or bizarre symptoms. Since conversion symptoms vary widely across patients, treatment needs to be individualized. Identifying precipitating stressors is an important treatment strategy so that patients can be taught more adaptive ways of coping with these stressors. Occasionally, manipulation of the patient's social environmental is necessary to reduce the influence of secondary gain. Partners and significant others may have to learn how to reinforce the patient's nonsymptomatic behavior. Favorable prognosis seems to be related to acute onset, massive stressor, and a good premorbid psychological and medical condition (Kent et al., 1995).

6.24.6.5 Body Dysmorphic Disorder

Behavioral interventions for BDD aim at changing avoidance behaviors, reassurance-seeking, checking, and excessive grooming. Exposure *in vivo* is used to counter avoidance of social situations (meeting people, having a conversation, being in the spotlight), body image situations (wearing camouflaging clothing such as baggy pants, long hair, sunglasses). Patients are encouraged to expose themselves to social situations, rather than avoiding them and to observe the reactions of other people to their imagined deformity. Rosen (1995) suggests exposure assignments such as wearing trendy clothes, using make-up to accentuate features, standing closer to people, undressing in front of one's spouse. Response prevention may involve stopping looking in the mirror for excessive periods of time and refraining from the use of make-up and from inspecting skin blemishes.

There are few controlled treatment outcome studies for BDD, but several case studies and uncontrolled trials have been reported that have used systematic desensitization (e.g., Munjack, 1978), exposure and response prevention (Neziroglu & Yaryura-Tobias, 1993), and social skills training (Braddock, 1982). Schmidt and Harrington (1995) describe a successful cognitive-behavioral therapy in a 24-year-old male who was preoccupied with having small hands. Nine one-hour sessions with behavioral experiments aimed at challenging beliefs about size of his own hands and other people's attention to his hands. Treatment resulted in a decrease of BDD-related cognitions, distress and avoidance behaviors, as well as a decrease in Beck Depression and Beck Anxiety Inventory scores.

Recently, a first controlled study (Rosen, Reiter, & Orosan, 1995) compared the effectiveness of cognitive-behavioral therapy with a wait-list control condition in 54 female BDD patients. Treatment was conducted in small groups and encompassed eight two-hour sessions. Therapy was aimed at modifying dysfunctional thoughts about the patients' body image, the reduction of appearance-checking, and exposure to avoided situations. Results at post-test and four month follow-up indicated improvement in the active treatment condition on several measures of body image, whereas no such changes occurred in the control condition. Although the results of this study are promising, more controlled studies are needed to corroborate existing data and to examine the role of the various treatment components.

6.24.7 CONCLUSIONS

Unexplained and unexplainable somatic symptoms are very common in the general population. These problems are costly to the individuals concerned in terms of distress and financial expense as well as to society in terms of lost productivity and healthcare costs. Compared with other common psychological dysfunctions (e.g., anxiety and depression), our present conceptual understanding of the somatoform disorders is poor and satisfactory comprehensive models are still lacking. Moreover, for most of these disorders we do not have basic and reliable information on issues such as prevalence, gender differences, etiology, and treatment outcome.

One factor that has impeded a better understanding of the somatoform disorders is the unsatisfactory and somewhat arbitrary current *DSM* classification. We have noted a considerable degree of phenomenological and functional overlap in the problems of persons diagnosed with somatization disorder, hypochondriasis, and chronic pain disorder as well as overlap between somatoform and anxiety disorders. In fact, comorbidity of somatoform disorders with an anxiety disorder, depression, and general medical conditions are the rule rather than the exception. In view of the existing conceptual and diagnostic confusion, vagueness, and imprecision, we question not only the utility of the current criteria for and distinctions between somatoform disorders but also the wisdom of keeping hypochondriasis in particular in a section separate from the anxiety disorders.

Difficulty in relating the physical and psychological aspects of somatoform disorders has led to further confusion. Although we caution against an over-reliance upon medical diagnostic procedures and the explanatory power of current medical theory, diagnoses of the somatoform disorders are based upon the sophistication and accuracy of medical diagnostic procedures. Research and service delivery for patients would benefit from a more balanced approach that is not just focused on finding or excluding somatic abnormalities but one that combines current medical knowledge and assessment techniques with assessments of a patient's behavior, cognitive processes, and social relationships (cf. Fink, 1996). In our work with cardiac patients (Eifert et al., 1997) we observed how a simple reliance on one source of information (medical or psychological) was inadequate for many patients. Instead, it was the combination of sophisticated medical tests and psychological information that yielded the type of knowledge that was useful for recommending and designing the most appropriate treatment for the individual patient.

Hence, one of the most compelling conclusions arising from this chapter is that somatoform disorders cannot be adequately understood, assessed, and treated from a single perspective. Both the *DSM* classification system as well as research in the area could be improved by adopting a multidisciplinary approach and an integrated biopsychosocial perspective. For example, Mayou, Bass, and Sharpe (1995) propose a multidimensional classification of patients with functional somatic symptoms along five dimensions: (i) number and type of somatic symptoms, (ii) mental state (mood and psychiatric disorder), (iii) cognitions (e.g., symptom misinterpretations, disease conviction), (iv) behavioral and functional impairment (illness behavior, avoidance, use of health services), and (v) pathophysiological disturbance (organic diseases, physiological mechanisms such as hyperventilation). As indicated in the section on health anxiety, individuals may be assigned different positions on all of these dimensions across the various types of somatoform disorders. Rather than attempting to find the "correct diagnosis," we recommend assessing patients along the crucial dimensions involved in the regulation of maladaptive illness behavior and devising treatment programs based on such functional dimensional assessments. Beyond the number and type of physical complaints, some of these dimensions include the presence and extent of preoccupation with body and health, symptom misinterpretation, disease suspicion or conviction, disease fear, safety-seeking approach and avoidance behavior, focus and modulation of worry and fear, and pathophysiological processes.

Based on a multicausal and multidisciplinary perspective, there are several potentially fruitful lines for future psychological research into somatoform disorders such as an increased focus on information processing behavior (attribution, attention, and memory) and environmental contingencies for illness behavior (e.g., social, occupational, medical). Psychoneuro-immunological studies may help clarify particular aspects of the nature of the interface between pathophysiological changes, and individual responses to such changes as exemplified in some chronic pain research (cf. Flor et al., 1990). Although the past emphasis on the problems of patients with no demonstrable physical pathology was worthwhile and deserves continued attention, the gray area of persons with some organic pathology, bodily symptoms, and psychological distress deserves greater recognition and needs to be investigated more carefully.

Treatment programs and outcomes are likely to be enhanced further by an improved understanding of these problems from a multidisciplinary perspective. The need for an improved understanding and better treatments is even more pressing for those somatoform problems that have been particularly neglected in the past, such as somatization disorder, BDD, and conversion problems. The relative success of recent cognitive-behavioral treatment programs for persons with unexplained physical symptoms, health anxiety, or chronic pain is promising. These treatment successes may help change the common perception of healthcare providers that people with such problems are just a "pain in the neck" and invariably difficult, or maybe even impossible, to treat and may provide more hope for patients as well.

6.24.8 REFERENCES

Adler, R. H., Zlot, S., Hurny, C., & Minder, C. (1989). Engel's psychogenic pain and the pain-prone patient: A retrospective controlled clinical study. *Psychosomatic Medicine, 51,* 87–101.

Agras, S., Sylvester, D., & Oliveau, D. (1969). The epidemiology of common fears and phobias. *Comprehensive Psychiatry, 10,* 151–156.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Apley, T. (1958). Common denominators in the recurrent pains of childhood. *Procedural Research and Sociological Medicine, 51,* 1023–1027.

Armentrout, D. P., Moore, J. E., Parker, J. C., Hewett, J. E., & Feltz, C. (1982). Pain patient MMPI subgroups: The psychological dimensions of pain. *Journal of Behavioral Medicine, 5,* 201–211.

Asmundson, G. J., & Norton, G. R. (1995). Anxiety sensitivity in patients with physically unexplained chronic back pain. *Behaviour Research and Therapy, 33,* 771–777.

Bach, M., & Bach, D. (1995). Predictive value of alexithymia: A prospective study in somatizing patients. *Psychotherapy and Psychosomatics, 64,* 43–48.

Barlow, D. H. (1988). *Anxiety and its disorders: The nature and treatment of anxiety and panic.* New York: Guilford Press.

Barsky, A. J., Brener, J., Coeytaux, R. R., & Cleary, P. D. (1995). Accurate awareness of heartbeat in hypochondriacal and non-hypochondriacal patients. *Journal of Psychosomatic Research, 39,* 489–497.

Barsky, A. J., & Klerman, G. L. (1983). Overview: Hypochondriasis, bodily complaints, and somatic styles. *American Journal of Psychiatry, 140,* 273–283.

Bass, C. M., & Benjamin, S. (1993). The management of the chronic somatizer. *British Journal of Psychiatry, 162,* 472–480.

Bass, C. M., & Murphy, M. R. (1990). Somatization disorder: Critique of the concept and suggestions for future research. In C. M. Bass (Ed.), *Somatization. Physical symptoms and psychological illness* (pp. 301–332). Oxford: Blackwell.

Bass, C. M., & Murphy, M. R. (1995). Somatoform and personality disorders: Syndromal comorbidity and overlapping developmental pathways. *Journal of Psychosomatic Research, 39,* 403–427.

Baum, W. M. (1974). On two types of deviations from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior, 22,* 231–242.

Beck, J. G., Berisford, M. A., Taegtmeyer, H., & Bennett, A. (1990). Panic symptoms in chest pain without coronary artery disease: A comparison with panic disorder. *Behavior Therapy, 21,* 241–252.

Beitman, B. D., Basha, I., Flaker, G., DeRosear, L., Mukerji, I. V., Trombka, L. and Katon, W. (1987). Atypical or non-anginal chest pain: Panic disorder or coronary artery disease? *Archives of Internal Medicine, 147,* 1548–1552.

Berrios, G. E., & Kan C. S. (1996). A conceptual and quantitative analysis of 178 historical cases of dysmorphophobia. *Acta Psychiatrica Scandinavica, 94,* 1–7.

Bianchi, G. N. (1973). Patterns of hypochondriasis: A principal components analysis. *British Journal of Psychiatry, 122,* 541–548.

Braddock, L. E. (1982). Dysmorphophobia in adolescence: A case report. *British Journal of Psychiatry, 140,* 199–201.

Budzynski, T. H., Stoyva, J. M., Adler, C. S., & Mullaney, D. J. (1973). EMG biofeedback and tension headache: A controlled outcome study. *Psychosomatic Medicine, 35,* 484–496.

Cailliet, R. (1993). *Pain: Mechanisms and management*. Philadelphia: F. A. Davis

Cairns, D., & Pasino, J. (1977). Comparison of verbal reinforcement and feedback in the operant treatment of disability due to chronic low back pain. *Behavior Therapy, 8,* 621–630.

Chambless, D. L. (1995). Training in and dissemination of empirically validated psychological treatments: Report and recommendations by the Task Force on Psychological Procedures. *Clinical Psychologist, 48,* 3–23.

Chambless, D. L., & Graceley, E. J. (1989). Fear of fear and the anxiety disorders. *Cognitive Therapy and Research, 13,* 9–20.

Chaturvedi, S., & Michael, A. (1986). Chronic pain in a psychiatric clinic. *Journal of Psychosomatic Research, 30,* 347–354.

Conger, R., & Killeen, P. (1974). Use of concurrent operants in small group research: A demonstration. *Pacific Sociological Review, 17,* 399–416.

Craig, K. D. (1986). Social modeling influences: Pain in context. In R. A. Sternbach (Ed.), *The psychology of pain* (2nd ed., pp. 67–95). New York: Raven Press.

Craig, T. K. J., Boardman, A. P., Mills, K., Daly-Jones, O., & Drake H. (1993). The south London somatisation study. I: Longitudinal course and the influence of early life experiences. *British Journal of Psychiatry, 163,* 579–588.

Dallenbach, K. M. (1939). Pain: History and present status. *American Journal of Psychology, 52,* 331–347.

Eifert, G. H. (1992). Cardiophobia: A paradigmatic behavioral model of heart-focused anxiety and non-anginal chest pain. *Behaviour Research and Therapy, 30,* 329–345.

Eifert, G. H. (1996). More theory-driven and less diagnosis-based behavior therapy. *Journal of Behavior Therapy and Experimental Psychiatry, 27,* 75–86.

Eifert, G. H., Edwards, K., Thompson, R. N., Haddad, J., & Frazer, N. L. (1997). *Panic symptoms and chest pain in cardiology patients with and without coronary artery disease*. Manuscript submitted for publication.

Eifert, G. H., Evans, I. M., & McKendrick, V. (1990). Matching treatments to client problems not diagnostic labels: A case for paradigmatic behavior therapy. *Journal of Behavior Therapy and Experimental Psychiatry, 21,* 163–172.

Eifert, G. H., & Forsyth, J. F. (1996). Heart-focused and general illness fears in relation to parental medical history and separation experiences. *Behaviour Research and Therapy, 34,* 735–739.

Eifert, G. H., Hodson, S. E., Tracey, D. R., Seville, J. L., & Gunawardane, K. (1996). Heart-focused anxiety, illness beliefs, and behavioral impairment: Comparing healthy heart-anxious patients with cardiac and surgical inpatients. *Journal of Behavioral Medicine, 19,* 385–399.

Eifert, G. H., Seville, J. L., Antony, M. M., Brown, T. A., & Barlow, D. H. (1992, November). *Panic disorder and cardiophobia: An examination of heart-focused anxiety and illness beliefs and behaviors in persons with panic disorder*. Paper presented at the 26th Annual Conference of the Association for the Advancement of Behavior Therapy, Boston, MA.

Fava, G. A. (1992a). Morselli's legacy: Dysmorphophobia. *Psychotherapy and Psychosomatics, 58,* 117–118.

Fava, G. A. (1992b). The concept of psychosomatic disorder. *Psychotherapy and Psychosomatics, 58,* 1–12.

Fava, G., & Grandi, S. (1991). Differential diagnosis of hypochondriacal fears and beliefs. *Psychotherapy and Psychosomatics, 55,* 114–119.

Fava, G., Kellner, R., Zielezny, M., & Gurand, S. (1988). Hypochondriacal fears in agoraphobia. *Journal of Affective Disorders, 14,* 239–244.

Fernandez, E., & McDowell, J. J. (1995). Response–reinforcement relationshoips in chronic pain syndrome: Applicability of Hernstein's law. *Behaviour Research and Therapy, 33,* 855–863.

Fink, P. (1996). Somatization—Beyond symptom count. *Journal of Psychosomatic Research, 40,* 7–10.

Fishbain, D. A., & Goldberg, M. (1991). The misdiagnosis of conversion disorder in a psychiatric emergency service. *General Hospital Psychiatry, 13,* 177–181.

Flor, F., Birbaumer, N., & Turk, D. C. (1990). The psychobiology of chronic pain. *Advances in Behaviour Research and Therapy, 12,* 47–84.

Flor, F., Kerns, R. D., & Turk, D. C. (1987). The role of spouse reinforcement, perceived pain, and activity levels of chronic pain patients. *Journal of Psychosomatic Research, 31,* 251–259.

Ford, C. V. (1995). Dimensions of somatization and hypochondriasis. *Neurologic Clinics, 13,* 241–253.

Fordyce, W. E. (1976). *Behavioral methods for chronic pain and illness*. St. Louis, MO: Mosby.

Fordyce, W. E., Roberts, A. H., & Sternbach, R. A. (1985). The behavioural management of chronic pain: A response to critics. *Pain, 22,* 113–125.

Forsyth, J. P., & Eifert, G. H. (1998) . Anxiety disorders. In J. J. Plaud & G. H. Eifert (Eds.), *From behavior theory to behavior therapy* (pp. 38–67). Needham Heights, MA: Allyn & Bacon.

Forsyth, J. P., Eifert, G. H., & Thompson, R. N. (1996). Systemic alarms in fear conditining—II: An experimental methodology using 20% CO_2 inhalation as a UCS. *Behavior Therapy, 27,* 391–415.

France, R. D., Krishnan, K. R. R., Houpt, J. L., & Maltbie, A. A. (1984). Differentiation of depression from chronic pain with the dexamethason suppression test and DSM-III. *American Journal of Psychiatry, 141,* 1577–1579.

Freud, S. (1956). *Collected papers*. London: Hogarth.

Frost, R. O., Williams, N., & Jenter, C. (1995). *Perfectionism and body dysmorphic disorder*. Paper presented at the World Congress of Behavioral and Cognitive Therapies, Copenhagen, Denmark.

Gamsa, A. (1994). The role of psychological factors in chronic pain. I. A half century of study. *Pain, 57,* 5–15.

Goldberg, D., Gask, L., & O'Dowd, T. (1989). The treatment of somatization: Teaching techniques of reattribution. *Journal of Psychosomatic Research, 33,* 689–695.

Hanback, J. W., & Revelle, W. (1978). Arousal and perceptual sensitivity in hypochondriasis. *Journal of Abnormal Psychology, 87,* 523–530.

Harvey, R. F., Salih, W. Y., & Read, A. E. (1983). Organic and functional disorders in 2000 gastroenterology outpatients. *Lancet, i,* 632–634.

Hayes, S. C., & Wilson, K. G. (1994). Acceptance and committment therapy: Altering the verbal support for experiential avoidance. *Behavior Analyst, 17,* 289–303.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13,* 243–266.

Hitchcock, P., & Mathews, A. (1992). Interpretation of bodily symptoms in hypochondriasis. *Behaviour Research and Therapy, 30,* 223–234.

Hunter, R. C. A., Lohrenz, J. G., & Schwartzman, A. E. (1964). Nosophobia and hypochondriasis in medical students. *Journal of Nervous and Mental Disorders, 139,* 147–152.

Iezzi, A., & Adams, H. E. (1993). Somatoform and factitious disorders. In P. B. Sutker & H. E. Adams (Eds.), *Comprehensive handbook of psychopathology* (2nd ed., pp. 167–201). New York: Plenum.

Jacob, R. G., & Turner, S. M. (1984). Somatoform disorders. In S. M. Turner & M. Hersen (Eds.), *Adult psychopathology and diagnosis* (pp. 304–328). New York: Wiley.

Jensen, M. P., Turner, J. A., Romano, J. M., & Karoly, P. (1991). Coping with chronic pain: A critical review of the literature. *Pain, 47,* 249–283.

Kellner, R. (1985). Functional somatic symptoms and hypochondriasis: A survey of empirical studies. *Archives of General Psychiatry, 42,* 821–833.

Kellner, R. (1986). *Somatization and hypochondriasis*. New York: Praeger.

Kellner, R. (1991). *Psychosomatic syndromes and somatic symptoms*. Washington, DC: American Psychiatric Press.

Kent, D. A., Tomasson, K., & Coryell, W. (1995). Course and outcome of conversion and somatization disorders. *Psychosomatics, 36,* 138–144.

Kenyon, F. E. (1964). Hypochondriasis: A clinical study. *British Journal of Psychiatry, 110,* 478–488.

Kenyon, F. E. (1976). Hypochondriacal states. *British Journal of Psychiatry, 129,* 1–14.

Krasner, T. M., Rost, K., Cohen, B., Anderson, M., & Smith, G. R. (1995). Enhancing the health of somatization disorder patients. Effectiveness of short-term group therapy. *Psychosomatics, 36,* 462–470.

Lipowsky, Z. J. (1988). Somatization: The concept and its clinical applicability. *American Journal of Psychiatry, 145,* 1358–1368.

Love, A. W., & Peck, C. L. (1987). The MMPI and psychological factors in chronic low back pain: A review. *Pain, 28,* 1–12.

Mace, C. J. (1992). Hysterical conversion. I: A history. *British Journal of Psychiatry*, *161,* 369–377.

MacLeod, C., & Mathews, A. (1991). Cognitive–experimental approaches to the emotional disorders. In P. R. Martin (Ed.), *Handbook of behavior therapy and psychological science: An integrative approach* (pp. 116–150). New York: Pergamon.

Margolis, R. B., Zimny, G. H., & Miller, D. (1984). Internists and the chronic pain patient. *Pain*, *20,* 151–156.

Marks, I. M. (1987). *Fears, phobias, and rituals*. New York: Oxford University Press.

Martens, B. K., & Houk, J. L. (1989). The application of Herrnstein's law of effect to disruptive and on-task behavior of a retarded adolescent girl. *Journal of the Experimental Analysis of Behavior*, *51,* 17–27.

Mayou, R., Bass, C. M., & Sharpe, M. (Eds.) (1995). *Treatment of functional somatic symptoms*. Oxford, UK: Oxford University Press.

Mayou, R., Bryant, B., Forbar, C., & Clark, D. (1994). Non-cardiac chest pain and benign palpitations in the cardiac clinic. *British Heart Journal*, *72,* 548–553.

Mayou, R., & Sharpe, M. (1995). Patients whom doctors find difficult to help. *Psychosomatics*, *36,* 323–325.

McDowell, J. J. (1981). On the validity and utility of Herrnstein's hyperbola in applied behavior analysis. In C. M. Bradshaw, E. Szabadi, & C. F. Lowe (Eds.), *Quantification of steady-state operant behaviour* (pp. 311–324). Amsterdam: Elsevier.

McDowell, J. J. (1982). The importance of Herrnstein's mathematical statement of the law of effect for behavior therapy. *American Psychologist*, *37,* 771–779.

Mechanic, D. (1964). The influence of mothers on their children's health attitudes and behavior. *Pediatrics*, *33,* 444–453.

Melzack, R., & Wall, P. (1965). Pain mechanisms: A new theory. *Science*, *50,* 971–979.

Miller, N. E. (1977). The effects of learning on visceral functions. *New England Journal of Medicine*, *296,* 1274–1278.

Munjak, D. J. (1978). The behavioural treatment of dysmorphophobia. *Journal of Behavioural Therapy and Experimental Psychiatry*, *9,* 53–56.

Murphy, M. R. (1990). Classification of the somatoform disorders. In C. M. Bass (Ed.), *Somatization. Physical symptoms and psychological illness* (pp. 10–39). Oxford, UK: Blackwell.

Newmark, J., & Hochberg, F. H. (1987). "Doctor, it hurts when I play": Painful disorders among instrumental musicians. *Medical Problems of Performing Arts*, *22,* 93–97.

Neziroglu, F. A., & Yaryura-Tobias, J. A. (1993). Exposure, response prevention, and cognitive therapy in the treatment of body dysmorphic disorder. *Behavior Therapy*, *24,* 431–438.

Nicholas, M. K., Wilson, P. H., & Goyen, J. (1991). Operant–behavioural and cognitive–behavioural treatment of chronic low back pain. *Behaviour Research and Therapy*, *29,* 225–238.

Noyes, R., Reich, J., Clancy, J., & O'Gorman, T. W. (1986). Reduction in hypochondriasis with treatment of panic disorder. *British Journal of Psychiatry*, *149,* 631–635.

Parker, G., & Lipscombe, P. (1980). The relevance of early parental experiences to adult dependency, hypochondriasis, and utilization of primary physicians. *British Journal of Medical Psychology*, *53,* 355–363.

Parkes, C. M. (1972). *Bereavement: Studies of grief in adult life*. New York: International Universities Press.

Pennebaker, J. W. (1982). *The psychology of physical symptoms*. New York: Springer.

Pilowsky, I. (1967). Dimensions of hypochondriasis. *British Journal of Medical Psychology*, *113,* 89–93.

Pilowsky, I. (1993). Aspects of abnormal illness behaviour. *Psychotherapy and Psychosomatics*, *60,* 62–74.

Pilowsky, I., Chapman, C. R., & Bonica, J. J. (1977). Pain, depression, and illness behavior in a pain clinic population. *Pain*, *4,* 183–192.

Pilowsky, I., Smith, Q. P., Katsikitis, M. (1987). Illness behavior and general practice utilisation: A prospective study. *Psychosomatic Research*, *31,* 177–183.

Pribor, E. F., Yutzy, S. H., Dean, J. D., & Wetzel, R. D. (1993). Briquet's syndrome, dissociation, and abuse. *American Journal of Psychiatry*, *150,* 1507–1511.

Rachlin, H. C. (1985). Pain and behavior. *Behavioral and Brain Sciences*, *8,* 43–83.

Rief, W. (1996). Die somatoformen Störungen—Grosses unbekantes Land zwischen Psychologie und Medizin. *Zeitschrift für Klinische Psychologie*, *25,* 173–189.

Rief, W., Hiller, W., Geissner, E., & Fichter, M. M. (1995). A two-year follow-up study of patients with somatoform disorders. *Psychosomatics*, *36,* 376–386.

Robbins, J. M., & Kirmayer, L. J. (1991). Attributions of common somatic symptoms. *Psychological Medicine*, *21,* 1029–1045.

Rosen, J. C. (1995). The nature of body dysmorphic disorder and treatment with cognitive behavior therapy. *Cognitive and Behavioral Practice*, *2,* 143–166.

Rosen, J. C., Reiter, J., & Orosan, P. (1995). Cognitive behavioral body image therapy for body dysmorphic disorder. *Journal of Consulting and Clinical Psychology*, *63,* 263–269.

Sacerdote, P. (1970). Theory and practice of pain control in malignancy and other protracted or recurring painful illness. *International Journal of Clinical and Experimental Hypnosis*, *18,* 160–180.

Salkovskis, P. M. (1996). The cognitive approach to anxiety: Threat beliefs, safety-seeking behavior, and the special case of health anxiety and obsessions. In P. M. Salkovskis (Ed.), *Frontiers of cognitive therapy* (pp. 49–74). New York: Guilford Press.

Salkovskis, P. M., & Clark, D. M. (1993). Panic and hypochondriasis. *Advances in Behaviour Research and Therapy*, *15,* 23–48.

Salkovskis, P. M., & Warwick, H. M. C. (1986). Morbid preoccupations, health anxiety and reassurance: A cognitive–behavioural approach to hypochondriasis. *Behaviour Research and Therapy*, *24,* 597–602.

Salminen, J. K., Saarijävi, S., & Äärelä, E. (1995). Two decades of alexithymia. *Journal of Psychosomatic Research, 39,* 803–807.

Salmon, P., & Calderbank, S. (1996). The relationship of childhood physical and sexual abuse to adult illness behavior. *Journal of Psychosomatic Research, 40,* 329–336.

Schermelleh-Engel, K., Eifert, G. H., Moosbrugger, H., & Frank, D. (1997). Perceived competence and anxiety as determinants of maladaptive and adaptive coping strategies of chronic pain patients. *Personality and Individual Differences, 22*, 1–10.

Schmidt, N. B., & Harrington, P. (1995). Cognitive behavioral treatment of body dysmorphic disorder: A case report. *Journal of Behaviour Therapy and Experimental Psychitry, 26,* 161–167.

Schuldt, K., Ekholm, J., Harms-Ringdahl, K., Aborelius, U., & Nemeth, G. (1987). Influence of sitting postures on neck and shoulder EMG during arm–hand work movements. *Clinical Biomechanics, 2,* 126–139.

Sharma, P., & Chaturvedi, S. K. (1995). Conversion disorder revisited. *Acta Psychiatria Scandinavica, 92,* 301–304.

Sharpe, M., & Bass, C. M. (1992). Pathophysiological mechanisms in somatization. *International Review of Psychiatry, 4,* 81–97.

Sharpe, M., Mayou, R., & Bass, C. M. (1995). Concepts, theories, and terminology. In R. Mayou, C. M. Bass, & M. Sharpe (Eds.), *Treatment of functional somatic symptoms* (pp. 3–16). Oxford, UK: Oxford University Press.

Smith, G. R., Monson, R. A., & Ray, D. C. (1986). Psychiatric consultation in somatization disorder. A randomized controlled study. *New England Journal of Medicine, 314,* 1407–1413.

Speckens, A. E. M., van Hemert, A. M., Spinhoven, P., Hawton, K. E., Bolk, J. H., & Rooijimans, H. G. M. (1996). Cognitive behavioral therapy for medically unexplained physical symptoms: A randomised controlled trial. *British Medical Journal, 311,* 1328–1332.

Starcevic, V. (1988). Diagnosis of hypochondriasis: A promenade through the psychiatric nosology. *American Journal of Psychotherapy, 42,* 197–211.

Taylor, G. J., Bagby, R. M., & Parker, J. D. A. (1991). The alexithymia construct. A potential paradigm for psychosomatic medicine. *Psychosomatics, 32,* 153–164.

Turk, D. C., & Flor, H. (1984). Etiological theories and treatments for chronic low back pain. II. Psychological models—interventions. *Pain, 19,* 209–233.

Turk, D. C., Meichenbaum, D., & Genest, M. (1983). *Pain and behavioral medicine: A cognitive–behavioral perspective.* New York: Guilford Press.

Turk, D. C., Okifuji, A., & Scharff, L. (1995). Chronic pain and depression: Role of perceived impact and perceived control in different age cohorts. *Pain, 61,* 93–101.

Turk, D. C., & Salovey, P. (1984). Chronic pain as a variant of depressive disease: A critical reappraisal. *Journal of Nervous and Mental Disease, 172,* 398–403.

Tyrer, P. (1995). Somatoform and personality disorders: Personality and the soma. *Journal of Psychosomatic Research, 39,* 395–397.

Tyrer, P., Lee, I., & Alexander, J. (1980). Awareness of cardiac function in anxious, phobic and hypochondriacal patients. *Psychological Medicine, 10,* 171–174.

Van Hemert, A. M., Hengeveld, M. W., Bolk, J. H., Rooijmans, H. G. M., & Vandenbroucke, J. P. (1993). Psychiatric disorders in relation to medical illness among patients of a general outpatient clinic. *Psychological Medicine, 23,* 167–173.

Vingerhoets, A. J. J. M., Van Heck, G. L., Grim, R., & Bermond, B. (1995). Alexithymia: A further exploration of its nomological network. *Psychotherapy and Psychosomatics, 64,* 32–42.

Visser, S., & Bouman, T. (1992). Cognitive–behavioural approaches in the treatment of hypochondriasis: Six single case cross-over studies. *Behaviour Research and Therapy, 30,* 301–306.

von Korff, M., Dworkin, S. F., & LeResche, L. (1990). Graded chronic pain status: An epidemiologic evaluation. *Pain, 40,* 279–291.

Warwick, H. M. C., Clark, D. M., Cobb, A. M., & Salkovskis, P. M. (1996). A controlled trial of cognitive–behavioural treatment of hypochondriasis. *British Journal of Psychiatry, 169,* 189–195.

Warwick, H. M. C, & Marks, I. M. (1988). Behavioural treatment of illness phobia. *British Journal of Psychiatry, 152,* 239–241.

Warwick, H. M. C., & Salkovskis, P. M. (1990). Hypochondriasis. *Behaviour Research and Therapy, 28,* 105–117.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.25
The Treatment of Substance Abuse and Dependence

ROBIN J. DAVIDSON
Belvoir Park Hospital, Belfast, UK

6.25.1 INTRODUCTION 567
6.25.1.1 Scope of the Problem 568
6.25.1.2 A Longitudinal Perspective 569
6.25.1.3 Assessment 570
6.25.2 THE STAGES OF CHANGE 571
6.25.3 PSYCHOLOGICALLY BASED TREATMENT METHODS 573
6.25.3.1 Motivational Interviewing 573
6.25.3.2 Behavioral Interventions 574
6.25.3.3 Relapse Prevention 575
6.25.3.4 Psychosocial Interventions 578
6.25.4 ISSUES IN TREATMENT 579
6.25.4.1 Treatment Intensity 580
6.25.4.2 Treatment Matching 580
6.25.5 CONCLUSION 581
6.25.6 REFERENCES 582

6.25.1 INTRODUCTION

The use of substances that alter mood, behavior, or cognitions has been a part of human life across numerous social contexts throughout history. Invariably, there are some individuals whose use of such substances may lead to abuse and eventual psychological, social, or physical harm. The fourth edition of the *Diagnostic and statistical manual of mental disorders* (*DSM-IV*; American Psychiatric Association, 1994) and the 10th edition of the *International classification of diseases* (*ICD-10*, World Health Organization, 1992) criteria for substance dependence owe much to the original description of the alcohol dependence syndrome (Edwards & Goss, 1976). These authors proposed a biaxial concept with dependence forming one axis, the other being alcohol-related consequences or disabilities. The elements first proposed for the alcohol dependence syndrome provide a framework for the description of psycho-active substance dependence in both major diagnostic systems.

Dependence was seen by Edwards and Goss as a cluster of behavioral, cognitive, and physiological phenomena which develop after repeated use of alcohol. The elements were persistent use in the light of harmful circumstances, increasing salience of alcohol-seeking behavior, narrowing of drinking repertoire, difficulties in controlling intake, relief use, and neuroadaptive changes producing tolerance and withdrawal symptomatology.

The *ICD-10* criteria for substance dependence are outlined in Table 1. Most of the components of dependence in *ICD-10* are to be found in *DSM-IV* and Hasin, Li, McCloud, and Endicott (1996) report excellent agreement between the two systems. One problem, however, with any diagnosis of psychoactive substance dependence is that it gives the impression of an all or none classification. This is not the case as there is continuous variation within all the elements, and so dependence should be seen as a continuum ranging from mild to severe. Mild dependence on alcohol could, for example, be regarded as a statistically normal condition. Nonetheless, the dependence syndromes as defined in the two major diagnostic classifications are robust and there is a body of factor analytic evidence that indicates that the elements form a single dimension, thus emphasizing the homogeneity of the syndrome (Davidson, Bunting, & Raistrick, 1989; Feingold & Rounsaville, 1995). This latter study also confirms that the dependence syndrome, as measured by *DSM-IV* diagnostic criteria, is factorially distinct from measures of the consequences of substance abuse in all drug groups. In other words, some people use drugs safely. Others, who may or may not develop symptoms of dependence, abuse drugs and encounter problems.

Thus drug abuse and its debilitating *sequelae* are not necessarily synonymous with dependence. Abuse has been defined by the Royal College of Psychiatrists (1979) as the use of a substance that "harms or threatens to harm the physical or mental health or social well-being of an individual." Harmful consequences can arise from prolonged and excessive drug use, but can equally occur as a result of acute intoxication in an occasional user, or indeed can be associated with the mode of administration of the drug. For example the recreational use of ecstasy at a Saturday night dance can have fatal consequences, and HIV infection is not restricted to severely dependent, long-term abusers. This chapter will focus primarily on the treatment of addictive behavior itself, rather than the harmful consequences associated with drug and alcohol abuse.

6.25.1.1 Scope of the Problem

The extent and demography of substance use is constantly changing. However, Babor (1994) draws a few broad conclusions about current trends. Use tends to be concentrated among the young and, with the exception of nicotine, the more addictive the substance, the more socially marginalized the user. The gender differential is diminishing and while the overall use of illicit drugs has increased in developed countries since 1985, alcohol sales have declined slightly (Anderson, 1997).

Since 1973 alcohol consumption in Europe has decreased from a pure alcohol equivalent, annual, per capita intake of 14 litres to 10.4 litres (Gual & Colom, 1997). Consumption has since stabilized and alcohol abuse continues to be a significant drain on economic resources. For example, it has recently been estimated that the total cost of the alcohol problem to the US economy is $70 billion per annum. For every $1 collected in taxes about $7 are spent on problems arising from alcohol abuse. Alcohol continues to be the major drug of abuse within the USA, where there are 18 million alcoholics compared with five million addicted to all other drugs combined (Asbury, 1995). Gelder, Gath, Mayou, and Cowen (1996) outline the results of two major US surveys which suggest a one year prevalence rate for alcohol abuse and dependence of 7–10% with a lifetime risk of around 14–20%.

Ramsey and Percy (1997) summarize the UK prevalence literature on drug use rather than abuse or dependence among the 16–29 year age group. In summary, about one-half of this group report using a prohibited drug ever, around one quarter in the last year, and 14% in the last month. In the surveys, it was estimated that 1% of respondents used heroin at some point in their lives. More direct approaches suggest that opiate use among the general, inner city population in London is around 2%. The Natural Comorbidity Survey in the USA (Kessler, McGonagle, & Zhao, 1994) found that one year prevalence for drug dependence for all adults was 3.6% while lifetime prevalence was 11.9%. Just under 40% in the US sample report using at least one illicit drug in their lifetime. It must be acknowledged that there is a range of prevalence estimation methodologies, which can render rather different estimates and which have been critically evaluated by Sutton and Maynard (1992).

Rather than detail the extent and pattern of use for each substance for different groups of people, some of the problems faced by epidemiologists in this field will be outlined using cocaine by way of illustration. One can measure population *use* of a drug in terms of quantity and frequency, or the extent of drug *dependence* or alternatively drug-related *problems*, such as physical illness, accidents, or emergency admissions. It is important to note that each of these indices can present different distributions and demographic correlates in the population. For example, most of a sample of 100 Canadian cocaine users described it as an infrequent, self-limiting behavior (Erickson & Weber, 1994). This finding was mirrored in a group of Dutch

Table 1 *ICD-10* criteria for substance dependence.

Diagnostic guidelines

A definite diagnosis of dependence should usually be made only if three or more of the following have been experienced or exhibited at some time during the previous year:

(i) A strong desire or sense of compulsion to take the substance.

(ii) Difficulties in controlling substance-taking behavior in terms of its onset, termination, or levels of use.

(iii) A physiological withdrawal state ... when substance use has ceased or been reduced, as evidenced by: the characteristic withdrawal syndrome for the substance; or use of the same (or a closely related) substance with the intention of relieving or avoiding withdrawal symptoms.

(iv) Evidence of tolerance, such that increased doses of the psychoactive substance are required in order to achieve effects originally produced by lower doses.

(v) Progressive neglect of alternative pleasures or interests because of psychoactive subtance use, increased amount of time necessary to obtain or take the substance or to recover from its effects.

(vi) Persisting with substance use despite clear evidence of overtly harmful consequences, such as harm to the liver through excessive drinking, depressive mood states consequent to periods of heavy substance use, or drug-related impairment of cognitive functioning; efforts should be made to determine that the user was actually, or could be expected to be, aware of the nature and extent of the harm.

Source: World Health Organization (1992).

nondeviant cocaine users (Cohen, 1994), in which only one in five proceeded to high use, while sustained problematic use was very much the exception. This distinction between use and dependence would seem to be manifest when samples are not drawn from biased sources, like the treatment or criminal justice systems. At its peak in the mid 1980s, 25 million Americans reported having used cocaine, while three million of these were considered to be dependent on the drug (Kleber, 1988). Reinarman, Murphy, and Waldorf (1994) conclude that dependent, harmful use of cocaine is more contingent on the social circumstances of users and the conditions under which the drug is taken rather than its pharmacological action on human physiology. It is important then to acknowledge the differing prevalence of use, dependence, and problems across sociodemographic groups.

The demographic profile of a drug can also change according to fluctuating availability, price, dependence potential of different forms of the drug, mode of administration, or public attitude. This has been elegantly illustrated by Pickering and Stimpson (1994) in their analysis of two centuries of cocaine use in the developed world. They trace the drug's changing demographic profile from the nineteenth century, when it was mainly taken in the USA by middle class women. Cocaine lost its nineteenth century status as a "tonic" and became a "social problem" when its use by soldiers could have affected the First World War effort. By the early 1960s, users of the drug came largely from the middle-aged, male, bohemian classes. When different routes of administration such as smoking and injecting as well as sniffing became available, the drug began to appeal to a wider group of users. As noted above, the widespread use peaked in 1985, but later the advent of crack cocaine meant that the profile of the typical user became more homogeneous; notably inner-city, socially marginalized young people. This change was mediated by a combination of lower price, more adverse public attitudes and greater dependence-forming potential. The evolving demographic profile of cocaine illustrates the problem of changing fashions in drug use. Babor (1994) notes that while drugs that have abuse potential seem to become a public health problem as a function of the variables outlined above, the decline of drug epidemics is related to many of the same variables.

6.25.1.2 A Longitudinal Perspective

In order to interpret treatment studies it is important to take account of the natural progression of dependence over time, although there have been surprisingly few studies that chart a typical drug-using career. This is a psychosocial construct, which takes account of various environmental influences and developmental changes that shape drug or alcohol use and must be understood in terms of the culture in which the individual lives. Treatment can be defined as a "set of facilitative events which may have the effect of edging an individual into a less destructive long-term career path." Edwards (1989) has similarly noted that treatment should be conceived of being, at best, a timely nudge or whisper in a long-term life course. Most people who successfully work through an addiction problem do so on their own without recourse to treatment. Indeed Prochaska, Di Clemente, and Norcross (1992) found that so-called self-changers often spontaneously employ strategies described in formal therapeutic systems. Klingemann (1994) encapsulated the importance of assessing treatment effects within the context of the natural history of a drug using career in his succinct definition of treatment as "assisted spontaneous remission."

Vaillant (1996) summarizes the long-term relationship between developmental changes and the onset, maintenance, and possible resolution of drug and alcohol abuse. When discussing resolution in particular, he concludes his review of longitudinal studies by arguing that variables such as maturation, treatment, a stable premorbid personality, or even social adjustment may be important in the short term but play a relatively limited role in eventual recovery. Rather, positive long-term outcome depends more on the eventual severity of dependence and "the individual encountering the right kind of naturalistic healing experience." There are very few established psychosocial predictors that can distinguish between those alcohol and drug users who achieve stable outcomes and those who do not. However, there seems to be a slowly emerging consensus in the literature that the key naturalistic factors that contribute to stable outcome include acquisition of an alternative behavior that acts as an addiction substitute, the discovery of new sources of hope and self-esteem and the development of new, stable, supportive relationships. The effects of treatment then must be understood against the backdrop of naturally occurring influences on the process of recovery. Vaillant (1983) outlines a useful set of ground rules for the ideal evaluation of the natural history of any treatment intervention. These include a follow-up period of more than five years, outcome assessment at multiple different times, the inclusion of multiple indicators of social functioning, minimal attrition, and controlling for post-treatment, environmental variables. Unfortunately, many treatment outcome studies reported in the literature fall short of this ideal.

6.25.1.3 Assessment

Before embarking on a treatment program, a sound assessment is required that will inform the choice of treatment goal and content. Information should be attained on the evolution of drug/alcohol intake, family history, patterns of current use, degree of dependence, the extent of drug- and alcohol-related problems, reinforcement parameters maintaining the behavior, and the opportunities within the client's environment for developing more adaptive responses. It is also important to assess the extent of any co-existing psychopathology.

There are numerous scales that have been developed for the assessment of various aspects of drug and alcohol abuse. Normally these instruments are very general and assess a composite of variables from a number of conceptually distinct areas; for example, dependence on the drug, craving, reasons for drug use, attitudes towards self and others, outcome and efficacy expectations, personal, occupational, and social consequences of heavy drug use, drug-taking history, and behavior when intoxicated. The most widely used screening questionnaire for alcohol abuse is the four-item CAGE (Mayfield, McLeod, & Hall, 1974). Other common screening instruments include the Michigan Alcohol Screening Test (Selzer, 1971), the Alcohol Use Disorders Identification Test (Saunders, Aasland, Babor, De La Fuente, & Grant, 1992) and the Drug Abuse Screening Questionnaire (Skinner, 1982). Various semi-structured interviews have also been developed to assist in the evaluation of a variety of drug and alcohol abuse variables. Examples include the Halikas–Crosby drug impairment rating scale for cocaine (Halikas & Crosby, 1991) or the Addiction Severity Index which assesses the severity of problems arising as a result of addiction to one or multiple substances (McClellan et al., 1992).

As noted above these are examples of instruments that measure a broad range of variables and can be used for general screening or assessment purposes. However, they do not provide the clinician with a pure measure of degree of dependence. This is an important predictor of eventual outcome and a number of scales have been developed that are specifically based on the alcohol and drug dependence syndromes. For alcohol, the most commonly applied dependence questionnaires are the Severity of Alcohol Dependence Questionnaire (SADQ) (Stockwell, Murphy, & Hodgson, 1983) and the Short Alcohol Dependence Data (SADD) (Davidson & Raistrick, 1986) in which there are 15 items that enquire about most of the elements of the alcohol dependence syndrome.

The severity of opiate dependence questionnaire SODQ was designed specifically, as the name suggests, to assess opiate dependence (Sutherland, Edwards, Taylor, Phillips, & Gossop, 1988). This instrument primarily focuses on the assessment of withdrawal symptoms and withdrawal relief behavior. As a result it may be difficult to adapt the SODQ to the measurement of dependence on drugs that do not produce a clearly defined withdrawal syndrome. The severity of dependence scale (SDS) is a five-item questionnaire designed to assess the degree of dependence experienced by users of different types of drugs (Gossop et al., 1995). It was devised to provide a short, easily administered scale and the five items are concerned with the psychological components of the drug dependence syndrome. The time frame of inquiry is the past year and the items are as follows:

(i) Do you think your use of (named drug) is out of control?

(ii) Does the prospect of missing a fix (or dose) or not chasing make you anxious or worried?

(iii) Do you worry about your use of (named drug)?

(iv) Do you wish you could stop?

(v) How difficult do you find it to stop or go without (named drug)?

The SDS would seem to be a promising research instrument which can provide an assessment of the degree of dependence experienced by users of a variety of psychotropic drugs, although as yet it is too early to evaluate its usefulness in clinical settings.

The assessment process will assist the therapist in tailoring a treatment program to the needs of the individual as well as elucidating on the most appropriate treatment goal. Gossop (1996) has provided examples of treatment goals which might include the following:

(i) Reduction of psychosocial or physical problems either directly or indirectly related to the drug problem.

(ii) Reduction of risky behavior associated with the use of the drug.

(iii) Attainment of controlled or nondependent use.

(iv) Attainment of abstinence from the problem drug.

(v) Attainment of abstinence from all drugs.

The issue of controlled use rather than abstinence has been debated at length within the alcohol field for many years and is perhaps symbolic of the difference between the more traditional disease approach and explanations of alcohol abuse generated from social learning theory. For this reason the debate was at times conducted at a more personal and anecdotal level than is normally the case in scientific discourse. The question is not "can some alcoholics drink normally?" since the literature is replete with examples to show that they can (Heather & Robertson, 1985), but rather which client will benefit most from abstinence and who would best be suited to a controlled drinking goal. More recently the term "harm reduction" has referred to interventions for drug abusers which involve controlled use rather than abstinence. A good example is the resurgence of interest in the methadone maintenance for opiate abusers and the reported associated benefits (Ball & Ross, 1991). Whatever the treatment goal, a key issue is the client's commitment to implementing real and permanent change in his or her pattern of drug or alcohol use. The following section will review some models that assist in our understanding of the stages in this decision-making process.

6.25.2 THE STAGES OF CHANGE

Psychotherapists have long been interested in the process of change. Pentony (1981) noted that most therapeutically induced change in cognition, affect, or behavior involves an initial destructuring with resistance being a central feature, an intermediate stage of conversion, and a final stage of restructuring. Janis and Mann (1977) analyzed how all sorts of permanent life decisions are made and they note that such change progresses through a number of stages notably reappraisal, considering the options, evaluation, action, and consolidation.

Within the addictions literature the idea of transitional stages or dispositional states has been developed by various authors. Tuchfeld (1976) interviewed a large sample of individuals who had recovered from alcohol dependence without recourse to formal treatment regimes, and on the basis of these interviews he proposed a two-stage model of change. The first is when defensive avoidance becomes untenable and a commitment to change is made. The second stage is one of maintenance of a new behavioral repertoire, which is characterized by personal vigilance and during which time the individual develops coping strategies to attenuate the possibility of relapse. Kanfer and Grimm (1980) also suggest a number of critical transition points as an individual begins to recognize the need for change, prepares for change, and eventually works towards maintaining change and minimizing the risk of relapse.

Arguably, the most influential model of change within the addictions field has been the so-called transtheoretical model of Prochaska and Di Clemente (1984) and Prochaska et al. (1992). The stages of change circle is now probably as familiar to addiction workers as Glatts' U-shaped curve. Orford (1992) likened the development of the transtheoretical model to a Kuhnian paradigm shift and suggested that it has helped rationalize the diversity of psychological interventions and place them in the context of the evolution of personal change. He argued that it allows us to see things from a fresh perspective and casts the process of change in an entirely new mold. The model is summarized in Figure 1.

The transtheoretical model was originally developed after detailed interviews of almost 1000 successful exsmokers, with success being defined as a period of abstinence of four or more years. Essentially during *precontemplation* individuals do not feel impelled to do anything about their behavior, perhaps as a result of ambivalence, denial, or selective exposure to information. As they become aware a problem exists, they enter the *contemplation* stage which is characterized by conflict and dissonance. *Preparation* is defined as a time when the individual drug user formulates action plans and is serious about his or her intention to alter behavior. *Action* is a period when overt changes are made, after which successful individuals enter the *maintenance* stage when new behaviors are strengthened and consolidated. The individual who does not relapse during this stage eventually exits the change system to termination, or in other words favorable long-term outcome.

The authors concede that a linear progression through the stages could be a simplification and acknowledge that on occasions a more cyclical pattern may be in evidence. People may move back from action to contemplation and precontemplation before eventually achieving long-term resolution of the problem. A number of questionnaires have been developed to assess the stage of change for drug abusers, including the eponymous Stages of Change Questionnaire (Prochaska et al., 1992) and the 12-item Readiness to Change scale (Rollnick, Heather, Gould, & Hall, 1992). Essentially these scales yield a score corresponding to the primary stage of change, which is indicated by the person's self-perceived intentions and behavior. Heather (1992) argues that the stages of change can now be reliably assessed and have considerable predictive validity. The key heuristic value of the model is that various categories of intervention are said to be differentially effective at each stage.

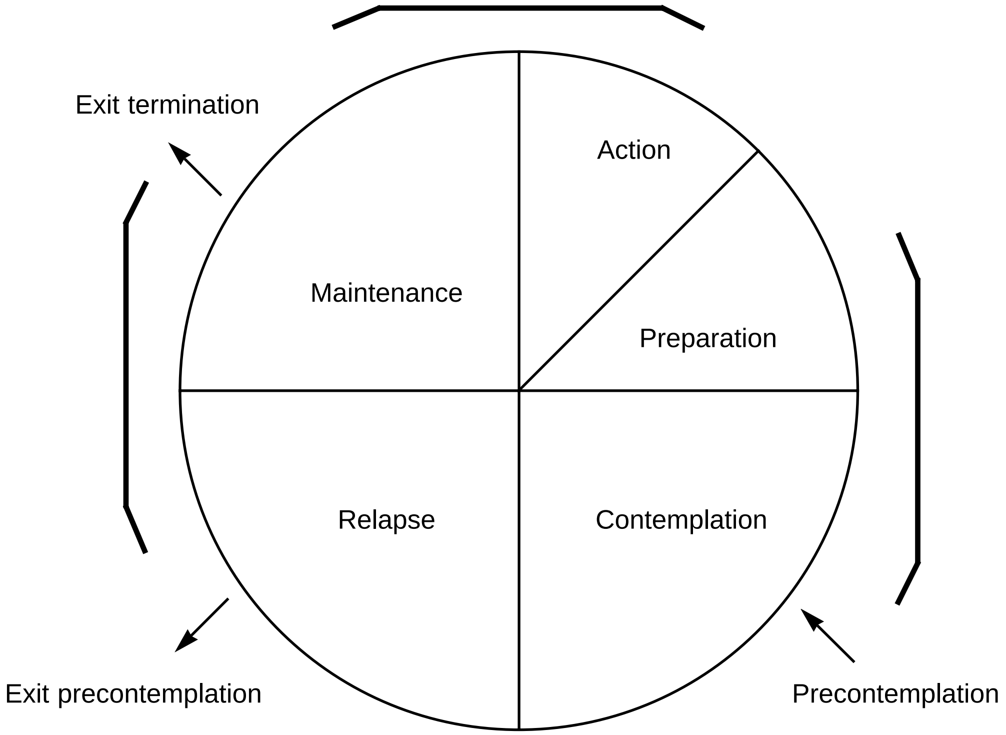

Figure 1 Stages of individual change. Source: adapted from Prochaska, DiClemente, and Norcross (1992).

Perhaps as a result of its intuitive appeal to practitioners and researchers alike the model escaped any considered criticism in the addictions literature for almost a decade. However, more recently a number of authors have questioned its validity and theoretical cohesiveness. Davidson (1992, 1998) commented that it is derivative and noted a number of empirical and conceptual weaknesses. Sutton (1996) felt that it was arbitrary to divide a behavioral/cognitive continuum into five mutually exclusive and exhaustive categories. He also reanalyzed some data on which the model is based and found that individuals did not necessarily progress through the stages of change in the sequence predicted by the model. Few subjects in fact showed a stable progression through three or more stages and, indeed, almost one-third of the sample remained in the same stage throughout the two year time frame. Budd and Rollnick (1996) applied a broadly similar analysis to scores on the Readiness to Change questionnaire, which revealed a continuum of readiness to change. These authors noted that while a stage model has greater intuitive appeal, a continuum model provides a better representation of the available data.

In a stinging criticism of the transtheoretical model, Bandura (1997) announced that stage theories are undergoing a dignified burial in psychology. He argued that this oversimplified stage view substituted a categorical approach for what is essentially a process model of human change. He suggested that people do not recycle through discrete stages but fluctuate in their struggle to exercise control over their health behavior. The stages of precontemplation and contemplation are simply differences in degree of intention while the subsequent stages are gradations of the very behavior that the model seeks to explain. For example, he noted that the action and maintenance stages are simply arbitrary subdivisions of the duration of the new abstinent behavior rather than differences in kind. In other words, this is simply a quantitative rather than a qualitative distinction. Less than six months abstinence from alcohol or drug use is said to define the action stage, while longer than six months defines the maintenance stage.

While all of these criticisms may be regarded as legitimate, the model continues to wield considerable influence in the addictions field and it is quite wrong for Bandura to argue that the stage models are undergoing a dignified burial. Terms such as contemplation and maintenance have become entrenched in addictions parlance and scarcely an article is written on treatment in addictions without reference to the stage of change model. Despite its shortcomings, the model of change does provide a structure to help us match different interventions with the albeit arbitrarily defined stages (Davidson, Rollnick, & MacEwan, 1991). Miller (1983) suggested that motivational interviewing is most useful for individuals in the contemplation stage, although it can prove beneficial for individuals in all stages of change. Behavioral and cognitive interventions can be optimally applied in the preparation and maintenance stages. In summary, despite the criticisms, this model remains a useful guide to assist practitioners in their choice of appropriate interventions.

6.25.3 PSYCHOLOGICALLY BASED TREATMENT METHODS

6.25.3.1 Motivational Interviewing

Motivational interviewing was described by Stockwell (1992) as the most important and influential therapeutic development within the field of addiction over the past decade. In his original account of motivational interviewing, Miller (1983) saw motivational problems as a result of the therapist/client dialogue, with the behavior of the therapist influencing the expectations, attributions, and behavior of the client. Denial was said to be a product of the more traditional, confrontational, therapeutic interaction. During motivational interviewing the individual is encouraged to reach his or her own decision about change, while the role of the therapist is simply to facilitate this process through clarification, advice when appropriate, accurate feedback, and empathy. The aim of therapy is to increase cognitive dissonance until a critical mass of motivation has been achieved and the individual is ready to move from precontemplation to eventual action. At this point commitment to real behavioral change is a likely outcome. Motivational interviewers operationally define motivation as the probability that a person will enter into, continue, and adhere to a specific change strategy and there is a strong emphasis on ambivalence resolution and the decisional balance. Essentially the client begins to present his or her own argument for change rather than being directed by a coercive therapist, while it is the therapist's role to set in place the optimum conditions for change. Specific motivational interviewing strategies and the treatment rationale have been detailed elsewhere (Miller & Rollnick, 1991).

Motivational interviewing is something of a misnomer in as much as it has little to do with contemporary cognitive theories of motivation. Rather it seems to be an example of the phenomenological approach to change and adapts the psychology of self-actualization to the promotion of personal change among alcohol and drug abusers. Miller and Rollnick (1991) contrast motivational interviewing with client-centered counseling by arguing, for example, that empathic reflection is invariably and noncontingently employed in client-centered counseling but used only selectively in motivational interviewing. Furthermore, they say that the good motivational interviewer is not afraid to proffer advice and will actively attempt to create discomfort and discrepancy rather than passively follow the client. Davidson (1996) suggests that this analysis is based on something of a caricature of the Rogerian position and that neo-Rogerian client-centered approaches are more active and task-focused than was hitherto the case. While the distinction between motivational interviewing and contemporary client-centered counseling may be little more than semantic, this is no bad thing. Motivational interviewing is an excellent example of a therapeutic system, squarely based on psychological principles, tailored to individual change in addictive behavior.

Rollnick, Heather, and Bell (1992) described a brief form of motivational interviewing that is beginning to be used to good effect in primary care settings. Miller and Baca (1993) found some evidence of better long-term outcome in a small group of patients who received a brief motivational interview over those who experienced a more directive, traditional style of interview. Baker, Kochan, Dixon, Heather, and Wodak (1994), however, could demonstrate no significant difference in HIV risk-taking behavior between a brief motivational interview and a nonintervention control condition. Kuchipudi, Hobein, Fleckinger, and Iber (1990) found brief motivational interviewing to be unsuccessful in reducing future drinking. Saunders, Wilkinson, and Phillips (1995) present a controlled trial of a brief motivational intervention with 122 drug abusers attending a methadone clinic. Subjects were randomly assigned to a motivational condition or a control, educational procedure. After six months, the motivational subjects showed significantly greater commitment to abstinence, reported more positive outcome expectancies and relapsed less quickly than the control group. The authors concluded that motivational interventions can be a useful adjunct to a methadone program. Heather, Rollnick, Bell, and Richmond (1996) reported a comparison between brief motivational interviewing, skill-based counseling, and a nonintervention control condition in a sample of heavy drinkers. While there were no significant differences between the intervention conditions in terms of a quantity/frequency measure at six months, clients who were evaluated at baseline as "not ready to change" responded better to motivational interviewing. This is interpreted as providing support for the view that motivational interviewing is most appropriate for those in the contemplation stage. Noonan and Moyers (1997) conducted a review of 11 trials that compared motivational interviewing with a range of other treatments. It seemed that motivational interviewing was uniquely effective if it succeeded in eliciting positive motivational responses without evoking resistance. It was, however, not particularly effective in the more severely dependent drinkers. Nonetheless the authors concluded that motivational interviewing was essential for all groups in the assessment interview as it reduced attrition rates.

The enormous popularity of motivational interviewing is in contrast to relative paucity of positive outcome studies. However, on balance it would seem that motivational interviewing strategies act in some way to resolve ambivalence and promote greater commitment to change.

6.25.3.2 Behavioral Interventions

Within the addictions literature the popularity of behavioral interventions has waxed and waned over the years. During the 1980s, there was a move away from such treatments, which were said to be reductionistic, mechanistic, and deterministic, and they were replaced by much more cognitively based multimodal packages of care. However, there has been a renewed interest in once-popular interventions such as cue-exposure and aversive conditioning which lend themselves to parsimonious explanation in the language of conditioning. Cue-exposure is based on the idea that a compulsion (a response to externally conditioned stimulus cues) will be reduced if the urge, which Marlatt (1978) defined as intention to carry out the behavior, is restricted. Cue-exposure can be understood in terms of both classical and operant conditioning. Drug responses, which can include withdrawal symptoms, antagonistic, compensatory, or agonostic effects, can become associated with the internal or external environment. Environmental cues are the conditioned stimuli which when presented in the absence of the unconditioned stimulus can produce this variety of responses. After classical conditioning, instrumental conditioning will occur if the conditioned responses become discriminative stimuli. Cue-exposure should in theory extinguish the conditioned response and so alter future behavior. In practice the client is exposed to cues which would usually trigger an episode of alcohol or drug use but if responses are prevented or controlled, this weakens the stimulus–response relationship.

Drummond, Cooper, and Glautier (1990) reviewed 20 years of cue-exposure research within addictions and concluded, rather disappointingly, that its efficacy must be questioned as there were so few methodologically sound studies that demonstrate its effectiveness in producing good quality, long-term change in the behavior of alcohol and drug users. More recently it has been suggested (Davidson, 1996) that this may be due to the possibility that treatment conditions employed by practitioners have not in fact taken account of the necessary contingencies required to promote maximum change in drug-using behavior. He suggests that there should be greater emphasis on individual differences in cue reactivity, only stimuli relevant to the individual should be included if generalization is to occur in the natural environment, and individuals should be instructed to focus on the most salient aspects of the stimulus. Finally there must be within-session habituation. In practice this means that there should be individually tailored cues of different duration with the termination of each session determined by clear evidence of habituation of the induced responses. Studies with opiate addicts (Powell, Bradley, & Gray, 1993), which have adhered to a theoretically driven treatment design, have produced promising results. Drummond and Glautier (1994) compared cue-exposure with a relaxation control treatment in a randomized control trial involving a sample of 35 severely dependent alcoholics. The cue-exposure clients had a total of 400 minutes of exposure to the sight and smell of preferred drinks over a 10 day period and showed significantly greater improvement on a range of consumption variables at six month follow-up. These recent positive results have tempered the rather pessimistic earlier conclusions of Drummond et al. Cue-exposure is an example of a treatment for which conditioning models neatly predict the optimum conditions for change, without recourse higher order cognitive explanations.

The second example of a treatment based explicitly on conditioning models is aversive conditioning, which was graphically described by Wilson (1991) as a strategy to decrease the appetitive allure of alcohol and drugs. The general criticism, that aversive procedures are inhumane and pejorative, is well known. However, in the context of the present discussion the importance of these interventions are that the conditions of maximum change can be theoretically specified. Cannon and Baker (1981) demonstrated the superiority of emetic therapy or chemical aversion over aversion produced to an electric shock and Elkins (1991) produced an elegant theoretical rationale for the consistent superiority of chemical aversion over electrical aversion for alcohol and drug abusers. Nevertheless, there have been surprisingly few well-controlled outcome studies that demonstrate the long-term efficacy of emetic therapy.

It is possible that most trials to date may not necessarily have adhered to the conditions that should promote maximum change. Of central importance is the temporal continuity between the conditioned and unconditioned response and it may be that some of the failures of emetic therapy are due to the fact that exposure to alcohol or drugs is introduced after the initial peak of nausea. Arguably, there is only one study in the addictions literature to date that is theoretically and methodological flawless (Cannon, Baker, & Wehl, 1981). In this study there was random assessment of clients to chemical aversion and nonaversion treatment conditions. Optimum contingency intervals were employed and a significant superiority of aversive conditioning was demonstrated at one year follow-up. Parloff, London, and Wolf (1986) found that for some alcohol abusers, aversive procedures were consistently better than more cognitively based interventions. Hester and Miller (1989) have outlined a treatment protocol for covert sensitization (verbal aversion). While this offers many practical advances, Elkins (1980) suggested that alcohol misusers who experience real, rather than suggested, nausea after the taste or smell of alcohol remain abstinent for longer than those who do not. However, it is probably fair to conclude, at the time of writing, that the use of aversive procedures for alcohol or drug misusers is relatively rare.

Heather and Stallard (1989) suggest that, generally, contemporary models of relapse perhaps underestimate the importance of classically conditioned craving in the relapse process. It is likely, therefore, that behavioural interventions, particularly cue-exposure, will become more important components of treatment programs for alcohol and drug abusers than they have been hitherto.

6.25.3.3 Relapse Prevention

Substance abuse has long been seen as a chronic, relapsing condition. A sound understanding of relapse management is perhaps the most fundamental issue in a discussion of the nature of treatment for addictive behavior. Relapse prevention is a generic term for a variety of approaches to the treatment of drug and alcohol abuse, primarily aimed at those in the maintenance stage of change. Edwards (1994) points out that it is important to remember that relapse prevention "does not imply a single theory of just one intervention technique but the deployment of a range of methods derived from a variety of behavioural and cognitive postulates" (p. 252).

The term relapse has been afforded a number of meanings by various authors. A lapse is seen by Gossop (1996) as being part of a transitional process which may ultimately lead to favorable long-term outcome, while a relapse is a return to the original pattern of drug intake. Relapse is not necessarily a discrete event that occurs at one moment in time, but rather it is the end point of the process of returning to former patterns of abuse. The process of relapse transcends the actual preferred substance of abuse. This was demonstrated in a seminal paper by Hunt, Barnett, and Branch (1971) who showed similar post-treatment, relapse/survival curves for alcohol, opiate, and nicotine abusers. Indeed, relapse prevention models have subsequently been applied to a range of behaviors about which an individual has made a commitment to behavioral control and has formulated a set of rules to govern this control. It was not until the 1970s that addiction researchers began systematically to examine the nature of the processes involved in relapse. Since then a number of leading models have emerged to guide our thinking and these have been reviewed by Connors, Maisto, and Donovan (1996).

Gloria Litman and her colleagues (e.g., Litman, Eiser, Rawson, & Oppenheim, 1977, 1979) saw relapse as a person–situation interaction. Of primary importance was the high-risk situation and the effectiveness and appropriateness of the individual's coping skills for dealing with it. Relapse was more likely in individuals who had few coping resources and who encountered a relatively large number of risk situations. Sanchez-Craig (1987) placed more emphasis on the individual's perception and appraisal of the risk situation rather than the situation itself. Relapse has also been systematically approached from what has been called a psychobiological perspective. Central to these models is the idea that relapse is precipitated by craving, which results in loss of control. Craving can be seen as a cognitive interpretation of the feelings of arousal associated with drug-related stimuli. Solomon's (1980) opponent process model defines craving as the interplay between positive and negative emotional states, induced by the substance. Tiffany (1990) formulated a model of drug use and relapse which highlighted the importance of automatic and nonautomatic cognitive processing. Automatic processes are characterized by speed, autonomy, effortlessness, and lack of conscious awareness. Examples are the urges and cravings generated by drug withdrawal or the reinforcing effects of drugs. Nonautomatic processes require conscious effort, the choice of competing strategies, the execution of preferred strategies, and the modulation of the strategy to take account of environmental changes. In order to deal with what Tiffany calls automatically induced urges and cravings, the individual must mobilize nonautomatic cognitive processes which require effort, attention, and intention and which can lead to overt behavioral and cognitive coping strategies.

Most relapse prevention programs therefore emphasize the importance of a broad range of cognitive and behavioral strategies which will attenuate the possibility of relapse. Arguably, the most influential model developed with the addictions field has been Marlatt's taxonomy of relapse precipitants.

Marlatt and Gordon (1985) presented relapse prevention as a set of principles broadly based on social learning theory. In a sample of over 300 individuals who used a variety of drugs, they were able to clarify the key relapse precipitants into two broad categories. The first includes intrapersonal precipitants such as negative and indeed positive physical and emotional states. For over one-third of alcohol and nicotine abusers and about one-fifth of heroine abusers the main relapse precipitant was a negative emotional state. The second category is interpersonal precipitants of relapse including relationship conflict and indirect or direct social pressure.

As part of a relapse prevention program, Marlatt and his colleagues suggest that individuals are taught to recognize the possibility of relapse. Essentially the client constructs a personal behavioral analysis and receives training in specific coping strategies. These can include broad-based skills training (behavioral rehearsal, assertiveness training), cognitive reframing (coping imaginary, reframing reactions to lapse), and lifestyle interventions (relaxation and exercise enhancement). All of these have been found to contribute to the effectiveness of a relapse prevention program. Clients are taught to recognize early warning signals and made aware of apparently irrelevant decisions that can increase the possibility of relapse. Emphasis is placed on the modification of cognitive distortions and the challenging of faulty beliefs or dysfunctional assumptions. For example, the *abstinence violation effect* is a distorted redefinition of lapse as relapse, so undermining the effectiveness of future coping behaviour. The Marlatt and Gordon relapse prevention program is therefore a combination of skills training, self-management, and cognitive interventions and the client is encouraged to practice these strategies using rehearsal, role-play, and homework tasks.

Carroll (1996) carried out a detailed review of 24 controlled trials that have evaluated the effectiveness of relapse prevention across a range of addictive behaviors including nicotine, alcohol, marijuana, opioid, and cocaine abuse. In the review, she included only those trials that specifically employed the relapse prevention model of Marlatt and Gordon (1985) and that were, in her opinion, methodologically sound. The outcome variables were primarily post-treatment patterns of substance use, rather than behavioral and cognitive indicators such as coping skills, self-efficacy enhancement, or social adjustment. The review concluded that there was some evidence of effectiveness of this type of relapse prevention approach across a range of substances, particularly smoking cessation, when compared with no treatment controls. However, its superiority over other active treatments and discussion control conditions was less consistent. The specific benefits over other treatments were in reducing the severity of relapses if they occurred and the approach seemed to be of most value for clients with more severe dependence or higher levels of psychopathology. The Relapse, Replication and Extension Project (RREP) was a major multicenter trial established to investigate the validity and reliability of Marlatt's taxonomy of relapse precipitants (Lowman, Allen, & Stout, 1996). Essentially the study confirmed the predictive validity of a modified version of Marlatt's taxonomy, although it did raise some questions about differences between proximal variables associated with time-limited relapse and variables that predict outcome in the longer term.

The relapse prevention model of Annis and Davis (1989) draws more explicitly on self-efficacy theory. The emphasis of this approach is on performance-based methods, notably the exposure to increasingly high-risk situations with continuing self-monitoring of efficacy expectations. In guiding the client through risk situations, Bandura's (1985) four factors that engender strong efficacy expectations are taken into account. First, the situation is challenging; second, to succeed in mastering the situation a moderate degree of effort is needed; third, the client is responsible and external help is kept to a minimum; and fourth, the success is described as part of a pattern of improved performance. The Inventory of Drink Situations helps identify the client's areas of high risk of relapse and the Situational Confidence Questionnaire measures perceived ability to cope effectively with a variety of drug-related risk situations. More recently Moser and Annis (1995) have emphasized the role of coping strategies in their cognitive behavioral model of relapse. They demonstrate that the survival of a relapse is strongly related to the number and type of coping strategies employed. Avoidance strategies are less effective in preventing relapse than active strategies, such as carrying out alternative activities, seeking support from others, positive self-talk, and cognitive problem-solving.

The relapse prevention models described above have been developed specifically within the addictions field, but their use has extended to a range of other clinical contexts, for example, anxiety management, eating disorders, and impulse control. However, there is also a series of models primarily used to inform research on health behavior decision making that emphasize the importance of cognitive processes as crucial determinants, not only in the abstention from risky behavior, but also the adoption of health beneficial behavior. These models have recently been employed within the addictions field.

One of the first and most influential was the Health Belief Model (Becker, 1974) which expressed the likelihood of positive health decision-making in terms of a cost–benefit analysis. This was superseded by the theory of reasoned action (Ajzen & Fishbein, 1980) and then came the theory of planned behavior (Ajzen, 1988). This latter model specified a number of variables that predict intention and behavior, namely control beliefs, subjective norms, and attitudes (a form of outcome expectations). These health belief models have been refined and developed by Ralph Schwarzer in his Health Action Process Approach (HAPA). This idea is now widely applied to the understanding of the fluctuating relapse/survival process of addictive behavior. Schwarzer (1992) comments that traditional relapse prevention approaches, such as that of Marlatt and Gordon (1985), have focused on the maintenance stage and have not promoted understanding of change strategies in the motivational and action phases of change. Essentially, within the motivation phase the individual forms an intention to adopt precautionary measures for changing risk behavior in favor of other behaviors. The action phase is an attempt to explain the relationship between intention and subsequent behavior, with emphasis being placed on action control and action plans. Estimates of the relative proportion of outcome variance explained by predictors of intention and behavior have been gleaned from a variety of sources. For example, outcome expectations and the perception of subjective norms contributed to about half of the variance in intention to stop smoking, but rather disappointingly stated intention to stop smoking accounts for only about 20–25% of the variance in actual observed long-term smoking abstention. However, when beliefs about personal efficacy are added to the model, its explanatory power increases substantially. Schwarzer concludes that self-efficacy expectation, or the belief that addictive behavior can be changed by mobilizing personal resources, is the most powerful predictor of intention and subsequent behavioral control. This view was supported by Kok, De Vries, Muddle, and Strecher (1990), who demonstrated that efficacy expectations alone could account for over two-thirds of the outcome variance of intention to stop and subsequent smoking abstinence. The total predictive effect of self-efficacy on the maintenance of health behavior exceeds the effect of any combination of other cognitive variables. Bandura (1997) agrees, and concludes his review of models of health behavior by noting that outcome expectations have variable effects, normative influences have little effect but efficacy expectations are constantly predictive.

In summary, the work of Annis, Schwarzer, Bandura, and others indicates that self-efficacy is emerging as the most powerful predictor of relapse/survival across a variety of addictive behaviours. Allsop, Saunders, Philips, and Cann (1997) suggest that training in coping skills is not enough, people must use the skills at the right time for them to be effective. The practice of coping skills improves self-efficacy so that the skills can be deployed in real-life risk situations. Furthermore, in recent years there has been much cross-fertilization of ideas between relapse prevention models developed primarily for substance abusers and more general models of health decision-making behavior.

6.25.3.4 Psychosocial Interventions

The prototypical, community self-group is Alcoholics Anonymous (AA) which was founded in 1935 by a serendipitous meeting between two American alcoholics, Bill W and Dr. Bob. Advice to attend AA remains part of many contemporary multimodal treatment packages for alcohol abuse. The principles of change espoused by AA are based on the Oxford group, a religious organization in the evangelical Protestant tradition (Bill W., 1957), which emphasized self-examination, the public admission of character deficits, restitution, pledge taking, and bible reading. Bill W. used this as a metaphor for the organizing principles of AA. At an AA meeting open sharing is encouraged and members admit their powerlessness over alcohol. The AA belief system is articulated in the 12 steps, which include the requirement of a searching personal inventory, commitment to a greater power, making amends to other people, and carrying the message to other alcoholics. Essentially, the AA model of addiction is one of an illness which can only be arrested by complete and life-long abstinence from alcohol; an illness that is said to be arrested but never cured. All attenders, no matter how long they have been alcohol-free, remain designated as recovering alcoholics. An analysis of AA-mediated processes of recovery suggests that motivation can be bolstered by the use of proximal goals (a day at a time) for everyone from the first attender to the 20 year abstainer. Emphasis is placed on role modeling, with constant recourse to a pay-off matrix in which the profit of staying sober is compared to the loss of being drunk. AA also offers, in the best relapse management tradition, alternative activity and new social networks. Much AA folk wisdom, such as not getting angry, tired, or bored; monitoring thinking tricks, and self-justifying statements, is the stuff of social learning theory. This prompted Edwards (1996) to compare an AA meeting to a cognitive behavioral workshop. In a small sample of AA attenders, Edwards et al. (1987) found that the most valued AA activity was hearing other people's stories and the least valued was the quasi-spiritual element.

Tonigan, Ashcroft, and Miller (1995) examined three AA groups in terms of perceived social dynamics and the relative emphasis on the 12 steps. The groups significantly differed in terms of cohesiveness, independence, and expressed aggression and it was found that 12 step discussion was lowest in the groups with the highest aggressiveness. In the light of these results it was suggested that AA need not be regarded as a homogeneous entity and a clear understanding of the variation among AA groups needs to be gained before we can make definitive statements about its effectiveness.

Davies (1992), in his attributional analysis of what he called the "myth of addition," argued that step one, which is an admission of powerlessness over alcohol, can absolve the individual of personal responsibility for his drinking. It minimizes the importance of a functional explanation of drinking behavior and Davies would say that it does not bode well for dealing with relapses should they occur. Ogbourne (1993) observed that we know little about the real long-term effectiveness of AA because as a matter of principle the organization keeps no records of the success and failures of its members. Furthermore as noted above, AA is less homogeneous than many people think. There is no doubt that AA, with its rigid explanations and emphasis on abstinence, does not suit most alcohol abusers and so attrition rate is high (Miller & McCrady, 1993). Edwards (1996) cautions "the evidence that AA works is suggestive and rests on the evidence of its popularity and seeming ability to meet need, rather than being a matter of proven fact" (p. 235). Nonetheless, for some people it is the belief that they are personally powerless over alcohol which can inspire faith and promote change (Keene & Raynor, 1993), and it is indisputable that the profound and international influence of AA belies its humble beginnings. Narcotics Anonymous (NA) was established in 1953. Christo and Franey (1995), in a six month follow-up of over 100 graduates from a residential program, found that NA attendance was inversely related to continued drug use and that spiritual beliefs or beliefs in the disease concept were not prerequisites for NA attendance.

The 12-step approach has been the basis of many community residential groups for alcohol and drug abusers. The 12 steps have been intermeshed with Maxwell Jones's (1956) pioneering models of therapeutic communities to produce concept houses, such as the original Phoenix House in New York, which provided long-term residential programs for addicts. Most of the emergent therapeutic communities in the North America and the UK can trace their roots back to the Phoenix House program. Contemporary Minnesota Model programs vary from group to group, but they are generally based on the first half of the 12 steps and include group work, lectures, and individual assignments, all of which can be preceded by detoxification. When appropriate there is family involvement, which will typically draw on the principles of Al Anon. DeLeon (1995) noted that not all residential drug abuse treatment programs are therapeutic communities and not all programs calling themselves therapeutic communities employ the same psychosocial models of treatment. However, the core themes of the drug and alcohol therapeutic communities are those identified over 30 years ago by Rapoport (1960), namely democratization, communalism, and reality confrontation. More recently, Norton (1996) has usefully reviewed the change process inherent in therapeutic communities and noted that it is difficult to specify the components of programs that best predict positive outcome. Smart (1976) reviewed outcome studies on community graduates and suggested that when recovery is defined solely in terms of patterns of drug use, outcome is fairly good. However when other outcome indices are employed, for example future employment, the results are less encouraging. Thorley (1981) reviewed a number of UK studies and concluded that therapeutic communities for drug abusers can greatly benefit some individuals but can actively harm others. It is also difficult to specify the optimum duration of care within such communities. For example, McCusker et al. (1995) compared a six month with a 12 month therapeutic community program for a mixed group of drug abusers and found no difference on a whole range of psychosocial variables and pattern of drug use at six month follow-up. While therapeutic communities may arguably be an effective rehabilitation system, they are relatively few in number and normally have quite strict entrance criteria, thus rendering them genuinely unavailable to all but a small minority of alcohol and drug abusers.

One final addendum to the AA story is the remarkable growth of the 12-step model within the USA for the "treatment" of a whole range of so-called appetitive disorders. Almost any potential problem from shopping to sex has a 12-step residential program for individuals prone to excess. This raises fundamental questions about choice, compulsion, and volition in our understanding of human behavior. Stanton Peele (1989) in his best-selling book, *Diseasing of America*, argues that there is now more "treatment" for everything and this inevitably undermines an individual's personal perception of control over his or her behavioral outcomes. Self-efficacy, so important in the treatment of any addictive behavior, is being diluted by what Peele calls the "therapeutic colonisation of our lives." One can only speculate how Bill W. and Dr. Bob would have reacted to the widespread use, perhaps abuse, of their simple approach. Maybe it has gone one step too far.

6.25.4 ISSUES IN TREATMENT

There are two key issues relating to the provision of treatment within addictions that have been the subject of much debate in recent years. The first relates to treatment intensity and the emergence of a plethora of so-called brief interventions for substance abusers. The second is the long-held assumption by most workers in the field that treatment should be tailored to the needs of the individual. In other words, which intervention for which client by which therapist is the most efficacious? Most of the ideas on parameters of treatment have been developed in studies of alcohol abuse and it is this literature that will primarily be covered in the following review.

6.25.4.1 Treatment Intensity

There have been over a dozen randomized controlled trials comparing the dose–response relationship in populations seeking psychological treatment for alcohol problems. In this context brief intervention is defined as one to three sessions of specialist treatment. A distinction should be drawn between minimal intervention or advice in a primary care setting and brief intervention within a specialist context.

The seminal study on brief intervention was reported by Edwards et al. (1977), who followed up a sample of 100 male alcoholics, all of whom had a three hour initial assessment. Half of the men then received a single session of direct advice and the other half were offered the standard intensive treatment package of the time which included medical interventions, introduction to AA, and hospital admission when appropriate. This study is well documented elsewhere but the key conclusion, which has influenced thinking on delivery of care to alcoholics, is that even after a 10 year follow-up there was no significant difference between the groups on a host of outcome measures. Subsequent, similar trials have more or less confirmed this finding. Bien, Miller, and Torigan (1993) reviewed this body of work and concluded that there are remarkably few differences between brief and extended interventions for treatment-seeking alcoholics. Well-planned, brief interventions produced similar outcomes when compared with more extended psychosocial interventions, with only two studies reporting a slight advantage for the latter.

The dose–response literature should, however, be interpreted with some caution. These studies are notoriously difficult to design and most of the controlled trails comparing brief with extended treatments have been subject to some criticism. Indeed Orford (1980) would concede that the Edwards' trial was by no means methodologically flawless. Heather (1995) has recently produced a useful evaluation of the literature. He argues that brief interventions tend to be seen as a homogeneous category and researchers, as well as practitioners, have paid insignificant regard to heterogeneity of this type of intervention. Furthermore, there are important differences between treatment-seeking populations as distinct from those that are selected as a result of opportunistic screenings. Heather also highlights the possibility of type two error, in that this family of studies is essentially proving the null hypothesis. It is an inferential error to conclude on the basis of a lack of significant difference between intensive and brief intervention that they are equally effective. Even with a large sample and sufficient statistical power, which is not the case in most of the trials, one can only conclude that there is no evidence of difference in effectiveness. Heather concludes his methodological critique by suggesting that it is at best premature, and at worst cavalier, to argue that highly dependent, treatment-seeking alcohol abusers gain as much from brief interventions as they would from more extended intervention.

There is, however, some merit in investigating the so-called active ingredients of brief interventions which seem to be, at least, relatively successful. Miller and Sanchez (1993) reviewed the protocols of a large number of brief interventions and isolated some critical conditions or key elements that facilitate change in addictive behavior. These can be summarized by the acronym FRAMES, representing feedback, responsibility, advice, menu, empathy, and self-efficacy. Essentially their findings would suggest that for brief interventions to be successful the client should be given realistic and accurate feedback about any drug-related impairment, emphasis should be placed on personal control and efficacy with corresponding attention to feelings of helplessness and powerlessness, good advice should be proffered, and the client should be given a choice of preferred change options. Finally, and arguably most important, therapeutic styles should be empathic rather than confrontational. While acknowledging Heather's caveats, the brief intervention literature does offer hope that for some drug and alcohol abusers, treatments that are structured along these lines could offer potential long-term benefits.

6.25.4.2 Treatment Matching

Almost invariably, any review of the treatment of addictive behavior concludes with a comment that intervention should be matched to the needs of the client. For example, Gossop (1996) notes that "for all types of drug problems that require treatment the intervention offered should be tailored to the needs and circumstances of the individual" (p. 159). Raistrick and Davidson (1985) say that "therapy should be tailored to the individual. What is useful for one person may be singularly inappropriate for another" (p. 154). Miller and Herster (1986) say that "clients should be matched to optimal interventions" (p. 27), and so it goes on.

However, these apparently obvious and laudable matching sentiments have recently been the subject of some scrutiny in the alcohol literature. Edwards and Taylor (1994) report a study that explored the interaction among treatment intensity, client characteristics, and 12 month outcome. There were no interactive effects and the authors conclude that the matching hypothesis was not confirmed. These data suggest that individual characteristics such as degree of dependence on alcohol or socioeconomic class, rather than client/treatment matching, may best predict outcome. Edwards and Taylor draw the tentative conclusion that relatively weak matching differentials are masked by developmental or environmental influences on the process of recovery.

Of particular interest in this regard is Project MATCH which Heather (1996) describes as the largest treatment trial ever mounted in the alcohol field. It was funded by the National Institute on Alcohol Abuse and Alcoholism and involved over 1600 clients tracked over a 15 month period. Meticulous attention was paid to all aspects of the trial design which essentially tested the relative merits of a variety of client/treatment combinations. Clients were randomly assigned to one of the three treatment modalities namely 12-step facilitation (TSF), cognitive behavioral coping skills (CB), and what was called motivational enhancement therapy (MET). TSF was based on the methods of Alcoholics Anonymous, but as it was only conducted for 12 sessions, did not include regular and life-long AA attendance as part of the treatment condition. CB was limited to 12 sessions of coping skills training, and MET was a four session program based on the principles of brief motivational counseling as outlined above. A variety of client characteristics were assessed, including motivational status, self-efficacy, degree of alcohol dependence, extent of social support, the number of alcohol-related problems, psychiatric morbidity, and cognitive impairment. Obviously the interactions between client characteristics and intervention were as important as any potential main effect between treatment conditions. The primary objective of the study was to determine which of various subgroups of alcohol-dependent clients would differentially respond to the three broad interventions. Additionally, the interaction of treatment location (inpatient vs. outpatient) with client characteristics and mode of intervention was examined.

In summary, the results indicated significant and sustained improvement on a range of outcome variables across the three main treatment conditions (Project MATCH, Research Group, 1997). In other words all three types of interventions were equally effective. With the exception of psychiatric morbidity, matching client characteristics with treatment did not enhance outcome. This one "match" seemed to indicate that outpatients with greater psychiatric morbidity responded better to CB than TSF. However, the overall conclusion of the study was that the general lack of robust matching effects means that practitioners need not take these individual characteristics into account when triaging clients to one or other of the three treatment interventions, despite their different philosophies. This is a counterintuitive and radical finding which will undoubtedly resonate in the alcohol treatment field for decades to come. Despite the methodological rigor, it may be argued that the 12-week protocol for CB and TSF and the four-session MET intervention do not reflect the range or richness of available interventions, particularly those based on social learning theory. The length of follow-up was only 15 months in this first Project MATCH publication and it could be that matching effects become manifest much later. Furthermore, it is unclear about the validity of cross-cultural application of these results. For example in the UK, Woodhouse and Davidson (1996) matched broadly similar interventions with motivational stage and while there was no significant difference in a quantity/frequency measure at one year, they found a small but significant matching effect at two years. Essentially contemplators responded relatively better to motivational interviewing, while those in the action stage were best suited to relapse prevention strategies. While this latter finding may not be clinically significant, it may add some weight to the possible criticisms of the Project MATCH trial regarding cross-cultural generalizability of the findings and the relatively short follow-up period.

6.25.5 CONCLUSION

A number of conclusions can be drawn from this review of cognitive and behavioral approaches to the treatment of addictive behavior. Most of the interventions described in the chapter have been offered to users of a range of substances. There has been a tendency to assume, not unreasonably, that the important component of any intervention is the cognitive and behavioral changes that accrue irrespective of the preferred drug of abuse. While this is intuitively appealing, it would seem that it is not always the full story. For example, at the risk of oversimplification, relapse prevention procedures are clearly useful for smokers, moderately effective for alcohol abusers, and have some variable effect, according to the available literature, as a treatment for cocaine abuse. There has, as yet, been no demonstrable significant improvement against a control condition for the relapse prevention treatment of marijuana use (Stephens, Roffman, & Simpson, 1994). Klingemann (1994) has speculated about possible reasons for these substance specific effects. He particularly notes that good quality, extensive social support may become progressively less accessible as the substance changes from cigarettes to alcohol to illicit drugs. While it is obviously important to elucidate on the general principles of change, there may be some merit in clarifying the possible value of substance specific interventions.

In answer to the question "Do psychologically based treatments of addictive behavior work?" most reviewers interpret the literature slightly differently. In the past there has been an overwhelming sense of what people termed therapeutic nihilism, based on the weak relationship between any type of intervention and eventual outcome. There is now, however, increasingly good evidence on the effectiveness of psychological approaches for addictive behavior. Moos, Finney, and Crankite (1990) in their studious review, concluded that treatment leads to substantial improvements in drinking behavior while acknowledging the importance of individual and social factors in the determination of outcome. The treatment outcome prospective study (TOPS) was a systematic evaluation of a heterogeneous group of treatment programs for a range of drug abusers. The TOPS study reported a general overall reduction in drug use over a three to five year period, as well as improvement in other psychosocial variables (Hubbard, Marsden, & Rachal, 1989).

While treatments based on psychological theory have been shown to be effective, no specific treatment approach has emerged as being consistently superior to the others. One problem with conventional, randomized controlled trials, particularly pertinent in the field of substance use disorder, is that the preference of an individual for one particular type of intervention over another is not taken into account in random allocation. A potential, future methodological advance may be greater use of designs like the partially randomized, patient-preference design advocated by Brewin and Bradley (1989) for the investigation of relative treatment effectiveness.

Finally, work on the interaction between treatment intensity and outcome in both inpatient and outpatient settings has been summarized above. Unfortunately at times this debate is conducted against a background of cost considerations rather than what is best for each individual. Room (1980) did well to remind us that the basic justification for providing treatment should be primarily for reasons of humanity rather than cost–benefit analysis.

6.25.6 REFERENCES

Ajzen, I. (1988). *Attitudes, personality and behaviour*. Milton Keynes, UK: Open University Press.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting behavior*. Englewood Cliffs, NJ: Prentice Hall.

Allsop, S., Saunders, B., Philips, M., & Cann, A. (1997). A trial of relapse prevention with severely dependent male problem drinkers. *Addiction, 92,* 61–74.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Annis, H. M., & Davis, C. S. (1989). Relapse prevention. In R. K. Hester & W. R. Miller (Eds.), *Handbook of alcoholism treatment approaches: Effective alternatives* (pp. 137–148). New York: Pergamon.

Anderson, P. (1997). Health, society and alcohol. *Addiction, 92,* S3–S6.

Asbury, W. (1995). Government's wrong war. *Addiction Counselling World, 6,* 20.

Babor, T. F. (1994). Overview: Demography, epidemiology and psychopharmacology—making sense of the connections. *Addiction, 89,* 1391–1396.

Baker, A., Kochan, N., Dixon, J., Heather, N., & Wodak, A. (1994). Controlled evaluation of a brief intervention for HIV prevention among injecting drug users not in treatment. *AIDS Care, 6,* 559–570.

Ball, J., & Ross, A. (1991). *The effectiveness of methadone maintenance treatment*. New York: Springer.

Bandura, A. (1985). The self and mechanisms of agency. In J. Suls (Ed.), *Psychological perspectives on the self* (Vol. 1, pp. 187–201). Hillsdale, NJ: Erlbaum.

Bandura, A. (1997). Health promotion from the perspective of social cognitive theory. *Psychology and Health, 6,* 1–27.

Becker, H. M. (1974). *The health belief model and personal health behavior*. Thorofare, NJ: N. J. Slack.

Bien, T., Miller, W., & Torigan, S. (1993). Brief interventions of alcohol problems: A review. *Addiction, 88,* 315–336.

Bill, W. (1957). *Alcoholics Anonymous comes of age: A brief history of AA*. New York: Alcoholics Anonymous World Services.

Brewin, C. R., & Bradley, C. (1989). Patient preferences and randomised clinical trials. *British Medical Journal, 299,* 313–315.

Budd, R., & Rollnick, S. (1996). The structure of the Readiness to Change questionnaire: a test of Prochaska and DiClemente's transtheoritical model. *British Journal of Health Psychology, 1,* 365–376.

Cannon, D. S., & Baker, T. B. (1981). Emetic and electric shock alcohol aversion therapy: Assessment of conditioning. *Journal of Consulting Clinical Psychology, 49,* 20–33.

Cannon, D. S., Baker, T. B., & Wehl, C. K. (1981). Emetic and electric shock alcohol aversion therapy: Six- and twelve-month follow-up. *Journal of Consulting Clinical Psychology, 49,* 360–368

Carroll, K. (1996). Relapse prevention as a psychosocial treatment: A review of controlled clinical trials. *Experimental and Clinical Psychopharmacology, 4,* 46–54.

Christo, G., & Franey, C. (1995). Drug users' spiritual beliefs, locus of control and the disease concept in relation to Narcotics Anonymous attendance and six month outcomes. *Drug and Alcohol Dependence, 38,* 51–56.

Cohen, P. (1994). Cocaine use in Amsterdam in non-deviant subcultures. *Addiction Research, 2*(1), 71–94.

Connors, J., Maisto, S., & Donovan, D. (1996). Conceptualisations of relapse: A summary of psychological and psychobiological models. *Addiction, 91,* 5–14.

Davidson, R. (1992). Prochaska and Di Clemente's model of change: A case study. *British Journal of Addiction, 87,* 821–822.

Davidson, R. (1996). Motivational issues in the treatment of addictive behaviour. In G. Edwards & C. Dare (Eds.), *Psychotherapy, psychological treatments and the addictions* (pp. 173–188). Cambridge, UK: Cambridge University Press.

Davidson, R. (1998). The process of change: A critical review. In W. Miller & N. Heather (Eds.), *Treating addictive behaviors*. New York: Plenum.

Davidson, R., Bunting, B., & Raistrick, D. (1989). The homogeneity of the Alcohol Dependence Syndrome: a factor analysis of the SADD questionnaire. *British Journal of Addiction, 84,* 907–915.

Davidson, R., & Raidstrick, D. (1986). The validity of the Short Alcohol Dependence Data (SADD) questionnaire: A short self report questionnaire for the assessment of alcohol dependence. *British Journal of Addiction, 81,* 217–222.

Davidson, R., Rollnick, S., & MacEwan, I. (1991) *Counselling problem drinkers*. London: Routledge.

Davies, J. B. (1992). *The myth of addiction*. Reading, MA: Harwood Academic.

DeLeon (1995). Residential therapeutic communities in the mainstream: Diversity and issues. *Journal of Psychoactive Drugs, 27,* 3–15.

Drummond, C., Cooper, T., & Glautier, S. (1990). Conditioned learning in alcohol dependence: Implications for cue exposure treatment. *British Journal of Addiction, 85,* 725–744.

Drummond, D., & Glautier, S. (1994). A controlled trial of cue-exposure treatment in alcohol dependence. *Journal of Consulting and Clinical Psychology, 62,* 809–817.

Edwards, G. (1989). As the years go rolling by: Drinking problems in the time dimension. *British Journal of Psychiatry, 154,* 18–26.

Edwards, G. (1994). Technical habiliments, science and the clinical front line. In G. Edwards & M. Lader (Eds.), *Addiction processes of change* (pp. 251–266). Oxford, UK: Oxford Medical Publications.

Edwards G. (1996). Alcoholics Anonymous as a mirror held up to nature. In G. Edwards and C. Dare (Eds.), *Psychotherapy, psychological treatments and the addictions* (pp. 220–239). Cambridge, UK: Cambridge University Press.

Edwards, G., Brown, D., Duckitt, A., Oppenheimer, E., Sheehan, M., & Taylor, C. (1987). Outcome of alcoholism: The structure of patient attribution as to what causes change. *British Journal of Addiction, 82,* 533–545.

Edwards, G., & Goss, M. (1976). Alcohol dependence syndrome: Provisional description of a clinical syndrome. *British Medical Journal, 1,* 1058–1061.

Edwards, G., Orford, S., Egert, S., Guthrie, S., Hawker, A., Hensman, C., Mitcheson, M., Oppenheimer, E., & Taylor, C. (1977). Alcoholism: A controlled trial of treatment and advice. *Journal of Studies on Alcohol, 38,* 1004–1031.

Edwards, G., & Taylor, C. (1994). A test of the matching hypothesis: Alcohol dependence, intensity of treatment and twelve month outcome. *Addiction, 89,* 553–545.

Elkins, R. (1980). Covert sensitisation treatment of alcoholism: contributions of successful conditioning to subsequent abstinence maintenance. *Addictive Behaviour, 5,* 67–89.

Elkins, R. (1991). An appraisal of chemical aversion (emetic therapy) approaches to alcoholism treatment. *Behaviour Research and Therapy, 29*, 387–413.

Erickson, P. G., & Weber, T. R. (1994). Cocaine careers, control and consequences: Results from a Canadian study. *Addiction Research, 2*(1), 37–50.

Feingold, A., & Rounsaville, B. (1995). Construct validity of the dependence syndrome as measured by *DSM-IV* for different psycho-active substances. *Addiction, 90,* 1661–1669.

Gelder, M., Gath, D., Mayou, R., & Cowen, P. (1996). *Oxford text book of psychiatry*. Oxford, UK: Oxford University Press.

Gossop, M. (1996). Cognitive and behavioural treatments for substance misuse. In G. Edwards & C. Dare (Eds.), *Psychotherapy, psychological treatments and the addictions* (pp. 139–157). Cambridge, UK: Cambridge University Press.

Gossop, M., Darke, S., Griffiths, P., Hando, J., Powis, B., Hall, W., & Strang, J. (1995). The severity of dependence scale (SDS): Psychometric properties of the SDS in English and Australian samples of heroin, cocaine and amphetamine users. *Addiction, 90,* 607–614.

Gual, A., & Colom, J. (1997). Why alcohol consumption declined in countries of southern Europe. *Addiction, 92,* S21.

Halikas, J. A., & Crosby, R. D. (1991). Measuring outcome in the treatment of cocaine abuse: The drug impairment rating scale for cocaine. *Journal of Addictive Disorders, 11,* 121–138.

Hasin, D., Li, Q., McCloud, S., & Endicott, J. (1996). Agreement between *DMS-III, DMS-IIIR, DMS-IV* and *ICD-10* alcohol diagnoses in a US community—sample of heavy drinkers. *Addiction, 91*(10), 1517–1527.

Heather, N. (1992). Addictive disorders are essentially motivational problems. *British Journal of Addiction, 87,* 828–830.

Heather, N. (1995). Interpreting the evidence of brief interventions for excessive drinkers: The need for caution. *Alcohol and Alcoholism, 30,* 287–296.

Heather, N. (1996). Waiting for a match: The future of psychosocial treatment for alcohol problems. *Addiction, 94,* 469–472.

Heather, N., & Robertson, I. (1985). *Controlled drinking*, (2nd ed.). New York: Methuen.

Heather, N., Rollnick, S., Bell, A., & Richmond, R. (1996). Effects of brief counselling among male heavy drinkers identified on general hospital wards. *Drug and Alcohol Review, 15,* 29–38.

Heather, N., & Stallard, A. (1989). Does the Marlatt model underestimate the importance of conditioned craving in the relapse process? In M. Gossop (Ed.), *Relapse and addictive behaviour* (pp. 137–151). London: Routledge.

Hester, R. K., & Miller, W. R. (1989). *Handbook of alcoholism treatment approaches: effective alternatives.* New York: Pergamon.

Hubbard, R., Marsden, M., & Rachal, V. (1989). *Drug abuse treatment. A national study of effectiveness.* Chapel Hill, NC: University of North Carolina Press.

Hunt, W. A., Barnett, W., & Branch, L. G. (1971). Relapse rates in addiction programs. *Journal of Clinical Psychology, 27,* 455–456.

Janis, I., & Mann, L. (1977). *Decision-making: A psychological analysis of conflict, choice and commitment.* New York: Free Press.

Jones, M. (1956). The concept of the therapeutic community. *American Journal of Psychiatry, 112,* 647–650.

Kanfer, F. H., & Grimm, G. L. (1980). Managing clinical change: A process model of therapy. *Behaviour Modification, 4,* 419–444.

Keene, J., & Raynor, P. (1993). Addiction as a "social weakness." The influence of client and therapist beliefs. *Addiction Research, 1,* 77–87.

Kessler, R., McGonagle, K., & Zhao, S. (1994). Life time and twelve month prevalence of *DSM-III-R* psychiatric disorders in the United States: Results from the national comorbidity survey. *Archives of General Psychiatry, 51,* 8–20.

Kleber, H. (1988). Epidemic cocaine abuse: America's present, Britain's future? *British Journal of Addiction, 83,* 1359–1371.

Klingemann, H. K. (1994). Environmental influences which promote or impede change in substance behaviour. In G. Edwards & M. Lader (Eds.), *Addiction: Processes of change* (pp. 131–162). New York: Oxford University Press.

Kok, G., De Vries, H., Mudde, A. & Stretcher, V. (1990). Planned health education and the role of self efficacy: Dutch research. *Health Education Research, 5,* 101–109.

Kuchipudi, V., Hobein, K., Fleckinger, A., & Iber, F. L. (1990). Failure of a two-hour motivational intervention to alter recurrent drinking behaviour in alcoholics with gastrointestinal disease. *Journal of Studies on Alcohol, 51,* 356–360.

Litman, G. K., Eiser, J. R., Rawson, N. S. B., & Oppenheim, A. N. (1977). Towards a typology of relapse: A preliminary report. *Drug and Alcohol Dependence, 2,* 157–162.

Litman, G. K., Eiser, J. R., Rawson, N. S. B., & Oppenheim, A. N. (1979). Towards a typology of relapse: Differences in relapse and coping behaviours between alcoholic relapsers and survivors. *Behaviour Research and Therapy, 17,* 89–94.

Lowman, C., Allen, J., & Stout, R. (1996). Replication and extension of Marlatt's taxonomy of relapse precipitants: Overview of procedures and results. *Addiction, 91,* 51–72.

Marlatt, G. A. (1978). Craving for alcohol, loss of control, and relapse: a cognitive/behavioural analysis. In P. Nathan, G. Marlatt & T. Loberg (Eds.), *Alcoholism: new directions in behavioral research and treatment* (pp. 271–314). New York: Plenum.

Marlatt, G. A., & Gordon, J. R. (1985). *Relapse prevention: maintenance strategies in the treatment of addictive behavior*. New York: Guilford.

Mayfield, D., McLeod, G., & Hall, P. (1974). The CAGE questionnaire: Validation of a new alcoholism screening instrument. *American Journal of Psychiatry, 131,* 1121–1123.

McClellan, A. T., Kushner, H., Metzger, D., Peters, R., Smith, I., Grissom, G., Pettinati, H., & Argeriou, M. (1992). The fifth edition of the addiction severity index. *Journal of Substance Abuse Treatment, 9,* 199–213.

McCusker, J., Vickers-Lahti, M., Stoddard, A., Hindin, C., Bigelow, C., Zom, M., Garfield, F., Frost, R., Love, C., & Lewis, B. (1995). The effectiveness of alternative planned durations of residential drug abuse treatment. *American Journal of Public Health, 85,* 1426–1429.

Miller, W. R. (1983). Motivational interviewing with problem drinkers. *Behavioural Psychotherapy, 11,* 147–172.

Miller, W. R., & Baca, L. M. (1993). Two year follow-up of bibliotherapy and therapist-directed controlled drinking training for problem drinkers. *Behaviour Therapy, 14,* 441–448.

Miller, W. R., & Herster, R. (1986). The effectiveness of alcoholism treatment methods. what research reveals. In W. R. Miller & R. Herster (Eds.), *Treating addictive behaviours, processes of change* (pp. 4–37). New York: Plenum.

Miller, W. R., & McCrady, B. (1993). The importance of research in Alcoholics Anonymous. In W. R. Miller & B. McCrady (Eds.), *Research on alcoholics anonymous: opportunities and alternatives* (pp. 201–222). New Brunswick, NJ: Rutgers Center of Alcohol Studies.

Miller, W. R., & Rollnick, S. (1991). *Motivational interviewing: Preparing people to change addictive behavior*. New York: Guilford.

Miller, W. R., & Sanchez, (1993). Motivating young adults for treatment and lifestyle change. In G. Howard (Ed.), *Issues in alcohol use and misuse by young adults*. Paris: University of Notre Dame Press.

Moos, R., Finney, J. & Crankite, R. (1990). *Alcoholism treatment: Contact, process and outcome*. New York: Oxford University Press.

Moser, A., & Annis, H. (1995). The role of coping in relapse crisis outcome: A prospective study of treated alcoholics. *Addiction, 91*, 1101–1114.

Noonan, W., & Moyers, T. (1997). Motivational interviewing. *Journal of Substance Misuse, 2,* 8–16.

Norton, K. (1996). How therapeutic communities work. In G. Edwards & C. Dare (Eds.), *Psychotherapy, psychological treatments and the addictions* (pp. 240–256). Cambridge, UK: Cambridge University Press.

Ogbourne, A. (1993). Assessing the effectiveness of AA: Meeting the challenges. In B. McCrady & W. R. Miller (Eds.), *Research on Alcoholics Anonymous: Opportunities and alternatives* (pp. 76–85). New Brunswick, NJ: Rutgers Center of Alcohol Studies.

Oxford, J. (1980). Understanding treatment: Controlled trials and other strategies. In G. Edwards & M. Grant (Eds.), *Alcoholism treatment in transition* (pp. 71–92). London: Groomhelm.

Orford, J. (1992). Davidson's dilemma. *British Journal of Addiction, 27,* 832–833.

Parloff, M., London, P., & Wolf, B. (1986). Individual psychotherapy and behaviour change. *Annual Review of Psychology, 37,* 321–349.

Peele, S. (1989). *Diseasing of America: Addiction treatment out of control.* Lexington, MA: Lexington Books.

Pentony, P. (1981). *Models of influence in psychotherapy*. New York: Free Press.

Pickering, H., & Stimpson, G. (1994). Prevalence and demographic factors of stimulant use. *Addiction, 89*(11), 1385–1390.

Powell, T., Bradley, B., & Gray, J. (1993). Subjective craving for opiates: Evaluation of a cue exposure protocol for use with detoxified opiate addicts. *British Journal of Clinical Psychology, 32,* 39–53.

Prochaska, J. O., & Di Clemente, C. C. (1984). *The transtheoretical approach: Crossing traditional boundaries of therapy*. New York: Dow-Jones Irwin.

Prochaska, J. O., Di Clemente, C. C., & Norcross, J. (1992). In search of how people change: Applications to addictive behaviors. *Americian Psychologist, 7,* 1102–1114.

Project MATCH Research Group (1997). Matching alcoholism treatments to client heterogeneity: Project MATCH post-treatment drinking outcomes. *Journal of Studies in Alcohol, 58,* 7–29.

Raistrick, D., & Davidson, R. (1985). *Alcoholism and drug addiction*. London: Churchill Livingstone.

Ramsay, M., & Percy, A. (1997). A national household survey of drug misuse in Britain: A decade of development. *Addiction, 92*(8), 931–938.

Rapoport, R. (1960). *The community as a doctor*. London: Tavistock.

Reinarman, C., Murphy, S., & Waldorf, D. (1994). Pharmacology is not destiny: The contingent character of cocaine abuse and addiction. *Addiction Research, 2*(1), 21–36.

Rollnick, S., Heather, N., & Bell, A. (1992). Negotiating behaviour change in medical settings: The development of brief motivational interviewing. *Journal of Mental Health, 1,* 25–37.

Rollnick, S., Heather, N., Gould, R., & Hall, W. (1992). Development of a short "readiness to change" questionnaire for use in brief opportunistic interventions among excessive drinkers. *British Journal of Addiction, 87,* 743–754.

Room, R. (1980). New curves on the course a comment on Polich, Armor and Braiker, "The Course of Alcoholism." *British Journal of Addiction, 75,* 351–360.

Royal College of Psychiatrists (1979). *Report of a special committee on alcohol and alcoholism*. London: Tavistock.

Sanchez-Craig, M. (1987). Cognitive and behavioural coping strategies in the reappraisal of stressful social situations. *Journal of Counselling Psychology, 23,* 7–12.

Saunders, J. B., Aasland, O. G., Babor, T. F., De La Fuente, J. R., & Grant, M. (1993). Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption. *Addiction, 88,* 791–804.

Saunders, J. B., Wilkinson, C., & Phillips, M. (1995). The impact of a brief motivational intervention with opiate users attending a methodine program. *Addiction, 90,* 415–424.

Schwarzer, R. (1992). Self efficacy in the adoption and maintenance of health behaviors: Theoretical approaches and a new model. In R. Schwarzer (Ed.), *Self efficacy: Thought control of action* (pp. 223–243). Washington, DC: Hemisphere.

Selzer, M. L. (1971). The Michigan alcoholism screening test: The quest for a new diagnostic instrument. *American Journal of Psychiatry, 127,* 1653–1658.

Skinner, H. A. (1982). The drug abuse screening test. *Addictive Behaviour, 7,* 363–371.

Smart, R. (1976). Outcome studies of therapeutic community and halfway house treatment for addicts. *International Journal of Addictions, 11,* 143–159.

Solomon, R. L. (1980). The opponent-process theory of acquired motivation: The costs of pleasure and the benefits of pain. *American Psychologist, 35,* 691–712.

Stephens, R., Roffman, R., & Simpson, E. (1994). Treating adult marijuana dependence: A test of the relapse prevention models. *Journal of Consulting and Clinical Psychology, 62,* 92–99.

Stockwell, T. (1992). Models of change, heavenly bodies and weltanschauugs. *British Journal of Addiction, 87,* 830–831.

Stockwell, T., Murphy, D., & Hodgson, R. (1983). The severity of alcohol dependence questionnaire: Its use, reliability and validity. *British Journal of Addiction, 78,* 145–155.

Sutherland, G., Edwards, G., Taylor, C., Phillips, G., & Gossop, M. (1988). The opiate dependence syndrome: Replication study using the SODQ in a New York clinic. *British Journal of Addiction, 83,* 755–760.

Sutton, M., & Maynard, A. (1992). *What is the size and nature of the drug problem in the UK?* York, UK: University of York.

Sutton, S. (1996). Can "stages of change" provide guidance in the treatment of addictions? A critical examination of Prochaska and DiClemente's model. In G. Edwards & C. Dare (Eds.), *Psychotherapy, psychological treatments and the addictions* (pp. 189–205). Cambridge, UK: Cambridge University Press.

Thorley, A. (1981). Longitudinal studies of drug dependence. In G. Edwards, & Busch (Eds.), *Drug problems in Britain: A review of ten years* (pp. 110–121). London: Academic Press.

Tiffany, S. T. (1990). A cognitive model of drug urges and drug-use behaviour: role of automatic and nonautomatic processes. *Psychological Review, 97,* 147–168.

Tonigan, J., Ashcroft, G. & Miller, W. R. (1995). AA group dynamics and 12-step activity. *Journal of Studies on Alcohol, 56,* 616–621.

Tuchfeld, B. (1976). *Changes in the patterns of alcohol use without the aid of formal treatment*. Research Triangle Park, NC: Research Triangle Institute.

Vaillant, G. E. (1983). *The natural history of alcoholism*. Cambridge, MA: Harvard University Press.

Vaillant, G. E. (1996). Addictions over the life course: Therapeutic implications. In G. Edwards & C. Dare (Eds.), *Psychotherapy, psychological treatments and the addictions* (pp 3–18). Cambridge, UK: Cambridge University Press.

Wilson, T. (1991). Chemical aversion conditioning in the treatment of alcoholism: Further comments. *Behaviour Research and Therapy, 29,* 415–419.

Woodhouse, L., & Davidson, R. J. (1996). *Stage matched treatment and brief intervention*. London: Alcohol Education and Research Council.

World Health Organization (1992). *The ICD-10 classification of mental and behavioral disorders: Clinical descriptions and diagnostic guidelines*. Geneva, Switzerland: Author.

Hester, R. K.
Miller, W. R.
Edwards, G.
Lader, M.
Suls, J.
Dare, C.
Heather, N.
Marlatt, G. A.
Gordon, J. R.
Gossop, M.
Nathan, P.
Loberg, T.
McCrady, B.
Howard, G.
Grant, M.
Schwarzer, R.
Busch,

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.26
Cognitive Approach to Understanding and Treating Pathological Gambling

ROBERT LADOUCEUR
Université Laval, Québec, PQ, Canada
and
MICHAEL WALKER
University of Sydney, NSW, Australia

6.26.1 INTRODUCTION 588
6.26.2 MOTIVATION TO GAMBLE 588
6.26.3 DEPARTURES FROM NORMATIVE DECISION MAKING IN GAMES OF SKILL AND GAMES OF CHANCE 590
6.26.4 DEVELOPMENT OF GAMBLING PROBLEMS 590
6.26.5 DEFINITION AND ASSESSMENT OF PATHOLOGICAL GAMBLING 592
6.26.6 PREVALENCE OF PATHOLOGICAL GAMBLING 592
6.26.7 PSYCHOLOGICAL TREATMENT OF PATHOLOGICAL GAMBLERS 593
  6.26.7.1 Changing Erroneous Beliefs Concerning Randomness 594
  6.26.7.2 Problem Solving Training 596
  6.26.7.3 Social Skills Training 596
  6.26.7.4 Relapse Prevention 596
6.26.8 EFFICACY OF A COGNITIVE/BEHAVIORAL TREATMENT FOR PATHOLOGICAL GAMBLERS 597
  6.26.8.1 Case History 597
    6.26.8.1.1 Identification and correction of faulty cognitions toward gambling 598
    6.26.8.1.2 Problem solving and social skills training 598
    6.26.8.1.3 Relapse prevention 598
6.26.9 EVALUATION OF A COGNITIVE TREATMENT FOR PATHOLOGICAL GAMBLING 598
6.26.10 COMMON DIFFICULTIES IN THE TREATMENT OF PATHOLOGICAL GAMBLERS 599
6.26.11 CONCLUDING REMARKS 599
6.26.12 REFERENCES 600

6.26.1 INTRODUCTION

The twentieth century has been described as the "hinge of history." This description follows from the wide range of indications of social, economic, and environmental measures showing catastrophic change. One of these indications is the general public's level of involvement in gambling activities. Rowntree (1941) compared the standard of living of the residents of New York City in 1899 and 1936. In his analysis of leisure in everyday life, he reached the conclusion that gambling had grown enormously over that period of time. It is well-documented that, since 1936, gambling has continued its rapid growth not only in the USA, but throughout the world (Frey & Eadington, 1984; McMillen, 1996). In explaining the upsurge in gambling, the British Royal Commission on Lotteries and Betting (1933, p. 60) stated that, "One of the main causes, perhaps the most potent in the growth of gambling, has been the increased facilities for organized gambling." Similar views have repeatedly been expressed by contemporary observers of gambling (Connor, 1983; Cornish, 1978; Dickerson, 1984: Orford, 1985). However, the claim that the rapid growth in gambling is attributable to the legalization and the accessibility of gambling outlets may be considered false, or at least incomplete. Betting shops, gambling machines, lotteries, and casinos would be useless if ordinary people did not want to gamble. Even in societies where gambling has been prohibited, gambling games have nevertheless flourished (Dixon, 1996).

Prior to discussing why people want to gamble, it is important to specify what we are referring to when using the word "gambling." Research on gambling does not incorporate all risk-taking behavior, but only a limited range of such behavior. The essential nature of gambling is that money (or its equivalent) is risked on the uncertain outcome of an event, subject to certain conditions: (i) gambling occurs in a group context whereby after costs, taxes, and profits, the money wagered by the losers is redistributed to the winners; (ii) the redistribution of money is independent of any other commercial enterprise related to the gambling event. This definition excludes insurance such as life or property insurance. This definition of gambling is similar to those put forth elsewhere (Perkins, 1950).

It should be noted that all definitions of gambling appear to differ with respect to what is included or excluded. This problem can be overcome if the activities to be included are listed explicitly. In this context, three broad categories of gambling have been identified: betting, gaming, and lotteries. Betting refers to staking money on the outcome of an uncertain or future event; gaming refers to playing for money in a game of chance; and lotteries refers to the distribution of prizes by drawing lots. Although these three kinds of gambling have much in common at the structural level, they are completely different sociologically.

Claiming that people want to gamble does not advance our knowledge of gambling to any great extent, but immediately raises the question as to why people want to gamble. Understanding the motivation to gamble would appear to be central to understanding why people may gamble excessively, and can be expected to have an important bearing on how gambling-related suffering may be alleviated, diminished, or avoided.

This chapter argues that the motivation to gamble is seen as the acquisition of wealth and that the real problems in explaining gambling concern why it is that gamblers believe that money can be won. Because the belief of gambers that money can be won is erroneous, it follows that therapeutic methods which change the false beliefs of the gambler are more likely to be effective. This proposition is examined in detail.

6.26.2 MOTIVATION TO GAMBLE

Gambling explicitly involves the attempt to win money by staking money on an uncertain event. As a starting point in the attempt to understand the motivation to gamble, the acquisition of wealth can be assumed to be the prime motivation. The problem with this assumption is that all legalized forms of gambling are constructed so that the expected return is less than the sum wagered. For example, a roulette wheel with one zero takes in, on average, 1/37 of the money staked. Totalizators typically take in approximately 20% of the money wagered in racing (Ladouceur, Giroux, & Jacques, 1998) and lotteries typically take in approximately 40% of the revenue from ticket sales. These percentages vary from place to place and according to the structure of the distribution of prizes or returns, but in all cases the expected return on money invested constitutes a loss for the gambler. Thus, if the acquisition of wealth is the individual's goal, rational economic considerations would lead people to avoid gambling. This is the principal paradox of gambling: people, in attempting to gain wealth, engage in an activity which is expected to decrease wealth.

The gambling paradox can be resolved in two different ways: (i) accepting that the acquisition of wealth is the motivation, but the gambler misjudges the chance of winning; or (ii) rejecting the acquisition of wealth as the sole or central motivation involved.

The majority of theories of gambling behavior reject the acquisition of wealth as being the fundamental motivation for gambling behavior. However, cognitive theories of gambling assume that the acquisition of wealth is the primary motivation involved, and that people do not behave as they normally would with respect to that motivation. According to cognitive theories, gamblers hope to win money or believe that they will win money. Why people should hope to win or expect to win in the face of the adverse odds involved is the central concern of such theories.

By contrast, other theories of gambling assume that winning money is not the principal motivation for gambling. One cluster of such theories assumes that it is the amusement and excitement (the change in arousal level) that motivates gambling behavior. In the behaviorist form of these theories, the changes in arousal reinforce the gambling behavior (Dickerson & Adcock; 1987; McConaghy, 1980), whereas in more purposive theories, gambling has the function of changing mood (Brown, 1996) as when the excitement of gambling overcomes boredom. In some cognitive behavioral explanations, arousal retains a dominant role but must be coupled with appropriate cognitions (Sharpe & Tarrier, 1993). A second cluster of theories places more emphasis on personality dimensions and, more specifically, on impulsiveness (Blaszczynski & McConaghy, 1994). It is assumed that risk-oriented individuals are attracted to gambling and that problems arising from excessive gambling can be attributed to impulsiveness. Yet other explanations see the gambling as largely irrelevant: gambling is simply one possible means of escaping or avoiding stresses and associated anxiety elsewhere in the individual's life. All such theories assume that the principal reward for gambling is something other than money. Some support for this view comes from the gambling industry which regards gambling as a leisure activity in which the money expended by the gambler buys the gambling product. The product is understood to be the amusement and excitement of the gambling venture. Thus, gambling is viewed as a desirable activity in itself, similar to eating in a restaurant or playing a round of golf. The money spent by the gambler pays for the enjoyment received.

It might seem that these alternative explanations for gambling have a certain face validity. If, after all, it is unreasonable to expect to win money by gambling, then surely some other factor must be involved. When gamblers are questioned about why they gamble, the majority of answers concern amusement, excitement, and relief of boredom. By contrast, only a minority of individuals involved in betting and gaming state that they gamble in order to win money. The important question is whether or not the opinions of gamblers as to why they gamble should be accepted indiscriminately. A certain amount of validation of the data derived from other sources is a minimum requirement. Consider, for example, the claim that gambling is exciting and that the rewarding value of the excitement is the main reason why most people gamble. If this claim is true, then several implications follow:

(i) Behavioural observation should support the self-report data. Gamblers should appear to be excited when gambling. Excitement should be apparent on faces (laughing and smiling), in exclamations of delight, and in the general bodily tension and alertness.

(ii) Physiological measures should support the self-report data. Gamblers should exhibit raised levels of heart rate, blood pressure, and palmar sweating. Neurophysiological measures should also suggest general arousal.

(iii) Indications of excitement should not be highly correlated with winning. If excitement always accompanies winning and no other events within the gambling cycle, then it is not possible to know whether it is the winning or the excitement which maintains the behavior.

(iv) If an explanation of gambling in terms of the reward value of excitement is to be generalized, then the indicators of excitement should be present for all forms of gambling.

Surprisingly, there is very little evidence supporting any of these specific implications. Here, we briefly review the evidence.

Excitement is most easily observed in certain forms of betting and gaming, and among the winners of lottery prizes. Thus, the noise of the crowd at a race track reaches a crescendo as the leading horses reach the winning post. Similarly, certain casino games, such as craps and two-up, have been described as suitable for the extrovert in view of the manifest excitement of the players (Allcock & Dickerson, 1986). However, such observable excitement is not sufficient evidence. Observable excitement is less in off-course betting shops and may be largely absent among home gamblers with phone accounts. Furthermore, certain casino games, such as blackjack and pai gow, are traditionally played with a minimum display of emotion. In poker, the ideal involves an absence of genuine emotion. Thus, there are many examples of gambling situations in which no excitement is expressed.

Different measures of physiological arousal may be used in gambling research. Increases in heart rate have been reported for blackjack (Anderson & Brown, 1984) and changes in skin conductivity for slot machines (Sharpe, Tarrier, Schotte, & Spence, 1995) when players begin a session. It is likely that arousal increases at the beginning of all gambling sessions. However, excitement prior to gambling cannot be a reward for gambling. More importantly, arousal must be demonstrated to correlate with events within the gambling session. In a detailed study of slot machine gambling through time, Dickerson (1993) did not find any correlation between player arousal and the events within a session. Although variations in arousal have been demonstrated for certain gambling-related events, the failure by Dickerson to show that arousal is correlated with machine events in slot machines is an important obstacle for an excitement-based explanation of gambling. Furthermore, the low numbers of positive reports of arousal in relation to the wide range of gambling, casts doubt on the ability of excitement-based theories to have wide explanatory power. However, as discussed in detail later, arousal has been found to be associated with erroneous perceptions that create an illusion of control and result in the overestimation of the probabilities of winning. Interestingly, arousal decreased when erroneous perceptions were corrected (Giroux, Ladouceur, & Jacques, 1998).

By contrast with arousal and personality theories, cognitive theories of gambling assume that the acquisition of wealth is the primary motivation involved, and that people do not behave as they normally would with respect to that motivation. According to cognitive theories, gamblers hope to win money or believe that they will win money. Why people should hope to win or expect to win in the face of the adverse odds involved is the central concern of such theories.

6.26.3 DEPARTURES FROM NORMATIVE DECISION MAKING IN GAMES OF SKILL AND GAMES OF CHANCE

Some gambling games allow the players a range of decisions which can affect the outcomes. In such games, players may depart from optimal play by choosing alternatives with lower expected values. Skill then refers to the extent to which a player's strategy of choices approaches the optimal strategy. Gambling games involving skill include blackjack, poker, and sports betting. In all of these games, systematic departures from optimal play have been reported.

In blackjack, with optimal play, the expected value of casino blackjack is approximately -0.7% with variations from that figure depending on the specific rules of play that apply. Nevertheless, most players do not approach this figure in their play (Griffin, 1987; Wagenaar, 1988; Walker, Sturevska, & Turpie, 1995). Analysis of the errors in play shows that in general the players adopt strategies which are too conservative. Players tend to sit when they should hit, avoid splitting pairs when it is desirable, and take out insurance when it is unnecessary (Wagenaar, 1988; Walker, 1995).

Interestingly, the departures from normative play in poker tend towards risk rather than avoidance of risk. For example, players in five card stud tend to stay in the pot with a small pair when they should fold (Yardley, 1957). In draw poker, players will draw to hands with very small probabilities of success, and professional poker players prefer games in which the action is loose, that is, risky (Hayano, 1977; Yardley, 1959).

There are many systems for betting on horse races and, as the variations between the systems are great (Allcock, 1987; Beyer, 1993; Drapkin & Forsyth, 1987; Scott, 1982) and no more than one can be accurate, it follows that the majority of systems yield less than optimal decisions. More generally, betting patterns have been observed to vary across a race meeting with betting on the last race at the meeting being more risk oriented than betting on earlier races (Bird & McCrae, 1985). Recently, Ladouceur, Giroux and Jacques (1998) showed that regular punters, defining themselves as experts in horse races, could not provide a better rate of return than a random selection of horses.

In games of chance all alternatives have equal expectations, and thus there is no optimal method of play. Well known games of chance include lotteries, roulette, and slot machines. Since no optimal strategy exists for games of chance, players would be expected to expend no effort in trying to choose the best action—the most likely winning ticket in lotteries, the most likely winning number in roulette, or the machine most likely to pay out in slot machines. Nevertheless, players can be observed expending considerable effort in making these decisions (Griffiths, 1994; Wagenaar, 1988; Walker, 1992). To the extent that players believe that some alternatives are more likely than others in games of chance, their approach to betting can be regarded as not normative.

6.26.4 DEVELOPMENT OF GAMBLING PROBLEMS

Although a wide range of gambling-related problems have been documented (Walker, 1992), the classification of such problems has not been adequately developed. Classification of gambling problems can proceed from different perspectives and be based on different criteria. Most commonly, gambling-related problems are classified by the area in the gambler's life that is affected. Thus, Dickerson et al. (1995) divided gambling problems into those associated with the individual, the family, financial status, employment, and criminal activity. Lorenz and Shuttleworth (1983) divided the problems into personal, relationship, and financial. Similarly, Custer and Milt (1985) divided the problems into gambling, alienation, marital problems, boredom, legal problems, indebtedness, needs, and goalessness. Categorization of problems in this way has value at the level of assessment, but does not clarify the nature or source of the problems. Although overlaps must exist, it remains possible for a new researcher to divide the gambling-related problems differently into another, possibly equally useful, set based on areas affected.

An alternative approach, which places more emphasis on the genesis of the problems, assumes that the main cause of the problems is persistence with gambling despite the losses. Cognitive theories seek to explain why the gambler may persist with gambling until the losses become excessive. The next step in understanding problem gambling, which ultimately becomes labeled "pathological," involves analyzing the consequences of extreme persistence in the face of large losses. The central consequence, and possibly the core factor in causing gambling problems, is the financial loss. Although it may seem obvious that financial loss is a fundamental aspect of gambling problems, this perspective is sometimes not given the emphasis that would seem appropriate. For example, only four of the 10 criteria defining pathological gambling in the *Diagnostic and statistical manual of mental disorders* (4th. ed.; *DSM-IV*; American Psychiatric Association [APA], 1994) explicitly refer to the loss of money and the problems caused thereby. If the financial cost of gambling is emphasized, then many of the criteria for identifying pathological gambling can be understood as consequences of this common cause. Walker (1992), in his description of a sociocognitive theory of gambling, shows how the false beliefs of gamblers can lead to chasing losses, changes in mood, withdrawal and secretiveness, deceitfulness, irritation and anger, and foolish financial transactions. These changes at the individual level, coupled with the large loss in income, would be expected to impact on the family life, employment, and social life of the gambler.

Persistence with gambling causes not only financial loss, but also absorbs large amounts of the gambler's time. The time away can be expected to impact heavily on the family and on employment. Thus, to the group of spouses that includes "golfing widows" and "fishing widows," should be added "gambling widows." The impact on the family of excessive involvement in leisure activities or employment is common across activities and may be a cause of family argument and distress. However, it is likely that time away is for most gamblers and their families a minor factor compared to the financial losses suffered by the persistent gambler.

Apart from the loss of time and money, there is one further area of loss that is more difficult to quantify. Gambling can be characterized as a background of failure broken only by occasional success. According to cognitive accounts of persistence with gambling, the gambler holds a set of erroneous beliefs about the nature of gambling and the role of the gambler in relation to the gambling. Persistence with gambling increases the likelihood of overall loss. Thus the gambler is continually engaged in searching for explanations that maintain the core beliefs. The mass of evidence suggesting that the gambler's beliefs are erroneous is a continuing stress that can be expected to cause loss of self-esteem and, ultimately, depression. It would not be surprising to find that some gamblers show evidence of this stress in aspects of their physiology and biology (Blaszczynski, Winter, & McConaghy, 1986; Carlton & Manowitz, 1987; Sharpe et al., 1995).

One problem that general theories of gambling must confront involves specifying why only a minority of regular gamblers suffer problems to the extent that they ultimately seek counseling and treatment. Individual differences in persistence with gambling have been explained in terms of personality differences (Zuckerman, 1979), biological differences (Jacobs, 1986), and learning differences (Dickerson, 1984). However, perhaps the most valuable insights concerning individual differences in gambling have been provided by Orford (1985) and Oldman (1978). Orford asked the important question as to why not all gamblers continue gambling until their money is exhausted. If gambling is intrinsically rewarding, progression to gambling problems and pathology would be expected. Yet the majority of gamblers control their gambling sufficiently to avoid the potential problems. Thus, inability to exercise control over the desire to gamble is an important aspect of the genesis of gambling problems. Orford suggests that gambling problems may involve the conjunction of excessive appetites, incomplete socialization of control over appetites, and the availability of opportunities to gamble. Evidence for such a view of gamblers comes from observational studies of regular gamblers that show that most are able to modify their approaches to gambling when demanded by changed financial circumstances (Rosecrance, 1986). Oldman (1978) took the argument one step further by pointing out that gambling problems were a natural consequence of persistence with gambling. The label "pathological gambling" may thus be a means by which society negotiates the counseling and treatment of gamblers who are sufficiently unlucky that they lose too much (Oldman, 1978; Walker, 1995).

6.26.5 DEFINITION AND ASSESSMENT OF PATHOLOGICAL GAMBLING

Pathological gambling was officially recognized in 1980 with the publication of *DSM-III* (APA, 1980), and was classified as an impulse control disorder. The *DSM-IV* (APA, 1994) defined 10 criteria reflecting different aspects of pathological gambling. To assign the diagnostic of pathological gambling, the individual must meet at least five of these criteria. As mentioned above, if most individuals gamble without it being a problem, some will eventually become overwhelmed by the desire to gamble, will gamble more than they planned, and will eventually spend more money than they can afford to lose. Pathological gambling is characterized by a loss of control over gambling, lies about the extent of involvement with gambling, family and job disruption, stealing money, and continuous chasing of losses. From a clinical perspective, two elements are the most representative of a pathological gambler: continuous or obsessional chasing of losses; and family, job, and social disruption caused by gambling.

In our clinic (R. L.), we have recently adopted a multistep evaluation procedure. When an individual calls for help for a gambling problem, we return the call within 24 hours. During this first call, we ask the gambler to describe his main complaints and then administer the South Oaks Gambling Screen (SOGS). The SOGS is a 20 item instrument used in many prevalence studies around the world to identify the number of pathological or problem gamblers in the general population. It has been translated into many languages including French, Chinese, German, and Spanish. If preliminary data collected during the phone call suggests that the individual is a pathological gambler, a formal semistructured interview is immediately scheduled in order to identify the nature and history of the problem. Before starting this interview, the individual will be asked to complete questionnaires evaluating the following areas: depression, anxiety beliefs about gambling (Ladouceur, Arsenault, Dubé, Jacques, & Freeston, 1997), superstitous behavior (Ladouceur, Giroux, & Jacques, 1988), and problem solving abilities. The interview is divided into two sections and covers the following aspects: history of the gambling activities; motivation for the consultation; first contact with gambling; first problems with gambling; familial, professional, and marital problems; money lost and criteria of pathological gambling. The last step in our evaluation is the second administration of the *DSM-IV* in order for the therapist to confirm the diagnostic of pathological gambling. If the individual is diagnosed as a pathological gambler, treatment is offered and usually starts the next session after the evaluation procedure.

This procedure has many advantages. First, the gambler is contacted within the 24 hours after his call. Second, the telephone interview focuses on the gambling problem (description of the main complaints and the administration of the SOGS), thus setting the purpose of the consultation and the subject matter of further treatment, if necessary. Third, this assessment procedure provides relevant information about the different aspects of the gambler's life problems (family, job, social, legal, financial, etc.). Fourth, this procedure will provide useful data for the validation of various instruments, and on the characteristics of the patients who refuse or drop out of treatment. As will be discussed later, adherence to treatment is a major concern for professionals working with pathological gamblers. We need to identify the characteristics and the reasons of individuals who refuse treatment, drop out, or simply do not show up for the first session.

6.26.6 PREVALENCE OF PATHOLOGICAL GAMBLING

The prevalence of pathological gambling is the percentage of the members of a society at a given point in time whose gambling is pathological according to some agreed criterion. Nearly all studies of the prevalence of pathological gambling have used one or other of the two measures described: the *DSM-IV* criteria and the SOGS. Most of the research has been conducted in the USA by Rachael Volberg on a state by state basis (Volberg, 1996), although substantial numbers of studies have been conducted outside the USA, such as in Canada (Ladouceur, 1996), Spain (Becona, 1996), Australia (Dickerson, Baron, Hong, & Cottrell, 1996), and New Zealand (Abbott & Volberg, 1996).

Comparison of the data from the studies reviewed by the authors listed above is made difficult by the fact that the details of the designs varied considerably across research groups. Volberg and Ladouceur used telephone surveys, whereas the surveys in Spain, Australia, and New Zealand used door knock surveys. Becona used *DSM-III-R* criteria whereas other researchers based their conclusion on the SOGS. Although most research has adopted a cut-off of five on the SOGS for the identification of pathological gamblers, Dickerson has argued that the cut-off should be higher. Finally, the SOGS itself may not be a sufficiently accurate indicator of pathological gambling for use in the prevalence research (Walker & Dickerson, 1996). Thus, current estimates of the prevalence of pathological gambling must be treated with a degree of caution. Nevertheless, the available evidence suggests that the occurrence of pathological gambling varies from country to country. Walker and Dickerson note that the prevalence figures for pathological gambling are correlated with the average expenditure on gambling across countries. Thus, in countries in which a higher percentage of personally expendable income is spent on gambling, there is a higher reported prevalence of pathological gambling (Table 1).

Evidence of this kind strengthens the argument that gambling-related problems are primarily associated with the loss of excessive amounts of money. Theories of pathological gambling must explain why the gambler persists in gambling despite such losses. Cognitive theories assume that it is erroneous beliefs and inferences about gambling and the likelihood of favorable outcomes which maintains the behavior in the face of serious monetary losses. It follows that a cognitive approach to therapy is the one that will attempt to correct the erroneous thinking involved.

6.26.7 PSYCHOLOGICAL TREATMENT OF PATHOLOGICAL GAMBLERS

The central assumption of cognitive approaches to treatment is that the pathological gambler continues to gamble because they maintain an unrealistic hope that they will recover their losses if they persevere with the gambling. It is assumed that their erroneous beliefs about gambling, about the nature of predictability, and about their own special skills and knowledge in relation to predicting gambling events, conspire to maintain the gambling far beyond any reasonable limits. It follows that any correction of erroneous perceptions weakens the belief that losses can be recouped. However, alternative approaches to the treatment of pathological gambling are not based on these assumptions. There are in fact two main alternatives to the erroneous thinking approach: (i) the behaviorist orientation, based on the use of extinction processes; and (ii) the problem solving approach where the gambler is counseled in methods appropriate to solving problems causing the gambling. Both orientations assume that the central motivation is not avarice, but some other factor altogether. These two approaches differ in the level at which they assume the relevant processes are operating: molecular and below awareness for the behavioristic approach; conscious planned processes in the problem solving approach.

The behavioral approach assumes that gambling-based arousal is the central factor in the reinforcement process. It is assumed that the increase in arousal associated with gambling is positively reinforcing. With repeated gambling a whole range of associated stimuli become conditioned reinforcers. Approach to, and participation in, gambling in regular gamblers are triggered by a wide range of environmental features: the sight of the newspaper, driving the car, leaving work, the sight of money in the wallet, and so on. The wide range of factors associated with gambling is often referred to, at the macro level, as preoccupation with gambling. Treatment makes use of established learning theory principles involving extinction of the association between arousal and central conditioned elicitors. The most effective specific treatment program using the behavioral approach appears to be imaginal desensitization (McConaghy,

Table 1 SOGS scores and gambling expenditure across countries.

Country	*SOGS scores (% scoring 5+)*	*Expenditure on gambling (% of personal consumption)*[c]
Australia	7.1	1.6
New Zealand	2.7	0.9
Spain	1.5[a]	0.7
Canada	1.2[b]	0.5

[a] Means of estimates provided by Becona. [b] Ladouceur's data only. [c] Haig (1985).

Armstrong, Blaszczynski, & Allcock, 1983; Blaszczynski, McConaghy, & Frankova, 1991). Imaginal desensitization involves creating a list of specific gambling triggers for the individual. The gambler is taught standard muscle relaxation techniques. Finally, the gambler is asked to imagine the trigger situations one by one, each time accompanied by the relaxation procedure. In this way, the association of arousal with each of the triggers is extinguished. McConaghy, Blaszczynski and co-workers treated 60 pathological gamblers with the imaginal desensitization procedure, and a further 60 pathological gamblers by a range of other techniques (aversion therapy, relaxation therapy, and cue exposure). Of the 60 pathological gamblers treated by imaginal desensitization, only 33 could be followed up two or more years later. However, of these 33, 26 (79%) were gambling in a controlled way or not at all. In the control group, only 16 (53%) of the 30 followed up had achieved control or abstinence.

The problem solving approach refers to a general orientation towards treatment rather than a coherent orientation towards the causes of pathological gambling. The actual causes of the gambling are assumed to be a variety of factors determined primarily according to the implicit theories of the counselor. According to addiction counselors, the gambler is driven by intense urges to gamble repetitively and in a maladaptive way. Thus the problem solving approach is oriented to increasing the gamblers ability to cope with the urges and to provide action alternatives that can be used to redirect the energies involved into alternative less hazardous directions. Many counselors believe that the gambling is an escape from crises and dilemmas which cause the individual great anxiety. Thus the problem solving approach is oriented to dealing with these other issues for which the gambling is an escape. There is a considerable body of research suggesting that counseling approaches yield improvement within a before-and-after design (Walker, 1992), little research has been done in which alternative therapies are compared or control groups used. The evidence available suggests that significant improvement may occur in a group of untreated pathological gamblers in a six month period (Echeburua, Baez, & Fernandez-Montalvo, 1996). Thus the necessity for controlled trials with longer-term follow ups (two years or more, ideally) is apparent.

The cognitive approach to the treatment of gambling is based on experimental work demonstrating a wide range of cognitive errors made by gamblers in relation to gambling (Ladouceur & Walker, 1996). The research of Ladouceur and his colleagues in Canada suggests that the core cognitive error lies in the gambler's notions concerning randomness. The illusion of control and belief in the predictability of events that depend on the misconception of randomness are assumed to lead ultimately to the bizarre beliefs documented by Coventry (1997), Walker (1992), and others. Gamblers try to control and predict outcomes of games that are objectively uncontrollable. The illusion of control motivates them to elaborate strategies to win more money. However, all gambling is based on the inherent unpredictability of gambling events either through inadequate information, as in sports betting, or through the incorporation of randomness as in slot machines, casino games, and lotteries. It follows that if the erroneous perceptions and understanding of randomness in the gambler can be corrected, then the motivation to gamble should decrease dramatically. Our treatment programs have focused on erroneous cognitions concerning randomness as the most important targets for change. The whole range of erroneous cognitions, sometimes labeled "irrational thinking" and which constitute the illusion of control, are also important targets for change. Since persistent gambling induces a range of other problems these are also treated. The loss of money is associated with many of these problems. Training in problem solving techniques appears to be appropriate and necessary in some cases. Also, many gamblers often lie and isolate themselves in order to gamble, and so social skills training may be necessary to help the client to reestablish adequate social relationships.

In order to evaluate the effectiveness of cognitive theory for pathological gambling, a controlled study has been undertaken at Laval University. Four components were included in the therapy: (i) cognitive correction; (ii) problem solving training; (iii) social skills training; and (iv) relapse prevention. These components are now described and are followed by a case study which illustrates how these elements have been integrated in the treatment of a pathological gambler. Treatment is administered on an individual basis with one 60–90 minute session per week, over a period of 12 weeks.

6.26.7.1 Changing Erroneous Beliefs Concerning Randomness

Correction of the misunderstanding of randomness is the first goal of the treatment. Different approaches to this task are possible according to the characteristics of the individual. However, only the general guidelines will be described here. First of all, most gamblers are not aware of their erroneous perceptions of randomness. They spontaneously deny that they maintain such false conceptions. Increasing awareness of the actual way in which gambling events occur is a first step in enabling gamblers to recognize their misconceptions concerning the predictability of the game (Ladouceur & Dubé, 1997; Ladouceur, Paquet, Lachance, & Dubé, 1996). The patient will be asked to describe the evolution of their gambling habits, how they were betting at first, the changes in their betting as they became more familiar with the games, and to what extent they feel they have some potential control over certain games. By asking about the way to get an edge in the particular form of gambling involved, the gambler is invited to expose some of their errors in thinking. Inevitably, the gambler will describe strategies of play which assume that there is more predictability present than is in fact the case (Ladouceur, Paquet, & Dubé, 1996).

The therapist will ask the client to describe what they are saying to themselves when they gamble. In doing so, the therapist may ask the patient to answer the following questions: Why did you place one particular bet instead of another? How did you determine this bet? Are you trying to control the game by avoiding certain bets? Would you agree to switch poker machines in the middle of a session when the machine you are playing has not paid out for a long time? Would you agree to bet on any number at the roulette table? How did you pick the numbers that you did on the lottery ticket you bought this week? The main goal of these questions is to clarify the fact that the gambler is using some sort of information to predict an event which is independent of all other events and essentially unpredictable beyond its chance probability.

The therapist makes a distinction between gambling that involves events that are inherently unpredictable because of inadequate information (such as horse racing) and events which are random (such as lotteries or slot machines). The therapist then shows that inherently unpredictable events are essentially the same as random events. Then, explanation of the concept of randomness follows, focusing on the most crucial element: each turn is an independent event (Gaboury & Ladouceur, 1989; Walker, 1992). Since each event is independent, there can be no influence from one event to the next and no predictability across events. Furthermore, as the events cannot be influenced legally, there can be no strategies to control the outcomes of the game. The therapist will focus on the strategies used by the gambler in their preferred game, and draw attention to the specific strategies that imply a sequential relationship between outcomes or that the probabilities of specific outcomes can be altered. The therapist should focus attention on any verbalizations made by the gambler that suggest the existence of links between the outcomes of the games. We often tape record these sessions in order to capture and analyze all of the patient's verbalizations suggesting links between the outcomes of the games. This recorded material is used to increase the patient's awareness, and later on to correct faulty perceptions.

Another useful way to illustrate the erroneous links inferred between events is by tossing a coin with the patient. First the patient is asked to predict whether the next event will be "heads" or "tails" and to explain and justify his choice (Ladouceur & Dubé, 1997). Most patients will say that their choice is based on a 50/50 probability of each possible outcome, which is indeed correct. This exercise is carried out a few times in order to demonstrate that predicting heads or tails is such a simple game, and that all of the outcomes of the toss are independent. Gamblers will generally agree with the therapist. Then, in order to demonstrate the presence of erroneous cognitions during a gambling session, a simple test is performed.

The therapist writes down six consecutive outcomes of heads and covers them with a piece of paper. Once again, the patient is asked to predict the outcome of the next toss. After their choice has been made, the six previous outcomes are revealed and the patient is asked if they would like to change their prediction before the coin is tossed again. Whether they change their prediction or not, patients will examine this series of outcomes. The therapist then points out the fact, that although the gambler knew that every outcome of a coin toss is independent, they spontaneously examined past outcomes even though these are completely irrelevant. This simple behavioral exercise has proven to be very helpful for demonstrating to the patient how this tendency to link irrelevant events is very powerful. (Ladouceur & Walker, 1996, give an extensive discussion of this phenomenon.)

The notion of randomness is then explained in detail, illustrated by examples of the games played by the patient. The fundamental error is in believing that information may be used to establish links between outcomes and then used to place winning bets. Gamblers will erroneously perceive some elements of skill that, if used appropriately, enhance their probability of winning. This illusion of control explains why people bet more money as they become more familiar with a game, firmly believing that they have developed some skills that can be used profitably. During this first stage of treatment, it is brought to the patient's attention that in many studies, conducted with different games in different countries, more than 75% of the players' verbalizations are erroneous (Ladouceur & Walker, 1996). Literature detailing the frequent misconceptions of gamblers is also provided to the client.

The pathological gambler is then asked to identify their own erroneous perceptions. This is achieved through a variety of methods such as: asking the patients to describe what they are saying to themselves when gambling (see example below); simulating a game and having the gambler describe how they proceed in choosing their bets; and asking the patient to imagine a gambling session and describe out loud what they are thinking, using the "thinking out loud method" (Gaboury & Ladouceur, 1989). These sessions are also usually recorded. Some examples of erroneous cognition are: "If I lose four times in a row, I will win the next time," or, after one or two wins, "I am really getting better at this game, I know how to bet." We will often start by replaying the tape obtained within an earlier treatment session, to demonstrate to the clients their false beliefs about the notion of randomness. The clients will also be asked to listen to the tape at home and identify every erroneous perception, or irrational statement. It is important to note that the basic cognitive error involves the linking of independent events.

Finally, the last phase involves the correction of inadequate verbalizations and faulty beliefs. Patients will monitor their own verbalizations when they are thinking about gambling, when they have the urge to gamble, or when they are gambling if they have not yet stopped. The patients will (i) identify erroneous perceptions, (ii) evaluate and challenge the adequacy of these perceptions, (iii) replace these inadequate cognitions by adequate verbalizations, and (iv) assess the strength of their belief in the new cognitions. The recording of their own verbalizations made during a simulated gambling session may be used during this corrective phase by asking the patient to reformulate erroneous perceptions by an adequate verbalization. The success of this phase is normally required before addressing further issues. Our clinical experience and empirical data (see below) support the fact that other therapeutic components are likely to be unsuccessful unless the gamblers have developed an adequate conception of the notion of randomness and can apply this notion to their own behavior. If the illusion of predictability in gambling events is allowed to remain, relapse is likely to occur.

6.26.7.2 Problem Solving Training

Problem solving training is a second therapeutic intervention, used if the gambler shows poor problem solving skills when coping with excessive gambling activities. Problem solving training becomes an integral aspect of the treatment of the pathological gambler if the therapist and the patient identify that additional skills are needed to solve the actual problems related to excessive gambling. The therapist will introduce a problem solving technique (Goldfried & Davison, 1976) that involves the following five steps: (i) defining the problem, (ii) collecting information about the problem, (iii) generating different solutions, (iv) listing advantages and disadvantages for each solution, and (v) implementing and evaluating the solution. The patient learns how to cope with the difficulties related to gambling. For example, in order to have better control over spending, they may decide to pay their bills immediately after they are issued, create a budget, and carry only the amount of money they need.

6.26.7.3 Social Skills Training

If necessary, gamblers are also given social skills training in order to improve their social competence. The potential link between poor social skills and gambling activities is discussed with the patients. It is important to recognize that pathological gamblers may need more than usual social skills to deal with their relationship conflicts. For example, some gamblers needed assertiveness training in order to increase their ability to refuse invitations to gamble with friends. Role playing can be used to improve communication skills. This training focuses on the negative consequences of gambling and how the lack of good social skills is a contributing factor.

6.26.7.4 Relapse Prevention

Relapse prevention is based on and adapted from the relapse prevention model developed by Marlatt (1985) for alcoholics. The possibility of relapsing is always discussed with the participants. They learn to become aware of high risk situations (present or past) and the reasons why people return to gambling. Patients will describe their relapses, identify high-risk situations, and develop specific ways to deal with the situations. For example, carrying cash (as on pay days), stress, loneliness, and lack of social activities are common high-risk situations. Each situation is discussed in terms of the erroneous perceptions associated with gambling. Specific attention is drawn to the debt and to the likelihood that the debt will be increased if gambling is resumed.

6.26.8 EFFICACY OF A COGNITIVE/ BEHAVIORAL TREATMENT FOR PATHOLOGICAL GAMBLERS

Single case experimental studies have been conducted over several years to evaluate the cognitive/behavioral approach to treatment for adults and adolescents suffering from pathological gambling. The results were quite encouraging (Bujold, Ladouceur, Sylvain, & Boisvert, 1994; Ladouceur, Boisvert, & Dumont, 1994). The following control group comparison study was conducted to further assess the treatment's efficacy (Sylvain, Ladouceur, & Boisvert, 1997). Twenty-nine pathological gamblers participated in the study. The majority of gamblers were video poker players, whereas others gambled on horse races or casino games. Subjects were randomly assigned to a treatment or a waiting list control group.

The following dependent variables were used:

(i) *DSM-III-R*;

(ii) SOGS;

(iii) perception of control over the gambling problem rated on a scale of 0–10;

(iv) desire to gamble indicated on a 0–10 scale;

(v) self-efficacy perception evaluating their belief that they could refrain from gambling in high risk situations;

(vi) frequency of gambling in terms of the number of gambling sessions, the number of hours spent gambling, and the total amount of money spent on gambling during the previous week.

Results showed that treated subjects improved significantly compared to the control group. Treated individuals met fewer diagnostic criteria, reported less desire to gamble, and had a lower SOGS score. They also reported a significantly higher perception of control and self-efficacy. In order to provide clinically relevant results, the percentage of change and end state functioning (comparing post-test scores to a criterion score) were calculated. Among the treatment group, 12 of the 14 participants improved by 50% or more on three dependent variables, and on the end state functioning criteria, in comparison to one of the 15 participants in the control group (85% success rate). Finally, six and 12 month follow-up measures indicated that the therapeutic gains were still present, confirming the long-term effects of this therapeutic program. The following case study illustrates the procedure described above.

6.26.8.1 Case History

Peter is 43 years old, married, and the father of two adolescents. He is currently working as a civil servant. He has been playing video poker for four years. At the time of his consultation, he was playing three times a week on average, which resulted in monetary losses of $350–500 per week (Canadian).

Peter started gambling when two of his colleagues invited him for a drink after work. As soon as they entered the bar, the two colleagues, both video poker players, showed Peter how to play the game. After an initial bet of $10, Peter won $125. During the following weeks, he developed an interest for video poker and played often. He occasionally made significant wins, and started believing that video poker was a good way to make money. After six months, his gambling became so intense that he had to use his personal credit and borrow money to cover his losses.

He progressively became obsessed with gambling. He accumulated increasing debts and was constantly preoccupied by the need to recover his money. He neglected his wife and children more and more each day. He lost his motivation to work, he often arrived late after having spent the lunch hour gambling or, he left early in the afternoon in order to play. His spouse did not know about her husband's gambling habits. She began to worry about his repeated lateness in the evening and his incapacity to pay the bills before the due date. She questioned and doubted Peter's justifications for his prolonged absences. Tension and conflicts became commonplace within their relationship. Peter's urge to play increased with his need to recover the lost money so that he could pay his bills, and in order to escape the climate of tension reproaches in his relationship with his wife.

After four years of excessive gambling, the losses were enormous. His absences from home became no longer justifiable and lying was frequent. One day, his boss, worried about Peter's diminished productivity, asked him the reasons for this change. Confronted with the possibility of losing his job, Peter had no choice but to admit that he had a gambling problem. He also decided to tell everything to his wife. Finally, he decided to seek treatment. A diagnostic of pathological gambling was made by the therapist and confirmed by a second experienced therapist who listened to a tape of the first session. The goal of the first intervention was to evaluate his motivation to change. Peter listed the advantages of stopping gambling and the negative consequences the game has had on his life. Monetary losses, professional and marital problems, and stress were examples of the negative consequences. Also, the return to a normal financial situation, general well-being, and better relationships with the people around him were the main advantages formulated by Peter to stop gambling.

6.26.8.1.1 Identification and correction of faulty cognitions toward gambling

Considering that one of Peter's motivations to gamble was to make money rapidly, it was important to provide him with factual information concerning gambling, and to identify his erroneous cognitions. The therapist asked him to imagine his last gambling session and try to identify his thoughts and the triggering events. Peter reported the following sequence. He left the office and drove toward his usual gambling place instead of going home as he had initially planned. He reported having this reaction instantly, without thinking, as if someone had told him what to do. The therapist had Peter focus his attention on the sequence of events that took place immediately before leaving his office. He was asked to remember what was going on and what he was thinking at that specific moment. Peter remembered putting his documents in order and, by accident, seeing the bills for his credit card that were overdue. The therapist inquired about the thoughts associated with finding unpaid bills. Peter answered that he suddenly felt panic and became very tense. When thinking back on the feeling of panic, he identified the links between this feeling and his unpaid bills. Peter panicked because he did not have the money to pay his bills. This made him so uncomfortable that he wanted to get money to pay them as fast as he could. The therapist finally asked him what he said to himself following this discomfort, and how was he going to solve his problem of obtaining money? At that moment, Peter became aware that he strongly believed that the best way to get a lot of money in a short period of time was by gambling. He realized that this was his motivation to gamble on this last occasion.

The therapist discussed the probabilities of winning, and the negative monetary expectancies of the games which would inevitably result in losses in the long run. The therapist explained that no strategies could be used by the player to win, and that the urge to recover the lost money was, in fact, just an illusion that leads to a vicious circle which can only end in loss. By questioning and confronting Peter's erroneous cognitions over several sessions, he started to realize that thinking he will win money is not realistic. Peter learnt to replace erroneous cognitions with appropriate ones, and to question his gambling habits (e.g., "Why do I feel like playing?" "I want to make money ... is this realistic?" "What could happen to me if I play? I risk further loss." "And if I lose, what will happen? I want to recover my money." "Can I really win?" "Even if I win tonight, it will never be enough to compensate for all the money I have lost." "The longer I try to win, the more likely I am to lose.").

6.26.8.1.2 Problem solving and social skills training

This component was used to modify behaviors related to excessive gambling. Peter mentioned going to the bank machine to check how much money he had left in his banking account. This compulsive checking triggered cognitions related to his lack of money, and the need to win more and gamble. Peter became aware of the sequence of these behaviors and realized the links between this checking behavior and his gambling activities. He resolved to stop checking his banking account. Peter also had to modify other relevant behaviors that stimulated gambling such as keeping less cash money with him, not carrying his banking and credit cards, and avoiding bars with video poker machines. Finally, the involvement with new activities, such as sports and family activities, helped Peter to reorganize his timetable and to replace the time spent gambling with less financially damaging activities.

6.26.8.1.3 Relapse prevention

By learning to identify situations with a high risk of triggering a relapse, and by analyzing situations which led him to gamble, Peter developed cognitive and behavioral strategies needed to refrain from gambling.

6.26.9 EVALUATION OF A COGNITIVE TREATMENT FOR PATHOLOGICAL GAMBLING

Following the positive results obtained through this controlled study, it was decided to evaluate the specific role of correcting the fundamental cognitive error about the notion of randomness. From a theoretical and clinical perspective, it was believed that this component was the crucial variable in the maintenance of excessive gambling. The efficacy of a cognitive treatment for pathological gamblers was assessed based on the correction of erroneous cognitions concerning the notion of randomness and, more specifically, through the modification of the gambler's tendency to link independent events when gambling.

Five pathological gamblers from the population described above participated in this study (Ladouceur, Sylvain, Letarte, Giroux, & Jacques, in press). A single case experimental design across subjects was used to assess the efficacy of the treatment. Cognitive correction included four components:

(i) Understanding the concept of randomness—the therapist explains the concept of randomness, independence among events, the impossibility to control the game.

(ii) Understanding the gamblers' erroneous cognitions, mainly the difficulty to apply the principle of independence among events. The therapist explains how the illusion of control contributes to the maintenance of gambling habits.

(iii) Awareness of inaccurate perceptions.

(iv) Cognitive correction of erroneous perceptions.

The dependent variables were the same measures used in the study described previously. Results indicated that all subjects, except one, increased their perception of control and reduced their urge to gamble, thus supporting the prediction that a cognitive treatment, based on the correction of erroneous perception about the notion of randomness, decreases pathological gambling. The treatment outcome of this intervention provided positive results equivalent to those obtained by a multicomponent intervention. Therefore, cognitive correction of erroneous perceptions toward the notion of randomness is likely to be the key element in the treatment of pathological gambling. We are now replicating this study with a greater number of subjects, using a controlled group comparison design. In conclusion, these therapeutic interventions, with an 85% success rate, open new avenues for the treatment of pathological gambling.

6.26.10 COMMON DIFFICULTIES IN THE TREATMENT OF PATHOLOGICAL GAMBLERS

Clinicians who have treated pathological gamblers have been confronted with a number of difficulties. We will briefly mention the most common and difficult ones encountered over the years in our clinic where approximately 200 pathological gamblers have been interviewed. The first and most important issue is treatment compliance. Very often, individuals will ask for an appointment but, without canceling, will not show up. They frequently drop out after one or two sessions, or will simply decide to terminate treatment after a few weeks of abstinence, thinking that their problem has been solved. There are no obvious solutions to these matters. Our clinical experience has led to the development of a series of procedures used to improve treatment compliance.

First of all, when an individual contacts us, we return the call within 24 hours, conduct a brief telephone interview, and administer the SOGS. If the individual appears to be a pathological gambler, an appointment is scheduled within a week at which point we conduct a structured interview and administer a series of questionnaires. If therapy is undertaken, the therapist will inform the patient that compliance is a major difficulty for pathological gambling and explore ways to overcome this problem. Furthermore, the patient will sign a contract indicating that they agree to participate for at least 10 sessions, and that if they are unable to attend, they will call ahead of time to cancel the session, that they will pay one session in advance, and that after two nonmotivated absences, treatment will be terminated.

A second difficult issue is determining whether the goal of the treatment should be controlling the gambling or abstinence. Many pathological gamblers ask that the goal of treatment be controlled gambling. Some will be very convincing and will put forth appropriate and rational arguments. We have often caught ourselves spending many sessions discussing this difficult subject without finding a solution. Gamblers suggest that controlled gambling is the main goal, simply because they cannot see themselves abstaining from this activity, or simply because they are not ready to stop gambling. Our approach on this matter is to consider abstinence as the main goal, and then once it is achieved, the matter can be reconsidered. In a few cases, instead of losing the patient, we have accepted controlled gambling as the main goal of treatment. However, this solution should never be accepted except as a last option.

The third difficulty is identifying erroneous perceptions towards the notion of randomness as discussed above. Often, pathological gamblers will simply deny that they have these misconceptions. The best way to pursue this matter is to design some behavioral experiment to facilitate the identification of erroneous perceptions. Gamblers deny these perceptions because they are not aware of them.

6.26.11 CONCLUDING REMARKS

This account of pathological gambling takes a cognitive perspective. It is assumed that pathological gambling occurs when the gambler persists in gambling, despite the losses involved, believing that ultimately the losses will be recovered and money will be won. These false beliefs of the gambler are maintained by the inadequate knowledge of the probabilities of success and the erroneous interpretation of the notion of randomness. If this account of the motivation to gamble is correct, then it follows that pathological gambling can be treated effectively by correcting erroneous beliefs of the patients. This chapter presents data supporting this theoretical position.

6.26.12 REFERENCES

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorder* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorder* (4th ed.). Washington, DC: Author.

Abbott, M., & Volberg, R. (1996). The New Zealand national survey of problem and pathogical gambling. *Journal of Gambling Studies, 12,* 143–160.

Allcock, C. C. (1987). An analysis of a successful racing system. In M. B. Walker (Ed.), *Faces of gambling*. Sydney, Australia: National Association for Gambling Studies.

Allcock, C. C., & Dickerson, M. G. (1986). *The guide to good gambling*. Australia: Social Sciences Press.

Anderson, G., & Brown, R. I. F. (1984). Real and laboratory gambling, sensation-seeking and arousal. *British Journal of Psychology, 75,* 401–410.

Becona, E. (1996). Prevalence surveys of problem and pathological gambling in Europe: The of Germany, Holland and Spain. *Journal of Gambling Studies, 12,* 179–192.

Beyer, A. (1993). *Beyer on speed: New strategies for racetrack betting*. Boston: Houghton Mifflin.

Bird, R., & McCrae, M. (1985). Gambling markets: A survey of empirical evidence. In G. T. Caldwell, M. G. Dickerson, B. Haig, & L. Sylvan (Eds.), *Gambling in Australia*. Sydney, Australia: Croom Helm.

Blaszczynski, A., & McConaghy, N. (1994). Antisocial personality disorder and pathological gambling. *Journal of Gambling Studies, 10,* 129–146.

Blaszczynski, A., McConaghy, N., & Frankova, A. (1991). Control versus abstinence in the treatment of pathological gambling: A two to nine year follow-up. *British Journal of Addiction, 86,* 299–306.

Blaszczynski, A. P., Winter, S. W., & McConaghy, N. (1986). Plasma endorphin levels in pathological gambling. *Journal of Gambling Behavior, 2,* 3–14.

Brown, R. I. F. (1996). Arousal and sensation-seeking components in the general explanation of gambling and gambling addictions. *International Journal of the Addictions, 21,* 1001–1016.

Bujold, A., Ladouceur, R., Sylvain, C., & Boisvert, J.-M. (1994). Cognitive behavioral treatment of pathological gamblers. *Journal of Behavior Therapy and Experimental Psychiatry, 25,* 275–282.

Carlton, P. L., & Manowitz, P. (1987). Physiological factors as determinants of pathological gambling. *Journal of Gambling Behavior, 3,* 274–285.

Connor, X. (1983). *Report of the inquiry into a casino for Victoria*. Melbourne, Australia: Victorian Government Printer.

Cornish, D. B. (1978). *Gambling: A review of the literature and its implication for policy and research*. London: HMSO.

Coventry, K. R. (1997). Rationality and decision making: The case of gambling and the development of gambling addiction. *The proceedings of the Ninth International Conference on Gambling and Risk Taking*, Las Vegas, NE, 1994.

Custer, R. L., & Milt, H. (1985). *When luck runs out*. New York: Facts on File Publications.

Dickerson, M. G. (1984). *Compulsive gamblers*. London: Longman.

Dickerson, M. G. (1993). Internal and external determinants of persistent gambling: Problems in generalizing from one form to another. In W. R. Eadington & J. A. Cornelius, (Eds.), *Gambling behavior and problem gambling*. Reno, NV: Institute for the study of gambling and commercial gaming.

Dickerson, M. G., & Adcock, S. (1987). Mood, arousal and cognitions in persistent gambling: Preliminary investigation of a theoretical model. *Journal of Gambling Behavior, 3,* 3–15.

Dickerson, M. G., Allcock, C., Blaszczynski, A., Nicholls, B., Williams, J., & Maddern, R. (1995). *An examination of the socio-economic effects of gambling on individuals, families and the community, including research into the costs of problem gambling in New South Wales*. Sydney, Australia: Report to the Casino Community Benefit Fund Trustees in New South Wales.

Dixon, D. (1996). Illegal betting in Britain and Australia: Contrasts in control strategies and cultures. In J. McMillen (Ed.), *Gambling cultures: Studies in history and interpretation* (pp. 86–100). London: Routledge.

Drapkin, T., & Forsyth, R. (1987). *The punter's revenge*. London: Chapman and Hall.

Echeburua, E., Baez, C., & Fernandez-Montalvo, J. (1996). Comparative effectiveness of three therapeutic modalities in the psychological treatment of pathological gambling: Long-term outcome. *behavioral and Cognitive Psychotherapy, 24,* 51–72.

Frey, J. H., & Eadington, W. R. (1984). *Gambling: Views from the social sciences*. The Annals of the American Academy of Political and Social Science, 474 (Special Issue).

Gaboury, A., & Ladouceur, R. (1989). Erroneous perceptions and gambling. *Journal of Social Behavior and Personality, 4,* 411–420.

Giroux, I., Ladouceur R., & Jacques, C. (1998). *Correction of erroneous perceptions and arousal*. Manuscript submitted for publication.

Griffin, P. A. (1987, August). *Mathematical expectation for the public's play in casino blackjack*. Paper presented at the Seventh International Conference on Gambling and Risk Taking, University of Nevada, Reno, NV.

Goldfried, M. R., & Davison, G. C. (1976). *Clinical behavior therapy*. New York: Holt Rinehart and Winston.

Griffiths, M. D. (1994). The role of cognitive bias and skill in fruit machine gambling. *British Journal of Psychology, 85,* 351–369.

Haig, J. (1985). Gambling expenditure. In G. Caldwell, M. Dickerson, J. Haig, & M. Sylvain (Eds.), *Gambling in Australia* (pp. 76–85). Sydney, Australia: Croom Helm.

Hayano, D. M. (1977). The professional poker player: Career identification and the problem of respectability. *Social Problems, 24,* 556–564.

Jacobs, D. F. (1986). A general theory of addictions: A new theoretical model. *Journal of Gambling Behavior, 2,* 15–31.

Ladouceur, R. (1996). Prevalence of pathological gamblers in Canada and related issue. *Journal of Gambling Studies, 12,* 129–142.

Ladouceur, R., Arsenault, C., Dubé, D., Jacques, C., & Freeston, M. H. (1997). Pathological gamblers among volunteers in experiments on gambling. *Journal of Gambling Studies, 13,* 69–84.

Ladouceur, R., Boisvert, J.-M., & Dumont, J. (1994). Cognitive behavioral treatment of adolescent pathological gamblers. *Behavior Modification, 18,* 230–242.

Ladouceur, R., & Dubé, D. (1997). Erroneous perceptions in generating random sequences: Identification and strength a basic misconception. *Swiss Journal of Psychology, 56,* 256–259.

Ladouceur, R., Giroux, I., & Jacques, C. (1998). Winning on the horses: How much strategy and knowledge is needed? *Journal of Psychology, 132,* 133–142.

Ladouceur, R., Paquet, C., & Dubé, D. (1996). Erroneous perceptions in generating sequences of random events. *Journal of Applied Social Psychology, 26,* 2157–2166.

Ladouceur, R., Paquet, C., Lachance, N., & Dubé, D. (1996). Mécanismes psychologiques de la perception du hasard. *International Journal of Psychology, 31,* 93–99.

Ladouceur, R., Sylvain, C., Letarte, H., Giroux, I., & Jacques, C. (in press). Cognitive treatment of pathological gamblers. *Behavior Research and Therapy.*

Ladouceur, R., & Walker, M. (1996). Cognitive perspective on gambling. In P. M. Salkovskis (Ed.), *Trends in cognitive therapy* (pp. 89–120). Chichester, UK: Wiley.

Lorenz, V. C., & Shuttleworth, D. E. (1983). The impact of pathological gambling on the spouse of the gambler. *Journal of Community Psychology, 11,* 67–76.

Marlatt, G. A. (1985). Relapse prevention: General overview. In G. A. Marlatt & J. R. Gordon, (Eds.). *Relapse prevention. Maintenance strategies in the treatment of addictive behaviors* (pp. 33–44). New York: Guilford.

McConaghy, N. (1980). Behavior completion mechanisms rather than primary drives maintain behavioral patterns. *Activita Nervosa Supplement, 22,* 138–151.

McConaghy, N., Armstrong, M. S., Blaszczynski, A., & Allcock, C. (1983). Controlled comparison of aversive therapy and imaginal desensitisation in compulsive gambling. *British Journal of Psychiatry, 142,* 366–372.

McMillen, J. (1996). *Gambling cultures.* London: Routledge.

Oldman, D. (1978). Compulsive gamblers. *Sociological Review, 26,* 349–371.

Orford, J. (1985). *Excessive appetites: A psychological view of addictions.* Chichester, UK: Wiley.

Perkins, E. B. (1950). *Gambling in English life.* London: Epworth Press.

Rosecrance, J. (1986). Why regular gamblers don't quit: A sociological perspective. *Sociological Perspectives, 29,* 357–378.

Rowntree, B. S. (1941). *Poverty and progress; a second social survey of York, London, and New York.* London: Longmans and Green.

Royal Commission (1933). *Royal Commission on Lotteries and Betting, 1932–33: Findings.* London: Her Majesty's Stationary Office.

Scott, D. (1982). *The winning way.* Sydney, Australia: Puntwin.

Sharpe, L., & Tarrier, N. (1993). Towards a cognitive-behavioral theory of problem gambling. *British Journal of Psychiatry, 162,* 407–412.

Sharpe, L., Tarrier, N., Schotte, D., & Spence, S. H. (1995). The role of autonomic arousal in problem gambling. *Addiction, 90,* 1529–1540.

Sylvain, C., Ladouceur, R., & Boisvert, J.-M. (1997). Cognitive and behavioral treatment of pathological gambling: A controlled study. *Journal of Consulting and Clinical Psychology, 65,* 727–732.

Volberg, R. (1996). Prevalence of pathological gambling in the United States. *Journal of Gambling Studies, 12,* 111–128.

Wagenaar, W. A. (1988). *Paradoxes of gambling behavior.* London: Erlbaum.

Walker, M. B. (1992). A sociocognitive theory of gambling involvement. In W. R., Eadington & J. A. Cornelius (Eds.), *Gambling and commercial gaming: Essays in business, economics, philosophy and science* (pp. 371–398). Reno, NV: University of Nevada.

Walker, M. B. (1995). Pathological gambling: The fundamental error. In J. O'Connor (Ed.), *High stakes in the nineties* (pp. 97–98). Fremantle, Australia: National Association for Gambling Studies.

Walker, M. B., & Dickerson, M. G. (1996). The prevalence of problem and pathological gambling: A critical analysis. *Journal of Gambling Studies, 12,* 233–249.

Walker, M. B., Sturevska, S., & Turpie, D. (1995). The quality of blackjack play in Australian casinos. In J. O'Connor (Ed.), *High stakes in the nineties* (pp. 173–182). Fremantle, Australia: National Association for Gambling Studies.

Yardley, H. O. (1957). *The education of a poker player.* Glasgow, UK: Fontana/Collins.

Zuckerman, M. (1979). *Sensation seeking: Beyond the optimal level of arousal.* Hillsdale, NJ: Erlbaum.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.27

Sexual Problems: Dysfunction

W. P. DE SILVA

Institute of Psychiatry, University of London, UK

6.27.1 SEXUAL DYSFUNCTION 604

6.27.2 CLASSIFICATION AND DIAGNOSTIC ISSUES 604

- *6.27.2.1 Sexual Desire Disorders* 605
  - *6.27.2.1.1 Hypoactive sexual desire disorder* 605
  - *6.27.2.1.2 Sexual aversion disorder* 605
- *6.27.2.2 Sexual Arousal Disorders* 605
  - *6.27.2.2.1 Female sexual arousal disorder* 605
  - *6.27.2.2.2 Male erectile disorder* 605
- *6.27.2.3 Orgasmic Disorders* 605
  - *6.27.2.3.1 Female orgasmic disorder* 605
  - *6.27.2.3.2 Male orgasmic disorder* 605
  - *6.27.2.3.3 Premature ejaculation* 605
- *6.27.2.4 Sexual Pain Disorders* 605
  - *6.27.2.4.1 Dyspareunia* 605
  - *6.27.2.4.2 Vaginismus* 605

6.27.3 ETIOLOGICAL FACTORS 605

6.27.4 ASSESSMENT OF SEXUAL DYSFUNCTION 606

- *6.27.4.1 Aims of Assessment* 606
- *6.27.4.2 The Interview* 606
  - *6.27.4.2.1 Special problems* 606
  - *6.27.4.2.2 Areas to enquire about* 607
  - *6.27.4.2.3 Individual and couple interviews* 607
  - *6.27.4.2.4 Motivation and selection* 607
  - *6.27.4.2.5 Physical examination and investigations* 607
  - *6.27.4.2.6 Questionnaires and inventories* 608
  - *6.27.4.2.7 Subjective ratings* 608
  - *6.27.4.2.8 Physiological measures* 609

6.27.5 FORMULATION 609

6.27.6 TREATMENT OF SEXUAL DYSFUNCTION 609

- *6.27.6.1 Behavioral Sex Therapy* 609
  - *6.27.6.1.1 The behavioral approach* 609
  - *6.27.6.1.2 Present-day conjoint cognitive-behavioral sex therapy* 610
  - *6.27.6.1.3 Some general considerations on therapy* 613
- *6.27.6.2 Some Practical Issues* 614
- *6.27.6.3 Therapy for Those Without Partners* 614
- *6.27.6.4 Group Therapy* 615
- *6.27.6.5 Efficacy of Therapy* 615
- *6.27.6.6 A Note on Physical Treatments* 615
- *6.27.6.7 Scope of Cognitive-behavioral Sex Therapy* 616

6.27.7 TREATMENT ISSUES WITH SPECIAL POPULATIONS 616

- *6.27.7.1 The Elderly* 617
- *6.27.7.2 Those with Physical Handicaps and Chronic Illnesses* 617
- *6.27.7.3 Gay Clients* 617

6.27.7.4 Women with Eating Disorders 618
6.27.8 SUMMARY AND COMMENTS ON FUTURE DEVELOPMENTS 618
6.27.9 REFERENCES 619

6.27.1 SEXUAL DYSFUNCTION

The term "sexual dysfunction" refers to a disturbance or impairment in sexual functioning. Sexual dysfunctions are usually contrasted with sexual deviations, or called "paraphilias." The latter are seen as a disturbance in the *direction* of one's sexual interest or desire, rather than of *functioning*. Thus paraphilias include paedophilia (sexual interest in children), fetishism (sexual interest in inanimate objects), and so on. Dysfunctions, on the other hand, are failures or problems in the actual sexual functioning, for example, difficulty in obtaining an erection, or failure to reach an orgasm.

It must be noted, however, that there can be overlap and interdependence between sexual dysfunctions and paraphilias. A man complaining of erectile difficulties in coital activity with adult partners may get strong erections in paedophiliac acts. Equally, a man presenting with, say, fetishistic desires for rubber or leather may also have erectile difficulty when he is not in contact with rubber or leather garments (de Silva, 1995). This overlap between these two types of sexual problems is important to recognize and take into account in clinical settings.

A further consideration is the link between sexual dysfunction and marital or relationship problems. Both clinical and research data show that the two areas are often related. Rust, Golombok and Collier (1988) reported that sexual dissatisfaction in the male partner strongly correlated with a poor marital relationship as perceived by both partners. The relationship between the two areas is not an invariant one, nor is there any clear evidence of the direction of causation when there is a link.

The prevalence of sexual dysfunction is hard to establish, as epidemiological studies are scarce. Masters and Johnson (1970) estimated that half of all couples in the USA have a sexual dysfunction. Spector and Carey (1990), who undertook a full review of the sexuality, psychopathology, and epidemiology literature, concluded that sexual dysfunctions may be very common. A major problem is that the presence of a sexual dysfunction is not always distressing to, or seen as a problem by, couples or individuals. The much-cited study of Frank, Anderson and Rubinstein (1978) provides some interesting data. Well over half the women and nearly half the men in this survey of 100 happily married American couples reported sexual difficulties, yet most were satisfied with their sexual life. Similar findings have been reported in a Swedish study (Nettelbladt & Uddenberg, 1979) and a British investigation (Reading & Wiest, 1984). This highlights another key issue. Sexual dysfunction is not an all-or-none affair; there are varying degrees of difficulties. The way they are perceived varies; occasional erectile failure may be seen as a perfectly acceptable situation by some men and their partners, while it may bring other men to a clinic seeking help. There are also major differences in sexual functioning in relation to age. Inability to obtain an erection increases progressively with age (Feldman, Goldstein, Hatzichristou, Krane, & McKinlay, 1994; Weizman & Hart, 1987). In the Kinsey survey (Kinsey, Pomeroy, & Martin, 1948), it was found that less than 1% of males had erectile dysfunction before the age of 19, increasing to 25% by the age of 75.

6.27.2 CLASSIFICATION AND DIAGNOSTIC ISSUES

Sexual dysfunctions have been classified in different ways by different authors in the past (e.g., Bancroft, 1989; Hawton, 1985), but in the 1990s a consensus has been emerging. The current classification offered in the *Diagnostic and statistical manual of mental disorders* (4th ed., *DSM-IV*) (American Psychiatric Association [APA], 1994) links dysfunctions to the phases of the sexual response cycle: (i) desire (this phase includes desire to have sexual activity and sexual fantasies); (ii) excitement (this phase consists of a subjectively felt sense of sexual pleasure and related physiological changes, for example, penile erections and vaginal lubrication); (iii) orgasm (this phase consists of the climax or peaking of sexual pleasure, with release of sexual tension and the rhythmic contraction of the perineal muscles and reproductive organs; in the male there is also ejaculation of seminal fluid, and in the female contraction of the wall of the outer third of the vagina); and (iv) resolution (this phase consists of a sense of muscular relaxation and general well-being, in the male this also includes a physiologically refractory period in which further erection and orgasm do not take place).

The dysfunctions, linked to the first three of these phases, are classified below.

6.27.2.1 Sexual Desire Disorders

6.27.2.1.1 Hypoactive sexual desire disorder

Here, there is an absence or deficiency of sexual fantasies and desire for sexual activity. The terms "low libido" and "reduced libido" are also commonly used to refer to this.

6.27.2.1.2 Sexual aversion disorder

In this disorder, there is an aversion to, and avoidance of, genital sexual contact with a sexual partner. Often the aversion is focused on a particular aspect of sexual experience. There may also be anxiety, fear, or disgust.

6.27.2.2 Sexual Arousal Disorders

6.27.2.2.1 Female sexual arousal disorder

Here, there is a persistent inability to attain or maintain adequate vaginal lubrication and vaginal expansion in response to sexual excitement.

6.27.2.2.2 Male erectile disorder

In this condition, the main feature is the persistent or recurrent inability to attain or sustain an adequate penile erection. There is much variability in this. Some attain very strong erections without being able to sustain them, some have sustainable erections which are not strong enough for penetration. Loss of erection during thrusting is also reported by some.

6.27.2.3 Orgasmic Disorders

6.27.2.3.1 Female orgasmic disorder

This is also referred to as "inhibited female orgasm," "anorgasmia," and more rarely, "inorgasmia." The essential feature is the persistent delay in, or absence of, orgasm following a normal sexual excitement phase.

6.27.2.3.2 Male orgasmic disorder

Here, the male has difficulty in reaching orgasm following a normal sexual excitement phase, despite adequate stimulation. This condition has also been called "inhibited male orgasm." An earlier term, "retarded ejaculation" is problematic, as it focuses on ejaculation rather than orgasm. In some conditions a male can reach an orgasm without discharging any semen. The new term refers specifically to the orgasmic experience.

6.27.2.3.3 Premature ejaculation

Here, the essential feature is the persistent onset of orgasm with minimal sexual stimulation. Men with this problem reach an orgasm prior to, or just after, penetration, thus making any meaningful coital activity impossible. Attempts to define premature ejaculation by linking it to partner's orgasm (cf. Masters & Johnson, 1970) is problematic.

6.27.2.4 Sexual Pain Disorders

In addition the *DSM-IV* includes the following:

6.27.2.4.1 Dyspareunia

Here, there is genital pain associated with sexual intercourse. This is most commonly experienced during coitus, but it can also occur before or after coitus. It can happen in both men and women, but it appears from clinical settings that it is much more common in the latter (Bancroft, 1989; Renshaw, 1988).

6.27.2.4.2 Vaginismus

This consists of the persistent involuntary contraction of the muscles surrounding the outer third of the vagina when penetration is attempted. When severe, it makes penetration impossible.

Each of the above dysfunctions can be either *lifelong* or *acquired.* A lifelong dysfunction is said to exist when the problem has been present since the beginning of sexual functioning, for example, a man who has never had an erection as an adult. In the acquired type, the dysfunction develops only after a period of normal functioning.

Sexual dysfunctions are also categorized into *generalized* and *situational* types. In the former, the dysfunction occurs in all situations, and is not linked to certain types of partners, situations, or stimulation. In the latter, the dysfunction is limited to certain types of stimulation, situations, or partners. For example, a man may have an erectile difficulty when attempting sex with a partner, but not when masturbating.

It is also important to note that there can be comorbidity in sexual dysfunctions. The co-occurrence of erectile dysfunction and premature ejaculation is not uncommonly seen in clinics.

6.27.3 ETIOLOGICAL FACTORS

Sexual dysfunctions may be due to a variety of etiological factors, and clinicians usually look for organic and psychological factors. The state of knowledge with regard to organic factors is not easy to summarize, as different dysfunctions can have different etiologies. The most work has been done on male erectile disorder. Endocrine, vascular, and neurological systems have been seen as crucial for erectile functioning (Bancroft, 1997). Alcohol and other drugs of abuse, and certain prescribed medications, can also have a role to play. For female dysfunctions, hormonal factors are sometimes strongly implicated, as in sexual arousal disorder where decreases in estrogen can cause vaginal dryness. Such changes are not uncommon following menopause or ovariectomy. Local infections, presence of pelvic tumor, and pelvic inflammatory disease are among the factors that can be associated with dyspareunia. Psychological factors in the etiology of sexual dysfunctions have been extensively written about (e.g., Bancroft, 1989; de Silva, 1994a; Wincze & Carey, 1991). The role of anxiety, including "performance anxiety," is commonly found to be linked to erectile disorder, and indeed some other dysfunctions. Stresses of various kinds, depressed affect, anger, and relationship factors have been implicated. In women, lack of learning to relinquish control has been considered by some authors as an important variable (McConaghy, 1993). Cognitive factors, including attitudes and beliefs, both individual and cultural, are also seen as relevant (e.g., Baker, 1993; Bhugra & de Silva, 1993; Spence, 1991). The overall picture from many studies with inconsistent results clearly points in the direction of multifactoral etiology for sexual dysfunctions. Biological factors, and cultural and psychological factors in the individual and relationship domains, seem to determine the origins of sexual dysfunctions, as indeed they determine and influence sexual functioning in general. These issues are well reviewed in McConaghy (1993) and Wincze & Carey (1991) among others.

6.27.4 ASSESSMENT OF SEXUAL DYSFUNCTION

Assessment of sexual dysfunctions should be comprehensive, and cover in detail the presenting problem(s) and the relevant associated factors. While the principles of clinical assessment of sexual dysfunctions is in theory no different from the clinical assessment of other problems, in practice it is often a more challenging task. Since sex is a very private aspect of one's life, detailed and open discussion of it, however relevant clinically, can be difficult for the patient and the inexperienced therapist. Further, some of the usual methods of clinical assessment, such as direct observation by therapist and reports from independent observers, are not possible for obvious reasons. Some of these matters, and other general issues in the assessment of sexual dysfunction, are discussed in detail in the literature (e.g., Bancroft, 1989; Hawton, 1985; Wincze & Carey, 1991).

6.27.4.1 Aims of Assessment

The purpose of assessment is to obtain as clear a picture as possible of the problem(s) and factors that may be related to it, so that a good formulation can be arrived at. From the perspective of psychological therapy, one would want to gain information about the dysfunction in behavioral terms, also including relevant cognitive aspects. What happens in the actual sexual situation needs to be enquired about, and this is usually facilitated by focusing on the most recent attempt. Information about the behavior of both partners will be elicited, including their antecedents and consequences. Relevant past experiences, past partners, traumatic events, and so forth are part of the picture. A full behavioral analysis (e.g., Kanfer & Saslow, 1969) should be aimed for.

6.27.4.2 The Interview

6.27.4.2.1 Special problems

The main source of information in assessment is the clinical interview. In the case of sexual dysfunction assessment, there are special problems associated with it (cf. de Silva, 1994a; Wincze & Carey, 1991). These include the following:

(i) There may be embarrassment on the part of the patient. The therapist needs to be sensitive to the difficulties the patient may have in talking about intimate matters. Starting with more general questions may help build rapport. Acknowledging that talking about sexual matters could be difficult can also help.

(ii) The language must be simple and easy to understand. The jargon that comes easily to a professional is often unfamiliar to the patient. Every effort must be made to ensure that the patient understands the questions.

(iii) Precise details need to be obtained about the problem presenting, going beyond general descriptions such as "I don't seem to want it any more" or "I have lost interest." The therapist, however, needs to be patient in enquiring about details. Careful probing is often needed.

(iv) There should be an attitude of nonjudgmental acceptance with regard to the patient's behavior, likes, and dislikes. Even unintended display of disapproval or surprise may discourage the patient from giving relevant details.

(v) It is possible that patients may not be able to disclose some information in the initial assessment interview. This may be due to embarrassment, or due to the therapist being still unfamiliar to them. It is important to allow patients to give more information in later sessions.

6.27.4.2.2 Areas to enquire about

The areas to be covered in the interview include the following:

(i) The nature of the problem in as much detail as required to obtain a full picture of the difficulty and all its associated factors, including anxiety and situational variations; in men with impotence, particular enquiry should be made about whether early morning erections are present or not, as their presence usually helps to rule out the possibility that the problem is organic.

(ii) The history of the problem, its beginnings and course, and present sexual activities including masturbation.

(iii) The partner's reactions to the problem, both in the sexual situation and in general.

(iv) The patient's sexual knowledge, beliefs, and attitudes, including those determined by religion, culture, and subculture.

(v) The patient's sexual likes, dislikes, preferences, and fantasies.

(vi) Past sexual history including relevant early experiences, first experience of intercourse, and so on.

(vii) Psychiatric and medical factors, including drugs, alcohol, and so forth; current depression is particularly important to assess.

(viii) Menstrual history and relation of problem to the menstrual cycle.

(ix) Contraception and past pregnancies, and attitude to the possibility of conception.

(x) General relationship factors.

(xi) Background factors, such as job, income, accommodation, extended family and so on, which can be sources of stress.

(xii) Previous treatment, if any.

More details may be needed on some of the areas than on others in a given case, and this is a matter of clinical judgment as the interview proceeds. Needless to say, therapists must be prepared to vary their enquiry to suit each patient, as needed. Most clinicians rely on a history schedule or a general checklist of topics to enquire about as a matter of routine. Examples of such checklists are available in, among others, de Silva (1994a), Hawton (1985), and Spence (1991). In practice, while such detailed checklists are indeed useful, flexibility is needed in their use; such a list or schedule should serve as a general guide rather than as something to be rigidly adhered to.

6.27.4.2.3 Individual and couple interviews

One major issue in interviewing patients with sexual dysfunction is whether the partners—when a couple present themselves for help—should be interviewed together or not. Different viewpoints have been expressed about this (e.g., Bancroft, 1989; Masters & Johnston 1970; Wincze & Carey, 1991). In general, a good arrangement is to see the couple jointly to start with, and then to conduct assessment interviews separately. If two therapists are available, this may be done in parallel sessions; if not, more time should be spent with the partner with the presenting problem, followed by a briefer interview with the other partner.

It is important that individual interview sessions are undertaken in all cases. This provides an opportunity for each partner to give his or her version of the problem, and to discuss with the therapist relevant matters, including feelings about the partner, without inhibition. It also gives an opportunity for individuals to divulge information that might have been kept away from the partner, such as an extramarital relationship, or a particular aspect of the individual's past history.

6.27.4.2.4 Motivation and selection

Assessment of the motivation of the patient/couple for therapy is an important aspect of the interview, although this may not prove easy in a single interview except perhaps to identify those who are clearly unwilling to accept the treatment offered. As for suitability for therapy, an assessment comprising two interviews at most is usually sufficient to identify those who are clearly not likely to benefit from the therapy that can be offered. For example, presence of clear psychiatric illness will often require appropriate treatment of that condition first; serious marital difficulties may require referral to a specialist clinic, unless therapists feel competent to deal with this themselves; and, when physical factors are probably involved, investigations of these, and a physical examination, will need to be arranged prior to acceptance for therapy.

6.27.4.2.5 Physical examination and investigations

An important issue for the clinician dealing with sexual dysfunctions is that of whether or not a physical examination and/or investigations should be asked for. Nonmedical clinicians cannot always assume that the clients referred to them have been adequately screened for possible organic factors. Some clinicians take the view that all patients presenting with sexual dysfunction should routinely be examined physically (e.g., Kolodny, Masters, & Johnson, 1979; Spence, 1991). This is unnecessary, but it is essential to have the facilities to get this and relevant investigations done if required in a given case. Bancroft (1989) has given an extremely useful set of indications for physical examination. These are:

(i) complaints of pain or discomfort during sex;

(ii) recent history of ill-health or physical symptoms other than the sexual problem;

(iii) recent onset of loss of sex drive with no apparent cause;

(iv) when the patient believes that a physical cause is most likely, or is concerned about the genitalia (e.g., a man complaining that his penis is too small or bent, or a woman suspecting that there is something abnormal about her sexual organs);

(v) history of abnormal puberty or other endocrine disorder;

(vi) in men, being aged over 50; and

(vii) in women, being in the pre- or post-menopausal age group or having a history of marked menstrual irregularities or infertility.

Where appropriate, the medical practitioner carrying out the physical examination will also carry out, or arrange for, relevant laboratory investigations. Details of the common investigations may be found in several sources including Bancroft (1989), Hawton (1985), and Wincze & Carey (1991).

6.27.4.2.6 Questionnaires and inventories

Data obtained from the interview can profitably be supplemented by the use of questionnaires and inventories. These help to cover some important areas quickly, but more importantly they provide quantitative data that are particularly useful in assessing differences between before and after treatment.

Several useful instruments are available for the measurement of sexual experiences, attitudes, dysfunctions, and other related matters. Hoon, Wincze, and Hoon (1976a) provide an inventory (Sexual Arousability Index) for the assessment of female sexual arousal. Lo Piccolo and Steger (1974) have developed an inventory (Sexual Interaction Inventory) to assess sexual interaction and satisfaction of a couple. Lief and Reed (1972) provide a questionnaire to assess both sexual knowledge and sexual attitudes, while Wilson (1978) has produced a useful fantasy questionnaire, which measures fantasies, desires, and actual behaviours. The Derogatis Sexual Functioning Index (Derogatis & Melisaratos, 1979) is a wide-ranging scale of sexual functioning covering 10 domains (e.g., information, desire, attitudes). While its comprehensiveness is no doubt an asset, its prohibitive length (245 items) makes it somewhat unwieldy for routine clinical use. A brief measure for assessing erectile functioning, developed recently, is the International Index of Erectile Function (Rosen et al., 1997).

A relatively recent British development is the Golombok–Rust Inventory of Sexual Satisfaction (GRISS; Rust & Golombok, 1986). This 28-item questionnaire, intended for use with heterosexual couples or individuals with a current heterosexual relationship, yields an overall score, for men and women separately, of the quality of sexual functioning. In addition, the following subscores can also be obtained: erectile dysfunction, premature ejaculation, anorgasmia, vaginismus, infrequency, poor communication, dissatisfaction, nonsensuality, and avoidance. This instrument has good reliability, and is easy to use. In view of the multiplicity of scores it yields, the GRISS is an economical instrument that can be used routinely.

A parallel, and equally economical, instrument for the assessment of the quality of the overall relationship is the Golombok–Rust Inventory of Marital State (GRIMS; Rust, Bennun, Crowe, & Golombok, 1988). Other established marital questionnaires include the Locke–Wallace Marital Adjustment Scale (Locke & Wallace, 1959) and the Dyadic Adjustment Scale (Spanier, 1976). The recently developed Marriage and Relationship Questionnaire (Russell & Wells, 1993) is another useful instrument to be considered.

When depression is a relevant factor needing to be assessed, a standard depression inventory may be used for this purpose. The same applies to anxiety.

6.27.4.2.7 Subjective ratings

Self-rating scales may be used as part of an assessment of the major variables in question for a given patient. For example, anxiety in sex, desire, and sexual arousal may each be rated by the patient on a 0 to 100 scale indicating subjective estimates (cf. de Silva, 1994a). Patients usually find these simple scales easy to use. Equally simple are frequency charts, recording the frequency of target behaviors on a daily basis. A predesigned diary provided by the assessor/therapist, specifying the targets to be recorded, is an effective way of obtaining baseline data as well as of monitoring change. Conte (1986) and Spence (1991), among others, provide discussions of these.

6.27.4.2.8 Physiological measures

Physiological techniques have been used in the assessment of sexual function for many years, receiving impetus from the work of Masters and Johnson (1966, 1970). Measuring techniques are available for both male and female arousal. Penile plethysmography for the assessment of erection is widely used in research and can be used in clinical practice where needed and practicable (e.g., Wagner & Green, 1981; Wincze et al., 1988). The measure may be of either penile volume or penile circumference changes. Penile plethysmography has also been used to assess nocturnal erections in an attempt to distinguish psychogenic from organically caused impotence (Karacan, 1978). The injection of a chemical vasodilator, such as papaverine, into one of the corpora cavernosa of the penis, is also used in the assessment of erectile disorder. If the patient develops a full erection following the injection, severe vascular insufficiency can be excluded as a cause of the problems (Brindley, 1983; Wagner, 1993). As for female arousal, the best established method is the photoplethysmography technique (Hoon, Wincze, & Hoon, 1976b), in which vasocongestion in the vaginal walls is measured with the help of a probe. Detailed discussions of these, and other laboratory techniques, for the assessment of sexual problems are provided by Bancroft (1989), Meisler and Carey (1990), McConaghy (1993), and Schiavi (1992).

In a clinical setting, however, the use of physiological methods for routine assessment is not feasible. Also, the interpretation of their results in the clinical context is not always clear-cut (Bancroft, 1989; Conte, 1986).

6.27.5 FORMULATION

All the information gained from the different aspects of assessment need to be carefully integrated. This should lead to a case formulation, the equivalent of a working hypothesis (Carey, Flasher, Maisto, & Turkat, 1984; Hawton, 1985). A good formulation will include the following:

(i) a detailed description of the problem;

(ii) predisposing factors (e.g., cultural beliefs, history of diabetes, strict upbringing, childhood traumatic events);

(iii) precipitating or triggering factors (e.g., childbirth, physical illness, redundancy, bereavement, discovery of partner's adultery, failure while attempting sex in insecure circumstances);

(iv) perpetuating factors (e.g., anxiety about possible failure, ongoing marital conflict, continuing stress at work); and

(v) prognostic factors (e.g., good overall relationship and commitment of both to the marriage as good prognostic factors).

The formulation will provide a tentative explanation of the problem, and also establish a basis for treatment. The formulation is fed back to the patient or couple, and discussed. Revision of the formulation may follow this discussion.

6.27.6 TREATMENT OF SEXUAL DYSFUNCTION

6.27.6.1 Behavioral Sex Therapy

6.27.6.1.1 The behavioral approach

Historically, there have been several psychological approaches to sexual problems (de Silva, 1994b). However, the major approach since the 1960s has been the behavioral one. Behavioral sex therapy became established as the treatment of choice for these problems in the 1970s, and began to be practiced widely in various parts of the world (Gillan, 1987).

In the modern folklore of sex therapy, it is widely held that the basic techniques of behavioral sex therapy derive from the work of William Masters and Virginia Johnson (Masters & Johnson, 1970). It is certainly true that Masters and Johnson developed, established, and popularized a treatment package that has had a major impact. However, the roots of behavioral sex therapy can be traced to the work of Joseph Wolpe in South Africa. Wolpe is regarded as the pioneer of modern behavior therapy, his work from the late 1940s on the application of learning principles to the analysis and modification of behavioral problems culminating in his book *Psychotherapy by reciprocal inhibition* (Wolpe, 1958). While much of the work reported in this book related to phobias and other anxiety disorders, Wolpe also reported behavioral work on sexual dysfunctions. He recognized the crucial role of anxiety in sexual problems, and used anxiety reduction as a key element in the treatment of these problems. The principle of the couple undertaking cooperative therapy, the need to remove pressure to perform, and the use of a graded series of activities eventually leading up to the resumption of sexual intercourse, which are seen as three major elements of present-day sex therapy, were key aspects of Wolpe's approach.

Wolpe (1958) cites cases of both erectile dysfunction and premature ejaculation with whom he had successfully used this approach. It is perhaps worth quoting a passage that summarizes Wolpe's approach:

> The patient is told to inform his sexual partner (quoting the therapist if necessary) that his sexual difficulties are due to absurd but automatic fears in the sexual situation, and that he will overcome them if she will help him, i.e. if she will participate on a few occasions in situations of great sexual closeness without expecting intercourse or exerting pressure toward it. He is to ask her to be patient and affectionate and not to criticize. Assured of her cooperation, he is to lie in bed with her in the nude in a perfectly easy, relaxed way, and thereafter to do just what he really feels like doing *and no more*. He has no duty at any stage to reach any criterion of performance. It is found that from one love session to the next there is a decrease in anxiety and an increase in sexual excitation and therefore in the extent of the caresses to which the patient feels impelled. He has increasingly strong erections, and usually after a few sessions coitus is accomplished and then gradually improves. (Wolpe, 1958, p. 131)

6.27.6.1.2 Present-day conjoint cognitive-behavioral sex therapy

While the work of Wolpe (1958) represents the early use of standard behavior therapy techniques for sexual dysfunctions (see also Lazarus, 1963), it is profitable to consider these and other behavioral techniques as elements that may be incorporated into a wider treatment package. The obvious reason for this is that sexual problems are often multifaceted so that to deal with them effectively one needs several methods. The general package most widely used in this way is the conjoint therapy of Masters and Johnson (1970), modified and further developed by subsequent writers (e.g. Bancroft, 1989; Gillan, 1987; Hawton, 1985; Spence, 1991; Wincze & Carey, 1991). One major new development is the incorporation of cognitive principles and techniques (e.g. Baker, 1993; Spence, 1991). This approach may be termed "conjoint cognitive-behavioral sex therapy."

In the Masters and Johnson conjoint therapy approach, the presenting partner and spouse are seen as a couple for therapy. In the original program, each couple was seen by a male and female cotherapist team. Therapy was carried out on an intensive basis: daily sessions over a two-week period (Masters & Johnson, 1970). Other clinicians and researchers found that the involvement of a second therapist added little to the program and that sessions did not need to be so closely massed together (Arentewicz & Schmidt, 1983; Bancroft, 1989; Crowe, Gillan, & Golombok, 1982).

The main elements of the conjoint cognitive-behavioral sex therapy approach may be summarized as follows:

(i) Treating the dysfunction as a joint problem. This helps to reduce worry and guilt in the presenting partner and also emphasizes that the need is to learn, or relearn, how to have satisfactory sex jointly.

(ii) Reduction of anxiety. This is achieved by several strategies, including banning any attempt at intercourse. This removes the pressure to perform. Relaxation training may also be used as an extra help. Graded steps are recommended, so that progress is achieved without a resurgence of anxiety.

(iii) Setting sexual tasks or assignments to be carried out at home. They are specific behavioral tasks and involve touching, caressing, and so on. The two main stages of this are "nongenital sensate focus," where the touching excludes genitals and breasts, and "genital sensate focus" where these are included. These basic tasks aim to help the couple to learn giving and receiving pleasure by touch, with no pressure for performance, moving from less sexual to more intimate interactions.

(iv) Work on eliciting and changing negative or problematic cognitions.

(v) Educating the couple in sexual knowledge, for example, anatomy, physiology, coital positions.

(vi) Helping the couple to develop sexual and other communication skills.

(vii) Using specific additional techniques for specific dysfunctions.

While (i) to (vi) above are common to all, there are specific interventions designed to deal with specific presenting problems. The timing of their introduction is flexible; they are usually introduced after some progress has been made in the common stages.

This program is a cognitive-behavioral package in that there is no attempt to interpret the presenting symptoms in terms of psychodynamic constructs, and that behavioral tasks and cognitive work are a major part of the package. The degree to which an approach geared towards unraveling conflicts and problems in the overall relationship is incorporated into this, varies from therapist to therapist and from case to case (Beck, 1992; Crowe & Ridley, 1990; Woody, 1992). Anxiety reduction is a key part of this package. The prohibition of attempts at intercourse helps to achieve this, as immediately the performance anxiety in the male and fears of pain and so on in the female are removed. Other techniques of anxiety reduction may be added as required. McCarthy (1977, 1992), for example, has listed a number of ways, both general and specific, in which anxiety can be reduced in sexual problems.

The sensate focus assignments help a couple to learn to relax in each other's company and enjoy physical contact and interaction without worries of failure. In this relaxed, mutually pleasuring stage, they can acquire the confidence to move towards more intimate interactions. It is perhaps worth noting here that the progression from nongenital to genital sensate focus, and from there to more specific and more explicitly sexual acts, is similar in many ways to an *in vivo* desensitization program (Wolpe, 1958). Self-exploration may also be incorporated into the program where required. Communication, both verbal and nonverbal, on matters of pleasure, sensations, and sexual responses is encouraged and taught (e.g., how to indicate to the partner where and how to touch and how to express pleasure at what the partner is doing). The verbal aspects of this kind of simple, but to many couples new, interaction may be role-played and rehearsed during therapy sessions. The meetings with the therapist are crucial in discussing progress or otherwise of the homework assignments, and difficulties and problems are discussed fully. Not infrequently in these feedback discussions relevant new material about the relationship emerges for the first time (Bancroft, 1989; Hawton, 1985).

Cognitive therapy principles have a key part to play. The elicitation and modification of cognitions related to the dysfunction are the main feature of this, and individual work along these lines is often a key part of therapy. Where maladaptive cognitions have a role in the genesis and/or maintenance of a problem, these need to be challenged and changed. Common myths, attitudes, and idiosyncratic beliefs, and so forth, often contribute to sexual difficulties (Baker & de Silva, 1988; Bishay, 1988; de Silva & Dissanayake, 1989). Spence (1991) has provided a useful discussion of the use of cognitive strategies in the treatment of sexual difficulties. She includes fantasy training and attention-focusing skills as possible aspects of therapy. The major cognitive interventions, however, consist of identifying and restructuring the relevant cognitions. Maladaptive or negative thoughts need to be identified and changed as needed (Zilbergeld, 1978, 1992). This approach is well represented in the work of Baker (1993). She highlights the role of the individual's core assumptions about gender role and accompanying expected behaviors. Such assumptions are reflected in automatic thoughts, commonly reported by patients in sex therapy, such as: "My partner will leave me if I continue to lose my erections" or "I can't be very masculine if I don't last long with a partner." Baker (1993) argues that these negative thoughts play a crucial role in maintaining, and in some cases even causing, sexual dysfunctions. The cognitive approach she uses aims to challenge and modify these dysfunctional cognitions, and to break the vicious circle of negative thoughts that contribute to difficulties. The role of cognitive work has also been emphasized by Bancroft (1997) and Wincze and Carey (1991). The need to focus on the partner's cognitions as well as those of the presenting patient has also to be emphasized. Typical partner cognitions include: "I am no longer exciting to him," "He is just not trying," and "She must be having an affair with someone" (cf. Wincze & Carey, 1991).

Specific techniques are incorporated into this general program in the treatment of specific disorders. Comments on the treatment of some specific dysfunctions are made below.

(i) Hypoactive sexual desire disorder

Low interest usually begins to respond to the general treatment package if nothing more serious is underlying. Additional techniques are self-focusing, self-stimulation, use of vibrators, and stimulation with erotic material such as pictures, videos, and audiotapes. (Gillan, 1979, 1987). It is important, however, that the kind of erotic material recommended or provided is not distasteful to the patients, and so the choice is best left to them. Fantasy training is a related technique, sometimes useful for those whose fantasies are minimal (Spence, 1991). To encourage fantasies, published fantasy materials may be used, such as the volumes of fantasies published by Nancy Friday (1976, 1991). Masturbation training, usually with fantasy, may also be considered. The use of agreed timetables for sex are also found to be useful (e.g., Crowe & Ridley, 1990). The therapist would negotiate with the couple a timetable for sex; for example, intercourse will take place only on certain days of the week, or sex will be initiated by each partner on certain days only. This approach is often quite effective in cases where the partner with the low level of desire finds the demands made by the spouse too much, thus making the problem worse. The timetable helps to establish an acceptable pattern or schedule, within which further progress can be made.

(ii) Male erectile disorder

In the genital sensate focus stage, "teasing" is introduced, that is, periods of penile stimulation alternating with absence of stimulation. While erections may spontaneously occur, these are not considered the aim of therapy at this stage and the couple are encouraged to let the erection subside before restimulating. This helps in training them not to rush to intercourse once an erection is there, and also demonstrates that erections, when lost, can reappear. The next stage is vaginal entry, in the female superior position but with no movement. In the following stage the female makes slow movements, eventually leading to the male participating in and/or initiating movement, and using different positions.

Cognitive work has a particularly valuable role to play in the treatment of this disorder. Positive sexual thoughts should be identified and encouraged. The patient needs to be helped to refocus his thinking onto more positive cognitions. One way of doing this is to encourage recall of the thought content during past satisfactory sexual experiences. The therapist can also help the patient to develop positive thoughts by prompting "typical" helpful thoughts (Wincze & Carey, 1991). There is also a need to appraise and reappraise the meaning of specific types of sexual acts. Negative meanings or appraisals that inhibit the sexual functioning need to be altered (cf. Bancroft, 1997).

Some cognitive work may also be needed with the partner. Partners of men with erectile disorder not uncommonly develop negative cognitions, and they need to be dealt with.

(iii) Female orgasmic disorder

For females with orgasmic dysfunction the main additional element in the therapy package is a good deal of self-focusing and self-stimulation (Barbach, 1980; Gillan, 1987; Heiman & Lo Piccolo, 1988; Lo Piccolo & Stock, 1986). This helps the patient to learn to enjoy the sexual sensations in a relaxed manner. These self-sessions, including stimulation until orgasm is achieved, can be built into the basic program as parallel assignments. It has been shown that the use of vibrators can help these women to achieve orgasm; this may be done by the patient herself first, and the partner may help her to achieve orgasm with the vibrator in later stages. The use of fantasy and erotic materials may also be used as an adjunct (Gillan, 1979, 1987; Spence, 1991). Exercises to achieve control over the pubococcygeal muscle and to strengthen its tone, commonly referred to as "Kegel exercises" (Kegel, 1952), have also been recommended as an aid to achieving orgasm. For females who are unable to achieve orgasm with the partner but have no problem in masturbation, other elements may need to be added to the program. For example, orgasmic reconditioning may be attempted, in which the patient is taught to pair the positive pleasurable aspects of self-stimulation or other sexual situations with images of the partner in a fantasy-based graded program (Asirdas & Beech, 1975; Gillan, 1987).

In some, perhaps many, instances a woman's inability or difficulty in reaching an orgasm is linked to her inability to relinquish control (McConaghy, 1993, 1996). Some help specifically directed towards this, involving education, counseling and practice exercises, is necessary in such cases.

(iv) Male orgasmic disorder

The aim is to work towards intravaginal ejaculation, in a series of steps gradually approaching this goal. Again the instructions are introduced at the genital touching stage. For those males who do not reach an orgasm easily in any situation, vigorous stimulation ("superstimulation") with the aid of a lubricant is recommended. The use of a vibrator may also be considered. Once orgasm can be achieved in this way, vaginal entry, after some initial manual stimulation, may be attempted. In those men whose problem is that they cannot reach orgasm in the vagina but can do so with manual stimulation, a graded program in which orgasm is achieved by manual stimulation, with the penis increasingly close to the vagina, is recommended. In the next stage vaginal entry is achieved, after stimulation by hand close to orgasm. Even then, some manual stimulation may be needed to achieve orgasm once the penis is in the vagina. Suitable coital positions that facilitate these steps need to be recommended. Subsequently, vaginal stimulation alone will be sufficient for orgasm (Kaplan, 1987).

(v) Premature ejaculation

There are two, closely similar, techniques used for premature ejaculation. Masters and Johnson (1970) recommend what is called the "squeeze" technique. The couple are asked to practice this in the genital touching state. The female stimulates the penis of her partner with her hand, and when the man feels he is about to reach a climax, he indicates this to her with a prearranged signal. She then squeezes the penis hard for two or three seconds. For squeezing, the penis is held with the thumb on the frenulum and the first and second fingers on the opposite surface, one on each side of the coronal ridge. The squeeze makes the man lose his urge to ejaculate, and also perhaps some of the erection. This process of stimulation and squeeze is repeated several times in a session. Several sessions of this leads to gradual increase in ejaculatory control, and the couple is then asked to effect vaginal entry, in the woman-above position. At first, entry is not followed by movement. If the man feels he is about to ejaculate during vaginal containment, he communicates this to his partner who then lifts herself off him lightly and applies the squeeze. Kolodny et al. (1979) recommend a basilar squeeze technique at this stage so that the penis does not have to be completely disengaged. The penis is held at the base, anterior to posterior, and the pressure applied. The couple may eventually revert to preferred positions.

The squeeze technique is, in fact, a variant of the start–stop method described in 1956 by Semans, which is used by many therapists (e.g. Kaplan, 1974). This consists of stimulating the penis and stopping at the point of near-climax, and repeating the process several times. Initially, the stimulation is with a dry hand; later a lubricant is used to increase sensitivity and make the sensations more like the experience of vaginal entry (Gillan, 1987). The rest of the program consists of vaginal containment without movement, followed by movement in the woman-above position.

(vi) Vaginismus

This is treated by helping the patient, as a first step, to learn to relax and to explore her own genitals. Following this, a graded series of steps involving penetration using own fingers, and dilators (or trainers) of increasing sizes, is undertaken (Gillan, 1987; Scholl, 1988). The use of a vaginal lubricant may be advised. After the initial stages, the partner may be included in the procedure. This is done, however, with the patient retaining control. In many cases, this graduated dilation training leads to quick improvement.

6.27.6.1.3 Some general considerations on therapy

Some general points about this treatment approach need to be made.

(i) The therapy package

The package in its basic form is for use with all sexual dysfunctions, with the specific elements added to suit specific difficulties. There are, however, many instances where there is no need to apply the whole package, where perhaps sexual counseling and basic education are all that is needed (Bancroft, 1989; Hawton, 1985; Kaplan, 1987). The therapist needs to have a flexible approach in applying the therapy package or parts of it.

(ii) Psychogenic and organic problems

The program is intended mainly for psychogenic problems, as patients with clearly organic problems may need different forms of therapy. On the other hand, there is no reason why some aspects of the program should not be used even with those whose problems are not psychogenic, as a way of helping them to relax in their sexual activity and enhance their sexual experience, thus facilitating whatever physical treatment may be used. For example, in elderly men with erectile dysfunction who, due to aging and other physical causes, may not be able to have erectile function restored, much can still be achieved by a program aimed at enhancing their enjoyment of sexual activity, enabling them to accept that sex need not always mean vaginal intercourse (Gibson, 1992). It is also important to remember that the psychogenic–organic distinction is not always clear cut, and there are often multiple causal factors (Bhugra & Crowe, 1995). A problem caused, or triggered, by a physical condition, may often be aggravated or maintained by psychological factors, such as anxiety, invalidism, and diminished self-esteem. In many cases, substantial psychological help is needed in addition to a physical intervention.

(iii) Flexibility

The program is meant to be applied flexibly. The needs of each couple determine what changes to aim for, and the therapist must be prepared to change the direction of a program as and when required; for example, a couple presenting with a premature ejaculation problem may turn out to be one where the basic problem is lack of responsiveness in the female. Some couples require a good deal of direct education; shy patients may not be able to participate fully in therapy sessions until their embarrassment is overcome, and will need considerable time and effort to reach the point where therapy can proceed. The need for flexibility is very well illustrated by Lobitz, Lo Piccolo, Lobitz, and Brockway (1976) and Wincze and Carey (1991).

(iv) Resistance to the program

The program can sometimes meet with resistance. A couple may not carry out homework assignments, or do them only infrequently or cursorily. Their difficulties will need to be fully explored. Previously unacknowledged marital or relationship problems may come to the surface at this stage, in which case efforts should be directed towards resolving these (Bancroft, 1989; Crowe & Ridley, 1990).

(v) Additions to the program

Additions to the basic program need not be confined to the specific techniques mentioned above. Any suitable cognitive and/or behavioral technique for aspects of the problem may be incorporated as required. For example, when a strong phobic element is present, an intervention to deal with the phobia may be used; for a man whose dysfunction is bound up with paraphiliac desires, a parallel program aimed at dealing with these may be required (de Silva, 1995).

(vi) Functional analysis

The notion of functional analysis (Kanfer & Saslow, 1969) is extremely useful in devising individualized treatment programs. When the assessment shows clearly identifiable factors related to the problem behavior, the systematic manipulation of these can achieve impressive results, usually as part of the program, but sometimes also as the main intervention (Lobitz et al. 1976).

6.27.6.2 Some Practical Issues

As noted earlier, Masters and Johnson (1970) used male and female teams of cotherapists in their program, but others have shown that a second therapist adds little to the program effectiveness (e.g., Crowe et al., 1982). On the other hand, in certain cases an additional therapist of the opposite sex to the main therapist can be an advantage, for example, by making it easier for both of the partners to communicate about the problem and their feelings about it.

The individuals should be given an opportunity to report to the therapist separately, that is, in the absence of the partner, at least in some of the sessions. This enables the therapist to gain a balanced account of the progress and related matters.

Giving written instructions, with illustrations, is a valuable addition to the therapy program. They should be used as a supplement to verbal instructions. Equally useful is to recommend a well-written basic book on sex. There are many good books that may be recommended, for example, Barnes and Rodwell (1992), Delvin (1974) and Yaffe and Fenwick (1986a, 1986b). Films, videos, and slides may also be used (Gillan, 1987).

6.27.6.3 Therapy for Those Without Partners

When a patient comes for therapy without a partner, what help can be offered? If the patient has a steady partner who, though unwilling to come to the clinic, will co-operate with a therapy program, a "remote control" approach may be used, with the presenting partner also acting as communicator of instructions. The use of written material will be particularly useful in such cases. Clinical experience shows that this is a less than ideal substitute for conjoint therapy, but should be considered in the right circumstances. If the refusal of the partner to attend the clinic reflects a poor relationship, and/or an attitude that it is all the other's problem, then clearly the chances of joint work being successfully carried out are slim.

A somewhat different problem arises when the patient has no partner available. Many young men with erectile or ejaculatory problems not only do not have a steady partner but also avoid developing relationships, through fear of failure and rejection. Individual therapy is the only option available in such cases, unless group therapy is considered and facilities are available for it (see below).

The basic principles of individual sex therapy are largely the same as those for couples. Education, counseling, anxiety reduction in various ways including relaxation, self-focusing and self-stimulation, and fantasy training are all possible and useful elements in such programs. Cognitive principles have a key role. As for specific techniques, premature ejaculators will find the start-stop technique easier to use than the squeeze. Imaginal desensitization may be used for fear or anxiety, while role play and social skills training can be useful where required (McCarthy, 1992). The sexual reeducation program of Zilbergeld considers it important to dispel some widespread male "myths" about sex (such as that sex always means intercourse, that the male must always take the active role, and so on) in helping these persons, and exploration—and correction—of the individual's misconceptions about sex will be a useful element in an individual sex therapy program (Zilbergeld, 1978, 1992). For females, masturbatory exercises, Kegel exercises, fantasy training, vibrator use, and other such techniques that are possible for the individual to use without a partner, may be used as needed.

The treatment of individuals without partners presenting with sexual problems is discussed by, among others, Anson (1995) and Catalan, Hawton, and Day (1991).

The use of surrogate partners in the treatment of patients who come without partners has been reported and recommended by some therapists.

Several sex therapists in the USA use surrogate therapy (Dauw, 1988; Sommers, 1980), while in the UK Cole (1988) has described such therapy with 425 patients, 390 men and the rest women. Of these, 316 (74.4%) completed therapy; unfortunately, the follow-up was possible in only 13.3% of these. There are serious problems with the use of surrogate therapy, including obvious legal and ethical issues. For example, in the Cole sample, the spouse in many cases was not aware that the patient had come for therapy. A serious clinical question is whether patients who have been treated with a surrogate partner will be able to generalize their gains to other situations. In view of these reservations, it is difficult to recommend surrogate therapy as an option in the management of sexual dysfunction.

6.27.6.4 Group Therapy

In recent years, an increasing number of therapists have treated patients with sexual dysfunction in groups. There are reports in the literature of male groups, female groups, couple groups, and mixed-single groups; and there have been groups for patients with similar problems and groups for heterogeneous problems (e.g., Barbach & Flaherty, 1980; Kayata & Szydlo, 1988; O'Gorman, 1978; Spence, 1985; Zilbergeld, 1975). The groups have used a variety of techniques, including education, task-setting, relaxation, desensitization, instructions to use masturbation and vibrators, and open discussion of problems. Particularly for young and sexually diffident persons, the experience of well-conducted group therapy can be of much benefit (Bancroft, 1989).

6.27.6.5 Efficacy of Therapy

What about the efficacy of the therapeutic approaches and techniques that have been discussed above? The very high success rates reported by Masters and Johnson (1970; also Kolodny et al., 1979) have not been matched by later investigators but there is, generally, evidence that the conjoint cognitive-behavioral approach is beneficial to many patients (Arentewicz & Schmidt, 1983; Bancroft, 1989; Hawton, 1992; Hawton, Catalan, Martin, & Fagg, 1986; Spence, 1991). A thorough evaluation of sex therapy outcome is not easy to undertake because of various confounding factors: heterogeneity of samples, lack of uniformity in outcome measures, ambiguity of criteria of improvement, preponderance of single-case reports, poor description of patient characteristics, absence of data on drop-outs and failures, and so on (Beck, 1992). Within these limitations, the available data are encouraging. There are no reliable comparative outcome data, but on the whole 50–60% of couples who enter treatment show satisfactory outcome. Marked differences in outcome exist for specific sexual dysfunctions. For example, excellent results are obtained for vaginismus (e.g., Hawton & Catalan, 1990), whereas the treatment of hypoactive sexual desire appears to be less effective, especially in the long term (e.g., Hawton, Catalan, & Fagg, 1991). While the conjoint cognitive-behavioral therapy approach is the best option when it is feasible, it is up to clinicians to use their ingenuity and judgment in choosing, for each case, those elements of therapy that are particularly suitable for the problems they are called upon to deal with.

A related issue is that of the mechanisms of change. Even when various techniques and strategies achieve clinical success, the mechanisms through which change has been achieved is not always clear. Bancroft (1997) has recently highlighted this issue and offered a valuable, if brief, discussion. The relative lack of understanding of mechanisms of change also reflects the uncertainty about the etiological factors in these dysfunctions.

The outcome literature also points to factors associated with prognosis. Good prognostic factors include the quality of the overall relationship. Motivation for therapy also appears to be a key factor. There is also some evidence that the attractiveness of the partners to each other predicts satisfactory outcome. Current psychiatric illness is a predictor of poor outcome. These issues have been reviewed by Hawton (1992).

6.27.6.6 A Note on Physical Treatments

A brief note is necessary about physical treatment. A variety of physical interventions are used in medical settings for sexual dysfunctions. It is beyond the scope of this chapter to review these. Hormonal therapy, surgery, vacuum devices, and penile prosthetic devices have been used, and certain drugs have also been claimed to be of value. Much of this work has been reviewed by Bancroft (1989), Bhugra and Crowe (1995), McConaghy (1996) and others. In recent years, many reports have appeared on the use of injections of smooth muscle relaxants such as papaverine and prostaglandin E into one of the corpora cavernosa of the penis to induce erections (e.g. Brindley, 1986; Linet & Ogring, 1996). The erections induced in this way may last from one to four hours and are usually adequate for satisfactory intercourse. The patient is taught to give these injections himself, and the partner may also be taught to play this role. The use of this method to help men with organically caused erectile problems, or those who have failed to respond to an adequate trial of psychological methods, is advocated by several authorities. There is, however, an urgent need for further research and the use of this approach needs caution (Althof & Turner, 1992; Bhughra & Crowe, 1995; Crowe & Qureshi, 1991). The procedure also has some side-effects, especially parapism (prolonged erection) in a small number of cases. An even more recent development is the intraurethral administration of prostaglandin E (e.g., Wolfson, Pickett, Scott, De Kernion, & Rajfer, 1993).

When used appropriately after careful investigation, physical methods do have a useful role in the treatment of sexual dysfunctions. However, when the problem is largely psychologically determined, the use of physical treatments can be counterproductive. Even when there are physical causes of the presenting problem, and the main treatment is of a physical nature, additional psychological help along the lines discussed above is often useful and necessary (Wincze & Carey, 1991). The reliance on a physical treatment alone in the treatment of sexual dysfunction is not recommended (Bhugra & Crowe, 1995).

6.27.6.7 Scope of Cognitive-behavioral Sex Therapy

It was noted above that success rates have declined in recent years when compared to the very high rates reported in early years. This may be due to changes in culture and to changes in the nature of referrals for specialist help. Lo Piccolo (1994) has argued that the behavioral approach of the 1970s and the 1980s needs to be augmented, and that a "postmodern" sex therapy is required. In this approach, while the behavioral outlook is still the primary feature, the role that sexual dysfunctions may play in maintaining a systemic homeostasis for the couple in crucial areas of their relationship (e.g., intimacy, trust, power) is recognized, and is appropriately dealt with. Second, it is recognized that a sexual problem may have individual psychodynamic significance. Third, it is recognized that a sexual dysfunction may be a mechanism for the patient to deal with unresolved family or origin issues, such as problems of closeness associated with an alcoholic father, or an aversive reaction associated with incestuous abuse. Finally, the operant reinforcing value of a sexual dysfunction for clients in their larger environment is looked into; for example, "the financial success that comes to a man who works especially hard to compensate for his erectile failure" (Lo Piccolo, 1994, p. 6).

It is clear that all but one of these aspects of sexual dysfunction highlighted by Lo Piccolo (1994) can be accommodated easily within the broad framework of the cognitive-behavioral approach. In fact this happens almost routinely in actual clinical practice. As noted above, relationship factors are always an important area to investigate and work on in sex therapy. A systemic approach may be of much value in this endeavor in some cases. Such an approach can be accommodated easily within a broad cognitive-behavioral framework, as demonstrated by Crowe and Ridley (1990). The reinforcing value of the sexual problem would naturally be investigated and dealt with in the cognitive-behavioral approach, as indeed it is done for all behavioral difficulties (Kanfer & Saslow, 1969). The historical aspects of a case need to be considered in any approach, and a good cognitive-behavioral formulation would necessarily include recognition of these.

As for the psychodynamic value and significance of a symptom, this may appear at first sight to be an aspect that does not easily fit within a cognitive-behavioral framework. Kaplan (1974, 1987) has argued for the use of a psychodynamic approach if and when behavioral procedures produce no further change in a patient, in other words when there is resistance. However, it can be argued that a broadly conceived cognitive-behavioral approach, imaginatively applied, should be able to deal with these matters. It appears that what is needed in such cases is for the individual to be helped to reappraise the meanings of certain sexual acts or responses (Bancroft, 1997). Negative meanings acquired in childhood, persisting into adult life and inhibiting appropriate sexual responses, need to be reappraised. As Bancroft (1997) points out, there is no doubt that this endeavor "lends itself particularly well to the cognitive-behavioral approach" (p. 247). What is required is for the therapist to lead the patient towards this necessary reappraisal with the skillful and sensitive use of cognitive-behavioral principles.

6.27.7 TREATMENT ISSUES WITH SPECIAL POPULATIONS

Sexual dysfunctions occurring in various special populations need some consideration. The approach to treatment described in the above sections is a robust one, the principles of which are applicable to all groups. However, certain additional considerations are sometimes needed, and treatment needs to be adapted as required. A few commonly encountered special clinical populations are discussed below.

6.27.7.1 The Elderly

Elderly couples and individuals increasingly attend services for sexual problems (Gibson, 1992; Spence, 1991). This is not because of any real increase of sexual dysfunctions in the group, but because of the awareness of the availability of services and also the increasing acceptance that sexual activity in later life is not something to be ashamed of (Gibson, 1992). Erectile disorder is one of the commonest presentations (Mulligan, 1989); male and female orgasmic disorder also present for help. While physiological influences have a clear role to play in some of these cases, problems of ill-health can also contribute to, and/or complicate, the difficulties. Changing life circumstances, including the nature of the overall relationship between a couple, are also relevant factors. Psychological factors, especially attitudinal ones, may also play a part. It is not uncommon for elderly people to take the view that sexual activity should no longer be part of their lives, and this can be a problem when the partner continues to desire an active sex life.

In the treatment of elderly clients, all of the above need to be considered in the assessment, and should be part of the formulation. When needed, advice, information, and reassurance need to be given. The normality of continuing sex in later years may need to be emphasized. If there are fears or concerns about one's loss of attractiveness to the partner, some specific cognitive work on this will need to be undertaken.

The therapy program needs to take into account the clients' preferences and longstanding habits and practices, and any instructions for new and adventurous sexual activity should be given only after careful consideration and discussion. Greater emphasis on the nongenital aspects of sexual activity may be needed, and many elderly couples can learn to find much sexual pleasure in this way. Specific practical advice may include the use of lubricants, when there is a female arousal disorder including lack of lubrication—a very common presentation—and on how vaginal penetration may be achieved even with a not fully erect penis (Gillan, 1987). Finally, there are instances when the presentation of a sexual dysfunction in this age group is linked to anxiety about not being able to perform as well as in previous years. This is particularly so with males who may get concerned about their declining erectile response. Part of the treatment would then be providing information about the common changes in sexual functioning with increasing age, thus allaying any undue anxiety. The variability of the effects of aging on sex needs to be emphasized. Further, increasing the need for direct physical stimulation of the partner's penis in order to obtain an erection has to be emphasized, and the instructions for touching and foreplay need to include this as a major part.

6.27.7.2 Those with Physical Handicaps and Chronic Illnesses

Many clients with physical handicaps and chronic illnesses seek help for sexual problems. Cardiac problems, epilepsy, arthritis, cancer, and prostrate problems are among the difficulties that are commonly encountered. Assessment should include full details of the problems, including the study of medical reports about the client's difficulties. The medication the client may be on is also important to ascertain, as it may well have an effect on sexual functioning and the presenting dysfunction may even be just a secondary result of medication. A third area to pay attention to is how the physical illness or handicap has affected the person psychologically, mainly in the areas of self-image and role in relation to the partner. Chronic illness and handicaps often lead to invalidism, which affects the person's functioning including confidence and role in the area of sex. These issues need to be addressed in therapy, and it is best done using cognitive techniques with the individual and in joint counseling (cf. Malloy & Herold, 1986).

When there is chronic handicap, the aim of therapy for sexual dysfunction should be to maximize pleasurable sexual activity that is possible within the limits imposed by the handicap. A couple may need to use only certain coital positions, and variations within these. Those who are paraplegic or quadraplegic, or who have a limited range of motion due to back problems or arthritis, need education and advice on comfortable positions (cf. Bullard, 1988). Sometimes the focus needs to be on noncoital sexual activities. Couples may need to be given carefully graded behavioral homework exercises, in which they can proceed at a slow pace and without fear of pain or risk of fatigue. Some of these issues are discussed in detail by Schover (1992), Schover and Jensen (1988), and Spence (1991).

6.27.7.3 Gay Clients

Presentation of gay men for help in sexual dysfunction is common (Reece, 1988). The problems they present are varied, and cover the full spectrum that is seen in heterosexual men (Bhugra & Wright, 1995). Etiological factors often include fear of HIV/AIDS, uncertainty about sexual orientation, and feelings of being shunned or oppressed by society at large (cf. Bhugra, 1987; Dupras & Morriset, 1993). In the treatment of these clients, the general principles of cognitive-behavioral sex therapy are applied, with special focus on the issues noted above. Therapists need to be nonjudgmental in their attitude. When fear of sexually transmitted diseases are a part of the picture, clear advice on safe sex practices needs to be provided as part of the therapy package.

In the treatment of gay clients without partners, a group approach may be considered (cf. Reece, 1985). This format provides an economical and effective way of dealing with issues of stigma, ambivalence about sexual orientation and so on, in addition to work on specific dysfunctions. Another issue that needs to be dealt with, in the treatment of gay men without partners, is how one might get over the fear—and avoidance—of developing new relationships, out of embarrassment about the sexual dysfunction. Here, a behavioral rehearsal may be used to provide practice in disclosing the dysfunction to a new partner (cf. McCarthy, 1992).

6.27.7.4 Women with Eating Disorders

Females with eating disorders, especially anorexia nervosa, present a special challenge to therapists. There is a very high incidence of sexual problems in women with anorexia nervosa (Beaumont, Abraham, & Simson, 1981; Bancroft, 1989; de Silva & Todd, 1998; Raboch & Faltus, 1991). The reasons for this are complex. Clearly, in severe anorexia nervosa endocrine factors are involved, and this affects sexual functioning. There is also evidence that many anorexic patients have anxieties about sexuality and womanhood (cf. Anderson, 1985). When they seek treatment, it is often for hypoactive female sexual disorder, or for arousal disorder. Anorgasmia is also found as a problem, and not uncommonly sexual aversion disorder (see de Silva & Todd, 1998).

There are major issues in the treatment of these clients. While the basic cognitive-behavioral approach described in this chapter forms the basis of therapy for them, its application has to be cautious and carefully tailored to the individual. Often they come for treatment with some ambivalence; they may come either due to the need to please their partner if they are in a relationship, or in order to "become normal." Any attempt to treat sexual dysfunction while the client is still at a very low body weight is a mistake. The pace of therapy should be determined by the client's readiness. Preliminary work on sex education and attitudes may be needed (Van Vreckem & Vandereycken, 1994). Individual work preceding joint work is often advisable. Specific cognitive work on the significance of sex and sexuality, the client's self-concept, and so on may also be needed. Essentially, the therapy with this group of clients takes a slower, more finely graded form than in general sexual dysfunction therapy. Examples of the treatment of females with anorexia nervosa along these lines are found, among others, in Guile, Horne, and Dunston (1978) and de Silva and Todd (1998).

6.27.8 SUMMARY AND COMMENTS ON FUTURE DEVELOPMENTS

This chapter has reviewed the nature, diagnosis, etiology, assessment, and treatment of sexual dysfunction. The field has developed rapidly since the 1970s, with many practicing clinical psychologists getting actively involved in the clinical management of these problems, and in research aimed at elucidating them. Work in this area is now very much part of mainstream clinical psychology practice.

The current picture is, in essence, a promising one. There is much in the literature to guide the practitioner in the clinical assessment and treatment of these difficulties. The outcome literature is relatively weak, due to a variety of factors. However, there is enough in the literature on the outcome of therapy for sexual dysfunction to enable practitioners confidently to undertake the treatment of those patients who seek help for sexual problems.

As for the future, several developments can be predicted. Many therapists have begun to use and develop cognitive techniques more extensively in recent years, and this trend will continue to grow. The incorporation of the systemic approach into a broad cognitive-behavioral framework is another recent development that is likely to progress further. An integrated approach deriving concepts and strategies from the cognitive-behavioral and the systemic traditions may well emerge as the favored model in the near future. The other major area of development is likely to be in the physical treatments for sexual dysfunctions, mainly but not exclusively male erectile dysfunction. Extensive and ambitious research is currently taking place in this area, and the results of these investigations and trials are likely to have an impact on sexual dysfunction therapy. It is probable that more effective physical treatments will become available, but this is unlikely to reduce the need or demand for psychological treatments. First, the proportion of patients who will derive significant benefit from physical treatments is likely to be limited. Second, even when an effective physical treatment is used for a patient or couple, there will remain a clear need for additional psychological help. For these cases, it is likely that an integrated approach combining physical and psychological interventions will become the treatment of choice.

6.27.9 REFERENCES

Althof, S. F., & Turner, L. A. (1992). Self-injection therapy and vacuum devices in the treatment of erectile dysfunction: Methods and outcome. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: Assessment and treatment* (pp. 283–309). New York: Guilford.

Anderson, A. E. (1985). *Practical comprehensive treatment of anorexia nervosa and bulimia*. London: Edward Arnold.

Anson, M. (1995). Non-couple therapy for sexual dysfunction. *International Review of Psychiatry, 7,* 205–216.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Arentewicz. G., & Schmidt, G. (1983). *The treatment of sexual disorder: Concepts and techniques of couple therapy*. New York: Basic Books.

Asirdas, S., & Beech, H. R. (1975). The behavioural treatment of sexual inadequacy. *Journal of Psychosomatic Research, 19,* 345–353.

Baker, C. D. (1993). A cognitive-behavioural model for the formulation and treatment of sexual dysfunction. In J. M. Ussher & C. D. Baker (Eds.), *Psychological perspectives on sexual problems* (pp. 110–128). London: Routledge.

Baker, C. D., & de Silva, P. (1988). The relationship between male sexual dysfunction and belief in Zilbergeld's myths: An empirical investigation. *Sexual and Marital Therapy, 3,* 229–238.

Bancroft, J. (1989). *Human sexuality and its problems* (2nd ed.). Edinburgh, UK: Churchill Livingstone.

Bancroft, J. (1997). Sexual problems. In D. M. Clark & C. G. Fairburn (Eds.), *Science and practice of cognitive behaviour therapy* (pp. 243–257). Oxford, UK: Oxford University Press.

Barbach, L. (1980). *Women discover orgasm*. New York: Free Press.

Barbach, L., & Flaherty, M. (1980). Group treatment of situationally anorgasmic women. *Journal of Sex and Marital Therapy, 6,* 19–29.

Barnes, T., & Rodwell, L. (1992). *A Woman's guide to loving sex*. London: Boxtree.

Beaumont, P. J. V., Abraham, S. F., & Simson, K. (1981). The psychosexual histories of adolescent girls and young women with anorexia nervosa. *Psychological Medicine, 11,* 131–140.

Beck, J. G. (1992). Behavioural approaches to sexual dysfunction. In S. M. Turner, K. S. Calhoun, & H. E. Adams (Eds.), *Handbook of clinical behavior therapy* (2nd ed.). New York: Wiley.

Bhugra, D. (1987). Homophobia: A review of the literature. *Sexual and Marital Therapy, 2,* 169–177.

Bhugra, D., & Crowe, M. J. (1995). Physical treatments of erectile disorders. *International Review of Psychiatry, 7,* 217–223.

Bhugra, D., & de Silva, P. (1993). Sexual dysfunction across cultures. *International Review of Psychiatry, 5,* 243–252.

Bhugra, D., & Wright, P. (1995). Sexual dysfunction and gay men: Diagnosis and management. *International Review of Psychiatry, 7,* 247–252.

Bishay, N. R. (1988). Cognitive therapy for sexual dysfunction: A preliminary report. *Sexual and Marital Therapy, 3,* 83–90.

Brindley, G. S. (1983). Cavernosal alpha-blockade: A new technique for investigating and treating erectile impotence. *British Journal of Psychiatry, 143,* 332–337.

Brindley, G. S. (1986). Maintenance treatment of erectile impotence by cavernosal unstriated muscle relaxant injection. *British Journal of Psychiatry, 149,* 210–215.

Bullard, D. G. (1988). The treatment of desire disorders in the medically ill and physically disabled. In S. Leiblum & R. Rosen (Eds.), *Sexual desire disorders* (pp. 109–120). New York: Guilford.

Carey, M. P., Flasher, L. V., Maisto, S. A., & Turkat, I. D. (1984). The apriori approach to psychological assessment. *Professional Psychology: Research and Practice, 15,* 515–527.

Catalan, J., Hawton, K., & Day, A. (1991). Individuals presenting without partners at a sexual dysfunction clinic: Psychological and physical morbidity and treatment offered. *Sexual and Marital Therapy, 6,* 15–23.

Cole, M. (1988). Sex therapy for individuals. In M. Cole & W. Dryden (Eds.), *Sex therapy in Britain* (pp. 272–299). Milton Keynes, UK: Open University Press.

Conte, H. R. (1986). Multivariate assessment of sexual dysfunction. *Journal of Consulting and Clinical Psychology, 54,* 149–157.

Crowe, M. J., Gillan, P., & Golombok, S. (1982). Form and content in the conjoint treatment of sexual dysfunction: A controlled study. *Behaviour Research and Therapy, 19,* 47–54.

Crowe, M. J., & Qureshi, M. J. H. (1991). Pharmacologically induced penile erections (PIPE) as a maintenance treatment of erectile impotence: A report of 41 cases. *Sexual and Marital Therapy, 6,* 273–285.

Crowe, M. J., & Ridley, J. (1990). *Therapy with couples*. Oxford, UK: Blackwell.

Dauw, D. C. (1988). Evaluating the effectiveness of the SECS surrogate-assisted sex therapy model. *Journal of Sex Research, 24,* 269–275.

Delvin, D. C. (1974). *The book of love*. London: New English Library.

Derogatis, L. R., & Melisaratos, N. (1979). The DFSI: A multidimensional measure of sexual functioning. *Journal of Sex and Marital Therapy 5,* 244–281.

de Silva, P. (1994a). Sexual dysfunction: Investigation. In S. J. E. Lindsay & G. E. Powell (Eds.), *The handbook of clinical adult psychology* (2nd ed., pp. 199–212). London: Routledge.

de Silva, P. (1994b). Sexual dysfunction: Treatment. In S. J. E. Lindsay & G. E. Powell (Eds.), *The handbook of clinical adult psychology* (2nd ed., pp. 213–228) London: Routledge.

de Silva, P. (1995). Paraphilias and sexual dysfunction. *International Review of Psychiatry, 7,* 225–229.

de Silva, P., & Dissanayake, S. A. W. (1989). The loss of semen syndrome in Sri Lanka: A clinical study. *Sexual and Marital Therapy, 4,* 195–204.

de Silva, P., & Todd, G. (1998). Sexual dysfunction in women with anorexia nervosa: Nature and treatment. *Sexual and Marital Therapy, 13,* 21–36.

Dupras, A., & Morriset, R. (1993). Sexual dysfunction among HIV positive gay males. *Sexual and Marital Therapy, 8,* 37–46.

Feldman, H. A., Goldstein, I., Hatzichristou, D. G., Krane, R. J., & McKinlay, J. B. (1994). Impotence and its medical and psychosexual correlates: Results of the Massachusetts Male Aging Study. *Journal of Urology, 151,* 54–61.

Frank, E., Anderson, C., & Rubinstein, D. (1978). Frequency of sexual dysfunction in normal couples. *New England Journal of Medicine, 299,* 111–115.

Friday, N. (1976). *My secret garden: Women's sexual fantasies*. London: Quartet.

Friday, N. (1991). *Women on top*. London: Hutchinson.

Gibson, N. B. (1992). *Love, sex and power in later life*. London: Freedom Press.

Gillan, P. (1979). Stimulation therapy for sexual dysfunction. *British Journal of Sexual Medicine, 6,* 13–14.

Gillan, P. (1987). *Sex therapy manual.* Oxford, UK: Blackwell.

Guile, L., Horne, M., & Dunston, E. (1978). Anorexia nervosa: Sexual behaviour modification as an adjunct to an integrated treatment programme. *Australian and New Zealand Journal of Psychiatry, 12,* 165–167.

Hawton, K. (1985). *Sex therapy: A practical guide*. Oxford, UK: Oxford University Press.

Hawton, K. (1992). Sex therapy: For whom is it likely to be effective? In K. Hawton & P. Cowen (Eds.) *Practical problems in clinical psychiatry* (pp. 93–104). Oxford, UK: Oxford University Press.

Hawton, K., & Catalan, J. (1990). Sex therapy for vaginismus: Characteristics of couples and treatment outcome. *Sexual and Marital Therapy, 5,* 39–48.

Hawton, K., Catalan, J., & Fagg, J. (1991). Low sexual desire: Sex therapy results and prognostic factors. *Behaviour Research and Therapy, 29,* 217–224.

Hawton, K., Catalan, J., Martin, P., & Fagg, J. (1986). Long-term outcome of sex therapy. *Behaviour Research and Therapy, 24,* 377–385

Heiman, J. R., & Lo Piccolo, J. (1988). *Becoming orgasmic.* Englewood Cliffs, NJ: Prentice-Hall.

Hoon, E., Wincze, J., & Hoon, P. (1976a). The SAI: An inventory for the measurement of female sexual arousal. *Archives of Sexual Behavior, 5,* 291–300.

Hoon, E., Wincze, J., & Hoon, P. (1976b). Psychological assessment of sexual arousal in women. *Psychophysiology 13,* 196–208.

Kanfer, F. H., & Saslow, G. (1969). Behavioral diagnosis. In C. M. Franks (Ed.), *Behavior therapy: Status and appraisal* (pp. 417–444). New York: McGraw-Hill.

Kaplan, H. S. (1974). *The new sex therapy.* London: Ballière Tindall.

Kaplan, H. S. (1987). *The illustrated manual of sex therapy* (2nd ed.) New York: Brunner/Mazel.

Karacan, I. (1978). Advances in the psychophysiological evaluation of male erectile impotence. In L. Lo Piccolo & J. Lo Piccolo (Eds.), *Handbook of sex therapy* (pp. 137–145). New York: Plenum.

Kayata, L., & Szydlo, D. (1988). Sex therapy in groups. In M. Cole & W. Dryden (Eds.), *Sex therapy in Britain* (pp. 184–203). Milton Keynes, UK: Open University Press.

Kegel, A. (1952). Sexual functions of the pubococcygeus muscle. *Western Journal of Surgery, Obstetrics and Gynaecology, 60,* 521–524.

Kinsey, A. C., Pomeroy, W. B., & Martin, C G. (1948). *Sexual behavior in the human male*. Philadelphia: Saunders.

Kolodny, R. C., Masters, W. H., & Johnson, V. E. (1979). *A textbook of sexual medicine*. Boston: Little Brown.

Lazarus, A. A. (1963). The treatment of chronic frigidity by systematic desensitization. *Journal of Nervous and Mental Disease, 136,* 272–278.

Lief, H. I., & Reed, D. M. (1972). *Sexual knowledge and attitude test (SKAT)* (2nd ed.). Centre for the Study of Sex Education in Medicine, University of Pennsylvania.

Linet, O. I., & Ogring, F. G. (1996). Efficacy and safety of intracavernosal alpostradil in men with erectile dysfunction. *New England Journal of Medicine, 334,* 873–877.

Lobitz, W. C., Lo Piccolo, J., Lobitz, G. K., & Brockway, J. (1976). A closer look at simplistic behaviour therapy for sexual dysfunction: Two case studies. In H. J. Eysenck (Ed.), *Case studies in behaviour therapy* (pp. 237–271). London: Routledge and Kegan Paul.

Locke H. J., & Wallace, K. M. (1959). Short marital adjustment and prediction tests: Their reliability and validity. *Marriage and Family Living, 21,* 251–255.

Lo Piccolo, J. (1994). Leading comment: The evolution of sex therapy. *Sexual and Marital Therapy, 9,* 5–7.

Lo Piccolo, J., & Steger, J. C. (1974). The Sexual Interaction Inventory: A new instrument for assessment of sexual dysfunction. *Archives of Sexual Behavior 3,* 585–595.

Lo Piccolo, J., & Stock, W. (1986). Treatment of sexual dysfunction. *Journal of Consulting Clinical Psychology, 54,* 158–167.

Malloy, G. L., & Herold, E. S. (1986). Factors related to sexual counselling of physically disabled adults. *Journal of Sex Research, 24,* 200–227.

Masters, W. H., & Johnson, V. E. (1966). *Human Sexual Response*. Boston: Little Brown.

Masters, W. H., & Johnson, V. E. (1970). *Human Sexual Inadequacy*. Boston: Little Brown.

McCarthy, B. W. (1977). Strategies and technique for the reduction of sexual anxiety. *Journal of Sex and Marital Therapy, 3,* 243–248.

McCarthy, B. W. (1992). Treatment of erectile dysfunction with single men. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: assessment and treatment* (pp. 313–348). New York: Guilford.

McConaghy, N. (1993). *Sexual behavior: Problems and management*. New York: Plenum.

McConaghy, N. (1996). Treatment of sexual dysfunctions. In V. B. Van Hessalt & M. Hersen (Eds.), *Sourcebook of psychological treatment manuals for adult disorders* (pp. 333–373). New York: Plenum.

Meisler, A. W., & Carey, M. P. (1990). A critical re-evaluation of nocturnal penile tumescence monitoring in the diagnosis of erectile dysfunctions. *Journal of Nervous and Mental Disease, 178,* 78–89.

Mulligan, T. (1989). Impotence in the older man. *Medical Aspects of Human Sexuality, 23* 32–36.

Nettelbladt, P., & Uddenberg, N. (1979). Sexual dysfunction and sexual satisfaction in 58 married Swedish men. *Journal of Psychosomatic Medicine, 23,* 141–147.

O'Gorman, E. (1978). Treatment of frigidity: A comparative study of group and individual desensitization. *British Journal of Psychiatry, 132,* 580–584.

Raboch, J., & Faltus, F. (1991). Sexuality in women with anorexia nervosa. *Acta Psychiatrica Scandinavica, 84,* 9–11.

Reading, A. E., & Weist, W. M. (1984). An analysis of self-reported sexual behavior in a sample of normal males. *Archives of Sexual Behavior, 13,* 69–83.

Reece, R. (1985). Group treatment of sexual dysfunction in gay men. In J. C. Gonsianek (Ed.), *A guide to psychotherapy with gay and lesbian clients* (pp. 113–129). New York: Haworth.

Reece, R. (1988). Causes and treatment of sexual desire discrepancies in male couples. In E. Coleman (Ed.), *Psychotherapy with homosexual men and women* (pp. 157–172). New York: Haworth.

Renshaw, D. C. (1988). Profile of 2376 patients treated at Loyola Sex Clinic between 1972 and 1987. *Sexual and Marital Therapy, 3,* 117.

Rosen, R. C., Riley, H., Wagner, G., Osterloh, I. H., Kirkpatrick, J., & Mishra, A. (1997). The International Index of Erectile Function (IIEF): A multidimensional scale for the assessment of erectile dysfunction. *Urology, 49,* 822–830.

Russell, R. J. H., & Wells, P. A. (1993). *Marriage and Relationship Questionnaire*. London: Hodder & Stoughton.

Rust, J., Bennun, I., Crowe, M. J., & Golombok, S. (1988). *The Golombok–Rust Inventory of Marital State*. Windsor, UK: NFER-Nelson.

Rust, J., & Golombok, S. (1986). *The Golombok–Rust Inventory of Sexual Satisfaction*. Windsor, UK: NFER-Nelson.

Rust, J., Golombok, S., & Collier, J. (1988). Marital problems and sexual dysfunctions: Are they related? *British Journal of Psychiatry, 152,* 629–631.

Schiavi, R. C. (1992). Laboratory methods for evaluating erectile dysfunction. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: Assessment and treatment* (pp. 141–170). New York: Guilford.

Scholl, G. M. (1988). Prognostic variables in treating vaginismus. *Obstetrics and Gynaecology, 72,* 353–358.

Schover, L. R. (1992). Erectile failure and chronic illness. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders* (pp. 341–367). New York: Guilford.

Schover, L. R., & Jensen, S. B. (1988). *Sexuality and chronic illness: A comprehensive approach*. New York: Guilford.

Semans, J. H. (1956). Premature ejaculation: A new approach. *Southern Medical Journal, 49,* 353–358.

Sommers, F. G. (1980). Treatment of male sexual dysfunction in psychiatric practice integrating the sexual therapy practitioner (surrogates). In R. Farleo & W. P. Asini (Eds.), *Medical sexology* (pp. 261–270). Amsterdam: Elsevier.

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family, 38,* 15–23.

Spector, I. P., & Carey, M. P. (1980). Incidence and prevalence of the sexual dysfunctions: A critical review of the empirical literature. *Archives of Sexual Behavior, 19,* 389–408.

Spence, S. H. (1985). Group versus individual treatment of primary and secondary orgasmic dysfunction. *Behaviour Research and Therapy, 23,* 539–548.

Spence, S. H. (1991). *Psychosexual therapy: A cognitive-behavioural approach*. London: Chapman and Hall.

Van Vreckem, E., & Vanderycken, W. (1994). A sexual education programme for women with eating disorders. In B. Dolan & J. Gitzinger (Eds.), *Why women? Gender issues and eating disorders* (pp. 110–116). London: Athlose Press.

Wagner, G. (1993). Injection therapy for impotence. In G. Wagner & H. S. Kaplan (Eds.), *The new injection treatment for impotence* (pp. 75–93). New York: Brunner Mazel.

Wagner, G., & Green, R. (1981). *Impotence: Physiological, psychological, surgical diagnosis and treatment*. New York: Plenum.

Weizman, R., & Hart, J. (1987). Sexual behavior in healthy married American men. *Archives of Sexual Behavior, 16,* 39–44.

Wilson, G. D. (1978). *The secrets of sexual fantasy* London: Dent.

Wincze, J. P., Bansal, S., Malhotra, C., Balko, A., Susset, J. G., & Malamud, M. (1988). A comparison of nocturnal penile tumescence and penile response to erotic stimulation during working states in comprehensively diagnosed groups of males experiencing erectile difficulties. *Archives of Sexual Behavior, 17,* 333–348.

Wincze, J. P., & Carey, M. P. (1991). *Sexual dysfunction: A guide for assessment and treatment*. New York: Guilford.

Wolfson, B., Pickett, S., Scott, N. E., DeKernion, J. B., & Rajfer, R. (1993). Intraurethral prostaglandin E-2 creams: A possible alternative treatment for erectile dysfunction. *Urology, 42,* 73–75.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition* Stanford, CA: Stanford University Press.

Woody, J. D. (1992). *Treating sexual distress*. Newbury Park, CA: Sage Publications.

Yaffe, M., & Fenwick, E. (1986a). *Sexual happiness for men*. London: Dorling Kindersley.

Yaffe, M., & Fenwick, E. (1986b). *Sexual happiness for women*. London: Dorling Kindersley.

Zilbergeld, B. (1975). Group treatment of sexual dysfunction in men without partners. *Journal of Sex and Marital Therapy, 1,* 204–214.

Zilbergeld, B. (1978). *Men and sex*. Boston: Little Brown.

Zilbergeld B. (1992). The man behind the broken penis: Social and psychological determinants of erectile failure. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: Assessment and treatment* (pp. 27–51). New York: Guilford.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.28
Relationship Problems

W. KIM HALFORD
Griffith University, Brisbane, Qld, Australia
HOWARD J. MARKMAN
University of Denver, CO, USA
and
PETER FRAENKEL
New York University, NY, USA

6.28.1 INTRODUCTION 623
6.28.2 NATURE AND SIGNIFICANCE OF MARITAL DISTRESS 624
  6.28.2.1 Prevalence of Marital Problems 624
  6.28.2.2 Effects of Relationship Problems 624
6.28.3 NATURE AND ETIOLOGY OF RELATIONSHIP PROBLEMS 626
6.28.4 BEHAVIORAL, COGNITIVE, AND AFFECTIVE COMPONENTS OF RELATIONSHIP PROBLEMS 626
  6.28.4.1 The Etiology of Relationship Distress 628
6.28.5 COUPLES THERAPY 629
  6.28.5.1 The Different Approaches to Couples Therapy 630
  6.28.5.2 Assessment of Relationship Problems and the Process of Couples Therapy 630
  6.28.5.3 Cognitive-behavioral Couples Therapy 633
  6.28.5.4 The Effectiveness of Cognitive-behavioral Couples Therapy 635
  6.28.5.5 Couples Therapy and Psychological Disorder 636
6.28.6 PREVENTION OF RELATIONSHIP PROBLEMS 638
  6.28.6.1 Prevention and Early Intervention 638
  6.28.6.2 Content of Prevention Programs 639
6.28.7 CONCLUSIONS 642
6.28.8 REFERENCES 643

> Marriage is one of the most nearly universal of human institutions. No other touches so intimately the life of practically every member of the earth's population. (Terman, 1938, p.1)

6.28.1 INTRODUCTION

An overwhelming percentage of people become involved in intimate couple relationships at some point in their lives. In Western countries over 90% of the population marry by age 50 years (DeGuilbert-Lantione & Monnier, 1992; McDonald, 1995). Even among those who choose not to marry, the vast majority engage in "marriage-like" relationships by living together in committed, couple relationships (McDonald, 1995). Expectations of couple relationships are high. In Western cultures the vast majority of adults perceive their marital relationship as their primary source of support and affection (Levinger & Huston, 1990). Most young unmarried adults expect to marry at some point in their lives, for that marriage to be lifelong, and expect their partners to show sexual monogamy, honesty, expressions of affection, intimacy, and support (Millard, 1990).

Almost all couples that marry report high levels of relationship satisfaction early in their relationship (Markman & Hahlweg, 1993). Many couples sustain a mutually satisfying relationship, and seem to derive many benefits from that relationship. Relative to other people, those in satisfying marriages have lower rates of psychological distress, higher rated life happiness, and greater resistance to the detrimental effects of negative life events (Glenn & Weaver, 1981; Gore, 1978; Gove, Hughes, & Style, 1983; Halford, Kelly, & Markman, 1997). Unfortunately, for many other couples their relationship quality and satisfaction erodes over time. Relationship problems spill over to negatively affect partners' ability to cope with parenting and work outside the home, and this has major detrimental effects on them and their children (Halford & Markman, 1996; Sanders, Nicholson, & Floyd, 1997).

Seeking help about relationship problems is very common. Relationship preparation and enhancement programs are widely becoming accepted and attended by increasing proportions of the population entering marriage (Bradbury & Fincham, 1990; Stanley, Markman, Leber, & St. Peters, 1995; Simons, Harris, & Willis, 1994). In fact, relationship preparation programs are more widely attended than just about any other form of psychoeducational program (Simons et al., 1994). Couple relationship problems are one of the most common presenting problem of adults seeking psychological assistance (Veroff, Kulka, & Douvan, 1981), and couples often seek divorce mediation when their relationships end (Emery & Wyer, 1987; Walsh, Jacob, & Simons, 1995). In this chapter we analyze the nature and significance of relationship problems, assessment of couple relationships, the conduct of couples therapy, and prevention of relationship problems.

6.28.2 NATURE AND SIGNIFICANCE OF MARITAL DISTRESS

6.28.2.1 Prevalence of Marital Problems

The most statistically reliable index of marital distress is divorce rates, and divorce has reached epidemic proportions in most Western societies. About 45% of Australian marriages, 55% of American, 42% of UK, and 37% of German end in divorce (De Gulbert-Lantoine & Monnier, 1992; McDonald, 1995). As painful as the experience of divorce is for many people, most divorcees still aspire to be in a committed relationship. In the US about 75% of divorced people remarry within five years of the end of their first marriage (Glick, 1989; Martin & Bumpass, 1989). The divorce rates of second marriages are even higher than for first marriages (Glick, 1989).

Divorce rates represent only a portion of couples experiencing relationship problems. Many couples have significant relationship distress, but opt to stay together for various reasons such as the financial implications of divorce, and personal and cultural expectations about marriage (Gottman, 1993, 1994). Surveys of representative samples of married adults show that between 80% and 85% report they are very satisfied with their current relationship (Beach, Arias, & O'Leary, 1986; Eddy, Heyman, & Weiss, 1991, Gallup, 1989; Reynolds, Rizzo, Gallagher, & Speedy, 1979; Stanley & Markman, 1996). However, satisfied partners tend to make unrealistically positive comments and predictions about their relationship functioning (Fowers, Lyons, & Montel, 1996). For example, the majority of maritally satisfied partners believe there is zero probability that they will ever divorce, despite the well-publicized evidence of how common divorce is (Fowers et al., 1996)

Furthermore, of the married people who report high relationship satisfaction, 40% also report having seriously considered leaving their current partners at some point (Gallup Poll, 1989), and over 50% are characterized by interaction patterns that erode satisfaction and predict future divorce (Stanley & Markman, 1996). Thus, even those couples who report that they are very satisfied in their relationship at any point in time are at risk for future problems and divorce. All these figures converge on the point that significant relationship problems are common in most Western societies.

6.28.2.2 Effects of Relationship Problems

Given the central importance placed upon couple relationships, it is not surprising to find that relationship distress and dissolution are experienced as extremely stressful events. In fact, after a death in the immediate family, marital distress and divorce are rated as the most severe of commonly occurring stresses experienced by adults (Bloom, Asher, & White, 1978). Relationship distress is associated with increased risk for development of a range of individual psychological disorders including depression, particularly in women (Bebbington, 1987; Coyne, Kahn, & Gotlib, 1987; Hooley, Orleay, & Teasdale, 1986), alcohol abuse, particularly in men (O'Farrell, 1989), and anxiety disorders (Craske & Zoellner, 1995), and sexual dysfunction in both sexes (Zimmer, 1983). Marital conflict is also associated with increased behavior problems and poorer psychological adjustment in the couples' children (Emery, 1982; Emery, Joyce, & Fincham, 1987; Grych & Fincham, 1990).

The causal connections between marital distress and individual psychological disorder are complex. Long-standing, psychological disorders reduce the chance of having a satisfactory marriage. For example, patients diagnosed with schizophrenia or severe personality disorders are much less likely than the rest of the population to get married, and are much more likely to get divorced if they do marry (Lange, Schaap, & van Widenfelt, 1993; Reich & Thompson, 1985). However, the association between marital problems and psychiatric disorder is not simply psychological disorder causing marital problems. For example, in cases of coexisting marital problems and depression, the marital problems often antedate the onset of depression (Birchnall & Kennard, 1983). Even when treatment produces an improvement in depressed mood there is limited effect on marital distress (Dobson, 1987; Klerman & Weissman, 1982; O'Leary & Beach, 1990), and the ongoing marital problems are associated with poor prognosis for the depression (Rousanville, Weissman, Prusoff, & Herceg-Baron, 1979). Similarly, marital problems stimulate excessive drinking (Davis, Berenson, Steinglass, & Davis, 1974), precipitate relapse by people with high alcohol dependence who have been abstinent from alcohol consumption (Maisto, O'Farrell, Connors, McKay, & Pelcovits, 1988), and are predictive of poor prognosis in alcohol treatment programs (Billings & Moos, 1983; Vannicelli, Gingereich, & Ryback, 1983). Thus, a simple unidirectional model of causality is inadequate. Marital distress and psychological disorder reciprocally influence each other.

In addition to the psychological maladjustment associated with relationship problems, relationship problems are also correlated with poorer physical health. Individuals in satisfying and supportive marriages are less likely to have major illness, and they recover better when they do become ill than individuals in distressed relationships (Burman & Margolin, 1992; Schmaling & Sher, 1997). The mechanisms linking poor physical health to relationship problems are complex and only partially understood. Relationship distress has some indirect effects mediated through health-related behaviors, such as low adherence to medical treatment regimens by those in distressed relationships (Schmaling & Sher, 1997).

Inadvertent reinforcement of illness behavior in distressed marriages and inadequate support of coping in chronic illness can have a negative impact upon partners' health (Schmaling & Sher, 1997). There are also some direct effects of relationship distress on physiological processes which impact upon health. For example, relationship conflict is associated with immunosuppression (Kiecolt-Glaser et al., 1993), elevated stress hormone levels (Keicolt-Glaser et al., 1996), elevated blood pressure in people with essential hypertension (Ewart, Taylor, Kraemer, & Agras, 1991), and possibly atherosclerosis (Gottman, 1990; Medalie & Goldbourt, 1976). Each of these physiological effects are likely to increase risk for serious health problems.

Relationship problems at home spill over to work. More specifically, for men relationship conflict at home is associated with days lost at work and poor work performance (Thompson, 1997). The economic cost of the effects of relationship problems are extensive.

Relationship violence has a major effect on the health and well-being of many partners in committed relationships. About 25% of all marriages have at least one episode of interspousal physical aggression at some point (O'Leary et al., 1989; Strauss, Gelles, & Steinmetz, 1980), and this violence almost always is associated with relationship distress. Over half of maritally distressed couples presenting for therapy report having experienced interpartner violence in the previous year (O'Leary & Vivian, 1990). Whilst men and women are equally likely to engage in aggressive acts such as hitting, pushing, or slapping (Strauss & Gelles, 1986; O'Leary et al., 1989), men are much more likely to engage in severe violence, and women are at particular risk for being injured or even killed by their partner (Koop, 1985; Stets & Strauss, 1990). Even when physical injuries are less severe, violence can have severe negative consequences for female victims. Women repeatedly assaulted by their partners are at high risk of developing depression, alcohol abuse, psychosomatic disorders, and are high users of the healthcare system (Cascardi, Langhinrichsen, & Vivian, 1992; Jaffee, Wolfe, Wilson, & Zak, 1986; Stets & Strauss, 1990). Relationship aggression is also linked to child abuse (Grych & Fincham, 1990), development of antisocial behavior in male offspring (Grych & Fincham, 1990), and increased risk of children entering a violent relationship as an adult (Widom, 1989).

6.28.3 NATURE AND ETIOLOGY OF RELATIONSHIP PROBLEMS

The common element to clinical presentations of relationship problems is almost always dissatisfaction by at least one partner with the relationship. Attempts to define and measure relationship problems have taken relationship (marital) satisfaction, as assessed by self-report inventories, as the ultimate criterion (Weiss & Heyman, 1997). The most widely used self-report inventories are the Locke–Wallace Marital Adjustment Test (Locke & Wallace, 1959), and an expanded revision known as the Dyadic Adjustment Scale (Spanier, 1976). These two scales, and many similar scales, have been subjected to repeated criticism (e.g., Heyman, Sayers, & Bellack, 1994) for confounding relationship satisfaction, measured by items such as "overall how would you rate your marital happiness," with adjustment processes alleged to influence satisfaction, measured by items such as "how often do you and your partner disagree about finances?" While the collapsing of the constructs of relationship satisfaction and adjustment seems conceptually unsound, factor analyses of measures of marital adjustment and dissatisfaction consistently show that partners make unidimensional, global evaluations of their relationship problems (Eddy, Heyman, & Weiss, 1991; Heyman et al., 1994). In essence, partners who are dissatisfied with their relationships tend to report that just about anything that could be negative about their relationship is negative.

The finding that reported relationship distress seems to reflect an overriding global evaluation of negative aspects of the relationship does not mean that relationship quality is unidimensional. In the late 1990s work shows that couples distinguish between dissatisfaction about negative aspects of the relationship (e.g., the level of distress associated with conflict) and satisfaction with positive aspects of the relationship (e.g., the sense of satisfaction with expressions of love) (Fincham, Beach, & Kemp-Fincham, 1997). Some couples present with severe relationship distress and little positive feeling about their partner or the relationship, while other couples are distressed by the negative aspects of the relationship associated with dissatisfaction but still retain some positive feelings about the partner. The latter group of couples seem to benefit more from couples therapy (Hahlweg, Schindler, Revenstorf, & Brengelman, 1984). In either group of couples their initial focus often is on reducing negative behaviors associated with dissatisfaction, but successful therapy also needs to prompt couples to attend to developing positive aspects of the relationship.

Active consideration of separation is common in couples with relationship problems, (Weiss & Cerreto, 1980). However, the association of relationship problems, consideration of separation, and divorce is not straightforward. Most couples who eventually divorce go through a protracted process which proceeds from worsening relationship problems, through active consideration of separation, to actual separation, and eventually to divorce (Gottman, 1990, 1993). Often in couple presentations one or both partners are ambivalent about the relationship and its future. Ambivalence can be a major challenge to the partners. They risk neither proceeding with separation nor working to improve a distressed relationship, and by default the distress of the status quo is maintained. The systematic assessment of relationship problem areas and strengths, combined with conjoint goal setting, often serves to overcome this inertia (Halford, Osgarby, & Kelly, 1996; Weiss & Halford, 1995). We will say more on this in the section on relationship assessment (Section 6.28.5.2).

6.28.4 BEHAVIORAL, COGNITIVE, AND AFFECTIVE COMPONENTS OF RELATIONSHIP PROBLEMS

Problems in communication are the most frequently cited specific complaint by couples seeking therapy, with up to 90% of distressed couples citing these difficulties as a major issue in the relationship (Bornstein & Bornstein, 1986). Both independent observers and spouses report communication deficiencies are associated with relationship distress (Weiss & Heyman, 1997). When discussing problem issues, distressed partners are often hostile, and criticise and demand change of each other (Christensen & Shenk, 1991; Gottman & Krokoff, 1989; Gottman, 1994; Halford, Hahlweg, & Dunne, 1990; Heavey, Christensen, & Malmuth, 1995; Notarius & Markman, 1993). Distressed couples also do not actively listen to their partner when discussing problems (Halford et al., 1990; Jacobson, McDonald, Follette, & Berley, 1985; Weiss & Heyman, 1990), and tend to withdraw from problem discussions (Christensen & Shenk, 1991; Gottman, 1994; Gottman & Krokoff, 1989; Halford, Gravestock, Lowe, & Scheldt, 1992; Heavey, Christensen, & Malmuth, 1995). Contentious relationship issues are less likely to be resolved by discussion in distressed couples than non-distressed couples (Halford et al., 1992).

Distressed couples are highly reactive at an emotional level to their partners' behavior, and show significantly higher rates of negative reciprocity during interaction than nondistressed couples (e.g., Gottman, Notarius, & Markman, 1977; Schaap, 1984). In observational studies of communication, the conditional probabilities of distressed partners responding negatively to their partner's negativity is much higher than the conditional probabilities for nondistressed partners (e.g., Halford et al., 1990). In addition to this negative reciprocity, relationship distress is also associated with high levels of psychophysiological arousal during interaction (e.g., Gottman & Levenson, 1988). This arousal is aversive, which may explain the higher rates of withdrawal during problem-focused discussions by maritally distressed partners (Christensen & Shenk, 1991; Gottman & Krokoff, 1989). In any case, both the extent of arousal and the frequency of withdrawal prospectively predict deterioration in marital satisfaction (Gottman, 1993; Gottman & Krokoff, 1989; Heavey, Layne, & Christensen, 1993; Heavey et al., 1995).

Another common complaint of couples seeking relationship therapy is the negativity of their day-to-day interactions (Halford, in press). Using behavioral checklists in which partners monitor their spouses' behavior there is a well-replicated finding that monitored daily behaviors correlate with relationship satisfaction (Birchler, Weiss, & Vincent, 1975; Halford & Sanders, 1988; Jacobson, Follette, & McDonald, 1982; Johnson & O'Leary, 1996). More specifically, relative to maritally satisfied couples, distressed couples report higher rates of negative, displeasing behaviors by their spouse and fewer positive, pleasing behaviors (Birchler et al., 1975; Halford & Sanders, 1988; Jacobson et al., 1982; Johnson & O'Leary, 1996). Furthermore, distressed couples tend to reciprocate on a *quid pro quo* basis the behaviors of their spouse. In other words, in a distressed relationship partners tend only to be positive if their partner recently has been positive, and if one partner behaves negatively the other often responds negatively immediately (Birchler et al., 1975; Jacobson et al., 1982). In contrast, satisfied couples' behavior is less contingent on the preceding partners' behaviors; satisfied couples tend to be positive irrespective of their partners' prior actions.

There also is evidence that distressed couples' perceptions of their partners' behavior are negatively biased. Distressed couples disagree to a greater extent with both objective observers (Robinson & Price, 1980) and with each other (Christensen & Nies, 1980; Jacobson & Moore, 1981) regarding the occurrence of particular behaviors in their relationship. More specifically, they tend to overestimate the frequency of negative partner behaviors. The nature of relationship distress cannot be understood simply in terms of the behaviors occurring during marital interaction, but also requires attention to the cognitive appraisal by partners of that interaction.

Distressed couple have a number of characteristic cognitions about their relationships (Beaucom, Epstein, Sayers, & Sher, 1989). Maritally distressed couples selectively attend to their partner's negative behavior (Eidelson & Epstein, 1982; Floyd & Markman, 1983; Jacobson & Moore, 1981), and selectively recall such negative behavior (Halford & Osgarby, 1996). In contrast, maritally, satisfied partners tend to overlook negative behaviors by their spouses (Gottman et al., 1977; Notarius, Benson, Sloane, Vanzetti, & Horyak, 1989), to have an unrealistically positive view of their partners and relationships (Fowers, Applegate, Olson, & Pomerantz, 1994), and to selectively recall positive aspects of relationship interaction (Halford & Osgarby, 1996).

Another characteristic of distressed couples in the cognitive domain is holding unrealistic beliefs about relationships and partners. More specifically, relative to happy couples, distressed couples are more likely to believe that any form of disagreement is destructive, that change by partners is not possible, and that rigid adherence to traditional gender roles is desirable (Baucom & Epstein, 1990; Eidelson & Epstein, 1982). Distressed couples also report that their relationships often violate standards about how they think their relationship should be (Baucon et al., 1996). For example, distressed women report that their partners do not share power within the relationship in the manner the women believe they should, and men believe their partners should invest more time and energy in the relationship than they do (Baucom et al., 1996).

Distressed couples attribute the causes of relationship problems to stable, internal, negative, and blameworthy characteristics of the partners (Bradbury & Fincham, 1990; Fincham & Bradbury, 1992). For example, a partner arriving home late from work may be perceived as "a generally selfish person who doesn't care about the family" by a maritally distressed partner. The same behavior may be attributed by a maritally satisfied partner as the spouse "struggling to keep up with a heavy load at work, and being subject to lots of pressure from the boss." The process of attributing much or all of the relationship problems to their partners leaves most people with relationship distress feeling powerless to improve their relationship (Vanzetti, Notarius, & NeeSmith, 1992). A key element of couples therapy is to enhance each partners' sense of relational efficacy or capacity to improve their relationship through their own actions (Halford, Sanders, & Behrens, 1994).

One additional cognitive characteristic of a distressed couple is that they expect negative outcomes from interaction with their partners. Distressed couples report that prior to a discussion they expect not to be able to resolve problem issues in their relationships (Vanzetti et al., 1992). In anticipation of a problem-solving discussion, maritally distressed partners show high physiological arousal (Gottman, 1994), negative affect, and become primed to access negative evaluative judgements about their partner and the relationship (Fincham, Garnier, Gano-Phillips, & Osborne, 1995).

The cognitive characteristics of distressed couples mediate their subsequent behavior toward their partners. For example, the occurrence of negative attributions is associated with subsequent behavioral negativity (Bradbury & Fincham, 1992). In unhappy couples, negative thoughts about the partner predict future negative behaviors better than predictions from previous behavior (Halford & Sanders, 1990), suggesting these cognitions are more than just the consequences of negative marital behavior. In other words, relationally distressed partners seem to respond to their subjective perceptions and memories of relationship interactions, and these perceptions and memories are negatively biased.

6.28.4.1 The Etiology of Relationship Distress

There are over 100 published studies assessing the longitudinal course of couple relationship satisfaction and stability (Karney & Bradbury, 1995). This comprehensive literature can be usefully summarized in terms of three broad classes of variables which impact upon the etiology of relationship problems: adaptive processes within the couple system, stressful events impinging upon the couple system, and enduring individual vulnerabilities of the partners (Bradbury, 1995).

Adaptive processes refer to the cognitive, behavioral, and affective processes that occur during couple interaction. Certain deficits in these adaptive processes seem to predispose couples to relationship problems. More specifically, deficits in communication and management of negative affect and conflict observed in engaged couples prospectively predict divorce and relationship dissatisfaction over the first 10 years of marriage (Markman & Hahlweg, 1993). Dysfunctional communication and negative affect regulation in engaged couples also predicts the development of relationship verbal and physical aggression in the first few years of marriage (Murphy & O'Leary, 1989; O'Leary et al., 1989), at least for mild to moderate severity aggression. Relationship aggression is often established early in the relationship, and usually continues and escalates once established (Murphy & O'Leary, 1989; O'Leary et al., 1989).

It is noteworthy that the communication and conflict management deficits observed in some engaged couples do not correlate with their reported relationship satisfaction or commitment at the time (Markman & Hahlweg, 1993; Sanders, Halford, & Behrens, 1996). It seems that these communication difficulties do not stop couples from forming committed relationships, but the difficulties do predispose couples to develop relationship problems later. In couples who have been married for some time, these same communication difficulties predict deterioration in relationship satisfaction and decreased relationship stability (Gottman, 1993, 1994)

The beliefs and expectations individuals have when entering into relationships and marriage predict the risk of divorce in the first few years of marriage (Olson & Fowers, 1986; Olson & Larsen, 1989). Couples characterized by unrealistic expectations and beliefs in areas such as importance of communication, appropriate methods of conflict resolution, importance of family and friends, and gender roles, have higher rates of erosion in relationship satisfaction than couples not so characterized. Negative attributional patterns in which partners attribute blame for relationship problems to stable, negative characteristics of their spouse also prospectively predict deterioration in relationship satisfaction (Fincham & Bradbury, 1990). Thus, certain communication and cognitive characteristics of the couple's adaptive processes predate, and prospectively predict, relationship problems.

Stressful events refer to the developmental transitions and acute and chronic circumstances which impinge upon the couple or individual partners. Relationship problems are more likely to develop during periods of high rates of change and stressful events (Karney & Bradbury, 1995). For example, the early stages of marriage, including transition to parenthood, is often associated with decline in couple relationship satisfaction (Cowen & Cowen, 1992), as is an increase in work demands (Thompson, 1997). Retirement is another major transition for couples which can be associated with relationship distress (Dickson, 1997). One partner developing a major health problem also puts couples at increased risk for relationship and sexual problems (Schmaling & Sher, 1997).

A common stressful transition worthy of special mention is entering a second marriage. Second marriages in which there are dependant children from an earlier relationship break down at very high rates (Booth & Edwards, 1992; Martin & Bumpass, 1989). Negotiating parenting roles in step-families is a common source of interpartner conflict, and unresolved differences in this area are the most common stated reason for relationship breakdown in step-families (Lawson & Sanders, 1994). Moreover remarital partners tend to repeat the same patterns of negative interactions with their new partners despite efforts to the contrary (Prado & Markman, in press).

In general, couples with less robust adaptive processes are believed to be particularly vulnerable to the negative effects of a range of stressful events (Markman, Halford, & Cordova, 1997). In particular, couples who lack communication skills, or who have inflexible or unrealistic expectations of relationships, find it hard negotiate the changes required to adapt to major life transitions (Markman, Stanley, & Blumberg, 1994). For example, one of us (WKH) is studying couples where the women were recently diagnosed with breast or gynecological cancer. In couples with good communication and effective mutual support the adversity of cancer diagnosis and treatment seems to bring the couples closer together and reinforce the relationship bonds. In contrast, couples with poor adaptive processes show deterioration in their relationships and poor individual coping with the cancer.

Enduring vulnerabilities refer to the stable historical, personal, and experiential factors which each partner brings to a relationship (Bradbury, 1995). Family of origin experiences have been widely studied as historical factors which correlate with risk of relationship problems. For example, the adult offspring of divorce are more likely than the rest of the population to divorce (Glenn & Kramer, 1987), and interparental aggression is associated with increased risk for having an aggressive relationship as an adult (Widom, 1989). The mechanisms by which exposure to parental divorce or aggression may impact upon subsequent adult relationships is becoming clearer. Exposure to parental divorce is associated with more negative expectations of marriage (Black & Sprenkle; 1991; Gibardi & Rosen, 1991; Van Widenfelt, Schaap, & Hosman, 1996), and with observable deficits in communication and conflict management in couples prior to marriage (Halford et al., 1994). Adult offpsring of parents who were aggressive also show deficits in communication, and conflict management skills in dating and marital relationships (Sanders, Halford, & Behrens, 1998; Skuja & Halford, 1998). Negative expectations and communication deficits may well be learned from the parents' relationships and subsequently these learned behaviors impact negatively upon the adult relationships of the offspring. The argument that communication difficulties may be acquired through observation and interaction with parents is supported by a finding from Howes and Markman (1991). They found couple communication style assessed premaritally predicted subsequent communication style when the partners become parents and were interacting with their children (Howes & Markman, 1991).

The association between personality variables and relationship problems has been widely studied. Normal personality variations do not seem to contribute much variance to relationship satisfaction (Gottman, 1994; Karney & Bradbury, 1995; Notarius & Markman, 1993). One exception is that low ability to regulate negative affect (high neuroticism) consistently has been found to predict higher risk for relationship problems and divorce (Karney & Bradbury, 1995). How this personality characteristic may impact upon relationship problems is not yet understood.

Another major risk indicator for relationship distress and divorce is past or present history of psychological disorder. Higher rates of relationship problems and divorce consistently have been reported in populations with severe psychiatric disorder (Halford, 1995), and in people with depression, alcohol abuse, and some anxiety disorders (Emmelkamp, De Haan, & Hoogduin, 1990; Halford, Kelly, Bouma, & Young, in press; Halford & Osgarby, 1993; O'Farrell & Birchler, 1987; Reich & Thompson, 1985; Ruscher & Gotlib, 1988; Weissman, 1987). As described earlier in this chapter, relationship problems and individual problems can both exacerbate each other (Halford et al., in press). In addition, certain personal vulnerabilties may dispose people to both psychological disorders and relationship problems. For example, deficits in interpersonal communication and negative affect regulation are risk factors that predict the onset of both alcohol abuse (Block, Block, & Keyes, 1988) and relationship problems (Markman & Hahlweg, 1993). This common risk factor might be part of the explanation for the common co-occurrence of relationship and alcohol problems.

6.28.5 COUPLES THERAPY

Intervention with relationship problems can potentially target couples' adaptive processes, stressful events, or enduring vulnerabilities, as each of these classes of variables influence relationship problems. The focus of all interventions evaluated in research has been on modifying couples' adaptive processes, either through conjoint therapy for couples with existing relationship problems (e.g., Baucom & Epstein, 1990; Halford, in press), or through brief programs to prevent the development of relationship problems (Halford & Behrens, 1996; Markman & Hahlweg, 1993). Often interventions are adapted to take account of the enduring vulnerabilities of the partners. For example, there has been considerable research on therapy with couples in which one partner has depression (Beach, Sandeen, & O'Leary, 1992) or alcohol abuse (O'Farrell & Rotunda, 1997). In the 1990s some work has focused on developing couples' adaptive processes to manage particular life transitions and stressful events, such as being in a step-family (Lawton & Sanders, 1994). In the rest of this chapter we focus on interventions targeting change in couple adaptive processes.

6.28.5.1 The Different Approaches to Couples Therapy

A number of different approaches to couples therapy, including behavioral, cognitive, cognitive-behavioral, emotion-focused, and insight-oriented couples therapy, have been demonstrated to improve the relationship satisfaction of the majority of couples who present for therapy (Baucom & Epstein, 1990; Greenberg & Johnson, 1988; Hahlweg & Markman, 1988; Halford, Sanders, & Behrens, 1993; Shadish et al., 1993; Snyder, Wills, & Grady-Fletcher, 1991a). All these approaches share the characteristics of the use of systematic assessment and goal setting at the beginning of therapy, and a focus on conjoint therapy targeting changes in the couples' adaptive processes.

The different theoretical approaches to couples therapy which have empirically validated efficacy have placed different emphases on the processes alleged to produce change in couples' adaptive processes. In insight-oriented and emotion-focused couples therapy emphasis was placed on changing partners' subjective experience of relationship interactions, usually through insights gained into the subjective significance of particular relationship events (Greenberg & Johnson, 1988; Snyder, Wills, & Grady-Fletcher, 1991b). In contrast, in behavioral approaches to couples therapy the original emphasis was on changing relationship interaction through procedures such as behavioral contracting and skills training in areas such as communication and problem-solving (Jacobson & Margolin, 1979). Cognitive behavioral couples therapy combined attempts to change the subjective experience of relationship interactions through cognitive change procedures and attempts to change relationship interactions with behavioral procedures (e.g., Baucom & Epstein, 1990; Halford et al., 1993).

The distinctions between these different empirically validated approaches to couples therapy are becoming somewhat blurred. Treatment manuals written for insight-oriented couples therapy (Snyder & Wills, 1989) include considerable use of behaviour change procedures that seem identical to those used in cognitive-behavioral couples therapy (Jacobson, 1991; Markman, 1991). Treatment manuals published by cognitive-behavioral couples therapists (Christensen, Jacobson, & Babcock, 1995; Weiss & Halford, 1995) describe the limitations of relying on changing relationship interactions and incorporate procedures acknowledged as very similar to those advocated by emotion-focused therapists to change the subjective experience of relationship interactions.

While there is some blurring of the distinctions between different theoretical approaches to couples therapy, there are still some important differences in such approaches. The most widely used approaches to couples therapy in clinical practice are systemic and strategic approaches (Boughner, Hayes, Bubenezer, & West, 1994), and there are a diversity of procedures used within these approaches (Fraenkel, 1997). There is little evidence by which to judge how useful these procedures are. In contrast, relatively few practitioners use the most extensively researched approach to couples therapy: cognitive-behavioral couples therapy (CBCT) (Markman, Halford, & Cordova, 1997; Shadish et al. 1993). As we believe in practice being guided by research, the couples-therapy approach outlined in this chapter is based on the CBCT approach, which has developed from the research foundation we reviewed earlier in this chapter.

6.28.5.2 Assessment of Relationship Problems and the Process of Couples Therapy

Clinical descriptions of cognitive-behavioral couples therapy emphasize the importance of developing therapeutic alliances with each partner, developing a shared understanding of relationship problems and goals which promotes adaptive change, and explicit negotiation with couples about their roles in therapy (e.g., Beach, Sandeen, & O'Leary, 1990; Baucom & Epstein, 1990; Christensen et al., 1995; Weiss & Halford, 1995). Assessment is crucial in achievement of these process outcomes in couples therapy. In essence, the therapist seeks through assessment both to establish an empathic understanding of each partner's experience of the relationship and to promote a shared conceptualization of problems in terms of relationship interactions. The shared conceptualization is the basis for the participants to negotiate the goals of therapy and to determine how those goals are to be achieved.

Establishing a shared conceptualization of relationship problems sometimes is a difficult therapeutic task. Therapists tend to see relationship problems in terms of the adaptive processes occurring between the partners. In contrast, couples often enter therapy conceptualizing their relationship problems as due to stable, global negative characteristics of their partners (Fincham & Bradbury, 1990), and believe there is little they can do individually to improve their relationship.

To achieve a shared, productive relationship focus in couples therapy the therapist initially assesses the presenting concerns from the perspective of each partner. Often these initial descriptions by clients are critical and blaming of the partner (e.g., "he does not communicate," "she is too demanding"). The therapist then uses a variety of strategies to promote a relationship focus. For example, strategic use of questioning can promote attention to relationship interaction (e.g., "How do you two resolve conflict?" "As a couple how do you ensure you have quality time together?"). Particular assessment tasks also prompt attention to relationship interaction. For example, undertaking certain communication tasks can be used to identify strengths and weaknesses in communication. Reframing summaries also can foster a relationship focus. For example, suppose a couple argue about parenting. One partner may present this issue as "he is too soft on the children when they misbehave." The spouse might describe the issue as "she is too harsh in her discipline of the children." The therapist may summarize this as: "as a couple you struggle to agree on the best ways to manage your children's behavior." The therapist's summary reframes the issue as a relationship challenge the couple can work on together.

In CBCT there is a strategic separation of assessment and therapy. At the completion of assessment the therapist provides the couple with structured feedback on the results of the assessment. This feedback summarizes and integrates the assessment information and focuses the results on how the problems can be conceptualized within a relationship framework. The process typically involves the therapist presenting the results of the assessment to the couple, one assessment instrument at a time. The therapist continually checks that the descriptions and conclusions being reached are accurate according to each partner. After presenting all the key findings the therapist summarizes the results and discusses with the couple possible goals for therapy. The most commonly identified goals are improving communication, controlling conflict, enhancing quality time together, and renegotiating key relationship responsibilities. Other common goals include working together to improve parenting difficulties, enhancing support for each other, and enhancing the expression of intimacy and closeness.

After developing a shared relationship focus, the next step in couples therapy is to have each partner define what she or he can do to change problematic interactions. This emphasis on self-change (also known as self-regulation) focuses therapy on that which each client has the most direct control over, namely their own behavior (Halford, Sanders, & Behrens, 1994). This is not to say that changes in the partner are unimportant, but rather that the most productive method of achieving change is for each partner to focus on their own opportunities to change.

For some couples the process of assessment, plus one or two sessions focused on self-directed goal setting and change, can be sufficient to improve a distressed relationship (Halford et al., 1996). For other couples, assessment is just the beginning of a more extensive therapeutic process, but serves the crucial function of determining the directions for therapy.

Assessment provides two kinds of information: content and process. Content information refers to the reports of the concerns of the partners, such as the specific content of conflicts, or their current thinking or actions about separation. Process information samples how the couple respond to the various probes or tasks the therapist initiates. For example, if the couple are asked to discuss an issue of relevance to the relationship, the therapist may be particularly interested in either the process of interchange between the partners (e.g., the degree of effective listening to each other's perspective) or the actual content of the discussion.

Many therapists rely exclusively on conjoint interviews with the couple to assess the content of relationship problems (Boughner, Hayes, Bubenezer, & West, 1994). This is a mistake. Content information obtained from conjoint interviews is less reliable than information derived from either individual interviews or self-report inventories (Haynes, Jensen, Wise, & Sherman, 1981). For example, physical abuse is dramatically under-reported in conjoint interviews relative to either self-report inventories or individual interviews (O'Leary, Vivian, & Malone, 1992). Furthermore, encouraging couples to recount lengthy lists of complaints about their partners in the presence of the partner rarely promotes a collaborative set to move therapy forward. Consequently, we recommend that a combination of self-report inventories and individual interviews be used to assess content information, and that conjoint interviews be focused on assessment of process.

A structure for an initial interview which illustrates the combination of process and content assessment is outlined in Table 1. The session begins with a brief conjoint interview in which the agenda for the session is negotiated. At the beginning the therapist should ask the couple an open-ended question about the reasons they are seeking assistance. When posing the question it is often helpful for the therapist to get eye contact with both partners, and then look at a point midway between the partners as the question is finished. The goal at this point is to establish if the couple can collaborate and conjointly tell a coherent story about their relationship problems. If each partner allows the other to speak, and there is mutual respect, then a conjoint intake interview may be possible. More commonly, in couples with relationship problems, the response to this initial probe is escalating conflict in the session, with each partner being blaming and negative toward the other. The therapist should intervene promptly to stop such escalation. The therapist also needs to note that the couple need to develop a mutually acceptable way of talking about their problems if therapy is to proceed successfully.

Table 1 Structure of couples initial interview.

Initial orientation (10 minutes)
  Introduction
  Present structure of session
  Ask for a brief/joint statement of problem

Individual interviews (30 minutes each)
  Current status of relationship
  Sources of difficulty
  Specific areas:
    behavior changes for self/partner
    communication
    time use
    sex
    external stressors
    individual problems
    domestic violence
    separation
    commitment

Joint session (20 minutes)
  Review individual material
  Describe behavioral couples therapy approach
  Review suitability of behavioral couples therapy
  Logistics of therapy

The second phase of the intake interview is then done with each individual partner, covering the key areas identified in Table 1. The goal in this phase is to identify key areas of concern, and to establish an empathic relationship with each partner. The third phase brings the couple back together again. In the third phase the therapist summarizes the key concerns identified by each partner. These summaries need to express the concerns in a manner consistent with the experience of each partner, and in a manner that allows mutual agreement that particular areas need attention. For example, the therapist might describe areas of disagreement as identified areas requiring further negotiation. Both partners can agree these are issues that they need to resolve.

Table 2 is a summary of key areas of relationship problems which need assessment, and some useful assessment instruments for each of those areas. The process of assessment typically involves each partner completing the various self-report inventories and behavioral monitoring tasks between therapy sessions. The therapist reviews and discusses the information each partner provides. In addition, the therapist often sets the couple particular tasks within sessions. A common assessment procedure is to have couples undertake one or more communication tasks, such as discussing a source of disagreement within the relationship. This is best done with the therapist outside the room observing via one-way screen or videorecording where available. The goal is to assess how the couple communicate and manage conflict. Further detail on these assessment tasks are available in O'Leary (1977), Baucom and Epstein (1990), and Halford (in press).

The assessment process is more than just passive collection of information. In the process of assessment the therapist is working with the partners to develop an agreed on concept of the relationship problems which facilitates change. The completion of assessment tasks often prompts each partner to consider aspects of their relationship in new ways. For example, as noted earlier, many partners in distressed relationships attribute the causes of their relationship problems to their partner, and feel powerless to change the things that distress them. In a version of the Areas of Change Questionnaire (Weiss & Halford, 1995), each partner is asked to identify behaviors they personally need to change to enhance their relationship. Just posing the question of self-change in a structured way can help the person to generate change ideas. Similarly, the experience of a communication assessment task can prompt partners to identify aspects of their own communication behavior they wish to change.

Table 2 Key areas of assessment for couples therapy.

Area	*Example measure*	*Explanation*
Relationship satisfaction	Dyadic Adjustment Scale (Spanier, 1976): 31-item self-report inventory	Level of distress, global evaluations of relationship
Divorce potential	Marital Status Inventory (Weiss & Ceretto, 1980); a 14-item rating scale	Steps toward separation, high steps predict poor prognosis
Behavior exchange	Areas of Change Questionnaire (Weiss & Perry, 1983); 34 items rating extent of requested behavior change by a partner	Current patterns of daily interaction, identify behavior change preferences of each client
Communication and conflict management		The most common presenting concern, and a key area requiring attention in much couples therapy
Aggression and violence	Conflict Tactics Scale (Strauss, 1979)	Verbal, psychological, and physical aggression between partners
Relationship standards and beliefs	Baucom et al. (1996)	The expectations partners have of each other and relationships. Often unrealistic beliefs can be a source of relationship problems

The assessment process needs to move couples stuck in a partner-blaming focus for distress to a relationship focused view of the problems. Ultimately the goal is to have each partner take responsibility for self-directing personal change to enhance the relationship (Halford et al., 1994).

6.28.5.3 Cognitive-behavioral Couples Therapy

Based upon the premise that a relationship is defined by the exchange of behaviors between partners, CBCT began as the application of behavioral contracting to the treatment of relationship problems. Couples were trained to monitor their partners' behavior and, based on such assessments, contingency contracts were developed to reduce displeasing and increase pleasing behaviors within the relationship (e.g., Azrin, Naster, & Jones, 1973; Stuart, 1969; Weiss, Hops, & Patterson, 1973). Such contracting initially stressed tightly structured *quid pro quo* agreements, in which spouses were taught systematically and immediately to reward desired behavior from the partner (Azrin et al., 1973; Stuart, 1969). This was later replaced by unilateral "good faith" contracts in which partners were asked to undertake positive change for the good of the relationship (e.g., Gottman, Notarius, Markman, & Gonso, 1976; Weiss, Birchler, & Vincent, 1974). Subsequently, emphasis has been placed on each partner actively seeking out information in order to self-select and implement self-change goals to enhance the relationship (Halford et al., 1994; Weiss & Halford, 1995). Although the details of the procedures have been refined, an emphasis on changing relationship behaviors remains an important element of how CBCT attempts to alter couples' adaptive processes.

In Table 3 some of the key behaviors identified as promoting long-term satisfying relationships are identified. In couples therapy partners are helped to self-identify behaviors that they wish to change within these different domains to enhance their intimacy with their partner. There are several ways of doing this. For example, partners can be asked to self-monitor behaviors within a given class and then to identify behaviors they wish to increase in that domain. A second example is the Caring Days Exercise (Weiss & Halford, 1995). In this procedure each person is encouraged to identify some small specific behaviors which they can do which demonstrate caring for their partner. To enhance their ability to self-select appropriate caring behaviors, partners can be assisted to be more creative in generating ideas. For example, to enhance creativity people can brainstorm ideas, they can ask their partner what would demonstrate caring effectively to them, they can ask friends about caring behaviors that they engage in within their relationships, they can observe others demonstrating caring in their normal day-to-day lives, or they can read through checklists of ideas and suggestions provided by the therapist. The self-regulatory focus encourages each person to take individual responsibility for being creative in identifying behaviors that they will engage in and to set themselves tasks of implementing behavioral changes.

A second element of CBCT is communication and problem-solving skills training. These skills are conceptualized as providing couples with the means to enhance intimate communication and resolve their current and future sources of conflict (Jacobson & Margolin, 1979; Notarius & Markman, 1994). In most applications of CBCT the communication skills targeted in training have been identified by the therapist based upon contrasting the couple's current communication with a model of adaptive marital communication. The models of adaptive marital communication were derived, in large part, from research contrasting the communication behaviors of maritally distressed and nondistressed couples in problem-solving interactions within research laboratories (see Weiss & Heyman, 1990, 1997, for reviews of this literature). Often CBCT would teach couples a relatively fixed curriculum of skills (e.g., paraphrasing, asking open-ended questions, behavioral pin-pointing), based on the assumption that each of these skills were adaptive as communication skills.

Research shows that there is no clear relationship between the use of particular, specific communication behaviors (e.g, paraphrasing, minimal encouragers) and relationship problems (Halford et al., 1990, 1993). Rather, it seems that there are a few broad classes of adaptive relationship communication behaviors, such as validation (active, positive listening to partner) and positive engagement, and use of any of a wide variety of behaviors within these broad classes is associated with improvements in relationship problems (Sayers, Baucom, Sher, Weiss, & Heyman, 1991). The specific behaviors within the broad classes which are functional vary across relationships, time and settings (Halford, Gravestock, Lowe, & Scheldt, 1991). In other words, different communication styles suit different relationships and circumstances, but we can be reasonably confident that an adaptive communication style will need to involve validation and positive engagement. Furthermore, successful couples do not always use these skills after intervention, but do so when needed (Hahlweg & Markman, 1996).

A crucial skill in couple interaction is management of negative affect. In successful long-term relationships, partners must both be able to manage their own negative feelings and respond constructively to their partner's negative feelings (Markman, Stanley, & Blumberg, 1944). One negative interaction erodes the effect of 5–10 previous positive interactions (Notarius & Markman, 1994), so preventing destructive negativity is crucial.

Table 3 Classes of behavior most strongly related to marital satisfaction.

Class of behavior	*Examples*
Affection	Saying "I love you" Giving a hug or kiss Enjoying a shared laugh or joke Saying they enjoy partner's company
Respect	Listening to the partner's opinion Telling partner of admiration/respect Showing confidence in partner's abilities Introducing partner to others with pride
Support	Doing errands for partner Making self available to do work for partner Asking partner about their day Doing something to save partner time/energy
Communication of ideas	Telling partner about their day Discussing topical events Giving an opinion Talking about mutual interest(s)
Shared quality time	Spending an hour or more just talking Work together on a project Take a drive or walk Go out together, just the two of you Discuss personal feelings

Applying the self-regulation approach to communication skills, each partner is assisted to self-select goals for changing his or her own communication, and to self-evaluate his or her own communication. For example, clients review their own couples discussions (e.g., an audiotape of a conversation at home on a difficult topic) with the therapist in sessions, with partners each focusing on their own communication. This maintains the therapeutic focus on what the client can change (i.e., his or her own behavior). Based on the assumption that adaptive marital communication is defined by its functional impact within the relationship, self-directed attempts at changes in communication are seen as behavioral experiments. Consequently, when a partner makes a change in communication that produces a negative outcome (e.g., making an assertive rather than aggressive request for change by the partner when discussing a particular issue elicits verbal abuse from the spouse), this shows that assertion was not adaptive in this context, and a different approach is needed.

A third element of CBCT is altering subjective experience of relationship interaction. Although CBCT always included recognition of the importance of internal mediators of external experience (Weiss, 1984), more recent developments have placed greater emphasis on cognitive and affective change strategies. Many of these strategies apply standard cognitive therapy procedures to relationship problems. For example, self-instructional strategies are used to modify negative attributions (Baucom & Lester, 1986) or control anger (Schindler & Vollmer, 1984), and guided discovery, Socratic dialogue and behavioral experiments are used to challenge irrational relationship beliefs (Baucom & Epstein, 1990; Halford, in press).

The greater emphasis on changing subjective experience in CBCT is particularly evident in recent changes in managing destructive conflict within distressed couples. Earlier versions of CBCT emphasized teaching communication and problem-solving skills to reduce conflict. More recently, emphasis has been placed on exploring with each partner the attributions, meaning, and significance attached to the issues which are sources of conflict (e.g., Christensen et al., 1995; Weiss & Halford, 1995). While there are variations in the details of the therapeutic process used by different authors, the common emphasis is on altering how partners respond to behaviors of their spouses which they dislike. Christensen et al. (1995) describe the goal of this process as promoting acceptance, which they define behaviorally as the reduction of attempts to get the other person to change. In essence, there is an attempt to balance a combination of behavior change to alter relationship interaction and change in the subjective experience of existing relationship interactions.

The CBCT approach to relationship dissatisfaction usefully can be summarized within an extension of the self-regulation framework proposed by Halford et al. (1994). Subjective dissatisfaction with your relationship can be responded to in one of five ways. First, you can alter the ways in which you attempt to persuade your partner to change, so that you get change. Second, you can alter your own behavior to enhance relationship functioning. These first two options were the focus of traditional behavioral couples therapy, producing behavior change through procedures such as behavioral contracting and negotiation. Third, you can alter your own subjective response to negative aspects of relationship interactions so that those aspects are less stressful and you feel no pressing need for change. This is what Christensen et al. (1995) call acceptance. Fourth, you can decide that particular existing behaviors are unacceptable, intolerable, and unchangeable, and that you will therefore leave the relationship. Finally, you can do nothing, and maintain the status quo. In our experience few clients have explored all these options for individual action.

6.28.5.4 The Effectiveness of Cognitive-behavioral Couples Therapy

CBCT consistently has been shown to be superior to no treatment or therapist contact control conditions in reducing marital distress (Hahlweg & Markman, 1988; Markman & Hahlweg, 1993). More specifically, CBCT improves couples' communication skills, reduces destructive conflict, enhances positivity of day-to-day interactions, and increases the positivity of couples' cognitions about their partners and their relationships (Hahlweg & Markman, 1988; Halford et al., 1993). In other words, CBCT is effective in changing the key identified risk factors for marital distress.

Among couples who present for couples therapy, the negative effects of relationship distress usually are already evident (Bloom, 1985). For example, alcohol abuse and depression are common in partners seeking marital therapy, as are behavior problems in the couples' children (Halford, Kelly, & Markman, 1997). These individual problems often result from the prolonged effects of marital distress, and make marital therapy less effective (Halford & Bouma, 1997). Furthermore, patterns of dysfunctional marital interaction typically become entrenched over time and resistant to change (Markman, Floyd, Stanley, & Storaasli, 1988; Raush, Barry, Hertel, & Swain, 1974). Eventually there is erosion of any positive affect toward the spouse, which is predictive of very poor response to couples therapy (Hahlweg, Schindler, Revenstorf, & Brengelmann, 1984). Approximately 25–30% of couples show no measurable improvement with CBCT, and as many as a further 30% improve somewhat from therapy but still remain significantly maritally distressed after treatment (Bray & Jouriles, 1995; Halford et al., 1993; Jacobson, 1989; Jacobson et al., 1984). Even among those couples who initially respond well to CBCT, there is substantial relapse toward marital distress over the next few years (Jacobson, Schmaling, & Holtzworth-Munroe, 1987; Snyder, Mangrum, & Wills, 1993; Snyder et al., 1991a). The consistent finding is that the longer couples have been maritally distressed, and the more severe their relationship dissatisfaction, the poorer their response to marital therapy (Whisman & Snyder, 1997).

The effectiveness of CBCT relative to other approaches to couples therapy is unclear. There are relatively few studies which compare approaches to couples therapy (Snyder et al., 1991a). One study compared insight-oriented and behavioral marital therapies. The findings were that each treatment was effective, with few significant differences in effects of the treatments in the short term (Snyder & Wills, 1989). At four-year follow-up the insight-oriented therapy was associated with lower rates of divorce and higher rated marital satisfaction than the behavioral marital therapy (Snyder et al., 1991a). However, this study has been the source of considerable controversy (Jacobson, 1991; Snyder et al., 1991b). A major criticism of the study was that the so-called insight-oriented therapy was not really psychodynamic in approach, but rather represented state of the art CBCT (Jacobson, 1991).

An important index of the value of couples therapy is the satisfaction of clients with the service they receive in that therapy. The magazine *Consumer Reports*, which is very widely read in the US, has its readers complete a survey every year. In one survey they asked about satisfaction with psychotherapy, including couples therapy. The results were very positive in terms of the effects of psychotherapy from the consumers' point of view. However, consumers expressed less satisfaction with couples' counseling than any other form of therapy. From one perspective, this makes the field of couples' counseling look ineffective. However, when one thinks about the goals of couples therapy, it actually makes sense that there will be less satisfaction from the consumers' point of view. Often one or sometimes both partners in couples therapy are using the therapy as a way of exiting the relationship—leaving the hurt spouse on the therapist's doorstep. Some marital therapists describe how a spouse who is interested in leaving the relationship often brings in the other spouse with the hope that the therapist will help wean the more committed spouse from the marriage and ease the break-up of the relationship.

In a survey of therapists conducted with practitioners in Colorado in the US (Stanley, Lobitz, & Markman, 1989), couples' therapists were asked to indicate the percentage of couples who were significantly improved or not improved, how many of the couples seen in the practice were still together, and how many decided to end their relationship. The results indicated that about 33% of the relationships ended in separation or divorce. Interestingly, the therapists rated 80% of these relationships that ended in divorce as a successful outcome. The *Consumer Reports* survey suggests that this 80% probably did not rate the therapy as a success from their own perspective (Seligman, 1995). Consumer satisfaction with couples therapy may be low because people unrealistically expect therapy to persuade their partners to stay with them.

6.28.5.5 Couples Therapy and Psychological Disorder

Couples therapy can be useful in the treatment of individual psychological disorders, either as the entire treatment or as an adjunct to individual treatment. More specifically, there have been a number of studies evaluating couples therapy and its effects on depression and alcohol abuse. In this section we evaluate this evidence.

As noted previously, people with co-existing marital problems and depression often do not respond well to traditional individual therapies for depression. Beach and O'Leary (1986) and Jacobson, Dabson, Fruzetti, Schmaling, and Salusky (1991) showed that CBCT was effective in reducing both marital distress and depression where wives were depressed and the couples were maritally distressed. Individual cognitive therapy also was somewhat effective with these clients, but not as effective as CBCT (Beach & O'Leary, 1986; Jacobson et al., 1991). Of particular interest was Jacobson's finding that, while both CBCT and cognitive therapy were helpful, adding cognitive therapy to CBCT did not enhance outcome. In Jacobson's study the treatment conditions were matched for total hours of therapy contact. As a consequence the combined condition got less CBCT than the

CBCT alone condition, and this may explain why the combined treatment condition did not do as well.

The Jacobson et al. (1991) study highlights that couples therapy should not be used for all presentation for depression by married clients. For depressed clients where there was no marital distress, individual cognitive therapy was superior to couples therapy in its effect on depression. Even when relationship problems and depression coexist there is a need to make careful judgments as to when couples therapy is most helpful. In couples where the individual depression is severe, the marital problems postdated the onset of depression, and neither partner attributes the cause of the depression to the marital problems, then marital therapy seems less effective than individual therapy (Jacobson et al., 1991). On the other hand, where the marital problems antedate the depression, the marital problems are severe, and the source of the depression is attributed by at least one of the partners to the marital problems, then marital therapy seems to be the treatment of choice (Jacobson et al., 1991).

There is a large body of literature attesting to the value of various forms of couples therapy in the treatment of alcohol problems. The best established approach is the use of CBCT to complement individual therapy in the treatment of alcohol abuse. In the case of people with heavy dependence on alcohol, there has been successful use of conjoint contracting to promote use of antebuse (an oral drug which induces severe illness if alcohol is consumed) to establish abstinence (O'Farrell & Bayog, 1986). Once sobriety is established, conjoint couples therapy when used in combination with ongoing use of antabuse promotes both improvements in relationship satisfaction and enhanced maintenance of drinking control (O'Farrell, Cutter, Choquette, Floyd, & Bayog, 1992; O'Farrell, Cutter, & Floyd, 1985). When CBCT is combined with relapse prevention training, there is even better maintenance of reduced alcohol consumption (O'Farrell, Choquette, Cutter, Brown, & McCourt, 1993). Furthermore, there is also a substantial reduction in the prevalence of marital violence after CBCT, at least in those couples with mild to moderate severity of violence (O'Farrell & Murphy, 1995).

There are two significant limitations on the evidence of the impact of couples treatments for problem drinking. First, all successful trials have used antabuse, and the effectiveness of CBCT without antabuse is unclear. Second, up to 50% of men in alcohol treatment refuse offers of couples-based treatment (O'Farrell, Kleinke, Thompson, & Cutter, 1986). This may be associated with more general reluctance of many people who abuse alcohol to accept treatment. Partners of male problem drinkers often present to treatment agencies reporting that their partners refuse to seek treatment (Halford & Osgarby, 1993). The frequent unwillingness of the person abusing alcohol to accept help, with consequent strains on the relationship and individual distress, have prompted the development of therapies to assist the spouses of heavy drinkers. For example, Sisson and Azrin (1986) and Thomas and Ager (1993) both developed therapies aiming to help clients reduce the negative impacts of their partner's drinking. Both these programs were aimed at teaching women to manage stress associated with their husbands' drinking and to encourage the men to seek individual therapy (Sisson & Azrin, 1986; Thomas & Ager, 1993). These programs do assist the women to reduce their individual distress, and there is some evidence that these approaches can increase the chance of the male drinker presenting for individual treatment (Halford, Price, Bouma, Kelly, & Young, 1998; Sisson & Azrin, 1986; Thomas & Ager, 1993).

Some programs for the wives of male problem drinkers were also aimed at helping the women to influence the men to reduce drinking. For example, Sisson and Azrin (1986) described helping women to identify high-risk settings for problem drinking, and suggested that wives schedule activities incompatible with drinking for high-risk times. Partners have also been encouraged to praise sobriety, and not to inadvertently reduce the negative consequences of drinking (e.g., by refusing to ring your partner's supervisor to say your partner is sick when he really has a hangover). Similar strategies were used by Halford et al. (1998). The only systematic evaluation of this approach was in the Halford study, which found that these strategies had little impact upon the drinker's alcohol consumption. The degree to which women actually implemented the suggested strategies was not evaluated, so it is unclear if women were unconvinced to implement the suggestions or if women changing these behaviors had little effect on the men's drinking.

In summary, CBCT is a well-established treatment for relationship problems which has demonstrated efficacy. CBCT is also very useful in the treatment of depression and alcohol problems of adults in committed relationships. However, common to all applications of CBCT and other approaches to couples therapy is a substantial input of time from therapists and couples. As described in the research and clinical literature, couples therapy typically involves anything from 12 to 30 sessions of conjoint therapy (e.g., Baucom & Epstein, 1990;

Greenberg & Johnson, 1988; Snyder et al., 1991a; Weiss & Halford, 1995). The time commitment and costs involved in all these approaches probably are a major disincentive to couples to seek assistance; only a small minority of distressed couples ever seek out therapy (Sanders, 1995; Wolcott & Glazer, 1989). Thus, even if CBCT were optimally effective for those who present for therapy, it would have modest impact on the community prevalence of relationship problems. This has led to attempts to prevent relationship problems.

6.28.6 PREVENTION OF RELATIONSHIP PROBLEMS

6.28.6.1 Prevention and Early Intervention

Programs to prevent relationship problems vary greatly along several dimensions, including the phase of the relationship at which they are targeted, the settings in which they are offered, the number and duration of sessions, the level of training of the service providers, the extent to which didactic vs. experiential forms of learning are emphasized, and the amount of group vs. dyadic interaction that occurs (Guerney, Guerney, & Cooney, 1985; Levant, 1986; Van Widenfelt, Markman, Guerney, Behrens, & Hosman, 1997). Existing programs can be usefully classified into two broad categories: (i) counseling and educational programs (e.g., Mace & Mace, 1976; Rutledge, 1968), and (ii) skill acquisition and cognitive restructuring programs (e.g., Guerney, 1977; Markman, 1984; Markman, Floyd, Stanley, & Lewis, 1986). Programs are based on a wide range of theoretical perspectives including social learning theory, Rogerian theory, and systems theory (Bradbury & Fincham, 1990). Regardless of theoretical orientation, all programs include efforts aimed at improving communication and problem-solving, negotiation of roles and responsibilities, clarification of values and expectations, sexuality and intimacy, developmental changes and transitions (e.g., parenting), and awareness of relationship dynamics.

The rationale for skills-based prevention of relationship problems is that acquisition of crucial relationship competencies are presumed to prevent problems. As noted earlier, research has identified these crucial competencies to include constructive communication and problem-solving skills, management of negative affect and conflict, realistic relationship beliefs and expectations, and promotion of positive behavior exchange (Markman, 1984; Notarius & Vanzetti, 1984). Programs designed to improve couples' abilities in these areas have been conducted both with couples planning marriage (e.g., Markman, 1981; Halford et al., 1994) and with married couples not experiencing relationship distress (e.g., Guerney, 1977; Miller et al., 1975).

Skills training prevention programs do increase couples' communication and problem-solving skills (Hahlweg & Markman, 1988; Markman et al., 1988; Miller, Nunnally, & Wackman, 1975). In their meta-analysis of controlled trials of such interventions, Hahlweg and Markman (1988) found an effect size of 1.51 on communication skills. Given that an effect size of 0.8 is generally considered a large effect size, this is a very impressive result. Couples in prevention trials are, by definition, currently satisfied with their relationships. Consequently, it is unlikely that prevention programs could produce large, immediate increases in relationship satisfaction above the pre-existing high levels. Consistent with this interpretation, Hahlweg and Markman (1988) report a mean effect size across seven premarital programs of 0.51 on relationship satisfaction, which is still a moderate size effect.

Over time many couples show declines in relationship satisfaction (Markman & Hahlweg, 1993), and the real test of prevention programs' effectiveness is whether the onset of relationship problems can be prevented. A significant methodological problem associated with evaluating prevention programs is devising appropriate comparison conditions for long-term controlled trials. If couples are randomly assigned to a no-intervention control when they desire intervention, they may seek intervention outside the study. On the other hand, failure to randomly assign to conditions can make interpretation of results difficult. Perhaps because of these difficulties few studies have assessed long-term outcome. A survey of 85 couple and family prevention and enrichment studies found that only 40% included follow-up measures at some point after post-test (mean follow-up was 12 weeks), with the longest follow-up assessment occurring at 12 months (Giblin, Sprenkle, & Sheehan, 1985). Only Markman and co-workers (e.g., Markman et al., 1988; Markman, Renick, Floyd, Stanley, & Clements, 1993) have reported follow-up data beyond two years.

Markman and co-workers found that skills-based marital distress prevention programs reduce marital violence (Markman et al., 1993), and the prevalence of marital separation and divorce over the first 5–10 years of marriage (Markman et al., 1993). However, these series of studies have a methodological problem. Subjects initially were recruited for a long-term follow-up study of marriage, and subsequently randomly assigned subjects to matched groups to be offered, or not offered, an intervention. Sixty percent of couples offered the relationship preparation program declined to participate. The two treatment conditions consist of those who accepted the program and their matched controls. As the treatment group were self-selected for agreeing to participate in the program, and the matched controls were not offered the program, there is confound between the conditions of self-selection.

Despite the limitations of the existing research, it is clear that skills-based prevention programs do modify aspects of couples' adaptive processes which are identified risk factors for relationship problems. The long-term maintenance of these behavioral changes needs further investigation, as do the long-term effects of these interventions on relationship problems, but preliminary results are encouraging.

6.28.6.2 Content of Prevention Programs

The content covered in skills-based prevention programs can be usefully considered as falling into six modules: (i) behavior exchange and positive intimacy enhancement, (ii) communication skills enhancement, (iii) management of conflict and negative affect, (iv) sexual enrichment, (v) gender role flexibility, and (vi) adaption to life transitions and maintenance of relationship satisfaction. We will consider each of these in turn.

Behavior exchange and positive intimacy enhancement procedures have been adapted from those used in couples therapy. As for therapy, the emphasis is on each partner self-selecting behavior change goals which enhance relationship functioning. In prevention programs this is often much easier to achieve than in therapy. In therapy, couple's negative feelings toward each other and possible ambivalence about the relationship often inhibit partners from making relationship enhancing efforts. Prevention program couples, by definition, have positive feelings about their partner and relationship.

Communication skills training involves individuals self-selecting goals for enhancing their own communication from an array of available skills and self-evaluating their own communication. This can be done by asking the couple to have a discussion. Immediately after the discussion both partners are asked to self-assess their communication using a checklist of potentially helpful communication behaviors set out on the form in Table 4. After completing the form, partners self-identify specific communication behaviors that each would like to increase in their repertoire in order to enhance their communication. This does not mean that the leader avoids responsibility for helping the partners to determine goals. Rather, the role of the leader is to help both partners to accurately self-evaluate their current communication and to develop specific, self-selected goals for enhancing communication.

Another exercise often used in communication skills training is to ask the couple to have a discussion at home which they audiotape. At the end of the discussion both partners self-evaluate their communication skills, formulate self-change goals, and then repeat the conversation building on the selected communication skills. Each partner then evaluates the impact of changing communication skills on the communication process. This procedure allows the partners to directly test the hypothesis that changing their communication behaviors in a particular way will enhance the relationship communication process. It also has the partners practice active self-correction of ineffective communication attempts.

Communication tasks that we recommend in relationship enhancement initially focus on nonconflictual topics. For example, we emphasize to couples the importance of discussing day-to-day occurrences on a regular basis, so as to keep informed and interested in the experiences of your partner. Consequently, we usually ask couples to have a series of discussions about what each does during the time they are apart. We also ask the couple to address ways in which they can support and show interest in each other's interests. This support might be in work, parenting, household chores, or in particular interests or hobbies. The relevant communication task is for one partner to initiate a conversation with their spouse about the spouse's interests and needs. Both partners are asked to identify ways in which they might be more supportive.

Conflict management is another important element of prevention programs. Once couples have a reasonable level of communication skills, they can then use these communication skills to handle difficult conflictual issues. In discussion of conflict management we describe the different settings in which conflictual topics may be discussed, and the fact that there are some settings in which it is easier to have a productive discussion than others. Couples are asked to consider the times, places, and circumstances in which they could most productively talk about topics which are difficult. Also couples are educated about the common maladaptive patterns in interaction around conflictual topics, such as the demand-withdraw and mutual avoidance patterns. Couples self-select goals which will help them to avoid those unhelpful patterns of interaction.

Table 4 Communication skills self-evaluation form.

Name: ______________________ Date: __________

The aim of this form is for you to identify your strengths and weaknesses in communication and to specify goals for improvement. Rate each of the skills below using this code:

0 — Very poor use of skill
1 — Unsatisfactory use of skill
2 — Satisfactory use of skill, but room for improvement
3 — Good use of skill
N/A — Not applicable

Skill	*0*	*1*	*2*	*3*	*N/A*
Specific descriptors					
Self-disclosure					
Clear expression of positives					
Assertive expression of negatives					
Attending to partner					
Minimal encouragers					
Reserving judgement					
Asking questions					
Summarizing content					
Paraphrasing feelings					

Self-identified strengths in communication: ______________________

____

Self-identified weaknesses in communication: ______________________

____

Another useful notion in relationship preparation programs is the idea of relationship rules for managing conflict. Relationship rules refer to the implicit assumptions and processes which couples develop for handling conflict. For exmple, some couples may believe that either partner should be able to raise any issue that is troubling them at any time. Other couples may feel that there is some responsibility to raise difficult issues at times and in places which maximize the chance of productive discussion. Table 5 lists some common relationship rules that couples may apply to managing conflict. Through discussion and negotiation couples can be asked to agree on relationship rules that increase their chance of managing conflict effectively.

Sexual enrichment is also an important element of many prevention programs. We make use of the widely available video tape *The lovers guide*. We would play this to couples, ask them to talk about their general reactions as a group, and then have the couples talk together in privacy about any goals they may wish to set in order to enhance their sexual experience. The video that we show is very sexually explicit, demonstrating ways in which one can enhance sensual and sexual pleasure, and identifies and describes some common sexual difficulties. We find that showing this tape is useful in helping couples to talk intimately about their desires to enhance their sexual relationship.

Gender role flexibility is another element of some prevention programs. Gender role flexibility refers to each partner taking some responsibility for a wide range of relationship maintaining activities, and transcending very rigid or traditional gender-based definitions of responsibilities. The notion of gender role flexibility can be introduced by asking both partners to identify which of a range of different behaviors, which potentially can be stereotyped as belonging to a woman or a man, they take responsibility for within the relationship. The potential value of being flexible in gender roles is in helping the couple adapt to changing life circumstances. For example, if both partners are able to earn an income, then the couple can adapt better if one person has difficulty obtaining work. Partners can be helped to self-identify specific behavior changes they wish to make in order to become more gender role flexible. For example, males may set themselves the target of learning to cook better, and women to learn how to maintain a car. An additional rationale for this process is that often the domestic responsibilities in families fall primarily on women, and this is a major source of relationship dissatisfaction for women. We point out to our male participants that women initiate approximately three-quarter of all divorce proceedings (Wolcott & Glezer, 1989). If men want to remain in happy relationships, then they need to ensure that they are domestically competent to contribute to the household.

In relationship preparation and enhancement programs it is important that couples prepare themselves for future changes in their relationship likely to result from major life transitions (Hahlweg & Markman, 1993). In one such exercise couples complete an adapted form of the life change event scale and are asked to rate the likelihood of these events occurring in the next one or two years and the probability that if those events occur they might have a negative effect on their relationship. For example, couples are asked to identify the ways in which events such as birth of a child, loss of a job, or a change of work circumstances might impact on their relationship. Partners are asked to self-select goals that they believe would help them adapt in a relationship-maintaining way to these life transitions. For example, a number of couples have identified that they may have reduced opportunities for having time together as a couple once they have children. They have

Table 5 Suggested ground rules for handling conflict.

1. We can bring up issues at any time, but the "listener" can say "this is not a good time" If listener does not want to talk at that time, he or she takes responsibility for setting up a time to talk in the near future (you need to decide on how "the near future" is defined)	_____ Yes	_____ No
2. When conflict is escalating we will call a "stop action" and (i) try it again, using the speaker/listener technique, (ii) agree to talk later at a specified time about the issue, using the speaker/listener technique	_____ Yes	_____ No
3. When we are having trouble communicating we will "engager" the speaker/listener technique	_____ Yes	_____ No
4. When using the speaker/listener technique, we will completely separate problem-*discussion* from problem-*solution*	_____ Yes	_____ No
5. We will have weekly "couple meetings" (you should set up a time now)	_____ Yes	_____ No
6. We will state when we have a problem, if we want to negotiate a relationship, or just solve the immediate problem	_____ Yes	_____ No

Source: Markman, Stanley, and Blumberg (1995).

then set individual goals, such as ensuring child minding is available or that they cultivate activities they can do at home, which increase the chance of having mutual shared enjoyable activities when they have young children.

In Table 6 we have set out the typical content of each of five 2 hour group sessions of a relationship preparation program. This is not a rigid schedule that must be followed, but rather a schematic representation about how a group might work. Sessions can be combined and homework resequenced to cover the same content in a full day, with one or two follow-up sessions.

6.28.7 CONCLUSIONS

Relationship difficulties are very common, and have a huge impact on adults and their offspring. The most common relationship problems are lack of relationship satisfaction, ambivalence about the future of the relationship, and problems in the adaptive interactional processes between partners. In the domain of the adaptive processes between the partners the most important factors contributing to relationship problems are poor communication, poor management of negative affect and conflict, unrealistic beliefs and expectations, and lack of positive shared experiences. Other important contributors to relationship problems are enduring vulnerabilities of the partners, such as individual psychological disorders and stressful life events. Couples lacking robust adaptive couple processes are likely to develop relationship problems when challenged by major life transitions and stressful events.

Cognitive-behavioral couples therapy is a well-established and effective means of helping many couples with relationship problems. Effective couples therapy uses the assessment process to help partners build a shared and constructive model of their relationship problems, and to define individual action that can improve the relationship. Therapy itself assists partners to self-direct a change process focused on enhancing positive exchange between partners, improving communication, and controlling destructive conflict. Adaptions of these procedures have been used successfully to teach currently satisfied couples to enhance their relationship and to prevent future relationship problems.

ACKNOWLEDGMENTS

Preparation of this chapter was supported by an Australian Research Council Grant on

Table 6 Overview of typical content in a five session premarital skills training program.

Session	*Detail of content*
1	Introduction to group members; explain rationale for self-directed, skills training focused group program; identification of key behavioral domains promoting relationship intimacy; self-directed intimacy enhancement through self-directed goal setting and definition of homework task of behavior change; identification and modeling of key communication skills to enhance intimacy
2	Review of intimacy enhancement behavioral homework tasks; self-directed selection of further behavior change goals; review of key communication skills; guided self-evaluation of current communication skills; self-directed selection of communication enhancement goals and practice of implementation of those skills; self-directed goal setting and definition of homework task to enhance communication
3	Review of communication homework tasks, and self-directed further goals selection and definition of further homework task; introduction to the concept of the patterns of conflict and effective conflict management; negotiation with partner about relationship rules for managing conflict; self-directed goal setting for effective management of conflict; introduction to the concept of flexible gender roles, couple review of current gender roles, self-directed goal setting for future gender role flexibility
4	Review of communication homework task; review of the role of sexuality in relationship intimacy; couple discussion and goal setting to enhance sexual intimacy; introduction to the concept of partner support, self-directed goal setting to enhance partner support; self-directed definition of homework tasks to implement selected goals in areas of sexuality or partner support
5	Review of homework tasks; self-directed selection of any further goals to enhance relationship functioning; introduction of issue of maintenance of relationship functioning; self-directed identification of future life events impacting upon relationship; planning to promote relationship adaptation to predictable life events Closure

"Prediction and prevention of marital distress" to the first author. Thank you to Rhoda Richardson, Kathy Eadie, and Amanda Thomas for help with manuscript preparation.

6.28.8 REFERENCES

Azrin, N. H., Naster, B. J., & Jones, R. (1973). A rapid learning based procedure for marital counselling. *Behavior Research and Therapy, 11,* 365–382.

Baucom, D. H., & Epstein, N. (1990). *Cognitive behavioral marital therapy.* New York: Brunner/Mazel.

Baucom, D. H., Epstein, N., Daituo, A. D., Carels, R. A., Rankin, L. A., & Burnett, C. K. (1996). Cognitions in marriage: The relationship between standards and attributions. *Journal of Family Psychology, 10,* 209–222.

Baucom, D. H., Epstein, N., Sayers, S., & Sher, T. G. (1989). The role of cognitions in marital relationships: Definitional, methodological and conceptual issues. *Journal of Consulting and Clinical Psychology, 57,* 31–38.

Baucom, D. H., & Lester, G. W. (1986). The usefulness of cognitive restructuring as an adjunct to behavioral marital therapy. *Behavior Therapy, 17,* 385–403.

Beach, S. R., Arias, I., & O'Leary, K. D. (1986). The relationship of marital satisfaction and social support to depressive symptomatology. *Journal of Psychopathology and Behavioral Assessment, 8,* 305–316.

Beach, S. R., & O'Leary, K. D. (1986). The treatment of depression occurring in the context of marital discord. *Behavior Therapy, 17,* 43–49.

Beach, S. R. H., Sandeen, E. E., & O'Leary, K. D. (1990). *Depression in marriage: A model for etiology and treatment.* New York: Guilford Press.

Beach, S. R. H., Sandeen, E. E., & O'Leary, K. D. (1992). *Depression in marriage.* New York: Guilford Press.

Bebbington, P. (1987). Marital status and depression: A study of English national admission statistics. *Acta Psychiatrica Scandanavica, 75,* 640–650.

Billings, A. G., & Moos, R. H. (1983). Psychosocial processes of recovery among alcoholics and their families: Implications for clinicians and program evaluators. *Addictive Bahavior, 8,* 79–88.

Birchler, G. R., Weiss, R. L., & Vincent, J. P. (1975). Multimethod analysis of social reinforcement exchange between maritally distressed and nondistressed spouse and stranger dyads. *Journal of Personality and Social Psychology, 31,* 349–360.

Birchnall, J., & Kennard, J. (1983). Does marital maladjustment lead to mental illness? *Social Psychiatry, 18,* 79–88.

Black, L. E., & Sprenkle, D. H. (1991). Gender differences in college student's attitudes toward divorce and their willingness to marry. *Journal of Divorce and Remarriage, 15,* 47–60.

Block, J., Block, J. H., & Keyes, S. (1988). Longitudinally foretelling drug usage in adolescence. Early childhood personality and environmental precursors. *Child Development, 59,* 336–355.

Bloom, B. (1985). *Community mental health: A general introduction* (2nd ed.). Monterey, CA: Brooks-Cole.

Bloom, B., Asher, S. J., & White, S. W. (1978). Marital disruption as a stressor: A review and analysis. *Psychological Bulletin, 85,* 867–894.

Booth, A., & Edwards, J. N. (1992). Starting over: why remarriages are more unstable. *Journal of Family Issues, 13,* 179–194.

Bornstein, P. H., & Bornstein, M. T. (1986). *Marital therapy: a behavioral communications approach.* New York: Pergamon.

Boughner, S. R., Hayes, S. F., Bubenezer, D. L., & West, J. D. (1994). Use of standardised assessment instruments by marital and family therapists: A survey. *Journal of Marital and Family Therapy, 20,* 69–75.

Bradbury, T. N. (1995). Assessing the four fundamental domains of marriage. *Family Relations, 44,* 459–468.

Bradbury, T. N., & Fincham, F. D. (1990). Preventing marital dysfunction: review and analysis. In: F. D. Fincham & T. N. Bradbury (Eds.), *The psychology of marriage: basic issues and applications* (pp. 375–401). New York: Guilford Press.

Bradbury, T. N., & Fincham, F. D. (1992). Attribution and behavior in marital interaction. *Journal of Personality and Social Psychology, 63,* 613–628.

Bray, J. H., & Jouriles, E. N. (1995). Treatment of marital conflict and prevention of divorce. *Journal of Marital and Family Therapy, 21,* 461–473.

Burman, B., & Margolin, G. (1992). Analysis of the association between marital relationships and health problems: An interactional perspective. *Psychological Bulletin, 112,* 39–63.

Cascardi, M., Langhinrichsen, J., & Vivian, D. (1992). Marital aggression, impact, injury and health correlates for husbands and wives. *Archives of Internal Medicine, 152,* 1178–1184.

Christensen, A., Jacobson, N. S., & Babcock, J. (1995). Integrative behavioral couple therapy. In N. S. Jacobson & A. S. Gurman (Eds.), *Clinical handbook of couple therapy.* New York: Guilford Press.

Christensen, A., & Nies, D. C. (1980). The spouse observation checklist: Empirical analysis and critique. *American Journal of Family Therapy, 8,* 69–79.

Christensenn, A., & Shenk, J. L. (1991). Communication, conflict and psychological distance in nondistressed, clinic, and divorcing couples. *Journal of Consulting and Clinical Psychology, 59,* 458–463.

Cowen, C. P., & Cowen, P. A. (1992). *When partners become parents.* New York: Basic Books.

Coyne, J. C., Kahn, J., & Gotlib, I. H. (1987). Depression. In T. Jacob (Ed.), *Family interaction and psychopathology* (pp. 509–534). New York: Plenum.

Craske, M. G., & Zoellner, L. A. (1995). Anxiety disorders: the role of marital therapy. In N. S. Jacobson & A. S. Gurman (Eds.) *Clinical handbook of couple therapy.* (pp. 369–393). New York: Guilford Press.

Davis, D. I., Berenson, D., Steinglass, P., & Davis, S. (1974). The adaptive consequences of drinking. *Psychiatry, 37,* 209–215.

DeGuilbert-Lantione, C., & Monnier, A. (1992). La conjuncture demognishique: L'Europe et les pays developpes d'Outre-Mer. *Population, July–August.*

Dickson, F. (1997). Ageing and marriage. In W. K. Halford & H. J. Markman (Eds.), *Clinical handbook of marriage and couples interventions* (pp. 255–272). Chichester, UK: Wiley.

Dobson, K. S. (1987). Marital and social adjustment in depressed and married women. *Journal of Clinical Psychology, 43,* 261–265.

Eddy, J. M., Heyman, R. E., & Weiss, R. L. (1991). An empirical evaluation of the Dyadic Adjustment Scale: Exploring the differences between marital "satisfaction" and "adjustment." *Behavioral Assessment, 13,* 199–220.

Eidelson, R. J., & Epstein, N. (1982). Cognition and relationship maladjustment: Development of a measure of dysfunctional relationship beliefs. *Journal of Consulting and Clinical Psychology, 50,* 715–720.

Emery, R. E. (1982). Interparental conflict and the children of discord and divorce. *Psychological Bulletin, 92,* 310–330.

Emery, R. E., Joyce, S. A., & Fincham, F. D. (1987). The assessment of child and marital problems. In K. D. O'Leary (Ed.), *Assessment of marital discord.* Hillsdale, NJ: Erlbaum.

Emery, R. E., & Wyer, M. M. (1987). Divorce mediation.

American Psychologist, 42, 472–480.

Emmelkamp, P. M. G., De Haan, E., & Hoogduin, C. A. I. (1990). Marital adjustment and obsessive-compulsive disorder. *British Journal of Psychiatry, 156,* 55–60.

Ewart, C. K., Taylor, C. B., Kraemer, H. C., & Agras, S. (1991). High blood pressure and marital discord; not being nasty matters more than being nice. *Health Psychology, 10,* 155–163.

Fincham, F. D., Beach, S. R. H., & Kemp-Fincham, S. I. (1997). Marital quality: A new theoretical perspective. In R. J. Sternberg & M. Hojjat (Eds.), *Satisfaction in close relationships*. New York: Guilford Press.

Fincham, F. D., & Bradbury, T. D. (1990). *The psychology of marriage*. New York: Guilford Press.

Fincham, F. D., & Bradbury, T. N. (1992). Assessing attributions in marriage: The Relationship Attribution Measure. *Journal of Personality and Social Psychology, 62,* 457–468.

Fincham, F. D., Garnier, P. C., Gano-Phillips, S., & Osborne, L. N. (1995). Preinteraction expectations, marital satisfaction and accessibility: a new look at sentiment override. *Journal of Family Psychology, 9,* 3–14.

Floyd, F. J., & Markman, H. J. (1983). Observational biases in spouse observation: Toward a cognitive-behavioral model of marriage. *Journal of Consulting and Clinical Psychology, 51,* 450–457.

Fowers, B. J., Applegate, B., Olson, D. H., & Pomerantz, B. (1994). Marital conventionalization as a measure of marital satisfaction: a confirmatory factor analysis. *Journal of Family Psychology, 8,* 98–103.

Fowers, B. J., Lyons, E R. M., & Montel, K. H. (1996). Positive marital illusions: self-enhancement or relationship enhancement? *Journal of Family Psychology, 10,* 192–208.

Fraenkel, P. (1997). Systematic and strategic couples therapy. In W. K. Halford & H. J. Markman (Eds.), *Clinical handbook of marriage and couples interventions* (pp. 379–414). Chichester, UK: Wiley.

Gallup (1989). *Marriage satisfaction*. Los Angeles: Los Angeles Times Syndicate.

Gibardi, L., & Rosen, L. A. (1991). Differences between college students form divorced and intact families. *Journal of Divorce and Remarriage, 15,* 175–191.

Giblin, P., Sprenkle, D. H., & Sheehan, R. (1985). Enrichment outcome research: A meta-analysis of premarital, marital and family interventions. *Journal of Marital and Family Therapy, 11,* 257–271.

Glenn, N. D., & Kramer, K. B. (1987). The marriages and divorces of the children of divorce. *Journal of Marriage and the Family, 49,* 811–825.

Glenn, N. D., & Weaver, C. N. (1981). The contribution of marital happiness to global happiness. *Journal of Marriage and the Family, 43,* 161–168.

Glick, P. C. (1989). Remarried families, stepfamilies and stepchildren: Brief demographic profile. *Family Relations, 38,* 24–27.

Gore, S. (1978). The effect of social support in moderating the health consequences of unemployment. *Journal of Health and Social Behavior, 19,* 157–165.

Gottman, J. M. (1990). How marriages change. In G. R. Patterson (Ed.), *Depression and aggression in family interaction* (pp. 75–102). Hillsdale, NJ: Erlbaum.

Gottman, J. M. (1993). A theory of marital dissolution and stability. *Journal of Family Psychology, 7,* 57–75.

Gottman, J. M. (1994). *What predicts divorce? The relationship between marital processes and marital outcomes*. Hillsdale, NJ: Erlbaum.

Gottman, J. M., & Krokoff, L. J. (1989). Marital interaction and satisfaction: A longitudinal view. *Journal of Consulting and Clinical Psychology, 57,* 47–52.

Gottman, J. M., & Levenson, R. W. (1988). The social psychophysiology of marriage. In P. Noller & M. A. Kirkpatrick (Eds.), *Perspectives on marital interaction* (pp. 82–120). Philadelphia: Multilingual Matters.

Gottman, J. M., Notarius, C. I., Markman, H. J., & Gonso, J. (1976). *A couple's guide to communication*. Champaign, IL: Research Press.

Gottman, J. M., Markman, H. J., & Notarius, C. I. (1977). The topography of marital conflict: a sequential analysis of verbal and nonverbal behavior. *Journal of Marriage and the Family, 39,* 461–477.

Gove, W. R., Hughes, M., & Styles, C. B. (1983). Does marriage have positive effects on the psychological well-being of the individual? *Journal of Health and Social Behavior, 24,* 122–132.

Greenberg, L. S., & Johnson, S. M. (1988). *Emotionally focussed therapy for couples*. New York: Guilford Press.

Grych, J. H., & Fincham, F. D. (1990). Marital conflict and children's adjustment: A cognitive-contextual framework. *Psychological Bulletin, 108,* 267–290.

Guerney, B. G. (1977). *Relationship enhancement*. San Francisco: Jossey-Bass.

Guerney, B. G., Guerney, L., & Cooney, T. (1985). Marital and family problems prevention and enrichment programs. In L. L'Abate (Ed.), *The handbook of family psychology and therapy* (Vol. 2, pp. 1179–1217). Homewood, IL: Dorsey.

Hahlweg, K., & Markman, H. J. (1988). Effectiveness of behavioral marital therapy: Empirical status of behavioral techniques in preventing and alleviating marital distress. *Journal of Consulting and Clinical Psychology, 56,* 440–447.

Hahlweg, K., Schindler, L., Revenstorf, D., & Brengelmann, J. C. (1984). The Munich marital therapy study. In K. Hahlweg & N. S. Jacobson (Eds.), *Marital interaction: analysis and modification* (pp. 3–19). New York: Guilford Press.

Halford, W. K. (1995). Behavior therapy and schizophrenia in context: Challenges and opportunities provided within the changing mental health system. *Behavior Change, 12,* 41–50.

Halford, W. K. (in press). *Couples therapy: Helping partners manage change*. New York: Guilford Press.

Halford, W. K., & Behrens, B. C. (1996). Prevention of marital difficulties. In P. Cotton & H. J. Jackson (Eds.), *Early intervention and preventative mental health application of clinical psychology* (pp. 21–58). Brisbane, Australia: Australian Academic Press.

Halford, W. K., & Bouma, R. (1997). Marital distress and individual psychopathology. In W. K. Halford & H. J. Markman (Eds.), *Clinical handbook of marriage and couples interventions* (pp. 291–322). Chichester, UK: Wiley.

Halford, W. K., Gravestock, F., Lowe, R., & Scheldt, S. (1992). Toward a behavioral ecology of stressful marital interactions. *Behavioral Assessment, 13,* 135–148.

Halford, W. K., Hahlweg, K., & Dunne, M. (1990). The cross-cultural consistency of marital communication associated with marital distress. *Journal of Marriage and the Family, 52,* 109–122.

Halford, W. K., Kelly, A., Bouma, R., & Young, R.McD. (1996). Individual psychopathology and marital distress: Analysing the association and implications for therapy. *Behavior Modification.*

Halford, W. K., Kelly, A., & Markman, H. J. (1997). The concept of a healthy marriage. In W. K. Halford & H. J. Markman (Eds.), *Clinical handbook of marriage and couples interventions*. Chichester, UK: Wiley.

Halford, W. K., & Markman, H. J. (Eds.) (1997). *Clinical handbook of marriage and couples interventions*. Chichester, UK: Wiley.

Halford, W. K., & Osgarby, S. M. (1993). Alcohol abuse in clients presenting with marital problems. *Journal of Family Psychology, 6,* 1–11.

Halford, W. K., & Osgarby, S. M. (1996). *Memories,*

memories: negative bias in recall of relationship events in maritally distressed couples. Paper presented to the 30th Annual Convention of the Association for the Advancement of Behavior Therapy, New York, November.

Halford, W. K., Osgarby, S. M., & Kelly, A. (1996). Brief behavioral couples therapy: A preliminary investigation. *Cognitive and Behavioral Psychotherapy*.

Halford, W. K., Price, J., Bouma, R. O., Kelly, A., & Young, R. (1998). The wives of untreated men misusing alcohol: Their marriages and coping. Unpublished manuscript.

Halford, W. K., & Sanders, M. R. (1988). Assessment of cognitive self statements during marital problem solving: A comparison of two methods. *Cognitive Therapy and Research, 12,* 515–530.

Halford, W. K., & Sanders, M. R. (1990). The relationship of cognition and behavior during marital interaction. *Journal of Social and Clinical Psychology, 9,* 489–510.

Halford, W. K., Sanders, M. R., & Behrens, B. C. (1993). A comparison of the generalisation of behavioral marital therapy and enhanced behavioral marital therapy. *Journal of Consulting and Clinical Psychology, 61,* 51–60.

Halford, W. K., Sanders, M. R., & Behrens, B. C. (1994). Self-regulation in behavioral couples therapy. *Behavior Therapy, 25,* 431–452.

Haynes, S. N., Jensen, B. J., Wise, E., & Sherman, D. (1981). The marital intake interview: A multi-method criterion validity assessment. *Journal of Consulting and Clinical Psychology, 49,* 379–387.

Heavey, C. L., Christensen, A., & Malmuth, N. M. (1995). The longitudinal impact of demand and withdrawal during marital conflict. *Journal of Consulting and Clinical Psychology, 63,* 797–801.

Heavey, C. L., Layne, C., & Christensen, A. (1993). Gender and conflict structure in marital interaction: a replication and extension. *Journal of Consulting and Clinical Psychology, 61,* 16–27.

Heyman, R. E., Sayers, S. L., & Bellack, A. S. (1994). Global marital satisfaction versus marital adjustment: An empirical comparison of three measures. *Journal of Family Psychology, 8,* 432–446.

Hooley, J. M., Orleay, J., & Teasdale, J. D. (1986). Levels of expressed emotion and relapse in depressed patients. *British Journal of Psychiatry, 148,* 642–647.

Howes, P., & Markman, H. J. (1991). Marital quality and child functioning: a longitudinal study. *Child Development, 60,* 1044–1051.

Jacobson, N. S. (1989). The politics of intimacy. *Behavior Therapist, 12,* 29–32.

Jacobson, N. S. (1991). Toward enhancing the efficacy of marital therapy and marital therapy research. *Journal of Family Psychology, 4,* 373–393.

Jacobson, N. S., Dobson, K., Fruzetti, A. E., Schmaling, K. B., & Salusky, S. (1991). Marital therapy as treatment of depression. *Journal of Consulting and Clinical Psychology, 59,* 547–557.

Jacobson, N. S., Follette, W. C., & McDonald, D. W. (1982). Reactivity to positive and negative behavior in distressed and nondistressed married couples. *Journal of Consulting and Clinical Psychology, 50,* 706–714.

Jacobson, N. S., Follette, W. C., Revenstorf, D., Baucom, D. H., Hahlweg, K., & Margolin, G. (1984). Variability in outcome and clinical significance of behavioral marital therapy: A reanalysis of outcome data. *Journal of Consulting and Clinical Psychology, 52,* 497–504.

Jacobson, N. S., & Margolin, G. (1979). *Marital therapy: Strategies based on social learning and behavior exchange principles*. New York: Brunner/Mazel.

Jacobson, N. S., McDonald, D. W., Follette, W. C., & Berley, R. A. (1985). Attribution processes in distressed and nondistressed married couples. *Cognitive Therapy and Research, 9,* 35–50.

Jacobson, N. S., & Moore, D. (1981). Spouses as observers of the events in their relationship. *Journal of Consulting and Clinical Psychology, 49,* 269–277.

Jacobson, N. S., Schmaling, K. B., & Holtzworth-Munroe, A. (1987). Component analysis of behavioral marital therapy: Two year follow-up and prediction of relapse. *Journal of Marital and Family Therapy, 13,* 187–195.

Jaffee, P., Wolfe, D. A., Wilson, S., & Zak, L. (1986). Emotional and physical health problems of battered women. *Canadian Journal of Psychiatry, 31,* 625–629.

Johnson, P. L., & O'Leary, K. D. (1996). Behavioral components of marital satisfaction: an individualized assessment approach. *Journal of Consulting and Clinical Psychology, 64,* 417–423.

Karney, B. R., & Bradbury, T. N. (1995). The longitudinal course of marital quality and stability: A review of theory, method, and research. *Psychological Bulletin, 118,* 3–34.

Kiecolt-Glaser, J. K., Malarkey, W. B., Chee, M. A., Newton, T., Cacioppo, J. T., Mao, H. Y., & Glaser, R. (1993). Negative behavior during marital conflict is associated with immunological down-regulation. *Psychosomatic Medicine, 55,* 395–409.

Kiecolt-Glaser, J. K., Newton, T., Cacioppo, J. T., MacCallum, R. C., Glaser, R., & Malarkey, W. B. (1996). Marital conflict and endocrine function: are men really more physiologically affected than women? *Journal of Consulting and Clinical Psychology, 64,* 324–332.

Klerman, G. L., & Weissman, M. M. (1982). Interpersonal psychotherapy: Theory and research. In A. J. Rush (Ed.), *Short-term psychotherapy for depression: Behavioral, interpersonal, cognitive and psychodynamic approaches*. New York: Guilford Press.

Koop, C. E. (1985). *The Surgeon General's workshop on violence and public health*. Washington, DC: Government Printing Office.

Lange, A., Schaap, C., & Van Widenfelt, B. (1993). Family therapy and psychotherapy: Developments in research and approaches to treatment. *Journal of Family Therapy, 15,* 113–146.

Lawton, J. M., & Sanders, M. R. (1994). Designing effective behavioral family interventions for stepfamilies. *Clinical Psychology Review, 14,* 463–496.

Levant, R. F. (1986). An overview of psychoeducational family programs. In R. F. Levant (Ed.), *Psychoeducational approaches to family therapy and counselling* (pp. 1–51). New York: Springer.

Levinger, G., & Huston, T. L. (1990). The social psychology of marriage. In F. D. Fincham & T. N. Bradbury (Eds.), *The psychology of marriage* (pp. 19–58). New York: Guilford Press.

Locke, H. J., & Wallace, K. M. (1959). Short marital adjustment and prediction tests: Their reliability and validity. *Marriage and Family Living, 2,* 251–255.

Mace, D. R., & Mace, V. C. (1976). Marriage enrichment: A preventive group approach for couples. In D. H. L. Olson (Ed.), *Treating relationships* (pp. 321–336). Lake Mills, IA: Graphic Publications.

Maisto, S. A., O'Farrell, T. J., Connors, G. J., McKay, J. R., & Pelcovits, M. (1988). Alcoholics' attributions of factors affecting their relapse to drinking and reasons for terminating relapse episodes. *Addictive Behavior, 13,* 79–82.

Markman, H. (1981). The prediction of marital distress: A five-year follow-up. *Journal of Consulting and Clinical Psychology, 49,* 760–762.

Markman, H. J. (1984). The longitudinal study of couples" interactions: Implications for understanding and predicting the development of marital distress. In K. Hahlweg & N. S. Jacobson (Eds.), *Marital interaction: Analysis and modification* (pp. 253–281). New York: Guilford Press.

Markman, H. J. (1991). Backwards into the future of couples therapy and couples therapy research: a comment on Jacobson. *Journal of Family Psychology, 4,* 416–425.

Markman, H. J., Floyd, F., Stanley, S. M., & Lewis, H. C. (1986). Prevention. In N. S. Jacobson & A. S. Gurman (Eds.), *Clinical handbook of marital therapy* (pp. 173–195). New York: Guilford Press.

Markman, H. J., Floyd, F., Stanley, S. M., & Storaasli, R. D. (1988). Prevention of marital distress: A longitudinal investigation. *Journal of Consulting and Clinical Psychology, 56,* 210–217.

Markman, H. J., & Hahlweg, K. (1993). The prediction and prevention of marital distress: An international perspective. *Clinical Psychology Review, 13,* 29–43.

Markman, H. J., Halford, W. K., & Cordova, A. (1997). Through a scientist-practitioner's lens: A grand tour of the study and promotion of healthy relationships. In W. K. Halford & H. J. Markman (Eds.), *Clinical handbook of marriage and couples interventions* (pp. 695–716). Chichester, UK: Wiley.

Markman, H. J., Renick, M. J., Floyd, F. J., Stanley, S. M., & Clements, M. (1993). Preventing marital distress through communication and conflict management training. A 4- and 5-year follow-up. *Journal of Consulting and Clinical Psychology, 61,* 70–77.

Markman, H. J., Stanley, S., & Blumberg, S. L. (1994). *Fighting for your marriage: Positive steps for preventing divorce and preserving a lasting love.* San Francisco: Jossey-Bass.

Martin, T. C., & Bumpass, L. (1989). Recent trends in marital disruption. *Demography, 26,* 37–51.

McDonald, P. (1995). *Families in Australia.* Melbourne, Australia: Australian Institute of Family Studies.

Medalie, J. H., & Goldbourt, U. (1976). Angina pectoris among 10,000: II. Psychosocial and other risk factors as evidenced by a multivariate analysis of a five year incidence study. *American Journal of Medicine, 60,* 910–921.

Miller, S., Nunnally, E. W., & Wackman, D. (1975). *Alive and aware: Improving communication in relationships.* Minneapolis, MN: Interpersonal Communication Program.

Millard, C. (1990). What marriage means to young adults. *Family Matters, 29,* 26–28.

Murphy, C. M., & O'Leary, K. A. (1989). Psychological aggression predicts physical aggression in early marriage. *Journal of Consulting and Clinical Psychology, 57,* 579–582.

Notarius, C. I., Benson, P. R., Sloane, D., Vanzetti, N. A. & Horyak, L. M. (1989). Exploring the interface between perception and behavior. An analysis of marital interaction in distressed and nondistressed couples. *Behavioral Assessment, 11,* 39–64.

Notarius, C., & Markman, H. J. (1994). *We can work it out: Making sense of marital conflict.* New York: Putnam.

Notarius, C., & Markman, H. J. (1994). *We can work it out: how to resolve conflict, save your marriage, and strengthen your love for each other.* New York: Berkley.

Notarius, C., & Vanzetti, N. (1984). The marital agenda protocol. In E. Filsinger (Ed.), *Marital and family assessment* (pp. 209–227). Newbury Park, CA: Sage.

O'Farrell, T. J. (1989). Marital and family therapy in alcoholism treatment. *Journal of Substance Abuse Treatment, 6,* 23–29.

O'Farrell, T. J., & Bayog, R. D. (1986). Antabuse contracts for married alcoholics and their spouses: A method to maintain Antabuse ingestion and decrease conflict about drinking. *Journal of Substance Abuse Treatment, 3,* 1–8.

O'Farrell, T. J., & Birchler, G. R. (1987). Marital relationships of alcoholic, conflicted, and nonconflicted couples. *Journal of Marital and Family Therapy, 13,* 259–274.

O'Farrell, T. J., Choquette, K. A., Cutter, H. S. G., Brown, E. D., & McCourt, W. F. (1992). Behavioral marital therapy with and without additional couples relapse prevention sessions for alcoholics and their wives. *Journal of Studies on Alcohol, 53,* 652–666.

O'Farrell, T. J, Cutter, H. S. G., & Floyd, F. J. (1985). Evaluating behavioral marital therapy for male alcoholics: Effects on marital adjustment and communication from before to after treatment. *Behavior Therapy, 16,* 147–167.

O'Farrell, T. J., Cutter, H. S. G., Choquette, K. A., Floyd, F. J. & Bayog, R. D. (1992). Behavioral marital therapy for male alcoholics: Marital and drinking adjustment during the two years after treatment. *Behavior Therapy, 23,* 529–549.

O'Farrell, T. J., Kleinke, C. L., Thompson, D. L., & Cutter, H. S. G. (1986). Differences between alcoholic couples accepting and rejecting an offer of outpatient marital therapy. *American Journal of Drug and Alcohol Abuse, 12,* 285–294.

O'Farrell, T. J., & Murphy, C. M. (1995). Marital violence before and after alcoholism treatment. *Journal of Consulting and Clinical Psychology, 63,* 256–262.

O'Farrell, T. J., & Rotunda, R. J. (1997). Couple interventions and alcohol abuse. In W. K. Halford & H. J. Markman (Eds.), *Clinical handbook of marriage and couples interventions* (pp. 555–618). Chichester, UK: Wiley.

O'Leary, K. D. (Ed.) (1977). *Assessment of marital discord: an integration for research and clinical practice.* Hillsdale, NJ: Erlbaum.

O'Leary, K. D., Barling, J., Arias, I., Rosenbaum, A., Malone, J., & Tyree, A. (1989). Prevalence and stability of physical aggression between spouses: A longitudinal analysis. *Journal of Consulting and Clinical Psychology, 57,* 263–268.

O'Leary, K. D., & Beach, S. R. H. (1990). Marital therapy: A viable treatment for depression and marital discord. *American Journal of Psychiatry, 147,* 183–186.

O'Leary, K. D., & Vivian, D. (1990). Physical aggression in marriage. In F. D. Fincham & T. N. Bradbury (Eds.), *The psychology of marriage* (pp. 323–348). New York: Guilford Press.

O'Leary, K. D., Vivian, D., & Malone, J. (1992). Assessment of physical aggression against women in marriage: The need for multimodal assessment. *Behavioral Assessment, 14,* 5–14.

Olson, D. H., & Fowers, B. J. (1986). Predicting marital success with PREPARE: A predictive validity study. *Journal of Marital and Family Therapy, 12,* 403–413.

Olson, D. H., & Larsen, A. S. (1989). Predicting marital satisfaction using PREPARE: A replication study. *Journal of Marital and Family Therapy, 15,* 311–322.

Prado, L., & Markman, H. J. (in press). For better or worse the second time around: analyzing the communication patterns of remarried couples. In M. Cox & J. Brooks-Gunn (Eds.), *Conflict and cohesion.* Hillsdale, NJ: Erlbaum.

Raush, H. L., Barry, A., Hertel, R. K., & Swain, M. A. (1974). *Communication, conflict, and marriage.* San Francisco: Jossey-Bass.

Reich, J., & Thompson, W. D. (1985). Marital status of schizophrenic and alcoholic patients. *Journal of Nervous and Mental Disease, 173,* 499–502.

Reynolds, I., Rizzo, D., Gallagher, H., & Speedy, B. (1979). *Psychosocial problems of Sydney adults.* Sydney, Australia: Health Commission of New South Wales.

Robinson, E. A., & Price, M. G. (1980). Pleasurable behavior in marital interaction: An observational study. *Journal of Consulting and Clinical Psychology, 48,* 117–118.

Rousanville, B. J., Weissman, M. M., Prusoff, B. A., & Herceg-Baron, R. L. (1979). Marital disputes and treatment outcome in depressed women. *Comprehensive Psychiatry, 20,* 483–490.

Ruscher, S. M., & Gotlib, I. H. (1988). Marital interaction patterns of couples with and without a depressed partner. *Behavior Therapy, 19,* 455-470.

Rutledge, A. L. (1968). A systematic approach to premarital counselling. In J. C. Heston & W. B. Frich (Eds.), *Counselling for the liberal arts campus.* Yellow Springs, OH: Antioch Press.

Sanders, M. R. (Ed.) (1995). *Healthy families healthy nation.* Brisbane, Australia: Australian Academic Press.

Sanders, M. R., Halford, W. K., & Behrens, B. C. (1998). Parental divorce and premarital interactions. An observational study of interactional correlates of parental divorce. Paper under review.

Sanders, M. R., Nicholson, J. M., & Floyd, F. (1997). Couples relationships and children. In W. K. Halford & H. J. Markman (Eds.), *Clinical handbook of marriage and couples interventions* (pp. 507–536). Chichester, UK: Wiley.

Sayers, S. L., Baucom, D. H., Sher, T. G., Weiss, R. L., & Heyman, R. E. (1991). Constructive engagement, behavioral marital therapy, and changes in marital satisfaction. *Behavioral Assessment, 13,* 25–49.

Schaap, C. (1984). A comparison of the interaction of distressed and nondistressed married couples in a laboratory situation: Literature survey, methodological issues, and an empirical investigation. In K. Hahlweg & N. S. Jacobson (Eds.), *Marital interaction: Analysis and modification* (pp. 133–158). New York: Guilford Press.

Schindler, L., & Vollmer, M. (1984). Cognitive perspectives in behavioral marital therapy: Some proposals for bridging theory, research, and practice. In K. Hahlweg & N. S. Jacobson (Eds.), *Marital interaction: Analysis and modification.* New York: Guilford Press.

Schmaling, K. B., & Sher, T. G. (1997). Physical health and relationships. In W. K. Halford & H. J. Markman (Eds.), *Clinical handbook of marriage and couples interventions* (pp. 323–348). Chichester, UK: Wiley.

Seligman, M. E. P. (1995). The effectiveness of psychotherapy: The Consumer Reports study. *American Psychologist, 50,* 965–974.

Shadish, W. R., Montgomery, L. M., Wilson, P., Wilson, M. R., Bright, I., & Okwumabua, T. (1993). Effects of family and marital psychotherapies: A meta-analysis. *Journal of Consulting and Clinical Psychology, 61,* 992–1002.

Simons, M., Harris, R., & Willis, R. (1994). *Pathways to marriage: learning for married life in Australia.* Adelaide, Austrialia: University of South Australia Press.

Sisson, R. W., & Azrin, N. H. (1986). Family member involvement to initiate and promise treatment of problem drinkers. *Journal of Behavior Therapy and Experimental Psychiatry, 17,* 15–21.

Skuja, K., & Halford, W. K. (1998). Violence in the family of origin and communication in dating relationships. Paper under review.

Snyder, D. K., Mangrum, L. F., & Wills, R. M. (1993). Predicting couples response to marital therapy: A comparison of short- and long-term predictors. *Journal of Consulting and Clinical Psychology, 61,* 61–69.

Snyder, D. K., & Wills, R. M. (1989). Behavioral versus insight-oriented marital therapy: Effects on individual and interspousal functioning. *Journal of Consulting and Clinical Psychology, 57,* 39–46.

Snyder, D. K., Wills, R. M., & Grady-Fletcher, A. (1991a). Long term effectiveness of behavioral versus insight orientated marital therapy: A four year follow-up study. *Journal of Consulting and Clinical Psychology, 59,* 138–141.

Snyder, D. K., Wills, R. M & Grady-Fletcher, A. (1991b). Risks and challenges of long-term psychotherapy outcome research: Reply to Jacobson. *Journal of Consulting and Clinical Psychology, 59,* 146–149.

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family, 38,* 15–28.

Stanley, S., Lobitz, C., & Markman, H. J. (1989). Marital therapy in Colorado. *Colorado Psychological Association Bulletin, May,* 12–18.

Stanley, S., & Markman, H. J. (1996). *National survey of couples relationships in the United States.* University of Denver.

Stanley, S., Markman, H. J., Leber, D., & St. Peters, M. (1995). Strengthening marriages and predicting divorce: new directions in prevention research. *Family Relations, 44,* 392–401.

Stets, J. E., & Strauss, M. A. (1990). The marriage licence as a hitting licence: A comparison of dating, cohabiting and married couples. In M. A. Strauss & R. J. Gelles (Eds.), *Physical violence in American families: Risk factors and adaption to violence in 8415 families* (pp. 131–164). NJ: Transaction Press.

Strauss, M. A. (1979). Measuring intrafamily conflict and violence: The Conflict Tactics Scale. *Journal of Marriage and the Family, 41,* 75–88.

Strauss, M. A., & Gelles, R. (1986). Societal change and change in family violence from 1975 to 1985 as revealed by two national surveys. *Journal of Marriage and the Family, 48,* 465–479.

Strauss, M. A., Gelles, R., & Steinmetz, S. K. (1980). *Behind closed doors: Violence in the American family.* New York: Doubleday.

Stuart, R. B. (1969). Operant-interpersonal treatment for marital discord. *Journal of Consulting and Clinical Psychology, 33,* 657–682.

Terman L. M. (1938). *Psychological factors in marital happiness.* New York: McGraw-Hill.

Thomas, E. J., & Ager, R. D. (1993). Unilateral family therapy with spouses of uncooperative alcohol drinkers. In T. J. O'Farrell (Ed.), *Treating alcohol problems: Marital and family interventions* (pp. 3–33). New York: Guilford Press.

Thompson, B. T. (1997). Couples and the work-family interface. In W. K. Halford & H. J. Markman (Eds.), *Clinical handbook of marriage and couples interventions* (pp. 273–290). Chichester, UK: Wiley.

Van Widenfelt, B., Markman, H. J., Guerney, B., Behrens, B., & Hosman, C. (1997). Prevention of relationship problems. In W. K. Halford & H. J. Markman (Eds.), *Clinical handbook of marriage and couples interventions* (pp. 651–678). Chichester, UK: Wiley.

Van Widenfelt, B. M., Schaap, C., & Hosman, C. (1996). The prevention of relationship distress: a controlled evaluation with nine month and two year follow-ups. *Family Relations, 45,* 156–165.

Vanicelli, M., Gingereich, S., & Ryback, R. (1983). Family problems related to the treatment and outcome of alcoholic patients. *British Journal of Addiction, 78,* 193–204.

Vanzetti, N. A., Notarius, C. I., & NeeSmith, D. (1992). Specific and generalized expectancies in marital interaction. *Journal of Family Psychology, 6,* 171–183.

Veroff, J., Kulka, R. A. M., & Douvan, E. (1981). *Mental health in America: Patterns of help-seeking from 1957 to 1976.* New York: Basic Books.

Walsh, F., Jacob, L., & Simons, V. (1995). Facilitating healthy divorce processes: therapy and mediation issues. In N. S. Jacobson & A. S. Gurman (Eds.), *Clinical handbook of couples therapy* (pp. 340–368). New York: Guilford Press.

Weiss, R. L. (1984). Cognitive and behavioral measures of marital interaction. In K. Hahlweg & N. S. Jacobson (Eds.), *Marital interaction: Analysis and modification* (pp. 232–252). New York: Guilford Press.

Weiss, R. L., Birchler, G. R., & Vincent, J. P. (1974).

Contractual models for negotiation training in marital dyads. *Journal of Marriage and the Family, 36,* 321–330.

Weiss, R. L., & Cerreto, M. S. (1980). The Marital Status Inventory: Development of a measure of dissolution potential. *American Journal of Family Therapy, 8,* 80–85.

Weiss, R. L., & Halford, W. K. (1995). Managing couples therapy. In V. Van Hasselt & M. Hersen (Eds.), *Sourcebook of psychological treatment manuals for adult disorders* (pp. 312–341). New York: Plenum.

Weiss, R. L., & Heyman, R. (1990). Marital distress and therapy. In A. S. Bellack, M. Hersen, & A. Kazdin (Eds.), *International handbook of behavior modification* (2nd ed., pp. 475–502). New York: Plenum.

Weiss, R. L., & Heyman, R. (1997). A clinical-research overview of couples' interactions. In W. K. Halford & H. J. Markman (Eds.), *Clinical handbook of marriage and couples therapy* (pp. 13–42). Chichester, UK: Wiley.

Weiss, R. L., Hops, H., & Patterson, G. R. (1973). A framework for conceptualizing marital conflict: A technology for altering it, some data for evaluating it. In L. A. Hamerlynck, L. C. Handy, & E. J. Marsh (Eds.), *Behavior change: Methodology, concepts, and practice* (pp. 309–342). Champaign, IL: Research Press.

Weiss, R. L., & Perry, B. A. (1983). The spouse observation checklist. In E. E. Filsinger (Ed.), *A source book of marriage and family assessment*. Beverly Hills, CA: Sage.

Weissman, M. M. (1987). Advances in psychiatric epidemiology: Rate and risks for major depression. *American Journal of Public Health, 77,* 445–451.

Whisman, M. A., & Snyder, D. (1997). Evaluating and improving the efficacy of conjoint couple therapy. In W. K. Halford & H. J. Markman (Eds.), *Clinical handbook of marriage and couples interventions* (pp. 679–694). Chichester, UK: Wiley.

Widom, C. S. (1989). Does violence beget violence? A critical examination of the literature. *Psychological Bulletin, 106,* 3–28.

Wolcott, I., & Glazer, H. (1989). *Marriage counselling in Australia: An evaluation*. Melbourne, Australian: Australian Institute of Family Studies.

Zimmer, D. (1983). Interaction patterns and communication skills in sexually distressed, maritally distressed, and normal couples: Two experimental studies. *Journal of Sex and Marital Therapy, 9,* 251–266.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

6.29
Eating Disorders

ANITA JANSEN
Universiteit Maastricht, The Netherlands

6.29.1 INTRODUCTION 649
6.29.2 ANOREXIA NERVOSA 650
6.29.3 BULIMIA NERVOSA 651
6.29.4 OBESITY AND BINGE EATING DISORDER 652
6.29.5 ATYPICAL EATING DISORDERS 653
6.29.6 DIFFERENTIAL DIAGNOSES AND COMORBIDITY 653
  6.29.6.1 Differential Diagnoses 653
  6.29.6.2 Comorbidity 654
6.29.7 PREVALENCE AND PROGRESS 654
6.29.8 ETIOLOGY 655
  6.29.8.1 Genetics 655
  6.29.8.2 Childhood Sexual Abuse 656
  6.29.8.3 Dieting for the Perfect Figure 656
  6.29.8.4 Serotonin Depletion 657
  6.29.8.5 The Learned Nature of Binge Eating 659
  6.29.8.6 Dysfunctional Cognitions 659
6.29.9 TREATMENT 661
  6.29.9.1 Anorexia Nervosa 661
  6.29.9.2 Bulimia Nervosa 662
  6.29.9.3 Obesity and Binge Eating Disorder 663
  6.29.9.4 Predictors of Treatment Outcome 663
6.29.10 IN CONCLUSION 663
6.29.11 REFERENCES 664

6.29.1 INTRODUCTION

Nowadays, more than half of all women diet because they believe they are too fat (e.g., Agras, 1990). In the late 1970s, one in every five of a sample of British women binged once a month, and 10% of these women regurgitated after bingeing in order to get the quantity of food eaten out of the system as quickly as possible (Wardle, 1980). Eating behaviors of this type can be problematic, but they are seen so frequently that they can no longer justifiably be categorized as "abnormal."

This chapter discusses eating disorders. Eating disorders differ from the problems mentioned above in the frequency and seriousness of the symptoms and the extent to which they interfere with the individual's functioning. With an eating disorder, the symptoms are much more serious and frequent than with an eating problem. Unlike an eating problem, an eating disorder is often dysfunctional: the patient suffers from it to an extent that it is virtually impossible to function normally. This is not always evident to others; the eating disorder bulimia nervosa, for example, can be concealed by sufferers for many years.

The eating disorders anorexia nervosa and bulimia nervosa, as well as atypical variations of these disorders and the recently identified binge eating disorder, are discussed in the present chapter. Because eating disorders are primarily seen among women, patients will be referred to as women. After a clinical description of the eating disorders, the diagnostic features, differential diagnoses, comorbidity and prevalence are discussed. Next, empirically validated theories of the origin and maintenance of eating disorders are considered. A final section discusses the effectiveness of the various treatments currently available.

6.29.2 ANOREXIA NERVOSA

The characteristic of a patient with anorexia nervosa is that she continually assumes that she is too fat. This conviction fuels an intense drive to be thinner, sometimes referred to as a relentless pursuit of thinness. The fervent desire to lose weight continually dominates the patient's thoughts, feelings, and behavior (Garfinkel, 1995). She eats only low-calorie food and exerts an enormous amount of physical effort, resulting in a loss of weight. Sleeping problems and depression can result from malnutrition. It is striking that neither the idea of being too fat nor the fear of gaining weight disappear as the weight loss continues. On the contrary, as weight decreases, the fear of becoming fat intensifies. Another striking aspect is that the patient usually denies the seriousness of the loss of weight.

Incessant attempts to lose an extreme amount of weight, and the belief among these underweight patients that they are still too fat, resulted in recent decades in the suspicion by numerous specialists and researchers that a distorted perception of the body is characteristic of patients with anorexia nervosa. Patients with anorexia nervosa were presumed to overestimate the size of their body. However, this clinical impression could not be irrefutably confirmed. Researchers asked patients, for example, if they could estimate the size of their body while viewing themselves on a video monitor. By turning knobs they could show on the screen how big they thought they were. It was learned that some patients with anorexia nervosa overestimate the size of their body more than others, while other patients with anorexia nervosa did not. Moreover, it was determined that people without an eating disorder also often overestimate their size (Garfinkel, 1995; Smeets, Smit, Panhuysen, & Ingleby, 1997).

The group of anorexia nervosa patients can be divided into two subgroups: the binge/purging type and the restricting type (American Psychiatric Association [APA] 1994). Purging refers to behavior oriented towards getting the food out of the body as quickly as possible: self-induced vomiting and the use of laxatives and diuretics (substances that stimulate the excretion of fluids: "pee pills"). Dieting, fasting, and sports activities are not considered purging behavior. An average of nine months after anorexia nervosa commences, about half of the patients start to suffer from uncontrollable bingeing and purging behavior (Polivy & Herman, 1985; Wardle & Beinart, 1981).

Purgers often stimulate the vomiting reflex by sticking a finger, toothbrush, or other oblong object into the desired position in the throat. After some time, some patients can stimulate this reflex simply by exerting a small amount of pressure on the stomach under the ribs. Vomiting is then more or less "automatic." Some chronic vomiters have a hard spot or callus on the back of the hand or under the ribs. This callus results from the hand rubbing along the teeth or from exerting pressure on the stomach.

The use of laxatives is not as effective as the patient often believes. Laxatives rinse food through the intestines at an accelerated speed. This rapid journey through the digestive track can prevent complete absorption of the nutrients, as a result of which somewhat less energy (kilocalories) is absorbed. The weight loss achieved in this manner, however, is small and short-lived because the laxatives primarily stimulate the excretion of fluids. The weight lost is regained after drinking a few glasses of fluids. A paradoxical effect of the long-term use of laxatives is that constipation occurs because of the reduced or complete loss of activity of the intestines.

Patients of the other, restricting type do not suffer from binges. These patients continually restrict their intake of food and therefore do not turn as easily to the purging measures explained above. Which factors determine when and why a patient changes from a restricting type to a purging type with binges is unknown. The two groups do, however, differ in a number of other ways: on average, the bingeing/purging type has a higher weight before the anorexia nervosa started and suffers more often from disorders involving controlling impulses, excessive use of alcohol or drugs, mood disorders, and personality disorders (Garfinkel, 1995).

Malnutrition and purging behavior have medical consequences. The menstrual cycle is usually interrupted, and when underweight becomes extreme the menses cease. The extremities (fingers, toes) are cold and turn purple-blue (this is referred to as cyanosis), the body temperature is low (hypothermia), which can result in a downy layer of hair growing on the body (lanugo), the blood pressure is low (hypotension), and the heartbeat is slow (often lower than 60 beats per minute: in medical terms bradycardia). The metabolism is also often slow. The physical consequences of malnutrition often disappear once the eating pattern and weight have returned, but this can take some time.

The diagnostic criteria for anorexia nervosa according to the *Diagnostic and statistical manual of mental disorders* (4th edn., *DSM-IV*) are listed in Table 1.

6.29.3 BULIMIA NERVOSA

A chaotic eating pattern, weight-control measures, and irrational ideas about one's figure, appearance, and weight are characteristic of the eating disorder bulimia nervosa (APA, 1994). Normally, periods of sparse eating alternate with binge eating episodes (bulimia literally means "as hungry as an ox"). During a binge, in an uninterrupted period of time an amount of food is consumed that is considerably larger than most people would normally eat in the same period of time. Typical is the subjective feeling of loss of control over eating: during the binge, the patient feels as if she can no longer stop or control what or how much she eats (Fairburn & Wilson, 1993). The binge can be induced by a variety of stimuli: fluctuating moods (e.g., depression, fear, euphoria), the breaking of idiosyncratic diet rules, for example, by eating "forbidden" food (i.e., food rich in calories), and a craving for food stimulated by the smell of palatable food (Jansen & van den Hout, 1991). Some patients describe a type of dissociative amnesia: they experience their binge as a dream and cannot remember the details. After bingeing, in the very short term, the binge improves affect, but very soon thereafter (often even before the end of the binge) the subject feels guilty and self-disgusted. The patient is afraid of gaining weight and seeks refuge in the compensational behavior described above: purging and dieting.

The chaotic eating pattern (alternating between much and little) results from extreme worry about the figure, appearance, and weight. Like patients with anorexia nervosa, despite the fact that their weight is often normal, patients with bulimia nervosa believe they are heavy, slow, and fat (Phelan, 1987). These irrational ideas about their figure and weight determine to a significant degree their self-esteem, which is often low.

Patients with bulimia nervosa also have irrational cognitions about things other than their body shape and weight. Their ideas about food and eating are also dysfunctional (Dritschel, Williams, & Cooper, 1991). Everything that is edible is either "good" or "bad." Good food is food that has a low amount of kilocalories (e.g., skimmed yogurt, bouillon, and lettuce); bad food is food with a high amount of calories, and is "forbidden" (e.g., chocolate, pizza, and fried meat). Eating a bit of forbidden food is "completely" bad. This dichotomous (black-and-white) thinking style is an invitation to failure: the stricter the rules they apply to themselves, the more likely the rules are to be broken.

Bulimia nervosa, like anorexia nervosa, is often accompanied by medical complaints that result from the binges, vomiting, fasting, and use of laxatives and diuretics. Most of these complaints are reversible and disappear as soon as the eating behavior has normalized. Some, however, are life threatening. When vomiting

Table 1 *DSM-IV* diagnostic criteria for 307.1 anorexia nervosa.

A. Refusal to maintain body weight at or above a minimally normal weight for age and height (e.g., weight loss leading to maintenance of body weight less than 85% of that expected; or failure to make expected weight gain during period of growth, leading to body weight less than 85% of that expected)
B. Intense fear of gaining weight or becoming fat, even though underweight
C. Disturbance in the way in which one's body weight or shape is experienced, undue influence of body weight or shape on self-evaluation, or denial of the seriousness of the current low body weight
D. In postmenarcheal females, amenorrhea, i.e., the absence of at least three consecutive menstrual cycles (a woman is considered to have amenorrhea if her periods occur only following hormone, e.g., estrogen, administration)

Specify type:

Restricting type: during the current episode of anorexia nervosa, the person has not regularly engaged in binge-eating or purging behavior (i.e., self-induced vomiting or the misuse of laxatives, diuretics, or enemas)

Binge-eating/purging type: during the current episode of anorexia nervosa, the person has regularly engaged in binge-eating or purging behavior (i.e., self-induced vomiting or the misuse of laxatives, diuretics, or enemas) and using laxatives and diuretics, the patient loses a significant amount of fluids. This dehydration may result in thirst, dry skin, dizziness when rising from a sitting or horizontal position, light-headedness, and fainting. Chronic purging can also result in serious disorders of the electrolyte balance. Electrolytes (the most important of which are sodium, potassium, and chloride) are called just that because they are responsible for the transmission of electrical signals between the cells in our body. They play an important role in metabolic processes and ensure that the nerve and muscle cells function properly. As a result of purging, a considerable quantity of electrolytes, in particular sodium and potassium, are excreted. Serious shortages can be life threatening because the heart muscle is weakened and receives insufficient impulses to compress and then relax. A simple blood test can indicate whether the electrolyte levels are abnormal.

The *DSM-IV* criteria for bulimia nervosa are listed in Table 2.

6.29.4 OBESITY AND BINGE EATING DISORDER

Obesity refers to an excess of body fat. Because body weight and body fat are highly correlated, overweight is often used as a proxy for obesity (Brownell, 1995; Wardle, 1995). To determine whether someone's weight is in or outside the normal range, the body mass index (BMI) is generally used. The BMI refers to the ratio of height and weight; it is calculated by dividing body weight (in kilograms) by height^2 (in meters). For example, the BMI of a subject weighing 63 kg and 1.68 m tall is: $63/(1.68)^2 = 22.3$. A BMI between 20 and 25 shows a healthy weight. The more a BMI drops below 20, the more severe the underweight is, and, vice versa, the more the BMI rises above 25, the more severe the overweight (see Table 3). Levels of BMI greater than 30 are usually taken to indicate obesity.

In most cases, obesity develops in the absence of any medical disease (Foster, 1992), but it is a risk factor for a range of major health hazards, including cardiovascular diseases, hypertension, diabetes, stroke, hypercholesterolemia, gallbladder diseases, and osteoarthritis (Wardle, 1995; Willett & Manson, 1995). Statistics from life insurance companies show a significant increase of morbidity and mortality when the BMI rises above 30. The health consequences are more serious for the obese with fat around the waist (the "apple") than for obese subjects with fat around the hips and thighs (the "pear").

Stigmatization, prejudice, and discrimination of the obese, for example, in relation to acceptance for a job, makes obesity a psychological problem as well. Obesity is considered unattractive in Western cultures, leading to negative self-images and even self-hate in obese persons, especially women (Foster, 1992; Stunkard & Wadden, 1992). Children as young as six years of age describe silhouettes of obese children as "lazy, dirty, stupid, ugly, cheats, and lies," and they rate children with a variety of disabilities as more attractive than the obese children (Rodin, Silberstein, & Striegel-Moore, 1984). Low self-esteem, feelings of anxiety, depression, and guilt are frequently seen in the seriously obese (Stunkard & Wadden, 1992).

Following the definitions of *DSM-IV*, obesity is not defined as an eating disorder or even as a

Table 2 *DSM* diagnostic criteria for 307.51 bulimia nervosa.

A. Recurrent episodes of binge eating. An episode of binge eating is characterized by both of the following:
  (1) eating, in a discrete period of time (e.g., within any 2 hour period), an amount of food that is definitely larger than most people would eat during a similar period of time and under similar circumstances
  (2) a sense of lack of control over eating during the episode (e.g., a feeling that one cannot stop eating or control what or how much one is eating)
B. Recurrent inappropriate compensatory behavior in order to prevent weight gain, such as self-induced vomiting, misuse of laxatives, diuretics, enemas, or other medications; fasting; or excessive exercise
C. The binge eating and inappropriate compensatory behaviors both occur, on average, at least twice a week for 3 months
D. Self-evaluation is unduly influenced by body shape and weight
E. The disturbance does not occur exclusively during episodes of anorexia nervosa

Specify type:

Purging type: during the current episode of bulimia nervosa, the person has regularly engaged in self-induced vomiting or the misuse of laxatives, diuretics, or enemas

Nonpurging type: during the current episode of bulimia nervosa, the person has used other inappropriate compensatory behaviors, such as fasting or excessive exercise, but has not regularly engaged in self-induced vomiting or the misuse of laxatives, diuretics, or enemas

Table 3 Body Mass Index.

BMI	*Weight category*
<18	Severe underweight
18–20	Underweight
20–25	Normal weight
25–27	Slightly overweight
27–30	Overweight
>30	Obesity
30–35	Mild obesity
35–40	Moderate obesity
>40	Morbid or severe obesity

Body Mass Index (BMI) = weight/height^2; kg/m^2.

mental disorder. Until recently, obesity has generally been studied by doctors and biologists, whereas eating disorders have always been within the scope of psychologists and psychiatrists. However, in the last few years, interest in obesity among psychologists and psychiatrists has been growing. One of the main reasons for this renewed interest is the recognition of increased psychopathology in some obese people.

This subgroup of the obese population frequently overeat in a way which could be described as bingeing. Some of these obese binge eaters meet the diagnostic criteria of bulimia nervosa, but most of them do not. Those who do not meet the criteria binge regularly but do not engage in the characteristic compensatory purging behaviors of bulimia nervosa, that is, self-induced vomiting or use of laxatives. *DSM-IV* considered this a possible eating disorder and called it the Binge Eating Disorder. In the diagnostic handbook, binge eating disorder is included in an appendix listing clinical presentations which must be studied in detail before they can be definitely included in a diagnostic category in the handbook. In Table 4 the preliminary *DSM-IV* criteria for binge eating disorder are listed.

Obese subjects who meet the criteria for the binge eating disorder differ on several characteristics from obese persons who do not meet the binge eating disorder criteria (Brody, Walsh, & Devlin, 1994; Castonguay, Eldredge, & Agras, 1995; Fichter, Quadflieg, & Brandl, 1993; Lowe & Caputo, 1991; Spitzer et al., 1993). Obese binge eaters in general have a higher body mass index, report more weight fluctuations, show an earlier onset of obesity, spend more time dieting, are more restrained in their eating patterns, drop out of treatment more frequently, have poorer treatment outcomes, and show more psychopathology than the obese nonbingers.

Observed links between obesity and psychopathology (see Section 6.29.6.2) are mostly accounted for by binge eating; it is the severity of binge eating and not the severity of obesity which is correlated with the indices of psychopathology (Castonguay et al., 1995). Note, however, that the obese nonpurging binge eaters show less psychopathology than the purging bingers, i.e. subjects with bulimia nervosa.

6.29.5 ATYPICAL EATING DISORDERS

About one-third of the patients who presented in England for the treatment of an eating disorder failed to meet the diagnostic criteria for anorexia or bulimia nervosa (Fairburn & Walsh, 1995). They normally have an atypical eating disorder or, in *DSM-IV* terms, an eating disorder "not otherwise specified" (APA, 1994). These are people who, for example, meet all of the criteria for anorexia nervosa except that they still have a regular menstrual cycle, or they have lost a considerable amount of weight, but the weight is not low enough. A patient may also meet all of the criteria for bulimia nervosa except that the binges do not occur frequently enough.

6.29.6 DIFFERENTIAL DIAGNOSES AND COMORBIDITY

6.29.6.1 Differential Diagnoses

Considerable weight loss can occur as a result of certain physical illnesses (e.g., cancer, intestinal disorders, and AIDS), depression, schizophrenia, and others. Weight loss of this type must be distinguished from anorexia nervosa if the patient fails to meet the other diagnostic criteria for the eating disorder, such as an extreme fear of being fat and the desire to continue to lose weight.

Binges can occur as a result of schizophrenia, depression, a borderline personality disorder, and certain neurological disorders such as the Kleine–Levy syndrome. If these patients do not show compensation behavior and do not appear extremely overly concerned with their figure, they do not meet the criteria for bulimia nervosa.

If patients with an eating disorder handle food compulsively or are afraid of eating in social situations, they do not necessarily suffer from an obsessive-compulsive disorder or from social anxiety. It must first be determined whether the patient also meets the other criteria for these disorders and whether, for example, they also suffer from obsessions, compulsions, or social fears that are not related to food or the figure.

Table 4 *DSM-IV* research criteria for Binge Eating Disorder.

A. Recurrent episodes of binge eating. An episode of binge eating is characterized by both of the following:
  (1) eating, in a discrete period of time (e.g., within any 2 hour period), an amount of food that is definitely larger than most people would eat in a similar period of time under similar circumstances
  (2) a sense of lack of control over eating during the episode (e.g., a feeling that one cannot stop eating or control what or how much one is eating)
B. The binge eating episodes are associated with three (or more) of the following:
  (1) eating much more rapidly than normal
  (2) eating until feeling uncomfortably full
  (3) eating large amounts of food when not feeling physically hungry
  (4) eating alone because of being embarrassed by how much one is eating
  (5) feeling disgusted with oneself, depressed, or very guilty after overeating
C. Marked distress regarding binge eating is present
D. The binge eating occurs, on average, at least 2 days a week for 6 months
  Note: The method of determining frequency differs from that used for bulimia nervosa; future research should address whether the preferred method of setting a frequency threshold is counting the number of days on which binges occur or counting the number of episodes of binge eating
E. The binge eating is not associated with the regular use of inappropriate compensatory behaviors (e.g., purging, fasting, excessive exercise) and does not occur exclusively during the course of anorexia nervosa or bulimia nervosa

6.29.6.2 Comorbidity

Obese bingers more often than obese nonbingers are characterized by depression, anxiety, personality disorders, low self-esteem, impulsivity, body perception disturbances, neuroticism, and overall distress. Obese binge eaters significantly more often meet criteria for a psychiatric disorder than obese nonbinge eaters (60% vs. 28%; Castonguay et al., 1995). Some authors consider depression a risk factor for the occurrence of binges among people who are overweight (Marcus, 1995).

Nearly half of women suffering from anorexia nervosa and bulimia nervosa are depressive, whereas others show obsessions and compulsions or other anxiety disorders. Their first-degree relatives suffer from mood disorders more often than could be expected by chance (Cooper, 1995a). In particular, patients who regularly purge are clearly more depressed and show more anxiety than those who do not regularly purge (Garner, 1993). This is why from time immemorial there have been experts who believe that eating disorders are a variation of mood and/or anxiety disorders (Garfinkel, 1995), but the symptoms of anxiety disorders and depression occurring during the eating disorder are virtually always secondary to the eating disorder. When the eating disorder disappears, the depressive and/or phobic characteristics often disappear as well (Fairburn & Cooper, 1989).

Recently, Welch and Fairburn (1996b) found that deliberate self-harm by repeated cutting was much more common among a community sample of bulimia nervosa subjects (19%) than among a psychiatric (3%) and a normal (1%) control group. Although it is a general clinical impression that subjects with eating disorders show a high incidence of psychoactive substance abuse, controlled epidemiological studies do not confirm this (Welch & Fairburn, 1996b; Wilson, 1991). Also, a borderline personality disorder is considered to be much more prevalent among subjects with eating disorders, in particular bulimia nervosa, than among other psychiatric populations. Rates have ranged from 2% to 47% (Wonderlich, 1995). However, methodologically sound studies are rare; the main problems in these studies are the absence of (psychiatric) control groups and the diagnostic problems related to the definition of the borderline personality disorder. There is considerable overlap between the diagnostic criteria for eating disorders and the diagnostic criteria for the borderline personality disorder. The latter includes, for example, binge eating, mood fluctuations, unstable self-image, suicidal or self-mutilating behavior, feelings of boredom and emptiness. Because of the overlap in criteria, there is a risk of overdiagnosing the borderline personality disorder among (depressed) subjects with eating disorders. This idea is supported by data that indicate a decrease in the rate of borderline personality disorders with short-term, symptom-focused interventions (Wonderlich, 1995).

6.29.7 PREVALENCE AND PROGRESS

Anorexia and bulimia nervosa are seen significantly more often among women than men. About 10% of the patients are male (APA, 1994). Epidemiological research into the prevalence of the two eating disorders suggests that about 0.5–1% of girls and young adult women

(18–30 years) suffer from anorexia nervosa, while bulimia nervosa is seen in 1–3% of this group (APA, 1994; Fairburn & Beglin, 1990). Atypical eating disorders are considered to be much more common (APA, 1994; Fairburn & Walsh, 1995), but how common is currently unknown. It is difficult to collect exact data concerning the prevalence of eating disorders. In addition to the fact that women with anorexia nervosa often deny that they have an eating disorder and women with bulimia nervosa often keep their problem hidden because they are ashamed, research of this type requires precise, unambiguous application of the diagnostic criteria.

Some researchers believe that the prevalence of the two eating disorders has increased in recent decades (APA, 1994). Epidemiological studies, however, do not show an upward trend in rates (Fombonne, 1996). The alleged increase could be a result of the increased publicity about and familiarity with eating disorders, as a result of which cases are identified more quickly. Patients or relatives turn to their family physician or another expert more readily, and the diagnostic process and the available assistance have improved, so that an eating disorder is more readily recognized by all concerned.

Obesity is much more common than the eating disorders anorexia and bulimia nervosa. In the USA 12% of adult men and 15% of adult women are obese, in the UK 8% and 12%, respectively (Wardle, 1995). Binge eating disorder is primarily seen in overweight people and one and a half times more often in females than in males (APA, 1994; Marcus, 1995). About 2–5% of overweight nonpatient community subjects meet criteria for the binge eating disorder, compared with one-third of overweight people who present for treatment. The higher the BMI, the greater the likelihood of binge eating.

Obesity is more prevalent in lower than in higher socioeconomic groups (Foster, 1992), whereas anorexia and bulimia nervosa occur to about the same extent in all social classes (Fairburn & Beglin, 1990; Gard & Freeman, 1996; Rogers, Resnick, Mitchell, & Blum, 1997) and appear to be more common in industrialized countries, where there is an excess of food and the "perfect figure" culture prevails, than in undeveloped countries. However, we do not know this for certain, because systematic studies of the prevalence in other cultures are seldom made. Women in certain professions run a higher risk than other women: anorexia nervosa is seen relatively more frequently among ballet dancers, models, and athletes (professions in which a low body weight is an advantage), but it is not clear whether these professions stimulate the eating disorder or people with (a predisposition for) an eating disorder choose these professions.

Anorexia and bulimia nervosa usually start prior to the age of 18. In 10% of the cases in which hospitalization is required, anorexia nervosa results in death (APA, 1994). The progress of the two eating disorders is highly variable. Some people are cured after a short episode of anorexia nervosa, while some regularly relapse and others suffer chronically from the disorder. The younger anorexia nervosa starts and the shorter the time between onset and presentation for treatment, the better the prognosis. The progress of bulimia nervosa varies: 20% of the treated cases suffer chronically from the eating disorder and 30% relapse now and then, in which cases periods of bulimia nervosa alternate with periods of normal eating behavior (Hsu, 1995). Fifty per cent of bulimics remain completely free of symptoms up to 10 years after cognitive-behavior therapy (Hsu, 1995). Figures for untreated cases are not available.

6.29.8 ETIOLOGY

A number of influential theories concerning the origin and maintenance of eating disorders are discussed in this section. Their empirical validity is also examined.

6.29.8.1 Genetics

The fact that first-degree relatives of anorexia and bulimia nervosa patients suffer from an eating disorder more often than could be expected by chance is often presented as proof that eating disorders are genetically determined. This conclusion is not justified: a higher prevalence of eating disorders among relatives does not establish the genetic nature of the disorders. Such evidence can only be supplied by disentangling the influence of inherited factors from environmentally transmitted factors, for example, through studies of twins or adopted children. By studying concordance rates of eating disorders among monozygotic (MZ) and dizygotic (DZ) twins, Treasure and Holland (1995) found significantly more MZ pairs concordant for anorexia nervosa than DZ pairs. This was not the case for bulimia nervosa, suggesting that genetic factors play a significant role in the development of anorexia nervosa, whereas this may not be true for bulimia nervosa. Thus, both eating disorders tend to run in families, but, for the origin of bulimia nervosa, environmental factors also appear to be important. Genetic influences may be more important in the development of anorexia nervosa. The nature of inherited factors in anorexia and bulimia nervosa needs to be examined more closely.

The cause of obesity is still poorly understood. It is clear that obesity is always caused by a greater energy intake over output, but the cause of the energy imbalance is unknown. Obesity runs in families. Whether this results from genetic factors or family eating patterns or maybe both is yet unknown. Adoption studies point to a heredity factor for body weight: they show that adopted children have weights more closely related to their biological than to their adoptive parents, and monozygotic twins resemble each other more than dizygotic twins, whether they have grown with each other or not (Foster, 1992). According to some experts, the heritability factor is not a gene for excess body weight, but a vulnerability to obesity that is expressed under suitable environmental circumstances (Drewnowski, 1996; Wilson, 1994). There are also some indications that the obese are characterized by inherited metabolic efficiency which enables them to gain or maintain weight easily. For example, obese children tend to have obese parents, and research has shown that the energy expenditure of normal weight offspring of obese parents was only about 80% of that found in normal-weight children of nonobese parents (Foster, 1992). However, there is also evidence for distorted eating patterns in the obese. In general, obese persons are found to eat more than their lean counterparts, and there are indications that childhood obesity is triggered by overeating (Foster, 1992). Moreover, the obese show a preference for fat, and their diet indeed is richer in fat than diets of lean persons. Children's preferences for fat food were not only influenced by their own body fatness, but also by the fatness of their mother, which led Drewnowski (1996) to hypothesize that the genetic vulnerability is an increased appetite for energy-dense fat-containing foods.

6.29.8.2 Childhood Sexual Abuse

The fact that a large proportion of patients with eating disorders and people who have suffered from sexual abuse have low self-esteem and tend to feel shame about their body led some authors to relate childhood sexual abuse to the development of eating disorders (e.g., Andrews, 1997). Some recent large-scale and well-controlled studies show that self-reported childhood sexual abuse is present in about 25–35% of the eating-disordered subjects (Vize & Cooper, 1995; Welch & Fairburn, 1994; 1996a). Welch and Fairburn (1994) found a rate of 26% for sexual abuse in a community sample of subjects with bulimia nervosa. This was significantly more than the rate of sexual abuse in a normal control group (10%) but no different from the rate of sexual abuse in community controls with other psychiatric disorders (24%), like affective disorders or anxiety disorders. Thus, childhood sexual abuse is not directly relevant to most cases of eating disorders, and the data suggest that childhood sexual abuse is a nonspecific risk factor for the development of psychopathology in general and not for the development of eating disorders in particular. However, self-report data have serious limitations: over-reporting, under-reporting, and retrospective distortion may play parts in skewing the data, suggesting that care is needed in their interpretation (Esman, 1994), especially when the influence of neuroticism and depression is not controlled for.

6.29.8.3 Dieting for the Perfect Figure

It has been found in many studies that physically attractive people are considered to be overendowed with favorable qualities. Beautiful people are considered more interesting, stronger, friendlier, more intelligent, and more sexually exciting than less beautiful people; appearance is unrelated to these qualities. As a result of the wonderful qualities attributed to attractive people, their lives become even better than they would otherwise be: they are more likely to receive assistance during conflicts, are more likely to get the job they apply for, and are less likely to be found guilty in a court of law (Feingold, 1992).

It has therefore been suggested that the media systematically present distorted information about the ideal figure and the perfect body, dictating weight standards that are unobtainable for women or only obtainable by considerable effort; being beautiful makes life much more pleasant while being fat is generally not considered attractive; therefore many women use extreme measures to lose weight and consequently develop an eating disorder.

A number of facts appear to support this train of thought. Eating disorders are seen much more often among women and girls than among men and boys. Eating disorders are also reported to be more common in the Western, industrialized countries where women compare themselves to extremely skinny models than in non-Western cultures where this type of figure is not considered to be as ideal. Women with a typically female distribution of fat (broad hips, buttocks, and upper legs) suffer more often from eating disorders than women with a figure that is more male (narrow hips, buttocks, and legs). Lesbians suffer less often from eating disorders than heterosexual women, while homosexual men are more likely to develop eating disorders than heterosexual men (it is assumed that homosexual women are less influenced by the ideal female figure, while outward appearance is more important to homosexual men than to heterosexual men) (Brand, Rothblum, & Solomon, 1992). In our culture, a slim figure is appreciated and striving to lose weight causes a lot of women eating problems, simply because they inherited a figure which—according to our cultural standards—is not ideal.

The idea that our culture overvalues the "perfect figure" and therefore causes eating disorders is, nevertheless, an oversimplification of the facts. The perfect-figure culture is something that is applicable to all women in Western society. If the perfect-figure culture did in fact cause eating disorders, there would be epidemics of these disorders. This is not the case. It would appear more likely that the "perfect-figure culture" stimulates women to diet to lose weight, thereby increasing the risk of an eating disorder. Why?

Binges go hand in hand with the intention to limit the intake of food (Polivy & Herman, 1985); retrospectively, it was found that about 80% of patients with bulimia nervosa tried to lose weight by strict dieting prior to the onset of binge eating (Rossiter, Wilson, & Goldstein, 1989); about 50% of anorexia nervosa patients develop binge eating episodes about 9–18 months after the onset of strict dieting (Polivy & Herman, 1985); and nonclinical dieters report significantly more binge eating than nondieters (Wardle, 1980).

The set-point theory explains the relationship between dieting and binge eating. According to that theory, every organism has a certain amount of fat that has been predetermined for life. The idea is that the amount of fat present in an organism is continuously registered. A homeostatic mechanism is responsible for keeping the amount of fat balanced. If, in comparison with the predetermined amount of fat, a shortage of fat is identified, processes are activated that bring the quantity of fat up to the set-point level. If too much fat is identified, processes are activated that bring the quantity of fat back down to the set-point level (Keesey, 1995). According to this theory, people with a low set-point weight in our "perfect-figure" culture are just plain lucky. The people with bad luck—a high set-point weight—are condemned to being "overweight" for their entire lives. When these people—for instance, when influenced by the concept of a "perfect-figure'—start to diet and lose weight, their weight drops below their own set-point weight. Then the body responds with mechanisms designed to regain weight: bingeing and a slower metabolism can be triggered. According to the set-point theory, patients with bingeing urges have dropped below their set-point weight.

The theory is elegant, but there are a few anomalies. In the first place, patients with anorexia nervosa are the perfect example of malnourished dieters. Nevertheless, only half of these patients report suffering from bingeing urges. Second, there are dieters who never overeat; they are called "successful" dieters. A German study indicated that about one-half of every one thousand women diet without any problems (Westenhoefer, Pudel & Maus, 1990); in England and the USA, documented studies indicate that "only" 8–9% of dieting women and girls suffer from bingeing urges or develop an eating disorder (Agras, 1990; Patton, Johnson-Sabine, Wood, Mahn, & Wakeling, 1990). The set-point mechanism is therefore not nearly as dominant as suggested. Apparently, many people can diet and lose weight to below the set-point level without any problems. Furthermore, it does not appear likely that dieters with bingeing urges are below their set-point weight: a number of studies have indicated that dieters have a higher BMI and eat the same amount or even more kilocalories (and retain that energy) than nondieters (see, e.g., Jansen, 1996). Of course, one might put forward the untestable and circular argument that the bingeing dieters have a higher set-point than the normal eaters; advocates of the set-point theory still have the formidable task of discovering how a person's set-point weight can be determined.

Studies of the relationship between dieting and bingeing are usually retrospective or correlative and they thus do not render any certainties with reference to causality. It therefore cannot be proven that dieting is not a consequence of a tendency to regularly overeat or that there is not a third factor involved; an underlying mechanism that explains both the dieting and the bingeing.

Briefly summarized: the "perfect-figure" culture prescribes a female figure that most women will never be able to achieve. This is why many women diet. In about 9% of the cases, dieting results in an eating disorder. This makes it improbable that the "perfect-figure" culture and dieting are the key determinants of eating disorders; if this were the case, a considerably larger number of women would suffer from eating disorders.

6.29.8.4 Serotonin Depletion

First-degree relatives of patients with an eating disorder not only have eating disorders more often, but mood disorders are also more common than would be expected. Nearly half of patients with eating disorders are also diagnosed as suffering from depression (Laessle, 1990). The association of the two disorders within families and individuals led to the hypothesis that eating disorders are in fact a type of biological mood disorder. This will be discussed further below.

Various studies identified a relationship between eating carbohydrates and mood. Eating a proportionately large amount of carbohydrates results in a better mood (Fichter & Pirke, 1995; Jansen, van den Hout, & Griez, 1989; Wurtman & Wurtman, 1984). This mood improvement is assumed to be mediated by the neurotransmitter serotonin: by eating relatively large amounts of carbohydrates and little protein, serotonin production in the brain increases. The increased serotonergic activity results in mood improvement. A shortage of carbohydrates, by contrast, results in a decrease in the amount of serotonin in the brain. In turn, this results in a lowering of mood.

The amount of serotonin in the brain determines not only mood state but also the craving for carbohydrates. As the amount of active serotonin decreases, an increased craving for carbohydrates is observed. It can therefore be hypothesized that the production of serotonin in the brain is in part dependent on the amount of carbohydrates consumed. Eating a larger amount of carbohydrates results in more serotonin being produced in the brain. As a result, mood improves and the craving for carbohydrates diminishes. If, however, too few carbohydrates are consumed, less serotonin is produced, mood declines, and the craving for carbohydrates increases.

Animal experiments have shown that deprivation of carbohydrates seriously disrupts the carbohydrate regulation as described above. After rats were deprived of carbohydrates for a short period, cerebral sertonin level decreased as predicted. It was striking that serotonin did not reach the normal level again until after an excessive amount of carbohydrates was consumed (Wurtman & Wurtman, 1984).

These and other findings resulted in the following hypothesis concerning the origin of binges among patients with bulimia nervosa (see also Figure 1): for a variety of reasons, for example the desire to fulfill the cultural ideal of the female figure, some women subject themselves to a rigid eating regime. If this eating regime results in too few carbohydrates being consumed, successfully following the strict regime will result in disruption of the serotonergic system on the cerebral level. By depriving themselves of carbohydrates for longer periods of time, consumption of carbohydrates later will not immediately result in the required increase in serotonin. This increase will not be achieved unless excessive amounts of carbohydrates are consumed during a binge. A binge could, therefore, be considered a type of "self-medication."

Indeed it was found, analogous to the finding from the animal experiment of the Wurtmans (Wurtman & Wurtman, 1984), that the consumption of a standard quantity of carbohydrates by patients with bulimia nervosa did not result in the expected increase in the tryptophan ratio. Tryptophan is a large neutral amino acid, and its ratio to other large amino acids determines the amount of tryptophan entering the brain. A larger ratio means more tryptophan entering the brain, resulting in increased serotonin levels in the brain. What's more, eating carbohydrates did not stop until the tryptophan ratio in the blood plasma reached a normal level (see Pirke, 1995). Others found a lower concentration of the serotonin metabolite 5-hydroxyindoleacetic acid (5-HIAA) in the cerebrospinal fluid (CSF) of severe bulimic patients (Jimerson, Lesem, Kaye, & Brewerton, 1992). The quantity of 5-HIAA in the CSF reflects the amount of serotonin active in the brain; the data thus point to a low turnover of 5-hydroxytryptophan (5-HT) in subjects with severe bulimic symptoms. However, a study with weight-restored anorexics indicates that abnormalities in central serotonin function *follow* from weight loss and malnutrition and are not related primarily to the disorder itself (Odwyer, Lucey, & Russell, 1996).

Caution is also advised when the findings are presented as evidence supporting the theory that binges serve to increase the quantity of serotonin in the brain. In the above-mentioned studies, it is assumed that the food patients consume during binges is rich in carbohydrates, whereas this is not the case. It has repeatedly been found that patients with bulimia nervosa eat as much carbohydrates during the binges as they do in between binges (Jansen, van den Hout, & Griez, 1989; Walsh, 1993). Moreover, in both periods they consume the same amount of carbohydrates as a normal control group. However, it should also be noted that a recent study among healthy female subjects shows that even moderate dieting on a diet with a normal macronutrient composition (1000 kcal, 31% protein, 44% carbohydrate, 25% fat) causes increased sensitivity of 5-HT receptors in the brain (Cowen, Clifford, Walsh, Williams, & Fairburn, 1996), suggesting that the overall decrease in the neurotransmission of serotonin during dieting does not require a relative carbohydrate depletion.

The serotonin hypothesis concerning the origin of bingeing is important, but it must as

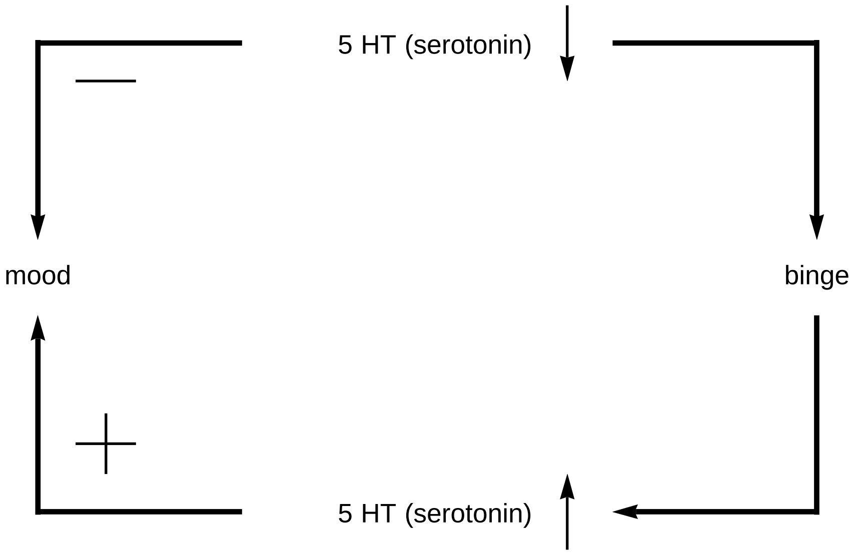

Figure 1 The serotonin model of binge eating.

yet be proven that the bingeing truly serves to increase the serotonin level in the brain. Findings suggesting that other psychiatric disorders may be characterized by low serotonin levels in the brain also generate the question of why these patients do not suffer from bingeing as a type of self-medication.

6.29.8.5 The Learned Nature of Binge Eating

Recent studies in the area of addiction suggest that the urge to consume a substance is a classically conditioned response (Drummond, Tiffany, Glautier, & Remington, 1995; Siegel, 1983). Addicts normally use a substance in the presence of certain stimuli, for example, their favorite pub, drinking or shooting buddies, the bottle, the syringe, the spoon and the silver foil. In a long series of experiments, it was shown that stimuli the nature of their exclusive presence eventually predict the use of a substance resulting in the addict having preparatory physiological reactions. Even before any substance is swallowed, injected, or inhaled, changes occur in the body. These anticipative physical changes are said to be experienced as a virtually uncontrollable urge to consume the substance, and striking is the fact that the occurrence of physical reactions after successful withdrawal accurately predicts a relapse (Drummond et al., 1995).

Patients with bulimia nervosa show some striking similarities with addicts. Both are preoccupied with the substance that they use excessively, both experience an urge to consume the substance, and both are characterized by loss of control during consumption. The analogy between bulimia nervosa and addiction may teach us something about the origin of binges. Excessive consumption of food during a binge is often related to a configuration of specific stimuli. Think in this respect of seeing, smelling, and tasting rich food, emotional confusion (a somber mood, fear, boredom, or loneliness), certain disinhibiting thoughts, and the time of the day. When repeatedly and exclusively combined with the bingeing, these stimuli become excellent predictors for a binge. Impressive data from experiments with animals indicate that eating behavior can be induced by stimuli that are repeatedly associated with the intake of food (Wardle, 1990; Woods, 1991).

According to this theory, the stimuli that predict the binge continually induce preparatory physiological responses in the eater (see Jansen, 1994, in press). The model predicts that the physiological responses are experienced as a virtually irresistible urge to consume food and therefore increase the chance of a binge. It remains, however, to be seen whether this model for the origin of binges will be confirmed by empirical studies.

6.29.8.6 Dysfunctional Cognitions

Excessive worry about the figure, appearance, and weight, as already indicated, is characteristic for patients with an eating disorder. The extreme worries about appearance and weight probably represent the core of the eating disorders. Thinness is idealized, and the feeling of being fat fuels the drive to lose weight. Eating-disordered subjects continually evaluate themselves in terms of outward appearance and weight. They tend to judge self-worth almost exclusively in terms of body shape and weight and, therefore, often show long-standing negative self-evaluation (Fairburn, 1997a). Being fat is associated with being worthless, unlikable, disgusting, and so on; being skinny means being attractive, successful, intelligent, happy, and so on (Vitousek & Hollon, 1990). This method of self-evaluation is remarkable because in all cases of anorexia nervosa and in most cases of bulimia nervosa, patients are not objectively overweight at all. Of course, the evaluation of a person is dependent on much more than his outward appearance alone. It is more usual to judge self-worth on the basis of performance in a variety of domains, like work, sports, friendships, relationships, and so on (Fairburn, 1997a). According to the cognitive view on the maintenance of anorexia nervosa and bulimia nervosa (Fairburn, 1997a), the overvalued ideas about the significance of body shape and weight are at the heart of the eating disorders. Self-evaluation is unduly influenced by how the subject perceives her body shape and weight. The preoccupation with food, eating, dieting, and weight control methods are supposed to be secondary to this overconcern.

The model states that binge eating follows from a particular type of dieting: intense and rigid dieting. It is postulated that eating-disordered subjects are characterized by perfectionism and black-and-white (all-or-nothing) thinking, which also becomes manifest in their way of dieting. The idea is that breaking one of their diet rules, even after a minor dietary transgression, results in the view that their diet is "totally broken," which ends up in a total loss of control and thus a binge (Fairburn, 1997a; Herman & Polivy, 1984). This is said to be accompanied by typical, disinhibiting automatic thoughts along the lines of "my day is already ruined, I might as well continue to eat" and "I'm trying not to eat. If I eat anyway, I might just as well continue to eat," reflecting their perfectionism (Garner, 1986). Thus, if eating-disordered subjects feel as if they have eaten too much, they consider their diet a failure for that day and eat too much the rest of the day. Disinhibiting, all-or-nothing thoughts are therefore supposed to be responsible for the onset of the binge.

The cognitive view on the maintenance of anorexia nervosa and bulimia nervosa is at the moment the most widely accepted model of eating disorders. Strikingly, considering the popularity and plausibility of the cognitive view on the maintenance of eating disorders, there is little direct empirical evidence for most of the hypotheses flowing from the model. Among other things, the model predicts that (i) subjects with eating disorders are characterized by negative self-evaluation or low levels of self-esteem, (ii) low self-esteem is closely related to the extreme concerns about body shape and weight, and (iii) long-term treatment success is related to a reduction of the overconcern about body shape and weight. Furthermore, it is predicted that (iv) a binge is caused by disinhibitive thoughts such as "my day is ruined now, I might as well continue to eat." Let us review the evidence.

The first hypothesis (subjects with eating disorders are characterized by low levels of self-esteem) has indeed been supported in diverse questionnaire and interview studies (see, e.g., Mizes & Christiano, 1995). Also a prospective study on the development of eating disorders showed that 11- and 12-year-old dieting girls with low self-esteem were at significantly greater risk of developing eating disorders at the age of 15–16 than dieting girls with normal and high self-esteem (Button, Sonagu-Barke, Davies, & Thompson, 1996). Considering hypothesis (ii) (the low self-esteem is closely related to the overconcern about body shape and weight), questionnaire studies indicate that subjects with eating disorders perceive weight and eating as the basis of approval from others, whereas control subjects do not (Mizes & Christiano, 1995). However, self-report data on the assumed link between mental concepts have serious limitations. There is little experimental research investigating the direct link between low self-esteem and overconcern with body shape and weight. In an analog study with restrained eaters, Eldredge, Wilson, and Whaley (1990) hypothesized that highly restrained eaters experiencing failure in an achievement task would react more negatively towards their own bodies than restrained eaters who experienced success. They found that the highly restrained eaters experiencing failure did not feel worse about their bodies than the highly restrained eaters experiencing success. Others, however, did find evidence for the link between self-evaluation and body concern. It was found that activation of thoughts about eating, weight, and shape led to an increase in negative self-statements (Cooper, Clark, & Fairburn, 1993) and, the other way around, presentation of negative events involving the self led to weight and shape explanations in eating-disordered subjects (Cooper, 1997). Furthermore, a recent study from our lab (Meijboom, Jansen, Kampman, & Schouten, in press) showed that priming low self-esteem in highly restrained subjects facilitated the accessibility of subliminally presented body shape and weight words. The latter three findings support the idea that concern with body shape and weight is linked to self-esteem in eating-disordered subjects. However, much more experimental research is needed before any definite conclusions about the link, and its causal relationship, between self-esteem and overconcern with body shape and weight can be drawn. It might, for example, be an interesting enterprise to test whether priming body shape and weight words will lead to higher accessibility of (low) self-esteem stimuli. Furthermore, it is necessary to test experimentally whether reducing the overconcern about body shape and weight will increase self-esteem, whereas an activation of the overconcern about body shape and weight in normal dieters will decrease their self-esteem.

There is some evidence for hypothesis (iii) (long-term treatment success is related to a reduction of the overconcern about body shape and weight). Fairburn, Peveler, Jones, Hope, and Doll (1993) showed that the severity of concerns about shape and weight at the end of treatment were directly related to the likelihood of relapse: 9% of the subjects with the least concerns relapsed vs. 19% of those with a moderate level of concern and 75% of those with the greatest level of concern.

Considering hypothesis (iv), clinical notes indeed indicate that a binge is caused by disinhibiting thoughts such as "my day is ruined now, I might as well continue to eat." However, such clinical impressions have received little empirical support. In our laboratory, the prediction was studied in bogus taste experiments, but it could not be demonstrated that disinhibiting thoughts play a significant role in the onset of overeating (Jansen, Merckelbach, Oosterlaan, Tuiten, & van den Hout, 1988; Jansen & van den Hout, 1991). There is, however, some circumstantial evidence consistent with the idea that the activation of disinhibiting thoughts in normal dieters can elicit bingeing. In a laboratory study with high and low restrained eaters, Spencer and Fremouw (1979) manipulated their subjects' beliefs concerning the caloric content of a preload and demonstrated that the restrained eaters overate after the preload when they only *thought* they had overeaten during the preload. It was not clear, however, whether the overeaters were characterized by more disinhibitive thoughts. Whether a straightforward activation of disinhibitive thoughts elicits a binge has never been tested. A final prediction following from the hypothesis is that a reduction of disinhibitive thoughts prevents a binge. There is an experimental finding which counters this prediction. Dieters were led into temptation to eat by smelling very palatable food. Because they merely smelled the food (their diet was not broken), it was assumed that the manipulation prevented the occurrence of disinhibitive thoughts. Contrary to what the cognitive model predicts, it was found that dieters who were merely exposed to, and intensively smelled the food, but whose diets were not actually broken or threatened, overate (Jansen & van den Hout, 1991). Contrary to hypothesis (iv), the absence of disinhibiting thoughts did not prevent dieters overeating.

All in all, there is evidence that subjects with eating disorders are characterized by low self-esteem, and there is some support for the hypothesis that low self-esteem is linked to extreme concerns about body shape and weight. However, much more experimental research on the precise link between self-esteem and body shape evaluation is needed before firm conclusions can be drawn. Furthermore, there is hardly any empirical evidence for the idea that disinhibiting thoughts cause binge eating.

There have been more studies on the cognitions of subjects with eating disorders than the studies cited above (for an overview see Vitousek & Hollon, 1990). Remarkably, however, the goal of most of those studies was to verify that eating-disordered subjects show cognitive distortions concerning food, eating behavior, body shape, and weight. For example, in the modified Stroop test, it has repeatedly been found that subjects with eating disorders selectively process stimuli related to food, eating, body shape, and weight. It has also been found that they show a memory bias for fatness words (Sebastian, Williamson, & Blouin, 1996). This kind of research may be based on circular reasoning; subjects are selected for their abnormal cognitive processes concerning body shape, weight, food, and eating, and it is concluded that they are characterized by an abnormal processing of body shape, weight, food, and eating information.

6.29.9 TREATMENT

6.29.9.1 Anorexia Nervosa

Traditionally, anorexia nervosa is treated on an inpatient basis with a program to increase weight, loosely based on operant conditioning. The patients are admitted to the clinic and it is agreed that they will gain a predetermined amount of weight each week. They are prescribed lots of bed rest (so that they will burn less energy), and a number of pleasant pastimes (receiving visitors and watching television, for example) are forbidden. They can gradually earn back these pleasant pastimes by gaining the desired amount of weight. The effectiveness of treatments of this type has never been established; the clinical impression is that patients gain weight during the treatment but quickly relapse after leaving the clinic. Perhaps this is caused by the overemphasis on the sole goal of gaining weight. Anorexia nervosa consists of more than simply being considerably underweight. Its core psychopathology consists of cognitive distortions concerning body shape and weight, which probably do not change during a treatment that is solely oriented towards gaining weight.

For this reason, today many (but not all) inpatient treatment programs are much more versatile. In addition to focusing on gaining weight, the patients learn to accept their body, for example, by video confrontation, and talk about emotions. They are often given nutritional counseling and family therapy. Focus is also placed on increasing self-esteem and improving social skills (Fichter, 1995). The effectiveness of treatments of this type, however, has not been empirically confirmed.

As long as the patient's physical condition makes it possible, outpatient treatment is preferred above intramural treatment. Recently, it has been argued that patients with anorexia nervosa would, like subjects with bulimia nervosa, benefit from outpatient cognitive behavior therapy, a treatment which is focused on changing maladaptive thoughts, feelings, and behavior (Fairburn, Marcus & Wilson, 1993; Vitousek, 1995). Again, however, the effectiveness of cognitive-behavior therapy in the treatment of anorexia nervosa also remains to be established.

Antidepressants are the most commonly prescribed medications for anorexia nervosa; however, in the long run they do not appear to be effective. The reason for this ineffectiveness may be that antidepressants do not affect the morbid attitudes, beliefs, and associated behaviors which constitute the central features of the illness (Wakeling, 1995).

6.29.9.2 Bulimia Nervosa

The effectiveness of treatments for bulimia nervosa have been studied in numerous controlled studies. Cognitive-behavior therapy, based on the cognitive model of eating disorders (see Section 6.29.8.6), has been shown to be the most effective therapy. For the present, it must be the treatment of choice for bulimia nervosa. A detailed manual was developed by Chris Fairburn (see, e.g., Fairburn, 1997a) focusing on disrupted eating behavior, the cognitive distortions concerning body shape and weight, and the long-standing negative self-evaluations of subjects with eating disorders. In about twenty 50-minute sessions over 20 weeks, the patient learns to eat and think more normally. After the therapy, a reduction in binge and purge frequency of about 80% has been observed and about 55% of the patients stopped bingeing altogether (Fairburn, 1995; Wilson, 1996a). The patients also show more normal attitudes towards their body shape and weight. In one study, patients were followed for a period of six years. The long-term results are also favorable: after six years there is virtually no relapse (Fairburn et al., 1995).

Cognitive-behavior therapy for eating disorders in fact is a multicomponent treatment package including, among others, self-control techniques, self-monitoring, education, diet management, cognitive restructuring, problem-solving training, interpersonal training, and relapse prevention. A relevant question of course is: which of these treatment components are necessary to effect positive response? For researchers on eating disorders, the time is here, now to identify the effective parts of cognitive-behavior therapy as well as to identify the mechanisms of action, that is, the mechanisms by which the effective treatment parts achieve changes.

There has been one treatment that may be as effective as cognitive-behavior therapy: interpersonal psychotherapy (Fairburn, 1997b). During this short-term focal psychotherapy, little attention is paid to the eating disorder symptoms, instead the treatment focuses on the identification and modification of current interpersonal problems. In the long run, interpersonal therapy proved to be as effective as cognitive-behavior therapy. The idea is that the treatment is successful because it improves interpersonal functioning and thereby self-worth, meaning that self-evaluation is less dependent on body shape and weight, which makes dieting less necessary and so the eating disorder gradually erodes (Fairburn, 1997b). However, the assumed mechanism by which the treatment brings about change still has to be studied, and more controlled trials are needed before any firm conclusions about the robustness of the effects can be drawn.

A rather new and promising approach is the cognitive-behavioral body image therapy (Rosen, Reiter, & Orosan, 1995). Considering the fact that the severity of concerns about shape and weight at the end of treatment are related to the likelihood of relapse (Fairburn et al., 1993), it is argued that maintenance of recovery without significant body image changes will be a continuous struggle (Rosen, 1995). The body image therapy is divided into two parts: (i) cognitive restructuring, and (ii) exposure. During the cognitive part, subjects keep a body image diary in which they record situations that provoke concerns about appearance, body image beliefs, and the effect of both on mood and behavior. Irrational, dysfunctional thoughts about the body and damaging beliefs concerning self-worth are challenged and restructured. Furthermore, the patient gradually exposes herself to a hierarchy of distressing aspects of appearance, and during the exposure the goal is to change negative self-talk in nonjudgmental self-descriptions (Rosen, 1995). A further goal is to reduce or prevent body-checking behavior. Body image therapy has been shown to be effective (Rosen, 1995), but it is surprisingly infrequently applied.

Medication for bulimia nervosa has mainly focused on the use of antidepressant drugs. Considering binge frequency, the drugs establish a reduction in binge frequency of about 60% with an abstinence rate of 22% (percentage of subjects that stopped binge eating completely). However, the effects are only maintained when subjects remain on the drugs. When medication is stopped, recurrence of binge eating is frequently seen, so the maintenance of change is poor (Fairburn, Agras, & Wilson, 1992). Furthermore, it seems that the drugs do not influence attempts to diet, cognitive distortions concerning body shape and weight, nor the longstanding negative self-evaluations of subjects with bulimia nervosa. Moreover, comparative trials show that cognitive-behavior therapy is superior to medication, and a combination of medication and cognitive-behavior therapy is not convincingly more effective than cognitive-behavior therapy alone (Wilson, 1996a).

6.29.9.3 Obesity and Binge Eating Disorder

Obesity has, so far, proved to be refractory to most treatment methods (Wilson, 1994). The behavioral weight loss programs, which are focused on dieting and self-control strategies, produce short-term weight loss but are ineffective in the long run. Recently, Wilson (1996b) has argued that the narrow focus on weight loss only may no longer prove acceptable, and he proposes a change of treatment goals. He argues that treatment of the obese should focus on the enhancement of self-acceptance, predicting that increased self-acceptance of the obese might lead to more lasting changes in eating and exercise behavior.

Treatment of the binge eating disorder has been studied only minimally. Cognitive-behavior therapy appears to be partly effective for patients suffering from a binge eating disorder. Loss of weight is not achieved, but no more weight is gained, binge frequency decreases, mood improves, dissatisfaction with the body decreases, self-esteem increases, interpersonal skills improve, and marriages become more fulfilling (Smith, Marcus, & Kaye, 1992; Wilson, 1994). Binge eaters may also benefit from cognitive-behavioral self-help treatment manuals (Cooper, 1995b; Fairburn, 1995a). In a couple of studies they have proved to be an effective treatment method to combat binge eating. They can be used independently by sufferers on their own, or with guidance from a nonspecialist therapist (Fairburn, 1997a; Fairburn & Carter, 1997).

Finally, there are preliminary data pointing to cue exposure as a successful method for the treatment of binge eating (Jansen, in press; Jansen, Broekmate, & Heymans, 1992). *In vivo* exposure with response prevention is a therapy which has been successfully applied for many years in the treatment of anxiety disorders. Subjects are repeatedly exposed to the stimuli they are afraid of, and during the exposure their anxiety gradually extinguishes. From the learning model on binge eating (see Section 6.29.8.5), it follows that the urge to binge will also be extinguished during exposure to the cues which predict a binge. By repeatedly exposing patients for long periods to the stimuli that predict a binge with response prevention (eating is not allowed), the stimuli eventually lose their predictiveness. Once the stimuli are no longer predictors for a binge, there is no point in preparing the body for consumption. The preparatory physical reactions no longer occur and the urge to binge has disappeared. The data from five, largely uncontrolled and small-scale studies are promising (see Jansen, 1998).

6.29.9.4 Predictors of Treatment Outcome

An important clinical issue is whether it can be predicted who will respond to treatment. However, reliable predictors of treatment outcome have not yet been identified. In some but not all studies, weight, premorbid and paternal obesity, self-esteem, personality disorders, and binge eating have been identified as predictor variables (Wilson, 1996a). These studies show that the worst outcome is found in subjects with lower weight, higher premorbid and paternal weight, lower self-esteem at the start of the treatment, comorbid borderline personality disorders, an early onset of binge eating, and more severe binge eating. However, the data on these negative prognostic factors are not conclusive; more and better research is needed.

6.29.10 IN CONCLUSION

The eating disorders anorexia nervosa and bulimia nervosa are characterized by: (i) abnormal eating behavior, (ii) the use of unusual methods to control weight, and (iii) irrational and dysfunctional ideas about the patient's own body and weight, leading to negative self-evaluation. Because the patients feel fat and are afraid to gain weight, they continually strive to lose weight. In the case of anorexia nervosa, this results in a decrease in weight of at least 15% compared with the person's normal weight. People with bulimia nervosa usually have a normal weight, often because of binges that regularly disrupt attempts to lose weight. During binges, considerable quantities of food are consumed, and after the binge the patient vomits, uses laxatives, or starts extreme dieting. Weight control measures of this type are partly effective: the patient gains virtually no weight. However, they do not prevent some of the food from being absorbed by the body, as a result of which the patient does not lose an extreme amount of weight.

Obesity is not an eating disorder as presently defined, but a subgroup of the obese is found to fulfill the criteria of the binge eating disorder. Like subjects with bulimia nervosa, they are characterized by objective binges during which they eat a large amount of food and lose control over intake. However, contrary to subjects with bulimia nervosa, they do not engage in the characteristic compensatory purging behaviors of bulima nervosa.

It is (as yet) unknown how the eating disorders originate. A number of influential theories have been discussed here. There is evidence for a genetic component in anorexia nervosa and obesity, but not in bulimia nervosa. However, what exactly is inherited is still unknown.

The "perfect-figure" culture appears to be more of an explanation for the common dieting among women. This does not explain the origin of eating disorders because "only" a fraction of the dieters (9%) develop an eating disorder. Dieting is related to binge eating; however, their cause and effect relationship is unclear. Empirical data show that dieting is not necessarily required for binge eating to occur, and it is also clear that dieting is not sufficient for eating disorders, including binge eating, to occur.

The serotonin hypothesis offers an explanation for the isolated act of bingeing and for the coherence between eating disorders and depression. Too little serotonin in the brain is said to result in a virtually uncontrollable craving for carbohydrates and low mood. Eating an excessive amount of carbohydrate is said to restore the serotonin level in the brain and, as a type of "self-medication," to improve mood. Empirical findings suggest that patients with bulimia nervosa do in fact have a brain serotonin deficit. However, other psychological disorders may also be accompanied by a lack of serotonin. Moreover, it has been repeatedly shown that the food patients with bulimia nervosa consume during a binge does not contain more carbohydrates than either the food they consume between binges or the food consumed by a normal control group, findings which do not support the self-medication hypothesis.

Various experimental findings suggest that overeating is a learned behavior. Predictors or cues of excessive eating trigger a craving for food for as long as they continue to predict a binge, meaning as long as they are systematically followed by excessive food consumption. They will always trigger a physical preparatory response which is experienced as craving or an urge to eat. The model is confirmed *post hoc* by a variety of findings, but there is no direct empirical evidence that supports this theory on the origin of binges.

Irrational cognitions concerning one's body shape, weight, eating, and food are characteristic for patients with anorexia and bulimia nervosa. The precise role they play in the etiology and maintenance of the eating disorders is, however, as yet unknown. In terms of the proposed cognitive origin of binge eating, it appears at this time that binges do not automatically follow irrational disinhibiting thoughts, for example, "my day is ruined anyway, I may as well continue to eat."

Despite the fact that little is known about the exact causes of eating disorders, bulimia nervosa can be effectively treated. Cognitive behavioral therapies are the most successful methods. They have a behavioral and a cognitive component; the focus is on both learning to eat normally and changing irrational thinking (about body shape and weight as well as food intake). New cognitive-behavioral approaches such as the exposure to binge provoking cues as well as the exposure to appearance, are promising and need to be studied further. The treatment of obesity is disappointing; most treatment methods are ineffective in the long run. Obese binge eaters may benefit from cognitive-behavior therapy; although weight loss is not achieved, weight gain is prevented, binge frequency decreases, and well-being improves. Finally, less is known about the effectiveness of treatments for anorexia nervosa: unfortunately, to date, virtually no properly controlled studies have been performed.

6.29.11 REFERENCES

Agras, W. S. (1990). Is restraint the culprit? Commentary. *Appetite, 14,* 111–112.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Andrews, B. (1997). Bodily shame in relation to abuse in childhood and bulimia: a preliminary investigation. *British Journal of Clinical Psychology, 36,* 41–49.

Brand, P., Rothblum, E., & Solomon, L. (1992). A comparison of lesbians, gay men, and heterosexuals on weight and restrained eating. *International Journal of Eating Disorders, 3,* 253–259.

Brody, M. L., Walsh, B. T., & Devlin, M. J. (1994). Binge eating disorder: Reliability and validity of a new diagnostic category. *Journal of Consulting and Clinical Psychology, 62*(2), 381–386.

Brownell, K. D. (1995). Definition and classification of obesity. In K. D. Brownell & C. G. Fairburn (Eds.), *Eating disorders and obesity: A comprehensive handbook* (pp. 386–390). New York: Guilford Press.

Button, E. J., Sonuga-Barke, E. J. S., Davies, J., & Thompson, M. (1996). A prospective study of self-esteem in the prediction of eating problems in adolescent schoolgirls: Questionnaire findings. *British Journal of Clinical Psychology, 35,* 193–203.

Castonguay, L. G., Eldredge, K. L., & Agras, W. S. (1995). Binge eating disorder: Current state and future directions. *Clinical Psychology Review, 15*(8), 865–890.

Cooper, M. (1997). Bias in interpretation of ambiguous scenarios in eating disorders. *Behaviour Research and Therapy, 35*(7), 619–626.

Cooper, M., Clark, D. M., & Fairburn, C. G. (1993). An experimental study of the relationship between thoughts and eating behaviour in Bulimia Nervosa. *Behaviour Research and Therapy, 31*(8), 749–757.

Cooper, P. J. (1995a). Eating disorders and their relationship to mood and anxiety disorders. In K. D. Brownell & C. G. Fairburn (Eds.), *Eating disorders and obesity: A comprehensive handbook* (pp. 159–164). London: Guilford Press.

Cooper, P. J. (1995b). *Bulimia nervosa and binge-eating: A guide to recovery*. London: Robinson.

Cowen, P. J., Clifford, E. M., Walsh, A. E. S., Williams, C., & Fairburn, C. G. (1996). Moderate dieting causes 5-HT2c receptor supersensitivity. *Psychological Medicine, 26,* 1155–1159.

Drewnowski, A. (1996). The behavioral phenotype in human obesity. In: E. Capaldi (Ed.), *Why we eat what we eat: The psychology of eating* (pp. 291–308). Washington, DC: American Psychological Association.

Dritschel, B., Williams, K., & Cooper, P. (1991). Cognitive distortions amongst women experiencing bulimic episodes. *International Journal of Eating Disorders, 5,* 547–555.

Drummond, D. C., Tiffany, S. T., Glautier, S., & Remington, B. (1995). *Addictive behaviour: Cue exposure theory and practice*. Chichester, UK: Wiley.

Eldredge, K., Wilson, G. T., & Whaley, A. (1990). Failure, self-evaluation and feeling fat in women. *International Journal of Eating Disorders, 9,* 37–50.

Esman, A. H. (1994). "Sexual abuse", pathogenesis, and enlightened skepticism. *American Journal of Psychiatry, 151,* 1101–1103.

Fairburn, C. G. (1995a). *Overcoming binge eating*. New York: Guilford Press.

Fairburn, C. G. (1995b). Short-term psychological treatments for bulimia nervosa. In K. D. Brownell & C. G. Fairburn (Eds.), *Eating disorders and obesity: A comprehensive handbook* (pp. 344–348). London: Guilford Press.

Fairburn, C. G. (1997a). Eating disorders. In D. M. Clark & C. G. Fairburn (Eds.), *Science and practice of cognitive behaviour therapy* (pp. 209–241). Oxford: Oxford University Press.

Fairburn, C. G. (1997b). Interpersonal psychotherapy for Bulimia Nervosa. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 278–294). New York: Guilford Press.

Fairburn, C. G., Agras, W. S., & Wilson, G. T. (1992). The research on the treatment of Bulimia Nervosa: practical and theoretical implications. In G. H. Anderson & S. H. Kennedy (Eds.), *The biology of feast and famine: Relevance to eating disorders* (pp. 318–340). New York: Academic Press.

Fairburn, C. G., & Beglin, S. J. (1990). Studies of the epidemiology of Bulimia Nervosa. *The American Journal of Psychiatry, 4,* 401–408.

Fairburn, C. G., & Carter, J. C. (1997). Self-help and guided self-help for binge-eating problems. In D. M. Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders* (pp. 494–499). New York: Guilford Press.

Fairburn, C. G., & Cooper, P. J. (1989). Eating disorders. In K. Hawton, P. Salkovskis, J. Kirk, & D. M. Clark (Eds.), *Cognitive behaviour therapy for psychiatric problems*. Oxford: Oxford University Press.

Fairburn, C. G., Marcus, M. D., & Wilson, G. T. (1993). Cognitive-behavioral therapy for binge eating and Bulimia Nervosa: A comprehensive treatment manual. In C. G. Fairburn & G. T. Wilson (Eds.), *Binge eating: Nature, assessment, and treatment* (pp. 361–404). New York: Guilford Press.

Fairburn, C. G., Norman, P. A., Welch, S. L., O'Connor, M. E., Doll, H. A., & Peveler, R. C. (1995). A prospective study of outcome in bulimia nervosa and the long-term effects of three psychological treatments. *Archives of General Psychiatry, 52,* 304–312.

Fairburn, C. G., Peveler, R. C., Jones, R., Hope, R. A., & Doll, H. A. (1993). Predictors of twelve-month outcome in bulimia nervosa and the influence of attitudes to shape and weight. *Journal of Consulting and Clinical Psychology, 61,* 696–698.

Fairburn, C. G., & Walsh, B. T. (1995). Atypical eating disorders. In K. D. Brownell & C. G. Fairburn (Eds.), *Eating disorders and obesity: A comprehensive handbook* (pp. 135–140). London: Guilford Press.

Fairburn, C. G., & Wilson, G. T. (1993). Binge eating: definition and classification. In C. G. Fairburn & G. T. Wilson (Eds.), *Binge eating: Nature, assessment, and treatment* (pp. 3–14). New York: Guilford Press.

Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin, 2,* 304–341.

Fichter, M. (1995). Inpatient treatment of anorexia nervosa. In K. D. Brownell & C. G. Fairburn (Eds.), *Eating disorders and obesity: A comprehensive handbook* (pp. 336–343). London: Guilford Press.

Fichter, M. M., & Pirke, K. M. (1995). Starvation models and eating disorders. In G. Szmukler, C. Dare, & J. Treasure (Eds.), *Handbook of eating disorders: Theory, treatment, and research* (pp. 83–108). Chichester, UK: Wiley.

Fichter, M. M., Quadflieg, N., & Brandl, B. (1993). Recurrent overeating: An empirical comparison of binge eating disorder, bulimia nervosa, and obesity. *International Journal of Eating Disorders, 14,* 1–16.

Fombonne, E. (1996). Is Bulimia Nervosa increasing in frequency. *International Journal of Eating Disorders, 19,* 287–296.

Foster, D. W. (1992). Eating disorders: Obesity, Anorexia Nervosa, and Bulimia Nervosa. In J. D. Wilson & D. W. Foster (Eds.), *Williams textbook of endocrinology* (pp. 1335–1365). Philadelphia: Saunders.

Gard, M. C. E., & Freeman, C. P. (1996). The dismantling of a myth: A review of eating disorders and socio-economic status. *International Journal of Eating Disorders, 20,* 1–12.

Garfinkel, P. E. (1995). Classification and diagnosis of eating disorders. In K. D. Brownell & C. G. Fairburn (Eds.), *Eating disorders and obesity: A comprehensive* *handbook* (pp. 125–134). London: Guilford Press.

Garner, D. M. (1986). Cognitive therapy for Anorexia Nervosa. In K. D. Brownell & J. P. Foreyt (Eds.), *Handbook of eating disorders: physiology, psychology and treatment of obesity, anorexia and bulimia* (pp. 301–327). New York: Basic Books.

Garner, D. M. (1993). Binge eating in anorexia nervosa. In C. G. Fairburn & G. T. Wilson (Eds.), *Binge eating: Nature, assessment, and treatment* (pp. 50–76). New York: Guilford Press.

Herman, C. P., & Polivy, J. (1984). A boundary model for the regulation of eating. In A. J. Stunkard & E. Steller (Eds.), *Eating and its disorders*. New York: Raven Press.

Hsu, L. K. G. (1995). Outcome of bulimia nervosa. In K. D. Brownell & C. G. Fairburn (Eds.), *Eating disorders and obesity: A comprehensive handbook* (pp. 238–246). London: Guilford Press.

Jansen, A. (1994). The learned nature of binge eating. In D. Booth & C. Legg (Eds.), *Human appetites and their neural bases*. Oxford: Oxford University Press.

Jansen, A. (1996). How restrained eaters perceive the amount they eat. *British Journal of Clinical Psychology, 35,* 381–392.

Jansen, A. (in press). A learning model of binge eating: cue reactivity and cue exposure. *Behavior Research and Therapy*.

Jansen, A., Broekmate, J., & Heymans, M. (1992). Cue exposure vs self-control in the treatment of binge eating: a pilot study. *Behaviour Research and Therapy, 30,* 235–241.

Jansen, A., Merckelbach, H., Oosterlaan, J., Tuiten, A., & van den Hout, M. A. (1988). Cognitions and self-talk during food intake of restrained and unrestrained eaters. *Behaviour Research & Therapy, 26,* 393–398.

Jansen, A., & van den Hout, M. A. (1991). On being led into temptation: "counterregulation" of dieters after smelling a "preload'. *Addictive Behaviors, 5,* 247–253.

Jansen, A., van den Hout, M. A., & Griez, E. (1989). Does bingeing restore bulimics" alleged 5HT-deficiency? *Behaviour Research & Therapy, 5,* 555–560.

Jimerson, D. C., Lesem, M. D., Kaye, W. H., & Brewerton, T. D. (1992). Low serotonin and dopamine metabolite concentrations in cerebrospinal fluid from bulimic patients with frequent binge episodes. *Archives of General Psychiatry, 49,* 132–138.

Keesey, R. E. (1995). A set-point model of body weight regulation. In K. D. Brownell & C. G. Fairburn (Eds.), *Eating disorders and obesity: A comprehensive handbook* (pp. 46–50). London: Guilford Press.

Laessle, R. G. (1990). Affective disorders and bulimic syndromes. In M. Fichter (Ed.), *Bulimia nervosa: Basic research, diagnosis & therapy* (pp. 112–125). New York: Wiley.

Lowe, M. R., & Caputo, G. C. (1991). Binge eating in obesity: Toward the specification of predictors. *International Journal of Eating Disorders, 10,* 49–55.

Marcus, M. D. (1995). Binge eating and obesity. In H. D. Brownell & C. G. Fairburn (Eds.), *Eating disorders and obesity: A comprehensive handbook* (pp. 441–444). London: Guilford Press.

Meijboom, A., Jansen, A., Kampman, M., & Schouten, E. (in press). An experimental test of the relationship between self-esteem and concern about body shape and weight in restrained eaters. *International Journal of Eating Disorders.*

Mizes, J. S., & Christiano, B. A. (1995). Assessment of cognitive variables relevant to cognitive behavioural perspectives on Anorexia Nervosa and Bulimia Nervosa. *Behaviour Research and Therapy, 33,* 95–105.

Odwyer, A. M., Lucey, J. V., & Russell, G. F. M. (1996). Serotonin activity in anorexia nervosa after long term weight restoration: response to D-Fenfluramine challenge. *Psychological Medicine, 26,* 353–359.

Patton, G. C., Johnson-Sabine, E., Wood, K., Mahn, A. H., & Wakeling, A. (1990). Abnormal eating attitudes in London schoolgirls—a prospective epidemiological study: Outcome at twelve month follow-up. *Psychological Medicine, 20,* 383–394.

Phelan, P. (1987). Cognitive correlations of bulimia: The bulimic thoughts questionnaire. *International Journal of Eating Disorders, 5,* 593–607.

Pirke, K. M. (1995). Physiology of bulimia nervosa. In K. D. Brownell & C. G. Fairburn (Eds.), *Eating disorders and obesity: A comprehensive handbook* (pp. 261–265). New York: Guilford Press.

Polivy, J., & Herman, C. P. (1985). Dieting and bingeing: A causal analysis. *American Psychologist, 40,* 193–201.

Rodin, J., Silberstein, L., & Striegel-Moore, R. (1984). *Women and weight: A normative discontent*. Paper presented at the Nebraska Symposium on Motivation, Lincoln, NE.

Rogers, L., Resnick, M. D., Mitchell, J. E., & Blum, R. W. (1997). The relationship between socio-economic status and eating disordered behaviors in a community sample of adolescent girls. *International Journal of Eating Disorders, 22,* 15–23.

Rosen, J. C. (1995). Assessment and treatment of body image disturbance. In K. D. Brownell & C. G. Fairburn (Eds.), *Eating disorders and obesity: A comprehensive handbook* (pp. 369–373). New York: Guilford Press.

Rosen, J. C., Reiter, J., & Orosan, P. (1995). Assessment of body image in eating disorders with the body dysmorphic disorder examination. *Behaviour Research and Therapy, 33,* 77–84.

Rossiter, E. M., Wilson, G. T., & Goldstein, L. (1989). Bulimia Nervosa and dietary restraint. *Behaviour Research and Therapy, 27,* 465–468.

Sebastian, S. B., Williamson, D. A., & Blouin, D. C. (1996). Memory bias for fatness stimuli in the eating disorders. *Cognitive Therapy and Research, 20,* 275–286.

Siegel, S. (1983). Classical conditioning, drug tolerance and drug dependence. In Y. Israel & F. B. Smart (Eds.), *Research Advances in Alcohol and Drug Problems, 7,* 207–246.

Smeets, M. A. M., Smit, F., Panhuysen, G. E. M., & Ingleby, J. D. (1997). The influence of methodological differences on the outcome of body size estimation studies in anorexia nervosa. *British Journal of Clinical Psychology, 36,* 263–277.

Smith, D. E., Marcus, M. D., & Kaye, W. (1992). Cognitive-behavioral treatment of obese binge eaters. *International Journal of Eating Disorders, 12,* 257–262.

Spencer, J. A., & Fremouw, W. J. (1979). Binge eating as a function of restraint and weight classification. *Journal of Abnormal Psychology, 88,* 262–267.

Spitzer, R. L., Yanovski, S., Wadden, T., Wing, R., Marcus, M. D., Stunkard, A., Devlin, M., Mitchell, J., Hasin, D., & Horne, R. L. (1993). Binge eating disorder: its further validation in a multisite study. *International Journal of Eating Disorders, 13,* 137–153.

Stunkard, A. J., & Wadden, T. A. (1992). Psychological aspects of severe obesity. *American Journal of Clinical Nutrition, 55,* 524s–532s.

Treasure, J., & Holland, A. (1995). Genetic factors in eating disorders. In G. Szmukler, C. Dare, & J. Treasure (Eds.), *Handbook of eating disorders: Theory, treatment, and research* (pp. 65–82). Chichester, UK: Wiley.

Vitousek, K. B. (1995). Cognitive-behavioral therapy for anorexia nervosa. In K. D. Brownell & C. G. Fairburn (Eds.), *Eating disorders and obesity: A comprehensive handbook* (pp. 324–329). London: Guilford Press.

Vitousek, K. B., & Hollon, S. D. (1990). The investigation of schematic content and processing in eating disorders. *Cognitive Therapy and Research, 14*(2), 191–214.

Vize, C. M., & Cooper, P. J. (1995). Sexual abuse in patients with eating disorders, patients with depression, and normal controls. *British Journal of Psychiatry, 167,* 80–85.

Wakeling, A. (1995). Physical treatments. In G. Szmukler, C. Dare, & J. Treasure (Eds.), *Handbook of eating disorders: Theory, treatment, and research* (pp. 351–362). Chicester, UK: Wiley.

Walsh, T. B. (1993). Binge eating in bulimia nervosa. In C. G. Fairburn & G. T. Wilson (Eds.), *Binge eating: Nature, assessment and treatment* (pp. 37–49). New York: Guilford Press.

Wardle, J. (1980). Dietary restraint and binge eating. *Behaviour Analysis and Modification, 4,* 201–209.

Wardle, J. (1990). Conditioning processes and cue exposure in the modification of excessive eating. *Addictive Behaviors, 15,* 387–393.

Wardle, J. (1995). The assessment of obesity: Theoretical background and practical advice. *Behaviour Research and Therapy, 33,* 107–117.

Wardle, J., & Beinart, H. (1981). Binge eating: A theoretical review. *British Journal of Clinical Psychology, 20,* 97–109.

Welch, S. L., & Fairburn, C. G. (1994). Sexual abuse and bulimia nervosa: three integrated case control comparisons. *American Journal of Psychiatry, 151,* 402–407.

Welch, S. L., & Fairburn, C. G. (1996a). Childhood sexual and physical abuse as risk factors for the development of Bulimia Nervosa: A community-based case control study. *Child Abuse & Neglect, 20,* 633–642.

Welch, S. L., & Fairburn, C. G. (1996b). Impulsivity or comorbidity in Bulimia Nervosa: A controlled study of deliberate self-harm and alcohol and drug misuse in a community sample. *British Journal of Psychiatry, 169,* 451–458.

Westenhoefer, J., Pudel, V., & Maus, N. (1990). Some restrictions on dietary restraint. Commentary. *Appetite, 14,* 137–141.

Willett, W., & Manson, J. E. (1995). Epidemiologic studies of health risks due to excess weight. In K. D. Brownell & C. G. Fairburn (Eds.), *Eating disorders and obesity: A comprehensive handbook*. New York: Guilford Press.

Wilson, G. T. (1991). The addiction model of eating disorders: A critical analysis. *Advances in Behaviour Research and Therapy, 13,* 27–72.

Wilson, G. T. (1994). Behavioral treatment of obesity: Thirty years and counting. *Advances in Behavioral Research and Therapy, 16,* 31–75.

Wilson, G. T. (1996a). Treatment of Bulimia Nervosa: When CBT fails. *Behaviour Research and Therapy, 34,* 197–212.

Wilson, G. T. (1996b). Acceptance and change in the treatment of eating disorders and obesity. *Behavior Therapy, 27,* 417–439.

Wonderlich, S. A. (1995). Personality and eating disorders. In K. D. Brownell & C. G. Fairburn (Eds.), *Eating disorders and obesity: A comprehensive handbook* (pp. 171–176). New York: Guilford Press.

Woods, S. C. (1991). The eating paradox: how we tolerate food. *Psychological Review, 4,* 488–505.

Wurtman, R., & Wurtman, J. (1984). Nutritional control of central neurotransmitters. In K. Pirke & D. Ploog (Eds.), *The psychobiology of anorexia nervosa.* Berlin/Heidelberg: Springer.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.01

Physiological Aspects of Aging: Relation to Identity and Clinical Implications

SUSAN K. WHITBOURNE
University of Massachusetts–Amherst, MA, USA

7.01.1 INTRODUCTION 2

7.01.2 IDENTITY AND PHYSICAL CHANGES IN LATER LIFE 2
  7.01.2.1 Identity Processes 2
  7.01.2.2 A Multiple Threshold Model 3
  7.01.2.3 Physical Identity and Aging 4

7.01.3 THE EFFECTS OF AGE ON APPEARANCE AND MOBILITY 5
  7.01.3.1 Skin 5
    7.01.3.1.1 Age changes 5
    7.01.3.1.2 Clinical implications 5
  7.01.3.2 Hair 6
    7.01.3.2.1 Age changes 6
    7.01.3.2.2 Clinical implications 6
  7.01.3.3 Body Build 6
    7.01.3.3.1 Age changes 6
    7.01.3.3.2 Clinical implications 7
  7.01.3.4 Mobility 7
    7.01.3.4.1 Age changes 7
    7.01.3.4.2 Clinical implications 8

7.01.4 THE EFFECTS OF AGE ON VITAL FUNCTIONS 9
  7.01.4.1 Cardiovascular System 9
    7.01.4.1.1 Age changes 9
    7.01.4.1.2 Clinical implications 9
  7.01.4.2 Respiratory System 10
    7.01.4.2.1 Age changes 10
    7.01.4.2.2 Clinical implications 11

7.01.5 THE EFFECTS OF AGE ON REGULATORY SYSTEMS 11
  7.01.5.1 Excretory System 11
    7.01.5.1.1 Age changes 11
    7.01.5.1.2 Clinical implications 12
  7.01.5.2 Digestive System 12
    7.01.5.2.1 Age changes 12
    7.01.5.2.2 Clinical implications 12
  7.01.5.3 Immune System 13
    7.01.5.3.1 Age changes 13
    7.01.5.3.2 Clinical implications 13

7.01.6 AGE CHANGES IN THE REPRODUCTIVE SYSTEM AND SEXUALITY 13
  7.01.6.1 Female Reproductive System 13
  7.01.6.2 Male Reproductive System 14

7.01.6.3 Clinical Implications 14
7.01.7 THE NERVOUS SYSTEM 15
7.01.7.1 Central Nervous System 15
7.01.7.1.1 Age changes 15
7.01.7.1.2 Clinical implications 15
7.01.7.2 Autonomic Nervous System 16
7.01.7.2.1 Age changes 16
7.01.7.2.2 Clinical implications 16
7.01.8 AGE CHANGES IN SENSORY FUNCTIONING 17
7.01.8.1 Vision 17
7.01.8.1.1 Age changes 17
7.01.8.1.2 Clinical implications 17
7.01.8.2 Hearing 18
7.01.8.2.1 Age changes 18
7.01.8.2.2 Clinical implications 18
7.01.8.3 Balance 19
7.01.8.3.1 Age changes 19
7.01.8.3.2 Clinical implications 19
7.01.9 CONCLUDING OBSERVATIONS 19
7.01.10 REFERENCES 20

7.01.1 INTRODUCTION

Understanding the normal physiological changes associated with aging is of essential importance to clinical psychologists who work with the elderly. Issues related to depression, anxiety, and other psychological disorders are seen from new perspectives when interpreted in terms of the changes within the body that older clients are likely to experience. Of particular importance is the necessity of distinguishing between the processes of normal aging and alterations in functioning that are the result of specific diseases. Although it is true that older individuals are at heightened risk of developing chronic health problems such as arthritis, cardiovascular disease, cancer, and diabetes, these are diseases rather than normal age-related changes. Clinicians should be aware that they need to persist in the treatment of an elderly client rather than assume that the symptoms of chronic disease represent inevitable losses due to aging.

After reading this chapter, it should be clear that the physical changes associated with the aging process are complex, multifaceted, and multidirectional. Individuals age in their own unique ways, due to their genetic inheritance and lifestyle choices. Clinicians can be of tremendous help to an older person by being aware of expectable age-related changes that affect the aging population as a whole but also by planning interventions that take into account the individual's unique characteristics, life history, and resources.

In general, with regard to physical functioning, it is beneficial to advocate the "use it or lose it" principle; that is, the individual maximizes functioning by maintaining maximum levels of activity. This philosophy gives the clinician concrete tools to design treatment plans and helps to communicate a sense of optimism rather than hopelessness about aging. Furthermore, this approach often results in demonstrable gains in the older individual's daily functioning. However, the clinician is best advised to tailor the application of this principle to the client's preferences, lifestyle, and resources. It is of little help to prescribe an exercise regimen for an individual who has no access to supervised exercise facilities, has never had an interest in exercise, and who cannot afford even the proper pair of sneakers. Apart from such restrictions, however, clinicians may expect to encounter resistance of a different nature when proposing lifestyle changes to older clients. Those who have incorporated into their own identities the belief that there is in fact nothing they can do to alter their own aging will not be led by even the most effective clinician down the activity path unless this resistance can be approached and sensitively overcome. Therefore, in this chapter, a perspective incorporating the notion of identity will be discussed in regard to physical changes. With this perspective, clinicians can target interventions that take into account the ways that older individuals incorporate views about aging into their identities.

7.01.2 IDENTITY AND PHYSICAL CHANGES IN LATER LIFE

7.01.2.1 Identity Processes

The construct of identity has been used in the area of adult development to refer to the individual's sense of self over time. It is conceptualized as incorporating various content areas, including physical functioning, cognition, social relationships, and experiences in the world. In this model (Whitbourne, 1986, 1996a), identity is theorized to form an organizing schema through which the individual's experiences are interpreted. The affective content of identity, for psychologically healthy adults, takes the positive self-referential form encapsulated in the expression "I am a competent, loving, and good person;" that is, competent at work, loving in family life, and good in the sense of being morally and ethically righteous. In this model, such self-attributions are seen as contributing to the individual's emotional well-being. The developmental imperative for normal, nondepressed adults to regard themselves as having these qualities may be seen as providing a driving force in development in the adult years.

The individual's experiences, both past and present, are postulated to relate to identity through processes of assimilation and accommodation. Identity assimilation is defined as the interpretation of life events relevant to the self in terms of the cognitive and affective schemas that are incorporated in identity. The events and experiences to which the assimilation function applies can include major life events, cumulative interactions with the environment over time, or minor incidents that can have a potential impact on identity. Each type of experience can have different meanings depending on the nature of the individual's current self-conceptualizations in identity. The process of identity accommodation, by contrast, involves changing identity in response to identity-relevant experiences. An event that reflects unfavorably on identity is likely to be processed first through assimilation and, only after such efforts prove unsatisfactory, will identity accommodation follow (Brandtstadter & Renner, 1990). As a general principle, it is theorized that a healthy state of emotional adaptation involves a balance or equilibrium between the two identity processes so that the individual maintains a sense of consistency over time but is able to change when there are large discrepancies between the self and experiences (Kiecolt, 1994).

The aging process is a constant challenge to the maintenance of a stable sense of identity over time. Events caused by physical aging are qualitatively different from other experiences that are of a more transitory or fleeting nature. An unpleasant encounter in a chance meeting with a stranger that reflects unfavorably upon one's social identity will fade through time as it is dismissed through identity assimilation as lacking direct personal relevance. Aging, however, is a process that does not fade with time but may in fact grow with importance as the individual must find ways to integrate age-related changes into identity throughout adulthood.

7.01.2.2 A Multiple Threshold Model

Integrating these factors into a comprehensive framework is a "multiple threshold model" of aging (Whitbourne, 1996a, 1996b), that relates age changes to identity, coping processes, adaptational outcomes, and age-related control behaviors. The term "threshold" in this model refers to the point at which an age-related change is recognized by the individual. Before this threshold is reached, the individual does not think of the self as "aging" or "old," or even perhaps as having the real potential to be "aging" or "old." After crossing the threshold, the individual recognizes the possibility that functions may be lost through aging (or disease), and begins to accommodate to this possibility by changing identity accordingly.

The term "multiple" in this model refers to the fact that the aging process involves potentially every system in the body, so that there is in actuality no single threshold leading to the view of the self as aging. The individual may feel "old" in one domain of functioning, such as in the area of mobility, but feel "not old," "middle-aged," or possibly "young" in other domains, such as in the area of sensory acuity or intellectual functioning. Whether a threshold is crossed depends in part on the actual nature of the aging process and whether it has affected a particular area of functioning. However, it is also the case that individuals vary widely in the areas of functioning that they value. Mobility may not be as important to an individual whose major source of pleasure is derived from sedentary reading activities. Changes in the area of mobility will not have as much relevance to the individual's direct adaptation to the environment or to identity as would be the case for losses in vision or memory. In the multiple threshold model, it is assumed that changes in areas important to adaptation and competence have the greatest potential for affecting identity.

Not only are changes in central functions likely to have a greater impact on adaptation and identity, but it is possible to assume further that changes in these functions are watched for most carefully by the individual and early signs of age-related changes will be noticed in these areas. For example, the individual who takes great pride in a full and vibrant head of hair will be on the lookout for gray hairs and thinning, and the individual who values intellectual skills will scrutinize performance on activities involving memory for signs of deterioration. The increased vigilance for areas of functioning central to identity will theoretically heighten the impact of age changes.

On the other hand, with increased vigilance may come increased motivation to take advantage of control activities that reduce the impact of aging. The adult who values being in peak physical shape will work devotedly each day to avoid loss of muscle strength and aerobic capacity. Having valued these qualities, he or she will be more likely to hold onto them into old age. This situation presents a double-edged sword, in which heightened vigilance leads to behavior that offsets the aging process, thus reducing the dreaded outcomes, but also leads the individual to become acutely sensitive to signs that loss has occurred. It may be in this situation, then, that identity accommodation becomes important. The individual who recognizes that he or she is aging is less likely to become discouraged and frustrated at evidence that certain age-related changes cannot be completely reversed. The assimilative individual who desperately attempts to hang onto a youthful identity has the potential to be demolished when an event occurs that makes it abundantly clear that the aging process has taken its inevitable toll.

The multiple threshold concept is most easily viewed as a set of linear processes, with the outcome of passing through a threshold an alteration in the individual's level of well-being. The process is, however, cyclical. Each time a threshold is passed, the individual's identity has the potential to be altered through identity accommodation. When this happens, the crossing of the threshold serves as an event relevant to the self that becomes integrated into identity. The resulting change in identity then alters the individual's subsequent vigilance regarding future age-related in changes in other areas of functioning and the behaviors relevant to the threshold just crossed.

7.01.2.3 Physical Identity and Aging

The substantive content of identity incorporates the domains of physical, psychological, and social functioning. Within the domain of physical functioning are the attributes of appearance, competence, and health (Whitbourne & Primus, 1996). Physical identity may be regarded as one of the fundamental, yet overlooked, areas within psychological gerontology. For example, in a preliminary test of the multiple threshold model, it was found that the aspect of physical functioning in particular that seems to trigger the crossing of a threshold with regard to aging is the recognition of loss of physical competence, an area that is generally not given a great deal of attention by clinicians (Whitbourne, 1996b).

Individuals who use identity assimilation tend to deny as much as possible the importance of age-related changes in valued aspects of physical functioning. For these individuals, it might be speculated that changes in functioning that are related to age are not acknowledged as a result of the aging process, but instead are attributed to transitory states of health or health-related behaviors. This type of denial may be seen as an effort to preserve and protect the individual's sense of the self as competent and consistent over time.

Those older adults who rely more or less exclusively on identity accommodation overreact to small events or experiences and draw overly broad and sweeping conclusions from one instance or situation. With regard to aging, they may conclude prematurely that they are "over the hill" and fail to take preventative actions (Woodward & Wallston, 1987) because they adopt the self-image of being too weak or feeble to continue to maintain their participation in activities that were formerly rewarding to them (Janelli, 1993). As a result, they can become unduly depressed and hopeless (Parmalee, Katz, & Lawton, 1991). In a test of the multiple threshold model (Whitbourne, 1996b), identity accommodation was found to predict awareness of age-related physical changes. Thus, individuals who are highly responsive to their experiences and do not have a firm sense of self are more likely to be sensitive to the effects of normal aging on their bodies. Greater sensitivity to age-related changes, in turn, appears to be predictive of lower self-esteem.

A balanced approach to adapting to age-related changes in valued aspects of physical functioning involves taking precautions and attempting to preserve functions for as long as possible, and adapting to them when they occur by finding ways to compensate for losses. The individual may find a compromise between assimilation and accommodation by recognizing the reailty of aging, but not letting this knowledge interfere with daily functioning or future plans. There may be emotional costs involved in the recognition of impending losses, but individuals who maintain a balanced approach rebound from setbacks with renewed vigor and optimism (Labouvie-Vief, Hakim-Lawson, & Hobart, 1987; Ryff, 1989).

With this model as a backdrop, research on the physical aspects of aging will now be summarized. The concepts of the multiple threshold model and the identity processes of assimilation and accommodation will be used as the basis for suggesting clinical implications. The underlying assumption is not that the aging process must or should be slowed, but that the individual's adaptation to the aging process can be maximized by working with the assimilative or accommodative functions of identity. Ultimately, the value of these interventions can be measured in terms of improved physical health and in terms of an improved or perhaps more realistic outlook on the part of the individual toward the aging process.

7.01.3 THE EFFECTS OF AGE ON APPEARANCE AND MOBILITY

7.01.3.1 Skin

7.01.3.1.1 Age changes

A variety of changes contribute to the phenomena of wrinkling and sagging, the most apparent age-related effects on the skin. The epidermis, the outer layer of the skin, becomes flattened (Kligman, Grove, & Bolin, 1985), and as cells in the epidermis are replaced through cell renewal processes, their patterns become less organized and their arrangment less regular. These changes at the microscopic level are reflected in visible changes of the geometric furrows visible on the surface of exposed areas of the skin (Lavker, Kwong, & Kligman, 1980). Wrinkling and sagging can be explained in large part by changes in the middle, dermal layer of the skin, notably a decrease in the connective tissue collagen (Kligman et al., 1985). Elastin fibers become more brittle so that the ability of the skin to conform to movement of the bones is compromised. The subcutaneous fat on the limbs decreases and instead collects in areas of fatty deposits, such as around the waist and hips. A decrease in muscle mass, as will be discussed later, further adds to the loss of firmness in the skin's appearance.

There are also significant changes in the sweat and oil-producing glands within the dermis that maintain body temperature and lubricate the skin surface (Bolin & Pratt, 1989; Kurban & Bhawan, 1990). In areas such as the palms and underarms, the sweat glands become less active. The sebaceous glands that lubricate the skin with their secretions diminish in activity, and consequently the skin becomes rougher, dryer, and more vulnerable to surface damage. These changes lead to heightened risk of medical problems such as dermatitis (Grove, 1989) and pruritis (excessive itching), as well as to general discomfort (Kligman, 1989).

Age-related alterations also occur in the coloring of the skin. There are fewer melanocytes, and those that remain develop irregular areas of dark pigmentation. Under the skin surface, capillaries and small arteries become dilated, creating small irregular colored lines. Varicose veins may also form, creating an uneven surface of knotty, bluish, and cordlike irregularities. The loss of capillaries means that fair-skinned people look paler, and in general the blood vessels and bones can be seen more easily under what is a more transparent skin surface.

7.01.3.1.2 Clinical implications

Age-related changes in the skin and skin-related structures alter the skin's protective functions as well as its appearance. There is less insulation provided against extremes of temperature, and less of a barrier against environmental agents that can irritate the skin and cause dermatitis (Grove, 1989). Even if the individual is not particularly concerned about the effects of age on appearance, these changes in comfort level can trigger the crossing of an aging threshold. However, for many individuals, particularly in middle age (Whitbourne, 1996a, 1996b), age changes in the skin that lead to wrinkling are very noticeable and can have an impact on the individual's identity. In particular, comparisons of present appearance with pictures or memories of early adulthood can trigger identity accommodation in a negative way to older people who valued their youthful image (Fenske & Albers, 1990; Kleinsmith & Perricone, 1989)

Having crossed this threshold in terms of appearance, there are many steps that older individuals can take to slow, compensate for, or correct the changes in appearance caused by aging of the skin. The primary method of prevention is for fair-skinned people to avoid direct exposure to the sun and to use sunblocks when exposure cannot be avoided (Gilchrest, 1989). There are also many ways for the individual to compensate for age-related changes in the skin once they have become manifest. To counteract the fragility, sensitivity, and dryness of the skin, the individual can use sunscreens, emollients, and fragrance-free cosmetics (O'Donoghue, 1991). If a client raises concerns about these aspects of appearance, interventions here, as with hair color (see below), must attempt to strike a judicious balance between intrusiveness into matters of personal taste and the concerns of preventive practice.

7.01.3.2 Hair

7.01.3.2.1 Age changes

With increasing age in adulthood, for men and women, the hair on the head and body loses pigmentation and takes on a white appearance due to a decrease in melanin production in the hair follicles. The rate at which hair color changes varies from person to person due to variations in the timing of onset and rate of melanin production decrease across the surface of the scalp. Interestingly, although gray hair is thought to be a universal feature of aging, the degree of hair grayness is not as reliable an indicator of a person's age as is the extent to which hair on the body (axillary and pubic) has turned gray (Kenney, 1989).

Gradual and general thinning of scalp hair occurs in both sexes over the years of adulthood, although loss of hair is popularly regarded as a concern of men. Hair loss results from destruction or regression of the germ centers that produce the hair follicles underneath the skin surface. However, in men with pattern baldness, the hair follicles do not actually die but change the type of hair they produce from coarse terminal hair to fine translucent vellus hair. Men may also experience thinning of the hairs in their whiskers, and a growth of coarse hair on the eyebrows and inside the ear. Patches of coarse terminal hair may develop on the face of women, particularly around the chin.

7.01.3.2.2 Clinical implications

As is true for wrinkles, the threshold for age-related changes in the hair, at least for graying, appears to be crossed at some point in the 30s or 40s (Whitbourne, 1996b). Compared to wrinkles, gray hairs are relatively easy to detect; however, unlike wrinkles, gray hairs can be returned to virtually their original state through the use of hair dye. Changes in hair thickness, however, are not so easily reversed. The desire to disguise or stop the apparent signs of aging through surgery or the wearing of hairpieces is widespread, as is evident in the many advertisements for hair loss replacement processes.

There are many ways in which clinical intervention can be of use when changes in hair amount and color becomes a concern. Apart from providing information about commercial products that can be used if the individual chooses to try to combat the aging process, interventions can focus on ways the individual can learn to reframe age changes in the hair's appearance, particularly with regard to hair thinning, in a manner that allows for a balance between identity assimilation and accommodation. Since there are no practical implications of hair loss or hair graying, interventions in this area can be entirely at the discretion of the client and clinician. Unlike cardiovascular functioning, in which denial of age-related changes can prove fatal, the denial of changes in the hair does not have severe health consequences. The only exception is when the individual is at risk of ostracism or ridicule by others due to inappropriate attempts to maintain what he or she thinks is a youthful head of hair. In these cases, the accommodation process may need to be brought to bear on the situation so that the individual can accept that patch of skin on the top of the head or can find a hair dye that does not attempt to recreate the look of a 20-year-old. Of course, these issues quickly become matters of taste and preference, and short of major concerns or problems, should ultimately be left to the individual's discretion.

7.01.3.3 Body Build

7.01.3.3.1 Age changes

It is a well-established fact that over the course of adulthood there is a consistent pattern of a reduction in standing height, occurring at a greater rate after the 50s and particularly pronounced in women (Shephard, 1978). The major cause is loss of bone mineral content in the vertebrae, which leads them to collapse and cause compression in the length of the spine (Garn, 1975). Changes in the joints and flattening of the arches of the feet can further contribute to height loss (Kenney, 1989).

Total body weight increases from the 20s until the mid-50s, after which total body weight declines. Most of the weight gain in middle adulthood is due to an accumulation of body fat, particularly around the waist and hips. The weight loss that occurs in the later years of adulthood is not due to a slimming of the torso but to a loss of lean body mass consisting of muscle and bone (Chien et al., 1975; Ellis, Shukla, Cohn, & Pierson, 1974). Consequently, very old adults may have very thin extremities but fatty areas in the chin, waist, and hips. Middle-aged and older women are particularly likely to experience this accumulation of body fat around the torso, with a gain of abdominal girth amounting to 25–35% across the adult years compared to 6–16% for men over a comparable time period (Shephard, 1978). However, contemporary changes in lifestyle and improved nutrition may be altering this trend, with current cohorts of older women showing a decrease rather than an increase in body fat (Rico, Revilla, Hernandez, Gonzalez-Riola, & Villa, 1993).

7.01.3.3.2 Clinical implications

Changes in body fat that lead to the appearance of a sagging or heavier body shape can result in increased identification of the self as moving away from the figure of youth. The development of "middle-aged spread" is one of the first occurrences to trigger identity accommodation of the self as aging, perhaps even before the first gray hairs have sprouted. Indeed, awareness of changes in body composition can occur surprisingly early in adulthood, perhaps by the 30s (Whitbourne, 1996b).

Fortunately, the crossing of the body fat threshold may be readily compensated, and before the individual gives up on trying to slow down this particular process, there are many activities that the clinician can promote if this is an area of concern for the individual. It is well established that participation in active sports and exercise activity can offset the deleterious effects of aging on body fat accumulation. For instance, endurance athletes do not gain weight and they maintain their muscular physiques throughout adulthood for as long as they continue to train (Suominen, Heikkinen, Parkatti, Forsberg, & Kiiskinen, 1977). Participation in exercise training programs can even be of value for middle-aged and elderly adults who were sedentary throughout their lives. By engaging in vigorous walking, jogging, or cycling for 30–60 minutes a day for 3–4 days a week, the sedentary adult can expect to achieve positive results in a period as short as 10–20 weeks (Whitbourne, 1996a). Furthermore, the same activities that a middle-aged person might engage in to combat body fat changes can also have very positive cardiovascular and general health benefits. Thus, vanity might actually serve a protective function in this area.

Changes in body height are less easily compensated, but identity processes are also less likely to be engaged in this area unless the individual has had concerns about height from the early years and on. Clinical interventions here may focus on the maintenance of bone strength through weight resistance training and although these activities might not have an enhancing effect on height, they can be very important in promoting mobility (see below), another major concern for many older adults.

7.01.3.4 Mobility

7.01.3.4.1 Age changes

The individual's ability to move around in the physical environment is a function of the integrity of the muscles, bones, joints between the bones, tendons and ligaments that connect muscle to bone, and contractility of flexor and extensor muscles. Mobility changes in important ways over the course of the adult years such that movement becomes more difficult, more painful, and often less effective for the older adult.

Between ages 40–70 years, the loss of muscle strength amounts to approximately 10–20% with more severe losses of 30–40% after ages 70 to 80 years; losses that are more pronounced in the muscles of the lower extremities (McArdle, Katch, & Katch, 1991). This pattern of diminishing muscle strength appears to be due to atrophy of muscle fibers, particularly the "fast twitch" fibers important in developing the rapidly accelerating powerful contractions normally associated with strength. The "slow twitch" fibers involved in maintaining posture and muscular contractions over protacted periods of exertion remain constant over the adult years. The atrophy of fast-twitch fibers is thought to result from the loss of the motoneurons that activate the muscles. After the muscle fiber dies, it is replaced initially by connective tissue and in its final stages, by fat (Fiatrone & Evans, 1993). However, declines in muscle strength are not completely accounted for by loss of muscle mass, and there are other influences on age effects on muscle strength as yet to be determined (Kallman, Plato, & Tobin, 1990; Overend, Cunningham, Kramer, Lefeve, & Paterson, 1992).

There are individual variations that can lead to important deviations from such a general pattern of decline. The extent to which aging affects loss of muscle strength depends in part on which gender is being tested, the general level of activity in which the individual has typically engaged, the particular muscle group being tested, and whether the type of muscle strength being assessed is static (isometric) or dynamic.

Although the process of bone growth and remodeling continues, the overall course of bone development in adulthood is toward loss of bone strength, resulting in diminished ability of the bones to withstand mechanical pressure and to show greater vulnerability to fracture. The decrease in various measures of bone strength ranges from 5 to 12% per decade from the 20s through the 90s (McCalden, McGeough, Barker, & Court-Brown, 1993). Loss of bone strength is generally explained as a function of the loss of bone mineral content, meaning that the bone becomes increasingly porous and unable to support the loads it must bear. The extent to which bone mass decreases over the years of adulthood amounts to approximately 20–30% of total bone mineral content in women, and approximately half that amount in men (Mazess, 1982; Riggs et al., 1981). The period of maximum bone loss is between the 50s and the 70s (Sparrow, Beausoleil, Garvey, Rozner, & Silbert, 1982). The explanation of the underlying process that causes loss of bone mineral content is that the rate of resorption exceeds that of new bone growth in later adulthood, giving a net result of a reduction in bone mass (Sherman et al., 1992). In part, bone density and the loss of bone mass in the later years of adulthood may be a function of genetic factors (Dargent & Breart, 1993; Kelly et al., 1993). Lifestyle also seems to play a role, however, including factors such as physical activity, smoking, alcohol use, and diet which can account for 50–60% of the variation in bone density (Krall & Dawson-Hughes, 1993). There are also hormonal influences on bone mass, as indicated by the observation that bone mineral loss in women proceeds at a higher rate in postmenopausal women who are no longer producing estrogen in monthly cycles (Nuti & Martini, 1993).

Although the aging of joints is most commonly associated with the later years of life, degenerative processes that reduce the functional efficiency of the joints begin even before the individual reaches skeletal maturity. Restrictions of movement and discomfort are therefore a potential problem for adults of any age, but they occur with increasing frequency as age progresses. The decline in joint functioning can be accounted for by age losses in virtually every structural component of the joint. Starting in the 20s and 30s, the arterial cartilage begins to thin, fray, shred, and crack (Adrian, 1981). Unprotected by cartilage, the underlying bone eventually begins to wear away. At the same time, outgrowths of cartilage develop and these interfere with the smooth movement of the joint. Age-related weakening of the muscles further contributes to restrictions in range of movement due to changes in the joints themselves (Vandervoort et al., 1992). Changes at the cellular level in the structure of collagen and elastin comparable to those that occur in the dermis are thought to contribute to loss of flexibility, strength, and resiliency of connective tissue. Diminished efficiency of circulation may contribute further to deteriorative changes. Since the cartilage receives little vascular supply to begin with, any reductions in adulthood due to aging or arterial disease will further reduce its reparative ability. Furthermore, the joints are subjected to an extreme amount of trauma throughout life, including the strains and sprains encountered during everyday activities and strenuous exercise, activities that further contribute to their deterioration.

7.01.3.4.2 Clinical implications

For many individuals, thresholds of aging become painfully crossed with each newly-discovered joint ache or mobility restriction, sometimes beginning in the 40s (Whitbourne, 1996b). Changes in the structures that support movement have many pervasive effects on the individual's life, resulting in restrictions in activity and pain which can interfere with the individual's psychological adaptation and sense of well-being (Hughes, Edelman, Singer, & Chang, 1993). In part, the effects of mobility loss can be attributed to the enforced reliance on others that is created, leading to other deleterious processes such as being treated like a dependent child (Whitbourne & Wills, 1993).

One of the most serious outcomes of reduced muscle strength, bone strength, and joint mobility is the heightened susceptibility of older individuals, particularly women, to falls (Dargent & Breart, 1993; Lord, Clark, & Webster, 1991). If a broken hip results, the individual is more likely to suffer long-term disability and dependency (Roberto, 1992). Unfortunately, individuals might overreact to falls by developing a "fear of falling" (Downton & Andrews, 1990) or lowered sense of self-efficacy regarding the ability to avoid a fall (Tinetti & Powell, 1993). As a result, they become less stable on their feet and avoid physical activities that might benefit their strength and stability. Understandably, these individuals are vulnerable to feelings of depression and anxiety. Other adults may react to falls through identity assimilation, by repressing their occurrence (Wright, Aizenstein, Vogler, Rowe, & Miller, 1990), which might maintain their sense of self-efficacy but place them at risk of further serious injury.

This summary of age-related changes in mobility describes a fairly downhill process, but there are many interventions clinicians can pursue that will have practical advantages. As has been mentioned already, it is important to structure an intervention around the nature of the individual's identity processes, particularly as balanced against the individual's actual age-related losses and possiblities for compensation. Given such an assessment, older individuals can be encouraged to participate in a variety of physical exercises that can prove to be extremely helpful in maintaining or restoring mobility. A regular program of exercise can help the middle-aged and older adult compensate substantially for the loss of muscle fibers. Although there is nothing that can be done to stop the loss of muscle cells, the remaining fibers can be strengthened and work efficiency increased through exercise training even in persons as old as 90 years (Fiatrone et al., 1990). Inactivity can accelerate the loss of muscle strength, and the same is true of bone loss. Older individuals can also benefit from resistance training exercises that, within limits, increase the stress placed upon the bone (Dalsky et al. 1988; Rikli & McManis, 1990; Smith, 1981). Although degenerative changes in the joints are not reversible, it is possible for older individuals to benefit from exercise training, particularly if it is oriented toward promoting flexibility (Jirovec, 1991), strengthening the muscles that support the joints, increasing the circulation of blood to the joints thereby promoting the repair of injured tissues (Brooks & Fahey, 1984), and decreasing the risk of injury (Stamford, 1988).

7.01.4 THE EFFECTS OF AGE ON VITAL FUNCTIONS

7.01.4.1 Cardiovascular System

7.01.4.1.1 Age changes

The aging process results in serious limitations of the heart's functional requirement to pump blood continuously through the circulatory system at a rate that adequately perfuses the body's cells. Although a distinct entity in terms of underlying processes, fatal cardiovascular diseases do become more probable with advancing age in adulthood. These diseases can have widespread effects on daily life in addition to providing constant sources of reminders of the individual's mortality. For example, the chest pains associated with angina, a chronic cardiac illness, are not only uncomfortable but provide clear warning signals of the heart's impending failure.

The reduction in the heart's pumping capacity is due to a variety of changes affecting the structure and function of the heart muscle walls, particularly the left ventricle which becomes progressively thicker with each decade in adulthood and less elastic (Kitzman, Scholz, Hagen, Ilstrup, & Edwards 1988; Kitzman & Edwards, 1990) as the number of myocardial cells decreases and the remaining cells become hypertrophied (Olivetti, Melissari, Capasso, & Anversa, 1991). The decreased capacity of the ventricle walls to expand results in a reduced and delayed filling of the left ventricle and the ejection of less blood into the aorta. The cardiac muscle also becomes less responsive to the neural stimulation of the "pacemaker" cells in the heart that initiate each contraction (Montamat & Davies 1989; Schulman & Gerstenblith, 1989).

Effects of aging on the arteries further compromise the system's ability to distribute blood to the body's cells. Although it is difficult to separate the effects of aging from those of atherosclerosis (a disease in which the arteries become rigid and narrowed by fatty accumulation), the aging process itself appears to cause changes in the aorta and arteries (Shimojo, Tsuda, Iwasaka, & Inada, 1991). The wall of the aorta becomes less flexible, so that the blood leaving the left ventricle of the heart is faced by more resistance and cannot travel as far. The walls of the arteries throughout the body become thicker and less flexible. Impedence of blood flow through the arteries is further augmented by the accumulation of lipids in the blood (Heiss et al., 1980).

Although age-related changes in the heart and arteries reduce the amount of blood available to the cells of the body, under normal conditions these changes are not particularly pronounced or noticable. The effects of aging of the circulatory system are most apparent while the individual is engaging in aerobic exercise, when there is a reduction both in maximum oxygen consumption (aerobic capacity) and the maximum attainable heart rate (Lakatta, 1987; Morley & Reese, 1989). Aerobic capacity decreases in a linear fashion throughout the adult years, so that the average 65-year-old individual has 30–40% of the aerobic capacity of the young adult (McArdle et al., 1991). Other functional variables decrease to a commensurate degree, such as the amount of blood pumped at each beat of the heart (Weisfeldt & Gerstenblith, 1986), and the cardiac output per minute. These decreases in the functional capacity of the heart mean that less oxygen reaches the muscles during exercise. Changes in the heart and arteries can also account, in part, for the observation that older individuals have higher levels of systolic and diastolic blood pressure, both at rest and during exercise.

7.01.4.1.2 Clinical implications

Although not as apparent in its effects on identity as appearance, the functioning of the cardiovascular system is an important influence on the individual's feelings of well-being and identity. The efficiency of the cardiovascular system is essential to life so that threats to the integrity of this system are perceived as highly dangerous. Awareness of reduced cardiovascular efficiency can therefore serve as reminders of one's own personal mortality.

In working with middle-aged and older adults, clinicians can turn to the wealth of research pointing to the effectiveness of exercise in slowing or reversing the effects of the aging process. The results of research on the effectiveness of exercise training have consistently revealed improved functioning in long-term endurance athletes, master athletes, exercisers, and previously sedentary adults (Whitbourne 1996a; McArdle et al., 1991). Previously sedentary adults can improve to levels that approximate the improvements achieved by younger adults (Govindasamy, Paterson, Poulin, & Cunningham 1988; Hagberg et al., 1989). Even moderate or low-intensity exercise can have beneficial effects on healthy sedentary elderly (Foster, Hume, Byrnes, Dickinson, & Chatfield 1989; Hamdorf, Withers, Penhall, & Haslam, 1992). Further, aerobic exercise training has the positive effect of lowering the heart rate and improving work load intensity and duration in submaximal exercise, (Morey et al., 1991; Poulin, Paterson, Govindasamy, & Cunningham, 1988), ultimately placing less stress upon the heart during exertion. Other benefits of exercise training are improvements in the peripheral vasculature (Blumenthal et al., 1989), lipid metabolism (Tamai et al., 1988), and blood pressure during or immediately after exertion (Webb, Poehlman, & Tonino, 1993). That exercise can have such positive effects on physical functioning in old age, even if it is begun after a lifetime of sedentary patterns, provides strong evidence in favor of the view that the rate of the aging process can be altered significantly through active lifestyle choices made by the individual.

The main advantage that exercise seems to hold as a means of retaining a higher level of cardiovascular functioning is that it provides a continued potent stimulus for the muscle cells of the heart to undergo strong contractions so that they retain or gain contractile power. The greater strength of the myocardial muscle improves the functioning of the left ventricle and as a result, more blood can be ejected from the left ventricle (Ehsani, Ogawa, Miller, Spina, & Jilka, 1991). The other advantage of exercise training is that it makes it possible for the individual to "save" energy during aerobic work that is less than maximal by fulfilling the demands of the work load but placing less stress on the heart. Because more blood is ejected with each cardiac muscle contraction, the same output of blood can be pumped per minute but at a lower heart rate. The effects of training on cardiac functioning under submaximal conditions are of interest in that these performance situations are closer to the conditions under which people exert themselves in their daily lives. Exercise training also has favorable effects on the body's performance by increasing the efficiency of metabolism in the working muscles (Meredith et al., 1989).

In addition to the advantages of exercise training for the cardiovascular system, it is well established that adults who become involved in aerobic activities experience a variety of positive effects on mood, anxiety levels, and particularly feelings of mastery and control, leading to enhanced feelings of self-esteem (Hill, Storandt, & Malley 1993; McAuley, Lox, & Duncan, 1993). Although some researchers have demonstrated there to be positive effects of exercise on cognitive functioning (Chodzko-Zajko, Schuler, Solomon, Heiul, & Ellis, 1992; Stevenson & Topp, 1990), this effect is not observed consistently (Hill et al., 1993). Nevertheless, it appears clear that older adults who enter exercise training programs "feel better," and this improvement in mood can be clinically useful for many aging clients.

A great deal has been written in the area of behavioral medicine on strategies for increasing the compliance of cardiac patients with dietary and exercise controls over the progress of their disease. In terms of the normal aging process, similar interventions might be suggested. Those individuals who deny the effects of aging through identity assimilation must be brought to recognize that they will run real risks to their lives if they either fail to engage in preventative behaviors or engage in strenuous activities without proper supervision or training. In contrast, middle-aged and older individuals who overreact to age changes in what they have heard about cardiovascular functioning, or perhaps to the appearance of slight symptoms, need to be given assurance that they can nevertheless continue to engage in exercise activities and that, indeed, such activities are necessary for the prevention of further and more serious health limitations.

7.01.4.2 Respiratory System

7.01.4.2.1 Age changes

The airways in the respiratory system permit gas exchange between the blood and air, making it possible for the body's cells to receive support for their metabolic activities. Aging's primary effect is to reduce the quality of gas exchange in the lungs so that less oxygen from the outside air reaches the blood (Kenney, 1989). Aging also reduces vital capacity, the amount of air that is moved into and out of the lungs at maximal levels of exertion (Reddon, 1981), and forced expiratory volume, the amount of air that can be breathed out during a short amount of time (Smith, Cunningham, Patterson, Rechnitzer, & Koval, 1992). These reductions in respiratory functioning result from changes in pulmonary structures such that the airways, and particularly lung tissue, lose the elastic ability to resist expansion as they fill with air. The changes in lung structures that account for decreased elastic recoil and increased compliance are alterations in the composition and structure of the elastin and collagen composition of lung tissue (Brandstetter & Kazemi, 1983; D'Errico et al., 1989). Increased rigidity of the chest wall further lower the lung's ability to be fully compressed during expiration and expanded during inspiration (Mahler, Rossiello, & Lohe, 1986). These changes mean that less than the maximal amount of air can be brought into and out of the lungs, particularly under conditions of exertion (Teramoto, Fukuchi, Nagase, Matsuse, & Orimo, 1995). The greater rigidity of the chest wall also increases the amount of work that must be performed by the respiratory muscles during the inspiratory and expiratory phases of the ventilation cycle.

7.01.4.2.2 Clinical implications

Age changes in respiration can lead to feelings of dyspnea and fatigue associated with exertion, sensations that can approach the intensity of a panic attack. Adding to age effects on cardiovascular functioning, changes in the respiratory system can lead the aging individual to avoid strenuous activities, a consequence that further impairs the individual's cardiovascular and respiratory efficiency. Because both of these functions are so crucial to life, and because shortness of breath is so frightening, the individual might conclude prematurely that death is around the corner.

The effects of exercise training on respiratory functioning are encouraging (Blumenthal et al., 1989) but the specific effects of exercise on respiratory functioning are not as dramatic as the effects on the cardiovascular system. Equally, if not more, beneficial to the respiratory function is the avoidance of cigarette smoking (Hermanson, Omenn, & Kronmal, 1988).

Taking these factors into account, clinicians may wish to treat instances of dyspnea leading to panic as instances of overaccommodation, in which the individual may be brought to a more balanced state through recognizing the fact that some age-related changes in breathing capacity are in fact inevitable. On the other hand, for individuals who are using identity assimilation to deny the importance of factors that impede respiratory efficiency, notably cigarette smoking, steps need to be taken to increase identity accommodation and the recognition that this behavior can accelerate the deleterious effects of aging on respiratory efficiency and feelings of comfort during exercise.

7.01.5 THE EFFECTS OF AGE ON REGULATORY SYSTEMS

7.01.5.1 Excretory System

7.01.5.1.1 Age changes

It has been known for years that there are significant and widespread changes in the structure of the kidneys, changes that are reflected in cross-sectional studies in impaired efficiency across adulthood on every measure of renal functioning studied. (Rowe, Shock, & DeFronzo, 1976). An early estimate of the loss of renal functions over the adult years was a figure of 6% per decade from age 20 and continuing into the 90s, a loss that exists in longitudinal studies as well (Shock, Andres, Norris, & Tobin, 1978). The cause of aging of the kidney is loss of structure and function of the basic cell of the kidney, the nephron (McLacklan, Guthrie, Anderson, & Fulker, 1977). There are also independent losses in mechanism within the kidney responsible for concentrating urine (Rowe, 1982).

Particularly challenging to the aging kidney is aerobic exercise because it diverts blood to the working skeletal muscles, and thereby causes a further reduction in the blood flow through the kidneys. Equally important, the urine concentrating mechanism begins to fail during exercise or under extreme conditions of heat when the individual begins to perspire. Fatigue, changes in body chemistry, and potentially harmful changes in bodily fluid levels occur more rapidly in the older adult who cannot conserve sodium and water adequately under these conditions.

The effects of aging on the bladder are of great importance to the individual's conscious experience of aging. Adults past the age of 65 years experience a reduction in the total amount of urine they can store before feeling a need to void, and more urine is retained in the bladder after the individual has attempted to empty it. These phenomena appear to be related to age changes in the connective tissue of the bladder causing the organ to lose its expandability and contractility somewhat like that which is seen in the lung. Furthermore, recognition of the need to void may not occur until the bladder is almost or even completely filled. This means that the individual has less or perhaps no time to reach a lavatory before leakage or spillage occurs.

The most significant effect of changes with age in the bladder is on patterns of urinary incontinence. The prevalence of incontinence among the population 60 years and older is estimated to be 19% for women and 8% for men (Herzog, Diokno, Brown, Normolle, & Brock, 1990) but can reach as high as 36% among community-dwelling elderly with dementia

(Ouslander, Zarit, Orr, & Muira, 1990). Women are more likely to suffer from stress incontinence, which refers to loss of urine at times of exertion such as when one is laughing, sneezing, lifting, or bending and is the result of weakness of the pelvic muscles. Urge incontinence, which is more prevalent in men, involves urine loss following an urge to void or lack of control over voiding with little or no warning (Diokno, Brock, Brown, & Herzog, 1986); it is related to prostatic disease or incomplete emptying of the bladder. Among the community-dwelling elderly, each of these conditions, particularly urge incontinence, is reversible and may disappear within a year or two of its initial development (Herzog et al., 1990)

7.01.5.1.2 Clinical implications

Although not insignificant in terms of the body's physiological efficiency, the aging of the kidney is unlikely to have a discernible impact on everyday functioning. However, age-related changes in the bladder are another matter, and can be important threshold phenomena (Whitbourne, 1996b). Incontinence can be highly disruptive to the older individual's everyday life, causing distress and embarrassment (Hunskaar & Vinsnes, 1991; Ouslander & Abelson 1990; Wyman, Hawkins, & Fantl, 1990). Not only do such occasions involve shame, but they feed into the association in many people's minds between "senility" and urinary incontinence. Other horrors of incontinence may also trigger overaccommodation, such as loss of independence and respect. Given the many associations to urinary continence, it would not be unreasonable to suppose that even a single episode of stress incontinence could create a threshold experience and overaccommodation. Unfortunately, advertisers prey on this vulnerability of older adults in commercials for adult diapers that suggest these changes to be expectable and unpreventable. Clinicians needs to work with older adult clients to dispel these myths.

On the positive side, once identified, incontinence is a condition that in many cases can be managed if the individual is able to learn new behavioral strategies. Through behavioral techniques, sometimes involving only very simple exercises, the problem can be held in check if not reversed (Burgio & Engel, 1990; Burns et al., 1993).

Age effects on renal functioning significantly affect the older person's ability to excrete medications, a fact that can be of great importance in a therapeutic context. When an older person is given the the same dosage of medication appropriate for a younger adult, more of the drug will remain in the bloodstream over the period of time between doses. With repeated dosages of the drug it is more likely that toxic levels will build up in the blood. Unless the dosage is adjusted to take into account this lower rate of tubular transport, drugs may have an adverse impact instead of their intended benefits (Lamy, 1988; Montgomery, 1990).

7.01.5.2 Digestive System

7.01.5.2.1 Age changes

The documented effects of aging on digestion in the empirical literature are remarkably minor compared to the lavish attention devoted by advertisers and the popular media to the supposed inefficiencies of digestive processes in older adults. One change with age that may have implications for digestive functioning is a reduction in the metabolism of certain nutrients in the stomach and small intestine, but other structures such as the esophagous, liver, pancreas, and large intestine appear to be relatively spared by the aging process (Whitbourne, 1996). More important than constipation, dyspepsia, and reduced sensory capacities for enjoying food are the beliefs that people hold as communicated through the media, the social context in which food is eaten, other physical and cognitive deficiencies, and the individual's lifelong patterns of nutrition (Costello & Moser-Veillon, 1992; Fischer & Iohson 1990; Ryan, Craig, & Finn, 1992; Wedman et al., 1991).

7.01.5.2.2 Clinical implications

Changes in the digestive system at the physiological level due to aging are not likely to propel the individual over an aging threshold (Whitbourne, 1996b). However, so many elders have been exposed to inaccurate conceptions in the media about the aging digestive system that clinicians have an important psychoeducational role to play. In the area of digestion and particularly elimination, embarrassment may lead the older individual to avoid discussions of anxiety-provoking issues that, if discussed and allayed, could ease these very uncomfortable feelings. Furthermore, questions about food intake are essential for clinicians to explore, as inadequate nutrition may lead to symptoms that mimic those of psychological or cognitive impairment disorders. Unchecked, a vicious cycle may be created, as depression can lead to loss of interest in food and food preparation (Rosenbloom & Whittington, 1993).

Another clinical issue related to digestion concerns fecal incontinence. Older adults, particularly those who overrely on identity accommodation, may associate irregularities in defecation with feared diseases and the prospects of institutionalization in later life (Holt, 1991; Wald, 1990). Older persons who believe that "senility" is an inevitable feature of aging may regard with alarm any indication that their patterns of elimination are changing, seeing in such changes a more ominous significance that can risk putting them over the threshold. The anxiety created by this concern may contribute further to gastrointestinal problems so that what originates as a temporary problem comes to have a more prolonged course. Clinicians can intervene in this process by sensitive discussion of this very personal and potentially frightening area of daily life.

7.01.5.3 Immune System

7.01.5.3.1 Age changes

There is substantial evidence of reduced immune system functioning across age groups of adults. T-cells, which destroy antigens (foreign substances that enter the body), lose effectiveness over the adult years (Bloom, 1994; Globerson, Eren, Abel, & Ben-Menahem, 1990), in part due to changes in the thymus gland. Autopsy studies have revealed that deterioration of the thymus gland begins shortly after sexual maturity is reached, so that by the time the individual is 45–50 years old, the thymus retains only 5–10% of its peak mass. As a result of these changes, there are more immature T-cells present both within the thymus gland and in the bloodstream. Other immune system cells, including natural killer (NK) cells, K cells, and macrophages, do appear to retain their functioning into old age (Bloom 1994; Kutza, Kaye, & Mwasko, 1995); however, another entire line of defenses is altered in the aging process through the reduced activity of helper T-cells, increased activity of suppresser T-cells, and the increasing development of certain autoimmune diseases.

7.01.5.3.2 Clinical implications

Investigations have been stimulated by progress in psychoneuroimmunology, a field in which the intricate connections are examined between affective states such as stress and depression, nervous system functioning, and operation of the immune system (O'Leary 1990; Vollardt, 1991). For example, elderly individuals with high levels of life stress have been found to experience lower T-cell functioning (McNaughton, Smith, Patterson, & Grant, 1990). Conversely, social support, at least among women, was found to be related positively to immune system competence measured in terms of lymphocyte numbers and response to mitogens (Thomas, Goodwin, & Goodwin, 1985). Older individuals are at least as susceptible to stress as younger adults in increases of T-suppressor cells and NK cell numbers, although not in terms of increased NK activity (Nabiloff et al., 1991). Stress also seems linked to the release of beta-endorphin, an opioid peptide released from the pituitary gland that has an analgesic effect. This process may play an important role in mediating the effects of emotions on the immune system (Antoni, 1987).

Apart from changes in the immune system that interact with psychological functioning, the lowered effectiveness of the immune system in older adults has important implications for health. The aging immune system has been linked to increased vulnerability to influenza, infections, cancer, and certain age-associated autoimmune disorders such as diabetes and possibly atherosclerosis and Alzheimer's disease. A less competent immune system can put the elderly individual at higher risk at least to certain forms of cancer and influenza (Ershler, 1990; Miller, 1993). Clearly, the development of severe health problems can have significant effects on the individual's psychological well-being.

There are many ways in which clinicians can take advantage of the potential interactions between identity and immune system functioning. The crossing of an aging threshold in any salient area of functioning, to the extent that it triggers a stressful reaction, can lead to deleterious effects on immune functioning. The individual then becomes more susceptible to immune-related conditions, further accentuating the effects of crossing the threshold. Overassimilation may have some beneficial aspects, in this regard, if the individual is not placed at risk by ignoring important physical changes.

7.01.6 AGE CHANGES IN THE REPRODUCTIVE SYSTEM AND SEXUALITY

7.01.6.1 Female Reproductive System

Throughout her 40s, a woman's reproductive capacity becomes gradually reduced until, by the age of 50–55 years, it ceases altogether. Associated with the ending of the monthly phases of ovulation and menstruation is a diminution of the hormones estrogen and progesterone. Other changes in sexual functioning are related to aging of the tissues in other bodily systems. For example, sagging of the breasts and torso results from decreased skin elasticity. Subcutaneous fat accumulates around the waist, leading to uneven bulges. The appearance and functioning of the genital organs also change after the menopause. The pubic hair on the mons veneris and around the vulva becomes thin and coarser. The labia majora and minora become thinner and wrinkled, the skin in the vulva atrophies, and the surface cells of the vaginal wall become thin, dry, pale, and smooth. The vagina also becomes narrower and shorter. These changes are significant not only for their effects on sexual functioning but for their effects on the woman's enjoyment of sexual intercourse. The older woman may experience discomfort during intercourse due to changes in the vagina and vulva, and the rhythmic contractions of the uterus during orgasm may become painful.

Despite changes in the reproductive system, the conclusion reached by Masters and Johnson (1966) in the mid 1960s is still generally accepted: that older women are limited in their sexual gratification more by their attitudes, values, and accessibility to partners than they are by physical changes involved in aging. The sexual response cycle might be slowed down somewhat in the older woman, but the possibility of achieving orgasm is not reduced.

7.01.6.2 Male Reproductive System

Just as women gradually lose reproductive capacity, men experience a climacteric of sorts in which there is a reduction in the number of viable sperm they produce due to degenerative changes in the seminiferous tubules of the testes (Harman, 1978). With increasing age, men may experience changes in the prostate gland that lead to a reduction of the volume and pressure of semen expelled during ejaculation. Age-related changes also include overgrowth or hypertrophy of the glandular and connective tissue in the parts of the prostate that surround the prostatic urethra. This condition, called benign prostatic hypertrophy, is increasingly prevalent in men past 50 years, rising to an estimated 50% of men 80 years and older. The adjacent penile urethra may become constricted due to this overgrowth, and urinary rentention may ensue. Discomfort and embarrassment may result from difficulties in urination and from the occurrence of involuntary penile erections (Masters & Johnson, 1966). If urinary retention becomes a chronic condition, kidney problems may develop, leading to more serious health threats.

A physiological index of reproductive function that has a decidedly noticeable effect on the older man's sense of his own sexuality is that of penile erectility. It is well documented that older men experience fewer nightly episodes of penile erections compared to younger males (Karacan, Willliam, Thornby, & Salis, 1975). By contrast, there are inconsistencies in the findings regarding increase in penile circumference during erection, with some decreases noted across age groups (Solnick & Birren, 1977), and in other research no age differences are observed (Schiavi & Schreiner-Engel, 1988).

As is true for aging women, there is a general slowing down in aging men of the progression through the human sexual response cycle. Compared to young adult men, orgasm is shorter, involving fewer contractions of the prostate, and ejection of a smaller amount of seminal fluid (Masters & Johnson, 1970). These findings may carry some negative implications for the aging male's sexual relations. However, the gains for the older man's ability to enjoy sexuality are also compelling. He may feel less driven toward the pressure to ejaculate, be able to prolong the period of sensual enjoyment prior to orgasm, and have the control to coordinate his pleasure cycle to correspond more to his female partner.

The main predictor of a man's pattern of sexuality in the earlier years of adulthood is by far the best predictor of his sexuality in old age (George & Weiler, 1985). The sexually active middle-aged man, given good health, has the potential to remain sexually active well into his later years.

7.01.6.3 Clinical Implications

The impact of the menopause depends heavily on how the woman interprets the significance of this transition, which may be met with relief or as a reminder of the inevitability of aging and mortality. Although the aging male is likely to be less preoccupied by his diminished (but not lost) reproductive capacity, he may also find age-related changes to have a negative impact on his enjoyment of sexual relations. He may overaccommodate to aging changes, believing that his masculine prowess has failed. Furthermore, the changes he perceives in his sexual functioning may be interpreted as movement toward deterioration and death. As is true for any man regardless of age, depression, heavy alcohol use, or late-life career pressures may also interfere with the aging male's ability to enjoy sexual relations. Difficulties in adjusting to age changes in the sexual response cycle may present a problem if the partners are unfamiliar with the fact that sexual responsivity naturally becomes altered in later adulthood. The woman may worry that she has lost her orgasmic capacity because it takes her longer to become aroused, excited, and stimulated. The aging male may be at high risk of developing symptoms of secondary (nonphysiological) impotence.

As is true of other sensitive areas of physical functioning, the clinician's responsibility in these areas is to explore in a patient and careful manner the level of concern the individual or couple may have about their functioning. Given cohort differences in attitudes toward sexuality, it is likely that the average older person finds it difficult to discuss specific details about problems in sexual functioning. The general level of anxiety associated with sexual matters can lead to heightened potential for either identity assimilation in the form of denial, or identity accommodation in the form of overreaction. Furthermore, in this area perhaps more so than in any other, it might be necessary for the clinician's own attitudes toward sexuality in the aged to become an area of personal focus and reflection before proceeding to the level of intervention.

7.01.7 THE NERVOUS SYSTEM

7.01.7.1 Central Nervous System

7.01.7.1.1 Age changes

In the central nervous system (CNS), as in the other major organ systems, changes that are due to aging alone are difficult to separate from changes that are the result of disease. Neurofibrillary tangles, amyloid plaques, and granulovacuolar degeneration are all deleterious changes observed in the brains of people with Alzheimer's disease, but are also thought to occur to a lesser extent in normal aging. There are thought to be decreases in acetylcholine levels in the hippocampus and a decline in the number of neurons in the hippocampus, but these are also well-known correlates of Alzheimer's disease. Similarly, decreased amounts of dopamine in the substantia nigra–basal ganglia pathway are regarded as age-related changes but are also observed in Parkinson's disease. It is not clear whether these losses occur normally in later life and are exaggerated in the case of disease, or whether they constitute discrete phenomena.

Keeping in mind the distinction between normal aging and disease processes, there is documentation that areas of the brain are differentially affected by neuronal loss in later adulthood. There appear to be signficiant losses of neurons from the hippocampus, the area affected by Alzheimer's disease and involved in short-term memory (Mouritzen Dam, 1979), and in the cerebellum (Hall, Miller, & Corsellis, 1975), the structure that regulates refined motor patterns. Age losses in these subcortical areas, however, are considered relatively minor compared to losses (as high as 50%) in the areas of the cerebral cortex responsible for processing sensory information, including the primary visual cortex (Devaney & Johnson, 1980), the primary somatosensory region (Henderson, Tomlinson, & Gibson, 1980), and the secondary auditory cortex which loses between one-third and one-half of its total number of neurons over the adult years (Brody, 1955; Henderson et al., 1980). Cells in the motor cortex show comparable losses of 20–50%, and these losses are thought to be a contributing factor to muscular stiffness, slowness, and joint pain (Scheibel, 1982). In contrast to these somewhat dramatic losses of neuron numbers is the pattern of relative stability in the prefrontal region of the cortex (Higatsberger, Budka, & Bernheimer, 1982; Huttenlocher, 1979). The stability of neurons in these regions of the cortex is important, as these are the parts of the nervous system involved in judgment, abstract thinking, and the ability to plan.

7.01.7.1.2 Clinical implications

Aging of the CNS has direct effects on a variety of sensory, motor, and cognitive capacities and behaviors including perception, short-term memory, fine motor coordination, and large muscle control. However, the view of the aging brain as a degenerating system does not take into account what is known about the compensatory processes of redundancy and plasticity (Buell & Coleman, 1979; Flood, Buell, Horowitz, & Coleman 1987; Flood & Coleman, 1988). The impact of these processes is most likely to occur in the association areas of the cerebral cortex, where higher order abstract thinking processes are mediated. In fact, these abilities may improve in the later years of adulthood as the individual stores more experiences into the long-term memory association areas on which decisions and judgment are based.

Furthermore, it is important to emphasize once again that there is tremendous variation among adults in the ways in which the aging process (apart from disease) affects CNS functions. Some of these functions in later adulthood are, in addition, affected by the extent to which the individual developed and used them during the adult years. Exposure to stimulation in the areas of perception, cognition, and motor activity through mental and physical exercise can influence, throughout old age, the degree to which the abilities that depend on these functions are maintained. The importance placed by the individual on mental and physical abilities contributes further to the personal meaning of age changes in any of these areas and whether a threshold is crossed in this domain.

Clinically, of course, it is essential that those working with the elderly are knowledgable about and sensitive to the symptoms of Alzheimer's disease, particularly as these differ from other forms of cognitive impairment or depression. Clinicians must also be aware of the extent to which elders fear the onset of Alzheimer's disease and interpret even minor changes in cognitive functions as signs that the end is in sight. Overaccommodation in the cognitive domain can prove extremely damaging to the older adult's identity as a sentient being, and vigilance in this area toward any signs of deterioration is likely to be high. This risk is perhaps greater than that of overassimilation, except when an individual who really does have early symptoms of Alzheimer's disease denies the memory or judgment problems that are interfering with daily life activities. Media attention to Alzheimer's disease, although positive in the sense of educating the public, may backfire in this domain, leaving many elders (and clinicians) with the erroneous impression that the disease is an inevitable outcome of the aging process.

7.01.7.2 Autonomic Nervous System

7.01.7.2.1 Age changes

Although in many aspects the autonomic nervous system (ANS) operates without significant age-related alterations throughout adulthood, there are important effects of aging on two functions served by the ANS that have a considerable impact on the individual's daily life: bodily temperature control and sleep patterns.

It is well established both through population health statistics and experimental studies that individuals over the age of 65 years have impaired adaptive responses to extremely hot and cold outside temperatures. The diminished response of the elderly to cold appears to be due to a diminished perception that the core body temperature is low, and to an impaired ability to raise core temperature when the body's peripheral temperature becomes lowered. Responses to extremes of heat are impaired due to decreased secretion by the sweat glands in the skin.

In the area of sleep regulation, older adults spend more time in bed relative to time spent asleep due to longer time taken to fall asleep, more periods of wakefulness during the night, and time spent lying awake before arising in the morning. The primary causes of sleep disturbance include sleep apnea (Ancoli-Israel & Kripke, 1991), periodic leg movements, heartburn, and frequent needs to urinate (Freidman, Bliwise, Tauke, & Salom, 1992). EEG sleep patterns show some corresponding age alterations, including a rise in Stage 1 sleep (drowsiness without actual sleep), and a large decrease in Stage 4 sleep (slow wave or heavy sleep). By the 60s and 70s, REM sleep starts to diminish as well, as do the observable behaviors associated with REM sleep.

7.01.7.2.2 Clinical implications

The knowledge gained from personal experience and media exposure that one's aging body is less adaptable to outside temperatures means that adults living in geographical areas with cold winters will restrict their outdoor activities and spend more time at home. They will also limit their outdoor exposure during the summer. Overaccommodation to the threshold experience of aging in this regard can reduce the well-being of the older adult who may feel forced to remain indoors unnecessarily even on days when the temperature would not pose a threat. Conversely, overassimilation can have deleterious effects when the older adult does not take precautions during physical exertion or when venturing out into a very hot or very cold day. Fortunately, the changes in the ANS described here occur gradually over a period of years. Consequently, there are many opportunities for the individual to learn to adjust to the effects of aging and find new behavioral accommodations as these become necessary. Clinicians can play an important role in this process through helping their clients establish a favorable balance between overreaction and denial of aging-related changes in temperature regulation.

Similarly, the function of sleep in everyday life is clearly crucial to the individual's sense of well-being and there is a strong relationship between quality of sleep and psychological symptoms (Bliwise, 1992). Given the sensitivity of sleep patterns to psychological distress, this is an area where identity–age change relationships might very profitably be discussed in therapy. Older adults who overreact through accommodation to slight sleep change patterns are perhaps the ones fated to experience the most significant changes in their ability to get a good night's sleep. They need to be given the information that a night's sleep need not consist of more than seven hours and that the longer the time spent in bed awake, the harder it will be for the individual to develop a normal nightly rhythm based on this more realistic sleep requirement. The clinician can also advise the client to avoid daytime naps and to follow good sleep habits.

7.01.8 AGE CHANGES IN SENSORY FUNCTIONING

7.01.8.1 Vision

7.01.8.1.1 Age changes

The optical structures in the eye focus light reflected from stimuli in the environment onto the sensory cells of the retina. Many, if not all, age effects on visual functions can be explained in terms of the effects of normal aging on these optical structures (Scheie & Albert, 1977). One effect of these changes is to reduce the clarity of the visual image reaching the retina. The image reaching the retina is clouded by increased density and opacity of the lens (Spector, 1982), and by the formation of opacities in the vitreous. Adding to these structural changes are changes in the retina itself, including declines in the number of photoreceptors (rods and cones) (Ordy, Brizzee, & Johnson, 1982), the accumulation of debris in the outermost layer of the retina (Marmor, 1980), and detachment of the vitreous from the surface of the retina (Scheie & Albert). The amount of light reaching the retina is diminished by the condition known as senile meiosis, reduction in the size of the pupil due to atrophy of the iris dilator (Carter, 1982).

In addition to these changes that reduce the quality of the visual image on the retina are changes in the lens that decrease its capacity to accommodate to necessary changes in focus as objects move closer or further away from the individual. In addition to becoming denser, the lens fibers become harder and less elastic (Paterson, 1979). The loss of accommodative power of the lens due to these changes is referred to as presbyopia, and it is a condition that typically requires correction between the ages of 40 to 50 years. By the age of 60 years, the lens is completely incapable of accommodating to focus on objects at close distance (Moses, 1981). The lens also becomes yellowed due to an accumulation of yellow pigment, and as a result the older adult is less able to discriminate colors in the green–blue–violet end of the spectrum (Mancil & Owsley, 1988).

These structural changes in the eye result in altered refractive power leading to hypermetropia (farsightedness), a condition that cannot be compensated by greater focusing of the lens due to presbyopia. There is a reduction in visual acuity which is especially severe at low levels of illumination, such as driving at night, and when tracking moving objects (Panek, Barrett, Sterns, & Alexander, 1977; Richards, 1977). Dark adaptation is reduced so that older adults have greater difficulty adjusting to movement from bright to dim lighting, and have lower absolute levels of ability to see in the dark (McFarland, Dorney, Warren, & Ward, 1960). There is also a reduction in the individual's ability to react to scotomatic glare, or sudden exposure to bright light, such as a flashbulb or the headlights of an oncoming car at night (Wolf & Gardiner, 1965). Although there are only minimal data on depth perception, there appear to be reductions in stereopsis, the perception of three-dimensional space resulting from the varying input that reaches the two eyes (Bell, Wolf, & Bernholz, 1972).

7.01.8.1.2 Clinical implications

Given the centrality of vision to many activities, changes with age in the eye's basic functions can have profound psychological effects. One set of problems relates to discomfort and frustration caused by poorer acuity and reduced focusing power. Even if visual problems can be corrected with glasses or contact lenses, there are residual symptoms that may remain in special circumstances such as when the individual overworks or is trying to read small print. Older adults report experiencing as problems sensitivity to glare, difficulty seeing in dim light, and problems focusing on near objects (Kosnik, Winslow, Kline, Rasinski, & Sekuler, 1988). When these problems become more persistent, they may propel the individual toward a threshold experience.

Presbyopia, although reached after a gradual process of changes in the lens, is often perceived with relative suddenness by the individual (Carter, 1982). The immediacy of this apparent change, given the association that many people have between presbyopia and the infirmities of age, makes it more likely that the change will be negatively interpreted. The necessity of wearing bifocals adds the complication of requiring the individual to adjust to a new and awkward way of using corrective lenses.

Difficulties in depth perception and dark adaptation can heighten the older adult's vulnerability to falls (Felson et al., 1989; McMurdo & Gaskell, 1991), adding to the tendency to overaccommodate due to concerns about bone fragility and stiffness. Visual losses can increase the individual's dependence on others (Hakkinen, 1984), interfering with the older adult's ability to complete basic tasks of living such as driving, housekeeping, grocery shopping, and food preparation (Branch, Horowitz, & Carr 1989; Rudberg, Furner, Dunn, & Cassel, 1993). Apart from the practical implications, these changes can further erode the individual's physical identity.

Finally, age changes in vision can detract from the older adult's ability to enjoy leisure and esthetic activities, and even to choose appropriate clothes (Morgan, 1988). Although not a serious problem compared with difficulties in driving, such changes can be annoying and perhaps lead to derision by others. Other leisure activities involving the perception of color and fine detail may be interfered with, such as needlepoint, gardening, word puzzles, and painting. Loss of the ability to enjoy these activities add further to the threshold experience which may be crossed each time a new visual function is altered or affected by aging.

A number of changes in visual functioning can, fortunately, be compensated by corrective lenses, increases in the ambient lighting, and efforts to reduce glare and heighten contrast between light and dark. Corrective surgery for cataracts can have a number of widespread positive effects on daily life (Brenner, Curbow, Javitt, Legro, & Sommer, 1993). The success of these efforts depends on the individual's willingness to persist in trying new ideas when the old methods no longer work, a process that might be enhanced through clinical intervention. Nevertheless, a point may be reached within each sphere of functioning in which the individual's range of movement becomes compromised. Furthermore, the situation may not permit compensation, as is true for night driving. Again, clinical sensitivity plays an important role in helping the older person accommodate to this age-related change in a valued function.

7.01.8.2 Hearing

7.01.8.2.1 Age changes

Presbycusis, the general term used to refer to age-related hearing loss, actually includes several specific subtypes reflecting different changes in the auditory structures. The most common form of hearing loss reduces sensitivity to high-frequency tones earlier and more severely than sensitivity to low-frequency tones (Van-Rooij & Plomp, 1990). The loss of high-frequency pitch perception is particularly pronounced in men (Lebo & Reddell, 1972). Speech perception is affected both by the various forms of presbycusis operating at the sensory level and by changes in the central processing of auditory information at the level of the brain stem and above (Van-Rooij & Plomp, 1992). Translated into speech discrimination, this means that older adults suffering from presbycusis have greater dificulty perceiving sibilants, such as the "s" in plurals.

In addition to the effects of aging on speech discrimination due to the loss of high-pitched tone sensitivity are the effects of aging on the ability to hear when there is interference or distraction. Age effects begin to appear even as early as 40 years in the understanding of sentences under a variety of distorting conditions, particularly when the speech signal is interrupted (Bergman et al., 1976). Other conditions known to impair speech perception in the aged include higher rate of presentation, deletion of parts of the message, competition from background noise or competing messages, and reverberation (Neils, Newman, Hill, & Weiler, 1991).

7.01.8.2.2 Clinical implications

Hearing deficits interfere greatly with interpersonal communication, leading to strained relationships and greater caution by the elderly in an attempt to avoid making inappropriate responses to uncertain auditory signals. They also reduce the older person's ability to hear noises such as a siren or a door knock (Gatehouse, 1990). These changes are almost impossible to avoid noticing (Slawinski, Hartel, & Kline, 1993), and it is perhaps for this reason that hearing loss forms a threshold for a large percentage of individuals over the age of 70 years and particularly those in their 80s (Whitbourne, 1996b). There is evidence linking hearing loss to impaired physical functioning (Bess, Lichenstein, Logan, & Burger, 1989) and psychological difficulties including loneliness (Christian, Dluky, & O'Neill, 1989) and depression (Kalayam, Alexopoulos, Merrell, & Young, 1991).

Those who interact with hearing-impaired elders can benefit from learning ways to communicate that lessen the impact of age-related changes (Slawinski et al., 1993). Modulating one's tone of voice, particularly for women, so that it is not too high, and avoiding distractions or interference, can be important aids to communicating clearly with older adults. The clinician can also use observations of the elderly client's reaction to communication difficulties in therapy as the basis for clinical recommendations. For example, identity accommodation can be encouraged in a client who denies the existence of an obvious hearing deficit; a client who has no apparent hearing deficit but appears preoccupied with this particular threshold will need to develop a more balanced approach to this area of functioning.

7.01.8.3 Balance

7.01.8.3.1 Age changes

The effects of age on the vestibular organs involve the losses occurring in the sensory structures as well as in the pathways to the higher levels of the nervous system (Johnsson & Hawkins, 1972; Rosenhall & Rubin, 1973). There are also reductions in the numbers of Purkinje cells in the cerebellum (Hall, Miller, & Corsellis, 1975) which could result in lowering of reflexive abilities to adjust the body's posture to changes in position. The apparent results of these changes in the vestibular system are increased dizziness and vertigo in older adults (Toglia, 1975). Not only are these sensations unpleasant, but they can increase the likelihood of accidental falls.

However, structural changes in the vestibular system do not account entirely for the phenomena of dizziness and vertigo. There is considerable plasticity within the vestibular system so that loss of receptor cells may be compensated by the activity of structures in other sensory systems (Teasdale, Stelmach, & Breunig, 1991). For example, the positional receptors in the somesthetic system appear to be less vulnerable to aging effects (Babin & Harker, 1982). Conversely, the loss of information from more than one sensory system, such as the visual and the vestibular systems, can make it difficult for the older person to compensate successfully (Manchester, Woollacott, Zederbauer-Hylton, & Marin, 1989). Slowing of central integrative processes responsible for maintaining postural stability can also contribute to increased likelihood of falls. It may take longer for the older adult to integrate information from the vestibular, visual, and somesthetic systems, resulting in less efficient control of posture under changing body positions (Teasdale et al., 1991).

7.01.8.3.2 Clinical implications

Balance is an essential element of moving about effectively in the environment. Aging of the vestibular system brings with it the potential for the individual to feel insecure in moving, particularly under conditions that are less than ideal, such as sloping, steep, or uneven surfaces. Fear of falling due to other changes in mobility can increase the individual's anxiety and perhaps exacerbate any true deficits in vestibular functioning. Conversely, individuals who through identity assimilation ignore signs of dizziness and vertigo may place themselves in danger as they may not be able to avoid a fall when and if they do lose their balance.

There are social consequences of aging of the vestibular system as well. The older adult who experiences dizziness and vertigo may fear appearing disoriented in front of other people, perhaps due to concern about appearing intoxicated or mentally confused. A desire to avoid such embarrassment may lead the individual to avoid leaving the home, creating unnecessary limitations on social opportunities.

Nevertheless, it remains the case that vestibular dysfunction is not an inevitable consequence of the aging of the vestibular system. Furthermore, compensation is possible if the individual is able to adapt other sensory systems to make up for vestibular losses (Lord, Clark, & Webster, 1991). Such compensation seems to be more likely to occur if the individual is able to react in a balanced manner to the experience of dizziness or vertigo. The clinician can encourage the older client to benefit from the coping strategies that involve seeking other cues, such as those provided by the somesthetic system. During episodes of dizziness or vertigo, the individual can learn to pay attention to stance and bodily orientation, learning to judge the position of the lower body limbs to make better use of feedback in adjusting posture (Hu & Woollacott 1994a, 1994b; Meenwsen, Sawicki, & Stelmack, 1993).

7.01.9 CONCLUDING OBSERVATIONS

In this chapter, a number of age-related changes have been described that occur throughout the body's organ systems and sensory processes. The multiple threshold model postulates that an individual's reactions to these changes varies according to how central the area of functioning is to identity as well as according to whether the individual approaches the age change through identity assimilation or accommodation. The problems involved in using overaccommodation or overassimilation have also been described.

Clinicians can learn from this approach to pay greater attention to the subtle and not-so-subtle cues that elderly clients give regarding the concerns they have about the functioning of their bodies. Assessments of older adults in particular, but also adults in middle age, can benefit from detailed history-taking not only of "health" problems but also of the individual's appraisal of and reaction to normal age-related changes in physical and sensory functioning. Such questions are particularly important for the clinician to raise because older clients are likely to be reticent in raising delicate and potentially embarrassing concerns about how their bodies are aging. Clinicians should also be encouraged to narrow their questions as specifically as possible to particular areas of functioning rather than explore issues related to "aging" or "health" in general. Such general questions are easily disregarded, as most older individuals tend to report that they are in good health even if they have a number of specific age-related limitations (Heidrich & Ryff, 1993).

Authors in the field of clinical geropsychology have, for years, advised clinicians to examine their own age biases and attitudes toward elderly people. This admonition definitely applies with regard to the aging of the body, as clinicians who are a product of Western culture have undoubtedly acquired a number of negative attitudes toward the loss of functioning that is so generally associated with old age. Less well recognized, though, is the need for the clinician to examine his or her own personal aging thresholds. To the extent that the clinician harbors concern about how aging will affect his or her own appearance, physical mobility, or cognitive functioning, that clinician might approach in a defensive and unhelpful way an elderly client with apparent deficits in that area.

Finally, it is crucial for clinicians to recognize the independence, autonomy, and vitality of spirit seen in many elders, even those with severe losses or age-related limitations. The older adults seen in therapy are "survivors" who have managed to reach this point in life through resiliency and an indefatigible ability to adapt to change. They are coping daily with physical changes that would daunt individuals many years their junior. Therapists who condescend to the elderly or patronize them (perhaps as a result of their own fears of aging) are missing important treatment opportunities as well as important opportunities to learn from the wisdom of their clients.

7.01.10 REFERENCES

Adrian, M. J. (1981). Flexibility in the aging adult. In E. L. Smith & R. C. Serfass (Eds.), *Exercise and aging: The scientific basis* (pp. 45–58). Hillsdale, NJ: Enslow.

Ancoli-Israel, S., & Kripke, D. F. (1991). Prevalent sleep problems in the aged. *Biofeedback and Self Regulation, 16,* 349–359.

Antoni, M. H. (1987). Neuroendocrine influences in psychoimmunology and neoplasia: A review. *Psychology and Health, 1,* 3–24.

Babin, R. W., & Harker, L. A. (1982). The vestibular system in the elderly. *Otolaryngolic Clinics of North America, 15,* 387–393.

Balin, A. K., & Pratt, L. A. (1989). Physiological consequences of human skin aging. *Cutis, 43,* 431–436.

Bell, B., Wolf, E., & Bernholtz, C. D. (1972). Depth perception as a function of age. *Aging and Human Development, 3,* 77–81.

Bergman, M., Blumenfeld, V. G., Cascardo, D., Dash, B., Levitt, H., & Margulies, M. K. (1976). Age-related decrement in hearing for speech: Sampling and longitudinal studies. *Journal of Gerontology, 31,* 533–538.

Bess, F. H., Lichtenstein, M. J., Logan, S. A., & Burger, M. C. (1989). Hearing impairment as a determinant of function in the elderly. *Journal of the American Geriatrics Society, 37,* 123–128.

Bliwise, N. G. (1992). Factors related to sleep quality in healthy elderly women. *Psychology and Aging, 7,* 83–88.

Bloom, E. T. (1994). Natural killer cells, lymphokine-activated killer cells, and cytolytic T lymphocytes: Compartmentalization of age-related changes in cytolytic lymphocytes? *Journal of Gerontology: Biological Sciences, 49*, B85–92.

Blumenthal, J. A., Emery, G. F., Madden, D. J., George, L. K., Coleman, R. E., Riddle, M. W., McKee, D. C., Reasoner, J., & Williams, R. S. (1989). Cardiovascular and behavioral effects of aerobic exercise training in healthy older men and women. *Journal of Gerontology: Medical Sciences, 44,* M147–157.

Branch, L. G., Horowitz, A., & Carr, C. (1989). The implications for everyday life of incident self-reported visual decline among people over age 65 living in the community. *Gerontologist, 29,* 359–365.

Brandtstadter, J., & Renner, G. (1990). Tenacious goal pursuit and flexible goal adjustment: Explication and age-related analysis of assimilative and accommodative strategies in coping. *Psychology and Aging, 5,* 58–67.

Brandstetter, R. D., & Kazemi, H. (1983). Aging and the respiratory system. *Medical Clinics of North America, 67,* 419–431.

Brenner, M. H., Curbow, B., Javitt, J. C., Legro, M. W., & Sommer, A. (1993). Vision change and quality of life in the elderly: Response to cataract surgery and treatment of other chronic ocular conditions. *Archives of Ophthalmology, 5,* 680–685.

Brody, H. (1955). Organization of cerebral cortex. III. A study of aging in the human cerebral cortex. *Journal of Comparative Neurology, 102*, 511–556.

Brooks, G. A., & Fahey, T. D. (1984). *Exercise physiology: Human bioenergetics and its applications.* New York: Wiley.

Buell, S. J., & Coleman, P. D. (1979). Dendritic growth in the aged human brain and failure of growth in senile dementia. *Science, 206,* 854–856.

Burgio, K. L., & Engel, B. T. (1990). Biofeedback-assisted behavioral training for elderly men and women. *Journal of the American Geriatrics Society, 38,* 338–340.

Burns, P. A., Pranikoff, K., Nochajski, T. H., Hadley, E. C., Levy, K. J., & Ory, M. G. (1993). A comparison of effectiveness of biofeedback and pelvic muscle exercise treatment of stress incontinence in older community-dwelling women. *Journal of Gerontology: Medical Sciences, 38,* M167–174.

Carter, J. H. (1982). The effects of aging on selected visual functions: Color vision, glare sensitivity, field of vision, and accommodation. In R. Sekuler, D. Kline, & K. Dismukes (Eds.), *Aging and human visual function* . New York: Alan R. Liss.

Carter, J. H. (1982). Predicting visual responses to increasing age. *Journal of the American Optometric Association, 53,* 31–36.

Chien, S., Peng, M. T., Chen, K. P., Huang, T. F., Chang, C., & Fang, H. S. (1975). Longitudinal studies on adipose tissue and its distribution in human subjects. *Journal of Applied Physiology, 39,* 825–830.

Chodzko-Zajko, W. J., Schuler, P., Solomon, J., Heinl, B., & Ellis, N. R. (1992). The influence of physical fitness on automatic and effortful memory changes in aging. *International Journal of Aging and Human Development, 35,* 265–285.

Christian, E., Dluhy, N., & O'Neill, R. (1989). Sounds of silence: Coping with hearing loss and loneliness. *Journal of Gerontological Nursing, 15*(11), 4–9.

Costello, R. B., & Moser-Veillon, P. B. (1992). A review of magnesium intake in the elderly. A cause for concern? *Magnesium Research, 5,* 61–67.

Dalsky, G. P., Stocke, K. S., Ehsani, A. A., Slatopolsky, E., Lee, W. C., & Birge, S. J. (1988). Weight-bearing exercise, training, and lumbar bone mineral content in post-menopausal women. *Annals of Internal Medicine, 108,* 824–828.

Dargent, P., & Breart, G. (1993). Epidemiology and risk factors of osteoporosis. *Current Opinions in Rheumatology, 5,* 339–345.

D'Errico, A., Scarani, P., Colosimo, E., Spina, M., Grigioni, W. F., & Mancini, A. M. (1989). Changes in the alveolar connective tissue of the ageing lung. An immunohistochemical study. *Archives of Pathology Anatomy and Histopathology, 415,* 137–144.

Devaney, K. O., & Johnson, H. A. (1980). Neuron loss in the aging visual cortex in man. *Journal of Gerontology, 35,* 836–841.

Diokno, A. C., Brock, B. M., Brown, M. B., & Herzog, A. R. (1986). Prevalence of urinary incontinence and other urological syptoms in the noninstitutionalized elderly. *Journal of Urology, 136,* 1022–1025.

Downton, J. H., & Andrews, K. (1990). Postural disturbance and psychological symptoms amongst elderly people living at home. *International Journal of Geriatric Psychiatry, 5,* 93–98.

Ehsani, A. A., Ogawa, T., Miller, T. R., Spina, R. J., & Jilka, S. M. (1991). Exercise training improves left ventricular systolic function in older men. *Circulation, 83,* 96–103.

Ellis, F. P., Shukla, K. K., Cohn, S. H., & Pierson, R. N. (1974). A predictor for total-body potassium in man based on height, weight, sex, and age: Application in metabolic disorders. *Journal of Laboratory and Clinical Medicine, 83,* 716–727.

Ershler, W. B. (1990). Influenza and aging. In A. L. Goldstein (Ed.), *Biomedical advances in aging* (pp. 513–521). New York: Plenum.

Felson, D. T., Anderson, J. J., Hannan, M. T., Milton, R. C., Wilson, P. W., & Kiel, D. P. (1989). Impaired vision and hip fracture: The Framingham Study. *Journal of the American Geriatrics Society, 37,* 495–500.

Fenske, N. A., & Albers, S. E. (1990). Cosmetic modalities for aging skin: what to tell patients. *Geriatrics, 45*(9), 59–67.

Fiatrone, M. A., & Evans, W. J. (1993). The etiology and reversibility of muscle dysfunction in the aged. *Journal of Gerontology, 48 (Special issue)*, 77–83.

Fiatrone, M. A., Marks, E. C., Ryan, N. D., Meredith, C. N., Lipsitz, L. A., & Evans, W. J. (1990). High-intensity strength training in nonagenarians. Effects on skeletal muscle. *Journal of the American Medical Association 263,* 3029–3034.

Fischer, J., & Johnson, M. A. (1990). Low body weight and weight loss in the aged. *Journal of the American Dietetic Association, 90,* 1697–1706.

Flood, D. G., Buell, S. J., Horwitz, G. J., & Coleman, P. D. (1987). Dendritic extent in human dentate gyrus granule cells in normal aging and senile dementia. *Brain Research, 402,* 205–216.

Flood, D. G., & Coleman, P. D. (1988). Cell type heterogeneity of changes in dendritic extent in the hippocampal region of the human brain in normal aging and in Alzheimer's disease. In T. L. Petit & G. O. Ivy (Eds.), *Neural plasticity: A lifespan approach* (pp. 265–281). New York: Alan R. Liss.

Foster, V. L., Hume, G. J. E., Byrnes, W. C., Dickinson, A. L., & Chatfield, S. J. (1989). Endurance training for elderly women: Moderate vs low intensity. *Journal of Gerontology: Medical Sciences, 44,* M184–188.

Friedman, L. F., Bliwise, D. L., Tanke, E. D., & Salom, S. R. (1992). A survey of self-reported poor sleep and associated factors in older individuals. *Behavior, Health, and Aging, 2*(1), 13–20.

Garn, S. M. (1975). Bone loss and aging. In R. Goldman & M. Rockstein (Eds.), *The physiology and pathology of aging* (pp. 39–57). New York: Academic Press.

Gatehouse, S. (1990). Determinants of self-reported disability in older subjects. *Ear and Hearing, 11* (Suppl.), 57S–65S.

George, L. K., & Weiler, S. J. (1985). Sexuality in middle and late life. In E. Palmore, J. Nowlin, E. Busse, I. Siegler & G. Maddox (Eds.), *Normal aging III* . Durham, NC: Duke University Press.

Gilchrest, B. A. (1989). Skin aging and photoaging: an overview. *Journal of the American Academy of Dermatology, 21*(3, Pt. 2), 610–613.

Globerson, A., Eren, R., Abel, L., & Ben-Menahem, D. (1990). Developmental aspects of T lymphocytes in aging. In A. L. Goldstein (Ed.), *Biomedical advances in aging* (pp. 363–373). New York: Plenum.

Govindasamy, D., Paterson, D. H., Poulin, M., & Cunningham, D. A. (1988). The time course of cardiorespiratory adaptations in elderly men. *Canadian Journal of Sport Sciences, 13,* 53–54.

Grove, G. L. (1989). Physiologic changes in older skin. *Clinics in Geriatric Medicine, 5,* 115–125.

Hagberg, J. M., Graves, J. E., Limacher, L., Woods, D. R., Leggett, S. H., Cononie, C., Gruber, J. J., & Pollock, M.L. (1989). Cardiovascular responses of 70– to 79–yr-old men and women to exercise training. *Journal of Applied Physiology, 66,* 2589–2594.

Hakkinen, L. (1984). Vision in the elderly and its use in the social environment. *Scandanavian Journal of Social Medicine, 35,* 5–60.

Hall, T. C., Miller, A. K. H., & Corsellis, J. A. N. (1975). Variations in the human Purkinje cell population according to age and sex. *Neuropathology and Applied Neurobiology, 1,* 267–292.

Hamdorf, P. A., Withers, R. T., Penhall, R. K., & Haslam, M. V. (1992). Physical training effects on the fitness and habitual activity patterns of elderly women. *Archives of Physical Medicine and Rehabilitation, 73,* 603–608.

Harman, S. M. (1978). Clinical aspects of the male reproductive system. In E. L. Schneider (Ed.), *Aging: Vol. 4. The aging reproductive system.* New York: Raven Press.

Heidrich, S. M., & Ryff, C. D. (1993). The role of social comparisons processes in the psychological adaptation of elderly adults. *Journal of Gerontology: Psychological Sciences, 48,* 127–136.

Heiss, G., Tamir, I., Davis, C. E., Tyroler, H. A., Rifkind, B. M., Schonfeld, G., Jacobs, D., & Frantz, I. D. (1980). Lipoprotein-cholesterol distributions in selected North American populations: The Lipid Research Clinics Program Prevalence Study. *Circulation, 61,* 302–315.

Henderson, G., Tomlinson, B., & Gibson, P. H. (1980). Cell counts in human cerebral cortex in normal adults throughout life using an image analysing computer. *Journal of the Neurological Sciences, 46,* 113–136.

Hermanson, B., Omenn, G. S., & Kronmal, R. A. (1988). Beneficial six-year outcome of smoking cessation in older men and women with coronary artery disease. *New England Journal of Medicine, 24,* 1365–1369.

Herzog, A. R., Diokno, A. C., Brown, M. B., Normolle, D. P., & Brock, B. M. (1990). Two–year incidence, remission, and change patterns of urinary incontinence in noninstitutionalized older adults. *Journal of Gerontology: Medical Sciences, 45,* M67–74.

Higatsberger, M. R., Budka, H., & Bernheimer, H. (1982). Neurochemical investigations of aged human brain cortex. In S. Hoyer (Ed.), *The aging brain: Physiological and pathophysiological aspects.* New York: Springer.

Hill, R. D., Storandt, M., & Malley, M. (1993). The impact of long-term exercise training on psychological function in older adults. *Journal of Gerontology: Psychological Sciences, 48,* P12–17.

Holt, P. R. (1991). General perspectives on the aged gut. *Clinics in Geriatric Medicine, 7,* 185–189.

Hu, M.-H., & Woollacott, M. H. (1994a). Multisensory training of standing balance in older adults: I. Postural stability and one-leg stance balance. *Journal of Gerontology: Medical Sciences, 49,* M52–61.

Hu, M.-H., & Woollacott, M. H. (1994b). Multisensory training of standing balance in older adults. II. Kinematic and electromyographic postural responses. *Journal of Gerontology: Medical Sciences, 49,* M62–71.

Hughes, S. L., Edelman, P. L., Singer, R. H., & Chang, R. W. (1993). Joint impairment and self-reported disability in elderly persons. *Journal of Gerontology: Social Sciences, 48,* S84–92.

Hunskaar, S., & Vinsnes, A. (1991). The quality of life in women with urinary incontinence as measured by the Sickness Impact Profile. *Journal of the American Geriatrics Society, 39,* 378–382.

Huttenlocher, P. (1979). Synaptic density in human frontal cortex—developmental changes and effects of aging. *Brain Research, 163,* 195–205.

Janelli, L. M. (1993). Are there body image differences between older men and women? *Western Journal of Nursing Research, 15,* 327–339.

Jirovec, M. M. (1991). The impact of daily exercise on the mobility, balance and urine control of cognitively impaired nursing home residents. *International Journal of Nursing Studies, 28,* 145–151.

Johnsson, L. G., & Hawkins, J. E., Jr. (1972). Sensory and neural degeneration with aging, as seen in microdissections of the human inner ear. *Annals of Otolaryngology, Rhinology, and Laryngology, 81,* 179–193.

Kalayam, B., Alexopoulos, G. S., Merrell, H. B., & Young, R. C. (1991). Patterns of hearing loss and psychiatric morbidity in elderly patients attending a hearing clinic. *International Journal of Geriatric Psychiatry, 6,* 131–136.

Kallman, D. A., Plato, C. C., & Tobin, J. D. (1990). The role of muscle loss in the age–related decline of grip strength: Cross-sectional and longitudinal perspectives. *Journal of Gerontology: Medical Sciences, 45,* M82–88.

Karacan, I., Williams, R. L., Thornby, J. I., & Salis, P. J. (1975). Sleep-related penile tumescence as a function of age. *American Journal of Psychiatry, 132,* 932–937.

Kelly, P. J., Nguyen, T., Hopper, J., Pocock, N., Sambrook, P., & Eisman, J. (1993). Changes in axial bone density with age: a twin study. *Journal of Bone Mineral Research, 8,* 11–17.

Kenney, A. R. (1989). *Physiology of aging* (2nd ed.). Chicago, IL: Yearbook Medical.

Kiecolt, K. J. (1994). Stress and the decision to change oneself: A theoretical model. *Social Psychology Quarterly, 57,* 49–63.

Kitzman, D. W., & Edwards, W. D. (1990). Age-related changes in the anatomy of the normal human heart. *Journal of Gerontology: Medical Sciences, 45,* M33–39.

Kitzman, D. W., Scholz, D. G., Hagen, P. T., Ilstrup, D. M., & Edwards, W. D. (1988). Age-related changes in normal human hearts during the first ten decades. Part II (Maturity): A quantitative anatomic study of 765 specimens from subjects 20 to 99 years old. *Mayo Clinic Proceedings, 63,* 137–146.

Kleinsmith, D. M., & Perricone, N. V. (1989). Common skin problems in the elderly. *Clinics in Geriatric Medicine, 5,* 189–211.

Kligman, A. M. (1989). Psychological aspects of skin disorders in the elderly. *Cutis, 43,* 498–501.

Kligman, A. M., Grove, G. L., & Balin, A. K. (1985). Aging of human skin. In C. E. Finch & E. L. Schneider (Eds.), *Handbook of the biology of aging* (2nd ed.). New York: Van Nostrand Reinhold.

Kosnik, W., Winslow, L., Kline, D., Rasinski, K., & Sekuler, R. (1988). Visual changes in daily life throughout adulthood. *Journal of Gerontology: Psychological Sciences, 43,* P63–70.

Krall, E. A., & Dawson-Hughes, B. (1993). Heritable and life-style determinants of bone mineral density. *Journal of Bone Mineral Research, 8,* 1–9.

Kurban, R. S., & Bhawan, J. (1990). Histologic changes in skin associated with aging. *Journal of Dermatology and Surgical Oncology, 16,* 908–914.

Kutza, J., Kaye, D., & Murasko, D. M. (1995). Basal natural killer cell activity of young versus elderly humans. *Journal of Gerontology: Biological Sciences, 50A,* B110–116.

Labouvie-Vief, G., Hakim-Larson, J., & Hobart, C. J. (1987). Age, ego level, and the life-span development of coping and defense processes. *Psychology and Aging, 2,* 286–293.

Lakatta, E. G. (1987). Why cardiovascular function may decline with age. *Geriatrics, 42,* 84–94.

Lamy, P. P. (1988). Actions of alcohol and drugs in older people. *Generations, 12*(4), 9–13.

Lavker, R. M., Kwong, F., & Kligman, A. M. (1980). Changes in skin surface patterns with age. *Journal of Gerontology, 35,* 348–354.

Lebo, C. P., & Reddell, R. C. (1972). The presbycusis component in occupational hearing loss. *Laryngoscope, 82,* 1399–1409.

Lord, S. R., Clark, R. D., & Webster, I. W. (1991). Postural stability and associated physiological factors in a population of aged persons. *Journal of Gerontology: Medical Sciences, 46,* M69–76.

Lord, S. R., Clark, R. D., & Webster, I. W. (1991). Physiological factors associated with falls in an elderly population. *Journal of the American Geriatrics Society, 39,* 1194–1200.

Mahler, D. A., Rossiello, R. A., & Loke, J. (1986). The aging lung. *Clinics in Geriatric Medicine, 2,* 215–225.

Manchester, D., Woollacott, M., Zederbauer-Hylton, N., & Marin, O. (1989). Visual, vestibular and somatosensory contributions to balance control in the older adult. *Journal of Gerontology: Medical Sciences, 44,* M118–127.

Mancil, G. L., & Owsley, C. (1988). 'Vision through my aging eyes' revisited. *Journal of the American Optometric Association, 59,* 288–294.

Marmor, M. F. (1980). Clinical physiology of the retina. In G. A. Reyman, D. R. Sanders, & M. F. Goldberg (Eds.), *Principles and practice of opthalmology* (Vol. 2, pp. 823–856). Philadelphia: Saunders.

Masters, W. H., & Johnson, V. E. (1966). *Human sexual response*. Boston, MA: Little, Brown.

Masters, W. H., & Johnson, V. E. (1970). *Human sexual inadequacy*. Boston, MA: Little, Brown.

Mazess, R. B. (1982). On aging bone loss. *Clinical Orthopaedics and Related Research, 165,* 239–252.

McArdle, W. D., Katch, F. I., & Katch, V. L. (1991). *Exercise physiology: Energy, nutrition, and human performance* (3rd ed.). Philadelphia: Lea & Febiger.

McAuley, E., Lox, C., & Duncan, T. E. (1993). Long-term maintenance of exercise, self-efficacy, and physiological change in older adults. *Journal of Gerontology: Psychological Sciences, 48,* P218–224.

McCalden, R. W., McGeough, J. A., Barker, M. B., & Court-Brown, C. M. (1993). Age-related changes in the tensile properties of cortical bone. The relative importance of changes in porosity, mineralization, and microstructure. *Journal of Bone and Joint Surgery, 75,* 1193–1205.

McFarland, R. A., Domey, R. G., Warren, A. B., & Ward, D. C. (1960). Dark adaptation as a function of age: I. A statistical analysis. *Journal of Gerontology, 15,* 149–154.

McLachlan, M. S. F., Guthrie, J. C., Anderson, C. K., & Fulker, M. J. (1977). Vascular and glomerular changes in the ageing kidney. *Journal of Pathology, 121,* 65–78.

McMurdo, M. E., & Gaskell, A. (1991). Dark adaptation and falls in the elderly. *Gerontology, 37,* 221–224.

McNaughton, M. E., Smith, L. W., Patterson, T. L., & Grant, I. (1990). Stress, social support, coping resources, and immune status in elderly women. *Journal of Nervous and Mental Disease, 178,* 460–461.

Meeuwsen, H. J., Sawicki, T. M., & Stelmach, G. E. (1993). Improved foot position sense as a result of repetitions in older adults. *Journal of Gerontology: Psychological Sciences, 48,* P137–141.

Meredith, C. N., Frontera, W. R., Fisher, E. C., Hughes, V. A., Herland, J. C., Edwards, J., & Evans, W. J. (1989). Peripheral effects of endurance training in young and old subjects. *Journal of Applied Physiology, 66,* 2844–2849.

Miller, R. A. (1993). Aging and cancer—Another perspective. *Journal of Gerontology: Biological Sciences, 48,* B8–9.

Montamat, S. C., & Davies, A. O. (1989). Physiological response to isoproterenol and coupling of beta-adrenergic receptors in young and elderly human subjects. *Journal of Gerontology: Medical Sciences, 44,* M100–105.

Montgomery, S. A. (1990). Depression in the elderly: Pharmacokinetics of antidepressants and death from overdose. *International Clinical Psychopharmacology, 5*(Suppl. 3), 67–76.

Morey, M. C., Cowper, P. A., Feussner, J. R., DiPasquale, R. C., Crowley, G. M., Kitzman, D. W., & Sullivan, R. J., Jr. (1991). Two-year trends in physical performance following supervised exercise among community-dwelling older veterans. *Journal of the American Geriatrics Society, 39,* 986–992.

Morgan, M. W. (1988). VIsion through my aging eyes. *Journal of the American Optometric Association, 59,* 278–280.

Morley, J. E., & Reese, S. S. (1989). Clinical implications of the aging heart. *The American Journal of Medicine, 86,* 77–86.

Moses, R. A. (1981). Accommodation. In R. A. Moses (Ed.), *Adler's physiology of the eye.* St. Louis, MO: Mosby.

Mouritzen Dam, A. (1979). The density of neurons in the human hippocampus. *Neuropathology and Applied Neurobiology, 5,* 249–264.

Naliboff, B. D., Benton, D., Solomon, G. F., Morley, J. E., Fahey, J. L., Bloom, E. T., Makinodan, T. et al. (1991). Immunological changes in young and old adults during brief laboratory stress. *Psychosomatic Medicine, 53,* 121–132.

Neils, J., Newman, C. W., Hill, M., & Weiler, E. (1991). The effects of rate, sequencing, and memory on auditory processing in the elderly. *Journal of Gerontology, 46*(2), 71–75.

Nuti, R., & Martini, G. (1993). Effects of age and menopause on bone density of entire skeleton in healthy and osteoporotic women. *Osteoporosis International, 3,* 59–65.

O'Donoghue, M. N. (1991). Cosmetics for the elderly. *Dermatology Clinics, 9,* 29–34.

O'Leary, A. (1990). Stress, emotion, and human immune function. *Psychological Bulletin, 108,* 363–382.

Olivetti, G., Melissari, M., Capasso, J. M., & Anversa, P. (1991). Cardiomyopathy of the aging human heart. Myocyte loss and reactive cellular hypertrophy. *Circulation Research, 68,* 1560–1568.

Ordy, J. M., Brizzee, K. R., & Johnson, H. A. (1982). Cellular alterations in visual pathways and the limbic system: Implications for vision and short–term memory. In R. Sekuler, D. Kline, & K. Dismukes (Eds.), *Aging and human visual function.* New York: Alan R. Liss.

Ouslander, J. G., & Abelson, S. (1990). Perceptions of urinary incontinence among elderly outpatients. *Gerontologist, 30*(3), 369–372.

Ouslander, J. G., Zarit, S. H., Orr, N. K., & Muira, S. A. (1990). Incontinence among elderly community-dwelling dementia patients: Characteristics, management, and impact on caregivers. *Journal of the American Geriatrics Society, 38,* 440–445.

Overend, T. J., Cunningham, D. A., Kramer, J. F., Lefcoe, M. S., & Paterson, D. H. (1992). Knee extensor and knee flexor strength: Cross-sectional areas ratio in young and elderly men. *Journal of Gerontology: Medical Sciences, 47,* M204–210.

Panek, P. E., Barrett, G. V., Sterns, H. L., & Alexander, R. A. (1977). A review of age changes in perceptual information ability with regard to driving. *Experimental Aging Research, 3,* 387–449.

Parmelee, P. A., Katz, I. R., & Lawton, M. P. (1991). The relation of pain to depression among institutionalized aged. *Journal of Gerontology: Psychological Sciences, 46,* P15–21.

Paterson, C. A. (1979). Crystalline lens. In R. E. Records (Ed.), *Physiology of the human eye and visual system.* New York: Harper & Row.

Poulin, M., Paterson, D. H., Govindasamy, D., & Cunningham, D. A. (1988). Endurance training of elderly men. Responses to submaximal exercise. *Canadian Journal of Sport Sciences, 13,* 78–79P.

Reddan, W. G. (1981). Respiratory system and aging. In E. L. Smith & R. C. Serfass (Eds.), *Exercise and aging: The scientific basis* (pp. 89–107). Hillside, NJ: Enslow.

Richards, O. W. (1977). Effects of luminance and contrast on visual acuity ages 16 to 90 years. *American Journal of Optometry and Physiological Optics, 54,* 178–184.

Rico, H., Revilla, M., Hernandez, E. R., Gonzalez-Riola, J. M., & Villa, L. F. (1993). Four-compartment model of body composition of normal elderly women. *Age and Aging, 22,* 265–268.

Riggs, B. L., Wahner, W. H., Dunn, W. L., Mazess, R. B., Offord, K. P., & Melton, L. J. (1981). Differential changes in bone mineral density of the appendicular and axial skeleton with aging. *Journal of Clinical Investigation, 67,* 328–335.

Rikli, R. E., & McManis, B. G. (1990). Effects of exercise on bone mineral content in postmenopausal women. *Research Quarterly in Exercise of Sport, 61,* 243–249.

Roberto, K. (1992). Coping strategies of older women with hip fractures: Resources and outcomes. *Journal of Gerontology: Psychological Sciences, 47,* 21–26.

Rosenbloom, C. A., & Whittington, F. J. (1993). The effects of bereavement on eating behaviors and nutrient intakes in elderly widowed persons. *Journal of Gerontology: Social Sciences, 48,* S223–229.

Rosenhall, U., & Rubin, W. (1973). Degenerative patterns in the aging human vestibular neuroepithelia. *Acta Otolaryngolica, 76,* 208–220.

Rowe, J. W. (1982). Renal function and aging. In M. E. Reff & E. L. Schneider (Eds.), *Biological markers of aging.* Bethesda, MD: National Institutes of Health, Publication No. 82–2221.

Rowe, J. W., Shock, N. W., & DeFronzo, R. A. (1976). The influence of age on the renal response to water deprivation in man. *Nephron, 17,* 270–278.

Rudberg, M. A., Furner, S. E., Dunn, J. E., & Cassel, C. K. (1993). The relationship of visual and hearing impairments to disability: An analysis using the longitudinal study of aging. *Journal of Gerontology: Medical Sciences, 48,* M261–265.

Ryan, A. S., Craig, L. D., & Finn, S. C. (1992). Nutrient intakes and dietary patterns of older Americans: A national study. *Journal of Gerontology: Medical Sciences, 47,* M145–150.

Ryff, C. D. (1989). In the eye of the beholder: Views of psychological well-being among middle-aged and older adults. *Psychology and Aging, 4,* 195–210.

Scheibel, A. B. (1982). Age-related changes in the human forebrain. *Neurosciences Research Progress Bulletin, 20,* 577–583.

Scheie, H. G., & Albert, D. M. (1977). *Textbook of opthalmology.* (9th ed.). Philadelphia: Saunders.

Schiavi, R. C., & Schreiner-Engel, P. (1988). Nocturnal penile tumescence in healthy aging men. *Journal of Gerontology: Medical Sciences, 43,* M146–150.

Schulman, S. P., & Gerstenblith, G. (1989). Cardiovascular changes with aging: The response to exercise. *Journal of Cardiopulmonary Rehabilitation, 9,* 12–16.

Shephard, R. J. (1978). *Physical activity and aging*. Chicago, IL: Yearbook Medical.

Sherman, S. S., Tobin, J. D., Hollis, B. W., Gundberg, C. M., Roy, T. A., & Plato, C. C. (1992). Biochemical parameters associated with low bone density in healthy men and women. *Journal of Bone and Mineral Research, 7,* 1123–1130.

Shimojo, M., Tsuda, N., Iwasaka, T., & Inada, M. (1991). Age-related changes in aortic elasticity determined by gated radionuclide angiography in patients with systemic hypertension or healed myocardial infarcts and in normal subjects. *American Journal of Cardiology, 68,* 950–953.

Shock, N. W., Andres, R., Norris, A. H., & Tobin, J. D. (1978). Patterns of longitudinal changes in renal function, *Proceedings XII International Congress of Gerontology, Tokyo, August 20–25* (pp. 525–527). Amsterdam: Excerpta Medica.

Slawinski, E. B., Hartel, D. M., & Kline, D. W. (1993). Self-reported hearing problems in daily life throughout adulthood. *Psychology and Aging, 8,* 552–561.

Smith, E. L. (1981). Physical activity: A preventive and maintenance modality for bone loss with age. In F. J. Nagle & H. J. Montoye (Eds.), *Exercise in health and disease* (pp. 196–202). Springfield, IL: Charles C. Thomas.

Smith, W. D. F., Cunningham, D. A., Patterson, D. H., Rechnitzer, P. A., & Koval, J. J. (1992). Forced expiratory volume, height, and demispan in Canadian men and women aged 55–86. *Journal of Gerontology: Medical Sciences, 47,* M40–44.

Solnick, R. L., & Birren, J. E. (1977). Age and male erectile responsiveness. *Archives of Sexual Behavior, 6,* 1–9.

Sparrow, D., Beausoleil, N. I., Garvey, A. J., Rosner, B., & Silbert, J. E. (1982). The influence of cigarette smoking and age on bone loss in men. *Archives of Environmental Health, 37,* 246–249.

Spector, A. (1982). Aging of the lens and cataract formation. In R. Sekuler, D. Kline, & K. Dismukes (Eds.), *Aging and human visual function*. New York: Alan R. Liss.

Stamford, B. A. (1988). Exercise in the elderly. In K. B. Pandolf (Ed.), *Exercise and Sports Sciences Reviews, 16,* 341–379.

Stevenson, J. S., & Topp, R. (1990). Effects of moderate and low intensity long-term exercise by older adults. *Research in Nursing and Health, 13,* 209–218.

Suominen, H., Heikkinen, E., Parkatti, T., Forsberg, S., & Kiiskinen, A. (1977). Effect of lifelong physical training on functional aging in men. *Scandanavian Journal of the Society of Medicine, 14* (Suppl.), 225–240.

Tamai, T., Nakai, T., Takai, H., Fujiwara, R., Miyabo, S., Higuchi, M., & Kobayashi, S. (1988). The effects of physical exercise on plasma lipoprotein and apolipoprotein metabolism in elderly men. *Journal of Gerontology: Medical Sciences, 43,* M75–79.

Teasdale, N., Stelmach, G. E., & Breunig, A. (1991). Postural sway characteristics of the elderly under normal and altered visual and support surface conditions. *Journal of Gerontology: Biological Sciences, 46,* B238–244.

Teramoto, S., Fukuchi, Y., Nagase, T., Matsuse, T., & Orimo, H. (1995). A comparison of ventilation components in young and elderly men during exercise. *Journal of Gerontology: Biological Sciences, 50A,* B34–39.

Thomas, P. D., Goodwin, J. M., & Goodwin, J. W. (1985). Effect of social support on stress-related changes in cholesterol, uric acid level, and immune function in an elderly sample. *American Journal of Psychiatry, 142,* 735–737.

Tinetti, M. E., & Powell, L. (1993). Fear of falling and low self-efficacy: A cause of dependence in elderly persons. *Journal of Gerontology, 48* (Special Issue), 35–58.

Toglia, J. U. (1975). Dizziness in the elderly. In W. Fields (Ed.), *Neurological and sensory disorders in the elderly*. New York: Grune & Stratton.

Vandervoort, A. A., Chesworth, B. M., Cunningham, D. A., Paterson, D. H., Rechnitzer, P. A., & Koval, J. J. (1992). Age and sex effects on mobility of the human ankle. *Journal of Gerontology: Medical Sciences, 47,* M17–21.

Van-Rooij, J. C., & Plomp, R. (1990). Auditive and cognitive factors in speech perception by elderly listeners: II. Multivariate analyses. *Journal of the Acoustical Society of America, 88,* 2611–2624.

Van-Rooij, J. C., & Plomp, R. (1992). How much do working memory deficits contribute to age differences in discourse memory? Special Issue: Cognitive gerontology. *Journal of the Acoustical Society of America, 91,* 1028–1033.

Vollardt, L. T. (1991). Psychoneuroimmunology: A literature review. *American Journal of Orthopsychiatry, 61,* 35–47.

Wald, A. (1990). Constipation and fecal incontinence in the elderly. *Gastroenterology Clinics of North America, 19,* 405–418.

Webb, G. D., Poehlman, E. T., & Tonino, R. P. (1993). Dissociation of changes in metabolic rate and blood pressure with erthrocyte Na–K pump activity in older men after endurance training. *Journal of Gerontology: Medical Sciences, 48,* M47–52.

Wedmann, B., Schmidt, G., Wegener, M., Coenen, C., Ricken, D., & Althoff, J. (1991). Effects of age and gender on fat-induced gallbladder contraction and gastric emptying of a caloric liquid meal: a sonographic study. *American Journal of Gastroenterology, 86,* 1765–1770.

Weisfeldt, M. L., & Gerstenblith, G. (1986). Cardiovascular aging and adaptation to disease. In J. W. Hurst (Ed.), *The heart*. New York: Macmillan.

Whitbourne, S. K. (1996a). *The aging individual: Physical and psychological perspectives*. New York: Springer.

Whitbourne, S. K. (1996b). *Identity processes and perceptions of physical functioning in adults: A test of the multiple threshold model*. American Psychological Association 104th Annual Meeting. Toronto, Ontario, Canada.

Whitbourne, S. K. (1986). *The me I know: A study of adult identity*. New York: Springer.

Whitbourne, S. K., & Primus, L. A. (1996). Physical identity in later adulthood. In J. E. Birren (Ed.), *Encyclopedia of aging*. San Diego, CA: Academic Press.

Whitbourne, S. K., & Wills, K.-J. (1993). Psychological issues in institutional care of the aged. In S. B. Goldsmith (Ed.), *Long-term care administration handbook* (pp. 19–32). Gaithersburg, MD: Aspen.

Wolf, E., & Gardiner, J. S. (1965). Studies on the scatter of light in the dioptric media of the eye as a basis of visual glare. *Archives of Opthalmology, 74,* 338–345.

Woodward, N. J., & Wallston, B. S. (1987). Age and health care beliefs: Self-efficacy as a mediator of low desire for control. *Psychology and Aging*, 3–8.

Wright, B. A., Aizenstein, S., Vogler, G., Rowe, M., & Miller, C. (1990). Frequent fallers: Leading groups to identify psychological factors. *Journal of Gerontological Nursing, 16*(4), 15–19.

Wyman, J. F., Harkins, S. W., & Fantl, J. A. (1990). Psychosocial impact of urinary incontinence in the community-dwelling population. *Journal of the American Geriatrics Society, 38*(3), 282–288.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.02
Cognition and Geropsychological Assessment

BOO JOHANSSON
Institute of Gerontology and Centre for Psychology, University College of Health Sciences, Jönköping, Sweden
and
ÅKE WAHLIN
Stockholm Gerontology Research Center and Karolinska Institute, Stockholm, Sweden

7.02.1 INTRODUCTION 26
7.02.2 NONCOGNITIVE INFLUENCES ON MEMORY AND COGNITION IN AGING 26
  7.02.2.1 Types of Aging 26
    7.02.2.1.1 Primary aging 27
    7.02.2.1.2 Secondary aging—physical health 28
    7.02.2.1.3 Summary 29
  7.02.2.2 The Individual 29
    7.02.2.2.1 Coping and self-attributions 29
    7.02.2.2.2 Psychological well-being 31
    7.02.2.2.3 Summary 32
  7.02.2.3 Sociocultural Context 33
    7.02.2.3.1 Cohort and period effects 33
    7.02.2.3.2 Sociocognitive development 34
    7.02.2.3.3 Summary 34
  7.02.2.4 Cognitive Demands and Everyday Functioning 35
    7.02.2.4.1 Cognitive mechanisms and tasks 35
    7.02.2.4.2 Tests of functional abilities 36
    7.02.2.4.3 Summary 36
7.02.3 MEMORY AND OTHER COGNITIVE ABILITIES 36
  7.02.3.1 Learning and Memory 36
    7.02.3.1.1 Types of learning 37
    7.02.3.1.2 Models of memory 37
    7.02.3.1.3 Episodic memory 37
    7.02.3.1.4 Memory tests and cognitive support 38
    7.02.3.1.5 Summary 39
  7.02.3.2 Intellectual Abilities 39
    7.02.3.2.1 Intelligence and aging 40
    7.02.3.2.2 Specific cognitive abilities 41
    7.02.3.2.3 Summary 45
7.02.4 COPING WITH COGNITIVE CHANGE IN AGING: PERSPECTIVES OF THE AGING INDIVIDUAL AND THE GEROPSYCHOLOGIST 45

7.02.4.1 "From Lab to Life" 45
  7.02.4.1.1 Individual differences 45
  7.02.4.1.2 Ecological validity 45
  7.02.4.1.3 Complex cognitive processing 46
  7.02.4.1.4 Everyday life experiences and learning 46
  7.02.4.1.5 Current knowledge and extralaboratory cognition 46
  7.02.4.1.6 Cognitive aging and the individual: an integrative model 47
  7.02.4.1.7 Cognitive reserves and plasticity 47
7.02.4.2 General Assessment Guidelines 48
7.02.4.3 Future Prospects and Cohorts 49
7.02.5 REFERENCES 49

7.02.1 INTRODUCTION

Understanding age-related changes, as well as stability, in highly integrated cognitive functions is of fundamental importance for the geropsychologist since memory and cognitive abilities are often focused in an assessment setting. The geropsychologist may be confronted with elderly individuals who seek help because of worries and experiences of cognitive failures, or who are referred to an assessment by family or health professionals due to impairments in memory and thinking.

A fundamental issue for an unbiased assessment of memory and other cognitive abilities is the geropsychologist's expertise in considering strengths or preserved ability as well as weaknesses or impaired ability relative to previous functioning, as well as the cognitive demands that are imposed in everyday life and how this relates to outcomes of cognitive testing. Other issues refer to the application of theoretical guidelines for analyzing performance, including the mechanisms involved, taking into account individual factors and the sociocultural context for cognitive functioning in aging.

This chapter discusses various aspects of cognitive functioning in aging in a context that is intended to correspond to that of the clinical geropsychologist, where considerations are based on detailed information about the individual in the assessment setting as well as outside the clinic. Other chapters in this volume address reasons for assessment of elderly people, including principles and methods. The present chapter provides a solid base for these more applied contributions. Given the limitation of a single chapter for discussing cognitive aspects of aging, the reader is directed to consult recent reviews for detailed information on current research on memory (e.g., Kausler, 1994), cognitive aging (e.g., Craik & Salthouse, 1992), and clinical geropsychology, including relevant chapters in the *Handbook of the psychology of aging* (Birren & Schaie, 1996 and previous editions).

This chapter is organized into subsections based on the heuristic scheme presented in Figure 1. The general idea with this scheme is that cognitive functioning must be understood and consequently analyzed within the context of noncognitive factors. The emphasis will be on aspects of memory and intellectual abilities that are likely to be addressed in an assessment setting (the term assessment is used throughout in its broadest sense and includes assessment in clinical settings as well as research approaches to various aspects of memory and cognitive aging).

The noncognitive factors depicted in Figure 1 are first discussed with reference to general aspects of memory and cognition. An overview of these domains of functioning follows, based on what is considered to be of relevance for the geropsychologist. The final section comprises an integration of cognitive aspects of aging, including messages to the geropsychologist.

7.02.2 NONCOGNITIVE INFLUENCES ON MEMORY AND COGNITION IN AGING

7.02.2.1 Types of Aging

Psychological aging is a complex interaction process involving influences from biological and social factors. In addition, the aging process is characterized by change as well as continuity and becomes manifest in a multidimensional and multidirectional manner. The former refers to the differential view on aging in which various abilities are affected differently, while the latter refers to the observation that certain abilities may deteriorate gradually, whereas still others remain stable or even improve.

The cascade metaphor of aging, proposed by Birren and Cunningham (1985), represents a heuristic simplification in the context of geropsychological assessment. The metaphor is based on the separation of primary, secondary,

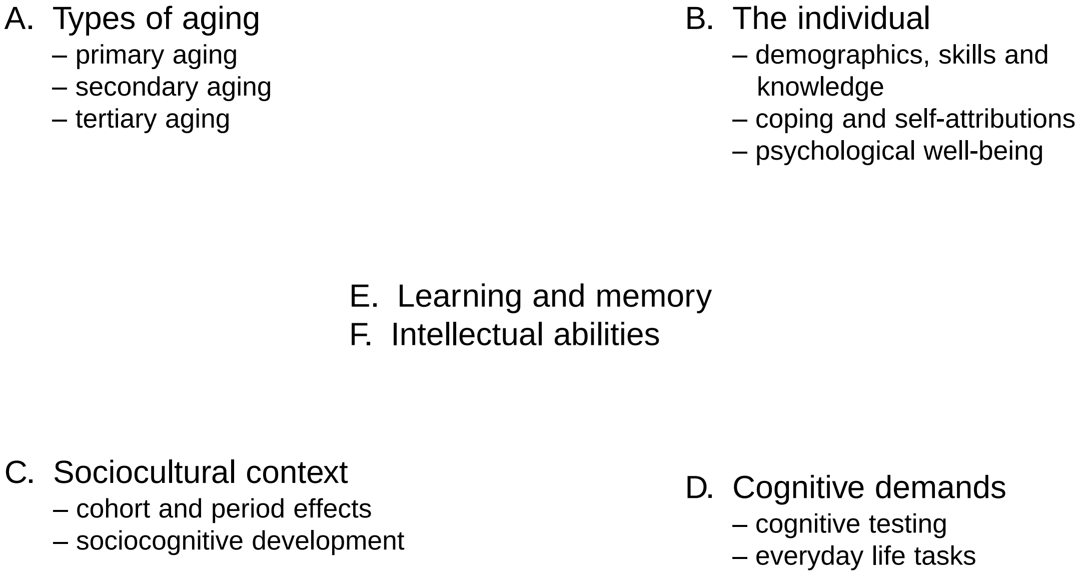

Figure 1 An analytical scheme for understanding and assessment of cognitive functioning in aging.

and tertiary aging patterns. Primary aging refers to intrinsic sources producing irreversible changes but otherwise a healthy or normal aging trajectory. Secondary aging refers to changes due to illnesses or pathology, adding to the inevitable primary aging. Tertiary aging refers to factors increasing the mortality risk and the pattern of terminal decline that becomes manifest preceding death. At the individual level this may be seen as a "curvilinear" dramatic drop while at the group level as a linear decline (cf. Berg, 1996).

It could easily be argued that it is impossible fully to separate the three theoretical aging types. However, in an assessment it is usually required that the geropsychologist concludes whether an observed impairment is to be diagnosed in terms of dysfunction or whether the impairment is to be considered within the borders of "normal aging." The phenomenon of terminal decline (e.g., Johansson & Berg, 1992) must also be taken into account, but is otherwise not an obvious diagnostic issue.

7.02.2.1.1 Primary aging

With increasing age we can expect less primary aging and more secondary and tertiary aging. Although the likelihood of sole primary aging decreases substantially with age because of increased morbidity, primary aging constitutes the bottom line for our understanding of cognitive aging.

Primary aging is, in practice, defined as the lack of illnesses that potentially could be diagnosed. In a population-based random sample of nondemented community-dwelling individuals aged 75 to 96 years, Wahlin et al. (1993) had to exclude 38% of the participants in order to meet their health criteria. Of these, 19% were excluded because of a Mini-Mental State score (MMSE) below 25 (Folstein, Folstein, & McHugh, 1975), 4% due to psychiatric disease and/or medication, and 15% due to values of serum vitamin B^{12} or thyroid-stimulating hormone indicative of disease.

Primary cognitive aging is principally characterized by a reduction in the speed at which information is processed in the cognitive system (e.g., Light, 1991; Salthouse, 1993). This slowing affects all cognitive processes where speed of behavior is an essential component, for example, in problem solving or executive functioning. In many tests, time limits are therefore more disadvantageous for an elderly person as compared to a younger individual.

A conceptual distinction within primary aging is that between usual and successful aging. According to Rowe and Kahn (1987), usual aging refers to normal aging heightened by the effects of extrinsic factors, both physical and psychosocial. In successful aging, these extrinsic influences have a neutral or positive role in maintaining functioning. As emphasized by Rowe and Kahn, the elucidation of factors that explain success is of great value for potential interventions, including preventive efforts, aiming at successful aging. Hill, Wahlin, Winblad, and Bäckman (1995) followed this route and examined the predictive strength of demographic and life-style variables for memory performance and ability to utilize cognitive support in very old age. The authors concluded that among healthy elderly persons, those who were younger, better educated, and more socially active not only improved more from cognitive support, but also demonstrated better overall performance. Further, level of exercise (in terms of walking or bicycling) was found to be a reliable predictor of episodic free recall performance. Although this was a cross-sectional study, and no conclusions can be drawn about causative relations, it was argued that the results are relevant for the selection of individuals who are likely to benefit from memory training programs. Identifying factors important for level of cognitive performance (such as level of social activity or exercise) may also have implications for other interventions.

7.02.2.1.2 Secondary aging—physical health

Secondary or disease-related aging becomes significantly more important with increasing age. It is well known that age is associated with an increased prevalence of a variety of diseases that may affect cognitive and memory functioning (e.g., Brody & Schneider, 1986; Fries & Crapo, 1981). Among persons over 65, more than 80% have at least one chronic illness, and many individuals have multiple diagnoses.

In addition to an even more pronounced slowing of speed (e.g., Salthouse, 1993), other intellectual abilites are also likely to become affected by various illnesses. A few examples are sufficient to illustrate the effects of secondary aging on cognitive abilities: the cardiovascular system is prone to age-related changes, the prevalence of heart disease among those 65 years of age and older in the Western world is estimated to be approximately 30% (Epstein et al., 1965), and the treatment of hypertension among elderly adults is considered by many geriatricians to be an almost normal and inevitable medication in aging. It has been demonstrated that cardiovascular disease has negative impacts on memory and cognition (e.g., Barclay, Weiss, Bond, & Blass, 1988; Grut, Forsell, Hill, Bäckman, & Winblad, 1995). Furthermore, the negative relation between elevated blood pressure and cognitive functioning has been documented in several longitudinal investigations (e.g, Farmer et al., 1990), and performance differences may also exist as a function of blood pressure medication (e.g., Elias, Wolf, D'Agostino, Cobb, & White, 1993). Also, older people may have health-related motor impairments that interfere with performance on psychological tests. Examples include tremor associated with Parkinson's disease, or difficulties associated with severe arthritis. Obviously, it may be necessary to modify tasks that put demands on hand motor functioning. Additional examples concern the negative relation between cognitive performance and such diverse conditions as diabetes (e.g., Bale, 1973), epilepsy (e.g., Dikman & Matthews, 1977), hyperthyroidism (Gambert, 1988), hypothyroidism (Mennemeier, Garner, & Heilman, 1993), or diseases related to vitamin deficiency (e.g., Wahlin, Hill, Winblad, & Bäckman, in press). It is well known that visual acuity (Lerman, 1983) and hearing, especially of high-frequency tones (Ordy, Brizzee, Beavers, & Medart, 1979), decline with age. In the assessment situation, it therefore becomes neccessary first to make sure that the older person has sufficient hearing or vision abilities to meet the sensory demands of the tasks at hand. It is notable that even within normal ranges, sensory functioning seems to be a strong late-life indicator of intellectual functioning (Lindenberger & Baltes, 1994).

Comprehensive health screening is rarely performed in cognitive aging studies. For the clinician, this should be kept in mind in using results from scientific reports as normative data. More recent studies of age-related effects on memory and other cognitive performance, however, have tended to screen for potentially confounded critical aspects of physical health (e.g., Bäckman & Wahlin, 1995; Hill et al., 1995; Robins Wahlin, Bäckman, Wahlin, & Winblad, 1993, 1996; Wahlin, Bäckman, & Winblad, 1995). Given the complexity of the potential impact of health-related factors, it is difficult to present clear-cut recommendations to the geropsychologist. Generally speaking, it is however important to evaluate performance, keeping in mind that influences other than what is addressed by the diagnostic question must be taken under serious consideration by the geropsychologist.

There exists at present no clear consensus about the differentiation between primary aging and pathology. This distinction becomes increasingly difficult with age, with the risk of both over- and underdiagnosis, that is, defining assessment results within the category of "normal aging" as disease/illness or ascribing an impairment to "normal aging." Also, the symptoms for many diseases change and tend to be more diffuse, in addition to prevalent comorbidity in old age (Fozard, Metter, & Brant, 1990). Furthermore, a diagnosis of disease is often accompanied by certain treatment efforts and with potential side-effects on cognitive functions. For example, Berg and Dellasega (1996), examined the relationships between cognitive function and psychotropic medication use in the H70 population-based sample in Sweden over a nine-year period (see Section 7.02.3.2.1). About one-third of the sample was on a drug regimen including anxiolytics, hypnotics, and antidepressant medications. The study showed that these psychotropics had a negative effect on most cognitive functions, although the effects generally were found to be moderate.

The dementia syndrome represents a diagnostic category of particular relevance for cognitive aging, since dementia is the single most devastating group of illnesses in which memory disturbances and cognitive impairments are cardinal outcomes. We will only address a few issues here. Dementia is extensively discussed in Chapter 10, this volume. The prevalence of dementia increases with age and the incidence rates are considerable (e.g., Fratiglioni et al., 1991; Johansson & Zarit, 1995). Dementia also demonstrates the principal difficulties in separating primary and secondary aging at an early stage, despite findings of significant impairments already present in the earliest, preclinical stage. Johansson and Zarit (1997) examined this "borderline issue" in a longitudinal study of a population-based sample of people aged 84 and older. They found that nondemented subjects who were diagnosed with "mild dysfunction" because of their performance were 11 times more likely to develop dementia within a two-year period, and seven times more likely to become demented over a four-year period. Their results also provide further evidence for terminal decline, as the likelihood for subsequent mortality was greater among individuals diagnosed with "mild dysfunction." The mortality risk was three times greater over two years and five times greater over four years compared with individuals who were rated as cognitively intact (see Figure 2).

The insidious onset of many diseases, like that of primary degenerative dementia (e.g., Alzheimer's disease) presents particular diagnostic problems. As pointed out by Sliwinski, Lipton, Buschke, and Stewart (1996), the identification of preclinical dementia is complicated by the overestimation of age effects on cognitive measures through an underestimation of the mean and an overestimation of variance. This notion becomes more important with increasing age, because of higher prevalence rates and increased likelihood of the incidence of dementia. Besides the neccessary criteria of impaired memory and other cognitive abilities, a functional criterion is generally employed for the definition of a clinically manifest dementia. This functional criterion, as in the case of the *Diagnostic and statistical manual for mental disorders* (*DSM*; American Psychiatric Association, 1987, 1994), requires that cognitive impairments need to be of sufficient severity to interfere with social or occupational functioning. A problem, however, is that cognitive demands may vary with age. The risk of underdiagnosis tends to be greater with the oldest-old due to lower cognitive demands in everyday life which makes it less likely that impairments will be uncovered. A more detailed discussion about inputs from the sociocultural context follows in Section 7.02.2.3.

7.02.2.1.3 Summary

Aging is not a unitary process but the complex outcome of numerous interactions with biological, psychological and social subprocesses. The outcome can roughly be categorized using three broad aging types: primary, secondary, and tertiary aging. The distinction between primary and secondary aging often corresponds to what is expected of the geropsychologist. A differential view of cognitive aging means that while performance in certain functions has declined and may be diagnosed as an expression of disease, for example dementia, performance in other functions may remain within the "normal aging" category or at the border of primary aging. This is especially the case when impairments are mild or changes are subtle and where there is lack of information on decline and course of change.

7.02.2.2 The Individual

Within each type of aging pattern, described above, one can discern features unique to the individual. Such individual or subject-related factors refer to exposures that have affected the individual throughout life. Whereas certain factors reflect basic demographics like age, gender, ethnicity, and socioeconomics, others refer to various aspects of personality and mental health. An understanding of the unique constellation of these influences requires the gathering of information necessary to evaluate present functioning in light of previous ability. In addition, this notion provides arguments for an interindividual difference perspective on aging.

7.02.2.2.1 Coping and self-attributions

The individual's unique repertoire of behaviors, including cognitive and emotional styles, represents another dimension to consider in an assessment. Coping and self-attributions are, in the late 1990s, more often identified as significant influences for memory and cognition. A number of coping definitions have been proposed. Most of them share the general idea

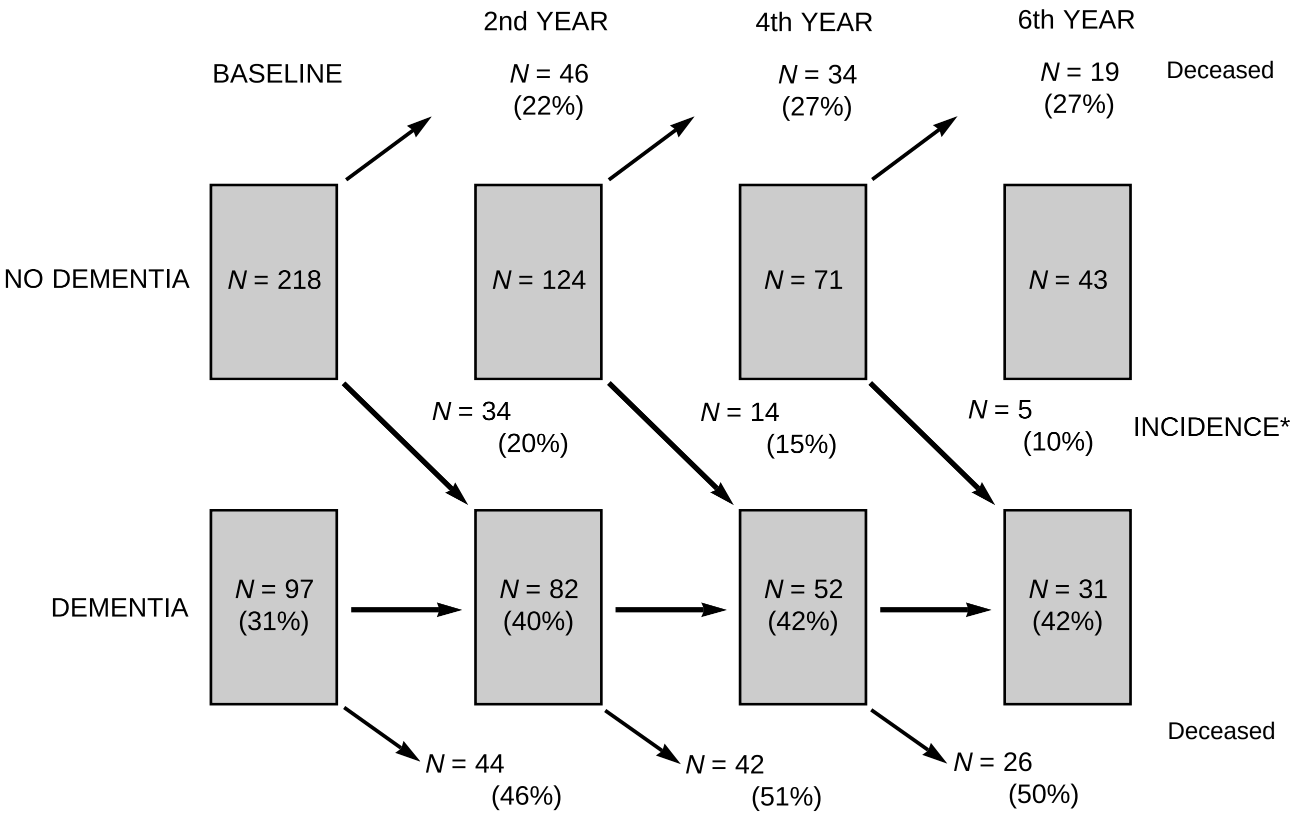
2nd YEAR
4th YEAR
6th YEAR
BASELINE
N = 46 (22%)
N = 34 (27%)
N = 19 (27%)
Deceased
NO DEMENTIA
N = 218
N = 124
N = 71
N = 43
N = 34 (20%)
N = 14 (15%)
N = 5 (10%)
INCIDENCE*
DEMENTIA
N = 97 (31%)
N = 82 (40%)
N = 52 (42%)
N = 31 (42%)
Deceased
N = 44 (46%)
N = 42 (51%)
N = 26 (50%)

* Incidence was calculated by taking the percentage of new cases of dementia from among surviving subjects not diagnosed as demented at the previous time of assessment. (At Time 2 there was incomplete information for diagnosis in 14 cases, at Time 3 in 8 cases and at Time 4 in one case.)

Figure 2 Prevalence of dementia and incidence of demention and mortality over a six-year period in a population-based sample of 84–90 year olds ("Early cognitive markers of the incidence of dementia and mortality: A longitudinal population-based study of the oldest old" by B. Johansson and S. H. Zarit, 1997. *Journal of Geriatric Psychiatry, 12,* pp. 53–59. Copyright 1997 by John Wiley & Sons Ltd. Adapted with permission).

of Lazarus and Folkman (1984), who described coping as "constantly changing cognitive and behavioral efforts to manage specific external and/or internal demands that are appraised as taxing or exceeding the resources of the person" (p. 14). Because appraisal is a crucial component, self-attributions and personal control come into focus. Personal control, locus of control, self-efficacy, and other similar concepts all refer to the individual's feeling of being in control or lacking control over life in general or in managing certain aspects, for example intellectual tasks (see, Baltes & Baltes, 1986). In this respect, coping and dimensions of personal control are also important for daily life adjustment and performance on certain cognitive tasks.

Of particular relevance for memory and cognition in this context is metamemory and metacognition, referring to people's awareness of their own strengths and weaknesses in performing memory and other cognitive tasks. For example, less effort than necessary may be invested in performance when belief exists that one's memory and cognitive capacity is too low for a certain task, and self-attributions given to actual success or failures in performance tend to reinforce these attributions. Hayslip and Panek (1993) describe the interrelations between performance experiences, self-statements, and emotional reactions in terms of a vicious cycle (see Figure 3).

Further, the relation between conceptions or self-attributions of, for example, intelligence, may be under multifactorial influence. Much of the work on conceptions of intelligence indicates that the dimensions that make up the prototypic intelligent person may vary at different periods of development and for individuals of different ages or experiental background (e.g., Siegler & Richards, 1982; Yussen & Kane, 1983). Berg and Sternberg (1992) examined whether young, middle-aged, and older adults viewed the concept of intelligence differently depending on their own age. They found that adults perceive differences in what constitutes exceptional intelligence during young, middle, and late adulthood, but also that adults' conceptions of intelligence are characterized by multidirectionality and modifiability.

In a study of actual and perceived change in intellectual performance, Schaie, Willis, & O'Hanlon (1994) identified three groups: realists (individuals who accurately estimated actual change), optimists (those who overestimated actual change), and pessimists (those who underestimated change). This classification was based on self-reports and actual performance on several tests. Older individuals were more realistic than younger subjects on a test of number ability, while more pessimistic about verbal meaning and reasoning tests. Section 7.02.2.2.2 examines further relations between psychological well-being, self-attributions, and cognitive performance.

7.02.2.2.2 Psychological well-being

Research into the psychological well-being of elderly persons tends to emphasize prior life satisfaction and mood as important predictors. Self-attributions, in terms of subjective memory and cognition, are also concurrent markers for adjustment and well-being. Beyond normal emotional fluctuations, the interrelationships between self-reports and performance become obvious in clinical depression. An inability in the depressed person to mobilize motivation will ultimately lead to deficient performance. Depression is a prevalent and disabling mental disorder in the elderly and the symptoms tend to be more indistinguishable and often involve impairments in memory and cognition (e.g., Bäckman & Forsell, 1994), in some cases with a severity that corresponds to depressive pseudodementia (see Katona, 1994). Severity of depression seems to be correlated with the magnitude of performance deficits in tests of memory and cognition (e.g., Cipioli, Neri, Andermarcher, Pinelli, & Lalla, 1990).

Numerous studies have reported strong associations between depression and self-reports of cognitive functioning. It was, however, early recognized that complaints must be differentiated from actual performance (Kahn, Zarit, Hilbert, & Niederehe, 1979). A diagnostic scheme for different configurations of performance, complaints, and affective status was presented by Gallagher, Thompson and Levy (1980). In a meta-analysis, Burt, Zembar, and Niederehe (1995) even suggested that complaints may have a stronger association with general psychopathology than with depressive symptoms. Further studies are needed to resolve the complex interrelationships between cognitive performance, mood, and self-attributions (see Kaszniak & Ditraglia Christenson, 1995).

Zarit (1979) suggested that memory complaints also reflect "a memory and aging stereotype" that ascribes poor memory to the elderly due to their age. Such a stereotype is likely to be internalized by the elderly. Studies on self-reports show, however, a great variation, typically in the range of 40–80%, in the extent to which older people feel that their memory performance is compromised. These discrepancies may reflect sampling differences and sociocultural values. In a population-based

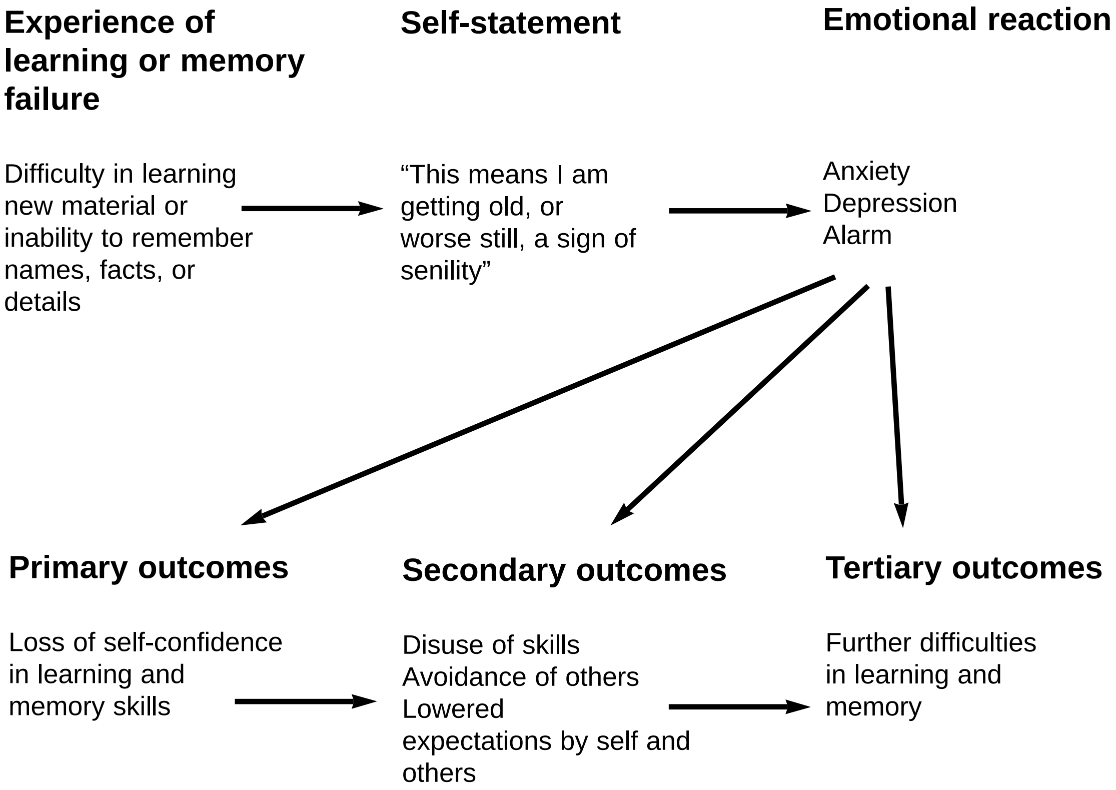
Experience of learning or memory failure
Difficulty in learning new material or inability to remember names, facts, or details
Self-statement
"This means I am getting old, or worse still, a sign of senility"
Emotional reaction
Anxiety
Depression
Alarm
Primary outcomes
Loss of self-confidence in learning and memory skills
Secondary outcomes
Disuse of skills
Avoidance of others
Lowered expectations by self and others
Tertiary outcomes
Further difficulties in learning and memory

Figure 3 A vicious cycle in which memory failures, self-statements, and emotional reactions tend to compromise future performance (*Adult Development and aging* (p. 158), by B. Hayslip Jr. and P. E. Panek, 1993, New York: HarperCollins. Copyright 1993 by HarperCollins College Publishers. Adapted with permission).

sample of very old Swedish people, participants were examined four times over a six-year period, and were asked to rate their own memory and cognitive ability (Johansson, Allen-Burge, & Zarit, in press). Participants typically evaluated their abilities in a positive manner (see Table 1). A subsample of 31% were, however, excluded because of dementia, which has to be taken into account when examining the figures.

These self-evaluations were scaled and further examined in relation to concurrent, prospective, and retrospective performance on a battery of memory and cognitive tasks (Johansson et al., 1997). Individuals who evaluated their memory more positively showed significantly better concurrent test performances. In addition, they also reported less depression. The study interestingly revealed that self-evaluations predicted subsequent decline in test performance, as well as onset of dementia within the next two to four years. In the retrospective analysis, more negative self-evaluations were found in individuals who two years later reported that they had experienced declining memory. The actual strength in the associations between test scores and self-evaluations were, however, quite weak. This suggests that self-evaluations are influenced by factors not covered by performance on tests. The few existing studies on the relationships between self-rated intellectual abilities and performance on intelligence and practical problem-solving tasks also show weak associations (e.g., Cornelius, 1990).

7.02.2.2.3 Summary

Present functioning should always be related to the individual's prior level of functioning. This will obviously present difficulties as we seldom have baseline information or any other detailed information about the individual's past function at the level of various abilities in focus in an assessment. It is also necessary to consider influences from demographics and to take into account the individual's own perception of memory and cognitive functioning, as well as other expressions of psychological well-being that may exert an impact on the effort invested in a certain performance. Although self-reports are closely related to psychological well-being, self-evaluations may also be formed by experiences of success and failure in performing cognitive tasks in daily life.

Table 1 Self-evaluations of memory in a population-based sample of 84–90 year olds in Sweden.

Self-evaluation	*On the whole, how would you rate your memory? (%)*	*On the whole, how would you rate your thinking and problem solving? (%)*
Good	44%	53%
Rather good	34%	33%
Neither good nor bad	8%	7%
Rather bad	5%	6%
Bad	9%	1%
	Do you experience memory failures in everyday life that makes life more difficult? (%)	*Do you experience problems in remembering what to do and say at a later occasion (prospective memory)? (%)*
No problems	82%	66%
Minor problems	17%	32%
Major problems	1%	2%
	Have you experienced any changes with your memory in the last two years? (%)	
Better now	1%	
No change	50%	
Somewhat worse now	44%	
Worse now	5%	

Source: Johansson et al. (1997).

7.02.2.3 Sociocultural Context

Aging is certainly not a context-free process (see Blanchard-Fields & Abeles, 1996). On the contrary, health status and level of functioning in later life are dependent on cumulative outcomes from numerous sources, including influences from the sociocultural context (e.g., dietary habits, transportation, and physical mobility, or the coverage and inputs by health care institutions like dentistry or preventive pediatrics). A list of potential factors would be very extensive. The differential exposure to environmental influences generally makes people different, and it is often claimed that interindividual differences increase with age (e.g., Krauss, 1980; Nelson & Daneffer, 1992). This is still an unresolved issue, but in assessing elderly persons, attending to sociocultural influences seems as important as attending to ethnicity or social class.

7.02.2.3.1 Cohort and period effects

Cohort is a concept often used in the aging literature. Cohort generally refers to people born during a certain time and therefore exposed to similar experiences that tend to make them more alike than generations of people born earlier or later (see Rosow, 1978). Thus, age differences should be interpreted taking cohort effects into account. This was noticed early when age differences found in cross-sectional studies could be ascribed to significant cohort effects. The classical example is a study by Birren and Morrison (1961), who found higher correlations between certain tests of intelligence and years of formal schooling, than with age. The important role of education has since been identified in many studies (e.g., Inouye, Albert, Mohs, Sun, & Berkman, 1993). The educational system also represents an important social institution where different methods and content, hopefully, have long-lasting but differential effects on various cohorts. Thus, the potential impact of educational differences must be considered in evaluating cognitive performance, particularly in the presence of broad age ranges.

The importance of cohort is easily observed in the educational system. In elderly Swedish cohorts most individuals have six years of basic formal schooling only, compared to more than 12 years for persons in their twenties. Even at the level of brief cognitive screening, employing devices like the MMSE, it has been suggested that different cut-offs have to be used for elderly people at different educational levels (e.g., Crum, Anthony, Basset, & Folstein, 1993). Employing more sophisticated tests, cohort-adjusted norms have to be appropriate for accurate interpretations. Other cohort influences may be identified in the political scene, for example the 1968 movement in Europe, corresponding to the Vietnam generation in the USA, or personal experiences of the world wars that affected people differently across ages.

Period effects refer, in the context of aging, to influences that affect people throughout ages. An example is the introduction of television. Some claimed that socially and intellectually rewarding activities, such as crosswords and jigsaw puzzles, hobbies and so forth were replaced by television. The Swedish expression *dumburken* ("couch potatoes") captures these assumed negative aspects of passive television watching. At the same time no one can deny the wealth of information and general world knowledge that is easily brought to the public through television. Although period effects tend to be pervasive across ages, certain influences may affect older persons differently from younger (see Elder, 1979).

7.02.2.3.2 Sociocognitive development

Memory and other cognitive abilities develop in a sociocultural context where environmental influences tend to be normatively age-graded, that is, highly correlated with chronological age. A life-span model of intellectual development based on significant inputs from the sociocultural context was formulated by Schaie (1977–78). The model distinguishes four major developmental stages. Acquisition is the major theme in childhood and adolescence, preparing the individual with basic skills and information for later life. This basic expertise includes what society defines as essential for successful socialization and crucial for the reproduction of society. In the next stage, the young adulthood period—or the achieving stage—acquired knowledge has to be applied, whether at work, in the family, or in other social contexts. The middle age period that follows is described basically as a stage where cognitive abilities involve more social responsibilities or executive roles. The reintegrative stage in old age is a period where one's interests, attitudes, and values become the important inputs for the individual. Personal meaningfulness is ascribed a major theme in this stage. The transitions between stages proceeds from "What should I know?" through "How should I use what I know?" to "Why should I know?" In aging, the individual requires more meaning in order to be motivated to engage in a certain task.

Disuse has been identified as a main source for decline in abilities that require continuous use in order to be preserved (see Kohn & Schooler, 1983; Lee, 1991). In addition to the general slowing in primary aging, disuse is generally thought of as an important input for decline in functions not practiced. Support for this hypothesis is offered by an increased performance–capacity gap with increasing age, and interventions demonstrating considerable plasticity or modifiability in primary aging (e.g., Baltes & Kliegl, 1992). Perlmutter (1986) describes a life-span view of memory in which the individual's memory capabilities reflect past experience, and where decline in capacity can be compensated for by growth in memory contents. This leads to an increasing environmental fitness and to an increasingly adapted memory system.

7.02.2.3.3 Summary

In a sociocognitive development context, aging can be understood as a change of the behavior repertoire in accordance with the individual's own preferences, besides the influences of cohort and generation. Cohort effects refer to age differences due to influences that variously have affected different birth cohorts. Period effects, on the other hand, emphasize the need to consider certain sociocultural influences as more pervasive across ages. The actual impact of these influences may, however, vary depending on the age of the individual. The specific organization of society also imposes certain age-graded roles and tasks. Although a fixed stage model like Schaie's (1977–78) tends to neglect interindividual differences, it directs attention to the fact that certain tasks may be perceived quite differently depending on the differential use of intellect at different ages. This view emphasizes a need to consider what types of cognitive tasks could be employed in examining elderly people in an age-fair manner. The latter includes an ecological aspect of memory and cognition.

7.02.2.4 Cognitive Demands and Everyday Functioning

In cognitive aging research, age-related decline in a variety of abilities such as visuospatial functioning or episodic memory performance is well documented. In other types of task one may find that performance remains relatively stable or even improves across the adult life span. Preserved abilities are likely in tasks tapping general knowledge or language-related skills. Our knowledge is sparse on how this translates into everyday functioning (e.g., Salthouse, 1990; Searight & Goldberg, 1991). In clinical practice, the geropsychologist may have to resolve whether it is likely that impaired performance on, for example, tests of visuospatial ability is related to difficulties in everyday functioning. In other words, to what extent is it possible to infer, from certain test results, that an individual is incapable of performing adequately in everyday life? A patient with extensive prefrontal damage may be significantly impaired in everyday life, yet perform within the normal range on neuropsychological tests (e.g., Stuss & Benson, 1986). Another patient may score below cut-off for normal performance, but demonstrate adequate performance on a variety of everyday tasks, perhaps because these tasks are overlearned and because adequate performance is defined according to lower demands because of old age.

Clearly, everyday performance has to be related to the demands of daily living and available support systems, as well as baseline capacity. One can discern two major explanations for the understanding of discrepancies between elderly people's daily-life performance and scores on test of memory and cognition. The first is based on the idea that standard tests are inadequate because they simply do not measure the relevant processes underlying daily-life functioning. Thus, the assumption is that tests and everyday life differ in terms of the mechanisms involved. The other, functional, explanation, relates to age fairness emphasizing unfamiliarity and task novelty as significant inputs for lower test performance in the elderly (e.g., Bruce, 1991).

7.02.2.4.1 Cognitive mechanisms and tasks

The experimental and clinical test traditions have been criticized for their predilection for using artificial materials for remembering, or as stimuli in problem solving. The claim is that these methods have restricted relevance for the understanding of memory and cognition in everyday life (see Willis, 1996). In this vein, Neisser (1978) even suggested that naturalistic studies of memory would be more productive than their laboratory counterparts.

Some authors have argued that standard experimental tasks, in which elderly people demonstrate worse performance than younger ones, are not valid as indicators of everyday performance for older adults (e.g., Baltes & Willis, 1979; Berg & Sternberg, 1985; Labouvie-Vief, 1982, 1985). However, in spite of many efforts to develop more naturalistic measures, such as memory for texts or television programs, the results do not confirm the expected elimination of age effects or even attenuation of age differences (e.g., Denney, 1989; Hartley, 1989; Salthouse, 1983). On the contrary, performance on standard memory tasks and on more ecologically relevant tasks seem to provide similar results. It has also been claimed that certain tests may have a potential disadvantageous impact because different age groups are not equally experienced with stimulus materials or experimental and clinical tests of cognitive functions. There is, however, rather convincing evidence that differences in experience or unfamiliarity with stimulus materials cannot solely account for observed age differences (e.g., Salthouse, 1987, 1990).

The claim that standard tasks do not accurately portray everyday performance led some researchers to construct tests better suited for elderly persons, so called age-fair tasks. An example is the brief battery of tests based on ideas of ecological relevance and used by Johansson, Zarit, and Berg (1992) in studies of the oldest old. The tests, Coin Test, Clock Test, and MIR ("Memory in Reality"), were deliberately designed to minimize floor effects and to differentiate performance in cognitively impaired individuals. Age-fair testing is highly interrelated with *ecological relevance*, referring to naturalistic studies of memory and cognition. Both concepts emphasize the functional aspects, not that different mechanisms are involved in everyday performance compared to an experimental or clinical context. Ecological relevance refers to the congruence between task demands and demands imposed in everyday life. In other words, to what extent do the obtained pattern of results using psychometric or laboratory tasks generalize to situations in the everyday life of the individual? In a study by Cockburn and Smith (1991), subjects were tested with various measures of fluid and crystallized intelligence (see Section 7.02.3.2.1) in order to examine the predictive value of these tests for performances on a range of everyday memory problems, including analogs of situations met in daily life, such as remembering names or asking for the return of a hidden possession. Crystallized intelligence showed low predictive power, while fluid intelligence was a better predictor, although the authors concluded that cognitive decline in aging involves more than a simple decline in fluid intelligence. In another study on everyday memory, West, Crook, and Barron (1992) found that a measure of crystallized intelligence, vocabulary, played a significant role for almost every memory measure examined.

7.02.2.4.2 Tests of functional abilities

In some cases it may be preferable to perform a functional assessment of the client's abilities in parallel with cognitive testing. For this, clinicians have at their disposal several standardized methods measuring activities of daily living (ADL; e.g., Kane & Kane, 1981). These instruments assess functions that are associated with basic self-maintenance such as dressing, feeding, or hygiene. Furthermore, there are other instruments assessing a broader domain of more complex abilities such as managing transportation, communication, self-medication, or financial transactions. This domain of abilities is often called instrumental activities of daily living (IADL; e.g., Kuriansky & Gurland, 1976). One example is the Community Competence Scale (see Grisso, 1986). Using a multidimensional battery, Zarit, Johansson, and Berg (1993) found strong associations between cognitive impairment, ADL, and IADL in the oldest old. Results like this demonstrate that cognitive abilities are necessary prerequisites for adequate performance, although the sensitivity may be low within normal ranges of functioning.

7.02.2.4.3 Summary

Our understanding of memory and cognition are greatly determined by procedures and methods used in research. There seems to be no reason to believe that experimental and psychological tests measure functions other than those also constituting the basis for everyday functioning. The relation between test performance and level of functioning in daily life appears, however, to be influenced by numerous other factors such as social demands, support-systems, or the individual's level of experience. Everyday functioning of the client is of importance. In assessing elderly people, the degree to which test demands overlap with demands of daily life ought to be taken into consideration. Much effort has been invested in constructing tasks more suitable for the oldest people. Research suggests that results on various psychometric tests may show weak associations with everyday functioning. The inclusion of an assessment of actual performance in everyday life, such as IADL measures, can provide additional information of value to the geropsychologist.

7.02.3 MEMORY AND OTHER COGNITIVE ABILITIES

This section reviews some select research on memory and cognition in relation to current theory. Some brief outlines of models of memory, as well as processes and components involved in other cognitive abilities, will be provided. The reader is reminded of the analytical scheme (Figure 1) presented in the introductory section when critically examining reported findings and the bases for interpretations. In doing so, it is important to keep in mind the multifactorial influences on cognitive performance in old age. Furthermore, apart from noncognitive influences, it is also important to recognize that most published studies are cross-sectional and consequently demonstrate age differences only, and that the bulk of research includes convenience samples recruited for studying cognitive mechanisms or potential age-related differences in various abilities. Most studies provide no detailed information concerning physical health (relating to the distinction between primary and secondary aging). To what degree the failure to do so has influenced the typical picture of cognitive aging remains yet to be elucidated. Last, few studies of memory and cognition have included individuals beyond 80 years, which limits our understanding of the very old and oldest old.

7.02.3.1 Learning and Memory

Learning and memory represent highly interrelated processes, and the study of either requires understanding of the other. Studies of learning tend to emphasize the acquisition of new behaviors and information, while memory refers to processes involved in the retention and retrieval of what has been acquired. Learning and memory are here treated separately from other aspects of cognition because of the wealth of information available and the extensive research in the area. This subsection provides first a framework, out of several possible, for the understanding of learning and memory. In addition, there is an examination of the significance of task properties, here defined as levels of cognitive support, for variations in episodic memory performance. Finally, some relevant research is reviewed with focus on the oldest. The review is largely restricted to episodic memory, because this memory system has proved to be most affected by age, in addition to its sensitivity to depression, dementia, and other illnesses that are prevalent in aging.

7.02.3.1.1 Types of learning

Learning may mean many quite different things. It may refer to the mastering of a new skill such as learning how to ride a bicycle. The resulting procedural or motor-based memories are formed without awareness. This general area of learning and memory has been reviewed by Chiarello and Hoyer (1988) and Graf (1990). Learning may also mean acquiring and changing habits. Behavior modification is an example of this type of learning in clinical practice. Learning may further mean the acquisition of new information. This is something we all do in school. We may still remember what date World War One started, but not when that information was acquired. This is generally what we refer to as semantic memory.

This subsection restricts itself to episodic memory, and learning is concerned with the remembering of personal incidents that are, in most cases, not very distant in time. Tested persons are presupposed to remember where they picked up the requested information. The typical example is a test of memory where the psychologist reads a list of words or sentences, and the client is later asked to recall as many of the words or sentences as possible of those that were presented some minutes ago.

7.02.3.1.2 Models of memory

Many researchers classify memory phenomena in terms of memory systems, although the results are often analyzed in terms of various processes operating within these systems. For the clinician, the system model is perhaps the most intuitive, since it alludes to specific tasks. Before proceeding to some of the systems identified in modern experimental psychology, we choose as the point of departure a model that has exerted an enormous impact in the development of later theoretical accounts. This is the "Multistore" model of memory proposed by Atkinson and Shiffrin (1968). In clinical practice, their line of reasoning is perhaps still sufficient. Basically, this model of memory distinguishes between the sensory store where new information enters through sensory stimulation, the short-term store where the information is subjected to conscious mental activity, and the long-term store which is thought of as a more or less permanent store of information.

As a general guideline, it may be sufficient to distinguish between recent and remote memories, or between short- and long-term memory. These subdivisions are, however, somewhat complicated in terms of test materials. We therefore now turn to selected parts of the system model which are somewhat more clear for clinical purposes. For more detailed discussions, consult Baddeley (1990) or Parkin (1993).

Short-term memory is analogous to primary memory in that it represents the locus of conscious mental activity. Early accounts of primary memory (e.g., Waugh & Norman, 1965) stated that this system registers and retains incoming information in a highly accessible form for a short period of time after input. This conception of primary memory has been elaborated in recent years (e.g., Baddeley, 1986). It may be useful to distinguish primary memory (holding information in consciousness) from working memory (simultaneously holding information in mind and elaborating on or using that information in combination with other incoming information). Roughly, in terms of tests, this corresponds to the difference between Digit span forward and backwards (Wechsler, 1981).

The long-term store may, in terms of memory systems, be further subdivided into episodic and semantic memory. Episodic memory refers to memory of personally experienced events with a reference both in time and space, that is, we remember something we have experienced, and that memory is associated with information about when and where the information was encoded. The command "recall as many of the words as possible from the list I presented to you 10 minutes ago" evokes episodic memory. Semantic memory is memory for general knowledge (Tulving, 1983, 1985). The information task of the Wechsler Adult Intelligence Scale (WAIS) battery corresponds to a test of semantic memory.

7.02.3.1.3 Episodic memory

Most researchers agree that age differences in performance are most likely to occur in episodic memory tasks, that is, in tasks requiring conscious retrieval of information acquired in a particular place at a particular time (see Kausler, 1994; Light, 1991; Salthouse, 1991, for reviews). However, memory performance in very old age has received relatively little empirical attention. The available evidence is mixed as to whether there are differences in episodic memory functioning among different cohorts of old and very old adults. Most studies of the late adult life span indicate a gradual decline in performance with age (e.g., Bäckman, 1991; Bäckman & Larsson, 1992; Crook & Larrabee, 1992). There are, however, studies indicating no age differences (e.g., Cohen & Faulkner, 1989; Larsson & Bäckman, 1993; Olofsson & Bäckman, 1993). With reference to the previous discussion on potential impacts of health-related factors on cognitive performance, it is important to note that a careful health screening has not been performed in these memory and aging studies. This may be an essential factor underlying the contradictory findings. In studies performed by Wahlin and colleagues (e.g., Bäckman & Wahlin, 1995; Wahlin et al., 1993, 1995), careful health screening was routinely performed. The outcome of these studies showed that performance in a variety of episodic memory tasks differs reliably in comparing elderly people between 75 and 96 years of age. This was the case in tests of episodic face recognition, free recall and recognition of random word lists, and free and cued recall of organizable words. The age-related differences were, however, relatively small in raw score numbers and more pronounced in the free recall tasks as compared to the recognition tasks. In a study of 1000 healthy participants ranging in age between 35 and 80 years, a battery of episodic and semantic memory tests (Nyberg, Bäckman, Erngrund, Olofsson, & Nilsson, 1996) was employed. It was found that after controlling for differences in demographic (education, gender), cognitive (visuospatial ability), and health-related (blood pressure, serum vitamin B^{12}) factors, age did not predict performance on tests of semantic memory. The impact of age on tests of episodic memory was substantially reduced when differences in the background factors were controlled for, although still significant. These authors proposed that the failure to account for the age effects on episodic memory is because it can be attributed to age-related neuronal changes.

7.02.3.1.4 Memory tests and cognitive support

In addition to the many influences on cognitive and memory performance described in previous sections, performance may also vary depending on level of support provided. Some research indicates that age differences are large in unsupported tasks such as free recall, while relatively small in more supported tasks such as recognition (Craik & Jennings, 1992).

Before elaborating on the concept of cognitive support, this subsection will briefly describe three types of memory tests that have dominated research on episodic memory. In these tests, the retrieval conditions vary, such that free recall represents the lowest level of support, cued recall represents a more supportive test condition, and recognition of the information to-be-remembered (TBR) is the most supportive retrieval condition.

In the free recall task, subjects are presented with the items and are then asked to recall without prompting what was previously presented. In the cued recall condition, cues of some sort are provided by the examiner, and test materials are then typically organizable in some way. Questions such as "what animals do you remember?" or "what kind of furniture do you remember?" constitute the cues. Finally, in the recognition tasks, the initially presented materials are presented intermixed with a number of distractors. Subjects are asked to state whether they recognize items from the presentation.

Tasks may also vary in terms of presentation rate, where a slow presentation rate represents a more supportive condition as compared to a rapid presentation rate. Performance is also dependent upon the nature of the TBR materials. It is well known that stimulus dimensions like modality (i.e., items may be presented in a way that one or more of the sensory systems are stimulated during encoding), richness (e.g., items may be presented as objects or words), and organizability (i.e., it may or may not be possible to combine items in a meaningful fashion) are important to goodness of remembering (see Bäckman, Mäntylä, & Herlitz, 1990). Further examples include the datedness of the TBR materials (e.g., items may be faces of persons who were famous 50 years ago or faces of persons who attained their fame recently). There is a consensus that performance of older adults is facilitated by increasing the support with respect to modality, richness, organizability, and prior knowledge (Bäckman et al., 1990; Craik & Jennings, 1992) (see Figure 4).

There is converging evidence from experimental research and intervention studies that episodic memory in normal aging is modifiable (Bäckman et al., 1990; Verhaeghen, Marcoen, & Goossens, 1992). Older adults possess a substantial cognitive reserve capacity, although it seems to be less than for younger individuals (e.g., Baltes & Kliegl, 1992). However, there is no evidence for a simple relationship among cognitive support, episodic memory, and age (see Bäckman, 1995). Also, the level of support needed for the optimization of memory performance in dementia differs considerably from that required of healthy elderly people. While in dementia, support is needed at both encoding and retrieval, normally aged adults seem to require support at either encoding or retrieval only in order to enhance their performance (e.g., Herlitz, Lipinska, & Bäckman, 1992).

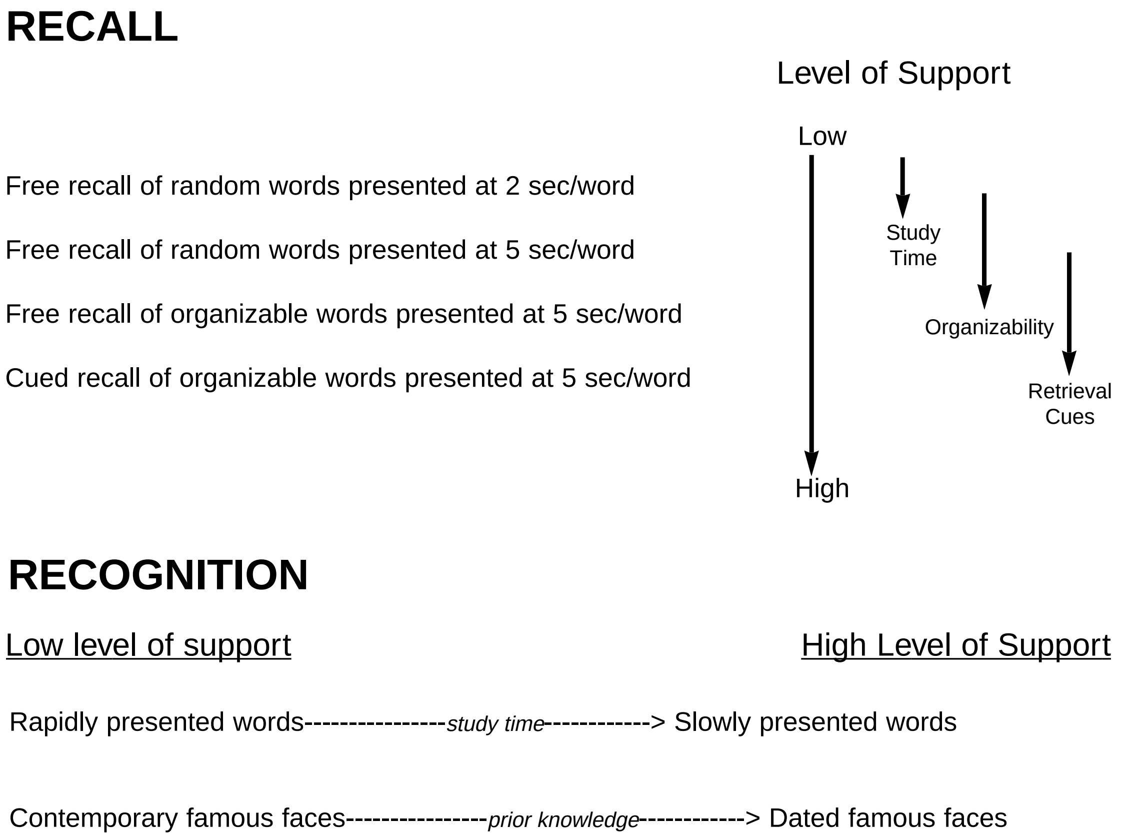
RECALL
Level of Support
Low
High
Free recall of random words presented at 2 sec/word
Free recall of random words presented at 5 sec/word
Free recall of organizable words presented at 5 sec/word
Cued recall of organizable words presented at 5 sec/word
Study Time
Organizability
Retrieval Cues
RECOGNITION
Low level of support
High Level of Support
Rapidly presented words----------------study time------------> Slowly presented words
Contemporary famous faces----------------prior knowledge------------> Dated famous faces

Figure 4 Memory tasks ordered along a continuum of cognitive support.

7.02.3.1.5 Summary

Learning and memory are highly interrelated and cannot be fully understood independently of each other. Most studies on memory in aging are cross-sectional and there is typically a lack of information necessary for ruling out the influence of the health status of the participants. The system model of memory employed in recent research usually distinguishes between episodic, semantic, primary, and working memory. Memory performance is seen as influenced by numerous properties acting at both encoding and retrieval stages of memory processing. Task properties may generally be recognized in terms of cognitive support (low–high), although there is no evidence of a simple relation between level of support and performance. The relatively few studies on episodic memory performance in very old age portray a gradual decline into the very late stages of the life span.

7.02.3.2 Intellectual Abilities

Learning and memory represent in a sense the input and output stages of information processing, while cognition, in a broad sense, also includes processes involved in the integration and reorganization of aquired knowledge for the sake of matching behavior in an appropriate manner, that is, intelligent behavior. Certain tasks are novel and require the execution of new behavior, while others may be practiced and require less cognitive effort. The outcome of matching task demands and cognitive resources is evaluated in the relative success, or failure, to perform a task whether measured as everyday performance or by cognitive tests. This section concentrates on abilities as expressed in aging and reported in the current literature, more than on the fundamental mechanisms involved in the mental operations necessary for intelligent behavior (see Craik & Salthouse, 1992; Rybash, Hoyer, & Roodin, 1986).

Studies on intelligence and aging have generally aimed at identifying normative patterns of age-related changes in the various ability components. The components are operationalized by the test battery. The following section describes such intelligence and aging patterns, while the next focuses on some of the specific cognitive components involved in intelligent behavior.

7.02.3.2.1 Intelligence and aging

The intelligence test tradition evolved at the beginning of the twentieth century as a differential psychology with focus on interindividual differences in intelligence. Age was recognized early as an important stratification variable in these studies. This psychometric tradition has since provided a basis for our understanding of normative aging and the typical age-related pattern. The most frequently used battery of tests in this research arena, as well as in the clinical setting, is the WAIS (Wechsler, 1955) and later modifications (WAIS-R, Wechsler, 1981) of this battery, encompassing tests of verbal ability (1, Vocabulary; 2, Information; 3, Comprehension), short-term learning and remembering (4, Digit-span forward and backwards), numerical abilities (5, Arithmetic), abstraction and generalizations (6, Similarities), and nonverbal abilities (7, Digit-symbol substitution; 8, Picture arrangement; 9, Picture completion; 10, Block design; 11, Object assembly). Wechsler divided the test battery into two broad categories: "the hold tests" (1, 2, 3, and 6) and "the don't hold tests" (7, 8, 9, 10, and 11) according to their resistance to age.

A number of cross-sectional studies over a broad age range have consistently confirmed the "classic aging pattern" (Botwinick, 1978), in which the verbal tests are identified as the hold tests and the nonverbal or performance tests as the don't hold tests. These findings were largely confirmed in later longitudinal studies, although the magnitude of age-related changes generally were found to be less than previously found by age differences. These design-related differences are intensively discussed in the gerontological literature and have supported the notion of including characteristics of the individual as well as recognizing influences from the sociocultural context when interpreting studies showing age effects.

The classical aging pattern was partly refined in a series of studies by Horn, Catell, and Donaldsson (see Horn, 1982, for an overview) who suggested a dichotomization between crystallized and fluid intelligence. This distinction was assumed to reflect different developmental trends. The label of fluid intelligence (Gf) was attached to basic processes involved in identifying complex relations among stimulus patterns and the drawing of inferences on the basis of comprehending relationships. The tests proposed to measure this broad ability component should not be greatly aided by the breadth of previous knowledge and include tests of logical reasoning and of figural and spatial relations. Crystallized intelligence (Cc) was defined as the cumulative end product of information acquired by the operation of fluid processes through interactions with the environment. The measurement of experience-based crystallized intelligence includes, for example, tests of vocabulary, information or general knowledge, comprehension, and arithmetic (Horn, 1982). The definition leads to expectations of preserved function, or even improvements with age.

The distinction between fluid and crystallized intelligence in aging differentiates intelligence components with regard to age. While development into adulthood is characterized by a differentiation of various abilities it is followed by a dedifferentiation in late adulthood and aging (e.g., Reinert, 1970). The broad categories of fluid and crystallized intelligence correspond to a dedifferentiated pattern of formerly more independent abilities. Some researchers claim that this dedifferentiation continues, and that there remains only a broad component of basic cognitive performance in the oldest old (eg., Marsiske, Lindenberger, & Baltes, 1994). The arguments for a more rudimentary intelligence system in advanced age awaits, however, further empirical support.

Although the results from some studies have been interpreted as favoring the general outcomes predicted by the fluid–crystallized dichotomy, the trend in recent research on intelligence and aging is to examine various intelligence components separately. This is consistent with the interest of a clinican when faced with separate scores for various tests. There are problems inherited in all higher-order approaches to intelligence, for example, the comparability between tests in different batteries. The greatest value, therefore, may be that higher-order approaches to intelligence can serve as heuristic frameworks for the clarification of relationships between complex abilities.

The most informative data on aging and intelligence patterns emerge from longitudinal studies using multiple measurements over longer time periods (see Berg, 1996; Schaie, 1996). In the ongoing H70 study in Gothenburg, Sweden (Rinder, Roupe, Steen, & Svanborg, 1975), a random population-based sample were included at the age of 70 and then followed over the years. Interesting patterns of age-related changes, as well as continuity, emerged. In this study, the effects of tertiary aging (terminal decline) were also examined separately for different survival intervals. That is, trends of change for individuals, for example measured at Time 1, Time 2, and Time 3, and then deceased, were compared with that of individuals who survived into the assessment at the next occasion, that is, Time 4. The main finding in the H70 study was that of a robust and pervasive terminal decline pattern in which performance levels were associated with survival. Figure 5 (a) and (b) (Berg, 1987) show this for a test of verbal meaning (Synonyms; SRB:1, Dureman & Sälde, 1959) and reasoning (Figure logic; SRB: 2, Dureman & Sälde, 1959).

The data from longitudinal studies typically show less change than is the case when examining age differences in cross-sectional studies. Although more accurate in providing a portrayal of intelligence and aging, longitudinal studies are affected by attrition and other selective influences that are likely to produce too optimistic a portrayal of aging trends.

7.02.3.2.2 Specific cognitive abilities

The "top-down" intelligence and aging perspective (described above) is here supplemented with a "bottom-up" approach focusing on specific cognitive abilites and some tests used for the measurement of these abilities.

(i) Speed of behavior: perceptual and psychomotor performance

Speed of behavior is a crucial ingredient in many memory and cognitive tasks because of time constraints, whether it is in everyday life, in the laboratory, or in clinical tests. While it has been argued by some that time is less important than accuracy, and that speeded tasks are unfair to elderly people, others have argued for the inclusion of speed of behavior as an important marker of intelligent behavior. The former position was supported by early assumptions that the reduced rate of processing was extrinsic or peripheral to the primary aging of the nervous system. Slowing has, however, been recognized as an intrinsic phenomenon that accompanies aging (e.g., Salthouse, 1993) caused by increased levels of noise in the nervous system (e.g., Cremer & Zeef, 1987), or broken or attenuated neural connections (MacKay & Burke, 1990). Most reseachers seem to have accepted speed as an inevitable component in various tests of intelligent behavior.

Psychomotor performance was frequently studied in the early days of gerontological inquiries. The prototype experiment compared performance of young and older people on tests of simple and complex reaction time (RT). The results of numerous RT studies consistently show that when the cognitive demands are increased by the introduction of multiple stimuli and response formats, the effects of age become more pronounced (Welford, 1977). The literature on perceptual speed also demonstrates a general agreement on age-related decline.

Attention and decision constitute two highly interwoven components in speed of behavior. Attention has been examined using experimental paradigms for divided and selective attention, vigilance, or the ability to identify less frequent signals occurring over long time periods, or in terms of concentration and search for certain signal patterns within an array of similar signals. Tests like digit-symbol in the WAIS battery represent a way to study concentration and search. The findings from this test demonstrate marked age effects. Elderly people show worse performance in terms of items accomplished per time unit, as compared to younger people. Accuracy is, however, not impaired in the elderly. The trail-making test is a frequently used clinical test for psychomotor performance and attention. Robins Wahlin et al. (1996) administered this test in a healthy sample of individuals aged 76–93. As expected, they found that the time needed to complete the test increased significantly with age. Noteworthy, however, is that they found no age differences in accuracy.

An important message is that lower speed is also related to deficient performance on a number of memory and other cognitive tasks. The pervasiveness of reduced speed of behavior in aging, thus requires that this phenomenon is taken into account in evaluating performance in tasks where speed is an ingredient (Salthouse, 1985; Stankov, 1988).

Studies in which elderly individuals' slower rate of processing is compensated for by elimination of time limits still show an overall pattern of performance decline with age, identical to that found when time constraints are present. The oldest, however, seem to benefit somewhat more from eliminated time limits (e.g., Botwinick, 1977; Robins Wahlin et al., 1993), but age and other interindividual differences remain. In everyday life, however, compensatory behaviors become essential for the functioning of the individual. Performance is better when time is less of a restriction.

(ii) Verbal abilities and language

The spontaneous use of language, as in unrestricted speech, typically reveals experience-based learning and thereby cohort differences and other sociodemographic characteristics. That is often evident in the communication across generations where younger and older persons may find themselves using different words, grammar, and syntax. In this sense, language serves as a "generation marker."

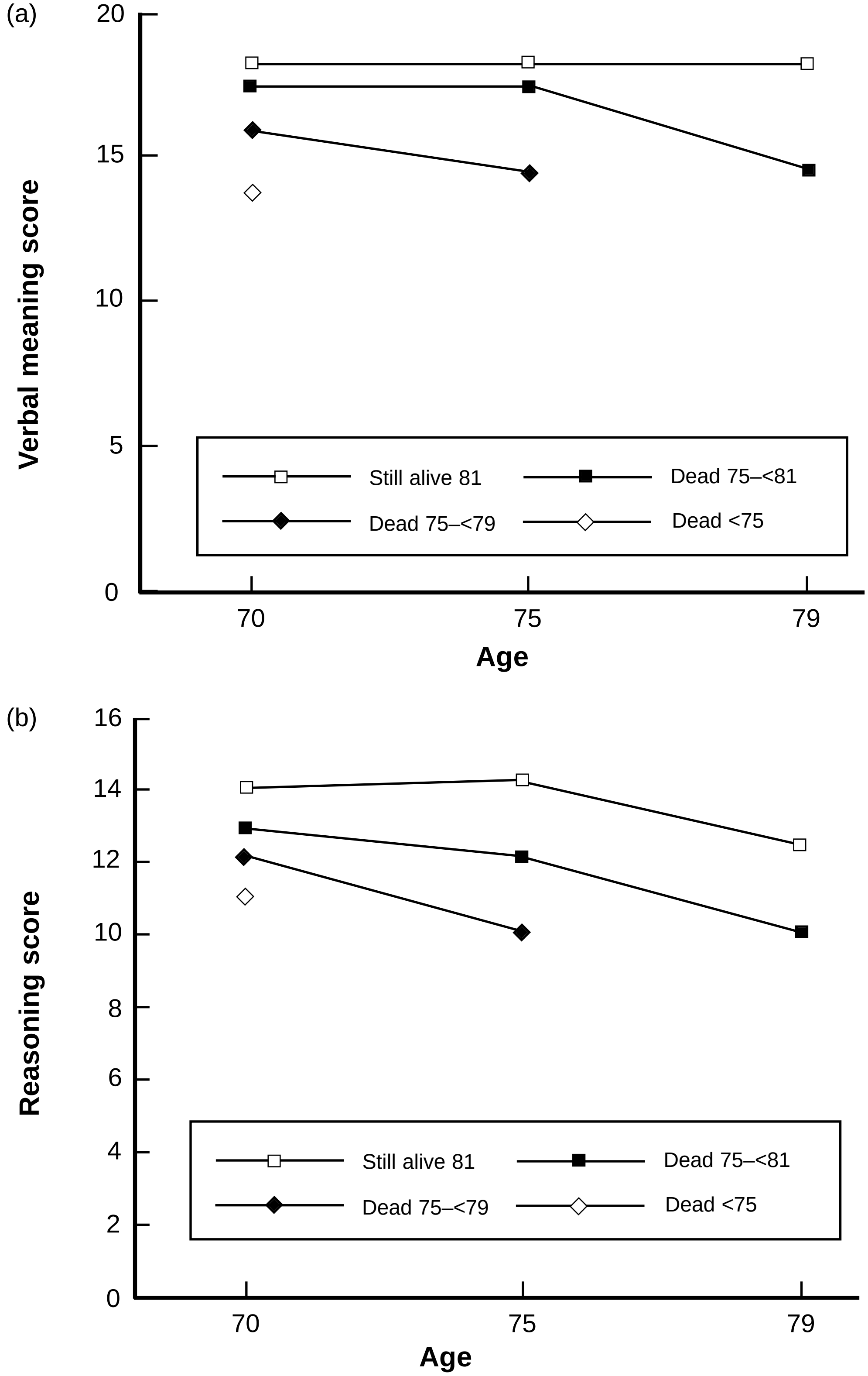

Figure 5 Longitudinal scores for a test of verbal meaning (a) and reasoning (b) showing terminal decline. ("Intelligence and terminal decline" by S. Berg, 1987. In G. L. Maddox and E. W. Busse (Eds.), *Aging: The universal experience* (pp. 414–415), New York: Springer. Copyright 1987 by Springer Publishing Company. Adapted with permission).

Besides speech comprehension, language also involves the domains of reading and writing. All these aspects of language are likely to be affected by sensory problems prevalent in the aged, including difficulties in visual acuity and hearing (presbycusis). In writing, hand motor ability also becomes important. The influences of these peripheral sources of impairment must be considered in evaluating potential cognition-related deficits in language. Despite sensory problems, language abilities are generally well preserved at the level of communication in aging (e.g., Bayles, Kaszniak, & Tomeda 1987). Furthermore, certain language difficulties may not become manifest because individuals may compensate for minor difficulties in unrestricted speech. Difficulties can be covered through the use of compensatory strategies, for example, simplifications or pauses filled with "Ums."

Language abilities require components of linguistic and lexical knowledge (vocabulary), word finding, and verbal fluency. The knowledge of linguistic structure, including syntax, seems to be a well-preserved ability in aging. In an extensive study by Emery (1985) on various aspects of language in younger individuals, normal elderly, and elderly individuals with dementia, results indicated that normal elderly people had greater difficulties with more complex syntactic structure. In the demented sample all aspects of language were affected (for details see Emery, 1985).

Vocabulary, or verbal meaning and verbal fluency have been extensively studied and tests of these aspects are often also included in comprehensive batteries. The general findings for vocabulary demonstrate continuity, whereas speeded fluency tests show lower performance in elderly individuals. In this sense, fluency becomes yet another marker of speed of behavior (Obler & Albert, 1985).

The ability of naming has been investigated using the tip-of-the-tongue (TOT) phenomenon. Studies generally show that younger individuals are more advantaged in their access to the lexical network needed for accurate naming of objects (Burke, Mackay, Worthley, & Wade, 1991). The close relationship between accuracy and speed, previously discussed, is also observed in word-finding tasks, where accurate speed tends to become a prerequsite for the primary naming. Difficulties in finding the name of objects may be described as a memory failure by subjects themselves. The literature indeed shows that memory mechanisms are involved in TOTs, although this involvement is more pronounced in the complex usage of language abilities, as for example in discourse (Burke, Worthley, & Martin, 1988).

While relatively robust in primary aging, language and verbal abilities are prone to cerebral lesions in dementia, where the "aphasia-agnosia" syndrome constitutes a diagnostic marker. Available information on various aspects of language are still relatively sparse, compared to other cognitive functions in aging (see Kemper, 1992; Light, 1988).

(iii) Visuospatial and visuoconstructive abilities

The components involved in visuospatial and visuoconstructive abilities include pattern recognition and the understanding of dimensionality. Reading a clock face, telling the time and setting the clock hands on certain times, reading maps and drawing maps represent some ecological examples of this cognitive domain.

The block design test in the WAIS battery is often thought of as a prototype test for visuospatial ability. The results from this test show worse performance in older individuls. Robins Wahlin et al. (1993) used block design and a clock test to study visuospatial functioning and spatial orientation in a healthy sample ranging from 75 to 96 years. In block design, they found significant age differences in which the youngest individuals showed superior performance. Interestingly, the test was administered both in the standard format including fixed time limits and in a format without time limits. The differences across ages remained in the no-time-limit condition, although performance levels were increased. In examining performance for each individual design, they found age effects only for the most complex designs. In the clock test, encompassing the tasks of setting hands on preprinted clock faces and time telling, they also found a slight but significant age effect in both. The oldest, aged 90 to 96 years, were less accurate than individuals in the age range of 75–89 years. It was concluded that there is an age-related deterioration in visuospatial abilities that continues into very old age.

Visuoconstructive tasks include components of perception, motor behavior, and spatial orientation. Such tasks are often included in clinical batteries, for example, in tasks where subjects are asked to copy figures at different levels of complexity. Such tests also provide qualitative information, including type of errors, that may be useful. The findings from studies on visuoconstruction generally show that the copying ability is well preserved in aging when motor and sensory difficulties are taken into account (e.g., Van Gorp, Satz, & Mirushina, 1990).

(iv) Problem solving and executive functions

A complex domain of cognitive functioning involves problem solving and logical reasoning, including inductive as well as deductive components. Problem solving encompasses all behaviors executed when facing old problems that we have learned how to solve, as well as novel problems requiring reorganization of already established modes of thinking and acting. Problem solving includes concrete as well as abstract thinking, and therefore constitutes a broad category of intelligent behavior.

Various tests of reasoning and problem solving are often included as subtests in comprehensive test batteries. For example, in Schaie's longitudinal Seattle study (Schaie, 1995), estimates of age changes in reasoning show a general decline with age. A more marked decline was found for those in their late sixties. This pattern resembles that of other abilities, suggesting that this aspect of problem solving is equally prone to aging as other cognitive abilities. In the Swedish H70 study, inductive reasoning remained remarkably stable across times of measurement in individuals with longer survival. That is, when retrospectively excluding those individuals who died after a certain occasion in the series of measurements, those who survived to the next occasion in the series showed an almost stable performance in reasoning, an ability often considered sensitive to age. This finding suggests that decline is more closely associated with subsequent mortality, rather than with age *per se* (see Berg, 1996, and Figure 5).

A well-known reasoning test is that of Raven's Progressive Matrices (Raven, 1982), including the original and the coloured matrices versions. Subjects are instructed to find a principle among certain patterns and to choose a pattern that corresponds to this principle in identifying a missing pattern. Significant age differences suggest lower performance with age in the original version of the test (see Lezak, 1983, 1995; Salthouse, 1983). This finding was also confirmed in a study by Denney and Heidrich (1990), who interestingly employed a simple training strategy before administering the test. Following training, equal improvements were found for young, middle-aged, and elderly individuals.

Another aspect of intelligent behavior is that of executive functions which refers to abilities that successfully enable a person to engage in "independent, purposive, self-serving behavior" (Lezak, 1983, p. 38). Questions about whether people will execute a certain behavior and if so, how they are doing it, capture the dimension of executive functioning. Lezak (1983) articulated four components of executive activity-related behaviors: goal formulation, planning, carrying out goal-directed plans, and effective performance. Given these components the question arises regarding how to evaluate the aspects of "whether" and "how." The interest in executive functions emerged in neuropsychology and is largely associated with interest in cerebral dysfunction, especially with dementia and frontal lobe lesions (e.g., Moss, Albert, & Kemper, 1992; Shallice & Burgess, 1991)

Symptoms like distractability, passivity and apathy, apraxia, and perseverations, affecting self-care behaviors and leading to impaired activities of daily living have often been the focus. In patients suffering from frontal lobe lesions, impairments become manifest in a more obvious manner. Executive functioning is, however, usually maintained in primary aging although affected by a slower speed of behavior. The complexity of the executive dimension requires close scrutinity of sensory and motor handicaps, and of other noncognitive factors that potentially may influence the individual's adaptation to demands in everyday life as well as performance on cognitive tests (LaRue, 1992).

The generation of alternate behaviors necessary to overcome and compensate for deficits in certain domains can be seen as an indicator of the flexibility aspect of executive functioning. Cognitive flexibility, or the ability to shift strategy in order to solve and overcome certain problems, is an essential component in many tests, as well as in real-life situations. Rigidity, the converse of flexibility, is a more visible marker for the lack of flexibility in the ecological context. Tests that require switching from one strategy to another or tasks in which overlearned responses are to be inhibited in favour of new responses generally show superior performance in younger individuals. The Wisconsin Card Sorting Test (WCST, Grant & Berg, 1948) is another test requiring cognitive flexibility. Subjects are here instructed to organize cards according to different principles, for example, color and form. Only slight age differences were found in a study of healthy, well-educated elderly individuals' performance on a modified WCST version (Haaland, Vranes, Goodwin, & Garry, 1987). The study provides support for the conclusion that executive functions in primary aging remain relatively unaffected.

Although verbal fluency and certain tests of construction and motor performance can serve as markers of executive functioning (Royall, Mahurin, & Gray, 1992; see also Lezak, 1995), noncognitive influences must be taken into account in evaluating performance. A powerful approach for examining executive functioning is the observation of naturalistic behaviors interpreted in the context of previous ability, sensory and motor functioning. An elderly person unwilling or unable to perform as before and exhibiting lower cognitive flexibility has to be assessed in a broad context, including information on coping, perceived personal control, and well-being, before any definite conclusions can be drawn about deficient executive functioning.

7.02.3.2.3 Summary

In the intelligence test tradition, the most frequently used battery is the WAIS. The subtests can be subdivided into "hold tests," assumed to measure crystallized intelligence, and "don't hold tests" measuring fluid abilities. Crystallized abilities are relatively stable across the entire adult age span, and some may improve even into old age. Fluid abilities, on the other hand, are more sensitive to age, as well as to various diseases. The findings of terminal decline (tertiary aging) in cognitive abilities demonstrate the close interrelationships between brain function and vital physiology in aging.

7.02.4 COPING WITH COGNITIVE CHANGE IN AGING: PERSPECTIVES OF THE AGING INDIVIDUAL AND THE GEROPSYCHOLOGIST

Some implications of reported findings on memory and other cognitive aspects of aging will be discussed in this section, including considerations relevant to the geropsychologist.

7.02.4.1 "From Lab to Life"

The research literature on memory and other cognitive functions conveys the general conclusion that decline and change become more likely with advancing age, although the age-related sensitivity differs for various tasks. At the same time we usually find that most elderly people maintain independence and successful functioning in everyday life even though this involves numerous tasks that require complex memory and cognitive abilities. These observations seem contradictory and address questions about the significance of reported findings on memory and cognitive changes in aging, revealed in the experimental or clinical setting. These different portrayals of aging also become evident to the geropsychologist. Salthouse (1983) suggested four explanations for the discrepancy between "lab and life": individual differences, ecological validity, cognitive demands in terms of complexity and difficulty, and everyday life experiences and learning.

7.02.4.1.1 Individual differences

The individual differences issue is a neglected aspect in the prevailing normative tradition in the gerosciences. The use of chronological age, as a proxy marker for aging, and the arithmetic mean as the typical statistic for describing performances, tends to emphasize general patterns over unique features, including that of interindividual differences.

In everyday life we may face individuals in advanced ages who show better performance than they did some years ago. The recovery from certain diseases and treatment effects are also seen in advanced ages. The normative design is, however, unlikely to show both sides of the coin, that is, while some individuals decline, others remain stable or even improve their performance over time. Besides generation and cohort differences, individuals of the same age typically manifest great differences in physical as well as psychological functioning. Elderly people are distributed between the extremes of those who experience early physical and cognitive impairments and, at the other end of the continuum, the successful agers who maintain their performance throughout life. This notion is important because it provides a basic frame of reference for assessment, as well as an accurate portrayal of aging.

An interindividual perspective is an essential element for the geropsychologist. The research literature provides the general norms for late-life functioning, and at best adequate norms for various tests. An assessment requires, however, detailed information about previous functioning of the individual. Such information, or more likely assumptions about "baseline performance" in earlier days, should provide the ultimate points of reference for the geropsychologist.

7.02.4.1.2 Ecological validity

Ecological validity refers to the congruence between laboratory and clinical tests, and everyday life tasks requiring memory and other cognitive resources. Although certain experimental tasks may be considered as esoteric, they surely activate cognitive subprocesses and components of relevance for life outside the laboratory. The overall aim in experimental aging and cognition studies is the understanding of age-related change in those subprocesses and components. In this sense, the complexity of everyday life tasks can never correspond to requirements in the laboratory setting. Although the range of available tests covers many dimensions of memory and cognitive functioning, it is important to remember that they still represent only a sample of markers for various cognitive processes.

Another aspect of ecology relates to compensatory behaviors that tend to be neglected in the laboratory. In everyday life individuals can compensate for weaknesses in certain abilities by profiting from strengths in other areas of functioning. In this sense difficulties may not become manifest in natural settings. In tests, however, individuals are usually inhibited from using compensatory strategies to overcome weaknesses. Impairments are then more likely to be evinced.

The "lab–life dimension" represents an important issue for the geropsychologist when confronted with an individual showing impaired test performance, but demonstrating relatively intact functioning in everyday life. The detection of cognitive problems on certain tests may be due to disease processes (secondary aging) and provide arguments for interventions aimed at retarding or improving performance. Low performance on tests is also indicative of the cognitive expenditure that is needed to manifest "intact" everyday functioning.

7.02.4.1.3 Complex cognitive processing

The notion that cognitive demands in everyday life are less complex and difficult than that of laboratory and clinical tests deserves attention. One might argue that the demands of the two situations are not identical (e.g., Baltes & Willis, 1979). Laboratory and clinical tests of memory and cognition are usually designed in accordance with psychometrics. This means that level of difficulty is deliberately introduced as a variable for evaluating performance. Everyday life is often highly predictable through recurrent tasks and events with which we have learned to cope. Tests of memory and cognition are also likely to introduce novel tasks, requirements, and a setting that is unfamiliar to the individual.

The familiarity with psychological testing differs across age cohorts. The oldest old tend to have fewer such experiences. Recommendations of "age-fairness" should be evaluated against this background. In addition, in real life people tend to seek those environments that fit their interests and available resources, including those that relate to cognition. This is a matching of abilities and task demands that does not occur in the laboratory or in the clinical test situation. In evaluating an individual's performance, geropsychologists have to consider how a task should be presented, how the instructions are administered, including feedback to the individual. Motivation and other noncognitive factors are also important. The importance of instructions and practice before administering a certain task is well known to the experienced clinician. Standard tests and batteries usually allow only brief instructions. Difficulties experienced by elderly people in comprehending these instructions, for example, in abstract tasks, are likely to interfere with later performance. Ensuring that the individual understands task requirements can therefore significantly improve performance (e.g. Baltes, Sowarka & Kliegl, 1989).

7.02.4.1.4 Everyday life experiences and learning

Although research on memory and cognition in aging typically confirms a deterioration view, the resources in most individuals are sufficient to cope with environmental demands of daily life. When the magnitude of decline exceeds the ecological demands it becomes a manifest deficit. For example, in the diagnosis of dementia, the impairment in memory and other cognitive abilities must interfere with daily life functioning according to diagnostic criteria (American Psychiatric Association, 1987, 1994). Many activities and tasks in everyday life are overlearned by their being practiced over and over again. As noted by Salthouse (1983), "since increased age in adulthood is generally positively correlated with experience, many highly practiced activities may be maintained at a constant level despite age-related declines in component ability because of the compensating effects of greater experience" (p. 203).

In experimental and clinical settings, memory and cognitive tasks are typically novel and the cues and instructions provided do not necessarily correspond to those available in natural, everyday life. Performance is then more unlikely to rely on previous experiences. The ability for adequate everyday functioning is also affected by more or less subtle social and cognitive support systems that constitute natural ingredients in daily-life activities (use of a calendar, reminders, and other external aids). In addition, individuals also tend to withdraw from activities that are viewed as too demanding, using excuses such as that they have lost interest (see Baltes's model below). The geropsychologist has to consider and be aware of the mechanisms of the fit between environmental demands and cognitive capabilities of the individual.

7.02.4.1.5 Current knowledge and extralaboratory cognition

As pointed out by Salthouse (1983), and by recent proponents of testing memory and cognition in daily-life contexts (e.g., Willis, 1996), "the greatest hindrance to knowledge about the relationships between laboratory performance and extralaboratory competence is simply lack of detailed information about the nature of real-world activities" (Salthouse, 1983, pp. 203–204). Clearly, more attention needs to be given to the external validity of results obtained in the laboratory or clinical setting. Improved means of evaluating everyday performances constitutes one avenue for this. Much of the literature on aging and cognitive functions can be divided into three parts on the basis of methodology. Research is conducted in the psychometric tradition, in the experimental tradition, or in the ecological tradition. It seems clear that there are possible gains by cross-fertilization between approaches. Indeed such efforts are not absent. One approach includes bringing everyday activities such as television watching into the laboratory (e.g., Stine, Wingfield, & Myers, 1990) for subsequent tests of, for example, memory. Another logical possibility is to perform clinical or laboratory-like tasks in the subject's everyday world. Weber, Brown, and Weldon (1978), conducted such a study where elderly residents in a nursing home were tested on spatial knowledge and memory for scenes and locations in their nursing home. An even broader approach may be warranted, taking into account such things as collaborative support provided by a spouse (Gould & Dixon, 1993; Gould, Trevithick, & Dixon, 1991). On the other hand, proponents of an applied approach aiming at intervention and adjustment of the environment argue that researchers have now developed sufficient understanding of older adults' cognitive systems to make it possible to improve the quality of life for older adults (see Park, 1992). The latter view relies mainly on knowledge about cognitive processes, as well as documented strengths and weaknesses of the elderly in a broader sense.

7.02.4.1.6 Cognitive aging and the individual: an integrative model

A comprehensive model for the understanding of primary cognitive aging in the real-life context that integrates much of the literature on cognitive aspects of aging was proposed by Baltes (1991). The model (see Figure 6) is based on a series of propositions. These propositions, or general statements, are to some extent already covered in previous sections of the present chapter. Briefly, it is assumed that aging is characterized by much interindividual variability; it is important to distinguish between pathological, normal, and optimal aging; there is much developmental reserve capacity in old age; aging is characterized by a loss of fluid intelligence; knowledge and crystallized abilities may compensate for the losses in fluid abilities; with aging the balance between gains and losses becomes less positive or even negative; and the self in old age continues to be a powerful system of coping and maintaining integrity.

Against the background of these general conclusions, some adaptional mechanisms were suggested for successful psychological aging. The model focuses on the dynamic interplay among selection, optimization, and compensation, all components in a general process of adaption, effective throughout life. However, because of the losses experienced in aging, whether biological, social, or mental, the model is especially applicable to elderly persons. By selection of activities from a previously more extensive repertoire of behaviours, maximizing the effort spent on maintaining certain activities by means of practice or acquisition of new knowledge, and finally fostering of compensatory behavior, performance may be maintained at high levels even in old age.

Without going into details, some conclusions apply to the present chapter. The model has obvious implications for the geropsychologist. In the clinical setting, a wide variety of functions or abilities are routinely tested. Since the purpose of an investigation is usually to arrive at a diagnosis, the focus is often on deficits and weaknesses. Planning for training or rehabilitation, on the other hand, also requires closer examination of strengths and the potential for improvements. However, with reference to the discussion about everyday cognitive performance, it may be justifiable to include a routine for evaluation of cognitive reserves and abilities to adopt compensatory behavior. The impact of interindividual variability is strongly acknowledged in Baltes's model as interindividual variability in old age applies not only to behavioral functioning, but also to indicators of biomedical functioning (Costa & Andres, 1986; Rowe & Kahn, 1987). This variability may in combination with other critical components, such as expertise in some areas, produce differential patterns in aging. Hence, both strengths and weaknesses are equally critical components in a full understanding of the individual's general level of functioning.

7.02.4.1.7 Cognitive reserves and plasticity

Some additional comments on cognitive reserve capacity may be warranted. Following the Baltes (1991) model, potential for cognitive improvement may be discussed in terms of baseline reserve capacity as well as developmental reserve capacity (see Kliegl & Baltes, 1987). Baseline reserve capacity refers to

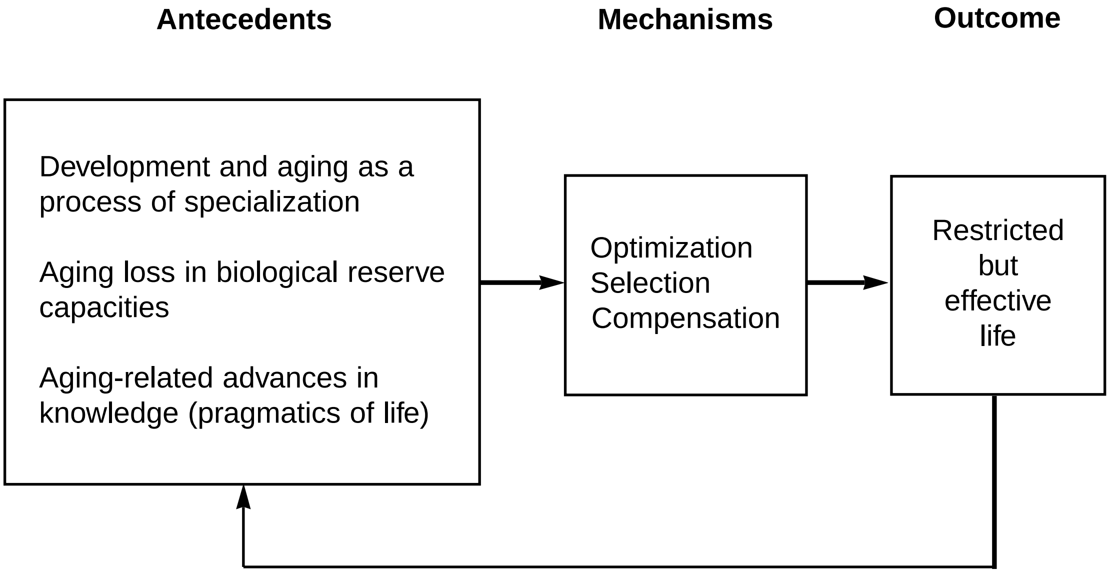
Antecedents
Mechanisms
Outcome
Development and aging as a process of specialization
Aging loss in biological reserve capacities
Aging-related advances in knowledge (pragmatics of life)
Optimization
Selection
Compensation
Restricted but effective life

Figure 6 Baltes's selective optimization with compensation ("The many faces of human ageing: Toward a psychological culture of old age" by P. B. Baltes, 1991, *Psychological Medicine, 21,* p. 85. Copyright by Cambridge University Press. Adapted with permisssion).

variations in performance due to varying levels of task demands and support, while developmental reserve capacity denotes the range of performance under extended periods of development or practice. Reserve capacity is related to such diverse examples as the acquisition of expertise, testing the limits of an individual's capacity, and memory training.

7.02.4.2 General Assessment Guidelines

Some guidelines in the assessment of elderly persons can be extracted from previous discussions. These general guidelines by no means constitute a complete coverage of all topics. Also, advice in relation to specific diagnostic issues is not dealt with. Certain suggestions are self-evident, but nevertheless important to address explicitly.

Throughout the chapter, it has been emphasized that the level of cognitive functioning is under multifactorial influence and that interindividual variability is substantial in aging. These notions have obvious implications for the geropsychologist. Chronological age has traditionally been the sole stratification variable in the psychometric test tradition. Using test norms to aid in the interpretation of performance ought to be considered in this context. Further, the distinction between primary and secondary aging requires the integration of information from several sources, including that of physical health and social life.

Before meeting the client, we would strongly recommend the geropsychologist to review medical records. Although literature on the relation between cognition and disease is scarce for illnesses in old age other than dementia or major depression, it may be worthwhile to look for the presence of, for example, cardiovascular disease including elevated blood pressure, hyper- or hypothyroidism, diabetes, epilepsy, vitamin deficiency, previous transient ischemic attacks or stroke. The medical literature also lacks, in general, age-graded norms for the elderly, for example, age-adjusted cut-offs for laboratory tests such as the test of folic acid. (For details, see the literature listed in the previous section on secondary aging.)

In the assessment setting it is also strongly recommended to check for vision and hearing. Does the client need certain glasses or a hearing aid? Secure the use of these aids, and, when vision or hearing is poor, consider using adjusted test materials.

Testing should start with a brief interview. During this conversation, much information crucial to later interpretation of test results may be obtained. The geropsychologist should, in addition to basic information concerning education and prior work or career, also collect data about previous and present skills and interests. When necessary, a relative may provide valuable information (e.g., Jorm et al., 1996). The client's general feeling of being in control of daily life as well as beliefs about their own capacity may be informative in evaluating subsequent test performance.

Motivational factors play an important role in the assessment setting. Therefore, one should explain carefully the purpose of the assessment and make sure of the client's comprehension. At this stage, it is possible that the geropsychologist has obtained a clinical impression of the client's general level of functioning. This is important for selecting the relevant tests and for avoiding floor and ceiling effects.

When introducing tests to the client, it may be worthwile to vary task demands systematically in order to investigate reserve capacity. Also, tasks may be varied in terms of time restraints. It is generally advised that tasks be selected in terms of age-fairness.

As noted, in testing older clients, norms are often not entirely reliable. In addition to problems concerning sample selections for these norms, they seldom cover the age range above 80 years. Most studies portray, however, a continuous decline of cognitive performance in the oldest, although performance may remain at relatively high levels. In the absence of norms for the very old and the introduction of other stratification variables than age, it becomes advisable to interpret raw scores carefully.

7.02.4.3 Future Prospects and Cohorts

In the years to come, geropsychologists will meet new cohorts of elderly people who are exposed to other experiences and influences than the elderly people of today. Future clients are likely to require efforts to promote independence and growth in aging, also for aspects of life related to memory and cognition. Future cohorts may require somewhat different diagnostic considerations than today and age-fair testing may refer to something different than today. Improvements in the educational system are likely to exert an impact on performance in new cohorts of elderly people. Furthermore, it may be that successful medical treatments for conditions such as Alzheimer's disease are introduced. In that case, geropsychological expertise may require improved methods for early, preclinical, detection of the disease.

Given increasing interest in geropsychological matters, new test norms adjusted for individuals in the range 80–100 years will hopefully be available to future cohorts of clinicians. The recognition of interindividual differences may end in norms that do not apply chronological age as the sole stratification variable for differentiation within normal and pathological aging. Educational level and gender are two probable candidates among other critical variables to be considered.

The impact of education may seem somewhat far-fetched, as formal education is quite distant in time for elderly people. Education is, however, related to verbal ability and strategy knowledge (e.g., Flavell & Wellman, 1977; Gillund & Perlmutter, 1988), both of which are prerequisites for successful performance on many memory and cognitive tasks.

Norms adjusted for gender are, to a certain extent, already available. Generally, women perform better than men on verbal memory tasks (e.g., Larrabee & Crook, 1993; Schaie & Willis, 1993; Zelinski, Gilewski, & Schaie, 1993). Women also appear to perform at a higher level on face recognition tasks (see Shepherd, 1981, for review). Social (e.g., McKelvie, Standing, St. Jean, & Law, 1993) as well as biological (e.g., Gur et al., 1991) explanations have been proposed for these differences.

Aging research is being conducted around the globe and there is a continuous stream of reports and an increasing wealth of information from the psychological, medical, and social gerosciences. Many projects are designed with an interdisciplinary approach which opens potential for cross-fertilizations and analyses of complex interactions among biological, social, and psychological variables. In cognitive aging research, structural modeling techniques are now applied more often in an effort to identify critical components for successful performance (see Salthouse, 1991). Although, not theory based, the approach truly recognizes the complexity of various aging processes.

The complexity of aging processes becomes apparent for the geropsychologist in the evaluation of performance and gathering of other information in a comprehensive assessment of memory and other cognitive abilities. The challenge is to end up with appropriate conclusions for diagnosis and treatment. For this, geropsychological expertise is indispensable.

7.02.5 REFERENCES

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. rev.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 89–195). New York: Academic Press.

Bäckman, L. (1991). Recognition memory across the adult life span: The role of prior knowledge. *Memory & Cognition, 19,* 63–71.

Bäckman, L. (1995). Cognitive support and episodic memory functioning in normal aging and Alzheimer's disease: A continuity view of the modifiability of memory performance in old age. In R. Cacabelos, H. Frey, T. Nishimura, & B. Winblad (Eds.), *Neuroger-*

ontology and neurogeriatrics (pp. 85–106). Barcelona, Spain: Prous Science.

Bäckman, L., & Forsell, Y. (1994). Episodic memory functioning in a community-based sample of very old adults with major depression: Utilization of cognitive support. *Journal of Abnormal Psychology, 103,* 361–370.

Bäckman, L., & Larsson, M. (1992). Recall of organizable words and objects in adulthood: Influences of instructions, retention interval, and retrieval cues. *Journal of Gerontology: Psychological Sciences, 47,* P273–P278.

Bäckman, L., Mäntylä, T., & Herlitz, A. (1990). The optimization of episodic remembering in old age. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 118–163). New York: Cambridge University Press.

Bäckman, L., & Wahlin, Å. (1995). Influences of item organizability and semantic retrieval cues on episodic recall in very old age. *Aging and Cognition, 2,* 312–325.

Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Oxford University Press.

Baddeley, A. D. (1990). *Human memory: Theory and practice*. Boston: Allyn and Bacon.

Bale, R. N. (1973). Brain damage in diabetes mellitus. *British Journal of Psychiatry, 122,* 337–341.

Baltes, M. M., & Baltes, P. B. (Eds.) (1986). *The psychology of control and aging*. Hillsdale, NJ: Erlbaum.

Baltes, P. B. (1991). The many faces of human ageing: Toward a psychological culture of old age. *Psychological Medicine, 21,* 837–854.

Baltes, P. B., & Kliegl, R. (1992). Further testing of limits of cognitive plasticity: Negative age differences in a mnemonic skill are robust. *Developmental Psychology, 28,* 121–125.

Baltes, P. B., Sowarka, D., & Kliegl, R. (1989): Cognitive training research on fluid intelligence in old age: What can older adults achieve by themselves? *Psychology and Aging, 4,* 217–221.

Baltes, P. B., & Willis, S. L. (1979). The critical importance of appropriate methodology in the study of aging: The sample case of psychometric intelligence. In F. Hoffmeister & C. Muller: (Eds.), *Brain function in old age* (pp. 164–187). Berlin: Springer-Verlag.

Barclay, L., Weiss, E., Bond, O., & Blass, J. (1988). Unrecognized cognitive impairment in cardiac rehabilitation patients. *Journal of the American Geriatrics Society, 36,* 22–28.

Bayles, K. A., Kaszniak, A. W., & Tomeda, C. K. (1987). *Communication and cognition in normal aging and dementia*. Boston: College-Hill/Little, Brown & Co.

Berg, C. A., & Sternberg, R. J. (1985). A triarchic theory of intellectual development during adulthood. *Developmental Review, 5,* 334–370.

Berg, C. A., & Sternberg, R. J. (1992). Adults' conceptions of intelligence across the adult life span. *Psychology and Aging, 7,* 221–231.

Berg, S. (1987). Intelligence and terminal decline. In G. L. Maddox & E. W. Busse (Eds.), *Aging: The universal experience* (pp. 414–415). New York: Springer.

Berg, S. (1996). Aging, behavior, and terminal decline. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 323–337). San Diego, CA: Academic Press.

Berg, S., & Dellasega, C. (1996). The use of psychoactive medications and cognitive function in older adults. *Journal of Aging and Health, 8,* 136–149.

Birren, J. E., & Cunningham, W. R. (1985). Research on the psychology of aging: Principles, concepts and theory. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 3–34). New York: Van Nostrand Reinhold.

Birren, J. E., & Morrison, D. F. (1961). Analysis of the WAIS subtests in relation to age and education. *Journal of Gerontology, 16,* 363–369.

Birren, J. E., & Schaie, K. W. (1996). *Handbook of the psychology of aging* (4th ed.). San Diego, CA: Academic Press.

Blanchard-Fields, F., & Abeles, R. P. (1996). Social cognition and aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 150–161). San Diego, CA: Academic Press.

Botwinick, J. (1977). Intellectual abilities. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 580–605). New York: Van Nostrand Reinhold.

Botwinick, J. (1978). *Aging and behavior: A comprehensive integration of research findings*. New York: Springer.

Brody, J. A., & Schneider, E. L. (1986). Disease and disorders of ageing: A hypothesis. *Journal of Chronic Disorders, 39,* 871–876.

Bruce, D. (1991). Mechanistic and functional explanations of memory. *American Psychologist, 46,* 46–48.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language, 30,* 542–579.

Burke, D. M., Worthley, J., & Martin, J. (1988). I'll never forget what's-her-name: Aging and the tip of the tongue experience. In M. M. Gruneberg, P. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 2, pp. 113–118). Chichester, UK: Wiley.

Burt, D. B., Zembar, M. J. & Niederehe, G. (1995). Depression and memory impairment: A meta-analysis of the association, its pattern, and specificity. *Psychological Bullentin, 117,* 285–305.

Chiarello, C., & Hoyer, W. J. (1988). Adult age differences in implicit and explicit memory: Time course and encoding effects. *Psychology and Aging, 3,* 358–366.

Cipioli, C., Neri, M., Andermarcher, E., Pinelli, E., & Lalla, M. (1990). Self-rating and objective memory testing of normal and depressed elderly. *Aging, 2,* 39–48.

Cockburn, J. & Smith, P. T. (1991). The relative influence of intelligence and age on everyday memory. *Journal of Gerontology: Psychological Sciences, 46,* P31–P36.

Cohen, G., & Faulkner, D. (1989). Age differences in source forgetting: Effects on reality monitoring and on eyewitness testimony. *Psychology and Aging, 4,* 10–17.

Cornelius, S. W. (1990). Aging and everyday cognitive abilities. In T. Hess (Ed.), *Aging and cognition: Knowledge, organization and utilization* (pp. 411–444). Amsterdam: North-Holland.

Costa, P. T., Jr., & Andres, R. (1986). Patterns of age changes. In I. Rossman (Ed.), *Clinical geriatrics* (pp. 23–30). Philadelphia: Lippincott.

Craik, F. I. M., & Jennings, J. M. (1992). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 51–110). Hillsdale, NJ: Erlbaum.

Craik, F. I. M., & Salthouse, T. A. (Eds.) (1992). *Handbook of aging and cognition*. Hillsdale, NJ: Erlbaum.

Cremer, R., & Zeef, E. J. (1987). What kind of noise increases with age? *Journal of Gerontology, 42,* 515–518.

Crook, T. H., & Larrabee, G. J. (1992). Changes in facial recognition memory across the adult life span. *Journal of Gerontology: Psychological Sciences, 47,* P138–P141.

Crum, R. M., Anthony, J. C., Basset, S. S., & Folstein, M. F. (1993). Population-based norms for Mini-Mental State Examination by age and educational level. *Journal of the American Medical Association, 269,* 2386–2391.

Denney, N. W. (1989). Everyday problem solving: Methodological issues, research findings, and a model. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 330–351). New York: Cambridge University Press.

Denney, N. W., & Heidrich, S. M. (1990). Training effects on Raven's Progressive Matrices in young, middle-aged, and elderly adults. *Psychology and Aging, 5,* 144–145.

Dikman, S., & Matthews, C. G. (1977). Effect of major motor seizure frequency upon cognitive-intellectual functions in adults. *Epilepsia, 18,* 21–29.

Dureman, I., & Sälde, H. (1959). *Psykometriska och experimental-psykologiska metoder för klinisk tillämpning*. Uppsala, Sweden: Almqvist & Wiksell.

Elder, G. H. (1979). Historical change in life patterns and personality. In P. B. Baltes & O. G. Brim (Eds.), *Life-span development and behavior* (Vol. 2, pp. 117–159). New York: Academic Press.

Elias, M. F., Wolf, P. A., D'Agostino, R. B., Cobb, J., & White, L. R. (1993). Untreated blood pressure level is inversely related to cognitive functioning: The Framingham Study. *American Journal of Epidemiology, 138,* 353–364.

Emery, O. B. (1985). Language and aging. *Experimental Aging Research, 11,* 3–60. Special Monograph Issue.

Epstein, F. H., Ostrander, L., Johnson, B. C., Payne, M. W., Hayner, N. S., Keller, J. B., & Francis, T., Jr. (1965). Epidemiological studies of cardiovascular disease in a total community: Tumseh, Michigan. *Annals of Internal Medicine, 62,* 1170–1187.

Farmer, M. E., Kittner, S. J., Abbott, R. D., Wolz, M. M., Wolf, P. A., & White, L. R. (1990). Longitudinally measured blood pressure, antihypertensive medication use, and cognitive performance: The Framingham Study. *Journal of Clinical Epidemiology, 43,* 475–480.

Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. V. Kail, Jr. & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 3–34), Hillsdale, NJ: Erlbaum.

Folstein, M. F., Folstein, S. E., McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive status of patients for the clinican. *Journal of Psychiatric Research, 12,* 189–198.

Fozard, J. L., Metter, E. F., & Brant, L. J. (1990). Next steps in describing aging and disease in longitudinal studies. *Journal of Gerontology: Psychological Sciences, 45,* P116–P127.

Fratiglioni, L., Grut, M., Forsell, Y., Viitanen, M., Grafström, M., Holmén, K., Eriksson, K., Bäckman, L., Ahlbom, A., & Winblad, B. (1991). Prevalence of Alzheimer's disease and other dementias in an elderly urban population: Relationship with age, sex and education. *Neurology, 41,* 1886–1892.

Fries, J. F., & Crapo, L. M. (1981). *Vitality and Ageing*. New York: Freeman.

Gallagher, D., Thompson, L., & Levy, S. (1980). Clinical psychological assessment of older adults. In L. W. Poon (Ed.), *Aging in the 1980s. Psychological issues* (pp. 19–40). Washington DC: American Psychological Association.

Gambert, S. R. (1988). Dementia secondary to metabolic and nutritional abnormalities. *Clinics in Geriatric Medicine, 4,* 831–839.

Gillund, G., & Perlmutter, M. (1988). Episodic memory and knowledge interactions across adulthood. In L. L. Light & D. M. Burke (Eds.), *Language, memory and aging* (pp. 191–208). Cambridge, UK: Cambridge University Press.

Gould, O. N., & Dixon, R. O. (1993). How we spent our vacation: Collaborative storytelling by young and old adults. *Psychology and Aging, 8,* 10–17.

Gould, O. N., Trevithick, L., & Dixon, R. O. (1991). Adult age differences in elaborations produced during prose recall. *Psychology and Aging, 6,* 93–99.

Graf, P. (1990). Life-span changes in implicit and explicit memory. *Bulletin of the Psychometric Society, 28,* 353–358.

Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of the degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem. *Journal of Experimental Psychology, 38,* 404–411.

Grisso, T. (1986). *Evaluating competencies: Forensic assessments and instruments*. New York: Plenum.

Grut, M., Forsell, Y., Hill, R. D., Bäckman, L., & Winblad, B. (1995). *The relationship between signs of cardiovascular deficiency and cognitive performance in old age*. Manuscript submitted for publication.

Gur, R. C., Mozley, P. D., Resnick, S. M., Gottlieb, G. E., Kohn, M., Zimmerman, R., Herman, G., Atlas, S., Grossman, R., Beretta, D., Erwin, R., & Gur, R. E. (1991). Gender differences in age affect on brain atrophy measured by magnetic resonance imaging. *Proceedings of the National Academy of Sciences, 88,* 2845–2849.

Haaland, K. Y., Vranes, L. F., Goodwin, J. S., & Garry, P. J. (1987). Wisconsin Card Sort Test performance in a healthy elderly population, *Journal of Gerontology, 42,* 345–346.

Hartley, A. A. (1989). The ecology of problem solving. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 300–309). New York: Cambridge University Press.

Hayslip, B., Jr., & Panek, P. E. (1993). *Adult development and aging* (2nd ed.). New York: HarperCollins.

Herlitz, A., Lipinska, B., & Bäckman, L. (1992). Utilization of cognitive support for episodic remembering in Alzheimer's disease. In L Bäckman (Ed.), *Memory functioning in dementia* (pp. 73–96). Amsterdam: Elsevier.

Hill, R. D., Wahlin, Å., Winblad, B., & Bäckman, L. (1995). The role of demographic and lifestyle variables in utilizing cognitive support for episodic remembering among very old adults. *Journal of Gerontology: Psychological Sciences, 50B,* P219–P227.

Horn, J. L. (1982). The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in adulthood. In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitve processes* (pp. 237–278). New York: Plenum.

Inouye, S. K., Albert, M. S., Mohs, R., Sun, K., & Berkman, L. F. (1993). Cognitive performance in a high-functioning community-dwelling elderly population. *Journal of Gerontolgy: Medical Sciences, 48,* M146–M151.

Johansson, B., Allen-Burge, R., & Zarit, S. H. (1997). Self-reports on memory functioning in a longitudinal study of the oldest old: Relation to current, prospective, and retrospective performance. *Journal of Gerontology: Psychological Sciences, 52B,* P139–P146.

Johansson, B., & Berg, S. (1992). The robustness of the terminal decline phenomenon: Longitudinal data from the digit span memory test. *Journal of Gerontology: Psychological Sciences, 44,* P184–P186.

Johansson, B. & Zarit, S. H. (1995). Prevalence and incidence of dementia in the oldest old: A longitudinal study of a population-based sample of 84–90 year-olds in Sweden. *International Journal of Geriatric Psychiatry, 10,* 359–366.

Johansson. B. & Zarit, S. H. (1997). Early cognitive markers of the incidence of dementia and mortality: A longitudinal population-based study of the oldest old. *Journal of Geriatric Psychiatry, 12,* 53–59.

Johansson, B., Zarit, S. H., & Berg, S. (1992). Changes in cognitive functioning of the oldest old. *Journal of Gerontology: Psychological Sciences, 47,* P75–P80.

Jorm, A. F., Christensen, H., Henderson, A. S., Jacomb, P. A., Korten, A. E., & Mackinnon, A. (1996). Informant ratings of cognitive decline in elderly people: Relationship to longitudinal change on cognitive tests. *Age and Ageing, 25,* 125–129.

Kahn, R. L., Zarit, S. H., Hilbert, N. M., & Niederehe, G. (1979). Memory complaint and impairment in the aged. The effect of depression and altered brain function. *Archives of General Psychiatry, 32,* 1569–1573.

Kane, R., & Kane, R. (1981). *Assessing the elderly: A* *practical guide to measurement*. Lexington, MA: Lexington.

Kaszniak, A. W., & Ditraglia Christenson, G. (1995). Differential diagnosis of dementia and depression. In M. Storandt & G. R. VandeBos (Eds.), *Neuropsychological assessment of dementia and depression in older adults: A clinician's guide* (pp. 81–117). Washington, DC: American Psychological Association.

Katona, C. L. E. (1994). *Depression in old age*. Chichester, UK: Wiley.

Kausler, D. H. (1994). *Learning and memory in normal aging*. San Diego, CA: Academic Press.

Kemper, S. (1992). Language and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 213–270). Hillsdale, NJ: Erlbaum.

Kliegl, R. L., & Baltes, P. B. (1987). Theory-guided analysis of mechanisms of development and aging through testing—The limits and research on expertise. In C. Schooler & K. W. Schaie (Eds.), *Cognitive functioning and social structure over the life course* (pp. 95–119). Norwood, NJ: Ablex.

Kohn, M. L., & Schooler, C. (1983). *Work and personality*. Norwood, NJ: Ablex.

Krauss, I. K. (1980). Between- and within-group comparisons in aging research. In L. W. Poon (Ed.), *Aging in the 1980s: Psychological issues* (pp. 542–551). Washington, DC: American Psychological Association.

Kuriansky, J., & Gurland, B. (1976). Performance test of activities of daily living. *International Journal of Aging and Human Development, 7,* 343–352.

Labouvie-Vief, G. (1982). Dynamic development and mature autonomy: A theoretical prologue. *Human Development, 25,* 161–191.

Labouvie-Vief, G. (1985). Intelligence and cognition. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of psychology and aging* (pp. 500–530). New York: Van Nostrand Reinhold.

Larrabee, G. J., & Crook, T. H. (1993). Do men show more rapid age-associated decline in simulated everyday verbal memory than do women? *Psychology and Aging, 8,* 68–71.

Larsson, M., & Bäckman, L. (1993). Semantic activation and episodic odor recognition in young and older adults. *Psychology and Aging, 8,* 582–588.

LaRue, A. (1992). *Aging and neuropsychological assessment*. New York: Plenum.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.

Lee, J. S. (1991). *Abstraction and aging*. New York: Springer.

Lerman, S. (1983). An experimental and clinical evaluation of lens transparency and aging. *Journal of Gerontology, 38,* 293–301.

Lezak, M. D. (1983). *Neuropsychological assessment* (2nd ed.). New York: Oxford University Press.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Light, L. L. (1988). Language and aging: Competence versus performance, In J. E. Birren & V. L. Bengtson (Eds.), *Emergent theories of aging* (pp. 177–213). New York: Springer.

Light, L. L. (1991). Memory and aging: Four hypotheses in search of data. *Annual Review of Psychology, 42,* 333–376.

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging, 9,* 339–355.

MacKay, D. G., & Burke, D. M. (1990). Cognition and aging: A theory of new learning and the use of old connections. In T. M. Hess (Ed.), *Aging and cognition: Knowledge organization and utilization* (pp. 213–264). Amsterdam: North Holland.

Marsiske, M., Lindenberger, U., & Baltes, P. B. (1994, April). *The dedifferentiation of intelligence hypothesis revisited: From early adulthood to very old age*. Paper presented at the Cognitive Aging Conference, Atlanta, GA.

McKelvie, S. J., Standing, L., St. Jean, D., & Law, J. (1993). Gender differences in recognition memory for faces and cars: Evidence for the interest hypothesis. *Bulletin of the Psychonomic Society, 31,* 447–448.

Mennemeier, M., Garner, R. D., & Heilman, K. M. (1993). Memory, mood and measurement in hypothyroidism. *Journal of Clinical and Experimental Neuropsychology, 15,* 822–831.

Moss, M. B., Albert, M. S., & Kemper, T. L. (1992). Neuropsychology of frontal lobe dementia. In R. F. White (Ed.), *Clinical syndromes in adult neuropsychology: The practioner's handbook*. Amsterdam: Elsevier.

Neisser, U. (1978). Memory. What are the important questions? In M. M. Gruneber, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 3–19). London: Academic Press.

Nelson, E. A., & Daneffer, D. (1992). Aged heterogeneity: Fact or fiction? The fate of diversity in gerontological research. *Gerontologist, 32,* 17–23.

Nyberg, L., Bäckman, L., Erngrund, K., Olofsson, U., & Nilsson, L.-G. (1996). Age differences in episodic memory, semantic memory, and priming: Relationships to demographic, intellectual, and biological factors. *Journal of Gerontology: Psychological Sciences, 51B,* P234–P240.

Obler, L. K., & Albert, M. L. (1985). Language skills across adulthood. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 463–473). New York: Van Nostrand Reinhold.

Olofsson, M., & Bäckman, L. (1993). Predictors of prose recall in adulthood and old age. *Archives of Gerontology and Geriatrics, 16,* 129–140.

Ordy, J. M., Brizzee, K. R., Beavers, T., & Medart, P. (1979). Age differences in the functional and structural organization of the auditory system in man. In J. M. Ordy & K. R. Brizzee (Eds.), *Aging: Vol. 10. Sensory systems and communication in the elderly* (pp. 153–166). New York: Raven Press.

Park, D. C. (1992). Applied cognitive aging research. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 449–493). Hillsdale, NJ: Erlbaum.

Parkin, A. J. (1982). Residual learning capability in organic amnesia. *Cortex, 18,* 417–440.

Perlmutter, M. (1986). A life span view of memory. In P. B. Baltes, D. Featherman, & R. Lerner (Eds.), *Advances in life-span development and behavior* (Vol. 7, pp. 271–313). Hillsdale, NJ: Erlbaum.

Raven, J. C. (1982). *Revised manual for Raven's Progressive Matrices and Vocabulary Scale*. Windsor, UK: NFER Nelson.

Reinert, G. (1970). Comparative factor analytic studies of intelligence through the human life-span. In L. R. Goulet & P. B. Baltes (Eds.), *Life-span development psychology: Research and theory* (pp. 467–484). New York: Academic Press.

Rinder, L., Roupe, S., Steen, B., & Svanborg, A. (1975). Seventy-year-old people in Gothenburg. A population study in an industrialized Swedish city. I. General presentation of the study. *Acta Medica Scandinavica, 198,* 397–407.

Robins Wahlin, T.-B., Bäckman, L., Wahlin, Å., & Winblad, B. (1993). Visuospatial functioning and spatial orientation in a community-based sample of healthy very old persons. *Archives of Gerontology and Geriatrics, 17,* 165–177.

Robins Wahlin, T.-B., Bäckman, L., Wahlin, Å., & Winblad, B. (1996). Trail Making Test performance in a community-based sample of healthy very old adults:

Effects on age on completion time, but not on accuracy. *Archives of Gerontology and Geriatrics, 22,* 87–102.

Rosow, I. (1978). What is a cohort and why? *Human Development, 21,* 65–75.

Rowe, J. W. & Kahn, R. L. (1987). Human aging: Usual and successful. *Science, 237,* 143–149.

Royall, D. A., Mahurin, R. K., & Gray, K. F. (1992). Bedside assessment of executive cognitive impairment: The executive interview. *Journal of the American Geriatrics Society, 40,* 1221–1226.

Rybash, J. M., Hoyer, W. J., & Roodin, P. (1986). *Adult cognition and aging: Developmental changes in processing, knowing and thinking*. New York: Pergamon.

Salthouse, T. A. (1983). *Adult cognition. An experimental psychology of human aging*. New York: Springer.

Salthouse, T. A. (1985). Speed of behavior and its implications for cognition. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 400–426). New York: Van Nostrand Reinhold.

Salthouse, T. A. (1987). The role of experience in cognitive aging. In K. W. Schaie (Ed.), *Annual review of gerontology and geriatrics* (Vol. 7, pp. 135–157). New York: Springer.

Salthouse, T. A. (1990). Cognitive competence and expertise in aging. In J. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 310–319). San Diego, CA: Academic Press.

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Erlbaum.

Salthouse, T. A. (1993). Speed mediation of adult age differences in cognition. *Developmental Psychology, 29,* 722–738.

Schaie, K. W. (1977–78). Towards a stage theory of adult cognitive development. *International Journal of Aging and Human Development, 8,* 129–138

Schaie, K. W. (1995). *Intellectual development in adulthood: The Seattle longitudinal study*. New York: Cambridge University Press.

Schaie, K. W. (1996). Intellectual development in adulthood. In J. E. Birren, & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 266–286). San Diego, CA: Academic Press.

Schaie, K. W., & Willis, S. L. (1993). Age difference patterns of psychometric intelligence in adulthood: Generalizability within and across ability domains. *Psychology and Aging, 8,* 44–55.

Schaie, K. W., Willis, S. H., & O'Hanlon, A. M. (1994). Perceived intellectual performance change over seven years. *Journal of Gerontology: Psychological Sciences, 49,* P108–P118.

Searight, H., & Goldberg, M. (1991). The Community Competence Scale as a measure of functional daily living skills. *Journal of Mental Health Administration, 18,* 128–134.

Shallice, T., & Burgess, P. (1991). Higher order cognitive impairments and frontal lobe lesions in man. In H. S. Levin, H. M. Eisenber, & A. I. Benton (Eds.), *Frontal lobe function and dysfunction* (pp. 125–138). New York: Oxford University Press.

Shepherd, J. (1981). Social factors in face recognition. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 55–79). New York: Academic Press.

Siegler, R. S., & Richards, D. D. (1982). The development of intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 897–971). Cambridge, UK: Cambridge University Press.

Sliwinski, M., Lipton, R. B., Buschke, H., & Stewart, W. (1996). The effects of preclinical dementia on estimates of normal cognitive functioning in aging. *Journal of Gerontology: Psychological Sciences, 51B,* P217–P225.

Stankov, L. (1988). Aging, attention, and intelligence. *Psychology and Aging, 3,* 59–74.

Stine, E. A. L., Wingfield, A., & Myers, S. D. (1990). Age differences in processing information from television news: The effect of bisensory augmentation. *Journal of Gerontology: Psychological Sciences, 45,* P1–P8.

Stuss, D., & Benson, D. (1986). *The frontal lobes*. New York: Raven.

Tulving, E. (1983). *Elements of episodic memory*. Oxford, UK: Oxford University Press.

Tulving, E. (1985). How many memory systems are there? *American Psychologist, 40,* 385–398.

Van Gorp, W. G., Satz, P., & Mirushina, M. (1990). Neuropsychological processes associated with normal aging. *Developmental Neuropsychology, 6,* 279–290.

Verhaeghen, P., Marcoen, A., & Goossens, L. (1992). Improving memory performance in the aged through mnemonic training: A meta-analytical study. *Psychology and Aging, 7,* 242–251.

Wahlin, Å., Bäckman, L., Mäntylä, T., Herlitz, A., Viitanen, M., & Winblad, B. (1993). Prior knowledge and face recognition in a community-based sample of healthy, very old adults. *Journal of Gerontology: Psychological Sciences, 48,* P54–P61.

Wahlin, Å., Bäckman, L., & Winblad, B. (1995). Free recall and recognition of slowly and rapidly presented words in very old age: A community-based study. *Experimental Aging Research, 21,* 251–271.

Wahlin, Å., Hill, R. D., Winblad, B. & Bäckman, L. (in press). Effects of serum vitamin B^{12} and folate status on episodic memory performance in very old age: A population-based study. *Psychology and Aging, 11,* 487–496.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72,* 89–104.

Weber, R. J., Brown, L. T., & Weldon, J. K. (1978). Cognitive maps of environmental knowledge and preference in nursing home patients. *Experimental Aging Research, 3,* 157–174.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale*. New York: Psychological Corporation.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence-Scale Revised*. New York: Psychological Corporation.

Welford, A. T. (1977). Motor performance. In J. E. Birren, & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 450–496). New York: Van Nostrand Reinhold.

West, R. L., Crook, T. H., & Barron, K. L. (1992). Everyday memory performance across the life span: Effects of age and noncognitive individual differences. *Psychology and Aging, 7,* 72–82.

Willis, S. L. (1996). Everyday problem solving. In J. E. Birren, & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 287–307). San Diego, CA: Academic Press.

Yussen, S. R., & Kane, P. T. (1983). Children's ideas about intellectual ability. In R. L. Leahy (Ed.), *The child's construction of social inequality* (pp. 109–134). New York: Academic Press.

Zarit, S. H. (1979). Helping an aging patient to cope with memory problems. *Geriatrics, April,* 82–90.

Zarit, S. H., Johansson, B., & Berg, S. (1993). Functional impairment and co-disability in the oldest old: A multidimensional approach. *Journal of Aging and Health, 5,* 291–305.

Zelinski, E. M., Gilewski, M. J., & Schaie, K. W. (1993). Individual differences in cross-sectional and three year longitudinal memory performance across the adult life span. *Psychology and Aging, 8,* 176–186.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.03
Social Relationships and Adaptation in Late Life

FRIEDER R. LANG
Freie Universität Berlin, Germany
and
LAURA L. CARSTENSEN
Stanford University, CA, USA

7.03.1 INTRODUCTION 55
7.03.2 AGE-RELATED CHANGES OF SOCIAL-RELATIONAL CONTEXTS 56
  7.03.2.1 Family Relationships 57
    7.03.2.1.1 Marriage and intimate relationships 58
    7.03.2.1.2 Adult children and grandchildren 58
    7.03.2.1.3 Older siblings 58
  7.03.2.2 Friendship Relationships 59
7.03.3 HEALTH-RELATED ASPECTS OF SOCIAL RELATIONSHIPS 59
  7.03.3.1 Social Relationships and Health: Evidence for Positive Effects 60
  7.03.3.2 The Negative Side of Social Relationships 61
    7.03.3.2.1 Dysfunctional social support 61
    7.03.3.2.2 Unpleasant and demanding social contacts 62
    7.03.3.2.3 Violence and neglect: the relational dynamics of caregiving 62
    7.03.3.2.4 Anticipation or experience of relational loss 62
  7.03.3.3 Managing Social Resources 63
7.03.4 THEORETICAL PERSPECTIVES ON THE MANAGEMENT OF SOCIAL RELATIONSHIPS 64
  7.03.4.1 Socioemotional Selectivity Theory: Emotion Regulation in Late Life 64
  7.03.4.2 Social Convoy and Support Efficacy: Maintaining Competence in Late Life 64
7.03.5 CONCLUSIONS AND OUTLOOK 65
7.03.6 REFERENCES 66

7.03.1 INTRODUCTION

The role of social relationships in late-life adaptation has stimulated much social gerontological and psychological research since the 1950s (Cumming & Henry, 1961; Havighurst & Albrecht, 1953; Lee, 1985; Osgood, 1989) and generated considerable debates about theoretical models of social aging. Disengagement theorists, for example, have claimed that social withdrawal reflects an older individual's adaptive response to personal loss and to social forces such as age stereotypes and restrictive social practices (Cumming & Henry, 1961; Maddox, 1963; Pratt & Norris, 1994). Activity theorists, in contrast, have argued that successful adaptation in later life requires older people to be sociall active (Havighurst, Neugarten, & Tobin, 1968; Lemon, Bengtson, & Peterson, 1972). From a life-course perspective, Atchley (1993) stated that individuals in old age prefer those social interactions and social activities that allow for continuity of their life-long roleidentities, experiences, and competencies. Rosow (1975, 1982) claimed that loss of social roles is inevitable in late life and adaptive social aging occurs when older people engage in social activities with age peers. Despite the great diversity in existing theoretical and empirical models of social aging, it has been widely acknowledged that the rules, mechanisms, and functions of social relationships in later life are better understood within a life-span perspective (Antonucci & Akiyama, 1995; Carstensen, 1993a; Marsiske, Lang, Baltes, & Baltes, 1995; Schulz & Rau, 1985). Socioemotional selectivity theory has emphasized and elaborated the role of life-long adaptive and motivational mechanisms that regulate social contact and social relationships (Carstensen, 1987, 1993a, 1995). According to the theory, during early phases of life, preferences for social contacts with others are instigated very frequently by information-seeking goals, yet social contacts in later life are more often determined by emotion-regulatory goals (Carstensen, 1993a, 1995). Empirical findings suggest that older people select and manage their social relationships such that attainment of emotional benefits and emotional goals is maximized or enhanced (Fredrickson & Carstensen, 1991; Lang & Carstensen, 1994; Lang, Staudinger, & Carstensen, in press; Lang & Tesch-Römer, 1993). In addition, structural and functional aspects of social networks appear to follow predictable life course patterns (Antonucci & Akiyama, 1995; Kahn & Antonucci, 1980; Lang & Carstensen, 1994). Social adaptation in later life can also be understood within the framework of the model of selective optimization with compensation (Baltes & Carstensen, 1996, Marsiske et al., 1995). In order to adapt successfully to the constraints of later life, older individuals have to first select meaningful social goals and social resources within the limits of their social-relational contexts. Such selection may then, subsequently, serve compensatory or optimizing purposes in achieving relevant social goals. Compensation, for example, pertains to seeking social support, for example, when adjusting to functional or personal loss. Optimization relates to relational enhancement mechanisms such as generativity or reciprocity (Lang & Baltes, 1997).

Social relationships are dynamic by nature. They vary from person to person, from situation to situation, and from one instance of interaction to the next instance of interaction (e.g., Caspi, Bolger, & Eckenrode, 1987). How adaptive a specific social transaction is thus depends on the situational and relational context as well as on the individual's needs and preferences. For example, a confidant's advice and support to someone who struggles with a decision may be quite supportive and useful during a specific decision-making process but the same advice may be inadequate and even disturbing after the decision has been made. In this example, the adaptive outcome of social contact depends largely on the confidant's empathy and knowledge of the individual's decisional process (e.g., Hogg & Heller, 1990). In more general terms, the adaptivity of social relationships is determined by the extent to which social-relational contexts reflect the individual's social needs and capacities. Such a match between person and context characteristics plays a critical role in understanding the older person's social adaptivity (e.g., Kahn, 1979; Lawton, 1989), and will be discussed in greater detail below.

In this chapter, normative changes in the characteristics of social relationships and social networks across adulthood focusing on close family and friendship relationships are addressed. Then empirical evidence for health-related functions of social support and social relationships underscoring the influence of the older individuals' social needs and preferences in adaptive relational outcomes is reviewed. Next the issue of negative aspects and negative outcomes of social relationships in later life is addressed. Negative aspects of social relationships often have been related to facets of the individual's social relationships, such as ineffective, inadequate, or unpleasant social resources (e.g., Rook & Pietromonaco, 1987) including even physical abuse (Pillemer & Finkelhor, 1988). However, negative aspects of social relationships also may be related to maladaptive social behaviors or characteristics of the older individual such as excessive autonomy needs (e.g., Baltes & Silverberg, 1994) or misanthropy (e.g., Schütze & Lang, 1995). Finally, in Section 7.03.4, findings on the age-, health-, and relation-specific aspects of social relationships are discussed in the light of some more recent theoretical perspectives on individuals' adaptive social behaviors and relational management in later life drawing on socioemotional selectivity theory (Carstensen, 1995) and the convoy model of social relationships (Antonucci & Akiyama, 1995).

7.03.2 AGE-RELATED CHANGES OF SOCIAL-RELATIONAL CONTEXTS

Many changes in the structural or functional characteristics in the social worlds of older people are associated with particular age-specific experiences such as retirement (Bosse, Aldwin, Levenson, Spiro, & Mroczek, 1993; Carp, 1972; Havighurst, Munnichs, Neugarten, & Thomae, 1969) and adult children leaving home, marrying, and eventually having children of their own (Lowenthal, Thurnher, & Chiriboga, 1976; Miller & Cavanaugh, 1990; White & Edwards, 1990). In addition, experiencing the loss of a spouse, close relative, or friend inevitably changes the resources available to the older person (Garrett, 1987; Lopata, 1986; Morgan, 1989). While there is strong evidence that people have fewer social contacts and social relationships in later life as compared to earlier life phases (Gordon & Gaitz, 1976; Harvey & Singleton, 1989; Heller & Mansbach, 1984; Lawton, Moss, & Fulcomer, 1987; Lee & Markides, 1990; Lowenthal et al., 1976; Palmore, 1981), there are also consistent reports suggesting that the majority of old and even very old people are well integrated into family and friendship networks (e.g., Antonucci & Akiyama, 1995; Bowling & Browne, 1991; Levitt, Weber, & Guacci, 1993).

In the United States and in Europe, most older people above 65 years report having 5–15 important or close social partners in their social network (Antonucci & Akiyama, 1987; 1995; Dykstra, 1993; House & Kahn, 1985; Lang & Carstensen, 1994; Lang, Staudinger, & Carstensen, 1996; Stacey-Konnert & Pynoos, 1992). Empirical findings regarding social network size in early and middle adulthood are less consistent but generally range from 15 to 35 network members (e.g., Antonucci & Akiyama, 1994; Carstensen & Lang, in preparation; Van Sonderen, Ormel, Brilman & Van Linden, 1990; Wellman, 1979). Compared to social networks of young or middle-aged adults, older people report a greater percentage of family members and fewer friends or acquaintances in their social networks (Antonucci & Akiyama, 1987; 1995; Babchuk, 1978–79; Levitt et al., 1993; Stacey-Konnert & Pynoos, 1992). Indeed, while many friendships with peers are lost due to death (Garrett, 1987; Johnson & Troll, 1994; Roberto & Stanis, 1994), there is also some indication that older people selectively maintain contacts with close friends rather than less meaningful friends or acquaintances in old age (Bowling & Browne, 1991; Carstensen, 1992; Lang & Carstensen, 1994).

Although the risk of social isolation or feeling lonely may increase in very old age (Larson, 1990; Mullins, Sheppard, & Andersson, 1991), this does not seem to characterize later life phases in general (Revensen & Johnson, 1984; Schultz & Moore, 1988; Shute & Howitt, 1990) and varies across individuals depending on specific relational contexts such as widowhood or marital break-up (Dykstra, 1995; Ferraro & Barresi, 1982; Keith, 1986), marital crisis (Tornstam, 1992), living alone (Henderson, Scott, & Kay, 1986), childlessness (Lang, 1994; Page & Cole, 1991), filial neglect (Connidis & McMullin, 1993; Lang, 1994), or death of close friends (Roberto & Stanis, 1994). Obviously, the social networks of people who are widowed, who have never married, or who are childless differ in many respects from those older people who have lived with their spouse for many decades or who have reared children that have eventually become independent adults (e.g., Connidis & McMullin, 1994; Dykstra, 1993; Lang, 1994). Some older individuals have divorced (e.g., Hammond & Muller, 1992), others have faced the loss of a spouse or close intimates (Dean, Matt, & Wood, 1992; Morgan, 1989), or even an adult child (Moss, Lesher, & Moss, 1986–87; Rubin, 1991–92). The majority of people over 65 years of age have experienced the death of a sibling (Moss & Moss, 1989). Some have focused their social worlds on the maintenance of friendship relationships (Armstrong & Goldsteen, 1990; Johnson & Troll, 1994; Roberto & Scott, 1987), while for others close ties with family members prevail (Field & Minkler, 1988; Field, Minkler, Falk, & Leino, 1993; Johnson & Troll, 1992). Such pervasive multifacetness and heterogeneity in the structure of interpersonal contexts in late life is also reflected in the specific demands and potentials of older peoples' family and friendship relationships.

7.03.2.1 Family Relationships

In general, family relations prove to be reliable and stable sources of social support and social embeddedness in old and very old age, even when relatively little effort is made to keep in touch or to maintain the relationship (Carstensen, 1993b; Crohan & Antonucci, 1989; Field & Minkler, 1993). The close family is most often named as a source of meaningful social contact or social support in old age (Caplan, 1982; Depner & Ingersoll-Dayton, 1985; Dykstra, 1993; Hogan, 1995; Sussman, 1985). Findings from the Berkeley Intergenerational Study have revealed that contacts and feelings of closeness in family relationships are stable or even increase in old age (Field & Minkler, 1988, 1993). This seems to be true for marital relationships and parent–child relationships although less so for sibling relationships (Cicirelli, 1989; Connidis, 1989; Scott, 1990), which are discussed later in this chapter.

7.03.2.1.1 Marriage and intimate relationships

When older people are married or live with an intimate partner, the relationship is mostly described as an important source of meaningful and satisfactory social contact in everyday life (Alwin, Converse, & Martin, 1985; Huyck, 1995). Partners are most often named as social supporters and caregivers when functional losses occur (Brody, 1985; Chappell, 1991; Connidis & McMullin, 1994; Hess & Soldo, 1985). In addition, there is consistent empirical evidence showing that older people are, on average, quite satisfied with their marriages (Brubaker, 1990; Herman, 1994; Levenson, Carstensen, & Gottman, 1993). Later life couples have fewer conflicts and experience more pleasure in sharing intimate and intergenerational experiences or activities than middle-aged couples (Levenson, Carstensen, & Gottman, 1993). Older couples also express more affection than younger couples (Carstensen, Gottman, & Levenson, 1995; Levenson, Carstensen, & Gottman, 1994). Such evidence may also suggest a survival of the fittest marriages, that is, those which have stood the test of time. Moreover, interindividual differences also matter. Empirical evidence suggests that extraversion and neuroticism are associated with marital stability (Cramer, 1993; Kurdek, 1991).

Overall, long-term marital relationships have been demonstrated consistently to be stable and reliable sources of positive and gratifying social experiences. Next to marital relationships, adult children and their families frequently are nominated as the second most important source of emotionally meaningful contact and social support in late life (Chappell, 1991; Dykstra, 1993; Eggebeen, 1992; Hogan, 1995).

7.03.2.1.2 Adult children and grandchildren

Contrary to the popular lore about older parents being abandoned and neglected by their adult children (Shanas, 1980; Steinmetz, 1988), there is considerable empirical evidence showing that adult children not only stay in contact with their parents (Hagestad, 1987), but they continue to share close emotional relationships with them (Field & Minkler, 1988). Frequently they exchange support (Eggebeen & Hogan, 1990; Hogan, 1995; Lang, 1996; Rossi & Rossi, 1990), and are, second only to spouses, the principle caregivers of frail elderly parents (Brody, 1985; Chappell, 1990; Horowitz, 1985; Mancini & Blieszner, 1989).

Typically, older parents also experience the developmental task of grandparenthood that demands the acquisition of a new social role (Kivnick, 1982) and the development of an individual style of grandparenting (Sticker, 1991). Grandparenthood is associated with many changes in the socioemotional worlds of older people such as increased contact and satisfaction with adult children (Aldous, 1987), increase in intergenerational giving (Eggebeen & Hogan, 1990; Tinsely & Parke, 1987), and becoming the keepers of meaning and conveyers of the family tradition (Ramirez-Barranti, 1985; Bengtson, 1985; Hagestad, 1985; Lang & Baltes, 1997). While grandparenthood is often associated with positive experiences and social satisfaction (Kennedy, 1992; Sticker, 1991), grandchildren are rarely named as support providers before they reach early adulthood (Hodgson,1992; Langer, 1990; Markides, Boldt, & Ray, 1986).

7.03.2.1.3 Older siblings

In many ways, social ties with siblings are unique in later adulthood. Older siblings share life-long biographical and historical experiences and are often the only social partners who have known the older individual as a child or adolescent (Cicirelli, 1995; Gold, 1987). Thus, older siblings share an exclusive biographical and sociohistorical knowledge of one another. However, compared to most other close family relationships, social contact frequencies and support exchanges among older siblings are typically limited to times of crises or increased need (Avioli, 1989; Connidis, 1994) unless other intimates such as spouses (O'Bryant, 1988; Wilson, Calsyn, & Orlofsky, 1994) or adult children are unavailable (Connidis & Davis, 1992; Ikels, 1988). As is true for most other social relationships as well, siblings interact more often when they live in close geographical proximity to one another (Connidis & Davis, 1992; Scott, 1983).

However, it is the quality and emotional closeness of sibling relationships that contributes to well-being rather than the quantitative aspects of contact or support exchange (Wilson, Calsyn, & Orlofsky, 1994). These associations are consistently found to be stronger for relationships with sisters or a same-sex sibling as opposed to relationships of sisters with their brothers (Cicirelli, 1989; Connidis, 1989; Pulakos, 1987). Although sibling relationships in later life are often continued sources of conflict and rivalry (Bedford, 1995; Goetting, 1986), older adults may be more discriminating in their choice of social partners in an effort to ensure emotionally gratifying contact with siblings.

7.03.2.2 Friendship Relations

In contrast to social relationships that are legally defined or determined by kinship, friendship is characterized by the voluntary involvement of two or more partners with one another (Adams & Blieszner, 1994; Blau, 1973; Ishii-Kuntz, 1990). Failures to abide by normative expectations in friendships (Roberto & Scott, 1986a, 1986b) in contrast to parent–child relationships, for example, leads to dissolution of the friendship. In addition, longitudinal findings suggest that while the relevance of family roles remains quite stable in old age, the significance of other social roles seems to decrease (Field & Minkler, 1993; Lehr & Minnemann, 1987). Some friendships dissipate and those that remain consist predominantly of friends who have known each other for a relatively long period of time (e.g., Lowenthal, Thurnher, & Chiriboga, 1976).

Notably, friendship relationships have been shown consistently to be associated with higher social satisfaction than relationships with family members who are not living in the same household (Chappell & Badger, 1989; Larson, Mannell, & Zuzanek, 1986, Mullins & Mushel, 1992). One possible explanation for this is that friends are often discharged from tasks that are related to instrumental support exchange and, thus, are less likely to be burdened by caregiving than are family members. In contrast, friends are often named as confidants of old people (Keith, Hill, Goudy, & Powers, 1984; Strain & Chappell, 1982) and are associated with pleasant interactions such as social companionship (Crohan & Antonucci, 1989; Rawlins, 1992). In this respect, friends might play a moderating role in older people's social networks by counterbalancing sometimes difficult or strained family relationships (Johnson, 1983) and by providing opportunities for social integration in the community (Peters & Kaiser, 1985).

When it comes to supportive transactions among friends, satisfaction with friendship is associated with the perceived equity of the exchange (Roberto, 1989; Roberto & Scott, 1986a; Rook, 1989). Nonreciprocal exchanges among friends seem to have even more detrimental effects than among family members (Antonucci & Jackson, 1990). Interestingly, though, close or "best" friends appear to be exceptions (Armstrong & Goldsteen, 1990). Older individuals who reflect their best friends' needs and welfare in their thinking (i.e., communal orientation) report greater satisfaction with their friendships than those who focus on the equity or comparibility of exchanged benefits with their friends (i.e., exchange orientation; Jones & Vaughan, 1990).

In sum, friendship and family relationships are associated with distinct social resources and specific functions that cannot be easily substituted for one another (Litwak, 1985). However, there is also evidence that when specific types of relationships such as spouse or children are unavailable, other relatives or friends appear to assume their functions (Cantor, 1979; Chatters, Taylor, & Jackson, 1986). In general, there is much evidence that older people are satisfied with both their family relationships and their friendships (Field & Minkler, 1988; Herzog & Rodgers, 1981; Ryff, 1989).

7.03.3 HEALTH-RELATED ASPECTS OF SOCIAL RELATIONSHIPS

There is widespread agreement that social relationships provide a fundamental resource for late-life adaptation. However, the underlying mechanisms and health functions of social resources are not fully understood (House, Umberson, & Landis, 1988). Social support has been proposed generally to exert a direct protective effect on health or well-being, while stressor-specific support of others may operate as a buffer when adverse or stressful events occur (Cohen & Wills, 1985; Schwarzer & Leppin, 1991). Others have proposed more complex patterns of associations in which specific types of social resources—such as emotional, instrumental, or tangible support (Rook & Pietromonaco, 1987)—are expected to have specific effects on health or well-being for specific individuals in specific situations (Gottlieb, 1983; House & Kahn, 1985; House et al., 1988; Kessler & McLeod, 1985). However, the age-specific needs and capacities of older support recipients and their social partners are often not well integrated in theoretical models of adaptivity in later life (see also Baltes & Carstensen, 1996; Carstensen & Lang, 1997; Staudinger, Marsiske, & Baltes, 1995). Indeed, considerable empirical evidence may serve to underscore the notion that age-specific social preferences and social belief systems (e.g., social self-efficacy; Lang, Featherman, & Nesselroade, in press) contribute to the adaptive functions of social support and social contact in later life.

The most convincing evidence for positive effects of social support and social networks on health comes from a variety of longitudinal studies of mortality in later life. The interested reader is referred to several excellent reviews of empirical findings about associations of social resources with mortality (Atkins, Kaplan, & Toshima, 1992; Schwarzer & Leppin, 1991), physical health (Antonucci & Jackson, 1987;

Schwarzer & Leppin, 1991), immunologic functioning (Antonucci & Jackson, 1987; Kennedy, Kiecolt-Glaser, & Glaser, 1990), disease course (Kriegsman, Penninx, & van Eijk, 1995), recovery from illness (Reifman, 1995), adjustment to cancer (Helgeson & Cohen, 1996), depression and mental health (Gotlib & Whiffen, 1992; Schwarzer & Leppin, 1992), and well-being (Bowling, 1994; Kessler & McLeod, 1985). It should be noted, however, that some reviews have reported minimal or even no impact of social support on health or well-being (Ganster & Victor, 1988; Schulz & Rau, 1985; Wallston, Alagna, DeVellis, & DeVellis, 1983) and some have also pointed to the potential negative effects of social relationships (Rook, 1984; Rook & Pietromonaco, 1987). In the following, health effects of social relationships are distinguished as they pertain to the relational contexts of older people, on the one hand, and as they pertain to relational behaviors, on the other, such as older individuals' styles or strategies of relating to others, such as selecting social partners, relying on others, reciprocating support, and enhancing social relationships.

7.03.3.1 Social Relationships and Health: Evidence for Positive Effects

Epidemiological studies in the United States and Europe have consistently found associations between the individual's risk of premature mortality and social resources or social contexts. For example, having close confidants or intimate social partners has been found repeatedly to be associated with increased longevity (Berkman, 1985; Berkman & Syme, 1979; Broadhead et al., 1983; Hammer, 1983; Hanson, Isacsson, Janzon, & Lindell, 1990; Orth-Gomer & Johnson, 1987; Steinbach, 1992; Sugisawa, Liang, & Liu, 1994). In the Terman Gifted Study, having a stable long-term marriage predicted increased longevity above and beyond gender, childhood personality, and health behaviors, while the risk of premature mortality was highest among those who had experienced divorce or separation before age 40 (Friedman et al., 1995). Separation and divorce are critical non-normative life events that may affect longevity. When individuals do not develop new intimate relationships after separation or divorce, they are at greater risk of premature mortality (Tucker, Friedman, Wingard, & Schwartz, 1996).

Furthermore, increased social network size and frequency of social contacts reduce the risk of mortality (Cohen, Teresi, & Holmes, 1986–87; Olsen, Olsen, Gunner-Svensson, & Waldstrom, 1991). However, social resources seem to affect morbidity differentially, among men and women (Bryant & Rakowski, 1992; Shye, Mullooly, Freeborn, & Pope, 1995). For example, while premature mortality is related to marital status among men (Berkman & Syme, 1979; Schoenbach, Kaplan, Fredman, & Kleinbaum, 1986), womens' chance of survival is better determined by size of social network (Shye et al., 1995), level of social activity, and church attendance (Berkman & Syme, 1979; House, Robbins, & Metzner, 1982).

In an Israeli five-year incidence study on adult men, high family conflict and low spousal support were related to increased risk of angina pectoris (Medalie & Goldbourt, 1976). In a Danish 14-year follow-up study of mortality, the perceived quality of social resources rather than quantitative characteristics of social networks were associated with increased longevity (Olsen et al., 1991). Generally, there is much empirical evidence that the perceived or subjective quality of social support or social contact are associated more strongly with adaptive outcomes than any other aspect or measure of social support (Antonucci & Jackson, 1987; Dunkel-Schetter & Bennett, 1990; Schwarzer & Leppin, 1991). This further underscores the qualification that social relationships are most effective when they are congruent with the individual's competence and social preferences (Carstensen, 1987; Carstensen & Lang, 1997). For example, in old age declines in health status are associated with receiving help and support from sources that are increasingly distant from the older person (Stoller & Pugliesi, 1988). This means that emotional or instrumental social support may often come from social partners who are not emotionally meaningful to the older person and may thus create unease or enhance feelings of dependence in the older individual.

When changes in health status occur, individuals may change their apprehension of available social relationships and social support (e.g., Gotlib & Whiffen, 1992; Wilcox, Kasl, & Berkman, 1994). When the individual's level of functioning and competence change, the individual may also adjust their preferences for specific social partners and for specific functions of social interactions. However, such changes in social preferences or needs may also affect the adequacy and satisfaction with previously available social resources. Research on social support has often neglected the fact that individuals change their social preferences and social behaviors in the course of physical or mental health crises, even when examined in prospective studies, where social support can be measured before and after health changes occur (e.g., Holahan, Moos, Holahan, & Brennan, 1995; Wilcox, Kasl, & Berkman, 1994). In fact, the benefits and positive outcomes of receiving social support may depend on the older individual's competence to utilize social resources in adaptive ways (Hansson & Carpenter, 1994; Lang, 1996; Thomae, 1994). Findings from empirical studies of health and well-being that do not consider changes in the individual's social needs and preferences may underestimate the actual impact of social resources.

7.03.3.2 The Negative Side of Social Relationships

Since the early 1980s, the perspective of social relationships as a uniformly positive resource has been challenged by conceptual and empirical work that demonstrates that positive relational outcomes can be neutralized or even reversed by negative aspects of social relationships, such as relational strain (Rook, 1984, 1987a), overprotective or unsolicited support (Thompson & Sobolow-Shubin, 1993a, 1993b), daily relational hassles (Bolger & Kelleher, 1993), and violence or physical abuse (Pillemer & Finkelhor, 1988; Pillemer & Suitor, 1992). Another, often overlooked, source of negative experience in social relationships pertains to the experience or anticipation of personal loss that is relatively common in later life (Garrett, 1987; Lopata, 1986; Morgan, 1989).

While negative aspects of social relationships, in theory, are often exclusively discussed in respect to the social behaviors of others (Rook & Pietromonaco, 1987), empirical studies on negative interactions have often focused exclusively on subjective evaluations of social support to the older individual (e.g., Schuster, Kessler, & Aseltine, 1990). Although an individual's negative relational experiences may be difficult to empirically disentangle from social (dis-)satisfaction or perceived support (in-)adequacy, this is crucial for an understanding of the individual's role in dealing with negative contents of social interactions. How an older individual responds to a specific threat or strain in a concrete social relationship may determine the individual's satisfaction or affect as much as the negative event itself. For example, conflict or friction among spouses may be easily resolved when both partners are satisfied with their marriage but may have detrimental effects and even predict separation when marital satisfaction is low (e.g., Carstensen et al., 1995; Gottman & Levenson, 1992; Levenson et al., 1993).

Most negative aspects of social relationships typical in later life may be associated with intensive emotional turmoil that demands efficient types of affective regulation on the part of the older individual. In the following, four sources of potential negative aspects of social relationships in later life are distinguished. They are: (i) dysfunctional social support, such as undesirable, ineffective, or excessive support; (ii) unpleasant or overdemanding social contact; (iii) violence or neglect; and (iv) anticipation or experience of personal loss of others.

7.03.3.2.1 Dysfunctional social support

While there is convincing empirical evidence that social support does contribute to life quality and well-being in later life, there is also a growing body of empirical studies showing that social relationships not only provide gratification but also entail risks and costs for the individual (for reviews see Jung, 1990; Rook & Pietromonaco, 1987; Shinn, Lehmann, & Wong, 1984). Failures to provide social support have been related to a variety of characteristics of the support process such as undesirable social contact (Cutrona, Cohen, & Igram, 1990; Rose, 1993), lack of empathy and sensitivity (Wortman & Lehman, 1985), grudgingly given support (Coyne, 1976), interpersonal betrayal (Hansson, Jones, & Fletcher, 1990), emotional overinvolvement (Coyne, Wortman, & Lehmann, 1988), overprotection (Thompson & Sobolow-Shubin, 1993a, 1993b), and the reinforcement of dependent behaviors (Baltes, 1995).

Negative aspects of social relationships have been shown consistently to operate relatively independently of personality or situational characteristics (Finch, Okun, Barrera, Zautra, & Reich, 1989). Furthermore, empirical evidence suggests that negative experiences in supportive interactions are more strongly associated with well-being or mental health than perceived positive instances of social support (Finch et al., 1989; Okun, Melichar, & Hill, 1990; Rook, 1984; Schuster et al., 1990). However, there is no consistent evidence that detrimental effects of negative social support are enhanced dramatically in the absence of positive social support (Schuster et al.) or in the presence of negative events (Okun et al., 1990). Typically, negative and positive social interactions are only mildly or not at all correlated (Finch et al.; Okun et al.; Rook, 1984; Shinn et al., 1984). It has been suggested (Schuster et al.; Shinn et al.) that such a dissociation of positive and negative aspects may parallel findings on independence of positive and negative affect (e.g., Watson, Clark, & Tellegen, 1988). However, no study, to the best of our knowledge, has examined the age-differential effects of positive and negative supportive interactions on the experience of positive and negative affect in daily life.

Empirical studies of the experience of negative supportive interactions typically have focused on the individual's subjective experience of such negative interaction. However, it seems plausible that older individuals may have idiosyncratic responses to negative events in their close relationships and may develop individual strategies of avoiding or managing negative aspects of close relationships (e.g., Carstensen et al., 1995). Moreover, many negative aspects of dysfunctional social support may also be related to the older individual's social contacts and social preferences. For example, a failure to communicate wishes or needs may be as much responsible for inefficient or inadequate social support as a social partners' lack of social abilities to "do the right thing" (e.g.,Query & James, 1989). It has also been suggested that individuals do not rely on others when they cannot reciprocate received benefits in order to avoid feelings of shame or guilt (Fisher, Nadler, & Whitcher-Alagna, 1982; Riley & Eckenrode, 1986; Shumaker & Jackson, 1979)

7.03.3.2.2 Unpleasant and demanding social contacts

Conflicts, demands, and criticism are part and parcel of even the best close social relationships and may even increase over time (Rook & Pietromonaco 1987; Antonucci, 1985). In addition, personal relationships are not always stable and reliable but to some extent are also characterized by unpredictability and fluctuations (e.g., Lang, Featherman, & Nesselroade, in press).

Although empirical research on unpleasant and demanding aspects of social relationships in later life remains scarce, there is some evidence that older adults may differ from younger adults in respect to how they react and deal with relational conflict or friction (e.g., Carstensen et al., 1995). Older people may be more skilled in avoiding unpleasant or overly demanding situations with social partners about whom they feel less strongly or positive. For example, it has been shown that lack of equity or reciprocity expectations in older adult's friendships are very likely to lead to dissatisfaction or even dissolution of the relationship (Roberto & Scott, 1986a, 1986b).

There is some evidence that when individuals hold negative views toward others they have a greater risk of premature mortality (Barefoot et al., 1987) and are more vulnerable to depression or other form of negative affect when confronted with demands, criticism, or rejection (Smith, 1992). For example, spouses of depressed partners are less supportive, more critical, hostile, and ambivalent towards their partners (Hautzinger, Linden, & Hoffman, 1982; Ruscher & Gotlib, 1988; for a review, see Gotlib & Hooley, 1988). Personality characteristics and relational styles may thus conjointly influence the way individuals react to unpleasant social encounters (Hansson & Carpenter, 1994). However, although personality characteristics are relatively stable over the life course (Costa & McCrae, 1994), their impact in determining outcomes of social interactions and social adaptation may well change with age and may account for fewer outcomes in later adulthood (Lang et al., in press). In other words, older people may be at greater risk of experiencing negative effects when confronted with unavoidable, unpleasant,or demanding social behaviors of others.

7.03.3.2.3 Violence and neglect: the relational dynamics of caregiving

Violent or neglectful behaviors are the most obvious and dramatic negative aspects of social relationships in later life. In the United States, estimates of incidence of elder abuse range from seven hundred thousand to one million people older than 65 years (Hansson & Carpenter, 1994). Violence or neglect occurs mostly in relationships with social partners who share a household (Pillemer & Finkelhor, 1988). Indeed, social contexts of caregiving children or spouses may be experienced as extremely stressful and frustrating for caregivers as well as older individuals (Franks & Stephens, 1996; Schulz, Tompkins, & Rau, 1988; Walker, Martin, & Jones, 1992). Caregiver stress (for reviews: Pearlin, Mullan, Semple, & Skaff, 1990; Schulz, Visintainer, & Williamson, 1990) has been reported to be associated with increased depression and emotional turmoil (Pruchno, Peters, & Burant, 1995; Townsend & Franks, 1995), social isolation, and the lack of social support (Johnson & Catalano, 1983; George & Gwyther, 1986).

In addition, the relational dynamics of caregiving may lead to further escalation of hostility and aggression in the care receiver as well as in the caregiver. Older individuals who receive care must also come to terms with their lack of autonomy and dependence on the caregiver (Baltes & Silverberg, 1994).

7.03.3.2.4 Anticipation or experience of relational loss

Finally, experiencing the deaths of close friends and relatives is normative in late life

(Holmes & Rahe, 1967; Johnson & Troll, 1994). Most individuals over 65 years have experienced losses of close friends, siblings, or of partners during late life and also struggle with the anticipatory loss of other close social partners (Garrett, 1987; Lopata, 1986; Morgan, 1989).

Despite empirical evidence for the negative impact of widowhood on the social world, however, research suggests a more differentiated perspective on adjustment to widowhood (e.g., Ferraro & Farmer, 1995). Prior to widowhood, many spouses endure extended periods when their partners suffer from severe illness and require considerable care (Bass & Bowman, 1990; Bass, Bowman, & Noelker, 1991; Mullan, 1992). Subsequently, as painful as widowhood may be, it can also entail a sense of relief that the spouse's suffering has finally come to an end and the widowed spouse regains freedom to engage in more pleasant activities.

Due to women's longer life expectancy and the traditional age differences among married couples, widowhood is a normative experience among married women. It is a less common experience for married men. In addition, widowhood seems to be related to gender-differentiated demands and needs. Adjustment to widowhood is more strongly related to social compensation among widows than among widowers (Ferraro & Farmer, 1995). Typically, widows are threatened by financial losses following widowhood, whereas men are at risk if they have difficulties with keeping up an independent household (Umberson, Wortman, & Kessler, 1992). In addition, widowhood is associated with a higher mortality risk for widowers than for widows (Bowling, 1988-89; Bowling & Windsor, 1995).

7.03.3.3 Managing Social Resources

A review of empirical studies on differential effects of positive and negative aspects of supportive relationships on late-life adaptation suggests that positive outcomes of social resources are largely determined by the extent to which social resources actually meet the older individual's social preferences and needs. For example, it has been reported consistently, that perceived adequacy of social support is more strongly associated with adaptive outcomes than any other aspects of social support. However, the cognitive and behavioral strategies in supportive interactions that determine and influence subjective support adequacy are not well understood. Few empirical studies have examined cognitive or behavioral correlates of perceived social support adequacy in adulthood (e.g., Dunkel-Schetter, Folkman, & Lazarus, 1987; Röhrle, 1994) or old age (Seeman & Berkman, 1988), let alone considered age-related changes in motivational or emotional determinants of managing social support. Empirical evidence for changed social preferences towards greater emphasis on emotionally meaningful social contacts in late adulthood (Carstensen, 1995; Fredrickson & Carstensen, 1990) suggests that the adequacy and impact of social contact and social support on well-being in later life depends on emotion-regulatory aspects of social relationships such as emotional closeness (Lang & Carstensen, 1994), intimacy (Babchuk & Anderson,1989), generativity (McAdams, de St. Aubin, & Logan, 1993), functional specificity (Felton & Berry, 1992), and trust (Holmes, 1991; Thomae, 1994). There is also empirical evidence that older people benefit from actively engaging in support-seeking behaviors (Filipp & Aymanns, 1987), and from reciprocating support to others (Jung, 1990; Krause, Herzog, & Baker, 1992; Lang, 1996). Others have focused on the personality characteristics of older people that were shown to be closely associated with increased levels of relational competence (Hansson & Carpenter, 1994) related to the initiation, maintenance, and enhancement of social relationships or relational outcomes (Carpenter, 1993; Mariske et al., 1995). Moreover, feelings of control in salient social roles (such as spouse, parent, or grandparent) have been shown to buffer deleterious effects of stressful life events on life satisfaction (Krause, 1994).

The adaptive management of social resources in later life is also illustrated by empirical evidence for the learned dependency model developed by Baltes and her colleagues (Baltes, 1995; Baltes, Neumann, & Zank, 1994). This line of observational research has revealed that dependent behaviors of older ,nursing home residents are more often positively reinforced by responsiveness and social contact than independent behaviors. Subsequent intervention studies demonstrated that independent behaviors of residents increased when staff members and other social partners were trained to be responsive to independent behaviors (Baltes et al.). The initiation of such dependence-support scripts in nursing home settings has been conceptualized as an adaptive behavioral strategy by which older people regulate and optimize the available resources in their social environment (Baltes, 1995). In this case, relinquishing independence is seen as a "proxy control" strategy where the individual gains control in particular domains of functioning through other persons, thereby exerting an indirect influence over the environment (Bandura, 1982; Baltes, 1995). Proxy control may be an important mechanism by which older individuals can compensate for the loss of primary control and enhance the perceived adequacy of social resources (Heckhausen & Lang, 1996).

Moreover, most personal relationships go through times of conflict or periods of relationship crises that must be mastered if social partners are to maintain their relationships (Antonucci, 1985; Duck, 1986). Successfully managing both the stable and varying, positive, and negative aspects of family ties or friendships is likely to be a fundamental requisite for maintaining social adaptivity throughout adulthood and late life (Lang et al., in press; Marsiske et al., 1995; Morgan, 1990).

7.03.4 THEORETICAL PERSPECTIVES ON THE MANAGEMENT OF SOCIAL RELATIONSHIPS

As people grow old, their social relationships and social networks undergo considerable change as a result of both changing relational contexts in later life as well as changes in the individual's social behaviors and preferences. In this chapter, literature showing age-related changes in relational contexts as well as age-related differences in individuals' social motivations and social contacts has been discussed. In the remainder of this chapter, we will summarize and integrate the emerging picture of normal aging in social relationships and propose a common theoretical framework of socioemotional selectivity and social adaptation. It is concluded that people select, mold, and manage their social relationships in accordance with their needs, preferences, and competencies. However, the extent to which an individual can exert control over their social relationships is constrained by the structural and functional characteristics of their social context.

Greater emphasis has been placed on the individual's goals and behaviors when dealing with others (Baltes & Carstensen, 1996; Carstensen, 1993a, 1995; Marsiske et al., 1995). It has been suggested, for example, that the adaptivity of social interactions does not depend solely on the behaviors and competencies of the individual's social partners but also on the individual's own social behaviors and competence (Hansson & Carpenter, 1994; Marsiske et al., 1995). Older individuals, as much as younger individuals, actively contribute to the positive and negative outcomes of their social relationships by selecting and maintaining relationships with particular social partners as well as influencing the specific social transactions that occur within a given social relationship (Lang & Baltes, in press). Thus, when dealing with others, older individuals actively contribute to the maximization of positive gratification or benefits and minimization of negative experience. For example, older individuals do not just receive social support, they also provide support to other people, often to those people with whom they have close relationships or friendships for many years.

7.03.4.1 Socioemotional Selectivity Theory: Emotion Regulation in Late Life

According to socioemotional selectivity theory (Carstensen, 1987, 1991, 1992, 1993a), basic functions of social interaction, such as regulating desired emotional states and acquiring knowledge about oneself or the world, differ in respect to their relative importance for determining social preferences across the life span. For example, Carstensen (1993a) argues that emotional regulation becomes more important in old age because a more limited future alters the likelihood that long-term goals will be achieved. Hence, emphasis is shifted to achieving short-term goals in the immediate social context that are mostly emotional rather than informational.

Socioemotional selectivity theory provides a theoretical framework in which most of the above-mentioned empirical findings on age-related changes and health-related functions of social relationships can be considered. Older people narrow the breadth of their social lives and social participation in order to intensify their engagement in a narrow range of social functions, focusing on close emotional relationships with peers as well as with younger social partners from whom they derive a sense of meaning in life. Reduced social contact in late life, according to this model, reflects the older individual's changed needs and motivations. Indeed, the profile of empirical findings reviewed suggests that it is not the frequency of social interaction or the number of network partners but rather the quality, emotional meaningfulness, and functional adequacy of enduring close ties that contribute to well-being in late life (Beckman, 1981; Ishii-Kuntz, 1990; Lee & Markides, 1990; Lowenthal & Haven, 1968; Rook, 1984).

7.03.4.2 Social Convoy and Support Efficacy: Maintaining Competence in Late Life

The social convoy model of social relations (Antonucci, 1985, 1990, 1991; Antonucci &

Akiyama, 1987a, 1995; Kahn & Antonucci, 1980) describes a conceptual framework for age-related changes in structural and functional characteristics of social networks. Social convoys serve as a metaphorical illustration of the dynamic and age-related changes in social networks over the life span. Throughout life, individuals construct and maintain social relationships with others who also undergo similar changes in social preferences and social behaviors. This also implies that individuals become increasingly aware of specific strengths and weaknesses of their convoy members over time.

This ever-increasing knowledge base allows people to better select particular social network members for particular functions. In other words, people learn over time who best to turn to when emotional support is needed, which social partners are best able to help with activities of daily living, and who is best suited to help in times of crisis. In short, accumulated expertise about others and about relationship rules grows over a lifetime, further guiding and refining access to potential functions of the social world. Through repeated interactions with others, for example, people learn how to avoid critical or unpleasant issues that could result in conflict or negative outcomes. The importance of social convoys to psychological status may be particularly important in old age when many social structural features of the environment (e.g., work) disappear and self-defining roles (e.g., parent) become less active at an instrumental level.

In addition to the social convoy model, the support-efficacy model of social relationships (Antonucci & Jackson, 1987; Lang, Featherman, & Nesselroade, in press) extends the perspective to the individual's competence and self-efficacy in utilizing social resources. Rather than construing social support as a unidirectional process, the support-efficacy model characterizes social support as a process that depends on the personal characteristics of a target individual and support-giver within a specific social opportunity structure. The support-efficacy model is truly relational in that social support is conceptualized as a social transaction of at least two interaction partners who share a specific relationship. In the support-efficacy model, the support-receiver is considered an active agent in the social support process who exerts influence on when and from whom social support is received. On the one hand, social support can compensate for deficits in competence and thus contribute to the older person's adjustment. For example, Hobfoll (Hobfoll, 1989; Hobfoll & Freedy, 1990) has shown that the more social and personal resources are available, the better older people compensate for age-related losses. On the other hand, relying on social support resources also places specific demands on the older individual such as the normative obligation to reciprocate received supports (Antonucci & Jackson, 1987; Gouldner, 1960; Rook, 1987b; Wentowski, 1981).

7.03.5 CONCLUSIONS AND OUTLOOK

In this chapter, it has been argued that the interplay of changes in relational contexts and relational behaviors in later life may determine negative and positive outcomes of social resources. Thus, social adaptivity in later life depends not just on characteristics of the older individual's relational contexts but also on the older individual's needs and preferences within a given structure of relational opportunities. Negative outcomes of social relationships may be a function of a mismatch between the older individual's relational preferences and individuals' actual relational contexts (Lang & Baltes, in press)

Older individuals contribute actively to the positive and negative outcomes of their social relationships, not only in terms of whom they choose to interact with but also in terms of how they engage in specific relational behaviors. Thus, an older individual may contribute actively to the maximization of positive relational outcomes and the minimization of negative relational outcomes. For example, older individuals are not simply active in help-seeking and solicitation of social support from others (Nadler, 1991), they are also support providers to others with whom they have entertained relationships for many years. Providing social support may be seen as the most pervasive illustration of the individual's active and optimizing role in social relationships. Giving support entails social rewards such as feeling needed, trusted or relied upon by others (Thomae, 1994). This also addresses the transactional nature of social relationships. In most cases, when social partners seek or utilize the older individual's advice or other forms of support, this also conveys some appraisal of the older person's competence and thus contributes to feelings of control and social connectedness. Finally, giving support may also enhance the receiver's willingness to reciprocate support in future and, thus, enhance the stability and closeness of the supportive relationship. In this sense, increasing opportunities for older people to provide social support to other people may also support social functioning in later life.

7.03.6 REFERENCES

Adams, R. G., & Blieszner, R. (1994). An integrative conceptual framework for friendship research. *Journal of Social and Personal Relationships, 11,* 163–184.

Aldous, J. (1987). New views on the family life of the elderly and the near-elderly. *Journal of Marriage and the Family, 49,* 227–234.

Alwin, D. F., Converse, P. E., & Martin, S. S. (1985). Living arrangements and social integration. *Journal of Marriage and the Family, 47,* 319–334.

Antonucci, T. C. (1985). Personal characteristics, social support, and social behavior. In R. H. Binstock & E. Shanas (Eds.), *Handbook of aging and the social sciences* (2nd ed., pp. 94–128). New York: Van Nostrand Reinhold.

Antonucci, T. C. (1990). Social supports and social relationships. In R. H. Binstock & L. K. George (Eds.), *Handbook of aging and the social sciences.* (3rd ed., pp. 205–226). San Diego, CA: Academic Press.

Antonucci, T. C. (1991). Attachment, social support, and coping with negative life events in mature adulthood. In E. M. Cummings, A. L. Greene, & K. H. Karraker (Eds.), *Life-span developmental psychology: Perspectives on stress and coping* (pp. 261–276). Hillsdale, NJ: Lawrence Erlbaum.

Antonucci, T. C., & Akiyama, H. (1987). Social networks in adult life and a preliminary examiation of the convoy model. *Journals of Gerontology, 42,* 519–527.

Antonucci, T. C., & Akiyama, H. (1994). Convoys of attachment and social relations in children, adolescents, and adults. In F. Nestmann & K. Hurrelmann (Eds.), *Social networks and social support in childhood and adolescence* (pp. 37–52). Berlin: de Gruyter.

Antonucci, T. C., & Akiyama, H. (1995). Convoys of social relations: Family and friendships within a life span context. In R. Blieszner & V. H. Bedford (Eds.), *Handbook of aging and the family* (pp. 355–371). Westport, CT: Greenwood Press.

Antonucci, T. C., & Jackson, J. S. (1987). Social support, interpersonal efficacy, and health: A life course perspective. In L. L. Carstensen & B. A. Edelstein (Eds.), *Handbook of clinical gerontology* (pp. 291–311) New York: Pergamon.

Armstrong, M. J., & Goldsteen, K. S. (1990). Friendship support patterns of older American women. *Journal of Aging Studies, 4,* 391–404.

Atchley, R. C. (1993). Continuity theory and the evolution of activity in later adulthood. In J. R.. Kelly (ed.), *Activity and aging: Staying involved in later life* (Suppl. pp. 5–16). Newbury Park, CA: Sage.

Atkins, C. J., Kaplan, R. M., & Toshima, M. T. (1992). Close relationships in the epidemiology of cardiovascular disease. In. W. H. Jones & D. Perlman (Eds.), *Advances in personal relationships* (Vol. 3, pp. 207–231). London: Kingsley.

Avioli, P. S. (1989). The social support functions of siblings in later life: A theoretical model. *American Behavioral Scientist, 33,* 45–57.

Babchuck, N. (1978–79). Aging and primary relations. *International Journal of Aging and Human Development, 9,* 137–151.

Babchuk, N., & Anderson, T. B. (1989). Older widows and married women: Their intimates and confidants. *International Journal of Aging and Human Development, 28,* 21–35.

Baltes, M. M. (1995). Dependency in old age: Gains and losses. *Current Directions in Psychological Science, 4,* 14–19.

Baltes, M. M., & Carstensen, L. L. (1996). The process of successful aging. *Ageing and Society, 16,* 397–422.

Baltes, M. M., Neumann, E. M., & Zank, S. (1994). Maintenance and rehabilitation of independence in old age: An intervention program for staff. *Psychology and Aging, 9,* 179–188

Baltes, M. M., & Silverberg, S. B. (1994). The dynamics between dependency and autonomy: Illustrations across the life-span. In D. L. Featherman, R. M. Lerner, & M. Perlmutter (Eds.), *Life-span development and behavior* (Vol. 12, pp. 41–90). Hillsdale, NJ: Erlbaum.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37,* 122–147.

Barefoot, J. C., Siegler, I. C., Nowlin, J. B., Peterson, B. L., Haney, T. L., & Williams, R. B. (1987). Suspiciousness, health, and mortality: A follow-up study of 500 older adults. *Psychosomatic Medicine, 49,* 450–457.

Bass, D. M., & Bowman, K. (1990). The transition from caregiving to bereavement: The relationship of care-related strain and adjustment to death. *The Gerontologist, 30,* 35–44.

Bass, D. M., Bowman, K., & Noelker, L. S. (1991). The influence of caregiving and bereavement support on adjusting to an older relative's death. *The Gerontologist, 31,* 32–42.

Beckman, L. J. (1981). Effects of social interaction and children's relative inputs on older women's psychological well-being. *Journal of Personality and Social Psychology, 41,* 1075–1086.

Bedford, V. H. (1995). Sibling relationships in middle and old age. In R. Blieszner & V. H. Bedford (Eds.), *Handbook of aging and the family* (pp. 201–222). Westport, CT: Greenwood Press.

Bengtson, V. L. (1985). Diversity and symbolism in grandparental roles. In V. L. Bengtson & J. F. Robertson (Eds.), *Grandparenthood* (pp. 11–26). Beverly Hills, CA: Sage.

Berkman, L. F. (1985). The relationship of social networks and social support to morbidity and mortality. In S. Cohen & S. L. Syme (Eds.), *Social support and health* (pp. 241–262). New York: Academic Press.

Berkman, L. F., & Syme, S. L. (1979). Social networks, host resistance, and mortality: A nine-year follow-up study of Alameda County residents. *American Journal of Epidemiology, 109,* 186–204.

Blau, Z.S. (1973). *Old age in a changing society*. New York: Franklin Watts.

Bolger, N., & Kelleher, S. (1993). Daily life in relationships. In S. Duck (Eds.), *Social context and relationships* (pp. 100–108). Newbury Park, CA: Sage.

Bosse, R., Aldwin, C. M., Levenson, M. R., Spiro A. III, & Mroczek, D. K. (1993). Change in social support after retirement: Longitudinal findings from the normative aging study. *Journal of Gerontology: Psychological Sciences, 48,* P210–P217.

Bowling, A. (1988–89). Who dies after widow(er)hood? A discriminant analysis. *Omega: Journal of Death and Dying, 19,* 135–153.

Bowling, A. (1994). Social networks and social support among older people and implications for emotional well–being and psychiatric morbidity. *International Review of Psychiatry, 6,* 41–58.

Bowling, A., & Browne, P. D. (1991). Social networks, health, and emotional well–being among the oldest old in London. *Journal of Gerontology: Psychological Sciences, 46,* S20–32.

Bowling, A., & Windsor, J. (1995). Death after widow(er)-hood: An analysis of mortality rates up to 13 years after bereavement. *Omega: Journal of Death and Dying, 31,* 35–49.

Broadhead, W. E., Kaplan, B. H., James, S. A., Wagner, E. H., Schoenbach, V. J., Grimson, R., Heyden, S., Tibblin, G., & Gehlback, S. H. (1983). The epidemiologic evidence for a relationship between social support and health. *American Journal of Epidemiology, 117,* 521–537.

Brody, E. M. (1985). Parent care as a normative family stress. *Gerontologist, 25,* 19–29.

Brubaker, T. H. (1990). Families in later life: A burgeoning research area. *Journal of Marriage and the Family, 52,* 959–981.

Bryant, S., & Rakowski, W. (1992). Predictors of mortality among elderly African-Americans. *Research on Aging, 14,* 50-67.

Cantor, M. H. (1979). Neighbors and friends: An overlooked resource in the informal support system. *Research on Aging, 1,* 434–463.

Caplan, G. (1982). The family as a support system. In H. I. McCubbin, A. E. Cauble & J. E. Patterson (Eds.), *Family stress, coping, and social support* (pp. 200–220). Springfield, IL: Thomas.

Carp, F. M. (Ed.) (1972). *Retirement*. New York: Behavioral Publications.

Carpenter, B. N. (1993). Relational competence. In D. Perlman & W. H. Jones (Eds.), *Advances in personal relationships* (Vol. 4, pp. 1–28). London: Kingsley.

Carstensen, L. L. (1987). Age-related changes in social activity. In L. L. Carstensen & B. A. Edelstein (Eds.), *Handbook of clinical gerontology* (pp. 222–237). New York: Pergamon.

Carstensen, L. L. (1991). Socioemotional selectivity theory: Social activity in life-span context. *Annual Review of Gerontology and Geriatrics, 11,* 195–217.

Carstensen, L. L. (1992). Social and emotional patterns in adulthood: Support for socioemotional selectivity theory. *Psychology and Aging, 7,* 331–338.

Carstensen, L. L. (1993a). Motivation for social contact across the life span: A theory of sociomotional selectivity. *Nebraska symposium on motivation* (pp. 209–254). Lincoln, NE: University of Nebraska Press.

Carstensen, L. L. (1993b). Perspective on research with older families: Contributions of older adults to families and to family theory. In P. Cowan, D. Field, D. Hansen, A. Skolnick, & G. E. Swanson (Eds.), *Family, self, and society: Toward a new agenda for family research* (pp. 353–359). Hillsdale, NJ: Lawrence Erlbaum.

Carstensen, L. L. (1995). Evidence for a life-span theory of socioemotional selectivity. *Current directions in Psychological Science, 5,* 151–156.

Carstensen, L. L., & Lang, F. R. (in preparation). *Social preferences and social contact behavior from early to late adulthood.* Stanford University.

Carstensen, L. L., & Lang, F. R. (in press). Social relationships in context and as context: Social support and the maintenance of competence in old age. In S. Willis, K. W. Schaie, & M. Hayward (Eds.), *Societal mechanisms for maintaining competence in old age* (pp. 207–222). New York: Springer.

Carstensen, L. L., Gottman, J. M., & Levenson, R. W. (1995). Emotional behavior in long-term marriage. *Psychology and Aging, 10,* 140–149.

Caspi, A., Bolger, N. & Eckenrode, J. (1987). Linking person and context in the daily stress process. *Journal of Personality and Social Psychology, 52,* 184–195.

Chappell, N. L. (1990). Aging and social care. In R. H. Binstock & L. K. George (Eds.), *Handbook of aging and the social sciences* (3rd ed., pp. 438–454). San Diego, CA: Academic Press.

Chappell, N. L. (1991). Living arrangements and sources of caregiving. *Journal of Gerontology: Social Sciences, 46,* S1–8.

Chappell, N. L., & Badger, M. (1989). Social isolation and well-being. *Journal of Gerontology: Social Sciences, 44,* 5169–5176

Chatters, L. M., Taylor, R. J., & Jackson, J. S. (1986). Aged blacks' choices for an informal helper network. *Journal of Gerontology, 41,* 94–100.

Cicirelli, V. G. (1989). Feelings of attachment to siblings and well-being in later life. *Psychology and Aging, 4,* 211–216.

Cicirelli, V. G. (1995). *Sibling relationships across the life span*. New York : Plenum.

Cohen, C. I., Teresi, J., & Holmes, D. (1986–87). Social networks and mortality in an inner-city elderly population. *International Journal of Aging and Human Development, 24,* 257–269.

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin, 98,* 310–357.

Connidis, I. A. (1989). Siblings as friends in later life. *American Behavioral Scientist, 33,* 81–93.

Connidis, I. A. (1994). Sibling support in older age. *Journals of Gerontology: Social Sciences, 49,* S309–S317.

Connidis, I. A., & Davis, L. (1992). Confidants and companions: Choices in later life. *Journal of Gerontology: Social Sciences, 47,* S115–122.

Connidis, I. A., & McMullin, J. A. (1993). To have or have not: Parent status and the subjective well-being of older men and women. *The Gerontologist, 33,* 630–636.

Connidis, I. A., & McMullin, J. A. (1994). Social support in older age: Assessing the impact of marital and parent status. *Canadian Journal on Aging, 13,* 510–527.

Costa, P. T., & McCrae, R. R. (1994). Set like plaster: Evidence for the stability of adult personality. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change?* (pp. 21–40). Washington, DC: APA.

Coyne, J. C. (1976). Depression and the response of others. *Journal of Abnormal Psychology, 85,* 186–193.

Coyne, J. C., Wortman, C. B., & Lehmann, D. R. (1988). The other side of support. Emotional overinvolvement and miscarried helping. In B. H. Gottlieb (Ed.), *Marshaling social support. Formats, processes, and effects* (pp. 305–330). Newbury Park, CA: Sage.

Cramer, D. (1993). Personality and marital dissolution. *Personality and Individual Differences, 14,* 605–607.

Crohan, S. E., & Antonucci, T. C. (1989). Friends as a source of social support in old age. In R. G. Adams & R. Blieszner (Eds.), *Older adult friendship: Structure and process* (pp. 129–146). Newbury Park, CA: Sage.

Cumming, E., & Henry, W. E. (1961). *Growing old: The process of disengagement*. New York: Basic Books.

Cutrona, C. E., Cohen, B. B., & Igram, S. (1990). Contextual determinants of the perceived supportiveness of helping behaviors. *Journal of Social and Personal Relationships, 7,* 553–562.

Dean, A., Matt, G. E., & Wood, P. (1992). The effects of widowhood on social support from significant others. *Journal of Community Psychology, 20,* 309–325.

Depner, C. E., & Ingersoll-Dayton, B. (1985). Conjugal social support: Patterns in later life. *Journal of Gerontology, 40,* 761–766.

Duck, S. (1986). *Human relationships: An introduction to social psychology*. London: Sage.

Dunkel-Schetter, C., & Bennett, T. L. (1990). Differentiating the cognitive and behavioral aspects of social support. In B. R. Sarason, I. G. Sarason, & G. R. Pierce (Eds.), *Social support: An interactional view* (pp. 267–296). New York: Wiley.

Dunkel -Schetter, C., Folkman, S., & Lazarus, R. S. (1987). Correlates of social support receipt. *Journal of Personality and Social Psychology, 53,* 71–80.

Dykstra, P. A. (1993). The differential availability of relationships and the provision and effectiveness of support to older adults. *Journal of Social and Personal Relationships, 10,* 355–370.

Dykstra, P. A. (1995). Loneliness among the never and formerly married: The importance of supportive friendships and a desire for independence. *Journals of Gerontology: Series B, Psychological and Social Sciences, 50B,* S321–S329.

Eggebeen, D. J. (1992). From generation unto generation: Parent-child support in aging American families. *Generations, 16,* 45–49.

Eggebeen, D. J., & Hogan, D. P. (1990). Giving between the generations in American families. *Human Nature, 1,* 211–232.

Felton, B. J., & Berry, C. A. (1992). Do the sources of the urban elderly's social support determine its psychological consequences? *Psychology & Aging, 7,* 89–97.

Ferraro, K. F., & Barresi, C. M. (1982). The impact of widowhood on the social relations of older people. *Research on Aging, 4,* 227–247.

Ferraro, K. F., & Farmer, M. M. (1995). Social compensation in adulthood and later life. In R. Dixon & L. Bäckman (Eds.), *Compensating for psychological deficits and declines* (pp. 127–145). Hillsdale, NJ: Erlbaum.

Field, D., & Minkler, M. (1988). Continuity and change in social support between young-old and old-old or very-old age. *Journals of Gerontology: Psychological Sciences, 43* (4), P100–106.

Field, D., & Minkler, M. (1993). The importance of family in advanced old age: The family is "forever" In P. Cowan, D. Field, D. Hansen, A. Skolnick, & G. E. Swanson (Eds.), *Family, self, and society: Toward a new agenda for family research* (pp. 331–351). Hillsdale, NJ: Lawrence Erlbaum.

Field, D., Minkler, M., Falk, R. F., & Leino, E. V. (1993). The influence of health on family contacts and family feelings in advanced old age: A longitudinal study. *Journal of Gerontology: Psychological Sciences, 48,* P18–P28.

Filipp, S. H., & Aymanns, P. (1987). Die Bedeutung sozialer und personaler Ressourcen in der Auseinandersetzung mit kritischen Lebensereignissen [The significance of social and personal resources in dealing with critical life events]. *Zeitschrift fur Klinische Psychologie, 16,* 383–396.

Finch, J. F., Okun, M. A., Barrera, M., Zautra, A., & Reich, J. W. (1989). Positive and negative social ties among older adults: Measurement models and the prediction of psychological distress and well-being. *American Journal of Community Psychology, 17,* 585–605.

Fisher, J. D., Nadler, A., & Whitcher-Alagna, S. (1982). Recipient reactions to aid. *Psychological Bulletin, 91,* 27–54.

Franks, M. M., & Stephens, M. A. P. (1996). Social support in the context of caregiving: Husbands' provision of support to wives involved in parent care. *Journals of Gerontology: Psychological Sciences, 51B,* P43–P52.

Fredrickson, B. L., & Carstensen, L. L. (1990). Choosing social partners: How old age and anticipated endings make people more selective. *Psychology and Aging, 5,* 335–347.

Friedman, H. S., Tucker, J. S., Schwartz, J. E., Tomlinson–Keasy, C., Martin, L. R., Wingard, D. L., & Criqui, M. H. (1995). Psychosocial and behavioral predictors of longevity: The aging and death of the Termites. *American Psychologist, 50,* 69–78.

Ganster, D. C. , & Victor, B. (1988). The impact of social support on mental and physical health. *British Journal of Medical Psychology, 61,* 17–36.

Garrett, J. E. (1987). Multiple losses in older adults. J*ournal of Gerontological Nursing, 13*, 8–12.

George, L. K., & Gwyther, L. P. (1986). Caregiver well-being: A multidimensional examination of family caregivers of demented adults. *Gerontologist, 26,* 253–259.

Goetting, A. (1986). The developmental tasks of siblingship over the life cycle. *Journal of Marriage and the Family, 48,* 703–714.

Gold, D. T. (1987). Siblings in old age: Something special. *Canadian Journal on Aging, 6,* 199–215.

Gordon, C., & Gaitz, C. (1976). Leisure and lives. In R. Binstock & E. Shanas (Eds.), *Handbook of aging and the social sciences* (Vol. 1, pp. 310–341). New York: Van Nostrand Reinhold.

Gotlib, I. H., & Hooley, J. M. (1988). Depression and marital distress: Current status and future directions. In S. Duck (Ed.), *Handbook of personal relationships: Theory, research and interventions* (pp. 543–570). Chichester, UK: Wiley.

Gotlib, I. H., & Whiffen, V. E. (1992). The interpersonal context of depression. In W. H. Jones & D. Perlman (Eds.), *Advances in personal relationships* (Vol. 3, pp. 177–206). London: Kingsley.

Gottlieb, B. H. (1983). Social support as a focus for integrative research in psychology. *American Psychologist, 38,* 278–287.

Gottman, J. M., & Levenson, R. W. (1992). Marital processes predictive of later dissolution: Behavior, physiology, and health. *Journal of Personality and Social Psychology, 63,* 221–233.

Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review, 25,* 161–178.

Hagestad, G. O. (1985). Continuity and connectedness. In V. L. Bengtson & J. F. Robertson (Eds.), *Grandparenthood* (pp. 31–48). Beverly Hills, CA: Sage.

Hagestad, G. O. (1987). Parent-child relations in later life: Trends and gaps in past research. In J. B. Lancaster, J. Altmann, A. S. Rossi, & L. R. Sherrod (Eds.), *Parenting across the life span: Biosocial dimensions.* New York: Aldine de Gruyter.

Hammer, M. (1983). "Core" and "extended" social networks in relation to health and illness. *Social Science and Medicine, 17,* 405–411.

Hammond, R. J., & Muller, G. O. (1992). The late-life divorced: Another look. *Journal of Divorce & Remarriage, 17,* 135–150.

Hanson, B. S., Isacsson, S.-O., Janzon, L., & Lindell, S.-E. (1990). Social network and social support influence mortality in elderly men. *Advances, 7,* 16–18.

Hansson, R. O., & Carpenter, B. N. (1994). *Relationships in old age: Coping with the challenge of transition.* New York: Guilford.

Hansson, R. O., Jones, W. H., & Fletcher, W. L. (1990). Troubled relationships in later life: Implications for support. *Journal of Social and Personal Relationships, 7,* 451–463.

Harvey, A. S., & Singleton, J. F. (1989). Canadian activity patterns across the life span: A time budget perspective. *Canadian Journal on Aging, 8,* 268–285.

Hautzinger, M., Linden, M., & Hoffman, N. (1982). Distressed couples with and without a depressed partner: An analysis of their verbal interaction. *Journal of Behavior Therapy and Experimental Psychology, 13,* 307–314.

Havighurst, R. J., & Albrecht, R. (l953). *Older people.* New York: Longmans.

Havighurst, R., Neugarten, B., & Tobin, S. (1968). Disengagement and patterns of aging. In B. Neugarten (Ed.), *Middle age and aging* (pp. 161–172). Chicago, IL: University of Chicago Press.

Havighurst, R., Munnichs, J. M. A., Neugarten, B., & Thomae, H. (Eds.) (1969), *Adjustment to retirement: A crossnational study.* Assen,The Netherlands: Van Gorcum.

Heckhausen, J., & Lang, F. R. (1996). Social construction in old age: Normative conceptions and interpersonal processes. In G. R. Semin & K. Fiedler (Eds.), *Applied social psychology* (pp. 374–398). London: Sage.

Helgeson, V. S., & Cohen, S. (1996). Social support and adjustment to cancer: Reconciling descriptive, correlational, and intervention research. *Health Psychology, 15,* 135–148.

Heller, K., & Mansbach, W. E. (1984). The multifaceted nature of social support in a community sample of elderly women. *Journal of Social Issues, 40,* 99–113.

Henderson, A. S., Scott, R., & Kay, D. W. (1986). The elderly who live alone: Their mental health and social relationships. *Australian and New Zealand Journal of Psychiatry, 20,* 202–209.

Herman, S. M. (1994). Marital satisfaction in the elderly. *Gerontology and Geriatrics Education, 14,* 69–79.

Herzog, R. A., & Rodgers, W. L. (1981). Age and satisfaction: Data from several large surveys. *Research on Aging, 3,* 142–165.

Hess, B. B., & Soldo, B. J. (1985). Husband and wife networks. In W. J. Sauer & R. T. Coward (Eds.), *Social support networks and care of the elderly: Theory, research, and practice* (pp. 67–92). New York: Springer.

Hobfoll, S. E. (1989). Conservation of resources. A new attempt at conceptualizing stress. *American Psychologist, 44,* 513–524.

Hobfoll, S. E., & Freedy, J. R. (1990). The availability and effective use of social support. *Journal of Social and Clinical Psychology, 9,* 91–103.

Hodgson, L. G. (1992). Adult grandchildren and their grandparents: Their enduring bond. *International Journal of Aging and Human Development, 34,* 209–225.

Hogan, D. P. (1995). Sources of emergency help and routine assistance in old age. *Social Forces, 73,* 917–936.

Hogg, J. R., & Heller, K. (1990). A measure of relational competence for community-dwelling elderly. *Psychology and Aging, 5,* 580–588.

Holahan, C. J., Moos, R. H., Holahan, C. K., & Brennan, P. L. (1995). Social support, coping, and depressive symptoms in a late-middle-aged sample of patients reporting cardiac illness. *Health Psychology, 14,* 152–163.

Holmes, J. G. (1991). Trust and the appraisal process in close relationships. In W. H. Jones & D. Perlman (Eds.), *Advances in personal relationships: A research annual* (Vol. 2, pp. 57–104). London, Kingsley.

Holmes, T., & Rahe, R. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research, 11,* 213–218.

Horowitz, A. (1985). Family caregiving to the frail elderly. In C. Eisdorfer (Ed.), *Annual review of gerontology and geriatrics* (Vol. 5, pp. 194–246). New York: Springer.

House, J. S., & Kahn, R. L. (1985). Measures and concepts of social support. In S. Cohen & S. L. Syme (Eds.), *Social support and health* (pp. 83–105). New York: Academic Press.

House, J. S., Robbins, C., & Metzner, H. M. (1982). The association of social relationships and activities with mortality: Prospective evidence from the Tecumset Community Health Study. *American Journal of Epidemiology, 116,* 123–140.

House, J. S., Umberson, D., & Landis, K. R. (1988). Structures and processes of social support. *Annual Review of Sociology, 14,* 293–318.

Huyck, M. H. (1995). Marriage and close relationships of the marital kind. In R. Blieszner & V. H. Bedford (Eds.), *Handbook of aging and the family* (pp. 181–200). Westport, CT: Greenwood Press.

Ikels, C. (1988). Delayed reciprocity and the support networks of the childless elderly. *Journal of Comparative Family Studies, 19,* 99–112.

Ishii-Kuntz, M. (1990). Social interaction and psychological well-being: Comparison across stages of adulthood. *International Journal of Aging and Human Development, 30,* 1, 15–36.

Johnson, C. L. (1983). Fairweather friends and rainy day kin: An anthropological analysis of old age friendship in the United States. *Urban Anthropology, 12,* 103–123.

Johnson, C. L., & Catalano, D. J. (1983). A longitudinal study of family support to impaired elderly. *The Gerontologist, 23,* 612–618.

Johnson, C. L., & Troll, L. E. (1992). Family functioning in late late life. *Journals of Gerontology: Social Sciences, 47,* S66–S72.

Johnson, C. L., & Troll, L. E. (1994). Constraints and facilitators to friendships in late life. *The Gerontologist, 34,* 79–87.

Jones, D. C., & Vaughan, K. (1990). Close friendships among senior adults. *Psychology and Aging, 5,* 451–457.

Jung, J. (1990). The role of reciprocity in social support. *Basic and Applied Social Psychology, 11,* 243–253.

Kahn, R. L. (1979). Aging and social support. In M. W. Riley (Ed.), *Aging from birth to death: Interdisciplinary perspectives* (pp. 77–91). Boulder, CO: Westview Press.

Kahn, R. L., & Antonucci, T. C. (1980). Convoys over the life course. Attachment, roles and social support. In P. B. Baltes & O. G. Brim (Eds.), *Life-span development and behavior* (pp. 254–283). New York: Academic Press.

Keith, P. M. (1986). Isolation of the unmarried in later life. *Family Relations: Journal of Applied Family and Child Studies, 35,* 389–395.

Keith, P. M., Hill, K., Goudy, W. J., & Powers, E. A. (1984). Confidants and well-being: A note on male friendship in old age. *The Gerontologist, 24,* 318– 320.

Kennedy, G. E. (1992). Quality in grandparent/grandchild relationships. *International Journal of Aging and Human Development, 35,* 83–98

Kennedy, S., Kiecolt-Glaser, J. K., & Glaser, R. (1990). Social support, stress, and the immune system. In B. R. Sarason, I. G. Sarason, G. R. Pierce (Eds.), *Social support: An interactional view* (pp. 253–266). New York: Wiley.

Kessler, R. C., & McLeod, J. D. (1985). Social support and mental health in community samples. In S. Cohen & S. L. Syme (Eds.). *Social support and health* (pp. 219–240). New York: Academic Press.

Kivnick, H. Q. (1982). *The meaning of grandparenthood.* Ann Arbor, MI: UMI Research Press.

Krause, N. (1994). Stressors in salient social roles and well-being in later life. *Journal of Gerontology: Psychological Sciences, 49,* P137–148.

Krause, N., Herzog, A. R., & Baker, E. (1992). Providing support to others and well-being in later life. *Journals of Gerontology, 47,* P300–P311.

Kriegsman, D. M. W., Penninx, B. W. H. J., & van Eijk, J. T. M. (1995). A criterion-based literature survey of the relationship between family support and incidence and course of chronic disease in the elderly. *Family Systems Medicine, 13,* 39–68.

Kurdek, L. A. (1991). Marital stability and changes in marital quality in newly wed couples: A test of the contextual model. *Journal of Social and Personal Relationships, 8,* 27–48.

Lang, F. R. (1994). *Die Gestaltung informeller Hilfebeziehungen im hohen Alter - Die Rolle von Elternschaft und Kinderlosigkeit* [Social support management in late life-The role of parenthood and childlessness]. Berlin: Edition Sigma (Max-Planck-Institut für Bildungsforschung, Studien und Berichte, Band 59).

Lang, F. R. (1996). Social support relationships of parents and non-parents in old and very old age. In H. Mollenkopf (Ed.), *Elderly people in industrialised societies: Social integration in old age by or despite technology* (pp. 43–54). Berlin: Edition Sigma.

Lang, F. R., & Baltes, M. M. (1997). Brauchen alte Menschen junge Menschen? Überlegungen zu den Entwicklungsaufgaben imhohen Lebensalter [Do older people need young people? Considerations about the developmental tasks in late life]. In L. Krappmann (Ed.), *Alt und Jung: Spannungen und solidaritaet zwischen den Generationen* (pp. 161–184). Frankfurt: Campus.

Lang, F. R., & Baltes, M. M. (in press). Being with people and being alone in late life: Costs and benefits for everyday functioning. *International Journal of Behavioral Development.*

Lang, F. R., & Carstensen, L. L. (1994). Close emotional relationships in late life: Further support for proactive aging in the social domain. *Psychology and Aging, 9,* 315–324.

Lang, F. R., Featherman, D. L., & Nesselroade, J. R. (in press). Social self efficacy and short-term variability in personal relationships: The MacArthur Successful Aging Studies. *Psychology and Aging.*

Lang, F. R., Staudinger, U. M., & Carstensen, L. L. (1996). Perspectives on socioemotional selectivity in late life: How personality does (and does not) make a difference. *Journal of Gerontology: Psychological Science.*

Lang, F. R., & Tesch-Römer, C. (1993). Erfolgreiches Altern und soziale Beziehungen: Selektion und Kompensation im sozialen Kontaktverhalten [Successful aging and social relationships: Selection and compensation in social contact behavior]. *Zeitschrift für Gerontologie*, *26,* 321–329.

Langer, N. (1990). Grandparents and adult grandchildren: What do they do for one another? *International Journal of Aging and Human Development, 31,* 101–110.

Larson, R. W. (1990). The solitary side of life: An examination of the time people spend alone from childhood to old age. *Developmental Review, 10,* 155–183.

Larson, R., Mannell, R., & Zuzanek, J. (1986). Daily well-being of older adults with friends and family. *Psychology and Aging, 1,* 117–126.

Lawton, M. P. (1989). Behavior-relevant ecological factors. In K. W. Schaie & C. Schooler (Eds.), *Social structure and aging: Psychological processes* (pp. 57–78). Hillsdale, NJ: Lawrence Erlbaum.

Lawton, M. P., Moss, M., & Fulcomer, M. (1986–87). Objective and subjective uses of time by older people. *International Journal of Aging and Human Development, 24,* 171–188.

Lee, G. R. (1985).Theoretical perspectives on social networks. In W. J. Sauer & R. T. Coward (Eds.), *Social support networks and the care of the elderly: Theory, research, and practice* (pp. 21–37). New York: Springer.

Lee, D. J., & Markides, K. S. (1990). Activity and mortality among aged persons over an eight-year period. *The Journals of Gerontology: Social Sciences, 45,* S39–S42.

Lehr, U., & Minnemann, E. (1987). Veränderung von Quantität und Qualität sozialer Kontakte vom 7. bis 9. Lebensjahrzehnt [Changes of quantity and quality of social contact from the 7th to the 9th decade of life]. In U.Lehr & H. Thomae (Eds.), *Formen seelischen Alterns. Ergebnisse der Bonner Gerontologischen Längsschnittstudie* (pp. 80–91). Stuttgart: Enke.

Lemon, B. W., Bengtson, V. L., Peterson, I. A. (1972). An exploration of the activity theory of aging: Activity types and life satisfaction among inmovers to a retirement community. *Journal of Gerontology, 27,* 511–523.

Levenson, R. W., Carstensen, L. L., & Gottman, J. M. (1993). Long-term marriage: Age, gender, and satisfaction. *Psychology and Aging, 8,* 301–313.

Levenson, R. W., Carstensen, L. L.. & Gottman, J. M. (1994). The influence of age and gender on affect, physiology, and their interrelations: A study of long-term marriages. *Journal of Personality and Social Psychology, 67,* 56–68.

Levitt, M. J., Weber, R. A., & Guacci, N. (1993). Convoys of social support: An intergenerational analysis. *Psychology and Aging, 8,* 323–326.

Litwak, E. (1985). *Helping the elderly.* New York: Guilford.

Lopata, H. Z. (1986). Becoming and being a widow: Reconstruction of the self and support systems. *Journal of Geriatric Psychiatry, 19,* 203–214.

Lowenthal, M. F., & Haven, C. (1968). Interaction and adaptation: Intimacy as a critical variable. *American Sociological Review, 33,* 20–30.

Lowenthal, M. F., Thurnher, M., & Chiriboga, D. (1976). *Four stages of life : A comparative study of women and men facing transitions.* San Francisco: Jossey-Bass.

Maddox, G. L. (1963). Activity and morale: A longitudinal study of selected elderly subjects. *Social Forces, 42,* 195–204.

Mancini, J. A. & Blieszner, R. (1989).Aging parents and adult children: Research themes in intergenerational relations. *Journal of Marriage and the Family, 51,* 275–290.

Markides, K. S., Boldt, J. S., & Ray, L. A. (1986). Sources of helping and intergenerational solidarity: A three-generations study of Mexican Americans. *Journal of Gerontology, 41,* 506–511.

Marsiske, M., Lang, F. R., Baltes, P. B., & Baltes, M. M. (1995). Selective optimization with compensation: Life-span perspectives on successful human development. In R. Dixon & L. Bäckman (Eds.), *Compensating for psychological deficits and declines* (pp. 35–79). Hillsdale, NJ: Erlbaum.

McAdams, D. P. , de St. Aubin, E., & Logan, R. L. (1993). Generativity among young, midlife, and older adults. *Psychology and Aging, 8,* 221–230.

Medalie, J., & Goldbourt, V. (1976). Angina pectoris among 10,000 men: Psychosocial and other risk factors as evidenced by a multivariate analysis of a five-year incidence study. *American Journal of Medicine, 60,* 910–924.

Miller, S. S., & Cavanaugh, J. C. (1990). The meaning of grandparenthood and its relationship to demographic, relationship, and social participation variables. *Journals of Gerontology: Psychological Sciences, 45,* P244–P246.

Morgan, D. L. (1989). Adjusting to widowhood: Do social networks really make it easier? *The Gerontologist. 29,* 101–107.

Morgan, D. L. (1990). Combining the strengths of social networks, social support, and personal relationships. In S. Duck & R. C. Silver (Eds.), *Personal relationships and social support* (pp. 190–215). Newbury Park, CA: Sage.

Moss, M. S., Lesher, E., & Moss, S. Z. (1986–87). Impact of the death of an adult child on elderly parents: Some observations. *Omega: Journal of Death and Dying, 17,* 209–218.

Moss, S. Z., & Moss, M. S. (1989). The impact of the death of an elderly sibling. *American Behavioral Scientist, 33,* 94–106.

Mullan, J. T. (1992). The bereaved caregiver: A prospective study of changes in well-being. *The Gerontologist, 32,* 673–683.

Mullins, L. C., & Mushel, M. (1992). The existence and emotional closeness of relationships with children, friends, and spouses: The effect of loneliness among older persons. *Research on Aging, 14,* 448–470.

Mullins, L. C., Sheppard, H. L., & Andersson, L. (1991). Loneliness and social isolation in Sweden: Differences in age, sex, labor force status, self-rated health, and income adequacy. *Journal of Applied Gerontology, 10,* 455–468.

Nadler, A. (1991). Help-seeking behavior: Psychological costs and instrumental benefits. In M. S. Clark (Ed.), *Prosocial behavior* (pp. 290–311). Newbury Park, CA: Sage.

O'Bryant, S. L. (1988). Sibling support and older widows' well-being. *Journal of Marriage and the Family, 50,* 173–184.

Okun, M. A., Melichar, J. F., & Hill, M. D. (1990). Negative daily events, positive and negative social ties, and psychological distress among older adults. *The Gerontologist, 30,* 193–199.

Olsen, R. B., Olsen, J., Gunner-Svensson, F., & Waldstrom, B. (1991). Social networks and longevity: A 14 year follow-up study among elderly in Denmark. *Social Science and Medicine, 33,* 1189–1195.

Orth-Gomer, K., & Johnson, J. V. (1987). Social network interaction and mortality: A six-year follow-up study of a random sample of the Swedish population. *Journal of Chronic Diseases, 40,* 949–957.

Osgood, N. J. (1989). Theory and research in social gerontology. In N. J. Osgood & A. H. L. Sontz (Eds.), *The science and practice of gerontology. A multidisciplinary guide* (pp. 55–87). New York: Greenwood Press.

Page, R. M., & Cole, G. E. (1991). Demographic predictors of self-reported loneliness in adults. *Psychological Reports, 68,* 939–945.

Palmore, E. (1981). *Social patterns in normal aging: Findings from the Duke Longitudinal Study*. Durham, NC: Duke University Press.

Pearlin, L. I., Mullan, J. T., Semple, S. J., & Skaff, M. M. (1990). Caregiving and the stress process: An overview of concepts and their measures. *The Gerontologist, 30,* 583–594.

Peters, G. R., & Kaiser, M. A. (1985). The role of friends and neighbors in providing social support. In W. J. Sauer & R. T. Coward (Eds.), *Social support networks and care of the elderly: Theory, research, and practice* (pp. 123–158). New York: Springer.

Pillemer, K., & Finkelhor, D. (1988). The prevalence of elder abuse: A random sample survey. *The Gerontologist, 28,* 51–57.

Pillemer, K., & Suitor, J. J. (1992). Violence and violent feelings: What causes them among family caregivers? *Journals of Gerontology, 47,* S165–S172.

Pratt, M. W., & Norris, J. E. (1994). *The social psychology of aging*. Oxford: Blackwell.

Pruchno, R. A., Peters, N. D., & Burant, C. J. (1995). Mental health of coresident family caregivers: Examination of a two-factor model. *Journal of Gerontology: Psychological Sciences, 50B,* P247–P256.

Pulakos, J. (1987). Brothers and sisters: Nature and importance of the adult bond. *Journal of Psychology, 121,* 521–522.

Query, J. L., & James, A. C. (1989). The relationship between interpersonal communication competence and social support among elderly support groups in retirement communities. *Health Communication, 1,* 165–184.

Ramirez-Barranti, C. C. (1985). The grandparent-grandchild relationship: Family resource in an era of voluntary bonds. *Family Relations: Journal of Applied Family and Child Studies, 34,* 343–352.

Rawlins, W. K. (1992). *Friendship matters. Communication, dialectics, and the life course*. New York: de Gruyter.

Reifman, A. (1995). Social relationships, recovery from illness, and survival: A literature review. *Annals of Behavioral Medicine, 17,* 124–131.

Revenson, T. A., & Johnson, J. (1984). Social and demographic corelates of loneliness in late life. *American Journal of Community Psychology, 12,* 71–85.

Riley, D., & Eckenrode, J. (1986). Social ties: Subgroup differences in costs and benefits. *Journal of Personality and Social Psychology, 51,* 770–778.

Roberto, K. A. (1989). Exchange and equity in friendships. In R. G. Adams & R. Blieszner (Eds.), *Older adult friendship: Structure and process* (pp. 147–165) London: Sage.

Roberto, K. A., & Scott, J. P. (1986a). Equity Considerations in the friendships of older adults. *Journal of Gerontology, 41,* 241–247.

Roberto, K. A., & Scott, J. P. (1986b). Friendships of older men and women: Exchange patterns and satisfaction. *Psychology and Aging, 1,* 103–109.

Roberto, K. A., & Scott, J. P. (1987). Friendships in late life: A rural–urban comparison. *Lifestyles, 8,* 16–26.

Roberto, K. A., & Stanis, P. I. (1994). Reactions of older women to the death of their close friends. *Omega: Journal of Death and Dying, 29,* 17–27.

Röhrle, B. (1994). *Soziale Netzwerke und soziale Unterstützung [Social networks and social support]*. Munich: Beltz.

Rook, K. S. (1984). The negative side of social interaction: Impact on psychological well-being. *Journal of Personality and Social Psychology, 46,* 1097–1108.

Rook, K. S. (1987a). Social support versus companionship: Effects on life stress, loneliness, and evaluations by others. *Journal of Personality and Social Psychology, 52,* 1132–1147.

Rook, K. S. (1987b). Reciprocity of social exchange and social satisfaction among older women. *Journal of Personality and Social Psychology, 52,* 145–154.

Rook, K. S. (1989). Strains in older adults' friendship. In R. G. Adams & R. Blieszner (Eds.), *Older adult friendship: Structure and process* (pp. 166–194). Newbury Park, CA: Sage.

Rook, K. S., & Pietromonaco, P. (1987). Close relationships: Ties that heal or ties that bind. In W. H. Jones & D. Perlman (Eds.), *Advances in personal relationships* (Vol. 1, pp. 1–35). Greenwich, CT: JAI Press.

Rose, J. H. (1993). Interactions between patients and providers: An exploratory study of age differences in emotional support. *Journal of Psychosocial Oncology, 11,* 43–67.

Rosow, I. (1975). *Socialization to old age*. Berkeley, CA: University of California Press.

Rosow, I. (1982). Intergenerational perspectives on aging. In G. Lesnoff-Caravaglia (Ed.), *Aging and the human condition* (pp. 40–53). New York: Human Sciences Press.

Rossi, A., & Rossi, P. (1990). *Of human bonding: Parent child-relations across the life course*. Hawthorne, NY: Aldine and Gruyter.

Rubin, S. S. (1991–92). Adult child loss and the two track model of bereavement. *Omega: Journal of Death and Dying, 24,* 183–202.

Ruscher, S. M., & Gotlib, I. H. (1988). Marital interaction patterns of couples with and without a depressed partner. *Behavior Therapy, 19,* 455–470.

Ryff, C. D. (1989). Beyond Ponce de Leon and life satisfaction: New directions in quest of successful aging. *International Journal of Behavioral Development, 12,* 35–55.

Schoenbach, V. J., Kaplan, B. H., Fredman, L., & Kleinbaum, D. G. (1986). Social ties and mortality in Evans County, Georgia. *American Journal of Epidemiology, 123,* 577–591.

Schultz, N. R. & Moore, D. (1988). Loneliness: Differences across three age levels. *Journal of Social and Personal Relationships, 5,* 275–284.

Schulz, R., & Rau, M. T. (1985). Social support through the life course. In S. Cohen & S. L. Syme (Eds.), *Social support and health* (pp. 129–149). New York: Academic Press.

Schulz, R., Tompkins, C. A., & Rau, M. T. (1988). A longitudinal study of the psychosocial impact of stroke on primary support persons. *Psychology and Aging, 3,* 131–141.

Schulz, R., Visintainer, P., Williamson, G. M. (1990). Psychiatric and physical morbidity effects of caregiving. *Journals of Gerontology, 45,* P181–P191.

Schuster, T. L., Kessler, R. C., & Aseltine, R. H. (1990). Supportive interactions, negative interactions, and depressed mood. *American Journal of Community Psychology, 18,* 423–438.

Schütze, Y., & Lang, F. R. (1995). Integration in family, kinship and friendship networks. In H. Mollenkopf (Ed.), *Elderly people in industrialized societies* (pp. 25 – 41). Berlin: Sigma.

Schwarzer, R., & Leppin, A. (1991). Social support and health: A theoretical and empirical overview. *Journal of Social and Personal Relationships, 8,* 99–127.

Schwarzer, R., & Leppin, A. (1992). Social support and mental health: A conceptual and empirical overview. In L. Montada, S.-H. Filipp & M. J. Lerner (Eds.), *Life crises and experiences of loss in adulthood* (pp. 435–458). Hillsdale, NJ: Erlbaum.

Scott, J. P. (1983). Siblings and other kin. In T. Brubaker (Ed.), *Family relationships in later life* (pp. 47–62). Newbury Park, CA: Sage.

Scott, J. P. (1990). Sibling interaction in later life. In T. H. Brubaker (Ed.), *Family relationships in later life* (2nd ed., Suppl., pp. 86–99). Beverly Hills, CA: Sage.

Seeman, T. E., & Berkman, L. F. (1988). Structural characteristics of social networks and their relationships with social support in the elderly: Who provides support? *Social Science and Medicine, 26,* 737–749.

Shanas, E. (1980). Older people and their families: The new pioneers. *Journal of Marriage and the Family, 42,* 9–15.

Shinn, M., Lehmann, S., & Wong, N. W. (1984). Social interaction and social support. *Journal of Social Issues, 40,* 55–76.

Shumaker, S., & Jackson, J. S. (1979). The aversive effects of nonreciprocated benefits. *Social Psychology Quarterly, 42,* 148–158.

Shute, R. & Howitt, D. (1990). Unraveling paradoxes in loneliness: Research and elements of a social theory of loneliness. *Social Behaviour, 5,* 169–184.

Shye, D., Mullooly, J. P., Freeborn, D. K., Pope, C. R. (1995). Gender differences in the relationship between social network support and mortality: A longitudinal study of an elderly cohort. *Social Science and Medicine, 41,* 935–947.

Smith, T. W. (1992). Hostility and health: Current status of a psychosomatic hypothesis. *Health Psychology, 11,* 139–150.

Stacey-Konnert, C., & Pynoos, J. (1992). Friendship and social networks in a continuing care retirement community. *Journal of Applied Gerontology, 11,* 298–313.

Staudinger, U. M., Marsiske, M., & Baltes, P. B. (1995). Resilience and reserve capacity in later adulthood: Potentials and limits of development across the life span. In D. Cicchetti & D. Cohen (Eds.), *Developmental psychopathology* (Vol. 2, pp. 801–847). New York: Wiley.

Steinbach, U. (1992). Social networks, institutionalization, and mortality among elderly people in the United States. *Journals of Gerontology: Social Sciences, 47,* S183–S190.

Steinmetz, S. K. (1988). Parental and filial relationships: Obligation, support and abuse. In S. K. Steinmetz (Ed.), *Family support systems across the life span* (pp. 165–182). New York: Plenum.

Sticker, E. J. (1991). The importance of grandparenthood during the life cycle in Germany. In P. K. Smith (Ed.), *The psychology of grandparenthood. An international perspective* (pp. 32–49). London: Routledge.

Stoller, E. P., & Pugliesi, K. L. (1988). Informal networks of community-based elderly: Changes in composition over time. *Research on Aging, 10,* 499–516.

Strain, L. A., & Chappell, N. L. (1982). Confidants. Do they make a difference in quality of life? *Research on Aging, 4,* 479–502.

Sugisawa, H., Liang, J., & Liu, X. (1994). Social networks, social support, and mortality among older people in Japan. *Journals of Gerontology: Social Sciences, 49,* S3–S13.

Sussman, M. B. (1985). The family life of old people. In R. H. Binstock & E. Shanas (Eds.), *Handbook of aging and the social sciences* (2nd ed., pp. 415–449). New York: Van Nostrand Reinhold.

Thomae, H. (1994). Trust, social support, and relying on others. *Zeitschrift für Gerontologie, 27,* 103–109.

Thompson, S. C., & Sobolew-Shubin, A. (1993a). Overprotective relationships: A nonsupportive side of social networks. *Basic and Applied Social Psychology, 14,* 363–383.

Thompson, S. C., & Sobolew-Shubin, A. (1993b). Perceptions of overprotection in ill adults. *Journal of Applied Social Psychology, 23,* 85–97.

Tinsley, B. J., & Parke, R. D. (1987). Grandparents as interactive and social support agents for families with young infants. *International Journal of Aging and Human Development, 25,* 259–277.

Tornstam, L. (1992). Loneliness in marriage. *Journal of Social and Personal Relationships, 9,* 197–217.

Townsend, A. L., & Franks, M. M. (1995). Binding ties: Closeness and conflict in adult children's caregiving relationships. *Psychology and Aging, 10,* 343–351.

Tucker, J. S., Friedman, H. S., Wingard, D. L., & Schwartz, J. E. (1996). Marital history at midlife as a predictor of longevity: Alternative explanations to the protective effect of marriage. *Health Psychology, 15,* 94–101.

Umberson, D., Wortman, C. B., & Kessler, R. C. (1992). Widowhood and depression: Explaining long-term gender differences in vulnerability. *Journal of Health & Social Behavior, 33,* 10–24.

Van Sonderen, E., Ormel, J., Brilman, E., & van Linden van den Heuvell, C. (1990). Personal network delineation: A comparison of the exchange, affective and role-relation approach. In K. C. P. M. Knipscheer & T. C. Antonucci (Eds.), *Social network research* (pp. 101–120). Amsterdam: Swets & Zeitlinger.

Walker, A. J., Martin, S. S., & Jones, L. L. (1992). The benefits and costs of caregiving and care receiving for daughters and mothers. *Journal of Gerontology: Social Sciences, 47,* S130–139.

Wallston, B. S., Alagna, S. W., DeVellis, B. M., DeVellis, R. F. (1983). Social support and physical health. *Health Psychology, 2,* 367–391.

Watson, D., & Clark, L. A. (1992). Affects separable and inseparable: On the hierarchical arrangement of the negative affects. *Journal of Personality and Social Psychology, 62,* 489–505.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales. *Journal of Personality and Social Psychology, 54,* 1063–1070.

Wellman, B. (1979). The community question: The intimate networks of East Yorkers. *American Journal of Sociology, 84,* 1201–1231.

Wentowski, G. J. (1981). Reciprocity and the coping strategies of older people: Cultural dimenions of network building. *The Gerontologist, 21,* 600–609.

White, L., & Edwards, J. N. (1990). Emptying the nest and parental well-being: An analysis of national panel data. *American Sociological Review, 55,* 235–242.

Wilcox, V. L., Kasl, S. V., & Berkman, L. F. (1994). Social support and physical disability in older people after hospitalization: A prospective study. *Health Psychology, 13,* 170–179.

Wilson, J. G., Calsyn, R. J., & Orlofsky, J. L. (1994). Impact of sibling relationships on social support and morale in the elderly. *Journal of Gerontological Social Work, 22,* 157–170.

Wortman, C. B., & Lehman, D. R. (1985). Reactions to victims of life crisis: Support attempts that fail. In I. G. Sarason & B. R. Sarason (Eds.), *Social support: Theory, research and application* (pp. 463–489). Dordrecht: Martinus Nijhoff.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.04
Sexuality

SUSANNE ZANK
Freie Universität Berlin, Germany

7.04.1 INTRODUCTION 73
  7.04.1.1 Reviews and Surveys 74
7.04.2 EPIDEMIOLOGY 75
  7.04.2.1 Sexual Behavior of Elderly Females 75
  7.04.2.2 Sexual Behavior of Elderly Males 76
  7.04.2.3 Critique of Epidemiological Studies 76
    7.04.2.3.1 Methodological problems 76
    7.04.2.3.2 Theoretical problems 77
7.04.3 AGE-RELATED CHANGES IN SEXUAL FUNCTIONING 78
  7.04.3.1 Physiological Changes 78
    7.04.3.1.1 Physiological changes in females 78
    7.04.3.1.2 Physiological changes in males 79
  7.04.3.2 Psychological Aspects 79
  7.04.3.3 Health-related Changes 81
    7.04.3.3.1 Diabetes mellitus 82
    7.04.3.3.2 Cardiovascular disease 82
    7.04.3.3.3 Cancer 82
    7.04.3.3.4 Renal failure 82
    7.04.3.3.5 Dementia 83
    7.04.3.3.6 Depression 83
    7.04.3.3.7 Use of medication 83
    7.04.3.3.8 Institutionalization 83
7.04.4 SEXUAL DISORDERS AND TREATMENT 83
  7.04.4.1 Sexual Desire Disorders 84
  7.04.4.2 Female Disorders 85
    7.04.4.2.1 Arousal disorder 85
    7.04.4.2.2 Orgasmic disorder 85
    7.04.4.2.3 Sexual pain disorders 86
  7.04.4.3 Male Disorders 86
    7.04.4.3.1 Arousal disorder 86
    7.04.4.3.2 Orgasmic disorders 87
  7.04.4.4 Special Issues in Treatment with Elderly Patients 87
7.04.5 CONCLUSION 88
7.04.6 REFERENCES 88

7.04.1 INTRODUCTION

Historical, societal, and demographic changes have led to the rise of gerontology. This comparatively new science already provides an impressive amount of empirical data, which contradict some common stereotypes of old age and aging. Results include, for example, developmental capacity even in the very old, tremendous interindividual differences, different developmental processes in distinct domains within an individual and the important role individual biography plays in the aging process (Baltes & Baltes, 1990; Birren & Schaie, 1985;

Kruse, 1989; Lehr & Thomae, 1987; Mayer & Baltes, 1996).

It is interesting to note, however, that there is a considerable lack of sound studies on sexuality and aging. The scientific community considers the study of sexual behavior a very difficult task which is fascinating but somewhat unseemly, especially with elderly subjects. This attitude might be caused to some extent by the unconscious resistance of scientists to exploring the sexual secrets of their parents. This problem was precisely described by the comedian Sam Levinson: "When I first found out how babies were born, I couldn't believe it. To think that my mother and father could do such a thing." Then, after a moment's reflection: "My father—maybe: but my mother—never!" (O'Donohue, 1987, p. 69).

This anecdote also reflects the conception of women as almost asexual beings, which might be one reason for the fact that gerontological knowledge about elderly sexuality refers predominantly to males. Since elderly females considerably outnumber males, due to the longer life expectancy of women, this focus on male sexual functioning is especially puzzling. It might mirror the societal "double standard of aging" (Sontag, 1972) which means that older women are perceived as being unattractive and asexual much earlier than men.

Many reviews comment on the myths of geriatric sexuality (O'Donohue, 1987; Spence, 1992). These myths suggest that elderly people are not sexually capable, not desirable, and not desirous. Such societal attitudes and stereotypes have various negative consequences for those who are affected. Mild sexual alterations might lead to panic and severe performance anxiety because they are misinterpreted as the beginning of the unavoidable end of sexuality. Many older people feel guilty or are ashamed because of their ongoing sexual interest and behavior, and this may result in total suppression of any sexual feelings (Renshaw, 1983).

Empirical data show that these attitudes are myths and do not reflect reality. Most elderly people are interested in sex and capable of sexual performance until very old age. Combating these myths, however, has led to the establishment of equally incorrect countermyths (Gatz, 1987). These countermyths almost deny any changes of sexual desire and performance with age. They establish an ideal which is a projection of middle-aged values and can be equally harmful (Thomas, 1982). Rappoport (1977) hints at this problem by saying "All we want from gerontology is eternal youth" (p. 207). Unfortunately, the unsatisfactory database on elderly sexuality provides support for truth in the myth as in the countermyth.

7.04.1.1 Reviews and Surveys

There are numerous review articles on sexuality and aging which seem to outnumber original research reports (e.g., Byers, 1983; Capuzzi, 1982; Corby & Zarit, 1983; Datan & Rodeheaver, 1983; Garza & Dressel, 1983; Genavay, 1982; Hermanova, 1983; Hodson & Skeen, 1994; Kay & Neelley, 1982; Kellett, 1993; Kirkpatrick, 1980, Laws, 1980; Libman, 1989; Ludeman, 1981; Luria & Meade, 1984; O'Donohue, 1987; Renshaw, 1983; Robinson, 1983; Starr, 1985; Sviland, 1981; Traupmann, 1984; Turner & Adams, 1983; Von Sydow, 1992a; Weg, 1983; Weiler, 1981; White, 1982). These authors usually review results of the most important studies, which were conducted in the USA and will be described in the following.

Alfred C. Kinsey and his colleagues interviewed 17 500 men and women (Kinsey, Pomeroy & Martin, 1948; Institute for Sex Research, 1953). Their sample was not representative but this was the first study to investigate sexual behaviors from childhood to old age, and it therefore represented a remarkable first attempt to understand human sexual behaviors. Cross-sectional data were collected with structured and standardized interviews. Unfortunately, only 126 males and 56 females over 60 were included in the study, and only 10 females were over 70. More recently, a new report from the Kinsey institute was published which described data gathered in 1970 from extensive interviews with more than 3000 men and women (Klassen, Williams, & Levitt, 1989). This study was restricted to sexual attitudes and did not report data on sexual behaviors.

Other important sources of information are the two Duke University Longitudinal Studies. These results are also rather old, but still very valuable, because they are the only ones which stem from longitudinal data (George & Weiler, 1981; Pfeiffer, Verwoerdt, & Wang 1968; Verwoerdt, Pfeiffer, & Wang, 1969a, 1969b). In the first study, data on sexuality were collected as part of a medical interview, and in the second study a self-administered questionnaire was used. Neither sample was representative (Weiler, 1981). Over the three measures of the first Duke Longitudinal Study the number of subjects aged over 60 declined from 254 to 126 (Verwoerdt et al., 1969a, 1969b). In the second Duke Longitudinal Study a sample of 502 men and women aged 46–71 years was drawn; 348 persons completed all four waves of data collection, but data were only reported about those 278 subjects who remained married throughout the study. Forty-eight subjects were older than 66 years at the beginning of the six-year inquiry (George & Weiler, 1981).

Most of the above mentioned reviews could not discuss the results of the more recent studies (Brecher, 1984; Janus & Janus, 1993; Laumann, Gagnon, Michael, & Michels, 1994; Starr & Weiner, 1981). Starr and Weiner (1981) investigated 800 subjects over 60 in senior centers who attended lectures on sexuality and aging. Of the 4246 subjects who participated in the Consumer Union survey, 2622 subjects were over age 60. Subjects were a self-selected sample of readers of the Consumer Report (Brecher, 1984). Janus and Janus (1993) investigated 2765 subjects with questionnaires "distributed to subjects at a wide variety of sites" (p. 3) in a cross-sectional study between 1988 and 1992; 441 of the subjects were older than 65. The samples of these three studies were also not representative. In a very carefully designed study with a representative sample Laumann et al. (1994) collected 3432 questionnaires about sexual practices in the USA. Unfortunately, they only included subjects aged between 18 and 59.

European studies are not very often reported in the US literature because they are rarely translated. Consequently, the considerable differences between results of US and European studies are not discussed (Von Sydow, 1992a). The following non-American studies are published in English. Persson (1980) selected a representative sample of 70-year-old people in Gothenburg, Sweden. Subjects were invited to take part in a health control, 85% accepted, and 166 men and 266 women were interviewed. Bergström-Walan and Nielsen (1990) drew a random sample of citizens aged 60–80 of Stockholm, Sweden. They mailed 1574 questionnaires, and 509 were returned. Antonovsky, Sadowsky, and Maoz (1990) interviewed 177 men and 121 women from a nonrepresentative sample in Israel. Data included current sexual activities and retrospective reports about past sexual experience.

Taken together, epidemiological knowledge on sexuality and aging is grounded on rather few studies. The database for men is somewhat better, because there are some studies on male sexual behavior only (e.g., Mulligan & Moss, 1991; Schiavi, 1990; Weizmann & Hart, 1987). White (1982) cited 10 studies which investigated exclusively males, while only three studies exclusively focused on females. The following section will briefly review the major epidemiological results on sexual activities. Data are extremely heterogeneous and therefore very difficult to interpret. Some scientific objections against uncautious generalization of available data will be raised.

7.04.2 EPIDEMIOLOGY

7.04.2.1 Sexual Behavior of Elderly Females

Studies in the USA have shown that 70–90% of females aged between 60 and 90 were sexually active (Adams & Turner, 1985; Brecher, 1984; Starr & Weiner, 1981). Sexual activity in these studies referred to intercourse and masturbation. German studies reported, however, that 45%–55% of females aged 50 to 60 were sexually active, and up to 3% of women over 70 years (Tümmers, 1976; Weissbach-Rieger & Franke 1987a; 1987b; Wickert, 1983).

Data on activity in sexual intercourse varied between 15 and 90% for women aged 50 to 59, and between 0 and 66% for women older than 70 years. Frequency of intercourse was reported between once a week and once per month, with a decline with age (Bergström-Walan & Nielsen, 1990; Brecher, 1984; Christenson & Johnson, 1973; George & Weiler, 1981; Janus & Janus, 1993; Institute for Sex Research, 1953; Pfeiffer, Verwoerdt, & Davis, 1972; Starr & Weiner, 1981). Bretschneider and McCoy (1988) investigated very old subjects between 80 and 102 with an average of 86.1 years; of the interviewed women, 70% no longer had intercourse.

Of women over 50 years, 10–50% masturbated, and between 0 and 49% females over 70 years reported this activity. Masturbation was reported much more frequently in US than in European studies (Adams & Turner, 1985; Aresin, 1980; Bergström-Walan & Nielsen, 1990; Brecher, 1984; Bretschneider & McCoy, 1988; Institute for Sex Research, 1953; Schneider, 1980; Starr & Weiner, 1981).

Questions about sexual enjoyment or achieving orgasm were rare. Of women in all age groups, 9–26% reported that they never achieved orgasm (Hite, 1976; Institute for Sex Research, 1953; Janus & Janus, 1993; Laumann et al., 1994). Studies about elderly females revealed that about 70% of the American subjects reported sexual enjoyment but only 20% of subjects aged over 80 (Brecher, 1984; Starr & Weiner, 1981). In an Israeli sample, 43% of women over 65 reported that they were not very satisfied. Of these subjects, 14% recalled that they were not very satisfied between 20 and 30 years, and 28% were not very satisfied in their fifties (Antonovsky et al., 1990). The rate of women who never or rarely reacted with orgasm remained stable at roughly 21% between 40 and 59 years (Laumann et al., 1994) but seemed to increase over 70 years (Bergström-Walan & Nielsen, 1990; Brecher, 1984).

Between 4 and 19% of the subjects reported that they had homosexual contacts at least once in their life and 1–5% stated an ongoing relationship (Adams & Turner, 1985; Brecher, 1984, Hagstad & Janson, 1984). Data from a representative study of subjects aged between 18 and 59 years showed that 4% had same-gender sex partners, 1.3% during the last year (Laumann et al., 1994).

7.04.2.2 Sexual Behavior of Elderly Males

Between 60% and 90% of men aged over 60 and between 48% and 79% over 70 report being sexual active. Over 80% of men in their sixties and 75% in their seventies reported a strong or moderate interest in sex (Brecher, 1984; Pfeiffer, Verwoerdt & Wang, 1968; Verwoerdt et al., 1969a, 1969b).

Between 64% and 89% of men aged over 60 reported having sexual intercourse, with 24%–69% reporting weekly or higher frequency sexual intercourse (Antonovsky et al., 1990; Bergström-Walan & Nielsen, 1990; Brecher, 1984; Janus & Janus, 1993; Weizmann & Hart, 1987). Between 46 and 62% of men over 70 years reported having intercourse (Brecher, 1984; Bretschneider & McCoy, 1988; Persson, 1980).

Between 45 and 59% of elderly males reported that they never masturbate, and 28%–44% reported masturbating at least weekly (Bergström-Walan, 1990; Brecher, 1984; Janus & Janus, 1993; Weizmann & Hart, 1987).

There seems to be a decline in orgasm with age in men. Janus and Janus (1993) reported that 77% of men between age 51 and 64 always experienced orgasm, but only 55% over age 65. Mulligan and Moss (1991) claimed that ability to experience orgasm declined from 61% in the 60–69 year olds to 21% in men aged over 90. Brecher (1984) reported that 85% of men in their sixties and 70% in their seventies always reached orgasm. In comparison, an average of 8% of men between 18 and 59 years were unable to experience orgasm (Laumann et al., 1994).

7.04.2.3 Critique of Epidemiological Studies

7.04.2.3.1 Methodological problems

(i) Sample

There are hardly any representative studies and samples are often very small. Taken together, our knowledge on sexuality and aging is based on roughly 7000 subjects. Discrepancies between European and US data cannot be interpreted because of small sample sizes. Furthermore, the age of subjects ranges between 60 and 100 years, spanning 40 years. This means that sample sizes (e.g., for 60 and 80 years olds) become extremely small for different age groups. From a gerontological perspective, 60-year-old subjects differ markedly from 80-year-olds.

Relying on epidemiological data which stem from nonrepresentative samples can be very misleading. Greeley (1994) compared results from Janus and Janus (1993) with data from the General Social Survey (GSS), which was based on a national household probability sample. The comparison was made about percentages of people having sex at least once a week. Janus and Janus (1993) reported that 81% of men aged between 51 and 64 and 69% of men over 65 had sex at least once a week. The GSS data claimed only 43% of men between age 51 and 64, and 17% over 65 had sex at least once a week. For women, Janus and Janus (1993) reported that 65% aged between 51 and 64 and 74% over 65 had sex at least once a week. GSS results showed that 25% of women between 51 and 64, and 6% over 65 had sex that often (Greeley, 1994, cited in Laumann et al., 1994).

(ii) Research design

Most of the studies used a cross-sectional design. Longitudinal designs are essential, however, to describe intra-individual changes. Cross-sectional data are very often interpreted as if they are longitudinal (Institute for Sex Research, 1953; Kinsey et al., 1948). George and Weiler (1981) reported from their longitudinal study that sexual behavior was remarkably stable over the six years of their study. Their data suggested that differences between age-groups reflected cohort-related rather than age-related effects. Liberation of sexual moral attitudes or changes in age-related stereotypes may even enhance cohort differences in the future.

(iii) Data collection

Most studies used questionnaires or standardized interviews. These methods assume that the meaning of questions will be the same for respondents and scientists. Von Sydow (1992b) conducted an interview study with 91 females and found that seven did not know the term orgasm. Five other women of this sample claimed that a woman needs to have an orgasm in order to become pregnant. Finally, the important Duke longitudinal study (George & Weiler, 1981; Pfeiffer et al. 1968; Verwoerdt et al., 1969a, 1969b) collected data on general medical problems. The medical context for exploring sexuality may not support open answers about sex, especially with many subjects suffering from chronic illness. Furthermore, structured interviews in the first study were conducted by male interviewers which may have caused some reluctance to self-disclose among older women in the 1950s (Pfeiffer & Davis, 1972; Starr, 1985; Verwoerdt et al., 1969a, 1969b).

7.04.2.3.2 Theoretical problems

Every study on human sexuality shows that men are sexually more active than women. This seems to be true for every age group, for intercourse and, especially, for masturbation (Brecher, 1984; Institute for Sex Research, 1953; Janus & Janus, 1993; Kinsey et al., 1948; Klassen et al., 1989; Laumann et al., 1994). Although physiological results indicate higher potency of female response in terms of orgasmic potential in comparison with men (Masters & Johnson, 1966, 1981), women seem not to use this potential. This appears to be particularly true for elderly females and raises the question whether research efforts have met the sexual, social, and psychological reality of older women.

(i) Intercourse

Older widowed women outnumber widowed men approximately 5:1, which means that in basically monogamous societies most women do not have many opportunities for intercourse (Malatesta, Chambless, Pollack, & Cantor, 1988). Therefore, higher frequencies of having intercourse in males do not prove age-related differences of sexual interests between men and women. Pfeiffer and Davis (1972) found that past sexual enjoyment, marital status, and age were significant determinants of female sexual behavior. Antonovsky et al. (1990) reported that marital status alone explained 25% of the variance of sexual activities in women. A representative study of 460 subjects aged over 60 in Germany showed that gender differences diminished when "availability of a partner" was statistically controlled (Brähler & Unger, 1994). Interestingly, investigations of couples revealed that both men and women agreed that they ceased intercourse because of the male (Persson, 1980; Pfeiffer et al., 1968; Von Sydow, 1993).

Differences in intercourse rates for younger age-groups are surprising because theoretically women and men should report the same amount of intercourse altogether. An explanation could be exaggeration of males or understatement of females. Another explanation could be visits to prostitutes by men.

(ii) Orgasm

Examining intercourse ignores information about individual feelings and quality of experience. Men and women can have intercourse without enjoyment or orgasm, but this happens far more often to women than to men (Laumann et al., 1994; Leiblum & Rosen, 1989b; McConaghy, 1993). The fact that important studies did not explore sexual enjoyment or orgasm but focus on intercourse (e.g., George & Weiler, 1981; Pfeiffer, Verwoerdt, & Wang 1969; Verwoerdt et al., 1969a, 1969b) reveals that research has been biased by a male point of view. The occurrence of orgasm would be a clear indicator of satisfying sexual experience and would provide solid ground for the prediction of future sexual interests (Laws, 1980). Fortunately, more recent studies have considered this aspect.

(iii) Interest in sexual activities

Many (male) scientists have investigated interest in intercourse and interpreted their results as representing interest in all sexual activities. Presumably, this has led to invalid results. Studies have shown that, in contrast to men, many women prefer other sexual activities to intercourse, like manual or oral stimulation (Denney, Field, & Quadagno, 1984; Von Sydow, 1993). The clitoris and penis are homologous organs which need stimulation for orgasm. During intercourse the penis is directly, and the clitoris indirectly stimulated. Most women need additional clitoral stimulation to achieve orgasm (Masters & Johnson, 1966; McConaghy, 1993), which explains their preferences. If true gender differences with regard to sexual interests exist they are probably due to different experiences of pleasure and orgasm during the life course (Laws, 1980).

(iv) Sexual experience and relationship

Research efforts have focused heavily on sexual performance and ignored the phenomenology of sexual experiences, namely, the subjective meaning, and relational quality of the experience. Starr (1985) reported that 75% of his subjects aged over 60 indicated that their sexual feelings were as good as or even better than in younger years. They stressed the quality of their sexual experiences instead of quantity. This seems especially true for women for whom the emotional aspect of a sexual relationship are more important than for many men (Luria & Meade, 1984).

Taken together, our knowledge about human sexual behavior and experience in the elderly is sparse. Most of the studies reveal considerable methodological and conceptual weaknesses. Nevertheless, results are often cited and generalized, supporting myths (especially for elderly women) and countermyths on aging and sexuality. Many studies have no theoretical framework. A developmental perspective of aging individuals and couples is clearly lacking in most studies. The following section will focus on physiological changes in healthy elderly people and on psychological aspects of aging and sexuality. Another focus will be on the effects of chronic illnesses on sexual functioning, because the prevalence of chronic illness increases with advancing age.

7.04.3 AGE-RELATED CHANGES IN SEXUAL FUNCTIONING

Most of the epidemiological data suggest that there is a decrease in sexual interest and sexual behavior as one ages. It is often argued that this decline is due to biological changes in sexual response capacity. Age-related biological changes in capacity are difficult to determine because changes in sexual functioning are confounded by the effects of chronic diseases, medication, and cultural expectations. Moreover, Winn and Newton (1982) have documented examples of cultural influences on sexual response capacity. They studied data of the Human Relations Area Files of about 106 preindustrial and traditional societies all over the world. For 28 cultures, data on sexual activities in males were available. In 70% of these societies, older men had little or no reduction in sexual behavior until the age of 75. Twenty-six cultures provided data about female sexuality, and 84% of these data reported continued sexual interest and activity in old females which often included having intercourse with younger men. This study gives some evidence that biological changes do not necessarily lead to a reduction in sexual interest and sexual behavior. Facts about physiological changes will be briefly reviewed in the following. Detailed information is provided by McConaghy (1993), Morley (1991), Sarrel (1990), Schiavi (1990), and Schiavi and Segraves (1995).

7.04.3.1 Physiological Changes

7.04.3.1.1 Physiological changes in females

Menopause is associated with marked hormonal changes, namely, a decrease in estrogen production by the ovaries. There is an ongoing controversy about how the hormonal changes affect female sexuality (for a methodological review and meta-analysis see Myers, 1995). It was thought that decline in sexual activity was related directly to decline in sexual hormones, but estrogen and progesterone levels did not predict women's sexual functioning in the laboratory (Roughan, Kaiser, & Morley, 1993). Estrogen deficiency is responsible, however, for decreased elasticity of the vaginal walls, atrophy of vaginal epithelium, diminished vaginal lubrication, and increased fragility of the vaginal skin (Goldstein & Teng, 1991; Leiblum, 1990; Masters & Johnson, 1966, 1981; Morley, 1991; Morley & Kaiser, 1989; Roughan et al., 1993; Schiavi & Segraves, 1995; Wise, 1989). These changes are probably responsible for coital pain, which is more often reported by postmenopausal women. Estrogen replacement therapy is very successful in treating these vaginal changes.

Although Masters and Johnson (1966, 1981) studied only 11 women aged over 60, their description of physiological changes in healthy women proved to be accurate and their data have been supported by findings from self-report surveys (Bachman & Leiblum, 1991; Dennerstein, Smith, Morse, & Burger, 1994; Hagstad, 1988; McCoy & Davidson, 1985; Morley, 1991). Physiological changes in female sexuality include decreased congestion of the labia minora, decreased separation of the labia majora, decreased vaginal lubrication, decreased orgasmic platform contraction, and occasional painful spastic contractions of the uterus. Furthermore, the breasts become less vasocongested and nipple erection is less likely to occur. Masters and Johnson (1981) reported that aging women require more time in precoital stimulation until lubrication develops sufficiently. Orgasmic duration and intensity may be reduced but the capacity for multiple orgasms remains.

There is some ongoing controversy as to whether women suffer sexual dysfunction following hysterectomy. Bernhard (1986) analyzed 18 studies and detected numerous methodological and measurement problems. Eicher (1994a, 1994b) reviewed research outcomes and concluded that older studies reported negative, and newer studies positive, consequences for female sexuality. He argued that this change is due to better information, sexual behavior advice, and historical changes in attitudes towards sexuality. Results from a recent longitudinal study with 104 women revealed that half of the women reported improvement after surgery and 21% reported poorer sexual functioning. Preoperative sexual activity was the best predictor of postoperative sexuality (Helström, Lundberg, Sörbom, & Bäckström, 1993).

Dennerstein et al. (1994) reported a study with 2001 randomly selected Australian women aged between 45 and 55 years. The natural menopause transition was clearly associated with declining sexual interest, decreased likelihood of intercourse, and increased likelihood of unusual pain during intercourse. Hällström and Samuelsson (1990) reported data from 497 subjects with a male partner who were interviewed twice within seven years. At first interviews, subjects were aged 38, 46, 50, or 54. Of the total sample, 27% reported a decrease in sexual desire and 10% experienced an increase in desire. Age-group differences were marginal. Antonovsky et al. (1990) conducted retrospective interviews with 121 women and 177 men between the ages of 65 and 85. The women reported a marked decrease in sexual desire and sexual satisfaction in comparison to their twenties, thirties, and fifties. McCoy and Davidson (1985) conducted a longitudinal study and reported a significant decrease in coital frequency and sexual thoughts in the postmenopausal compared to the premenopausal period. The decline in coital frequency was related to testosterone rather than to estrogen levels. Sarrel (1982) described sex problems after menopause in 50 couples (e.g., loss of orgasmic response in female partner, erectile difficulties in male partner) who participated in a sex counseling program. Dyspareunia in women was often followed by erectile difficulties in men.

In contradiction to the reviewed data, which support a decline hypothesis, Starr (1985) claimed that "the Starr–Weiner data indicate that women over 60 are more orgasmic than when they were younger" and that "most investigators report that sexual response remains the same or improves after menopause" (Starr, 1985, p. 102). Epidemiological data from a representative cross-sectional study clearly suggested that the number of women who never experience orgasm remained fairly stable until the age of 70 (Laumann et al., 1994). These data do not answer the question, however, whether orgasmic women increase their capacity to orgasm after menopause.

7.04.3.1.2 Physiological changes in males

Men do not undergo a clearly marked, obvious change such as females experience with menopause. Nevertheless, there is a remarkable change in hormone production as well; namely, a decrease in testosterone and bioavailable testosterone in the majority of males (Morley & Kaiser, 1989). Decrease in testosterone production has led to the hypothesis that androgen deficiency contributes to the decline in sexual desire and activity in older men (Schiavi, 1990). Remarkably few studies, however, assessed the role of gonadal hormones on age-related changes.

Schiavi and colleagues investigated 77 healthy, married men aged 45–74 and their partners. The sample was drawn via media announcements. Couples underwent an extensive interview together as well as separately. Data gathering included psychosexual, psychiatric, and marital information as well as a battery of psychological tests. The male subjects were additionally studied in a sleep laboratory for four nights. Electroencephalogram, eye movements, muscle tone, and penile tumescence were monitored. Data revealed that age was negatively correlated with bioavailable testosterone. Bioavailable testosterone showed significant associations with sexual desire, arousal, and activity measures. Furthermore, results showed age-related changes in nocturnal penile tumescence. Frequency, duration, and rigidity of this reaction decreased with increasing age. Finally, age was significantly negatively correlated with sexual desire, sexual arousal, and sexual activity, but there were no age differences in degree of enjoyment of marital sexuality. Increased age was also correlated with erectile difficulties (erectile rigidity, loss of erection during sex) and retarded ejaculation (Schiavi, 1990; Schiavi, Mandeli, & Schreiner-Engel, 1994; Schiavi, Mandeli, Schreiner-Engel, & Chambers, 1991; Schiavi, Schreiner-Engel, White, & Mandeli, 1991).

Masters and Johnson (1981) described four variations in basic sexual physiology in males who were past their mid-fifties. First, it takes longer to achieve full penile engorgement and requires more direct stimulation. Second, there is a reduction of expulsive pressure. Third, there is a reduction of the volume of seminal fluid expelled. Fourth, there is a reduction or loss of ejaculatory demand. Weiss and Mellinger (1990) stated that there are only one or two expulsive contractions of the urethra instead of four major contractions in younger men. Seminal fluid is expelled with less force, which diminishes the sensation of passage through the urethra. There is, rather, a seepage of semen which reduces the intense sensation of ejaculatory inevitability. After ejaculation, the refractory period increases with age. Many men cannot have another erection for 12–24 hours. Further physiological changes include less nipple erection; less, or absence of, flushing of neck, trunk, shoulders, thighs, and forearms; scrotal integument becomes thinner and inelastic; rapid detumescence after orgasm (Weiss & Mellinger, 1990).

7.04.3.2 Psychological Aspects

Modern developmental psychology emphasizes that development is not limited to childhood or youth but continues over the life course. This life-span perspective views development as a product of an interplay between biological, cultural, and psychological influences, that is, individuals interact with their environment continuously over the life course (Baltes, 1987; Baltes, Reese, & Lipsitt, 1980).

Cultural influences change considerably with time. This fact can be illustrated by a closer look at elderly subjects who were investigated, at age 70 in 1980. These subjects were born in 1910. Their parents grew up in the Victorian century. This means that these subjects were exposed to very strict societal norms regarding sexuality during their childhood and adolescence. Masturbation was forbidden and considered to be extremely harmful. Sexual feelings for decent women were unthinkable. Freud's theories and discoveries, especially his description of sexual feelings in children, were regarded as tremendously scandalous in Vienna at the beginning of the century (Freud, 1964). Despite short periods with less restricted sexual norms (e.g., in Berlin during the 1920s), sexuality remained a tabooed, guilt- and shame-laden topic for many years. Considerable changes began with the pioneering work of Kinsey and his colleagues (Institute for Sex Research, 1953; Kinsey et al., 1948) who provided the first published survey of sexual behavior. Masters and Johnson (1966) followed with serious information about "human sexual response." Their work was a breakthrough for the acknowledgment of female sexuality. At the same time the development of the contraceptive pill, the ideas of the hippie movement, and later feminism, contributed to the liberation of sexuality. During the 1980s, the threat of AIDS reduced sexual freedom. The above-mentioned subjects who were 70 years old in 1980 experienced all these influences, but the development of their sexual identity was strongly influenced by Victorian moral standards. The fading of sexuality over ontogenetic time might be, therefore, a product of socialization rather than of biology (Traupmann, 1984).

In addition, cultural influences are very important for age-graded behavior expectations and societal values in general. Modern Western societies are dominated by youth orientation, and sexuality is linked to it. Firm and slim bodies, smooth skin, health, energy, career orientation, money, and luxury are highly valued. All these characteristics are directly or indirectly related to youth, sexual attractiveness, and potency. Elderly people have the difficult task of coping with these societal norms which threaten their self-esteem.

Elderly homosexuals suffer from multiple myths and stereotypes. These stereotypes include that aging gays are oversexed, frustrated, unable to have a lasting relationship, and have an extreme youth orientation (Pope & Schulz, 1990). McDougall (1993) described these myths, and stressed cohort differences between younger and older gays. Elderly gays are used to hiding their sexual preferences, in contrast to younger ones who have benefited more from the gay liberation movement in the 1970s. Cruikshank (1990) reviewed gerontological studies about aging lesbians and gays, and concluded that there were more positive outcomes than expected. Friend (1990) presented a "theory of successful aging" (p. 99) for lesbian and gay individuals. He theorized that the tough lifelong reality with numerous experiences of discrimination leads to the development of a strong ego in gays. They are thus better prepared to cope with the demands and humiliating experiences of aging than the heterosexual elderly.

For women, cultural pressure is particularly damaging because esthetic and erotic feelings in Western societies lead to old women being considered as repulsive (Sontag, 1972). Whereas wrinkles and gray hair are an interesting and attractive sign of experience in men, they are a symbol of decaying beauty and femininity in women. A marriage of an older man and a young woman is a sign of vigor and vitality in the man, but when an older woman marries a young man, it is generally disapproved or even judged as pathological. An illustration of this humiliating experience for women is provided by Thomas Mann. Being himself in his late seventies, he wrote a story about a 50-year-old woman who fell in love with a young man. With his deep psychological understanding of human conflicts he described desire, restriction, and a "betrayal by nature" of this woman (Mann, 1953).

Men also have to face cultural pressures. Male identity and masculinity is directly linked to sexual performance demands. Books, stories, movies, and television shows project a "fantasy model of sex: It's two feet long, hard as steel, and can go all night" (Zilbergeld, 1992, p. 30). The penis is expected to function automatically whenever its owner wishes to use it, regardless of any other conditions. Older men who feel a normal need for more direct stimulation or experience erectile weaknesses can be threatened to the core of their identity. As a 51-year-old man put it: "You want to know how I feel, I'll tell you. I feel like an absolute nothing. I know I can satisfy her in other ways and I do, but that's not the point. I feel like shit, like the center has been taken out of me" (Zilbergeld, 1992, p. 36). Unfortunately, because of this culturally shared fantasy model of male sexuality, many women interpret occasional lack of erection as if they were rejected, not desired, not loved, or not attractive enough. Women have difficulties understanding the intense threat that erectile problems pose for men because they have no parallel experience. The female gender identity will not be questioned by any sexual dysfunction because it is not linked to sexual performance. In addition, women can almost always perform sexually: they can have intercourse without being aroused and can fake enjoyment, but there is no way to fake an erection.

These cultural influences, shared and gender-specific ones, are of considerable importance for psychosocial development and well-being in the elderly (Rosenmayr, 1995). Aging individuals as well as aging couples have to cope with various developmental tasks. Robert Peck (1956) described three phases of development in late adulthood: ego differentiation versus work-role preoccupation, body transcendence versus body preoccupation, and ego transcendence versus ego preoccupation. The first phase refers to coping with retirement. The crucial question is whether a person can derive self-esteem from activities other than work. The second phase means coping with physical decline and health problems. The last stage is very similar to Erikson's (1950) description of ego integrity vs. ego despair as the outcomes of the final developmental task in old age. Acceptance of one's unique life with unchangeable failures and the realization of death are the most important challenges of this stage. These difficult tasks represent a threat to normal narcissism and can have pathological consequences (Eckert, 1984; Liptzin, 1984). Kernberg (1977) described the difference as follows:

> It is part of normal narcissism to love oneself, to be satisfied with one's own appearance, and to wish to be accepted, liked, and loved for what one is. The problem begins when a highly ideal, beautiful and triumphant image of the self linked to childhood or youth becomes an indispensable precondition for self-acceptance and for trusting that one will be accepted by others: this usually reflects a deep devaluation or depreciation of one's self, a sense of loathing and distrust for one's present appearance and value, which can only be compensated for by the eternal looks of triumphant adolescence. (pp. 27–28)

Coping with the narcissistic wounds of aging is a major task for every individual. Fortunately, empirical data indicate that most elderly people manage well and retain positive self-esteem and life satisfaction (Bengtson, Reedy, & Gordon, 1985; Costa et al., 1987; Staudinger, Freund, Linden, & Maas, 1996).

Intimacy and intimate relationships seem to be especially important for coping with the demands of aging (Weiss, 1983, see also Chapter 3, this volume). Intrapsychic changes and role transitions additionally challenge the aging couple. Two examples are marital life after children have left the household, and the retirement of one or both partners. As one woman commented on her stress with the retirement of her husband: "I married my husband for better or for worse, but not for lunch" (Leiblum & Segraves, 1989, p. 377).

Life-span development of gender identity and the concept of androgynization with age has important implications for the individual as well as for the interindividual dynamic of couples and their sexuality (Livson, 1983). Jung (1933) reasoned that role demands in early adulthood require development of selective domains and that older people show greater awareness of aspects of their self which were suppressed earlier. "The man discovers his tender feelings and the woman her sharpness of mind" (p.108).

Colarusso and Nemiroff (1981) theorized that oedipal issues might reawaken because of illness and death of parents. Therefore, internal representations of parents may change and alter the adult's self-concept. This change in personality may also effect sexual feelings insofar that long-standing relationships are higher valued than short-term sexual gratification (Nadelson, 1984).

Finally, aging couples have to face possible dependency of one partner due to chronic illness, and separation through death. Coping with increasing losses of significant others is one of the most difficult tasks for elderly people. Death of the partner means for many traditionally oriented women, besides many other losses, the end of any sexual activity (Malatesta et al., 1988).

7.04.3.3 Health-related Changes

One major problem in examining age-related changes is that the distinction between age-related and health-related changes is almost impossible. Chronic illnesses increase markedly with age and most diseases have a direct or indirect effect on sexual functioning. Thus, it is not aging per se, but comorbid illness that may cause an increase in sexual dysfunction (Schover, 1989, 1992; Spence, 1992). Some authors claim, however, that health aspects tend to be overestimated with regard to sexuality (Fooken, 1994). Common chronic diseases will be briefly reviewed in this section. An excellent overview about geriatric sexuality and chronic diseases is provided by Mooradian (1991). Additional information is given by Burgener and Logan (1989), Gupta (1990), Kaiser (1991), Liptzin (1984), Majerovitz and

Revenson (1994), and Sjögren, Damber, and Liliequist (1983).

7.04.3.3.1 Diabetes mellitus

The prevalence of diabetes for Caucasians aged over 65 in the USA is 7% for men and 9% for women. For the black population aged over 65, it is 12% for men and 15% for women (Zemel, 1988). Diabetes is an important cause of sexual dysfunction. Orgasmic impairments are experienced by 18%–35% of women. Impotence, retrograde ejaculation, premature ejaculation, and oligospermia are reported for men. At 65 years of age, up to 75% of diabetic men are impotent (Mooradian, 1991; Newman & Bertelson, 1986; Zemel, 1988).

7.04.3.3.2 Cardiovascular disease

Prevalence of this disease increases with age, and the disease causes over 50% of deaths in individuals aged over 64 (Wise, 1989). Both organic and psychological factors affect the sexuality of individuals suffering from cardiovascular disease. Orgasmic difficulties in women as well as erectile dysfunction and premature ejaculation in men are reported (Wise, Epstein, & Ross, 1992). Following coronary bypass surgery, the sexual desire of women appears to decline, while an increase in performance anxiety is common in men. Survivors of myocardial infarction often report anxiety and depression. Fear of sudden death during intercourse often leads to impotence. Patients are assured that they can have intercourse without any risk when they can climb two flights of stairs or can vigorously walk around a block without chest discomfort (Mooradian, 1991; Relf, 1991; Wise, 1989). This rule of thumb ignores interindividual differences in patients and, additionally, poses a psychological dilemma. Patients are allowed to resume sex but they have to watch cautiously for any signs of recurring angina. This anxious self-observation contradicts the "letting go" which is part of successful sexual experiences (Thienhaus, 1988).

7.04.3.3.3 Cancer

This disease continues to be the most dreaded diagnosis, which often leads to depression and anxiety. For most patients their initial question concerns survival. Cancer and its treatment has an intense influence on individual identity and self-esteem, and it can have negative effects on desire, arousability, and orgasm in females and males (Gendel, 1986; Kaiser, 1992; Schover, 1988a, 1988b; Schover, Evans, & Von Eschenbach, 1987; Schover, Schain, & Montague, 1989; Shell & Smith, 1994).

In women, breast and gynecological cancer and its treatment can have tremendously deleterious effects on self-image and sexual functioning. One out of nine American and European women experience breast cancer and the incidence is increasing (Kaplan, 1992). Schover et al. (1995) compared psychosocial adjustment, body image, and sexual function for 72 women who had experienced partial mastectomy and 146 women who had undergone immediate breast reconstruction after partial mastectomy. Contrary to expectations, the two groups did not differ in psychosocial adjustment to the illness, with fewer than 20% of subjects reporting poor adjustment. Women who had been treated with chemotherapy had more sexual dysfunction, poorer body image, and more psychological distress. Margolis, Goodman, and Rubin (1990) compared 22 patients who had experienced radical mastectomy with 32 women who had undergone lumpectomy combined with radiation therapy. Patients who had experienced mastectomy felt less attractive, less desirable, and experienced less sexual enjoyment than prior to treatment. Patients who had undergone radiation therapy reported no changes in these domains. Kaplan (1992) stated that all current treatments for breast cancer can have negative effects on female sexuality but that these side-effects are widely ignored by clinicians and scientists.

In men, the effects of prostate cancer on sexual functioning vary with treatment. Radical prostatectomy results in impotence in up to 85% of cases, but only 5% of patients reported impotence after transurethral resection; 80% of these patients stated, however, retrograde ejaculation which can reduce enjoyment (Mooradian, 1991; Wise et al., 1992). Penile cancer is a rare malignancy. Opjordsmoen, Waehre, Aass, and Fossa (1994) reported Norwegian data on 30 men. Their results indicate that radiotherapy preserves sexuality better than surgery or laser treatment. Partial or total penectomy severely reduces sexual interest and ability of performance.

Colon cancer often reduces sexual functioning because of psychological problems. Surgery may lead to an ostomy, and fear of stool leakage and odor may inhibit sexual responses (Kinn & Öhmann, 1986; Mooradian, 1991; Wise et al., 1992).

7.04.3.3.4 Renal failure

Many patients with decreased renal functioning report sexual difficulties. This disease is associated with various hormonal changes which may influence sexual potency (Mooradian, 1991; Wise et al., 1992).

7.04.3.3.5 Dementia

The prevalence of dementia is about 5% in individuals aged over 65, and 20% in persons over 85 (Helmchen et al., 1996; Henderson, 1986). Cognitive and personality changes of patients lead to considerable conflict, burden, and grief in relationships. Patients with Alzheimer's or multi-infarct dementia, however, may have an increase in libido, very often accompanied by inappropriate sexual behavior like exposure, public masturbation, or sexual harassment of caregivers (Mooradian, 1991). Litz, Zeiss, and Davies (1990) described psychotherapy with a man who experienced erectile dysfunction because of dramatic changes in his wife's sexual behavior due to Alzheimer's disease. Fedoroff, Peyser, Franz, and Folstein (1994) reported high frequency of sexual disorders in patients with Huntington's disease.

7.04.3.3.6 Depression

Next to dementia, depression is the second important psychiatric disease in the elderly (Helmchen et al., 1996; Henderson, 1986). Loss of libido is a common symptom of depression. Depression can be secondary to sexual dysfunction especially in men. These patients recover from depressive symptoms when sexual potency is improved. Unfortunately, mental health professionals often ignore elderly sexuality (Liptzin, 1984).

7.04.3.3.7 Use of medication

Treatment of chronic diseases almost always requires medication which may have considerable side effects on sexuality. Of the prescribed drugs, 31% are consumed by only 11.7% of the population, (i.e., elderly patients aged over 65) (Mooradian, 1991). Antihypertensive agents, major tranquilizers, minor tranquilizers, antidepressants, and miscellaneous other drugs may cause sexual dysfunction. Antihypertensive agents are probably responsible for the majority of dysfunction. They can lead to impotence, anorgasmia, and ejaculatory disturbances (Segraves & Segraves, 1992).

7.04.3.3.8 Institutionalization

Nursing home residents are often chronically ill and, therefore, dependent on caregivers (Baltes, 1996). Opportunities to express sexual feelings are often restricted because of three factors: disease-related sexual dysfunction, disapproval of sexuality by staff members, and lack of privacy (Hodson & Skeen, 1994; McCartney, Izeman, Rogers, & Cohen, 1987;Mulligan & Modigh, 1991; White, 1982). Nevertheless, institutionalized elderly people do have sexual needs and engage in sexual behaviors (Bretschneider & McCoy, 1988; White, 1982). This fact has been increasingly accepted by nursing staff, and the linkage between knowledge and attitudes towards elderly sexuality in staff has been explored (Commons, Bohn, Godon, Hauser, & Gutheil, 1992; Hillman & Stricker, 1994).

Summarizing age-related changes in sexual functioning, it seems obvious that sexuality in the elderly is not the same as in younger adults. Shaping interests in and capacity for sexual performance of older people, biology does play a role as well as socialization, individual biography with its unique events and experiences, and societal expectations. On the one hand, data show that healthy elderly people can have a satisfying sex life. On the other hand, sexual dysfunction increases with age due to the increase of chronic diseases and its treatment.

7.04.4 SEXUAL DISORDERS AND TREATMENT

Sexual functioning is often classified around three sexual response phases: desire, arousal, and orgasm (American Psychiatric Association [APA], 1994; Schiavi & Segraves, 1995). Sexual desire and sexual arousal are closely connected and the distinction between them raises scientific difficulties. Desire refers to the intensity of sexual drive whereas sexual arousal includes a subjective sense of sexual excitement. Orgasm is defined as being the peak of sexual pleasure and the release of sexual tension. Women are capable of multiple orgasms which can be experienced at the clitoral or the vaginal site. Some women experience them both, sometimes in a blended form, which are reported to be the most intense (McConaghy, 1993).

The following refers to the *Diagnostical and statistical manual of mental disorders* (4th ed.) (*DSM-IV,* APA,1994) classification. It will focus on sexual dysfunction and its treatment. The second and third group of sexual disorders, namely the paraphilias and the gender identity disorders, are not discussed here. A key diagnostic issue is the distinction between sexual disorders as the primary diagnosis versus loss of sexual functioning due to other medical or psychiatric conditions. The *DSM-IV* stresses this point by defining exclusion criteria for sexual desire disorders, sexual arousal disorders, and orgasm disorders as "the sexual dysfunction is not better accounted for by another Axis 1 disorder (except another sexual dysfunction) and is not due exclusively to the direct physiological effects of a substance (e.g., a drug of abuse, a medication) or a general medical condition" (p. 498). Other Axis 1 disorders include major depressive disorder and post-traumatic stress disorder. Sexual dysfunction in the elderly patient is often a secondary diagnosis, caused by chronic illness (Gupta, 1990; Kaiser, 1991; Leiblum & Segraves, 1989; Renshaw, 1984; Weiss & Mellinger, 1990).

Treatment of sexual disorders has changed considerably since the pioneering reports of Masters and Johnson (1966) about their success rates after two weeks of treatment, which had led to tremendous optimism with regard to sex therapy. Unfortunately, reports of successful treatment procedures had often to be corrected because long-term evaluations of the techniques failed to reveal long-term positive outcomes (Dekker & Everaerd, 1983; Fichten, Libman, & Brender, 1986; Kilmann et al., 1986). Therefore, optimism has diminished and sexual difficulties are regarded as complex problems which are the consequence of biological, psychological, cultural, and interpersonal influences. Since the late 1980s, four trends have been identified (Leiblum & Rosen, 1989b).

(i) Increased attention towards biological and medical causes, especially in the diagnosis of erectile difficulties. Vascular, hormonal, and neurological procedures have been developed in order to detect organic factors (Davidson & Rosen, 1992; Melman, 1992; Schiavi, 1992; Segraves & Segraves, 1992; Tiefer & Melman, 1989). The increased age of patients asking for sex therapy has contributed to this trend (Leiblum & Rosen, 1989b).

(ii) Increased pharmacological interventions. Again, erectile difficulties are one focus of attention. Antidepressant and antianxiety drugs are prescribed even for men with psychogenic erectile failures, which was unthinkable in the past. Kaplan and colleagues suggest pharmacotherapy in the treatment of sexual phobias and sexual aversion in addition to her psychodynamically and behavioral-oriented sex therapy (Kaplan, Fyer, & Novick, 1982; Leiblum & Rosen, 1989b; Othmer & Othmer, 1987). This trend is in line with the medical disease model and societal expectations that a sick body is like a machine and should be made to function again as quickly as possible.

(iii) Concentration on desire disorders. On the one hand, clinicians report the growing number of couples presenting desire discrepancies, with usually one partner reporting low desire (Leiblum & Rosen, 1988). On the other hand, there are case reports about sexual addicts or hyperactive sexual interest (Goodman, 1993).

(iv) A therapeutic shift from behavioral to interpersonal system concepts and psychodynamic object-relations theories. Short-term behavioral sex therapies, although successful in some cases, failed with many individuals and with complex problems (Bagarozzi, 1987; Kaplan, 1974, 1977; Leiblum & Rosen, 1992). Intrapsychic difficulties in attachment, dependency, and commitment as well as interpersonal systemic issues have become the focus of attention (Bagarozzi, 1987; Brown & Field, 1988; Leiblum & Rosen, 1992; Levay, 1983; McCarthy, 1984; Talmadge & Talmadge, 1986).

7.04.4.1 Sexual Desire Disorders

These disorders are hypoactive sexual desire disorder (HSDD) and sexual aversion disorder. HSDD is defined as:

> persistently or recurrently deficient (or absent) sexual fantasies and desire for sexual activity. The judgment of deficiency or absence is made by the clinician, taking into account factors that affect sexual functioning, such as age, sex, and the context of the person's life. (APA, 1994, p. 498)

This definition has been frequently criticized because it does not give clear-cut criteria such as length of time, and is therefore highly subjective (LoPiccolo, 1985). Nevertheless, clinicians agree that low or absent desire is one of the most prevalent psychosexual problems seen by them (Leiblum & Rosen, 1988; Rosen & Leiblum, 1989). Desire disorders are believed to be highly refractory to treatment because they are supposed to be a symptom of severe intrapsychic conflicts. Radin (1989) provided a theoretical and clinical understanding of this disorder proposing that preoedipal factors and psychological defense mechanisms that originated during the pregenital years play a significant role in the etiology. Schreiner-Engel and Schiavi (1986) tried to prove the assumption of severe underlying psychopathology. They compared 46 couples with one partner who had inhibited sexual desire with 36 controls. Results showed that individuals with inhibited desire had a history of major and/or intermittent depression twice as high as controls. Furthermore, depressive disorder coincided or preceded sexual dysfunction.

Donahey and Carroll (1993) raised the question whether there is in fact an increase in hypoactive sexual desire, or whether individuals or couples define this as a problem because sexual functioning has become very important in Western cultures. They reported that biological, psychological, and interpersonal factors have been discussed in the etiology of hypoactive sexual desire and that these factors are assumed to affect men and women equally. Donahey and Carroll (1993) hypothesized that hypoactive sexual desire develops and manifests itself differently for men and women. They investigated 47 males and 22 females and found that women reported more stress, and that they were less satisfied with their marriages. The authors suggested that treatment for women should focus more on psychological symptomatology and relationship issues. LoPiccolo (1985) noted that low sexual desire is inconsistent with the cultural expectation and fantasy of male sexuality. Therefore, it is an interesting phenomenon but empirical data are lacking.

Kaplan & Owett (1993) contributed an important physiological reason for inhibited desire in females. They showed that a loss of androgens is associated with a marked decrease in sexual desire and libido. This observation should be further investigated, because the Kaplan and Owett data stem from cancer patients who underwent chemotherapy and, thus, other explanations for decreased interest in sex are also plausible.

The differentiation between sex therapy and psychotherapy has become increasingly difficult. One reason for this change is the growing interest in sexual desire disorders. Many authors have argued that chronic desire disorders can only be treated within individual psychotherapy, in particular psychodynamic approaches, psychoanalysis, and object-relations therapy (Kaplan, 1977; Levay, 1983; Levine, 1988; Radin, 1989; Scharff; 1988; Talmadge & Talmadge, 1986).

Other authors have argued for cognitive-behavioral or systemic-interactional approaches. The necessity of exploring sexual scripts in order to understand sexual fantasies, unreal ideals, and false beliefs has been stressed (Lazarus, 1988; LoPiccolo & Stock, 1986; McCarthy, 1984; Rosen & Leiblum, 1988; 1989). Furthermore, the couple's relationship has to be examined carefully. For example, inhibited sexual desire in one individual can be a very effective means for regulating power distribution (Bagarozzi, 1987).

Some authors have suggested pharmacotherapy for treatment of low desire. Despite some successful reports, results have not been convincing (Leiblum & Rosen, 1989; Segraves, 1988).

Sexual aversion disorder is defined as "persistent or recurrent extreme aversion to, and avoidance of, all (or almost all) genital sexual contact with a sexual partner" (APA, 1994, p. 500). This syndrome occurs in both females and males. Crenshaw (1985) differentiated between primary and secondary aversion. Primary aversion is characterized by a negative response to any sexual interactions from the earliest memories to the present and is often combined with other dysfunction. A psychosexual trauma like incest or child molestation is frequently found. Secondary aversion is acquired, which means that sexuality was found pleasurable in the past. In the beginning of the disorder individuals may function well sexually when they get involved in sexual activities. The differentiation between the two desire disorders is often difficult. There are no special treatment suggestions.

7.04.4.2 Female Disorders

LaFerla (1984) listed a number of factors contributing to female sexual dysfunction in desire and orgasmic response. He differentiated psychosocial and organic factors. Psychosocial factors are relationship issues, negative past experiences, and depression as well as drug and alcohol abuse. Organic factors are chronic illness, pregnancy, pharmacological agents, and endocrine alterations.

7.04.4.2.1 Arousal disorder

For females, arousal disorder means a "persistent or recurrent inability to attain, or to maintain until completion of the sexual activity, an adequate lubrication-swelling response of sexual excitement" (APA, 1994, p. 502). The excitement phase in elderly women is much slower than in younger ones. Whereas in young women lubrication occurs after 10–15 seconds, it requires at least five minutes in the postmenopausal women (Gupta, 1990). However, this change is normal and cannot be considered a disorder. Some women produce insufficient vaginal lubrication altogether, and they may need artificial lubricants.

7.04.4.2.2 Orgasmic disorder

The female orgasmic disorder is defined as a:

> persistent or recurrent delay in, or absence of, orgasm following a normal sexual excitement phase. Women exhibit wide variability in the type or intensity of stimulation that triggers orgasm. The diagnosis of female orgasmic disorder should be based on the clinician's judgment that the woman's orgasmic capacity is less than would be reasonable for her age, sexual experience, and the adequacy of sexual stimulation she receives. (APA, 1994, p. 506).

Orgasmic disorder is still one of the most frequent sexual dysfunctions in women. It is divided into primary and secondary anorgasmia. Primary anorgasmia refers to women who have never experienced an orgasm, secondary anorgasmia refers to women who sometimes or in specific situations (e.g., masturbation) react orgasmically. Empirical data from cross-sectional studies suggest that the rate of women who never or rarely react with orgasm decreases from 50% in young adulthood to 21% at the age of 40 (McConaghy, 1993), remains fairly stable with 21% between the ages of 40 and 59 (Laumann et al., 1994), and increases to 28% over the age of 70 (Bergström-Walan & Neilsen, 1990; Brecher, 1984). Nevertheless, these data are difficult to evaluate. Concerns about women's orgasm are a rather new phenomenon. Women can have an orgasm in the same amount of time as men, and they can achieve multiple orgasms. Personal history, relationship difficulties, lack of education, and lack of knowledge about sexuality may be reasons for the inhibition of the orgasmic response. Very little is known about factors that contribute to orgasmic capacity. The neurophysiological basis of human orgasm is underexplored (Heiman & Grafton-Becker, 1989).

Treatment approaches of this disorder depend on the theoretical orientation of the therapist. Three major approaches are psychoanalytic, cognitive-behavioral, and systemic theory. Psychoanalytic psychotherapy does not focus on symptom removal but on working through underlying conflicts that lead to the symptom. Once these conflicts are understood and worked through, symptoms disappear. Kaplan (1974, 1977) combined a psychodynamic approach with behavioral sex therapy elements.

Cognitive-behavioral therapy depends on learning theory and tries to promote cognitive as well as behavioral changes. Behavioral exercises, like directed masturbation, partner stimulation, and Kegel exercises, are frequently used. Whereas directed masturbation seems to be successful in treatment of primary anorgasmia, its long-term effectiveness in treating secondary anorgasmia is questionable (Fichten, Libman, & Brender, 1986; Kilmann et al., 1986; Kuriansky, Sharpe, & O'Connor, 1982; Nairne & Hemsley, 1983; Wakefield, 1987).

Systemic therapists concentrate on the couple as a system and do not focus on the individual woman with anorgasmia. Her dysfunction is regarded as a problem of the system. Thus, therapists focus on interactional patterns. Heimann and Grafton-Becker (1989) suggested an integrated theory and practice that combines the theoretical and practical elements of psychoanalytic, cognitive-behavioral, and systemic theories.

7.04.4.2.3 Sexual pain disorders

Dyspareunia and vaginismus are sexual pain disorders. Dyspareunia is a "recurrent or persistent genital pain associated with sexual intercourse in either a male or a female" (APA, 1994, p. 513). This disorder can have organic and/or psychological causes. Careful assessment and diagnosis is very important (Sandberg & Quevillon, 1987). Lazarus (1989) described three psychological factors, namely traumatic (e.g., rape), developmental (e.g., religious taboos) and relational (e.g., resentment, insufficient sexual stimulation through partner). Unsatisfactory relationships are the major reason for this dysfunction. After a careful multimodal assessment, treatment depends on the identified cause and may consist of different methods (e.g., relationship enhancement, social skills training, imagery techniques, systematic desensitization, and sensate focus).

Vaginismus is defined as a "recurrent or persistent involuntary spasm of the musculature of the outer third of the vagina that interferes with sexual intercourse" (APA, 1994, p. 515). Women with vaginismus may function sexually without difficulties except their inability to have intercourse. Different psychological theories agree that anxiety is the core of these spastic contractions. Treatment very often consists of the gradual insertion of objects in the vagina. Dilators with increasing size are frequently used. Examples of treatment with different outcomes are discussed by Leiblum, Pervin, and Campell (1989).

7.04.4.3 Male Disorders

7.04.4.3.1 Arousal disorder

For males, arousal disorder is defined as erectile disorder: "Persistent or recurrent inability to attain, or to maintain until completion of the sexual activity, an adequate erection" (APA, 1994, p. 504). Erectile difficulties increase with age and they are the major sexual problem of elderly males (Kaiser, 1991; Levine, 1992; Nelson, 1987; Rosen & Leiblum, 1992a, 1992b; Segraves & Segraves, 1992; Weiss & Mellinger, 1990).

For many years psychological factors have been perceived as major causes of erectile difficulties, and organic causes have been widely overlooked. Recent work has focused on differential diagnosis of physiological and psychological factors (Kinder & Curtiss, 1988; LoPiccolo, 1985). Levine (1992) suggested four categories for classification: psychogenic, organic, mixed, and idiopathic (when the professional cannot identify the cause with certainty). Disease or drug-related erectile failures are common in elderly men but psychogenic factors may work in addition.

Intensive research efforts have focused on the determinants of erectile dysfunction. In addition, numerous treatment approaches have been tried (Rosen & Leiblum, 1992). Modern psychotherapeutic treatment concentrates on the systemic function of erectile difficulties (Althof, 1989; Levine, 1992; Leiblum & Rosen, 1992; LoPiccolo, 1992; Speckens et al., 1995).

Surgical treatment has become very important in recent years. There is surgical correction of reversible dysfunction and, in cases of nonreversibility, penile prosthesis can be implanted. Implant devices have, however, considerable risks and disadvantages, and careful consideration has to precede surgery (LoPiccolo & Stock, 1986; Melman & Tiefer, 1992). Injection therapy and vacuum devices are also widely used (Althof & Turner, 1992; Kaplan, 1990).

7.04.4.3.2 Orgasmic disorders

In males, there are two orgasm disorders, namely the male orgasmic disorder and premature ejaculation. Male orgasmic disorder is defined as a "persistent or recurrent delay in, or absence of, orgasm following a normal sexual excitement phase during sexual activity that the clinician, taking into account the person's age, judges to be adequate in focus, intensity, and duration" (APA, 1994, p. 509). This disorder is not often seen in a clinical context. Apfelbaum (1989) proposed that the retarded ejaculator is more responsive to his own stimulation and experiences hostility towards his partner. The patient has to become aware of his distance and anger before he can make any progress.

Premature ejaculation is defined as a "persistent or recurrent ejaculation with minimal sexual stimulation before, on, or shortly after penetration and before the person wishes it" (APA, 1994, p. 511). This disorder is more common in younger men and its treatment seems to be very effective. Information about different treatment techniques are given by LoPiccolo (1985), LoPiccolo and Stock (1986), and McCarthy (1989).

7.04.4.4 Special Issues in Treatment with Elderly Patients

Basically, there are no differences between sexual difficulties and dysfunction between younger and older adults. Nevertheless, there are some issues which should be considered. First, biological changes with age and health-related factors have to be explored thoroughly. Second, cohort experiences and values have to be considered. Third, life events and circumstances have to be evaluated (Leiblum & Segraves, 1989; LoPiccolo, 1991).

Physiological and health-related changes have been previously described in this chapter. Clinical observations about the behavior of hundreds of elderly men, women, and couples who have sought help for sexual difficulties revealed adaptation processes to the major physiological changes. Kaplan (1990) summarized that couples who have a good and intimate relationship, who care for each other and each other's pleasure, and who are sexually open and free of severe conflicts, can adapt to these changes very successfully. She described the following coping mechanisms that may help to maintain sexual functioning:

(i) *Diminished sexual desire*. This problem may occur in either one or both partners and can be overcome with the help of fantasies or erotica. Testosterone replacement is helpful in some cases. Treatment of possible depression or substance abuse might be necessary.

(ii) *Higher penile threshold for mental and physical stimulation/delayed lubrication*. These problems can be solved with partner-provided, extended physical stimulation. Erotica or fantasy and more rapid love making are helpful. Morning sex is recommended because elderly men function better in the morning due to higher testosterone levels.

(iii) *Vaginal dryness and atrophy; dyspareunia*. Use of lubricants prior to coitus and/or estrogen replacement are beneficial. Intercourse on a regular basis and dilators may be useful.

(iv) *Softer penile erections*. Coital positions which facilitate intromission (e.g., stuffing) are recommended. Stimulation, enjoyment, and orgasms by manual or oral practices should be emphasized.

(v) *Lengthened male refractory period*. This causes less frequent intercourse. Quality vs. quantity of experience should be emphasized.

(vi) *Physical mobility and pain*. Analgesics taken before sexual activity or a hot bath to reduce stiffness may help. A side-by-side-position is often more comfortable.

Counseling or treatment should be done cautiously. Very often, information about age-related changes and "permission giving" of new experiences (e.g., more manual or oral stimulation, fantasies, new positions) are helpful. Keeping in mind cohort differences, the therapist should carefully evaluate the patient's value system and possible shame reactions before any suggestions are made. Middle-class values of younger adults are easily imposed and might substantially harm the elderly patient. For example, masturbation, oral sex, or the use of a vibrator may be common practices for sex therapists but not necessarily for a couple born in the countryside in the 1920s.

In some cases, sexual dysfunction of elderly patients has a severe underlying conflict. For example, meeting the psychosocial crises of ego integrity or ego despair may result in sexual dysfunction (Fagan, Meyer, & Schmidt, 1986). In these cases individual psychotherapy should be considered. Contrary to common beliefs and stereotypes of many psychotherapists of all theoretical backgrounds, psychotherapy can be very successful with elderly patients (Radebold, 1992; Smyer, Zarit, & Qualls, 1990). It is important, however, that psychotherapists of elderly people have some gerontological and historical knowledge, and that they are aware of special countertransference issues (Apfel, Fox, Isberg, & Levine, 1984; Radebold, 1992).

Finally, it has to be recognized that sexual behavior is not easily modifiable. Sexual anxiety, which is believed to be the core of most sexual dysfunction (Patterson & O'Gorman, 1989), has "an important adaptive feature of human sexual behavior" (Cole, 1985, p. 337). Cole presented a critical review about sex therapy and its outcome research. He concluded that sex therapy has intrinsic limitations which have to be accepted. Vandereycken (1987) argued in addition that sexual lust has turned from a taboo into a "must." The establishment of new behavioral demands should certainly not be the result of gerontological research and psychotherapy with elderly people. Therefore, careful and critical reflection of scientific data and therapeutic aims is crucial.

7.04.5 CONCLUSION

This chapter has examined literature on sexuality and aging. In spite of the growing body of research on old age and aging, there is a considerable lack of sound studies in the context of sexuality and aging, particularly with respect to elderly women. Epidemiological data have been presented and their methodological and theoretical problems described. Age-related physiological changes in women and men were examined. It is important to differentiate between normal age-related changes and health-related problems. Chronic disease increases with age, but deleterious effects on sexual enjoyment are due to disease, and not to aging *per se*. Sexual disorders of desire, arousal, and orgasm as well as treatment approaches have been described. Finally, sexuality of the elderly individual should be embedded in a life-span developmental perspective. The important role that the individual biography plays in the aging process is an accepted fact in modern gerontology. Further research should pay more attention to intra-individual development. Many research results (e.g., differences in sexual activities between men and women) may be due to cohort effects and not to biological factors.

7.04.6 REFERENCES

Adams, C. G., & Turner, B. F. (1985). Reported change in sexuality from young adulthood to old age. *The Journal of Sex Research*, *21*(2), 126–141.

Althof, S. E. (1989). Psychogenic impotence: Treatment of men and couples. In S. R. Leiblum & R. C. Rosen (Eds.), *Principles and practice of sex therapy: Update for the 1990s* (2nd ed., pp. 237–265). New York: Guilford Press.

Althof, S. E., & Turner, L. A. (1992). Self-injection therapy and external vacuum devices in the treatment of erectile dysfunction: Methods and outcome. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: Assessment and treatment* (pp. 283–309). New York: Guilford Press.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Antonovsky, H., Sadowsky, M., & Maoz, B. (1990). Sexual activity of aging men and women: An Israeli study. *Behavior, Health, and Aging, 1*(3), 151–161.

Apfel, R. J., Fox, M., Isberg, R. S., & Levine, A. R. (1984). Countertransference and transference in couple therapy: Treating sexual dysfunction in older couples. *Journal of Geriatric Psychiatry, 17*(2), 203–214.

Apfelbaum, B. (1989). Retarded ejaculation: A much-misunderstood syndrome. In S. R. Leiblum & R. C. Rosen (Eds.), *Principles and practice of sex therapy: Update for the 1990s* (2nd ed., pp. 168–206). New York: Guilford Press.

Aresin, L. (1980). Sexualverhalten nach dem 60. Lebensjahr. *Zeitschrift für Alternsforschung*, *35*, 55–58.

Bachmann, G. A., & Leiblum, S. R. (1991). Sexuality in sexagenarian women. *Maturitas*, *13*(1), 43–50.

Bagarozzi, D. A. (1987). Marital/family developmental theory as a context for understanding and treating inhibited sexual desire. *Journal of Sex and Marital Therapy, 13*(4), 276–285.

Baltes, M. M. (1996). *The many faces of dependency in old age*. Cambridge, UK: Cambridge University Press.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology, 23,* 611–626.

Baltes, P. B. & Baltes, M. M. (Eds.) (1990). *Successful aging: Perspectives from the behavioral sciences*. Cambridge, UK: Cambridge University Press.

Baltes, P. B., Reese, H. W., & Lipsitt, L. P. (1980). Life-span developmental psychology. *Annual Review of Psychology, 31,* 65–110.

Bengston, V. L., Reedy, M. N., & Gordon, C. (1985). Aging and self-conceptions: Personality processes and social contents. In J. E. Birren & K. W. Schaie (Eds.). *Handbook of the psychology of aging* (2nd ed., pp. 544–593). New York: Van Nostrand Reinhold.

Bergström-Walan, M.-B., & Nielsen, H. H. (1990). Sexual expression among 60–80-year-old men and women: A sample from Stockholm, Sweden. *The Journal of Sex Research, 27*(2), 289–295.

Bernhard, L. (1986). Methodology issues in studies of sexuality and hysterectomy. *The Journal of Sex Research*, *22*(1), 108–128.

Birren, J. E. & Schaie, K. W. (Eds.) (1985). *Handbook of the psychology of aging* (2nd ed.). New York: Van Nostrand Reinhold.

Brähler, E., & Unger, U. (1994). Sexuelle Aktivität im höheren Lebensalter im Kontext von Geschlecht, Familienstand und Persönlichkeitsaspekten—Ergebnisse einer repräsentativen Befragung. (Sexual activity in old age in context of gender, marital status and personality—Results of a representative study). *Zeitschrift für Gerontologie, 27* ,110–115.

Brecher, E. M. (1984). *Love, sex and aging: A Consumer Union report*. Boston: Little Brown.

Bretschneider, J. G., & McCoy, N. L. (1988). Sexual interest and behavior in healthy 80- to 102-year-olds. *Archives of Sexual Behavior, 7*(2), 109–129.

Brown, R. A., & Field, J. R. (Eds.) (1988). *Treatment of sexual problems in individual and couples therapy*. New York: PMA.

Burgener, S., & Logan, G. (1989). Sexuality concerns of the post-stroke patient. *Rehabilitation Nursing, 14*(4), 178–181.

Byers, J. P. (1983). Sexuality and the elderly. *Geriatric Nursing, 4*, 293–297.

Capuzzi, D. (1982). Sexuality and aging: An overview for counselors. *Personnel and Guidance Journal, 61*, 31–35.

Christenson, C., & Johnson, A. B. (1973). Sexual patterns in a group of older never married women. *Journal of Geriatric Psychiatry, 6*(1), 80–98.

Colarusso, C., & Nemiroff, R. (1981). *Adult development: A new dimension in psychodynamic theory and practice*. New York: Plenum.

Cole, M. (1985). Sex therapy: A critical appraisal. *British Journal of Psychiatry, 147*, 337–351.

Commons, M. L., Bohn, J. T., Godon, L. T., Hauser, M. J., & Gutheil, T. G. (1992). Professional's attitudes towards sex between institutionalized patients. *American Journal of Psychotherapy, 46*(4), 571–580.

Corby, N., & Zarit, J. M. (1983). Old and alone: The unmarried in later life. In R. B. Weg (Ed.), *Sexuality in the later years. Roles and behavior* (pp. 131–145). New York: Academic Press.

Costa, P. T., Jr., Zondermann, A. B., McCrae, R. R., Cornoni-Huntley, J., Locke, B. Z., & Barbano, H. E. (1987). Longitudinal analyses of psychological well-being in a national sample: Stability of mean levels. *Journal of Gerontology, 42*, 50–55.

Crenshaw, T. L. (1985). The sexual aversion syndrome. *Journal of Sex and Marital Therapy, 11*(4), 285–292.

Cruikshank, M. (1990). Lavender and gray: A brief survey of lesbian and gay aging studies. *Journal of Homosexuality, 20*(3/4), 77–87.

Datan, N., & Rodeheaver, D. (1983). Beyond generativity: Toward a sensuality of later life. In R. B. Weg (Ed.), *Sexuality in the later years. Roles and behavior* (pp. 280–288). New York: Academic Press.

Davidson, J. M., & Rosen, R. C. (1992). Hormonal determinants of erectile function. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: Assessment and treatment* (pp. 72–95). New York: Guilford.

Dekker, J., & Everaerd, W. (1983). A long-term follow-up study of couples treated for sexual dysfunctions. *Journal of Sex and Marital Therapy, 9*(2), 99–113.

Dennerstein, L., Smith, A. M. A., Morse, C. A., & Burger, H. G. (1994). Sexuality and the menopause. *Journal of Psychosomatic Obstetrics and Gynaecology, 15*, 59–66.

Denney, N. W., Field, J. K., & Quadagno, D. (1984). Sex differences in sexual needs and desires. *Archives of Sexual Behavior, 13*, 233–245.

Donahey, K. M., & Carroll, R. A. (1993). Gender differences in factors associated with hypoactive sexual desire. *Journal of Sex and Marital Therapy, 19*(1), 25–40.

Eckert, J. W. (1984). "Clinical perspectives on sexuality in older patients": Discussion. *Journal of Geriatric Psychiatry, 17*(2), 183–188.

Eicher W. (1994a). Hysterektomie und Sexualität (I). Eine Standortbestimmung. *Sexualmedizin, 5,* 144–148.

Eicher W. (1994b). Hysterektomie und Sexualität (II). Eine Standortbestimmung. *Sexualmedizin, 6,* 176–180.

Erikson, E. H. (1950). *Childhood and Society*. New York: Norton.

Fagan, P. J., Meyer, J. K., & Schmidt, C. W. (1986). Sexual dysfunction within an adult developmental perspective. *Journal of Sex and Marital Therapy, 12*(4), 243–257.

Fedoroff, J. P., Peyser, C., Franz, M. L., & Folstein, S. E. (1994). Sexual disorders in Huntington's disease. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *6*, 147–153.

Fichten, C. S., Libman, E., & Brender, W. (1986). Measurement of therapy outcome and maintenance of gains in the behavioral treatment of secondary orgasmic dysfunction. *Journal of Sex and Marital Therapy, 12*(1), 22–34.

Fooken, I. (1994). Sexuality in later years: The impact of health and body image in a sample of older women. *Patient Education and Counseling, 23*(3) 227–233.

Freud, S. (1964). Drei Abhandlungen zur Sexualtheorie. (Three contributions to the theory of sex). *Gesammelte Werke*. (Vol. V, pp. 27–145). Frankfurt, Germany: Fischer. (Original work published 1905).

Friend, R. A. (1990). Older lesbian and gay people: A theory of successful aging. *Journal of Homosexuality, 20*(3–4), 99–118.

Garza, J. M. & Dressel, P. L. (1983). Sexuality and later-life marriages. In: T. H. Brubaker (Ed.), *Family relationships in later life* (pp. 91–108). Beverly Hills, CA: Sage.

Gatz, M. (1987). *Adult development and aging: Growing older in a contingent and noncontingent world.* Presented at the Association for Advancement of Behavior Therapy annual convention, Boston.

Gendel, E. S. (1986). Self-esteem and sexuality in older patients with cancer. *Frontiers of radiation therapy and oncology, 20,* 166–172.

Genevay, B. (1982). In praise of older women. In M. Kirkpatrick (Ed.), *Women' s sexual experience. Explorations of the dark continent* (pp. 87–101). New York: Plenum.

George, L. K., & Weiler, S. J. (1981). Sexuality in middle and late life: The effects of age, cohort and gender. *Archives of General Psychiatry*, *38*, 919–923.

Goldstein, M. K., & Teng, N. N. H. (1991). Gynecological factors in sexual dysfunction of the older women, *Clinics in Geriatric Medicine, 7*(1), 41–61.

Goodman, A. (1993). Diagnosis and treatment of sexual addiction. *Journal of Sex and Marital Therapy, 19*(3), 225–251.

Greeley, A. M. (1994). Janus report. *Contemporary Sociology*. Forthcoming.

Gupta, K. (1990). Sexual dysfunction in elderly women. *Clinics in Geriatric Medicine, 6*(1), 197–203.

Hagstad, A. (1988). Gynecology and sexuality in middle-aged women. *Women and Health, 13*, 57–80.

Hagstad, A., & Janson, P. O. (1984). Sexuality among Swedish women around forty: An epidemiological survey. *Journal of Psychosomatic Obstetrics and Gynaecology, 3*, 191–203.

Hällström, T., & Samuelsson, S. (1990). Changes in women's sexual desire in middle life: The longitudinal study of women in Gothenburg. *Archives of Sexual Behavior, 19*(3), 259–268.

Heiman, J. R., & Grafton-Becker, V. (1989). Orgasmic disorders in women. In S. R. Leiblum & R. C. Rosen (Eds.), *Principles and practice of sex therapy: Update for the 1990s* (2nd ed., pp. 51–88). New York: Guilford.

Helmchen, H., Baltes, M. M., Geiselmann, B., Kanowski, S., Linden, M., Reischies, F. M., Wagner, M., & Wilms, H.-U. (1996). Psychische Erkrankungen im Alter. In K. U. Mayer & P. B. Baltes (Eds.), *Die Berliner Altersstudie* (pp. 185–219). Berlin, Germany: Akademie Verlag.

Helmström, L. , Lundberg, P. O., Sörbom, D., & Bäckstrom, T. (1993). Sexuality after hysterectomy: A factor analysis of women's sexual lives before and after subtotal hysterectomy. *Obstetrics and Gynecology*, *81*(3), 357–362.

Henderson, A. S. (1986). Epidemiology of mental illness. In H. Häfner, G. Moschel & N. Sartorius (Eds.), *Mental Health in the Elderly. A Review of the Present State of Research* (pp. 29–34). Berlin, Germany: Springer-Verlag.

Hermanova, H. M. (1983). Human sexuality and aging. In M. Bergener, U. Lehr, E. Lang, & R. Schmitz-Scherzer (Eds.), *Aging in the eighties and beyond* (pp. 324–332). New York: Springer.

Hillman, J. L., & Stricker, G. (1994). A linkage of knowledge and attitudes toward elderly sexuality: Not necessarily a uniform relationship. *The Gerontologist, 34*(2), 256–260.

Hite, S. (1976). *The Hite report*. New York: Dell.

Hodson, D. S., & Skeen, P. (1994). Sexuality and aging: The hammerlock of myths. *Journal of Applied Gerontology, 13*(3), 219–235.

Institute for Sex Research (1953). *Sexual behavior in the human female*. Philadelphia: Saunders.

Janus, S. S., & Janus, C. L. (1993). *The Janus report on sexual behavior*. New York: Wiley.

Jung, C. G. (1933). *Modern Man in Search of a Soul*. New York: Harcourt Brace.

Kaiser, F. E. (1991). Sexuality and impotence in the aging man. *Clinics in Geriatric Medicine, 7*(1), 63–72.

Kaiser, F. E. (1992). Sexual function and the older cancer patient. *Oncology*, 6(2), 112–118.

Kaplan, H. S. (1974). *The new sex therapy*. New York: Brunner/Mazel.

Kaplan, H. S. (1977). Hypoactive sexual desire. *Journal of Sex and Marital Therapy, 3,* 3–9.

Kaplan, H. S. (1990). The combined use of sex therapy and intrapenile injections in the treatment of impotence. *Journal of Sex and Marital Therapy, 16*(4), 195–207.

Kaplan, H. S. (1992). A neglected issue: The sexual side effects of current treatments for breast cancer. *Journal of Sex and Marital Therapy*, *18*(1), 3–19.

Kaplan, H. S., Fyer, A. J., & Novick, A. (1982). The treatment of sexual phobias: The combined use of antipanic medication and sex therapy. *Journal of Sex and Marital Therapy, 8*(1), 3–28.

Kaplan, H. S., & Owett, T. (1993). The female androgen deficiency syndrome. *Journal of Sex and Marital Therapy*, *19*(1), 3–24.

Kay, B., & Neelley, J. N. (1982). Sexuality and aging: A review of current literature. *Sexuality and Disability, 5*(1), 38–46.

Kellett, J. M. (1993). Sexuality in later life. *Reviews in Clinical Gerontology, 3*(3), 309–314.

Kernberg, O. F. (1977). Normal psychology of the aging process revisited-II. Discussion. *Journal of Geriatric Psychiatry*, *10,* 27–45.

Kilmann, P. R., Boland, J. P., Norton, S. P., Davidson, E., & Caid, C. (1986). Perspectives of sex therapy outcome: a survey of AASECT providers. *Journal of Sex and Marital Therapy, 12*(2), 116–138.

Kilmann, P. R., Mills, K. H., Caid, C., Davidson, E., Bella, B., Milan, R., Drose, G., Boland, J., Follingstad, D., Montgomery, B., & Wanlass, R. (1986). Treatment of secondary orgasmic dysfunction: An outcome study. *Archives of Sexual Behavior, 15*(3), 211–229.

Kinder, B. N., & Curtiss, G. (1988). Specific components in the etiology, assessment, and treatment of male sexual dysfunction: Controlled outcome studies. *Journal of Sex and Marital Therapy, 14*(1), 40–48.

Kinn, A.-C., & Öhman, U. (1986). Bladder and sexual function after surgery for rectal cancer. *Diseases of the Colon and Rectum, 29*, 43–48.

Kinsey, A. C., Pomeroy, W. B., & Martin, C. R. (1948). *Sexual behavior in the human male*. Philadelphia: Saunders.

Kirkpatrick, M. (1980). *Women's sexual development: Explorations of inner space*. New York: Plenum.

Klassen, A. D., Williams, C. J., & Levitt, E. E. (1989). *Sex and morality in the US. An empirical inquiry under the auspices of the Kinsey institute*. Middletown, CT: Wesleyan University Press.

Kruse, A. (1989). Psychologie des Alters. In: K. P. Kisker, H. Lauter, J.-E. Meyer, C. Müller & E. Strömgren (Eds.), *Psychiatrie der Gegenwart 8, Alterspsychiatrie*. (pp. 1–58). Berlin, Germany: Springer-Verlag.

Kuriansky, J. B., Sharpe, L., & O'Connor, D. (1982). The treatment of anorgasmia: Long-term effectiveness of a short-term behavioral group therapy. *Journal of Sex and Marital Therapy, 8*(1), 29–43.

LaFerla, J. J. (1984). Inhibited sexual desire and orgasmic dysfunction in women. *Clinical Obstetrics and Gynecology, 27*(3), 738–749.

Laumann, E. O., Gagnon, J. H., Michael, R. T., & Michels, S. (1994). *The social organization of sexuality. Sexual practices in the United States*. Chicago: University of Chicago Press.

Laws, J. L. (1980). Female sexuality through the life-span. In P. B. Baltes & O. G. Brim, (Eds.), *Life-span development and behavior* (Vol. 3, pp. 207–252). New York: Academic Press.

Lazarus, A. A. (1988). A multimodal perspective on problems of sexual desire. In S. R. Leiblum & R. C. Rosen (Eds.), *Sexual desire disorders* (pp. 145–167). New York: Guilford Press.

Lazarus, A. A. (1989). Dyspareunia: A multimodal psychotherapeutic perspective. In S. R. Leiblum & R. C. Rosen (Eds.), *Principles and practice of sex therapy: Update for the 1990s* (2nd ed., pp. 89–112). New York: Guilford.

Lehr, U. & Thomae, H. (Eds.) (1987), *Formen seelischen Alterns: Ergebnisse der Bonner Längsschnittstudie (BOLSA)*. Stuttgart, Germany: Enke.

Leiblum, S. R. (1990). Sexuality and the mid-life woman. Special issue: Women at midlife and beyond. *Psychology of Women Quarterly, 14*, 495–508.

Leiblum, S. R., Pervin, L. A., & Campbell, E. H. (1989). The treatment of vaginismus: Success and failure. In S. R. Leiblum & R. C. Rosen (Eds.), *Principles and practice of sex therapy: Update for the 1990s* (2nd ed., pp. 113–138). New York: Guilford Press.

Leiblum S. R., & Rosen, R. C. (Eds.) (1988). *Sexual desire disorders*. New York: Guilford Press.

Leiblum, S. R. & Rosen, R. C. (1989a). Introduction: Sex therapy in the age of Aids. In S. R. Leiblum & R. C. Rosen (Eds.), *Principles and practice of sex therapy: Update for the 1990s* (2nd ed., pp. 1–16). New York: Guilford Press.

Leiblum, S. R., & Rosen, R. C. (Eds.). (1989b). *Principles and practice of sex therapy: Update for the 1990s* (2nd ed.). New York: Guilford Press.

Leiblum, S. R., & Rosen, R. C. (1992). Couples therapy for erectile disorders: Observations, obstacles, and outcomes. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: Assessment and treatment* (pp. 226–254). New York: Guilford Press.

Leiblum, S. R., & Segraves, R. T. (1989). Sex therapy with aging adults. In S. R. Leiblum & R. C. Rosen (Eds.), *Principles and practice of sex therapy: Update for the 1990s* (2nd ed., pp. 352–381). New York: Guilford Press.

Levay, A. N. (1983). Personality change in sex therapy. *Journal of the American Academy of Psychoanalysis, 11*(3), 425–433.

Levine, S. B. (1988). Intrapsychic and individual aspects of sexual desire. In S. R. Leiblum & R. C. Rosen (Eds.), *Sexual desire disorders* (pp. 21–44). New York: Guilford Press.

Levine, S. B. (1992). Intrapsychic and interpersonal aspects of impotence: Psychogenic erectile dysfunction. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: Assessment and treatment*. (pp. 198–225). New York: Guilford Press.

Libman, E. (1989). Sociocultural and cognitive factors in aging and sexual expression: Conceptual and research issues. *Canadian Psychology/Psychologie Canadienne, 30*(3), 560–567.

Liptzin, B. (1984). Clinical perspectives on sexuality in older patients. *Journal of Geriatric Psychiatry, 17* (2), 167–188.

Litz, B. T., Zeiss, A. M., & Davies, H. D. (1990). Sexual concerns of male spouses of female Alzheimer's disease patients. *The Gerontologist*, *30*(1), 113–116.

Livson, F. B, (1983). Gender identity: A life-span view of sex-role development. In R. B. Weg (Ed.), *Sexuality in the later years: Roles and behavior* (pp. 105–127). New York: Academic Press.

LoPiccolo, J. (1985). Diagnosis and treatment of male sexual dysfunction. *Journal of Sex and Marital Therapy, 11*(4), 215–232.

LoPiccolo, J. (1991). Counseling and therapy for sexual problems in the elderly. *Clinics in Geriatric Medicine, 7*(1), 161–179.

LoPiccolo (1992). Postmodern therapy for erectile failure. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: Assessment and treatment* (pp. 171–197). New York: Guilford Press.

LoPiccolo, J., & Stock, W. E. (1986). Treatment of sexual dysfunction. *Journal of Consulting and Clinical Psychology, 54*(2), 158–167.

Ludeman, K. (1981). The sexuality of the older person: Review of the literature. *The Gerontologist, 21*(2), 203–208.

Luria, Z., & Meade, R. G. (1984). Sexuality in the middle-aged woman. In: G. Baruch & R. Brooks-Gunn (Eds.), *Women in mid-life* (pp. 371–397). New York: Plenum.

Majerovitz, S. D., & Revenson, T. A. (1994). Sexuality and rheumatic disease: The significance of gender. *Arthritis-Care Research, 7*(1), 29–34.

Malatesta, V. J., Chambless, D. L., Pollack, M, & Cantor, A. (1988). Widowhood, sexuality and aging: A life span analysis. *Journal of Sex and Marital Therapy, 14*(1), 49–62.

Mann, T. (1953). Die Betrogene. (The black swan). In T. Mann. *Gesammelte Werke*. Bd. 8. Frankfurt, Germany: Fischer.

Margolis, G., Goodman, R. L., & Rubin, A. (1990). Psychological effects of breast-conserving cancer treatment and mastectomy. *Psychosomatics, 31*(1), 33–39.

Masters, W. H., & Johnson, V. E. (1966). *Human sexual response*. Boston: Little Brown.

Masters, W. H., & Johnson, V. E. (1981). Sex and the aging process. *Journal of the American Geriatrics Society, 19*(9), 385–390.

Mayer, K. U., & Baltes, P. B. (Eds.) (1996) . *Die Berliner Altersstudie*. Berlin, Germany: Akademie Verlag.

McCarthy, B. W. (1984). Strategies and techniques for the treatment of inhibited sexual desire. *Journal of Sex and Marital Therapy, 10*(2), 97–104.

McCarthy, B. W. (1989). Cognitive-behavioral strategies and techniques in the treatment of early ejaculation. In S. R. Leiblum & R. C. Rosen (Eds.), *Principles and practice of sex therapy: Update for the 1990s* (2nd ed., pp. 141–167). New York: Guilford Press.

McCartney, J. R., Izeman, H., Rogers, D., & Cohen, N. (1987). Sexuality and the institutionalized elderly. *Journal of the American Geriatrics Society, 35*, 331–333.

McConaghy, N. (1993). *Sexual behavior: Problems and management*. New York: Plenum.

McCoy, N. L. , & Davidson, J. M. (1985). A longitudinal study of the effects of menopause on sexuality. *Maturitas, 7,* 203–210.

McDougall, G. J. (1993). Therapeutic issues with gay and lesbian elders. Special issue: The forgotten aged: Ethnic, psychiatric and societal minorities. *Clinical Gerontologist, 14*(1), 45–57.

Melman, A. (1992). Neural and vascular control of erection. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: Assessment and treatment*. (pp. 55–71). New York: Guilford Press.

Melman, A., & Tiefer, L. (1992). Surgery for erectile disorders: Operative procedures and psychological issues. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: Assessment and treatment* (pp. 255–282). New York: Guilford Press.

Mooradian, A. D. (1991). Geriatric sexuality and chronic diseases. *Clinics in Geriatric Medicine, 7*(1), 113–131.

Morley, J. E. (1991). Endocrine factors in geriatric sexuality. *Clinics in Geriatric Medicine, 7*(1), 85–93.

Morley, J. E., & Kaiser, F. E. (1989). Sexual function with advancing age. *Medical Clinics of North America, 73*(6), 1483–1495.

Mulligan, T., & Modigh, A. (1991). Sexuality in dependent living situations. *Clinics in Geriatric Medicine, 7*(1), 153–160.

Mulligan, T., & Moss, C. R. (1991). Sexuality and aging in male veterans: A cross-sectional study of interest, ability, and activity. *Archives of Sexual Behavior, 20*(1), 17–25.

Myers, L. S. (1995). Methodological review and meta-analysis of sexuality and menopause research. *Neuroscience and Biobehavioral Reviews, 19*(2), 331–341.

Nadelson, C. C. (1984). Geriatric sex problems: Discussion. *Journal of Geriatric Psychiatry, 17*(2), 139–148.

Nairne, K. D., & Hemsley, D. R. (1983). The use of directed masturbation training in the treatment of primary anorgasmia. *British Journal of Clinical Psychology, 22*, 283–294.

Nelson, R. P. (1987). Male sexual dysfunction: Evaluation and treatment. *Southern Medical Journal*, *80,* 69–74.

Newman, A. S., & Bertelson, A. D. (1986). Sexual dysfunction in diabetic women. *Journal of Behavioral Medicine, 9*(3), 261–270.

O'Donohue, W. T. (1987). The sexual behavior and problems of the elderly. In L. Carstensen & B. Edelstein (Eds.), *Handbook of clinical gerontology* (pp. 66–75). New York: Pergamon.

Opjordsmoen, S., Waehre, H., Aass, N., & Fossa, S. D. (1994). Sexuality in patients treated for penile cancer: patients' experience and doctors' judgment. *British Journal of Urology*, *73*, 554–560.

Othmer, E., & Othmer, S. C. (1987). Effect of buspirone on sexual dysfunction in patients with generalized anxiety disorder. *Journal of Clinical Psychiatry, 48*, 201–203.

Patterson, D. G., & O'Gorman, E. C. (1989). Sexual anxiety in sexual dysfunction. *British Journal of Psychiatry, 155*, 374–378.

Peck, R. C. (1956). Psychological developments in the second half of life. In B. L. Neugarten (Ed.), *Middle age and aging* (pp. 88–92). Chicago: University of Chicago Press.

Persson, G. (1980). Sexuality in a 70-year-old urban population. *Journal of Psychosomatic Research, 24*, 335–342.

Pfeiffer, E., & Davis, G. C. (1972). Determinants of sexual behavior in middle and old age. *Journal of the American Geriatrics Society, 20*(4), 151–158.

Pfeiffer, E., Verwoerdt, A. & Davis, G. C. (1972). Sexual behavior in middle life. *American Journal of Psychiatry, 128*, 1262–1267.

Pfeiffer, E., Verwoerdt, A., & Wang, H. S. (1968). Sexual behavior in aged men and women. Observations on 254 community volunteers. *Archives of General Psychiatry, 19*, 753–758.

Pfeiffer, E., Verwoerdt, A., & Wang, H. S. (1969). The natural history of sexual behavior in a biologically advantaged group of aged individuals. *Journal of Gerontology, 24*, 193–198.

Pope, M., & Schulz, R. (1990). Sexual attitudes and behavior in mid-life and aging homosexual males. *Journal of Homosexuality, 20*(3-4), 169–177.

Radebold, H. (1992). *Psychodynamik und Psychotherapie Älterer*. Berlin: Springer-Verlag.

Radin, M. M. (1989). Preoedipal factors in relation to psychogenic inhibited sexual desire. *Journal of Sex and Marital Therapy, 15*(4), 255–268.

Rappoport, L. (1977). Aging as oppression. *Human Development*, *20*, 204–209.

Relf, M. V. (1991). Sexuality and the older bypass patient. *Geriatric Nursing, 12*(6), 294–296.

Renshaw, D. C. (1983). Sex, intimacy, and the older woman. *Women and Health, 8*, 43–54.

Renshaw, D. C. (1984). Geriatric sex problems. *Journal of Geriatric Psychiatry, 17*(2), 123–138.

Robinson, P. K. (1983). The sociological perspective. In R. B. Weg, (Ed.), *Sexuality in the later years* (pp. 81–103). New York: Academic Press.

Rosen, R. C., & Leiblum, S. R. (1988). A sexual scripting approach to problems of desire. In S. R. Leiblum & R. C. Rosen (Eds.), *Sexual desire disorders* (pp. 168–191). New York: Guilford Press.

Rosen, R. C., & Leiblum, S. R. (1989). Assessment and treatment of desire disorders. In S. R. Leiblum & R. C. Rosen (Eds.), *Principles and practice of sex therapy: Update for the 1990s* (2nd ed., pp. 19–47). New York: Guilford Press.

Rosen, R. C., & Leiblum, S. R. (Eds.) (1992a). *Erectile disorders: Assessment and treatment*. New York: Guilford Press.

Rosen, R. C., & Leiblum, S. R. (1992b). Erectile disorders: An overview of historical trends and clinical perspectives. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: Assessment and treatment* (pp. 3–26). New York: Guilford Press.

Rosenmayr, L. (1995). Eros und Liebe im Alter. In M. M. Baltes & L. Montada. *Produktives Leben im Alter* (pp. 258–289). Frankfurt, Germany: Campus.

Roughan, P. A., Kaiser, F. E., & Morley, J. E. (1993). Sexuality and the older women. *Clinics in Geriatric Medicine, 9*(1), 87–106.

Sandberg, G., & Quevillon, R. B. (1987). Dyspareunia: an integrated approach to assessment and diagnosis. *Journal of Family Practice, 24*(1), 66–70.

Sarrel, P. M. (1982). Sex problems after menopause: A study of fifty married couples treated in a sex counseling programme. *Maturitas, 4*, 231–237.

Sarrel, P. M. (1990). Sexuality and menopause. *Obstetrics and Gynecology, 75* (Suppl. 4), 26S–35S.

Scharff, D. E. (1988). An object relations approach to inhibited sexual desire. In S. R. Leiblum & R. C. Rosen (Eds.), *Sexual desire disorders* (pp. 45–74). New York: Guilford Press.

Schiavi, R. C. (1990). Sexuality and aging in men. *Annual Review of Sexual Research, 1*, 227–249.

Schiavi, R. C. (1992). Laboratory methods for evaluating erectile dysfunction. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: Assessment and treatment*. (pp. 141–170). New York: Guilford Press.

Schiavi, R. C., Mandeli, J. & Schreiner-Engel, P. (1994). Sexual satisfaction in healthy aging men. *Journal of Sex and Marital Therapy, 20*(1), 3–13.

Schiavi, R. C., Mandeli, J., Schreiner-Engel, P., & Chambers, A. (1991). Aging, sleep disorders, and male sexual function. *Biological Psychiatry, 30*, 15–24.

Schiavi, R. C., Schreiner-Engel, P., White, D., & Mandeli, J. (1991). The relationship between pituitary–gonadal function and sexual behavior in healthy aging men. *Psychosomatic Medicine, 53*, 363–374.

Schiavi, R. C., & Segraves, R. T. (1995). The biology of sexual function. *Psychiatric Clinics of North America: Clinical Sexuality, 18*(1), 7–23.

Schneider, H. D. (1980). Sexualverhalten in der zweiten Lebenshälfte: Ergebnisse sozialwissenschaftlicher Forschung. Stuttgart, Germany: Kohlhammer.

Schover, L. (1988a). *Sexuality and cancer: For the man who has cancer and his partner*. New York: American Cancer Society.

Schover, L. (1988b). *Sexuality and cancer: For the woman who has cancer and her partner*. New York: American Cancer Society.

Schover, L. R. (1989). Sexual problems in chronic illness. In S. R. Leiblum & R. C. Rosen (Eds.), *Principles and practice of sex therapy: Update for the 1990s* (2nd ed., pp. 319–351). New York: Guilford Press.

Schover, L. R. (1992). Erectile failure and chronic illness. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: Assessment and treatment* (pp. 341–367). New York: Guilford Press.

Schover, L. R., Evans, R. B., & Von Eschenbach, A. C. (1987). Sexual rehabilitation in a cancer center: Diagnosis and outcome in 384 consultations. *Archives of Sexual Behavior, 16*(6), 445–461.

Schover, L. R., Schain, W. S., & Montague, D. K. (1989). Sexual problems of patients with cancer. In V. T. DeVita, S. Hellman, & S. A. Rosenberg (Eds.), *Cancer: Principles and practice of oncology* (3rd ed., pp. 2206–2225). Philadelphia: Lippincott.

Schover, L. R., Yetman, R. J., Tuason, L. J., Meisler, E., Esselstyn, C. B., Hermann, R. E., Grundfest-Broniatowski, S., & Dowden, R. V. (1995). Partial mastectomy and breast reconstruction: A comparison of their effects on psychosocial adjustment, body image, and sexuality. *Cancer, 75*(1), 54–64.

Schreiner-Engel, P., & Schiavi, R. C. (1986). Lifetime psychopathology in individuals with low sexual desire. *The Journal of Nervous and Mental Disease, 174*(11), 646–651.

Segraves, R. T. (1988). Drugs and desire. In S. R. Leiblum & R. C. Rosen (Eds.), *Sexual desire disorders* (pp. 313–347). New York: Guilford Press.

Segraves, R. T., & Segraves, K. B. (1992). Aging and drug effects on male sexuality. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: Assessment and treatment*. (pp. 96–138). New York: Guilford Press.

Shell, J. A., & Smith, C. K. (1994). Sexuality and the older person with cancer. *Oncology Nursing Forum, 21*(3), 553–558.

Sjögren, K., Damber, J.-E., & Liliequist, B. (1983). Sexuality after stroke with hemiplegia. I: Aspects of sexual dysfunction. *Scandinavian Journal of Rehabilitation Medicine, 15*, 55–61.

Smyer, M. A., Zarit, S. H., & Qualls, S. H. (1990). Psychological intervention with the aging individual. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the Psychology of Aging* (3rd ed., pp. 375–403). San Diego, CA: Academic Press.

Sontag, S. (1972). The double standard of aging. *Saturday Review*, *55*, 29–38.

Speckens, A. E. M., Hengeveld, M. W., Lycklama à Nijeholt, G., van Hemert, A. M., & Hawton, K. E. (1995). Psychosexual functioning of partners of men with presumed non-organic erectile dysfunction: Cause or consequence of the disorder? *Archives of Sexual Behavior, 24*(2), 157–172.

Spence, S. H. (1992). Psychosexual dysfunction in the elderly. *Behaviour Change, 9*(2), 55–64.

Starr, B. D. (1985). Sexuality and aging. *Annual Review of*

Gerontology and Geriatrics, 5, 97–126.

Starr, B. D., & Weiner, M. B. (1981). *The Starr–Weiner report on sex and sexuality in the mature years*. New York: McGraw-Hill.

Staudinger, U. M., Freund, A. M., Linden, M., & Maas, I. (1996). Selbst, Persönlichkeit und Lebensgestaltung im Alter: Psychologische Widerstandsfähigkeit und Vulnerabilität. In K. U. Mayer & P. B. Baltes. *Die Berliner Altersstudie* (pp. 321–350). Berlin, Germany: Akademie Verlag.

Sviland, M. (1981). Sexuality and intimacy in later life. In: R. Kastenbaum (Ed.), *Old age in the new scene* (pp. 15–26). New York: Springer.

Talmadge, L. D., & Talmadge, W. C. (1986). Relational sexuality: An understanding of low sexual desire. *Journal of Sex and Marital Therapy, 12*(1), 3–21.

Thienhaus, O. J. (1988). Practical overview of sexual function and advancing age. *Geriatrics, 43*(8), 63–67.

Thomas, L. E. (1982). Sexuality and aging: Essential vitamin or popcorn? *The Gerontologist, 22*(3), 240–243.

Tiefer, L., & Melman, A. (1989). Comprehensive evaluation of erectile dysfunction and medical treatments. In S. R. Leiblum & R. C. Rosen (Eds.), *Principles and practice of sex therapy* (2nd ed., pp. 207–236). New York: Guilford Press.

Traupmann, J. (1984). Does sexuality fade over time? A look at the question and the answer. *Journal of Geriatric Psychiatry, 17*(2), 149–159.

Tümmers, H. (1976). Sozialpsychologische Aspekte der Sexualität im Alter.Cologne, Germany: Bühlau.

Turner, B. F., & Adams, C. G. (1983) . The sexuality of older women. In E. W. Markson (Ed.), *Older women—Issues and perspectives* (pp. 55–72). Lexington, KY: Lexington.

Vandereycken, W. (1987). On desire, excitement, and impotence in modern sex therapy. *Psychotherapy and Psychosomatics, 47*, 175–180.

Verwoerdt, A., Pfeiffer, F. & Wang, H. S. (1969a). Sexual behavior in senescence. I. Changes in sexual activity and interest of aging men and women. *Journal of Geriatric Psychiatry, 2*, 163–180.

Verwoerdt, A., Pfeiffer, F. & Wang, H. S. (1969b). Sexual behavior in senescence. II. Patterns of sexual activity and interest. *Geriatrics, 24*, 137–154.

Von Sydow, K. (1992a). Weibliche Sexualität im mittleren und höheren Erwachsenenalter: Übersicht über vorliegende Forschungsarbeiten. (Female sexuality in middle and old age: Research review). *Zeitschrift für Gerontologie, 25*, 113–127.

Von Sydow, K. (1992b). Eine Untersuchung zur weiblichen Sexualität im mittleren und höheren Erwachsenenalter (A study on female sexuality in middle and old age). *Zeitschrift für Gerontologie, 25*, 105–112.

Von Sydow, K. (1993). *Lebenslust: Weibliche Sexualität von der frühen Kindheit bis ins hohe Alter*. Bern, Switzerland: Huber.

Wakefield, J. C. (1987). The semantics of success: Do masturbation exercises lead to partner orgasm? *Journal of Sex and Marital Therapy, 13*(1), 3–14.

Weg, R. B. (Ed.) (1983). *Sexuality in the later years: Roles and behavior*. New York: Academic Press.

Weiler, S. J. (1981). Aging and sexuality and the myth of decline. In R. W. Fogel, E. Hatfield, S. B. Kiesler, & E. Shanas (Eds.), *Aging: Stability and change in the family* (pp. 317–327). New York: Academic Press.

Weiss, L. J. (1983). Intimacy and adaptation. In R. B. Weg (Ed.), *Sexuality in the later years: Roles and behavior* (pp. 147–166). New York: Academic Press.

Weiss, J. N., & Mellinger, B. C. (1990). Sexual dysfunction in elderly men. *Clinics in Geriatric Medicine, 6*(1), 185–196.

Weissbach-Rieger, A., & Franke, B. (1987a). Partnerbeziehung, Sexualität und Sexualverhalten bei älteren Frauen und Männern nach dem 55. Lebensjahr (1. + 2. Mitteilung). *Zeitschrift für Alternsforschung, 42*(4), 203–209.

Weissbach-Rieger, A., & Franke, B. (1987b). Partnerbeziehung, Sexualität und Sexualverhalten bei älteren Frauen und Männern nach dem 55. Lebensjahr (3. Mitteilung). *Zeitschrift für Alternsforschung, 42*(4), 211–213.

Weizmann, R., & Hart, J. (1987). Sexual behavior in healthy married elderly men. *Archives of Sexual Behavior, 16*(1), 39–44.

White, C. B. (1982). Sexual interest, attitudes, knowledge, and sexual history in relation to sexual behavior in the institutionalized aged. *Archives of Sexual Behavior, 11*(1), 11–21.

Wickert, J. (1983). *Partnerschaft im Alter. Auswertung der Ergebnisse einer Längsschnittstudie* (Berichte aus dem Institut für Psychologie). Cologne, Germany: Universität Köln, Erziehungswissenschaftliche Fakultät.

Winn, R. L., & Newton, N. (1982). Sexuality in aging: A study of 106 cultures. *Archives of Sexual Behavior, 11*(4), 283–298.

Wise, T. N. (1989). Sexuality in the aging individual: A biopsychosocial perspective. *Issues in the Geriatric Psychiatry, 19*, 53–66.

Wise, T. N., Epstein, S., & Ross, R. (1992). Sexual issues in the medically ill and aging. *Psychiatric Medicine, 10*(2), 169–180.

Zemel, P. (1988). Sexual dysfunction in the diabetic patient with hypertension. *The American Journal of Cardiology, 61*, 27–33.

Zilbergeld, B. (1992). The man behind the broken penis: Social and psychological determinants of erectile failure. In R. C. Rosen & S. R. Leiblum (Eds.), *Erectile disorders: Assessment and treatment* (pp. 27–51). New York: Guilford Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.05
Bereavement

DALE A. LUND
University of Utah, Salt Lake City, UT, USA

7.05.1 INTRODUCTION 95
7.05.2 LIFELONG BEREAVEMENT EXPERIENCES 96
7.05.3 BEREAVEMENT IN A MULTIDIMENSIONAL CONTEXT 97
7.05.4 EXPERTS ON BEREAVEMENT 98
7.05.5 THEORIES ABOUT BEREAVEMENT 100
  7.05.5.1 Psychoanalytic Theories 100
  7.05.5.2 Stress, Appraisal, and Coping Theories 101
  7.05.5.3 Stage and Task Models 101
  7.05.5.4 Other Theories 102
7.05.6 FEATURES OF THE BEREAVEMENT PROCESS 103
7.05.7 PREDICTORS OF BEREAVEMENT ADJUSTMENTS 105
7.05.8 WORKING WITH THE BEREAVED 108
7.05.9 SUMMARY 110
7.05.10 REFERENCES 110

7.05.1 INTRODUCTION

Issues about death surround everyone's daily lives, yet most people neglect to use these opportunities to improve their understanding of grief. A plane carrying hundreds of passengers crashes in the Florida swamps and Americans watch with interest on television and read in newspapers and magazines about the loss of lives, grieving families, heroic workers trying to retrieve bodies, and the need for improved safety standards in the airline industry. Another assisted suicide is described in the media but very few people ever discuss the grief of the family members. The evening news regularly reports on the victims of gang-related murders that occurred the previous night. Local newspaper obituaries include another person who has died from AIDS. Each of these four examples contain in-depth stories of human emotion and relationships in the lives of the victims and survivors, yet most passive onlookers fail to realize that they will share many of the same emotional and social experiences that these distant "people in the news" have had to encounter. While most people never experience horrendous or public deaths nearly everyone will have several personal bereavement experiences during their lives. Individuals could benefit from learning more about death and grief by more closely examining some of these news events and people and assessing their own personal feelings and reactions. Instead, most people only learn about death, grief, loss, and bereavement when they have no other choice because it is now their own personal experience.

In order better to describe the complexities of the bereavement process and how we might best assist those who have the greatest difficulty, this chapter will summarize and highlight some of the most basic and essential knowledge available. The chapter begins with a review and discussion of the many common but highly diverse bereavement experiences that occur throughout the life course. Although the primary focus is on bereavement in later life it is necessary to place these losses in the context of the entire life course because earlier experiences continue to impact coping strategies used in late life. Bereavement is also examined in a multidimensional context because it can affect nearly every aspect of a person's life. Knowledge about grief and bereavement comes from a variety of experts and sometimes the information is contradictory or inconsistent, so it is important to understand the advantages and limitations from these diverse sources. This section is presented along with a discussion of the benefits from integrating theory, research, and clinical practice to achieve a better understanding of the bereavement process. Reviewed next are the most common features of the long-term process and predictors of who is most likely to have the greatest difficulty. The chapter ends with a review of the most promising ways to intervene to assist those in greatest need, and some additional comments about the need to be creative and collaborate with others from a variety of different disciplines.

Many people use the terms bereavement, grief, and mourning as though they were interchangeable or have the same meaning. This is not the case. All three of these terms are highly interrelated but it is important to note that each has a unique meaning.

Bereavement is the broadest or most general of the three terms because it refers to both the situation and the long-term process of adjusting to the death of someone to whom a person feels close (Lund, 1993). The term also has been used in reference to losses other than death, such as the loss of a home or personal belongings because of a fire or a natural disaster. It is more common, however, to use the term bereavement in situations following the death of a friend or relative, or even an animal companion. Grief and mourning are more specific in their meaning because they refer to different aspects of the overall bereavement process. Grief refers specifically to the affective or emotional responses that people often experience during the early phases of the bereavement process. These feelings frequently include disbelief, confusion, shock, numbness, sadness, and sometimes anger and guilt. Feeling abandoned, lonely, or depressed, and being preoccupied with thoughts of the deceased person and the events surrounding the death are often components of grief (Parkes, 1987).

It is common for bereaved persons to feel a wide range of emotions that come and go quickly and then reappear many months or even years later. These grief-related feelings do not come as a set package to all bereaved persons. Each person experiences a relatively unique array of emotions with varying levels of intensity and duration. Culture plays a major role in shaping how people experience and express bereavement because it provides the beliefs, values, and norms that guide behavior (Lund, 1989b). Most professionals use the term mourning to refer to the commonly accepted or culturally patterned ways people behave and express themselves during the adjustment process (Lund, 1993). Mourning refers to such behaviors as crying; missing work; altering daily routines; attending funerals, cremations, and burials; following standards for dress and appearance; and participating in other rituals that are influenced by the surrounding culture. Each culture has traditions to determine the most appropriate ways to behave (Chigier, 1988; Stroebe & Stroebe, 1987). Some cultures avoid open displays of sadness yet others expect the bereaved to be vocal and demonstrative (Irish, Lundquist, & Nelsen, 1993). Each of these three related terms will be further clarified later in this chapter along with discussion of the many commonalities and diverse ways that people experience bereavement.

7.05.2 LIFELONG BEREAVEMENT EXPERIENCES

Throughout the life course nearly everyone experiences many kinds of losses that require them to make adjustments. Infants become temporarily separated from parents and siblings when they leave the familiar surroundings of the home to attend day care and school. With increasing numbers of dual-career couples, children are leaving their homes at much younger ages to spend much of their weekday time away from both parents. Children also lose toys, clothes, pets, and valued belongings. With increased geographic mobility it is becoming common for children to move several times from one neighborhood to another, changing friends and familiar surroundings with each move. Young adults have lost relationships with teachers, clubs, organizations, and peers, as they advance from elementary grades, through high school and college. In later life adults continue to lose more friendships and relationships because of relocations, deaths, and retirement. With divorce rates near 50% in the USA (Hess, Markson, & Stein, 1996), millions are experiencing the loss of marital relationships, parents, and family life. Due to physiological changes associated with aging, people often lose hair, hearing, sight, muscle tone, independence, and mobility. In short, people who are fortunate to live long lives will experience many significant losses. Each loss has the potential of disrupting patterns of behavior, initiating grief-related feelings and requiring adjustments, but also adding to the course of human development.

Older adults have many opportunities to develop their coping skills or become overwhelmed by life's cumulative losses. It is much easier to understand how people cope with their current bereavement situations if one recognizes how they have managed their previous losses. For example, in the case of older bereaved spouses, effective interventions need to take into account how they experienced relocation, physical, and relationship losses earlier in their life (Lund, 1989a). If they relied on task-oriented strategies in handling these previous situations then a similar approach might be effective again. At the same time we need to appreciate the uniqueness of each loss or bereavement experience and not assume that the loss of personal belongings is the same as a divorce or death of a parent. While sadness might be a grief response common to both relocation and the death of a parent, it is unlikely that anger or guilt would be equally present in both situations (Farnsworth, Pett, & Lund, 1989).

The field of gerontology has been redefined since the late 1970s to give greater recognition to aging as a lifelong process (Hooyman & Kiyak, 1996). Rather than being exclusively focused on older adults, it recognizes that aging occurs throughout the life course but the primary focus still remains on later life. The most common death-related losses of interest to gerontologists are the deaths of parents, spouses, siblings, and friends. Much of the information in this chapter is based on spousal bereavement in later life because more gerontological research has focused on this particular loss. Children's death-related bereavement experiences are most likely to result from the death of a neighbor, parent, older relative, or school classmate. Some experts have described the death of a parent as the loss of our past, the death of a spouse or sibling as the loss of the present, and the death of a child as a loss of part of our future (Leming & Dickinson, 1990; Schiff 1977; Staudacher, 1987). Again, early and consistent experiences with death-related losses can shape subsequent attitudes and ways of coping well into later life (Lund, 1993). For example, the young boy who is told to be strong and not to cry after his father's death might easily grow up to be the unexpressive and isolated widower in later life.

Although statistics cannot capture the full essence of the bereavement process they are very helpful in documenting the commonality and expectedness of the experience. Each year approximately eight million Americans experience the death of an immediate family member. Even though suicides are underreported in the USA roughly 27 000 are reported each year. Over 400 000 of the deaths are children and youths under the age of 25. Also, each year there are at least 800 000 new widows and widowers (Osterweis, Solomon, & Green, 1984).

The death of a spouse is an increasingly common experience among aging adults, particularly women. The 1990 Census revealed that only 1% of the population between ages 18 and 54 were widowed, but 34% of those 65 and over and 67% of those 85 and over were widowed. A gender difference becomes increasingly apparent in the later years of life. After age 65, nearly 49% of the women but only 14% of men were widowed. After age 85, 80% of women and 43% of men were widowed (US Bureau of the Census, 1991). These high percentages actually underestimate the number of persons who experience the death of a spouse because they do not include those who remarried. In short, death-related bereavement is a common and expected experience for older adults.

7.05.3 BEREAVEMENT IN A MULTIDIMENSIONAL CONTEXT

One of the major reasons why bereavement has become such an important topic to clinicians, researchers, health professionals, clergy, and educators is the growing recognition that the adjustment process can impact virtually every aspect of a person's life (Lund, 1989b; Shuchter, 1986; Shuchter & Zisook, 1993). This pervasive and multidimensional nature of bereavement means that the coping process can affect emotions, spirituality, identity, social interactions and relationships, sexuality and intimacy, work productivity, financial stability, health, and even death. Clinicians and others who are interested in providing assistance to the bereaved need to recognize the broad range of possible effects and determine the extent to which bereavement is the cause of the difficulty. For example, patterns of interaction in family life can be radically altered when someone dies. Family meals, and other daily rituals related to getting ready for work, school, or church are affected by the absence of the deceased family member. Worden (1991) presents this difficulty as one of four tasks of mourning. He refers to this as the need to adjust to an environment in which the deceased is missing. These tasks of mourning are discussed in more detail later in this chapter. In the case of spousal bereavement survivors often change their self-concept to an "uncoupled identity" because they now carry out their daily life alone rather than as part of a couple (Lund, Caserta, Dimond, & Gray, 1986). Our society places great importance on adults doing many activities as couples rather than doing them alone. Imagine the difficulty of a bereaved spouse going alone to a movie or restaurant after many years of always doing these things with their husband or wife. Bereaved persons may also lose their appetite and motivation to remain physically active so nutrition, exercise, and physical well-being can be essential parts of the coping process (Lund, 1989a).

The multidimensional nature of bereavement is evident in the topics that professionals from different disciplines choose for special attention. Sociologists often examine bereavement's impact on family life, interactions, life satisfaction, and the role of culture and socioeconomic characteristics in shaping these bereavement outcomes. Finding differences among men and women, rich and poor, and diverse ethnic groups are common concerns for sociologists. Psychologists frequently examine the emotions of grief, depression, morale, and identity, and how previous relationships and the personality of the bereaved influence their coping strategies. Social workers are likely to assess family functioning and the need for, and effectiveness of, specific community services. Nurses and other health professionals are often concerned about the mental and physical well-being of bereaved persons. Economists are usually interested in the financial stability of bereaved persons and their families. Because many women in the USA do not have the same level of retirement and other financial benefits as men, they are much more likely to suffer economically during widowhood (Caserta, Lund, & Dimond, 1989). Each discipline and professional group adds a unique focus to examining bereavement experiences, and collectively, we better understand the many ways that bereavement can affect our lives. Clinicians interested in further reading about the many dimensions of grief and bereavement should read Shuchter (1986), Worden (1991), and Stroebe, Stroebe, and Hansson (1993b). These references also include recommendations based on international research studies.

The multidimensional nature of bereavement cannot be understood without an awareness that bereavement takes place within a broad context that varies from person to person. Although there are many common reactions and feelings that most people experience during the course of bereavement, there also are many unique experiences that may not be widely shared. For example, some bereaved persons feel intense guilt and anger while others feel a general sense of relief because the deceased is no longer suffering pain and discomfort. Some are overwhelmed with sadness and immobilized by depression while others confront the loss as a challenge and opportunity for growth. In order to understand why one person manages very well and another never seems to adjust to the loss, it is necessary to examine the multidimensional nature of the context in which bereavement occurs (Cleiren, 1993; Lund, 1989b). Previous life experiences, personality traits, social support received from others, circumstances surrounding the death, and a person's competency in managing the many tasks of daily living are some of the many factors that comprise the context of bereavement. Health professionals and clinicians who work with bereaved persons are better able to provide assistance when they know more about the specific aspects of each individual's bereavement context. What helps one person may not be of much value to another.

7.05.4 EXPERTS ON BEREAVEMENT

For thousands of years the subject of death has received considerable thought and discussion but much less attention has been devoted to the subject of how death affects those who are left to grieve (Stephenson, 1985). Since the Second World War, however, we have seen a steady increase in interest and writing about bereavement. Various kinds of experts have emerged during this time, each relying on different experiences, training, and sources of knowledge. Academic disciplines, especially psychology, sociology, social work, nursing, and other behavioral and health professions have contributed attention and knowledge to the subject through the development of theory and research (Stroebe et al., 1993b). Clinicians who work directly with the bereaved have added their knowledge about ways to help those who are having the greatest difficulty (Rando, 1993; Worden, 1991). Religious leaders and clergy from various denominations have shared their experiences regarding spiritual aspects of the bereavement process (Grollman, 1977; Kushner, 1982). Local bookstores are now replete with numerous autobiographical reports written by bereaved persons who have had experiences that they want to share (Caine, 1974; Lewis, 1961). These five sources of expertise and knowledge—theory, research, clinical reports, religious views, and autobiographical accounts—each make unique contributions to our understanding of bereavement. These different sources of knowledge, however, have unique advantages and limitations that need to be recognized because they can lead to different and competing conclusions.

Some bereaved persons will automatically turn to their religious leaders or their religious beliefs for comfort and guidance. Most religions have a set of beliefs regarding the questions of ultimate meaning. What is the purpose of life? Does God or a supernatural being have a plan for each specific person? Is there special meaning in the death of the loved one? Because most religions offer answers to these and other questions, religiously oriented people can find predictability, structure, and meaning to help them during the upheaval of bereavement. Research has shown that religion and spirituality are only part of the bereavement context and that the long-term adjustment process involves much more than a religious code to follow (Faletti, Gibbs, Clark, Pruchno, & Berman, 1989; Gass, 1989a; Lund, Caserta, & Dimond, 1988). In fact, some bereaved persons report disappointment that their religious beliefs did not provide greater comfort and relief. Research has not confirmed that religious beliefs and practices are universally helpful to bereaved persons but we do know that many people turn to their spiritual views as a source of knowledge and assistance (Lund, 1989a).

Personal accounts written by bereaved persons are abundant because bereaved persons want to know how others have managed, if their experiences are unique or common, and how they can cope better with their loss. These autobiographical accounts can be useful because they offer detailed descriptions of feelings, problems, and solutions that worked for the authors. Some bereaved persons, for example, might find comfort in knowing that their feelings of anger and guilt were shared by the author and may benefit by following some of the suggestions in the book. Kushner's (1982) *When Bad Things Happen to Good People* has continued for many years to be on the best-sellers list because bereaved readers often find comfort in his expressions of anger, sadness, and even in some disappointment in his previous long-standing religious beliefs. Caution should be exercised, however, in fully relying on these autobiographical accounts because the authors' experiences may not be widely shared and their credentials are most likely in journalism rather than bereavement.

Therapists, counselors, psychologists, psychiatrists, and other clinicians have provided valuable information and insights about bereavement. Like autobiographical accounts, clinical reports often offer detailed personal descriptions of the bereavement process and suggestions for helping to alleviate some of the most difficult problems. These in-depth reports about individuals' bereavement experiences are not only interesting but they help to organize experiences of more people, identify serious but common adjustment difficulties, and inform other professionals and bereaved persons of specific intervention strategies and their potential benefits. Knowledge that comes from clinical practice, however, must be interpreted in a way that recognizes the uniqueness of the clients or patients. Those who receive counseling or professional clinical services are likely to differ in some important ways from those who never seek help (Caserta & Lund, 1992a). Bereaved persons who receive these services are more likely to be experiencing the greatest difficulty in adjusting, have less social support available to them, have financial resources adequate to pay for the services, and have a tendency to rely on others for help (Caserta & Lund, 1992a). Knowledge about bereavement based solely on information from those who sought help from clinicians does not include the experiences of the majority of bereaved persons (Lund, Dimond, & Juretich, 1985; Stroebe & Stroebe, 1987). It is equally important to know how people manage their grief on their own and to what extent they want and need help even though they do not seek professional assistance.

Theory and research have been much slower in adding to our knowledge about bereavement but their contributions have been essential. Freud (1957) was one of the earliest to theorize about bereavement. In the early 1900s Freud practiced and wrote extensively about psychoanalytic theory and applied it to understanding the feelings and thoughts that follow separation or loss of attachment. His work was followed by other theorists and eventually systematic research studies were completed. Lindemann (1944) also used a psychoanalytic framework to guide perhaps the first bereavement study in the 1940s. His study interviewed over 100 relatives of 13 people killed in a fire at the Coconut Grove night club in Boston. Research and theory related to bereavement has steadily increased since these early beginnings (Stroebe, Stroebe, & Hansson, 1993a). Numerous theories and research findings have emerged that both support and challenge what we initially believed to be true about bereavement. These early theories and more recent conceptual frameworks are discussed in greater detail later in this chapter.

Research has generated knowledge on a much broader range of bereaved persons. By collecting data from people who have not sought professional help, and including persons from varied cultures and social classes, we have learned much more about what is common and what is unique during bereavement (Caserta & Lund, 1992a). Systematic research studies, using well-developed and standardized scales, representative samples of bereaved persons, and nonbereaved control groups, have revealed findings that are inconsistent with and challenge some of the previous views based upon the other sources of knowledge (Stroebe & Stroebe, 1991; Wortman & Silver, 1989). These research findings are presented in more detail in Section 7.05.7. One of the most prominent research findings is that while bereavement is a highly stressful experience and some people report never getting over it, the majority of bereaved persons are very resilient, especially older bereaved spouses. Many bereaved persons find personal and social resources to meet their needs and manage their adjustment process. Research has also not been able to confirm that people go through specific stages as suggested by many clinicians and some theorists (Lund, Caserta, & Dimond, 1986; Shuchter & Zisook, 1993; Wortman, Silver, & Kessler, 1993). The process is much more varied and complex than the uniform and orderly course depicted in the stage theories (Bowlby, 1969). Research has revealed that not all people need to express in equal amounts their emotions outwardly to others in order to avoid a build-up of problems that will come out even worse later (Wortman & Silver, 1989). Many people who appear in control and able to manage their early grief tend to be doing very well years later (Wortman et al., 1993). Conversely, those who were doing poorly and were most expressive of their difficulties in the first several weeks were still having great difficulty two years later. These findings appear to be similar for adults of all ages (Lund, Dimond, et al., 1986; Wortman et al., 1993). In short, all five types of experts and their unique experiences and knowledge have added to a more comprehensive and useful understanding of bereavement.

7.05.5 THEORIES ABOUT BEREAVEMENT

The bereavement process is multidimensional and requires professionals from different disciplines to contribute their specific theories of human behavior to our understanding, explanation, and prediction of how people will experience this complicated process. Various conceptual frameworks have been applied to grief and bereavement, each providing a unique focus on specific parts of the bereavement context. These theories are particularly valuable if they predict adjustments and guide the development of interventions to help those in greatest need. The following sections briefly summarize some of the most prominent theories about grief and bereavement.

7.05.5.1 Psychoanalytic Theories

Most of the variations in psychoanalytic theory that have applied to bereavement emphasize the importance of recognizing the conscious and subconscious ties or attachments that the bereaved person formed with the deceased (Cleiren, 1993; Parkes, 1987; Stroebe et al., 1993a). According to this view the bereaved person must express rather than repress all of the feelings associated with the detachment or loss. Anger, guilt, fear, and love, when expressed by the bereaved, help lead to normal instead of abnormal grief. These theories also stress the importance of early childhood relationships and separations because these first experiences remain throughout life. When children have traumatic losses in early life they often form ambivalent relationships later, repress their anxieties, and have greater difficulty with grief and bereavement. Separation anxiety is seen as an instinctive reaction to loss but it is essential that the bereaved person withdraw from the deceased and form new attachments.

Freud (1957), Lindemann (1944), Bowlby (1969), Parkes (1987), and Marris (1958) are the most prominent contributors to the application of psychoanalytic theory to bereavement (Cleiren, 1993). While all theorists have their own unique interpretations they have generally agreed that grief is largely an intrapsychic process that each bereaved person must cognitively resolve in order to avoid delayed, distorted, or abnormal grief. Some of the abnormal reactions include social withdrawal, hypochondriacal symptoms, psychosomatic illnesses, and hyperactivity. Normal grief is thought to be marked by distress and impaired functioning but recovery is to be expected (Parkes, 1987). Bereaved persons usually experience a period of upset, protest, and anxiety because of the separation. They are likely to search for the deceased in familiar settings. Despair and depression can follow because the deceased person does not return. A reorganization phase occurs when grievers accept the reality of the loss, resolve the feelings of guilt and anger, restructure their lives and return to normal functioning (Parkes, 1987; Worden, 1991). These theories suggest that bereavement follows a relatively ordered progression of phases and that unresolved grief occurs when the sequence is disrupted or distorted.

7.05.5.2 Stress, Appraisal, and Coping Theories

Theories of stress, appraisal, and coping have been applied to bereavement because the death of a spouse, parent, child, or sibling is generally considered to be among the most stressful events that can ever happen. Stress exists when the demands on the individual exceed the resources readily available to meet the demands. The death of a loved one, particularly when the deceased played an important part in the daily life of the bereaved, creates stress in many ways for the survivor. Daily life is radically changed because the bereaved can no longer rely on the deceased for on-going companionship, conversation, security, guidance, comfort, love, sharing, and listening; and instrumental assistance such as financial support, transportation, and caregiving. Bereavement requires the survivor to change and adjust to an environment where the deceased is absent (Worden, 1991).

When bereavement occurs people must cognitively appraise their situation, usually consistent with their previously established patterns of dealing with stress. Cognitive appraisal means that people consciously assess their loss, try to make sense out of it, and determine the consequences that are likely to result from different courses of action. The degree of stress and subsequent adjustment difficulties is largely dependent on the bereaved person's appraisal of the situation and the coping strategies used.

The most common types of appraisal are pain, threat, challenge, and relief (Gass, 1989b). Those who primarily see the pain often focus their attention on the damage already done and expend much of their energy expressing anger, sadness, and depression, and their coping strategy often involves withdrawal, self-blame, and feeling overwhelmed. Those who see their loss as a threat are aware of the stress but are uncertain about their ability to deal with it. The death of their loved one means that they see other losses that are likely to follow. Their coping strategies usually include fear and worry about future threats to their lives. Rather than focusing their energy in broader ways to make their adjustment easier, they often engage in wishful thinking. The bereaved who see their loss as a challenge recognize the stress and the need to adapt. They assess their strengths, limitations, and needs and pursue a course of action that will allow them to learn and possibly grow from the experience. Having pride, self-confidence, motivation, and some skills in carrying out the many tasks of daily living helps the bereaved to define the situation as a challenge and follow a more problem-focused approach to adjustment (Lund, Caserta, Dimond & Shaffer, 1989). Lazarus (1966), Folkman, Lazarus, Dunkel-Schetter, De Longis, and Gruen (1986), Pearlin and Mullan (1992), and many others have advanced theory and research on stress, appraisal, and coping, and the perspective continues to make very important contributions to understanding bereavement.

7.05.5.3 Stage and Task Models

Ever since Kubler-Ross (1969) proposed the well-known stage model of the dying process, it has been applied to many discussions of how people deal with problems and life crises. The dying person was seen to pass through a naturally occurring sequence of experiences beginning with denial and progressing through anger, bargaining, depression, and eventually culminating in acceptance. Because bereaved persons and many professionals became familiar with the stage model developed by Kubler-Ross and they could see many similar emotions, thoughts, and issues expressed during the course of bereavement, it seemed only natural to apply the sequential and progressive stages to bereavement.

While research has identified and confirmed that bereaved persons experience many of the same kinds of emotions, thoughts, and behaviors presented in the stage models (Aiken, 1994; Hardt, 1978; Parkes, 1987), the overall bereavement process appears to be much less orderly, sequential, and uniform than these models suggest (Lund, Caserta, & Dimond, 1986; Wortman et al., 1993). For example, some bereaved persons report many years after the loss that they sometimes feel the sadness, pain, and regret that they experienced in the first several weeks (Lund, 1989a). The stage models have contributed to our knowledge about bereavement because they have helped to identify parts of the complex process and provided a foundation for further study.

Similar to the stage models and building on some of the concepts from other theories is the idea that adaptation requires the bereaved person to accomplish or complete specific tasks. Worden's (1991) "tasks of mourning" approach is an example. His work is based on the assumption that grief is work, requiring both emotional and physical energy to complete. He suggests that it is important for the bereaved eventually to accept the reality of the loss, to experience the pain of grief, to adapt emotionally and socially to an environment in which the deceased is missing, and to relocate psychologically the deceased person in their life and move forward. The notion of "relocating the deceased" recognizes the value of keeping memories alive and not assuming that the bereaved should forget the importance of their past relationships. While many other specific tasks can be added to his model, it does provide a very useful approach in helping bereaved persons. The task-oriented approach highlights the importance of being active and learning new skills rather than simply being passive and allowing time to pass and heal the wounds of grief. Worden's book is highly recommended reading for clinicians and researchers.

7.05.5.4 Other Theories

A wide range of theories from social, behavioral, and health sciences have been used to understand specific aspects of grief and bereavement but most of them have not been systematically tested in research. Some of the theories have received only limited attention even though they have considerable promise in explaining and predicting bereavement behaviors and helping bereaved persons through the process. For example, family and systems theories, symbolic interactionism, the health belief model, and the theory of reasoned action have much to add but have not been extensively used in research or practice.

Family and systems theories (Olson, Sprenkle, & Russell, 1979) are relevant to bereavement because bereaved people are members of nuclear and extended families and their lives take place within functioning systems. People do not live in total isolation from families, groups, and organizations, and their experiences within these units are important parts of the broader bereavement context. An 80-year-old widow is impacted by how other family members grieve the death of her husband, who also was a father, grandfather, uncle, and brother. Just like individuals, families grieve. Some family members provide strength and support to one another and other families create more conflict and tension during bereavement (Worden, 1991). Systems theory recognizes that people are parts of larger systems and individual lives consist of various functioning parts as well. Anger and depression following bereavement can easily carry over from the home to the workplace and many other social settings, thereby impacting and further complicating other relationships.

Symbolic interaction theory, which comes from the discipline of sociology, has much to offer the topic of bereavement (Lund, Caserta, Dimond, & Gray, 1986). This theory emphasizes the importance of social interactions with others and how they shape a person's self-concept or identity which in turn, influence behavior. Bereavement losses have the potential of altering a person's interactions and self-concepts and ultimately bereavement adjustments (Worden, 1991). From the field of health education, the health belief model (Rosenstock, 1990) and theory of reasoned action (Ajzen & Fishbein, 1980) are examples of conceptual frameworks relevant to bereavement. Both are applicable to examining the way bereaved persons make decisions and what they take into account to decide a course of action. Bereavement can very easily lead to neglect of one's own health. These approaches suggest that a broad range of factors including messages in the media can influence bereavement behaviors by identifying healthy behaviors to pursue and stressing the likelihood of positive outcomes that will follow. Grief requires considerable energy so good nutrition and regular exercise can be very helpful during the bereavement process (Lund, 1989a).

Finally, the most prominent theories in the field of gerontology have relevancy to understanding bereavement but they have not received much attention. Disengagement (Cumming & Henry, 1961), activity (Havighurst, 1968), continuity (Atchley, 1989), and age stratification (Riley, 1971) theories have contributions to make and hopefully, in the future, they will be examined and tested in research. Disengagement theory would suggest that those who were already gradually withdrawing from society will have less difficulty during bereavement than those who have not yet begun the process of withdrawal. The bereavement experience would not seem as abrupt to them. Activity theory would argue in favor of maintaining relatively high levels of physical activity during bereavement because it will help create greater life satisfaction and well-being. Continuity theory would suggest that the best course of action during bereavement would be to continue doing most of the things that one had always enjoyed. Age stratification theory could be used for a better understanding of why age might influence bereavement adjustments and that people in different age cohorts have different needs during stressful life transitions. An excellent reference resource for more information about these and other gerontological theories is *Social Gerontology* (Hooyman & Kiyak, 1996). In short, theory helps to inform and guide research and we need considerably more of both theory and research to improve our knowledge about grief and bereavement.

7.05.6 FEATURES OF THE BEREAVEMENT PROCESS

Research on bereavement, and particularly studies which examine predictors of outcomes, has considerably increased since the early 1980s. Some research investigators have waited many weeks and months to interview bereaved persons and others have obtained information only days after the death occurred (Stroebe & Stroebe, 1987). Nearly all of the early research studies avoided contacting bereaved persons during the first six months of the process, assuming that the bereaved persons would be too upset to participate. Later studies revealed that was not necessarily the case and by waiting too long much valuable information was lost (Lund, 1989b). Some studies followed bereaved persons for several years to examine long-term outcomes but most projects obtained information at only one point in time, providing only a limited picture of the process (Stroebe et al., 1993b). These cross-sectional studies did not adequately capture the many dynamic aspects of the process.

Only few investigators have included nonbereaved control groups, even though it is valuable to be able to compare the thoughts, feelings, and behaviors of bereaved persons with those who are not experiencing death-related bereavement (Lund, 1989a; Stroebe, Hansson, & Stroebe, 1993). In fact, nonbereaved control groups are needed to help determine what effects are most likely attributed to bereavement rather than pre-existing conditions. The ideal study design would include before and after bereavement measures to assess the effects of bereavement most directly, but these studies are expensive and lengthy. Research has also included more women than men, more adults than children, and more Caucasians than non-Caucasians. Therefore, much of what we have learned from research has been limited to specific populations. Even with these and other limitations, research has improved our understanding about grief and bereavement. It is not possible to review all or even most of these findings but five of the most important general conclusions are presented here.

First, bereavement is very stressful but many people manage their losses with considerable resiliency. In the case of spousal bereavement, the majority of research participants (72%) reported that their loss situation was the worst and most stressful thing that had ever happened to them (Lund, 1989b). The death of a child also creates stress and is particularly difficult because children are expected to live longer than their parents (Rubin, 1993). Among adults, it is disturbing when both parents are deceased because it means that their own generation in the family is now the most senior. The death of a spouse, parent, or child almost always brings out thoughts of mortality and death among the survivors. In Western culture, thoughts about the inevitability of death are usually troubling because, as mentioned in the introduction to this chapter, we have not openly discussed death. Some researchers have attempted to identify how many people continue for several years having major difficulties that interfere with their daily lives. It has been estimated that between 15 and 25% will need some kind of professional help with their bereavement adjustments (Lund, 1989a; Lund, Dimond, et al., 1986).

Even though bereavement is extremely stressful, it is somewhat surprising to know that many people are very resilient and able to restructure their lives in ways that allow them to find enjoyment again (Caserta & Lund, 1992b; Cleiren, 1993; McCrae & Costa, 1993; Murrell, Himmelfarb, & Phifer, 1988). One study (Caserta & Lund, 1992b) found that older bereaved spouses rated their stress levels in the range of 5.0–5.5 (on a 1–7 scale) during the first two years of bereavement. However, the currently married, nonbereaved control subjects indicated that their expected stress levels would range from 5.8 to 6.0 if they were in that same spousal bereavement situation (Caserta & Lund, 1992b). These data suggest that it is possible that many older adults will cope more effectively with bereavement than they expect. These research findings also help direct our attention to those 15–25% who are most distressed and unable to manage their bereavement losses alone and also recognize that not all bereaved persons will need professional assistance (Murrell et al., 1988; Wortman & Silver, 1989). These issues are discussed in further detail in the next section of this chapter.

Second, there is considerable diversity among and within bereaved persons (Levy, Martinkowski, & Derby, 1994). Not all people experience the same feelings, thoughts, and actions as they move through the process. For example, a 65-year-old widow in a research project described herself in the following way four weeks after her husband's death: "I am a very lonely person. Lousy, all washed out, despondent, feel deserted, angry, hurt, hopeless, alone, mixed-up, cannot concentrate, very emotional, very tired, cry a lot, hateful, very bitter, misfit, nobody, very miserable, very much of a loner." This same woman described herself in much the same troubled way two years into the bereavement process. She added that she was concerned about herself, and really hurt inside. She noted that she was sick of living alone and didn't care about life. In contrast, another 69-year-old widow described herself at six months after her husband's death as being "independent, excited about keeping busy, enjoying the company of others, and working in my yard and doing handiwork." Similar diversity in bereavement reactions is also found within each individual (Cleiren, 1993; Lund, Caserta & Dimond, 1993). It is not unusual to find a person simultaneously experiencing a very broad range of feelings and behaviors. For example, bereaved persons can feel angry, guilty, and lonely, yet at the same time feel personal strength and pride in how they are coping. A 70-year-old woman described herself at length as being busy, enjoying many different activities, and doing things with other people, but her final self-descriptive comment was that she still feels lonely. It is common for bereaved persons to experience simultaneously both positive and negative feelings (Lund, Caserta, & Dimond, 1986). For these reasons, it is difficult accurately to describe a bereaved person as doing very well or coping poorly. It is much more accurate to identify specific aspects or dimensions of people's lives where they are in greatest need for assistance.

Third, a theme emerging from recent studies is that of being more cautious in using the terms "normal" and "abnormal" to describe grief or bereavement experiences (Lund, 1993). Individuals may experience and express themselves in a variety of ways during bereavement. Research has shown that what would have been considered outside the realm of "normal" grief several years ago may now be considered a normal part of the process. For example, talking to a deceased spouse was previously thought to be abnormal or even pathological, but recent research findings show that this is actually a very common practice among bereaved spouses, and it is now often encouraged as a therapeutic technique related to reminiscence or relocating the deceased in the life of the bereaved (Worden, 1991). When we learn from a community-based sample of bereaved research subjects that they regularly talk out loud to their deceased spouses and appear to be coping relatively well, it suggests that the behavior is not nearly so problematic (Lund, 1989b).

Rather than applying rigid labels such as "normal" or "abnormal," it is more appropriate to use terms such as "common" or "uncommon" to describe bereavement feelings and behaviors (Lund, 1996). Bereaved persons are far less stigmatized by being told that they are experiencing some uncommon feelings than by being told, "You are abnormal." In fact, Western culture values people for being unique and different from others. Clinicians, in particular, need to be cautious not to assume that being different is abnormal.

Fourth, recent studies have revealed that the bereavement process is more like a rollercoaster ride than an orderly progression of stages with clear time frames associated with each stage (Lund, 1996; Lund et al., 1993). The bereavement roller coaster is characterized by the rapidly changing emotions of grief; meeting the challenges of learning new skills; recognizing personal weaknesses and limitations; developing new patterns of behavior; experiencing fatigue, loneliness, and helplessness; and forming new friendships and relationships.

Research has been unable to identify clear time markers associated with these many ups and downs. In fact, the highs and lows can occur within minutes, the same day, week, or month. Fortunately, most bereaved persons experience a roller coaster ride that becomes more manageable over time, with fewer abrupt highs and lows. This gradual improvement may never lead to an end or resolution (Silverman, 1986), however, because many bereaved persons report: "You never get over it—you learn to live with it."

Fifth, loneliness and problems associated with managing the tasks of daily living are two of the most common and difficult adjustments for older adults (Lund et al., 1993). These problems are even more difficult for the spousally bereaved because their daily lives are closely connected with their spouses and they frequently become dependent on each other for conversation, love, and sharing of tasks. Loneliness is problematic because it involves missing, sadness, and a void that does not go away simply by being with or among other people (Lopata, 1973; Osterweis et al., 1984; Parkes, 1987). Many bereaved persons report feeling lonely but not being alone. In the case of experiencing the death of a parent, child, or sibling who may live many miles away, the bereaved often describe how lonely they feel when they realize they can no longer call their loved one on the phone or hear their voice again. In a study by the author, 70% of the bereaved spouses reported that loneliness was their single greatest difficulty during the two years following their partners' deaths (Lund, 1989a).

Unfortunately, most older adults have not learned to do many of the tasks of daily life that are important to maintaining their health, well-being, and happiness. Many older men have not learned how to prepare meals, wash clothes, and clean house because they left these tasks to their wives. Similarly, many older women have not learned how to do home repairs, maintain or sometimes even drive an automobile, and manage finances. Research has shown that these deficiencies further complicate the bereavement process because grief requires physical and mental energy which can be additionally depleted by worrying about how these tasks will be done (Lund, Caserta, et al., 1989). The lack of skills in daily living lends itself to a promising focus for future interventions because these skills can be easily taught. Hopefully, future cohorts of men and women will be better prepared and more skilled in a much broader range of tasks of daily living.

7.05.7 PREDICTORS OF BEREAVEMENT ADJUSTMENTS

It is very common to hear people say that the passage of time will heal the wound of grief, loss, or bereavement (Parkes, 1987; Silverman, 1986). This assumption is problematic and inaccurate because it implies that little or no effort is required to cope with the situation (Worden, 1991). Simply allowing time to pass is expected to bring about successful adjustment. Although it is a fact that with the passage of time most bereaved persons experience less difficulty and make more positive adjustments, it is important to recognize that it is what people do with their time that has the most direct impact on their outcomes (Lund, 1989a). Successful adjustments require active rather than passive coping strategies. The importance of being active and taking charge is integrated within Worden's task-oriented approach to grief (1991). Worden recommends this approach to clinicians because it involves an active coping strategy that fits well into a variety of interventions and it appears to be a more effective strategy than just letting time heal a wound (Silverman, 1986).

Another similar finding supported by research is that many bereaved persons have told others that it is most helpful to take one day at a time and to remain active, busy, and socially connected with other people (Caserta, Van Pelt, & Lund, 1989). Being physically and socially active during the process can help to reduce the feelings of despair, helplessness, loneliness, and being overwhelmed that often contribute to a long-term strategy of waiting—often in vain—for time to bring about healing. In addition to these active coping strategies, recent studies have documented the importance of having social support from others and opportunities for self-expression available, especially in the early months of bereavement (Dimond, Lund, & Caserta, 1987; Lund, Dimond et al., 1986; Schuster & Butler, 1989; Thompson, Gallagher, Cover, Gilewski, & Peterson, 1989). Some of the most difficult and stressful bereavement experiences occur in the first few months when grief issues are especially intense (Wortman et al., 1993). The quality of social support from friends, neighbors, and family is more important than the quantitative dimensions of support and it appears that older spousally bereaved persons tend to rely increasingly more on neighbors and friends because they are most convenient to contact (Dimond et al., 1987; Lund, Caserta, Van Pelt, & Gass, 1990; Schuster & Butler, 1989).

Research has also shown that how a person copes with loss early in the process appears to be one of the best predictors of long-term coping (Levy et al., 1994). Those who report that they are doing quite well during the first couple of months usually cope better than others a few years later (Lund, Dimond et al., 1986; Wortman & Silver, 1989). Bereaved persons therefore are likely to benefit most from social support that they receive during the first two to eight weeks and from opportunities to express how they feel at this time. The first couple of months are most critical because that is when new patterns of behavior are developed and the adjustments that are required can be overwhelming. Early social support, therefore, is most beneficial (Schuster & Butler, 1989; Trunnell, Caserta, & White, 1992). Bereaved persons most value the support provided by those who allow them to openly express anger, sadness, and other emotions without passing along advice and counsel (Caserta, Van Pelt, & Lund, 1989; Dimond et al., 1987; Vachon & Stylianos, 1988). It is important for the bereaved to know that their reactions are common and that others respect their feelings and care about their well-being. People can be most supportive by being good listeners. This also relieves pressure that many people impose on themselves by believing that they need to say things that will make everything better. Such bereavement-eliminating messages do not exist. Self-help groups for bereaved persons provide good opportunities for self-expression, particularly for those who do not already have someone in whom they can confide (Lund & Caserta, 1992). These support groups vary considerably in quality and effectiveness but they do have the potential of providing some valuable sources of support (Lieberman, 1993).

The importance of internal coping resources is supported by research (Ben-Sira, 1983), that has focused on the role of self-esteem (Lund, Dimond et al., 1986) and self-efficacy (Bandura, 1982; Goldfried & Robins, 1982; Parkes, 1987). Both concepts represent personal coping resources that individuals develop throughout their lives. Self-esteem refers to the positive or negative judgments or evaluations that one makes about self-worth. In addition to being a husband or wife and a father or mother, we usually make assessments about being good or bad in the way we perform these roles. Self-efficacy refers to one's ability to meet the changing demands of everyday life and feeling confident to do so. Not all of us feel equally able to manage difficult and challenging situations, particularly bereavement, which requires a broad range of skills and confidence in being able to develop and use them.

The way people feel about themselves and how skilled they are in managing the many tasks of daily—living maintaining a household, paying bills, driving a car, knowing how to access resources, and so forth—will influence how effectively they adjust to the loss of a loved one (Lund, Caserta, et al., 1989). Men and women in today's older age cohort are highly deficient in predictable skills, with men not being well prepared to do shopping, cooking, and cleaning and women being unskilled in financial matters and home repairs. People who develop positive self-esteem and competencies such as social, interpersonal, and instrumental skills are likely to have more favorable bereavement outcomes than those who develop negative self-images and lack self-efficacy. The need to be skilled in the tasks of daily living is important to bereaved people of all ages (Goldfried & Robins, 1982). Worden (1991) describes one of his four tasks of mourning as adjusting to an environment in which the deceased is missing. Learning new skills is considered to be essential to this part of the adjustment process.

Self-esteem and self-efficacy are highly interrelated because people who have confidence and pride in themselves are usually more motivated to learn new skills, and the process of becoming more competent in daily life itself creates more positive self-esteem (Lund, Caserta et al., 1989). Bereaved persons with these positive characteristics are likely to cope quite well because they will not be content with a passive approach to coping and simply waiting for things to get better. People with high self-esteem are more likely to believe that they deserve to feel better and have the desire to at least try to engage in new activities and pursue interests that they have always had. Conversely, bereaved people who never developed these personal coping resources are likely to experience long-term difficulties because they are more inclined to believe that they deserve to remain depressed, lonely, and incapacitated. One research subject said, "Things are terrible, I feel awful but who cares." These people tend to feel overwhelmed and take few, if any, constructive actions on their own behalf. In short, the three best predictors of how older adults experience bereavement are self-esteem, competence in managing tasks of daily living, and opportunities for self-expression (Lund, Dimond et al., 1986). When bereaved older adults are high on these three traits they are much more likely to manage quite well, and adults who are lacking all three are at considerable risk of having long-term difficulties. These factors are likely to be similarly important to people of all ages (Lund et al., 1993).

While it is extremely difficult to summarize briefly the growing and increasingly complicated research literature regarding the many factors that might predict bereavement outcomes, it is appropriate to note that many variables that were expected to be important have not been consistently related to the adjustment process. For example, research has not found large differences in adjustments between those who experienced expected or sudden death situations (Faletti et al., 1989; Hill, Thompson, & Gallagher, 1988). The general expectation has been that it is much easier to manage a bereavement loss if it was anticipated. Unfortunately, there do not appear to be substantial benefits from knowing that a loved one is dying because when the death actually occurs it still sets in motion the most profound parts of the bereavement process which include dealing with loneliness, depression, and managing the numerous tasks of daily living. When a death is sudden and unexpected, the bereaved survivors are likely to feel more intensely the grief-related emotions, but when these gradually subside, the long-term process is quite similar to those who expected the death (Lund, l989a). If there are benefits from anticipating the death, they are most likely to be attributed to what people do with their time while they anticipate the death. The key element to realizing any benefit is the preparation, not for the death itself, but preparation for learning to do some of the many things that will be required after the death such as learning to do the laundry, cooking, home repairs, and other tasks of daily living. Developing new skills and "having affairs in order" can be helpful to bereaved persons of all ages because energy can then be focused on the grief work that needs to be done (Worden, 1991).

Research comparing bereaved persons from different religious membership groups, with unique religious beliefs and various degrees of religiosity, has revealed that these variables might create some unique bereavement experiences but they do not predict very well the major outcomes such as depression, loneliness, morale, life satisfaction, and a variety of other physical and mental health measures (Falleti et al., 1989; Gallagher, Lovett, Hanley-Dunn, & Thompson, 1989; Lund et al., 1993). It is possible that with different measures of adjustment these religious factors might be more influential but current research points out that these and other factors may be less important than previously expected. The degree of religious activity did, however, emerge in some research to be associated with bereavement adjustment with the more religiously active coping more effectively than the less active (Lund et al., 1988). This finding is consistent with other research that has shown an advantage to bereaved persons who make an effort to remain more socially active and to keep busy (Caserta, Van Pelt, & Lund, 1989). For many older adults, their religious participation can be a very important, convenient, and consistent source of activity that offers social contact, informal support, and a meaningful use of time.

Another important finding to emerge from research is that many older spousally bereaved persons find themselves in a dilemma wherein they must make decisions either to please themselves or others (Caserta, Van Pelt, & Lund, 1989; Dimond et al., 1987; Lund et al., 1990). It is not uncommon for some adult children to exert considerable pressure on their grieving parent to sell the house, get away from the memories, move closer to others in the family, sell the car, or simply do what they expect of them. In most cases, the bereaved older adult does not want the parent/child roles to reverse where they feel pressure to do what their children want them to do (Lund et al., 1990). Most adult children typically have good intentions but sometimes get carried away with trying to take away their parent's grief and make everything instantly better. As research shows, there are no quick solutions to the many adjustments that need to be experienced.

The pressure to conform to the desires and wishes of others, including the deceased, happens often enough that it can be referred to as "grieving according to the expectations of others" or even, "grieving according to the expectations of the deceased." One very common example of doing or saying what others expect occurs when an adult child visits a grieving parent and asks how he or she is doing. Our research interviewers have heard many of these bereaved parents say that they simply tell their children that they are doing fine because they know that is what their children want and expect to hear. They say that they do not want to trouble their children with the truth. They also report feeling that they need to project an image of strength to their children, neighbors, and family. Another interesting example was a bereaved woman who told an interviewer that she wanted to continue a friendship that she had with another man but felt that others would not approve of her spending time with him. She and her deceased husband had been lifelong friends with this man and his deceased wife. They had done many things together as couples. Now that they had both been widowed they wanted to continue spending time together. This woman made it clear to the interviewer that they were "just friends, nothing sexual going on." Rather than living according to their own expectations they made the effort to conceal from others that they were spending time together. Her male friend would wait until dark to visit and would park his car around the block and sneak through a neighbor's backyard to get to her house. As she described the strategy, the interviewer noted that she was embarrassed about the behavior but did not know how she could be more open about the relationship for fear of what others might think.

The deceased often continue to play important influential roles in how the bereaved survivor makes decisions about what is appropriate or not (Troll, Miller, & Atchley, 1979). Some bereaved spouses participating in research studies have said that they were very reluctant to do things that their deceased partners would not approve and that they were struggling with a decision or had already decided in favor of their deceased spouse's wishes. Most commonly, the decisions were regarding what to do with personal belongings and household items, forming new friendships, selling the house and moving to another area, returning to old hobbies and interests, and the purchase of clothing, furniture, and automobiles. One of the more interesting examples was a widow in a research project who described the very serious dilemma that she was having while trying to decide what to do with her husband's favorite friend, the dog. She indicated that she had always hated the dog and would love to get rid of it. Now that her husband was dead she had the opportunity to give the dog away. But, as she considered her choices she now felt that her deceased husband would be very upset if she did not keep and care for his companion. The widow did not take this matter lightly and chose to keep the dog even though her preference was to end the relationship. Until they are actually in the bereavement process people do not realize how difficult some daily life decisions can become. What appear to be rather simple choices are in fact complicated by the reality that others' opinions and expectations create considerable pressure. Although statistical data are not available regarding this, it is very likely that those who are bold enough to live more according to their own expectations will make the many bereavement adjustments with greater satisfaction than those who favor doing things primarily to please others (Caserta, Van Pelt, & Lund, 1989). Many older adults will need validation that it is acceptable for them to do things consistent with their own wishes, but, at the same time, know that others may disagree with their choices. Bereaved people will need to work through these kinds of dilemmas many times and discover what works best for them.

The predictors of adjustment to bereavement are similar for both men and women and for young and older adults (Lund, Caserta, & Dimond, 1986). Age and gender are not the most influential factors in the course of bereavement (Lund et al., 1993). What is of greater importance is to develop positive feelings about oneself early in life, continue to enhance these views over the life course, and develop skills that help one to meet the changing circumstances and demands of daily living (Lund, 1989a; Worden, 1991). People with these traits are more likely than others to adjust well to nearly any major life stress or transition (Goldfried & Robins, 1982; Hooyman & Kiyak, 1996). Human development is a lifelong process. In the case of late life bereavement, the developmental process is challenged. During this transition period, the bereaved person can remain physically, psychologically, and socially disrupted or can emerge with a sense of growth from learning new skills, becoming more independent, and developing a clearer self-identity.

7.05.8 WORKING WITH THE BEREAVED

Although bereavement is one of the most stressful experiences that occur in our lives, we should not conclude that most bereaved persons will need help with their adjustment process. Research has shown that many bereaved persons do not want or need intervention services (Levy, Derby, & Martinkowski, 1993; Lund et al., 1985). Most researchers estimate that between 15 and 25% of the bereaved will experience considerable long-term difficulty and may require major forms of assistance from others (Lund, 1993). This point is important to recognize because many clinicians and other service providers assume that most older bereaved persons are depressed, socially isolated, and incapacitated by their loss. In fact, gerontologists have reminded us that the majority of older adults do not fit the general stereotypes of being dependent, frail, poor, lonely, and in poor health (Hooyman & Kiyak, 1996). The research evidence does not support these and other assumptions, although certainly some bereaved persons are depressed, isolated, and incapacitated (Wortman & Silver, 1989). In fact, as mentioned previously, older adults are quite resilient and find ways to manage the many losses that they experience throughout their lives (Cleiren, 1993; Lund, 1993; Murrell et al., 1988; Worden, 1991).

The tasks for those developing intervention services are many, and among their first should be an attempt to reach those who are at greatest risk—to seek them out rather than simply announcing the existence of the service—and encourage their participation by explaining how and why the service will be helpful. It is absolutely critical to recognize that many, perhaps most, of the potential population of bereaved persons will not want to participate (Caserta & Lund, 1993; Lund et al., 1985). This can be discouraging to those who are committed to the value of their services, but their motivation to continue their efforts will probably be less adversely affected if they anticipate the lack of enthusiasm among many potential clients. For example, only 44% of the bereaved spouses who completed a study by the author said that they would have liked the opportunity of attending self-help groups (Lund et al., 1985). Later in another study only 27% of those assigned to self-help groups actually agreed to participate (Lund, Redburn, Juretich, & Caserta, 1989). For those who need assistance but are reluctant, it is often beneficial to obtain encouragement from others in the bereaved person's network, particularly physicians and important family members and friends.

Interventions need to be available early in the bereavement process and continue over relatively long periods of time (Horacek, 1995; Schuster & Butler, 1989; Trunnell et al., 1992). As discussed previously, there is a good deal of research evidence suggesting that the first few months are the most difficult and that early adjustments will influence outcomes much later (Caserta & Lund, 1993; Wortman et al., 1993). Also, because bereavement may last for many years and some people may not be ready for early interventions, it would be most helpful to have services available over long periods of time so that they may obtain the appropriate help when they are most ready to receive and benefit from it (Faletti et al., 1989). Others may require intervention or assistance for long periods of time but dependency on any particular form of help is not desired (Silverman, 1986).

Because the impact of bereavement is multidimensional, it is imperative that interventions offer comprehensive and diverse sets of services. It is unlikely that any one intervention will be capable of providing all that is needed, but each intervention should clearly identify which needs are being targeted in relation to the overall multidimensional process so that there is an awareness of what help is not being provided. For example, the death of a spouse in later life can impact emotions, psychosocial functioning, health, family life, interpersonal relationships, work, recreation, and financial situations. Therefore, those designing interventions which provide primarily an opportunity for self-expression (such as self-help groups) should recognize that some dimensions are not likely to be addressed (Shuchter, 1986; Shuchter & Zisook, 1993). Clinicians need to be aware of the full range of services that are available in their communities so their clients can obtain a unique set of diverse services to meet their unique needs. Ideally, all communities would have available a variety of interventions or services so that each person's unique skills, resources, and circumstances could be matched to the most appropriate set of services. While this is unlikely, we can at least strive to offer interventions with the broadest scope of impact. Therefore, whenever possible, it would be worthwhile to impact several dimensions simultaneously (Shuchter, 1986) by providing opportunities for self-expression and the enhancement of self-esteem; by teaching new skills to complete the tasks of daily living; by enhancing and mobilizing already existing social support networks; by providing education and assistance regarding health, nutrition, and exercise; and by encouraging social participation. Each of the predictor variables reviewed in this chapter can be integrated within nearly every type of intervention.

Behavioral competency interventions (Ben-Sira, 1983; Hill, Lund, & Packard, 1996) that identify and provide specialized training in daily survival skills can be of great value for several reasons. First, such training equips bereaved individuals with the special expertise they need to manage the tasks that had been performed by their deceased loved one. Second, successful mastery of new skills builds self-esteem and reinforces beliefs about personal competence. Finally, the process of training provides informal opportunities for the individual to socialize and interact with others while learning important life skills. Skill training interventions offer an innovative approach with promise for helping bereaved older adults cope with issues of grief and mourning and prepare for healthy independent living.

Various intervention formats and professionals are needed to ensure that appropriate services are available. Not all people experience bereavement in the same way; similarly, not all people will use or benefit from the same interventions. In terms of format, some people will only want to have a one-on-one type of intervention. This might be because they feel the uniqueness of their situation can be dealt with more effectively one-to-one, because they are reluctant to express personal and sensitive feelings in group situations (Silverman, 1986). Others have reported that they particularly enjoyed being in a self-help group because they learned from others, recognized some commonalities in their situations, and enjoyed the socializing and friendships which developed (Lund & Caserta, 1992; Caserta & Lund, 1996).

Self-help support groups are a commonly used intervention with older adults (Lund, Redburn et al., 1989). Such groups generally consist of bereaved individuals who voluntarily meet to discuss issues and problems associated with their situation (Lieberman, 1993; Lieberman & Videka-Sherman, 1986). Participants can experience hope, learn new ideas and problem-solving strategies, improve skills for developing social relationships, become less lonely, explore new role definitions, discover that others have similar difficulties, and receive additional social support and encouragement. Belonging to a group and sharing feelings of fellowship and solidarity have been helpful to many older adults. Many of the general benefits of various self-help groups are especially relevant to bereaved older adults. Group members often discuss ways to alleviate the feelings of loneliness, social isolation, and hopelessness that accompany widowhood. Members also have opportunities to meet new people, participate in social activities, and form new friendships (Caserta & Lund, 1993). In addition, many bereaved spouses experience difficulties managing tasks of daily living such as meal preparation, house cleaning, yard maintenance, and transportation that were performed previously by the deceased spouse. Group members regularly share information, advice, and resources about these common problems. Finally, self-help groups are less expensive than most alternative services and are often attractive to individuals who are reluctant to use the services of mental health professionals and agencies. Support groups are not without their share of problems because they require clear goals, well trained leaders, and realistic expectations. Poorly run groups can become overly depressing and ineffective. Good outcomes do not often happen by chance but are the result of careful planning and good leadership (Lund, Redburn et al., 1989).

Many people have skills and expertise which are well suited for helping the bereaved. Silverman (1986), who developed the widow-to-widow program, has shown that widows are quite capable of assisting each other. An experienced widow can reveal to the new widow that she has been there and knows how it feels to grieve. Also, there are important contributions that can be made by researchers, gerontologists, psychologists, psychiatrists, physicians, social workers, nurses, occupational therapists, counselors, clergy, and many other professionals, including social scientists, lawyers, accountants, educators, and direct service providers. Again, because bereavement has a multidimensional impact, interventions can be developed by many different professional and trained team members. Other bereaved persons can provide a sharing of experiences while trained persons assist with legal, financial, health, spiritual, and educational issues. A multidimensional team approach is highly recommended because it will increase the likelihood of developing interventions which address the diversity of needs and lead to greater success.

7.05.9 SUMMARY

Bereavement is a common and natural experience that we all share. Unfortunately, many people are not well prepared for the pain, threats, and challenges that it presents. Because bereavement occurs in broader contexts, we do not all have the same personal and social resources available (Caserta & Lund, 1993), or have the same experiences and needs. What appears to help one person may have an entirely different impact on someone else. The best way to be prepared to help bereaved persons is to be educated about the process, assess the context in which it occurs, become familiar with a wide range of helping resources, and be patient, a good listener and nonjudgmental.

One final suggestion concerning future research and clinical practice is the need to be creative and work in interdisciplinary teams whenever possible. Being creative means that while looking at the same thing you are able to see something different, perhaps many different things. Good quality research and interventions are not always the result of following a "cookbook" approach where we simply look up the issue or problem and find the recipe to resolve it. In order to ask the most interesting and important research questions, decide on the best methodology, or develop and offer the most effective intervention, we need to be very creative. We need to be willing to take a risk and modify the way we mechanically make research and treatment decisions. One of the best ways to be more creative is to work in interdisciplinary teams.

Researchers and clinicians are trained within specific disciplines and learn to favor specific research designs and intervention strategies. Over time, each professional's training becomes his or her area of expertise and there is a natural tendency to define questions and issues in a way that leads to using the same methodologies and practices. Abraham Kaplan, a philosopher of science, refers to this rigidity as "the law of the instrument" (1964, p. 28). He states that if you give a young boy a hammer, everything he encounters will need pounding. If a researcher or clinician exclusively uses only one theory, one methodology, or one counseling strategy then other techniques and strategies are ignored. If they cannot become skilled with more than just a single approach to their work (the hammer) it is even more important for them to work collaboratively in interdisciplinary teams so that other tools, strategies, perspectives, and methodologies are used. One of the best ways to be creative is to assemble a team of investigators who represent diverse and needed areas of expertise. When clinicians and researchers from medicine, pharmacy, nursing, social work, social and behavioral sciences, and other disciplines work collaboratively the end results are much greater than just the sum of the parts. New methodologies, techniques, and strategies will result from the creativity that is enhanced by working in multidisciplinary teams.

Death and other bereavement-related losses will continue to surround our daily lives, offering us choices to learn from these experiences or to use them to be better prepared for our personal losses and our professional work. Although we have learned much from research and clinical practice there are substantial gaps in our knowledge. Much creative and pioneering work remains for those who have the interest, expertise and enthusiasm to learn more about grief and bereavement and to apply this knowledge to make a difference in people's lives.

7.05.10 REFERENCES

Aiken, L. R. (1994). *Dying, death and bereavement* (3rd ed.) Boston: Allyn & Bacon.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Atchley, R. C. (1989). A continuity theory of normal aging. *The Gerontologist, 29,* 183–190.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37,* 122–147.

Ben-Sira, Z. (1983). Loss, stress, and readjustment: The structure of coping with bereavement and disability. *Social Science Medicine, 17*, 1619–1632.

Bowlby, J. (1969). *Attachment and loss. Vol.1: Attachment*. New York: Basic Books.

Caine, L. (1974). *Widow*. New York: Bantam.

Caserta, M. A., & Lund, D. A. (1992a). Bereaved older adults who seek early professional help. *Death Studies, 16*(1), 17–30.

Caserta, M. S., & Lund, D. A. (1992b). Bereavement stress and coping among older adults: Expectations versus the actual experience. *Omega, 25*, 33–45.

Caserta, M. S., & Lund, D. A. (1993). Intrapersonal resources and the effectiveness of self-help groups for bereaved older adults. *The Gerontologist, 33*, 619–629.

Caserta, M. S., & Lund, D. A. (1996). Beyond bereavement support group meetings: Exploring outside social contacts among the members. *Death Studies, 20*, 537–556.

Caserta, M. S., Lund, D. A., & Dimond, M. F. (1989). Older widows' early bereavement adjustments. *Journal of Women and Aging, 1*(4), 5–27.

Caserta, M. S., Van Pelt, J., & Lund, D. A. (1989). Advice on the adjustment to loss from bereaved older adults: An examination of resources and outcomes. In D. A. Lund (Ed.), *Older bereaved spouses: Research with practical applications*. New York: Hemisphere.

Chigier, E. (1988). *Grief and bereavement in contemporary society*. London: Freund.

Cleiren, M. (1993). *Bereavement and adaptation: A comparative study of the aftermath of death*. Washington, DC: Hemisphere.

Cumming, E., & Henry, W. E. (1961) *Growing old*. New York: Basic.

Dimond, M.F., Lund, D. A., & Caserta, M. S. (1987). The role of social support in the first two years of bereavement in an elderly sample. *The Gerontologist,* 27, 599–604.

Faletti, M. V., Gibbs, J. M., Clark, M. C., Pruchno, R. A., & Berman, E. A., (1989). Longitudinal course of bereavement in older adults. In D. A. Lund (Ed.), *Older bereaved spouses: Research with practical applications* (pp. 37–51). New York: Hemisphere.

Farnsworth, J., Pett, M., & Lund, D. A. (1989). Predictors of loss management and well-being in later life widowhood and divorce. *Journal of Family Issues, 10,* 102–121.

Folkman, S., Lazarus, R. S., Dunkel-Schetter, C., De Longis, A., & Gruen, R. J. (1986). Dynamics of a stressful encounter: Cognitive appraisal, coping, and encounter outcomes. *Journal of Personality and Social Psychology, 50,* 992–1003.

Freud, S. (1957). *The Standard Edition of the Complete Psychological Works of Sigmund Freud*. London: Hogarth.

Gallagher, D., Lovett, S., Hanley-Dunn, P., & Thompson, L. W. (1989). Use of select coping strategies during late-life spousal bereavement. In D. A. Lund (Ed.), *Older bereaved spouses: Research with practical applications* (pp. 111–121). New York: Hemisphere.

Gass, K. A. (1989a). Appraisal, coping, and resources: Markers associated with the health of aged widows and widowers. In D. A. Lund (Ed.), *Older bereaved spouses: Research with practical applications* (pp. 79–94). New York: Hemisphere.

Gass, K. A. (1989b). Health of older widowers: Role of appraisal, coping, resources, and type of spouse's death. In D. A. Lund (Ed.), *Older bereaved spouses: Research with practical applications* (pp. 95–110). New York: Hemisphere.

Goldfried, M. R., & Robins, C. (1982). On the facilitation of self-efficacy. *Cognitive Therapy and Research*, 6, 361–380.

Grollman, E. A., (1977). *Living when a loved one has died.* Boston: Beacon.

Hardt, D. V. (1978). An investigation of the stages of bereavement. *Omega, 9*, 279–285.

Havighurst, R. J. (1968). Personality and patterns of aging. *The Gerontologist, 8,* 20–23.

Hess, B., Markson, E., & Stein, P. (1996). *Sociology*. Needham Heights, MA: Allyn & Bacon.

Hill, C. D., Thompson, L. W., & Gallagher, D. (1988). The role of anticipatory bereavment in older women's adjustment to widowhood. *The Gerontologist, 28,* 792–796.

Hill, R. D., Lund, D. A., & Packard, T. (1996). Bereavement. In J. I. Sheikh (Ed.), *Treating the elderly* (pp. 45–74). San Francisco: Jossey Bass.

Hooyman, N., & Kiyak, H. A. (1996). *Social gerontology: A multidisciplinary perspective*. Needham Heights, MA: Simon & Schuster.

Horacek, B. J. (1995). A heuristic model of grieving after high-grief deaths. *Death Studies, 19,* 21–31.

Irish, D. P., Lundquist, K. F., & Nelsen, V. J. (1993). *Ethnic variations in dying, death, and grief.* Washington, DC: Taylor & Francis.

Kaplan, A. (1964). *The conduct of inquiry*. San Francisco: Chandler.

Kubler-Ross, E. (1969). *On death and dying*. New York: Macmillan.

Kushner, H. S. (1982). *When bad things happen to good people*. New York: Avon.

Lazarus, R. S. (1966). *Psychological stress and the coping process*. New York: McGraw-Hill.

Leming, M. R., & Dickinson, G. E. (1990). *Understanding dying, death and bereavment*. New York: Holt, Rinehart & Winston.

Levy, L. H., Derby, J. F., & Martinkowski, K. S. (1993). Effects of membership in bereavement support groups on adaptation to conjugal bereavement. *American Journal of Community Psychology, 21*, 361–381.

Levy, L. H., Martinkowski, K.S., & Derby, J. F. (1994). Differences in patterns of adaptation in conjugal bereavement: Their sources and potential significance. *Omega, 29*, 71–87.

Lewis, C. S. (1961). *A grief observed.* Minneapolis, MN: Winston.

Lieberman, M. A. (1993). Bereavement self-help groups: A review of conceptual and methodological issues. In M. S. Stroebe, W. Stroebe, & R. O. Hansson (Eds.), *Handbook of bereavement: Theory, research, and intervention* (pp. 411–426). New York: Cambridge University Press.

Lieberman, M. A., & Videka-Sherman, L. (1986). The impact of self-help groups on the mental health of widows and widowers. *American Journal of Orthopsychiatry, 56,* 435–449.

Lindemann, E. (1944). The symptomatology and management of acute grief. *American Journal of Psychiatry, 101,* 141–48.

Lopata, H. Z. (1973). *Widowhood in an American city*. Cambridge, MA: Schenkman.

Lund, D. A. (1989a). Conclusions about bereavement in later life and implications for interventions and future research. In D. A. Lund (Ed.), *Older bereaved spouses: Research with practical applications* (pp. 217–231). New York: Hemisphere.

Lund, D. A. (1989b). *Older bereaved spouses: Research with practical applications*. New York: Hemisphere.

Lund, D. A. (1993). Widowhood: The coping response. In R. Kastenbaum (Ed.), *The encyclopedia of adult development* (pp. 537–541).Phoenix, AZ: Oryx.

Lund, D. A. (1996). Coping in later life. In K. J. Doka (Ed.), *Journeys* (pp. 1,3). Washington, DC: Hospice Foundation of America.

Lund, D. A., & Caserta, M. S. (1992). Older bereaved spouses' participation in self-help groups. *Omega, 25*, 47–61

Lund, D. A., Caserta, M. S., & Dimond, M. F. (1986). Gender differences through two years of bereavement among the elderly. *The Gerontologist, 26*(3), 314–320.

Lund, D. A., Caserta, M. S., & Dimond, M. F. (1988). A comparison of bereavement adjustments between Mormon and non-Mormon older adults. *Journal of Religion and Aging, 5*(1/2), 75–92.

Lund, D. A., Caserta, M. S., & Dimond, M. F. (1993). The course of spousal bereavement in later life. In M. S.

Stroebe, W. Stroebe, & R. O. Hansson (Eds.), *Handbook of bereavement* (pp. 240–254). New York: Cambridge University Press.

Lund, D. A., Caserta, M. S., Dimond, M. F., & Gray, R. M. (1986). Impact of bereavement on the self-conceptions of older surviving spouses: An application of the interactionist perspective. *Symbolic Interaction, 9*(2), 235–244.

Lund, D. A., Caserta, M. S., Dimond, M. F., & Shaffer, S. K. (1989). Competencies, tasks of daily living, and adjustments to spousal bereavement in later life. In D. A. Lund (Ed.), *Older bereaved spouses: Research with practical applications* (pp. 135–152). New York: Hemisphere.

Lund, D. A., Caserta, M. S., Van Pelt, J., & Gass, K. A. (1990). Stability of social support networks after later-life spousal bereavement. *Death Studies, 14*, 53–73.

Lund, D. A., Dimond, M. F., Caserta, M. S., Poulton, J. L., Johnson, R. J., & Connelly, R. J. (1986). Identifying elderly with coping difficulties after two years of bereavement. *Omega: Journal of Death & Dying, 16*(3), 213–224.

Lund, D. A., Dimond, M. F., & Juretich, M. (1985). Bereavement support groups for the elderly: Characteristics of potential participants. *Death Studies, 9*, 309–321.

Lund, D. A., Redburn, D. E., Juretich, M. S., & Caserta, M. S. (1989). Resolving problems implementing bereavement self-help groups. In D. A. Lund (Ed.), *Older bereaved spouses: Research with practical applications* (pp. 213–216). New York: Hemisphere.

Marris, P. (1958). *Widows and their families*. London: Routledge & Kegan Paul.

McCrae, R. R., & Costa, P. T. (1993). Psychological resilience among widowed men and women: A 10-year follow-up of a national sample. In M. S. Stroebe, W. Stroebe, & R. O. Hansson (Eds.), *Handbook of Bereavement: Theory, research and intervention* (pp. 196–202). New York: Cambridge University Press.

Murrell, S. A., Himmelfarb, S., & Phifer, J. F. (1988). Effects of bereavement/loss and pre-event status on subsequent physical health in older adults. *International Journal of Aging and Human Development, 27*, 89–107.

Olson, D. H., Sprenkle, D., & Russell, C. (1979). Circumplex model of marital and family systems: 1. Cohesion and adaptability dimensions, family types, and clinical applications. *Family Process, 18,* 3–28.

Osterweis, M., Solomon, F., & Green, M. (1984). *Bereavement reactions, consequences, and care*. Washington, DC: National Academy Press.

Parkes, C. M. (1987). *Bereavement: Studies of grief in adult life* (2nd ed.). Madison, WI: International Universities Press.

Pearlin, L. I., & Mullan, J. T. (1992). Loss and Stress in Aging. In M. L. Wykle, E. Kahana, & J. Kowal (Eds.) *Stress & health among the elderly* (pp. 117–132). New York: Springer.

Rando, T. A. (1993). *Treatment of complicated mourning*. Champaign, IL: Research Press.

Riley, M. W. (1971). Social gerontology and the age stratification of society. *The Gerontologist, 11,* 79–87.

Rosenstock, I. M. (1990). The health belief model: Explaining health behavior through expectancies. In K. Glanz, F. M. Lewis, & B. K. Rimer (Eds.), *Health behavior and health education* (pp. 39–62). San Francisco: Jossey-Bass.

Rubin, S. S. (1993). The death of a child is forever: The life course impact of child loss. In M. S. Stroebe, W. Stroebe, & R.O. Hansson (Eds.), *Handbook of bereavement: Theory, research, and intervention* (pp. 285–299). New York: Cambridge University Press.

Schiff, H. S. (1977). *The bereaved parent*. New York: Penguin.

Schuster, T. L., & Butler, E. W. (1989). Bereavement, social networks, social support, and mental health. In D. A. Lund (Ed.), *Older bereaved spouses: Research with practical applications* (pp. 55–68) New York: Hemisphere.

Shuchter, S. R. (1986). *Dimensions of grief: Adjusting to the death of a spouse*. San Francisco: Jossey-Bass.

Shuchter, S. R., & Zisook, S. (1993). The course of normal grief. In M. S. Stroebe, W. Stroebe, & R. O. Hansson, (Eds.), *Handbook of bereavement: Theory, research, and intervention* (pp. 23–43). New York: Cambridge University Press.

Silverman, P. R. (1986). *Widow to widow*. New York: Springer.

Staudacher, C. (1987). *Beyond grief: A guide for recovering from the death of a loved one*. Oakland, CA: New Harbinger.

Stephenson, J. S. (1985). *Death, grief and mourning: Individual and social realities*. New York: Free Press.

Stroebe, M. S., Hansson, R. O., & Stroebe, W. (1993). Contemporary themes and controversies in bereavement research. In M. S. Stroebe, W. Stroebe, & R. O. Hansson (Eds.), *Handbook of Bereavement: Theory, research, and intervention* (pp. 457–477). New York: Cambridge University Press

Stroebe, M. S., & Stroebe W. (1991). Does "grief work" work? *Journal of Consulting and Clinical Psychology, 59,* 479–482.

Stroebe, M. S., Stroebe, W, & Hansson, R. O. (1993a). Bereavement research and theory: An introduction to the handbook. In M. S. Stroebe, W. Stroebe, & R. O. Hansson (Eds.), *Handbook of Bereavement: Theory, research, and intervention* (pp. 3–22). New York: Cambridge University Press.

Stroebe, M. S., Stroebe, W., & Hansson, R. O. (Eds.) (1993b). *Handbook of Bereavement: Theory, research, and intervention.* New York: Cambridge University Press.

Stroebe, W., & Stroebe, M. (1987). *Bereavement and Health: The Psychological and physical consequences of partner loss*. London: Cambridge University Press.

Thompson, L. W., Gallagher, D., Cover, H., Gilewski, M., & Peterson, J. (1989). Effects of bereavement on symptoms of psychopathology in older men and women. In D. A. Lund. *Older bereaved spouses: Research with practical applications* (pp. 17–24). New York: Hemisphere.

Troll, L. E., Miller, S. J., & Atchley, R. C. (1979). *Families in later life*. Belmont, CA: Wadsworth.

Trunnell, E. P., Caserta, M. S., & White, G. L. (1992). Bereavement: Current issues in intervention and prevention. *Journal of Health Education, 23*, 275–280.

US Bureau of the Census (1991). Marital status and living arrangements: March 1990. *Current Population Reports, Series P–20, No. 450*. Washington, DC: Government Printing Office.

Vachon, M. L., & Stylianos, S. K. (1988). The role of social support in bereavement. *Journal of Social Issues, 44*, 175–190.

Worden, J. W. (1991) *Grief counseling and grief therapy* (2nd ed.). New York: Springer.

Wortman, C., & Silver, R. (1989). The myths of coping with loss. *Journal of Consulting and Clinical Psychology, 57,* 349–357.

Wortman, C., Silver, R., & Kessler, R. (1993). The meaning of loss and adjustment to bereavement. In M. S. Stroebe, W. Stroebe, & R. O. Hansson (Eds.), *Handbook of bereavement: Theory, research, and intervention* (pp. 349–366). New York: Cambridge University Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.06

Understanding Decisional Capacity of Older Adults

LORI FRANK
MEDTAP International, Bethesda, MD, USA
and
MICHAEL SMYER
Boston College, Chestnut Hill, MA, USA

7.06.1 INTRODUCTION 114

7.06.2 INFORMED CONSENT 114
- *7.06.2.1 Brief History of the Doctrine of Informed Consent* 114
- *7.06.2.2 Informed Consent and the Elderly* 115

7.06.3 COMPETENCE TO PROVIDE INFORMED CONSENT 115
- *7.06.3.1 Definitions of Competence* 115
  - *7.06.3.1.1 Rationality-dependent definitions* 116
  - *7.06.3.1.2 Combination definitions* 116
  - *7.06.3.1.3 Appreciation-dependent definitions* 117
- *7.06.3.2 Other Considerations for Definitions* 117

7.06.4 PSYCHOLOGICAL PERSPECTIVES ON COMPETENCE TO MAKE DECISIONS 119
- *7.06.4.1 Forensic Psychological Views of Competence* 119
- *7.06.4.2 Cognitive Psychological Views of Competence* 119

7.06.5 ASSESSMENT OF COMPETENCE 120
- *7.06.5.1 Assessment of Competence to Consent to Treatment* 120
- *7.06.5.2 Assessment of Other Types of Competence* 123

7.06.6 COMPETENCE IN ACTION: ADVANCE DIRECTIVES 124
- *7.06.6.1 The Patient Self-determination Act* 124
  - *7.06.6.1.1 Knowledge and attitudes of the elderly about ADs* 124
  - *7.06.6.1.2 Practical and ethical aspects of implementation of the PSDA* 126

7.06.7 OTHER ISSUES 127
- *7.06.7.1 Other Considerations for Decision Making* 127
  - *7.06.7.1.1 The nature of preferences* 127
  - *7.06.7.1.2 The role of the family* 127
- *7.06.7.2 Directions for Future Research* 128

7.06.8 REFERENCES 129

7.06.1 INTRODUCTION

With age comes increased interaction with the legal system in our society. Because of increased prevalence of cognitive impairment with age, psychologists are often expected to address legal concerns that result from potential cognitive impairment. Concerns about the capacity of the older individual to make autonomous decisions surface frequently, in part because of the greater decisional burden that age-related health problems place on older adults. Thus, assessing capacity for decision making is a large and growing area of geropsychology.

An understanding of decision-making ability of older adults is dependent upon a number of different fields of inquiry and several different areas of literature. This chapter is organized into five main sections. The first summarizes the development of the current doctrine of informed consent and the literature on informed consent with older adults is briefly reviewed. In the second, the concept of competence is examined more closely, beginning with a brief history of the development of theoretical and empirical work on decisional capacity. In section three, the forensic psychology and cognitive psychology contributions to the concept of capacity are compared. In the fourth, competence assessment is discussed. In the fifth section some of the issues raised by work with older adults is explored in the context of a specific policy, the Patient Self-determination Act. Finally, additional areas relevant to older adults' decisional capacity are highlighted.

7.06.2 INFORMED CONSENT

7.06.2.1 Brief History of the Doctrine of Informed Consent

Investigations of the adequacy of informed consent procedures usually involve reference to measurable components of the construct of "informed consent." Three components are identified in a widely accepted model of informed consent, a model which derives from legal requirements: there must be adequate disclosure of relevant information, the choice must be voluntary, and the individual must be competent to make the decision (Appelbaum and Grisso, 1988, 1995; Faden & Beauchamp, 1986). While some controversy surrounds all three of these elements, competence has been the most difficult to define and measure, and has generated the most discussion. Views of competence in the late 1990s will be examined in the next section. First, however, the development of the current view is described.

The legal history of informed consent in the USA can be traced back to eighteenth century Britain, where the King's Court ruled that medical surgery performed without the consent of the patient constituted a "tortious assault." More recently, a 1905 Minnesota case (*Mohr v. Williams*) established that surgery performed without permission constituted a battery. The significance of this ruling was that the patient, and not the physician, was to determine choices as to physical interference with an individual's body. Further affirmation of this principle of self-determination was recorded by Justice Benjamin Cardozo, in an opinion written in 1914: "every human being of adult years and sound mind has a right to determine what should be done with his own body" (*Schloendorff v. Society of New York Hospitals,* 1914).

Two cases refined the concept of "informed consent." In the 1957 case of *Salgo v. Leland Stanford Jr. Board of Trustees*, the court ruled that physicians had a duty to disclose information about the risks and alternatives of treatment. In the 1960 case of *Natanson v. Kline*, the court ruled that the patient must be provided with a certain amount of relevant information; it was no longer sufficient for the patient to consent to or decline treatment. At that time, the legal standard for the amount and kind of information provided was based on a professional practice standard and was, therefore, set by the medical community: "deviation from the standard of conduct of a reasonable and prudent medical doctor of the same school of practice as the defendant under similar circumstances."

In 1972, the case of *Canterbury v. Spence* shifted the determination of the legal standard for disclosure from the medical community to the patient. The patient was entitled to disclosure of information that any "average, reasonable patient" would require in deciding whether or not to consent to the procedure, and *not* disclosure based on a professional practice standard. This ruling is significant for its emphasis on patients' rights which, according to the court, "demands a standard set by law for physicians rather than one which physicians may or may not impose upon themselves."

By the end of the 1970s, legal developments required that patients be apprised of the nature and risks of a medical procedure by the physician. The concern of the law was that information was properly provided to the patient, the patient comprehended it, and the patient voluntarily consented to the procedure.

The cases reviewed above indicate the source of legal standards in the late 1990s for the information that providers are required to disclose, but the practical translation of these legal requirements to clinical situations is unclear. The law has paid less attention to the abilities an individual should possess in order to provide truly informed consent; such requirements are being explored more thoroughly in the field of psychology (and will be discussed more fully later).

7.06.2.2 Informed Consent and the Elderly

The elderly are a heterogeneous group (Stanley, Guido, Stanley, & Shortell, 1984), yet older adults are grouped together for the purposes of many policy initiatives. Besides policy precedent, there may be other reasons to examine older adults as a single group. In some studies, elderly subjects have demonstrated poorer comprehension of informed consent information than younger subjects (Stanley et al., 1984; Taub, 1980). Given historical differences in education level and vocabulary level between successive cohorts in the US, certain segments of the elderly population may be at increased risk for poorer comprehension of informed consent material.

To the extent that age differences continue to be manifest, and since end-of-life decisions occur more commonly among older adults than any other age group, elderly people may in fact present "unique ethical challenges" (Buehler, 1990) to health care providers. Before singling out a group for study, however, consideration must be made of the criteria involved in defining the group, and within-group heterogeneity must be acknowledged. While it is true that certain broad statements about a heterogeneous group like the elderly are of limited use (and may be ethically questionable (Cohen, 1996)), generalities do exist which justify the examination of elderly people as a separate group.

The process of obtaining informed consent from older adults may differ from obtaining informed consent from younger adults in important ways. For example, because of the prevalence of chronic medical conditions among older adults, disclosure and consent may be part of an ongoing process rather than just elements related to an acute episode. Chronic illness may change important features of the relationship between health care provider and patient for which informed consent laws were written.

Moody (1988) advocates a shift in emphasis from informed consent to the concept of "negotiated consent," in which emphasis shifts from a dichotomous variable view of consent (can the patient provide informed consent or not?) to a continuous variable view (to what extent can the patient understand this information and make an informed choice, over time?). Although originally conceived of in relation to long-term care settings, the concept of "negotiated consent" may be applicable in acute care settings as well, particularly for older adults with chronic conditions, re-admitted to the same facility multiple times.

7.06.3 COMPETENCE TO PROVIDE INFORMED CONSENT

Historically, legal applications of the term "competence" have provided inadequate specificity and have been underdetermined for research purposes (Appelbaum & Grisso, 1988). The limitations of legal conceptualizations of competence have prompted some researchers to examine the concept from a psychological perspective instead. Below, the evolution of definitions of the term "competence" are briefly reviewed. General aspects of definitions are presented first, drawing largely from the forensic psychology literature. In Section 7.06.4, contributions of forensic and cognitive psychology to thinking about competence are reviewed. The ways in which conceptualizations have been put into practice are discussed in Section 7.06.5.

7.06.3.1 Definitions of Competence

Definitions current in the late 1990s, are most readily understood when considered in relation to recent historical developments. These developments are briefly reviewed below, and Table 1 provides a summary of this discussion. As shown in the table, it is possible to trace a progression from early definitions which included a requirement that the choice be rational to later definitions which dropped the requirement for a rational choice ("Rationality" in the table) and added a requirement of awareness or insight ("Appreciation" in the table). Each of these individual definitional elements requires some explanation.

Definitions of competence have historically been tied to legal definitions of informed consent. Because proper disclosure of information is central to legal requirements for informed consent, early definitions of competence centered around understanding of the disclosed information. Roth, Meisel, and Lidz (1977) found most definitions required that "an individual must be able to comprehend the nature of the particular conduct in question and to understand its quality and its consequences" (p. 279). Appelbaum and Grisso (1988) identify understanding as "the concept referred to most often in legal cases and statutes" (p. 1635). As

Table 1 Elements of definitions of decisional competence from selected sources, 1977–1995.

	Definition Elements			
	Understanding	*Reasoning*	*Appreciation*	*Rationality*
Rationality-dependent definitions				
as noted in Roth et al. (1977)	X	X		X
Lidz et al. (1984)	X			X
Stanley et al. (1984)	X	X		X
Combination Definitions				
Miles (1989)	X		X	X
Draper and Dawson (1990)	X		X	X
Appreciation-dependent definitions				
Appelbaum and Roth (1982)	X	X	X	
Tepper and Elwork (1984)	X	X	X	
Kapp (1988)	X	X	X	
Appelbaum and Grisso (1995)	X	X	X	

Kelman (1982) states: "Voluntary, informed consent is illusory if the participants lack the capacity to evaluate the information" (p. 62). It is so central to competence, in fact, that "most researchers have implicitly assumed that an understanding of treatment information is the sole criterion of competence" (Tepper & Elwork, 1984, p. 205) . Understanding therefore appears as a standard in every definition of competence (see column 1 of Table 1). Of interest is the appearance of other standards and the rationale for their inclusion.

In general, earlier definitional systems included reference to a standard of reasonableness or rationality of the choice. Because of problems in separating politically or socially subjective standards of rationality from competence assessment, the rationality standard is often excluded from more recent systems, although rationality of the *reasons* for the choice is still included in many formulations of competence. Rationality of the reasons behind the choice requires both reasoning ability and some level of appreciation of the relevance of the decision to oneself. Rationality of reasons is a component of reasoning quality, and systems which include this criterion are listed under "Reasoning" (column 2) in Table 1. Systems which evaluate rationality of the decision itself are listed under "Rationality" (column 4). Those systems which attend to individuals' appreciation of the relevance of the information for their own situation (or at least acknowledgment of illness) are indicated under "Appreciation" (column 3).

As Table 1 indicates, definitions of competence tend to take one of two forms, labeled here as rationality-dependent and appreciation-dependent types.

7.06.3.1.1 Rationality-dependent definitions

One example of an early formulation is presented by Roth et al., (1977). They cite the following four standards as ones in use at the time of their writing: evidencing a choice, reasonableness of choice, rationality of reasons for choice (which requires an ability to understand), and actual understanding. Their presentation is not an endorsement, and in fact they point out some of the problems with each standard, particularly with the reasonableness standard.

Building from the earlier work of Roth and colleagues, Lidz et al., (1984) define competence as "the legal presumption that a patient has the capacity to comprehend the information the doctor discloses" (p.22). In addition to understanding of disclosed information, they cite rationality of the decision and rational reasoning ability as additional required elements.

Stanley et al., (1984) present a substantially similar definition to Lidz et al. they include comprehension, rationality of the choice, and quality of reasoning as required elements for competence. Of interest in their system is the criterion they use for evaluating quality of reasoning: evidence of weighing risks and benefits. This is a criterion that appears in later—and more sophisticated—systems as well.

7.06.3.1.2 Combination definitions

Like Stanley et al. (1984), Miles and Gomez Carlos (1989) define competence as "an ability to learn and repeat risks and benefits and make personal choices in light of that knowledge."

Also like Stanley et al., they include a decisional component in addition to an understanding component. Rationality of the choice is also part of his framework. An important difference, however, is that Miles and Gomez Carlos includes appreciation of the relevance of the information as a standard.

In 1990, Draper and Dawson presented their definition of competence to consent to treatment. Like Miles and Gomez Carlos's system, appreciation of the situation is included, but no reference is made to reasoning ability. Surprisingly, despite a warning to remember that "legal competence is not identical to rationality, that rationality is not a necessary condition of legal competence" (p. 286), rationality is included in their system, in which the rationality of the choice itself—rather than rationality of the reasons for the choice—is required for competence.

7.06.3.1.3 Appreciation-dependent definitions

Only those definitional systems that exclude rationality of the choice as a standard, and include appreciation of the relevance of the information for oneself are considered representative of the second definitional type, the appreciation-dependent definition. The appreciation requirement excludes as incompetent individuals with delusional thought about the relationship between the informed consent disclosure and their own lives. The appreciation-dependent definitional type is represented by the work of Appelbaum and Grisso (1995), Appelbaum and Roth (1982), Kapp (1988), and Tepper and Elwork (1984).

Most researchers agree that individuals out of touch with reality fall outside a credible definition of competence. The appreciation and rationality standards are both attempts to include a contact with reality requirement in a definition of competence. The rationality-dependent definition, by using rationality of the choice as the measure of contact with reality, allows for societal norms and evaluator subjectivity to intrude on the assessment of competence and it implicitly assumes that rationality can be objectively evaluated, or at least that there is an absolute standard of rationality. The appreciation-dependent type of definition requires no such assumptions. Competence is free to remain a relative term, accommodating idiosyncratic aspects of the individual's thinking, with the exception of delusional thinking. An individual needs to demonstrate some understanding of the disclosed information and some ability to think about that information in a rational manner. But individuals must also possess insight into the fact that informed consent represents a choice point: they must be aware that the disclosed information is relevant to the choice in question.

Kapp (1988) presents a system which includes rationality of reasons for the choice as a standard, but significantly he does not include a criterion of reasonableness of the choice. He labels his system a "functional approach" in which the focus is on individual ability. In addition to rationality of reasons, the following standards are cited as requirements: the ability to make and communicate a choice, the ability to reason about the choice, and understanding of relevant information and appreciation of the relevance of that information for oneself. Kapp's system is noteworthy because, like Appelbaum and Roth's (1982), it adds the dimension of appreciation of the situation.

Appelbaum and Grisso (1995) present an appreciation-dependent definition of competence as well. Understanding, rational manipulation of disclosed information (i.e., reasoning), and appreciation of the relevance of the disclosed information for one's own situation are all required elements. This particular definition will be discussed in greater detail in Section 7.06.5.1, below.

7.06.3.2 Other Considerations for Definitions

As discussed earlier, requiring that the reasons for a choice be rational still allows for objective assessment. Requiring that the *choice* be rational implicitly requires a subjective value judgment on the part of the assessor. Recognition of the interaction of social and legal requirements with definitions of competence represents an improvement in definitional schemes. For some, this interaction requires some means of minimizing the role of socially based values—like rationality—from competence determinations. One way is to use an individual-centered view of competence. An example of such a system is Drane's (1985) recommendation of the use of a sliding scale, in which the threshold for each of the relevant standards is a function of the nature of the situation. This relativistic position is also expressed by Katz, who recommends shared decision making (from Faden & Beauchamp, 1986); Moody's (1988) idea of "negotiated consent" (discussed above) is similarly individual-centered. In these views, an individual may be competent to make decisions at a low level, that is, when the decision does not require truly "autonomous action" (Faden & Beauchamp, 1986). Higher thresholds may be surpassed with assistance. Setting the appropriate threshold, however, involves moral judgments. While some researchers deal with the need for thresholds (e.g., Wang & Ennis, 1986), few acknowledge the dependence on socially determined, subjective standards.

A more complex view of competence is presented by Grisso (1994), in which he addresses the issue of threshold setting and subjectivity explicitly. He discusses four essential components of competence: functional, causal, interactive, and judgmental. The functional component centers around the extent to which an individual can "understand or do that which is relevant." The extent to which an individual fulfills this requirement can be assessed either functionally (in practice) or inferentially (based on clinical assessment). The causal component requires presence of a "causal explanation for any functional deficits." This component is an attempt to include objectivity in the definition. It stimulates exploration of potential causal disorders and it also brings attention to the remediability and relevance of those disorders. For example, if someone's competence is in question because of a diagnosis of Alzheimer's disease, to what extent do the consequences of that disease affect the individual's abilities? To what extent might some relevant consequences be corrected? The interactive component includes consideration of the individual's abilities in relation to the contextual demands on that individual. Grisso's interactive component is most similar to the context-specific competence described by Faden and Beauchamp (1986) and to the domain-specific competence of Sprung and Winick (1989).

Grisso's final component of competence recognizes the close relationship between individual-level and societal-level considerations. The judgmental component depends on societal norms, those criteria that society deems important in determinations of competence. The tension between societal obligations to protect individuals and societal obligations to respect individual autonomy characterize the judgmental component.

Legal and social context influences all competence-related activities. Like Grisso, Kopelman (1990) recognizes the impact of social and legal norms on definitions of competence; she states that competence cannot be defined without reference to those norms. The issue is best expressed by Faden and Beauchamp (1986) in their distinction between two senses of informed consent. To satisfy the first sense, an individual must be psychologically competent, meaning that relevant abilities are possessed. In the second sense, an individual must meet legal criteria (being over the age of majority, for example). The two senses may be assessed separately, but both are required under legal systems in the late 1990s. Their distinction between general and specific components of competence flows from this dichotomy.

Because of the confusion possible when applying general definitions of competence to specific situations, legal definitions are advancing from the general to the specific (Sullivan & Youngner, 1994). Grisso (1994) and Anderer (1990) note the trend in state statutes, and Appelbaum, Lidz, and Meisel (1987) observe that judicial decisions reflect this trend as well. The earlier definitions of competence are general, and thus are inadequate for application to specific situations. Attempts to refine and improve definitions of competence have resulted in attention to the dynamic nature of the abilities in question and their dependence on context for meaning (e.g., Grisso, 1994; Willis, 1996).

Other definitions also make the distinction between general and specific components of competence. Faden and Beauchamp (1986) say that competence is "the ability to perform a task," and this definition is invariant across all domains to which it is applied. This is the general component of the definition. They go on to note that criteria for determination of competence, however, vary by situation, and are therefore context-dependent. The dependence on context is their form of a specific component.

Sprung and Winick (1989) take a dual aspect approach to the definition of competence as well. Similar to Faden and Beauchamp's identification of a specific context-dependent competence, Sprung and Winick identify domain-specific incompetence. (Use of the term competence and incompetence is frequently mixed in most discussion, with incompetence implicitly assumed to represent the inverse of competence. This assumption remains unexamined, however, and arguably incompetence may be defined as something different from simple "noncompetence." For the purposes of this discussion, however, the focus remains on defining competence, and the term incompetence" is used only when the research in question uses that term.) However, Sprung and Winick's view differs from Faden and Beauchamp's in terms of the meaning of the general component. To Faden and Beauchamp, it is a characteristic of the definition which is general, that is, that all definitions of incompetence are similar because of their reliance on an inability. For Sprung and Winick, general incompetence means global inability. Individuals are incompetent in the general sense if they are unconscious or profoundly psychotic, for example. This latter view is more pervasive.

Lidz et al., (1984) make the distinction between two different types of incompetency in the same manner that Sprung and Winick do; general incompetence means global inability, and specific incompetence means inability to make a specific decision.

As Tepper and Elwork (1984) note, "development of an operational definition of any construct is, in part, a subjective endeavor" (p. 219), and as such will defy definitive resolution. Clearly, definitions of competence are evolving, accommodating increasing complexity and acknowledging the interaction between societal standards and individual abilities.

7.06.4 PSYCHOLOGICAL PERSPECTIVES ON COMPETENCE TO MAKE DECISIONS

In this section, psychological views of competence are examined in greater detail. The contributions of two distinct psychological perspectives to the question of competence are reviewed in detail: the forensic psychological view, which originates from legal requirements, and the cognitive psychological view, which originates from the study of human cognitive processes.

Both perspectives share a disciplinary emphasis of psychology, the science of the mind and behavior. The differences in the views will be reviewed further, but one critical distinction must be made before further examination. The forensic psychological perspective is concerned with "research that can lead to the development of improved and empirically based criteria for decision-making with respect to a number of difficult law and mental health issues" (S. Shah, quoted in Voit, 1995, p. 10). Competence questions are framed in terms of what relevant skills the law requires; the law is the springboard for investigation. In contrast, the cognitive psychological perspective is concerned with identification and measurement of cognitive processes such as problem solving, reasoning, memory, and knowledge representation (Craik & Salthouse, 1992; Datan, Rodeheaver, & Hughes, 1987). Most of the 1980s and 1990s definitions of competence reviewed above derive from a forensic psychological perspective. The next subsection expands on the meaning of a forensic psychological view, followed by a brief review of cognitive psychological perspectives on competence. Finally, the contributions of these two psychological views are compared.

7.06.4.1 Forensic Psychological Views of Competence

The forensic psychological view of competence is most often used by individuals with a combination of legal and psychological training. In most cases, these researchers have the express goal of enhancing psychological practice by attending to relevant legal issues and enhancing legal policy through empirical assessment. The forensic psychological view is distinguished from purely legal views by its emphasis on quantification of relevant variables.

One prominent feature of forensic psychological definitions is reliance on the term "decisional capacity" rather than "competence" as a means of improving specificity (Appelbaum & Grisso, 1988; President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, 1983). As discussed in the preceding review of definitions, a trend towards increased specificity is evident. Similarly, the American Bar Association (1991) notes the evolution in the legal concept of competence, from a global conception to a conception based on functional abilities; the American Bar Association report even recommends use of the specific term "decisional capacity" (instead of the more global term "competence") in legal contexts.

The contribution of forensic psychology to definitions of competence is evident in the review of definitions presented above. For example, the forensic psychological view has advanced definitions from a simple, global level centered around understanding to a more complex, hierarchical set of standards or component abilities. The work of Appelbaum and Grisso (1995), Appelbaum and Roth (1982), Kapp (1988), and Tepper and Elwork (1984) all represent forensic psychological views. All share a sensitivity to the process of mapping legal notions of informed consent to actual situations. One of the main contributions of this work is the foundation it lays for assessment of legally based constructs in clinical situations. The assessment approaches developed within this perspective will be reviewed later in the chapter.

7.06.4.2 Cognitive Psychological Views of Competence

Cognitive psychology has made distinct contributions to the clarification of competence. Several caveats must be noted before reviewing this literature, however. First, "competence" has numerous meanings in cognitive psychology, only some of which relate to the informed consent component of interest here. For example, the subject index of an edited volume in cognitive psychology includes a listing for "competence" but says "see also Self-efficacy" (Lachman & Barrack, 1993). Similarly, in Sternberg and Wagner's (1986) volume *Practical Intelligence*, several different types of competence are referenced in the subject index, including cognitive, social, intellectual, and interpersonal. Applying a cognitive psychology perspective to the examination of competence to provide informed consent therefore requires vigilance regarding the concepts actually being discussed.

The literature on everyday problem solving has some important parallels to the development of the concept of competence to provide informed consent. Although definitions vary slightly, in general, everyday problem solving is defined as intelligence directed at practical problems, distinct from abstract or "academic" intelligence (Sternberg & Wagner 1986). Measures of academic intelligence have been characterized as holding little intrinsic interest for the individual, being disconnected from the individual's daily experience, and often having a single correct answer (Sternberg & Wagner). In contrast, practical intelligence has as a goal "the solution of problems which challenge the well-being, needs, plans, and survival of the individual" (Charlesworth, 1976, quoted in Sternberg & Wagner, p. 52).

Willis and Schaie (1994) list the following five aspects of everyday problem solving: relevant mental abilities and skills; domain-specific knowledge; understanding of personal circumstances and of interpersonal context; attitudes, beliefs, and preferences; and integration of these dimensions. An important feature of this type of problem solving is recursiveness, that is, solutions or decisions may be arrived at in a nonlinear manner (Willis, 1996).

Two of cognitive psychology's everyday problem-solving dimensions are notably absent from forensic conceptualizations, however. Within the forensic framework, the role of attitudes, beliefs, and preferences is not addressed, nor is the potential for recursiveness in the decision process. Arguably, legal requirements necessitate a single response on a single occasion, and the importance of attitudes, beliefs, and preferences to the decision is potentially minimized by this single-time focus: each may change over time, but if a one-time decision is the focus of the law, then at most the law will be concerned with the influence of each on the present decision. A single-time focus also obviates a temporally recursive perspective.

Among the core issues addressed by cognitive psychology are memory and attention, both of which are relevant to a consideration of competence. Not surprisingly, findings from cognitive psychology are immediately applicable to competence concerns. Salthouse (1996) notes that mental screening tests are one method of assessing "cognitive competency, and indirectly the capacity to make responsible decisions" (p. 36). Using the Mini Mental State Examination (MMSE, Folstein et al., 1975) as an example, he notes that comprehension, short-term memory, and the ability to attend and concentrate are among the skills assessed by the MMSE. These skills are the minimum needed for "reasonable decision making," and thus are applicable to views of competence. Salthouse identifies the main similarities between forensic psychology and cognitive psychology views of competence when he discusses competence assessment, which he says "should not only be context-sensitive or situation-specific but should also incorporate evaluations of the individual's comprehension and reasoning relevant to the decision" (p. 13).

Advances in conceptualizations of competence have led to attempts to assess competence empirically, since assessment follows from, and in some cases feeds back to, conceptualization. Assessment of competence is considered next.

7.06.5 ASSESSMENT OF COMPETENCE

Older adults are increasingly confronted with serious health decisions, yet little normative information about the decision-making abilities of older adults exists. The growth of the elderly population in the USA and the widespread availability of life-sustaining technology underscore the need for proper assessment of medical decision-making ability.

Assessment schemes depend on definitions of competence and the standards used to operationalize the concept. Several assessment schemes are examined below, and some of their problems are explored. Table 2 presents an overview of this discussion. As the table demonstrates, all systems reviewed here include at least some aspect of "Understanding." The systems differ in their attention to the other conceptual elements. Mental status is arguably an implied component of most, especially those which address memory and language elements explicitly. The rationality and appreciation elements are discussed for each example.

7.06.5.1 Assessment of Competence to Consent to Treatment

An example of an early competence assessment scheme is that of Stanley et al., (1984). Comprehension figures prominently, as does

Table 2 Conceptual elements addressed in the empirical assessment of competence (organized by type of competence).

	Conceptual Element						
	Mental status	*Understanding*	*Memory*	*Language*	*Reasoning*	*Rationality*	*Appreciation*
Type of competence							
Consent to treatment							
Stanley et al., 1984		X			X	X	
Wang and Ennis, 1986		X	X	X			
Draper and Dawson, 1990		X				X	X
Fitten and Waite, 1990		X			X		
Freedman et al., 1991		X	X	X	X		X
Janofsky et al., 1992		X					
Appelbaum and Grisso, 1995		X	X	X	X		X
Competency to stand trial							
Golding et al., 1984	X	X	X	X			X
Civil competence							
Schwartz and Barone, 1992	X	X	X	X			
Testamentary capacity							
Spar and Garb, 1992	X	X					

"quality of reasoning" indicated by weighing of risks and benefits. Absent from their scheme, however, is attention to the appreciation of the information to the subject's own situation. Another problem is their inclusion of a reasonableness of choice standard.

Wang and Ennis (1986) developed an elaborate "cognitive competency test" which includes subtests designed to assess memory, reading skill, verbal reasoning, and spatial orientation. The test is based on their multidimensional view of cognitive skills required for competence. While their assessment scheme addresses only global competence, they do recommend combining this assessment with a more focused assessment of "mental status." However, little mention is made of matching specific situational demands with specific skill tests.

Draper and Dawson (1990) present an example of another assessment scheme: a flowchart for determination of competency. The main decision nodes occur regarding understanding of illness and treatment alternatives. Of interest is the inclusion of a question about acknowledgement of illness but one problem with this assessment scheme, in addition to extensive reliance on understanding, is the reliance on rationality of choice for competence determination.

Fitten and Waite (1990) assessed understanding of various elements of the disclosure (the condition, the treatment, risks, benefits, and alternatives) and the "quality of patient's or subject's reasoning process" (p. 1718). These authors hint at a distinction between "bureaucratic" disclosure forms and forms geared toward patient understanding, but they do not expand upon the distinction. Prompting (cuing) helped their control group more than their hospitalized group. The hospitalized sample was impaired relative to controls on weighing of alternatives.

Like Wang and Ennis (1986), Freedman, Stuss, and Gordon (1991) identify attention, language, memory, and frontal lobe function as the "cognitive functions underlying competency" and base their suggestions for assessment on these four areas. Once these abilities are established, "then issues related to judgment and awareness must be critically evaluated" (p. 207). The result is a measure of global competence to the exclusion of situation-specific abilities.

A more recent system also relies upon understanding and little else. Janofsky, McCarthy, and Folstein (1992) base their assessment scheme on a quick measure of understanding of a disclosure. Comparison is made between their test instrument, the Hopkins Competency Assessment Tool (HCAT), and a forensic psychiatrist's assessment. Several criticisms have been made of HCAT. While it does assess understanding, other domains relevant to competence are unaddressed. Both Lavin (1992) and Sales (1992) note that reasoning or rational weighing of choices is ignored. Lavin also argues that relatively stable personal values are a required component of competence, one which HCAT does not measure. Englehart (1992) notes the lack of attention to the risk–benefit ratio of the choice in question, and Kaye (1992) notes that prognosis if no treatment is undertaken is not included in the HCAT disclosure vignette. Another problem of their assessment method is the potential for subjective error with their "gold standard" of clinician judgment. Thus, at best HCAT can be considered a measure of understanding only, and not a measure of a more global concept of competence.

Some of the work reviewed above suggests another crucial element to consider when assessing competence: context. Appelbaum, Mirkin, and Bateman (1981) observed that "The stress surrounding any hospitalization, whether psychiatric or medical, may contribute to a failure of rational modes of thought, as may the necessity for making critical decisions from a position of relative ignorance about the medical issues involved" (p. 1175). Empirical evidence supports this claim. According to Fitten and Waite (1990), the context of hospitalization impaired performance in general: in their sample of older adults, 28% of the hospitalized subjects had impaired performance on a vignette decision-making task despite no evidence of cognitive impairment on a standardized mental status exam or via clinical assessment (Fitten and Waite, 1990). Dellasega, Smyer, Frank, and Brown (1996) and Frank (1995) also found evidence of decision-making impairment despite relatively high mental status exam scores among their elderly hospitalized subjects. The specific causes of impairment (e.g. the subject's illness, medical interventions) are unknown, but these studies suggest that hospitalized older adults are one group "most vulnerable to decisional impairments" (Fitten & Waite, 1990).

Consideration of context in competence assessment is also crucial in long-term care settings. Increasingly the nursing home is recognized as a setting in which decisional capacity assessments are often required, and research suggests that three-quarters of nursing home residents exhibit some decisional capacity impairments (Goodwin, Smyer, & Lair,1995). Pruchno, Smyer, Rose, Hartman-Stein, and Henderson-Laribee (1995) discuss methods that are objective, reliable, and easily administered to assess decision-making ability of residents in nursing homes.

Appelbaum and Grisso (1995) have developed three separate instruments for the purposes of assessing competence to provide informed consent, or more specifically, medical treatment decision-making capacity. The Understanding of Treatment Disclosure (UTD) assesses capacity to understand relevant information. The Perception of Disorder (POD) assesses appreciation of relevant information for the individual's own situation. The Thinking Rationally About Treatment (TRAT) assesses the individual's ability to manipulate relevant information rationally. The TRAT represents an expansion of the quality of reasoning standard that is part of Stanley et al.'s (1984) scheme. Stanley and colleagues use comparison of risks and benefits as their indicator of quality of reasoning; the TRAT goes beyond that and includes other abilities as well: consequential thinking, complexity of thinking, and the ability to generate consequences for choices. The TRAT, POD, and UTD instruments provide a consistent method for assessment based on relevant legal criteria, but have not yet been normed for use with older adults.

7.06.5.2 Assessment of Other Types of Competence

Some researchers have recognized the need to match assessment to the specific competence ability under question (Grisso, 1994; Roth et al., 1977). As demonstrated above, the law increasingly recognizes that competence is not a global construct. Instead, competence may vary by content area within an individual. It is possible to find "islands of competent functioning" (Gutheil & Bursztajn, 1986) in an otherwise incompetent individual. Attention to the context of the assessment is therefore desirable.

Civil competence, testamentary capacity, and competency to stand trial are all distinct concepts from competence to consent to treatment, but important work has been done in assessment of each of these types, and that work is relevant to the discussion here. The assessment of both civil competence and competency to stand trial share elements with several approaches to assessing capacity to consent to treatment. The Community Competence Scale-Revised (CCS-R, Schwartz & Barone, 1992) was developed to assess civil competence, and the Interdisciplinary Fitness Interview (IFI, Golding, Roesch, & Schreiber, 1984) was developed to assess competency to stand trial. Both attempt to tie mental health issues to legal requirements by including assessment of psychological functioning. The CCS-R contains 17 subscales and the IFI contains 16, so while they clearly address global aspects of competence, the high degree of detail offers the potential for quite specific applications. Further, by attending to both psychological and legal issues, they represent an attempt to bridge the divide between Faden and Beauchamp's (1986) two senses of consent and competence.

Spar and Garb (1992) state that to assess competence one must determine global mental status, assess specific aspects of mental functioning relevant to the ability in question, determine implications of those abilities for the specific context, and finally, form a diagnostic impression. Although it was developed for the assessment of testamentary capacity rather than treatment, their assessment fits well with the four components discussed by Grisso (1994): diagnostic impression addresses the causal component, determining abilities relative to context addresses the interactive component, and the functional component is implied in the assessment of mental functioning. Given their goal of presenting assessment consonant with current legal and social systems, the judgmental component is implied.

Several problems exist in most 1990s assessment schemes. First, rarely does assessment address all of the specific domains relevant to competence, that is, causal processes and interactive components as well as functional abilities. Instead, as noted above, many systems are based on assessment of global competence. Individual-centered and contextually sensitive systems are rare. Second, few authors attend to the subjective components of their assessment schemes. As with definitions of competence, some assessment schemes allow societally determined standards of reasonableness or rationality to intrude. Third, criteria for success on an ability are often set casually rather than following formal investigation. The relative weight each ability should be given in an assessment is largely neglected. Fourth, there is a dearth of norms for older populations. Thus, when gauging the functioning of older people, it is difficult to assess their relative standing.

Increasingly, psychological and legal conceptualizations of competence are being tied together. Linking the two increases complexity, but conceptualizations of competence have been improved by this increased attention to the relationship between psychological and legal issues. Rigorous investigation of definitions, standards, and assessment methods will further refine and improve our understanding of competence.

Miller and Cugliari (1990) surveyed long-term care facilities in New York State regarding treatment policies and found that only 12% of responding facilities had written guidelines for the determination of decisional capacity in 1986. By 1988 that number grew to 48%. The content of those policies varied, but of interest is the trend toward increased attention to formalizing the assessment process through guidelines. The question, however, is what should be contained in those guidelines? Additional empirical investigation of the concept of competence is required before that question can be answered.

7.06.6 COMPETENCE IN ACTION: ADVANCE DIRECTIVES

In multiple contexts, from law courts to medical clinics, competence assessment is central to the protection of autonomy and fulfillment of society's obligation to protect individuals. Every day, competence of individuals—especially older adults—is brought into question, but in the early 1990s a new area emerged making decisional competence even more prominent. Because of advances in life-sustaining technology, and because of the increased cost of these technologies, social and governmental forces recognized the need for increased attention to decision making about life-sustaining technologies. Particular attention was focused on decision making in the event that an individual becomes incompetent, increasing concern for the proper assessment of competence. Because of their relevance to decision making by older adults, developments in end-of-life decision making are reviewed below. The Patient Self-determination Act (PSDA) is used as an example of a social context in which decision-making concerns are played out.

7.06.6.1 The Patient Self-determination Act

The PSDA became effective in December 1991, and requires that all institutional Medicare and Medicaid participants address the question of advance directives (ADs) with patients (Omnibus Budget Reconciliation Act, 1990). According to regulations published by the Health Care Financing Administrations (Medicare and Medicaid Programs, 1992), all competent individuals have the right to issue ADs. The capacity of an individual to create an AD is explicitly equated with the capacity of that individual to provide informed consent for medical treatments. However, the extent to which the two capacities may differ is unknown. One study directly compared AD and medical treatment decision-making capacity (Frank, 1995). While few quantitative differences emerged, the relationships of variables like cognitive status and verbal ability to the two types of decision making differed, suggesting that the decision types may be distinct. Given the significant consequences of end-of-life decisions, closer empirical attention to differences between capacities is necessary.

The need for instruments to assess decisional capacity is acute and growing (Fitten and Waite, 1990; Stanley, Stanley, Guido & Garvin, 1988; Kapp, 1992). The hospital is the focus of most serious medical interventions and thus hospitalized older adults are often called upon to make serious medical treatment decisions. The PSDA has drawn attention to an often neglected area of treatment decisions, those centered on life-and-death concerns. The mandates of that law (specifically the requirement to address AD issues with newly admitted hospital patients) increase the prominence of the hospital as the site for such life-and-death decisions for patients. Developments in the area of assisted suicide have focused additional attention on end-of-life decisions.

7.06.6.1.1 Knowledge and attitudes of the elderly about ADs

Zweibel and Cassel (1989) report criteria that older adults use in making hypothetical end-of-life treatment decisions. Prominent among them are the desire to avoid being a burden to family, the desire to avoid living with impaired cognitive function, and the desire to avoid pain.

According to self-report data, individual characteristics that most influence preferences regarding life-sustaining treatments of hospitalized elders are personal values, religion, and personal experience with illness (self or others) (Cohen-Mansfield, Droge, & Billig, 1992). Multiple studies have found that most patients want to discuss ADs (Emanuel, Barry, Stoeckle, Ettelson, & Emanuel, 1991; Gamble, McDonald, & Lichstein 1991; Lo, McLeod, & Saika, 1986; Shmerling, Bedell, Lilienfeld, & Delbanco, 1988; Smucker et al., 1993; Wolf et al., 1991).

The validity of an individual's preference regarding life-sustaining treatment is of concern for both ethical and legal reasons. One index of the validity of patients' preferences is the stability of those preferences over time. Several studies indicate high stability over time (cited in Wolf et al., 1991). However, the effect of several specific potential influences on preferences and their stability have been studied in detail, especially depression, prognosis information, and information framing.

(i) Depression

Because of concern about the potential for depression to invalidate decisions to avoid life-sustaining treatment, the relationship of depression to preference stability has been examined. In their examination of 43 older adults admitted to an inpatient psychiatric facility for major depression, Ganzini, Lee, Heintz, Bloom, and Fenn (1994) found an association between treatment for depression and preference for life-sustaining treatment: 26% of the subjects preferred more life-sustaining treatment following treatment for depression. However, these subjects were initially rated as more severely depressed than the rest of the sample. Among individuals with less severe depression, preferences remained relatively stable. In a separate study comparing depressed and nondepressed elderly patients, Lee and Ganzini (1994) found higher variability in stated preference over time among the depressed group. Of interest is their finding that their depressed subjects did not differ from control subjects in actual choices made (in either frequency or kind). A limitation of the Lee and Ganzini study is that it does not examine the extent to which change in preference is related to relevant legal standards for informed consent, like change in level of understanding of the information and/or change in ability to manipulate the information rationally. If the mechanism of action of depression on preferences is through these abilities, then assessment of the abilities, rather than of depression *per se*, is appropriate. The relationship between the abilities and level of depression must therefore be explored.

Additional studies are required to address the role of depression in preference stability, but the influence of depression on decisional capacity itself remains to be explored.

(ii) Prognosis information

As noted by Torian, Davidson, Fillit, Fulop, and Sell (1992), Do-Not-Resuscitate (DNR) orders indicating that in the event of a cardiac arrest, health providers should not attempt to resuscitate the patient, are unique among potential advance directives because of the necessity for quick response in the case of cardiac arrest, leaving no time for a decision change. They found that the majority of frail elderly on an inpatient geriatric medicine unit express a preference to have cardiopulmonary resuscitation (CPR) performed rather than forego it. However, patients in this study were not presented with prognosis information, another potential influence on patient preferences.

In an important experimental study of patient preferences for the DNR type of advance directive, Murphy et al. (1994) report that probability of survival does in fact influence patient preferences for CPR. They questioned ambulatory geriatric patients about desire for DNR orders in three different hypothetical situations: following an accident, in the event of an acute medical illness, and in the event of chronic illness. Over 40% of their sample expressed a desire for CPR to be performed following an accident or acute illness. Only 11% would desire CPR in the event of chronic illness. Next, subjects were asked their preferences for CPR following presentation of probabilities of survival to discharge following CPR for each of the three situations. Presentation of this information reduced by half the number of patients who would request CPR. These authors conclude that once presented with probability of survival information, most patients would choose not to undergo CPR.

Murphy et al. (1994) demonstrated that preference for CPR decreased with age: once the probability of survival was discussed, only a few subjects over age 85 would opt for CPR in any of the three hypothetical situations. Clearly, potential prognosis is a crucial factor to be considered in any advance directive decision, but this study leaves the influence of mode of information presentation unexplored.

(iii) Information framing

Evidence suggests that the manner in which information is presented may exert a strong influence on preference (Mazur & Merz, 1993; Redelmeier, Rozin, & Kahneman, 1993; Tymchuk, Ouslander, Rahbar, & Fitten, 1988). For example, the following two explanations present the same outcome: (i) 90 of 100 patients survive an operation and 34 are alive after five years; (ii) 10 of 100 patients die after surgery, and 66 die by the end of five years. Subjects express a preference for the procedure associated with the first version; the procedure associated with the second version is nonpreferred. Thus subject preference for a medical procedure is dependent upon the way in which outcome statistics are framed (Redelmeier et al., 1993).

Similarly, subjects prefer zero risk to any risk when deciding between two risk reduction dilemmas, regardless of the magnitude of risk reduction for the nonzero choice (reported in Redelmeier et al., 1993). Subjects are also more averse to loss than they are enthusiastic about gains. These effects highlight just some of the ways in which framing of information can influence subject choice. Any exploration of decisional capacity must therefore attend to manner of presentation.

7.06.6.1.2 Practical and ethical aspects of implementation of the PSDA

This subsection uses the PSDA as an example to review some of the concerns relevant to older adults' decision making, including timing of decision, standards for acceptance of decision, assessment of the validity of the decision, and coercion.

As required by the PSDA, patients must be approached at admission with information about ADs. As expressed by Wolf and colleagues (1991), "Implementation may result in drowning patients in written materials on admission, [and] insensitive and ill-timed inquiry into patients' preferences"; therefore "Concern ... persists about whether the time of admission or initial receipt of treatment is an appropriate moment to broach the topic of directives" (p. 1667).

Another potential problem with PSDA implementation relates to standards for acceptance of ADs. Some authors maintain that the strictness of the standard for accepting informed consent should be directly proportional to the seriousness of the consequences of the decision (Drane, 1985).

Drane (1985) strongly supports a contextual definition of competence. He outlines three levels of standards to be used based on the situation at hand. At the lowest level, decisional consequences are minimal, and awareness and assent are all that are required of the individual. With more serious decisions, understanding and expression of a choice based on medical outcomes is required. For the most stringent level, appreciation of the situation and rationality of the decision are required.

Roth et al. (1977) also outline a context-specific standard for competency tests, based on the risk–benefit ratio of treatment and whether the patient is accepting or refusing the treatment in question. The sliding scale notion has an important limitation, however: the degree of seriousness of the decision is determined by the assessor and not by the individual being assessed.

In the course of describing a treatment refusal scenario, Drane (1985) writes: "Objectively, the standard medical treatment is preferable to what she decides" (p.20). The preference is the author's. Elsewhere, he describes a voluntary psychiatric hospitalization scenario, in which the patient's consent to hospitalization is accepted as valid at a low threshold. A "higher degree of competence" will be required, however, if "more dangerous" treatment is required (p. 18). Drane's view is that there are treatments more serious than limiting one's own freedom via hospitalization. This is the author's opinion, but certainly different individuals may have different, but potentially equally valid, ranking systems for "evils" (Culver & Gert, 1982). Fitten and Waite (1990) also acknowledge the problem of assessor perspective, stating that the threshold for competence is influenced by the assessors' view of "normality" (p. 1720).

How should clinicians rate the seriousness of the decisions included in ADs? Should a lower standard of acceptance be applied to DNR orders than to withdrawal of nutrition and hydration requests? Is the "seriousness" hierarchy of medical interventions applicable to end-of-life issues?

Discussion of attitudes of older adults towards advance directives must also include reference to the role of families in medical decision making. Families are usually included in critical decisions by health care professionals as well as by the patient (Watts, 1992). In fact, Ely, Peters, Zweig, Elder, and Schneider (1992) demonstrated that clinicians will defer to the family, even in the presence of an advance directive document. Also, most state laws regarding surrogate decision making set up explicit hierarchies based on family relationship to the patient. The role of the family in decision making by older adults will be addressed further in the next section.

Another concern about PSDA implementation relates to a unique feature of ADs: one may decide not to decide, and instead designate a proxy decision maker. Alexander Capron has written that consent requires disclosure of information because information is "necessary to make meaningful the power to decide" (quoted in Freedman, 1975). The case of proxy designation in advance directives presents a special challenge to this interpretation of informed consent. By specifying another individual to make decisions for them, individuals may be engaging in "valid yet ignorant consent" (Freedman, 1975). In addition, AD decisions may be invalid due to subtle impairments in judgment or reasoning.

A related ethical concern raised by PSDA is the threat of involuntary rationing through coercion (Callahan, 1988; Jecker, 1988). Assessment of the extent to which an older adult's decision about ADs is truly informed and voluntary is therefore an important means of avoiding coercion.

7.06.7 OTHER ISSUES

7.06.7.1 Other Considerations for Decision Making

7.06.7.1.1 The nature of preferences

Any complete consideration of medical decision making requires attention to the nature of preferences and choices. What exactly is a choice? What is the definition of a preference? Available research suggests that static preferences are a myth (Slovic, 1995). Instead, a more useful conceptualization is that of preferences that are constructed "on the spot by an adaptive decision maker" (Slovic, 1995, p. 369). As Slovic notes, "decision making is a highly contingent form of information processing, sensitive to task complexity, time pressure, response mode, framing, reference points, and numerous other contextual factors" (p. 369). Consideration of the role of the family (discussed further below) represents one step towards contextual sensitivity to the decision-making process. Other factors need to be addressed, including the formation of personal values and their relationship to medical decisions and the influence of information from different sources on values and decision outcomes.

"Negotiated consent" has been suggested as a means of supporting patient autonomy (Moody, 1988); a closely related concept is that of "participative decision making" in which patient and clinician share responsibility for each phase of the decision-making process (Ballard-Reisch, 1990). However, if one accepts the premise that individuals actually construct decisions based on a variety of factors, including information framing, information presentation, and method of response elicitation (Slovic, 1995), then negotiated consent does not necessarily bring us closer to the goal of autonomous, valid consent. The ideal of legally valid informed consent requires a full comprehension of the disclosed information by the individual. Negotiated consent puts the individual at risk of making less than legally informed and valid choices based on a variety of factors, including the individual's relationship to others involved in the decision, and the manner in which information is portrayed.

In practice, decisions are not purely rational (Appelbaum, et al. 1987). Explorations of ways to improve the informed consent process must acknowledge the multiple sources of choice construction. As discussed earlier, depression, prognosis information, and information framing are among the factors that might influence decisions. More importantly, the potential for improving the nature of informed consent must be explored. Simple remedial measures should be tested. For example, expanding explanation of alternatives and treatment choices during a disclosure may improve understanding of information (Grisso & Appelbaum, 1995), and may ultimately improve the quality of the choice rendered.

7.06.7.1.2 The role of the family

As noted above, some older adults defer consent to participate until they can consult with a family member (Gueldner & Hanner, 1989) and many studies document the tendency for older adults to rely on family members in medical treatment decision situations (High, 1993; Kapp, 1991; Zweibel & Cassel, 1989). One potential reason for this preference may be related to a general lack of familiarity among older adults with technological advances. Reasons have not been explored in depth, but the ability of older adults to consider the issues involved in end-of-life decisions needs to be measured.

Because of evidence of the preference of some older adults for family involvement, some commentators suggest that society's emphasis on individual autonomy is misplaced in the case of older adults; instead, courts should adopt a "rebuttable presumption in favor of family surrogates" (High, 1991). High argues that by favoring autonomy over family involvement, the current system denies the role of caring concern on the part of the family in decision-making, and it underestimates the role of values shared by the family unit.

Kapp (1989) points out additional problems with an individual autonomy focus for older adults. The current view simply transfers decisional authority from the institution to the patient, with the danger of condoning neglect when older adults do not exercise their autonomy sufficiently. Instead, he argues, autonomy is best blended with beneficence, with "negotiated sharing" the result. In this way the institution's obligations towards the patient are not dismissed in the name of autonomy. Kapp makes the additional point that autonomy "is the capacity for self-determination" and the individual is free to choose to exercise autonomy or not.

Of interest is that many states in the USA have a "family consent" statute in which authority for medical treatment decisions is designated to family members, in a specific order based on relationship to the patients, in the event that the patient becomes unable to make or communicate his or her own choices. The law in most states therefore recognizes the prime role of family in health care decisions for incapacitated individuals.

In practice, evidence suggests that many older adults make decisions in collaboration with others, particularly with family members (Kapp, 1991). Among the reasons why older adults may choose to share decision making is to minimize family burden (Kapp, 1991), and concern among older adults about family burden has been documented (e.g., Zweibel and Cassel, 1989). Thus, assessment approaches require sensitivity to the role of the family for a full understanding of decision making by older adults.

7.06.7.2 Directions for Future Research

Assessment of decisional capacity requires solid theoretical formulations of decisional capacity as well as construction of valid and reliable assessment instruments. In the late 1990s, most of the theoretical formulations about decisional capacity stem from the legal requirements for informed consent. Instrument development requires that theoretical formulations be translated into measurable psychological constructs. This translation has proven difficult, however, given the differences between legal traditions and psychological traditions.

Requirements of legal informed consent stipulate that treatment choices and alternatives be discussed, along with risks and benefits of treatments. The doctrine of informed consent was developed through the legal tradition, in the absence of empirical testing. If certain elements of the informed consent disclosure are more difficult than other elements (in terms of subject understanding), then the legal goal of valid, informed consent may not be being met. If portions of the informed consent disclosure are truly more difficult than other portions for subjects to understand, then remedial measures are indicated. Either the doctrine of informed consent must be reconsidered, or special attention must be given to the communication of certain elements of the disclosure. Rather than relying on a set of criteria for informed consent, perhaps health care professionals should consider a more focused question and ask “informed consent for what?”

For informed consent for research, a sliding scale has been suggested (Levine, 1995). It may be the case that such a sliding scale is appropriate for medical treatment decisions as well. For example, if individuals can understand the symptoms of their disorder and the risks and benefits of a low-risk treatment, they can be considered competent to decide whether to accept or reject that treatment. However, if understanding is intact for symptoms, risks, and benefits, but impaired for other choices or alternatives, then a limited decisional autonomy may be appropriate. The positive effect of corrected feedback on understanding of informed consent information has been documented (Taub, Kline, & Baker, 1981). The extent to which simple feedback could improve performance on these measures should be addressed.

Application of additional measurement methods could enhance understanding of decisional capacity. For example, item response theory and factor analysis could contribute to identification of the distinct latent traits measured by each instrument. The relevance of those traits to legal requirements for competence could then be quantified. These methods could also reduce respondent burden of many instruments. Improved conceptualization and measurement awaits application of advanced methodologies to this area.

In addition to making consent procedures more situation-specific, researchers should aim to make the procedures more subject-specific as well. Identifying specific predictors of performance for groups of subjects (hospitalized adults, cognitively impaired patients, etc.) represents one strategy for matching the subject and the task. Another strategy involves identification of individual-level differences, such as differences in verbal ability or level of depressive symptoms. Quantifying the effects of these differences on the informed consent process could lead to improvements in the match of individual to context.

When considering “older adults” it is necessary to remain aware of heterogeneities within the population. As Murphy et al. (1994) found, adults over the age of 85 were distinct from other older adults in terms of reaction to a CPR scenario (see Section 7.06.6 for a synopsis). Different age groups may have a different constellation of responses to decision-making tasks. Further research is needed to identify the range of responses across a broader age range.

Many studies explore decisional capacity at a single time point. However, a single-time study is inadequate for full characterization of decisional capacity for several reasons. First, many studies document the susceptibility of hospitalized older adults to transient shifts in cognitive functioning (e.g., Francis, Martin, & Kapoor, 1990). Thus the role of time in obtaining informed consent is particularly important for this population. Second, a single-time study provides no information about the developmental nature of the phenomenon under study. Both short-term changes (including those resulting from training or other intervention) and long-term changes should be explored. There may be important differences between an individual's decision making before and after a health crisis; a prospective study could address this question. A longitudinal examination of the same subjects could provide insight into another important aspect of decision-making: the relationship between the stability of patient preferences and the validity of the patient's choices. There is a growing literature on the stability of patient preference over time (e.g., Emanuel et al., 1991; Everhart and Pearlman, 1990), but few studies address both AD decision and medical treatment decision stability.

A body of research exists about adolescent treatment decision making (e.g., Ambuel & Rappaport, 1992; Ginsburg et al., 1995; Lewis, 1987; Mann, Harmoni, & Power, 1989; Scott, Repucci, & Woolard, 1995). Much could be gained theoretically and empirically by linking work on the development of decisional capacity with work on the expression of decisional capacity in later life. Also, some of the questions about the role of age and health status in decision making should be addressed.

The goal of obtaining legally and ethically valid informed consent for treatment is a worthy one, but many commentators agree that practices in the late 1990s fall short of this goal (Culver & Gert, 1982; Mazur, 1988; Rodwin, 1989; Taub, Baker, & Sturr, 1986; Tymchuk & Ouslander, 1990). Improving our understanding of the many issues involved in decision-making by older adults will improve the decision-making process. The aging of our society assures that these issues will have increasing personal and public policy relevance.

7.06.8 REFERENCES

Ambuel, B., & Rappaport, J. (1992). Developmental trends in adolescents' psychological and legal competence to consent to abortion. *Law and Human Behavior, 16*(2), 129–154.

American Bar Association. (1991). *Patient Self-determination Act. State law guide*. Washington DC: ABA Commission on Legal Problems of the Elderly.

Anderer, S. J. (1990). *Determining competency in guardianship proceedings*. Washington, DC: American Bar Association.

Appelbaum, P. S., & Grisso, T. (1988). Assessing patients' capacities to consent to treatment. *New England Journal of Medicine, 319,* 1635–1638.

Appelbaum, P. S., & Grisso, T. (1995). The MacArthur Treatment Competence Study. I. Mental illness and competence to consent to treatment. *Law and Human-Behavior, 19*(2), 105–126.

Appelbaum, P. S., Lidz, C. W., & Meisel, A. (1987). *Informed consent: Legal theory and clinical practice*. New York: Oxford University Press.

Appelbaum, P. S., Mirkin, S. A., & Bateman, A. L. (1981). Empirical assessment of competency to consent to psychiatric hospitalization. *American Journal of Psychiatry, 138*(9), 1170–1176.

Appelbaum, P., & Roth, L. H. (1982). Competency to consent to research. A psychiatric overview. *Archives of General Psychiatry, 39*, 951–958.

Ballard-Reisch, D. S. (1990). A model of participative decision making for physician–patient interaction. *Health Communication, 2*(2), 91–104.

Buehler, D. A. (1990). Informed consent and the elderly. An ethical challenge for critical care nursing. *Critical Care Nursing Clinical of North America, 2(3)*, 461–471.

Callahan, J. C. (1988). Acts, omissions, and euthanasia. *Public Affairs Quarterly, 2*(3), 21–36.

Canterbury v. Spence, 464 F.2d 772, (D.C. Cir. 1972).

Cohen E. S. (1996). Resolving ethical dilemmas arising from diminished decision-making capacity of the elderly. In M. Smyer, K. W. Schaie, & M. B. Kapp. (Eds.) *Older adults' decision-making and the law* (pp. 162–174). New York: Springer.

Cohen-Mansfield, J., Droge, J. A., & Billig, N. (1992). Factors influencing hospital patients' preferences in the utilization of life-sustaining treatments. *Gerontologist, 32*(1), 89–95.

Craik, F. I. M., & Salthouse, T. (Eds.) (1992). *The handbook of aging and cognition*. Hillsdale, NJ: Erlbaum.

Culver, C. M. & Gert, B. (1982). *Philosophy in medicine. Conceptual and ethical issues in medicine and psychiatry*. New York: Oxford University Press.

Datan, N., Rodeheaver, D., & Hughes, F. (1987). Adult development and aging. In M. R. Rosenzweig & L. W. Porter (Eds.). *Annual Review of Psychology,* (Vol. 38, pp. 153–80). Palo Alto, CA: Annual Reviews Inc.

Dellasega, C., Smyer, M., Frank, L., & Brown, R. (1996). Decision-making capacity in the acutely ill elderly. In M. Smyer, K. W. Schaie, and M. B. Kapp. (Eds.) *Older adults' decision-making and the law* (pp. 112–161). New York: Springer.

Drane, J. F. (1985). The many faces of competency. *Hastings Center Report, 19*(2),17–21.

Draper, R. J., & Dawson, D. (1990). Competence to consent to treatment: A guide for the psychiatrist. *Canadian Journal of Psychiatry*, *35*(4), 285–289.

Ely, J. W., Peters, P. G., Zweig, S., Elder, N., & Schneider, F. D. (1992). The physician's decision to use tube feedings: The role of the family, the living will, and the Cruzan decision. *Journal of the American Geriatrics Society, 40,* 471–475.

Emanuel, L. L., Barry, M. J., Stoeckle, J. D., Ettelson, L. M., & Emanuel, E. J. (1991). Advance directives for medical care—a case for greater use. *New England Journal of Medicine*, *324*, 889–95.

Englehart, J. (1992). Assessing competency. (Letter to the editor). *Hospital and Community Psychiatry*, *43*(6), 647–648.

Everhart, M. A., & Pearlman, R. A. (1990). Stability of patient preferences regarding life-sustaining treatment. *Chest, 97*(1), 159–164.

Faden, R. R., & Beauchamp, T. L. (1986). *A history and theory of informed consent*. New York: Oxford University Press.

Fitten, L. J., & Waite, M. S. (1990). Impact of medical hospitalization on treatment decision-making capacity in the elderly. *Archives of Internal Medicine*, *150*, 1717–1721.

Folstein, M., Folstein, S., & McHugh, P. (1975). Mini Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198.

Francis, J., Martin, D., & Kapoor, W. (1990). A prospective study of delirium in hospitalized elderly. *Journal of the American Medical Association, 263*(8), 1097–1101.

Frank, L. B. (1995). *Psychological and legal considerations* *in the assessment of decision-making capacity of hospitalized older adults*. Pennsylvania State University. Dissertation. State College, PA.

Freedman, B. (1975). A moral theory of informed consent. *Hastings Center Report*, *5(*4), 32–39.

Freedman, M., Stuss, D. T. and Gordon, M. (1991). Assessment of competency: the role of neurobehavioral deficits. *Annals of Internal Medicine, 115*(3), 203–208.

Gamble, E. R., McDonald, P. J., & Lichstein, P. R. (1991). Knowledge, attitudes, and behavior of elderly persons regarding living wills. *Archives of Internal Medicine*, *151*, 277–280.

Ganzini, L., Lee, M. A., Heintz, R. T., Bloom, J. D., & Fenn, D. S. (1994). The effect of depression treatment on elderly patients' preferences for life-sustaining medical therapy. *American Journal of Psychiatry*, *151*(11), 1631–1636.

Ginsburg, K. R., Slap, G. B., Cnaan, A., Forke, C. M., Balsley, C. M., & Rouselle, D. M. (1995). Adolescents' perceptions of factors affecting their decisions to seek health care. *Journal of the American Medical Association, 273*(24), 1913–1918.

Golding, S. L., Roesch, R., & Schreiber, J. (1984). Assessment and conceptualization of competency to stand trial. Preliminary data on the Interdisciplinary Fitness Interview. *Law and Human Behavior, 8*(3/4), 321–334.

Goodwin, P. E., Smyer, M. A., & Lais, T. I. (1995). Decision-making incapacity among nursing home residents: Results from the 1987 NMES survey. *Behavioral Sciences and the Law, 13,* 405–414.

Grisso, T. (1994). Clinical assessments for legal competency of older adults. In M. Storandt & G. P. Vanden Bos (Eds.), *Neuropsychological assessment of dementia and depression in older adults: A clinician's guide* (pp. 119–139). Washington, DC: American Psychological Association.

Grisso, T., & Appelbaum, P. S. (1995). The MacArthur Treatment Competence Study: III. Abilities of patients to consent to psychiatric and medical treatments. *Law and Human Behavior*, *19*(2), 149–174.

Gueldner, S. H., & Hanner, M. B. (1989). Methodological issues related to gerontological nursing research. *Nursing Research*, *38*(3), 183–185.

Gutheil, T. G., & Bursztajn, H. (1986). Clinicians' guidelines for assessing and presenting subtle forms of patient incompetence in legal settings. *American Journal of Psychiatry*, *143*, 1020–1023.

High, D. M. (1991). A new myth about families of older people? *Gerontologist*, *31*(5), 611–618.

High, D. M. (1993). Advance directives and the elderly: A study of intervention strategies to increase use. *Gerontologist*, *33*(3), 342–349.

Janofsky, J. S., McCarthy, R. J., & Folstein, M. F. (1992). The Hopkins Competency Assessment Test: A brief method for evaluating patients' capacity to give informed consent. *Hospital and Community Psychiatry*, *43*(2), 132–136.

Jecker, N. S. (1988). Disenfranchising the elderly from life-extending medical care. *Public Affairs Quarterly, 2*(3), 51–58.

Kapp, M. B. (1988). Decision making by and for nursing home residents: A legal view. *Clinics in Geriatric Medicine, 4*(3), 667–679.

Kapp, M. B. (1989). Medical empowerment of the elderly. *Hastings Center Report, 19*(4), 5–7.

Kapp, M. B. (1991). Health care decision making by the elderly: I get by with a little help from my family. *Gerontologist*, *31*(5), 619–623.

Kapp, M. B. (1992). *Geriatrics and the Law. Patient rights and professional responsibilities* (2nd ed.). New York: Springer.

Kaye, N. S. (1992). Assessing competency. (Letter to the editor). *Hospital and Community Psychiatry*, *43*(2), 132–136.

Kelman, H. C. (1982). Ethical issues in different social science methods. In T. L. Beauchamp, R. R. Faden, R. J. Wallace, Jr., & L. Walters (Eds.), *Ethical Issues in Social Science Research* (pp. 40–98). Baltimore: Johns Hopkins University Press.

Kopelman, L. M. (1990). On the evaluative nature of competency and capacity judgments. *International Journal of Law and Psychiatry, 13*(4), 309–329.

Lachman, M., & Barrack, O. R. (1993). Planning and control processes across the life span: An overview. *International Journal of Behavioral Development, Special Issue. 16*(2), 129–384.

Lavin, M. (1992). Assessing competency (Letter to the editor). *Hospital and Community Psychiatry*, *43*(2), 132–136.

Lee, M., & Ganzini, L. (1994). The effect of recovery from depression on preferences for life-sustaining therapy in older patients. *Journal of Gerontology: Medical Sciences*, *49*(1), M15–M21.

Levine, R. J. (1995, January). *Proposed regulations for research involving those institutionalized as mentally infirm: A consideration of their relevance in 1994*. Paper presented at Ethics in Neurobiological Research with Human Subjects, Baltimore.

Lewis, C. C. (1987). Minors' competence to consent to abortion. *American Psychologist*, *42*(1), 84–88.

Lidz, C. W., Meisel, A., Zerubavel, E., Carter, M., Sestak, R. M., & Roth, L. H. (1984). *Informed consent. A study of decision making in psychiatry*. New York: Guilford.

Lo, B., McLeod, G. A., & Saika, G. (1986). Patient attitudes to discussing life-sustaining treatment. *Archives of Internal Medicine, 146*(8), 1613–1615.

Mann, L, Harmoni, R., & Power, C. (1989). Adolescent decision making: The development of competence. *Journal of Adolescence*, *12*(3), 265–278.

Mazur, D. J. (1988). Why the goals of informed consent are not realized: Treatise on informed consent for the primary care physician. *Journal of General Internal Medicine, 3,* 370–380.

Mazur, D. J., & Merz, J. F. (1993). How the manner of presentation of data influences older patients in determining their treatment preferences. *Journal of the American Geriatrics Society*, *41*, 223–228.

Medicare and Medicaid Programs; Advance Directives, 57 Fed. Reg. 8194-01 (1992).

Miles, S. H., & Gomez Carlos, F. (Eds.) (1989). *Protocols for elective use of life-sustaining treatments: A design guide*. New York: Springer.

Miller, T., & Cugliari, A. M. (1990). Withdrawing and withholding treatment: policies in long-term care facilities. *Gerontologist*, *30*(4), 462–468.

Mohr v. Williams, 95 Minn. 261, 104 N.W. 12 (1905).

Moody, H. R. (1988). From informed consent to negotiated consent. *Gerontologist, 28*(Supp.), 64–70.

Murphy, D. J., Burrows, D., Santilli, S., Kemp, A. W., Tenner, S., Kreling, B., & Teno, J. (1994). The influence of the probability of survival on patients' preferences regarding cardiopulmonary resuscitation. *New England Journal of Medicine*, *330*, 545–549.

Natanson v. Kline, 186 Kan. 393, 350 (1960).

Omnibus Budget Reconciliation Act of 1990, Pub. L. No. 101-508, Sections 4206,4751. (1990).

President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research (1983). *Deciding to forego life-sustaining treatment*, 51–52. Washington, DC: Government Printing Office.

Pruchno, R. A., Smyer, M. A., Rose, M. S., Hartman-Stein, P. E., & Henderson-Laribee, D. L. (1995). Competence of long-term care residents to participate in decisions about their medical care: A brief objective assessment. *The Gerontologist, 35*(5), 622–629.

Redelmeier, D. A., Rozin, P., & Kahneman, D. (1993). Understanding patients' decisions. Cognitive and emotional perspectives. *Journal of the American Medical Association*, *270*(1), 72–76.

Rodwin, M. A. (1989). Physicians' conflicts of interest: The limitations of disclosure. *The New England Journal of Medicine, 321*(20), 1405–1408.

Roth, L. H., Meisel, A., & Lidz, C. W. (1977). Tests of competency to consent to treatment. *American Journal of Psychiatry*, *134*(3), 279–284.

Sales, G. N. (1992). Assessing competency. (Letter to the editor.) *Hospital and Community Psychiatry*, *43*(6), 646.

Salgo v. Leland Stanford Jr. University Board of Trustees, 317 P.2d 170, 181 (1957).

Salthouse, T. (1996). Commentary: A cognitive psychologist's perspective on the assessment of cognitive competency. In M. Smyer, K. W. Schaie, and M. B. Kapp (Eds.), *Older adults' decision making and the law*, (pp. 29–39). New York: Springer.

Schloendorff v Society of New York Hospitals. 211 NY 125, 126, 105, NE 92, 93. (1914).

Schwartz, J. J., & Barone, D. F. (1992). Assessing civil competence in the elderly. *Journal of Forensic Sciences*, *37*(3), 938–941.

Scott, E. S., Reppucci, N. D., & Woolard, J. L. (1995). Evaluating adolescent decision making in legal contexts. *Law and Human Behavior*, *19*(3), 221–244.

Shmerling, R. H., Bedell, S. E., Lilienfeld, A., & Delbanco, T. L. (1988). Discussing cardiopulmonary resuscitation: A study of elderly outpatients. *Journal of General Internal Medicine, 3*(4), 317–221.

Slovic, P. (1995). The construction of preference. *American Psychologist*, *50*(5), 364–371.

Smucker, W. D., Ditto, P. H., Moore, K. A., Druley, J. A., Danks, J. H., & Townsend, A. (1993). Elderly outpatients respond favorably to a physician-initiated advance directive discussion. *Journal of the American Board of Family Practice, 6*(5), 473–482.

Spar, J. E., & Garb, A. S. (1992). Assessing competency to make a will. *American Journal of Psychiatry*, *149*(2), 169–174.

Sprung, C. L., & Winick, B. J. (1989). Informed consent in theory and practice: Legal and medical perspectives on the informed consent doctrine and a proposed reconceptualization. *Critical Care Medicine, 17*(12), 1346–1354.

Stanley, B., Guido, J., Stanley, M., and Shortell, D. (1984). The elderly patient and informed consent. Empirical findings. *Journal of the American Medical Association*, *252*(10), 1302–1306.

Stanley, B., Stanley, M., Guido, J., and Garvin, L. (1988). The functional competency of elderly at risk. *Gerontologist*, *28*(Suppl.), 53–58.

Sternberg, R. J., & Wagner, R. K. (Eds) (1986). *Practical intelligence*. New York: Cambridge University Press.

Sullivan, M. D., & Youngner, S .J. (1994). Depression, competence, and the right to refuse lifesaving medical treatment. *American Journal of Psychiatry*, *151*(7), 971–978.

Taub, H. A. (1980). Informed consent, memory and age. *Gerontologist*, *20*(6), 686–690.

Taub, H. A., Baker, M. T., & Sturr, J. F. (1986). Informed consent for research. Effects of readability, patient age, and education. *Journal of the American Geriatrics Society, 34*(8), 601–606.

Taub, H. A., Kline, G. E., & Baker, M. T. (1981). The elderly and informed consent: effects of vocabulary level and corrected feedback. *Experimental Aging Research, 7*(2), 137–146.

Tepper, A. M.., & Elwork, A. (1984). Competence to consent to treatment as a psycholegal construct. *Law and Human Behavior*, *8*(3/4), 205–223.

Torian, L. V., Davidson E. J., Fillit, H. M., Fulop, G., & Sell, L.L. (1992). Decisions for and against resuscitation in an acute geriatric medicine unit serving the frail elderly. *Archives of Internal Medicine*, *152*(3), 561–565.

Tymchuk, A. J., & Ouslander, J. G. (1990). Optimizing the informed consent process with elderly people. *Educational Gerontology 16*, 245–257.

Tymchuk, A. J., Ouslander, J. G., Rahbar, B., & Fitten, J. (1988). Medical decision-making among elderly people in long term care. *Gerontologist*, *29*(Suppl.), 59–63.

Voit, E. S. (1995). Developing a research program: Saleem Shah's leadership role at the National Institute of Mental Health. *Law and Human Behavior 19*(1), 5–14.

Wang, P. L., & Ennis, K. E. (1986). Competency assessment in clinical populations: An introduction to the Cognitive Competency Test. In P. Uzzell & Y. Gross (Eds.), *Clinical neuropsychology of intervention* (pp. 119–133). Hingham, MA: Martinus Nijhoff.

Watts, A. T. (1992). The family's will or the living will: Patient self-determination in doubt. *Journal of the American Geriatrics Society, 40,* 533–534.

Willis, S. L. (1996). Assessing everyday competence in the cognitively challenged elderly. In M. Smyer, K. W. Schaie, and M. B. Kapp (Eds.), *Older adults' decision making and the law* (pp. 87–127). New York: Springer.

Willis, S. L., & Schaie, K. W. (1994). Assessing everyday competence in the elderly. In C. B. Fisher & R. M. Lerner (Eds.), *Applied developmental psychology* (pp. 339–372). New York: McGraw Hill.

Wolf, S. M., Boyle, P., Callahan, D., Fins, J. J., Jennings, B., Nelson, J. L., Barondess, J. A., Brock, D. W., Dresser, R., Emanuel, L., Johnson, S., Lantos, J., Mason, D. R., Mezey, M., Orentlicher, D., & Rouse, F. (1991). Special report: sources of concern about the patient self-determination act. *New England Journal of Medicine*, *325*, 1666–1671.

Zweibel, N. R., & Cassel, C. K. (1989). Treatment choices at the end of life: a comparison of decisions by older patients and their physician-selected proxies. *Gerontologist, 29*(5), 615–621.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.07

Neuropsychological Assessment of the Elderly

JOHN R. CRAWFORD and ANNALENA VENNERI
University of Aberdeen, UK

and

RONAN E. O'CARROLL
University of Stirling, UK

7.07.1 THE AIMS OF NEUROPSYCHOLOGICAL ASSESSMENT 134

7.07.2 SOME SPECIFIC CONSIDERATIONS IN NEUROPSYCHOLOGICAL ASSESSMENT OF THE ELDERLY 135

7.07.3 THE NEUROPSYCHOLOGICAL INTERVIEW 136

7.07.4 BASIC PSYCHOMETRIC ISSUES 137

- *7.07.4.1 The Rationale of Deficit Measurement in Clinical Practice* 137
- *7.07.4.2 Converting Scores to a Common Measurement Scale* 137
- *7.07.4.3 Reliability* 138
- *7.07.4.4 Reliable vs. Abnormal Differences* 138
- *7.07.4.5 Monitoring Change* 139

7.07.5 APPROACHES TO ASSESSMENT IN CLINICAL NEUROPSYCHOLOGY 140

7.07.6 SCREENING TESTS FOR MENTAL STATUS 141

7.07.7 MEASURES OF CURRENT INTELLECTUAL ABILITY 141

- *7.07.7.1 Use of the WAIS-R with the Elderly* 141
- *7.07.7.2 Analysis of Subtest Profiles* 142
- *7.07.7.3 WAIS-R Factor Scores* 144

7.07.8 SPECIFIC METHODS FOR THE ESTIMATION OF PREMORBID ABILITY 145

7.07.9 AGING AND MEMORY 146

- *7.07.9.1 Introduction to Aging and Memory* 146
- *7.07.9.2 Age-associated Memory Impairment* 147
- *7.07.9.3 Does Memory Functioning in Dementia Represent Accelerated Aging?* 147
- *7.07.9.4 Commonly Used Memory Tests: Implications for use with Elderly Subjects* 147
  - *7.07.9.4.1 The Auditory Verbal Learning Test* 147
  - *7.07.9.4.2 The California Verbal Learning Test* 148
  - *7.07.9.4.3 The Wechsler Memory Scale-Revised* 148
  - *7.07.9.4.4 The Rivermead Behavioural Memory Test* 149
  - *7.07.9.4.5 The Rey Complex Figure Test* 149
  - *7.07.9.4.6 The Recognition Memory Test* 149
  - *7.07.9.4.7 The Doors and People Test* 149

7.07.10 ASSESSING LANGUAGE DYSFUNCTION 150

7.07.10.1 Naming 150
7.07.10.2 Vocabulary 151
7.07.10.3 Discourse 151
7.07.10.4 Verbal Comprehension 152
7.07.11 ASSESSING VISUOPERCEPTUAL DYSFUNCTION 152
7.07.11.1 Visual Recognition 152
7.07.11.2 Visual Organization 153
7.07.11.3 Visuoconstructional Ability 153
7.07.11.4 Visual Interference 153
7.07.12 ASSESSMENT OF ATTENTION 153
7.07.12.1 Unilateral Neglect 153
7.07.12.2 The Anterior Attention System 154
7.07.13 ASSESSMENT OF EXECUTIVE FUNCTIONS 156
7.07.13.1 Frontal Hypotheses 156
7.07.13.2 Verbal Fluency 156
7.07.13.3 Nonverbal Fluency 157
7.07.13.4 Behavioral Assessment of the Dysexecutive Syndrome 158
7.07.14 THE DIFFERENTIAL DIAGNOSIS OF DEMENTIA VS. DEPRESSION 159
7.07.14.1 Delayed Verbal Recall Impairment as an Indicator of Early DAT? 159
7.07.14.2 Kendrick Cognitive Tests for the Elderly 160
7.07.14.3 Effortful vs. Automatic Processing and Memory 161
7.07.14.4 Is Poor Memory Test Performance in Elderly Depressed Patients an Artifact of Poor Motivation? 161
7.07.14.5 Future Directions 161
7.07.14.6 Recommendations and Conclusions for the Differential Diagnosis of Dementia vs. Depression 162
7.07.15 REFERENCES 162

7.07.1 THE AIMS OF NEUROPSYCHOLOGICAL ASSESSMENT

In the early days of clinical neuropsychology, most neuropsychological assessment was aimed at answering questions concerning differential diagnosis and lesion location. With the rapid advances in brain imaging technology the importance of this aim has diminished; the focus of assessment has moved to identifying the cognitive and behavioral consequences of cerebral dysfunction whether this be for rehabilitative or medico-legal purposes. However, the neuropsychologist continues to play a valuable role in these former areas, most notably in the diagnosis of dementing illnesses. In addition, although most neuropsychologists work in acute neurological/neurosurgical services or rehabilitation settings, most also take direct referrals from general medical and psychiatric services. As a result of this they are often the first to establish evidence for the presence of a neurological disorder.

Medico-legal assessments constitute a significant part of many neuropsychologists' workloads. Areas of particular relevance to the elderly in which a neuropsychological opinion is sought include issues of guardianship (e.g., is a client competent enough to manage their own financial affairs?) and fitness to hold a driving license. Other areas include personal injury litigation, fitness to plead in criminal cases (i.e., could a client with known or suspected neuropsychological deficits follow a legal argument against her/him and have an appreciation of its import), and evaluation of pleas of diminished responsibility (i.e., did pre-existing neuropsychological deficits impair a client's responsibility for her/his actions?).

In broad terms, neuropsychological assessment is aimed at answering the following questions:

(i) Which components of the cognitive system are dysfunctional, how severe is this dysfunction, and also, just as importantly, which components have been spared?

(ii) To what extent has there been change in mood, comportment, and personality; to what extent are these likely to be a direct effect of neurological damage as opposed to a psychological reaction to any injury or illness?

(iii) What are the implications of any changes in cognition, mood, and comportment for a client's every day functioning now and in the foreseeable future?

(iv) Based on the pattern of cognitive strengths and weaknesses and changes in mood and comportment, what practical advice can be given regarding the design of any formal rehabilitation and what advice can be given to the client and significant others to help them adjust to any deficits?

These questions are answered by integrating information gained from the medical notes, an interview with the client (and whenever possible a relative or close acquaintance), and the qualitative and quantitative findings from formal neuropsychological testing.

7.07.2 SOME SPECIFIC CONSIDERATIONS IN NEUROPSYCHOLOGICAL ASSESSMENT OF THE ELDERLY

There is a strong relationship between sensory deficits and cognitive function in elderly populations. For example, Lindenberger and Baltes (1994) conducted structural equation modeling of the relationship between age, cognitive test performance, and sensory functioning in an elderly sample (age range 70–103) and found that visual and auditory acuity accounted for 49% of the total and 93% of the age-related variance in cognitive performance. They evaluated a number of causal hypotheses to explain this relationship and concluded that it arose because both cognitive performance and acuity were indicators of the underlying physiological integrity of the aging brain. However, an additional possibility not considered by Lindenberger and Baltes is that sensory deficits impose an additional central processing load, resulting in poorer performance on cognitive tasks. Consistent with this possibility, Rabbit (1990) has reported that elderly subjects with 35–50 db hearing loss could repeat aloud flawlessly words read out to them but showed poor recall relative to age and IQ-matched controls with good hearing. In contrast, they had no such problem when the words were presented visually. This effect can be modeled in young subjects with good hearing when they have to process words heard through white noise. Rabbitt (1968) found that, under these conditions, the words were repeated successfully but recall was poor. It would appear, then, that the extra effort expended on making out the material places demands on processing resources which would otherwise be free for rehearsal or elaborative encoding, etc. This has far-reaching implications for the interpretation of neuropsychological test results in the elderly as evidence that clients are able to successfully to make out the stimulus materials does not rule out the possibility that their sensory deficits, rather than a higher level problem, has impaired their performance.

The presence of sensory loss should also be borne in mind when attempting to evaluate if there is dispositional change resulting from a central organic process. For example, hearing loss can lead to withdrawal from previous social activities because of frustration and the impatience of others.

Test fatigue can be an issue with the elderly and requires avoiding long test sessions. The neuropsychologist should also be sensitive to the possibility of effects arising from the order of test administration when examining test profiles. Clinicians experienced in working with individuals with dementia, particularly those in the early stages, will also be aware of their ability to "lift their game" for short periods sometimes to the frustration of their relatives who have sought an appointment because of concerns over cognitive deterioration.

As the result of more elderly people taking an increased interest in their health and the risks associated with aging, the neuropsychologist can play an important role in establishing objective evidence of adequate performance, thereby providing reassurance to clients.

Prior to 1990, there was a lack of adequate normative data on neuropsychological tests for the elderly. However, as governments have digested the implications of the change in the age structure of their populations, resources have been allocated to large-scale studies such as the Australian Longitudinal Study of Ageing (e.g., Bryan, Luszcz, & Crawford, 1997), the Cambridge Ageing Project (see Huppert, 1994), the Mayo Older Americans Normative Study (e.g., Ivnik et al., 1992), and the Canadian Study of Health and Aging (e.g., Tuokko & Woodward, 1996) among others. As a result, it will soon be possible to have better norms for the elderly than for other age groups.

The conventional approach to norming of tests is to obtain a sample which has been screened carefully to exclude potential participants with any medical conditions which may have a detrimental effect on cognitive performance. The case for an alternative approach has been well made by Huppert and colleagues. For example, Huppert (1994) has argued that "Rather than excluding individuals with risk factors which might affect performance, normalisation samples should be based on unselected population samples and the effects of these risk factors should be examined" (pp. 315–316). These two approaches are best seen as complimentary rather than competitors. Comparison of an individual's performance with conventional norms, or norms based on regression scores derived from demographic variables (i.e., age, education, current or former occupation, etc.), can address the issue of whether there is evidence of acquired impairment in a particular functional domain. The latter approach, presumably using multiple regression equations, can then potentially inform the clinician as to whether such impairment is likely to be unusual for someone with this particular combination of factors in their medical history. It should be noted, however, that this latter approach would be complex and would require very large samples.

7.07.3 THE NEUROPSYCHOLOGICAL INTERVIEW

Conducting a neuropsychological interview demands considerable empathy, tact, and intellectual effort. In many cases, the client will have multiple cognitive and physical disabilities and organically induced changes in mood and personality; they will also be attempting to cope with the often massive interpersonal and economic upheaval arising from their illness or injury. Added to this is the fact that a neuropsychological assessment can be anxiety provoking because of potential threats to self-esteem or because of its medico-legal ramifications; accordingly some clients can exhibit more concern at the prospect of such an assessment than over the lengthy and sometimes painful and bewildering physical investigations they may have to undergo.

As in all clinical interviewing, time must be spent establishing rapport and clarifying the nature and purpose of the investigation. General, open-ended, and nonthreatening questions should be employed, followed by a request for the client to describe any problems they have been experiencing in their everyday life. The client's description of these everyday problems constitutes a principal source for the generation and refinement of clinical hypotheses which will be tested in the course of the examination. Such descriptions and follow-up questions are also crucial in establishing the degree of insight the client has into the nature and severity of their deficits. Lack of insight is a major barrier to effective intervention and a client's successful adjustment. A client may exhibit a cheery disregard in the face of severe deficits coupled with grossly unrealistic expectations concerning a return to a former occupation or lifestyle; in other cases there may be a disproportionate concern with a relatively minor cognitive or physical problem when other deficits have much more serious implications.

Often, the most useful clinical information from the interview is gained by asking clients about their short-, medium-, and long-term goals and how they intend to achieve them. Answers to such questions are relevant to assessing drive, mood, planning ability, and level of insight. Inquiries about how they spend a typical day can also be employed to address these issues which have implications for the approach that should be adopted in any rehabilitation/remediation attempts. As Brooks (1989) states in this latter context, "the examiner is attempting to estimate whether the patient is a passive receptor of disability or an active fighter struggling to achieve a better outcome" (p. 66).

The neuropsychologist has a large agenda to cover in the interview and often will be under time pressure; this necessitates a structured, systematic inquiry. However, most formal neuropsychological tests also provide a structure for the client; the examiner sets out the concrete aims and rules, and largely controls the pace and order of the items administered (see the section on assessment of executive deficits for further discussion and exceptions to this). It is, therefore, important that enough of the interview is minimally structured to determine the extent to which the client initiates and organizes the discussion of topics.

A detailed educational history is important in building a picture of a client's premorbid abilities. The history should include years of schooling (with a check on whether any of these were repeat years), the nature of any tertiary education and any further study (evening classes, day release, etc.), and formal qualifications and grades achieved. The possibility that health, economic, social, or attitudinal factors may have prevented a client achieving their full educational and occupational potential should be carefully explored.

The client should be asked to describe the nature of their previous and/or current occupation with a detailed breakdown of the duties and functions performed. The nature of their work and responsibilities provides additional clues to general premorbid level of functioning but also an indication of specific premorbid cognitive skills that are likely to have been strongly developed/highly practiced. The clinician should inquire if there have been any recent changes in living circumstances, and if relevant, the work environment (i.e., introduction of new technology, reorganization of duties) as such changes can often be the catalyst which exposes problems in an individual who had previously been coping despite diminished cognitive resources.

Lezak (1995) sets out a number of fundamental questions which must be addressed prior to formal testing; these include: does the client understand the reason for referral and do they have specific questions of their own they want answered? Do they understand the uses to which the information gained from the examination may be put and who will and will not have access to it? (of particular importance in medico-legal contexts); do they know how and when feedback will be provided? Finally, do they appreciate that the assessment will largely be concerned with cognitive functioning (misperceptions abound) and are they aware of the general nature of the tests to be employed? In individuals with severe deficits the answers to these questions cannot be fully in the affirmative and it is therefore important that their relatives or guardians are briefed.

7.07.4 BASIC PSYCHOMETRIC ISSUES

7.07.4.1 The Rationale of Deficit Measurement in Clinical Practice

Attempting to detect and quantify cognitive deficits in the individual case is problematic because of the wide variability in cognitive abilities within the general population. Scores on neuropsychological measures which are average or even above average can still represent a significant impairment for an individual of high premorbid ability (and have serious implications for return to a previous lifestyle or occupation). Similarly, test scores which fall well below the mean do not necessarily reflect an acquired impairment for an individual whose premorbid resources were modest. Because of this, normative comparison standards are of limited utility in neuropsychological assessment and must be supplemented by individual comparison standards when assessing acquired deficits (Crawford, 1992; Lezak, 1983; Walsh, 1991). Ideally, this individual comparison standard can be obtained from psychological test scores obtained in the premorbid period. However, this is rarely a viable option; the amount of routine psychological testing conducted varies greatly between countries so that many individuals may have had no prior formal testing. Even where such test results exist they are often difficult or impossible to obtain, the content of the tests may have limited relevance, or they may have been administered so long ago that they are of questionable value. Because of these difficulties clinicians normally have to settle for an individual comparison standard which is based on a client's current performance. The explicit rationale here is that (i) cognitive ability measures are almost invariably positively correlated, therefore performance on one measure allows some level of prediction of performance on another, and (ii) some abilities will be preserved, or relatively so, following most neurological injuries or illnesses. Thus, the areas in which a client has performed best are used as the standards (i.e., estimates of premorbid ability) against which to compare performance on other measures. Large discrepancies between measures are taken as indicators of the presence and severity of acquired impairment (Lezak, 1991).

7.07.4.2 Converting Scores to a Common Measurement Scale

In constructing a neuropsychological profile of a client's strengths and weaknesses, most clinicians use instruments drawn from diverse sources. These instruments will differ from each other in the measurement scale used to express test scores; for some instruments no formal scaling will have been developed so that clinicians will be working from the means and standard deviations (SDs) of the raw score from normative samples. The process of assimilating the information from these tests is eased greatly if the scores are all converted to a common scale of measurement.

Converting all scores to percentiles has the important advantage that percentiles are comprehended easily by other health workers. However, because such a conversion involves an area transformation (i.e., the difference between a percentile score of 10 and 20 does not reflect the same underlying raw score difference as that between 40 and 50), they are not ideally suited to the rapid and accurate assimilation of information from a client's profile; percentiles are also inappropriate for use with many inferential statistical methods. Z scores do not suffer from these problems but have the disadvantage of including negative values and decimal places which can cause problems in communication.

One option is to convert all scores to have a mean of 100 and SD of 15 (McKinlay, 1992) as tests commonly forming a part of the neuropsychologist's armamentarium are already expressed on this scale, that is, IQs and factor score indices on the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981), memory indices from the Wechsler Memory Scale-Revised (WMS-R; Wechsler, 1987), and estimates of general premorbid ability such as the National Adult Reading Test (NART; Nelson & Willison, 1991). The most common alternative is to use *T* scores (mean 50; SD = 10); the meaning of such scores are easy to communicate and are free of the conceptual baggage associated with deviation IQs.

Regardless of which method is used, the clinician must be constantly aware that the validity of any inferences regarding relative strengths and weaknesses is heavily dependent on the degree of equivalence of the normative samples for the test results compared. Although the quality of normative data for neuropsychological tests has improved markedly since the 1980s, there are still tests used in practice which are standardized on small samples of convenience, the representativeness of which are largely unknown. Thus, discrepancies in an individual's profile may in some cases be more a reflection of differences between normative samples than differences in the individual's relative level of functioning in the domains covered by the tests.

7.07.4.3 Reliability

Adequate reliability is a fundamental requirement for any instrument used in neuropsychology regardless of purpose (Franzen, 1989). However, when the concern is with assessing the cognitive status of an individual its importance is magnified, particularly as the demands of clinical practice are such that decisions, must be commonly made based on information from single administrations of each instrument.

Information on test reliability is used to quantify the degree of confidence that can be placed in test scores, for example, when comparing an individual's scores with appropriate normative data or assessing whether discrepancies between scores on different tests represent genuine differences in the functioning of the underlying components of the cognitive system as opposed to simply reflecting measurement error in the tests employed. In the latter case, that is, where evidence for a differential deficit is being evaluated, it is important to consider the extent to which the tests are matched for reliability; an apparent deficit in function A with relative sparing of function B may simply reflect the fact that the measure of function B is less reliable. This point was well made in a classic paper by Chapman and Chapman (1973) in which the performance of a schizophrenic sample on two parallel reasoning tests was examined. By manipulating the number of test items the schizophrenic sample was made to appear to have a large differential deficit on one or other of the tests.

Particular care should be taken in comparing test scores when one of the measures is not a simple score but a difference or ratio score. Such scores are used quite commonly in neuropsychological assessment; for example, in the assessment of implicit memory functioning a priming score may be derived from the difference between completion of word fragments (e.g., c–pe– / carpet) from a list that has been primed in an earlier study phase and an unprimed list (performance on the unprimed list essentially controls for individual differences in verbal ability). This priming score will have poor reliability because the measurement error associated with the two lists is additive. For example, if both lists had reliabilities of 0.75 and an intercorrelation of 0.6, the formula for the reliability of a difference score (e.g., see Crocker & Algina, 1986) reveals that the reliability of this priming score would be only 0.37. This compares unfavorably with the reliability of most explicit memory tests and raises the danger that the clinician may conclude that an individual has impaired explicit memory coupled with preserved implicit memory when the pattern may simply be an artifact of differences in reliability.

7.07.4.4 Reliable vs. Abnormal Differences

The distinction between the reliability and the abnormality of differences between test scores is an important one in clinical neuropsychology. A difference between scores normally would be considered reliable if it exceeded the 95% confidence interval for the difference, that is, a difference of this magnitude is unlikely to have arisen from measurement error in the instruments. Establishing if a difference is reliable is only the first step in neuropsychological profile analysis. There is considerable intraindividual variability in cognitive abilities in the general elderly population such that reliable (i.e., statistically significant) differences between tests of these different abilities are common. In evaluating the probability that a discrepancy reflects acquired impairment, it is important to consider the abnormality or rarity of the difference, that is, what percentage of the general (i.e., unimpaired) elderly population would be expected to exhibit a difference of this magnitude?

Base rate data, which permits the clinician to assess the abnormality of test score discrepancies, is available for some neuropsychological measures. When such data are not available, a simple formula can be used to estimate the abnormality of any discrepancy from the correlation between the two tests of interest (see Payne & Jones, 1957). It is also possible to assess the abnormality of the discrepancy between a single test and a client's mean scores on a series of measures (Silverstein, 1984).

To highlight the distinction between the reliability and abnormality of a difference take the example of a discrepancy between the General Memory and Attention/Concentration indices of the WMS-R. A discrepancy of 17 points would be necessary for a reliable difference ($p < 0.05$). Based on the average correlation between these two indices in the WMS-R standardization sample ($r = 0.51$), approximately 25% of the general population would be expected to exhibit a discrepancy of this magnitude; to be abnormal (operationally defined for the present purpose as discrepancy which would occur in less than 5% of the general population), a 29-point discrepancy would be required.

The importance of evaluating the abnormality of discrepancies through base rate data or correlational techniques cannot be overstressed. Most clinical neuropsychologists have not had the opportunity to administer neuropsychological measures to significant numbers of individuals drawn from the general elderly population. It is possible, therefore, to form a distorted impression of the degree of intrasubject variability found in unimpaired individuals. The impression is that the degree of normal variability may be underestimated leading to a danger of overinterference when working with clinical populations. For example, Matarazzo, Daniel, Prifitera, and Herman (1988) have reported that in the WAIS-R standardization sample, a sample presumed to be free of neurological or psychiatric disorder, the mean difference between an individual's highest and lowest WAIS-R subtest score was 6.7. For this difference (i.e., the subtest range to be abnormal it would have to exceed 11 Scaled score points. As subtests have an SD of 3, it can be seen that a subtest range of around 3 SDs is not unusual in individuals without acquired impairments. These considerations are just as relevant when profile analysis is carried out with more specific neuropsychological instruments. The WMS-R and WAIS-R are used here as examples simply because of the availability and quality of the relevant data.

7.07.4.5 Monitoring Change

There are many situations in which the neuropsychologist needs to measure potential changes in cognitive functioning in the elderly. Common examples would be to determine whether cognitive decline is occurring in an individual in whom a degenerative neurological process is suspected, or to determine the extent of recovery of function following a stroke or traumatic brain injury. In both these cases, neuropsychological assessment will provide useful information to assist clients, relatives, and other health professionals to plan for the future. Monitoring the cognitive effects of surgical, pharmacological, or cognitive interventions in the individual case is also an important role for the neuropsychologist. Although the aim here is most commonly to determine if there has been any improvement, the possibility of detrimental effects can also be an issue. For example, many drugs can potentially impair cognitive functioning, particularly in the elderly. Anticholinergic agents provide a good example as they are widely used in the treatment of various conditions and yet can have serious negative effects on cognition, particularly memory (Crawford, Besson, & Ebmeier, 1990).

In assessing the effectiveness of a rehabilitation effort it is often possible to obtain multiple repeated measures of an individual's performance before, during, and after intervention. A number of inferential statistical techniques can be used in this situation because there are multiple data points for the different phases (see Barlow & Hersen, 1984 for a general treatment of single-case designs). However, in general clinical practice the neuropsychologist often must come to conclusions about change from only a single retesting. This situation will also arise in rehabilitation settings; although multiple measures may have been obtained on the training task(s), the issue of the generalizability of any improvement is often addressed by comparing single before-and-after scores on related, but separate tests.

Monitoring change on the basis of a single retesting is a formidable task except in cases where the level of change has been dramatic. It is rendered more formidable by the fact that many of the standard instruments used in clinical neuropsychology (e.g., the WMS-R) do not have parallel versions. The clinician must therefore differentiate changes resulting from systematic practice effects and random measurement error from change reflecting genuine improvement or deterioration. Among other complications are the fact that (i) the magnitude of practice effects vary with the nature of the task (e.g., the Performance subtests of the WAIS-R show larger practice effects than their Verbal counterparts), (ii) the length of time that has elapsed between test and retest will influence the magnitude of effects, and (iii) a diminution of practice effects is to be expected in neurological populations given the high prevalence of memory and learning deficits, but the expected diminution is difficult to estimate for individuals.

One approach to dealing with many of these considerable interpretive problems is to gather information from test–retest studies on the relevant neuropsychological tests. Provided that such studies report the test–retest correlations, means, and SDs, regression equations can be constructed to predict performance on retest from scores at initial testing. The predicted scores are then compared with the retest scores actually obtained by the client to assess whether the observed gain or decline significantly exceeds that expected (see Knight & Shelton, 1983; McSweeny, Naugle, Chelune, & Lüders, 1993). The utility of this approach is determined by the extent and nature of retest studies available for a particular test. For example, the test–retest scores on memory tasks for an elderly client with suspected dementia could be compared with estimated retest scores derived from a healthy, elderly sample retested after a similar period to gauge how atypical any decline may be. For some questions test–retest data from a clinical sample may be used. For example, a stroke patient's scores on measures of attention or speed of processing could be compared with estimated retest scores from a stroke sample if the clinician suspects that the extent of recovery is atypical.

7.07.5 APPROACHES TO ASSESSMENT IN CLINICAL NEUROPSYCHOLOGY

Approaches to the assessment process can be characterized as falling on a continuum between what has been referred to as the "fixed, big battery" approach and a flexible, hypothesis-testing approach (Brooks, 1989; Walsh, 1991). In the former, a large comprehensive battery is routinely administered to all clients. This approach is most common in the USA and is exemplified by the Halstead–Reitan battery which consists of a large number of specific neuropsychological measures, for example, tests of tactual performance, sensory extinction, finger tapping, categorization, language functioning, etc., which are often supplemented with a full-length WAIS or WAIS-R and personality inventories (see Reitan, 1986).

In the latter approach, measures are selected to test clinical hypotheses derived from the neuropsychological literature on a client's known or suspected disorder and from information gained from the interview. The process is a dynamic one; the results of preliminary testing are used to test or modify existing hypotheses and generate new ones. As a simple example, screening measures may detect problems with the organization of visual material which are consistent with spatial and/or planning deficits; follow-up measures can then be administered to evaluate these three competing possibilities.

Proponents of the flexible approach acknowledge that there is a potential danger of failing to detect a client's difficulties in some cognitive domains if the clinician focuses too rapidly on testing inappropriate hypotheses (Miller, 1992). However, they argue that the fixed, big battery approach may represent an inefficient use of resources and is simply not an option in some hard-pressed services (Brooks, 1989; Walsh, 1991). Further, because of the time taken to complete an assessment, many of the measures may be administered by psychology assistants rather than fully accredited clinicians. It has been argued that important clinical information may therefore be missed or misinterpreted (Brooks, 1989). In the fixed, battery approach all clients and normative groups are administered the same tests, therefore services employing this approach accumulate substantial data on their measures. Such databases are tremendous clinical assets. However, the very existence of this database, and the resources expended to obtain it, may promote a reluctance to substitute existing tests with newer measures designed to reflect developments in neuropsychological knowledge.

Approaches to assessment can also be characterized by the emphasis placed on qualitative vs. quantitative evidence. The extreme quantitative pole of this dimension is characterized by a strict statistical/actuarial methodology in which comparison of a client's test scores against various cut-offs is the principal focus. At the other extreme, tests are used to reveal qualitative features of a client's behavior, the emphasis being on how a task is approached rather than on any empirical analysis of the test scores obtained. In the service of this end there may be major deviations from standardized test instructions, thus precluding statistical analysis in any case. Although there is no fundamental reason why this dimension should not be orthogonal to the fixed/flexible dimension, in practice they have tended to be correlated. The fixed, battery approach arose from researchers working in the actuarial tradition and the sheer number of measures administered almost demands the employment of actuarial indices to assist in the assimilation of the information obtained.

A strict quantitative methodology could be employed with the flexible hypothesis-testing approach. Indeed, the term hypothesis-testing might be seen as implying this and indeed influential figures in the development of this approach devoted considerable attention to quantification (e.g., Shapiro, 1973). However, many clinical proponents of this approach have tended to emphasize qualitative analysis that is, hypotheses may be expressed and tested in qualitative as well as quantitative terms (Walsh, 1991). It should be stressed that very few clinical neuropsychologists are at the extreme ends of this qualitative/quantitative dimension; qualitative observations are verified with quantitative instruments whenever possible whilst recognizing that much important information is not easily amenable to such verification.

7.07.6 SCREENING TESTS FOR MENTAL STATUS

This section will briefly review screening tests for mental status. All these Scales include questions covering basic orientation for time and place but differ in the extent to which they address other competencies. Among the most common Scales are the Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975), the Mattis Dementia Rating Scale (MDRS; Mattis, 1976), and the cognitive section (CAMCOG) of the Cambridge Mental Disorders of the Elderly Examination (CAMDEX; Roth et al., 1986). These instruments produce a global score which summarizes a patient's cognitive competence. They are brief and easy to administer and provide sufficient information for epidemiological studies, or as an index of severity of decline. They cannot be viewed as exhaustive diagnostic instruments.

The MMSE (Folstein et al., 1975) is probably the most frequently used screening Scale and is often used as part of a larger battery for a comprehensive assessment of dementia. It is a very brief and easily administered instrument and takes about 5–10 minutes in total. It tests orientation, information, and visuoconstructive abilities; total scores can range between 0 and 30. In their original study, Folstein et al. reported a high test–retest reliability of 0.83 when retesting was conducted by a different examiner, and 0.89 when the same examiner was used. High test–retest reliability, ranging from 0.8 to 0.95, has been confirmed by later studies (Anthony, Le Resche, Niaz, Van Korff, & Folstein, 1982; Dick et al., 1984). An effect of age on MMSE scores has been reported by several studies (Magni, Binetti, Bianchetti, Rozzini, & Trabucchi, 1996; Pi, Olive, & Esteban, 1994; Tombaugh & McIntyre, 1992). Furthermore, caution is required when interpreting scores obtained from poorly educated individuals. Level of education has been shown consistently to have an effect on MMSE score (Kittner et al., 1986; Kukull et al., 1994; Uhlmann & Larson, 1991).

A useful instrument for the evaluation of the mental status is the Mattis Dementia Rating Scale (MDRS) (Mattis, 1976). This Scale can take about 30–45 minutes to administer, but since items are ordered in descending order of difficulty passes on early items can shorten the time required because of discontinuation rules. The MDRS tests orientation, attention, initiation and perseveration, visuoconstructive abilities, conceptualization, and memory. Split-half reliability for this Scale is high (0.90) (Gardner, Oliver-Munoz, Fisher, & Empting, 1981). Norms are available (Montgomery, 1982); however, according to Spreen and Strauss (1991) the normative sample cannot be considered representative of the elderly population as it contained a disproportionate number of high socioeconomic status (SES), highly educated participants.

As noted, the CAMDEX (Roth et al., 1986) includes a cognitive section (CAMCOG). This Scale contains the items of the MMSE with the addition of some additional coverage of perception, memory, and abstract thinking. A study (Huppert et al., submitted) has examined the psychometric properties of the CAMCOG; the test–retest reliability of the Scale as a whole was very high (0.86), as was its internal consistency (coefficients ranged from 0.82 to 0.89). A study carried out in a group of 222 elderly people reported that CAMCOG scores are affected by age, sociocultural factors, and hearing and visual deficits (Blessed, Black, Butler, & Kay, 1991). The cognitive subScale of the Alzheimer's Disease Assessment Scale (Rosen, Mohs, & Davis, 1984) is used increasingly widely in clinical trials in DAT because of its proven sensitivity to the effects of cholinesterase inhibitors and other drugs (Schneider, 1996). It takes around an hour to administer and includes tests of word recall, orientation, naming, spoken language and comprehension, and constructional praxis.

7.07.7 MEASURES OF CURRENT INTELLECTUAL ABILITY

7.07.7.1 Use of the WAIS-R With the Elderly

The WAIS-R (Wechsler, 1981) continued to be used widely in neuropsychological assessment of all individuals of all ages. Lezak (1988a), for example, describes the Wechsler as, "the workhorse of neuropsychological assessment" and identifies it as "the single most utilized component of the neuropsychological repertory" (p. 38). A number of points should be made about the use of the WAIS-R with the elderly. First, as will be discussed later, many clinicians base their analysis of WAIS-R performance on a client's subtest profile. It is important that age-graded Scaled scores are used in this process rather than the standard Scaled scores which are not corrected for age (Lezak, 1995). Failure to use these former scores can produce a totally misleading picture of individuals' strengths and weaknesses relative to their peers; (see Crawford, 1992 p. 27). Second, as aging is associated with reduced visual acuity, it is regrettable that the stimulus materials for some of the subtests (most notably Picture Completion and Picture Arrangement) are unnecessarily small and the artwork poor.

The most obvious problem in using the WAIS-R with the elderly is that the original norms only cover those up to age 74 years. For clients over this age two sets of independently collected norms are available (Ivnik et al., 1992; Ryan, Paolo, & Brungardt, 1990). The Mayo norms were established on a sample of 512 healthy individuals from Rochester and surrounding areas in the US and covers up to age 97 years. However, many of the total subject pool were under age 74 years and were used to examine the comparability of the MAYO norms with the existing WAIS-R norms. Thus, the norms for those over age 74 years were based on a subsample of 222. The sample was relatively highly educated and had an underrepresentation of ethnic minorities. Two features of these norms are potentially appealing. First the norms were established using the method of overlapping midpoint age ranges described by Pauker (1988); this takes maximum advantage of the data available and, because it permits more age bands to be formed, lessens the sometimes substantial apparent improvement in scores that can occur when an individual moves from one age band to the next. As an aside, it is worth noting that, using the existing WAIS-R norms and any spreadsheet package, it is possible to go one step further and derive continuous norms for those aged under 74 years (see Zachary & Gorsuch, 1985). Second, age corrections are carried out at the subtest level rather than when converting the sum of Scaled scores to IQs. This is the same approach used for the Wechsler children's Scales; why it was not adopted for the adult Scales remains a mystery.

Ryan et al's. (1990) norms were established for two age groups, 75–79 ($n=60$) and 80 and above ($n=70$). Although this sample was somewhat smaller than the Mayo sample, it was recruited to be broadly representative of the US elderly population in terms of those demographic variables that were considered to be most pertinent to cognitive tests. Paolo and Ryan (1995) discuss the pros and cons of both samples in detail. They also provided an example subtest profile for a case aged 75 years (i.e., just outside the standard WAIS-R norms) and compared the results obtained by both sets of elderly norms and by scoring the case as if he were 74 years using the original norms. The original and Ryan norms showed much greater evidence of convergence; the Mayo norms indicated significant deficits on some Performance subtests which were not present in the other two profiles. An additional advantage of the Ryan norms is that base rate data on subtest scatter and other aspects of performance are available (e.g., Ryan & Paolo, 1992).

7.07.7.2 Analysis of Subtest Profiles

Analysis of WAIS-R performance can be based on the IQs, the subtests or scores derived from factor analysis. Authorities on the Wechsler have divided opinions on the usefulness of the summary IQs. Matarazzo and Herman (1985), for example, have suggested that the discrepancy between Verbal and Performance IQs is the best validated Wechsler index of cerebral dysfunction, whereas Lezak (1995) is dismissive of its importance. Many clinicians primarily base their interpretation of the WAIS-R on the pattern of strengths and weaknesses in a client's subtest profile (e.g., Lezak 1995; Walsh, 1991). Lezak (1988b), for example, notes that IQs can obscure clinically important strengths and weaknesses and suggests that "although the traditional scoring scheme for the Wechsler exemplifies the IQ problem, these tests also show us a way out of the problem by providing a profile of (subtest) scores" (p. 359).

A starting position in conducting such a profile analysis is to assess whether any observed subtest differences are reliable (i.e., statistically significant) as opposed to simply reflecting measurement error. The only assistance for this process provided in the WAIS-R manual appears in Table 13 (Wechsler, 1981). This table provides the critical values required for significance when comparing any subtest with any other. It is appropriate to use these critical values if and only if the clinician has made an *a priori* decision to examine the difference between only one pair of subtests out of the 55 potential comparisons that could be made. However, in practice, clinicians will want to compare more than one pair; indeed, when the client's history and behavior at interview provide insufficient grounds for forming clinical hypotheses, such comparisons will be *post hoc*. In effect, all possible comparisons are made in this situation even though the clinician will focus attention on the subtest comparisons which yielded the largest discrepancies.

The principal problem with Wechsler's approach is that it does not introduce any correction for the inflation of the Type I error rate that occurs when multiple comparisons are conducted (Crawford, 1992; Silverstein, 1982). Any attempt to overcome this problem inevitably involves striking a balance between controlling the Type I error rate while retaining reasonable power to reject the null hypothesis of no subtest differences. For example, an unsatisfactory solution would be to produce a modified table in which a Bonferroni correction was applied to derive new critical values reflecting the fact that 55 comparisons were involved. Although such an approach would control for inflation of the Type I error rate, it would result in very low power to detect subtest differences, that is, the Type II error rate (i.e., the rate of false negatives) would be unacceptably high.

Knight and Godfrey (1984) and Silverstein (1982) independently suggested a method which achieves a useful compromise between the need to minimize both Type I and Type II errors; instead of potentially comparing all subtests with each other, each subtest is compared with a client's mean subtest score. In the case of a full-length WAIS-R this reduces the number of comparisons to 11 so that, when the Bonferroni correction is applied to maintain the Type I error rate at the specified level, the loss of power to detect subtest differences is much less acute; 11 appears in the denominator rather than 55. Both authors provided a table which recorded the size of discrepancy between each subtest and an individual's mean subtest scores required for varying levels of statistical significance. This approach, when combined with equivalent tables on the abnormality of any subtest differences (see below), provides a rapid and straightforward means of analyzing a client's profile of strengths and weaknesses and therefore provides information which can be integrated usefully with other quantitative and qualitative information available to the clinician.

Crawford (1997) presented a modified table which corrects a minor error common to both of the original tables. Further tables are provided for use with any of nine short-forms of the WAIS-R, including those specifically developed for neuropsychological purposes (e.g., Reynolds, Willson, & Clark, 1983; Walsh, 1991) and those in which the selection of subtests was guided by their suitability for use with the elderly and other groups in which reduced visual acuity can be a problem (e.g., Britton & Savage, 1966; Crawford, Allan, & Jack, 1992a).

If subtest differences are found to be reliable, it is often appropriate to determine if they are also abnormal, that is, to establish how rarely differences of the magnitude observed in an individual's profile occur in the general population (see Section 7.07.4.4). Many healthy individuals have reliable strengths and weaknesses in their subtest, thus reliable differences do not imply the presence of acquired impairment. Silverstein (1984) provided a table to assess the abnormality of subtest differences. As was the case for the reliability of subtest differences, this table is used to compare each subtest with an individual's mean subtest score. Crawford, Allan, McGeorge, and Kelly (1997) adopted the same method to prepare a set of abnormality tables for use with any of the nine WAIS-R short-forms referred to earlier.

The above methods allow clinicians to test hypotheses based on clinical experience or the research literature on a client's known or suspected neurological disorder. However, the methods themselves make no *a priori* assumptions about the nature of an individual's subtest pattern. In contrast, there have also been attempts in the past to identify fixed, specific subtest patterns for differential diagnostic purposes. Most of these attempts naively aimed to identify "organicity" and hence were based on a unitary concept of brain damage which is totally discredited. Not surprisingly, empirical studies of these profiles, such as Wechsler's "Hold–Don't Hold" pattern, have been consistently disappointing (e.g., Christensen & MacKinnon, 1992; Savage, Britton, Bolton, & Hall, 1973).

A more rational and interesting attempt to identify specific subtest patterns was conducted by Fuld (1984). This approach arose from evidence implicating cholinergic dysfunction in the cognitive deficits seen in dementia of the Alzheimer type (DAT) (e.g., Perry et al., 1978; see Crawford et al., 1990 for a review). Fuld (1984) examined Drachman and Leavitt's (1974) data on the effects of scopolamine on the WAIS performance of healthy adults and reported that the age-graded subtest profile could be typified by the formula $A > B$; $B > C$, $C \leqslant D$, $A > D$ where A is the mean of Information and Vocabulary, B of Similarities and Digit Span, C of Digit Symbol and Block Design, and D is Object Assembly. Fuld reported that around a half of patients with DAT exhibited this profile, whereas it was observed in less than 7% of controls and patients with multi-infarct dementia (MID). These results provoked much interest (see Massman & Bigler, 1993 for a review). Subsequent studies have reported rates of occurrence in DAT ranging from a high of 57% (Brinkman & Braun, 1984) to a low of 7% (Logsdon, Teri, Williams, Vitiello, & Prinz, 1989). Most subsequent data on the specificity of the profile indicated that it occurred with low frequency in the healthy elderly and in neurological disorders other than DAT. However, reasonably high rates of occurrence have been reported in depression (16%) and schizophrenia (15%). A particularly disappointing set of results was reported in the previously mentioned study by Logsdon et al. (1989) in which the 7% rate of occurrence in DAT was exactly equivalent to the rates in the healthy elderly and depressed patients. Gfeller & Rankin (1991) have also reported that the profile was exhibited by 12% of cases of multi-infarct dementia, a figure which is little different from the rate found in their DAT sample (15%).

With knowledge on the extent of intraindividual variability in subtest scores (e.g., Matarazzo et al., 1988), it should not be surprising that the profile has low sensitivity for DAT. For example, it would not be uncommon for individuals to have premorbid Block Design or Digit Symbol scores which were well over one SD higher than their average subtest score; similarly, it would not be uncommon for individuals to have Information or Vocabulary scores which were substantially below their average subtest score. Thus, although these sets of subtests may be differentially affected by the onset of DAT, a massive effect would be required to reverse the premorbid pattern; accordingly, these differential effects often would be obscured in the individual case. In conclusion, the Fuld profile has little to recommend its use in clinical practice. However, it is still of some theoretical interest as scopolamine-treated healthy subjects and DAT cases are apparently the only groups studied to date in which the profile is exhibited in their respective means. This provides some support for the cholinergic hypothesis in DAT. However, it should be noted that there has been no attempt to replicate the original report on the effects of scopolamine. It would also be important to demonstrate that the profile has pharmacological specificity, that is, that a similar profile would not be observed following administration of a benzodiazepine or some other centrally active agent.

7.07.7.3 WAIS-R Factor Scores

An alternative to the analysis of both IQs and subtest discrepancies is to examine scores obtained from factor analysis of the WAIS-R (Atkinson, 1991; Canavan, Dunn, & McMillan, 1986; Crawford, Allan, Stephen, Parker, & Besson, 1989a; Kaufman, 1990). Although this suggestion has a long history (Maxwell, 1960), factor scores have not been employed widely in clinical practice or research. Factor analyses of the WAIS and WAIS-R commonly have extracted three factors: a verbal (V) factor on which Information, Vocabulary, Comprehension, and Similarities have high loadings; a perceptual–organization (PO) factor defined by high loadings from Block Design and Object Assembly (other Performance subtests have more modest loadings); and a third factor termed attention/concentration (A/C) or freedom-from-distractibility consisting of Arithmetic, Digit Span, and often Digit Symbol (Crawford et al., 1989a; Leckliter, Matarazzo, & Silverstein; 1986). The arguments in favor of using factor scores when interpreting WAIS-R performance are many and compelling. First, the IQs do not have optimal construct validity, the subtests were allocated to a Verbal and Performance Scale on intuitive grounds. Second, the factor structure of the WAIS-R appears to be very robust in that it is consistently extracted from samples from countries outwith the USA. Most crucially, it emerges when the WAIS-R scores of clinical samples are factor analyzed (e.g., Atkinson, Cyr, Doxey, & Vigna, 1989). Third, the reliabilities of factor scores are much higher than those of the individual subtests from which they are formed. Fourth, although authorities on the Wechsler have laid stress on the analysis of subtest scatter, the WAIS-R could be viewed as primarily providing broad indicators of current functioning against which more specific neuropsychological measures are compared. Again, the superior construct validity of the factorially derived composites suggest they are better suited to provide these broad intraindividual comparison standards than the IQs.

Crawford, Johnson, Mychalkiw, and Moore (in press-a) have examined the clinical utility of IQs, indices of subtest scatter, and factor scores in discriminating between healthy ($n = 356$) and head-injured samples ($n = 233$). The scatter indices employed were the Profile Variability Index (McLean, Reynolds, & Kaufman, 1990), which simply records the variance of an individual's subtest scores, and The Mahalanobis Distance Index (Burgess, 1991); this latter index factors in the correlation between subtests (therefore, a discrepancy involving a subtest which has a relatively low average correlation with the other subtests is given less weight than a discrepancy involving a subtest with a high correlation). Factor scores were significantly better than both the IQ Scales and the scatter indices at discriminating between healthy and impaired performance; the scatter indices essentially performed at chance levels. The extent to which these findings will generalize to neurological conditions more directly associated with the elderly (e.g., DAT, stroke, etc.) remains to be determined.

The lack of a rapid method of factor scoring the WAIS-R, combined with the absence of data with which to assess the reliability and abnormality of discrepancies between factors, acted as barriers to the effective use of factor scores in clinical practice. Both these limitations have simultaneously been rectified in a very useful paper by Atkinson (1991).

7.07.8 SPECIFIC METHODS FOR THE ESTIMATION OF PREMORBID ABILITY

The National Adult Reading Test (NART) (Nelson & Willison, 1991) is the test most widely used to estimate premorbid ability. The NART is a single word, oral reading test consisting of 50 items. All the words are irregular, that is, they violate grapheme–phoneme correspondence rules (e.g., chord). Because the words are irregular, intelligent guess work should not provide the correct pronunciation, therefore the test taps previous word knowledge; as the test only requires the reading of single words, clients do not have to analyze a complex stimulus and it is argued that the test therefore makes minimal demands on current cognitive capacity (Nelson & O'Connell, 1978). Development of the NART arose from the clinical observation that oral reading is commonly preserved in dementia (whereas reading for meaning is commonly impaired). However, the test is used to estimate premorbid ability in a wide range of conditions.

To qualify for use as a measure of premorbid ability, a test must fulfill three criteria (Crawford; 1992; O'Carroll, 1995). First, as with any psychological test, it must possess adequate reliability. The NART has high split-half reliability/internal consistency, test–retest reliability, and inter-rater reliability (Crawford, 1992). Second, it must have high criterion validity. The NART is normally used to provide an estimate of general premorbid IQ against which current performance on the WAIS-R is compared. Thus, to meet the second requirement, the NART must be capable of predicting a substantial proportion of IQ variance. There has been some confusion in the literature on how to examine this aspect of putative measures of premorbid ability (e.g., Klesges, Wilkening, & Golden, 1981), so it is worth noting that it must necessarily be studied using unimpaired rather than clinical samples. In a clinical sample, NART performance and performance on the criterion (e.g., the WAIS-R) will commonly become dissociated, indeed it is the presence of such a dissociation (i.e., a large discrepancy between estimated premorbid ability and current ability in favor of the former) that is used to infer impairment. In most studies using the WAIS or WAIS-R as the criterion variable the NART predicted well over 50% of IQ variance. For example, in one study the NART predicted 66% of WAIS IQ variance in a sample of 151 healthy subjects (Crawford, 1992).

The final criterion for a putative measure of premorbid ability is that test performance be resistant to neurological or psychiatric disorder. NART performance appears to be largely resistant to the effects of many neurological and psychiatric disorder, for example, depression, acute schizophrenia alcoholic dementia, closed head injury, and Parkinson's disease (Crawford, 1992; O'Caroll, 1995) and compares favorably with previously suggested alternative measures of premorbid ability such as the Vocabulary subtest of the WAIS or WAIS-R (e.g., Crawford et al., 1988). Mixed findings have been found in samples with probable DAT. O'Carroll, Baike, and Whittick (1987) readministered the NART to a sample of dementia cases after one year and found no change in performance despite a significant decline on measures of dementia severity. Crawford et al. (1988) found that the NART performance of DAT cases did not differ significantly from matched controls despite the presence of severe deficits on other cognitive measures and the presence of marked morphological and blood flow abnormalities on brain imaging. Stebbins, Wilson, Gilley, Bernard, and Fox (1988), in contrast, found that NART performance was impaired significantly in DAT cases classified as moderate or severe, although impairment was not in evidence in mild cases. Subsequent studies have continued to produce conflicting results in DAT (e.g., Patterson, Graham, & Hodges, 1994; Sharpe & O'Carroll, 1991) but it is becoming increasingly clear that NART performance can be impaired substantially in many cases of moderate to severe dementia. Any findings of impaired NART performance poses a threat to the validity of this approach. However, the practical implications of impaired performance in cases of severe neurological disorder are not as serious as they may appear; in such cases the presence of deficits is unfortunately only too obvious, thereby largely obviating the need for the NART or a similar instrument to assist in its detection and quantification.

Most research on the NART's ability to estimate premorbid ability has used scores on IQ tests as the criterion variable. Although this is in keeping with the notion of obtaining an estimate of general level of premorbid functioning, the NART also has the potential to provide estimates of premorbid functioning for other WAIS-R indices and for more specific neuropsychological tests. For example, Crawford, Moore, and Cameron (1992c) built a regression equation which can be used to estimate premorbid performance on the FAS verbal fluency test (see Section 7.07.13). Similarly Crawford, Obonsawin, and Allan (in press-b) have provided an equation which uses NART and age to estimate premorbid scores on the Paced Auditory Serial Addition Test (PASAT; Gronwall, 1977).

Schlosser and Ivison (1989) derived a regression equation which incorporated age and NART scores to estimate performance on the Wechsler Memory Scale (WMS). They reported that the discrepancy between current WMS performance and estimated premorbid performance was a highly successful index for the detection and quantification of memory dysfunction in DAT. O'Carroll et al. (1994) extended this work by comparing a group of DAT patients, a group of depressed patients, and a group of healthy controls on the discrepancy between NART-estimated premorbid intellectual ability and current mnemonic ability as assessed using the various indices of the WMS-R. While large mean differences were demonstrated between groups, the degree of overlap between the DAT and depressed group was large, leading them to conclude that none of the NART/WMS-R discrepancy quotients were significantly discriminating between depression and DAT to be useful in clinical practice.

A proposed modification to the NART is the Cambridge Contextual Reading Test (CCRT; Beardsall & Huppert, 1994). As its name suggests, the NART words are embedded in sentences to provide context for the examinee. Baddeley, Emslie, and Nimmo-Smith (1993) proposed another alternative reading test which they termed the Spot The Word Test (STW). STW is a lexical decision task in which the examinee has to identify the legitimate words from a series word/pseudoword pairs (e.g., stamen/floxid). Both of these new tests have considerable potential as alternatives to the NART. However, their use is mainly limited to research applications as regression equations have yet to be developed to provide estimates of premorbid ability for individuals and there is insufficient evidence on their resistance to impairment.

An entirely different approach to the estimation of premorbid ability uses demographic variables (e.g., years of education, occupational classification) as predictors. This has the advantage that the estimates it provides are entirely independent of current level of functioning (care should, however, be taken to assess whether prodromal difficulties could have led to a client's failing to achieve their educational or occupational potential). As evidence of impaired NART performance in various clinical conditions accumulates, the demographic approach may be a useful alternative method in cases where use of the NART would be inappropriate. However, the obvious disadvantage of the demographic approach is its modest criterion validity. Where WAIS or WAIS-R scores have been used as the criterion variable, regression equations based solely on demographic variables account for only 30–54% of the variance (Barona, Reynolds, & Chastain, 1984; Crawford & Allan, 1997; Crawford et al., 1989b; Wilson et al., 1978). This compares unfavorably with the NART and its variants (Crawford, 1992; O'Carroll, 1995).

7.07.9 AGING AND MEMORY

7.07.9.1 Introduction to Aging and Memory

It is a well-established phenomenon that as humans age, memory complaints increase. These subjective complaints are reflected in objective measurements of memory test performance. For example, Davis and Bernstein (1992) examined verbal memory performance, as assessed by the Auditory Verbal Learning Test (AVLT) in a sample of 474 healthy individuals ranging from those in their twenties to those in their eighties; a steady decrease in performance was observed with increasing age in those over 30 years. Free recall performance often appears to be more detrimentally affected by the aging process than recognition, although other memory processes do appear to be sensitive to age effects, for example, priming skill learning and classical conditioning (see Binks, 1989 and Schonfield, 1980 for reviews).

Shimamura (1994) proposed that age-related effects on memory may be associated with regional brain changes, that is, that problems of attention, working memory, word finding, and source memory may be related to aging effects in frontal lobe functioning, whereas problems with new learning may be related to medial temporal lobe dysfunction. Superimposed on these changes, Shimamura (1994) argued that general changes in neuronal efficiency may pervade all neural systems as a consequence of age (e.g., reduced blood flow, inability to prevent damage from oxidation, and poor metabolic uptake). It is proposed that these metabolic changes account for the pervasive finding of cognitive slowing in older individuals. It can be imagined easily how these specific age-related memory effects may act synergistically with nonspecific slowing to result in impaired performance. Evidence has shown that differing brain regions undergo differing degrees of neuronal loss with age. Many areas, such as striate-cortex and parietal cortex, do not lose an appreciable number of neurons during life. However, 15–20% of the neurons in the neo-striatum and frontal cortex are lost between young adulthood and old age (Squire, 1987). Similarly, approximately 5% of hippocampal neurons are lost per decade, so that at the age of 80 years; approximately 20–30% of hippocampal pyramidal cells will have been lost (Squire (1987). It is tempting, therefore, to link the neuronal cell loss in frontal and hippocampal regions to frontal and temporal memory systems which Shimamura (1994) has implicated in the aging process.

7.07.9.2 Age-associated Memory Impairment

Crook et al. (1986) proposed the concept of age-associated memory impairment (AAMI). The following criteria were developed:

(i) at least 50 years of age;

(ii) gradual onset of complaints of memory loss in everyday activities;

(iii) memory test performance one standard deviation below the mean of young adults on standardized tests of recent memory with adequate normative data;

(iv) adequate intellectual function (e.g., Scaled score of at least nine on WAIS-R vocabulary);

(v) absence of dementia (i.e., Mini-Mental Status Examination score of 24 or above); and

(vi) exclusion criteria, delirium, neurological disorder, cerebrovascular pathology, head injury, etc.

However, it is clear that a vast number of elderly people would meet these criteria, particularly the reference to scoring one standard deviation below memory test scores for young adults; indeed by definition over 16% of young adults would meet this criterion. This led to the proposed diagnostic category being severely criticized (e.g., O'Brien & Levy, 1992). Partly as a consequence of this type of criticism, AAMI was not included in the *Diagnostic and Statistical Manual of Mental Disorders* (*4th ed., DSM-IV*) of the American Psychiatric Association. Some skeptics felt that the condition may have been invented in order to legitimize pharmaceutical treatment of this new disorder. However, it is clear that aging is associated with significant impairment of memory functioning and more needs to be known about this naturally occurring phenomenon rather than labeling it as a categorical disorder. Caine (1992) has proposed that instead of the term AAMI, age-related cognitive decline (ARCD) should be discussed and researched and this has been included in *DSM-IV* (code 780.9).

7.07.9.3 Does Memory Functioning in Dementia Represent Accelerated Aging?

In many studies, the finding that memory after a delay is poorer in the elderly than in young subjects has usually been attributed to an acquisition deficit rather than to faster rate of forgetting. For many years, a controversy has raged over whether dementia (particularly of the Alzheimer type) represents accelerated aging effects or whether it represents a quite separate clinical entity (Huppert, 1994). In an important study, Huppert and Kopelman (1989) found that normal aging produces a mild acquisition deficit as well as a significant increase in forgetting rate, and the rapid rate of forgetting in the elderly cannot be attributed simply to differences in the initial level of acquisition. In a comparison of normal aging with dementia, normal aging produced a mild acquisition deficit and a slight increase in the rate of forgetting, whereas dementia superimposes a severe acquisition defect but no further effect on forgetting rate (Huppert & Kopelman, 1989). In their study, Greene, Baddely, and Hodges (1996) employed the Doors and People test (Baddeley, Emslie, & Nimmo-Smith, 1994; see below) in a comparison of early DAT patients vs. elderly controls. They concluded that the episodic memory deficit in DAT is general in nature and primarily reflects impaired learning rather than accelerated forgetting or disrupted retrieval. Thus, memory functioning in dementia appears to be qualitatively different from that observed in "normal" aging.

7.07.9.4 Commonly Used Memory Tests: Implications for use with Elderly Subjects

In the following sections several commonly used clinical memory measures will be outlined briefly and evaluated, with a particular emphasis on the appropriateness of their use with the elderly. For further details, the interested reader is referred to Lezak (1995) and Spreen and Strauss (1991). Unfortunately, pressure of space precludes coverage of all but a few of the most widely used instruments.

7.07.9.4.1 The Auditory Verbal Learning Test

The auditory verbal learning test (AVLT; Rey, 1964) is one of the most widely used word learning tests in clinical research and practice. Five presentations of a 15-word list are given, each followed by attempted recall. This is followed by a second 15-word list (list B), followed by recall of list A, and delayed recall and recognition are also tested. A key feature of the AVLT (and its successor, the California Verbal Learning Test) is that it affords the opportunity to measure rate of learning, as opposed to recall of a single stimulus, or series of stimuli. An equivalent form of AVLT has been provided by Crawford, Stewart, and Moore (1989c). Because of the rather abstract nature of the test, that is, 15 unconnected words and repetition over five trials, this can be a rather stressful measure for an elderly individual of failing cognitive abilities. The initial normative data that was referred to widely was rather limited (Lezak, 1983). However, the situation has improved, with excellent norms for the elderly provided by Geffen, Moar, O'Hanlon, Clark, and Geffen (1990), Ivnik et al. (1992), and Tuokko and Woodward (1996). The normative data on this test has also been reviewed by D'Elia, Boone, and Mitrushimo (1995). For a detailed review of the AVLT, see Peaker and Stewart (1989).

7.07.9.4.2 The California Verbal Learning Test

In an attempt to expand upon the assessment of learning and retrieval strategies, the California Verbal Learning Test (CVLT) evolved from the AVLT and was developed to provide an instrument "reflecting the multifactorial ways in which examinees learn, or fail to learn, verbal material. The CVLT's assessment of learning strategies, processes, and errors is designed for use in both clinical and research practice" (Delis, Kramer, Kaplan, & Ober, 1987). A variety of memory measures are obtainable from this measure including short- and long-term free recall and recognition, serial learning curves, learning strategy (i.e., semantic vs. serial clustering), etc. The CVLT also has the added advantage of presenting the stimuli in an everyday, relatively nonthreatening manner, that is, learning two shopping lists. The reliability data for the principal measures appear to be adequate although some of the scores which are derived from them have poor reliability. The normative database consists of 273 neurologically intact individuals, 104 males, 169 females, with a mean age of 58.9 (15.4) years. This mean age for controls is considerably higher than that normally employed for psychometric test development. However, the control subjects were recruited from several research projects and the authors state "in order to make the best use of the data while giving due regard to the varying number of cases at each age and sex, smoothed age curves were fitted to the raw data using multiple regression" (Delis et al., 1987, p. 35). This curve-fitting procedure for normative data has been criticized (see Elwood, 1991).

7.07.9.4.3 The Wechsler Memory Scale-Revised

The WMS-R (Wechsler, 1987) is generally considered to be a considerable improvement on its predecessor, the WMS. The WMS-R includes new subtests such as figural memory, visual paired associates, and visual memory span. As Mayes and Warburg (1990) comment, "the revised WMS therefore appears (potentially) to be considerably more satisfactory in a number of respects than its older brother—particularly in respect of normative data, information of standardization, validity and reliability, scoring criteria (especially logical memory), range of memory functions sampled and provision for clinical interpretation of test rests. This has been achieved at the cost of increased administration time, particularly if the Delayed Recall is tested, abandonment of provision of an equivalent alternate form, and increased cost of the test." For elderly subjects, however, the Delayed Recall Test requires an attempt to reach a criterion on both verbal and visual paired associates. This can lead to repeated experience of failure over six trials which can provoke distress in elderly individuals. In addition, the scoring of the tapping forwards and backwards subtest leads to considerable anxiety on the part of many psychologists as they find it an extremely difficult subtest to score. However, preliminary data on the inter-rater reliability of this subtest appears to be satisfactory (O'Carroll & Badenoch, 1994).

The normative sample consists of 316 subjects in six age groups, selected to be representative of the US population in terms of race, geographic region, education, and IQ. Elwood (1991) points out that the normative sample of 316 "seems hardly adequate for the standardization sample, given the broad age range of its target population. In relative terms, comparison between the WMS-R sample and the 2000 subjects used for the WAIS-R is striking" (p. 185). It is important to note that the WMS-R standardization sample systematically excluded subjects in the age range of 45–54 years using interpolated values in their place.

Given that the WMS-R is one of the most widely used memory tests, it would be expected that the normative sample would be representative of the general population. However, Huppert (1994) has pointed out that examiners were instructed to screen potential examinees for a list of several medical factors which may affect cognitive performance and has suggested that, not only were these criteria poorly specified, they almost certainly excluded a high percentage of normal nondemented older individuals. As she states, "this has serious consequences for the use of these norms in assessing whether or not an individual's memory performance is below that expected for their age. The effect will be to over-estimate the number of elderly people with clinically significant memory impairment" (p. 315). A further brief discussion of this issue can be found in Section 7.07.2.

7.07.9.4.4 The Rivermead Behavioural Memory Test

The Rivermead Behavioural Memory Test (RBMT) was designed specifically to try to detect impairment of everyday memory function by providing test items that resembled activities in everyday life, for example, remembering to deliver a message, remembering to retrieve a personal belonging after an interval, etc. The RBMT has the advantage of having four matched parallel versions, thus allowing for repeated assessment to determine the effects of disease progression and/or clinical intervention. An additional important advantage of the RBMT is the careful work the authors have undertaken in order to ensure the measure's validity. They assessed validity in three ways: (i) by demonstrating a high correlation between the RBMT and other standard memory tests; (ii) by demonstrating a high correlation between RBMT scores and subjective ratings of memory impairment; and (iii) most importantly, by demonstrating a high correlation between RBMT score and observer ratings of memory lapses (Wilson, Cockburn, Baddeley, & Hiorns 1989). Such a rigorous approach to demonstrating the ecological validity of a memory measure is both unusual and welcome.

In general, the RBMT is well tolerated by elderly individuals on account of its "everyday feel" and the fact that it consists of a number of brief, relatively nonthreatening subtests. The initial normative data were provided from 118 subjects aged 16–69 years. However, a supplement to the manual provides details of elderly healthy subject's performance and is, therefore, a welcome addition to the clinical neuropsychologist's armamentarium for the assessment of memory in the elderly adult. Cockburn and Smith (1989) recruited 119 people aged between 70 and 94 years, with a mean age of 80.5 years. In order to provide as random a sample as possible, every fourth name of people born in or before 1917 was taken separately from anyone from the age/sex register of a five-doctor general practice that covered both urban and rural areas. Such a detailed attempt to recruit a truly "normative" elderly sample is welcome. Furthermore, the authors determined that age and general intellectual ability were significant correlates of Rivermead performance. Accordingly, they developed useful tables for determining abnormal memory scores in relation to age and estimated premorbid intellectual level. This additional work renders the RBMT one of the best memory measures for use in an elderly sample. Furthermore, the nonthreatening nature of the stimulus materials makes it more user-friendly for an elderly sample than many of its competitors. A limitation of the RBMT is that it may be rather insensitive to mild impairments of memory. Following on from this, ceiling effects may be encountered in mildly impaired patients, making any positive effects of intervention difficult to demonstrate.

7.07.9.4.5 The Rey Complex Figure Test

This test has been used widely (Rey, 1964), particularly because of the dearth of adequate visuospatial memory tests. A complicated figure is presented and the subject is requested to copy it. The original and copy are then removed and the subject is asked to draw the figure again from memory, after varying delay intervals. As Mayes and Warburg (1990) point out there have been many variants on administration and scoring criteria; and the available norms, particularly for elderly samples, have been rather limited. Spreen and Strauss (1991), however, provide normative data from the age ranges 60–85 years. While this is a welcome addition, the numbers in each respective age band are again limited. For example, in the 70+ age range, the sample size is $n=10$. This obviously makes clinical interpretation somewhat difficult.

7.07.9.4.6 The Recognition Memory Test

The recognition memory test (RMT) is a widely used instrument, particularly in amnesia research (Warrington, 1984). It consists of two subtests, recognition memory for 50 words and recognition memories for 50 faces. Following presentation of the words and faces, the testee is required to perform an immediate two-choice recognition task and is generally considered to be less stressful and anxiety provoking than many other memory measures. The normative data consists of 310 control in-patients without cerebral disease, aged 18–70 years. As Nelson (1990) points out in a review of the RMT, "This test can only be used with subjects up to the age of 70 because of the unknown effects of normal aging on the test scores past this age."

7.07.9.4.7 The Doors and People Test

The Doors and People Test was devised in order to provide comparable measures of visual and verbal memory that test both recall and recognition, that does not provide floor or ceiling effects and includes both learning and forgetting measures (Baddeley et al., 1994). The test comprises four sections: a doors test, where single doors have to be recognized against three competing distracters; a shapes test, where diagrams have to be reproduced from memory; a names test, which is a verbal recognition test; and finally, the peoples test, where the subject is required to learn an association between a series of photographs and names. Normative data from 238 subjects is provided, in five age bands (16–31, 32–47, 48–63, 64–79, 80–97). In fact, one of the stated aims was that doors and people should "provide an unstressful test that is acceptable to a wide range of subjects, extending from patients suffering from dementia or dense amnesia to healthy young normal subjects" (Baddeley et al., p. 4). However, some elderly, cognitively impaired subjects may find this test quite demanding.

7.07.10 ASSESSING LANGUAGE DYSFUNCTION

In normal aging, deterioration of language has been considered to be either very marginal or a consequence of deterioration of other cognitive functions. Although the majority of elderly people do not develop any gross language disturbance, close scrutiny suggests that some subtle changes are indeed present (Kemper, 1992). In contrast, language deficits are a very pronounced aspect of cognitive deterioration in DAT. Aspects of language may also be disrupted by other forms of cortical and subcortical dementias (e.g., fronto-temporal dementias, Huntington disease, Parkinson's disease, vascular dementia, etc.).

Given the frequency of occurrence of language disorders after degenerative neurological diseases associated with old age, language should be investigated in depth. In the following sections, the most frequent language disturbances encountered either in normal aging or as a consequence of degenerative disorders will be reviewed briefly and suggestions made about appropriate assessment procedures. The examination of language competence should include assessment of verbal expression (naming, vocabulary, discourse and verbal fluency, repetition, prosody), verbal academic skills (reading, writing, spelling), and verbal comprehension (syntax, idioms, and proverb interpretation). As yet there is no comprehensive battery specifically designed for assessing language abilities in the elderly, therefore clinicians often fall back on general aphasia batteries. This can be problematic as the norms for the elderly are limited. In addition there is a need to derive individual comparison standards when assessing language function because of the large degree of variability in premorbid competence (see Section 7.07.4.1). Therefore, examiners should devote time in their initial interview to inquire about factors such as educational level, previous occupation, and reading habits.

7.07.10.1 Naming

Everybody has personal experience of looking for the name of an object, having a very strong feeling of exactly knowing what one wants to say but being unable to recall the correct word. This phenomenon is often referred to as "tip of the tongue." Word finding difficulty is the most common language symptom experienced by healthy people over 60 (Burke, Mackay, Worthley, & Wade, 1991). It is experienced occasionally by young people, but aging seems to exacerbate the problem. The standard task for assessing word finding abilities is confrontation naming, in which subjects are required to name a given object, or more commonly, to name objects depicted in line drawings. One of the most popular tests is the Boston Naming Test (BNT) (Kaplan, Goodglass, & Weintraub, 1983). The original version of this test included 85 items (Kaplan, Goodglass, & Weintraub, 1978); however, the shortened version produced in 1983 is used more frequently. This test consists of 60 line drawings of objects of graded difficulty, ranging from very common objects (e.g., a tree) to less familiar objects such as an abacus. Although the original normative data are scanty, several studies have provided supplementary norms; these studies have been reviewed comprehensively by D'Elia et al. (1995). Data collected from a large sample of 750 older adults (over the age of 55 years) showed that performance on the BNT was influenced by age which accounted for more than 10% of the raw score variance (Ivnik, Malec, Smith, Tangalos, & Petersen, 1996).

A study carried out with patients showed that correlation between the two (long and short) forms of the test and two 42-item forms (obtained by dividing the original test into two nonoverlapping forms) ranged from 0.96 to 0.92 (Huff, Collins, Corkin, & Rosen, 1986). A study (Mitrushina & Satz, 1995) carried out with healthy elderly individuals has reported that repeated testing with the BNT reveals a high stability of scores over time, suggesting a lack of practice effects.

Naming may also be assessed using the Object Naming Test (Newcombe, Oldfield, Ratcliff, & Wingfield, 1971) or the Graded Naming Test (Warrington, 1984). A detailed examination of naming should also include naming of actions (verbs), body parts, and colors. Naming of actions may be tested by means of formal action naming tests (McCarthy & Warrington; 1986; Obler & Albert, 1978). As an alternative, the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1987) provides a comprehensive examination of naming abilities. However, this test is time-consuming and is best suited to highly experienced examiners; furthermore, there is limited normative data for the healthy elderly.

Confrontation naming studies based on the BNT indicate that there is an increase in the variability of scores among apparently healthy individuals in their 60s, thereby suggesting impairment in a subset of this age group (Van Gorp, Satz, Kiersch, & Henry, 1986). There is consistent evidence of a steep decline in the late 70s (Goodglass, 1980; Mitrushina & Satz, 1995). This decline is evident with both low-frequency nouns and verbs but is more marked in the former (Obler, Albert, & Goodglass, 1981). Poor scores on confrontation naming can arise because of a failure to give any response or because of incorrect responses. Increases in the number of errors on confrontation naming tasks were observed in a couple of studies that compared the performance of over 80-year-old subjects with younger subjects (60- and 70-year-olds) (Albert, Heller, & Milberg, 1988; Borod, Goodglass, & Kaplan, 1980). Obler and Albert (1985) analyzed the qualitative features of errors made by the elderly and noted that they tend to be characterized by the use of circumlocutions.

Naming can also be studied using verbal fluency tasks which require the generation of words from semantic or phonemic cues. These tasks are reviewed in Section 7.07.13.2. Two hypotheses have been put forward to explain the naming problems that occur with normal aging. One argues for a degradation in the structure of lexical knowledge, whilst the other argues for an impairment in retrieval mechanisms (see Light, 1992 for a review). Experimental evidence shows that semantic priming effects are virtually identical for young and healthy older adults (Balota & Duchek, 1988; Burke, White, & Diaz, 1987; Monti et al., 1996); further, in naming tasks, performance both of young and old subjects is enhanced more by phonological cues than by semantic cues (Nicholas, Obler, Albert, & Goodglass, 1985). These findings suggest that structural semantic knowledge of words is intact and that a deficit of retrieval or access to this knowledge may be responsible for word-finding difficulties.

The situation is more complicated when trying to understand the severe word-finding difficulties that occur in DAT. Conflicting explanations have been provided for this, and evidence in favor of an access or retrieval deficit (Kempler, 1988; Neils, Brennan, Cole, Boler, & Gerdeman, 1988), or evidence of a structural disruption (Hodges, Patterson, Graham, & Dawson, 1996; Kempler, Andersen, & Henderson, 1995; Nebes, Brady, & Huff, 1989) has been reported. Further, in demented people, failure on confrontation naming could be due to visuoperceptual deficits that result in failure to identify and recognize an object, as suggested either by the fact that performance improves when additional sensory cues are provided (Barker & Lawson, 1968) or by the fact that errors consist of naming objects which are visually rather than semantically related to the target.

7.07.10.2 Vocabulary

As discussed earlier, structural knowledge about words does not change greatly in the normal elderly population, and vocabulary scores do not typically decrease in the normal aged (Salthouse, 1988), rather they may increase, albeit modestly. This evidence supports the notion of a preserved structural knowledge of words in the normal aged population. The most widely used instrument for the assessment of vocabulary ability is the Vocabulary subtest of the WAIS-R (Wechsler, 1981) which consists of 35 items organized in order of increasing difficulty. Performance on this test is greatly influenced by education (Malec, Ivnik, Smith, & Tangalos, 1992) and socio-cultural factors (Huttenlocher, Haight, & Bryk, 1991) rather than age. A test–retest reliability coefficient of 0.71 has been reported in a study of the healthy elderly (Snow, Tierney, Zorzitto, Fisher, & Reid, 1989).

7.07.10.3 Discourse

The assembling of a meaningful and well-structured discourse requires lexical and syntactic competence, including a knowledge of pragmatics (i.e., discourse structure and rules), adequate sequencing skills, including the ability to integrate concepts within and between sentences, and adequate working memory capacity. Material from the clinical interview can be used to assess the quality of discourse and this can be supplemented by asking a patient to describe a standard picture or relate a well-known story (e.g., Little Red Riding Hood). As a standard picture, the Cookie Theft Picture of the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1983) may be used. The examiner should assess whether the production is assembled coherently, if elements are connected and sequenced in the appropriate order, whether an individual makes correct use of reference pronouns, and whether appropriate elements are produced. Performance should also be evaluated in terms of the complexity of sentences, grammar and syntax, and richness of vocabulary.

Aging does not have macroscopic effects on discourse; some subtle deficits, however, may be present in the use of pronouns in discourse, with older subjects using pronouns without specifying their referent in spontaneous speech, a situation that especially involves memory (Pratt, Boyes, Robins, & Manchester, 1989). It has also been reported that deficits on discourse production reflect changes in those macrolinguistic skills that require integration of linguistic and nonlinguistic cognitive processes rather than language-specific cognitive processes (Glosser & Deser, 1992). It was observed that elderly subjects did not differ from young subjects on microlinguistic aspects such as syntactic complexity, lexical production, and use of lexical cohesive ties; however, older individuals failed to refer coherently to the general topic. Discourse production is impaired significantly in patients with DAT, even in the very early stages, and is characterized by incoherence and emptiness of speech; spontaneous speech therefore remains fluent but carries less information (Giles, Patterson, & Hodges, 1996; Ripich & Terrell, 1988).

7.07.10.4 Verbal Comprehension

Comprehension is usually well preserved in healthy elderly individuals, and apparent deficits are often simply a reflection of poor auditory acuity (see Section 7.07.2). The most common test of verbal comprehension is the Token Test (De Renzi & Vignolo, 1962). The original version of this test consists of 62 items but a 36-item short-form (De Renzi & Faglioni, 1978) is more commonly used in clinical practice. The materials consist of tokens which differ in color, shape (squares and circles), and size (large and small). The examinee has to follow verbal instructions which increase in complexity from simple commands (e.g., "Touch a circle," "Touch the red circle") to commands such as "Before touching the yellow circle, pick up the red square." The test authors recommend that scores be adjusted for years of education. Adjusted scores of 25–28 are regarded as indicating mild comprehension problems and 17–27 moderate problems; scores below this are classified as severe or very severe. This test is extremely sensitive to impairment of linguistic processes. De Renzi and Faglioni (1978) reported a true positive classification rate of 93% when assessed against results from an extensive aphasia examination; the true negative rate was 95%.

Test–retest reliability with aphasic patients is very high, ranging from 0.92 to 0.96 (Gallagher, 1979; Orgass, 1976). There are inconsistent reports about the effects of age and education on performance, some suggesting a moderate increase in number of errors with aging (see Lezak, 1995).

Memory and attention play an important role in performance on comprehension tests. For example, working memory load is high on the later items of the Token Test. Therefore, results obtained from measures of attention and memory must be integrated with the results from tests of comprehension when reaching a formulation regarding the nature of an individual's difficulties.

7.07.11 ASSESSING VISUOPERCEPTUAL DYSFUNCTION

Visuoperceptual deficits represent a very common finding in sufferers of degenerative cortical illnesses. A comprehensive evaluation of visuoperceptual abilities should include assessment of color perception, visual recognition, visual organization, and visual interference. A first step in assessing visuoperceptual functions should be to test for the presence of unilateral neglect; this topic is covered in Section 7.07.12.1.

7.07.11.1 Visual Recognition

Evaluation of visual recognition abilities includes assessment of several aspects. It includes, first of all, assessment of perception of angulation. The most widely used test in this case is the Judgment of Line Orientation Test (Benton, Hannay, & Varney, 1975). This test includes 30 pairs of angled lines to be matched with the corresponding lines in a radially arranged 11-line display. An advantage of this test over many clinical tests used to assess visuospatial functioning is that it imposes minimal demands on praxic, mnestic, or executive functions. The norms include correction factors for age and education but are based on only 144 subjects. Age effects on this test have been confirmed by Eslinger and Benton (1983); test–retest reliability over a one-year delay in an elderly sample was modest ($r = 0.59$) (Levin et al., 1991).

The recognition of objects depicted under conventional and unusual views should be examined. Warrington and Taylor's (1973) Unusual Views Test can be used for the latter. Assessment of visual recognition also includes testing of face recognition abilities. For this purpose the Benton Facial Recognition Test (BFRT) (Benton, Hamser, Varney, & Spreen, 1983; Benton & Van Allen, 1968) may be administered. The original version of this test includes 22 items; it is a matching task rather than a true recognition task since it requires matching a front-view photograph of a target face to a matrix of faces photographed either in front-view, three-quarter-view, or under different lighting conditions. Norms are available up to the age of 74 years. Some degree of age-related decline in performance on this test has been observed (Benton, Eslinger, & Damasio, 1981) and the norms have correction factors for age and education. Further evidence of age-associated impairment at several stages of the face recognition process has been reported (Maylor, 1990). It should be noted that prosopagnosics can perform within normal limits on the BFRT. However, they take an inordinate amount of time per trial as their performance is characterized by a piecemeal strategy in which each feature of the target face is compared with each feature of the faces in the matrix (Ellis, 1992).

7.07.11.2 Visual Organization

Making sense out of fragmented, incomplete, or ambiguous figures requires intact visual organization abilities. Those abilities may be tested using tasks such as the Street Completion Test (Street, 1931) or the Hooper Visual Organization Test (HVOT) (Hooper, 1983). The HVOT requires examinees to name 30 fragmented objects. Test–retest reliability in an elderly sample was 0.68 (Levin et al., 1991) and the original norms cover up to age 68 years. Studies which have investigated this task in elderly subjects (55–92 years of age) indicate that performance drops off significantly with age (Farver & Farver, 1982; Montgomery & Costa, 1973; Villardita, Cultrera, Cupone, & Mejia, 1985).

7.07.11.3 Visuoconstructional Ability

Visuoconstructional abilities are usually assessed by means of copying or free drawing. The Rey-Osterreith Complex Figure (Osterreith, 1944) is the most widely used copying test (for further details see Section 7.07.9.4.5). The Block Design subtest of the WAIS-R may be used for assessing two-dimensional constructional abilities.

7.07.11.4 Visual Interference

Visual interference may be evaluated by means of the Hidden Figures Test (Gottschaldt, 1928) which requires a subject to recognize a simpler figure, hidden in a more complex one, by marking its outline. Normative data up to the age of 79 years are provided in Spreen and Strauss (1991). Another useful and well-known task for assessing visual interference is the Overlapping Figure Test (Poppelreuter, 1917). An effect of normal aging on performance on this test has been reported in a recent study (Della Sala, Laiacona, Trivelli, & Spinnler, 1995) in which severely impaired performance was observed in a sample of patients with DAT.

7.07.12 ASSESSMENT OF ATTENTION

7.07.12.1 Unilateral Neglect

Unilateral cerebral lesions (especially right parietal lesions) often results in unilateral neglect. Neglect labels a phenomenon that involves lack of awareness of stimuli appearing on the side contralateral to the cerebral lesion. It is accepted widely that neglect is not a unitary phenomenon and evidence of dissociations across modalities within the same patient and between patients have been reported (Halligan & Robertson, 1992).

Assessment of neglect should explore representational and motor abilities as well as visuospatial abilities. Personal (relating to the subjects body), peripersonal (within arms length), and extrapersonal space should be investigated. A dissociation between personal peripersonal space and extrapersonal space has been illustrated by Halligan and Marshall (1991) who described a case with severe neglect in the former spheres but who nevertheless failed to exhibit neglect for distant experimental stimuli and could aim successfully at areas on a dart board which were to the left of the midline.

The most frequently administered test for assessing visuospatial exploration is the Albert Test (Albert, 1973) which consists of 40 lines drawn out at various angles on a sheet of paper; the subject's task is to cross out all of them. Normal subjects usually perform flawlessly on this test. Neglect may also be assessed using the Bells Test (Gauthier, Dehaut, & Joanette, 1989) that includes 315 silhouettes of common objects, of which 35 are bells; the subject's task is to search for the bells. Normative data (including data from subjects aged 50–89 years) are available for this test, and no more than three bells are usually missed by normal elderly subjects. Test–retest reliability over a two-week period was reported as 0.69.

Line Bisection is also a widely used task for assessing neglect; it requires a subject to divide horizontal lines at the center point. Performance on this task can be highly variable, and it is advisable to acquire data from multiple trials. Patients with right hemisphere lesions show a right-deviation error of the bisection point, that increases as a function of the length of the line (Halligan & Marshall, 1989).

Representational neglect, that involves mental imagery, may be assessed by asking a patient to describe a well-known scene from a particular perspective. A patient may fail to recall objects on the left but recall the same objects when asked to imagine the scene from the opposite perspective (Bisiach & Luzzatti, 1978).

A comprehensive battery for testing visual neglect has been assembled by Wilson, Cockburn, and Halligan (1987). This battery is called the Behavioural Inattention Test (BIT) and includes conventional tests (star cancellation, letter cancellation, figure copying, line crossing, line bisection, and representational drawing) and behavioral subtests (picture scanning, telephone dialing, menu reading, article reading, telling and setting the time, coin sorting, address and sentence copying, map navigation, and card sorting). The test manual reports impressively high coefficients for inter-rater reliability (0.99), test–retest reliability (0.99) and parallel form reliability (0.91). Among the six conventional tests included in the battery, star cancellation had substantially greater sensitivity than the other tests (Halligan & Marshall, 1988) (see Figure 1).

Although many studies have explored the performance of normal subjects on some of the tests commonly used to assess neglect, relatively little is known about the effect of normal aging and dementia on these tasks. Neglect is seldom found in DAT, despite the frequency of visuoperceptual disorders in this condition; Huff et al. (1987) reported an incidence of 3% based on 95 cases. A high incidence of neglect has been reported by one study of visuospatial ability in DAT (Freedman & Dexter, 1991). These authors concluded that neglect in DAT may not be an unusual phenomenon but could often be missed because of a failure to conduct a detailed assessment. However, it should be noted that neglect was considered to be present if the discrepancies between errors on left- vs. right-sided stimuli on visuospatial tasks exceeded control values (these control values were simple difference scores). As there typically will be a high number of errors on visuospatial tasks in DAT cases, and often they will exhibit considerable intrasubject variability, these results may be an artifact of Freedman and Dexter's (1991) scoring methods. This alternative explanation is strengthened by the fact that the number of cases classified as exhibiting right-sided neglect was exactly equivalent to the number classified as exhibiting left-sided neglect (right-sided neglect is relatively rare). When looking at individual reports of cases of DAT displaying neglect, only three cases have been reported (Ishiai, Okiyama, Koyama, & Seki, 1996; Pentore & Venneri, (1994). Pentore and Venneri (1994) in their study reported two patients with DAT of moderate severity who manifested neglect in the later stages; one of these patients manifested motor neglect in addition to visuospatial neglect. Such limited evidence may lead to the conclusion that neglect in DAT is actually rather rare. This would be in accord with the fact that neglect is generally observed after unilateral lesions, while DAT, although it may present some degree of asymmetry at onset, involves bilateral degeneration (see Kirk & Kertesz, 1991, for further discussion).

7.07.12.2 The Anterior Attention System

The anterior attention system is considered to underlie performance on tasks of sustained, divided, and selective attention (Posner & Petersen, 1990). Such attentional problems are common and debilitating features of many neurological and psychiatric disorders associated with aging (Crawford, McKinlay, & Parker, 1992b; Lezak, 1995; Walsh, 1991). However, the assessment of attentional dysfunction is a problematic area in clinical neuropsychology as many potentially useful tasks used in experimental studies have not been translated into practical clinical instruments. Many existing tests do not reflect theoretical knowledge on attention and its fractionation (Posner & Petersen, 1990) and, in many cases, are poorly standardized (Lezak, 1995). An additional consideration is that tests such as the PASAT, which have proved valuable in work with younger clinical populations because of their demonstrated sensitivity and predictive validity, are too demanding to be suitable for use with the majority of elderly clients. Because of space limitations, we will limit coverage to a single battery, the Test of Everyday Attention

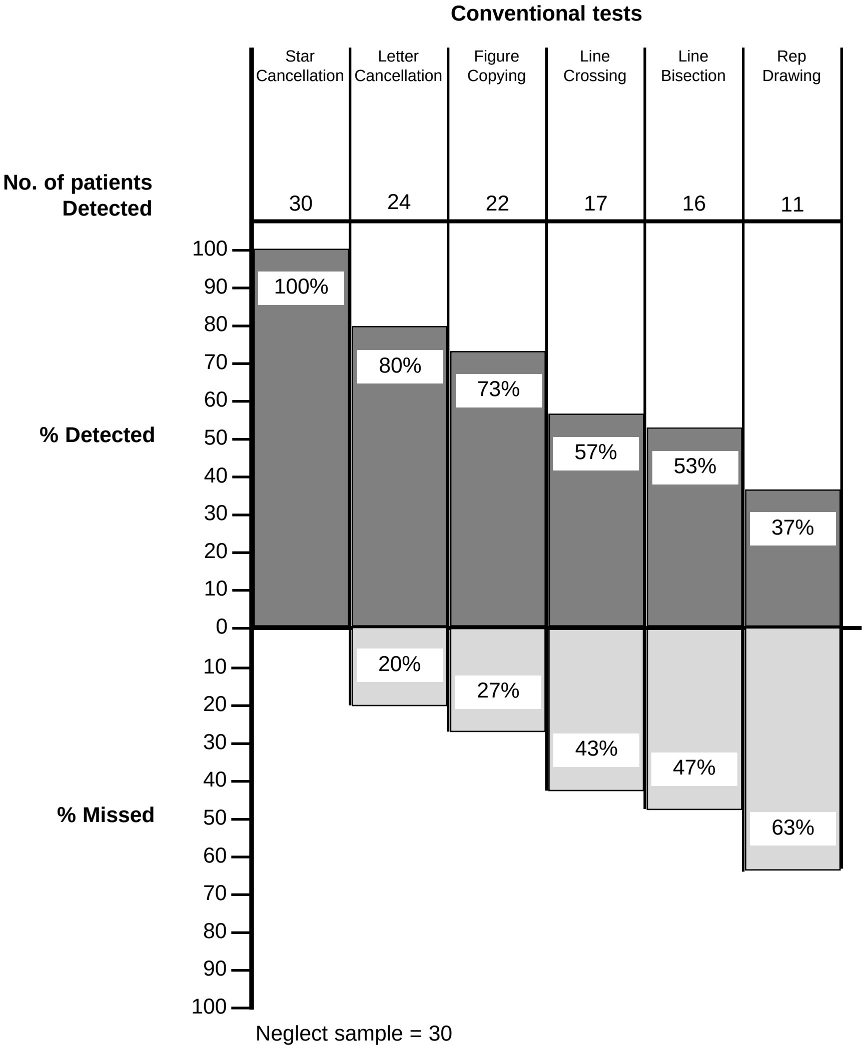

Figure 1 The relative sensitivities of the six conventional screening tests of the Behavioral Inattention Test (reproduced from Halligan & Robertson, 1992).

(TEA) (Robertson, Ward, Ridgeway, & Nimmo-Smith, 1994; Robertson, Ward, Ridgeway, & Nimmo-Smith, 1996).

TEA is a welcome development as it attempts to overcome some of the aforementioned problems in the assessment of attentional dysfunction. It consists of eight subtests, standardized to have means of 10 and SDs of 3, which measure sustained, selective, and divided attention. For example, in the Map Search subtest, which was designed as a measure of visual selective attention, the examinee has to search for symbols (e.g., a knife and fork representing eating facilities) on a tourist map of a city. The Elevator Counting subtest, which was designed to measure sustained attention, is a tone counting task in which the examinee is asked to imagine they are in an elevator in which the floor indicator has failed (thus counting the tones is the only way to establish which floor the elevator is on). A related subtest, designed to measure auditory selective attention, requires the examinee to ignore distracter tones which are of a higher pitch.

The TEA was normed on 154 healthy participants aged 18–80 years. In addition to reflecting current thinking on the fractionation of attention, it is designed to be ecologically valid, that is, many of the subtests are designed to mimic everyday activities. Another advantage is that it has three parallel versions; this avoids the interpretative problems encountered when attempting to measure change using the same test materials.

Crawford, Somerville, and Robertson (1997) have prepared tables which allow clinicians to examine a client's relative strengths and weaknesses on the TEA (in terms of the reliability and abnormality of subtests deviations from the subtest mean), thereby supplementing the conventional normative comparison standards presented in the test manual with individual comparison standards.

7.07.13 ASSESSMENT OF EXECUTIVE FUNCTIONS

Executive deficits arising from damage to the prefrontal cortex and related structures can have a much more profound effect on a client's prospects for successful adjustment and independent living than the more circumscribed deficits arising from posterior cortical lesions. However, these problems have proven difficult to quantify and as such can be regarded as constituting the most problematic area in neuropsychological assessment. For example, reviews of tests such as the Wisconsin Card Sorting Test (Grant & Berg, 1948; Heaton, 1981), which are aimed specifically at capturing the core cognitive problems (e.g., planning deficits, underutilization of feedback, etc.), indicate that existing indices are of limited utility when used with individuals (Mountain & Snow, 1993; Reitan & Wolfson, 1994). Thus, considerable reliance must be placed on careful scrutiny of a client's approach to a range of tasks to identify these cognitive difficulties.

7.07.13.1 Frontal Hypotheses

Executive deficits are thought to underlie many of the behavioral disturbances seen in neurological disorders associated with increasing age. One of the more obvious examples is Parkinson's disease (PD). Executive deficits would be expected in PD on pathophysiological grounds as the basal ganglia, which are deprived of dopaminergic innervation because of degeneration in the substantia nigra, have strong reciprocal connections with the frontal cortex. In addition, the ventral tegmental area, which provides a direct dopaminergic innervation of the prefrontal cortex, also deteriorates in PD (see Crawford, Besson & Ebmeier, 1990 for brief review). There is neuropsychological evidence which suggests that executive deficits are an early feature in PD and that these deficits qualify as differential deficits in executive functioning (Lees & Smith, 1983). Indeed, it has been argued that fronto-executive dysfunction provides a parsimonious explanation for all of the cognitive deficits observed in nondemented Parkinson's cases (Della Sala, 1988). This is an extreme position but it is interesting to note that Bondi, Kaszniak, Bayles, and Vance (1993) have demonstrated that deficits in other cognitive domains were no longer significant after partialing out performance on tests of executive function.

It has also been proposed that normal aging is associated with a selective decline in fronto-executive function (Dempster, 1992). Daigneault, Braun, and Whitaker (1992), for example, note that "The currently dominant neuropsychological model of normal brain aging postulates that cognitive functions dependent on the integrity of the prefrontal brain regions are among the first to deteriorate" (p. 99). This frontal hypothesis stems from indications that age-related changes in the neuroanatomy and neurochemistry of the brain are more pronounced in the frontal lobes than in other cortical areas (Fuster, 1989) and from demonstrations of age-related deficits on neuropsychological tasks sensitive to frontal dysfunction (Daigneault & Braun; 1993; Daigneault et al., 1992; Mittenberg, Seidenberg, O'Leary, & DiGiulio, 1989). However, there is only limited evidence that such deficits qualify as differential deficits (i.e., that they significantly exceed the deficits found in other cognitive domains).

7.07.13.2 Verbal Fluency

Verbal fluency tests, and in particular initial letter fluency, are widely used as tests of executive dysfunction (Parker & Crawford; 1992; Walsh, 1991). Initial letter fluency, also referred to as the Controlled Oral Word Association Test (COWAT) (Benton; 1969) or phonemic fluency, requires the generation of words from initial letters (normally F, A, and S or C, F, and L) under time constraints (normally 60 seconds per letter, although 90 seconds is also used). Fluency tests are reliable, quick to administer, and even patients with quite severe deficits can understand the task requirements. In an area replete with failures to replicate, studies of initial letter fluency have been remarkably consistent in demonstrating impaired fluency following left or bilateral damage (Benton, 1968; Bornstein, 1986; Miceli, Caltagirone, Gainotti, Masullo, & Silveri, 1981; Miller, 1984; Milner, 1964). In view of the above, fluency tests are among the most useful tests in the clinician's armamentarium.

Performance on verbal fluency tasks reflects the integrity of the semantic system (i.e., the system responsible for storage of words and their associations), the efficiency of retrieval strategies, and the efficiency of self-monitoring and inhibition of inappropriate responses. In an early influential paper, Perret (1974) reported significant deficits on initial letter fluency in patients with focal frontal lesions whilst noting that semantic or category fluency (i.e., generating exemplars from categories such as animals, fruits, etc.) often failed to expose deficits (Newcombe, 1969). Perret (1974) and others have argued that the differential sensitivity of initial letter fluency to frontal dysfunction stems from the fact that in everyday life words are retrieved predominantly on the basis of their meaning. Thus, initial letter fluency is relatively novel and places greater demands on strategy generation and inhibition of semantic associates. In contrast, semantic fluency places greater demands on the integrity of the semantic system. Hodges, Salmon, and Butters, (1990), for example, is one of many studies in which there was evidence of a differential deficit in semantic fluency vs. initial letter fluency in DAT. This is consistent with other evidence indicating that deterioration of the semantic system is a marked feature of DAT (e.g., Grober, Buschke, Kawas, & Fuld, 1985).

Further support for this distinction comes from dual task studies of initial letter and semantic fluency in healthy participants. Martin, Wiggs, Lalonde, and Mack (1994) demonstrated that a secondary task involving motor sequencing (designed to activate frontal cortex) produced a differential deficit in initial letter fluency, whereas a secondary task involving semantic decisions regarding visual stimuli (designed to activate posterior cortex) produced a differential deficit in semantic fluency.

Some very useful normative data for the elderly have been published. Ivnik et al., (1996) present age- and education-based norms based on 743 elderly participants using the COWAT (i.e., the letters C, F, and L with 60 seconds per trial). Particularly positive features of these data are the age range covered (55–97 years) and the fact that most participants were also administered other commonly used neuropsychological instruments. This latter feature makes the data useful when attempting to construct a profile of a client's relative strengths and weaknesses as raw scores are converted to a common Scale and there is no worry that observed discrepancies are artifacts of differences between normative samples (see Section 7.07.4.2). Kozora and Cullum (1995) have published elderly norms for initial letter fluency (F, A, and S) and semantic fluency covering the age range 50–90 years. The semantic fluency test used was the supermarket item list from the Mattis Dementia Rating Scale (Mattis, 1988).

Errors on verbal fluency tasks have been afforded great significance in both the research literature and in advice offered to clinicians (Walsh, 1991). Perseverative errors (repeating a word produced previously) and rule-break errors (e.g., producing a proper noun or a word beginning with the wrong letter on initial-letter fluency trials) have been regarded as evidence of disinhibition or a failure of self-monitoring (Crowe, 1992; Walsh, 1991). Crowe (1992) has produced data to suggest that such errors tend to be associated with orbito-frontal damage, whereas impoverished but legitimate output is more indicative of damage to the lateral convexity.

Given the importance attached to such errors, there is surprisingly little normative data on errors to assist clinicians interpret an individual client's performance. Kozora and Cullum's (1995) study provided means and ranges for perseverative and rule-break errors. Although the mean and modal rates were low in all age bands, a finding consistent with a COWAT normative study conducted by Ruff, Light, Parker, and Levin (1996), the ranges are quite substantial (e.g., 0–15 in the 50–59 age group). It would be useful to know how many errors are required to exceed the 85th or 95th percentiles of the healthy population but there is no relevant published data.

There is a strong relationship between verbal fluency and general verbal ability in the general population. For example, Miller (1984) reported a correlation of 0.87 between initial letter fluency and Verbal IQ in a healthy sample ($N = 30$). Crawford, Obonsawin, and Bremner (1993) reported a lower but still substantial correlation of 0.64 between Wechsler Verbal IQ in a sample of 144 healthy participants. This indicates that an individual's premorbid verbal ability must be considered as an important factor when interpreting fluency performance in clinical populations. This was graphically illustrated by Borkowski, Benton, and Spreen (1967) who reported that the fluency performance of a "brain-damaged" sample of above average verbal IQ was significantly higher than the performance of healthy subjects of below average verbal IQ. Crawford et al. (1992c) have provided a NART-based regression equation which provides an estimate of a client's premorbid fluency performance (see Section 7.07.8 for details of the NART). This equation was derived from a healthy sample ($N = 142$) with an age range of 16–88 years.

7.07.13.3 Nonverbal Fluency

A number of attempts have been made to develop methods of assessing nonverbal fluency. Jones-Gotman and Milner (1977) developed a task they termed Design Fluency which requires examinees to draw abstract (nonsense) designs under a time constraint. This test can capture the lack of creativity and spontaneity and the perseverative tendencies which are often striking clinical features of patients with frontal damage. The disadvantages of this test include the lack of adequate normative data, the highly subjective nature of scoring, and the fact that it can be difficult to maximize cooperation, either because clients fail to grasp the nature of the task or view it as odd and lacking in relevance.

Ruff and colleagues (Ruff, 1988; Ruff, Light, & Evans, 1987) have developed a figural fluency test (hereafter RFFT) which requires the examinee to join up a repeating sequence of dots such that each design formed is unique. The procedure removes subjectivity from scoring as each design is unequivocally unique, a perseveration, or (more rarely) an error. The RFFT was normed on a sample of 358 healthy individuals with an age range of 16–70 years. Group studies of clinical populations indicate that the RFFT is sensitive to anterior cerebral damage, particularly in the right hemisphere (Lee et al., 1997; Ruff, Allen, Farrow, Nieemann, & Wylie, 1994). For example, Lee et al. found that 35% of designs produced by a group with right frontal lesions were perseverations, a figure which substantially exceeded the rate in right temporal (14%) and left frontal (17%) cases.

Motor slowing is a common consequence of cerebral damage and should be considered as a factor when interpreting the number of legitimate designs produced on the RFFT. Ruff, Evans, and Marshall (1986) have reported that significant RFFT deficits are observed in neurological samples after controlling for motor speed using a finger-tapping test. However, the appropriateness of finger-tapping as a control measure could be questioned. Crawford, Wright, and Bate (1995) used a simple motor task which required subjects to join repeated pairs of dots; the motor demands of this task are very similar to the RFFT but it does not require the generation of designs. They reported that the RFFT deficit in a head-injured sample did not significantly exceed the deficit on the motor task.

7.07.13.4 Behavioral Assessment of the Dysexecutive Syndrome

A very useful battery for assessing executive problems has been developed (Wilson, Alderman, Burgess, Emslie, & Evans, 1996; Wilson, Evans, Emslie, Alderman, & Burgess, in press). The Behavioral Assessment of the Dysexecutive Syndrome (BADS) battery consists of six subtests and was normed on a sample of 216 healthy participants with an age range of 16–87 years. As was noted previously, many of the formal neuropsychological tests can fail to detect executive problems because they are highly structured. Shallice and Burgess (1991) note "The patient typically has a single explicit problem to tackle at any one time ... the trials tend to be very short ... task initiation is strongly prompted by the examiner and what constitutes successful trial completion is clearly characterized" (pp. 727–728). Although some of the tests reviewed above clearly attempt to deviate from this formula, the BADS represents the most ambitious and systematic attempt to date to capture the core elements of dysexecutive syndromes.

The BADS subtests include the Rule Shifts Cards Test in which a previously established response set (responding "yes" to red cards, "no" to black) has to be inhibited in favor of responding in terms of whether or not a card matches the color of the card immediately preceding it. The Action Program Test is a planning task in which the solution requires the client to utilize various everyday materials (e.g., plastic, cork, and wire). The Modified Six Elements Test (see Shallice & Burgess, 1991) assesses scheduling and time management by requiring clients to tackle three different tasks within the time limit; there are two versions of each task and the rules prohibit tackling these contiguously.

The inter-rater reliability of the BADS is excellent; correlations between raters ranged from a low of 0.88 on one index from the Modified Six Elements Test, to unity, or near unity, on most of the other tasks. Test–retest reliability coefficients obtained from a small subset of the standardization sample (retested over a 6–12 month interval) ranged from modest to poor but this, in part, reflects the effects of range restriction.

A useful supplement to the formal BADS subtests is the Dysexecutive Questionnaire (DEX). The DEX covers dispositional and cognitive changes and comes in two forms, one for completion by the client, the other by a relative or carer. A provisional factor analysis indicates that the DEX measures three broad factors labeled as "Behavior," "Cognition," and "Emotion." Large discrepancies between client and carer reports of change are common when working with clients who have executive problems. The DEX provides one means of quantifying this lack of insight; DEX results can also form a basis for discussion when counseling clients and their families.

The DEX has been used to examine the relationship between carers' ratings of executive problems and formal test scores. Multiple regression analyses found that BADS scores were the only significant predictors of rated problems from among a set of measures which included traditional "frontal" tasks and measures of current and premorbid general intellectual ability.

7.07.14 THE DIFFERENTIAL DIAGNOSIS OF DEMENTIA VS. DEPRESSION

The differential diagnosis of dementia vs. depression has been described as "probably the knottiest problem of differential diagnosis" (Lezak, 1995), yet is a common reason for elderly patient referral to clinical psychologists. Depression is the functional disorder which can most closely resemble the early stages of a dementing condition. It is vitally important for an accurate differential diagnosis to be made because as Woods and Britton (1985) point out, "Depression can be treated; dementia typically leads to therapeutic despair." Des Rosiers (1992) reviewed 18 studies and reported that, on average, 10% of cases initially diagnosed as organic dementia were later rediagnosed as depression. In practice, the diagnosis of dementia vs. depression is made by drawing on information from a wide variety of sources, that is, previous history, informant's account, phenomenology, brain scan measures, blood biochemistry, etc. Problems emerge, however, when an individual presents with no well-documented history and no useful informant information.

Huppert (1994) has pointed out that in most studies where an early DAT group has been compared with a control group, it is likely that the controls vary from the demented subjects in ways which are rarely specified. For example, as the normal controls are usually community volunteers, they are less likely to have mobility problems than sensory deficits and are more likely to be in good health. The patients usually have been referred to specialists which means that dementia is sufficiently severe to have come to specialist attention, and second, that the patients who see specialists may not be representative of all demented subjects. As Huppert (1994) states "one implication of these considerations is that observed differences are likely to be an exaggeration (if not a distortion) of the performance which would be seen in a genuinely representative sample" (p. 295).

7.07.14.1 Delayed Verbal Recall Impairment as an Indicator of Early DAT?

Over the years there have been many attempts at using memory tests to help distinguish between dementia and depression. In order to highlight the problems associated with this approach, in the following section the Delayed Word Recall test will be used as an example.

In two studies by the consortium to establish a registry for DAT (CERAD), the investigators reported that delayed verbal recall was a highly sensitive indicator of early DAT (Morris et al., 1989; Welsh, Butters, Hughes, Mohs, & Heyman, 1992). In DAT, there appears to be a rapid rate of forgetting within the first 5–10 minutes following acquisition (Butters et al., 1988; Hart, Smith, & Swash, 1988) but a relatively normal rate of forgetting thereafter (Corkin et al., 1984; Kopelman, 1985). Knopman and Ryberg (1989) suggested that elaborative encoding may provide a substantial benefit in normal elderly individuals but not to patients with DAT. They proposed that the optimum memory test procedure for the early diagnosis of DAT should involve: (i) a study phase in which subjects were required to engage in elaborative encoding; (ii) a filled delay interval; and (iii) a test phase which involved free recall. Accordingly, Knopman and Ryberg (1989) developed the Delayed Word Recall test (DWR). Ten nouns are presented on cards, one at a time (chimney, salt, harp, button, meadow, train, flower, finger, rug, book). For each word, the subject is asked to read out the word and try to remember it and then to make up a sentence using the word. When all 10 words have been presented, the complete procedure is repeated, with subjects instructed to produce an alternative sentence. Following a five minute filled delay, free recall for the 10 words in the list is tested. Based on the delayed free recall cut-off score of less than or equal to 2 out of 10 = DAT, Knopman and Ryberg (1989) reported an overall predictive accuracy of 95% in a group of DAT patients vs. elderly control subjects. This is an extremely impressive degree of between-group separation.

In a follow-up study, Coen et al. (1996) compared a group of mildly demented DAT patients with individually matched healthy controls and found that the DWR recall measure achieved 98% overall accuracy (i.e., almost perfect between-group separation). In an extension to the original DWR, Coen et al. included a four-choice recognition memory test and, using a cut-off score of less than or equal to 9 out of 10 = DAT, greater than nine = controls, they reported that the recognition measure yielded a sensitivity of 98%, specificity of 95%, and an overall accuracy of 96%. Taken together, the results of the Knopman and Ryberg (1989) and Coen et al. studies suggests that the DWR is an extremely promising measure to aid in the early detection of DAT.

However, for any neuropsychological test index to have clinical utility, the method must be shown clearly to identify patients suffering from DAT without producing false-positives from an elderly depressed comparison group. Unfortunately, too many studies have appeared which have focused on the simple comparison between DAT and very healthy controls, and this criticism can be leveled at the Knopman and Ryberg (1989) and Coen et al. studies.

O'Carroll, Conway, Ryman, and Prentice (1997) attempted to test the utility of the DWR in separating patients with DAT from patients with major depression. Fifty patients with DAT were compared with 50 patients who fulfilled *DSM-IV* criteria for a major depressive episode. While there were highly significant differences in mean free recall and recognition scores, the between-group overlap was large (e.g., see Figure 2 for free recall results—very similar results were obtained for recognition). Put simply, if the recommended cut-off point for free recall of 2 or less out of 10 = DAT is adopted, 44% depressed patients would be misclassified as suffering from DAT. Using the less than 10 recommended cut-off score for recognition score, 48% of depressed patients would be misclassified as suffering from DAT. This study highlights the degree of cognitive impairment which often presents in elderly depressed patients which can masquerade as "pseudo-dementia" (Robbins, Joyce, & Sahakian, 1992).

7.07.14.2 Kendrick Cognitive Tests for the Elderly

The Kendrick Cognitive Tests for the Elderly (KCTE) comprise two components, a pictorial memory task (the Object Learning Test or OLT) and the Digit Copying Test (DCT). The principle underlying the use of the KCTE in differential diagnosis lies in the expected pattern of visual memory and psychomotor speed in depression vs. dementia. According to Kendrick, Gibson, and Moyes (1979), both DCT and OLT should be affected in primary dementia and not in depression. An innovation of the Kendrick test is the requirement that the measures be readministered six weeks following the initial assessment. desRosiers (1992) describes this requirement as somewhat less than practical. However, in our opinion, repeat neuropsychological assessment in cases where diagnosis is in doubt represents good clinical practice.

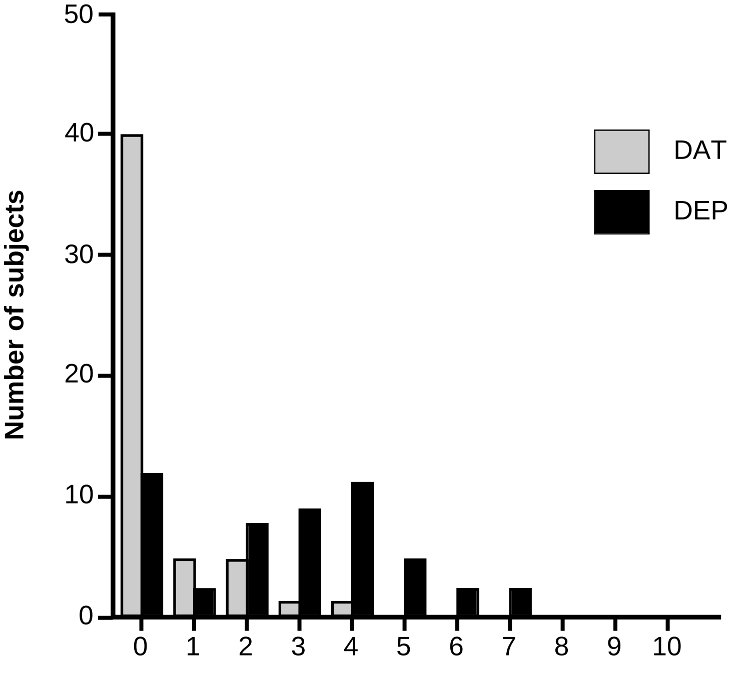

Figure 2 DWR Recall results. Using < 3 recommended cut-offs, only 2 (4%) DAT (Alzheimer) sujects would be misclassified, but 22 (44%) DEP (Depressed) would be classed as DAT (reproduced from O'Carroll et al., 1997).

One of the strongest claims regarding the KCTE is that "it has been found possible to discriminate absolutely between dementias and non-dementing subjects" (Kendrick et al., 1979, p. 329). The literature on the Kendrick test is well reviewed by desRosiers (1992). Some reports have confirmed the utility of the KCTE in differentiating a depressed sample from a demented sample. However, as is usually the case in studies of this sort, the samples consist of very well-defined groups where differential diagnosis is not a problem (Knight & Moroney, 1985). As desRosiers (1992) cautions, the validity of the KCTE should best be examined in patients for whom diagnosis is still ambiguous. A further point to consider when considering the theoretical foundation of the KCTE is the assumption that in depression, psychomotor speed, and mnemonic function are not as detrimentally affected as in early DAT. However, in depressive pseudodementia, it is precisely because both domains are seriously impaired that diagnostic difficulties arise.

7.07.14.3 Effortful vs. Automatic Processing and Memory

Weingartner, Cohen, Bunney, Ebert, and Kaye (1982) proposed that tasks that are normally accomplished automatically and require little cognitive capacity are particularly useful in distinguishing progressive dementia patients from depressed patients. Weingartner (1984, 1986) reported that depressed patients experienced difficulty on memory tasks that require elaborate or effortful organization and processing of material to be remembered, but that they do not show problems in memory tasks on which "automatic" memory encoding is presumed to be involved. In contrast, it is postulated that demented patients have difficulty with both automatic and effortful processing. In line with this proposal, Weingartner, Cohen, Murphy, Martello, and Gerdt (1981) reported that depressed patients demonstrate impairments on delayed free recall (effortful) with no impairment on recognition memory (relatively automatic), whereas demented patients perform poorly in both conditions. However, Lachner, Satzger and Engel (1994) have provided evidence which challenges the Weingartner model. They compared three groups (demented, depressed, and healthy subjects) who were in their 70s on five verbal recall and two recognition memory tasks. They found that both delayed recall and recognition after long delay were the measures which best discriminated between the demented and the depressed patient groups. It is interesting to note that the recognition measure was more discriminating than free recall, selective reminding, or serial learning, all presumably more "effortful" than a recognition task.

Lachner et al. (1994) point out that the labeling of some cognitive tasks as effortful vs. automatic is often ambiguous, and propose sensibly that "In future research, the tasks cognitive capacity requirements should be examined empirically to avoid inconsistencies by subjectively estimating capacity demands" (Lachner et al., p. 10). A particular advantage of the Lachner et al. study is their comparison of severely depressed psychiatric inpatients with a demented patient sample where 60% were still capable of independently running their household, as it is this type of comparison of mildly demented and severely depressed patients that is of particular clinical interest with regard to the problem of differential diagnosis.

7.07.14.4 Is Poor Memory Test Performance in Elderly Depressed Patients an Artifact of Poor Motivation?

In commenting on the relatively poor performance of many depressed patients with standard neuropsychological measures, many writers have attributed this impairment to the nonspecific effects of motivational deficits. Few studies have tested this hypothesis experimentally. One notable exception is the study of Richards and Ruff (1989), who randomly assigned two groups of subjects, depressed and nondepressed, to either motivation or nonmotivation conditions (motivation involved encouragement, a monitory incentive and performance feedback). Richards and Ruff (1989) reported that motivation (as measured by improvement on a simple card-sorting task) was enhanced for the subjects in the motivation condition. Crucially, motivation did not significantly affect neuropsychological test performance and the authors concluded that although depressed patients may be less motivated, this reduced motivation does not fully account for the observed cognitive impairments in depression.

7.07.14.5 Future Directions

The literature indicates that demented subjects are impaired at any task which requires the conscious recollection of newly learned material. However, much work has been carried out attempting to evaluate whether specific aspects of memory functioning are disproportionately impaired in early DAT. Divided attention tasks are particularly sensitive to DAT, that is, tasks where subjects are required to divide their attention during learning (e.g., performing a digit span test while keeping a stylus on a target on a pursuit rotor task). This would appear to be a particular deficit in "working memory," that is, the ability to hold information in a short-term store or carrying out the other processing operations (Baddeley & Wilson, 1986). Additionally, prospective memory tasks appear to be particularly impaired in early DAT. Huppert and Beardsall (1993) administered a variety of memory tasks and found that even patients with minimal dementia were disproportionately impaired on prospective memory tasks (e.g., remembering to ask for a belonging back at the end of the Rivermead Behavioural Memory Test). In the search for particularly dementia-sensitive screening tasks, a combination of divided attention and prospective memory tasks may well be worthy of further investigation.

7.07.14.6 Recommendations and Conclusions for the Differential Diagnosis of Dementia vs. Depression

Sweet, Newman, and Bell (1992) made several useful recommendations to facilitate the differential diagnosis of depressed patients from demented patients. First, they recommend administering neuropsychological measures that assess domains which the literature indicate are particularly discriminating, notably psychomotor retardation, recognition vs. recall, and easy vs. difficult paired associated learning. Second, they recommend utilization of measures of affective status (e.g., the Beck Depression Inventory or the Geriatric Depression Rating Scale). They also propose seeking collateral information through interviewing the spouse or significant other. Third, when evidence of depression is found, customary cognitive "cut-off" scores may be misleading and they propose using more stringent criteria (e.g., an additional standard deviation away from normal performance). Fourth, with depressed patients, they caution against diagnosing brain dysfunction based only on findings such as mild attentional or memory problems. For example, as psychomotor retardation in depression can easily result in a markedly impaired performance on the Stroop task (Trenerry, Crosson, DeBoe, & Leber, 1989), many depressed patients fall within the recommended cut-off score used to indicate "brain damage."

Sweet et al. (1992) also urge against differentiating depression from brain damage after only a single session with the patient and propose seeing the patient on more than one day and, if in doubt, readministering at least a portion of the measures on a second testing session (this would be in accord with the protocol of the Kendrick Cognitive Tests for the Elderly). Sweet et al. conclude "In the event that a diagnostic conclusion regarding brain dysfunction in the depressed patient appears unclear, despite all your best efforts, recommend a re-evaluation of the depressed patient after a period of pharmacological or psychotherapeutic intervention" (p. 41). This is a recommendation that is fully endorsed. Neuropsychological measures should be used in conjunction with information derived from a variety of sources (e.g., relatives and clinical history). Developments in neuroimaging precision suggest that an *in vivo* maker for DAT may become available (Jobst et al., 1994).

Given the potentially catastrophic outcome of misclassifying a treatable, depressed patient as having DAT, it has to be concluded that the available evidence indicates that no single neuropsychological measure can be recommended for use in the individual case to confidently separate early DAT from major depression in an elderly population. However, neuropsychological measures may offer useful information that can be integrated with data gathered from a variety of sources. In particular, accurate objective quantification of change over time is one of the key areas where detailed assessment of cognitive functioning can provide important diagnostic information. Furthermore, changes in neuropsychological status can offer a useful measure of treatment response in both dementia and depression.

7.07.15 REFERENCES

Albert, M. L. (1973). A simple test of visual neglect. *Neurology*, *23*, 658–664.

Albert, M. S., Heller, H. S., & Milberg, W. (1988). Changes in naming ability with age. *Psychology and Aging*, *3*, 173–178.

Anthony, J. C., Le Resche, L., Niaz, U., Von Korff, M. R., & Folstein, M. F. (1982). Limits of the Mini-Mental-State as a screening test for dementia and delirium among hospital patients. *Psychological Medicine*, *12*, 397–408.

Atkinson, L. (1991). Some tables for statistically based interpretation of WAIS-R factor scores. *Psychological Assessment*, *3*, 288–291.

Atkinson, L., Cyr, J. J., Doxey, N. C. S., & Vigna, C. M. (1989). Generalizability of WAIS-R factor structure within and between populations. *Journal of Clinical Psychology*, *45*, 124–129.

Baddeley, A., Emslie, H., & Nimmo-Smith, I. (1993). The Spot-the-Word Test: A robust estimate of verbal intelligence based on lexical decision. *British Journal of Clinical Psychology*, *32*, 55–65.

Baddeley, A., Emslie, H., & Nimmo-Smith, I. (1994). *Doors* *and people: A test of visual and verbal recall and recognition*. Bury St. Edmunds, UK: Thames Valley Test Company.

Baddeley, A., & Wilson, B. (1986). Dementia and working memory. *Quarterly Journal of Experimental Psychology*, *38*(A), 603–618.

Balota, D. A., & Duchek, J. M. (1988). Age related differences in lexical access, spreading activation, and simple pronunciation. *Psychology and Aging*, *3*, 84–93.

Barker, M. G., & Lawson, J. S. (1968). Nominal aphasia in dementia. *British Journal of Psychiatry*, *114*, 1351–1356.

Barlow, D. H., & Hersen, M. (1984). *Single case experimental designs: strategies for studying behavior change* (2nd ed.). New York: Pergamon.

Barona, A., Reynolds, C. R., & Chastain, R. (1984). A demographically based index of premorbid intelligence for the WAIS-R. *Journal of Consulting and Clinical Psychology*, *52*, 885–887.

Beardsall, L., & Huppert, F. A. (1994). Improvement in NART word reading in demented and normal older persons using the Cambridge Contextual Reading Test. *Journal of Clinical and Experimental Neuropsychology*, *16*, 232–242.

Benton, A. L. (1968). Differential behavioral effects in frontal lobe disease. *Neuropsychologia*, *6*, 53–60.

Benton, A. L. (1969). Development of a multilingual aphasia battery: Progress and problems. *Journal of the Neurological Sciences*, *9*, 39–48.

Benton, A. L., Eslinger, P. J., & Damasio, A. R. (1981). Normative observations of neuropsychological test performances in old age. *Clinical Neuropsychology, 3*, 33–42.

Benton, A. L., Hamser, K. D. S., Varney, N. R., & Spreen, O. (1983). *Contributions to neuropsychological assessment*. New York: Oxford University Press.

Benton, A. L., Hannay, H. J., & Varney, N. R. (1975). Visual perception of line direction in patients with unilateral brain disease. *Neurology*, *25*, 907–910.

Benton, A. L., & Van Allen, M. W. (1968). Impairment in facial recognition in patients with cerebral disease. *Cortex, 4*, 344–358.

Binks, M. (1989). Changes in mental functioning associated with normal ageing. In R. Tallis (Ed.), *The clinical neurology of old age* (pp. 27–39). Chichester, UK: Wiley.

Bisiach, E., & Luzzatti, C. (1978). Unilateral neglect of representational space. *Cortex, 14*, 129–133.

Blessed, G., Black, S. E., Butler, T., & Kay, D. W. K. (1991). The diagnosis of dementia in the elderly a comparison of CAMCOG (the cognitive section of CAMDEX), the AGECAT program, DSM-III, The Mini-Mental State Examination and some short rating Scales. *British Journal of Psychiatry, 159*, 193–198.

Bondi, M. W., Kaszniak, A. W., Bayles, K. A., & Vance, K. T. (1993). Contributions of frontal system dysfunction to memory and perceptual abilities in Parkinson's disease. *Neuropsychology*, *7*, 89–102.

Borkowski, J. G., Benton, A. L., & Spreen, O. (1967). Word fluency and brain damage. *Neuropsychologia, 5*, 135–140.

Bornstein, R. A. (1986). Contribution of various neuropsychological measures to detection of frontal lobe impairment. *International Journal of Clinical Neuropsychology, 8*, 18–22.

Borod, J. C., Goodglass, H., & Kaplan, E. (1980). Normative data on the Boston Diagnostic Aphasia Examination, Parietal Lobe Battery, and the Boston Naming Test. *Journal of Clinical Neuropsychology, 2*, 209–215.

Brinkman, S. D., & Braun, P. (1984). Classification of dementia patients by a WAIS profile related to central cholinergic deficiencies. *Journal of Clinical Neuropsychology, 6*, 393–400.

Britton, P. G., & Savage, R. D. (1966). A short form of the WAIS for use with the aged. *British Journal of Psychiatry, 112*, 417–418.

Brooks, D. N. (1989). Closed head trauma: assessing the common cognitive problems. In M. D. Lezak (Ed.), *Assessment of the behavioral consequences of head trauma. Frontiers of Clinical Neuroscience* (Vol. 7, pp. 61–85). New York: Liss.

Bryan, J., Luszcz, M. A., & Crawford, J. R. (1997). Verbal knowledge and speed of information processing as mediators of age differences in verbal fluency performance among older adults. *Psychology and Aging, 12,* 473–478.

Burgess, A. (1991). Profile analysis of the Wechsler Intelligence Scales: a new index of subtest scatter. *British Journal of Clinical Psychology, 30*, 257–263.

Burke, D. M., Mackay, D. G., Worthley, J. A., & Wade, E. (1991). On the tip of the tongue: what causes word finding failures in young and older adults? *Journal of Memory and Language, 30*, 542–579.

Burke, D. M., White, H., & Diaz, D. L. (1987). Semantic priming in young and older adults: evidence for age-constancy in automatic and attentional processes. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 79–88.

Butters, B., Salmon, D. P., Cullum, C. M., Cairns, P., Troster, A. I., & Jacobs, D. (1988). Differentiation of amnesic and demented patients with the Wechsler Memory Scale-Revised. *Clinical Neuropsychologist, 2*, 133–148.

Caine, E. D. (1992). *Nomenclature and diagnosis of cognitive disorders: a US perspective*. Paper presented at the Abstract from Age-related Cognitive Disorders Conference, Nice, France.

Canavan, A. G. M., Dunn, G., & McMillan, T. M. (1986). Principal components of the WAIS-R. *British Journal of Clinical Psychology, 25*, 81–35.

Chapman, L. J., & Chapman, J. P. (1973). Problems in the measurement of cognitive deficit. *Psychological Bulletin, 79*, 380–385.

Christensen, H., & MacKinnon, A. (1992). Wechsler intelligence scale profiles in Alzheimer type dementia and healthy ageing. *International Journal of Geriatric Psychiatry, 7*, 241–246.

Cockburn, J., & Smith, P. T. (1989). *The Rivermead Behavioural Memory Test (Suppl. 3: Elderly people*). Bury St. Edmunds, England: Thames Valley Test Company.

Coen, R. F., Swanwick, G. R. J., Maguire, C., Kirby, M., Lawlor, B. A., Walsh, J. B., & Coakley, D. (1996). Memory impairment in Alzheimer's disease: replication and extension of the delayed word recall (DWR) test. *Irish Journal of Psychological Medicine, 13*, 55–58.

Corkin, S., Growdon, J., Nissen, M. J., Huff, F. J., Freed, D. M., & Sagar, H. J. (1984). Recent advances in the neuropsychological study of Alzheimer's disease. In R. J. Wurtman, S. Corkin, & J. H. Growdon, (Eds.), *Advances in basic research and therapies* (pp. 75–94). New York: Centre for Brain Sciences and Metabolism Trust.

Crawford, J. R. (1992). Current and premorbid intelligence measures in neuropsychological assessment. In J. R. Crawford, D. M. Parker & W. McKinlay, (Eds.), *A handbook of neuropsychological assessment* (pp. 21–49). Hove, UK: Erlbaum.

Crawford, J. R. (1997). WAIS-R short-forms: Assessing the statistical significance of subtest differences. *British Journal of Clinical Psychology, 36,* 601–608.

Crawford, J. R., & Allan, K. M. (1997). Estimating premorbid IQ with demographic variables: regression equations derived from a U.K. sample. *The Clinical Neuropsychologist, 11,* 192–197.

Crawford, J. R., Allan, K. M., & Jack, A. M. (1992a). Short-forms of the UK WAIS-R: regression equations and their predictive accuracy in a general population sample. *British Journal of Clinical Psychology, 31*, 191–202.

Crawford, J. R., Allan, K. M., McGeorge, P., & Kelly, S. M. (1997). Base rate data on the abnormality of subtest scatter for WAIS-R short-forms. *British Journal of Clinical Psychology, 36,* 433–444.

Crawford, J. R., Allan, K. M., Stephen, D. W., Parker, D. M., & Besson, J. A. O. (1989a). The Wechsler Adult Intelligence Scale-Revised (WAIS-R): Factor structure in a UK sample. *Personality and Individual Differences, 10*, 1209–1212.

Crawford, J. R., Besson, J. A. O., & Ebmeier, K. P. (1990). Dementia: The role of neurotransmitter deficits. In T. G. Dinan (Ed.), *Principles and practice of biological psychiatry* (pp. 105–133). London: Clinical Neuroscience Publishers.

Crawford, J. R., Besson, J. A. O., & Parker, D. M., (1988). Estimation of permorbid intelligence in organic conditions. *British Journal of Psychiatry, 153*, 178–181.

Crawford, J. R., Johnson, D. A., Mychalkiw, B., & Moore, J. W. (in press-a). WAIS-R performance following closed head injury: A comparison of the clinical utility of summary IQs, factor scores and subtest scatter indices. *The Clinical Neuropsychologist, 11*.

Crawford, J. R., Moore, J. W., & Cameron, I. M. (1992c). Verbal fluency: a NART-based equation for the estimation of premorbid performance. *British Journal of Clinical Psychology, 31*, 327–329.

Crawford, J. R., Obonsawin, M. C., & Allan, K. M. (in press-b). PASAT and components of WAIS-R performance: Convergent and discriminant validity. *Neuropsychological Rehabilitation.*

Crawford, J. R., Obonsawin, M. C., & Bremner, M. (1993). Frontal lobe impairment in schizophrenia: Relationship to intellectual functioning. *Psychological Medicine, 23*, 787–790.

Crawford, J. R., Parker, D. M., & McKinlay, W., & (Eds.). (1992b). *A handbook of neuropsychological assessment*. Hove, UK: Erlbaum.

Crawford, J. R., Sommerville, J., & Robertson, I. H. (1997). Assessing the reliability and abnormality of subtest differences on the Test of Everyday Attention. *British Journal of Clinical Psychology, 36,* 609–617.

Crawford, J. R., Stewart, L. E., Cochrane, R., Foulds, J., Besson, J. A. O., & Parker, D. M. (1989b). Estimating premorbid IQ from demographic variables: a regression equation derived from a UK sample. *British Journal of Clinical Psychology, 28*, 275–278.

Crawford, J. R., Stewart, L. E., & Moore, J. W. (1989c). Demonstration of savings on the AVLT and development of a parallel form. *Journal of Clinical and Experimental Neuropsychology, 11*, 975–981.

Crawford, J. R., Wright, R., & Bate, A. (1995). Verbal, figural and ideational fluency in CHI. *Journal of the International Neuropsychological Society, 1*, 321.

Crocker, L. M., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rhinehart & Winston.

Crook, T., Bartus, R. T., Ferris, S. H., Whitehouse, P., Cohen, G. D., & Gershon, S. (1986). Age-associated memory impairment: proposed diagnostic criteria and measures of clinical change: Report of a National Institute of Mental Health Work Group. *Developmental Neuropsychology, 2*, 261–276.

Crowe, S. F. (1992). Dissociation of two frontal lobe syndromes by a test of verbal fluency. *Journal of Clinical and Experimental Neuropsychology, 14*, 327–339.

Daigneault, S., & Braun, C. M. J. (1993). Working memory and the self-ordered pointing task: Further evidence of early prefrontal decline in normal aging. *Journal of Clinical and Experimental Neuropsychology, 15*, 881–895.

Daigneault, S., Braun, C. M. J., & Whitaker, H. A. (1992). Early effects of normal aging on perserverative and non-perseverative prefrontal measures. *Developmental Neuropsychology, 8*, 99–114.

Davis, H. P., & Bernstein, P. A. (1992). Age-related changes in explicit and implicit memory. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (2nd ed., pp. 249–261). New York: Guilford Press.

De Renzi, E., & Faglioni, P. (1978). Normative data and screening power of a shortened version of the Token Test. *Cortex, 14*, 41–49.

De Renzi, E., & Vignolo, L. A. (1962). The Token Test: A sensitive test to detect disturbances in aphasics. *Brain, 85*, 665–678.

D'Elia, L. F., Boone, K. B., & Mitrushima, A. M. (1995). Handbook of normative data for neuropsychological assessment. New York: Oxford University Press.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Learning Test—Adult Version: Research Edition Manual Version 1*. San Antonio, CA: The Psychological Corporation.

Della Sala, S. (1988). Cognitive deficits of Parkinsonians and Occam's Razor. *Europa Medicophysica, 24*, 1–22.

Della Sala, S., Laiacona, M., Trivelli, C., & Spinnler, H. (1995). Poppelreuter-Ghent overlapping figures test. Its sensitivity to age, and its clinical use. *Archives of Clinical Neuropsychology, 10*(6), 511–534.

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Development Review, 12*, 45–75.

desRosiers, G. (1992). Primary or depressive dementia: Psychometric assessment. *Clinical Psychology Review, 12*, 307–343.

Dick, J. P. R., Guiloff, R. J., Stewart, A., Blackstock, J., Bielawska, C., Paul, E. A., & Marsden, C. D. (1984). Mini Mental State Examination in neurological patients. *Journal of Neurology, Neurosurgery, and Psychiatry, 47*, 496–499.

Drachman, D. A., & Leavitt, J. (1974). Human memory and the cholinergic system: A relationship to aging? *Archives of Neurology, 30*, 113–121.

Ellis, H. D. (1992). Assessment of deficits in facial processing. In J. R. Crawford, D. M. Parker, & W. W. McKinlay (Eds.), *A handbook of neuropsychological assessment* (pp. 141–150). Hove, UK: Lawrence Erlbaum.

Elwood, R. W. (1991). The Wechsler Memory Scale-Revised: Psychometric characteristics and clinical application. *Neuropsychology Review, 2*, 179–201.

Eslinger, P. J., & Benton, A. L. (1983). Visuoperceptual performance in aging and dementia. Clinical and theoretical implications. *Journal of Clinical Neuropsychology, 5*, 213–220.

Farver, P. F., & Farver, T. B. (1982). Performance of normal older adults on tests designed to measure parietal lobe functions. *American Journal of Occupational Therapy, 36*, 444–449.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental-State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatry Research, 12*, 189–198.

Franzen, M. D. (1989). *Reliability and validity in neuropsychological assessment*. New York: Plenum.

Freedman, L., & Dexter, L. E. (1991). Visuospatial ability in cortical dementia. *Journal of Clinical and Experimental Neuropsychology, 13*(5), 677–690.

Fuld, P. A., (1984). Test profile of cholinergic dysfunction and of Alzheimer-type dementia. *Journal of Clinical Neuropsychology, 6*, 380–392.

Fuster, J. M. (1989). *The prefrontal cortex* (2nd ed.). New York: Raven Press.

Gallagher, A. J. (1979). Temporal reliability of aphasic performance on the Token Test. *Brain and Language, 7*, 34–41.

Gardner, R., Oliver-Munoz, S., Fisher, L., & Empting, L.

(1981). Mattis Dementia Rating Scale: Internal reliability study using a diffusely impaired population. *Journal of Clinical Neuropsychology, 3*, 271–275.

Gauthier, L., Dehaut, F., & Joanette, Y. (1989). The Bells Test: A quantitative and qualitative test for visual neglect. *International Journal of Clinical Neuropsychology, 11*, 49–54.

Geffen, G., Moar, K. J., O'Hanlon, A. P., Clark, C. R., & Geffen, L. B. (1990). Performance measures of 16- to 86-year-old males and females on the Auditory Verbal Learning Test. *Clinical Neuropsychologist, 4*, 45–63.

Gfeller, J. D., & Rankin, E. J. (1991). The WAIS-R profile as a cognitive marker of Alzheimer's Disease: A misguided venture? *Journal of Clinical and Experimental Neuropsychology, 13*, 629–636.

Giles, E., Patterson, K., & Hodges, J. R. (1996). Performance on the Boston Cookie Theft Picture Description Task in patients with early dementia of the Alzheimer's type: Missing information. *Aphasiology, 10*, 395–408.

Glosser, G., & Deser, T. (1992). A comparison of changes in macrolinguistic and microlinguistic aspects of discourse production in normal aging. *Journals of Gerontology, 47*, 266–272.

Goodglass, H. (1980). Naming disorders in aphasia and ageing. In L. K. Obler & M. L. Albert (Eds.), *Language and communication in the elderly* (pp. 37–45). Lexington, MA: Heath.

Goodglass, H., & Kaplan, E. (1983). *Boston Diagnostic Aphasia Examination (BDAE)*. Philadelphia: Lea & Febiger.

Goodglass, H., & Kaplan, E. (1987). *The assessment of aphasia and related disorders* (2nd ed.). Philadelphia: Lea & Febiger.

Gottschaldt, K. (1928). Uber den Einfluss der Erfahrung auf die Wahrnehmung von Figuren. *Psychologische Forschung, 18*, 18–317.

Grant, D. A., & Berg, E. A. (1948). A behavioural analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology, 38*, 404–411.

Greene, J. D. W., Baddeley, A. D., & Hodges, J. R. (1996). Analysis of the episodic memory deficit in early Alzheimer's disease: Evidence from the Doors and People Test. *Neuropsychologia, 34*, 537–551.

Grober, E., Buschke, H., Kawas, C., & Fuld, P. (1985). Impaired ranking of semantic attributes in dementia. *Brain and Language, 26*, 276–286.

Gronwall, D. (1977). Paced auditory serial-addition task: a measure of recovery from concussion. *Perceptual and Motor Skills, 44*, 367–373.

Halligan, P. W., & Marshall, J. C. (1988). How long is a piece of string? A study of line bisection in a case of visual neglect. *Cortex, 24*, 321–328.

Halligan, P. W., & Marshall, J. C. (1989). Line bisection in visuo-spatial neglect: Disproof of a conjecture. *Cortex, 25*, 517–521.

Halligan, P. W., & Marshall, J. C. (1991). Left neglect for near but not far space in man. *Nature, 350*, 498–500.

Halligan, P. W., & Robertson, I. H. (1992). The assessment of unilateral neglect. In J. R. Crawford, D. M. Parker, & W. W. McKinlay (Eds.), *A handbook of neuropsychological assessment* (pp. 151–175). Hove, UK: Lawrence Erlbaum.

Hart, S., Smith, C. M., & Swash, M. (1988). Word fluency in patients with early dementia of Alzheimer type. *British Journal of Clinical Psychology, 27*, 115–124.

Heaton, R. K. (1981). *Wisconsin Card Sorting Test manual*. Odessa, FL: Psychological Assessment Resources.

Hodges, J. R., Patterson, K., Graham, N., & Dawson, K. (1996). Naming and knowing in dementia of Alzheimer's type. *Brain and Language, 54*, 302–325.

Hodges, J. R., Salmon, D. P., & Butters, N. (1990). Differential impairment of semantic and episodic memory in Alzheimer's and Huntington's diseases: A controlled prospective study. *Journal of Neurology, Neurosurgery and Psychiatry, 53*, 1089–1095.

Hooper, H. E. (1983). *Hooper Visual Organization Test (VOT)* Los Angeles: Western Psychological Services.

Huff, F. J., Boller, F., Lucchelli, F., Querriera, R., Beyer, J., & Belle, S. (1987). The neurologic examination in patients with probably Alzheimer's disease. *Archives of Neurology, 44*, 929–932.

Huff, F. J., Collins, C., Corkin, S., & Rosen, T. J. (1986). Equivalent forms of the Boston Naming Test. *Journal of Clinical and Experimental Psychology, 8*, 556–562.

Huppert, F. A. (1994). Memory function in dementia and normal aging: Dimension or dichotomy? In F. A. Huppert, C. Brayne, & D. W. O'Connor (Eds.), *Dementia and normal aging* (pp. 291–330). Cambridge: Cambridge University Press.

Huppert, F. A., & Beardsall, L. (1993). Prospective memory impairment as an early indicator of dementia. *Journal of Clinical and Experimental Neuropsychology, 15*, 805–821.

Huppert, F. A., Jorm, A. F., Brayne, C., Girling, D. M., Barkley, C., Beardsall, L., & Paykel, E. S. (submitted). Psychometric properties of the CAMCOG and its efficacy in the diagnosis of dementia.

Huppert, F. A., & Kopelman, M. D. (1989). Rates of forgetting in normal ageing: A comparison with dementia. *Neuropsychologia, 27*, 849–860.

Huttenlocher, J., Haight, W., & Bryk, A. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*, 236–248.

Ishiai, S., Okiyama, R., Koyama, Y., & Seki, K. (1996). Unilateral spatial neglect in Alzheimer's disease. A line bisection study. *Acta Neurologica Scandinavica, 93*, 219–224.

Ivnik, R. J., Malec, J. F., Smith, G. E., Tangalos, E. G., & Petersen, R. C. (1996). Neuropsychological tests' norms above age 55: COWAT, BNT, MAE Token, WRAT-R reading, AMNART, STROOP, TMT, and JLO. *The Clinical Neuropsychologist, 10*, 262–278.

Ivnik, R. J., Malec, J. F., Smith, G. E., Tangalos, E. G., Peterson, R. C., Kokmen, E., & Kurland, L. T. (1992). Mayo's older Americans normative studies: WAIS-R norms for ages 56 to 97. *The Clinical Neuropsychologist, 6*, 1–30.

Jobst, K. A., Smith, A. D., Szatmari, M., Esiri, M. M., Jaskowski, A., Hindley, N., McDonald, B., & Molyneux, A. J. (1994). Rapidly progressing atrophy of medial temporal lobe in Alzheimer's disease. *Lancet, 343*, 829–830.

Jones-Gotman, M., & Milner, B. (1977). Design fluency: The invention of nonsense drawings after focal frontal lesions. *Neuropsychologia, 15*, 653–674.

Kaplan, E. F., Goodglass, H., & Weintraub, S. (1978). *The Boston Naming Test*. (Experimental ed.). Boston: Kaplan & Goodglass.

Kaplan, E. F., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test* (2nd ed.). Philadelphia: Lea & Febiger.

Kaufman, A. S. (1990). *Assessing adolescent and adult intelligence*. Boston: Allyn & Bacon.

Kemper, S. (1992). Language and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 213–272). Hillsdale, NJ: Erlbaum.

Kempler, D. (1988). Lexical and pantomime abilities in Alzheimer's disease. *Aphasiology, 2*, 147–159.

Kempler, D., Andersen, E. S., & Henderson, V. W. (1995). Linguistic and attentional contributions to anomia in Alzheimer's disease. *Neuropsychiatry, Neuropsychology, and Behavioural Neurology, 8*(1), 33–37.

Kendrick, D. C., Gibson, A. J., & Moyes, I. C. A. (1979). The revised Kendrick Battery: Clinical studies. *British*

Journal of Social and Clinical Psychology, 18, 329–340.

Kirk, A., & Kertesz, A. (1991). On drawing impairment in a Alzheimer's disease. *Archives of Neurology, 48*, 73–77.

Kittner, S. J., White, L. R., Farmer, M. E., Wolz, M, Kaplan, E., Moes, E., Brody, J. A., & Feinleib, M. (1986). Methodological issues in screening for dementia: The problem of education adjustment. *Journal of Chronic Diseases, 39*, 163–170.

Klesges, R. C., Wilkening, G. N., & Golden, C. J. (1981). Premorbid indices of intelligence: A review. *Clinical Neuropsychology, 3*, 32–39.

Knight, R. G., & Godfrey, H. P. D. (1984). Assessing the significance of differences between subtests on the Wechsler Adult Intelligence Scale-Revised. *Journal of Clinical Psychology, 40*, 808–810.

Knight, K. R., & Moroney, B. (1985). An investigation of the validity of the Kendrick Battery for the detection of dementia in the elderly. *International Journal of Clinical Psychology, 7*, 147–151.

Knight, R. G., & Shelton, E. J. (1983). Tables for evaluating predicted retest changes in Wechsler Adult Intelligence Scale scores. *British Journal of Clinical Psychology, 22*, 77–81.

Knopman, D. S., & Ryberg, S. (1989). A verbal memory test with high predictive accuracy for dementia of the Alzheimer type. *Archives of Neurology, 46*, 141–145.

Kopelman, M. D. (1985). Rates of forgetting in Alzheimer-type dementia and Korsakoff's syndrome. *Neuropsychologia, 23*, 623–638.

Kozora, E., & Cullum, C. M. (1995). Generative naming in normal aging: Total output and qualitative changes using phonemic and semantic constraints. *The Clinical Neuropsychologist, 9*, 313–320.

Kukull, W. A., Larson, E. B., Teri, L., Bowen, J., McCormick, W., & Pfanschmidt, M. L. (1994). The Mini-Mental State Examination score and the clinical diagnosis of dementia. *Journal of Clinical Epidemiology, 47*, 1061–1067.

Lachner, G., Satzger, W., & Engel, R. R. (1994). Verbal memory tests in the differential diagnosis of depression and dementia: discriminative power of seven test variations. *Archives of Clinical Neuropsychology, 9*, 1–13.

Leckliter, I. N., Matarazzo, J. D., & Silverstein, A. B. (1986). A literature review of factor analytic studies of the WAIS-R. *Journal of Clinical Psychology, 42*, 332–342.

Lee, G. P., Strauss, E., Loring, D. W., McCloskey, L., Haworth, J. M., & Lehman, R. A. W. (1997). Sensitivity of figural fluency on the Five Point Test to focal neurological dysfunction. *The Clinical Neuropsychologist, 11*, 59–68.

Lees, A. J., & Smith, E. (1983). Cognitive deficits in the early stages of Parkinson's disease. *Brain, 106*, 257–270.

Levin, B. E., Llabre, M. M., Reisman, S., Weiner, W. J., Sanchez Ramos, J., Singer, C., & Brown, M. C. (1991). Visuospatial impairment in Parkinson's disease. *Neurology, 41*, 365–369.

Lezak, M. D. (1983). *Neuropsychological assessment* (2nd ed.). New York: Oxford University Press.

Lezak, M. D. (1988a). Neuropsychological tests and assessment techniques. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 1, pp. 47–68). Amsterdam: Elsevier.

Lezak, M. D. (1988b). IQ: R.I.P. *Journal of Clinical & Experimental Neuropsychology, 10*, 351–361.

Lezak, M. D. (1991). Identifying neuropsychological deficits. In R. G. Lister & H. J. Weingartner (Eds.), *Perspectives on cognitive neuroscience* (pp. 357–367). New York: Oxford University Press.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Light, L. L. (1992). The organization of memory in old age. In F. I. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 111–166). Hillsdale, NJ: Erlbaum.

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging, 3*, 339–355.

Logsdon, R. G., Teri, L., Williams, D. E., Vitiello, M. V., & Prinz, P. N. (1989). The WAIS-R profile: a diagnostic tool for Alzheimer's disease? *Journal of Clinical and Experimental Neuropsychology, 11*, 892–898.

Magni, E., Binetti, G., Bianchetti, A., Rozzini, R., & Trabucchi, M. (1996). Mini Mental State Examination. A normative study in Italian elderly population. *European Journal of Neurology, 3*(3), 198–202.

Malec, J. F., Ivnik, B. J., Smith, G. E., Tangalos, G. G. (1992). Mayo's older American normative studies: Utility of corrections for age and education for the WAIS-R. *The Clinical Neurospychologist, 6* (Suppl.), 31–47.

Martin, A., Wiggs, C. L., Lalonde, F., & Mack, C. (1994). Word retrieval to letter and semantic cues: a double dissociation in normal subjects using interference tasks. *Neuropsychologia, 32*, 1487–1494.

Massman, P. J., & Bigler, E. D. (1993). A quantitative review of the diagnostic utililty of the WAIS-R Fuld profile. *Archives of Clinical Neuropsychology, 8*, 417–428.

Matarazzo, J. D., Daniel, M. H., Prifitera, A., & Herman, D. O. (1988). Inter-subtest scatter in the WAIS-R standardization sample. *Journal of Clinical Psychology, 44*, 940–950.

Matarazzo, J.D., & Herman, D. O. (1985). Clinical uses of the WAIS-R: Base rates of differences between VIQ and PIQ in the WAIS-R standardisation sample. In B. Wolman (Ed.), *Handbook of intelligence: Theories, measurement and applications* (pp. 899–932). New York: Wiley.

Mattis, S. (1976). Mental status examination for organic mental syndrome in the elderly patient. In L. Bellack & T. B. Karasu (Eds.), *Geriatric psychiatry* (pp. 77–121). New York: Grune & Stratton.

Mattis, S. (1988). *Dementia Rating Scale*. Odessa, FL: Psychological Assessment Resources.

Maxwell, A. E. (1960). Obtaining factor scores on the Wechsler Adult Intelligence Scale. *Journal of Mental Science, 106*, 1060–1062.

Mayes, A., & Warburg, R. (1990). Memory assessment in clinical practice and research. In J. R. Crawford, D. M. Parker, & W. W. McKinlay (Eds.), *A handbook of neuropsychological assessment* (pp. 73–101). Hove, UK: Erlbaum.

Maylor, E. A. (1990). Recognizing and naming faces. Aging, memory retrieval, and the tip of the tongue state. *Journal of Gerontology, 45*(6), 215–226.

McCarthy, R. A., & Warrington, E. K. (1986). Visual associative agnosia: A clinico-anatomical study of a single case. *Journal of Neurology, Neurosurgery and Psychiatry, 49*, 1233–1240.

McKinlay, W. W. (1992). Assessment of the head-injured for compensation. In J. R. Crawford, D. M. Parker, & W. W. McKinlay (Eds.), *A handbook of neuropsychological assessment* (pp. 381–392). Hove, UK: Erlbaum.

McLean, J. E., Reynolds, C. R., & Kaufman, A. S. (1990). WAIS-R subtest scatter using the profile variability index. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2*, 289–292.

McSweeny, A. J., Naugle, R. I., Chelune, G. J., & Lüders, H. (1993). "T Scores for Change": an illustration of a regression approach to depicting change in clinical neuropsychology. *The Clinical Neuropsychologist, 7*, 300–312.

Miceli, G., Caltagirone, C., Gainotti, G., Masullo, C., & Silveri, M. C. (1981). Neuropsychological correlates of localized cerebral lesions in nonaphasic brain-damaged patients. *Journal of Clinical Neuropsychology, 3*, 53–63.

Miller, E. (1984). Verbal fluency as a function of a measure of verbal intelligence and in relation to different types of cerebral pathology. *British Journal of Clinical Psychology, 23*, 53–57.

Miller, E. (1992). Some basic principles of neuropsychological assessment. In J. R. Crawford, D. M. Parker, & W. W. McKinlay (Eds.), *A handbook of neuropsychological assessment* (pp. 7–20). Hove, UK: Lawrence Erlbaum.

Milner, B. (1964). Some effects of frontal lobectomy in man. In J. M. Warren & K. Alkert (Eds.), *The frontal granular cortex and behaviour* (pp. 313–331). New York: McGraw-Hill.

Mitrushina, M., & Satz, P. (1995). Repeated testing of normal elderly with the Boston Naming Test. *Aging: Clinical and Experimental Research, 7*(2), 123–127.

Mittenberg, W., Seidenberg, M., O'Leary, D. S., & DiGiulio, D. V. (1989). Changes in cerebral functioning associated with normal aging. *Journal of Clinical and Experimental Neuropsychology, 11*, 918–932.

Montgomery, K., & Costa, L. (1973). *Neuropsychological test performance of a normal elderly sample*. Paper presented at the Annual Meeting of the International Neuropsychological Society, Mexico City, Mexico.

Montgomery, K. M. (1982). *A normative study of neuropsychological test performance of a normal elderly sample*. Unpublished Master's thesis, University of Victoria, British Columbia, Canada.

Monti, L. A., Reminger, S. L., Gabrieli, J. D. E., Rinaldi, J. A., Wilson, R. S., & Fleischman, D. A. (1996). Differential effects of aging and Alzheimer's disease on conceptual implicit and explicit memory. *Neuropsychology, 10*, 101–112.

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., Mellits, E. D., & Clark, C. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD) 1. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology, 39*, 1159–1165.

Mountain, M. A., & Snow, W. G. (1993). Wisconsin Card Sorting Test as a measure of frontal pathology: A review. *The Clinical Neuropsychologist, 7*, 108–118.

Nebes, R. D., Brady, C. G., & Huff, F. J. (1989). Automatic and attentional mechanisms of semantic priming in Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 11*, 219–230.

Neils, J., Brennan, M. M., Cole, M., Boller, F., & Gerdeman, B. (1988). The use of phonemic cueing with Alzheimer's disease patients. *Neuropsychologia, 26*, 351–354.

Nelson, H. E. (1990). Review of the Recognition Memory Test. In J. R. Beech & L. Harding (Eds.), *Assessment of the elderly* (pp. 126–127). Windsor, UK: NFER-Nelson.

Nelson, H. E., & O'Connell, A. (1978). Dementia: the estimation of pre-morbid intelligence levels using the new adult reading test. *Cortex, 14*, 234–244.

Nelson, H. E., & Willison, J. (1991). *National Adult Reading Test manual* (2nd ed.). Windsor: NFER-Nelson.

Newcombe, F. (1969). *Missile wounds of the brain: A study of psychological deficits*. London: Oxford University Press.

Newcombe, F., Oldfield, R. C., Ratcliff, G. G., & Wingfield, A. (1971). Recognition and naming of object-drawing by men with focal brain wounds. *Journal of Neurology, Neurosurgery and Psychiatry, 34*, 329–340.

Nicholas, M., Obler, L., Albert, M., & Goodglass, H. (1985). Lexical retrieval in healthy aging. *Cortex, 21*, 595–606.

Obler, L., & Albert, M. (1978). *Action Naming Test*. (Experimental ed.). Boston: VA Medical Center.

Obler, L. K., & Albert, M. L. (1985). Language skills across adulthood. In J. Birren & K. W. Schaie (Eds.), *The psychology of aging* (pp. 134–156). New York: Van Nostrand Reinhold.

Obler, L., Albert, M., & Goodglass, H. (1981). *Word finding difficulties in aging and dementia*. Paper presented at the Proceedings of the Gerontology Society of America Meetings. Toronto, ON.

O'Brien, J. T., & Levy, R. (1992). Age associated memory impairment. *British Medical Journal, 304*, 5–6.

O'Carroll, R. (1995). The assessment of premorbid ability: A critical review: *Neurocase, 1*, 83–89.

O'Carroll, R. E., & Badenoch, L. D. (1994). The inter-rater reliability of the Wechsler Memory Scale-Revised Visual Memory test. *British Journal of Clinical Psychology, 33*, 208–210.

O'Carroll, R. E., Baikie, E. M., & Whittick, J. E. (1987). Does the National Adult Reading Test hold in dementia? *British Journal of Clinical Psychology, 26*, 315–316.

O'Carroll, R. E., Conway, S., Ryman, A., & Prentice, N. (1997). Performance on the Delayed Word Recall Test (DWR) fails to clearly differentiate between depression and Alzheimer's disease in the elderly. *Psychological Medicine, 27,* 967–971.

O'Carroll, R. E., Curran, S. M., Ross, M., Murray, C., Riddle, W., Moffoot, A. P. R., Ebmeier, K. P., & Goodwin, G. M. (1994). The differentiation of major depression from dementia of the Alzheimer-type using within-subject neuropsychological discrepancy analysis. *British Journal of Clinical Psychology, 33*, 23–32.

Orgass, B. (1976). Eine Revision des Token Test II. Validitatsnachweis, Normierung und Standardisierung. *Diagnostica, 22*, 141–156.

Osterreith, P. A. (1944). Le test de copie d'une figure complexe. *Archives de Psychologie, 30*, 206–256.

Paolo, A. M., & Ryan, J. J. (1995). Selecting WAIS-R norms for persons 75 years and older. *The Clinical Neuropsychologist, 9*, 44–49.

Parker, D. M., & Crawford, J. R. (1992). Assessment of frontal lobe function. In J. R. Crawford, D. M. Parker, & W. W. McKinlay (Eds.), *A handbook of neuropsychological assessment* (pp. 267–291). Hove, UK: Lawrence Erlbaum.

Patterson, K., Graham, N., & Hodges, J. R. (1994). Reading in dementia: A preserved ability? *Neuropsychology, 8*, 395–407.

Pauker, J. D. (1988). Constructing overlapping cell tables to maximise the clinical usefulness of normative test data: rationale and an example from neuropsychology. *Journal of Clinical Psychology, 44*, 930–933.

Payne, R. W., & Jones, G. (1957). Statistics for the investigation of individual cases. *Journal of Clinical Psychology, 13*, 115–121.

Peaker, A., & Stewart, L. E. (1989). Rey's Auditory Verbal Learning Test: A review. In J. R. Crawford & D. M. Parker (Eds.), *Developments in clinical and experimental neuropsychology* (pp. 219–236). New York: Plenum.

Pentore, R., & Venneri, A. (1994). Visuo-spatial neglect in Alzheimer's disease. *Journal of Neurology, 241*(Suppl. 1), 57–58.

Perret, E. (1974). The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. *Neuropsychologia, 12*, 323–330.

Perry, E. K., Tomlinson, B. E., Blessed, G., Bergmann, K., Gibson, P. H., & Perry, R. H. (1978). Correlation of cholinergic abnormalities with senile plaques and mental test scores in senile dementia. *British Medical Journal, 2*, 1457–1459.

Pi, J., Olive, J. M., & Esteban, M. (1994). Mini Mental State examination: Association of the score obtained with the age and degree of alphabetization in an elderly population. *Medicina Clinica, 103*(17), 641–644.

Poppelreuter, W. (1917). *Die psychischen Schadigungen durch Kopfschuss im Kriege*. Leipzig: Leopold Voss.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*, 25–42.

Pratt, M. W., Boyes, C., Robins, S., & Manchester, J. (1989). Telling tales: Aging, working memory, and the narrative cohesion of story retellings. *Developmental Psychology, 25*, 628–635.

Rabbitt, P. (1990). Applied cognitive gerontology: Some problems, methodologies and data. *Applied Cognitive Psychology, 4*, 225–246.

Rabbitt, P. M. A. (1968). Channel capacity, intelligibility and immediate memory. *Quarterly Journal of Experimental Psychology, 20*, 241–248.

Reitan, R. M. (1986). Theoretical and methodological bases of the Halstead-Reitan Test Battery. In I. Grant & K. M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (pp. 3–30). New York: Oxford University Press.

Reitan, R. M., & Wolfson, D. (1994). A selective and critical review of neuropsychological deficits and the frontal lobes. *Neuropsychology Review, 4*, 161–198.

Rey, A. (1964). *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.

Reynolds, C. R., Willson, V. L., & Clark, P. L. (1983). A four-test short form of the WAIS-R for clinical screening. *Clinical Neuropsychology, 5*, 111–116.

Richards, P. M., & Ruff, R. M. (1989). Motivational effects on neuropsychological functioning: Comparison of depressed versus nondepressed individuals. *Journal of Consulting and Clinical Psychology, 57*, 396–402.

Ripich, D. N., & Terrell, B. Y. (1988). Patterns of discourse: Cohesion and coherence in Alzheimer's disease. *Journal of Speech and Hearing Disorders, 53*, 8–15.

Robbins, T. W., Joyce, E. M., & Sahakian, B. J. (1992). Neuropsychology and imaging. In E. S. Paykel (Ed.), *Handbook of affective disorders* (2nd ed., pp. 289–309). Edinburgh, UK: Churchill Livingstone.

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1994). *The Test of Everyday Attention*. Bury St. Edmunds, UK: Thames Valley Test Company.

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1996). The structure of normal human attention: The Test of Everyday Attention. *Journal of the International Neuropsychological Society, 2*, 525–534.

Rosen, W. G., Mohs, R. C., & Davis, K. L. (1984). A new rating scale for Alzheimer's disease. *American Journal of Psychiatry, 141*, 1356–1364.

Roth, M., Tym, E., Mountjoy, C. Q., Huppert, F. A., Hendrie, J., Verma, S., & Goddard, R. (1986). CAMDEX: A standardized instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia. *British Journal of Psychiatry, 149*, 698–709.

Ruff, R. M. (1988). *Ruff Figural Fluency Test administration manual.* San Franscisco: Neuropsychological Resources.

Ruff, R. M., Allen, C. C., Farrow, C. E., Nieemann, H., & Wylie, T. (1994). Figural fluency: Differential impairment in patients with left versus right frontal lobe lesions. *Archives of Clinical Neuropsychology, 9*, 41–56.

Ruff, R. M., Evans, R., & Marshall, L. F. (1986). Impaired verbal and figural fluency after head injury. *Archives of Clinical Neuropsychology, 1*, 87–102.

Ruff, R. M., Light, R. H., & Evans, R. (1987). The Ruff Figural Fluency Test: A normative study with adult. *Developmental Neuropsychology, 3*, 37–51.

Ruff, R. M., Light, R. H., Parker, S. B., & Levin, H. S. (1996). Benton Controlled Oral Word Association Test: Reliability and updated norms. *Archives of Clinical Neuropsychology, 11*, 329–338.

Ryan, J. J., & Paolo, A. M. (1992). Base rates of intersubtest scatter in the old age standardization sample of the WAIS-R. *Archives of Clinical Neuropsychology, 7*, 515–522.

Ryan, J. J., Paolo, A. M., & Brungardt, T. M. (1990). Standardization of the Wechsler Adult Intelligence Scale-Revised for persons 75 years and older. *Psychological Assessment, 2*, 404–411.

Salthouse, T. A. (1988). Effects of aging on verbal abilities: Examination of the psychometric literature. In L. L. Light & D. M. Burke (Eds.), *Language, memory, and aging* (pp. 17–35). Cambridge: Cambridge University Press.

Savage, R. D., Britton, P. G., Bolton, N., & Hall, E. H. (1973). *Intellectual functioning in the aged.* London: Methuen.

Schlosser, D., & Ivison, D. (1989). Assessing memory deterioration with the Wechsler Memory Scale, the National Adult Reading Test, and the Schonell Graded Word Reading Test. *Journal of Clinical and Experimental Neuropsychology*, *11*, 785–792.

Schneider, L. S. (1996). An overview of rating scales used in dementia research. *Alzheimer Insights*, *2*(3), 1–7.

Schonfield, A. E. D. (1980). Learning, memory and aging. In J. E. Birren & R. B. Sloane (Eds.), *Handbook of mental health and aging* (pp. 214–244). Englewood Cliffs, NJ: Prentice-Hall.

Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, *114*, 727–741.

Shapiro, M. B. (1973). Intensive assessment of the single case. In P. E. Mittler (Ed.), *The psychological assessment of mental and physical handicaps* (pp. 645–666). London: Tavistock.

Sharpe, K., & O'Carroll, R. (1991). Estimating premorbid intellectual level in dementia using the National Adult Reading Test: A Canadian study. *British Journal of Clinical Psychology*, *30*, 381–384.

Shimamura, A. P. (1994). Neuropsychological perspectives on memory and cognitive decline in normal human aging. *The Neurosciences*, *6*, 387–394.

Silverstein, A. B. (1982). Pattern analysis as simultaneous statistical inference. *Journal of Consulting and Clinical Psychology*, *50*, 234–240.

Silverstein, A. B. (1984). Pattern analysis: The question of abnormality. *Journal of Consulting and Clinical Psychology*, *52*, 936–939.

Snow, W. G., Tierney, M. C., Zorzitto, M. L., Fisher, R. H., & Reid, D. W. (1989). WAIS-R test-retest reliability in a normal elderly sample. *Journal of Clinical and Experimental Neuropsychology*, *11*, 423–428.

Spreen, O., & Strauss, E. (1991). *A compendium of neuropsychological tests. Administration, norms, and commentary*. New York: Oxford University Press.

Squire, L. R. (1987). *Memory and brain.* Oxford: Oxford University Press.

Stebbins, G. T., Wilson, R. S., Gilley, D. W., Bernard, B. A., & Fox, J. H. (1988). Use of the National Adult Reading Test to estimate premorbid IQ in dementia. *The Clinical Neuropsychologist*, *4*, 18–24.

Street, R. F. (1931). *A Gestalt Completion Test*. New York: Bureau of Publications, Teachers College, Columbia University.

Sweet, J. J., Newman, P., & Bell, B. (1992). Significance of depression in clinical neuropsychological assessment. *Clinical Psychology Review*, *12*, 21–45.

Tombaugh, T. N., & McIntyre, N. J. (1992). The Mini Mental State Examination. A comprehensive review. *Journal of American Geriatrics Society*, *40*(9), 922–935.

Trenerry, M. R., Crosson, B., DeBoe, J., & Leber, W. R. (1989). *Stroop Neuropsychological Screening Test Manual.* Odessa, FL: Psychological Assessment Resources Inc.

Tuokko, H., & Woodward, T. S. (1996). Development and validation of a demographic correction system for neuropsychological measures used in the Canadian Study of Health and Aging. *Journal of Clinical and Experimental Neuropsychology*, *18*, 479–616.

Uhlmann, R. F., & Larson, E. B. (1991). Effect of education on the Mini-Mental State Examination as a screening test for dementia. *Journal of American Geriatric Society*, *39*, 876–880.

Van Gorp, W. G., Satz, P., Kiersch, M. E., & Henry, R. (1986). Normative data on the Boston Naming Test for a group of normal older adults. *Journal of Clinical and Experimental Neuropsychology*, *8*, 702–705.

Villardita, C., Cultrera, S., Cupone, V., & Mejia, R. (1985). Neuropsychological test performances and normal aging. *Archives of Gerontology and Geriatrics*, *4*, 311–319.

Walsh, K. W. (1991). *Understanding brain damage: A primer of neuropsychological evaluation*. (2nd ed.). Edinburgh, Scotland: Churchill Livingstone.

Warrington, E. K. (1984). *Recognition memory test*. Windsor: Nelson.

Warrington, E. K., & Taylor, A. M. (1973). Contribution of the right parietal lobe to object recognition. *Cortex*, *9*, 152–164.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. New York: The Psychological Corporation.

Wechsler, D. (1987). *Wechsler Memory Scale-Revised*. New York: The Psychological Corporation.

Weingartner, H. (1984). Psychobiological determinants of memory failures. In L. R. Squire & N. M. Butters (Eds.), *Neuropsychology of Memory* (pp. 203–212). New York: Guildford Press.

Weingartner, H. (1986). Automatic and effort-demanding cognitive processes in depression. In L. W. Poon (Ed.), *Handbook for clinical memory assessment of older adults* (pp. 218–225). Washington, DC: American Psychological Association.

Weingartner, H., Cohen, R. M., Bunney, W. E., Ebert, M. H., & Kaye, W. (1982). Memory-learning impairments in progressive dementia and depression. *American Journal of Psychiatry*, *139*, 135–136.

Weingartner, H., Cohen, R. M., Murphy, D. L., Martello, J., & Gerdt, C. (1981). Cognitive processes in depression. *Archives of General Psychiatry*, *38*, 42–47.

Welsh, K. A., Butters, N., Hughes, J. P., Mohs, R. C., & Heyman, A. (1992). Detection and staging of dementia in Alzheimer's disease. Use of the neuropsychological measures developed for the consortium to establish a registry for Alzheimer's disease. *Archives of Neurology*, *49*, 448–452.

Wilson, B. A., Alderman, N., Burgess, P., Emslie, H., & Evans, J. J. (1996). *Behavioural Assessment of the Dysexcutive Syndrome*. Bury St Edmunds, England: Thames Valley Test Company.

Wilson, B. A., Cockburn, J., & Halligan, P. (1987). *Behavioural Inattention Test*. Bury St. Edmunds, UK: Thames Valley Test Company.

Wilson, B. A., Cockburn, J. M., Baddeley, A. D., & Hiorns, R. (1989). The development and validation of a test battery for detecting and monitoring everyday memory problems. *Journal of Clinical and Experimental Neuropsychology*, *11*, 855–870.

Wilson, B. A., Evans, J. J., Emslie, H., Alderman, N., & Burgess, P. (in press). The development of an ecologically valid test for assessing patients with a dysexecutive syndrome. *Neuropsychological Rehabilitation*.

Wilson, R. S., Rosenbaum, G., Brown, G., Rourke, D., Whitman, D., & Grisell, J. (1978). An index of premorbid intelligence. *Journal of Consulting and Clinical Psychology*, *46*, 1554–1555.

Woods, R. T., & Britton, P. G. (1985). *Clinical psychology with the elderly*. London: Croom Helm.

Zachary, R. A., & Gorsuch, R. L. (1985). Continuous norming: Implications for the WAIS-R. *Journal of Clinical Psychology*, *41*, 86–94.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.08
Anxiety Disorders

MELINDA A. STANLEY
University of Texas Medical School, Houston, TX, USA
and
J. GAYLE BECK
State University of New York at Buffalo, NY, USA

7.08.1 INTRODUCTION 171
7.08.2 PREVALENCE 172
  7.08.2.1 Anxiety Disorders 172
  7.08.2.2 Anxiety Symptoms 173
7.08.3 ONSET AND ASSOCIATED RISK FACTORS 173
  7.08.3.1 Age of Onset 173
  7.08.3.2 Associated Risk Factors 174
    7.08.3.2.1 Gender 174
    7.08.3.2.2 Ethnicity 175
    7.08.3.2.3 Other risk factors 175
7.08.4 PSYCHOPATHOLOGY AND DIFFERENTIAL DIAGNOSIS 175
  7.08.4.1 Nature of Anxiety 176
  7.08.4.2 Anxiety and Medical Problems 177
  7.08.4.3 Anxiety and Sleep Disturbance 178
  7.08.4.4 Anxiety and Depression 179
  7.08.4.5 Anxiety and Cognitive Impairment 180
7.08.5 MEASUREMENT 181
  7.08.5.1 Clinician-rated Measures 181
  7.08.5.2 Self-report Measures 182
  7.08.5.3 Behavioral Observation 184
  7.08.5.4 Psychophysiological Assessment 184
7.08.6 TREATMENT 184
  7.08.6.1 Utilization of Services 184
  7.08.6.2 Pharmacological Interventions 185
  7.08.6.3 Psychosocial Interventions 186
7.08.7 SUMMARY 187
7.08.8 REFERENCES 188

7.08.1 INTRODUCTION

Significant progress has been made since the late 1960s with regard to understanding and treating anxiety disorders (Barlow, 1988; Beidel & Turner, 1991). It is now well known that these conditions can have a profound impact on social and occupational functioning, and lifetime prevalence rates have indicated that anxiety disorders are one of the two most common forms of psychiatric disturbance (Robins et al, 1984). The body of knowledge currently available regarding anxiety disorders, however, is focused largely on the younger adult population. Although the need to expand this work to older adults has been emphasized repeatedly (Cartensen, 1988; Hersen & Van Hasselt, 1992; Salzman & Lebowitz, 1991), very little attention had been paid to the nature and treatment of anxiety disorders in this segment of the population. In the late 1990s, a small body of literature has begun to accumulate (Beck & Stanley, 1997; Stanley & Beck, in press). This work will be reviewed here, with particular focus on the prevalence, onset and associated risk factors, psychopathology and differential diagnosis, measurement, and treatment of anxiety in elderly people.

7.08.2 PREVALENCE

7.08.2.1 Anxiety Disorders

The National Institute of Mental Health (NIMH) Epidemiological Catchment Area (ECA) survey documented a one-month prevalence rate of 5.5% for anxiety disorders in older adults (age 65 years or older) (Regier et al., 1988; Regier, Narrow, & Rae, 1990). This figure is slightly lower than the 7.3% rate noted for younger adults, a pattern that mirrors another epidemiological report from Canada (Bland, Newman, & Orn, 1988). Although these figures suggest gradually decreasing rates of clinical anxiety disorders over the life span, they nonetheless indicate that anxiety in elderly people is a significant mental health problem. In fact, in the ECA study, anxiety disorders occurred more than twice as often as affective disorders in older adults, with a frequency four to eight times greater than major depression (Regier et al., 1988; Weissmann et al., 1985). Given the significant attention that has already been focused on depression as a major health problem in the elderly (Reynolds, Lebowitz, & Schneider, 1993), the ECA data emphasize the need for further study of older adults with anxiety problems, particularly those with symptoms which interfere with daily functioning.

The majority of anxiety disorders assigned to older adults within the ECA survey were classified as phobias (4.8%), including agoraphobia, social phobia, and simple phobia. Much lower prevalence rates were documented for obsessive-compulsive disorder (OCD; 0.8%) and panic disorder (PD; 0.1%). At the Yale University site, wherein older adults were oversampled, prevalence rates were similar (Weissmann et al., 1985). In particular, six-months figures showed an overall rate of 4.6% for anxiety disorders among elderly people, again with the majority classified as patients with phobic disorders (3.9%) and many fewer individuals with OCD (0.7%) or PD (0.1%). Additional analyses of the New Haven data indicated that prevalence rates were higher among homebound elderly people (that is, those confined to homes, beds, or chairs) relative to community residents with more freedom of movement, although these differences may have been related to poorer physical health status among the homebound individuals (Bruce & McNamara, 1992).

Even though the ECA data suggested that anxiety disorders pose a significant mental health problem for elderly people, it is important to note that these figures may have underestimated the percentage of older adults with significant anxiety problems. First, many individuals in this age group tend to underreport or deny psychological symptoms (Lasoski; 1986; Oxman, Barrett, Barrett, & Gerber, 1987). Second, prevalence figures were based on interviews with individuals residing in the community, and it has been demonstrated that rates of anxiety disorders may be much higher among institutionalized elderly people (Bland et al., 1988). Given the significant percentage of elders who reside in institutionalized settings, prevalence figures may seriously underrepresent the scope of anxiety disorders in older populations. Finally, Wave 1 of the survey, which produced the overall 5.5% rate, failed to consider prevalence of post-traumatic stress disorder (PTSD) and generalized anxiety disorder (GAD). GAD was excluded because it had not been recognized as a distinct psychiatric disorder at the time (American Psychiatric Association, 1980). Wave 2 of the study, however, which occurred approximately one year later, queried a subset of the original participants about the symptoms of PTSD and GAD. Only total population estimates for PTSD have been published (Helzer, Robins, & McEvoy, 1987), but six-month and lifetime prevalence rates of 1.9% and 4.6% were noted for GAD in the elderly (Blazer, George, & Hughes, 1991). These rates, although striking as they are presented, also may have underestimated the true prevalence of GAD given that hierarchical diagnostic criteria required exclusion of GAD if any other *Diagnostic and statistical manual of mental disorders* (*3rd ed.*) (*DSM-III*) diagnosis was present. As such, all individuals with diagnoses of major depression or PD, for example, who may also have met criteria for GAD, were not included in the published prevalence rates.

In addition to the data provided by the ECA study, at least seven other random sample community surveys from the US, Canada, and the UK have provided prevalence figures for anxiety problems in older adults (defined here as age 60 years or older; see review by Flint, 1994). Some of these studies surveyed all adults over age 18 but reported separate prevalence figures for older and younger groups, while others examined anxiety only in elderly samples. Across the studies, overall prevalence rates for anxiety disorders ranged from 0.7% to 18.6%. Figures presented separately for specific groups of disorders varied as follows: phobic disorders, 0.0%–10.0%; GAD, 0.7%–7.1%; OCD, 0.0%–1.5%; and PD, 0.1%–1.0% (Flint, 1994). These data generally are consistent with ECA findings that phobias and GAD are most common, although specific figures varied widely given the range of survey methods, case definitions, and diagnostic procedures used to establish diagnoses. For example, some diagnoses were based on self-report measures and others on semistructured clinical interviews. Some studies utilized widely accepted standards for case definition, while others relied on more idiosyncratic diagnostic criteria. In light of these methodological variations, ECA data probably represent the best figures currently available to estimate the prevalence of anxiety disorders in elderly people, despite criticisms that have been leveled at ECA methodology including variations in interview procedures at different sites, the use of lay interviewers rather than trained clinicians, and low sensitivity for diagnostic procedures (Beidel & Turner, 1991; McNally, 1994).

7.08.2.2 Anxiety Symptoms

Another collection of studies has addressed the prevalence of distressing anxiety symptoms in elderly people rather than anxiety disorder diagnoses *per se*. These reports have assessed the frequency and/or severity of symptoms using standardized self-report inventories such as the Hopkins Symptom Checklist (Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1973) and the State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970). In these surveys, cut-off scores based on standard deviations from mean scores in normal and psychiatric samples of older adults were established to define cases of elevated anxiety. This type of methodology has suggested that up to 20% of elderly people in the community report significant levels of anxiety (Feinsen & Thoits, 1986; Himmelfarb & Murrell, 1984). Significant anxiety has also been reported for 8–11% of elders assessed in primary care settings in both the US (Oxman et al., 1987) and Hong Kong (Wong & Pan, 1994). When relevant comparisons have been made, patterns again fail to confirm the notion of increasing anxiety over the life span given lower prevalence of distressing anxiety in older than younger adults (Feinsen & Thoits, 1986; Oxman et al., 1987). Although these figures disallow differential diagnosis that would separate anxiety symptoms resulting from various medical and/or psychiatric conditions (see Section 7.08.4), they nonetheless confirm that a substantial number of elders report distressing anxiety symptoms.

7.08.3 ONSET AND ASSOCIATED RISK FACTORS

7.08.3.1 Age of Onset

It has been suggested from the younger adult literature that age of onset for anxiety disorders may be based at least to some degree on developmental stage (Beidel & Turner, 1991). For example, the onset of specific phobias seems to mirror the progression of developmentally appropriate normal fears through childhood; that is, animal phobias are often first noticed between the ages of four and seven, when animal fears are common (Marks & Gelder, 1996; Ost, 1987), and phobias of natural disasters and health-related situations frequently begin among older elementary school children at the ages when these fears are commonly reported (Barrios, Hartmann, & Shigetomi, 1981). Likewise, the age of onset for social phobia is typically in early to late adolescence when the establishment of social relationships is a critical developmental task (Beidel & Turner, 1991). These patterns suggest that some phobias reported by older adults may represent fears which originated in earlier developmental stages and continued through later adulthood, while others may have developed more recently from concerns that are developmentally appropriate for older adults, for example, fears of falling (Downton & Andrews, 1990). In fact, available data regarding age of onset for phobias in elderly people are mixed, suggesting the possibility of varying periods of onset for different fears. In a British survey of older adults, Lindesay (1991) reported that specific phobias generally had their onset during childhood, a finding that is consistent with data from the younger adult literature (Thyer, Parrish, Curtis, Nesse, & Cameron, 1985). Alternatively, data from the the US ECA study implied that many phobias reported by older adults were of recent onset (Eaton et al., 1989). Similarly, Lindesay reported that agoraphobic fears (e.g., of enclosed places crowds, public transportation, going away from home) among elderly people most often were of recent onset. However, many of these "agoraphobic" fears began after an episode of physical illness or trauma. As such, the fears may have represented realistic or excessive fears of physical limitations rather than anticipation of panic, suggesting that a diagnosis of specific phobia might have been more appropriate for consideration than agoraphobia (see McNally (1994) for a discussion of this diagnostic issue in younger adults). In addition, a high incidence of depression in Lindesay's sample suggested that "agoraphobic" fears may have been associated with a social withdrawal characteristic of a depressive clinical picture (Flint, 1994).

In light of these inconsistent findings and important differential diagnosis issues, as well as the influential role played by developmental stages in the origin of phobias, further investigation is needed to understand patterns of onset for different types of phobias reported by elderly people. In this area, the important differentiation of realistic versus excessive fear is central, particularly given developmentally appropriate fears that may be characteristic of normal aging.

The age of onset for GAD among elderly people appears to have a bimodal distribution. In Wave 2 of the US ECA study, 39% of participants reported duration of GAD for 21 years or more (Blazer et al., 1991). For these individuals, GAD may have been conceptualized best as a personality disturbance rather than an Axis I disorder, a perspective that has received some attention in the literature of the 1990s (Sanderson & Wetzler, 1991). For another 52% of participants, the onset of GAD was reported within five years of the interview. For this subset of older adults, GAD may have begun as a reaction to stressful life events, a pattern that has been documented in both younger and older adults (Blazer, Hughes, & George, 1987; Ganzini, McFarland, & Culter, 1990). In particular, younger adult men and women interviewed in Wave 2 of the ECA survey experienced a threefold increase in risk for GAD in the year following one or more unexpected, negative, and important life events (Blazer et al., 1987). In a separate sample of older adults, an increased incidence of GAD was reported during the 20 months following a catastrophic financial loss (Ganzini et al., 1990). In this sample, 27% of participants reported symptoms meeting criteria for GAD during the postloss period relative to 10% of elders in a control sample during this same time.

A bimodal pattern of onset for GAD was also reported in a sample of older adults recruited for participation in a treatment study (Beck, Stanley, & Zebb, 1996a). A comparison of clinical features (e.g., anxiety, depression, specific fears) in patients with onset of GAD in childhood (before age 15) and those reporting onset in middle adulthood (after age 39) revealed very few differences. The impact of age of onset on treatment response was not studied, however, given inadequate sample sizes.

Given that prevalence rates for PD and OCD in older adults are quite low, few data are available to address the onset of these syndromes. Although ECA data indicated that neither of these disorders typically began in older age (Eaton et al., 1989), a retrospective chart review of 51 elderly clinic patients with panic disorder revealed that over half reported onset at age 60 years or later (Raj, Covea, & Dagon, 1993). Related case reports have also suggested that panic disorder may begin in later life (Luchins & Rose, 1989; Sheikh, King, & Taylor, 1991). However, conclusions about time and mode of onset for PD and OCD will require greater study.

The potential role of stressful life events in the development of anxiety in later life has already been noted (Ganzini et al., 1990). It has also been demonstrated in studies of Holocaust survivors and World War Two prisoners of war that stress-related symptoms of PTSD can persist into old age (Kluznick, Speed, Van Valkenberg, & Magraw, 1986; Kuch & Cox, 1992). However, no data are available regarding onset of PTSD in elderly people. Given the number of traumatic life events that can occur for older adults (e.g., elder abuse, muggings, motor vehicle accidents), onset of PTSD symptoms in later life is of significant interest and importance. Relevant to this issue are published findings suggesting that prior experience with trauma can serve to "inoculate" older adults against the anxiety-related effects of natural disasters (Norris & Murrell, 1988). The more general role of prior stress as a buffer against the development of PTSD in late life will need to be considered.

7.08.3.2 Associated Risk Factors

7.08.3.2.1 Gender

As has been shown repeatedly among younger adults, ECA data demonstrated that elderly women are at greater risk of experiencing anxiety disorders than elderly men, with a ratio of approximately 2:1 among adults over age 65 (Regier et al., 1988; Regier et al., 1990; Weissman et al., 1985). In a similar Canadian survey (Bland et al., 1988), older women residing in the community were 1.5 times more likely to be diagnosed with an anxiety disorder than men in the same age group. In an institutionalized subgroup of participants surveyed for this study, women were five times as likely to be assigned an anxiety disorder diagnosis as men. However, figures from both of these studies failed to address the prevalence and/or gender distribution of GAD. Two studies relevant to this issue have suggested that elderly women are approximately two to nine times more likely than elderly men to meet criteria for this disorder (Flint, 1994). In Wave 2 of the ECA survey, prevalence rates for GAD within gender were presented separately for black and nonblack participants (Blazer et al., 1991). These figures indicated that elderly black women were 12 times more likely to be assigned a diagnosis of GAD than black men in the same age group. Nonblack elderly women also were more likely to be diagnosed with GAD than their nonblack male counterparts, with a ratio of 4:1.

With regard to the prevalence of notable anxiety symptoms instead of disorders, women have also been identified more often than men to have significant anxiety according to pre-established cutoff scores on standardized self-report instruments (Himmelfarb & Murrell, 1984). Similarly, elderly women reported greater overall fearfulness than elderly men on a standardized measure of specific fears (Liddell, Lockes & Burman, 1991). Thus, epidemiological data overall indicate that women are at greater risk of experiencing problematic anxiety than men among the older adult population.

7.08.3.2.2 Ethnicity

Another potential risk factor for the experience of anxiety in elderly people is ethnicity, although the impact of this variable has been understudied. In the ECA survey, problems with sampling strategies for some minority groups were noted across all age groups (Beidel & Turner, 1991), making it difficult to draw even general conclusions from these data. Moreover, very few ECA data relevant to ethnicity have been provided separately by age. Among the elderly people interviewed through the New Haven site, Weissman et al. (1985) noted that rates of cognitive impairment were higher in blacks than whites, although other *DSM-III* disorders were less prevalent among blacks, relative to whites. However, no data were presented regarding the distribution of anxiety disorders among these groups. In Wave 2 of the survey, prevalence rates for GAD in elderly people were presented by both ethnicity and gender, as noted earlier, with figures suggesting the highest rate among black women (3.7%). Rates for the other groups were as follows: nonblack women (2.7%), black men (0.3%), and nonblack men (0.7%).

In studies examining prevalence rates for anxiety symptoms rather than disorders, nonsignificant correlations between ethnicity and anxiety severity have been reported (Feinson & Thoits, 1986; Himmelfarb & Murrell, 1984). However, in one of these reports, minority groups were underrepresented (Feinson & Thoits, 1986), and in the other, the ethnic distribution of the sample was unspecified (Himmelfarb & Murrell, 1984). Given the very limited data available to address this issue, as well as significant methodological problems evident in the reports that are accessible, the potential role of ethnicity as a risk factor for the experience of anxiety in later life is unclear and requires future study.

7.08.3.2.3 Other risk factors

Other potential risk factors for the experience of anxiety in elderly people relate to marital status, income, and level of education. In particular, fewer anxiety symptoms have been reported by married older adults relative to unmarried older persons, including all individuals not married at the time of assessment (Feinsen & Thoits, 1986). A more complicated interaction between marital status and gender was reported in another study, with highest levels of anxiety occurring among never-married men and the least amount of anxiety within a group of never-married women (Himmelfarb & Murrell, 1984). Significant negative correlations have also been noted between income and anxiety among the elderly (Feinsen & Thoits, 1986; Himmelfarb & Murrell, 1984; Liddell et al., 1991), and in some reports, between level of education and anxiety severity (Himmelfarb & Murrell, 1984; Liddell et al., 1991). In general, these patterns are similar to those reported in younger adults with anxiety disorders (Regier et al., 1990).

7.08.4 PSYCHOPATHOLOGY AND DIFFERENTIAL DIAGNOSIS

Significant advances have been made in understanding the nature and experience of anxiety in younger adults (see Barlow, 1988; Beidel & Turner, 1991; Rapee & Barlow, 1991 for reviews of this literature). However, whether this knowledge applies directly to the experience of anxiety among elderly people is unknown. Of crucial importance to this issue are descriptions of the specific nature of anxiety in older adults with and without psychiatric disorders, the relationship between anxiety and medical problems common in older adults, and the overlap between anxiety and sleep disturbance, depression, and cognitive impairment, in particular, memory. These issues are reviewed here.

7.08.4.1 Nature of Anxiety

A small body of literature has begun to address potential differences in the experience of anxiety across the life span. As noted earlier, the prevalence of anxiety disorders seems to decrease with age, but also of interest are potential differences in the qualitative experience of anxiety from younger to older adulthood. In one relevant empirical examination, Lawton, Kleban, and Dean (1993) examined potential differences in the factor structure and prevalence of affect across three age groups: younger adults (ages 18–30), middle-aged adults (ages 31–59), and older adults (age 60 and over). Although a six-factor structure was evident for each group in initial exploratory analyses, subsequent confirmatory analyses revealed differences in the loading sizes, factor correlations, and residual correlations across the three groups. The primary differences occurred between younger and older adults with regard to factors assessing positive affect and depression, but differences between these two groups were also evident on the anxiety-guilt factor. In particular, affect terms assessing guilt (e.g., ashamed, guilty, worried, blamed) were more salient to this factor for younger adults than for the older group. As suggested by the authors, this difference may indicate that older adults grew up before "guilt became fashionable" (p. 174). Consequently, they may experience anxiety more directly than younger adults who make cognitive connections between anxiety and guilt (Lawton et al., 1993). Other analyses revealed that older adults reported less of almost every negative emotional state, including anxiety, than participants in the other two groups. Fewer differences between groups were evident with regard to the experience of positive affect. These data suggest that affect terms may have different meanings across the life span and that the experience of anxiety may vary in important ways with age. As such, theoretical and empirical knowledge regarding the experience of anxiety among younger adults cannot be extrapolated directly to an older population.

Another group of studies has addressed the nature of fears and worries among older adults living in the community. As noted earlier, a developmental examination of normal fears is needed as a baseline against which to understand the experience of problematic anxiety. For older adults, as with individuals in any other age group, it is necessary to differentiate maladaptive anxiety from realistic fears that accompany a particular developmental stage, in this case, aging. It is also important, however, for practitioners not to overattribute anxiety and fears in elderly people to the normal aging process and in so doing, overlook or ignore treatable psychological symptoms (Weiss, 1994). In either case, the experience of anxiety as it accompanies the normal aging process needs to be understood clearly.

The majority of work addressing the nature of normal fears among elderly people has focused on worry, probably because generalized anxiety is one of the most common forms of anxiety in this age group (Blazer et al., 1991). In one community survey of adults ages 25–64, Person and Borkovec (1995) classified participants' worries or concerns into five categories: family–home–interpersonal, illness–health–injury, work–school, finances, and miscellaneous. Comparisons of worries reported by younger (age 25–64) and older (age 65 and over) adults revealed that elderly people worried most about health and least about work, whereas younger adults were most concerned about family and finances. These data suggested that worry content across age groups was representative of changing life circumstances. Similar findings were reported in another community survey of worries in younger and older adults (Powers, Wisocki, & Whitbourne, 1992). In particular, the elderly reported few worries overall, with significantly fewer worries about social events and financial issues than a college-age comparison group. Older adults worried most about health, although scores in this domain were not significantly different from the younger sample.

In an assessment study with 94 older adults characterized via semistructured interview to be free of any psychiatric diagnoses, scores on standardized self-report measures of worry, state and trait anxiety, specific fears, and obsessive-compulsive symptoms were significantly lower than normal mean scores reported in other samples of younger adults (Stanley, Beck, & Zebb, 1996). These data accentuate the need for age-appropriate norms to interpret scores on self-report measures of anxiety symptoms and are consistent with prior findings suggesting that severity of anxiety seems to decrease with age.

The fact that older age is not necessarily associated with increasing anxiety should not detract from the impact that problematic anxiety can have in this population. In one of the first studies to examine the clinical characteristics of anxiety disorders in elderly people, 44 older adults with GAD (diagnosed according to semistructured interview and *DSM-III-R* criteria) were compared with a sample free of psychiatric diagnoses who were equivalent with regard to age, gender, level of education, and ethnicity (Beck et al., 1996a). Analyses revealed that GAD was associated with increased levels of worry, anxiety, social fears, and depression, with scores on these measures comparable to younger samples of GAD patients. Similarly, Raj et al. (1993) reported that clinical characteristics of a group of elderly patients with PD were similar to those reported in younger samples with the same diagnosis. As such, although the overall severity and frequency of fears and worries may decrease over the life span, anxiety disorders appear to be associated with similar clinical features regardless of age. However, additional data with groups of elderly patients meeting criteria for various anxiety disorders are needed to replicate and expand these findings.

7.08.4.2 Anxiety and Medical Problems

The overlap between medical and psychological problems is a highly salient issue for practitioners working with older adults. In fact, it has been suggested that these two areas are so closely interwoven that treatment of mental health problems is impossible without knowledge of medical illnesses common among the elderly (Frazer, 1995). As noted by Frazer, the relationship between physical and psychological symptoms or disorders can occur for at least three reasons: medical illnesses can mimic and/or cause psychological difficulties, the two types of symptoms can coexist, and psychological disturbance can potentially leave one vulnerable for development of physical difficulties. Katz (1996), in a 1995 presidential address to the American Association for Geriatric Psychiatry, noted that the significant degree of interrelatedness between psychological and medical difficulties among elderly people should serve a central role in the development of public policy and the design of health care services for the elderly. In this domain, a reasonable amount of attention has focused on the comorbidity between medical disease and depression (Katz, 1996), but less work has addressed the relationship between physical illness and anxiety.

It is of particular importance to consider the overlap of anxiety and medical disease, given that older adults often attribute anxiety-related symptoms such as agitation, fears, or aches and pains to physical illness, thereby denying or underreporting psychological difficulties (Gurian & Miner, 1991). Furthermore, a range of medical conditions can create anxiety-like symptoms, in particular, cardiovascular disease, chronic obstructive pulmonary disease, hyperthyroidism, and visual and auditory impairments. (Cohen, 1991). Limited empirical data have addressed the nature of the overlap between physical illness and anxiety, although a few studies have explored the prevalence of anxiety symptoms in samples of older adults presenting to medical clinics or hospitals.

As noted earlier, prevalence rates of 8–11% for anxiety disorders were reported among elderly outpatients (age 55 years and over) in a rural US group practice (Oxman et al., 1987). Case identification in this study was based on symptom checklist scores that were at least two standard deviations above a mean. Given apparent differences in prevalence rates for anxiety in urban and rural settings (Blazer et al., 1991), along with the restricted ability of self-report inventories to identify diagnosable cases, the data from this report are somewhat limited. In another survey of Italian medical inpatients, 40% of adults over age 60 reported moderate to high levels of distress on an anxiety symptom checklist (Magni, Schifano, deDominicis, & Belloni, 1988). Severity of anxiety was not associated with the classification of medical problems as being acute or chronic, but scores on a phobia subscale were significantly higher for patients with central nervous system disturbances than for those with other diseases. Again, although these data provide an interesting pattern of results, conclusions are limited given reliance solely on a self-report symptom checklist to evaluate anxiety.

In two other studies, elderly medical patients were evaluated to examine the relationship between complaints of postural disturbance (dizziness, fear of falling) and anxiety. In one of these, 203 adults age 75 and over were recruited from four general medical practices in Manchester, UK (Downton & Andrews, 1990). In this sample, dizziness and fear of falling were associated with higher levels of anxiety, and symptoms were more severe for participants who had actually experienced a fall within the previous 12 months. It is interesting to note that 42% of the sample reported restricted activity due to fear of falling, a behavioral pattern that requires differentiation from agoraphobic fears reported in older adults (Beck & Stanley, 1997). In another report, 37.5% of 56 older adults (age 60 and over) who presented to a dizziness clinic were diagnosed with at least one *DSM-III-R* Axis I disorder (Sloane, Hartman, & Mitchell, 1994). The most frequent diagnoses in this group were anxiety disorders, depression, and adjustment disorders. Older adults with chronic dizziness also scored significantly higher on a symptom checklist of anxiety than a matched group of healthy adults. In a follow-up study of stroke survivors, prevalence rates of 5% and 19% for anxiety disorders were reported for elderly men and women at four months post-stroke (Burvill et al., 1995). When hierarchical rules for diagnosing *DSM-III* disorders were suspended, rates were 12% and 28% for men and women, respectively. Participants with anxiety disorders at the time of stroke were identified via retrospective interview. When these individuals were excluded, prevalence rates for men and women were 9% and 20%, figures that far outweigh general population estimates for older adults.

These data suggest that elderly patients presenting for medical treatment frequently have disturbing symptoms of anxiety. As such, it is important for medical practitioners to be aware of the signs and symptoms of anxiety and to provide appropriate treatment or referrals when needed. Likewise, it is important for mental health care providers to be cognizant of medical causes or correlates of anxiety (Stanley & Beck, in press). In this latter domain, epidemiological data from a British survey documented that older adults with phobic disorders reported significantly more physical symptoms than control participants (Lindesay, 1991). In particular, adults with phobic disorders more frequently reported the experience of palpitations, dyspnea, dysphagia, abdominal pain, bowel problems, giddiness, tinnitus, and joint swelling within the three months prior to interview. In another community survey, a strong positive correlation ($r=0.60$) was observed between severity of anxiety and physical health problems (Himmelfarb & Murrell, 1984). Given a reported low correlation between age and physical health, however, the authors concluded that the observed relationship between anxiety and physical condition was not mediated only by age. Multivariate examination of this issue appears warranted, as bivariate correlations do not reveal complex patterns of interrelationships among variables. In a study of institutionalized elders living in nursing homes or congregate housing, significant associations were also observed between anxiety and physical health, with more anxious residents reporting more severe renal problems and other gastrointestinal and urninary difficulties (Parmelee, Katz, & Lawton, 1992).

Taken together, these data strongly support the importance of considering the overlap between anxiety symptoms and medical health problems. To determine appropriate diagnoses and treatment approaches, close attention to both of these areas is paramount, with careful focus on differential diagnosis issues. It seems clear that any evaluation of older adults presenting for either psychological or medical help should include detailed assessments in both domains.

7.08.4.3 Anxiety and Sleep Disturbance

Sleep disturbance, in particular insomnia, is a frequent problem among older adults (Engle-Friedman & Bootzin, 1991) and in fact, is the topic of Chapter 14, this volume. Of interest here, however, is the potential overlap between sleep disturbance and anxiety. Sleep difficulties in elderly people are often viewed as part of the normal aging process and thereby overlooked or underdiagnosed, as are many other symptoms reported frequently in this group (Morin & Gramling, 1989). However, sleep disturbance can lead to changes in mood and conversely, stress and psychopathology can create disruptions in normal sleep patterns (Engle-Friedman & Bootzin, 1991). Thus, an examination of the overlap between sleep difficulties and anxiety is of relevance.

Insomnia has been associated with increases in a variety of self-reported psychological symptoms, including introversion, anxiety, neuroticism, and depression (Engle-Friedman & Bootzin, 1991). When relevant comparisons of older and younger adults have been undertaken, elders reporting sleep disturbance have endorsed fewer symptoms of anxiety, depression, and other psychopathology than younger individuals with similar difficulties (Bliwise, Bliwise, & Dement, 1985; Roehrs, Lineback, Zorick, & Roth, 1982; Roehrs, Zorick, Sicklesteel, Wittig, & Roth, 1983). Morin and Gramling (1989) suggested a number of possible explanations for this pattern, including differences in typical problematic sleep patterns across the two groups (i.e., older adults more often report trouble maintaining rather than initiating sleep) and the tendency of older adults to focus on physical rather than psychological symptoms. Nonetheless, at least two recent studies have demonstrated a significant relationship between sleep difficulties and increased anxiety among the elderly.

In the first of these, Morin and Gramling (1989) assessed 42 adults age 60 and over who responded to media advertisements for participants with chronic sleep difficulties. These individuals were compared on a variety of sleep and mood measures with a control group of 30 individuals matched for age, gender, education, and marital status. Results demonstrated increased levels of anxiety and depression in the self-reported insomniac group relative to the comparison group without sleep difficulties. However, scores on mood measures within the poor sleeper group were not in the pathological range, suggesting to the authors that sleep disturbance alone may not be a good indicator of psychological problems among elderly people. In a more recent study, Friedman, Brooks, Bliwise, and Yesavage (1993) extended these findings by including an objective measure of sleep, in particular a wrist actigraph. Twenty-seven older community volunteers with sleep complaints were assessed on measures of state and trait anxiety and depression. They were also asked to complete daily sleep logs for a period of two weeks and to wear a wrist actigraph continuously for a period of three days. Results demonstrated no correlation between state anxiety and any measure of sleep. However, trait anxiety correlated positively with two self-reported sleep measures: amount of time spent awake after first falling asleep, and total time spent in bed. Trait anxiety also correlated in a significant positive way with total sleep time as determined by the actigraph. The positive correlation between anxiety and total sleep time was unexpected, although the authors noted that the actigraph may have overestimated sleep time for participants who spent a lot of time lying awake and motionless in bed. They also suggested that more anxious adults may have perceived their sleep to be worse than it actually was.

In addition to the mixed data regarding anxiety in older adults with sleep complaints, it is of interest to note that some treatment studies for sleep disorders in elderly people have examined the impact of relaxation treatment (Friedman, Bliwise, Yesavage, & Salom, 1991) or have included relaxation as part of a multicomponent treatment package (Morin, Colecchi, Ling, & Sood, 1995). The impact of relaxation for reducing anxiety is well-known, although studies addressing the utility of these interventions for sleep disturbance failed to include measures of anxiety that might have addressed the impact of treatment on coexistent psychological symptoms. In another study of psychosocial treatment for insomnia that omitted an anxiety-reduction component (Morin, Kowatch, Barry, & Walton, 1993), no impact of treatment on anxiety was noted. However, given the apparent overlap between symptoms of anxiety and sleep difficulties for at least some patients, it will be of interest for future treatment studies of both anxiety and insomnia to examine more regularly transfer effects to the nontarget set of symptoms.

7.08.4.4 Anxiety and Depression

The overlap between anxiety and depression is an issue of longstanding importance within psychology and psychiatry given the high rates of coexistence for these conditions regularly reported among younger adults (Cloninger, 1990; Copp, Schwiderski, & Robinson, 1990; Feldman, 1993). In fact, six-month prevalence data from the ECA survey documented that one-fifth of individuals with an anxiety disorder also had symptoms that met criteria for an affective disorder (Regier et al., 1990). Conversely, one-third of those assigned an affective disorder diagnosis also were given at least one anxiety disorder diagnosis (Regier et al., 1990). Given these high rates of coexistence, further understanding of the nature of this overlap is warranted among elderly people.

ECA data regarding the coexistence of anxiety and affective disorders were not presented separately for the elderly. Nonetheless, recent reviews have highlighted the significant overlap between these sets of diagnoses and symptoms in this segment of the population (Alexopoulos, 1991; Flint, 1994). With regard to data addressing coexistence of assigned diagnoses, it has been reported that over 90% of older adults from the community who reported symptoms appropriate for a diagnosis of GAD also met criteria for depression (Lindesay, Briggs, & Murphy, 1989). In this same community sample, almost 40% of individuals with a phobia diagnosis were also assigned a diagnosis of depression according to the *International classification of disease* (*8th ed.*) (*ICD-8*) criteria assessed via semi-structured interview. In another study conducted via retrospective chart review (Raj et al., 1993), 45% of elderly clinic patients with PD diagnosed according to *DSM-III-R* also met criteria for coexistent major depression. In an examination of 44 older adults with a principal *DSM-III-R* diagnosis of GAD (Beck et al., 1996a), approximately 15% met criteria for a coexistent affective disorder. When older adults with affective disorders have been evaluated, over one-third have met criteria for a coexistent anxiety disorder diagnosis (Alexopoulos, 1991). Thus, rates of overlap between anxiety and affective disorders among the elderly are high.

Other reports have addressed the degree of overlap between symptoms of anxiety and depression, rather than prevalence of coexistent disorders *per se*. In these studies, significant positive associations between symptoms of anxiety and depression have been observed in community samples of older adults (Lindesay et al., 1989; Ben-Arie, Swartz, & Dickman, 1987), institutionalized elderly people (Parmelee et al., 1992), and older clinic patients (Alexopoulos, 1991). In the authors' own investigation of older adults with GAD, higher levels of depressive symptoms were also evident relative to a matched control sample (Beck et al., 1996a).

In a longitudinal study of 1070 older adults from the community, 702 of whom were interviewed twice over a three-year interval, Larkin et al (1992) reported some covariance between anxiety and depressive symptoms assessed in the context of "neurotic" conditions. These data highlight further the significant overlap between anxiety and depression among the elderly, and may even be interpreted to support some indistinguishability among these two sets of symptoms (Beck & Stanley, 1997). However, data reviewed below attest to the ability of clinicians to differentiate and diagnose both types of disorders reliably among elderly samples (see Section 7.08.5).

7.08.4.5 Anxiety and Cognitive Impairment

Normal age-related changes in cognitive functioning have been studied extensively (see Chapter 2, this volume), and a relatively large body of work has examined the impact of depression on cognitive performance among elderly people (see Chapters 9 and 10, this volume). Results of this latter research have demonstrated moderate relationships between depression and cognitive impairment that are influenced by sample characteristics and the measures selected to evaluate cognitive function (Burt, Zembar, & Niederehe, 1995; Poon, 1992). Less work has examined the overlap between anxiety and cognitive functioning, and results of available studies are mixed and limited by a number of methodological problems. In particular, wide ranges of both anxiety and cognitive measures have been used, limiting the ability to generalize findings across studies. In addition, the majority of studies have been conducted with community volunteers recruited for participation in studies of memory. No studies to date have examined cognitive function among older adults who report significant anxiety symptoms or who meet diagnostic criteria for anxiety disorders. Nonetheless, among older adult volunteers, some relationships have been demonstrated between increased anxiety and decreased cognitive performance, most often related to memory.

In one early report, 67 women volunteers over age 65 were administered measures of anxiety, verbal and nonverbal memory, and self-reported memory complaints in a single testing session (West, Boatwright, & Schleser, 1984). Results demonstrated significant negative correlations between state anxiety and measures of nonverbal memory (recall for sequences of related or unrelated numbers), although no similar relationship was demonstrated between anxiety and memory for lists of related or unrelated words. Higher state anxiety was also related to increased memory complaints, although life satisfaction was a stronger predictor of perceived memory abilities.

In another study of community volunteers, ages 60–79, significant negative correlations were noted between state anxiety and cognitive performance measured by a Piagetian problem-solving task and the Wechsler Adult Intelligence Scale Similarities subtest (LaRue & D'Elia, 1985). In this report, correlations between anxiety and cognitive function were stronger among younger adults ($r = 0.24$–0.46), ages 40–59, relative to older adults ($r = 0.09$–0.29). Hill and Vandervoort (1992) reported a significant negative association between state anxiety and performance on a verbal memory task for a group of 74 older adult volunteers, age 60 years and older. However, it is important to note that the measure of anxiety in this study was administered after participants were given feedback about their cognitive performance. As such, feedback that performance had been poorer than expected might have created elevated levels of anxiety. In another comparison, significant negative correlations were observed between measures of verbal recall and anxiety, depression, and social withdrawal for older adult volunteers, ages 60–78 (Deptula, Singh, & Pomara, 1993). In this group, ratings of social withdrawal accounted for a significantly larger proportion of the variance in memory scores than both anxiety and depression. In a younger group of adults assessed in the same study (ages 19–35), no significant relationships were observed between any measure of affect and memory.

In addition to these studies demonstrating significant, albeit moderate, relationships between anxiety and cognitive performance, another small set of reports has examined the impact of anxiety reduction treatment on memory. In the earliest of these, 26 older adult volunteers, ages 59–85, were recruited for a memory improvement study (Yesavage, Rose, & Spiegel, 1982). Participants with evidence of dementia or significant depression were excluded, and four weekly sessions of relaxation training were provided. Anxiety measures were taken at pretreatment, and verbal recall was measured before and after relaxation induction during session three. Results revealed significant positive correlations between two measures of anxiety and improvement in recall following relaxation. In other words, participants with increased anxiety before treatment improved more on the recall task following relaxation. In a second study, Yesavage and Jacob (1984) provided training in both relaxation and mnemonic procedures for a group of 25 older adults, ages 61–82. Results revealed improvement in memory following the combined intervention and significant negative correlations between anxiety and recall at post-treatment.

In general, this small body of literature suggests some potentially important overlap between anxiety and cognitive function. However, further studies are needed to examine cognitive performance in older adults with well-diagnosed anxiety disorders relative to carefully selected control participants. In addition, broad-based and well-standardized measures of both anxiety and cognitive function need to be used.

7.08.5 MEASUREMENT

Severity of anxiety can be measured in a variety of ways, including clinician ratings, standardized self-report instruments, behavioral observation, and psychophysiological recording. Among younger adults, strategies for the evaluation of anxiety within each of these domains have received extensive study, and many well-established measures are currently available (Nietzel, Bernstein, & Russell, 1988). However, as has been suggested elsewhere (Hersen & Van Hasselt, 1992, Hersen, Van Hasselt, & Goreczny, 1993), the same type of work is needed to establish the utility of a range of anxiety measures for use with elderly people. In fact, the development of psychometrically sound assessment tools for evaluating anxiety in this population will be the cornerstone of future research examining the psychopathology and treatment of older anxious adults.

At least two general approaches have been taken in the initial development of anxiety measures for older adults. In one of these, the utility of measures developed with younger samples has been assessed with older participants. This approach is economical and allows for direct comparisons of clinical features in younger and older groups. However, such a strategy may fail to take into account phenomenological features of anxiety that are particularly salient for older adults (e.g., worries about health), and may overlook potentially important discriminations between anxiety and related difficulties (e.g., medical problems, depression) (Hersen et al., 1993; Sheikh,1991). Given these limitations, an alternative approach involves the development of unique measures of anxiety for use among the elderly. This method is more costly in terms of time and expense, but it allows for more specialized assessment. Both of these strategies have been used to produce a small amount of work focused on the development of assessment technques for measuring anxiety in older adults.

7.08.5.1 Clinician-rated Measures

Clinician-rated instruments have been used routinely among younger adults to diagnose anxiety disorders and evaluate the severity of anxiety symptoms. With regard to establishing anxiety disorder diagnoses, recent data have suggested that two semistructured interviews used routinely among younger adults can also be used reliably in elderly samples. First, the Anxiety Disorders Interview Schedule-Revised (ADIS-R; DiNardo, Moras, Barlow, Rapee, & Brown, 1993) was used by the authors and colleagues in a set of studies designed to examine the nature and treatment of GAD among adults ages 55 and over (Beck et al., 1995, 1996a; Stanley, Beck, & Glassco, 1996; Stanley et al., 1996). Estimates of interrater reliability for use of this measure indicated 100% agreement for principal diagnoses of *DSM-III-R* GAD (although this may have been due at least in part to extensive telephone screening of participants prior to ADIS-R administration), as well as perfect agreement for coexistent diagnoses of social phobia, simple phobia, and panic disorder ($\kappa = 1.00$). Moderate reliability was noted for the diagnosis of major depression ($\kappa = 0.58$). Although these data were obtained with only 28% of the total sample ($n = 14$ of 50), the figures nonetheless suggest that the ADIS-R can be a useful tool for the diagnosis of anxiety and affective disorders among older samples. Future studies will need to address the utility of the ADIS-IV (Brown, DiNardo, & Barlow,1994), which was recently created to conform with *DSM-IV* criteria (American Psychiatric Association, 1994).

Another set of studies examined the utility of the structured clinical interview for *DSM-III-R* (SCID, Spitzer, Williams, Gibbon, & First, 1988) among older adults (Segal, Hersen, Van Hasselt, Kabacoff, & Roth, 1993; Segal, Kabacoff, Hersen, Van Hasselt, & Ryan, 1995). This interview covers a broader range of potential psychiatric disorders than the ADIS, and data have supported its use for the diagnosis of anxiety, affective, and somatoform disorders in older samples. In particular, the SCID was administered to two sets of adults ages 55 years and older, with all interviews reviewed by a second clinician to evaluate interrater agreement for any diagnostic category with a base rate over 10%. Kappa coefficients indicated adequate agreement for the broad categories of anxiety disorders (0.73–0.77) and somatoform disorders (0.84–1.00), as well as the more specific diagnoses of major depression (0.70–0.79) and panic disorder (0.80). These data suggest that the SCID can also be useful for the diagnosis of anxiety disorders among the elderly. It is important to note, however, that all data examining the reliability of the ADIS and SCID have utilized videotaped interviews which may heighten interrater agreement given that reliability raters are able to hear the interviewer's questions at key decision points of the interview. In this regard, future studies should address the reliability of the interviews when administered separately by two different clinicians.

Other clinician-rated instruments focus on the severity of anxiety symptoms, irrespective of psychiatric diagnosis. The most well-known of these instruments is the Hamilton Anxiety Rating Scale (HARS, Hamilton, 1959), a 14-item measure used routinely in studies of anxiety in younger adults. Although the potential utility of the HARS among the elderly has been questioned due to its emphasis on somatic symptoms (Sheikh, 1991), a report by the authors and colleagues provided support for its use among older adults (Beck, Stanley, & Zebb, 1996b). In particular, interrater agreement for the HARS was adequate among a subgroup of older adults with GAD. Additionally, the instrument successfully discriminated the patient sample from a matched group of normal control participants without psychiatric diagnoses, with near perfect classification. The patient and control groups were both selected based on administration of the ADIS-R, and normative data for the HARS were provided for each. As such, this instrument is a potentially viable measure for evaluating anxiety severity in older adults. However, correlations between the HARS and the Hamilton Rating Scale for Depression (HRSD, Hamilton, 1960) were high (Beck et al, 1996b), and the utility of these measures for discriminating between anxiety and depression among the elderly warrants further study. Revisions in both the HARS and HRSD have been proposed in an attempt to improve internal consistency and discriminant validity (Riskind, Beck, Brown, & Steer, 1987). However, these reconstructed scales continue to share considerable variance among older adults (Beck et al. 1996b), a finding that has also been documented in younger samples (Moras, DiNardo, & Barlow, 1992).

7.08.5.2 Self-report Measures

A wide range of psychometrically sound self-report measures are available to assess anxiety in young and middle-aged adults, although only a small body of literature has examined the utility of these same measures with older samples. Moreover, much of this work is limited due to heterogeneous samples of participants, omission of standardized diagnostic procedures for group assignment, and potentially confounding variables including medication use and poor physical health (Hersen & Van Hasselt, 1992).

The earliest studies attempting to document the utility of standardized anxiety questionnaires among older adults evaluated the psychometric properties of the Spielberger State-Trait Anxiety Inventory (STAI; Spielberger et al., 1970). In an initial report, the STAI was administered along with the STAI for Children (STAIC; Spielberger, Edwards, Lushene, Montouri, & Platzek, 1973) to a group of adults, ages 55–87, who were participating in a mental health gerontology program. (Patterson, O'Sullivan, & Spielberger, 1980). The STAIC was included given that its simpler verbage and format were thought to be appropriate for some individuals who experienced difficulty with the STAI. Results provided support for the convergent and divergent validity of both instruments. Given that the STAI and STAIC were significantly and positively correlated, the authors recommended that these instruments could be used as parallel forms for the evaluation of anxiety in older adults (Patterson et al., 1980). However, the small and heterogeneous nature of the sample, as well as the inclusion of some participants who were heavily medicated, limits these conclusions (Hersen & van Hasselt, 1992). In addition, variations in item content and format between the STAI and STAIC limit the potential utility of these measures as parallel forms.

In another study with the STAI, the trait scale was administered to two large samples of geriatric inpatients and community residents (Himmelfarb & Murrell, 1983). Adequate internal consistency was evident in both groups, and convergent validity was demonstrated by significant correlations between the STAI trait scale and other measures of negative affect and life satisfaction. STAI scores were significantly different between the inpatient and community groups, even when variance accounted for by age, gender, health, and marital status was removed. However, the most potent predictors of group membership were measures of depression and life satisfaction, with the STAI trait scale providing no additional contribution to the differentiation of groups. The implications of these findings for the utility of the STAI in discriminating anxious and nonanxious elders are limited, however, given that the patient group was not selected to represent individuals with primary anxiety difficulties. In fact, both groups were heterogeneous and standardized diagnostic procedures were not used.

More recently, the authors and colleagues investigated the psychometric properties of the STAI in 50 older adults with GAD and a control group of 94 older adults with no diagnosable psychiatric difficulties (Stanley et al., 1996). Participants in both groups were identified based on the ADIS-R, and data provided support for the internal consistency of both STAI subscales in each group. Test–retest reliability, assessed in a subset of the normal control group, was strong for the trait subscale, but only moderate for the state subscale, as expected. Convergent validity for the STAI was evident in the control group, but mixed results in this regard were apparent among participants with GAD, suggesting some degree of independence among different types of anxiety symptoms within this group. In general, although methodological limitations of these reports preclude definitive conclusions, the STAI appears to be a potentially useful measure of anxiety among elderly people.

Other self-report measures that have been examined in samples of older adults include Hopkins Symptom Checklist (SCL-90, Derogatis et al., 1973), the Self-Rating Anxiety Scale (SAS, Zung, 1971), the Penn State Worry Questionnaire (PSWQ, Meyer, Miller, Metzger, & Borkovec, 1990), the Padua Inventory (PI, Sanavio, 1988), and the Fear Questionnaire (FQ, Marks & Mathews, 1979). Data addressing the SCL-90 and SAS are limited to descriptive reports of SCL-90 scores in a group of 178 elderly hospital patients (Magni & De Leo, 1984) and SAS scores for 92 older adults from the community (Zung, 1980). Thus, conclusions about the utility of these measures are tentative at best.

More thorough psychometric data addressing the potential use of the PSWQ, PI, and FQ are available from data collected by the authors and colleagues in their study of the nature and treatment of GAD among the elderly (Beck et al., 1995; Stanley et al., 1996). The PSWQ is a 16-item scale designed to evaluate the tendency to worry, irrespective of worry content; the PI is a 60-item questionnaire that evaluates obsessive-compulsive symptoms; and the FQ is a 15-item instrument measuring severity of specific fears. All of these instruments have strong psychometric properties when used to assess various aspects of anxiety among younger adults. One purpose of the authors' work was to assess whether psychometric characteristics of these measures would be similar in older adult groups. As noted previously, two samples of older adults were selected via diagnostic interview for this set of studies, one characterized by the presence of GAD and the other with no diagnosable psychiatric complaints. Participants with GAD were administered the questionnaires as part of a larger pretreatment battery; normal control participants received financial compensation for completion of the instruments.

In general, results supported the utility of the PSWQ and the PI (Beck et al., 1995; Stanley et al., 1996). In particular, estimates of internal consistency were adequate for both measures in each participant group, with the exception of the PI behavior control subscale in the normal control group. This subscale measures fears of losing control over motor behavior, a symptom that was only seldom endorsed in this group. Thus, a restricted range of scores in the control group probably created subscale instability in this instance. Adequate convergent validity was demonstrated for the PSWQ and the PI in both participant groups. Factor analysis of the PSWQ suggested a two-factor structure in both groups, with one factor assessing a tendency to worry and a second factor estimating the absence of worry (Beck et al., 1995). Similar analyses were not conducted for the PI given inadequate sample sizes (Stanley et al., 1996). The test–retest reliability of the PI, assessed in a subsample of the normal control group, demonstrated that the total score and scores on the contamination and behavior control subscales assessed stable, trait-like features ($r > 0.70$), while scores on the mental control and checking subscales ($r = 0.61$, 0.64) estimated more statelike clinical symptoms. In general, however, these initial data supported the utility of the PSWQ and the PI in anxious and nonanxious elders.

On the other hand, data for the FQ suggested that alternative measures of specific fears should be investigated for use with the elderly (Stanley et al., 1996). In particular, estimates of internal consistency were mixed in both groups of participants, a finding that is unexpected given the construction of the FQ and the stable nature of phobic fear. Test–retest reliablility (as above, assessed in a subgroup of the normal control sample) was poor and suggested the possibility of regression to the mean over a two to four week interval. In addition, the FQ failed to correlate significantly with other measures of anxiety, indicating poor convergent validity. Furthermore, item content was not selected to probe specific fears of particular relevance for the elderly. As such, the FQ does not appear to be a viable measure of specific fears among older adults.

As an alternative to examining the psychometric properties of measures that are well-established for younger groups, at least one measure of anxiety has been developed specifically with older adults in mind. Wisocki and colleagues created the Worry Scale (WS, Wisocki, Handen, & Morse, 1986), a 35-item questionnaire, to assess worries of particular significance for older adults, namely, concerns about financial, health, and social issues. Initial normative data and support for concurrent validity of the WS were provided by Wisocki and colleagues (Powers et al., 1992; Wisocki, 1988; Wisocki et al., 1986). In the authors' own work described above (Stanley et al., 1996), the WS also demonstrated good internal consistency and generally adequate test–retest reliability. The only exception to the latter was poor stability for the health subscale, potentially due to the sensitivity of this measure to real physical complaints which may vary over time. Finally, the WS demonstrated adequate convergent validity. In this regard, the measure seems to hold promise for the evaluation of worry in older adults.

7.08.5.3 Behavioral Observation

Behavioral observation has been described as an essential component for the evaluation of fear and anxiety (Last & Hersen, 1988). Despite the availability of various systematic schemes for behavioral observation of anxiety among younger adults (Nietzel et al., 1988), only case studies have provided any such approaches for the assessment of anxiety in elderly people. In one report, Turner, Hersen, Bellack, and Wells (1979) described a behavioral observation scheme for monitoring ritualistic handwashing and associated symptoms in a 66-year-old woman with OCD. Behavioral observations were made on an inpatient unit by research assistants three times a day, with measures estimating mean time spent washing and percentage of observation periods in which washing occurred. In another case report, Junginger and Ditto (1984) measured the frequency and nature of checking behaviors in a 65-year-old woman receiving inpatient psychiatric treatment for OCD and major depression. However, no standardized approaches for the behavioral observation of anxiety symptoms in older adults have been developed, and the need for further research in this area has been highlighted (Hersen et al., 1993). In particular, reliable systems for the observation of behavior in clinical and naturalistic settings are needed.

7.08.5.4 Psychophysiological Assessment

As noted by Hersen et al., (1993) psychophysiological responses to stress may be expected to change with increasing age. In particular, changes in cardiovascular variables over the life span have been reported, with older anxious adults showing increased cardiovascular reactivity (Hersen et al., 1993). However, hypothesized age-related changes in physiological functioning are not well-documented, and psychophysiological assessment of anxiety is an issue that has yet to be addressed within the older adult literature. Of interest here will be comparisons of baseline arousal in older anxious and nonanxious adults, as well as potentially differential responses to stressors and feared stimuli (Hersen et al., 1993). In general, the inclusion of physiological measures should be considered carefully, as the use of psychophysiological assessment has not always yielded useful information about the nature and treatment of anxiety (Barlow & Maser, 1984).

7.08.6 TREATMENT

7.08.6.1 Utilization of Services

It is well-documented that mental health resources are severely underutilized by the elderly (Arean, 1993; Lasoski, 1986; Smyer & Gatz, 1995). Overall, it appears that only about one-third of older adults in need of mental health care actually see a mental health professional; the other two-thirds either receive no care or obtain services from primary care physicians (Smyer & Gatz, 1995). Low rates of referral for psychiatric or psychological services have also been reported among primary care physicians who treat older adults (Smyer & Gatz, 1995).

More specific data relevant to treatment of anxiety disorders among elderly people are similarly disconcerting. For example, Blazer et al. (1991) reported that only 38% of older adults with GAD indicated use of outpatient mental health services during the year before they were interviewed. Even lower rates of mental health care were reported in the ECA study, with less than 1% of older adults with agoraphobia and only 2% of those with other phobias indicating that they had sought sevices from a mental health professional during the previous six months (Thompson et al., 1988). A larger percentage of elders with phobic disorders reportedly had sought help for mental health problems from other sources (e.g., clergy, general physicians), but rates of help-seeking overall were still significantly lower among older adults than in the group of adults between the ages of 25 and 64 (Thompson et al., 1988).

A number of explanations have been posited to account for these low rates of mental health care among the elderly. First, older adults themselves are reluctant to define their difficulties in terms of psychological problems, and they are often hesitant to seek social services of any kind (Lasoski, 1986; McCarthy, Katz, & Foa, 1991). Additionally, older adults may be fearful of the implications of their symptoms and perceive that mental health services will be of no real benefit (Smyer & Gatz, 1995). Problems in daily living, such as transportation difficulties and financial limitations, may also contribute to low rates of mental health care services sought by the elderly. Second, problems within the health care community can also be identified. For example, relatively few treatment providers with adequate training are available to provide services to older adults, and outreach and/or publicity of services are often limited (Lasoski, 1986; McCarthy et al., 1991). In addition, negative attitudes about the aging process, and perceptions that the signs and symptoms of anxiety are due merely to normal aging and are therefore irreversible, may create a tendency to ignore or misdiagnose mental health problems among elderly people (Smyer & Gatz, 1995; Weiss, 1994).

Despite the low rates of services sought by elders, a growing body of literature has begun to address the efficacy of both pharmacological and psychosocial interventions for anxiety problems. Much of this literature is limited by serious methodological problems, but growing interest in the nature and treatment of anxiety in this population suggests that notable advances in the area will be made in the early years of the twenty-first century.

7.08.6.2 Pharmacological Interventions

Given that much of the mental health care sought by older adults is provided through primary care physicians, (Blazer et al., 1991; McCarthy et al., 1991; Smyer & Gatz, 1995), it is not surprising that treatment for anxiety most often involves the use of medication, usually the benzodiazepines. Among older adults, use of the benzodiazepines occurs far more frequently than in younger samples. (Blazer et al., 1991; Salzman, 1991). Clinical recommendations for the use of benzodiazepines among the elderly, however, are not based on well-controlled treatment trials. Although some data have suggested the utility of a variety of benzodiazepines, all studies available in the late 1990s are limited by serious methodological problems (Cohn, 1984; Salzman, 1991). For example, patients generally were not selected via any systematic diagnostic procedure. As such, generalization of conclusions is limited by the heterogeneous nature of the study samples. Relatedly, patients with anxiety symptoms or disorders were not well-differentiated from participants with anxiety secondary to depression or medical illness, nor was anxiety always distinguished from behavioral agitation associated with cognitive impairment. Finally, the studies varied significantly with regard to outcome criteria, treatment duration, and the use of concommitant medications.

Despite the severe methodological limitations of the outcome literature, recommendations regarding the use of benzodiazepines are fairly consistent, with the suggestion that compounds with short half-lives be used for as brief a duration as possible (Fernandez, Levy, Lachar, & Small, 1995; Markovitz, 1993; Weiss, 1994). Benzodiazepine doses recommended for older adults are also generally lower than those prescribed for younger adults given altered pharmacokinetic properties that lead to increased therapeutic and adverse effects (Lader, 1986). In general, strategies for benzodiazepine use have been recommended based on studies addressing the utility of anxiolytics in younger adults, the pharmacokinetic properties of the medications in nonsymptomatic older samples, and anxiolytic toxicity among the elderly (Salzman, 1991).

Problems with the use of benzodiazepines among the elderly are noted routinely, with some authors even suggesting that psychosocial interventions be considered in conjunction with a pharmacological approach (Lader, 1986; Wengel, Burke, Ranno, & Roccaforte, 1993). Given the increased sensitivity to these medications generally observed in older adults, adverse events of particular concern include cognitive impairment (e.g., amnesia, delayed recall, confusion, disorientation), agitation (which may exacerbate anxiety symptoms), respiratory depression, and psychomotor slowing which may lead to increased rates of falls and hip fractures (Weiss, 1994; Wengel et al., 1993). Increased rates of medical illness among older adults also raise the issue of potentially harmful drug interactions.

Thus, although the benzodiazepines are recommended routinely for the treatment of anxiety among the elderly, the literature upon which clinical recommendations are made is quite limited and the increased risk of potentially serious adverse effects warrants caution in using these compounds. Other medications may be considered as alternatives, although again, the relevant treatment literature is quite limited in scope. For example, clinical recommendations have suggested the use of buspirone as an alternative to benzodiazepines for the treatment of anxiety in older adults, given its apparently equivalent therapeutic effects without significant potential for serious adverse effects, tolerance, or withdrawal (Weiss, 1994). However, the treatment literature has documented the efficacy of buspirone only in a number of open trials (Salzman, 1991), and other more well-controlled research will need to substantiate these clinical claims. Likewise, clinical recommendations have suggested the potential utility of beta blockers and antidepressants for the treatment of anxiety among the elderly (Fernandez et al., 1995; Markowitz, 1993), but these suggestions are based only on literature documenting the efficacy of these compounds in younger anxious adults. Future studies will need to address directly the efficacy of these medications in well-diagnosed samples of older adults.

7.08.6.3 Psychosocial Interventions

The consideration of psychosocial treatments as an adjunct or alternative to the use of medication for the treatment of psychiatric disturbance in the elderly is an important but understudied issue. In general, it has been noted that many of the basic clinical skills useful for the treatment of younger adults are appropriate for therapeutic work with elderly people (Taylor & Hartman Stein, 1995). However, variations in technique or approach are often necessary to address the special needs of an older population. In particular, the importance of assessing both patients' and family members' beliefs about mental health problems and their treatment has been noted, as well as the need to educate both of these groups about the nature of psychosocial treatment (Arean, 1993; Gallagher-Thompson & Thompson, 1995). In addition, differences in learning strategies and sensory function between older and younger adults need to be considered, with recommendations suggesting that therapy with elderly people occur at a slower pace and include increased use of visual and written aids (Gallagher-Thompson & Thompson, 1995; Smyer, Zarit, & Qualls, 1990) Group approaches have also been recommended to reduce cost and increase social contact and support (Arean, 1993; Yost, Beutles, Corbishley & Allender, 1986) and regular contact with the patient's primary care physician has been suggested (Arean, 1993; Gallagher-Thompson & Thompson, 1995).

In addition to these general recommendations, the older adult literature includes a number of well-controlled clinical trials documenting the specific efficacy of psychosocial treatment for depression (Arean et al., 1993, Thompson, Gallagher, & Breckenridge, 1987; see Chapter 9, this volume). However, much less attention has been paid to the development of psychosocial treatments for anxiety among the elderly, and the need for further work in this domain has been noted repeatedly (Cartensen, 1988, Hersen & Van Hasselt, 1992; Niederehe, Cooley, & Teri, 1995). Given the voluminous body of work documenting the efficacy of cognitive behavioral treatments for anxiety in younger adults (Barlow, 1988, Beidel & Turner, 1991) it is not surprising that initial attempts to establish psychosocial interventions for older adults have focused on these approaches. However, the available literature is limited largely to clinical case studies and group comparisons with community volunteers reporting general anxiety symptoms.

A relatively large group of case studies has examined the use of cognitive behavioral interventions that have been demonstrated efficacious in younger adult samples. In particular, interventions including exposure, response prevention, relaxation training, and cognitive restructuring have been utilized to treat elderly patients with a range of anxiety disorders, such as GAD (King & Barrowclough, 1991), PD (Swales, Solfvin, & Sheikh, 1996), specific phobias (Thyer, 1981), and OCD (Calamari, Faber, Hitsman; & Poppe, 1994; Carmin, Ownby, & Pollard, 1995; Rowan, Holborn, Walker, & Siddiqui, 1984) The efficacy of cognitive behavioral interventions with a wide range of anxiety difficulties suggests strongly the utility of further work in this domain. However, conclusions from these reports alone can be drawn only cautiously given the absence of appropriate methodological controls.

Other more controlled research has examined the utility of cognitive behavior therapy (CBT) for the reduction of anxiety symptoms in nonclinical, community volunteers. In one of the earliest of these studies, DeBerry (1982) compared the effects of two versions of relaxation training with an attention control condition in a sample of older adult women volunteers with various anxiety complaints. Results demonstrated the efficacy of both relaxation training conditions, relative to the control group, on measures of state and trait anxiety. In a similar comparison, 24 older community volunteers reporting both anxiety and depressive symptoms were assigned randomly to receive relaxation training, CBT for depression, or a nonspecific control condition (Sallis, Lichstein, Clarkson, Stalgaitis, & Campbell, 1983). Results revealed significant reductions in both anxiety and depression across all treatment groups.

In a subsequent study, another community group of older adults with anxiety symptoms received relaxation training, cognitive restructuring, or an attention control condition

(DeBerry, Davis, & Reinhard, 1989). In this comparison, only relaxation training was effective in decreasing state anxiety, although trait anxiety and symptoms of depression were reduced at post-treatment across all conditions. In a more recent comparison with a larger group of community volunteers, the effects of progressive and imaginal relaxation training were examined relative to a wait-list control condition (Scogin, Rickard, Keith, Wilson & McElreath, 1992). Both types of relaxation training led to significant decreases in anxiety, relative to the control group, and treatment gains were maintained at one year follow-up (Rickard, Scogin, & Keith, 1994).

These studies together suggest that CBT, in particular relaxation training, may be beneficial for the treatment of anxiety in older adults. However, the sample sizes in many of these reports were small, and participants were not evaluated for the presence of diagnosable anxiety pathology. Only recently has a preliminary empirical trial examined the utility of cognitive behavioral interventions in older adults with well-diagnosed anxiety disorders.

In this comparison, the efficacy of CBT was evaluated relative to a nondirective supportive psychotherapy condition in a sample of 48 older adults with a principal diagnosis of GAD (Stanley et al., 1996). Treatment in both conditions was administered in a small group format over a period of 14 weeks. CBT followed the treatment approach described by Craske, Barlow, and O'Leary (1992) and included relaxation training, cognitive therapy, and systematic exposure to worry-related situations. Modifications to this treatment model included the use of a group format, as well as increased use of visual and written aids (e.g., outlines of session topics, written instructions for homework and monitoring assignments). Results revealed significant decreases at post-treatment for participants in both treatment conditions on self-report and clinician-rated measures of worry, anxiety, and depression. CBT also produced significant improvement in severity of specific fears, although psychometric limitations of the FQ restrict the conclusions which can be made in this regard. Results of a six month follow-up evaluation revealed maintenance of treatment gains in both conditions, with some continued improvement in severity of worry over time. Correlational analyses suggested that pretreatment severity of anxiety and ratings of treatment credibility were significant predictors of outcome. As such, these data further substantiate the potential utility of CBT (and other psychosocial interventions) for the treatment of anxiety among the elderly. However, study conclusions are limited given the omission of a no-treatment or wait-list control condition. Future research will need to replicate these findings in a better controlled design.

7.08.7 SUMMARY

Anxiety disorders are the most prevalent of all psychiatric conditions among elderly people and they pose a major public health problem for this population. Nevertheless, research addressing the prevalence, nature, and treatment of these conditions among elderly is in an infancy stage. No data are available regarding the prevalence of PTSD among the elderly, and the clinical picture of many anxiety syndromes in this segment of the population is as yet unclear. Much of the available literature is limited due to the significant dearth of well-established measures of anxiety among elderly people and the failure to utilize strict diagnostic criteria in the selection of clinical groups. The utility of measures already established for use with younger groups deserves further study in this regard. Additionally, new instruments that target specifically the unique features of anxiety among elderly people will need to be developed. In this domain, the focus will need to be not only on self-report instruments, but also on systematic methods of behavioral observation, clinician rating, and psychophysiological assessment. The development of more refined measures will assist in clarifying the difference between anxiety disorders and other syndromes that may produce similar clinical features, including depression, various medical conditions, sleep disorders, and cognitive dysfunction.

With the establishment of psychometrically sound assessment tools, more sophisticated treatment research should emerge. Pharmacological trials might focus on the relative efficacy of various antianxiety and antidepressant medication, and psychosocial treatment comparisons should evaluate further the utility of CBT and alternative interventions in well-diagnosed groups of older patients. Emphasis in all of these investigations should be given not only to the effects of treatment on symptom reduction, but also to the transfer effects of the interventions on associated clinical features, physical health, and social functioning. Durability of treatment response following both pharmacological and psychosocial treatments will need to be examined with well-controlled evaluations over long-term follow-up intervals, and findings will need to be generalized to patients recruited from a variety of settings (e.g., psychiatric clinics, primary care clinics) and from diverse sociocultural backgrounds.

7.08.8 REFERENCES

Alexopoulos, G. S. (1991). Anxiety and depression in the elderly. In C. Salzman & B. D. Lebowitz (Eds.), *Anxiety in the elderly: Treatment and research* (pp. 63–77). New York: Springer.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Arean, P. A. (1993). Cognitive behavioral therapy with older adults. *The Behavior Therapist*, *16*, 236–239.

Arean, P. A., Perri, M. G., Nezu, A. M., Schein, R. L., Christopher, F., & Joseph, T. X. (1993). Comparative effectiveness of social problem-solving therapy and reminiscence therapy as treatments for depression in older adults. *Journal of Consulting and Clinical Psychology, 61*, 1003–1010.

Barlow, D. H. (1988). *Anxiety and its disorders: The nature and treatment of anxiety and panic*. New York: Guilford.

Barlow, D. H., & Maser, J. D. (1984). Psychopathology in anxiety disorders. *Journal of Behavioral Assessment*, *6*, 331–348.

Barrios, B. A., Hartmann, D. B., & Shigetomi, C. (1981). Fears and anxieties in children. In E. J. Mash & L. G. Terdal (Eds.), *Behavioral assessment of childhood disorders* (pp. 259–304). New York: Guilford Press.

Beck, J. G., & Stanley, M. A. (1997). Anxiety disorders in the elderly: The emerging role of behavior therapy. *Behavior Therapy, 28,* 83–100

Beck, J. G., Stanley, M. A., & Zebb, B. J. (1995). Psychometric properties of the Penn State Worry Questionnaire in older adults. *Journal of Clinical Geropsychology*, *1*, 33–42.

Beck, J. G., Stanley, M. A., & Zebb, B. J. (1996a). Characteristics of generalized anxiety disorder in older adults: A descriptive study. *Behaviour Research and Therapy*, *34*, 225–234.

Beck, J. G., Stanley, M. A., & Zebb, B. J. (1996b, August). *How do the Hamilton Scales perform with geriatric GAD patients*? Presented at the Annual Convention for the American Psychological Association, Toronto, ON, Canada.

Beidel, D. C., & Turner, S. M. (1991). Anxiety disorders. In M. Hersen & S. M. Turner (Eds.), *Adult psychopathology and diagnosis* (pp. 226–278). New York: Wiley.

Ben-Arie, O., Swartz, L., & Dickman, B. J. (1987). Depression in the elderly living in the community: Its presentation and features. *British Journal of Psychiatry*, *150*, 169–174.

Bland, R. C., Newman, S. C., & Orn, H. (1988). Prevalence of psychiatric disorders in the elderly in Edmonton. *Acta Psychiatrica Scandanavia (Suppl.)*, *338*, 57–63.

Blazer, D., George, L. K., & Hughes, D. (1991). The epidemiology of anxiety disorders: An age comparison. In C.. Salzman & B. D. Lebowitz (Eds.), *Anxiety in the elderly: Treatment and research* (pp. 17–30). New York: Springer.

Blazer, D., Hughes, D., & George, L. K. (1987). Stressful life events and the onset of a generalized anxiety syndrome. *American Journal of Psychiatry*, *144*, 1178–1183.

Brown, T. A., DiNardo, P. A., & Barlow, D. H. (1994). *Anxiety Disorders Interview Schedule for DSM-IV*. Albany, NY: Graywinds.

Bliwise, N. G., Bliwise, D. L., & Dement, W. C. (1985). Age and psychopathology in insomnia. *Clinical Gerontologist*, *4*, 3–9.

Bruce, M. L., & McNamara, R. (1992). Psychiatric status among the homebound elderly: An epidemiologic perspective. *Journal of the American Geriatrics Society*, *40*, 561–566.

Burt, D. B., Zembar, M. J., & Niederehe, G. (1995). Depression and memory impairment: A meta-analysis of the association, its pattern, and specificity. *Psychological Bulletin*, *117*, 285–305.

Burvill, P. W., Johnson, G. A., Jamrozik, K. D., Anderson, C. S., Stewart-Wynne, E. G., & Chakera, T. M. H. (1995). Anxiety disorders after stroke: Results from the Perth Community Stroke Study. *British Journal of Psychiatry*, *166*, 328–332.

Calamari, J. E., Faber, S. D., Hitsman, B. L., & Poppe, C. J. (1994). Treatment of obsessive compulsive disorder in the elderly: A review and case example. *Journal of Behavior Therapy and Experimental Psychiatry, 25*, 95–104.

Carmin, C. N., Ownby, R. L., & Pollard, C. A. (1995, November). *Effects of cognitive behavioral treatment of obsessive compulsive disorder in geriatric versus younger adult patients.* Presented at the Annual Convention of the Association for Advancement of Behavior Therapy, Washington, DC.

Carstensen, L. L. (1988). The emerging field of behavioral gerontology. *Behavior Therapy*, *19*, 259–281.

Cloninger, C. R. (1990). Comorbidity of anxiety disorders. *Journal of Psychopharmacology*, *10*, 43–46.

Cohen, G. D. (1991). Anxiety and general medical disorders. In C. Salzman & B. D. Lebowitz (Eds.), *Anxiety in the elderly: Treatment and research* (pp. 47–62). New York: Springer.

Cohn, J. B. (1984). Double-blind safety and efficacy comparison of alprazolam and placebo in the treatment of anxiety in geriatric patients. *Current Therapeutic Research*, *35*, 100–112.

Copp, J. E., Schwiderski, U. E., & Robinson, D. S. (1990). Symptom comorbidity in anxiety and depressive disorders. *Journal of Clinical Psychopharmacology*, *10*, 52–60.

Craske, M. G., Barlow, D. H., & O'Leary, T. (1992). *Mastery of your anxiety and worry*. Albany, NY: Graywind.

DeBerry, S. (1982). The effects of meditation–relaxation on anxiety and depression in a geriatric population. *Psychotherapy: Theory, Research, and Practice, 19*, 512–521.

DeBerry, S., Davis, S., & Reinhard, K. E. (1989). A comparison of meditation–relaxation and cognitive behavioral techniques for reducing anxiety and depression in a geriatric population. *Journal of Geriatric Psychiatry, 22*, 231–247.

Deptula, D., Singh, R., & Pomara, N. (1993). Aging, emotional states, and memory. *American Journal of Psychiatry*, *150*, 429–434.

DiNardo, P. A., Moras, K., Barlow, D. H., Rapee, R. M., & Brown, T. A. (1993). Reliability of DSM-III-R anxiety disorder categories: Using the Anxiety Disorders Interview Schedule—Revised (ADIS-R). *Archives of General Psychiatry*, *50*, 251–256.

Downton, J. H., & Andrews, K. (1990). Postural disturbance and psychological symptoms amongst elderly people living at home. *International Journal of Geriatric Psychiatry*, *5*, 93–98.

Eaton, W. M., Kramer, M., Anthony, J. C., Dryman, A., Shapiro, S., & Locke, B. Z. (1989). The incidence of specific DIS/DSM-III mental disorders: Data from the NIMH Epidemiologic Catchment Area Program. *Acta Psychiatrica Scandanavia*, *79*, 163–178.

Engle-Friedman, M., & Bootzin, R. (1991). Insomnia as a problem for the elderly. In P. A. Wisocki (Ed.), *Handbook of clinical behavior therapy with the elderly client: applied clinical psychology* (pp. 273–298). New York: Plenum.

Feinson, M. C., & Thoits, P. A. (1986). The distribution of distress among elders. *Journal of Gerontology*, *41*, 225–233.

Feldman, L. A. (1993). Distinguishing depression and anxiety in self-report: Evidence from confirmatory factor analysis on nonclinical and clinical samples. *Journal of Consulting and Clinical Psychology*, *61*, 631–638.

Fernandez, F., Levy, J. K., Lachar, B. L., & Small, G. W. (1995). The management of depression and anxiety in the elderly. *Journal of Clinical Psychiatry*, *56*, 20–29.

Flint, A. J. (1994). Epidemiology and comorbidity of anxiety disorders in the elderly. *American Journal of Psychiatry*, *151*, 640–649.

Frazer, D. W. (1995). The medical issues in geropsychology training and practice. In B. G. Knight, L. Teri, P. Wohlford, & J. Santos (Eds.), *Mental health services for older adults: Implications for training and practice in geropsychology* (pp. 63–71). Washington, DC: American Psychological Association.

Friedman, L., Bliwise, D. L., Yesavage, J. A., & Salom, S. R. (1991). A preliminary study comparing sleep restriction and relaxation treatments for insomnia in older adults. *Journal of Gerontology*, *46*, 1–8.

Friedman, L., Brooks, J. O., Bliwise, D. L., & Yesavage, J. A. (1993). Insomnia in older adults: Relations to depression and insomnia. *American Journal of Geriatric Psychiatry*, *1*, 153–159.

Gallagher-Thompson, D., & Thompson, L. W. (1995). Efficacy of psychotherapeutic interventions with older adults. *The Clinical Psychologist, 48*, 24–30.

Ganzini, L., McFarland, B. H., & Culter, D. (1990). Prevalence of mental disorders after catastrophic financial loss. *Journal of Nervous and Mental Disease*, *178*, 680–685.

Gurian, B. S., & Miner, J. H. (1991) Clinical presentation of anxiety in the elderly. In C. Salzman & B. D. Lebowitz (Eds.), *Anxiety in the elderly: Treatment and research* (pp. 31–44). New York: Springer.

Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology*, *32*, 50–55.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, & Psychiatry*, *23*, 56–62.

Helzer, J. E., Robins, L. N., & McEvoy, L. (1987). Posttraumatic stress disorder in the general population: Findings of the Epidemiologic Catchment Area survey. *The New England Journal of Medicine*, *317*, 1630–1634.

Hersen, M., & Van Hasselt, V. B. (1992). Behavioral assessment and treatment of anxiety in the elderly. *Clinical Psychology Review*, *12*, 619–640.

Hersen, M., Van Hasselt, V. B., & Goreczny, A. J. (1993). Behavioral assessment of anxiety in older adults: Some comments. *Behavior Modification*, *17*, 99–112.

Hill, R. D., & Vandervoort, D. (1992). The effects of state anxiety on recall performance in older learners. *Educational Gerontology*, *18*, 597–605.

Himmelfarb, S., & Murrell, S. A. (1983). Reliability and validity of five mental health scales in older persons. *Journal of Gerontology*, *38*, 333–339.

Himmelfarb, S., & Murrell, S. A. (1984). The prevalence and correlates of anxiety symptoms in older adults. *The Journal of Psychology*, *116*, 159–167.

Junginger, J., & Ditto, B. (1984). Multitreatment of obsessive-compulsive checking in a geriatric patient. *Behavior Modification*, *8*, 379–390.

Katz, I. R. (1996). On the inseparability of mental and physical health in aged persons: Lessons from depression and medical comorbidity. *The American Journal of Geriatric Psychiatry*, *4*, 1–16.

Lader, M. (1986). The use of hypnotics and anxiolytics in the elderly. *International Clinical Psychopharmacology*, *1*, 273–283.

King, P., & Barrowclough, C. (1991). A clinical pilot study of cognitive-behavioral therapy for anxiety disorders in the elderly. *Behavioural Psychotherapy, 19*, 337–345.

Kluznick, J. C., Speed, N., Van Valkenberg, C., & Magraw, R. (1986). Forty-year follow-up of United States prisoners of war. *American Journal of Psychiatry*, *143*, 1443–1446.

Kuch, K., & Cox, B. J. (1992). Symptoms of PTSD in 124 survivors of the Holocaust. *American Journal of Psychiatry*, *149*, 337–340.

Larkin, B. A., Copeland, J. R. M., Dewey, M. E., Davidson, I. A., Saunders, P. A., Sharma, V. K., McWilliam, C., & Sullivan, C. (1992). The natural history of neurotic disorder in an elderly urban population: Findings from the Liverpool Longitudinal Study of Continuing Health in the Community. *British Journal of Psychiatry*, *160*, 681–686.

LaRue, A., & D'Elia, L. F. (1985). Anxiety and problem solving in middle-aged and elderly adults. *Experimental Aging Research*, *11*, 215–220.

Lasoski, M. C. (1986). Reasons for low utilization of mental health services by the elderly. In T. L. Brink (Ed.), *Clinical gerontology: A guide to assessment and intervention* (pp. 1–18). New York: Haworth.

Last, C. G., & Hersen, M. (Eds.) (1988). *Handbook of anxiety disorders*. New York: Pergamon.

Lawton, M. P., Kleban, M. H., & Dean, J. (1993). Affect and age: Cross-sectional comparisons of structure and prevalence. *Psychology and Aging*, *8*, 165–175.

Liddell, A., Locker, D., & Burman, D. (1991). Self-reported fears (FSS-II) of subjects aged 50 years and over. *Behaviour Research and Therapy*, *29*, 105–112.

Lindesay, J. (1991). Phobic disorders in the elderly. *British Journal of Psychiatry*, *159*, 531–541.

Lindesay, J., Briggs, K., & Murphy, E. (1989). The Guy's/Age Concern Survey: Prevalence rates of cognitive impairment, depression and anxiety in an urban elderly community. *British Journal of Psychiatry*, *155*, 317–329.

Luchins, D. J., & Rose, R. P. (1989). Late-life onset of panic disorder with agoraphobia in three patients. *American Journal of Psychiatry*, *146*, 920–921.

Magni, G., & De Leo, D. (1984). Anxiety and depression in geriatric and adult medical inpatients: A comparison. *Psychological Reports*, *55*, 607–612.

Markovitz, P. J. (1993). Treatment of anxiety in the elderly. *Journal of Clinical Psychiatry*, *54*, 64–68.

Marks, I. M., & Gelder, M. G. (1966). Different ages of onset in varieties of phobia. *American Journal of Psychiatry*, *123*, 218–221.

Marks, I. M., & Mathews, A. M. (1979). Brief standardized self-rating for phobic patients. *Behaviour Research and Therapy*, *17*, 263–267.

McCarthy, P. R., Katz, I. R., & Foa, E. B. (1991). Cognitive-behavioral treatment of anxiety in the elderly: A proposed model. In C. Salzman & B. D. Lebowitz (Eds.), *Anxiety in the elderly: Treatment and research* (pp. 197–214). New York: Springer.

McNally, R. J. (1994). *Panic disorder: A critical analysis*. New York: Guilford Press.

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, *28*, 487–495.

Moras, K., DiNardo, P. A., & Barlow, D. H. (1992). Distinguishing anxiety and depression: Re-examination of the reconstructed Hamilton Scales. *Psychological Assessment*, *4*, 224–227.

Morin, C. M., Colecchi, C. A., Ling, W. D., & Sood, R. K. (1995). Cognitive behavior therapy to facilitate benzodiazepine discontinuation among hypnotic-dependent patients with insomnia. *Behavior Therapy*, *26*, 733–745.

Morin, C. M., & Gramling, S. E. (1989). Sleep patterns and aging: Comparison of older adults with and without insomnia complaints. *Psychology and Aging*, *4*, 290–294.

Morin, C. M., Kowatch, R. A., Barry, T., & Walton, E. (1993). Cognitive-behavior therapy for late-life insomnia. *Journal of Consulting and Clinical Psychology*, *61*, 137–136.

Niederehe, G., Cooley, S. G., & Teri, L. (1995). Research and training in clinical geropsychology: Advances and current opportunities. *The Clinical Psychologist, 48*, 37–44.

Nietzel, M. T., Bernstein, D. A., & Russell, R. L. (1988). Assessment of anxiety and fear. In A. S. Bellack & M. Hersen (Eds.), *Behavioral assessment: A practical handbook*, (3rd ed., pp. 280–312). New York: Pergamon.

Norris, F. H., & Murrell, S. A. (1988). Prior experience as a moderator of disaster impact on anxiety symptoms in older adults. *American Journal of Community Psychology*, *16*, 665–683.

Ost, L. G. (1987). Age of onset in different phobias. *Journal of Abnormal Psychology*, *96*, 223–229.

Oxman, T. E., Barrett, J. E., Barrett, J., & Gerber, P. (1987). Psychiatric symptoms in the elderly in a primary care practice. *General Hospital Psychiatry*, *9*, 167–173.

Parmelee, P. A., Katz, I. R., & Lawton, M. P. (1992). Anxiety and its association with depression among institutionalized elderly. *American Journal of Geriatric Psychiatry*, *1*, 46–58.

Patterson, R. L., O'Sullivan, M. J., & Spielberger, C. D. (1980). Measurement of state and trait anxiety in elderly mental health clients. *Journal of Behavioral Assessment*, *2*, 89–97.

Person, D. C., & Borkovec, T. D. (1995, August). *Anxiety disorders among the elderly: Patterns and issues*. Paper presented at the annual meeting of the American Psychological Association, New York.

Poon, L. W. (1992). Toward an understanding of cognitive functioning in geriatric depression. *International Psychogeriatrics*, *4*, 241–266.

Powers, C. B., Wisocki, P. A., & Whitbourne, S. K. (1992). Age differences and correlates of worrying in young and elderly adults. *The Gerontologist*, *32*, 82–88.

Raj, B. A., Covea, M. H., & Dagon, E. M. (1993). The clinical characteristics of panic disorder in the elderly: A retrospective study. *Journal of Clinical Psychiatry*, *54*, 150–155.

Rapee, R. M., & Barlow, D. H. (1991). *Chronic anxiety: Generalized anxiety disorder and mixed anxiety-depression*. New York: Guilford Press.

Regier, D. A., Boyd, J. H., Burke, J. D., Rae, D. S., Myers, J. K., Kramer, M., Robins, L. N., George, L. K., Karno, M., & Locke, B. Z. (1988). One-month prevalence of mental disorders in the United States: Based on five epidemiologic catchment area sites. *Archives of General Psychiatry*, *45*, 977–986.

Regier, D. A., Narrow, W. E., & Rae, D. S. (1990). The epidemiology of anxiety disorders: The Epidemiologic Catchment Area (ECA) experience. *Journal of Psychatric Research*, *24*, 3–14.

Reynolds, C. F., Lebowitz, B. D., & Schneider, L. S. (1993). The NIH Consensus Development Conference on the diagnosis and treatment of depression in late life: An overview. *Psychopharmacology Bulletin*, *29*, 83–85.

Rickard, H. C., Scogin, F., & Keith, S. (1994). A one-year follow-up of relaxation training for elders with subjective anxiety. *The Gerontologist, 34*, 121–122.

Riskind, J. H., Beck, A. T., Brown, G. B., & Steer, R. A. (1987). Taking the measure of anxiety and depression: Validity of the reconstructed Hamilton Scales. *Journal of Nervous and Mental Disease*, *175*, 474–479.

Robins, L. N., Helzer, J. E., Weissman, M. M., Orvaschel, H., Gruenberg, E., Burke, J. D., & Regier, D. A. (1984). Lifetime prevalence of specific psychiatric disorders in three sites. *Archives of General Psychiatry*, *41*, 949–958.

Roehrs, T., Lineback, W., Zorick, F., & Roth, T. (1982). Relationship of psychopathology to insomnia in the elderly. *Journal of the American Geriatrics Society*, *30*, 312–315.

Roehrs, T., Zorick, F., Sicklesteel, J., Wittig, R., & Roth, T., (1983). Age-related sleep–wake disorders at a sleep disorder center. *Journal of the American Geriatrics Society*, *31*, 364–370.

Rowan, V. C., Holborn, S. W., Walker, J. R., & Siddiqui, A. R. (1984). A rapid multi-component treatment for an obsessive-compulsive disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 15*, 347–352.

Sallis, J. F., Lichstein, K. L., Clarkson, A. D., Stalgaitis, S., & Campbell, M. (1983). Anxiety and depression management for the elderly. *International Journal of Behavioral Geriatrics, 1*, 3–12.

Salzman, C. (1991). Pharmacologic treatment of the anxious elderly patient. In C. Salzman & B. D. Lebowitz (Eds.), *Anxiety in the elderly: Treatment and research* (pp. 149–173). New York: Springer.

Salzman, C., & Lebowitz, B. D. (1991). *Anxiety in the elderly: Treatment and research*. New York: Springer.

Sanavio, E. (1988). Obsessions and compulsions: The Padua Inventory. *Behaviour Research and Therapy*, *26*, 169–177.

Sanderson, W. C., & Wetzler, S. (1991). Chronic anxiety and generalized anxiety disorder: Issues in comorbidity. In R. M. Rapee & D. H. Barlow (Eds.), *Chronic anxiety: Generalized anxiety disorder and mixed anxiety-depression* (pp. 119–135). New York: Guilford Press.

Scogin, F., Rickard, H. C., Keith, S., Wilson, J., & McElreath, L. (1992). Progressive and imaginal relaxation training for elderly persons with subjective anxiety. *Psychology and Aging, 7*, 419–424.

Segal, D. L., Hersen, M., Van Hasselt, V. B., Kabacoff, R. I., & Roth, L. (1993). Reliability of diagnosis in older psychiatric patients using the Structured Clinical Interview for DSM-III-R. *Journal of Psychopathology and Behavioral Assessment*, *15*, 347–356.

Segal, D. L., Kabacoff, R. I., Hersen, M., Van Hasselt, V. B., & Ryan, C. F. (1995). Update on the reliability of diagnosis in older psychiatric outpatients using the Structured Clinical Interview for DSM-III-R. *Journal of Clinical Geropsychology, 1,* 313–321.

Sheikh, J. I. (1991). Anxiety rating scales for the elderly. In C. Salzman & B. D. Lebowitz (Eds.), *Anxiety in the elderly: Treatment and research* (pp. 251–265). New York: Springer.

Sheikh, J. I., King, R. J., & Taylor, C. B. (1991). Comparative phenomenology of early onset versus later onset panic attacks: A pilot survey. *American Journal of Psychiatry*, *148*, 1231–1233.

Sloane, P. D., Hartman, M., & Mitchell, C. M. (1994). Psychological factors associated with chronic dizziness in patients aged 60 and older. *Journal of the American Geriatric Society*, *42*, 847–852.

Smyer, M. A., & Gatz, M. (1995). The public policy context of mental health care for older adults. *The Clinical Psychologist*, *48*, 31–36.

Smyer, M. A., Zarit, S. H., & Qualls, S. H. (1990). Psychological intervention with the aging individual. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 375–403). San Diego, CA: Academic Press.

Spielberger, C. D., Edwards, C. D., Lushene, R. E., Montouri, J., & Platzek, D. (1973). *STAIC: Preliminary manual for the State-Trait Anxiety Inventory for children*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C., Gorsuch, R., & Lushene, R. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Spitzer, R. L., Williams, J. B. W., Gibbon, M., & First, M. B. (1988). *Structured Clinical Interview for DSM-III-R—Patient Edition (SCID-P 6/1/88)*. New York: Biometrics Research Department, New York State Psychiatric Institute.

Stanley, M. A., Beck, J. G., & Glassco, J. D. (1996).

Generalized anxiety in older adults: Treatment with cognitive-behavioral and supportive approaches. *Behavior Therapy, 27,* 565–581.

Stanley, M. A., & Beck, J. G. (in press). Anxiety disorders. In M. Hersen and V. B. Van Hasselt (Eds.). *Handbook of clinical geropsychology*, New York: Plenum.

Stanley, M. A., Beck, J. G., & Zebb, B. J. (1996). Psychometric properties of four anxiety measures in older adults. *Behaviour Research and Therapy, 34,* 827–838.

Swales, P. J, Solfvin, J. F., & Sheikh, J. I. (1996). Cognitive-behavioral therapy in older panic disorder patients. *The American Journal of Geriatric Psychiatry, 4*, 46–60.

Taylor, G. P., & Hartman-Stein, P. E. (1995). Providing psychological services to older adults: Opportunities for psychologists to sail beyond the sunset. *The Clinical Psychologist*, *48*, 5–11.

Thompson, J. W., Burns, B. J., Bartko, J., Boyd, J. H., Taube, C. A., & Bourdon, K. H. (1988). The use of ambulatory services by persons with and without phobia. *Medical Care*, *26*, 183–198.

Thompson, L. W., Gallagher, D., & Breckenridge, J. S. (1987). Comparative effectiveness of psychotherapies for depressed elders. *Journal of Consulting and Clinical Psychology, 55*, 385–390.

Thyer, B. A. (1981). Prolonged *in vivo* exposure therapy with a 70-year-old woman. *Journal of Behavior Therapy and Experimental Psychiatry, 12*, 69–71.

Turner, S. M., Hersen, M., Bellack, S., & Wells, K. C. (1979). Behavioral treatment of obsessive-compulsive neurosis. *Behaviour Research and Therapy*, *17*, 95–106.

Weiss, K. J. (1994). Management of anxiety and depression syndromes in the elderly. *Journal of Clinical Psychiatry*, *55*, 5–12.

Weissman, M. M., Myers, J. K., Tischler, G. L., Holzer, C. E., Leaf, P. J., Orvaschel, H., & Brody, J. A. (1985). Psychiatric disorders (DSM-III) and cognitive impairment among the elderly in a US urban community. *Acta Psychiatrica Scandanavia*, *71*, 366–379.

Wengel, S. P., Burke, W. J., Ranno, A. E., & Roccaforte, W. H. (1993). Use of benzodiazepines in the elderly. *Psychiatric Annals*, *23*, 325–331.

West, R. L., Boatwright, L. K., & Schleser, R. (1984). The link between memory performance, self-assessment, and affective status. *Experimental Aging Research*, 10, 197–200.

Wisocki, P. A. (1988). Worry as a phenomenon relevant to the elderly. *Behavior Therapy*, *19*, 369–379.

Wisocki, P. A., Handen, B., & Morse, C. K. (1986). The Worry Scale as a measure of anxiety among homebound and community active elderly. *The Behavior Therapist*, *9*, 91–95.

Wong, M. T. H., & Pan, P. C. (1994). Patterns of psychogeriatric referral and attendance at three different settings in Hong Kong. *International Psychogeriatrics*, *6*, 199–208.

Yesavage, J. A., & Jacob, R. (1984). Effects of relaxation and mnemonics on memory, attention and anxiety in the elderly. *Experimental Aging Research*, *10*, 211–214.

Yesavage, J. A., Rose, T. L., & Spiegel, D. (1982). Relaxation training and memory improvement in elderly normals: Correlation of anxiety ratings and recall improvement. *Experimental Aging Research*, *8*, 195–198.

Yost, E. B., Beutler, L. E., Corbishley, M. A., & Allender, J. R. (1986). *Group Cognitive Therapy: A treatment approach for depressed older adults*. New York: Pergamon.

Zung, W. W. K. (1971). A rating instrument for anxiety disorders. *Psychosomatics*, *12*, 371–379.

Zung, W. W. K. (1980). How normal is anxiety? *Current Concepts*. Kalamazoo, MI: Scope.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.09

Mood Disorders in Late Life

AMY FISKE, JULIA E. KASL-GODLEY, and MARGARET GATZ
University of Southern California, Los Angeles, CA, USA

7.09.1 INTRODUCTION 194

7.09.2 DEFINITION AND CLASSIFICATION 194

- *7.09.2.1 Standardized Classification Systems* 194
  - *7.09.2.1.1* DSM-IV 194
  - *7.09.2.1.2* ICD-10 195
  - *7.09.2.1.3 Earlier versions of* DSM *and* ICD 195
- *7.09.2.2 Symptom Presentation in Late Life* 195
- *7.09.2.3 Subtypes* 196
  - *7.09.2.3.1 Melancholic* 196
  - *7.09.2.3.2 Delusional depression* 196
  - *7.09.2.3.3 Minor depression* 197
  - *7.09.2.3.4 Mixed anxiety–depressive disorder* 197
  - *7.09.2.3.5 Masked depression* 197
  - *7.09.2.3.6 Depletion syndrome* 197
  - *7.09.2.3.7 Hopelessness depression* 198
- *7.09.2.4 Categorical vs. Dimensional Systems* 198

7.09.3 PREVALENCE 198

- *7.09.3.1 Rates of Mood Disorders* 198
- *7.09.3.2 Presence of Depressive Symptoms* 198
- *7.09.3.3 Age and Depression* 198
- *7.09.3.4 Gender* 202
- *7.09.3.5 Ethnicity/Race* 202
- *7.09.3.6 Special Populations* 202
  - *7.09.3.6.1 Medical inpatients and outpatients* 202
  - *7.09.3.6.2 Long-term care facilities* 202
  - *7.09.3.6.3 Caregivers* 203

7.09.4 ETIOLOGICAL AND CORRELATED FACTORS 203

- *7.09.4.1 Physical Health* 203
- *7.09.4.2 Cognitive Dysfunction* 204
- *7.09.4.3 Interpersonal Relationships and Stressful Life Experiences* 205
- *7.09.4.4 Genetic Risk* 205

7.09.5 ASSESSMENT 206

- *7.09.5.1 Self-report Instruments* 206
  - *7.09.5.1.1 Depression scales developed with mixed-age samples* 206
  - *7.09.5.1.2 Instruments devised specifically for use with older adults* 208
  - *7.09.5.1.3 Multidimensional measures with depression subscales* 208
- *7.09.5.2 Observer-rated Inventories* 209
- *7.09.5.3 Structured Interviews* 209
  - *7.09.5.3.1 Instruments designed for mixed ages* 209
  - *7.09.5.3.2 Instruments designed for older adults* 209
- *7.09.5.4 Assessing Depression in Cognitively-impaired Older Adults* 210
- *7.09.5.5 Assessing Depression in Older Adults with Poor Physical Health* 210

7.09.6 SUICIDE 211

- *7.09.6.1 Rates* 211

7.09.6.2 Methods 211
7.09.6.3 Depression and Suicide 211
7.09.6.4 Hopelessness 212
7.09.6.5 Issues in Late Life Suicide 212
7.09.6.6 Assessment and Treatment 212
7.09.7 COURSE OF DEPRESSION 212
7.09.8 TREATMENT APPROACHES 213
7.09.8.1 Psychotherapeutic Interventions 213
7.09.8.1.1 Cognitive-behavior therapy 213
7.09.8.1.2 Brief psychodynamic, interpersonal, and life review therapies 215
7.09.8.1.3 Support groups and psychoeducational interventions for depression in caregivers of dementia patients 215
7.09.8.2 Pharmacological Interventions 216
7.09.8.3 Electroconvulsive Therapy 217
7.09.8.4 Concluding Comment 217
7.09.9 PREVENTION 218
7.09.9.1 Indicated Preventive Measures 218
7.09.9.2 Selective Preventive Interventions 219
7.09.9.3 Universal Preventive Interventions 219
7.09.10 CONCLUSION 219
7.09.11 REFERENCES 220

7.09.1 INTRODUCTION

This chapter provides an overview of mood disorders in older adults. Issues addressed include: How well does the current diagnostic system map onto the types of affective disorders that are observed among older adults? How frequent are mood disorders among older adults? What is the connection of depression to suicide among older adults? What interventions are effective for preventing depression in old age, for alleviating episodes of depression, and for maintaining improvement and thwarting recurrence? Additionally, what is known about etiology is explored, emphasizing the complex influences that link affective symptoms, physical illness, and cognitive dysfunction. This material is then used to inform the discussion of assessment and treatment.

Although the relative proportions are unclear, a key differentiation is between depressive disorders that represent a continuation of conditions from earlier in the life cycle and disorders with first onset in old age. Throughout the chapter, therefore, this differentiation will be maintained, although regrettably the literature frequently does not provide sufficient information to divide the data in this manner.

The mood disorders include depressive disorders, both major depressive disorder and dysthymic disorder, and bipolar disorders. There is minimal literature on mania or bipolar illness in old age. Instead, gerontological researchers have paid most attention to either major depression or subsyndromal symptom constellations. While the task is to encompass all mood disorders, the emphases reflect the differential availability of literature.

Finally, attention is called to the importance of this material, given the common image of depression as characteristic of old age and the frequency with which depressive symptoms are used as outcome measures in research on aging.

7.09.2 DEFINITION AND CLASSIFICATION

Definitional problems underlie debates about prevalence and etiology. For this reason, different classification systems are reviewed and the effect of differences in criteria on decisions about diagnosis and rates of disorder is explored. The most commonly accepted classifications of mood disorders are those in the *Diagnostic and statistical manual of mental disorders* (*4th Edition; DSM-IV*; American Psychiatric Association, 1994) and the *International classification of disease* (*10th Edition*; *ICD-10*; World Health Organization, 1992). These include depressive disorders (major depressive disorder and dysthymic disorder) and bipolar disorders. It is not clear, however, how fully these conventional diagnostic categories represent the experience of affective disorders among older adults. For this reason, descriptions of several subtypes of depression are also presented that may better capture the syndromes prevalent in late life.

7.09.2.1 Standardized Classification Systems

7.09.2.1.1 DSM-IV

Major depressive disorder is defined in the *DSM-IV* by the presence most of the day, nearly every day, of dysphoria (depressed mood), anhedonia (diminished interest or pleasure in all or nearly all activities), appetite or weight disturbance, sleep disturbance, psychomotor agitation or retardation, loss of energy, worthlessness or guilt, inability to concentrate, and/or recurrent thoughts of death or suicide. At least five symptoms must be present, of which one must be either dysphoria or anhedonia. Symptoms attributed to uncomplicated bereavement, the direct physiological effects of a medical condition, or use of a substance (including medications), are excluded. A diagnosis of depressive disorder is made, however, if symptoms in a bereaved person persist longer than two months or include marked functional impairment, morbid preoccupation with worthlessness, suicidal ideation, psychotic symptoms, or psychomotor retardation.

Dysthymic disorder is characterized by chronicity and a lower level of severity than major depression. Depressive symptoms must be present for at least two years. If criteria for both major depressive and dysthymic disorder are met, both diagnoses are assigned, a condition commonly referred to as double depression.

Bipolar-I disorder is marked by alternating major depressive and manic episodes, resulting in impaired functioning. If the manic episodes are briefer and do not interfere with functioning (i.e., hypomanic episodes), the diagnosis of bipolar-II is assigned. Hypomanic episodes alternating with dysthymic symptoms are classified as cyclothymic disorder.

Depressive symptoms that do not meet the above diagnostic criteria, but can be attributed to a specific stressor, may be classified as adjustment disorder with depressed mood.

7.09.2.1.2 ICD-10

There is much similarity between *ICD-10* and *DSM-IV* for mood disorders. Depressive disorders are classified by severity, with moderate and severe equivalent to the *DSM-IV* category of major depressive disorder, and mild episode more closely resembling the provisional category of minor depression in the appendix of *DSM-IV*. *ICD-10* criteria for bipolar affective disorder, dysthymia, and cyclothymia are equivalent to the corresponding *DSM-IV* categories.

7.09.2.1.3 Earlier versions of DSM *and* ICD

Much of the currently available research on mood disorders is based on earlier versions of *DSM* and *ICD*, including *DSM-III* (American Psychiatric Association, 1980), *DSM-III-R* (American Psychiatric Association, 1987), *ICD-9-CM* (Commission on Professional Hospital Activities, 1978), and *ICD-9* (World Health Organization, 1977), or on the Research Diagnostic Criteria (RDC; Spitzer, Endicott, & Robins, 1978). The RDC influenced changes appearing in *DSM-III*, which was published shortly thereafter. RDC mood disorder classifications include: major depressive disorder, minor depressive disorder, bipolar with mania (bipolar-I), bipolar with hypomania (bipolar-II), and intermittent depressive disorder. Changes in *DSM* categories since 1980 have been minimal, while some *ICD* classifications have been completely reworked during this time frame.

In the *DSM*, criteria for mood disorder have become somewhat more stringent with later revisions. A requirement that symptoms be clinically significant or cause impaired social, occupational, or other functioning was introduced in the *DSM-III-R* for dysthymia and in the *DSM-IV* for major depressive disorder. This criterion could affect diagnosis in older adults disproportionally, who may be less likely to engage in social or occupational roles in which impaired functioning could be observed. There have been significant changes in the specification of melancholic features, with *DSM-IV* criteria more closely resembling *DSM-III* than *DSM-III-R*. Bipolar-II disorder was added to *DSM-IV* as a new classification.

Perhaps the greatest changes have occurred between *ICD-9* and *ICD-10*. In the earlier version, mood disorders were classified as either a neurotic disorder or an affective psychosis. Since diagnostic criteria were not specified in detail in *ICD-9*, precise comparisons of classifications are not possible. Affective psychosis approximates the *ICD-10* category of depressive disorder with somatic symptoms, while a neurotic depression would be diagnosed in *ICD-10* as either depressive disorder without somatic symptoms or dysthymia. Due to the differences between classification systems (particularly between *ICD-9* and all other systems), caution should be exercised when comparing results of studies that were not based on a common rubric.

7.09.2.2 Symptom Presentation in Late Life

Although some scholars emphasize similarities between older and younger individuals in the experience of depressive symptoms, there is also an accumulating body of research identifying differences in symptom presentation in late life (for a review, see Caine, Lyness, King, & Connors, 1994). The most frequent differences include: lower prevalence of dysphoria, fewer ideational symptoms (e.g., guilt, suicidal ideation), and increases in selected somatic symptoms.

Several authors have reported a decreased likelihood of dysphoria as a symptom of depression in old age (e.g., Kongstvedt & Sime, 1992). In the Epidemiologic Catchment Area (ECA) sample, Gallo, Anthony, and Muthén (1994) found that older individuals were less likely to endorse an item comprising dysphoria or anhedonia than younger respondents with the same level of overall depression. Using a self-report symptom checklist, Gatz and Hurwicz (1990) found that depressed mood was neither higher nor lower in older adults; however, the older adults in their sample were less likely than other age groups to endorse items on a well-being subscale. Consequently, these authors speculate that older adults may be more likely to experience a lack of positive feelings than active negative feelings.

Age differences in ideational symptoms among depressed patients have been described by Musetti et al. (1989) and Small, Komanduri, Gitlin, and Jarvik (1986). They found lower levels of guilt and lower levels of a factor including self-blame and suicidal ideation. Even though older adults are more likely to complete acts of suicide than younger individuals (Gulbinat, 1996), they appear to be less likely to display suicidal ideation (Blazer, Bachar, & Hughes, 1987; Brown, Sweeney, Loutsch, Kocsis, & Frances, 1984). Brown et al. noted that lower levels of guilt were characteristic of late-onset cases.

The central question about somatic symptoms in depressed older adults is whether they are reliable indicators of depression or merely an artifact of the high comorbidity of physical illness in this age group. Blazer et al. (1987) found that older adults with melancholia were more likely than younger adults to have weight loss and constipation. In patients with major depression, older adults have been found more likely to have delusions, agitation, appetite disturbance, initial insomnia, and hypochondriasis than younger patients (Brodaty et al. 1991; Brown et al. 1984). No difference in somatic symptom presentation was found between early- and late-onset cases. Other investigators have found no age differences in somatic symptoms or preoccupation with physical health, either in depressed psychiatric patients (Kongstvedt & Sime, 1992) or among male patients with medical problems (Koenig, Cohen, Blazer, Krishnan, & Silbert, 1993). Using self-report depression symptom instruments, Kessler, Foster, Webster, and House (1992) found similar relationships with age for both mood and somatic symptom subscales. Finally, Davidson, Feldman, and Crawford (1994) found that physically disabled older adults who scored high on the somatic factor tended to score high on the other factors as well.

Regardless of how common or uncommon specific symptoms are in older adults, it is also important to know which of them are useful prognostically. The particular symptoms that best distinguish depressed from nondepressed older individuals seem to include loss of interest, lack of energy, sleep disturbance, suicidal thoughts, and feeling blue (Koenig et al., 1993; Norris, Snow-Turek, & Blankenship, 1995). Fatigue and changes in appetite and sexual activity did not, and may reflect physical comorbidity or normal aging.

7.09.2.3 Subtypes

Knowledge of the depressive symptoms most frequently experienced by older adults informs the discussion of several subtypes of depression or symptom constellations thought to be relevant in later life. The subtypes reviewed here include ones that can be accommodated within the diagnostic rubric (melancholia and delusional depression), experimental diagnostic categories (minor depression and mixed anxiety–depression), and other syndromes that may occur in older adults (masked depression, depletion syndrome, and hopelessness depression). There is also considerable convergence across many of these entities.

7.09.2.3.1 Melancholic

In the *DSM-IV*, individuals with a unique pattern of symptoms, including nearly complete loss of pleasure in most activities, lack of reactivity to pleasurable stimuli, a distinct quality of mood, depression worse in the morning, early waking, psychomotor retardation or agitation, weight loss, and excessive guilt, are diagnosed with major depressive disorder with melancholic features. At least one of the first two symptoms must be present. The criteria for melancholia in *ICD-10*, referred to as "somatic symptoms," includes the same symptom pattern as "melancholic features" in *DSM-IV* with the exception of distinct quality of depressed mood and excessive guilt. The term "endogenous" has also been used to refer to melancholic depression.

7.09.2.3.2 Delusional depression

Delusional depression is a depressive disorder with the presence of mood-incongruent delusions. It is diagnosed in *DSM-IV* and *ICD-10* by specifying psychotic features/symptoms. In

ICD-10, psychotic symptoms are further classified as mood-congruent or incongruent. There is mixed evidence concerning the relationship between delusional depression and age of onset. Meyers, Greenberg, and Mei-Tal (1985) found delusional depression to be more common in later-onset cases and in late-onset endogenous depression (Meyers, Kalayam & Mei-Tal, 1984). As mentioned earlier, Brodaty et al. (1991) found delusions to be more common in older adults than in younger adults. Nelson, Conwell, Kim, and Mazure (1989) and Baldwin (1995), on the other hand, found no such association.

7.09.2.3.3 Minor depression

As a result of concerns regarding the validity of diagnostic categories for older adults, the 1991 Consensus Conference on Depression in Late Life issued a call for research on "minor (subsyndromal) depression" (Blazer, 1994). Standardized study criteria for minor depression, included in the appendix of *DSM-IV*, are the same as those for major depressive disorder but require fewer symptoms. A similar definition has been referred to as "subdysthymic depression" (Mossey, Knott, Higgins, & Talerico, 1996) or simply "depressive symptoms" (Newmann, 1989; Wells et al., 1989). Several authors have used the term "minor depression," but defined it more inclusively, capturing dysthymia and adjustment disorder with depressed mood as well as dysphoria (Beekman et al., 1995; Koenig & Blazer, 1992; Parmelee, Katz, & Lawton, 1989).

Beekman et al. (1995) found that risk factors for major depression (family history of depression, previous episodes of depression) differed from those for minor depression (being unmarried, having smaller networks of social contacts, living in a large city, chronic disease, and functional limitations). These authors concluded that major depression may be more chronic, while minor depression may be more reactive to stressors. Although minor depression requires fewer symptoms than major depression, and the name implies a syndrome with lesser severity, the Medical Outcomes Study (Wells et al., 1989) revealed that even depressive symptoms in the absence of a depressive disorder were associated with more functional disability than most physical illnesses studied.

7.09.2.3.4 Mixed anxiety–depressive disorder

Another experimental category added to *DSM-IV* that may be relevant to older adults is mixed anxiety–depressive disorder. This proposed disorder is characterized by a combination of subsyndromal depression and anxiety symptoms. The research criteria require a persistent or recurrent dysphoric mood lasting at least one month, with at least four of these additional symptoms: difficulty concentrating, sleep disturbance, fatigue or low energy, irritability, worry, tearfulness, hypervigilance, anticipating the worst, hopelessness, or low self-esteem. Although mixed anxiety-depression was found in one study to be more prevalent in younger than in older age groups (Blazer et al., 1988), this subtype merits further research in an older population due to the obvious overlap between the specified symptoms and those found commonly in depressed older adults.

7.09.2.3.5 Masked depression

A further term used for older individuals is masked depression (Blumenthal, 1980). Also known as somatic depression, masked depression refers to a predominance of somatic symptoms without mood-related symptoms. As the term "masked" implies, diagnosis can be difficult due to the absence of the dysphoric mood that characterizes depression.

7.09.2.3.6 Depletion syndrome

Fogel and Fretwell (1985) proposed depletion syndrome as a new subtype of depression, characterized by predominant somatic and vegetative symptoms in depressed older adults with good premorbid psychosocial functioning and no evidence of psychosis. Newmann, Klein, Jensen, and Essex (1996) have pursued this concept, identifying a set of symptoms that occur together more frequently in older adults than in younger individuals. This constellation of symptoms differs from the depression syndrome in that self-deprecating feelings (self-blame and inappropriate guilt) are not present, dysphoric mood items are less prominent, and enervation, loss of interest in things, hopelessness, and loss of appetite are more prominent. Although sleep disturbance is associated with depletion syndrome, it does not include early morning awakening. Newmann et al. characterize this syndrome as "quieter" than major depression. In a longitudinal study, they found that individuals with depletion syndrome were more likely to die before the next time of measurement than those with depressive syndrome, suggesting that depletion syndrome may be related to a wearing down process. It is important to note that the only research on this subtype to date has been in a female population.

7.09.2.3.7 Hopelessness depression

This subtype was developed in younger populations (Abramson, Metalsky, & Alloy, 1989). Like depletion syndrome, it is characterized by decreased motivation, lack of energy, apathy, and psychomotor retardation, as well as by sad affect. Thus, applicability of this conceptualization to older adults seems warranted.

7.09.2.4 Categorical vs. Dimensional Systems

Although the focus has been on various categories used to describe mood disorders in older adults, depression has also been conceptualized as a continuum. Caine et al. (1994) recommend the use of a dimensional descriptive system encompassing the full range of mood syndromes and comorbid medical conditions. Unlike categorical diagnostic methods, depressive symptom checklists yield a rating on a continuous scale and are inclusive in approach. In general, scores relating to the presence or severity of individual symptoms are summed to provide a single rating of level of depressive symptomology. Since no symptoms are excluded due to etiology (e.g., bereavement, physical illness), a more complete picture of distress is provided.

7.09.3 PREVALENCE

7.09.3.1 Rates of Mood Disorders

Studies measuring depressive disorders have found a relatively low prevalence in late life. In a review of earlier studies, Kay and Bergmann (1980) concluded that the prevalence of depressive neuroses ranged from 4 to 9%, and that field surveys on affective psychoses were too limited to yield reliable information. Table 1 summarizes studies of the prevalence of mood disorders in late life that employed representative sampling methods and large sample sizes. The one-year prevalence of major depression among individuals aged 65 years and older ranges from a reported low of 0.9% in the ECA survey in the US (Weissman, Bruce, Leaf, Florio, & Holzer, 1991) to a high of 5.7% in a study of ethnic Chinese living in Singapore (Kua, 1992), with more studies finding rates at the low end of this range. Fewer studies have specifically measured rates of dysthymia in an older population. Due to the chronic nature of this disorder, only lifetime rates have been reported. Weissman and Myers (1978) found a 1.8% lifetime prevalence of dysthymic disorder among older adults using RDC. In the more recent ECA study, based on *DSM-III* criteria, Weissman et al. (1991) reported an overall rate of dysthymic disorder of 1.7% among those aged 65 years or older. Studies of bipolar disorder in older adults are even more rare. According to ECA data, one-year prevalence rates of bipolar disorder ranged from 0 to 0.4% across study sites, for an overall prevalence of 0.1%. Establishing the prevalence of other diagnostic subtypes, such as minor depression, is complicated by varying definitions of these entities.

7.09.3.2 Presence of Depressive Symptoms

Several studies have measured the prevalence of clinically significant depressive symptoms, without tying symptoms to diagnostic criteria for a depressive disorder. Rates of mood disorders suggested through the use of clinical cutoffs on depressive symptom checklists are shown in Table 2 for adults aged 65 years and older. These rates range from 9.0 to 16.9%. Rates found by these studies are generally high compared to studies using diagnostic criteria, as would be expected, since summary scores cannot distinguish between individuals with a disorder and those with subsyndromal levels of symptoms.

7.09.3.3 Age and Depression

In general, epidemiological studies have found lower rates of depressive disorders in older adults compared to other age groups, whereas studies measuring depressive symptoms have found elevated levels in late life, particularly after age 75 years (see Newmann, 1989, for a review). This outcome lends support to the notion that the symptoms experienced by older individuals may not map well onto existing diagnostic criteria.

The association between age and depression is further complicated by the consistent finding that older individuals report a lower lifetime prevalence of depression compared to younger individuals (Weissman et al., 1991). Since older adults have had more years in which to experience depressive episodes, this finding is remarkable. Several possible explanations have been posited. Some researchers have suggested that later-born generations in the twentieth century have experienced progressively higher rates of depression than those born earlier (e.g., Klerman et al., 1985). This cohort effect could explain relatively low rates of depression among those now old, and would predict that rates of depression in late life would rise as successive generations enter old age. A second hypothesis suggests that lifetime prevalence rates are

Table 1 Prevalence of affective disorders in late life.

Authors	*Study*	*Sample*	*N*	*Age range*	*Measures*	*Diagnostic criteria*	*Major depression*	*Minor depression*	*Dysthymia*	*Bipolar disorder*
Baltes et al. (1993); Wernicke and Linden (1994)	Berlin Aging Study	Stratified random (West Berlin)	156	70+	GMS-A, HAS	*DSM-III-R*	4.8%			
Beekman et al. (1995)	Longitudinal Aging Study, Amsterdam	Stratified random (Netherlands)	3056	55–85	DIS, CES-D (16+)	*DSM-III*	2.0% m = 1.4% f = 2.6%	12.9% m = 9.8% f = 15.7%		
Copeland, Dewey, et al. (1987)		Random[b] (Liverpool)	1070	65+	GMS-A AGECAT		2.9% m = 1.7% f = 3.6%	8.3% m = 5.9% f = 9.9%		
Copeland, Gurland, et al. (1987)	US–UK Cross-National Project	Random (London[b] New York)	841	65+	GMS-A AGECAT		2.5% m = 1.3% f = 3.2%	m = 11.8% f = 17.3%		
Henderson et al. (1993)	Canberra-Queanbeyan Study of the Elderly	Random community and institutional samples (Australia)	848	70+	Canberra Interview for the Elderly	*DSM-III-R* *ICD-10*	*DSM*: 1.0% m = 0.4% f = 1.5% *ICD*: 3.3%[a] m = 2.2%[a] f = 3.9%[a]			
Kua (1992)		Stratified random (Singapore)	612	65+	AGMS-A AGECAT	*ICD-9*	5.7%			
Matt, Dean, Wang, and Wood (1992)		Stratified probability (Albany, New York)	1131	50+	SCL-90	*DSM-III-R*	3.9% m = 3.2% f = 4.4%			

Table 1 (continued).

Authors	*Study*	*Sample*	*N*	*Age range*	*Measures*	*Diagnostic criteria*	*Major depression*	*Minor depression*	*Dysthymia*	*Bipolar disorder*
O'Hara, Kohout, and Wallace (1985)		Entire population of two rural counties (Iowa)	3159	65+	CES-D	RDC	1.2%	9.0%		
Uhlenhuth, Balter, Mellineger, Cisin, and Clinthorne (1983)		National probability (US)	798	65–79	HSCL	*DSM-III*	5.1%[c]			
Weissman and Myers (1978)		Stratified random (New Haven)	110	66+	SADS, project specific scale	RDC	5.4%	2.7%	1.8%	
Weissman et al. (1988, 1991)	Epidemiologic Catchment Area (ECA)	Area probability		65+	DIS	*DSM-III*	0.9%, m = 0.4% f = 1.4%		m = 1.0% f = 2.3%	0.1%
		(New Haven)		2576			m = 0.8% f = 1.8%		m = 1.4% f = 3.1%	0.1%
		(Baltimore)		923			m = 0.3% f = 1.5%		m = 0.6% f = 1.3%	0.0%
		(St. Louis)		576			m = 0.1% f = 0.6%		m = 0.5% f = 1.6%	0.4%
		(Central NC)		1240			m = 0.1% f = 1.2%		m = 0.3% f = 2.5%	0.1%
		(Los Angeles)		386			m = 0.7% f = 1.2%		m = 2.5% f = 2.4%	0.0%

[a] Includes mild, moderate, and severe depression. [b] Sampling frame—general medical practice patient lists. [c] Includes dysthymia.

GMS-A, Geriatric Mental State Schedule-Abbreviated; HAS, History and Aetiology Schedule; DIS, Diagnostic Interview Schedule; CES-D, Center for Epidemiological Studies—Depression Scale; AGECAT, Automated Geriatric Examination for Computer-Assisted Taxonomy; SCL-90, Symptom Checklist 90; RDC, Research Diagnostic Criteria; HSCL, Hopkins Symptom Checklist; SADS, Schedule for Affective Disorders and Schizophrenia; m, male; f, female.

Table 2 Prevalence of clinically significant depressive symptoms in late life.

Authors	*Study*	*Sample*	*N*	*Age range*	*Measures*	*Above cutoff*
Berkman et al. (1986)	EPESE/Yale Health and Aging Study	Probability, noninstitutionalized (New Haven)	2806	65+	CES-D (16+)	16.4% m = 11.3% f = 19.2%
Blazer et al. (1991)	EPESE/Duke	Stratified random (North Carolina)	4163	65+	modified CES-D (9+)	9.0% m = 6.4% f = 10.7%
Eaton and Kessler (1981)	HANES	National cluster probability (US)	338[a]	65+	CES-D (16+)	14.8%
Frerichs et al. (1981)		Area probability (Los Angeles)	126	66+	CES-D (16+)	16.7%
Fuhrer et al. (1992)	Paquid	Stratified random, noninstitutionalized (France)	2792	65+	CES-D (m = 17+, f = 23+)	13.4% m = 12.8% f = 13.8%
Gatz Johansson, Pedersen, Berg, and Reynolds (1993)	SATSA	Population-based, twins (Sweden)	1193	50–59 60–69 70–79 80+	CES-D (16+)	20% 24% 25% 28%
Kennedy et al. (1989)		Random (urban community, USA)	2137	65+	CES-D (16+)	16.9% m = 11.1% f = 19.9%
Lindesay Briggs, and Murphy (1989)	Guy's/Age Concern	Random, noninstitutional (inner city London)	890	65+	CARE (13+)	4.3% m = 2.8%, f = 5.3%
					(8+)	13.5%, m = 8.4%, f = 16.9%
Murrell et al. (1983)		Stratified area probability (Kentucky)	2517	55+	CES-D (20+)	m = 13.7% f = 18.2%
O'Hara et al. (1985)		Entire population of two rural counties (Iowa)	3159	65–105	CES-D (16+)	9.0% m = 7.3% f = 10.0%
Stallones et al. (1990)		Stratified probability, noninstitutionalized (US)	1232	65+	CES-D (16+)	9.9%

[a] Reported total $n = 2\,867$, 11.8% aged 65+ years.

EPESE, Established Populations for Epidemiologic Studies of the Elderly; CES-D, Center for Epidemiologic Studies—Depression Scale; HANES, Health and Nutrition Examination; SATSA, Swedish Adoption/Twin Study of Aging; CARE, Comprehensive Assessment and Referral Evaluation; m, male; f, female.

selectively biased in old age, with forgetting of symptoms exacerbated by cognitive impairment. In fact, research has found lifetime prevalence rates to be fairly unreliable at all ages (Kendler, Neale, Kessler, Heath, & Eaves, 1993). A third factor is selective mortality. Since depression has been associated with increased risk of mortality (Burvill, & Hall, 1994), depressed individuals would be less likely than nondepressed individuals to survive into old age. However, Klerman and Weissman (1889) argued that selective mortality is insufficient to explain observed differences, since a strong association between depressive disorders and subsequent death has not been consistently demonstrated.

It is not clear what proportion of depression in older adults within the community represents late-onset cases. Two studies of consecutive psychiatric admissions found approximately half to have first onset after the age of 60 years (Meyers et al., 1985; Steingart & Herrmann, 1991).

7.09.3.4 Gender

Studies of mixed age populations have consistently found the prevalence of major depressive disorders to be higher in women than in men. This sex difference is also evident for older adults. In both the ECA data (Weissman et al., 1991) and the Canberra study (Henderson et al., 1993), the rate of major depressive disorder in older women was 3.5 times the rate for older men. Evidence indicates, however, that this difference diminishes, or even reverses, in extreme old age. Equal prevalence rates for men and women with major depression were found by Girling et al. (1995) for individuals over 77 years and by Forsell, Jorm, von Strauss, and Winblad (1995) for nonagenarians. Wallace and O'Hara (1992) found that the sex difference on the Center for Epidemiologic Studies—Depression scale (CES-D) diminished in older age and was reversed in individuals aged 85 years and older.

In contrast to the sex ratio for major depressive disorder, bipolar disorder seems to occur with equal frequency in men and in women in mixed-age samples and in old age (Weissman et al., 1984).

7.09.3.5 Ethnicity/Race

Few studies report rates of mood disorders in older adults by race or ethnicity. Blazer, Burchett, Service, and George (1991) reported no differences between Black and non-Black older adults on a modified CES-D. In a national probability sample of older adults, Stallones, Marx, and Garrity (1990) found no differences in rates of major depression for Black and White men, but Black women were nearly twice as likely to have major depressive disorder as White women. Both Frerichs, Aneshensel, and Clark (1981) and Kemp, Staples, and Lopez-Aqueres (1987) found Latino older adults in Los Angeles to have especially elevated rates of depression. The authors suggest that income and education levels may account for ethnic differences. In the latter report, rates were highest among non-English speaking individuals, raising issues both of translation of terms and cultural understanding of the meaning of symptoms.

7.09.3.6 Special Populations

7.09.3.6.1 Medical inpatients and outpatients

The prevalence of mood disorders is higher in medical outpatients, medical inpatients, and residents of long-term care facilities than in the general population (see review by Koenig & Blazer, 1992). The few studies examining depressive disorders among older adult outpatients have reported widely varying rates. Barrett, Barrett, Oxman, and Gerber (1988) found major depression or minor episodic depression in 5.6% of patients over age 65 years in a general medical practice, with an additional 7.9% showing masked depression. In contrast, two studies of Veterans Administration (VA) outpatient clinic populations found rates of major depression of 23% (Norris, Gallagher, Wilson, & Winograd, 1987) and 31% (Okimoto et al., 1982).

Studies of older medical inpatients have yielded rates of major depression ranging generally from 6 to 13%, with an additional 15–30% displaying dysthymia, adjustment disorder, depressive disorders not otherwise specified, or dysphoria (Koenig & Blazer, 1992). As with outpatient samples, the highest rates of depression among older adult inpatients have been found in a VA setting. Kitchell, Barnes, Veith, Okimoto and Raskind (1982) found major depressive disorder in 45% of a sample of veterans.

7.09.3.6.2 Long-term care facilities

Studies of depression in institutional settings have found 6–16% meeting *DSM* criteria for major depression, and an additional 25–35% diagnosed with another type of depressive disorder or significant dysphoria (Henderson et al., 1993; Koenig & Blazer, 1992). Both Parmelee et al. (1989) and Mann, Graham, and Ashby 1984) found an approximately doubled risk among newly admitted residents compared to veteran residents or intrainstitution transfers.

Although few studies have examined rates of bipolar disorder in institutional settings, the ECA found a rate of 9.7% in nursing homes compared to a rate of major depression of 16.2% (Weissman et al., 1991).

7.09.3.6.3 Caregivers

Depressive disorders and depressive symptoms have been found in a high proportion of older adult caregiver samples, although methodological issues make comparisons with the general population difficult. In studies using diagnostic criteria ascertained by clinical interview, rates of depressive disorders among caregivers range from 23 to 46%, while virtually all studies measuring depressive symptoms have found elevated levels in this group (Schulz & Williamson, 1994). A problem with research on caregivers is lack of representative sampling, since most samples are drawn from individuals who have sought help (e.g., support group members) and therefore may be more distressed.

7.09.4 ETIOLOGICAL AND CORRELATED FACTORS

This section reviews factors that are related to clinically diagnosed depression and severity of depressive symptoms. Some of these variables are generally regarded as risk factors. However, the terminology of risk factors implies a temporal and causal relationship, whereas the terminology of correlated factors allows for reversed causal order, feedback mechanisms, or that both depression and the respective factor are the outcome of the same process. Variables encompassed in this section include physical health, cognitive dysfunction, interpersonal relations and stressful life events, and genetic risk.

When considering etiological and correlated factors, it is important to keep in mind that not all older adults who experience one or more of these factors necessarily become depressed. Why are some individuals vulnerable and others apparently resilient? Coping styles, perceived control and efficacy, and social support are among the variables that have been shown to mediate the relationship. These are relevant protective factors when designing preventive interventions and useful variables in accounting for individual differences in response to physical illness, interpersonal loss, and cognitive impairment.

7.09.4.1 Physical Health

Physical health status, which can be defined in terms of medical conditions and associated medication use, functional disability, pain, and subjective perceptions of poor health, has been found to be correlated with depression (Forsell et al., 1995; Heikkinen, Berg, & Avlund, 1995; Kennedy et al., 1989; Stallones et al., 1990). This is one of the most well-established correlations in studies of mood and aging. Several explanations for this association exist: physical illness and/or its treatment induce biological changes that trigger depressive symptoms; the same physiological factors underlie physical illness and depression; depression leads to physical illness by creating vulnerabilities, possibly neuroimmune, or by undermining self-care efforts such as medication compliance; physical illness is a stressful experience with which people have difficulty coping; or physical illness and depression co-occur as independent conditions. Thus, the powerful relationship between physical illness, associated conditions such as reduced activity, functional disability and perceptions of health, and depressive symptoms could reflect interdependence and continuing feedback loops of biological changes and psychological factors.

Biological changes induced by neurochemical, neuroanatomical, and functional alterations can create a liability for depression or for mania (see Ferrier & McKeith, 1991, and Leuchter, 1994, for reviews; Conn, 1989; Shulman, Tohen, Satlin, Mallya, & Kalunian, 1992). Preexisting disease processes, such as Parkinson's disease, medical conditions, such as hypothyroidism, or focal lesions may also induce affective changes (see Rubin, 1994 for review; Ferrier & McKeith, 1991). Cerebrovascular disease, including stroke and more minor insults such as deep white matter changes and lacunar infarcts, are particularly likely to precipitate depression (Lawlor & Anderson, 1995; Robinson & Starkstein, 1990), and vascular disease has a notable relationship with late-onset depression (Baldwin & Tomenson, 1995). Use of prescription medications for these conditions may precipitate or aggravate depression (Kinzie, Lewinsohn, Maricle, & Teri, 1986), as may psychotropic medications (Forsell et al., 1995).

Once present, depressive symptoms can worsen morbidity in medically ill individuals (Rapp, Smith, & Britt, 1990). This worsening may reflect a depression-induced vulnerability, a shared process, or a reaction to illness and disability. For example, stroke-induced paralysis could reduce mobility and increase dependence that has a negative impact on social and emotional functioning. Lyness, Caine, Conwell, King, and Cox (1993) found that in both younger and older depressed inpatients, physical disability was associated with depressive symptoms independent of, and likely additive to, medical illness. This finding is consistent with the conclusion that physical disability is associated with depressive symptomology across a range of physical conditions (Turner & Beiser, 1990), some of which do not have a purported biological basis for depression.

However, depression has also been found to increase risk of disability in individuals in the absence of significant health problems or disability (Bruce, Seeman, Merrill, & Blazer, 1994). Perceptions of health may be an important moderating variable given that perceived poor health or disability, irrespective of objective health status, can create depressogenic effects (Williamson & Schulz, 1992b). However, depressive symptoms also can influence self-perceived physical function (Cress et al., 1995).

The presence of pain contributes to the association between depression, disability, and medical illness. The association among illness, pain, disability, and depression has been examined in both community and institutionalized samples. In a community sample, Williamson and Schulz (1992a) found pain related to depression through its association with activity restriction. In institutionalized older adults, Parmelee, Katz, and Lawton (1991) found pain to serve as an independent predictor of depression.

7.09.4.2 Cognitive Dysfunction

Depression has been found to be associated with varying degrees of cognitive dysfunction. Three primary associations seem to exist: (i) memory loss can lead to depression; (ii) depression can be accompanied by cognitive impairment that is reversible, at least temporarily, in that it lifts when the depression is treated (referred to as "pseudodementia" or "depression related cognitive dysfunction"); and (iii) depression can co-occur with dementia, where depression represents a prodromal symptom, a comorbid condition possibly reflecting similar physiological factors, or a reaction to the dementia. The possibility of these variations, expanded below, underscores the difficulty in differentiating depression and dementia diagnostically.

Depression does not seem to be associated with a global decline in cognition. When an effect is observed, it tends to be limited to circumscribed deficits in specific domains, that is, executive functions and right-hemisphere tasks, although findings are equivocal across studies and among different measures of the same domain (see La Rue, 1992, for a review). For example, depressed middle-aged and older adults have been found to show deficits in nonverbal memory, select frontal lobe skills (Boone et al., 1994), and visual-spatial tasks (Lyness, Eaton, & Schneider, 1994), which Boone et al. interpret as premature "aging" of cognitive abilities. In contrast, O'Hara, Hinrichs, Kohout, Wallace, and Lemke (1986) found no significant verbal memory impairment in depressed elderly community-residing older adults. These findings are important because complaints about memory frequently accompany depression in older adults (O'Connor, Pollitt, Roth, Brook, & Reiss, 1990), and depression can affect concentration and hence performance on cognitive test batteries (Poon, 1992). Demonstrable cognitive deficits are observed more often among very severely depressed patients than in community samples. One study of older adult depressed inpatients found that impairment in learning and memory was related to longer duration of depression (La Rue, Goodman, & Spar, 1992). Further, elderly inpatients with psychotic features show poorer cognitive functioning than depressed patients without psychotic features (Kunik, Champagne, Harper, & Chacko, 1994).

The added difficulty with these studies is that it is possible that some of the more severe depressed patients are actually early dementia cases. Longitudinal follow-up suggests that a majority of cognitively impaired depressed patients who regain premorbid cognitive function after the depression is treated will develop dementia if followed long enough (Kral & Emery, 1989) and that degree of cognitive impairment in depression may predict who will develop dementia (Alexopolous, Meyers, Young, Mattis, & Kakuma, 1993). Various characteristics have been suggested to distinguish patients whose cognitive symptoms are due to depression alone from patients who are becoming demented, especially motivational disturbances and vegetative symptoms (Forsell, Jorm, & Winblad, 1994; Lazarus, Newton, Cohler, Lesser, & Schweon, 1987). History may be even more helpful than current features (see Kaszniak & Christenson (1994) for a comprehensive discussion of differential diagnosis).

Finally, there can be multiple, shared, or reciprocal causes. Although little is known about the cognitive dysfunction of depressed, demented individuals, there is some suggestion that depression worsens the cognitive impairment (Fitz & Teri, 1994). Moreover, demented patients with co-occurring depression show some cognitive improvement when depression is treated (Greenwald et al., 1989). Gross cognitive impairment concomitant with later age of onset of depression may suggest a shared biological vulnerability (Steingart & Herrmann, 1991). Awareness of deficits caused by dementia may engender dysthymia (Migliorelli et al., 1995; Seltzer, Vasterling, & Buswell, 1995). Some observers feel that depressed mood correlates negatively with degree of impairment, which intimates the role of awareness, although others have found no association between awareness and illness (Ballard, Cassidy, Bannister, & Mohan, 1993).

7.09.4.3 Interpersonal Relationships and Stressful Life Experiences

Both diagnosed depression and depressive symptoms have been found to be related to interpersonal difficulties (Hinrichsen & Zweig, 1994). This association may reflect three primary processes: (i) depression strains and disrupts interpersonal relationships, (ii) depression influences an individual's perception of social relationships; and (iii) unsupportive, demanding, sparse or nonreciprocal, nonequitable interpersonal interactions precipitate depression. Whether or not depression is an outcome of these processes seems to be determined, in part, by an individual's level of assertiveness (Kogan, Van Hasselt, Hersen, & Kabacoff, 1995), self-efficacy, or personal control (Krause, Hertzog, & Baker, 1992), degree of social embeddedness (Adelmann, 1994) and concurrent stressors.

What is known about interpersonal difficulties of depressed older adults tends to be derived from studies of social support that often are examined in the context of stressful life events. These events provide the occasion for emotional support and tangible assistance (Dean, Kolody, & Wood, 1990; Russell & Cutrona, 1991). Stressful life experiences include discrete, disruptive negative life events, chronic strains, and daily hassles. They may or may not be predictable or under a person's control. Often a discrete event precipitates chronic strains or hassles. Some common events with which older adults often must cope are widowhood, deaths of family and friends, caregiving, and the need for institutionalization. These events have been found to be associated with depression (see Murrell & Meeks, 1992 for a review). Declining health, increasing disability, and cognitive dysfunction are also stressful experiences that have been discussed earlier.

As with declining health and cognitive dysfunction, the causal direction of the association between stressful experiences, disrupted interpersonal relationships, and depression is controversial. However, a likely shared mechanism for depression is the depletion of resources (Hobfoll & Wells, in press). Identity, independence, and perceived control or efficacy can be threatened. Stressful experiences may result in disrupted relationships. Depressed individuals may act in ways that alienate their support network, or losses in a support network can precipitate depression. These processes all have implications for psychotherapeutic intervention.

One of the most important loss events is bereavement. Depressive symptoms following loss of spouse vary in relation to characteristics of the death (Gilewski, Farberow, Gallagher, & Thompson, 1991), quality of the relationship and degree of attachment with the deceased (Levy, Martinowski, & Derby, 1994; Murrell & Himmelfarb, 1989), concurrent stressors (Levy et al., 1994), timing and type of social support received (Schuster & Butler, 1989), and perceived competencies (Lund, Caserta, Dimond, & Shaffer, 1989).

Depressive symptoms or bereavement-related depression have been distinguished from grief, or the psychological and physiological reactions to the loss (Marwit, 1991), although both responses are time-limited. However, sometimes grief becomes complicated and the individual fails to return to preloss levels of well-being (Prigerson et al., 1995). Symptoms include crying, being stunned by loss, searching and yearning for the deceased, disbelief about the death, and preoccupation with thoughts of the deceased.

The distinction between grief and depression may explain some inconsistencies in previous studies on depression and loss events. Few investigations have examined the relationship between complicated grief and depression. Gilewski et al. (1991) found that levels of depression at the time of bereavement predicted subsequent adjustment. Hays, Kasl, and Jacobs (1994) and Lichtenstein, Gatz, Pedersen, Berg, and McClearn (1996) observed elevated depressive symptoms prior to the actual loss, with no significant difference between the period directly before the death and the period right after the loss. This pattern seems to be consistent with anticipatory grieving, a process that others have suggested may facilitate adaptation to bereavement (Levy et al., 1994).

7.09.4.4 Genetic Risk

Heritability of major depressive disorder has been studied using family, twin, and adoption designs in younger adults but not in older adults. It is generally accepted that there is some genetic involvement in etiology, that recurring episodes of depression suggest a more heritable form of the disorder, and that early-onset depression has greater genetic determination than late-onset depression (Moldin, Reich, & Rice, 1991). Baldwin and Tomenson (1995) found family history of depression in about a third of early-onset cases compared with 5% of late-onset cases (mean age of onset = 74 years).

In adult samples, Torgersen (1986) reported concordance rates of 27% for monozygotic (MZ) and 12% for dizygotic (DZ) twin pairs from a Norwegian sample and he estimated the heritability of major depression to be 54%. In a Danish twin sample, Bertelsen, Harvald, and Hauge (1977) found concordance rates for unipolar depression of 39% for MZ and 12.5% for DZ pairs. In the Swedish Adoption/Twin Study of Aging, with twin pairs all aged 50 and older, concordance rates for scoring 16 or higher on the CES-D were 33% in MZ and 12% in DZ pairs (Gatz, 1992). Using a British twin register, McGuffin, Katz, Watkins, and Rutherford (1996) reported concordance rates of 46% in MZ and 20% in DZ pairs for *DSM-IV* diagnosed major depression, resulting in heritability estimates of 48–75% dependent upon true population prevalence.

Heterogeneity in etiology is likely. Torgersen (1986) found that, while partners of probands with major depression were apt to have major depression, dysthymic disorder, or depressive adjustment disorder, concordance rates for dysthymic disorder were low in both MZ and DZ twin pairs. Kendler, Neale, Kessler, Heath, and Eaves (1992), using adult twins, suggest that genetic liability to major depression and to generalized anxiety disorder is influenced by the same genetic factors but different environmental experiences.

Consistent with the notion that genetic factors for early-onset depression are distinct from those for late-onset depression, van Ojen et al. (1995a, 1995b) found that older adults with depression accompanied by cognitive impairment were likely to have no history of psychiatric illness themselves and a family history of dementia or of Down's syndrome rather than a family history of depressive disorder. Those depressed older adults who reported both personal and family history of mental illness did not have impaired cognitive performance.

More than major depression, manic-depressive disorder has generally been regarded as running in families, and twin studies have suggested greater genetic influence for bipolar than for unipolar disorder (Bertelsen et al., 1977; Torgersen, 1986). Shulman et al. (1992) found that elderly patients admitted to a hospital for mania were more likely to have a family history suggestive of genetic predisposition compared with elderly admissions for major depression.

7.09.5 ASSESSMENT

In this section the focus is on psychometric properties, measurement efficiency, and measurement invariance of instruments used to assess depression. There follows a discussion on how validity is influenced by uncertainties about the diagnostic criteria in older adults and whether depression is the same across age and within different subgroups, defined by gender, ethnicity, or comorbid conditions (for other reviews and recommendations about measuring depression in older adults, see Pachana, Gallagher-Thompson, & Thompson, 1994, and Scogin 1994).

The efficiency of an assessment instrument can be gauged in terms of several metrics. Sensitivity is proportion of correctly identified depressed cases; specificity is proportion of nondepressed cases correctly identified; positive predictive power (PPP) is probability that a score at or above the clinical cutoff will correspond to a diagnosis of depression; negative predictive power (NPP) is probability that a score below the cutoff will correspond to the absence of a diagnosis of depression. PPP and NPP are influenced by population prevalence. The efficiency of any given cutoff will change with the base rate for depression. Because base rates are known to vary with cognitive impairment, physical illness and gender, efficiency can be affected. Variability in efficiency will also result when there are differences between groups in the underlying construct.

Assessment instruments are grouped on the basis of whether they are self-report or interview-based and by the population with whom they were developed. Interview-based measures include observer-rated measures and structured interviews. Structured interviews have two general formats: one applies diagnostic algorithms upon completion; the second incorporates the algorithms into the structure of the interview (i.e., if a required criterion is not met, the remaining questions are skipped).

7.09.5.1 Self-report Instruments

7.09.5.1.1 Depression scales developed with mixed-age samples

The Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) includes 21 items, each including four alternative statements ranging in order of severity from zero to three. Conventional cutoffs are 0–9 for normal range, 10–18 for mild to moderate depression, 19–29 for moderate to severe depression, and 30–63 for severe depression (Beck, Steer, & Garbin, 1988). For the 13-item short form (Beck & Beck, 1972), 0–4 corresponds to none or minimal depression, 5–7 to mild depression, 8–15 to moderate depression, and 16 or higher to severe depression.

The BDI has demonstrated test–retest and split-half reliability as well as concurrent validity with older adults (see Gallagher, 1986). With a cutoff of 10, the BDI has shown good efficiency with older depressed psychiatric outpatients (100% sensitivity, 96% specificity, 100% PPP, 100% NPP; Olin, Schneider, Eaton, Zemansky, & Pollock, 1992). In a sample seeking outpatient mental health services, Kogan, Kabacoff, Hersen, and Van Hasselt (1994) found that a cutoff of 22 maximized efficiency, yielding 64% sensitivity and 73% specificity.

The BDI has good efficiency in older medical outpatients (Rapp, Parisi, Walsh, & Wallace, 1988) and medical inpatients (Norris et al., 1987). Cognitively impaired individuals have been found to have difficulty comprehending the format (Lyons, Strain, Hammer, Ackerman & Fulop, 1989) and the BDI may be less sensitive to detecting major depression in older men when compared with older women (Allen-Burge, Storandt, Kinscherf, & Rubin, 1994).

The other conspicuous feature of the BDI lies in its inclusion of somatic symptoms. Zemore and Eames (1979) demonstrated that age differences on the BDI could be explained entirely by the somatic symptoms, with no age differences on the more cognitive symptoms.

The Zung Self-rating Depression Scale (SDS; Zung, 1965) contains 20 items, 10 worded positively, and 10 negatively. Items are on a one to four scale; raw scores are converted into a percentage by dividing the sum of all questions by 0.80. Conventional cutoffs are 25–43 for no depression, 50–59 for mild to moderate depression, 60–69 for moderate to severe depression, and 70 or over for severe depression (Zung, 1967). Older adults have shown higher normal baseline scores than the validation sample (Zung, 1967) which may be attributed to the large number of items about somatic symptoms (Berry, Storandt, & Coyne, 1984). Reliability has been found to be moderate or low, particularly among the oldest old, and sex and age differences in factor structure have been suggested (Kivela & Pahkala, 1986; Kivela & Pahkala, 1987; McGarvey, Gallagher, Thompson, & Zelinski, 1982). In medically ill inpatients, the SDS has been found to have a 58% sensitivity and 87% specificity using a cutoff score of 60 (Kitchell et al., 1982) in comparison with an 82% sensitivity–87% specificity in medical outpatients (Okimoto et al., 1982). A 10-item interviewer-assisted form has been developed and validated for older adults (Tucker, Ogle, Davison, & Eilenberg, 1986). A cutoff score above 70 is recommended for the 10-item form.

The Center for Epidemiological Studies—Depression Scale (CES-D; Radloff, 1977) is a 20-item measure designed for large-scale epidemiological surveys. It can be self-administered, by paper-and-pencil or using computer speech recognition (Muñoz, Gonzales, & Starkweather, 1995), or used in an interview, face-to-face or by telephone (George, Blazer, Hughes & Fowler, 1989). Four items are worded positively and are reverse scored; these items have a particularly high rate of omission or response bias (Callahan & Wolinsky, 1994; Gatz & Hurwicz, 1990). Scores range from zero to 60. A score of 16 or more is used as the conventional cutoff for clinical depression, although a cutoff score of 20 may have increased efficiency with older adults (see Radloff & Teri, 1986; Murrell, Himmelfarb, & Wright, 1983). Several shorter forms have been developed specifically for use with older adults by collapsing response categories and giving only 10 or 11 of the items (Andresen, Malmgren, Carter, & Patrick, 1994; Kohout, Berkman, Evans, & Cornoni-Huntley, 1993). The CES-D has probable measurement invariance across age (Hertzog, Van Alstine, Usala, Hultsch, & Dixon, 1990), culture (Heikkinen et al., 1995), and physical health or disability (Davidson et al., 1994). Although the same broad factors emerge, there are differences across gender and ethnic or racial subgroups with respect to factor loadings and placement of specific items on factors. Garcia and Marks (1989) and Callahan and Wolinsky (1994) emphasize that these differences imply sociocultural differences in the expression of depression.

Two self-report forms of the observer-rated Hamilton Rating Scale for Depression (HRSD) have been developed: the Carroll Rating Scale for Depression (CRS; Carroll, Feinberg, Smouse, Rawson, & Greden 1981; Feinberg, Carroll, Smouse, & Rawson, 1981; Smouse, Feinberg, Carroll, Park, & Rawson, 1981) and the Hamilton Depression Inventory (HDI; Reynolds & Kobak, 1991). In the CRS, 52 "yes/no" items replace the HRSD severity ratings. A score of 10 is the recommended clinical cutoff. Koenig, Meador, Cohen, and Blazer (1988, 1992) examined a 12-item short version of the CRS in medical inpatients

(BCDRS; Duke Depression Evaluation Schedule for the Elderly, 1988). A cutoff of six yielded the best efficiency.

The Inventory to Diagnose Depression (IDD; Zimmerman & Coryell, 1987; Zimmerman, Coryell, Corenthal, & Wilson, 1986) is a 22-item instrument designed to diagnose major depressive disorder according to *DSM-III* criteria and to provide information about symptom severity and duration. Although the IDD has been used with older adults, psychometric properties are unknown.

7.09.5.1.2 Instruments devised specifically for use with older adults

The Geriatric Depression Scale (GDS; Yesavage et al., 1983) includes 30 "yes/no" items, 10 reverse scored. A cutoff of 11 demonstrates good sensitivity (84%) and specificity (95%; Yesavage et al.) and adequate hit rates (17% false-negative rate, 18% false-positive rate; Dunn & Sacco, 1989) in community-dwelling older adults; good overall efficiency in psychiatric outpatients preselected for major depression (96% sensitivity, 96% specificity, 96% PPP, 96% NPP; Olin et al., 1992); and good sensitivity (89%) but poorer specificity (75%) and positive predictive value (27%) in older medical inpatients when compared with psychiatric interview (Koenig et al., 1988; Koenig, et al., 1992). Others suggest a cutoff of 14–16 for best efficiency (Burke, Houston, Boust, & Roccaforte, 1989; Kogan et al., 1994). Other evidence suggests that the GDS maintains its reliability and validity when administered over the telephone (Burke, Roccaforte, Wengel, Conley, & Potter, 1995).

The GDS has been found less effective in the detection of major depression in older men when compared with older women (Allen-Burge et al., 1994). It appears to be reliable and valid for older, medically ill outpatients (Norris et al., 1987), hospitalized individuals (Lyons et al., 1989; Rapp et al., 1988), and noncognitively impaired nursing home residents (Lesher, 1986), although some items are sensitive to health status (Fischer, Rolnick, Jackson, Garrard, & Luepke, 1996). There is some suggestion that it is insensitive to detecting minor depression (Harper, Kotik-Harper, & Kirby, 1990; Parmelee et al., 1989). Some older respondents have expressed dissatisfaction with the "yes/no" format (Fischer et al., 1996).

Use of the GDS in cognitively-impaired older adults is equivocal. When identifying depression in mildly to moderately demented individuals, Lichtenberg, Marcopulos, Steiner, and Tabscott (1992) found that a cutoff of 11 yielded good sensitivity (82%) and specificity (88%). With geriatric assessment center clients, Burke, Nitcher, Roccaforte, and Wengel (1992) found that a cutoff of 14 yielded the highest combined sensitivity (70%) and specificity (78%). Other studies show that the GDS does little better than chance as a screening measure for depression in cognitively-impaired individuals (Burke et al., 1989; Kafonek et al., 1989).

A 15-item short form of the GDS (Lesher & Berryhill, 1994; Sheikh & Yesavage, 1986) retains the same "yes/no" format and the same ratio of positively- and negatively-worded items. A cutoff of five is recommended. The GDS short form has good efficiency with medically ill skilled nursing facility residents (Baker & Miller, 1991) and older primary care surgery attenders (sensitivity 91%, specificity 72%, NPP 94%; D'ath, Katona, Mullan, Evans, & Katona, 1994), but poorer efficiency in ethnically heterogeneous populations (Cwikel & Ritchie, 1988) and in cognitively-impaired individuals (Burke, Roccaforte, & Wengel, 1991; Lesher & Berryhill, 1994).

7.09.5.1.3 Multidimensional measures with depression subscales

The 567-item Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) includes a 57-item depression scale. Little is known about the measurement properties of the MMPI-2-D with older adults. One 15-item short form of the MMPI-D was developed and construct-validated using a sample of older adult male inpatients (Dush & Hutzell, 1986). Another short form has been used for detecting depression in cognitively-impaired older adults (Ward, Wadsworth, & Peterson, 1994).

The Symptom Checklist 90—Revised (SCL-90-R; Derogatis, 1977), as well as its precursors, the SCL-90 and the Hopkins Symptom Checklists (HSCL), assess distress on a wide range of psychiatric symptoms. The depression subscale on the SCL-90-R comprises 13 items. A 53-item short version of the SCL-90, entitled the Brief Symptom Inventory (BSI; Derogatis & Melisaratos, 1983), includes the depression dimension, reduced to six items. Norms have been published for use with community-dwelling older adults (Hale, Cochran, & Hedgepeth, 1984) and with caregivers (Anthony-Bergstone, Zarit, & Gatz, 1988). The BSI depression subscale excludes all somatic items. For medically ill patients, high false-negative rates have been reported, particularly those with minor depression (Harper et al., 1990).

Other self-report mood instruments that sometimes appear in the literature are the Multiple Affect Adjective Check List (MAACL:

Zuckerman & Lubin, 1965; MAACL-R: Zuckerman & Lubin, 1985) and the Profile of Mood States (POMS; McNair, Lorr, & Droppleman, 1971/1981/1992) or a shortened form (Guadagnoli & Mor, 1989). Properties with older adults are not well known.

7.09.5.2 Observer-rated Inventories

The Hamilton Rating Scale for Depression (Hamilton, 1960; Hamilton, 1967) is a 17-item instrument that was designed to measure frequency and intensity of depressive symptoms in individuals with major depressive disorder. Ratings are made using either a five- or a three-point scale, yielding total scores from zero to 61. Conventional clinical cutoff scores range from 14 to 20, with 17 being the most common in pharmacological research (see Reynolds & Kobak, 1991). The HRSD has demonstrated reliability, validity, and efficiency in older adults (Stukenberg, Dura, & Kiecolt-Glaser, 1990).

Cognitively-impaired, medically ill older adults have shown difficulty understanding, responding to, and completing the questions (Baker & Miller, 1991) and use of the HRSD for identifying depression in demented individuals has mixed results (Lichtenberg et al., 1992; Vida, Des Rosiers, Carrier, & Gauthier, 1994a). With a cutoff of 17, the HRSD showed 9% sensitivity and 92% specificity in mild to severely demented hospitalized patients. However, a cutoff of seven maximized efficiency, yielding 90% sensitivity and 63% specificity. A structured interview guide has been developed for the HRSD (SIGH-D; Williams, 1988) to standardize the way in which rating information is obtained; its use in older adults is unknown.

Another observer-rating inventory that more often appears in European literature is the Montgomery–Åsberg Depression Rating Scale (MADRS; Montgomery & Åsberg, 1979).

The Geriatric Depression Rating Scale (GDRS; Jamison & Scogin, 1992) was developed specifically for older adults. It includes 29 of the 30 items on the GDS and six somatic items to conform to RDC criteria. A cutoff of 20 can be used to discriminate depressed and nondepressed individuals (88% sensitivity, 82% specificity).

7.09.5.3 Structured Interviews

7.09.5.3.1 Instruments designed for mixed ages

The Present State Examination (PSE; Wing, Birley, Cooper, Graham, & Issacs, 1967) uses a standardized question format from which a computer program (CATEGO) yields behavioral syndromes that approximate the *ICD* diagnostic categories.

The Schedule for Affective Disorders and Schizophrenia (SADS; Endicott & Spitzer, 1978) is a semistructured interview developed in conjunction with the RDC. There are three versions: regular SADS, which queries severity of symptoms during the preceding week and at the time when they were the most severe during the current episode; lifetime version (SADS-L); and the version for measuring change (SADS-C). The SADS has been used with older adults. Efficiency is good (82% sensitivity and 93% specificity; Gallagher, Breckenridge, Steinmetz & Thompson, 1983) when compared with the BDI as a criterion measure.

The National Institute of Mental Health Diagnostic Interview Schedule (DIS; Robins, Helzer, Croughan, & Ratcliff, 1981) is a structured interview used in the ECA. Both current and lifetime diagnoses using *DSM-III* and RDC criteria can be made by computer.

The Composite International Diagnostic Interview (CIDI; Robins et al., 1988) was designed for cross-cultural epidemiologic and comparative studies of psychopathology in general populations, whereas its companion, the Schedules for Clinical Assessment in Neuropsychiatry (SCAN), was intended for use in clinical samples. The CIDI is constructed in diagnostic modules that can be extracted and administered on their own. These modules yield both *DSM-III-R* and PSE principal diagnoses, each with its own set of diagnostic computer programs. CIDI short screening scales are available for quicker administration (Kessler & Mroczek, 1993).

The Structured Clinical Interview for *DSM-III-R* (SCID; Spitzer, Williams, Gibbon, & First, 1992) provides current and lifetime diagnoses. It has been revised for *DSM-IV*. It is intended to be administered by clinicians and makes allowance for clinical judgement in modifying and supplementing the interview (Spitzer et al.). Use of information from multiple sources to inform rating decisions is encouraged. There are two versions of the SCID: the SCID-P (Patient) and the SCID-NP (Nonpatient). The major depression module of the SCID-P has been studied with older adults (Segal, Herzen, Van Hasselt, Kabacoff, & Roth, 1993).

7.09.5.3.2 Instruments designed for older adults

The Geriatric Mental State examination (GMS; Copeland et al., 1976) is a semistructured interview from which the AGECAT computer program is used to assign diagnoses and confidence ratings for each diagnosis.

Many of the 541 items were derived from the PSE. AGECAT depression diagnoses have been shown to have a sensitivity of 80%, a specificity of 89%, and a PPP of 71% in physically ill hospitalized older adults (Ames, Flynn, Tuckwell, & Harrigan, 1994).

The Geriatric Mental State Schedule-Depression Scale (GMSS-DS; Ravindran, Welburn, & Copeland, 1994) is a 33-item rating scale derived from the GMS, designed to be sensitive to symptomatic changes in depression with treatment in older adults.

The Comprehensive Assessment and Referral Evaluation (CARE; Gurland et al., 1977–78) uses a shortened form of the GMS as part of a comprehensive assessment package of health and social problems of elderly community residents. There is a 19-item depression/demoralization scale on the CORE–CARE (Golden, Teresi, & Gurland, 1984). The GMS–AGECAT and CARE are used in differential diagnosis of depression and dementia.

7.09.5.4 Assessing Depression in Cognitively-impaired Older Adults

Measures developed specifically to assess depression in cognitively-impaired individuals tend to use simple formats and to incorporate informant information and direct behavior observations. Information from collaterals has been demonstrated to provide an accurate, independent indicator of depression (Gilley et al., 1995; Logsdon & Teri, 1995), although discrepancies from patient report or unavailability of a collateral can pose problems in interpretation (Gilley et al.; Moye, Robiner, & Mackenzie, 1993; Ott & Fogel, 1992; Teri & Wagner, 1991).

Measures include the Cornell Scale for Depression in Dementia (CSDD; Alexopoulos, Abrams, Young, & Shamoian, 1988), the National Institute of Mental Health Dementia Mood Assessment Scale (DMAS; Sunderland et al., 1988), Columbia University Scale for Psychopathology in Alzheimer's disease (CUSPAD; Devanand et al., 1992), Canberra Interview for the Elderly (CIE; MacKinnon et al., 1993), Cambridge Mental Disorders of the Elderly Examination (CAMDEX; Roth et al., 1986), the Neurobehavioral Rating Scale (NRS; Sultzer, Levin, Mahler, High, & Cummings, 1992), and the Psychogeriatric Assessment Scales (PAS; Jorm et al., 1995).

The overlap of symptoms of depression and dementia raises difficult questions for the diagnostician. Vida, Des Rosiers, Carrier, and Gauthier (1994b) examined the validity of RDC criteria for major and minor depression in individuals with Alzheimer's disease (AD). They found that three criteria were not associated with major depression (difficulty with concentration or thinking, excessive self-reproach or guilt, anhedonia or apathy), whereas nonverbal manifestations of depression, resentfulness, irritability, anger or complaining, and demandingness or clinging dependency, showed the strongest association with major depressive disorder. Diagnostic criteria for minor depression remained valid.

7.09.5.5 Assessing Depression in Older Adults with Poor Physical Health

Assessment of depression in individuals with comorbid medical illness or disability has been an area of focused attention because of ambiguity in the interpretation of somatic symptoms. Several strategies have been undertaken: (i) development of new instruments, (ii) development of instruments that eliminate somatic symptoms or substitute cognitive/affective symptoms for somatic ones, and (iii) use of alternative diagnostic schemes that count all criterion-related symptoms and do not require judgments about etiology.

Koenig, Cohen, Blazer, Meador, and Westlund (1992) developed a scale with 11 items derived from the GDS and the CRS, selected for their predictive validity in medically ill individuals. When examined in a sample of men over age 65 years admitted to a medical and neurological VA service, a cutoff score of three yielded a sensitivity of 83% and specificity of 77%. The scale showed moderate correlations with other self-report measures. Evans (1993) developed the 15-item Evans Liverpool Depression Rating Scale (ELDRS) for detection of depression in physically ill older adults. It obtains both patient and collateral information in "yes/no" format. The ELDRS has been found to have adequate interrater reliability and good efficiency using a cutoff of 5/6 (sensitivity of 88–91% and specificity of 85–96%).

Rapp and Vrana (1989) examined the sensitivity and specificity of conventional SADS/RDC depression diagnoses, which exclude somatic symptoms judged to be secondary to organic condition, compared with a modification of the RDC that reassigned diagnoses after substituting nonsomatic symptoms for somatic symptoms. Fully 96% of individuals were classified the same way with both sets of criteria; the alternative criteria showed 87% sensitivity and 100% specificity. Eliminating the somatic symptoms without replacement, however, resulted in poorer classification. Koenig, Pappas, Holsinger, and Bachar (1995) evaluated the reliability of six diagnostic strategies varying in inclusivity. Some strategies eliminated fatigue and anorexia. Some replaced four somatic symptoms (fatigue, anorexia/weight loss, psychomotor retardation, impaired concentration) with four cognitive/affective symptoms (irritability, tearfulness, social withdrawal, feeling punished). When structured interviews were used, clinicians were able to determine reliably whether major depression was present regardless of diagnostic strategy. These studies suggest that there is no advantage to abandoning the *DSM-IV* in medically ill older adults, but that modified criteria may be easier and more accurate than the conventional system requiring judgments about the cause of somatic symptoms.

7.09.6 SUICIDE

A discussion of affective disorders in late life would be incomplete without consideration of suicide among this population. Death by suicide is more common among older adults than at any other age. Although most depressed older adults do not commit suicide, many older adults who do commit suicide are depressed.

7.09.6.1 Rates

The suicide rate for older adults in most countries is higher than that of any other age group (Gulbinat, 1996). Among nations reporting to the World Health Organization, rates of suicide for older men are up to seven times higher than adolescent male suicide rates (Gulbinat, 1996). Male suicides among all ages are two to three times more frequent than female suicides (Gulbinat, 1996). Rates also vary by race and ethnic background. In the US, suicide rates are highest among Native Americans and Whites, followed by people of Asian origin and Blacks, with people of Latino origin having lower rates of suicide than non-Latino origin. Gender, racial, and ethnic differences in suicide rates are most pronounced among older adults (McIntosh, 1992).

7.09.6.2 Methods

Methods used by older adults to commit suicide are generally less amenable to intervention than those used by younger age groups (De Leo & Diekstra, 1990). In the US, guns are used to commit suicide more frequently than any other method, both by men and women (Kaplan, Adamek, & Johnson, 1994; McIntosh, 1992). This is true of all ages, but to the greatest extent in late life. In Finland, hanging is the leading method of suicide for all ages, with the proportion of suicides using this method greatest in old age. In most Western countries, over 85% of older adult suicides involve hanging, poisoning, or drowning, with hanging more common in men and poisoning more characteristic of women (De Leo & Diekstra, 1990). In Japan, the most frequently used methods in late life are hanging, drowning, and jumping.

A unique feature of older adult suicide is the small window of opportunity for intervention. The ratio of attempted to completed suicides among older adults has been estimated at 4:1, whereas this ratio among all ages is estimated at 20:1, and among youth it is thought to be as high as 200:1 (McIntosh, 1992; Parkin & Stengel, 1965). This difference could reflect several factors: the selection of less reversible means, which in itself can be a sign of increased lethality of intent, decreased powers of recuperation (Lindesay, 1991), or isolation from rescue. Older adults have also been found to be less likely to express suicidal ideation than younger adults (Blazer et al., 1987), a factor further complicating assessment efforts.

7.09.6.3 Depression and Suicide

Depression and suicide are linked closely throughout the lifespan. Individuals who are depressed are more likely to commit suicide, and individuals who committed suicide were more likely to have been depressed. On a neurobiological level, the same serotonin pathways have been implicated in depression and suicide, particularly in violent suicides (Åsberg, Nordström, & Träskman-Bendz, 1986). Moreover, the association between depression and suicide is strongest in late life.

In a review of follow-up studies, Miles (1977) concluded that 15% of depressed individuals of all ages would eventually die by suicide. Gardner, Bahn, and Mack (1964) calculated suicide rates for individuals with affective psychosis to be 207/100 000 for those under age 55 years and 475/100 000 for those aged 55 years or older. Psychological autopsy studies in England and the US have found evidence for affective disorders in 83–87% of older adults who committed suicide (Barraclough, Bunch, Nelson, & Sainsbury, 1974; Conwell, Olsen, Caine, & Flannery, 1991), while a psychological autopsy study in Finland found depressive disorders in 67% of suicides who were over age 60 years (Henriksson, Marttunen et al., 1995). Substance use disorders are also prevalent in older suicides (either alone or comorbid with affective disorders), but they are less common than among younger suicides (see Conwell & Brent, 1995, for a review; Henriksson, Marttunen et al., 1995). Although differential mortality could be an alternative explanation for the lower rates of substance use disorders among older adult suicides, this would not explain the higher rates of affective disorders. As Conwell and Brent (1995) concluded, since affective disorders are less prevalent in late life but more common among older completed suicides, older adults who are depressed seem to be at higher risk of suicide than depressed younger adults.

7.09.6.4 Hopelessness

Hopelessness has been studied extensively as a possible correlate of suicide. Hopelessness has been linked to suicidal ideation (Wetzel, Margulies, Davis, & Karam, 1980) and completed suicides (Beck, Steer, Kovacs, & Garrison, 1985; Fawcett et al., 1990).

7.09.6.5 Issues in Late Life Suicide

A growing literature has documented the phenomenon that older adults who commit suicide frequently have seen a physician within a short interval before the suicidal death. In psychological autopsy studies of older adult suicides, Conwell et al. (1991) found that 75% had seen their physician within a month of the suicide. Takahashi et al. (1995) found that older adult inpatients on a psychiatric ward who were admitted after a suicide attempt were three times more likely to have seen a physician within the previous two weeks than older psychiatric inpatients who had not made a suicide attempt.

Another important issue concerns suicide among the medically ill. Suicides in this population have been described as rational, particularly in the presence of chronic pain or a poor prognosis (Battin, 1991). Although few studies have explored suicidal behavior in chronic or terminally ill individuals, the existing research indicates that depressive disorders are prominent in these cases (Henriksson, Isometsä, et al., 1995; Leibenluft & Goldberg, 1988). This result highlights the need to assess carefully for affective disorders and for suicide risk in medically ill individuals. Since pain has been associated with depressive disorders (Parmelee et al., 1991), effective treatment of pain may be an avenue for reducing both depression and suicide risk among medically ill older adults.

7.09.6.6 Assessment and Treatment

Assessment of suicide risk in older adults is an inexact science. Although older adults are at highest risk of suicide, most suicide research has focused on younger populations. There is ample reason, however, to doubt that findings would necessarily translate to an older suicidal population. As discussed earlier, older adult suicides differ in several important ways from younger suicides, including a higher degree of lethality and decreased likelihood of expressed suicidal ideation. There are no widely accepted assessment instruments validated with older individuals (for approaches to suicide risk assessment in this population, see Wolfe, Morrow, & Frederickson, 1995).

Even less research is available on the efficacy of interventions for suicidal older adults, or indeed for suicidal individuals of any age. In the absence of such information, the most promising approach in working with suicidal individuals with affective disorders is the effective treatment of these affective disorders.

7.09.7 COURSE OF DEPRESSION

In monitoring the course of depression, individuals can be classified as recovered (partial or full remission), relapsed, or chronic (nonremitted). Some have suggested that older adults show worse prognosis both in terms of lower rates of recovery and greater chance of relapse, but differences in study designs or sample characteristics make alternative explanations replete. When the differences are controlled, age in and of itself seems to have little prognostic value. Hinrichsen (1992) examined recovery and relapse rates in elderly depressed inpatients and compared them with those observed in a mixed age sample, using the same diagnostic and follow-up methods and the same uncontrolled treatment approach. He found that the rates were comparable: by one year after hospital admission, 59% of the elderly individuals achieved full or partial sustained recovery, 13% recovered and subsequently relapsed into the same condition as the index episode, and 28% never recovered. The figures for the mixed age sample were 58%, 16%, and 26%, respectively.

Kennedy, Kelman, and Thomas (1991) examined factors that contributed to chronicity, remission, and emergence of depressive symptoms in elderly community residents. Baseline health status was less predictive than change in health. Declining health and increasing disability were associated with chronicity, whereas improved health was associated with remission. This association could reflect a psychological reaction to decline or a more general wearing down process.

Diagnosed depression and depressive symptoms are associated with increased risk of mortality in institutionalized and inpatient samples (Evans, 1993; Parmelee, Katz, & Lawton, 1992) and in community residents (Thomas, Kelman, Kennedy, Ahn, & Yang, 1992). However, the explanation remains unclear. For example, the increased mortality may be related to the presence of cognitive impairment, that is, a comorbid dementia (Snowdon & Lane, 1995).

Burvill and colleagues examined mortality rates in a cohort of elderly depressed inpatients at the time of the index depressive episode, one year later, and then at five years (Burvill & Hall, 1994). Poor outcome of the index episode after one year (i.e., "depressive invalids" who had moderate symptoms of depression but not enough to merit diagnosis of major depression, continuously ill, or demented) was associated with increased mortality over the subsequent four years. Individuals most at risk of death were those who, on entry into study, were male, aged 75 years or older, had a diagnosis of major depression without melancholia or psychosis, and had substantial health and mobility problems.

Evans (1993), in contrast, found that one year death rates of depressed patients 12 months after inpatient treatment did not differ significantly from those of psychologically well individuals. Only timing of deaths varied, with most deaths occurring in the first three months for depressed individuals compared to being evenly distributed in the nondepressed group.

7.09.8 TREATMENT APPROACHES

The array of treatments provided to depressed older adults largely parallels those used to treat depression in adults generally. These can be divided into psychosocial treatments, predominantly psychotherapy, and somatic treatments, that is, psychopharmacological treatments and electroconvulsive therapy (ECT). There do appear to be some age differences in the distribution of treatments. For example, psychosocial interventions based on life review or reminiscence are exclusive to older adults. ECT has registered an increase in its use, disproportionately accounted for by older adults (Warren & Levy, 1991).

The literature is just beginning to have sufficient information to give good guidance. In the past, studies have not always distinguished between treating a disorder and endeavoring to raise scores on depression self-report scales for community volunteers who were not necessarily depressed to begin with. In fact, this failure may contribute to a misperception perpetuated by the Consensus Development Conference Statement (Friedhoff, 1994) that psychotherapy is for mild (nonserious) depression, while somatic treatments are for true depression. Because the topic of the chapter is mood disorders, for the purposes of discussing efficacy of treatment, the focus is on studies with individuals who have a diagnosed disorder. However, to reflect interest in variations of disorder observed in older adults, literature concerning subdysthymic disorder is included.

7.09.8.1 Psychotherapeutic Interventions

Psychosocial treatments for depression encompass psychodynamic psychotherapy, brief psychodynamic therapy, interpersonal therapy, behavior therapy, cognitive and cognitive behavior therapy, life review, marital and family therapy, and group therapy (which can be psychodynamic or cognitive-behavioral). Empirical evidence to date suggests that efficacy of treatments for older adults is basically similar to efficacy for adults generally. However, the fact that older adults are not routinely included in evaluations of interventions is disturbing, for example, the Treatment of Depression Collaborative Research Program (Elkin et al., 1989).

There are several available reviews of psychological treatments for depression, including Scogin and McElreath's (1994) meta-analysis and chapters by Futterman, Thompson, Gallagher-Thompson, and Ferris (1995), Niederehe (1994), Smyer, Zarit, and Qualls (1990), Teri, Curtis, Gallagher-Thompson, and Thompson (1994), and Wolfe, Morrow, and Frederickson (1995). Clinical psychologists have placed more explicit attention on identifying those interventions that are supported empirically, and on establishing criteria by which to make this judgment (Chambless et al., 1996; Gatz et al., in press). While not directly employing this framework, treatments that are supported by good empirical designs are emphasized (e.g., randomized controlled trial).

7.09.8.1.1 Cognitive-behavior therapy

The rationale for behavioral therapy for depression is found in Lewinsohn, Muñoz, Youngren, and Zeiss (1996), and the rationale for cognitive therapy in Beck, Rush, Shaw, and Emery (1979). Behavior therapy teaches the role of reduced reinforcement in creating and maintaining depression, methods to increase pleasant events, relaxation techniques (Bernstein & Borkovec, 1973), and problem-solving skills (Nezu, Nezu, & Perri, 1989). Cognitive therapy teaches the importance of negative, pervasive attributions about oneself, the world, and the future in determining one's mood, and offers methods to monitor and challenge these thoughts. In both cases, the treatment was developed for adults in general. In studying the effectiveness of these therapies, sometimes the two are contrasted with one another; sometimes a combined cognitive-behavioral therapy is compared with another treatment. In practice, it is probably most common to combine elements of both behavioral and cognitive treatments, and the outcome research does not lead us to differentiate among these approaches.

Adaptations for older adults include changing pleasant events lists to be more appropriate to the activities in which older adults are likely to engage and resolving practical barriers such as mobility and vision limitations (Rybarczyk et al., 1992). These adaptations seem unlikely to interfere with comparisons of treatment outcome across age groups.

Gallagher and Thompson have conducted a series of studies that provide the clearest evidence for the efficacy of cognitive and behavioral treatments for patients meeting diagnostic criteria for major depressive disorder. The first study (Gallagher & Thompson, 1982) compared individual cognitive, behavioral, and supportive/insight therapy; the second (Thompson, Gallagher, & Breckenridge, 1987) compared individual cognitive, behavioral, and brief dynamic therapy with a waiting list control group also included in the design. Treatment lasted 20 sessions. At post-test, all therapies registered improvement on self-report and observer-rated scales and on proportions of patients diagnosed as depressed; all therapies were more effective than the control group. At follow-up assessment a year after termination of treatment, in the first study, maintenance of gains was superior for the behavioral and cognitive groups compared with those who had received supportive/insight therapy (Gallagher & Thompson, 1983). In the second study, there was no difference among treatments (Gallagher-Thompson, Hanley-Peterson, & Thompson, 1990).

Two other research groups have included psychopharmacological conditions, but permit evaluation of the psychotherapy component. Beutler et al. (1987) compared various combinations of cognitive therapy and alprazolam (Xanax). Cognitive therapy was administered in small groups. All four groups (including the placebo) improved on both BDI and HRSD. The placebo group and the medication-only group showed less steep improvement on the BDI and a reversal to baseline levels at follow-up.

Steuer et al. (1984) compared cognitive-behavioral to psychodynamic psychotherapy administered in groups. Participants had volunteered for a drug study (Jarvik, Mintz, Steuer, & Gerner, 1982), and psychotherapy groups were by and large filled by those for whom pharmacologic treatment was contraindicated due to a chronic physical disease or to a medication being used to treat such a condition. Periodic assessments over the nine months of psychotherapeutic treatment showed significant improvement on self- and observer-rated scales for both modalities. On the BDI, cognitive-behavioral treatment was superior to psychodynamic. Riskind, Beck, and Steer (1985) used Steuer et al.'s data to demonstrate that nearly 50% of the cognitive therapy patients were either in remission or only mildly depressed at the conclusion of treatment.

These studies have not been replicated by large numbers of other laboratories. Campbell (1992) found significant reductions in depression after individual cognitive therapy combined with reminiscence in comparison with a psychological placebo (a crafts group) and to an untreated control group. Improvement was measured with the Zung SDS and with thought diaries. Arean et al. (1993) compared problem-solving group therapy, group reminiscence therapy, and a waiting list control group. Individuals in both the problem-solving and reminiscence group showed reductions in depression (assessed with HRSD, BDI, GDS) in comparison with the control group, with extent of improvement at post-test and maintenance of gain better for the problem-solving therapy than for reminiscence. Problem-solving therapy was also applied successfully by Hussian and Lawrence (1981) with mildly to severely depressed nursing home residents.

All of these outcome studies had similar exclusion criteria: nonsuicidal, nonpsychotic, no cognitive impairment, or alcohol abuse. In addition, many potential participants with chronic physical illnesses were *de facto* excluded because these studies typically relied on community samples who could come to the facility where treatment was being administered. However, it should not be concluded that depression was not severe. For example, in Thompson et al. (1987), mean BDI was 23.5 at pretest; mean HSRD was 19.1. In Beutler et al. (1987), mean HSRD at pretest was 22.0 across all groups, mean BDI (on the 13-item form) was 11.3. In Steuer et al. (1984), mean BDI (on the 13-item form) was 15.3; mean HSRD was 18.7.

Dropout rates are not always easy to discern from published descriptions. Thompson et al. (1987) report that they had four patients drop out of behavioral (14%), four from psychodynamic (14%), and 10 from cognitive therapy (27%). Dropping out reflected expressed dissatisfaction or more ambiguous causes, such as the client's declaring transportation problems or physical illness. Beutler et al. (1987) reported dropout at 13% in the first four weeks; thereafter, it was 67% for medication and 31% for cognitive therapy, with or without medication. Side effects of the drugs was a key factor for most. Steuer et al. (1984) found dropout rate to be 39% across both types of psychotherapies, with no difference between them, but only 9% before the 12th week.

7.09.8.1.2 Brief psychodynamic, interpersonal, and life review therapies

Brief psychodynamic and interpersonal therapies have resulted in a clinical literature describing special aspects of their application to older adults. For interpersonal therapy, a manual has been prepared for older adults (described in Reynolds et al., 1992). The most salient difference is in the nature of the interpersonal problems. Life review traces its conceptual roots to Erik Erikson's description of the eighth stage in the life cycle, where the individual deals with the achievement of ego integrity. Butler (1974) has provided a framework for making this task more intentional and more therapeutic. There is no comparable therapy for younger adults.

Thompson et al. (1987), Gallagher and Thompson (1982), and Steuer et al. (1984), already described, all showed brief psychodynamic therapy or supportive/insight therapy to be generally equivalent to cognitive-behavioral approaches, sometimes not on all measures or after extended follow-up. Schneider, Sloane, Staples, and Bender (1986) found interpersonal therapy (IPT) to be as effective as nortriptyline and superior to placebo control, indicated by 58% improvement on HDRS for drug and 64% for psychotherapy; in addition, dropout rate was lower for IPT than for medication.

Reynolds et al. (1992) tested the efficacy of IPT combined with nortriptyline, IPT alone, nortriptyline alone, and placebo as a maintenance treatment for depression. Participants were older adult patients who had already been treated in the acute and continuation phases by combined IPT and nortriptyline. No one started maintenance treatment unless they had already improved. Of patients who began the study, there was 19% attrition before the end of the continuation phase, and 16% of those who completed treatment did not improve. Preliminary results showed that 24% of those shifted to placebo suffered relapse, with no significant difference for those in IPT or not in IPT.

A variant of IPT, interpersonal counseling (IPC), has also been used with subsyndromal dysthymia. Mossey et al. (1996) compared IPC to usual care for recently discharged medical patients with GDS scores in excess of 10 at two occasions before treatment began, but who did not meet diagnostic criteria for major depressive disorder. Follow-up at 2–3 months after completing 10 sessions of IPC indicated that those receiving IPC had significantly lower depressive symptoms than the control group.

Fry (1983) compared life review to unstructured reminiscence and found both superior to an attention placebo. Haight (1992) compared reminiscence, using structured life review devised by the author to standardize the intervention, to a friendly visitor program (attention control) and to a no-treatment group. She found improved psychological well-being that was maintained at follow-up but no significant change on Zung SDS. In neither of these studies were *DSM* diagnoses used to screen participants. Haight's sample was homebound and therefore assumed to be at elevated risk of depression.

With participants diagnosed with major depressive disorder, Arean et al. (1993), as has been described, found life review superior to a waiting list control but not to problem-solving therapy. In a nursing home sample, Dhooper, Green, Huff, and Austen-Murphy (1993) found a combined life review and problem-solving group superior to a control group, with depression measured with the Zung SDS. Niederehe (1994) cautions not only has most of the research on life review been conducted with minimally symptomatic samples, but also the technique could have negative effects with individuals who do have major depressive disorder.

7.09.8.1.3 Support groups and psychoeducational interventions for depression in caregivers of dementia patients

A population that has received special attention are caregivers of older adults, especially of older adults with AD. As noted already, caregivers are subject to increased risk of depression as well as more general strain and burden. Moreover, depression often accompanies dementia; thus, both the older adult with dementia and the caregiver are appropriate targets for treatment.

A meta-analysis of interventions for caregiver distress (Knight, Lutzky, & Macofsky-Urban, 1993) concluded that these interventions did have a respectable effect on mood and burden. Although outcome measures often include measures of depression, participants do not always meet criteria for depressive disorder. Elements of intervention with caregivers typically include emotional support and sharing of common experiences, training in problem solving and stress reduction, and education about the disease and community resources. The component concerned with problem solving and stress reduction tends to build on a behavioral model of treatment for depression. Using this framework, for example, Gallagher-Thompson (1994) developed a treatment manual and curriculum, called "Coping with the Blues." She and her colleagues reported significant reduction of depression on the BDI (Lovett & Gallagher, 1988).

An additional element is added by Teri, Logsdon, Wagner, and Uomoto (1994), whose model is to involve both the caregivers and dementia patients. The intervention explicitly teaches the caregivers to use a behavioral model with the dementia patient, focused on the role of pleasant events in maintaining a positive emotional state.

7.09.8.2 Pharmacological Interventions

The most common psychopharmacological treatment for depression in older adults has been the tricyclic antidepressants, amitriptyline (Elavil), imipramine (Tofranil), desipramine (Norpramin), protriptyline (Vivactyl), and nortriptyline (Pamelor). Tricyclics all have anticholinergic side effects, related to degree of muscarinic blockage, manifesting both centrally (delirium, confusion) and peripherally (urinary retention, tachycardia, orthostatic hypotension). The older drugs in this category, such as amitriptyline and imipramine, have more severe side effects, while nortriptyline is regarded as perhaps causing the least orthostatic hypotension, and desipramine is least anticholinergic (Jenike, 1989; Salzman, 1994). Age-related changes in drug metabolism by the liver and in neurotransmitters may also lead to greater sensitivity to both therapeutic and side effects (Naranjo, Herrmann, Mittmann, & Bremner, 1995). The effect of concurrent medications can also be substantial (Reynolds, Schneider, Lebowitz, & Kupfer, 1994). The reason that orthostatic hypotension is a concern lies in risk of falls. One report suggests that risk of hip fracture is tripled by the use of tricyclic antidepressants (Ray, Griffin, Schaffner, Baugh, & Melson, 1987). Cardiac side effects are also a concern (Glassman & Roose, 1994).

Newer, or second-generation, antidepressants with atypical structure, such as trazodone (Desyrel), may be especially sedating although having fewer anticholinergic effects and cardiac side effects. Other newer medications are selective serotonin-reuptake inhibiting (SSRI) antidepressants, for example, fluoxetine (Prozac) and sertraline (Zoloft). These do not have the same cholinergic side effects as tricyclics, however, their cardiac side effects have not received sufficient study (Reynolds et al., 1994). It has been reported that substitution of SSRIs for tricyclics is occurring at a slower rate for patients aged 65 years and older than for patients younger than 65 years (Klawansky, 1994), although the picture may be changing and different settings may have different prescribing habits.

Monoamine oxidase inhibitors also present side effects, including orthostatic hypotension (Naranjo et al., 1995). Lithium, which is commonly used to treat bipolar disorder, requires particularly careful regulation of dosage in older adults, and creates a particular risk of neurotoxicity when the patient is demented (Jenike, 1989).

In their review of drug studies through 1986, Gerson, Plotkin, and Jarvik (1988) conclude that remission rates on the HRSD among those with major depressive disorder averaged 51% across different drugs, predominantly tricyclic antidepressants, with no obvious advantage for one drug over another. This contrasts with a mean improvement of 20% for those on placebo. Dropout rates were 19% for those in the active drug conditions and 31% for those in the placebo condition (average = 21%). Side effects were described as a major impediment.

Klawansky (1994) collected all available randomized controlled trials through 1991 for meta-analysis. Drugs were more effective than placebo and also had more side effects. There was no overall difference in dropout rates between tricyclics and placebo. Those on medication mainly dropped out because of side effects and those in the placebo condition, due to lack of benefit. In comparisons of fluoxetine to tricyclics, dropout was less for those assigned to the fluoxetine condition.

As with psychotherapeutic studies, psychopharmacologic studies tend to feature time-limited treatments and to exclude patients who have various complications. By default, those excluded are often the oldest patients. There is almost no literature on patients aged 80 years and older (Salzman, 1994). In these respects, research may match poorly with actual practice. Thus, it is of interest to look at Katz, Simpson, Curlik, Parmelee, and Muhly (1990), who conducted a double-blind study of nortriptyline with nursing home and congregate housing residents. After excluding those who had contraindications (22%), or refused, or met some other exclusion criterion, patients who completed the drug treatment improved significantly on the HRSD but not on the GDS relative to the placebo control group. Still, of patients given active medication, 11 of 32 (34%) had some sort of adverse event, acknowledging that it is impossible to accurately distinguish side effects from unrelated medical events. Nortriptyline has also been studied with patients who have both depression and dementia (Salzman, 1994), a group for whom cognitive side effects are a particular concern.

Comparisons of pharmacotherapy to psychotherapy are limited. In Jarvik et al. (1982), although assignment to drug and to psychotherapy was not random, the authors note that (i) more patients dropped out of the drug conditions than group therapy, (ii) improvement for psychotherapy tended to be more homogeneous, while improvement for the drug conditions was bimodal, with patients either responding or failing to respond, (iii) group therapy was significantly more effective than placebo drug. Given the concern with side effects among older adults, the recommendation of Wexler and Cicchetti (1992) seems sensible; that is, begin with psychotherapy, and if it is not effective, add medication.

7.09.8.3 Electroconvulsive Therapy

Treatment with electroconvulsive therapy (ECT) is attractive when side effects of medications are of particular concern or the patient has not responded to medication, or when there is particular urgency. Many argue that ECT is actually safer than drugs. The short-term efficacy of ECT with older adults is well accepted (Sackheim, 1994). Relapse does tend to be high, which has led to studies of continuation and maintenance ECT (e.g., Zorumski, Rubin, & Burke, 1988). While drugs are used commonly to prevent relapse after ECT, the same reasons for not turning to drugs in the acute phase would also be applicable when considering drugs for this role.

Rate of complications does increase with age, although ECT is used in those over age 80 years (Cattan et al., 1990). The main side effects are cardiovascular, with these complications generally accounting for ECT-related mortality. Cognitive side effects are a major concern, including the fact that confusion can lead to falls. Brief confusion directly after treatment is standard. Duration of amnesia varies. Severity of cognitive sequelae appears to be related to mode of administration. It is also not established whether those who already have cognitive dysfunction are particularly at risk of cognitive complications, especially lasting or severe effects, if their depression is treated with ECT. However, ECT is used in mixed depression and dementia (Liang, Lam, & Ancill, 1988).

While ECT advocates argue that the treatment is underused, there are three other implications of this literature. First, improved recognition of depression could lead to its earlier treatment, so that cases do not take on the urgency that more often results in referral for ECT. Second, since side effects do seem so dependent on technique, monitoring of side effects and relapse is advisable not only in controlled trials in academic clinical centers, but also at sites where ECT is often administered, such as private not-for-profit and private for-profit hospitals (Sackheim, 1994). Third, treatments to prevent relapse should incorporate more psychotherapy and more environmental manipulation, including work with the family and social networks.

7.09.8.4 Concluding Comment

It is difficult to discuss treatment without also discussing problems in getting the older adult to treatment (Estes, 1995). Older adults tend to be seen by a primary care physician who either does not notice depression or elects not to treat it. The older adult may contribute by failing to seek treatment for symptoms of depression, or by not mentioning the symptoms to the primary care physician. Reasons suggested for failure to recognize depression include: comorbidity with physical health problems, which become the focus of treatment; lack of training in geriatrics; older adults' reticence to discuss emotional distress; older adults' attribution of the symptoms to normal aging. As a consequence, much depression may go untreated, may not be treated until it becomes severe, or may be mistreated, for example, with anxiolytics or with sleeping pills, until it becomes severe and gives the false impression of intractability and hopelessness. If the depression is recognized by the primary care physician, it may be easiest to treat it with drugs, like the patient's physical conditions are being treated, particularly because the depression may have notable somatic manifestations. Referral to specialists is uncommon. Because of delays in initiating treatment, by the time that treatment is begun, the depressed older adult may more often seem like a psychiatric emergency, posing a danger to him or herself explicitly via creating fears of suicide or implicitly by extreme lack of self-care (e.g., not eating). These delays are combined with a greater background chance of physical frailty, so that the individual may not be physically as capable of enduring an episode of depression. This systemic context may lead to greater emphasis on somatic solutions.

7.09.9 PREVENTION

Prevention of depression is of keen interest because research results about risk and protective factors can inform preventive intervention strategies. Risk and protective factors, both established and putative, are suggested by the literature on correlated and etiological influences. In turn, this literature betokens key points of intervention for any preventive efforts.

The Institute of Medicine (Mrazek & Haggerty, 1994) recommends distinguishing universal, selective, and indicated preventive measures. Indicated interventions are directed toward individuals who may have symptoms that foreshadow disorder or some kind of marker that may indicate a predisposition. Selective preventive interventions are targeted to subgroups who may have an elevated risk. Universal measures are targeted to the general public. These distinctions can usefully be applied to older adults (Smyer, 1995).

7.09.9.1 Indicated Preventive Measures

Screening and outreach are core tools for indicated preventive intervention. One of the best indicators of risk of future depression is depressive symptoms. Using Epidemiologic Catchment Area survey data, Horwath, Johnson, Klerman, and Weissman (1992) demonstrated that incidence of first-onset major depression was significantly increased in individuals who reported depressive symptoms when interviewed the previous year. Indeed, over half of new cases of major depression were associated with a history of depressive symptoms. These findings provide a strong empirical rationale for depression symptom screening, for example, national depression awareness days.

Unfortunately, screening for depression, especially in older adults, is not practiced widely. Dorfman et al. (1995) used a computer-assisted questionnaire with depression items from the CES-D supplemented by a second instrument. Among a sample of generally healthy community-dwelling adults aged 65 years and older, they found that nearly three-quarters of those with major depression were not being treated. However, when these individuals were referred to a mental health professional, two-thirds did seek help. These results illustrate both the failure to screen, since these individuals had not heretofore been identified as having depressive disorder, and the fact that screening can be effective in getting people with significant distress linked to treatment.

The high rate of depressive symptoms and disorder among medical patients suggests that primary care practitioners are uniquely well positioned to conduct screening for depression. However, underrecognition of psychiatric conditions by primary care physicians in outpatient and inpatient settings is well demonstrated (Jorm, 1995; Rapp et al., 1988). In a sample of older individuals of Latino origin in Southern California, Kemp et al. (1987) found that none of the respondents who scored as depressed on the screening instrument reported receiving any treatment for depression. Notably, 90% had seen a physician within the past year, where there could have been an opportunity for screening to occur. Mann (1995) determined that primary care physicians were unaware of their patient's depression in two out of every three cases where the patient was in fact pervasively depressed. The pivotal role of the primary care practitioner is further supported by evidence that some three-quarters of elderly persons who kill themselves have recently seen a physician (Conwell et al., 1991).

These observations have led to strong recommendations that health service providers be trained to screen for depression as a condition of licensure or certification (White House Mini-Conference on Emerging Issues in Mental Health and Aging, 1995). Methods for improving the recognition of depression by primary care physicians have been evaluated, including visits to the physician's office analogous to visits by drug industry representatives, during which the physician is presented with a list of questions to use in making a diagnosis (Pond, Mant, Kehoe, Hewitt, & Brodaty, 1994).

Outreach programs to older adults aim to make it easier for the older adult who does have symptoms of depression to connect with someone who can offer assistance. Waters (1995) emphasizes the use of alternative settings, not only health care clinics, but also neighborhood centers, senior centers, religious institutions, community organizations, educational institutions, as locations where mental health care is available. Raschko (1985) established the Gatekeepers program, in which he trained meter readers, letter carriers, bank tellers, and others who might have contact with older individuals who lived alone, so that they became referral sources.

Once individuals with symptoms of depression are identified, Horwath et al. (1992) caution that most will not go on to become clinically depressed. Therefore, these authors suggest, it would be premature to use antidepressant medication to treat depressive symptoms. There may also be reluctance to see a mental health professional. Thus, Breckenridge, Zeiss, and Thompson (1987) developed a nine-week Life Satisfaction Course for older adults who have depressive symptoms but do not meet criteria for major depression. The course is built on Lewinsohn's psychoeducational model (Lewinsohn et al., 1986). This group (Thompson, Gallagher, Nies, & Epstein, 1983) has also offered Coping with Depression classes for older adults, targeted towards those who were clinically depressed. Participation in these classes resulted in improvement on measures of depression and life satisfaction.

7.09.9.2 Selective Preventive Interventions

Selective interventions are directed toward subgroups who have an elevated risk of depression. Research on risk factors for depression suggests what some of those subgroups may be. Phifer and Murrell (1986) found that an increase in CES-D scores from below 16 to 20 or higher was predicted by depressive symptoms six months earlier, poor health, loss events, and low social support, with support functioning both additively and interactively. In the Western Collaborative Group Study of older men, increase on the CES-D was predicted by poor perceived health and stroke, while physical exercise was protective (McElroy, Carmelli, Swan, & LaRue, 1995). It is often observed that the rate of mood disorders in stroke patients is high, and that depression can interfere with cognitive and functional recovery.

A major loss event is spousal bereavement. One of the best known prevention programs is "Widow-to-Widow," developed by Silverman (1969) and sponsored by the American Association of Retired Persons (AARP). Vachon et al. (1980) showed, in a controlled study, that fewer depressive symptoms were experienced by program participants than by the control group. Vachon and Stylianos (1988) have suggested ways to identify which bereaved individuals may be at greater risk of distress based on the adequacy of their social support network, and they propose that friends and family of the bereaved could be taught to distinguish helpful from unhelpful methods of support.

Among medical patients, strategies to prevent depression can include treatment of pain and maintenance of good nutrition (Katz, Streim, & Parmelee, 1994). In cognitively intact or mildly demented patients admitted to long-term care facilities, it has been suggested that prevention of depression may be accomplished through fostering social support in the setting, maximizing opportunities for autonomy and control, and promoting coping skills with respect to institutional life (Foster & Cataldo, 1994). However, there is not good research on the effectiveness of psychotherapeutic interventions with the physically ill or with those in long-term care settings (Jorm, 1995).

A final typical target for selective preventive interventions are caregivers of demented or physically incapacitated older adults. Meta-analysis suggests that psychoeducational support groups are effective in alleviating burden (Knight et al., 1993).

7.09.9.3 Universal Preventive Interventions

Universal preventive programs include public education campaigns and various wellness programs (Waters, 1995). Prevention research programs under this heading present major methodological problems, particularly with respect to demonstrating decreased incidence of a specific mental disorder (Mrazek & Haggerty, 1994). A central consideration should be linking the knowledge base about risk and protective factors to educational strategies for older adults and those who work with them (Smyer, 1995).

7.09.10 CONCLUSION

Old age is not characterized by depression. When depression is experienced, the most frequent mood disorders seem to be those about which least is known. Some of the newer diagnostic categories or symptom clusters, that is, minor depression, depletion syndrome, hopelessness depression, are promising descriptors.

Gerontological literature often blurs the distinction between depression as a clinical entity and depression as an expression of negative affect and hence a convenient indicator of well-being in old age. This blurring may contribute to the misconception that depression is an inevitable accompaniment of aging, brought on by the various losses and indignities, and also by the conviction that depression is underdiagnosed among older adults. Blurring also occurs in the literature on prevention, where distinctions often are poorly drawn between programs to augment well-being and programs where prevention of mood disorder is legitimately realized.

Stress does predict depression. Yet, despite the chronic illnesses and psychosocial strains that frequently distinguish old age, most older people manage without becoming depressed. Physical illness and cognitive dysfunction are major correlated factors for depression in old age. These factors help to account for age differences in presentation and appear to be of etiological significance. For example, there may be a general wearing down process underlying depression, physical symptoms (often cardiovascular), and functional decrements. For late-onset depression, with no previous episodes at younger ages, the depression often foreshadows dementia. Depression in old age also has important consequences with respect to both morbidity and mortality.

Numerous assessment instruments are available, with relative levels of use and sanction not necessarily reflective of existing empirical evidence as to measurement efficiency. The problems with depression scales reflect the nature of depression in old age, that is, the need to evaluate syndromes other than major depression and issues of differentiating depressive disorders among those who are physically or cognitively frail.

Although suicide is a relatively rare event, and most people who are depressed do not commit suicide, the risk of a completed suicide is highest in old age. This is an area where proper assessment and intervention can make a difference.

Psychotherapy works rather well with older adults, and there is empirical and practical guidance available about the best approaches. Somatic approaches have been overemphasized in the treatment of depression in late life. Although they appear to be effective for acute treatment, evidence is not unambiguous that they are more effective than psychotherapy. The guideline that psychopharmacology should be used with severe depression while psychotherapy should be used with mild depression does not accurately reflect the research. In studies of both drugs and psychotherapy, the patients tend to be young old, less physically ill, and more compliant. Risks of pharmacotherapy are higher in older individuals due to greater likelihood of comorbid medical conditions, greater likelihood of interaction with other medications, and greater sensitivity to dosage levels. Moreover, somatic treatments have no lasting effect on dysfunctional thoughts and nonhelpful or depopulated social networks. There is need for evaluations of combined treatments that reflect actual clinical practice, greater attention to dropouts and those who relapse, and studies of depressed patients with concurrent medical conditions. Better outreach and early assessment can also reduce the need for urgent intervention.

Finally, promising approaches to prevention of depression among older adults lie in training those in a position to screen for symptoms, targeting the most vulnerable segments of the older population, and disseminating knowledge about risk and protective factors.

7.09.11 REFERENCES

Abramson L. Y., Metalsky G. I., & Alloy L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review, 96,* 358–372.

Adelmann, P. K. (1994). Multiple roles and psychological well-being in a national sample of older adults. *Journal of Gerontology: Social Sciences, 49,* 277–285.

Alexopoulos, G. S., Abrams, R. C., Young, R. C., & Shamoian, C. A. (1988). Cornell Scale for Depression in Dementia. *Biological Psychiatry, 23,* 271–284.

Alexopoulos, G. S., Meyers, R. S., Young, R. C., Mattis, S., & Kakuma, T. (1993). The course of geriatric depression with "reversible dementia": A controlled study. *American Journal of Psychiatry, 150,* 1693–1799.

Allen-Burge, R., Storandt, M., Kinscherf, D. A., & Rubin, E. H. (1994). Sex differences in the sensitivity of two self-report depression scales in older depressed inpatients. *Psychology and Aging, 9,* 443–445.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., Rev.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Ames, D., Flynn, E., Tuckwell, V., & Harrigan, S. (1994). Diagnosis of psychiatric disorder in elderly general and geriatric hospital patients: AGECAT and *DSM-III-R* compared. *International Journal of Geriatric Psychiatry, 9,* 627–633.

Andresen E. M., Malmgren J. A., Carter W. B., & Patrick, D. L. (1994). Screening for depression in well older adults: Evaluation of a short form of the CES-D. *American Journal of Preventive Medicine, 10,* 77–84.

Anthony-Bergstone, C. R., Zarit, S. H., & Gatz, M. (1988). Symptoms of psychological distress among caregivers of dementia patients. *Psychology and Aging, 3,* 245–248.

Arean, P. A., Perri, M. G., Nezu, A. M., Schein, R. L., Christopher, F., & Joseph, T. X. (1993). Comparative effectiveness of social problem-solving therapy and reminiscence therapy as treatments for depression in older adults. *Journal of Consulting and Clinical Psychology, 61,* 1003–1010.

Åsberg, M., Nordström, P., Träskman-Bendz, L. (1986). Cerebrospinal fluid studies in suicide: An overview. *Annals of the New York Academy of Sciences, 487,* 243–255.

Baker, F. M., & Miller, C. L. (1991). Screening a skilled nursing home population for depression. *Journal of Geriatric Psychiatry and Neurology 4,* 218–221.

Baldwin, R. C. (1995). Delusional depression in elderly patients: Characteristics and relationship to age at onset. *International Journal of Geriatric Psychiatry, 10,* 981–985.

Baldwin, R. C., & Tomenson, B. (1995). Depression in later life: A comparison of symptoms and risk factors in early and late onset cases. *British Journal of Psychiatry, 167,* 649–652.

Ballard, C. G., Cassidy, G., Bannister, C., & Mohan, R. N. C. (1993). Prevalence, symptom profile and aetiology of depression in dementia sufferers. *Journal of Affective Disorders, 29,* 1–6.

Baltes, P. B., Mayer, K. U., Helmchen, H., & Steinhagen-Thiessen, E. (1993). The Berlin Aging Study (BASE): Overview and design. *Ageing and Society, 13,* 485–515.

Barraclough, B., Bunch, J., Nelson, B., & Sainsbury, P. (1974). A hundred cases of suicide: Clinical aspects. *British Journal of Psychiatry, 125,* 355–373.

Barrett, J. E., Barrett, J. A., Oxman, T. E., & Gerber, P. D. (1988). The prevalence of psychiatric disorders in a primary care practice. *Archives of General Psychiatry, 45,* 1100–1106.

Battin, M. P. (1991). Rational suicide: How can we respond to a request for help? Special Issue: Suicide in the Elderly. *Crisis, 12,* 73–80.

Beck, A. T., & Beck, R. W. (1972). Screening depressed patients in family practice: a rapid technique. *Postgraduate Medicine, 52*(6), 81–85.

Beck, A. T., Rush, A. J. Shaw, B. F., & Emery, G. (1979). *Cognitive therapy for depression.* New York: Guilford Press.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8,* 77–100.

Beck, A. T., Steer, R. A., Kovacs, M., & Garrison, B. (1985). Hopelessness and eventual suicide: A 10-year prospective study of patients hospitalized with suicidal ideation. *American Journal of Psychiatry, 142,* 559–563.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 53–63.

Beekman, A. T. F., Deeg, D. J. H., van Tilburg, T., Smit, J. H., Hooijer, C., & van Tilburg, W. (1995). Major and minor depression in later life: A study of prevalence and risk factors. *Journal of Affective Disorders, 36,* 65–75.

Berkman, L. F., Berkman, C. S., Kasl, S., Freeman, D. H., Jr., Leo, L., Ostfeld, A. M., Cornoni-Huntley, J., & Brody, J. A. (1986). Depressive symptoms in relation to physical health and functioning in the elderly. *American Journal of Epidemiology, 124,* 372–388.

Bernstein, D. A., & Borkovec, T. D. (1973). *Progressive relaxation training: A manual for the healing professions.* Champaign, IL: Research Press.

Berry, J. M., Storandt, M., & Coyne, A. (1984). Age and sex differences in somatic complaints associated with depression. *Journal of Gerontology, 39,* 465–467.

Bertelsen A., Harvald B., & Hauge M. A. (1977). A Danish twin study of manic-depressive disorders. *British Journal of Psychiatry, 130,* 330–351.

Beutler, L. E., Scogin, F., Kirkish, P., Schretlen, D., Corbishley, A., Hamblin, D., Meredith, K., Potter, R., Bamford, C. R., & Levenson, A. I. (1987). Group cognitive therapy and alprazolam in the treatment of depression in older adults. *Journal of Consulting and Clinical Psychology, 55,* 550–556.

Blazer, D. G., (1994). Epidemiology of late life depression. In L. S. Schneider, C. F. Reynolds, III, B. D. Lebowitz, & A. J. Friedhoff (Eds.), *Diagnosis and treatment of depression in late life: Results of the NIH Consensus Development Conference* (pp. 9–19). Washington, DC: American Psychiatric Press.

Blazer, D., Bachar, J. R., & Hughes, D. C. (1987). Major depression with melancholia: A comparison of middle-aged and elderly adults. *Journal of the American Geriatrics Society, 35,* 927–932.

Blazer, D. G., Burchett, B., Service, C., & George, L. K. (1991). The association of age and depression among the elderly. An epidemiological exploration. *Journals of Gerontology: Medical Sciences, 46,* M210–M215.

Blazer, D., Swartz, M., Woodbury, M., Manton, K. G., Hughes, D., & George, L. K. (1988). Depressive symptoms and depressive diagnoses in a community population: Use of a new procedure for analysis of psychiatric classification. *Archives of General Psychiatry, 45,* 1078–1084.

Blumenthal, M. D. (1980). Depressive illness in old age: Getting behind the mask. *Geriatrics, 35,* 34–43.

Boone, K. B., Lesser, I., Miller, B., Wohl, M., Berman, N., Lee, A. & Palmer, B. (1994). Cognitive functioning in a mildly to moderately depressed geriatric sample: Relationship to chronological age. *Journal of Neuropsychiatry and Clinical Neurosciences, 6,* 267–272.

Breckenridge, J. S., Zeiss, A. M., & Thompson, L. W. (1987). The Life Satisfaction Course: An intervention for the elderly. In R. F. Muñoz (Ed.), *Depression prevention: Research directions* (pp. 185–196). Washington, DC: Hemisphere Publishing.

Brodaty, H., Peters, K., Boyce, P., Hickie, I., Parker, G., Mitchell, P., & Wilhelm, K. (1991). Age and depression. *Journal of Affective Disorders, 23,* 137–149.

Brown, R. P., Sweeney, J., Loutsch, E., Kocsis, J., & Frances, A. (1984). Involutional melancholia revisited. *American Journal of Psychiatry, 141,* 24–28.

Bruce, M. L., Seeman, T. E., Merrill, S. S., & Blazer, D. G. (1994). The impact of depressive symptomatology on physical disability: MacArthur studies of successful aging. *American Journal of Public Health, 84,* 1796–1799.

Burke, W. J., Houston, M. J., Boust S. J., & Roccaforte, W. H. (1989). Use of the Geriatric Depression Scale in dementia of the Alzheimer's type. *Journal of the American Geriatrics Society, 37,* 856–860.

Burke, W. J., Nitcher, R. L., Roccaforte, W. H., & Wengel, S. P. (1992). A prospective evaluation of the Geriatric Depression Scale in an outpatient geriatric assessment center. *Journal of the American Geriatrics Society, 40,* 1227–1230.

Burke, W. J., Roccaforte, W. H., & Wengel, S. P. (1991). The short form of the Geriatric Depression Scale: A comparison with the 30-item form. *Journal of Geriatric Psychiatry and Neurology, 4,* 173–178.

Burke, W. J., Roccaforte, W. H., Wengel, S. P., Conley, D. M., & Potter, J. F. (1995). The reliability and validity of the Geriatric Depression Rating Scale administered by telephone. *Journal of the American Geriatric Society, 43,* 674–679.

Burvill, P. W., & Hall, W. D. (1994). Predictors of increased mortality in elderly depressed patients. *International Journal of Geriatric Psychiatry, 9,* 219–227.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring.* Minneapolis, MN: University of Minnesota Press.

Butler, R. N. (1974). Successful aging and the role of the life review. *Journal of the American Geriatric Society, 22,* 529–535.

Caine, E. D., Lyness, J. M., King, D. A., & Connors, B. A. (1994). Clinical and etiological heterogeneity of mood disorders in elderly patients. In L. S. Schneider, C. F. Reynolds, III, B. D. Lebowitz, & A. J. Friedhoff (Eds.), *Diagnosis and treatment of depression in late life: Results of the NIH Consensus Development Conference* (pp. 23–53). Washington, DC: American Psychiatric Press.

Callahan, C. M., & Wolinsky, F. D. (1994). The effect of gender and race on the measurement properties of the CES-D in older adults. *Medical Care, 32,* 341–356.

Campbell, J. M. (1992). Treating depression in well older adults: Use of diaries in cognitive therapy. *Issues in Mental Health Nursing, 13,* 19–29.

Carroll, B. J., Feinberg, M., Smouse, P. E., Rawson, S. G., & Greden, J. F. (1981). The Carroll Rating Scale for Depression I. Development, reliability and validation. *British Journal of Psychiatry, 138,* 194–200.

Cattan R. A., Barry P. P., Mead G., Reefe W. E., Gay A., & Silverman M. (1990). Electroconvulsive therapy in octogenerians. *Journal of the American Geriatrics Society, 38,* 753–758.

Chambless, D. L., Sanderson, W. C., Shoham, V., Johnson, S. B., Pope, K. S., Crits-Christoph, P., Baker, M., Johnson, B., Woody, S. R., Suie, S., Beutler, L., Williams, D. A., & McCurry, S. (1996). An update on empirically validated therapies. *The Clinical Psychologist, 49*(2), 5–18.

Commision on Professional Hospital Activities (1978). *ICD-9-CM. The international classification of diseases: Clinical modifications* (9th Rev.). Ann Arbor, MI: Author.

Conn, D. K. (1989). Neuropsychiatric syndromes in the elderly: An overview. In D. K. Cohen, A. Grek, & J. Sadavoy (Eds.), *Psychiatric consequences of brain disease in the elderly: A focus on management* (pp. 1–49). New York: Plenum.

Conwell, Y., & Brent, D. (1995). Suicide and aging I: Patterns of psychiatric diagnosis. *International Psychogeriatrics, 7,* 149–164.

Conwell, Y., Olsen, K., Caine, E. D., & Flannery, C. (1991). Suicide in later life: Psychological autopsy findings. *International Psychogeriatrics, 3,* 59–66.

Copeland, J. R. M., Dewey, M. E., Wood, N., Searle, R., Davidson, I. A., & McWilliam, C. (1987). Range of mental illness among the elderly in the community: Prevalence in Liverpool using the GMS-AGECAT package. *British Journal of Psychiatry, 150,* 815–823.

Copeland, J. R. M., Gurland, B. J., Dewey, M. E., Kelleher, M. J., Smith, A. M. R., & Davidson, I. A. (1987). The distribution of dementia, depression and neurosis in elderly men and women in an urban community: Assessed using the GMS-AGECAT package. *International Journal of Geriatric Psychiatry, 2,* 177–184.

Copeland, J. R. M., Kelleher, M. J., Kellett, J. M., Gourlay, A. L., with Gurland, B. J., Fleiss, J. L., & Sharpe, L. (1976). A semi-structured clinical interview for the assessment of diagnosis and mental state in the elderly: the Geriatric Mental State Schedule I. Development and reliability. *Psychological Medicine, 6,* 439–449.

Cress, M. E., Schechtman, K. B., Mulrow, C. D., Fiatarone, M. A., Gerety, M. B., & Buchner, D. M. (1995). Relationship between physical performance and self-perceived physical function. *Journal of the American Geriatrics Society, 43,* 93–101.

Cwikel, J., & Ritchie, K. (1988). The short GDS: Evaluation in a heterogeneous, multilingual population. *Clinical Gerontologist, 8,* 63–83.

D'Ath, P., Katona, P., Mullan, E., Evans, S., & Katona, C. (1994). Screening, detection and management of depression in elderly primary care attenders. I: The acceptability and performance of the 15 item Geriatric Depression Scale (GDS15) and the development of short versions. *Family Practice, 11,* 260–266.

Davidson, H., Feldman, P. H., & Crawford, S. (1994). Measuring depressive symptoms in the frail elderly. *Journal of Gertontology: Psychological Sciences, 49,* 159–164.

Dean, A., Kolody, B., & Wood, P. (1990). Effects of social support from various sources on depression in elderly persons. *Journal of Health and Social Behavior, 31,* 148–161.

De Leo, D., & Diekstra, R. F. W. (1990). *Depression and suicide in late life*. Lewiston, NY: Hogrefe & Huber.

Derogatis, L. R. (1977). *SCL-90: Administration, scoring and procedures manual I for the R (revised) version.* Baltimore, MD: Johns Hopkins University School of Medicine.

Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine, 13,* 595–605.

Devanand, D. P., Miller, L., Richards, M., Marder, K., Bell, K., Mayeux, R., & Stern, Y. (1992). The Columbia University Scale for Psychopathology in Alzheimer's Disease. *Archives of Neurology, 49,* 371–376.

Dhooper, S. S., Green, S. M., Huff, M. B., & Austen-Murphy, J. (1993). Efficacy of a group approach to reducing depression in nursing home elderly residents. *Journal of Gerontological Social Work, 20,* 87–100.

Dorfman, R. A., Lubben, J. E., Mayer-Oakes, A., Atchison K., Schweitzer S. O., DeJong, F. J., & Matthias, R. E. (1995). Screening for depression among a well elderly population. *Social Work, 40,* 295–304.

Duke Depression Evaluation Schedule for the Elderly (1988). Center for the Study of Depression in Later Life, Department of Psychiatry, Duke University Medical Center, Durham, NC.

Dunn, V. K., & Sacco, W. P. (1989). Psychometric evaluation of the Geriatric Depression Scale and the Zung Self-Rating Scale using an elderly community sample. *Psychology and Aging, 4,* 125–126.

Dush, D. M., & Hutzell, R. R. (1986). A brief MMPI depression scale for the elderly. *Clinical Gerontologist, 5,* 175–185.

Eaton, W. W., & Kessler, L. G. (1981). Rates of symptoms of depression in a national sample. *American Journal of Epidemiology, 114,* 528–538.

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., Glass, D. R., Pilkonis, P. A., Leber, W. R., Doherty, J. P., Fiester, S. J., & Parloff, M. B. (1989). NIMH Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Archives of General Psychiatry, 46,* 971–982.

Endicott, J., & Spitzer, R. L. (1978). A diagnostic interview. The Schedule for Affective Disorders and Schizophrenia. *Archives of General Psychiatry, 35,* 837–844.

Estes, C. L. (1995). Mental health services for the elderly: Key policy elements. In M. Gatz (Ed.), *Emerging issues in mental health and aging* (pp. 303–327). Washington, DC: American Psychological Association.

Evans, M. E. (1993). Depression in elderly physically ill inpatients: A 12-month prospective study. Special issue: Affective disorders: Current and future perspectives. *International Journal of Geriatric Psychiatry, 8,* 587–592.

Fawcett, J., Scheftner, W. A., Fogg, L., Clark, D. C., Young, M. A., Hedeker, D., & Gibbons, R. (1990). Time-related predictors of suicide in major affective disorder. *American Journal of Psychiatry, 147,* 1189–1194.

Feinberg, M., Carroll, B. J., Smouse, P. E., & Rawson, S. G. (1981). The Carroll Rating Scale for Depression: III. Comparison with other rating instruments. *British Journal of Psychiatry, 138,* 205–209.

Ferrier, I. N., & McKeith, I. G. (1991). Neuroanatomical and neurochemical changes in affective disorders in old age. Special Issue: Affective disorders in old age. *International Journal of Geriatric Psychiatry, 6,* 445–451.

Fischer, L. R., Rolnick, S. J., Jackson, J., Garrard, J., & Luepke, L. (1996). The Geriatric Depression Scale: A content analysis of respondent comments. *Journal of Mental Health and Aging, 2,* 125–135.

Fitz, A. G., & Teri, L. (1994). Depression, cognition, and functional ability in patients with Alzheimer's disease. *Journal of the American Geriatric Society, 42,* 186–191.

Fogel, B. S., & Fretwell, M. (1985). Reclassification of depression in the medically ill elderly. *Journal of the American Geriatrics Society, 33,* 446–448.

Forsell, Y., Jorm, A. F., von Strauss, E., & Winblad, B.

(1995). Prevalence and correlates of depression in a population of nonagenarians. *British Journal of Psychiatry, 167,* 61–64.

Forsell, Y., Jorm, A. F., & Winblad, B. (1994). Association of age, sex, cognitive dysfunction, and disability with major depressive symptoms in an elderly sample. *American Journal of Psychiatry, 151,* 1600–1604.

Foster, J. R., & Cataldo, J. K. (1994). Protection from clinical depression in medical long-term care facilities: Evidence for psychologic adaptation in cognitively intact patients. *International Journal of Geriatric Psychiatry, 9,* 115–125.

Frerichs, R. R., Aneshensel, C. S., & Clark, V. A. (1981). Prevalence of depression in Los Angeles County. *American Journal of Epidemiology, 113,* 691–699.

Friedhoff, A. J. (1994). Consensus Development Conference Statement: Diagnosis and treatment of depression in late life. In L. S. Schneider, C. F. Reynolds, III, B. D. Lebowitz, & A. J. Friedhoff (Eds.), *Diagnosis and treatment of depression in late life: Results of the NIH Consensus Development Conference* (pp. 493–511). Washington, DC: American Psychiatric Press.

Fry, P. S. (1983). Structured and unstructured reminiscence training and depression among the elderly. *Clinical Gerontologist, 1,* 15–37.

Fuhrer, R., Antonucci, T. C., Gagnon, M., Dartigues, J.-F., Barberger-Gateau, P., & Alperovitch, A. (1992). Depressive symptomatology and cognitive functioning: An epidemiological survey in an elderly community sample in France. *Psychological Medicine, 22,* 159–172.

Futterman, A., Thompson, L., Gallagher-Thompson, D., & Ferris, R. (1995). Depression in later life: Epidemiology, assessment, etiology, and treatment. In E. E. Beckham & W. R. Leber (Eds.), *Handbook of depression* (2nd ed., pp. 494–525). New York: Guilford Press.

Gallagher, D. (1986). The Beck Depression Inventory and older adults: Review of its development and utility. *Clinical Gerontologist, 5,* 149–163.

Gallagher, D., Breckinridge, J., Steinmetz, J., & Thompson, L. (1983). The Beck Depression Inventory & Research Diagnostic Criteria: Congruence in an older population. *Journal of Consulting and Clinical Psychology, 51,* 945–946.

Gallagher, D. E., & Thompson, L. W. (1982). Treatment of major depressive disorder in older adult outpatients with brief psychotherapies. *Psychotherapy: Theory, Research and Practice, 19,* 482–490.

Gallagher, D. E., & Thompson, L. W. (1983). Effectiveness of psychotherapy for both endogenous and nonendogenous depression in older adult outpatients. *Journal of Gerontology, 38,* 707–712.

Gallagher-Thompson, D. (1994). Clinical intervention strategies for distressed caregivers: Rationale and development of psychoeducational approaches. In E. Light, G. Niederehe, & B.D. Lebowitz (Eds.), *Stress effects on family caregivers of Alzheimer's patients: Research and interventions* (pp. 260–277). New York: Springer.

Gallagher-Thompson, D. E., Hanley-Peterson, P., & Thompson, L. W. (1990). Maintenance of gains versus relapse following brief psychotherapy for depression. *Journal of Consulting and Clinical Psychology, 58,* 371–374.

Gallo, J. J., Anthony, J. C., & Muthén, B. O. (1994). Age differences in the symptoms of depression: A latent trait analysis. *Journal of Gerontology: Psychological Sciences, 49,* P251–P264.

Garcia, M., & Marks, G. (1989). Depressive symptomatology among Mexican-American adults: An examination with the CES-D Scale. *Psychiatry Research, 27,* 137–148.

Gardner, E. A., Bahn, A. K., Mack, M. (1964). The relationship between premature deaths and affective disorders. *British Journal of Psychiatry, 115,* 1277–1282.

Gatz, M. (1992, December 14). *Depression and dementia studies in twins.* Berlin: Department of Psychiatry, Free University, and Max Planck Institute.

Gatz, M., Fiske, A., Fox, L., Kaskie, B., Kasl-Godley, J. E., McCallum, T. J., & Wetherell, J. (in press). Empirically-validated psychological treatments for older adults. *Journal of Mental Health and Aging.*

Gatz, M., & Hurwicz, M. (1990). Are old people more depressed? Cross-sectional data on Center for Epidemiological studies depression scale factors. *Psychology and Aging, 5,* 284–290.

Gatz, M., Johansson, B., Pedersen, N., Berg, S., & Reynolds, C. (1993). A cross-national self-report measure of depressive symptomatology. *International Psychogeriatrics, 5,* 147–156.

George, L. K., Blazer, D. G., Hughes, D. C., & Fowler, N. (1989). Social support and the outcome of major depression. *British Journal of Psychiatry, 154,* 478–485.

Gerson, S. C., Plotkin, D. A., & Jarvik, L. F. (1988). Antidepressant drug studies, 1964 to 1986: Empirical evidence for aging patients. *Journal of Clinical Psychopharmacology, 8,* 311–322.

Gilewski, M. J., Farberow, N. L., Gallagher, D. E., & Thompson, L. W. (1991). Interaction of depression and bereavement on mental health in the elderly. *Psychology and Aging, 6,* 67–75.

Gilley, D. W., Wilson, R. S., Fleischman, D. A., Harrison, D. W., Goetz, C. G., & Tanner, C. M. (1995). Impact of Alzheimer's Type Dementia and information source on the assessment of depression. *Psychological Assessment, 7,* 42–48.

Girling, D. M., Barkley, C., Paykel, E. S., Gehlhaar, E., Brayne, C., Gill, C., Methewson, D., & Huppert, F. A. (1995). The prevalence of depression in a cohort of the very elderly. *Journal of Affective Disorders, 34,* 319–329.

Glassman, A. H., & Roose, S. P. (1994). Risks of antidepressants in the elderly: Tricyclic antidepressants and arrhythmia-revising risks. *Gerontology, 40*(Suppl. 1), 15–20.

Golden, R. R., Teresi, J. A., & Gurland, B. J. (1984). Development of indicator scales for the comprehensive assessment and referral evaluation (CARE) interview schedule. *Journal of Gerontology, 39,* 138–146.

Greenwald, B. S., Kramer-Ginsberg, E., Marin, D. B., Laitman, L. B., Hermann, C. K., Mohs, R. C., & Davis, K. L. (1989). Dementia with coexistent major depression. *American Journal of Psychiatry, 146,* 1472–1478.

Guadagnoli, E., & Mor, V. (1989). Measuring cancer patients' affect: Revision and psychometric properties of the profile of mood states (POMS). *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2,* 150–154.

Gulbinat, W. H. (1996). The epidemiology of suicide in old age. *Archives of Suicide Research, 2,* 31–42.

Gurland, B., Kuriansky, J., Sharpe, L., Simon, R., Stiller, P., & Birkett, P. (1977–78). The Comprehensive Assessment and Referral Evaluation (CARE): Rationale, development and reliability. *International Journal of Aging and Human Development, 8,* 9–42.

Haight, B. K. (1992). Long-term effects of a structured life review process. *Journal of Gerontology, 47,* P312–P315.

Hale, W. D., Cochran, C. D., & Hedgepeth, B. E. (1984). Norms for the elderly on the Brief Symptom Index. *Journal of Consulting and Clinical Psychology, 52,* 321–322.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry, 23,* 56–62.

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology, 6,* 278–296.

Harper, R. G., Kotik-Harper, D., & Kirby, H. (1990). Psychometric assessment of depression in an elderly general medical population: Over- or underassessment?

The Journal of Nervous and Mental Disease, 178, 113–119.

Hays, J. C., Kasl, S. V., & Jacobs, S. C. (1994). The course of psychological distress following threatened and actual conjugal bereavement. *Psychological Medicine, 24,* 917–927.

Heikkinen R.-L., Berg S., & Avlund K. (1995). Depressive symptoms in late life. *Journal of Cross-Cultural Gerontology, 10,* 315–330.

Henderson, A. S., Jorm, A. F., MacKinnon, A., Christensen, H., Scott, L. R., Korten, A. E., & Doyle, C. (1993). The prevalence of depressive disorders and the distribution of depressive symptoms in later life: A survey using Draft *ICD-10* and *DSM-III-R*. *Psychological Medicine, 23,* 719–729.

Henriksson, M. M., Isometsä, E. T., Hietanen, P. S., Aro, H., & Lönnqvist, J. K. (1995). Mental disorders in cancer suicides. *Journal of Affective Disorders, 36,* 11–20.

Henriksson, M. M., Marttunen, M. J., Isometsä, E. T., Heikkinen, M. E., Aro, H. M., Kuoppasalmi, K. I., & Lönnqvist, J. K. (1995). Mental disorders in elderly suicide. *International Psychogeriatrics, 7,* 275–286.

Hertzog, C., Van Alstine, J., Usala, P. D., Hultsch, D. F., Dixon, R. (1990). Measurement properties of the Center for Epidemiological Studies Depression Scale (CES-D) in older populations. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2,* 64–72.

Hinrichsen, G. A. (1992). Recovery and relapse from major depressive disorder in the elderly. 144th Annual Meeting of the American Psychiatric Association (1991, New Orleans, LA.). *American Journal of Psychiatry, 149,* 1575–1579.

Hinrichsen, G. A., & Zweig, R. (1994). Family issues in late-life depression. *Journal of Long-Term Home Health Care, 13,* 4–15.

Hobfoll, E., & Wells, J. D. (in press). Conservation of resources, stress, and aging: Why do some slide and some spring? In J. Lomranz (Ed.), *Handbook of aging and mental health.* New York: Plenum.

Horwath, E., Johnson, J., Klerman, G. L., & Weissman, M. M. (1992). Depressive symptoms as relative and attributable risk factors for first-onset major depression. *Archives of General Psychiatry, 49,* 817–823.

Hussian, R. A., & Lawrence, P. S. (1981). Social reinforcement of activity and problem-solving training in the treatment of depressed institutionalized elderly patients. *Cognitive Therapy and Research, 5,* 57–69.

Jamison, C., & Scogin, F. (1992). Development of an interview-based geriatric depression rating scale. *International Journal of Aging and Human Development, 35,* 193–204.

Jarvik, L. F., Mintz, J., Steuer, J. L., & Gerner, R. (1982). Treating geriatric depression: A 26-week interim analysis. *Journal of the American Geriatric Society, 30,* 713–717.

Jenike, M. A. (1989). Treatment of affective illness in the elderly with drugs and electroconvulsive therapy. *Journal of Geriatric Psychiatry, 22,* 77–112.

Jorm, A. F. (1995). The epidemiology of depressive states in the elderly: Implications for recognition, intervention, and prevention. *Social Psychiatry and Psychiatric Epidemiology, 30,* 53–59.

Jorm, A. F., MacKinnon, A. J., Henderson, A. S., Scott, R., Christensen, H., Korten, A. E., Cullen, J. S., & Mulligan, R. (1995). The Psychogeriatric Asssessment Scales: a multidimensional alternative to categorical diagnoses of dementia and depression in the elderly. *Psychological Medicine, 25,* 447–460.

Kafonek, S., Ettinger, W. H., Roca, R., Kittner, S., Taylor, N., & German, P. S. (1989). Instruments for screening for depression and dementia in a long-term care facility. *Journal of the American Geriatrics Society, 37,* 29–34.

Kaplan, M. S., Adamek, M. E., & Johnson, S. (1994). Trends in firearm suicide among older American males: 1979–1988. *The Gerontologist, 34,* 59–65.

Kaszniak, A. W., & Christenson, G. D. (1994). Differential diagnosis of dementia and depression. In M. Storandt & G. R. Vanden Bos (Eds.), *Neuropsychological assessment of dementia and depression in older adults: A clinician's guide.* Washington, DC: American Psychological Association.

Katz, I. R., Simpson, G. M., Curlik, S. M., Parmelee, P. A., & Muhly, C. (1990). Pharmacologic treatment of major depression for elderly patients in residential care settings. *Journal of Clinical Psychiatry, 51* (Suppl. 7), 41–47.

Katz, I. R., Streim, J., & Parmelee, P. (1994). Prevention of depression, recurrences, and complications in late life. *Preventive Medicine: An International Journal Devoted to Practice and Theory, 23,* 743–750.

Kay, D. W. K., & Bergmann, K. (1980). Epidemiology of mental disorders among the aged in the community. In J. E. Birren & R. B. Sloane (Eds.), *Handbook of mental health and aging* (pp. 34–56). Englewood Cliffs, NJ: Prentice Hall.

Kemp, B. J., Staples, F., & Lopez-Aqueres, W. L. (1987). Epidemiology of depression and dysphoria in an elderly Hispanic population: Prevalence and correlates. *Journal of the American Geriatrics Society, 35,* 920–926.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992). Major depression and generalized anxiety disorder: Same genes, (partly) different environments? *Archives of General Psychiatry, 49,* 716–722.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1993). The lifetime history of major depression in women: Reliability of diagnosis and heritability. *Archives of General Psychiatry, 50,* 863–870.

Kennedy, G. J., Kelman, H. R., & Thomas, C. (1991). Persistence and remission of depressive symptoms in late life. *American Journal of Psychiatry, 148,* 174–178.

Kennedy, G. J., Kelman, H. R., Thomas, C., Wisniewski, W., Metz, H., & Bijur, P. E. (1989). Hierarchy of characteristics associated with depressive symptoms in an urban elderly sample. *American Journal of Psychiatry, 146,* 220–225.

Kessler, R. C., Foster, C., Webster, P. S., & House, J. S. (1992). The relationship between age and depressive symptoms in two national surveys. *Psychology and Aging, 7,* 119–126.

Kessler, R., & Mroczek, D. (1993). *UM-CIDI Short-Form* [Memo]. Ann Arbor, MI: University of Michigan.

Kinzie, J. D., Lewinsohn, P., Maricle, R., & Teri, L. (1986). The relationship of depression to medical illness in an older community population. *Comprehensive Psychiatry, 27,* 241–246.

Kitchell, M. A., Barnes, R. F., Veith, R. C., Okimoto, J. T., & Raskind, M. A. (1982). Screening for depression in hospitalized geriatric medical patients. *Journal of the American Geriatrics Society, 30,* 174–177.

Kivela, S., & Pahkala, K. (1986). Sex and age differences of factor pattern and reliability of the Zung Self-Rating Depression Scale in a Finnish elderly population. *Psychological Reports, 59,* 587–597.

Kivela, S., & Pahkala, K. (1987). Factor structure of the Zung Self-rating Depression Scale among a depressed elderly population. *International Journal of Psychology, 22,* 289–300.

Klawansky, S. (1994). Meta-analysis on the treatment of depression in late life. In L. S. Schneider, C. F. Reynolds, III, B. D. Lebowitz, & A. J. Friedhoff (Eds.), *Diagnosis and treatment of depression in late life: Results of the NIH Consensus Development Conference* (pp. 333–352). Washington, DC: American Psychiatric Press.

Klerman, G. L., Lavori, P. W., Rice, J., Reich, T., Endicott, J., Andreasen, N. C., Keller, M. B., & Hirschfield, R. M. A. (1985). Birth-cohort trends in rates of major depressive disorder among relatives of patients with affective disorder. *Archives of General Psychiatry, 42,* 689–693.

Klerman, G. L., & Weissman, M. M. (1989). Increasing rates of depression. *Journal of the American Medical Association, 261,* 2229–2235.

Knight, B. G., Lutzky, S. M., & Macofsky-Urban, F. (1993). A meta-analytic review of interventions for caregiver distress: Recommendations for future research. *Gerontologist, 33,* 240–248.

Koenig, H. G., & Blazer, D. G. (1992). Epidemiology of geriatric affective disorders. *Clinics in Geriatric Medicine, 8,* 235–251.

Koenig, H. G., Cohen, H. J., Blazer, D. G., Krishnan, K. R. R., & Sibert, T. E. (1993). Profile of depressive symptoms in younger and older medical inpatients with major depression. *Journal of the American Geriatrics Society, 41,* 1116–1176.

Koenig, H. G., Cohen, H. J., Blazer, D. G., Meador, K. G., & Westlund, R. (1992). A brief Depression Scale for use in the medically ill. *International Journal of Psychiatry in Medicine, 22,* 183–195.

Koenig, H. G., Meador, K. G., Cohen, H. J., & Blazer, D. G. (1988). Self-rated depression scales and screening for major depression in the older hospitalized patient with medical illness. *Journal of the American Geriatrics Society, 36,* 699–706.

Koenig, H. G., Meador, K. G., Cohen, H. J., & Blazer, D. G. (1992). Screening for depression in hospitalized elderly medical patients: Taking a closer look. *Journal of the American Geriatrics Society, 40,* 1013–1017.

Koenig, H. G., Pappas, P., Holsinger, T., & Bachar, J. R. (1995). Assessing diagnostic approaches to depression in medically ill older adults: How reliably can mental health professionals make judgments about the cause of symptoms? *Journal of the American Geriatrics Society, 43,* 472–478.

Kogan, E. S., Kabacoff, R. I., Hersen, M., & Van Hasselt, V. B. (1994). Clinical cutoffs for the Beck Depression Inventory and the Geriatric Depression Scale with older adult psychiatric outpatients. *Journal of Psychopathology and Behavioral Assessment, 16,* 233–242.

Kogan, E. S., Van Hasselt, V. B., Hersen, M., & Kabacoff, R. I. (1995). Relationship of depression, assertiveness, and social support in community-dwelling older adults. *Journal of Clinical Geropsychology, 1,* 157–163.

Kohout, F. J., Berkman, L. F., Evans, D. A., & Cornoni-Huntley, J. (1993). Two shorter forms of the CES-D depression symptoms index. *Journal of Aging and Health, 5,* 179–193.

Kongstvedt, S. J., & Sime, W. (1992). Diagnosis of depression in early- and late-onset geriatric depression: An evaluation of commonly used self-assessment measures. *Behavior, Health and Aging, 2,* 133–148.

Kral, V. A., & Emery, O. B. (1989). Long-term follow-up of depressive pseudodementia of the aged. *Canadian Journal of Psychiatry, 34,* 445–446.

Krause, N., Herzog, A. R., & Baker, E. (1992). Providing support to others and well-being in later life. *Journal of Gerontology: Psychological Sciences, 47,* 300–311.

Kua, E. H. (1992). A community study of mental disorders in elderly Singaporean Chinese using the GMS-AGE-CAT package. *Australian and New Zealand Journal of Psychiatry, 26,* 502–506.

Kunik, M. E., Champagne, L., Harper, R. G., & Chacko, R. C. (1994). Cognitive functioning in elderly depressed patients with and without psychosis. *International Journal of Geriatric Psychiatry, 9,* 871–874.

La Rue, A. (1992). *Aging and neuropsychological assessment.* New York: Plenum.

La Rue, A., Goodman, S., & Spar, J. E. (1992). Risk factors for memory impairment in geriatric depression. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 5,* 178–184.

Lawlor, B. A., & Anderson, M. C. (1995). The neurobiology of late-life depression: Impact of silent cerebrovascular disease. *Human Psychopharmacology, 10,* S223–S227.

Lazarus, L. W., Newton, N., Cohler, B., Lesser, J., & Schweon, C. (1987). Frequency and presentation of depressive symptoms in patients with primary degenerative dementia. *American Journal of Psychiatry, 144,* 41–45.

Leibenluft, E., & Goldberg, R. L. (1988). The suicidal, terminally ill patient with depression. *Psychosomatics, 29,* 379–386.

Lesher, E. L. (1986). Validation of the Geriatric Depression Scale among nursing home residents. *Clinical Gerontologist, 4,* 21–28.

Lesher, E. L., & Berryhill, J. S. (1994). Validation of the Geriatric Depression Scale—Short form among inpatients. *Journal of Clinical Psychology, 50,* 256–260.

Leuchter, A. F. (1994). Brain structural and functional correlates of late life depression. In L. S. Schneider, C. F. Reynolds, III, B. D. Lebowitz, & A. J. Friedhoff (Eds.), *Diagnosis and treatment of depression in late life: Results of the NIH Consensus Development Conference* (pp. 117–130). Washington, DC: American Psychiatric Press.

Levy, L. H., Martinkowski, K. S., & Derby, J. F. (1994). Differences in patterns of adaptation in conjugal bereavement: Their sources and potential significance. *Omega: Journal of Death and Dying, 29,* 71–87.

Lewinsohn, P., Muñoz, R., Youngren, M. A., & Zeiss, A. (1986). *Control your depression.* Englewood Cliffs, NJ: Prentice Hall.

Liang, R. A., Lam, R. W., & Ancill, R. J. (1988). ECT in the treatment of mixed depression and dementia. *British Journal of Psychiatry, 152,* 281–284.

Lichtenberg, P. A., Marcopulos, B. A., Steiner, D. A., & Tabscott, J. A. (1992). Comparison of the Hamilton Depression Rating Scale and the Geriatric Depression Scale: Detection of depression in dementia patients. *Psychological Reports, 70,* 515–521.

Lichtenstein, P., Gatz, M., Pedersen, N. L., Berg, S., & McClearn, G. E. (1996). A cotwin-control study of response to widowhood. *Journal of Gerontology: Psychological Sciences, 51B,* P279–P289.

Lindesay, J. (1991). Suicide in the elderly. *International Journal of Geriatric Psychiatry, 6,* 355–361.

Lindesay, J., Briggs, K., & Murphy, E. (1989). The Guy's/Age Concern Survey: Prevalence rates of cognitive impairment, depression and anxiety in an urban elderly community. *British Journal of Psychiatry, 155,* 317–329.

Logsdon, R. G., & Teri, L. (1995). Depression in Alzheimer's Disease patients: Caregivers as surrogate reporters. *Journal of the American Geriatrics Society, 43,* 150–155.

Lovett, S., & Gallagher, D. (1988). Psychoeducational interventions for family caregivers: Preliminary efficacy data. *Behavior Therapy, 19,* 321–330.

Lund, D. A., Caserta, M. S., Dimond, M. F., & Shaffer, S. K. (1989). Competencies, tasks of daily living, and adjustments to spousal bereavement in later life. In D. A. Lund (Ed.), *Older bereaved spouses: Research with practical applications* (pp. 135–152). Series in Death Education, Aging, and Health Care. New York: Hemisphere Publishing.

Lyness, J. M., Caine, E. D., Conwell, Y., King, D. A., & Cox, C. (1993). Depressive symptoms, medical illness, and functional status in depressed psychiatric inpatients. *American Journal of Psychiatry, 150,* 910–915.

Lyness, S. A., Eaton, E. M., & Schneider, L. S. (1994). Cognitive performance in older and middle-aged depressed outpatients and controls. *Journal of Gerontology: Psychological Sciences, 49,* 129–136.

Lyons, J. S., Strain, J. J., Hammer, J. S., Ackerman, A. D., & Fulop, G. (1989). Reliability, validity, and temporal stability of the Geriatric Depression Scale in hospitalized elderly. *International Journal of Psychiatry in Medicine, 19,* 203–209.

MacKinnon, A., Christensen, H., Cullen, J. S., Doyle, C. J., Henderson, A. S., Jorm, A. F., Korten, A. E., & Scott, L. R. (1993). The Canberra Interview for the Elderly: Assessment of its validity in the diagnosis of dementia and depression. *Acta Psychiatrica Scandinavica, 87,* 146–151.

Mann, A. H. (1995). Future directions for the specialty of old age psychiatry: Some suggestions for debate. *International Journal of Geriatric Psychiatry, 10,* 87–91.

Mann, A., Graham, N., & Ashby, D. (1984). Psychiatric illness in residential homes for the elderly: A survey in one London borough. *Age and Ageing, 13,* 257–265.

Marwit, S. J. (1991). *DSM-II-R,* grief reactions and a call for revision. *Professional Psychology: Research and Practice, 22,* 75–79.

Matt, G. E., Dean, A., Wang, B., & Wood, P. (1992). Identifying clinical syndromes in a community sample of elderly persons. *Psychological Assessment, 4,* 174–184.

McGuffin, P., Katz, R., Watkins, S., & Rutherford, J. (1996). A hospital-based twin register of the heritability of *DSM-IV* unipolar depression. *Archives of General Psychiatry, 53,* 129–136.

McElroy, M. R., Carmelli, D., Swan, G. E., & LaRue, A. (1995). Risk factors for the onset of depressive symptoms in older men: A 5-year follow-up of the Western Collaborative Group Study. Paper presented at the Gerontological Society of America, Los Angeles, CA.

McGarvey, B., Gallagher, D., Thompson, L. W., & Zelinski, E. (1982). Reliability and factor structure of the Zung Self-rating Depression Scale in three age groups. *Essence, 5,* 141–158.

McIntosh, J. L. (1992). Epidemiology of suicide in the elderly. In A. A. Leenaars, R. Maris, J. L. McIntosh, & J. Richman (Eds.), *Suicide and the older adult* (pp. 15–35). New York: Guilford Press.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1971/1981/1992). *Manual for the Profile of Mood States.* San Diego, CA: Educational and Industrial Testing Service.

Meyers, B. S., Greenberg, R., & Mei-Tal, V. (1985). Delusional depression in the elderly. In C. A. Shamoian (Ed.), *Treatment of affective disorders in the elderly.* Washington, DC: American Psychiatric Press.

Meyers, B. S., Kalayam, B., & Mei-Tal, V. (1984). Late-onset delusional depression: A distinct clinical entity? *Journal of Clinical Psychiatry, 45,* 347–349.

Migliorelli, R., Teson, A., Sabe, L., Petracchi, M., Leiguarda, R., & Starkstein, S. E. (1995). Prevalence and correlates of dysthymia and major depression among patients with Alzheimer's disease. *American Journal of Psychiatry, 152,* 37–44.

Miles, C. P. (1977). Conditions predisposing to suicide: A review. *Journal of Nervous and Mental Disease, 164,* 231–246.

Moldin, S. O., Reich, T., & Rice, J. P. (1991). Current perspectives on the genetics of unipolar depression. *Behavior Genetics, 21,* 211–242.

Montgomery, S. A., & Åsberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, 134, 382–389.

Mossey, J. M., Knott, K. A., Higgins, M., & Talerico, K. (1996). Effectiveness of a psychosocial intervention, interpersonal counseling, for subdysthymic depression in medically ill elderly. *Journal of Gerontology: Medical Sciences, 51A,* M172–M178.

Moye, J., Robiner, W. N., & Mackenzie, T. B. (1993). Depression in Alzheimer patients: Discrepancies between patient and caregiver reports. *Alzheimer Disease and Associated Disorders, 7,* 187–201.

Mrazek, P. J., & Haggerty, R. J. (Eds.). (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research.* Washington, DC: National Academy Press.

Muñoz, R. F., Gonzalez, G. M., & Starkweather, J. (1995). Automated screening for depression: Toward culturally and linguistically appropriate uses of computerized speech recognition. *Hispanic Journal of Behavioral Sciences, 17,* 194–208.

Murrell, S. A., & Himmelfarb, S. (1989). Effects of attachment, bereavement, and pre-event conditions on subsequent depressive symptoms in older adults. *Psychology and Aging, 4,* 166–172.

Murrell, S. A., Himmelfarb, S., & Wright, K. (1983). Prevalence of depression and its correlates in older adults. *American Journal of Epidemiology, 117,* 173–185.

Murrell, S. A. & Meeks, S. (1992). Depressive symptoms in older adults: Predispositions, resources, and life experiences. In K. W. Schaie & M. P. Lawton (Eds.), *Annual review of gerontology and geriatrics* (Vol. 11, pp. 261–275). New York: Springer.

Musetti, L., Perugi, G., Soriani, A., Rossi, V. M., Cassano, G. B., & Akiskal, H. S. (1989). Depression before and after age 65: A re-examination. *British Journal of Psychiatry, 155,* 330–336.

Naranjo C. A., Herrmann N., Mittmann N., & Bremner K. E. (1995). Recent advances in geriatric psychopharmacology. *Drugs and Aging, 7,* 184–202.

Nelson, J. C., Conwell, Y., Kim, K., & Mazure, C. (1989). Age at onset in late-life delusional depression. *American Journal of Psychiatry, 146,* 785–786.

Newmann, J. P. (1989). Aging and depression. *Psychology and Aging, 4,* 150–165.

Newmann, J. P., Klein, M. H., Jensen, J. E., & Essex, M. J. (1996). Depressive symptom experiences among older women: A comparison of alternative measurement approaches. *Psychology and Aging, 11,* 112–126.

Nezu, A. M., Nezu, A. M., & Perri, M. G. (1989). *Problem-solving therapy for depression: Theory, research, and clinical guidelines.* New York: Wiley.

Niederehe G. T. (1994). Psychosocial therapies with depressed older adults. In L. S. Schneider, C. F. Reynolds, III, B. D. Lebowitz, & A. J. Friedhoff (Eds.), *Diagnosis and treatment of depression in late life: Results of the NIH Consensus Development Conference* (pp. 293–315). Washington DC: American Psychiatric Press.

Norris, J. T., Gallagher, D. E., Wilson, A., & Winograd, C. H. (1987). Assessment of depression in geriatric medical outpatients: The validity of two screening measures. *Journal of the American Geriatrics Society, 35,* 989–995.

Norris, M. P., Snow-Turek, A. L., & Blankenship, L. (1995). Somatic depressive symptoms in the elderly: Contribution or confound? *Journal of Clinical Geropsychology, 1,* 5–17.

O'Connor, P. W., Pollitt, P. A., Roth, M., Brook, C. P. B., & Reiss, B. B. (1990). Memory complaints and impairment in normal depressed and demented elderly persons identified in a community survey. *Archives of General Psychiatry, 47,* 224–227.

O'Hara, M. W., Hinrichs, J. V., Kohout, F. J., Wallace, R. B., & Lemke, J. H. (1986). Memory complaint and memory performance in the depressed elderly. *Psychology and Aging, 3,* 208–214.

O'Hara, M. W., Kohout, F. J., & Wallace, R. B. (1985). Depression among the rural elderly: A study of prevalence and correlates. *Journal of Nervous and Mental Disease, 173,* 582–589.

Okimoto, J. T., Barnes, R. F., Veith, R. C., Raskind, M. A., Inui, T. S., & Carter, W. B. (1982). Screening for depression in geriatric medical patients. *American Journal of Psychiatry, 139,* 799–802.

Olin, J. T., Schneider, L. S., Eaton, E. M., Zemansky, M. F., & Pollock, V. E. (1992). The Geriatric Depression Scale and the Beck Depression Inventory as screening instruments in an older adult outpatient population. *Psychological Assessment, 4,* 190–192.

Ott, B. R. & Fogel, B. S. (1992). Measurement of depression in dementia: Self vs. clinician rating. *International Journal of Geriatric Psychiatry, 7,* 899–904.

Pachana, N. A., Gallagher-Thompson, D., & Thompson, L. W. (1994). Assessment of depression. In M. P. Lawton & J. A. Teresi (Eds.), *Focus on assessment techniques*. New York: Springer.

Parkin, D., & Stengel, E. (1965). Incidence of suicidal attempts in an urban community. *British Medical Journal, 2,* 133–138.

Parmelee, P. A., Katz, I. R., & Lawton, M. P. (1989). Depression among institutionalized aged: Assessment and prevalence estimation. *Journal of Gerontology: Medical Sciences, 44,* M22–M29.

Parmelee, P. A., Katz, I. R., & Lawton, M. P. (1991). The relation of pain to depression among institutionalized aged. *Journal of Gerontology: Psycholoical Sciences, 46,* P15–P21.

Parmelee, P. A., Katz, I. R., & Lawton, M. P. (1992). Depression and mortality among institutionalized aged. *Journal of Gerontology: Psychological Sciences, 47,* P3–P10.

Phifer, J. F., & Murrell, S. A. (1986). Etiologic factors in the onset of depressive symptoms in older adults. *Journal of Abnormal Psychology, 95,* 282–291.

Pond, C. D., Mant, A., Kehoe, L., Hewitt, H., & Brodaty. (1994). General practitioner diagnosis of depression and dementia in the elderly: Can academic detailing make a difference? *Family Practice, 11,* 141–147.

Poon, L. W. (1992). Toward an understanding of cognitive functioning in geriatric depression. *International Psychogeriatrics, 4* (Suppl. 2), 241–266.

Prigerson, H. G., Frank, E., Kasl, S. V., Reynolds, C. F., Anderson, B., Zubenko, G. S., Houck, P. R., George, C. J., & Kupfer, D. J. (1995). Complicated grief and bereavement-related depression as distinct disorders: Preliminary empirical validation in elderly bereaved spouses. *American Journal of Psychiatry, 152,* 22–30.

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1,* 385–401.

Radloff, L. S., & Teri, L. (1986). Use of the Center for Epidemiological Studies-Depression Scale with older adults. *Clinical Gerontologist, 5,* 119–136.

Rapp, S. R., Parisi, S. A., Walsh, D. A., & Wallace, C. E. (1988). Detecting depression in elderly medical inpatients. *Journal of Consulting and Clinical Psychology, 56,* 509–513.

Rapp, S. R., Smith, S. S., & Britt, M. (1990). Identifying comorbid depression in elderly medical patients: Use of the extracted Hamilton Depression Rating Scale. *Psychological Assessment, 2,* 243–247.

Rapp, S. R., & Vrana, S. (1989). Substituting nonsomatic for somatic symptoms in the diagnosis of depression in elderly male medical patients. *American Journal of Psychiatry, 146,* 1197–1200.

Raschko, R. (1985). Systems integration at the program level: Aging and mental health. *Gerontologist, 25,* 460–463.

Ravindran, A. V., Welburn, K., & Copeland, J. R. (1994). Semi-structured depression scale sensitive to change with treatment for use in the elderly. *British Journal of Psychiatry, 164,* 522–527.

Ray, W. A., Griffin, M. R., Schaffner, W., Baugh, D. L., & Melson. (1987). Psychotropic drug use and the risk of hip fracture. *New England Journal of Medicine, 316,* 363–369.

Reynolds, C. F., III, Frank, E., Perel, J. M., Imber, S. D., Cornes, C., Morycz, R. K., Mazumdar, S., Miller, M. D., Pollock, B. G., Rifai, A. H., Stack, J. A., George, C. J., Houck, P. R., & Kupfer, D. J. (1992). Combined pharmacotherapy and psychotherapy in the acute and continuation treatment of elderly patients with recurrent major depression: A preliminary report. *American Journal of Psychiatry, 149,* 1687–1692.

Reynolds, W. M., & Kobak, K. A. (1991). *Hamilton Depression Inventory: A self-report version of the Hamilton Depression Rating Scale*. Odessa, FL: Psychological Assessment Resources.

Reynolds, C. F., Schneider, L. S., Lebowitz, B. D., & Kupfer, D. (1994). Treatment of depression in elderly patients: Guidelines for primary care. In L. S. Schneider, C. F. Reynolds, III, B. D. Lebowitz, & A. J. Friedhoff (Eds.), *Diagnosis and treatment of depression in late life: Results of the NIH Consensus Development Conference* (pp. 463–490). Washington, DC: American Psychiatric Press.

Riskind, J. H., Beck, A. T., & Steer, R. A. (1985). Cognitive-behavioral therapy in geriatric depression: Comment on Steuer et al. *Journal of Consulting and Clinical Psychology, 53,* 944–945.

Robins, L. N., Helzer, J. E., Croughan, J., & Ratcliff, K. S. (1981). National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics, and validity. *Archives of General Psychiatry, 38,* 381–389.

Robins, L. N., Wing, J., Wittchen, H. U., Helzer, J. E., Babor, T. F., Burke, J., Farmer, A., Jablenski, A., Pickens, R., Regier, D. A., Sartorius, N., & Towle, L. H. (1988). The Composite International Diagnostic Interview. An epidemiological instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry, 45,* 1069–1077.

Robinson, R. G., & Starkstein, S. E. (1990). Current research in affective disorders following stroke. *Journal of Neuropsychiatry, 2,* 1–14.

Roth, M., Tym, E., Mountjoy, C. Q., Huppert, F. A., Hendrie, H., Verma, S., & Goddard, R. (1986). CAMDEX: A standardized instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia. *British Journal of Psychiatry, 149,* 698–709.

Rubin, R. T. (1994). Neuroendocrine aspects of stress in major depression. In R. P. Lieberman & J. Yager (Eds.), *Stress in psychiatric disorders*. New York: Springer.

Russell, D. W., & Cutrona, C. E. (1991). Social support, stress, and depressive symptoms among the elderly: Test of a process model. *Psychology and Aging, 6,* 190–201.

Rybarczyk, B., Gallagher-Thompson, D., Rodman, J., Zeiss, A., Gantz, F. E., & Yesavage, J. (1992). Applying cognitive-behavioral psychotherapy to the chronically ill elderly: Treatment issues and case illustration. *International Psychogeriatrics, 4,* 127–140.

Sackeim, H. A. (1994). Use of electroconvulsive therapy in late-life depression. In L. S. Schneider, C. F. Reynolds, III, B. D. Lebowitz, & A. J. Friedhoff (Eds.), *Diagnosis and treatment of depression in late life: Results of the NIH Consensus Development Conference* (pp. 259–277). Washington, DC: American Psychiatric Press.

Salzman, C. (1994). Pharmacological treatment of depression in elderly patients. In L. S. Schneider, C. F. Reynolds, III, B. D. Lebowitz, & A. J. Friedhoff (Eds.), *Diagnosis and treatment of depression in late life: Results of the NIH Consensus Development Conference* (pp. 181–244). Washington, DC: American Psychiatric Press.

Schneider L. S., Sloane R. B., Staples F. R., & Bender M. (1986). Pretreatment orthostatic hypotension as a predictor of response to nortriptyline in geriatric depression. *Journal of Clinical Psychopharmacology, 6,* 172–176.

Schulz, R., & Williamson, G. M. (1994). Health effects of caregiving: Prevalence of mental and physical illness in Alzheimer's caregivers. In E. Light, G. Niederehe, & B. D. Lebowitz (Eds.), *Stress effects on family caregivers of*

Alzheimer's patients (pp. 38–63). New York: Springer.

Schuster, T. L., & Butler, E. W. (1989). Bereavement, social networks, social support, and mental health. In D. A. Lund (Eds.), *Older bereaved spouses: Research with practical applications* (pp. 55–68). New York: Hemisphere Publishing.

Scogin, F. R. (1994). Assessment of depression in older adults: A guide for practitioners. In M. Storandt & G. R. Vanden Bos (Eds.), *Neuropsychological assessment of dementia and depression in older adults: A clinician's guide* (pp. 61–80). Washington, DC: American Psychological Association.

Scogin, F., & McElreath, L. (1994). Efficacy of psychosocial treatments for geriatric depression: A quantitative review. *Journal of Consulting and Clinical Psychology, 62,* 69–74.

Segal, D. L., Hersen, M., Van Hasselt, V. B., Kabacoff, R. I., & Roth, L. (1993). Reliability of diagnosis in older psychiatric patients using the Structured Clinical Interview for DSM-III-R. *Journal of Psychopathology and Behavioral Assessment, 15,* 347–356.

Seltzer, B., Vasterling, J. J., & Buswell, A. (1995). Awareness of deficit in Alzheimer's disease: Associations with psychiatric symptoms and other disease variables. *Journal of Clinical Geropsychology, 1,* 79–87.

Sheikh, J. I., & Yesavage, J. A. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clinical Gerontologist, 5,* 165–173.

Shulman, K. I., Tohen, M., Satlin, A., Mallya, G., & Kalunian, D. (1992). Mania compared with unipolar depression in old age. *American Journal of Psychiatry, 149,* 341–345.

Silverman, P. R. (1969). The Widow-to-Widow Program: An experiment in preventive intervention. *Mental Hygiene, 53,* 333–337.

Small, G. W., Komanduri, R., Gitlin, M., & Jarvik, L. (1986). The influence of age on guilt expression in major depression. *International Journal of Geriatric Psychiatry, 1,* 121–126.

Smouse, P. E., Feinberg, M., Carroll, B. J., Park, & Rawson, S. G. (1981). The Carroll Rating Scale for Depression: II. Factor Analysis of the feature profiles. *British Journal of Psychiatry, 138,* 201–204.

Smyer, M. A. (1995). Prevention and early intervention for mental disorders of the elderly. In M. Gatz (Ed.), *Emerging issues in mental health and aging* (pp. 163–182). Washington, DC: American Psychological Association.

Smyer, M. A., Zarit, S. H., & Qualls, S. H. (1990). Psychological intervention with the aging individual. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed.). San Diego, CA: Academic Press.

Snowdon, J., & Lane, F. (1995). The Botany Survey: A longitudinal study of depression and cognitive impairment in an elderly population. *International Journal of Geriatric Psychiatry, 10,* 349–358.

Spitzer, R. L., Endicott, J., & Robins, E. (1978). Research Diagnostic Criteria: Rationale and reliability. *Archives of General Psychiatry, 35,* 773–782.

Spitzer, R. L., Williams, J. B. W., Gibbon, M., & First, M. B. (1992). The structured clinical interview for *DSM-III-R* (SCID) I: History, rationale, and description. *Archives of General Psychiatry, 49,* 624–629.

Stallones, L., Marx, M. B., & Garrity, T. F. (1990). Prevalence and correlates of depressive symptoms among older U. S. adults. *American Journal of Preventive Medicine, 6,* 295–303.

Steingart, A., & Herrmann, N. (1991). Major depressive disorder in the elderly: The relationship between age of onset and cognitive impairment. *International Journal of Geriatric Psychiatry, 6,* 593–598.

Steuer, J. L., Mintz, J., Hammen, C. L., Hill, M. A., Jarvik, L. F., McCarley, T., Motoike, P., & Rosen, R. (1984). Cognitive-behavioral and psychodynamic group psychotherapy in treatment of geriatric depression. *Journal of Consulting and Clinical Psychology, 52,* 180–189.

Stukenberg, K. W., Dura, J. R., & Kiecolt-Glaser, J. K. (1990). Depression screening scale validation in an elderly, community-dwelling population. *Psychological Assessment, 2,* 134–138.

Sultzer, D. L., Levin, H. S., Mahler, M. E., High, W. M., & Cummings, J. L. (1992). Assessment of cognitive, psychiatric, and behavioral disturbances in patients with dementia: the Neurobehavioral Rating Scale. *Journal of the American Geriatrics Society, 40,* 549–555.

Sunderland, T., Alterman, I. S., Yount, D., Hill, J. L., Tariot, P. N., Newhouse, P. A., Mueller, E. A., Mellow, A. M., & Cohen, R. M. (1988). A new scale for the assessment of depressed mood in demented patients. *American Journal of Psychiatry, 145,* 955–959.

Takahashi, Y., Hirasawa, H., Koyama, K., Asakawa, O., Kido, M., Onose, H., Udagawa, M., Ishikawa, Y., & Uno, M. (1995). Suicide and aging in Japan: An examination of treated elderly suicide attempters. *International Psychogeriatrics, 7,* 239–251.

Teri, L., Curtis, J., Gallagher-Thompson, D., & Thompson, L. (1994). Cognitive-behavior therapy with depressed older adults. In L. S. Schneider, C. F. Reynolds, III, B. D. Lebowitz, & A. J. Friedhoff (Eds.), *Diagnosis and treatment of depression in late life: Results of the NIH Consensus Development Conference* (pp. 279–291). Washington, DC: American Psychiatric Press.

Teri L., Logsdon R., Wagner A., & Uomoto J. (1994). The caregiver role in behavioral treatment of depression in dementia patients. In E. Light, G. Niederehe, & B. D. Lebowitz (Eds.), *Stress effects on family caregivers of Alzheimer's patients: Research and interventions* (pp. 185–204). New York: Springer.

Teri, L., & Wagner, A. W. (1991). Assessment of depression in patients with Alzheimer's disease: Concordance among informants. *Psychology and Aging, 6,* 280–285.

Thomas, C., Kelman, H. R., Kennedy, G. J., Ahn, C., & Yang, C. (1992). Depressive symptoms and mortality in elderly persons. *Journal of Gerontology: Social Sciences, 47,* 80–87.

Thompson, L. W., Gallagher, D. E., & Breckinridge, J. S. (1987). Comparative effectiveness of psychotherapies for depressed elders. *Journal of Consulting and Clinical Psychology, 55,* 385–390.

Thompson, L. W., Gallagher, D. E., Nies, G., & Epstein, D. (1983). Evaluation of the effectiveness of professionals and nonprofessionals as instructors of "Coping with Depression" classes for elders. *Gerontologist, 23,* 390–396.

Torgersen, S. (1986). Genetic factors in moderately severe and mild affective disorders. *Archives of General Psychiatry, 43,* 222–226.

Tucker, M. A., Ogle, S. J., Davison, J. G., Eilenberg, M. D. (1986). Development of a brief screening test for depression in the elderly. *Journal of Clinical Experimental Gerontology, 8,* 173–190.

Turner, R. J., & Beiser, M. (1990). Major depression and depressive symptomology among the physically disabled: Assessing the role of chronic stress. *Journal of Nervous and Mental Disease, 178,* 343–350.

Uhlenhuth, E. H., Balter, M. B., Mellinger, G. D., Cisin, I. H., & Clinthorne, J. (1983). Symptom checklist syndromes in the general population: Correlations with psychotherapeutic drug use. *Archives of General Psychiatry, 40,* 1167–1173.

Vachon, M. L. S., Sheldon, A. R., Lancee, W. J., Lyall, W. A. L., Roger, J., & Freeman, S. J. J. (1980). A controlled study of self-help intervention for widows. *American Journal of Psychiatry, 137,* 1380–1384.

Vachon, M. L. S., & Stylianos, S. K. (1988). The role of social support in bereavement. *Journal of Social Issues, 44,* 175–190.

van Ojen, R., Hooijer, C., Bezemer, D., Jonker, C., Lindeboom, J., & van Tilburg, W. (1995a). Late-life depressive disorder in the community: I. The relationship between MMSE score and depression in subjects with and without psychiatric history. *British Journal of Psychiatry, 166,* 311–315.

van Ojen, R., Hooijer, C., Bezemer, D., Jonker, C., Lindeboom, J., & van Tilburg, W. (1995b). Late-life depressive disorder in the community: II. The relationship between psychiatric history, MMSE and family history. *British Journal of Psychiatry, 166,* 316–319.

Vida, S., Des Rosiers, P., Carrier, L., & Gauthier, S. (1994a). Depression in Alzheimer's disease: Receiver operating characteristics analysis of the Cornell Scale for Depression in Dementia and the Hamilton Depression Scale. *Journal of Geriatric Psychiatry and Neurology, 7,* 159–162.

Vida, S., Des Rosiers, P., Carrier, L., & Gauthier, S. (1994b). Prevalence of depression in Alzheimer's disease and validity of research diagnostic criteria. *Journal of Geriatric Psychiatry and Neurology, 7,* 238–244.

Wallace, J., & O'Hara, M. W. (1992). Increases in depressive symptomatology in the rural elderly: Results from a cross-sectional and longitudinal study. *Journal of Abnormal Psychology, 101,* 398–404.

Ward, L. C., Wadsworth, A. P., & Peterson, L. P. (1994). Concurrent validity of measures of anxiety, depression, and somatization in elderly, demented, male patients. *Clinical Gerontologist, 15,* 3–13.

Warren, C. A., & Levy, K. A. (1991). Electroconvulsive therapy and the elderly. *Journal of Aging Studies, 5,* 309–327.

Waters, E. (1995). Let's not wait 'till it's broke: Interventions to maintain and enhance mental health in late life. In M. Gatz (Ed.), *Emerging issues in mental health and aging* (pp. 183–209). Washington, DC: American Psychological Association.

Weissman, M. M., Bruce, M. L., Leaf, P. J., Florio, L. P., & Holzer, C., III. (1991). Affective disorders. In L. N. Robins & D. A. Regier (Eds.), *Psychiatric disorders in America* (pp. 53–80). New York: Free Press.

Weissman, M. M., Leaf, P. J., Holzer, C. E., Myers, J. K., & Tischler, G. L. (1984). The epidemiology of depression: An update on sex differences in rates. *Journal of Affective Disorders, 7,* 179–188.

Weissman, M. M., Leaf, P. J., Tischler, G. L., Blazer, D. G., Karno, M., Bruce, M. L., & Florio, L. P. (1988). Affective disorders in five United States communities. *Psychological Medicine, 18,* 141–153.

Weissman, M. M., & Myers, J. K. (1978). Affective disorders in a US urban community: The use of Research Diagnostic Criteria in an epidemiological survey. *Archives of General Psychiatry, 35,* 1304–1311.

Wells, K. B., Stewart, A., Hays, R. D., Burnam, M. A., Rogers, W., Daniels, M. Berry, S., Greenfield, S., & Ware, J. (1989). The functioning and well-being of depressed patients: Results from the Medical Outcomes Study. *Journal of the American Medical Association, 262,* 914–919.

Wernicke, T. F., & Linden, M. (1994). Clinical recognition and drug treatment of depression in cases found by standardized assessment. *Pharmacopsychiatry, 27* (Suppl.), 54–57.

Wetzel, R. D., Margulies, T., Davis, R., & Karam, E. (1980). Hopelessness, depression, and suicide intent. *Journal of Clinical Psychiatry, 41,* 159–160.

Wexler, B. E., & Cicchetti, D. V. (1992). The outpatient treatment of depression: Implications of outcome research for clinical practice. *Journal of Nervous and Mental Disease, 180,* 277–286.

White House Mini-Conference on Emerging Issues in Mental Health and Aging. (1995). Resolutions of the White House Mini-Conference on Emerging Issues in Mental Health and Aging. In M. Gatz (Ed.), *Emerging issues in mental health and aging* (pp. 329–364). Washington DC: American Psychological Association.

Williams, J. B. (1988). A structured interview guide for the Hamilton Depression Rating Scale. *Archives of General Psychiatry, 45,* 742–747.

Williamson, G. M., & Schulz, R. (1992a). Pain, activity restriction, and symptoms of depression among community-residing elderly adults. *Journal of Gerontology: Psychological Sciences, 47,* 367–372.

Williamson, G. M. & Schulz, R. (1992b). Physical illness and symptoms of depression among elderly outpatients. *Psychology and Aging, 7,* 343–351.

Wing, J. K., Birley, J. L. T., Cooper, J. E., Graham, P., & Isaacs, A. D. (1967). Reliability of procedure for measuring and classifying "Present Psychiatric State." *British Journal of Psychiatry, 113,* 499–515.

Wolfe, R., Morrow, J., & Fredrickson, B. L. (1995). Mood disorders in older adults. In L. L. Carstensen, B. A. Edelstein, & L. Dornbrand (Eds.), *The practical handbook of clinical gerontology*. Thousand Oaks, CA: Sage.

World Health Organization (1977). *Manual of the international statistical classification of diseases, injuries, and causes of death* (9th rev., Vol. 1). Geneva: Author.

World Health Organization (1992–1993). *International statistical classification of diseases and related health problems* (10th rev.). Geneva: Author.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., Leirer, V. O. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research, 17,* 37–49.

Zemore, R., & Eames, N. (1979). Psychic and somatic symptoms of depression among young adults, institutionalized aged and noninstitutionalized aged. *Journal of Gerontology, 34,* 716–722.

Zimmerman, M., & Coryell, W. (1987). The inventory to diagnose depression, lifetime version. *Acta Psychiatrica Scandinavica, 75,* 495–499.

Zimmerman, M., Coryell, W., Corenthal, C., & Wilson, S. (1986). A self-report scale to diagnose major depressive disorder. *Archives of General Psychiatry, 43,* 1076–1081.

Zorumski, C.F., Rubin E. H., & Burke, W. J. (1988). Electroconvulsive therapy for the elderly: A review. *Hospital and Community Psychiatry, 39,* 643–647.

Zuckerman, M., & Lubin, B. (1965). *Manual for the Multiple Affect Adjective Checklist*. San Diego, CA: EdITS/Educational and Industrial Testing Service.

Zuckerman, M., & Lubin, B. (1985). Manual for the MAACL-R: The Multiple Affect Adjective Checklist Revised. San Diego, CA: Educational and Industrial Testing Service.

Zung, W. W. K. (1965). A self-rating depression scale. *Archives of General Psychiatry, 12,* 63–70.

Zung, W. W. K. (1967). Depression in the normal aged. *Psychosomatics, 13,* 287–292.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.10
Cognitive Development in Alzheimer's Disease: Charting the Decline Process

BRENT J. SMALL
Stockholm Gerontology Research Center, Sweden and Karolinska Institute, Stockholm, Sweden
AGENTA HERLITZ
Stockholm Gerontology Research Center, Sweden and Stockholm University, Sweden
and
LARS BÄCKMAN
Stockholm Gerontology Research Center, Sweden and Göteborg University, Sweden

7.10.1 INTRODUCTION 231
7.10.2 SAMPLE COMPOSITION 233
7.10.3 INCIDENCE OF DEMENTIA 234
7.10.4 DETECTION OF DEMENTIA 236
7.10.5 STAGING OF DEMENTIA 237
7.10.6 PROGRESSION OF DEMENTIA 238
7.10.7 TRAJECTORY OF DECLINE 239
7.10.8 PREDICTORS OF DEMENTIA-RELATED DIFFERENCES AND CHANGES 240
7.10.9 CONCLUSIONS AND EXTENSIONS 241
7.10.10 SUMMARY 242
7.10.11 REFERENCES 242

7.10.1 INTRODUCTION

Several reviews agree that Alzheimer's disease (AD) results in a global cognitive impairment that involves deficits in perception, attention, psychomotor speed, memory, visuospatial functioning, intelligence, and language (Bäckman, 1992; Morris, 1996; Nebes, 1992).

This extension of cognitive impairment in AD is to be expected, given that this disease is associated with marked lesions in limbic structures as well as in the neocortical association areas (Braak & Braak, 1991, 1995; Frackowiak et al., 1981; Grady et al., 1990; Haxby et al., 1990).

However, the view that AD affects our cognitive functions in a global manner does not necessarily imply that all functions decline at the same time and at the same rate. By contrast, some variability in the degree to which different cognitive functions are compromised should be expected, given that (i) the neuropathological trajectory in AD may begin with lesions in medial-temporal regions that gradually radiate into neocortical (e.g., temporal, parietal, and frontal) areas (Braak & Braak, 1991, 1995; Fewster, Griffin-Brooks, MacGregor, Ojalvo-Rose, & Ball, 1991); and (ii) different cognitive tasks vary with regard to the extent to which overlearned skills, procedures, and knowledge could be utilized. Conceivably, this variability is likely to decrease as a function of increasing severity of dementia. Knowledge regarding whether AD related cognitive deficits occur in a lawful manner across the pathogenesis would seem to be imperative not only for advancing our scientific understanding of the disease itself, but also to facilitate early clinical diagnosis and be useful in behavioral intervention.

The purpose of this chapter is to provide an overview of recent research addressing the aforementioned issues. The overview is organized so as to parallel the disease development. That is, we begin by reviewing the evidence pertaining to potential preclinical cognitive markers of AD. This is followed by an examination of the cognitive tasks that may or may not be useful in differentiating early AD patients from normal old adults. The next step is to determine whether the same cognitive tasks that are particularly effective at detecting AD are also most sensitive in discriminating between patients at different phases in the disease process.

Most of the work examining the efficacy of cognitive measures in detecting and staging AD is cross-sectional in nature; this is reflected in the present review. However, an increasing number of longitudinal studies charting the cognitive progression in AD are available. Examining the results from these studies not only allows for an interesting comparison with the cross-sectional research on detection and staging of AD but also it provides the necessary means by which different cognitive functions can be compared with regard to rate of pathological change.

Although a major goal of the chapter is to examine whether there is a predictable development of cognitive decline in AD, there may also be individual differences among patients with regard to both the magnitude and rate of change in various cognitive functions. Individual difference variables (e.g., demographic, psychometric, biological) have been found to be powerful mediators of cognitive performance in normal aging (e.g., Elias, Wolf, D́Agostino, Cobb, & White, 1993; Wahlin et al., 1993; West, Crook, & Barron, 1992). Thus, an additional objective of this chapter is to determine whether variability in cognitive performance among AD patients can systematically be linked to specific subject-related factors.

Before describing the changes in cognitive performance that accompany AD, it is important to address a number of practical problems associated with studying this population of individuals. The most basic problem concerns the uncertainty of the diagnosis itself. Although the diagnostic criteria in the last few years have become increasingly strict (American Psychiatric Association, 1980, 1987, 1994; McKhann et al., 1984), there are still no specific clinical diagnostic markers for AD. Presently, the clinical diagnostic criteria are descriptive and exclusive, and pathological confirmation remains the current "gold standard" for AD diagnosis (Galasko et al., 1994; Huppert, Brayne, & O'Connor, 1994). Unfortunately, in AD research most studies have relied on the clinical diagnosis as the gold standard, which means that the criterion variable is far from perfect. It is possible that the clinical diagnostic criteria are not sensitive enough to detect very old subjects in a very early phase of AD. For example, what defines normality (e.g., in terms of physical, sensory, and cognitive status) may be rather equivocal in very advanced age (Brody & Schneider, 1986; Fries & Crapo, 1981), making diagnostic decisions uncertain (e.g., Fratiglioni, Grut, Forsell, Viitanen, & Winblad, 1992). Nevertheless, because these clinical criteria are one of the only means, at the present time, for rendering dementia diagnoses we must utilize what is available to us.

In addition to problems associated with dementia diagnosis, there are a number of issues regarding sample composition, such as sample selection, sample size, and dementia etiology that require consideration. These three facets all have the potential to influence the results that are found from studies that have examined cognitive performance in demented persons. The following section addresses these aspects of sample composition.

7.10.2 SAMPLE COMPOSITION

The source from which researchers select their populations of demented and nondemented (control) subjects has important implications for the severity of impairment that is observed, as well as the extent to which the results that are produced are generalizable to the population of adults at large. Typically, samples of demented and nondemented subjects are selected from clinical or population settings. One of the advantages of selecting patients from a clinical setting is often that relatively large numbers of subjects are available for study. Further, the accuracy of the dementia diagnosis is likely to be more precise, because clinicians are allowed to spend more time interviewing the patient and relevant family members as compared to population-based studies (Brayne et al., 1995). In addition, various diagnosis-related assessment procedures, including radiological assessment (e.g., computerized tomography, magnetic resonance imaging) as well as other medical procedures (e.g., blood sampling) are easier to implement in a clinical setting, thus enhancing diagnostic accuracy. Alternatively, there are also several disadvantages in selecting subjects from clinical populations. For example, it is likely that these subjects are biased towards being more demented than subjects from the community. In addition, the controls used in a clinical setting are likely to be healthier and cognitively superior compared with a random sample from the community. Both of these factors may result in an overestimation of the differences between patients and controls. Moreover, the greater initial impairment of the clinical population may make investigating mild changes in dementia difficult, because these populations are dominated by persons with more obvious signs of cognitive deficiencies. This may have negative repercussions concerning the generalizability of the findings that are observed. That is, when individuals self-select themselves for assessment in a clinical setting, the composition of this group, both demographically and cognitively, may be very different from studies who use patients that are sampled in a more representative fashion.

One of the chief advantages of sampling adults from the community is the ability to make generalizations back to the larger population. In addition, more diversity in terms of cognitive performance and demographic characteristics may be present in the samples as compared to clinical populations. This diversity is represented not only as milder levels of impairment among the demented subjects, but also in terms of increased variability in performance among those who are nondemented. Although there are benefits in examining cognitive performance in community-based samples of adults, there are also several disadvantages. Chief among them is the small sample size of demented adults that is often available. The relatively small sample size of demented persons in population-based studies is largely a byproduct of prevalence and incidence rates of AD. Several studies have shown that between 1.3 and 11.2% of adults over the age of 65 have AD (see Ebly, Parhad, Hogan, & Fung, 1994, for a review). Although prevalence rates increase with age, it remains that the population that must be sampled to achieve sufficient numbers of AD patients needs to be relatively large. Similar problems of small sample sizes are evident in studies that have examined incidence of dementia (e.g., Katzman et al., 1989; Tierney et al., 1996).

To combat problems with small sample sizes of AD patients, some authors have resorted to combining different dementia etiologies, typically AD and vascular dementia (VaD), into one group (e.g., Almkvist & Bäckman, 1993a, Herlitz, Hill, Fratiglioni, & Bäckman, 1995; Small & Bäckman, 1996). Although this procedure does reduce problems associated with sample size, it can be criticized on the grounds of potentially obscuring relationships with cognitive performance because of the different dementia types. For example, Gottfries, Blennow, Wallin, and Gottfries (1995) found that AD and VaD patients significantly differed in a number of cognitive dimensions. Specifically, on the one hand, AD patients experienced more parietal symptoms (e.g., sensory aphasia, visuospatial dysfunction), whereas VaD patients experienced more symptoms related to confusion (e.g., clouded consciousness, attentional problems) and subcortical impairments (e.g., psychomotor retardation). However, there is a growing body of evidence indicating nearly identical patterns of cross-sectional performance in tasks assessing various aspects of episodic memory for those individuals with AD and those with VaD (e.g., Almkvist, 1994; Almkvist, Bäckman, Basun, & Wahlund, 1993; Erkinjuntti, Laaksonen, Sulkava, Syrjäläinen, & Palo, 1986). A reasonable recommendation is that studies utilizing more than one dementia type should incorporate a test of the differences between the groups in their analyses. Only after different dementia groups have been found to be indistinguishable in the measures of relevance, is collapsing across dementia etiology justifiable.

In this chapter, we review studies that have focused exclusively on subjects with AD, but also on studies that have combined different dementia etiologies. This has been done for purely pragmatic reasons; the pool of available studies is so much larger when accounts that have used both subject selection methods are employed. Nevertheless, we will orient the reader to the use of these different subject groups and, where possible, indicate whether between-group comparisons were made.

7.10.3 INCIDENCE OF DEMENTIA

A number of studies have examined potential preclinical cognitive markers of dementia. In these studies, subjects are followed longitudinally and incident cases of dementia are identified. Investigators then compare the baseline cognitive performance of these two groups of individuals to determine whether a certain profile of cognitive performance is predictive of who will develop dementia. Identifying certain patterns of performance that may be particularly diagnostic of future dementia status has important clinical implications. Persons may be identified relatively early in the disease process and appropriate behavioral or pharmacological therapies may be designed to best deal with the impending disease process. Further, supportive opportunities (e.g., support groups, day care arrangements) for family members can be developed in attempts to lessen the impact of the dementia on loved ones.

Several studies have investigated the relationship between baseline performance on global indices of cognitive functioning and dementia incidence (Aronson et al., 1990; Katzman et al., 1989; Yoshitake et al., 1995). These measures, such as the Mini-mental state examination (MMSE; Folstein, Folstein, & McHugh, 1975), the Mattis dementia rating scale (MDRS; Mattis, 1976), and the Blessed information-concentration test (BIMC; Blessed, Tomlinson, & Roth, 1968) are designed to index a person's global level of cognitive performance. In general, research indicates that individuals who initially score at low levels on these cognitive indices are more likely to develop dementia than the remaining study sample. For example, Aronson et al. (1990) examined this relationship with the BIMC and found baseline performance on this measure to be a strong predictor of dementia incidence (AD, VaD, and mixed dementia). Similar results were reported by Yoshitake et al. (1995), whereby baseline performance on a Japanese test of global cognitive functioning, Hasegawa's dementia scale, was predictive of subsequent development of AD.

Although some investigators have focused on measures of global cognitive functioning as predictors of dementia incidence, others have examined specific aspects of cognitive performance. In these studies, larger batteries of cognitive tests, comprised of indices assessing domains such as episodic memory, visuospatial functioning, and verbal ability were administered at baseline. Among the specific measures of performance that have been found to be related to dementia incidence, evidence has indicated that markers of verbal ability, including the Boston naming test and measures of category fluency (Jacobs et al., 1995; Masur, Sliwinski, Lipton, Blau, & Crystal, 1994) may predict dementia incidence. In addition, Masur et al. (1994) found the digit symbol subtest from the Wechsler adult intelligence scale-revised (WAIS-R; Wechsler, 1981) to be predictive of dementia incidence, whereas Jacobs et al. found another WAIS-R subtest, word similarities, to be significantly related to incidence of AD.

Although these indices of verbal ability (category fluency, Boston naming), speeded psychomotor skill (digit symbol), and abstract reasoning (word similarities) have been found to be predictive of dementia incidence, the most salient predictor of who will become afflicted with dementia is baseline episodic memory performance. Numerous studies have found that individuals who perform poorly on measures of recall and recognition are more likely to develop dementia after a follow-up interval than their superior-performing counterparts. For example, Masur and colleagues (Masur, Fuld, Blau, Crystal, & Aronson, 1990; Masur et al., 1994) have found the delayed recall measure from the Buschke selective reminding test to be an effective predictor of dementia incidence. Similar relationships have been reported with the Fuld object memory test (Fuld, Masur, Blau, Crystal, & Aronson, 1990; Masur et al., 1994), and the logical memory subtest of the Wechsler memory scale (Linn, et al., 1995).

Tierney et al. (1996) addressed the relative importance of different memory tasks in predicting incident cases of AD. They examined 123 memory-impaired patients and identified 29 probable AD patients after a two-year follow-up period. At baseline, all subjects completed a comprehensive neuropsychological battery including multiple indices of memory, verbal fluency, and psychomotor speed. The results of a logistic regression analysis, predicting AD status (yes/no), indicated that two tests (delayed recall from the Rey auditory verbal learning test and the mental control subtest of the WMS) were predictive of dementia incidence. Furthermore, these variables were fairly successful at correctly classifying the proportion of patients diagnosed with dementia (sensitivity 76%), as well as correctly identifying those who were free of dementia (specificity 94%). This resulted in an overall accuracy rate of the model of 89%. The additional measures of verbal fluency and psychomotor speed added little diagnostic accuracy in predicting which individuals would develop AD. Taken together, the results from this research indicate that episodic memory measures may be the most salient predictors of dementia incidence.

Although the relationship between episodic memory performance and dementia incidence is well established, it is also apparent that not all memory tasks are created equal. For example, tasks may differ in terms of the amount of cognitive support that is offered the individual (see Bäckman, Mäntylä, & Herlitz, 1990; Herlitz, Lipinska, & Bäckman, 1992, for reviews). Cognitive support refers to characteristics of memory tasks that make successful retrieval more or less likely. For example, free recall of unrelated words can be considered to be a task with relatively little cognitive support. This is due to the fact that the to-be-remembered items have little inherent organizability, and the recall condition offers no cues as to which items should be remembered. One can increase the level of cognitive support in this task by selecting words from several superordinate categories (e.g., articles of clothing, musical instruments). This change makes the task more supportive by supplying organizational structure to the materials. In this manner, the subject can arrange the words into their respective categories facilitating memory for these items (Bäckman & Wahlin, 1995). Cognitive support can also be manipulated at retrieval. Instead of the relatively impoverished retrieval conditions prevailing in free recall, a cued recall task may be given whereby the subject is prompted by word stems or category names in order to improve recall performance. In addition, a recognition task, in which subjects are asked to choose among target and distractor items, could be used to further increase the supportive nature of the retrieval task (Craik, 1983).

Several authors have suggested that the supportive nature of episodic memory tasks can influence the success of these tasks in detecting early cases of dementia. For example, Branconnier, Cole, Spera, and DeVitt (1982) found that tasks that are well supported (e.g., recognition) were the best discriminators between mild cases of dementia and normal aging. A possible reason for this finding is that many older adults, especially very old adults, may have deficits in cognitively demanding tasks such as free recall, whereas recognition and cued recall are relatively well preserved (Bäckman & Larsson, 1992; Craik & McDowd, 1987; Schonfield & Robertson, 1966). Thus, when deficits in the well-supported task are present, this may indicate conditions other than normal aging (e.g., dementia, depression).

The role of cognitive support in episodic memory tasks in predicting incidence of dementia has received less empirical attention. Although several studies have included multiple measures of episodic memory (e.g., Fuld et al., 1990; Masur et al., 1990), the lack of systematic variation among memory tasks has made firm conclusions regarding the relationship between cognitive support and dementia status difficult. Small, Herlitz, Fratiglioni, Almkvist, and Bäckman (1997) attempted to examine the issue of cognitive support in episodic memory and incidence of dementia by including multiple episodic memory tasks that varied in terms of a number of encoding and retrieval demands. Specifically, tasks were chosen that varied in terms of encoding demands (e.g., unrelated vs. organizable items), retrieval conditions (e.g., recall vs. recognition), as well as the nature of the materials that were examined (e.g., faces vs. words). In this study, 26 incident cases of AD were identified after a three-year follow-up interval, whereas 179 remained nondemented. In addition to the memory tasks, measures of visuospatial ability (block design, clock test, Poppelreuter's figures), verbal fluency (letter and category fluency), and primary memory (digit span) were included to examine the predictive relationship between these measures and incidence of AD.

Results from a logistic regression analysis indicated that three measures, recall of organizable words, face recognition, and letter fluency, were significant predictors of AD. Other less supported memory tasks (e.g., free recall of unrelated words), as well as other measures of visuospatial ability, category fluency, and primary memory were not predictive of dementia incidence. By examining the sensitivity and specificity of the model, it was determined that the two memory measures, organized word recall and face recognition, provided the greatest ability to detect who would become afflicted with AD (sensitivity 62%). On the other hand, although letter fluency added significantly to the logistic regression equation, its role was in improving the ability of the model to detect individuals who would remain nondemented (specificity 99%). Thus, among the predictors of dementia incidence, it was the two episodic memory tasks that were most effective at predicting incident cases of AD. In addition, the results of this study confirmed the hypothesis that more supported memory tasks (i.e., face recognition, recall of organizable words) rather than less supported tasks (e.g., recall of unrelated words) are most effective at predicting incident cases of AD.

In summary, the results from studies that have examined preclinical cognitive markers of dementia provide important information as to the changes that accompany the pathogenesis of AD. Specifically, tests of episodic memory performance appear to be the most powerful predictors of dementia incidence. On the one hand, recent research suggests that episodic memory tasks that are cognitively well supported are the best predictors of development of dementia. On the other hand, the results from all relevant studies have shown that markers of other abilities, including verbal and visuospatial skill, are less effective predictors of dementia incidence. Thus, it seems wise to include episodic memory tasks, and especially those that are well supported (e.g., recognition, recall of organizable words) when developing a predictive model for AD.

7.10.4 DETECTION OF DEMENTIA

In the previous section, we described the effectiveness of cognitive performance measures, especially those that index episodic memory functioning, in predicting who will become demented. In this section, we change our focus slightly in that it is not the development of dementia that we are interested in; rather, the issue of interest is the ability to differentiate between adults who are already mildly demented compared with nondemented adults. For the purpose of the present chapter, we are interested in whether those same measures of cognitive performance that are able to predict development of dementia would be effective at discriminating between nondemented adults and persons who are in an early phase of the disease.

Two main statistical methods have been used to differentiate normal old from AD individuals, namely the algorithm approach (e.g., Fuld, 1984) and the discriminant analysis approach (e.g., Storandt, Botwinick, Danziger, Berg, & Hughes, 1984). In the following, we will limit our review to the more commonly used discriminant analysis method. In this method, researchers have tried to determine the cognitive tasks or tests that best discriminate between normal old and mild AD patients. As is true with research attempting to delineate preclinical markers of AD, the uncertainties regarding dementia diagnoses are problematic for attempts to detect dementia. Thus, it may be that the differentiation between normal old adults and AD patients is based on a less than perfect "answer key," which also means that we do not get a fair evaluation of the discriminative power of the cognitive tests.

With this problem fresh in mind, we now turn to the results from studies investigating how well cognitive measures can discriminate between normal old adults and early AD patients. Most studies have found that tasks assessing episodic memory are particularly sensitive in detecting early dementia (e.g., Almkvist & Bäckman, 1993a; Herlitz et al., 1995; Storandt et al., 1984; Welsh, Butters, Hughes, Mohs, & Heyman, 1991, 1992). For example, in groups of AD patients and community volunteers, matched on age and education, Welsh et al. (1991) found that a measure of episodic memory, delayed word recall, correctly classified 96% of the control subjects and 86% of the patients. However, the authors of this study pointed out that the AD patients were compared with high-functioning control subjects, and that the results cannot be generalized to older adults in the community at large.

A study investigated the relative importance of episodic memory tasks and visuospatial tasks in the detection of dementia (Herlitz et al., 1995). Subjects were sampled from the general population in Stockholm, Sweden. Thus, all subjects were randomly sampled from the same population, and were generally older than subjects in other studies. In addition to standardized neuropsychological tests, the test battery included specialized measures derived from the experimental literature to examine various features of episodic memory. Three-hundred and ninety-eight normal older adults and 63 mildly demented individuals participated in this study. The group of mildly demented adults consisted of patients diagnosed with AD, VaD, mixed dementia (AD and VaD), Parkinson's disease, and alcohol dementia. Separate analysis was conducted for the AD group and the entire dementia group, but it was found that the results were similar irrespective of dementia etiology.

The results of this study demonstrated that it is possible to differentiate those with mild dementia from normal old adults using mainly measures of episodic memory. Of the five measures that contributed to the differentiation of dementia from normal aging, four were indicators of episodic memory and, specifically, measures of recognition and cued recall. Thus, in contrast to the Welsh et al. (1992) study, no free recall task was included in the function, but rather highly supportive measures of episodic memory. As with the research on dementia incidence, it was hypothesized that a possible reason for this result is that many nondemented older adults, especially those in advanced age, already have deficits in free recall (see Craik & Jennings, 1992, for a review), whereas recognition and cued recall may be relatively well preserved. In the Welsh et al. study (1992), subjects were recruited from a clinical population with healthy volunteers from the community. Thus, differences in cognitive ability between the nondemented subjects in Welsh et al. (1992) and those in Herlitz et al. (1995) may have contributed to the differences in results.

Although episodic memory tasks provide a good classification of normal old and mildly demented subjects, a number of subjects are misclassified by these measures. Correct classification rates, ranging from 92%–96% and 67%–89%, have been achieved for normal old and mildly demented subjects, respectively (Almkvist & Bäckman, 1993a; Herlitz et al., 1995; Welsh et al., 1991, 1992). Herlitz et al. (1995) found an overall classification rate of 90.7%, misclassifying 22 normal controls as demented and 21 demented subjects as normals. Further analysis showed that the misclassified normal controls were older and had lower MMSE scores than the correctly classified normal older adults. By contast, the group of mildly demented subjects that were misclassified as normal controls, were younger, had higher MMSE scores, and consisted of more males than the group of correctly classified mildly demented subjects.

In a later study, Herlitz, Small, Fratiglioni, Almkvist, and Bäckman (1997) were able to provide follow-up information regarding dementia incidence of the subjects that were incorrectly classified as demented, by reassessing all subjects three years after the initial examination. It was found that out of the 14 surviving subjects (eight subjects died from time 1 to time 2), 10 had developed AD. This suggests that the misclassification of normal old as demented, by the cognitive tests administered at time 1, was fairly accurate in selecting subjects who would eventually develop dementia. Hence, the majority of the misclassified subjects may not have been misclassified at the baseline assessment, but rather been in an early phase of dementia already at time 1.

The reasons for the classification of subjects, who had a clinical diagnosis of dementia at time 1 as nondemented, were different. At time 2, it was found, among the 15 surviving subjects, that two subjects had their diagnosis changed from demented to nondemented. Thus, these subjects had received, incorrectly, a clinical dementia diagnosis at time 1 that was later corrected. However, the majority of subjects in this group, of course, remained demented, and it is reasonable to assume that the cognitive tests were unable to discriminate between the group of correctly classified dementia patients and this group of subjects who, were younger, had a higher MMSE scores, and consisted of more males. Thus, it appears that the clinical diagnosis of dementia is easier in younger, male, and highly educated subjects than it is in older, female, and less educated persons. These results indicate that cognitive tests, and especially tests of episodic memory, are effective at detecting subjects in a very early stage of dementia who are easily missed in the clinical diagnosis due to confounding variables such as age, education, and gender. Yet, although cognitive tests may be useful in the preclinical detection of dementia, these tests may also miss a sizeable proportion of prevalent dementia cases due to insensitivity to the same demographic indices.

Taken together, the results on cognitive detection of dementia show that episodic memory measures are particularly effective at discriminating between normal old adults and AD patients. In addition, this line of research demonstrates that memory tasks may be effective in the detection of AD at a very early phase—a finding which is consistent with research conducted on the incidence of AD (e.g., Small et al., 1997; Tierney et al., 1996). In particular, results indicate that performance on highly supported tests of episodic memory performance are especially salient predictors of future cases of dementia, as well as good discriminators between normal elderly adults and those in mild stages of the disease process.

7.10.5 STAGING OF DEMENTIA

It is of considerable interest whether the same or different cognitive tasks are particularly useful in the detection and staging of AD, respectively. Similarities or differences with respect to this issue are relevant to the understanding of the progression of AD. On the one hand, if the same cognitive functions would be involved in both detection and staging of dementia, this would mean that decline in these functions would be linear across the pathogenesis and, ultimately, that some cognitive abilities are relatively well preserved throughout the disease. On the other hand, if different functions are involved in the detection and staging of dementia, it would imply that different cognitive abilities may decline at different times in the disease process and at different rates.

In studies investigating staging of AD, the most common method has been to use a global cognitive assessment instrument (e.g., the MMSE) as the criterion variable. There are, of course, problems related to this method, the most important being that the global assessment instrument used consists of a given set of cognitive tasks, which means that how well a certain set of cognitive test measures can stage dementia, by means of discriminant analysis, is dependent on how well they are correlated with the criterion measure. Only a few studies have investigated the ability of cognitive tests to stage dementia, that is, to separate mildly demented subjects from moderately demented subjects (Almkvist & Bäckman, 1993a; Herlitz et al., 1995; Hill, Storandt, & LaBarge, 1992; Welsh et al., 1992). These studies have found that, although episodic memory tasks are critical in detecting dementia, tasks drawing on verbal skills and visuospatial abilities are more effective at staging dementia. For example, Hill et al. (1992) used the clinical dementia rating scale (George et al., 1990) as the criterion variable, and found that tasks assessing verbal naming, visuospatial copying, and episodic recognition separated between mildly and moderately demented AD patients, with an overall correct classification rate of 86.2%. Almkvist and Bäckman (1993a) staged very mildly, mildly, and moderately demented patients, with the MMSE as the criterion variable. This study found that tasks assessing visuospatial ability, semantic memory, psychomotor speed, and episodic memory separated the three groups of patients with an overall correct classification rate of 64%.

Herlitz et al. (1995) also used the MMSE as the criterion variable and found an overall classification rate of 84.3% in the differentiation of mildly and moderately demented subjects. As in the corresponding analysis on detection of dementia, all analyses were done with the patient group consisting of either AD patients or a mixture of dementia subjects. However, such a division made little difference to the outcome. Four test measures were included in the discriminant function, a free recall task, backward digit span, block design (WAIS-R; Wechsler, 1981), and modified trail making A (Armitage, 1946). Thus, in contrast to the tasks involved in the detection of dementia, two of the tasks in the discriminant function were visuospatial in nature and one was a measure of primary memory. Only one of the tasks in the discriminant function was an episodic memory measure. These results suggest that visuospatial and primary memory abilities may deteriorate later and/or at a slower rate than episodic memory functions in the early stages of dementia (see also Almkvist & Bäckman, 1993b).

The results on staging of dementia are less consistent than are the studies on detection, probably because different criterion variables have been used across studies. In addition, comparatively little research has addressed the efficacy of cognitive measures in the staging of dementia, hence making firm conclusions difficult. However, the majority of studies addressing this issue indicate that episodic memory plays a less dominant role than other cognitive abilities in the differentiation between mildly and moderately demented patients. We contend that the most parsimonious interpretation of these results is that cognitive abilities such as visuospatial functioning, semantic memory, and primary memory deteriorate later and/or at a slower rate than episodic memory in the early stages of dementia.

7.10.6 PROGRESSION OF DEMENTIA

In addition to studies that have examined the cognitive performance of mildly to moderately demented adults through attempts to stage dementia severity, research has focused on longitudinal changes in cognitive performance in these patients. Ultimately, information regarding rates of change, and whether there is a differential nature of cognitive change at later stages of dementia severity can only be inferred from studies that have examined data from more than one time of measurement. Studies that have attempted to stage and detect dementia are hampered by the lack of longitudinal evidence when deriving cognitive profiles of disease progression. The studies that we review in this section are based on data from more than one measurement point.

Grady et al. (1988) examined changes in memory, attentional, verbal, and visuospatial functioning in a group of 11 AD patients. They noted that among all of the measures of cognitive performance, impaired episodic memory performance was the first sign of cognitive dysfunction in dementia. Further, these investigators reported that visuospatial functioning and verbal ability were among the last abilities to show deficits in their sample of AD patients. Similar results were reported by Almkvist and Bäckman (1993b). In this case, longitudinal changes in a range of cognitive functions in AD were examined over a two-year period. Almkvist and Bäckman found that episodic memory measures exhibited the least amount of decline over the retest interval, but when the estimated onset of decline was examined, these measures were the first to exhibit decline in AD. Moreover, visuospatial functioning, indexed by block design and object assembly from the WAIS-R, exhibited a moderate rate of decline over the retest interval but were among the last measures to decline in the sample of AD patients.

Additional evidence for verbal and visuospatial functioning declining at a later point in the disease process comes from a study by

Haxby et al. (1990). This study examined changes in cognitive functioning over a two-year follow-up period in a group of 11 mildly demented AD patients. Haxby and colleagues found that the impairment that dominated at the initial time of measurement were localized to tasks measuring memory and ability to sustain attention. However, "neuropsychological impairment in mildly impaired AD patients became more severe and widespread as the disease progressed" (Haxby et al., 1990, p. 757).

Small and Bäckman (1997) also examined longitudinal changes in episodic memory, visuospatial functioning, and verbal fluency in a group of 26 AD and 10 VaD patients. In this study, longitudinal changes were less pronounced in the episodic memory measures, whereas performance in the visuospatial and verbal domains exhibited dramatic negative changes. Further, the changes in performance varied minimally as a function of dementia type, indicating approximately equivalent declines in performance for both the AD and VaD patients.

Taken together, the results from longitudinal studies that have examined changes in cognitive performance of mildly demented persons suggest a differential pattern of change depending on what domain of cognitive functioning is being measured. Specifically, changes in episodic memory functioning of mildly demented adults are less pronounced, probably as a result of the bulk of the dementia-related declines having occurred already. Certainly, this hypothesis is supported by the fact that impairment of episodic memory may be seen in persons who will later develop dementia several years before a clinical diagnosis is possible (e.g., Jacobs et al., 1995; Linn et al., 1995; Tierney et al., 1996). On the other hand, visuospatial and verbal functioning exhibit stronger declines at later stages of the progression of the disease. The results from the longitudinal studies thus converge with those seen in cross-sectional research attempting to stage dementia severity through cognitive performance indices. In particular, measures of visuospatial and verbal abilities may be particularly effective in staging dementia, because these abilities tend to decline at a later stage of the dementing illness than episodic memory skills.

7.10.7 TRAJECTORY OF DECLINE

Although there is some evidence to suggest that not all cognitive abilities decline at the same time, there is additional evidence to suggest that the shape of the decline function may also vary as a function of dementia severity. Several authors have suggested that the cognitive changes accompanying dementia follow a nonlinear pattern. For example, Haxby et al. (1990) examined the cognitive performance of 16 AD patients. They noted that changes in the WAIS, MDRS, and MMSE followed a biphasic pattern of change, that is, there was an initial plateau of performance followed by a sharp linear decline in cognitive functioning.

Additional evidence for nonlinear changes in cognitive performance in dementia comes from research reported by Brooks and his colleagues (Brooks et al., 1993; Brooks & Yesavage, 1995). These investigators proposed a trilinear model of cognitive change in AD, in which the cognitive decline function in AD is conceptualized as a tripartite curve. According to this view, there is an initial period of stability before detectable decline, followed by a period of decline, and ending with a final period of relative stability. Longitudinal evidence from Helmes et al. (1995) demonstrates that performance on their scale of global cognitive functioning, the extended scale for dementia, follows a trilinear pattern of change in dementia patients much like Brooks et al. have postulated.

The possibility of a nonlinear pattern of cognitive change in dementia has several implications for the way in which we study progression of cognitive changes across the disease process. For example, the use of statistical methodologies that assume linear changes in performance (e.g., multiple regression, analysis of variance) may be inappropriate because of the nonlinear changes that are embedded in performance. Perhaps the greatest implication deals with the number of measurement occasions that we choose to follow our longitudinal samples for. Obviously, in order to describe nonlinear patterns of change, at least three measurement occasions are required. Brooks et al. (1993) suggested that in order to adequately describe the trilinear decline function in AD, five measurement occasions should be planned for. In addition, if we wish to examine the trilinear shape of changes in memory and other cognitive abilities, even more measurement points will need to be examined. Recall from our reviews of incidence, detection, staging, and longitudinal changes, that episodic memory functioning may be the first cognitive ability to decline in dementia, whereas other cognitive functions, such as verbal and visuospatial skill, appear to decline at a later stage. The differential onset of changes in individual cognitive measures implies that persons will be at different stages in the trilinear model depending on what domain of functioning is being indexed. For example, some individuals may be experiencing decline in episodic memory functioning, but be at the initial plateau stage of verbal ability because they are relatively early in the course of dementia. By contrast, other persons may be experiencing decline in verbal ability, but be at the later plateau stage in terms of episodic memory functioning. Furthermore, this argument assumes that performance in each domain of functioning follows a similar trilinear pattern of decline, with comparable duration in each part of the function. Because much of the work on non-linear change in dementia has focused on more global indices of cognitive functioning (e.g., Brooks et al., 1993; Helmes et al., 1995), rather than individual domains of performance, it is unclear whether every domain of cognitive function (e.g., verbal ability, visuospatial functioning) will follow a similar trajectory of decline. Clearly, although the results presented here are suggestive of nonlinear changes in the cognitive performance of demented adults, additional evidence is necessary before firm conclusions can be made regarding the onset, shape, and duration of the respective change functions.

7.10.8 PREDICTORS OF DEMENTIA-RELATED DIFFERENCES AND CHANGES

We have argued that the changes accompanying dementia are somewhat predictable. For example, on the one hand, we have marshalled evidence to indicate that declines in episodic memory performance are among the first signs of a dementing illness. Indices of verbal and visuospatial functioning, on the other hand, begin to deteriorate markedly only after the diagnosis of dementia has been made and proceed to decline with increasing severity of dementia. In this section, we look again at the ability to predict dementia-related differences and changes in cognitive performance. However, although our previous attempts to examine the cognitive transitions in dementia have focused on the temporal progression of the disease, we focus here on whether cognitive performance is predictable by different demographic, biological, and psychometric indices.

In research on normal aging, there is a great tradition of using individual-difference variables to predict age-related differences, and age-related changes, in cognitive functioning (see Hultsch & Dixon, 1990; Salthouse, 1991, for reviews). For example, a number of studies have found age-related differences being predicted by demographic characteristics such as age, gender, and education (e.g., Herlitz, Nilsson, & Bäckman, in press; Inouye, Albert, Mohs, Sun, & Berkman, 1993; Wiederholt et al., 1993). Similarly, other studies have focused on the ability of biological indices, such as general health or vitamin status, in predicting cognitive performance (e.g., Hultsch, Hammer, & Small, 1993; Wahlin, Hill, Winblad, & Bäckman, 1996). Finally, investigators have identified a number of basic information processing indices, including measures of processing speed (e.g., Salthouse, 1991) and working memory (e.g., Hultsch, Hertzog, & Dixon, 1990; Salthouse, 1991) as important predictors of cognitive performance among normal elderly adults.

Although the role of these types of individual-difference variables in predicting age-related differences and changes in performance is well documented in the normal aging literature, it is less clear whether these same classes of variables can predict dementia-related differences or dementia-related changes in cognitive performance. Using global performance indices as target variables, a number of factors have been found to be associated with a more rapid rate of progression, including early onset age (e.g., Burns, Jacoby, & Levy, 1991), language deficits (e.g., Boller, Becker, Holland, & Forbes, 1991), visuospatial deficits (e.g., Becker, Huff, Nebes, Holland, & Boller, 1988), and psychomotor slowing (e.g., Berg et al., 1984). However, other studies have found the relationship between these individual-difference predictors and cognitive progression to be equivocal (e.g., Almkvist & Bäckman, 1993b; Mortimer, Ebbit, Jun, & Finch, 1992).

Bäckman, Hill, Herlitz, Fratiglioni, and Winblad (1994) reported that the role of individual-difference variables in predicting performance on episodic memory tasks is substantially reduced in dementia. In this study, Bäckman et al.(1994) examined the relationship between a variety of subject-related variables corresponding to the demographic (e.g., age, education, onset age, activities of daily living), biological (ferritin, albumin, folic acid, fructoseamine/albumin, vitamin B12, and thyroid stimulating hormone), and psychometric (e.g., block design, digit span) domains and performance on two episodic memory factors, recall and recognition. They found that after controlling for differences in disease severity, only two predictors contributed to performance. In one case, a marker of fluid intelligence (block design) was positively related to performance on a factor of word recall, and in the other, age was negatively related to performance on a measure of word recognition. However, the relative paucity of individual-difference variables that could predict dementia-related differences in episodic memory functioning stands in contrast to the variety of such measures that were significant predictors of episodic memory performance in the sample that contained nondemented adults from the same study population (e.g., Bäckman & Wahlin, 1995; Hill, Wahlin, Winblad, & Backman, 1995; Wahlin et al., 1993). In a related study, Hill, Bäckman, Wahlin, and Winblad (1995) demonstrated evidence for a limited number of predictors being related to visuospatial performance in dementia, once severity of dementia was statistically controlled. Specifically, using Poppelreuter's figures, the clock test, and block design as criterion variables, only one relationship out of a total of 15 possible predictor variables was observed, namely that education was positively related to block design.

Using the study of Bäckman et al. (1994) as a guide, we recently investigated whether the relatively paucity of individual-difference predictors also held when changes in the cognitive performance of demented adults were examined (Small & Bäckman, 1997). In this study, using the same constellation of predictor variables, we sought to determine whether longitudinal changes in episodic memory, visuospatial functioning, and verbal fluency could be predicted by these indices. Results indicated that none of the demographic, biological, or psychometric indices of performance were significant predictors of change in any domain of cognitive functioning. We concluded that although the rate of cognitive deterioration in dementia is highly variable, this variability in change appears to generalize across a variety of demographic, psychometric, and biological characteristics. A possible reason for the lack of effects is that the role of individual-difference variables in cognitive functioning in dementia is overshadowed by the pathogenetic process itself. Specifically, much of the systematic variation in performance may be absorbed by the severity of the disease (as indexed by, for example, the MMSE), hence leaving little predictive power to other subject-related factors.

In summary, it appears that although the role of individual-difference variables in predicting cognitive performance of nondemented adults has a long and successful heritage, the importance of these same variables appears to be substantially reduced in dementia when contrasted against the severity of the disease itself.

7.10.9 CONCLUSIONS AND EXTENSIONS

In this chapter, we have seen that as individuals make the transition from normal aging to dementia, as well as the transition from milder to more severe stages of dementia, there is a relatively predictable pattern of cognitive changes. The results from studies that have examined incident cases of dementia (e.g., Small et al., 1997; Tierney et al., 1996) indicate that baseline measures of episodic memory performance are the best predictors of who will become demented. Similarly, results from studies that have attempted to discriminate between nondemented and mildly demented adults on the basis of cognitive performance measures suggest that episodic memory performance is an important factor in making these discriminations (e.g., Herlitz et al., 1995; Welsh et al., 1992). However, the predictive role of verbal ability, visuospatial functioning, and primary memory at these early stages of dementia is considerably smaller. It is only after the disease has progressed that visuospatial functioning, verbal ability, and primary memory begin to deteriorate markedly. This is evidenced in studies that have found these measures to be effective at staging disease progression (Herlitz et al., 1995; Hill et al., 1992; Welsh et al., 1992), as well as in studies that have assessed longitudinal changes directly (Almkvist & Bäckman, 1993b; Grady et al., 1988; Small & Bäckman, 1996).

Although we still lack a complete understanding of why the different domains of cognitive functioning follow such a pattern of deterioration, analyzing the trajectory of cognitive decline in light of the biological changes that accompany AD may provide useful information. Specifically, it may be argued that the differential change in cognitive functioning is consistent with models of the pathogenetic process in AD based on both neuropathological (Arrigada, Growdon, Hedley-White, & Hyman, 1992, Braak & Braak, 1991, 1995) and brain-imaging (Grady et al., 1988, 1990) data. For example, Braak and Braak contend that initial changes in the brain structures of AD patients occur in the medial temporal regions (e.g., the entorhinal cortex and the hippocampus) and, as the disease process advances, these changes radiate into neocortical structures. It is well known from research with amnesic patients that the hippocampal formation plays an important role in episodic remembering (e.g., Mayes, 1988; Squire, 1987). Recent brain activation research using positron emission tomography supports the view that episodic memory is heavily dependent on the integrity of the medial–temporal cortex (e.g., Grady et al., 1995; Nyberg, McIntosh, Houle, Nilaaon, & Tulving, 1996). It is also known that the neocortical association areas are particularly critical to performance across a variety of verbal and visuospatial tasks (e.g., Raichle et al., 1994; Ungerleider & Haxby, 1994). Thus, the pattern of dementia-related cognitive decline across the pathogenesis makes perfect sense in view of independent research addressing the biological basis of human cognition. From a somewhat different perspective, it may further be noted that episodic memory performance has been found to be an especially sensitive cognitive indicator of a variety of conditions, including herpes simplex encephalitis (Williams, Watts, Macleod, & Matthews, 1988), hypoxic damage (Hashtroudi, Parker, DeLisi, & Wyatt, 1983), toxic changes (Teitelbaum et al., 1990), depression (Bäckman & Forsell, 1994), alcohol intoxication (Nilsson, Bäckman, & Karlsson, 1989), and sleep deprivation (Polzella, 1975). Tulving (1993; see also Schacter & Moscovitch, 1984) has speculated that episodic memory evolved relatively late phylogenetically and develops relatively late ontogenetically relative to other forms of memory and cognitive functions. Thus, following Ribot's law (Ribot, 1882), it may perhaps be expected that episodic memory should be especially prone to brain insufficiencies such as dementia.

7.10.10 SUMMARY

In summary, although the evidence we have presented here was gathered from a variety of sources that often differed in terms of sample composition, dementia severity, and selection of the specific cognitive tests that were employed, a relatively predictable pattern of cognitive deterioration in dementia was identified. On the one hand, it appears that episodic memory is the first ability to deteriorate at or even before the dementia diagnosis is rendered. On the other hand, visuospatial and verbal functioning decline somewhat later in the disease process, and this is likely to reflect the fact that additional brain structures are being affected by the disease process. We hope that this chapter, with its insights into the pattern of cognitive deterioration in dementia, may generate some considerations in the clinical setting. For example, knowledge pertaining to which specific cognitive tasks that may be particularly sensitive in detecting very early dementia-related changes (e.g., highly supported episodic memory tasks) should be helpful in improving the accuracy of early clinical diagnoses. Although we currently lack effective pharmacological or behavioral intervention procedures for AD and related dementia disorders, such procedures may be developed in the years to come. If that were to be the case, it is conceivable that treatment efficacy would be greatest at the earliest possible phase in the disease process. In addition, even in the present situation, accurate diagnostic information early in the disease process may be useful for the patient, practitioners, and caregivers alike in order to adjust optimally to the impending disease process.

ACKNOWLEDGMENTS

Preparation of this chapter was supported by grants from the Swedish Council For Social Research, and the Swedish Council For Research in the Humanities and the Social Sciences.

7.10.11 REFERENCES

Almkvist, O. (1994). Neuropsychological deficits in vascular dementia in relation to Alzheimer's disease: Reviewing evidence for functional similarity or divergence. *Dementia, 5,* 203–209.

Almkvist, O., & Bäckman, L. (1993a). Detection and staging of early clinical dementia. *Acta Neurologica Scandinavica, 88,* 10-15.

Almkvist, O., & Bäckman, L. (1993b). Progression in Alzheimer's disease: Sequencing of neuropsychological decline. *International Journal of Geriatric Psychiatry, 8, 755–763.*

Almkvist, O., Bäckman, L., Basun, H., & Wahlund, L.-O. (1993). Patterns of neuropsychological performance in Alzheimer's disease and vascular dementia. *Cortex, 29,* 661–673.

American Psychiatric Association (1980). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed. rev.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: Author.

Armitage, S. G. (1946). An analysis of certain psychological tests used for the evaluation of brain injury. *Psychological Monograph, 60, 1–48.*

Aronson, M. K., Ooi, W. L., Morgenstern, H., Hafner, A., Masur, D., Crystal, H., Frishman, W. H., Fisher, D., & Katzman, R. (1990). Women, myocardial infarction, and dementia in the very old. *Neurology, 40,* 1102–1106.

Arragada, P. V., Growdon, J. H., Hedley-White, E. T., & Hyman, B. T. (1992). Neurofibrillary tangles but not senile plaques parallel duration and severity of Alzheimer's disease. *Neurology, 42,* 631–639.

Bäckman, L. (Ed.) (1992). *Memory functioning in dementia.* Amsterdam: North-Holland.

Bäckman, L., & Forsell, Y. (1994). Episodic memory functioning in a community-based sample of old adults with major depression: Utilization of cognitive support. *Journal of Abnormal Psychology, 103,* 361–370.

Bäckman, L., Hill, R. D., Herlitz, A., Fratiglioni, L., & Winblad, B. (1994). Predicting episodic memory performance in dementia: Is severity all there is? *Psychology and Aging, 9,* 520–527.

Bäckman, L., & Larsson, M. (1992). Recall of organizable words and objects in adulthood: Influences of instructions, retention interval, and retrieval cues. *Journal of Gerontology: Psychological Sciences, 47,* P273–P278.

Bäckman, L., Mäntylä, T., & Herlitz, A. (1990). The optimization of episodic remembering in old age. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 118–163). Cambridge, UK: Cambridge University Press.

Bäckman, L., & Wahlin, Å. (1995). Influences of item organizability and semantic retrieval cues on word recall in very old age. *Aging and Cognition, 2,* 312–325.

Becker, J. T., Huff, J. P., Nebes, R. D., Holland, A., & Boller, F. (1988). Neuropsychological function in Alzheimer's disease: Pattern of impariment and rates of progression. *Archives of Neurology, 45,* 263–268.

Berg, L., Danziger, W. L., Storandt, M., Coben, L. A., Gado, M., Hughes, C. P., Knesevich, J. W., & Botwinick, J. (1984). Predictive features in mild senile dementia of the Alzheimer type. *Neurology, 34,* 563–569.

Blessed, G., Tomlinson, B. E., & Roth, M. (1968). The association between quantitative measures of dementia and of senile changes in the cerebral grey matter of elderly subjects. *British Journal of Psychiatry, 114,* 797–811.

Braak, H., & Braak, E. (1991). Neuropathological staging of Alzheimer-related changes. *Acta Neuropathologica, 82, 239–259.*

Braak, H., & Braak, E. (1995). Staging of Alzheimer's disease-related neurofibrillary changes. *Neurobiology of Aging, 16, 271–284.*

Branconnier, R. J., Cole, J. O., Spera, K. F., & DeVitt, D. R. (1982). Recall and recognition as diagnostic indices of malignant memory loss in senile dementia: A Bayesian analysis. *Experimental Aging Research, 8,* 189–193.

Brayne, C., Gill, C., Huppert, F. A., Barkley, C., Gehlhaar, E., Girling, D. M., O'Connor, D. W., & Paykel, E. S. (1995). Incidence of clinically diagnosed subtypes of dementia in an elderly population: Cambridge Project for Later Life. *British Journal of Psychiatry, 167,* 255–262.

Brody, J. A., & Schneider, E. L. (1986). Disease and disorders of aging: A hypothesis. *Journal of Chronic Disorders, 39,* 871–876.

Brooks, J. O., Kraemer, H. C., Tanke, E. D., & Yesavage, J. A. (1993). The methodology of studying decline in Alzheimer's disease. *Journal of the American Geriatrics Society, 41,* 623–628.

Brooks, J. O., & Yesavage, J. A. (1995). Identification of fast and slow decliners in Alzheimer disease: A different approach. *Alzheimer Disease and Associated Disorders, 9*(Suppl. 1), S19–S25.

Burns, A., Jacoby, R., & Levy, R. (1991). Progression of cognitive impairment in Alzheimer's disease. *Journal of the American Geriatrics Society, 39,* 39–45.

Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London, B302,* 341–359.

Craik, F. I. M., & Jennings, J. M. (1992). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 51–110). Hillsdale, NJ: Erlbaum.

Craik, F. I. M., & McDowd, J. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 474–479.

Ebly, E. M., Parhad, I. M., Hogan, D. B., & Fung, T. S. (1994). Prevalence and types of dementia in the very old: Results from the Canadian Study of Health and Aging. *Neurology, 44,* 1593–1600.

Elias, M. F., Wolf, P. A., D́'Agostino, R. B., Cobb, J., & White, L. R. (1993). Untreated blood pressure level is inversely related to cognitive functioning: The Framingham Study. *American Journal of Epidemiology, 138, 353–364.*

Erkinjuntti, T., Laaksonen, R., Sulkava, R., Syrjäläinen, R., & Palo, J. (1986). Neuropsychological differentiation between normal aging, Alzheimer's disease, and vascular dementia. *Acta Neurologica Scandinavica, 74,* 393–403.

Fewster, P. H., Griffin-Brooks, S., MacGregor, J., Ojalvo-Rose, E., & Ball, M. J. (1991). A topographical pathway by which histopathological lesions disseminate through the brain of patients with Alzheimer's disease. *Dementia, 2,* 121–132.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state:" A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198

Frackowiak, R. S. J., Pozzilli, C., Legg, N. J., Du Boulay, G. H., Marshall, J., Lenzi, G. L., & Jones, T. (1981). Regional cerebral oxygen supply and utilization in dementia: A clinical and physiological study with oxygen-15 and positron tomography. *Brain, 104,* 753–778.

Fratiglioni, L., Grut, M., Forsell, Y., Viitanen, M., & Winblad, B. (1992). Clinical diagnosis of Alzheimer's disease and other dementias in a population survey: Agreement and causes of disagreement in applying DSM-III-R criteria. *Archives of Neurology, 49,* 927–932.

Fries, J. F., & Crapo, L. M. (1981). *Vitality and aging: Implications of the rectangular curve.* New York: Freeman.

Fuld, P. A. (1984). Test profile of cholinergic dysfunction and of Alzheimer-type dementia. *Journal of Clinical Neuropsychology, 6,* 380–392.

Fuld, P. A., Masur, D. M., Blau, A. D., Crystal, H., & Aronson, M. K. (1990). Object-memory evaluation for prospective detection of dementia in normal functioning elderly: Predictive and normative data. *Journal of Clinical and Experimental Neuropsychology, 12,* 520–528.

Galasko, D., Hansen, L. A., Katzman, R., Wiederholt, W., Masliah, E., Terry, R., Hill, L. R., Lessin, P., & Thal, L. J. (1994). Clinical-neuropathological correlations in Alzheimer's disease and related dementias. *Archives of Neurology, 51,* 888–895.

George, A. E., deLeon, M. J., Stylopoulos, L. A., Miller, J., Kluger, A., Smith, G., & Miller, D. C. (1990). CT diagnostic features of Alzheimer disease: Importance of the choroidal/hippocampal fissure complex. *American Journal of Neuroradiology, 11,* 101–107.

Gottfries, J., Blennow, K., Wallin, A., & Gottfries, C. G. (1995). Diagnosis of dementias using partial least squares discriminant analysis. *Dementia, 6,* 83–88.

Grady, C. L., Haxby, J. V., Horwitz, B., Sundaram, M., Berg, G., Schapiro, M., Friedland, R. P., & Rapoport, S. I. (1988). Longitudinal study of the early neuropsychological and cerebral metabolic changes in dementia of the Alzheimer type. *Journal of Clinical and Experimental Neuropsychology, 10,* 576–596.

Grady, C. L., Haxby, J. V., Schapiro, M. B., Gonzales-Aviles, A., Kumar, A., Ball, M. J., Heston, L., & Rapaport, S. I. (1990). Subgroups in dementia of the Alzheimer type identified using positron emission tomography. *Journal of Neuropsychiatry and Clinical Neurosciences, 2,* 373–384.

Grady, C. L., McIntosh, A. R., Horwitz, B., Maisog, J. M., Ungerleider, L. G., Mentis, M. J., Pietrini, P., Schapiro, M. B., & Haxby, J. V. (1995). Age-related reductions in human recognition memory due to impaired encoding. *Science, 269,* 218–221.

Hashtroudi, S., Parker, E. S., DeLisi, L. E., & Wyatt, R. J. (1983). On elaboration and alcohol. *Journal of Verbal Learning and Verbal Behavior, 22,* 164–173.

Haxby, J. V., Grady, C. L., Koss, E., Horwitz, B., Heston, L., Schapiro, M., Friedland, R. P., & Rapoport, S. I. (1990). Longitudinal study of cerebral metabolic asymmetries and associated neuropsychological patterns in early dementia of the Alzheimer type. *Archives of Neurology, 47,* 753–760.

Helmes, E., Merskey, H., Fox, H., Fry, R. N., Bowler, J. V., & Hachinski, V. C. (1995). Patterns of deterioration in senile dementia of the Alzheimer type. *Archives of Neurology, 52,* 306–310.

Herlitz, A., Hill, R. D., Fratiglioni, L., & Bäckman, L. (1995). Episodic memory and visuospatial ability in detecting and staging dementia in a community-based sample of very old adults. *Journal of Gerontology: Medical Sciences, 50a,* M107–M113.

Herlitz, A., Lipinska, B., & Bäckman, L. (1992). Utilization of cognitive support for episodic remembering in Alzheimer's disease. In L. Bäckman (Ed.), *Memory functioning in dementia* (pp. 73–96). Amsterdam: North-Holland.

Herlitz, A., Nilsson, L.-G., & Bäckman, L. (in press). Gender differences in episodic memory. *Memory and Cognition*.

Herlitz, A., Small, B. J., Fratiglioni, L., Almkvist, O., & Bäckman, L. (1997). Detection of mild dementia in community surveys: Is it possible to increase the accuracy of our diagnostic instruments? *Archives of Neurology, 54,* 319–324

Hill, R. D., Bäckman, L., Wahlin, Å., & Winblad, B. (1995). Visuospatial performance in very old demented persons: An individual difference analysis. *Dementia, 6,* 49–54.

Hill, R. D., Storandt, M., & LaBarge, E. (1992). Psychometric discrimination of moderate senile dementia of the Alzheimer type. *Archives of Neurology, 49,* 377–380.

Hill, R. D., Wahlin, Å., Winblad, B., & Bäckman, L. (1995). The role of demographic and life style variables in utilizing cognitive support for episodic remembering among very old adults. *Journal of Gerontology: Psychological Sciences, 50b,* P219–P227.

Hultsch, D. F., & Dixon, R. A. (1990). Learning and memory in aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 258–274). San Diego, CA: Academic Press.

Hultsch, D. F., Hammer, M., & Small, B. J. (1993). Age differences in cognitive performance in later life: Relationships to self-reported health and activity life style. *Journal of Gerontology: Psychological Sciences, 48,* P1–P11.

Hultsch, D. F., Hertzog, C., & Dixon, R. A. (1990). Ability correlates of memory performance in adulthood and aging. *Psychology and Aging, 5,* 356–368.

Huppert, F. A., Brayne, C., & O'Connor, D. W. (1994). *Dementia and normal aging*. Cambridge, UK: Cambridge University Press.

Inouye, S. K., Albert, M. S., Mohs, R., Sun, K., & Berkman, L. F. (1993). Cognitive performance in a high-functioning community-dwelling elderly population. *Journal of Gerontology: Medical Sciences, 48,* 146–151.

Jacobs, D. M., Sano, M., Dooneief, G., Marder, K., Bell, K. L., & Stern, Y. (1995). Neuropsychological detection and characterization of preclinical Alzheimer's disease. *Neurology, 45,* 957–962.

Katzman, R., Aronson, M., Fuld, P., Kawas, C., Brown, T., Morgenstern, H., Frishman, W., Gidez, L., Eder, H., & Ooi, W. L. (1989). Development of dementing illnesses in an 80-year-old volunteer cohort. *Annals of Neurology, 25,* 317–324.

Linn, R. T., Wolf, P. A., Bachman, D. L., Knoefel, J. E., Cobb, J. L., Belanger, A. J., Kaplan, E. F., & D'Agostino, R. B. (1995). The 'preclinical phase' of probable Alzheimer's disease: A 13-year prospective study of the Framingham cohort. *Archives of Neurology, 52,* 485–490.

Masur, D. M., Fuld, P. A., Blau, A. D., Crystal, H., & Aronson, M. K. (1990). Predicting development of dementia in the elderly with the Selective Reminding test. *Journal of Clinical and Experimental Neuropsychology, 12,* 529–538.

Masur, D. M., Sliwinski, M., Lipton, R. B., Blau, A. D., & Crystal, H. A. (1994). Neuropsychological prediction of dementia and the absence of dementia in healthy elderly persons. *Neurology, 44,* 1427–1432.

Mattis, S. (1976). Mental status examination for organic mental syndromes in the elderly patient. In L. Bellack & T. B. Kanasu (Eds.), *Geriatric psychiatry*. New York: Grune & Strattan.

Mayes, A. R. (1988). *Human organic memory disorders*. New York: Cambridge University Press.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of the Department of Health and Human Service Task Force on Alzheimer's Disease. *Neurology, 34,* 939–944.

Morris, R. G. (Ed.) (1996). *The neuropsychology of Alzheimer's disease*. Oxford: Oxford University Press.

Mortimer, J. A., Ebbit, B., Jun, S. P., & Finch, M. D. (1992). Predictors of cognitive and functional progression in patients with probable Alzheimer's disease. *Neurology, 42,* 1689–1696.

Nebes, R. D. (1992). Cognitive dysfunction in Alzheimer's disease. In F. I. M. Craik & T. A. Salthouse (Eds.), *The Handbook of aging and cognition* (pp. 373–446). Hillsdale, NJ: Erlbaum.

Nilsson, L.-G., Bäckman, L., & Karlsson, T. (1989). Priming and cued recall in elderly, alcohol intoxicated and sleep deprived subjects: A case of functionally similar memory deficits. *Psychological Medicine, 19,* 423–433.

Nyberg, L., McIntosh, A. R., Houle, S., Nilsson, L. G., & Tulving, E. (1996). Activation of medial temporal structures during episodic memory retrieval. *Nature, 380,* 715–717.

Polzella, D. J. (1975). Effects of sleep deprivation on short-term recognition memory. *Journal of Experimental Psychology: Human Learning and Memory, 104,* 194–200.

Raichle, M. E., Fiez, J. A., Videen, T. O., MacLeod, A. M. K., Pardo, J. V., Fox, P. T., & Petersen, S. E. (1994). Practice-related changes in human brain functional anatomy during non-motor learning. *Cerebral Cortex, 4,* 8–26.

Ribot, T. (1882). *Diseases of memory*. New York: Appleton.

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Lawrence Erlbaum.

Schacter, D. L., & Moscovitch, M. (1984). Infants, amnesics, and dissociable memory systems. In M. Moscovitch (Ed.), *Infant memory* (pp. 173–216). New York: Plenum.

Schonfield, D., & Robertson, B. A. (1966). Memory storage and aging. *Canadian Journal of Psychology, 20,* 228–236.

Small, B. J., & Bäckman, L. (1997). *Predictors of longitudinal changes in memory, visuospatial, and verbal functioning in very old demented adults*. Manuscript submitted for publication.

Small, B. J., Herlitz, A., Fratiglioni, L., Almkvist, O., & Bäckman, L. (1997). Cognitive predictors of incident Alzheimer's disease: A prospective longitudinal study. *Neuropsychology, 11,* 413–420.

Squire, L. R. (1987). *Memory and brain*. New York: Oxford University Press.

Storandt, M., Botwinick, J., Danziger, W. L., Berg, L., & Hughes, C. P. (1984). Psychometric differentiation of mild senile dementia of the Alzheimer type. *Archives of Neurology, 41,* 497–499.

Storandt, M., & Hill, R. D. (1989). Very mild senile dementia of the Alzheimer type: II. Psychometric test performance. *Archives of Neurology, 46,* 383–386.

Teitelbaum, J. S., Zatorre, R. J., Carpenter, S., Gendron, D., Evans, A. C., Gjedde, A., & Cashman, N. R. (1990). Neurologic sequelae of domoic acid intoxication due to the ingestion of contaminated mussels. *New England Journal of Medicine, 322,* 1781–1787.

Tierney, M. C., Szalai, J. P., Snow, W. G., Fisher, R. H., Nores, A., Nadon, G., Dunn, E., & St. George-Hyslop, P. H. (1996). Prediction of probable Alzheimer's disease in memory-impaired patients: A prospective longitudinal study. *Neurology, 46,* 661–665.

Tulving, E. (1993). Human memory. In P. Anderson, O. Hvalby, O. Paulsen, & B. Hökfelt (Eds.), *Memory concepts: Basic and clinical aspects* (pp. 27–46). Amsterdam: Elsevier.

Ungerleider, L. G., & Haxby, J. V. (1994). 'What' and 'where' in the human brain. *Current Opinions in Neurobiology, 4,* 157–165.

Wahlin, Å., Bäckman, L., Mäntylä, T., Herlitz, A., Viitanen, M., & Winblad, B. (1993). Prior knowledge and face recognition in a community-based sample of healthy, very old adults. *Journal of Gerontology: Psychological Sciences, 48,* 54–61.

Wahlin, Å., Hill, R. D., Winblad, B., & Bäckman, L. (1996). Effects of serum vitamin B12 and folate status on episodic memory performance in very old age: A population-based study. *Psychology and Aging, 11,* 487–496

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale Revised: Manual.* New York: The Psychological Corporation.

Welsh, K., Butters, N., Hughes, J., Mohs, R., & Heyman, A. (1991). Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. *Archives of Neurology, 48,* 278–81.

Welsh, K., Butters, N., Hughes, J., Mohs, R., & Heyman, A. (1992). Detection and staging of dementia in Alzheimer's disease: Use of the neuropsychological measures developed for the Consortium to Establish a Registry for Alzheimer's disease. *Archives of Neurology, 49,* 448–452.

West, R. L., Crook, T. H., & Barron, K. L. (1992). Everyday memory performance across the life span: Effects of age and noncognitive individual differences. *Psychology and Aging, 7,* 72–82.

Wiederholt, W. C., Cahn, D., Butters, N. M., Salmon, D. P., Krtiz-Silverstein, D., & Barrett-Connor, E. (1993). Effects of age, gender and education on selected neuropsychological tests in an elderly community cohort. *Journal of the American Geriatrics Society, 41,* 639–647.

Williams, J. M. G., Watts, F. N., MacLeod, C., & Matthews, A. (1988). *Cognitive psychology and emotional disorders.* New York: Wiley.

Yoshitake, T., Kiyohara, Y., Kato, I., Ohmura, T., Iwamoto, H., Nakayama, K., Ohmori, S., Nomiyama, K., Kawano, H., Ueda, K., Sueishi, K., Tsuneyoshi, M., & Fujishima, M. (1995). Incidence and risk factors of vascular dementia and Alzheimer's disease in a defined elderly Japanese population: The Hisayama Study. *Neurology, 45,* 1161–1168.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.11
Suicidal Behavior in Late Life

JANE L. PEARSON
National Institute of Mental Health, Rockville, MD, USA

7.11.1 INTRODUCTION 248
7.11.2 SOURCES OF DATA FOR STUDIES OF SUICIDAL BEHAVIOR IN THE ELDERLY 248
  7.11.2.1 Mortality Data 248
  7.11.2.2 Community Studies Using the Psychological Autopsy Method 249
  7.11.2.3 Clinical Follow-up Studies 249
  7.11.2.4 Suicide Attempter Studies 250
7.11.3 EPIDEMIOLOGY OF SUICIDAL BEHAVIOR IN LATE LIFE 250
  7.11.3.1 Suicide Completions 250
    7.11.3.1.1 Age, gender, and race 250
    7.11.3.1.2 Marital status 251
    7.11.3.1.3 Methods of suicide 251
  7.11.3.2 Suicide Attempts among the Elderly 252
7.11.4 CORRELATES OF LATE LIFE SUICIDE COMPLETION 253
  7.11.4.1 Psychopathology 253
  7.11.4.2 Previous Suicide Attempts 253
  7.11.4.3 Physical Illness 253
  7.11.4.4 Personality Dimensions 254
  7.11.4.5 Life Events 254
    7.11.4.5.1 Physical illness and loss 254
    7.11.4.5.2 Living alone 254
  7.11.4.6 Neurobiological Correlates 255
    7.11.4.6.1 Serotonergic system 255
    7.11.4.6.2 Serum cholesterol 255
  7.11.4.7 Health Service Use Patterns 255
7.11.5 STUDIES OF ELDERLY SUICIDE ATTEMPTERS 256
7.11.6 SUBGROUPS OF ELDERLY AT RISK FOR SUICIDE 256
  7.11.6.1 Intolerance for Physical Illness or Dependency 256
  7.11.6.2 Older Alcoholics 257
  7.11.6.3 Immigrants and Ethnic Minority Elderly 257
  7.11.6.4 Older Women 257
  7.11.6.5 Nursing Home Patients 258
7.11.7 BIOETHICAL ISSUES 258
  7.11.7.1 Physician Assisted Suicide 258
  7.11.7.2 Competency and Depression 259
  7.11.7.3 Bioethical Research Needs 259
7.11.8 CLINICAL RESOURCES FOR TREATING SUICIDAL OLDER ADULTS 260
  7.11.8.1 Psychotherapeutic Resources 260
  7.11.8.2 Pharmacologic and Electroconvulsive Therapies 261
    7.11.8.2.1 Pharmacologic therapies 261
    7.11.8.2.2 Electroconvulsive therapy 261
7.11.9 SUMMARY 261
7.11.10 REFERENCES 262

7.11.1 INTRODUCTION

The research and clinical treatment of suicidal persons among any age group presents enormous challenges. One aspect of those challenges is to consider the influence of how societies view and tolerate suicidal behavior among their members. Stillion and McDowell (1996) have provided an historical overview of how suicide has been viewed in the Western world. Views on suicidal acts have ranged from Socrates' description of "honorable exits" in 399 BC to the Council of Toledo's declaration that instituted excommunication for the deceased suicide victim or attempter in AD 693. The recent US public health perspective (*Healthy People 2000,* 1991), which began in the 1960s (Shneidman, Farberow, & Litman, 1961), views suicide as an outcome of intolerable psychological distress and inadequately treated mental disorders. This more recent perspective, however, is not unanimously held in the USA. The Hemlock Society (Humphry, 1991, 1992) and a recent ruling by the 9th US District Circuit Court of Appeals (Compassion in Dying vs. State of Washington, 1996) consider suicide a right of individuals. Moreover, many persons hold the view that some persons have more of a right to this act than others; suicide in youth is still considered a tragedy relative to suicides by older, disabled, and less healthy persons (Marks, 1988–1989).

This perception, along with the higher percentage of media coverage devoted to suicides among young and middle-aged persons relative to the elderly could be interpreted as higher rates among youthful groups. However, in 1993 the suicide rate in the USA for persons age 85 and older was twice the rate than persons age 15–19 (10.9 per 100 000 compared to 22.8 per 100 000). If we consider the rates for white males, the highest risk group in the USA, the ratio is more than tripled: adolescents age 15–19 have a rate of 18.5 per 100 000, and older adults age 85+ have a rate of 73.6 per 100 000. This latter rate is over six times the rate of all age-adjusted suicides (National Center for Health Statistics, 1996).

In addition to agist views that consider an older life less valued than a young life, psychological research and clinical interest in elderly suicide may have been limited due to the fact that psychologists (as well as other mental health professionals) rarely see older patients who are suicidal. Generally speaking, older persons with mental disorders rarely seek help from mental health professionals, preferring to visit their primary care physician instead (German, Shapiro, & Skinner, 1985). In addition, psychologists may be less likely to see suicidal patients than psychiatrists: psychologists have a one-in-five chance, and psychiatrists a one-in-two chance of losing a patient to suicide during their careers (Chemtob, Hamada, Bauer, Kinney, & Torigoe, 1988; Chemtob, Hamada, Bauer, Torigoe, & Kinney, 1988). However, as depression is better recognized and effective treatments for depression are increasingly applied to the elderly population (NIH Consensus Development Panel, 1992), as more older persons receive mental health treatment, and as more psychologists become specialists in treating elderly patients (Niederehe, Cooley, & Teri, 1995), the probability that psychologists will encounter suicidal older adults will inevitably increase.

It is hoped that this review will highlight the incompleteness of our current understanding of late life suicide, and make apparent the need for more psychological research on this tragedy. In particular, the epidemiology and putative "risk factors" for completed suicides, and suicide completers' interface with the health system, are surveyed. Also reviewed are studies on elderly attempters, and research on the bioethical issues involved in assisted suicides. Suggestions for further research, including the consideration of possible subgroups at risk, and recommended clinical resources, are also discussed.

7.11.2 SOURCES OF DATA FOR STUDIES OF SUICIDAL BEHAVIOR IN THE ELDERLY

7.11.2.1 Mortality Data

The US national data on gender, age, race, marital status, and method of death characteristics of completed suicides are based on vital statistics information gathered by the Center for Disease Control and Prevention's National Center for Health Statistics. The source of this information is derived from death certificates that come from each state, each state being responsible for its own vital registry system. In addition to the variation by state policies, the qualifications for the coroner, medical examiner, or others responsible for completing death certificates also varies widely. The person may be an elected or appointed layman, lawyer, sheriff, mortician, general practitioner, or forensic pathologist. It is generally agreed that what systematic error exists, it will be toward conservative estimates: there are more "missed" suicide deaths than deaths of other causes miscounted as suicides (see O'Carroll, 1989 for a review). Clark (1992) has suggested that these undercounts are probably around 5%. Despite these sources of variation, mortality data has been used in epidemiologic methods to measure the magnitude of suicide as a public health problem, and assess the effects of public health policies (Healthy People 2000, 1991).

7.11.2.2 Community Studies Using the Psychological Autopsy Method

Although determining the cause of death in suspected suicides usually requires forensic work, such as in-depth interviews with the victims' associates, these efforts are not routinely recorded and available for research use. Therefore, although information such as recent life events, history of hospitalizations, or mental state of the person prior to the suicide may have been assessed, it is not systematically recorded at a state or national level. The psychological autopsy (PA) method was developed to integrate such information in order to achieve a better understanding of the psychological circumstances contributing to a death, and has been applied to a number of communities and populations. Clark and Horton-Deutsch (1992) describe the evolution of this method since the 1930s. As used today, the PA constructs a detailed picture of the victim's psychological state prior to death, including psychiatric symptomatology, behavior, and life circumstances in the weeks or months before death. The gathering of this information includes interviewing knowledgeable informants, reviewing available clinical records, and comprehensive case formulation by one or more mental health professional with expertise in postmortem studies (Beskow, Runeson & Asgard, 1990; Clark & Horton-Deutch, 1992). Clark and Horton-Deutch outline 12 standards for using this method in suicide research.

The application of the PA to suicides among the elderly has been slow in coming, due to the belief that older suicide victims are socially isolated from both formal and informal contacts and therefore there would be few knowledgeable informants to obtain data from. To the contrary, Younger, Clark, Oehmig-Lindroth, and Stein, (1990) described ample numbers of informants available: In Cook County, Illinois, 70% of elderly suicides were married; 90% had one knowledgeable informant and 50% had two informants. Several studies have further demonstrated that the PA method can be applied to elderly suicides: Conwell and Brent (1995) have reviewed a number of PA studies that compare younger and older adults. It is important to note that few psychological autopsy studies in general, and none that are focused on elderly suicides, have yet included appropriate case-controls that would maximize the benefits of this methodology.

7.11.2.3 Clinical Follow-up Studies

As will be discussed below, the majority of suicide victims have some sort of mental disorder. Among the elderly, first episode, unipolar depression is the predominating disorder among completed suicides (Conwell, 1994). A relatively efficient approach to obtaining data on suicides is to follow-up persons most at risk; those with depressive disorders. This has been most frequently conducted through retrospective cohort studies. In addition to the challenge of tracking of persons after treatment, this method also relies on medical records that are notoriously variable in accuracy and completeness. Although this approach has been used to examine suicides among young adult and middle-aged persons, few studies have examined what characteristics, in addition to clinical treatment for depression, place the elderly at risk for suicide. Of the studies that have considered "elderly" persons (those 50 and older), some have suggested that delusional depression is a risk factor, where other have not found this type of depression to be related to increased risk for later suicide (Coryell & Tsuang, 1982; Robinson & Spiker, 1985; Roose, Glassman, Walsh, Woodring, & Vital-Herre, 1983). When specific age groups such as the elderly are examined in these retrospective studies, limited sample sizes and variable follow-up periods often limit both the power to reliably discern risk rates and generalizability to other clinic populations.

Since the elderly are less likely to seek or obtain mental health services:(German et al., 1985), the biases in this approach are likely to favor the documentation of those more likely to see mental health services: women (Kessler, Brown & Broman, 1981; Nathanson, 1977), those with more severe depression, comorbid personality disorder, and persons who have used an overdose method (Lindesay, 1986; Vieta, Nieto, Gastó, & Cirera, 1992). Although it is important to study these biased samples in order to understand what are safe and effective treatments in such service contexts, they minimally overlap with community samples described by epidemiologic and PA methods: Characteristics of the later life suicides in epidemiologic populations and community samples include men, those with late onset, unipolar depression, persons unlikely to have received mental health care, and those who have used a firearm as a suicide method.

7.11.2.4 Suicide Attempter Studies

The PA approach, although providing a wealth of information, is most vulnerable to the retrospective reports of informants who may be biased by knowledge of the type of death of the subject of the inquiry. The PA method is also limited with regard to providing information on the subject's private self-perceptions, coping, or cognitive distortions. Such data can be obtained by interviewing persons who have survived a suicide attempt. However, attempters and completers have been described as distinct but overlapping populations (Linehan, 1986). Among younger groups, the ratio of attempts to completions is high, indicating that many more persons attempt than complete suicide, with variability in the intention for completion (Meehan, Lamb, Saltzman, & O'Carroll, 1992). In addition, women outnumber men in suicide attempts, but men exceed women in suicide completions in the general population both in the USA and the UK (Barraclough, Bunch, Nelson, & Sainsbury, 1974; Mościcki et al.,1988).

Among the elderly, however, the ratio of attempts to completions is estimated to be around 4:1 (Stenback, 1980). This has been interpreted to suggest that the elderly who attempt are more intent on completing suicide, and therefore may be more similar to completers than younger populations (Frierson, 1991; Merrill & Owens, 1990). In addition, previous studies of elderly suicide attempters find higher rates of later suicide completions relative to younger groups (Gardner, Bahn, & Mack, 1964; Kreitman, 1976a; Mościcki, 1995; O'Neal, Robins, & Schmidt, 1956). Because elderly attempters may be very similar to completers, and because studying attempters may offer valuable insight into the risk factors for late life suicide, this method would seem to be among the more promising methods to pursue to increase our knowledge about risk factors for completed suicide. In addition, elderly suicide attempters are an important group to study due to the significant morbidity and mortality that follows an attempt.

7.11.3 EPIDEMIOLOGY OF SUICIDAL BEHAVIOR IN LATE LIFE

7.11.3.1 Suicide Completions

7.11.3.1.1 Age, gender, and race

In the USA, like most other industrialized nations, the most consistent demographic factors associated with risk for suicide are older age and male gender (McIntosh, Santos, Hubbard, & Overholser, 1994; Mościcki, 1995; Pearson & Conwell, 1995). In the USA, older white males have by far the highest rate, followed by African-American men of all adult years. White females have the third highest rate among older persons, and African-American women having the lowest rate. Figure 1 illustrates these differences in rates.

A US Center for Disease Control and Prevention report declared that elderly suicides rates were on the increase (Suicide Among Older Persons, 1996). Fluctuations in suicide rates, such as these, have been the subject of cohort and period effect studies since the mid 1970s. Manton, Blazer and Woodbury (1987) examined age, period, and cohort effects separately on middle age and later life suicides; they concluded that when age and period effects were controlled, cohorts born before World War I and persons born after World War II (baby boomers) were at most risk relative to cohorts born between the World Wars. Murphy and Wetzel (1980), examining US rates from 1949 to 1974 by birth cohort, concluded that with more recent cohorts, rates were rising, regardless of race or gender, and raised an alarm that suicide rates would continue to increase in the USA unless major public health efforts were employed. Furthermore, given that the absolute numbers of the elderly will be increasing as the baby boomers age, the absolute numbers of suicides have been predicted to increase (Conwell, 1994).

However, Newman and Dyck (1988) proposed that the many national trends for "cohort effects" of increasing suicide rates could also be explained by period effects that may or may not be long-lasting. More recently, Fombonne (1994), in considering the evidence for increasing rates of depression and suicide, raised several issues that make it difficult to clearly conclude when changes in rates reflect cohort, period or age effects. These include the interaction among cohort, age and period, and the instability of effects due to selected age groupings. In studies of suicide rates, "the elderly" are often those considered age 65 and older (e.g., Kaplan, Adamek, & Johnson, 1994; McCall, 1991), but within this large age span, some age groups have shown more fluctuation or steeper rises in rates. White males age 65–69 had a suicide rate in 1980 of 30 per 100 000, and a 1993 rate of 27.8 per 100 000, which is a very slight decrease. In contrast, white males age 85 and older had a rate of 52.8 in 1980 but a rate of 73.6 in 1993, a very significant increase (National Center for Health Statistics, 1996). McCall and Land (1994) attempted to account for the increases in older suicides with social structural factors such as social security, cohort size, and widowhood. Although there was some evidence

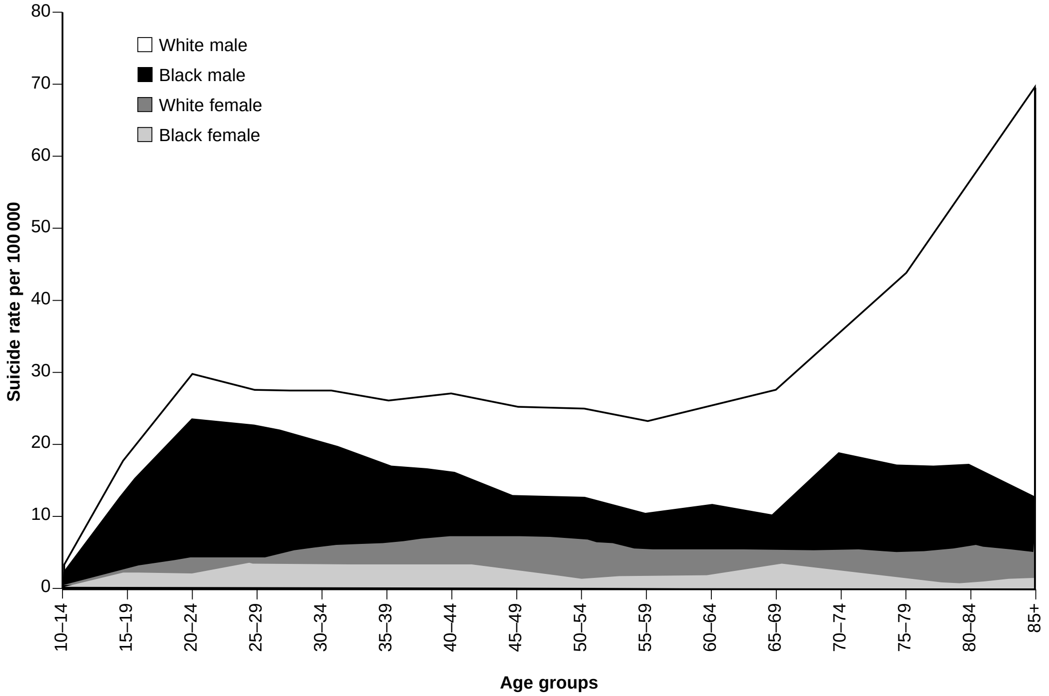

Figure 1 US suicide rates by age, gender, and racial groups, 1995. Source: National Institute of Mental Health, 1997. Data are from the Center for Disease Control and Prevention, National Center for Health Statistics.

for these factors being related to elderly suicides among the "young old" (ages 65–74), they found that none of the factors were related to suicides among the oldest old (ages 75 and older).

7.11.3.1.2 Marital status

In 1992, the rate for divorced and widowed men age 65 and older was 76.4 per 100 000. This was 2.7 times the rate for married men, and 1.4 times the rate for never-married men of this large age group (Suicide Among Older Persons, 1996). In a study comparing older widowed men and women, Li (1995) also reported that the risk of suicide for widowed men was three times greater than for widowed women over a 12 year period in a Maryland county sample. Buda and Tsuang (1990), using 1985 national statistics, reported age differences in specific rates for widowed and divorced males. Up to age 54, males who were widowed had higher rates than divorced males, but beginning with the age 55 group and higher, males who were divorced had higher rates of suicide compared to widowed males. Cohler and Jenuwine (1995) have considered the effects of age and marital status from a life course perspective. As an "off time" event, widowhood is less expected in earlier stages of life and therefore may be more stressful. This may be especially the case for men, where higher mortality rates occur during widowhood for younger men compared to younger women (Kaprio, Koskenvuo, & Rita, 1987), including suicide.

The higher rates of divorce among older men who suicided may reflect long-term family conflict or lack of supportive family interactions which increase risk for suicide. Indeed, family conflicts have been reported to play a role in elderly suicide attempts (Draper, 1994; Schmid, Manjee, & Shah, 1994; Zweig & Hinrichsen, 1993). These studies, along with the patterns of marital status and completed late life suicide rates highlight the need for more focused study on family process factors that may affect older persons' risk for suicidal behavior.

7.11.3.1.3 Methods of suicide

Death by firearms is the most common suicide method in the USA, and the elderly are more likely than any other age group to commit suicide with firearms (Kachur, Potter, James, & Powell, 1995; Kaplan et al., 1994; Meehan, Saltzman, & Sattin, 1991). Figure 2 illustrates the rate of suicides by method and by five-year age groups for 1992. Elderly males (those 65 and older) showed the most proportional increase in the use of firearms as a method of suicide from 1979 to 1992 (Kaplan et al., 1994; Meehan, et al., 1991; Suicide Among Older Persons, 1996).

Kaplan, Adamek, and Geling (1996) examined sociodemographic correlates of firearm use in white elderly male suicides from 1989 to 1991. Consistent with gun ownership patterns (Davis & Smith, 1989–1991), they found that nonmetropolitan residents and men with less than high school education were more likely to use firearms. With regard to marital status, single, never married, males were somewhat less likely than married, widowed, or divorced men to use firearms as a method.

Because efforts to reduce availability of methods of suicide have been reported to result in fewer deaths by suicide (Kreitman, 1976b; Nowers & Irish, 1988), it would seem logical that limiting access to firearms would reduce the rate of elderly suicide (Kaplan et al., 1996). For example, Dennis and Lindesay (1995) describe how a combination of limiting suicide methods on several fronts may have served to reduce elderly suicide rates in the UK. The detoxification of coal gas in the early 1960s was linked to a pronounced and prolonged drop in elderly suicides rates (Kreitman, 1976b), and restriction of prescribed barbiturates has also been associated with fewer deaths by self-poisoning in the elderly (Nowers & Irish). The introduction of exhaust emission laws led to a reduction of suicides by this method in the USA and abroad (Clarke & Lester, 1987).

Studies on the effects of reduced access to handguns and firearms have focused mainly on younger suicides. The degree to which such limitations have had significant effects remains controversial (Rich, Young, Fowler, Wagner, & Black, 1990), where rates of other methods (e.g., hanging) have been found to rise when guns (primarily handguns) are restricted. Sloan and his associates compared suicide rates between King County (Seattle) with the city of Vancouver BC, where the former has fewer restrictions on handgun ownership (Sloan, Rivara, Reay, Ferris, & Kellermann, 1990). Although persons aged 15 to 24 had higher rates of suicide by handguns in King County, the rate of death by firearms among middle-aged and elderly persons (age 45+) was virtually the same in these two communities. There has been no test of the limitations of firearms and its association with decreased suicides by firearms specifically among the elderly.

7.11.3.2 Suicide Attempts among the Elderly

Few epidemiologic studies have focused on the rates of suicide attempts in the elderly. Using data from the National Institute of

Figure 2 Percentage of suicides by age and method for 1992. From Kachur et al. (1995). Source: National Center for Health Statistics.

Mental Health Epidemiologic Catchment Area study (ECA; Regier et al., 1984), Mościcki et al. (1988) found a much lower prevalence of lifetime suicide attempts than younger populations. For persons age 65 and older, the lifetime prevalence for suicide attempt was 1.1%. In comparison, for persons age 25 to 44, the rate was 4%. There are no published ECA rates that examine older adults in smaller age groups, nor in combination with gender or race.

7.11.4 CORRELATES OF LATE LIFE SUICIDE COMPLETION

7.11.4.1 Psychopathology

The primary method for studying the psychopathology associated with suicide in late life is the PA. Structured diagnostic interviews with knowledgeable informants allow for comparisons of risk for various diagnoses. Various estimates indicate that psychiatric disorder and/or substance use were present in about 90% of suicides, with affective disorder the most common psychopathology, followed by substance use and schizophrenia (Conwell, 1994). The diagnostic profile of suicides, however, differs with age. Conwell and Brent (1995) reviewed 14 PA studies from the USA and Europe that examined psychiatric disorders across age groups. Schizophrenia, substance use, and conduct disorder are more common among younger persons, while affective disorder, primarily major depression of a late onset, is more characteristic of older suicides. In addition, the proportion of suicides associated with substance use disorders also declines with age. Although the disorders of dementia and delirium increase substantially in late life, few studies have found these diagnoses associated with risk for suicide (Conwell & Brent).

Conwell and his colleagues (Conwell, 1994; Conwell, Duberstein et al., 1996) make a strong case for late onset depression as the most powerful diagnostic risk factor for late life suicide. Conwell (1994) also cites a clinical follow-up study that illustrates this point. Gardner et al. (1964) linked psychiatric case register information with death certificates for Monroe County, New York. Age effects were considered among persons less than 55 compared to those 55 and older. Among persons with affective disorders, one year rates for suicide among the younger group were 207 per 100 000, while for the older group, the rate was 475 per 100 000.

7.11.4.2 Previous Suicide Attempts

Although elderly suicide victims are much less likely than younger victims to have a history of past suicide attempts (Conwell, 1994; Kreitman, 1976a; Merrill & Owens, 1990; Pierce, 1987), and as noted above, the elderly in general have low prevalence of lifetime suicide attempts compared to younger age groups (Mościcki et al., 1988), previous suicide attempts appears to be a potent risk factor for later completion. Although one could argue that competitive mortality and physical frailty may work to produce this effect, greater suicide intent and lethality among the elderly would suggest that these effects are not the major forces behind this pattern (Conwell). Moreover, in the clinical registry follow-up study by Gardner et al. (1964), of the persons age 54 or less who attempted suicide, 2% completed suicide within a year. In contrast, among the persons age 55 and older, 6% completed suicide in the same period. Nordentoft et al. (1993) reported more recently on a 10 year follow up of Danish patients admitted to a poison treatment center for suicide attempts, where greater age was among the significant risk factors for subsequent suicide.

7.11.4.3 Physical Illness

Several studies (Conwell, Rotenberg, & Caine, 1990; Sainsbury, 1955, Whitlock, 1978) have indicated that physical illness is a common diagnosis among elderly suicides. When making age comparisons with suicide rates, older persons are more likely to have experienced a physical illness as a “life event” relative to younger persons (Rich, Young, & Fowler, 1986). The British studies (Sainsbury and Whitlock) noted above, found that cancer, in particular, was a frequent correlate of suicide in late life. However, a case controlled study by Henriksson and his colleagues found depressive syndromes to be equally present (80% and 82%) among cancer patients who suicided (two-thirds were age 60 and older) and age- and sex-matched comparison suicides without cancer histories (Henriksson, Isometsä, Hietanen, Aro, & Lönnqvist, 1995).

Other studies have indicated that the actual rates of physical illness among elderly suicides are lower than expected. Clark (1992) reported that among 73 consecutive elderly suicides in Cook County, Illinois, only 14% were terminally ill and 23% were chronically or severely medically ill at the time of death. Given the high rate of physical illness (chronic, acute, terminal) among older persons, it is imperative to further pose this question in case controlled studies.

7.11.4.4 Personality Dimensions

Research on younger samples of completed suicides has indicated an association with personality disorders and certain personality characteristics (Angst & Clayton, 1986; Marttunen, Aro, Henriksson, & Lönnqvist, 1991; Paffenbarger, Lee, & Leung, 1994; Runeson & Beskow, 1991). Interest in personality characteristics and disorders in late life suicide may have been limited due to the challenge of assessing them through informants, as well as the belief that diagnosed personality disorders have less validity in the elderly (Sadavoy & Fogel, 1992).

Duberstein and his colleagues (Duberstein, 1995; Duberstein, Conwell, & Caine, 1994) have used an informant-based measure of the NEO-Personality Inventory (Costa & McCrae, 1985) to examine possible personality traits among older and younger suicides, relative to controls. The NEO inventory measures five general personality traits: neuroticism, extroversion, openness to experience, agreeableness, and conscientiousness. Suicides were found to have higher neuroticism scores, and older suicide victims had lower openness to experience scores than both younger suicides and normal controls. Clark (1993) has theorized that a "narcissistic" crisis, based on a personality predisposition, leaves the older person with increasingly rigid, inflexible thoughts under stressful life events.

Duberstein (1995) has theorized about the behavioral correlates of low openness to experience (OTE) with implications for clinical assessment, the possible gender differences in OTE and depression, biological correlates, and association between OTE and attempted suicide. Each of these offer important clinical implications and plausible research directions in suicidal behavior in late life.

Some reports suggest that it may be the confluence of personality, poor coping skills, fear of illness and the hopelessness of depression that influences the association between physical illness and suicide in the elderly. Conwell, Caine, and Olsen (1990) presented a series of eight cases where the victims' fears that they had cancer played a major role in committing suicide. Among these eight cases, only one was reported to have cancer upon autopsy, and five had diagnoses of major depression. Similarly, Horton-Deutsch, Clark, and Farran (1992), reporting on 14 cases from the study cited by Clark (1992) above, describe the co-occurrence of chronic dyspnea, increasing limitations in functional status, life events, and depression and anxiety among elderly men who suicided. Common to both of these studies is the description of the personality characteristics of the majority of elderly male suicides: appearance of being overtly strong and self-sufficient, yet covertly or overtly having difficulty in being dependent and accepting help, and having a rigid, inflexible coping style, not unlike a lack of OTE.

7.11.4.5 Life Events

7.11.4.5.1 Physical illness and loss

Significant age differences in recent life events among completed suicides have been reported in several studies (Rich et al., 1986; Heikkinen & Lönnqvist, 1995), where the proportion of victims with current somatic illness increases with age, and marital separations, family discord, and job problems are more common among younger suicides. When physical illness is excluded, few life events are found among elderly suicide victims. Outside physical illness, the most commonly cited life event associated with late life suicide is "loss." Loss in this context has referred to loss of spouse, family members, friends, job, physical autonomy, and impending nursing home placement (Conwell, 1994). Unfortunately, such losses are common in late life in general. Once again, without adequate control groups, it is not clear in what ways these events become risk enhancing for suicidal behavior.

7.11.4.5.2 Living alone

Another frequently cited risk factor is social isolation and living alone, which may or may not be linked to life events such as widowhood or family estrangement. Barraclough (1971) reported that among 30 suicides of persons age 65 and older, 50% lived alone, compared to 20% in the community from which they lived. More recent studies conducted in the USA have not found similar patterns. For example, Conwell et al. (1990) did not find an increase in living alone among complete suicides age 50 and older. Only 35% were found to be living alone; more than one-third of elderly suicide victims were married, and more than half lived with a spouse or family member.

Perhaps more important are the distinctions to be made among the constructs of living alone, social isolation, and loneliness. Living alone does not imply social isolation or loneliness. Younger et al. (1990) reported that among the nearly 50% of people living alone who had committed suicide, 60% had weekly visits with others outside the home and 98% had weekly contact with friends and family. The degree to which these suicide victims felt lonely, despite their contact with others, is not known and suggests the need for separating these constructs in both clinical assessment and future research.

7.11.4.6 Neurobiological Correlates

7.11.4.6.1 Serotonergic system

To date, there are still no specific biological markers for either depression or suicidal behavior. However, there are hopes that such biological information, in conjunction with psychologically assessed data, can improve detection of high risk persons, with prevention and treatment implications. If biological subtypes of suicidal behavior were to be discovered, for example, more targeted pharmacologic interventions might be developed. With regard to the elderly, it is conceivable that a "neurobiological vulnerability" to suicide might undergo modulation by age-related changes in neurobiological systems (Rifai, Reynolds, & Mann, 1992; Schneider, 1996). The consistency of suicide risk increasing with age and male gender across nations also suggests a possible neurobiological process (Conwell, Raby, & Caine, 1995). Decreased brain concentrations of serotonin, dopamine, norepinephrine and their metabolites (HVA, 5-HIAA); increased brain MAO-B activity; increased hypothalamic–pituitary–adrenal activity; and increased sympathetic nervous system activity are associated with both depression and normal aging (Schneider, 1996). Although it would seem that these correlates would have led to more neurobiological studies of elderly suicides, few have been conducted.

Several reviews have examined the evidence for neurobiologic abnormalities among elderly suicide victims relative to controls and have found too few studies that had included sufficient "elderly" subjects (older than 60 years) to draw any conclusions (Conwell & Brent, 1995; Rifai et al., 1992). This is particularly true of the subgroup of the elderly most at risk; those 85 and older. Conwell and Brent have also described the challenges to be met in obtaining sufficient cases to test various neurobiological hypotheses, and have outlined a protocol for standardization of data for postmortem research protocols.

7.11.4.6.2 Serum cholesterol

The association between serum cholesterol concentrations and late life comes from the heart disease outcomes literature. In clinical trial studies aimed at lowering serum cholesterol in persons either at risk for, or having heart disease, gains in lowered heart-related deaths have been off-set by increases in suicide and accidental deaths (e.g., Muldoon, Manuck, & Matthew, 1990). Among men age 70 and older, depression is also three times more common when plasma cholesterol concentrations are low (Morgan, Palinkas, Barrett-Connor, & Wingard, 1993). However, other studies have not found this association in free cholesterol levels, but have in the esterification of cholesterol (the reverse cholesterol transport; Maes et al.,1994). Clearly, further research in the cardiovascular and neurobiological correlates of suicidal behavior in later life are needed.

7.11.4.7 Health Service Use Patterns

Clark (1992) has noted that suicides in the USA are not higher for those with less access to health care. To the contrary, suicide rates are lower for African-Americans and Hispanics (groups who underuse health care) relative to their white counterparts. Indeed, a phenomenon common among elderly suicide completers and attempters is that the majority have recently visited a health care provider. Importantly, however, virtually none of these elderly had sought mental health services. This is in contrast to younger aged adults, where 25–30% were under the care of a mental health professional at the time of their death (Clark & Fawcett, 1992).

Instead, over 70% of older suicides have visited their primary care physician within the month of their suicide in the USA (Conwell, 1994) and the UK (Vassilas & Morgan, 1994). A similar pattern appears for elderly attempters in Japan (Takahashi et al., 1995), where more attempters relative to depressive controls sought out primary care physicians. This pattern may be in part due to physical concerns that reflect actual acute or chronic illness, as rates of comorbid physical illness among elderly suicides are also high compared to younger groups (Cattell, 1988; Conwell, 1994). This pattern may also be related to hypochondriacal concerns (Barraclough et al., 1974), and fear of pain, physical illness, and concomitant dependency as described above (Clark, 1993; Conwell et al., 1990; Duberstein, 1995; Horton-Deutsch et al., 1992).

The failure to recognize and adequately treat older adults with depression is probably due to several causes. First, the elderly themselves and their family members may incorrectly attribute depressive symptoms to the aging process, and see depression as a "normal" aspect of aging, and an expectable side effect of physical illness. Primary care physicians tend to spend less time with older patients and focus on physical problems to the exclusion of mental morbidity (German et al., 1985; Keeler, Solomon, Beck, Mendenhall & Kane, 1982). Second, there may be hesitancy or inability to report psychiatric symptoms even when specifically asked. Gallo, Anthony, and Muthén, (1994) have found that older depressed adults are less likely to endorse dysphoric mood as part of their symptom clusters. Older men, in particular, are less likely to report depressive symptoms (Allen-Burge, Storandt, Kinscherf, & Rubin, 1994; Lutsky & Knight, 1994). Third, primary care physicians, who are the most sought after health care providers among all persons with depression in the USA, typically do not recognize (e.g., Perez-Stable, Miranda, Munoz, & Ying, 1990) nor adequately treat depression among any age group (Sturm & Wells, 1995).

7.11.5 STUDIES OF ELDERLY SUICIDE ATTEMPTERS

Given the limited data on the epidemiology of late life suicide attempts, the bulk of this discussion focuses on studies of hospital based samples. In a review of attempted suicide in late life, Draper (1996) reviewed 12 studies published between 1985 and 1994. Despite variation in sampling contexts and approaches to measurement, and lack of adequate control groups, there were several consistent factors associated with attempted suicide in late life: depression, social isolation, and being unmarried. The degree to which physical health was a risk factor was unclear; in some studies it appeared to play a major role (Draper, 1994; Frierson, 1991; Nowers, 1993), while in another only about a third of the patients identified health as a factor (Pierce, 1987). Draper suggested that the interaction between psychosocial factors such as social isolation, loneliness, family conflict, and physical factors such as presence of pain and perceived limitations may be a more critical combination for risk of suicide attempt than physical illness alone. A gender difference by age pattern was also noted across the studies: the proportion of males attempting suicide increased with age, particularly among those age 75 and older. Without adequate control groups, however, these findings remain interesting leads rather than true advances in our understanding of risk factors for suicide attempts in late life.

Among the few studies with control groups, elderly depressives admitted for a suicide attempt have been more likely to have a later age of onset than elderly depressives who were not suicidal (Lyness, Conwell, & Nelson, 1992). In a German study an age by sex interaction in age of onset was noted, where men age 50 and older were more likely to have had an affective illness of late onset, and women had a younger age of onset and were more likely to have had previous suicide attempts (Bron, Strack, & Rudolph, 1991).

Psychosis and somatization have characterized the psychiatric symptoms of elderly suicide attempters. Lyness et al. (1992) and Roose et al., (1983) found psychotic depression to be more related to suicidal behavior than other affective disorders in retrospective cohort studies. Psychotic depression is interpreted by some to include delusions about physical symptoms. In a Japanese study of urban elderly attempters, Takahashi et al. (1995) examined the rates of somatization (defined as a separate entity from depression) among attempters and controls. Consistent with earlier descriptive studies of high rates of hypochondriasis among elderly attempters (Barraclough et al., 1974; De Alarcon, 1964), they found higher rates of somatization among attempters relative to controls. Although somatization can be defined as a type of psychosis, its unique role and cooccurrence with depression in late life appears worthy of further investigation.

7.11.6 SUBGROUPS OF ELDERLY AT RISK FOR SUICIDE

Suicide is a multidetermined, low base rate event. Despite the many correlates and posited risk factors identified, the prediction of suicidal behavior is next to impossible (Goldstein, Black, Nasrallah, & Winokur, 1991; Pokorny, 1983). However, since suicide risk in the elderly is strongly associated with certain demographic (white males) and psychological variables (depressive diagnoses and rigid coping), it follows that certain subgroups of the elderly are more or less likely to be at risk for suicidal behavior. The primary intent of proposing subgroups here is to stimulate research questions that consider the interaction of such demographic characteristics with other known risk or protective factors that may help the field move beyond studies that focus on main effects.

7.11.6.1 Intolerance for Physical Illness or Dependency

The discussions of somatization and findings pertaining to the role of physical illness in elderly suicide would suggest that there are persons, who, when faced with what they perceive to be an illness that may limit autonomy, become risks for suicide. White males, living in the community, who have taken pride in being independent and self-sufficient would be the prototype of this subgroup. Stated fears of losing what are considered markers of independence, such as driving or living independently, may reflect underlying intolerable anxiety and depression. These fears, in combination with suspicion of having an incurable illness with functional limitations, may create a lethal combination. In contrast, studies of patients with terminal illnesses, who have learned to readjust their expectations about life have a lower risk of suicidality (Marshall, Burnett, & Brasure, 1983). Measures of hopelessness (Beck, Brown, Berchick, Steward, & Steer, 1990; Beck, Steer, Kovacs, & Garrison, 1985; Hill, Gallagher, Thompson, & Ishida, 1988; Rifai, George, Stack, Mann, & Reynolds, 1994), limited openness to new experience (Duberstein, 1995), perfectionism (Hewitt, Flett, & Turnbull-Donovan, 1992), and willingness to adapt may help screen those elderly with physical illness who may be at risk for depression and suicidal ideation.

7.11.6.2 Older Alcoholics

Another group at risk for suicide are older male alcoholics with early onset who have survived to be among the "young old." Murphy and his associates described a number of factors associated with suicide among alcoholics. For the older male alcoholics, loss of the last social support has been found to be a pivotal event in suicide risk (Murphy, Wetzel, Robins, & McEvoy, 1992). Similarly, Duberstein et al. (1994) reported that interpersonal stressors and recent alcohol use combined to increase risk for suicide. In a Finnish sample, Heikkinen et al. (1994) compared nonalcoholic depressives and alcoholics who had committed suicide. Unlike depressives, alcoholics were likely to have recent interpersonal loss, unemployment, and financial trouble. Drinking patterns appeared to lead to these concurrent stressors. Assessment of depression as well as current social support among older alcoholics would seem to be essential in understanding risk for suicidal behavior in this subgroup.

7.11.6.3 Immigrants and Ethnic Minority Elderly

Studies conducted in Hong Kong (Yap, 1958), Singapore (Ko & Kua, 1995), Australia (Burvill, 1995); Burvill, Woodings, Stenhouse, & McCall, 1982), and Canada (Kliewer & Ward, 1988) indicate that suicide rates are higher among immigrants than natives of the host country. Such research is complicated as it must consider factors that may influence suicide risk in the country of origin (e.g., religion, the individuals' reason for leaving and mental health history); factors within the country of migration; and the society of settlement (e.g., similar or different in ethnicity, and degree of acculturation) (Murphy, 1977). With regard to elderly suicide, the study of these factors are even more challenging to consider as age of immigration can vary widely (e.g., the elderly migrant may have emigrated within the year of suicide, or 80 years prior), and familial networks will vary with the degree to which acculturation is desired or avoided. Despite these challenges, research on elderly suicide in countries and cities where immigration is high (e.g., USA, Hong Kong, Australia), comparing immigrant groups of varying rates of suicide, may help explicate the role of culture, religion, and family interactions in suicidal risk in the elderly (see Corin, 1995).

Very little empirical data exists about the risk or protective factors for suicidal behavior among African-American, non-white Hispanic, and Asian ethnic minority groups in the USA. However, theory development about risk and protective factors among ethnic minority Americans can be found (Griffith et al., 1989; Sorenson & Golding, 1988). Studies on the risk for depression in older ethnic minorities should provide clues as to risk for suicidal behavior in later life (e.g., Baker, 1996). As the number of American ethnic minority elderly are projected to increase substantially in numbers in the next few decades (US Bureau of the Census, 1996), much must be learned to best assess and treat suicidal behavior in these subpopulations in order to adequately provide care and service in the future.

7.11.6.4 Older Women

A fourth subgroup needing further study are elderly women. Canetto has discussed what is known about gender issues in suicide, including what little is known about older female suicides (Canetto, 1992, 1994, 1995). Of the many sociologic risk factors proposed to increase risk for suicide, including retirement, financial stress, widowhood, living alone, and poor physical health, older women suffer more than older men in each of these categories, yet have lower suicide rates (Canetto, 1995). Moreover, of the mental disorders thought to increase risk for late life suicide, depression is the most frequent disorder cited and older women have higher rates of depression than older men (Kennedy et al., 1989; Livingston, Hawkins, Graham, Blizard, & Mann, 1990; Romanoski et al., 1992), as well as more functional impairment with depression (Kornstein, et al., 1995).

Because of these contradictions, Canetto has recommended that discussions of elderly suicide become more precise with regard to gender. Clearly, applying typical risk factors that reflect older male suicides in studies of older women will not yield great benefits. What may be of more utility is to examine gender differences in acceptability of completed suicide as a behavior (Canetto, 1995), as well as examining gender differences in coping and cognitive styles. Rumination, a cognitive style more common among women, may prolong depressive episodes and serve to increase prevalence rates (Nolen-Hoeksema, 1987; Nolen-Hoeksema, Morrow, & Fredrickson, 1993). Interpersonal sensitivity and a relational interactional style more common among women (Markus & Oyserman, 1989) makes them more sensitive and responsive to others' problems. Although this gender-typed style may contribute to prolonged depressive episodes among women, it may also protect them from acting on suicidal thoughts (Canetto, 1992). Little is known about how interpersonal sensitivity or a ruminative cognitive style are related to depression or suicidal ideation in older women, or what mediating factors such as social support influence these possible associations.

With regard to the older women who do complete suicide, one could hypothesize that older female suicides have had more previous attempts, based on epidemiologic data of middle-aged populations (Mościcki et al., 1988). Higher rates of previous attempts is also related to borderline personality disorder and less lethal attempts (Rich et al, 1986; Runeson & Beskow, 1991), which also are more common among women. Nolen-Hoeksema (1995) has articulated the biological and psychosocial theories of the evolution of gender differences in depression. There is evidence to suggest that some of the factors that increase rates for depression in women are similar to those for suicidal behavior, particularly histories of abuse in childhood and adulthood (Davidson, Hughes, George, & Blazer, 1996; Kaplan, Asnis, Lipschitz, & Chorney, 1995).

7.11.6.5 Nursing Home Patients

If nursing home settings are considered (where the vast majority of residents are female), less lethal and more passive attempts, such as refusal of food or adherence to medications may characterize suicidal behavior in this subgroup. Indeed, Osgood surveyed long-term care facilities and reported that most self-destructive behaviors in nursing homes were indirect, and that setting characteristics were related to rates of self-destructive behavior (Osgood, 1992; Osgood & Brant, 1990). Building on this work, Conwell, Pearson, and DeRenzo (1996) have summarized several research issues on self-destructive behaviors in nursing home patients. Included are patient-related factors (psychiatric illness, physical health, personality and coping styles, religious beliefs, and neurobiological beliefs), setting influences (nature and size of facility, quality of care provided) and their interactions, as well as ethical issues involved. Given the high prevalence rates of mental disorders in nursing homes that go undetected and untreated (Rovner, Kafonek, Filipp, Lucas, & Folstein, 1986; Strahan, 1991), the need for further study of self-destructive behavior, both direct and indirect, is clearly needed in these settings.

7.11.7 BIOETHICAL ISSUES

7.11.7.1 Physician Assisted Suicide

The question of risks for suicidal behavior in nursing homes brings to the surface the many complicated issues involved in passive and active euthanasia, advance directives, and physician-assisted suicide (see Chapter 6, this volume). Although there is consensus that decisions regarding the limitation of life-saving procedures are not "suicidal" in terminal illness, such as requesting do-not-resuscitate orders as a part of acute health care treatment, the role of the physician to hasten the process by active means remains controversial. The US public, health care guilds, and health care providers continue to debate patients' rights and physician-assisted suicide and in some cases, act on their beliefs (e.g., see Gibbs, 1993 on Kervorkian; Quill, 1991). In March 1996, the 9th US Circuit Court of Appeals (Compassion in Dying vs. State of Washington, 1996) ruled that *mentally-competent, terminally-ill* adults have a constitutionally protected liberty interest in hastening what might otherwise be a protracted, undignified and extremely painful death. The guiding principle here is that of patient autonomy. Patients have the right to control their own bodies and to decide for themselves the course of their own medical treatment.

The salience and complications of these issues vary with the health care context, perspective and health status of the patient, and age and perspective of the clinician (Marks, 1988–1989). Among adults with terminal illnesses who are cognitively intact, refusal of treatment or requests for physician-assisted suicide appears less controversial. With regard to the elderly, many of whom may not be competent to advocate for themselves, or may or may not have family members who can be trusted to represent their wishes, the issue is very controversial (Osgood, 1995). Competency is a pivotal issue, because cognitive impairment and depression often co-occur with terminal illness. Sullivan and Youngner (1994) highlight differences in clinicians' views on competency in the elderly by examining the refusal of life-sustaining treatment by patients with serious medical illness and comorbid depression. In medical settings, competence to refuse life-sustaining treatment is typically presumed unless the physician proves otherwise. In psychiatric settings, a desire to die is considered evidence of a mental disorder and proves justification of treatment.

7.11.7.2 Competency and Depression

Sullivan and Youngner (1994) recommended that depression in medical patients be treated to improve competency. If treatment is not successful, a residual diagnosis of depression is neither necessary nor sufficient for determining that the patients' medical decision is impaired. In their argument, depression and competency are separate issues. Yet others argue that depression can render limitations on consciousness, where alternatives other than suicide cannot be clearly considered by the patient (Clark, 1992; Koenig, 1993). Shneidman (1986) represents this same sentiment in his assessment of the cognitive narrowing and inflexibility of the suicidal patient: "It is not a thing to do while one is not in one's best mind. Never kill yourself when you are suicidal" (p. 8).

Studies on how depression affects decision making in late life are limited, but some evidence suggests that depression can affect the consistency in preferences for end of life decisions. Lee and Ganzini (1992) studied elderly persons seeking medical treatment, and found mild to moderate depression not to affect preference for life-sustaining medical therapy (intravenous fluids with medications, feeding tube, blood transfusions, intensive care, kidney dialysis, ventilatory support, and cardiopulmonary resuscitation). Among severely depressed elderly, however, there was less consistency in preferences over time. Lee and Ganzini (1994) have also documented little consistency among depressed older medical inpatients who recover from their depression. When depressed subjects who recovered changed their preferences, they were as likely to change from refusing to desiring a treatment as vice versa. Those who became depressed, and remained depressed, were likely to refuse an intervention they had initially asked for.

Sullivan and Youngner (1994) have argued that studies of nursing home residents provide equivocal evidence for the effects of depression on life-sustaining treatments, where depression in patients has been related to both decreased and increased desire for life-sustaining treatments (Garrett, Harris, Norburn, Patrick, & Davis, 1993) compared to preferences among nondepressed residents. What may contribute to inconsistencies in findings may be the mixing of research methods (hypothetical future illness versus current state of illness). Attempting to draw comparisons across diverse groups of elderly may be misleading due to the demographic and health differences among the elderly: community elderly seeking medical care are likely to differ significantly from nursing home residents with regard to age, gender, general health status, and perspectives on death.

Consistent with Lee and Ganzini's work, along with commonly found risk factors for late life suicide (male gender, white ethnicity), other evidence suggests that older persons' attitudes towards suicide may be related to their perceived social support and health and mental state (i.e., depression). Seidlitz, Duberstein, Cox, and Conwell (1995), using data from a Gallup Poll, found that nonreligious white men, especially those with strained family relationships and self-perceived failing health, were the most likely to endorse lenient attitudes toward suicide. Although owning such an attitude does not predict suicidal behavior, it does lower the threshold for tolerance of this option. Moreover, since lenient attitudes towards suicide may reflect hopelessness about family relations and self-perceived health—aspects that may be subject to change with intervention—treatment of depression or anxiety among patients holding such attitudes may affect lenient attitudes toward suicide.

7.11.7.3 Bioethical Research Needs

Future research that considers both patient-related and setting influences may help reconcile the seemingly contradictory evidence regarding depression, end-of-life decisions, and suicide. In considering these issues in the nursing home setting, Conwell, Pearson, and DeRenzo (1996) identified five patient-related factors relevant to self-harming behaviors: assessment of psychiatric symptomatology (dementia, depression, hopelessness, suicide intent, ideation); physical health (including somatic complaints and pain); personality and coping style (rigid coping style, openness to experience); religious beliefs; and neurobiological influences. Although setting influences described by Conwell, Pearson and DeRenzo (1996) were directed towards nursing homes, the context of access and willingness to seek health care, social supports, and attitudes of family members and health care providers can be assessed among the hospitalized and community dwelling elderly, and should be factored in the investigations of the individuals' decisions about end-of-life decisions.

An important implication of implementing right to refuse treatment policies or enacting physician-assisted suicide laws may be how such options affect older persons' attitudes about suicide. Increased availability and use of hospice care could conceivably affect older persons' attitudes towards suicide. If the end of life seems less likely to be characterized by pain, suffering, and dependence, dying by a "natural course" may seem more tolerable, and suicide a less preferred suicide option (Amenta & Shubert, 1995). For others, physician-assisted suicide may hasten death and reduce the time of enduring intolerable pain and suffering. Posner (1995) has argued that the perception that physician-assisted suicide can be available as needed, could be an effective means of suicide prevention in later life. These are researchable issues that can and should be addressed to best inform legislators and policy makers.

7.11.8 CLINICAL RESOURCES FOR TREATING SUICIDAL OLDER ADULTS

Suicide is an attempt to solve a problem of unbearable psychological pain, where few alternative solutions are seen or considered by the patient (Shneidman, 1992). The inability to consider that solutions, even though not obvious now, may be present in the future, reflects the serious hopelessness felt by suicidal patients (Beck et al.,1990; Beck et al., 1985; Rifai et al., 1994). Aspects of approaches to suicidal patients that appear particularly applicable to the elderly include the assessment of what Shneidman termed "cognitive constriction" (Shneidman, 1985). Constriction can be defined generally as dichotomous thinking, tunnel vision, or a narrowing of the range of options. As long as alternative solutions are not seen as plausible, suicide as a viable option remains and risk for acting on suicidal ideation is high. Clinical interventions for suicidal persons, whether they are individual, family, group, or pharmacologic in nature, must be effective in providing the alternative solutions to suicide.

As multiple psychological autopsy studies have shown, more than 90% of cases of adult suicides have had a recent major psychiatric disorder (Clark & Fawcett, 1992). Moreover, the lifetime probability of suicidal death is 15% for persons with a major affective disorder (Klerman, 1987). Therefore, the prevention of suicide in late life in both a crisis situation, and in the long run, will be best accomplished by the increased recognition and adequate treatment of depression (NIH Consensus Conference, 1992). As such, most of the clinical resources suggested below address treatment of depression primarily, with secondary benefits to reducing risk for suicide.

Unfortunately, there are few studies that address the effectiveness of various interventions with suicidal persons in general, and virtually none focusing on the elderly. Crisis hotlines are nearly impossible to evaluate (Lester, 1993), and little information exists on the elderly's use of such services in the USA (Conwell, 1994). Adamek and Kaplan (1996), in a survey of American Association of Suicidology listed crisis prevention centers, found that even though about a third had specialty programs for the elderly or specific outreach efforts, two-thirds of the staff did not know that older adults are the age group with the highest suicide rate. Although crisis lines may not be effective with the elderly, telephone outreach programs do show some promise of reducing psychological distress and suicidal ideation, if not completed suicide (see McIntosh, 1995, for a review).

With regard to comparative interventions with suicidal elderly, no studies exist, highlighting the need for this type of research. Therefore, the clinical resources described below are not proven to be effective treatments, but do reflect the state-of-the-art and limited science.

7.11.8.1 Psychotherapeutic Resources

There are a growing number of guidelines and descriptions of psychotherapeutic approaches to suicidal outpatients (e.g., Bongar, 1992; Leenaars, Maltsberger, & Neimeyer, 1994). Increasingly among these guidelines are articles, chapters, and even books, devoted solely to the suicidal elderly (e.g., Lester & Tallmer, 1994; McIntosh, et al., 1994; Richman, 1993). The most frequent author of chapters and papers on psychotherapeutic approaches to suicidal elderly is Joseph Richman, a family therapist who integrates research findings with nearly 30 years of clinical experience with suicidal elderly

(Richman, 1996). Richman's approach typically includes the family as well as the patient, as he considers the family an important context of the older person. The family can both contribute to the risk, and to the treatment of the suicidal elderly. His book, *Preventing elderly suicide,* reviews research that he has grounded his approach in, assessment, crisis intervention, and individual, family and group modalities (Richman, 1993).

Cognitive-behavior therapy has also been used successfully to treat suicidal persons (Beck, Rush, Shaw, & Emery, 1979; Beck & Weishaar, 1990), including the elderly. Generally described, cognitive behavior therapy, as applied to the suicidal individual, identifies the dysfunctional beliefs that are typically helpless and hopeless in character, the feelings associated with the beliefs, and then directs the client to problem solving methods. By forming a more precise description of the problem separate from depressed mood, breaking down the problem into components, and considering possible solutions, suicide as an only option begins to dissipate. DeVries and Gallagher-Thompson (1994) have described how this approach can be applied specifically to older adults in crisis, including suicidality.

Maltsberger (1991) applies a more psychoanalytic, "ego deficit" model of suicide in the elderly, where the suicidal person is narcissistically vulnerable, fearing loss of mental and physical capacity, as well as abandonment. If the older patient is unable to ask or obtain support for these functions and self-contempt and anxiety increase, the risk of suicide increases. In this approach, the therapist can play a critical role during this narcissistic crisis by providing resources through his or her warmth, encouragement, and availability.

7.11.8.2 Pharmacologic and Electroconvulsive Therapies

7.11.8.2.1 Pharmacologic therapies

Serotonergic abnormalities in patients who are suicidal would suggest that treatments using serotonin reuptake inhibitors (SSRIs, e.g., fluoxetine, paroxetine, sertraline) would be effective (Montgomery, 1993). Since the SSRIs are just as effective as the earlier tricyclic antidepressants but have lower side-effect profiles among the elderly (Schneider, 1996), acceptability and treatment compliance with SSRIs among the elderly may be improved. Furthermore, because fluoxetine has been shown to decrease suicidal ideation relatively early in treatment among adults ages 18–65 (Worthington et al., 1995), its use with older suicidal adults may be life saving. Although there is no research pointing to particularly effective pharmacologic treatments for suicidal elderly, there are several reviews that describe the many effective pharmacologic approaches to treating late life depression (Hay & Hay, 1996; Kennedy, 1996; Salzman, 1994). More studies that track the presence and change and continuity of suicidal ideation over time of treatment and follow-up are also needed (see Rifai et al., 1994 for one example).

7.11.8.2.2 Electroconvulsive therapy

Electroconvulsive therapy, although still controversial, can be an extremely effective treatment for elderly patients with depression who are not able to tolerate the side effects of pharmacotherapies. Moreover, it is often effective where patients have not responded to traditional pharmacologic interventions. The speed of onset of its action, which is typically more efficient than medication, also makes it an indicated treatment for suicidal persons and those with severe depressions in late life. The efficacy, applications, adverse effects, and contra-indications for use in the elderly have been reviewed elsewhere (Hay & Hay, 1996; Sackheim, 1994). Although there are no outcome studies specific to suicidal elderly, Avery and Winokur (1976, 1978) have demonstrated that electroconvulsive therapy does decrease the number of suicidal outcomes in depressed patients.

7.11.9 SUMMARY

Suicide among the elderly is one of the most critical public health problems facing us today. We have sufficient information on what appear to be strong correlates if not risk factors to address several levels of preventive interventions (Conwell & Duberstein, 1995; Rihmer, 1996). Although suicide in the elderly as a public health problem has probably been ignored due to agism, like other areas of psychopathology in late life, clinical, research, and policy interests in late life suicide (Suicide and the elderly: A population at risk, 1996) have grown with the increasing numbers of older persons in the USA, and other countries. Perhaps current older cohorts and younger cohorts will become more active in their interests in maintaining and increasing their own well-being (McIntosh, Pearson, & Lebowitz, 1997). As "baby boomers" age and demand more and better health care, the interest in the prevention and treatment of psychopathology of later life will certainly grow. It is hoped that this chapter has provided some useful groundwork, and has raised some potentially important research directions for the much needed information that will be required to meet the future needs and interests of the aging population and their family members.

ACKNOWLEDGMENTS

Dr. Paul Duberstein provided invaluable suggestions and comments on this chapter. His expert advice and criticism is greatly appreciated.

7.11.10 REFERENCES

Adamek, M. E., & Kaplan, M. S. (1996). Managing elder suicide: A profile of American and Canadian crisis prevention centers. *Suicide and Life-Threatening Behavior, 26,* 122–131.

Allen-Burge, R., Storandt, M., Kinscherf, D. A., & Rubin, E. H. (1994). Sex differences in the sensitivity of two self-report depression scales in older depressed inpatients. *Psychology and Aging, 9,* 443–445.

Amenta, M. O., & Shubert, V. (1995). Hospice care can obviate the need for suicide. *Home Healthcare Nurse, 13,* 72–74.

Angst, J. & Clayton, P. (1986). Premorbid personality of depressive, bipolar, and schizophrenic patients with special reference to suicidal issues. *Comprehensive Psychiatry, 27,* 511–531.

Avery, D., & Winokur, G. (1976). Mortality in depressed patients treated with electroconvulsive therapy and antidepressants. *Archives of General Psychiatry, 33,* 1029–1037.

Avery, D., & Winokur, G. (1978). Suicide, attempted suicide, and relapse rates in depression. *Archives of General Psychiatry, 35,* 749–753.

Baker, F. M. (1996). Suicide among ethnic elders. In G. J. Kennedy (Ed.), *Suicide and depression in late life: Critical issues in treatment, research and public policy* (pp. 51–79). New York: Wiley.

Barraclough, B. M. (1971). Suicide in the elderly: Recent developments in psychogeriatrics. *British Journal of Psychiatry, special supplement, 6,* 87–97.

Barraclough, B. M., Bunch, J., Nelson, B., & Sainsbury, P. (1974). A hundred cases of suicide: Clinical aspects. *British Journal of Psychiatry, 125,* 355–373.

Beck, A. T., Brown, G., Berchick, R. J., Stewart, B. L., & Steer, R. A. (1990). Relationship between hopelessness and ultimate suicide: A replication with psychiatric outpatients. *American Journal of Psychiatry, 147,* 190–195.

Beck, A.T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression.* New York: Guilford.

Beck, A. T., Steer, R. A., Kovacs, M., & Garrison, B. (1985). Hopelessness, depression, and attempted suicide. *American Journal of Psychiatry, 142,* 559–563.

Beck, A. T., & Weishaar, M. E. (1990). Suicide risk assessment and prediction. *Crisis, 11,* 22–30.

Beskow, J., Runeson, B., & Asgard, U. (1990). Psychological autopsies: Methods and ethics. *Suicide and Life-Threatening Behavior, 20,* 307–323.

Bongar, B. (1992). *Suicide: Guidelines for assessment, management and treatment.* New York: Oxford University Press.

Bron, B., Strack, M., & Rudolph, G. (1991). Childhood experiences and loss and suicide attempts: Significance in depressive states of major depressed and dysthymic or adjustment disordered patients. *Journal of Affective Disorders, 23,* 165–172.

Buda, M. & Tsuang, M. T. (1990). The epidemiology of suicide: Implications for clinical practice. In S. J. Blumenthal & D. J. Kupfer (Eds.), *Suicide over the life cycle: Risk factors, assessment, and treatment of suicidal patients* (pp. 17–37). Washington, DC: American Psychiatric Press.

Burvill, P. W. (1995). Suicide in the multiethnic elderly population of Australia, 1979–1990. *International Psychogeriatrics, 7,* 319–333.

Burvill, P. W., Woodings, T. L., Stenhouse, N. S. & McCall, M. G. (1982). Suicide during 1961–1970 of migrants in Australia. *Psychological Medicine, 12,* 295–308.

Canetto, S. S. (1992). Gender and suicide in the elderly. *Suicide and Life-Threatening Behavior, 22,* 80–97.

Canetto, S. S. (1994). Gender issues in the treatment of suicidal individuals. In A. A. Leenaars, J. T. Maltsberger, & R. A. Neimeyer (Eds.), *Treatment of suicidal people: A series in death education, aging, and healthcare.* (pp.115–126). Washington, DC: Taylor & Francis.

Canetto, S. S. (1995). Elderly women and suicidal behavior. In S. S. Canetto & D. Lester (Eds.), *Women and suicidal behavior* (pp. 215–233). New York: Springer.

Cattell, H. R. (1988). Elderly suicide in London: An analysis of coroners' inquests. *International Journal of Geriatric Psychiatry, 3,* 251–261.

Chemtob, C. M., Hamada, R. S., Bauer, GB., Kinney, B., & Torigoe, R. Y. (1988). Patients' suicides: Frequency and impact on psychiatrists. *American Journal of Psychiatry, 145,* 224–228.

Chemtob, C. M., Hamada, R. S., Bauer, GB., Torigoe, R. Y. & Kinney, B. (1988). Patient suicides: Frequency and impact on psychologists. *Professional Psychology: Research and Practice, 19,* 421–425.

Clark, D. C. (1992). "Rational" suicide and people with terminal conditions or disabilities. *Issues in Law & Medicine, 8*(2), 147–166.

Clark, D. C. (1993). Narcissistic crises of aging and suicidal despair. *Suicide and Life-Threatening Behavior, 23,* 21–26.

Clark, D. C., & Fawcett, J. (1992). Review of empirical risk factors for evaluation of the suicidal patient. In B. Bongar (Ed.), *Suicide: Guidelines for assessment, management, and treatment* (pp.16–48). New York: Oxford University Press.

Clark, D. C., & Horton-Deutsch, S. L. (1992). Assessment *in Absentia:* The value of the psychological autopsy method for studying antecedents of suicide and predicting future suicides. In R. W. Maris, A. L. Berman, J. T. Maltsberger, & R. I. Yufit (Eds.), *Assessment and prediction of suicide* (pp. 145–181). New York: Guilford.

Clarke, R. V., & Lester, D. (1987). Toxicity of car exhausts and opportunity for suicide: Comparison between Britain and the United States. *Journal of Epidemiology and Community Health, 41,* 114–120.

Cohler, B. J. & Jenuwine, M. J. (1995). Suicide, life course, and life story. *International Psychogeriatrics, 7,* 199–219.

Compassion in Dying v. State of Washington. (1996). 96 C.D.O.S. 1507. [On-line]. Available: http://www.law.vill.edu/fed-ct/circuit/9th/opinions.

Conwell, Y. (1994). Suicide in elderly patients. In L. S. Schneider, C. F. Reynolds, B. D. Lebowitz, & A. J. Friedhoff (Eds.), *Diagnosis and treatment of depression in late life: Results of the NIH consensus development conference* (pp. 397–418). Washington, DC: American Psychiatric Press.

Conwell, Y. & Brent, D. (1995). Suicide and aging I: Patterns of psychiatric diagnosis. *International Psychogeriatrics, 7,* 149–164.

Conwell, Y., Caine, E. D., & Olsen, K. (1990). Suicide and cancer in late life. *Hospital and Community Psychiatry, 41,* 1334–1339.

Conwell, Y., & Duberstein, P. R. (1995). Prevention of late life suicide: When, where, why and how. *Psychiatry and Clinical Neurosciences, 49*(Suppl), S79–S83.

Conwell, Y., Duberstein, P. R., Cox, C., Herrmann, J. H., Forbes, N. T., & Caine, E. D. (1996). Relationships of age and Axis I diagnoses in victims of completed suicide: A psychological autopsy study. *American Journal of Psychiatry, 153,* 1001–1008.

Conwell, Y., Pearson, J., & DeRenzo, E. G. (1996). Indirect self-destructive behavior among elderly patients in nursing homes. *American Journal of Geriatric Psychiatry, 4,* 152–163.

Conwell, Y., Raby, W. N., & Caine, E. D. (1995). Suicide and aging II: The psychobiological interface. *International Psychogeriatrics, 7,* 165–181.

Conwell, Y., Rotenberg, M., & Caine, E. D. (1990). Completed suicide at age 50 and over. *Journal of the American Geriatrics Society, 38,* 640–644.

Corin, E. (1995). From a cultural stance: Suicide and aging in a changing world. *International Psychogeriatrics, 7,* 335–355.

Coryell, W., & Tsuang, M. T. (1982). Primary unipolar depression and the prognostic importance of delusions. *Archives of General Psychiatry, 39,* 1181–1184.

Costa, P. T., & McCrae, R. R. (1985). *The NEO Personality Inventory: Manual.* Odessa, Fl: Psychological Assessment Resources.

Davidson, J. R. T., Hughes, D. C., George, L. K., & Blazer, D. G. (1996). The association of sexual assault and attempted suicide within the community. *Archives of General Psychiatry, 53,* 550–555.

Davis, J. A. & Smith, T. W. (1989–1991). *General social surveys, 1989–1991.* Chicago: National Opinion Research Center.

De Alarcon, R. (1964). Hypochondriasis and depression in the aged. *Gerontology Clinics, 6,* 266–277.

Dennis, M. S., & Lindesay, J. (1995). Suicide in the elderly—The United Kingdom perspective. *International Psychogeriatrics,* 7, 263–274.

DeVries, H. M., & Gallagher-Thompson, D. (1994). Managing psychological crises of older adults. In F. M. Dattilio & A. M. Freeman (Eds.), *Cognitive-behavior strategies in crisis intervention* (pp. 200–218). New York: Guilford.

Draper, B. (1994). Suicidal behaviour in the elderly. *International Journal of Geriatric Psychiatry, 9,* 655–661.

Draper, B. (1996). Attempted suicide in old age. *International Journal of Geriatric Psychiatry, 11,* 577–587.

Duberstein, P. R. (1995). Openness to experience and completed suicide across the second half of life. *International Psychogeriatrics, 7,* 183–198.

Duberstein, P. R., Conwell, Y., & Caine, E. D. (1994). Age differences in the personality characteristics of suicide completers: Preliminary findings from a psychological autopsy study. *Psychiatry: Interpersonal and Biological Processes, 57,* 213–224.

Fombonne, E. (1994). Increased rates of depression: Update of epidemiological findings and analytical problems. *Acta Psychiatrica Scandinavica, 90,* 145–156.

Frierson, R. L. (1991). Suicide attempts by the old and the very old. *Archives of Internal Medicine, 151,* 141–144.

Gallo, J. J., Anthony, J. C., & Muthén, B. G. (1994). Age differences in the symptoms of depression: A latent trait analysis. *Journal of Gerontology: Psychological Sciences, 49,* P251–P264.

Gardner, E. A., Bahn, A. K., & Mack, M. (1964). Suicide and psychiatric care in the aging. *Archives of General Psychiatry, 10,* 547–553.

Garrett, J. M., Harris, R. P., Norburn, J. K., Patrick, D. L., & Danis, M. (1993). Life-sustaining treatments during terminal illness: Who wants what? *Journal of General Internal Medicine, 8,* 361–368.

German, P. S., Shapiro, S., & Skinner, E. A. (1985). Mental health of the elderly: Use of health and mental health services. *Journal of the American Geriatric Society, 33,* 246–252.

Gibbs, N. (1993, May 31). R for death: The return of Dr. Death. *Time, 141*(22), 35–39.

Goldstein, R. B., Black, D. W., Nasrallah, A., & Winokur, G. (1991). The prediction of suicide: Sensitivity, specificity, and predictive value of a multivariate model applied to suicide among 1906 patients with effective disorders. *Archives of General Psychiatry, 48,* 418–422.

Griffith, E. E. H., Delgado, A. K., Foulks, E., Ruiz, P., Spiegel, J., Wintrob, R., & Yamamoto, J. (1989). *Group for the Advancement of Psychiatry: Report no. 128. Suicide and ethnicity in the United States.* New York: Brunner/Mazel.

Hay, D. P., & Hay, L. K. (1996). Biological treatment of severe late-life depression: Pharmacotherapy and electroconvulsive therapy. In G. J. Kennedy (Ed.), *Suicide and depression in late life: Critical issues in treatment, research, and public policy* (pp. 119–132). New York: Wiley.

Healthy People 2000 (1991). US Department of Health and Human Services, Public Health Service. DHHS Publication no. (PHS) 91-50212. Washington, DC: US Government Printing Office.

Heikkinen, M. E., Aro, H. M., Henriksson, M. M., Isometsä, E. T., Sarna, S. J., Kuoppasalmi, K. I., & Lönnqvist, J. K. (1994). Differences in recent life events between alcoholic and depressive nonalcoholic suicides. *Alcoholism: Clinical and Experimental Research, 18,* 1143–1149.

Heikkinen, M. E., & Lönnqvist, J. K. (1995). Recent life events in elderly suicide: A nationwide study in Finland. *International Psychogeriatrics, 7,* 287–300.

Henriksson, M. M., Isometsä, E. T., Hietanen, P. S., Aro, H.M. & Lönnqvist, J. K. (1995). Mental disorders in cancer suicides. *Journal of Affective Disorders, 36,* 11–20.

Hewitt, P. L., Flett, G. L., & Turnbull-Donovan, W. (1992). Perfectionism and suicide potential. *British Journal of Clinical Psychology, 31,* 181–190.

Hill, R. D., Gallagher, D., Thompson, L. W., & Ishida, T. (1988). Hopelessness as a measure of suicidal intent in the depressed elderly. *Psychology and Aging, 3,* 230–232.

Horton–Deutsch, S. L., Clark, D. C., & Farran, C. J. (1992). Chronic dyspnea and suicide in elderly men. *Hospital and Community Psychiatry, 43,* 1198–1203.

Humphry, D. (1991). *Final exit: The practicalities of self-deliverance and assisted suicide for the dying.* New York: Dell.

Humphry, D. (1992). Rational suicide among the elderly. *Suicide and Life Threatening Behavior, 22,* 125–129.

Kachur, S. P., Potter, L. B., James, S. P., Powell, K. E. (1995). *Suicide in the United States, 1980–1992. Violence Surveillance Summary Series, No. 1.* Atlanta, GA: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control.

Kaplan, M. L., Asnis, G. M., Lipschitz, D. S., & Chorney, P. (1995). Suicidal behavior and abuse in psychiatric outpatients. *Comprehensive Psychiatry, 36,* 229–235.

Kaplan, M. S., Adamek, M. E., & Geling, O. (1996). Sociodemographic predictors of firearm suicide among older white males. *The Gerontologist, 36,* 530–533.

Kaplan, M. S., Adamek, M. E., & Johnson, S. (1994). Trends in firearm suicide among older American males: 1979–1988. *The Gerontologist, 34,* 59–65.

Kaprio, J., Koskenvuo, M., & Rita, H. (1987). Mortality after bereavement: A prospective study of 95 647 widowed persons. *American Journal of Public Health, 77,* 283–287.

Keeler, E. B., Solomon, D. H., Beck, J. C., Mendenhall, R. C., & Kane, R. L. (1982). Effect of patient age on duration of medical encounters with physicians. *Medical Care, 20,* 1101–1108.

Kennedy, G. J. (1996). A brief antidepressant prescribing guide for the generalist. In G. J. Kennedy (Ed.), *Suicide and depression in late life: Critical issues in treatment, research and public policy* (pp. 133–138). New York: Wiley.

Kennedy, G. J., Kelman, H. R., Thomas, C., Wisniewski, W., Metz, H. & Bijur, P. E. (1989). Hierarchy of characteristics associated with depressive symptoms in an urban elderly sample. *American Journal of Psychiatry, 146,* 220–225.

Kessler, R. C., Brown, R. L., & Broman, C. L. (1981). Sex differences in psychiatric help-seeking: Evidence from four large-scale surveys. *Journal of Health and Social Behavior, 22,* 49–64.

Klerman, G. L. (1987). Clinical epidemiology of suicide. *Journal of Clinical Psychiatry, 48*(suppl.), 33–38.

Kliewer, E. V., & Ward, R. H. (1988). Convergence of immigrant suicide rates to those in the destination country. *American Journal of Epidemiology, 127,* 640–653.

Ko, S. M., & Kua, E. H. (1995). Ethnicity and elderly suicide in Singapore. *International Psychogeriatrics, 7,* 309–317.

Koenig, H. G. (1993). Legalizing physician-assisted suicide: Some thoughts and concerns. *The Journal of Family Practice, 37,* 171–179.

Kornstein, S. G., Schatzberg, A. F., Yonkers, K. A., Thase, M. E., Keitner, G. I., Ryan, C. E. & Schlager, D. (1995). Gender differences in presentation of chronic major depression. *Psychopharmacology Bulletin, 31,* 711–718.

Kreitman, N. (1976a). Age and parasuicide. *Psychological Medicine, 6,* 113–121.

Kreitman, N. (1976b). The coal gas story: United Kingdom suicide rates, 1960–1971. *British Journal of Preventive and Social Medicine, 30,* 86–93.

Lee, M., & Ganzini, L. (1992). Depression in the elderly: Effect on patient attitudes toward life-sustaining therapy. *Journal of the American Geriatrics Society, 40,* 983–988.

Lee, M., & Ganzini, L. (1994). The effect of recovery from depression on preferences for life-sustaining therapy in older patients. *Journal of Gerontology, 49,* M15–M21.

Leenaars, A. A., Maltsberger, J. T., & Neimeyer, R. A. (1994). *Treatment of suicidal people.* Washington, DC: Taylor & Francis.

Lester, D. (1993). The effectiveness of suicide prevention centers. *Suicide and Life-Threatening Behavior, 23,* 263–267.

Lester, D., & Tallmer, M. (Eds.) (1994). *Now I lay me down: Suicide in the elderly.* Philadelphia: Charles Press.

Li, G. (1995). The interaction effect of bereavement and sex on the risk of suicide in the elderly: An historical cohort study. *Social Science and Medicine, 40,* 825–828.

Lindesay, J. (1986). Trends in self-poisoning in the elderly 1974–1983. *International Journal of Geriatric Psychiatry, 1,* 37–43.

Linehan, M. M. (1986). Suicidal people: One population or two? *Annals of the New York Academy of Sciences, 487,* 16–33.

Livingston, G., Hawkins, A., Graham, N., Blizard, R., & Mann, A. H. (1990). The Gospel Oak study: Prevalence rates of dementia, depression and activity limitations among elderly residents in inner London. *Psychological Medicine, 20,* 137–146.

Lutzky, S. M., & Knight, B. G. (1994). Explaining gender differences in caregiver distress: The roles of emotional attentiveness and coping styles. *Psychology and Aging, 9,* 513–519.

Lyness, J. M., Conwell, Y., & Nelson, J. C. (1992). Suicide attempts in elderly psychiatric inpatients. *Journal of the American Geriatrics Society, 40,* 320–324.

Maes, M., Delange, J., Meltzer, H. Y., Scharpé, S., D'Hondt, P., & Cosyns, P. (1994). Lower degree of esterification of serum cholesterol in depression: Relevance for depression and suicide research. *Acta Psychiatrica Scandinavica, 90,* 252–258.

Maltsberger, J. T. (1991). Psychotherapy with older suicidal patients. *Journal of Geriatric Psychiatry, 24,* 217–234.

Manton, K. G., Blazer, D. G., & Woodbury, M. A. (1987). Suicide in middle age and later life: Sex and race specific life table and cohort analyses. *Journal of Gerontology, 42*(2) 219–227.

Marks, A. (1988–1989). Structural parameters of sex, race, age, and education and their influence on attitudes toward suicide. *Omega: Journal of Death and Dying, 19,* 327–336.

Markus, H., & Oyserman, D. (1989). Gender and thought: The role of the self-concept. In M. Crawford and M. Gentry (Eds.), *Gender and thought: Psychological* Perspectives (pp. 100–127). New York: Springer-Verlag.

Marshall, J. R., Burnett, W., & Brasure, J. (1983). On precipitating factors: Cancer as a cause of suicide. *Suicide and Life-Threatening Behavior, 13,* 15–27.

Marttunen, M. J., Aro, H. M., Henriksson, M. M., & Lönnqvist, J. K. (1991). Mental disorders in adolescent suicide: DSM-III-R Axes I and II diagnoses in suicides among 13- to 19-year olds in Finland. *Archives of General Psychiatry, 48,* 834–839.

McCall, P. L. (1991). Adolescent and elderly white male suicide trends: Evidence of changing well-being? *Journal of Gerontology: Social Sciences, 46,* S42–S51.

McCall, P. L. & Land, K. C. (1994). Trends in white male adolescent, young-adult, and elderly suicide: Are there common underlying structural factors? *Social Science Research, 23,* 57–81.

McIntosh, J. L. (1995). Suicide prevention in the elderly (Age 65–99). *Suicide and Life-Threatening Behavior, 25,* 180–192.

McIntosh, J. L., Pearson, J. L., & Lebowitz, B. D. (1997). Mental disorders of elderly men. In J. I. Kosberg & L. W. Kaye (Eds.), *Elderly men: Special problems and professional challenges* (pp. 193–215). New York: Springer

McIntosh, J. L., Santos, J. F., Hubbard, R. W., & Overholser, J. C. (1994). *Elder suicide: Research theory and treatment.* Washington, DC: American Psychological Association.

Meehan, P. J., Lamb, J. A., Saltzman, L. E., & O'Carroll, P. W. (1992). Attempted suicide among young adults: Progress toward a meaningful estimate of prevalence. *American Journal of Psychiatry, 149,* 41–44.

Meehan, P. J., Saltzman, L. E., & Sattin, R. W. (1991). Suicides among older United States residents: Epidemiologic characteristics and trends. *American Journal of Public Health, 81,* 1198–1200.

Merrill, J., & Owens, J. (1990). Age and attempted suicide. *Acta Psychiatrica Scandinavica, 82,* 385–388.

Montgomery, S. A. (1993). Suicide prevention and serotonergic drugs. *International Clinical Psychopharmacology, 8*(Suppl 2), 83–85.

Morgan, R. E., Palinkas, L. A., Barrett-Connor, E. L., & Wingard, D. L. (1993). Plasma cholesterol and depressive symptoms in older men. *Lancet, 341,* 75–79.

Mościcki, E. K. (1995). Epidemiology of suicide. *International Psychogeriatrics, 7,* 137–148.

Mościcki, E. K., O'Carroll, P., Rae, D. S., Roy, A., Locke, B. Z., & Regier, D. A. (1988). Suicide attempts in the Epidemiologic Catchment Area Study. *Yale Journal of Biology and Medicine, 61,* 259–268.

Muldoon, M. F., Manuck, S. B., & Matthews, K. A. (1990). Lowering cholesterol concentrations and mortality: A quantitative review of primary prevention trials. *British Medical Journal, 301,* 309–314.

Murphy, G. E., & Wetzel, R. D. (1980). Suicide risk by birth cohort in the United States, 1949 to 1974. *Archives of General Psychiatry, 37,* 519–523.

Murphy, G. E., Wetzel, R. D., Robins, E., & McEvoy, L. (1992). Multiple risk factors predict suicide in alcoholism. *Archives of General Psychiatry, 49,* 459–463.

Murphy, H. B. M. (1977). Migration, culture and mental health. *Psychological Medicine, 7,* 677–684.

Nathanson, C. A. (1977). Sex, illness and medical care: A review of data, theory and methods. *Social Science and Statistics, 11,* 13–25.

National Center for Health Statistics (January 18, 1996). Table 292A. Death rates for 282 selected causes, by 5-year age groups, color, and sex: United States, 1979–93.

Newman, S. C., & Dyck, R. J. (1988). On the age–period–cohort analysis of suicide rates. *Psychological Medicine, 18,* 677–681.

Niederehe, G., Cooley, S. G., & Teri, L. (1995). Research and training in clinical geropsychology: Advances and current opportunities. *The Clinical Psychologist, 48,* 37–44.

NIH Consensus Development Panel on Depression in Late Life. (1992). Diagnosis and treatment of depression in late life. *Journal of the American Medical Association, 268,* 1018–1024.

Nolen-Hoeksema, S. (1987). Sex differences in unipolar depression: Evidence and theory. *Psychological Bulletin, 101,* 259–282.

Nolen-Hoeksema, S. (1995). Epidemiology and theories of gender differences in unipolar depression, in M. V. Seeman (Ed.), *Gender and psychopathology* (pp. 63–87). Washington, DC: American Psychiatric Press.

Nolen-Hoeksema, S., Morrow, J., & Fredrickson, B. L. (1993). Response styles and the duration of episodes of depressed mood. *Journal of Abnormal Psychology, 102,* 20–28.

Nordentoft, M., Breum, L, Munck, L., Nordestgaard, A. G., Hunding, A., & Lawsen-Bjaeldager, P. A. (1993). High mortality by natural and unnatural causes: A 10 year follow up study of patients admitted to a poisoning treatment centre after suicide attempts. *British Medical Journal,* 306, 1637–1641.

Nowers, M. (1993). Deliberate self-harm in the elderly: A survey of one London borough. *International Journal of Geriatric Psychiatry, 8,* 609–614.

Nowers, M., & Irish, M. (1988). Trends in reported rates of suicide by self-poisoning in the elderly. *Journal of the Royal College of General Practitioners, 38,* 67–69.

O'Carroll, P. (1989). A consideration of the validity and reliability of suicide mortality data. *Suicide and Life Threatening Behavior, 19,* 1–16.

O'Neal, P., Robins, E., & Schmidt, E. H. (1956). A psychiatric study of attempted suicide in persons over sixty years of age. *Archives of Neurology and Psychiatry, 75,* 275–284.

Osgood, N. J. (1992). Environmental factors in suicide in long-term care facilities. *Suicide and Life-Threatening Behavior, 22,* 98–106.

Osgood, N. J. (1995). Assisted suicide and older people—A deadly combination: Ethical problems in permitting assisted suicide. *Issues in Law & Medicine, 10,* 415–435.

Osgood, N. J., & Brant, B. A. (1900). Suicidal behavior in long-term care facilities. *Suicide and Life-Threatening Behavior, 20,* 113–122.

Paffenbarger, R. S., Jr., Lee, I-M., & Leung, R. (1997). Physical activity and personal characteristics associated with depression and suicide in American college men. *Acta Psychiatrica Scandinavica, 377* (Suppl.), 16–22.

Pearson, J. L., & Conwell, Y. (1995). Suicide in late life: Challenges and opportunities for research. *International Psychogeriatrics, 7,* 131–136

Perez-Stable, E. J., Miranda, J., Muñoz, R., & Ying, Y. (1990). Depression in medical outpatients: Underrecognition and misdiagnosis. *Archives of Internal Medicine, 150,* 1083–1088.

Pierce, D. (1987). Deliberate self-harm in the elderly. *International Journal of Geriatric Psychiatry, 2,* 105–110.

Pokorny, A. D. (1983). Prediction of suicide in psychiatric patients: Report of a prospective study. *Archives of General Psychiatry, 40,* 249–257.

Posner, R. (1995). *Aging and old age*. Chicago: University of Chicago Press.

Quill, T. E. (1991). Death and dignity—a case of individualized decision making. *New England Journal of Medicine, 324,* 691–694.

Regier, D. A., Myers, J. K., Kramer, M., Robins, L. N., Blazer, D. G, Hough, R. L., Eaton, W. W., & Locke, B. Z. (1984). The NIMH Epidemiologic Catchment Area (ECA) Program: historical context, major objectives, and study population characteristics. *Archives of General Psychiatry, 41,* 934–941.

Rich, C. L., Young, D., & Fowler, R. C. (1986). San Diego suicide study: I. Young vs. old subjects. *Archives of General Psychiatry, 43,* 577–582.

Rich, C. L., Young, J. G., Fowler, R. C., Wagner, J., & Black, N. A. (1990). Guns and suicide: Possible effects of some specific legislation. *American Journal of Psychiatry, 147,* 342–346.

Richman, J. (1993). *Preventing elderly suicide: Overcoming personal despair, professional neglect, and social bias.* New York: Springer.

Richman, J. (1996). Psychotherapeutic approaches to the depressed and suicidal older person and family. In G. J. Kennedy (Ed.), *Suicide and depression in late life: Critical issues in treatment, research, and public policy* (pp. 103–117). New York: Wiley.

Rifai, A. H., George, C. J., Stack, J. A., Mann, J. J., & Reynolds, C. F. (1994). Hopelessness in suicide attempters after acute treatment of major depression in late life. *American Journal of Psychiatry, 151,* 1687– 1690.

Rifai, A. H., Reynolds, C. F., & Mann, J. J. (1992). Biology of elderly suicide. *Suicide and life-threatening behavior, 22,* 48–61.

Rihmer, Z. (1996). Strategies of suicide prevention: Focus on health care. *Journal of Affective Disorders, 39,* 83–91.

Robinson, D. G., & Spiker, D. G. (1985). Delusional depression: A one year follow-up. *Journal of Affective Disorders, 9,* 79–83.

Romanoski, A. J., Folstein, M. F., Nestadt, G., Chahal, R., Merchant, A., Brown, C. H., Gruenberg, E. M., & McHugh, P. R. (1992). The epidemiology of psychiatrist-ascertained depression and DSM-III depressive disorders. *Psychological Medicine, 22,* 629–655.

Roose, S. P., Glassman, A. H., Walsh, B. T., Woodring, S., & Vital-Herre, J. (1983). Depression, delusions, and suicide. *American Journal of Psychiatry, 140,* 1159–1162.

Rovner, B. W., Kafonek, S., Filipp, L., Lucas, M. J., & Folstein, M. F. (1986). Prevalence of mental illness in a community nursing home. *American Journal of Psychiatry, 143,* 1446–1449.

Runeson, B., & Beskow, J. (1991). Borderline personality disorder in young Swedish suicides. *Journal of Nervous and Mental Disease, 179,* 153–156.

Sackeim, H. A. (1994). Use of electroconvulsive therapy in late-life depression. In L. S. Schneider, C. F. Reynolds, B. D. Lebowitz, & A. J. Friedhoff (Eds.), *Diagnosis and treatment of depression in late life* (pp. 259–277). Washington, DC: American Psychiatric Press.

Sadavoy, J., & Fogel, B. (1992). Personality disorders in old age. In J. E. Birren, R. B. Sloane, & G. D. Cohen (Eds.), *Handbook of mental health and aging* (pp. 433–463). New York: Academic Press.

Sainsbury, P. (1955). *Suicide in London*. Maudsley Monographs no. 1. London: Chapman & Hall.

Salzman, C. (1994). Pharmacological treatment of depression in elderly patients. In L. S. Schneider, C. F.

Reynolds, B. D. Lebowitz, & A. J. Friedhoff (Eds.), *Diagnosis and treatment of depression in late life* (pp. 181–244). Washington, DC: American Psychiatric Association.

Schmid, H., Manjee, K., & Shah, T. (1994). On the distinction of suicide ideation versus attempt in elderly psychiatric inpatients. *Gerontologist, 34,* 332–339.

Schneider, L. (1996). Pharmacological considerations in the treatment of late-life depression. *Americal Journal of Geriatric Psychiatry, 4*(suppl.), S51–S65.

Schneider, L. S. (1996). Biological commonalities among aging, depression, and suicidal behavior. In G. J. Kennedy (Ed.), *Suicide and depression in late life: Critical issues in treatment, research and public policy* (pp. 39–50). New York: Wiley.

Seidlitz, L., Duberstein, P. R., Cox, C., & Conwell, Y. (1995). Attitudes of older people toward suicide and assisted suicide: An analysis of Gallup Poll findings. *Journal of the American Geriatrics Society, 43,* 993–998.

Shneidman, E. (1985). *Definition of suicide.* New York: Wiley.

Shneidman, E. (1986). Some essentials of suicide and some implications for response. In A. Roy (ed.), *Suicide* (pp. 1–16). Baltimore: Williams & Wilkins.

Shneidman, E. (1992). What do suicides have in common? A summary of the psychological approach. In B. Bongar (Ed.), *Suicide: Guidelines for assessment, management and treatment.* New York: Oxford University Press.

Shneidman, E. S., & Farberow, N. L. (1961). Statistical comparisons between attempted and committed suicides. In N. L. Farberow, & E. S. Shneidman (Eds.), *The cry for help* (pp. 19–47). New York: McGraw-Hill.

Shneidman, E. S., Farberow, N., & Litman, R. (1961). The suicide prevention center. In N. Farberow, & E. S. Shneidman (Eds.), *The cry for help* (pp. 6–18). New York: McGraw-Hill.

Sloan, J. H., Rivara, F. P., Reay, D., Ferris, J. A. J., & Kellerman, A. L. (1990). Firearm regulations and rates of suicide. *New England Journal of Medicine, 322,* 369–373.

Sorenson, S. B., & Golding, J. M. (1988). Prevalence of suicide attempts in a Mexican–American population: Prevention implications of immigration and cultural issues. *Suicide and Life-threatening Behavior, 18,* 322–333.

Stenback, A. (1980). Depression and suicidal behavior in old age. In J. E. Birren & R. B. Sloane, (Eds.), *Handbook of mental health and aging* (pp. 616–652). Englewood Cliffs, NJ: Prentice-Hall.

Stillion, J. M., & McDowell, E. E. (1996). *Suicide across the life span: Premature exits* (2nd ed.). Washington, DC: Taylor & Francis.

Strahan, G. W.(1991). *Mental illness in nursing homes: United States, 1985.* Vital Health Statistics 13 (105). DHHS Publication No. (PHS) 91-1766. Hyattsville, MD: National Center for Health Statistics.

Sturm, R., & Wells, K. B. (1995). How can care for depression become more cost-effective. *Journal of the American Medical Association, 273,* 51–58.

Suicide Among Older Persons—United States, 1980̆n1992. (January 12, 1996). *Morbidity and Mortality Weekly Report, 45*(1), 3–6.

Suicide and the elderly: A population at risk, Senate Hearing, July 30, 1996.

Sullivan, M. D., & Youngner, S. J. (1994). Depression, competence, and the right to refuse lifesaving medical treatment. *American Journal of Psychiatry, 151,* 971–978.

Takahashi, Y., Hirasawa, H., Koyama, K., Asakawa, O., Kido, M., Onse, H., Udagawa, M., Ishikawa, Y., & Uno, M. (1995). Suicide and aging in Japan: An examination of treated elderly attempters. *International Psychogeriatrics, 7,* 239–251.

US Bureau of the Census (1996). *65 + in the United States.* Current Population Reports, Special Studies, P23-(No. 190). Washington, DC: US Government Printing Office.

Vassilas, C. A., & Morgan, H. G. (1994). Elderly suicides' contact with their general practitioner before death. (Letter to the editor) *International Journal of Geriatric Psychiatry, 9,* 1008–1009.

Vieta, E., Nieto, E., Gastó,C., & Cirera, E. (1992). Serious suicide attempts in affective patients. *Journal of Affective Disorders, 24,* 147–152.

Whitlock, F. A. (1978). Suicide, cancer and depression. *British Journal of Psychiatry, 132,* 269–274.

Worthington, J., Fava, M., Davidson, K., Alper, J., Nierenberg, A. A., & Rosenbaum, J. F. (1995). Patterns of improvement in depressive symptoms with fluoxetine treatment. *Psychopharmacology Bulletin, 31,* 223–226.

Yap, P. M. (1958). Suicide in Hong Kong. *Journal of Mental Science, 104,* 266–301.

Younger, S. C., Clark, D. C., Oehmig-Lindroth, R., & Stein, R. J. (1990). Availability of knowledgeable informants for a psychological autopsy study of suicides committed by elderly people. *Journal of the American Geriatrics Society, 38,* 1169–1175.

Zweig, R. A., & Hinrichsen, G. A. (1993). Factors associated with suicide attempts by depressed older adults: A prospective study. *American Journal of Psychiatry, 150,* 1687–1692.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.12
Personality Disorders

DANIEL L. SEGAL and FREDERICK L. COOLIDGE
University of Colorado, Colorado Springs, CO, USA

7.12.1 INTRODUCTION 267
7.12.1.1 Evolution of the Diagnoses 267
7.12.1.1.1 Diagnostic and Statistical Manual of Mental Disorders 268
7.12.1.1.2 DSM-II 268
7.12.1.1.3 DSM-III 268
7.12.1.1.4 DSM-III-R 268
7.12.1.1.5 DSM-IV 269
7.12.1.2 Age and the DSM 269
7.12.1.3 Dimensional vs. Categorical Classification 269
7.12.2 ETIOLOGY 270
7.12.2.1 Rutter's Theory of Temperaments, Personality, and Personality Disorders 270
7.12.2.2 A Cognitive Theory of Personality Disorders 271
7.12.2.3 An Evolutionary Theory of Personality Disorders 272
7.12.2.4 Cloninger's Biosocial Theory and Personality Disorders 272
7.12.2.5 Relevance of Personality Disorder Theories to the Elderly 273
7.12.3 EPIDEMIOLOGY 274
7.12.4 DIAGNOSIS 275
7.12.4.1 General Issues 275
7.12.4.2 Personality Disorder Diagnosis and the Elderly 276
7.12.5 ASSESSMENT 277
7.12.5.1 Self-report Personality Inventories 277
7.12.5.1.1 Millon Clinical Multiaxial Inventory-III 278
7.12.5.1.2 Personality Diagnostic Questionnaire-4 278
7.12.5.1.3 Coolidge Axis II Inventory 278
7.12.5.1.4 Application to older adults 279
7.12.5.2 Structured Clinical Interviews 279
7.12.5.2.1 Structured Interview for DSM-IV *Personality* 280
7.12.5.2.2 International Personality Disorder Examination 281
7.12.5.2.3 Structured Clinical Interview for DSM-IV *Axis II Personality Disorders* 282
7.12.5.2.4 Application to older adults 282
7.12.6 TREATMENT 283
7.12.7 SUMMARY 285
7.12.8 REFERENCES 286

7.12.1 INTRODUCTION

7.12.1.1 Evolution of the Diagnoses

Personality disorders (PDs) have been defined as long-standing and long-lasting maladaptive behaviors that seriously impair an individual's ability to perform adequately in a variety of settings including social, occupational, educational, and interpersonal. They have also been called character disorders, and one classic hallmark of the diagnosis has been that the patients are typically unaware or unperturbed about their disorder. Although anecdotal and literary descriptions of individuals who may have had PDs have been with us for thousands of years, the official diagnosis of PDs as a unique and separate diagnostic entity did not officially appear until 1952 with the publication by the American Psychiatric Association (APA) of the *Diagnostic and Statistical Manual of Mental Disorders (DSM).*

7.12.1.1.1 Diagnostic and Statistical Manual of Mental Disorders

The *Diagnostic and Statistical Manual of Mental Disorders (DSM-I*; APA, 1952) contained seven major diagnostic categories including one named "Personality disorders." PDs were defined as developmental defects or pathological trends in the personality, with little or no sense of subjective distress. They were thought of as lifelong patterns of behavior and not likely to go into remission. Of the 12 PDs in *DSM-I,* four were "cardinal" types, (inadequate, schizoid, cyclothymic, and paranoid) indicating that individuals were rarely, if ever, altered by therapy; four were personality trait disturbances (emotionally unstable, passive–aggressive, compulsive personality, and an "other" personality trait disorder); and finally, four were types of sociopathic personality disturbance (antisocial, dyssocial, sexual deviation, and addiction).

The *DSM-I* descriptions for all 12 main types of PD consisted of, at most, two paragraphs and, most typically, four or five sentences. Although these types provided a major evolution in the diagnosis of PDs, psychological tests for the assessment of PDs lagged for more than 25 years. The *DSM* was widely used and no rival system existed in the USA. The international World Health Organization had produced the only worldwide alternative to the *DSM* and that was the *International Classification of Diseases (ICD),* which included a section on mental disorders. The eighth edition *(ICD-8)* was approved in 1966 and published in 1968. The second edition of the *DSM (DSM-II)* was published in 1968, and was then aligned with the *ICD-8*.

7.12.1.1.2 DSM-II

The *DSM-II* (APA, 1968) listed 10 specific types of PD. Their general description was very brief (two sentences only): "This group of disorders is characterized by deeply ingrained maladaptive patterns of behavior that are perceptibly different in quality from psychotic and neurotic symptoms. Generally, these are lifelong patterns, often recognizable by the time of adolescence or earlier." The 10 specific PDs were as follows: paranoid, cyclothymic, schizoid, explosive, obsessive compulsive, hysterical, asthenic, antisocial, passive–aggressive, and inadequate. One subtle change between *DSM-I* and *DSM-II* was the simple addition that PDs are recognizable by adolescence or earlier; *DSM-I* had mentioned that PDs tended to be lifelong, but there was no mention of when the onset occurred. Both the *DSM-I* and *DSM-II* noted the chronicity but neither noted any changes in PDs as a differential function of lifespan, or type of PD as a function of age.

7.12.1.1.3 DSM-III

The diagnosis of PDs took a quantum leap with the publication of *DSM-III* (APA, 1980). Its diagnostic, paradigmatic center shifted from a psychoanalytic to a more behavioral perspective. It also became multiaxial: major clinical syndromes and symptoms were to be coded on "Axis I" while PDs were to be coded on "Axis II." Also, a specific list of numbered criteria was presented for each PD, and patients had to meet a specified minimum number of criteria before receiving a diagnosis. *DSM-III* also mentioned aging as a factor. It stated "The manifestations of Personality Disorders are generally recognizable by adolescence or earlier and continue throughout most of adult life, though they often become less obvious in middle or old age" (p. 305). For the first time, in the history of the *DSM*, it was hypothesized that PDs may diminish with age.

Another interesting aspect of PD diagnosis in *DSM-III* was that they were grouped into three clusters. The odd or eccentric cluster comprised paranoid, schizoid, and schizotypal types. The dramatic, emotional, or erratic cluster contained histrionic, narcissistic, antisocial, and borderline types. The anxious or fearful cluster was composed of avoidant, dependent, compulsive, and passive–aggressive types.

Overall, *DSM-III* listed 11 specific PDs. Only four of these had remained essentially unchanged from the earlier editions. These were the paranoid, schizoid, antisocial, and passive–aggressive types. Two PDs had name changes (histrionic was formerly hysterical, and compulsive was formerly obsessive compulsive). Four specific disorders were eliminated: cyclothymic, inadequate, explosive, and asthenic. Five new PDs were also added since *DSM-II*: schizotypal, narcissistic, borderline, avoidant, and dependent.

7.12.1.1.4 DSM-III-R

A revised edition of *DSM-III* appeared only seven years later (APA, 1987). The 11 PDs remained, and again they were coded on axis II.

The odd or eccentric cluster of PDs from *DSM-III* officially became "Cluster A" in *DSM-III-R* and still comprised the paranoid, schizoid, and schizotypal types. There were also changes to the criteria for determining each type, although each PD diagnosis still consisted of a group of polythetic criteria. The dramatic, emotional, or erratic cluster became "Cluster B" and still consisted of the histrionic, narcissistic, antisocial, and borderline types. The anxious or fearful cluster became "Cluster C" and still consisted of the avoidant, dependent, and passive–aggressive types. The compulsive PD of *DSM-III* remained in Cluster C but was renamed the obsessive-compulsive PD. That change was highly ironic since in *DSM-I* it was the compulsive PD, in *DSM-II* it was the obsessive-compulsive PD, and in *DSM-III* it was again the compulsive PD. *DSM-III-R* also included two new and potentially controversial PDs that were placed in Appendix A for diagnostic categories needing further study. These were the sadistic and self-defeating types.

In terms of aging and developmental issues, *DSM-III-R* essentially repeated the same hypotheses as *DSM-III*, that is, PDs are generally recognizable by adolescence or earlier, continue through most of adult life, and become less prominent with age. These hypotheses were far too simplistic, however, and raised many more questions than they answered. For example, did all PDs diminish with increasing age? This certainly does not appear to be the case, at least intuitively, for the schizoid and avoidant PDs. Also, why do some PDs diminish with age? Is it simply an artifact of activity level or sociability, or is the phenomenon more complicated, that is, does the abnormal behavior diminishment associated with aging reflect underlying changes in neural substrate or chemical neurotransmitter functioning in the brain?

7.12.1.1.5 DSM-IV

The *DSM* was revised and published again after a further seven year interval (APA, 1994). This edition contained only 10 PDs on Axis II. The passive–aggressive PD was removed from Cluster C and was placed in Appendix B (for further study). An alternative name, negativistic PD, was given to the passive–aggressive PD. A new PD appeared in Appendix B, the depressive PD. The sadistic and self-defeating PDs were dropped completely from the diagnostic manual. The current list of PDs in *DSM-IV* is as follows: Cluster A—the odd or eccentric (paranoid, schizoid, schizotypal); Cluster B—the dramatic, emotional, or erratic (antisocial, borderline, histrionic, narcissistic); Cluster C—the anxious or fearful (avoidant, dependent, obsessive-compulsive); and Appendix B—criteria sets and axes provided for further study (depressive, passive–aggressive [negativistic]).

7.12.1.2 Age and the *DSM*

DSM-IV was the first version of the manual to include a specific section (along with cultural and gender topics) on the developmental issue of aging in the diagnosis of PDs. Again it recorded that a PD is usually recognized by adolescence or early adult life, and, by definition, the onset of a PD must occur before middle age for an official diagnosis to be made although, of course, an actual diagnosis might not be given until later adult life. Also noted in *DSM-IV* is that a PD may be exacerbated by social upheaval, such as the loss of a spouse or a job, although a clear change in personality in middle age or later adult life would more likely be due to a medical condition or substance-related disorder. Since, and again by definition, PDs are "enduring patterns of thinking, feeling and behaving" (APA, 1994, p. 632) they were also hypothesized to be consistent and stable throughout adulthood. However, the *DSM-IV* also noted that the antisocial and borderline PDs "tend to become less evident or to remit with age, whereas this appears to be less true for some other types (e.g., obsessive compulsive and schizotypal PDs)" (p. 632). Empirical research, however, in support or against these latter specific hypotheses has been scarce.

7.12.1.3 Dimensional vs. Categorical Classification

The diagnostic philosophy contained within *DSM-IV* is that PDs are distinct clinical categories or syndromes. Theoretically, these clinical syndromes form discrete entities of abnormal behavior patterns, and again, theoretically, they would be unrelated to normal personality dimensions, for example, the extraversion–introversion dimension. Currently, a more attractive and parsimonious theory is the dimensional diagnostic approach to PDs in which the PDs are viewed as maladaptive variations of normal personal traits. Thus, in the dimensional classification approach, normal and abnormal personality behaviors are simply points along a behavioral continuum. Typically, the extreme points at either end of a continuum might be reflective of a PD. For example, an extreme extrovert might be characteristic of the histrionic PD while an extreme introvert might be characteristic of a schizoid PD.

In the late twentieth century, dimensional trait research approaches have focused on finding the number of factors minimally required to describe an individual's personality, and whether those same dimensions can describe both normal and abnormal behavior. Eysenck (1978) originally proposed a two-factor model, extroversion and neuroticism, and then expanded his model to include a third factor, psychoticism. His model clearly covers the abnormal personality, although perhaps at the expense of normal personality constructs.

The most famous model of personality is that of Carl Jung from 1924 (1971). He conceptualized the personality as being fourfold: (i) one's preference to focus on the inner or outer world, (ii) how one perceives or acquires information, (iii) how one makes decisions, and (iv) how one orients one's self towards the outer world. This model was operationalized in a psychometric test, the Myers–Briggs Type Indicator (Myers, 1962), with four bipolar dimensions: (i) extraversion–introversion, (ii) sensing–intuiting, (iii) thinking–feeling, and (iv) judging–perceiving. This model, however, appears to be biased oppositely to Eysenck's model. The Myers–Briggs model appears to be more useful in explaining the normal personality rather than the abnormal one.

Norman (1963) developed and organized a pool of 2800 personality descriptors into sets such as states, traits, and roles. Based on his work, McCrae and Costa (1986, 1989) began a prodigious flow of research into a five-factor model describing both normal and abnormal personality; specifically, this included neuroticism, extraversion, openness, agreeableness, and conscientiousness. However, their model has been criticized for lacking diagnostic utility with regard to the *DSM-IV* PDs, for being unclear as to whether they were describing five, six, or even seven factors to account for PDs (Coolidge et al., 1994), for a lack of empirical research to support their claim that their model accounts for abnormal behavior, and the overgeneralization from statistical techniques (Millon & Davis, 1996). Furthermore, Kernberg (1996) wrote: "To develop factorial profiles for each personality disorder on the basis of those five factors has an eerie quality of unreality for the experienced clinician" (p. 107).

Cloninger (1987) originally proposed a three-factor model of personality based on three brain neurotransmitters: dopamine, serotonin, and norepinephrine. Later, he expanded his model to seven factors, operationalizing the measurement of his model with the Temperament and Character Inventory (TCI; Cloninger 1986, 1987; Cloninger, Przybeck, Svrakic, & Wetzel, 1994). This theory will be discussed further later in this chapter. By contrast, Cattell identified 16 source traits from data that he identified from surface traits, and he operationalized his theoretical personality structure into the Sixteen Personality Factor Questionnaire (16PF; Cattell, Saunders, & Stice, 1950). The 16PF has been a popular instrument in personality research, although more popular in nonclinical, rather than clinical, settings.

Thus, although *DSM-IV* approaches clinical diagnoses from a categorical perspective, most empirical personality research of abnormal behavior, particularly PDs, has entailed the use of dimensional models, typically ranging from 3 to 16 core factors or personality traits. However, little research has been devoted to how aging might affect these core traits and concomitant behaviors.

7.12.2 ETIOLOGY

Enduring, inflexible, and pervasive patterns of maladaptive behavior, coupled with an early onset, suggest that PDs may have an inheritable component. Thus, most conceptions about the etiology of PDs are based upon both biological and environmental contributions, although the theories vary in the emphasis placed upon these two components. Four major etiological theories, reflecting these ideas, will now be reviewed.

7.12.2.1 Rutter's Theory of Temperaments, Personality, and Personality Disorders

Rutter (1987) noted that temperaments are hypothesized to be biologically inherited predispositions for behaving in particular ways, and they emerge early in life. He specifically hypothesized, that temperaments were limited. According to Rutter, there are not a particularly large number of temperaments influencing behavior; such temperaments are nonmotivational and noncognitive. The search for the core number of temperaments may parallel the search for the core number of personality traits or factors, with the number varying from as few as three to as many as 16. There is a fairly substantial body of empirical research in support of heritability of temperaments and their enduring influence on behavior (e.g., McGue, Bacon, & Lykken, 1993; Plomin, Chipuer, & Loehlin, 1990; Thomas & Chess, 1977). Relatively undisputed is the hypothesis that the extraversion–introversion dimensional trait is reflective of a prominent temperament. Emotionality and activity levels have also received strong support as temperaments.

According to Rutter's theory, this limited number of temperaments helps to form what comes to be recognized as the individual's unique personality, although the personality is much broader than a simple collection of temperamental traits. Furthermore, Rutter proposes that as the individual interacts with family and society, these temperamental features may become reinforced by the social environment, thus producing the long-standing, enduring patterns of behavior that are characteristic of adult personalities. However, when particular temperamental features appear in the extreme or provoke unusual responses from the individual's environment, then PDs may result. The temperamental basis of a PD, along with its reinforcement history, therefore accounts for its enduring and inflexible nature.

Rutter's theory is confined to speculations about early childhood through adulthood, but there is little mention of issues related to later adulthood and the aged. Nevertheless, his theory can be used to generate a few hypotheses. Rutter argued that several longitudinal studies have shown "heterotypic" continuity, meaning that the phenotypic quality of the behavior changes over time while individual differences remain stable. Thus, early childhood passivity is predictive of noncompetitiveness in later adult life (Kagan & Moss, 1962). With respect to PDs, therefore, it is possible that the elderly should perhaps be evaluated according to different criteria or, at the least, the PD criteria should be examined for changes as a function of middle and later adulthood.

7.12.2.2 A Cognitive Theory of Personality Disorders

Pretzer and Beck (1996) have described difficulties in the psychotherapeutic treatment of patients with PDs, and they argue that even cognitive–behavioral treatments may prove counterproductive or ineffective if the therapist fails to recognize the specific nature of the PD and fails to modify treatment accordingly. In terms of the etiology of PDs from a cognitive perspective it is assumed that PDs develop phenomenologically, that is, through the person's view of their world. Thus, an individuals' perception of stimuli interacts with their prior beliefs and assumptions. An individual may react consciously and logically resulting in the observed emotional and behavioral responses to the perceived stimuli. However, prior misbeliefs (delusions) and false assumptions about the world may lead to inappropriate or maladaptive emotional responses and behavior. An individual may also commit "systematic errors in reasoning" which Pretzer and Beck labeled "cognitive distortions" that also result in inappropriate and maladaptive responses. They noted the prominence of at least 12 common cognitive distortions, such as viewing events in terms of mutually exclusive categories (dichotomous thinking) or focusing on one aspect of a complex situation to the exclusion of others (selective abstraction).

According to the theory, individuals are also likely to respond unconsciously or automatically to perceived stimuli. These ways of acting may have once been adaptive in an evolutionary sense. Beck (1992) has termed them "primeval strategies," and further, Pretzer and Beck (1996) have proposed that "human social change has outstripped physiological evolution [such that] we may well have also retained some biases in information processing that now tend to be maladaptive" (pp. 44–45). Thus, they argue that our inherited predispositions poorly fit some of the demands of modern society. For example, an earlier adaptive primeval strategy of predatory behavior may lead to the development of the antisocial PD, the earlier help-eliciting behavior may now result in the dependent PD, and earlier competitive behavior may now result in the narcissistic PD.

Nevertheless, Pretzer and Beck (1996) have argued that PD phenomena can be understood and treated without regard to inherited or inborn defects, and that inherited predispositions towards the development of PDs has not been clearly established. Moreover, their contribution appears neither simple nor straightforward. Overall, cognitive theory tends to emphasize that interactions between predispositions, the environment, and significant life events shape development rather than predispositions alone, and it is not external events that shape the personality but an individuals perception of those events that forms an individual's thinking and consequent behavior.

Although Pretzer and Beck's cognitive theory has a developmental perspective, it emphasizes early childhood influences upon the adult personality. Even though they hypothesize that traumatic experiences, either singular or recurrent, may alter an individual's personality over time, they do not address the influence of traumatic experiences upon the later adult personality. In fact, it could be easily hypothesized that 80 years of traumatic experiences should far outweigh a singular early traumatic experience. However, dramatic adult personality change is highly unlikely except for clear physiological reasons (e.g., brain disease, toxins, etc.). Thus, this aspect of their theory neglects personality issues of middle and late adulthood.

7.12.2.3 An Evolutionary Theory of Personality Disorders

Millon and Davis (1996) have proposed a highly complex and sophisticated evolutionary theory of PDs that attempts to link both normal and abnormal behavior to both phylogenetic and ontogenetic processes. Their theory explains why early experiences play a more profound role in the development of the personality than later experiences, drawing upon diverse sources of evidence such as Freud's writings about the influence of early childhood on the later adult personality and upon ethologists' observations of critical imprinting periods in maturational development.

Central to Millon and Davis' evolutionary theory are the hypotheses that four evolutionarily based "bipolarities" interact with four stages of early neuropsychological development to produce stable and consistent patterns of behavior that can either be adaptive or maladaptive. These bipolarities are pain–pleasure, active–passive, self–other, and thinking–feeling. The four bipolarities also reflect the parallel evolutionary processes of existence, adaptation, replication, and abstraction, respectively. The four parallel neuropsychological stages are sensory–attachment, sensorimotor–autonomy, pubertal–gender identity, and intracortical–initiative, respectively. Although their theory is comprehensive and far-reaching, they do not address PD issues of the elderly. Nor, despite the magnitude of their theory, does it offer much heuristically for the differential waxing or waning of PDs in the elderly.

7.12.2.4 Cloninger's Biosocial Theory and Personality Disorders

Cloninger (1987) has proposed a biosocial theory which asserts that phenotypic personality traits are the product of the interaction of genetically determined brain systems and environmental influences. He offered an operationalization of his theory in the TCI (Cloninger et al., 1994). This comprises seven major dimensions. There are four, purportedly genetically independent, trait dimensions: novelty seeking, harm avoidance, reward dependence, and persistence. The first three dimensions form the physiological component of the biosocial theory, and are identified with three primary monamine neurotransmitters: dopamine, serotonin, and norepinephrine. The first three traits are linked with these three neurotransmitters, respectively. Cloninger hypothesized that these three form neural temperaments which serve as discrete systems for the activation, inhibition, and maintenance of behavior in response to novelty, danger, and reward. The fourth character dimension, persistence, was not thought to be associated with a specific neurotransmitter. The decision to classify it as an autonomous factor in the model was based upon empirical statistical factor analyses.

The TCI also assesses three character dimensions: self-directedness, cooperativeness, and self-transcendence. These dimensions became later additions to the TCI to account for behavior not otherwise described by the initial three physiological dimensions and in order to determine whether a PD was present or not. The four temperament dimensions were designed to delineate clinical subtypes of personality but it appeared that they did not discriminate those individuals with PDs from those without PDs.

Another interesting feature of Cloninger's theory is that it attempts to link two primary components of human memory, procedural and declarative, to the formation of the seven trait and character dimensions. His theory postulates that the development of the four trait dimensions involves preconceptual responses to perceptual stimuli. Learned behavior patterns require the stimulus input of perceptual experience, but do not require conscious recall to replicate the behavior, which is the definition of procedural memory. The declarative memory system processes experience that is represented as words, images, or symbols. Once encoded, concept-based memories can be consciously retrieved, declared verbally, and willfully acted upon. Cloninger hypothesized that his three character dimensions reflected, in behaviorally expressed form, the information processed by the declarative memory system.

To account for PDs, Cloninger proposed that particular combinations of trait dimensions were predictive of various PDs. For example, he thought that the combination of high novelty seeking (high dopamine), low harm avoidance (low serotonin), and high reward dependence (high norepinephrine) was indicative of the histrionic PD, while the antisocial PD would be high on novelty seeking, and low on both harm avoidance and reward dependence. Presumably, all of the PDs would be low on the three character dimensions of self-directedness, cooperativeness, and selftranscendence.

Because the TCI is a relatively new test, there is a current lack of empirical evidence in support of Cloninger's biosocial theory. However, the previous version of the TCI (with only three trait dimensions) did receive modest but not overwhelming support in a few studies (e.g., Pfohl, Black, Noyes, Kelly, & Blum, 1990; Ruegg et al., 1993). Coolidge and Griego's (1996) study of the TCI with a new PD test (Coolidge Axis II Inventory; Coolidge & Merwin, 1992) provided moderate support for Cloninger's theory in a sample of 164 college students.

To date, Cloninger's biosocial theory makes no predictions about PDs as a function of the aging process. However, his biosocial theory is certainly fruitful for speculation. For example, there is substantial evidence (e.g., Morgan, May, & Finch, 1987) that there are normal age-related changes in the dopamine systems in humans and other mammals. There is an age-related, but not disease-related, loss in dopamine and its metabolites, a loss of dopaminergic neurons, and a decrease in particular types of dopamine receptors. Severe dopamine losses (greater than 75%) are characteristic of Parkinson's disease. Interestingly and anecdotally, these age-related losses of dopamine may, in part, account for the observation that some elderly persons do not typically engage in high novelty seeking behaviors. It appears that there are few or no age-related changes in the serotonin systems in humans. Although there appears to be a modest reduction in particular serotonin receptor sites, serotonin levels appear to be stable across the lifespan. Major serotonin level declines and serotonin receptor site losses are, however, indicative of Alzheimer's disease. Evidence regarding age-associated decreases in norepinephrine is scant.

In summary, it appears that Cloninger's biosocial theory is provocative and capable of generating a plethora of hypotheses concerning both pathological (PDs and brain syndromes) and nonpathological behaviors in the elderly. However, neurotransmitter systems in the brain are extremely complex and interactive, and it is doubtful whether singular descriptions of human behavior (e.g., novelty seeking) are based upon singular neurotransmitting systems. Furthermore, it is popularly suspected that many more neurotransmitters than have presently been discovered are yet to be discovered. Because of this, Cloninger's theory, at least in regard to PDs, may only be speciously complex and sophisticated.

7.12.2.5 Relevance of Personality Disorder Theories to the Elderly

It appears that many of the popular theories of PDs directly ignore issues related to aging and the elderly. Most of the theories, however, are useful in generating hypotheses about PDs in the elderly, regarding etiology, diagnosis, and treatment. For example, Rutter's hypothesis of heterotypic continuity in the lifespan is useful diagnostically, and it appears that comparisons of younger and older adults in terms of the *DSM* criteria for PDs may be useful. Pretzer and Beck's cognitive theory is provocative in that it appears to minimize the influences of traumatic experiences in the elderly. It appears that it would be fruitful to examine how, why, or if traumatic experiences vary from childhood to later adulthood. Also, it might be useful to determine whether and to what extent traumatic experiences differentially affect older people, that is, what characteristics do individuals possess that make them vulnerable or invulnerable to traumatic experience? Are there ways of inoculating the vulnerable? Are the elderly, as a group, more or less vulnerable? Also, Pretzer and Beck's primeval strategies are fascinating as a basis for the development of differences among the various PDs. Are there PDs unique to the elderly? If there are, do they have a primeval strategic basis?

Millon and Davis' evolutionary theory also does not address aging related issues. However, the four bipolarities (pain–pleasure, active–passive, self–other, and thinking–feeling) should be examined with respect to aging. Are there differences between younger and older persons? Also, evolutionary theories, in general, tend to ignore the latter end of the lifespan. After adaptation and replication, are humans simply devoid of evolutionary purpose? Do dramatic behaviors such as memory loss in the elderly have evolutionary meaning? What evolutionary purpose does personality stability possess?

Cloninger's biosocial theory can also be used to generate hypotheses about personality stability and change, and PDs in the elderly. To what extent do the chemical neurotransmitters in the brain affect behavior differentially in the elderly? How do declines in the neurotransmitter systems affect behavior in the elderly? Can classes of behaviors be directly linked to neurotransmitter systems in the elderly? If PDs are created by varying levels of neurotransmitters, can chemical interventions be developed to "cure" PDs?

This section on the etiology of PDs is far from an exhaustive review. However, a number of trends clearly emerge. The popular theories concerning PD attempt to integrate the biological and evolutionary factors with environmental influences. Most of these theories virtually, if not totally, ignore issues related to aging. It appears, however, that one can generate a host of provocative hypotheses from the theories with regard to aging issues, and psychologists should be encouraged to do so.

7.12.3 EPIDEMIOLOGY

Despite the growing body of literature about PDs in the elderly, there is a striking paucity of epidemiological and prevalence data for these interesting and challenging disorders in the older population. Moreover, reported rates for these conditions are far more inconsistent and controversial than estimates for more widely studied disorders, such as schizophrenia, cognitive impairment, or major depression. Such inconsistencies illustrate the difficulty clinicians and researchers have in accurately diagnosing PDs in the elderly.

Overall, surveys conducted in the 1980s show that prevalence rates for PDs in the older adult community range from as low as 2.8% to as high as 11% (Cohen, 1990). Prevalence figures are even more varied when mental health settings are considered. For example, Mezzich, Fabrega, Coffman, and Glavin (1987) conducted a retrospective chart review, and found that 5.1% of a sample of mixed inpatient and outpatient elders ($N = 494$) manifested PDs. Casey and Schrodt (1989) reviewed charts of 100 consecutive admissions to a geriatric psychiatry unit and found that 7% received a diagnosis of PD as determined by psychiatrist rating. In the largest retrospective study to date, Fogel and Westlake (1990) ascertained prevalence of characterological disorder in 2332 older inpatients suffering from major depression. Again, diagnoses were assigned by attending psychiatrists, with results showing a 15.8% PD rate. In a related study, Kunik et al. (1993) evaluated depressed elderly inpatients and found that 37 of 154 (24%) met criteria for a PD based on consensus diagnosis among psychiatrists. Similarly, Thompson, Gallagher, and Czirr (1988) showed that 33% of depressed elderly clients participating in a psychotherapy outcome study were diagnosed with a PD. More recently, Molinari and Marmion (1995) used a structured PD scale to assess PDs in elderly patients with mood disorders, revealing a 63% prevalence rate.

Limitations of many older epidemiological studies include their use of retrospective chart assessments, reliance on clinical judgment in assigning diagnoses, and the absence of reliability or validity checks of PD diagnoses. Although the National Institute of Mental Health Epidemiological Catchment Area programs were successful at providing quality epidemiological data for many common mental disorders, the PDs (with the exception of antisocial PD) were neglected in these surveys. Clearly, large-scale cross-sectional and longitudinal studies that employ standardized, objective, and validated diagnostic instruments, such as structured interviews, are warranted to fully investigate the prevalence of PD in the aged.

Other factors have also hampered epidemiological studies of PDs. For example, estimates of PDs have fluctuated greatly depending on the type of sample (e.g., general population, medical patients, psychiatric patients), method of diagnosis (e.g., clinician diagnosis, self-report personality inventory, structured interview), and type of research design (e.g., retrospective chart review, crosssectional, longitudinal). Furthermore, different diagnostic criteria have been employed in different studies, including research diagnostic criteria (Spitzer, Endicott, & Robins, 1978) and various versions of the *DSM* system. Spar and LaRue (1990) suggest that "Few data have been gathered on Axis II disorders in elderly patients, in part because most of the personality disorders in *DSM-III-R* have not been defined long enough for thorough studies to take place" (p. 180). Indeed, diagnostic systems prior to *DSM-III* did not provide well-operationalized or specific criteria for each particular PD, and reliability of diagnosis was notoriously poor. Since early investigations of PDs in the elderly utilized classification systems that are dissimilar from contemporary *DSM-III-R* and *DSM-IV* conceptualizations (Kroessler, 1990), meaningful comparison of prior and more current data is difficult. Despite gaps in the epidemiological research base, it can be concluded that PDs in the elderly are not uncommon, and that many clients so disordered are presently on the caseloads of diverse mental health practitioners.

Another epidemiological issue concerns differential rates for PDs among older males and females. According to the *DSM-IV*, borderline, dependent, and histrionic PDs are diagnosed more frequently in females, while antisocial and obsessive-compulsive PDs, and, to a lesser extent paranoid PD, are more common in males. However, it is unclear to what extent these differences apply to older individuals who typically were not adequately represented in the research studies from which conclusions about sex differences were drawn. Hard data are also lacking about the stability or change in PD symptoms as one enters old age. As noted earlier, *DSM-IV* describes a PD as a lifelong, stable pattern of thinking, feeling, and behaving. However, it also suggests that some disorders (e.g., antisocial and borderline) tend to remit with age, while others may exacerbate (e.g., obsessive-compulsive and schizotypal). However, these hypotheses should be empirically tested.

Tyrer (1988) investigated empirically the relationship between PD and increasing age and suggested that "mature" forms of PD such as obsessive-compulsive, paranoid, schizoid, and schizotypal remain stable with age, while more "immature" or flamboyant disorders such as antisocial, borderline, histrionic, and narcissistic are likely to decrease with age. This observation certainly has some anecdotal support, culled from clinicians with vast experience in the diagnosis and assessment of older persons. For some elderly clients, their current personality and social functioning appears markedly improved compared to the interpersonal chaos and occupational disaster characteristic of their earlier years. On the other hand, some older clients show exacerbation in their personality styles. In an interesting study, Coolidge, Burns, Nathan, and Mull (1992) examined these questions about course of PDs by directly comparing PD rates in younger and older subjects. Their results showed that the elderly were significantly more obsessive-compulsive and schizoid than their younger counterparts (based on mean T-scores), and were equal on avoidant and dependent PD scales. Although this cross-sectional study is intriguing, definitive research about the long-term course of individual PDs is lacking. Longitudinal studies will provide the clearest evidence about this issue, but these studies are obviously time-consuming, expensive, and difficult to carry out. However, information is sorely needed about this important conceptual and clinical issue if we are to better understand our elderly clients.

7.12.4 DIAGNOSIS

7.12.4.1 General Issues

The awareness and importance of PDs as a diagnostic category took a major leap in 1980 with the publication of *DSM-III*, with its innovative multiaxial diagnostic system. Clinical syndromes (located on Axis I) were separated from long-standing PDs (located on Axis II). Since that time, PD diagnosis has been fraught with controversy and difficulty, with the psychiatric community expressing grave (and justified) concerns about the apparent lack of reliability for these disorders. In general, the reliability of measurement refers to consistency, replicability, or stability. In the context of psychiatric diagnosis, reliability often refers to the extent of agreement between two raters concerning presence or absence of a psychiatric diagnosis. If different diagnosticians cannot agree about particular diagnoses, then those diagnostic categories are, at best, of limited value and, at worst, virtually meaningless. A full discussion of reliability issues and procedures is provided by Segal (1997) and Segal, Hersen, and Van Hasselt (1994). In fact, classic investigations of inter-rater reliability for diagnosis of *DSM-III* PDs produced results that were unacceptably poor (see Mellsop, Varghese, Joshua, & Hicks, 1982; Spitzer, Forman, & Nee, 1979). Moreover, these values were considerably lower than values for major axis I disorders, such as depression and schizophrenia.

The pervasive lack of agreement between clinicians and researchers when diagnosing PDs is particularly troublesome, given that reliability is a necessary, but insufficient, prerequisite for validity (Grove, 1987). Fortunately, there recently has been much improvement in this area with the publication of *DSM-III-R* and *DSM-IV*. In an effort to enhance reliability (and validity) of diagnosis, these classification systems have provided more specific and behaviorally-operationalized trait criteria for each particular PD. Following the development of better diagnostic criteria, several structured diagnostic interviews (discussed later) were created and refined; this further enhanced the ability to diagnose PDs consistently.

Whereas reliability refers to reproducibility and stability, the validity of psychiatric diagnosis refers to the extent to which a diagnosis actually measures what it purports to measure. The issue here is one of diagnostic accuracy and meaningfulness. Unfortunately, the validity of PD diagnosis is much harder to gauge and document accurately; thus, it represents a formidable challenge for clinicians and researchers alike. A common strategy employed is to compare PD diagnoses generated by a structured interview with diagnoses from another source, such as a clinical interview, chart review, or other testing procedure. However, such practice is far from ideal, mainly because these other diagnostic measures often suffer from questionable reliability and validity in the first place. Indeed, a major problem encountered by clinical researchers interested in the validity of PDs concerns the lack of a solid criterion diagnosis, a sort of "gold standard," against which to compare an interview generated diagnosis. This inadequacy severely restricts conclusions drawn from many validity studies.

The reader should be aware that the current *DSM-IV* Axis II system, while considerably better than its predecessors, is far from ideal. PD diagnoses generally are still less reliable than those of the major Axis I conditions, and most Axis II disorders show extensive overlapping, with individuals typically meeting criteria for several PDs at one point in time (Oldham et al., 1992). This extensive pattern of comorbidity and the fact that "pure" types are rarely encountered has fueled debate and discussion among clinical researchers as to the validity of categorical diagnosis of PDs in the first place (Frances, 1982; Oldham et al.). As cogently noted by Oldham et al. most studies

> have found that at least 50% of patients have two or more coexisting PDs. It may be that a dimensional approach is better suited for Axis II disorders than a categorical one that results in an unwieldy list of presumably comorbid diagnoses to describe "character pathology" in a given patient. (p. 213)

Usefulness of the three clusters (odd–eccentric, dramatic–erratic, anxious–fearful) proposed by the *DSM* system has also been criticized (e.g., Morey, 1988). Unquestionably, accurate diagnosis of PDs in any age group is still widely viewed as a complex and challenging task.

7.12.4.2 Personality Disorder Diagnosis and the Elderly

In the elderly population, diagnosis of PDs is complicated by several additional factors. First, diagnosis is affected by the longtime, widely held (albeit erroneous) belief that personality dysfunction either was not present or was not a significant mental health problem for older people. With this diagnostic mindset, PD symptoms may be minimized or overlooked altogether by the clinician. In fact, some researchers have suggested that diverse service providers to older individuals are biased against diagnosing PDs or defining problems as characterological (Kroessler, 1990; Rose, Soares, & Joseph, 1993; Rosowsky & Gurian, 1992). According to Rose et al. (1993) the chronicity, rigidity, and resistance to intervention attributed to individuals carrying a PD diagnosis often make the client seem "hopeless" and service providers feel "helpless." Professionals may lack appropriate knowledge about the manifestations of PDs in older adults, and consequently may fail to address this component in their assessment procedures.

Second, misdiagnoses often occur because individuals with PDs often present clinically with comorbid symptoms of major axis I disorders, most notably depression and anxiety, which are more easily recognized by both clients and service providers. Typically, clients view such acute symptoms as distressing, while lacking insight into their maladaptive personality style. In some cases, the blatant psychiatric symptoms may be consequences of the chronic interpersonal and occupational difficulties experienced by those suffering from dysfunctional personalities, but this cause–effect relationship is difficult to evaluate. Moreover, many clinicians simply are more experienced with the diagnosis and treatment of major mental illnesses, not PDs.

A third obstacle to the accurate diagnosis of PDs in older adults is that some signs of personality disturbance (e.g., avoidance, dependency, social withdrawal, odd thinking, irritability, eccentric behavior, or emotional lability) are wrongly perceived by some young and old individuals as "normal" in the aged population, or they are dismissed as being attributed to "old age." Such faulty assumptions almost certainly preclude appropriate assessment and diagnosis.

A fourth impediment to diagnosis is that *DSM-IV* criteria for some PDs have limited relevance or applicability to aged clients given their unique physical, cognitive, and social circumstances. Just as somatic problems may confound diagnosis of Axis I disorders in older individuals (e.g., depression, or panic disorder), Kroessler (1990) notes similar difficulties with the diagnosis of PDs in the aged. For example, he points out that many symptoms of a dysfunctional personality (e.g., aggressiveness, impulsiveness, brawling, child-beating, recklessness, and promiscuity) require physical activity. These criteria may not be applicable to some older adults for purely biological reasons in terms of the expected declines in energy, strength, and mobility due to normal aging. Despite this, however, such symptoms do appear in some elderly persons. For borderline PD in particular, many classic symptoms of this condition in younger persons (e.g., impulsive acting-out, sexual promiscuity, self-mutilating behaviors, and substance abuse) are likely to decrease in intensity with increasing age and physical/social limitations, although the underlying psychodynamics remain present (Fogel & Westlake, 1990; Rosowsky & Gurian, 1991, 1992). Reflecting on the fact that older clients meet fewer criteria for borderline PD than they did as younger clients, Rosowsky and Gurian (1991) propose that this difference may actually be due to inadequate criteria for the elderly rather than the lack of the disorder. We suggest that these "age-biases" in diagnostic criteria also apply to other personality types. Alternatively, some symptoms of apparent personality dysfunction in older individuals may actually be due to age-related social and biological factors, rather than a disordered personality. For example: social withdrawal (characteristic of schizoid and avoidant PD) may increase due to physical disabilities and transportation problems experienced by some elderly people; perceptions of exploitation by others (characteristic of paranoid PD) may be attributable to realistic concerns about crimes and scams directed at the elderly; and impulsivity, emotional lability, poor social judgment, and uncontrollable rage (characteristic of borderline and histrionic PDs) may actually be due to frontal lobe damage caused by vascular accidents in some older persons.

Fifth, a major requirement for the diagnosis of any PD is the experience of clinically significant distress or impairment in social or occupational functioning (*DSM-IV*, criterion C). However, as individuals age, they work less frequently and have smaller social networks due to sensory and physical disabilities, as well as the cumulative loss of family and friends through death or relocation. Thus, the requirement of social/occupational impairment may preclude accurate PD diagnosis in older adults whose social world and work opportunities are much different from their younger counterparts upon whom many taxonomies are based. As age-related behavior changes may either mimic or mask PD symptoms we concur with Kroessler (1990) and also argue for inclusion of age-associated criteria in the diagnostic guidelines.

Several case reports have appeared in the literature documenting the unique diagnostic and intervention challenges faced by mental health professionals. Specifically, a vivid case description has been reported for borderline PD (Siegel & Small, 1986). More recently, case examples of antisocial (Howard, Bandyopadhyay, & Cook, 1992), histrionic, borderline, and avoidant PDs (Rose et al., 1993) in old age have been presented. It is obvious, however, that systematic research is necessary to clarify clinical manifestations and develop clearer profiles of each PD in the aged. As noted previously, longitudinal studies following personality-disordered individuals into later life are needed to assess changes over time in symptom patterns and personality characteristics of older individuals. It is probable that some symptoms are robust, some appear in muted format in the aged, and some are irrelevant. Given the diagnostic problems highlighted in this account, the need for elder-specific diagnostic criteria is underscored. For example, the number needed to assign a diagnosis (e.g., five out of eight for histrionic PD) may be modified (usually lowered) to improve diagnostic accuracy and application to the elderly. This issue can be determined empirically by validity studies assessing cut-off rates as the criteria requirements are adjusted. When evaluating and diagnosing PDs in the elderly, clinicians are also encouraged to recognize and be sensitive to potential age-biases in the *DSM-IV* system. In conclusion, a significant ramification of these diagnostic difficulties is that determination of the most appropriate form of intervention is often obfuscated. Treatment is often not indicated if the behavior is considered a "normal" concomitant of the aging process; or, when intervention is applied, it solely targets the blatant anxiety or depressive symptoms, while the etiologically significant underlying personality dysfunction is minimized, or even worse, completely ignored.

7.12.5 ASSESSMENT

Accurate assessment of PDs is a critical component of patient care given the fact that most clinicians working with older individuals will have some clients with PDs on their caseload. As diagnostic criteria for PDs has become more refined in recent years, the assessment of PDs has also flourished. Several instruments are available to the clinician, and these can be divided into two broad categories: self-report inventories and structured clinical interviews. Both assessment approaches are widely applied in clinical and research contexts, and both have their own advantages and disadvantages. On the one hand, structured interviews generally are more reliable and allow consideration of important observational data, but they require extensive training and experience on the part of the interviewer, and can be lengthy to conduct especially with severely disordered clients. On the other hand, self-report personality inventories are easy to administer and take less time, but in some cases tend to overpathologize and are prone to deviant test-taking response behaviors. This chapter now focuses on the description and development of several popular self-report inventories and structured interviews, as well as the application of these devices to older adults.

7.12.5.1 Self-report Personality Inventories

Instruments to be reviewed in this chapter include the *Millon Clinical Multiaxial Inventory-III (MCMI-III*; Millon, 1994), the *Personality Diagnostic Questionnaire-4 (PDQ-4*; Hyler, 1994), and the *Coolidge Axis II Inventory (CATI*; Coolidge, 1984; Coolidge & Merwin, 1992). These instruments are among the most commonly utilized tools for measuring *DSM-IV* PDs. Other self-report instruments have been developed to measure PDs, including the Wisconsin Personality Disorders Inventory (Klein et al., 1993). Also, Morey, Waugh, and

Blashfield (1985) have derived PD scales from the original Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1983). The MMPI, and its revision, the MMPI-II (Hathaway & McKinley, 1989), are not true measures of *DSM*-based PDs. Rather, they evaluate personality traits, symptom profiles, and axis I disorders. A full review of all measures is beyond the scope of this chapter. Therefore, we will restrict our reviews to *MCMI-III, PDQ-4,* and *CATI.*

7.12.5.1.1 Millon Clinical Multiaxial Inventory-III

The original *MCMI* (Millon, 1983) has undergone two revisions, that is, *MCMI-II* (Millon, 1987) and *MCMI-III* (Millon, 1994). *MCMI-III* is a 175-item, true-false, self-report inventory specifically designed to assess PDs. This instrument is widely used in research and clinical activities, and its taxonomy is "aligned" with *DSM-IV*. The *MCMI-III* includes 14 personality pattern scales (Axis-II related) including: schizoid, avoidant, depressive, dependent, histrionic, narcissistic, antisocial, aggressive (sadistic), compulsive, passive–aggressive (negativistic), self-defeating, schizotypal, borderline, and paranoid. Additionally, 10 symptom scales (Axis-I related) are provided, covering moderate clinical presentations (e.g., anxiety, dysthymia, alcohol dependence) and severe symptoms (e.g., thought disorder, major depression). It has three validity scales designed to detect random or confused response patterns and defensiveness. The *MCMI* is designed to be given only to those suspected of having a PD, and thus, should not be used for differential diagnosis in nonclinical populations. The *MCMI* has been the focus of numerous publications since its development in 1967, and earlier versions have been shown to have good psychometric properties, although it is said to be more reliable in diagnosing among PDs than in differentiating between clinical syndromes (Millon, 1992). Psychometric properties of the *MCMI-III* are currently being evaluated.

7.12.5.1.2 Personality Diagnostic Questionnaire-4

The original *PDQ* was revamped to match changes in the *DSM-III-R* system, and was renamed the *PDQ-R* (Hyler & Rieder, 1987). A newer version of *PDQ-R* has been created (*PDQ-4*; Hyler, 1994) and this is now consistent with *DSM-IV*. The *PDQ-4* assesses the 10 standard PDs in the *DSM-IV*. It contains 85 self-report, true-false items and takes about 30 minutes to complete. It has two validity scales: one detects underreporting of psychopathology and the other identifies those who are responding randomly or inconsistently. A supplementary clinical significance scale was designed for *PDQ-4*; to control for an excess of false positives. Also, a summed "total" score of all pathological items representing an overall level of personality dysfunction can be generated. Interestingly, an ancillary version, *PDQ-4+*, can be used to evaluate passive–aggressive and depressive PDs as listed in the appendix of *DSM-IV*. A computer-assisted version is available for administration and scoring. The *PDQ-R* has been demonstrated to be an efficient instrument for screening personality disorders according to *DSM-III-R* (Hyler, Skodol, Kellman, Oldham, & Rosnick, 1990; Hyler, Skodol, Oldham, Kellman, & Doidge, 1992), and has been found to have high sensitivity and moderate specificity for most axis II disorders (Hyler et al., 1990; Hyler et al., 1992). Most recently, Trull and Larson (1994) have conducted a validity study of the *PDQ-R*, comparing it to a structured clinical interview. They concluded that the *PDQ-R* was overly sensitive to psychopathology (false positives), although it was suggested that the *PDQ-R* might be helpful in ruling out PDs. The new clinical significance scale of the *PDQ-4* may help in this regard.

7.12.5.1.3 Coolidge Axis II Inventory

The *CATI* (Coolidge, 1984; Coolidge & Merwin, 1992) is a 225-item, self-report and significant-other report inventory developed and revised to assess PDs according to the diagnostic criteria of *DSM-IV*. The inventory currently contains 14 PD scales, retaining the *DSM-III-R* sadistic and self-defeating PDs, and offering the new *DSM-IV* depressive PD. The items on each PD scale were designed to cover all of the criteria listed in *DSM-IV* (or *DSM-III-R*). The *CATI* also provides assessment for 13 Axis I scales including depression, anxiety, schizophrenia, and post-traumatic stress disorder. The new *DSM-IV,* Personality Change Due to a General Medical Condition is also evaluated, with scores provided for each of the five subtypes: apathy, disinhibition, emotional lability, aggression, and paranoia. One unique feature of *CATI* is that it has an 18 item neuropsychological dysfunction scale for assessing symptoms of brain disease, trauma, and dysfunction, with three subscales assessing language dysfunction, memory and concentration difficulties, and somatic complaints related to brain dysfunction. The *CATI* also includes a 16 item executive functions scale with three subscales measuring poor planning, decision-making difficulty, and task incompletion. The test–retest reliability (one week intervening) for the *CATI* is good (mean r = 0.90; range = 0.78–0.97), and internal consistency is moderate (median r = 0.76; range = 0.66–0.87). With regard to validity, a 50% concordance rate with clinicians' diagnoses (based on *DSM-III-R* criteria) for 24 personality-disordered patients was found, and the *CATI* showed moderate concurrent validity with the *MCMI-II* (median r = 0.58; range = 0.10–0.87) with the same patients (Coolidge & Merwin, 1992).

7.12.5.1.4 Application to older adults

Application of the *MCMI* to older psychiatrically impaired adults has been explored in only one study. Hyer and Harrison (1986) administered the inventory to 60 elderly psychiatric inpatients. Results showed that rates of dependent and avoidant PDs were highest in this older group, a finding that mirrors results from younger age groups. In contrast, histrionic, narcissistic, and antisocial PDs had the lowest rates. These disorders were characterized as higher intensity personality styles. Notably, Hyer and Harrison suggested that these high energy personality types tend to become less prominent with increasing age.

Literature regarding the *CATI* and older adults is also limited. Coolidge and colleagues (1992) administered the *CATI* to normal younger (N = 573; mean age 24.0 years) and older (N = 36; mean age 69.4 years) individuals and compared results. They found that, for Axis I syndromes, older adults were significantly less anxious, but reported more signs of brain dysfunction. No group differences were reported on the depression scale. As for Axis II, the elderly scored significantly lower on antisocial, borderline, histrionic, narcissistic, paranoid, passive–aggressive, schizotypal, sadistic, and self-defeating scales than their younger counterparts. By contrast, the older adult sample was significantly more obsessive–compulsive and schizoid. No group differences were obtained for avoidant or dependent PD scales. Coolidge and colleagues (1992) noted that, in general, the older sample scored low on questions associated with impulsivity and endorsed items consistent with restricted emotional expressiveness. In this study, age differences on the *CATI* were assessed in nonclinical, normal subjects. To date, age comparisons have not been examined in psychiatric populations. We found no studies of the *PDQ* with older persons.

The dearth of studies applying the *PDQ, MCMI*, and *CATI* to older populations represents a major gap in the research literature. Recently, however, the concurrent validity of the *MCMI-II* and *CATI* has been evaluated in a pilot study (Silberman, Roth, Segal, & Burns, in press) targeting chronically mentally ill elderly inpatients (N = 30; mean age 63.3 years). Patients were diagnosed by a psychiatrist as having a chronic mental illness (e.g., schizophrenia, bipolar disorder). The *MCMI-II* and *CATI* were concurrently administered in counterbalanced fashion and the data submitted to correlational analyses. The median correlation coefficient (raw score sums for the *CATI* and *MCMI-II*) for the 13 PD scales was 0.55 (range - 0.13 for schizoid to 0.88 for borderline), which was similar to the moderate results obtained by Coolidge and Merwin (1992) with a younger and less severe sample (24 psychiatric outpatients, mean age 38.0 years).

7.12.5.2 Structured Clinical Interviews

Since the 1960s, several structured diagnostic interviews have been developed for clinical, research, and training applications. Notably, structured interviews have been devised to assist in diagnosis of all standard Axis II PDs and all major Axis I clinical syndromes. Prominent interviews for Axis I conditions include the Schedule for Affective Disorders and Schizophrenia (Endicott & Spitzer, 1978), the *Diagnostic Interview Schedule* (original *DIS*; Robins, Helzer, Croughan, & Ratcliff, 1981; *DIS-IV*; Robins, Cottler, Bucholz, & Compton, 1996), and the Structured Clinical Interview for *DSM-IV* (First, Spitzer, Gibbon, & Williams, 1995a). (See recent review by Segal & Falk [in press] for further description of interviews for Axis I disorders.) Principal interviews for Axis II PDs include the Structured Interview for *DSM-IV* Personality (SIDP-IV), the International Personality Disorder Examination (IPDE), and the Structured Clinical Interview for *DSM-IV* Axis II Personality Disorders (SCID-II). Each PD interview is closely aligned to the *DSM* system, and subsequently has been revamped to match refinements in that classification scheme as it has evolved. Overall, structured interviews for PDs have improved clinical and research endeavors by providing a more standardized, scientific, and quantitative approach to the evaluation of these disorders. Indeed, these instruments have strongly contributed to the advancement in diagnostic clarity and precision for PDs. For each instrument, the purpose and construction will be briefly discussed, followed by some thoughts about the clinical applications of these instruments with older adults. PDs covered by the *DSM-IV* version of each instrument are presented in Table 1.

Table 1 *DSM-IV* personality disorders diagnoses covered by the SCID-II, IPDE, and SIDP-IV.

	Structured interview		
Personality disorders	*SCID-II*	*IPDE*	*SIDP-IV*
Cluster A			
Paranoid	X	X	X
Schizoid	X	X	X
Schizotypal	X	X	X
Cluster B			
Antisocial	X	X	X
Borderline	X	X	X
Histrionic	X	X	X
Narcissistic	X	X	X
Cluster C			
Avoidant	X	X	X
Dependent	X	X	X
Obsessive-compulsive	X	X	X
Other personality disorders			
Mixed or PD NOS	X	X	X
Sadistic	no	no	no
Self-defeating	no	no	X
Depressive	X	no	X
Negativistic	X	no	X

7.12.5.2.1 ***Structured Interview for*** **DSM-IV** ***Personality***

The SIDP-IV (Pfohl, Blum, & Zimmerman, 1995) is a comprehensive, semistructured diagnostic interview for *DSM-IV* PDs. The SIDP-IV covers 14 axis II diagnoses, including the 10 standard PDs, mixed PD, as well as self-defeating, depressive, and negativistic PDs which are proposed for "further study" in *DSM-IV*. The current edition of this popular instrument (SIDP-IV) has its roots in earlier versions that date back over a decade ago. In fact, the original Structured Interview for *DSM-III* Personality Disorders (SIDP; Stangl, Pfohl, Zimmerman, Bowers, & Corenthal, 1985) was the first structured interview designed to diagnose the full range of *DSM-III* PDs. Soon after, the Structured Interview for *DSM-III-R* Personality Disorders (SIDP-R; Pfohl, Blum, Zimmerman, & Stangl, 1989) was published to match changes in *DSM-III-R*.

Prior to the SIDP-IV structured interview, a full evaluation of current mental state or axis I conditions is required (Pfohl et al., 1995). This is not surprising given that self-report of enduring personality characteristics can be seriously compromised in a respondent experiencing acute distress. Indeed, the aim of all personality assessment measures is to rate the respondents typical, habitual, and lifelong personal functioning rather than acute or temporary states. For the most part, each SIDP-IV question corresponds to a unique *DSM-IV* axis II criterion. To help guide the interviewer, the specific *DSM* criterion associated with each question is provided on the form for easy reference. Items are not presented in a disorder-by-disorder fashion; rather, items are grouped according to 10 "topical sections" that reflect different dimension of personality functioning. Examples of these groupings include: interests and activities, close relationships, emotions, self-perception, and social conformity (Pfohl et al.). Questions typically are open-ended to encourage respondents to talk about their *usual* behaviors and long-term functioning. Based on patient responses, each criterion is rated on a scale with four anchor points, ranging from not present to strongly present and debilitating. The SIDP-IV requires that a trait be prominent for most of the last five years to be considered a part of the respondents' personality. This "five year rule" helps ensure that the particular personality characteristic is "stable and of long duration" as required by the "general diagnostic criteria" for a personality disorder described in *DSM-IV*.

Like the IPDE and SCID-II, significant clinical judgment is required to properly administer the SIDP-IV as interviewers are expected to ask additional questions to clarify patient responses when necessary. Also data are not limited to self-report; rather, chart records, and significant others such as relatives and friends who know the patient well should be consulted when available. Such collateral information is particularly prized when evaluating personality disordered individuals, who tend to lack insight into their own maladaptive personality traits and distort facts about their strengths and limitations. Moreover, informants can also provide diagnostic data that can help resolve the state/trait distinction about specific criterion behaviors. Overall, the SIDP typically requires 60 to 90 minutes for the patient interview, 20 minutes for interview of significant informants, and approximately 20 minutes to tabulate the responses and make diagnostic conclusions.

Reliability (inter-rater and internal consistency) of SIDP diagnoses has been evaluated in many studies (e.g., Stangl et al., 1985; Nazikian, Rudd, Edwards, & Jackson, 1990; Hogg, Jackson, Rudd, & Edwards, 1990; Jackson, Gazis, Rudd, & Edwards, 1991) with generally favorable results. Convergent validity has been assessed in several investigations in which SIDP diagnoses were compared to diagnoses from self-report Axis II instruments, such as the *MCMI*, and results typically have been modest at best (Hogg et al., 1990; Jackson et al., 1991; Miller, Streiner, & Parkinson, 1992; Nazikian et al., 1990; Turley, Bates, Edwards, & Jackson, 1992). It should be noted, however, that these studies may not be a fair test of the validity of the SIDP due to the differing assessment methods (interview vs. self-report) and the questionable validity of self-report PD measures.

7.12.5.2.2 International Personality Disorder Examination

The IPDE (Loranger et al., 1994) is an extensive, semistructured diagnostic interview for PDs. Although the 1994 version is compatible with the *DSM-III-R* and *ICD-10*, a version aligned with *DSM-IV* criteria has been produced by the World Health Organization (WHO; in press). The IPDE is an outgrowth of the Personality Disorder Examination (PDE) which was developed and refined over a five year period from 1983 to 1988 (Loranger, Hirschfeld, Sartorius, & Regier, 1991). Impetus for the development of the IPDE came from WHO and the US Alcohol, Drug Abuse, and Mental Health Administration (ADAMHA) in their joint effort aimed at producing standardized assessment instruments to measure mental disorders on an worldwide basis. The IPDE was selected by the WHO/ADAMHA for use as an instrument capable of assessing PDs in different languages, cultures, and countries. As such, the IPDE allows for diagnosis of PDs according to both the *DSM-IV* and *ICD-10* systems.

Items reflecting PD criteria according to the *DSM* and *ICD* are grouped into six thematic headings: work, self, interpersonal relationships, affects, reality testing, and impulse control (Loranger et al., 1991). Like the SIDP-IV, disorders are not covered on a one-by-one basis. IPDE sections typically begin with open-ended prompts to encourage respondents to elaborate about themselves in a less structured fashion. Then, specific questions are asked to evaluate each *DSM* and *ICD* criterion. There are separate modules for the *DSM-IV* and *ICD-10* classification systems. This division serves to reduce administration time, as interviewers typically follow one system or the other. Still, the IPDE typically requires 3–4 hours.

The IPDE requires that a trait be prominent during the last five years to be considered a part of the respondent's personality. Information about age of onset of particular behaviors is also explored to determine if a late-onset diagnosis is appropriate. When a respondent acknowledges a particular trait, interviewers follow-up by asking for examples and anecdotes to clarify the trait or behavior, gauge impact of the trait on the person's functioning, and fully substantiate the rating. Such probing requires significant clinical judgment and knowledge about each criterion on the part of interviewers. Moreover, some items are rated based on observation of the respondents behavior during the session, and this too requires a certain level of clinical expertise. To supplement self-report, interview of informants is encouraged. All items are rated on a three-point scale with the following definitions: "O" indicates that the behavior or trait is denied or within normal limits, "1" refers to present to an accentuated degree, and "2" signifies pathological/meets criterion (Loranger et al., 1994). The final output of the IPDE is extensive, including presence or absence of each criterion, number of criteria met for each disorder, a dimensional score (sum of individual scores for each criteria for each disorder), and a categorical diagnosis (definite, probable, or negative) (Loranger et al., 1994).

Like the SIDP-IV, inter-rater reliability of the IPDE (or PDE) generally is good, from report to report (Loranger et al., 1994; Loranger, Susman, Oldham, & Russakoff, 1987; Standage & Ladha, 1988). Not surprisingly, convergent validity coefficients typically have been weaker, ranging from 0.14 to 0.66 (Skodol, Oldham, Rosnick, Kellman, & Hyler, 1991; also see Hunt & Andrews, 1992; Pilkonis, Heape, Ruddy, & Serrao, 1991; Soldz, Budman, Demby, & Merry, 1993).

7.12.5.2.3 Structured Clinical Interview for DSM-IV *Axis II Personality Disorders*

To complement the Axis I version of the SCID, Spitzer and colleagues endeavored to develop a semistructured interview that would facilitate improved diagnosis of Axis II PDs. Their efforts resulted in publication of the SCID-II (version 1.0; Spitzer, Williams, Gibbon, & First, 1990), which targeted 11 PDs according to the then current *DSM-III-R* system. In 1994, the SCID-II was updated (SCID-II, version 2.0; First, Spitzer, Gibbon, Williams, & Benjamin, 1994) so that questions correspond to the new diagnostic criteria in *DSM-IV*. In this *DSM-IV* version, the 10 standard PDs are evaluated, along with the proposed depressive and negativistic PDs, while self-defeating and sadistic PDs have been deleted.

The SCID-II is usually employed in conjunction with the SCID-I, which typically is administered prior to personality assessment. This is encouraged so that the respondents present psychological state can be considered when judging accuracy of self-reported personality traits (First, Spitzer, Gibbon, & Williams, 1995b). A unique feature of the SCID-II is that it includes a 117 item, self-report, forced choice, yes/no screening component (called the Personality Questionnaire) that is completed prior to the interview portion and requires about 20 minutes to complete. The purpose of this screening measure is to reduce overall administration time. In the questionnaire, initial diagnostic items for each PD are presented, so that later sections of the structured assessment can be omitted if particular personality traits are denied.

During the structured interview component, only positively endorsed screening responses are further pursued to ascertain whether the symptom criteria is actually at clinically significant levels. Each SCID-II question taps a unique PD criterion, and the actual *DSM-IV* diagnostic criteria are printed on the interview page for easy review. Responses are coded on a three-point system (i.e., 1 = absent, 2 = subthreshold, 3 = threshold, ? = inadequate information). Each disorder is assessed completely and diagnoses are made before proceeding to the next disorder (First et al., 1995b). The modular format permits researchers and clinicians to tailor the SCID-II to their specific needs and reduce administration time. SCID-II administrators are expected to use their clinical judgment to clarify responses, gently challenge inconsistencies, and ask for additional information as required to accurately assess each disorder. Like the SIDP-IV and IPDE, collection of diagnostic information from ancillary sources is permitted. Complete administration of the SCID-II typically takes under 50 minutes (First et al., 1995b), which is substantially shorter than other structured measures. Training requirements and interviewer qualifications are similar to that of the other structured interviews for PDs.

With regard to psychometric properties of the SCID-II, an impressive large-scale multisite test–retest reliability project has recently been completed (First et al., 1995) with reasonably good results (overall weighted kappa = 0.53; range = 0.34–0.74). Other studies have reported even higher inter-rater reliability values for the SCID-II (Arntz et al.,1992; Brooks, Baltazar, McDowell, Munjack, & Bruns, 1991; Fogelson, Nuechterlein, Asarnow, Subotnik, & Talovic, 1991; Malow, West, Williams, & Sutker, 1989; O'Boyle & Self, 1990). Like the SIDP-IV and IPDE, less empirical attention has been devoted to validity of SCID-II diagnoses, and the studies that have been conducted generally produce only fair validity coefficients across different methodologies and criterion diagnoses used (O'Boyle & Self, 1990; Skodol et al., 1991; Skodol, Rosnick, Kellman, Oldham, & Hyler, 1988).

It is apparent that application of structured interviews for PDs over the last 15 years has greatly facilitated the task of psychiatric diagnosis for these conditions. Reliability values for the structured interviews generally indicate modest to excellent agreement with diverse patient populations and experimental designs. Most importantly, these values greatly surpass results from early investigations with unstructured interviews (Mellsop et al., 1982). Unfortunately, although reliability rates are now more acceptable for most PDs, it remains to firmly establish the validity. Most studies in this area have focused on basic reliability, differentiation of separate PDs from each other, and description of these conditions. Firm establishment of construct and predictive validity for the PDs is lacking in the literature. As such, investigation in this area should be the major focus of future research attention.

7.12.5.2.4 Application to older adults

As previously noted, each major structured interview has been subjected to extensive evaluation by clinical researchers. Unfortunately, the literature regarding specific application of these instruments with the elderly is sparse, although several clinical studies have been reported. For example, Abrams, Alexopoulos, and Young (1987) administered the

PDE to 21 older adults who had recovered from depression and 15 without a history of depression. The purpose of this study was to investigate the relationship between personality traits and late-life depression. Only two study participants met criteria for a PD and no reliability estimates were reported. In a more recent study also employing the PDE in a sample of 30 recovered depressed older adults, inter-rater reliability was high for all PDs, except histrionic (Abrams, Rosendahl, Card, & Alexopoulos, 1994). In a similar investigation, Schneider, Zemansky, Bender, and Sloane (1992) used the SCID-II to diagnose PDs in a group of older adults with history of major depression, although no reliability data were reported. Thompson et al. (1988) obtained Axis II formulations by using the SIDP in a sample of 79 depressed older adults, and found high inter-rater reliability. These clinical reports provide some evidence that the structured interviews can be successfully applied to older adults. However, larger reliability and validity studies are warranted especially with more diverse elder populations (e.g., medical patients, minorities, psychiatric inpatients).

Several additional concerns about the application of structured interviews with older adults should be noted. A common feature among all three structured interviews is that respondents are encouraged to describe their "typical or usual" functioning, rather than their often altered personality functioning during times of acute psychiatric illness. As such, during the course of the structured interview, respondents are sometimes asked to describe their behaviors at different points in time, possibly spanning several years. Unfortunately, such distinctions may be difficult for some elderly clients to make, as memory problems can preclude accurate self-report about behavior patterns many years ago. Another problem with the utilization of structured interviews with older clients is the considerable length of time required for completion. With young clients, administration time for the various interviews typically ranges from 1.5 to 3.5 hours. However, administration may take considerably longer with older persons. One reason for this is that older individuals simply have more extensive and complicated histories to review. Some older persons also require several "breaks" that may not be needed by younger individuals. Indeed, many older adults fatigue more quickly, and this must be monitored as responses will become less accurate if the client becomes inattentive or unmotivated. Older adults typically will perform their best when given frequent opportunities to stretch, take a brief walk, rest, or use the bathroom. Sometimes, it may be helpful to divide the structured interview into several shorter sessions.

Sensory impairments can also affect the evaluation. Older persons with hearing difficulties may misunderstand parts of the query or fail to answer the question entirely. It is often helpful to sit closer to the older client, face him or her directly, speak slowly and clearly, and reduce background noise in the evaluation room. The interested reader is referred to Edelstein, Staats, Kalish, and Northrop (1996) for full discussion of the effects of sensory deficits on the assessment of older adults. Finally, some older adults may respond negatively to the structure imposed by the interviews. In these cases, it is advisable to spend more time developing rapport, reflect more feelings during the interview, allow for more elaboration and venting from the client about troublesome symptoms, and explain the purpose and format of the structured interview. In all, flexibility and sensitivity are required when assessing and interviewing the personality-disordered older adult.

7.12.6 TREATMENT

The personality-disordered older adult is often ill-equipped to handle the challenges and changes associated with the normal aging process. A lifetime of social and occupational failures typically exacerbates the clinical picture, and increases the older adults sense of disappointment and hopelessness. As such, these individuals are often in need of counseling or supportive services to reduce their frustrations and difficulties. Unfortunately, little is currently known about the treatment of various PDs in the elderly. And, most of the literature is in the form of clinical case studies with a striking lack of group studies. Much of the literature we will discuss pertains to treatment models developed for younger individuals. The reader should be aware, however, that specific literature about treatment of PDs in the elderly is in short supply.

Evidence suggests that the majority of older adults with psychiatric problems do not seek or receive appropriate mental health services. Lazarus, Sadavoy, and Langsley (1991) have aptly described the many barriers to psychotherapy facing the current cohort of elderly persons: their attribution of psychiatric symptoms to physical rather than psychiatric causes, emphasis on medical treatment, stigmatization associated with psychiatric care, practical deterrents such as cost, transportation, and physical disabilities, and negative family and physician attitudes about geriatric mental health treatment. Unfamiliarity with current mental health treatment and a desire to remain independent also preclude intervention in some cases.

Two other factors specifically impede appropriate treatment of PDs in the elderly. The first concerns the lack of appropriate diagnosis. As we have discussed earlier, PDs are sometimes underdiagnosed by mental health professions due to their unfamiliarity with the disorders, the belief that PDs are rare in the elderly, an overemphasis on the Axis I symptoms that the client perceives as the problem, and a masking of Axis II diagnoses by the more flagrant or pressing Axis I symptoms. Although the underlying personality may not be taken into account, it certainly affects treatment outcome for these clients. Consider the treatment of an older male adult who is clinically depressed because he has recently retired and subsequently lost his wife to a fatal illness. Another elderly male may present with similar symptoms of depression, but his symptoms are not reflective of recent losses but rather are the effects of a lifetime of interpersonal chaos, self-destructive actions, and a poorly developed sense of self characteristic of borderline PD. Appropriate treatment for these two individuals most likely will take remarkably different forms. However, failure to recognize the moderating impact of personality dysfunction on psychiatric symptoms can result in poor treatment planning and therapeutic outcome. Indeed, in their classic report, Thompson et al. (1988) investigated whether presence of PDs affect outcome in treatment of 75 depressed older adults. They found that older adults with PDs did considerably worse in treatment than those without PDs, and these effects were still present at six month follow-up. Similar results have been reported in related studies with younger depressed clients with comorbid PD (Frank, Kupfer, & Jarret, 1987; Pilkonis & Frank, 1988).

The second factor hindering treatment of PD in older adults is that individuals (both older and younger) with a significant PD often fail to recognize their symptoms as signs of a psychiatric disturbance. In other words, they lack insight into the maladaptive nature of their behavior. Imagine an older adult suffering from narcissistic PD. This individual may have great difficulty maintaining adequate employment because of his sense of entitlement, condescending attitude, and difficulty following directions from his bosses whom the person regards as inferior. This older individual may blame his problems on his bosses who do not properly appreciate his talents or he may believe he is being discriminated against based on his age, when the real problem is the dysfunctional attitudes and behaviors characteristic of his personality type. Such individuals, who lack awareness of their own personality disturbance, typically see no need for personal change or psychiatric intervention. In general, individuals with a PD seek therapy when they begin to manifest troubling symptoms of an Axis I disorder, such as anxiety or depression. They see these symptoms as ego-dystonic and present at the clinician's office expecting relief. One tenet in treating older adults in this condition is to focus one's therapeutic attention initially on reduction of Axis I clinical symptoms until maximum therapeutic gain has been achieved (Abrams, 1991). At this early point, the therapeutic goal negotiated between therapist and client converge.

The early stage of treatment focusing on symptom relief is important for developing a therapeutic alliance and laying the foundation for later, more difficult, and more intensive treatment of the underlying character pathology. It is necessary to spend considerable time forming a therapeutic relationship with the older client. Indeed, a strong bond and safe environment is crucial to help older clients begin the process of viewing themselves and their behaviors more realistically (often more negatively). The older personality-disordered client may have few skills in this area and may have difficulty tolerating the negative feelings that often arise during this process. It is often helpful to proceed with caution and to be sensitive to increases in symptoms and distress. Typically, long-term psychotherapy is needed to address a lifetime of long-standing maladaptive behavior patterns. And, in some cases, treatment goals should be modest (Abrams, 1991).

Several models of treatment specific to PDs have appeared in the literature. For example, Young (1994) combines psychodynamic, object relations, and gestalt therapy concepts with standard cognitive therapy techniques in his schema-focused approach for treatment of PDs. Millon, one of the leaders in formulation and assessment of PDs, has developed a personologic and integrative treatment approach for PDs (Millon & Davis, in press), and Benjamin (1996) has presented her interpersonal approach to case conceptualization and treatment. These approaches have generated a great deal of anecdotal support, but little empirical validation has been reported. Other theorists have concentrated their attentions on one particular PD, with borderline PD receiving the most attention. For example, Paris (1994) describes a multidimensional treatment, while Horwitz et al. (1996) explicate a modified psychodynamic approach for borderlines. An innovative treatment for borderline PD, dialectical behavior therapy (DBT; Linehan, 1993) combines tenets from humanistic therapies and cognitive behavioral problem solving techniques. Several well-designed controlled outcome studies have been conducted to evaluate effectiveness of DBT with promising results (e.g., Linehan, Armstrong, Suarez, Allmon, & Heard, 1991; Linehan, Heard, & Armstrong, 1993), and DBT has recently been approved as an empirically validated treatment for borderline PD. The question remains whether this approach (as well as the others mentioned above) can be adapted for use with *older* individuals with severe personality dysfunction. In summary, although psychotherapy is generally considered the treatment of choice for PDs, few studies have actually tested various therapy models with older persons. The complexity of these disorders and well-established difficulty forming therapeutic relationships with personality-disordered individuals no doubt contribute to this lack of investigation.

Besides psychotherapeutic interventions, the psychopharmacological management of clients with PDs has been the focus of increasing attention in the past decade. Pharmacotherapy is sometimes used as an adjunctive strategy to ameliorate some acute depressive or anxiety symptoms and, in fact, can help some personality-disordered patients better engage in psychotherapy (Soloff, 1990). Such treatment is often welcomed by older clients who typically have been socialized to the medical model of intervention that requires little effort on their part. Unfortunately, the medication, while valuable in some instances in helping the patient regulate distressing affective, anxious, and cognitive symptoms, is not addressing the underlying pathology (i.e., the personality dysfunction) but is only providing symptom relief. Research on pharmacotherapy of PDs has mostly focused on Cluster B (dramatic, erratic, emotional) PDs, with generally promising, but preliminary results. For example, antidepressant medications have been found to show some positive effects for Cluster B patients (Ruegg & Frances, 1995), and antipsychotics have been found helpful for individuals with borderline PD, as they tend to reduce hostility, anxiety, and depression (Links & Steiner, 1988). The interested reader is referred to Soloff (1990) for an engaging and detailed review of pharmacological therapy for the PDs. Regrettably, pharmacotherapy studies typically have focused on younger and middle aged individuals. Despite encouraging findings, it remains to be seen how results with younger clients apply to geriatric clients. Indeed, older clients are particularly sensitive to interactive effects of medications and usually have greater side-effect profiles than younger persons. Pharmacotherapy studies with older clients are sorely needed.

7.12.7 SUMMARY

Initially in this chapter, we described the major evolution of the diagnosis of PDs from *DSM-I* to *DSM-IV*. By the placement of PDs on their own Axis in *DSM-III* and listing of specific and requisite criteria for each disorder in *DSM-III-R*, there was a tremendous growth in our knowledge of PDs and accompanying assessment tools and treatment techniques. Although there has been a parallel increasing interest in the psychology of older adults, clinical and investigative interest in PDs in this population is at the nascent stage. Many theories of the etiology of PDs have also been proposed, and we have reviewed several major biological, cognitive, evolutionary, and biosocial theories. It is clear from our review that most of these theories have ignored lifespan and aging issues, and, thus, there has been little or no attention to the nature or stability of PD symptomatology in the elderly.

Despite recent advances in the field, accurate diagnosis of PDs in the elderly is still a significant concern for clinicians and researchers. We have noted that some *DSM* diagnostic criteria may be inadequate for older individuals, given the unique physical, cognitive, and social complications encountered by the aged. We have suggested that the apparent decline in prevalence rates of some disorders (i.e., borderline, histrionic) may not be due to a maturation of the individual but rather the inability of the current diagnostic system to identify the age-related manifestations of these disorders. Some symptoms of PDs are erroneously attributed to old age and subsequently dismissed. Clients and professionals may also focus on Axis I symptoms and neglect to appreciate fully the role of the underlying personality dysfunction.

We have also reviewed the major self-report personality inventories and structured interviews employed to evaluate PDs. In clinical practice, all of the devices have been applied, no doubt, to older persons. However, formal reliability and validity studies of these instruments with the elderly are lacking in the literature. Recently, it has been demonstrated that the SCID for Axis I disorders can be reliably administered to older psychiatric populations (Segal, Hersen, Van Hasselt, Kabacoff, & Roth, 1993; Segal, Kabacoff, Hersen, Van Hasselt, & Ryan, 1995), and similar studies for Axis II instruments are needed. In the absence of empirical verification, it cannot be assumed that self-report devices and structured interviews for Axis II disorders have acceptable operating characteristics for this unique population. Definitive comparisons between instruments are also needed to advance the field. Our knowledge about the treatment of PDs in the elderly is presently limited, and much work remains to be conducted in this domain.

We hope that in this chapter we have delineated the problems and issues in the etiology, diagnosis, assessment, and treatment of PDs in the elderly. There have been some promising studies on these topics, and we feel confident that the growing general interest in aging and in PDs will lead to new contributions to this inchoate area.

7.12.8 REFERENCES

Abrams, R. (1991). Anxiety and personality disorders. In J. Sadavoy, L. W. Lazarus, & L. F. Jarvik,. (Eds.), *Comprehensive review of geriatric psychiatry* (pp. 369–386). Washington, DC: American Psychiatric Press.

Abrams, R. C., Alexopoulos, G. S., & Young, R. C. (1987). Geriatric depression and DSM-III-R personality disorder criteria. *Journal of the American Geriatrics Society, 35,* 383–386.

Abrams, R. C., Rosendahl, E., Card, C., & Alexopoulos, G. S. (1994). Personality disorder correlates of late and early onset depression. *Journal of the American Geriatrics Society, 42,* 727–731.

American Psychiatric Association. (1952). *Diagnostic and statistical manual of mental disorders*. Washington, DC: Author.

American Psychiatric Association. (1968). *Diagnostic and statistical manual of mental disorders* (2nd ed.). Washington, DC: Author.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. Rev.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Arntz, A., van Beijsterveldt, B., Hoekstra, R., Hofman, A., Eussen, M., & Sallaerts, S. (1992). The interrater reliability of a Dutch version of the Structured Clinical Interview for DSM-III-R Personality Disorders. *Acta Psychiatrica Scandinavica, 85,* 394–400.

Beck, A. T. (1992). Personality disorders (and their relationship to syndromal disorders). *Across-Species Comparisons and Psychiatry Newsletter, 5,* 3–13.

Benjamin, L. S. (1996). *Interpersonal diagnosis and treatment of personality disorders* (2nd ed.). New York: Guilford.

Brooks, R. B., Baltazar, P. L., McDowell, D. E., Munjack, D. J., & Bruns, J. R. (1991). Personality disorders cooccurring with panic disorder with agoraphobia. *Journal of Personality Disorders, 5,* 328–336.

Casey, D. A., & Schrodt, C. J. (1989). Axis II diagnoses in geriatric inpatients. *Journal of Geriatric Psychiatry and Neurology, 2,* 87–88.

Cattell, R. B., Saunders, D. R., & Stice, G. F. (1950). *The 16 Personality Factor Questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.

Cloninger, C. R. (1986). A unified biosocial theory of personality and its role in the development of anxiety states. *Psychiatric Developments, 3,* 167–226.

Cloninger, C. R. (1987). A systematic method for clinical description and classification of personality variants: A proposal. *Archives of General Psychiatry, 44,* 573–588.

Cloninger, C. R., Przybeck, T. R., Svrakic, D. M., & Wetzel, R. D. (1994). *The Temperament and Character Inventory (TCI): A guide to its development and use*. St. Louis, MO: Center for the Psychobiology of Personality, Washington University.

Cohen, G. D. (1990). Psychopathology and mental health in the mature and elderly adult. In J. E. Birren & K. W. Schaie (Eds.) *Handbook of the psychology of aging* (3rd ed., pp. 359–371). San Diego, CA: Academic Press.

Coolidge, F. L. (1984). *Coolidge Axis II Inventory*. US Copyright TXU 182 026, Washington, DC.

Coolidge, F. L., Becker, L. A, DiRito, D. C., Durham, R. L., Kinlaw, M. W., & Philbrick, P. B. (1994). On the relationship of the five-factor personality model to personality disorders: Four reservations. *Psychological Reports, 75*, 11–21.

Coolidge, F. L., Burns, E. M., Nathan, J. H., & Mull C. E. (1992). Personality disorders in the elderly. *Clinical Gerontologist, 12*, 41–55.

Coolidge, F. L., & Griego, J. (1996). *A construct validity investigation between the Temperament and Character Inventory and the Coolidge Axis II Inventory*. Manuscript submitted for publication.

Coolidge, F. L., & Merwin, M. M. (1992). Reliability and validity of the Coolidge Axis II Inventory: A new inventory for the assessment of personality disorders. *Journal of Personality Assessment, 59,* 223–238.

Edelstein, B., Staats, N., Kalish, K. D., & Northrop, L. R. (1996). Assessment of older adults. In M. Hersen & V. B. Van Hasselt (Eds.), *Psychological treatment of older adults: An introductory text* (pp. 35–68). New York: Plenum.

Endicott, J. & Spitzer, R. L. (1978). A diagnostic interview: The schedule for affective disorders and schizophrenia. *Archives of General Psychiatry, 35,* 837–844.

Eysenck, H. J. (1978). Superfactors P, E, N in a comprehensive factor space. *Multivariate Behavioral Research, 13,* 475–481.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1995a). *Structured Clinical Interview for Axis I DSM-IV Disorders—Patient Edition (SCID-I/P, Version 2.0)*. New York: Biometrics Research Department, New York State Psychiatric Institute.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1995b). The Structured Clinical Interview for DSM-III-R Personality Disorders (SCID-II), Part I: Description. *Journal of Personality Disorders, 19,* 83–91.

First, M. B., Spitzer, R. L., Gibbon, M., Williams, J. B. W., & Benjamin, L. (1994). *Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II, Version 2.0)*. New York: Biometrics Research Department, New York State Psychiatric Institute.

First, M. B., Spitzer, R. L., Gibbon, M., Williams, J. B. W., Davies, M., Borus, J., Howes, M. J., Kane, J., Pope, H. G., & Rounsaville, B. (1995). The Structured Clinical Interview for DSM-III-R Personality Disorders (SCID-II), Part II: Multisite test–retest reliability study. *Journal of Personality Disorders, 19,* 92–104.

Fogel, B. S., & Westlake, R. (1990). Personality disorder diagnoses and age in inpatients with major depression. *Journal of Clinical Psychiatry, 51,* 232–235.

Fogelson, D. L., Nuechterlein, K. H., Asarnow, R. F., Subotnik, K. L., & Talovic, S. A. (1991). Interrater reliability of the structured clinical interview for DSM-III-R, Axis II: Schizophrenia spectrum and affective spectrum disorders. *Psychiatry Research, 39,* 55–63.

Frances, A. (1982). Categorical and dimensional systems of personality diagnosis: A comparison. *Comprehensive Psychiatry, 23,* 516–527.

Frank, E., Kupfer, D. J., & Jarret, D. (1987). Personality features and response to acute treatment in recurrent depression. *Journal of Personality Disorders, 1,* 14–26.

Grove, W. M. (1987). The reliability of psychiatric diagnosis. In C. G. Last & M. Hersen (Eds.), *Issues in diagnostic research* (pp. 99–119). New York: Plenum.

Hathaway, S. R., & McKinley, J. C. (1983). *The Minnesota Multiphasic Personality Inventory manual.* New York: Psychological Corporation.

Hathaway, S. R., & McKinley, J. C. (1989). *MMPI-2: Minnesota Multiphasic Personality Inventory-2: Manual for administration and scoring.* Minneapolis, MN: University of Minnesota Press.

Hogg, B., Jackson, H. J., Rudd, R. P., & Edwards, J. (1990). Diagnosing personality disorders in recent-onset schizophrenia. *Journal of Nervous and Mental Disease, 178,* 194–199.

Horwitz, L., Gabbard, G. O., Allen, J. G., Frieswyk, S. H., Colson, D. B., Newsom, G. E., & Coyne, L. (1996). *Borderline personality disorder: Tailoring the psychotherapy to the patient.* Washington, DC: American Psychiatric Press.

Howard, R., Bandyopadhyay, D., & Cook, M. (1992). Dissocial personality disorder in the elderly: A case history. *International Journal of Geriatric Psychiatry, 7,* 527–532.

Hunt, C., & Andrews, G. (1992). Measuring personality disorder: The use of self-report questionnaires. *Journal of Personality Disorders, 6,* 125–133.

Hyer, L., & Harrison, W. R. (1986). Later life personality model: Diagnosis and treatment. *Clinical Gerontologist, 5,* 399–416.

Hyler, S. E. (1994). *Personality Diagnostic Questionnaire, (PDQ-4)* (4th ed.). New York: New York State Psychiatric Institute.

Hyler, S. E., & Rieder, R. O. (1987). *PDQ-R: Personality Diagnostic Questionnaire—Revised.* New York: New York State Psychiatric Institute.

Hyler, S. E., Skodol, A. E., Kellman, H. D., Oldham, J. M., & Rosnick, L. (1990). Validity of the Personality Diagnostic Questionnaire-Revised: Comparison with two structured interviews. *American Journal of Psychiatry, 147,* 1043–1048.

Hyler, S. E., Skodol, A. E., Oldham, J. M., Kellman, H. D., & Doidge, N. (1992). Validity of the Personality Diagnostic Questionnaire-Revised: A replication in an outpatient sample. *Comprehensive Psychiatry, 33,* 73–77.

Jackson, H. J., Gazis, J., Rudd, R. P., & Edwards, J. (1991). Concordance between two personality disorder instruments with psychiatric inpatients. *Comprehensive Psychiatry, 32,* 252–260.

Jung, C. G. (1971). *Psychological types.* (A revision by R. F. C. Hull of H. G. Baynes, Trans.). Princeton, NJ: Princeton University Press. (Original work published in 1924).

Kagan, J. & Moss, H. A. (1962). *Birth to maturity.* New York: Wiley.

Kernberg, O. F. (1996). A psychoanalytic theory of personality disorders. In J. F. Clarkin & M. F. Lenzenweger (Eds.), *Major theories of personality disorder* (pp. 106–140). New York: Guilford.

Klein, M. H., Benjamin, L. S., Rosenfeld, R., Treece, C., Husted, J., & Griest, J. H. (1993). The Wisconsin Personality Disorders Inventory: Development, reliability, and validity. *Journal of Personality Disorders, 7,* 285–303.

Kroessler, D. (1990). Personality disorder in the elderly. *Hospital and Community Psychiatry, 41,* 1325–1329.

Kunik, M. E., Mulsant, B. H., Rifai, A. H., Sweet, R. A., Pasternak, R., Rosen, J., & Zubenko, G. S. (1993). Personality disorders in elderly inpatients with major depression. *American Journal of Geriatric Psychiatry, 1,* 38–45.

Lazarus, L.W., Sadavoy, J., & Langsley, P. R. (1991). Individual psychotherapy. In J. Sadavoy, L. W. Lazarus, & L. F. Jarvik, (Eds.), *Comprehensive review of geriatric psychiatry* (pp. 487–512). Washington, DC: American Psychiatric Press.

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder.* New York: Guilford.

Linehan, M. M., Armstrong, H. E., Suarez, A., Allmon, D., & Heard, H. L. (1991). Cognitive-behavioral treatment of chronically parasuicidal borderline patients. *Archives of General Psychiatry, 48,* 1060–1064.

Linehan, M. M., Heard, H. L., & Armstrong, H. E. (1993). Naturalistic follow-up of a behavioral treatment of chronically parasuicidal borderline patients. *Archives of General Psychiatry, 50,* 971–974.

Links, P. S., & Steiner, M. (1988). Psychopharmacologic management of patients with borderline personality disorder. *Canadian Journal of Psychiatry, 33,* 355–359.

Loranger, A. W., Hirschfeld, R. M., Sartorius, N., & Regier, D. A. (1991). The WHO/ADAMHA international pilot study of personality disorders: Background and purpose. *Journal of Personality Disorders, 5,* 296–306.

Loranger, A.W., Sartorius, N., Andreoli, A., Berger, P., Buchleim, P., Channabasavanna, S. M., Coid, B., Dahl, A., Diekstra, R. F. W., Feguson, B., Jacobsberg, L. B., Mombour, W., Pull, C., Ono, Y., & Regier, D. (1994). The International Personality Disorder Examination: The World Health Organization/Alcohol, Drug Abuse, and Mental Health Administration international pilot study of personality disorders. *Archives of General Psychiatry, 51,* 215–224.

Loranger, A. W., Susman, V. L., Oldham, J. M., & Russakoff, L. M. (1987). The Personality Disorder Examination: A preliminary report. *Journal of Personality Disorders, 1,* 1–13.

Malow, R. M., West, J. A., Williams, J. L., & Sutker, P. B. (1989). Personality disorders classification and symptoms in cocaine and opioid addicts. *Journal of Consulting and Clinical Psychology, 57,* 765–767.

McCrae, R. R., & Costa, P. T., Jr. (1986). Clinical assessment can benefit from recent advances in personality psychology. *American Psychologist, 41,* 1001–1003.

McCrae, R. R., & Costa, P. T., Jr. (1989). More reasons to adopt the five-factor model. *American Psychologist, 44,* 451–452.

McGue, M., Bacon, S., & Lykken, D. T. (1993). Personality stability and change in early adulthood: A behavioral genetic analysis. *Developmental Psychology, 29,* 96–109.

Mellsop, G., Varghese, F., Joshua, S., & Hicks, A. (1982). The reliability of Axis II of DSM-III. *American Journal of Psychiatry, 139,* 1360–1361.

Mezzich, J. E., Fabrega, H., Coffman, G. A., & Glavin, Y. (1987). Comprehensively diagnosing geriatric patients. *Comprehensive Psychiatry, 28,* 68–76.

Miller, H. R., Streiner, D. L., & Parkinson, A. (1992). Maximum likelihood estimates of the ability of the MMPI and MCMI personality disorder scales and the SIDP to identify personality disorders. *Journal of Personality Assessment, 59,* 1–13.

Millon, T. (1983). *Millon Clinical Multiaxial Inventory.* Minneapolis, MN: Interpretive Scoring Systems.

Millon, T. (1987). *Millon Clinical Multiaxial Inventory II (MCMI II) manual.* Minneapolis, MN: National Computer Systems.

Millon, T. (1992). Millon Clinical Multiaxial Inventory: I & II. *Journal of Counseling and Development, 70,* 421–426.

Millon, T. (1994). *Millon Clinical Multiaxial Inventory III*

(MCMI III) manual. Minneapolis, MN: National Computer Systems.

Millon, T., & Davis, R. D. (1996). An evolutionary theory of personality disorders. In J. F. Clarkin & M. F. Lenzenweger (Eds.), *Major theories of personality disorder* (pp. 221–346). New York: Guilford.

Millon, T., & Davis, R. D. (in press). *Synergistic psychotherapy*: *Integrative syndromic and personologic treatment*. New York: Wiley-Interscience.

Molinari, V., & Marmion, J. (1995). Relationship between affective disorders and Axis II diagnoses in geropsychiatric patients. *Journal of Geriatric Psychiatry and Neurology, 8,* 61–64.

Morey, L. C. (1988). The categorical representation of personality disorder: A cluster analysis of DSM-III-R personality features. *Journal of Abnormal Psychology, 97,* 314–321.

Morey, L. C., Waugh, M. H., & Blashfield, R. K. (1985). MMPI scales for DSM-III personality disorders: Their derivation and correlates. *Journal of Personality Assessment, 49,* 245–251.

Morgan, D. G., May, P. C., & Finch, C. E. (1987). Dopamine and serotonin systems in human and rodent brain: Effects of age and neurogenerative disease. *Journal of the American Geriatric Society, 35,* 334–345.

Myers, I. B. (1962). *Manual: Myers–Briggs Type Indicator*. Palo Alto, CA: Consulting Psychological Press.

Nazikian, H., Rudd, R. P., Edwards, J., & Jackson, H. J. (1990). Personality disorder assessment for psychiatric inpatients. *Australian and New Zealand Journal of Psychiatry, 24,* 37–46.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology, 43,* 574–583.

O'Boyle, M., & Self, D. (1990). A comparison of two interviews for DSM-III-R personality disorders. *Psychiatry Research, 32,* 85–92.

Oldham, J. M., Skodol, A. E., Kellman, H. D., Hyler, S. E, Rosnick, L., & Davies, M. (1992). Diagnosis of DSM-III-R personality disorders by two structured interviews: Patterns of comorbidity. *American Journal of Psychiatry, 149,* 213–220.

Paris, J. (1994). *Borderline personality disorder: A multidimensional approach*. Washington, DC: American Psychiatric Press.

Pfohl, B., Black, D. W., Noyes, R., Kelly, M., & Blum, N. (1990). A test of the tridimensional personality theory: Association with diagnosis and platelet imiprimine binding in obsessive-compulsive disorder. *Biological Psychiatry, 28,* 41–46.

Pfohl, B., Blum, N., & Zimmerman, M. (1995). *Structured Interview for DSM-IV Personality SIDP-IV*. Iowa City, IA: University of Iowa.

Pfohl, B., Blum, N., Zimmerman, M., & Stangl, D. (1989). *Structured Interview for DSM-III-R Personality (revised) (SIDP-R)*. Iowa City, IA: University of Iowa.

Pilkonis, P. A., & Frank, E. (1988). Personality pathology in recurrent depression: Nature, prevalence, and relationship to treatment response. *American Journal of Psychiatry, 145,* 435–441.

Pilkonis, P. A., Heape, C. L., Ruddy, J. & Serrao, P. (1991). Validity in the diagnosis of personality disorders: The use of the LEAD standard. *Psychological Assessment, 3,* 46–54.

Plomin, R., Chipuer, H. M., & Loehlin, J. C. (1990). Behavioral genetics and personality. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 225–243). New York: Guilford.

Pretzer, J. L., & Beck, A. T. (1996). A cognitive theory of personality disorders. In J. F. Clarkin & M. F. Lenzenweger (Eds.), *Major theories of personality disorder* (pp. 36–105). New York: Guilford.

Robins, L. N., Cottler, L. B., Bucholz, K., & Compton, W. (1996). *The Diagnostic Interview Schedule, (Version IV)*. St. Louis. MO: Washington University School of Medicine.

Robins, L. N., Helzer, J. E., Croughan, J., & Ratcliff, K. S. (1981). National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics, and validity. *Archives of General Psychiatry, 38,* 381–389.

Rose, M. K., Soares, H. H., & Joseph, C. (1993). Frail elderly clients with personality disorders: A challenge for social work. *Journal of Gerontological Social Work, 19,* 153–165.

Rosowsky, E., & Gurian, B. (1991). Borderline personality disorder in late life. *International Psychogeriatrics, 3,* 39–52.

Rosowsky, E., & Gurian, B. (1992). Impact of borderline personality disorder in late life on systems of care. *Hospital and Community Psychiatry, 43,* 386–389.

Ruegg, R., & Frances, A. (1995). New research in personality disorders. *Journal of Personality Disorders, 9,* 1–48.

Ruegg, R. G., Gilmor, J., Ekstrom, R. D., Corrigan, M., Knight, B., Tancer, M., Leatherman, M. E., Carson, S. W., & Golden, R. N. (1993). A psychoneuroendocrine study of the unified biosocial theory of personality. *Biological Psychiatry, 34,* 885–888.

Rutter, M. (1987). Temperament, personality and personality disorder. *British Journal of Psychiatry, 150,* 443–458.

Schneider, L. S., Zemansky, M. F., Bender, M., & Sloane, R. B. (1992). Personality in recovered depressed elderly. *International Psychogeriatrics, 4,* 177–185.

Segal, D. L. (1997). Structured interviewing and DSM classification. In S. M. Turner & M. Hersen (Eds.), *Adult psychopathology and diagnosis* (3rd. ed., App. 24–57). New York: Wiley.

Segal, D. L., & Falk, S. B. (in press). Structured diagnostic interviews and rating scales. In M. Hersen & A. S. Bellack (Eds.), *Behavioral assessment: A practical handbook* (4th ed.). New York: Allyn & Bacon.

Segal, D. L., Hersen, M., & Van Hasselt, V. B. (1994). Reliability of the Structured Clinical Interview for DSM-III-R: An evaluative review. *Comprehensive Psychiatry, 35,* 316–327.

Segal, D. L., Hersen, M., Van Hasselt, V. B., Kabacoff, R. I., & Roth, L. (1993). Reliability of diagnoses in older psychiatric patients using the Structured Clinical Interview for DSM-III-R. *Journal of Psychopathology and Behavioral Assessment, 15,* 347–356.

Segal, D. L., Kabacoff, R. I., Hersen, M., Van Hasselt, V. B., & Ryan, C. F. (1995). Update on the reliability of diagnosis in older psychiatric outpatients using the Structured Clinical Interview for DSM-III-R. *Journal of Clinical Geropsychology, 1,* 313–321.

Siegel, D. J., & Small, G. W. (1986). Borderline personality disorder in the elderly: A case study. *Canadian Journal of Psychiatry, 31,* 859–860.

Silberman, C. S., Roth, L., Segal, D. L., & Burns, W. (in press). Relationship between the Millon Clinical Multiaxial Inventory-II and Coolidge Axis II Inventory in chronically mentally ill older adults: A pilot study. *Journal of Clinical Psychology*.

Skodol, A. E., Oldham, J. M., Rosnick, L., Kellman, H. D., & Hyler, S. E. (1991). Diagnosis of DSM-III-R personality disorders: A comparison of two structured interviews. *International Journal of Methods in Psychiatric Research, 1,* 13–26.

Skodol, A. E., Rosnick, L., Kellman, H. D., Oldham, J. M., & Hyler, S. E. (1988). Validating structured DSM-III-R personality disorder assessments with longitudinal data. *American Journal of Psychiatry, 145,* 1297–1299.

Soldz, S., Budman, S., Demby, A., & Merry, J. (1993). Diagnostic agreement between the Personality Disorder

Examination and the MCMI-II. *Journal of Personality Assessment, 60,* 486–499.

Soloff, P. H. (1990). What's new in personality disorders?: An update on pharmacological treatment. *Journal of Personality Disorders, 4,* 233–243.

Spar, J. E., & LaRue, A. (1990). *Concise guide to geriatric psychiatry*. Washington, DC: American Psychiatric Press.

Spitzer, R. L., Endicott, J., & Robins, E. (1978). Research diagnostic criteria: Rationale and reliability. *Archives of General Psychiatry, 35,* 773–782.

Spitzer, R. L., Forman, J. B. W., & Nee, J. (1979). DSM-III field trials: I. Initial interrater diagnostic reliability. *American Journal of Psychiatry, 136,* 815–817.

Spitzer, R. L., Williams, J. B. W., Gibbon, M., & First, M. B. (1990). *Structured Clinical Interview for DSM-III-R Personality Disorders (SCID-II, Version 1.0)*. Washington, DC: American Psychiatric Press.

Standage, K., & Ladha, N. (1988). An examination of the reliability of the Personality Disorder Examination and a comparison with other methods of identifying personality disorders in a clinical sample. *Journal of Personality Disorders, 2,* 267–271.

Stangl, D., Pfohl, B., Zimmerman, M., Bowers, W., & Corenthal, C. (1985). A structured interview for the DSM-III personality disorders: A preliminary report. *Archives of General Psychiatry, 42,* 591–596.

Thomas, A. & Chess, S. (1977). *Temperament and development*. New York: Brunner/Mazel.

Thompson, L. W., Gallagher, D., Czirr, R. (1988). Personality disorder and outcome in the treatment of late-life depression. *Journal of Geriatric Psychiatry, 21,* 133–146.

Trull, T. J. & Larson, S. L. (1994). External validity of two personality disorder inventories. *Journal of Personality Disorders, 8,* 96–103.

Turley, B., Bates, G. W., Edwards, J., & Jackson, H. J. (1992). MCMI-II personality disorders in recent-onset bipolar disorders. *Journal of Clinical Psychology, 48,* 320–329.

Tyrer, P. (1988). *Personality disorders: Diagnosis, management, and course*. London: Wright/Butterworth.

World Health Organization. (1995). *The International Personality Disorder Examination (IPDE) DSM-IV module*. Washington, DC: American Psychiatric Press.

Young, J. E. (1994). *Cognitive therapy for personality disorders: A schema-focused approach* (Rev. ed.). Sarasota, FL: Professional Resources Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.13
Schizophrenia and Related Conditions in Late Life

PHILIP D. HARVEY
Mount Sinai School of Medicine, New York, NY, USA

7.13.1 INTRODUCTION 291
7.13.2 CONCEPTIONS OF SCHIZOPHRENIA AND IMPLICATIONS FOR RESEARCH ON SCHIZOPHRENIA IN LATE LIFE 292
7.13.3 DIAGNOSIS OF SCHIZOPHRENIA AND RELATED CONDITIONS IN LATE LIFE 292
7.13.4 EPIDEMIOLOGY OF SCHIZOPHRENIA AND RELATED CONDITIONS IN LATE LIFE 293
  7.13.4.1 New Incidence of Schizophrenia in Late Life 293
  7.13.4.2 New Incidence of Related Conditions in Late Life 294
7.13.5 GENERAL OUTCOME OF SCHIZOPHRENIA IN LATE LIFE 294
7.13.6 OUTCOME OF ACUTE PSYCHIATRIC ADMISSION IN LATE LIFE 294
7.13.7 CLINICAL SYMPTOMS IN LATE-LIFE SCHIZOPHRENIA 295
7.13.8 COGNITIVE FUNCTIONING IN LATE-LIFE SCHIZOPHRENIA 296
7.13.9 ADAPTIVE IMPAIRMENT IN LATE-LIFE SCHIZOPHRENIA 300
7.13.10 TREATMENT OF SCHIZOPHRENIA IN LATE LIFE 301
7.13.11 CONCLUSION 302
7.13.12 REFERENCES 302

7.13.1 INTRODUCTION

Schizophrenia has been studied in detail from a psychological perspective. This research has focused on symptomatology, adaptive functioning, and cognitive impairments during the premorbid, early, and middle stages of the illness. In contrast to the huge volume of psychological research on schizophrenia in these earlier stages, there is considerably less research on schizophrenia and related conditions in late life. Understanding schizophrenia in late life is important for several reasons, including developing knowledge of the impact of aging on the illness, developing new treatments specifically targeted at older individuals, and understanding any differences between individuals with schizophrenia that has been chronic throughout the lifespan and individuals who develop schizophrenia late in life. With the aging of the population of the United States, the average age of psychiatric patients, including persons with schizophrenia, will increase as well. It is estimated that in 1997, there are 300 000 individuals over the age of 65 with a history of a schizophrenia diagnosis (Blazer, 1996), with many (20 000) still receiving institutional care of some sort (Hafner & Hambrecht, 1994). In addition, late-onset schizophrenia has been studied in considerably more detail in the early 1990s, with a number of interesting findings emerging regarding the similarities and differences between same-aged individuals with schizophrenia with early onset and those with a lifelong course of illness (Jeste et al., 1995).

This chapter focuses on the full range of psychological characteristics of schizophrenia, including clinical symptoms and phenomenology, cognitive functioning, adaptive skills, and treatments. Since schizophrenia has proven to be the most heterogeneous of all psychiatric conditions, the effects of aging on this condition would be expected to be very diverse and the developing research findings confirm this expectation. Symptoms, cognitive functions, and outcomes appear to vary greatly across elderly patients with a history of schizophrenia, ranging from complete disappearance of symptoms with stable residence in the community (Harding, Brooks, Ashikaga, Strauss, & Breier, 1987a, 1987b; Harding, Zubin & Strauss, 1992) to stays in chronic care hospitals of more than seven consecutive decades with complete dependence on others for all of their needs (Davidson et al., 1995). Thus, the extreme variability across individuals seen in schizophrenia in earlier life appears magnified in late life.

7.13.2 CONCEPTIONS OF SCHIZOPHRENIA AND IMPLICATIONS FOR RESEARCH ON SCHIZOPHRENIA IN LATE LIFE

Schizophrenia was initially described as a condition that developed in late adolescence and led to a deterioration in cognitive functions and behavior. The specifics of the deterioration have been in dispute since that time. Although Kraepelin (1919) referred to schizophrenia as "dementia praecox," it is not clear that he meant that schizophrenia was a condition which led to progressive deterioration in cognitive and behavioral functioning. It is clear, however, that his general perspective was that after the development of the illness, a full return to the previous level of functioning is almost never expected. In contrast, Bleuler's (1911/1950) conception of the illness allowed not only for a full remission of symptoms but also for considerable variation in symptom severity during the active illness itself. Thus, from the very first detailed descriptions of schizophrenia there has been controversy regarding the course of the illness and, by implication, the level of adaptive and cognitive impairment expected in older individuals with the illness.

Onset age has also been a controversial issue in the conceptions of the illness. Kraepelin delineated dementia praecox from paraphrenia, thus subdividing the illness on the dimension of age of onset. Paraphrenia was described as an illness of late onset, with predominantly persecutory (i.e., paranoid) delusions and minimal additional impairments. Hallucinations, if present, were not a central feature of the illness and delusions were nonbizarre. In the *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)* (American Psychiatric Association [APA], 1993), all hallucinations are required to be tactile or olfactory, not auditory, and must be related to the theme of the delusion. Patients with paraphrenia are classically described as lacking major changes in their affective response, and their level of motivation, and having preserved cognitive functions (Kraeplin, 1919). Their symptoms were seen as also remitting spontaneously with considerably more frequency than in dementia praecox. Bleuler disputed this conception, arguing instead that schizophrenia could develop in later life and that it was possible for a late onset of schizophrenia. Thus, the Bleulerian approach to the conception of the illness was that there was minimal difference in the characteristics of the illness seen as a function of onset age.

There has been controversy and evolution in diagnostic criteria for schizophrenia. Since the late 1970s, there has been considerable revision in diagnostic criteria for schizophrenia, with one of the major changes focusing on the issue of age at onset and the diagnosis for the disorder. The first of the new American diagnostic criteria, Research Diagnostic Criteria (RDC; Spitzer, Endicott, & Robins, 1978), required onset of symptoms between the ages of 16 and 45 for a diagnosis of schizophrenia. The rule-out for late onset was also present in *DSM-III* (APA, 1980), but was removed from the revision of these criteria in *DSM-IIIR* (APA, 1987). *DSM-IV* (APA, 1993) criteria also allow for schizophrenia to be diagnosed in individuals with an onset of symptoms at any age. No differences in symptomatology are required as a function of age of onset. One of the problems, of course, with diagnosing schizophrenia in later life is that many dementing conditions, including subcortical dementias such as Parkinson's disease and cortical dementias such as Alzheimer's disease, often have psychotic symptoms as a component of their symptom presentation (Jeste, Wragg, Salmon, Harris, & Thal, 1992). This situation can lead to complications in the initial diagnosis of schizophrenia.

7.13.3 DIAGNOSIS OF SCHIZOPHRENIA AND RELATED CONDITIONS IN LATE LIFE

Schizophrenia, as established by US and international criteria, is defined by the presence of characteristic symptoms, a clear decline in overall adaptive functioning, a minimum duration of illness, and the absence of other factors that could have caused the symptoms, such as affective disorder, substance abuse, brain dysfunction, or other general medical conditions (seizure disorders, etc.). Related conditions include schizophreniform disorder (illness meeting all criteria for schizophrenia other than the duration requirement), delusional disorder (a disorder with a single prominent, nonbizarre, delusion but not enough additional symptoms to meet criteria for schizophrenia), and schizoaffective disorder. As noted above, a major problem in diagnosing a new onset of any of these disorders in late life is the fact that degenerative dementing illnesses often have psychotic features as a symptom (Jeste et al., 1992). Thus, a careful cognitive assessment, including follow-up evaluation, is required at the time of diagnosis to rule out the presence of an alternative dementing condition. Neurological or neuroimaging assessment to rule out a cerebrovascular accident is also prudent.

One of the complexities of confirming a continued diagnosis of any of these disorders in individuals with a lifelong course of illness is that of ruling out alternative influences on symptoms. Aging-related physical illnesses and their consequences, including vascular disease, diabetes, and pulmonary problems, may cause cognitive impairments similar to those seen in schizophrenia. Thus, knowledge of current medical status is more important in generating a diagnosis of schizophrenia and related conditions in late life.

7.13.4 EPIDEMIOLOGY OF SCHIZOPHRENIA AND RELATED CONDITIONS IN LATE LIFE

Schizophrenia is well accepted to have a lifetime prevalence of around 1% of the population, across all cultures. Data suggest that there are slightly more male than female patients who meet diagnostic criteria for schizophrenia (Lewine, 1981), although onset age is quite important in determination of sex ratios of newly incident cases. Female patients are known to have a generally later age of onset for the illness (Faraone, Chen, Goldstein & Tsuang, 1994), although this finding may not hold true for patients who have a positive family history for schizophrenia in a first-degree relative (Albus & Maier, 1995). Prevalence of schizophrenia in late life is also a function of survival of early-onset cases, since schizophrenia has significant mortality associated with it. Most estimates of mortality associated with schizophrenia suggest that the first few years of illness are most dangerous, with accidents, suicide, and homicide accounting for mortality in these early periods. The large numbers of very old long-stay patients found in public psychiatric facilities in both the US and the UK suggest that long-stay patients as a group do not necessarily have an extremely abbreviated life expectancy. Two large-scale whole-hospital studies have found that female long-stay patients outnumber male patients (Davidson et al., 1995; Leff, Thornicroft, Coxhead, & Crawford, 1994), but differences in gender-related life expectancy might account for these findings.

There are an estimated 20 000 early-onset geriatric schizophrenic patients residing in chronic psychiatric hospitals in the US. In addition, there are another 200 000 living in nursing home care and an additional unknown number residing in community care (Meeks et al., 1990). It is not known how many former schizophrenic patients are living in relative remission in the community, but the results of some studies (Harding et al., 1987a, 1987b, 1992) suggest that this number may be substantial. Thus, the prevalence of diagnosable schizophrenia will be less than the sum of all cases still surviving who at one time carried a diagnosis of schizophrenia.

7.13.4.1 New Incidence of Schizophrenia in Late Life

Studies reviewed by Harris and Jeste (1988) suggested that approximately 15% of patients with a first episode of schizophrenia were over the age of 60 at the time of first psychiatric admission. In nearly every one of these studies, the proportion of females was greater than that of males in the new onset cases. These studies are often limited by virtue of being hospital admission studies and therefore fail to identify those cases who manifest the disorder but do not seek or otherwise obtain treatment. In a large-scale population-based prospective study, Castle and Murray (1993) examined the new-onset cases of schizophrenia in a London borough. They found that the average age of onset for males (31.2 years) was ten years younger than that of females (41.1 years) and that the prevalence and associated characteristics of the illness did not differ markedly as a function of the breadth of the schizophrenia criteria that were used. The male/female ratio changed as a function of age of onset, with the ratio at 1.6 to 1 for onset from ages 16 to 25 and 0.2 to 1 for onset age after 76. Patients with late onset were also much less likely to have a family history of schizophrenia, but some increase in familial prevalence of affective disorders was found in these patients. Previous studies of late-onset schizophrenia have suggested that late-onset schizophrenic patients typically have good adaptive functioning before the onset of illness, with some increase over normative expectations in the prevalence of schizoid or paranoid personality (Harris & Jeste, 1988).

7.13.4.2 New Incidence of Related Conditions in Late Life

A few studies have examined the characteristics of paraphrenia (i.e., late-onset delusional disorder) and compared it with schizophrenia. Parsons (1964) reported a prevalence of "late paraphrenia" of 0.17 to 0.24 cases per hundred community-dwelling elderly. Christenson and Blazer (1984) reported a 4% prevalence of paranoid delusions in community-dwelling elderly. Flint, Rifat, and Eastwood (1991) found a high incidence of "organicity" in the paraphrenia cases as compared with patients with late-onset schizophrenia. Yassa and Suranyi-Cadotte (1993) compared 20 patients with late-onset schizophrenia with 20 patients with late-onset paraphrenia. The incidence of paraphrenia was 4% of the total admissions over a 7-year period to a geriatric inpatient service, which was the same prevalence as for first-episode schizophrenia. Patients with paraphrenia were more likely to have had a concomitant physical illness and were less likely to have had a premorbid picture of personality abnormalities. These patients were also less likely to have a relapse after successful neuroleptic treatment.

Thus, many cases of schizophrenia apparently develop late in life, with several distinct epidemiological characteristics that separate them from early-onset cases and from related disorders with late-life onset. The prevalence of late-onset paraphrenia is more difficult to estimate, but is probably much less than that of schizophrenia. Since one-third of all cases of schizophrenia may have onset after age 50, the expected population prevalence of late-onset schizophrenia may be 0.33%, which is at least twice that of estimates of the prevalence of paraphrenia.

7.13.5 GENERAL OUTCOME OF SCHIZOPHRENIA IN LATE LIFE

Schizophrenia has long been noted for its variability in outcome, with a tendency toward negative outcomes dominating. Only about one-third of patients who meet diagnostic criteria for schizophrenia are able to resume adaptive functioning with any success and as many as one-third are chronically impaired and largely dependent on others for care on a lifelong basis (Wyatt, 1991). Comprehensive outcome studies suggest a wide variance in overall outcome (Ciompi, 1980). Although regional trends in mental health care may have as great an influence as the disorder itself on prevalence of patients with extremely long inpatient stay, a considerable number of schizophrenic patients have a very negative outcome with extended stays at high levels of care. Hospital-based studies have suggested that large numbers of geriatric patients with schizophrenia still reside in chronic psychiatric hospitals with low levels of functioning (Arnold et al., 1995; Davidson et al., 1995). These patients often have a relatively early age of onset, severe negative symptoms, poor adaptive functioning (both before and after the onset of initial symptoms), and severe cognitive impairments. Although some of these patients may be discharged from the hospital, many long-stay patients are simply referred to nursing home care (White et al., in press). Among the characteristics of patients who are difficult to discharge in late life are belligerent and hostile behaviors (White et al., in press), including both verbal and behavioral aggressiveness.

In contrast to hospital-based studies which tend to paint a fairly negative picture of outcome in the disorder, one comprehensive study has suggested that many individuals with an early diagnosis of schizophrenia go on to be discharged from the institution and to live in the community for decades with modest to no residual symptoms. The Vermont longitudinal study (Harding et al., 1987a, 1987b, 1992) suggested that many individuals with a diagnosis of schizophrenia in early adulthood are residing in the community, with modest to adequate adaptive, occupational, and interpersonal functions. Since this study was based in a rural state, which varies from urban areas in terms of levels of support and intactness of the community, the applicability of this study to patients from urban and suburban catchment areas is unclear. A study of community-dwelling geriatric schizophrenic patients (Bartels, Mueser, & Miles, 1997) based in the nearby state of New Hampshire found that there was considerably more impairment in adaptive functioning and clinical symptoms than reported in the Vermont study.

7.13.6 OUTCOME OF ACUTE PSYCHIATRIC ADMISSION IN LATE LIFE

While studies of chronic hospital populations may paint a gloomy picture of outcome in the disorder, there are some data from recent studies of acute care of geriatric patients with an acute admission for schizophrenia that also raise some questions about the outcome of even acute psychiatric care in the elderly patient with schizophrenia. Studies of nongeriatric first-admission patients (Loebel et al., 1992) suggest that over 90% of first-admission patients are in a state of remission within one year. Statistics for readmissions of patients with a previous history of schizophrenia are typically similar. While approximately 60% of acutely admitted nongeriatric psychiatric patients are discharged to the care of relatives, with another 30% being referred to some community residence, and only about 5% referred directly to chronic care (Keefe & Harvey, 1994), the two studies previously published on outcome of acute psychiatric admissions in geriatric patients suggested that only 50% at most of all acutely admitted patients who had been residing in the community returned to their former residence (Mulsant et al., 1993; Rockwood, Stolee, & Brahim, 1991). The geriatric schizophrenic patients in those two studies had a very high rate of referral to nursing home care (33% in Mulsant et al.) or long-term psychiatric hospitals (33% in Mulsant et al. and 20% in Rockwood et al.). These data suggest that there is a high risk for any single acute psychiatric treatment of a geriatric patient with schizophrenia to result in a referral to long-term care.

Explanations for this situation are complex. It is possible that acute psychiatric admissions in geriatric patients with schizophrenia are precipitated as much by medical or social crises as they are by exacerbation of symptoms of the illness. If there is an illness or other adverse event affecting a caregiver, a psychiatric admission may be precipitated, with the change in social environment precluding a return to the prior residence. An additional possibility is that what appears to be an admission for psychiatric reasons is actually precipitated by the onset of a degenerative dementia, such as Alzheimer's disease. The current set of short-term outcome studies do not allow for this possibility to be ruled out. Thus, it is not clear if schizophrenia alone, or if additional neurological or medical conditions are responsible for the relatively adverse outcome of geriatric acute psychiatric care.

7.13.7 CLINICAL SYMPTOMS IN LATE-LIFE SCHIZOPHRENIA

Clinical symptoms have been studied in both early- and late-onset cases of late-life schizophrenia. Symptoms of the illness are typically characterized in terms of three different dimensions: positive symptoms (i.e., florid), including delusions and hallucinations; negative symptoms (i.e., deficit symptoms), including reductions in affective response, language output, and motivated behaviors; and disorganized symptoms, including formal thought disorder, inappropriate affective, and bizarre behavior (Andreasen, Arndt, Alliger, Miller, & Flaum, 1995). For years, clinical lore held that positive and disorganized symptoms were "burned out" and that only negative or deficit symptoms would be detected in late life, even in long-stay institutionalized patients (Bridge, Cannon, & Wyatt, 1978). Several studies have suggested that this conception is partially to mostly incorrect.

In a large-scale study of chronically institutionalized schizophrenic patients ranging in age from 25 to 95 (Davidson et al., 1995), negative symptoms manifested an increase in severity over the age-range studied, while positive symptoms manifested a decrease. The severity scores for negative symptoms in this population were very elevated compared with those seen in younger patients. The decrease in severity of positive symptoms, however, was such that even the oldest of the patients studied manifested positive symptom severity scores at the 40th percentile for the standardization sample for the rating scale used (the Positive and Negative Syndrome Scale: PANSS; Kay, 1991). The severity of negative symptoms in the geriatric patients in this sample was found to increase slightly at a one-year follow-up of the sample, with no statistically significant decrease in positive symptoms (Putnam et al., 1996). This result is in contrast to studies of younger chronic schizophrenic patients (Mueser, Douglas, Bellack, & Morrison, 1991; Pogue-Geille & Harrow, 1984, 1985) where negative symptoms did not worsen in severity at similar follow-up intervals. In a study of formal thought disorder (Harvey, Lombardi et al., 1997), roughly parallel findings were seen. Geriatric chronically institutionalized patients manifested more severe poverty of speech and less severe "positive" thought disorders (e.g., derailment, tangentiality, loss of goal, and circumstantiality) than younger chronically institutionalized patients.

In a comparative study of late-onset, acutely ill patients with schizophrenia and early-onset chronic patients, paranoid symptoms were significantly more common on the late-onset cases (Jeste et al., 1995). In addition, late-onset patients were more likely to have better lifelong functioning, including marriage and a stable work history. A previous study (Harris, Jeste, Krull, Montague, & Heaton, 1991) had suggested that the deficit syndrome was present in approximately 30% of older patients with schizophrenia, with this prevalence probably not elevated from that seen in younger patients. When the stability of the deficit syndrome was examined in a large sample of geriatric chronic patients (Putnam et al., 1996), a high degree of cross-temporal stability was noted. Finally, White et al. (1994) used confirmatory factor analysis to compare the structure of symptoms in poor-outcome geriatric schizophrenic patients, divided on the basis of their level of cognitive dysfunction, with younger poor-outcome patients. No differences in the factor structure of clinical symptoms were found between older and younger patients or between the geriatric patients who varied in their level of cognitive impairment. The authors concluded that since schizophrenic cognitive impairment did not affect the structure of symptoms, it may be an instrinsic aspect of poor-outcome schizophrenia.

Any assessment of clinical symptoms in chronic schizophrenic patients, especially those with a long and often uninterrupted institutional stay, must seriously consider the role of iatrogenic factors. While life in a chronic care institution may not be likely to sustain the presence of positive symptoms, the level of impoverishment of the environment, years of treatment with antipsychotic medication, and the low staff-to-patient ratio could easily be argued to influence negative symptom severity. Although length of institutional stay had no relationship with the severity of negative symptoms in the Davidson et al. (1995) study, all patients had already been in the hospital for at least a decade. As in other discussions of cognitive functioning, chronic institutional care cannot be ruled out as a contributory factor to the severity of negative symptoms as measured in those studies. The results of longitudinal studies of early-stage patients with schizophrenia (Knight, Roff, Barrnet, & Moss, 1979) suggest that more severe negative symptoms early in the course of illness predict a more adverse outcome, and multiple cross-sectional studies of patients at approximately the same age suggest that negative symptoms are more severe in patients with a chronic course of illness and greater deficits in life skills. Positive symptom severity has been demonstrated to have no cross-sectional or longitudinal relationship with outcome (Green, 1996). More severe symptoms of disorganization, especially communication disorder, are associated cross-sectionally with poor outcome (Keefe et al., 1987). Thus, the most tempered interpretation of the data suggests that negative symptoms may start out worse in those patients with a lifelong stay in a chronic psychiatric hospital, with a late-life worsening where the institutional stay itself may affect symptom severity in certain patients.

7.13.8 COGNITIVE FUNCTIONING IN LATE-LIFE SCHIZOPHRENIA

Cognitive impairment is a very common symptom in schizophrenia, being present across the entire course of the illness (Kraeplin, 1919; Saykin et al., 1994). While a full discussion of the characteristics of cognitive impairment is beyond the scope of this chapter (see Gold & Harvey 1993 for a specific review of that symptom), it should be noted that deficits in academic functioning, relative to classmates, are identified in preschizophrenic individuals as early as the first grade (Crow, Done, & Sackler, 1995). Cognitive impairment appears to be stable within patients over time, across various follow-up intervals, and to predict outcome in the illness (Gold & Harvey, 1993). It should also be noted that there are many changes in cognitive functioning with aging in the normal population. While the domain of aging-related changes is enormous and beyond the scope of this chapter, several general points can be made. Major aging-related changes occur in motor speed and in memory functions, while the level of change in long-term memory and ability to access and utilize previously learned material is much more modest. As a consequence, baseline functioning can be estimated from performance on tests of old learning such as reading and vocabulary, because of the minimal aging-related changes in these areas. For a complete review of aging and neuropsychology, the reader is referred to Albert and Moss (1988). Since schizophrenic patients at their first episode often score 2 to 3 standard deviations below normal comparison groups (Saykin et al., 1994) or standardization samples, the possibility of detecting cognitive decline with aging is reduced because of floor effects. At the same time, it should be noted that the types of cognitive operations that change with normal aging (new learning and recall, concentration and attention, psychomotor and visuomotor speed) are the deficits most commonly noted in schizophrenic patients early in life. Thus, aging-related changes are likely to maximally impact those functions most impaired in schizophrenia earlier in life.

In younger schizophrenic patients, the typical assessment of cognitive functioning has either used assessment with clinical neuropsychological tests or with experimental measures of attention and memory. Across these studies, significant impairments have been found in learning and memory, attention, and executive functioning (see Gold & Harvey, 1993 for a review of these findings in younger patients), with lesser impairments noted in most other cognitive functions (Braff et al., 1991). In older schizophrenic patients, prominent cognitive impairments resembling dementia have also been reported. For example, in the Davidson et al. (1995) study, over 60% of the geriatric chronic schizophrenic patients would have met *DSM-IIIR* criteria for dementia in addition to schizophrenia. Similar findings were reported by Arnold, Franz, and Trojanowski (1994) in a separate sample. As demonstrated below, one of the issues that has been addressed in research on cognitive functioning is a comparison of the cognitive performance profiles of geriatric patients with schizophrenia, both early and late-onset, to the cognitive impairments seen in degenerative dementing conditions such as Alzheimer's disease (AD). Spurred by preliminary reports of an elevated prevalence of AD in postmortem brain specimens from patients who have a hospital chart diagnosis of schizophrenia (Prohovnik, Dwork, Kaufman, & Wilson, 1993), a number of studies have compared the cognitive profiles of patients with these two disorders.

One of the problems that develops in the assessment of geriatric patients is that of test selection. Patients who meet criteria for dementia may not be able to meaningfully complete many clinical neuropsychological tests, with their scores reflecting floor effects (Harvey, Mohs, & Davidson, 1993). One of the solutions that has been employed (Harvey et al., 1992) is the use of global screening measures, such as the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975 or a global assessment measure such as the Clinical Dementia Rating (CDR; Berg, 1988). These instruments have the advantage of being able to provide a global estimate of the level of cognitive functioning of the individual. They are limited by their focus on the cognitive impairments seen in degenerative dementia and their failure to include many of the content areas that are relevant to schizophrenia, such as executive functioning.

Studies of global estimates of cognitive functioning in patients with late-life schizophrenia have suggested that as many as 60% of geriatric chronically hospitalized patients have MMSE scores in the range suggestive of dementia. Davidson et al. (1995) used a cross-sectional design to compare MMSE scores in chronically hospitalized patients ranging in age from 25 to 95. There was a consistent linear relationship between MMSE scores and age, with the mean MMSE of patients who were aged 25 to 35 at 27 ("questionable impairment") and the mean MMSE of 85 to 95-year-olds only 9.6 ("severe dementia"). Even when low levels of premorbid education and age are considered (Crum, Anthony, Bassett, & Folstein, 1993), these scores are extremely low. There was a linear 3 MMSE points per decade reduction over the age-span of the patients studied. In contrast, in AD MMSE scores change an average of 3 points per year (Morris et al., 1993). When the geriatric patients who were still in the hospital were followed up 1 and 2 years later (Harvey, White et al., 1995) there was no measurable change in MMSE scores for the sample as a whole, the IntraClass Correlation Coefficient (ICC) was over 0.85 at each assessment, and one-month stability of the MMSE was found to be over 0.90 (ICC). When the sample of patients in the Davidson et al. sample was compared with similar-age, long-stay patients from the UK, there were no differences in their MMSE scores (Harvey, Leff, Trieman, Anderson, & Davidson, in press). In addition, MMSE scores of chronic schizophrenic patients in other US hospitals are strikingly similar across institutions (Gur et al., 1996).

These data suggest that globally estimated cognitive functions appear to manifest age-related differences, although these scores do not decline at a rate consistent with a degenerative dementia such as AD (Mohs, 1995). Further evidence for dissimilarity to AD was provided by Harvey, Lombardi et al.(1996), who used a battery of neuropsychological tests that was developed to assess the nature and progression of cognitive impairments in AD. At a one-year follow-up, none of these cognitive functions showed any evidence of decline. In addition, the profile of impairment was quite different from that seen in AD. A delayed recall memory deficit is the first reliable sign of mild AD (Welsh et al., 1991), with this deficit not progressing over time and being uncorrelated with overall MMSE scores (Morris et al., 1993). In the Harvey, Lombardi et al. study, delayed recall deficits were correlated with MMSE scores more strongly than any other measure. Further evidence of the dissimilarity of cognitive performance profiles in AD and geriatric schizophrenia is provided by the results of a cross-sectional study comparing late and early onset geriatric schizophrenic patients, nongeriatric schizophrenic patients, patients with AD, and normal controls on a expanded version of the Halstead—Reitan neuropsychological battery (Heaton et al., 1994). Patients with AD were found to perform worse than all schizophrenic patients on a measure of delayed recall, with the schizophrenic patients performing worse on several other summary measures. A very similar result was obtained by Davidson et al. (1996), who compared schizophrenic and AD patients who were matched on global cognitive impairment as measured by MMSE scores. Patients with AD were more impaired on delayed recall and less impaired on verbal and visuo-motor skills than schizophrenic patients who had similar MMSE scores. No differences in rate of learning were detected. Both of these studies suggest that schizophrenic patients are differentially impaired compared with patients with AD and both were careful to rule out alternative causes of cognitive impairment in their samples.

The results of these studies, combined with prospective studies of postmortem characteristics of geriatric schizophrenic patients failing to find evidence of AD-related neuropathology in the majority of cognitively impaired patients with schizophrenia (Arnold et al., 1995; Powchik et al., in press; Purohit et al., 1993), can be interpreted to conclusively demonstrate that the profound cognitive impairments seen in geriatric long-stay schizophrenic patients are not always due to degenerative dementia. What is considerably less clear, are the characteristics of cognitive impairment in geriatric schizophrenic patients and any possible differences between late- and early-onset cases, as well as the associations between outcome and cognitive impairments. In the Heaton et al. (1994) study, no differences in global neuropsychological summary scores were found between similar-aged early- and late-onset geriatric schizophrenic patients and no cognitive performance differences were found as a function of age in the schizophrenic sample. Similarly, Hyde et al. (1994) reported no age-related differences in cognitive functions in a cross-sectional study of schizophrenic patients ranging in age from the 20s to the late 60s. In contrast, several other studies of a large sample of extremely chronic institutionalized patients have found major age-related differences in both global indices of cognitive functions (Davidson et al., 1995) and in specific neuropsychological measures (Harvey, Leibman et al., in press; Harvey, Lombardi et al., 1995).

In terms of the specific characteristics of cognitive impairments seen in geriatric schizophrenic patients, Harvey, Powchik, Mohs, and Davidson (1995) compared geriatric chronic schizophrenic patients with either absent-mild or moderate cognitive impairments, defined by global ratings of a general clinical severity measure for dementia, on a battery of memory tests that have been demonstrated in brain-injured populations to be differentially sensitive to dysfunctions in the temporal and frontal lobes. They found that there were large differences in memory performance associated with global cognitive impairments and that patients, regardless of their level of cognitive impairment, tended to manifest patterns of performance consistent with either a temporal-lobe or frontal-lobe memory deficit. Paulsen, Heaton, Sadek, Perry, and Jeste (1995) also performed a comprehensive neuropsychological evaluation of profiles of memory impairment in schizophrenic patients, dividing their geriatric patients on the basis of cortical, subcortical, or unimpaired profiles. Cortical profiles include deficits in learning, recall, and recognition memory, while subcortical profiles include preserved recognition memory and increased benefits from cuing and encoding assistance. They found about one-third of the patients manifested an unimpaired profile, with 50% manifesting a subcortical profile. Only 15% of the patients manifested a cortical pattern of impairment. These results are quite consistent with the findings of Harvey, Powchik et al., who found that recognition memory performance did not vary across patients who differed considerably in their global level of cognitive impairment. In a study of verbal fluency performance, geriatric chronically hospitalized schizophrenic patients were found to be more impaired on phonemic fluency than younger chronic patients, with no age-related differences in category-based fluency when education differences between these two samples were considered (Harvey, Leibman et al., in press). Both groups of patients performed at least 2 to 3 standard deviations below age and education-corrected mean on category fluency and 1.5 to 2 standard deviations below the mean on phonemic fluency. An additional finding of some importance is that estimates of premorbid intellectual attainment (i.e., the Wide Range Achievement Test-Revised [WRAT-R; Jastak, 1984] reading subtest) were correlated with overall cognitive status in geriatric patients, but not nongeriatric patients. This suggests a possible bias effect that should be considered in later studies, as the WRAT-R or other similar tests are often employed as estimates of premorbid cognitive status in studies of cognitive functions in younger schizophrenic patients (O'Carroll, Walker, & Dunan, 1992).

In an examination of the long-term effects of psychosurgery in chronic schizophrenia, Harvey et al. (1993) studied 72 chronically hospitalized schizophrenic patients on a neuropsychological battery, comparing patients who had experienced a leukotomy with a sample matched for age, education, and overall cognitive status who had not. Consistent with earlier studies of leukotomized patients at younger ages (see Joschko, 1986 for a review), no major differences in cognitive functioning were found. These data suggest that a frontal lobe lesion that may disconnect the cortex and striatum does not interact with aging in chronic schizophrenia to produce even greater cognitive impairment, although there are some interpretive problems with the results. The geriatric chronic schizophrenic patients without leukotomy performed so poorly on tests sensitive to frontal lobe lesions, including the Wisconsin Card Sorting Test (WCST; Heaton, Chelune, Talley, Kay, & Curtiss, 1993), that it would be almost statistically impossible for the leukotomized patients to perform more poorly and reveal a differential performance deficit.

The results of these studies highlight the problem in studying cognitive performance in schizophrenic patients across wide ranges of age, chronicity, and severity of illness. Some of the inconsistent results may simply be the result of subject selection processes. The subjects in the studies conducted by Heaton and colleagues (Heaton et al., 1994; Paulsen et al., 1995) were largely outpatients residing in the community, in contrast to the subjects we studied, who had been continuously hospitalized for decades. Since more severe cognitive impairment is a powerful correlate of poor outcome in younger schizophrenic patients (Addington, Addington, & Maticka-Tyndale, 1991; Perlick, Mattis, Stastny, & Teresi, 1992), there is no reason to expect that extremely bad-outcome geriatric patients would not also have more severe impairments than those patients who are functioning adaptively in community. In the Hyde, Nawroz, Goldberg, Kleinman, and Weinberger (1994) study, relatively poor-outcome patients were studied, but younger and older patients were matched in WRAT-R scores in order to equalize premorbid intellectual functioning. As noted above, if WRAT-R scores are correlated with cognitive impairment in geriatric patients, matching on this variable would create a sample of geriatric patients who are preselected for reduced levels of cognitive impairment.

The previous research has revealed several important cohort effects when younger and older chronically hospitalized patients are compared. Geriatric chronic patients are consistently found to have a later age of first psychiatric admission and lower levels of formal education than younger patients (Arnold et al., 1995; Davidson et al., 1995). These differences have the potential of obscuring the interpretation of differences in cognitive functioning between the samples, because education and onset age have both been previously demonstrated to influence performance on cognitive tests. Performance differences between early and late-onset patients have the potential of being influenced by the correlates of chronic schizophrenia, including treatment and lifelong histories of social isolation, ostracism, and failure.

One area where age effects were not detected was in a set of studies examining the prevalence of age disorientation in chronic schizophrenia and affective disorder (Harvey, Lombardi et al., 1995; Lombardi et al., 1996). Age disorientation has been studied extensively by Crow and colleagues (Crow & Stevens, 1978; Liddle & Crow, 1984), with the consistent finding that approximately 25 to 33% chronic schizophrenic patients underestimate their age by 5 years or more. This finding has been demonstrated not to be caused by history of treatment, premorbid functioning, length of current hospital stay, age of onset, or age of the patients (Buhrich, Crow, Johnstone, & Owens 1988). Harvey, Lombardi et al. found that the prevalence of age disorientation was relatively stable, at approximately 40% from age 35 to age 75 in a sample of 393 chronically hospitalized schizophrenic patients. Age-related differences in cognitive functioning were found only in the patients who manifested age disorientation. Lombardi et al. found a similar rate of age disorientation (33%) in chronically hospitalized patients with affective disorders. These data suggest that one of the principal features of the cognitive impairment in schizophrenia may be orientation deficits. These deficits may be caused by impairments in episodic memory functions, which lead to faulty updating of personal information, in the same way that these memory impairments lead to poor performance on formal tests of memory functioning. Since only a portion of the patients manifest age disorientation, it cannot be due to some global factor associated with institutionalization. Furthermore, age disorientation was not associated, in either schizophrenic or affective patients, with history of premorbid functioning, including education, or history of somatic treatment, including leukotomy, insulin coma, or ECT, replicating previous findings (Buhrich et al., 1988). These data suggest that certain types of memory impairment may be correlated with poor outcome, regardless of diagnosis, and also clearly indicate that not all of the aspects of cognitive impairment in geriatric institutionalized schizophrenic patients could be attributed to aspects of institutional stay.

The main findings associated with the study of cognitive functioning in late-life schizophrenia is that there is very little evidence for the development of types of cognitive impairments that are qualitatively different from those seen in younger schizophrenic patients. While some studies find no age-related differences in cognitive performance, others do, with the principal variation between the studies being the nature of subject samples. The primary change appears to be a worsening of memory performance, in line with expected changes associated with normal aging, although this worsening still does not lead to performance as poor as seen in Alzheimer's disease. The association between extremely poor outcome and profound cognitive impairment leads to interpretive problems because of the very long length of stay and extensive treatment experienced by these patients. Since only a subset of these patients manifest age disorientation or other profound cognitive impairments, it is not possible that all aspects of their impairment are caused by institutional stay. Assuming that institutional stay and its correlates (e.g., medication treatment) have the potential to cause some of the cognitive impairments seen in geriatric chronically institutionalized patients, the variability across patients in level of impairment suggests that there may be individual difference factors that predict vulnerability to these effects. Identification of these factors is a useful target for further research.

7.13.9 ADAPTIVE IMPAIRMENT IN LATE LIFE SCHIZOPHRENIA

One of the principal sources of negative outcome in schizophrenia is a deficit in adaptive life functioning (Perlick, Stastny, Mattis, & Teresi, 1992; Serban & Gidynski, 1979; Tessler & Manderscheid, 1994). Many schizophrenic patients manifest severe and persistent disability because they are unable to care for themselves and become dependent on others for their care. Some of these impairments are more often present in patients with the negative symptoms (Mueser, Bellack, Douglas, & Wade, 1991) of schizophrenia as well as in patients with more severe cognitive impairment (Breier, Schreiber, Dyer, & Pickar, 1991). Patients who are dependent on others for their care are typically not more seriously affected by the positive symptoms of the disorder (Pogue-Geile & Harrow, 1985). Previous behavioral interventions in schizophrenia have often focused on remediation of these deficits in adaptive life functions (Hogg & Hall, 1992; Liberman, Mueser, & Wallace, 1986; Paul & Lenz, 1977), with some suggestions that pharmacological interventions should also be targeted at these types of deficits (Davidson & Keefe, 1995). As noted above, in younger chronic schizophrenic patients the presence of severe cognitive impairments is often correlated with the presence of these adaptive deficits (Green, 1996).

Limited information is available on adaptive functioning deficits in geriatric patients with schizophrenia. The available information suggests that the range of adaptive functioning is quite wide and that the level of care received by the patient is often correlated with their adaptive skill. Kincaid et al. (1995) studied 339 geriatric chronically hospitalized inpatients with schizophrenia, finding that clinical ratings of self-care and personal hygiene accounted for 9% of the variance in global ratings of severity of overall functional, cognitive, and behavioral impairment. Harvey, Davidson et al. (1997) developed the Social-Adaptive Functioning Evaluation (SAFE) scale, which was designed to evaluated overall adaptive dysfunction in geriatric institutionalized patients. They reported that total scores on the measure were strongly associated with the overall severity of cognitive impairment, as measured by the MMSE. In addition, the overall level of impairment was quite high in their population. For example, very severe impairments were found, on the average, for neatness and maintenance, money management, orientation and mobility, communication skills, and friendships, while impairments at a level that was considered "severe" were found on nine other measures of social, instrumental, and self-care skills. When a sample of 140 patients were followed prospectively for up to 3 years, it was found that deficits in the general areas of impulse control and self-care were quite predictive of discharge and placement. Patients with poor impulse control were likely to remain hospitalized, while patients with better self care were more likely to be discharged to a community setting. Social functioning did not discriminate between three different placement outcomes: remaining hospitalized, being referred to the community, or being transferred to a nursing home. In a study of over 800 geriatric psychiatric patients followed for 3 years during a period of downsizing of their psychiatric hospital, White et al. (in press) found that the presence of severe adaptive deficits was less strongly correlated with discharge from chronic care than the presence of hostile and belligerent behaviors. In patients who were discharged, however, severe deficits in adaptive functioning were more common in patients referred to a nursing home than in those discharged to the community.

In another study examining geriatric psychiatric patients residing outside of chronic hospitals and receiving community-based residential care, Bartels et al. (1997) evaluated patients with schizophrenia, major depression, and bipolar affective disorder. The severity of clinical symptoms was evaluated with the Brief Psychiatric Rating Scale-Expanded version (Lukoff, Liberman,& Nuechterlein, 1986) and with the Specific Levels of Functioning Scale (SLOF; Schneider & Streuening, 1983). The SLOF is a 43-item rating scale that comprehensively evaluates social, self-care, and community living skills, as well as disruptive behaviors. Schizophrenic patients, in addition to having more severe psychotic symptoms, were more impaired on all aspects of social and instrumental functioning than community-dwelling patients with affective disorders. These data suggest that the impairments in adaptive skills that characterize younger patients with schizophrenia are also present in late life, even in community-dwelling cases. Thus, severe adaptive deficits are present in many community-dwellers as well as in chronically hospitalized patients with schizophrenia.

In contrast, the results of the Vermont longitudinal study (Harding et al., 1987a, 1987b, 1992) suggest that some geriatric individuals with a history of schizophrenia are able to successfully adapt to life in the community and live there for extended periods without ever receiving any mental health services. Early identification of which patients are likely to have this type of adaptive success is not yet possible. When this constellation of characteristics predicting adaptive success is identified, early interventions for patients with a poor prognosis will be easier to target.

7.13.10 TREATMENT OF SCHIZOPHRENIA IN LATE LIFE

There is no specialized literature on the treatment of late-life schizophrenia with psychological interventions. At the same time, there is no information suggesting that current psychological treatments would be less effective in this type of population. Principal psychological interventions for schizophrenia with demonstrated efficacy are skills-based "rehabilitation" approaches and family intervention approaches designed to reduce expressed emotion (EE). Since EE approaches are applied to acutely ill patients, their applicability to geriatric schizophrenic patients with an early age of onset is limited. No specific EE research for late-life acute schizophrenia has been published up to 1997, although this may be a critical area for investigation. In particular, since many late-onset schizophrenic patients are referred to nursing home care directly from an acute psychiatric admission, assessment of EE characteristics at home may provide useful information in identifying those acutely ill late-onset patients who are at high risk to be referred to nursing home care because of negative aspects of their home life.

In contrast, skills-training approaches appear to be indicated for many of the geriatric patients (Bellack & Mueser, 1993). Since deficits in adaptive functioning are apparently extremely common in institutionalized and community-dwelling geriatric populations, this intervention would appear, at least superficially, to be likely to provide benefit to many of these patients. It is unclear how many of these geriatric patients have received such treatment earlier in their illness, without benefit, and how many have simply received "warehouse" institutional treatment or have been "dumped" into the community without follow-up plans. Previous studies suggest that even very severely ill younger patients may benefit from this treatment (Corrigan, Wallace, Schade, & Green, 1994), indicating that this population could be targeted for such an intervention.

A cautionary note is provided by the clear description by Green (1996) of cognitive barriers to such treatment. Since these interventions are essentially learning-based, deficits in the ability to learn with exposure, in performance on tests of vigilance, and in executive functioning have been demonstrated to markedly interfere with the success of these interventions in younger chronic schizophrenic patients. Since many geriatric patients with a chronic course of institutionalization manifest profound deficits in learning and memory, as described above, these patients would appear to be at high risk of failure in rehabilitation-type treatments (Mueser et al., 1991). Until techniques for cognitive enhancement are better supported with empirical research (see Vol. 18 (1992) of *Schizophrenia Bulletin* for a discussion of limitations in this regard), the application of social-instrumental skills-training approaches with geriatric patients with cognitive impairments appears to be quite likely to be unsuccessful.

Psychotropic medication for geriatric patients with schizophrenia is the most common treatment. While these interventions are not presently in the domain of clinical psychology, understanding of the issues involved in this type of treatment is important for any professional involved in the treatment of this population. As with many other drugs, dosing of geriatric patients is most useful at lower levels, a fact of which some treating physicians are apparently unaware. Geriatric patients are apparently more susceptible to the side effects of these medications as well. The reader is referred to a review paper by Jeste, Lacro, Gilbert, Kline, and Kline (1993) which addresses these issues, and others, in detail.

7.13.11 CONCLUSION

Schizophrenia is a lifelong illness that does not disappear with aging. Although some patients with schizophrenia have a good outcome in late life, the majority, both institutionalized and community dwelling, manifest severe deficits in adaptive functions. Clinical symptoms appear to be severe in most patients studied, with the level of symptoms seen across care settings arguing strongly against the idea that these symptoms burn out with age. Cognitive impairment appears to be a strong predictor of adaptive deficits and patients with a particularly poor outcome may actually show cognitive decline into the range of functioning commonly associated with dementia.

Schizophrenia appears to have a fairly high new case incidence in late life, with as many as 15% of all geriatric cases having had initial onset after age 65. These patients appear to resemble good-prognosis early-onset cases, other than for the high prevalence of females. Data on specific psychological treatments for late-life schizophrenic patients are completely lacking and are clearly necessary. As is clear from the data in this chapter, schizophrenia in late life is not benign and is associated with considerable disability in adaptive functioning and a clear burden on society or relatives. Development of effective treatments, psychological and pharmacological, for late-life schizophrenia is a clear goal for the future.

7.13.12 REFERENCES

Addington, J., Addington, D., & Maticka-Tyndale, E. (1991). Cognitive functioning and positive and negative symptoms in schizophrenia. *Schizophrenia Research,4*, 123–134.

Albert, M. S., & Moss, M. B. (1988). *Geriatric neuropsychology.* New York: Guilford.

Albus, M., & Maier, W. (1995). Lack of gender differences at age of onset in familial schizophrenia. *Schizophrenia Research. 18*, 51–57.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd. Ed.). Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd Ed., revised). Washington, DC: Author.

American Psychiatric Association (1993). *Diagnostic and statistical manual of mental disorders* (4th. Ed.). Washington, DC: Author.

Andreasen, N. C., Arndt, S., Alliger, R., Miller, D., & Flaum, M. (1995). Symptoms of schizophrenia: Methods, meanings, and mechanisms. *Archives of General Psychiatry, 52,* 341–351

Arnold, S. E., Franz, B. R., & Trojanowski, J.Q. (1994). Elderly patients with schizophrenia exhibit infrequent neurodegenerative lesions. *Neurobiology of Aging, 15,* 299–303.

Arnold, S. E., Gur, R. E., Shapiro, R. M., Fisher, K. R., Moberg, P. J., Gibney, M. R., Gur, R. C., Blackwell, P., & Trojanowski, J. Q. (1995). Prospective clinicopathological studies of schizophrenia: Accrual and assessment of patients. *American Journal of Psychiatry, 152,* 731–737.

Bartels, S. J., Mueser, K. T., & Miles, K. M. (in press). Functional impairments in elderly patients with schizophrenia and major affective disorders living in the community: Social skills, living skills, and behavior problem. *Behavior Therapy.*

Bellack, A. S., & Mueser, K. T. (1993). Psychosocial treatment for schizophrenia. *Schizophrenia Bulletin,19*, 317–336.

Berg, L. (1988). Clinical dementia rating (CDR). *Psychopharmacology Bulletin,24*, 637–639.

Blazer, D. G. (1996). Epidemiology of psychiatric disorders in late life. In E. W. Busse & D. G. Blazer (Eds), *American Psychiatric Press handbook of geriatric psychiatry* (2nd ed., pp 155–171). Washington, DC: American Psychiatric Press.

Bleuler, E. (1911/1950). *Dementia praecox; or the group of schizophrenias.* New York: International Universities Press.

Braff, D. L., Heaton, R. K., Kuck, J., Cullum, M., Moranville, J., Grant, I., & Zisook, S. (1991). The generalized pattern of neuropsychological deficits in outpatients with chronic schizophrenia with heterogenous Wisconsin Card Sorting Test results. *Archives of General Psychiatry, 48,* 891–898.

Breier, A., Schreiber, J. L., Dyer, J., & Pickar, D. (1991). National Institute of Mental Health longitudinal study of chronic schizophrenia: Prognosis and predictors of outcome. *Archives of General Psychiatry,48*, 239–246.

Bridge, T. P., Cannon, E. H., & Wyatt, R. J. (1978). Burned-out schizophrenia: Evidence for age effects on schizophrenic symptomatology. *Journal of Gerontology, 33,* 835–839.

Buhrich, N., Crow, T. J, Johnstone, E. C., & Owens, D. G. (1988). Age disorientation in chronic schizophrenia is not associated with pre-morbid intellectual impairment or past physical treatment. *British Journal of Psychiatry, 152,* 466–469.

Castle, D. J., & Murray, R. M. (1993). The epidemiology of late-onset schizophrenia. *Schizophrenia Bulletin, 19,* 691–700.

Christenson, R., & Blazer, D. (1984). Epidemiology of persecutory ideation in an elderly population in the community. *American Journal of Psychiatry,141,* 1088–1091.

Ciompi, L. (1980). Catamnestic long-term study on the course of life and aging of schizophrenics. *Schizophrenia Bulletin, 6,* 606–618.

Corrigan, P. W., Wallace, C. J., Schade, M. L., & Green, M. F. (1994). Learning medication self-management skills in schizophrenia: Relationships with cognitive deficits and psychiatric symptoms. *Behavior Therapy, 25,* 5–15.

Crow, T. J., Done, D. J., & Sackler, A. (1995). Childhood precursors of psychosis as clues to its evolutionary origins. *European Archives of Psychiatry and Clinical Neuroscience, 245,* 61–69.

Crow, T. J., & Stevens, M. (1978). Age disorientation in chronic schizophrenia: The nature of the cognitive deficit. *British Journal of Psychiatry,133,* 137–142.

Crum, R. M., Anthony, J. C., Bassett, S. S., & Folstein, M. F. (1993). Population based norms for the mini-mental state examination by age and education level. *Journal of the American Medical Association, 269,* 2386–2391.

Davidson, M., Harvey, P. D., Powchik, P., Parrella, M., White, L., Knobler, H. Y., Losonczy, M., Keefe, R. S. E., Katz, S., & Frecksa, E. (1995). Severity of symptoms in chronically institutionalized geriatric schizophrenic patients. *American Journal of Psychiatry, 152,* 197–207.

Davidson, M., Harvey, P. D., Putnam, K., Welsh, K., & Mohs, R. C. (1996). Cognitive impairment in old-age schizophrenia: A comparative study of schizophrenia and Alzheimer's disease. *American Journal of Psychiatry, 153,* 1274–1279.

Davidson, M., & Keefe, R. S. E. (1995). Cognitive impairment as a target for pharmacological treatment in schizophrenia. *Schizophrenia Research, 17,* 123–129.

Faraone, S. V., Chen, W. J., Goldstein, J. M., & Tsuang, M. T. (1994). Gender differences in age at onset of schizophrenia. *British Journal of Psychiatry,164,* 625–629.

Flint, A. J., Rifat, S. L., & Eastwood, R. (1991). Late-onset paranoia: Distinct from paraphrenia. *International Journal of Geriatric Psychiatry, 6,* 103–109.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Green, M. F. (1996). What are the functional consequences of neurocognitive deficits in schizophrenia? *American Journal of Psychiatry, 153,* 321–330.

Gold, J. M., & Harvey, P. D. (1993). Cognitive deficits in schizophrenia. *Psychiatric Clinics of North America, 16,* 295–312.

Gur, R. E., Arnold, S. E., Moberg, P. J., Petty, R. G., Arnold, S., & Gur, R. C. (1996). The dementia of schizophrenia: Antemortem and postmortem studies. *Biological Psychiatry, 39,* 612.

Hafner, H. & Hambrecht, M. (1994). The elderly with schizophrenia. In E. Chiu & D. Ames (Eds.),*Functional psychiatric disorders of the elderly*. (pp. 287–302). Cambridge, UK: Cambridge University Press.

Harding, C. M., Brooks G. W., Ashikaga T., Strauss, J. S., & Breier, A. (1987a). The Vermont longitudinal study of persons with severe mental illness: I. Methodology, study sample, and overall current status 32 years later. *American Journal of Psychiatry,144,* 718-726.

Harding, C. M., Brooks, G. W., Ashikaga, T., Strauss, J. S., & Breier, A. (1987b). The Vermont longitudinal study of persons with severe mental illness: II. Long term outcome of subjects who retrospectively met *DSM-III* criteria for schizophrenia. *American Journal of Psychiatry, 144,* 727–735.

Harding, C. M., Zubin, J., & Strauss, J. S. (1992). Chronicity in schizophrenia: revisited. *British Journal of Psychiatry, 161* (Suppl.18), 27–37.

Harris, M. J., & Jeste, D. V. (1988). Late-onset schizophrenia: An overview. *Schizophrenia Bulletin, 14*, 39–55.

Harris, M. J., Jeste, D. V., Krull, A., Montague, J., & Heaton, R. K. (1991). Deficit syndrome in older schizophrenic patients. *Psychiatry Research, 39,* 285–292.

Harvey, P. D., Davidson, M., Mueser, K., Parrella, M., White, L., & Powchik, P. (1997). The social adaptive functioning evaluation: An assessment measure for geriatric psychiatric patients. *Schizophrenia Bulletin, 23,* 131–146.

Harvey, P. D., Davidson, M., Powchik, P., Parella, M., White, L., & Mohs, R. C. (1992). Assessment of dementia in elderly schizophrenics with structured rating scales. *Schizophrenia Research, 7,* 85–90.

Harvey, P. D., Leff, J., Trieman, N., Anderson, J., & Davidson, M. (in press). Cognitive impairment in geriatric chronic schizophrenic patients. A cross national study in New York and London. *International Journal of Geriatric Psychiatry.*

Harvey, P. D., Leibman, M., Lombardi, J., Parrella, M., White, L., Powchik, P., Mohs, R. C., & Davidson, M. (in press). Verbal fluency in geriatric and nongeriatric chronic schizophrenic patients. *Journal of Neuropsychiatry and Clinical Neurosciences.*

Harvey, P. D., Lombardi, J., Kincaid, M., Parrella, M., White, L., Powchik, P., & Davidson, M. (1995). Cognitive functioning in chronically hospitalized schizophrenic patients: Age-related changes and age disorientation as a predictor of impairment. *Schizophrenia Research, 17,* 15–24.

Harvey, P. D., Lombardi, J., Leibman, M., White, L., Parrella, M., Powchik, P., Mohs, R. C., & Davidson, M. (1996). Performance of geriatric chronic schizophrenic patients on cognitive neuropsychological measures sensitive to dementia. *International Journal of Geriatric Psychiatry, 11,* 621–627.

Harvey, P. D., Lombardi, J., Leibman, M., Parrella, M., White, L., Powchik, P., Mohs, R. C., Davidson, M., & Davis, K. L. (1997). Age-related changes in formal thought disorder in chronic schizophrenic patients: A cross-sectional study across 9 decades. *American Journal of Psychiatry, 154,* 205–210.

Harvey, P. D., Mohs, R. C., & Davidson, M. (1993). Leukotomy and aging in chronic schizophrenia: A followup study 40 years after psychosurgery. *Schizophrenia Bulletin, 19,* 723–732.

Harvey, P. D., Powchik, P., Mohs, R. C., & Davidson, M. (1995). Memory functions in geriatric chronic schizophrenic patients: A neuropsychological study. *Journal of Neuropsychiatry and Clinical Neurosciences, 7,* 207–212.

Harvey, P. D., White, L., Parrella, M., Putnam, K. M., Kinkaid, M. M., Powchik, P., Mohs, R. C., & Davidson, M. (1995). The longitudinal stability of cognitive impairment in schizophrenia: Mini-mental state scores at one and two year follow-ups in geriatric inpatients. *British Journal of Psychiatry, 166,* 630–633.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. C., & Curtiss, G. (1993). Wisconsin Card Sorting Test (WCST) manual-revised and expanded. Odessa, FL: Psychological Assessment Resources.

Heaton, R., Paulsen, J. S., McAdams, L. A., Kuck J., Zisook, S., Braff, D., Harris, M. J., & Jeste, D. V. (1994). Neuropsychological deficits in schizophrenics. Relationship to age, chronicity, and dementia. *Archives of General Psychiatry, 5,* 469–476.

Hogg, L., & Hall, J. (1992). Management of long-term impairments and challenging behaviour. In M. J. Birchwood & N. Tarrier (Eds), *Innovations in the psychological management of schizophrenia: Assessment, treatment, and services* (pp. 171–203). Chichester, UK: Wiley.

Hyde, T. M., Nawroz, S., Goldberg, T. E., Kleinman, J. E., & Weinberger, D. R. (1994). Is there cognitive decline in schizophrenia? A cross-sectional study. *British Journal of Psychiatry, 164,* 494–500.

Jastak, S. (1984). The wide-range achievement test-revised. Wilmington DE: Jastak Associates.

Jeste, D. V., Harris, M. J., Krull, A., Kuck, J., McAdams, L. A., & Heaton, R. (1995). Late onset schizophrenia: Clinical and neuropsychological characteristics. *American Journal of Psychiatry, 152,* 722–730.

Jeste, D. V., Lacro, J. P., Gilbert, P. L., Kline, J., & Kline, N. (1993). Treatment of late-life schizophrenia with neuroleptics. *Schizophrenia Bulletin, 19,* 817–830.

Jeste, D. V., Wragg, R. E., Salmon, D. P., Harris, M. J., & Thal, L. J. (1992). Cognitive deficits of patients with Alzheimer's disease with and without delusions. *American Journal of Psychiatry, 149,* 184–189.

Joschko, M. (1986). Clinical and neuropsychological outcome following psychosurgery. In I. Grant & K. Adams (Eds), *Neuropsychological assessment in neuropsychiatric disorder* (pp. 300-320) New York: Oxford University Press.

Kay, S. R. (1991). *Positive and negative syndromes in schizophrenia.* New York: Brunner/Mazel.

Keefe, R. S. E., & Harvey, P. D. (1994). *Understanding schizophrenia*. New York: Free Press.

Keefe, R. S., Mohs, R. C., Losonczy, M. F., Davidson, M., Silverman, J. M., Kendler, K. S., Horvath, T. B., Nora, R., & Davis, K. L. (1987). Characteristics of very poor outcome schizophrenia. *American Journal of Psychiatry, 144,* 889–895.

Kincaid, M. M., Harvey, P. D., White, L., Parrella, M., Powchik, P., & Davidson, M. (1995). Validity and utility of the Alzheimer's Disease Assessment Scale-Late Version (ADAS-L) for the measurement of cognitive and functional impairment in geriatric chronic schizophrenic in patients. *Journal of Neuropsychiatry and Clinical Neuroscience, 7,* 76–81.

Knight, R. A., Roff, J. D., Barrnet, J., & Moss, L. (1979). Concurrent and predictive validity of thought disorder and affectivity: A 22 year followup of acute schizophrenics. *Journal of Abnormal Psychology, 88,* 1–12.

Kraepelin, E. (1919). *Dementia praecox and paraphrenia*. Edinburgh, UK: Livingstone.

Leff, J. P., Thornicroft, G., Coxhead, N., & Crawford, C. (1994). The TAPS Project. 22: A five-year follow-up of long stay psychiatric patients discharged to the community. *British Journal of Psychiatry, 165* (Suppl 25), 13–17.

Lewine, R. R. J. (1981). Sex differences in schizophrenia: Timing or subtypes? *Psychological Bulletin, 90,* 432–444.

Liberman, R. P., Mueser, K. T., & Wallace, C. J. (1986). Social skills training for schizophrenic individuals at risk for relapse. *American Journal of Psychiatry, 143,* 523–526.

Liddle, P. F., & Crow, T. J. (1984). Age disorientation in chronic schizophrenia is associated with global intellectual impairment. *British Journal of Psychiatry, 144,* 193–199.

Loebel, A. D., Lieberman, J. A., Alvir, J. M., Mayerhoff, D. I., Deisler, S. H., & Szymanski, S. R. (1992). Duration of psychosis and outcome in first episode schizophrenia. *American Journal of Psychiatry, 149,* 1183–1188.

Lombardi, J., Harvey, P. D., White, L., Parrella, M., Powchik, P., & Davidson, M. (1996). Age disorientation in chronically hospitalized patients with mood disorders. *Psychiatry Research, 60,* 87–90.

Lukoff, D., Liberman, R. P., & Nuechterlein, K. H. (1986). Symptom monitoring in the rehabilitation of schizophrenic patients. *Schizophrenia Bulletin, 12,* 549–602.

Meeks, S., Carstensen, L. L., Stafford, P. B., Brenner, L. L., Weathers, F., Welch, R., & Oltmanns, T. F. (1990). Mental health needs of the chronically mentally ill elderly. *Psychology and Aging, 5,* 163–171.

Mohs, R. C. (1995). Neuropsychological assessment of patients with Alzheimer's disease In F. E. Bloom & D. J. Kupfer (Eds.), *Psychopharmacology: The fourth generation of progress*. (pp. 1377–1388) New York: Raven Press.

Morris, J. C., Edland, S., Clark, C., Galasko, D., Koss, E., Mohs, R., van Belle, G., Fillenbaum, G., & Hayman, A. (1993). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part IV. Rates of cognitive change in the longitudinal assessment of probable Alzheimer's disease. *Neurology, 43,* 2457–2465.

Mueser, K. T., Bellack, A. S., Douglas, M. S., & Wade, J. H. (1991). Prediction of social skills acquisition in schizophrenic and major affective disorder patients from memory and symptomatology. *Psychiatry Research, 37,* 281–296.

Mueser, K. T., Douglas, M. S., Bellack, A. S., & Morrison, R. L. (1991). Assessment of enduring deficit and negative symptom subtypes in schizophrenia. *Schizophrenia Bulletin, 17,* 565–582.

Mulsant, B. H., Stergiou, A., Keshavan, M. S., et al. (1993). Schizophrenia in late life: Elderly patients admitted to an acute psychiatric care hospital. *Schizophrenia Bulletin, 19,* 709–722.

O'Carroll, R. E, Walker, M., & Dunan, J. (1992). Selecting controls for schizophrenia research studies: The use of the National Adult Reading Test (NART) as a measure of premorbid ability. *Schizophrenia Research, 8,* 137–141.

Parsons, P. L. (1964). Mental health of Swansea's old folk. *British Journal of Preventive and Social Medicine, 19,* 43–47.

Paul, G. T., & Lenz, R. J. (1977). *Psychosocial treatment of chronic mental patients: Milieu versus social-learning programs*. Cambridge, MA: Harvard University Press.

Paulsen, J. S., Heaton, R. K., Sadek, J. R., Perry, W., & Jeste, D. V. (1995). The nature of learning and memory impairments in schizophrenia. *Journal of the International Neuropsychologocal Society, 1,* 88–99.

Perlick, D., Mattis, S., Stastny, P., & Teresi, J. (1992). Neuropsychological discriminators of long-term inpatient or outpatient status in chronic schizophrenia. *Journal of Neuropsychiatry and Clinical Neurosciences, 4,* 428–434.

Perlick, D., Stastny, P., Mattis, S., & Teresi, J. (1992). Contribution of family, cognitive, and clinical dimensions to long-term outcome in schizophrenia. *Schizophrenia Research, 6,* 257–265.

Pogue-Geile, M. F., & Harrow, M. (1984). Negative and positive symptoms in schizophrenia and depression: A followup. *Schizophrenia Bulletin, 10,* 371–387.

Pogue-Geile, M. F., & Harrow, M. (1985). Negative symptoms in schizophrenia: Their longitudinal course and prognostic significance. *Schizophrenia Bulletin, 11,* 427–439.

Powchik, P., Haroutunian, V., Davidson, M., Gabriel, S., Harvey, P. D., & Davis, K. L. (in press). Postmortem studies in schizophrenia. *Schizophrenia Bulletin.*

Prohovnik, I., Dwork, A. J., Kaufman, M. A., & Wilson, N. (1993). Alzheimer-type neuropathology in elderly schizophrenia patients. *Schizophrenia Bulletin, 19,* 805–816.

Purohit, D. P., Davidson, M., Perl, D. P., Powchik, P., Haroutunian, V. H., Bierer, L. M., McCrystal, J., Losonczy, M., & Davis, K. L. (1993). Severe cognitive impairments in elderly schizophrenic patients: A clinicopathological study. *Biological Psychiatry, 33,* 255–260.

Putnam, K. M., Harvey, P. D., Parrella, M., White, L., Powchik, P., & Davidson, M. (1996). Symptom stability in geriatric chronic schizophrenic inpatients: A one-year follow-up study. *Biological Psychiatry, 39,* 92–99.

Rockwood, K., Stolee, P., & Brahim, A. (1991). Outcomes of admission to a psychogeriatric service. *Canadian Journal of Psychiatry, 36,* 275–279.

Saykin, A. J., Shtasel, D. L., Gur, R. E., Kester, D. B., Mozley, L. H., Stafiniak, P., & Gur, R. C. (1994). Neuropsychological deficits in neuroleptic naive patients with first episode schizophrenia. *Archives of General Psychiatry, 51,* 124–131.

Schneider, L. C., & Streuening, E. L. (1983). SLOF: A behavioral rating scale for assessing the mentally ill. *Social Work Research and Abstracts, 6,* 9–21.

Serban, G., & Gidynski, C. B. (1979). Relationship between cognitive defect, affect response, and community adjustment in chronic schizophrenics. *British Journal of Psychiatry, 134,* 602–608.

Spitzer, R., Endicott, J., & Robins, L. (1978). Research diagnostic criteria: Rationale and reliability. *Archives of General Psychiatry, 35,* 773–784.

Tessler, R. C., & Manderscheid, R. W. (1994) Factors affecting adjustment to community living. *Hospital and Community Psychiatry, 33,* 203–207.

Welsh, K. A., Butters, N., Hughes, J., Mohs, R. C., & Heyman, A. (1991). Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. *Archives of Neurology, 48,* 278–281.

White, L., Harvey, P. D., Parrella, M., Sevy, S., Knobler, H., Powchik, P., & Davidson, M. (1994). Empirical assessment of the factorial structure of clinical symptoms in schizophrenic patients: Symptom structure in geriatric and nongeriatric samples. *New Trends in Experimental and Clinical Psychiatry, 10,* 75–83.

White, L., Parrella, M., McCrystal-Simon, J., Masiar, S., Harvey, P. D., & Davidson, M. (1997). Characteristics of elderly psychiatric patients retained in a state hospital during downsizing: A prospective study with replication. *International Journal of Geriatric Psychiatry, 12,* 474–480.

Wyatt, R. J.(1991) Neuroleptics and the natural course of schizophrenia. *Schizophrenia Bulletin, 17,* 325–351.

Yassa, R., Suranyi-Cadotte, B. (1993). Clinical characteristics of late-onset schizophrenia and delusional disorder. *Schizophrenia Bulletin, 19,* 701–707.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.14
Sleep Disorders

SONIA ANCOLI-ISRAEL, RUTH PAT-HORENCYZK, and
JENNIFER MARTIN
*University of California and Veterans Affairs Medical Center,
San Diego, CA, USA*

7.14.1 INTRODUCTION 308
- *7.14.1.1 Measurement of Sleep* 308
  - *7.14.1.1.1 Sleep stages* 308
  - *7.14.1.1.2 Sleep cycles and architecture* 308
- *7.14.1.2 Theories on the Function of Sleep* 309
- *7.14.1.3 Changes in Sleep Associated with Aging* 309
  - *7.14.1.3.1 Objective sleep studies* 309
  - *7.14.1.3.2 Subjective reports* 309
  - *7.14.1.3.3 Gender differences* 310

7.14.2 DISORDERS OF SLEEP 310
- *7.14.2.1 Sleep-disordered Breathing* 310
  - *7.14.2.1.1 Symptoms* 310
  - *7.14.2.1.2 SDB and hypertension* 311
  - *7.14.2.1.3 SDB and cognitive impairment* 311
  - *7.14.2.1.4 Mortality* 311
  - *7.14.2.1.5 Prevalence* 311
  - *7.14.2.1.6 Treatments* 311
- *7.14.2.2 Periodic Limb Movements in Sleep* 312
  - *7.14.2.2.1 Symptoms* 312
  - *7.14.2.2.2 Clinical evaluation* 312
  - *7.14.2.2.3 Prevalence* 313
  - *7.14.2.2.4 Treatments* 313
- *7.14.2.3 Insomnia* 313
  - *7.14.2.3.1 Clinical evaluation* 314
  - *7.14.2.3.2 Etiology of insomnia* 315
  - *7.14.2.3.3 Pharmacological treatments* 315
  - *7.14.2.3.4 Psychological treatments* 316

7.14.3 CIRCADIAN RHYTHM CHANGES WITH AGE 319
- *7.14.3.1 Bright-light Treatment* 319
- *7.14.3.2 Melatonin* 320

7.14.4 DEMENTIA AND SLEEP 320
- *7.14.4.1 Laboratory Studies* 321
- *7.14.4.2 Dementia and Depression* 321
- *7.14.4.3 Nursing Home Patients* 321

7.14.5 SUMMARY 322

7.14.6 REFERENCES 322

7.14.1 INTRODUCTION

Since the 1950s, increasing attention has focused on changes in sleep with advancing age. This chapter reviews what sleep is, how sleep is measured, and the major theories of sleep, and then focuses on the changes that occur in sleep with advancing age. Prevalent sleep disorders found in elderly populations are described, current methods of evaluation and treatments of these disorders discussed, and special emphasis is placed on adapting the interventions to the specific needs of elderly populations.

7.14.1.1 Measurement of Sleep

Sleep is an active process with changes occurring in physiological systems. Traditionally, sleep is studied by recording brain waves (EEG or electroencephalography), eye movement (EOG or electro-oculography), and muscle tension (EMG or electromyography). Information gained from these three channels of information allow the identification of the different stages of sleep. When additional clinical information is necessary, other systems such as heart rate, respiration, level of oxygen in the blood, limb movements, body position, and blood pressure may also be recorded. The recording of these variables is called a polysomnograph (PSG).

For diagnostic purposes, sleep is usually recorded for one full night. Since there may be some adaptation problems to the laboratory environment (referred to as "the first night effect"), for further assessment or for research purposes, sleep may be recorded for additional nights.

There have been many technical advances in the field of sleep with portable PSG now being available. Many of the same variables that are recorded in the laboratory can now be recorded in the patient's own home. Two major advantages of portable PSG are that the patient is more comfortable and relaxed at home and that it is less expensive than laboratory PSG. Much research is being conducted to determine the minimum number of channels of information necessary to screen for specific sleep disorders (Ferber et al., 1994).

An alternative approach for objectively measuring sleep/wake is the wrist actigraph. An actigraph measures wrist movement and significantly correlates with EEG for distinguishing wake from sleep (Sadeh, Hauri, Kripke, & Lavie, 1995). Wrist activity has been shown to provide valid and reliable measurements of the sleep/wake rhythm in elderly patients and has been used for elderly populations who might have difficulty with traditional EEG recordings such as elderly patients with insomnia (Hauri & Wisbey, 1994) and nursing home patients (Ancoli-Israel, Jones, et al., 1991). An additional advantage of wrist activity is the ability to record for consecutive 24-hour periods. The limitation of wrist actigraphy, however, is that it cannot distinguish between different sleep stages.

7.14.1.1.1 Sleep stages

Sleep is traditionally divided into two major states: rapid eye movement (REM) sleep and non-rapid eye movement (NREM) sleep. NREM sleep is further divided into four sleep stages. Each of the stages differs from one another in the amplitude and frequency of the brain waves, in the type of eye movements, and in the presence or absence of muscle tone. The four NREM stages roughly parallel a "depth of sleep" continuum, with arousal thresholds generally increasing from a low threshold in stage 1 to the highest threshold in stage 4. Stage 1 is the transitional phase between wakefulness and sleep and is characterized by low-voltage, mixed-frequency EEG and slow eye movements. Stage 1 sleep usually comprises about 2–5% of the night. Stage 2 sleep is considered the onset of sleep and is characterized by sleep spindles and k-complexes. Stage 2 usually comprises about 45–75% of the night. Stages 3 and 4 sleep are called deep sleep or slow-wave sleep and are characterized by high-amplitude, slow-frequency delta waves. Slow-wave sleep usually comprises 25% of the night in younger adults but almost 0% in the elderly. REM sleep is characterized by low-voltage, high-frequency EEG with minimal or no muscle tone and with rapid eye movements. REM sleep, which is also called dream sleep, comprises 20–25% of the night.

REM sleep is often called paradoxical sleep since it is characterized by high-frequency brain activity which resembles wakefulness. However, the arousal threshold is high (i.e., the subject is difficult to awaken) and there is motor paralysis or muscle atonia. This paralysis is believed to be a protective mechanism which keeps the dreamer from acting out dreams. REM sleep is also characterized by episodic bursts of rapid eye movements and irregular heart rate, respiration, and blood pressure.

7.14.1.1.2 Sleep cycles and architecture

NREM sleep and REM sleep alternate through the night in a cyclic fashion. The individual moves through the four NREM stages, then gradually returns to stage 2 sleep from which REM sleep is entered. NREM sleep generally comprises 75–80% of the sleep period, whereas REM sleep comprises 20–25% of the sleep period.

7.14.1.2 Theories on the Function of Sleep

Although we sleep for about one-third of our lives, the function of sleep remains enigmatic and controversial. The generally held assumption is that sleep is an essential and unavoidable condition. The strongest research evidence for the essential role of sleep comes from long-term sleep deprivation studies. For example, animal studies have shown that after 2–3 weeks of total sleep deprivation, rats die, suggesting that sleep is critical for survival (Glotzbach & Heller, 1996). These types of observable changes that occur during the process of sleep deprivation stimulated the development of several theories which have attempted to explain the underlying mechanisms of sleep.

There are four major theories on the function of sleep. The thermoregulatory theories of sleep suggest that sleep serves as a time for the organism to produce or to dissipate heat in order to maintain optimal body temperature (Nakao, McGinty, Szymusiak, & Yamamoto, 1995). The homeostatic drive hypothesizes that the sleep/wake system is driven by a homeostatic process (Borbely, 1994), i.e., as the duration of wakefulness (and therefore the length of time since the last sleep period) increases, the propensity for sleep increases. While the organism sleeps, this drive gradually decreases. The model suggests, however, that independent of the duration of the wake period, sleep is more likely at some circadian times than others. A third theory, that of memory consolidation, suggests that during some periods of sleep (i.e., REM sleep), certain types of memories are consolidated for the purpose of long-term storage (Carskadon & Dement, 1994). The fourth theory is that of the restorative function of sleep. This theory suggests that the physiological and biochemical degradation that occurs during the day is restored during deep sleep (Rechtschaffen, 1994). Each theory has data supporting it as well as data contradicting the models. The only conclusion that can be drawn at this time is that, although the full function is not yet known, sleep is crucial for survival.

7.14.1.3 Changes in Sleep Associated with Aging

Various epidemiological studies and population-based surveys have consistently indicated that up to 50% of people over age 65 experience some disruption of sleep (Foley et al., 1995). However, studies have indicated that these sleep disturbances were associated with poor physical health and depression (Monjan & Foley, 1996). Studies of healthy older people have shown that sleep complaints were infrequent, notwithstanding the presence of increased night-time wakefulness and decreased time in deeper stages of sleep (Prinz, Vitiello, Raskind, & Thorpy, 1990).

7.14.1.3.1 Objective sleep studies

The general findings, based on objective PSG measurements, are that in contrast to younger people, older people spend more time in bed but sleep less, and are more easily aroused from sleep. The most prominent finding is the profound decrease in slow-wave sleep (stages 3 and 4).

As people age, sleep becomes increasingly fragmented, resulting from both a rise in the number of transient arousals and an increase in the total amount of time awake throughout the night. Increased night-time wakefulness may cause increased daytime fatigue, daytime napping, and increased likelihood of falling asleep during the day. The question of whether daytime naps are the cause or the consequence of disturbed night-time sleep needs to be further explored.

Daytime napping is common among elderly populations. While surveys vary in their estimations of the frequency of napping in elderly populations, there is consensus that the amount of napping increases with age. Prinz (1977) reported that 25% of elderly people studied took a daytime nap, while Wauquier and colleagues reported that 80% of healthy elderly subjects napped on average for 30 minutes a day (Wauquier, Van Sweden, Lagaay, Kemp, & Kamphuisen, 1992).

Results obtained from objective studies which focused on daytime sleepiness also showed an increase in daytime sleepiness as measured by the multiple sleep latency test (MSLT). The MSLT consists of 4–5 opportunities to go to sleep at 2 hour intervals and is considered to be the standard for evaluating daytime sleepiness. Older people fell asleep faster than younger adults on all nap opportunities (Dement, Seidel, & Carskadon, 1982).

7.14.1.3.2 Subjective reports

Subjective complaints of the elderly tend to be consistent with PSG findings. Older people report taking longer to fall asleep, sleeping less at night, waking more often during the night, waking up early in the morning, and napping more during the day (McGhie & Russel, 1962).

Using a sample of over 9000 community-dwelling individuals over the age of 65, Foley et al. (1995) found that over half of the sample reported at least one of five of the most common complaints: trouble falling asleep, waking up during the night, waking up too early, needing a nap, and not feeling rested. Sleep complaints have also been found to relate to measures of depression (Benca, Obermeyer, Thisted, & Gillin, 1992).

7.14.1.3.3 Gender differences

Gender-related differences have been noted in sleep of the healthy elderly. A meta-analysis of PSG data conducted by Benca et al. (1992) showed that older women may sleep better than older men. Women also appeared to maintain sleep better than men (Rediehs, Reis, & Creason, 1990). It is noteworthy that older women were more likely than older men to complain of sleeping difficulties and thus receive hypnotics more often (Ohayon & Caulet, 1995).

Hormonal changes associated with menopause impact sleep. Menopause has been associated with both subjective complaints of insomnia, and objective changes including increased sleep latency, decreased REM sleep, and decreased total sleep time. Treatment with supplemental estrogens may help to alleviate some of these sleep disturbances (Wooten, 1994).

7.14.2 DISORDERS OF SLEEP

Sleep disturbances in elderly people may be due to primary sleep disorders (e.g., sleep-related respiratory disturbances or periodic movements in sleep)—the frequency and severity of which tend to increase with age—or they may be secondary to a variety of medical, psychiatric, pharmacologic, and psychosocial factors. Chronic sleep disturbances may also exacerbate illness, cause discomfort and frustration, and may be related to depression. The daytime effects of chronic sleep disturbances include impaired cognition, confusion, slowed psychomotor functioning, and increased risks of injuries due to falls, thus impairing the quality of life of the patients and their families.

7.14.2.1 Sleep-disordered Breathing

Sleep-disordered breathing (SDB), also called sleep apnea, is a repetitive process of respiratory cessation during sleep, often resulting in arousals during which respiration resumes. There are three types of SDB: (i) obstructive sleep apnea involves a collapse of the airway which partially or completely blocks the air flow, (ii) central sleep apnea results from failure of the respiratory neurons to stimulate the motor neurons that mediate respiration, (iii) mixed sleep apnea involves both a central and an obstructive component. Respiratory events involve either a complete cessation of breathing (apnea), or a partial obstruction of the airway resulting in a reduction in breathing (hypopnea). The number of apneas per hour of sleep (apnea index; AI) plus the number of hypopneas per hour of sleep (hypopnea index; HI) determine the respiratory disturbance index (RDI). SDB is generally diagnosed when a patient has 10 or more respiratory events per hour of sleep (i.e. $\text{RDI} \geqslant 10$) or five or more apneas per hour of sleep (i.e., $\text{AI} \geqslant 5$) (Kryger, 1989).

Several factors, acting alone or in concert, have been hypothesized to contribute to the presence and severity of SDB. These factors include age; obesity; micrognathia; jaw or nasal deformities; thyroid, pituitary, or neurologic impairments; and alterations of respiratory reflexes during sleep (Shepard, 1989). It has been demonstrated that the effect of weight on SDB is at least as great as chronological age, even above age 70 (Bliwise et al., 1987).

7.14.2.1.1 Symptoms

Snoring and daytime sleepiness are the two most common symptoms of SDB. Snoring itself is a partial obstruction of the airway, and is associated with a number of health problems including hypertension and cardiovascular disease (Fairbanks, Fujita, Ikematsu, & Simmons, 1987). The snoring of sleep apnea patients, especially those with OSA, is often so loud that it disrupts the sleep of the bed partner. Therefore, the bed partner is often the first to suspect a breathing disturbance.

Another common symptom of SDB is excessive daytime sleepiness which likely results from the frequent nocturnal awakenings caused by the apnea episodes as well as from the hypoxia experienced at night. The resulting daytime sleepiness can have serious implications as patients with SDB may have difficulty sustaining wakefulness while working, conversing with others, or driving a car.

With repeated decreases in respiration, patients with SDB often experience hypoxemia which may lead to other health problems such as hypertension, cardiac arrhythmias, myocardial infarction, and stroke, which have long-term health implications (Shepard, 1992).

7.14.2.1.2 SDB and hypertension

Many studies have shown a relationship between snoring and hypertension with those reporting frequent snoring reporting a higher prevalence of diagnosed hypertension (Hoffstein, 1990). In fact, research has shown that about one-third of all hypertensives have SDB (Ancoli-Israel, Bliwise, & Mant, 1993). Apneas are accompanied by a transient elevation of blood pressure suggesting that SDB may contribute to sustained systemic hypertension.

7.14.2.1.3 SDB and cognitive impairment

In addition to excessive daytime sleepiness, sleep apnea is also associated with problems in concentration, attention, and memory during the day (Bliwise, 1993). Numerous researchers have demonstrated a relationship between oxygen desaturations in patients with sleep apnea and cognitive functioning (Ancoli-Israel & Coy, 1994). Bliwise (1991) demonstrated that SDB has an effect on cognitive function independent of age and other risk factors; however, some researchers have not successfully replicated these findings (Findley, Presty, Barth, & Suratt, 1991). In elderly sleep apnea patients with no other health problems, this relationship may not be as strong (Hayward et al. 1992). It is important to note that extremely healthy elderly are not representative of the older population in general who typically suffer from one or more health problems in addition to sleep apnea. Although controlled studies have not been done, it appears that treatment of SDB may result in some improvement in cognitive functioning (Derderian, Bridenbaugh, & Rajagopal, 1988).

7.14.2.1.4 Mortality

Many investigators have studied the relationship between SDB and mortality (He, Kryger, Zorick, Conway, & Roth, 1988; Lavie et al., 1995), but few have studied the relationship in elderly. Ancoli-Israel, Kripke et al. (1996) reported that, in a population of community-dwelling elderly men and women, those with severe SDB had a significantly shorter survival time on follow-up than those with mild–moderate SDB or those with minimal or no SDB. Further, over 50% of the patients with SDB were more likely to die from cardiovascular disease, while only 35% of the subjects with minimal SDB died of the same cause. However, SDB was not predictive of death, while cardiovascular disease and pulmonary disease were. The relationship between SDB, cardiovascular disease, pulmonary disease, and mortality needs further exploration.

7.14.2.1.5 Prevalence

SDB is a common sleep disorder in elderly men and women, with most studies confirming that sleep apnea is more common in the elderly than in younger populations (Ancoli-Israel & Coy, 1994). In middle-aged adults (ages 30–60), the prevalence of SDB is estimated at 4% in men and 2% in women (Young et al., 1993). Ancoli-Israel and her colleagues have shown that 81% of those over 65 years had 10 or more respiratory events per hour of sleep, and 24% had five or more apneas per hour of sleep (Ancoli-Israel et al., 1991a). They also demonstrated that 33% of comparably aged inpatient populations, and 42% of an elderly nursing home population had at least five apneas per hour of sleep (Ancoli-Israel, 1989).

Few studies have been done examining the relationship between SDB and race. In a sample of older adults, it was found that although the prevalence of SDB was not different between Caucasians and African–Americans, the older African–Americans had more severe SDB (Ancoli-Israel et al., 1995).

7.14.2.1.6 Treatments

Some suggest that an RDI $\geqslant 20$ should be aggressively treated since significant mortality risks have been demonstrated at this level. Until about the late 1980s, tracheostomy was the only available treatment for sleep apnea. Since that time, a variety of therapeutic options have become available. The appropriateness and efficacy of each treatment varies from one patient to another, and should be assessed on an individual basis.

(i) Positive airway pressure

Nasal continuous positive airway pressure (CPAP) and bilevel positive airway pressure (BiPAP) are highly effective treatments of OSA, and these are the most commonly used nonsurgical treatments. Nasal CPAP acts by providing continuous pressure to the upper airway which prevents airway collapse (Sullivan & Grunstein, 1989). The CPAP system involves a mask worn over the nose at night which is connected via a hose to the CPAP machine. The pressure is adjusted by the prescribing physician until no more or very few respiratory events are seen.

(ii) Surgery

Some patients may be unwilling or unable to comply with CPAP or other treatments to control OSA. For such patients a surgical intervention may be appropriate. One of these surgical treatments is uvulopalatopharyngoplasty (UPPP) which is intended to widen the airway passages by shortening the uvula, removing unnecessary tissue, and tightening the flabby muscles of the airway. This surgery may improve snoring, and has been shown to decrease the number of apneas during the night to the level achieved by CPAP in about 50% of patients (Powell, Guilleminault, & Riley, 1994). Since it is possible for snoring to decrease without improvement of SDB, follow-up evaluation of sleep is extremely important.

One of the advances in the surgical treatment of snoring is laser-assisted uvulopalatoplasty (LAUP) (Standards of Practice Committee of the American Sleep Disorders Association, 1994). A laser beam is used to shorten or remove the uvula and trim the soft palate which vibrates while breathing during sleep. This procedure may alleviate snoring without improving the severity of SDB, and sleep should be reassessed following treatment. LAUP is not recommended as a treatment for SDB.

(iii) Dental devices

Numerous dental devices have been developed to open the airway by holding the tongue or jaw forward in order to reduce snoring and OSA. These devices are quiet, cost-effective, simple to use, and may be particularly appropriate for patients who have difficulty using CPAP or are not appropriate for surgery. Although these devices successfully increase the size of the airway, the actual reduction in RDI varies greatly between patients (Lowe, 1994).

(iv) Medications

Tricyclic antidepressants such as protriptyline (Vivactil) and flouxetine (Prozac) are respiratory stimulants and may improve sleep apnea. The side effects of these medications and their interaction with other drugs a patient may be taking should be considered, especially in the elderly who may be taking multiple medications.

(v) Weight loss

Since obstructive sleep apnea results from a closing of the airway, excess fatty tissue in the neck area may be a contributing factor. Ancoli-Israel, Kripke et al. (1996b) reported that elderly men and women with sleep apnea had significantly higher body mass indices than those with minimal or no sleep apnea, supporting evidence that weight plays a significant role in the presence and severity of SDB.

(vi) Other treatment strategies

Several other factors which may exacerbate SDB should be addressed in the treatment of afflicted patients. Alcohol, which is sometimes used by men and women suffering from insomnia, has been shown to exacerbate sleep apnea (Taasan et al., 1981). Supine body position was found to increase SDB and snoring in some patients (Cartwright, Diaz, & Lloyd, 1991). Taking measures to avoid sleeping on the back may reduce the number of respiratory events. Some medications such as hypnotics, narcotics, and barbiturates may also exacerbate respiratory problems and reduce the patient's ability to wake up and breathe (Ancoli-Israel et al., 1993). These medications should be used with extreme care in the elderly who are more likely to suffer from SDB, and who also use sleeping pills more frequently than any other age group.

7.14.2.2 Periodic Limb Movements in Sleep

Periodic limb movements in sleep (PLMS), also called nocturnal myoclonus, is characterized by repetitive leg kicks during sleep, occurring every 20–40 seconds, lasting 0.5–5 seconds and causing brief arousals (Coleman, 1982). The precise mechanisms underlying PLMS are not well understood although numerous hypotheses have been proposed.

7.14.2.2.1 Symptoms

The primary symptom of PLMS is leg kicking during sleep. The kick may be as mild as a muscle twitch, or may be an involuntary jerk. An uncomfortable sensation builds up into a need to kick which relieves the sensations. Sometimes the patients themselves are not aware of the leg kicks, although bed partners usually are. Other symptoms of PLMS include insomnia, motor restlessness, excessive daytime sleepiness, and sometimes cold feet (Ancoli-Israel, Seifert, & Lemon, 1986; Coleman et al., 1982).

7.14.2.2.2 Clinical evaluation

Patients complaining of night-time awakenings or sleep-onset trouble may suffer from PLMS. Patients with PLMS may also report daytime sleepiness as a result of the night-time sleep fragmentation resulting from arousals associated with the leg jerks. Interviewing the bed partner along with the patient provides additional information.

The only way to make an accurate diagnosis of PLMS is to record leg movements as part of a complete PSG. The number of leg jerks recorded per hour of sleep is called the myoclonus index (MI). A diagnosis is made when there are five or more leg kicks per hour of sleep ($MI \geqslant 5$) each with an accompanying arousal. In patients with PLMS, periods of normal sleep alternate with disrupted sleep. The presence and severity of PLMS may vary from night to night (Mosko, Dickel, & Ashurst, 1988).

A related disturbance, restless legs syndrome (RLS), occurs more during the relaxed waking state. As summarized by Walters et al. (1995), patients experience "creeping" sensations in their legs. Moving the legs relieves this uncomfortable sensation, but the sensations return when the movement stops. Patients with RLS report difficulty falling asleep and symptoms may recur throughout the day or night. RLS is associated with caffeine intake, fatigue, exceptionally warm environment, prolonged exposure to cold, anemia, rheumatoid arthritis, and the last trimester of pregnancy. Almost all patients with RLS have PLMS, although the converse is not true.

7.14.2.2.3 Prevalence

The prevalence of PLMS in adults is estimated at 5–6% in the general population (Bixler et al., 1982). The prevalence increases with age, up to 45% in older populations, with no difference between men and women (Ancoli-Israel et al., 1991b).

Patients with PLMS sleep less than those with no leg jerks, and there is evidence for some sleep disturbance being associated with the leg kicks (Ancoli-Israel et al., 1991b). There is also evidence that patients with PLMS often suffer from SDB or narcolepsy (Coleman, Pollak, & Weitzman, 1979).

7.14.2.2.4 Treatments

PLMS is treated with medications and each treatment regimen has its advantages and disadvantages. Common treatments of PLMS include the use of sedative-hypnotics such as clonazepam or temazepam (Mitler, Browman, Menn, Gujavarty, & Timms, 1986), opiates such as Tylenol with codeine (Kavey, Walters, Hening, & Gidro-Frank, 1988), or levidopa/carbidopa (e.g., Sinemet or Sinemet CR) (Kaplan, Allen, Buchholz, & Walters, 1993). The fact that dopamine precursors and opiate drugs may alleviate PLMS in some patients suggests that one or both of these systems play a role in the pathophysiology of this disorder; however, both of these treatments are ineffective in some patients. Since many patients with PLMS also suffer from sleep apnea, a careful assessment of the sleep complaints should be conducted to determine whether SDB coexists with the PLMS since treatment of the PLMS with sedative-hypnotic medications may exacerbate sleep apnea. Whether the beneficial effects of long-term treatment with medications outweighs the risks associated with chronic use is yet unresolved.

7.14.2.3 Insomnia

The term "insomnia" (lack of sleep) refers to the complaint of an inability to initiate and/or to maintain sleep. Insomnia can be further divided into three subtypes: sleep-onset insomnia (difficulty in initiating sleep), sleep-maintenance insomnia (difficulty remaining asleep), and early morning insomnia (waking early and being unable to fall back to sleep). Many individuals experience a combination of the insomnia complaints. Younger adults are more likely to complain of sleep-onset insomnia whereas older adults are more likely to have sleep-maintenance or early morning insomnia, (Lichstein & Reidel, 1994). Insomnia is also defined by the duration of the disorder: transient insomnia refers to a very short-term insomnia that usually occurs in response to a sudden stress, short-term insomnia refers to a state that occurs over weeks and is often associated with acute medical or psychosocial factors, and chronic insomnia is a disturbance lasting months or longer.

Insomnia is the most common sleep complaint reported by elderly people. All epidemiological surveys that have evaluated the effect of age on sleep patterns have shown a higher prevalence of insomnia in older persons than in younger persons (Ford & Kamerow, 1989). Estimates of complaints of difficulties in falling asleep, of frequent or prolonged wakening, and of generally light and disturbed sleep range up to 40% for those 60 years of age and older (Miles & Dement, 1980), and estimates of severe insomnia complaints are more than 20% (Mellinger, Balter, & Uhlenhuth, 1985).

In addition to advancing age, female gender and concomitant health problems have been identified as important risk factors for insomnia (Klink, Quan, Kaltenborn, & Lebowitz, 1992). Women also tend to use more sedative medications to treat their insomnia than men (Mellinger et al., 1985). Individuals with chronic insomnia experience more psychological distress, are more preoccupied with somatic complaints, and utilize health care resources more often than good sleepers (Morin, Culbert, & Schwartz, 1994).

7.14.2.3.1 Clinical evaluation

Clinical evaluation of insomnia should encompass the various facets of the complaint. Insomnia should never be evaluated strictly as a night-time problem; rather the daytime effects on cognition and behaviors and the daytime behaviors that may contribute to the insomnia should also be examined (Morin, 1993).

(i) Clinical interview

As with any sleep complaint, a comprehensive clinical assessment includes a detailed medical and sleep history. The sleep interview usually examines the nature and the severity of the complaint, the daytime consequences of the insomnia, the environmental factors, the medication use, the dietary and smoking habits, the medical history, and the possibility for other sleep disorders.

Following the interview, Morin (1993) suggested that a functional behavioral analysis be conducted. The purpose of the functional analysis is to examine the antecedents, consequences, and the other controlling variables of the insomnia, and to identify the perpetuating factors. This understanding is essential for tailoring an intervention that is adapted to the specific difficulties of the individual. The assessment needs to be carried out with support and empathy and should focus on the patient's prebedtime routine, sleep incompatible activities, secondary gains, and typical responses to sleeplessness. The clinician may use and integrate information obtained by subjective daily sleep reports and by objective measurements. Compliance issues need to be addressed directly in individual discussions with the patient and, when possible, should involve the bed partners as they are most helpful in providing additional information.

(ii) Sleep diaries

Self-report measures are frequently used for assessment of the disturbance and for the evaluations of the treatment outcomes. Sleep diaries, also called sleep logs, are usually employed daily to collect descriptive data on sleep. A typical sleep diary is shown in Table 1. From the reported data, the total time in bed, total sleep time, and sleep efficiency can be calculated. Sleep diaries are useful for tracking treatment progress. Another advantage is that they demand active participation of the client. Despite the subjective nature of the sleep diaries, they tend to correlate significantly with PSG measures (Carskadon et al., 1976). Since insomnia is a subjective complaint, the treatment success, as defined by the patient, is more dependent on the patient's favorable perception of sleep than on objective changes.

(iii) Objective measurements

The usefulness of PSG for diagnosis of insomnia has been questioned. First, insomnia may be setting specific, thus laboratory PSG might not adequately represent the patient's insomnia as experienced at home (Hauri, 1983). Second, a single PSG might not detect chronic insomnia, given the nature of the disorder, since insomnia does not necessarily present every night, and repeated recordings are very costly. Conversely, it has been argued that PSG may be valuable for older patients since the presenting symptoms of insomnia may be secondary to sleep apnea, PLMS, or other sleep disorders that may have a similar clinical presentation to insomnia (Edinger et al., 1989). Lichstein and Reidel (1994) also conclude that PSG is indicated not only when there is a suspicion that other sleep disorders are present but also when insomnia treatments fail.

Wrist activity is an alternative objective measurement of the sleep–wake schedule. However, the results of actigraphy with insomnia patients may exhibit substantial disagreement with PSG results (Hauri & Wisbey, 1992). Nevertheless, actigraphy can be used effectively as part of psychological treatment to measure changes in sleep within subjects and can provide the patient with some reflection on the changes occurring in his or her sleep patterns.

Table 1 Sample sleep diary.

1. What time did you get into bed last night?
2. What time did you fall asleep last night?
3. How long do you think it took you to fall asleep?
4. How many times did you wake up during the night?
5. How long did it take you to fall back to sleep?
6. What time did you wake up in the morning?
7. What time did you get out of bed?
8. How did you feel when you woke up in the morning?
9. How much caffeine did you have yesterday?
10. How much alcohol did you have yesterday?
11. Did you nap? If yes, how many times and for how long?

Note. This questionnaire should be completed in the morning.

7.14.2.3.2 Etiology of insomnia

It is widely accepted that insomnia is a multifactorial symptom resulting from medical, behavioral, psychosocial, and environmental factors. *DSM-IV* (American Psychiatric Association, 1994) provides three categories for insomnia: primary insomnia, insomnia related to another mental disorder, and sleep disorders due to a general medical condition, insomnia type. The diagnoses are made on the basis of interview data. Chronic insomnia can have debilitating daytime consequences on mood, fatigue, performance, and motivation, as well as an adverse affect on an elderly person's quality of life and therefore on the public health system.

(i) Medical factors

The discomfort of medical conditions may disturb sleep, and some disorders, such as neurological and respiratory diseases, will directly alter sleep structure. Medical problems that are commonly associated with insomnia include neurological disorders, chronic obstructive pulmonary disease (COPD), cerebrovascular disease, headache, asthma, and any conditions that produce persistent pain (Wooten, 1994).

(ii) Psychiatric conditions

Sleep is also influenced by psychiatric conditions which can alter sleep patterns and architecture. It has been estimated that approximately 90% of depressed individuals exhibit some abnormality in sleep architecture (Reynolds, 1989). Depression has been related to early morning insomnia, although the most common finding was decreased REM latency (Benca et al. 1992). When sleep variables of elderly depressed individuals were compared with those of younger counterparts, the elderly show decreased sleep efficiency, decreased REM latency, decreased total sleep time, increased intermittent wakefulness, and decreased slow-wave sleep (Gillin et al., 1981).

(iii) Alcohol, caffeine, and nicotine

Alcohol, caffeine, and nicotine are all substances that are widely used (and abused), yet each can have a negative effect on sleep. Alcohol is the most commonly used nonprescription sleep medication. While it may initially decrease sleep latency, later in the night it causes sleep fragmentation, REM deprivation, and REM rebound. Alcohol use decreases with age, however, the sedating effects of alcohol may be more pronounced in elderly due to their decreased metabolic rate (Dufour, Archer, & Gordis, 1992). Furthermore, the combination of alcohol and hypnotics can also play a role in exacerbating existing insomnia.

Stimulant substances such as caffeine and food or drugs that contain caffeine (such as coffee, tea, certain sodas, chocolate, analgesics, bronchodilators, steroids, some antihypertensives, decongestants, and appetite suppressants) may produce insomnia as a side effect (Espie, 1991). Caffeine has been associated with an increase in sleep latency, decrease in total sleep time, and increase in spontaneous awakenings. Discontinuation of these substances may be associated with withdrawal symptoms of sleepiness, irritability, and depression (Diagnostic Classification Steering Committee, 1990). Nicotine may also produce insomnia and is related to lighter and more fragmented sleep.

(iv) Psychosocial factors

Life events (such as bereavement) and lifestyle behavior can affect sleep and instigate insomnia. Retirement and the lack of scheduled activities may cause more daytime napping (Kronholm & Hyyppa, 1985). Evidence shows that adhering to an active and structured schedule may be an important factor in maintaining satisfactory sleep (Hoch, 1991).

Physical activity has been related to improvement of both sleep onset and sleep maintenance. PSG reports of older adults who participate in regular exercise revealed more sleep consolidation (e.g., less wake time, more continuous and deeper sleep) than sedentary subjects (Edinger et al., 1993).

A diverse group of behaviors, sleep habits, and environmental variables tend to promote or impair sleep. Poor sleep habits, irregular sleep schedule, noise, room temperature, and changes in the sleep environment are related to sleep disruptions. Sensitivity to environmental factors increases with age and is also subject to individual difference.

(v) Cognitive factors

Patients with insomnia are more prone to cognitive arousals and anxiety than those with no insomnia (Lichstein & Fanning, 1990). Cognitive factors were identified by patients with insomnia as a more frequent source of sleep disruption compared with somatic complaints (Espie, Brooks, & Lindsay, 1988).

7.14.2.3.3 Pharmacological treatments

Older patients frequently take multiple prescription medications, often prescribed by more than one physician, as well as over-the-counter medications. Some of these medications may cause insomnia, contribute to daytime sleepiness, or exacerbate sleep apnea. Alcohol, beta blockers, bronchodilators, corticosteroids, CNS stimulants, decongestants, gastrointestinal drugs, cardiovascular drugs, and neurological drugs can all cause insomnia. Analgesics and benzodiazepines are known to cause daytime sleepiness. Antidepressants also affect sleep. Amitriptyline, sinequan, and trazodone are known to cause sleepiness, while buproprion, fluoxetine, paroxetine, sertraline, and venlafaxine are known to cause insomnia. Many antidepressants, such as desipramine, imipramine, and nortriptyline, can cause both insomnia and/or daytime sleepiness (Ancoli-Israel, 1996).

Pharmacotherapy is still the most commonly used therapy for insomnia in the elderly (Prinz et al., 1990), although it is complicated by age-related changes in pharmacokinetics and pharmacodynamics (Hicks et al., 1981). Sleep medication usage is disproportionally higher among older people, particularly among women.

There may also be adverse daytime effects including performance deficits (e.g., poor motor coordination, visual-motor problems), increased daytime sleepiness, and accidental injury (Roth, Roehrs, & Zorick, 1988). In addition, some hypnotics may alter sleep architecture by reducing slow-wave sleep and REM sleep. With continued use, tolerance can develop after only a few weeks, resulting in the return of the insomnia symptoms or requiring an increase in dosage in order to achieve the previous efficacy. Sleeping pills, therefore, when used inappropriately, can play a role in preserving and exacerbating insomnia rather than improving it. Withdrawal from hypnotics (especially abrupt cessation from long-acting drugs) may produce rebound insomnia, which is characterized by marked worsening of sleep and heightened anxiety. In addition, elderly are more likely to have disorders that can be aggravated by hypnotics such as respiratory, hepatic, renal, or cardiac disorders. As mentioned, hypnotics prescribed to patients who suffer from sleep apnea could exacerbate their oxygen desaturation and create serious medical risks (Gillin & Ancoli-Israel, 1992).

Elderly patients who are prescribed hypnotics should be given the lowest possible dose of a shorter-acting hypnotic for a short time period. The patient should be closely followed to decide if a renewal of the prescription is warranted.

All hypnotics impair performance while active; therefore, those with shorter half-lives and shorter duration of action have less impairment. Anterograde amnesia is a potential side effect of all hypnotics; however, the dose and duration will determine how severe the memory loss is. Since sleep itself produces amnesia, the interaction of the hypnotic with sleep onset will determine how much amnesia is experienced.

Rebound insomnia can also be caused by hypnotics, particularly those with short or intermediate half-lives. When the hypnotic is suddenly withdrawn, insomnia returns for several nights and the patient is tempted to resume taking the drug. For all these reasons, hypnotics are not recommended to be used on a nightly basis and should only be used for short periods of time.

All hypnotics, if given in adequate doses, will improve insomnia. The more commonly prescribed hypnotics include estazolam, flurazepam, quazepam, temazepam, triazolam, and zolpidem. The task is to determine the lowest adequate dose that promotes sleep with the fewest side effects. In addition, drug therapy should always be in conjunction with a nonpharmacological plan directed toward the underlying disorder.

7.14.2.3.4 Psychological treatments

The recognition of the role of psychological factors in creating, precipitating, or enhancing insomnia combined with acknowledgment of the limitations and the risks associated with long-term use of hypnotics prompted the development of nonpharmacological approaches for the management of insomnia. These alternative approaches include educational, behavioral, and cognitive interventions aimed at changing the maladaptive behaviors and attitudes toward sleep and insomnia.

(i) Sleep hygiene

Sleep hygiene is an educational approach for treatment of insomnia that was developed by Hauri and Linde, (1990). The basic premise of sleep hygiene is that poor behavioral habits (i.e., inadequate sleep hygiene) interfere with the maintenance of good quality of sleep and daytime alertness. Several habits may lead to disturbed sleep, for example, taking daytime naps; spending extended time in bed awake; keeping irregular sleep/wake schedules; excessive use of caffeine, nicotine, or other stimulants; exercise at inappropriate times; poor sleep environment; and engaging in exciting or emotionally upsetting activities close to bedtime. Providing basic knowledge about the structure of sleep and the factors that affect sleep constitute the core component of sleep hygiene. Table 2 lists the major rules of good sleep hygiene for the elderly.

Sleep hygiene has been incorporated into most psychological interventions for insomnia. Clinically, these instructions provide a good start for treatment. Hauri (1991) cautioned against giving a complete written list of sleep hygiene rules as that might overwhelm the patient. Instead, he recommended targeting a few rules which are pertinent to the patients' specific problem.

(ii) Stimulus-control therapy

Stimulus-control therapy, developed by Bootzin and Nicassio (1978), is based on the premise that sleep disturbances are behaviorally conditioned and therefore need to be reconditioned. The purpose of the intervention is to establish a consistent sleep/wake pattern, recondition the bed and bedroom as cues for sleep, and extinguish the maladaptive associations. Table 3 lists instructions for stimulus-control therapy.

The focus of the instructions is primarily to shorten sleep onset. For sleep-maintenance difficulties, the instructions are to be followed after awakening when the patient has difficulty falling back to sleep. Compliance is enhanced when the instructions are discussed individually and a rationale is provided for each rule. Two modifications of the rules have been suggested for older patients. For rules 2 and 3, those over the age of 60 are instructed to get out of bed after about 20 minutes, instead of 15 minutes. This recognizes that older individuals usually take longer to fall asleep than do younger individuals. The goals of rule 8 are to keep patients with insomnia from disrupting their sleep patterns by irregular napping and to prevent them from losing the advantage of the sleep loss from the previous night which increases the likelihood of faster sleep onset the following night. However, a short daily nap at the same time would be permissible. For those elderly patients with insomnia who feel that they need to nap, a short daily nap (30 minutes or less) or relaxation period as a nap substitute may be recommended (Bootzin & Engle-Friedman, 1987).

Table 2 Rules of sleep hygiene for the elderly.

1. Limit naps to less than 30 minutes per day
2. Take a walk both to exercise and to increase light exposure
3. Check the effect of medications on sleep
4. Avoid caffeine, alcohol, and tobacco
5. Limit liquids in the evenings
6. Keep regular hours

Table 3 Instructions for stimulus-control therapy.

1. Patient goes to bed only when sleepy
2. If not asleep within 15 minutes, patient gets out of bed, and does not return to bed until sleepy
3. When patient returns to bed, if not asleep within 15 minutes, gets out of bed again
4. Pattern is repeated until patient can fall asleep within a few minutes
5. Must avoid looking at the clock
6. Must get up at the same time each morning (even if only slept 2 hours)
7. Bed is used only for sleeping (not for watching television, reading exciting books, etc.)
8. All naps during the day must be avoided

Stimulus-control techniques have been shown to be effective with a variety of insomnia patients. Comparisons of stimulus-control interventions with placebo interventions have demonstrated superior sleep-latency curtailment for stimulus-control (Lacks, Bertelson, Gans, & Kunkel, 1996). Several studies reported improvement in sleep-maintenance insomnia in older patients (Morin & Azrin, 1988).

(iii) Sleep-restriction therapy

The approach to restrict time in bed and sleep duration in order to increase sleep efficiency was developed as a technique for the treatment of insomnia by Spielman, Saskin, and Thorpy (1987). It is based on the observation that many patients with insomnia spend excessive time in bed in futile attempts to achieve more sleep. Over time, the frustration and worry produced by excess awake time in bed serve to perpetuate the insomnia. Sleep-restriction therapy consists of curtailing the time spent in bed to the actual amount of time slept. The effects of restricting time in bed are twofold: first, this excess worry time is eliminated, and second, restriction may create an initial state of sleep deprivation, which itself produces more consolidated and satisfying sleep. Prescribed sleep-period durations from bedtime to arising time are based on self-reported sleep durations that are obtained from sleep diaries of the previous week. At first, patients may experience some sleep deprivation since they usually underestimate their amount of sleep, but this sleep deprivation may eventually work to reduce the sleep latency and to consolidate sleep. For example, if a patient reported sleeping 5 hours per night (out of 8 hours spent in bed) then the initial prescribed "sleep window" would be 5 hours per night. After sleep efficiency increases and the patient is "using effectively" the 5 hours for sleep, the allowed time in bed is increased by about 20 minutes each week, and this gradual increase is contingent on improved sleep efficiency. Implementation of the technique requires flexibility and adjustment to the individual patient and obviously a high level of cooperation and compliance. Table 4 lists the rules for sleep-restriction therapy.

Further modifications were added by Glovinsky and Spielman (1991) in order to increase compliance and to adjust the instructions to specific problems. For example, they included some flexibility with regard to setting the time in bed for patients with subjective perception of little or no sleep. For elderly patients, the sleep-efficiency criteria is lower by 5%. Research on sleep-restriction therapy, focused on older adults, showed good results (Friedman, Bliwise, Yesavage, & Salom, 1991). It has also been reported that sleep restriction combined with sleep hygiene and stimulus-control instructions can be effective for the treatment of sleep-maintenance insomnia in older patients (Gillin & Byerley, 1990). Lichstein and Reidel (1994) concluded that sleep-restriction therapy is probably the treatment of choice for geriatric insomnia.

A study examined the effects of sleep compression (restricted time in bed) combined with sleep education for older patients with insomnia (Riedel, Lichstein, & Dwyer, 1995). The study also looked at the differences between self-help techniques (using a guiding video) and therapist guidance. They found that the self-help insomniac group exhibited improvement on multiple sleep variables, but the addition of therapist guidance appeared to enhance treatment outcome for sleep latency, wake time after sleep onset, and sleep satisfaction. Interestingly, the sleep knowledge was equivalent for older patients with insomnia and no insomnia, suggesting that knowledge of age-related sleep changes may have a weaker influence than actual sleep-pattern differences on whether individuals consider themselves as having insomnia or not.

Table 4 Instructions for sleep-restriction therapy.

1. Patient is only allowed to stay in bed for the amount of time they think they sleep each night, plus 15 minutes. For example, if patient reports sleeping only 5.75 hours, they are allowed to stay in bed for 6 hours
2. Patient must get up at the same time each day. If normal waking time is 6.30 a.m., patient is allowed to go to bed at 12.30 a.m
3. Napping is not allowed
4. When sleep efficiency has reached 85%, the patient can go to bed 15 minutes earlier
5. This procedure is repeated until the patient can sleep for the desired amount of time

(iv) Cognitive interventions

Cognitive interventions for the treatment of insomnia are applications of the general principles of the cognitive model to psychotherapy as developed by Beck, Rush, Shaw, and Emery (1979), and others and further adapted specifically to the treatment of insomnia. The general principle of cognitive interventions is to focus on clarification and change of the maladaptive expectations, beliefs, and attitudes that the individual holds toward his or her problems; in this case, toward his or her sleep and sleeplessness. The cognitive change is ensued by using cognitive restructuring, reframing of the problem, challenging misconceptions, and enhancing the sense of self-control.

Morin, Stone, Trinkle, Mercer, and Remsberg (1993) described cognitive behavioral therapy for insomnia as an integration of behavioral, educational, and cognitive components which target the different facets of insomnia. The cognitive component is based on the principle of cognitive restructuring, applied by Morin (1993) to the special needs of the treatment of insomnia. The purpose of the cognitive intervention is to focus on dysfunctional beliefs and attitudes toward sleep and insomnia and to alter maladaptive attitudes that lead to or exacerbate sleep difficulties. Common misconceptions are held about sleep requirements, the causes for sleep difficulties, the consequences of insomnia and about the ability to control, predict, and promote sleep. These attitudes and misconceptions about sleep and sleeplessness may perpetuate a vicious cycle of insomnia, and may lead to or exacerbate existing sleeping difficulties. The fear of sleeplessness may then cause increased emotional arousal and anxiety that may create more problems in initiating and maintaining sleep.

Cognitive-behavior therapy sessions, as described by Morin, Kowatch, Barry, and Walton (1993), follow a three-step process: identification of the patient's specific dysfunctional beliefs, confrontation and challenging of those thoughts, and implementation of methods for revising and altering these thoughts with more rational and adjusted beliefs. A study conducted on elderly patients with insomnia showed that they exhibited more unrealistic sleep expectations, misattributed the causes of their sleep disturbances, and engaged in catastrophic thinking about the consequences of their insomnia, compared with normal-sleeper controls (Morin, Stone et al., 1993). Thus with elderly patients, the cognitive restructuring techniques should address the issue of normal changes in sleep with age versus pathological changes.

(v) Relaxation and biofeedback

A basic assumption behind various techniques of relaxation, meditation, and biofeedback is that the patients with insomnia are overaroused and anxious, and this arousal interferes with their ability to fall asleep and to resume sleep after being awakened in the middle of the night (Hauri, 1981). The purpose of the relaxation procedure is to teach the patients to relax so that they can fall asleep faster. It has been argued that the benefit may extend beyond the sleep problem and will facilitate a more general improvement in coping skills.

(vi) Summary

Two meta-analyses, conducted to examine the efficacy of psychological treatments for insomnia, summarized 59 and 66 treatment outcome studies, respectively (Morin et al., 1994; Murtagh & Greenwood, 1995). Both analyses concluded that nonpharmacological interventions produce a reliable and durable change in both sleep patterns and the subjective experience of sleep, compared with control groups. The improvement obtained by the psychological interventions was also maintained at short- and long-term follow-ups. Whereas Morin et al. (1994) concluded that the most effective single-therapy procedure was stimulus control and sleep restriction, Murtagh and Greenwood (1995) concluded that psychological treatments did not differ greatly in efficacy. Greater therapeutic gains were available for participants who were clinically referred and were not regular users of sedative hypnotics. It is possible that the drug users, because of their psychological reliance on hypnotics, had low self-efficacy for falling asleep naturally, which may have reduced expectations for improvement and consequently undermined treatment (Murtagh & Greenwood, 1995).

Although only a small number of studies focused on the elderly, the results show that insomnia in old age responds to psychological techniques. However, in a review, only five of the intervention studies with older adults included a no-treatment control group and the role of placebo has not been tested in geriatric insomnia (Lichstein & Reidel, 1994).

Studies have compared the efficacy of psychological vs. pharmacological therapies (McClusky, Milby, Switzer, Williams, & Wooten, 1991). Frequently, psychological treatments are used in combination with hypnotic therapy. In fact, there is a growing advocacy for a package approach combining both pharmacologic and behavioral interventions (Lichstein & Reidel, 1994).

7.14.3 CIRCADIAN RHYTHM CHANGES WITH AGE

The sleep/wake cycle is controlled by the body's circadian rhythms. The circadian rhythms, in turn, are controlled by the suprachiasmatic nucleus (SCN) located in the anterior hypothalamus. Studies have shown that surgical removal of the SCN results in animals taking multiple brief naps during the day and night rather than having one long sleep period at night.

Circadian rhythms are synchronized by external cues, or zeitgebers, such as light. Light enters the retina of the eye through the retinohypothalamic tract to the SCN. As the light intensity varies, our sleep/wake cycle varies. The circadian core body temperature rhythm and the circadian melatonin rhythm also interact with the timing and quality of sleep (Myers & Badia, 1995).

Much of the increased napping, daily fatigue, and daytime sleepiness in the elderly can be partially attributed to the breakdown in the body's circadian rhythm, manifested in the advancement of the rhythm and the decline in the rhythm amplitude (Bliwise, 1993; Vitiello, 1996). Czeisler et al. (1989) confirmed a phase advance of about 90 minutes and reduced amplitude of about 30% in the body temperature rhythms of older subjects compared with younger adults.

Circadian rhythm disturbances occur when the body's rhythm no longer matches that of the environment. One such common problem in the elderly is advanced sleep phase syndrome (ASPS). People with ASPS get sleepy earlier in the evening, around 8.00 p.m. If they were to go to bed at that point, they would most likely sleep for about 8 hours, waking up at 4.00 a.m. Most older people, though, do not go to bed early even if they are tired. Instead, they stay up until 10–11.00 p.m. Their bodies however, are still waking up at 4.00 a.m. Figure 1 illustrates the ASPS. In this way, older people are essentially depriving themselves of sleep by not allowing enough time in bed during the body's correct circadian phase to get sufficient sleep.

7.14.3.1 Bright-light Treatment

Studies of light exposure have shown that elderly living in the community are only exposed to 60 minutes of bright light a day, i.e., of 2000 lux or more. Alzheimer's disease patients living at home are exposed to 30 minutes of bright light a day (Campbell, Kripke, Gillin, & Hrubovcak, 1988). Studies on nursing home patients have shown that no

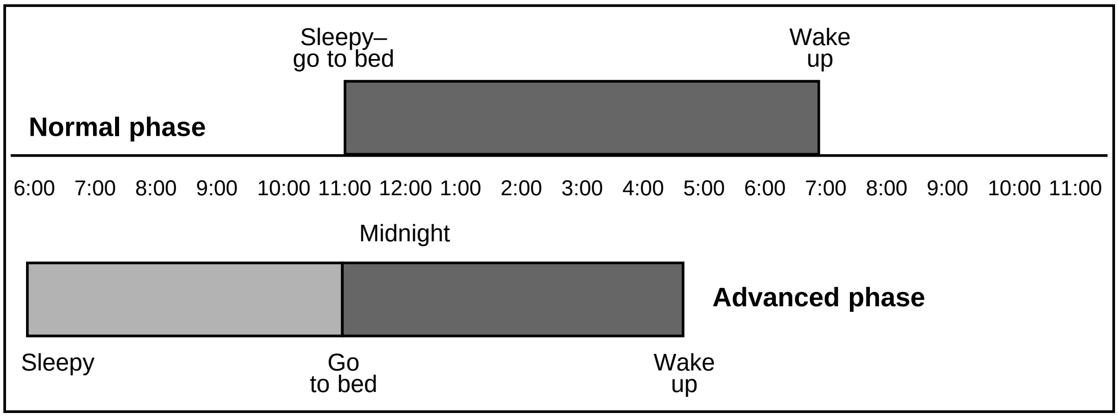

Figure 1 Advanced sleep phase syndrome. (Reprinted with permission of Mosby from Ancoli-Israel (1996).)

patient is ever exposed to over 2000 lux a day and on average, only 11 minutes are spent at levels over 1000 lux per day (Ancoli-Israel, Jones et al., 1991). Since light is such a strong zeitgeber, appropriately timed exposure to bright light can influence phase and amplitude of the circadian timing system (Czeisler et al., 1989). Specifically, bright-light exposure in the evening delays the circadian rhythm, i.e., pushes the rhythm forward so that the individual is no longer sleepy early in the evening. It has been shown that 2 hours of evening bright-light exposure (of at least 2500 lux) is beneficial for sleep maintenance insomnia in healthy older individuals (Campbell et al., 1995).

7.14.3.2 Melatonin

Melatonin, which is produced by the pineal gland at night, can reset sleep onset through its synchronizing effect on the internal biological clock (Dawson & Encel, 1993). With increasing age, the pineal gland secretes less melatonin (Reiter, 1995). The efficacy of administration of melatonin as a replacement therapy is a focus of growing interest. Haimov et al. (1995) reported a correlation between poor sleep and lower peak levels of melatonin in elderly patients with insomnia, suggesting that melatonin insufficiency may be a key variable in sleep disturbance of these older people. Garfinkel, Laudon, Nuf, and Zisapel (1995) conducted a follow-up study on 12 elderly patients with insomnia using a controlled-release formulation that provided effective concentrations of melatonin throughout the night. They found an improvement in sleep efficiency, and a reduction of the time awake after sleep onset during treatment, although total sleep time was not significantly different. The authors concluded that melatonin deficiency is a common cause of insomnia in the elderly and that replacement with controlled-release melatonin may effectively improve sleep quality in this population; however, patients should be screened for "melatonin deficiency" before administration of exogenous melatonin. These results were confirmed by Haimov et al. (1995).

Since the research on the efficacy of melatonin for improving sleep in general and for its specific efficacy on sleep consolidation in the elderly is in its first stages, more data are needed. Moreover, since melatonin is not regulated by the FDA and sufficient data are unavailable, caution should be taken with regard to the purity and dosage of the melatonin currently being sold over-the-counter and the time at which the melatonin is taken.

7.14.4 DEMENTIA AND SLEEP

Dementing illnesses include several nonreversible conditions (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, multi-infarct dementia), and there is considerable evidence that dementia impacts sleep differently than the normal aging process (Bliwise, 1993). The sleep/wake rhythm in Alzheimer's disease (AD) patients has been found to be extremely disturbed (Ancoli-Israel, Klauber et al., 1996; Bliwise, Hughes, McMahon, & Kutner, 1995). As many as 70% of demented patients suffer from AD, and due to this high proportion of AD patients, most sleep studies include a majority of patients suffering from AD and often fail to distinguish between the specific type of dementia (Bliwise, 1993). Prinz et al. (1990) describe the sleep of AD patients as showing increased duration and frequency of night-time awakenings, decreased slow-wave and REM sleep, and daytime napping.

7.14.4.1 Laboratory Studies

Compared with nondemented control subjects, the waking EEG of dementia patients typically shows abundant diffuse slow activity, decreased alpha frequency, and other specific abnormalities associated with pathophysiological changes in the brain (Prinz & Vitiello, 1989). This slowing of EEG activity in dementia may cause distinguishing wake from sleep and differentiating between sleep stages to be very difficult. Temporal asymmetry frequently appears in both the waking and REM sleep EEG (Petit, Montplaisir, Lorrain, & Gauthier, 1992). Additionally, abnormalities in REM sleep measures (including REM percent and REM latency) are often found in AD patients, and may present only in the more severe stages of the disease (Reynolds, Kupfer et al., 1988).

7.14.4.2 Dementia and Depression

Demented patients presenting with depression may suffer from even more serious sleep disturbances than those with dementia in the absence of depression. This has been demonstrated with a PSG study comparing AD patients with age-matched depressives. This study showed that the depressed patients had even more disturbed sleep than the demented patients (Reynolds, Kupfer et al., 1988). In fact, elderly men and women with severe depression may appear to have symptoms of dementia (Reynolds, Hoch et al., 1988).

7.14.4.3 Nursing Home Patients

Sleep fragmentation has been shown to be extreme in institutionalized elderly (Ancoli-Israel, Parker, Sinaee, Fell, & Kripke, 1989). Disturbed sleep with frequent awakenings, night wandering, wakefulness, and confusion has been identified as one of the most common symptoms leading to the institutionalization of the elderly (Pollak & Perlick, 1991). Table 5 lists suggestions for dealing with sleep disturbances in this population.

Several attempts to objectively measure the sleep of nursing home patients with traditional EEG measures have had multiple problems. One drawback of EEG measurement is that the EEG wires themselves may exacerbate confusion and agitation in demented elderly (Cohen-Mansfield, Werner, & Freedman, 1995). A second problem is that the waking EEG in these patients shows diffuse or slow activity, often making it difficult to distinguish wake from sleep (Bliwise, 1993). In addition, both objective and observational sleep studies have demonstrated that the night-time sleep and the sleep/wake cycles of nursing home patients are extremely disrupted (Ancoli-Israel et al. 1989; Bliwise et al. 1995). Bliwise et al. (1995), using observational data collected by nursing home staff members, demonstrated that the level of night-time sleep fragmentation in AD patients was related to the severity of their dementia.

Ancoli-Israel and her colleagues (Ancoli-Israel, Klauber et al., 1996; Jacobs, Ancoli-Israel, Parker, & Kripke, 1989) conducted a series of studies using wrist actigraphy with nursing home patients. They computed the minutes of sleep and wake hour-by-hour, and found that, on average, over 24 hours, there was not a single hour that consisted of complete sleep or of complete wake; rather, a period of some sleep and some awakening occurred during each hour. Severely demented patients slept more in every hour of the 24-hour period than mild–moderately demented patients (Pat-Horenczyk & Ancoli-Israel, 1996). Times of most alertness were confirmed to be the meal times (Jacobs et al., 1989).

Previous results have shown great individual variation in sleep patterns and sleep pathology of cognitively impaired nursing home populations (Regestein & Morris, 1987). An exploration of the relationship between the severity of dementia and sleep found that overall, severely demented patients slept more both during the

Table 5 Dealing with sleep disturbances in nursing home patients.

1. Limit time in bed
2. Keep regular schedule of wake/sleep, meals, etc.
3. Avoid caffeine; make sure patients are not buying caffeinated snacks in the nursing home gift shop or getting gifts of chocolate from family and friends
4. Encourage time cues: serve meals at a table at the same time each day and avoid eating meals in bed; place large clocks and calendars around the home
5. Avoid frequent naps; limit to one daily nap
6. Keep night-time noise to a minimum
7. Keep resident rooms as dark as possible at night
8. Keep the nursing home as brightly lit as possible during the day
9. Match room-mates on basis of night-time as well as daytime behaviors and routines
10. Encourage exercise (appropriate to the patient's level of function)
11. Keep bedtime routines similar to prior routines at home whenever possible

day and during the night when compared with mild–moderately demented patients (Ancoli-Israel, Klauber et al., 1996). Systematic observations of sleep and wakefulness in a nursing home population also yielded similar results (Bliwise, Bevier, Bliwise, Edgar, & Dement, 1990).

A 1996 study explored the differences in profiles of alertness and sleepiness between severely and mild–moderately demented patients (Pat-Horenczyk, Jones, Martin, & Ancoli-Israel, 1996). Patients with mild–moderate dementia showed more alertness during the night, while patients with severe dementia showed more sleepiness during the day.

A multidimensional scaling analysis performed on the percentage sleep for each hour in severely demented patients and in mild–moderately demented patients suggested that in the mild–moderate dementia group, there was greater similarity in the level of sleep/wake between the daytime hours, whereas in the night-time hours the amount of sleep/wake varied greatly from hour to hour. In the severe dementia group, sleep/wake was equally disturbed during both the day and night. The authors concluded that sleep/wake patterns in severe dementia and in mild–moderate dementia were mainly different in the daytime hours. Mild–moderately demented patients had disturbed night-time sleep (sleep fragmentation), whereas the severely demented patients had both day and night disturbances (wake fragmentation during the day and sleep fragmentation during the night).

7.14.5 SUMMARY

In summary, sleep does change with age. Subjective complaints and objective measurements of sleep indicate an increase of sleep and wakefulness disturbances with old age. In general, elderly people sleep less deeply, have more fragmented sleep and tend to wake up earlier than younger adults. A central and controversial question is whether these changes constitute a decreased need for sleep with advanced age or whether it is the ability to sleep that diminishes with age.

Some of the changes in sleep are secondary to circadian rhythm changes. The circadian rhythm naturally advances a few hours with age, so older people may feel sleepy earlier in the evening and wake up earlier in the morning. There is also a high prevalence of sleep disorders, such as SDB and periodic limb movements in sleep, which contribute to sleep disruption. These disruptions in sleep lead to daytime sleepiness which has been objectively documented.

As mentioned, older people have a shorter sleep-onset latency on MSLT than younger adults, indicating more daytime drowsiness. Daytime drowsiness is caused by sleeplessness at night. This increase in daytime sleepiness seen in the elderly has been interpreted as support of the notion that it is the ability to sleep that changes with age and not the need for sleep. If the actual need for sleep decreased, there would be no daytime sleepiness. The fact that healthy elderly have normal sleep and that sleep complaints are associated with medical illness also supports the notion that it is the ability to sleep that decreases with age, rather than the need to sleep.

When evaluating an older person's sleep complaints, it is imperative to consider primary causes of sleep disturbances including medications, illness, depression, circadian rhythm changes, and sleep disorders. It is also important to remember that sleep of the elderly can be improved. Physical activity, natural light and sticking to a regular sleep/wake schedule may improve sleep and help to reset or stabilize the circadian clock. Not sleeping well at night and then napping during the day can cause a disturbed sleep/wake cycle.

Therefore, no matter what the age of the individual, a careful diagnosis needs to be made based on normal changes in sleep and on the complaints and history of that patient. Once the proper diagnosis is made, any sleep problem, regardless of the age of the individual, can be treated.

In conclusion, aging by itself does not cause sleep problems. Older adults do not need less sleep although sleep patterns do change with age. Rather, it is the ability to sleep that changes with age.

7.14.6 REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders (4th ed.).* Washington, DC: APA.

Ancoli-Israel, S. (1989). Epidemiology of sleep disorders. In T. Roth & T. A. Roehrs (Eds.), *Clinics in geriatric medicine* (pp. 347–362). Philadelphia: Saunders.

Ancoli-Israel, S. (1996). *All I Want is a Good Night's Sleep.* Chicago: Mosby-Year Book.

Ancoli-Israel, S., Bliwise, D. L., & Mant, A. (1993). Sleep and breathing in the elderly. In N. A. Saunders & C. E. Sullivan (Eds.), *Sleep and breathing* (pp. 673–693). New York: Dekker.

Ancoli-Israel, S., & Coy, T. V. (1994). Are breathing disturbances in elderly equivalent to sleep apnea syndrome? *Sleep, 17,* 77–83.

Ancoli-Israel, S., Jones, D. W., Hanger, M. A., Parker, L., Klauber, M. R., & Kripke, D. F. (1991). Sleep in the nursing home. In S. T. Kuna, P. M. Suratt, & J. E.

Remmers (Eds.), *Sleep and respiration in aging adults* (pp. 77–84). New York: Elsevier.

Ancoli-Israel, S., Klauber, M. R., Jones, D. W., Kripke, D. F., Martin, J., Mason, W., Pat-Horenczyk, R., & Fell, R. (1997). Variations in circadian rhythms of activity, sleep and light exposure related to dementia in nursing-home patients. *Sleep, 20*(1) 18–23.

Ancoli-Israel, S., Klauber, M. R., Stepnowsky, C., Estline, E., Chinn, A., & Fell, R. (1995). Sleep-disordered breathing in African–American elderly. *American Journal of Respiratory and Critical Care Medicine, 152*(6). 1946–1949.

Ancoli-Israel, S., Kripke, D. F., Klauber, M. R., Fell, R., Stepnowsky, C., Estline, E., Khazeni, N., & Chinn, A. (1996). Morbidity, mortality and sleep disordered breathing in community dwelling elderly. *Sleep, 19* (4) 277–282.

Ancoli-Israel, S., Kripke, D. F., Klauber, M. R., Mason, W. J., Fell, R., & Kaplan, O. (1991a). Sleep disordered breathing in community-dwelling elderly. *Sleep, 14*(6), 486–495.

Ancoli-Israel, S., Kripke, D. F., Klauber, M. R., Mason, W. J., Fell, R., & Kaplan, O. (1991b). Periodic limb movements in sleep in community-dwelling elderly. *Sleep, 14(6)*, 496–500.

Ancoli-Israel, S., Parker, L., Sinaee, R., Fell, R., & Kripke, D. F. (1989). Sleep fragmentation in patients from a nursing home. *Journal of Gerontology, 44*(1), M18–M21.

Ancoli-Israel, S., Seifert, A. R., & Lemon, M. (1986). Thermal biofeedback and periodic movements in sleep: Patients' subjective reports and a case study. *Biofeedback and Self-Regulation, 11,* 177–188.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression.* New York: Guilford Press.

Benca, R. M., Obermeyer, W. H., Thisted, R. A., & Gillin, J. C. (1992). Sleep and Psychiatric Disorders: A Meta-analysis. *Archives of General Psychiatry, 49,* 651–668.

Bixler, E. O., Kales, A., Vela-Bueno, A., Jacoby, J. A., Scarone, S., & Soldatos, C. R. (1982). Nocturnal myoclonus and nocturnal myoclonic activity in the normal population. *Research Communications in Chemical Pathology and Pharmacology, 36,* 129–140.

Bliwise, D. L. (1991). Cognitive function and sleep disordered breathing in aging adults. In S. T. Kuna, J. E. Remmers, & P. M. Suratt (Eds), *Sleep and respiration in aging adults* (pp. 237–244). New York: Elsevier.

Bliwise, D. L. (1993). Review: Sleep in normal aging and dementia. *Sleep, 16,* 40–81.

Bliwise, D. L., Bevier, W. C., Bliwise, N. G., Edgar, D. M., & Dement, W. C. (1990). Systemic 24-hr behavior observations of sleep and wakefulness in a skilled-care nursing facility. *Psychology and Aging, 15,* 16–24.

Bliwise, D. L., Feldman, D. E., Bliwise, N. G., Carskadon, M. A., Kraemer, H. C., North, C. S., Petta, D. E., Seidel, W. F., & Dement, W. C. (1987). Risk factors for sleep disordered breathing in heterogeneous geriatric populations. *Journal of the American Geriatrics Society, 35,* 132–141.

Bliwise, D. L., Hughes, M., McMahon, P. M., & Kutner, N. (1995). Observed sleep/wakefulness and severity of dementia in an alzheimer's disease special care unit. *Journal of Gerontology, 50A*(6), M303–M306.

Bootzin, R. R., & Engle-Friedman, M. (1987). Sleep disturbances. In B. A. Edelstein, & L. L. Carstensen (Eds), *Handbook of clinical gerontology*. New York: Pergamon.

Bootzin, R. R., & Nicassio, P. M. (1978). Behavioral treatments for insomnia. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds), *Progress in behavior modification* (Vol. 6, pp. 1–45). New York: Academic Press.

Borbely, A. A. (1994). Sleep homeostasis and models of sleep regulation. In M. H. Kryger, T. Roth, & W. C. Dement (Eds), *Principles and practice of sleep medicine* (pp. 309–320). Philadelphia: Saunders.

Campbell, S. S., Kripke, D. F., Gillin, J. C., & Hrubovcak, J. C. (1988). Exposure to light in healthy elderly subjects and Alzheimer's patients. *Physiology and Behavior, 42,* 141–144.

Campbell, S. S., Terman, M., Lewy, A. J., Dijk, D. J., Eastman, C. I., & Boulos, Z. (1995). Light treatment for sleep disorders: Consensus report. V. Age-related disturbances. *Journal of Biological Rhythms, 10*(2), 151–154.

Carskadon, M. A., & Dement, W. C. (1994). Normal human sleep: An overview. In M. H. Kryger, T. Roth, & W. C. Dement (Eds), *Principles and practice of sleep medicine*. Philadelphia: Saunders.

Carskadon, M. A., Dement, W. C., Mitler, M. M., Guilleminault, C., Zarcone, V. P., & Spiegel, R. (1976). Self-reports versus sleep laboratory findings in 122 drug-free subjects with complaints of chronic insomnia. *American Journal of Psychiatry, 133,* 1382–1388.

Cartwright, R. O., Diaz, F., & Lloyd, S. (1991). The effects of sleep posture and sleep stage on apnea frequency. *Sleep, 14*(4), 351–353.

Cohen-Mansfield, J., Werner, P., & Freedman, L. (1995). Sleep and agitation in agitated nursing home residents: An observational study. *Sleep, 18*(8), 674–680.

Coleman, R. M. (1982). Periodic movements in sleep (nocturnal myoclonus) and restless legs syndrome. In C. Guilleminault (Ed.), *Sleeping and waking disorders: Indications and techniques* (pp. 265–296). Menlo Park, CA: Addison-Wesley.

Coleman, R. M., Bliwise, D. L., Sajben, N., Boomkamp, A., de Bruyn, L. M., & Dement, W. C. (1982). Daytime sleepiness in patients with periodic movements in sleep. *Sleep, 5,* S191-S202.

Coleman, R. M., Pollak, C. P., & Weitzman, E. D. (1979). Periodic nocturnal myoclonus in a wide variety of sleep-wake disorders. *Transactions of the American Neurological Association, 103,* 230–233.

Czeisler, C. A., Kronauer, R. E., Allan, J. S., Duffy, J. F., Jewett, M. E., Brown, E. N., & Ronda, J. M. (1989). Bright light induction of strong (type 0) resetting of the human circadian pacemaker. *Science, 244,* 1328–1333.

Dawson, D., & Encel, N. (1993). Melatonin and sleep in humans (Review). *Journal of Pineal Research, 15,* 1–12.

Dement, W. C., Seidel, W., & Carskadon, M. A. (1982). Daytime alertness, insomnia and benzodiazepines. *Sleep, 5,* S28-S45.

Derderian, S. S., Bridenbaugh, R. H., & Rajagopal, K. R. (1988). Neuropsychologic symptoms in obstructive sleep apnea improve after treatment with nasal continuous positive airway pressure. *Chest, 94*(5), 1023–1027.

Diagnostic Classification Steering Committee (1990). *International classification of sleep disorders: Diagnostic and coding manual.* Rochester, MN: American Sleep Disorders Association.

Dufour, M. C., Archer, L., & Gordis, E. (1992). Alcohol and the elderly. *Clinics in Geriatric Medicine, 8*(1), 127–141.

Edinger, J. D., Hoelscher, T. J., Webb, M. D., Marsh, G. R., Radtke, R. A., & Erwin, C. W. (1989). Polysomnographic assessment of DIMS: empirical evaluation of its diagnostic value. *Sleep, 12*(4), 315–322.

Edinger, J. D., Morey, M. C., Sullivan, R. J., Higginbotham, M. B., Marsh, G. R., Dailey, D. S., & McCall, W. (1993). Aerobic fitness, acute exercise and sleep in older men. *Sleep, 16*(4), 351–359.

Espie, C. A. (1991). *The psychological treatment of insomnia.* Chicester, UK: Wiley.

Espie, C. A., Brooks, D. N., & Lindsay, W. R. (1989). An evaluation of tailored psychological treatment of insomnia. *Journal of Behavior Therapy and Experimental Psychiatry, 19,* 51–56.

Fairbanks, D. N. F., Fujita, S., Ikematsu, T., & Simmons, F. B. (1987). *Snoring and obstructive sleep apnea.* New York: Raven Press.

Ferber, R., Millman, R., Coppola, M., Fleetham, J., Murray, C. F., Iber, C., McCall, V., Nino-Murcia, G., Pressman, M., Sanders, M., Strohl, K., Votteri, B., & Williams, A. (1994). Portable recording in the assessment of obstructive sleep apnea. *Sleep, 17*(4), 378–392.

Findley, L. J., Presty, S. K., Barth, J. T., & Suratt, P. M. (1991). Impaired cognition and vigilance in elderly subjects with sleep apnea. In S. T. Kuna, P. M. Suratt, & J. E. Remmers (Eds), *Sleep and respiration in aging adults* (pp. 259–265). New York: Elsevier.

Foley, D. J., Monjan, A. A., Brown, S. L., Simonsick, E. M., Wallace, R. B., & Blazer, D. G. (1995). Sleep complaints among elderly persons: an epidemiologic study of three communities. *Sleep, 18*(6), 425–432.

Ford, D. E., & Kamerow, D. B. (1989). Epidemiologic study of sleep disturbances and psychiatric disorders: An opportunity for prevention? *Journal of the American Medical Association, 262*(11), 1479–1484.

Friedman, L., Bliwise, D. L., Yesavage, J. A., & Salom, S. R. (1991). A preliminary study comparing sleep restriction and relaxation treatments for insomnia in older adults. *Journal of Gerontology, 46*(1), P1-P8.

Garfinkel, D., Laudon, M., Nuf, D., & Zisapel, N. (1995). Improvement of sleep quality in elderly people by controlled-release melatonin. *Lancet, 346,* 541–549.

Gillin, J. C., & Ancoli-Israel, S. (1992). The impact of age on sleep and sleep disorders. In C. Salzman (Ed.), *Clinical psychopharmacology* (pp. 213–234). Baltimore: Williams & Wilkins.

Gillin, J. C., & Byerley, W. F. (1990). The diagnosis and management of insomnia. *New England Journal of Medicine, 322,* 239–248.

Gillin, J. C., Duncan, W. C., Murphy, D. L., Post, R. M., Wehr, T. A., Goodwin, F. K., Wyatt, R. J., & Bunney, W. E. J. (1981). Age-related changes in sleep in depressed and normal subjects. *Psychiatry Research, 4,* 73–78.

Glotzbach, S. F., & Heller, H. C. (1996). Temperature regulation. In M. H. Kryger, T. Roth, & W. C. Dement (Eds), *Principles and practice of sleep medicine* (pp. 260–275). Philadelphia: Saunders

Glovinsky, P. B., & Spielman, A. J. (1991). Sleep restriction therapy. In P. J. Hauri (Ed.), *Case studies in insomnia* (pp. 49–63). New York: Plenum.

Haimov, I., Lavie, P., Laudon, M., Herer, P., Vigder, C., & Zisapel, N. (1995). Melatonin replacement therapy of elderly insomniacs. *Sleep, 18*(7), 598–603.

Hauri, P. J. (1981). Treating psychophysiologic insomnia with biofeedback. *Archives of General Psychiatry, 38,* 752–758.

Hauri, P. J. (1983). A cluster analysis of insomnia. *Sleep, 6*(4), 326–338.

Hauri, P. J. (1991). Sleep hygiene, relaxation therapy, and cognitive interventions. In P. J. Hauri (Ed.), *Case studies in insomnia* (pp. 65–84). New York: Plenum.

Hauri, P., & Linde, S. (1990). *No more sleepless nights.* New York: Wiley.

Hauri, P. J., & Wisbey, J. (1992). Wrist actigraphy in insomnia. *Sleep, 15*(4), 293–301.

Hauri, P. J., & Wisbey, J. (1994). Actigraphy and insomnia: A closer look, Part 2. *Sleep, 17*(5), 408–410.

Hayward, L., Mant, A., Eyland, A., Hewitt, H., Purcell, C., Turner, J., Goode, E., Le Count, A., Pond, D., & Saunders, N. (1992). Sleep disordered breathing and cognitive function in a retirement village population. *Age and Ageing, 21*(2), 121–128.

He, J., Kryger, M. H., Zorick, F. J., Conway, W., & Roth, T. (1988). Mortality and apnea index in obstructive sleep apnea: Experience in 385 male patients. *Chest, 94,* 9–14.

Hicks, R., Dysken, M. W., Davis, J. M., Lesser, J., Ripeckyj, A., & Lazarus, L. (1981). The pharmacokinetics of psychotropic medication in the elderly: A review. *Journal of Clinical Psychiatry, 42,* 374–385.

Hoch, C. C. (1991). Sleep in old age. In E. Murrow-Baines (Ed.), *Perspectives on gerontological nursing* (pp. 59–77). Newbury Park, CA: Sage.

Hoffstein, V. (1990). Snoring and hypertension. In F. G. Issa, P. M. Suratt, & J. E. Remmers (Eds), *Sleep and Respiration* (pp. 371–376). New York: Wiley-Liss.

Jacobs, D., Ancoli-Israel, S., Parker, L., & Kripke, D. F. (1989). Twenty-four-hour sleep-wake patterns in a nursing home population. *Psychology and Aging, 4*(3), 352–356.

Kaplan, P. W., Allen, R. P., Buchholz, D. W., & Walters, J. K. (1993). A double-blind, placebo-controlled study of the treatment of periodic limb movements in sleep using carbidopa/levidopa and propoxyphene. *Sleep, 16*(8), 717–723.

Kavey, N., Walters, A. S., Hening, W., & Gidro-Frank, S. (1988). Opioid treatment of periodic movements in sleep in patients without restless legs. *Neuropeptides, 11*(4), 181–184.

Klink, M. E., Quan, S. F., Kaltenborn, W. T., & Lebowitz, M. D. (1992). Risk factors associated with complaints of insomnia in a general adult population. Influence of previous complaints of insomnia. *Archives of Internal Medicine, 152*(8), 1634–1637.

Kronholm, E., & Hyyppa, M. T. (1985). Age-related sleep habits and retirement. *Annals of Clinical Research, 17,* 257–264.

Kryger, M. H. (1989). Management of obstructive sleep apnea: overview. In M. H. Kryger, T. Roth, & W. C. Dement (Eds), *Principles and practice of sleep medicine* (pp. 584–590). Philadelphia: Saunders.

Lacks, P., Bertelson, A. D., Gans, L., & Kunkel, J. (1983). The effectiveness of three behavioral treatments for different degrees of sleep onset insomnia. *Behavior Therapy, 14,* 593–605.

Lavie, P., Herer, P., Peled, R., Berger, I., Yoffe, N., Zomer, J., & Rubin, A. E. (1995). Mortality in sleep apnea patients: A multivariate analysis of risk factors. *Sleep, 18*(3), 149–157.

Lichstein, K. L., & Fanning, J. (1990). Cognitive anxiety in insomnia: An analogue test. *Stress Medicine, 6,* 47–51.

Lichstein, K. L., & Riedel, B. W. (1994). Behavioral assessment and treatment of insomnia: A review with an emphasis on clinical application. *Behavior Therapy,* 25, 659–688.

Lowe, A. A. (1994). Dental appliances for the treatment of snoring and obstructive sleep apnea. In M. H. Kryger, T. Roth, & W. C. Dement (Eds), *Principles and practice of sleep medicine* (pp. 722–735). Philadelphia: Saunders.

McClusky, H. Y., Milby, J. B., Switzer, P. K., Williams, V., & Wooten, V. (1991). Efficacy of behavioral versus triazolam treatment in persistent sleep-onset insomnia. *American Journal of Psychiatry, 148,* 121–126.

McGhie, A., & Russell, S. M. (1962). The subjective assessment of normal sleep patterns. *Journal of Mental Science, 108,* 642–654.

Mellinger, G. D., Balter, M. B., & Uhlenhuth, E. H. (1985). Insomnia and its treatment. Prevalence and correlates. *Archives of General Psychiatry, 42,* 225–232.

Miles, L., & Dement, W. C. (1980). Sleep and aging. *Sleep, 3,* 119–220.

Mitler, M. M., Browman, C. P., Menn, S. J., Gujavarty, K., & Timms, R. M. (1986). Nocturnal myoclonus: Treatment efficacy of clonazepam and temazepam. *Sleep, 9,* 385–392.

Monjan, A., & Foley, D. (1996). Incidence of chronic insomnia associated with medical and psychosocial factors: An epidemiologic study among older persons. *Sleep Research, 25,* 108.

Morin, C. M. (1993). *Insomnia: Psychological assessment and management*. New York: Guilford Press.

Morin, C. M., & Azrin, N. H. (1988). Behavioral and cognitive treatments of geriatric insomnia. *Journal of Consulting and Clinical Psychology, 56*(5), 748–753.

Morin, C. M., Culbert, J. P., & Schwartz, S. M. (1994). Nonpharmacological interventions for insomnia: A meta-analysis of treatment efficacy. *American Journal of Psychiatry, 151*(8), 1172–1180.

Morin, C. M., Kowatch, R. A., Barry, T., & Walton, E. (1993). Cognitive-behavior therapy for late-life insomnia. *Journal of Consulting and Clinical Psychology, 61*(1), 137–146.

Morin, C. M., Stone, J., Trinkle, D., Mercer, J., & Remsberg, S. (1993). Dysfunctional beliefs and attitudes about sleep among older adults with and without insomnia complaints. *Psychology and Aging, 8*(3), 463–467.

Mosko, S. S., Dickel, M. J., & Ashurst, J. (1988). Night-to-night variability in sleep apnea and sleep in sleep-related periodic leg movements in the elderly. *Sleep, 11,* 340–348.

Murtagh, D. R., & Greenwood, K. M. (1995). Identifying effective psychological treatments for insomnia: A meta-analysis. *Journal of Consulting Clinical Psychology, 63*(1), 79–89.

Myers, B. L., & Badia, P. (1995). Changes in circadian rhythms and sleep quality with aging–-mechanisms and interventions. *Neuroscience and Biobehavioral Reviews, 19*(4), 553–571.

Nakao, M., McGinty, D., Szymusiak, R., & Yamamoto, M. (1995). Dynamical features of thermoregulatory model of sleep control. *Japanese Journal of Physiology, 45,* 311–326.

Ohayon, M. M., & Caulet, M. (1995). Insomnia and psychotropic drug consumption. *Progress in Neuro-Psycharmacology and Biological Psychiatry, 19,* 421–431.

Pat-Horenczyk, R., & Ancoli-Israel, S. (1996). 24 Hour sleep/wake in demented patients: A multidimensional scaling analysis. *Proceedings of Society for Research on Biological Rhythms,* 130.

Pat-Horenczyk, R., Jones, D. W., Martin, J., & Ancoli-Israel, S. (1996). Hour-by-hour sleep in demented nursing home residents. *Sleep Research, 25,* 130.

Petit, D., Montplaisir, J., Lorrain, D., & Gauthier, S. (1992). Temporal asymmetry in the awake and REM sleep EEG in mild to moderate Alzheimer's disease. *Sleep Research, 21,* 305.

Pollak, C. P., & Perlick, D. (1991). Sleep problems and institutionalization of the elderly. *Journal of Geriatric Psychiatry and Neurology, 4,* 204–210.

Powell, N. B., Guilleminault, C., & Riley, R. W. (1994). Surgical therapy for obstructive sleep apnea. In M. H. Kryger, T. Roth, & W. C. Dement (Eds), *Principles and practice of sleep medicine* (pp. 706–721). Philadelphia: Saunders.

Prinz, P. N. (1977). Sleep patterns in the healthy aged: Relationship with intellectual function. *Journal of Gerontology, 32,* 179–186.

Prinz, P. N., & Vitiello, M. V. (1989). Dominant occipital (alpha) rhythm frequency in early stage Alzheimer's disease and depression. *Electroencephalography and Clinical Neurophysiology, 73,* 427–432.

Prinz, P. N., Vitiello, M. V., Raskind, M. A., & Thorpy, M. J. (1990). Geriatrics: Sleep disorders and aging. *New England Journal of Medicine, 323*(8), 520–526.

Rechtschaffen, A. (1994). Mammalian sleep. In M. H. Kryger, T. Roth, & W. C. Dement (Eds), *Principles and practices of sleep medicine* (pp. 69–80). Philadelphia: Saunders.

Rediehs, M. H., Reis, J. S., & Creason, N. S. (1990). Sleep in old age: Focus on gender differences. *Sleep, 13*(5), 410–424.

Regestein, Q. R., & Morris, J. (1987). Daily sleep patterns observed among institutionalized elderly residents. *Journal of the American Geriatrics Society, 35,* 767–772.

Reiter, R. J. (1995). The pineal gland and melatonin in relation to aging: A summary of the theories and of the data. *Experimental Gerontologist, 30,* 199

Reynolds, C. F.I. (1989). Sleep in affective disorders. In M. H. Kryger, T. Roth, & W. C. Dement (Eds), *Principles and practice of sleep medicine* (pp. 413–415). Philadelphia: Saunders.

Reynolds, C. F., Hoch, C. C., Kupfer, D. J., Buysse, D. J., Houck, P. R., Stack, J. A., & Campbell, D. W. (1988). Bedside differentiation of depressive pseudodementia from dementia. *American Journal of Psychiatry, 145,* 1099–1103.

Reynolds, C. F., Kupfer, D. J., Houck, P. R., Hoch, C. C., Stack, J. A., Berman, S. R., & Zimmer, B. (1988). Reliable discrimination of elderly depressed and demented patients by electroencephalographic sleep data. *Archives of General Psychiatry, 45,* 258–264.

Riedel, B. W., Lichstein, K. L., & Dwyer, W. O. (1995). Sleep compression and sleep education for older insomniacs: Self-help versus therapist guidance. *Psychology and Aging, 10*(1), 54–63.

Roth, T., Roehrs, T., & Zorick, F. (1988). Pharmacological treatment of sleep disorders. In R. L. Williams, I. Karacan, & C. A. Moore (Eds), *Sleep disorders: Diagnosis and treatment* (pp. 373–395). New York: Wiley.

Sadeh, A., Hauri, P. J., Kripke, D. F., & Lavie, P. (1995). The role of actigraphy in the evaluation of sleep disorders. *Sleep, 18,* 288–302.

Shepard, J. W. J. (1989). Cardiorespiratory changes in obstructive sleep apnea. In M. H. Kryger, T. Roth, & W. C. Dement (Eds), *Principles and practice of sleep medicine* (pp. 537–551). Philadelphia: Saunders.

Shepard, J. W. J. (1992). Hypertension, cardiac arrhythmias, myocardial infarction, and stroke in relation to obstructive sleep apnea. *Clinics in Chest Medicine, 13*(3), 437–458.

Spielman, A. J., Saskin, P., & Thorpy, M. J. (1987). Treatment of chronic insomnia by restriction of time in bed. *Sleep, 10,* 45–56.

Standards of Practice Committee of the American Sleep Disorders Association. (1994). Practice parameters for the use of laser-assisted uvulopalatoplasty. *Sleep, 17*(8), 744–748.

Sullivan, C. E., & Grunstein, R. R. (1989). Continuous positive airway pressure in sleep-disordered breathing. In M. H. Kryger, T. Roth, & W. C. Dement (Eds), *Principles and practice of sleep medicine* (pp. 559–570). Philadelphia: Saunders.

Taasan, V. C., Block, A. J., Boysen, P. G., & Wynne, J. W. (1981). Alcohol increases sleep apnea and oxygen desaturation in asymptomatic men. *American Journal of Medicine, 71,* 240–245.

Vitiello, M. V. (1996). Sleep disorders and aging. *Current Opinion in Psychiatry, 9*(4), 284–289.

Walters, A. S., Aldrich, M. S., Allen, R., Ancoli-Israel, S., Buchholz, D. W., Chokroverty, S., Coccagna, G., Earley, C., Ehrenberg, B., Feest, T. G., Hening, W., Kavey, N., Lavigne, G., Lipinski, J., Lugaresi, E., Montagna, P., Montplaisir, J., Mosko, S. S., Oertel, W., Picchietti, D., Pollmacher, T., Shafor, R., Smith, R. C., Telstad, W., Trenkwalder, C., von Scheele, C., Ware, J. C., & Zucconi, M. (1995). Toward a better definition of the restless legs syndrome. *Movement Disorders, 10*(5), 634–642.

Wauquier, A., Van Sweden, B., Lagaay, A. M., Kemp, B., & Kamphuisen, H. A.C. (1992). Ambulatory monitoring of sleep-wakefulness patterns in healthy elderly males and females (>88 years): The "Senieur" protocol. *Journal of the American Geriatrics Society, 40,* 109–114.

Wooten, V. (1994). Medical causes of insomnia. In M. H. Kryger, T. Roth, & W. C. Dement (Eds), *Principles and Practice of sleep medicine*. (pp. 456–475). Philadelphia: Saunders.

Young, T., Palta, M., Dempsey, J., Skatrud, J., Weber, S., & Badr, S. (1993). Occurrence of sleep disordered breathing among middle-aged adults. *New England Journal of Medicine, 328,* 1230–1235.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.15

Aging with Intellectual Disabilities and Later-life Family Caregiving

TAMAR HELLER

University of Illinois at Chicago, IL, USA

7.15.1 INTRODUCTION 328
  7.15.1.1 Overview of Chapter 328
  7.15.1.2 Context of Family Caregiving 328
7.15.2 AGING IN ADULTS WITH INTELLECTUAL DISABILITIES 329
  7.15.2.1 Life Expectancy and Demographic Trends 329
    7.15.2.1.1 Growing numbers of older adults with intellectual disabilities 329
    7.15.2.1.2 Increased life expectancy 329
  7.15.2.2 Age-related Changes 329
    7.15.2.2.1 Biological and psychosocial changes 329
    7.15.2.2.2 Earlier aging for some conditions 330
7.15.3 TRENDS IN SERVICES AND SUPPORTS 330
  7.15.3.1 Models of Care and Residential Services 330
    7.15.3.1.1 Growth of community vs. institutional programs 330
    7.15.3.1.2 Waiting lists and unmet needs 331
  7.15.3.2 Community Inclusion Models for Older Adults 331
    7.15.3.2.1 Integration into the aging service network 332
    7.15.3.2.2 Inclusion into the general community 332
7.15.4 FAMILY CAREGIVING ACROSS THE LIFE-SPAN 333
  7.15.4.1 Changes across the Life-span 333
    7.15.4.1.1 Theories of family caregiving over the life-span 333
    7.15.4.1.2 Findings on families of adults with intellectual disabilities 333
    7.15.4.1.3 Predictors of caregiver well-being in later life 335
  7.15.4.2 Roles of Mother, Fathers, and Siblings 335
    7.15.4.2.1 Comparison of maternal and paternal caregiving 335
    7.15.4.2.2 Sibling caregiver roles 336
  7.15.4.3 Future Planning and Placement 337
    7.15.4.3.1 Residential plans 338
    7.15.4.3.2 Predictors and outcomes of placement 338
    7.15.4.3.3 Legal and financial planning 339
  7.15.4.4 Sociocultural Issues 339
    7.15.4.4.1 Comparisons across countries 339
    7.15.4.4.2 Culturally diverse groups in the USA 340
7.15.5 FAMILY SUPPORT INTERVENTIONS AND POLICIES 340
  7.15.5.1 Family Needs and Service Use 341
  7.15.5.2 Progam Models 341
    7.15.5.2.1 Support and psycho-educational groups 341
    7.15.5.2.2 Outreach through the aging network 342
    7.15.5.2.3 Family support public programs 342
7.15.6 RESEARCH DIRECTIONS 343
  7.15.6.1 Summary of Research 343
  7.15.6.2 Research Agenda 343
7.15.7 REFERENCES 344

7.15.1 INTRODUCTION

7.15.1.1 Overview of Chapter

Adults with intellectual disabilities have needs for lifetime support and may require personal care, long-term health and mental health care, economic support, and social and housing services. The most consistent source of this care and support is the family. Nearly 80% of adults with intellectual disabilities live with their families (Fujiura, 1996). Although the likelihood of placement out of the home increases with age, family care still predominates for older adults with intellectual disabilities, at least while the parents are still alive (Meyers, Borthwick, & Eyman, 1985).

This chapter examines later-life caregiving for adults with intellectual disabilities from a life-course perspective that takes into account the social trends and changes in service patterns that effect these families. Older families have experienced pervasive changes in the context of caregiving. This life-course perspective aids in understanding their attitudes towards caregiving and their interface with the service system. The term "caregiving" includes both meeting socioemotional and instrumental needs. While we most often think of caregiving as unidirectional, in the case of older families caregiving can also be bidirectional encompassing support provided by the adult with intellectual disabilities to parents who are aging.

The introduction will present the sociodemographic trends affecting family caregiving. Subsequent sections will include: (i) description of life expectancy, demographic trends, and age-related changes for adults with intellectual disabilities; (ii) trends in services and supports and model of community inclusion; (iii) family caregiving across the life-course for mothers, fathers, and siblings and for various culturally diverse groups; (iv) support interventions that help adults with intellectual disabilities and their families cope with later-life transitions, including planning for the future; and (v) future research agenda.

7.15.1.2 Context of Family Caregiving

In the USA, the composition and structure of families have undergone major transformations since the 1950s that have changed families' abilities to provide care for dependent adults with disabilities. The traditional nuclear family consisting of a father who works out of the home and a mother at home providing care is in the minority. Rather, we see an increase in dual-income, single-parent, and step-families. We also see growing numbers of minority families and families living in poverty. At the same time increases in life expectancy extend the number of years for which families provide care for other members.

Increased life expectancy and lower fertility rates have contributed to the "aging" of the US population, with the median age rising from 28 years in 1970 to 33 in 1990. In 1950 the average life-course for persons born that year was 68 years. By 1990, average life expectancy had reached 78.8 years for women and 71.9 years for men. Projections are that average life expectancy will rise to 78 years by 2005 (US Bureau of the Census, 1992f). The percentage of adults aged 60 years and older is expected to increase from 17% of the US population in 1990 to 25% by the year 2030. The increase will be most dramatic for the oldest-old, with those over age 85 years projected to increase their numbers from 3.1 million in 1990 to 8 million by the year 2030 (US Bureau of the Census, 1992c). This growth of the elderly population will have a major impact on caregiving responsibilities within families across multiple generations as persons may need to provide care for parents, grandparents, and potentially, great-grandparents. This trend coupled with lower fertility rates, which have dropped from 3.6 children during the "Baby Boom Period" (1952–1964) to 1.8 in 1988, results in more demands for caregiving in later life and fewer siblings and children from which to draw support.

With an increasing number of single women bearing children, rising divorce rates, and higher life expectancy for women (Rossi, 1987), there has been a large increase in single parent households. Between 1970 and 1990 the percentage of US households with children under 18 years old headed by a single parent increased from 10% to 30%. This rise in single family households is most marked among African-American households, where by 1990 61% comprised single-parent households (US Bureau of the Census, 1992d).

Labor rate participation among women with children has increased in the last two decades. From 1975 to 1990 the percentage of married women with children under the age of 6 years who worked in paid employment increased from 37% to 59%. Across all ages 57% of married women and 68% of single women are in the workforce (US Bureau of the Census, 1992a). Hence, while the responsibilities of families across generations continue to increase, both the extent to which families can draw upon other family members for support and the time available for women caregivers has decreased. Furthermore, increased economic pressures on families have contributed to stress on caregivers. The economic situation has worsened for lower-income families, as the gap between the rich and poor has widened. As of 1990 13.5% of families in the USA were below the poverty level (US Bureau of the Census, 1992b).

There has been a rapid growth rate among racial/ethnic minority groups in the USA. Between 1980 and 1990 the growth rate was 108% for Asians, 53% for Hispanics, 38% for Native Americans, 13% for African-Americans, and only 6% for white non-Hispanics (Braddock, 1991). It is estimated that by the year 2050 the non-Hispanic white population will decline to only 53% of the US population (US Bureau of the Census, 1992e). Hence, any analysis of the caregiving experience for family members with intellectual disabilities needs to take into account differences among families attributed to cultural diversity and diverse family structures.

7.15.2 AGING IN ADULTS WITH INTELLECTUAL DISABILITIES

7.15.2.1 Life Expectancy and Demographic Trends

7.15.2.1.1 Growing numbers of older adults with intellectual disabilities

Intellectual disabilities (commonly referred to as mental retardation in the USA) is defined by the American Association on Mental Retardation as significant subaverage general intellectual functioning existing concurrently with deficits in adaptive behaviors which are first manifested before the age of 18 years (Luckasson et al., 1993). About 20% of adults with intellectual disabilities require assistance in performing basic activities of daily living such as dressing and bathing (Center for Vulnerable Populations, 1993). Generally, persons with intellectual disabilities show stability in their functional and cognitive levels in adulthood, at least prior to age-related changes in later life (Eyman & Widaman, 1987). Estimates of the prevalence of intellectual disabilities range from 1% to 3% of the population, depending on whether persons with mild forms of intellectual disabilities are included. The prevalence appears to be similar across ages, with some decrease at the oldest ages. However, the proportion of persons with intellectual disabilities who show up in the service system varies, with the greatest proportion served in the school age (McClaren & Bryson, 1987). It has been estimated that as many as 40% of adults with intellectual disabilities are not known to the service system (Krauss, 1986).

The number of older people with intellectual disabilities is not known. However, it is estimated that the age distribution will mirror that of the general population, with a large growth in the elderly population. Using the prevalence rate of 1.2% (which includes both persons known and not known to the service system) recommended by McClaren and Bryson (1987) it is estimated that the number of adults with intellectual disabilities aged 60 years and over in the US will increase from 503 000 in 1990 to 1 065 000 in 2030.

7.15.2.1.2 Increased life expectancy

Historically, persons with intellectual disabilities have had short life-spans, with many dying in childhood (Eyman & Borthwick-Duffy, 1994). With improved medical interventions, better social conditions, and technological advances, the life expectancy of adults with intellectual disabilities has improved substantially. From 1931 to 1980 the life expectancy increased from 15 years for males and 22 years for females to 58 years and 60 years, respectively, in 1976–1980 (Carter & Jancar, 1983). The increase had been most marked for persons with Down syndrome. Their life expectancy has risen from 9 years in 1929 to 55 years by 1995 (Adlin, 1993). Adults with milder forms of intellectual disabilities who do not have other physical disabilities, generally have the same life-span as the general population (Eyman & Borthwick-Duffy, 1994). Exceptions are people with Down syndrome, cerebral palsy, and more severe forms of intellectual disabilities (Heller, 1997).

With the longer life-spans of both adults with intellectual disabilities and their relatives, the period of family caregiving is much longer than in the past. Also, the likelihood that adult children with intellectual disabilities will outlive their parents has increased, with most expected to outlive their parents by several decades.

7.15.2.2 Age-related Changes

7.15.2.2.1 Biological and psychosocial changes

It is commonly assumed that persons with intellectual disabilities experience precocious aging. Yet, the research indicates that this phenomenon is only true for some subgroups, such as those with Down syndrome, severe levels of cognitive disability, cerebral palsy, or epilepsy (Heller, 1997). Persons with milder forms of intellectual disabilities experience age-related problems similar to those of the general population. In a large population study of adults with intellectual disabilities in Great Britain (Hogg & Moss, 1993) chronic diseases increased from age 50 years on, hearing loss from age 60 on, and mobility loss and incontinence from age 70 on. When compared with a general elderly group, older adults (age 50 years and over) with moderate to severe intellectual disabilities had a higher rate of incontinence and obesity, and a lower rate of angina (Wilkin, Durie, Wade, Jolley, & Stout, 1984). Cross-sectional analyses of age-related differences among adults with intellectual disabilities generally find a bias due to differential mortality, with those surviving to old age generally having better health and higher levels of cognitive ability.

Generally, studies that have examined adaptive behavior, maladaptive behavior, and cognitive changes among adults with intellectual disabilities have found few age-related changes, when persons with Down syndrome were not included. Researchers have noted some gradual decline in overall intellectual functioning after the age of 65 years (Fisher & Zeaman, 1970; Walz, Harper, & Wilson, 1986).

Losses in social support and in physical functioning can contribute to psychiatric disorders for older adults with intellectual disabilities. Often issues of death and bereavement are not adequately addressed in this population. Also, psychiatric problems tend to be underdiagnosed among adults with intellectual disabilities (Patel, Goldberg, & Moss, 1993). After thoroughly screening a community population of older adults with intellectual disabilities (age 50 and over) Patel et al. (1993) diagnosed 11.4% of the sample as having a psychiatric disorder (mostly depression and anxiety) other than dementia. Only one third of these cases had been identified and treated for their disorder.

7.15.2.2.2 Earlier aging for some conditions

Earlier aging occurs for persons with Down syndrome and to some extent for persons with cerebral palsy and epilepsy. Often, age-related changes can be due to secondary conditions related to the disability. For example, long-term use of antiseizure medication appears to be associated with an increased risk of osteoporosis and fractures (Adlin, 1993). For adults with Down syndrome there is a higher prevalence of Alzheimer's disease at a younger age (starting with age 50 years), which also is related to decrements in adaptive behaviors and increased maladaptive behavior (Janicki, Heller, Seltzer, & Hogg, 1996). Persons with Down syndrome also experience other age-related health problems, including hypothyroidism, depression, and sleep apnea, which can mimic the symptoms of dementia (Adlin, 1993).

For adults with cerebral palsy, aging-related changes have been noted starting in the 40s. These include decreased muscle strength and mobility, problems with choking and swallowing, increased pain, and incontinence (Turk, Overeynder, & Janicki, 1995). It is not uncommon for persons with cerebral palsy to begin using a wheelchair in middle age, as walking becomes more difficult and fatigue is greater.

7.15.3 TRENDS IN SERVICES AND SUPPORTS

7.15.3.1 Models of Care and Residential Services

7.15.3.1.1 Growth of community vs. institutional programs

Older adults with intellectual disabilities and their families have witnessed major changes in the service system and in ideologies concerning best models of care and support. For over 120 years from 1850 to the 1970s the dominant paradigm of care for persons with intellectual disabilities emphasized institutionalization and segregation. They often were sent to far-away institutions and special schools. Medical professionals dominated and made most of the decisions regarding the medical and educational needs of the children. Many older parents recall their early encounters with physicians who told them that their child had to be institutionalized because of the intellectual disabilities diagnosis. During this era there were no publicly funded schools or day/vocational programs for children with intellectual disabilities and few supports for families who chose to maintain their relative at home. At the height of this "institutional" era in 1967 there were 194 650 children and adults with intellectual disabilities and developmental disabilities in institutions (Braddock, Hemp, Bachelder, & Fujiura, 1995).

Beginning in the 1950s parents began to challenge the institutional and segregated model. They created the national Association for Retarded Children (renamed The Arc) which is the fourteenth-largest nonprofit organization in the USA. Largely through the efforts of the parent movement and class action lawsuits, legislation was passed that provided public education services to children with disabilities, including Public Law 194-142 in 1975, which guaranteed free and appropriate public education for all children with disabilities aged 3 to 21 years. Although older adults with disabilities were not able to benefit directly from this law as they already were adults by the time this law was passed, they have benefitted from its influence on families' expectations and advocacy for adult services.

The period 1976–1986 has been characterized as one of deinstitutionalization and community development (Bradley & Knoll, 1990). This phase emphasized a shift from the medical model to the developmental model, with specialized services that trained people for integration into more independent settings. During this era there were efforts to phase down or close large institutions and to develop other residential programs in the community, including large group homes and living centers. Concurrently, there was a major growth in schools, sheltered workshops, and activity centers located within the community. The model implied that there was a "continuum of services" and that persons needed to learn skills to advance to the less restrictive level of care. A major problem with this model was that without appropriate environmental supports most adults never "progressed" to the next level.

The 1990s model which is being promulgated is one of community membership. This approach emphasizes individualized functional supports to enhance community inclusion and quality of life. Services are being reconfigured to incorporate more emphasis on the role of informal family and community supports, consumer control, and individualized approaches. Rather than emphasizing skill development learned in specialized settings, its emphasis is on providing supports to individuals with disabilities in the natural setting, such as in supported employment, supported living arrangements, assistive technology and environmental accommodations, and family supports. There has also been an emphasis on civil rights for people with disabilities. This has been particularly evident with the passage of the Americans with Disabilities Act of 1990 (PL 101-336), which prohibits discrimination against individuals with disabilities in enjoyment of goods, services, privileges, and public accommodation.

Since the peak era of institutionalization there has been a steady decline in the institutional population and a concomitant increase in community-based services. By 1992 the majority of persons with intellectual disabilities living in out-of-home placements were living in small versus larger residences (170 077 vs. 165 119) (Braddock et al., 1995). However, older adults are much more likely than younger persons to be living in larger residential settings (Heller, 1997). Furthermore, there have also been federally mandated transfers of adults with developmental disabilities out of nursing homes and into other community-based settings. The Nursing Home Reform Act, part of the Omnibus Budget Reconciliation Act of 1987, required that all residents with a developmental disability who did not have skilled nursing medical needs move out, unless they had lived in the nursing home for over 30 months and chose to stay there. Between 1988 and 1992 their numbers decreased from 50 606 to 41 429 (Braddock et al., 1995).

7.15.3.1.2 Waiting lists and unmet needs

With the massive phase down of persons with intellectual disabilities from institutions and nursing homes and the lack of sufficient growth of alternative options to meet growing needs for residential supports, there has been a concomitant increase in the waiting list for residential services. A national survey indicated that in 1992, 60 876 people with developmental disabilities were on waiting lists for residential services in 37 states (Hayden, 1992). In addition, there is a big demand for day programs, with the same survey indicating that over 35 095 adults with developmental disabilities were on waiting lists for day habilitation or vocational services.

Services tend to go to persons already in the service system. Often, the older families who have chosen to rear their children at home and who have continued to care for them into adulthood face long waiting lists when they begin to look for placements. Furthermore, the actual demand is probably underestimated by waiting lists, as many families are unknown to the system (Krauss, 1986) and many other families, especially older ones, are wary of formal services. Some of this can be attributed to their earlier experiences with poor or nonexistent services when their children were younger. Other families may be unaware of the available options. Several outreach studies have identified numerous families who previously had not received any formal day or residential services for their relative with intellectual disabilities (e.g., Smith, Fullmer, & Tobin, 1994; Wood & Skiles, 1992). In a 1992 study of families age 55 and over caring for children with intellectual disabilities, only 33% had any involvement with any provider agencies (Wood & Skiles, 1992).

7.15.3.2 Community Inclusion Models for Older Adults

Paralleling the changing concepts of care for persons with intellectual disabilities, models for community inclusion of older adults with intellectual disabilities also have been rapidly changing. In the late 1970s through the 1980s there was a growing recognition that the population of older adults with intellectual disabilities was increasing and that they may require specialized services targeted to them. This segregated model focused on special senior programs within workshops or residential facilities. These programs were often modeled after seniors programs in the general population. The next phase emphasized community inclusion, either in terms of being integrated into the generic aging services programs or in terms of being included in the general community.

7.15.3.2.1 Integration into the aging service network

Integration into the aging service network has occurred on many levels: (i) federal and statewide legislation and memorandums of agreements in which the aging and disability networks agree to work together, (ii) integration of adults with intellectual disabilities into seniors residential and day programs, and (iii) combined outreach to families.

Amendments to the Developmental Disabilities Assistance and Bill of Rights Act (PL 100-142) and the Older American Act (PL 89-73) in the USA have fostered collaborative program planning and development across the aging and intellectual disabilities service systems. Examples of such programs include senior center/senior companion programs, nutrition programs, and senior adult day programs. The senior companion programs, in which staff, volunteers, or friends involve adults with intellectual disabilities in these various senior services and in other community activities has generally documented improvements in the social integration and self-esteem of participants (Roberts, 1990).

Outreach to older adults and their families has been successfully implemented through collaboration of both networks. For example, the Caregiver Resource Program in New York has outreached to numerous older family caregivers of adults with intellectual disabilities through the Area Agencies on Aging (Janicki, 1996). Many of the families who had never been linked to any formal services were more responsive to the aging network than to the intellectual disabilities service network. This could be attributed to lack of knowledge, previous negative experiences, or to feelings that their needs were primarily related to aging-related changes rather than to their child's intellectual disabilities.

7.15.3.2.2 Inclusion into the general community

Fitting in with the emerging model of community membership, other approaches to services and supports to older adults with intellectual disabilities focus on individualized programs that offer a wide variety of options. These can include making use of intellectual disabilities services, aging services, or services and activities offered to the general community. Examples of innovative programs using such an approach are the Community Entry Project (Blaney, 1995) and the Person-Centered Later Life Planning Project (Heller, Sterns, Sutton, & Factor, 1996). The Community Entry Project worked with older adults in creating a circle of support, which could include parents, guardians, staff members, friends, and service coordinators. This group then helped the individual connect with community activities in natural settings, such as YMCA programs, church groups, volunteer activities, and art classes. Another example, the Person-Centered Later Life Planning Project has developed a training program which helps older adults with intellectual disabilities become more informed about community options and better able to make choices and goals regarding their future. Also, it provides opportunities for greater choice-making through training of staff and families. Research examining the effectiveness of this program has indicated its success in increasing participants' knowledge of later-life options and increasing participation in leisure activities for participants living at home. Additionally, the training resulted in greater consideration of the preferences of older adults with intellectual disabilities in the actual goals developed in planning meetings with staff and families.

With the emerging models of community membership there is a greater emphasis on involving natural supports, including family and friends, than in previous segregated models that emphasized professional involvement. Hence, families are expected to be integrally involved in developing service plans, providing support, and in some cases directing services. This chapter more specifically discusses family support programs and new models of family consumer-direction in a later section.

In summary, the first three sections of this chapter provide a general overview of the historical changes that have effected the context of family caregiving for adults with intellectual disabilities. These have included changes in demographics and in service ideologies and trends. In addition, this chapter summarizes major biological and social psychological changes that adults with intellectual disabilities face as they age. These changes along with age-related changes of family caregivers affect their caregiving experience.

7.15.4 FAMILY CAREGIVING ACROSS THE LIFE-SPAN

7.15.4.1 Changes across the Life-span

7.15.4.1.1 Theories of family caregiving over the life-span

A life-span perspective on family caregiving recognizes that caregiving can span many decades and is usually a lifetime commitment for families of adults with intellectual disabilities. Unfortunately, the research examining caregiving across the life-span is sparse. Rather, most studies focus on particular ages, with the vast majority focusing on childhood. However, there have been studies focusing on older caregivers. In examining caregiving across the life-span, the terminology differs. Studies focusing on children typically refer to the "parenting" role; whereas studies focusing on adults refer to the "caregiving" role. A few studies have begun to examine caregiving reciprocity as adults with intellectual disabilities begin to serve caregiving roles for older parents (Heller & Factor, 1993a).

Over time, caregiving roles, stressors, and rewards are likely to have both stable and dynamic aspects. In examining the effects of caregiving over time, researchers need to take into account developmental transitions occurring for the person with a disability and the caregiver, changes in the availability of services and supports, and changes in societal attitudes towards the individual with a disability. Another important dimension could be the changing nature of societal expectations regarding caregiving and opportunities for fulfillment of other life roles. Several theories from the gerontological and intellectual disabilities literature address some of these dimensions in understanding the dynamic nature of caregiving over time.

Two contrasting theories emanating from the gerontological literature offer a framework for understanding family adaptation over the life-span. The "wear and tear" hypothesis predicts that the constant caregiving demands and accumulation of stressors will deplete the caregiver's physical and emotional resources (Johnson & Catalano, 1983). Several cross-sectional studies of caregivers of aging relatives have provided some support for this theory (e.g., Morycz, 1985; Zarit, Reever, & Bach-Peterson, 1980).

On the other hand, the "adaptational" hypothesis suggests that, with experience, caregiving becomes easier (Townsend, Noelker, Deimling, & Bass, 1989). The assumption is that when caregiving is a new role it is most stressful and that over time families learn to cope. Several longitudinal studies of caregiving for elderly relatives have reported less caregiving burden and depression over time despite deterioration in the condition of the elderly relative (Townsend et al., 1989; Zarit, Todd, & Zarit, 1986).

7.15.4.1.2 Findings on families of adults with intellectual disabilities

In the intellectual disabilities literature there has been support for both hypotheses, with reports of increases in both positive and negative adaptations over time and with much variability in family adaptations over time. The theories and research on intellectual disabilities also examine adaptation to caregiving from a developmental perspective that takes into account the normative life stage tasks faced by families at different life phases. Birenbaum (1971) and Farber (1975) have addressed the degree to which caregiving is normative vs. non-normative as the child with intellectual disabilities grows older. They suggest that as the child moves towards adulthood, caregiving becomes less normative, particularly if the child is still highly dependent on the parents. Birenbaum has noted that as the child gets older the rights and obligations of typical parenthood become more taxing and more divergent from normative roles. Similarly, Farber discusses the notion of "arrest in the life cycle," referring to the fact that as the child reaches adolescence and adulthood, the parents often are still involved in parenting tasks more typical in families with younger children. As children with intellectual disabilities become adults, most continue to live with parents, unlike what occurs in most families. Hence, parents of adults with intellectual disabilities are much less likely to experience an empty nest.

In addition to fulfilling roles that are less normative with age, parents of adults vs. children with intellectual disabilities are less likely to receive as much formal service support for their child. Suelzle and Keenan (1981) reported that families of adults were more stressed than families of children with intellectual disabilities. They attributed this greater stress to the lack of adult services. While there is US federal legislation mandating educational services for all children with disabilities, there are no mandates for educational, vocational, or residential services for adults with disabilities. Hence, in some families, the parents are the sole providers for their adult child with intellectual disabilities who may not be receiving any services or who may be on a waiting list for services. Furthermore, many of these adults lack friends and rely on their parents to meet their social recreational needs (Heller & Factor, 1988).

In discussing life-time accommodations to caregiving of a child with intellectual disabilities, Todd and Shearn (1996) discuss the meaning that parents give to the caregiving experience and its effect on their identity. They note that socialization into the parental role continues over the life course and entails a series of incremental changes and re-evaluations. They found different reactions of families depending on their interpretations of past experiences, anticipated futures, and personal aspirations. In their ethnographic study of 17 parents of adults with intellectual disabilities in Wales, Todd and Shearn describe how some of these older parents come to experience heightened stress as a consequence of their inability to end their parental responsibilities for their child with a disability. "In the beginning you know that you're going to be responsible for the rest of your life in one way or another, but you never really know in what way. It's like you just roll up your sleeves and get stuck in" (p. 8). Parents reported that adolescence was a turning point as their normalizing strategies were becoming less successful when the character of their own lives departed more from that of their peers. In the process of comparing themselves to others, caregivers (mothers in most cases) who had given up careers and other pursuits to care for their child with a disability, experienced greater discrepancy from the lives of their peers as their children got older.

On the other hand, there is also evidence supporting the adaptational hypothesis that, with time, caregiving for family members becomes less stressful and more rewarding. In a large study of 450 mothers aged 55 and over caring for an adult child with intellectual disabilities Seltzer and Krauss (1989) found, on average, favorable well-being. A higher proportion of these mothers rated their physical health as good or excellent (78%) than did a national probability sample of women their age (60%; Sweet & Bumpass, 1987) and they were less stressed than in other samples of caregivers of elderly relatives (e.g., Zarit et al., 1980) or relatives with intellectual disabilities (e.g., Friedrich, Greenberg, & Crnic, 1983).

In one of the few studies to examine family adjustment throughout the life-span Heller, Rowitz, and Farber (1992) compared 489 family caregivers grouped into five age phases based on the age of the person with intellectual disabilities: preschooler, young child, adolescent, young adult, and older adult. They found that perceived burden was lowest for caregivers of the youngest (under six years of age) and oldest (30 years and older) members with intellectual disabilities. They suggested that perhaps at this stage there is greater acceptance of the family member and greater reciprocity in caregiving as the child with intellectual disabilities takes on caregiving roles with aging parents. Seltzer and Ryff (1994) also suggest that for some mothers the experience of rasing a child with intellectual disabilities might lead to the development of new capacities and strengths that foster adaptational responses. Although the Heller et al. (1992) study was limited by its cross-sectional design and potential cohort effects, parents' retrospective reports further supported the notion of adaptational responses over the years. Families reported how over the years it became easier to adjust to the routines; they became more knowledgable about what was expected of them, and they understood their child better. Another frequent theme was that parents grew to love and appreciate their child more. These affectional bonds between the parent and child can nurture both the parent and the child (Seltzer & Ryff, 1994).

Todd and Shearn (1996) note that some parents make a career of parenting and redefine the centrality of parenting as the major line of activity open to them, beginning to interpret it more positively. Over time, parenting became a stronger source of self meaning and purpose.

While most of the research on adult children with intellectual disabilities has portrayed them as sources of burden to their primary caregiver, studies are beginning to document the positive contributions they can make to their families. Heller and Factor's (1988) study of 100 parents caring for adults with intellectual disabilities (aged 30 years and older) found that these children were an important source of support for them. In addition to providing meaning to their lives, many families reported that their child did chores for them, contributed financially through their Social Security Income, and kept them from feeling lonely. In a study of 85 family caregivers of adults (aged 30 years and over) that specifically addressed caregiving reciprocity, Heller and Factor (1993b) also found that adult children with intellectual disabilities provided considerable companionship support to their parents and help with household chores. Support from the child with intellectual disabilities was a key factor predicting parents' degree of perceived caregiving satisfaction and caregiving burden. Parents who received support from their adult child with intellectual disabilities experienced greater caregiving satisfaction and less caregiving burden. This reciprocal support may be particularly important as parents experience frailty due to aging or widowhood.

Both the "wear-and-tear" and "adaptational" hypotheses assume a linear relationship between time and caregiver well-being. Neither account for periodic variations in levels of stress as families experience major transitions, such as when the child enters or completes school or when the child moves out of the family home. Often the least stressful times are periods of stability in family roles (Wikler, 1986). These hypotheses also do not account for variability in characteristics of the child and in the family's coping and support resources.

7.15.4.1.3 Predictors of caregiver well-being in later life

Both characteristics of the child and of the family and coping and support resources can explain the wide variability in responses of families to caring for a child with intellectual disabilities over the lifespan. The major risk factors for stressful reactions in later life are having a child with more severe intellectual disabilities and more maladaptive behaviors (Heller & Factor, 1993b; Pruchno, Patrick, & Burant, in press).

The extent that formal services can reduce caregiver burden and increase families' ability to maintain caregiving in the home is unclear. However, in their two year longitudinal study of parents caring for adults with intellectual disabilities aged 30 years and over, increase in the number of perceived unmet formal needs was associated with increased caregiving burden and increased need for out-of-home placements (Heller & Factor, 1993b). However, greater use of services has not been associated with improved parental well-being. This chapter discusses the impact of family support services in greater detail in the section on "family support interventions and policies."

There is some evidence that informal support from family and friends can reduce burden (Heller & Factor, 1994; Seltzer & Krauss, 1989). In particular, a high level of sibling involvement with the adult child with retardation has been associated with positive maternal well-being. Seltzer, Begun, Seltzer, and Krauss (1991) found that when support was provided by the sibling to the adult child, mothers reported better health and life satisfaction than when mothers had no other children or received no such help from their other children. The critical variable was not the amount of support given to the mother, but the amount given to the adult child with a disability. They note that perhaps sibling involvement indicates potential continuity of family caregiving into the future.

7.15.4.2 Roles of Mother, Fathers, and Siblings

Much of the literature on family caregiving has focused on mothers, with little mention of the role of fathers or other siblings. Yet, each of the members' roles are interrelated and affect the ability of the family to cope with raising and caring for a child with intellectual disabilities.

7.15.4.2.1 Comparison of maternal and paternal caregiving

One of the most consistent findings in family research is that mothers are the primary caregivers across the lifespan (Hooyman & Gonyea, 1995). However, in the intellectual disabilities literature there have been only a few studies that compare the maternal and paternal roles and most of these studies have focused on families of young children. These studies have generally found that parental roles tend to be even more traditional in families that have a child with intellectual disabilities than in families that do not have a child with a disability. After the birth of a child with a disability, mothers are much less likely to return to work than are other mothers. In a study of two-parent families in which the mother worked prior to the birth of a child with a disability, Mardiros (1985) found that only 8% returned to work afterwards. Fathers of young children with a disability tend to assume less task involvement with the care of their child than do the other fathers. Willoughby and Glidden (1995) found that among parents of children with intellectual disabilities the only tasks in which fathers participated nearly equally with the mothers were those of play and helping with activities.

Several studies have found higher subjective caregiving burden among mothers than fathers of children (Beckman, 1991; Kazak, 1987) and of young adults with a disability (Hallum & Krumboltz, 1993), although some have found little differences in burden (Bruce & Schultz, 1994; Essex et al., 1992; Friedrich et al., 1983), or only in some dimensions of burden (e.g., pessimism) (Rousey, Best, & Blacher, 1992). In a cross-sectional study of 113 two-parent families of children and adults with intellectual disabilities, Heller, Hsieh, and Rowitz (in press) found that mothers spend more time in caregiving tasks for their child, offer more types of support, and perceive more caregiving burden than do fathers. These gender differences persist from childhood into adulthood and whether or not the child continues to reside in the family home.

In addition to the explanation that women have been socialized to play nurturant roles, another explanation for these gender differences in caregiving could be attributed to women having fewer competing roles and time demands, enabling more time to provide care. However, in a study of caregivers for elderly relatives (Stoller, 1983) employment has been shown to affect men's assistance to aging parents, but not women's care to parents. Among older fathers of children with intellectual disabilities, the father's retirement has not been related to increased involvement in the care of the child (Essex, Seltzer, & Krauss, 1992). Even when the mother is employed, men are not more likely to help with the household tasks or with care for the child with intellectual disabilities (Heller et al., in press; Marcenko & Meyers, 1991; Willoughby & Glidden, 1995). Also, women's employment does not seem to be related to the amount or type of help she provides her child with a disability. This could be one explanation for women reducing or giving up paid employment after the birth of a child with a disability. Both Heller et al. (in press) and Todd and Shearn (1996) have noted that many older mothers report that having a child with a disability hurt their job opportunities.

In examining predictors of maternal and paternal involvement and perceived burden, Heller et al. (in press) found that both mothers and fathers provided less support to the adult vs. the child with a disability and less to offspring who lived out of the home. Both mothers and fathers perceived less caregiving burden when children remained in the home and adult offspring with a disability were placed out of the home. Hence, placement out of the home seemed to be associated with reduced burden for parents of adults, but not for parents of children. One reason could be that launching out of the home is a normative function for adults, but not for children.

Given that mothers spend more time caring for the child having a disability than do the fathers, they appear to be more affected by the extent of the child's disability. Hence, in the Heller et al. (in press) study both the child's maladaptive behavior and physical health were related to higher perceived caregiving burden for the mother but not for the father. Also, mothers and fathers were differentially affected by their other role commitments and by the role commitments of their spouses. Mothers perceived the least burden when they were involved in organizational activities and when their husbands worked fewer hours. On the other hand, for fathers, neither their nor their wife's involvement in other roles had much effect on their perceived caregiving burden. However, given the correlational nature of the research it is also possible that fathers work more when their wives are more stressed either to provide better for the family or to avoid stressful family life (Farber, 1959).

Although family caregiving is predominantly a woman's task, fathers play a significant role in providing support. This support continues throughout the life-span, whether the offspring with a disability lives at home or out of the home. Often, fathers play an increasingly large role as their wife dies or becomes incapacitated, assuming the role of primary caregiver, previously held by the mother. Gordon, Seltzer, and Krauss (1996) found that upon the mother's death very few fathers sought out-of-home placement for their child with intellectual disabilities. These authors suggest that it is easier for fathers to move into the role of primary caregiver than to adjust to the empty nest after placement. Placement of their child also may remind fathers of their own mortality.

7.15.4.2.2 Sibling caregiver roles

Siblings provide the most durable relationships for adults with intellectual disabilities. They often share both common genes and the common experiences of growing up in the same family. Relationships among siblings without a disability are most typically egalitarian (Baskett & Johnson, 1982) and often intimate. The long-term impact of having a sibling with intellectual disabilities has been rarely studied.

There has been a large body of literature on the social and psychological effects of having a sibling with intellectual disabilities on children (Lobato, 1983). Generally, the research indicates that some siblings are at risk for greater emotional problems than other children. These siblings are more likely to have increased caregiving responsibilities, lower socioeconomic status, and anxiety or guilt surrounding the sibling with a disability (Begun, 1989). McHale and Gamble (1989) found that siblings of children with intellectual disabilities took on more household and caregiving responsibilities, but had similar well-being to that of siblings without a brother or sister with a disability. Other studies have noted the positive and valued aspects of having a sibling with a disability (Cleveland & Miller, 1977; Wilson, Blacher, & Baker, 1989). These studies have attributed the positive adjustment of siblings to parental attitudes and behaviors that fostered a family environment attentive to the needs of all members and that did not overly focus on the child with a disability.

There is some evidence that adult siblings of persons with intellectual disabilities continue to maintain involvement with their brother or sister with intellectual disabilities into adulthood and throughout the lifespan. Generally, these studies have found that these adult sibling relationships differ from those of other sibling dyads in that the relationships are less egalitarian and less intimate. Zetlin's (1986) study of sibling relationships of 35 adults with intellectual disabilities found that the relationship was hierarchical, with the disabled sibling being more emotionally dependent on the nondisabled sibling. Two larger studies of siblings of adolescents and adults (Heller, 1994; Seltzer et al., 1991) found that the nondisabled sibling performed nearly twice as many support functions for the other siblings than they received back from them. Heller (1994) also compared sibling involvement across ages. Among 175 siblings aged 11 years and older of siblings with moderate to profound intellectual disabilities, adults were more likely than adolescents to spend time helping their relative and to perform more support functions. This was despite the fact that adults were more likely than adolescents to be living out of the family home.

In examining the nature of supports provided by 326 siblings of adults with intellectual disabilities, Krauss (1993) found that these siblings were more likely to provide affective support (86%) than instrumental support (25%) to their sibling with a disability. The factors determining support functions provided differed for the two types of support. For affective support the key factor was past interpersonal climate of the family of origin, including greater cohesion and expressiveness; for instrumental support the key factor was the objective need for support, such as when the mother was older and widowed. The sibling was most likely to provide support when the sibling with a disability had less severe intellectual disabilities and the needs were seen as less demanding.

Several researchers have noted the positive but not intimate relationship between siblings with intellectual disabilities and their nondisabled siblings (Begun, 1989). Begun notes that these relationships were less intimate because it is more difficult to interact intimately with persons having limited language and social communication skills than it is to interact with nondisabled siblings. However, when the functioning level of the sibling with intellectual disabilities was higher the sibling relationship more closely resembled normative sibling relationships.

As parents can no longer care for their child with intellectual disabilities either due to frailty or death, other siblings begin to take on the role of primary caregiver. Griffiths and Unger (1994) reported that nearly 44% of 41 adult siblings interviewed were willing to take on future caregiving responsibilities if needed. Willingness to assume responsibility depended on the degree to which parents desired it and family communication was positive. In the Krauss, Seltzer, Gordon, and Friedman (1996) study of 140 adult siblings, 36% intended to coreside with their bother or sister with intellectual disabilities. These siblings were more likely to be sisters, as is typical of caregiving across disabilities (Hooyman & Gonyea, 1995). They also had siblings with less dependency needs and had engaged in more joint activities with their sibling. Hence, a sibling was more willing to take on the responsibility if the task did not seem too demanding. Furthermore, they were more likely to anticipate coresidence if their mother was already in poor health, making the transfer of responsibility more eminent.

Despite the fact that siblings are often expected to take on the primary caregiving role, often they are not consulted about future plans. For example, in the Krauss et al. (1996) study, only half of the siblings who expected to coreside with their sibling had discussions with their parents regarding future plans for their brother or sister. One explanation for families' reluctance to discuss these future plans is ambivalence about wanting another sibling to take on the responsibility. As a mother in one of our studies noted "I just do not want to burden our daughter." Another mother noted "My daughter has offered to care for her, but I know she works and it would be difficult for her." Hence, many siblings assume caregiving upon the death of their parents without sufficient planning and information about the needs and preferences of their siblings and the support services available to them.

In an examination of actual placements over a 3 year period following the death of parents, Gordon et al. (1996) found that in nearly one-quarter of the 21 families the adult with intellectual disabilities moved in with his or her sibling. They found that even when the adult with intellectual disabilities moved into an out-of-home placement, there was considerable sibling involvement, with many taking on the roles of guardian and advocate. In some cases they assumed these roles out of a sense of obligation to the deceased parents, and in others they did so out of personal desires to help their sibling.

7.15.4.3 Future Planning and Placement

As family caregivers age they increasingly face the prospect that their child with intellectual disabilities will outlive them. With increased frailty it may also be more difficult for parents to provide care. While planning for the future is a task faced by all persons as they age, the need to plan for their child is particularly salient and often stressful for these parents. Key aspects which usually need to be addressed include future living arrangements, guardianship, and financial planning. However, many of these older adults have great difficulty making future plans as do older adults in general (Heller & Factor, 1994).

7.15.4.3.1 Residential plans

Several studies have noted that only between one-third and one-half of aging parents make concrete plans for the their relative's future living arrangements (Freedman, Krauss, & Seltzer, 1997; Heller & Factor, 1994; Roberto, 1988; Smith, Fullmer, & Tobin, 1994). Residential planning is one of the most anxiety-provoking aspects of future planning as it reminds families of their mortality and it can invoke unspoken concerns regarding family responsibility, separation, and independence (Freedman et al., 1997). Furthermore, parents who have cared for their child at home for decades may be reluctant to relinquish their role. Older families often report that providing care for their relative within the family home provides meaning to their lives. The launching stage is a transition point which can result in considerable family accommodations, disequilibrium, and decreased well-being (Seltzer, Krauss, Choi, & Hong, in press).

While many families do not make concrete plans for living arrangements, most have preferences regarding the type of care they would like, and most would like to have plans in place in their lifetime (Heller & Factor, 1991; Roberto, 1988). About half of the families prefer continued care with another family member, usually a sibling of the adult with intellectual disabilities, and the other half would like an out-of-home placement. However, when families do begin to consider residential programs, they face long waiting lists (Hayden, 1992) and have significant concerns about the quality of housing options. A common sentiment of these families is that staff could not possibly provide the same degree of caring and understanding that a parent could. As one parent commented "It would break our heart to place her; there is no place like home and your parents; they know what their child is like" (Heller & Factor, 1988, p. 40). Wood and Skiles (1992) have noted that the caregiver's reluctance to plan may reflect both limited awareness of services and actual gaps in the service system. Apprehension about the residential service system may be particularly strong for older parents who may have had negative experiences when their children were young and when adequate community services were nonexistent.

Lack of planning by families can lead to emergency placements, difficulties for other family members who suddenly assume responsibility for their relative upon the death or frailty of a parent, pressures on the service system, and major disruptions in the daily life of the adult with intellectual disabilities. Also, without planning and discussions with other family members it is not possible to consider the views and preferences of various family members. Many parents do not discuss future plans with other family members or with their son or daughter with intellectual disabilities (Heller & Factor, 1994; Smith & Tobin, 1989). In a study that interviewed adults with intellectual disabilities (Heller & Factor, 1993a), we found that nearly half of the adults aged 30 years and older wanted to move out of the family home either to gain more independence or to live with friends. Those who wanted to remain with families stressed the desire for stability or the reciprocal supports they provided to and received from parents.

7.15.4.3.2 Predictors and outcomes of placement

Generally the parents who are most likely to make residential plans are those of higher socioeconomic status (Heller & Factor, 1991). Freedman et al. (1997) differentiate short-term (wanting a change in settings in less than two years) from long-term planners and note that the short-term planners are more likely to have health problems, worries about the future, and lower sense of purpose than longer-term planners or persons who had not made plans. The long-term planners had the most positive well-being of all the groups. A three-year follow-up indicated that the short-term planners were more likely to have achieved out-of-home placements.

The variables predicting preferences for future out-of-home placement versus actual placements differ. Heller and Factor's (1993b) longitudinal study indicated that over a two year period the major predictors of both outcomes were high perceived unmet need for formal services. High maladaptive behavior was associated with desire for placement, but not with actual placement. Actual placement was related to parents' perceptions that caring for their relative would become more difficult and to having placed their relative on a waiting list for residential services. In a study spanning three years, Seltzer, Greenberg, Krauss, and

Hong (1997) similarly found that placement was most likely to occur when the mother was in poor health and had placed her child on a waiting list. This pattern is similar to the factors predicting nursing home placement for elderly persons (Colerick & George, 1986; Pruchno, Michaels, & Potashnik, 1990) with characteristics of the caregiver playing a larger role than characteristics of the care recipient.

In examining outcomes of placement, Seltzer et al. (1997) found that parents' level of perceived caregiving burden was lower after placement. Although nearly 94% of the mothers maintained at least weekly contact with their child and at least one-third of the mothers continued providing them with some care, the mother's level of perceived burden was lower. With placement the daily caregiving responsibilities were reduced. In addition to reducing burden, out-of-home placement prior to the death of either or both parents can result in better outcomes for the adult with intellectual disabilities, in contrast with placements made after parental death. Gordon et al. (1996) found that the latter placements often result in more disruption in lifestyle, including not only a change in residence but also a change in day program and social life. These findings point to the importance of planning in advance of major parental declines in health.

7.15.4.3.3 Legal and financial planning

Legal and financial planning are other challenges for families who want to assure financial resources for their relative with intellectual disabilities without jeopardizing government benefits and who want to obtain guardianship if their relative is incompetent. In the Heller and Factor (1991) study nearly two-thirds of families of adults aged 30 years and over had made financial plans and nearly one-quarter had made guardianship plans. As with planning for living arrangements, the extent that families make plans is tied to their socioeconomic status. Certainly, families who are wealthier are more likely to make financial plans (Heller & Factor, 1991). However, lack of information and misinformation among families and professionals has often lead to inappropriate plans that do not protect government benefits or that deprive adults with intellectual disabilities of their rights. For example, some families will money or property to another child with the hope that the child with intellectual disabilities will be taken care of. There are many reasons why this approach is fraught with problems, including the high divorce rate and the encouragement of dependency of the adult child with a disability on another family member. Another example is the tendency of families to obtain full guardianship rather than partial guardianship where some of the rights of the adult with disability are retained. There has been very little research on advance directives: that is, plans regarding future health care needs.

7.15.4.4 Sociocultural Issues

The caregiving experience for families cannot be understood without taking into consideration their sociocultural context. We are just beginning to examine the commonalities and differences in the later-life experiences of families from different countries, ethnic backgrounds, racial heritage, and socioeconomic level.

7.15.4.4.1 Comparisons across countries

While this chapter primarily describes the service trends and research findings from the USA, many of the same trends are occurring in other countries, particularly in developed countries. For example in Germany, Adam (1992) has estimated that, from 1988 to 1998, the percentage of adults with intellectual disabilities receiving services in the Wilhlemshaven region will increase from 17% to 26% in the 50–64 year age group and from 3% to 4% in the over 60 years age group. The trend to close or phase down institutions and to develop community programs has been proceeding rapidly in Europe. For example, between 1980 and 1993 in Scandinavian countries and in the United Kingdom there has been a decrease in the institutional population at an average rate of 4.25% per year (Hatton, Emerson, & Kieman, 1995). As in the USA many countries have long waiting lists for residential services. Breitenbach (1995) has noted that in France there are few openings in residential programs for older adults who are living with families, as priority is given to persons already in the residential service system. Van Walleghem (1993) noted the long waiting lists awaiting older parents seeking residential services for their relative with intellectual disabilities in Belgium.

In a cross-national comparison Heller, Haveman, and Van Berkum (1996) report that in comparison with older caregivers in the USA, parents in the Netherlands are more likely to make future plans, have fewer unmet needs for supports, and perceive more caregiving burden. These results could be explained by differences in the two service systems. In the Netherlands the fact that families made more plans could be attributed to their greater use of case management services and less apprehension about the service system. The higher stipend given to persons with intellectual disabilities in the Netherlands than in the USA may explain the lower need for information about future financial planning. In both countries high unmet service needs and poor health of the adult with intellectual disabilities were related to higher caregiving burden.

In a study comparing older caregivers in the USA with caregivers in Ireland Seltzer et al. (1995) found that in Ireland caregivers were less likely to make future plans, had poorer health, a smaller support network, more parenting stress, and less life satisfaction. There were also differences between caregivers in the Republic of Ireland and in Northern Ireland, with caregivers from the Republic of Ireland reporting more social isolation. Both of these studies demonstrate the potential role that culture, society, and service availability and expectations play in shaping the caregiving experience for families.

7.15.4.4.2 Culturally diverse groups in the USA

Only in the 1990s have studies begun to examine caregiving issues among families of persons with intellectual disabilities from various minority groups. These studies have focused on residential placement and future planning, use of informal and formal supports, and perceived caregiving burden and satisfaction. Comparisons of African-American and white families have found that African-American families are less likely than other families to place their relative out of the family home (Markwardt, Heller, Rowitz, & Farber, 1993; Pruchno et al., in press). One reason could be attributed to difficulties minority families experience in accessing services, either due to lack of awareness or lack of available services in their communities (Vidal de Haymes, 1989). In a study of Hispanic families, Heller, Markwardt, Rowitz, and Farber (1994) found that in comparison with white non-Hispanic families these families reported significantly more problems with the service system, especially in finding information and finding support groups. Others have attributed low service participation to discrimination, lack of accessibility, and lack of bilingual services (Rosenthal & Carty, 1988).

Another explanation for the higher rate of coresidence with families for African-American adults with intellectual disabilities could be that African-American families tend to have more positive attitudes towards family responsibility and greater involvement with extended families. In one study of caregivers of adults with intellectual disabilities (Heller, Rowitz, & Farber, 1992) it was found that African-American families report a similar degree of caregiving burden to that of other families, but that they report less burden on the siblings of the adult with intellectual disabilities. On the other hand, Pruchno et al. (in press) found that African-American families of adults with intellectual disabilities or mental illness experienced less caregiving burden and greater caregiving satisfaction than did white caregivers, despite being poorer and in worse health.

Among African-American families, many express an interest in having a sibling take over responsibility when the parents can no longer provide care. In one focus group, many parents have stated that their other children have been more than willing to eventually take over the care. They noted that they did not need to ask the other siblings, as they came to them to offer their support: “My son has four sisters. All of them said that they would take over.”

It has been noted that many older family caregivers of adults with intellectual disabilities do not make plans for the future care of their relative. When asked about reasons for not making future plans, two major themes have emerged: low socioeconomic status and religiosity contributed to a present-day rather than future focus. Examples of comments by African-American caregivers include: “We are close to God. God will lead us”; “We have no money with which to plan. I live from day to day. I can’t afford big dreams; I just hope to stay healthy.”

Given the changing structure and composition of the American family, and in particular the African-American family, with more single-parent households and females in the workforce, the reliance on the extended family may not be sufficient. Families will need increased help in providing support to their relative with intellectual disabilities and more information on residential services and financial planning so that they can secure a desirable future for their relative. A critical aspect for many minority families will be outreach to families not only through the aging or disability networks, but also through churches and other community resources.

7.15.5 FAMILY SUPPORT INTERVENTIONS AND POLICIES

As adults with intellectual disabilities and their parents age, they likely face some functional declines and greater difficulty in maintaining independence. Yet this group of families tends to make low use of services. Most of these families have been quite self-reliant in providing care for their child over the years. This section describes their service needs and use, the importance of formal services in sustaining their ability to provide care, and different models of service provision.

7.15.5.1 Family Needs and Service Use

As their children make the transition from childhood to adulthood, parents of adults with intellectual disabilities face the unique challenge of dealing with their own age-related declines in health and energy and the decreased formal services available for adults with a disability. Older parents have major unmet needs for information regarding residential programs, financial plans, and guardianship (Heller & Factor, 1994; Kelly & Kropf, 1995). Also, they have major needs for case management and advocacy (Kelly & Kropf, 1995) and financial assistance (Fujiura & Roccoforte, 1990). Fujiura and Roccoforte found that families spent an average of 20% of their pretax annual income on unreimbursed expenses for their adult relative with intellectual disabilities residing with them. In addition to needing services for their child, parents may also need services for themselves through the aging network.

Families often are unsure of appropriate goals and programs for their adult children and they may have reservations about using either the aging or intellectual disabilities formal services system. Many older caregivers have not had formal services available in the past and express fear that placement is the main aim of the formal service system (Janicki, 1996). Todd and Shearn (1996) note that for some families who have made a career of parenting, use of services reduces their roles and reduces their sense of purpose and identity. Other reasons why some families are reluctant to accept agency services are: (i) previous negative experiences, (ii) difficulties talking with providers who offer advice families do not agree with or who are not sensitive to their concerns, (iii) complex eligibility criteria for some services, and (iv) inflexibility of services. On the other hand, some families may have been long-term users of services. They may have been active in developing parent organizations, support groups, or provider agencies. The service families are most likely to use for their child is day programming. While many families desire respite, frequently the public financial resources for respite care are insufficient to meet demand.

The extent to which formal services improve the well-being of the caregiver is unclear. In our studies (Heller & Factor, 1991, 1993b) perceived unmet formal service needs was a key factor influencing the well-being of parents, with high unmet needs being related to caregiving burden. Use of services also has been related to parents' perceived current and future ability to provide care (Englehardt, Brubaker, & Lutzer, 1988).

7.15.5.2 Progam Models

Several models of support interventions help families in sustaining care and in coping with their relative. These include: (i) support and psycho-educational groups, (ii) outreach through the aging network, and (iii) family support programs.

7.15.5.2.1 Support and psycho-educational groups

At least one-quarter of the older family caregivers interviewed (Heller & Factor, 1991) expressed an unmet need for a support group. Support groups are one type of intervention that can provide aging families with information, skills, and support for helping the adult with intellectual disabilities, for helping themselves cope with the emotional aspects of caregiving, and for developing future plans. Smith, Majeski, and McClenny (1996) report that older families who participated in psycho-educational groups geared to caregivers of adults with intellectual disabilities most valued the information on future planning and formal services and the opportunity to interact with other parents having similar concerns. Parents who did not coreside with their adult child having intellectual disabilities reported benefitting the most from information on quality of residential programs. Although Smith et al. (1996) report family satisfaction with psycho-educational groups, few older family caregivers attend such events. Even when they would like to participate, older family members may have difficulty with transportation or scheduling. While these short-term interventions may ease stress on a short-term basis or provide some needed information, they do not tend to have long-lasting impacts. Also, Abel and Nelson (1990) caution that educational programs for caregivers can undermine the caregiver's own problem-solving capacity and sense of control over the situation.

Support groups for caregivers of elderly relatives have had modest benefits as measured by rating of participants' well-being and caregiving burden (Zarit & Toseland, 1989). Among families of persons with mental illness, they have been shown to help caregivers gain a sense of self-control and lower burden in some cases (Anthony & Blanch, 1989). No research has examined the effectiveness of support groups that are not psycho-educational groups on older caregivers of adults with intellectual disabilities.

7.15.5.2.2 Outreach through the aging network

Since US federal legislation was passed that targeted collaborative projects between the aging and developmental disabilities networks, there have been several model projects that targeted elderly caregivers of adults with intellectual disabilities. These projects entail coordination, network building, and internetwork coalition building (Factor, 1996; Janicki, 1996) between not only the aging and disabilities service agencies, but also other community agencies. One example of such a project is the Hispanic Coalition Network (Factor, 1996) that was developed in Chicago, which organized an Hispanic Fair that brought in 120 providers, elderly Hispanic family caregivers, and other community resources in a local community. In order to attract the families, health screenings and food were offered. Many families who had not previously received needed services became aware and linked up with a variety of services. Another example of a caregiver project is the New York Caregiver Assistance Project (Janicki, 1996), mentioned earlier in this chapter. That project operated from a mode which assigned outreach and assistance functions to a host agency. Major components included a needs assessment, targeting of special households, and a program of promoting community awareness. The project found that area agencies on aging could provide a safety valve for potential crises that would have resulted in immediate placement demands and referral to potentially more costly developmental disability programs. Another major finding was that working with cultural organizations within the local community resulted in greater willingness of families to accept services.

7.15.5.2.3 Family support public programs

In the 1980s and 1990s there has been a growing expansion of family support programs for families of persons with intellectual disabilities in the USA. Most of the earlier programs provided discrete types of services which were directed by agencies and by professionals. These included primarily respite, counseling, and parent training services. Later, cash subsidies were added in some states. However, with the emerging movement towards community empowerment and recognition of the need for persons with disabilities to remain in their communities and "natural" settings, new family-driven models have been promulgated. Generally, family support programs have been based on the goals of strengthening family functioning and enhancing growth and development of individuals and the family. Ultimately, most of these programs have the goals of maintaining persons with intellectual disabilities within the family home. As of 1995, 21 states had family support legislation (Bergman & Singer, 1996). Between 1988 and 1992 there has been an increase in intellectual disabilities state expenditures for family support services from 171 million to 277 million dollars allocated and from 163 330 to 198 827 families served (Braddock et al., 1995). While there has been an expansion of family support services, funding for them only represents 2% of expenditure for intellectual disabilities services (Braddock et al.) and many of these programs only target children.

Most of the innovative family support programs have developed models of consumer control in which families determine how resources are spent. These are based on the premise that families usually know best what it takes to maintain their relative with a disability in the family home. They also acknowledge the importance of using the informal network, including relatives, neighbors, and friends by allowing families to pay them for support services.

Most of the studies evaluating the impact of these family support programs have focused on families of children. These studies have documented their success in reducing stress and increasing life satisfaction for caregivers (Herman, 1991; Meyers & Marcenko, 1989). In a study of a family support program in Illinois geared to adults, it was found that family subsidies of about $1 300 combined with service facilitation and consumer control of services and expenditures resulted in many positive benefits over a 4 year period (Heller, Ruch-Ross, & Kopnick, 1995). In comparison with families who did not receive these services through a lottery selection process, the participant families reported greater caregiving satisfaction and greater feelings of competence and self-efficacy in helping their relative with intellectual disabilities. In addition, benefits included more community integration and improved interpersonal relations of the adult with intellectual disabilities, fewer unmet services needs, more satisfaction with services received, and less need for out-of-home placement. Also, families who participated were more likely to engage in legal and financial planning for their relative.

With the trend to community inclusion and consumer control of resources, and with pressures to cut costs by reducing residential services, there is likely to be an increase in family support programs. For older family caregivers this can result in added resources for maintaining caregiving within the family. However, with the advent of managed care approaches there is the danger that the funds will be too small or that sufficient residential options will not be in place for those families not able to provide care in the home setting. The Illinois family support program provides twice more funds per family than most of the other states' programs and is one of the few that targets adults.

7.15.6 RESEARCH DIRECTIONS

7.15.6.1 Summary of Research

The research on family caregiving for adults with intellectual disabilities is indicating tremendous variability in families' responses and the importance of understanding the life course and social context within which families operate. Some families devoted much of their lives to parenting a child with a disability through years when there were few services. Other families placed their relative out of the home and had various levels of continued involvement with their relative. With the advent of parent advocacy programs some families became active in promulgating new and expanded services in their communities. Certainly, during these eras there have been major demographic trends that have affected families' ability to provide care.

While expectations of family caregiving are higher, the resources within the family have often been diminished. This has been most acutely felt by women who most often are the primary caregivers. Yet, other members, such as fathers and siblings, also play crucial roles in supporting the adult with intellectual disabilities, not only during the parents' lifetime, but also afterwards.

Support programs and interventions have been one resource helping family caregivers. These have included support groups, counseling, parent training, respite, and cash subsidies. The effectiveness of these programs in helping families care for their relative and in enhancing the lives of the family members is being studied. Little is known about the best methods for providing supports to families. However, it is clear that, as they age, many families who have relied on their own resources will require some help from the formal service systems.

7.15.6.2 Research Agenda

To improve theoretical understanding of the caregiving experience and of ways to promote positive coping among families of adults with intellectual disabilities, comparative research across disability groups, across developmental periods, and across culturally diverse groups is needed. Most of the research up to 1997 focuses on one type of disability at one developmental period, usually that of parenting young children or of caregiving in later life. There are only a handful of studies that span different age phases, and these studies are hampered by their cross-sectional methodology.

With the exception of the research conducted by Seltzer and Krauss and their colleagues and by Heller and Factor there have been no longitudinal studies examining family caregiving among older family caregivers of adults with intellectual disabilities. With more longitudinal data it will be easier to delineate causal directions.

Another aspect that needs to be pursued further is the reciprocal nature of family caregiving and the impact of reciprocal supports on various family members. This line of research needs to take into account the perspectives of various family members—including the adult with intellectual disabilities, who may view the situation differently. Fathers and siblings also have been largely neglected in research; yet they play a critical role in supporting the adult with a disability and the mother.

There is little data on the impact of cultural, national and local differences on families' caregiving experience. There is a need for more studies that include various cultural groups and that examine the impact of differences in cultures and service networks in different communities, states, and countries. The challenge in this type of research is to disentangle the effects of socioeconomic status, culture, and service network differences on families.

Finally, more needs to be known about ways to help families cope. One way is to learn from families who have coped successfully, have nurtured a positive family climate, and have enhanced the development of the member with intellectual disabilities. These families have been able to successfully negotiate both the fulfillment of their own roles and the development of the adult with disabilities. In addition to quantitative data, there is a need for qualitative, ethnographic data to more richly elucidate this coping process .

Another way to learn more about ways to help families is to study the effectiveness of interventions and public programs. Most of the research is anecdotal or lacks empirically-sound comparison groups. In studying these interventions, a longer-term perspective needs to be taken. For example, to understand the impact on actual placement, a span of more than just 1 year must be looked at. As service models evolve examinations of both the mode in which services are delivered and the outcomes of these approaches are needed. Satisfaction of adults with disabilities and their families with the process and outcomes should be a major determinant of the program's success.

ACKNOWLEDGMENT

The preparation of this chapter was supported in part by funds provided by the National Institute on Disability and Rehabilitation Research and Training Center on Aging with Mental Retardation, Grant No. H133B30069.

7.15.7 REFERENCES

Abel, E. K., & Nelson, M. K. (1990). *Circles of care: Work and identity in women's lives* (pp. 65–91). Albany, NY: State University of New York Press.

Adam, G. A. (1992). German perspective. In S. Moss (Ed.), *Aging and developmental disabilities: Perspectives from nine countries* (pp. 37–44). Durham, NH: World Rehabilitation Fund University of New Hampshire.

Adlin, M. (1993). Health care issues. In E. Sutton, A. Factor, B. Hawkins, T. Heller, & G. Seltzer (Eds.), *Older adults with developmental disabilities: Optimizing choice and change* (pp. 49–60). Baltimore: Brookes.

Anthony, W., & Blanch, A. (1989). Research on community support services: What have we learned? *Psychosocial Rehabilitation Journal, 12,* 55–81.

Baskett, L. M., & Johnson, S. M. (1982). The young child's interactions with parents versus siblings: A behavioral analysis. *Child Development, 53,* 643–650.

Beckman, P. J. (1991). Comparison of mothers' and fathers' perceptions of the effect of young children with and without disabilities. *American Journal on Mental Retardation, 95,* 585–595.

Begun, A. L. (1989). Sibling relationships involving developmentally disabled people. *American Journal of Mental Retardation, 93,* 566–574.

Bergman, A., & Singer, G. H. S. (1996). The thinking behind new public policy. In G. H. S. Singer, L. Powers, & A. L. Olson (Eds.), *Redefining family support: Innovations in public–private partnerships* (pp. 435–464). Baltimore: Brookes.

Birenbaum, (1971). The mentally retarded in the home and the family cycle. *Journal of Health and Human Behavior, 12,* 55–65.

Blaney, B. (1995). *Aging with mental retardation.* Presentation at the Presidents Committee on Mental Retardation Collaborative Academy, Washington, DC.

Braddock, D. (1991). Technical note to census data. Chicago: Illinois University Affiliated Program, University of Illinois at Chicago.

Braddock, D., Hemp, R., Bachelder, L., & Fujiura, G. (1995). *The state of the states in developmental disabilities* (4th ed.*)* Washington, DC: American Association on Mental Retardation.

Bradley, V., & Knoll, J. (1990). *Shifting paradigms in services to people with developmental disabilities.* Cambridge, MA: Human Services Research Institute.

Breitenbach, N. (1995). *The lost generation.* Paris: Fondation De France.

Bruce, E. J., & Schultz, C. L. (1994). A cross-sectional study of parenting perceptions: Caring for children with intellectual disability. Special Issue: Family caregiving. *Australian Journal of Marriage and Family, 15,* 56–65.

Carter, G., & Jancar, J. (1983). Mortality in the mentally handicapped: A 50 year study at the Stokes Park group of hospitals. *Journal of Mental Deficiency Research, 27,* 143–156.

Center for Vulnerable Populations (1993). *Families focus: The status of America's vulnerable populations: A chartbook.* Waltham, MA: Brandeis University Press.

Cleveland, D.W., & Miller, N. (1977). Attitudes and life commitments of older siblings of mentally retarded adults: An exploratory study. *Mental Retardation, 15,* 38–41.

Colerick, E. J., & George, L. K. (1986). Predictors of institutionalization among caregivers of patients with Alzheimer's disease. *Journal of American Geriatrics Society, 34,* 493–498.

Englehardt, J., Brubacker, T. H., & Lutzer, V. (1988). Older caregivers of adults with mental retardation: Service utilization. *Mental Retardation, 26,* 191–195.

Essex, E. L., Seltzer, M. M., & Krauss, M. W. (1992). *Aging fathers as caregivers for adult children with developmental disabilities.* Paper presented at the National Institute on Aging Seminar on "Men's Caregiving Roles in an Aging Society," Washington, DC.

Eyman, R. K., & Borthwick-Duffy, S. A. (1994). Trends in mortality rates and predictors of mortality. In M. M. Seltzer, M. W. Krauss, & M. Janicki (Eds.), *Life-course perspectives on adulthood and old age* (pp. 93–105). Washington, DC: American Association on Mental Retardation Monograph Series.

Eyman, R. K., & Widaman, K. F. (1987). Life-span development of institutionalized and community-based mentally retarded persons, revisited. *American Journal of Mental Deficiency, 91,* 559–569.

Factor, A. R. (1996). *Final Report: Innovative internetwork service models serving older adults with developmental disablities and older family caregivers.* Chicago: Institute on Disability and Human Development, University of Illinois at Chicago.

Farber, B. (1959). Effects of a severely mentally retarded child on family integration. *Monographs of the Society for Research in Child Development, 24*(2, Serial No. 71).

Farber, B. (1975). Family adaptations to severely mentally retarded children. In M. J. Begab & S. A. Richardson (Eds.), *The mentally retarded and society.* Baltimore: University Park Press.

Fisher, M., & Zeaman, D. (1970). Growth and decline of retardate intelligence. In N. R. Ellis (Ed.), *International Review of Research in Mental Retardation* (Vol. 4, pp. 151–191). New York: Academic Press.

Freedman, R. I., Krauss, M. W., & Seltzer, M. M. (1997). Aging parents' residential plans for adult children with mental retardation. *Mental Retardation, 35,* 114–123.

Friedrich W. N., Greenberg, M. T., & Crnic, K. (1983). A short-form of the questionnaire on resources and stress. *American Journal of Mental Deficiency, 88,* 41–48.

Fujiura, G. (1996). *The demography of family care for adults with developmental disabilities.* Chicago: Institute on Disability and Human Development, University of Illinois at Chicago.

Fujiura, G., & Roccoforte, J.R. (1990, May). *Cost of disability to the American family.* Public Policy Forum, American Association on Mental Retardation, Washington, DC.

Gordon, R. M., Seltzer, M. M., & Krauss, M. W. (1996). The aftermath of parental death: Changes in the context and quality of life. In R. L. Schalock (Ed.), *Quality of life: Its applications to persons with disabilities* (pp. 23–40). Washington, DC: American Association on Mental Retardation.

Griffiths, D. L., & Unger, D. G. (1994). Views about planning for the future among parents and siblings of adults with mental retardation. *Family Relations, 43,* 221–227.

Hallum, A., & Krumboltz, J. D. (1993). Parents caring for young adults with severe physical disabilities: Psychological issues. *Developmental Medicine and Child Neurology, 35,* 24–32.

Hatton, C., Emerson, E., & Kieman, C. (1995). Trends and milestones. *Mental Retardation, 33,* 132.

Hayden, M. F. (1992). Adults with mental retardation and other developmental disabilities waiting for community-based services in the United States. *Policy Research Brief, 4,* 1–18.

Heller. T. (1994). *Contemporary research on sibling issues.* Paper presented at the Brother and Sister: A Lifelong Connection conference of the New York Association for the Help of Retarded Children, New York.

Heller, T. (1997). Older adults with mental retardation and their families. In N. Bray (Ed.), *International review of research in mental retardation* (Vol. 20, pp. 99–136). San Diego, CA: Academic Press.

Heller, T., & Factor, A. (1988). *Development of a transition plan for older adults with developmental disabilities residing in the natural home.* Springfield, IL: Governor's Planning Council on Developmental Disabilities.

Heller, T., & Factor, A. (1991). Permanency planning for adults with mental retardation living with family caregivers. *American Journal of Mental Retardation, 96,* 163–176.

Heller, T., & Factor, A. (1993a). Support systems, well-being, and placement decision-making among older parents and their adult children with developmental disabilities. In E. Sutton, A. Factor, B. Hawkins, T. Heller, & G. Seltzer (Eds.), *Older adults with developmental disabilities: Optimizing choice and change* (pp. 107–122). Baltimore: Brookes.

Heller, T., & Factor, A. (1993b). Aging family caregivers: Changes in burden and placement desire. *American Journal on Mental Retardation, 98,* 417–426.

Heller, T., & Factor, A. (1994). Facilitating future planning and transitions out of the home. In M. M. Seltzer, M. W. Krauss, & M. Janicki (Eds.), *Life-course perspectives on adulthood and old age* (pp. 39–50). Washington, DC: American Association on Mental Retardation Monograph Series.

Heller, T., Sterns, H., Sutton, E., & Factor, A. (1996). Impact of person-centered later life planning training program. *Journal of Rehabilitation, 62,* 77–83.

Heller, T., Haveman, M., & Van Berkum, G. H.(1996). *Subjective and objective caregiving intensity and future planning among aging parents of adults with mental retardation: Comparison of United States and Dutch findings.* Annual Meeting of the International Association for the Scientific Study of Intellectual Disability, Helsinki, Finland.

Heller, T., Hsieh, K., & Rowitz, R. (in press). Maternal and paternal caregiving of persons with mental retardation across the lifespan. *Family Relations.*

Heller, T., Markwardt, R., Rowitz, L, & Farber, B. (1994). Adaptations of Hispanic families to a member with mental retardation. *American Journal on Mental Retardation, 99,* 289–300.

Heller, T., Rowitz, L., & Farber, B. (1992). *The domestic cycle of families of persons with mental retardation.* Chicago: University of Illinois Affiliated Program in Developmental Disabilities and School of Public Health, University of Illinois at Chicago.

Heller, T., Ruch-Ross, H., & Kopnick, N. M. (1995). *The Illinois Home-Based Support Services Program evaluation report.* Chicago: Institute on Disability and Human Development, University of Illinois at Chicago.

Herman, S. (1991). Use and impact of a cash subsidy program. *Mental Retardation, 29,* 253–258.

Hogg, J., & Moss, S. (1993). Characteristics of older people with intellectual disabilities in England. In N. Bray (Ed.), *International review of research in mental retardation* (Vol. 19). New York: Academic Press.

Hooyman, N. R., & Gonyea, J. (1995). *Feminist perspectives on family care: Policies for gender justice.* Thousand Oaks, CA: Sage.

Janicki, M. P. (1996). *Help for caring—For older people caring for adults with a developmental disability.* Albany, NY: New York State Developmental Disablities Planning Council.

Janicki, M., Heller, T., Seltzer, G., & Hogg, J. (1996). Practice guidelines for the clinical assessment and care management of Alzheimer's disease and other dementias among adults with intellectual disability. *Journal of Intellectual Disability Research, 40,* 374–382.

Johnson, C., & Catalano, D. (1983). A longitudinal study of family supports to impaired elderly. *The Gerontologist, 23,* 612–618.

Kazak, A. E. (1987). Families with disabled children: Stress and social networks in three samples. *Journal of Abnormal Psychology, 15,* 137–146.

Kelly, T. B., & Kropf, N. P. (in press). Stigmatized and perpetual parents: Older parents caring for adult children with lifelong disabilities. *Journal of Gerontological Social Work, 23.*

Krauss, M. W. (1986). *Long-term care issues in mental retardation.* Paper presented at the National Institute of Child Health and Human Development and Kennedy Foundation Conference, Mental Retardation: Accomplishments and New Frontiers, Bethesda, MD.

Krauss, M. W. (1993). *Role variation among adult siblings of persons with mental retardation.* Paper presented at the Annual Meeting of the American Association on Mental Retardation. Washington, DC.

Krauss, M. W., Seltzer, M. M., Gordon, R., & Friedman, D. H. (1996). Binding ties: The roles of adult siblings of persons with mental retardation. *Mental Retardation, 34,* 83–93.

Lobato, D. (1983). Siblings of handicapped children: A review. *Journal of Autism and Developmental Disorders, 13,* 347–364.

Luckasson, R., Coulter, D. L., Polloway, E. A., Reiss, S., Schalock, R. L., Snell, M. E., Spitalnik, D. M., & Stark, J. A. (1993). *Mental retardation: Definition, classification, and systems of supports.* Washington, DC: American Association on Mental Retardation.

Marcenko, M. O., & Meyers, J. C. (1991). Mothers of children with developmental disabilities: Who shares the burden? *Family Relations, 40,* 186–190.

Mardiros, M. (1985). Role alterations of female parents having children with disabilities. *Canada's Mental Health, 33*(4), 24–26.

Markwardt, R., Heller, T., Rowitz, L., & Farber, B. (1993). *Adaptation of African-American families to a member with mental retardation.* Paper presented at the Annual Meeting of the American Association on Mental Retardation. Washington, DC.

McClaren, J., & Bryson, S. E. (1987). Review of recent epidemiological studies of mental retardation: Prevalence, associated disorders, and etiology. *American Journal of Mental Retardation, 92,* 243–254.

McHale. S. M., & Gamble, W. C. (1989). Sibling relationships of children with disabilities and nondisabled brothers and sister. *Developmental Psychology, 25,* 412–429.

Meyers, C. E., Borthwick, S. A., & Eyman, R. (1985). Place of residence by age, ethnicity, and level of retardation of the mentally retarded/developmental disabled population of California. *American Journal of Mental Deficiency, 90,* 266–270.

Meyers, J. C., & Marcenko, M. O. (1989). Impact of a cash subsidy program for families of children with severe developmental disabilities. *Mental Retardation, 27,* 383–386.

Morycz, R. K. (1985). Caregiving strain and the desire to institutionalize family members with Alzheimer's disease. *Research on Aging, 7,* 329–361.

Patel, P., Goldberg, D. P., & Moss, S. C. (1993). Psychiatric morbidity in older people with moderate and severe learning disability (mental retardation), Part II: The prevalence study. *British Journal of Psychiatry, 163,* 481–491.

Pruchno, R. A., Michaels, J. E., & Potashnik, S. L. (1990). Predictors of institutionalization among Alzheimer disease victims with caregiving spouses. *Journal of Gerontology: Social Sciences, 45,* S259–S266.

Pruchno, R. A., Patrick, J. H., & Burant, C. J. (in press). African-American and white mothers of adults with chronic disabilities: Caregiving burden and satisfaction. *Family Relations.*

Roberto, K. (1988). *Caring for aging developmentally disabled adults: Perspectives and needs of older parents.* (Final report presented to the Colorado Developmental Disabilities Planning Council). Greeley, CO: Department of Human Services, University of North Colorado.

Roberts, R.S. (1990). *Use of quality of life scales to measure outcomes of a project using volunteers with elderly* persons with mental retardation. Paper presented at the Growing Up and Growing Old symposium, Institute for Research on Mental Retardation and Brain Aging, Troina, Italy.

Rosenthal, E., & Carty, L. A. (1988). *Impediments to services and advocacy for black and Hispanic people with mental illness.* Mental Health Law Project, National Institute on Mental Health.

Rossi, A. (1987). Parenthood in transition: From lineage to child to self-orientation. In J. Lancaster, J. Altmann, A. Rossi, & L. Sherrod (Eds.), *Parenting across the life span: Biosocial dimensions.* New York: de Gruyter.

Rousey, A. M., Best, S., & Blacher, J. (1992). Mothers' and fathers' perceptions of stress and coping with children who have severe disabilities. *American Journal on Mental Retardation 97,* 99–110.

Seltzer, G., Begun, A. L., Seltzer, M. M., & Krauss, M. W. (1991). Adults with mental retardation and their aging mothers: Impact of siblings. *Family Relations, 40,* 310–317.

Seltzer, M. M., Greenberg, J. S., Krauss, M. W., & Hong, J. (1997). The end of co-resident caregiving: Predictive factors and changes in well-being in aging mothers of adults with disabilities. *Family Relations, 46,* 13–22

Seltzer, M. M., & Krauss, M. W. (1989). Aging parents with mentally retarded children: Family risk factors and sources of support. *American Journal on Mental Retardation, 94,* 303–312.

Seltzer, M. M., Krauss, M. W., Choi, S. C., & Hong, J. (in press). Midlife and later life parenting of adult children with mental retardtion. In C. D. Ryff & M. M. Seltzer (Eds.), *The parental experience at midlife.* Chicago: University of Chicago Press.

Seltzer, M. M., Krauss, M. W., Walsh, P., Conliffe, C., Larson, B., Birkbeck, G., Hong, J., & Choi, S.C. (1995). Cross-national comparisons of aging mothers of adults with intellectual disabilities. *Journal of Intellectual Disability Research, 39,* 408–418.

Seltzer, M. M., & Ryff, C. D. (1994). Parenting across the lifespan: The normative and nonnormative cases. In D. L. Featherman, R. Lerner, & M. Perlmutter (Eds.), *Life-span development and behavior* (Vol. 12, pp. 1–39). Hillsdale, NJ: Erlbaum.

Smith, G. S., Fullmer, E. M., & Tobin, S. S. (1994). Living outside the system: An exploration of older families who do not use day programs. In M. M. Seltzer, M. W. Krauss, & M. P. Janicki (Eds.), *Life-course perspectives on adulthood and old age* (pp. 19–37). Washington, DC: American Association on Mental Retardation Monograph Series.

Smith, G., Majeski, R., & McClenny, B. (1996). Psychoeducational support groups for aging parents: Development and preliminary outcomes. *Mental Retardation, 34,* 172–181.

Smith, G., & Tobin, S. (1989). Permanency planning among older parents of adults with lifelong disabilities. *Journal of Gerontological Social Work, 14,* 35–59.

Stoller, E. P. (1983). Parental caregiving by adult children. *Journal of Marriage and the Family, 45,* 851–858.

Suelzle, M., & Keenan, V. (1981). Changes in family support networks over the life of mentally retarded persons. *American Journal of Mental Deficiency, 86,* 267–274.

Sweet, J. A., & Bumpass, L. (1987). *American families and households.* New York: Russell Sage Foundation.

Todd, S., & Shearn, J. (1996). Time and the person: The impact of support services on the lives of parents of adults with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities, 9,* 40–60.

Townsend, A., Noelker, L., Deimling, G., & Bass, D. (1989). Longitudinal impact of interhousehold caregiving on adult children's mental health. *Psychology and Aging, 4,* 393–401.

Turk, M. A., Overeynder, J. C., & Janicki, M. P. (1995). *Uncertain future: Aging and cerebral palsy: Clinical concerns.* New York: New York State Developmental Disabilities Planning Council.

US. Bureau of the Census (1992a). *Households, families, and children: A 30-year perspective* (P-23, No. 181). Washington, DC: Government Printing Office.

US Bureau of the Census (1992b). How we're changing—demographic state of the nation: 1992. *Current Population Reports* (Special Studies, Series P-23, No. 177). Washington, DC: Government Printing Office.

US Bureau of the Census (1992c). Marital status and living arrangements: March 1992. *Current Population Reports* (Series P-20, No. 468). Washington, DC: Government Printing Office.

US Bureau of the Census (1992d). Marriage, divorce and remarriage in the 1990s. *Current Population Reports* (Series P-23, No. 180). Washington, DC: Government Printing Office.

US Bureau of the Census (1992e). Population projections on the United States, by age, sex, race and Hispanic origin: 1992 to 2050. *Current Population Reports* (Series P-25, No. 1092). Washington, DC: Government Printing Office.

US Bureau of the Census (1992f). Sixty-five plus in America. *Current Population Reports* (Series P-23, No. 178). Washington, DC: Government Printing Office.

Van Walleghem, M. (1993). Families with an elderly mentally handicapped child: Results of a qualitative study in Flanders Belgium. In M. K. Haveman & W. H. E. Buntinx (Eds.), *Family needs and family support in mental retardtion: An international perspective* (pp. 73–80). Nijmegen, The Netherlands: Kavanah Press.

Vidal de Haymes, M. (1989). *The Illinois Hispanic community living plan: Final report.* Springfield, IL: Illinois Hispanic Community Living Plan Advisory Board, Illinois Planning Council on Developmental Disabilities.

Walz, T., Harper, D., & Wilson, J. (1986). The aging developmentally disabled person: A review. *The Gerontologist, 26,* 622–629.

Wikler, L. M. (1986). Family stress theory and research on families of children with mental retardation. In J. Gallagher & P. Vietze (Eds.), *Families of handicapped persons: Research programs and policy issues* (pp. 167–196). Baltimore: Brookes.

Wilkin, D., Durie, A., Wade, G., Jolley, D., & Stout, I. (1984). *Specialist services for the elderly: A study of* *referrals to geriatric, psychiatric, and social services* (Research Rep. No. 5). Manchester, UK: University of Manchester.

Willoughby, J. C., & Glidden, L. M. (1995). Fathers helping out: Shared child care and marital satisfaction of parents of children with disabilities. *American Journal on Mental Retardation, 99,* 399–406.

Wilson, J., Blacher, J., & Baker, B. L. (1989). Siblings of children with severe handicap. *Mental Retardation, 27,* 167–173.

Wood, J. B., & Skiles, L. L. (1992). Planning for the transfer of care. *Generations, 16*(1), 61–62.

Zarit, S. H., Reever, K. E., & Bach-Peterson, J. (1980). Relatives of the impaired elderly: Correlates of feelings of burden. *The Gerontologist, 20,* 649–655.

Zarit, S. H., Todd, P. A., & Zarit, J. M. (1986). Subjective burden of husbands and wives as caregivers: A longitudinal study. *The Gerontologist, 26,* 260–266.

Zarit, S. H., & Toseland, R. W. (1989). Current and future directions in family caregiving research. *The Gerontologist, 29,* 481–483.

Zetlin, A. (1986). Mentally retarded adults and their siblings. *American Journal of Mental Deficiency, 91,* 217–225.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.16
Falls Among Older Adults

BARRY A. EDELSTEIN and LISA W. DROZDICK
West Virginia University, Morgantown, WV, USA

7.16.1 INTRODUCTION 350

7.16.2 CONCEPTUAL AND METHODOLOGICAL ISSUES 350

7.16.3 CONSEQUENCES OF FALLS 351

- *7.16.3.1 Physical Consequences* 351
- *7.16.3.2 Psychological Consequences* 352
- *7.16.3.3 Economic Consequences* 353

7.16.4 RISK FACTORS AND THE CONTEXT OF FALLS 354

- *7.16.4.1 Extrinsic Factors* 354
  - *7.16.4.1.1 Time of fall* 354
  - *7.16.4.1.2 Location of fall* 355
  - *7.16.4.1.3 Walking aids* 355
  - *7.16.4.1.4 Restraints* 355
  - *7.16.4.1.5 Footwear* 356
  - *7.16.4.1.6 Environmental hazards* 356
- *7.16.4.2 Intrinsic Risk Factors* 356
- *7.16.4.3 Demographic Factors* 356
  - *7.16.4.3.1 Gender* 356
  - *7.16.4.3.2 Age* 357
  - *7.16.4.3.3 Race* 358
- *7.16.4.4 Sensory Impairments* 358
- *7.16.4.5 Muscle Weakness* 358
- *7.16.4.6 Gait* 359
- *7.16.4.7 Balance* 359
- *7.16.4.8 Medications* 360
  - *7.16.4.8.1 Antidepressants* 360
  - *7.16.4.8.2 Sedatives/hypnotics/anxiolytics* 361
  - *7.16.4.8.3 Neuroleptics/antipsychotics* 361
  - *7.16.4.8.4 Cardiovascular medications* 361
  - *7.16.4.8.5 Analgesics and anti-inflammatory medications* 362
- *7.16.4.9 Medical Disorders* 362
- *7.16.4.10 Psychiatric Disorders* 363

7.16.5 PREVENTION OF FALLS 363

- *7.16.5.1 Medical Prevention Strategies* 364
- *7.16.5.2 Rehabilitative Strategies* 364
- *7.16.5.3 Environmental Strategies* 364
- *7.16.5.4 Combined Prevention Strategies* 365

7.16.6 SUMMARY 366

7.16.7 REFERENCES 366

7.16.1 INTRODUCTION

Falls are a common experience among older adults. Approximately one-third of community-dwelling older adults (Studenski et al., 1994; Tinetti, Speechley, & Ginter, 1988; Walker & Howland, 1991), 20% of older adults in acute care facilities (Kay & Tideiksaar, 1990), and half of older adults in long-term care facilities (Speechley and Tinetti, 1991; Tinetti, Baker, et al., 1994; Wright, Aizenstein, Volger, Rowe, & Miller, 1990) fall each year. Within a year, half of these fallers will have fallen again (Hernandez & Miller, 1986; Wolinsky, Johnson, & Fitzgerald, 1992). Studies suggest that these estimates may underestimate true fall rates. For example, Cummings, Nevitt, and Kidd (1988) found poor recall of past falls among community-dwelling older adults. Rates among institutionalized elderly may be higher as well, due to under reporting of noninjurious falls. This is a particular concern in studies utilizing chart reviews as their primary means of data collection (Kanten et al., 1993).

While the probability of falling is greater among nursing home residents, the vast majority of falls (85%) occur among community-dwelling older adults (Tideiksaar, 1993b). This is due to the greater number of older adults residing in the community; only 5–10% of older adults reside in long-term care facilities (Brock, Guralnik, & Brody, 1990). Nearly 50% of falls result in minor or soft tissue damage (Nevitt, Cummings, Kidd, & Black, 1989), while 10–20% result in more serious injury (Tinetti et al., 1988). Hospital admission follows 5–10% of falls (Gibson, Andres, Isaacs, Radebaugh, & Wore-Peterson, 1987) and falls account for 10 – 15% of emergency medical service (EMS) transports and trauma admissions among older adults (Mosenthal, Livingston, Elcavage, Herritt, & Stucker, 1995; Wofford, Heuser, Moran, Schwartz, & Mittelmark, 1994). Hip fractures result in tremendous physical and emotional consequences and an increased risk of death. Over 90% of the approximately 215 000 people who experience hip fractures each year in the USA (National Institutes of Health, 1991) occur as a result of a fall (cited in Graafmans et al., 1996). They are a leading cause of decreased independence and entrance into long-term care facilities (National Center for Health Statistics, cited in Greenhouse, 1994). Nearly half of older adult patients who enter a hospital due to a fall become long-term care patients following their injuries (Kay & Tideiksaar, 1990). Increased morbidity is also associated with falls (Rubenstein, Josephson, & Robbins, 1994). In 1987, 75% of all deaths due to falls in the USA occurred among older adults, even though they made up only 12% of the population at that time (Greenhouse, 1994).

The goal of this chapter is to provide an overview of the empirical literature on falls among older adults, beginning with a consideration of a few conceptual and methodological issues that plague this research area, and proceeding through the consequences of falls, factors associated with falls, and finally fall prevention. This is not intended to be an exhaustive review of the falls literature. Interested readers are referred to Tideiksaar (1997) for additional consideration of this literature.

7.16.2 CONCEPTUAL AND METHODOLOGICAL ISSUES

What constitutes a fall? The answer to this question varies according to who is being asked. One of the major problems with falls research is the failure to adopt a single definition of falls. Most definitions are consistent with the notion that a fall is an unintentional event that leads to an individual coming to rest on the ground or a lower level. However, specific aspects of a fall may be excluded from one study's definition while included in another. For example, Tinetti et al. (1988) defined a fall as "an event which results in a person coming to rest on the ground or other lower level, not as the result of a major intrinsic event (such as a stroke or syncope) or overwhelming hazard" (p. 1702). However, Hernandez and Miller (1986) included in their definition any abrupt accidental or purposeful change in position which could potentially result in an injury. While both definitions are acceptable, they make comparable results between the two studies difficult to interpret. In an effort to establish guidelines into research on falls, the Kellogg International Work Group (1987) proposed the following definition of a fall: "A fall is an event which results in a person coming to rest inadvertently on the ground or other lower level and other than as a consequence of the following: sustaining a violent blow, loss of consciousness, sudden onset of paralysis, as in a stroke, or an epileptic seizure" (Gibson et al., 1987, p. 4). Since this definition was proposed, it has been utilized in several studies. The use of such a single definition allows for comparisons across various studies, which has been difficult with the majority of published fall studies.

The definition of a fall is not the only problem found in the falls research literature. Diverse subject pools, different criteria for fallers, and disparate outcome measures also preclude comparisons among studies. Subjects in different studies vary along a variety of demographic variables (e.g., housing status, age, and health). Very few studies have included a wide age range of subjects (i.e., middle age and oldest-old adult subjects) and none have included samples across both community-dwelling older adults and nursing home residents. While studies have reported the populations investigated, sample differences may account for differences in the results rather than the experience of a fall. Studies have defined fallers both as individuals who have fallen at least once and individuals who have fallen two times or more. Since most results are based on differences between fallers and non fallers, discrepancies in criteria for group inclusion may account for variance in results. Finally, the use of a variety of outcome measures complicates the interpretation of results from various studies. Fall rates, total number of falls, and percentage of group experiencing a fall have all been used as outcome measures in studies of fall risk factors. These three outcome measures address different aspects of falls and are not always directly comparable. For example, a study showing a high total number of falls may have a low percentage of fallers and several repeat fallers may account for the large total number of falls. In contrast, a high total number of falls may be due to a relatively small number of falls among a large number of fallers. When interpreting or comparing the results of various studies, it is imperative that these fundamental methodological differences be considered.

7.16.3 CONSEQUENCES OF FALLS

7.16.3.1 Physical Consequences

Falls can result in a variety of physical consequences, ranging from minor injuries (e.g., abrasions, contusions, sprains, lacerations) to serious injuries (e.g., head trauma, spinal cord injuries, fractures, internal injuries). Thirty-one percent of falls result in minor injury and another 6% result in serious injury (Morse, Tylko, & Dixon, 1987). Though the percentage of falls resulting in fractures ranges from only 1–6% for the USA (Tinetti et al., 1988), France (Vellas, 1985, cited in Gibson et al., 1987) and Canada (Gryfe, Amies, & Ashley, 1977), "most deaths and long-term disability [related to falls] are … due to the complications of fall-related fractures, especially hip fractures and other fractures among the elderly" (Melton & Riggs, 1985, p. 525). In view of the significance of fractures, falls related to fractures will be the focus of this section.

The most common fractures resulting from falls are to the wrist, upper arm, and pelvis (Melton & Riggs, 1985). The occurrence rate of fractures to the hip, wrist, pelvis and upper arm increases with age (Melton & Riggs, 1985), with such fractures among US, UK, and Scandinavian women outnumbering men between two and three to one (Evans, Prudham, & Wandless, 1979; Kreutzfeld, Haim, & Bach, 1984, cited in Gibson et al., 1987). The incidence of hip fractures increases after age 50 (Cummings, Kelsey, Nevitt, & Dowd, 1985), with a steep increase after age 70 (Gibson et al., 1987). Falls in the USA result in an estimated 215 000 hip fractures each year, resulting in 100 000 new cases of severe long-term disability (National Institutes of Health, 1991). On average, hip fractures result in a 21-day acute-care hospitalization (Baker & Harvey, 1985). The National Institutes of Health (1991) estimate that, by the year 2030, there will be 400 000 hip fractures incurred each year.

Of those individuals who are independently ambulatory prior to incurring a hip fracture, up to 75% will require ambulatory assistance for 6–12 months following the fracture (Tideiksaar, 1997). Of those hospitalized for a fall, only 50% survive one year (Tideiksaar & Kay, 1986). For those over the age of 85 residing in a nursing home, one fall out of five ends in death (Baker et al., 1984 as cited in Tideiksaar, 1989). Each year in the USA, 1800 falls are recorded as immediately fatal (Baker, O'Neill, & Karpf, 1984 as cited in Baker & Harvey, 1985).

Head trauma is a relatively common consequence of falls. Most head injuries resulting in acute hospitalization are from falls (Saywell, Woods, Rappaport, & Allen, 1989). Curtis, Wofford, and Branch (1995), cited in Tideiksaar, (1997), found that 54% of older adults hospitalized for subdural hematomas had fallen shortly before they were hospitalized.

Falls sometimes result in the faller remaining immobile for a period of time. Immobility secondary to a fall can result in hypothermia, broncho pneumonia, pulmonary emboli, and dehydration, all of which may lead to death (Mitchell, 1984; Tideiksaar, 1993c; Tideiksaar & Kay, 1986). The amount of time a person remains immobile immediately following a fall is termed "lie time" and is a marker for weakness, illness, and social isolation (Gibson et al., 1987). Moreover, lie time is a mortality risk factor. British studies indicate that lie times of at least 60 min are associated with death within six months (Wild, Nayak, & Isaacs, 1981, cited in Gibson et al., 1987). Unfortunately, data from France suggest that long lies are not uncommon. In Toulouse, 10 of 48 elderly patients hospitalized for falls had laid on the ground for an hour or more, and one for three days (Vellas, 1985, cited in Gibson et al., 1987).

7.16.3.2 Psychological Consequences

Older adults experiencing a fall also face potential emotional consequences. However, the psychological consequences following a fall are often ignored in falls research. Psychological consequences may lead to further physical decline. For example, an individual who experiences a fall may lose confidence in his/her ability to perform activities without falling again. Therefore, in an effort to avoid further falls, he/she restricts his/her activities. The decline in activities leads to a decline in physical health and muscle tone. If an individual remains inactive, he/she can become weak to the point of becoming unable to perform activities. Thus, psychological factors may result in decreased physical abilities and increased physical dependence.

Psychological factors have been examined as both risk factors and consequences of falls. Due to the retrospective nature of most falls research, it is difficult to determine if psychological problems related to a fall were present prior to, or appeared as a consequence of, a fall. This issue is further complicated by the fact that psychological factors can be both consequences and risk factors of falls. Many older adults experience a loss of confidence, depression, or fear of falling after a fall (Downton & Andrews, 1990; Tinetti, Richman, & Powell, 1990; Tinetti et al., 1988). The presence of any of these psychological problems may lead to decreased mobility and further physical decline as described previously.

The failure of many falls researchers to collect information on psychological consequences of falls reveals an emphasis on medical factors rather than psychological factors within falls researchers. Psychological factors are related to the occurrence and outcome of falls. Simpson and Mandelstam (1995) found that about 23% of hospital patients who had experienced a fall refused to participate in physical therapy training even though it was determined that they were capable of participation. Given that the patients were physically capable of participation, psychological factors likely account for the patient's refusals. This finding and others discussed below suggest that confidence in abiding abilities, self-perceived health status, depression, and fear of falling can have detrimental effects on health, activity level, and life satisfaction. It is important that research into psychological factors be conducted and prevention and treatment programs be developed.

Tinetti, Richman, and Powell (1990) developed the Falls Efficacy Scale (FES) to measure an individual's confidence in his/her ability to avoid a fall while performing specific activities. This scale was developed initially as a measure of fear of falling. Fear was defined as "low perceived self-efficacy or confidence at avoiding falls" (p. P239). Therefore, it is utilized more accurately as a measure of confidence. Decreased confidence has been shown to be related to a history of falls (Tinetti, Mendes deLeon, Doucette, & Baker, 1994). Poor ability to estimate mobility was a significant predictor of falls in a study conducted by Morse, Morse, and Tylko (1989). Poorer ability to accurately predict overall physical abilities appears to place one at risk of experiencing a fall.

Falls influence health ratings in otherwise healthy individuals (Cwikel, Kaplan, & Barell, 1990). People who have experienced a fall rate their health as being poorer after a fall than nonfallers (Cwikel et al.). Subjective health ratings are also related to severity of injury due to a fall. People requiring emergency room visits after a fall rate their health as poorer than fallers with less severe physical consequences (Cwikel et al.). Further, longitudinal studies demonstrate that subjective health rating is a predictor of morbidity and mortality (Kaplan, Barell, & Lutzky, 1988 cited in Cwikel et al.). Intuitively, it makes sense that people who report poor health and ability would fall more frequently, especially if their self-report is accurate. However, individuals who are actually in poor health but report good health (i.e., high confidence in a poor ability) may be more at risk. It is difficult to determine if subjective health ratings reflect actual health status. A complex interplay among reported health, actual health, and falls may be involved and warrants further investigation.

Depression has been found to be associated with falls (Tinetti et al., 1988), recurrent falls (Luukinen, Koski, Laippala, & Kivela, 1995), and to distinguish between groups of fallers and nonfallers (Tinetti, Liu, & Claus, 1993). The presence of depression significantly and independently predicted falls in community-dwelling psychogeriatric patients (Asada, Kariva, Kitarjima, Kakuma, & Yoshioka, 1994) and hospital patients (Soja, Kippenbrock, Hendrick, & Nyhius, 1992). Furthermore, antidepressant medications have been found to be associated with decreased risk of falling (Asada et al.), possibly due to the alleviation of depressive symptoms. These findings suggest that depression is a risk factor for falls. However, the majority of studies are retrospective in nature and should be interpreted with caution. Depression is a common diagnosis applied to older adults; therefore, it is difficult to determine from cross-sectional data whether depression is related to the occurrence of a fall or was a premorbid condition. Depression is often measured within fallers and nonfallers and then the two groups are compared. Researchers then conclude from these comparisons that the presence of depression is predictive of falls. However, a study which finds depression as a risk factor for falls by utilizing data collected after a fall may not be accurate. Findings may indicate that falls are predictive of depression. Prospective studies of depression and falls are needed to clarify the impact falls have on the expression of depression.

Fear of falling is reported by both older adults who have and have not fallen. Between 20 and 33% of nonfallers (Downton & Andrews, 1990; Howland, Peterson, Levin, & Fried, 1993) and 48 to 56% of fallers (Downton & Andrews, 1990; Tinetti et al., 1988) reported being afraid of falling. Walker and Howland (1991) found fear of falling to be the most prevalent fear reported by older adults living in a housing community for the elderly. Studies investigating the fear of falling are hindered greatly by the lack of a single definition or measure of fear of falling. Every study utilizes a different measure and differences in definition influence the results. For example, the FES discussed earlier is often used as a measure of fear of falling when it actually measures confidence in the ability to avoid a fall. Results from studies utilizing this scale cannot be compared with studies utilizing a direct measure of fear of falling. Using an open-ended question on fears, Null-Kogan (1996) found fear of falling to be the second most frequently reported fear among community-dwelling older adults. Of the 50 participants in this study, 68% reported at least a little fear of falling. Howland et al. (1993) demonstrated that injurious falls may create higher rates of fear of falling. Forty-one percent of people who experienced injurious falls reported a fear of falling as opposed to 29% of those who experienced noninjurious falls. Fear of falling is associated with higher dependency, greater physical symptomatology (Downton & Andrews, 1990), greater drug consumption (Downton & Andrews, 1990), higher reported depression and anxiety (Arfken, Lach, Birge, & Miller, 1994; Downton & Andrews, 1990; Howland et al.), restriction of movement and activities (Howland et al.), and increased morbidity. People with a severe fear of falling after a fall are more likely to have long hospital stays and die (Murphy & Isaacs, 1982).

One of the most serious consequences of falling is the subsequent restriction in movement accompanying the fear of future falls. In a study of nursing home residents using wheelchairs, 76% reported they used wheelchairs voluntarily because they were afraid of falling if they tried to walk (Pawlson, Goodwin, & Keith, 1986). Between one quarter and one half of fallers stated that they restricted or avoided activities due to fear of falling (Howland et al., 1993; Nevitt et al., 1989; Tideiksaar, 1993c; Tinetti, Mendes de Leon, et al., 1994). Restriction in movement can lead to social isolation and withdrawal as well as an increased risk of falling due to decreased muscle strength and mobility. Prolonged bed rest can lead to bone demineralization, lung and peripheral circulation stasis, skin and soft tissue damage, and altered supply of oxygen to the brain resulting in confusion and cognitive problems (Commodore, 1995). It can also lead to joint contractions, balance loss, pneumonia, pulmonary emboli, and death. Fear of falling and restriction of movement appear to be more common in people with a history of past falls.

7.16.3.3 Economic Consequences

It is estimated that the direct and indirect costs of accidental injury in the USA is between 75 and 100 billion dollars annually (Committee on Trauma Research, Commission on Life Sciences, National Research Council, & Institution on Medicine as cited in Urton, 1991). Falls are the leading cause of accidental injury. The direct costs include medical treatment, rehabilitation, nursing home care, medical equipment, and loss of productivity by the faller and their caregivers. A study conducted in Seattle in 1983 found that 1740 fall-related injuries cost US $500 000 for emergency room services alone (Hongladarom, 1983 cited in Commodore, 1995). The estimated cost of acute care for fractures is roughly 10 billion dollars each year (Sattin, 1992 cited in Tinetti, Baker et al., 1994). Hip fractures are the most serious fractures in terms of cost. In 1987–1988, a study conducted in the UK reported expenses of 160 million pounds for the medical treatment of hip fractures (Gibson et al., 1987). The annual cost of hospitalization and long-term care related to fall injuries is 2.5 billion dollars in the USA (Tideiksaar, 1997). Each year billions of dollars are spent to care for people after they experience a fall.

The economic cost of falls can also be indirect. For example, the additional costs of staff time and allotted hospital days for inpatient falls can mount quickly following a fall. Moreover, the additional care required of a fall victim is time taken away from other residents in the facility. Since falls are the leading cause of accidental traumatic injury in individuals over the age of 65, emergency medical costs are also affected by falls and fall-related injuries (Tinetti & Speechley, 1989). Campbell, Reinken, Allan, and Martinez (1981) found that patients admitted to a hospital for traumatic injury experienced longer stays than patients admitted for a nontrauma-related reason (43 days vs. 25 days). Individuals experiencing numerous falls are more likely to have longer stays and be placed in nursing homes (Wolinsky et al., 1992). In the UK and France, one of the leading causes of admission to geriatric hospitals, residential homes, and nursing homes is falls not resulting in fractures (Gibson et al., 1987). The demands of frequent fall-related trauma influence the allocation of resources.

Falls also impact psychological, familial, and societal resources. Loss of independence, depression, and fear of falling require emotional and psychological resources of health care providers, fallers, and families. For example, pain resulting from a fall can have far-reaching effects on an individual's functioning. Families may experience disruption in routines as family members adjust to living and interacting with a physically diminished individual. Once-independent members may require assistance from family members who must rearrange schedules or living situations to accommodate the new demands of caregiving. Finally, community resources may be allotted differently to meet the needs of physically limited older adults. For example, more funds may be designated for use in senior centers or elder care programs rather than building a new school. Falls can have far reaching effects on many diverse areas.

7.16.4 RISK FACTORS AND THE CONTEXT OF FALLS

A substantial number of extrinsic and intrinsic factors associated with falls have been identified. It is important to note that when considering these individual factors, the reader should be aware of the general finding that multiple risk factors are involved with any particular fall.

7.16.4.1 Extrinsic Factors

Every fall occurs within an environment. However, external or extrinsic factors are rarely the sole cause of individual falls. In a study conducted by Rubenstein et al. (1988), no fall could be completely explained by environmental factors. Most falls occur as a result of both intrinsic and extrinsic factors. For example, an older adult experiencing depression may trip over the edge of a rug overlooked due to inattentiveness. This might not have occurred had the individual been mentally healthy and attentive. Environmental factors are associated with approximately one-third of falls occurring within the home and two-thirds of falls occurring outside of the home (Nevitt et al., 1989). This may be an underestimation of the role of environmental factors due to the retrospective nature of research into falls. A faller's memory of factors involved in a fall may be faulty due to injuries, personal beliefs about the cause of the fall, or the passage of time. Moreover, a fall can alter the appearance of the physical environment (e.g., an obstacle may be moved following a fall), thereby influencing the recall of relevant environmental information. In a study investigating the prevention of falls, 53.8% of residents in a skilled nursing care facility and 35.1% of residents in a board and care facility were given recommendations regarding changes to their environment (Rubenstein et al.). The high rate of environmental hazards led the investigators to project that environmental factors may be involved in as many as 50% of falls. Some of the extrinsic factors considered in previous research are considered below.

7.16.4.1.1 Time of fall

Research into time-related factors has produced conflicting results. Several studies have found peak times for the occurrence or nonoccurrence of falls (Brown & Edelstein, 1995; Graffmans et al., 1996). Several studies suggest higher occurrences of falls in institutions during the morning and evening shifts (Marx, Cohen-Mansfield, & Werner, 1990; Walker & Howland, 1991), but no consensus exists across studies. Quail (1994) and Walker and Howland (1991) reported that increased fall rates occurred at times of increased activity. In their studies, falls were more likely to occur during toileting at night and while performing activities during the day. In addition, Heslin et al. (1992) found that falls from beds were more likely to occur at night. Thus, the location of specific activities influenced where falls occurred. It appears that falls are more likely to occur at periods of increased activity, and time of day influences which activities are being performed. The activities most frequently reported at the time of a fall were ambulation and transfer of position (Heslin et al.). These studies suggest that the differential effects of time of day reported in individual studies may be indicative of increased periods of activity.

Seasonal influences also impact fall rates. Commodore (1995) reported an increased number of falls during the winter months. This was likely due to increased numbers of slippery surfaces (e.g., ice) and outdoor environmental hazards (e.g., snow covering holes). Mobility is also restricted by bulky clothing during the colder winter months. Although it has not been examined in the literature, increased physical activity during the summer months may lead to a similar increase in falls during the summer months. More information is needed on seasonal fall rates to investigate seasonal contributions.

7.16.4.1.2 Location of fall

Previously mentioned fall rates differ across residential status. For example, nursing home residents are more likely to experience falls than community-dwelling older adults. In addition, recently admitted patients to a nursing home are more likely to fall than long-term residents (Quail, 1994), with over 60% of nursing home residents experiencing a fall during their first year (Tinetti, 1987). The unfamiliar environment of the nursing home is likely to contribute to many falls occurring with newly admitted residents.

Rubenstein et al. (1988) reported that 80% of falls occurring in a multilevel care facility occurred indoors. Similar results were found in a prospective study in residents of homes and apartments for the elderly in Amsterdam (Graffmans et al., 1996). Among acute care patients, 65% of falls occur in the bedroom, half of these occurring from the bed (Heslin et al., 1992). It is unclear if fall rates coincide with length of time spent in a specific location. For example, Brown and Edelstein (1995) found falls among psychogeriatric inpatients to occur more frequently in the bedroom, hallway, dining room, dayhall, bathroom, and outside. However, no overall percentage of time spent in these locations was reported. If an individual spends 75% of the day in bed, the chances of a fall occurring in his/her bedroom is higher than one occurring in another location. Lipsitz, Jonsson, Kelley, and Koestner (1991) found that older adults frequently fell when they ambulated or transferred position (e.g., stood up from a chair). It is likely that the frequency of activities occurring in specific locations influence fall rates in those locations.

Stairs are hazardous for many older adults. In one study, 10% of falls among community-dwelling adults occurred around stairs (Tinetti et al., 1988). Decreased visual perception may lead to miscalculation of the length or height of individual steps. Archea (1985), cited in Gibson et al. (1987), found that most accidents involving stairs occurred while the individual was descending the stairs and were often related to poor architectural design. Stair-related accidents occur more frequently among young-old adults (65–75 years) (Abrams, 1984 cited in Gibson et al.). This may be due to decreased use of stairs by old-old adults or increased caution while using them. However, old (75–85 years) and oldest-old (85 + years) adults experience more serious consequences from falls on stairs than young-old adults (Gibson et al.). Where a fall occurs can be influenced by the amount of time spent in a location, activities that occur in a location, and particular environmental characteristics of the location.

7.16.4.1.3 Walking aids

Improper use of a walking aid (e.g., cane, walker) is a common contributory factor to a fall. Fallers are more likely to use an ambulatory aid (Mahoney, Sager, Dunham, & Johnson, 1994; Morse et al., 1989; Robbins et al., 1989a, 1989b). It is unclear whether this is a direct result of previous falls or unrelated conditions. Tinetti (1987) found the use of a walking aid to be the most common external factor associated with falls in a nursing home population. Improper use of walking aids is also associated with the inability to get up following a fall (Tinetti et al., 1993). As previously discussed, individuals who are unable to get up after a fall are more likely to experience serious physical consequences than individuals who are able to get up after a fall. The increased risk of falling due to use of walking aids may be due to reasons related to the use of the walking aid. Campbell, Borrie, and Spears (1989) stated that the increased risk due to the use of walking aids was a reflection of frailty or poor mobility requiring the use of such aids. If this is found to be true, it is unknown if improper use of these aids would be an additional risk factor for experiencing a fall.

7.16.4.1.4 Restraints

Restraints have been used to prevent frail or confused older adults from falling and incurring injury. In a study conducted by Heslin et al. (1992), 98% of acute care patients who fell from a bed had restrictive devices and 2% had restraints on them at the time of the fall. This suggests that many falls may occur as a result of failed attempts to remove or escape from restraints. Increasing evidence suggests that restraints do not prevent falls (Evans & Strumpf, 1990) and may be associated with injurious falls (Heslin et al.). A study investigating the impact of removing restraints found that removing them led to a significant increase in the overall fall rate but not in the rate of serious falls when compared with a nonrestrained control group (Ejaz, Jones, & Rose, 1994). Given these results, restraints should be used with caution in individuals at risk of falling.

7.16.4.1.5 Footwear

In humans, contact between the foot and the ground is necessary to provide balance and stability. Loose or poor fitting shoes contribute to instability among older adults (Quail, 1994) by reducing the stability normally provided by contact between the foot and ground. Several studies implicate loose or poor fitting footwear (e.g., slippers) as major contributors to falls (Kinsman, 1983, cited in Gibson et al., 1987). Shoes with limited traction or slippery soles may also increase the risk of falling. Research into footwear appropriate for older adults needs to be conducted.

7.16.4.1.6 Environmental hazards

Numerous articles have attempted to list hazards commonly encountered in the home. Table 1 lists many of these potential hazards. Tinetti et al. (1988) found that 25% of falls occurred as a result of tripping over an object. Environmental hazards may be more of a factor in falls by active older adults. However, with increasing age and infirmity, environmental factors may play a significant role. Few studies have investigated the impact of specific environmental hazards or the utility of making environmental changes in an attempt to decrease falls.

7.16.4.2 Intrinsic Risk Factors

An extensive array of intrinsic variables associated with falls among older adults has been examined in the literature. The factors discussed include: demographic characteristics (e.g., gender, age); physical impairment (e.g., muscle weakness, hearing loss, visual acuity loss); the results of various physical impairments (e.g., ambulation, gait, and balance problems); number, types, and amount of medications (e.g., antidepressants, sedatives), medical disorders (e.g., Parkinson's disease, cardiovascular disease); and psychological factors (e.g., denial of limitations, cognitive status). A comprehensive survey of the research addressing intrinsic factors is beyond the scope of this chapter. Therefore, representative research findings regarding the most commonly examined intrinsic variables will be discussed.

7.16.4.3 Demographic Factors

Demographic information assembled from numerous studies allows clinicians to identify populations at risk for falls. Identification of high-risk groups helps clinicians predict the risk levels of specific individual clients and can aid in the development of treatment and prevention programs.

7.16.4.3.1 Gender

The findings of research on gender as a risk factor for falling have been contradictory (Brown & Edelstein, 1995), with some studies revealing gender differences and others finding no significant differences. However, the majority of studies suggest that women fall more frequently than men (Cwikel et al., 1990; Nevitt, Cummings, & Hudes, 1991; Poster, Pelletier, & Kay, 1991) and have more risk factors associated with falls than men (Campbell et al., 1989). For example, two studies conducted in US nursing homes (Gross, Shimamoto, Rose,

Table 1 Environmental hazards.

Location	*Hazard*
Outdoors	Ice, water, holes, cracks, steps
Bedroom	Improper bed height, wheels on furniture, nonlocking bed wheels, poor lighting
Bathroom	Mats, slippery floors, low or high toilet, poor lighting, glare from floor, lack of grab bars, sliding carpets, inaccessible tub or shower stall
Kitchen	Unsturdy step stools, poor lighting, slippery linolium floors, slippery bathtubs, glare from floor, high cabinets
Living room	Low furniture, extension cords, wheels on furniture, loose rugs, poor lighting, inappropriate chair height, high steps, worn stair treads, lack of arm rests
Everywhere	Objects in pathway, sliding throw rugs, worn carpets, poor access to switches and lamps, raised door sills, lack of night lights, clutter, limited walking space, waxed or slippery floors

& Frank, 1990; Kalchthaler, Bascon, & Quintos, 1978) found a greater rate of falls among women than men. Similarly, Campbell et al. found that among community dwellers 70 years of age and older, women were at greater risk than men of falling. Moreover, a French longitudinal study of community dwellers (mean age 73) conducted by Ousset et al. (1996) found that 48% of the women and 29.8% of the men reported one or more falls during a two-year period. A Japanese longitudinal study of community dwellers conducted by Yasumura et al. (1996) found the same pattern of results as Ousset et al. However, they found a lower incidence of falls with 14.7% of men and 23.5% of women reporting at least one fall during a one-year period. The shorter time period used in the Yasumura et al. study may account for the smaller fall rate than that reported by Ousset et al.

Women incur injury as a result of falls more frequently than men (Hornbrook et al., 1994), with elderly women over the age of 75 years most likely to experience a serious injury as a result of a fall (Gibson et al., 1987). Further, Abrams (1984), cited in Gibson et al., noted that while women in the UK represent only 25% of the 65+ population, they account for over 50% of hospital-treated falls within this population. The incidence of hip fractures among women increases around age 40 and doubles every five to six years after that (Brody, Farmer, & White, 1984, cited in Gibson et al.). Differences in risks factors have also been related to gender. For example, in their prospective study, Campbell et al. (1989) found total number of drugs taken, use of drugs that can cause postural hypotension, use of psychotropic drugs, use of hypnotic drugs, mental status, and joint replacement were risk factors for women but not for men.

In contrast to the previous studies, Van Dijk, Meulenberg, van de Sande, & Habbema (1993) found a significantly higher rate of falling among male than female demented nursing home residents. Other studies have found no relation between gender and falling among patients in a general hospital (Poster et al., 1991), patients in a veteran's center and geriatric unit (Morse et al., 1987), and individuals discharged from an acute-care hospital (Mahoney et al., 1994). The different findings among these studies may be attributable to subject differences (Brown & Edelstein, 1995), with community dwelling women showing a relatively consistent higher rate of falling than community dwelling men. In the nursing home studies, higher fall rates were found among male residents with dementia (Van Dijk et al.). The two studies that found higher fall rates among women in nursing homes (Gross et al., 1990; Kalchthaler et al., 1978) appeared to include a more cognitively heterogeneous population. The confusion and disorientation that often accompanies dementia could place dementia sufferers at greater risk of falling.

Studies suggest that gender differences dissipate with age and fall rates become roughly equal after the age of 70 (Luukinen et al., 1995; Poster et al., 1991). Moreover, gender does not appear to differentially influence fall rates in elderly nursing home patients (Luukinen et al.).

7.16.4.3.2 Age

The relation between age and falls is also unclear, with numerous studies supporting and failing to support a relation. Campbell et al. (1989) found fallers to be older than nonfalling inpatients, and Kalchthaler et al. (1978) found higher frequencies of falling among older nursing home residents 40 years of age and older. Members of the Kellogg International Work Group on the Prevention of Falls by the Elderly report that most falls occur among older women (Gibson et al., 1987) even when the fall figures were adjusted for the proportion of women in the population (Prudham & Evans, 1981). Graafmans et al. (1996) found age to be related significantly to recurrent falls (two or more falls) but not to single falls among a group of community-dwelling adults in Amsterdam. The participants were 1.7 times more likely to fall for each 10-year increase in age. Wofford et al. (1994) found that adults over the age of 85 were seven times more likely to experience a fall than those aged 65–74. A study conducted by Studenski et al. (1994) found that while the probability of falling increased with age, the probability of incurring an injury as a result of a fall did not increase. This may be due to the decreased number of vigorous activities performed by older adults. For example, Speechley and Tinetti (1991) found that "vigorous" older adults were more likely to incur an injury as a result of a fall than "frail" older adults. Despite conflicting reports on risk, older adults appear more likely to experience a fall (Cwikel et al., 1990; Greenhouse, 1994; Poster et al., 1991) than younger adults.

In contrast, several studies have failed to support the relation between falling and age. Lipsitz et al. (1991) studied 70 fallers and 56 nonfallers in a nursing home and found no significant difference in age between fallers and nonfallers. Similarly, Morse et al. (1987) did not find age to be a risk factor for classifying fallers and nonfallers. Van Dijk et al. (1993) found no relationship between fall rates and age among older adults with dementia. In a study of 306 community-dwelling adults aged 70-years and older, Studenski et al. (1994) found no significant age difference between fallers and nonfallers. Moreover, Rosen, Campbell, Villanueva, and Morgan (1985) found the mean age between nonfalling and falling community-dwelling residents age 70 or greater not statistically different. Mahoney et al. (1994) compared fallers and nonfallers after discharge from a community hospital and found no significant differences in age. Again, differences in subject pools complicate the interpretation of these results.

While age appears to influence risk of falling, other variables (e.g., health and mental status) often associated with aging may be more pertinent as moderator variables. These variables will be discussed later in the chapter. Carefully conducted studies, controlling for the presence of moderator variables, need to be conducted to determine the direct influence of age on falls.

7.16.4.3.3 Race

Tentative findings suggest that race is also associated with the risk of falling. Specifically, Wofford et al. (1994) and Nevitt et al. (1989, 1991) found that Caucasians have a higher risk of falling and requiring EMS transport following a fall than African-Americans. Further studies of the relations between falls and ethnic and racial variables are needed. The falls literature has not addressed racial and ethnic group differences. The relations found between falling and these demographic variables should be interpreted with care because of possible differences in the reporting of falls by the different groups. Many variables may influence the reporting of falls and these issues need to be addressed.

7.16.4.4 Sensory Impairments

Impairments of vision, the vestibular system, and the somatosensory system can contribute directly and indirectly to falls. For example, visual impairments can preclude the detection of obstacles that may lead to a fall. Inaccurate or inadequate visual or proprioceptive information or a malfunctioning vestibular system can all lead to gait or balance problems. Attenuated sensations of touch and pressure on the bottom of the feet (sometimes due to peripheral neuropathy secondary to diabetes mellitus) can contribute to poor control of balance during stance and gait. The ability of the somatosensory system to function properly declines with age, as does the ability of the sensory-motor system after age 55 (Lord, 1996) and the ability to correctly register joint movement and position (Simoneau & Lebowitz, 1996). On a more positive note, good vision can sometimes compensate for other sensory deficits (Whipple, Wolfson, Derby, Singh, & Tobin, 1993).

The few studies that have examined the relationship between sensory impairment and falls have produced inconsistent findings. Mahoney et al. (1994) compared fallers and nonfallers with regard to visual impairments and found a significantly higher rate of visual impairment among fallers. Similarly, Granek et al. (1987) found a significantly higher rate of falls among blind than sighted individuals. Nevitt et al. (1989) found both decreased visual acuity and depth perception to place individuals at increased risk of multiple falls. Similarly, Felson et al. (1989) found a relationship between risk of hip fracture and poor visual acuity. Interestingly, they also found that older adults with good vision in one eye and moderately good vision in the other had an increased risk of hip fracture. Lord (1996) suggests that this could be due to poor stereoscopic vision that, in turn, reduces depth perception and increases fall risk. Lord (1996) concluded from his studies that contrast sensitivity, “detection of large visual stimuli under low contrast conditions,” is more important than visual acuity in predicting falls. The significance of these findings is particularly important when one considers that contrast sensitivity is essential for detecting edges, such as steps and misalignments of walking surfaces.

In contrast to the foregoing findings, Morse et al. (1987) found no significant differences between fallers and nonfallers with regard to visual impairments or hearing deficits. Similarly, Brocklehurst, Robertson, and James-Groom (1982) and Robbins et al. (1989a, 1989b) found no relation between visual acuity and falls among older adults beyond the age of 65. Though a comparison of findings regarding vision and falls is precluded due to methodological differences among studies, it is clear that the majority of studies suggest a relationship between falls and various measures of vision. More comparative studies utilizing similar subject populations and methodologies are sorely needed.

7.16.4.5 Muscle Weakness

The musculoskeletal system is essential for maintaining postural control and movement through the environment. Declines in muscle mass begin at approximately 25 years of age.

Muscle strength losses of 10–20% occur between the ages of 40 and 70, with losses of 30–40% occurring between the ages of 70 and 80 (Whitbourne, in press). The greatest declines in muscle strength in later years occur in muscle groups of the lower extremities, including the muscle groups involved in ambulation, transferring to and from seated positions, and climbing stairs.

Several studies have found hip and lower extremity muscle weakness to be related to falls (e.g., Campbell et al.,1989; Lipsitz et al., 1991; Robbins et al., 1989a, 1989b; Whipple, Wolfson,& Amerman, 1987; Wolfson, Judge, Whipple, & King, 1995). For example, Campbell et al., in a one-year prospective study, examined multiple risk factors for falling among community-based older adults. Female fallers had significantly less leg strength than female nonfallers, with a relative risk factor of 1.6. No differences were found in muscle strength between male fallers and nonfallers. Robbins et al. (1989) also examined multiple risk factors using both community-dwelling and institutionalized older adults. Hip weakness was one of the three risk factors identified. This risk factor yielded a 68% chance of falling within a one-year period, those without this risk factor had a 19% chance of falling within the same time period. Knee and ankle weakness have been found significantly more prevalent among fallers than nonfallers (Whipple et al.). Similarly, Wolfson et al. found that residents of a nursing home with histories of falls had less than half of the knee and ankle strength of nonfallers.

In summary, muscle weakness in the lower extremities is consistently found to be a risk factor for falling. This relation, however, may not always be direct since muscle strength is associated with a variety of other risk factors (e.g., gait, balance, speed). Fortunately, as later noted, muscle strength training can contribute to a decrease in falls.

7.16.4.6 Gait

Gait reflects the characteristics of a person's walking, which is controlled and influenced by a multitude of factors. Moreover, gait is a multidimensional construct that is measured in a variety of ways (see Duncan & Studenski, 1994, for a review). Disorders of gait affect 20–50% of older adults (Sudarsky, 1990). Impaired gait is considered to be a major contributor to falls, which can also contribute to negative psychological consequences (e.g., fear of falling) and the concomitants of fear previously discussed (e.g., immobility, further functional decline).

Numerous studies have found a relation between gait problems and falls. Morse et al. (1987), for example, found a greater percentage of gait problems among fallers (88%) than nonfallers (43%). Tinetti et al. (1988) found community-dwelling adults with balance and gait abnormalities to be as much as 1.9 times as likely to fall as those without these problems. Whipple et al. (1987) found fallers to have significantly slower gait speed and shorter stride length than nonfallers in a nursing home population. Also, Guimaraes and Isaacs (1980) found that fallers displayed slower walking speeds, shorter step length, and greater variability in their step length than nonfallers. Similarly, Wolfson, Whipple, Amerman, and Tobin (1990) revealed that fallers in nursing homes differed from nonfallers with regard to stride length and walking speed. Overall, studies of gait have found a high prevalence of gait disorders among older adults, placing them at significant risk of falling.

7.16.4.7 Balance

A balanced position requires the coordination and muscular control of individual body parts. Walking has been characterized as successive loss of balance episodes in which a fall is averted with each step by repositioning each foot (Simoneau & Leibowitz, 1996). Such postural control results from the coordination of sensory and musculo-skeletal systems. The ability to transfer from bed to floor, walk, and be seated requires balance obtained through these systems and their coordination. Therefore, problems with these systems can result in a loss of righting reflexes and an increase in body sway (Kay & Tideiksaar, 1990), leading to falls.

Poor control of posture (postural dyscontrol), as evidenced by poor balance, has been associated with falls for both community-dwelling and institutionalized older adults. Wild et al. (1981), as cited in Funk, Tornquist, Champagne and Wiese (1992), found that postural dyscontrol among community-dwelling older adults contributed to at least 67% of falls. In a prospective study of falls, Tinetti, Williams, and Mayewski (1986) found that residents recently admitted to an intermediate care facility with balance impairment were 5.4 times as likely to fall as those without a balance impairment. Similarly, Lipsitz et al. (1991), in a retrospective study of 126 nursing home residents, found subjects with a balance impairment to be 3.9 times as likely to fall as those without such an impairment. Tinetti et al. (1988), in a prospective study of community-dwelling adults 75 years of age and older, found that individuals with abnormalities of balance were 1.9 times as likely to fall when compared with individuals without such abnormalities. In a study of institutionalized and noninstitutionalized older adults, Robbins et al. (1989a, 1989b) found that fallers were significantly more likely than nonfallers to have impaired gait and balance. Balance was one of three risk factors that significantly and independently predicted falling among both populations.

Older adults may not lose their balance more often than younger individuals, but are simply less capable of correcting the imbalance (Overstall, 1992). One of the principal indices of imbalance is body sway. With increasing age there is a slowing of the ability to detect changes in posture or sway, which is associated with greater risk of falling. Brocklehurst et al. (1982) found a relation between postural sway, decreased proprioception, and falls. Similarly, Wild et al. (1981) found a relation between responses to pressure applied to the sternum (balance response) and risk of future falls. Several other studies have reported differences between fallers and nonfallers with regard to postural sway or unsteadiness (e.g., Lichtenstein, Burager, Shields, & Shiari, 1990; Overstall, Exton-Smith, & Johnson, 1977; Wolfson, Whipple, Amerman, & Kleinberg, 1986). Balance and gait tests have also been found useful in discriminating between residents of intermediate care facilities who fell two or more times from those who fell once or never fell (Tinetti et al., 1986). The findings of Tinetti et al. are particularly interesting because the situations in which balance was measured were those for which balance is particularly important (e.g., when rising, sitting, and turning around). Overall, the results of numerous studies suggest that poor balance, which increases with age, is a potentially significant risk factor for falls.

7.16.4.8 Medications

Medications have often been examined as potential factors influencing falls. However, research into the effects of drugs on fall rates is equivocal. Many studies have found that drugs influence fall risk (Cooper, 1994; Lipsitz et al., 1991; Tinetti et al., 1986), while others have found no significant effect of drugs on fall risk (Lord, Clark, & Webster, 1991; Mahoney et al., 1994; Morris, Rubin, Morris, & Mandel, 1987; Tinetti, 1987). Total number of medications taken (Lipsitz et al.; Robbins et al., 1989a, 1989b; Studenski et al., 1994), addition of a new medication (Tinetti et al., 1986), and recency of drug administration (Aisen, Deluca, & Lawlor, 1992; Lipsitz et al.; Nevitt et al., 1989) have all been associated with falls. Further, the use of more than six medications was associated with the inability to get up following a fall (Tinetti et al., 1993). Methodological differences between studies may explain disparate results. Drug variables entered into data analysis differ greatly between studies. Many studies group all drug classes together within a single category while others examine specific classes of drugs. While many studies fail to find a relation between a global medication measure and falls, the influence of particular classes of drugs have been found to influence fall risk. For example, antidepressants have been identified as a risk factor for further falls (Lipsitz et al.). In addition, there may be gender differences in drug influences on falls. Cambell et al. (1989) found the total number of drugs taken to influence fall rates among women but not among men.

Specific classes of drugs influence the risk of falls. Drugs directly influencing cognitive states or alertness may increase the risk of falling. For example, Quail (1994) found alcohol to be related to the occurrence of falls. Side effects of drugs also influence the risk of falling. For example, Downton and Andrews (1990) found sedatives associated with subjective reports of dizziness, which is related to increased risk of falling. Drug classes which have been investigated in the literature include antidepressants, sedative/hypnotics, anti-inflammatories and analgesics, cardiovascular agents, neuroleptic/antipsychotics, antihistamines/anticholinergics, antidiabetics, diuretics, anticonvulsants, nitrates, and anti-Parkinsonian medications. Several of these will be discussed in detail below. As new medications are introduced and utilized in the older adult population, the influence of medications on fall risk needs to be reevaluated.

7.16.4.8.1 Antidepressants

Antidepressants are utilized for their ability to elevate mood and increase energy. As with all the drug classes, research on antidepressants has produced mixed results. Antidepressant use has been associated with increased risk of falling among nursing home residents (Lipsitz et al., 1991), recurrent falls in long-term care patients (Luukinen et al., 1995), and distinguished between fallers and nonfallers among community-dwelling males and nursing home residents (Studenski et al., 1994; Lipsitz et al.). Moreover, it has been associated with decreased risk of falls in community-dwelling psychogeriatric patients (Asada et al., 1994). In contrast, other researchers have not found antidepressant use to be associated with falls among residents of homes for the elderly (Graffmans et al., 1996) or community-dwelling older adults (Nevitt et al., 1989). Differences in results may be attributable to differing classes of antidepressants.

Three major types of antidepressant are prescribed: tricyclics, monoamine oxidase inhibitors (MAOI), and selective serotonin reuptake inhibitors (SSRI). Each class will be discussed separately due to their various effect and side effect profiles. While each drug produces numerous side effects, only those relevant to falling will be mentioned. The exact mechanism by which tricyclic antidepressants produce their antidepressant effect is unknown but it is assumed that they block the reuptake of norepinephrine, dopamine, and/or serotonin. Tricyclics may produce orthostatic hypotension, heart palpitations, tremor, sedation, and delirium. Research into tricyclic antidepressants and falling is equivocal. Quail (1994) reported that tricyclics increased the risk of falling in an outpatient population, especially when they were combined with other drugs. In contrast, Mahoney et al. (1994) found no significant effect of tricyclics among discharged hospital patients. While no research has been conducted on MAOIs and SSRIs, they are important to mention due to their possible relation to falls. Both produce anticholinergic side effects such as dizziness, blurred vision, or nausea which can influence the possibility of falls. More research into the exact types of antidepressants producing positive relations with falling is needed to determine the risk of prescribing these drugs to the elderly.

7.16.4.8.2 Sedatives/hypnotics/anxiolytics

Sedatives/hypnotics are typically prescribed as sleep aids or calming agents. They may also be prescribed as anticonvulsants. However, they can have serious side effects including disrupted sleep patterns, confusion, light-headedness, and drowsiness. Sedatives have also been associated with subjective dizziness in community-dwelling older adults (Downton & Andrews, 1990). Sedative use has been associated with falls in community-dwelling older adults (Tinetti et al., 1988, 1993), recurrent falls of these individuals (Nevitt et al., 1989), and the inability of these individuals to get up following a fall (Tinetti et al., 1993). While a relationship has been found between sedative/hypnotic use among women (Campbell et al., 1989), many studies have failed to find a relation between falls and sedative/hypnotic use in male community-dwelling older adults (Studenski et al., 1994), ambulatory nursing home residents (Lipsitz et al., 1991), or community-dwelling psychogeriatric patients (Asada et al., 1994). The relation between alcohol use and falls has been unclear in the literature. Campbell et al. and Nevitt et al. (1989) found no relation between alcohol and falls in community-dwelling older adults, while Quail (1994) found alcohol to be related to falls among outpatients. Similar to research findings regarding antidepressants, disparate results may be attributable to different specific drugs within the classes. However, it is also possible that drugs are exacting different effects on different subjects.

Anxiolytics, of which the most common class is benzodiazepines, are utilized to reduce anxiety symptoms. Unfortunately, they are also addictive and can be fatal in high doses. No relation between benzodiazepines and falls in community-dwelling males (Studenski et al., 1994), residents of homes for the elderly (Graafmans et al., 1996), or discharged hospital patients (Mahoney et al., 1994) has been found. However, Aisen et al. (1992) found that PRN administration of benzodiazepines was two times more frequent in fallers. Moreover, in 24 of 33 falls among inpatient geropsychiatric patients, benzodiazepines were administered within 24 hours of the occurrence of a fall.

7.16.4.8.3 Neuroleptics/antipsychotics

Neuroleptics are sometimes referred to as the major tranquilizers due to their heavy sedating effect in high doses. Typically, they are prescribed for the treatment of psychotic disorders or agitation in elderly patients. They can result in extrapyramidal effects and blurred vision, both of which can increase the risk of falling. Brown and Edelstein (1995) found that the use of antipsychotic medications significantly increased the risk of falls among psychogeriatric inpatients. Lord, McLean, and Stathers (1992) found similar results with antipsychotic use being a significant risk factor for falls among community-dwelling older adults. Quail (1994) found similar results among outpatients with the use of major tranquilizers increasing the risk of falling by 50%. In contrast, Lipsitz et al. (1991) found no relation between antipsychotics and falls in ambulatory nursing home residents. Moreover, no relation was found between antipsychotics and falls among community-dwelling psychogeriatric patients (Asada et al., 1994) or discharged hospital patients (Mahoney et al., 1994).

7.16.4.8.4 Cardiovascular medications

Although a large portion of older adults are prescribed cardiovascular medications, very little research has been conducted on their relation to fall among older adults. Graafmans et al. (1996) is the only study which has found a positive relation between a type of cardiac medication, peripheral vasodilator, and the recurrence of falls. Cardiac medications were not associated with falls in nursing home or community-dwelling older adults (Robbins et al., 1989a, 1989b), psychogeriatric patients (Brown & Edelstein, 1995), or ambulatory nursing home patients (Lipsitz et al., 1991). Moreover, although antihypertensives led to severe postural symptoms among outpatients, they were not associated with falls (Graafmans et al.; Quail, 1994). Finally, diuretics, digitalis, and vasodilator were not significantly related to falls among long-term care patients (Luukinen et al., 1995).

7.16.4.8.5 Analgesics and anti-inflammatory medications

Medicines prescribed for the alleviation of pain or inflammation are often prescribed to older adults. In a study of ambulatory nursing home residents, fallers were significantly more likely than nonfallers to be taking analgesics (Lipsitz et al., 1991). However, the taking of analgesics was not found to be significantly related to falls among psychogeriatric inpatients (Brown & Edelstein, 1995) or residents of homes for the elderly (Graafmans et al., 1996). Ingestion of anti-inflammatory medications has not been found to be significantly related to falls among nursing home residents (Graffmans et al.; Lipsitz et al.).

The previous discussions regarding drugs demonstrate the need for research examining the relation between drugs and falls. Although some classes of drugs have been identified as potentially increasing the risk of falling, drug interactions may also account for increased fall rates. The majority of research on drugs has involved retrospective reviews. Unfortunately, very few subjects are prescribed only one drug. Therefore, differing results may be due to drug interactions. For example, a person taking an antidepressant may not be at increased risk of falling but an individual taking an antidepressant and a cardiovascular medication may be at increased risk. Medications have multiple effects on many systems within the body. It is likely that individual differences account for differential drug effects. The complex interplay of drugs and physiological systems requires more knowledge.

7.16.4.9 Medical Disorders

The relationship between the type of medical disorder and fall risk is unclear, as studies have yielded mixed results. Numerous studies have found relations between falls and underlying medical diseases (e.g., Campbell et al., 1990; Nevitt et al., 1989; Rubenstein et al., 1990). Rubenstein et al., following their review of several falls studies, estimated that 55% of falls are related to medical disorders, with the greatest percentage of these being due to chronic diseases. Among the disorders related to falls are arthritis of the knee, stroke (Campbell et al.,1989), Parkinson's disease (Davie, Blumenthal, & Robinson-Hawkins, 1981; Studenski et al, 1994), osteoarthritis, unspecified arthritis (Davie et al.), diabetes, and hypotension (Davie et al.; Lipsitz et al., 1991). Though considerable support has been found for relations between medical disorders and falls, several studies have failed to find this relation for some of the same disorders (e.g.,Lipsitz et al.; Morse et al., 1987; Studenski et al.). Brown and Edelstein (1995) noted that the only apparently consistent finding in this literature is the strong relation between musculoskeletal disorders and falling.

The mostly mixed and contradictory findings noted above are likely attributable to differences in numerous factors across studies (e.g., subject population, physical and mental status, age, nature and number of diseases, research procedures, length or severity of disorder) that can directly or indirectly influence reported fall rates. For example, Lipsitz et al. (1991) based their fall rate on falls witnessed by staff or self-reported by nursing home residents. In contrast, Davie et al. (1981) and Studenski et al. (1994) obtained data via telephone calls and interviews of community-dwelling adults. Davie et al. may have obtained results different from Lipsitz et al. and Morse et al. (1987) regarding falls among individuals with diabetes or hypotension because of participant differences. Davie et al. studied community-dwelling residents, whereas Lipsitz et al. and Morse et al. studied inpatients. Hospitalized older adults are generally at a greater risk of falling than community-dwelling adults, as they tend to be more cognitively impaired and less ambulatory. In conclusion, it may be naive to believe that the effects of individual medical conditions can be isolated and that findings across studies can be compared. It is not uncommon for older adults to simultaneously experience multiple medical disorders. When considering the problems of matching participants across studies on the many factors believed to be related to falls, the prospect of obtaining reliable results appears improbable.

The relationship between the number of medical disorders and fall risk also remains unclear. Studies by Kalchthaler et al. (1978) and Morse et al. (1987) found a positive relationship between number of medical disorders and falls.

Kalchthaler et al. found the largest number of falls among nursing home residents with three to six medical disorders. Similarly, Morse et al. found fallers had significantly more medical disorders than nonfallers. Although the medical disorder in fall literature is fraught with inconsistencies in methodologies and results, there is certainly sufficient support for the contention that medical disorders, particularly those of a chronic nature, place older adults at risk of falls.

7.16.4.10 Psychiatric Disorders

The effects of psychiatric disorders are difficult to separate from their treatment (e.g., psychotropic medications), the overall mental status of the participants, and the setting in which falls are examined. Institutionalized individuals are generally more likely to suffer from psychiatric disorders, be receiving psychotropic medications, and have poorer mental status than individuals still living in the community. In addition, they are more likely to fall. A small number of researchers have examined the relationship between psychiatric disorders and falls, some of which have found support for such a relationship. Hernandez and Miller (1986), for example, examined fallers on a psychogeriatric unit and found fallers to have the following psychiatric diagnoses: dementia (77%), depression (43%), and agitation or paranoia (21%). Similarly, Poster et al. (1991) found older adults in a psychiatric hospital with psychotic disorders to be at greater risk of falling than those without psychotic disorders. Granek et al. (1987) found significantly more depressed fallers than nonfallers in a long-term care facility.

Similar findings have been revealed among community-dwelling adults. Davie et al. (1981) found the presence of a psychiatric diagnosis to be related to falling among community-dwelling older adults. Similarly, Asada et al. (1994) found depressed community-dwelling individuals at a significantly greater fall risk than nondepressed individuals.

In contrast to the foregoing results, several studies have failed to find positive relations between psychiatric disorders and falls. For example, Hernandez and Miller (1986) and Granek et al. (1987) found no significant differences in the incidences of Parkinson's disease between fallers and nonfallers. Brown and Edelstein (1995) also found no relation between psychiatric disorders and falling. Moreover, there have been a few studies suggesting a possible protective effect of psychiatric disorders. For example, Poster et al. (1991) found patients with Bipolar Disorder to be at lower risk of falling than other patients. The only apparently consistent finding across several studies is the positive relation between depression and falling. Unfortunately, it cannot be determined whether this is primarily due to the disorder, cognitive deficits resulting from the disorder, the medications used to treat the disorder, or some combination of these and other factors that could serve to moderate this relationship.

In conclusion, there appear to be few intrinsic factors that reliably predict falls. However, a few factors have emerged as predictors in a sufficient number of studies to certainly warrant further investigation and exploration in prevention programs, particularly with regard to the type of intervention that is most effective for each particular risk factor.

7.16.5 PREVENTION OF FALLS

While the prevention of all falls is virtually impossible, numerous authors have offered suggestions for the development of prevention programs and limited empirical support has been found for some prevention practices. Recommendations for fall prevention have typically been based upon "common sense" (Gibson et al., 1987) and identified risk factors. However, Rubenstein and Josephson (1993) have aptly noted that while the literature is replete with such recommendations for preventing falls, few of the recommendations are based upon sound, controlled research. Similarly, Tinetti (1994) argued that fall prevention "studies reported in the literature to date have shown either no or only marginal reduction in fall rate" (p. 758). However, there has been evidence of successful prevention programming. In the following section approaches to prevention will be discussed and a few of the more promising studies published will be reviewed.

Falls can have multiple causes that vary across individuals. Thus, effective fall prevention programs must take into consideration a myriad of potential risk factors for each individual. Prevention programs might include risk factor identification, assessment following falls, exercise programs, alarm systems, and protective devices (Rubenstein & Josephson, 1993). Because the prevention programs often contain multiple interventions, it is difficult to compare studies in any meaningful way. Nevertheless, an attempt will be made to consider elements of prevention programs by following Tideiksaar's (1993a) categorization of prevention strategies, which include medical, rehabilitative, and environmental interventions.

7.16.5.1 Medical Prevention Strategies

Medical prevention strategies typically address risk factors through an initial risk factor assessment that might include mental status examination (to assess for delirium, dementia, depression), postural blood pressure measurement (to test for orthostatic hypotension), cardiac evaluation (to check for arrhythmias, valvular disorders), neurological evaluation, mobility assessment (e.g., ambulation, transfer from chair, bed or toilet), visual acuity testing, lower extremity muscle testing, balance testing (e.g., postural stress test), gait testing (e.g., Gait Abnormality Rating Scale), functional status evaluations, and a review of medications (Rubenstein & Josephson, 1993; Tideiksaar, 1993a). While many of these "conditions" cannot be prevented, many can be treated successfully, thereby reducing fall risk. In some cases the impact of these factors can be attenuated. For example, improved prescription practices (e.g., appropriate medication use, limiting number of medications, attention to drug interaction effects) can reduce fall rate. Several researchers have explored the effects of improving prescription practices. Goals included the reduction in amount of medication taken, reduction in the use of sedatives (particularly long-acting sedatives), and reduction in the use of tranquilizers. Rane-Szostak (cited in Gibson et al., 1987) attempted to reduce the number of falls in a large rehabilitation-oriented nursing home by asking physicians to reduce the use of sedatives (particularly long-acting sedatives) and tranquilizers. Nurses offered alternatives to medication for sleep, including changes in bedtime routines, using mild analgesics for pain, and providing evening snacks. A 70% reduction in the number of fall-related fractures has been attributed to the decreased use of sedatives and tranquilizers.

Once a fall has occurred, a detailed description of the circumstances surrounding the fall is recommended. The medically-oriented questions address such factors as loss of balance, dizziness, and other medically-related symptoms that have preceded the fall. The results of these inquiries can then be used in the development of prevention programs.

Tideiksaar (1993a) also suggested that these same questions following near-falls be explored. A near-fall involves a loss of balance without the occurrence of a fall due to gaining support from the environment (e.g., grabbing a hand rail). Near-falls are important to explore as they appear to portend future falls and may involve many of the same risk factors. Knowledge of these factors can be incorporated into prevention programs that could preclude or at least reduce the likelihood of future falls.

7.16.5.2 Rehabilitative Strategies

Fall risk can potentially be reduced through a variety of rehabilitative strategies. These might include muscle strengthening exercises, balance and gait training, and training in the proper use of ambulatory assistive devices. As previously noted, the relationship between lower extremity muscular strength and falls is quite strong. Weakened muscular strength can affect gait and balance, all of which can contribute to falling and fear of falling. Numerous studies have focused on increasing muscle strength with the goal of improving joint flexibility, gait, and balance (e.g., Fiatarone et al., 1994; Province et al., 1995). Such programs can include resistance exercises, weight-lifting, and leg stretching exercises. Flexibility and balance training, and even Tai Chi exercises, have shown positive results through improved gait and balance (e.g., Province et al.). Fiatarone et al. increased muscle strength by 113%, yielding a 12% increase in walking speed and a 28% increase in stair climbing ability through progressive resistance training among frail nursing home residents with a mean age of 87 years. Promising results have also been reported by Wolf, Kutner, Green, and McNeely (1993), who conducted a prospective, randomized controlled clinical trial comparing the effects of Tai Chi and computerized balance training with community-dwelling older adults. By 15 weeks, participants in the Tai Chi course reduced their fall rates by 47%, although the risk of falling was actually higher for those Tai Chi participants that experienced falls within the previous year. Fear of falling also decreased in the Tai Chi group.

7.16.5.3 Environmental Strategies

Environmental approaches to reduce fall risk attempt to do so by modifying the physical environment. Falls occur when there is a mismatch between an individual's performance level and the performance demands of environmental tasks (Hogue, 1992). Although physically fit individuals can lose their balance and fall, very frail individuals are probably the most dependent upon the environment to support the performance of tasks (Hogue, 1982, 1991; Waller, 1978). Frail individuals "may be unable to maintain postural control (prevent a fall) with virtually no external provocation" (Hogue, 1992, p. 45). The physical environment can often be modified to accommodate individuals whose abilities do not match the current environment. For example, higher seated chairs and toilets can replace lower seated ones to reduce the demands of transfer. Lighting can be intensified to reduce the risk of falls resulting from poorly illuminated areas, nonskid floor surfaces can replace slippery surfaces, stairways can be equipped with handrails for support, grab bars can be placed near toilets, bathtubs, and showers for support during transfers, and objects situated around the living environment that block or obstruct walk areas can be moved.

Tideiksaar (1997) noted that there is no definitive evidence of the efficacy of these environmental modifications in reducing falls. However, a few studies have offered promising results. Tideiksaar (1990), as reported in Tideiksaar (1997), eliminated various environmental hazards in the homes of older adults with multiple slip and trip falls. Slippery and irregular ground surfaces in the home were modified resulting in a decrease in number of falls over the following year.

Taking a more comprehensive approach, Wolf-Klein et al. (1988) provided medical management, home visits, and a falls prevention educational program for the participants in their study and their families, yielding a significant decrease in falls over the following year. Such comprehensive approaches have also been successful in institutional settings (e.g., Heslin et al., 1992; Hollinger & Patterson, 1992). Heslin et al. identified fall risk factors in acute, long-term, and residential health care settings and developed prevention strategies for each setting. Participants included 581 acute, 138 long-term care, and 247 residential care patients in a 966-bed teaching hospital. Fall rates at the beginning of the study were 2.16% for acute care, 1.46% for long-term care, and 9.38% for residential care. All but 17% of the fallers were 60 years of age or older. Prevention interventions targeted specific care settings and types of falls. Prevention in the acute care setting was directed at falls related to beds. The program included a fall risk assessment, nursing care plan, spot-the-dot (color-coded markers placed in risk areas), and bed sensors to alert staff to movement from beds. The number of bed-based falls decreased significantly (85%) over a three-month period using the sensor system. The spot-the-dot procedure was completely successful, with no bed-based falls following initial adjustments to the program on a medical unit. The long-term care interventions yielded similarly successful outcomes, with bed falls and transfers being targeted for interventions. When Huntington's Chorea patients who had experienced falls from bed had their restraints removed, their bed rails padded with cushions, and the bed surrounded with mattresses, injuries were reduced to zero. Bed to chair transfers, which account for approximately 50% of bed-based falls in the long-term care setting, were reduced by 92% through changes in transfer procedures and transfer devices.

Hollinger and Patterson (1992) developed and evaluated a falls precaution assessment and prevention program in an acute care setting. All participants were evaluated initially with a Fall Assessment Tool that addressed patient and environmental factors known to be risk factors for falling. Individualized prevention programs were then developed for individual patients based upon the findings with the Fall Assessment Tool. These programs included such precautions as use of assistive devices, night lights, and nonskid footwear. The fall rate over a six-month period was reduced significantly from a preintervention mean of 4.06 to a postintervention mean of 2.63.

7.16.5.4 Combined Prevention Strategies

Many of the reported prevention studies have employed a variety of intervention methods that involved medical, environmental, and rehabilitative approaches or some combination of two or more of these. For example, Tinetti et al. (1994) examined the effects of three different interventions with 301 men and women living in the community. All participants were at least 70 years of age and had at least one of the following risk factors: "postural hypotension; use of sedatives; use of at least four prescription medications; and impairment in arm or leg strength or range of motion, balance, ability to move safely from bed to chair or to the bathtub or toilet (transfer skills), or gait" (Tinetti et al., 1994, p. 821). The interventions were either usual health care or the combination of medication adjustment, behavioral change recommendations, and home-based exercise programs. Participants were followed for one year, during which 35% of the intervention group and 47% of the control group fell. The authors also demonstrated significant reduction in risk factors among the intervention group.

Although a broad range of prevention programs have been implemented, they have yielded mixed results. As Tinetti (1994) noted, the handful of clinical trials conducted prior to 1994 had somewhat discouraging results (see Hornbrook et al., 1994; Mulrow et al., 1994; Reinsch, MacRae, Lachenbruch, & Tobis, 1992; Rubenstein et al., 1990; Rubenstein, Robbins, & Josephson, 1991; Vetter, Lewis, & Ford, 1992). The recent "Frailty and Injuries:

Cooperative Studies of Intervention Techniques" (FICSIT) trials, however, offer more promising data from a large series of clinical trials (e.g., Tinetti, Baker et al., 1994; Wolf et al., 1993). Moreover, if empirical studies that do not meet the criteria for "prospective, controlled trials" are considered, the evidence in support of prevention programs is even more encouraging.

Perhaps what is most discouraging about the prevention literature is the failure to identify interventions that reliably reduce fall frequency or severity. Researchers may be at the point in prevention research that they must match interventions to the participants, with the hope of establishing which interventions are most effective with participants with particular characteristics and risk factors. Similarly, more effective prevention programs may be developed by considering such factors as whether the potential fallers are at little or great risk of falling, with different preventive approaches being most effective with one or the other of these two groups (Gibson et al., 1987). Perhaps attention should be focused on the prevention of particular types of falls. Tinetti (1994) has raised the provocative question of whether the goal of prevention programs should be the reduction in risk of all falls or the reduction in risk of injurious falls.

In conclusion, there is reason for optimism about the prevention of falls. Researchers have identified factors that place older adults at risk, have demonstrated that some of these factors can be changed through a variety of interventions, and have begun to demonstrate that some prevention programs can be effective with some individuals. The optimal approach for individuals with particular characteristics experiencing particular types of falls remains to be determined.

7.16.6 SUMMARY

This chapter has provided an overview of the empirical literature on falls among older adults. It began with considerations of conceptual and methodological issues, including how falls are conceptualized and measured. The many consequences of falls, including physical, medical, psychological, and economic consequences, were then discussed. Numerous intrinsic and extrinsic risk factors for falls were then considered, followed by a discussion of prevention approaches based upon these risk factors. Finally, the chapter concluded with a cautiously optimistic view of the prevention literature, suggesting that future preventive efforts focus on optimal matches of prevention programs to individual faller characteristics.

7.16.7 REFERENCES

Aisen, P. S., Deluca, T., & Lawlor, B. A. (1992). Falls among geropsychiatry inpatients are associated with PRN medications for agitation. *International Journal of Geriatric Psychiatry, 7,* 709–712.

Arfken, C. L., Lach, H. W., Birge, S. J., & Miller, J. P. (1994). The prevalence and correlates of fear of falling in elderly persons living in the community. *American Journal of Public Health, 84,* 565–570.

Asada, T., Kariva, T., Kitajima, E., Kakuma, T., & Yoshioka, M. (1994). Falls among community-dwelling psychogeriatric patients. *International Journal of Geriatric Psychiatry, 9,* 17–23.

Baker, S. P., & Harvey, A. H. (1985). Fall injuries in the elderly. *Clinics in Geriatric Medicine, 1,* 501–512.

Brocklehurst, J. C., Robertson, D., & James-Groom, P. (1982). Clinical correlates of sway in old age-sensory modalities. *Age and Aging, 11,* 1–10.

Brown, S. A., & Edelstein, B. A. (1995, November). *Factors associated with falls among psychogeriatric patients.* Paper presented at meeting of the Association for the Advancement of Behavior Therapy.

Campbell, A. J., Reinken, J., Allan, B. C., & Martinez, G. S. (1981). Falls in old age: A study of frequency and related clinical features. *Age and Ageing, 10,* 264–270.

Campbell, A. J., Borrie, M. J., & Spears, G. F. (1989). Risk factors for falls in a community-based prospective study of people 70 years and older. *Journal of Gerontology: Medical Sciences, 44,* M112–117.

Campbell, A. J., Borrie, M. J., Spears, G. F., Jackson, S. L., Brown, J. S., & Fitzgerald, J. L. (1990). Circumstances and consequences of falls experienced by a community population 70 years and over during a prospective study. *Age and Ageing, 19* (2), 136–141.

Commodore, D. I. B. (1995). Falls in the elderly population: A look at incidence, risks, healthcare costs, and preventive strategies. *Rehabilitation Nursing, 20,* 84–89.

Cooper, J. W. (1994). Falls and fractures in nursing home patients receiving psychotropic drugs. *International Journal of Geriatric Psychiatry, 9,* 975–980.

Cummings, S. R., Kelsey, J., Nevitt, M., & Dowd, K. (1985). Epidemiology of osteoporosis and osteoporotic fractures. *Epidemiologic Reviews, 7,* 178–208.

Cummings, S. R., Nevitt, M. C., & Kidd, S. (1988). Forgetting falls: The limited accuracy of recall of falls in the elderly. *Journal of the American Geriatrics Society, 36,* 613–616.

Cwikel, J., Kaplan, G., & Barell, V. (1990). Falls and subjective health rating among the elderly: Evidence from two Israeli samples. *Social Science and Medicine, 31,* 485–490.

Davie, J. W., Blumenthal, M. D., & Robinson-Hawkins, S. (1981). A model of risk of falling for psychogeriatric patients. *Archives of General Psychiatry, 38,* 463–467.

Downton, J. H., & Andrews, K. (1990). Postural disturbance and psychological symptoms amongst elderly people living at home. *International Journal of Geriatric Psychiatry, 5,* 93–98.

Duncan, P. W., & Studenski, S. (1994). Balance and gait measures. In M. P. Lawton & J. A. Teresi (Eds.), *Annual review of gerontology and geriatrics: Focus on assessment techniques, 14* (pp. 76–92). New York: Springer.

Ejaz, F. K., Jones, J. A., & Rose, M. S. (1994). Falls among nursing home residents: An examination of incident reports before and after restraint reduction programs. *Journal of the American Geriatrics Society, 42,* 960–964.

Evans, J. G., Prudham, D. A., & Wandless, I. (1979). A prospective study of fractured proximal femur: Incidence and outcome. *Public Health, 93*, 235–241.

Evans, L. K. & Strumpf, N. E. (1990). Myths about elder restraint. *Image: Journal of Nursing Scholarship, 22,* 124–128.

Felson, D. T., Anderson, J. J., Hannan, M. T. , Milton, R. C., Wilson, P. W. F., & Keil, D. P. (1989). Impaired vision and hip fracture: The Framingham study. *Journal of the American Geriatrics Society, 37,* 495–500.

Fiatarone, M. A., O'Neill, E. F., Ryan, N. D., Clements, K. M., Solares, G. R., Nelson, M. E., Roberts, S. B., Kehayias, J. J., Lipsitz, L. A., & Evans, W. J. (1994). Exercise training and nutritional supplementation for physical frailty in very elderly people. *New England Journal of Medicine, 330,* 1769–1775.

Funk, S. G., Tornquist, E. M., Champagne, M. T., & Wiese, R. A. (1992). *Key aspects of elder care: Managing falls, incontinence, and cognitive impairment*. New York: Springer.

Gibson, M. J., Andres, R. O., Isaacs, B., Radebaugh, T., & Worm-Peterson, J. (1987). The prevention of falls in later life. *Danish Medical Bulletin, 34* (Suppl. 4), 1–22.

Graafmans, W. C., Ooms. M. E., Hofstee, H. M. A., Bezemer, P. D., Bouter, L. M., & Lips, P. (1996). Falls in the elderly: A prospective study of risk factors and risk profiles. *American Journal of Epidemiology, 143,* 1129–1136.

Granek, E., Baker, S. P., Abbey, H., Robinson, E., Myers, A. H., Samkoff, J. S., & Klein, L. E. (1987). Medications and diagnoses in relation to falls in a long-term care facility. *Journal of the American Geriatrics Society, 35,* 503–511.

Greenhouse, A. H. (1994). Falls among the elderly. In M. L. Albert & J. E. Knoefel (Eds.). *Clinical neurology of aging* (2nd ed., pp. 611–626). New York: Oxford University Press.

Gross, Y. T., Shimamoto, Y., Rose, C. L. & Frank, B. (1990). Why do they fall?: Monitoring risk factors in nursing homes. *Journal of Gerontological Nursing, 16,* 20–25.

Gryfe, C. I., Amies, A., & Ashley, M. J. (1977). A longitudinal study of falls in an elderly population: I. Incidence and morbidity. *Age and Ageing, 6,* 201–210.

Guimaraes, R.M., & Isaacs, B. (1980). Characteristics of the gait in old people who fall. *International Rehabilitation Medicine, 2,* 177–180.

Hernandez, M., & Miller, J. (1986). How to reduce falls. *Geriatric Nursing, 7,* 97–102.

Heslin, K., Towers, J., Leckie, C., Thornton-Lawrence, H., Perkin, K., Jacques, M., Mullin, J., & Wick, L. (1992). Managing falls: Identifying population-specific risk factors and prevention strategies. In S. G. Funk, E. M. Tornquist, M. T. Champagne, & R. A. Wiese (Eds.), *Key aspects of elder care: Managing falls, incontinence, and cognitive impairment* (pp. 70–88). New York: Springer.

Hogue, C. C. (1982). Injury in late life. Part II: Prevention. *Journal of the American Geriatrics Society, 30,* 276–280.

Hogue, C. C. (1991). A person-environment model for understanding fall risk. In R. Weindruch & M. Ory (Eds.), *Frailty reconsidered: Reducing frailty and fall-related injuries in the elderly* (pp. 96–105). Springfield, IL: Charles C. Thomas.

Hogue, C. C. (1992). Managing falls: The current bases for practice. In S. G. Funk, E. M. Tornquist, M. T. Champagne, & R. A. Wiese (Eds.), *Key aspects of elder care: Managing falls, incontinence, and cognitive impairment* (pp. 41–57). New York: Springer.

Hollinger, L. M., & Patterson, P. A. (1992). A falls prevention program for the acute care setting. In S. G. Funk, E. M. Tornquist, M. T. Champagne, & R. A. Wiese (Eds.), *Key aspects of elder care: Managing falls, incontinence, and cognitive impairment* (pp. 110–117). New York: Springer.

Hornbrook, M. C., Stevens, V. J., Wingfield, D. J., Hollis, J. F., Greenlick, M. R., & Ory, M. G. (1994). Preventing falls among community-dwelling older persons: Results from a randomized trial. *The Gerontologist, 34,* 16–23.

Howland, J., Peterson, E. W., Levin, W. C., & Fried, L. (1993). Fear of falling among the community dwelling elderly. *Journal of Aging and Health, 5,* 229–243.

Kalchthaler, T., Bascon, R. A., & Quintos, V. (1978). Falls in the institutionalized elderly. *Journal of the American Geriatrics Society, 26,* 424–428.

Kanten, D.N., Mulrow, C. D., Gerety, M. B., Lichtenstein, M. J., Aguilar, C., & Cornell, J. E. (1993). Falls: An examination of three reporting methods in nursing homes. *Journal of the American Geriatrics Society, 41,* 662–666.

Kay, A. D., & Tideiksaar, R. (1990). Falls and gait disorders. In W. B. Abrams & R. Berkow (Eds.). *Merck manual of geriatrics* (pp. 52–68). Rahway, NJ: Merck Sharp & Dohme Research Laboratories.

Kreutzfeldt, J., Haim, M., & Bach, E. (1984). Hip fracture among the elderly in a mixed urban and rural population. *Age and Ageing, 13,* 111–119.

Lichtenstein, M. J., Burger, M. C., Shields, S. L., & Shiavi, R. G. (1990). Comparison of biomechanics platform measures of balance and videotaped measures of gait with a clinical mobility scale in elderly women. *Journal of Gerontology, 45,* M49–M54.

Lipsitz, L. A., Jonsson, P. V., Kelley, M. M., & Koestner, J. S. (1991). Causes and correlates of recurrent falls in ambulatory frail elderly. *Journal of Gerontology, 46,* M1114–M1122.

Lord, S. R. (1996). Instability and falls in elderly people. In C. Lafont, A. Baroni, M. Allard, R. Tideiksaar, B.J. Vellas, P. J. Garry, & J. L. Albarede (Eds.). *Falls, gait, and balance disorders in the elderly: From successfull aging to frailty*. New York: Springer.

Lord, S. R., Clark, R. D., & Webster, I. W. (1991). Physiological factors associated with falls in an elderly population. *Journal of the American Geriatrics Society, 39,* 1194–1200.

Lord, S. R., McLean, D., & Stathers, G. (1992). Physiological factors associated with injurious falls in older people living in the community. *Gerontology, 38,* 338–346.

Luukinen, H., Koski, K., Laippala, P., & Kivela, S. L. (1995). Risk factors recurrent falls in the elderly in long-term institutional care. *Public Health, 109,* 57–65.

Mahoney, J., Sager, M., Dunham, N. C., & Johnson, J. (1994). Risk of falls after hospital discharge. *Journal of the American Geriatrics Society, 42,* 269–274.

Marx, M. S., Cohen-Mansfield, J., & Werner, P. (1990). Agitation and falls in institutionalized elderly persons. *Journal of Applied Gerontology, 9,* 106–117.

Melton, L. J., & Riggs, B. L. (1985). Risk factors for injury after a fall. *Clinics in Geriatric Medicine, 1,* 525–539.

Mitchell, R. G. (1984). Falls in the elderly. *Nursing Times, January 11–17,* 51–53.

Morris, J. C., Rubin, E. H., Morris, E. J., & Mandel, S. A. (1987). Senile dementia of the Alzheimer's type: An important risk factor for serious falls. *Journal of Gerontology, 42,* 412–417.

Morse, J. M., Morse, R. M., & Tylko, S. J. (1989). Development of a scale to identify the fall-prone patient. *Canadian Journal on Aging, 8,* 366–377.

Morse, J. M., Tylko, S. J., & Dixon, H. A. (1987). Characteristics of the fall-prone patient. *Gerontologist, 27,* 516–522.

Mosenthal, A. C., Livingston, D. H., Elcavage, J., Merritt, S., & Stucker, S. (1995). Falls: Epidemiology and strategies for prevention. *Journal of Trauma: Injury, Infection, and Critical Care, 38,* 753–756.

Mulrow, C. D., Gerety, M. B., Kanten, D., Cornell, J. E., DeNino, L. A., Chiodo, L., Aguilar, C., O'Neil, M. B., Rosenberg, J., & Solis, R. M. (1994). A randomized trial of physical rehabilitation for very frail nursing home residents. *Journal of the American Medical Association, 271,* 519–524.

Murphy, J., & Isaacs, B. (1982). The post-fall syndrome: A study of 36 elderly patients. *Gerontology, 28,* 265–270.

National Institutes of Health (1991). *Physical frailty: a reducible barrier to independence for older Americans* (NIH Publication No. 91–37). Washington, DC: National Institutes of Health.

Nevitt, M. C., Cummings, S. R., & Hudes, E. S. (1991).Risk factors for injurious falls: A prospective study. *Journal of Gerontology: Medical Sciences, 46,* M164–M170.

Nevitt, M. C., Cummings, S. R., Kidd, S., & Black, D. (1989). Risk factors for recurrent nonsyncopal falls: A prospective study. *Journal of the American Medical Association, 261,* 2663–2668.

Null-Kogan, J. (1996). *Self-reported fear in middle-aged and older adults and its relation to daily functioning and satisfaction with life*. Unpublished master's thesis, West Virginia University, Morgantown, WV.

Ousset, P.J., Hennia, A. O., Faisant, C., Sedeuilh, M., Reyes, G., Guiterez, L., Allard, M., Lafont, C., Vellas, B., & Albarede, J. L. (1996). A two years longitudinal study of falls, gait, and balance disorders in 267 healthy elderly persons living in the community: The ICARE study. In C. Lafont, A. Baroni, M. Allard, R. Tideiksaar, B. J. Vellas, P. J. Garry, & J. L. Albarede (Eds.), *Falls, gait and balance disorders in the elderly: From successful aging to frailty*. Paris: Serdi Publishing Company.

Overstall, P. W. (1992). Falls. *Reviews in Clinical Gerontology, 2,* 31–38.

Overstall, P. W., Exton-Smith, F. J., & Johnson, A. L. (1977). Falls in the elderly related to postural imbalance. *British Medical Journal, 1,* 261–264.

Pawlson, L. G., Goodwin, M., & Keith, K. (1986). Wheelchair use by ambulatory nursing home residents. *Journal of the American Geriatrics Society, 34,* 860–864.

Poster, E. C., Pelletier, L. R., & Kay, K. (1991). A retrospective cohort study of falls in a psychiatric inpatient setting. *Hospital and Community Psychiatry, 42,* 714–720.

Province, M. A., Hadley, E. C., Hornbrook, M. C., Lipsitz, L. A., Miller, J. P., Mulrow, C. D., Ory, M. G., Sattin, P. W., Tinetti, M. E., Wolf, S. L., & FICSIT Group (1995). The effects of exercise on falls in elderly patients: A preplanned meta-analysis of the FICSIT trials. *Journal of the American Medical Association, 273,* 1341–1347.

Prudham, D., & Evans, J. (1981). Factors associated with falls in the elderly: A community study. *Age and Ageing, 10,* 141–146.

Quail, G. G. (1994). An Approach to the assessment of falls in the elderly. *Australian Family Physician, 23,* 873–882.

Reinsch, S., MacRae, P., Lachenbruch, P. A., & Tobis, J. S. (1992). Attempts to prevent falls and injury: A prospective community study. *The Gerontologist, 32,* 450–456.

Robbins, A. S., Rubenstein, L. Z., Josephson, K. R., Schulman, B. L., Osterwell, D., & Fine, G. (1989a). Predictors of falls among elderly people: Results of two population-based studies. *Archives of Internal Medicine, 149,* 1628–1633.

Robbins, A. S., Rubenstein, L. Z., Josephson, K. R., Schulman, B. L., Osterwell, D., & Fine, G. (1989b). Predictors of falls among elderly people. *American Journal of Psychiatry, 143,* 1446–1449.

Rosen, A. M., Campbell, R. J., Villanueva, J., & Morgan, K. (1985). Factors affecting falling by older psychiatric inpatients. *Psychosomatics, 26,* 117–123.

Rubenstein, L. Z., & Josephson, K. R. (1993). Clinical research on falls in the nursing home. In L. Z. Rubenstein & D. Wieland (Eds.), *Improving care in the nursing home: Comprehensive reviews of clinical research* (pp. 216–240). Newbury Park,CA: Sage.

Rubenstein, L. Z., Josephson, K. R., & Robbins, A. S. (1994). Falls in the nursing home. *Annals of Internal Medicine, 121,* 442–451.

Rubenstein, L., Robbins, A. S., & Josephson, K. (1991). Falls in the nursing-home setting: Causes and preventive approaches. In P. R. Katz, R. L. Kane, & M. D. Mezey (Eds.), *Advances in long-term care* (Vol. 1, pp. 28–42.). New York: Springer.

Rubenstein, L. Z., Robbins, A. S., Josephson, K. R., Schulman, B. L., & Osterweil, D. (1990). The value of assessing falls in an elderly population. *Annals of Internal Medicine, 113,* 308–316.

Rubenstein, L. Z., Robbins, A. S., Schulman, B. L., Rosado, J., Osterweil, D., & Josephson, K. R. (1988). Falls and instability in the elderly. *Journal of the American Geriatrics Society, 36,* 266–278.

Saywell, R. M., Woods, J. R., Rappaport, S. A., & Allen, T. L. (1989). The value of age and severity as predictors of costs in geriatric head trauma patients. *Journal of the American Geriatrics Society, 37,* 625–630.

Simoneau, G. G., & Leibowitz, H. W. (1996). Posture, gait, and falls. In J. E. Birren, K. Warner Schaie, R. P. Abeles, M. Gatz, & T. A. Salthouse (Eds.), *Handbook of the psychology of aging* (4th ed. pp. 204–217). New York: Academic Press.

Simpson, J. M., & Mandelstam, H. (1995). Elderly people at risk of falling: Do they want to be taught how to get up again? *Clinical Rehabilitation, 9,* 65–69.

Soja, M. E., Kippenbrock, T. A., Hendrich, A. L., & Nyhuis, A. (1992). A risk model for patient fall prevention. In S. G. Funk, E. M. Tornquist, M. T. Champagne, & R. A. Wiese (Eds.), *Key aspects of elder care: Managing falls, incontinence, and cognitive impairment*. New York: Springer.

Speechley, M., & Tinetti, M. E. (1991). Falls and injuries in frail and vigorous community elderly persons. *Journal of the American Geriatrics Society, 39,* 46–52.

Studenski, S., Duncan, P. W., Chandler, J., Samsa, G., Prescott, B., Hogue, C., & Bearon, L. B. (1994). Predicting falls: The role of mobility and nonphysical factors. *Journal of the American Geriatrics Society, 42,* 297–302.

Sudarsky, L. (1990). Geriatrics: Gait disorders in the elderly. *New England Journal of Medicine, 322,* 1441–1446.

Tideiksaar, R. (1989). *Falling in old age: Its prevention and treatment*. New York: Springer.

Tideiksaar, R. (1990). The biomechanical and environmental characteristics of slips, stumbles, and falls. In B. E. Gray (Ed.), *Slips, stumbles, and falls: Pedestrian footwear and surfaces* (pp. 17–27). Philadelphia: American Society for Testing and Materials.

Tideiksaar, R. (1993a). Falls in older persons. *The Mount Sinai Journal of Medicine, 60,* 515–521.

Tideiksaar, R. (1993b). Mechanisms of falls in community residing older persons. *Pride Institute Journal of Long Term Home Health Care, 12,* 3–11.

Tideiksaar, R. (1993c). *Falls in older persons: Prevention and management in hospitals and nursing homes*. Boulder, CO: Tactilitics, Inc.

Tideiksaar, R. (1997). *Falling in old age: Prevention and management* (2nd ed.). New York: Springer.

Tideiksaar, R., & Kay, A. (1986). What causes falls? A logical diagnostic procedure. *Geriatrics, 41,* 32–50.

Tinetti, M. E. (1987). Factors associated with serious injury during falls by ambulatory nursing home residents. *Journal of the American Geriatrics Society, 35,* 644–648.

Tinetti, M. E. (1994). Prevention of falls and fall injuries in elderly persons: A research agenda. *Preventive Medicine, 23,* 756–762.

Tinetti, M. E., Baker, D. I., McAvay, G., Claus, E. B., Garrett, P., Gottschalk, M., Koch, M. L., Trainor, K., & Horwitz, R. I. (1994). A multifactorial intervention to reduce the risk of falling among elderly people living in the community. *New England Journal of Medicine, 331,* 821–827.

Tinetti, M. E., Liu, W. L., & Claus, E. B. (1993). Predictors and prognosis of inability to get up after falls among elderly persons. *Journal of the American Medical Association, 269,* 65–70.

Tinetti, M. E., Mendes de Leon, C. F., Doucette, J. T., & Baker, D. I. (1994). Fear of falling and fall-related efficacy in relationship to functioning among community-living elders. *Journal of Gerontology: Medical Sciences. 49,* M140–M147.

Tinetti, M. E., Richman, D., & Powell, L. (1990). Falls efficacy as a measure of fear of falling. *Journal of Gerontology: Psychological Sciences, 45*(6), P239–P243.

Tinetti, M. E., & Speechley, M. (1989). Current concepts in geriatrics: Prevention of falls among the elderly. *New England Journal of Medicine, 320,* 1055–1059.

Tinetti, M. E., Speechley, M., & Ginter, S. F. (1988). Risk factors for falls among elderly persons living in the community. *The New England Journal of Medicine, 319,* 1701–1707.

Tinetti, M. E., Williams, T. F., & Mayewski, R. (1986). Fall risk index for elderly patients based on number of chronic disabilities. *American Journal of Medicine, 80,* 429–434.

Urton, M. M. (1991). A community home inspection approach to preventing falls among the elderly. *Public Health Reports, 106,* 192–195.

Van Dijk, P. T., Meulenberg, O. G., Van de Sande, H. J., & Habbema, J. D. (1993). Falls in dementia patients. *Gerontologist, 33,* 200–204.

Vetter, N. J., Lewis, P. A., & Ford, D. (1992). Can health visitors prevent fractures in elderly people? *British Medical Journal, 304,* 888–890.

Walker, J. E., & Howland, J. (1991). Falls and fear of falling among elderly persons living in the community: Occupational therapy interventions. *American Journal of Occupational Therapy, 45,* 119–122.

Waller, J. A. (1978). Falls among the elderly: Human and environmental factors. *Accident Analysis and Prevention, 10,* 21–23.

Whipple, R. H., Wolfson, L. I., & Amerman, P. M. (1987). The relationship of knee and ankle weakness to falls in nursing home residents: An isokinetic study. *Journal of the American Geriatrics Society, 35,* 13–20.

Whipple, R., Wolfson, L., Derby, C., Singh, D., & Tobin, J. (1993). Altered sensory functions and balance in older persons. *The Journals of Gerontology, 48* (Special Issue), 71–76.

Whitbourne, S. K. (in press). Physiological aspects of aging: Relation to identity and clinical implications. In B. Edelstein (Ed.), *Clinical geropsychology*. Oxford: Elsevier.

Wild, D., Nayak, U. S., & Isaacs, B. (1981). Prognosis of falls in old people at home. *Journal of Epidemiology and Community Health, 35,* 200–204.

Wofford, J. L., Heuser, M. D., Moran, W. P., Schwartz, E., & Mittelmark, M. B. (1994). Community surveillance of falls among the elderly using computerized EMS transport data. *American Journal of Emergency Medicine, 12,* 433–437.

Wolf, S. L., Barnhart, H. X., Kutner, N. G., McNeely, E., Coogler, C., & Xu, T. (1996). Reducing frailty and falls in older persons: An investigation of Tai Chi and computerized balance training. *Journal of the American Geriatrics Society, 44,* 489–497.

Wolf, S. L., Kutner, N. C., Green, R. C., & McNeely, E. (1993). The Atlanta FICSIT Study: Two exercise interventions to reduce frailty in elders. *Journal of the American Geriatrics Society, 41,* 329–332.

Wolf-Klein, G. P., Silverstone, F. A., Basavaraju, N., Foley, C. T., Pascaru, A., & Ma, P. H. (1988). Prevention of falls in the elderly population. *Archives of Physical Medicine and Rehabilitation, 69,* 689–691.

Wolfson, L., Judge, J., Whipple, R., & King, M. (1995). Strength is a major factor in balance, gait and the occurrence of falls. *The Journals of Gerontology, 50a* (Special Issue), 64–67.

Wolfson, L. I., Whipple, R., Amerman, P., & Kleinberg, A. (1986). Stressing the postural response: A quantitative method for testing balance. *Journal of the American Geriatrics Society, 34,* 845–850.

Wolfson, L. I., Whipple, R., Amerman, P., & Tobin, J. N. (1990).Gait assessment in the elderly: A gait abnormality rating scale and its relation to falls. *Journal of Gerontology, 45,* M12–M19.

Wolinsky, F. D., Johnson, R. J., & Fitzgerald, J. F. (1992). Falling, health status, and the use of health services by older adults: A prospective study. *Medical Care, 30,* 587–597.

Wright, B. A., Aizenstein, S., Vogler, G., Rowe, M., & Miller, C. (1990). Frequent fallers: Leading groups to identify psychological factors. *Journals of Gerontological Nursing, 16,* 15–19.

Yasumura, S., Arai, H., Haga, H., Niino, N., Nagai, H., Suzuki, T., Amano, H., & Shibata, H. (1996). Risk factors for falls among elderly people living in an urban community in Japan: A longitudinal study. In C. Lafont, A. Baroni, M. Allard, R. Tideiksaar, B. J. Vellas, P. J. Garry, & J. L. Albarede (Eds.), *Falls, gait and balance disorders in the elderly*. Paris: Serdi Publishing Company.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.17
Wandering

DONNA L. ALGASE
University of Michigan, Ann Arbor, MI, USA

7.17.1 WANDERING: A MODEL CASE 372

7.17.2 DEFINITIONS AND TYPOLOGIES OF WANDERING 372
- *7.17.2.1 Clinical Views* 372
- *7.17.2.2 Conceptual Schemes and Typologies* 373
- *7.17.2.3 Empirically-based Descriptions and Features* 374

7.17.3 EPIDEMIOLOGICAL DATA 376
- *7.17.3.1 Prevalence* 376
- *7.17.3.2 Incidence* 376

7.17.4 PUTATIVE RISK FACTORS FOR WANDERING 376
- *7.17.4.1 Personal Factors* 376
- *7.17.4.2 Potential Environmental Contributors* 377

7.17.5 PUTATIVE OUTCOMES OF WANDERING 379
- *7.17.5.1 Falls* 379
- *7.17.5.2 Elopements* 379
- *7.17.5.3 Weight Loss* 380
- *7.17.5.4 Sleep Disturbance and Fatigue* 380
- *7.17.5.5 Other Possible Outcomes* 380

7.17.6 ETIOLOGIC THEORIES AND FACTORS 383
- *7.17.6.1 Wandering as Spatial Disorientation* 383
- *7.17.6.2 Wandering as a Movement Disorder* 384
  - *7.17.6.2.1 The HPC and locomotor output* 384
  - *7.17.6.2.2 Perseveration* 385
  - *7.17.6.2.3 Effect of neurotransmitters* 385
- *7.17.6.3 Wandering as a Circadian Rhythm Disruption* 385
- *7.17.6.4 Wandering as an Expression of Mood or Affect* 386
- *7.17.6.5 An Emerging Multifactorial Explanation for Wandering* 386

7.17.7 ASSESSMENT TECHNIQUES AND ISSUES 389
- *7.17.7.1 Assessment Techniques for Wandering* 389
  - *7.17.7.1.1 Behavior mapping* 389
  - *7.17.7.1.2 Counting and timing strategies* 389
  - *7.17.7.1.3 Activity monitors* 391
- *7.17.7.2 Assessment for Associated Clinical Concerns* 391
  - *7.17.7.2.1 Foundational abilities* 391
  - *7.17.7.2.2 Mediating factors* 393
- *7.17.7.3 Assessment of Wandering Outcomes* 394
  - *7.17.7.3.1 Falls* 394
  - *7.17.7.3.2 Elopements* 395
  - *7.17.7.3.3 Social functioning* 395
  - *7.17.7.3.4 Weight loss and nutritional status* 396
  - *7.17.7.3.5 Fatigue and sleep disruption* 396

7.17.8 INTERVENTION STRATEGIES 397
- *7.17.8.1 Goals for Treatment and Intervention* 397
- *7.17.8.2 General Health and Safety Considerations* 397
- *7.17.8.3 Tested Techniques for Altering Wandering Behavior* 397

7.17.8.3.1 Exercise 397
7.17.8.3.2 Cueing 398
7.17.8.3.3 Lighting 398
7.17.8.3.4 Social interaction 398
7.17.8.3.5 Drug therapy 399
7.17.8.4 Tested Techniques for Addressing Undesired Outcomes 400
7.17.8.4.1 Behavior modification 400
7.17.8.4.2 Camouflage 400
7.17.8.4.3 Perimeter control devices 401
7.17.8.5 Wandering Paths or Tracks 402
7.17.8.6 Activities Programming 402
7.17.8.6.1 Physical and chemical restraint 402
7.17.9 RECOMMENDATIONS FOR FURTHER RESEARCH 404
7.17.9.1 Definition, Measurement, and Description 404
7.17.9.2 Etiologic Factors 405
7.17.9.3 Intervention Studies 405
7.17.10 SUMMARY 406
7.17.11 REFERENCES 407

7.17.1 WANDERING: A MODEL CASE

Mrs. G. walks intently toward the dining hall of the nursing home, a purse draped over her arm and golf hat sitting slightly askew atop her graying hair. When asked of her destination, she talks pleasantly in partial sentences of a luncheon date at the club. Mrs. G. enters the dining area and paces back and forth in front of a windowed wall overlooking a grassy courtyard. Appearing upset at times, she repeatedly attempts to slide the glass panels open. Passersby can hear her express determination to meet her husband, now long deceased, and observe her distress about his "overdue" arrival.

Most clinicians providing long-term care to cognitively-impaired persons encounter behavior such as this and call it wandering. Compared to its frequency in the clinical setting, the number of articles and studies of this behavior is small. Wandering has generally been viewed as an accompaniment to dementia or, worse, is seen as a nuisance. Perhaps its very commonness has worked against serious investigation of the phenomenon until the past decade.

This chapter presents a comprehensive review of the research and clinical literature on wandering. Sections addressing definitions and typologies, epidemiological data, putative risk factors and outcomes, etiologic factors and theories, assessment, and intervention are offered. Throughout each section, findings are summarized, gaps identified, suggestions for further research offered, and implications for clinical practice identified.

7.17.2 DEFINITIONS AND TYPOLOGIES OF WANDERING

Despite increasing interest in this perplexing behavior, the term "wandering" lacks a consistent meaning. Multiple definitions have been proposed from the clinical, conceptual, and empirical viewpoints of several disciplines, most notably psychology, nursing, and social work. Some definitions aim for objectivity while others impute a basis or cause for the behavior. Similarly, some attempt to circumscribe or limit the meaning to a narrow range of behavior while others use it loosely to include behaviors that accompany wandering, such as requests to leave, or are consequences of it, such as eloping or absconding from an expected or authorized area. The only universals in various characterizations of wandering discussed in this chapter are the presence of cognitive impairment and movement through large-scale space.

The term "wanderer" has also been applied to sleepwalkers (nocturnal wanderers) (Guilleminault, Kushida, & Leger, 1995; Guilleminault, Moscovitch, & Leger, 1995; Kushida, Clerk, Kirsch, Hotsin, & Guilleminault, 1995; Plazzi, Tinuper, Montague, Provini, & Lugaresi, 1995), vagrants (Weller & Jauhar, 1987), and patients who switch repeatedly from one physician to another (Pankratz & Jackson, 1994). These meanings of wandering are excluded from this review.

7.17.2.1 Clinical Views

Wandering poses a quandary for clinicians, inducing a conflict between the goals of affording both maximum freedom and safety for the afflicted. Consequently, clinicians usually advocate for a humane and tolerant approach to the behavior within a protected area. However, their definitions of wandering are among the most broad and imprecise and, thus, may inadvertently result in restriction of those whose ambulation poses no threat of harm. Coons (1988) defined wandering simply as self-initiated ambulation occurring independently of environmental cues to move, such as a ringing dinner bell or an invitation to an activity by a caretaker. While objective in that it is observable and infers no specific cause, Coons' definition is not particularly useful since virtually all spontaneous ambulation of cognitively-impaired individuals is included. In an early effort to clarify the behavior, Rader, Doan, and Schwab (1985) identified wandering as one end of a continuum of behaviors reflecting some effort to fulfill a social, emotional, or physical agenda or need that may or may not be rooted in current reality. While useful for pointing to possible constructive avenues for intervention, this definition fails to differentiate harmless from worrisome wandering and, thus, offers no clear guide for when to intervene.

7.17.2.2 Conceptual Schemes and Typologies

In efforts to more clearly delineate the behavior and explicate possible explanations for it, a number of others have taken more analytical approaches to defining wandering and classifying possible subtypes. Although data are offered in support of some of these typologies, most are more anecdotally than empirically derived.

Snyder, Rupprecht, Pyrek, Brekhus, and Moss (1978) defined wandering "as a tendency to move about, either in a seemingly aimless or disoriented fashion, or in pursuit of an indefinable or unobtainable goal" (p. 272). They also proposed the earliest typology of wandering based on chart reviews of eight nurse-identified wanderers. Three types were reported: (i) overtly goal-directed/searching behavior, associated with calling out repeatedly or approaching others in pursuit of a goal; (ii) overtly goal-directed/industrious behavior, typified by keeping busy with work-like actions, such as dusting; and (iii) apparently nongoal-directed behavior, described as ambulation that is aimless or shifts direction abruptly in response to changing environmental stimuli.

"Getting lost" or the inability to return to one's point of origin has also been suggested as a definition for wandering (Hussian, 1981, 1982). Greene, Smith, Gardiner, and Timbury (1982), who also equated wandering with getting lost, expanded this view by differentiating wandering types according to time of day (day or night) and location (outside vs. inside the home). However, wandering also entails a change in activity level that does not necessarily correlate with getting lost (Fairburn & Hope, 1988).

Based on observations of 13 residents in the geriatric unit of a state mental institution, Hussian and Davis (1985) proposed four independent types of wanderers: (i) akathesiacs, associated with extended hospitalization and long-term neuroleptic use and characterized by episodes of higher frequency and duration and lower attempts to leave or self-stimulate; (ii) exit seekers, associated with recent admission and characterized by frequent requests to leave, a high number and proportion of attempts to open exit (rather than other internal) doors, and a varied frequency and duration of ambulation episodes; (iii) modelers, who join with another ambulating resident and display little contact with door knobs, a low degree of self-stimulating behavior, and varying frequency and duration of episodes; and (iv) self-stimulators, who show other perseverative and stereotypical behaviors and no clear preference for exit doors. Analysis to support these categories was lacking.

More recently, a classification was proposed to differentiate dangerous from harmless wandering and to better target intervention strategies (Butler & Barnett, 1991). Based on the person's whereabouts and the patient's and nurse's knowledge of the same, four classes were delineated: purposeful (person and nurse know), escapist (person knows, nurse doesn't), aimless (nurse knows, person doesn't), and critical (neither know). While aimless wandering may be most stressful to caregivers, escapist and critical types are most dangerous due to associated elopements and injuries.

A time-in-motion approach was suggested to delineate wandering categories (Thomas, 1995). Continuous wanderers were said to be "moving constantly." Characteristics offered to describe this category included a higher level of cognitive impairment and inefficient travel behavior, a dissociation from feelings of fatigue or lack of awareness of the need to rest, and minimal social activity. On the other hand, sporadic wanderers move about occasionally. In addition to higher levels of cognitive function than continuous wanderers, sporadic wanderers were described as having greater ability to communicate verbally. While this classification scheme does reflect some clinically observed variations among wanderers also reported by others, Thomas used the cut-off point of 50% of time-in-motion to distinguish the category types without affording a rationale for this criterion.

One typology of wandering is clearly derived from an empirical base, where the travel patterns of wanderers ($n = 24$) and nonwanderers in a sample of 40 nursing home residents were evaluated (Martino-Saltzman, Blasch, Morris, & McNeal, 1991). Subjects' ambulation episodes occurring over a four-week period in pre-set locations were captured on videotape; 5768 events initiated without assistance of staff, visitors, or other residents were recorded. Each was described and then categorized by coders and the first author. Four patterns evolved: direct, travel between two points without diversion; pacing, repetitive back and forth travel within a limited area; lapping, repetitive travel encircling large areas; and random, haphazard travel between several locations without repetition. Coders achieved an acceptable inter-rater reliability (kappa = 0.79, $z = 13.2$, $p < 0.0001$) in a subsample of 231 events.

7.17.2.3 Empirically-based Descriptions and Features

Serious effort to evolve empirical descriptions and characterizations of wandering is limited to only a few recent studies. Two independent lines of research, differentiated by target population and methodology, are emerging. Although each approach has its limitations, these investigations are beginning to clarify what constitutes wandering behavior and to elucidate differences among wandering types.

Investigators in the UK are using community-based samples to uncover the inherent structure of wandering. Using a portion of the Present Behavioural Examination (Hope & Fairburn, 1992), a semistructured guide, they interviewed 29 caregivers of wanderers who were referred from the caseload of three community psychiatric nurses. The proportion of subjects showing each of the nine discrete behaviors at anytime since the onset of dementia was determined along with their gender, duration of dementia, and score on the Mini-Mental Status Examination (MMSE) (Folstein, Folstein, & McHugh, 1975). The following frequencies were reported: increased walking activity (17), wandering away from home (16), aimless walking (15), walking for an inappropriate purpose (14), checking/trailing (11), attempting to leave home (10), overactivity at night (9), pottering (6), and walking with appropriate purpose but an inappropriate frequency (4). No association with concomitant physical illness, medications, or premorbid activity levels was found. From these behaviors, a typology of wandering that could point to potential underlying etiologies was posed. These were: (i) increased overall walking activity, (ii) avoidance of being alone, (iii) diurnal rhythm disturbance, (iv) navigational problem, and (v) faulty goal-directed behavior (Hope & Fairburn, 1990). Although each subject exhibited more than one behavior, the time order of behaviors or any concurrent expression of them were not discussed. In a secondary analysis of the same data using Guttman scaling, Albert (1992) demonstrated that these nine behaviors were not a typology, but reflected a unidimensional hierarchical structure. Although the coefficient of reproducibility (0.78) fell slightly short of the 0.90 suggested by Guttman (1974), Albert argues that it is notable given the small sample size.

To challenge Albert's conclusion, Hope et al. (1994) conducted a factor analysis upon data from 83 subjects who were rated for 11 behaviors, that is, those listed earlier with the separation of checking/trailing and the addition of reduced walking. Only nine behaviors were used in analyses as wandering away from home and attempting to leave home did not correlate well with other behaviors. A four-factor solution explained 54.5% of the variance. Checking and trailing composed Factor 1, more and aimless walking constituted Factor 2, pottering stood alone as Factor 3, and inappropriate and overappropriate walking constituted Factor 4. Less walking and night-time walking did not load on any factor. In a scaling analysis of this data, neither the coefficients of reproducibility nor scalability were sufficient to support the conclusion of a hierarchical structure.

Investigators in the USA are using nursing home samples to describe various patterns of wandering in rhythmic terms (Algase, Beattie, Kupferschmid, & Beel-Bates, 1997; Algase, Beattie, Kupferschmid, & Therrien, 1997; Algase & Cheng, 1992a; Algase & Tsai, 1991). In their first study, a descriptive survey design and cluster sampling were used to screen 31 wanderers from a larger sample of 198 ambulatory, cognitively-impaired persons from eight randomly selected nursing homes. Wandering was viewed as a nonlinear, ultradian, biphasic rhythm with locomoting and nonlocomoting phases. It was operationalized by counting and timing cycles and their locomoting phases; each cycle was coded for impetus (self- or other-directed), location, and activity at onset and cessation of the locomoting phase using stopwatches and paper–pencil coding techniques. Without a typology for differentiating wandering from other ambulating, self-initiated cycles with locomoting phases ending in activity other than self-care (e.g., toileting, eating) were considered wandering. In accord with then current views of wandering as being aimless activity, "purposeful" ambulation, such as going to the toilet, was not counted as wandering. Wanderers were observed 1:1 for two 24-hour periods one week apart by coders trained on videotapes to time cycles to within 10 seconds of each other (Algase & Cheng, 1992a; Algase & Tsai, 1991).

Subjects averaged 79.24 wandering cycles (range = 19.5–162; SD = 39.33) and 110.07 minutes of locomotion (range = 69.04–257.02; SD = 61.32) daily. A distinct 24-hour pattern for hourly cycle frequency and hourly sum of locomoting phase durations was also identified (Algase, 1992a; Algase & Cheng, 1992a; Algase & Tsai, 1991), with the period from 1400 to 1600 representing the period of greatest wandering by both parameters for the sample as a whole. Night-time wandering was minimal, averaging less than one cycle per hour per subject.

Repeated measures ANOVA demonstrated that cycle frequency, locomoting and nonlocomoting phase durations, and percent of cycle locomoting were stable over a one-week period with no day effect or day by hour interaction effects. However, there was a significant hour effect on all cycle variables (Algase & Cheng, 1992a; Algase & Struble, 1992; Algase & Tsai, 1991). Further analysis also demonstrated that, of four start and stop types, solitary/unfocused starts (52%) and stops (45%) were most common, while social/unfocused starts (5%) and stops (5%) were least likely. Though relatively rare, cycles with social/unfocused starts had longer locomoting phases ($F = 3.76$; $p = 0.01$) and a higher percent of cycle locomoting ($F = 3.25$; $p = 0.02$). Variability in cycle period and nonlocomoting phases by start type and in all cycle variables by stop type were not significant (Algase & Cheng, 1992b). Though measures for wandering were cutting edge at that time, distinctions between wandering and nonwandering were crude and no classification to distinguish among any subtypes was used.

In a subsequent study, Algase, Beattie, Kupferschmid, and Beel-Bates (1997) estimated daily distribution, stability of direct and wandering ambulation cycles, and whether wandering occurring in a two-hour epoch is representative of that occurring in a 24-hour day. A correlational design incorporating direct observation and repeated measures over selected time periods was employed. Twenty-five ambulatory, cognitively-impaired residents from two long-term care settings served as the nonprobability sample.

Measurement of wandering was refined using Martino-Saltzman et al.'s (1991) typology to code the locomoting phase of each independent ambulation cycle. Accordingly, "inefficient" patterns (random, lapping, and pacing) were considered wandering, while "efficient" travel episodes (direct pattern) were not. Further, stopwatch methods were replaced with a portable terminal programmed to time entries, prompt coders, and store and transmit data for analysis. One subject per day was observed 1:1 over an initial 24-hour period beginning at 0700. On the third day following observation the subject was observed again for two two-hour periods: one standard (1400–1600) for all subjects, based on outcome of the study presented above, and a second unique period based on each subject's own 24-hour distributions of hourly ambulation cycle frequency and locomoting phase sums.

Distinct from previous studies, where researchers had classified subjects as wanderers or nonwanderers, this study revealed that all cognitively-impaired ambulatory subjects displayed wandering behavior as operationalized in this study. Although three subjects did not wander at all during the 24-hour period, they did wander during subsequent observation periods. All subjects, regardless of level of cognitive impairment, also displayed direct (nonwandering) ambulation. Descriptive statistics for wandering and direct cycle parameters are summarized in Table 1. Random was the dominant wandering pattern across all subjects, accounting for 77.2% of all wandering cycles, with lapping and pacing constituting 13.2% and 9.6%, respectively. Examination of the hourly frequency distributions of all ambulating revealed that the two-hour period of highest activity, 1700–1900 hours, was also the two-hour period in which the total wandering cycle frequency was greatest at 62 cycles per hour for the group as a whole.

To examine stability of wandering over a three-day period, wandering and direct cycle variables were compared at standard and individual periods. Wandering cycle frequency and locomoting phase sums were moderately and similarly correlated at both individual and standard periods; no direct cycle variable had a significant correlation at either period. For the standard period, all variables were stable for wandering and direct cycles. However, in the individual period, only the locomoting phase sum for wandering cycles was stable and there were more significant differences and, thus, less stability, among direct cycle variables.

The universal expression of wandering among all cognitively-impaired, ambulatory subjects has only been demonstrated in this one sample. Although subjects were drawn from two settings, the sample was small and nonrandom. Thus, further research is needed to validate and generalize this finding. However, it suggests that wandering may be a marker for disease progression or staging and could possibly serve as a measure for the efficacy of dementia treatment strategies in the future.

Table 1 Descriptive statistics for wandering and direct cycle variables over 24 hours ($N = 25$).

	Cycle frequency			*Sum locomoting phase*			*Mean % cycle locomoting*		
	Mean	*Range*	*SD*	*Mean*	*Range*	*SD*	*Mean*	*Range*	*SD*
Wandering	19.7	120	27.3	42.9	199.1	53.2	36.1	88	22.8
Random	15.2	104	23.2	34.6	173.1	46.4	40.1	88	26.2
Lapping	2.6	18	4.3	5.2	37.4	9.1	19.3	81.7	25.7
Pacing	1.9	9	2.6	3.1	26.6	5.8	20.8	85.5	27.7
Direct	22.0	64	13.2	29.0	92.3	18.8	28.9	34.9	10.2

7.17.3 EPIDEMIOLOGICAL DATA

7.17.3.1 Prevalence

Several estimates of wandering prevalence have been published, but all are flawed by unclear and inconsistent criteria for wandering and by mostly nonrandom, small, and incomparable samples. The basis for dementia may also vary by sample, a few being limited solely to subjects meeting diagnostic criteria for Alzheimer's disease (AD). Estimates are also likely influenced by other unknown or uncontrolled factors such as physical restraints and medications.

Among institutionalized demented elderly, estimates of the number of wanderers range from 6 to 24% (Dawson & Reid, 1987; Hiatt, 1980; National Center for Health Statistics, 1979; Wilkinson & Graham-White, 1980). In a random sample of eight nursing homes (Algase, 1988), wherein 747 residents (84% of those eligible) consented to cognitive screening, 5% were classified as wanderers by the nursing staff. As a percentage of those with cognitive impairments, the estimate doubled to 10%. Further, as a proportion of residents who were both cognitively impaired and ambulatory, the ratio was 1:4. Although not intended as an estimate of prevalence *per se,* Algase, Beattie, Kupferschmid, and Beel-Bates (1997) later found that in two long-term care settings, all ambulatory, cognitively-impaired residents, who had not necessarily been labeled as wanderers, displayed the behavior when directly observed over an extended time period. They concluded that if a person's overall ambulation activity was low or the proportion of their ambulation that was wandering was small, nurses tended to *underidentify* the phenomenon. However, in comparing nurses' ratings to observation, nurses have also *overestimated* the behavior (Milke, 1989). Conflicting conclusions may be due to different operational definitions for wandering in these studies.

In community samples of demented elderly, usually based on caregiver informants, estimates of wandering range between 19 and 56% (Burns, Jacoby, & Levy, 1990; Rabins, Mace, & Lucas, 1982; Teri, Larson, & Reifler, 1988). Among 92 consecutive patients with suspected dementia who received a home visit, approximately 37% were reported to get lost when outside the home and 28% when inside the home. These were most often persons with AD, as compared to vascular dementia, and level of cognitive impairment did not affect prevalence (Ballard, Mohan, Bannister, Handy, & Patel, 1991). In a survey of county sheriff's departments, sheriffs were called to locate one missing wanderer per year for every 1000 persons over the age of 65 years (Butler, 1990). Of these, five deaths occurred in 450 incidents. The cost to search for a missing wanderer was US$1200 per incident in law officer time alone. These estimates are conservative, as local police and volunteer fire departments are often the first line of assistance in such searches and therefore a number of events were likely excluded from this survey.

7.17.3.2 Incidence

No estimates for the incidence of wandering, either for the dementias in general or for any single type of dementia, were located. However, in staging dementia, "getting lost" is a criterion for stages 3 and 4 of the Global Cognitive Deterioration Scale (Reisberg, Ferris, deLeon, & Crook, 1982). By these instruments, at least one definition of wandering is accepted as an inevitable part of dementia, to be expected in every instance of the disease. However, the empirical basis for viewing this behavior as representative of a stage of dementia within these systems is unspecified.

7.17.4 PUTATIVE RISK FACTORS FOR WANDERING

7.17.4.1 Personal Factors

Currently, among those with cognitive impairments, knowledge is insufficient to predict who might wander or to specify the nature of the wandering behavior that one is likely to display. Although newer work has cast doubt on whether wandering is confined to a subgroup of cognitively-impaired ambulatory persons (Algase, Beattie, Kupferschmid, & Beel-Bates, 1997), older studies which compare wanderers and other nursing home residents may begin to illuminate the profile of a "problematic" wanderer. Since most of these older studies used nurses' or caregivers' judgments to identify wanderers, they likely represent comparisons between those with the most obvious display of the behavior and those with a low level or more subtle form of wandering. Variables most often reported include demographics, cognitive deficits, mobility, and psychosocial characteristics. Data from 12 samples are summarized in Table 2.

Most of these samples were not randomly chosen and, when comparisons were made with nonwandering controls, controls were generally also cognitively impaired. Wandering was noted in a wide age range of demented persons who were not consistently older or younger than controls. As might be expected, except for the two studies conducted on subjects from the Veterans' Administration, samples were largely female, representing a ratio of approximately 3:1. This gender ratio was maintained among wanderers. Samples were primarily ambulatory, with a few subjects in wheelchairs. When mobility was a variable of interest, wanderers with freedom to walk had better overall mobility than controls. Although wanderers were found among persons with all levels of cognitive impairment, as a group, their performance on global measures of impairment were generally worse than controls. On discrete cognitive tasks, poorer performance was found in memory and recall, abstraction, orientation, and visuospatial skills. Wanderers also had poorer language skills and were less able to communicate even though their personalities were generally more agreeable and extroverted. While none of these studies were strictly aimed at identifying risk factors, their overall results reveal that advanced cognitive impairment, good mobility, and impaired ability to communicate, particularly among individuals with a history of sociability, may portend a tendency to wander.

While the question of whether medical diagnosis, as a basis for dementia, might help to identify wanderers is an interesting one, few studies cast light on this possibility. Several studies do report various types of dementia among wanderers with a range of pathologies including AD, multi-infarct dementia (MID), and mixed dementia (Algase, 1992a; Hope et al., 1994; Thomas, 1997). However, due largely to small sample sizes and the limited nature of medical documentation in long-term care settings (from which most subjects are drawn), little analysis of the relationship between medical diagnosis and wandering has been done. In a random sample of 163 cognitively-impaired nursing home residents, 10% of one sample ($n = 163$) had only an irreversible basis for their dementia, 31% had only a reversible basis, and 59% had a mixed basis (both reversible and irreversible etiologies) noted in their medical records (Algase, 1992a). Wanderers tended to have their dementia due to mixed or irreversible causes. In a sample of 68 nursing home residents, drawn from a 911-bed nursing facility, wanderers ($n = 31$) tended to have a diagnosis of AD while nonwanderers were more likely to have a vascular dementia (Thomas, 1997). However, studies to examine the relative frequency, overall output, and type of wandering associated with various medical bases of dementia have yet to be conducted.

7.17.4.2 Potential Environmental Contributors

Although behavior in general (Lewin, 1935; Murray, 1938), among the aged (Whedon & Shedd, 1989), and within wanderers (Algase, 1992b; Algase & Struble, 1992; Hussian, 1981), is understood to arise as a function of person–environment interaction, few studies have yet explicated physical or social factors in the environment that may precipitate or otherwise influence wandering.

In a case study, observation of the intrusions of one wanderer into the rooms of other residents revealed a consistent pattern: attraction to a limited number of rooms which shared a common view of the outdoors and dalliance at a particular dresser with a mirror (Donat, 1984). By reassigning the wanderer to the room most frequently entered and containing the dresser, intrusions into other rooms decreased from an average of 16 per hour to less than two. While manipulating the environment had an effective outcome for those intruded upon, it made no real change in, but rather accommodated to, the subject's behavior. Thus, both the theoretical and practical value of this study is limited as no attempt to analyze attracting environmental features from a conceptual standpoint was made.

In analyses of the nursing home environment for its impact on agitated behaviors, observations of pacing (including "aimless" wandering) under naturally occurring conditions were also reported (Cohen-Mansfield & Werner, 1985; Cohen-Mansfield, Werner, & Marx, 1989;

Table 2 Characteristics of 12 samples of wanderers.

Author	*N*	*Control group*	*%wan.*	*Age*	*%fe*	*Cognitive status of wanderers*	*Mobilty status of wanderers*	*Communication/sociability of wanderers*
Algase (1992)	163	CI NHR[a]	20.0	mean = 83.7	77.0	more global impairment; $p<0.001$ poorer abstract thinking; $p<0.01$ poorer judgment, $p<0.001$ poorer spatial skills; $p<0.001$	ambulatory	poorer language skills, $p<0.001$
Cornbleth (1977)	48	CI NHR	62.5	68.7	0.0[b]	global impairment, ns	better if free to walk; $p<0.05$	worse social skills; $p<0.05$
deLeon et al. (1984)	16	CI NHR	31.3	72–90	nr	global impairment, ns poor parietal signs, $p<0.001$	nr	nr
Henderson et al. (1989)	28	AD	39.0	54–89; ns	ns	worse recall; $p<0.002$ poor visuoconstruction; $p<0.05$	ambulatory	nr
Hope and Fairburn (1990)	29	CI	100.0	55–92	69.0	MMSE 0–21	59% excessive	21% avoided being alone
Jackson et al. (1989)	8983	G NHR[b]	6.0	older; $p<0.01$	ns	more global impairment; $p<0.01$	better, $p<0.001$	poor communic, $p<0.0001$
Martino-Saltzman et al. (1991)	40	CI NHR	60.0	56–104	63.0	more impairment correlated with poorer travel efficiency		
Mayer and Darby (1990)	9	na	100.0	mean = 77.8	nr	100% severely impaired	walk or wc	nr
Monsour and Robb (1982)	44	CI NHR	50.0[c]	53–89	0.0[b]	55% severely impaired	walk or wc	nr
Ryan et al. (1995)	18	CI	33.3	70–92; ns	72.2; ns	more perseveration; $p<0.05$ attention, concentration; ns	ambulatory	nr
Snyder et al. (1978)	8	CI NHR	50.0[c]	ns	ns	worse memory, $p=0.001$ worse orientation, $p<0.001$	nr	alone more, $p<0.002$ poor conversing, $p<0.005$
Thomas (1997)	40	CI NHR	50.0[c]	younger $p<0.01$	ns	global, ns primarily AD; $p<0.05$	ambulatory	more extroverted; $p<0.01$ more agreeable; $p<0.01$

[a] Cognitively impaired nusing home residents. [b] General nursing home residents. [c] Matched sample. [d] VA setting.
ns = not significantly different from controls.
nr = not reported.

Cohen-Mansfield, Werner, Marx, & Freedman, 1991). Features of the physical environment in which wandering increased included location (corridor, nurses' station, and another resident's room) and normal lighting, noise, and temperature conditions. With colder temperatures, wandering decreased. Wandering also decreased significantly under certain social conditions, namely during structured activities and activities at daily living (ADLs), but increased in the presence of a few people (2–5) or when housekeeping staff or residents from another unit were present. However, these researchers made no attempt to manipulate environmental factors to elucidate effects on wandering.

Work on wandering in controlled environments wherein aspects of the environment are specifically manipulated is just beginning. To test the premise that the low light conditions associated with sundown increase confusion and wandering, a study of 10 wanderers was conducted in a specially controlled nursing home lounge. Preliminary analysis revealed no impact of various lighting conditions (dim, bright, or average ambient lighting) on the frequency of wandering episodes or total time wandering in the lounge (Algase, 1997a). However, although subjects were observed during the time of day when each was normally most active, they displayed very little wandering in the highly simplified, controlled environment of this study, suggesting that other aspects of the environment may have a greater potential to reduce wandering than light.

7.17.5 PUTATIVE OUTCOMES OF WANDERING

Outcomes of wandering have not been extensively studied and most knowledge concerning them has been derived from clinical experience. Falls and elopements are the direct, more immediate, and, due to their association with high morbidity (e.g., hip fractures) and mortality, the most serious outcomes of wandering. Weight loss, sleep disturbances, and fatigue are more indirect outcomes thought to reflect wandering's cumulative effects (Algase & Struble, 1992). An outcome that is not well documented, but frequently observed in the course of my clinical practice and research, is the tendency of other nursing home residents to shun or berate wanderers, most likely in an effort to protect themselves, their space, and their possessions from unwelcomed intrusions. Although effects of such negative responses to wandering on morale of wanderers is unknown, logic suggests that it is not a pleasant experience. In fact, it is not uncommon for such interactions to be verbally hostile, and they can escalate into physical aggression on the part of either party.

7.17.5.1 Falls

While cognitive impairment is the most important and frequently documented risk factor for falls (Buchner & Larson, 1987; Janken, Reynolds, & Swiech, 1986; Morse, Tylko, & Dixon, 1987; Nevitt, Cummins, Kidd, & Black, 1989; Reinboth & Gleidenvand, 1982; Rubenstein, Josephson, & Robbins, 1994; Tinetti, Speechley, & Ginter, 1988; Walshe & Rosen, 1979), little empirical data exists to support the clinical contention that wandering compounds the risk of falling (Cancro, 1968; Soverini & Borghesi, 1968). However, any such effects are probably mediated by disturbances in gait and balance associated with cognitive impairment itself (Kippenbrock & Soja, 1993; Morse, Tylko, & Dixon 1985; Morse et al., 1987). Characteristics of wanderers or types of wandering most likely to result in falls are unknown.

The higher frequency of and time spent in ambulation among wanderers may simply increase opportunity to fall. Wanderers as a whole are thought to be more physically able than less active persons with cognitive impairments. However, as dementia progresses, wandering increases at the same time as physical vigor declines for a given individual. Thus, the ability to apprehend and interpret environmental hazards worsens in tandem with deterioration in one's physical stamina. Based on a study of 60 institutionalized women with AD, researcher's concluded that falls were related to a decline in overall vigor in those who had been active (Brody, Kleban, Moss, & Kleban, 1984). Together, these factors may interact to increase the incidence of falls among wanderers.

7.17.5.2 Elopements

The most potentially dangerous consequence of wandering is elopement (Henderson, 1993–4; Mulrooney, 1993–4). Elopement, the unattended leaving of the premises, has been linked to out-of-facility deaths (Foxwell, 1993; Morino, 1989). While reported elopements have declined, largely due to growth of dementia units equipped with sophisticated security systems, they remain a serious and costly problem (Henderson, 1993–4). Death or serious injury occurring during elopement accounts for an estimated 17% of malpractice suits directed toward long-term care facilities (Kapp, 1986).

Risk management programs and in-service education, while essential in training personnel, have not demonstrated a reduction in the incidence of elopement (Kapp, 1986). Due to nursing home reforms contained in the Omnibus Budget Reconciliation Act (OBRA), the use of restraints to ensure safety for persons who wander is increasingly challenged, legally and empirically, because of known undesirable side effects of restraint use (Evans & Strumpf, 1989; Miller, 1975). Thus, the need for alternative approaches to reduce or control elopements, based upon a clearer understanding of particular wandering behaviors and circumstances precipitating egress, remains.

7.17.5.3 Weight Loss

Weight loss is associated with progressive dementia. In one study of hospitalized patients who were no longer mobile, those with AD were an average 11.3 kg (21%) lighter than those who were not demented and 6.6 kg lighter than those with MID despite comparable nutritional indices and food intake for all three groups (Singh, Mulley, & Lusowsky, 1988). In a retrospective comparison of weights recorded in the medical records of these same patients, body weight fell significantly for those with AD but did not change appreciably for the others. Reasons for these differences are poorly understood.

No studies examining the effects of wandering on weight, food intake, or nutritional indices were located. However, whatever the basis for weight loss in dementia and particularly in AD, wanderers may be prone to even greater losses for two reasons. First, their activity level and, therefore, their energy expenditure is high. Second, due to their activity level, wanderers are often difficult to keep seated and focused during mealtimes. Although they are usually capable of feeding themselves, wanderers often walk away from the table before having consumed their entire meal and, thus, may be at greater risk for insufficient caloric intake to offset their high level of energy expenditure.

7.17.5.4 Sleep Disturbance and Fatigue

Sleep patterns in dementia are altered to a degree even greater than that attributable to aging (Prinz, Peskind et al., 1982). The sleep of persons with dementia is punctuated with frequent arousals and nearly devoid of the deep sleep and REM sleep needed for physical and mental restoration (Vitiello, Poceta, & Prinz, 1991). Brief daytime naps are also common.

While the effects of wandering on sleep have not been reported, an examination of the individual 24 hour distributions of wandering episodes in 58 nursing home residents with cognitive impairment does confirm that wandering occurs during the night (Algase, 1997b). However, by comparison to daytime wandering, the frequency of night-time episodes and their total duration is quite small (see Figures 1 and 2). Distributions vary widely among subjects and not all daytime wanderers display the behavior at night. In an analysis of six cases, three with severe levels of cognitive impairment displayed wandering behavior during more hours in a day (9–19) than three with moderate levels of impairment (6–8). The longest period of time without a wandering episode was shorter for those with severe impairment (5–10 hours) than it was for those whose impairment was moderate (14–16 hours) (Algase, Beel-Bates, & Kupferschmid, 1995). Thus, wanderers, particularly those with more advanced cognitive loss, receive little uninterrupted rest.

What is unknown is whether the propensity to wander interferes with sleep to a greater degree than dementia itself or whether the fractured sleep patterns accompanying dementia inadvertently provide greater opportunity to wander. Clearly, more studies are needed to explicate relationships among sleep, fatigue, and wandering, as rest periods may be a potentially useful intervention (Hall & Buckwalter, 1987).

7.17.5.5 Other Possible Outcomes

Longitudinal effects of unrestrained wandering have been the focus of only one study (Cornbleth, 1977). Physical, psychosocial, and cognitive outcomes were assessed in a sample of 48 veterans prior to and at 6 and 12 weeks following their transfers between an experimental and other wards. Controlling for baseline performance, wanderers *on* the experimental ward, where they were free to walk about at will, showed greater physical ability (range of motion) at both postmove time points as did nonwanderers *off* the experimental ward. However, regardless of ward assignment, wanderers showed lower psychosocial functioning over time, reflected primarily in terms of friendliness, self-control, and communication skill than nonwanderers. Neither ward nor group had an effect on measures of cognitive performance.

Thus, not all outcomes of wandering are negative. Although no measures of wandering behavior were employed to demonstrate an actual increase in ambulation on the experimental ward, the above study reveals that

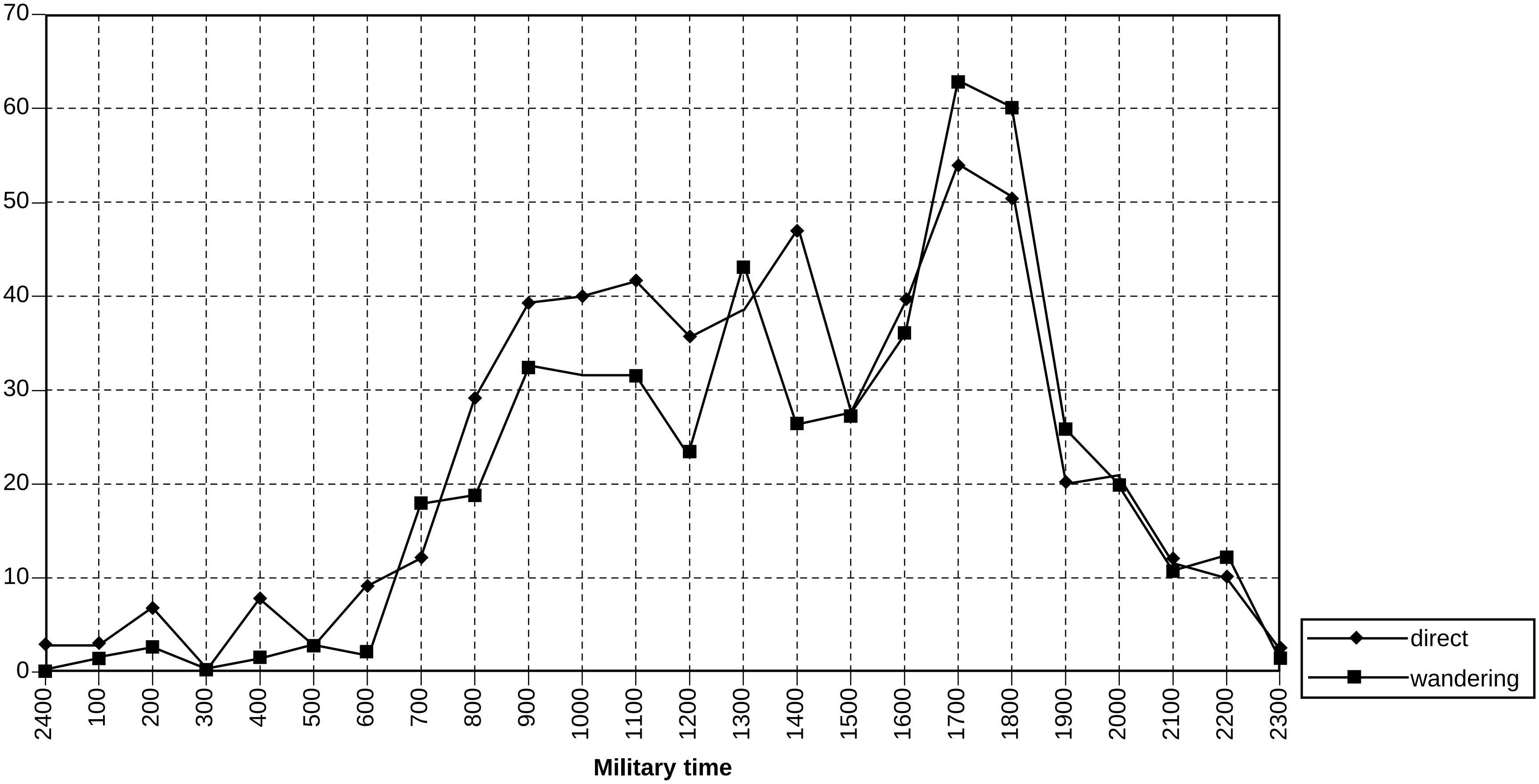

Figure 1 Sum of wandering and direct episode frequencies by hour (N = 25).

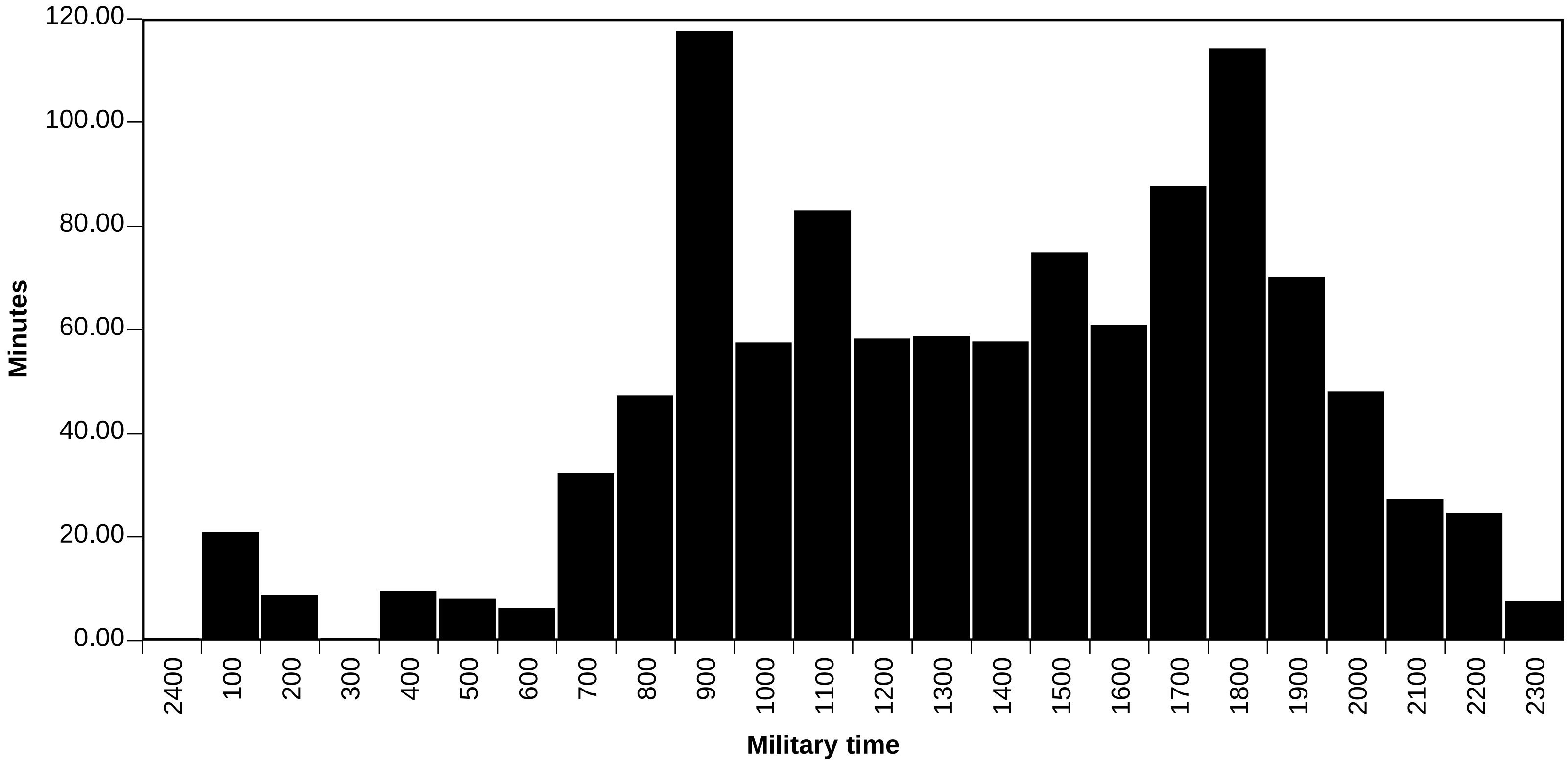

Figure 2 24-hour distribution of minutes spent wandering ($N = 25$).

physical functioning may be enhanced by freedom to wander. Unfortunately, on wards housing both wanderers and nonwanderers, such freedom for wanderers may adversely affect nonwanderers, who may inhibit their movements as a protective strategy. Further, wandering may have a negative impact on the psychosocial skills of those afflicted.

7.17.6 ETIOLOGIC THEORIES AND FACTORS

While several etiologies for wandering have been suggested, none as yet fully explain it. Most explanations are speculative, lacking adequate testing, yet all remain viable as hypotheses. The most glaring deficiency in efforts to test these "theories" is a reliance on proxy variables for wandering, rather than on any direct, valid, and reliable measure of the behavior itself. When eventually tied to observable features of wandering, each of these various theories may be able to explain some dimension or type of the behavior, thereby lending direction to more targeted intervention strategies for the most troubling aspects of wandering.

7.17.6.1 Wandering as Spatial Disorientation

Based on multiple findings from lesion and other studies on rats, the hippocampus (HPC), situated deep in the parietal lobe, is known to capture memories of spatial environments (Nadel, 1991). In humans, the right HPC has a special role in spatial learning (Goldsten, Canavan, & Polkey, 1989) and HPC damage is a classic neuroanatomical feature of dementia, particularly in AD. Further evidence supporting specific involvement of the right parietal lobe in wandering is drawn from a recent study of 18 subjects with vascular dementia, nine of whom were identified as wanderers on the basis of behaviors observed by staff in a long-term care setting (Meguro et al., 1996). Using positron emission tomography (PET), these researchers demonstrated that wanderers had significantly higher cerebral metabolic rate for glucose (CMR_{glc}) than nonwanderers in the left and right frontal lobe, left parietal lobe, temporo-parieto-occipital region, left occipital lobe, and cerebellum. Thus, while the CMR_{glc} of the right parietal lobe of wanderers was neither higher nor lower than nonwanderers, it was relatively lower than other brain regions for wanderers.

Impaired spatial memory is an expected feature of dementia and "getting lost" in familiar or unfamiliar surroundings has been used to characterize particular stages of cognitive decline (Reisberg et al., 1982). Several researchers have demonstrated the presence and nature of spatial disorientation in AD and other dementias. In tests of parietal lobe function, wanderers ($n = 5$) were shown to perform more poorly than similarly demented controls ($n = 16$) (deLeon, Potegal, & Gurland, 1984). Using neuropsychological measures of visuoconstructive (right parietal) performance, long-term declarative memory, attention, language performance, and dementia severity, Henderson, Mack, and Williams (1989) examined spatial disorientation, evaluated by caregiver ratings on four items (wandering, getting lost indoors, getting lost on familiar streets, unable to recognize familiar surroundings) from the Memory and Behavior Checklist (S. Zarit, Orr, & J. Zarit, 1985) in a sample of 28 ambulatory persons who met criteria for probable AD. Eleven subjects (39%) showed substantial spatial disorientation yet did not differ from remaining subjects by age, education, duration of symptoms, or gender. Correlations of neuropsychological measures to spatial disorientation ranged from 0.21 (attention) to 0.56 (long-term memory). Long-term memory and visuoconstructive performance together explained 41% of the variance in spatial disorientation with memory being the stronger predictor.

Visuospatial ability has also been examined in relation to right–left orientation (R/L-O) in AD and MID subjects with comparable dementia severity. Operationalized by putting geometric objects through holes of corresponding shape, visuospatial ability was significantly correlated to R/L-O (as applied both to self and external to self) in MID with focal lesions, but not in AD (Fischer, Marterer, & Danielczyk, 1990). In a study comparing AD and Parkinsons' dementia (PD), both groups, which were matched for dementia severity, age, and education, performed substantially worse on visuospatial tests than nondemented controls (Mohr, Litvan, Williams, Fedio, & Chase, 1990). Further, these PD subjects showed a selective deficit on the Object Assembly subscale of the WAIS-R, a task not principally associated with memory and thought to depend on frontal lobe functioning. However, the AD subjects performed more poorly on tasks with principal memory components, likely reflecting involvement of the HPC which, while always affected in AD, is not consistently so in PD (Chiu et al., 1986; Jellinger, 1987).

Still unknown from these works is the correspondence between these measures of visuospatial ability, which are done in grasping space, and performance of spatial tasks, such as way-finding, which occur in large-scale space. A study by Liu, L. Gauthier, and S. Gauthier (1991) lends support to linkages in performance at these two spatial levels. These researchers administered an array of tasks reflecting perceptual, cognitive, and functional spatial skills to 30 subjects, 15 with early AD and 15 age-matched normals. Perceptual spatial skills included figure-ground perception, visual and tactile shape recognition, visual and tactile size discrimination, position in space, spatial relations, and left–right orientation. Cognitive spatial skills were: spatial planning, decision-making, problem-solving, mental direction rotation, immediate spatial memory, spatial learning, and mental representation of both a familiar and a new environment. While all skills in the perceptual category and most in the cognitive category were performed within grasping space, one (problem solving) in the latter group did require walking in a large-scale space that afforded total visual accessibility. Orientation within a familiar and a new environment were the functional spatial skills examined. Both of these required walking to specified locations that were not immediately visible from one's point of origin.

Some perceptual skills were intact for AD subjects, namely visual shape discrimination, both visual and tactile size discrimination, and left–right discrimination. However, the measure of left–right discrimination favored application to self, as opposed to application external to self. AD subjects displayed gross disturbance of cognitive skills, performing significantly worse than normal controls on all tasks in this category. In finding their way about a highly familiar environment, AD subjects did as well as controls, but in a new setting they performed only half as well as controls.

While results of these studies indicate that spatial orientation is disturbed in dementia, the relationship of performance on various visuospatial tasks to observed travel patterns and ambulation efficiency and to the frequency, duration, distribution, and location of wandering remains unknown. Effective and efficient ambulation depends on spatial memories, also known as cognitive maps. In fact, building spatial memories requires some degree of exploration, which includes ambulation. As a kind of ambulation occurring concomitant with AD and other dementias, it seems reasonable to postulate that at least some aspects of wandering, most likely its geographic or spatial patterns, may reflect deterioration of such maps (i.e., difficulty remembering one's way around), or reduced capacity to form new maps (i.e., difficulty learning one's way around). Evidence for the link between spatial disorientation and the spatial patterns of wandering is found in the swim patterns of rats. Rats with unilateral and bilateral HPC damage swim in patterns resembling, respectively, the lapping and random travel patterns of demented elderly observed by Martino-Saltzman et al. (1991) (Holden & Therrien, 1993).

7.17.6.2 Wandering as a Movement Disorder

7.17.6.2.1 The HPC and locomotor output

Three presently unrelated views have been offered to explain the high level of movement in wanderers. The first again involves the HPC. In addition to its role in spatial orientation, researchers have demonstrated that the HPC also affects the amount of locomotion. Classical work (Ranck, 1973; Vanderwolf, 1969; Vanderwolf, Kramis, Gillespie, & Bland, 1975) first established that the electrical activity of cells in the HPC is correlated with locomotion. Work in 1994 by Whishaw, Cassel, Majchrzak, Cassel, and Will suggests that animals with damage to HPC shift towards more locomotion (i.e., shorter stop periods) on what is termed the mobility gradient. Specifically, theta units (groups of neurons, most likely HPC interneurons and/or dentate granule cells) produce a characteristic theta rhythm (7–10 Hz; also known as rhythmic slow wave activity). Rhythmical firing of theta units occurs when animals are walking, running, jumping, or moving the head. In contrast, predominantly irregular wave patterns and low rates of irregular firing or absence of theta is seen during motionless periods or with appetitive behaviors like grooming, chewing, etc. (Vanderwolf & Leung, 1983; Whishaw, 1991), suggesting that theta activity is involved in generating rhythmical locomotion and modulating movements and that intrinsic pacers of this rhythm basic to locomotion exist (Morgenson & Nielsen, 1984; Muller & Kubie, 1989; Traub, Miles, & Wong, 1989; Whishaw & Nikkel, 1975).

In humans, case reports of HPC theta exist from EEG recordings, although theta is difficult to record without invasive procedures and special electrode placements. However, Halgren (1991) has recorded action potentials related to movements requiring intention or high effort, indicating that human HPC neural activity strongly influences movement. Further he suggests (Halgren et al., 1980) that these are endogenous potentials or pacers. Thus, the well-established involvement of the HPC in AD and some other dementias may also explain the overall increase in locomotion observed among wanderers.

7.17.6.2.2 Perseveration

Perseveration, or difficulty in stopping or shifting behavior, is a second possible characterization for the overall high activity levels of wanderers. In other words, once walking is initiated, wanderers may have difficulty terminating an episode. Present with focal damage to the parietal and temporal lobes (Goldberg, 1986; Sandson & Albert, 1984) and with more diffuse central nervous system pathology (Goldberg, 1986), perseveration manifests as repetitive behavior, usually without a discernible stimulus. Perseveration is a documented feature in many cases of AD, MID, and PD (Fuld, Katzman, Davies, & Terry, 1982; Marley, 1982) and can involve repetitious speech patterns (e.g., echolalia, word intrusions) and motor behaviors.

This possible view of wandering was the basis of a study by Ryan et al. (1995). Eighteen community-residing subjects (72% female) with mild to moderate dementia were classified as wanderers ($n = 6$) or nonwanderers on the basis of caregiver ratings. Three types of perseverative motor behavior were quantified from subjects' drawings on the Bender Gestalt and clock drawing tests. The groups did not differ with regard to attention and concentration, spatial orientation, age, gender, or dementia severity. Wanderers demonstrated significantly greater perseveration than nonwanderers, although nonwanderers did perseverate. Wanderers were especially likely to have an ongoing, continuous, or repeated response to a singular stimulus, a form of perseveration associated with the right hemisphere and with attention or motor output deficits (Marley, 1982) or to carry over elements of one drawing into another, a form of perseveration associated with bilateral frontal or frontotemporal impairments (Sandson & Albert, 1987). Interestingly, authors of this study did not confirm spatial orientation deficits of wanderers reported above. They attributed this lack of difference to either their method for scoring the clock drawing test or to the earlier stage of dementia of their subjects.

7.17.6.2.3 Effect of neurotransmitters

A third hypothesis addressing the high degree of motor output of wanderers is not based on localization to particular brain structures or regions, but involves neurotransmitter systems. Weller (1987) suggested that the reduction of acetylcholine characteristic of AD may result in a relative excess of dopamine. Seeing wandering as indicative of restlessness, he further proposed that, analogous to its role in schizophrenia, this dopamine "excess" may account for wandering. However, no studies to evaluate this hypothesis could be located.

7.17.6.3 Wandering as a Circadian Rhythm Disruption

The suprachiasmatic nucleus (SCN) is thought the be the "biological clock" that governs daily fluctuations occurring in, at least, some physiological and behavioral phenomena in humans. Because the SCN is located atop and connected to major optic tracks in the brain, phenomena that naturally change in accord with the light–dark cycle, such as activity and rest, are more likely affected by it. The SCN decreases in size with aging and, to an even greater extent, in AD (Swaab, Fliers, & Partiman, 1985). Loss of neurons in the SCN accounts to a large degree for degradation in the normal sleep–wake cycle observed in both aging and AD (Prinz, Peskind et al., 1982; Prinz, Vitaliano et al., 1982; Witting, Kwa, Eikelenboom, Mirmiran, & Swaab, 1990). Further, if neuronal loss causes the SCN to fall below a certain critical mass, circadian periodicity may be completely obliterated (Brock, 1991).

Along with wandering, day–night reversal and "sundowning" are other clinical concerns observed in dementia that may coincide with shrinkage of the SCN. In fact, several authors have actually linked sundowning with wandering (Eisdorfer & Faan, 1982; Rindlisbacher & Hopkins, 1991). One group (Rindlisbacher & Hopkins, 1992; Rindlisbacher, Hopkins, & Grant, 1992) studying sundowning has used methodology (activity monitoring) similar to that used by researchers studying wandering (Algase, Beattie, Kupferschmid, & Beel-Bates, 1997; Satlin et al. 1991).

As sleep and physical activity are mutually exclusive, an inverse relationship between sleep and wandering would be expected. Thus, changes in overall circadian periodicity associated with AD and other dementias are likely to affect the daily distribution of wandering episodes. Figure 3 illustrates this likelihood as the data depicted for severely impaired cases, where SCN shrinkage would be advanced, show more night-time wandering and wandering occurring in more hours of the day than the data shown for moderately impaired cases. Further evidence supporting the role of the SCN in affecting wandering distribution is found in a comparison of AD subjects classified as "pacers" and "nonpacers" (Satlin et al., 1991). These researchers reported that relative amplitude in 24 hour activity patterns was lower for pacers, suggesting a weakening and loss of integrity in circadian periodicity that exceeds that of nonpacing AD victims.

7.17.6.4 Wandering as an Expression of Mood or Affect

Terming it "agenda behavior," clinicians have attributed wandering to affective states, such as anxiety or fear, underlying the here-and-now experience of a cognitively impaired person (Rader et al., 1985). However, while an appealing explanation because of its potential for intervention, little research yet exists to substantiate this claim. While Cohen-Mansfield and Marx (1990) demonstrated a connection between depression and screaming behaviors, no such relationship between depression and wandering was revealed. In a more recent work wherein agitation was viewed as a reflection of negative affective states such as anxiety or anger, researchers (Algase, Beattie, Kupferschmid, & Beel-Bates, 1997) used Cohen-Mansfield's Agitation Inventory (CMAI) to explain rhythmic parameters for various patterns of wandering. The physical aggression subscale (PA) of the CMAI, which contains behaviors suggestive of anger or frustration, accounted for 37% of the variance in cycle frequency and 43% of the variance in the sum of locomoting phases for all wandering patterns together (random, lapping, pacing) over a 24 hour day, with most of this attributable to the predominance of random-pattern wandering. The PA factor also accounted for 40% of the variance in the mean locomoting phase duration and 54% of the variance in the percentage of a cycle spent in locomotion for the pacing pattern. The physically nonaggressive subscale (PNA), which contains behaviors suggestive of anxiety or restlessness, accounted for 18% of the variance in cycle frequency, 19% of the variance in the 24-hour locomoting phase sum, 23% of the variance in the nonlocomoting phase mean, and 30% of the variance in the percentage of cycle locomoting for all wandering patterns together. Again, cycle frequency and locomoting phase sum results are largely attributable to the preponderance of random-pattern wandering displayed by subjects. However, nonlocomoting phase mean results were most influenced by lapping, for which the PNA predicted 39% of the variance. Results for the percentage of cycle spent in locomotion were largely influenced by pacing, for which the PNA predicted 61% of the variance. The verbal aggression subscale (VA), which contains behaviors suggestive of anger and general confusion, predicted only the nonlocomoting phase mean (19%) for random-pattern wandering. While promising, results should be viewed with caution as no controls for the effects of cognitive impairment were considered in these analyses. Although preliminary and based only on a convenience sample of 25 long-term care residents, these results do suggest that affect may be important in the expression of particular parameters of wandering and, with further study, these may be useful in identifying what kind and amount of wandering might be explained by affect and, thus, be potentially influenced by interventions targeted to change affect.

7.17.6.5 An Emerging Multifactorial Explanation for Wandering

A group of investigators devised a putative explanatory model for wandering and other dementia-related behaviors based on an extensive review of the literature (Algase et al., 1996). Potential etiologies were grouped and organized as either background factors (i.e., the relatively stable or slowly changing characteristics of persons with dementia) and proximal factors (i.e., the more fluid and dynamic features of the person or the environment). Background factors included dementia-compromised functions (e.g., memory, circadian rhythms), health state, demographics, and psychosocial variables (e.g., personality, response to stress); proximal factors included physiological need states (e.g., pain, hunger), psychological need states (e.g., affect), physical environment (e.g., light level), and social environment (e.g., staff stability). In addition to factors thought operative for any dementia-related behavior, a particular set of additional factors were also delineated for selected behaviors, including wandering. Additional variables posed as relevant to wandering were level of social engagement, complexity of environmental design, and crowding.

Figure 4 is a more recent schematic of the model as it pertains to wandering alone. Major propositions of the model are that: (i) background and proximal variables interact or combine in some way to produce wandering, (ii) proximal variables are perceived in the context of existing background variables, and (iii) some background variables may have a direct influence on wandering independent of any proximal variable. Accordingly, unless it is a direct result of an underlying neuropathology, wandering arises from some immediate perceived need and is expressed in a way that is compromised by dementia and mediated by environmental considerations.

The authors acknowledge that the model is tentative and requires empirical testing. How-

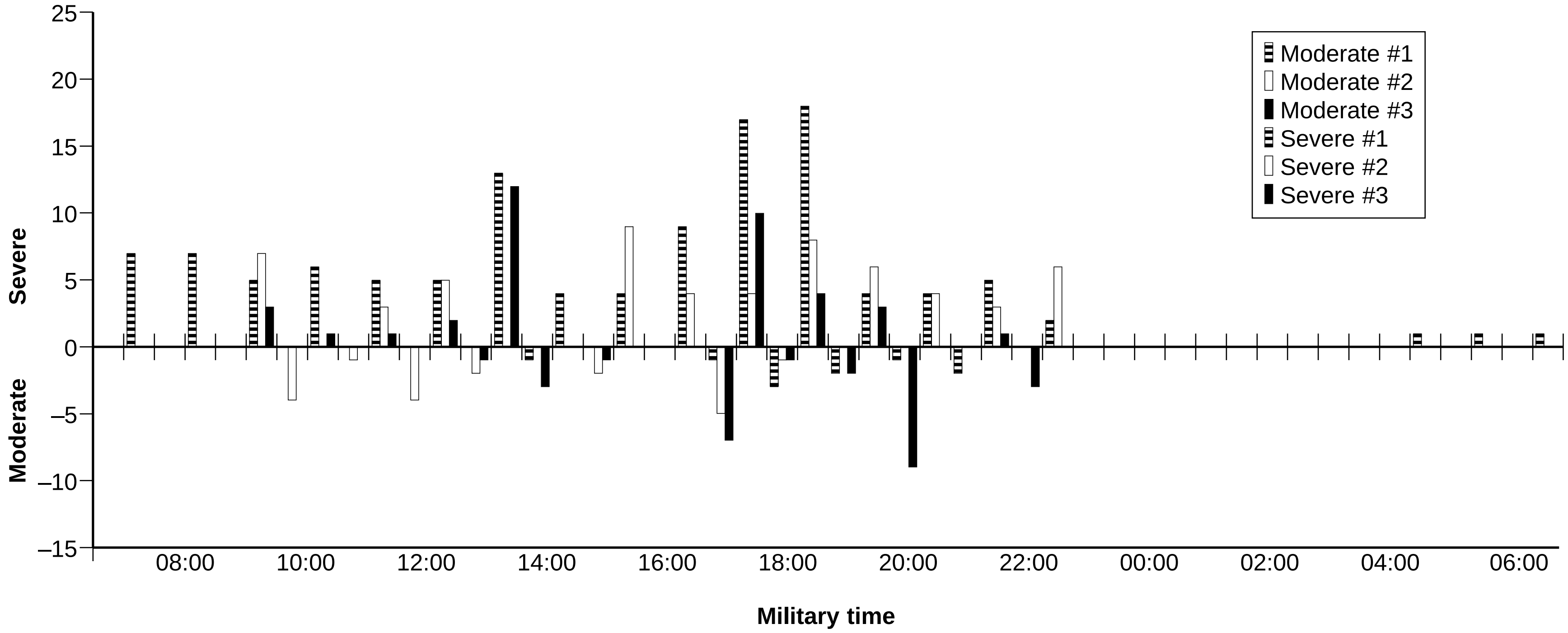

Figure 3 24-hour distribution of wandering episodes comparing moderate ($n = 3$) and severe ($n = 3$) cases.

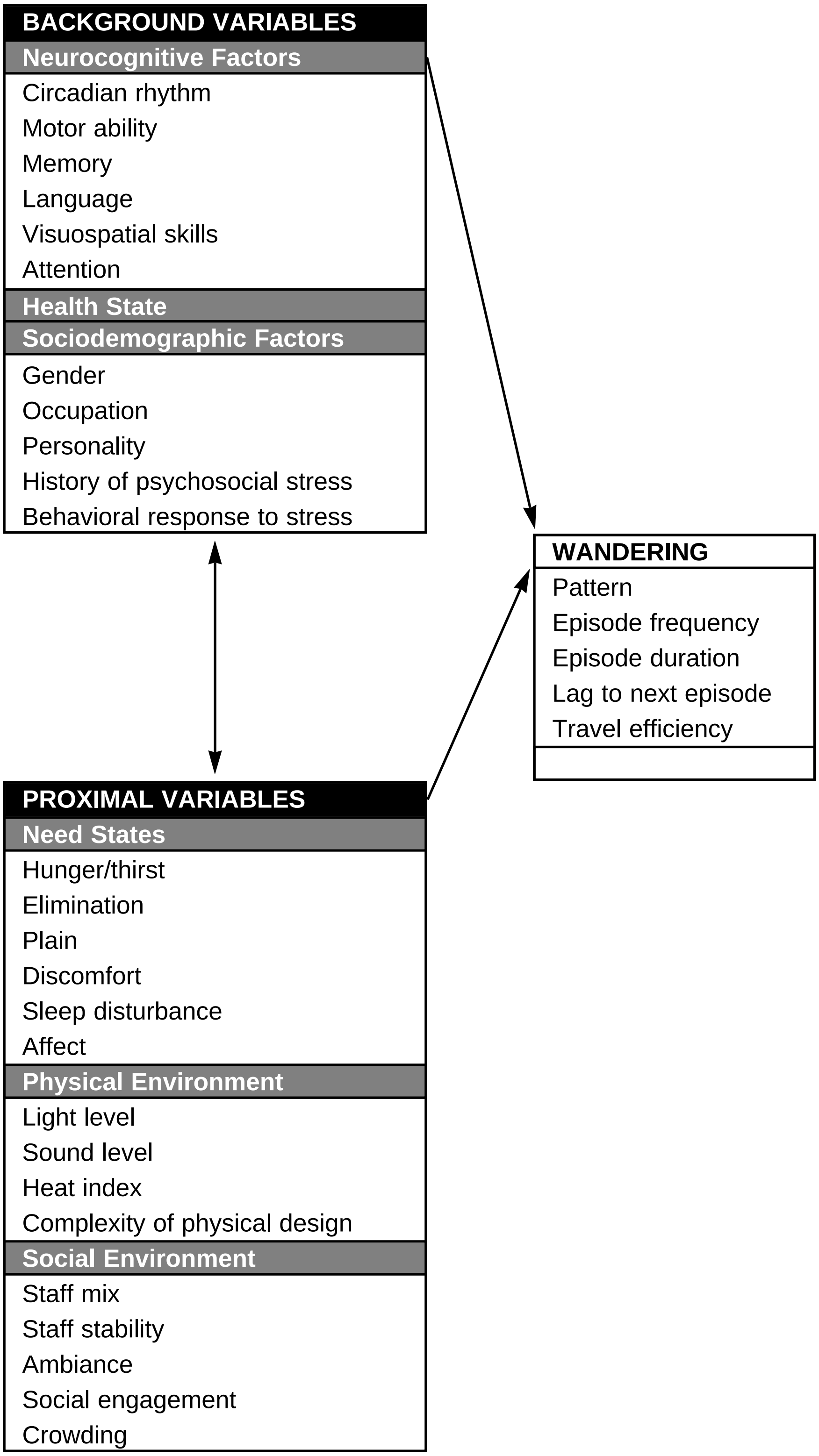
BACKGROUND VARIABLES
Neurocognitive Factors
Circadian rhythm
Motor ability
Memory
Language
Visuospatial skills
Attention
Health State
Sociodemographic Factors
Gender
Occupation
Personality
History of psychosocial stress
Behavioral response to stress
WANDERING
Pattern
Episode frequency
Episode duration
Lag to next episode
Travel efficiency
PROXIMAL VARIABLES
Need States
Hunger/thirst
Elimination
Plain
Discomfort
Sleep disturbance
Affect
Physical Environment
Light level
Sound level
Heat index
Complexity of physical design
Social Environment
Staff mix
Staff stability
Ambiance
Social engagement
Crowding

Figure 4 A tentative multifactorial model for wandering behavior.

ever, it is noteworthy in that it represents effort to advance a more comprehensive, integrated, and detailed view of this complex behavior. Further, it does so in a way that can lend support to development of specific intervention strategies for particular parameters or aspects of wandering that present problems or issues in particular contexts.

7.17.7 ASSESSMENT TECHNIQUES AND ISSUES

Most studies of wandering are actually studies of wanderers; few focus on the behavior itself and many of those actually target only exit behavior. Yet, research does provide some guidance for the clinical assessment of wandering behavior. Three measurement strategies used by researchers for quantifying wandering could be applied or adapted for the purposes of clinical assessment: mapping techniques, counting and timing strategies, and motion meters. Each is discussed in the sections that follow. In addition, sections addressing the assessment of clinical concerns and potential outcomes from wandering are offered.

7.17.7.1 Assessment Techniques for Wandering

7.17.7.1.1 Behavior mapping

The earliest study in which wandering behavior was quantified employed a behavior mapping strategy (Snyder et al., 1978). For designated time periods, episodes of locomotion were observed and traced on a floor plan of the unit where the wanderer resided. This approach captured approximate distance covered, areas of the unit where locomotion occurred, and number and heading of turns or direction changes. Robb (1987) extended this approach by adding recordings of time spent in each location and observation of several associated motor behaviors. Clinically, such data could be used to estimate relative amount of locomotion; to identify times of day when locomotion is greater or lesser than others; to document changes over time, including comparisons pre- and postintervention; and possibly to provide clues regarding both environmental stimuli that precipitate or repel wandering activity and personal factors that may influence the nature of one's wandering, such as lateral neglect. Supplemented by notes that capture other concurrent events or behaviors, mapping could reveal much about an individual's wandering to assist the selection or development of an individualized treatment program. However, this mapping approach does not differentiate wandering from nonwandering locomotion.

7.17.7.1.2 Counting and timing strategies

Researchers have also attempted to quantify wandering by counting and timing individual walking episodes (Algase & Cheng, 1992a; Algase & Tsai, 1991). Using a rhythm perspective, they cast each episode of wandering as a cycle with locomoting and nonlocomoting phases. By timing each phase, various parameters, such as cycle period (locomoting plus nonlocomoting phase) or percentage of cycle locomoting (locomoting phase divided by cycle period) can be calculated. By aggregating observations over specific time periods (e.g., hourly), additional parameters, such as cycle frequency or mean locomoting phase, can also be determined and graphed for a picture of variation over time. In later work, Algase combined her rhythm view with Martino-Saltzman and associates' typology of travel patterns to differentiate wandering from nonwandering locomotion and to compare cycle parameters for different wandering patterns (Algase, Beattie, Kupferschmid, & Beel-Bates, 1997). An example of a 24 hour graph of data resulting from this approach is shown in Figure 5. This strategy has also been combined with the recording of related behaviors, such as activity prior to and at the conclusion of the locomoting phase of each cycle (Algase & Cheng, 1992b).

The counting and timing of locomotion episodes, combined with the use of Martino-Saltzman's patterns and observation of antecedent and consequent activity or events, can provide powerful information, both scientifically and clinically. While researchers have observed continuously over time periods ranging from 2 to 24 hours and recorded data on portable terminals, a useful clinical picture can be obtained using less costly and technologically sophisticated approaches. By limiting observations to ten minutes per hour over a three-day period, a good baseline can still be documented for identifying times of relatively high and low activity and for comparing changes over time or in response to interventions (Algase & Struble, 1992). Low-tech stop watches and paper and pencil recording approaches can replace more sophisticated and costly computerized approaches. An individual's dominant wandering pattern, which may give cues to an underlying etiology, can also be determined. Further, by including observations of antecedents and consequences, assessment data can be used to inform the design of individualized behavior management strategies.

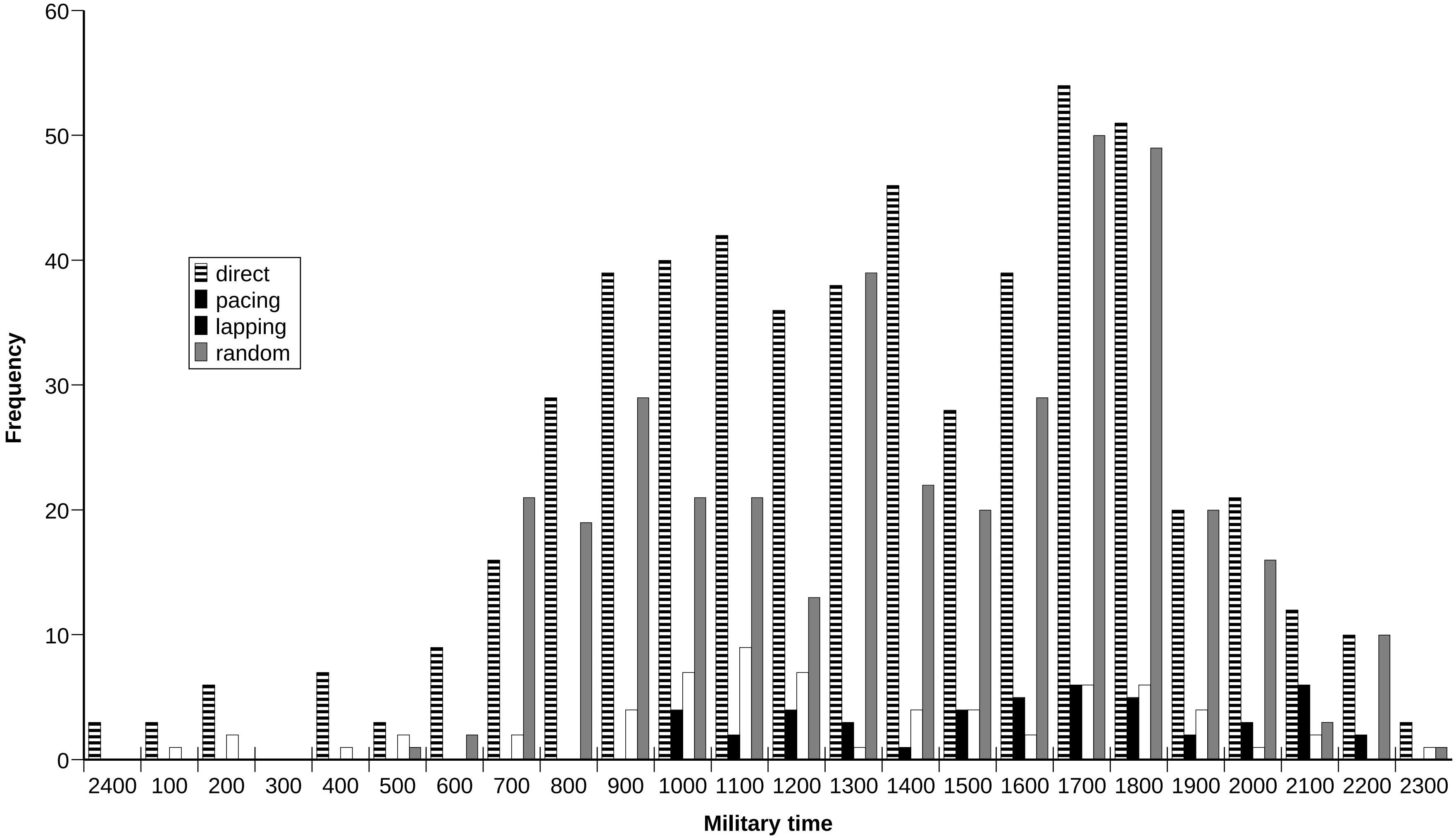

Figure 5 24-hour distribution of wandering episodes by pattern (N = 25).

7.17.7.1.3 Activity monitors

Cognizant of the expense and reliability issues associated with direct observation, researchers are beginning to test biomechanical motion detectors and pressure-sensitive counters for their ability to estimate wandering behavior (Algase, Beattie, Kupferschmid, & Beel-Bates, 1997; Algase & Cheng, 1992a; Madson, 1990). Although only in the preliminary stages of investigation, these devices hold promise for clinical as well as research purposes.

Large-scale integrated (LSI) monitors, watch-sized devices worn on the wrist, waist, or ankle, record changes in position by tripping an internal mercury switch. Movements are aggregated over time and registered in increments of 10 units or 160 movements. LSIs, originally developed for epidemiological research to quantify physical activity in the general population, are effective for aggregating frequency, but not intensity of movement (LaPorte et al., 1979). Because of their sensitivity, LSI monitors are considered well suited to assessing activity associated with skeletal muscle movement in elderly and/or disabled persons (LaPorte et al., 1983). Reliability estimates for LSIs (correlating day-to-total and evening-to-total readings) under laboratory and field conditions across a variety of populations, including postmenopausal women, are stable at values exceeding 0.70 (LaPorte et al., 1982).

In studies of wandering rhythm spanning 24 hours (Algase, Beattie, Kupferschmid, & Beel-Bates, 1997; Algase & Tsai, 1991), LSIs were used reliably; interunit reliability ranged from 0.76 to 0.93. They provided a relatively strong index of wandering over a 24-hour period but their ability to index wandering over a two-hour period was weaker; ankle applications were more valid than waist applications (Algase, Beattie, Kupferschmid, & Beel-Bates, 1997). LSIs were more effective for indexing parameters of wandering cycles alone than for approximating these for direct cycles alone. The use of LSIs as a proxy for direct observation over short time periods cannot be recommended confidently without further study using larger samples. Further, LSIs are not capable of detecting pattern.

Since 1990 new devices have been developed for the measurement of motor output. Examples are the Step-sensor (Madson, 1990) and Actillume (Ambulatory Monitoring Incorporated). Only the Step-sensor was developed specifically to quantify wandering. Composed of a heel pad, worn in the shoe, and a hardware component attached to the outer aspect of the shoe, the Step-sensor was highly reliable for recording actual number of steps taken when tested in a sample of elderly nursing home residents (Madson, 1990). These devices, alone or in combination, may estimate rhythmic aspects of wandering by demented persons more specifically and sensitively than LSI monitors. However, none yet posses the capacity to identify wandering patterns. The reliability, validity, and clinical acceptability of these devices for assessing wandering in a sample of demented elderly is unknown and is currently under investigation by Algase and associates in a nationally funded US study.

7.17.7.2 Assessment for Associated Clinical Concerns

A number of clinical concerns are associated with wandering and these deserve thorough assessment both for their contribution to and impact on this behavior. These are discussed below in relation to necessary foundational abilities, mediating factors, and antecedents and consequences.

7.17.7.2.1 Foundational abilities

Good gait, balance, and sensory capabilities undergird the ability of persons to walk safely. Regardless of deficiencies in these abilities, cognitively-impaired persons are known to move about their environment. Because of cognitive limitations, it is seldom effective to simply remind them to use special caution or assistive devices. Thus, good assessment and attention to these abilities is needed.

In contrast to declining intellectual abilities, systematic studies have established that motor function remains normal until the final stages of AD, the most prevalent dementia (Fayden & Townsend, 1976; Koller, Wilson, Glatt, & Fox, 1984; Molsa, Martilla, & Rinne, 1984). Progressive rigidity and limb flexion eventually lead to loss of balance, impaired ambulation and immobility, but psychomotor speed, posture, gait, and coordination are generally preserved. This has important clinical implications for the assessment and management of wandering behavior. Gait and balance are the fundamental physical skills which support independent mobility and thus wandering. Any patient with a history of wandering requires both gait and balance assessment. The purpose of assessment is to document the ability of the patient to ambulate safely. The extent of preserved safe ambulation behavior is the fundamental skill, second only to memory, underpinning "safe" wandering (i.e., the ability of the patient to negotiate the environment without harm to self or others).

With the exception of PD, normopressure hydrocephalus, and cerebral infarction with focal damage to relevant motor areas, abnormal gait or balance is not likely to be dementia-related until late in the course of the disease. Thus, examination for alternative explanations should be sought and resulting causes treated if such abnormalities are identified.

In the cognitively impaired patient gait and balance assessment may require modification to compensate for the inability of the individual to respond appropriately to verbal direction or gestures. In such a circumstance, discrete observation of the patient in a natural daily setting can substitute for specific testing of gait and balance. Since evidence suggests that wandering occurs in time-frames unique to each individual, gait and balance assessment timed to coincide with the patient's most active daily period will yield the most useful information.

Few gait and balance assessment tools with established reliability and validity exist for use with the elderly, and none have been specifically constructed for assessment of the cognitively impaired elder. In the absence of specific tools, those developed by Tinetti (1986) are widely used in clinical contexts and can be readily adapted to the needs of the cognitively impaired. While comprehensive gait and balance assessment will alert staff to specific mobility skills deficits, such testing cannot predict the success of preventative and adaptive safety measures for patients who wander. Anecdotal evidence suggests that many wanderers have superior mobility skills, can negotiate complex obstacles, and will persist in hazardous wandering despite safety measures. Recent empirical evidence suggests that, while a regular exercise program can increase muscle strength and stamina in AD, regular exercise has no significant effect on overall functional ability (Francese, Sorrell, & Butler, 1997).

The Tinetti gait assessment tool (Tinetti, 1986) covers nine areas of function basic to walking, including step height, length, symmetry and continuity, path deviation and turning while walking. The Tinetti balance assessment tool (Tinetti, 1986) covers sitting and standing balance, eyes closed balance, turning balance, rising, sitting, bending, and reaching. When testing gait and balance, observe the patient in an obstacle-free hallway and continue natural observation long enough for the severely cognitively impaired patient to demonstrate the required skills without verbal cueing. Less impaired patients may be able to respond to several of the simple commands required in both tools.

With a wandering patient it is critical to document as much of the data from formal assessment as possible, along with anecdotal accounts from staff and family of preserved mobility skills. For example, many wanderers, very late in their disease process, can climb and will be observed by staff climbing over chairs, up ladders, or onto ledges and tables. In addition, some patients had occupations in which climbing, balancing, and getting underneath or between structures was part of their daily work and remains part of their activity in institutions. Documented accounts of these preserved skills can help in the planning of individual environments free from known hazards.

With the cognitively impaired patient, balance and gait testing needs to be repeated at regular intervals or when cognitive status deteriorates from previous levels. While no standard time frame for re-evaluation has been established, periodic testing ensures the regular and systematic assessment of gait and balance as cognitive ability declines and provides a baseline against which to compare current function. Since the performance of motor skills is closely related to visuospatial ability, any deterioration in sensory capabilities or environmental change will also impact gait and balance, at least in the short term, and repeat testing should be considered.

In cognitively impaired wandering patients, pre-existing sensory loss of any type can complicate assessment. In addition, sensory loss interferes with communication and leaves fewer uncompromised options available to both subject and assessor. In those with cognitive impairment, even minimal assessment of sensory function is challenging and is based more on the subject's ability to function in the environment than on the results of objective testing. Wandering patients may not be able to remain still for long enough to complete formal testing of sensory function.

People with AD experience visual losses, specifically problems with color vision, depth perception, eye movements, contrast sensitivity, and higher-order visual perception (M. F. Mendez, M. A. Mendez, Martin, Smyth, & Whitehouse, 1990). In assessing vision, one looks for indications that the patient may be reading signs or books aloud, or that they appear clumsy and are bumping into furniture.

Hearing assessment in the wandering cognitively impaired patient relies heavily on knowledge of the patient's prediagnostic hearing ability. Wanderers frequently do not respond to voice instructions, even when their hearing appears to be intact. This seeming oblivion may have to do with lack of recognition of their name, the inability to distinguish individual sound amongst background noise, or a high level of preoccupation with their wandering behavior. None of these possibilities have been empirically tested. Clues about hearing ability can be obtained by watching for signs that the wanderer is hearing and responding to sound. For example, the patient may obviously enjoy music by tapping the feet or hands or singing along. Where there is existing hearing loss, three simple clinical tests using a tuning fork may help determine the type of hearing loss—the Rinne, Weber, and Schwabach tests, but each of these is highly dependent on the patient's ability to respond verbally to commands (Gallo, Reichel, & Andersen, 1995).

Cognitively impaired patient may experience deficits with taste and smell, but are often unable to report them. Wanderers will frequently appear with bruising related to bumping into things and have lost the ability to determine danger where there is hot water or electrical appliances. Some cognitively impaired persons also develop pica and many persons with late-stage AD will eat voraciously without regard for taste.

Crude assessment of kinesthetic sense in wanderers is possible during gait and balance testing by observing for ability to maintain sitting and standing balance and to recover from a gentle push against the shoulder. Fear of falling is apparent when a patient reaches out or braces, even when induced disequilibrium is slight. Wanderers are generally at less risk of kinesthetic deficits than less active or immobile patients because frequent walking stimulates proprioceptive and vestibular function.

Although assessment of sensory abilities in all modalities is important for safety considerations, anecdotal evidence gained from observations during our research suggests that some deficits may actually contribute to particular patterns or rhythmic parameters of wandering. However, the specific contribution of various sensory deficits to aspects of wandering behavior remains to be determined.

7.17.7.2.2 Mediating factors

Unlike foundational abilities, which are necessary for safe ambulation but may not actually affect the amount or type of wandering, mediating factors may play an active role in shaping this behavior. These factors include: cognitive strengths and deficits; mood, affect, agitation, physical discomfort; and personal patterns and preferences. As assessment of cognition, mood, affect, and agitation are the focus of other chapters, our discussion here is limited to personal patterns and preferences and physical discomfort.

To determine a wanderer's personal patterns and preferences, a personal interview and discussion with the patient and the primary informant is essential. Since patterns are laid down over time and stored in remote or procedural memory, many persist in the presence of recent memory loss and can be understood in the context of the patient's prediagnostic lifestyle and occupation. For example, a patient who once owned a chicken farm spent the early morning out in the yard making "shooshoo-ing" sounds and gesturing as if throwing seed down. Any attempt to interrupt the patient during this time resulted in an aggressive outburst. This type of habit behavior has been described as "agenda" behavior (Rader et al., 1985). Knowledge of such patterning can assist in understanding the origin and role of a behavior in a patient's life and bring new meaning to seemingly irrational actions.

Preferences, too, are long-standing and thus likely preserved late into a dementia. One of the most enduring preferences is that of food and many patients in advanced stages of the disease cannot choose between foods but will reject food they have always disliked. The role of personal patterns and preferences for daily routines, food, clothing, and any other aspect of the patient's environment is needed to understand their role in shaping wandering behavior. Such evaluation can greatly inform the design of effective targeted intervention strategies.

Although evaluations of efforts to preserve enduring life patterns and preferences for effects on parameters of wandering behavior are lacking, attention to such preferences has been used to reduce the problems arising from wandering. In a case study by Donat (1984), often repeated behavioral patterns were accommodated by reassigning a patient to the room most frequently entered and moving its then current occupant to new quarters. For the wanderer, this room reassignment eliminated exposure to the unhappy response her intrusive behavior precipitated from staff and the other resident; the gain for the other resident was exemption from unwanted intrusions into her private space.

Ability to accurately identify, personally address, or effectively communicate physical discomfort is hampered by cognitive impairment in wanderers. Increased restless or pacing behavior can signify such meaning. Common discomforts can be avoided or reduced in most cognitively impaired patients using systematic toileting, feeding and resting schedules. Food acceptance charting and routine weighing can aid in assessing hunger. Enforced daytime resting and napping, while offering respite from walking, may exacerbate circadian rhythm disturbance and create an artificial "need" for restraints. Observation and recording of the frequency and length of daytime napping and night sleep may assist in determining fatigue.

Although many wanderers are in better physical condition than nonwandering counterparts, they are not immune to arthritis, angina, and other acute and chronic conditions which cause pain. Additionally, wandering, by virtue of high activity levels, can aggravate certain conditions, such as arthritis, and precipitate others, such as foot blisters. Pain may account for noticeable changes in activity level.

Two instruments can assist in the assessment of discomfort in wanderers. (i) The Discomfort Scale-Dementia of the Alzheimer's type (DS-DAT), developed to assess for nonspecific discomfort (Hurley, B. Volicer, Hanrahan, Houde, & L. Volicer, 1992). The DS-DAT contains nine behaviors which reflect presence or absence of discomfort. Good construct validity (coefficient alpha = 0.77, $p < 0.001$) and test–retest reliability (0.60, $p < 0.001$) were reported on data from serial observations of 20 subjects with conditions known to cause discomfort. Inter-rater reliability ranges from 0.86 to 0.98 across four observations of five to nine subjects (Beck, personal communication, 1997). Again, this instrument was developed for the purpose of evaluating behavior captured in natural settings on videotape, but could be readily adapted for clinical use. (ii) An instrument for assessing pain that may be appropriate for use with wanderers is the Wong–Baker Faces Pain Rating Scale (FPRS). Useful for rating both presence and severity of pain, this series of six line-drawn faces was developed to evaluate pain in children who are at least three years old (Wong & Baker, 1988). In a sample of 97 adults (aged 21–65 years), Cason and Grissom (1994) report that the FPRS correlated with perceived pain intensity ($r = 0.79$) and a visual analog scale (0.67). Although reports of use with demented adults were not located, the fact that the FPRS is targeted to three-year olds suggests that it does not exceed the cognitive capacity of most demented elders (Thornbury, 1992).

7.17.7.3 Assessment of Wandering Outcomes

Outcomes most frequently associated with wandering are falls and elopement. Other outcomes of clinical interest include social functioning, weight loss and nutritional status, sleep and fatigue. Assessment concerns applicable to each are discussed in the following sections.

7.17.7.3.1 Falls

The management of highly mobile, wandering nursing home residents is a particular challenge because their increased mobility levels predispose them to increased extrinsic (e.g., tripping) and intrinsic (e.g., dehydration, fatigue) hazards (Berg & Cassels, 1990; Whall, Gillis, Yankou, Booth, & Beel-Bates, 1989). While mobility is crucial to functional ability and continued well-being during aging, associated falls are critical incidents for the elderly because of their serious immediate and long-term effects (Hogue, 1984; Rubenstein et al., 1994; Vallarino & Sherman, 1981). Falls by the elderly, "or events in which the patient came to rest on the floor from a lying, standing or sitting position" (Heslin et al., 1992) are widespread and well documented. Etiology, risk factors, and management are common research foci.

Limitations in mobility place cognitively intact individuals at higher risk of falls and fractures and the complications associated with them. A higher incidence of falls has been demonstrated among the institutionalized aged (25–61%) than among the community-based aged (Campbell, Reinken, Allan, & Martinez, 1981; Foerster, 1981; Hogue, Studenski, & Duncan,1990; Morris & Isaacs,1980; Overstall, 1978; Sehested & Severin-Nielsen, 1977; Tinetti, 1987; Tinetti et al., 1988; Wild, Nayak, & Isaacs, 1981). However, fall rates and fall-associated injuries vary greatly among and between hospitals and long-term care facilities and are highly dependent on fall reporting methods (Kanten et al., 1993). An epidemiological review of all large-scale published nursing home studies determined a mean incidence rate of 1.5 falls per bed per year (range 0.2–3.6 falls) (Rubenstein et al., 1994). Nursing home studies of the outcome of falls identified serious injuries such as hip fractures, cerebral contusions and hematomas as occurring in 3–6% of the subjects who fell (Morse et al., 1985, 1987; Tinetti, 1987).

The etiology of falls can be characterized as either extrinsic or intrinsic (Berg & Cassels, 1990). Extrinsic factors are general environmental factors, for example, lighting levels, time of day, placement of furniture, type of floor coverings, cleanliness of floors, and type of footwear worn. Critical intrinsic factors include the aging process itself, the presence of disease, increased general frailty, impaired postural control, limitations in mobility, poor vision, psychotropic medication intake, and gender (Barbieri, 1983; Cook, Exton-Smith, Brocklehurst, & Lampert-Barber, 1982; Louis, 1993; Tinetti, 1986; Wild et al., 1981). The relative contribution of each of these intrinsic factors to both fall risk and frequency has not been clearly elucidated.

Dementia can contribute to falls (Buchner & Larson, 1987; Janken et al., 1986; Morse et al., 1987; Nevitt et al., 1989). In a number of studies related to falls in nursing homes, cognitive impairment is the most important and frequently documented risk factor identified (Buchner & Larson, 1987; Reinboth & Gleidenvand, 1982; Rubenstein et al., 1994; Tinetti et al., 1988; Walshe & Rosen, 1979). In several studies, cognitive changes are associated with impaired gait, impaired balance, and wandering (Kippenbrock & Soja, 1993; Morse et al., 1985, 1987). Although thought to be an important factor in falls, the association between wandering and falls has not been well documented.

Conversely, elderly residents rated by nursing staff as alert had the greatest number of falls (Kalchthaler, Bascon, & Quintus, 1978) and a low incidence of mental confusion was found among residents involved in falls. Based on a study of 60 institutionalized women with Alzheimer's disease, researcher's concluded that falls were related to a decline in overall "vigor" in those who had been active (Brody et al., 1984).

While the relative contribution of critical intrinsic factors in the etiology of falls has not been clearly elucidated, the presence of a dementing process remains the most important and frequently documented intrinsic risk factor identified. Early work by Algase (1992b) examined antecedents and consequences as pacers for wandering cycles and determined that wandering is most likely to occur when the residents' activity is unfocused and solitary (52% of 2405 wandering episodes, $n=31$).

Retrospective chart auditing can establish incidence rates only for reported/recorded falls. While such data will provide a baseline, it is critical to determine actual rates more accurately by the systematic observation of at-risk patients. Examination of the at-risk patient for evidence of bruising and cuts caused by falling, and the comprehensive assessment of reports of pain-related behavior in cognitively impaired wandering patients, for example, wincing and limping, can reveal that a patient has had an unseen fall. The early identification of at-risk patients at the time of their admission to a facility, including previous fall history, drug history, and physical examination, is recommended (Brungardt, 1994).

7.17.7.3.2 Elopements

Virtually all clinicians who work with wanderers are aware of elopement as a general hazard and have experienced the wiliness and persistence of some wanderers in successfully escaping the boundaries of their prescribed area. However, there are also those wanderers whose exits are accidental and whose spatial disorientation prevents them from recognizing that they are lost or disables retracing their steps for a safe return. While researchers have counted a variety of behaviors to signify elopement attempts (approaching exit doors, touching or manipulating door knobs, asking to leave, etc.) (Gaffney, 1986; Hussian, 1982; Namazi, Rosner, & Calkins, 1989), specific behavioral predictors for either "type" of elopement are unknown. Observing both antecedents and consequences of an individual's exit attempts can illuminate their unique pattern and thus inform the design of individualized interventions. Due to nursing home reforms contained in the Omnibus Budget Reconciliation Act (OBRA), the use of restraints to prevent elopement is increasingly challenged, legally and empirically (Evans & Strumpf, 1989). Strategies that are less restrictive and operate more passively, such as alarms, are preferred.

7.17.7.3.3 Social functioning

While it has been suggested that wanderers were formally more sociable than nonwanderers (Monsour & Robb, 1982), the social functioning of wanderers is severely inhibited by their marked memory problems, impaired linguistic communication skills, and disinhibition. Assessing social behavior in the severely cognitively impaired is best accomplished by observation of the frequency and quality of interactions with staff, visitors, and other residents. The Minimal Social Behavior Scale (Dastoor et al., 1975) can be used to determine preserved aspects of 1:1 social interactive ability but does not address aspects of social behavior such as the ability to listen and respond appropriately to conversation, the ability to share an activity with others, and the ability to initiate social behavior.

As wanderers become more severely impaired, their ability to reach out and connect with others diminishes. Both staff and other patients frequently find wanderers irritating to deal with because they are forgetful, no longer adhere to social conventions, and unwittingly invade the personal space and privacy of others. Wanderers were the least favored patients in several surveys of caregivers (Lanceley, 1985; May, 1993). Surveys of caregivers and intimates of dementia sufferers indicate that visiting and spending time with a wanderer is a frustrating and difficult experience. The outcome of this potent combination of factors is the further isolation of the wanderer from others in the environment, which may impact the length of time a wanderer spends alone, which may in turn impact parameters of wandering behavior.

The comprehensive assessment of social functioning includes a detailed account of the patient's previous social network and family system and descriptions of the nature and purpose of current relationships (Zarit, Perlin, & Schaie, 1993). A wanderer is limited socially by several factors: (i) the inability to control or modulate their own activity levels and (ii) the loss of linguistic communication ability to the extent that others have difficulty listening to, interpreting, and responding to the wanderer's speech. Thus it is difficult for a wanderer to initiate and maintain social interaction unless another person is able to keep pace with them and tolerate a generally increased activity and a high level of impaired communication.

To date, most research on linguistic communication in dementia has focused on the identification and etiology of deficit rather than preserved functional skills. Several studies support the position that severely cognitively impaired nursing home residents may be capable of greater functional communication competence than has been thought previously (Murray, Marquardt, Richardson, & Nalty, 1984; Slauson, Bayles, & Tomoeda, 1987). Assessment of the preserved linguistic communication skills of the patient will assist in determining the type and nature of interaction the patient has the capacity to take part in. Several scales have been specifically developed for this purpose, including the Communication Observation Scale for the Cognitively Impaired (Tappen, 1988) which focuses on broad categories of impaired function such as the presence or absence of aphasia, blocking, perseveration, talkativeness, meaningless sentence production, and poverty of speech. Since problematic wandering appears to be associated with more severe levels of cognitive impairment, the majority of persistent wanderers will tend to have marked linguistic communication deficits.

A prime function of staff is to nurture the enduring relationships the patient has and to instruct others in ways to spend meaningful time with the patient. One simple way of doing this is to encourage others to walk with and accompany a wanderer, rather than expecting a wanderer to sit down and visit. Maintaining simple conversation focused in the here and now can assist in orientating the wanderer and increasing the interactive climate and expectations of the patient to remain social. However, little is known about the impact of expected social interaction and interactive pressure on levels of anxiety in wanderers, or on pattern and rhythm of wandering.

7.17.7.3.4 Weight loss and nutritional status

Patients with dementia have difficulty obtaining, preparing, and eating food. Patients who wander tend to have a higher lean muscle mass than less active patients, while requiring a higher calorie intake to support wandering activity. While overnutrition can be a problem early in the dementing process, the more likely scenario for a wanderer in the long term is a profile of undernutrition. Overnutrition may result from the inability of the patient to choose appropriate foods and uncontrolled eating. Undernutrition related to wandering is multifactorial resulting from the inevitable wasting related to dementia, the inability of the patient to sit still long enough to eat and drink adequately, changes in the condition of the mouth and teeth, increased activity levels occurring simultaneously with interrupted sleep patterns, and the inability of the patient to manipulate utensils. Severely cognitively impaired individuals may also spit, gag and stuff their mouths with food.

It is essential to calculate the weight and body mass index of the wandering patient and compare these measures to normative measures. The weight of a wanderer requires monitoring at regular intervals to prevent significant weight loss of going unreported. The nursing home standard of monthly weightings may be insufficient for wanderers. A dietary history including known allergies, food preferences, and usual eating times will assist in food selection and meal planning. Wanderers require increased calorie and fluid intake to balance the energy depleted in excessive motor output. A simple pedometer secured on the patient's clothing can give an estimate of miles traveled during waking hours and permit the calculation of calorie requirements for the patient in order to maintain current body weight. However, gait quality may affect the accuracy of pedometers as most work on impact and a more shuffling gait is common among wanderers. Assessment of the mouth and teeth, and correction, where possible, of mechanical problems associated with eating, will facilitate intake.

7.17.7.3.5 Fatigue and sleep disruption

Sleep disruption routinely occurring over extended time periods results in fatigue. In the wandering patient, fatigue levels may also be increased by persistent levels of physical overactivity. A characteristic anecdotal activity of wanderers is day-time napping but little empirical evidence exists to support this (Snyder et al., 1978). Sleep observation charting, at intervals of 30 minutes, will provide a reasonably accurate indication of the pattern of night waking, total uninterrupted sleep time, and daytime napping experienced by a wanderer (Bliwise, 1994), however, this method is staff intensive. Other methods of sleep assessment are both costly and cumbersome and therefore not yet clinically viable with this population.

7.17.8 INTERVENTION STRATEGIES

While the literature on wandering is certainly not vast, almost as many papers address interventions as discuss all other aspects of the behavior. Many suggested interventions have only clinical evidence to support their value. However, some have been tested empirically, but only a few of these have been refined through replication studies. Unfortunately, whether based on clinical experience or research, few interventions are theoretically driven or tied to putative etiologies for wandering. In the following sections, goals for treatment and intervention of wandering, as well as general health and safety considerations for wanderers, are outlined. Information on the range of tested and untested strategies and interventions are then summarized.

7.17.8.1 Goals for Treatment and Intervention

Overall goals for treatment and intervention with wanderers are: (i) assuring safety, (ii) preserving abilities, (iii) maximizing comfort, (iv) minimizing restriction, and (v) eliminating or reducing negative effects of wandering for the wanderer, other residents, and staff. Achieving these goals may entail efforts to reduce, redirect, contain, or redistribute wandering behavior itself. Further, strategies for achieving one goal can, at times, present challenges for attaining another. Prioritization of goals, which should include known personal preferences of the patient and involvement of both family and caregivers, can assist in resolving such conflicts. However, since knowledge of the etiology and purpose of the behavior remain limited at this time, measures designed to either restrict or facilitate the amount of wandering have little empirical basis and should be undertaken with extreme caution. A general rule of thumb is that such measures may be practiced if they do no harm, but careful monitoring to assure such is mandatory.

7.17.8.2 General Health and Safety Considerations

Sustained effort needs to be made to maintaining the physical condition and emotional well-being of the wandering patient. All possible measures for environmental safety need to be instituted to minimize or eliminate hazards associated with falling and elopement. In addition some simple environmental manipulation may assist the wanderer in way-finding and environmental interpretation.

Overall goals for maintaining the physical condition of wanderers are (i) ensuring adequate hydration, (ii) ensuring adequate nutrition, (iii) preventing exhaustion, (iv) minimizing circadian rhythm disturbance, (v) preserving skin integrity, and (vi) monitoring the dosage and undesirable effects of prescribed medications. Overall goals for promoting emotional well-being for the wanderer include (i) engaging the person in meaningful, supervised, and structured activity which provides sensory stimulation; (ii) creating regular opportunities for 1:1 purposeful communication at the person's level of comprehension; and (iii) preserving elements of previous meaningful life activities.

Overall goals for ensuring the physical safety of the wanderer are (i) removing safety hazards from the environment, (ii) enhancing lighting, (iii) preventing elopement, and (iv) balancing the level of stimulation in the environment. Since wanderers are cognitively impaired, aspects of the environment will grow increasingly unclear to them. Measures which may improve the functional ability of wanderers in negotiating their environment include (i) providing redundant cueing; (ii) decorating rooms with relevant personal items and photo identification; (iii) displaying orienting information in multiple, prominent places; and (iv) using rooms for a single, clear purpose.

7.17.8.3 Tested Techniques for Altering Wandering Behavior

7.17.8.3.1 Exercise

On the premise that directed energy expenditure could dissipate or reduce spontaneous wandering, Robb (1987) tested a program of exercise. In a quasiexperiment involving a multiple pre- and post-test design, the hypotheses that exercise would act to reduce the length of time spent in wandering, reduce overall energy expenditure, and improve night time sleep behavior in subjects with a medical diagnosis of dementia were tested. Twenty male subjects (aged 50–87 years) received 12 weeks of programmed outdoor exercise for two hours per day five days per week. Elements of wandering behavior, for example, change in location, time spent in each location, and limb and head movement, were measured during the day and at night prior to the intervention, at eight weeks into the intervention, and six weeks postintervention using behavioral mapping. All subjects received initial fitness and cognitive testing. Effort-induced physical harm was prevented with pulse checks and close observation.

Planned exercise had no significant effect on daytime wandering. However, reductions in night-time wandering were significant for number of locations visited and total time in motion. Serendipitous improvements in range of motion and energy expenditure were also reported. Despite its promise for reducing night-time wandering, no replication studies or clinical reports using exercise were found.

7.17.8.3.2 Cueing

The use of signs, whether composed of words, pictures, graphics, or familiar objects or pictures, has been widely advocated clinically to improve the spatial orientation of cognitively impaired persons. While this strategy has not been examined with regard to effect on overall amount or type of wandering behavior, it has been evaluated for its ability to enhance spatial orientation in the cognitively impaired.

In one experimental study, baseline route navigation and four location-to-location comparative trials were used to determine the effect of sign-post cueing on navigational ability in a sample of 20 spatially-disoriented long-term care residents (Cooke, Algase, & Terpenning, 1993). The sign-posting consisted of large, primary-colored, eye-level signs in simple shapes, such as circles, squares, and hexagons, or in shapes favored by or meaningful to the resident, such as cats or flowers, and labeled in large type. Sign-posts were strategically placed in doorways and corridors directing subjects to relevant locations, for example, the dining room, bathrooms, and their own rooms. Subjects navigated each of four routes sequentially without directional prompting but with occasional verbal reminders of the target location. Time to navigate the route, hesitancies, errors, and overall points scored were recorded for each route and all routes combined. Results of the study demonstrated that sign-posting, regardless of design, did not increase navigational ability.

7.17.8.3.3 Lighting

Several empirical studies have examined the effect of light on activity levels, defining wandering as one of a number of expressions of agitation in dementia. Early work by Evans (1987) identified the phenomenon of "sundowning" in which demented individuals become more agitated and restless with declining light levels in the late afternoon. Cohen-Mansfield et al. (1989) detected temporal patterns of activity disturbance over a 24-hour day in demented patients. Seasonal variation exists in agitated behavior (Bliwise, 1994; Bliwise, Carol, Lee, Nekich, & Dement, 1993), with agitation increasing in the darker winter months. In a Japanese study (Okawa et al., 1989), both morning and evening bright light treatment deceased agitation behavior. More recently, Lovell, Ancoli-Israel, and Gevirtz (1995) administered bright light at 2500 lux levels to six demented subjects (mean age = 82 years) for two hours in the morning for repeated 10-day periods in an ABABA design. Results indicated a significant difference in agitation levels on light-intervention days.

In the only known study of light directed at wandering, three light levels were specifically manipulated (baseline, bright light, and dim light simulating sundown) in a nursing home lounge outfitted for the study (Algase, 1997a). Ten cases were studied for three two-hour periods weekly over a six-week period in random assignment to either an ABAACA or ACAABA design; the second baseline week was intended as a washout period for any carryover effects. The time period was individually selected to match each case's most active period of the day. Although data analysis is incomplete at this time, preliminary results indicate that, regardless of lighting condition, wandering in the lounge was minimal. Logically, factors other than light account for the quite dramatic decrease in wandering observed. Simplicity of the lounge environment, which was completely accessible visually from any vantage point in the room and controlled most sources of stimulation, is the speculated explanation.

7.17.8.3.4 Social interaction

Interventions for wandering addressing aspects of social interaction are primarily related to restraint-reducing efforts. Many programs designed to increase and facilitate interactive social behaviors in the demented give anecdotal descriptions of the reduction in agitation and wandering produced by increased social opportunity and stimulation (Arno & Frank, 1994; Gerdner & Swanson, 1993; McGrowder-Lin & Bhatt, 1988; Mintzer et al., 1993). However, only one empirical study has so far explicitly sought to impact wandering.

Goldsmith, Hoeffer, and Rader (1995) report a test of Rader's (1987) "approach" to wandering (which includes steps to identify potential wanderers, programs and activities for safe ambulation, decreasing use of restraints and psychotropic medications, interaction strategies for managing exiting behavior, creating a safe and supportive physical and psychosocial environment) taught to nursing staff. The study used a single case design with a four-week preintervention phase, a four-week intervention phase, and a six-week postintervention phase. Use of medications and restraints decreased during the intervention phase, but medication use began to rise postintervention. Decreases were associated with increased Physical Self-maintenance scores. This approach had no demonstrable effect on problematic wandering (ambulating alone, exit-seeking while alone, alone in unsafe locations). During all phases of the study, 80% of observations documented a lack of social interaction between staff and subject, except during ADLs. The study was poorly controlled in that once staff were trained, it is theoretically possible, but unlikely, that they did not revert to their pretraining behavior during the postintervention phase of the study. The extent of reversion to pretraining habits and the amount of retraining required over time were not considered.

A recent study utilized a multiple case study design to examine the effect of a systematic, structured nurse–resident communication style on pattern and rhythm of wandering behavior in severely cognitively impaired nursing home residents (Beattie, 1996). An identical quasi-experiment was embedded within each case study as an interrupted time-series with multiple replications and nonequivalent dependent measures. The intervention was a modification of the Heron Six-category Intervention Analysis (Heron, 1976). Results were highly individual across cases, showing no consistent significant effect of the intervention.

7.17.8.3.5 Drug therapy

Older people have qualitative and quantitative differences from younger people in absorption, distribution, metabolism, and excretion of drugs. The wide use of drugs to control difficult behaviors associated with dementia has increased the potential for negative drug reactions. In the cognitively impaired patient, the management of drug therapy is a particular challenge because patients are frequently unable to report undesirable effects, and many patients have multiple pathology requiring concurrent medication.

The dangers of polypharmacy in the elderly are well known. In the cognitively impaired patient the general rule of "less is better" usually applies. A minimalist approach to medications is particularly important when trying to manage difficult behavioral symptoms and tease apart factors contributing to them. In addition, many cognitively impaired patients have difficulty accepting and swallowing medication and compliance must be carefully monitored. Regular and systematic review of medications can prevent some problems and minimize the impact of others. Wherever behavioral change occurs unexpectedly, or unusual behavior persists over a time period longer than a week, it is essential to consider the role of current drug therapy in the clinical presentation of the patient.

Two groups of drugs that are commonly prescribed in to patients with cognitive impairments may have a therapeutic effect on wandering: neuroleptics and benzodiazepines. The neuroleptics are most commonly prescribed for the treatment of agitation and hyperactivity, and are often differentiated on the basis of their undesirable effect profiles. The two major classes of neuroleptics are the phenothiazines (e.g., chlorpromazine and thioridazine) and the butyrophenones (e.g., haloperidol). In general the phenothiazines tend to be quite sedating and are associated with hypotension. Thus in the wandering patient the risk of circadian rhythm disruption, increased gait instability and fall risk, and hypothermia secondary to accidental elopement must be considered. The butyrophenones cause less sedation and few hypotension problems but are more frequently associated with neuroleptic-induced Parkinsonism, characterized by drooling, mask-like facial expression, muscular rigidity, and gait changes. Tardive dyskinesia (TD), an irreversible syndrome involving involuntary movements of the tongue, mouth, face, and neck, is usually associated with long-term neuroleptic use. In some cases, agitated behavior can be exacerbated by neuroleptic treatment so drug treatment should be conservative and time-limited. Over 90 studies of neuroleptic treatment of agitation in the elderly have been reported. Salzman (1987) and Sunderland and Silver (1988) report very similar levels of acceptance and behavioral control, with 70–80% of patients having none or minimal undesirable effects.

The benzodiazepines (e.g., diazepam, oxazepam, alprazolam, lorazepam) are frequently prescribed for symptoms of insomnia and anxiety in the elderly. To the extent that wandering is based on anxiety, they may be effective in its management. There are many age-related changes in response to this group of drugs, primarily due to pharmicokinetic alterations in drug metabolism which result in extended duration of action and lead to an accumulation of side effects (Greenblatt, Divoll, Abernethy, Harmatz, & Shader, 1982). Lorazepam has an uncomplicated metabolic pathway, but has been associated with specific memory impairments (Preston et al., 1988) and is thus a poor choice for the cognitively impaired patient.

7.17.8.4 Tested Techniques for Addressing Undesired Outcomes

Although several putative outcomes of wandering, discussed earlier, are thought hazardous, only one has received the attention of researchers. Reduction or control of egress has been the chief objective of such research. While not concerned with affecting the overall amount, distribution, or type of wandering, these studies focused mainly on reducing successful exits or exit attempts of known wanderers from their assigned unit or permitted range for free movement. The extent to which such exits or attempts occur during actual wandering, as compared to other self-directed locomotion, i.e., direct ambulation, remains unexamined. Three techniques have been tested: behavior modification, camouflage, and perimeter control devices.

7.17.8.4.1 Behavior modification

Principles of learning are frequently incorporated into planning wandering management, for example, rewarding constructive behavior and changing the consequences of the behavior (Hussian & Davis, 1985; Stokes, 1986). However, there are few examples in the literature of controlled, systematic behavior modification. Early work by Hussian (1982) investigated the effect of controlled stimuli and rewards on wandering patterns. The wandering behavior of three subjects (mean age = 73.4 years; mean length of hospitalization = 13 months) was conditioned using two sets of cardboard stimuli: orange rectangles and blue circles. Attention to the orange rectangles resulted in reinforcement with food; attention to the blue circles resulted in the aversive stimulant of an unpleasant noise. During conditioning, baseline data on ambulation behavior were obtained.

Following conditioning, the stimuli were affixed to walls where ambulation was either encouraged (orange) or discouraged (blue). Subjects were systematically observed for periods of 30 minutes. If no contact with the blue areas was made, subjects were rewarded with food. If contact with the blue areas occurred, subjects were returned to their point of origin without consequence. Over time the stimuli were faded by size reduction and eventually removed from the environment.

Mean number of entries into the designated blue trespass area was reduced across all subjects following conditioning by rates ranging from 13 to 100% per 30-minute period. However, booster sessions were required by all subjects at three months and the stimuli were replaced on the walls. Four booster conditionings were required over the 12-month period following initial conditioning. Eventually fading resulted in re-emergence of wandering to preintervention levels. While results of this study clearly demonstrate that wandering can be altered by environmental cueing, the effect on wandering will persist only as long as the cueing is maintained. Given the pattern of known decline in memory and visuospatial deficits in AD, the frequent reconditioning needed to maintain cueing effectiveness is not surprising. Using this method of behavioral control, cueing may need to be increased in intensity and frequency as cognitive functioning declined in order to achieve the same behavioral change. Further studies replicating Hussian's work would benefit from greater specification of level and stage of AD and finer gradation criteria for the fading of stimuli.

7.17.8.4.2 Camouflage

Problems with color vision, depth perception, eye movements, contrast sensitivity, and higher-order visual perception are known visual deficits associated with Alzheimer's disease (Mendez et al., 1990). Hussian and Brown (1987) proposed that these deficits might result in a misperception whereby a two-dimensional grid pattern, fixed to the floor with tape, would be seen as three-dimensional and, thus, unstable. If so, demented subjects may be hesitant to walk upon it and, if strategically placed in front of doorways, grids might reduce exits and exit attempts.

The effect of several different grid patterns made of low-contrast tape when applied to the floor in front of emergency doors were examined for effect on exiting behavior in a sample of eight male subjects with senile-onset, primary, degenerative dementia. Baseline data on attempted exits indicated that an attempt was associated with exit-door contact in 98% of observations. All grid patterns were effective in reducing exit attempts for all but one subject, who was never deterred by any pattern. Subjects did not cross the grid in 70% of their exit attempts.

Several researchers have attempted to replicate Hussian and Brown's results using a variety of color combinations and tape patterns and improvements in design of the original work (Chafetz, 1990; Hamilton; 1993; Hewawasam, 1996). Most were ineffective or results were equivocal. Inability to achieve success was attributed to factors such as confounding effects of medication or strength or "dose" effect of the grid intervention (i.e., grid insufficiently aversive to overcome enticing view through glass doors). However, grid patterns may be effective for patients diagnosed with AD due to associated problems in depth perception, but are not successful with other dementias (Hewawasam, 1996).

Based on the visual agnosia associated with dementia, visual barriers (i.e., door knob coverings or other disguising techniques that appear to create an obstruction, but do not actually impede egress through a door) have been suggested as an alternative camouflage strategy. Namazi et al. (1989) tested the effect of seven different visual barriers on resident exit behavior on an Alzheimer's unit. The subjects ($n = 9$; 55% female; mean time since diagnosis = 5.75 years) had similar cognitive abilities as measured by the MMSE (Folstein et al., 1975) and the Short Portable Mental Status Questionnaire (PMSQ) (Johnson & Grant, 1985). Barriers, which varied in color, material, width, angle of placement, and purpose, disguised both the exit and the door knob. Doorknob concealment was 100% effective in preventing exit behavior in all subjects. However, visual-spatial testing of subjects was not standardized and the exit behavior of subjects and their anxiety levels were not discussed.

In a replication of Namazi's work, the possible attraction of windowed doorways, which may have confounded efforts to replicate the original grid studies, was taken into account (Dickinson, McLain-Kark, & Marshall-Baker, 1995). A cloth panel matching the emergency door color and covering its panic bar was the visual barrier tested. In addition, a miniblind was used to control the view from a window in the door. A sample of seven ambulatory demented subjects (five male; two female; aged 70–87 years) with a history of exiting behavior were chosen from one dementia care unit; no other cognitive characteristics of the sample were reported. Three conditions were tested during a fixed time frame: closed miniblind alone, cloth panel alone, and both miniblind and cloth panel. After obtaining a one-week baseline count of exit attempts at the door, each test condition, which lasted for two hours per day over a one-week period, was followed by a one-week interval during which the blind was left open, no barrier was in place, and no counts were taken.

Overall, the miniblind alone reduced exit attempts by 44% from the baseline while the cloth panel reduced attempts by 96%. Together, the blind and panel were somewhat less effective (88%), suggesting a learning effect. When examining cases individually, four of the seven were found to have had five or fewer exits attempts during the baseline period. Of these, the blind alone was 100% effective for two cases while the panel alone was 100% effective in all cases. However, two of these four subjects returned to or exceeded their baseline count when the blind and panel were used together. The remaining three subjects accounted for 87% of the baseline exit attempts. The miniblind alone was 43% effective; however, the combined miniblind and panel remained nearly as effective (94%) as the panel alone (95%). Thus, the learning effect among these three cases was not as great. Only one of the three demonstrated an increase in exit attempts from the panel to the combined condition, but the increase was small (from 0 to 4) and still represented less than 10% of the subject's baseline count. Characteristics distinguishing those who made many exit attempts from those who made few were not reported. However, these characteristics may be important to delineate. Since these data suggest that those making fewer attempts may be more difficult to dissuade over time, such characteristics may help identify those whose exit attempts may persist and lead to harm.

7.17.8.4.3 Perimeter control devices

Reports of devices such as alarms, fences, and locks to control wandering began appearing in the literature in the 1970s (Miller, 1976). Many commercial electronic and computerized products now exist to provide perimeter control for wanderers, on the assumption that such devices will warn staff of an impending unsafe exit, or prevent such an exit. Devices generally operate in one of three ways: sounding an alarm when a patient crosses an electronic barrier, creating an electronic barrier by automatic closing of gates, doors, and windows, or tracking the patient via a unit worn on the body (Leach, 1985).

In tandem with the development of perimeter control devices, there has been an increase in sensitivity to issues in the humane and compassionate confinement of individuals with cognitive impairments, such as the right to the Least Restrictive Alternative (LRA), that is, to create a balance where individuals are confined at the lowest level considered appropriate to preserve their safety while retaining the highest possible levels of freedom, independence, and dignity (Ebersole & Hess, 1990). Many early perimeter control devices were large, cumbersome and conspicuous. Because these devices contradicted LRA principles, they were unpopular with relatives and staff. Further, they were likely to attract and be tampered with or removed by wanderers. In addition, concerns for the rights of patients have been expressed by ethicists and patient advocates who view electronic monitoring and alarms as violating the rights of privacy, dignity, confidentiality, and the right to give informed consent (Warrington & Eagles, 1995). Further, many practical problems have been associated with perimeter control devices over time, including inadvertent and purposeful triggering, multiple exiting, keeping devices comfortable and functional on patients, and the generally high costs of installation and maintenance.

Several studies have focused specifically on the effect of an alarm system designed to deter the wanderer and alert staff to a possible elopement, however, these studies did not employ definitions or criteria for the selection of wanderers and were uncontrolled (e.g., Gaffney, 1986). One study by Negley, Molla, and Obenchain (1990) focused on changes on resident behaviors when an alarm system is utilized to prevent elopement. Five ambulatory male subjects (mean age = 68.6 years), with various medical diagnoses and considered wanderers by the subjective assessment of senior medical and nursing staff, comprised the sample. Subjects wore wrist tags that activated an alarm concealed under a mat at the entrance to the ward. Observations totaling 120 hours were conducted at 30 hours per week for 4 weeks, with the test conditions varied each week as follows: (i) door unclosed and no sensor tags worn; (ii) door open, sensor tags worn, alarm off; (iii) door open, tags off, alarm on; and (iv) door open, tags off, alarm off. Results indicated a 67% decrease in attempted door exits between weeks 2 and 4.

7.17.8.5 Wandering Paths or Tracks

Many special care environments designed for dementia sufferers now incorporate specific paths or tracks, both indoor and outdoor, where patients can wander; see Zeisel, Hyde, and Levkoff (1994) for an example. By facilitating wandering in a safe environment, rather than restraining a wanderer, proponents suggest that wanderers would be more content and less aggressive, sleep better, and fall and elope less. While the shape of these walking paths varies from geometric to circular, the design is usually simple and the surrounds peaceful and interesting, with plants, pools, fountains, sculptures, and seats for resting. The pleasant atmosphere and accompanied walking which occurs within such environments are assumed to affect both mood and activity levels positively in wandering patients. Clearly, relatives and staff appreciate being able to visit in pleasant surrounds. However, the approach is atheoretical and disregards potential negative effects of encouraging or heightening wandering. Little research exists that tests the impact of tracks on level of wandering behavior or on wanderers, other residents, visitors, or staff.

An empirical study involving 14 subjects in a specially-designed dementia care unit (Kovach, Weisman, Chaudhury, & Calkins, 1997) explored the association of physical environment with behaviors and care of the demented. The study incorporated a "meaningful wandering path" in the facility design. The path circled a central area allowing continuous movement overlooking activities and involved small travel distances and landmarks, but the specific effect of this design element was not reported. To date, there is no evidence to suggest that such environmental manipulation alters parameters of wandering or benefits the wanderer.

7.17.8.6 Activities Programming

Most intervention studies using programmed activities have been global in intent and eclectic in content. Few have focused on a single systematic intervention targeting a single problem behavior such as wandering. These studies have assumed an unsubstantiated homogeneity among problematic behaviors associated with dementia and thus their clinical applicability and utility is limited (Cleary, Clamon, Price, & Shullaw, 1988).

Only one empirical study to date has explicitly sought to impact only wandering behavior (Goldsmith et al., 1995) with little success. The rationale for using activities to moderate wandering remains atheoretical. Anecdotal evidence suggests that it is difficult to find activities that are age-appropriate, span the abilities of a group of cognitively-impaired people, engage the interest of wanderers, and assure the success of wanderers with the activity.

7.17.8.6.1 Physical and chemical restraint

Physical restraint use by belts, wrist ties, vests, lap trays, and other mechanical devices remains common in wandering and agitation management (Werner, Cohen-Mansfield, Braun, & Marx, 1989), even though no published studies document the supposed reduction of wandering or agitation resulting from the use of physical restraints. Rather, a rebound effect in activity level may occur upon restraint release (Miller, 1975). Prevalence of restraint use is documented at between 18% and 85% in nursing homes (Folmar & Wilson, 1989;

Morrison, Crinklaw-Wianko, King, Thibeault, & Wells, 1987; Zimmer, Watson, & Treat, 1984). While these rates are from studies made before Federal criteria and restrictions for restraint use in long-term care settings were established to improve the quality of life and safety of the residents (OBRA, 1987), no more recent estimates were located. High levels of restraint use have also been found when residents are over age 75 years have a medical diagnosis of organic brain disease (Folmar & Wilson, 1989). Problems associated with restraints include: loss of dignity, urinary incontinence, constipation, muscle atrophy, limb ischemia, increased agitation, strangulation, and, in some cases, death (Beardsley et al., 1989; Werner et al., 1989).

In a post-OBRA implementation study, effects of physical restraints on psychotropic medication use in nursing homes were examined (Werner, Cohen-Mansfield, Karoknay, & Braun, 1994). A retrospective and prospective longitudinal design was used to obtain information about medication use and restraints over a two-year period at seven points in time: four, two, and one month prior to restraint removal; at the time of restraint removal and one, two, and four months after. Subjects (n = 88, 75% female, mean age = 86.3 years, mean length in residence = 5.2 years) were those physically restrained residents of a large nursing home and constituted 62% of the 140 residents who were restrained in the facility. Sixty-one subjects (69.3%) had a medical diagnosis of dementia and 18 had a specific diagnosis of AD.

Results indicated that the proportion of all residents receiving psychotropic drugs decreased from 83% to 70% during the study period, the most significant period of decrease being one month after restraint removal. Those with dementia evidenced the greatest decrease. In this group, medication use decreased from 87% four months prior to restraint removal to 57% four months after restraint removal. These results should be considered cautiously because they represent only one institution. However, adaptations to legal control of restraint use and medication withdrawal was clearly beneficial to those with dementia.

Neuroleptics and benzodiazepines are the most frequently used agents and psychotropic medications are often combined with other centrally acting agents. Zimmer et al. (1984) documented the frequency of interventions for agitation (often including wandering) used in skilled nursing facilities: psychoactive drugs (58% of patients), physical restraints (47%), reality orientation strategies (47%), and specialist psychiatric consultation (38%). In a survey of 1276 geriatric inpatients at 12 Veterans' Administration hospitals, 61% were receiving psychoactive drugs (Prien, Haber, & Caffey, 1985). Fifty-five percent of patients with organic brain syndrome were receiving psychotropics and the most frequently prescribed agents were neuroleptics (44%), antidepressants (11%), and antianxiolytics (10%). Another study of 101 demented inpatients reported that 38% received one psychotropic and 20% received two such medications, with neuroleptics being the agent of choice (Michel & Kolakowska, 1981).

An extensive body of literature, most of it descriptive and anecdotal, addresses the effectiveness of general pharmacological interventions for the control of wandering, agitation, and other behavioral symptoms. Barnes, Veith, Okimoto, Raskind, and Gumbrecht (1982) reported on the effectiveness of antipsychotic drugs for specific behavioral problems in a study of 60 demented subjects. In contrast, Druckenbrod, Rosen, and Cluxton (1993) advised limited use of psychoactive drugs because little data were available documenting the efficacy of their use in the elderly. Nevertheless, neuroleptics remain the mainstay treatment of behavioral disturbance in dementia (including wandering) and the intervention of choice in agitation management, although efficacy studies conclude that the effects are modest at best (Schneider, Pollock, & Lyness, 1990). Haloperidol is the most frequently chosen medication for agitation, restlessness, and wandering in older adults because of its reportedly benign cardiovascular and anticholinergic profile (Weiner, Bruhn, Svetlik, Tintner, & Hom, 1991). On the other hand, many elders are extremely sensitive to its extrapyramidal effects and are already at increased risk of drug-induced TD (Marsden, Milson, Parkes, Pycock, & Tarsy, 1975). For example, Prien et al. (1985) reported that in a small group of AD outpatients with psychotic symptoms, patients could not be maintained on more than 4 mg four times per day of haloperidol because of its extrapyramidal effects.

Neuroleptics have been more widely studied than any other class of pharmacological agent. Most controlled studies of neuroleptic use in dementia have suffered from major methodological flaws, for example, diagnostic heterogeneity was a common feature, methods of assessment were often global, and samples were comprised predominantly of severely demented subjects. In addition, the effects of neuroleptics on core psychotic features as well as behavioral disturbance (including wandering) were not assessed. Only a few published studies used a randomized, double-blind, placebo-controlled design.

Schneider et al. (1990) conducted a meta-analysis of placebo-controlled trials of neuroleptics in wandering, agitated patients with any form of organic brain syndrome. The methodological flaws mentioned previously compromised the utility of the meta-analytic approach to an extent. However, neuroleptics were found to be moderately more efficacious than the placebo (one-tailed $p = 0.004$), with effect size reduced by the high placebo response rate observed in some studies. These response rates may have been due to the inclusion of subjects with very mild behavior change only in order to increase sample size (Devanand & Levy, 1995).

To date no published randomized, double-blind, placebo-controlled study of neuroleptic use (without the concurrent use of other psychoactive agents) to treat wandering in a "pure" AD sample of either institutionalized or community-dwelling subjects have been published. A few reports of neuroleptic–benzodiazepine comparisons in the treatment of behavioral symptoms in demented samples are available.

Several other medications have been reported to be efficacious: trazodone (Pinner & Rich, 1988; Tingle, 1986); lithium (Schneider & Sobin, 1992); carbamazepine (Gleason & Schneider, 1990; Leibovici & Tariot, 1988); valproate (Mellow, Solano-Lozez, & Davis, 1993); propranolol (Pauszeck, 1991); and zolpidem, a nonbenzodiazepine hypnotic (Shelton & Hocking, 1997). Published data from randomized, double-blind, placebo-controlled study of all these medications are sparse. Szwabo, Woodward, Grossberg, and Shenn (1991) postulated that wandering is a response to anxiety and used Alprazolam, a short-acting antianxiolytic, in their 12-week study of 16 wanderers in a nursing home. Wanderers received cognitive testing pre- and postintervention and were placed on four weeks of 0.25 mg twice daily. Alprazolam, titrated to 0.5 mg if needed by a subject. A four-week period without Alprazolam followed the intervention. Behavioral change was monitored using daily ratings of total time spent dealing with wandering behavior and with the Nurse's Observation Scale for Inpatient Evaluation (Honigfeld, Gillis, & Klett, 1966). Results indicated statistically significant improvements in speech and irritability while on the drug, however, drowsiness increased.

Kamei, Oishi, and Takasu (1996) evaluated the use of fasudil hydrochloride, a novel cerebral vasodilator, for night wandering symptoms in two patients with a medical diagnosis of cerebrovascular dementia. Results indicated that wandering behavior disappeared in both patients during eight weeks of treatment, while memory function improved mildly. Wandering began again immediately after the cessation of treatment. Regional cerebral blood flow, measured by P-magnetic resonance spectroscopy and Xe-computed tomography, remained unaltered throughout.

7.17.9 RECOMMENDATIONS FOR FURTHER RESEARCH

7.17.9.1 Definition, Measurement, and Description

Lack of a standard, widely-accepted definition remains a primary challenge in research on wandering. While multiple authors have offered definitions of wandering, the scientific community has made little advance in the discussion and acceptance of a universal definition of the behavior. Reaching such a conceptual consensus would assist researchers in communicating effectively about the phenomenon and move them to issues of measurement and description which are required before interventions can be theoretically supported and rigorously tested. Fundamental questions such as: (i) Do all cognitively impaired individuals wander? (ii) What is the point at which an individual is classified "a wanderer"? (iii) Given that men and women are known to have different spatial abilities, are there gender differences in wandering behavior? and (iv) What are the natural limits to parameters of wandering, such as frequency and time in motion? will remain unanswered until consensus on a definition is reached.

In line with a need for a standard definition is that for valid and reliable measurement techniques for wandering. If there is any "gold standard" it is direct observation, the parameters of which are supported by some theoretical rationale and/or valid typology for wandering behaviors. An additional need, given the known costs and inter-rater reliability concerns with multiple raters in several settings, is for a biomechanical measure that reliably approximates at least some wandering parameters. One aim of our current study is to validate a range of mechanical devices for their utility in estimating frequency and total time in motion. However, there is yet no meter developed for capturing variations among any types or patterns of wandering. While our work has concentrated on the use of patterns (Martino-Salzman et al., 1991) as an organizer, there are other typologies which remain untested, for example, those developed by Hussian and Davis (1985) and Hope et al. (1994). The further testing and refining of mechanical measures, within a variety of typologies, would assist in moving us closer to understanding how best to quantify the phenomenon.

Studies of wandering tend to utilize small samples. While this issue relates to difficulties in the recruitment of aged subjects generally, and the frail elderly with dementia specifically, the result is that the overall pool of subjects about whom we have information remains small. The protocols for subject selection have tended to be loose, meaning that subject comparison across studies is limited and generalization is difficult, if not impossible. Refinement of protocols for subject selection using more sophisticated medical diagnosis parameters and neuropsychological testing, for example, the NINCDS-ADRDA (McKhann et al., 1984) for AD or for MID would result in more homogenous samples about which more robust generalizations could be made than those currently possible. Studies with larger and random samples will also facilitate breakdown of data into diagnostic groups.

To date we have few data describing the stability and normal variation for any parameter of wandering over time. Thus, studies incorporating baseline and repeated measures would assist in future methodological decision-making. In studies designed to evaluate episode frequency or time in motion, it is critical that both the period of intervention is in effect, as well as the distribution of wandering over the rest of a 24 hour period is determined. In other words, did the intervention change the volume or the distribution of wandering? Without such clarity, the overall utility of interventions cannot be meaningfully interpreted in terms of clinical management.

7.17.9.2 Etiologic Factors

Research to date has proffered multiple explanatory hypotheses and employed a wide variety of research paradigms, none of which adequately explain the phenomenon. Thus, there is a need for well-controlled studies to test competing hypotheses or to evaluate the relative contribution of various putative etiologies for explaining wandering. While a number of potential etiologies have been suggested, measurement poses a problem in evaluating some of these in persons with severe dementia. For example, while proxy measures of circadian rhythm disturbance such as sleep and body temperature are reasonable, it is not directly possible to measure the function of the suprachiasmatic nucleus given current technology. Further exploration of the role of previous personality in disturbing behaviors associated with dementia is indicated, as demonstrated by the results of research utilizing the NEO five-factor personality inventory (Kolanowski, Strand, & Whall, 1997; Thomas, 1997). If, as suggested, the basic features of an individual's personality are preserved predominantly intact in AD, patterns of motor activity integral to the expression of personality may also be preserved.

While a number of proximal factors thought to be associated with wandering have been proposed in the NDB model (Algase et al., 1996), the extent of the impact of individual and amalgamated proximal factors in either moderating or exacerbating wandering is not known. Likewise, the conditions under which wandering can be altered have not been elucidated as yet. A number of studies have recognized wandering as influenced by a variety of environmental factors, yet studies that link type and amount of wandering to particular outcomes are few. Studies that use a theoretical perspective to isolate and manipulate individual or conceptually-grouped environmental factors are needed.

Looking more broadly in relation to other or related behavioral issues in dementia, an overlap and lack of clarity persists in current understandings of restlessness, agitation, wandering, and confusion. Initial concept clarification, followed by further development, testing, and refinement of measures for each concept in a severely impaired population is necessary. Existing measures for agitation, for example, do not adequately capture the frequency and intensity of the behavior in many wanderers. Likewise, measures for evaluating physiological need states such as thirst, hunger, fatigue, and elimination in the cognitively impaired require development.

Few tests for assessing areas of specific cognitive deficits in individuals with severe impairments have been developed and most are quite dependent on language skills. The results of such testing, especially in combination with neuroimaging and behavioral measures, may prove invaluable in further understanding the nature and origin of intellectual deficits and the impact of specific deficits on wandering. In studies that have attempted to examine specific aspects of cognitive function in relation to wandering, such as by Liu et al. (1991), most testing was done in grasping space (space within one's reach). Yet we do not know how tests of visual-spatial ability done in grasping space correspond with actual functioning in large-scale space.

7.17.9.3 Intervention Studies

While the clinical need for such studies is appreciated, the majority of intervention studies designed to impact wandering behavior are both premature and atheoretical. It is unclear in many studies what dimension of wandering the researcher is trying to manipulate. In many studies, wandering is only one of multiple behaviors examined and is not the single dependent variable of interest. Wandering is considered in an homogenous cluster with other behaviors such as agitation and elopement. Focusing research specifically on the phenomenon of wandering isolates the behavior and assists in clarity of definition. Intervention studies would also be more useful if derived from theory or based on putative etiologies for wandering.

Case study designs are an underutilized strategy in wandering research. Case study research has the advantage of small sample size, a particular advantage when working with a fragile and difficult-to-recruit population. Case studies can yield rich and complex descriptive data. Those designs using multiple cases with theoretically-structured embedded experiments afford generalization to theoretical propositions based on replication logic (Yin, 1989). Such replication offers compelling support for the initial set of propositions and the observations for each case are considered independent, that is, the intervention for one case does not affect the intervention for the other cases. More systematic use of case studies to accomplish descriptive and explanatory work and to provide a preliminary test of interventions is recommended.

In addition, certain institutional level studies would also be useful to answer questions such as these: (i) Are there characteristics of social environments that affect wandering? (ii) Does a match between the personality of the wanderer and that of the staff impact the behavior? (iii) Do certain institutional policies or physical design elements actually induce or exacerbate wandering?

The relative risk of undesirable outcomes associated with specific aspects or parameters of wandering behavior requires exploration. Which kinds and amounts of wandering predict falls? Or elopements? Is there a threshold at which wandering escalates weight loss? While not exhaustive, answers to these questions and response to issues raised above will do much to advance our understanding of wandering and the development of rational, humane, and effective care for the afflicted.

7.17.10 SUMMARY

This chapter has aimed to present a comprehensive overview of the state of the science on wandering behavior and its implications for research and clinical practice. Review of definitions and typologies revealed a lack of standardization and few attempts of researchers to build on the work of others. Thus, knowledge is disjointed and the ability to interpret results and generalize findings is hampered. Although some estimates of prevalence and incidence have been reported, their value is questionable for the same reasons.

A profile of the cognitively impaired person who may be prone to wandering was derived from a summary of the characteristics of wanderers across a number of studies. In short, neither age, gender, nor medical basis for dementia appears to differentiate wanderers from other cognitively impaired persons. However, differences in the characteristics of wandering behavior among these subgroups is unknown. In general, wanderers have better mobility and worse cognitive impairment on global measures than nonwandering counterparts. Language and visual-spatial skills are particularly affected. Long-standing active behavior patterns and an outgoing personality may also characterize the wanderer.

In an effort to clarify and advance conceptualizations of wandering, the behavior was differentiated from its outcomes. The following putative outcomes were isolated: falls, elopements, weight loss, sleep disturbance and fatigue, improved physical stamina, and worsened social and cognitive skills.

Several possible etiologies for wandering were proposed and the evidence in support of each were summarized. Disturbances in visual-spatial ability and evidence of parietal lobe damage (particularly, the right parietal lobe, which is associated with spatial learning) were offered to support the hypothesis that wandering is a reflection of spatial disorientation. Hippocampal damage, perseveration, and neurotransmitter imbalances involving acetylcholine and dopamine were suggested as the basis for wandering characterized as increased movement. Deterioration of the suprachiasmatic nucleus and the fractured sleep and activity patterns that accompany it provided the rationale for viewing night-time wandering as a reflection of circadian rhythm disruption. While less evidence exists to support the roles of affect and needs in explaining wandering, they cannot be eliminated as etiologies based on existing knowledge.

Behavior mapping, counting and timing strategies, and activity monitoring devices were reviewed for their value in assessing wandering behavior. Multiple assessment techniques for related factors, such as sensory abilities, personal patterns and preferences, and physical discomfort, and for outcomes were discussed.

A range of tested and untested approaches to intervention were evaluated. Few proven and effective techniques were found. Exercise may mitigate night-time wandering, but further study is needed. The effectiveness of cueing and lighting have not been demonstrated. Social interaction approaches are promising, but also require further study before specific protocols can be advocated. Drug therapy is the mainstay of treatment, but effects are global rather than specifically targeted to wandering and multiple issues can arise with the use of medications.

Techniques for managing the elopement attempts were also reviewed. Behavior modification can be effective if conditioned stimuli are continuously applied. However, replication and further evaluation of useful protocols is needed. Two types of camouflage have been evaluated for reducing exit attempts. Results are conflicting, but approaches based on visual agnosia may prove effective if strategies to overcome learning effects are incorporated. A range of perimeter control devices was also discussed but lack of data prevents development of clear recommendations at this time. Several additional strategies such as wandering paths and tracks, activities programming, and use of restraints (both physical and chemical) were also considered. While they are popular in clinical settings, their efficacy and value are unsubstantiated.

Finally, suggestions for further research were offered in the areas of definition and measurement, explanation, and intervention.

ACKNOWLEDGMENTS

The author is grateful for the assistance of Elizabeth R. A. Beattie, Ph.D., Senior Lecturer at James Cook University, Queensland, Australia, for her assistance in the initial draft of sections on assessment and intervention.

7.17.11 REFERENCES

Albert, S. M. (1992). The nature of wandering in dementia: A Guttman scaling analysis of an empirical classification scheme. *International Journal of Geriatric Psychiatry, 7,* 783–787.

Algase, D. (1988). Cognitive and social discriminants of wandering among cognitively-impaired nursing home residents. Unpublished doctoral dissertation, Case Western Reserve University, Cleveland.

Algase, D. (1992a). Cognitive discriminants of wandering among nursing home residents. *Nursing Research, 41,* 78–81.

Algase, D. (1992b). A century of progress: Today's strategies for responding to wandering behavior. *Journal of Gerontological Nursing, 18,* 28–34.

Algase, D. (1997a). Impact of light levels on wandering. Unpublished data.

Algase, D. (1997b). Daily distributions of wandering. Unpublished data.

Algase, D., Beattie, E., Kupferschmid, B., & Beel-Bates, C. (1997). Estimates of stability of daily wandering behavior among cognitively-impaired long-term care residents. *Nursing Research, 46,* 172–178.

Algase, D., Beattie, E., Kupferschmid, B., & Therrien, B. (1997). Cognitive impairment and agitation as predictors of rhythmic parameters and patterns of wandering: A pilot study, unpublished data.

Algase, D., Beck, C., Kolanowski, A., Whall, A., Berent, S., Richards, K., & Beattie, E. (1996). Need-driven dementia compromised behavior: An alternative view of disruptive behavior. *American Journal of Alzheimer's Disease, 11,* 10–19.

Algase, D., Beel-Bates, C., & Kupferschmid, B. (1995). Wandering rhythms and patterns of cognitively-impaired elderly. Unpublished manuscript.

Algase, D., & Cheng, J. (1992a). *Wandering as a rhythm* [abstract]. In Proceedings of the 16th Annual Midwest Nursing Research Society Conference. Indianapolis, IN: Midwest Nursing Research Society.

Algase, D., & Cheng, J. (1992b). Antecedent and consequent activity as pacers for wandering cycles. *The Gerontologist* (Program Abstracts), *31,* 144.

Algase, D., & Struble, L. (1992). Wandering: What, why, and how? In K. Buckwalter (Ed.), *Geriatric mental health nursing: Current and future challenges* (pp. 61–74). Thorofare, NJ: Slack.

Algase, D., & Tsai, J. (1991). Wandering as rhythm. *The Gerontologist* (Program Abstracts), *30,* 140.

Arno, S., & Frank, D. (1994). A group for "wandering" institutionalized clients with primary degenerative dementia. *Perspectives in Psychiatric Care, 30,* 13–16.

Ballard, C. G., Mohan, R. N. C., Bannister, C., Handy, S., & Patel, A. (1991). Wandering in dementia sufferers. *International Journal of Geriatric Psychiatry, 6,* 611–614.

Barbieri, E. (1983). Patient falls are not patient accidents. *Journal of Gerontological Nursing, 9,* 165–173.

Barnes, R., Veith, R., Okimoto, J., Raskind, M., & Gumbrecht, G. (1982). Efficacy of antipsychotic medications in behaviorally disturbed dementia patients. *American Journal of Psychiatry, 139,* 1170–1174.

Beardsley, R. S., Larson, D. B., Lyons, J. S., Gottlieb, G. L., Rabins, R., & Rovner, B. W. (1989). Health services research in nursing homes: A systematic review of three clinical geriatric journals. *Journal of Gerontology, 44,* M30–M35.

Beattie, E. R. A. (1996). *Wandering and agitation behavior in severely cognitively-impaired nursing home residents: The impact of nurse–resident communication style.* Townsville, Australia: James Cook University.

Berg, R., & Cassels, J. (Eds.) (1990). Falls in older persons. Risk factors and prevention. In *The second fifty years: promoting health and preventing disability* (pp. 263–298). Washington, DC: National Academy Press.

Bliwise, D. L. (1994). What is sundowning? *Journal of the American Geriatrics Society, 42,* 1009–1011.

Bliwise, D. L., Carol, J. S., Lee, K. A., Nekich, J. C., & Dement, W. C. (1993). Sleep and "sundowning" in nursing home patients with dementia. *Psychiatry Research, 48,* 277–292.

Brock, M. A. (1991). Chronobiology and aging. *Journal of the American Geriatrics Society, 39,* 74–91.

Brody, E., Kleban, M. & Moss, M., & Kleban, F. (1984). Predictors of falls among institutionalized females with Alzheimer's disease. *Journal of the American Geriatrics Society, 32,* 877–882.

Brungardt, G. S. (1994). Patient restraints: New guidelines for a less restrictive approach. *Geriatrics, 49,* 43–50.

Buchner, D., & Larson, E. (1987). Falls and fractures in patients with Alzheimers type dementia. *Journal of the American Medical Association, 257,* 1492–1495.

Burns, A., Jacoby, R., & Levy, R. (1990). Psychiatric phenomena in Alzheimer's disease. IV: Disorders of behavior. *British Journal of Psychiatry, 157,* 86–94.

Butler, J. (1990). *The anatomy of a critical wandering episode*. Paper presented at the Arkansas Gerontological Society, Little Rock, AK.

Butler, J. P., & Barnett, C. A. (1991). Window of wandering. *Geriatric Nursing, 12,* 226–227.

Campbell, A., Rienken, J., Allan, B., & Martinez, G. (1981). Falls in old age: A study of frequency and related clinical factors. *Age and Ageing, 10,* 264–270.

Cancro, R. (1968). Elopements from the C. F. Menninger Memorial Hospital. *Bulletin Menninger Clinic, 32,* 228–238.

Cason, C. L., & Grissom, N. (1994). *Distraction as an intervention for short-term pain.* Intramural Grant Final Report. University of Arkansas for Medical Sciences, College of Nursing, Little Rock, AK.

Chafetz, P. K. (1990). Two dimensional grid is ineffective against demented patients' exiting through glass doors. *Psychology and Aging, 5,* 146–147.

Chui, H. C., Mortimer, J. A., Slager, U., Zarow, C., Bondareff, W., & Webster, D. D. (1986). Pathologic correlates of dementia in Parkinson's disease. *Archives of Neurology, 43,* 991–995.

Cleary, T. A., Clamon, C., Price, M., & Shullaw, G. (1988). A reduced stimulation unit: Effects in patients with Alzheimer's disease and related disorders. *The Gerontologist, 28,* 511–514.

Cohen-Mansfield, J., & Marx, M. S. (1990). Do past experiences predict agitation in nursing home residents? *International Journal of Aging and Human Development, 28,* 285–294.

Cohen-Mansfield, J., & Werner, P. (1985). Environmental influences on agitation: An integrative summary of an observational study. *The American Journal of Alzheimer's Care and Related Disorders and Research, 10,* 32–39.

Cohen-Mansfield, J., Werner, P., & Marx, M. S. (1989). An observational study of agitation in agitated nursing home residents. *International Psychogeriatrics, 1,* 153–165.

Cohen-Mansfield, J., Werner, P., Marx, M. S., & Freedman, L. (1991). Two studies of pacing in the nursing home. *Journal of Gerontology: Medical Sciences, 46,* M77–M83.

Cook, P., Exton-Smith, A., Brocklehurst, J., & Lampert-Barber, S. (1982). Fractured femurs, falls and bone disorders. *Journal of the Royal College of Physicians London, 16,* 45–49.

Cooke, D. A., Algase, D. A., & Terpenning, M. (1993). *Signposts and the institutionalized elderly: Effectiveness and principles of design.* Unpublished paper.

Coons, D. H. (1988). Wandering. *The American Journal of Alzheimer's Care and Related Disorders & Research, 3,* 31–36.

Cornbleth, T. (1977). Effects of a protected hospital ward area on wandering and non-wandering geriatric patients. *Journal of Gerontology, 32,* 573–577.

Dastoor, D. P., Norton, S., Boillat, J., Minty, J., Papadoupoulou, F., & Mullzur, H. F. (1975). A psychogeriatric assessment program. 1. Social functioning and ward behavior. *Journal of the American Geriatrics Society, 23,* 465–471.

Dawson, P., & Reid, D. W. (1987). Behavioral dimensions of patients at risk of wandering. *The Gerontologist, 27,* 104–107.

deLeon, M. J., Potegal, M., & Gurland, B. (1984). Wandering and parietal lobe signs in senile dementia of Alzheimer's type. *Neuropsychobiology, 11,* 155–157.

Devanand, D. P., & Levy, S. R. (1995). Neuroleptic treatment of agitation and psychosis in dementia. *Journal of Geriatric Psychiatry and Neurology, 8,* S518–S524.

Dickinson, J. I., McLain-Kark, J., & Marshall-Baker, A. (1995). The effects of visual barriers on exiting behavior in a dementia care unit. *The Gerontologist, 35,* 127–130.

Donat, D. C. (1984). Modifying wandering behavior: A case study. *Clinical Gerontologist, 3,* 41–43.

Druckenbrod, R. W., Rosen, J. & Cluxton, R. J. Jr. (1993). As-needed dosing of antipsychotic drugs: Limitations and guidelines for use in the elderly agitated patient. *Annals of Pharmacotherapy, 27,* 645–648.

Ebersole, P., & Hess, P. (1990). *Toward healthy aging: Human needs and nurses response*. St. Louis, MO: Mosby.

Eisdorfer, C., & Faan, W. (1982). Psychotropic drug side effects in the elderly. In *Treatment of psychopathology in the aging* (p. 211). New York: Springer.

Evans, L. (1987). Sundown syndrome in institutionalized elderly. *Journal of the American Geriatrics Society, 35,* 101–108.

Evans, L., & Strumpf, N. (1989). Tying down the elderly: A review of the literature on physical restraints. *Journal of the American Geriatrics Society, 37,* 65–74.

Fayden, A. I., & Townsend, J. J. (1976). Myoclonus in Alzheimer's disease. A confusing sign. *Archives of Neurology, 33,* 278–280.

Fairburn, C. G., & Hope, R. A. (1988). Changes in behaviour in dementia: A neglected research area, *British Journal of Psychiatry, 152,* 406–407.

Fischer, P., Marterer, A., & Danielczyk, W. (1990). Right–left disorientation in dementia of the Alzheimer's type. *Neurology, 40,* 1619–1620.

Foerster, J. (1981). A study of falls: The elderly nursing home resident. *Journal of New York State Nurses Association, 12,* 9–17.

Folmar, S., & Wilson, H. (1989). Social behavior and physical restraints. *The Gerontologist, 29,* 650–653.

Folstein, J., Folstein, S., & McHugh, P. R. (1975). " Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Foxwell, L. (1993). Risk management decreases incidence of falls, wandering. *Provider, 14,* 51.

Francese, T., Sorrell, J., & Butler, F. R. (1997). The effects of regular exercise on muscle strength and functional abilities of late stage Alzheimer's residents. *American Journal of Alzheimer's Disease, 12,* 122–127.

Fuld, P. A., Katzman, R., Davies, P., & Terry, R. D. (1982). Intrusions as a sign of Alzheimer's dementia: Chemical and pathological verification. *Annals of Neurology, 11,* 155–159.

Gaffney, J. (1986). Toward a less restrictive environment. *Geriatric Nursing, 7,* 94–95.

Gallo, J., Reichel, W., & Andersen, L. M. (1995). *Handbook of geriatric assessment* (2nd ed.). Gaisthersburg, MD: Aspen.

Gerdner, L., & Swanson, E. (1993). Effects of individualized music on confused and agitated elderly patients. *Archives of Psychiatric Nursing, 7,* 284–291.

Gleason, R. P., & Schneider, L. S. (1990). Carbamazepine treatment of agitation in Alzheimer's outpatients refractory to neuroleptics. *Journal of Clinical Psychiatry, 51,* 115–118.

Goldberg, E. (1986). Varieties of perseveration: A comparison of two taxonomies. *Journal of Clinical and Experimental Neuropsychology, 8,* 710–726.

Goldsmith, S. M., Hoeffer, B., & Rader, J. (1995). Problematic wandering behavior in the cognitively-impaired elderly: A single-subjected case study. *Journal of Psychosocial Nursing and Mental Health Services, 33,* 6–12.

Goldsten, L., Canavan, A., & Polkey, C. (1989). Cognitive mapping after unilateral temporal lobectomy. *Neuropsychologia, 27,* 167–177.

Greenblatt, D. J., Divoll, M., Abernethy, D. R., Harmatz,

J. S., & Shader, R. I. (1982). Antipyrine kinetics in the elderly: Prediction of age-related changes in benzodiazepine oxidizing capacity. *Journal of Pharmacol Expimental Therapeutics, 220,* 120–126.

Greene, J., Smith, R., Gardiner, M., & Timbury, G. C. (1982). Measuring behavioural disturbance of elderly demented patients in the community and its effects on relatives: A factor analytic study. *Age & Ageing, 11,* 121–126.

Guilleminault, C., Kushida, C., & Leger, D. (1995). Forensic sleep medicine and nocturnal wandering. *Sleep, 18,* 721–723.

Guilleminault, C., Moscovitch, A., & Leger, D. (1995). Forensic sleep medicine: Nocturnal wandering and violence. *Sleep, 18,* 740–748.

Guttman, L. L. (1974). The basis for scalogram analysis. In G. M. Maranell (Ed.), *Scaling: A sourcebook for behavioral scientists* (pp. 43–45). Chicago: Aldine.

Halgren, E. (1991). Firing of human hippocampal units in relation to voluntary movements. *Hippocampus, 1,* 153–162.

Halgren, E., Squires, N. K., Wilson, C. L., Rohrbaugh, J. W., Babb, T. L., & Crandall, P.H. (1980). Endogenous potentials generated in the human hippocampal formation and amygdala by infrequent events. *Science, 210,* 803–805.

Hall, G. R., & Buckwalter, K. C. (1987). Progressively lowered stress threshold: A conceptual model for care of adults with Alzheimer's Disease. *Archives of Psychiatric Nursing, 1,* 399–406.

Hamilton, C. (1993). *The use of tape patterns as an alternative method for controlling wanderers' exiting behavior in a dementia care unit.* Unpublished master's thesis, Virginia Polytechnic Institute and State University.

Henderson, J. N. (1993–4). Protecting the "frustrated eloper". *Journal of Long Term Care Administration, 21,* 13–15.

Henderson, V., Mack, W., & Williams, B. W. (1989). Spatial disorientation in Alzheimer's Disease. *Archives of Neurology, 46,* 391–394.

Heron, J. (1976). *Six-category intervention analysis.* London: Croom-Helm.

Heslin, K., Towers, J., Leckie, C., Thornton-Lawrence, H., Perkin, K., Jaques, M., Mullin, J., & Wick, L. (1992). Managing falls: Identifying population-specific risk factors and prevention strategies. In S. Funk, E. Tornquist, M. Champagne, L. Copp, & R. Wiese (Eds.), *Key aspects of elder care: managing falls, incontinence and cognitive impairment* (pp. 70–88). New York: Springer.

Hewawasam, L. (1996). Floor patterns limit wandering of people with Alzheimer's. *Nursing Times, 92,* 41–44

Hiatt, L. G. (1980). The happy wanderer. *Nursing Homes, 29,* 27–31.

Hogue, C. (1984). Falls and mobility in later life: An ecological model. *Journal of the American Geriatrics Society, 32,* 858–861.

Hogue, C., Studenski, S., & Duncan, P. (1990). The first step in preventing falls. In S. Funk, E. Tornquist, M. Champagne, L. Copp, & R. Wiese (Eds.), *Key aspects of recovery: Improving nutrition, rest and mobility* (pp. 275–280). New York: Springer.

Holden, J., & Therrien, B. (1993). Cue familiarity reduces spatial disorientation following hippocampal damage. *Nursing Research, 42,* 338–343.

Honigfeld, G., Gillis, R. D., & Klett, C. J. (1966). NOSIE-30: A treatment sensitive ward behavior scale. *Psychological Reports, 19,* 180–182.

Hope, T., & Fairburn, C. G. (1990). The nature of wandering dementia: A community based study. *International Journal of Geriatric Psychiatry, 5,* 239–245.

Hope, T., & Fairburn, C. G. (1992). The Present Behavioral Examination (PBE): The development of an interview to measure current behavioral abnormalities. *Psychological Medicine, 22,* 223–230.

Hope, T., Tilling, K. M., Gedling, K., Keene, J. M., Cooper, S. A., & Fairburn, C. G. (1994). The structure of wandering in dementia. *International Journal of Geriatric Psychiatry, 9,* 149–155.

Hurley, A. C., Volicer, B., Hanrahan, P. A., Houde, S., & Volicer, L. (1992). Assessment of discomfort in advanced Alzheimer's patients. *Research in Nursing and Health, 15,* 369–377.

Hussian, R.A. (1981). Psychotherapeutic intervention: Organic mental disorders. *In geriatric psychology: A behavioral perspective* (chap. 7, pp. 123–151). New York: Van Nostrand Reinhold.

Hussian, R. A. (1982). Stimulus control in the modification of problematic behavior in elderly institutionalized patients. *International Journal of Behavioral Geriatrics, 1,* 33–40.

Hussian, R. A., & Brown, D. C. (1987). Use of two dimensional grid patterns to limit hazardous ambulation in demented patients. *Journal of Gerontology, 42,* 558–560.

Hussian, R. A., & Davis, R. L. (1985). *Responsive care behavioral interventions with elderly persons.* Champaign, IL: Research Press.

Jackson, M. E., Drugovich, M. L., Fretwell, M. D., Spector, W. D., Sternberg, J., & Rosenstein, R. B. (1989). Pervalence and correlates of disruptive behavior in the nursing home. *Journal of Aging and Health, 1,* 349–369.

Janken, J., Reynolds, B., & Swiech, K. (1986). Patient falls in the acute care setting: Identifying risk factors. *Nursing Research, 35,* 215–219.

Jellinger, K. (1987). Neuropathological substrates of Alzheimer's disease and Parkinson's disease. *Journal of Neural Transmittors Supplement, 24,* 109–129.

Johnson, C. L., & Grant, L. A. (1985). *The nursing home in American society.* Baltimore: Johns Hopkins University Press.

Kalchthaler, T., Bascon, R., & Quintos, V. (1978). Falls in the institutionalized elderly. *Journal of the American Geriatrics Society, 26,* 424–428.

Kamei, S., Oishi, M., & Takasu, T. (1996). Evaluation of fasudil hydrochloride treatment for wandering symptoms in cerebrovascular dementia with 31P-magnetic resonance spectroscopy and Xe-computed tomography. *Clinical Neuropharmocology, 19,* 428–438.

Kanten, D., Mulrow, C., Gerety, M., Lichenstein, M., Aguilar, C., & Cornell, J. (1993). Falls: An examination of three reporting methods in nursing homes. *Journal of the American Geriatrics Society, 41,* 662–666.

Kapp, M. (1986). Preventing malpractice suits in long-term care facilities. *Quality Review Bulletin, 12,* 109–113.

Kippenbrock, T., & Soja, M. (1993). Preventing falls in the elderly: Interviewing patients who have fallen. *Geriatric Nursing, 14,* 205–209.

Kolanowski, A. M., Strand, G., & Whall, A. (1997). A pilot study of the relation of premorbid characteristics to behavior in dementia. *Journal of Gerontological Nursing, 23,* 21–30.

Koller, W. C., Wilson, R. S. , Glatt, S. L., & Fox, J. H. (1984). Motor signs are infrequent in dementia of the Alzheimer type. *Annals of Neurology, 16,* 514–516.

Kovach, C., Weisman, G., Chaudhury, H., & Calkins, M. (1997). Impacts of a therapeutic environment for dementia care. *American Journal of Alzheimer's Disease, 12,* 99–110.

Kushida, C. A., Clerk, A. A., Kirsch, C. M., Hotson, J. R., & Guilleminault, C. (1995). Prolonged confusion and nocturnal wandering arising from NREM and REM sleep: A case report. *Sleep, 18,* 757–764.

Lanceley, A. (1985). Use of controlling language in the rehabilitation of the elderly. *Journal of Advanced Nursing, 10,* 125–135.

LaPorte, R. E., Black-Sandler, R., Cauley, J. A., Link, M., Bayles, C., & Marks, B. (1983). The assessment of physical activity in older women: Analysis of the interrelationship of activity monitoring, activity surveys, and caloric intake. *Journal of Gerontology, 38,* 394–397.

LaPorte, R. E., Cauley, B., Kinsey, C., Corbett, W., Robertson, R., Black-Sandler, R., Kuller. L., & Falkel, J. (1982). The epidemiology of physical activity in children, college students, middle-aged men, menopausal females and monkeys. *Journal of Chronic Diseases, 35,* 787–795.

LaPorte, R. E., Kuller, L. H., Kupfer, D. J., McPartland, R. J., Matthews, G., & Caspersen, C. (1979). An objective measure of physical activity for epidemiological research, *American Journal of Epidemiology, 109,* 158–167.

Leach, E. (1985). Safeguarding wandering patients. *Health Law Canada, 6,* 37–44.

Leibovici, A., & Tariot, P. N. (1988). Agitation associated with dementia: A systematic pproach to treatment. *Psychopharmacological Bulletin, 24,* 49–53.

Lewin, K. (1935). *A dynamic theory of personality: Selected papers by Kurt Lewin.* New York: McGraw-Hill.

Liu, L., Gauthier, L., & Gauthier, S. (1991). Spatial disorientation in persons with early senile dementia of the Alzheimer type. *American Journal of Occupational Therapy, 45,* 67–74.

Louis, M (1993). Falls and causes. *Journal of Gerontological Nursing, 9,* 143–156.

Lovell, B. B., Ancoli-Israel, S., & Gevirtz, R. (1995). Effect of bright light treatment on agitated behavior in institutionalized elderly subjects. *Psychiatry Research, 57,* 7–12.

Madson, J. (1990). *The development of a device to measure wandering in persons with senile dementia.* Unpublished master's thesis. University of Dubuque, Iowa.

Mardsen, C. D., Milson, J., Parkes, J. D., Pycock, C., & Tarsy, D. (1975). Proceedings: The effect of cholinergic and anticholinergic drugs on rotational behaviour in mice with destruction of one nigrostriatal pathway. *Journal of Physiology (London), 249,* 64P–65P.

Marley, M. L. (1982). *Organic brain pathology and Bender-Gestalt Test: A differential diagnostic scoring system.* New York: Grune & Stratton.

Martino-Saltzman, D., Blasch, B. B., Morris, R. D., & McNeal, L. W. (1991). Travel behavior of nursing home residents perceived as wanderers and nonwanderers. *The Gerontologist, 31,* 666–672.

May, C. (1993). Subjectivity and culpability in the constitution of nurse–patient relationships. *International Journal of Nursing Studies, 30,* 181–192.

Mayer, R., & Darby, S. (1991). Does a mirror deter wandering in demented older people? *International Journal of Geriatric Psychiatry, 6,* 607–609.

McGrowder-Lin, R., & Bhatt, A. (1988). A wanderer's lounge program for nursing home residents with Alzheimer's disease. *The Gerontologist, 28,* 607–609.

McKhann, G., Drachman, D., Folstein, M., Katz, R., Price, D., & Stadlan, E. M. (1984). Report of the NINCDS-ADRDA work group under the auspices of department of human and health services task force of Alzheimer's disease. *Neurology, 34,* 939–944.

Meguro, K., Yamaguchi, S., Yamazaki, H., Itoh, M., Yamaguchi, T., Matsui, H., & Sasaki, H. (1996). Cortical glucose metabolism in psychiatric wandering patients with vascular dementia. *Psychiatric Research Neuroimaging, 67,* 71–80.

Mellow, A. M., Solano-Lopez, C., & Davis, S. (1993). Sodium valproate in the treatment of behavioral disturbance in dementia. *Journal of Geriatric Psychiatry and Neurology, 6,* 205–209.

Mendez, M. F., Mendez, M. A., Martin, R., Smyth, K. A., & Whitehouse, P. J. (1990). Complex visual disturbances in Alzheimer's disease. *Neurology, 40,* 439–443.

Michel, K., & Kolakowska, T. (1981). A survey of prescribing psychotropic drugs in two psychiatric hospitals. *British Journal of Psychiatry, 138,* 217–221.

Milke, D. (1989). Comparisons of caregivers evaluations to systematic observations of wanderers. *The Gerontologist, 29,* 143A. Program for the 42nd Annual Scientific Meeting, Minneapolis, MN.

Miller, M. (1975). Iatrogenic and nurisgenic effects of prolonged immobilization of the ill aged. *Journal of the American Geriatrics Society, 23,* 360–369.

Miller, J. (1976). Control of the ambulatory senile. *Concern in Care of the Aging, 2,* 34–37.

Mintzer, J. E., Lewis, L., Pennypaker, L., Simpson, W., Bachman, D., Wohlreich, G., Meeks, A., Hunt, S., & Samson, W. (1993). Behavioral Intensive Care Unit (BICU): A new concept in the management of acute agitated behavior in elderly demented patients. *The Gerontologist, 33,* 801–806.

Mohr, E., Litvan, I., Williams, J., Fedio, P., & Chase, T. (1990). Selective deficits in Alzheimer and Parkinsonian dementia: Visuospatial function. *Canadian Journal of Neurological Sciences, 17,* 292–297.

Molsa, P. K., Martilla, R., J., & Rinne, U. K. (1984). Extrapyramidal signs in Alzheimer's disease. *Neurology, 34,* 1114–1116.

Monsour, N., & Robb, S. (1982). Wandering behavior in old age: A psychosocial study. *Social Work, 27,* 411–416.

Morgenson, G., & Nielsen, M. (1984). A study of the contribution of hippocampal- accumbens-subpallidal projections to locomotor activity. *Behavioral and Neural Biology, 42,* 38–51.

Morino, M. (1989). Preventing patient elopement at an outpatient Alzheimer's facility. *Contemporary Longterm Care, 12,* 84–86.

Morris, E., & Isaacs, B. (1980). The prevention of falls in a geriatric hospital. *Age and Ageing, 9,* 181–185.

Morrison, J., Crinklaw-Wiancko, D., King, D., Thibeault, S., & Wells, D. L. (1987). Formulating a restraint use policy. *Journal of Nursing Administration, 17,* 39–42.

Morse, J., Tylko, S., & Dixon, H. (1985). The patient who falls—and falls again: Defining the aged at risk. *Journal of Gerontological Nursing, 11,* 15–18.

Morse, J., Tylko, S., & Dixon, H. (1987). Characteristics of the fall-prone patient. *The Gerontologist, 27,* 516–522.

Muller, R. U., & Kubie, J. L. (1989). The firing of hippocampal place cells predicts the future position of freely moving rats. *Journal of Neuroscience, 9,* 4101–4110.

Mulrooney, C. (1993–4). Controlling exposure and exposing control. *Journal of Long Term Care Administration, 21,* 14–16.

Murray, H. A. (1938). *Explorations in personality.* New York: Oxford University Press.

Murray, J., Marquardt, T. P., Richardson, A., & Nalty, D. (1984). Differential diagnosis of aphasia and dementia from aphasia test battery scores. *Journal of Neurological Communication Disorders, 1,* 33–39.

Nadel, L. (1991). The hippocampus and space revisited. *Hippocampus, 1,* 221–229.

Namazi, K. H., Rosner, T. T., & Calkins, M. P. (1989). Visual barriers to prevent ambulatory Alzheimer's patients from exiting through an emergency doors. *The Gerontologist, 29,* 699–702.

National Center for Health Statistics (1979). *The national nursing home survey: 1977 summary for the United States: Vital and health statistics (1979).* Hyatsville, MD: Author.

Negley, E., Molla, P. M., & Obenchain, J. (1990). No exit: The effects of an electronic security system on confused patients. *Journal of Gerontological Nursing, 16,* 21–25.

Nevitt, M., Cummins, S., Kidd, S., & Black, D. (1989). Risk factors for recurrent nonsyncopal falls: A prospective study. *Journal of the American Medical Association, 261,* 2663–2668.

OBRA (Omnibus Budget Reconciliation Act) (1987). US Congress. Public Law 101–508.

Okawa, M., Mishima, K., Shimizu, T., Iijuma, S., Hishikawa, Y., Hozumi, S., & Hori, H. (1989). Sleep–waking rhythm disorders and their phototherapy in elderly patients with dementia. *Japanese Journal of Psychiatry and Neurology, 43,* 293–295.

Overstall, P. (1978). Falls in the elderly—epidemiology, aetiology and management. In B. Isaacs (Ed.), *Recent advances in geriatric medicine*. London: Churchill-Livingstone.

Pankratz, L., & Jackson, J. (1994). Habitually wandering patients. *New England Journal of Medicine, 331,* 1752–1755.

Pauszek, M. E. (1991). Propranolol for treatment of agitation in senile dementia. *Indiana Medicine, 84,* 16–17 .

Pinner, E., & Rich, C. L. (1988). Effects of trazodone on aggressive behavior in seven patients with organic mental disorders. *American Journal of Psychiatry, 145,* 1295–1296.

Plazzi, G., Tinuper, P., Montagna, P., Provini, F., & Lugaresi, E. (1995). Epileptic nocturnal wanderings. *Sleep, 18,* 749–756.

Preston, G. C., Broks, P., Traub, M., Ward, C., Poppleton, P., & Stahl, S. M. (1988). Effects of lorazepam on memory, attention, and sedation in man. *Psychoparmalogical Bulletin, 95,* 208–215.

Prien, R. F., Haber, E., & Caffey, D. (1985). Research overview of drug therapy in recurrent and chronic depression. *Advanced Biochemical Psychopharmacology, 40,* 121–130.

Prinz, P. N., Peskind, E. R., Vitaliano, P. P., Raskind, M. A., Eisdorfer, C., Zemcuznikov, N., & Gerber, C. J. (1982). Changes in the sleep and waking EEGS of nondemented and demented elderly subjects. *Journal of American Geriatric Society, 30,* 86–94.

Prinz, P. N., Vitaliano, P. P., Vitiello, M. V., Bokan, J., Raskind, M., Peskind, E., & Gerber, C. (1982). Sleep, EEG and mental function changes in senile dementia of the Alzheimer's type. *Neurobiology of Aging, 3,* 361–370.

Rabins, P. V., Mace, N. L., & Lucas, M. J. (1982). The impact of dementia on the family. *Journal of the American Medical Association, 248,* 333–335.

Rader, J. (1987). A comprehensive approach to problem wandering. *The Gerontologist 27,* 756–760.

Rader, J., Doan, J., & Schwab, M. (1985). How to decrease wandering, a form of agenda behavior. *Geriatric Nursing, 6,* 196–199.

Ranck, J. B. (1973). Studies on single neurons in dorsal hippocampal formation and septum in unrestrained rats, part I: Behavioral correlates and firing repertoires. *Experimental Neurology, 41,* 462–531.

Reinboth, J., & Gleidenvand, T. (1982). Pilot study of falls in an extended care setting. University of Iowa.

Reisberg, B., Ferris, S. H., deLeon, M. J., & Crook, T. (1982). The Global Deterioration Scale for assessment of primary degenerative dementia. *American Journal of Psychiatry, 139,* 1136–1139.

Rindlisbacher, P., & Hopkins, R. W. (1991). Sundowning syndrome: A conceptual analysis and review. *American Journal of Alzheimer Care and Related Disorders and Research, 6,* 2–9.

Rindlisbacher, P., & Hopkins, R. W. (1992). An investigation of the Sundowning syndrome. *International Journal of Geriatric Psychiatry, 7,* 15–23.

Rindlisbacher, P., Hopkins, R. W., & Grant, N. T. (1992). An investigation of the sundowning syndrome and ambient light. *American Journal of Alzheimer* Care and Research, 7, 22–27.

Robb, S. S. (1987). Exercise treatment for wandering. In H. J. Altman (Ed.), *Alzheimer's disease: Problems, prospects, and perspectives* (pp. 213–218). New York: Plenum.

Rubenstein, L., Josephson, K., & Robbins, A. (1994). Falls in the nursing home. *Annals of Internal Medicine, 121,* 442–451.

Ryan, J. P., McGowan, J., McCaffrey, N., Ryan, G. T., Zandi, T., & Brannigan, G. C. (1995). Graphomotor perseveration and wandering in Alzheimer's disease. *Journal of Geriatric Psychiatry and Neurology, 8,* 209–212.

Salzman, C. (1987). Treatment of the elderly agitated patient. *Journal of Clinical Psychiatry, 48,* S19–S22.

Sandson, J., & Albert, M. (1984). Varieties of perseveration. *Neuropsychologia, 22,* 715–732.

Sandson, J., & Albert, M. (1987). Perseveration in behavior neurology. *Neurology, 37,* 1736–1741.

Satlin, A., Teicher, M. H., Lieberman, H. R., Baldessarini, R. J., Volicer, L., & Rheaume, Y. (1991). Circadian locomotor activity rhythms in Alzheimer's Disease. *Neuropsychopharmocology, 5,* 115–126.

Schneider, L. S., Pollock, V. E., & Lyness, S. A. (1990). A meta-analysis of controlled trails of neuroleptic treatment in dementia. *Journal of American Geriatrics Society, 38,* 553–563.

Schneider, L. S., & Sobin, P. B. (1992). Non-neuroleptic treatment of behavioral symptoms and agitation in Alzheimer's disease and other dementia. *Psychopharmocology Bulletin, 28,* 71–79.

Sehested, P. & Severin-Nielsen, T. (1977). Falls by hospitalized elderly patients: causes, prevention. *Geriatrics, 32,* 101–108.

Shelton, P., & Hocking, L. B. (1997). Zolpidem for dementia-related insomnia and nighttime wandering. *Annals of Pharmacotherapy, 31,* 319–322.

Singh, S., Mulley, G., & Lusowsky, M. (1988). Why are Alzheimer's patients thin. *Age and Aging, 17,* 21–28.

Slauson, T. J., Bayles, K. A., & Tomeoda, C. K. (1987). *Communication disorders and late stage Alzheimer's disease*. Unpublished manuscript.

Snyder, L. H., Rupprecht, P., Pyrek, J., Brekhus, S., & Moss, T. (1978). Wandering. *The Gerontologist, 18,* 272–280.

Soverini, S., & Borghesi, E. (1968). On a strange case of wandering in an arteriosclerotic demented patient. *Gerontology, 16,* 846–851.

Stokes, G. (1986). *Wandering, common problems with the elderly confused*. Bicester, UK: Winslow Press.

Sunderland, T., & Silver, M. A. (1988). Neuroleptics in the treatment of dementia. *International Journal of Geriatric Psychiatry, 3,* 79–88.

Swaab, D. F., Fliers, E., & Partiman, T. S. (1985). The suprachiasmatic nucleus of the human brain in relation to sex, age and senile dementia. *Brain Research, 342,* 37–44.

Szwabo, P., Woodward, V., Grossberg, G., & Shenn, W. W. (1991). The use of Alprazolam for decreasing problem wandering in geriatric patients. *American Journal of Alzheimer's Care and Related Disorders and Research, 6,* 31–36.

Tappen, R. (1988). Revitalization: a developing approach to nursing care of the cognitively impaired. *Florida Nurse, 36,* 3.

Teri, L., Larson, E. B., & Reifler, B. V. (1988). Behavioral disturbance in dementia of the Alzheimer's type. *Journal of the American Geriatrics Society, 36,* 1–6.

Thomas, D. W. (1995). Wandering: A proposed definition. *Journal of Gerontological Nursing, 21,* 35–41.

Thomas, D. W. (1997). Understanding the wandering patient. *Journal of Gerontological Nursing, 23,* 16–24.

Thornbury, J. (1992). Cognitive performance on Piagetian tasks by Alzheimer's disease patients. *Research in Nursing and Health, 15,* 11–18.

Tingle, D. (1986). Trazodone in dementia [letter]. *Journal of Clinical Psychiatry, 47,* 482.

Tinetti, M. (1986). Performance-oriented assessment of mobility problems in elderly patients. *Journal of the American Geriatric Society, 34,* 119–126.

Tinetti, M. (1987). Factors associated with serious injury during falls by ambulatory nursing home residents. *Journal of the American Geriatrics Society, 35,* 644–648.

Tinetti, M., Speechley, M., & Ginter, S. (1988). Risk factors for falls among elderly persons living in the community. *New England Journal of Medicine, 319,* 1701–1707.

Traub, R. D., Miles, R., & Wong, R. (1989). Model of the origin of rhythmic population oscillations in the hippocampal slice. *Science, 243,* 1319–1326.

Vallarino, R., & Sherman, F. (1981). Stroke, fractured hip, amputation, pressure sores and incontinence: Principles of rehabilitation treatment. In L. Libow & F. Sherman (Eds.), *The core of geriatric medicine*. St. Louis, MO: Mosby.

Vanderwolf, C. H. (1969). Hippocampal electric activity and voluntary movement in the rat. *Electroencephalography and Clinical Neurophysiology, 26,* 407–418.

Vanderwolf, C. H., Kramis, R., Gillespie, L. A., & Bland, B. H., (1975). Hippocampal rhythmical slow activity and neocortical low voltage fast activity: Relation to behavior. In R. L. Isaacson & K. H. Pribram (Eds.), *The hippocampus: Neurophysiology* (pp. 101–128). New York: Plenum.

Vanderwolf, C. N., & Leung, L. (1983) Hippocampal rhythmic slow activity: A brief history and the effects of entorhinal lesions and phenycyclindine in neurobiology of the hippocampus. In W. Siefert (Ed.), *Neurobiology of the hippocampus* (pp. 275–302). London: Academic Press.

Vitiello, M. V., Poceta, J. S., & Prinz, P. M. (1991). Sleep in Alzheimer's disease and other dementing disorders. *Canadian Journal of Psychology, 45,* 221–239.

Walshe, A., & Rosen, H. (1979). A study of patient falls from bed. *Journal of Nursing Administration, 79,* 31–35.

Warrington, J., & Eagles, J. M. (1995). Day care for the elderly mentally ill: Diurnal confusion? *Health Bulletin, 53,* 99–104.

Weiner, M. F., Bruhn, M., Svetlik, D., Tintner, R., & Hom, J. (1991). Experiences with depression in a dementia clinic. *Journal of Clinical Psychiatry, 52,* 234–238.

Weller, M. P. I. (1987). A biochemical hypothesis of wandering. *Medicine, Science, and the Law, 27,* 40–41.

Weller, M. P. I., & Jauhar, P. (1987). Wandering at Heathrow airport by the mentally unwell. *Medicine, Science, and the Law, 27,* 37–39.

Werner, P., Cohen-Mansfield, J., Braun, J., & Marx, M. (1989). Physical restraints and agitation in nursing home residents. *Journal of the American Geriatrics Society, 37,* 1122–1126.

Werner, P., Cohen-Mansfield, J., Koroknay, V., & Braun, J. (1994). The impact of a restraint-reduction program on nursing home residents. *Geriatric Nursing, 15,* 142–146.

Whall, A., Gillis, G., Yankou, D., Booth, D., & Beel-Bates, C. (1989). *Journal of Gerontological Nursing, 18,* 13–17.

Whedon, M. B., & Shedd, P. (1989). Prediction and prevention of patient falls. *Image Journal of Nursing Scholarship, 21,* 108–114.

Whishaw, I.Q. (1991). Locale and taxon systems: no place for neophrenology. *Hippocampus, 1,* 272–274.

Whishaw, I. Q., Cassel, J. C., Majchrzak, M., Cassel, S., & Will, B. (1994). "Short stops" in rats with fimbria-fornix lesions: Evidence for change in the mobility gradient. *Hippocampus, 4,* 577–582.

Whishaw, I. Q., & Nikkel, R. W. (1975). Anterior hypothalamic electrical stimulation and hippocampal EEG in the rat: suppressed EEG, locomotion, self-stimulation, and inhibition of shock avoidance. *Behavioral Biology, 13,* 1–20.

Wild, D., Nayak, U., & Isaacs, B. (1981). How dangerous are falls in old people at home? *British Medical Journal, 282,* 226.

Wilkinson, I. M., & Graham-White, J. (1980). Psychogeriatric dependency rating scales (PGDRS): A method of assessment for use by nurses. *British Journal of Psychiatry, 137,* 558–565.

Witting, W., Kwa, I. H., Eikelenboom, P., Mirmiran, M., & Swaab, D. F. (1990). Alterations in the circadian rest activity rhythm in aging and Alzheimer's disease. *Biological Psychiatry, 27,* 563–572.

Wong, D., & Baker, C. (1988). Pain in children: Comparison of assessment scales. *Pediatric Nursing, 14,* 9–17.

Yin, R. K. (1989). *Case study research: Design and methods*. Newbury Park, CA: Sage.

Zarit, S., Orr, N., & Zarit, J. (1985). *The hidden victims of Alzheimer's Disease: Families under stress*. New York: New York University Press.

Zarit, S. H., Perlin, L. I., & Schaie, K. W. (Eds.) (1993). *Caregiving systems: Informal and formal helpers*. Hillsdale, NJ: Erlbaum.

Zeisel, J., Hyde, J., & Levkoff, S. (1994). Best practices : An environment behavior model (E-B) for Alzheimer special care units. *The American Journal of Alzheimer's Care and Related Disorders and Research, 9,* 4–21.

Zimmer, J. G., Watson, N., & Treat, A. (1984). Behavioral problems among patients in skilled nursing facilities. *American Journal of Public Health, 74,* 1118–1121.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.18
Agitated and Aggressive Behavior

JANE E. FISHER, DIANE N. SWINGEN, and COLLEEN M. HARSIN
University of Nevada at Reno, NV, USA

7.18.1 INTRODUCTION 414
  7.18.1.1 Agitated and Aggressive Behavior 414
7.18.2 ISSUES OF DEFINITION 415
7.18.3 PREVALENCE 416
7.18.4 ETIOLOGICAL FACTORS 416
  7.18.4.1 Physiological Factors 417
    7.18.4.1.1 Pain 417
    7.18.4.1.2 Medication 417
    7.18.4.1.3 Sensory and perceptual deficits 417
  7.18.4.2 Psychological Factors 418
    7.18.4.2.1 Cognitive factors 418
    7.18.4.2.2 Environmental factors 418
7.18.5 ASSESSMENT 420
  7.18.5.1 Physiological Assessment of Agitation and Aggression 420
    7.18.5.1.1 Pain 421
    7.18.5.1.2 Sensory impairment 422
  7.18.5.2 Topographical Assessment of Agitation and Aggression 422
    7.18.5.2.1 Cohen-Mansfield Agitation Inventory 422
    7.18.5.2.2 Pittsburgh Agitation Scale 422
    7.18.5.2.3 Overt Aggression Scale 422
    7.18.5.2.4 Blessed Dementia Scale 423
  7.18.5.3 Assessment of Functional/Contextual Features 423
    7.18.5.3.1 Setting events 423
    7.18.5.3.2 Antecedent environmental stimuli 423
    7.18.5.3.3 Response characteristics 424
    7.18.5.3.4 Environmental consequences 424
    7.18.5.3.5 Motivation Assessment Scale 424
    7.18.5.3.6 Behavior Environment Taxonomy of Agitation 424
    7.18.5.3.7 Functional Analysis of Behavior Recording and Tracking Forms 425
7.18.6 INTERVENTIONS 425
  7.18.6.1 Pharmacological Interventions 425
    7.18.6.1.1 Neuroleptics 425
    7.18.6.1.2 Benzodiazepines 426
    7.18.6.1.3 Beta blockers 426
    7.18.6.1.4 Other antianxiety agents 426
    7.18.6.1.5 Antidepressants 426
    7.18.6.1.6 Anticonvulsants 426
  7.18.6.2 Psychological Interventions 427
    7.18.6.2.1 Environmental stimuli 427
7.18.7 REFERENCES 428

7.18.1 INTRODUCTION

7.18.1.1 Agitated and Aggressive Behavior

Among the behavioral disturbances experienced by older adults with cognitive impairment, agitation and aggression present the most serious consequences for both patients and their caregivers. These behaviors increase the likelihood of an individual being institutionalized (Moak & Fisher, 1990) and experiencing physical or chemical restraint (Sloane, Mathew, & Scarborough, 1991). Numerous reports from studies of family and professional caregivers indicate that aggression in cognitively impaired individuals is the most stressful problem that caregivers confront (Fisher, Fink, & Loomis, 1993).

Within institutional settings physical and chemical restraints have historically been among the most common forms of managing aggression and agitated behaviors (Evans & Strumpf, 1989; Swearer, Drachman, O'Donnell, & Mitchell, 1988). Reliance on these intervention strategies with Alzheimer's disease (AD) patients is wrought with problems due to side effects such as increased risk of falls (Werner, Cohen-Mansfield, Braun, & Marx, 1989), concomitant reduction of desirable behaviors, and the risk of impairment in cognitive functions. Several of the medications commonly prescribed for the treatment of agitation and aggression are psychotropic drugs developed for problems quite different from those experienced by the cognitively impaired elderly. While medications such as haloperidol may be effective in reducing agitated and aggressive behavior they can also have undesirable effects on other behaviors. In addition, as an older adult's physiology changes due to normal (e.g., metabolism rate) or pathological processes (e.g., Alzheimer's disease, hypertension) the effects of medication may be altered. The etiology of the behavioral disturbances themselves may also change, and therefore they may not be responsive to formerly effective interventions.

While the use of physical and chemical restraints may decrease or temporarily eliminate agitation, it is becoming increasingly apparent that restraint-based interventions are not likely to lead to long-term alleviation of behavioral disturbances, and the risks involved in their use often outweigh the benefits.

Interest in the development of restraint-free alternatives for managing behavior problems has been prompted by US Federal legislation affecting nursing home care. In 1987 the US Congress passed the Omnibus Reconciliation Act of Nursing Home Reform (OBRA) requiring that facilities receiving Medicare or Medicaid funds adhere to a treatment program requiring psychotropic drug holidays and the implementation of behavioral programming in an effort to reduce reliance on psychotropic medication within nursing homes. While the essence of OBRA represents a commendable improvement in the care of the elderly within institutions, the requirements for alternative interventions were implemented prior to the development and empirical validation of such treatment alternatives.

Despite the obvious need for alternatives to physical and chemical restraints, and the increase in research addressing behavioral disturbances, there are still remarkably few established treatment alternatives. Three factors have hindered the development of alternative interventions. First, the study of behavioral disturbances in dementia patients has historically occurred within a medical model. The vast majority of studies on behavior problems in cognitively impaired older adults are pharmacological studies. Research on the environmental context of behavioral functioning in dementia has only emerged since the 1980s. Second, there is a lack of consensus within the field regarding basic issues of definition. This lack of consensus has impeded advances in explanation, prevention, and treatment of these problems, and limits comparison of findings across studies. Third, the large majority of nonmedical studies of aggression and agitation are descriptive accounts that have been conducted in an atheoretical context. Most research reports do not reference or employ any theoretical framework within which to consider findings. The atheoretical nature of the research has limited advances in the explanation, prediction, and treatment of behavioral disturbances in the elderly.

The development of restraint-free methods for managing behavioral disturbances in cognitively impaired older adults entails questions regarding the extent to which behavior continues to be influenced by environmental stimuli. Given the limited predictive validity of available classification systems of dementia, it is believed that answers to these questions will be found most effectively through idiographically based assessment and treatment planning methods.

This chapter describes a theoretical approach to the assessment and treatment of behavioral disturbances that considers the context in which behavioral disturbances occur, including physiological and environmental setting events, antecedent environmental stimuli, and the consequences of the behavior. Information on the epidemiology, etiology, assessment, and treatment of agitation and aggression in the elderly is reviewed. A discussion of definitional issues is presented first.

7.18.2 ISSUES OF DEFINITION

A requisite condition of efforts to explain, predict, and modify behavior is the reliable measurement of key elements of the behavior. The operationalization of a target behavior is a critical first step in progress toward reliable, consistent measurement (Johnston & Pennypacker, 1980). A difficulty in formulating an adequate definition of agitation or aggression involves addressing the multidimensional nature of these behaviors and the dynamic nature of the interaction between physiological functioning and response to environmental stimuli inherent in a degenerative disorder such as Alzheimer's disease. Selection of the relevant dimension depends upon the ultimate goal of a measurement or assessment process. Definitions of behavioral disturbances may be based on etiology, topographical features, natural history, environmental dependence, impact on others, or treatment response. Intra- and interindividual variation in these dimensions has been a consistent finding in the literature despite efforts to identify common elements associated with behavioral disturbance in older adults grouped on the basis of factors such as diagnosis, level of cognitive impairment or functional skill impairment, or duration of illness. Within individuals with degenerative forms of dementia, variation in the etiology, environmental dependence, and response to treatment of topographically similar behaviors may change over time, further complicating the reliable identification of agitated of aggressive behaviors.

Any definition of agitation or aggression must be evaluated on the basis of reliable measurement of the behavior as operationalized and the accumulation of evidence for criterion-related concurrent and predictive validity and ultimately construct validity. Despite difficulties inherent in identifying the dimensions relevant for reliable and valid definitions of behavior in dementia, it is evident that attention to this issue is mandated to facilitate advances in the field.

In defining agitation and aggression, researchers have proposed both molar definitions that represent broad inclusive categories and highly specific descriptions of behaviors that are accessible to direct observation. Molar definitions typically require higher inference judgements (Foster & Cone, 1986). The utility of a particular definition will be influenced by the population to which it is applied. This is due to the significant heterogeneity of the population characterized by the label "elderly."

Perhaps the most widely cited molar definition of agitation was proposed by Cohen-Mansfield and Billig (1986). Agitation was defined as "inappropriate verbal, vocal, or motor activity that is not explained by needs or confusion per se" (p. 712). The authors' statement that agitated behaviors are always inappropriate with regard to accepted social norms or the needs of the caregiver reflects the fact that this definition was developed for the general population of elderly. There are at least two problems with this definition of agitation when applied to dementia patients. First, applying the concept of "inappropriate" behavior to dementia patients may itself be inappropriate as it implies that the behavior is unsuitable or incompatible in a given context. When the behavior of dementia patients is considered within the context of a patient's declining competencies and, hence, an increasingly challenging environment, it can be argued that their behavior actually is adaptive rather than inappropriate. Second, when the concept of social norms is applied to dementia patients, agitation emerges as the norm, thus making the normative reference within this definition unsuitable for the population of cognitively impaired older adults.

Other authors have proposed molar definitions that incorporate reference to the motivation of agitation. Barnes and Raskind (1980) defined agitation as a behavioral term addressing excessive motor activity that is often nonpurposeful in nature and is commonly associated with feelings of internal tension, irritability, hostility, and belligerency on the part of the individual engaging in the behaviors. Other definitions of agitation incorporate reference to the *consequences* of behavior. Rosen, et al. (1994), for example, defined agitation as consisting of vocal or motor behavior that is disruptive, unsafe, or interferes with the delivery of care in a particular environment. When treatment outcome is a primary concern, measures which incorporate the consequences of a particular behavior for patients and caregivers may be particularly useful.

A limitation of molar definitions is that the higher-inference judgments required can result in reduced reliability. Two approaches have been employed in attempting to increase the specificity of the definitions of "agitation" and "aggression." First, some researchers have focused on the frequency, duration, and intensity of discrete behaviors that are accessible to direct observation (e.g., Yudofsky, Silver, Jackson, & Endicott, 1986). Other researchers have provided lists specifying which behaviors are encompassed by the larger class of "agitation," and have grouped behaviors into various categories based on characteristics such as response modality (e.g., motoric vs. vocal) or dangerousness. For example, Rosen et al. (1994) described four behavioral classifications of agitation: (i) aberrant vocalizations (repetitious requests or complaints, and nonverbal vocalizations); (ii) motor agitation (pacing and wandering); (iii) aggressiveness (verbal or physical threats, harm to self or others, and physical behavior toward property); and (iv) resisting care (not taking medications or not cooperating with activities of daily living). Another example of a categorical approach to agitated behaviors is found in a study by Cohen-Mansfield (1986). Cohen-Mansfield included pacing, aimless wandering, inappropriate dressing or undressing, spitting, hitting, kicking, throwing things, screaming or making strange noises, biting, complaining, harming self or others, cursing, verbal aggression, constant unwarranted requests or statements, and general restlessness within the class of "agitated behaviors." A factor analysis produced three factors that have since been repeatedly cited as general forms of agitation and have been used to classify and study various aspects of agitation. The first factor is described as physically aggressive behavior, including spitting, biting, hitting, and throwing objects. The second factor includes verbal aggressive behavior such as cursing, threats, and verbal aggression. The third factor is described as nonaggressive behavior, including pacing, wandering, inappropriately dressing or undressing, and constant requests for attention.

Factor analytic methods of studying agitation and aggression may prove useful for understanding the etiology, natural history, or response to treatment of agitated or aggressive behavior. It is possible that behaviors associated with the various factors may emerge as having a similar etiology or response to treatment. Identifying behaviors that are functionally related will ultimately be the most effective for treatment development.

The difficulties inherent in defining agitated behaviors have had important implications for several areas of research, including epidemiological research.

7.18.3 PREVALENCE

Reports of prevalence rates of behavioral disturbances vary widely, depending upon the definition of agitated behavior used (e.g., summary terms such as "agitation" or "aggression" vs. specific behaviors such as screaming or hitting), the research methodology employed (e.g., caregiver report vs. direct observation), and characteristics of the population studied (e.g., nursing home residents vs. Alzheimer's disease patients). Prevalence rates of agitated behaviors also vary depending on whether the behavior is perceived to be a problem by others (Deutsch & Rovner, 1991). The following discussion of prevalence rates should be interpreted in light of these issues.

In a study of 408 residents in a large suburban nursing home, 93% of the residents were found to exhibit one or more of 29 agitated behaviors at least once a week as reported by caregivers using the Cohen-Mansfield Agitation Inventory (CMAI) (Cohen-Mansfield, 1986; Cohen-Mansfield, Marx, & Rosenthal, 1989). Participants in this study were identified as having various levels of cognitive impairment and physical limitations. General restlessness was the most commonly manifested behavior. Other high-frequency behaviors included pacing, repetitious sentences, constant requests for attention, complaining, negativism, and cursing. Physically nonaggressive behavior and verbally agitated behaviors were considerably more common than aggressive behaviors. In general, the frequency of agitated behaviors decreased throughout the day as the evening approached, and occurred least frequently at night when the majority of residents were sleeping. In contrast, a study by Teri, Larson, and Reifler (1988) reported that 26% of the total sample of 127 individuals with Alzheimer's disease exhibited agitation. Participants in this study were outpatients referred to a geriatric and family services clinic for a comprehensive geriatric evaluation. "Agitation" was considered separately from "wandering" and "restlessness." Levels of agitation varied with level of cognitive impairment as assessed with the Folstein Mini-Mental State Exam (MMSE; Folstein, Folstein, & McHugh 1975). Behavior problems were rated as present or not present by a physician based on observations or information provided by family of the participants. Ten percent of individuals with mild, 27% of those with moderate, and 38% of those with severe, cognitive impairment were identified as exhibiting agitation. Wandering and restlessness were evidenced in 26% and 45% of the entire sample, respectively.

7.18.4 ETIOLOGICAL FACTORS

Several factors have been proposed to contribute to the development of agitation and aggression in older adults. These factors include medication (Addonizio, 1987), illness such as infection (Yoshikawa & Norman, 1996), cognitive impairment (Cooper, Mungas, & Weiler, 1990), functional status (Cohen-Mansfield & Billig, 1986), environmental stimuli (Burgio et al., 1994; Cohen-Mansfield, Werner, & Marx, 1992; Fisher, Goy, Swingen, & Szymanski, 1994) and emotional states (Swearer, et al., 1988).

Based on the large variation in the physical and psychological functioning of older adults, the following guidelines may be useful for generating hypotheses regarding antecedent or etiological factors contributing to agitation or aggression in an older adult: (i) Given the significant variation in the health and cognitive status of older adults, identification of the etiology of aggression and agitation is best considered within an interactional model that includes physiological, psychological, and environmental factors; (ii) While an individual may exhibit topographically similar behaviors at different points in time, it should not be assumed that the same factors that contributed to the behavioral disturbance in the past are causing the current problem. Variation in etiology is a particularly salient issue when an individual has a degenerative condition such as Alzheimer's disease, in that topographically similar behaviors may have different etiologies over the course of the disease. Physiological and psychological factors that have been found to be associated with agitation and aggression in older adults are reviewed.

7.18.4.1 Physiological Factors

Physiological factors contributing to agitation and aggression can be conceptualized as falling into two categories: (i) direct precipitants of; or (ii) setting events for agitation or aggression. The term "setting events" refers to conditions of the individual or the environment in which the individual functions that affect the function of the antecedent or consequent stimuli associated with a particular response (Bijou & Baer, 1966; Kantor, 1959). Physiologically based setting events that may increase the risk of agitation or an aggressive response include patient fatigue, hunger, mood, health status, cognitive impairment, and sensory functioning. The categories of physiological factors described below may serve as a guide for generating hypotheses regarding factors influencing aggression or agitation in older adults.

Physiological conditions that have been identified as predisposing patients to become agitated or to behave aggressively include pain, physical discomfort, and sensory impairment. Evidence for their role in the etiology of agitation and aggression will be reviewed.

7.18.4.1.1 Pain

There are several painful physical conditions that may influence the occurrence of agitation and aggression. Infectious processes (Gerdner & Buckwalter, 1994), such as urinary tract infection, may produce pain in a cognitively impaired patient and may go undetected because of the inability of the patient to verbally express discomfort. Other sources of pain, such as heartburn, sinus congestion, arthritis, neuropathy from diabetes, or injury (e.g., stress fracture), may lead to agitation (Ferrell, Ferrell, & Rivera, 1995) or function as setting events in that they increase the likelihood of agitation or aggression occurring when stress is placed on the patient (e.g., when environmental stimulation is high or demands are placed on the patient in the context of caregiving).

7.18.4.1.2 Medication

Due to age-associated decreases in metabolism, older adults are at increased risk for iatrogenic effects associated with drug toxicity, drug interactions, and drug withdrawal (Gerdner & Buckwalter, 1994). Problems associated with polypharmacy have received extensive attention in the geriatrics literature for several years (Buck, 1988; Ray, Federspiel, & Schaffner, 1980; Williams & Betley, 1995). Surveys indicate that both psychotropic and nonpsychotropic medications continue to be widely over prescribed for older adults in nursing homes (Williams & Betley, 1995) and board and care homes (Spore, Mor, Hiris, Larrat, & Hawes, 1995). While polypharmacy often plays an important role in treatment of older adults, with increased numbers of medications, the risk of one medication negating, potentiating, or in some way altering the effects of another medication increases. Medications with side effects that may contribute to agitation in an elderly individual include but are not limited to neuroleptics (Addonizio, 1987), benzodiazapines (Wolf, Grohmann, Beiber, Brenner, & Luther, 1989) and nonprescription medication with anticholinergic properties (Hall, Beresford, Stickney, Nasdahl, & Coleman, 1985; see Callahan, 1993 for a review).

7.18.4.1.3 Sensory and perceptual deficits

Visual and hearing impairments may contribute to the occurrence of agitation and aggression in that they may produce or increase confusion and alter perception of the dangerousness of a situation. Perceptual impairments such as visual neglect due to stroke may also contribute to violent behavior.

7.18.4.2 Psychological Factors

7.18.4.2.1 Cognitive factors

The relationship between cognitive functioning and behavioral disturbance has received the largest amount of attention by researchers attempting to determine the cause of behavioral disturbance. Studies have examined the relationship between the number and/or types of behavioral disturbance and either cognitive status as measured by neuropsychological test performance and/or functional/adaptive skill level as measured by behavioral checklists. The results have varied. Some studies have reported a positive correlation (Cooper et al., 1990) while other researchers have found no relationship (Bolger & Dougherty, 1992; Teri, Borson, Kiyak, & Yamagishi, 1989) or a nonlinear relationship (Cohen-Mansfield, Marx, & Rosenthal, 1990) between overall number of behavior problems and cognitive status. Studies that have found a statistically significant relationship between cognitive status and behavioral disturbance have reported relatively low correlations accounting for little of the variance (e.g., Cooper et al., 1990).

In a study of Alzheimer's disease patients residing in a nursing home, Teri et al. (1988) found that the number of behavioral problems was positively correlated with cognitive impairment. The types of behavior problems manifested varied with the level of cognitive impairment. Wandering, agitation, and incontinence were found to be most heavily associated with level of cognitive impairment.

In a related study designed to further clarify the relationship between behavioral problems and level of cognitive and functional impairment in 56 moderately impaired Alzheimer's disease patients, Teri et al. (1989) found that the problems most strongly associated with level of cognitive functioning were confusion, disorientation, and memory loss. These problems constituted at least 50% of the total for problems that occurred more than twice per week as recorded by caregivers using the Behavioral Problems Checklist (BPC). The next most common problems were underactivity, loss of interest in activities, tension, depression, and apathy. No caregiver reported an absence of problems for an individual patient, and an average of seven of the 48 problems on the BPC were reported to have occurred more than twice per week. This study also identified gender differences in problem behaviors, as men were found to have significantly higher rates of the following problems: preoccupation with bodily functioning, loss of interest in activities, verbal aggression towards others, loss of interest in others, and poor personal hygiene. Perhaps the most important finding to emerge from this study is that no specific behavioral problem or pattern of behavior problems was associated with overall level of cognitive impairment.

7.18.4.2.2 Environmental factors

Research provides compelling evidence for the role of environmental influence on behavior in even very severely cognitively impaired older adults (Bourgeois, Burgio, & Schulz, 1995; Burgio et al., 1994; Vitaliano, Young, Russo, Romano, & Magana-Amato, 1993). Findings from observational research have been particularly informative in describing the influence of environmental factors on the occurrence of agitation and aggressive behavior (Anderson, 1970; Colenda & Hamer, 1991; Hamel et al., 1990; Vitaliano et al., 1993). Lindsley (1964) delineated four components for the functional analysis of behavior that can be applied to the behavior of dementia patients. These include (i) the antecedent stimulus or stimulus setting in which the behavior occurs, (ii) the behavior of interest (e.g., aggression or agitation), (iii) the contingencies of reinforcement, and (iv) the environmental events that follow the behavior. To this list others have added the concept of "setting events" (Kantor, 1959) which are conditions of the individual or the environment that affect the function of antecedent and consequent events.

(i) Setting events

Environmentally based setting events may include factors such as noise level, the presence of others, or a change in caregiver that affects the function of antecedent or consequent stimuli associated with agitation or aggression. Variation in the occurrence or intensity of aggression or agitation when the antecedent stimuli (e.g., caregiver's behavior) appear constant, may be associated with a change in environmental setting events (see Carr, Levin, McConnachie, Carlson, Kemp, & Smith, 1994). Time of day may also function as a setting event in that it has been shown that the frequency of aggressive acts decreased from day to evening to night (Cohen-Mansfield et al., 1989). Using the Cohen-Mansfield Agitation Inventory as a measure of agitated behaviors in 408 nursing home residents with varying degrees of physical and cognitive impairment, Cohen-Mansfield et al.

(1989) reported that the majority of behaviors occurred at significantly different rates across day, evening, and night shifts with frequency decreasing from day to evening to night. There are at least three possible explanations for this finding. First, patients may become increasingly fatigued as the day progresses, and therefore the decrease in agitated behaviors is a function of energy level. Second, the administration of medications and/or the application of physical restraints may decrease or preclude a patient's ability to engage in agitated or aggressive behaviors. Or third, the environment within the nursing homes changes across the day (e.g., due to changes in the caregiving or activity schedules). In short, the finding may reflect changes in environmental setting events (e.g., activities in the nursing home) or physiological setting events (e.g., fatigue or the use of medications).

The role of emotion as a setting event associated with agitation and aggression has been receiving increasing empirical attention. Research suggests that a patient's emotional state, such as level of anxiety or depression (Swearer et al., 1988) may be predictive of agitation and aggression. In a 1996 study Lawton, Van Haitsma, and Klapper (1996) reported that AD patients who were rated as "less content" were also considered more agitated by staff members. Swearer et al. (1988) found that 63% of caregivers reported that the patient for whom they provided care exhibited one or more symptoms of anxiety, including general fearfulness, a fear of specific objects or events, restlessness, pacing, and appearing to worry over insignificant matters and that these behavioral symptoms of anxiety tended to be significantly associated with angry outbursts

In a rare investigation of the association of premorbid history and agitation in nursing home residents, Cohen-Mansfield and Marx (1989) found that residents who experienced a major stressful event during their lives tended to exhibit a greater number of agitated behaviors—in particular, physically nonaggressive behavior. Residents who had not experienced relocation prior to being admitted to the nursing home exhibited more aggressive behavior and a greater number of agitated behaviors than did residents who had relocated at least once. The results of this study suggest that exposure to stressful events earlier in life may result in the development of coping skills or an increase in the threshold for perceiving events as stressful in later life. The changes produced by exposure to stress earlier in life may continue to influence behavior after the onset of a dementing process.

(ii) Antecedent stimuli

Several studies have reported that antecedent environmental stimuli such as task demands influence the occurrence of agitation and aggression. The finding of an association between environmental stimulation or caregiver demands and behavioral disturbance in cognitively impaired elderly indicates that these behaviors are not random or purposeless but instead may be functionally related to environmental stimuli.

Attempting to engage patients in activities of daily living, including being bathed, changed, or fed, has been described by nursing home staff as a primary cause of agitation (Cohen-Mansfield, 1986). These findings are consistent with those of Fisher et al. (1994) that aggressive behavior in Alzheimer's disease patients was consistently associated with a nurse's request to perform a task. In a study of 42 institutionalized dementia patients, Fisher et al. (1994) found that nurses described all physically assaultive AD patients as appearing distressed during episodes of aggressive behavior. For 80% of these patients, physically aggressive behaviors were described as being associated with an effort to escape from a stimulus, such as a nurse's request.

The concept of a "catastrophic reaction" (Goldstein, 1952) is frequently mentioned within the context of efforts to account for aggressive outbursts in dementia patients (e.g., Hall, Gerdner, Zwygart-Stuffacher, & Buckwalter, 1995). A "catastrophic reaction" occurs when demands presented by the environment exceed the patient's ability to cope. The individual becomes increasingly anxious and may attempt to terminate the anxiety-provoking stimulus by expressing a desire to be alone, screaming, or asking to go home. Continued stimulation may result in the individuals becoming extremely agitated or violent. The reports that aggression is often associated with an effort to escape from an aversive situation, described above, are consistent with the concept of catastrophic reaction.

The presence of other people has also been shown to be associated with agitation (Ryden, 1988; Cohen-Mansfield, Werner, & Marx, 1992). In a study involving 24 dementia patients described as highly agitated on the basis of CMAI scores, Cohen-Mansfield, Werner, and Marx (1992) reported that the agitated behaviors of pacing, strange noises, requests for attention, strange movements, aggressive behaviors, picking at things, and repetitious mannerisms were differentially affected by the presence of others. For example, strange noises were shown to increase when the patient was left alone while pacing decreased in the same situation.

Other environmental stimuli reported to influence behavioral disturbances include the caregiver's emotional expression (Vitaliano et al., 1993), lack of structure, and lack of attention (Cohen-Mansfield, 1986).

These findings support the hypothesis that certain behaviors that emerge during the course of dementia may serve a function for dementia patients in that they may be an adaptive (and the only available) means of increasing or decreasing stimulation. As a patient's behavioral repertoire diminishes over the course of a dementing process, behaviors characterized as "agitated" or "disturbances" may emerge to serve the function of previously available behaviors (e.g., those associated with language, leisure activities, or social interaction).

(iii) Environmental consequences

The consequences of agitated behaviors and aggression have received the least amount of attention by researchers. Little is known regarding how caregiver behavior following aggression functions to maintain or reduce aggression. Several possible relationships between behavior and consequences may be relevant. Aggressive behavior may be influenced by positive reinforcement if the former's occurrence is associated with the presentation of a consequent stimulus (e.g., caregiver attention) that increases the probability of the behavior occurring in the future. Alternatively, negative reinforcement may serve to strengthen aggressive tendencies through the contingent removal of an aversive stimulus (e.g., task demand). A few case studies (e.g., Bakke et al., 1994) discussed in detail in the intervention section, have described functional relationships between agitated behavior and environmental consequences.

(a) Social consequences. Aggression associated with positive reinforcement in the form of attention appears to be rare in dementia patients, although a few case reports have described an association between agitation and social contact (e.g., Rosberger & MacLean, 1983). When social contact is implicated in an aggressive episode, it is more often associated with an effort to terminate social contact (Fisher, 1995).

(b) Tangible consequences. Aggressive behaviors exhibited by cognitively impaired older adults may be associated with an effort to gain a tangible consequence such as food or an object. The role of tangible consequences in the development and maintenance of aggression within community settings has not been explored. Within institutional settings such as nursing homes, access to tangible objects is limited for safety purposes. The most frequent type of tangible consequence for a behavior that appears aggressive involves obtaining food from another resident (e.g., grabbing food from another resident) (Fisher et al., 1994).

In another scenario, a patient may grab onto or pinch the caregiver in a painful fashion during caregiving tasks. While the behavior may appear aggressive in that it causes pain for the caregiver, it may function to provide security or stimulation. Identifying an alternative response that is incompatible with the seemingly aggressive one, such as placing an object (e.g., a towel or small pillow) in the patient's hand prior to initiating the caregiving activity may prevent this form of aggression (Lundervold & Lewin, 1992).

7.18.5 ASSESSMENT

Interest in the assessment of behavioral disturbances in AD has increased since the mid-1980s. However, there are several limitations to our understanding of agitation and aggression and the reliability and validity of assessment procedures. First, there is a lack of consensus in the field regarding the definitions of agitation and aggression. Second, as discussed earlier in this chapter, the majority of available assessment instruments focus on the topographical features of agitation and aggression and have limited utility for treatment planning. Third, few of the available instruments have been validated for use with dementia patients. Psychometric studies have focused primarily on the reliability of instruments. The predictive validity and treatment utility of the majority of instruments have not been empirically evaluated.

The physiological, topographical, and functional characteristics of agitation and aggression in cognitively impaired older adults and the significance of each of these areas in assessment and treatment planning are explored in the following sections.

7.18.5.1 Physiological Assessment of Agitation and Aggression

For the purposes of determination of etiology and treatment planning, assessment of agitation should begin at the physiological level. Assessing and treating physical symptoms contributing to agitated behavior may eliminate the need for the more time-consuming assessment of environmental causes of agitation. While there have been few empirical investigation regarding the physiological correlates of agitation in dementia patients, pain and sensory impairment have been implicated in the etiology of agitation and is reviewed.

7.18.5.1.1 Pain

A particularly difficult challenge in caring for cognitively impaired individuals involves their limited verbal abilities. As verbal abilities decline, caregivers must increasingly rely on observation to determine whether a patient is experiencing pain or distress. It is crucial for adequate treatment planning to rule out physical pain as a precipitant to aggression. Regular physical examinations are a necessary component of adequate care.

Although assessment of pain is complicated by the impaired verbal abilities of the dementia patients, Ferrell et al. (1995) reported that 83% of a sample of 217 cognitively impaired dementia patients was able to complete at least one of the five unidimensional measures used to assess pain. The unidimensional measures used included: (i) a verbally administered questionnaire in which subjects were asked to choose among descriptors of pain listed in one-inch letters on large index cards (McGill Pain Questionnaire; Melzack, 1975); (ii) a verbally administered 0–10 Likert scale; (iii) a 100 mm visual analog scale with word anchors of "no pain" and "worst possible pain" (p. 593); (iv) a subscale of the Memorial Pain Card (Fishman et al., 1987) involving the visual and verbal presentation of eight words describing various intensities of pain ranging from no pain to excruciating; and (v) the Rand Coop Chart for pain (Nelson et al., 1987) that requires subjects to choose one of five cartoon characters accompanied by descriptive words depicting various levels of pain, ranging from no pain to severe pain.

Ferrell et al. (1995) report that 62% of subjects who were able to complete the assessment reported the presence of pain and 40% of the sample reported pain in more than one location. Further, for 37% of patients who reported pain, no etiology for the described pain was documented in the medical records. These findings indicate the high prevalence or pain among dementia patients.

While Ferrell et al. (1995) indicate that this study provides support for the ability of cognitively impaired dementia patients to report pain, several factors may limit the generalizability of the findings to moderately and severely impaired patients. First, the authors did not report the reasons why 17% of the sample was unable to complete the interview. The inability of a significant subset of the sample to complete the interview suggests that the results may not be generalizable to more impaired patients. Second, it is unclear whether the discrepancy between subject reports and medical records reflects inaccurate reporting by the dementia patients or insufficient documentation in patient charts. Research indicates that while interviewing is a helpful assessment tool with mildly impaired individuals, it is not effective as cognitive decline progresses (Davis & Robins, 1989). The limited utility of interviews with dementia patients is reflected by the elimination of 70 subjects prior to selection of the final sample in the Ferrell et al. study. The authors report that these subjects were eliminated, because they were "essentially mute and unresponsive, and no meaningful information could be obtained from the patient interview" (p. 594). The omission of severely impaired patients reflects the difficulty in clarifying the relationship between pain and agitation or aggression in dementia patients, since these behaviors are generally more likely to appear when cognitive impairment is moderate to severe (Kurita, Blass, Nolan, Black, & Thaler, 1993; Swearer et al., 1988).

Given the limitations of the use of interviews for assessing pain in severely impaired AD patients, it is usually necessary to rely on direct observation methods. Hurley, Volcier, Hanrahan, Houde, and Volcier (1992) developed the Discomfort Scale for Dementia of the Alzheimer's type (DS-DAT) to assess discomfort in AD patients through direct observation of facial expressions and body posture. The authors defined discomfort as "a negative emotional and/or physical state subject to variation in magnitude in response to internal or environmental conditions" (p. 369). Behavioral signs of discomfort on the DS-DAT include noisy breathing; negative vocalization; having a frown; fidgeting; and looking sad, frightened, or tense.

The construct validity of the DS-DAT was evaluated by administering the scale monthly to AD patients who were highly dependent on the nursing staff for their activities of daily living. During the course of the study 20 subjects were identified as having an acute illness typically manifested by fever. It was hypothesized that discomfort would be higher during a fever episode than during baseline. The results indicated that the DS-DAT was sensitive to changes in discomfort associated with fever episodes. A limitation of this study is that the observers were not blind to the patients' fever episodes. While this knowledge may have biased observations, the results indicate that reliable, systematic observation of discomfort in AD patients is possible. Such observation may provide caregivers with information necessary to assist in assessment of physiological factors in the etiology of agitation.

While no studies, up to 1997, have directly assessed the relationship between pain and agitation, it is important to continue to explore this relationship. The promising developments in the assessment of pain in dementia patients described above, support the need to explore this venue. Reliable methods for assessing pain in the cognitively impaired individuals may help in the identification of physiological factors contributing to agitation and aggression and, hence, inform treatment.

7.18.5.1.2 Sensory impairment

Sensory impairment may also contribute to agitation (Gerdner & Buckwalter, 1994). Dementia patients, who are unable to correct for inadequacies in sensory perception by gathering more information or by altering their behavior (Fisher & Carstensen, 1990) may experience significant confusion and subsequent distress associated with sensory decrements.

Deficits in vision and hearing may contribute to a patient's inaccurate perception of situational demands. Patients with severe sensory impairment may overestimate the dangerousness of a situation and behave aggressively out of fear and a perceived need for self-defense. Assessing the status of a dementia patient's sensory functioning may provide information regarding how they perceive the environment and whether sensory deficits are contributing to a tendency toward agitation or aggression.

7.18.5.2 Topographical Assessment of Agitation and Aggression

The majority of research regarding the assessment of agitation and aggression has addressed the topographical features of the behaviors. Topographical assessment is informative for the identification of target behaviors and the frequency and intensity of these behaviors. Topographical information is also critical for assessment of baseline levels of aggressive or agitated behavior and for monitoring the effects of interventions. Four assessment instruments designed to record the frequency and intensity of agitation or aggression will be reviewed.

7.18.5.2.1 Cohen-Mansfield Agitation Inventory

The Cohen-Mansfield Agitation Inventory (CMAI; Cohen-Mansfield, 1986) is a 29-item questionnaire designed to measure the types and frequencies of agitated behaviors exhibited by elderly nursing home residents. Respondents rate the frequency on a seven-point scale ranging from "1—never engages in" to "7—manifests the behavior on the average of several times an hour." Examples of items include "pacing," "cursing or verbal aggression," and "repetitious sentences or questions." Cohen-Mansfield et al. (1989) reported adequate interrater agreement for each behavior on the CMAI (average $r = 0.92$). Studies of the reliability of the CMAI by other groups have produced varying levels of support for interrater reliability. Finkel, Lyons, & Anderson (1992) found evidence for high internal consistency but high variation in interrater reliability across items. In 1995, Miller, Snowdon, and Vaughan (1995) reported adequate interrater reliability for aggressive ($r = 0.85$) and physically nonaggressive behaviors ($r = 0.73$) but low interrater reliability for agitated verbal behaviors ($r = 0.47$).

7.18.5.2.2 Pittsburgh Agitation Scale

The Pittsburgh Agitation Scale (PAS; Rosen et al., 1994) is designed to assess four categories of agitated behavior in dementia patients, including aberrant vocalization, motor agitation, aggressiveness, and resisting care. Behaviors are rated using a five-point scale ranging from "0" indicating that the behavior was not observed during the specified observation period, to "4" indicating severe intensity, frequency, and/or resistance to redirection of the behavior. The authors report adequate intraclass correlations and concurrent validity based on findings from studies of dementia patients from an acute-care geropsychiatric unit and a chronic-care nursing facility.

7.18.5.2.3 Overt Aggression Scale

The Overt Aggression Scale (OAS; Yudofsky et al., 1986) is a checklist designed to document specific aspects of aggressive behavior based on observable criteria. It is designed to be completed immediately following each aggressive incident. In addition to indicating which of 16 behaviors occurred during an incident, raters also record the intervention strategies used and the time of onset and duration of the incident. The 16 behaviors on the OAS are divided into four categories: verbal aggression, physical aggression against objects, physical aggression against self, and physical aggression against others. Sample behavioral items include "Makes loud noise, shouts angrily," "Bangs head, hits fists into objects, throws self onto floor or into objects (hurts self without serious injury)," and "Slams doors, scatters clothing, makes a mess." Sample intervention items on the OAS include "Talking to patient," "Closer observation," and "Use of restraints." Silver and Yudofsky (1991) reported that the OAS captured 98% of aggressive incidents documented in medication books, hospital charts, and ward communication books combined, indicating adequate concurrent validity.

A limitation of the OAS is that it may be logistically difficult to complete the form for each incident of aggression when they occur at a high frequency (Ratey & Gutheil, 1991; Sorgi, Ratey, Knoedler, Markert, & Reichman, 1990). In an effort to eliminate this problem, Sorgi et al. (1990) developed the retrospective OAS using the original 16 items contained in the OAS, but with a modified scoring and temporal interval. The frequency of occurrence for each item is rated on a five-point Likert-type scale ranging from "0" to "greater than 10 times." The retrospective OAS is designed to reflect the patient's behavior over the previous week. The authors report adequate interrater reliability ($r = 0.96$) and internal consistency ($\alpha = 0.75$). Ratey and Gutheil (1991) also developed a modified version of the OAS to reduce respondent burden. However, they did not report information regarding its psychometric properties.

While empirical reports of the use of the OAS and the retrospective OAS with dementia patients have not emerged, given that the behaviors and interventions included on the OAS are observed in dementia patients, it is reasonable to assume that the instruments are appropriate for use with this population.

7.18.5.2.4 Blessed Dementia Scale

The Blessed Dementia Scale (BDS; Erkinjuntti, Hokkanen, Sulkava, & Palos 1988) is a 22-item scale that assesses cognitive functions, activities of daily living, instrumental activities of daily living, and personality characteristics. The BDS is administered within the context of interview with an informed respondent. Activities of daily living are rated on the basis of the frequency with which the client is able to perform the task (e.g., daily, occasional, always). Changes in personality, interests, and drive are rated as present or absent. Examples of items from the activities of daily living category include "Inability to perform household tasks" and "Inability to recall recent events." Personality items include "Increased egocentricity," "Impairment of emotional control," and "Sexual misdemeanor." The BDS has been shown to discriminate between institutionalized dementia patients and elderly community residents with no organic impairment.

7.18.5.3 Assessment of Functional/Contextual Features

Information regarding the frequency, duration, and intensity of agitation and aggression is necessary but insufficient for treatment planning. Research has consistently indicated that environmental factors influence behavioral disturbances in dementia patients, and therefore systematic identification of the environmental characteristics associated with agitation or aggression is essential for the determination of whether a behavior is environmentally dependent, and if it is, the features of the environment on which it is dependent. Specific factors that merit attention in the functional analysis of agitation include setting events, antecedent environmental stimuli, response characteristics, and response consequences.

7.18.5.3.1 Setting events

Setting events include, but are not limited to, factors such as patient mood, fatigue, level of cognitive impairment, physical status, or sensory functioning that affect the relationship between antecedent stimuli and consequences associated with aggression or agitation. A challenge in developing interventions for patients with a degenerative dementia is the inherently dynamic nature of the relationship between setting events and behavior as neuropathology progresses. An idiographic approach to assessment will be the most effective for addressing variation in the interaction between cognitive impairment, environmental stimuli, and behavior.

7.18.5.3.2 Antecedent environmental stimuli

Evidence for the importance of identifying antecedent stimuli associated with aggression and agitation has emerged from several studies that have documented a systematic relationship between the occurrence of these behaviors and various environmental stimuli (Cohen-Mansfield, Marx, Werner, 1992; Ryden, 1988). When, through repeated pairing of a stimulus and a particular reinforcement contingency, a stimulus comes to signal that reinforcement is available contingent on a particular response, the behavior is considered to be under stimulus control. The stimulus which signals that reinforcement is available contingent on a particular response is termed a discriminative stimulus.

Factors associated with the development and maintenance of the stimulus control of behavior in dementia patients have not be delineated. Idiographically based methods for determining whether a particular behavior exhibited by a dementia patient is under stimulus control are described below.

7.18.5.3.3 Response characteristics

In consideration of response characteristics, the topography of aggression and agitation in dementia patients has been the primary focus of empirical attention. Patients' emotional responses to environmental stimuli have not been systematically investigated. Cohen-Mansfield (1986) reported that lack of attention can be a precursor to agitation. It is possible that lack of attention leads to boredom, thus producing a functionally adaptive self-stimulatory behavior in dementia patients, such as stereotypic movement.

Another emotional response characteristic that may influence response to environmental stimuli and the subsequent occurrence of agitation or aggression is anxiety. Hall et al. (1995) suggested that AD patients may overestimate the dangerousness of a caregiver session, for example, as a result of confusion caused by cognitive impairment. Perceiving the environmental stimuli as dangerous may result in high levels of anxiety. Due to cognitive limitations the patient's repertoire of behaviors for coping with an anxiety-provoking stimulus is reduced (e.g., the patient is unable to verbalize a request to terminate the interaction), and the patient may strike the caregiver. The aggressive behavior may be negatively reinforced by termination of the caregiver's demands and subsequent reduction in the patient's anxiety. In this situation assessment of emotion as a response characteristic may suggest a proactive intervention such as presenting the patient with a distracting stimulus such as soothing music or a soft object to carry.

7.18.5.3.4 Environmental consequences

The consequences of agitation and aggression have received the least amount of attention by researchers. Little is known about how patient and caregiver behavior following behavioral disturbances function to maintain or reduce these behaviors. It is possible that termination of caregiver's demands or the reduction of an aversive emotional state may function to negatively reinforce agitated or aggressive behavior.

The instruments described below may be useful for formulating hypotheses regarding the functional context of agitated behaviors.

7.18.5.3.5 Motivation Assessment Scale

The Motivation Assessment Scale (MAS; Durrand & Crimmins, 1988) is a 16-item questionnaire designed to identify variables maintaining self-injurious behavior. It addresses four categories of situational determinants of self-injurious behavior including social attention, tangible consequences, escape from unpleasant situations, and sensory consequences. Respondents are asked to rate the likelihood of the specified target behaviors occurring in each of the situations listed using a seven-point Likert-type scale ranging from "0—never" to "6—always." Sample questions include, "Does this behavior occur repeatedly, over and over, in the same way?" and "Does this behavior occur when you are talking to other persons in the room?" The four categories on the MAS were supported by a factor analysis conducted by Bihm, Kienlen, Ness, and Poindexter (1991).

Durrand and Crimmins (1988) investigated the psychometric properties of the MAS on a sample of 50 developmentally disabled children. They found that the scale had adequate internal consistency, test–retest, and interrater reliability. The scale also demonstrated adequate concurrent validity. The reliability and validity of this scale for use with dementia patients have not been empirically investigated.

Subsequent studies on the interrater reliability of the MAS have produced varied results. Zarcone, Rodgers, Iwata, Rourke, and Dorsey (1991) applied the MAS to a population of subjects with severe disabilities using self-injurious behavior as the target behavior and reported interrater agreement ranging from -0.24 to 0.44. Sigafoos, Kerr, and Roberts (1994) also found low interrater reliability ($r = 0.03$) of the MAS when applied to subjects with profound intellectual disability and using aggressive behavior as targets. However, Kearney (1994) reported adequate interrater reliability ($r = 0.55$–0.79) when using the scale with a sample of mentally retarded adults and a variety of agitated behaviors (e.g., aggression, self-injury, screaming) as targets.

7.18.5.3.6 Behavior Environment Taxonomy of Agitation

The Behavior Environment Taxonomy of Agitation (BETA; Fisher, 1995) is a 16-item checklist that asks respondents to describe the likelihood of specific agitated behaviors occurring in any of four classes of situations. The BETA represents a modification of the MAS. A preliminary study of the utility of the MAS for use with AD patients, involving 17 AD patients, found that the category "tangible consequences" was ranked last (i.e., least relevant) for 58% of cases with a mean ranking of four out a possible five. Items associated with tangible consequences were endorsed as relevant almost exclusively for the behavior "Grabbing." This association was judged to be uninformative. Therefore, in an effort to reduce respondent burden associated with completing the BETA and to improve its utility, items pertaining to tangible consequences were eliminated. The situations described in the BETA have been identified in the literature (Anderson, 1970; Colenda & Hamer, 1991; Hamel et al., 1990) and by caregivers (Fisher et al., 1994) as being associated with the occurrence of agitation in dementia patients and in other clinical populations, including developmentally disabled persons and psychiatric patients (Durand & Crimmins, 1988; Paul & Lentz, 1977). The classes of situations on the BETA include: (i) social attention; (ii) escape from unpleasant situation; (iii) increase sensory consequences; and (iv) reduce sensory consequences Nurses and CNAs rate the likelihood of occurrence of four examples of each of these classes of situations on a seven-point Likert-type scale ranging from "0—never" to "6—always."

To evaluate the interrater reliability of the BETA, mean scores for the four BETA categories were ranked. The approach of ranking of category scores was used, because this approach is most consistent with how the categories are used clinically. The highest mean score was given a rank of one. The other mean scores were given ranks of two, three, and four, respectively. Tie ranks were averaged. Spearman rank-order correlation coefficients for the category scores indicated consistency in BETA ratings between respondents (median $r = 0.83$, $p < 0.01$). Test–retest reliability was evaluated using Spearman rank-order correlation coefficients for the category scores. Results indicated consistency in BETA ratings over time (two to three week interval) (median Spearman $r = 0.80$, $p < 0.01$).

Initial investigations regarding the factor structure of the BETA using a normalized varimax rotation factoring program revealed that the four factors that emerged are essentially equivalent to the subscales proposed in the revision of the BETA. The exceptions appear to be associated with the category "Reduce sensory stimulation." Items which include a reference to "stimulation" may load more heavily on a separate factor. This finding may also be associated with respondents' interpretation of the term "stimulation."

7.18.5.3.7 Functional Analysis of Behavior Recording and Tracking Forms

The Functional Analysis of Behavior (FAB; Lundervold and Lewin, 1992) is designed to assess the antecedents and consequences of verbal abuse and combativeness. It is divided into two parts each of which is completed when the target behavior occurs. Respondents record the date and time of each occurrence and place a checkmark next to the situation that immediately preceded the behavior. Examples of situations described on the form include "Meal time," "Bath time," "Dressing," and "Command." A brief description of the intervention that was implemented to address the behavior problem is also recorded.

The FAB Tracking Form (Lundervold & Lewin, 1992) is designed for monitoring the frequency with which several target behaviors occur and treatment progress. Respondents record (i) the type of behavior exhibited and (ii) the time interval within which the problem behavior occurred. Behavioral categories include aggression, property destruction, self-injurious behavior, and noncompliant behavior. No information is available on the reliability and validity of these instruments.

7.18.6 INTERVENTIONS

7.18.6.1 Pharmacological Interventions

A search of the medical, nursing, and psychological literature on treatment studies for aggression and agitation reveals that the vast majority of treatment studies on agitation and aggression are pharmacological. Several types of medication have been used in the treatment of agitated behaviors and aggression in older adults. These include neuroleptics, anticonvulsants, antidepressants, and antianxiety agents. A detailed consideration of the pharmacology of these agents is beyond the scope of this chapter. Studies of treatment efficacy are briefly reviewed.

7.18.6.1.1 Neuroleptics

Neuroleptics are the most commonly prescribed medications for reducing behavioral symptoms in older adults (Devenand & Levy, 1995). Their effects on agitation have been widely documented (e.g., Barnes, Veith, Okimoto, Raskind, & Gumbrecht, 1982; Coccaro et al., 1990; Schneider, 1990) although their mechanism of action has not been clearly established. In a meta-analytic review of 17 placebo-controlled studies of the use of neuroleptics with dementia patients, Schneider, Pollock, & Lyness (1990) reported that 18% of patients improved with neuroleptic treatment.

The use of neuroleptics for the management of aggression and agitation remains controversial because of the severity of associated side effects, including sedation, orthostatic hypotension, and akathesia (Finkel, 1993). Age-associated changes in metabolism increase the risk and severity of the side effects of neuroleptic use in older adults. Extrapyramidal reactions and tardive dyskinesia are also risk factors. Anticholinergic effects may exacerbate symptoms of dementia resulting in increased confusion, memory loss, and disorientation.

7.18.6.1.2 Benzodiazepines

Benzodiazepines are commonly used for the treatment of insomnia, anxiety, and agitation in older adults. In comparing the efficacy of benzodiazepines and neuroleptics in reducing agitation in cognitively impaired nursing home residents, Coccaro et al. (1990) reported a greater treatment response for neuroleptics. Side effects associated with benzodiazepines include sedation, cognitive impairment (Larson, Kukull, Buchner, & Reifler, 1987) and cerebellar ataxia resulting in an increased risk of falls (Ray, Griffin, Schaffner, Bough, & Milton, 1987). Paradoxically, when benzodiazepines are administered for agitation they may result in daytime sedation which, when coupled with the cognitive impairment associated with dementia, can result in increased irritability and agitation (Finkel, 1993). Further, withdrawal from benzodiazapines can result in serious disturbance, including insomnia, agitation, and psychosis (Wolf et al.,1989).

7.18.6.1.3 Beta blockers

Several studies describe the use of beta blockers for the treatment of agitation in elderly dementia patients (e.g., Greendyke, Kantor, & Schuster, 1986; Weiler & Goodman, 1987; Weiler, Mungas, & Burnick, 1988). Beta blockers such as propranolol affect agitation levels by blocking beta-adrenergic receptors and thereby decreasing sympathetic activity. Up to 1997, no well-controlled studies have evaluated the effects of beta blockers on agitation in cognitively impaired older adults. In a review of research on beta blockers for the treatment of agitation, Schneider and Sobin (1991) describe preliminary support for efficacy based on small samples of patients with various dementia diagnoses. Side effects involve both cardiovascular and central nervous systems. Propanolol is contraindicated in patients with sinus bradycardia, congestive heart failure, heart block greater than first degree, and bronchial asthma.

7.18.6.1.4 Other antianxiety agents

Anxiolytics such as buspirone are receiving increasing attention in the management of agitation and aggression in cognitively impaired older adults. Buspirone has fewer side effects than other antianxiety agents and the advantages of not producing memory impairment, psychomotor deficits, sedative, or addictive effects (Gelenberg, 1994). When side effects do occur they tend to be associated with higher initial doses (e.g. 10 mg t.i.d. rather than 5 mg t.i.d.). These include dizziness, nausea, and headache. Well-controlled treatment outcome studies of the effects of buspirone with dementia patients have not been conducted to date. However, the few studies that have been reported indicate that buspirone is effective in reducing aggression in dementia patients, with few side effects (Colenda, 1988; Herrmann & Eryavec, 1993; Jeanblanc & Davis, 1995).

7.18.6.1.5 Antidepressants

Several antidepressants have been used in efforts to reduce agitation in older adults. Serotonin uptake inhibitors have been receiving increased attention in the treatment of behavioral disturbance in dementia patients. Nyth and Gottfries (1990) evaluated the efficacy of citalopram for the treatment of emotional disturbance in 98 AD patients within a placebo-controlled group design. Significant improvement in agitation was reported for patients in the citalopram group.

7.18.6.1.6 Anticonvulsants

Evidence for the effectiveness of anticonvulsants for the treatment of agitation is limited to clinical case studies. A review of reports by Schneider and Sobin (1991) found 22 cases of agitated AD patients treated with anticonvulsants. Four patients were described as improved, with the remaining 18 not responding or developing toxicity. No well-controlled studies have evaluated the efficacy of these medications for the treatment of agitation in dementia patients. The most severe reactions associated with the use of carbamazine include dizziness, drowsiness, unsteadiness, nausea, and vomiting (Physician's desk reference, 1996).

7.18.6.2 Psychological Interventions

Empirical accounts of psychological interventions for managing behavior disturbances in older adults provide compelling support for preventative or proactive behavioral interventions (Lundervold & Lewin, 1992). Preventative approaches emphasize environmental design and the stimulus control of agitated behaviors and aggression. Given the paucity of information regarding the influence of the dementing disorders on response to consequences, interventions that attempt to *prevent* the occurrence of these behaviors may be the most judicious. Stimulus control interventions are described first.

7.18.6.2.1 Environmental stimuli

There are four components of the setting and sequence of events that warrant attention in determining whether agitated behavior or aggression are associated with environmental stimuli. The four components are: setting events, antecedent stimuli, patient response characteristics, and the consequences of the agitation or aggressive response.

(i) Setting events

When physiological or environmental setting events are identified as being reliably associated with aggressive responding, the context of the problem behavior can be modified to reduce the probability of the behaviors occurring. To illustrate, if aggression tends to occur during caregiving in the evening and not early in the day, then a reasonable intervention may be to reduce demands over the course of the day. Completing tasks early in the day will reduce the influence of fatigue and minimize the risk of a catastrophic reaction. If assessment reveals an association between specific environmental conditions and aggression (e.g., aggression occurs only when certain caregivers are providing care) the level of intervention will involve restructuring the physical or social context (e.g., altering the communication style of the caregiver).

In considering the role of antecedent stimuli associated with aggression and agitation the behavior of the caregiver clearly warrants attention. Specifically, the manner in which demands/requests are presented, even when quite benign in appearance, may provoke agitation or aggression. When verbal information or physical prompts are considered within the context of a patient's limited competencies it is clear that they should be presented individually and at a slower pace than normal. Instructions or requests should be made within the context of short sentences to accommodate impairment in language abilities and short-term memory. Caregiver eye contact and touch may also influence the likelihood of an undesirable response.

(ii) Response characteristics

A patient's emotional response before and during aggressive episodes may have implications for treatment planning. In a study of behavioral disturbance in dementia patients by Swearer et al. (1988), 63% of caregivers surveyed reported that the patient for whom they provided care exhibited one or more symptoms of anxiety, including general fearfulness, a fear of specific objects or events, restlessness, pacing, and appearing to worry over insignificant matters. Behavioral symptoms of anxiety were also found to be significantly associated with angry outbursts. Based on the observed association between aggression and anxiety, two treatment approaches warrant consideration. These involve the reduction of the anxiety-provoking characteristics of demand tasks and/or altering the consequences of the aggression.

The reduction of anxiety in cognitively impaired older adults can involve two strategies: (i) modifying the setting events or antecedent stimuli associated with the behavioral disturbance as described above, or (ii) implementing a distraction-based intervention to draw the patient's attention away from the distressing stimulus (e.g., instructing the caregiver to sing during caregiving episodes). Distraction-based techniques have been described for the reduction of anxiety in noncognitively impaired individuals (e.g., Beck, 1995; Barlow, 1988). The utility of such methods for dementia patients warrants further consideration by researchers.

If aggressive behavior tends to occur when demands to perform a task are placed on dementia patients, the aggressive response may be negatively reinforced by termination of demands. Further, if anxiety precipitates the aggression and the aggressive behavior results in a reduction of anxiety through termination of the anxiety-provoking stimulus (e.g. caregiver) the aggressive behavior can be viewed as adaptive since it may be the only readily accessed response in the patient's repertoire (Fisher & Swingen, 1996).

(iii) Social consequences

Assessment may reveal that the aggressive behavior of a dementia patient is contingently associated with social consequences (e.g., attention by caregivers). Differential reinforcement of incompatible behavior (DRI) may be a reasonable option to consider in this situation. To illustrate, Roseberger and MacLean (1983) described the implementation of this procedure in the treatment of a 79 year-old stroke patient residing in a nursing home. The patient was aphasic and partially paralyzed. Problem behaviors included physically aggressive behavior directed at staff, other residents, and visitors, throwing objects, smearing feces, and disrobing in front of visitors. Mechanical restraint and medication (e.g. Haldol and Mellaril) resulted in unacceptable side effects. Following assessment of the behaviors identified as the most disruptive, a DRI intervention was implemented wherein staff were instructed to greet the resident and praise her cooperation whenever she was observed to be behaving appropriately. In addition, problem behaviors resulted in a brief time-out from social attention. The intervention proved effective in that the resident's aggressiveness was reduced from 16 episodes per week to zero by the third week. Treatment effects were maintained at a 5 month follow-up.

A case study involving a 72 year-old man diagnosed with AD also demonstrates the effective contingent application of social consequences for reducing agitated behavior (Bakke et al., 1994). The patient had been prescribed buspirone for the agitation and had been expelled from two day treatment centers previously because of the agitation. On the basis of interview and observational data it was hypothesized that this man's frequent efforts to leave a work program and statements about going home were maintained by consistent attention from staff, while on-task, in-seat behavior resulted in social attention only 15% of the time. The intervention involved increasing the rate of tangible and social reinforcement contingent on working while incorporating efforts to reduce fatigue, discomfort, and boredom through restructuring of the work setting and schedule. Results indicated that this multicomponent program produced a significant decrease in agitated behavior while maintaining work productivity.

Understanding the contingencies of reinforcement associated with the behavior of dementia patients presents unique challenges for researchers and caregivers in that the progression in neuropathology limits the adequacy of traditional standards of treatment generalization and maintenance (Fisher & Carstensen, 1990). As is the case for physiologically based interventions, the effects of environmentally based interventions for dementia patients require continuous monitoring of changes in the reciprocal influence between behavior and the environment associated with deterioration in brain function. Stimulus control procedures, designed to prevent the occurrence of problem behaviors may be the most reasonable approaches to pursue until information regarding the efficacy of other environmentally based interventions is available (Fisher & Carstensen, 1990; Hussian, 1988).

ACKNOWLEDGMENT

Preparation of this chapter was supported by Grant R29 AG11241 from the National Institute on Aging.

7.18.7 REFERENCES

Addonizio, G. (1987). Neuroleptic malignant syndrome in elderly patients. *Journal of the American Geriatrics Society, 35,* 1011–1012.

Anderson, J. F. (1970). A study of disturbed behavior in patients with dementia in two hospital populations. *Gerontology Clinics, 12*, 49–64.

Bakke, B. L., Kvale, S., Burns, T., McCarten, J. R., Wilson, L., Maddox, M., & Cleary, J. (1994). Multiple component intervention for agitated behavior in a person with Alzheimer's disease. *Journal of Applied Behavior Analysis, 27,* 175–176.

Barlow, D. H. (1988). *Anxiety and its disorders: The nature and treatment of anxiety and panic*. New York: Guilford Press.

Barnes, R., & Raskind, M. (1980). Strategies for diagnosing and treating agitation in the aging. *Geriatrics, 35*(3), 111–119.

Barnes, R., Veith, R., Okimoto, J., Raskind, M., & Gumbrecht, G. (1982). Efficacy of antipsychotic medications in behaviorally disturbed dementia patients. *American Journal of Psychiatry, 139*, 1170–1174.

Beck, J. S. (1995). *Cognitive therapy: Basics and beyond.* New York: Guilford Press.

Bihm, E. M., Kienlen, T. L., Ness, M. E., & Poindexter, A. R. (1991). Factor structure of the motivation assessment scale for persons with mental retardation. *Psychological Reports, 68,* 1235–1238.

Bijou, S. W., & Baer, D. M. (1966). Methods in child behavior and development. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application.* Englewood Cliffs, NJ: Prentice-Hall.

Bolger, J. P., & Dougherty, L. M. (1992). Psychological determinants of aggressive behavior in geriatric inpatients. *Behavior, Health, and Aging, 2*(1), 3–11.

Bourgeois, M. S., Burgio, L. B., & Schulz, R. (1993). Teaching caregivers of spouses with Alzheimer's disease to modify problem behaviors in the home. *The Gerontologist, 33*, 260.

Buck, J. A. (1988). Psychotropic drug practice in nursing homes. *Journal of the American Geriatrics Society, 36,* 409–418.

Burgio, L. D., Scilley, K., Hardin, J. M., Janosky, J., Bonino, P., Slater, & Engberg, R. (1994). Studying disruptive vocalization and contextual factors in the nursing home using computer-assisted real-time observation. *Journal of Gerontology: Psychological Sciences, 49*, P230–P239.

Callahan, C. M. (1993). Psychiatric symptoms in elderly patients due to medications. In J. Rowe & J. C. Ahronheim (Eds.), *Annual review of gerontology and geriatrics: Focus on medications and the elderly* (Vol. 12). New York: Springer.

Carr, E. G., Levin, L., McConnachie, G., Carlson, J.I., Kemp, D. C., & Smith, C. E. (1994). *Communication-based interventions for problem behavior: A users guide for producing positive change*. Baltimore: Paul H. Brookes.

Coccaro, E. F., Kramer, E., Zemishlany, Z., Thorne, A. Rice, C.M., Giordani, B., Duvvis, K., Patel, B.M., Tores, J., & Nora, R. (1990). Pharmacologic treatment of noncognitive behavioral disturbances in elderly demented patients. *American Journal of Psychiatry, 147*, 1640–1645.

Cohen-Mansfield, J. (1986). Agitated behaviors in the elderly, II: Preliminary results in the cognitively deteriorated. *Journal of the American Geriatrics Society, 34,* 722–727.

Cohen-Mansfield, J., & Billig, N. (1986). Agitated behaviors in the elderly, I: A conceptual review. *Journal of the American Geriatrics Society, 34,* 711–721.

Cohen-Mansfield, J., & Marx, M. S. (1989). Do past experiences predict agitation in nursing home residents? *International Journal of Aging and Human Development, 28*(4), 285–294.

Cohen-Mansfield, J., Marx, M. S., & Rosenthal, A. S. (1989). A description of agitation in a nursing home. *Journal of Gerontology: Medical Sciences, 44*(3), M77–M84.

Cohen-Mansfield, J., Marx, M. S., & Rosenthal, A. S. (1990). Dementia and agitation in nursing home residents: How are they related? *Psychology of Aging, 5*(1), 3–8.

Cohen-Mansfield, J., Marx, M. S., & Werner, P. (1992). Agitation in elderly persons: An integrative report of findings in a nursing home. *International Psychogeriatrics, 4*(Suppl. 2), 221–240.

Cohen-Mansfield, J. Werner, P., & Marx, M. S. (1992). The social environment of the agitated nursing home resident. *International Journal of Geriatric Psychiatry, 7,* 789–798.

Colenda, C. C. (1988). Buspirone in treatment of agitated demented patient. *The Lancet, ii*, 510.

Colenda, C. C., & Hamer, R. M. (1991). Antecedents and intervention for aggressive behavior of patients at a geropsychiatric state hospital. *Hospital and Community Psychiatry*, *42*, 287–292.

Cooper, J. K., Mungas, D., & Weiler, P. G. (1990). Relation of cognitive status and abnormal behaviors in Alzheimer's disease. *Journal of the American Geriatrics Society, 38*, 867–870.

Davis, P. B., & Robins, L. N. (1989). History-taking in the elderly with and without cognitive impairment: How useful is it? *Journal of the American Geriatrics Society, 37,* 249–255.

Deutsch, L. H., & Rovner, B. W. (1991). Agitation and other noncognitive abnormalities in Alzheimer's disease. *The Psychiatric Clinics of North America, 14*(2), 341–351.

Devanand, D. P., & Levy, S. R. (1995). Neuroleptic treatment of agitation and psychosis in dementia. *Journal of Geriatric Psychiatry and Neurology, 8*, S18–S27.

Durand, V. M. & Crimmins, D. B. (1988). Identifying the variables maintaining self-injurious behavior. *Journal of Autism and Developmental Disorders, 18*(1), 99–117.

Erkinjuntti, T., Hokkanen, L., Sulkava, R., & Palos, J. (1988). The Blessed Dementia Scale as a screening test for dementia. *International Journal of Geriatric Psychiatry, 3,* 267–273.

Evans, L. K., & Strumpf, N. E. (1989). Tying down the elderly: A review of the literature on physical restraint. *Journal of the American Geriatrics Society, 37,* 65–74.

Ferrell, B. A., Ferrell, B. R., & Rivera, L. (1995). Pain in cognitively impaired nursing home patients. *Journal of Pain and Symptom Management, 10*(8), 591–598.

Finkel, S. I. (1993). The pharmacological management of agitation in demented nursing home elderly. In M. Bergener, R. H. Belmaker, & M. S. Tropper (Eds.), *Psychopharmacotherapy for the elderly: Principles of research and clinical implications*. New York: Springer.

Finkel, S. I., Lyons, J. S., & Anderson, R. L. (1992). Reliability and validity of the Cohen-Mansfield Agitation Inventory in institutionalized elderly. *International Journal of Geriatric Psychiatry, 7*, 487–490.

Fisher, J. E. (1995, August). *Agitation and adaptation: Functional characteristics of the behavior of dementia patients*. Paper presented at the Annual Convention of the American Psychological Association, New York.

Fisher, J. E., & Carstensen, L. L. (1990). Behavior management of the dementias. *Clinical Psychology Review, 10*, 611–629.

Fisher, J. E., Fink, C. M., Loomis, C. C. (1993). Frequency and management difficulty of behavior problems among dementia patients in long-term care. *Clinical Gerontologist, 13*, 3–12.

Fisher, J. E., Goy, E. R., Swingen, D. N., & Szymanski, J. (1994, November). *Functional characteristics of behavioral disturbances in Alzheimer's disease patients*. Presented at the Annual Meeting for the Association for the Advancement of Behavior Therapy, San Diego, CA.

Fisher, J. E., & Swingen, D. N. (1996). *Contextual features of behavioral disturbance in Alzheimer's Disease*. Presented at the 30th annual meeting of the Association for Advancement of Behavior Therapy, New York.

Fishman, B., Pasternak, S., Wallenstein, S. L., Houde, R. W., Holland, J. C., & Foley, K. M. (1987). The Memorial Pain Assessment Card: A valid instrument for the evaluation of cancer pain. *Cancer, 60,* 1151–1158.

Folstein, M. F., Folstein, S.E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Foster, S. L., & Cone, J. D. (1986). Design and use of direct observation procedures. In A. R. Ciminero, K. S. Calhoun, & H. E. Adams (Eds.), *Handbook of behavioral assessment* (2nd ed.). New York: Wiley.

Gelenberg, A. J. (1994). Buspirone: Seven-year update. *Journal of Clinical Psychiatry, 55*(5), 222–229.

Gerdner, L. A., & Buckwalter, K. C. (1994). A nursing challenge: Assessment and management of agitation in Alzheimer's patients. *Journal of Gerontological Nursing, 20,* 11–20.

Gilley, D. W., & Wilson, R. S. (1991). Nuisance parameters in the assessment of noncognitive manifestations of dementia. *Psychology and Aging, 6*(4), 528–532.

Goldstein, K. (1952). The effect of brain damage on the personality. *Psychiatry, 15*, 245–260.

Greendyke, R. M., Kantor, D. R., & Schuster, D. B. (1986). Propranolol treatment of assaultive patients with organic brain disease: A double-blind crossover placebo controlled study. *Journal of Nervous and Mental Disorders, 174*, 290–294.

Hall, G. R., Gerdner, L., Zwygart-Stauffacher, M., & Buckwalter, K. C. (1995). Principles of nonpharmacological management: Caring for people with Alzheimer's disease using a conceptual model. *Psychiatric Annals, 25*(7), 432–440.

Hall, R. C. W., Beresford, T. P., Stickney, S. K., Nasdahl, C. S., & Coleman, J. H. (1985). Psychiatric reactions produced by respiratory drugs. *Psychosomatics, 26*, 605–608; 615–616.

Hamel, M., Gold, D. P., Andres, D., Reis, M., Dastoor, D., Grauer, H., & Bergman, H. (1990). Predictors and consequences of aggressive behavior by community-based dementia patients. *The Gerontologist, 30*(2), 206–211.

Herrmann, N. & Eryavec, G. (1993). Buspirone in the management of agitation and aggression associated with dementia. *American Journal of Geriatric Psychiatry, 1*(3), 249–253.

Hurley, A. C., Volcier, J., Hanrahan, P. A., Houde, S., and Volcier, L. (1992). Assessment of discomfort in advanced Alzheimer patients. *Research in Nursing and Health, 15*(5), 369–377.

Hussian, R. A. (1988). Modification of behaviors in dementia via stimulus manipulation. *Clinical Gerontologist, 8*, 37–43.

Jeanblanc, W., & Davis, Y. B. (1995). Risperidone for treating dementia associated aggression. *American Journal of Psychiatry, 152*(8), 1239.

Johnston, J. M., & Pennypacker, H. S. (1980). *Strategies and tactics of human behavioral research*. Hillsdale, NJ: Erlbaum.

Kantor, J. R. (1959). *Interbehavioral psychology: A sample of scientific system instruction*. Bloomington, IN: Principia Press.

Kearney, C. A. (1994). Interrater reliability of the Motivational Assessment Scale: Another, closer look. *Journal of the Association for Persons with Severe Handicaps, 19*(2), 139–142.

Kurita, A., Blass, J. P., Nolan, K. A., Black, R. S., & Thaler, H. T. (1993). Relationship between cognitive status and behavioral symptoms in Alzheimer's disease and mixed dementia. *Journal of the American Geriatrics Society, 41,* 732–736.

Larson, E. B., Kukull, W. A., Buchner, D., & Reifler, B. V. (1987). Adverse drug reactions associated with global cognitive impairment in elderly persons. *Annals of Internal Medicine, 107*, 169–173.

Lawton, M. P., Van Haitsma, K., & Klapper, J. (1996). Observed affect in nursing home residents with Alzheimer's disease. *Journal of Gerontology: Psychological Sciences, 51*(1), P3–P14.

Lindsley, O. R. (1964). Geriatric behavioral prosthetics. In R. Kastenbaum (Ed.), *New thoughts on old age*. New York: Springer.

Lundervold, D. A. & Lewin, L. M. (1992). *Behavior analysis and therapy in nursing homes*. Springfield, IL: Charles C. Thomas.

Melzack, R. (1975). The McGill Pain Questionnaire: Major properties and scoring methods. *Pain, 1*, 277–299.

Miller, R. J., Snowdon, J., & Vaughan, R. (1995). The use of the Cohen-Mansfield Agitation Inventory in the assessment of behavioral disorders in nursing homes. *Journal of the American Geriatrics Society, 43*, 546–549.

Moak, G. S. & Fisher, W. H. (1990). Alzheimer's disease and related disorders in state mental hospitals: Data from a nationwide survey. *The Gerontologist, 30*(6), 798–802.

Nelson, E., Wasson, J., Kirk, J., Keller, A., Clark, D., Dietrich, A., Stewart, A., & Zubkoff, M. (1987). Assessment of function in routine clinical practice: Description of the COOP chart method and preliminary findings. *Journal of Chronic Disease, 40*, 55S–63S.

Nyth, A. L., & Gottfries, C. G. (1990). The clinical efficacy of citalopram in the treatment of emotional disturbance in dementia disorders: A Nordic multicentre study. *British Journal of Psychiatry, 157*, 894–901.

Omnibus Reconciliation Act of 1987, Subtitle of nursing home reform, Part C. Washington, DC: US Government Printing Office.

Paul, G. L., & Lentz, R. J. (1975). *Psychosocial treatment of chronic mental patients*. Champaign, IL: Research Press.

Physician's desk reference. (1996). Oradell, NJ: Medical Economics Company.

Ray, W. A., Federspiel, C. F., & Schaffner, W. (1980). Antipsychotic drug use in nursing homes: Epidemiological evidence suggesting misuse. *American Journal of Public Health, 70*, 485–491.

Ray, W. A., Griffin, M. R., Schaffner, W., Baugh, G. H., & Milton, L. J. (1987). Psychotropic drug use and the risk of hip fracture. *New England Journal of Medicine, 316*, 363–369.

Ratey, J.J., & Gutheil, C. M. (1991). The measurement of aggressive behavior: Reflections on the use of the Overt Aggression Scale and the modified Overt Aggression Scale. *Journal of Neuropsychiatry and Clinical Neurosciences, 3*(2), S57-S60.

Roper, J. M., Shapira, J., & Chang, B. L. (1991). Agitation in the demented patient: A framework for management. *Journal of Gerontological Nursing, 17*(3), 17–21.

Rosberger, Z., & MacLean, J. (1983). Behavioral assessment and treatment of "organic" behaviors in an institutionalized geriatric patient. *International Journal of Behavioral Geriatrics, 1*, 33–46.

Rosen, J., Burgio, L., Kollar, M., Cain, M., Allison, M., Fogleman, M., Michael, M., & Zubendo, G. S. (1994). The Pittsburgh Agitation Scale: A user-friendly instrument for rating agitation in dementia patients. *American Journal of Geriatric Psychiatry, 2*(1), 52–59.

Ryden, M. B. (1988). Aggressive behaviors in persons with dementia who live in the community. *The Alzheimer Disease and Associated Disorders: International Journal, 2*(4), 342–355.

Schneider, L. S., Pollock, V. E., & Lyness, S. A. (1990). A metaanalysis of controlled trials of neuroleptic treatment in dementia. *Journal of the American Geriatrics Society, 38*, 553–563.

Schneider, L. S., & Sobin, P. B. (1991). Non-neuroleptic medications in the management of agitation in Alzheimer's disease and other dementia: A selective review. *International Journal of Geriatric Psychiatry, 6*, 691–708.

Sigafoos, J., Kerr, M., & Roberts, D. (1994). Interrater reliability of the Motivation Assessment Scale: Failure to replicate with aggressive behavior. *Research in Developmental Disabilities, 15*(5), 333–342.

Silver, J. M., & Yudofsky, S. C. (1991). The Overt Aggression Scale: Overview and guiding principles. *Journal of Neuropsychiatry and Clinical Neurosciences, 3*(2), S22–S29.

Sloane, P. D., Mathew, L. J., & Scarborough, N. (1991). Physical and pharmacologic restraint of nursing home patients with dementia: Impact of specialized units. *Journal of the American Medical Association, 265*, 1278–1282.

Sorgi, P., Ratey, J., Knoedler, D. W., Markert, R. J., & Reichman, M. (1990). Rating aggression in the clinical setting: A retrospective adaptation of the overt aggression scale: Preliminary results. *Journal of Neuropsychiatry and Clinical Neurosciences, 3*(2), S52–S56.

Spore, D., Mor, V., Hiris, J., Larrat, P., & Hawes, C. (1995). Psychotropic use among older residents of board and care facilities. *Journal of the American Geriatrics Society, 43*, 1403–1409.

Swearer, J. M., Drachman, D. A., O'Donnell, B. F., & Mitchell, A. L. (1988). Troublesome and disruptive behaviors in dementia: Relationships to diagnosis and disease severity. *Journal of the American Geriatrics Society, 36,* 784–790.

Teri, L., Larson, E. B., & Reifler, B. V. (1988). Behavioral disturbance in dementia of the Alzheimer's type. *Journal of the American Geriatrics Society, 36*(1), 1–6.

Teri, L., Borson, S., Kiyak, A., & Yamagishi, M. (1989). Behavioral disturbance, cognitive dysfunction, and functional skill: Prevalence and relationship in Alzheimer's disease. *Journal of the American Geriatrics Society, 37,* 109–116.

Vitaliano, P. P., Young, H. M., Russo, J., Romano, J., & Magana-Amato, A. (1993). Does expressed emotion in spouses predict subsequent problems among care recipients with Alzheimer's disease. *Journal of Gerontology: Psychological Sciences, 48*, P202–P209.

Weiler, P. G., & Goodman, T. A. (1987). The use of propranolol in Alzheimer's's disease patients with disruptive behavior. *Current Therapeutic Research, 42*, 364–374.

Weiler, P. G., Mungas, D., & Bernick, C. (1988). Propranolol for the control of disruptive behavior and senile dementia. *Journal of Geriatric Psychiatry and Neurology, 1*, 226–230.

Werner, P., Cohen-Mansfield, J., Braun, J., & Marx, M. S. (1989). Physical restraints and agitation in nursing home residents. *Journal of the American Geriatrics Society, 37*, 1122–1126.

Williams, B., & Betley, C. (1995). Inappropriate use of nonpsychotropic medications in nursing homes. Journal of the American Geriatrics Society, 43, 513–519.

Wolf, B., Grohmann, R., Beiber, D., Brenner, P. M., & Luther, E. (1989). Benzodiazapine abuse and-dependence in psychiatric patients. *Pharmacopsychiatry, 22*, 54–60.

Yoshikawa, T. T., & Norman, D. C. (1996). Approach to fever and infection in the nursing home. *Journal of the American Geriatrics Society, 44*, 74–82.

Yudofsky, S. C., Silver, J. M., Jackson, W., & Endicott, J. (1986). The Overt Aggression Scale for the objective rating of verbal and physical aggression. *American Journal of Psychiatry, 143,* 35–39.

Zarcone, J. R., Rodgers, T. A., Iwata, B. A., Rourke, D.A., & Dorsey, M. F. (1991). Reliability analysis of the Motivational Assessment Scale: A failure to replicate. *Research in Developmental Disabilities, 12,* 349–360.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.19
Incontinence

JOHN F. SCHNELLE
University of California, Los Angeles, CA, USA

7.19.1 INTRODUCTION 433
7.19.2 PREVALENCE, MORBIDITY, AND CONSEQUENCES 434
7.19.3 TYPES OF URINARY INCONTINENCE 435
  7.19.3.1 Urge Incontinence 435
  7.19.3.2 Stress Incontinence 436
  7.19.3.3 Overflow Incontinence 436
  7.19.3.4 Functional Incontinence 436
7.19.4 BASIC ASSESSMENT 436
7.19.5 TREATMENT OPTIONS 438
  7.19.5.1 Pharmacological Interventions 438
    7.19.5.1.1 Drug interventions for urge incontinence 438
    7.19.5.1.2 Drug interventions for stress incontinence 439
  7.19.5.2 Surgery 439
  7.19.5.3 Behavioral Interventions 439
    7.19.5.3.1 Bladder retraining 440
    7.19.5.3.2 Pelvic muscle rehabilitation 440
    7.19.5.3.3 Toileting assistance protocols 442
7.19.6 IMPLEMENTATION OF AN INCONTINENCE MANAGEMENT PROGRAM IN NURSING HOMES 445
  7.19.6.1 Prompted Voiding Assessment 445
  7.19.6.2 Postvoid Residual Determination and Urinalysis 446
  7.19.6.3 Pad Test for Stress Incontinence 447
  7.19.6.4 Mobility Assessment 447
  7.19.6.5 Who are Candidates for Further Evaluation? 448
  7.19.6.6 Treatment for Mobility Problems that Limit Independent Toileting 448
  7.19.6.7 Staff Management and Continence Maintenance 449
7.19.7 FUTURE DIRECTIONS 450
7.19.8 REFERENCES 452

7.19.1 INTRODUCTION

Urinary incontinence (UI) affects approximately 13 million Americans (National Kidney and Urologic Diseases Board, 1994). It is a common problem among 10–35% of all adults, affecting 30% of persons over age 60 years and more than 50% of nursing home residents (Diokno, Brock, Herzog, & Bromberg, 1990). Despite its high prevalence, urinary incontinence in the community often goes undiagnosed and underreported due to the stigma it holds (United States Agency for Health Care Policy and Research [USAHCPR], 1996). Outside and in the nursing home, incontinence is most often treated by physicians and nurses as a result of federal clinical practice guidelines that focus on medical assessment and treatment. Psychologists and behavioral therapists have played a role in treatment involving biofeedback strategies for use with cognitively intact persons suffering from incontinence. The purpose of the present chapter is to present how and why psychologists should play a broader role in the treatment and management of incontinence, particularly among cognitively and physically impaired frail people. Due to a number of associated conditions, this is particularly true for the large population of people who suffer from incontinence in nursing homes.

The primary role of physicians and nurses in incontinence care is dictated by the fact that best-practice guidelines specify that basic medical assessment of this condition is the first step to good clinical practice (USAHCPR, 1996). In the case of the community dwelling elderly, surgery and pharmacological treatment are often suggested by these assessments, and despite their invasiveness, are acceptable to relatively healthy people who suffer from incontinence. As noted above, historically, psychologists and other behavioral professionals have been responsible for developing and implementing incontinence interventions based on biofeedback strategies with people who are cognitively intact and capable of following the protocols that underlie biofeedback training. This work has been an important contribution to the management of incontinence in community dwelling populations. The present chapter will describe these traditional approaches to incontinence care.

The case will be presented in this chapter that, in addition to traditional roles, psychologists could and should play a broader role in the treatment and management of UI among cognitively and physically impaired persons. This point holds particularly true for the large population of people who suffer from UI in nursing homes, since behavioral deficits, depression, motivation, and a variety of other nonphysiological factors play a much larger role in the UI problems of these institutionalized elderly than for those who live in the community. Psychologists are specially trained to assess and manage these problems. Frail institutionalized elderly are not good candidates for many of the available invasive treatments, even if medical justification for these interventions are present. The most effective treatment of UI problems in the frail elderly has involved noninvasive behavioral interventions that have been directed towards both the patient and the caregiver. The skills of the behavioral scientist both for assessing the need for these interventions and for successfully implementing them is critical.

The present chapter first describes the prevalence and sequela of UI problems in both community dwelling and institutionalized populations. A simple classification scheme to describe the different kinds of incontinence symptoms is presented, along with a brief description of three major classes of treatment: pharmacological, surgical, and behavioral. Evidence describing the efficacy of these treatments is provided, with particular emphasis on the behavioral interventions. The final section of the chapter is devoted to the assessment and management of UI among institutionalized elderly. The potential role behavioral healthcare professionals could play in this assessment and treatment is also highlighted.

7.19.2 PREVALENCE, MORBIDITY, AND CONSEQUENCES

Urinary incontinence is most prevalent among the elderly in both community and institutional settings (National Kidney and Urologic Diseases Advisory Board, 1994). It affects 30% of persons over age 60 years who live in the community, about 35% of older patients in acute care hospitals, and over 50% of those living in institutions (Help for Incontinent People, 1994). Urinary incontinence has significant adverse physical, psychological, and financial effects. It is a leading cause of infection among frail older persons, is considered an embarrassing stigma by many of its sufferers, and has an estimated annual cost of over US$15 billion (National Kidney and Urologic Disease Advisory Board, 1994). Coupled with its role as a primary predictor of institutionalization, these negative outcomes clearly justify more aggressive efforts to identify, evaluate, and treat UI. These efforts are further justified by growing evidence that appropriate management can reduce the morbidity and cost of incontinence, particularly in institutionalized populations (Ouslander, Palmer, Rovner, & German, 1993).

Reported prevalence rates of UI vary considerably, depending on the population studied, how UI is defined, and how the prevalence information is obtained (Diokno et al., 1990). Among the population aged between 15 and 64 years the prevalence of UI in men is in the range 1.5–5% and in women is 10–30% (Burgio, Matthews, & Engel, 1991; Harrison & Memel, 1994).

Prevalence of UI ranges from 15% to 35% for noninstitutionalized adults above the age of 60 years, with women having twice the prevalence of men. Of these incontinent older adults, 25–30% have frequent (i.e., daily or weekly) incontinence episodes (Burgio et al., 1991;

Diokno, Brock, Brown, & Herzog, 1986). Approximately 53% of homebound elderly are incontinent (Noelker, 1987); this being similar to the prevalence rates reported for institutionalized elderly.

UI among nursing home residents is more severe and more commonly associated with fecal incontinence than in the community populations. Incontinent nursing home residents generally have multiple incontinence episodes throughout the day and night, and approximately 50% are also incontinent of stool (Ouslander & Fowler, 1985; Ouslander, Kane, & Abrass, 1982). The reported prevalence of UI in institutionalized populations ranges from 40% to 70%, and even higher, varying considerably between individual facilities depending on the case mix of functional impairment found within the resident population (Ouslander & Schnelle, 1995).

UI is associated with substantial morbidity and costs in the nursing home. It can predispose residents to skin irritation, exacerbating pressure ulcers and making them difficult to heal (USAHCPR, 1996); and can result in symptomatic urinary tract infections (Ouslander, Greengold, & Chen, 1987; Warren et al., 1987). It is also associated with falls among residents with nocturia and impaired balance or gait (Stewart, Moore, May, Marks, & Hale, 1992). The adverse psychological effects of UI among nursing home residents have been difficult to document due to issues of cognitive status and lack of willingness to report. However, incontinent residents who do not have severe dementia have reported being often embarrassed and frustrated by their condition (Ouslander, Morishita, et al., 1987).

Nursing home staff generally consider UI to be one of the most onerous and difficult conditions for which they provide care, and they perceive that a disproportionate amount of time is spent on the care of incontinent residents. The economic costs of UI in the nursing home have been estimated to be close to US$5 billion annually, including the cost of staff time, laundry, and supplies (Hu, 1994).

Thus, UI is a severe problem for large numbers of older people. Yet, this condition becomes even more devastating when it is considered that UI is often associated with a variety of more severe conditions. Among community dwelling older adults, UI is listed as a common geriatric syndrome associated with (i) a general decline in upper and lower body strength, (ii) sensory and emotional decline, and (iii) a high risk of institutionalization (Tinetti, Inouye, Gill, & Doucette, 1995).

Among nursing home residents, UI is associated with a high rate of hospitalization (Murtaugh & Freiman, 1995). Incontinent nursing home residents are susceptible to a high rate of infection requiring costly medical treatment both in the hospital and within the nursing home. Furthermore, the act of hospitalizing a nursing home resident is expensive from both quality of life and economic viewpoints. When older persons are hospitalized, they are at risk for experiencing a deluge of dependencies (Creditor, 1993).

Elderly nursing home residents often deteriorate in their functional status during hospital stays and develop problems unrelated to the cause of the admission (e.g., pressure sores). It seems clear that reducing hospitalization rates in nursing home residents will both improve the quality of life of residents and reduce the overall costs of healthcare. In view of these data, it has been argued that prevention programs targeting incontinent nursing home residents should be developed (Murtaugh & Freiman, 1995). The intent would be to reduce serious side effects connected with UI (e.g., urinary tract infections, skin infections), or prevent the condition from occurring. One example of such a prevention program is described in Section 7.19.6.6

To launch preventative programs and productive interventions effectively, it is essential first to distinguish among specific types of UI and to design efforts that are appropriately based on etiology and symptoms.

7.19.3 TYPES OF URINARY INCONTINENCE

7.19.3.1 Urge Incontinence

Urge incontinence is characterized by the involuntary loss of urine associated with a strong (or urgent) desire to void. Normally, as the bladder (detrusor) fills with urine, stretch receptors send signals to the spinal cord indicating the need for contraction. Under normal circumstances these contractions are inhibited through the central nervous system control. The person with urge incontinence has a reduced ability either to sense the urgency to empty the bladder or to inhibit contractions from the point that such signals are recognized.

This condition is usually associated with involuntary detrusor contractions referred to as detrusor instability (DI). This response is urodynamically confirmed by filling a person's bladder slowly with fluid via a catheter and measuring bladder contractions. In the case of urge incontinence, bladder contractions occur before the patient can report that they are unable to hold further fluid. DI is at times associated with neurologic disorders, but can also occur in individuals who appear to be neurologically intact.

In cases where neurological lesions are established, the DI is distinguished as detrusor hyperreflexia (DH). The most common neurological disorder associated with DH is stroke. A second neurodynamic diagnosis associated with the symptom of urge incontinence in frail elderly people is detrusor hyperactivity with impaired bladder contractility (DHIC) (Resnick & Yalla, 1985). DHIC results in involuntary bladder contractions. However, symptomatically, patients must strain actually to empty their bladders. Patients with DHIC generally have symptoms of urge incontinence and elevated postvoiding residual volumes (i.e., the amount of urine left in the bladder after a person voluntarily attempts to empty his or her bladder). Clinically, the amount left in the bladder can be determined either by ultrasound or by catheterizing the person after they void and draining what remains in the bladder.

7.19.3.2 Stress Incontinence

Stress incontinence is characterized by the involuntary loss of urine during physical movements that increase intra-abdominal pressure, such as coughing, sneezing, laughing, or other physical activities. Stress incontinence is established if the urine loss is coincident with an increase in intra-abdominal pressure, but occurs in the absence of a detrusor contraction or an overdistended bladder (USAHCPR, 1996). The most common cause of stress incontinence occurs in women as intra-abdominal pressure results in a significant displacement of the urethra and bladder during exertion.

Stress incontinence may also be caused by a deficiency in the sphincter muscle that closes off the bladder outlet. Normally, the sphincter is relaxed when voluntary bladder emptying begins, but closes off the bladder outlet during other times. In women, this incontinence is commonly a result of multiple surgical procedures, low estrogen levels due to age or other causes, or both. In these cases, the urethra sphincter is unable to generate enough resistance to retain urine in the bladder when bladder pressure is raised. Patients with stress incontinence can leak continuously or with minimal exertion.

It is not unusual for people to have a combination of urge and stress incontinence. When both symptoms are present, the incontinence is called mixed. Mixed incontinence is particularly common in older women.

7.19.3.3 Overflow Incontinence

Overflow incontinence is the involuntary loss of urine associated with overdistension of the bladder due to a high volume of urine. Overflow incontinence may manifest itself in a variety of symptoms similar to those of stress incontinence. It may be caused by an underactive bladder (unable to contract efficiently) or a bladder-outlet obstruction leading to overdistension and overflow. The bladder may be underactive due to drugs, neurologic conditions, spinal cord injury, or as a side effect of diabetes. In men, overflow incontinence is often associated with obstruction caused by the prostate; outlet obstructions are rare in women. Overflow incontinence is best diagnosed by measuring a patient's postvoid residual volume, as described in Section 7.19.3.1.

7.19.3.4 Functional Incontinence

Functional incontinence is characterized by involuntary urine loss due to factors outside the lower urinary tract. For example, if a patient has no stress or urge symptoms or a high postvoid residual, but is incontinent because he or she cannot reach the bathroom in a timely manner, it may be termed functional incontinence. Functional incontinence is typically diagnosed in people with physical or cognitive impairment. Caution is recommended in making this diagnosis, in that it should only be made after other causes have been excluded. Because mobility and cognitive impairment may occur in patients suffering from other types of UI, it is essential first to rule out those types that could respond to specific therapies.

The most typical treatment for functional incontinence involves improving the patient's functional status, modifying transient causes for the incontinence (e.g., fecal impaction), and reducing environmental barriers to toileting or providing sufficient toileting assistance to avoid incontinence episodes. These latter treatments are suggested for consideration even if a lower urinary tract abnormality is present, since incontinence may be exaggerated by immobility and cognitive factors. The assessment and treatment of incontinence is now discussed.

7.19.4 BASIC ASSESSMENT

Best-practice guidelines recommend that all incontinent residents should undergo a basic evaluation that includes a history, physical

Table 1 Potentially reversible conditions that may contribute to urinary incontinence in nursing home residents, and the management of these conditions.

Conditions	*Management*
Impaired ability or willingness to reach a toilet	
Delirium	Diagnosis and treatment of underlying cause(s)
Illness, injury or restraint interfering with mobility	Regular toileting assistance. Use of toilet substitutes Environmental alterations.
Depression	Appropriate pharmacological and/or nonpharmacological treatment
Irritation or inflammation in or around the lower urinary tract	
Urinary tract infection (symptomatic with frequency, urgency, sudden onset, or worsening of UI, unexplained fever or functional decline)	
Antimicrobial therapy[a]	
Atrophic vaginitis/urethritis	Oral or topical estrogen[b]
Stool impaction	Disimpaction Use of stool softeners and laxatives as appropriate Adequate mobility and fluid intake.
Increased urine production	
Metabolic (hyperglycemia, hypercalcemia	Improved control of diabetes Therapy for hypercalcemia if indicated by underlying cause
Excess fluid intake	Reduction in diuretic fluid intake (e.g., caffeinated beverages)
Volume overload	
Venous insufficiency with edema	Support stockings Leg elevation Sodium restriction Diuretic therapy
Congestive heart failure	Medical treatment
Drug side-effects	
Rapid-acting diuretics causing frequency and urgency	Discontinuation of offending medication as possible Dosage reduction or modification of scheduling
Rapid-acting diuretics causing frequency and urgency	
Anticholinergics, narcotics, calcium channel blockers, α adrenergic agonists (in men) causing urinary retention	
α Adrenergic antagonists causing urethral relaxation and stress incontinence	
Psychotropic drugs with sedative or extrapyramidal effects causing sedation and immobility that interfere with toileting	

[a] For nursing home residents with bacteriuria (with or without pyuria) who have chronic, stable symptoms of UI and no other symptoms of urinary tract infection, antimicrobial therapy is not indicated (see text). [b] See text for specific recommendations.

examination, measurement of postvoiding residual volume, and urine analysis. This basic evaluation has three objectives:

(i) to identify potentially reversible factors for incontinence (e.g., stool impaction, drug side effects, and other factors, as listed in Table 1);

(ii) to identify serious underlying conditions that may require further urologic, gynecologic, or urodynamic evaluation (see Table 2 for listed conditions); and

(iii) to determine the types of incontinence and the appropriate plan for their management.

Depending on the outcomes of this basic evaluation, three general classes of treatment and several combinations of treatments may be justified. The three primary classes of treatment are now discussed.

Table 2 Conditions that may require further urologic, gynecologic, or urodynamic evaluation.

Conditions	*Comments*
History	
Incontinence with recurrent symptomatic urinary tract infections	Must first exclude structural abnormality or pathologic condition in the urinary tract predisposing the patient to symptomatic infection
Physical examination	
Marked pelvic prolapse protruding through vaginal introitus	Consider surgical repair or pessary management to prevent discomfort and tissue erosion
Possible prostate cancer	Surgical intervention to cure cancer would not be appropriate for most nursing home residents; however, cancer diagnosis may be important in managing the incontinence and other complications
Urinalysis	
Hematuria (sterile)	Consider urologic evaluation to identify pathologic conditions of the urinary tract
Determination of postvoid residual volume	
Residual volumes > 200 ml	No precise cutoff point can be recommended; however, residual volumes > 200 ml are abnormal and should prompt consideration of further evaluation of the urinary tract to identify complications (e.g., hydronephrosis or renal function impairment) and determine the cause of the retention (e.g., obstruction, impaired bladder contractility)

7.19.5 TREATMENT OPTIONS

7.19.5.1 Pharmacological Interventions

The most common pharmacological interventions involve medications given for detrusor instability with urge symptoms and for stress incontinence. Evidence for the efficacy of these medications is most complete for anticholinergic drugs, which are designed to limit the contractility of the normal bladder and which probably also reduce instable contractions in the bladder.

7.19.5.1.1 Drug interventions for urge incontinence

The pharmacological agent which has been most often tested in trials to reduce urge incontinence symptoms is oxybutynin. Oxybutynin proved superior to placebos in six studies of middle-aged patients, reducing UI frequency by 15–58% more than did the placebo (Holmes, Montz, & Stanton, 1989; Moore, Hay, Imrie, Watson, & Goldstein, 1990; Riva & Casolati, 1984; Tapp, Cardozo, Versi, & Cooper, 1990; Thuroff et al., 1991; Zeegers, Kiesswetter, Kramer, & Jonas, 1987). In two of these studies, 43–67% of the subjects became subjectively continent with oxybutynin, a significantly larger proportion than found in those receiving placebo.

To test its efficacy with a nursing home population, in a randomized clinical trial of nursing home residents, oxybutynin was administered only to those residents who failed to respond to a noninvasive behavioral intervention (prompted voiding) (Ouslander et al., 1995a). Oxybutynin was given as a supplement to prompted voiding for those residents who failed to reduce their UI level with the prompted voiding intervention to less than one episode per day. Oxybutynin showed very modest effects in this target group of residents, and did not reduce UI to levels which were regarded as clinically significant. Thus, given the cost of the medication and its potential negative side effects (dry mouth, constipation, urine retention), it is not recommended for the general treatment of UI among nursing home residents.

An expert consensus panel convened by the USAHCPR summarized the research findings for the use of pharmacological agents to treat detrusor instability as follows:

> Pharmacotherapy, at least in short-term trials, appears to benefit some patients with UI due to detrusor instability. However, involuntary bladder contractions are usually not abolished, the warning time between the appreciation of the need to

> void and the onset of bladder contraction is usually not affected, the degree of improvement is modest and cure is uncommon. In addition, many of these medications are costly and most have bladder side effects. These drugs should be used only in conjunction with a voiding schedule or behavioral intervention and only if other factors contributing to UI have been addressed. (USAHCPR 1996, p. 47)

7.19.5.1.2 Drug interventions for stress incontinence

Medications have also been used for stress incontinence, the purpose being to enhance muscle control in the bladder neck, bladder base, and proximal urethra. It is anticipated in these cases that the drug will stimulate muscle contraction in these areas and increase bladder outlet resistance. Increasing resistance should result in less urine loss when intra-abdominal pressure is raised. The data from controlled trials of medications that stimulate such muscle contraction are less clear than for drugs given for detrusor instability. The expert USAHCPR panel concluded that pharmacological therapy treatment for stress incontinence appears to result in few cures, but may cause subjective improvement in 20–60% of patients over placebo responses.

Estrogen replacement therapy is also considered a potential treatment for stress incontinence. Presumably estrogen replacement therapy will increase the vascularity, tone, and responsiveness of the urethra muscle. This, in turn, should increase bladder resistance and decrease stress incontinence. The exact mechanism of how estrogen may work, however, is unknown.

Fantl, Cardozo, & McClish, (1994), for The Hormones and Urogenital Therapy Committee, provided a meta-analysis of estrogen effects from six randomized controlled trials and 17 uncontrolled clinical series. Positive results were found for all subjects in the trials, representing a significant number of patients who presented with stress incontinence symptoms.

7.19.5.2 Surgery

Surgical intervention for UI is a consideration for a small but important subgroup of nursing home residents and a much larger group of community dwelling elderly. Since UI is not a life-threatening problem, elective surgery should be considered only if the patient reports strong quality-of-life concerns. It is also recommended that in most, if not all, cases surgery be considered only after less invasive pharmacological or behavioral treatments have proven ineffective.

Two major types of surgery are used to treat UI. First, women with stress incontinence associated with significant pelvic prolapse and urethral hypermobility may benefit from bladder-neck suspension and repair of the pelvic prolapse. Although studies have not been done in the nursing home population, the short-term (1–5 years) success of this type of surgery for properly selected older women is approximately 75%, success being defined as a significant reduction or elimination of wetness USAHCPR, 1996). Surgical placement of artificial urinary sphincters has been used for younger patients with stress incontinence associated with intrinsic urethral deficiency. This requires a relatively major surgical procedure and is probably not appropriate for nursing home residents or very frail older people. Periurethral injections of collagen have recently been approved for patients with intrinsic urethral deficiency. It is unknown what the mechanism of the effects are, but periurethral injection has reportedly been effective in younger patients.

The other type of surgical intervention for UI is the removal of anatomical obstruction, most commonly an enlarged prostate in men or a urethral stricture. Newer, less invasive techniques are now available for the surgical relief of obstruction in men and may be especially appropriate for consideration in frail, older nursing home residents. Detailed discussion of the evaluation and surgical treatment of lower urinary tract obstruction is beyond the scope of the present chapter. However, it is important to note that candidates for surgery who are suspected of having obstruction should be referred for complex urodynamic evaluation.

7.19.5.3 Behavioral Interventions

Behavior-based treatment strategies are a relatively noninvasive form of treatment for UI and can be effective in many cases. Behavioral therapies can be divided into (i) caregiver-dependent techniques for patients with cognitive and motor deficits, and (ii) those requiring active rehabilitation and education of the patient. These distinctions are arbitrary and require a careful analysis of how the patient's abilities match the performance requirements of a specific behavioral protocol. For example, most biofeedback interventions require that a patient follow multistep instructions, which may be beyond the ability of a cognitively impaired frail older adult. Therefore, in order to characterize the full array of available behavioral techniques for the large range of cognitive and functional abilities represented among incontinent people, this section of the present chapter describes techniques ranging from those that require the most active involvement on the part of the patient to those that require the least.

7.19.5.3.1 Bladder retraining

Bladder retraining is strongly recommended for the management of urge and mixed incontinence. The goal of this training is to build bladder capacity or to increase intervals between voiding. This, in turn, will reduce the probability of involuntary contractions. Bladder training has been used primarily to manage incontinence caused by DI. This form of training has few side effects, but is difficult to implement in cognitively impaired persons.

Bladder retraining (also termed "bladder training") has many variations, but generally consists of three primary components: (i) education, (ii) scheduled voiding with systematic delay of voiding, and (iii) positive reinforcement.

The education component usually involves a combination of written, visual, and verbal instructions that describe the physiology and pathophysiology of the lower urinary tract. This provides the background understanding considered useful in promoting adherence to the behavioral protocol.

Scheduled voiding and delay of voiding involves retraining the patient in resisting or inhibiting the sensation of urgency, to postpone voiding, and to urinate according to a timetable rather than according to urge (McCormick & Burgio, 1984). Behavioral approaches are often designed to distend the bladder by adjusting fluid loads and/or delaying voiding. The goal is to progressively increase the volumes retained in the bladder as well as the intervals between voids (Keating, Schulte, & Miller, 1988). If the patient is unable to meet the delay goal, the schedule is adjusted and the interval reset. Generally, the interval goal is initially set between two and three hours, with consideration given to the patient's baseline. This technique is not usually applied during sleeping hours. Once established, the goal interval is progressively increased and new goals accomplished through the use of distraction or relaxation techniques. The patient is thus "trained" to delay voiding when the urge to void occurs.

Several reports demonstrate that bladder training is effective in reducing UI (Fantl et al., 1991; Ferrie, Smith, Logan, Lyle, & Paterson, 1984; Pengelly & Booth, 1980). Fantl et al. (1991) conducted a controlled, randomized study of 131 women with sphincteric incompetence and unstable detrusor function. Subjects participated in a bladder-training program that included patient education, behavioral strategies to decrease urge, and scheduled voiding. Of the 60 women in the treatment group, 12% became dry and 75% had at least a 50% reduction in the number of UI episodes. Control subjects did not have significant changes in incontinent episodes. The magnitude of the effect was somewhat larger in cases of DI and the effect of bladder training was maintained after six months.

7.19.5.3.2 Pelvic muscle rehabilitation

Pelvic muscle rehabilitation is designed to strengthen the voluntary muscles that contribute to the closing force of the urethra and to support the pelvic structures. Underlying these technologies is a strategy directed toward stress symptoms; but there is some evidence that it can also be successful with DI and urge incontinence as well. When pelvic muscle rehabilitation exercises are done alone they often are called Kegel exercises. However, these exercises are also frequently supplemented by biofeedback training designed to increase awareness of pelvic muscle functioning.

The first step in pelvic muscle exercises (PMEs) is to help the patient develop a better awareness of pelvic muscle function and control. PMEs are performed by voluntarily contracting the perivaginal muscles and anal sphincter as if to control urination or defecation (Rose, Baigis-Smith, Smith, & Newman, 1990). The goal is to accomplish this with minimal contraction of abdominal, buttock, or inner thigh muscles. Patients are trained to sustain a contraction for at least 10 seconds, followed by an equal period of relaxation. It is a goal to progressively increase the contraction time. PMEs should be performed about 30–80 times a day for at least eight weeks, and may need to be continued indefinitely. Patients should be taught to contract the pelvic muscles both before and during an incident when leakage may occur.

The most effective use of PME to date has been reported in an uncontrolled study by Benvenuti et al. (1987). Patients were provided with 10 hours of training, including 20 sessions for practice of PMEs and other behavioral strategies, and an intensive home program. Subjects were 20 women aged 35–65 years (mean age 50.8 years). This protocol resulted in a 95% reduction in incontinence episodes. PME also has proven effective in the treatment of older adults in a study by Flynn, Cell, and Luisi (1994) that involved treatment for transient UI in 37 older adults receiving home nursing care (mean age 76 years).

The intensity of the PME program significantly affects physiological and functional outcomes. Bo, Hagen, Kvarstein, Jorgensen, and Larsen (1990) found significantly greater reductions in UI among a group of patients receiving ongoing training of increasing intensity over a period of six months as compared to a group that received only a single session of instruction and a home exercise program. The group receiving the ongoing instruction reported a 60.1% cure/improvement rate compared with 17.3% among the home exercise group. This study also found significant differences between these groups in physiological measurements of improved pelvic strength. This finding indicates a possible systematic relationship between symptom reduction and objective physical changes. Benvenuti et al. (1987), Elia and Bergman (1993), and Mouritsen, Frimodt, and Moller (1991) also report findings that PME has the potential to change muscle physiology.

In all cases of PME training, patients must show the capability of memorizing multistep instructions and be motivated to practice exercises when not with the therapist. These performance requirements limit the use of pelvic floor muscle rehabilitation for cognitively impaired individuals.

(i) Pelvic muscle rehabilitation: biofeedback

Biofeedback therapy generally uses electronic or mechanical instruments to sensitize patients to their physiological activity. This sensitization is generally accomplished through the use of auditory, visual, or other feedback. The goal is to train patients to identify and modify physiologic responses that mediate bladder control. Studies reporting various applications of biofeedback combined with behavioral treatment report a range of 54–87% improvement in UI across various patient groups (USAHCPR, 1996).

Biofeedback protocols differ in their use of reinforcement or feedback, as well as in the measurement system used to establish this feedback. The biofeedback protocol that has achieved the largest and most consistent reduction in UI episodes (75.9–82%) simultaneously reinforces pelvic muscle contraction, inhibition of abdominal contraction, and inhibition of detrusor contraction (Burgio, Robinson, & Engel, 1986; Burgio, Stutzman, & Engel, 1989; Burgio, Whitehead, & Engel, 1985; Burton, Pearce, Burgio, Engel, & Whitehead, 1988; McDowell, Burgio, Dombrowski, Locher, & Rodriguez, 1992). The presumed benefit of this approach, defined by a multimeasurement procedure, is that it reinforces pelvic floor contraction with moment-to-moment feedback, thus providing the patient with highly specific internal references to upgrade muscle contraction. Combining bladder and sphincter biofeedback also allows teaching pelvic muscle contraction in response to increasing bladder volume and observed detrusor activity.

In comparing biofeedback to PME, the approaches have been found to have comparable results. In a randomized controlled study, Burns et al. (1993) studied 135 women (age range 55–75 years) with primary stress incontinence. One treatment group received PME, and another group received single-measurement biofeedback for 20 minutes per week. The two treatment groups demonstrated a 54–61% reduction in incontinent episodes, compared with a 6% reduction of UI in a control group. No significant difference between the two treatment groups was found. This study demonstrated that both PME and biofeedback have the potential to reduce UI, and provided evidence that symptom severity has a role in response.

(ii) Pelvic muscle rehabilitation: weight training

Specially designed vaginal weights to strengthen the pelvic muscles have also been used to supplement PME. In this therapy, the patient receives a set of vaginal weights of identical shape and volume but of increasing mass. As part of a progressive resistance program, women insert a weight intravaginally and are instructed to attempt to retain it by contracting the pelvic muscle for up to 15 minutes at least twice daily. It is to be worn while the patient is ambulatory. The mechanism of action that the sustained contraction requires to retain the weight increases the strength of the pelvic muscles in a way that provides heightened feedback as to the desired pelvic muscle contraction. Available literature on this technique includes observations made in women with stress incontinence, but no controlled trials have been conducted. Initial observation from four anecdotal studies indicate a subjective cure rate or improved status in 68–80% of women after 4–6 weeks of treatment (Bridges, Denning, Olah, & Aorrar, 1988). The preliminary data from vaginal weight training, as well as the simplicity of the procedure, make it an attractive treatment option. However, the procedure is not yet well evaluated and further research is needed.

7.19.5.3.3 Toileting assistance protocols

Toileting assistance interventions are necessary for many physically impaired people who cannot independently toilet and for cognitively impaired people who cannot remember the behaviors necessary to maintaining continence (e.g., finding the bathroom). Most of the estimated one million nursing home residents who suffer from UI have both physical and cognitive problems that prevent independent toileting. Therefore, it is necessary that their caregivers provide some level of assistance to compensate for the resident's deficits. Unfortunately, many of these cognitive and physical problems are not reversible in the nursing home population, and it is necessary that caregivers continue to provide toileting assistance indefinitely. Thus, the goal for most cognitively impaired patients is to increase dryness under conditions where staff or other caregivers implement a routine protocol. This is in contrast to less impaired populations, in which the goal is to rehabilitate patients to the point where they are capable of maintaining dryness independently. Several toileting assistance protocols have been described for application in nursing homes.

(i) Routine or scheduled toileting

Routine or scheduled toileting is conceptually the most simple toileting assistance procedure, and involves the caregiver placing the patient on the toilet every 2–4 hours. No systematic effort is made to have the patient delay voiding or to have him or her initiate or ask for toileting assistance, and there are relatively few mobility requirements. There are no controlled trials of scheduled toileting procedures, even though there have been some reports that it has positive effects (Burgio et al., 1989).

(ii) Habit training

Habit training is a modification of the scheduled toileting protocol in which caregivers first attempt to identify a resident's voiding pattern. A toileting schedule is then set based on this voiding pattern. Presuming predictable voiding habits can be identified, placing a patient on a toilet at the point that they are most likely to void should prevent UI episodes. One study has reported positive results with this procedure (Colling, Ouslander, Hadley, Eisch, & Campbell, 1992).

(iii) Prompted voiding

The toileting assistance protocol that has been most thoroughly evaluated is prompted voiding. The major objective for prompted voiding is to establish an awareness on the part of the patients as to their continence status (wet or dry) and to increase the probability that they will ask for toileting assistance. Residents are not placed on the toilet unless they respond to a prompt for assistance or self-initiate a request for toileting assistance. The major steps in prompted voiding are listed in Table 3. It should be noted that these steps are deceptively simple. Many nursing home residents show severe communication and attention problems, and therefore a high level of clinical skill is required to induce them to respond to prompts. A behavioral training program in which caregivers are taught to orient residents to questions with a series of verbal and manual prompts greatly increases the ability of caregivers to teach cognitively impaired residents to respond affirmatively to prompts.

Prompted voiding has been documented to be highly effective in nursing home residents when skilled staff implement the protocol, even with residents who have different types of urinary tract dysfunction (Schnelle, Newman, Abbey, et al., 1990). Data collected during three major clinical trials indicate that 25–45% of incontinent residents will respond to prompted voiding, with a reduction in their UI frequency from three to four episodes per day to one per day (Schnelle, 1991; Schnelle & Ouslander, 1996). Approximately 34–40% of these residents do not respond, either because of their refusal to respond to prompts, or their inability to initiate voiding once they are assisted to the toilet. Severe depression and cognitive impairment often characterize those residents who refuse to respond to prompts. Problems in the lower urinary tract characterize residents who cannot initiate toileting once on the toilet. Procedures to identify residents who are capable of responding to prompts, initiating bladder emptying when on the toilet, and maintaining dryness have been described, and responsive residents can be identified in a three-day run-in trial (Ouslander et al., 1995b). This run-in assessment procedure is described in more detail in Section 7.19.6.1.

Unfortunately, there is strong evidence that prompted voiding procedures are not consistently implemented by indigenous nursing home staff, even with residents who are highly responsive (Schnelle, Newman, & Fogarty, 1990a). The inconsistent application of the prompted voiding procedure invariably results, even with the most responsive residents, in an increase in UI frequency, since few nursing home residents are capable of independent toileting. The primary dynamic that causes staff to gradually withdraw toileting assistance and decrease implementation of prompted voiding relates to time

Table 3 Prompted voiding protocol.

1. Contact the residents every two hours from 7 a.m. to 7 p.m.
2. Focus residents' attention on voiding by asking them whether they are wet or dry
3. Check resident for wetness, record on bladder record, and give feedback on whether the resident's self-report was correct or incorrect
4. Whether wet or dry, ask the residents if they would like to use the toilet (or urinal).
   (a) If they say yes:
      (i) Assist them
      (ii) Record the results on the bladder record
      (iii) Give them positive reinforcement by spending an extra minute or two talking with them
   (b) If they say no:
      (i) In the event they have not attempted to void in the last four hours, repeat the request to use the toilet once or twice before leaving, and follow step 4(a) if an affirmative response is received
      (ii) Inform them that you will be back in two hours and request they try to delay voiding until then

efficiency (Schnelle, 1991; Schnelle, Sowell, Hu, & Traughber, 1988a). It has been documented many times that changing residents when they are wet is less time consuming for staff than is prompting them with sufficient frequency to maintain dryness (Schnelle, Sowell, Hu, & Traughber, 1988b). Furthermore, a resident is typically changed one to two times per eight-hour shift, despite best practice guidelines, and a successful prompted voiding program would require that a resident be prompted four times and provided toileting assistance approximately three of those times. It has been calculated that the time staff spend in managing incontinent patients is approximately seven minutes per eight-hour shift, whereas to implement the prompted voiding program so as to maintain dryness would take approximately 26 minutes per eight-hour shift (Schnelle et al., 1995).

Due to the time dynamic and other factors, behavioral management systems that reinforce nursing home staff to maintain prompted voiding protocols have proven necessary if the protocols are to be maintained. The most successful of these management systems is based on applied behavioral analysis strategies and work sampling systems that are used in the area of total quality management. This management system and issues relevant to implementing the system will be described later in this chapter (Schnelle, Newman, et al., 1993).

The successful implementation of the prompted voiding protocol in nursing homes within the framework of the best-practice guidelines, describing how UI should be assessed, generates major opportunities for the involvement of a behaviorally trained professional. An algorithm that describes how UI can be successfully assessed and treated within the nursing home setting is provided in Figure 1. This algorithm is described below, with particular emphasis on those points in the algorithm that require the expertise of behaviorally trained professionals.

The first section of the algorithm covers the basic UI evaluation that is recommended by best-practice guidelines for all patients. This basic assessment was summarized earlier in this chapter, and it is fully described in the USAHCPR incontinence guidelines. A trained nurse can complete this part of the algorithm, and we will not discuss these basic assessment procedures again at this point.

However, we should note that experience indicates that many nursing home residents will show transient causes for their incontinence at this stage, but also that the management of these conditions will not result in the elimination of incontinence. Immobility and dementia are such powerful causative factors for UI in the nursing home population that a program to compensate for these problems is almost always necessary if UI is to be improved. The term "compensate" is used in the above sentence as opposed to "rehabilitate" or "reverse" since there is strong evidence that the cognitive and physical impairments that exacerbate UI are not reversible in most cases, given the current treatment technologies and resource availability.

The most important section of the algorithm for the treatment and management of UI is the quality assurance monitoring system. This system requires that nursing supervisors and staff be trained in basic principles of behavior management and quality control, and is

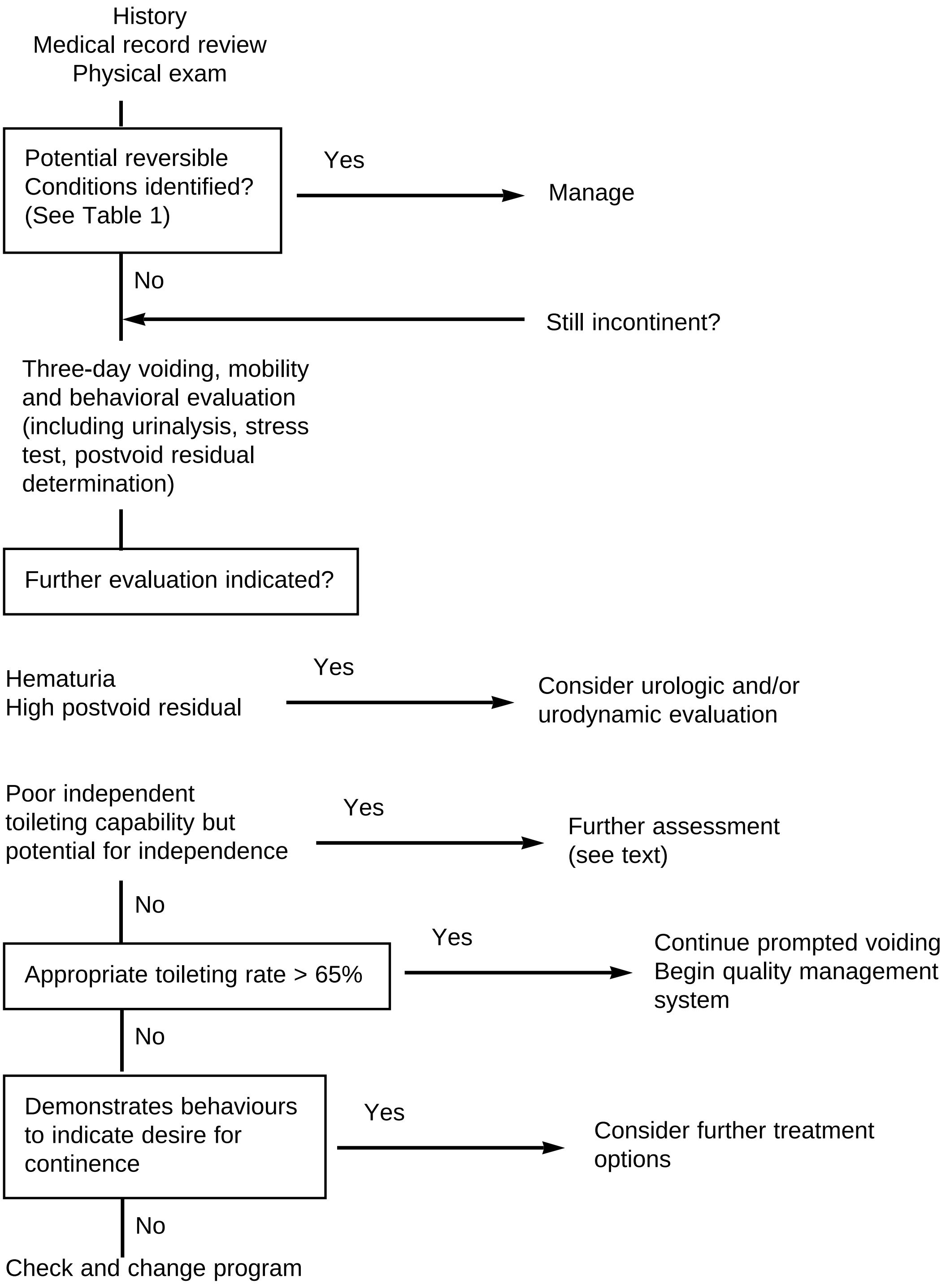
History
Medical record review
Physical exam
Potential reversible Conditions identified? (See Table 1)
Yes
Manage
No
Still incontinent?
Three-day voiding, mobility and behavioral evaluation (including urinalysis, stress test, postvoid residual determination)
Further evaluation indicated?
Hematuria
High postvoid residual
Yes
Consider urologic and/or urodynamic evaluation
Poor independent toileting capability but potential for independence
Yes
Further assessment (see text)
No
Appropriate toileting rate > 65%
Yes
Continue prompted voiding
Begin quality management system
No
Demonstrates behaviours to indicate desire for continence
Yes
Consider further treatment options
No
Check and change program

Figure 1 Assessment and treatment steps for urinary incontinence: basic evaluation.

described later in the present chapter. First, we discuss the critical assessment stage associated with prompted voiding.

7.19.6 IMPLEMENTATION OF AN INCONTINENCE MANAGEMENT PROGRAM IN NURSING HOMES

7.19.6.1 Prompted Voiding Assessment

The most important section of the algorithm for the assessment of UI in nursing home patients is a three-day voiding, mobility, and behavioral evaluation. This three-day evaluation can best be done by nurses in combination with the behavioral health professional who is expert in the assessment of behavioral competencies and motivation.

Critical information is collected during this stage regarding the following issues.

(i) Will a resident be responsive to a prompted voiding program?

(ii) Is the resident motivated to be continent, but not responsive to the prompted voiding protocol because of problems with the lower urinary tract?

(iii) Can the UI symptoms and problems with the lower urinary tract be identified and treated in the resident who is motivated to be continent, but who cannot maintain dryness?

(iv) Does a resident have mobility or cognitive problems which preclude safe independent toileting and can these problems be treated?

(v) What form of UI measurement is best for the resident who does not appear to be either a candidate for further treatment or who appears to not want further treatment?

The major activity during this three-day trial stage involves exposing the residents to a prompted voiding protocol every two hours between 8 a.m. and 4 p.m. (four times per day for three days = 12 prompts). Data are recorded at each prompt relevant to a resident's incontinent status (wet or dry); the resident's response to prompt (attempted to toilet successfully, attempted to toilet unsuccessfully, did not attempt to toilet), and the resident's behavioral response to prompt (shows resistance, agitation, or cooperation).

Prior to this three-day trial, the resident should be asked the following questions.

(i) Does it bother you to wet in your diaper?

(ii) Do staff help you to the toilet as much as you would like?

(iii) Do you want to be helped to the bathroom more often?

If the resident answers no to question (iii):

(iv) Do you want to be helped to the toilet less often?

(v) Do you want to be changed more often? If the answer is no, ask question (vi);

(vi) Do you want to be changed less often? Immediately after the three-day trial, the resident should be asked the same questions plus the question:

(vii) Do you like the amount of changing and toileting assistance you have received in the last three-days?

The most critical statistic calculated during this trial is the resident's appropriate toileting percentage. Appropriate toileting is calculated by dividing the total number of successful toilets by the total number of toilets plus the number of incontinence voids. Thus, a 50% appropriate toileting statistic will be derived if a resident appropriately toileted during six of the prompts on the three days and was wet on six of the prompts. Two separate major trials have determined that residents who have appropriate toileting percentages above 66% will very likely continue to be continent if exposed to prompted voiding over longer intervals (Schnelle, 1990; Schnelle, Ouslander, & Simmons, 1993); thus, it is recommended that all residents who show an appropriate toileting percentage of 65% or above be placed on a prompted voiding program. Residents who have appropriate toileting percentages of less than 65% seldom show responsiveness with longer term applications of prompted voiding. Based on previous studies in several major clinical trials, approximately 30–40% of residents will show a 66% appropriate toileting percentage or higher during this three-day trial.

The resident's response to the six questions asked prior to and after the three-day trial coupled with his or her behavioral responses to the prompts provided during the trial form the information base used to evaluate a resident's motivation to be continent. Approximately 80% of incontinent residents can provide answers to these questions. Surprisingly, however, approximately 10–20% of incontinent residents in nursing homes show no behavioral or verbal willingness to be either changed or toileted more frequently. Such residents respond to the questions with answers which indicate that they are not bothered by wetness and do not want to be changed or toileted more often. Moreover, when efforts are made to change or toilet such residents on a two-hour schedule during the three-day assessment period, there is often both behavioral and verbal evidence that this UI care is unwanted. For example, direct quotes taken during conduct of the protocol are as follows: "It hurts. I don't want to use the bedpan." "Leave me alone. I don't want changed." Similar comments are heard with noticeable frequency. Residents who express this type of resistance will typically not attempt to toilet even if consistently prompted by skilled and clinically sensitive staff. It is to be remembered that in prompted voiding, residents are not placed on the toilet unless they respond affirmatively to a prompt.

Even though a case can be made that the behavior of these resistive residents is not appropriate and may be caused by psychological conditions such as conditioned helplessness or depression, further treatment for these residents is not indicated at the present time due to three factors:

(i) there are no leads as to what types of behavioral, psychological, or medical interventions could be successful with these resistive patients;

(ii) there are so many incontinent nursing home residents that we currently strain resources to treat and manage UI with even those residents who are motivated to be continent; and

(iii) there is a distinct chance that some residents have such a high sickness burden that they do not care about staying dry and would prefer to not go through the activities involved with toileting, in which case prompted voiding would hinder rather than benefit quality of life.

Research investigating why some residents no longer care about their UI condition is highly appropriate, but treatment should not be considered until such research is completed.

Approximately 10–20% of residents will not attain a 66% appropriate toileting statistic but will show behavioral and verbal evidence that they are motivated to be continent and thatthey want further treatment. Further evaluation for these residents is indicated. The most important behavioral evidence to identify residents in potential need of further treatment is a statistic that reflects their positive response to prompts and their attempts to toilet during the three-day period. This statistic is different from the appropriate toileting statistic which requires not only an attempt to toilet, but also a successful void. If the resident attempts to toilet two times per day, even if they are not successful, it is an excellent sign that they are motivated to be continent. Many residents who attempt to toilet unsuccessfully cannot initiate the voiding process once they are placed on the toilet. Others will attempt to toilet frequently and void small amounts, but will continue to be wet between the checks and hence will not attain a 66% appropriate toilet ratio. These residents very often show stress incontinence, detrusor instability, or detrusor instability with impaired contractility, all of which were discussed in Section 7.19.3.1. Further information is collected during this three-day period which is highly relevant to determining which of these types of incontinence a resident may have. The purpose is to identify all problems appropriately treatable by other interventions. These procedures will be described next.

7.19.6.2 Postvoid Residual Determination and Urinalysis

During the three-day run-in evaluation for prompted voiding, an attempt is made to collect urine for analysis and also to measure a resident's postvoid residual, as is recommended for the basic assessment of people who suffer from incontinence. The basic prompted voiding protocol is temporarily modified in order to obtain this information efficiently. For further efficiency purposes, an attempt is made to collect these data the first time the resident successfully toilets during the three days of prompted voiding rounds. This toileting event offers the opportunity to collect urine for analysis with either a "clean catch" technology or by catheterization if the resident has to be catheterized to measure the amount of urine left in the bladder after urination (postvoid residual).

Clean urine may be difficult to obtain from incontinent, cognitively impaired, immobile nursing home residents. However, one can often noninvasively obtain a "clean catch" urine that accurately matches culture results from a catheterized specimen. For a woman, the technique involves careful cleaning of the perineum with a sterile preparation and then having her void into a disinfected collection device (Ouslander, Schapira, & Schnelle, 1995). For a man, after the penis has been cleaned with a sterile preparation such as betadine, a clean condom catheter can be applied (Ouslander, Schapira, Fingold, & Schnelle, 1995).

The primary purpose of the urinalysis is to exclude hematuria unless the purpose is to investigate symptoms other than UI which are suggestive of a urinary tract infection (e.g., dysuria, recent onset or worsening of UI, unexplained fever or functional decline). The urinalysis can also exclude sterile hematuria, which may be a sign of bladder or upper urinary tract pathology, and glucosuria resulting from poorly controlled diabetes. In nursing home residents it is important to note that bacteriuria should not be treated, because such treatment consistently has been found to have no effect on the severity of UI in this population (Ouslander, Schapira, Schnelle, Uman et al., 1995).

Although rapid urine screening methods are generally thought of as inaccurate for nursing home residents, they can be used to exclude bacteriuria with reasonable accuracy. The combination of negative dipstick tests for both nitrite and leukocyte esterase and a negative enzyme-based screen for bacteriuria (Uriscreen) has a negative predictive value of over 95% for significant bacteriuria among incontinent nursing home residents (Ouslander, Schapira, Fingold, & Schnelle, 1995).

Significant degrees of urinary retention can occur in the absence of symptoms or signs of obstruction or bladder contractility problems. In addition, physical examination (suprapubic palpation, bi-manual examination) is neither sensitive nor specific in detecting significant residual volumes.

Postvoid residual volume is usually determined by catheterization, but a portable ultrasound device is now available that can provide a noninvasive measurement with reasonable accuracy (Ouslander et al., 1994). It is important to note that postvoid residual volume demonstrates substantial intra-individual variability. The timing of the determination in relation to a void, the degree to which the void was natural, the nature of instructions provided to the patient, the amount of straining, and body position can all influence this measurement. Thus, when moderately elevated volumes are detected (e.g., 150–300 ml), repeated determinations may be appropriate. In these situations in particular, the use of ultrasound may be indicated, in order to avoid the need for repeated catheterization. Consistent residual volumes of more than 200 ml should prompt consideration of whether the patient should be referred for further evaluation.

7.19.6.3 Pad Test for Stress Incontinence

The second diagnostic modification of the prompted voiding run-in protocol is designed to evaluate signs of stress incontinence. Although the presence or absence of stress incontinence does not predict responsiveness to prompted voiding, this diagnostic measure is helpful in cases of residents who fail a trial of prompted voiding and who are candidates for other therapeutic interventions for stress incontinence. This test can be performed during the same toileting episodes in which a clean urine and postvoid residual are being obtained. It is performed by having the resident cough several times while standing. The test is positive if there is urine leakage simultaneous with the cough. If the test is negative and if the total bladder volume (voided volume plus residual) is less than 200 ml, the test should be repeated since a low bladder volume can cause a false negative result. In many cases, it is impossible to obtain an adequate stress test because of either low bladder volumes due to involuntary bladder contraction, low toileting frequency, or difficulty following instructions. Attempts to obtain a stress test should be terminated if it cannot be completed successfully during the first few toileting episodes.

7.19.6.4 Mobility Assessment

The third modification to the prompted voiding protocol evaluates the resident's mobility skills to toilet independently and to do so safely. Two critical questions are addressed by this assessment: (i) will a resident require ongoing prompted voiding management because of physical inability to safely toilet? and (ii) should a resident be placed on a rehabilitation program to improve mobility skills related to toileting?

The extent to which a resident can safely perform the mobility tasks involved in toileting can be evaluated most efficiently during the prompted voiding episodes, during which the resident must move for either changing or toileting assistance. This offers a time-efficient and natural opportunity to evaluate the resident's ability to perform three critical skills for independent toileting: standing, moving (either walking or wheeling a wheelchair) to the toilet, and transferring. A standardized behavioral assessment protocol has been designed to assess these skills (Schnelle, MacRae, Ouslander, Simmons, & Nitta, 1995). The instructions and prompting provided to the resident during this protocol are specifically designed to elicit the maximum possible independent effort. A resident is determined to be capable of independent toileting if he or she performs all three of the tasks either when instructed to do so or with instructions plus gentle manual guidance. Research conducted by the present author has demonstrated that 30–40% of nursing home residents are capable of independently performing all three skills necessary for independent toileting. However, many of these otherwise independent residents can complete these tasks only with major effort, due to deconditioning, muscular weakness, and balance problems, which place them at higher risk for falls. Many of the residents are aware of their safety risk and are so afraid to fall that they do not attempt to toilet independently because of this fear (Schnelle et al., 1994). These residents are candidates for further treatment of mobility problems, and it is anticipated that an unknown percentage of these residents may become capable of independent toileting. Mobility interventions that have been found to have positive outcomes for this group are described in Section 7.19.6.6.

The many residents who are not capable of toileting independently and who cannot be taught to do so safely will have to be maintained on prompted voiding indefinitely so as to receive toileting assistance even if other treatments for UI are attempted and successful. For example, even if a resident is capable of maintaining dryness for several hours at a time due to an intervention such as drug therapy, this resident will only be continent to the extent that staff offer him or her an opportunity to toilet on a regular basis.

Similarly, even residents who are capable of performing all toileting tasks independently must be maintained on prompted voiding if they are either very slow or show poor strength and endurance. Findings indicate that the average time it takes an incontinent resident to stand is six seconds; the average time span during which incontinent residents can sustain walking in a standardized endurance test is 2.6 minutes (Schnelle et al., 1994). These mobility problems limit the ability to stay dry, and the muscle weakness associated with the slowness and poor endurance place the resident at risk for falling. Thus, many such residents will have to be maintained on prompted voiding so that toileting assistance can be provided for the purpose of getting the residents to the bathroom quickly and safely. The issues involved in the maintenance of prompted voiding are complex and require an effective quality assurance system (described in Section 7.19.6.7). However, a decision regarding whether further evaluation is to be done should be made prior to determining if residents are to be managed on an ongoing basis with prompted voiding. The issue of who are candidates for further evaluation is discussed next.

7.19.6.5 Who are Candidates for Further Evaluation?

Urologic and/or urodynamic evaluation should be considered for residents with sterile hematuria or high (> 200 ml) postvoid residual volumes. Only residents who are candidates for specific interventions based on the results of these evaluations should be referred. Further evaluation may also be indicated for those residents who show positive stress symptoms or high wetness frequencies (two or more times per day), despite attempts to toilet two or more times a day. In these cases, either a stress or urge diagnosis is suspected, and the treatment options available for these conditions (described in Section 7.19.3) could be attempted. Given their attempts to toilet, these residents clearly have demonstrated at least the minimal cognitive capacity and motivation necessary to be appropriate candidates for these types of treatment. Recommendations for treatment with alternative behavioral, surgical, or pharmacological interventions should be qualified by the fact that there is no strong evidence that such interventions are successful, even with carefully targeted nursing home residents.

Further evaluation by a physical therapist is recommended for residents who either cannot independently perform all three mobility skills related to toileting or who do so very poorly and who are felt to have some potential for improving these skills. Based on previous work, we estimate that most incontinent residents will be in need of either a physical therapist consultation or a rehabilitative nursing exercise program designed to improve mobility, endurance, and muscle strength. Moreover, as discussed in Section 7.19.2, we reviewed data which indicates that UI and immobility problems are jointly predictive of a high rate of infection and hospitalization to treat these infections. An intervention program that addresses both risk factors is needed, even if the resident cannot attain toileting independence because of reversible physical or cognitive problems, since such a program may also prevent infections. Mobility interventions that may be effective with incontinent nursing home residents are now described.

7.19.6.6 Treatment for Mobility Problems that Limit Independent Toileting

It is not known to what degree physical therapy assessment and treatment is appropriate for incontinent nursing home residents who are chronically ill, demented, and who require maximal assistance to perform the three mobility skills related to independent toileting. Two separate exercise rehabilitation trials, each lasting for nine-week periods, have not been successful in significantly improving the independence level of these very dependent residents (Schnelle, MacRae et al., 1995). However, improvements have been found in endurance, speed, and safety levels for all three mobility skills among residents capable of performing these skills independently or with minimal physical assistance (Schnelle et al., 1996; Schnelle, MacRae et al., 1995). These findings are a result of a mobility intervention utilizing a functional incidental training (FIT) technique.

FIT is implemented concurrently with prompted voiding and involves two primary additions. The two additional FIT steps are done each time prompted voiding occurs when the resident is out of bed. First, the resident is encouraged to perform the toileting-related activities with minimal staff assistance. This is in contrast to standard care since, as a result of staff time constraints, aides typically provide maximal assistance to the residents in completing the prompted voiding sequence. In the absence of functional considerations, it is more efficient to provide such maximal assistance and to minimize the resident's involvement; however, the FIT intervention promotes functional activity. In the second FIT step, the resident is given repeated practice in each functional activity (e.g., standing, moving, transferring) and works toward an individualized goal, which is increased progressively as the resident demonstrates improvement.

The application of the FIT protocol three times per day for nine weeks has been shown in a controlled trial to result in significant improvements in walking endurance, standing speed, wheelchair endurance, and safety measures for a group of 38 residents (Schnelle, MacRae et al., 1995). Unfortunately, only 10% of the residents showed improvement on these mobility tasks to the point where they were rated as being at a low risk for falls. A modification of the FIT protocol is currently being tested which includes resistance training designed to further improve muscle strength. There is strong evidence from several trials that even very old nursing home residents will show strength improvements with resistance training (Fiatarone et al., 1994; Mulrow, Gerety, Kanten, & Cornell, 1994). In addition, FIT is being implemented for longer time periods to explore further its potential for producing a higher level of mobility independence and a lower risk for injury and infections which may require hospitalization.

However, even if these mobility interventions are successful, most residents will require routine prompting and some physical assistance to maintain continence and to maintain their improved mobility performance. A quality assurance system to assure that staff continue to provide such assistance is critical; therefore, the behavioral management skills that staff must learn to successfully implement such a system are discussed next. This behavioral management involves a system of accurately monitoring the resident incontinence status. It is in this critical arena that the role of the behavioral psychologist is most needed.

7.19.6.7 Staff Management and Continence Maintenance

Three major issues must be addressed if toileting assistance and the continence maintenance aspects of the prompted voiding program are to be accomplished. These issues, although discussed here specifically with regard to prompted voiding, are essential to any intervention requiring caregiver assistance. First, procedures are needed to target those residents who will respond well if ongoing toileting assistance is provided. Second, an information system must be developed to monitor incontinence outcomes effectively so as to facilitate quality control. Third, supervisors and direct care staff who actually implement the prompted voiding protocol must be taught how to use the outcome data to identify and correct problems. A comprehensive description of how these issues must be addressed to facilitate the implementation of prompted voiding can be found in several places (Ouslander & Schnelle, 1993; Schnelle, Ouslander, Osterwell, & Blumenthal, 1993).

The methodology to target residents who can be responsive to the prompted voiding toileting program has been described earlier in this chapter; specifically a three-day run-in assessment is described for this purpose. This same three-day assessment also provides the outcome data needed to maintain the incontinence system effectively. In review, during this three-day period residents are checked for wetness and prompted every two hours. Those residents who have a 66% appropriate toileting ratio or better are designated as responsive and are recommended for ongoing maintenance. The wetness data from these responsive residents collected during the two-hour checks can now be used to generate a control chart, which is the basic information tool used to monitor the prompted voiding program, thereby addressing the need for effective quality control (Fogarty, Schnelle, & Newman, 1989; Schnelle, Newman, Fogarty, Wallston, & Ory, 1991).

A control chart is a graphic display that allows the viewer visually to ascertain expected versus actual outcomes (Figure 2). The control chart in this case is constructed by graphing wetness data collected during each two-hour block during the three-day run-in period and calculating the average and standard deviation of this wetness. A warning limit is set two standard deviations above the average. This warning limit signifies unexpected variability and indicates either a failure of the prompted voiding work process to be implemented consistently or, perhaps, a change in the resident's status. For example, the control chart shown in Figure 2 was generated for a group of 22 residents. This control chart shows that 12% of the residents who showed $\geqslant 66\%$ appropriate toileting rate during the three-day assessment period were found wet on the two-hour checks. A warning limit was set above this 12% average at 40%. This warning limit is two standard deviations above the average. An upper warning limit is established three standard deviations above the average. Supervisory nurses and direct care staff can now effectively monitor the prompted voiding program by randomly checking at least 10 residents per week for wetness. If the findings of this check, termed a control check, are within the average and two standard deviations of the expected average, the prompted voiding system can be said to be in control. If the wetness rate for the 10 patients exceeds the warning limit, then an analysis should be done to identify the causes of this excessive variability. A failure of staff to implement the prompted voiding program consistently or a deterioration in a resident's status are two potential explanations for an unusually high wetness check.

Software has been developed to facilitate the analysis and interpretation of the control checks, and supervisory and direct care staff have been taught how to use the software to identify problems and to provide feedback to the staff who are implementing prompted voiding. The implementation of this computerized system has resulted in successful maintenance of high levels of resident dryness during six-month to one-year evaluation periods in eight nursing homes (Schnelle, McNees, Crooks, & Ouslander, 1995). The failure of more traditional employee training methods to sustain such improvements have been described in several places (Schnelle, Newman, & Fogarty, 1990a). Unfortunately, the staff training and management interventions that have been evaluated have only been successful as long as research staff were present to provide support and social incentives to both supervisory and direct care staff to maintain the system.

The fact that maintaining resident dryness with a toileting program is more time consuming than simply changing residents when they are wet produces strong natural pressures that reduce staff motivation to maintain the program. The development of incentives to increase the motivation of nursing home staff to implement the more labor intense toileting program is an important new area to be investigated, and this is discussed in the next section.

7.19.7 FUTURE DIRECTIONS

Numerous areas related to incontinence treatment that are in need of research can be identified for both community dwelling and nursing home populations. Furthermore, many of these areas have direct relevance to behavioral scientists. For example, in the case of community dwelling populations major questions remain as to why most people with incontinence symptoms do not seek treatment. Behavioral interventions to increase people's willingness to be treated could possibly be developed if such information were available. Furthermore, such interventions to increase people's awareness of treatment options are clearly justified by the evidence that urinary incontinence is predictive of general functional decline in the community dwelling elderly. The fact that urinary incontinence is associated with a decline in measures of emotional functioning in the community dwelling elderly is also an area of interest for future behavioral research (Tinetti et al., 1995).

In the case of nursing home residents, major improvements have been made with incontinence problems by utilizing the assessment, treatment, and staff management system described in this chapter. In addition, the work to date suggests future research directions that should lead to further improvements in how incontinence care might be done in nursing homes. Two of these most important future directions have direct implications for behavioral scientists.

The first issue relates to the development of incentives that will motivate staff to maintain the quality control and information feedback system that is so critical to the maintenance of the prompted voiding program. The reality that maintaining continence is more time consuming than simply changing residents when they are wet contributes to the maintenance problem. However, the fact that even residents who are responsive to the prompted voiding program are not good self-advocates because of high sickness burden, dementia, or other unknown factors has not been discussed, and makes the need for incentives even more imperative. Unfortunately, most residents do not provide sufficient positive verbal feedback to nursing home staff when they provide toileting assistance so as to maintain the more labor intense toileting activities, and many residents can be convinced that they should accept their incontinence condition and a changing schedule as opposed to a prompting and toileting schedule.

Several approaches could be taken in reaction to this problem. First, external incentives that link the maintenance of the prompted voiding protocol to job advancement or salary increases could be attempted. These reinforcement strategies have been used successfully with other institutional populations that are not good self-advocates or are incapable of socially reinforcing staff for better care. A description of this behavioral research is available (Burgio & Burgio, 1990; Burgio et al., 1990).

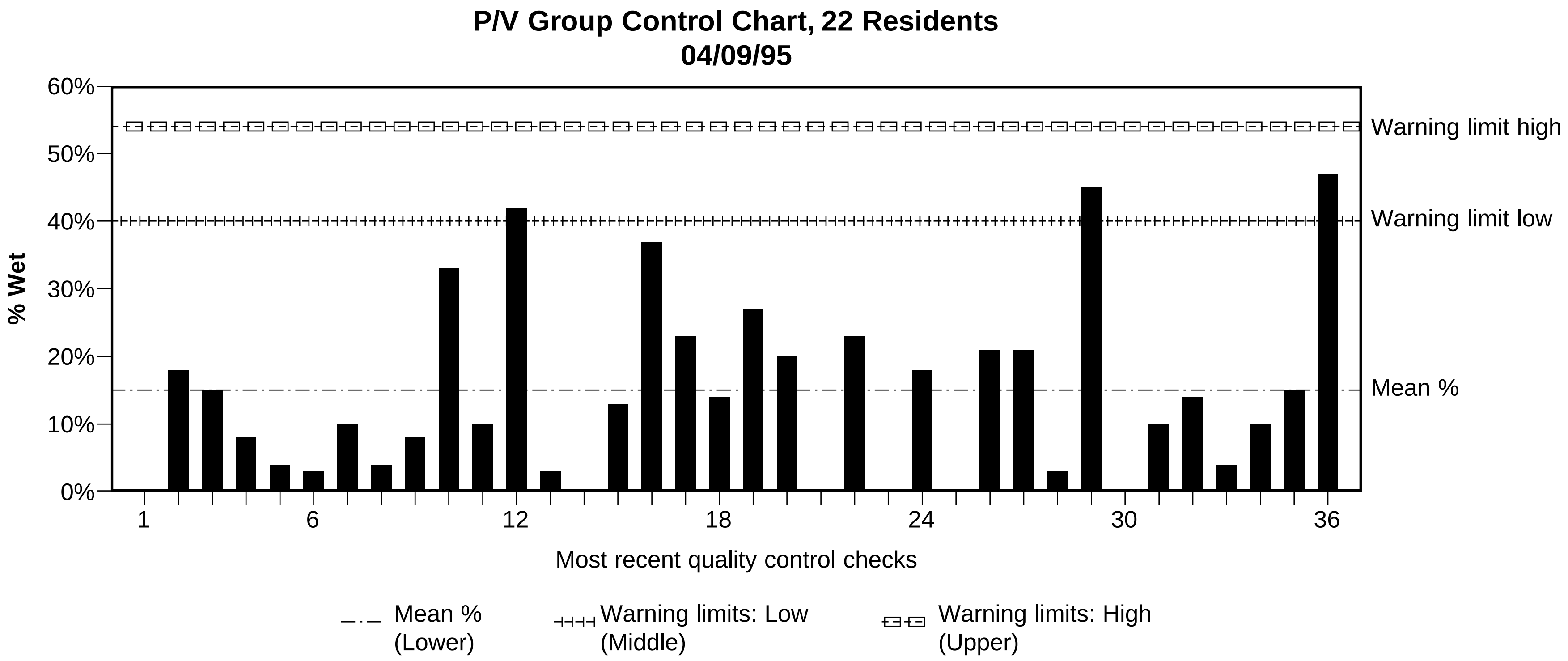

Figure 2 Sample control chart.

A second approach would be to develop interventions designed to make either residents or their families better advocates for their own care. With this approach, families and residents would have to be taught how to recognize good incontinence care and to interact with nursing staff in a manner so as to reinforce good care when it occurs. This form of patient/family involvement has been discussed by several consumer groups, but there have been no examples where this type of intervention has been done successfully.

A second direction for future behavioral research is related to the motivation of residents to be continent. The fact that many residents (estimated 15–30%) react negatively to the prompted voiding procedures and appear to prefer to be changed as opposed to toileted has been discussed previously in this chapter. These residents are not even carried into the staff management phase of the prompted voiding program because of their unresponsiveness to the prompting procedure during the run-in evaluation period. Research to measure residents' preferences for care and to determine the factors that contribute to these preferences would be important innovations that could guide how incontinence care should be delivered. For example, are preferences to urinate in pads and be changed as opposed to toileted influenced by the residents' reduced expectations of what is possible in the nursing home, or by true preferences for changing because it is less bothersome or painful than the toileting sequence? There are reasons to believe that both hypotheses may be plausible. The fact that many residents have experienced incontinence symptoms that were untreated before they entered the nursing home has been discussed in earlier sections in this chapter. It is thus plausible to predict that many residents may have adapted to the condition and sincerely are no longer bothered by wetness. Alternatively, it is also plausible to predict that some residents may have been convinced by nursing home staff that they should accept their incontinence condition because staff do not have time to help with toileting and thus have resigned themselves to the condition. Preferences based on the patient's acceptance and adaption to the condition versus preferences created by nursing home staff persuasion would require different interventions. In the first case it is arguable if anything should be done to convince residents of their need to be concerned about toileting, whereas in the second case an intervention to reverse the dependency and low expectations about care created by nursing staff is certainly necessary. Unfortunately, we know almost nothing about how to assess accurately the preferences of demented nursing home residents. The development of such a technology will allow us to begin the process of truly individualizing resident care according to their needs, not only in the case of incontinence care but also in other care areas.

7.19.8 REFERENCES

Benvenuti, F., Caputo, G. M., Bandinelli, S., Mayer, F., Biagini, C., & Sommavilla, A. (1987). Re-educative treatment of female genuine stress incontinence. *American Journal of Physical Therapy*, *66*, 155–168.

Bo, K., Hagen, R. M., Kvarstein, B., Jorgensen, J., & Larsen, S. (1990). Pelvic floor muscle exercises for the treatment of female stress urinary incontinence. III. Effects of two different degrees of pelvic floor muscle exercises. *Neurourological Urodynamics*, *9*(5), 489–502.

Bridges, N., Denning, J., Olah, K. S., & Aorrar, D. J. (1988). A prospective trial comparing interferential therapy and treatment using cones in patients with symptoms of stress incontinence. *Neurourological Urodynamics*, *7*(3), 267–268.

Burgio, K., Whitehead, W. E., & Engel, B. T. (1985). Urinary incontinence in the elderly: Bladder–sphincter biofeedback and toileting skills training. *Annals of Internal Medicine*, *103*, 507–515.

Burgio, K. L., Matthews, K. A., & Engel, B. T. (1991). Prevalence, incidence and correlates of urinary incontinence in healthy, middle-aged women. *Journal of Urology*, *146*, 1255–1259.

Burgio, K. L., Robinson, J. C., & Engel, B. T. (1986). The role of biofeedback in Kegel exercise training for stress urinary incontinence. *American Journal of Obstetrical Gynecology*, *154*(1), 58–64.

Burgio, K. L., Stutzman, R. E., & Engel, B. T. (1989). Behavioral training for post-prostatectomy urinary incontinence. *Journal of Urology*, *141*, 303–306.

Burgio, L. D., & Burgio, K. L. (1990). Institutional staff training and management: a review of the literature and a model for geriatric long-term care facilities. *International Journal of Aging and Human Development*, *30*(4), 287–302.

Burgio L. D., Engel, B. T., Hawkins, A., McCormick, K., Scheve, A., & Jones, L. T. (1990). A staff management system for maintaining improvements in continence with elderly nursing home residents. *Journal of Applied Behavior Analysis*, *23*, 111–118.

Burns, P. A. Pranifoff, K., Nochajski, T. H., Hadly, E. C., Levy, K. J., & Ory, M. G. (1993, July). A comparison of effectiveness of biofeedback and pelvic muscle exercise treatment of stress-incontinence in older community-dwelling women. *Journal of Gerontology*, *48*(4), 167–174.

Burton, J., Pearce, L., Burgio, K. L., Engel, B. T., & Whitehead, W. E. (1988). Behavioral training for urinary incontinence in elderly, ambulatory patients. *Journal of the American Geriatrics Society*, *36*(8), 693–698.

Colling, J., Ouslander, J., Hadley, B. J., Eisch, J., & Campbell, E. (1992). The effects of patterned urge response toileting (PURT) on urinary incontinence among nursing home residents. *Journal of the American Geriatrics Society*, *40*, 135–141.

Creditor, M. C. (1993). Hazards of hospitalization of the elderly. *Annals of Internal Medicine*, *118*, 219–223.

Diokno, A. C., Brock, B. M., Brown, M. B., & Herzog, A. R. (1986). Prevalence of urinary incontinence and other urological symptoms in the non-institutionalized elderly. *Journal of Urology*, *136*, 1022–1025.

Diokno, A. C., Brock, B. M., Herzog, A. R., & Bromberg, J. (1990). Medical correlates of urinary incontinence in the elderly. *Urology*, *36*, 129–138.

Elia, G. & Bergman, A. (1993, February). Pelvic muscle exercises: When do they work? *Obstetrical Gynecology*, *81*(2), 283–286.

Fantl, J. A., Cardozo, L., & McClish, D. K. (1994). Estrogen therapy in the management of urinary incontinence in postmenopausal women: A meta-analysis. First report of the hormones and urogenital therapy committee. *Obstetrical Gynecology*, *83*, 12–18.

Fantl, J. A., Wyman, J. F., McClish, D. K., Harkins, S. W., Elswick, R. K., Taylor, J. R., & Hadley., E. C. (1991). Efficacy of bladder training in older women with urinary incontinence. *Journal of the American Medical Association*, *265*(5), 609–613.

Ferrie, B. G., Smith, J. S., Logan, D., Lyle, R., & Paterson, P. J. (1984). Experience with bladder training in 65 patients. *British Journal of Urology*, *56*, 482–484.

Fiatarone, M. A., O'Neill, E. F., Ryan, N. D., Clements, K. M., Solares, G. R., Nelson, M. E., Roberts, S. B., Kehayias, J. J., Lipsitz, L. A., & Evans, W. J. (1994). Exercise training and nutritional supplementation for physical frailty in very elderly people. *New England Journal of Medicine*, *330*, 1769–1775.

Flynn, L., Cell, P., & Luisi, E. (1994). Effectiveness of pelvic exercises in reducing urge incontinence among community residing elders. *Journal of Gerontological Nursing*, *20*(5), 23–27.

Fogarty, T. E., Schnelle, J. F., & Newman, D. R. (1989). Statistical quality control in nursing homes: A key indicator to evaluate patient incontinence care. *Quality Review Bulletin*, *15*(9), 273–278.

Harrison, G. L. & Memel, D. S. (1994). Urinary incontinence in women: Its prevalence and its management in a health promotion clinic. *British Journal of General Practice*, *44*(381), 149–152.

Help for Incontinent People (HIP) & the UCLA-JHA/Anna & Harry Borun Center for Gerontological Research (1994). *Managing incontinence in frail elderly dependent institutionalized and community dwelling persons: An agenda for research and care*. Monograph.

Holmes, D. M., Montz, F. J., & Stanton, S. L. (1989). Oxybutinin versus propantheline in the management of detrusor instability. A patient-regulated variable dose trial. *British Journal of Obstetrical Gynaecology*, *96*(5), 607–612.

Hu, T. (1994). *The cost impact of urinary incontinence on health care services*. Paper presented at the National Multi-Specialty Nursing Conference on Urinary Continence, Phoenix AZ.

Keating, J. C., Jr., Schulte, E. A., & Miller, E. (1988). Conservative care of urinary incontinence in the elderly. *Journal of Manipulative Physiological Therapy*, *11*(4), 300–308.

McCormick, K. A., & Burgio, K. (1984). Incontinence: An update on nursing care measures. *Journal of Gerontological Nursing*, *10*(10), 16–23.

McDowell, B. J., Burgio, K. L., Dombrowski, M., Locher, J. L., & Rodriguez, E. (1992). An interdisciplinary approach to the assessment and behavioral treatment of urinary incontinence in geriatric outpatients. *Journal of the American Geriatrics Society*, *40*, 370–374.

Moore, K. H., Hay, D. M., Imrie, A. E., Watson, A., & Goldstein, M. (1990, November). Oxybutyning hydrochloride (3 mg) in the treatment of women with idiopathic detrusor instability. *British Journal of Urology*, *66*(5), 479–485.

Mouritsen, L., Frimodt, M. C., & Moller, M. (1991). Long-term effect on pelvic floor exercises on female urinary incontinence. *British Journal of Urology*, *68*(1), 32–37.

Mulrow, C. D., Gerety, M. B., Kanten, D., & Cornell, J. E. (1994). A randomized trial of physical rehabilitation for very frail nursing home residents. *Journal of the American Medical Association*, *271*, 519–524.

Murtaugh, C. M., & Freiman, M. P. (1995). Nursing home residents at risk of hospitalization and the characteristics of their hospital stays. *The Gerontologist*, *35*(1), 35–43.

National Kidney and Urologic Diseases Advisory Board (1994, March). *Barriers to rehabilitation of persons with end-stage renal disease or chronic urinary incontinence*. Workshop summary report, Bethesda, MD.

Noelker, L. (1987). Incontinence in elderly cared for by family. *Gerontologist*, *27*, 194–200.

Ouslander, J. G. & Fowler, E. (1985). Management of urinary incontinence in veterans administration nursing homes. *Journal of the American Geriatrics Society*, *33*, 33–40.

Ouslander, J. G., Kane, R. L., & Abrass, I. B. (1982). Urinary incontinence in elderly nursing home patients. *Journal of the American Medical Association*, *248*, 1194–1198.

Ouslander, J. G., Morishita, L., Blaustein, J., Orzeck, S., Dunn, S., & Sayre, J. (1987). Clinical, functional and psychosocial characteristics of an incontinent nursing home population. *Journal of Gerontology*, *42*, 631–637.

Ouslander, J. G., Palmer, M. H., Rovner, B. W., & German, P. S. (1993). Urinary incontinence in nursing homes: Incidence remission and associated factors. *Journal of the American Geriatrics Society*, *41*(10), 1083–1089.

Ouslander, J. G., Schapira, M., Fingold, S., & Schnelle, J. F. (1995). Accuracy of rapid urine screening tests among incontinent nursing home residents with asymptomatic bacteriuria. *Journal of the American Geriatrics Society*, *43*, 772–775.

Ouslander, J. G., Schapira, M., & Schnelle, J. F. (1995). Urine specimen collection from incontinent female nursing home residents. *Journal of the American Geriatrics Society*, *43*, 279–281.

Ouslander, J. G., Schapira, M., Schnelle, J. F., Uman, G., Fingold, S., Tuico, S., & Nigam, J. G. (1995). Does eradicating bacteriuria affect the severity of chronic urinary incontinence in nursing home residents? *Annals of Internal Medicine*, *122*, 749–754.

Ouslander, J. G., & Schnelle, J. F. (1993). Assessment, treatment and management of urinary incontinence in the nursing home. In L. Rubenstein & D. Wieland (Eds.), *Improving care in the nursing homes: Comprehensive reviews of clinical research*. Newbury Park, CA: Sage.

Ouslander, J. G., & Schnelle, J. F. (1995). Incontinence in the nursing home. *Annals of Internal Medicine*, *122*, 438–449.

Ouslander, J. G., Schnelle, J. F., Uman, G., Fingold, S., Nigam, J. G., Tuico, E., & Bates-Jensen, B. (1995a) Does oxybutynin add to the effectiveness of prompted voiding for urinary incontinence among nursing home residents? A placebo controlled trial. *Journal of the American Geriatrics Society*, *43*, 610–617.

Ouslander, J. G., Schnelle, J. F., Uman, G., Fingold, S., Nigam, J. G., Tuico, E., & Bates-Jensen, B. (1995b). Predictors of successful prompted voiding among incontinent nursing home residents. *Journal of the American Medical Association*, *273*(17), 1366–1370.

Ouslander, J. G., Simmons, S. F., Tuico, E., Nigam, J. G., Fingold, S., Bates-Jensen, B., & Schnelle, J. F. (1994). Use of a portable ultrasound device to measure post-void residual volume among incontinent nursing home residents. *Journal of the American Geriatrics society*, *42*, 1189–1192.

Pengelly, A. W. & Booth, C. M. (1980). A prospective trial of bladder training as treatment for detrusor instability. *British Journal of Urology*, *52*, 463–466.

Resnick, N. M. & Yalla, S. V. (1985). Management of urinary incontinence in the elderly. *New England Journal of Medicine*, *313*, 800–805.

Riva, D. & Casolabi, E. (1984). Oxybutynin chloride in the treatment of female idiopathic bladder instability.

Clinical Experimental Obstetrics and cynecology, 11(1–2), 37–42.

Rose, M., Baigis-Smith, J., Smith, D., & Newman, D. (1990). Behavioral management of urinary incontinence in homebound older adults. *Home Healthcare Nurse, 8*, 10–15.

Schnelle, J. F. (1991). *Managing urinary incontinence in the elderly*. New York: Springer.

Schnelle, J. F., Keeler, E., Hays, R. D., Simmons, S. F., Ouslander, J. G., & Siu, A. L. (1995). A cost and value analysis of two interventions with incontinent nursing home residents. *Journal of the American Geriatrics Society, 43*(10), 1112–1117.

Schnelle, J. F., MacRae, P. G., Giacobassi, K., MacRae, H. S-H., Simmons, S. F., & Ouslander, J. G. (1996). Exercise with physically restrained nursing home residents: Maximizing benefits of restraint reduction. *Journal of the American Geriatrics Society, 44*, 507–512.

Schnelle, J. F., MacRae, P. G., Ouslander, J. G., Simmons, S. F., & Nitta, M. (1995). Functional incidental training, mobility performance, and incontinence care with nursing home residents. *Journal of the American Geriatrics Society, 43*(12), 1356–1362.

Schnelle, J. F., MacRae, P. G., Simmons, S. F., Uman, G., Ouslander, J. G., Rosenquist, L. L., & Chang, B. (1994). Safety assessment for the frail elderly: A comparison of restrained and unrestrained nursing home residents. *Journal of the American Geriatrics Society, 42*, 586–592.

Schnelle, J. F., McNees, P., Crooks, V., & Ouslander, J. G. (1995). The use of a computer-based model to implement an incontinence management program. *The Gerontologist, 35*(5), 656–665.

Schnelle, J. F., Newman, D. R., Abbey, J. C., Dray, R. J., Petrilli, C. O., & Hadley, E. C. (1990). Urodynamic and behavioral analysis of incontinence in nursing home patients. *Behavior, Health, and Aging, 1*(1), 43–49.

Schnelle, J. F., Newman, D. R., Fogarty, T. E., Wallston, K., & Ory, M. (1991). Assessment and quality control of incontinence care in long-term nursing facilities. *Journal of the American Geriatrics Society, 39*, 165–171.

Schnelle, J. F., Newman, D. R., & Fogarty, T. E. (1990a). Management of patient continence in long-term care nursing facilities. *The Gerontologist, 30*(3), 373–376.

Schnelle, J. F., Newman, D. R., & Fogarty, T. E. (1990b). Statistical quality control in nursing homes: Assessment and management of chronic urinary incontinence. *Health Services Research, 25*(4), 627–637.

Schnelle, J. F., Newman, D. R., White, M., Abbey, J., Wallston, K. A., Fogarty, T., & Ory, M. G. (1993). Maintaining continence in nursing home residents through the application of industrial quality control. *The Gerontologist, 33*, 114–121.

Schnelle, J. F., & Ouslander, J. G. (1996). Management of incontinence in long-term care. *Urinary incontinence*. St. Louis: MO: Mosby-Year Book.

Schnelle, J. F., Ouslander, J. G., Osterweil, D., & Blumenthal, S. (1993). Total quality management: Administrative and clinical applications in nursing homes. *Journal of American Geriatrics Society, 41*, 1259–1266.

Schnelle, J. F., Ouslander, J. G., & Simmons, S. F. (1993). Predicting nursing home residents' responsiveness to a urinary incontinence treatment protocol. *International Uro-Gynecology Journal, 4*, 89–94.

Schnelle, J. F., Sowell, V. A., Hu, T. W., & Traughber, B. T. (1988a). A behavioral analysis of the labor cost of managing continence and incontinence in nursing home patients. *Journal of Organizational Behavior Management, 9*(2), 137–153.

Schnelle, J. F., Sowell, V. A., Hu, T. W., & Traughber, B. T. (1988b). Reduction of urinary incontinence in nursing homes: Does it reduce or increase cost? *Journal of the American Geriatrics Society, 36*, 34–39.

Stewart, R. B., Moore, M. T., May, F. E., Marks, R. G., & Hale, W. E. (1992). Nocturia: a risk factor for falls in the elderly. *Journal of the American Geriatric Society, 40*, 1217–1220.

Tapp, A. J., Cardozo, L. D., Versi, E., & Cooper D. (1990, June). The treatment of detrusor instability in post-menopausal women with oxybutynin chloride: A double-blind placebo controlled study. *British Journal of Obstetrical Gynaecology, 97*(6), 521–526.

Thuroff, J. W., Bunke, B., Ebner, A., Faber, P., de Geeter, P., Hannappel, J., Heidler, H., Madersbacher, H., Melchior, H., Schafer, W., Schwenzer, T., & Stockle, M. (1991). Randomized, double-blind, multicenter trial on treatment of frequency, urgency and incontinence related to detrusor hyperactivity: Oxybutynin versus propantheline versus placebo. *Journal of Urology, 145*(4), 813–817.

Tinetti, M. E., Inouye, S. K., Gill, T. M., & Doucette, J. T. (1995). Shared risk factors for falls, incontinence, and functional dependence. Unifying the approach to geriatric syndromes. *Journal of the American Medical Association, 273*(17), 1348–1353.

United States Agency for Health Care Policy and Research (1996). *Urinary incontinence in adults: Acute and chronic management*. Rockville, MD: US Department of Health and Human Services, Public Health Services, Agency for Health Care Policy and Research.

Warren, J. W., Damron, D., Tenney, J. H., Hoopes, J. M., Deforge, B., & Muncie, H. L., Jr. (1987). Fever, bacteremia and death as complications of bacteriuria in women with long-term urethral catheters. *Journal of Infectious Diseases, 155*, 1151–1158.

Zeegers, A. G. M., Kiesswetter, H., Kramer, A. E. J. L., & Jonas, U. (1987). Conservative therapy of frequency, urgency and urge incontinence: A double-blind clinical trial of flavoxate hydrochloride, oxybutinin chloride, emepronium bromide and placebo. *World Journal of Urology, 5*(1), 57–61.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.20

Community Mental Health Services in the United States and the United Kingdom: A Comparative Systems Approach

BOB G. KNIGHT and BRIAN KASKIE
University of Southern California, Los Angeles, CA, USA
BOB WOODS
University of Wales, Bangor, UK
and
EMILY PHIBBS
University College London, UK

7.20.1 INTRODUCTION 456

7.20.2 MULTIPLE SYSTEMS OF MULTILEVEL CARE 456

7.20.3 THE STATUS QUO ANTE: LONG-TERM INSTITUTIONAL PSYCHIATRIC CARE 457
  7.20.3.1 Deinstitutionalization of Mentally Ill 457
  7.20.3.2 Do The Old Institutions Still Exist? 458
  7.20.3.3 Who Was Discharged from Psychiatric Institutions? 458

7.20.4 ACUTE PSYCHIATRIC INPATIENT SERVICES 458
  7.20.4.1 Who is Served in Geropsychiatric Inpatient Units? 459

7.20.5 MENTALLY ILL OLDER ADULTS IN NURSING HOME CARE 459

7.20.6 OLDER ADULTS IN COMMUNITY MENTAL HEALTH SERVICES 461
  7.20.6.1 Community-based Mental Health Outpatient Care in the UK 461
    7.20.6.1.1 Multidisciplinary teams 461
    7.20.6.1.2 Memory clinics 462
    7.20.6.1.3 Family therapy clinics 463
  7.20.6.2 Community Mental Health Outpatient Services in the US 464
    7.20.6.2.1 What types of clients are served? 464
    7.20.6.2.2 Were CMHCs effective in outreach to elderly? 464
    7.20.6.2.3 What comes next? 465
  7.20.6.3 Summary: Community-based Outpatient Services 465

7.20.7 PARTIAL HOSPITAL SERVICES 465
  7.20.7.1 British Models 465

7.20.7.2 American Models 466
7.20.7.3 Summary 466
7.20.8 RESPITE CARE 466
7.20.9 CAREGIVER SUPPORT 467
7.20.10 MENTAL HEALTH SERVICES IN THE HEALTH CARE SYSTEM 468
7.20.11 DISCUSSION 469
7.20.11.1 Cross-national Systems Comparisons 469
7.20.11.2 Dementia Services 469
7.20.11.3 Mental Health and Primary Care 469
7.20.11.4 Roles for Clinical Psychology in Community Care 470
7.20.11.4.1 Development of clinical geropsychology 471
7.20.12 SUMMARY 471
7.20.13 REFERENCES 471

7.20.1 INTRODUCTION

Community-based mental health services assist older adults with serious psychological disorders who have been moved out of institutional care facilities, as well as seeking to prevent older adults with such disorders from entering institutional care. In addition, community-based services assist older adults with acute psychological distress who are not at risk of placement in institutions. In this chapter, we describe community-based mental health services in the United Kingdom and the United States, we highlight similarities and differences, and we explore the historical and political–economic contexts that shaped these service systems.

7.20.2 MULTIPLE SYSTEMS OF MULTILEVEL CARE

The notion that community-based services for older adults with mental illness form one coordinated system is not a sufficient description. Knight (1991; Knight & Kaskie, 1995) argued that older adults with mental health problems are embedded in a variety of distinct care systems at both institutional and community-based levels: mental health systems, medical care services, long-term care for the elderly, and aging network services. Our understanding of community-based mental health services necessitates a systems level analysis where each system is considered independently and in relation to each other. This systems analysis departs from the more typical "coordination of services" and "gaps in the system" arguments in four distinct ways.

First, we contend that systems of care operate on different principles and create a legitimately confusing service world for both the older client and the professional service provider. Gaps in service provision exist because different systems were created for different purposes (mental health treatment, medical care, or social support). These systems have developed in response to different historical pressures rather than because someone overlooked elements while rationally laying out a system for older adults.

Second, failures to coordinate are more often dictated by economic pressures than by muddled planning or by providers' personal failings. Economic resources for mental health services are scarce and have become even more meager since the Reagan–Thatcher era of tax cuts and cost reductions to public services. Older people, especially those who need mental health services, continue to suffer from the "last-in, first-out" principle. In the UK, financial pressures and the increasing costs of acute medicine have drawn health services funding away from the chronic problems typical of those faced by older adults. In the US, support for older adults' mental health services by the Carter administration was rescinded by Reagan era cutbacks. Older adult services receive a minimal amount of funding, and continue to be subjected to further cutbacks. With such severely limited resources, the economic pressure is for systems to compete rather than to cooperate.

Third, service use is further complicated because older adults experience coexisting physical health and/or social difficulties, making it all too easy for each system to displace responsibility to another "more appropriate" system of care. A system may limit service delivery to existing clients and may not accept clients who "belong" to some other system. Further, limited financial support may cause a system to not provide adequate care even to those older adults who "belong" to the system. As noted by Gatz and Smyer (1992, p. 744), "everyone and no one is responsible for the clients." Indeed, the task of assuring appropriate community-based care for older adults with mental health problems becomes one of securing appropriate services from multiple competing systems, each of which has distinctive operating principles and incentives to provide services.

Fourth, the delivery of community-based services is further complicated by the fact that the different systems do not share the same definition of "community." For example, Knight and Lower-Walker (1985) noted that the mental health system moved older adults into nursing homes, which it defined as a transition from state hospital care to "the community." The aging network in the US has, in marked contrast, tended to define nursing homes as the prototypical institution from which older adults must be moved "to the community."

Our objective is to place the provision of mental health services for older adults in the context of the multilevel mental health services system as well as within other systems of care: long-term care, medical health care, and social services. In what follows, we begin with the long-term hospitals for older adults with mental disorders which are the historical antecedents of community care. The remainder of the chapter will consider community-based interventions.

7.20.3 THE STATUS QUO ANTE: LONG-TERM INSTITUTIONAL PSYCHIATRIC CARE

In the UK, county asylums emerged from the seventeenth and eighteenth century workhouses and institutions of confinement which were common throughout industrial–urban Europe (Foucault, 1965). The nineteenth century saw the development of psychiatry and the redefinition of these institutions as places for treatment of the insane (Foucault, 1965). Many county asylums became large institutions with several thousand inmates. Arie and Jolley (1982) commented that in the mid-1970s half of all asylums were occupied by people over 65, reflecting that:

> Mental hospitals which were built to provide asylum for the mentally ill before more active treatments were available, are nowadays increasingly occupied by old people whose numbers would grow even further were it not for enforced government rationing and not uncommon reluctance on the part of psychiatrists to admit them. (Arie & Jolley, 1982, p. 225)

Institutional psychiatric care developed in the US in the late nineteenth century as a humane alternative to imprisonment, to mistreatment in family homes, and to vagabond status for the mentally ill (Rosen, Pancake, & Rickards, 1995). The US had a shorter, less urban–industrial history, and so institutional care emerged as an alternative to no care rather than a redefining of existing institutions. Older adults were often over-represented in US inpatient psychiatric institutions in comparison to their representation in the general population. Between 1910 and 1950, for example, the state and county psychiatric hospital network expanded substantially in the US and admissions of older adults doubled (Grob, 1987; Rosen et al., 1995). Adults over the age of 65 represented 25% of the total patient population, and state and county psychiatric hospitals were the principal points of care for a population with an assortment of diagnostic, economic, and social attributes (American Psychiatric Association, 1993).

7.20.3.1 Deinstitutionalization of Mentally Ill

In the UK, the move to close large psychiatric institutions and replace them with local, community-based services began with the development of acute psychiatric units which were attached to local general hospitals during the 1960s and 1970s. The continued closure of mental hospitals was underpinned by rhetoric regarding the benefits of community-based care and driven by several Public Inquiry Reports of the scandalously poor conditions existing in such institutions. There were clear financial incentives also. The cost of maintaining and refurbishing the hospitals was growing and they became increasingly inefficient to operate with the falling number of patients. Also, there was an opportunity during the 1980s to transfer financial responsibility for patients from health budgets to the social security budget, through the transinstutionalization of patients from asylums to residential and nursing homes.

The movement of younger patients into the community led to older people forming an increasing proportion of the mental hospital population. Yet, the development of age-appropriate services varied greatly from area to area, often depending on the presence of a charismatic psychiatrist. By 1986, about 70% of the UK population had access to a specialist psychiatrist in old age mental health care (Ineichen, 1989) and developments have subsequently continued. This growth has helped to focus attention on the inadequacy of mental hospitals for older people with dementia or other mental health problems.

Beginning in 1955, older Americans were included in the mass "deinstitutionalization" of psychiatric patients from state hospitals to the community. Publicly supported psychiatric hospitals began an extended process of transferring older patients to the community (i.e., deinstitutionalization) or to other facilities such as nursing homes (i.e., transinstitutionalization). Moreover, state and county psychiatric hospitals applied more stringent admission criteria that reduced older adult admissions altogether (Kahn, 1975). This movement is generally attributed to the discovery of more effective psychopharmacological interventions and to a desire to provide more humane care in the least restrictive setting (Rosen et al., 1995), although cost control undoubtedly played a part. Hospitals were full and state legislatures were faced with the choice of building more or doing something else (Rounds, 1984). Moak and Fisher (1991) reported that 96% of the state hospitals had a floor, unit, or wing designated for geriatric patients. Yet, just 60% of the state hospitals retained a professional staff member who was trained in geriatric psychiatry and less than 20% of the hospitals provided any form of psychosocial treatment. These findings support the American Psychiatric Association's (1993) position that state and county hospitals rely on psychotropic medications and rarely use psychosocial interventions—even though well-designed, multidisciplinary treatment principles have been formulated.

7.20.3.2 Do The Old Institutions Still Exist?

The care of older adults who reside within institutional settings has changed. The image that county asylums and state hospitals provide only chronic, custodial care is contradicted by recent studies. In the areas of the UK where final closure has yet to occur, older patients are likely to be admitted to specialist inpatient assessment and rehabilitation units with a very few continuing care beds. Most of the long-term care admissions occur in areas where the independent sector has been slow to develop nursing home alternatives (Tomenson, Benbow, & Jolley, 1994; Wattis & Fairbairn, 1996).

In the US, Clark and Travis (1994) reported that just 20% of their older adult sample remained in the hospital for more than one year, while most of the rest were discharged. At least 15% of these older patients returned to the community while the remainder were released to another community-based facility such as a foster home, board and care facility, or nursing home (10% died in the hospital).

7.20.3.3 Who Was Discharged from Psychiatric Institutions?

In the UK, the older people affected by the closure of county asylums fit into two principal groups: (i) Chronic patients, usually with a long history of psychosis and having been hospitalized for many years. Arie and Jolley (1982) estimated they formed approximately 20% of the mental hospital population. They have often been discharged to hostels, residential and nursing homes (e.g., Dean, Briggs, & Lindesay, 1993). Anderson and Trieman (1995) suggested that over half this group may additionally have a moderate to severe degree of cognitive impairment. (ii) People with severe dementia—usually admitted because of behavioral problems and mainly discharged to nursing homes (e.g., Lyons, Cohen, Henderson, & Walker, 1995). In some cases, they are discharged to homes that are specially designed for people with dementia, for example, domus homes (Dean et al. 1993; Lindesay, Briggs, Lawes, MacDonald, & Herzberg, 1991; Skea & Lindesay 1996). These are small, "homely" living units, with specially trained staff and a philosophy encouraging attention to psychosocial aspects of care, to balance the traditional emphasis on physical care.

Based on the 1994 State and County Mental Hospital Inpatient Survey, Semke, Fisher, Goldman, and Hirad (1996) suggested that the US state hospital population could be sorted into three prominent groups: older adults with schizophrenia comprise 40.2% of the population, older adults with dementia total 30.0%, and those with depression and other late-onset disorders account for 13.1% of elderly patients. Another 10% were diagnosed with alcoholism, mental retardation, and other psychoses.

In summary, the image of county asylums and state hospitals as custodial warehouses where older patients are admitted and abandoned appears to be out of date. Rather, most hospitals provide limited inpatient services to older adults who are placed back to the community or another facility within a year of admission. In some places, the state or county hospital may be the only point of inpatient service. In considering policies surrounding community care, it is important to recognize that the county asylum and state hospital systems have changed dramatically. These hospitals have emerged as a form of community-based care, especially for older adults affected by psychoses, dementia, and severe affective disorders.

7.20.4 ACUTE PSYCHIATRIC INPATIENT SERVICES

In the UK, many districts have an inpatient old age psychiatry assessment unit on a general hospital site, with varying degrees of collaboration with geriatric medicine units (see Arie & Jolley, 1982; Ineichen, 1989). Typically, these are the central component of a community-based service for older people. Some of these inpatient units provide separately for those presumed to have dementia and those with an affective disorder. The advantages of having the inpatient unit on a general hospital site include increased access to the general hospital, access to specialists for coexisting physical health problems which might complicate the person's presenting state and course, and the opportunity to develop liaison services between mental health and general medical wards.

Over the last 20 years in the US, adult psychiatric inpatients have moved from state and county facilities, and now predominantly use general medical and private psychiatric inpatient hospitals (Kiesler, 1991). Narrow, Regier, Rae, Manderscheid, and Locke (1993) reported that 43.3% of all adult inpatients were admitted to either a general hospital psychiatric unit or scatter bed ("scatter bed" refers to individuals who are admitted to a hospital that does not have a special psychiatric unit or department), while 10.8% were admitted to a private psychiatric hospital. Atay, Whitkin, and Manderscheid (1995) examined the Inventory of Mental Health Organizations survey and, after excluding those older adults who resided in nursing homes, found that 21.6% of all elderly inpatients were treated in general hospitals, and 10.6% were treated in psychiatric hospitals. The remainder were treated in state or county hospitals, Veteran's Administration Hospitals, or by inpatient services affiliated with multiservice community mental health centers.

7.20.4.1 Who is Served in Geropsychiatric Inpatient Units?

Are these the psychotic and demented elderly who were the principal patient populations of state hospitals? Reports from individual hospital programs suggest that this may not be the case. In the UK, Domken, Bothwell, and McKeith (1995) analyzed discharge data from a well-established old age psychiatry unit. Comparison of the years 1985, 1988, and 1991 illustrated the growth in discharges to independent sector homes, together with increased admission rates and shorter lengths of stay. The proportion of patients having a functional disorder without organic impairment fell from 50% in 1985 to 25% in 1991, reflecting increasing reliance on day-hospital and community support for the functional disorders group.

In the US, Conwell, Nelson, Kim, and Mazure (1989) found that of 168 consecutively admitted patients to a geriatric inpatient unit in New Haven, 75% were diagnosed with an affective disorder and 25% with dementia. Knight and Carter (1990) reported that most geriatric admissions to a psychiatric inpatient unit in southern California were diagnosed with psychoses (47%) or with affective disorders (36%), with about 14% diagnosed with an "organic brain syndrome."

In summary, the role of acute psychiatric inpatient care has been overlooked when discussing community-based services, reflecting a tendency to equate inpatient care with custodial institutional care. However, impatient treatment of mentally disordered elderly appears successful using both psychopharmacologic (Schneider, 1995) and psychosocial (Gatz, Popkin, Pino, & VandenBos, 1985) interventions. Moreover, Knight (1983, 1986) reported that the utilization rates of inpatient and outpatient care by older adults in the counties of California were statistically independent of one another, suggesting that the systems serve different community-based populations. Short-term inpatient care close to home appears in fact, to be an important and distinct component of community-based care for older adults.

7.20.5 MENTALLY ILL OLDER ADULTS IN NURSING HOME CARE

In the UK, around 5% of older adults live in permanent institutional care. The majority (approximately 72%) of these older adults in 24 hour care live in residential or nursing homes which are not provided by the National Health Service (NHS), housing authority, or by local government. Many of these people are funded, at least in part, by a combination of Social Services and Social Security payments. The high prevalence of psychiatric morbidity in residential homes in the UK is well established. For example, Neville et al. (1995) assessed older people in four local authority residential homes. The great majority (82%) had some degree of dementia. Over half (53%) of the residents were assessed as having severe dementia. Of those (45%) able to complete a self-report depression scale, 17% scored in the depressed range. Of the whole sample, only 17% showed no evidence of dementia or depression. The rate of depression reported here is much lower than in previous studies (see Ames, 1991), where over a third of residents without severe dementia have been depressed.

The lower prevalence of depressed older adults in residential care may be due to improved assessment as well as more creative community treatment and support packages which enable older people with depression to remain in their own homes. An alternative hypothesis would be that the closure of hospital beds for severe dementia places pressure on local authority homes to accept a greater number of people with severe dementia, while independent sector nursing and residential homes—who have more choice over which residents to accept—absorb the less physically dependent and behaviorally disturbed group of depressed older adults. This illustrates the need to examine trends across the multiple systems (health, mental health, social) of institutional care. Examining one component in isolation gives an incomplete, and perhaps misleading, picture.

In the UK, the National Health Service and Community Care Act (1990) was motivated by the need to control the flow of resources into residential and nursing homes; previously, funding had been available without assessment of need. Although wrapped in rhetoric about providing better community support and services, assessments of need form the basis for targeting the provision of institutional care (see Gilleard, 1996a). "Need" for continuing hospital care is more likely to depend on availability than on any rational assessment; some districts have very few hospital beds, others retain a significant number (Tomenson et al. 1994).

Between 1963 and 1990, nursing homes expanded tremendously in the US from 445 000 to 1 772 000 individual beds as Medicare and Medicaid were initiated and included financial support for long-term care services (Darney, 1994; Schick & Schick, 1994). Moreover, the nursing home population grew as patients were supplied by the transinstitutionalization of older adults from state psychiatric hospitals (Kahn, 1975), a move which represented a cost-cutting strategy for state legislatures who moved the elderly mentally ill from totally state-funded psychiatric hospitals to partly federally-funded nursing homes (Estes, 1995).

In 1963, 48% of institutionalized mentally ill elderly were in state psychiatric hospitals with the remainder in nursing homes; in 1969, 75% were in nursing homes (Dorwart & Epstein, 1993). The nursing home setting would appear to be where many of the former psychiatric hospital population would now receive care: 50% of nursing home residents are demented, 30% have psychoses (Shadish, 1989).

More recently, the US Federal Government responded to the relocation of older mentally ill with the Nursing Home Act of 1987. This act recognized that older adults with psychiatric conditions were over-represented in nursing homes and stipulated that nursing homes must screen patients for psychiatric illness. If patients were diagnosed with a mental illness, then the nursing home was required to (i) provide appropriate psychiatric care, or (ii) deny admission to them and/or transfer them to a more appropriate facility (Smyer, 1989). Smyer, Shea, and Streit (1994) reported that at the time of the passage of this law, 65% of older adults in nursing homes who had mental health problems did not receive treatment for these problems. They also reported that for-profit homes tended to rely on bachelor's level counselors and government homes used psychiatrists, whereas private nonprofit homes used psychologists and social workers as well as psychiatrists. Early estimates suggested that between 2.4% and 18% of residents would be excluded from nursing home care by this act (Freiman, Arons, Goldman & Burns, 1990; Eichmann, Griffin, Lyons, Larson, & Finkel, 1992). However, later reports suggest sharp increases in mental health service provision in nursing homes after the passage of this law (Knight & Kaskie, 1995; Office of Inspector General, 1996), rather than discharges of patients with mental health problems. However, there are concerns about service standards and fraud: approximately 24% of Medicare billings in 1993 for mental health services to nursing home residents were estimated to be medically unnecessary (Office of Inspector General, 1996). Two-fifths of the insurance carriers for the Medicare program considered part of the problem to be "excessive entrepreneurialism" on the part of mental health professionals (Office of Inspector General, 1996, p. 13).

On the whole, the issue of providing adequate treatment to the transinstitutionalized older mentally ill resident of nursing homes in the US is still unresolved. For these individuals, "return to the community" has meant a shift from the mental health system to the long-term care system, reduced access to mental health treatment, a shift of care from the states to private sector institutions, and a shift of costs from states to a mix of federal-state (Medicaid) and private (patient and family) financing.

Lomranz and Bar-Tur (see Chapter 21, this volume) provide more comprehensive coverage of the nursing home population and service models for mental health treatment. Our point is that many, probably most, of those moved to "the community" were transferred to nursing homes from state hospitals, and that long-term psychiatric hospital treatment and care has been replaced by such homes in both the US and UK. In the remainder of this chapter, we explore community-based outpatient mental health services for older adults.

7.20.6 OLDER ADULTS IN COMMUNITY MENTAL HEALTH SERVICES

In this section, we cover the provision of mental health services to older adults outside of 24-hour care institutional environments. The range of services and settings described includes traditional and nontraditional outpatient services, partial hospitalization and related day care services, respite care, and support for caregivers.

7.20.6.1 Community-based Mental Health Outpatient Care in the UK

All community-based services in the UK operate in the context of Department of Social Services community care assessment and care management, intended to ensure that needs are properly assessed and that service provisions are targeted appropriately. Challis (1992) reviewed good models of providing community care through care management schemes. Some success is reported with mental health problems as well as with physical frailty. They have tended to include older people with dementia, rather than to focus specifically on them. Although apparently intended to be the basis of UK community care provision, important elements of the demonstration programs have often not been present in mainstream practice. The key elements included a high degree of flexibility, an available budget per person to "buy in" services, availability of paid volunteers, and a skilled assessment of need. There is widespread disappointment with the very basic packages of care being offered, and there is concern to ensure that specialist mental health services are involved more fully in assessment and care planning (Department of Health, 1996).

7.20.6.1.1 Multidisciplinary teams

The key element of such specialist community services in most areas of the UK would be a multidisciplinary team (MDT) usually seeing the older person in his/her own home. Most teams aim to respond quickly—within a few days—to referrals, more rapidly for urgent cases. The range of disciplines involved varies greatly; at a minimum, most members may be community psychiatric nurses, perhaps with an old age psychiatrist having a part-time input to the team. Elsewhere, occupational therapists, physiotherapists, social workers, speech therapists, clinical psychologists, and workers from voluntary agencies for older people may be team members. Typically the teams have been established by redirecting resources from mental health hospital provision; most teams liaise closely with inpatient old age psychiatry units and day hospitals. Often the team will share a hospital base with these services, although increasingly efforts are being made to locate such teams more accessibly, for example, in primary health care centers.

Most teams have a defined population to serve, and will usually see clients referred from residential and nursing homes within this area, as well as those living at home (Brown, Challis, & von Abendorff, 1996). Many teams will accept referrals from any agency, or from families or self-referrals; other teams maintain the traditional, less accessible, stance of only accepting referrals from the person's primary care physician. Some teams insist that all referrals are seen by a medical member of the team; others have a policy that a team member from any discipline may carry out the initial assessment meeting, where medical input would be available.

Coles, von Abendorff, and Herzberg (1991) evaluated the initial impact of a new community mental health team in an Inner London borough. This service replaced a traditional hospital-based system in which a consultant psychiatrist made domiciliary visits. The new community team was multidisciplinary, and used a keyworker scheme, where one member of the team would be allocated primary responsibility for the coordination and delivery of the service provided to each case. In a period of six months the team assessed over 90 more referrals than the consultant had seen in the equivalent period. The primary diagnosis for both the MDT and the consultant was dementia, which made up approximately 40% of all referrals. Depressive disorder accounted for 30.3% of diagnoses by the MDT and 18.3% of diagnoses by the consultant. The team was able to provide increased levels of input including treatment for depression, coordinating a range of services, and counseling and support for caregivers.

Junaid and Bruce (1994) investigated the effectiveness of community psychiatric nurse services for the elderly in a particular district. They found wide variation in the numbers employed per thousand people over 65 (from one per 2255 to one per 9200; The Royal College of Psychiatrists recommends one nurse per 5000 of the population over 65). They found large differences in the skill mix of the community teams. In some teams nurses would specialize in patients with specific diagnoses, for example, cognitive impairment, neurosis, or psychosis. Other teams would share case loads equally. In conclusion, Junaid and Bruce stated "Current provision appears to be based primarily on historical trends of resource allocation, and also on a varying commitment to community development which is evident in the philosophy or approach of the teams" (p. 718). It would seem that while the development of MDTs may have led to an improvement in services at a local level, the variation among teams' working practices and funding allocations has led to wide regional and inter-regional differences.

In the UK, some mental health teams (see Dening, 1992) specialize exclusively in working with people with dementia, whereas others cover all mental health problems of older people. Brown et al. (1996) provide details of referrals, caseloads, contact history, and outcomes of two well-established and pioneering teams which work with both organic and functional disorders. They took a random sample of 120 open cases on caseloads of two MDTs for elderly people, both with an open access policy, and followed up these cases over an 18-month follow-up. The teams' caseloads contained a greater proportion of patients with affective disorders, and less with dementia, in comparison with referrals to the team. The response to dementia appeared to be a briefer episode of contact, emphasizing assessment and diagnosis more than continuing support. A higher proportion of those with affective disorders remained at home at follow-up. This tendency for people with dementia and their caregivers to be rapidly discharged from the active caseload has led these teams to explore the use of specific long-term workers attached to the team to act as case managers for people with dementia and their families. Elsewhere, an innovative response has been the development of a specialist mental health nursing service to work specifically with family caregivers (see Woods, 1995), or teams working only with people with dementia, offering assessment, ongoing support, interventions to relieve caregiver strain, and facilitation of appropriate services.

Outcome studies of the work of such teams with people with dementia are yet to be completed. O'Connor, Pollitt, Brook, Reiss, and Roth (1991) have reported a controlled trial of the effect of an MDT on admission rates of people with dementia over the age of 75 over a two-year period. The team offered a wide range of help to people with dementia and their families identified in an ascertainment survey—including advice on and help in arranging services and benefits and respite admissions. No difference in admission rates, compared with control patients, was noted in those living with a caregiver. Those living alone were much more likely to enter residential care than controls. However, this study does not necessarily reflect outcomes in usual practice where the person or their family seeks help, rather than being identified as someone in need, as was the case here. More positive results are reported by Hinchliffe, Hyman, Blizard, and Livingston (1995) in a study where caregivers were offered an intervention devised by an MDT, which included components of behavioral, practical, and pharmacological input depending on the assessed needs of the situation. Caregiver mental health and the patient's behavioral difficulties improved compared with a waiting-list control group.

Although around 10% of the teams' caseload in Brown et al. (1996) had a diagnosis of schizophrenia, it is not clear what proportion of these had developed a psychotic illness such as paraphrenia in late life, as opposed to having a long-standing mental health problem. Generally, patients who have been receiving mental health services for many years are not transferred to old age services when 65, unless these services are more appropriate (e.g., because of the development of a dementia or physical health complications). Little has been documented regarding the effects of this approach.

In summary, community teams have developed widely, with very little in the way of evaluation, or even detailed descriptions of practice, since the early publications (Lindesay, 1991). Their service coordination role has, to an extent, been taken on formally by Social Services through the community care assessment and care planning process, and there are significant differences between areas in the extent to which the skills of the MDTs are utilized in carrying out specialist assessments and planning programs of care for older people with mental health problems. Some teams have very close links with social worker team members, and carry out many community care assessments; others find themselves much less involved, at best acting as an advocate for the patient's service needs. Although diversity of team functioning can be seen as flexibility to respond to local circumstances, the lack of consistency makes it difficult to know whether such teams do offer an effective way of organizing the response to mental health needs of older people in the community. Most teams attract high levels of referrals, and so appear to be meeting a demand, but this is of course an inadequate measure of efficacy. Two further developments—Memory Clinics and Family Therapy Clinics—in the UK are worthy of mention. Although their implementation is not yet as widespread as community teams, they promise to significantly shape the direction of future services.

7.20.6.1.2 Memory clinics

Early diagnosis of memory disorders is regarded as fundamentally important when treating and counseling individuals with dementia. The memory clinics have been developed to provide community-based services to older adults with dementia. All memory clinics provide an outpatient diagnostic assessment, and, in many centers, advice and counseling is offered to the person and their family, along with any appropriate treatment.

The first British memory clinic was established in 1983 at St. Pancras Hospital, London. A recent review of memory clinics in the British Isles (Wright & Lindesay, 1995) identified 20 clinics running in the UK and Ireland. Memory clinics provide diagnostic assessments for dementia and other memory disorders. Typically, patients receive a cognitive assessment involving standardized tests of cognitive functioning (including the Mini Mental State Examination, MMSE; Folstein, Folstein, & McHugh, 1975) and mood. A wide variety of physical investigations are used including full physical examination, computerized tomography scanning, magnetic resonance imaging, and blood tests. On average, around three-quarters of referrals to the clinics were diagnosed as having a dementing condition. Clinics vary in the cognitive tests and physical investigations used. The time taken to carry out full assessments was reported to vary from one to six hours. It is not clear whether the differences in practice have any clinical impact.

The number and type of staff running memory clinics vary. The professionals commonly involved are psychiatrists, geriatricians, neurologists, psychologists, and nurses, as well as speech therapists, social workers, and occupational therapists. Staffing levels vary as does the amount of any specific type of professional input to the service. Many of the memory clinics in the UK are actively involved in recruiting participants for research into dementia and specifically Alzheimer's disease. The research is both pharmaceutical and nondrug orientated.

Memory clinics provide an important service within the communities they serve. They are a useful point of referral for general practitioners, and may provide a starting point for many individuals to find help and information about memory impairment, as well as access to other community services. In a few areas, the follow-up from the clinic is extensive and caregivers receive considerable support over an extended time period (Bayer, Richards, & Phillips, 1990; Gilliard & Wilcock, 1993).

7.20.6.1.3 Family therapy clinics

Family therapy for older adults has only a brief history of implementation in the UK, although old age mental health service providers have long recognized the role of the family as being a significant factor in the psychological and physical well-being of older adults (Colwell & Post, 1959). UK research suggested that family-based work with older adults had useful clinical implications (Ratna & Davies, 1984). However, it is only in the last few years that family therapy for older families has begun to be implemented in practice. Traditionally, families with young children or adolescent members have been the prime focus for family therapy practice. In 1992, Gilleard, Lieberman, and Peeler set up a team for older adults in an established family therapy center in Inner London. The service was underutilized and referrals soon began to "dry up." Research revealed that professionals working with elderly people were unaware of the existence of the team, or had limited knowledge of family therapy with older people in general. Some professionals expressed doubts as to the usefulness of such a service, believing that there was little potential for change in the families of older people. Although misconceptions about older people did have some bearing on the lack of referrals to the new service, practical problems also played a large part. Professionals (especially community nurses and social workers) did not know that they were able to make referrals.

The overwhelming majority of professionals working with elderly people expressed positive attitudes towards family therapy. They suggested ways of making the service more accessible to their particular client groups, by limiting the amount of traveling and by holding all or some of the sessions in the client's home. Altering times of sessions to accommodate working hours was also suggested. By disseminating information regarding the service and by implementing changes suggested by health professionals, family therapy with older adults may become an important resource available to the community.

Research has highlighted the efficacy of family therapy to different elderly client groups (Gilleard, 1996b). Benbow, Marriott, Morley, and Walsh (1993) assessed the usefulness of family therapy with older adults suffering from dementia. The basic therapeutic techniques remained the same as for those older adults with other mental health problems, however, a few differences were found. Family therapy based treatment with dementia patients was often seen in terms of psychological management of the disorder and an increased focus on caregiver stress. Families were crisis-oriented and attended for fewer sessions. These difficulties present new challenges in providing family therapy for dementia sufferers. Continuing development indicates that family therapy with older adults will prove an important addition to older adult mental health services in the UK.

Curiously, family therapy for the elderly has grown slowly in the US. While there are services for caregivers (see Chapter 22, this volume) and well-known advocates for family therapy with the elderly (e.g., Qualls, 1996), there is no effort to establish family clinics as such.

In summary, for the foreseeable future, community MDTs seem likely to form the basis of specialist community mental health services for older people in the UK. Developments which reflect the importance of careful assessment and working with families, such as memory clinics and family therapy clinics, are likely to augment the basic service in some areas, and to influence practice within the teams more broadly.

7.20.6.2 Community Mental Health Outpatient Services in the US

In the US, the Community Mental Health Centers (CMHC) Act of 1963 financed the development of comprehensive community mental health centers. These centers were intended to provide services to the deinstitutionalized mentally ill and to provide services to persons who would have been institutionalized in the past. CMHC services target state hospital patients who do not require intensive supervision and CMHCs also serve community residents who need intermittent care (US Congress, Senate Committee on Labor and Welfare, 1973). CMHCs provide comprehensive mental health services including individual assessment, diagnosis, and treatment in acute inpatient, day treatment, or outpatient settings. Unlike the UK approach, the US CMHCs were newly created agencies rather than programs reaching out from existing hospital services.

As noted in Section 7.20.2, the focus on community-based services for the elderly did not begin in the mental health system until shortly before the national initiative was terminated and responsibility was returned to the states. Perhaps for this reason, the efforts at national planning of community mental health services for the elderly that are seen in the UK are lacking in the US. However, there are model programs that can provide guidance in conceptualizing what community mental health services could offer older adults.

Knight, Rickards, Rabins, Buckwalter, and Smith (1995) reviewed a variety of mental health programs for older adults with a range of mental health problems. The programs included community mental health services in Spokane, Washington, Cedar Rapids, Iowa, and Ventura County, California, as well as a multiservice program at Long Island Jewish Hospital in New York and a psychiatric consultation service to a federal housing program for the elderly in Baltimore. The services spanned urban and rural areas and used a variety of approaches to the coordination of care. These programs have shown that it is possible to provide high-quality, community-based mental health services to the mentally ill elderly. The programs also share some common features: (i) All emphasize accurate diagnosis of older adults. (ii) All are interdisciplinary and treatment focused. (iii) All use active case-finding methods and community education approaches to bring clients in. They also collaborate actively with other agencies that serve older adults and in three cases with postal carriers, meter readers, and other "community gatekeepers" (Raschko, 1990). (iv) All of them deliver mental health services to older adults at home. Mobile geriatric outreach teams, which use home delivered services and interdisciplinary teams (see Chapter 24, this volume) have proven effective in increasing service use by older adults in both urban (Knight, 1983) and rural areas (Buckwalter, Smith, Zevenbergen, & Russell, 1991). Mosher-Ashley (1994) also has highlighted the importance of home delivered services in reaching this population.

7.20.6.2.1 What types of clients are served?

In a survey of 349 older adult CMHC clients in Miami, Pruchno et al. (1983) reported that 54% were diagnosed with a unipolar or bipolar mood disorder, 21% had dementia, and 8.9% had schizophrenia or psychoses. Speer, Williams, West, and Dupree (1991) determined that 29% of their Tampa sample was diagnosed with dysthymia, 20% had mood disorder, and 10% had maritial problems. If these local samples are representative, it would appear that CMHCs serve high proportions of depressed older adults, but may be less successful in reaching the demented and psychotic populations that were a large part of the state hospital population.

7.20.6.2.2 Were CMHCs effective in outreach to elderly?

While the CMHCs did not develop sufficient services for older adults with mental illness (Lebowitz, 1988), they did not completely overlook older adults. In fact, the use of CMHC services by older adults grew steadily through the 1970s and 1980s. The Government Accounting Office (1982) reported that 4% of all CMHC clients were over the age of 65 in 1978, whereas by 1983 the elderly comprised 6% of the users of community mental health (Fleming, Rickards, Santos, & West 1986), with lower representation in other clinics and fewer still in private practices. In the 1990s, more of the elderly have begun to use mental health services. In 1990, the Inventory of Mental Health Organizations showed that 9.72% of all outpatient mental health clinic clients were older adults; 11.87% of clients in publicly supported clinics were older adults (Stiles, 1994).

Lebowitz, Light, and Bailey (1987) gathered information from 45% of the 624 CMHCs and found that 52% of those responding to the survey provided both professional mental health services to older adults and employed a geriatric staff specialist. However, 28% of their sample had neither age-targeted services nor a staff specialist. They concluded that older adults were offered more services by those CMHCs with trained staff or specific programs. In these centers, the older adult population represented 15.8% of the entire population; among the CMHCs with no staff or services, the older adult census averaged 8%. Also, Knight (1986) reported that for county mental health programs in California, service use by older adults was predicted by the number of specialist staff and was uncorrelated with percentage of elderly in the county population.

7.20.6.2.3 What comes next?

In the mid-1980s, Fox, Swan, and Estes (1986) reported that more than 60% of CMHCs were maintaining or increasing their services for older adult populations. More recently, they reported in a follow-up study that older adults services are more likely to be dropped than added to community-based mental health programs (Estes, Linkins, & Kane, 1994). Since the control of mental health programs was shifted back to the states, they have increasingly left individual CMHCs with a reduced budget, and without any stipulation concerning how the money should be spent (Christianson & Gray, 1994). In the mid-1990s, California continued the displacement of responsibility from federal to local governments by shifting control of mental health programs from states to counties, where these services compete with the medical care system for the indigent as well as with police and fire services for funding.

It seems unlikely that comprehensive CMHCs will be a major part of mental health services for older adults in the future. As mental health services in the US move to the private sector and to managed care organizations, active casefinding and other community mental health outreach strategies are likely to be inverted with probable negative impact on the accessibility of mental health services to older adults with depression, anxiety, substance abuse, and especially with long-term mental disorders like schizophrenia.

7.20.6.3 Summary: Community-based Outpatient Services

At this level, outpatient mental health services within the UK and US diverge most markedly. The UK's history of universal coverage has led to relatively systematic programming for the elderly, including MDTs, and more recently in some areas memory clinics and family therapy clinics. In the US, the federal role in mental health services has been more research-oriented than services-oriented with the individual states playing the major role in mental health services provision (Rosen et al., 1995). The history of federal–state relationships regarding mental health in the US has been characterized by shifting responsibility back and forth between the two with neither level of government taking clear responsibility at present. As noted by Liebig (1993) with regard to health and social policies for the aged, the federal system tends to lead to wide variation in policy among states, and in some cases the absence of policy. In the mental health sector, this is perhaps most clearly the case with dementia care, which some states would include in mental health services, while others would define it as in a different system, as its own system, or as not a state responsibility (Knight & Macofsky-Urban, 1990). The US has a few model programs for community-based outpatient services for the elderly, but lacks a system or a coherent national policy.

7.20.7 PARTIAL HOSPITAL SERVICES

7.20.7.1 British Models

In the 1970s and 1980s, day hospitals developed as the central component of many local older adults mental health services. By 1990, 20 000 older people were attending such day hospitals regularly, with over 1.7 million day places available per year (Fasey, 1994). Day hospitals have served a variety of functions: preventing admission to inpatient facilities, providing relief to caregivers of people with dementia, assessment, monitoring, and treatment of people with depression, facilitating discharge from an inpatient unit back to the community, and for observing the person's behavior and mental state for a longer period than is possible in the person's own home.

There is considerable debate concerning how effective the day hospital is in fulfilling these functions, and whether it represents the most effective use of resources (Fasey, 1994; Howard, 1994; Murphy, 1994). There is relatively little in the way of evidence to inform the debate. Day hospitals have not been effective in delaying admission to long-term care (Diesfeldt, 1992; Woods & Phanjoo, 1991), but do reduce strain for many (but not all) caregivers (Gilleard, 1987). However, some caregivers find the strain of getting the person ready for day care an additional burden, or find that most of their difficulties are at night and so they are not helped by day care. A few "night hospital" programs have been developed in response to this need. The role of day hospitals in the treatment of depression has not been evaluated in the UK context.

The UK also makes the distinction between "day hospitals" and "day centers," the former are health facilities, with a team including doctors, nurses, and therapists, and the latter are primarily "social care" resources but may have input from visiting health professionals. In reality, the distinction is not always clear, and those opposed to day hospitals argue that resources spent upon them are better used to provide more health input to day centers, many of which work well in providing relief for caregivers or as a social base for older people with depression needing to expand their social network.

Although Tester (1989) argued that day hospitals generally provide respite care for those older people whose behavioral problems or level of dementia proves "too severe" to be managed in a day center setting, Warrington and Eagles (1996) reported from a specific sample that the overall level of behavioral disturbance was no greater in cognitively impaired day hospital attenders than in those attending day centers.

7.20.7.2 American Models

Partial hospital programs have expanded significantly in the US in the last 10 years and have also become a more common form of mental health service for older adults. These partial hospital programs are organized under many different models that offer a range of social and medical services (Whitelaw & Perez, 1987).

Culhane and Culhane (1993) reported that 63% of day treatment programs in the US are located in CMHCs, 17% are affiliated with general hospitals, and 12% are found in private psychiatric hospitals. As fiscal incentives promote alternatives to inpatient care, partial psychiatric services seem to be expanding rapidly, especially in the private sector. For example, Stiles (1994) reported that partial hospital services increased by 50% between 1986 and 1990. Culhane and Culhane (1993) added that older adults represent between 4.8% and 7.9% of all partial hospital patients. The American Psychiatric Association (1993) noted that partial hospitals serve 18% of those older adults who actually use professional mental health services.

Since the history of partial hospitalization is directly related to substitution for hospital-based care, it is perhaps not surprising that the population served is more similar to the state hospital population. Eskew, Sexton, Tars, and Wilcox (1983) indicated that 34% of partial hospital patients were diagnosed with dementia, 32% had an affective disorder, and 24% were schizophrenic. The patient population appears to vary by program though; Plotkin and Wells (1993) reported that 60% of their partial hospital sample were depressed, 13% had a cognitive impairment, and 9% had an adjustment disorder. The majority of these patients received treatment subsequent to an inpatient hospital stay. In a program designed and implemented at the Florida Institute for Mental Health, Roger Patterson and associates designed a behaviorally-oriented social skills training approach to day treatment which assisted chronically mentally ill older adults to return to the community from state mental hospitals. As described and evaluated in Patterson et al. (1982), the program used applied behavioral analysis, behavioral intervention, and psychotropic medication to overcome effects of chronic psychosis and chronic institutionalization. Program attenders and graduates were more successful in community placements with lower recidivism rates.

7.20.7.3 Summary

The day program model has been used quite successfully for older adults in both the UK and the US and appears to be well suited to the needs of the older adults who formerly used long-term institutional care. However, this modality of treatment has neither been extensively reported in the clinical literature nor evaluated. We would hope that it would receive more attention in future research, clinical reports, and public policy consideration.

7.20.8 RESPITE CARE

As developed in the UK, respite care generally means a short stay in hospital or in another 24-hour care setting. In the US, respite has been used to refer to community-based social services such as social day care and in-home care workers. In the UK, respite can be considered as another partial hospitalization program. As noted above, this may be one of the functions of a day care program. Often a rationale will be made as to why a particular patient "requires" respite in a hospital rather than a residential or nursing home, in terms of severity of disturbance, or of need for nursing care. The availability of resources is of course an over-riding factor in these considerations, but the need to support caregivers in this way is widely recognized.

Residential respite care involves the patient staying in a residential/nursing home/hospital setting for a given period of time, thus allowing their caregiver to have a sustained period of time off from caring. Some beds are designated as "rota beds," in that a number of patients share the resource, taking turns to come into the facility. For example, four patients might each have two weeks in hospital and six weeks at home on a rotational basis, filling one bed continuously. Some relief schemes do not run on a rota system, but allow admission over a period when the caregiver may wish to go on holiday. Relief care may also include crisis admissions, say when a caregiver is taken ill.

Nolan and Grant (1992) carried out an extensive evaluation of such a hospital rota bed scheme, concluding that respite care fulfills a deeply felt need in caregivers, but also raises issues about caregivers' attitude towards institutional care. Caregivers possess unique knowledge of their relative, and often feel that the needs of the care recipient could not possibly be met by those working in an institution who have only limited knowledge of the care recipient (Worcester & Quayhagen, 1983). It is interesting to note that although this type of respite care is generally reported as being beneficial to caregivers, they also report feeling sad and lonely when their relative is away (Pearson, 1988). Guilt and worry may also be pervasive emotions felt by a caregiver who places a relative in residential respite care (Nolan & Grant, 1992).

Evaluations of respite care have shown that it has considerable benefit for caregivers. Levin, Moriarty, and Gorbach (1994) found respite was valued highly by caregivers. When respite services were introduced, they formed an ongoing source of support. Levin et al. emphasize the need for services to be targeted to the individual, noting that one "care package" may not suit all caregivers' needs. Flexibility of service is important, especially where disease progression is an element.

Some caregivers report that the person with dementia returns from respite more confused than when he/she was admitted. This may be understood in terms of the effects of a change in environment and routine on the person with dementia, and the difficulty of staff in ascertaining quickly the person's needs. Preadmission visits to a familiar place might reduce such effects, but an alternative strategy is to provide the respite care in the person's own home to minimize the disruption for them. This service is referred to as domiciliary respite where an outside agency will dispatch trained caregivers (i.e., a "sitter" or care attendant) to the home to take over the role of the caregiver for a given amount of time—a few hours or several days. Some care workers may even spend the night at the person's home, as this is often the most difficult time. There are some indications that caregivers find these services beneficial (Turvey & Toner, 1991).

In the US, day care and in-home respite services are not typically viewed as a part of mental health service systems but rather as a part of community-based long-term care or of home health services. Curiously, these services seem more effective at meeting mental health goals (i.e., improving the caregivers' life satisfaction or reducing emotional distress) than meeting community-based services goals of reducing care costs or preventing institutional care (Weissert, Cready, & Pawelak, 1988; Knight, Lutzky, & Macofsky-Urban, 1993). However, it has long been noted that respite may not be provided in sufficient quantity to meet these other objectives, and recent research tends to support the notion that longer respite may lead to reduced institutional care (Koslowski & Montgomery, 1995)

7.20.9 CAREGIVER SUPPORT

In the UK, caregiver support groups have been a popular development from the early 1980s. They take many different forms: some are organized as "training courses"; some are akin to therapy groups, being closed and having a facilitator; many are open-ended, self-help groups. Some are organized by voluntary agencies (especially the Alzheimer's Disease Society); others are based in day centers, day hospitals, residential homes, and primary health care centers. Few evaluations have been reported of their impact, although they are positively regarded by the caregivers attending. Morris, Woods, Davies, Berry, and Morris (1992) showed no differential impact of a coping skills-based group compared with a conventional course with invited speakers on strain and depression, although knowledge and use of coping skills did increase in both groups.

Services for caregivers in mental health settings have been a prominent feature of programs for older adults in the US since 1980. In a meta-analytic summary of program evaluation literature in the US, Knight et al. (1993) found that individual psychosocial interventions with caregivers appeared to be more effective than group interventions. Fully randomized evaluations of caregiver interventions have provided mixed results regarding the effectiveness of either individual or group interventions, a finding that is more curious given the effectiveness of psychological interventions with other groups of older adults (Scogin & McElreath, 1994).

7.20.10 MENTAL HEALTH SERVICES IN THE HEALTH CARE SYSTEM

In the UK, most mental health problems (including those of older people) are managed in primary care, without recourse to specialist mental health services. It is important to note that mental health problems are common in patients admitted to geriatric units (i.e., those designated for physical health care). Burn, Davies, McKenzie, Brothwell, and Wattis (1993) reported that 43% of such admissions who could be assessed had a diagnosable psychiatric condition (30% were depressed or anxious); many others could not be assessed because of a severe cognitive impairment. Conversely, significant physical health problems are also common in those admitted to old age psychiatry units (34% in one prospective study; Perry, Milner, & Krishnan, 1995).

In the UK, the problems center on the inaction of primary care physicians rather than the lack of recognition of mental disorders in older adults by physicians, which is more common among physicians in the US (Rapp, Parisi, Walsh, & Wallace, 1988). Jenkins and Macdonald (1994) indicated that general practitioners do recognize depression in their older patients, but rather rarely respond. A randomized controlled trial of depressed patients referred from primary care to a multidisciplinary team did not show any significant benefits—but the sample was small; the authors concluded that male patients might benefit from being referred more frequently, as their outcomes were slightly better.

Blanchard, Waterreus, and Mann (1995) achieved slightly more positive results. They identified older adults with depression from a community screening survey. The older patients were randomly allocated to intervention from a primary care nurse or to "usual care" from the general practitioner. Interventions were individually designed by a MDT, and included increasing the person's social network, antidepressant medication, counseling, and the involvement of relatives. Depression scores improved in this intervention group.

General practitioners in the UK refer people with dementia to specialist services at times of crisis, or when the dementia is well advanced—hence the need for open-access services described above. Wolff, Woods, and Reid (1995) have shown that GPs mainly saw dementia as a social rather than a health problem, and felt that early referral was unhelpful. They felt that they had little to offer patients with dementia, and this sense of helplessness may contribute to the failure to refer. Primary care teams are now encouraged to carry out an annual screening of individuals over the age of 75. This normally includes a rudimentary assessment of mental state. Consequently, the awareness of patients with dementia in the primary care context could grow.

In the US, there has been an increasing recognition that the health care system in general, and physicians' offices in particular, are a *de facto* part of the system of mental health care in the community. Analyses of service utilization data by George, Blazer, Winfield-Laird, Leaf, and Fischbach (1988) call attention to the fact that adults of all ages with a mental health problem are more likely to go to a health care provider than to a mental health professional. This tendency to seek mental health care from primary care physicians rather than from mental health professionals also appears to increase with age after the age of 55.

What happens during these visits is less clear. Physicians have been shown to be poor at detecting depression in medical impatients (Rapp et al., 1988) and outpatient service provision does not appear to be any better. Katz (1996) reported that primary care doctors often misdiagnose or overlook depression among older adults—one survey revealed that 26.4% of primary care physicians never evaluate the emotional status of individuals over the age of 75 (Mann, 1995). Further, primary care physicians tend to treat mental health problems only with psychotropic medications (Burns & Taube, 1990). These physicians are not likely to conduct clinical histories, involve family members with treatment, or to use psychosocial interventions.

Haley, Salzberg, and Barrett (1993) reported that geriatric specialists in medicine tend to be more open to psychological interventions and to working with psychologists as part of the care team than primary care physicians without geriatric training. Haley (1996) has also argued for the importance of staffing medical clinics with psychologists as part of the outreach effort to assure that older adults in need receive adequate services.

In short, it appears that on both sides of the Atlantic, mental health services for older adults are moving into the medical care system. While the move is conceptually unassailable and has the potential to improve access to mental health services and the coordination of mental health and medical services, short-term results suggest that the result may be decreased access to mental health services due to ineffective screening and intervention by primary care physicians. The multiple systems perspective would suggest that the medical system will, of course, recognize and treat medical disorders rather than mental health disorders and so the elimination of the separate mental health system may be undesirable.

7.20.11 DISCUSSION

7.20.11.1 Cross-national Systems Comparisons

The contrasting histories of mental health service systems in the UK and US are reflected in much of our previous discussion. While both systems have seen the demise of the large institutions, in the UK a long-term tradition of universal health care coverage and parity of coverage between mental health and health services created an environment in which old age psychiatry has developed as a new and influential speciality in psychiatry. New-style, community mental health services have been established with the resources released by hospital closure. However, as Murphy (1994) points out, not all the services that the asylums offered have been replaced. Extensive social support, health care, and social care resources have not been developed extensively. In the UK, institutional care still absorbs most of the available resources, despite preferences of families and older people and the stated aims of services.

In the US, community-based programs for older adults have never been well-funded and continue to face major threats to funding. Historically, federal policy in the US has placed restrictions on mental health services as supported by Medicare and Medicaid health insurance programs. These restrictions are beginning to be eliminated in the 1990s, but it is too early to assess the impact of increasing Medicare reimbursements to more providers or of recent moves toward parity in mental health insurance reimbursement. In summary, both nations have been more successful in closing down the institutions than in creating substitute systems in the community.

Research in both nations has focused on dementia and depression (rather than anxiety, alcohol abuse, or older schizophrenics) and on outpatient services rather than on inpatient or partial hospital services. These tendencies need correction to provide an empirical basis for comprehensive services in clinical geropsychology.

7.20.11.2 Dementia Services

The nature of services for older adults with dementia also differs. In the UK there is considerable focus on services to demented people themselves within the mental health sector with active concern for improving the well-being of the person with dementia. Kitwood (1996) provides a timely reminder of the "personhood" of the patient with dementia—psychological approaches to dementia must entail ensuring that the full human needs and potential of older people with severe cognitive impairments are recognized and nurtured. He contrasts his approach with the "neurological determinism" of the traditional model of working with the demented. The US, perhaps caught up in the "neurological determinism" viewpoint, has tended to conceptualize services for dementing elderly as services for the family caregiver and often places dementia services outside of the mental health sector: in social services, aging services, and so forth.

7.20.11.3 Mental Health and Primary Care

In the UK, mental health services for older people are likely to be provided in primary healthcare settings (see Mann, 1995) where most older people with mental health problems are actually found. A new relationship with the independent sector providing consultation outside the traditional NHS and Social Services settings needs to be established. This is a remarkably difficult change for clinicians brought up to feel that the NHS should provide care "from the cradle to the grave," at no cost at the point of service delivery.

In the US, the same change is likely to occur, driven in large part by the increasing organization of both health and mental health care into various types of managed care networks of providers. In the US, where the private sector is the traditional provider of health care and of much outpatient mental health care, resistance to these arrangements has more to do with the traditional independence of mental health providers and the differing organizational cultures of medical and mental health care. The latter reason has been offered as an explanation for the tendency to rely on separate managed care provider organizations for mental health services (Wholey, Christianson, & Peterson, 1996).

Indeed, as both nations move toward versions of managed care, primary care physicians will assume increased control over the provision of health and mental health services. For different reasons, this change may not bode well for access to required services since primary care physicians have not shown a great willingness to refer older adults for required mental health services (the UK) or even to recognize the need for such services (the US). In both cases, the problems may stem from the distinct world views of physical medicine and the mental health disciplines, a gap which must be bridged in order to effectively serve older adults. Creating the required levels of communication and collaboration will require orienting the two worlds to one another's languages, professional practices, and folkways (see Chapter 24, this volume; Qualls & Czirr, 1988).

However, as noted in the beginning of this chapter, older adults have a variety of medical, social service, and mental health needs which have traditionally been served by multiple systems of care, each with its own expertise, organizational structure, and economically driven imperatives. Mental health policy for the last few decades has tended to assume that people in general and older adults in particular would be better served by a single, formally organized system. Yet, participant observation would suggest that formal attempts to coordinate services have a tendency to serve the needs of the organization rather than those of the client; that is, groups of providers come together to discuss how to prevent clients from overusing services or to impose organizationally defined models of "appropriate" service use on the client. Myrtle and Wilber (1994) argued that a modern organizational theory, "loosely coupled systems," may better describe the functioning of services for the aged and also be a better model for assuring that the client gets the best services. Interlocking multiple systems can assure that specialized services are available for the client's multiple needs; restricted access systems (e.g., "single point of entry" and so forth) are likely to create barriers to some needs being met.

A salient example is the attempts in both the UK and the US to funnel older adults with mental health problems through primary care medicine, a gatekeeper strategy that seems likely (perhaps unintentionally) to deny care rather than assure it. An empirical test of an organized system of mental health care for children in the US (the Fort Bragg Demonstration) showed that the organized system provided better continuity of care but that clinical outcomes were no better and the cost was higher than in the business as usual fee-for-service system (Bickman, 1996).

Older adults often need a wide range of services from several different service systems and their needs may be met more effectively by an informed consumer who negotiates with the loosely coupled systems and interacts with those professionals who collaborate with other systems ("interpersonal coordination," Myrtle & Wilber, 1994). Their needs are less likely to be met within artificially constrained "megasystems" whose organizations make the individual systems less adaptable to specific, changing problems. Myrtle and Wilber (1994) argued that, over time, the informed consumer and gerospecialists would be expected to evolve their own norms for coordination of care and perhaps create a new aging-centered network of services in which current boundaries (e.g., between health and mental health or between mental health and social services) would have withered away.

7.20.11.4 Roles for Clinical Psychology in Community Care

There are many areas in community care where a key role could be assumed by a clinical psychologist. (i) In assessment, it is clear that individualized approaches are required to provide care in the community—in memory clinic contexts and elsewhere, neuropsychological assessment and professional consultation with other professionals is valued. The identification of neuropsychological problems underlying behavioral difficulties is important as is greater attention to differential assessment of emotional distress (anxiety versus depression) and personality disorders in older clients. (ii) Depression and anxiety are responsive to psychological treatment approaches (see Scogin & McElreath, 1994; Woods & Roth, 1996), and their adequate treatment would reduce a great deal of suffering. This treatment could be provided across a number of community-based systems. (iii) In support for family caregivers, there is evidence that psychological factors mediate strain and depression in family caregivers, and the development of more effective programs with psychological underpinning is desired (see Knight et al., 1993; Woods & Roth, 1996). (iv) In prevention and systems-level analysis and intervention, there is a strategic role in evaluating and developing more effective community services, and a need exists for a systemic approach to thinking, practice, and evaluation. For example, outcomes for caregivers and for their relatives with dementia cannot be viewed separately and in isolation. Also, the realities of elder abuse need to be considered alongside caregiver strain. The risk factors for institutionalization need to be understood more fully, along with better insight as to when entry to a nursing home is a good outcome for all concerned.

7.20.11.4.1 Development of clinical geropsychology

Clinical psychology with older people has had contrasting histories in the two countries. In the UK, in the 1960s and 1970s contact of clinical psychologists with older people concerned assessment, usually to attempt to answer the depression vs. dementia question. Very few psychologists specialized in aging. This changed during the 1970s when a small group developed a special interest in the field and explored a much broader range of roles, including treatment approaches to depression and dementia. By 1980, a formal interest group was formed and membership has continued to grow dramatically. The training of all clinical psychologists in the UK now requires experience in working with older adults, which includes working with people with depression and dementia and with family caregivers.

In the US, specialization in aging within clinical psychology developed later and more slowly. The first meeting on training psychologists to work with the elderly was held in 1981 (Santos & VandenBos, 1982). A follow-up conference in 1992 provided mixed impressions of the growth of the specialty during the 1980s (Knight, Santos, Teri, & Lawton, 1995); university-based training showed little expansion although internship training, especially in the Veterans Affairs hospital system, had expanded (Cooley, 1995). The national organization for clinical geropsychology in the US was formed in 1992 (Knight, Santos et al., 1995). Training for psychologists in aging is not widespread in the US and not a requirement for all psychologists as in the UK. The training models in the US emphasize university teaching and research, Veterans Affairs or university medical center employment, or independent practice as eventual employment opportunities in clinical psychology. Training and working within community-based mental health programs has been relatively ignored, except for the use of community practica in training geropsychologists (Qualls, Duffy, & Crose, 1995). Few psychologists with expertise in clinical geropsychology are found in community settings (Knight, Rickards et al., 1995; Qualls et al., 1995). Psychologists in the US are also facing increased competition from master's level providers not identified with psychology (e.g., clinical social work, marriage and family therapists). These other mental health professionals are employed as a less expensive provider of psychological services, especially under managed care approaches.

Training programs need to rethink their goals to train psychologists for the jobs of the twenty-first century. Clinical geropsychology tends to be progressive in its emphasis on the interface between medicine and mental health, between social services and mental health, and placing psychologists into nontraditional settings. The future of clinical geropsychology depends upon developments in public policy and in the management of health care services over the next several decades. Certainly, it would be helpful for clinical geropsychologists to study the successes and failures of colleagues in other countries.

7.20.12 SUMMARY

In both nations, clinical geropsychology can offer a great deal to the delivery of services to older adults. Clinical geropsychology focuses on the clinical psychology of medical illnesses, on services to demented older adults, and involves psychologists in services and research across a wide variety of hospital, partial hospital, clinic, and community settings. Training and practice in community-based clinical geropsychology is inherently multidisciplinary in focus. Clinical geropsychologists, like geriatricians and other specialists in aging services, are likely to be well-positioned on the mental health side for the kind of collaboration required to make the loosely coupled systems of services work for older adults.

7.20.13 REFERENCES

American Psychiatric Association (1993). *State mental hospitals and the elderly: A task force report of the American Psychiatric Association.* Washington, DC: Author.

Ames, D. (1991). Epidemiological studies of depression among the elderly in residential and nursing homes. Special Issue: Affective disorders in old age. *International Journal of Geriatric Psychiatry, 6.* 347–354.

Anderson, J., & Trieman, N. (1995). The TAPS Project. 21: Functional and organic comorbidity and the effect of cognitive and behavioural disability on the placement of elderly psychiatric inpatients: A whole hospital survey. *International Journal of Geriatric Psychiatry, 10,* 959–966.

Arie, T., & Jolley, D. (1982). Making services work: Organization and style of psychogeriatric services. In R. Levy & F. Post (Eds.), *The psychiatry of late life* (pp. 222–251). Oxford, UK: Blackwell.

Atay, J., Whitkin, M., & Manderscheid, R. (1995). *Data highlights on: Utilization of mental health organizations by elderly persons*. Rockville, MD: Center for Mental Health Services.

Bayer A. J., Richards V., & Phillips G. (1990). The Community Memory Project: A multidisciplinary approach to patients with forgetfulness and early dementia. *Care of the Elderly, 2,* 5–6.

Benbow, S. M., Marriott, A., Morley, M., & Walsh, S. (1993). Family therapy and dementia: Review and clinical experience. *International Journal of Geriatric Psychiatry, 8,* 717–725.

Bickman, L. (1996). A continuum of care: More is not always better. *American Psychologist, 51,* 689–701.

Blanchard, M. R., Waterreus, A., & Mann, A. H. (1995). The effect of primary care nurse intervention upon older people screened as depressed: *International Journal of Geriatric Psychiatry, 10,* 289–298.

Brown, P., Challis, D., & von Abendorff, R. (1996). The work of a community mental health team for the elderly: Referrals, caseloads, contact history, and outcomes. *International Journal of Geriatric Psychiatry, 11,* 29–39.

Buckwalter, K. C., Smith, M., Zevenbergen, P., & Russell, D. (1991). Mental health services of the rural elderly outreach program. *Gerontologist, 31,* 408–412.

Burn, W. K., Davies, K. N., McKenzie, F. R., Brothwell, J. A., & Wattis, J. P. (1993). The prevalence of psychiatric illness in acute geriatric admissions. *International Journal of Geriatric Psychiatry, 8,* 171–174.

Burns, B., & Taube, C. (1990). Mental health services in general medical care and in nursing homes. In B. Fogel, A. Furino, & G. Gottlieb (Eds.), *Mental health policy for older Americans: Protecting minds at risk* (pp. 63–84). Washington, DC: American Psychiatric Press.

Challis, D. (1992). Providing alternatives to long stay hospital care for frail elderly patients: Is it cost effective? *International Journal of Geriatric Psychiatry, 7,* 773–781.

Christianson, J., & Gray, D. (1994). What CMHCs can learn from two states' efforts to capitate Medicaid benefits. *Hospital and Community Psychiatry, 45,* 777–781.

Clark, W., & Travis, S. (1994) Elderly admissions to a state psychiatric hospital: Cohort characteristics, after care needs, and discharge destinations. *Journal of Gerontological Social Work, 21,* 101–115.

Coles, R. J., von Abendorff, R., & Herzberg, J. L. (1991). The impact of a new community mental health team on an inner city psychogeriatric service. *International Journal of Geriatric Psychiatry, 6,* 31–39.

Colwell, C., & Post, F. (1959). Community needs of elderly psychiatric patients. *British Medical Journal, 270,* 214–217.

Conwell, Y., Nelson, J. C., Kim, K., & Mazure, C. M. (1989). Elderly patients admitted to the psychiatric unit of a general hospital. *Journal of the American Geriatrics Society, 37,* 35–41.

Cooley, S. (1995). Geropsychology services and training in the US Department of Veterans Affairs. In B. Knight, L. Teri, P. Wohlford, & J. Santos (Eds.), *Mental health services for older adults: Implications for training and practice in geropsychology* (pp. 11–20). Washington, DC: American Psychological Association.

Culhane, D., & Culhane, J. (1993). The elderly's underutilization of partial hospitalization for mental disorders and the history of Medicare reimbursement policies. *Journal of Geriatric Psychiatry, 26,* 95–112.

Darney, A. J. (Ed.) (1994). *Statistical record for older Americans*. Detroit: Gale Research.

Dean, R., Briggs, K., & Lindesay, J. (1993). The domus philosophy: A prospective evaluation of two residential units for the elderly mentally ill. *International Journal of Geriatric Psychiatry, 8,* 807–817.

Dening, T. (1992). Community psychiatry of old age: A UK perspective. *International Journal of Geriatric Psychiatry, 7,* 757–766.

Department of Health (1996). *Assessing older people with dementia living in the community: Practice issues for social and health issues*. London: Social Services Inspectorate/Department of Health.

Diesfeldt, H. F. (1992). Psychogeriatric day care outcome: A five-year follow-up. *International Journal of Geriatric Psychiatry, 7,* 673–679.

Domken, M., Bothwell, R., & McKeith, I. (1995). Discharge patterns from a psychogeriatric inpatient unit. *International Journal of Geriatric Psychiatry, 10,* 41–46.

Dorwart, R. A., & Epstein, L. (1993). *The privatization of mental health care: A fragile balance,* Westport, CT: Auburn House.

Eichmann, M. A., Griffin, B. P., Lyons, J. S., Larson, D. B., & Finkel, S. (1992). An estimation of the impact of OBRA 1987 on nursing home care in the United States. *Hospital and Community Psychiatry, 43,* 781–789.

Eskew, R., Sexton, R., Tars, S., & Wilcox. F. (1983). Day treatment program evaluation. In M. Smyer & M. Gatz (Eds.), *Mental health and aging: Programs and Evaluations* (pp. 63–84). Beverly Hills, CA: Sage Publications.

Estes, C. L. (1995). Mental health services for the elderly: Key policy elements. In M. Gatz (Ed.), *Emerging issues in mental health and aging* (pp. 303–328). Washington, DC: American Psychological Association.

Estes, C. L., Linkins, K. W., & Kane, K. (1994, November). *Mental health services and the elderly: Organizations in transition*. Paper presented at the meeting of the Gerontological Society of America, Atlanta.

Fasey, C. (1994). The day hospital in old age psychiatry: The case against. *International Journal of Geriatric Psychiatry, 9,* 519–523.

Fleming A. S., Rickards, L. D., Santos, J. F., & West, P. R. (1986). *Mental health services for the elderly* (Vol. 3). Washington, DC: Retirement Research Foundation.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Foucault, M. (1965). *Madness and civilization: A history of insanity in the Age of Reason*. R. Howard (Trans.). New York: Random House (Original work published 1961).

Fox, P., Swan, J., & Estes, C. L., (1986). Trends in CMHC services to elderly populations. *Hospital and Community Psychiatry, 67,* 58–102.

Freiman, M. P., Arons, B. S., Goldman, H., & Burns, B. J. (1990). Nursing home reform and the mentally ill. *Health Affairs, 9,* 47–60.

Gatz, M., Popkin, S. J., Pino, C. D., & VandenBos, G. S. (1985). Psychological interventions with older adults. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging*. (2nd ed., pp. 455–488). New York: Van Nostrand.

Gatz M., & Smyer, M. A. (1992). The mental health system and older adults in the 1990's. *American Psychologist, 47,* 741–751.

General Accounting Office (1982). *The elderly remain in need of mental health services*. Pub. No. GAO/HRD 82–112. Washington, DC: Government Printing Office.

George, L., Blazer, D., Winfield-Laird, I., Leaf, P., & Fischbach, R. (1988). Psychiatric disorders and mental health service use in later life: Evidence from the epidemiological catchment area programs. In J. Brody & G. Maddox (Eds.), *Epidemiology and aging: An international perspective* (pp. 189–219). New York: Springer.

Gilleard, C. (1987). Influence of emotional distress among supporters on the outcome of psychogeriatric day care. *British Journal of Psychiatry, 150,* 219–223.

Gilleard, C. (1996a). Community care: psychological perspectives. In R. T. Woods (Ed.), *Handbook of the Clinical Psychology of Ageing* (pp. 319–332). Chichester, UK: Wiley.

Gilleard, C. (1996b). Family therapy with older clients. In R. T. Woods (Ed.), *Handbook of the Clinical Psychology of Ageing* (pp. 561–573). Chichester, UK: Wiley.

Gilleard, C., Lieberman, S., & Peeler, R. (1992). Family therapy for older adults: A survey of professionals' attitudes. *Journal of Family Therapy, 14,* 413–422.

Gilliard, J., & Wilcock, G. (1993). Counseling for the carers. *Elderly Care, 5,* 24–26.

Grob, G. (1987). Mental health policy in post World War II America. *New Directions for Mental Health Services, 36,* 15–32.

Haley, W. E. (1996). The medical context of psychotherapy with the elderly. In S. H. Zarit & B. G. Knight (Eds.), *A guide to psychotherapy and aging: Effective interventions in a life stage context* (pp. 221–240). Washington, DC: American Psychological Association.

Haley, W. E., Salzberg, L. E., & Barrett, J. J. (1993). Psychologists in geriatric medicine: Results of a national survey. *Professional Psychology: Research and Practice, 24,* 491–495.

Hinchliffe, A. C., Hyman, I. L, Blizard, B., & Livingston, G. (1995). Behavioural complications of dementia: Can they be treated? *International Journal of Geriatric Psychiatry, 10,* 839–848.

HMSO (1990). *NHS and Community Care Act 1990*. London: HMSO.

Howard, R. (1994). Day hospitals: The case in favour. *International Journal of Geriatric Psychiatry, 9,* 525–529.

Ineichen, B. (1989). *Senile dementia: Policy and services*. London: Chapman & Hall.

Jenkins, D., & Macdonald, A. (1994). Should general practitioners refer more of their elderly depressed patients to psychiatric services? *International Journal of Geriatric Psychiatry, 9,* 461–466.

Junaid, O., & Bruce, J.M. (1994). Providing a community psychogeriatric service: Models of community psychiatric nursing provision in a single health district. *International Journal of Geriatric Psychiatry, 9,* 715–720.

Kahn, R. (1975). The mental health system and the future aged. *The Gerontologist, 15,* 24–31.

Katz, I. (1996). On the inseparability of mental and physical health in aged persons: Lessons from depression and comorbidity. *American Journal of Geriatric Psychiatry, 4,* 1–16.

Kiesler, C. A. (1991). Changes in general hospital psychiatric care, 1980–1985. *American Psychologist, 46,* 416–421.

Kitwood, T. (1996). A dialectical framework for dementia. In R. T. Woods (Ed.), *Handbook of the Clinical Psychology of Ageing* (pp. 267–282). Chichester, UK: Wiley.

Knight, B. G. (1983). Assessing a mobile outreach team. In M. Smyer & M. Gatz (Eds.), *Mental health and aging: Programs and evaluations* (pp. 23–40). Beverly Hills, CA: Sage Publications.

Knight, B. G. (1986). Management variables as predictors of service utilization by the elderly in mental health. *International Journal of Aging and Human Development, 23,* 141–148.

Knight, B. (1991). Outreach to older adults: Matching programs to specific needs. In N. Cohen (Ed.), *Psychiatric outreach to the mentally ill.* A monograph in the New Directions for Mental Health Services series. San Francisco: Jossey-Bass.

Knight, B., & Carter, P. M. (1990). Reduction of psychiatric inpatient stay for older adults by intensive case management. *The Gerontologist, 30,* 510–515.

Knight, B. G., & Kaskie, B. (1995). Models for mental health service delivery to other adults. In M. Gatz (Ed.), *Emerging issues in mental health and aging* (pp. 231–255). Washington, DC: American Psychological Association.

Knight, B. G., & Lower-Walker, D. (1985). Toward a definition of alternatives to institutionalization for the frail elderly. *The Gerontologist, 25,* 358–363.

Knight, B. G., Lutzky, S. M., & Macofsky-Urban, F. (1993). A meta-analytic review of interventions for caregiver distress: Recommendations for future research. *The Gerontologist, 33,* 240–249.

Knight, B., & Macofsky-Urban, F., (1990). Toward a policy for Alzheimer's disease in California. In P. Liebig & W. Lammers (Eds), *California policy choices for long term care* (pp. 145–170). Los Angeles: Andrus Gerontology Center.

Knight, B. G., Rickards, L., Rabins, P., Buckwalter, K., & Smith, M. (1995). Community-based services for mentally ill elderly. In B. Knight, L. Teri, P. Wohlford, & J. Santos (Eds.), *Mental health services for older adults: Implications for training and practice* (pp. 21–30). Washington, DC: American Psychological Association.

Knight, B. G., Santos, J., Teri, L., & Lawton, P. (1995). The development of training in clinical geropsychology. In B. Knight, L. Teri, P. Wohlford, & J. Santos (Eds.), *Mental health services for older adults: Implications for training and practice* (pp. 1–8). Washington, DC: American Psychological Association.

Kosloski, K., & Montgomery, R. (1995). The impact of respite use on nursing home placement. *The Gerontologist, 25,* 67–74.

Lebowitz, B. D. (1988). Mental health policy and aging. *Generations, 12,* 53–56.

Lebowitz, B. D., Light, E., & Bailey, F. (1987). Mental health center services for the elderly: The impact of coordination with area agencies on aging. *Gerontologist, 27,* 699–702.

Levin, E., Moriarty, J., & Gorbach, P. (1994). *Better for the break*. London: HMSO.

Liebig, P. S. (1993). The effects of federalism on policies for care of the aged in Canada and the United States. *Journal of Aging and Social Policy, 5,* 13–37.

Lindesay, J., Briggs, K., Lawes, M., MacDonald, A., & Herzberg. (1991). The domus philosophy comparative evaluation of a new approach to residential care for the demented elderly. *International Journal of Geriatric Psychiatry, 6,* 727–736.

Lindesay, J. (1991). *Working out: Setting up and running community psychogeriatric teams*. London: Research & Development in Psychiatry.

Lyons, D., Cohen, M., Henderson, T., & Walker, N. (1995). Partnership nursing home care for dementia: The Glasgow experience: The first twelve weeks. *International Journal of Geriatric Psychiatry, 10,* 557–560.

Mann, A. H. (1995). Future directions for the specialty of old age psychiatry: Some suggestions for debate. *International Journal of Geriatric Psychiatry, 10,* 87–92.

Moak, G., & Fisher, W. (1991). Geriatric patients and services in state hospitals: Data from a national survey. *Hospital and Community Psychiatry, 42,* 273–276.

Morris, R. G., Woods, R. T., Davies, K. S., Berry, J., & Morris, L. W. (1992). The use of a coping strategy focused support group for carers of dementia sufferers. *Counseling Psychology Quarterly, 5,* 337–348.

Mosher-Ashley, P. (1994). Therapy termination and persistence patterns of elderly clients in a community mental health center. *The Gerontologist, 34,* 180–189.

Murphy, E. (1994). A clinical role in service development: What should geriatric psychiatrists do? *International Journal of Geriatric Psychiatry, 9,* 947–955.

Myrtle, R., & Wilber, K. (1994). Designing service delivery systems: Lessons from the development of community based systems of care for the elderly. *Public Administration Review, 54,* 245–252.

Narrow, W. E., Regier, D. A., Rae, D. S., Manderscheid, R., & Locke, B. Z. (1993). Use of services by persons with mental illness and addictive disorders: Findings from the National Institute of Mental Health epidemiologic catchment area programs. *Archives of General Psychiatry, 50,* 95–107.

Neville, P. G., Boyle, A., Brooke, S., Baillon, S., Scothern, G., & Broome, C. (1995). Time for change: Psychiatric morbidity in residential homes for the elderly. *International Journal of Geriatric Psychiatry, 10,* 561–567.

Nolan, M. R., & Grant, G. (1992). *Regular respite: An evaluation of a hospital rota bed scheme for elderly people*. London: Age Concern England.

O'Connor, D. W., Pollitt, P. A., Brook, C. P. B., Reiss, B. B., & Roth, M. (1991). Does early intervention reduce the number of elderly people with dementia admitted to institutions for long-term care? *British Medical Journal, 302,* 871–875.

Office of Inspector General, Department of Health and Human Services. (1996). *Mental health services in nursing homes* (OEI–02–91–00860). New York: Author.

Patterson, R. L., Dupree, L. W., Eberly, D. A., Jackson, G. M., O'Sullivan, M. J., Penner, L. A., & Kelly, C. D. (1982). *Overcoming deficits of aging: A behavioral approach*. New York: Plenum.

Pearson, N. D. (1988). An assessment of relief hospital admissions for elderly patients with dementia. *Health Trends, 20,* 120–121.

Perry, D. W., Milner, E., & Krishnan, V. H. R. (1995). Physical morbidity in a group of patients referred to a psychogeriatric unit: A six month prospective study. *International Journal of Geriatric Psychiatry, 10,* 151–154.

Plotkin, D., & Wells, K. (1993). Partial hospitalization for psychiatrically ill elder patients. *American Journal of Psychiatry, 150,* 266–271.

Pruchno, R., Boswell, P., Wolff, D., & Faletti, M. (1983). A community mental health program: Evaluating outcomes. In M. Smyer & M. Gatz (Eds.), *Mental health and aging: Programs and evaluations* (pp. 41–62). Beverly Hills, CA: Sage Publications.

Qualls, S. H. (1996). Family therapy with aging families. In S. H. Zarit & B. G. Knight (Eds.), *A guide to psychotherapy and aging: Effective interventions in a life stage context* (pp. 121–138). Washington, DC: American Psychological Association.

Qualls, S. H., & Czirr, R. (1988). Geriatric health teams: Classifying models of professional and team functioning. *Gerontologist, 28,* 372–376.

Qualls, S., Duffy, M., & Crose, R. (1995). Clinical supervision and practicum placements in graduate training. In B. Knight, L. Teri, P. Wohlford, & J. Santos (Eds.), *Mental health services for older adults: Implications for training and practice in geropsychology* (pp. 119–128). Washington, DC: American Psychological Association.

Rapp, S. R., Parisi, S. P., Walsh, D. A., & Wallace, C. E. (1988). Detecting depression in medical inpatients. *Journal of Consulting and Clinical Psychology, 56,* 509–513.

Raschko, R. (1990). The gatekeeper model for the isolated, at-risk elderly. In N. Cohen (Ed.), *Psychiatry takes to the streets: Outreach and crisis intervention for the mentally ill*. New York: Guilford.

Ratna, L., & Davis, J. (1984). Family therapy with the elderly mentally ill: Some strategies and techniques. *British Journal of Psychiatry, 145,* 311–315.

Rosen, A. L., Pancake, J. A., & Rickards, L. (1995). Mental health policy and older Americans: Historical and current perspectives. In M. Gatz (Ed.), *Emerging issues in mental health and aging* (pp. 1–18). Washington, DC: American Psychological Association.

Rounds, J. (1984). Information and ambiguity in organizational change. In L. Sproull & P. Larkey, (Eds.), *Advances in information processing in organizations*. Greenwich, CT: JAI Press.

Santos, J. F., & VandenBos, G. R. (1982). *Psychology and the older adult: Challenges for training in the 1980s*. Washington, DC: American Psychological Association.

Scogin, F., & McElreath, L. (1994). Efficacy of psychosocial treatments for geriatric depression: A quantitative review. *Journal of Consulting and Clinical Psychology, 62,* 69–74.

Schick, F. L., & Schick, R. (Eds.) (1994). *Statistical handbook on aging Americans*. Phoenix, AZ: Oryx Press.

Schneider, L. S. (1995). Efficacy of clinical treatment for mental disorders among older persons. In M. Gatz (Ed.), *Emerging issues in mental health and aging* (pp. 19–71). Washington, DC: American Psychological Association.

Semke, J., Fisher, W., Goldman, H., & Hirad, A. (1996). The evolving role of the state hospital in care and treatment of older adults: Cross state trends, 1984–1993. *Psychiatric Services, 47,* 1082–1087.

Shadish, W. R. (1989). Private sector care for chronically mentally ill individuals: The more things change, the more they stay the same. *American Psychologist, 44,* 1142–1147.

Skea, D., & Lindesay, J. (1996). An evaluation of two models of long-term residential care for elderly people with dementia. *International Journal of Geriatric Psychiatry, 11,* 233–241.

Smyer, M. (1989). Nursing homes as a setting for psychological practice: Public policy perspectives. *American Psychologist, 44,* 1307–1314.

Smyer, M. A., Shea, D. G., & Streit, A. (1994). The provision and use of mental health services in nursing homes: Results from the National Medical Expenditure Survey. *American Journal of Public Health, 84,* 284–287.

Speer, D., Williams, J., West, H., & Dupree, L. (1991). Older adult users of outpatient mental health services. *Community Mental Health Journal, 27,* 69–76.

Stiles, P. G. (1994, November). *Utilization of ambulatory mental health services by the elderly: A preliminary report on the effects of recent Medicare reimbursement changes*. Paper presented at the meeting of the Gerontological Society of America, Atlanta.

Tester, S. (1989). *Caring by Day: A Study of Day Care Services for Older People*. Policy Studies in Ageing. London: Centre for Policy on Ageing.

Tomenson, B., Benbow S. M., & Jolley, D. J. (1994). Traffic Lights: Extracting order from the chaos of long-term care. *International Journal of Geriatric Psychiatry, 9,* 985–987.

Turvey, T., & Toner, H. (1991). The evaluation of a home based respite service for dementia sufferers and their carers. *Generations, 16,* 14–17.

United States Congress: Senate Committee on Labor and Public Welfare (1973). *Community Mental Health Center Act: History of the program and current problems and issues*. Washington DC: Government Printing Office.

Warrington, J., & Eagles, J. M. (1996). A comparison of cognitively impaired attenders and their coresident carers at day hospitals and day centres in Aberdeen. *International Journal of Geriatric Psychiatry, 11,* 251–256.

Wattis, J. P., & Fairbairn, A. (1996). Towards a consensus on continuing care for older adults with psychiatric disorder. *International Journal of Geriatric Psychiatry, 11,* 163–168.

Weissert, W. G., Cready, C. M., & Pawelak, J. (1988). The past and future of home and community based long term care. *The Milbank Quaterly, 66,* 309–388.

Whitelaw, C., & Perez, E. (1987). Partial hospitalization programs: A current perspective. *Administration in Mental Health, 15,* 62–72.

Wholey, D., Christianson, J., & Peterson, M. (1996). Organization of mental health care in HMOs. *Administration and Policy in Mental Health, 23,* 307–328.

Wolff, L. E., Woods, J. P., & Reid, J. (1995). Do general practitioners and old age psychiatrists differ in their attitudes to dementia? *International Journal of Geriatric Psychiatry, 10,* 63–70.

Woods, J. P., & Phanjoo, A. L. (1991). A follow-up study of psychogeriatric day hospital patients with dementia. *International Journal of Geriatric Psychiatry, 6,* 183–188.

Woods, R. (1995). Dementia care: Progress and prospects. *Journal of Mental Health, 4,* 115–124.

Woods, R. T., & Roth, A. (1996). Effectiveness of psychological interventions with older people. In A. Roth & P. Fonagy (Eds.), *What works for whom? A critical review of psychotherapy research* (pp. 321–340). New York: Guilford.

Worcester, M. I., & Quayhagen, M. P. (1983). Correlates of caregiving satisfaction: Prerequisites to elder home care. *Research Nursing Health, 6,* 61–67.

Wright, N., & Lindesay, J. (1995). A survey of memory clinics in the British Isles. *International Journal of Geriatric Psychiatry, 10,* 379–385.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.21
Nursing Home Care and Interventions

JACOB LOMRANZ and LIORA BAR-TUR
The Herczeg Institute on Aging, Tel Aviv University, Israel

7.21.1 INTRODUCTION 478
7.21.2 MENTAL HEALTH IN OLD AGE 478
  7.21.2.1 Definition and Approach 478
  7.21.2.2 Mental Health in Nursing Homes 479
7.21.3 NURSING HOMES: DIFFERENT SETTINGS OF LIVING AND CARE FOR DIFFERENT POPULATIONS 480
  7.21.3.1 Mental Health and the Role of Nursing Home 480
  7.21.3.2 The Effects of Institutionalization 481
7.21.4 MENTAL HEALTH IN LIGHT OF THE TRANSITION AND ADAPTATION TO A NURSING HOME: COGENT ISSUES 482
  7.21.4.1 The Decision-making Process and its Impact 482
  7.21.4.2 Preadmission Screening 483
  7.21.4.3 Adaptational Tasks in Nursing Homes 483
    7.21.4.3.1 Limited living space and seperation from personal effects 483
    7.21.4.3.2 A novel geographic ecology 483
    7.21.4.3.3 Structured and leisure time 484
    7.21.4.3.4 Food and communal dining 484
    7.21.4.3.5 Interpersonal and social adjustment 484
    7.21.4.3.6 Reframing family relationships 485
7.21.5 MENTAL DISORDERS IN THE NURSING HOME 486
  7.21.5.1 Prevalence of Selected Syndromes and Interventions 486
    7.21.5.1.1 Dementia 486
    7.21.5.1.2 Depression 487
    7.21.5.1.3 Suicide 488
    7.21.5.1.4 Psychotic symptoms and paraphrenia 488
    7.21.5.1.5 Personality disorders 489
7.21.6 A CLINICAL GEROPSYCHOLOGICAL SERVICE IN NURSING HOMES: A PROPOSED MODEL 490
  7.21.6.1 Setting and Clientele 490
  7.21.6.2 Orientation and Philosophy 490
  7.21.6.3 Methods 491
  7.21.6.4 Applications 491
    7.21.6.4.1 Preadmission screening 491
    7.21.6.4.2 Follow-up services 491
    7.21.6.4.3 Severe psychopathology 492
    7.21.6.4.4 Family interventions 492
    7.21.6.4.5 From education to psychotherapy 492
    7.21.6.4.6 Working with nursing home staff 493
    7.21.6.4.7 A university clinic 493
7.21.7 FUTURE PROSPECTS 494
7.21.8 REFERENCES 495

7.21.1 INTRODUCTION

Should a separate approach be taken to mental health for the aged in nursing homes? Could we not comprehend treatment of the aged in nursing homes as we comprehend the aged in the community? The answer is no. Furthermore, elaboration of this issue has been neglected, especially in light of the growing elderly population residing in nursing homes. This causes damage and reduces the quality of life for the elderly. While there certainly are similarities in the mental health approach to nursing home and community residents, there also are clear differences and unique demands of psychological treatment in nursing homes, based on specific knowledge of conceptualizations and clinical methodology. The absence of such knowledge hinders the function of the geropsychologist. The question whether elderly should age in organized and institutional frameworks of one kind or another has concerned modern society for many decades. But while debates continue as to the best, or necessary environment for the elderly to age, the fact is that in the postwar era organizational and institutional frameworks for the elderly have grown steadily. This growth parallels increased longevity and what is termed the "demographic revolution," that is, the colossal change in the population age structure, patterns of habitation, and the implications of these in all domains of life. Demographic projections suggest that the number of nursing home residents will triple in the 1990s (US Senate, 1987) since at least one person out of five will be institutionalized (Palmore, 1976). Keeping in mind that the needs of the elderly for mental health are extremely high, that psychiatric and medical institutionalization is unsatisfactory, and that the question of the appropriate residency in the later years is unsolved, there is a real danger that nursing homes will turn into centers inhabited by the most maladjusted elderly.

At the outset the reader was asked to bear in mind that many elderly do not suffer but enjoy their residency in nursing homes and experience well-being, development, and growth. Unfortunately, however, the problem of mental health in nursing homes is colossal. In this chapter, after defining the approach to mental health in the elderly, the issues of mental health as they appear in homes for the aged are addressed. The threats to adjustment of elderly people before, after, and during their relocation to homes for the aged are then delineated. To this end, the different nursing homes are noted as well as their inhabitants and then the utility of the nursing home is discussed, as well as the effects of institutionalization. This leads to an examination of the transitional process, the implication and effects of relocation, as well as the problems to overcome and the tasks to be achieved, in order for satisfactory adjustment to the new environment and for the preservation of mental health. While the initial focus is on the psychosocial process of adjustment, the prevailing psychopathology in these homes is discussed and then a psychological community-oriented clinical model is proposed, before a discussion of its organization and implementation as it applies to methods of treatment, including consultation, psychotherapy, and education, to staff and residents in nursing homes. The chapter ends with a future perspective on clinical geropsychology in nursing homes.

7.21.2 MENTAL HEALTH IN OLD AGE

7.21.2.1 Definition and Approach

The focus here is on the elderly person's ability to cope with personal aging processes in a specific, changing environment. Mental health is approached in the same way as those perceiving it beyond the absence of pathology. The modes of intervention and treatment are in line with such a broad perspective. The World Health Organization (1964) defines health as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" (p. 1). While many scholars are concerned with psychopathological aspects of mental health, the "positive mental health" approach (Jahoda, 1958) attempts to better understand the role of adaptive processes. Birren and Renner (1980) relate to four components in the concept of mental health: (i) the presence or absence of mental disease, (ii) the presence or absence of deficits or limitations in behavior, (iii) the satisfaction or contentment that is derived from one's life as it is being lived or has been lived, and (iv) some approximation of an ideal person. While the relevancy of (iii) and (iv) to the elderly could be understood in light of history, socialization processes, and life-review of the elderly, (i) raises problems since many elderly suffer from "social pathology," for example, loneliness (Pfeiffer, 1977) even in adjusted living, and (ii) is certainly complex since old age goes along with deficits and limitations in behavior. While mental health for the elderly may include many of the components applied to the young, such as adequate life satisfaction, self-acceptance, mastery of the self and environment, the ability to work, love, and play, the manner in which these criteria can be met changes in old age (Gatz, 1989).

While older adults experience the complete range of psychopathology seen in younger age groups (Knight, Santos, Teri, & Lawton, 1995), three points should be emphasized (Schaie & Willis, 1996). First, older people perform differently than younger on virtually all psychodiagnostic instruments, often resulting in inadequate assessment of their pathology. Second, in old age there is a high rate of physical or psychosocial loss which may constitute risk factors in mental illness. The third point concerns the major individual variability in old age. Some elderly may increase their psychological vulnerability through repeated inability to cope successfully with life stressors, while others may increase their sense of mastery as they improve resilience and become more resistant to mental illness. One should therefore be aware of the fact that applying various standards of mental health to the elderly in general, and to those residing in nursing homes in particular, is complex and requires clarification on the part of the clinical gerontologist.

A development-growth and not a depletion model is adopted to the elderly, in line with Gutmann (1994) who states that the second half of life should be viewed "neither catastrophically nor sentimentally; it is neither an exclusive tale of losses nor a golden age. Like any other sector in the life cycle, it has its own distinctive depletion, it has its own gains, and makes its own special contribution to individual development, social cohesion, and species continuity" (p. xviii). However, while adhering to this approach there is an awareness that due to, between others, lessened income, physical losses, poor health, enforced isolation, environmental relocation, retirement, interpersonal losses, widowhood, or the effects of drugs, many older persons cannot function according to the various standards of mental health. Elderly, especially in nursing homes, cannot always be autonomous and master their environments. They may lack the financial or personal and interpersonal resources to make decisions about their lives and, for certain individuals in specific situations, depression, hostility, apathy, or anxiety may be legitimate or even helpful (Gurland, 1973). Internal and environmental stress in the elderly may be more apt to produce mental disorder in the face of reduced physical and psychological reserves (Gatz, 1998). Therefore, when assessing the mental health of the elderly, it is more crucial to take into account the complete organism total life ecology, relocation being is a point in case. Exposition and explication of one's theory is crucial. Adherence to an adult developmental life-span approach is advocated (Baltes & Baltes, 1990), embedded in positive mental health as noted above and urge caution in defining and juxtaposing the different manifestation of "mental health" vs. "mental illness."

7.21.2.2 Mental Health In Nursing Homes

The problems in later life are singularly appropriate for psychological treatment. They include organic and cognitive loss, coping with stress, identity roles, interpersonal and family behavior, neurosis, psychosis and, indeed, every category to be found in any version of the *DSM* (Bienefeld, 1990). Although the magnitude and severity of the mental health problems in the elderly and their need for care are obvious, the extent of the treatment provided is severely limited and most needs are not met. The aged receive disproportionately less mental health services. It is estimated that only about 3% of all clinical services provided by psychologists go to the older adult population (Gatz, Popkin, Pino, & VandenBos, 1985). Given the higher rates of mental illness in homes for the aged, mental health care is proportionately even less then in the general population. Longevity may have increased, but not the quality of living.

Although many elderly persons enter long-term care institutions to improve their quality of life, unfortunately, the result is, at times, the opposite, with detrimental effects, often due to the inadequate quality of care (Kane & Kane, 1982). While mental disorders and symptoms are estimated between 15 and 25% (Schaie & Willis, 1996) for the population over age 65 years, in medical and institutional settings, such as nursing homes, the figure is dramatically higher, generally estimated to be in the range 60–90% (Cohen, 1991; Gatz & Smyer, 1992). A major criteria for mental health is a person's ability to cope with their environment. For many who are unable to live independently in the community, this criteria has to be met in adjusting to a nursing home. The absence of a caregiver or the lack of social support networks distinguishes most nursing home residents from equally impaired elders dwelling in the community. Twenty-five percent of nursing home placements are precipitated by the illness or death of the caregiver (Colerick & George, 1986). In the early 1980s, a national nursing home survey (Eisdorfer & Cohen, 1982) found that 30% of nursing home residents had "diagnosable" psychiatric disorders and that 61% had one or more mental impairments. While more older adults have dropped out of, or never entered, the mental health system, other organizational settings, notably nursing homes, have become the service destination for the mentally impaired elderly (Johnson &

Grant, 1985), especially since the Deinstitutionalization Act (Goplerud, 1979). Donahue (1978) estimated that 40% of the patients discharged from public mental hospitals eventually found their way into nursing homes. Despite the fact that a majority of nursing home residents are in need of mental health services (Newman, Griffin, Black, & Page, 1989), these settings are designed and staffed primarily for physical, rather than mental, health care (Smyer, 1989). In most nursing homes the staff is also untrained and too overburdened to deal with emotional problems, some of which are exacerbated or developed as a result of living in the home.

Psychotherapeutic mental health programs and provisions are rare. Psychiatric visits or consultations are the exception to the rule (Johnson & Grant, 1985). It has been noted that "the tragic irony about nursing homes is that they represent a major setting for the mentally ill, with grossly inadequate mental health attention" (Cohen, 1980, p. 983). There is a strong urgent need for interventions by mental health professionals such as psychiatrists, psychologists,or trained mental health social workers in order to handle the severe problems of mental health in homes for the aged.

7.21.3 NURSING HOMES: DIFFERENT SETTINGS OF LIVING AND CARE FOR DIFFERENT POPULATIONS

Although the generic term "nursing homes" has been adopted, the reader must realize the variability of these institutions. These habitats range from "sheltered homes" to "homes for the aged" to "convalescent homes." They differ in segregation (isolation from the outside world), policies, and control (the degree of autonomy permitted to patients). They range from "open" to "closed" institutions or floors, providing partial to total care, from secluded rural to urban community settings. They vary widely in quality, regulations, in the focus on instrumental and/or also social and affective needs, and offer various degrees of help in daily living. Some offer basic nursing and custodial care while others provide a variety of services to restore, or maintain, the residents' ability to function. Support in various homes may be related to the amenities and regulations of the facility and to the availability of more than minimal staff and care resources (Timko & Moos, 1990). Smaller facilities may, in certain instances, be more constructive since they resemble a home environment, better meet the resident's needs, and minimize the impact of institutionalization. As the size of the institution increases, the quality of communication, satisfaction, and active involvement of residents may decline (Schaie & Willis, 1996). It should be emphasized that the different structures and kinds of nursing homes bear different implications for their residents' behavior in terms of social life, humanistic or dehumanizing atmosphere, adequate health care, independence, control, well-being, and mental health.

Just as there are major difference between the various nursing homes, so do the residents differ in personality, coping abilities, mental health and cultural background. However, some generalities may be drawn. Generally, in contrast to the elderly dwelling in the community, most nursing home residents have the following in common: (i) their age is higher: 75% over 75 years of age and 35% of these are 85 or above (forecasts for the year 2003 suggest that 87% of nursing home residents would be over 75; National Center for Health Statistics, 1983); (ii) the majority are widowed; (iii) more have no surviving immediate family member living nearby; (iv) they are more functionally disabled; and (v) many more suffer from physical or mental problems as well as from chronic physical conditions involving pain, threat of hospitalization, and gradually increasing impairments.

7.21.3.1 Mental Health and the Role of Nursing Home

The function and utility of the nursing home in society is influenced by social forces, policy, economics, family structure, commitment, and social values. However, the climate and physical environment in which people live and work is related to life expectancy, stress, and various medical, physical, and mental disorders (Wantz & Gay, 1981). Nursing homes, therefore, determine to a great extent the well-being and mental health of their residents (Lawton, 1985).

For the elderly person with some deficit in self-care ability, the nursing home answers the major need for tending and care. It provides security, protection, and safety and assists in the activities of daily living. Usually nursing homes also attempt to satisfy the additional basic needs for shelter and human comfort. The family expects the nursing home to become a surrogate fulfilling the dual aspects inherent in the notion of care: providing concern, support and morale, as well as performing the direct work of tending. As illness, frailty, or loss in resources and support systems are often the background leading to nursing home admission (Hirschfeld & Fleishman, 1990), they highlight the fact that those admitted are vulnerable and in need of psychological support, in addition to tending. One task of the psychologist in nursing homes is to ensure that residents also receive mental support and care along with physical care.

Various studies indicated psychological and social improvements following the move to public housing facilities or retirement community centers (Lawton & Cohen, 1974), especially in homes which were designed specifically for the elderly and the residents themselves choose the home. However, disagreement exists with respect to effects of rehousing on health, and many studies found negative effects of relocation (e.g., Lieberman & Tobin, 1983). Butler (1975) claimed that most elderly people view the move to a nursing home with fear and hostility and believe that the move to an institution is a prelude to death. Farrari (1963) compared two groups moving from home to an institution, one voluntarily and the other with no option, and found that 10 weeks later mortality among the latter was to be extremely high. Lawton (1980) suggests that the differences in findings may be due to the variability in the quality of housing, sampling, economic status of inhabitants, and personality variables as well as to research design differences. Determining, in varying degrees, are also the detrimental effects of institutionalization.

7.21.3.2 The Effects of Institutionalization

Nursing homes are institutions and as such their inhabitants are subject to the effects of institutionalization. The very nature of an institution contradicts the nature of a "home," which implies intimacy, autonomy, privacy, and control. As Tobin and Lieberman (1976) state "it is difficult to avoid the implication that even the finest institutions cause the deleterious sequale commonly refered to as *institutional effects*" (p. 5). Basic factors leading to institutional effects include: physical and human ecology, territory and space constructions and limitations, the structuring of time, sharing of space and interaction with neighbors, regimes, inspections, routines, the burden of authority, use and abuse of dependency, ritualizations, stigmation, the impact of crowding, lack of privacy, group pressure, and engulfment (Goffman, 1961). Negative effects upon inmates of institutions include: changes in self-perception, feelings of inferiority, dependency, alienation, stress, depression, somatic and mental illness.

Such general factors have specific manifestations, impacts, and detrimental effects in nursing homes. The negative stereotyping of society incorporated by nursing home staff (Wright, 1988) may result in a negative atmosphere in the home and a lower self-concept of its residents. Insecurity and insufficient privacy (Firestone, Lichman, & Evans, 1980) may cause a sense of intrusion and fluidity of personal boundaries. Limitations in providing care may result in loss of self-esteem, estrangement, isolation, and deindividuation (Tobin & Lieberman, 1976). Low quality of care (Johnson & Grant, 1985; Kane & Kane, 1982), lack of sufficient personal attention, and respect may all induce depressive moods. Constricting freedom and control may be detrimental. Rodin and Langer (1977) have demonstrated how reducing restriction of choice and control in a home for the aged resulted in increased alertness, activity, and general well-being, while increasing restrictions have the opposite effect. They conclude that "some of the negative consequences of aging may be retarded, reversed, or possibly prevented by returning to the aged the right to make decisions and a feeling of competence" (pp. 897–902). Perlmuter and Monty (1979) suggest that the major problem for residents is the inability to exercise choice, control, and personal responsibility which then has a harmful effect on well-being and adaptation. Residents should be given greater control of their daily lives as well as participate in decisions affecting various aspects of the home's programs and policy (Moos & Lemke, 1985). Studies in social ecology of nursing homes show that reinforcing independence not only reduces dependence but also reduces the incidence of disease and death (Baltes, 1982).

Assembled in a common environment, the newcomer to the nursing home cannot escape the dreaded identification with their surroundings which include being sick, yearning for care, and feeling closer to death. In order to adapt, they have to participate with others in the cannons of the institution, but in doing so they may lose uniqueness, and become discontent, bitter, and hostile. The institutional structure and atmosphere may restrict or promote interpersonal and social contact (Hansson & Carpenter, 1994). It may decrease family visits leading to loss of contact with family members and friends. As family members recede into the background, the position of the institution as the last and "only" home for the elderly person is reinforced. The environment then signifies vulnerability, mutilation, and death, but at the same time protection and survival. Given all the above it should come as no surprise that individuals in nursing homes are significantly lower in psychological well-being than individuals with similar levels of physical and social function living in the community.

The impact of institutional processes, including staff burnout, may be so great as to undo and render useless the effects of professional therapeutic interventions (Lomranz, 1991).To counteract this impact, nursing homes should be viewed as communities whose subsystems mutually interact with the broader surrounding community and its social policy dimensions. At the same time, investment in nursing home staff should be aimed at active preventive programs to counteract institutionalization.

The chapter will now discuss the specific issues of the processes and mental health implications of moving to a nursing home and the factors involved in adjustment and adaptation.

7.21.4 MENTAL HEALTH IN LIGHT OF THE TRANSITION AND ADAPTATION TO A NURSING HOME: COGENT ISSUES

The transition to a nursing home involves significant change, including separation, loss, and readjustment. For many it means "closing" a 60+ year chapter of their lives. The new location encompasses a novel territory, neighborhood, social group, entire community, and access to the outside, larger world. The location chosen, its atmosphere and culture may be crucial. For most people a house is a meaningful life-space: "home is where the heart is." Home signifies stability, roots, a feeling of belonging, security, identity, intimacy, personal space, and generativity. Permanently relinquishing one's home can be a painful step. This surrender is especially hard for the elderly person who experiences many additional changes and may cling to the familiar for security and permanency. Thus, the need to adjust to a new framework involves the ability to separate emotionally from a previously meaningful stable world, that is, from habits and lifestyle held throughout many years, from defined identities, and from a familiar society and environment. Successful relocation, therefore, is a complex twofold process that requires reconciliation with loss as well as adaptation to a novel environment. The rationale to move to a nursing home, adjusting and maintaining well-being along time, all depend on a wide range of personal, interpersonal, and environmental resources. These include a complex range of factors, such as, individual characteristics and ego resources, past history, family structure, quality of relationships, economic status, social background, environment, and culture. Once in the nursing home, a variety of adaptational task are expected. Given the variability in nursing homes and in coping patterns of the individuals, it is almost impossible to predict adjustment patterns. Nevertheless, some basic factors are highlighted which are deemed critical to the adjustment process.

7.21.4.1 The Decision-making Process and its Impact

The decision to leave home and move to a nursing home is an emotionally-loaded process involving multiple aspects, phases, and actors (Groger, 1994). This major life decision often involves anxiety, stress, ambivalence, fears, doubts, hopes, and risk taking, feelings which may be unavoidable considering the life changes involved and the necessary uncertainty about the new habitat and the future. The family often plays a major role prior to admission and, thereafter, in the readjustment of the elderly. Overt and covert individual and family conflicts may often emerge. Family members frequently share the elderly person's ambivalence (Bar-Tur, 1993; Halpert, 1991). Institutionalization of an elderly parent may also create family crisis. Moreover, data suggests that above 50% of those entering homes for the aged do so without having taken a meaningful part in the decision-making process (Reinardy, 1992).

The decision-making process affects the person's adjustment to the nursing home. Elderly who exercised more control over the decision adjust better than those who have less control over the move. Often an elderly person may feel "dumped" by their family and consequently become alienated and depressed. Family structure has major impact. Having, for instance, one daughter or sibling reduces an older person's chances of admission by about one-fourth (Freedman, 1996). Sometimes family members act out conflicts in their relationships with the nursing home staff, further complicating the adjustment process. Children may seek professional help in the process of placing their parent in a nursing home. Guilt, anger, regret, and resentment are almost ubiquitous. Planning and implementing the decision with the family in a context of safety, nonthreatening, and open communication can minimize distortion and promote healing of the family (Bienenfeld, 1990). The psychological loaded milieu described here obviously indicates the possibilities for clinical geropsychologist to assist in conflict resolution and stress reduction for all those involved in this transition and relocation. At stake is not merely "adjustment" to the nursing home, but solving existential dilemmas to improve the quality of life.

Over one-half of the persons admitted to a nursing home are discharged within three months. This figure includes those who die as well those transferred to hospitals or to the community. Most of the hospitalized, as well as those discharged to the community, eventually return to the nursing home (Lewis, Cretin, & Kane, 1985). Cooperation between hospitals, the community, and nursing homes is crucial, especially when dealing with the chronically sick elderly.

7.21.4.2 Preadmission Screening

Elderly adults admitted to a nursing home usually present a combination of mental and physical health problems. Potential patterns of interaction between physical and mental health conditions among nursing home residents may be expressed in the four categories of (i) "classical" psychosomatic disorders; (ii) somatizing responses (e.g., conversion syndrome); (iii) physical illness causing psychiatric symptoms; and (iv) psychological responses to illness. Family and social issues are often embedded in all these. The preadmission screening procedure is therefore very important for certifying the type of care, mentally and physically, the new resident requires (Symer & Garfein, 1988). It may add clarification to the decision-making process as to the kind of home to seek and may lay the basis for the kinds of interventions needed to assist the individual's adaptation in the nursing home. It is recommended (Newman et al.,1989) that psychologists become more involved in the substance and process of preadmission screening, assessment being one of the applied psychologist's strengths. Depending on the results of the prescreening, the clinical psychologist may further refer the candidate to a medical, family, or consulting psychologist as well as to social workers.

7.21.4.3 Adaptational Tasks in Nursing Homes

The transition to a nursing home requires various adjustments. Some of the major tasks, in specific areas of daily living, that must be accomplished to promote an appropriate quality of life in the nursing home, are outlined in what follows.

7.21.4.3.1 Limited living space and seperation from personal effects

Exchanging a macro- for a microenvironment implies radical reorientation of spacial experience. Personal and interpersonal space has to be modified. Transition to diminished space, sometimes sharing one room with someone else, requires residents to limit drastically their living space and separate from many objects they have accumulated throughout many years. In such objects, a person's history, identity, and lifestyle are encapsulated. Separation from such meaningful objects, in addition to other losses experienced in old age, may create a painful sense of discontinuity. Relocation sometimes involves rapid renunciation, resignation, and separation, whereas mourning over the losses may be prolonged. For example, about one year after admission, Mrs. B. realized at her birthday party that she didn't have a single flower vase. She burst into tears at the memory of all the beautiful things she had been forced to leave behind, expressing a sense of deprivation. Mr. Z., a widower without relatives, negotiated for two years with a nursing home for them to accept him with his 200 books; he passed away, bitterly, in midst of negotiations. Neither had any contact with a mental health professional. Stripped of personal belongings and yearning for them may cause depressive mood. Attachment to familiar objects lends a sense of security and may serve as an adjustment mechanism. While the importance of transitional objects in young children (Winnicott, 1953) is recognized, they should be considered significant throughout the life-span. Just as the child therapist may utilize transitional objects to reinforce growth and development, the geropsychologist may do the same, but usually on the basis of the absence of such objects.

7.21.4.3.2 A novel geographic ecology

Adjustment to ecology includes the physical environment, the world of plants and animals, the human inhabitants, and the larger geographical surroundings. It also signifies the importance of environmental planning for older people (Moos & Lemke, 1985) and the behavioral and affective outcomes in the person-environment fit. Geographic experiences, defined as "the individual's involvement within the spaces and places of his life" (Rowles, 1979, p. 81), represent meaningful dynamics in the individual. Organization of a new living space is difficult. Residents must familiarize themselves with a whole new world. Mrs. B., a 90-year-old lady, said sadly "I lived in the same four-room apartment for 60 years, and everything had its place. Here, in one room, I have hardly anything and I still can't find nothing." The required adjustment not only concerns one's room, but also the facilities in the building and the surrounding area. Uncertainty as to what is usually posited firmly in our minds (e.g., rooms, buildings, streets, or bus schedules) may cause disorientation, especially for those who are confused, leading to anxieties and nightmares. Coping and adjusting to the ecological aspects consumes physical as well as mental energy but is crucial for enabling the residents to "feeling in place" and enhances belongingness.

7.21.4.3.3 *Structured and leisure time*

New residents face a transition from controlling and structuring their time, in the past, to a milieu in which almost all the time is structured for them and they also may be discontent about the way time is filled. Life in many nursing homes goes by a strict time schedule. There are specific hours for such things as eating, social activities, and receiving medication. Moreover, often these time-bound activities become the most significant events, the markers and content of existence and, as such, being after all very simple routines, may render the totality of life for the resident as boring and meaningless.

Transition to a nursing home frees time and energy that in the past were devoted to activities of independent daily living. Potentially individuals may enjoy, as many do, and use their time for a variety of creative and educational activities. However, although being released from burden and given freedom for pleasure and personal enrichment, many residents may find it difficult to shift from fulfilling duties to enjoying leisure time. Such an inability to shift relates to the difficulty in giving up roles, values, identities, and lifestyles, all molded during one's life-span. Neulinger (1981) postulated that the ability to enjoy leisure time depends on a certain mental state that includes a feeling of freedom together with basic motivation to experience and enjoy. Physically or emotionally frail residents may find it hard to learn new outlets for satisfaction and as a result experience low morale. The passage of time may also be experienced as monotonous, slow, and dispersonalized, features associated with the depressed elderly.

Further implications of temporality refer to the personal experience of time (Lomranz, Shmotkin, & Vardi, 1991) such as time perspective or the meaning of future time. Such aspects retain specific significance in the existence of nursing home residents and they present furtile ground for the work of the clinical geropsychologist.

7.21.4.3.4 *Food and communal dining*

Gutmann (1994) found, cross-culturally, food and oral preoccupation to be a major avenue of joy and pleasure and of highest significance for elderly people. Many nursing homes operate on the basis of communal dining and the issue of food becomes a major adaptational task for the resident. Coupled with changes in sense of taste, as well as the loss of familiar "home food," losing the role, especially for women, of "cook," and communal dining, can all be extremely frustrating and may become a major source of friction between residents as well as with staff. For an elderly resident, oral intake sometimes constitutes the only source of gratification and a main channel of communicating, interacting, and coping with the caretakers at the home and the world at large. Food, its quality and associated eating rituals, the conflicts as well as memories around it, all indicate personal and interpersonal significance. The psychological features associated with food: satisfying, receiving, giving, and caring, the ability to manipulate food as a means to interpersonal contact, are all severely restricted in communal dining and bear psychodynamic implications regarding mental health of residents, that the geropsychologist should be aware of.

7.21.4.3.5 *Interpersonal and social adjustment*

Adjustments in the interpersonal sphere and the social environment are perhaps the most crucial for the resident. The social circle, including residents as well as staff, is large, unknown, and unstable. Yet, new relationships have to be established. Many nursing homes are planned to have more than one resident per room. For some, the hardest tasks in adjusting to a nursing home are sharing personal space with a roommate, and forming a long-term intimate partnership. This requires a great deal of flexibility, compromise, and willingness to forego independence and freedom of choice regarding daily issues (e.g., when to turn out lights, cleaning and tidying, use of toilet, etc.) which take on increasing significance since they represent basic needs such as territory, belonging, or boundary intrusions, and become the ground for feelings of rejection, insult, or elicit aggressive behaviors.

Stripped of their previous acquaintances and social markers, the resident may feel insignificant and worthless. Loss of identity, together with the need to integrate into a new society, can cause insecurity, anxiety, and a sense of alienation. Mrs. B. said sadly, "for 60 years I lived in the same neighborhood and everyone knew me. When I came here, I suddenly lost my identity. I became anonymous, nobody knew who I am, what I have done all my life. I have become inferior in the world. I am just one old lady out of 300 old people." For many residents the absence of friends often renders the atmosphere in the home as distant and gloomy. Finding new friends and establishing personal relationships is a difficult task, especially in old age. Getting to know and be known requires motivation, risk-taking, confidence, learning, and adaptability, psychological qualities which the elderly resident may require professional help in utilizing for their well-being. The clinical geropsychologist may find here the appropriate ground upon which to help residents utilize their past experience, their social skills, and use ego strength to establish interpersonal and social relationships (Hansson & Carpenter, 1994). Since compatibility renders adjustment easier, the geropsychologist should be involved in decision-making about roommates. At times rearranging the environment may be useful. However, often they have to deal with already existing conflicts, in and between residents, and use them as analytic avenues towards adjustment and emotional security.

7.21.4.3.6 Reframing family relationships

Entering a nursing home necessitates the reformulation of relationships between the residents and their family. This sensitive process is emotionally loaded and may be the ground for deep conflict, as well reciprocal respect and love. The role of the family in the elderly's adjustment and their quality of life in the home is paramount. Relationships may become engulfed in ambivalence and stress, and past family dynamics may become the cornerstone for present conflicts or sound relationships. Although many families experience a heavy burden providing daily care for the older person at home, such stress often continues after the person's institutionalization (Stephens, Crowther, Hobfoll, & Tennenbaum, 1990). There is an ongoing launching process (Lomranz, 1995) in which, in contrast to the stage of adolescence, it is the parent who is being launched. Nevertheless, it bears similar launching characteristics, for example, guilt, ambivalence, letting go, holding on, transfer of responsibility, and mourning (Lomranz, Shmotkin, Eyal, & Zohar, 1996). Tobin (1987) describes children's mixed feelings, guilt and shame coupled with relief, in relocating a parent to a nursing home. Parents, however, may feel frustration stemming from their experience as they in the past shared lodgings with their own parents and grandparents. Although parents recognize the present difference in cohorts, housing arrangements, culture, and lifestyle, some may feel rejected and unwanted, believing that cohabitation would be possible if only desired by their children (Bar-Tur, 1993). Such beliefs may, in response, elicit guilt feelings and hostility in children. Zivoni (1993) found that positive intergenerational relationships had an indirect negative influence on application to a nursing home. Tobin (1987) described the importance of considering institutionalization as a family process. Tobin suggests that there will be a "good" figure (e.g., social worker or intern) satisfying the family's desire for individualized care for the resident, and a "bad" figure (e.g., administration) who could tolerate the family's anger. An elderly person's feeling of disconnection from their family may be detrimental to their mental health. Mutual ambivalence should be worked through. Intergenerational relations and the kind of daily involvement children have with the parent, family visits, and adequate contact with administration are extremely important for the well-being of the elderly and the family. In attending to their clients' problems, clinical geropsychologists sometimes have to involve the children (similar to the case of child therapy). Adult children should also be conceived as part of the social support network and in the role of caregivers, including the difficulties caregivers encounter (Stephens et al., 1990).

To summarize, nursing homes, the effects of institutionalization, and some core issues in a resident's adjustment to a nursing home, have been characterized from a mental health perspective. The elderly living in a nursing home should not be mainly preoccupied with coping and survival. Their quality of life should be the highest possible according to their potential. The nursing home should provide conditions which enhance physical and mental health and promote development and growth. This, however, is not always obtained and hence it is here that adaptational issues and the role and functions of the clinical geropsychologist are dealt with.

For growing numbers of aged people, especially (but not only) those who cannot live independently in the community, adjustment to a nursing home may be a major objective. The above explicated adaptational tasks actually refer to coping with aging processes and institutional effects. These tasks constitute the stage upon which the drama of life for the elderly resident is played out. They may continuously serve as "explosive potentials" as well as issues around which elderly may mask aggression, frustration, and depression. Thus, it is important for the clinical geropsychologist to be aware of them. Diminishment of these factors (e.g., spacial or geographical orientations, social circles) does not necessarily depend on length of residency, but may relate to lifestyle (e.g., a resident who is not mobile enough, or has no contact with the community), coping strategies, and quality of life in the home. For many elderly, life in nursing homes is dull and purposeless. Stripped of external responsibilities and social roles, the importance of the outlined existential components in the daily routine grows unproportionally, consuming physical and mental energy. Yet, at the same time, developmental processes fully operate, preoccupations with health, bodily concerns, life review, family issues, social surroundings, and processes relating to identity and the self. However, these processes may not find direct expression. Conflicts are expressed through the experience and events in daily life, be they tragedies or consumption of food. The clinical geropsychologist must be familiar with these ongoing adaptational tasks, their implications, and the personal meanings attached to them so as to be able to detect and utilize them as avenues to internal life.

If the foregoing may be comprehended in light of existential coping and "social pathology" (Pfeiffer, 1977), the chapter now focuses on more severe psychopathology.

7.21.5 MENTAL DISORDERS IN THE NURSING HOME

The extensive need for psychiatric treatment of nursing home residents has been noted above. Nursing homes report that they are admitting residents who are "sicker" and more frail than in the past. The prevalence of psychiatric behavioral disorders in nursing homes has been estimated to range from 68 to 94% (Grossberg, Hassan, & Szwabo, 1990). Burns et al. (1993) found mental illness among elderly nursing home residents at a rate of 64%; however, only about 4.5% received any mental health treatment in a given one-month period. When they did receive treatment, it was likely to be psychotropic medication only.

Manifestations of psychopathology in residents of long-term care facilities may include complaints of pain, anxieties, physical illness, impulse control problems, personality disorders, deliriums, thought disorders related to dementia, and chronic neurotic or psychotic behaviors. Evaluation of these behaviors and their differential diagnosis is complex and often not performed. Adequate treatment is prevented due to the lack of professional knowledge and facilities, limited resources, and staff. Norris, Snow-Turek, and Blankenship (1995) suggest that the environmental impact of institutionalization, as well as declining health, contributes to somatic and depressive symptomatology, as does frequent napping and use of psychotropic and sleeping medications prescribed (sometimes inappropriately) by general practitioners. The nursing home staff also suffer from such conditions. Everitt, Field, Soumerai, and Auron (1991) note that only minor attention has been focused on the effect on nursing home staff of agitating and disordered residents' behavior which might persist in spite of ongoing pharmacological therapy, although perhaps inappropriately managed. Staff, patients, physicians, and institutions are all leverage points for further intervention to improve the delicate balance between patient behavior, staff distress, and medication use.

7.21.5.1 Prevalence of Selected Syndromes and Interventions

The range of psychopatology in a nursing home is similar to that of the general population. Rovner and Rabins (1985) have divided psychiatric disorders seen in nursing home into three major groups: (i) cognitive disorders (e.g., delirium and dementias), (ii) depression, and (iii) behavioral disorders.

7.21.5.1.1 Dementia

Dementia is thoroughly presented in Chapter 10, this volume; here are highlighted some of their implications in homes for the aged. Estimates of incidence of the disease in individuals who are 80+ range from 10.8 to 47.2% (Gatz, Kasl-Godley, & Karel, 1996), hence its prevalence in many nursing homes is relatively high. As Chapter 10, this volume explicates, dementias appear in various patterns and modes of progression, a fact which carries many implications and difficulties in institutional management and care. Understanding emotions and moods of the older person suffering from dementia, correctly interpreting their behavior, and clarifying the impact of existing care practices are difficult tasks for the staff and pertain to the quality of care the institution is able to provide and the well-being of the demented (Cohen-Mansfield, 1996). Many of the demented elderly in nursing homes suffer from agitation reflected in a range of inappropriate behaviors: repetitive actions (e.g., pacing back and forth, repetition of words); socially inappropriate behavior (e.g., entering someone else's room and handling their belongings, or exposing themself nakedly); and aggressive behavior towards self or others. Often such patients walk out of the institution and cannot find their way back. Different agitating behaviors are associated with different individual characters, organicity, as well as situational characteristics. Such behavior is not only detrimental to the patient and their family, but creates fear and panic in the home where interaction is close and dense. "Healthy" and fearful residents express discontent of not being separated from persons even at early stages of Alzheimer. Each nursing home has its own covert norms and culture which are important for staff to be able to identify it in order to comprehend residents' behavior and reduce their own distress.

At various levels of dementia, aggressive behavior may be associated with poor-quality social interaction (e.g., responding to a roommate's minor invasion of personal space), or result from the patient's unsuccessful effort to interact verbally with their neighbor. Less aggressive behaviors tend to be manifested by persons who suffer from moderate to severe cognitive decline but are in relatively good physical health and under positive environmental conditions promoted by the institution. For some older persons, destructive behaviors may present a form of exercise or stimulation in the face of cognitive deterioration and reduced opportunities for other activities in the home. Various behavioral problems may occur frequently in a home inhabited by individuals suffering from a variety of illnesses, each at various levels of progression. It becomes an extremely difficult task for management to differentiate successfully and separate between such coexisting illnesses among residents.

A broad range of therapeutic functions and interventions are available to the geropsychologist. Proper neuropsychological assessment (see Chapter 7, this volume) is paramount to better understanding the cause of acute behavioral problems and in recommending appropriate treatment and case management. Delirium may be common in nursing homes and differentiation should be made whether or not it results from progressive demetia. The resident with multiple medical problems and receiving many medications may be particularly vulnerable to delirium. It takes special sensitivity to assess some deteriorated behaviors but also to reveal potential abilities and preserved areas. Some patients, for example, who do not respond otherwise, may respond to music even employing their memory. While often residents and staff avoid contact with a demented patient, it is important to provide them with social contact, a comfortable environment, and appropriate stimulation in order to maintain a proper quality of life. Psychotherapeutic interventions should address the different stages of dementia. Dynamic and primarily behavioral orientations ranging from personal to group interventions, including specified groups such as Reality Orientation, have all been used successfully. Other major avenues include working with the family and attending staff. Couples in the nursing home often have to be separated due to the deterioration of one of them, a traumatic event which often requires psychological intervention. The geropsychologist may advise staff regarding issues ranging from environmental safety of the wandering patient to impulse control problems, relevant neuropsychological presentations, and elevating staff distress. The various therapies should be promoted as important components of dignified management, in which the patient's entire life is respected and considered. This may also provide clues to the patient's resources that can to be drawn upon.

7.21.5.1.2 Depression

Depression has been considered the major mental health problem in aging. However, a more cautious approach is being seen. Differentiation in this category is related to issues such as prevalence, etiology, relation to physical illness, biology, and stress, as well as to methodology, screening, and diagnostic approaches (Blazer, Hughes, & George, 1987; Gatz, in press; Katz, in press; Teri, 1991). Affective disorders and depression are elaborated on in Chapter 9, this volume. This disorder in the nursing home setting should be emphasised. The prevalence of depression, its symptoms, or dysphoric mood rank high among residents of nursing homes. They may be subtle and often go unnoticed and untreated leading to excess disability, morbidity, and mortality. While the signs and symptoms are often ascribed to underlying medical problems and advanced age, in fact, some of depression's more prominent characteristics may be observed in many residents: loss of appetite and weight loss; severe fatigue; sleep disturbances; changes in bowel habits and somatization; sadness; feelings of loneliness; generalized lack of interest in life; apathy; increased time spent in bed; avoiding visits; nonparticipation in communal activity; low self-esteem; increased preoccupation with death; and a gloomy evaluation of one's present and future situation. More subtle signs in residents could include: memory problems (pseudodementia); difficulty in making even minor decisions; dreams of being lost and lonely in frightening, desolate places; and crying out for help with no one responding (Pachana, Gallagher-Thompson, & Thompson, 1994). Multiple losses appear in high concentration in nursing homes, perhaps rendering depressive symptoms as almost inevitable in some such institutions. Admission to a nursing home, especially involuntary, transition, and relocation, is often accompanied by an increase in depressive symptoms and lowlife satisfaction (Rovner et al., 1991), although confounding factors such as declining physical health, often the cause for the move, may be involved. Newmann, Engel, and Jensen (1991) have suggested the term depletion syndrome to describe a form of depression commonly observed in older adults, primarily involving lack of interest and feeling that everything is an effort. Blazer, Burchett, Service, & George (1991) term similar symptoms minor depression. These symptoms are often observed among residents and reinforce the search for more subtle signs of depression among nursing home residents. Etiology awaits further differentiation whether such symptoms constitute late-onset disorder, are secondary to residents' deteriorating health or reactive to being institutionalized.

Depressive disorders cause major distress to home and health care personnel. Residents at high risk would include those with a history of depression, poststroke individuals, new residents, those returning from hospitalization or admitted for rehabilitation, and those grieving or having severe adjustment problems. Differentiating between normal and pathological grief is critical. Also, several commonly prescribed medications can have mood-altering side effects, a significant factor when dealing with an institutionalized population which has, perhaps, the highest in intake of medical drugs. Often secondary depressions as manifestations of medical disorders are encountered. In order to provide treatment plans and direct the staff, the geropsychologist should administer diagnostic procedures, via clinical interview, symptom checklist, or more extensive testing. Although in most cases residents are referred to a psychiatrist for medication, the full range of causes for depression should be considered. These could include a variety of intrapsychic and psycho-social aspects, factors in which the geropsychologist's contribution is needed for evaluation and intervention. Effective therapies with depressive people, which also lend themselves well to the nursing setting, are behavior modification, cognitive, and psychodynamic therapies (Gallagher & Thompson, 1982).

7.21.5.1.3 Suicide

Suicide, its epidemiology, risks, and prevention, have been discussed in Chapter 11, this volume. This chapter emphasizes its special danger in the nursing home population. Many nursing homes constitute an assembly of suicide risk factors, as the discussion of depression indicates (Glass & Reed, 1993). Feeling helpless and hopeless, elderly candidates for suicide are less likely to make manipulative or suicidal gestures and they often signal their intentions, which are not always noticed. Some may, however, become unusually quiet and withdrawn. Recognizing the subtle signs of depression may require a high level of professional sensitivity. One clue is giving away valuable personal possessions. In either case, staff are often (it seems too often) "surprised" by the suicidal act.

The impact of suicide in an institutional perspective should be considered. Suicide in a nursing home can have an extreme impact on the residents and the staff. This may also be combated by defensive distancing. Usually it arouses anxiety, anger, and often a feeling of helplessness and inadequacy among those who were involved in the caregiving. Staff and management are often confused as death occurs (e.g., should obituaries be placed in the nursing home, should busing be arranged for funerals, etc.). There is always the threat of "contagion" as one suicide triggers suicidal thoughts or encourages other residents to attempt as well. This produces a threatening atmosphere in which fear of the "next" suicide prevails and is felt by most individuals in the home. At such times, interventions by staff may be crucial. In one nursing home six suicides were witnessed in a period of about two months. The atmosphere in the home became extremely tense. Both staff and residents were very anxious and many expressed guilt, helplessness, and desperation. Inspection revealed feelings of frustration, anger, and despair. A special intervention program was designed, including individual and group discussions in which residents were helped to deal with stressful life events and current obstacles, to ventilate, and were taught self-control and self-relaxation techniques. The psychologist's major task was to identify residents at risk and to provide immediate help for some residents, for example, those who had been friends of the suicidees or witnessed their death. A prevention program was also designed in which staff members were trained to identify signals of threat among residents, and certain security measures and technical devices (e.g., preventing accumulation of drugs, closing "risky" places such as roofs) were taken to prevent future suicidal events.

7.21.5.1.4 Psychotic symptoms and paraphrenia

A variety of delusions, often related to progressive dementia, are seen in patients in nursing homes. More frequent are the paranoid or accusatory behaviors. Late-onset schizophrenics, in contrast to early-onset schizophrenics, have an average life expectancy that seems to exceed that of the regular elderly population (Bienenfeld, 1990). Pharaphrenia, often regarded as late-onset schizophrenia, is usually described as having onset after the age of 60 and entailing paranoid delusions with or without auditory hallucinations, less thought disorder, and less flattening of affect.

Paraphrenia, which has a relatively high incidence among nursing home residents, can also be regarded as a delusional disorder characterized by an isolated delusion and no other symptoms, and with the delusion not necessarily signaling a neurological disorder (American Psychiatric Association, 1994). Delusions in older patients, according to Post (1987), take the form of "unshakable false belief" with remarkably "banal" content. Psychological diathesis for paraphrenia often include premorbid schizoid personality disorder (Pearlson et al., 1989). Traumas earlier in life, such as parental separation, being a refugee, and childlessness, may play a role in increasing vulnerability (Gurian, Wexler, & Baker, 1992), whereas recent stress may trigger the onset of the disorder. Residents who have immigrated from other countries seem susceptible to developing delusional disorders (Ruskin, 1990). Molinari and Chacko (1983) suggest that the development of paranoid symptoms in the elderly patient is usually the result of increased use of projection as a defense against grief or perceived loss of autonomy and control, typical occurrences in nursing homes. Many residents' ideation was often related to past experiences or unresolved conflictual relationships. Mr. G., aged 82, a frail elderly, blamed his wife of adultery since she became involved in the home's activities, while he remained in the room. In an other case, a widowed woman whose two daughters died in the last 10 years and who also suffered from hearing loss and eye sight deterioration, developed a paranoid ideation that her neighbor was stealing from her. These symptoms may defensively protect the resident from painful depression that some individuals experience as a reaction to loss and dependency. Some researchers have reported that hearing loss is associated with late-onset schizophrenia or paraphrenia and frequently among dementia patients with paranoid symptoms (Gurian et al., 1992). Paranoid symptoms may thus develop as a result of mental deterioration reflecting prethreshold organic brain change, sometimes in order to fill the gap created by impaired memory (Gatz, et al., 1996).

Late-onset psychotic symptoms often have an organic etiology. Therefore, a thorough medical and laboratory evaluation, together with a psychological one, should be obtained. A useful rule of thumb when dealing with psychotic symptoms in the nursing home is to ask to what degree these symptoms are disturbing to the patient or to the environment. Often such states are best treated nonchemically, late-onset schizophrenia does have more preservation of personality and less formal thought disorder then early-onset schizophrenia. If psychotic symptoms are disturbing, antipsychotic medication and ongoing psychiatric involvement is certainly called for. Case management interventions, such as encouraging the patient to remove from people or situations which stimulate the paranoid ideation, could be helpful. The role of the geropsychologist to could be to establish an alliance with the patient (Fried & Agassi, 1976), thus encouraging them to share thoughts about present and past experiences. When such communication can be established, the patient becomes less anxious and an ongoing treatment plan, with the consulting psychiatrist at the home, can be established.

7.21.5.1.5 Personality disorders

Personality disorders are discussed in Chapter 12, this volume, This discussion is limited to highlighting aspects relevant to nursing homes. Given the residents' advanced age, difficulties in diagnosis become apparent. Properly diagnosing personality disorders is further encumbered by the impact of the institutional environment, changes in cognitive processes, sorting out chronic from acute behaviors, and difficulties in obtaining history. A lifetime pattern of pathological interpersonal relationships, the present old-age experience of losses, and institutionalizational rules, all further promote regressive psychological functioning and may elicit primitive defenses such as splitting, all of which gain a special intense quality in an environment with multiple caregivers. Personality disorders are often seen in conjunction with other disorders, in particular major or masked depressions, frequent in nursing homes. Residents with personality disorders are especially ill-equipped to cope with the stress of institutionalization. Many of the "difficult" or "uncooperative" residents in a nursing home can be diagnosed as personality disorders, sometimes coupled with cognitive decline. Such residents may manifest inappropriate emotional responses or inconsistent control of impulse behavior, including aggression and violence. Sadavoy and Fogel (1992) suggest that these patients tend to represent a vulnerable personality that becomes increasingly less adaptive with new stressful events in old age due to eight critical age-specific stressors: interpersonal loss, physical disability, loss of external manifestations of the self (beauty, strength), loss of role, loss of defensive outlets; forced reliance on caretakers, confrontation with death; and conflict over the wish to live vs. the wish to die. Another major loss for such residents is the loss of resources of narcissistic gratification from others. These difficulties are often manifested in distorted ways such as self-perception oblivious to one's impact. This affects interaction with other residents and staff, and may lead to constant conflict with management. Difficulties in these individuals often result from the state in which they have to be dependent on others and to rely on them to supply their needs. They may fight for independence and autonomy in order to maintain their self-esteem by manipulating the staff, by refusing to obey rules, and acting out their anxieties and conflicts.

In spite of the difficulties with such patients, they may benefit from regular psychotherapeutic contact and are able to establish an alliance with the therapist. The relationship with the psychotherapist itself may gratify narcissistic needs which otherwise might have been expressed destructively. Work with the patient's family and staff caretakers is also indicated to introduce possible case management procedures shielding such a patient from stress or encouraging real and/or perceived autonomy and control (Perlmuter & Monty, 1979).

Although the major symtomatology has been outlined, it should be emphasized that the more "benign" disturbances and neurotic behaviors do not receive due attention, as if neuroticism disappears in old age and the elderly suffer only from cognitive difficulties, dementias, and psychosis. Reality is completely different; neurotic behavior and existential concerns are "there to stay," can be treated (including with psychotherapy), and present a major component in the mental health of the elderly in general and in nursing homes in particular. Other disturbances prevailing in nursing homes, and not discussed due to lack of space, include anxiety and somatoform disorders (Stanley and Beck, this vol.), drug abuse (Blow, this vol.), sleep disorders (Ancoli-Israel, this vol.), impulse control wandering, behavioral problems, family issues, further neurotic and psychotic disturbances, as well as surrogate decision-making for nursing home residents. Having emphasized that health problems do not receive adequate treatment, a comprehensive treatment model is presented to address all the explicated mental health issues discussed.

7.21.6 A CLINICAL GEROPSYCHOLOGICAL SERVICE IN NURSING HOMES: A PROPOSED MODEL

A geropsychological clinic operating in a nursing home in Isreal is described (Lomranz, 1991) and presented as a model clinic.

7.21.6.1 Setting and Clientele

The clinic operates within a nursing home housing a population of 750 residents. A psychiatrist serves as consultant. The elderly residents range from those functioning independently to those needing constant nursing care. The clinic is a recognized internship center for clinical psychology interns and is associated with the Tel Aviv University's Herczeg Institute on Aging. The target clientele consists of two main categories: the residents themselves and associated significant figures, mainly family, and the professional and paraprofessional staff. The clinic's framework includes regular staff meetings, case conferences, supervision, study groups, and seminars, as well as participation in many of the nursing home conferences, including referral conferences, admission committees, and meetings planning activities for residents.

7.21.6.2 Orientation and Philosophy

The philosophy constitutes the *zeitgeist* which prevails in the clinic and should be expressed in interaction with clientele, above and beyond the "professional approach." It is reflected in: (i) a dual focus on the image of a person, a Bo-psycho-social holistic view (Engel, 1977) embedded in a humanistic approach and respect for personal dignity, and a life-span developmental psychology orientation (P. Baltes & M. Baltes, 1993). The two orientations are eminently compatible. (ii) Comprehension of the environment and the significance of person/environment interaction; (iii) Aim: to improve mental health and upgrade the quality of life; (iv) Methods: The service adheres to a dynamic theoretical approach based on elaborations of clinical adult developmental theory (Nemiroff & Colarusso, 1990) and other eclectic clinical techniques of diagnosis and treatment. It has often been observed that mental health problems in nursing homes may go unnoticed for long periods of time. It is often only when a resident "becomes a nuisance" that they are dealt with. Early detection may improve the utility of treatment. It is therefore crucial that the service adheres to outreach, prevention, and community orientations.

7.21.6.3 Methods

The eclectic approach is tailored to the specific needs of the client. These approaches include dynamic, cognitive, and behavior modification, group methods, family and milieu systems, existential therapy, organizational, educational, and community approaches. Specific techniques found useful with the elderly are employed, including life review therapy, reminisence groups (Haight & Webster, 1995), family counseling (Brink, 1979), and educational programs and training (Garrison, 1978). These are applied to individuals, families, groups, community, staff, and interdisciplinary networks. Group work (e.g., psychotherapy, family, task groups, life review) is a major methodological component (MacLennan, Shura, & Weiner, 1988). Intervention are geared to eliminate psychopathology as well as preventive work to increase clients' sense of control and hopefulness about change. Reserve capacities are utilized, based on selection, optimization, and compensation (Baltes & Baltes, 1993; Ryff, 1989). Considerable attention is paid to assessment and diagnostics (Storandt, Siegler, & Elias, 1978). Existing diagnostic tools are adapted or modified, and specialized tools are developed in cooperation with the University Herczeg Institute on Aging. Thus, traditional and nontraditional approaches for intervention and assessment are employed. The psychologists work in cooperation with social workers, occupational, art- and physiotherapists. Special attention is given to psychosoma interaction, and ongoing communication is maintained with the medical and paramedical staff. Regular contact is held with the "Home mother," the caregivers, and technical departments (e.g., maintenance, cleaning, kitchen). Informal contact and interaction (discussions in the elevator, garden, etc.) are part of the approach and highly valued. All these activities also contribute to the integration of the clinic as an integral part of the nursing home.

7.21.6.4 Applications

Most referrals originate from interdisciplinary staff meetings, medical and social services, or individuals recommending neighbors. A minority are self-refered. Three principles of psychological interventions of a curative, preventive or palliative nature are categorized and for the sake of convenience, termed: (i) "treatment," basically consisting of a direct, short-term, or long-term systematic psychotherapeutic approach applied to different kinds of psychopathologies; (ii) "consultation," attending transitional, normative, or situational crises. This is mostly accomplished directly, but often indirectly (e.g., instructing a nurse, neighbor): short consultation or direction of other staff are central here; (iii) "prevention," consisting of two main activities of working with residents or staff on problematic issues or risk populations (e.g., groups work with isolated residents just before major holidays, or tackling staff problems of burnout) or of educational activities aimed at enriching the residents or furthering the professional development of the staff (e.g., lectures to residents on mental health or to staff on psycholpathology).

Thus, the psychological services in the nursing home comprises seven domains with their respective working goals of mental health.

7.21.6.4.1 Preadmission screening

Every applicant is interviewed and assessed to obtain information on cognitive as well as affective functioning and status. Thus, the service can be advise management in decisions to accept or reject an applicant. For those accepted, preadmission screening may be paramount in certifying the type of care and living facilities that the elderly requires (Smyer, 1989). While most people admitted to a nursing home have mental concerns and some physical problems, it is generally the latter which are documented and communicated to the staff, while mental or behavioral problems are rarely discussed professionally. At times, there is a tendency for the family and the older client to withhold information regarding deterioration of mental state. Such clients, once admitted, exhibit serious adjustment difficulty and, moreover, their cognitively impaired as well as affective state may deteriorate more rapidly. The psychologist can support the client and the family and can help reduce their initial anxiety, obtain their cooperation in assessing adequately the client and as a result prevent unnecessary misery and rapid decline in mental health.

7.21.6.4.2 Follow-up services

Working on the basis of data obtained in prescreening, the psychologist is later involved in assessing the follow-up of the new resident. Some residents need supportive sessions or crisis intervention, short-term therapy, focused on issues of separation and mourning over the losses they experience at the early phases of their transition to the nursing home. Others may be directed to task groups or activities in the home. The service conducts open-ended adjustment groups which focus on relocation effects during the first three to six months, a group in which every newcomer participates for at least several sessions.

7.21.6.4.3 Severe psychopathology

Here the clinic functions as any other psychiatric setting. Patients are refered and diagnosed, and treatment is planed and carried out. Often, combined psychopharmacological and psychological treatment is administered. Individual psychotherapy is increasingly being requested by residents.

7.21.6.4.4 Family interventions

A resident's family always occupies them in practice or in thought. The elderly person usually arrives at the nursing home accompanied by family members, with the adult children formally assuming "filial" responsibility, acting as "parents" as they relate to their parent and deal with staff and management. This role reversal is often likely to heighten the anxieties and difficulties of all parties concerned: children, parents, and staff. As some residents experience separation anxieties and relocation symptoms are manifested, the anxious, or guilt ridden, children may respond by engaging in acting-out behaviors. They may cause inconvenience, aggravate staff, and exhibit overdo concern. Parents, in turn, increase their fears and guilt by complaining, crying, expressing helplessness, and despair. This was the case with Mr. R., aged 88, who, prior to entering the home was functioning well, spending his time painting and sculpturing. He was forced to move to the nursing home after falling and became unable to take care of himself. He was the father of two devoted sons who encouraged him and helped him move to the home. Mr. R. seemed quiet and reserved in the admission interview. However, as soon as his sons left he became extremely anxious and restless. During his first week in the home he spent most of the day wandering around the phone box crying and calling his sons, asking them to visit. He nagged staff and other passers-by near the phone box to call his sons. His sons became concerned and paid visits a few times a day. Family sessions were held in which the children comprehended the father's separation anxieties and his intensive need for support. The sons were encouraged to talk about their own anxieties and work through their ambivalence at "launching" their father. The psychologist presented a treatment plan for the father with guidelines for the family and the nursing staff to help the father settle down and slowly shift towards self-control instead of enacting his need to control his sons. He gradually learned how to master his anxieties, his sadness declined, and his phone calls ceased.

Mr. G., 80-years old, lived in the nursing home for five years. His relationship with his only son was completely cut off, each accusing the other of ill-treating the mother/wife who passed away prior to Mr. G's move to the home. On memorial day each would come to the cemetery at a different hour. Mr. G. appeared to be well adjusted in the home until he developed a prostrate condition and feared he would die. At that point he began to long for his son who, stubbornly, refused to see him. The son had 20 hours of telephone discussions with one of the psychologists, during which he voiced his anger, discussed the role his mother played in his life, and his hate and ambivalence to his father. Only afterwards was a meeting arranged in which the two brought flowers to the mother's grave and joint treatment began.

7.21.6.4.5 From education to psychotherapy

Extending educational programs to residents and their families provides a channel leading residents to self-development. Psychological education is often a condition for various ethnic, economic, and less-educated residents to become utilizers of the psychological services.

Many elderly suffering in nursing homes may find it difficult to discuss personal issues and emotional problems with others, will not take the initiative to obtain psychological help, and may also refuse it when offered. One way to combat such difficulties is through offering relevant information, developing the proper atmosphere, and providing opportunities to meet therapists. Different psychologists and interns of the service provide a monthly lecture, open to residents, their families, and staff, on various aspects of the aging process, general psychology and adjustment issues in nursing homes. The psychologist then leads a discussion encouraging the audience to ask questions and to share ideas, thoughts, and feelings.

Mrs. J. approached the psychologist after a lecture on intergenerational relationships, seeking advice for problems her granddaugther has with her father. In the discussion she revealed her intense identification with the granddaughter voicing anger at her son. This became a basis for her visits to the service. She then focused on personal problems and remained in psychotherapy for eight months. Mrs. W., a highly educated 92-year-old, wished to continue the lecture on "successful aging" with the "young professor" (professional status being usually important, and overrated by elderly). This provided her with the opportunity to gain respect for her own knowledge and experience and eventually led to an extended and structured life review, after which she became more accepting of her present condition and learned new social skills. Obstacles in approaching the psychologist also reside in prejudice and stigmas about mental care, common in the older population. Public lectures, or special workshops (e.g., in communication skills, memory techniques) help the elderly, prospective client reduce anxieties and shame. Such activities, to most of which family members are also invited, may also serve as a vehicle of outreach to enhance constructive family involvement.

7.21.6.4.6 Working with nursing home staff

This is a primary goal in the work of the clinic. The clinic administers intervention programs aimed at reducing staff distress, professional upgrading, promoting team work (e.g., to eliminate stress between workers in the kitchen), assisting conflict resolution between various teams and between them and management, and preventing burnout.

Nursing homes often have high turnover rates and low staff morale, resulting from work pressures in dealing with threatening issues (aging, loss, illness, and death), distorted perceptions, and low intrinsic and extrinsic rewards. Psychologists in the service work with the interdisciplinary teams providing guidance, information, knowledge, and help in management of residents (e.g., how to address residents' hostile behaviors towards them), in coping with personal problems, and symptoms of burnout (increased fatigue, apathy, etc.). Workshops are conducted on improving interpersonal communication among the staff aiding managers to establish positive reinforcement techniques (Sbordone & Sterman, 1983), improving problem-solving on team building, and preventing burnout. Working educationally with different ranks of staff, the psychologist also conducts discussions, lectures, and case conferences with the various teams (e.g., medical, maintenance workers). Most interventions are based on short-term programs of two to six sessions.

The geropsychologist should project an independent, often "neutral," professional position in the home, thus promoting authority and the ability to work successfully with the various teams, management, and residents. The clinic has witnessed turnovers and shifts in staff roles as well as changes of management often resulting in unclarity and uncertainty as to structure and management. In one such instance staff had major difficulties adjusting to a transition period between managers. Discussions with the psychologist, working as an organizational consultant, revealed an accumulation of covert anger and coalitions between different team members detrimental to ongoing work. Often, in meetings, issues concerning complains about residents reveal conflicts between different teams of the staff. In another case, a nursing team became almost ineffective. It turned out that the team was understaffed, frustrated by hard work and lack of appreciation and recognition by management, and rapidly revealed burnout syptomatology. Group discussions combined with private sessions helped them release their feelings, work through some of the communication patterns with management, and address their complaints effectively. Some of the clinic's staff are also called for consultation to other nursing homes in the vicinity.

7.21.6.4.7 A university clinic

Coordinating links with the university's institute of aging and the nursing home, teaching and supervising students and interns, and cooperating in research are all part and parcel of the clinic. The staff includes postgraduate clinical psychologists who work as interns in the nursing homes, and provide supervision to undergraduate and graduate students of clinical psychology. Undergraduate students participate in a special practicum program involving them in activities which may range from friendly visits to more structured intervention, participation in group work, and providing lectures to the residents. Such liaisons with academic settings are beneficial to all parties involved and upgrade the quality of the home (Smyer & Garfein, 1988). They also encourage young students and interns to choose aging as their field of specialization (Kahana, 1986).

This section concludes with some general evaluative remarks. During two years of operation, residents have changed their attitudes to psychological services. They became less skeptic, have come to recognize the importance of the clinic, and make greater use of its services. Experience suggests the recommendation that as geropsychologists, even intermediately until clinics are set up, working as individuals in nursing homes, they should: (i) establish their position as consultants in a variety of roles; (ii) dwell on diagnostics and assessment; (iii) become major referal directors; (iv) operate as coordinators between teams; (v) be clinical-community oriented; (vi) advance milieu treatment; (vii) assert their impact on policy-making in the institute; (viii) plan and utilize interdisciplinary interventions; (ix) be liberal in using eclectic, traditional, as well as untraditional, intervention approaches, including psychotherapy; and (x) invest in educational programs and systematically accumulating knowledge through documentation and research.

The possibilities of psychotherapy in nursing homes should be emphasized as a challenge for every clinical psychologist. The goals of psychotherapy should be modified and theoretically they should consider gerontology, adult developmental processes, cognitive and medical losses and the effects of institutionalization, as well as clinical methodology. The dialectics between "case specificity" and "age specificity" have to be analyzed. Short-term treatment is often required (Brink, 1979; Gallagher & Thompson, 1982), and a relational approach to transference seems conducive (Frankel, 1995). In long-term therapy, sessions could be of short duration and may be conducted in the residents' personal environment (their room, or in the ward). Long-term psychotherapy often provides the resident with continuity assuring them of the therapist's availability as a secure base (Bowlby, 1980). Many problems of the aged person are fit for intrapsychic and interpersonal exploration and lend themselves to life review, development, and growth. While psychotherapy is always tailored to the specific case, its possible gains with the aged are (in the proper perspective) as unlimited as they are with younger populations. Psychotherapy should be considered as a viable option for the elderly (Shmotkin, Eyal, & Lomranz, 1989). The pleasure and growth in working with the suitable aged can be tremendously gratifying for all parties.

7.21.7 FUTURE PROSPECTS

There is undoubtedly an increasing need for mental health services in nursing homes. As the nursing home industry develops, this need intensively grows. Existing mental health facilities need to be expanded and improved, urgently, to answer neglected and increasing needs. Unfortunately, psychologists, not to mention geropsychological clinics, rarely are found in nursing homes. Also, the nursing home setting is ill-prepared to accept and employ psychologists. The absence of systematic assessment procedures, diagnosis and evaluation, consultations, therapy, and other appropriate interventions have a high cost. Unfortunately, there is a pronounced shortage of psychologists in gerontological clinical work in general (Gatz, 1995; Gatz et al., 1985). Clinical psychology as a discipline is still largely child-oriented and continues to reflect "professional ageism" (Butler & Lewis, 1982), especially with respect to nursing homes. The many misconceptions about the the older person (e.g., lack of energy, motivation, rigidity, and resistance to change) still remain after many decades and affect the "reluctant psychologist" who is unwilling to work with the elderly. Working with the elderly still seems to be rated by colleagues as having low professional status (Knight, Reinhart, & Field, 1982), all the more so in a nursing home. Psychologists may also tend to shrink from interdisciplinary collaboration and conflict (e.g., with psychiatrist, social workers, or other paraprofessionals). Working in a nursing home is often regarded as depressing and difficult since illness and death, helplessness, and hopelessness are confronted continuously. What then are the future prospects?

To meet the crucial challenge, a broader perspective should be adopted which associates geropsychology more strongly with clinical psychology at large and considers policy issues (Kane & Kane, 1990). Basic changes are recommended as stated in the following: (i) the discipline of clinical psychology has to further modify its developmental approach to include adulthood and aging; (ii) theory construction in clinical gerontology while thoroughly based on gerontology also needs to further integrate prevalent clinical theory (e.g., object relations); (iii) clinical methodology in geropsychology, on the aforementioned basis, should be eclectic and modified in goals and methods to fit the elderly population; (iv) policy should be further planned and policy makers have to present, perhaps more aggressively, the damage to nursing homes in cost-effective terms and develop reimbursement systems; (v) academic education and training has to be expanded intensively to include clinical geropsychology; (vi) governmental regulations should include the request for in-home psychological services in nursing homes as obligatory; and (vii) nursing homes should become academic-associated institutions, and clinics, as certified training placements and research units, should operate in each.

The major goal of the clinical gerontologist in nursing homes should be to contribute to a fundamental change which would render the various institutions as authentic "homes" in which residents could actualize their potentials for positive mental health, live with dignity, enhance their well-being, and live creatively. For that purpose, which to a certain extent will determine the kind of society that is built, the discipline of psychology, on its various domains, is expected to mobilize itself to change and meet that challenge.

7.21.8 REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Baltes, M. M. (1982). Environmental factors in dependency among nursing home residents: A social ecology analysis. In T. A. Wills (Ed.), *Basic processes in helping relationships*. New York: Academic Press.

Baltes, P. B., & Baltes, M. M. (Eds.) (1993). *Successful aging: Perspectives from the behavioral sciences*. Cambridge, UK: Cambridge University Press.

Bar-Tur, L. (1993). Adjustment to home for the aged as a developmental crisis in old age. *Gerontologia, 62,* 48–57.

Bienenfeld, D. (Ed.) (1990). *Verwoerdt's clinical geropsychiatry* (3rd ed.). Baltimore: Williams & Wilkins.

Birren, J. E., & Renner, V. J. (1980). Concepts and issues of mental health and aging. In J. E. Birren (Ed.), *Handbook of mental health and aging* (pp. 3–33). Englewood Cliffs, NJ: Prentice-Hall.

Blazer, D., Burchett, B., Service, C., & George, L. K. (1991). The association between age and depression among the elderly: An epidemiological exploration. *Journal of Gerontology: Medical Sciences, 46,* M210–215.

Blazer, D. G., Hughes, D. C., & George, L. K. (1987). The epidemiology of depression in an elderly community population. *Gerontologist, 27,* 281–287.

Bowlby, J. (1980). *Loss*. New York: Basic Books.

Brink, T. L. (1979). *Geriatric psychotherapy*. New York: Human Sciences Press.

Burns, B., Wagner, H. R., Taube, J. E., Magaziner, J., Permutt, T., & Landerman, L. R. (1993). Mental health service use by the elderly in nursing homes. *American Journal of Public Health, 83,* 331–337.

Butler, R. N. (1975). *Why survive? Being old in America*. New York: Harper & Row.

Butler, R. N., & Lewis, M. (1982). *Aging and mental health*. London: Merrill.

Cohen, G. D. (1980). Prospects for mental health and aging. In J. E. Birren & R. Sloane (Eds.), *Handbook of mental health and aging* (pp. 971-993*)*. Englewood Cliffs, NJ: Prentice-Hall.

Cohen, C. (1991). The chronic mentally ill elderly: Service research issues. In E. Light & B. Lebowitz (Eds.), *The elderly with chronic mental illness* (pp. 194–219). New York: Springer.

Cohen-Mansfield, J. (1996). *Trends in the treatment of frail elderly persons*. Paper presented in the Herczeg Research Institute on Aging, Tel Aviv University, Israel.

Colerick, E. J., & George, L. K. (1986). Predictors of institutionalization among caregivers of patients with Alzheimer's disease. *Journal of the American Geriatrics Society, 34,* 493–498.

Donahue, W. (1978). What about our responsibility for the abandoned elderly? *The Gerontologist, 18,* 102–111.

Eisdorfer, C., & Cohen, D. (1982). *Mental health care of the aging: A multidisciplinary curriculum for professional training*. New York: Springer.

Engel, G. (1977). The need for a new medical model: A challenge for biomedicine. *Science, 196,* 129–136.

Everitt, D. E., Field, D. R., Soumerai, S. S., & Avron, J. (1991). Resident behavior and staff distress in the nursing home. *Journal of American Geriatrics Society, 39,* 792–798.

Farrari, N. (1963). Freedom of choice. *Social Work, 8,* 105–106.

Firestone, I. J., Lichman, C. M., & Evans, J. R. (1980). Privacy and solidarity: Effects of nursing home accommodation on environmental perception and sociability preferences. *International Journal of Aging and Human Development, 1,* 221–241.

Frankel, S. (1995). *Intricate engagements*. London: Jason Aronson.

Freedman, V. (1996). Family structure and the risk of nursing home admission. *Journal of Gerontology: Social Sciences. 51B,* S61–S69.

Fried, Y., & Agassi, J. (1976). *Paranoia: A study in diagnosis*. Dordrecht, The Netherlands: Reidel.

Gallagher, D. E., & Thompson, L. W. (1982). Treatment of major depressive disorders in older outpatients with brief psychotherapies. *Psychotherapy: Theory, Research and Practice, 19,* 482–490.

Garrison, J. (1978). Stress management training for the elderly: A psychoeducational approach. *Journal of the American Geriatrics Society, 26,* 397–403.

Gatz, M. (1989). Clinical psychology and aging. In M. Storandt & G. R. VandenBos (Eds.), *The adult years: Continuity and changes* (pp. 79-114). Washington, DC: American Psychological Association.

Gatz, M. (Ed.) (1995). *Emerging issues in mental health and aging*. Washington, DC: American Psychological Association.

Gatz, M. (1998). Towards a developmentally-informed theory of mental disorder in older adults. In J. Lomranz (Ed.), *Handbook of mental health and aging*. New York: Plenum.

Gatz, M., Kasl-Godley, J. E., & Karel, M. J. (1996). Aging and mental disorders. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed.). San Diego, CA: Academic Press.

Gatz, M., Popkin, S., Pino, C., & VandenBos, G. R. (1985). Psychological interventions with older adults. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 755–785). New York: Van Nostrand Reinhold.

Gatz, M., & Smyer, M. A. (1992). The mental health system and older adults in the 1990s. *American Psychologist, 47,* 741-751.

Glass, J. C., & Reed, S. E. (1993). To live or to die: A look at elderly suicide. *Educational Gerontology, 19* (8), 767–778.

Goffman, E. (1961). *Asylums: Essays on the social situation of mental patients and other inmates*. New York: Anchor.

Goplerud, E. (1979). Unexpected consequences of deinstitutionalization of the mentally disabled elderly. *Community Psychology, 7,* 315–328.

Groger, L. (1994). Decision as a process: A conceptual model of black elders nursing home placement. *Journal of Aging Studies, 8,* 77–94.

Grossberg, G. T., Hassan, R., & Szwabo, P. A. (1990). Psychiatric problems in the nursing home. St. Louis University geriatric grand rounds. *Journal of the American Geriatric Society, 38* (8), 907–917.

Gurland, B. J. (1973). Abroad clinical assessment of psychopathology in the aged. In C. Eisdorfer & M. P. Lawton (Eds.), *The psychology of adult development and aging* (pp. 343–377). Washington, DC: American Psychological Association.

Gurian, B. S., Wexler, D., & Baker, E. H. (1992). Late-life paranoia: Possible association with early trauma and infertility. *International Journal of Geriatric Psychiatry, 7,* 277–284.

Gutmann, D. (1994). *Reclaimed powers: Men and women in later life*. Evanston, IL: Northwestern University Press.

Haight, B., & Webster, J. (Eds.) (1995). *The art and science of reminiscing: Theory, research methods and applications*. London: Taylor & Francis.

Halpert, E. (1991). Aspects of a dilemma of middle age: Whether or not to place aged failing parents in nursing home. *Psychoanalytic Quarterly, 60,* 426–449.

Hansson, R., & Carpenter, B. (1994). *Relationships in old age. Coping with the challenge of transition*. New York: Guilford.

Hirschfeld, M., & Fleishman, R. (1990). Nursing home care for the elderly. In R. L. Kane, G. Evans, & D. Macfadyn (Eds.), *Improving the health of older people—A* *world view* (pp. 473–490). Oxford, UK: Oxford University Press.

Jahoda, M. (1958). *Current concepts of positive mental health*. New York: Basic Books.

Johnson, C. L., & Grant, L. A. (1985). *The nursing home in American society*. Baltimore: Johns Hopkins University Press.

Kahana, B. (1986). *A proposal to establish a Ph.D program in clinical-research gerontology*. Cleveland, OH: Cleveland State University.

Kane, R. L., & Kane, R. A. (Eds.) (1982). *Values and long term care*. Lexington, MA: Lexington.

Kane, R. L., & Kane, R. A. (1990). Health care for older people: Organizational and policy issues. In R. H. Binstock & L. K. George (Eds.) *Handbook of aging and social sciences* (3rd ed., pp. 415–437) San Diego, CA: Academic Press.

Katz, I. (in press). Depression as a pivotal component of secondary aging. In J. Lomranz (Ed.), *Handbook of mental health and aging*. New York: Plenum.

Knight, B., Resnhart, R., & Field, D. (1982). Senior outreach services: A treatment oriented outreach team in community mental health. *The Gerontologist, 22,* 544–547.

Knight, B. G., Santos, J., Teri, L., & Lawton, M. P. (1995). Introduction: The development of training in clinical geropsychology. In B. G. Knight, L. Teri, P. Wohlford, & J. Santos (Eds.), *Mental health services for older adults: Implications for training and practice in geropsychology*. (pp. 1–8). Washington, DC: American Psychological Association.

Lawton, M. P. (1980). Housing the elderly: Residential quality and residential satisfaction. *Research on Aging, 2,* 309–328.

Lawton, M. P. (1985). Housing and living environments of older people. In R. Binstock & E. Shanas (Eds.), *Handbook of aging and social sciences* (pp. 450–478). New York: Reinhold.

Lawton, M. P., & Cohen, J. (1974). The generality of housing impact on the well-being of older people. *Journal of Gerontology, 29,* 194–204.

Lewis, M. A., Cretin, S., & Kane, R. L. (1985). The natural history of nursing home patients. *Gerontologist, 25,* 382–388.

Lieberman, M., & Tobin, S. (1983). *The experience of old age. Stress, coping and survival*. New York: Basic Books.

Lomranz, J. (1991). Mental health in homes for the aged and the clinical psychology of aging: Implementation of a model service. *Clinical Gerontologist, 10,* 47–71.

Lomranz, J. (1995). Thus we live forever taking leave: The launching stage in the Israeli family. *Social Behavior and Personality*, 28, 287–302.

Lomranz, J., Shmotkin, D., Eyal, N., & Zohar, Y. (1996). Launching themes in Israeli fathers and mothers. *Journal of Adult Development, 3,* 159–170.

Lomranz, J., Shmotkin, D., & Vardi, R. (1991). The equivocal meanings of time. *Current Psychology: Research and Reviews,* 10 (1 & 2), 3–20.

MacLennan, B., Shura, S., & Weiner, M. (Eds.) (1988). *Group psychotherapies for the elderly*. Madison, WI: International Universities Press.

Molinari, V., & Chacko, R. (1983). The classification of paranoid disorders in the elderly: A clinical problem. *Clinical Gerontologist, 1,* 31–37.

Moos, R., & Lemke, S. (1985). Specialized living environments for the older people. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 864–889). New York: Reinhold.

National Center for Health Statistics (1983). *The national nursing home survey: 1977. Summary for the United States*. Washington, DC: Government Printing Office.

Nemiroff, R., & Colarusso, C. (Eds.) (1990). *New dimensions in adult development*. New York: Basic Books.

Neulinger, J. (1981). *To leisure: An introduction*. Boston, MA: Allyn and Bacon.

Newman, F. L., Griffin, B. P., Black, R. W., & Page, S. E. (1989). Linking level of care to level of need: Assessing the need for mental health care for nursing home residents. *American Psychologist, 44,* 1315–1324.

Newmann, J. P., Engel, R. J., & Jensen, J. E. (1991). Changes in depressive symptom experiences among older women. *Psychology and Aging, 6,* 212–222.

Norris, M. P., Snow-Turek, A. L., & Blankenship, L. (1995). Somatic depressive symptoms in the elderly: Contribution or confound? *Journal of Clinical Geropsychology, 1,* 5–17.

Pachana, N., Gallagher-Thompson, D., & Thompson, L. W. (1994). Assessment of depression. In M. P. Lawton & J. A. Teresi (Eds.), *Annual review of gerontology and geriatrics: Focus on assessment techniques* (vol. 14, pp. 234–256). New York: Springer.

Palmore, E. (1976). Total risk of institutionalization among the aged. *Gerontologist, 16,* 504–507

Pearlson, G. D., Kreger, L., Rabins, P. V., Chase, G. A., Cohen, B., Wirth, J. B., Sclaepfer, T. B., & Tune, L. E. (1989). A chart review study of late-onset and early-onset schizophrenia. *American Journal of Psychiatry, 146,* 1568–1574.

Perlmuter, L., & Monty, R. (Eds.) (1979). *Choice and perceived control*. Hillsdale, NJ: Erlbaum.

Pfeiffer, E. (1977). Psychopathology and social pathology. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 650–671). New York: Van Nostrand.

Post, F. (1987). Paranoid and schizophrenic disorders among the aging. In L. L. Carstensen & B. A. Edelstein (Eds.), *Handbook of clinical gerontology* (pp. 43–56). New York: Pergamon.

Reinardy, J. R. (1992). Decisional control in moving to a nursing home: Postadmission adjustment and well-being. *The Gerontologist, 32,* 96–103.

Rovner, B. W., German, P. S., Brant, L. J., Clark, R., Burton, L., & Folstein, M. F. (1991). Depression and mortality in nursing homes. *Journal of the American Medical Association, 265,* 993–996.

Rovner, B. W., & Rabins, P. V. (1985). Mental illness among nursing home patients. *Hospital Community Psychiatry, 36,* 119–120.

Rowles, G. D. (1979). The last new home. In S. M. Golant (Ed.), *Location and environment of elderly population.* Washington, DC: Winston

Rodin, J., & Langer, E. (1977). Long term effects of a control-relevant intervention with the institutional aged. *Journal of Personality and Social Psychology, 35,* 897–902.

Ruskin, P. (1990). Schizophrenia and delusional disorders. In D. Bienfeld (Ed.), *Verwoerdt's clinical geropsychiatry* (pp. 125–136). Baltimore: Williams & Wilkins.

Ryff, C. D. (1989). Beyond Ponc de Leon and life satisfaction: New directions in quest of successful aging. *International Journal of Behavioral Development, 12,* 35–55.

Sadavoy, J., & Fogel, B. (1992). Personality disorders in old age. In: J. E. Birren, R. B. Sloan, & G. D. Cohen (Eds.). *Handbook of mental health and aging* (2nd ed.) (pp. 433–462). New York: Academic Press.

Sbordone, R. J., & Sterman, L. T. (1983). The psychologist as a consultant in a nursing home: Effect on staff morale and turnover. *Professional Psychology: Research and Practice, 14,* 240–250.

Schaie, K. W., & Willis, S. L. (1996). *Adult development and aging* (4th ed). New York: HarperCollins.

Shmotkin, D., Eyal, N., & Lomranz, J. (1989). *Motivations and attitudes of clinical psychologist regarding treatment of the elderly*. Research report, The Sapir Foundation, Tel Aviv University, Israel.

Smyer, M. A. (1989). Nursing homes as a setting for psychological practice: Public policy perspectives. *American Psychologist, 44,* 1307–1314.

Smyer, M. A., & Garfein, A. (1988). Effective assessment approaches. In M. A. Smyer, M. Cohn, & D. Brannon (Eds.), *Mental health consultation in nursing homes* (pp. 64–81). New York: New York University Press.

Stephens, M. A., Crowther, J., Hobfoll, S., & Tennenbaum, D. (Eds.) (1990). *Stress and coping in later-life families*. New York: Hemisphere.

Storandt, M., Siegler, I., & Elias, M. (Eds.) (1978). *The clinical psychology of aging*. New York: Plenum.

Teri, L. (1991). Behavioral assessment and treatment of depression in older adults. In P. A. Wisocki (Ed.), *Handbook of clinical behavior therapy with the elderly client* (pp. 225–244). New York: Plenum.

Timko, C., & Moos, R. H. (1990). Determinants of interpersonal support and self-direction in group residential facilities. *Journal of Gerontology: Social Sciences, 45,* S184–S192.

Tobin, S. S. (1987). A structural approach to families. In T. H. Brubaker (Ed.), *Aging, health and the family: Long-term care* (pp. 42–55). Newbury Park, CA: Sage.

Tobin, S., & Lieberman, M. (1976). *Last home for the aged: Critical implications of institutionalization*. San Francisco: Jossey-Bass.

US Senate, Special Committee on Aging. (1987). *Developments in Aging, 1986* (Vols. 1–3). Washington, DC: US Government Printing Office.

Wantz, M. S., & Gay, J. E. (1981). *The aging process: A health perspective*. Cambridge, MA: Winthrop Publishers.

Winnicot, D. W. (1953). Transitional objects and transitional phenomena. *Collected papers: Through pediatrics to psychoanalysis, 1958,* (pp. 229–242). New York: Basic Books.

World Health Organization (1964). *Basic documents* (15th ed.). Geneva, Switzerland: Author.

Wright, L. K. (1988). A reconceptualization of the "negative staff attitudes and poor care in nursing homes" assumption. *The Gerontologist, 28,* 813–820.

Zivoni, M. (1993). Intergenerational relationships of the elderly and attitudes to residential facilities. *Gerontologia, 61,* 18–32.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.22

Family Caregiving: Research Findings and Clinical Implications

STEVEN H. ZARIT, ADAM DAVEY, ANNE B. EDWARDS, ELIA E. FEMIA, and SHANNON E. JARROTT
Pennsylvania State University, University Park, PA, USA

7.22.1 INTRODUCTION 499
7.22.2 BASIC ISSUES IN FAMILY CAREGIVING 500
7.22.3 CAREGIVING AS AN ONGOING PROCESS: CAREERS OF CAREGIVERS 501
7.22.4 CAREGIVING STRESS: PROCESSES OF PROLIFERATION AND CONTAINMENT 502
7.22.5 CAREGIVING IN A FAMILY CONTEXT 503
7.22.6 THE PERSPECTIVE OF THE CARE RECIPIENT 503
7.22.7 RESEARCH ON FAMILY CAREGIVING 504
7.22.8 CAREGIVING AND FAMILY RELATIONSHIPS 504
7.22.9 NORMATIVE FAMILY RELATIONSHIPS IN LATER LIFE 504
  7.22.9.1 Family Structure and Caregiving 505
  7.22.9.2 Caregiver–Care Recipient Relationships 506
  7.22.9.3 Social Support from Family and Other Informal Sources 507
7.22.10 SECONDARY STRESSORS: COMPETING ROLES AND STRESS PROLIFERATION 508
7.22.11 OUTCOMES OF CAREGIVING: HEALTH AND MENTAL HEALTH EFFECTS 510
7.22.12 CLINICAL RESEARCH AND SERVICES 513
7.22.13 RESPITE SERVICES 513
  7.22.13.1 Adult Day Care 514
  7.22.13.2 In-Home Respite 514
  7.22.13.3 Institutionalization and its Consequences 515
7.22.14 CLINICAL SERVICES 516
7.22.15 CONCLUSIONS 518
7.22.16 REFERENCES 519

7.22.1 INTRODUCTION

Family care of disabled older people has emerged as one of the most important clinical and social issues in an aging society. Historically the family has been the main source of help to the elderly, but in previous eras when most people did not survive to old age, families' involvement in providing care to parents or grandparents tended to be limited and brief. The demographic revolution of the late twentieth century has resulted in more people surviving to ages when rates of disability are high. As a result, increasing numbers of older people need at least some help to manage activities of daily living. The most frequent source of that help is the family. Providing care to a disabled elder poses considerable challenges to the family. Care needs may place physical strain on caregivers or may impinge on already full schedules of work and family responsibilities. The dependency of a spouse or parent can also intensify long-standing family problems, or overwhelm the psychological resources of caregivers. Formal social and health services can reduce the burden of care on families, but can also introduce new strains or problems. Clinicians can make timely and effective interventions to mobilize the family's resources, address family conflict, and maximize the ability of both the family and the disabled older person to meet the challenge of disability.

This chapter will present an overview of research on family caregiving and clinical interventions to relieve strain on the family. The chapter begins by introducing basic concepts to provide a foundation for examining research findings and clinical approaches. Three main areas of research on family caregiving are reviewed: how family relationships set the context for caregiving and in turn are affected by it, role conflict in caregiving, particularly, between employment and caregiving, and how the stresses of caregiving influence mental and physical health. The chapter concludes by examining clinical interventions and services designed to relieve stress on families and research on their effectiveness.

7.22.2 BASIC ISSUES IN FAMILY CAREGIVING

In this section the emergence of family caregiving as a major social issue is reviewed, and four issues are presented which guide understanding of this phenomenon: (i) caregiving as a career; (ii) stress proliferation and containment; (iii) the family context of caregiving; and (iv) the care recipient's perspective.

Caregiving has become an important concern because of the convergence of several social trends. The developed countries of the world are aging societies, that is, increasing proportions of their populations are 65 years of age and older. In fact, the fastest growing segment of the population is the "oldest old," people aged 80 years or over (Kinsella & Taeuber, 1993). Most older people live independently with few or no physical limitations, but the likelihood of disability increases with advancing age. By age 85 years, 60% of the population has at least some degree of disability which requires regular, ongoing assistance (Zarit, Johansson, & Berg, 1995). Among 85-year-olds with no disabilities, approximately 30% develop problems over the next two years, and another 30% do so in the following two years. Clearly, there is a growing risk of disability at very advanced ages.

How the aging of the population leads to increased risk of disability can be illustrated by the example of Alzheimer's disease. Although this disease can occur in people as young as 40 years of age, the greatest risk of incidence is after 75 (Johansson & Zarit, 1995; Jorm, Kortem, & Henderson, 1987; Kokman, Chandra, & Schoenberg, 1988). When most people in a society did not survive to 65, let alone 75 or 85, a disease such as Alzheimer's had a trivial impact. The dementia syndrome has been described since antiquity, but it has only assumed importance since the 1950s as large numbers of people have survived to the ages of greatest risk. For people surviving to their mid-80s, the cumulative risk of developing dementia due to any cause before death may be as high as 47% (Johansson & Zarit, 1995). The result is that many older people require considerable assistance for a period of time before their death.

Extension of the average life span has meant that people have, on average, a longer period of active life. They are able to live independently into their 70s, 80s, and even in some cases 90s. But the factors which have led to improved life expectancy, better medical care and improved public health, may also be contributing to a prolonging of the period of disability at the end of life. The Expansion of Morbidity Hypothesis (Cassel, Rudberg, & Olshansky, 1992) posits that people now survive longer with chronic diseases and with greater amounts of disability. While unequivocal data to support this hypothesis are not available, it is widely apparent that people with severe disabilities can survive for long periods of time. To cite one example, 50% of a sample of dementia patients (with varying diagnoses) were alive 11 years after the first onset of symptoms, and just under 40% survived more than 14 years (Aneshensel, Pearlin, Mullan, Zarit, & Whitlatch, 1995). This long duration of morbidity is an important factor in the heavy caregiver burden placed on families and society.

The net result of these trends is that more older people need help for longer periods of time than ever before. At the same time, families' resources for providing help have probably diminished somewhat. The most likely person to give assistance is a surviving husband or wife, but that person will also usually be elderly and may have disabilities which limit the amount or type of help that can be provided. Because of historic trends toward smaller family size, there are, on average, fewer children available to share the care of a disabled parent. Daughters and daughters-in-law have previously provided much of the assistance to elders, but with women's participation in the workforce, fewer full-time homemakers are available to take on caregiving responsibilities. Finally, historically high divorce rates in many Western countries weaken family ties and may diminish the amount of help available when a parent becomes disabled. Divorce in the parents' generation may have left strained or weakened ties with children, while divorce among children may result in their having fewer resources to commit to the care of an aging parent (Webster & Herzog, 1995).

These sociodemographic trends provide the context for caregiving. Four important concepts set the stage for discussion of research and clinical interventions: (i) caregiving as an ongoing, evolving process; (ii) stress proliferation and containment in caregiving; (iii) family systems theory; and (iv) the importance of addressing the needs of both caregiver and care recipient.

7.22.3 CAREGIVING AS AN ONGOING PROCESS: CAREERS OF CAREGIVERS

Caregiving involves an ongoing process, in which problems change and the caregiver's perceptions of problems also evolve. Just as the care recipient's disorder can span several years, caregiving also has a long duration. The tendency for studies to capture caregivers only at one point in time often obscures the temporal structure of what has come before and what lies ahead.

To describe this process, Pearlin (1993) has introduced the construct of caregiving careers, that is, that caregiving is a role involving certain predictable expectations and transitions. Using findings from a longitudinal study of caregivers of dementia patients (Aneshensel et al., 1995; Pearlin, 1993), Pearlin and his colleagues suggest three distinct phases in a caregiver's career. First is the period of role acquisition, during which caregivers recognize the growing disability in their relative and make the decision to take on the caregiving role. In progressive disorders such as dementia, this recognition develops gradually. Families frequently will ignore the growing signs of disability or hope that their relative will snap out of it and return to normal. A physician's diagnosis or a critical event, such as the patient getting lost or mishandling a sum of money, often changes how the family views the situation, although caregivers may still fluctuate in their acknowledgment of disability and their relative's need for assistance. The tendency in dementing illnesses for patients to deny their problems or disabilities, especially as the disease progresses (e.g., Kaszniak, 1996), undoubtedly contributes to families' ambivalence in recognizing the disorder and taking on care responsibilities.

The second phase in the caregiving career is role enactment, during which caregivers provide regular, ongoing assistance to their relative. Role enactment includes providing care to people in the home, and also in institutional settings. As we will see below, caregiving does not end at the institution's door. Families remain involved in care. While placement relieves some stresses, it introduces other problems and challenges for the family (Zarit & Whitlatch, 1992). The transition from home to institution may also be one of the most stressful times for families. Most of the available research addresses the period of role enactment, especially concerning caregivers providing care in the home.

The third phase in caregiving is disengagement, which is marked by the patient's death. Although psychological disengagement sometimes occurs sooner, death of the patient is a critical turning point that has complex effects on caregivers. Death of the patient is frequently welcomed when it brings release from a long, painful or deteriorated condition. Yet it also forces caregivers to re-examine and reorient their lives. Grief follows a predictable course, with symptoms greatest following the patient's death, and then gradually diminishing, although some caregivers remain distressed for a sustained period of time (Aneshensel et al., 1995).

Caregiving, then, should not be viewed as a static event that takes place for a limited duration of time. It is an ongoing process with some predictable demands and transitions. The pressures on the family depend to some extent on the phase of caregiving. Likewise, the type of clinical intervention will depend on whether caregivers are assuming the role, are in the midst of performing extensive caregiving activities, or are trying to reconstitute their lives following their relative's death. (For a description of varying clinical interventions over the full course of caregiving, see Zarit & Zarit, in press.)

7.22.4 CAREGIVING STRESS: PROCESSES OF PROLIFERATION AND CONTAINMENT

Caregiving has generally been viewed from the perspective of stress theory. Stressors represent all of the conditions, experiences, and activities that are problematic for people by thwarting their efforts, threatening their well-being, and disrupting otherwise stable systems (Aneshensel et al., 1995; Pearlin, Mullan, Semple, & Skaff, 1990). Chronic stress creates a state of perpetual arousal that taxes the individual's usual ability to adapt to the environment, and so compromises their ability to reach sought-after goals (Lazarus & Folkman, 1984).

A specific model is used to guide exploration of caregiving. The stress process model, developed by Pearlin and his colleagues (Aneshensel et al., 1995; Pearlin et al., 1990) has been used to guide a major longitudinal investigation of caregivers of dementia patients. This model identifies three major components in caregivers' experiences: primary stressors, secondary stressors and outcomes (see Figure 1). Primary stressors evolve directly from the needs of the patient and represent the actual tasks of caregiving. These stressors are called primary because they are intimately associated with the root of the stress, namely, the chronic disabilities of the care recipient (Pearlin et al., 1990). Examples of primary stressors are behavioral problems, memory problems, and difficulties carrying out activities of daily living (ADLs).

Coupled with the objective conditions of caregiving are the caregiver's subjective appraisals of the memory problems, behavioral disturbances, and ADL deficits. Some primary stressors are more problematic or challenging than others and so can affect coping skills and erode caregiver self-efficacy (Kinney & Stephens, 1989). Pearlin et al. (1990) specifically identify feelings of overload and role captivity, that is, a sense of being trapped in the caregiving role, as primary subjective stressors. They also propose that the caregiver's sense of loss of the relationship with the care recipient should be considered a primary stressor.

Evolving out of the primary stressors are the secondary stressors. Although termed secondary, these stressors are of no less importance. They represent the way in which primary stressors encroach upon and disrupt other areas of the caregiver's life, and thus are more variable and individual in their manifestation. Secondary stressors include family conflict, financial strain (e.g., reduced financial resources as a result of caregiving) and work conflict. They also involve psychological dimensions, such as diminished self-esteem and a feeling of loss of self, that the caregiver's identity has been absorbed totally in the caregiving role (Skaff & Pearlin, 1992).

Long-term exposure to stressors can produce a variety of physiological and psychological changes, or what are termed outcomes in the

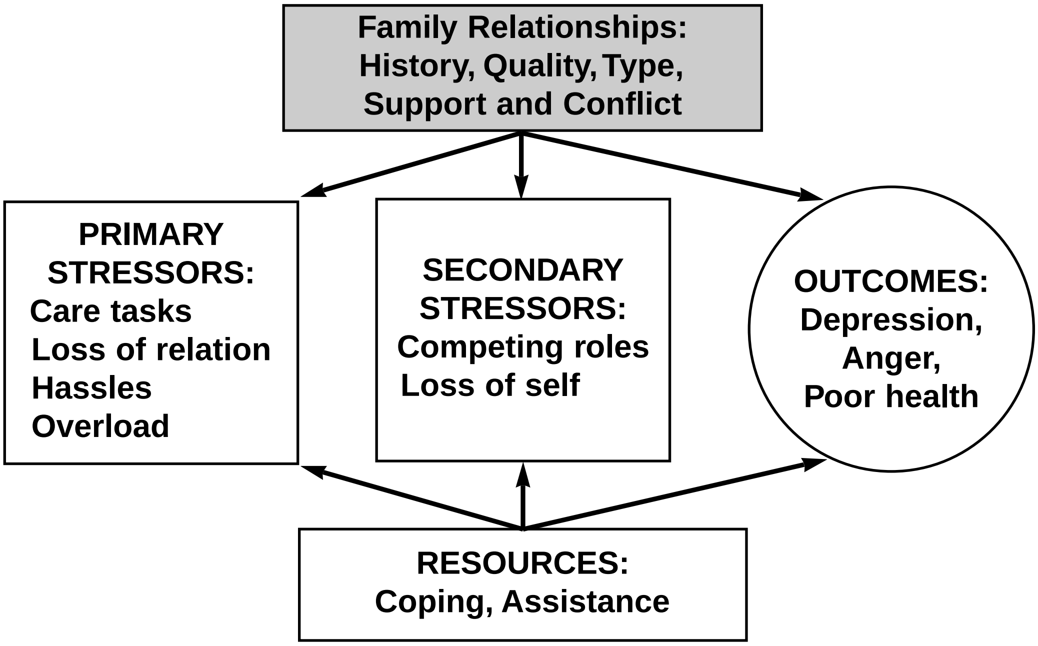
Family Relationships: History, Quality, Type, Support and Conflict
PRIMARY STRESSORS: Care tasks Loss of relation Hassles Overload
SECONDARY STRESSORS: Competing roles Loss of self
OUTCOMES: Depression, Anger, Poor health
RESOURCES: Coping, Assistance

Figure 1 A model of caregiving stress and family relationships.

stress process model. Among the outcomes that have been studied are depression, anger, well-being and a variety of indicators of health. These outcomes occur as chronic stress which results in disruptions of people's homeostatic mechanisms that overwhelm their coping resources.

Stressors, however, do not have a fixed or constant effect on outcomes. The caregiving literature has documented considerable variability in how caregivers adapt to similar circumstances (e.g., Aneshensel et al., 1995; Haley, Levine, Brown, & Bartolucci, 1987; Wright, Clipp, & George, 1993; Zarit, Todd, & Zarit, 1986). The relation of stressors and outcomes is described in the stress process model in terms of two mechanisms: stress proliferation and stress containment (Aneshensel et al., 1995). Stress proliferation describes how the stressors involved in caregiving infiltrate and disrupt other areas of the caregiver's life. Primary stressors lead in varying degrees to secondary stressors, which in combination lead to negative outcomes. Stress proliferation is greatest when caregivers are assisting someone with higher levels of behavior problems and when there is a greater subjective effect of these stressors, for example, when caregivers feel overwhelmed or trapped in their role (Aneshensel et al., 1995; Haley et al., 1987; Zarit et al., 1986). Stress containment, in turn, describes how the caregiver's use of resources may limit or circumscribe the effects of stressors. How caregivers cope with specific problems as well as the amount of social and financial resources available to them can protect them from some of the adverse effects of stress.

The outcome of a particular configuration of caregiving stressors is variable. The amount of variance in caregiver outcomes, such as depression, that is accounted for by any particular stressor, for example, the patient's behavior problems or the amount of family conflict, is statistically significant but relatively modest (Aneshensel et al., 1995). Instead, outcomes depend on the extent of proliferation or containment.

The implications of this framework for clinical intervention are twofold. First, clinicians need to assess which primary and secondary stressors are present in a given situation. Caregiving and caregivers' resources are sufficiently varied so that the constellation of problems and stressors will differ considerably from one family to another. Rather than assuming that a constellation of stressors will be overwhelming or that a caregiver will necessarily experience work or family conflict, clinicians need to conduct an assessment to identify the specific proliferation of stress in the caregiver's life. Second, the goal of clinical interventions is to reduce sources of stress proliferation and enhance the caregiver's resources that lead to containment. Although it may not be possible to change the patient's underlying disability or to alter the primary stressors on caregivers, interventions can modify their impact on the caregiver's life (Whitlatch, Zarit, & von Eye, 1991). In the sections below, research is highlighted which suggests which factors are associated with greater or less stress for family caregivers and which may be amenable to clinical intervention.

7.22.5 CAREGIVING IN A FAMILY CONTEXT

Stress theory is useful for characterizing the problems and challenges faced by caregivers, but it does not take into account the specific family context in which caregiving occurs. The relationship between caregiver and care recipient and their relationships with other family members should be considered. A family systems perspective focuses on what roles and functions a person performed within the family as well as the quality of these relationships (Zarit, Orr, & Zarit, 1985). Because of illness and disability, the care recipient can no longer perform certain activities and functions for the family. A spouse caregiver, for example, may experience the loss of a confidant and sexual partner, and may have to take over activities that his/her spouse previously performed in the household. Besides these specific losses, care recipients and caregivers have varying amounts of prestige and influence in the family, which affects how much help and assistance other people will provide. Children who view a caregiving mother as extremely competent may not view her as needing help. Conversely, a mother who is viewed as needy or lacking competency may get too much help or a mixture of criticism and assistance. Caregiving can also reawaken or intensify long-standing conflicts in the family.

7.22.6 THE PERSPECTIVE OF THE CARE RECIPIENT

Lost in much of the caregiving literature is the perspective of the person receiving care. In part, this omission is due to the extensive focus on providing care to people with dementia, who lose the ability to report their experiences reliably as the disease progresses. This exclusion, however, means that the care recipient's perspective often is overlooked. Care recipients can continue to provide some types of help, assistance, or comfort to the caregiver, even in cases of moderate to severe dementia. The care recipient's well-being and needs should be considered along with the caregiver's, in decisions about living arrangements and care. It is important to consider whose interests are being represented when making a clinical intervention. What is good for the family may, in some circumstances, not be in the patient's best interest. The clearest example is nursing home placement, which relieves strain on families but puts the dementia patient at risk for a variety of adverse outcomes, including an increased risk of mortality (Aneshensel et al., 1995). Though often omitted from research, the patient's perspective needs to be incorporated consistently into clinical intervention.

7.22.7 RESEARCH ON FAMILY CAREGIVING

The following review of the research literature focuses on the stress process and family relationships. It starts by emphasizing caregiving in its family context, and the relationships of caregiver and care recipient to one another and to other family members. One of the main secondary stressors, is then considered, that is, when there is strain or conflict between caregiving and employment. Finally, consideration is given to what is known about the effects of caregiving on physical and mental health and some of the correlates of negative outcomes.

7.22.8 CAREGIVING AND FAMILY RELATIONSHIPS

In this section, attention is focused on caregivers and their family relationships. Two primary themes underlie the organization of this section. First, family caregiving arises within a long-standing, often life-long, relationship. As a result, the caregiving relationship can be seen as an extension of the prior relationship history and its associated pattern of interactions. Second, informal caregiving is best understood when considered as a system, rather than as the characteristics of a single individual or a single relationship. As suggested by the stress-process model outlined above, events within the caregiver–care recipient relationship can have consequences, both positive and negative, for other familial relationships and caregiver well-being.

Despite the vast literature on family caregiving, surprisingly little is known about its effects on family relationships. In this section, the effects of caregiving on the relationship between caregiver and care recipient are considered, as well as the effects of caregiving on relationships outside the primary caregiving dyad, particularly between the caregiver and his or her spouse and siblings. Most is known about caregivers who are the spouses or children of the care recipient. Siblings sometimes take care of a brother or sister, and other relatives and even nonkin may take on the caregiving role (e.g., Cicirelli, 1985, 1992), but little is known about how caregiving affects these relationships.

7.22.9 NORMATIVE FAMILY RELATIONSHIPS IN LATER LIFE

Recognition of the importance of family relationships in later life has spawned a large body of research examining normative patterns of intergenerational exchanges. Overwhelmingly, this literature has demonstrated that older adults typically remain both active donors and recipients of support until very late in life. In this intergenerational context, older adults generally give more to their adult offspring than they, in turn, receive (Eggebeen & Wilhelm, 1995; Zarit & Eggebeen, 1995).

Reciprocity and the balance of exchanges have received particular attention because of their potential influence on psychological well-being (e.g., Ingersoll-Dayton & Antonucci, 1988). Using a large, nationally representative Canadian sample, Hirdes and Strain (1995) found a curvilinear relationship between age and the balance of instrumental exchanges between generations with the greatest balance occurring at midlife. Morgan, Schuster, and Butler (1991) examined the balance of exchange between older parents and their adult offspring as a function of parents' age. They found that parental giving and receiving declined with increasing parent age, but that older adults typically continued to give more than they received beyond age 85. Unlike Hirdes and Strain (1995), who focused only on instrumental exchanges, Morgan et al. (1991) also considered emotional support, which partially accounts for these observed differences.

Related to intergenerational exchanges are norms of filial responsibility. Hamon and Blieszner (1990) report that older parents and their children strongly endorse expectations of filial responsibility in times of need. There is ample evidence that these norms are robust across a wide variety of cultures and family structures (e.g., Brody, Johnsen, & Fulcomer, 1984; Chow, 1983; Fadel-Girgis, 1983; Hanson, Sauer, & Seelbach, 1983; Kendig & Rowland, 1983; Kivett & Atkinson, 1984; Maeda, 1983; Weihl, 1983; Worach-Kardas, 1983).

It appears that these cultural norms are readily translated into action, with relationships with adult children assuming greater importance for older adults in later life. Field and Minkler (1988) report that frequency of contact with adult children remains high from early to late old age. Whereas interactions with friends and their perceived importance decreases, satisfaction with relationships with children increases across the same period.

7.22.9.1 Family Structure and Caregiving

Caregiving grows out of these normative patterns of intergenerational exchange and interaction. Strong social norms govern who is most likely to become a caregiver. By any estimation, gender is the single strongest predictor of who will become a caregiver, with women comprising approximately 70% of all caregivers (Aneshensel et al., 1995; Stone, Cafferata, & Sangl, 1987). Most research has suggested that caregivers assume their role hierarchically within the family structure. Married individuals will turn first to a spouse for assistance in times of need. In the absence of a spouse, adult children are the next choice, followed by siblings, more distant relatives, and finally neighbors and friends (e.g., Penning, 1990). Wives, daughters, and daughters-in-law are the most common female caregivers, and husbands the most common male caregivers, followed distantly by sons (Aneshensel et al., 1995; Stone et al., 1987), who are only likely to assume the caregiving role in the absence of a female sibling (Horowitz, 1985).

In most caregiving situations, one person takes over the primary responsibility, providing the most help or spending the most time with the care recipient. Other family members may also provide considerable assistance. Although sons do not frequently assume the role of primary caregiver, they often give considerable help to their sisters. Similarly, a parent caring for his or her spouse will typically receive help from children. The relationships between primary and secondary caregivers are an important source of support, but can also be a focal point for conflict and disagreement (Semple, 1992).

In studies of how siblings take responsibility for care tasks, Matthews (1987; Matthews & Rosner, 1988) found that family structure (birth order, number of siblings, gender composition), history of parent-child affectional ties, and extrafamilial roles (employment status, marital status, and geographic proximity) all influence which sibling provides parent care (see also Brody, 1990). In families with two or more adult daughters, shared care was the norm, whereas sons consistently appeared to play less central roles.

According to Matthews (1995), both adult daughters and sons draw upon cultural norms for gender-appropriate behavior in shaping their caregiving behavior. Adult daughters are viewed as being primarily responsible for caregiving activities regardless of the actual division of labor, and their brothers' contributions to caregiving are seen as peripheral and unimportant.

Another important distinction in the caregiving situation is whether or not the caregiver and care recipient share the same household (Chappell, 1990; Tennstedt, Crawford, & McKinley, 1993; Townsend, Noelker, Deimling, & Bass, 1989). Spouse caregivers usually live in the same household, but children may or may not live with a parent. An important source of cross-cultural differences is the extent to which aged parents and their adult children share a household (Sundström, 1992; Davey et al., 1996). In some countries such as the USA, rates of shared households have been relatively high, with children frequently taking in a disabled parent. In other countries, such as Sweden, two-generation households have been quite rare.

The implications of shared or separate households are considerable. A shared household probably places more continuous strain on family members and has been found to result in higher levels of stress (Townsend et al., 1989). In contrast, some different problems confront families who provide care to a parent in a different household, including needing to coordinate help that may be brought into the home, and worrying about the safety of the elder when no one is around. A special situation is trying to provide care at a long distance, which can be a source of considerable anxiety to children.

Marital status of parents and their children have emerged as important predictors of caregiving and intergenerational exchange. Specifically, parental divorce has been shown to affect the quality of relationships and frequency of contact between fathers and their adult children and the frequency of contact with mothers (Webster & Herzog, 1995). In general, the effects of divorce were much greater for fathers than for mothers (Cooney & Uhlenberg, 1990; White, 1992).

Disruption in adult children's marital status has also been found to relate to lower levels of help given to elderly parents, perceptions that parents need less assistance, and a lower sense of filial obligation to provide assistance (Cicirelli, 1983). A study by Brody and her colleagues (Brody, Litvin, Albert, & Hoffman, 1994) elaborates upon these findings, suggesting that daughters' marital status can have considerable effects on the amount of care they provide to an older parent. In particular, never-married and widowed daughters were found to provide the greatest proportion of care needed by a frail older parent by themselves. Never-married daughters, in particular, were most likely to be sole caregivers for their elderly parents. Conversely, married daughters were more likely to receive assistance with care tasks from informal sources, and divorced daughters were more likely to report using formal services.

7.22.9.2 Caregiver–Care Recipient Relationships

The relationship between caregiver and care recipient is an important and often overlooked aspect of these situations. Probably the most important distinction is between cognitively-intact people who can maintain some continuity in their relationships with caregivers, and those suffering from dementia.

Providing care to an elderly relative with dementia presents stressors which differ in fundamentally important ways from situations in which cognition is spared, or recovery is possible. When comparing caregivers of people with dementia but no physical limitations to caregivers of people with functional disabilities but no cognitive deficits, Birkel (1987) found that the former experienced much higher levels of care-related stress. In a comparison of caregivers of relatives with dementia and of relatives with cancer, Clipp and George (1993) found that dementia caregivers reported poorer self-rated health, greater substance use, poorer emotional health, reduced social functioning, and poorer financial status, even when the effects of employment status, age, and duration of illness were controlled.

Wright (1993) compared a small, nonrepresentative sample of Alzheimer's disease (AD) caregiving couples with well couples across a variety of dimensions. The study is unique in that she considered the perspective of both members having each report on specific aspects of their marriage. AD caregivers reported decreased companionship and difficulty dealing with changes in sexual intimacy. Wright also found that ratings of past marital happiness were associated with current valuing of the afflicted spouse as a unique individual, but not with future commitment, which was lower among caregiving spouses.

Majerovitz (1995) found a relationship between family adaptability and outcomes for caregivers of dementia patients. Caregivers who reported high adaptability in their families had lower depression, even when there were high levels of care-related stressors. She also reported that low-adaptability spouses were more likely to cling to an unrealistic role for their cognitively-impaired spouse, expecting their partner to continue to meet former expectations. While these data are cross-sectional, they suggest that caregivers with more realistic expectations of their demented spouse are likely to experience fewer negative consequences as a result of the care they provide.

In spousal relationships, relationship quality between the caregiver and care recipient may also play a role in the decision to institutionalize. Pruchno, Michaels, and Potashnik (1990b) found that spouses reporting a better quality of relationship with the care recipient reported less desire to institutionalize, but this finding did not extend to actual decisions regarding institutionalization. By the same token, however, desire to institutionalize was a strong predictor of actual institutionalization, suggesting that an indirect link may exist.

These findings offer support to the idea that there is something uniquely debilitating about losing a spouse to dementia as compared to another terminal illness in which cognitive functioning is typically spared. Because dementing illness results in deterioration in cognition and behavior, patients lose the ability to participate in the family or to have a say about the course of their own care. Caregiving becomes completely one-sided, with the patient unable to reciprocate in any way and, often, unable even to acknowledge the help that is provided. The progressive course of these disorders means that families face a somewhat predictable pattern of deterioration. Care of patients who are cognitively intact is not without its stresses; rather, care demands do not have the unrelenting pattern found in dementia and care recipients can acknowledge and show appreciation for the assistance they receive. Families caring for someone with another type of major psychiatric disorder may experience similar levels of distress as caregivers of dementia patients (Pearson, Verma, & Nellett, 1988). It may be that disorders which involve behavioral and emotional disturbances, and which are more disruptive of the usual family relationships, have the greatest impact on caregivers. The past relationship between children and their parents has considerable influence on caregiving experiences. Feelings of attachment and obligation have both been found to predict the amount of help daughters provided to their older parents (Cicirelli, 1993). Those who reported stronger feelings of attachment experienced lower levels of burden, whereas those who reported higher levels of obligation experienced higher levels of burden. Thus, children providing care out of a sense of obligation rather than affection may experience greater negative consequences as a result.

There are other important implications of the quality of prior relationships for the caregiving experience. Caregivers with a poor relationship with the care recipient are less likely to assume the caregiving role and, if they do, to experience more stress within it (Whitlatch & Noelker, in press). In an intriguing study by Uchino, Kiecolt-Glaser, and Cacioppo (1994), higher ratings of pre-illness affection for the care recipient were associated with lower heart rate reactivity, a physiological indicator of stress, measured two years later.

Walker, Martin, and Jones (1992) examined the role of intimacy in predicting adult daughters' and mothers' perceived costs and benefits of caregiving. Adult daughters' feelings of intimacy were associated with fewer feelings of having insufficient time for extra-caregiving activities, less frustration with caregiving, and less anxiety about providing care. For mothers receiving care, intimacy was related to less anger and resentment about the care they received. Other research has found conflict in the caregiver–care recipient relationship to be associated with greater caregiver strain and negative affect (Sheehan & Nuttall, 1988).

Townsend and Franks (1995) found that adult childrens' perceptions of relationship quality with their parent was an important mediator between the parent's cognitive impairment and the adult child's well-being. In their study of 90 adult–child caregivers, perceived closeness was associated with lower levels of reported caregiver burden and a greater sense of caregiving efficacy. Their study also underscores the importance of conflict in the relationship, which was associated with higher levels of stress, efficacy, and depression.

While most research has examined quality of the relationship, Aneshensel and her colleagues (1995) looked at the extent to which caregivers reported the loss of their relationship with a relative suffering from dementia. As their relative's disease progressed, caregivers reported a greater loss of the relationship. This sense of loss continued to play a role even after the patient's death. Caregivers who had previously reported a greater loss recovered more quickly than those who still held onto the relationship.

One extreme outcome of a poor relationship between caregiver and care recipient is violence directed toward the elderly care recipient. In a study of 236 caregivers to a relative with dementia, Pillemer and Suitor (1992) examined predictors of violence and violent feelings toward the care recipient. With regard to care recipient characteristics, they found that both violent behaviors and disruptive behaviors were associated with caregivers' greater fear of becoming violent with their relative. Violent behaviors by the care recipient were also associated with a greater likelihood of actual violence on the part of the caregiver. Caregivers who lived with their relative were more likely to report being afraid of acting violently toward their relative, and spouse caregivers were more likely to have acted violently toward their relative than adult children.

Thus, there is evidence to suggest that aspects of the past and present relationship quality between caregivers and care recipients play an important role in the caregiving experience, and that these characteristics have an effect, for both spouses and adult children, on the burden experienced as a result of caring.

7.22.9.3 Social Support from Family and Other Informal Sources

One of the most important sources of help to caregivers is the assistance and emotional support provided by other family members. Although actual estimates vary widely, there is general consensus that a majority of caregivers have support in their role from a secondary caregiver (e.g., Aneshensel et al., 1995; Stone et al., 1987; Tennstedt, McKinlay, & Sullivan, 1989). Social relationships can also involve conflict, and it has been suggested that negative social interactions can have greater impact on satisfaction with social support than measures of support received from others (e.g., Krause, 1995). In many instances, the same relationship can be a source of support and conflict.

Clipp and George (1990a) report considerable stability in the social support caregivers receive over a one-year period. They also report a lack of relationship between objective needs (dementia symptoms, duration of caregiving) and social support received. In other words, some families provide support whether the need for help is great or small, while other families do not help much at all, whatever the circumstances.

Thompson and colleagues (Thompson, Futterman, Gallagher-Thompson, & Lovett, 1993) examined potential differential effects of various types of social support in predicting measures of caregiver burden. Reduced social participation with friends and family was the most consistent social support predictor of the burden indicators they considered (see also, Haley et al., 1995). Specifically, higher levels of social participation were associated with lower levels of burden, having a less restricted social life, and a less negative relationship with the care recipient. That social support may also come with associated costs to caregivers was also indicated by their research. Caregivers who received material aid or financial assistance tended to report more negative interactions with family and friends.

Using longitudinal data, Aneshensel and her colleagues (1995) expanded on this theme that different types of support are related to different outcomes. Using the stress process model described earlier, they found that instrumental support from informal helpers (family and friends) can contain the experience of overload that results from primary stressors, whereas emotional support is not effective to this end. Similarly, instrumental support reduces proliferation of stress into other important roles held by the caregiver, such as work. Emotional support appears to be most effective in containing the effects of stress on outcomes such as depression.

For caregivers who are married, their relationship with a husband or wife can be both a source of support or a point of increased strain, as they experience conflicting demands on their time and energy. In an intriguing study, Creasey and his colleagues (Creasey, Meyers, Epperson, & Taylor, 1990) investigated how caring for a parent or parent-in-law with AD affected the relationship between husband and wife with each other and with their parents and children. Women caring for a parent (or parent-in-law) experienced a loss of support both from their husband and parent. Support from their own children, however, was not affected by caregiving, nor did their involvement lead to increases in negative interactions with their spouse or parent. Their husbands reported less support from parents, but not from wives or children.

Suitor and Pillemer (1992; see also Franks & Stephens, 1996) found, not surprisingly, that when husbands provided emotional support to caregiving wives, marital satisfaction was higher. Conversely, marital satisfaction was lower when husbands hindered their wives' caregiving in some way.

Caregiving grows out of a relationship with a long history and previously well-established expectations for interactions. There is evidence that this history has a substantial influence on the caregiving experience. As the nature of the caregiver–care recipient relationship changes, positive and negative interactions in other important family relationships can also be found. How much the support from others helps caregivers depends on how and what type of support are given, in a complex, but comprehensible fashion. Support from family and friends is a readily available resource for caregivers, which is not limited to the working hours of formal service agencies, nor bound by bureaucratic rules. Finding ways to increase support to caregivers while reducing conflict is an important goal for clinical intervention, as we shall discuss later.

7.22.10 SECONDARY STRESSORS: COMPETING ROLES AND STRESS PROLIFERATION

Caregivers have dimensions in their lives outside the demands of caregiving. These other roles can create additional stress on the caregiver, or they can help to alleviate the stress of caregiving, depending on the nature and quality of these outside roles. In the model guiding this chapter, these outside influences fall under the term secondary stressors.

Secondary stressors contain both role strains and intrapsychic strains. Role strains may arise from family, work or informal responsibilities outside of the caregiving situation, but are generated or aggravated by the demands of caregiving (Aneshensel, Pearlin, & Schuler, 1993). These stressors include family conflict, work-caregiving conflict, and economic strain. Intrapsychic strains involve dimensions of self-concept and psychological states.

There are two main hypotheses regarding the effect of different roles on the well-being of a person. The scarcity hypothesis proposes that additional roles create strain. Goode (1960) introduced the concept of role strain, or the difficulty in meeting expected role demands. In this view, role strain is seen as normal and even inevitable; therefore, the task for everyone is to discover ways to reduce this strain. This corresponds to the competing commitment perspective within the caregiving literature. This concept focuses on the time dimension of the conflicts among multiple roles, and assumes that the other roles of caregivers create further demands on the caregivers' time (Stoller & Pugliesi, 1989).

In contrast to the scarcity hypothesis, Sieber (1974) introduced the expansion hypothesis, that there are positive effects of accumulating additional roles. Even though the obligations of a role may be taxing, the rights gained within that role may negate any detrimental outcomes that might occur because of the acquisition of the role. One way in which multiple roles might compensate for role strain is by allowing people to fall back on successes in one role if they fail within another. A number of studies have found that occupying multiple roles is associated with better overall well-being (Thoits, 1983; Verbrugge, 1983).

One important focus of multiple roles is on daughters who balance caregiving with being a wife and mother. Studies by Stephens and her colleagues (Stephens, Franks, & Townsend, 1994; Stephens & Franks, 1995) have identified both positive and negative consequences of these multiple roles. An important feature of these studies is that they examine how each role has stresses and rewards, which affect how one role might affect the other. Both stress and rewards in the role of mother were associated, respectively, with positive and negative affect reported by caregivers. Stress and rewards in the role of wife were also associated with ratings of physical health.

Quality of the marital relationship can also have important consequences for the effects of caregiving on adult daughters, and vice versa. Stephens and Franks (1995) examined positive and negative perceptions that one role (either caregiving or spousal) had affected the other ("spillover"). Whereas positive spillover from the role of wife to caregiver was only related to higher levels of marital satisfaction, positive spillover from the caregiver to wife role was associated with lower levels of depression and greater positive affect. Conversely, negative spillover from wife to caregiver roles was associated with less positive affect and lower marital satisfaction; negative spillover from caregiver to wife roles was associated with significantly greater depression.

The main area in which these competing hypotheses about multiple roles have been explored is work and caregiving. Work has a central role in many people's lives and the difficulties that balancing work and caregiving responsibilities pose make it a particularly challenging issue. The rapid increase in the number of working women in first-world nations has made work and caregiving conflict an important social issue, as well as a consideration in clinical situations. An unusual longitudinal study of work and caregiving documents the increased rate at which women are combining both roles. This study examined a group of women who were originally interviewed in 1956 and then interviewed again 30 years later about caregiving histories (Moen, Robison, & Fields, 1994). Over 60% of the women in the study were caregivers to relatives with disabilities at some time in their lives. In this study, women were as likely to combine working and caregiving as they were to be only caregivers. Women seem to be adding to their roles rather than giving up a role when a new role must be acquired.

Working and caring for an elderly relative produces some special challenges and issues. When studying mothers and their daughter caregivers, Brody and Schoonover (1986) found that mothers received the same amount of help whether their daughters worked or not. But when employed women were compared with their nonemployed sisters, significant differences were found in the amount of help given when their parents' functional status was poor, with working women providing less assistance (Matthews, Werkner, & Delaney, 1989). Working women, then, are often providing extensive help in the caregiving role, but there are limits to their ability to help, which become evident in situations involving more extensive disabilities.

The question of whether an employed woman should make adjustments to her employment situation or quit her job is one that many caregivers face. A sizable number of caregivers leave their jobs within three months of becoming a caregiver (Franklin, Ames, & King, 1994). Leaving a job or reducing work hours is the most dramatic change that a caregiver can make in order to deal with the competing demands of work and caregiving. Workers generally try to rearrange their schedules or reduce their hours before leaving the job (Mutschler, 1994). There are many different factors that contribute to the decision or need to leave a job because of caregiving responsibilities. Caregivers who leave their jobs or reduce their work hours tend to have relatives who have more severe disabilities (Brody, Kleban, Johnsen, Hoffman, & Schoonover, 1987) and have faced work constraints for a longer period of time (Mutschler, 1994).

Reports of strain are common among working caregivers (Mutschler, 1994; Scharlach & Boyd, 1989). But there is the question of what contributes to that strain. The four main predictors of conflict between work and caregiving are: (i) the elderly family member's emotional health; (ii) the number of caregiving tasks performed by the caregiver; (iii) the presence of children in the caregiver's residence; and (iv) having caregiving responsibilities when beginning at the current job (Gibeau & Anastas, 1989). In addition, sociodemographic variables such as the employee's gender, marital status, household income, job status, occupation, and flexibility of work schedule are related to employees' levels of absenteeism and stress (Neal, Chapman, Ingersoll-Dayton, Emlen, & Boise, 1990). Other studies have found that the severity of disability of the elderly relative and the intensity of the caregiving responsibilities are related to work and caregiving conflict (Gibeau & Anastas, 1989).

Beyond the simple distinction between working and nonworking caregivers, the type of occupation has an effect on strain or burden. Mutschler (1994), for example, reports that production workers have less opportunity to adjust their work and care demands than workers in executive, sales, or clerical positions.

The relationship between experiences in the employment role and caregiving stress may be more complicated than originally thought. As was the case when balancing roles of caregiver, mother, and wife, the dual commitments of work and caregiving involve rewards as well as stress. High satisfaction with work seems to be associated with fewer depressive symptoms, except for those caregivers who experience high levels of caregiving stress (Martire, Stephens, & Atienza, 1997). The number of hours that a caregiver worked is associated with negative outcomes under conditions of high caregiver stress. Martire, et al. (1997) also found that women working part-time experienced poorer physical health and less positive affect than those women who were working full-time. One reason for this may be the increased amount of time that full-time workers are away from the impaired relative (Brody, 1990; Scharlach, 1994). Another reason may be that full-time work provides greater knowledge, benefits, and support (Matthews et al., 1989).

Workers tend to be interested in work benefits that are designed to make work and caregiving easier. The ideas for benefits that generated the most interest were respite care, flexible work hours, reduced hours without reduced benefits, adult day care, and job sharing. Caregivers using formal services understood that the availability of those services was critical if they were to maintain their employment and meet their family member's need for care (Gibeau & Anastas, 1989).

In summary, the findings from research on competing roles indicate that the additional roles that a caregiver occupies can result in both stress proliferation and containment. In order to get a complete understanding of the challenges that the caregiver faces, it is necessary to understand the quality of the roles and the rewards as well as stresses that the various responsibilities bring. Advising a caregiver, for example, to reduce her work responsibilities before having a clear picture of the entire caregiving situation would be ill advised. Rather, the challenge in clinical situations can be to find ways to balance competing roles more effectively and to minimize the spillover of stress from one activity to another.

7.22.11 OUTCOMES OF CAREGIVING: HEALTH AND MENTAL HEALTH EFFECTS

Caregiving is a uniquely stressful event. Being a long-term caregiver has been found to be associated with a variety of negative outcomes. The consensus among studies examining physical and mental health outcomes is that a majority of caregivers report feeling physically and emotionally distressed (Anthony-Bergstone, Zarit, & Gatz, 1988; Brodaty & Hadzi-Pavlovic, 1990; Clipp & George, 1990b; George & Gwyther, 1986; Haley, Levine, Brown, Berry, & Hughes, 1987; Kinney & Stephens, 1989). Several studies report higher levels of psychological morbidity among caregivers compared to noncaregiving samples of the population. Anthony-Bergstone and her colleagues (1988) noted the increased feelings of hostility, anger, and anxiety among caregivers as compared to the general population. An Australian study showed that mean scores on the General Health Questionnaire (GHQ), a measure of psychological morbidity, were higher in the caregiving sample and among caregiving spouses than the general population or other subsamples of caregivers (Brodaty & Hadzi-Pavolovic, 1990). Women caregivers generally report higher levels of psychological distress than men (Anthony-Bergstone et al., 1988; Zarit & Whitlatch, 1992).

Schulz, Visintainer, and Williamson (1990) reviewed several studies showing the psychiatric morbidity effects of caregiving. In studies using self-report measures such as the Beck Depression Inventory (BDI) or the Center for Epidemiological Studies Depression Scale (CES-D), three general conclusions emerged. First, greater severity of the care recipient's problems is linked to more depressive symptomatology. Second, women have higher rates of depression than men. Third, the more representative the caregiving population, the lower the rate of depression. Thus, caregivers drawn from specific subsamples such as daughters taking care of severely demented relatives or those seeking help from a service program show higher rates of depression symptoms.

A second group of studies has used clinical assessment tools (e.g., the Schedule for Affective Disorders and Schizophrenia Interview) which yield rates of clinical diagnoses of depression and other problems. These studies have found that caregivers frequently meet diagnostic criteria for depression or other diagnoses (e.g., Cohen & Eisdorfer, 1988; Drinka, Smith, & Drinka, 1987; Goldman & Luchins, 1984). In one of the most comprehensive studies of this kind, Gallagher and colleagues (Gallagher, Rose, Rivera, Lovett, & Thompson, 1989) found prevalence rates for depression ranging from 31% for men to 46% for women. These studies provide strong evidence that caregivers are at risk of increased psychiatric symptoms and clinical disorders.

The model of the stress process described earlier in the chapter illustrates how primary and secondary stressors can have a cumulative effect on caregiver mental well-being and physical health. Among the primary stressors, behavior problems are most likely to lead to adverse outcomes for family caregivers. The common behavior problems identified by caregivers as difficult or stressful for them are incontinence of the bowel and/or bladder, crying easily, hiding things, acting restlessly or agitated, and becoming irritable, angry, or accusatory (Haley & Pardo, 1989). These behavioral problems require vigilance, surveillance, and control on the part of the caregiver. Pruchno and Resch (1989) found that asocial and disoriented behaviors on the part of the care recipient are associated with greater burden, decreased social participation, and less satisfaction with caregiving. For forgetful behaviors, the consequences for caregiver's mental health increase at first, and then diminish as the severity of behavior problems increases, suggesting a curvilinear relationship between the two.

In a study of depression in dementia caregivers, Redinbaugh, MacCallum, and Kiecolt-Glaser (1995) found that caregivers who experienced chronic depression over a three-year period reported their relative's behavior as more problematic, and had higher levels of upsetting social support than non-depressed caregivers or caregivers who had experienced a transient depressive episode during the course of the study.

Vitaliano and colleagues (Vitaliano, Young, Russo, Romano, & Magana-Amato, 1993) were interested in examining the possibility that caregiver behaviors, specifically in the form of expressed emotion, were predictive of subsequent care recipient behaviors one and a half years later. Controlling for initial levels of problem behaviors, dementia patients who were cared for by an individual with high levels of expressed emotion displayed significantly more negative behaviors at follow-up. These findings are significant because caregiver expressed emotion is correlated with higher levels of depression, suppressed anger, and lower life satisfaction (Bledin, MacCarthy, Kiupers, & Woods, 1990; Vitaliano et al., 1993).

Burden, guilt, inadequacy, and negative perceptions of patient mood and functioning are words that describe the motivation to engage in caregiving activity (Hadjistavropoulos, Taylor, Tuokko, & Beattie, 1994). Lack of time, lack of help, and difficulty in managing patient behavior describe the caregiver's ability to engage in the caregiving process. These stressors have a cumulative impact on caregiver well-being.

Two broad domains have been investigated for the extent to which they buffer or contain the stressors of caregiving: social support and coping. The role of social support was reviewed earlier. Coping resources potentially represent an important way of containing stress. Research in this area has identified some promising leads on what type of coping strategies are more effective, but limitations in current research make it difficult to draw firm conclusions.

Pruchno and Kleban (1993) studied the relative effectiveness of emotion-focused vs. problem-focused coping strategies in mediating stress experienced by caregivers to an institutionalized relative. Their findings suggest that, contrary to coping research in other caregiving contexts (e.g., Stephens, Norris, Kinney, Ritchie, & Grots, 1988), emotion-focused coping was associated with better mental health than problem-focused coping. These findings highlight the importance of recognizing caregiving as a chronic stressor which is less amenable to change than many other situations. As a result, caregivers may be more successful if they focus on dealing with their responses to the situation than if they take a more traditional, instrumental approach.

Williamson and Schulz (1993) examined the relationship between a number of coping strategies and the burden experienced across three domains of the caregiving relationship: dealing with memory deficits, loss of communication, and the decline of a loved one. Their results suggest that different strategies are more and less successful, depending on the specific domain. Use of wishfulness as a coping strategy was related consistently to the experience of greater burden for all three stressors considered. Similarly, use of direct action strategies in response to memory deficits was associated with greater depression. Strategies which are more successful for caregivers included using relaxation as a way of coping with memory problems and seeking support to deal with loss of the relationship with the care recipient. Finally, acceptance of the situation was associated with lower depression related to loss of a loved one and loss of communication.

Gignac and Gottlieb (1996) considered caregivers' use of 12 kinds of coping strategies in dealing with specific behavior problems exhibited by the care recipient. They found considerable variation in the kinds of strategies caregivers used over a five-month period, suggesting that caregivers modify their approach to dealing with stressors over time. Caregivers who focused on gaining insights into how their coping strategies affected the relationship with the care recipient, both in attributing behavior problems to the disease process and in understanding the contingencies between caregiver and care recipient behaviors reported lower levels of stress at follow-up, even when initial levels of stress were controlled.

Thus, the effects of coping are complex. Strategies which attempt to change elements of the disease process are unsuccessful and perhaps even harmful. Conversely, strategies which focus on acceptance of changes in the caregiver–care recipient relationship appear much more helpful. Again, support seeking can be helpful in this regard.

One of the most comprehensive examinations of the determinants of mental health problems of caregivers was conducted by Aneshensel et al. (1995). Using a longitudinal panel of caregivers of dementia patients, they were able to examine how changes in stressors were related to increased or decreased mental health symptoms over a four-year period of time. Stress proliferation took place when primary stressors, such as behavior problems or ADL disabilities, led to feelings of overload or role captivity (see above). Thus, there needed to be both the occurrence of specific care-related problems and a subjective sense that those problems were overwhelming. In those instances, secondary stressors such as work or family conflict were more likely to emerge, as were mental health problems. In turn, receiving more informal social support from family and better coping, specifically higher feelings of mastery, resulted in some containment of stress. These results suggest that both objective indicators of the patient's disabilities and caregivers' subjective evaluation of their impact are needed to understand the processes leading to depression and other mental health problems.

Reports that caregiving is associated with high rates of psychiatric symptoms are probably not surprising, given the multiple pressures experienced. More intriguing are the results of studies which suggest that caregivers may be at risk of poorer health or even increased risk of mortality. There is less consistency across studies in reporting health effects, but sufficient evidence has been found to suggest there is at least some risk.

Studies of the effects of caregiving on health generally focus on self-reported problems, although a few investigations have used more objective indicators. Caregivers generally report they are in poorer health than age-matched controls (e.g., Pruchno & Potashnik, 1989; Schulz et al., 1990). In a nationally-representative sample from the USA, one-third of women caregivers reported fair or poor health compared to only one-fifth of age-matched peers (Stone et al., 1987). Other studies (e.g., Haley et al., 1987; Pruchno & Potashnik, 1989) have found that caregivers visit their physicians more often and use more prescription drugs compared to noncaregivers. These results, however, have not been confirmed in other studies of medical service use (e.g., George & Gwyther, 1986; Kiecolt-Glaser et al., 1987). Finally, several studies have found specific increases in health problems, such as respiratory illnesses, hypertension, and cardiovascular disease (Moritz, Kasl, & Ostfeld, 1992; Schulz et al., 1990; Wright et al., 1993).

Undoubtedly, the response of caregivers to chronic caregiving stress can follow the patterns of the general adaptation syndrome (Selye, 1976). Caregivers are exposed to a great deal of rage, grief, despair, and other intense emotions that can have effects on the blood and heart as well as clinical depression and anxiety. Caregivers may also exert a great deal of physical effort that can raise blood pressure and initiate atherosclerotic responses. The most common risk factors for health problems include being a spouse caregiver (Wright et al., 1993), living arrangements (Pratt, Wright, & Schmall, 1987), financial strain, and the intensity of primary stressors and secondary stressors. As stated before, these risk factors become detrimental to health when they are chronic.

One of the most intriguing findings linking chronic stress to health problems is that caregivers of dementia patients may experience changes in immune system functioning. In investigations which examined caregivers and a sample of matched controls, Kiecolt-Glaser and her colleagues (Keicolt-Glaser et al., 1987; Kiecolt-Glaser, Dura, Speicher, Trask, & Glaser, 1991) found that caregivers had poorer immune function, and these changes were linked to higher rates of illness in some subjects.

Caregivers' health problems may increase due to compromised self-care and health behavior (Brodaty & Hadzi-Pavlovic, 1990; Kwee, Zarit, Edwards, Davey, & Jarrott, 1995; Lieberman & Fisher, 1995). Caregivers may not go to the doctor when ill, or may not get enough rest or exercise. Nutritional intake can be affected, with some caregivers not getting adequate nutrition and others eating too much. Other caregivers may smoke more, use alcohol excessively, or even take tranquilizing medications prescribed for the patient. Consistent with the stress process model, compromised health behavior follows a model of competing demand whereby caregivers are faced with dwindling resources of energy, time, and money, and so must choose between self-care and care of their relative (Kwee et al., 1995). Often, caregivers feel no choice but to neglect their own health to continue to give care to their loved one (Bunting, 1989). Neglected self-care then acts as a mediating variable between stressors and the outcome of physical and mental health.

The process by which mental and physical health are affected as a result of caregiving stressors is admittedly complex. Uncertainties in causal directionality of independent and dependent variables and the possible existence of feedback loops due to the dynamic nature of an elder's disabilities add to the difficulty of making causal statements. Based on findings of the research, a theoretical model can be summarized as follows: the relative's illness creates the environment of primary stressors that proliferate or spillover to the creation of secondary stressors. Primary stressors, as represented by the mental and physical hardships of caregiving (e.g., motivation and ability), have both a main effect and a mediated effect (through secondary stressors) on caregiver mental and physical health. The secondary stressors that result are manifested through the caregiver's inability to take proper physical and mental care of him/herself because of a lack of energy or time. Prolonged exposure to primary and secondary stressors affects well-being. Research suggests a lagged relationship between depression and physical health, with depression leading to greater levels of physical illness (Pruchno, Kleban, Michaels, & Dempsey, 1990a).

In this framework, healthfulness is the result of continually achieving the best possible adaptation to the environment and involves interdependent physical, psychological, and social aspects (Capra, 1982). Illness occurs if stress pushes the system away from balance in response to environmental influences. To regain balance, the individual depends on his/her physical and mental response mechanisms. Because of the extensive demands that caregiving places on primary caregivers, it is not surprising that some fail to adapt to changing environmental situations and so experience severe disruption and breakdown. The challenge now is twofold: to identify the factors that protect good health in caregivers, and to design interventions that alleviate chronic stress at points in time before the body becomes ill.

Although the negative outcomes of caregiving have been emphasized, these adverse effects are sometimes balanced by positive experience. Feelings of gain, a growing sense of competency, or satisfaction with having met one's obligations or done a job well can offset the negative consequences (Aneshensel et al., 1995; Lawton, Brody, & Saperstein, 1989a). For some caregivers, fulfilling an obligation in the ways they define it is more important than putting themselves at risk of health problems or depression. Clinicians need to recognize that these commitments represent deeply-held beliefs, and not neurotic guilt (although that can be a motivator for some people). Successful clinical interventions involve helping caregivers explore and sometimes redefine their commitments, rather than confrontation over what the clinical decides is inappropriate or irrational. Solutions may also involve finding better ways of fulfilling a commitment, as described in the following sections.

7.22.12 CLINICAL RESEARCH AND SERVICES

The unprecedented stress associated with family caregiving warrants an organized response to help share a portion of the burden. At the broadest level, it is in the interest of society to encourage families to provide care, since that reduces the demand for more expensive institutional care. From the care recipient's perspective, being cared for at home for as long as possible is usually preferred. Families as well usually prefer to do as much as they can to help older family members.

Clinical interventions can be organized with two goals in mind. Interventions can be used to bring in new resources in ways that directly relieve families of a portion of their responsibilities. Respite services such as adult day care or home help serve this function. The most extensive new resource is an institutional setting, which takes over most of the care. Interventions can also build up the caregiver's own abilities for managing the situation more effectively. The review that follows, will look first at respite services, including institutional care, and the benefits and drawbacks for families. It will then consider clinical interventions for caregivers.

7.22.13 RESPITE SERVICES

Respite services for caregivers of elderly relatives with dementia are widely advocated by professionals and sought after by caregivers. While there is no clear definition of what services comprise respite, day care, in-home help, and short- and long-term institutionalization are some of the most frequently used services. Respite is designed to serve both caregiver and patient. With the growing awareness that institutional care is often not optimal for older people, respite services have been promoted as a means by which impaired elders could remain in the community longer while reducing the cost of care provision.

Despite reports of high satisfaction with respite services (e.g., Burdz, Eaton, & Bond, 1988; Kosloski, Montgomery, & Borgotta 1990;

Lawton et al., 1989a, 1989b; Montgomery & Berkeley-Caines, 1989; Montgomery & Borgatta, 1985), there has not been much evidence of the benefits for caregivers or patients. Some investigations have found little impact of respite services, leading to the suggestion that these programs may be of no value to the family (Callahan, 1989). These negative findings, however, may be the consequence of poor programing or low or inconsistent levels of service utilization. This section describes different respite services and considers outcomes of and obstacles to utilization.

7.22.13.1 Adult Day Care

Adult day care programs are designed for individuals with dementia or other mental or physical disabilities. They provide the patient with a few hours of supervised, structured activity, and give the family caregiver time to rest or perform other chores and responsibilities. Typically an individual will attend a day care program two to three days each week for about five hours each day (Weissert et al., 1990). Program activities include orientation exercises, socializing, and reminiscing. Programs frequently provide services such as health screening, podiatrist and eye doctor visits, and prescription administration. Out-of-home care allows the caregiver to stay at home and enjoy their privacy or have time with other family members (Montgomery, 1995). Day care gives the patient the opportunity to meet others and have a change of scenery and activities while remaining in a safe environment.

Despite reports that caregivers are very satisfied with day care (Burdz et al., 1988; Lawton et al., 1989a, 1989b; Miller, Gulle, & McCue, 1986; Sands & Suzuki, 1983), benefits for the caregiver are equivocal. Some research supports the idea that day care usage results in improved physical and mental health, better relationships, and increased confidence in caregiving abilities among family caregivers (Montgomery & Borgatta, 1985; Scharlach & Frenzel, 1986). Other studies, however, have found that day care has little or no effect on caregiver burden or well-being (Burdz et al., 1988; Lawton et al., 1989b; Montgomery & Borgatta, 1989).

The effects, or lack thereof, of day care programs on caregiver burden may depend on the characteristics of the individual in the caregiving situation (Montgomery, 1995). Caregivers' employment, health, or responsibility for another family member could affect utilization and mediate the benefits of a day care program. Failure to find benefits may also be due to low utilization of programs. For example, Lawton et al. (1989a) found no effects of program participation on caregiver burden, but the caregivers in the program utilized day care less than once a month. Inability to control how much families utilize day care prevents firm conclusions being drawn about the benefits of adult day care. Some caregivers find the programs too inaccessible (MaloneBeach, Zarit, & Spore, 1992), and some relatives would rather remain at home.

Preliminary results from a large-scale trial of adult day care for dementia are more promising (Zarit, Stephens, Townsend, & Greene, 1996). In this study, caregivers received adequate amounts of respite (at least two days per week) from well-developed programs with expertise in care of dementia patients. Measures were also developed to be specific to the strains likely to be addressed by this kind of respite care. Results showed that caregivers using day care experienced lower feelings of overload, decreased anger and depression, and increased well-being compared to a control sample not receiving day care or other respite services. These results suggest that well-designed day care programs may relieve stress on caregivers.

7.22.13.2 In-Home Respite

Respite services provided in the home include attendance to the patient's physical needs, such as bathing and feeding, and "sitting" with the patient, either for a short period or overnight. Providers may include a companion, homemaker, home health aide, or a nurse (Isett, Kraus, & Malone, 1984). Levels of in-home care can be low or high, for a short or an extended period of time. Home respite is intended for patients who still live alone and need assistance, and for the coresiding caregiver of a more dependent patient.

Reports on benefits of in-home respite for the caregiver are generally positive. In-home care eliminates the need for transportation and may be scheduled with some flexibility (Montgomery, 1995). Benefits of in-home respite for the caregiver include improved mood (Gwyther, 1989) and decreased time spent in caregiving activities (Berry, Zarit, & Rabatin, 1991). In addition to being satisfied with in-home help, caregivers usually want more of it. At the conclusion of a 12-month program during which respite services were available, Lawton et al. (1989b) found that all caregivers in the study wanted more respite than any other class of services. MaloneBeach et al. (1992) also reported that caregivers wanted more help in the home.

Caregivers using in-home help have also voiced concerns with the services provided. Complaints about in-home help have focused on the qualifications, reliability, and continuity of respite providers, and availability and flexibility of scheduling (MaloneBeach et al., 1992; Montgomery, 1995). Obstacles to use of in-home respite include bureaucracy and disengagement of the caregiver from the service system (MaloneBeach et al.). Caregivers report that the structure of the service system is too narrow as it often does not meet all of the patient's or caregiver's needs. Caregivers also report that the system is unaccommodating of changes in patients' conditions unless they fit into the service levels of a schedule fixed by the respite program. These studies, which were conducted in the USA, may not capture caregivers' experiences in countries that provide better access to and coordination of services.

Although many consider in-home respite a means by which family caregivers may keep their relative with dementia in the community longer, use of respite is sometimes associated with short- or long-term institutionalization (Montgomery 1995; Montgomery & Borgatta, 1985; Montgomery & Kosloski, 1990). Lawton et al. (1989b) reported that caregivers who spent the most time and effort caregiving were most likely to use respite, particularly nursing home respite. As caregivers learn to be more comfortable sharing the burden of caregiving, it may become more acceptable to them to place their relative in an institution (Montgomery, 1995). An alternative explanation is that respite services are a stepping stone to institutionalization in the progression of the patient's illness.

7.22.13.3 Institutionalization and its Consequences

Short-term institutionalization offers overnight respite for family caregivers and provides skilled nursing for the patient. A short-term admission may be scheduled every couple of weeks for two to three nights or less frequently for one to two weeks (Montgomery, 1995). While short-term institutionalization offers the benefit of an extended period of respite, it tends to be less flexible than in-home care. Institutions cannot usually offer emergency respite because of the cost of keeping a bed free (Scharlach & Frenzel, 1986). Some patients have difficulty adjusting to a new environment (Mace & Rabins, 1981), and some will resist institutional respite for fear that it is a permanent placement (Montgomery, 1995). In a study of two weeks of institutional respite, Burdz et al. (1988) expected dementia patients to show increased problems due to the confusion of being in a new environment. The patients, however, demonstrated significant decreases in problem behaviors, and caregivers rated the program as highly beneficial, though few said their situation had improved after the program. Even with an extended break from caregiving, short-term institutionalization frequently precedes long-term placement (Montgomery & Borgatta, 1989).

Placement of a relative in a long-term care facility may be viewed as a final release of caregiver stress and responsibility, but it is accompanied by its own stressors (Townsend, 1990; Zarit & Whitlatch, 1992). Institutionalization may ease some of the strain of everyday care, but many caregivers continue to provide assistance even after transferring their relative to an institution. While primary stressors such as feelings of overload decrease following placement, feelings of depression and other mental health indicators are often unchanged (Zarit & Whitlatch, 1992). MaloneBeach et al. (1992) found that family caregivers with a relative in a nursing home reported that their relative still needed additional services not provided in the nursing home. Caregiving strains may persist for as long as four years after placement (Aneshensel et al., 1995).

The high levels of satisfaction which caregivers report with all respite services is encouraging. However, the results of empirical studies indicate that respite services do not reduce caregiver burden and stress consistently. A major problem in these studies has been low rates of use of services, which is often related to barriers to or problems in service delivery. Availability and accessibility is another obstacle to service utilization. For example, waiting lists at US day cares are more common than in Sweden where the municipality provides sufficient access to day programs (Jarrott, Zarit, Berg, & Johansson, 1996). Flexibility of scheduling and services are essential, along with personal contact with a case worker and regular care providers. Once caregivers are taken off a waiting list and admitted to a program, many patients attend sporadically because hours or transportation are inconvenient (Montgomery, 1995). Cost and clear guidelines on qualifications for different services need to be provided along with the support which permits all those in need of assistance to receive it.

Despite caregivers' concerns about respite programs and the inconsistent findings of benefits, families overwhelmingly say it is important to them and most would like more respite. Respite services may need to be considered for their long-term effects

(Fortinsky & Hathaway, 1990). Greene and Coleman (1990) suggest that available services may serve a preventive function, so that stress on the family does not reach an overwhelming level. Other families, however, consider caregiving a personal responsibility and only look for assistance when they have reached their own limits. More active recruitment of family members early in their caregiving career may result in greater long-term relief from burden and stress. Perhaps when caregivers feel more comfortable turning to respite services for assistance, utilization will be greater and effects will be evident.

7.22.14 CLINICAL SERVICES

Clinical protocols have been developed to lower stress on family caregivers by increasing their ability to manage care-related problems and improving their access to help from other family members and from formal services (e.g., Zarit, Orr, & Zarit, 1985; Zarit, 1996). These interventions build on existing research by identifying modifiable aspects of the stress process. As observed, behavior problems do not have a fixed or constant effect on caregiver outcomes. By improving how caregivers manage behavior disturbances, the impact of these problems will be reduced and caregivers will experience less stress. Similarly, increasing support from and reducing conflict with other family members should lead to improvements in the caregiver's situation.

A model for clinical interventions is shown in Figure 2. As with any intervention, the starting point is assessment. Assessment focuses on both care recipient and caregiver and serves several functions. First, it is important to confirm the care recipient's disabilities and to make sure that treatable components have not been overlooked. Symptoms of memory loss sometimes have treatable causes and should not be assumed to be due to dementia, unless an appropriate medical and neuropsychological examination has been conducted. Second, the clinician can assess the specific problems, disabilities, and strengths of the care recipient. Although the focus is often on decrements, care recipients may retain some abilities and resources, even when suffering from dementia or other progressive disorders. Strengths can include activities the person currently is able to perform, or past interests and activities that may be possible to enhance in the present situation. Third, the assessment should indicate which problems are stressful for the caregiver. As emphasized throughout this chapter, caregivers experience potential stressors in different ways. What is difficult for one person may not be so for another. That includes management of specific behaviors, such as incontinence or wandering, or handling potential conflicts between caregiving and work or other roles. Finally, the clinician should consider the extent to which the care recipient can participate in a clinical intervention, either jointly with the caregiver, or separately. There have been reports that short-term counseling and support groups are helpful for dementia patients in the earlier stages of their illness (Teri, 1994; Zarit & Zarit, in press). For cognitively intact people, a wide range of clinical interventions might be appropriate.

Following the initial assessment, clinicians can draw upon three treatment strategies when working with caregivers: increasing the caregiver's understanding of the patient's illness and disabilities; problem solving to manage daily care-related problems; and support. A starting point for interventions is often to increase caregiver's understanding of the care recipient's illness. Families often have questions that have not been answered by medical personnel, or they have trouble understanding the implications of a diagnosis. Often they will have questions about the possibilities for treatment.

An important facet of understanding the disorder is learning to view changes in the patient's behavior as part of the disease process. In disorders like dementia that have considerable impact on behavior, families often misunderstand or misinterpret common problems. They may, for example, believe that a patient asks the same question over and over to annoy the caregiver or to get attention, or because of laziness. Families should be encouraged to look at why someone with a memory-impairing disease would behave this way. They should also avoid reasoning or arguing with patients over factual issues. Patients may insist that they want to see their long-deceased mothers, or to go home, when they are already at home. Experience shows that caregivers cannot persuade patients they are mistaken. In fact, reminding patients that their mothers are dead is often upsetting, since they experience it as a new loss, and the result is more agitation. Instead, families are encouraged to think about why someone with memory loss would behave this way, and what they must be feeling at this time. Although families cannot respond effectively to the content of these types of communications, they can respond to the underlying emotions, for example, by providing comfort or reassurance, or by reminiscing about the patient's mother.

Increasing the caregiver's understanding of problem behaviors flows naturally into a more formal approach to problem solving. Problem

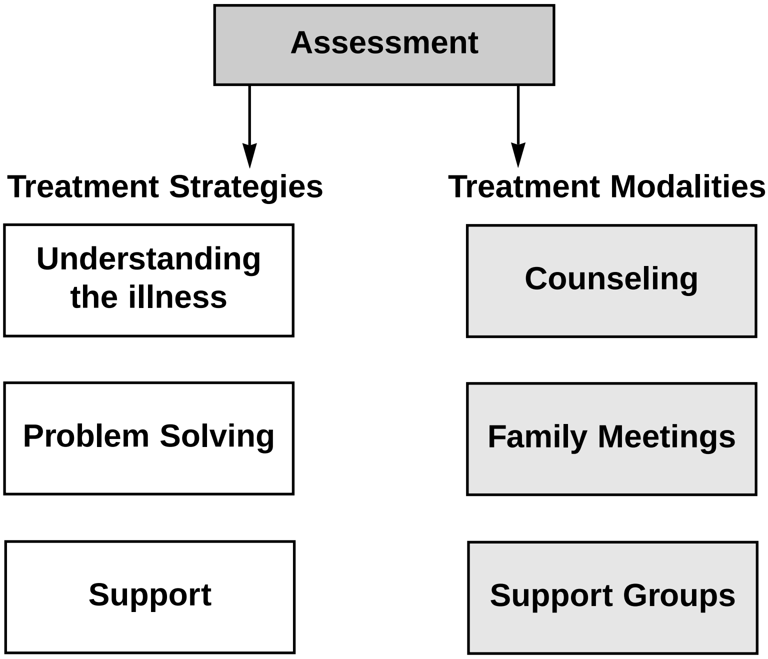

Figure 2 A model of intervention with caregivers.

solving can be used to address persistent troubling behaviors. Drawn from contemporary behavior therapy, problem solving involves a series of steps to identify causes and reinforcers of current problems and to generate and implement possible solutions (see Table 1). The starting point is to pinpoint a specific behavior or problem (e.g., agitation, depressed behaviors), and to have the caregiver observe it over a period of several days. Caregivers should note when the problem occurs, what went before it (antecedents), and what followed the behavior (reinforcers). This step provides a baseline of how often the problem occurs, against which interventions can be measured. It also identifies how the behavior may function. In dementia, for example, agitated behavior typically follows a period of inactivity or napping and is reinforced by attention.

Once a pattern is identified, the caregiver can be encouraged to think about strategies to change or head off the problem. Strategies need to be specific and practical and within the ability of the caregiver to manage. Caregivers may have difficulty selecting a solution. In that case, they can compare different strategies with the method of pros and cons, by which they write down advantages and drawbacks of each approach (including doing nothing). Planning and rehearsing the solutions before trying them out heads off problems if the plan is not realistic. After implementing a solution, caregivers should continue monitoring how often the problem occurs in order to evaluate how successful this strategy has been.

Although it is a behavior technique, problem solving is compatible with the use of medications. In fact, medications to control behavior or mood should be considered as one of the possible strategies for intervention. If the caregiver obtains a baseline of the frequency of the targeted problem before a medication is introduced, then it will be possible to get an accurate evaluation of the effectiveness of treatment. As clinicians know, medications often have uncertain and variable effects. Neuroleptics are widely used for control of behavior problems in the elderly, but they are not consistently effective or sometimes work for a while and then lose their effectiveness (Schneider, Pollock, & Lyness, 1990). By viewing medications as one of several possible

Table 1 The problem solving process.

	Stage
1	Pinpoint a behavior and assess when it occurs and how often
2	Identify antecedents and consequences
3	Identify possible strategies
4	Select a strategy (use pros and cons)
5	Plan and rehearse implementing the strategy
6	Try out strategy and evaluate

strategies and carefully monitoring their effectiveness, clinicians can develop more rational treatment approaches.

Problem solving can also be used to identify antecedents and consequences of caregiver stress. Rather than pinpointing a specific behavior problem, caregivers can note when they feel the most stress, and what has led up to that. Sometimes, it will be revealed that no one problem or behavior is overwhelming. Instead, the accumulation of small stressors without a break leads to difficulties. Solutions can then focus on how the caregiver can prevent the build up of stress, such as arranging for respite care.

The third treatment strategy is support. Clinicians can provide support directly with empathy for the caregiver and a nonjudgmental attitude. Interventions can also work to increase support and understanding from family and friends, and to gain more help from formal service providers. Clinicians can help families address the perceived and actual obstacles in respite services.

Three intervention modalities have been used with caregivers: one-to-one counseling; family meetings; and support groups. Each brings a different element to treatment. Short-term counseling can provide caregivers with the opportunity to explore their situation in depth and to consider the options and alternatives available to them in a supportive and nonjudgmental situation. Family meetings offer the opportunity to address directly areas of conflict and misunderstanding in the family and to build support and assistance for the primary caregiver. Support groups give caregivers the opportunity to discuss their situation with other people in similar situations. Groups may be an efficient means of sharing information and learning about new approaches and strategies.

Several studies document the effectiveness of these approaches. Whitlatch, Zarit, and von Eye (1991) found that caregivers receiving individual counseling and a family meeting had lower stress than people receiving only a support group or who were on a waiting list. In a series of studies, Toseland and his colleagues (Toseland, Rossiter, Peak, & Smith, 1990) have also found that counseling is more effective in the short run for relieving a caregiver's distress than support groups.

Mittelman and her colleagues (1993; 1995) developed a comprehensive intervention program that combines individual and family counseling during an initial four-month period, followed by sustained treatment in a support group. Caregivers in the treatment group had lower burden and distress and placed their relatives in nursing homes at a lower rate than people in a control group (see also, Whitlatch, Zarit, Goodwin, & von Eye, 1995).

These findings demonstrate the effectiveness of planned interventions designed to lower the stress of family caregivers. By increasing caregivers' skills in managing problems and stressors and bringing in new resources to support the caregiver whenever possible, clinicians can provide substantial assistance even in the face of a deteriorating condition. Through these interventions, caregivers can function as well as possible, while providing an optimal care environment for their relative. Timely use of these interventions may be cost-effective by preventing the breakdown of the family caregiving system, and encouraging a rational use of respite services, including institutional care when that is appropriate.

7.22.15 CONCLUSIONS

One of the major features of an aging society is that families are caring for disabled elders more often and for longer periods of time than ever before. There is considerable individual variability in how caregiving is handled and experienced, but it is typically stressful for the primary family helper and sometimes for the extended family network. The phenomenon of caregiving cannot be captured within a single measure of burden or stress. Rather, the many dimensions involved in the family's responses and the types of different stressors or problems caregivers experience have been observed. The stress of caregiving is an ongoing process, which evolves as the demands of care change and as caregivers develop new understanding, utilize different resources, or become overwhelmed by the unrelenting problems they face.

The challenge for the future is to find ways to support families so that they can provide high-quality, personalized care without excessive strain on themselves. The cost of care for social and health programs for the elderly has become an increasing problem for those countries with a significant proportion of their populations over age 65 years. As the number of elderly increase, cost-efficient means of providing care that do not sacrifice quality are needed. The responsibility for care cannot fall exclusively on families, because the burdens are too great, nor on society, since the cost will be too high. There need, instead, to be partnerships that find creative ways of sharing care, and which address the needs of both care recipient and caregiver. What shape those alliances will take, or the best models of working to support families, remain to be discovered. A small, but growing, clinical literature suggests that timely and well-conceived clinical interventions can be an important part of this partnership, helping families develop better strategies for managing a patient's disabilities and regulating the stress on themselves.

7.22.16 REFERENCES

Aneshensel, C. S., Pearlin, L. I., Mullan, J. T., Zarit, S. H., & Whitlatch, C. J. (1995). *Profiles in caregiving: The unexpected career*. New York: Academic Press.

Aneshensel, C. S., Pearlin, L. I., & Schuler, R. H. (1993). Stress, role captivity, and the cessation of caregiving. *Journal of Health and Social Behavior, 34,* 54–70.

Anthony-Bergstone, C. R., Zarit, S. H., & Gatz, M. (1988). Symptoms of psychological distress among caregivers of dementia patients. *Psychology and Aging, 3,* 245–248.

Berry, G. L., Zarit, S. H., & Rabatin, V. X. (1991). Caregiver activity on respite and nonrespite days: A comparison of two service approaches. *The Gerontologist, 31,* 830–835.

Birkel, R. C. (1987). Toward a social ecology of home-care households. *Psychology and Aging, 2*(3), 294-301.

Bledin, K. D., MacCarthy, B., Kuipers, L., & Woods, R. T. (1990). Daughters of people with dementia: Expressed emotion, strain, and coping. *British Journal of Psychiatry, 157,* 221-227.

Brodaty, H.,. & Hadzi-Pavlovic, D. H. (1990). Psychosocial effects on carers of living with persons with dementia. *Australian and New Zealand Journal of Psychiatry, 24,* 351–361.

Brody, E. M. (1990). *Women in the middle: Their parent-care years*. New York: Springer.

Brody, E. M., Johnsen, P. T., & Fulcomer, M. C. (1984). What should adult children do for elderly parents? Opinions and preferences of three generations of women. *Journal of Gerontology, 39,* 736–746.

Brody, E. M., Kleban, M. H., Johnsen, P. T., Hoffman, C., & Schoonover, C. B. (1987). Work status and parent care: A comparison of four groups of women. *The Gerontologist*, *27*, 201–208.

Brody, E. M., Litvin, S. J., Albert, S. M., & Hoffman, C. J. (1994). Marital status of daughters and patterns of parent care. *Journals of Gerontology, 49,* S95-S103.

Brody, E. M., & Schoonover, C. B. (1986). Patterns of parent-care when adult daughters work and when they do not. *The Gerontologist, 26,* 372–381.

Bunting, S. M. (1989). Stress on caregivers of the elderly. *Advances in Nursing Science, 11,* 63-73.

Burdz, M. P., Eaton, W. O., & Bond, J. B., Jr. (1988). Effect of respite care on dementia and nondementia patients and their caregivers. *Psychology and Aging, 3,* 38–42.

Callahan, J.J., Jr. (1989). Play it again, Sam—There is no impact. *Gerontologist, 29,* 5–6.

Capra, F. (1982). Wholeness & health. In *The turning point: Science, society and the rising culture*. (pp. 305–330). New York: Bantam Books.

Cassel, C. K., Rudberg, M. A., & Olshansky, S. J. (1992). The price of success: Health care in an aging society. *Health Affairs, (Summer),* 87–99.

Chappell, N. L. (1991). Living arrangements and sources of caregiving. *Journals of Gerontology, 46,* S1–S8.

Chow, N. W. (1983). The Chinese family and support of the elderly in Hong Kong. *The Gerontologist, 23,* 584–588.

Cicirelli, V. G. (1983). A comparison of helping behavior to elderly parents of adult children with intact and disrupted marriages. *The Gerontologist, 23,* 619–625.

Cicirelli, V. G. (1985). The role of siblings as family caregivers. In W. J. Sauer & R. T. Coward (Eds.), *Social support networks and the care of the elderly: Theory, practice, and research* (pp. 93–107). New York: Springer.

Cicirelli, V. G. (1992). Siblings as caregivers in middle and old age. In J. W. Dwyer & R. T. Coward (Eds.), *Gender, families, and elder care* (pp. 84–104). Newbury Park, CA: Sage.

Cicirelli, V. G. (1993). Attachment and obligation as daughters' motives for caregiving behavior and subsequent effect on subjective burden. *Psychology and Aging, 8,* 144–155.

Clipp, E. C., & George, L. K. (1990a). Caregiver needs and patterns of social support. *Journals of Gerontology, 45,* S102–S111.

Clipp, E. C., & George, L. K. (1990b). Psychotropic drug use among caregivers of patients with dementia. *Journal of the American Geriatrics Society, 38,* 227–235.

Clipp, E. C., & George, L. K. (1993). Dementia and cancer: A comparison of spouse caregivers. *The Gerontologist, 33,* 534–541.

Cohen, D., & Eisdorfer, C. (1988). Depression in family members caring for a relative with Alzheimer's disease. *Journal of the American Geriatric Society*, 36, 885–889.

Cooney, T. M., & Uhlenberg, P. (1990). The role of divorce in men's relations with their adult children after midlife. *Journal of Marriage and the Family, 52,* 677–688.

Creasey, G. L., Myers, B. J., Epperson, M. J., & Taylor, J. (1990). Couples with an elderly parent with Alzheimer's disease: Perceptions of familial relationships. *Psychiatry, 53,* 44–51.

Davey, A., Kwee, E. E., Shea, D. G., Zarit, S. H., Sundström, G., Berg, S., & Smyer, M. A. (1996, November). *How much do families help? Cultural and policy perspectives*. Paper presented at the annual meeting of the Gerontological Society of America, Washington, DC.

Drinka, T. J., Smith, J., & Drinka, P. J. (1987). Correlates of depression and burden for informal caregivers of patients in a geriatrics referral clinic. *Journal of the American Geriatrics Society, 35,* 522–525.

Eggebeen, D. J., & Wilhelm, M. O. (1995). Patterns of support given by older Americans to their children. In S. A. Bass (Ed.), *Older and active: How Americans over 55 are contributing to society* (pp. 122–168). New Haven, CT: Yale University.

Fadel-Girgis, M. (1983). Family support for the elderly in Egypt (1983). *The Gerontologist, 23,* 589–592.

Field, D., & Minkler, M. (1988). Continuity and change in social support between young-old and old-old or very-old age. *Journals of Gerontology, 43,* P100–P106.

Fortinsky, R. & Hathaway, T. (1990). Information and service needs among active and former family caregivers of persons with Alzheimer's Disease. *The Gerontologist*, *30*, 604–609.

Franklin, S. T., Ames, B. D., & King, S. (1994). Acquiring the family eldercare role: Influence on female employment adaptation. *Research on Aging*, *16*, 27–42.

Franks, M. M., & Stephens, M. A. P. (1996). Social support in the context of caregiving: Husbands' provision of support to wives involved in parent care. *Journals of Gerontology, 51B,* P43–P52.

Gallagher, D., Rose, J., Rivera, P., Lovett, S., & Thompson, L. W. (1989). Prevalence of depression in family caregivers. *The Gerontologist, 29,* 449–456.

George, L. K., & Gwyther, L. P. (1986). Caregiver well-being: A multidimensional examination of family caregivers of demented adults. *The Gerontologist, 26,* 253–259.

Gibeau, J. L., & Anastas, J. W. (1989). Breadwinners and caregivers: Interviews with working women. *Journal of Gerontological Social Work*, *14*, 19–40.

Gignac, M. A. M., & Gottlieb, B. H. (1996). Caregivers' appraisals of efficacy in coping with dementia. *Psychology and Aging, 11,* 214–225.

Goldman, L. S., & Luchins, D. J. (1984). Depression in the spouses of demented patients. *American Journal of Psychiatry, 14,* 1467–1468.

Goode, W. J. (1960). A theory of strain. *American Sociological Review*, *25*, 483–496.

Greene, V. L., & Coleman, P. D. (1995). Direct services for family caregivers: Next steps for public policy. In R. A. Kane & J. D. Penrod (Eds.) *Family are giving in an aging society: Policy perspectives* (pp. 46–63). Thousand Oaks, CA: Sage

Gwyther, L. P. (1990). *Barriers to service utilization among Alzheimer's patients and their families* (Report to the US Congress, Office of Technical Assessment, National Technical Information Service #PB89–225205. Durham, NC: Duke University Center on Aging.

Hadjistavropoulos, T., Taylor, S., Tuokko, H., & Beattie, B. L. (1994). Neuropsychological deficits, caregivers' perception of deficits, and caregiver burden. *Journal of the American Geriatrics Society, 42*, 308–314.

Haley, W. E., Levine, E. G., Brown, S. L., & Bartolucci, A. A. (1987). Stress, appraisal, coping and social support as predictors of adaptational outcome among dementia caregivers. *Psychology and Aging*, *2*(4), 323–330.

Haley, W. E., Levine E. G., Brown, S. L., Berry, J. W., & Hughes, G. H. (1987). Psychological, social, and health consequences of caring for a relative with senile dementia. *Journal of the American Geriatrics Society, 35*, 405–411.

Haley, W. E., & Pardo, K. M. (1989). Relationship of severity of dementia to caregiving stressors. *Psychology and Aging, 4*, 389–392.

Haley, W. E., West, C. A. C., Wadley, V. G., Ford, G. R., White, F. A., Barrett, J. J., Harrell, L. E., & Roth, D. L. (1995). Psychological, social, and health impact of caregiving: A comparison of black and white dementia family caregivers and noncaregivers. *Psychology and Aging, 10,* 540–552.

Hamon, R. R., & Blieszner, R. (1990). Filial responsibility expectations among adult child-older parent pairs. *Journals of Gerontology, 45,* P110–P112.

Hanson, S. L., Sauer, W. J., & Seelbach, W. C. (1983). Racial and cohort variations in filial responsibility norms. *The Gerontologist, 23,* 626–631.

Hirdes, J. P., & Strain, L. A. (1995). The balance of exchange in instrumental support with network members outside the household. *Journals of Gerontology, 50B,* S134–S142.

Horowitz, A. (1985). Sons and daughters as caregivers to older parents: Differences in role performance and consequences. *The Gerontologist, 25,* 612–617.

Ingersoll-Dayton, B., & Antonucci, T. C. (1988). Reciprocal and nonreciprocal social support: Contrasting sides of intimate relationships. *Journals of Gerontology, 43,* S65–S73.

Isett, R. D. Krauss, C., & Malone, M. (1984). *A study of the Intercommunity Actions, Inc. In-home respite care project*. Unpublished manuscript. Mid-Atlantic Long Term Care Gerontology Center.

Jarrott, S. E., Zarit, S. H., Johansson, L., & Berg, S. (1996, June). *A comparison of day care in Sweden and the United States*. Paper presented at the Nordic Gerontology Congress, Helsinki, Finland.

Johansson, B., & Zarit, S. H. (1995). Prevalence and incidence of dementia in the oldest old: A longitudinal study of a population-based sample of 84-90 year-olds in Sweden. *International Journal of Geriatric Psychiatry*, *10*, 359–366.

Jorm, A. F., Korten, A. E., & Henderson, A. S. (1987). The prevalence of dementia: A quantitative integration of the literature. *Acta Psychiatry Scandinavia*, *76*, 465–479.

Kaszniak, A. W. (1996). Preparing for treatment of the elderly: Assessment. In S. H. Zarit & B. Knight (Eds.), *A guide to psychotherapy and aging: Clinical intervention in a life stage context*. Washington, DC: APA Books.

Kendig, H. L., Rowland, D. T. (1983). Family support of the Australian aged: A comparison with the United States. *The Gerontologist, 23,* 643–649.

Kiecolt-Glaser, J. K., Dura, J. R., Speicher, C. E., Trask, O. J., & Glaser, R. (1991). Spousal caregivers of dementia victims: Longitudinal changes in immunity and health. *Psychosomatic Medicine, 53,* 345–362.

Kiecolt-Glaser, J. K., Glaser, R., Shuttleworth, E. C., Dyer, C. S., Ogrocki, P., & Speicher, C. E. (1987). Chronic stress and immunity in family caregivers of Alzheimer's Disease victims. *Psychosomatic Medicine, 49*, 523–535.

Kinney, J. M., & Stephens, M. A. P. (1989). Hassles and uplifts of giving care to a family member with dementia. *Psychology and Aging, 4,* 402–408.

Kinsella, K., & Taeuber, C. M. (1993). *An aging world II. International Population*. Reports P95/92–3. US Department of Commerce, Economics, and Statistics Administration, Bureau of the Census, Government Printing Office, Washington, DC.

Kivett, V. R., & Atkinson, M. P. (1984). Filial expectations, association, and helping as a function of number of children among older rural-transitional parents. *Journal of Gerontology, 39,* 499–503.

Kokmen, E., Chandra, V., & Schoenberg, B. S. (1988). Trends in incidence of dementing illness in Rochester, Minnesota, in three quinquennial periods, 1960–1974. *Neurology*, *38*, 975–980.

Koslowski, K. D., Montgomery, R. J. V., & Borgotta, E. F. (1990). Predicting nursing home utilization: Is this the best we can do? In S. M. Stahl (Ed.), *The legacy of longevity: Health and health care in later life*. Newbury Park, CA: Sage.

Krause, N. (1995). Negative interaction and satisfaction with social support among older adults. *Journals of Gerontology, 50B,* P59–P73.

Kwee, E. E., Zarit, S. H., Edwards, A. B., Davey, A., & Jarrott, S.E. (1995, November). *Caregiver's health-seeking behavior: Impact of the intensity of stressors.* Poster presented at the Fortyeighth Annual Meeting of the Gerontological Society of America, Los Angeles, CA.

Lawton, M. P., Brody, E. M., & Saperstein, A. R. (1989a). A controlled study of respite services for caregivers of Alzheimer's patients. *The Gerontologist, 29*(1), 8–16.

Lawton, M. P., Brody, E. M., Saperstein, A. R., & Grimes, M. (1989b). Respite services for caregivers: Research findings for service planning. *Home Health Care Services Quarterly, 10*(1/2), 5–32.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.

Lerner, M. J., Somers, D. G., Reid, D., Chiriboga, D., & Tierney, M. (1991). Adult children as caregivers: egocentric biases in judgements of sibling contributions. *The Gerontologist, 31,* 746–755.

Lieberman, M. A., & Fisher, L. (1995). The impact of chronic illness on the health and well-being of family members. *The Gerontologist*, 35, 94–102.

Mace, N. L. & Rabins, P. V. (1981). *The 36-hour day: A family guide to caring for persons with Alzheimer's disease, related dementing illness and memory loss in later life*. Baltimore: Johns Hopkins University Press.

Maeda, D. (1983). Family care in Japan. *The Gerontologist, 23,* 579–583.

Majerovitz, S. D. (1995). Role of family adaptability in the psychological adjustment of spouse caregivers to patients with dementia. *Psychology and Aging, 10,* 447–457.

MaloneBeach, E. E., Zarit, S. H., & Spore, D. L. (1992). Caregivers' perceptions of case management and community-based services: Barriers to service use. *The Journal of Applied Gerontology*, *11*(2), 146-159.

Martire, L. M., Stephens, M. A. P., & Atienza, A. A. (in press). The Interplay of work and caregiving: Relationships between role satisfaction, role involvement and caregivers' well-being. *Journals of Gerontology: Social Sciences, 52B,* S279-S289.

Matthews, S. H. (1987). Provision of care to old parents: Division of responsibility among adult children. *Research on Aging, 9,* 45–60.

Matthews, S. H. (1995). Gender and the division of filial responsibility between lone sisters and their brothers. *Journals of Gerontology, 50B,* S312–S320.

Matthews, S. H., & Rosner, T. T. (1988). Shared filial responsibility: The family as the primary caregiver. *Journal of Marriage and the Family, 50,* 185–195.

Matthews, S. H., Werkner, J. E., & Delaney, P. J. (1989). Relative contributions of help by employed and nonemployed sisters to their elderly parents. *Journal of Gerontology: Social Sciences*, *44*, S36–S44.

Miller, D. B., Gulle, N. & McCue, F. (1986). The realities of respite for families, clients and sponsors. *The Gerontologist*, *26*, 467–470.

Mittelman, M. S., Ferris, S. H., Shulman, E., Steinberg, G., Ambinder, A., Mackell, J., & Cohen, J. (1995). A comprehensive support program: Effect on depression in spouse-caregivers of AD patients. *The Gerontologist, 35,* 792–802.

Mittelman, M. S., Ferris, S. H., Steinberg, G., Shulman, E., Mackell, J. A., Ambinder, A., & Cohen, J. (1993). An intervention that delays institutionalization of Alzheimer's disease patients: Treatment of spouse-caregivers. *Gerontologist, 33,* 730–740.

Moen, P., Robison, J., & Fields, V. (1994). Women's work and caregiving roles: A life course approach. *Journal of Gerontology, Social Sciences*, *49*, S176–S185.

Monahan, D. J., & Hooker, K. (1995). Health of spouse caregivers of dementia patients: The role of personality and social support. *Social Work, 40,* 305–314.

Montgomery, R. J. V. (1995). Examining respite care: Promises and limitations. In D. E. Biegel, & R. Schulz (Series Eds.) & R. A. Kane & J. D. Penrod (Vol. Eds.), *Family Caregiver Applications Series: Vol. 5. Family caregiving in an aging society* (1st ed., pp. 29–45). Thousand Oaks, CA: Sage.

Montgomery, R. J., & Berkeley-Caines, L. (1989). Effects of in-home volunteer respite services on caregivers: A final report to Detroit Catholic Social Services. Detroit, IL: Wayne State University Institute of Gerontology.

Montgomery, R. J. V., & Borgatta, E. (1985). *Family support project: Final report to the administration on aging*. Seattle, WA: University of Washington

Montgomery, R. J. V., & Borgatta, E. (1989). The effects of alternative support strategies on family caregiving. *The Gerontologist*, *29*, 457–464.

Montgomery, R. J. V., & Kosloski, K. (1990, November). *Adult children and the nursing home decision.* Paper presented at the Fifty-second Annual Meeting of the National Council on Family Relations, Seattle, WA.

Morgan, D. L., Schuster, T. L., & Butler, E. W. (1991). Role reversals in the exchange of social support. *Journals of Gerontology, 46,* S278–S287.

Moritz, D. J., Kasl, S. V., & Berkman, L. F. (1989). The health impact of living with a cognitively impaired elderly spouse: Depressive symptoms and social functioning. *Journals of Gerontology, 44,* S17–S27.

Moritz, D. J., Kasl, S. V., & Ostfeld, A. M. (1992). The health impact of living with a cognitively impaired elderly spouse: Blood pressure, self-rated health, and health behaviors. *Journal of Aging and Health, 4,* 244–267.

Mutschler, P. H. (1994). From executive suite to production line: How employees in different occupation manage elder care responsibilities. *Research on Aging*, *16*, 7–26.

Neal, M. B., Chapman, N. J., Ingersoll-Dayton, B., Emlen, A. C., & Boise, L. (1990). Absenteeism and stress among employed caregivers of the elderly, disabled adults, and children. In D. E. Biegel & A. Blum (Eds.), *Aging and caregiving: Theory, practice, and research* (pp. 160–183). Newbury Park, CA: Sage.

Pearlin, L. I. (1993). The social contexts of stress. In L. Goldberger & S. Breznitz (Eds.), *Handbook of stress: Theoretical and clinical aspects* (2nd ed., pp. 303–315). New York: Free Press.

Pearlin, L., Mullan, J., Semple, S., & Skaff, M. (1990). Caregiving and the stress process: An overview of concepts and their measures. *The Gerontologist, 30,* 583–594.

Pearson, J., Verma, S., & Nellett, C. (1988). Elderly psychiatric patient status and caregiver perceptions as predictors of caregiver burden. *The Gerontologist, 28,* 79–83.

Penning, M. J. (1990). Receipt of assistance by elderly people: Hierarchical selection and task specificity. *The Gerontologist, 30,* 220–227.

Pillemer, K., & Suitor, J. J. (1992). Violence and violent feelings: What causes them among family caregivers? *Journals of Gerontology, 47,* S165–S172.

Pratt, C. P., Wright, S., & Schmall, V. (1987). Burden, coping, and health status: A comparison of family caregivers to community dwelling and institutionalized Alzheimer's patients. *Journals of Gerontological Social Work, 20,* 99–112.

Pruchno, R., & Kleban, M. H. (1993). Caring for an institutionalized parent: The role of coping strategies. *Psychology and Aging, 8,* 18–25.

Pruchno, R. A., Kleban, M. H., Michaels, E., & Dempsey, N. P. (1990a). Mental and physical health of caregiving spouses: Development of a causal model. *Journals of Gerontology: Psychological Sciences, 45,* P192–199.

Pruchno, R. A., Michaels, J. E., & Potashnik, S. L. (1990b). Predictors of institutionalization among Alzheimer disease victims with caregiving spouses. *Journals of Gerontology: Social Sciences, 45,* S259–S266.

Pruchno, R. A., Peters, N. D., Kleban, M. H., & Burant, C. J. (1994). Attachment among adult children and their institutionalized parents. *Journals of Gerontology: Social Sciences, 49,* S209–S218.

Pruchno, R. A., & Potashnik, S. L. (1989). Caregiving spouses physical and mental health in perspective. *Journal of the American Geriatrics Society, 37*, 697–705.

Pruchno, R. A., & Resch, N. L. (1989). Aberrant behaviors and Alzheimer's disease: Mental health effects on spouse caregivers. *Journals of Gerontology, 44,* S177–S182.

Redinbaugh, E. M., MacCallum, R. C., & Kiecolt-Glaser, J. K. (1995). Recurrent syndromal depression in caregivers. *Psychology and Aging, 10,* 358–368.

Reis, M. F., Gold, D. P., Gauthier, S., Andres, D., & Markiewicz, D. (1994). Personality traits as determinants of burden and health complaints in caregiving. *International Journal of Aging and Human Development, 39,* 257–271.

Sands, D., & Suzuki, T. (1983). Adult day care for Alzheimer's patients and their families. *The Gerontologist*, *23*(1), 21–23.

Scharlach, A. E. (1994). Caregiving and employment: Competing or complementary roles? *The Gerontologist*, *34*, 378–385.

Scharlach, A. E., & Boyd, S. L. (1989). Caregiving and employment: Results of an employee survey. *The Gerontologist*, *29*, 382–387.

Scharlach, A. E., & Frenzel, C. (1986). An evaluation of institution-based respite care. *The Gerontologist*, *26*, 77–82.

Schneider, L.S., Pollock, V. E., & Lyness, S. A. (1990). A metaanalysis of controlled trials of neuroleptic treatment in dementia. *Journal of American Geriatrics Society, 38,* 553–563.

Schulz, R., Visintainier, P., & Williamson, G. M. (1990). Psychiatric and physical morbidity effects of caregiving. *Journals of Gerontology: Psychological Sciences, 45,* P181–191.

Selye, H. (1976). *The stress of life* (Rev. ed.) (Preface and pp. 1–39). New York: McGraw Hill.

Semple, S.J. (1992). Conflict in Alzheimer's caregiving families: Its dimensions and consequences. *The Gerontologist*, *32*, 648–655.

Sheehan, N. W., & Nuttall, P. (1988). Conflict, emotion, and personal strain among family caregivers. *Family Relations, 37,* 92–98.

Sieber, S. D. (1974). Toward a theory of role accumulation. *American Sociological Review*, *39*, 567–578.

Skaff, M. M., Pearlin, L. I. (1992). Caregiving: Role engulfument and the loss of self. *The Gerontologist*, *32* (5), 656–664.

Stephens, M. A. P., & Franks, M. M. (1995). Spillover between daughters' roles as caregiver and wife: Interference or enhancement? *Journals of Gerontology, 50B,* P9–P17.

Stephens, M. A. P., Franks, M. M., & Townsend, A. L. (1994). Stress and rewards in women's multiple roles: The case of women in the middle. *Psychology and Aging*, *9*, 45–52.

Stephens, M. A. P., Norris, V. K., Kinney, J. M., Ritchie, S. W., & Grotz, R. C. (1988). Stressful situations in caregiving: Relations between caregiver coping and well-being. *Psychology and Aging, 3,* 208–209.

Stoller, E. P., & Pugliesi, K. L. (1989). Other roles of caregivers: Competing responsibilities or supportive resources. *Journals of Gerontology: Social Sciences*, *44*, S231–S238.

Stone, R., Cafferata, G. L., & Sangl, J. (1987). Caregivers of the frail elderly: A national profile. *The Gerontologist, 27,* 616–626.

Strawbridge, W. J., & Wallhagen, M. I. (1991). Impact of family conflict on adult child caregivers. *The Gerontologist, 31,* 770–777.

Suitor, J. J., & Pillemer, K. (1992). Status transitions and marital satisfaction: The case of adult children caring for elderly parents suffering from dementia. *Journal of Social and Personal Relationships, 9,* 549–562.

Sundström, G. (1992). Care by families: An overview of trends. *Caring for Frail Elderly People, Social Policy Studies, 14.* Paris: OECD.

Tennstedt, S. L., Crawford, S., & McKinlay, J. B. (1993). Determining the pattern of community care: Is coresidence more important than caregiver relationship? *Journals of Gerontology: Social Sciences, 48,* S74–S83.

Tennstedt, S. L., McKinlay, J. B., & Sullivan, L. M. (1989). Informal care for frail elders: The role of secondary caregivers. *The Gerontologist, 29,* 677–683.

Teri, L. (1994). Behavioral treatment of depression in patients with dementia. *Alzheimer Disease and Associated Disorders, 8*(3), 66–74.

Thoits, P. A. (1983). Multiple Identities and psychological well-being: A reformulation and test of the social isolation hypothesis. *American Sociological Review*, *48*, 174–187.

Thompson, E. H., Jr., Futterman, A. M., Callagher-Thompson, D., Rose, J. M., & Lovett, S. B. (1993). Social support and caregiver burden in family caregivers of frail elders. *Journals of Gerontology, 48,* S245–S254.

Toseland, R. W., Rossiter, C. M., Peak, T., & Smith, G. C. (1990). Comparative effectiveness of individual and group interventions to support family caregivers. *Social Work, 35,* 209–217.

Townsend, A. (1990). Nursing home care and family caregivers' stress. In M. A. P. Stephens, J. H. Crowther, S. E. Hobfall, & D. L. Tennenbaum (Eds.), *Stress and coping in later life families* (pp. 267–285). Washington, DC: Hemisphere.

Townsend, A. L., & Franks, M. M. (1995). Binding ties: Closeness and conflict in adult children's caregiving relationship. *Psychology and Aging, 10,* 343–351.

Townsend, A., Noelker, L., Deimling, G., & Bass, D. (1989). Longitudinal impact of interhousehold caregiving on adult children's mental health. *Psychology and Aging*, *4*, 393–401.

Uchino, B. N., Kiecolt-Glaser, J. K., & Cacioppo, J. T. (1994). Construals of preillness relationship quality predict cardiovascular response in family caregivers of Alzheimer's disease victims. *Psychology and Aging, 9,* 113–120.

Verbrugge, L. M. (1983). Multiple roles and physical health of women and men. *Journal of Health and Social Behavior*, *24*, 16–30.

Vitaliano, P. P., Young, H. M., Russo, J., Romano, J., & Magana-Amato, A. (1993). Does expressed emotion in spouses predict subsequent problems among care recipients with Alzheimer's disease? *Journals of Gerontology, 48,* P202–P209.

Walker, A. J., Martin, S. S. K., & Jones, L. L. (1992). The benefits and costs of caregiving and care receiving for daughters and mothers. *Journals of Gerontology, 47,* S130–S139.

Webster, P. S., & Herzog, A. R. (1995). Effects of parental divorce and memories of family problems on relationships between adult children and their parents. *Journals of Gerontology, 50B,* S24–S34.

Weihl, H. (1983). Three issues from the Israeli scene. *The Gerontologist, 23,* 576–578.

Weissert, W. G., Elston, J. M., Bolda, E. J., Zelman, W. N., Mutran, E., & Mangum, A. B. (1990). *Adult day care: Findings from a national survey*. Baltimore, MD: Johns Hopkins University Press.

White, L. (1992). The effect of parental divorce and remarriage on parental support for adult children. *Journal of Family Issues, 13,* 234–250.

Whitlatch, C. J., & Noelker, L. S. (in press). Caregiving and caring. In J. E. Birren (Ed.), *Encyclopedia of gerontology*. New York: Academic Press.

Whitlatch, C.J., Zarit, S.H., & von Eye, A. (1991). Efficacy of interventions with caregivers: A reanalysis. *The Gerontologist*, *31*, 9–14.

Whitlatch, C. J., Zarit, S. H., Goodwin, P. E., & von Eye, A. (1995). Influence of the success of psychoeducational interventions on the course of family care. *Clinical Gerontologist, 16,* 17–30.

Williamson, G. M., & Schulz, R. (1993). Coping with specific stressors in Alzheimer's disease caregiving. *The Gerontologist, 33,* 747–755.

Worach-Kardas, H. (1983). The Polish family tradition. *The Gerontologist, 23,* 593–596.

Wright, L. K. (1993). *Alzheimer's disease and marriage: An intimate account*. Newbury Park, CA: Sage.

Wright, L., Clipp, E., & George, L. (1993). Health consequences of caregiver stress. *Medicine Exercise Nutrition and Health, 2,* 181–195.

Zarit, J. M., & Zarit, S. H. (in press). Til death do us part. In R. P. Halgin & S. K. Whitbourne (Eds.) *Partners in change: Growth through the therapeutic process.* New York: Oxford University Press.

Zarit, S.H. (1996). Clinical interventions for family caregiving. In S. H. Zarit & B. Knight (Eds.), *A guide to psychotherapy and aging: Clinical intervention in a life stage context*. Washington, DC: APA Books.

Zarit, S. H., & Eggebeen, D. J. (1995). Parent child relationships in adulthood and old age. In M. H. Bornstein (Ed.), *Handbook of parenting: Vol. 1. Children and parenting* (pp. 119–140). Mahwah, NJ: Lawrence Erlbaum.

Zarit, S. H., Johansson, B., & Berg, S. (1993). Functional impairment and co-disability in the oldest old. *Journal of Aging and Health*, *5*, 291–305.

Zarit, S. H., Orr, N. K., & Zarit, J. M. (1985). *The hidden victims of Alzheimer's Disease: Families under stress*. New York: NYU Press.

Zarit, S. H., Stephens, M. A. P., Townsend, A. & Greene, R. (1996, November). Adult day care and the relief of caregiver strain. Paper presented at the annual meeting of the Gerontological Society of America, Washington, DC.

Zarit, S. H., Todd, P. A., & Zarit, J. M. (1986). Subjective burden of husbands and wives as caregivers: A longitudinal study. *The Gerontologist*, *26*, 260–266.

Zarit, S. H. & Whitlatch, C. J. (1992). Institutional placement: Phases of the transition. *The Gerontologist*, *32*, 665–672.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.23

Rehabilitation in Old Age: Psychosocial Issues

CLEMENS TESCH-RÖMER
Ernst Moritz Arndt University, Greifswald, Germany
and
HANS-WERNER WAHL
German Center for Research on Aging at the University of Heidelberg, Germany

7.23.1 INTRODUCTION 526

7.23.2 GENERAL PRINCIPLES AND GEROPSYCHOLOGICAL ESSENTIALS OF REHABILITATION IN OLD AGE 527

- *7.23.2.1 General Principles* 527
- *7.23.2.2 Geropsychological Essentials of Rehabilitation in Old Age* 529

7.23.3 PSYCHOSOCIAL ISSUES AND STRATEGIES OF REHABILITATION IN OLD AGE 530

- *7.23.3.1 Assessment Issues* 530
  - *7.23.3.1.1 Activities of daily living and instrumental activities of daily living* 530
  - *7.23.3.1.2 Performance vs. reserve capacity* 531
- *7.23.3.2 Restoration of Functional Independence* 531
  - *7.23.3.2.1 Functional independence as basic objective of geriatric rehabilitation* 531
  - *7.23.3.2.2 Contributions from clinical geropsychology* 532
- *7.23.3.3 Cognitive Retraining* 532
  - *7.23.3.3.1 Intelligence training* 533
  - *7.23.3.3.2 Memory training* 533
  - *7.23.3.3.3 Reality orientation training* 533
- *7.23.3.4 Psychotherapeutic Intervention* 534
- *7.23.3.5 Considering Motivation and Compliance Issues* 535
  - *7.23.3.5.1 Motivation* 535
  - *7.23.3.5.2 Compliance* 535
  - *7.23.3.5.3 Impact of the social network on motivation and compliance* 535
- *7.23.3.6 The Role of Social Support* 536
  - *7.23.3.6.1 Effects of social support* 536
  - *7.23.3.6.2 Mobilizing social support* 536
- *7.23.3.7 Ecological Aspects of Rehabilitation* 537
  - *7.23.3.7.1 Person–environment fit as critical for rehabilitation* 537
  - *7.23.3.7.2 Considering objective and subjective environmental aspects within the everyday world* 537
  - *7.23.3.7.3 New technologies and their rehabilitation potential* 538

7.23.4 PSYCHOSOCIAL REHABILITATION ISSUES WITH RESPECT TO PARTICULAR CLIENT GROUPS 538

- *7.23.4.1 Stroke Patients* 538
  - *7.23.4.1.1 Critical issues of rehabilitation* 538
  - *7.23.4.1.2 Effects of rehabilitation and challenges in the future* 539
- *7.23.4.2 Arthritis Patients* 539

7.23.4.2.1 Physical therapy 540
7.23.4.2.2 Psychological interventions 540
7.23.4.3 Sensory Impaired Elderly 540
7.23.4.3.1 Hearing impairment 540
7.23.4.3.2 Visual impairment 541
7.23.4.4 Demented Elderly 542
7.23.5 PRACTICAL ISSUES, RESEARCH QUESTIONS, AND ETHICAL PROBLEMS 542
7.23.5.1 Practical Issues 542
7.23.5.2 Research Questions 543
7.23.5.3 Ethical Problems 543
7.23.6 SUMMARY 544
7.23.7 REFERENCES 544

7.23.1 INTRODUCTION

Aging is associated with an increase in acute and chronic illness with serious consequences for daily living. To give the reader an idea of how prevalent these self-limiting conditions are in old age, the following empirical findings (Ebrahim, 1992; Kemp,1985) are used. The percentage of persons with arthritis increases from 1% for the less than 17 years old group to about 25% for the (noninstitutionalized) over 65 years old group. Heart conditions increase from 2.4% for the under 17 group to 23.4% for the over 65 group (these rates are significantly higher depending on the mode of data collection [self-reports versus medical expert rating; Jette, 1996). Visual and hearing disorders also increase dramatically with age. About one third of all persons 65 and older exhibits a significant hearing loss; about 8% suffer from severe age-related visual impairment. The prevalence of mental disorders also increases with age. The percentage of persons with major depression is about 11%; the percentage of elderly with at least some degree of dementia also hovers around 11%. Strokes, degenerative neurological disorders, pulmonary disorders, and most orthopedic disabilities increase with age at a similar rate. Also, there is convergent evidence that these prevalence rates increase dramatically from the young-old to the old-old and the oldest-old and are especially high in institutionalized settings (e.g., Jette, 1996; Suzman, Willis, & Manton, 1992; Schneekloth & Müller, 1995).

Whereas acute illnesses normally can be treated quite adequately in many old and very old persons, the burden of chronic illness (as well as the disabilities and functional losses in daily life that result from them) have become one of the most dramatic challenges for society, for gerontology in general, and for rehabilitation with elderly clients in particular. If our basic goal is to engender a "good" life into old age by fostering "active life expectancy" among the elderly (e.g., Crimmins, Hayward, & Saito, 1996), then rehabilitative work as the process of maximizing abilities after a disabling event (Kemp, 1996) gains crucial importance for the individual as well as for society.

Nevertheless, the burden imposed by chronic conditions, and the rehabilitative efforts to counteract them, deserve a highly differentiated analysis. Oversimplifying statements, whether positive or negative with respect to the power of rehabilitation to improve the lives of the elderly, are detrimental to research and can wreak havoc with the hopes and fears of the frail and chronically ill elderly and their families. On the one hand, the metamorphosis of gerontology from a decline and deficit view of old age toward a competence view that stresses the potentials and reserve capacities of old age has influenced significantly the field of geriatric rehabilitation. The body of empirical evidence underscores the success of rehabilitation in the elderly with respect to stroke, arthritis, cardiovascular disease, and falls (e.g., Andrews, 1987; Harris, O'Hara, & Harper, 1995; Kemp, 1985, 1996; Meier-Baumgartner, Nerenheim-Duscha, & Görres, 1992; Rubenstein et al., 1984; Schütz, 1993; Tallis, 1992). In addition, there is evidence that rehabilitation leads to better outcomes in the short and long run as compared to traditional medical treatment. On the other hand, convergent evidence from longitudinal and catamnestic studies show that the success of rehabilitation is tempered by everyday factors; further deterioration due to the underlying disease process, a lack of rehabilitation capacity in the homecare setting, poor ecological circumstances, or dependence-supporting behavior of the social environment (e.g., Horgas, Wahl, & Baltes, 1996; Meier-Baumgartner et al., 1992; Wahl & Baltes, 1993) can have a lasting, negative impact on the rehabilitation process.

With this in mind, it is accepted broadly in the academic field that rehabilitation with the elderly deserves additional research efforts in order to maximize its efficiency. For example, special problems that arise when treating the elderly must be considered, such as multimorbidity, concomitant cognitive impairments, reduced social networks, poor ecological conditions, longer time periods to regain certain functions, and lowered motivation ("I am too old; this is not worth the effort."). There is also general consensus within the professional field that rehabilitation truly deserves a multidisciplinary approach. The multiprofessional team should include high quality medical personnel, a psychiatrist, well-educated nursing staff, physical therapists, occupational therapists, speech pathologists, social workers, nutritionists, audiologists, and optometricians as well as experts in housing, design, and technical devices (Kemp, 1996; Saltz, 1995; Zeiss & Steffen, 1996). There is additional consent that rehabilitation of the elderly should take place in different settings such as the hospital, day-care center, home health care, and nursing home settings.

What is stressed in this chapter is that psychology as a research and applied field is essential to geriatric rehabilitation in all of these settings. The critical importance of addressing geropsychological and psychosocial issues in the field of geriatric rehabilitation is underscored by the two contrasting theoretical approaches briefly alluded to above. With respect to the competence view of rehabilitation in old age, the emergence and application of psychological intervention techniques such as behavioral or cognitive training programs has broadened significantly the scope of rehabilitation in different settings (in the nursing home setting in particular; e.g., Hussian & Davis, 1985; Wisocki, 1991). Also, there is general consensus that the differential success of rehabilitation programs has as much to do with underlying psychological and social issues (e.g., motivational issues, personality traits, sense of self-efficacy, ability to mobilize social support) as with the "objective" severity of the disease process. With respect to the limits and the fragility of rehabilitation success, the psychosocial perspective points to crucial aspects such as inadequate housing, certain neighborhood characteristics, and the attitudes and behaviors of significant others who may undermine the elderly's striving toward independent functioning (Baltes, 1996).

This chapter was written by two geropsychologists and thus presupposes that rehabilitation with the elderly can be effectively provided if a multidisciplinary and multiprofessional approach is applied. The task is to highlight the psychological and psychosocial perspective within geriatric rehabilitation. This is done in four sections: (i) general principles and geropsychological essentials of rehabilitation, (ii) psychosocial issues and strategies of rehabilitation with older adults (assessment issues, restoration of functional independence, cognitive retraining, psychotherapeutic intervention, motivation and compliance, role of social support, ecological aspects), (iii) psychosocial issues with respect to particular client groups (stroke, arthritis, the sensory impaired, demented patients), and (iv) final consideration of practical, research-related, and ethical problems of rehabilitation in old age.

7.23.2 GENERAL PRINCIPLES AND GEROPSYCHOLOGICAL ESSENTIALS OF REHABILITATION IN OLD AGE

7.23.2.1 General Principles

Rehabilitation is required whenever an impairment or a disability causes permanent, widespread disruption of the normal fit between person and environment. Central to understanding rehabilitation is the distinction between impairment, disability, and handicap (World Health Organization, 1980). Impairment is any loss or abnormality of psychological, physiological, or anatomical structure or function. Disability is any restriction, due to impairment, which limits the ability to perform an activity in the manner or within the range considered normal for a human being. Handicap is defined as a disadvantage resulting from an impairment or disability that limits or prevents the fulfillment of a role appropriate to a person's age, gender, social, and cultural circumstances. For example, a stroke might lead to pathological changes in the brain and quite often to unilateral paralysis (impairment). These stroke-induced impairments are responsible for deteriorations in walking, manual dexterity, memory, and speech performance (disability). These disabilities in turn lead to social consequences for the individual, preventing the person from participating in social roles and activities (handicap).

Rehabilitation can be regarded as a problem-solving and educational process aimed at reducing the disability and handicap experienced by someone as a result of disease, one that is always limited by the resources available (World Health Organization, 1980). Hence, rehabilitation is a set of procedures located between acute medical treatment and prevention. Of course, curable diseases should be treated medically, but whenever possible one should prevent disease instead of dealing with disabling and handicapping consequences. This statement is by no means trivial, as epidemiological data on stroke shows. Preventive measures are responsible for decreasing incidence rates for cardiovascular stroke (e.g., Svanborg, 1988; see also Fries, 1990). Nevertheless, rehabilitation is necessary in those cases where prevention or medical treatment are not effective. Hence, the focus of rehabilitative interventions is not the (underlying) disease but the functional level (e.g., Kemp, 1996). A medical diagnosis, that is knowledge about a certain impairment, does not necessarily imply knowledge about resulting disabilities and handicaps. Indeed, pronounced discrepancies in the emotional or functional outcomes between different patients suffering from "objectively" comparable diseases are more often the rule than the exception and require careful attention.

Rehabilitation is not only aimed at restoring the patients' rights, privileges, reputation, and/or proper condition, but also at realizing their potential, reablement, and resettlement (Ebrahim, 1992). However, rehabilitation with the elderly does not necessarily mean *restitutio ad integrum* (complete cure of disease, full elimination of handicap) but quite often *restitutio ad optimum* (improving functional capability and maximizing wellness). According to Williams (1984), rehabilitation in old age may be characterized as a set of techniques as well as a point of view or a philosophy. In accordance with this philosophy, the primary aim of rehabilitation is to restore an individual to their former functional status, or to maintain or maximize independent functioning (Lee & Itoh, 1990; Williams, 1994). Getting elderly clients to transfer themselves from the bed to a wheelchair may thus be a great rehabilitation success and have tremendous impact on the elderly persons' perception of life. At younger ages, the main goal of rehabilitation is the restoration of the individual's capacity to resume family and work roles. However, these are not necessarily the main goals of rehabilitation in old age. Rather, functional independence (self care and instrumental activities of daily living), the cultivation of leisure activities, and a sense of well-being, that is, improvement of life quality, are the main objectives of rehabilitation in old age (Andrews, 1991; Kemp, 1996; Mosqueda, 1996).

The general goals of geriatric rehabilitation are to control the underlying disease process, prevent secondary disabilities, restore functional abilities, facilitate the patient's adaptation, educate the patient regarding impairment, and modify or otherwise prepare the environment (Kemp, 1985; Saudan, 1990). The process of rehabilitation can be divided into five formal components (Ebrahim, 1992; Wade, 1992):

(i) *Assessment*. In addition to medical diagnosis, the identification, analysis, and measurement of problems in terms of disabilities and handicaps is an essential first step in the rehabilitation process.

(ii) *Planning*. This is extremely important in setting long-term, intermediate, and short-term goals. It is necessary to involve the patient as early and as intensively as possible in this planning procedure, since the acceptance of rehabilitation goals is important for motivation and compliance.

(iii) *Treatment*. Having identified the problems and formulated goals, the proper intervention should take place in order to reduce disability and minimize handicap. Treatment might consist of interventions aimed at the restoration of functional independence, cognitive retraining, psychotherapy or mobilization of social support.

(iv) *Care*. Although rehabilitation is aimed at restoring maximal functional capacity, very often care is necessary to support elderly persons and alleviate the consequences of disability.

(v) *Evaluation*. Great importance, although not a standard procedure in many rehabilitation settings, is the careful evaluation of the effectiveness of intervention which should consist of individual reassessment during or after the rehabilitation process (Görres, 1994).

Psychosocial and clinical-psychological issues play an important role in all of these components. Thus, rehabilitation should be based on a comprehensive psychological profile of the elderly client that includes neuropsychology evaluation as well as the assessment of personality, emotional-motivational functions and personal goals ((i) and (ii) above, assessment and planning). Individuals also need psychological support in order to cope with the existential crises normally following the onset of disability; treatment should focus on fostering behavioral, cognitive, and emotional changes, working on specific coping strategies, promoting self-esteem, improving family relations, and mobilizing other social network resources ((iii) and (iv) above; treatment and care). Finally, everyday life after rehabilitation should not only be seen in terms of functional gains but also in terms of life quality (subjective well-being, self-efficacy, and positive future expectations; (v) above; evaluation).

The goals of geriatric rehabilitation call for a thorough understanding of clinical psychology. Likewise, concepts from geropsychology and lifespan developmental psychology are fundamental to any rehabilitative endeavor. In the following pages, the theoretical foundations of geriatric rehabilitation are reviewed in some detail, before turning to psychosocial issues and strategies of rehabilitation with the elderly.

7.23.2.2 Geropsychological Essentials of Rehabilitation in Old Age

In practice, rehabilitation efforts will only be effective when essential concepts from gerontological and geropsychological research have been given due consideration. Because aging is a lifelong process, it is suggested that lifespan developmental psychology is crucial to gerontology and can contribute greatly to these conceptual essentials (e.g., Baltes, 1987; Lehr, 1993; 1996; Thomae, 1979).

First of all, the rehabilitation of the elderly should not be based on chronological age (or old age) per se but on aging. That is, the emphasis should lie on aging as a dynamic process which can be characterized, like any period of the lifespan, as the simultaneous occurrence of potentials and limits, gains and losses, continuity and change (Baltes, 1987; Thomae, 1979). A variety of studies within geropsychology and developmental psychology stresses the elderly person's potential for growth and development, contrasting starkly with a one-sided picture of mere decline (for overview, see Kruse, 1987; Lehr, 1996). Even when faced with severe competence losses, elderly persons are not reactive, helpless, or passive; instead, the majority of these elderly use a whole scope of coping strategies to deal actively and differentially with threats to their independence and life quality. Rehabilitation of the elderly should thus count on these coping capabilities to address severe illnesses or disabilities. Also, a major task of rehabilitation lies in the reactivation of the elderly person's potential to deal actively with their health threats, thereby regaining as much independent functioning in everyday life as possible. The strategy of selective optimization with compensation is a general adaptational tool to counteract detrimental health changes (e.g., Baltes & Baltes, 1990). From this point of view, rehabilitation in old age can be understood as the transmission of powerful compensation techniques and a mutual search for selective optimization opportunities (particularly in those cases in which the impairment and resulting disabilities can only be altered marginally). To conclude, rehabilitation in old age must base itself upon the insight that elderly persons possess certain fundamental physical and mental potentials (Kruse, 1992); agism must be scoured thoroughly from the rehabilitation effort.

Furthermore, the pronounced interindividual variability of the elderly must also be considered. Persuasive empirical indications from cognitive, personality, and social network research, show that heterogeneity is one of the defining criteria of the aged population (e.g., Nelson & Dannefer, 1992). The general implication for rehabilitation is that a highly individualized approach is necessary. In particular, there is reason to assume that the estimation of the rehabilitation potential deserves an individual point of view and can be significantly different even when the underlying disease (such as a certain type of stroke) is relatively similar between cases (Kruse, 1992). The mere fact that some individuals possess lifelong expertise in particular areas (whereas other individuals do not) may have tremendous implications for rehabilitation. The acknowledgment of variability also points to the need that rehabilitation efforts should not only focus on impairments and disabilities but also on remaining capacities and competencies.

Rehabilitative work must also view aging as an inherently biographical process (Lehr, 1996; Kruse, 1992; Thomae, 1996). Old age is a significant part of the individual biography and life course. Over the course of their lives, individuals develop particular attitudes and coping styles in dealing with everyday challenges and critical life events, such as the onset of severe health problems. Rehabilitation must not neglect these personal styles in dealing with illnesses; it may mean the difference between gaining a powerful tool to understand the idiosyncrasies of the elderly individual and jeopardizing the success of the rehabilitative effort. Individual development, biographical experiences, and evaluative standards also stress the importance of understanding the subjective interpretations assigned to competence losses in everyday life. It is commonplace in clinical work for subjective evaluations of illnesses, impairments, and disability to differ significantly from the ratings of physicians, clinical psychologists, and other professionals involved in the rehabilitation process. One consequence of these discrepancies can be that different goal attainment priorities between the elderly in rehabilitation and their professional surroundings emerge which may delay or even undermine rehabilitation success.

According to leading scholars in the field of lifespan developmental psychology (e.g., Baltes, 1987), aging must also be seen as a contextual phenomenon. One set of powerful factors is social influence. It is known from a body of classical work in gerontology that the course of aging is influenced by the way social networks have been arranged and maintained. Also that social contact means, on the one hand, mutual exchange, personal warmth, the opportunity to give and to receive, and in a sense even physical and cognitive training. On the other hand, social surroundings may hinder the elderly individual from using or exercising capabilities and may thus contribute to a decline in functioning. Rehabilitation should always consider the social embeddedness of aging processes, especially with regard to severe health problems. Thus, there is a need to integrate significant others in the rehabilitation process. These persons represent concrete links with the everyday world in which the elderly person will return; they may deserve training and education about how to interact with the elderly person after the rehabilitation program.

Aging is, however, also aging in place, that is, aging is also influenced by the physical environment (e.g., Lawton & Nahemow, 1973; Parmelee & Lawton, 1990; Saup, 1993; Wahl, 1992). Rehabilitation has to consider the simple fact that elderly faced with severe losses in competence do not normally age in clinical or institutional settings but in their own "world" at home. From this perspective, it is one of the challenges of rehabilitation to intervene as effectively as possible (which, in many cases, can only be done in a formal rehabilitation setting) but nevertheless as close to the everyday world of the elderly person as possible.

Finally, a very important construct in the rehabilitation of the aged stems from lifespan developmental psychology: plasticity (e.g., Baltes, 1987; Kruse, 1992). This term signifies that the course of aging is not predetermined or fixed, regardless of whether the process in question is one of "normal" aging or aging coupled with serious illnesses and disabilities. For example, cognitive training studies have underlined clearly the potential for significant gains which can sometimes exceed the level of untrained younger persons (Baltes, 1993; see also below). So-called control-enhancing interventions have proven their physical and psychological benefits even in frail elderly populations like nursing home residents (e.g., Rodin, 1986). Aging obviously is not a fixed and predetermined process but is subject to many different types of influences which may lead to differential outcomes. Rehabilitation may thus be regarded as one of the most powerful attempts to systematically influence aging organisms challenged by chronic disease processes.

In conclusion, the fact is emphasized that the fundamental principles and concepts of geriatric rehabilitation rest upon clinical as well as geropsychological insights. This should be kept in mind as some of the main psychosocial issues and strategies of rehabilitation in old age are discussed.

7.23.3 PSYCHOSOCIAL ISSUES AND STRATEGIES OF REHABILITATION IN OLD AGE

7.23.3.1 Assessment Issues

There is wide agreement that only comprehensive geriatric assessment (CGA) can serve as a basis for successful rehabilitative work with the elderly (e.g., Stuck, Siu, Wieland, Adams, & Rubenstein, 1992). Although not expressed explicitly in this widely used term, the assessment of psychological functioning is as a very important part of CGA. Furthermore, geropsychological knowledge and research can contribute to the shortcomings of CGA as it is commonly conducted in most rehabilitation settings. It should be mentioned that the aim of this section is not to give a complete overview of cognitive and neuropsychological assessment (which undoubtedly is also critical with respect to geriatric rehabilitation and its outcomes; see Chapter 7, this volume; see also Franzen & Martin, 1996; Rentz, 1991); instead, issues beyond the assessment of cognitive functioning will be discussed (cf. Lawton & Teresi, 1994). Comprehensive geriatric assessment has at least three major tasks: (i) to measure individual functional capacity, that is, losses as well as competencies at the beginning and during the course of the rehabilitation process; (ii) to contribute to the planning and prognosis of rehabilitation; and (iii) to serve as a tool for quality assurance and evaluation. To fulfill these tasks, CGA is basically aimed at the differential description of disease-related impairments to everyday life and functioning (see again the functional approach as one general principle of rehabilitation, above).

7.23.3.1.1 Activities of daily living and instrumental activities of daily living

One core issue of CGA is thus the estimation of decline and competence in activities of daily living (ADL) of varying complexity. In most research and practical work, basic ADL such as the capacity for eating, washing, or dressing are differentiated from instrumental activities of daily living (IADL), such as shopping, phone use, banking, or the intake of medications. A whole range of rather similar scales for the measurement of ADL and IADL can be found in the psychological literature. With respect to the basic ADL, the Barthel index has found wide use in geriatric rehabilitation (Mahoney & Barthel, 1965); prototypical IADL are represented in the widely applied Lawton and Brody scale (Lawton & Brody, 1969).

All of these scales are short and easy to apply. However, there are also problems: for one, the assessment of functional capacity is based mostly on face validity and still lacks theoretical foundation. It is thus strongly argued for closer links between the ADL/IADL literature (including their respective assessment devices in the clinical setting) and emerging theoretical and empirical research on everyday competence (e.g., Baltes, Maas, Wilms, & Borchelt, 1996; Willis, 1991, 1996). In particular, this latter work supports the notion that everyday competence depends heavily on personal and environmental variables that lie beyond chronic illness conditions such as cognitive performance and environmental barriers. One assessment implication of this insight is the need to "diagnose" the physical environment and always interpret a given degree of competence or impairment in everyday functioning (as assessed based on the common ADL/IADL approach) as the result of a transaction between personal capabilities and environmental circumstances. Furthermore, the existing everyday competence literature suggests that basic ADL and IADL are fundamentally different: the former are closely linked to basic physical capacity, whereas the latter depend heavily on motivation and socialization (e.g., preparing meals in elderly men). Second, functional capacity and thus the assessment of everyday performance should not only be focused on ADL and IADL; the capacity to solve everyday problems effectively (Willis, 1996) and to perform leisure activities in the social and nonsocial domain deserve to be assessed as well (e.g., Baltes et al., 1996) in order to gain a comprehensive picture of the everyday implications of chronic disablement.

7.23.3.1.2 *Performance vs. reserve capacity*

ADL/IADL scales can measure performance in functioning but not remaining reserve capacity, that is the individual's potential for further growth, which can be activated during the rehabilitation process. Although there is no *via regia* to approach this issue (which is a crucial one for planning the rehabilitation process), one important aspect in this regard is to consider the personality characteristics of the elderly client (see below).

Obviously, it would not make very much sense to expect that the reactivation of remaining reserve capacities always occurs under ideal personal and environmental conditions. Although it is always possible to alter personality (e.g., by psychotherapeutic intervention) or modify the social and physical surroundings, an adequate and realistic estimation of the current personal and living situation of the elderly client is indispensable. For example, anamnestic exploration should not only focus on the medical history but also on personal motivations, biographical key experiences with relevance for the coping with the current rehabilitation process, subjective evaluations of personal competencies, and coping styles (see the geropsychological essentials of rehabilitation described above). Another important area encompasses personality variables such as introversion—extraversion or anxiety. Both of these variables probably influence the rehabilitation process in different ways. For example, the personality characteristic of introversion versus extraversion may strongly influence the motivation to regain mastery over daily living tasks. Also, trait as well as state anxiety may undermine rehabilitation and transferral to the home situation if not taken seriously.

The results of all of these assessments can be crucial for the development of the final rehabilitation plan. Rehabilitation should always be done in close cooperation with the individual because the implementation of relevant goals and strategies depends heavily on the client's insight, motivation, and compliance. According to the contextuality of the aging process and the docility of the chronically ill organism to "pressing" environmental conditions (as stated quite early on in the so-called Environmental Docility Hypothesis, see Lawton & Simon, 1968), assessment should also include the description of the physical home environment (see, for example, Regnier & Pynoos, 1987). Also, the measurement of rehabilitation outcome should be regarded as more than mere ADL/IADL improvement. The basic aim of rehabilitation to contribute to life quality in general deserves the assessment of subjective well-being, depressivity, or of mood states as well (Kemp, 1996).

7.23.3.2 Restoration of Functional Independence

7.23.3.2.1 *Functional independence as basic objective of geriatric rehabilitation*

The restoration of functional independence can be regarded as one of the most important, though not the only, objective of geriatric rehabilitation. Independent functioning in daily life is highly valued within (Western) society; hence, the preservation or restoration of independence represents one of the most important life goals of most elderly. Basic ADL functions to be considered are self-care (such as eating, drinking, dressing, grooming, washing, and controlling urination and bowel movements) and mobility (getting in and out of bed, chair, or toilet, walking, and climbing stairs). For community dwelling elderly persons, being capable of performing instrumental activities of daily living is extremely important (doing household chores, shopping, cooking). There is also reason to assume that independent functioning in these daily living tasks is associated closely with related personal variables such as perceptions of control, self-efficacy, and competency; thus, personal autonomy can contribute much to good psychological adaptation in the course of aging.

Obviously, the majority of efforts within the multiprofessional team, such as physical or occupational therapy, is aimed toward improving functional capacity. It must once again be emphasized here that this goal does not necessarily entail the restoration of full independence. As Kemp (1996) has put it, to manage the step from "maximum assist" to "partial assist" can mean a dramatic gain for the elderly individual as well as their social environment. Most importantly, this "small" step can enable the individual to remain in a private household situation and to avoid relocation to a nursing home setting. Reliance upon functional criteria in rehabilitation is reliance upon personal variables, and the latter repeatedly have proven themselves superior to mere medical diagnoses in the prediction of important outcome variables (e.g., nursing home entry, daily life quality, life satisfaction, and even mortality; e.g., Kemp, 1996).

7.23.3.2.2 Contributions from clinical geropsychology

It would be shortsighted, however, to view the restoration of independence solely from a physical training perspective. Although bodily functions such as gross and fine mobility are important for performance of ADL and IADL (and should, thus, be trained by different means during geriatric rehabilitation), there is reason to assume that psychological factors are important as well. One factor is cognitive performance (e.g., Baltes et al., 1996; Willis, 1991); cognitive retraining can contribute significantly to the regaining of everyday competence. Also, since striving toward independence has much to do with the motivational and goal-directedness of each client, counseling and psychotherapeutic interventions can maximize the effects of medically-oriented treatments such as physical therapy and training in basic mobility (see respective sections below).

The restoration of functional independence in daily living tasks is also a prototypical case for the effectiveness of behavioral modification techniques based on social-learning principles, that is, it underscores the worth of a genuinely psychological intervention strategy. A whole body of research has demonstrated that such techniques (e.g., the use of different reinforcement procedures) can contribute significantly to the maintenance or restoration of functional independence in the elderly (e.g., Hasselkus, 1989; Smyer, Zarit, & Qualls, 1990; Wisocki, 1991). These approaches have particular rehabilitative value in institutionalized settings. Not long ago, the institutionalized elderly more or less led passive, sheltered lives under custodial care. In the 1990s, there is a basic consensus that institutions should be able to promote successful aging; they are not only expected to provide a secure and prosthetic environment, but also to stimulate autonomy (Baltes, Wahl, & Reichert, 1991). As was shown in a research program, the etiology and maintenance of dependence in institutions is related closely to the social environment, that is, the dependence-supportive behavior of staff and others in institutionalized settings. The systematic change of these interaction patterns by means of cognitive and behavioral training with staff led to a significant increase in the independent functioning of nursing home residents (e.g., Baltes & Wahl, 1996). Because a strong tendency has been found toward independence-supportive behavior within the home health setting (e.g., Baltes & Wahl, 1992), rehabilitative efforts based on cognitive and behavioral motivations should not be reserved solely for institutional settings.

Functional loss is thus not only a characteristic of the person but always a feature person–environment interchange. Naturally, this fact carries implications for rehabilitation. Disability may be reduced by increasing personal competence, changing situational demands, or employing effective compensation strategies (e.g., simplification or redesign of the task; see also Dixon & Bäckman, 1995). Due to practical constraints in terms of time, effort, or money, rehabilitation aimed toward the restoration of functional independence thus is forced frequently to make a decision of what is more promising, that is, to change the person, to change the environment, or to change the task performance itself. Assessment issues like the ones described above (i.e., information pertaining to the elderly client as well as their social and physical environment) is particularly helpful in selecting the best possible from these alternatives.

7.23.3.3 Cognitive Retraining

Cognitive capacity can be divided into two broad ability categories; fluid intelligence and crystallized intelligence (Baltes, 1984; Cattell, 1971; Horn & Hofer, 1992). Fluid intelligence, also called cognitive mechanics, encompasses the most basic cognitive processes and is measured mostly using memory tasks, problem-solving tasks, or perceptual speed tasks. Crystallized intelligence, or cognitive pragmatics, encompasses the experience and knowledge-related parts of cognitive capacity and is measured using knowledge or fluency tasks. Several studies have shown a decline in fluid intelligence already by middle age, while crystallized intelligence shows stability in adulthood and declines markedly only in old age (Salthouse, 1991; Schaie & Willis, 1993; Smith & Baltes, 1996). Such cognitive decline in old age seems characteristic of normal aging; it is not confined to elderly persons with a diagnosis of dementia (Reischies & Lindenberger, 1996), although it is more severe in such cases. Dementia is a syndrome due to brain disease in which there is impairment of multiple higher cortical functions, including memory, thinking, orientation, comprehension, learning capacity, and language (see also Chapter 10, this volume). The prevalence rates for dementia range from 3–7% for the population over 65 years, and increases dramatically with advancing age (Cooper, 1991a). Given cognitive decline in old age, and especially in the case of dementia, it is worth investigating whether (re)training might possibly help with these age- and/or disease-related losses.

7.23.3.3.1 Intelligence training

Experimental studies have demonstrated that cognitive interventions in old age are successful (Baltes & Lindenberger, 1988; Baltes,1993). In several studies, it could be shown that the cognitive performance of elderly people is modifiable (Baltes, Sowarka, & Kliegl, 1989; Blieszner, Willis, & Baltes, 1981; Willis, Blieszner, & Baltes, 1981). In these studies, the emphasis of training was on explaining and practicing rules for solving figural relations tasks and inductive reasoning problems. Even with few (5–10) training sessions, elderly participants increased their test scores in post-intervention assessments. Despite the gains achieved in cognitive training, it seems necessary to ensure the generalizability of these interventions (e.g., training in inductive reasoning, deductive reasoning, fluency, and memory techniques should help elderly people solve practical, everyday problems). Furthermore, the rehabilitative power of combining cognitive training with other intervention strategies (e.g., psychomotor training components, training in everyday problem solving) deserve consideration (see for instance Butler & Namerow, 1988, and their differentiation between cognitive retraining, aimed at improving cognitive "micro-processes," and skill acquisition training as a functional "macro-treatment" with greater rehabilitation impact; e.g., Oswald, Rupprecht, & Gunzelmann, 1998). Elderly persons at risk of dementia or even at low to moderate levels of dementia also can profit from cognitive training, although their training gains are much lower than those of healthy persons (Bäckman, Josephsson, Herlitz, Stigsdottir, & Viitanen, 1991; Baltes & Sowarka, 1995). Plasticity appears to be greatly reduced in dementia, even at the onset. Therefore, it has been argued that cognitive training might be a useful diagnostic tool for the early detection of dementia processes (e.g., the testing-the-limits approach; Baltes, Kühl, Gutzmann, & Sowarka, 1995).

7.23.3.3.2 Memory training

Decrease in memory capacity seems to be a marker of aging, at least in the subjective conceptions of elderly people (Abson & Rabbitt, 1988). More equivocal support for memory decline in old age has been reported (Fleischmann, 1989; Kausler, 1991). Nevertheless, memory training with older participants has been tested carefully. A wide variety of memory tasks (recognition, free recall, or pair-association learning) and memory strategies (including diverse mnemotechniques like verbal categorization, elaboration, visual imagery, and the method of loci; Fleischmann, 1982; Yesavage, Lapp, & Sheikh, 1989) have been examined. In the majority of these experiments, the effects of memory training in elderly people are positive. Interestingly, memory training which combines memory strategies with reattribution training, in order to reduce exaggerated achievement levels and change attribution patterns for success and failure, might help elderly people most (Lachman, Weaver, Bandura, Elliott, & Lewkowicz 1992; Weinert & Knopf, 1990). Memory training can be effective in mildly demented elderly persons as well (Ermini-Fünfschilling & Meier,1995). Still, the generalizability of memory training from laboratory settings to everyday memory functioning has yet to be proven (see also Camp, et al., 1993).

7.23.3.3.3 Reality orientation training

Persons suffering from dementia can profit greatly from training in time and space orientation (Kaschel, Zeiser-Kaschel, & Mayer, 1992). Dementia not only affects higher order problem solving but, in later stages, it also impedes basic functions like orienting in time (year, season, day of the month), space (rooms of apartment or institution), and contacting communication partners (staff members, family). Basic cognitive training encompasses orientation in time and space; also, reinforcement of the mechanics and pragmatics of intelligence may be included depending on the cognitive status of the rehabilitation client (e.g., Bäckman et al., 1991; Koh et al., 1994; Little, 1991; Rentz, 1991). Speech and communication therapy might also be very helpful in demented patients or patients suffering from stroke (Sarno, 1984; Jordan, Worral, Hickson, & Dodd, 1993). However, persons with dementia do not only profit from cognitive training in the strict sense of the word. A reactivating occupational therapy program can complement functional rehabilitation in long-term geriatric inpatients with slight to moderate dementia (Bach, Bach, Böhmer, Frühwald, & Grilc, 1995).

To summarize, it can be stated that although younger persons profit even more from cognitive interventions, intelligence and memory training are effective in the healthy elderly. Relatively smaller training effects have been demonstrated in the demented elderly, also. Despite the successes reported in such research, cognitive interventions which fit the everyday needs of elderly people are still the exception. Practical problem solving like reading (and understanding) public transportation schedules might be a good example (Schaie & Willis, 1986). It is also unclear whether cognitive interventions have substantial, long-term consequences. Especially in the case of dementia, however, such interventions might support the person and their caregivers in getting along with the ever-increasing complexity of everyday life.

7.23.3.4 Psychotherapeutic Intervention

Disability and handicap mean a more or less sudden decrease in the functional capacity of the elderly person; this may severely threaten mental balance and stability. Most threatening for elderly persons is the anticipation that living conditions must be changed dramatically in the near future (such as the relocation to an institutional setting) or that future life will only be possible at the cost of great dependency upon others. The psychological burden of losing one's own independence is compounded by the anticipation of becoming a burden to others, thereby negatively influencing the life course of loved ones. There is also broad evidence in the depression literature that severe physical chronic conditions are correlated significantly with depressive episodes (e.g., Schulz & Williamson, 1993; see also Gatz, Kasl-Godley, & Karel, 1996); in other words, besides functional impairments, the identity and self-esteem of the handicapped person are also threatened.

Geriatric rehabilitation seldom calls upon psychotherapy to play the role it should in addressing the psychological crisis experienced by the disabled elderly. Rehabilitation should, however, rise to meet the challenge of helping the elderly person to cope successfully with their disabilities and handicaps. To restate the argument: psychologists have more than only diagnostic tasks to fulfill in geriatric rehabilitation.

Because of the time limitations of normal rehabilitation efforts, individual and group psychotherapy based on a cognitive-behavioral philosophy seem particularly promising to deal with these problems, especially with depressive episodes (Hartke & Horowitz, 1991; Smyer, et al., 1990). However, treatment of depression is only one focus. There are further challenges having to do with the optimization of the rehabilitation success as well as with the maintenance of outcomes, particularly when transitioning from the hospital to the home situation. One task is to enhance motivation in certain clients, particularly in those who have started rehabilitation based on an externally induced motivation ("My wife suggested to try it."). Another approach with special importance for the maintenance of rehabilitation success in the "real world" employs coping skills training (see also Rodin, 1986). This powerful preventive tool serves to develop strategies and train the individual to deal with daily hassles, stress, and possible failure, all of which occur frequently after returning home.

Also, counseling with the goal of enhancing feelings of self-efficacy and competence (despite severe disabilities) may be an important task for clinical psychologists in geriatric rehabilitation. In some cases, psychological support with the goal of learning to accept irreversible losses is critical as well. Social competence training may be particularly helpful in those elderly living alone and running the risk of becoming still more isolated after the onset of chronic illnesses and the disabilities and handicaps they entail. Finally, counseling the relatives of the elderly clients (particularly those living with the client) may be crucial for the maintenance of rehabilitation success. If it is true, as research shows, that a dependence-supportive interaction pattern prevails in the private household setting (as well as in institutional settings), then the counseling and education of the persons in the elderly client's social sphere obviously becomes most important (Baltes & Wahl, 1992; Wahl & Baltes, 1993).

A point frequently not taken as seriously as it deserves should be mentioned at the end of this section, namely, rehabilitation with the mentally ill elderly (e.g., Häfner, 1986; see also Carstensen, Edelstein, & Dornbrand, 1996). The special rehabilitation needs of the mentally ill (e.g., those suffering from major depressive disorders or psychotic disturbances) are seldom acknowledged, even among professionals. Researchers and professionals in geriatric rehabilitation should be encouraged to give this subgroup of the chronically ill elderly the attention they rightfully deserve.

7.23.3.5 Considering Motivation and Compliance Issues

It is a truism to state that the motivation and compliance of the elderly person are necessary conditions for successful rehabilitation (King & Barrowclough, 1989; Ellul, Watkins, Ferguson, & Barer, 1992; Hartke, 1991; Trezona, 1991). The elderly person has to be convinced that participating in a rehabilitation program is worth the effort and cost. Hence, it is necessary to understand the personal attitudes and intentions of the person before starting the rehabilitation process (see, for example Wiesner & Tesch-Römer, 1996).

7.23.3.5.1 Motivation

Although the problem of motivating the elderly patient is often mentioned in the rehabilitation literature, rarely are specific measures for increasing motivation described. Quite often, elderly people are convinced that rehabilitation is not worthwhile, that improvement is impossible, that too much effort is required, or that the results of rehabilitation might threaten the current safe (albeit dependent) life situation. Improving motivation can be achieved through communication with the elderly person before and during rehabilitation, giving credence to individual goals, and simplifying steps needed to achieve rehabilitation objectives (Kemp, 1985). Specifically, changing pessimistic self-efficacy beliefs, strengthening self-esteem, increasing optimism, and exploring "possible selves" might be helpful in increasing motivation for rehabilitation (Guthrie & Harvey, 1994). Important aspects of the rehabilitation process that might enhance motivation are information (What should one know about the intervention?), choice (Which options are available?), goal-setting (What are long- and short-term goals?), attention to emotional needs (What does the disability/handicap mean to the person?), and dealing with social anxieties (What do other people think/expect?). Finally, it should be noted that lack of motivation could also be due to an endogenous depressive disorder. Similarly, motivational deficits can occur due to reactive depression following loss sustained in the initial illness or even a specific brain disease (like frontal lobe disorder, Vanderpool, 1984). In these cases, specific therapeutic interventions might be necessary.

7.23.3.5.2 Compliance

It is not enough that the person is motivated to start rehabilitation, they must also comply with the rehabilitative process, which can take weeks or months. Problems during the intervention process may result from low motivation, psychosocial crises, or cognitive decline (e.g., memory problems). Hence, elderly persons in rehabilitation should be helped to maintain a schedule if necessary. This can be achieved by keeping short-term goals clear and giving simple instructions for rehabilitation procedures. Moreover, motivation should be maintained by regularly encouraging and reinforcing the rehabilitation progress. Examples for the importance of compliance issues can be found in prosthetics use (Cameron & Quine, 1994), implementing safety features at home (El-Faizy & Reinsch, 1994), and continence training (Fonda, 1991).

7.23.3.5.3 Impact of the social network on motivation and compliance

The social network of the patient is very important in rehabilitation motivation and compliance. Motivation is not a personal characteristic, but a product of interpersonal interactions. Hence, motivation can be influenced significantly by rehabilitation staff, family members, and community services (Tucker, 1984). An example might be shaping the self-care behavior of elderly nursing home residents via social feedback from staff members (Baltes, Burgess, & Stewart, 1980; Baltes, Kindermann, & Reisenzein, 1986). It should be noted, however, that it is difficult for health professionals in institutional settings to rate rehabilitation motivation; rehabilitation decisions can be facilitated by describing specific behaviors, like mobility, self-care, independence, and socialization (King & Barrowclough, 1989).

Before beginning a rehabilitative intervention, it would seem useful to instill a sense of confidence in the elderly client, to persuade them that the long road ahead is not so very difficult and well worth the effort. Motivation is a complex product of the interaction between a person's values, the values held by significant others, history, talents and interests, personality, and treatment setting. As a rule of thumb, one should emphasize that if a derogatory, devalued label is applied to the person (either by the person or by the social network), rehabilitative efforts will fail; conversely, if the person is given positive feedback, the individual's motivation for rehabilitation will be enhanced.

7.23.3.6 The Role of Social Support

Social relations change with advancing age. The chance of living with a partner, especially for women, decreases, the size of the social network shrinks, and loneliness increases with advancing age (Wagner, Schütze, & Lang, 1996). Elderly people living in institutions are especially vulnerable to accumulated risk factors for social isolation and loneliness. Nevertheless, important aspects of the "social convoy" remain stable even in old age (Antonucci, 1985). Most elderly people have good relations and frequent contact with children and grandchildren, and elderly people at all ages name friends, acquaintances, or neighbors as parts of their social network. Moreover, elderly people are not solely receivers of social support; they also give support to members of their social network. At this point, one should distinguish between emotional and instrumental support. In rehabilitation, one might inquire regarding what role social support plays in the well-being of the disabled elderly, and how social support could be mobilized for the purpose of rehabilitation.

7.23.3.6.1 Effects of social support

There can be no doubt that social support, whether instrumental or emotional, helps the elderly individual cope with chronic illness and disabilities that remain even after the best rehabilitation efforts (e.g., Antonucci & Jackson, 1987; Silverstone, 1984). Several studies show that social support has positive effects upon a broad range of handicaps and in various rehabilitation settings. Chronic pain patients with supportive families tend to report less pain intensity, less reliance on medication, and greater activity levels than patients with non-supportive, disharmonic families (Jamison & Virts, 1990). In stroke patients, emotional support and moderate degrees of instrumental support are related to positive outcomes and recovery (Glass & Maddox, 1992). However, social support might be dysfunctional as well. Too much instrumental support might hinder the recovery of stroke patients. Hence, the positive and potentially negative sides of social support deserve consideration in this context. On the one hand, social partners should be integrated into the rehabilitation process as early as possible, since they are able to contribute significantly to the restoration and maintenance of functional independence of the elderly client. On the other hand, social interactions can also become detrimental to the elderly when social partners tend toward overcare (e.g., Baltes & Wahl, 1996), or care giving situations evoke an overflow of problematic psychodynamic elements (see Pearlin, Aneshensel, Mullan, & Whitlatch, 1996).

7.23.3.6.2 Mobilizing social support

Social relations (and the sources of social support) can be divided into informal and formal support givers (Silverstone, 1984). Among the informal support givers are spouses (especially wives), children (especially daughters), friends, and neighbors. The existence of a possible support giver does not necessarily mean that an elderly handicapped person receives help from network members. This is of some importance, because without some type of informal support, a return to the patient's home in the community might be very difficult. The burden (and the support) of the caregivers should also be acknowledged (see Chapter 22, this volume). The support offered by family (and friends) should be supplemented by formal support givers, such as primary health care, rehabilitation services, social workers, homemakers, meals-on-wheels, and so on. It should also be acknowledged that the (possibly short) acute phase of illness is usually followed by a recuperative phase and a (probably long) maintenance phase. Allocation of formal community services should depend on the needs and prospects of the care recipient and amount of resources available. Even in the case of psychiatric disorders, day-care, outpatient care, and community-based care can be implemented (Cooper, 1991b).

Hence, whenever possible, the global goal of rehabilitative efforts is to discharge elderly patients into the community, that is, back to their former, everyday living situation. Therefore, the mobilization of social support is important for helping the elderly person to live as independently as possible but with as much help from social partners at home as necessary. Living in the community and in one's own home means getting permanent and everyday support, even if this support takes the form of "small" gestures and signs from the neighborhood (Wagner, Schütze, & Lang, 1996).

7.23.3.7 Ecological Aspects of Rehabilitation

Although a "gestalt switch" from the elderly person to the environment took place in the 1960s and 1970s (see, for example, Lawton & Nahemow, 1973), gerontology in general and geriatric rehabilitation in particular is still very "person centered." From its inception, environmental gerontology persistently has considered the elderly person's surroundings as protective or risk factors for independent functioning. Of particular relevance here is the Environmental Docility Hypothesis, according to which the elderly suffering from competence losses and chronic conditions are particularly prone to environmental characteristics (e.g., Lawton, 1982; Lawton & Simon, 1968). A few simple statements can be emphasized. Independence in self-care always means independence in a certain bathroom environment (with certain supportive or constraining conditions). Independence in meal preparation always means independent meal preparation in a certain kitchen environment with specific features and equipment. Finally, independence in shopping always means independence within a certain neighborhood or meso-environment with supermarkets of varying proximity.

7.23.3.7.1 Person–environment fit as critical for rehabilitation

There is thus good reason to assume that discrepancies between the competence profile of the elderly person and the environment can undermine even the best person-centered rehabilitation effort (e.g., Lawton, 1982; Olbrich & Diegritz, 1995). This is true, for instance, with respect to the eastern part of Germany (the former German Democratic Republic), in which housing conditions are still significantly worse compared to those in the western part. As empirical analyses have revealed, discrepancies on measures of daily competence are at least in part caused by poor living environments within and outside the home (e.g., long distances to supermarkets and pharmacies) that are prevalent in the eastern regions of the country (Olbrich & Diegritz, 1995; Schneekloth & Potthoff, 1993).

Thus, rehabilitation should always be aimed at both training the elderly client to cope with environmental challenges and changing environmental features to suit their profile of deficits and competencies (e.g., Regnier & Pynoos, 1987; 1992; Parmelee & Lawton, 1990). Rehabilitation may thus be defined as the attempt to reduce all possible discrepancies between personal competence and environmental challenges by focusing on the person and their physical environment.

7.23.3.7.2 Considering objective and subjective environmental aspects within the everyday world

One should also note that the physical environments of the elderly are more than mere objective surroundings which vary in the amount of support they offer. Environments may also be viewed as materializations of personal biographies and as extensions of the personal self (Wahl & Oswald, in press); this concept is particularly relevant for the great number of elderly individuals who do not live for decades at the same location. Thus, alterations may indeed be helpful and necessary for the rehabilitation outcome from an objective perspective, but not from the elderly client's point of view. Geriatric practitioners should take these subjective evaluations seriously; the elderly individual's decision to refrain from modifying their home environment must be considered to be a valid choice, even though it might lower independence in daily functioning in the long run. Furthermore, elderly persons should not be seen as mere pawns of their environmental conditions. The creativity with which the elderly (particularly those suffering from chronic conditions) use their potential for compensation is very impressive; they demonstrate consistently a high degree of "proactivity" in adapting themselves to their living environments and vice versa (Lawton, 1989; see also Wahl & Oswald, in press).

The ecological approach to geriatric rehabilitation may be used to illustrate another basic issue. Environmental gerontology and psychology have always stressed the importance of remaining as close as possible to the everyday world of their subjects (see for example the concept of behavior setting proposed by Barker, 1968); likewise, rehabilitation efforts should also demonstrate a good measure of "ecological validity" (Wahl & Baltes, 1993). Rehabilitative work, particularly when it takes place in a hospital setting, should do its utmost remain close to the "life world" of the elderly individual; this is the most promising way to ensure the maintenance of rehabilitation outcomes in the home world.

7.23.3.7.3 New technologies and their rehabilitation potential

Finally the issue of old age and technology should be mentioned briefly, particularly so-called new technologies (e.g., Bouma & Graafmans, 1992; Charness & Bosman, 1990). What is emphasized here, is that there is no reason to assume that elderly persons are unable or unwilling to use new technologies. Instead, these devices could become major parts of the rehabilitative process in a variety of ways. For example, shopping via a home computer may be advantageous to the elderly person with significant impairment in mobility, and contribute greatly to self-efficacy and feelings of control. E-mail services or fax devices can facilitate the maintenance of social networks and social exchange, particularly when social isolation results from disablement.

7.23.4 PSYCHOSOCIAL REHABILITATION ISSUES WITH RESPECT TO PARTICULAR CLIENT GROUPS

Rehabilitation must change to suit different kinds of handicaps. With this in mind, four client groups, namely stroke patients, arthritis patients, the sensory impaired elderly, and demented elderly, are discussed in more detail in this subsection. Although it is clear that rehabilitation programs exist for other age-related handicaps as well (e.g., amputation, Marmann, 1994; falls and hip fracture, Campbell, 1992; Felsenthal & Stein, 1990), there are at least three reasons to concentrate on these four target groups. First, each of these illnesses represents a chronic condition that is highly prevalent in the population at large. Second, arthritis, sensory impairments, dementia, and to a lesser extent, stroke, do not receive the attention within geriatric rehabilitation which they deserve. Third, each patient group poses a significant, yet unique challenge to the rehabilitation professional.

7.23.4.1 Stroke Patients

Cerebrovascular diseases belong to the most prevalent diagnoses in old age (Steinhagen-Thiessen & Borchelt, 1996). Cerebral vascular accident (stroke) constitutes the third leading cause of death in the United States (following heart disease and cancer). About one third of stroke victims will die within the first three months after the stroke. Symptoms of stroke can be partial or complete paralysis, urinary incontinence, difficulties in swallowing, aphasia, and losses in memory and other cognitive functions (Wade, 1992). The functional consequences of stroke are often serious and concern overall physical ability (activities of daily living), walking, dexterity, and communication (Gibson & Caplan, 1984). Very often mood changes occur; many stroke victims develop depressive symptoms which might negatively influence the rehabilitation process (Goldberg, 1984). Early medical treatment of the stroke is extremely important, and it seems that specialized stroke units reduce morbidity and mortality among stroke victims (Langhorne, Williams, Gilchrist, & Howie, 1993). Once the clinical state has settled, one can expect a period of rapid recovery over the first 2–3 weeks, followed by slower but continuing improvement over the next five months (Wade & Langton-Hewer,1987). There is much debate whether functional improvement occurs after that, but some studies show that continued rehabilitative efforts might be helpful (Dam, et al., 1993). More information on pathophysiological changes can be found in geriatric textbooks (Abrams, Berkow, & Fletcher, 1990; Brocklehurst, Tallis, & Fillit, 1992; Evans & Williams, 1992; Pathy, 1991).

7.23.4.1.1 Critical issues of rehabilitation

Rehabilitation should start immediately after the onset of stroke. From the beginning, the patient should be mobilized and encouraged to be as independent as possible. Rehabilitation efforts should also be directed toward communication disorders, upper extremity function, and bowel and bladder care. Assessment is also crucial to stroke rehabilitation; detecting disabilities, discovering the relevant impairments, determining prognosis, planning the intervention process, and evaluating rehabilitation progress are the tasks of the interdisciplinary stroke rehabilitation team (Kruse, 1992, 1995; Rustemeyer, 1983). Physicians are responsible for medical diagnosis and treatment, although no effective medical therapy can ensure complete recovery from stroke. Hence, a variety of professions are involved in stroke rehabilitation. Physiotherapeutic exercises, such as those which improve range of motion, muscle strength, and endurance, are necessary for dealing with muscle spasticity and improving functional performance (Lee & Itoh, 1992). The Bobath method of treatment has been very influential in stroke rehabilitation, emphasizing the suppression of abnormal postures and movements before normal motor patterns are practiced (Bobath, 1971). Although physiotherapy should start early, even late physiotherapeutic intervention (12 months after stroke) can improve mobility (Wade, Collen, Robb, & Warlow, 1992). Occupational therapy is directed at training compensatory techniques for performing activities of daily living (Hasselkuss, 1989). Speech therapy is aimed at supporting stroke patients to overcome communication problems (Lehmann, 1992; Sarno, 1984; Vignolo, 1964).

Psychologists might carry out several important tasks in stroke rehabilitation. Changes concerning alterations in intellect, personality, and emotional behavior deserve attention in the diagnostic process as well as in psychotherapeutic interventions aimed at the elderly person or their social environment (Kruse, 1984). First, psychological and neuropsychological diagnostics are genuinely psychological tasks and highly relevant for assessing the specific problems of stroke patients (Lawton & Teresi, 1994). Since many patients suffer from depressive symptoms, psychological counseling might help the patient to cope with the losses resulting from the stroke and to recover from depressed mood. Clinical psychologists working with stroke patients need to plan realistic short-term goals, provide structure for working sessions, and carefully confront the patient with the denial of disability (Goldberg, 1984). Also group therapy with other stroke patients and family therapy, involving spouse and other family members, might be helpful.

Ideally, stroke patients should be able to return home after rehabilitation. It is necessary to plan carefully the discharge of the patient, involving family and community-based services as far as possible. If necessary, the patient's home should be altered to accommodate the needs of the stroke patient. For family members, caring for a relative suffering from stroke and its consequences is highly distressing; hence, community services should not only focus attention toward the neuropsychological aspects of stroke patients, but also toward the social functioning of caregivers who support them (Anderson, Linto, & Stewart-Wynne, 1995).

7.23.4.1.2 Effects of rehabilitation and challenges in the future

Since the treatment of stroke in older patients is the rule rather than the exception, research has been devoted to demonstrating the effectiveness of rehabilitation over spontaneous recovery. Meta-analyses show that stroke rehabilitation programs may improve the functional performance of stroke patients (Ottenbacher & Jennall, 1993). The extent of improvement appears to be related to the timely initiation of treatment, but not to the duration of intervention. Improvements are also associated with the patient's age, older patients taking longer to recover from stroke (Falconer, Naughton, Strasser, & Sinacore, 1994; Kalra, 1994). Regarding the institutional setting of stroke rehabilitation, it seems that early care in specialized stroke units might be most effective (Jørgensen et al., 1995; Kaste, Palomaki, & Sarna, 1995). In the further rehabilitation process, day hospitals may be more effective (and also economically advantageous) compared to conventional hospital management (Gladman, Forster, & Young, 1995; Hui, Lum, Woo, Or, & Kay, 1995).

The 1980s and 1990s have seen the establishment of special stroke units within hospitals. What is needed now are efforts for continuing rehabilitative efforts in the community, cooperation between professions, and better gerontological and geriatric education for professionals working in stroke rehabilitation. In particular, clinical psychologists working in geriatrics commonly report scarce academic preparation and training experiences and relate continuing education to higher satisfaction with geriatric work experiences (Lichtenberg & Rosenthal, 1994). The rehabilitation of a patient who has suffered a stroke is a challenge for the geriatric rehabilitation team and requires diagnostic, management, and counseling resources. Since so many psychomotor functions and psychosocial aspects are affected by a stroke, it can be taken as a general model for rehabilitation in old age. Although the medical treatment of risk factors for stroke (especially hypertension) has proven itself to be an effective prevention, stroke rehabilitation will continue to be an important part of geriatric medicine.

7.23.4.2 Arthritis Patients

The most common chronic condition in old age results from arthritis and rheumatism (US Department of Health and Human Services, 1992). Rheumatoid diseases are characterized by chronic inflammation of joints and soft tissues. Degenerative arthritis results from wearing out of joints (e.g., hip, knee) or from degeneration of bones (e.g., osteoporosis). Age at onset has an aggravating effect on disease expression in rheumatoid arthritis (Keitel, 1993). Symptoms of arthritic diseases are pain (like back pain or pain in the affected joints) and decreasing flexibility of the affected joints (e.g., hip, knee, foot, or fingers). Emotional consequences of arthritis can be increased depressive symptoms or dysphoria (Frank et al., 1992). Functional consequences of arthritic conditions are reduced mobility and losses in certain ADL functions (Badley & Tennant, 1993). Treatment of the several forms of rheumatoid and degenerative arthritis involves anti-inflammatory and analgesic drugs, although medical and physical treatment modalities are closely interrelated and should always be considered together (Baum, 1984). Hence, physical therapy and exercise programs are also extremely important. Sometimes surgery might be effective (like artificial hip joints). In some cases, dietary restrictions might be helpful (degenerative arthritis: weight loss; gout: restriction of calories and alcohol; osteoporosis: additional supply of calcium).

7.23.4.2.1 Physical therapy

In patients with degenerative joint disease, physical exercise programs should be used. Exercises to maintain muscle function are as important as exercises to strengthen specific muscle groups around the affected joint. In contrast, in elderly patients with arthritic conditions, it is necessary to control the underlying inflammatory disease before starting with exercise programs (Baum, 1984). Moreover, morning stiffness and afternoon fatigue make it necessary to schedule active therapy between the end of morning stiffness and the onset of fatigue. Exercise programs have been shown to be effective in increasing mobility, endurance, and functional status (Kovar et al., 1992). Also occupational therapy has positive effects on functional status (Helewa et al., 1991). Moreover, physiotherapy also seems to protect bone mass which may have implications for avoiding bone fractures in the long run (Shawe, Hesp, Gumpel, Sambrook, & Reeve, 1993).

7.23.4.2.2 Psychological interventions

Physical exercise programs not only improve physical health but can also improve mood and reduce anxiety and pain (Frobose & Schule, 1992). Hence, it is an important task for gerontopsychologists to strengthen the compliance of elderly arthritic patients with the therapeutic regimen and exercise programs. The perceived benefits of exercise programs can be an important determinant of exercise participation and compliance (Neuberger, Kasal, Smith, Hassanein, & DeViney, 1994). Moreover, certain sociodemographic characteristics (lower income, lower education, and higher chronological age) might lead to underuse of physical or occupational therapy provided by community services (Mayer-Oakes et al., 1992). The particular challenge for geropsychology with respect to arthritic patients lies in the chronic pain which may undermine life quality severely and permanently (Smith, Wallston, & Dwyer, 1995). The stress and coping literature in general, as well the intervention literature in particular, has not put enough emphasis on the question of what it means to age in the face of chronic and uncontrollable pain. Intervention may thus be aimed at direct alteration of the pain (e.g., by using biofeedback procedures) as well as supportive measures that enable the individual to live better with this chronic stressor (e.g., relaxation techniques, client-centered counseling). Controlled studies show that behavior-oriented interventions might influence knowledge about the disease and pain coping strategies, even if health status deteriorates because of the progressive nature of rheumatoid arthritis (Barry, Purser, Hazleman, McLean, & Hazleman, 1994; Kraaimaat, Brons, Geenen, & Bijlsma, 1995; Matussek, 1992; Simeoni, Bauman, Stenmark, & O'Brien, 1995). Since arthritic diseases affect mobility, the body self-image of an elderly person might be threatened. Clinical psychologists might help arthritic patients to cope with these losses (Jordan & Schmidt, 1988; Lehr, Kruse, & Kruse, 1992).

The musculoskeletal diseases account for the most frequent complaints among elderly persons. Since it is rarely possible to provide the patient with complete symptomatic relief, it is necessary to inform them about adequate health habits, concepts of preventive care, the value of physiotherapeutic and occupational exercises, and the gains to be achieved through behavioral and cognitive therapy.

7.23.4.3 Sensory Impaired Elderly

Losses in hearing and vision are very common in old age (Tesch-Römer & Wahl, 1996b). Hearing impairment undermines the elderly person's everyday communication, whereas losses in vision impede orientation within the physical environment. Aural rehabilitation with elderly suffering from hearing impairments includes technological devices like hearing aids and behavioral communication strategies (Kaplan, 1989; McCarthy & Sapp, 1993; Wahl & Tesch-Römer, 1996). Rehabilitation with the visually impaired elderly consists in orientation and mobility training, exercises in activities of daily living, and the use of magnifying optical devices (LaGrow & Blasch, 1992; Wahl & Oswald, 1996; Wahl & Tesch-Römer, 1996).

7.23.4.3.1 Hearing impairment

Hearing impairment is the second most prevalent chronic condition in old age; about 30% of those over 65 are so severely hearing-impaired that they would benefit from a hearing aid (Davis, 1989; US Department of Health and Human Services, 1992). Presbycusis (or hearing loss in old age) is a high-frequency symmetric hearing loss in old age beginning gradually around the fifth decade and leading progressively and irreversibly to moderate and severe losses of hearing capacity in higher ages (Corso, 1981; Fozard, 1990; Kline & Scialfa, 1996; Willott, 1991; Olsho, Harkins, & Lenhardt, 1985). Age effects for auditory and speech processing independent from hearing sensitivity have also been documented (Humes & Christopherson, 1991). Hearing loss in old age can have negative consequences in the areas of communication, social integration, well-being, and cognitive capacity (Marsiske, Delius, Lindenberger, Scherer, & Tesch-Römer, 1996; Rott, Wahl, & Tesch-Römer, 1996; Weinstein & Ventry, 1983).

There is no medical cure for presbycusis. The only way to help elderly hearing-impaired adults is through rehabilitation with hearing aids (Glass, 1986; Stephens & Goldstein, 1983). Hearing aids are small technical devices which amplify environmental sounds. Although there have been advances in hearing aid technology, hearing aids are problematic devices. Four problem areas in the use of hearing aids have been identified. Elderly hearing aid users complain of poor amplification, stigmatization, financial costs, and handling problems (e.g., inserting the hearing aid in the earmold or manipulating small levers; Franks & Beckmann, 1985; May, Upfold, & Battaglia, 1990). Nevertheless, the use of hearing aids reduces the subjective communication problems of elderly persons with hearing impairment (Tesch-Römer, 1997).

Elderly hearing-impaired patients should be supported in several ways (Tesch-Römer & Wahl, 1996a). Patients should be informed carefully about the functioning and handling of technical aids (in addition to hearing aids, a variety of listening devices exists; Kaplan, 1987). Since hearing with hearing aids is a new experience for the impaired person, they should adapt systematically to the hearing aid (e.g., use the hearing aid first in one-to-one communication and only later in group situations with background noise). Moreover, it could be useful to train the hearing impaired elderly in communication strategies and "hearing tactics," such as asking for well-articulated repetitions, occupying acoustically favorable positions, and lip reading (Vognsen, 1976).

7.23.4.3.2 Visual impairment

As is the case with respect to the elderly with hearing impairments, age-related low vision and its rehabilitation has only received due acknowledgment in the 1990s. The empirical work (Wahl, 1997; see also Wahl & Oswald, in press) has shown clearly that visual impairment has negative emotional and behavioral consequences. For example, significantly higher rates of clinically relevant depression and lower subjective well-being have been reported repeatedly (for overview, see Wahl, 1997; see also Horowitz, 1995). It should be mentioned here, however, that the negative impact on emotional adaptation does not substantially correlate with the objective loss in visual capacity. Regarding behavior, there can be no doubt that the visual loss has profound negative consequences with respect to daily living competence and independent functioning. Moreover, the visually impaired elderly are challenged by the fact that at least some of their preferred leisure activities (reading activities in particular) may become very difficult or even impossible without a good rehabilitation program. Finally, visual impairment is a prototypical environment-relevant loss in competence (Wahl, 1994). Thus, age-related loss in vision may be seen as a prototypical case for an ecological approach to geriatric rehabilitation.

The negative impact of significant (and irreversible) vision loss (e.g., caused by macular degeneration or diabetic retinopathia) makes a multidimensional rehabilitation approach necessary. The main goal of rehabilitation is the restoration of reading capability by the appropriate use of technical devices. A whole range of technical tools, from magnifying glasses to computer-based reading devices, exists (see, for example, Tesch-Römer & Wahl, 1996a). The challenge here is not the device itself but the implementation of these devices in the everyday world of the visually impaired elderly. Moreover, good training with a high degree of "ecological validity" is also needed. A second rehabilitation issue is training in orientation and mobility. The challenge with respect to these rehabilitation techniques lies in the fact that elderly persons often have concomitant impairments of a chronic nature, for example, coexisting hearing, motor, or cognitive impairments. A third issue is training in the domain of daily living tasks. In all of these rehabilitation efforts, training must take place within the elderly person's everyday world. Interventions in the environment of visually impaired elderly are frequently supportive and helpful. With respect to the social environment, it is very important to educate social partners to take the disabilities of their relatives and friends seriously while resisting the temptation to be overprotective. With respect to the physical environment, a whole range of alterations has been suggested (e.g., Hiatt, 1987) such as the enhancement of light conditions (without producing too much glare), the use of contrast and color, and the employment of technical innovations such as special telephones or talking appliances for the kitchen.

To conclude this section, the rehabilitative needs of sensory impaired persons have only been articulated in the 1990s. The aim of this section was thus not only to elaborate on the special needs of these groups of elderly (those with impairments in vision and hearing), but also to emphasize that the rehabilitation of these elderly must consist of more than the mere adoption of technical devices. Truly adequate interventions must implement the devices properly and facilitate the elderly person's adaptation to the new technology. And of course, rehabilitation must enable the individual to better cope with the psychosocial losses caused by the sensory impairment (e.g., Tesch-Römer & Nowak, 1995; Wahl & Oswald, 1996).

7.23.4.4 Demented elderly

Dementia is a very serious problem in the elderly, especially in the very old (see Chapter 10, this volume). In the most common form of dementia, Alzheimer's disease, aphasia, agnosia, and apraxia are present in various combinations as well as emotional and personality changes, memory disorders, slowness in rate of information processing, and the incapacity to fulfill daily living activities. It would therefore be reasonable to question the merit of rehabilitation efforts with the demented elderly. However, professionals as well as society as a whole, cannot neglect the burden and challenge of the tremendous and fast growing number of elderly suffering from dementia. Custodial care is obviously not the right answer. The responsibility for this problem should be shared by all and should address the challenge with those rehabilitative tools which promise at least some measure of success. In this context, it may not be wise to talk about "dementia" in general. There are different degrees of severity, and each patient exhibits a unique profiles of competencies; both of these factors crucially affect rehabilitation outcome.

Rehabilitation efforts undoubtedly should include reality orientation programs and training to cope, at least in part, with daily living challenges. Data suggest that rehabilitative work of this kind may be particularly promising when it starts in the early stages of the cognitive decline process. Intervention studies hint that further deterioration can be delayed by early cognitive training intervention (e.g., Oswald et al., 1988). Family members and other caregivers should be part of rehabilitative efforts, too (Pearlin et al., 1996; see also Chapter 22, this volume). Thus, the comprehensive care of the caregivers, as suggested in the literature on caregiver burden, may also be regarded as a very important part of the geriatric rehabilitation system. In a sense, it also serves important preventive purposes. As is known, being a caregiver for a demented individual heightens the risk of suffering from chronic physical or mental conditions in the long run. Finally, the treatment of dementia provides a prototypical example of how, given unchangeable person characteristics, the possibility of altering the physical environment always remains. This is particularly true in institutional settings, but also relevant in private home living. Some persuasive and empirically proven solutions for modifying and adjusting the environment have been suggested in the literature (e.g., Regnier & Pynoos, 1992).

7.23.5 PRACTICAL ISSUES, RESEARCH QUESTIONS, AND ETHICAL PROBLEMS

The last section of this chapter is concerned with some practical, research-oriented, and ethical problems associated with rehabilitation in old age. Practical problems concern the institutional organization of geriatric rehabilitation, the qualification of staff, and the assurance of rehabilitation quality. Research questions concern the effectiveness of interventions and the development and evaluation of innovative rehabilitation concepts. Finally, ethical problems deal for instance with the distribution of scarce resources within the health care system.

7.23.5.1 Practical Issues

Regarding the institutional organization of geriatric rehabilitation, it has been shown that specialized geriatric units or geriatric hospitals are more effective and more cost efficient than general medical wards (Görres, 1992b). An important problem for elderly patients after rehabilitation is returning home and maintaining independent living in the community. It could be helpful to view this problem from the perspective that individuals with disabilities are handicapped more by barriers in the environments than by their specific impairments (DeJong, Batavia, & McKnew, 1992) and that rehabilitative interventions should be realized in conjunction with community-based long-term care (Miller, 1991; see also Chapter 20, this volume). Another important practical problem is quality assurance (Görres, 1994; Meier-Baumgartner et al., 1992). Conceptions of rehabilitation programs might be very different from their practical implementation. Hence, it is necessary to control the quality of rehabilitative interventions. Three aspects of quality in medical care and rehabilitation can be distinguished (Donabedian, 1966): quality of structure, process quality, and outcome quality. Quality of structure concerns characteristics of the institution, the educational profile of staff, and quality of technical equipment. Process quality concerns characteristics of the rehabilitation process (diagnostics, development of rehabilitation goals, implementation of therapy and interventions, documentation of changes, and team communication). Outcome quality concerns the evaluation of changes stemming from rehabilitative efforts (reduced morbidity and mortality, increased functional independence, ability and motivation for independent living, patient satisfaction). Quality of rehabilitation structure and process are necessary conditions for outcome quality; but outcome quality certainly is what counts in evaluating rehabilitation services (Görres, 1992b).

7.23.5.2 Research Questions

For policy planning and program development, it seems necessary to gather epidemiological data on chronic diseases in old age. The lack of adequate epidemiological data might lead to underestimating the need for rehabilitation efforts. A second major research problem concerns the effectiveness of rehabilitation programs (Görres, Tschubar, & Meier-Baumgartner, 1991). Rehabilitation programs should be tested empirically regarding their effectiveness and efficiency by considering a psychosocial view of life quality as well as a cost-benefit ratio. Although it may be ethically questionable to evaluate rehabilitation programs involving experimental and control groups in pre/post-treatment designs, nevertheless controlled studies should be the standard procedure in rehabilitation research (Wottawa & Thierau, 1990). Finally, the development and evaluation of innovative rehabilitation programs is an important research task. It should be emphasized again that not only dramatic diseases (like stroke) should be the focus of future gerontological rehabilitation research. There are a variety of age-related diseases which urgently require rehabilitation (e.g., arthritis, vision problems and blindness, hearing loss).

7.23.5.3 Ethical Problems

There are several ethical issues concerning rehabilitation in old age (e.g., Spicker, Ingman, & Lawson, 1987). Thinking about rehabilitation implies reference to norms (What is the goal of rehabilitation?), making a decision between rehabilitation and care (When does rehabilitation promise progress in the individual?), and weighing individual and social benefits and costs (What are the gains of rehabilitation compared to the costs?). Considering rehabilitation goals, normative standards (e.g., independent living) could come in conflict with individual values (e.g., security and social contact). Hence, in developing rehabilitation goals, individual values should be considered carefully. Second, before an individual rehabilitation process is initiated, a decision has to be made regarding the rehabilitation potential of the particular individual; sometimes this decision concerns the alternatives rehabilitation vs. care. But even in geriatric, long-term care facilities, nursing does not necessarily have to fulfill a primarily custodial function; rather, it could also be performed according to a therapeutic model of rehabilitation (Burgio & Burgio, 1990). However, sometimes the general goal of rehabilitation has to be given up, and a decision to commit to long-term care be made. This decision might also involve considerations of costs and benefits (Brody, 1987; Hausman, 1987). There are individual costs connected with rehabilitation (e.g., time and effort), but these costs might be outweighed by the potential benefits of rehabilitation outcomes. Additionally, there are also material costs (e.g., financial support for rehabilitation programs). It should be emphasized, though, that rehabilitation is based on humanistic values. This basic ethical orientation stresses that the disabled elderly deserve, as much as anyone else, the opportunity to participate in all aspects of life (Kemp, 1985). Beyond humanistic arguments, however, one might add that current economic conditions make rehabilitation more efficient in the long run than long-term care.

Rehabilitation in old age is a promising field. Empirical research demonstrates the effectiveness of psychosocial interventions in a variety of chronic diseases. Moreover, rehabilitation in old age should not necessarily be dominated by the medical profession. Instead, clinical psychology should demonstrate its rich reservoir of intervention and therapeutic strategies, and clinical psychologists might more often lay hold of the leadership in interdisciplinary rehabilitation teams. Despite unsolved practical issues, open research questions, and difficult ethical problems, rehabilitation in old age is a gratifying field of expertise for clinical psychologists.

7.23.6 SUMMARY

In this chapter, general principles and geropsychological essentials of rehabilitation in old age are reviewed from a geropsychological perspective. Rehabilitation aims at reducing the disability and handicap experienced as a result of disease and impairment. Since aging is associated with increased chronic illness, the goal of rehabilitation in old age is not necessarily complete cure of disease but improving functional capability and maximizing wellness. An overview concerning psychosocial issues and strategies of rehabilitation in old is given, including comprehensive geriatric and psychosocial assessment, restoration of functional independence, cognitive and reality orientation training, therapeutic intervention, motivational and compliance problems, the role of social support, and the ecology of aging. Since different geriatric client groups have different rehabilitation needs, four highly prevalent and serious chronic conditions in old age are treated specifically, that is, stroke, arthritis, sensory impairment, and dementia. The chapter closes with perspectives for psychologists working in rehabilitative settings discussing practical issues, research questions, and ethical considerations.

7.23.7 REFERENCES

Abrams, W. B., Berkow, R., & Fletcher, A. J. (Eds.) (1990). *The Merck Manual of Geriatrics*. Rahway, NJ: Merck, Sharp, & Dohme Research Laboratories.

Abson, V., & Rabbitt, P. M. (1988). What do self-rating questionnaires tell us about changes in competence in old age? In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (pp. 188–191). Chichester, UK: Wiley.

Anderson, C. S., Linto, J., & Stewart-Wynne, E. G. (1995). A population-based assessment of the impact and burden of caregiving for long-term stroke survivors. *Stroke*, *26*, 843–849.

Andrews, K. (1987). *Rehabilitation of the older adult*. London: Edward Arnold.

Andrews, K. (1991). Rehabilitation of the elderly. In M. S. J. Pathy (Ed.), *Principles and practice of geriatric medicine* (pp. 1341–1365). Chichester, UK: Wiley.

Antonucci, T. C. (1985). Personal characteristics, social support, and social behavior. In R. H. Binstock & E. Shanas (Eds.), *Handbook of aging and the social sciences* (2nd ed., pp. 94–128). New York: Van Nostrand Reinhold.

Antonucci, T. C., & Jackson, J. S. (1987). Social support, interpersonal efficacy, and health: A life course perspective. In L. L. Carstensen & B. A. Edelstein (Eds.), *Handbook of clinical gerontology* (pp. 291–311). New York: Pergamon.

Bach, D., Bach, M., Böhmer, F., Frühwald, T., & Grilc, B. (1995). Reactivating occupational therapy: a method to improve cognitive performance in geriatric patients. *Age and Ageing*, *24*, 222–226.

Bäckman, L., Josephsson, S., Herlitz, A., Stigsdotter, A., & Viitanen, M. (1991). The generalizability of training gains in dementia: Effects of an imagery-based mnemonic on face-name retention duration. *Psychology and Aging*, *6*, 489–492.

Badley, E. M., & Tennant, A. (1993). Disablement associated with rheumatic disorders in a British population: problems with activities of daily living and level of support. *British Journal of Rheumatology*, *32*, 601–608.

Baltes, M. M. (1996). *The many faces of dependency in old age*. Cambridge, UK: Cambridge University Press.

Baltes, M. M., Burgess, R. L., & Stewart, R. B. (1980). Independence and dependence in self-care behaviors in nursing home residents: An operant-observational study. *International Journal of Behavioral Development*, *3*, 489–500.

Baltes, M. M., Kindermann, T., & Reisenzein, R. (1986). Die Beobachtung von unselbständigem und selbständigem Verhalten in einem deutschen Altenheim: Die soziale Umwelt als Einflußgröße [The observation of dependent and independent behavior in a German old age home: the social environment as an influencing factor]. *Zeitschrift für Gerontologie*, *19*, 14–21.

Baltes, M. M., Kühl, K. P., Gutzmann, H., & Sowarka, D. (1995). Potential of cognitive plasticity as a diagnostic instrument: a cross-validation and extension. *Psychology and Aging*, *10*, 167–172.

Baltes, M. M., Maas, I., Wilms, H.-U., & Borchelt, M. (1996). Alltagskompetenz im Alter: Theoretische Überlegungen und empirische Befunde [Everyday competence in old age: Theoretical issues and empirical findings]. In K. U. Mayer & P. B. Baltes (Eds.), *Die Berliner Altersstudie* [The Berlin Aging Study] (pp. 525–542). Berlin, Germany: Akademie Verlag.

Baltes, M. M., & Sowarka, D. (1995). Kognitive Intervention im Alter [Cognitive intervention in old age]. In R. Oerter & L. Montada (Eds.), *Entwicklungspsychologie* [Developmental psychology] (pp. 1116–1126). Weinheim, Germany: Psychologie Verlags Union

Baltes, M. M., & Wahl, H.-W. (1992). The dependency-support script in institutions: Generalization to community settings. *Psychology and Aging*, *7*, 409–418.

Baltes, M. M., & Wahl, H.-W. (1996). Patterns of communication in old age: The Dependency-Support and Independence-Ignore Script. *Health Communication*, *8*, 217–231.

Baltes, M. M., Wahl, H.-W., & Reichert, M. (1991). Successful aging in long-term care institutions. *Annual Review of Gerontology and Geriatrics*, *11*, 311–338.

Baltes, P. B. (1984). Intelligenz im Alter [Intelligence in old age]. *Spektrum der Wissenschaft*, Mai 1984, 46–60.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology*, *23*, 611–626.

Baltes, P. B. (1993). The aging mind: Potential and limits. *The Gerontologist*, *33*, 580–594.

Baltes, P. B., & Baltes, M. M. (1990). Psychological perspectives on successful aging: the model of selective optimization with compensation. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging. Perspectives from the behavioral sciences* (pp. 1–34). Cambridge, UK: Cambridge University Press.

Baltes, P. B. & Lindenberger, U. (1988). On the range of cognitive plasticity in old age as a function of experience: 15 years of intervention research. *Behavior Therapy*, *19*, 283–300.

Baltes, P. B., Sowarka, D., & Kliegl, R. (1989). Cognitive training research on fluid intelligence in old age: What can older adults achieve by themselves? *Psychology and Aging*, *4*, 217–221.

Barker, R. G. (1968). *Ecological psychology*. Stanford, CA: Stanford University Press.

Barry, M. A., Purser, J., Hazleman, R., McLean, A., & Hazleman, B. L. (1994). Effect of energy conservation and joint protection education in rheumatoid arthritis. *British Journal of Rheumatology*, *33*, 1171–1174.

Baum, J. (1984). Rehabilitation aspects of arthritis in the elderly. In T. F. Williams (Ed.), *Rehabilitation in the aging* (pp. 177–197). New York: Raven Press.

Blieszner, R., Willis, S. L., & Baltes, P. B. (1981). Training research in aging on the fluid ability of inductive reasoning. *Journal of Applied Developmental Psychology*, *2*, 247–265.

Bobath, B. (1971). *Adult hemiplegia: evaluation and treatment*. London: Heinemann.

Bouma, H., & Graafmans, J. A. M. (Eds.) (1992). *Gerontechnology*. Amsterdam: IOS Press.

Brocklehurst, J. C., Tallis, R. C., & Fillit, H. M. (Eds.). (1992). *Textbook of geriatric medicine and gerontology* (4th ed.). Edinburgh, UK: Churchill Livingstone.

Brody, B. (1987). Wholehearted and halfhearted care: national policies vs. individual choice. S. F. Spicker, S. R. Ingman, & I. R. Lawson (Eds.), *Ethical dimensions of geriatric care* (pp. 79–93). Dordrecht, The Netherlands: Reidel.

Burgio, L. D., & Burgio, K. L. (1990). Institutional staff training and management: A review of the literature and a model for geriatric, long-term-care facilities. *International Journal of Aging and Human Development*, *30*, 287–302.

Butler, R. W., & Namerow, N. S. (1988). Cognitive retraining in brain-injury rehabilitation: A critical review. *Journal of Neurologic Rehabilitation*, *2*, 97–101.

Cameron, I. D., & Quine, S. (1994). External hip protectors: Likely non-compliance among high risk elderly people living in the community. *Archives of Gerontology and Geriatrics*, *19*, 273–281.

Camp, C. J., Foss, J. W., Stevens, A. B., Reichard, C. C., McKitrick, L. A., & O'Hanlon, A. M. (1993). Memory training in normal and demented elderly populations: the E-I-E-I-O model. *Experimental Aging Research*, *19*, 277–290.

Campbell, A. J. (1992). Role of rehabilitation in fall recovery and prevention. *Reviews in Clinical Gerontology*, *2*, 53–65.

Carstensen, L. L., Edelstein, B. A. , & Dornbrand, L. (Eds.) (1996). *The practical handbook of clinical gerontology*. Thousand Oaks, CA: Sage.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.

Charness, N., & Bosman, E. A. (1990). Human factors and design for older adults. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 446–463). San Diego, CA: Academic Press.

Cooper, B. (1991a). The epidemiology of dementia. In R. Jacoby & C. Oppenheimer (Eds.), *Psychiatry in the elderly* (pp. 574–585). Oxford, UK: Oxford University Press.

Cooper, B. (1991b). Principles of service provision in old age psychiatry. In R. Jacoby & C. Oppenheimer (Eds.), *Psychiatry in the elderly* (pp. 274–300). Oxford, UK: Oxford University Press.

Corso, J. F. (1981). *Aging sensory systems and perception*. New York: Praeger.

Crimmins, E. M., Hayward, M. D., & Saito, Y (1996). Differentials in active life expectancy in the older population of the United States. *Journal of Gerontology: Social Sciences*, *51B*, S111–S120.

Dam, M., Tonin, P., Casson, S., Ermani, M., Pizzolato, G., Iaia, V., & Battistin, L. (1993). The effects of long-term rehabilitation therapy on poststroke hemiplegic patients. *Stroke*, *24*, 1186–1191.

Davis, A. C. (1989). The prevalence of hearing impairment and reported hearing disability among adults in Great Britain. *International Journal of Epidemiology, 18,* 911–917.

DeJong, G., Batavia, A. I., & McKnew, L. B. (1992). The independent living model of personal assistance in national long-term-care policy. *Generations, 16,* 89–95.

Dixon, R. A., & Bäckman, L. (Eds.) (1995).*Compensating for psychological deficits and declines. Managing losses and promoting gains.* Mahwah, NJ: Erlbaum.

Donabedian, A. (1966). Evaluating the quality of medical care. *Millbank Memorial Fund Quarterly*, *44*, 166–206.

Ebrahim, S. (1992). Rehabilitation. In J. C. Brocklehurst, R. C. Tallis, & H. M. Fillit (Eds.), *Textbook of geriatric medicine and gerontology* (4th ed., pp. 1038–1054). Edinburgh, UK: Churchill Livingstone.

El-Faizy, M., & Reinsch, S. (1994). Home safety intervention for the prevention of falls. *Physical and Occupational Therapy in Geriatrics*, *12*, 33–49.

Ellul, J., Watkins, C., Ferguson, N., & Barer, D. (1993). Increasing patient engagement in rehabilitation activities. *Clinical Rehabilitation*, *7*, 297–302.

Ermini-Fünfschilling, D., & Meier, D. (1995). Gedächtnistraining: Wichtiger Bestandteil der Milieutherapie bei seniler Demenz [Memory training: An important constituent of milieu therapy in senile dementia]. *Zeitschrift für Gerontologie und Geriatrie*, *28*, 190–194.

Evans, J. G., & Williams, T. F. (Eds.). (1992). *Oxford textbook of geriatric medicine*. Oxford, UK: Oxford University Press.

Falconer, J. A., Naughton, B. J., Strasser, D. C., & Sinacore, J. M. (1994). Stroke inpatient rehabilitation: A comparison across age groups. *Journal of the American Geriatrics Society*, *42*, 3–44.

Felsenthal, G., & Stein, B. D. (1990). Rehabilitation of dysmobility in the elderly: A case study of the patient with a hip fracture. In S. J. Brody & L. G. Pawlson (Eds.), *Aging and rehabilitation II: the state of the practice* (pp. 110–129). New York: Springer.

Fleischmann, U. M. (1982). Gedächtnistraining im höheren Lebensalter—Ansatzpunkte und Möglichkeiten [Memory training in old age—Foci and possibilities]. *Zeitschrift fr Gerontologie*, *15*, 53–62.

Fleischmann, U. M. (1989). *Gedächtnis und Alter. Multivariate Analysen zum Gedächtnis alter Menschen* [Memory and old age. Multivariate analyses of the memory of the elderly]. Bern, Switzerland: Huber.

Fonda, D. (1991). Management of the incontinent disabled patient. *Clinical Rehabilitation*, *5*, 331–336.

Fozard, J. L. (1990). Vision and hearing in aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 150–170). San Diego, CA: Academic Press.

Frank, R. G., Chaney, J. M., Clay, D. L., Shutty, M. S., Beck, N. C., Kay, D. R., Elliott, T. R., & Grambling, S. (1992). Dysphoria: a major symptom factor in persons with disability or chronic illness. *Psychiatry Research*, *43*, 231–241.

Franks, J. R., & Beckmann, N. J. (1985). Rejection of hearing aids: Attitudes of a geriatric sample. *Ear and Hearing*, *6*, 161–166.

Franzen, M. D. & Martin, R. C. (1996). Screening for neuropsychological impairment. In L. L. Carstensen, B. A. Edelstein, & L. Dornbrand (Eds.), *The practical handbook of clinical gerontology* (pp. 188–216). Thousand Oaks, CA: Sage.

Fries, J. F. (1990). Medical perspectives upon successful aging. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging. Perspectives from the behavioral sciences* (pp. 35-49). Cambridge, UK: Cambridge University Press.

Frobose, I., & Schule, K. (1992). Befindlichkeits- und Einstellungsänderungen beim Sport mit Rheumatikern [Mood and attitude change as a consequence of sports with rheumatic patients]. *Die Rehabilitation*, *31*, 198–201.

Gatz, M., Kasl-Godley, J. E., & Karel, M. J. (1996). Aging and mental disorders. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 365–382). San Diego, CA: Academic Press

Gibson, C. J. & Caplan, B. M. (1984). Rehabilitation of the patient with stroke. In T. F. Williams (Ed.), *Rehabilitation in the aging* (pp. 145–159). New York: Raven Press.

Gladman, J., Forster, A., & Young, J. (1995). Hospital- and home-based rehabilitation after discharge from hospital for stroke patients: analysis of two trials. *Age and Ageing*, *24*, 49–53.

Glass, L. E. (1986). Rehabilitation for deaf and hearing-impaired elderly. In S. J. Brody & G. E. Ruff (Eds.), *Aging and rehabilitation* (pp. 218–237). New York: Springer.

Glass, T. A., & Maddox, G. L. (1992). The quality and quantity of social support: Stroke recovery as psychosocial transition. *Social Science and Medicine*, *34*, 1249–1261.

Goldberg, R. (1984). Psychosocial aspects of stroke. In D. W. Krueger (Ed.), *Rehabilitation psychology* (pp. 293–300). Rockville, MD: Aspen.

Görres, S. (1992a). Stationäre Rehabilitation und Nachsorge älterer Menschen [Hospital rehabilitation and post-hospital care of the elderly]. In A. Mühlum & H. Oppl (Eds.). *Handbuch der Rehabilitation* [Handbook of rehabilitation] (pp. 401-423). Neuwied, Germany: Luchterhand.

Görres, S. (1992b). Qualitätssicherung in der Geriatrie [Quality assurance in geriatrics]. *Archiv für Wissenschaft und Praxis der Sozialen Arbeit*, *23*, 90–111.

Görres, S. (1994). Qualitätssicherung in der Rehabilitation älterer Menschen [Quality assurance in the rehabilitation of the elderly]. *Die Rehabilitation*, *33*, 64–68.

Görres, S., Tschubar, H., & Meier-Baumgartner, H.-P. (1991). Zur Bedeutung klinikinterner Erfolgskriterien in der geriatrischen Rehabilitation. Ergebnisse eines Nachsorgeprojekts zum Problem der Qualitätssicherung [On the importance of outcome criteria in geriatric rehabilitation: Findings from a post-hospital care project on the problem of quality assurance]. *Die Rehabilitation*, *30*, 63–68.

Guthrie, S., & Harvey, A. (1994). Motivation and its influence on outcome in rehabilitation. *Reviews in Clinical Gerontology*, *4*, 235–243.

Harris, R. E., O'Hara, P. A., & Harper, D. W. (1995). Functional status of geriatric rehabilitation patients: A one-year follow-up study. *Journal of the American Geriatrics Society*, *43*, 51–55.

Hartke, R. J. (1991). The older adult's adjustment to the rehabilitation setting. In R. J. Hartke (Ed.), *Psychological aspects of geriatric rehabilitation* (pp. 73-96). Gaithersburg, MD: Aspen.

Hartke, R. J., & Horowitz, M. (1991). Psychotherapy with the older rehabilitation patient. In R. J. Hartke (Ed.). *Psychological aspects of geriatric rehabilitation* (pp. 229–248). Gaithersburg, MD: Aspen.

Hasselkus, B. R. (1989). Occupational and physical therapy in geriatric rehabilitation. *Physical and Occupational Therapy in Geriatrics, 7,* 3–20.

Hausman, D. M. (1987). Health care: efficiency and equity. In S. F. Spicker, S.R. Ingman, & I.R. Lawson (Eds.), *Ethical dimensions of geriatric care* (pp. 67–78). Dordrecht The Netherlands: Reidel.

Häfner, H. (1986). *Psychische Gesundheit im Alter* [Mental health in old age]. Stuttgart, Germany: Fischer.

Helewa, A., Goldsmith, C. H., Lee, P., Bombardier, C., Hanes, B., Smythe, H. A., & Tugwell, P. (1991). Effects of occupational therapy home service on patients with rheumatoid arthritis. *Lancet*, *337*, 1453–1456.

Hiatt, L. G. (1987). Designing for the vision and hearing impairments of the elderly. In V. Regnier & J. Pynoos (Eds.), *Housing the aged* (pp. 341–371). New York: Elsevier.

Horgas, A., Wahl, H.-W., & Baltes, M. M. (1996). Dependency in late life. In L. L. Carstensen, B. A. Edelstein, & L. Dornbrand (Eds.), *The practical handbook of clinical gerontology*. Newbury Park, NJ: Sage.

Horn, J. L., & Hofer, S. M. (1992). Major abilities and development in the adult period. In R. J. Sternberg & C. A. Berg (Eds.), *Intellectual development* (pp. 44–99). New York: Cambridge University Press.

Hui, E., Lum, C. M., Woo, J., Or, K. H., & Kay, R. L. (1995). Outcomes of elderly stroke patients. Day hospital versus conventional medical management. *Stroke*, *26*, 1616–1619.

Humes, L. E., & Christopherson, L. (1991). Speech identification difficulties of hearing-impaired elderly persons: The contributions of auditory processing deficits. *Journal of Speech and Hearing Research*, *34*, 686–693.

Hussian, R. A. & Davis, R. L. (1985). *Responsive care*. Champain, IL: Research Press.

Jamison, R. N., & Virts, K. L. (1990). The influence of family support on chronic pain. *Behaviour Research and Therapy*, *28*, 283-287.

Jette, A. M. (1996). Disability trends and transitions**.** In R. H. Binstock & L. K. George (Eds.), *Handbook of aging and the social sciences* (4th ed., pp. 94–116). San Diego, CA: Academic Press.

Jordan, F. M., Worral, L.E., Hickson, L. M., & Dodd, B. J. (1993). The evaluation of intervention programmes for communicatively impaired elderly people. *European Journal of Communication Disorders*, *28*, 63–85.

Jordan, J., & Schmidt, M. (1988). Rehabilitationspsychologische Aspekte bei Erkrankungen des rheumatischen Formenkreises [Psychological aspects of rheumatic diseases relevant for rehabilitation]. In U. Koch, G. Lucius-Koene, & R. Stegie (Eds.), *Handbuch der Rehabilitationspsychologie* [Handbook of rehabilitation psychology] (pp. 455–478). Berlin, Germany: Springer.

Jørgensen, H. S., Nakayama, H., Raaschou, H. O., Larsen, K., Hubbe, P., & Olsen, T. S. (1995). The effect of a stroke unit: reductions in mortality, discharge rate to nursing home, length of hospital stay, and cost. A community-based study. *Stroke*, *26*, 1178–1182.

Kalra, L. (1994). The influence of stroke unit rehabilitation on functional recovery from stroke. *Stroke*, *25*, 821–825.

Kaplan, H. (1987). Assistive communication devices. In H. G. Mueller & V. C. Geoffrey (Eds.), *Communication disorders in aging: Assessment and management* (pp. 464–496). Washington, DC: Gallaudet University Press.

Kaplan, H. (1989). Communication problems of the hearing-impaired elderly: What can be done? *Pride Institute Journal of Long Term Health Care*, *7*, 10–22.

Kaschel, R., Zaiser-Kaschel, H., & Mayer, K. (1992). Realitäts-Orientierungs-Training: Literaturüberblick und Implikationen für die neuropsychologische Gedächtnisrehabilitation [Reality orientation training: Literature review and implications for neuropsychological memory rehabilitation]. *Zeitschrift für Gerontopsychologie und -psychiatrie, 5,* 223–235.

Kaste, M., Palomaki, H., & Sarna, S. (1995). Where and how should elderly stroke patients be treated? A randomized trial. *Stroke*, *26*, 249–253.

Kausler, D. H. (1991). *Experimental psychology, cognition, and human aging* (2nd ed.). New York: Springer.

Keitel, W. (1993). Verlauf der Rheumatoid-Arthritis bei Beginn im hheren Lebensalter [The course of rheumatoid arthritis beginning in advanced age]. *Zeitschrift für Rheumatologie*, *52*, 138–141.

Kemp, B. (1985). Rehabilitation and the older adult. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 647–663). New York: Van Nostrand Reinhold.

Kemp, B. (1996). Rehabilitation. In J. E. Birren (Ed.), *Encyclopedia of Gerontology* (Vol 2, pp. 381–386). San Diego, CA: Academic Press.

King, P., & Barrowclough, C. (1989). Rating the motivation of elderly patients on a rehabilitation ward. *Clinical Rehabilitation*, *3*, 289–291.

Kline, D. W., & Scialfa, C. T. (1996). Visual and auditory aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 181–203). San Diego, CA: Academic Press.

Koh, K., Ray, R., Lee, J., Nair, A., Ho, T., & Ang, P.C. (1994). Dementia in elderly patients: can the 3R mental stimulation programme improve mental status? *Age and Ageing*, *23*, 195–199.

Kovar, P. A., Allegrante, J. P., MacKenzie, C. R., Peterson, M. G., Gutin, B., & Charlson, M. E. (1992). Supervised fitness walking in patients with osteoarthritis of the knee. A randomized, controlled trial. *Annals of Internal Medicine*, *116*, 529–534.

Kraaimaat, F. W., Brons, M. R. Geenen, R., & Bijlsma, J. W. (1995). The effect of cognitive behavior therapy in patients with rheumatoid arthritis. *Behavior Research and Therapy*, *33*, 487–495.

Kruse, A. (1984). Der Schlaganfallpatient und seine Familie [The stroke patient and her/his family]. *Zeitschrift für Gerontologie*, *17*, 359–366.

Kruse, A. (1987). Kompetenz bei chronischer Krankheit im Alter [Competence in chronic disease in the aged]. *Zeitschrift für Gerontologie*, *20*, 355–366.

Kruse, A. (1992). Rehabilitation in der Gerontologie - theoretische Gundlagen und empirische Forschungsergebnisse [Rehabilitation in gerontology - theoretical base and empirical findings]. In A. Mühlum & H. Oppl (Eds.). *Handbuch der Rehabilitation* [Handbook of rehabilitation] (pp. 333–355). Neuwied, Germany: Luchterhand.

Kruse, A. (1995). Ergebnisse aus einer Längsschnittstudie zu Effekten der ambulanten Rehabilitation [Findings of a longitudinal study aimed at the outcome of home-based rehabilitation]. *Zeitschrift für Gerontologie und Geriatrie*, *28*, 104–112.

Lachman, M. E., Weaver, S. L., Bandura, M., Elliott, E., & Lewkowicz., C. J. (1992). Improving memory and control beliefs through cognitive restructuring and self-generated strategies. *Journals of Gerontology: Psychological Sciences*, *47*, P293–P299.

LaGrow, S. J., & Blasch, B. B. (1992). Orientation and mobility services for older persons. In A. Orr (Ed.), *Vision and aging: Crossroads for service delivery* (pp. 255–288). New York: American Foundation for the Blind.

Langhorne, P., Williams, B. O., Gilchrist, W., & Howie, K. (1993). Do stroke units save lives? *Lancet*, *342*, 395–398.

Lawton, M. P. (1982). Competence, environmental press, and the adaption of older people. In M. P. Lawton, P. G. Windley, & T. O. Byerts (Eds.), *Aging and the environment. Theoretical approaches* (pp. 33–59). New York: Springer.

Lawton, M. P. (1989). Behavior-relevant ecological factors. In K.W. Schaie & C. Schooler (Eds.), *Social structure and aging: Psychological processes* (pp. 57–78). Hillsdale, NJ: Erlbaum.

Lawton, M. P. & Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *The Gerontologist*, *9*, 179–186.

Lawton, M. P., & Nahemow, L. (1973). Ecology and the aging process. In C. Eisdorfer & M. P. Lawton (Eds.), *Psychology of adult development and aging* (pp. 619–674). Washington, DC: American Psychological Association.

Lawton, M. P., & Simon, B. (1968). The ecology of social relationships in housing for the elderly. *The Gerontologist*, *8*, 108–115.

Lawton, M. P., & Teresi, J. A. (Eds.). (1994). *Annual Review of Gerontology and Geriatrics. Focus on assessment techniques*. New York: Springer.

Lee, M. H. M., & Itoh, M. (1990). Rehabilitation. In W. B. Abrams, R. Berkow, & A. J. Fletcher (Eds.), *The Merck Manual of Geriatrics* (pp. 249–273). Rahway, NJ: Merck, Sharp, & Dohme Research Laboratories.

Lehmann, W. (1992). Diagnose und Therapie von geriatrischen Aphasiepatienten nach einem zerebrovaskulären Insult [Diagnosis and therapy of geriatric aphasic patients following cerobro-vascular stroke]. *Zeitschrift für Gerontologie*, *25*, 271–277.

Lehr, U. (1993). Erkenntnisse psychologischer Grundlagenforschung für Prävention und Rehabilitation im Alter [Results of psychological basic research relevant for prevention and rehabilitation in old age]. In A. Niederfranke (Ed.), *Fragen geriatrischer Rehabilitation* [Questions of geriatric rehabilitation]. (pp. 39–46). Stuttgart, Germany: Kohlhammer.

Lehr, U. (1996). Psychologie des Alterns [Psychology of aging] (8th ed.). Heidelberg, Germany: Quelle & Meyer.

Lehr, U., Kruse, A., & Kruse, M. (1992). Rheumatische Erkrankungen im höheren Lebensalter - Psychologische Aspekte [Rheumatic diseases in advanced age - psychological aspects]. *Zeitschrift für Rheumatologie*, *51*, 290–294.

Lichtenberg, P. A., & Rosenthal, M. (1994). Characteristics of geriatric rehabilitation programs: A survey of practitioners. *Rehabilitation Psychology*, *39*, 277–281.

Little, A. (1991). Psychological treatments. In R. Jacoby & C. Oppenheimer (Eds.), *Psychiatry in the elderly* (pp. 400–415). Oxford, UK: Oxford University Press.

Mahoney, F. I., & Barthel, D. W. (1965). Functional evaluation: The Barthel Index. *Maryland Medical Journal*, *14*, 61–65.

Marmann, C. (1994). Der geriatrisch Amputierte nach Entlassung aus erfolgreicher Rehabilitation [The geriatric amputee after discharge from successful rehabilitation]. *Die Rehabilitation*, *33*, 69–75.

Marsiske, M., Delius, J., Lindenberger, U., Scherer, H., & Tesch-Römer, C. (1996). Sensorische Systeme im Alter [Sensory systems in old age]. In K. U. Mayer & P. B. Baltes (Eds.). *Die Berliner Altersstudie* [The Berlin Aging Study] (pp. 379–403). Berlin, Germany: Akademie Verlag.

Mattussek, S. (1992). Die "cP-Schule Hannover": Entwicklung eines Curriculums und Durchführung einer kontrollierten Studie zur Schulung von Patienten mit einer chronischen Polyarthritis [The "Hannover Chronic Polyarthritis School": Development of a curriculum and conductance of a controlled study of education of patients with chronic arthritis]. *Zeitschrift für Rheumatologie*, *51*, Suppl. 1: 41–50.

May, A. E., Upfold, L. J., & Battaglia, J. A. (1990). The advantages and disadvantages of ITC, ITE and BTE hearing aids: Diary and interview reports from elderly users. *British Journal of Audiology*, *24*, 301–309.

Mayer-Oakes, S. A., Hoenig, H., Atchison, K. A., Lubben, J.E., De-Jong, F., & Schweitzer, S. O. (1992). Patient-related predictors of rehabilitation use for community-dwelling older Americans. *Journal of the American Geriatrics Society*, *40*, 336–342.

McCarthy, P. A., & Sapp, J. V. (1993). Rehabilitative considerations with the geriatric population. In J. G. Alpiner & P. A. McCarthy (Eds.), *Rehabilitative audiology: Children and adults* (2nd ed., pp. 331–373). Baltimore: Williams & Wilkins.

Meier-Baumgartner, H.-P., Nerenheim-Duscha, I., & Görres, S. (1992). *Die Effektivität von Rehabilitation* bei älteren Menschen unter besonderer Bercksichtigung psychosozialer Komponenten bei ambulanter, teilstationärer und stationärer Betreuung [Outcome of rehabilitation with the elderly]. Stuttgart, Germany: Kohlhammer.

Miller, J. A. (1991). *Community-based long-term care:*

Innovative models. Newbury Park, NJ: Sage.

Mosqueda, L. (1996). Geriatric rehabilitation. In L. L. Carstensen, B. A. Edelstein, & L. Dornbrand (Eds.), *The practical handbook of clinical gerontology* (pp. 468–480). Thousand Oaks, CA: Sage.

Nelson, E. A., & Dannefer, D. (1992). Aged heterogeneity: Fact or fiction? the fate of diversity in gerontological research. *The Gerontologist*, *32*, 17–23.

Neuberger, G. B., Kasal, S., Smith, K. V., Hassanein, R., & DeViney, S. (1994). Determinants of exercise and aerobic fitness in outpatients with arthritis. *Nursing Research*, *43*, 11–17.

Olbrich, E., & Diegritz, U. (1995). Das Zusammenwirken von Person- und Umweltfaktoren im Alltag: Eine kritische Diskussion von Aktivitäten des täglichen Lebens und instrumentellen Aktivitäten des täglichen Lebens [On the interaction of person and environment in everyday life: A critical discussion of activities of daily living and instrumental activities of daily living]. *Zeitschrift für Gerontopsychologie und -psychiatrie*, *8*, 199–212.

Olsho, L. W., Harkins, S. W., & Lenhardt, M. L. (1985). Aging and the auditory system. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 332–377). New York: Van Nostrand Reinhold.

Oswald, W. D., Rupprecht, R., & Gunzelmann, T. (1998). Effekte eines einjährigen Gedächtnis-, Kompetenz- und psychomotorischen Trainings auf die Leistungsfähigkeit im höheren Lebensalter [Effects of a one-year memory, competence and psychomotor training on performance in old age]. In A. Kruse (Ed.), *Jahrbuch der Medizinischen Psychologie. Psychosoziale Gerontologie, Band II: Intervention* [Yearbook of Medical Psychology. Psychosocial gerontology, Vol II: Intervention] (pp. 94–107). Göttingen, Germany: Hogrefe.

Ottenbacher, K. J., & Jannell, S. (1993). The results of clinical trials in stroke rehabilitation research. *Archives of Neurology*, *50*, 37–44.

Parmelee, P. A., & Lawton, M. P. (1990). The design of special environments for the aged. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 464–488). San Diego, CA: Academic Press.

Pathy, M. S. J. (Ed.). (1991). *Principles and practice of geriatric medicine* (2nd ed.). Chichester, UK: Wiley.

Pearlin, L. I., Aneshensel, C. S., Mullan, J. T., & Whitlatch, C. J. (1996). Caregiving and its social support. In R. H. Binstock & L. K. George (Eds.), *Handbook of aging and the social sciences* (4th ed., pp. 283–302). San Diego, CA: Academic Press.

Regnier, V., & Pynoos, J. (Eds.) (1987). *Housing the aged: Design directives and policy considerations.* New York: Elsevier.

Regnier, V., & Pynoos, J. (1992). Environmental intervention for cognitively impaired older persons. In J. E. Birren, R. B. Sloane, & G. D. Cohen (Eds.), *Handbook of mental health and aging* (2nd ed., pp. 763–792). San Diego, CA: Academic Press.

Reischies, F. M., & Lindenberger, U. (1996). Grenzen und Potentiale kognitiver Leistungsfähigkeit im Alter [Limits and potentials of cognitive performance in old age]. In K. U. Mayer & P. B. Baltes (Eds.), *Die Berliner Altersstudie* [The Berlin Aging Study] (pp. 351–377). Berlin, Germany: Akademie Verlag.

Rentz, D. M. (1991). The assessment of rehabilitation potential: Cognitive factors. In R. J. Hartke (Ed.), *Psychological aspects of geriatric rehabilitation* (pp. 97–114). Gaithersburg, MD: Aspen.

Rodin, J. (1986). Aging and health: Effects of the sense of control. *Science*, *233*, 1271–1276.

Rott, C., Wahl, H.-W., & Tesch-Römer, C. (1996). Zur Bedeutung von Seh- und Hörveränderungen für den Alternsprozess [On the meaning of visual and hearing impairment or the aging process]. In C. Tesch-Römer & H.-W. Wahl (Eds.), *Seh- und Höreinbussen älterer Menschen: Herausforderungen in Medizin, Psychologie und Rehabilitation* [Visual and hearing impairment in old age: Challenges within medicine, psychology, and rehabilitation] (pp. 89–106). Darmstadt, Germany: Steinkopff.

Rubenstein, L. Z., Josephson, K. R., Wieland, D., English, P. A., Sayre, J. A., & Kane, R. L. (1984). Effectiveness of a geriatric evaluation unit. *New UK Journal of Medicine*, *311*, 1664–1670.

Rustemeyer, J. (1983). Rehabilitative Massnahmen [Rehabilitation strategies]. In D. Platt (Ed.), *Handbuch der Gerontologie. Band 1: Innere Medizin* [Handbook of gerontology, Vol. 1, Internal Medicine] (pp. 13–32). Stuttgart, Germany: Fischer.

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging.* Hillsdale, NJ: Erlbaum.

Saltz, C. C. (1995). Rehabilitation. In G. L. Maddox (Ed.), *The encyclopedia of aging* (2nd ed., pp. 798–799). New York: Springer.

Sarno, M. T. (1984). Communication disorders in the elderly. In T.F. Williams (Ed.), *Rehabilitation in the aging* (pp. 161–176). New York: Raven Press.

Saudan, Y. (1990). Prinzipien der funktionellen Rehabilitation im Alter [Principles of functional rehabilitation in old age]. In F. Martin & J.-P. Junod (Eds.), *Lehrbuch der Geriatrie* [Textbook of geriatrics] (pp. 696–707). Bern, Switzerland: Huber.

Saup, W. (1993). *Alter und Umwelt. Eine Einführung in die Ökologische Gerontologie.* [Aging and environment. An introduction to environmental gerontology]. Stuttgart, Germany: Kohlhammer.

Schaie, K. W., & Willis, S. L. (1986). Can decline in adult intellectual functioning be reversed? *Developmental Psychology*, *22*, 223–232.

Schaie, K. W., & Willis, S. L. (1993). Age difference patterns of psychometric intelligence in adulthood: Generalizability within and across ability domains. *Psychology and Aging*, *8*, 44–55.

Schneekloth, U., & Müller, U. (1995). *Hilfe- und Pflegebedürftige in Heimen. Endbericht zur Repräsentativerhebung im Rahmen des Forschungsprojekts "Möglichkeiten und Grenzen selbständiger Lebensführung in Einrichtungen."* [Persons with basic care needs in institutions. Research report]. München, Germany: Infratest Sozialforschung und Infratest Epidemiologie und Gesundheitsforschung.

Schneekloth, U., & Potthoff, P. (1993). *Hilfe- und Pflegebedürftige in privaten Haushalten. Bericht zur Repräsentativerhebung im Forschungsprojekt "Möglichkeiten und Grenzen selbständiger Lebensführung."* [Persons with basic care needs in private households. Research Report]. München, Germany: Infratest Sozialforschung und Infratest Epidemiologie und Gesundheitsforschung.

Schulz, R., & Williamson, G. M. (1993). Psychosocial and behavioral dimensions of physical frailty. *The Journals of Gerontology*, *48* (Special Issue), 39–43.

Schütz, R. M. (1993). Geriatrische Rehabilitation—eine Standortbestimmung [Geriatric rehabilitation—A point of view]. In A. Niederfranke (Ed.), *Fragen geriatrischer Rehabilitation* [Questions of geriatric rehabilitation] (pp. 16–20). Stuttgart, Germany: Kohlhammer.

Shawe, D., Hesp, R., Gumpel, J. M., Sambrook, P. N., & Reeve, J. (1993). Physical activity as a determinant of bone conservation in the radial diaphysis in rheumatoid arthritis. *Annals of Rheumatic Diseases*, *52*, 579–581.

Silverstone, B. (1984). Social aspects of rehabilitation. In T. F. Williams (Ed.), *Rehabilitation in the aging* (pp. 59–79). New York: Raven Press.

Simeoni, E., Bauman, A., Stenmark, J., & O'Brien, J. (1995). Evaluation of a community arthritis program in

Australia: Dissemination of a developed program. *Arthritis Care Research*, *8*, 102–107.

Smith, C. A., Wallston, K. A., & Dwyer, K. A. (1995). On babies and bathwater: disease impact and negative affectivity in the self-reports of persons with rheumatoid arthritis. *Health Psychology*, *14*, 64–73.

Smith, J., & Baltes, P. B. (1996). Altern aus psychologischer Perspektive: Trends und Profile im hohen Alter [Aging from a psychological point of view: Trends and profile in the oldest old]. In K.U. Mayer & P.B. Baltes (Eds.), *Die Berliner Altersstudie* [The Berlin Aging Study] (pp. 221–250). Berlin, Germany: Akademie Verlag.

Smyer, M. A., Zarit, S. H. & Qualls, S. H. (1990). Psychological intervention with the aging individual. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 375–403). San Diego, CA: Academic Press.

Spicker, S. F., Ingman, S.R., & Lawson, I.R. (Eds.) (1987). *Ethical dimensions of geriatric care; Value conflicts for the 21st century*. Dordrecht, The Netherlands: Reidel.

Steinhagen-Thiessen, E., & Borchelt, M. (1996). Morbidität, Medikation und Funktionalität im Alter [Morbidity, medication, and functionality in old age]. In K. U. Mayer & P. B. Baltes (Eds.), *Die Berliner Altersstudie* [The Berlin Aging Study] (pp. 151–183). Berlin, Germany: Akademie Verlag.

Stephens, S. D. G., & Goldstein, D. P. (1983). Auditory rehabilitation for the elderly. In R. Hinchcliffe (Ed.), *Hearing and balance in the elderly* (pp.174–200). Edinburgh, UK: Churchill Livingstone.

Stuck, A. E., Siu, A. L., Wieland, G. D., Adams, J., & Rubenstein, L. Z. (1993). Comprehensive geriatric assessment: A meta-analysis of controlled trials. *Lancet*, *342*, 1032–1036.

Suzman, R. M., Willis, D. P., & Manton, K. G. (Eds.) (1992). *The oldest old*. New York: Oxford University Press.

Svanborg, A. (1988). The health of the elderly population: Results from longitudinal studies with age-cohort comparisons. In D. Evered & J. Whelan (Eds.), *Research and the ageing population* (pp. 3–16). Chichester, UK: Wiley.

Tallis, R. (1992). Rehabilitation of the elderly in the 21st century. The F. E. Williams Lecture 1992. *Journal of the Royal College of Physicians of London*, *26*, 413–422.

Tesch-Römer, C. (1997). Psychological effects of hearing aid use in older adults. *Journal of Gerontology. Psychological Sciences, 52B,* P127–P138.

Tesch-Römer, C., & Nowak, M. (1995). Bewältigung von Hör- und Verständnisproblemen bei Schwerhörigkeit [Coping with hearing and communication problems in aural impairment]. *Zeitschrift für Klinische Psychologie*, *24*, 35–45.

Tesch-Römer, C., & Wahl, H.-W. (Eds.) (1996a). *Seh- und Höreinbussen älterer Menschen: Herausforderungen in Medizin, Psychologie und Rehabilitation* [Visual and hearing impairment in old age: Challenges within medicine, psychology, and rehabilitation]. Darmstadt, Germany: Steinkopff.

Tesch-Römer, C., & Wahl, H.-W. (1996b). Was es bedeutet, (nicht) hören und sehen zu können [What it means (not) to be able to hear or see]. In C. Tesch-Römer & H.-W. Wahl (Eds.), Seh- und Höreinbussen älterer Menschen: Herausforderungen in Medizin, Psychologie und Rehabilitation [Visual and hearing impairment in old age: Challenges within medicine, psychology, and rehabilitation] (pp. 1–12). Darmstadt, Germany: Steinkopff.

Thomae, H. (1979). The concept of development and life span developmental psychology. In P. B. Baltes & O. G. Brim (Eds.), *Life span developmental psychology* (Vol. 2, S282–S312). New York: Academic Press.

Thomae, H. (1996). *Das Individuum und seine Welt: Eine Persönlichkeitstheorie* [The individual and her/his world] (2nd ed.). Göttingen, Germany: Hogrefe Verlag.

Trezona, R.R., Jr. (1991). The assessment of rehabilitation potential: Emotional factors. In R. J. Hartke (Ed.), *Psychological aspects of geriatric rehabilitation* (pp. 115–134). Gaithersburg, MD: Aspen Publishers.

Tucker, S.J. (1984). Patient staff interaction with the spinal cord patient. In D.W. Krueger (Ed.), *Rehabilitation psychology* (pp. 257–266). Rockvill, MD: Aspen.

US Department of Health and Human Services (1992). *Vital and health statistics: Current estimates from the National Health Interview, 1991. Series 10: Data from the National Health Survey, No. 184*. Hyattsville, MD: National Center for Health Statistics.

Vanderpool, J. P. (1984). Stressful patient relationships and the difficult patient. In D. W. Krueger (Ed.), *Rehabilitation psychology* (pp. 167–181). Rockvill, MD: Aspen.

Vignolo, L. A. (1964). Evolution of aphasia and language rehabilitation. *Cortex*, *1*, 344–367.

Vognsen, S. (Ed.) (1976). *Hearing tactics: How hearing impaired persons may solve their hearing problems in their daily life*. Copenhagen, Denmark: State Hearing Centre.

Wade, D. T. (1992). Stroke rehabilitation. In J. G. Evans & T. F. Williams (Eds.), *Oxford textbook of geriatric medicine* (pp. 321–328). Oxford, UK: Oxford University Press.

Wade, D. T., Collen, F. M., Robb, G. F., & Warlow, C. P. (1992). Physiotherapy intervention late after stroke and mobility. *British Medical Journal*, *304*, 609–613.

Wade, D. T., & Langton-Hewer, R. (1987). Functional abilities after stroke: measurement, natural history, and prognosis. *Journal of Neurology, Neurosurgery, and Psychiatry*, *50*, 177–182.

Wagner, M., Schütze, Y., & Lang, F. (1996). Soziale Beziehungen alter Menschen [Social relations of the elderly]. In K. U. Mayer & P.B. Baltes (Eds.), *Die Berliner Altersstudie* [The Berlin Aging Study] (pp. 301–319). Berlin, Germany: Akademie Verlag.

Wahl, H.-W. (1992). Ökologische Perspektiven in der Gerontopsychologie: Ein Blick in die vergangenen drei Jahrzehnte und in die Zukunft [Ecological perspectives in geropsychology: Viewing three decades backwards and into the future]. *Psychologische Rundschau*, *43*, 232–248.

Wahl, H.-W. (1994). Visual impairment in later life: A challenge for environmental gerontology. *Aging and Vision News, 6,* 1–2.

Wahl, H.-W. (1997). *Ältere Menschen mit Sehbeeinträchtigung: Eine empirische Untersuchung zur Person-Umwelt-Transaktion*. [Age-related visual impairment: An empirical study of person-environment transaction]. Frankfurt, Germany: Peter Lang.

Wahl, H.-W., & Baltes, M. M. (1993). Ökopsychologische Aspekte geriatrischer Rehabilitation [Ecological aspects of geriatric rehabilitation]. In A. Niederfranke (Ed.), *Fragen geriatrischer Rehabilitation* [Questions of geriatric rehabilitation]. (pp. 63–69). Stuttgart, Germany: Kohlhammer.

Wahl, H.-W., & Oswald, F. (1996). Schwere Seheinbussen im Alter aus psychologischer Sicht: Belastung und Bewältigungsmöglichkeiten [Severe visual impairment in old age from a psychological point of view: Stress and coping potential].. In C. Tesch-Römer & H.-W. Wahl (Eds.), *Seh- und Höreinbussen älterer Menschen: Herausforderungen in Medizin, Psychologie und Rehabilitation* [Visual and hearing impairment in old age: Challenges within medicine, psychology, and rehabilitation] (pp. 127–148). Darmstadt, Germany: Steinkopff.

Wahl, H.-W., & Oswald, F. (in press). The person/environment perspective of visual impairment. In B. Silverstone, M. A. Lang, B. Rosenthal, & E. Faye (Eds.), *The Lighthouse handbook on vision and rehabilitation*. New York: Oxford University Press.

Wahl, H.-W., & Tesch-Römer, C. (1996). Über die Technik hinaus - Psychosoziale Aspekte der Rehabilitation bei Hör- und Seheinbussen im Alter [Beyond technique - Psychosocial aspects of rehabilitation of hearing and vision loss in old age]. In C. Tesch-Römer & H.-W. Wahl (Eds.), *Seh- und Höreinbussen älterer Menschen: Herausforderungen in Medizin, Psychologie und Rehabilitation* [Visual and hearing impairment in old age: Challenges within medicine, psychology, and rehabilitation] (pp. 203–231). Darmstadt, Germany: Steinkopff.

Weinert, F. E., & Knopf, M. (1990). Gedächtnistraining im höheren Erwachsenenalter - Lassen sich Gedächtnisleistungen verbessern, während sich das Gedächtnis verschlechtert? [Memory training in old age - Improvement of memory performance and memory impairment at the same time?] In R. Schmitz-Scherzer, A. Kruse, & E. Olbrich (Eds.), *Altern - Ein lebenslanger Prozess der sozialen Interaktion* [Aging - A lifespan process of social interaction] (pp. 91–102). Darmstadt, Germany: Steinkopff.

Weinstein, B. E., & Ventry, I. M. (1983). Audiometric correlates of the Hearing Handicap Inventory for the Elderly. *Journal of Speech and Hearing Disorders*, *48*, 379–384.

Wiesner, M., & Tesch-Römer, C. (1996). Hörgerätebenutzung im Alter: Der Zusammenhang zwischen Intention und Verhalten [Use of an hearing aid in old age—The link between intention and behavior]. *Zeitschrift für Gerontologie*, *26*, 273–279.

Williams, T. F. (Ed.) (1984a). *Rehabilitation in the aging*. New York: Raven Press.

Williams, T. F. (1994). Rehabilitation in old age. In R.P. Abeles, H.C. Gift, & M.G. Ory (Eds.), *Aging and quality of life* (pp. 121–132). New York: Springer.

Willis, S. (1991). Cognition and everyday competence. *Annual Review of Gerontology and Geriatrics*, *11*, 80–109.

Willis, S. L. (1996). Everyday problem solving. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 287–307). San Diego, CA: Academic Press.

Willis, S. L., Blieszner, R., & Baltes, P. B. (1981). Intellectual training research in aging: Modification of performance on the fluid ability of figural relations. *Journal of Educational Psychology*, *73*, 41–50.

Willott, J.F. (1991). *Aging and the auditory system: Anatomy, physiology, and psychophysics*. San Diego, CA: Singular Publishers.

Wisocki, P. A. (Ed.) (1991). *Handbook of clinical behavior therapy with the elderly client*. New York: Plenum.

World Health Organisation. (1980). *International classification of impairments, disabilities and handicaps*. Geneva, Switzerland: World Health Organization.

Wottawa, H., & Thierau, H. (1990). *Lehrbuch Evaluation* [Textbook evaluation]. Bern, Switzerland: Huber.

Yesavage, J. A., Lapp, D., & Sheikh, J. L. (1989). Mnemonics as modified for use by the elderly. In L.W. Poon, D.C. Rubin, & B.A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 598–611). Cambridge, UK: Cambridge University Press.

Zeiss, A. M., & Steffen, A. M. (1996). Interdisciplinary health care teams: The basic unit of geriatric care. In L. L. Carstensen, B. A. Edelstein, & L. Dornbrand (Eds.), *The practical handbook of clinical gerontology* (pp. 423–450). Thousand Oaks, CA: Sage.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

7.24
Interdisciplinary Health Care Teams in Geriatrics: An International Model

ANTONETTE M. ZEISS,
VA Palo Alto Health Care System, CA, USA
and
ANN M. STEFFEN
University of Missouri at St. Louis, MO, USA

7.24.1 INTRODUCTION 552
7.24.2 WHY TEAMS? 552
7.24.3 INTERDISCIPLINARY TEAM THEORY 553
  7.24.3.1 Types of Teams 553
  7.24.3.2 Advantages of Interdisciplinary Teams 555
  7.24.3.3 Disadvantages of Interdisciplinary Teams (and Ways to Deal With Them) 556
7.24.4 NATURAL HISTORY OF TEAM DEVELOPMENT 556
  7.24.4.1 Processes of Team Development 557
    7.24.4.1.1 Forming 557
    7.24.4.1.2 Storming 557
    7.24.4.1.3 Norming 557
    7.24.4.1.4 Performing 558
    7.24.4.1.5 Team evolution 558
  7.24.4.2 Dilemmas of Team Work 558
7.24.5 DEVELOPMENT OF INTERDISCIPLINARY TEAM SKILLS 559
  7.24.5.1 Formal Educational Settings 559
  7.24.5.2 On-the-job Training 559
  7.24.5.3 Role Definition on Interdisciplinary Teams 560
7.24.6 INTERDISCIPLINARY TEAM TREATMENT PLANNING 561
7.24.7 INTERDISCIPLINARY TEAM CONCEPTS IN ACTION: TEAM RESPONSE TO A GERIATRIC PATIENT WITH MULTIPLE MEDICAL PROBLEMS AND DEPRESSION (SAM P.) 562
7.24.8 RESOURCES AND TRAINING MATERIALS 566
  7.24.8.1 Outcomes and Descriptions of Interdisciplinary Teams 566
  7.24.8.2 Team Training Guides 566
  7.24.8.3 Communication Training 567
  7.24.8.4 Factors that Influence Team Functioning 567
7.24.9 HELPING WITH INTERDISCIPLINARY TEAM DEVELOPMENT 568
7.24.10 SUMMARY 569
7.24.11 REFERENCES 569

7.24.1 INTRODUCTION

> As a mental health practitioner, you often receive referrals for the evaluation and treatment of patients who attend a variety of specialty clinics at your medical center. Your most recent referral is to assess and treat Sam P., a 76 year-old man exhibiting symptoms of depression. The Medical Resident treating Sam's cardiovascular disease and arthritis in the hospital's Geriatric clinic had become alarmed and made the referral after Sam began to cry and then refused to answer questions during his most recent medical appointment.

Your clinical decisions regarding the assessment and treatment of Sam P. will be influenced by your model of depression and how it is treated in older adults. Sam's response to treatment, however, will also depend upon your ability to work effectively with professionals from a variety of disciplines who are involved in his care. What difference in your work would there be if you were an official member of the interdisciplinary geriatric clinic team, as opposed to operating separately as a referral service? Few of us have received formal training in the advantages and tasks involved in working in interdisciplinary team settings. This chapter is designed to introduce both the conceptual and practical issues that are a part of interdisciplinary team work.

7.24.2 WHY TEAMS?

Geriatric patients are characterized by two pervasive qualities: their problems are complex and chronic (Binstock & George, 1990; Birren & Schaie, 1990; Dornbrand, Hoole, & Pickard, 1992; Schneider & Rowe, 1990). Problem complexity has many dimensions, including the following: patients may have multiple problems, each exacerbating the others; problems may present differently in geriatric patients than in younger patients; geriatric patients are likely to present inconsistent information about their health concerns; current problems are influenced in complex ways by a lengthy history of medical and/or psychosocial problems; and patients bring complex attitudes towards health care based on their past involvement with other health care providers (Hickey & Rakowski, 1981). Chronic health problems, such as cardiovascular complaints, arthritis, or pulmonary decline, are more typical of older individuals than younger ones (Federal Council on Aging, 1981). While older patients can, of course, also present with acute problems, the typical health care visit is more likely to be prompted by chronic concerns among elders (Woolf, Kamerow, Lawrence, Medalie, & Estes, 1990; Zawadski & Eng, 1988).

Older patients are also more heterogeneous than younger patients. While age typically brings decline, as measured by average capacity, on many biological functions (e.g., visual acuity, cardiac efficiency, speed of male erection, lung capacity), age is also typically characterized by increased variability (Fries & Crapo, 1981). Looking within a group of older adults, the range of capacities will be highly variable for any function, with some individuals displaying better function than the average younger person, whereas others display severe decline. There is also variability within individuals, with some functions remaining at high levels and others showing decline of varying degrees.

Because of these issues, the model of patient care which is the best fit for the management of such patients (i.e., those with chronic, complex problems which present in diverse ways and where stereotyped expectations about patients are inappropriate) is the interdisciplinary team model. This chapter will review basic aspects of interdisciplinary team theory, the natural history of team development, skills necessary for interdisciplinary teamwork, strategies for interdisciplinary team treatment planning, and models of effective interdisciplinary teams.

There has been a growing international awareness of the role of interdisciplinary care in a variety of settings. In Europe, concerns about the care of cancer patients led the Council of Europe in 1987 to create a Select Committee of experts to consider coordination of care for patients. The Committee recommended "the provision of joint professional education as a means of improving teamwork in cancer care ..." (Jones, 1992). The European Health Committee has funded a review of interdisciplinary training of health care staff in member states, and the World Health Organization has supported the development of a network for interprofessional education, based in Linkoping, Sweden (Goble, 1994). In the USA, a network of geriatric education centers (GECs) has been created to provide didactic and clinical training to a wide range of health professionals. These GECs are sponsored by the Bureau of Health Professions of the US Health Resources and Services Administration. Also in the US, the Interdisciplinary Team Training Program (ITTP) is funded by the Division of Veterans Affairs in 12 medical centers to promote direct service and training in a variety of inpatient and outpatient settings.

7.24.3 INTERDISCIPLINARY TEAM THEORY

To say that a group of health care providers is a "team" says very little, as there are many different kinds of teams, organized in different ways to perform different functions. A group that is a team shares a common workplace and set of patients, but teams differ among themselves in their membership composition, commitment to common goals, degree of collaboration in accomplishing team-related tasks, handling of leadership, and the kind of attention paid to team processes. While these factors theoretically could be combined in a large number of ways, in practice the five kinds of teams shown in Table 1 are commonly identified (Takamura; 1985; Takamura, Bermost, & Stringfellow, 1979). Interdisciplinary teams will be described in some detail, along with briefer descriptions of the other kinds of health care team organization.

7.24.3.1 Types of Teams

Unidisciplinary teams (from the Latin "uni," or one) are made up of many providers from a single background, such as a group of public health nurses providing home care. All team members share the same professional skills and training, speak a common language of health care, and function in the same role within the group.

Intradisciplinary teams (from the Latin "intra," or within) are also composed of professionals from one discipline, but include team members from different levels of training and skills within the discipline. For example, a geriatric mental health outpatient clinic team might be composed of a Licensed Psychologist, a postintern, prelicensing psychological assistant, a psychological technician, and two psychology practicum students.

The term *Multidisciplinary team* is often used as if it were synonymous with *interdisciplinary team*, but the two are actually different in very important ways. A multidisciplinary team (from the Latin "multi," or many) is composed of members from more than one discipline, so that the team can offer a greater breadth of services to patients. However, in a multidisciplinary team, each discipline does its own assessment, generates its own treatment plan, implements the plan, evaluates progress, and refines the plan based on their own evaluation (see Figure 1). Team members share information with each other, typically at meetings attended by all team members. There is no attempt to generate or implement a common plan, although members may choose to utilize information obtained in the meeting to revise their own goals or intervention strategies. Multidisciplinary teams are hierarchically organized, that is, there is a designated program "Chief," who is usually the highest ranking professional (commonly an MD). The designated leader is responsible for overseeing the program, leading meetings, resolving conflicts, allocating case load, and so on. Other team members feel responsible only for the clinical work of their discipline, and need not share a sense of responsibility for program function and team effectiveness.

Interdisciplinary teams (from the Latin "inter," or among) also are composed of members from more than one discipline, making a breadth of resources available to patients. Interdisciplinary teams work much more collaboratively, however, as shown in Figure 2. Individual team members do their own assessment, but generally have worked out overall assessment strategies so that patients are not asked redundant questions or submitted to

Table 1 Types of health care teams.

Type of team	*Defining characteristics*
Unidisciplinary team	Members from a single discipline; all members have the same role
Intradisciplinary team	Members from a single discipline, but at different levels of training; roles and responsibilities assigned by training level
Multidisciplinary team	Members from different disciplines; team members have independent roles and meet to share information
Interdisciplinary team	Members from different disciplines; team members have interdependent, collaborative roles and meet to plan treatment and evaluate response
Transdisciplinary team	Members from different disciplines, but team members largely have similar roles and responsibilities

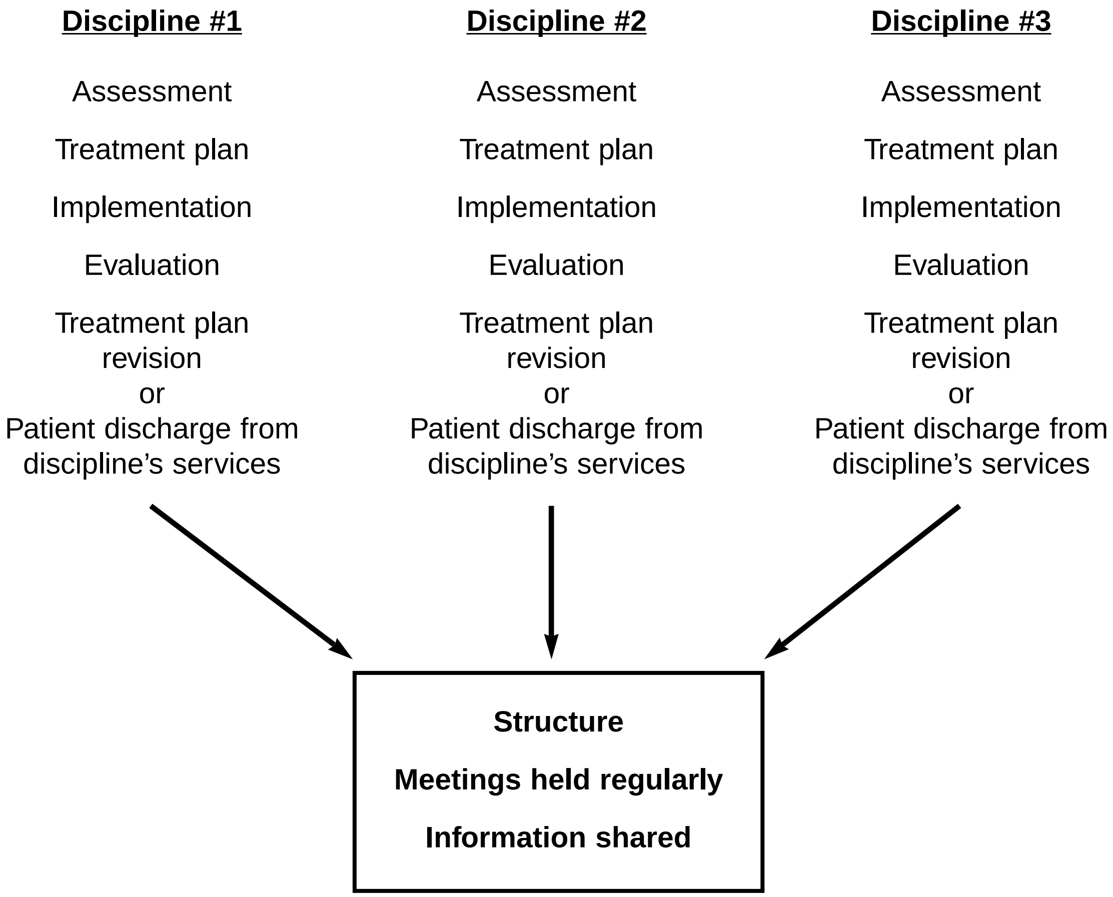
Discipline #1
Assessment
Treatment plan
Implementation
Evaluation
Treatment plan revision or Patient discharge from discipline's services
Discipline #2
Assessment
Treatment plan
Implementation
Evaluation
Treatment plan revision or Patient discharge from discipline's services
Discipline #3
Assessment
Treatment plan
Implementation
Evaluation
Treatment plan revision or Patient discharge from discipline's services
Structure
Meetings held regularly
Information shared

Figure 1 Organization of multidisciplinary team care.

repeated tests. Assessment information is shared with the team as a whole to generate an overall conceptualization of the relationships among biological, psychological, and social aspects of the case (the "biopsychosocial" model). This model is used to generate shared team goals for overall outcomes and to delineate how different disciplines will work together to create desired outcomes. Team members then can implement interventions, either individually or collaboratively as necessary, and evaluate progress. Those evaluations become new assessment data, which are brought back to the team to revise the team's goals and/or strategies for reaching them. This process is repeated as often as necessary until goals are achieved and the patient can either be discharged or a maintenance plan can be put in place.

In an interdisciplinary team, the group as a whole takes responsibility for program effectiveness and team function. Leadership functions are shared among members, with the expectation that everyone must be equally committed to both the clinical content of the program and the process in which team members work together. All team members are assumed to be colleagues, and there is no hierarchical team organization. One member may be designated team "coordinator," but this is understood to be an administrative role and does not imply that the team member has higher status, a stronger say in resolving conflict, or the ability to make unilateral decisions. Because interdisciplinary teamwork requires so much collaboration and consensus decision-making, team members must have a high degree of interpersonal skill. The team must make a commitment to spending time developing effective working strategies and to attending consistently to the process of working together as well as to the quality of the outcomes of their work. Communication skills, role development, leadership skills, conflict resolution skills, and skills for comprehensive conceptualization of cases are necessary for all team members.

This approach has also been labeled "interprofessional care," especially in Europe. A new journal which is rapidly becoming a primary source of articles on interdisciplinary care is titled the *Journal of Interprofessional Care*. Interdisciplinary and interprofessional seem to be interchangeable terms; there do not seem to

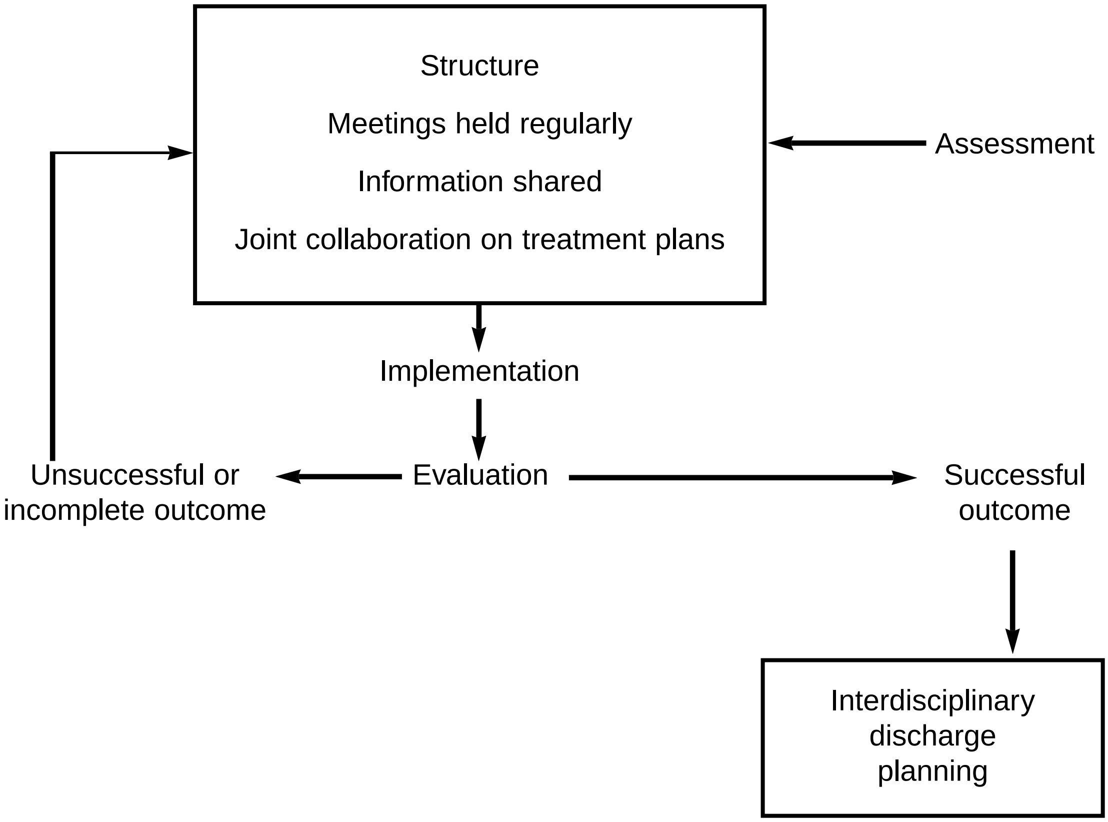
Structure
Meetings held regularly
Information shared
Joint collaboration on treatment plans
Assessment
Implementation
Evaluation
Unsuccessful or incomplete outcome
Successful outcome
Interdisciplinary discharge planning

Figure 2 Organization of interdisciplinary team care.

be important conceptual distinctions between them.

In *transdisciplinary teams* (from the Latin "trans," or beyond), disciplinary lines are blurred, and team members share role functions to a high degree. For example, nurses, psychologists, social workers, and occupational therapists in an adult day health care program might all drive the van to bring in patients, run groups, see individual cases for supportive counseling, supervise lunchtime interactions, work with patients on increasing physical and cognitive activity, and so on. This model does not have wide application in geriatric settings, since patients generally need the special skills of each discipline, obtained through extensive specialized training, not just generic skills that many disciplines might share.

7.24.3.2 Advantages of Interdisciplinary Teams

The major reason to adopt a particular model of health care is that it provides the most cost-effective results. In cost-effectiveness analyses, effectiveness should be considered first: for example, does the approach reduce symptoms, increase quality of life, promote health, and prevent future illness? If more than one approach is effective, then it is meaningful to ask what the relative costs of each approach are and to select the program that provides the greatest effectiveness for the least cost. Well-functioning interdisciplinary teams provide cost-effective care by increasing the effectiveness of outcomes, thereby decreasing costs for continued treatment (e.g., Hendriksen, Lund, & Stromgard, 1984). There are four usual ways in which this occurs. First, interdisciplinary teams reduce duplication of services, compared to multidisciplinary teams, because team members plan carefully how to allocate resources and to cover necessary tasks. This may involve the collaboration of more than one discipline but precludes duplication. For example, on a multidisciplinary team both the social worker and the psychologist might begin individual psychotherapy for depression with a patient; this would not happen on an interdisciplinary team. Second, a related cost-saving impact is that problems do not fall through the cracks when care is organized by an interdisciplinary team. It is not uncommon on units staffed by multidisciplinary teams for discharge to be planned and then delayed repeatedly as the staff discover that not all the necessary work has been done; for example, the home care coverage is arranged before the physical therapist has become aware of the need to see the patient, and then home care plans have to be rearranged. A third advantage is that interdisciplinary teams can generate more creative interventions, since the ideas and knowledge of a group can be brought to bear and team members can stimulate each other's ideas. Finally, interdisciplinary teams can reduce institutional costs, because they increase staff morale and reduce staff turnover (Lichtenberg, Strzepek, & Zeiss, 1990; Sbordone & Sherman, 1983). As a result, more continuous care can be provided and the cost of recruiting and training replacement staff is reduced.

While there is a need for more research to examine each of the elements of this argument in favor of the cost-effectiveness of teams, the evidence available does support the interdisciplinary team approach. For example, Zeiss & Okarma (1985) compared interdisciplinary team approaches to multidisciplinary care in an outpatient rheumatology clinic setting. Interdisciplinary team approaches were shown to improve patient outcomes on measures of quality of life (such as life satisfaction, self-rated pain, and depression) and measures of health care utilization (e.g., inpatient stays, the most costly care delivery modality). Another study in Sweden compared the outcomes of two groups of female outpatients with rheumatoid arthritis; one group had interdisciplinary care and the other standard care by a rheumatologist. While there were no significant differences in joint function or disease activity, overall health improved significantly over one year only in the interdisciplinary group (Ahlmen, Sullivan, & Bjelle, 1988). Other research showing the effectiveness and/or cost-effectiveness of interdisciplinary teams is reviewed in Lichtenberg (1994) and West and Poulton (1997).

Concern is sometimes expressed that interdisciplinary teams may be too costly, because of the number of disciplines involved. However, interdisciplinary teams have no specific complement of personnel; it is not necessary that every team have a physician, nurse, psychologist, social worker, pharmacist, occupational therapist, and so on. Teams need to include members of those professions who are essential to provide effective care for the identified patient population. Thus, a well-designed interdisciplinary team involves no more personnel than would normally staff a clinic or inpatient unit for that population. The staff are simply organized differently than they would be if the unit had some other type of team (or no team framework at all).

7.24.3.3 Disadvantages of Interdisciplinary Teams (and Ways to Deal With Them)

Even though interdisciplinary teams can be effective, they are not necessarily the treatment of choice in all care settings. Settings where patients have simple, routine problems would probably benefit more from some other form of care, since in-depth discussion of each case is unnecessary if everyone is going to receive essentially the same straightforward intervention. A clinic offering acute care to generally healthy adults, for example, is best staffed by nurses working with physicians; adding staff and meetings for the majority of the presenting concerns would serve no useful purpose. If a 40-year-old woman comes down with bronchitis, adding collaborative care from a social worker, pharmacist, psychologist, and other disciplines to a routine medical visit is unlikely to be cost-effective or satisfying to either the patient or the providers. It is rare, however, that elderly individuals have straightforward problems or benefit from routinized care. If an 80-year-old woman comes down with bronchitis, it may be essential to know that she is caregiver for a demented husband, what other medications she is taking, what her immune status is, and other issues pertinent to her physical and emotional well-being. Social work, pharmacy, psychology, psychiatry, and/or other team members may be essential in planning not only acute treatment for her bronchitis but also backup care for her demented husband.

The other potential drawback of interdisciplinary teams is that they are not organized to provide rapid response in crisis situations; they are organized to provide care for chronic, complex problems. If a patient calls with an acute problem on Tuesday morning, and the next team meeting is not scheduled until the following Monday afternoon, there must be a way for individual team members, or a smaller set of team members, to make emergency decisions. Interdisciplinary teams can handle this by allocating such authority to team members, with the understanding that the issues will be raised at the next team meeting so that the implications of the crisis can be factored into the team's overall plan and the treatment responsibilities of other team members.

7.24.4 NATURAL HISTORY OF TEAM DEVELOPMENT

Groups do not immediately become well-functioning interdisciplinary teams by deciding to make a commitment to that model of care. Much learning and development within the team must occur, over a period of time which lasts at least several months and sometimes longer. The term "team building" is sometimes used to describe activities during this period, but it is a particularly inappropriate term, because of the implications of the word "building," which could imply that a solid structure, impervious to change, is being built. In fact, good teams learn to change continually and to respond fluidly to new demands, new patients, new staff, new health care policies, and so on, through a coordinated, collaborative process. Thus, the term "team development" is preferable, since it implies an ongoing, fluid process.

7.24.4.1 Processes of Team Development

The sequence of experiences that groups go through in the process of developing cohesion and a high level of function has been described by many authors (e.g., Center for Interdisciplinary Education in Allied Health, 1980; Gersick, 1989; Scholtes, 1988), but most influentially by Tuckman (1965). He identified four processes which usually occur in sequence in the process of team development. He labeled these four processes Forming, Storming, Norming, and Performing; they will be described in more detail. It is important to note that these are processes, not stages. If these four processes are seen as four stages that all teams must go through, the opportunity to understand each team's unique pattern can be obscured. In addition, since these are fluid processes, it is not helpful to think of the team as being in only one stage at any given time. For example, a team may be still Storming with regard to some issues while it has already progressed to Performing in other areas. Furthermore, teams can change by leaping forward or falling back in response to personnel changes, the development of new treatment options, stresses or positive events in team members' lives, or a variety of other events. Nonetheless, these processes can be examined to help in conceptualizing the changes and skills necessary for effective team functioning without falling into a rigid stage theory framework.

7.24.4.1.1 Forming

Teams may be formed when programs develop, expand, or are consolidated. A new team must deal with a number of issues, including which staff are part of the team; which patients will be served; what resources are available, such as hospital beds or outpatient visit rooms; how the team treatment planning process will be organized; and so on. Early in the life of the team, these decisions about content matters take priority, and little attention is paid to the process of working together. Team members typically feel a bit awkward with each other, and are willing to develop team relationships but unsure of what that involves. Team members may have concerns about trust, wondering whether it is all right, for example, to express disagreement or concern about how the group is working together.

7.24.4.1.2 Storming

As the team continues to work together, process issues become unavoidable. It becomes clear that team members do not always agree about content issues, for example, how the team treatment planning should be set up. Differences in personal style may need to be faced: one person may handle concerns directly in ways that others see as too brusque, or a team member may try so hard to be tactful that no one can tell what he or she is saying. Team members may make different assumptions about basic working structure, for example, some may assume that starting the meeting at 10 a.m. means 10 on the dot, whereas others assume it means any time before 10.15 a.m. More substantive issues may arise, for example, different beliefs about whether the first-level intervention for depression should be antidepressant medication or psychotherapy. The group will have to learn to voice disagreement in a way that communicates a desire to stay engaged and work out the issues fairly in a collaborative process. If it cannot, the group will likely disband or fall back to multidisciplinary functioning, where individuals can make autonomous decisions without the need for collaboration and consensus.

7.24.4.1.3 Norming

If the group can work on areas of disagreement constructively, a working strategy will evolve. This strategy can be thought of as "ground rules" or norms which describe the way the group will work together and resolve conflict. Norms can be very specific (e.g., "starting at 10 a.m. means any time before 10.05 a.m.") or more general (e.g., "When team members disagree we will engage in a complex process to show our respect and trust for each other and remind each other that good people can disagree; then we will start to work on the substance of the issue"). If the norms developed are reasonable and effective, the group will set aside time to check on its progress (often referred to as "process time") and to continue clarifying its norms. In addition, content decisions will be made and acted on, and the group will begin to feel cohesive.

If the norms are ineffective, the group will feel stuck and rising frustration will throw the group back into Storming. For example, if the team develops a norm that polite agreement will be voiced on most occasions when disagreement occurs, team meetings may seem peaceful at first, but team members who disagree privately are likely to implement their own ideas rather than the ideas ostensibly agreed on in the meeting. This leads to hard feelings and the belief that meetings waste time; the group will have to learn to express disagreement openly and honestly before effective norms can be established that really resolve issues from the Storming process.

7.24.4.1.4 Performing

Over time, the group will develop a high level of effectiveness, called "performing," in which the team becomes expert at using its process to make decisions and get work done. Ground rules will rarely need to be explicitly stated; the group might seem to be on automatic pilot, especially to outsiders who see a seamless performance with deceptively little effort being needed to make complex decisions. The group will feel cohesive and morale will be high.

7.24.4.1.5 Team evolution

Unfortunately, maintaining function in Performing mode consistently over time is unlikely and not a realistic team goal. Key staff members leave; a needed new discipline is added to the team; the program is moved to a different building; or trainees rotate through at four-month intervals. The group must be able to change accordingly, and it is likely that the team must go back through the four processes of team development before settling into Performing mode again. The new team and its resources and responsibilities must be examined in a Forming phase. Old norms may no longer work (e.g., if the team member whose role it was to draw out quieter team members has left), so there will be a Storming process as the implications of change for the day-to-day life of the team become clear. The team will develop new norms, which maintain as much of what worked before as possible, but which also enthusiastically embrace the different skills and personal style of any new team members. If continual change is a reality, for example, if new trainees appear at frequent intervals, the team will develop metanorms for how to allow trainees to have some impact on the group, without constantly having to redefine the group from scratch.

7.24.4.2 Dilemmas of Team Work

One helpful way to think about the processes groups need to go through to become performing teams is summarized in Table 2. Groups must learn to deal effectively with three inevitable challenges: interdependence, complexity, and disagreement.

The first issue the team faces is interdependence; by definition, an interdisciplinary team seeks to share responsibility for patients by working collaboratively. Early in the life of the team, this usually leads to turf battles, as each team member realizes that his or her skills overlap, often substantially, with the skills of other team members. In well-functioning teams, this will be a source of camaraderie and confidence, but early in the team's development, members can fight over who "should" be doing various activities. For example, teams might argue about whether social work or psychology "should" do family therapy, or whether pharmacy or nursing "should" do home visits to check patients' medicine cabinets and prevent mistakes in the home drug regimen. In the Storming process, these issues will be highlighted in order to develop team-specific agreements (norms) about how such tasks will be distributed and/or shared. The section on team role maps below (7.24.5.3) gives more detail on this process.

The second issue is complexity. By definition, interdisciplinary teams are established to deal with complexity, and case discussions, to be useful, must increase the complexity of the

Table 2 Dilemmas of teamwork.

Theme	*Problem*	*Opportunity*	*Work to be done*
Interdependence	Turf battles	Collaboration	Knowledge of and respect for diversity and shared abilities
Complexity	Endless assessment	Comprehensive care	Conceptualization, key issues, language of function
Disagreement	Destructive conflict	Integration of multiple perspectives	Communication skills

whole group's understanding of the interrelationships of each patient's biological, functional, social, and psychological problems. On a well-functioning team, this leads to a breadth of understanding that allows effective comprehensive care. Early in the life of the team, it can lead to information overload and frustration that cases are discussed seemingly without end. In the Storming process, the team must develop strategies for integrating information and capturing the team's complex understanding of the patient in a written format that facilitates treatment planning and implementation. Section 7.24.6 provides some examples of formats teams can use for developing treatment plans that capture complexity and promote collaborative intervention.

The third issue is disagreement. Interdisciplinary teams rely on healthy conflict, in which the different perspectives and knowledge of team members can be presented. It is illusory to believe that these different perspectives will always be complementary; sometimes they are in direct contradiction. For example, the patient may have told one team member that he is sleeping better since starting a behavioral sleep hygiene program, but told another team member that his sleep problems are unchanged. Teams that are having problems are likely to treat these differences as issues in which one team member is right and the other is wrong, leading to conflict destructive to team cohesion and effective patient care. However, in a functioning team, the team members will recognize that there is direct conflict and see it as an opportunity to learn something more about the patient and themselves. The patient may be trying to manipulate one team member to obtain pain medications. Or the patient may be more demented than the team realized and generating unreliable answers to questions. It is also possible that different team members have different interview styles that elicit different information from the patient, or team members may attend selectively to different kinds of information patients present. The important point to remember is that the ability to consider multiple possibilities and integrate multiple perspectives in handling conflict is the hallmark of a highly functioning team. A detailed analysis of strategies for learning to handle conflict effectively are beyond the scope of this chapter, but important sources giving such analysis are presented in Section 7.24.8.

Thus, interdisciplinary teams must develop the capacity not only to deal with the patient issues that the program was originally designed to treat, but also to deal with the interpersonal world of collaborative care. Skills that are necessary for all team members include knowledge and respect for other team members' abilities, the capacity to conceptualize cases holistically, the expertise to develop written team treatment plans which capture the clinical conceptualization, communication skills, leadership skills, conflict resolution skills, and the capability to anticipate and respond to change.

7.24.5 DEVELOPMENT OF INTERDISCIPLINARY TEAM SKILLS

7.24.5.1 Formal Educational Settings

Since the mid-1980s there has been an increase in university-based gerontology programs. However, these programs may not result in the development of skills needed for working in interdisciplinary environments. Clark (1993) argues that the basic departmental structure of universities can create problems for maintaining a truly interdisciplinary perspective. Interdisciplinary principles and practice cannot be taught in a classroom with didactic methods over a short period of time. Instead, interdisciplinary experiences must be integrated into the program at a variety of points, so that students can master the conceptual frameworks and acknowledge the differing value orientations of different disciplines. In discussing interdisciplinary education embedded in specific university departments, Mitchell (1993) states that exposure to clinical settings is essential for beginning health care students if they are to change their traditional orientation and embrace an interdisciplinary perspective of patient care. For example, only 9% of first-year physiotherapy students acknowledged the importance of being able to communicate with other health professionals (Mitchell, 1993); it appears that these students had already developed specific beliefs about therapeutic practice.

Clark's (1993) guidelines for interdisciplinary education involved the following recommendations for educators who need to: (i) clarify the educational objectives of interdisciplinary programs, (ii) attend to team process issues by including a substantial experiential component, (iii) develop educational methods and materials for communicating across disciplines, (iv) focus on obtaining high-level institutional support for interdisciplinary programs, and (v) recognize that interdisciplinary education represents a true educational revolution.

7.24.5.2 On-the-job Training

The majority of professionals are exposed to interdisciplinary team concepts for the first time when at work. Most interdisciplinary health care teams are not formed with the intent to be time limited, and teams have the additional challenge of continuously changing group membership. Because of these factors, team skills need to be taught on an ongoing basis to individuals with very different beginning skill levels. Drinka (1994) points out that it is often the task of ongoing members to teach new members how to collaborate and assume leadership within the team. This process may be difficult at times, since developing a collaborative set involves integrative problem-solving skills in professionals who may be used to working independently. All team members must feel empowered to participate in conflict resolution. While individual members will have varying degrees of formal power, informal power is also important, particularly for new team members. Informal power can be gained through organizational skills, developing increased knowledge about the other disciplines, demonstrating commitment to the team's mission, and increased tenure (Drinka, 1994). Teams can facilitate this process through inservices and retreats that foster both an increased knowledge of other disciplines and a sense of ownership of the team's mission. Individual members may also protect the rights of new members by drawing them into discussions, for example, by asking their opinion before soliciting the view of established team members. Team members also need to feel comfortable enough with the team to disagree in order to avoid poor decision-making (Heinemann, Farrell, & Schmitt, 1994). In addition to team retreats that provide opportunities for informal social activities, team members can practice brain-storming exercises and rotate the role of "devil's advocate," thus formalizing the importance of alternative views (Heinemann et al., 1994).

7.24.5.3 Role Definition on Interdisciplinary Teams

One of the activities which helps a team progress is the development of a shared understanding of the roles of each team member. An example of such a role map is shown in Figure 3; this role map is adapted from one generated for a hospital-based home care team (HBHC). This Department of Veterans Affairs program serves home-bound older patients and their caregivers; all team members drive to each patient's home to assess needs and provide services. Team members shown in the figure are a physician, four nurse practitioners, a social worker, an occupational therapist, a pharmacist, and a psychologist. Each discipline is represented by an oval, and each oval overlaps with others. Within each discipline's oval are listed the components of their role. In the part of the oval that does not overlap with other ovals appear clinical responsibilities assumed only by that discipline. In areas where two circles overlap, clinical responsibilities appear which either discipline might perform or which the two team members might perform as co-therapists. For example, either psychology or medicine might do cognitive screening.

There are also areas where three, or even more, ovals overlap. In those intersections appear tasks which several team members might perform, for example, nursing, medicine, or pharmacy might all educate patients about the effects and side effects of medications. Figure 3 has been simplified for clarity and does not include all the roles of any of the disciplines, but a more complex version could be generated that did demonstrate the complete range of tasks performed.

Generating and/or examining such a diagram provides several kinds of information. First, it can delineate all the clinical services which patients could receive from this team. Second, it clarifies for each discipline what they offer to the team and what the team will expect them to provide. Third, it clarifies areas which, on a well-functioning team, will be opportunities for conjoint efforts or, in a poorly-functioning team, will be sources of conflict that can be termed "turf battles." If pharmacy and nursing each want to hold the unique right in this setting to educate the patient about medications, for example, the two disciplines are likely to be in conflict about this issue. If each sees the other as a collaborative resource instead, team work (and patient care) will be enhanced. Fourth, Figure 3 suggests which disciplines may have the most difficulty speaking each other's language. The roles of pharmacy and psychology, for example, overlap very little in this setting. In a well-functioning team, members will see these as complementary functions and take pride in being able to provide a range of services well beyond the purview of any discipline. In a poorly functioning team, members will complain about the time taken to discuss issues outside their own scope of practice and compete for time to discuss the issues most closely connected to their own role.

To generate a role map which captures the realities of a specific program and the disciplines on its team takes some time, but is usually well worth the effort. Each team member needs to specify his/her role and present it for team discussion; often this results in adding tasks to the list which the team member did not specify but recognized when mentioned by other team

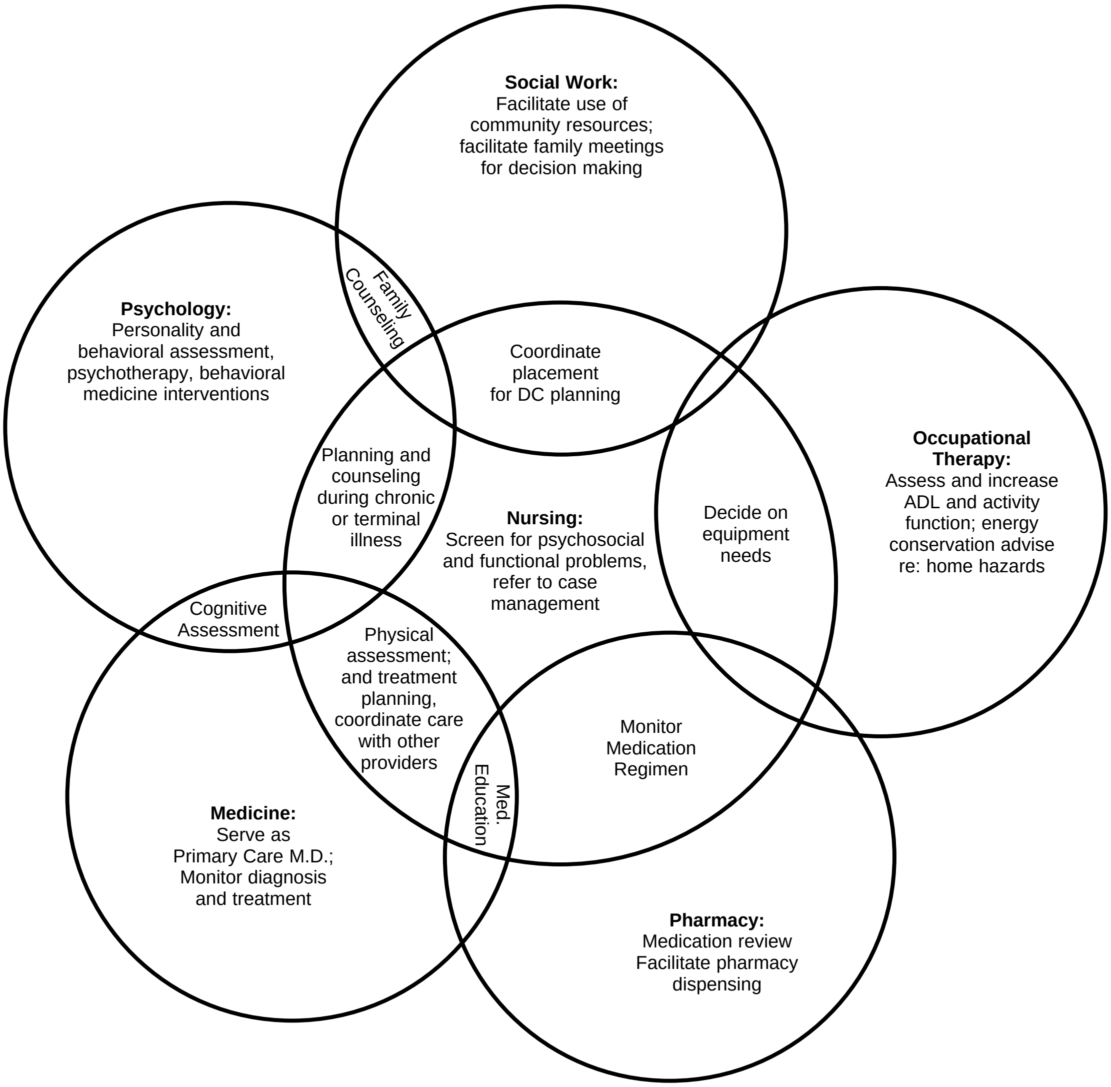
Social Work:
Facilitate use of community resources; facilitate family meetings for decision making
Psychology:
Personality and behavioral assessment, psychotherapy, behavioral medicine interventions
Family Counseling
Coordinate placement for DC planning
Occupational Therapy:
Assess and increase ADL and activity function; energy conservation advise re: home hazards
Planning and counseling during chronic or terminal illness
Nursing:
Screen for psychosocial and functional problems, refer to case management
Decide on equipment needs
Cognitive Assessment
Physical assessment; and treatment planning, coordinate care with other providers
Monitor Medication Regimen
Med. Education
Medicine:
Serve as Primary Care M.D.; Monitor diagnosis and treatment
Pharmacy:
Medication review Facilitate pharmacy dispensing

Figure 3 Geriatric interdisciplinary team role map.

members. This also begins to highlight the areas of overlap, as team members recognize some of their own tasks on colleagues' lists. After all disciplines have presented their roles, someone on the team, or a small group, needs to organize the lists to capture graphically areas of overlap and areas of unique function. This document can then become a tool to help in shaping norms and to help in the process of socializing new team members and reaching or returning to the Performing level of team process.

7.24.6 INTERDISCIPLINARY TEAM TREATMENT PLANNING

Interdisciplinary teams generate treatment plans through two interlocking processes. First, the team must be able to conceptualize patients broadly, incorporating all the relevant information and the relationships between different problems. Thus, the team must go beyond simply listing patient concerns on a problem list, to look at the relationships among problems. For example, a patient's depression may be triggered by loss of independence in activities of daily living (ADL) and the depression may be the primary reason (because of the patient's apathy and hopelessness) for medication nonadherence. Psychology or psychiatry could list depression on the problem list; occupational therapy could list difficulties with ADL; pharmacy could list medication nonadherence. But the efforts of each discipline to deal with these problems will be enriched if the linkages among problems are understood.

This type of conceptualization will be demonstrated in more detail in the case study in Section 7.24.7.

Second, the team must have a structured format for guiding and capturing on paper the treatment planning process. That structure should, at minimum, specify which problems will be treated, what the goals of treatment are, what treatment strategies will be used, and which team members will be involved. It is helpful to indicate which team members will be involved by name as well as discipline, to emphasize accountability. The plan may indicate time frames, specifying which interventions should begin immediately, which will be implemented after progress is made on initial problems, and the expected time frame to accomplish the goals. The plan could also include space to update status of progress generally, such as noting if the treatment goals are accomplished

The plan should be available to all team members to guide their activities in implementing their own roles, and it should be available at team meetings to guide discussion. To ensure continuous availability, teams can circulate copies to all members or keep charts in an easily accessible place. To guide team discussions, plans can be shown using overhead transparencies, or circulated at the meeting, or a team member can review highlights of the plan before case discussion.

7.24.7 INTERDISCIPLINARY TEAM CONCEPTS IN ACTION: TEAM RESPONSE TO A GERIATRIC PATIENT WITH MULTIPLE MEDICAL PROBLEMS AND DEPRESSION (SAM P.)

In order to further clarify the use of the team treatment planning process, we return now to the case which opened the chapter, Sam P., a man referred to a mental health practitioner for the assessment and treatment of depression. Let's look at this case in more detail to see how an interdisciplinary perspective influences his treatment. This case also offers the opportunity to contrast how an interdisciplinary team would approach the planning and treatment processes differently from a multidisciplinary team.

> Sam P. is a man in his mid-70s of Italian-American background. He has been married to his wife, Pauline, for 40 years: they have been in the same two-storey home for their entire marriage and enjoy gardening, visiting with family and friends, and participating in some church activities. For the last 15 years, Sam has had progressively greater trouble with day-to-day function as a result of his medical problems. His arthritis results in joint pain, especially in his hands, knees, and hips. This joint pain also negatively affects his sexual function. Sam's cardiovascular disease results in impaired blood flow, especially to his legs, to the extent that he gets cramping pain in his legs after walking as little as half a block. Impaired blood flow also has impaired Sam's ability to get and keep an erection. Taken together, these problems have affected almost every aspect of Sam's life. He has trouble getting up and downstairs in his home, he cannot do the gardening he enjoys, he has trouble joining family and friends on outings, and he and his wife have stopped any form of physical intimacy because of his frustration with his sexual function problems.
>
> Sam is taking various medications for his medical problems and comes to the geriatric clinic every three months on average. At the clinic, the medical residents rotate every four months; Sam feels frustrated when he has to retell information to a new doctor. He sometimes gets disgusted and angry with clinic staff, and forgets the questions he has about his medications. On several occasions, Sam has lost track of his medication schedule, became overmedicated, and developed severe nausea, diarrhea, and cognitive confusion. At other times, he has failed to take medication for days at a time, resulting in increased pain and decreased energy.
>
> Sam says that he sometimes feels depressed and useless. He and his wife used to baby-sit their infant granddaughter, but Sam can no longer keep up with a toddler and gets rattled by too much noise and activity. Sam used to enjoy activities in his neighborhood, such as pot-lucks and reading groups, but has stopped going because of his physical problems. His major independent forms of recreation before his decline were driving his friends on outings and attending Elks Club meetings; he has also stopped trying to do these things. He reports that he is often "blue" and tearful, and describes a poor appetite and waking early, about 4 a.m., several days a week.
>
> During the 40 years of their marriage, Sam and Pauline enjoyed physical intimacy and Sam describes feeling sad and frustrated that he "can't perform." Besides problems with erections, Sam often feels tired and is afraid that having intercourse would put too much physical stress on his joints. Sam also admits to feeling self-conscious about the way he looks and has a hard time believing that his wife still finds him sexually attractive.

You can see from this case description that a competent professional in any one discipline cannot address the many issues involved in Sam's physical and mental health. Like many older adults, his treatment should involve the cooperation of a number of disciplines, including nursing, social work, medicine, psychiatry, pharmacy, occupational therapy, and psychology. The psychologist involved in the case would probably conduct a careful mood assessment, then contact the psychiatry consult service to determine if Sam's medication regimen is commonly associated with depression. Several key issues likely to be related to Sam's negative mood would be examined: decreased participation in activities that are both meaningful and pleasurable for him; beliefs that he is no longer useful to others; and disruption of his marital intimacy, both physically and emotionally, as a result of his sexual dysfunction. The psychologist may consider some neuropsychological assessment to evaluate Sam's short-term memory, given that he has forgotten to take some doses of medication in the past and taken too much medicine at other times. There are additional areas that could be targeted for change, such as his frustration and communication difficulties with clinic staff. Psychiatry and psychology may discuss together the usefulness of prescribing antidepressant medications.

The other disciplines involved in his care would have their own concerns and areas of interest. The geriatric medical resident assigned to Sam would continue to evaluate choice of medications with the goals of reducing pain and minimizing joint damage, and maximizing cardiac function and circulation. The nursing staff may want to further educate Sam and his family about his medical diagnoses and the reasons for daily fluctuations in pain and function. Occupational therapy might work with Sam on selecting assistive devices to aid him in difficult tasks (e.g., buttons, zippers, jar lids), create hand splints to protect his hands from further arthritic damage, and plan ways for him to continue important daily activities (e.g., design a table-top garden for which he does not have to kneel). The social worker involved in the case would be aware that Sam and Pauline's two-storey house is a problem because of Sam's swollen knees and peripheral neuropathy, and would discuss housing and long-term plans with them. Sam's episodes of overdosing on medications would be concerning to the pharmacist, who may consider ways to reduce the likelihood of this re-occurring.

So far, it may appear that each professional could execute her or his own treatment plans without the input of the other disciplines (i.e., as a multidisciplinary team). Upon closer inspection, the evaluation and treatment goals in each area overlap in a way that creates the potential for conflict as well as cooperation. For example, the psychologist's suggestions for increasing pleasant activities may appear to directly contradict what Sam has been told by the occupational therapist and the geriatric medical resident. In addition, the psychologist and social worker might be convinced that his medication overdose and seeming confusion during appointments are related to complex issues such as anxiety about his illness, some recent decline in short-term memory that should be monitored, his response to the staffing patterns of the clinic, and a lack of support from his family. This may contradict the pharmacist's approach that the main issue is regulating medication regimens and providing more education about the purpose of each medication. From a different perspective, the medical resident may believe that too much time is spent in team meetings discussing Sam's emotional and psychosocial reactions to his illness. After all, the severity of his symptoms would decline significantly if he would just carefully follow his medication regimen and not engage in strenuous physical activity. These conflicts might be traced to several aspects of the team setting: poor patterns of communication during and outside of team meetings, a decision-making process that minimizes opportunities for discussion or dissent, and a lack of understanding and respect for each discipline's norms and perspectives for patient care.

In contrast to how a multidisciplinary team would approach the patient, the interdisciplinary team in this case would spend more time and attention coordinating Sam's evaluation and care. Through regularly scheduled opportunities for education during team meetings, the interdisciplinary team members would develop a greater understanding of the competencies and perspectives of the other disciplines. This would result in the team's enhanced ability to identify shared treatment goals and the specific needs for collaboration necessitated by each patient.

What team response could be expected in Sam's case? The following section describes an interdisciplinary effort focused on several interwoven goals and strategies. From the information provided in the case description plus some additional assessments, a treatment plan was developed. An important issue that became clear during interdisciplinary discussion was that Sam's care had been fragmented and spread across several specialty clinics, rather than being consolidated in a primary care setting and coordinated. Sam's loss of mobility was a particular focus. He had been seen by orthopedics, who planned hip replacement surgery with the expectation that he would be able to return to most of his prior activities after surgery. However, when seen by cardiology, he was told that hip replacement surgery was unnecessary and that he would be more mobile if his cramping pain (known as claudication) was treated with a medication to improve cardiovascular performance. Sam tried that medication, but stopped using it because of side effects, especially headaches. When he stopped using it, he did not go back to cardiology to try another medication, nor did he go back to orthopedics, because of his pain, difficulty getting to the doctor for appointments, and his general discouragement and depression. Thus, his treatment plan was based on the idea that his care needed to be coordinated, not fragmented, and that regaining mobility would be key to resolving many of the other issues that Sam presented. The three interwoven issues that were targeted as initial goals thus became: (i) decrease depression; (ii) reduce joint pain; and (iii) improve cardiovascular performance. It was expected that improvements in these three areas would result in improved function across a variety of activities, including increased mobility, return to valued social and recreational activities, and improved sexual function with his wife. These functional changes were also considered important aspects of the treatment plan, and are woven into it. The frst two elements of the treatment plan are shown in more detail in Table 3. This also shows how the goals for the first two problems are related to specific intervention strategies, the disciplines working together to provide services, and the time frame for completing each intervention. The reader will also notice from Table 3 that the suggested interventions were linked closely to a jointly developed conceptualization of each problem.

After targeting Sam's main problems and agreeing on etiological models for each, the team identified goals that could be linked to time-specific interventions facilitated by team members. The first, central issue was addressing Sam's discouragement about treatment and alienation from health care providers. A plan was developed to provide Sam with more personalized and consistent care by establishing a stable staff person on the interdisciplinary team as his primary contact person. This staff member (the social worker in this case) could help him use his medical appointments more effectively and efficiently and provide support and encouragement to pursue more vigorous treatment of his medical problems. This intervention also improved the team's overall knowledge of Sam's medical and psychosocial status. The team believed that improving Sam's involvement with care and motivation to participate fully in care were crucial in meeting the second goal of reducing joint pain in his hips by rescheduling hip replacement surgery. In turn, reducing that joint pain was essential to work on increasing mobility.

The team viewed Sam's depression as heavily influenced by his progressively decreasing involvement in pleasurable activities due to chronic pain and decreased mobility. The team encouraged Sam to participate in brief psychotherapy with the team psychologist, with the further expectation that the psychologist and Sam would work with the occupational therapist on selecting activities and environments that would match his interests and degree of physical disability. As his physical situation improved, with surgery and medication, the level of activity could increase, but they could also develop strategies for increased pleasure, satisfaction, and sense of competence even before the surgery. This could help combat his discouragement and passive withdrawal from treatment, just as the closer relationship with the team did. Antidepressant medications were not recommended by the psychiatrist because of the strong situational components tied to his recent depression (i.e., decreased involvement in pleasant activities, chronic pain) and Sam's difficulty in managing his other medications.

Sam was initially surprised at the team's feedback that his feelings of depression were important and could be treated. He had expected someone in the clinic to tell him to "not be so emotional and just deal with it." After talking to the psychologist about the team's view of his depression, Sam attended four individual sessions with the psychologist and learned to identify the activities that were most closely linked with feelings of happiness and satisfaction. Sam and the therapists carefully planned ways to retain the daily activities most meaningful to him and to adapt activities, like gardening, that he had believed must be given up. He reported that, while still feeling sad at times, he was less tearful and more energetic. In addition, his appetite and sleep schedule returned to a more normal basis.

The team also arranged an appointment for Sam and his wife to meet with the geriatric medical resident and the psychologist. During this visit, Sam and Pauline were able to discuss their concerns about the impact of Sam's pain and physical functioning on sexual activity. They agreed to try different body positions and touching that resulted in less pressure on his affected joints. Sam also began to use a warm bath and pain medication just prior to making love and reported that this was helpful. After Sam's hip replacement surgery and improved function with cardiovascular treatment, the therapists helped Sam and his wife work out additional positions and timing for sexual activity and to continue communicating clearly

Table 3 Treatment plan for Sam P.

Problem area	*Treatment goal*	*Intervention*	*Disciplines involved*	*Timeframe*
(1) Depression (linked to discouragement about treatment, reduced pleasant activities, and loss of marital and sexual intimacy)	(1) Increase	(1) Offer following therapies:	(1)	(1)
	(i) Understanding of and confidence in treatment providers and potential treatment outcomes	(i) Consolidate care in geriatrics clinic; appoint a long-term staff member to meet with client every visit to assist in clarifying issues and questions for the medical resident; preclinic discussion of case with resident to ensure knowledge of treatment plan and consistent follow-up	(i) Social worker as long term staff contact; entire team involved	(i) Ongoing
	(ii) Pleasant activities	(ii) Brief psychotherapy using OT as consultant on adapting activities	(ii) Psychology and	(ii) 4 months
	(iii) Level of physical intimacy	(iii) Joint appointments with patient and wife for problem-solving re. sexual activity	(iii) Medicine and psychology	(iii) 6 months
(2) Joint pain, particularly in hips, knees, and hands (believed by the team to be primarily a result of osteoarthritis, exacerbated by lack of joint protection activities)	2(i) Decrease pain	2(i)(a) Reschedule hip surgery, with close communication between geriatrics clinic and orthopedics	2(i)(a) Medicine	2(i)(a) 1 month
		(i)(b) Increase consistent use of pain medication through education and increasing motivation	(i)(b) Pharmacy and psychology	(i)(b) 2 months
		(i)(c) Use hand splints to reduce pain and prevent deformity	(i)(c) OT	(i)(c) 2 months
	(ii) Prevent further problems	(ii) Patient and family education about arthritis self-management	(ii) Nursing and psychology	(ii) 3 months
	(iii) Reduce stress on joints in day to day life	(iii) Assist patient and wife make long range plans about housing if still needed after hip surgery; adapt gardening and daily tasks using joint protection techniques and assistive devices	(iii) Social work and OT	(iii) 6 months, if needed

OT, occupational therapy.

about their expectations and occasional problems.

Sam was able to keep better track of his medications through the suggestions and encouragement of the pharmacist and the nurse, especially after his depression began to resolve with treatment. They helped him understand the differences between different medications in his regimen and set up a procedure for organizing his weekly medication regimen. There were no further instances of taking too much medication, and Sam was better able to report to the geriatric resident on whether different medications were having their intended impact.

Ultimately, Sam did have hip replacement surgery, which resulted in a significant improvement in his pain and mobility. During his postsurgery recovery, the work on depression and maintaining a strong team relationship continued, resulting in a rapid recovery and reestablishment of a wide spectrum of activities, including a return to more social contact. After the surgery, Sam was able to stay in his two-storey house, so the social work intervention which was considered was not, in the end, necessary.

7.24.8 RESOURCES AND TRAINING MATERIALS

While the case of Sam P. demonstrated an interdisciplinary team at work in a hospital setting, this type of case formulation and treatment is important when working with older adults in a variety of mental and physical health care settings. This chapter has introduced the concepts and practice of interdisciplinary teamwork, with the expectation that more resources and experience will be needed for those wanting to apply this information to their own work groups. This section discusses additional written materials and training guides that may be useful.

The disciplines of industrial/organizational psychology, social psychology, and business management have traditionally devoted the most attention to studying group behavior, decision-making, and conflict resolution. However, a separate literature on interdisciplinary team functioning in health care settings has also developed. Some basic early work is covered in an annotated bibliography (Czirr & Rappaport, 1984). Early work will be mentioned, but the chapter will emphasize more recent publications and training approaches. While it is beyond the scope of this chapter to review these literatures in depth, the following section outlines several articles and books that provide a context for understanding interdisciplinary team issues.

7.24.8.1 Outcomes and Descriptions of Interdisciplinary Teams

The literature on team functioning in geriatrics is dominated by articles describing and evaluating interdisciplinary programs (e.g., Rubenstein et al., 1984; Zimmer, Eggert, & Schiverton, 1990) and discussing barriers to effective team functioning (e.g., Saltz, 1992; Chafetz, West, & Ebbs, 1988). Qualls and Czirr (1988) discuss how models for professional functioning differ across disciplines and influence team participation and collaboration. Goldstein (1989) looks at team functioning from the perspective of physicians and offers practical advice for developing team approaches in geriatric health care settings. Drinka and Streim (1994) provide case examples of teams functioning ineffectively and suggest intervention strategies to help teams escape from "health care purgatory." In his book *A Guide to Psychological Practice in Geriatric Long-term Care*, Lichtenberg (1994) does an excellent job advocating for an interdisciplinary approach in long-term care settings. The chapter on creating effective interdisciplinary teams is particularly useful for those needing more information on models for training nursing assistants, organizational interventions in long-term care settings, and specific examples of team-development approaches (including the Department of Veterans Affairs' Interdisciplinary Team Training Program). Lichtenberg also reviews the empirical literature on the relationship between team functioning and patient outcome, noting studies such as the work by Feiger and Schmitt (1979) that demonstrate a relationship between collegiality in interdisciplinary health teams and improved physical, social, and emotional functioning in patients.

7.24.8.2 Team Training Guides

There are a variety of "hands-on" training manuals for those interested in practical advice and examples of team development exercises. Some of these guides target work groups in business, while others are specific to health care settings. Several are particularly helpful. While not specific to health care environments, *The Team-Building Sourcebook* (Phillips & Elledge, 1989) is a comprehensive resource for practicing team consultants and others who are involved in team development. This guide discusses the steps of the team development process and provides numerous examples and activities to be used in addressing the specific needs of a given team. A strength of this resource text is the inclusion of agendas and workshop suggestions for a variety of tasks, giving specific amounts of time for team development activities. For example, the implementation section of the source book outlines the possible contents of team development workshops lasting from four hours to two days. Eleven modules are presented with step-by-step instructions for conducting team development sessions, complete with handouts and materials to be used in flip charts.

The Team Handbook: How to Use Teams to Improve Quality (Scholtes, 1988) is another excellent team-training guide. Although it was also developed with business consulting in mind, the handbook is relevant for health-care professionals. This work discusses quality improvement concepts, then shows how productive meetings and cooperative work groups are important components of increasing team efficacy. Scholtes addresses common problems in team settings (e.g., dominating or reluctant participants, feuding team members) and makes specific recommendations on how to work through these difficulties. The handbook also includes an entire chapter of team development activities and exercises, from handling introductions to helping team members make observations about group processes.

A Guide to Team Development was developed by an interdisciplinary team (Center for Interdisciplinary Education in Allied Health, 1980) at The University of Kentucky's Center for Interdisciplinary Education in Allied Health. This field manual was designed for faculties who train students in team development. It identifies tasks and exercises focused on organizing team meetings, goal setting, decision-making, generating an open and creative environment for team problem-solving, and evaluating the team process.

An assessment tool for evaluating interdisciplinary training efforts is The Interdisciplinary Education Perception Scale (IEPS) (Luecht, Madsen, Taugher, & Petterson, 1990). The IEPS is an attitude inventory that assesses perceptions of professionals, and has four factors: beliefs about competence and autonomy; perceived need for cooperation; perceptions of actual cooperation among different disciplines; and understanding others' values.

The Centre for the Advancement of Interprofessional Postgraduate Education has identified interprofessional training initiatives throughout the United Kingdom (CAIPE, 1989). The University of Newcastle Health Care Research Unit also developed a report that suggested practical steps for increasing collaboration among health care professionals. More broadly in Europe, an organization, the European Network for Development of Multiprofessional Education in Health Sciences (EMPE), has been developed to promote interprofessional education (Goble, 1994). This organization offers expert resources and opportunities for meeting to share educational plans, implementation strategies, and data. Resources for promoting interprofessional collaboration have also been developed in Finland (Lamska, Hietanen, & Lamsa, 1994) and Romania (Stephenson & McCreery, 1994).

7.24.8.3 Communication Training

There are occasions when a group has already accomplished many of the tasks needed to function as an interdisciplinary team, yet team members have difficulty resolving conflict or communicating clearly. Two books are recommended that specifically address conflict resolution and communication skills. Gottman, Notarius, Gonso, and Markman's *A Couple's Guide to Communication* (1976) is a classic in the psychotherapy and counseling literature for very good reasons. Although it targets communication difficulties in married couples, this work provides the rationale and description of exercises that can improve communication in a variety of contexts and settings. The first five chapters in the book, covering skills such as listening, validation, constructively expressing grievances, negotiating agreements, and handling hidden agendas, are especially helpful. A highly recommended book by Fisher, Ury, and Patton (1991), *Getting to YES: Negotiating Agreement Without Giving In*, is based on work done by the Harvard Negotiation Project. *Getting to YES* addresses both the conceptual orientation and practical implications of principled negotiation, a method of negotiation that can be used in a wide variety of circumstances. In principled negotiation (also called negotiation on the merits), negotiators are encouraged to separate the people from the problem, focus on interests instead of positions, generate a variety of alternatives before making decisions, and base results on some objective standard. These authors also discuss handling several common dilemmas in negotiations, including negotiating with someone who is more powerful and working with others who continue to rely on positional bargaining or unprincipled tactics.

7.24.8.4 Factors that Influence Team Functioning

Interdisciplinary teams are based on nonhierarchical interactions, but team members come from a larger social structure that often emphasizes rank and status. Work has been done in areas such as the behavior of low-ranking group members, and cultural and gender differences, as they influence the decision-making strategies of groups. Lichtenberg, Strzepek, and Zeiss (1990) discuss factors that limit the participation of psychiatric aides in treatment teams. They use a particular program, the Veterans Administration's Interdisciplinary Team Training in Geriatrics Program (ITTG; now known as the Interdisciplinary Team Training Program or ITTP), as an example of how to bring about change.

Mechanic (1962) discussed the attributes of the social structure of an organization that led to greater power and influence by lower ranking group members. When these concepts are applied to team members from traditionally lower-ranked disciplines (i.e., recreational therapists, nursing assistants, clerks), Mechanic's work has implications for team interventions. For example, teams can increase opportunities for these staff to interact and have access to other professionals, which will increase their influence in teams and improve the quality of team decisions. The team can also develop ways to recognize explicitly that individuals from low-ranking disciplines often have important expert knowledge that is not readily available to others in the team.

On a related topic, Kirchmeyer and Cohen (1992) discuss the concept of constructive conflict and then demonstrate empirically how specific behaviors in multicultural groups may increase the participation of ethnic minorities. This approach is described in more detail by Tjosvold (1991), who presents some actions that can be taken by designated group leaders to encourage the active involvement of all team members. Kelley and Streeter (1992) review issues and research on gender in organizations, covering such wide topics as trends in work patterns and perceived gender differences in performance and job attitudes.

7.24.9 HELPING WITH INTERDISCIPLINARY TEAM DEVELOPMENT

It should be clear that all interdisciplinary teams, no matter how motivated and talented the members, will have difficulty from time to time. The material presented thus far is intended to provide a theoretical background for how teams function, an example of good interdisciplinary team treatment planning, and information on resources where more can be learned about team functioning and essential skills for teamwork. It is also helpful to consider how to try to have a positive impact as a team member or how to help with program development.

First, there is a need to respect a fundamental tenet of interdisciplinary team development: by definition, no single person can make a group become an interdisciplinary team. Even with enthusiasm about interdisciplinary team ideas and a desire to see them put into action, no single person can make this occur, even if he or she has the institutional power to define program goals and organizational structure. Interdisciplinary teams ultimately depend on and are the embodiment of a powerful principle: people support what they help create. A group can only become an interdisciplinary team if the group members decide themselves that they want to take on the responsibilities of shared leadership and interdependent functioning. An individual's role can be one of bringing the concept to the table and trying to generate shared enthusiasm for it, but it cannot be mandated from above. If the group must function in a way mandated by authority, it is, by definition, not an interdisciplinary team.

If the group embraces the interdisciplinary concept, an individual can help create experiences that will move the group forward. It is usually essential early in the life of the team to generate a vision statement that articulates team members' understanding of program goals, including what patients are to be served, what resources are available to serve them, and the expected impact of the program. This can help the team negotiate the Storming process by providing a grounding in a commonly held set of values.

When the team has a shared sense of purpose, other training experiences largely mirror the organization of this chapter. Team members need to learn about interdisciplinary concepts and to confirm their commitment to shared responsibility for the success of the team. They need to recognize how much there is to know, for example, about each others' professions, about sharing leadership, and about communication skills and conflict resolution. The group optimally will decide together on strategies for appropriate additional training. Team members can train each other, or it may be necessary to bring in a team consultant. It will be especially important to bring in a consultant if the team cannot handle successfully the Storming process, either because the group avoids all conflict or because conflict escalates uncontrollably. There are systems level factors that encourage team members to ignore conflict. For example, factors such as low staff-to-patient ratios, lack of proximate office space that would allow informal staff contact, and few direct incentives for team process meetings

(Drinka, 1994) all point to the benefits of having a consultant assist in the process of team development.

In the training process, it is particularly helpful to find a balance between content and process. During process time, team members can share personal reactions to the tensions and pleasures of interdisciplinary work and discuss experiences with communication styles in the group. During content training, it is very helpful to have the group develop its own unique role map. This also provides an opportunity for each discipline to educate the others about its skills, training experiences, licensure status, and model of health care. The group can also look at different approaches to team treatment planning and choose one that best fits the team's mission and style of working.

Any team member can suggest such strategies and contribute to implementing the group's decisions. If the group does not develop effectively, or one individual seems to make all the suggestions or guide all the training experiences, bringing in a team consultant would be a good idea. The more one person tries to carry a reluctant team, the less interdisciplinary it will be, and the more that person will become the *de facto* leader of a multidisciplinary team. In such cases, the most helpful thing to be done is to relinquish power in order to encourage other team members to join in the process of assessing the team's problems and generating solutions.

7.24.10 SUMMARY

Interdisciplinary teams are a particularly appropriate way to organize health care for geriatric patients, who typically have complex, chronic problems. The interdisciplinary team provides care in which several disciplines coordinate assessment and treatment, so that problems can be dealt with consistently and comprehensively. Interdisciplinary teams are characterized by a nonhierarchical organization, in which responsibility for the effective functioning of the team is shared by all team members. This requires that team members be excellent representatives of and advocates for their disciplines. It also requires that team members have training in team theory, leadership skills, and communication skills, including conflict resolution skills. Interdisciplinary teams go through a process of development. Newly formed teams work out goals and basic models of function; this process usually highlights areas of disagreement among team members in strategies of care and/or personal interaction styles. As groups work out these disagreements, they develop shared norms which define the team's working strategies. With experience, most interdisciplinary teams reach a high level of function and deliver cost-effective care while supporting staff morale.

Interdisciplinary teams come about through the efforts of health care providers who are committed to the idea that patients will be best served when their care is coordinated and provided by team members who challenge, learn from, and rely on each other. Interdisciplinary teams depend on the wise and creative integration of diverse viewpoints. Sometimes those are complementary, but sometimes they are conflicting. The team must learn to value diversity, remain cohesive when viewpoints conflict, and negotiate agreement to which all team members are committed.

7.24.11 REFERENCES

Ahlmen, M., Sullivan, M., & Bjelle, A. (1988). Team versus non-team outpatient care in rheumatoid arthritis. *Arthritis and Rheumatism, 31,* 471–479.

Binstock, R. H., & George, L. K. (1990). *Handbook of aging and the social sciences.* San Diego, CA: Academic Press.

Birren, J. E., & Schaie, K. W. (1990). *Handbook of the psychology of aging.* San Diego, CA: Academic Press.

Center for Interdisciplinary Education in Allied Health (1980). *A guide to team development.* Lexington, KY: College of Allied Health Professions, University of Kentucky.

Centre for the Advancement of Interprofessional Postgraduate Education (CAIPE) (1989). A report of a national survey on interprofessional education in primary health care.

Chafetz, P., West, H., Ebbs, E. (1987). Overcoming obstacles to cooperation in interdisciplinary long term care teams. *Journal of Gerontological Social Work, 11,* 131–140.

Clark P. (1993). A typology of interdisciplinary education in gerontology and geriatrics: Are we really doing what we say we are? *Journal of Interprofessional Care, 7,* 217–227.

Czirr, R., & Rappaport, M. (1984). Toolkit for teams: Annotated bibliography on Interdisciplinary Health Teams. *Clinical Gerontologist, 2,* 47–54.

Dornbrand, L., Hoole, A. J., & Pickard, C. G. (1992). *Manual of clinical problems in adult ambulatory care.* Boston, MA: Little, Brown.

Drinka, T. J. K. (1994). Interdisciplinary geriatric teams: Approaches to conflict as indicators of potential to model teamwork. *Educational Gerontology, 20,* 87–103.

Drinka, T. J. K., & Streim, J. E. (1994). Case studies from purgatory: Maladaptive behavior within geriatric health care teams. *The Gerontologist, 34,* 541–547.

Federal Council on Aging, United States Department of Health and Human Services (1981). The need for long term care: Information and issues. *DHHS Publication (OHDS) 81-20704.* Washington, DC: United States Government Printing Office.

Feiger, S. M., & Schmitt, M. H. (1979) Collegiality in interdisciplinary health teams: Its measurement and its effects. *Social Science Medicine, 13,* 217–229.

Fisher, R., Ury, W., & Patton, B. (1991). *Getting to YES: Negotiating agreement without giving in.* New York: Penguin.

Fries, J. F., & Crapo, L. M. (1981). *Vitality and aging: Implications of the rectangular curve*. San Francisco: Freeman.

Gersick, C. J. (1989). Marking time: Predictable transitions in task groups. *Academyof Management Journal, 32,* 274–309.

Goble, R. (1994). Multiprofessional education: European Network for Development of Multiprofessional Education in Health Sciences (EMPE). *Journal of Interprofessional Care, 8,* 85–92.

Goldstein, M. (1989). Physicians and teams. In R. J. Ham (Ed.) *Geriatric medicine annual* (pp. 256–275), Oradell, NJ: Medical Economics Books.

Gottman, J., Notarius, C., Gonso, J., & Markman, H. (1976). *A couple's guide to communication*. Champaign, IL: Research Press.

Heinemann, G. D., Farrell, M. P., & Schmitt, M. H. (1994). Groupthink theory and research: Implications for decision making in geriatric health care teams. *Educational Gerontology, 20,* 71–85.

Hendriksen, C., Lund, E., & Stromgard, E. (1984). Consequences of assessment and intervention among elderly people: A three year randomised controlled trial. *British Medical Journal, 289,* 1522–1524.

Hickey, T., & Rakowski, W. (1981). Consistency in patients' health and treatment expectations at a Geriatric Clinic. *Journal of the American Geriatrics Society, 29,* 278–284.

Jones, R. V. H. (1992). Teamwork in primary care: How much do we know about it? *Journal of Interprofessional Care, 6,* 25–2 9.

Kelley, K., & Streeter, D. (1992). The roles of gender in organizations. In K. Kelley (Ed.), *Issues, Theory, and Research in Industrial/Organizational Psychology* (pp. 285–337). Oxford, UK: Elsevier.

Kirchmeyer, C., & Cohen, A. (1992). Multicultural groups: Their performance and reactions with constructive conflict. *Group & Organization Management, 17,* 153–170.

Lamska, A., Hietanen, I., & Lamsa, J. (1994). Education for holistic care: A pilot programme in Finland. *Journal of Interprofessional Care, 8,* 31–43.

Lichtenberg, P. (1994). *A guide to psychological practice in geriatric long-term care*. New York: Haworth Press.

Lichtenberg, P., Strzepek, D., & Zeiss, A. (1990). Bringing psychiatric aides into the treatment team: An application of the Veterans Administration's ITTG model. *Gerontology and Geriatrics Education, 10,* 63–73.

Luecht, R. M., Madsen, M. K., Taugher, M. P., & Petterson, B. J. (1990). Assessing professional perceptions: Design and validation of an interdisciplinary education perception scale. *Journal of Allied Health, 19,* 181–191.

Mechanic, D. (1962). Sources of power of lower participants in complex organizations. *Administrative Science Quarterly, 7,* 349–364.

Mitchell, M. (1993). Models of therapeutic practice amongst first year students: A problem for educators in the health professions? *Journal of Interprofessional Care, 7,* 247–252.

Phillips, S., & Elledge, R. (1989). *The Team-Building Sourcebook*. San Diego, CA: University Associates.

Qualls, S., & Czirr, R. (1988). Geriatric health teams: Classifying models of professional and team functioning. *The Gerontologist, 28,* 372–376.

Rubenstein, L., Wieland, D., English, P., Josephson, K., Sayre, J., & Abrass, I. (1984). The Sepulveda VA Geriatric Evaluation Unit: Data on four–year outcomes and predictors of improved patient outcomes. *Journal of the American Geriatrics Society, 32,* 503–512.

Saltz, C. (1992). The interdisciplinary team in geriatric rehabilitation. Special Issue: Geriatric Social Work Education. *Journal of Gerontological Social Work, 18,* 133–142.

Sbordone, R. J., & Sherman, L. T. (1983). The Psychologist as a consultant in a nursing home: Effect on staff morale and turnover. *Professional Psychology: Research and Practice, 14,* 240–250.

Schneider, E. L., & Rowe, J. W. (1990). *Handbook of the biology of aging*. San Diego, CA: Academic Press.

Scholtes, P. R. (1988). *The Team Handbook: How to Use Teams to Improve Quality*. Madison, WI: Joiner Associates.

Stephenson, P., & McCreery, R. (1994). Two heads are better than one: Interprofessional collaboration for capacity building in health systems research in Romania. *Journal of Interprofessional Care, 8,* 57–61.

Takamura, J. (1985). Introduction—Health teams. In L. J. Campbell & S. Vivell (Eds.), *Interdisciplinaryteam training for primary care in geriatrics: An educational model forprogram development and evaluation* (pp. II.64—II.67). Washington, DC: United States Government Printing Office.

Takamura, J., Bermost, L., & Stringfellow, L. (1979). *Health team development team*. University of Hawaii, HI: John A. Burns School of Medicine.

Tjosvold, D. (1991). *The conflict-positive organization: Stimulate diversity and create unity*. Reading, MA: Addison-Wesley.

Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological Bulletin, 64,* 384–399.

West, M. A. & Poulton, B. C. (1997). A failure of function: Teamwork in primary health care. *Journal of Interprofessional Care, 11,* 205–216.

Woolf, S. H., Kamerow, D. B., Lawrence, R. S., Medalie, J. H., & Estes, E. H. (1990). The periodic health examination of older adults: The recommendations of the US Preventive Services Task Force. *Journal of the American Geriatrics Society, 38,* 817–823.

Zawadski, R. T., & Eng, C. (1988). Case management in capitated long-term care. *Health Care Financing Review,* Annual Suppl., 75–81.

Zeiss, A. M., & Okarma, T. (1985). Effects of interdisciplinary health care teams on elderly arthritic patients: A randomized prospective trial of alternative team models. *Proceedings of the Sixth Annual Conferenceon Interdisciplinary Health Team Care.* (pp. 21–35). Hartford, CN: University of Connecticut School of Allied Health Professions, Nursing, Pharmacy, & Medicine.

Zimmer, J., Eggert, G., & Schiverton, P. (1990). Individual versus team case management in optimizing community care for chronically ill patients with dementia. *Journal of Aging and Health, 2,* 357–372.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.01
Health Behavior

MARK CONNER
University of Leeds, UK

and

PAUL NORMAN
University of Sheffield, UK

8.01.1 INTRODUCTION 2

8.01.2 KEY BEHAVIORS INFLUENCING HEALTH, MORBIDITY, AND MORTALITY 2
- *8.01.2.1 Introduction* 2
- *8.01.2.2 Smoking* 3
- *8.01.2.3 Diet* 3
- *8.01.2.4 Exercise* 3
- *8.01.2.5 Other Behaviors* 4

8.01.3 UNDERSTANDING THE BASIS OF HEALTH BEHAVIOR 5
- *8.01.3.1 Introduction* 5
- *8.01.3.2 Health Belief Model* 6
  - *8.01.3.2.1 Model description* 6
  - *8.01.3.2.2 Review of research* 7
  - *8.01.3.2.3 Commentary* 8
- *8.01.3.3 Theory of Planned Behavior* 8
  - *8.01.3.3.1 Model description* 8
  - *8.01.3.3.2 Review of research* 10
  - *8.01.3.3.3 Commentary* 11
- *8.01.3.4 Health Locus of Control* 12
  - *8.01.3.4.1 Model description* 12
  - *8.01.3.4.2 Review of research* 12
  - *8.01.3.4.3 Commentary* 13
- *8.01.3.5 Protection Motivation Theory* 14
  - *8.01.3.5.1 Model description* 14
  - *8.01.3.5.2 Review of research* 15
  - *8.01.3.5.3 Commentary* 17
- *8.01.3.6 Self-efficacy* 17
  - *8.01.3.6.1 Model description* 17
  - *8.01.3.6.2 Review of research* 18
  - *8.01.3.6.3 Commentary* 18

8.01.4 ENCOURAGING THE ADOPTION OF HEALTH BEHAVIORS 19
- *8.01.4.1 Intervention Studies* 19

8.01.5 LIMITATIONS AND EXTENSIONS OF EXISTING MODELS 21
- *8.01.5.1 Limitations of Current Models* 21
- *8.01.5.2 Additional Theoretical Constructs* 22
- *8.01.5.3 Processes by which Cognitions Influence Behavior* 24
- *8.01.5.4 Stage Models of Health Behavior* 24
- *8.01.5.5 Theoretical Integrations and Future Directions* 27

8.01.6 CONCLUSIONS 28
8.01.7 REFERENCES 28

8.01.1 INTRODUCTION

Interest in health behaviors is derived from two assumptions; that a substantial proportion of the mortality from the leading causes of death is attributable to the behavior of individuals, and that the behavior is modifiable (Stroebe & Stroebe, 1995). It is now widely recognized that individuals can influence their own health and well-being through the adoption of health-enhancing behaviors (e.g., exercise) and the avoidance of health-compromising behaviors (e.g., smoking). The identification of the factors which predict who engages in health behaviors has become a focus of research in health psychology and other health-related disciplines in recent years (e.g., Adler & Matthews, 1994; Conner & Norman, 1996a; Glanz, Lewis & Rimmer, 1990; Hockbaum & Lorig, 1992; Rodin & Salovey, 1989; Winett, 1985).

The health behaviors examined have been many and varied; from health enhancing behaviors such as exercise and healthy eating, on the one hand, to avoidance of health harming behaviors such as smoking and excessive alcohol consumption, on the other. Each of these behaviors have immediate or long-term effects upon the individual's health and are to varying extents within the individual's control. Epidemiological studies reveal great variability in who performs these behaviors. The approaches taken to understanding such individual differences have been equally varied. One can classify these approaches into those which examine factors intrinsic to the individual (e.g., sociodemographic factors, personality, social support, cognitions) and those which examine factors extrinsic to the individual. This second group of approaches can be further divided into those based on incentive structures (e.g., taxing tobacco and alcohol, subsidizing sports facilities) and those based on legal restrictions (e.g., banning dangerous substances, financial penalties for not wearing seat-belts). The first approach (factors intrinsic to the individual) has received most attention from psychologists, with a particular focus on cognitive factors as the most important proximal determinants of behavior. A variety of models of how such cognitive factors produce various "social" behaviors have been developed and are commonly referred to as social cognition models (SCMs). SCMs are recognized to have made a valuable contribution to the understanding of both who performs health behaviors (Conner & Norman, 1996b; Marteau, 1989) and how extrinsic factors produce behavior change (e.g., Rutter & Quine, 1996).

This chapter examines the key health behaviors and ways in which their adoption might be encouraged. The chapter then considers the contribution psychology has made to understanding and changing health behaviors through the development of SCMs. The most widely applied SCMs are described and reviewed along with recent developments and future prospects for this field of research.

8.01.2 KEY BEHAVIORS INFLUENCING HEALTH, MORBIDITY, AND MORTALITY

8.01.2.1 Introduction

We might define health behavior as any activity taken for the purpose of preventing or detecting disease or for improving general well-being (Conner & Norman, 1996b). The behaviors within this definition include medical service usage (e.g., physician visits, vaccination, screening), compliance with medical regimens (e.g., dietary, diabetic, antihypertensive regimens), and self-directed health behaviors (e.g., diet, exercise, smoking, alcohol consumption).

Numerous studies have examined the relationship between health behaviors and health outcomes (e.g., Black Report, 1988; Blane, Smith, & Bartley, 1990; Blaxter, 1990; Cox, Huppert, & Whichelow, 1993; Doll, Peto, Wheatley, Gray, & Sutherland, 1994; Gottlieb & Green, 1984). Such studies demonstrate the role of a variety of behaviors for both morbidity and mortality. One of the first such studies in Alameda County identified seven features of lifestyle: not smoking, moderate alcohol intake, sleeping 7–8 hours per night, exercising regularly, maintaining a desirable body weight, avoiding snacks, and eating breakfast regularly which together were associated with lower morbidity and higher subsequent long-term survival (Belloc, 1973; Belloc & Breslow, 1972; Breslow & Enstrom, 1980). Such results have been replicated in a variety of different populations (e.g., Brock, Haefner, & Noble, 1988; Metzner, Carman, & House, 1983). The impact of health behaviors upon individuals' quality of life, via delaying the onset of chronic disease and extending active life span, should also be noted (Fries, Green, & Levine, 1989; Stroebe & Stroebe, 1995). Smoking, alcohol consumption, diet, gaps in primary care services and low screening uptake are all significant determinants of poor health (Amler & Dull, 1987). Such findings have led to a focus by those interested in health promotion on changing such behaviors in order to improve health. For example, in the USA, Healthy People 2000 (USDHHS, 1990) lists increased physical activity, changes in nutrition and reductions in tobacco, alcohol, and drug use as important for health promotion and disease prevention.

Below we examine several health behaviors in more detail, focusing on their prevalence and relationship to health outcomes.

8.01.2.2 Smoking

Smoking is the behavior most closely linked with long-term negative health outcomes. Both morbidity and mortality from coronary heart disease (CHD) are increased among smokers (Doll et al., 1994; Friedman, Dales, & Ury, 1979). Moreover, there is a strong positive relationship between the number of cigarettes smoked per day and the incidence of CHD (Friedman et al., 1979). In addition, smoking has been linked to a number of cancers including cancer of the lung, throat, stomach, and bowel as well as a number of more immediate negative health effects such as reduced lung capacity and bronchitis (Royal College of Physicians, 1983). Despite the array of negative health outcomes, smokers often report positive mood effects from smoking and the use of smoking as a coping strategy (Graham, 1987).

The number of people smoking in the USA and UK has shown a steady decline over the past 20 years. Data from the General Household Survey (1992) showed that 28% of people over the age of 16 smoke in the UK. Smoking is more common among men and among unskilled manual workers (General Household Survey, 1994). A similar pattern is evident in the USA, with smoking more common among less educated, lower income, and minority groups (Rigotti, 1989). Those who quit smoking reduce the risk to their health, particularly if they quit before 35 years of age (Doll et al., 1994).

8.01.2.3 Diet

The impact of various aspects of diet upon health, morbidity, and mortality are well established (USDHHS, 1988). Whilst in the Third World the problems related to diet and health are ones of undernutrition, in the First World, the problems are predominantly linked to overconsumption of food. In Western industrialized countries the major problems are linked to excessive fat consumption and insufficient fiber, fruit, and vegetable consumption. In addition, excess consumption of calories combined with insufficient exercise has made obesity a major health problem. Diet has been implicated in cardiovascular diseases (CVDs), strokes and high blood pressure, cancer, diabetes, obesity, osteoporosis, and dental disease.

It is generally agreed that elevated blood cholesterol level is a major risk factor for the development of CVD (Consensus Development Conference on Lowering Blood Cholesterol to Prevent Heart Disease, 1985). Nutbeam and Catford (1990) estimate that 26% of men and 25% of women in the UK have cholesterol levels greater than 6.5 mmol l^{-1} (a level considered to be excessive). While in the USA, it is estimated that 50% of the adult population is at risk of CHD by virtue of elevated blood cholesterol levels (Sampos, Fulwood, Haines et al., 1989). International studies have clearly demonstrated an association between saturated fat consumption (one source of cholesterol) and blood cholesterol levels. For example, Keys (1970) reported a correlation of 0.89 across the seven countries studied. The reduction of blood cholesterol via dietary change is now widely accepted as an important way of tackling CHD (Expert Panel, 1993). Drug treatment to reduce blood cholesterol levels is generally seen as only advisable if dietary change is ineffective (Wardle, 1995). Dietary recommendations include reducing fat in the diet and increasing soluble fiber intake (Committee on Medical Aspects of Food Policy, 1991; Expert Panel, 1993). However, their impact upon cholesterol levels may be limited.

8.01.2.4 Exercise

Engaging in regular exercise is seen to be another key component of a healthy lifestyle. The potential health benefits of engaging in regular exercise are many and include reduced cardiovascular morbidity and mortality (Oberman, 1985), lowered blood pressure (Blair, Goodyear, Gibbons, & Cooper, 1984) and the increased metabolism of carbohydrates (Lennon et al., 1983) and fats (Rosenthal, Haskell, Solomon, Widstrom, & Reavan, 1983), as well as a range of psychological benefits such as improved self-esteem (Sachs, 1984), positive mood states (Folkins & Sime, 1981), reduced life stress (Brown, 1991), and reduced levels of anxiety (Singer, 1992).

However, despite the various health benefits of exercise, a significant proportion of the population lead a sedentary lifestyle. For example, the General Household Survey (1989) indicated that only one in three men and one in five women in the UK participate in any sport or recreational physical activity. Moreover, the Allied Dunbar Fitness Survey (1992) of 6000 English adults reported that one in six adults had done no exercise (i.e., for 20 minutes or more at a moderate or vigorous level) in the previous four weeks. Participation in regular exercise is strongly related to a number of socio-demographic variables. In particular, young people and males are more likely to engage in regular exercise. For example, the 1988 Welsh Heart Health Survey (Health Promotion Authority for Wales, 1990) reported that among 18–34 year olds, 61% of men engaged in moderately vigorous exercise at least two times a week compared with only 35% of women. For 35–64 year olds, the percentages drop to 37% for men and 17% for women. Overall, the typical exerciser is likely to be young, well-educated, affluent, and male (King et al., 1992).

8.01.2.5 Other Behaviors

A number of other behaviors show clear links to health. Below we briefly consider recreational drug use, safe sex, alcohol use, and health screening as behaviors with important health consequences.

The use of recreational drugs has long been recognized as a potential health problem which needs to be tackled (Aguirre-Molina & Gorman, 1996). The most commonly used drugs are alcohol and tobacco, however, here we consider what are usually referred to as psychoactive drugs (i.e., a drug that alters mood or behavior). There are a large number of such drugs including hypnotic drugs such as barbiturates which reduce anxiety and produce sedation; stimulants such as amphetamines which elevate mood, increase wakefulness, and give an enhanced sense of mental and physical energy; opiates such as opium and heroin which produce pleasant mood states; antipsychotic agents such as chlorpromazine which diminish the symptoms of psychoses; and psychedelics/hallucinogens such as LSD, cannabis, and MDMA (3,4-methylenedioxymethamphetamine) which cause visual and auditory hallucinations. Most of these drugs have medical uses as well as being used recreationally.

Data from various drug use surveys in the UK (e.g., Leitner, Shapland, & Wiles, 1993) suggest that opiate use has remained stable (at around 1%), but the use of cannabis and other nonopiates has been on the increase over the last 25 years. The 1968 OPCS survey reported that 2% of the 16 years and over age group had used cannabis at some stage, while the Leitner et al. (1993) study reported that 10% of the same age group had used it. In general, drug use in the UK does not differ substantially from the rest of the Western world. Reported lifetime usage in the USA is slightly higher than in the UK, although estimates for injecting use (1–2% of the general population), alcohol use (around 85%), and lifetime tobacco use (around 75%) are roughly similar.

Sexual behaviors have also long been considered health behaviors because of their impact upon the spread of sexually transmitted diseases (STDs) such as gonorrhoea and syphilis. More recently, the role of sexual behaviors in the spread of the human immunodeficiency virus (HIV) has been a focus of attention (O'Leary & Raffaelli, 1996). Whilst early health education campaigns emphasized the need to reduce the number of sexual partners or avoid particular sexual practices (e.g., anal sex, penetrative sex), more recently the focus has been upon the use of condoms during penetrative sex to reduce the risk of HIV transmission (Reiss & Leik, 1989). Condom use is particularly recommended for those with multiple partners or those who do not know their partners' sexual history. For these reasons, much of the health advice concerning condom use has been focused on young people.

There seems to be considerable variation in the use of condoms in response to the threat of HIV/AIDS. For example, among heterosexuals, Richard and van der Pligt (1991) reported that 50% of their sample of Dutch teenagers with multiple partners consistently used condoms. While other studies in the UK and USA report rates of between 24% and 58% (Fife-Schaw & Breakwell, 1992; Gerrard, Gibbons, & Bushman, 1996). Among homosexuals, Weatherburn, Hunt, Davies, Coxon, and McManus (1991) reported that 39% of their sample always used a condom during anal sex. Whilst among bisexuals, Boulton, Schram Evans, Fitzpatrick, and Hart (1991) report that 25% of their male sample used condoms with their current male partner and only 12% with their current female partner. The General Household Survey (1993) in the UK reported changes by age group in the use of condoms for the period 1983–1991. Among 16–24-year-olds, condom use increase from around 6% to around 12% during this period, whilst among 40–49-year-olds it dropped from around 18% to around 12% over the same period. Thus, health messages aimed at increasing condom use may be having some impact among younger people, although the overall rates of use are still worrying low.

High alcohol consumption has been linked to a range of negative health outcomes including high blood pressure (Shaper et al. 1981), heart disease (Sherlock, 1982), and cirrhosis of the liver (Colliver & Malin, 1986), although there is some evidence to suggest that low levels of alcohol consumption may have slightly beneficial effects on health (Hennekens, 1983). High levels of alcohol consumption have also been associated with accidents, injuries, suicides, crime, domestic violence, rape, murder, and unsafe sex (*British Medical Journal*, 1982). While many of the adverse effects of high alcohol consumption are due to continued heavy drinking (e.g., cirrhosis of the liver, heart disease), others are more specifically related to excessive alcohol consumption in a single drinking session (e.g., accidents, violence) (Honkanen et al., 1983).

The General Household Survey (1992) reported that the average weekly consumption of alcohol in the UK was 15.9 units (approximately 8 pints of beer) for men and 5.4 (approximately 2.5 pints of beer) for women. Of more interest was the finding that about 27% of men and 11% of women were drinking more than the recommended weekly sensible limits (21 units for men, 14 units for women). Heavy drinking is also more likely among younger age groups. In a survey of 12 000 Welsh adults, Moore, Smith, and Catford (1994) reported that 31.1% of drinkers aged 18–24 engaged in binge drinking (i.e., drinking half the recommended weekly consumption of alcohol in a single session) at least once a week.

Finally, individuals may seek to protect their health by participating in various screening programs which attempt to detect disease at an early, or asymptomatic, stage. In the UK, screening programs have been set up for various diseases including anemia (Ashworth, 1963), diabetes (Redhead, 1960), bronchitis (Gregg, 1966), cervical cancer (Freeling, 1965), and breast cancer (Forrest, 1986). Considering breast cancer, it has been estimated that breast screening programs which include mammograms can reduce breast cancer mortality by up to 40% among women aged 50 and over (Strax, 1984). However, participation rates in breast screening programs show great variability across different countries, ranging from 25% to 89% (Vernon, Laville, & Jackson, 1990). Participation tends to be negatively related to age and positively related to education level and socioeconomic status (Vernon et al., 1990).

8.01.3 UNDERSTANDING THE BASIS OF HEALTH BEHAVIOR

8.01.3.1 Introduction

A clearer understanding of why individuals perform health behaviors might assist in the development of interventions to help individuals gain the benefits of improved health and well-being. A variety of factors have been found to account for individual differences in the performance of various health behaviors, including demographic factors, social factors, emotional factors, perceived symptoms, factors relating to access to medical care, personality factors, and cognitive factors (Adler & Matthews, 1994; Rosenstock, 1974; Taylor, 1991).

Demographic variables show reliable associations with the performance of health behaviors. For example, there is a curvilinear relationship between many health behaviors and age, with high incidences of many health-risking behaviors such as smoking in young adults and much lower incidences in children and older adults (Blaxter, 1990). Such behaviors also vary by gender, with females being generally less likely to smoke, consume large amounts of alcohol, engage in regular exercise but more likely to monitor their diet, take vitamins, and engage in in dental care (Waldron, 1988). Differences by socioeconomic status and ethnic group are also apparent for behaviors such as diet, exercise, alcohol consumption, and smoking (e.g., Blaxter, 1990). Generally speaking, younger, wealthier, better educated individuals, under low levels of stress, with high levels of social support, are more likely to practice health enhancing behaviors. Higher levels of stress and/or fewer resources are associated with health-compromising behaviors such as smoking and alcohol abuse (Adler & Matthews, 1994).

Social factors seem to be important in instilling health behaviors in childhood. Parent, sibling, and peer influences are important, for example, in the initiation of smoking (e.g., McNeil et al., 1988). Cultural values also have a major impact, for instance in determining the number of women exercising in a particular culture (e.g., Steptoe & Wardle, 1996). For example, Steptoe and Wardle (1992) report that between 34% and 95% of women in their European student sample had exercised in the past 14 days. Perceived symptoms control health habits when, for example, smokers regulate their smoking on the basis of sensations in the throat. Access to medical care has been found to influence the use of such health services (e.g., Black Report, 1988). Personality factors have also been associated with health behaviors (Adler & Matthews, 1994; Steptoe et al., 1994).

Cognitive factors also determine whether or not an individual practices health behaviors. Knowledge about behavior–health links is an important factor in an informed choice concerning a healthy lifestyle. Various other cognitive variables have been studied including perceptions of health risk, efficacy of behaviors in influencing this risk, social pressures to perform the behavior, and control over performance of the behavior. The relative importance of various cognitive factors in determining who performs various health behaviors constitutes the basis of different models. Such models have been labeled SCMs because of their focus on cognitive variables as the primary determinant of individual social behaviors.

Two types of SCMs have been applied in health psychology, predominantly to explain health-related behaviors and response to treatment (Conner, 1993). The first type focus on individuals' understanding of the causes of health-related events and are best typified by attribution models (e.g., King, 1982). The second type are more diverse in nature and attempt to predict future health-related behaviors and outcomes. These include the health belief model (HBM; e.g., Becker, 1974; Janz & Becker, 1984; Sheeran & Abraham, 1996), health locus of control (HLOC; Norman & Bennett, 1996; Seeman & Seeman, 1983; Wallston, Wallston, & De Vellis, 1978), protection motivation theory (PMT; e.g., Boer & Seydel, 1996; Rogers, 1983; van der Velde & van der Pligt, 1991), theory of reasoned action/theory of planned behavior (TRA/TPB; e.g., Ajzen, 1988; 1991; Ajzen & Fishbein, 1980; Conner & Sparks, 1996), and self-efficacy (SE; e.g., Bandura, 1982, 1991; Schwarzer, 1992; Schwarzer & Fuchs, 1996). Other models include self-regulation theory (Leventhal, Nerenz, & Steele, 1984), the transtheoretical model of change (Prochaska & DiClemente, 1984), the precaution-adoption process (Weinstein, 1988), and the model of goal achievement (Bagozzi, 1992). However, none of these latter models have been widely applied to the prediction of health behaviors at present.

These social cognition models provide a basis for understanding the determinants of behavior and behavior change. Each of these models emphasize the rationality of human behavior, although they do not assume that all behavior is based upon careful thought (Ajzen, 1996). Most assume that behavior and decisions are based upon elaborate, but subjective, cost/benefit analysis of the likely outcomes of differing courses of action. As such they have roots going back to expectancy-value theory (Peak, 1955) and subjective expected utility theory (SEU; Edwards, 1954). It is assumed that individuals generally aim to maximize utility and so prefer behaviors which are associated with the highest expected utility (Van der Pligt & de Vries, 1998). Whilst such considerations may well provide good predictions of which behaviors are selected, it has been noted by several authors that they do not necessarily provide an adequate description of the way in which individuals make decisions (e.g., Edwards, 1992; Feather, 1982; Frisch & Clemen, 1994; Jonas, 1993).

In the section which follows, the most widely used of these models (HBM, TPB, HLOC, PMT, SE) are outlined, the research using them described, and their use reviewed.

8.01.3.2 Health Belief Model

8.01.3.2.1 Model description

The HBM is probably the most widely used social cognition model in health psychology (Becker, 1974; Rosenstock, 1966; Sheeran & Abraham, 1996). It was originally developed by US public health researchers attempting to develop models upon which to base health education programs (Hochbaum, 1958; Rosenstock, 1966). The model attempts to conceptualize the health beliefs which make a behavior more or less attractive. In particular, the key health beliefs were seen to be the likelihood of experiencing a health problem, the severity of the consequences of the health problem, and the perceived costs and benefits of the health behavior. Thus, the HBM employs two aspects of individuals' representations of health behavior in response to threat of illness: perceptions of the threat of illness and evaluation of the effectiveness of behaviors to counteract this threat (see Figure 1). Threat perceptions depend upon two beliefs: the perceived susceptibility to the illness and the perceived severity of the consequences of the illness. Together these two variables determine the likelihood of the individual following a health-related action, although their effect is modified by individual differences in demographic variables, social pressure, and personality. The particular action taken is determined by evaluation of the possible alternatives. This behavioral evaluation depends upon beliefs concerning the benefits or efficacy of the health behavior and the perceived costs or barriers to performing the behavior. Hence, individuals are likely to follow a particular health action if they believe themselves to be susceptible to a particular condition or illness which they consider to be serious, and believe the benefits of the action taken to counteract the condition or illness outweigh the costs. Cues to action and health motivation are two other variables commonly included in the model. Cues to action include a diverse range of triggers to the individual taking action and are commonly divided into factors which are internal (e.g., physical symptom) or external (e.g., mass media campaign, advice from others such as physicians) to the individual (Janz & Becker, 1984).

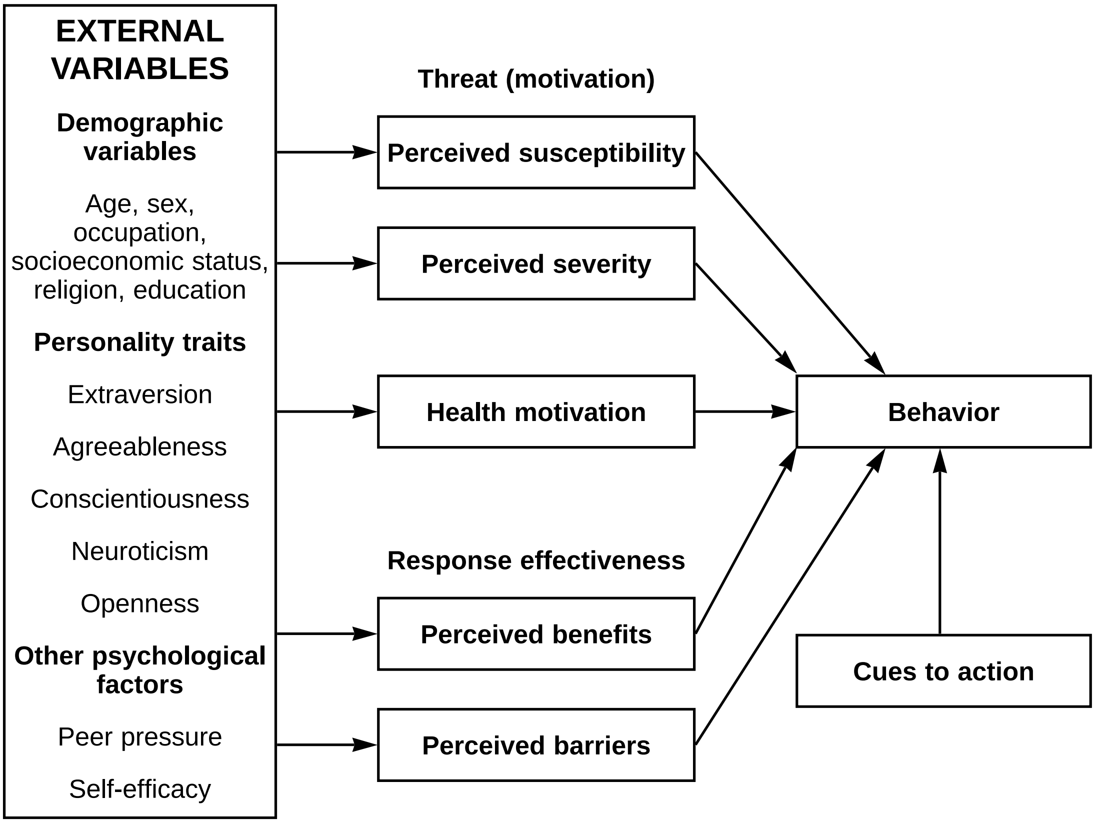
EXTERNAL VARIABLES
Demographic variables
Age, sex, occupation, socioeconomic status, religion, education
Personality traits
Extraversion
Agreeableness
Conscientiousness
Neuroticism
Openness
Other psychological factors
Peer pressure
Self-efficacy
Threat (motivation)
Perceived susceptibility
Perceived severity
Health motivation
Response effectiveness
Perceived benefits
Perceived barriers
Behavior
Cues to action

Figure 1 Health belief model.

Becker (1974) has argued that the HBM should also contain a measure of health motivation (readiness to be concerned about health matters) because certain individuals may be predisposed to respond to cues to action because of the value they place on their health. Other influences upon the performance of health behaviors, such as demographic factors or psychological characteristics (e.g., personality, peer pressure, perceived control over behavior), are assumed to exert their effect via changes in the six components of the HBM. This is a potentially important issue if the HBM is to claim to be a complete model of health behavior. However, this has not been widely addressed in empirical studies and where it has the evidence has been equivocal. Orbell, Crombie, and Johnson (1995), for example, reported that HBM components did mediate the effects of social class upon uptake of cervical screening, but did not mediate the effects of marital status or sexual experience.

8.01.3.2.2 Review of research

The HBM has been applied to a very broad range of health behaviors and populations. Sheeran and Abraham (1996) distinguish three broad areas of research. First, the HBM has been applied to various preventive health behaviors. These include health-risk behaviors such as smoking (Gianetti, Reynolds, & Rihen, 1985; Stacey & Lloyd, 1990) and alcohol use (K. H. Beck, 1981; Gottlieb & Baker, 1986), as well as health-promoting behaviors such as diet (Aho, 1979), exercise (Langlie, 1977), genetic (Becker, Kaback, Rosenstock, & Ruth, 1975) and health screening (Conner & Norman, 1994; King, 1982), vaccination (Oliver & Berger, 1979), breast self-examination (Champion, 1984; Ronis & Harel, 1989), contraceptive use (Hester & Macrina, 1985), and dental behaviors (Chen & Land, 1986). A second area the HBM has been applied to is various sick role behaviors which refer to compliance with professionally recommended medical regimens in response to illness. These include compliance with antihypertensive regimens (Taylor, 1979), diabetic regimens (Harris & Lynn, 1985), and renal disease regimens (Hartman & Becker, 1978), and regimens adhered to by parents for a child's condition (Becker, Radius, & Eveland, 1978). Third, the HBM has been applied to clinic use, which includes physician visits for a variety of reasons including preventative (Aiken, West, Woodward, Reno, & Reynolds, 1994), psychiatric (Connelly, 1984), and parent and child conditions (Kirscht, Becker, & Eveland, 1976). There is no strong evidence that the HBM has been more predictive of behavior in any one of these behaviors compared with any other (Sheeran & Abraham, 1996).

There have been two quantitative reviews of research with the HBM (Harrison, Mullen, & Green, 1992; Janz & Becker, 1984). The first, conducted by Janz and Becker (1984), examined the proportion of times each of the HBM's components showed a significant relationship with health behavior through the use of a "significance ratio." Across the 46 studies reviewed, the barriers component was found to have the most consistent relationship with health behavior (89%), followed by the susceptibility (81%), benefits (78%), and severity (65%) components. However, while Janz and Becker (1984) suggest that the HBM components are consistent predictors of health behavior, they fail to estimate the strength of the relationships. This question was addressed in the second quantitative review, conducted by Harrison et al. (1992). Over 200 published studies on the HBM were identified although only 16 of these were found to measure each of the components adequately. Harrison et al's (1992) meta-analysis on these 16 studies produced a similar pattern of results to Janz and Becker's (1984) earlier review with the barriers components having the highest average correlation with health behavior ($r = -0.21$), followed by the susceptibility ($r = 0.15$), benefits ($r = 0.13$), and severity ($r = 0.08$) components. The predictive power of individual components is therefore relatively modest, accounting for only 0.5–4% of variance in behavior. However, it should be noted that it is the combined effects of the six health beliefs which is generally of interest and this is commonly in excess of the sum of the effects of the individual components.

8.01.3.2.3 Commentary

The HBM has provided a useful framework for investigating health behaviors and identifying key health beliefs, has been widely used, and has met with moderate success in predicting a range of health behaviors (for reviews see Harrison et al., 1992; Janz & Becker, 1984; Sheeran & Abraham, 1996). The strength of the HBM lies in the fact that it was developed by researchers working directly with health behaviors and so many of the concepts possess face-validity to those working in this area. This commonsense operationalization of a number of cognitive variables relevant to the performance of health behavior partly account for the model's popularity.

However, compared to other similar social cognitive models of health behaviors, the HBM suffers from a number of weaknesses. The way in which the variables in the HBM combine to produce behavior has not been precisely specified (but see Becker & Rosenstock, 1987) and so the HBM is thus frequently tested as six independent predictors of behavior. In addition, various researchers have used somewhat different operationalizations of the six constructs (see Becker & Maiman, 1983; Rosenstock, 1974). Together these factors have weakened the status of the HBM as a coherent SCM of health behavior (Conner, 1993; Sheeran & Abraham, 1996). Moreover, key social cognitive variables, found to be highly predictive of behavior in other models, are not incorporated in the HBM. For example, intentions to perform a behavior and social pressure are key components of the TRA/TPB which do not appear in the HBM. Also, perceptions of personal control over the performance of the behavior (self-efficacy beliefs) which have been found to be such powerful predictors of behavior in models based upon social cognitive theory (Bandura, 1982; Schwarzer & Fuchs, 1996) are not explicitly included in the HBM. In addition, in not specifying a causal ordering among the variables, as is done in other models, more powerful analysis of data and clearer indications of how interventions may have their effects are precluded in the HBM. Several authors have noted, for example, that threat is perhaps best seen as a more distal predictor of behavior acting via influences upon outcome expectancies. Finally, the model is static; there is no distinction between a motivational stage dominated by cognitive variables and a volitional phase where action is planned, performed, and maintained (Schwarzer, 1992). Such distinctions are thought to be important in understanding various health behaviors. Hence, while an extremely popular SCM for use in understanding health behavior, it is also in a number of ways limited and may receive relatively less attention in the future.

8.01.3.3 Theory of Planned Behavior

8.01.3.3.1 Model description

The TPB was developed by social psychologists and has been widely applied to the understanding of a variety of behaviors including health behaviors (Ajzen, 1988, 1991; Conner & Sparks, 1996) (see Figure 2). The TPB details how the influences upon an individual determine that individual's decision to follow a particular behavior. This theory is an extension of the widely applied TRA (Ajzen & Fishbein, 1980; Fishbein & Ajzen, 1975). The TPB suggests that the proximal determinants of behavior are intentions to engage in that behavior and perceived behavioral control over that behavior. Intentions represent a person's motivation in the sense of his or her conscious plan or decision to exert effort to perform the behavior. Perceived behavioral control is a person's expectancy that performance of the behavior is within his/her control. The concept is similar to Bandura's (1982) concept of self-efficacy (see Schwarzer & Fuchs, 1996). Control is seen as a continuum with easily-executed behaviors at one end and behavioral goals demanding resources, opportunities, and specialized skills at the other.

Intentions are determined by three variables. The first is attitudes, which are the overall evaluations of the behavior by the individual. The second is subjective norms, which consist of a person's beliefs about whether significant others think he/she should engage in the behavior. The third is perceived behavioral control (PBC), which is the individual's perception of the extent to which performance of the behavior is within his/her control. In addition, to the extent that PBC reflects actual control, it is predicted to directly influence behavior.

Just as intentions are held to have determinants, so the attitude, subjective norm, and perceived behavioral control components are also held to have determinants. The attitude component is a function of a person's salient behavioral beliefs, which represent perceived likely consequences of the behavior. Following expectancy-value conceptualizations (Peak, 1955), the model quantifies consequences as being composed of the multiplicative combination of the judged likelihood that performance of the behavior will lead to a particular outcome and the evaluation of that outcome. These expectancy-value products are then summed over the salient consequences. It is not claimed that individuals perform such calculations each time they are faced with a decision about whether to perform a behavior or not, but rather the results of such considerations are maintained in memory and retrieved and used when necessary (Eagly & Chaiken, 1993). However, it is also possible for the individual to retrieve the relevant beliefs and evaluations when necessary.

Subjective norm is a function of normative beliefs, which represent perceptions of specific salient others' preferences about whether one should or should not engage in a behavior. In the model, this is quantified as the subjective likelihood that specific salient groups or individuals (referents) think the person should or should not perform the behavior, multiplied by the person's motivation to comply with that referent's expectation. Motivation to comply is the extent to which the person wishes to comply

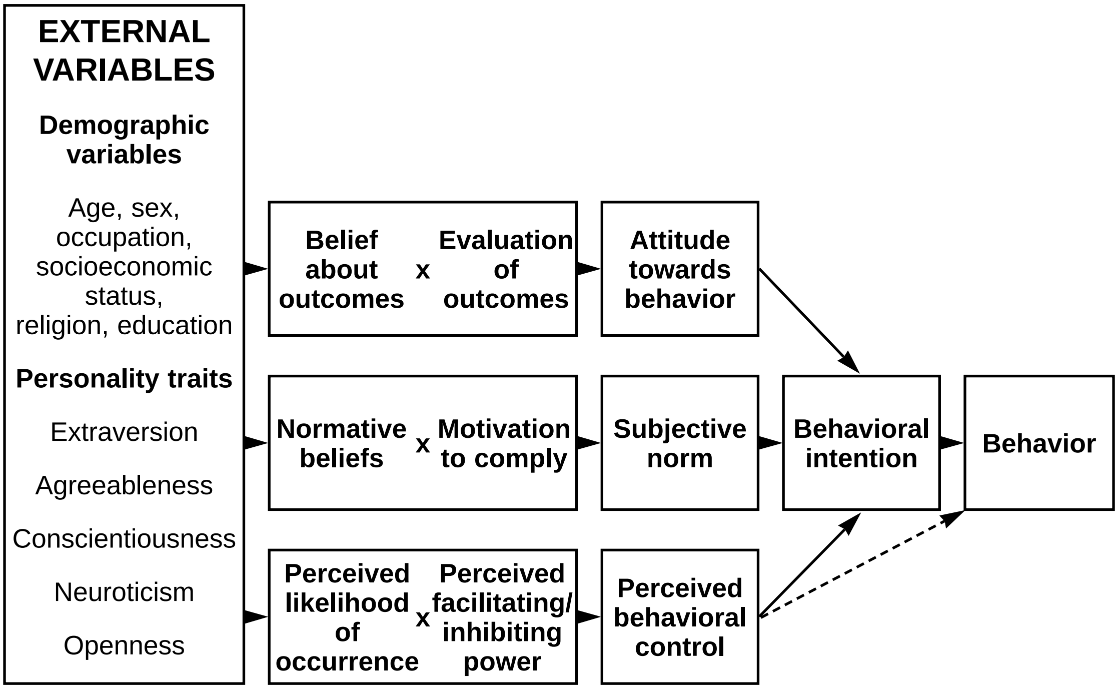
EXTERNAL VARIABLES
Demographic variables
Age, sex, occupation, socioeconomic status, religion, education
Personality traits
Extraversion
Agreeableness
Conscientiousness
Neuroticism
Openness
Belief about outcomes
x
Evaluation of outcomes
Normative beliefs
x
Motivation to comply
Perceived likelihood of occurrence
x
Perceived facilitating/inhibiting power
Attitude towards behavior
Subjective norm
Perceived behavioral control
Behavioral intention
Behavior

Figure 2 Theory of planned behavior.

with the specific wishes of the referent on this issue. These products are then summed across salient referents. Judgments of perceived behavioral control are influenced by beliefs concerning access to the necessary resources and opportunities to perform the behavior successfully, weighted by the perceived power of each factor (Ajzen, 1988, 1991). The perception of factors likely to facilitate or inhibit the performance of the behavior are referred to as control beliefs. These factors include both internal control factors (information, personal deficiencies, skills, abilities, emotions) and external control factors (opportunities, dependence on others, barriers). People who perceive they have access to the necessary resources and perceive that there are the opportunities (or lack of obstacles) to perform the behavior are likely to perceive a high degree of behavioral control (Ajzen, 1991). Ajzen (1991) has suggested that each control factor is weighted by its perceived power to facilitate or inhibit performance of the behavior. The model quantifies these beliefs by multiplying the frequency or likelihood of occurrence of the factor by the subjective perception of the power of the factor to facilitate or inhibit the performance of the behavior.

So, according to the TPB, individuals are likely to follow a particular health action if they believe that the behavior will lead to particular outcomes which they value, if they believe that people whose views they value think they should carry out the behavior, and if they feel that they have the necessary resources and opportunities to perform the behavior.

8.01.3.3.2 Review of research

The TPB has been applied to the prediction of a number of different behaviors including health-relevant behaviors with varying degrees of success (Ajzen, 1991; see Conner & Sparks, 1996 for a review of the application of the TPB to health behaviors). For example, smoking has been a focus of several studies (Babrow, Black, & Tiffany, 1990; Godin, Valois, Lepage, & Desharnais, 1992). Godin et al. (1992) looked at the prediction of the frequency of smoking in the general public over a six-month period. The prediction of intentions was significantly improved by the addition of the PBC component, and actual smoking behavior appeared to be primarily related to PBC. One study has examined drinking alcohol (Schlegel, D'Avernas, Zanna, & DeCourville, 1992) and found that PBC contributed to the predictions of intentions but not the frequency of getting drunk in nonproblem drinkers, while in problem drinkers, the PBC also contributed to predictions of frequency of getting drunk.

Several studies examined condom use (Boldero, Moore, & Rosenthal, 1992; Wilson, Zenda, McMaster, & Lavelle, 1992). Nucifora, Gallois, and Kashima (1993), for example, examined undergraduates' use of condoms using the TPB. PBC was found to make a small but significant contribution to the predictions of intentions to use condoms and actual condom use. However, intentions appeared to be principally determined by attitudes and subjective norms, while behavior was mainly influenced by intentions. Exercise has also been examined in several studies (Dzewaltowski, Noble, & Shaw, 1990; Godin & Shepherd, 1987; Norman & Smith, 1995). Dzewaltowski et al. (1990) reported the application of the TPB to exercise participation. Intentions were based both upon attitudes and PBC, but not subjective norms, whilst actual behavior seemed to be principally determined by intentions. Breast or testicle self-examination has been the focus of a couple of studies (McCaul, Sandgren, O'Neill, & Hinsz, 1993; Young, Lierman, Powell-Cope, & Kasprzyk, 1991). McCaul et al. (1993) showed the TRA components to predict breast/testicle self-examination intentions and behaviors, with PBC adding significantly to predictions of intentions but not behavior.

A range of other behaviors have been examined using the TPB, including health screening attendance (DeVellis, Blalock, & Sandler, 1990; Norman & Conner, 1993), food choices (Beale & Manstead, 1991; Sparks & Shepherd, 1992), kidney donation (Borgida, Conner, & Manteufel, 1992), drug compliance (Hounsa, Godin, Alihonou, & Valois, 1993), patient education (Kinket, Paans, & Verplanken, 1992), and weight control (Netemeyer, Burton, & Johnston, 1991; Schifter & Ajzen, 1985).

The published studies applying the TRA have been reviewed by Sheppard, Hartwick, and Warshaw (1988) and van den Putte (1993), with Ajzen (1991) reviewing 16 studies using the TPB. The findings are generally supportive of the TRA/TPB. Ajzen (1991) reports the multiple correlation between intentions and attitude, subjective norm and PBC to be 0.71 across the 16 studies he reviewed. Van den Putte (1993) computes a value of $r = 0.64$, but notes the large variation in results between behaviors. Ajzen reports the mean correlation between intentions, PBC and behavior to be 0.51, while van den Putte computes a value of 0.46. Ajzen (1991) and Madden, Ellen, and Ajzen (1992) report empirical evidence that PBC significantly improves predictions of both intentions and behavior. Hence, in summary, the evidence is broadly supportive of the TPB in helping to understand and predict health behaviors. The relative importance of the different predictors is largely an empirical matter. However, in terms of predictors of intentions, it has been argued that attitude may be more important than subjective norms for health behaviors performed in private (e.g., breast self-examination), while subjective norm may be more important than attitudes where the behavior is performed in public (e.g., safety helmet use) (Quine, Rutter, & Arnold, 1998).

8.01.3.3.3 Commentary

The TPB has been widely tested and successfully applied to the understanding of a variety of behaviors (for reviews see Ajzen, 1991; Conner & Sparks, 1996; Sheppard et al., 1988). The theory incorporates a number of important cognitive variables which appear to determine health behaviors (intentions, expectancy values, perceived behavioral control). Also the role of social pressure from others is incorporated in the model in the form of subjective norms. However, perhaps because the model was developed outside the health arena, the model does not make an assessment of health threat as is included in models such as the HBM. Finally, the theory states a clear causal ordering among variables in how they relate to behavior, allowing sophisticated analysis techniques to be applied to assessing the model.

Sheppard et al. (1988), in a review of the TRA, have argued for the need to consider both behavioral intentions and self-predictions when predicting behavior. Warshaw and Davis (1985) noted a number of different ways in which intentions had been measured, and distinguished measures of behavioral intentions (e.g., "I intend to perform behavior x") and from measures of self-predictions (e.g., "How likely is it that you will perform behavior x?"). This distinction is important when considering the prediction of health behavior because while, for example, David might intend to quit smoking, he might also think that it is unlikely that he will do so. Sheppard et al. went on to argue that self-predictions should provide better predictions of behavior as they are likely to include a consideration of those factors which may facilitate or inhibit performance of a behavior as well as a consideration of the likely choice of other competing behaviors. Sheppard et al.'s meta-analysis supports this view; measures of self-predictions were found to have stronger relationships with behavior than behavioral intentions. However, Norman and Smith (1995) found no difference in the extent to which the two measures correlated with exercise behavior. Furthermore, the measures of behavioral intentions and self-predictions are strongly correlated, suggesting that the discriminant validity for the two concepts may be weak. Bagozzi (1992) notes that the causal path may begin with the formation of desires which then develop into intentions, which in turn inform self-predictions. However, Conner and Sparks (1996) note that while theoretically these concepts may be distinguishable, empirically there is little to distinguish the three concepts. Clearly, more work is needed to further disentangle these and other related constructs that have appeared in the literature such as planning and commitment (Bagozzi, 1992, 1993), need to change (Paisley & Sparks, 1998), and behavioral willingness (Gibbons, Gerrard, Ouelette, & Burzette, 1998).

Self-efficacy is a powerful predictor of the performance of a range of health behaviors (Schwarzer & Fuchs, 1996). Ajzen (1991) argues that the PBC and self-efficacy constructs are interchangeable. However, several authors (e.g., Terry & O'Leary, 1995) have suggested that self-efficacy and PBC are not entirely synonymous. Bandura (1986), for example, argues that control and self-efficacy are quite different concepts. Although Bandura (1986) accepts that some external factors (e.g., task difficulty) will have an influence on self-efficacy, it may still be argued that self-efficacy is more concerned with perceptions of control based on internal control factors. In contrast, PBC is likely to reflect more external factors, and may be more usefully described as "perceptions of control over the behavior" (Armitage & Conner, in press). For example, Terry and O'Leary (1995) measure self-efficacy over exercising by items such as "For me to exercise would be ... easy–difficult" and perceived control over the behavior by "How much control do you have over exercising? no control–complete control."

De Vries, Dijkstra, and Kuhlman (1988) have advocated the use of measures of self-efficacy as opposed to PBC in the prediction of intentions and behavior. Further, Dzewaltowski et al. (1990), in a comparison of the TRA, TPB, and Bandura's (1986) Social Cognitive Theory, found that self-efficacy rather than PBC had a direct impact on behavior. Terry and O'Leary (1995) examined exercise behavior and found that self-efficacy only predicted intentions, while PBC had main and interactive effects on exercise behavior. Crucially, a combined measure of PBC and self-efficacy failed to moderate the effect of intention on behavior, suggesting that the two constructs are not synonymous. This issue warrants further empirical study. It seems plausible, however, that perceptions of control and self-efficacy are two separable constructs which not only have differential effects on intentions and behavior but may act differently for different behaviors.

The TPB is correctly regarded as a theory of the proximal determinants of behavior. Indeed, Ajzen (1991) describes the model as open to further elaboration if further important proximal determinants are identified. A number of potential candidate variables for addition to the TRA/TPB have been suggested. In each case both theoretical and empirical justifications are necessary (Fishbein, 1993). Some of the most promising of these "additional variables" are considered in Section 8.01.5.2.

8.01.3.4 Health Locus of Control

8.01.3.4.1 Model description

The HLOC construct is one of the most widely researched constructs in relation to the prediction of health behavior (K. A. Wallston, 1992). Its origins can be traced back to Rotter's (1954) social learning theory which states that the likelihood of a behavior occurring in a given situation is a function of the individual's expectancy that the behavior will lead to a particular reinforcement and the extent to which the reinforcement is valued. As well as being applied on a specific level, Rotter argued that social learning theory could be applied on a general level such that individuals may have generalized expectancy beliefs which cut across situations. It was at this generalized level that Rotter introduced the distinction between internal and external locus of control orientations, with "internals" believing that events are a consequence of their own actions and thereby under personal control and "externals" believing that events are unrelated to their actions and thereby beyond their personal control.

Early work examining the relationship between locus of control and health behavior with Rotter's (1966) Internal–External (I–E) scale produced mixed results. However, the I–E scale was criticized for being too generalized to predict health behavior and for conceptualizing locus of control as a unidimensional construct. In particular, Levenson (1974) argued that internal locus of control beliefs are orthogonal to external locus of control beliefs, and that within external locus of control a distinction can be made between external control exerted by powerful others and the influence of chance or fate. The development of the multidimensional health locus of control (MHLC) Scale (K. A. Wallston et al., 1978) addressed both these criticisms, and has since become the most popular locus of control measure in research on health behavior (Wallston & Wallston, 1981, 1982).

The MHLC scale measures generalized expectancy beliefs with respect to health along three dimensions. These measure the extent to which individuals believe their health is a function of their own actions (i.e., internal HLOC), the actions of powerful others such as health professionals (i.e., powerful others HLOC), and the influence of chance or fate (i.e., chance HLOC). According to HLOC theory, individuals who have strong internal HLOC beliefs should be more likely to engage in health-promoting behaviors. Conversely, those who believe that their health is due to chance or fate should be less likely to engage in health-promoting behaviors. The prediction for powerful others HLOC is less clear cut. Strong powerful others HLOC beliefs may reflect a receptivity to health messages endorsed by health professionals. Alternatively, strong powerful others HLOC beliefs may indicate a strong belief in the ability of health professionals to cure subsequent illnesses and may be unrelated or negatively related to the performance of health-promoting behaviors.

According to social learning theory, the above relationships should only hold for individuals who place a high value on their health as behavior is a function of both expectancy beliefs (e.g., HLOC) and the value attached to certain outcomes (e.g., health value). As K. A. Wallston (1991) argues, individuals are unlikely to engage in health-promoting behaviors if they place a low value on their health, whatever their HLOC beliefs. Thus, health value should act as a moderator of the relationship between HLOC and health behavior.

8.01.3.4.2 Review of research

The majority of the research using the HLOC construct has correlated HLOC beliefs with the performance of health behavior, without paying attention to the potential moderating role of health value. This may be due to a lack of appreciation of the complexity of social learning theory (Wallston, 1991) and/or an unchallenged assumption that all people value their health (Lau, Hartman, & Ware, 1986).

Several studies have examined the relationship between HLOC beliefs and general indices of health behavior. Most of these have reported a positive correlation between internal HLOC beliefs and the performance of health-promoting behaviours (Duffy, 1988; Mechanic & Cleary, 1980; Seeman & Seeman, 1983; Waller & Bates, 1992; Weiss & Larsen, 1990), although other studies have failed to find such a relationship (Brown, Muhlenkamp, Fox, & Osborn, 1983; Muhlenkamp, Brown, & Sands, 1985; Norman, 1995; Steptoe et al., 1994; Wurtele, Britcher, & Saslawsky, 1985). Some studies have found a negative relationship between chance HLOC beliefs and health behavior indices (Brown et al., 1983; Duffy, 1988; Muhlenkamp et al., 1985; Steptoe et al., 1994). Finally, powerful others HLOC beliefs have rarely been found to predict the performance of health-promoting behaviors (Brown et al., 1983; Duffy, 1988; Muhlenkamp et al., 1985; Steptoe et al., 1994; Waller & Bates, 1992; Weiss & Larsen, 1990). The above results are generally in line with predictions, although the relationship between HLOC and health behavior is typically a weak one.

A similar pattern of results is obtained for the relationship between HLOC beliefs and specific health behaviors. For example, studies have reported a link between internal HLOC beliefs and exercise (Carlson & Petti, 1989; O'Connell & Price, 1982; Slenker, Price, & O'Connell, 1985), while other studies have found only a weak link or no link (Burk & Kimiecik, 1994; Calnan, 1989; Liao, Hunter & Weinman, 1995; Norman, 1990, 1995; Speake, Cowart, & Stephens, 1991). In a large-scale representative sample, Calnan (1989) found significant negative correlations between the powerful others and chance HLOC dimensions and exercise.

In relation to alcohol consumption, a number of early studies used the locus of control construct to compare alcoholics with nonalcoholics producing mixed results (Butts & Chotlas, 1973; Costello & Manders, 1974). More recent work with the HLOC construct has produced similar mixed results with a number of studies finding no relationship between HLOC beliefs and drinking behavior (Dean, 1991; Liao et al., 1995; Norman, 1990, 1995) and Calnan (1989) only finding weak negative correlations between powerful others and chance HLOC beliefs and alcohol consumption.

A number of studies have applied the HLOC construct to the prediction of smoking cessation. A few of these studies have found internal HLOC beliefs to be related to smoking cessation (Horwitz, Hindi-Alexander, & Wagner, 1985; Rosen & Shipley, 1983). Other studies have failed to find a relationship between internal HLOC beliefs and smoking cessation (Kaplan & Cowles, 1978; Segall & Wynd, 1990; Wojcik, 1988). Smokers who believe that their health is under the control of powerful others might be expected to be more successful in giving up smoking after attending a formal smoking cessation program. However, a couple of studies suggest that strong powerful others HLOC beliefs are related to a greater likelihood of relapse following attendance at a smoking cessation program (Segall & Wynd, 1990; Wojcik, 1988).

Studies examining the relationship between HLOC beliefs and the performance of breast self-examination among women have found a positive relationship with internal HLOC beliefs (Redeker, 1989) and a negative relationship with powerful others HLOC beliefs (Hallal, 1982; Nemeck, 1990). The negative relationship with powerful others HLOC beliefs may reflect a belief that breast examination is the responsibility of health professionals. In support of this view, Bundek, Marks, and Richardson (1993) found a positive relationship between gynecological screening including physician breast examination and powerful others HLOC beliefs, and a positive relationship between self breast examination and internal HLOC beliefs. However, other studies have failed to find any relationship between HLOC beliefs and breast self-examination (Lau et al., 1986; Liao et al., 1995; Seeman & Seeman, 1983).

Those studies that have tested the moderating role of health value have generally produced positive results. Considering the prediction of indices of health behavior first, Weiss and Larsen (1990) found a significant correlation between internal HLOC beliefs and a health behavior index among individuals placing a high value on their health, but a nonsignificant correlation among individuals placing a low value on their health. Similar results have been reported by a number of researchers (Lau, 1982; Lau et al., 1986; Seeman & Seeman, 1983; K. A. Wallston & Wallston, 1980), although other studies have failed to find evidence for the moderating role of health value (Norman, 1995; Wurtele et al., 1985). Studies looking at the interaction between health value and the powerful others and chance dimensions have generally produced nonsignificant results (Lau et al., 1986; Wurtele et al., 1985).

A similar pattern of results emerges when the performance of specific health behaviors is considered. Evidence for an interaction between internal HLOC beliefs and health value has been found for a range of behaviors including dietary behavior (Hayes & Ross, 1987), smoking cessation (Kaplan & Cowles, 1978), breast self-examination (Lau et al., 1986), and information seeking (K. A. Wallston, Maiders, & Wallston, 1976). However, other studies have failed to find evidence for such an interaction when considering cancer-preventive behavior (McCusker & Morrow, 1979), exercise (Burk & Kimiecik, 1994), attendance at health checks (Norman, 1991), and information seeking (De Vito, Bogdanowicz, & Reznikoff, 1982).

8.01.3.4.3 Commentary

Reviews of research with the HLOC construct have concluded that HLOC is a weak predictor of health behavior, even when considered in conjunction with health value (Norman & Bennett, 1996; K. A. Wallston, 1991, 1992). There are a number of reasons for this poor performance which are outlined below.

K. A. Wallston and Wallston (1981, 1982) have argued that certain combinations of HLOC beliefs may be important in predicting health behavior. For example, when being advised to quit smoking by a health professional, it may be advantageous to have a combination of strong powerful others *and* internal HLOC beliefs. K. A. Wallston and Wallston (1981) have therefore proposed a 2 × 2 × 2 typology based on median splits on the three HLOC dimensions (see also Waller and Bates, 1992). Using this typology, "believers in control" (i.e., high internal high powerful others, low chance HLOC beliefs) have been found to show better adjustment in relation to rheumatoid arthritis (Roskam, 1986) and diabetes (Bradley, Lewis, Jennings, & Ward, 1990). To date, this approach has not been applied to the prediction of health behavior. However, despite the promise of this approach, its utility may be limited given that large sample sizes are required in order to compare all eight types. In addition, this approach may lead to results which are difficult to interpret, given the large number of comparisons.

A number of researchers have questioned the way in which health value is measured. Typically, health value has been measured as an *absolute* value using Lau et al.'s (1986) four-item scale. However, when deciding whether or not to perform a specific health behavior, individuals are often faced with more appealing alternatives. As a result, values other than health may be important in determining behavior. In such cases it may be more appropriate to use *relative* measures of health value. Kristiansen (1986) followed this approach and found that a measure of the value placed on health relative to value of an exciting life was more predictive of health behavior among young people than an absolute measure of health value. However, Wurtele et al. (1985) reported the opposite pattern of results in a sample of female undergraduates. Clearly, more work is needed comparing different approaches to the measurement of health value.

The need to consider behavior-specific control beliefs has been highlighted by a number of researchers. In particular it is possible to argue that one reason for the relatively poor performance of the HLOC construct is that it measures generalized expectancy beliefs with respect to health. In other words, while HLOC is specific to a given goal (i.e., health), it cuts across many situations (e.g., smoking, exercise, diet, etc.). As a result, the development of behavior-specific HLOC scales has been advocated. Georgiou and Bradley's (1992) smoking-specific locus of control scale is a good example of such a scale. This scale was found to have stronger correlations with smokers' behaviors and intentions than the more generalized MHLC scale. Other scales have been developed in relation to exercise (Burk & Kimiecik, 1994), AIDS risk behavior (Kelley et al., 1990), weight loss (Saltzer, 1982), alcohol use (Donovan & O'Leary, 1978) and a range of specific conditions including diabetes (Bradley et al., 1990), arthritis (Nicassio, Wallston, Callahan, Herbert, & Pincus, 1985), cancer (Prwun et al., 1988), hypertension (Stanton, 1987) and heart and lung disease (Allison, 1987). Generally, these scales have been found to be more predictive of health behavior than more generalized measures (Lefcourt, 1991).

In conclusion, the amount of variance in health behavior explained by the HLOC construct is low, even when considered in conjunction with health value (Norman & Bennett, 1996; K. A. Wallston, 1991, 1992). Nevertheless, the pattern of results obtained are generally in line with predictions suggesting that HLOC beliefs may have a distal influence on health behavior. In line with this position, K. A. Wallston (1992) has proposed a "modified social learning theory" in which health behavior is a function of HLOC beliefs, health value, and self-efficacy. Importantly, internal HLOC beliefs are seen to be a necessary, but not sufficient, condition for performing a health behavior. To date, there have been no formal tests of this modified theory, although the attempt to embed HLOC beliefs into a broader theoretical perspective is an encouraging development.

8.01.3.5 Protection Motivation Theory

8.01.3.5.1 Model description

Protection motivation theory (Rogers, 1983) was originally developed as a framework for understanding the effectiveness of health-related persuasive communications, although more recently it has also been used to predict health protective behavior. It has its origins in early work on the persuasive impact of fear appeals, which was concerned with the conditions under which fear appeals may influence attitudes and behavior. In an extension of the fear-drive model, Janis (1967) proposed that if a persuasive communication successfully arouses fear, usually through emphasizing the severity of a threat and the likelihood of its occurrence, individuals will be motivated to reduce this unpleasant emotional state. If the message also contains recommendations for action, then one way in which individuals can reduce this state of arousal is to follow the communicator's advice. If the message does not contain effective behavioral advice, then maladaptive coping reactions may follow such as denial or avoidance. Janis proposed that fear appeals may be most effective when a medium level of fear is evoked. Under such conditions the cognitive responses that promote adaptive reactions (e.g., following behavioral advice) outweigh those that promote maladaptive reactions (e.g., denial). However, later work has failed to confirm this hypothesis (Sutton, 1982).

Leventhal (1970) made a similar distinction between adaptive and maladaptive reactions in the parallel response model which differentiates between two independent control processes that are initiated by a fear appeal. The first, fear control, focuses on attempts to reduce the emotional threat (e.g., avoidance) while the second, danger control, focuses on attempts to reduce the threatened danger (e.g., following behavioral advice). The parallel response model is important in proposing that protection motivation results from danger control processes (i.e., cognitive responses) rather than from fear control processes (i.e., emotional responses).

Roger's (1983) (PMT) outlines the cognitive responses resulting from fear appeals in more detail (see Figure 3). It is argued that various environmental (e.g., fear appeals) and intrapersonal (e.g., personality variables) sources of information can initiate two appraisal processes: threat appraisal and coping appraisal. Threat appraisal, which is similar to Lazarus and Launier's (1978) primary appraisal, focuses on the source of the threat and the factors that may increase or decrease the probability of the maladaptive response. Both the perceived severity of the threat and the individual's perceived vulnerability to the threat are seen to inhibit maladaptive responses. However, there may be a number of intrinsic (e.g., pleasure) and extrinsic (e.g., social approval) rewards which may serve to increase the likelihood of maladaptive responses. Coping appraisal, which is similar to Lazarus and Launier's (1978) secondary appraisal, focuses on one's ability to cope with the threat and the factors that may increase or decrease the probability of an adaptive response. Both the belief that the recommended action will be effective in reducing the danger (i.e., response efficacy) and the belief that one is capable of performing the recommended action (i.e., self-efficacy) are likely to increase the probability of an adaptive response, although various response costs (e.g., financial cost) associated with performing an adaptive response will serve to inhibit such a response.

Protection motivation results from the two appraisal processes and is a positive function of beliefs about severity, vulnerability, response efficacy, and self-efficacy, and a negative function of beliefs about the rewards associated with the maladaptive response and the response costs of the protective behavior. Moreover, for protection motivation to be elicited, it is necessary for the rewards associated with the maladaptive response to be outweighed by perceptions of severity and vulnerability, and the response costs of the protective behavior to be outweighed by perceptions of response efficacy and self-efficacy. Protection motivation, which is usually measured by behavioral intentions, is seen to arouse, direct, and sustain protective behavior.

8.01.3.5.2 Review of research

Protection motivation theory provides a framework for understanding both the effects of fear appeals and the social cognitive variables underlying health behavior. As a result, tests of PMT have taken two forms. In the first, the key components of PMT are manipulated in persuasive communications and their effects on protection motivation tested (see Section 8.01.4.1). In the second, PMT is considered as a general attitude–behavior model and its components used as predictors of health behavior. This work is reviewed below.

One area in which PMT has been used as a general attitude–behavior model is in relation to cancer-related preventive behavior. In one of the few longitudinal tests of PMT, Hodgkins and Orbell (1998) examined the social cognitive predictors of breast self-examination (BSE) in a sample of young women (17–40 year olds) over a one-month period. Each of the main components of PMT was measured (i.e., severity, vulnerability, response efficacy, self-efficacy) as well as the response costs of performing BSE (e.g., I would feel awkward examining my breasts). In a path analysis, only self-efficacy was related to intentions to perform BSE. Time one behavioral intention was in turn found to be the most important predictor of performance of BSE at one month follow-up. Similar results have been reported by Seydel, Taal, and Wiegman (1990) who found response efficacy and self-efficacy to be predictive of intentions to engage in, and concurrent performance of, a number of cancer-related preventive behaviors. Boer and Seydel (1996) also found response efficacy and self-efficacy to be predictive of intentions to participate in mammography screening.

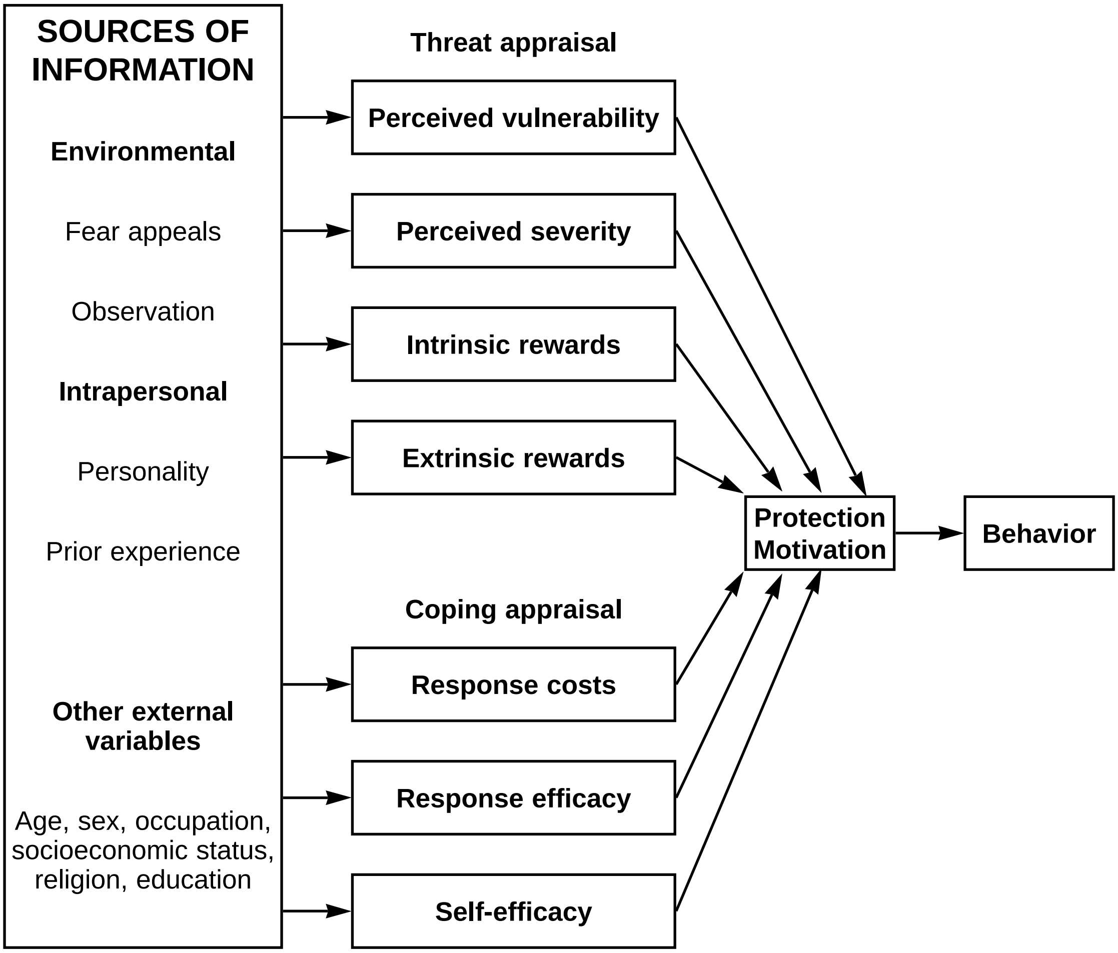
SOURCES OF INFORMATION
Environmental
Fear appeals
Observation
Intrapersonal
Personality
Prior experience
Other external variables
Age, sex, occupation, socioeconomic status, religion, education
Threat appraisal
Perceived vulnerability
Perceived severity
Intrinsic rewards
Extrinsic rewards
Coping appraisal
Response costs
Response efficacy
Self-efficacy
Protection Motivation
Behavior

Figure 3 Protection motivation theory.

A second area of application of PMT has been in relation to AIDS risk-reducing behaviors such as condom use. Aspinwall, Kemeny, Taylor, Schneider, and Dudley (1991) examined the ability of the PMT components to predict reductions in a number of AIDS risk-reducing behaviors in a sample of gay men over a six-month period. The results showed that strong levels of self-efficacy and a high level of perceived vulnerability at time one were predictive of a reduction in the number of sexual partners over the six-month follow-up period. In addition, self-efficacy also emerged as the most important predictor of reductions in the number of anonymous sexual partners.

Van der Velde and van der Pligt (1991) used PMT as a framework for assessing the coping responses of heterosexual men and women and homosexual men with multiple partners. Considering the heterosexual sample first, it was found that vulnerability, response efficacy, and self-efficacy all had a direct positive effect on behavioral intentions to use condoms. In addition, severity had an indirect effect on behavioral intentions through a measure of fear. Similar results were found with the homosexual sample with response efficacy, self-efficacy, and severity having a positive effect on behavioral intentions to engage in safe sex. However, contrary to expectations, a negative relationship was found between vulnerability and behavioral intentions.

In a sample of male and female adolescents, Abraham, Sheeran, Abrams, and Spears (1994) found that self-efficacy had a positive influence and response costs (i.e., concern about reputation) had a negative influence on behavioral intentions to use a condom. In addition, vulnerability had a negative effect on behavioral intentions to limit the number of sexual partners. In a cross-sectional study of male and female heterosexuals, Bengel, Beltz-Merk, and Farin (1996) found that self-efficacy was related to a greater use of condoms and fewer sexual partners. However, perceptions of vulnerability were related to greater use of condoms, but also higher number of sexual partners.

Taken together, the above results suggest that PMT is a useful framework for understanding HIV-related protective behavior. Self-efficacy emerges as the most important predictor of such behavior, with response efficacy and severity also emerging as significant predictors in some studies. A conflicting pattern of results has been found with the vulnerability component. Similar conclusions have been reached by Farin (1994) in a meta-analysis of PMT and HIV-protective behavior, in which self-efficacy and response efficacy emerged as the best predictors of protective behavior. However, these two components were only able to explain 2.2% and 1.8% of the variance in such behavior. Severity was seen to be less important, and vulnerability had a conflicting pattern of results.

8.01.3.5.3 Commentary

Despite relatively few studies in the area, the above review highlights the potential utility of PMT as a framework for considering the social cognitive predictors of health protective behavior. However, there are a number of issues which future work needs to address.

First, some studies have reported a positive relationship between perceptions of vulnerability and protection motivation, while others have reported a negative relationship. Seydel et al. (1990) suggest that the negative relationship between vulnerability and intentions to engage in cancer-related preventive behavior in their study may be due to a "defensive avoidance" style of coping, in which perceptions of vulnerability to cancer may lead to feelings of anxiety which may inhibit adaptive responses and promote avoidance. However, it is more likely that the mixed pattern of results is due to measurement issues. As Weinstein and Nicolich (1993) argue, a negative correlation may be expected between perceptions of vulnerability and concurrent protective behavior given that one's current behavior may be used to make vulnerability judgments. In contrast, a positive correlation may be expected between perceptions of vulnerability and future protective behavior to the extent that perceptions of vulnerability motivates protective behavior. Considering behavioral intentions, individuals may feel vulnerable and therefore intend to engage in a protective behavior (i.e., positive correlation), or may feel vulnerable because they do not intend to engage in a protective behavior (i.e., negative correlation). Clearly, more consideration needs to be given to the measurement of perceptions of vulnerability and the potential use of conditional measures of risk.

Second, while PMT provides a framework for considering the social cognitive predictors of protective behavior, a number of researchers have commented on its lack of specification in terms of the nature of its components and the relationships between them (Bengel et al., 1996). For example, Rogers (1983) states that the response costs of the protective behavior need to be outweighed by perceptions of response efficacy and self-efficacy for protection motivation to be elicited. However, no guidance is given as to how these variables are to be measured and combined in order to predict protection motivation.

Overall, PMT identifies many of the social cognitive variables which are important predictors of health behavior. It shares a number of similarities with the HBM (i.e., measures of perceived susceptibility, severity, benefits, and barriers) although it also includes self-efficacy, which has been found to be one of the most powerful explanatory constructs in relation to health behavior (Schwarzer & Fuchs, 1996), and a measure of behavioral intention, which is seen to mediate the influence of threat appraisal and coping appraisal. It is also important for providing a synthesis between social cognitive approaches and coping models as outlined by Lazarus and Launier (1978). However, to date there have been relatively few longitudinal tests of PMT in relation to health behavior. Such studies should help clarify some of the issues raised above.

8.01.3.6 Self-efficacy

8.01.3.6.1 Model description

Self-efficacy is one of the most powerful predictors of health behavior (K. A. Wallston, 1992). It has its origins in Bandura's (1977) social cognitive theory which states that behavior is a function of both incentives (i.e., reinforcements) and expectancies. Three kinds of expectancies can be identified, these being situation-outcome expectancies which refer to beliefs about how events are connected, outcome expectancies which refer to beliefs about the consequences of performing a behavior, and self-efficacy expectancies which refer to beliefs about one's ability to perform the behavior. Thus in order to perform a health behavior, individuals must value their health (i.e., incentive), believe that their current lifestyle poses a threat to their health (i.e., situation-outcome expectancy), believe that adopting the new behavior will reduce the threat to their health (i.e., outcome expectancy) and believe that they are capable of performing the behavior (i.e., self-efficacy expectancy). While all these beliefs are seen to be important in the initiation and maintenance of health behavior, self-efficacy expectancies are seen to be the most important. Individuals with strong self-efficacy beliefs are believed to develop stronger intentions to act, to expend more effort to achieve their goals, and to persist longer in the face of barriers and impediments (Bandura, 1991).

Self-efficacy beliefs are therefore believed to play a crucial role in the determination of health behavior. According to Bandura (1977, 1982), such beliefs can be conceptualized and measured in terms of three parameters; magnitude, strength, and generality. The first parameter refers to the level of difficulty of the behavior. Individuals with low-level expectations feel capable of performing only very simple behaviors, whereas individuals with high-level expectations feel capable of performing even the most difficult of behaviors. In this way it is possible to assess individuals' expectations about their level, or magnitude, of performance. The second parameter refers to individuals' confidence that they could perform a specific behavior, while the third parameter refers to the generality of expectations across situations or domains. The measurement of self-efficacy usually focuses on the strength of the self-efficacy expectation (e.g., "I am confident that I can refrain from smoking"), although it will often incorporate the magnitude of expectation (e.g., "I am confident that I can refrain from smoking, even if someone offers me a cigarette").

8.01.3.6.2 Review of research

The self-efficacy construct has been successfully applied to the prediction of a range of health behaviors (see Bandura, 1991; O'Leary, 1985; Schwarzer & Fuchs, 1996). One of the main areas of application has been in relation to smoking cessation (Condiotte & Lichtenstein, 1981; Colletti, Supnick, & Payne, 1985; Kavanagh, Piere, Lo, & Shelley, 1993). These studies have found self-efficacy to be a consistent predictor of smoking cessation. For example, Condiotte and Lichtenstein (1981) found that post-treatment self-efficacy beliefs were predictive of both the probability of relapse and the amount of time before relapse. In addition, a close correspondence was noted between self-efficacy beliefs for a range of specific tempting situations (e.g., after a meal, when drinking coffee) and the actual situation in which the relapse occurred. The results therefore indicate that self-efficacy is important not only in predicting the likely success of smokers who are trying to quit, but also the situations in which they are most likely to relapse.

Self-efficacy has also been related to AIDS risk-reducing behaviors, such as condom use (Basen-Engquist & Parcel, 1992; Boyd & Wandersman, 1991; Kok, de Vries, Mudde, & Strecher, 1991; Morrison, Gillmore, & Baker, 1995). While the majority of studies examining the relationship between self-efficacy and AIDS risk-reducing behaviors have reported significant results, a few studies have failed to do so (Boyd & Wandersman, 1991; Morrison et al., 1995). Considering exercise behavior, self-efficacy has been found to be an important predictor of both intentions to engage in regular exercise and actual exercise behavior (Desharnais, Bouillon, & Godin, 1986; McAuley, 1993; Sallis, Howell, Hofsteffer, & Barrington, 1992). Finally, self-efficacy has also been related to dieting and weight control (Bernier & Avard, 1986; Hofstetter, Sallis, & Howell, 1990; Jeffrey et al., 1984). For example, Jeffrey et al. examined the relationship between self-efficacy and weight loss following participation in a behavioral treatment program over a two-year period. Making a distinction between "emotional" self-efficacy (i.e., confidence in one's ability to refrain from eating during various emotional states) and "situational" self-efficacy (i.e., confidence in one's ability to refrain from eating in various situations), they found pre-treatment measures of both types of self-efficacy to be predictive of weight loss at one and two years. However, post-treatment measures of self-efficacy were not so powerful with only "situational" self-efficacy predictive of weight loss at one year.

8.01.3.6.3 Commentary

The self-efficacy construct has been found to be one of the most important predictors of health behavior (K. A. Wallston, 1992). However, there are a number of issues which future work needs to address. First, the relationship between outcome and self-efficacy expectancies has been a source of some debate (Corcorcan, 1991; Kirsch, 1986, Maddux, 1993). As Schwarzer (1992) argues, there may be a temporal and causal order among the two types of expectancy beliefs inasmuch as individuals are unlikely to consider their ability to perform a behavior before first considering the efficacy of the behavior. Bandura (1991) argues that the effect of outcome expectancies on intentions and behavior are partly governed by self-efficacy expectancies; even if outcome expectancy is high, performance of a behavior is unlikely if self-efficacy is low. As a result, when self-efficacy is partialed out, any relationship between outcome expectancy and behavior should disappear. However, some research has indicated that outcome and self-efficacy expectancies can be independent predictors of intentions and behavior (Maddux, 1993).

Second, the role of incentives, or outcome values, has tended to be overlooked in research with the self-efficacy construct. This is despite the fact that a large body of research in expectancy value theory has found outcome value to be an important predictor (Kirsch, 1986; McCelland, 1985). Unfortunately, those studies which have considered the role of outcome and self-efficacy expectancies in conjunction with incentives or outcome values have produced mixed results (Maddux, Norton, & Stoltenberg, 1986; Manning & Wright; 1983).

Third, there is some disagreement over whether it is appropriate to consider generalized self-efficacy expectancies. According to Bandura's (1977) original conceptualization, self-efficacy beliefs should focus on specific behaviors in specific situations. Self-efficacy is not seen to be a personality trait as self-efficacy beliefs can be seen to vary across behaviors and situations, although self-efficacy beliefs may generalize to other behaviors and situations to the extent that the new behaviors require similar skills and the new situations have similar features (Bandura, 1986). Nevertheless, Schwarzer and Fuchs (1996) have argued that self-efficacy can be viewed as a generalized trait reflecting a personal resource factor, pointing to studies which have successfully employed generalized measures of self-efficacy to predict behavior (Mittag & Schwarzer, 1993; Snyder et al., 1991; K. A. Wallston, 1992). Generalized self-efficacy may be closely related to dispositional optimism (Scheier & Carver, 1992), although on an empirical level Schwarzer (1994) found a correlation of only 0.60 between dispositional optimism and generalized self-efficacy.

In conclusion, self-efficacy has been found to be one of the most powerful and consistent predictors of health behavior. For this reason alone, it is not surprising that it has been incorporated into a number of the main social cognition models of health behavior; for example, Rogers' (1983) PMT and Ajzen's (1991) TPB. In addition, Rosenstock, Strecher, and Becker (1988) have called for the inclusion of self-efficacy in the HBM. It is likely that self-efficacy will continue to attract considerable interest and continue to be a key predictor of health behavior.

8.01.4 ENCOURAGING THE ADOPTION OF HEALTH BEHAVIORS

8.01.4.1 Intervention Studies

The models presented earlier in this chapter outline some of the key social cognitive determinants of health behavior. In this section we consider the use of these models in the design of interventions to encourage new health behaviors. As Fishbein (1993) has argued, the ultimate test of the utility of these models lies in their ability to inform the design of effective interventions. To the extent that these models outline the key social cognitive determinants of health behavior, interventions which target these variables should lead to associated changes in behavior. However, to date there have been relatively few theoretically driven interventions. As a result there is still a need for more tests of these models "in action" (Fishbein, 1993).

Brawley (1993) argues that we need to take account of the practicality of employing SCMs when designing interventions. In short, a model which offers a high level of practicality must be shown to have predictive utility, to describe the relationships between key constructs, to offer guidelines for the assessment of these constructs, to allow the translation of these constructs into operational manipulations, and to provide the basis for detecting the reasons why an intervention succeeds or fails. The SCMs considered here can be been seen to have a high level of practicality and therefore should provide a good framework for the design of effective interventions.

One model which has been used widely to design interventions is the PMT. A good example of the use of PMT in this respect is provided by Wurtele and Maddux (1987) in their study on exercise intentions and behavior. In this study essays recommending beginning a regular exercise program were presented to a sample of nonexercising female undergraduates. The essays were designed so that each of the PMT's main components (i.e., severity, vulnerability, response-efficacy, and self-efficacy) were independently manipulated, resulting in a 2 × 2 × 2 × 2 between-subjects factorial design with two levels (present vs. absent) of each factor. For example, the severity message emphasized the seriousness of the threat of a sedentary lifestyle by vividly describing the negative effects of a heart attack (e.g., "Nauseated, the victim vomits; pink foam comes out of the mouth. The face turns ashen grey, sweat rolls down the face . . ."). After reading an essay, subjects completed a questionnaire containing measures of the main components of PMT and were followed up two weeks later to chart any changes in exercise behavior.

Manipulation checks revealed that each of the messages successfully manipulated their corresponding component from PMT, as is the case in most PMT intervention studies. Of more interest was the effect of the messages on protection motivation as measured by behavioral intentions to start a regular exercise program. The results showed that only the vulnerability and self-efficacy messages had a significant effect on behavioral intentions, although a three-way interaction between vulnerability, response efficacy, and self-efficacy was also found. In relation to changes in exercise behavior over the two-week follow-up period, a significant interaction was found between severity and self-efficacy, such that the self-efficacy message only had a significant effect when the severity message was absent.

Overall, studies which have used PMT to design interventions have shown that self-efficacy is the most powerful component that can be manipulated in persuasive messages (Boer & Seydel, 1996). In addition, manipulating response efficacy has been found to have a significant effect on intentions in a majority of studies. These two variables have been shown to influence intentions to engage in a range of health behaviors including exercise (Stanley & Maddux, 1986), quit smoking (Maddux & Rogers, 1983), dietary intake (Wurtele, 1988), and breast self-examination (Rippetoe & Rogers, 1987). Manipulating perceptions of vulnerability has been found to influence intentions in other studies (Maddux & Rogers, 1983; Wurtele & Maddux, 1987), while manipulating perceptions of severity has not been found to have an influence on behavioral intentions. However, many PMT intervention studies have combined the vulnerability and severity components so that the potential threat of a maladaptive behavior is emphasized. This has been found to have a significant effect on intentions in relation to alcohol use (Stainback & Roger, 1983), dental flossing (K. H. Beck & Lund, 1981), dietary behavior (Wurtele, 1988), information seeking (Brouwers & Sorrentino, 1993), and breast self-examination (Rippetoe & Rogers, 1987).

Some studies have focused more specifically on enhancing feelings of self-efficacy as a means for encouraging health behavior change. As Bandura (1986) outlines, there are four main sources of self-efficacy, each of which could be addressed in interventions. First, individuals can develop feelings of self-efficacy from personal mastery experience. For example, it may be possible to split a behavior into various subgoals, such that the easiest subgoals are achieved before more difficult tasks are attempted. Second, individuals may develop feelings of self-efficacy through observing other people succeed on a task (i.e., vicarious experience). Third, it is possible to use standard persuasive techniques to try to instil feelings of self-efficacy. Finally, one's physiological state may be used as a source of information, such that high levels of arousal or anxiety may indicate to the individual that he or she is not capable of performing a given action. As a result, relaxation techniques may be employed to help maintain feelings of self-efficacy.

Each of these techniques have been used in intervention studies to try to enhance feelings of self-efficacy. Maibach, Flora, and Nass (1991) report the results of a year-long community health campaign to encourage the adoption of health behaviors. The campaign materials were all designed to reflect the main principles of Bandura's (1986) social cognitive theory and used a number of strategies for enhancing feelings of self-efficacy and encouraging behavior change attempts. These included encouraging participants to set behavior change goals, using community members who had successfully changed their behavior as role models, using health experts to give advice about behavior change and focusing on the skills needed to support behavior change. The campaign was found to successfully increase feelings of self-efficacy which, in turn, were related to the adoption of new health behaviors.

A number of interventions have focused on more situation-specific feelings of self-efficacy. Stevens and Hollis (1989) designed an intervention to help smokers quit smoking which built on the results of earlier research which had shown that situation-specific ratings of self-efficacy were predictive of the circumstances in which relapses occurred (Condiotte & Lichtenstein, 1981). Abstinent smokers have identified potential relapse situations in which they perceived low levels of self-efficacy and then developed and rehearsed appropriate coping strategies over three weekly meetings. This intervention led to a greater abstinence rate at one year than both a discussion-only intervention and a no-treatment control. Other studies which have attempted to improve behavioral skills to enhance feelings of self-efficacy have produced positive results in relation to alcohol use (Baer et al., 1992) and dental hygiene (McCaul, Glasgow, & O'Neil, 1992).

Few studies have attempted to use the TRA/TPB as a framework for developing interventions, despite quite clear guidelines outlined by Ajzen and Fishbein (1980). In fact, van den Putte (1993) reports a mere five studies which have followed such an approach, with only limited evidence for success compared to approaches not based upon the model. For example, Brubaker and Fowler (1990) examined the effect of persuasive messages upon men's intentions to perform testicular self-examination. A persuasive message based on the theory of reasoned action was found to increase intentions to perform testicular self-examination compared to a no-message control, but was no more effective than a knowledge-only message. More recently, Parker, Manstead, and Stradling (1996) have developed intervention videos based on the TPB to discourage speeding in residential areas by car drivers. The results showed that the normative belief video had a significant effect on a postintervention measure of normative beliefs, while the behavioral belief video had no effect on behavioral beliefs, and the perceived behavioral control video had a negative effect on perceptions of control. Overall, the videos had no effect on subjects' expectations of speeding in residential areas over the next year.

Finally, there is some evidence that tailoring interventions to fit in with individuals' existing belief orientations may lead to more effective interventions. Chambliss and Murray (1979a) devised a weight control program in which participants were given placebo medication to help control their metabolism. After two weeks participants in one group were debriefed about the placebo medication and encouraged to attribute any weight loss to their own efforts over the previous two weeks. Participants in a second group were given further information about the efficacy of the medication and encouraged to attribute any weight loss to the medication. At two-week follow-up, a significant interaction was found between the giving of information and participants' preprogram locus of control orientation, such that "internals" lost more weight than "externals" in the self-efficacy information group, while the opposite pattern of results was found for the drug information group. Similar results have been reported by Chambliss and Murray (1979b) in relation to smoking cessation. These results led to further work exploring the match between the control orientation of the intervention and participants' existing health locus of control beliefs. For example, Quadrel and Lau (1989) found an interaction between health locus of control beliefs and the control orientation of a message to encourage breast self-examination among female students. In particular, those females with strong internal health locus of control beliefs who received a message in a "control" frame were more likely to perform breast self-examination at follow-up, although this effect was reversed if a neutral reminder was sent. Further evidence for a "matching hypothesis" has been provided in relation to weight reduction (B. S. Wallston, Wallston, Kaplan, & Maides, 1976) and smoking cessation (Best, 1975).

Despite offering a high level of practicality, there have been relatively few studies testing SCMs "in action." As the above review shows, practical applications of these models have met with some success and suggest that health behavior change interventions may have a lot to gain from using these models as a guiding framework. As Lewin (1951) concludes, "there is nothing so practical as a good theory" (p. 169).

8.01.5 LIMITATIONS AND EXTENSIONS OF EXISTING MODELS

8.01.5.1 Limitations of Current Models

The SCMs described above represent one widely used approach to understanding health behaviors. Here we provide a critique of this approach to understanding health behavior and outline ways in which research might develop through consideration of additional theoretical constructs, the processes by which cognitions influence behavior, stage models of behavior change, potential integrations, and future directions for work in this area.

There are several advantages of using social cognition models in health psychology (Conner, 1993; Conner & Norman, 1996b). First, they provide a clear theoretical background to research, guiding the selection of variables to measure, the procedure for developing reliable and valid measures, and how these variables are combined in order to predict health behaviors and outcomes. Second, to the extent that the models identify the important variables in predicting health outcomes and behaviors, they enable us to develop effective behavioral interventions. Third, the models provide us with a description of the cognitive processes determining individuals' motivation to perform different behaviors.

There are parallel disadvantages in too exclusive a focus upon social cognition models as the way to understand health behaviors. First, in providing such an explicit general theoretical framework, these models may lead us to neglect variables (cognitive and noncognitive) potentially important in understanding a particular health behavior or outcome. For example, the decision to use a condom may be a function of cognitions, emotional reactions, and also a complex interaction between the individuals involved. Social cognition models on their own are unlikely to provide considerable predictive power in these situations. In addition, SCMs are open to extension when empirically and theoretically justified (Fishbein, 1993).

Second, while such models provide us with targets for interventions to produce behavior change, they do not specify how such cognitions are best changed. Effective interventions need to consider both the targets (e.g., cognitions) and the persuasion process itself. This process of persuasion is described by other models of social cognitive processes (e.g., the elaboration likelihood model: Petty & Cacioppo, 1986; the systematic-heuristic model: Chaiken, Lieberman, & Eagly, 1989). In addition, applications of SCMs should not lead to neglect of alternatives to persuasion in producing behavior change such as extrinsic changes to the rewards and costs of a given behavior. For example, increased taxation and legal restrictions can be effective in producing change in health behaviors either in isolation or in tandem with persuasion.

Third, although SCMs have furthered our understanding of motivational processes and their influence upon behavior, they have neglected other aspects of behavior change. For example, few of the models consider volitional processes beyond attempting to explain intentions (Bagozzi, 1993; Gollwitzer, 1990). However, many individuals who intend to change fail to do so. Hence, we need to consider the other important volitional processes associated with attempts to change and maintain behavior change (see Norman & Conner, 1996a; Schwarzer & Fuchs, 1996).

8.01.5.2 Additional Theoretical Constructs

While the social cognition models outlined here provide an important framework for considering the social psychological determinants of health behavior, it is clear that in some instances they only account for a modest amount of the variance in health behavior. For example, Sheppard et al. (1988) noted that about 10% of studies they reviewed reported correlations between behavioral intentions and behavior below 0.2. This suggests that key variables have failed to be included in these models. We review here the most promising of such variables appearing in the literature: measures of norms, anticipated affective reactions, self-identity, and past behavior.

A number of researchers have argued that further attention needs to be paid to the concept of normative influences (e.g., Cialdini, Reno & Kallgren, 1990; Conner, Martin, Silverdale, & Grogan, 1996). Of the major SCMs, only the TPB incorporates perceived social pressures to perform a behavior as a predictor of intentions. Based upon social identification theory, Terry and Hogg (1996) suggest that such (injunctive) normative measures might be more predictive of behavior if they employed a measure of group identification rather than motivation to comply (e.g., I identify with my friends with regard to smoking). Similarly, Gibbons and Gerrard (1995, 1997) have noted the need to consider social comparison processes and how the social image or prototype of the person who performs a particular behavior influences the performance of various health behaviors (particularly among young people). Other researchers have suggested other forms of normative influence we might consider. These include descriptive norms and personal or moral norms. Descriptive norms are perceptions of the behavior of salient others. For example, Jane's eating behavior may be influenced not only by her perceptions of what others think she should eat, but also perceptions of what they actually do eat. Several studies have reported that perceptions of others' behavior contributed to the prediction of intentions independently of perceived injunctive norms (e.g., Conner et al., 1996; De Vries, Backbier, Kok, & Dijkstra, 1995; Grube, Morgan, & McGee, 1986). Personal, or moral, norms are the individual's perception of the moral correctness or incorrectness of performing a behavior (Ajzen, 1991; Sparks, 1994) and take account of, "... personal feelings of ... responsibility to perform, or refuse to perform, a certain behavior" (Ajzen, 1991, p. 199). Moral norms should have an important influence on the performance of those behaviors with a moral or ethical dimension (L. Beck & Ajzen, 1991; Gorsuch & Ortberg, 1983; Kurland, 1995). A number of studies have found measures of moral norms to be predictive of blood donating behavior (Pomazal & Jaccard, 1976; Zuckerman & Reiss, 1978) as well as intentions to donate organs (Schwartz & Tessler, 1972), eat genetically produced food (Sparks, Shepherd, & Frewer, 1995), buy milk (Raats, 1992), use condoms (Nucifora et al., 1993), and commit driving violations (Parker, Manstead, & Stradling, 1995).

Anticipated affective reactions to the performance or nonperformance of a behavior may also be an important determinant of behavior (Triandis, 1977; Van der Pligt & de Vries, 1998), especially in situations where the consequences of the behavior are unpleasant or negatively affectively laden. In the 1990s research has focused on the influence of anticipated regret (Parker et al., 1995; R. Richard & van der Pligt, 1991; Richard, Van der Pligt, & de Vries, 1995, 1996a, 1996b). It is argued that if individuals anticipate feeling regret after performing a behavior then they will be unlikely to perform the behavior. Richard et al. (1995, 1996a) investigated the role of anticipated regret in relation to condom use among adolescents and found such feelings to be an important predictor of intentions. Richard et al. (1996b) examined the influence of anticipated regret on subsequent behavior. Participants in their study were asked to either focus on their anticipated feelings following safe and unsafe sexual behavior or on their present feelings about these behaviors. At follow-up, participants in the anticipated feelings condition were more likely to have used condoms in casual sexual encounters in the intervening five months. The effects of anticipated affective reactions have been confirmed in studies of driving (Parker, Manstead, Stradling, Reason, & Baxter, 1992) and consumer behavior (Simonson, 1992). However, in terms of developing social cognition models of health behavior, it is possible to argue that anticipated affective reactions may be incorporated into constructs that focus on the consequences of behavior (e.g., behavioral beliefs in the TPB; see Van der Pligt & de Vries, 1998).

The concept of self-identity has also been suggested as a predictor of behavior (Biddle, Bank, & Slavings, 1987; Charng, Piliavin, & Callero, 1988). For example, the extent to which individuals think of themselves as "healthy eaters" should predict their dietary intentions and behavior. In support, Sparks and Shepherd (1992) found that respondents who thought of themselves as "green consumers" had stronger intentions to consume organic vegetables. Sparks (1994) noted that self-identity may simply be a proxy for past behavior, although Sparks and Shepherd (1992) found that the relationship between self-identity and future intentions remained when past consumption of organic vegetables was controlled for. Self-identity as someone who is concerned about the health consequences of one's diet has also been related to intentions to reduce fat consumption (Sparks, Shepherd, Wieringa, & Zimmermanns, 1995), although in an earlier study Sparks, Shepherd, Wieringa, and Zimmermanns (1994) failed to find an independent effect for self-identity. Role identity (regarded as synonymous with self-identity) was measured by Theodorakis (1994) and found to be a significant predictor of exercise behavior. Future work needs to assess the influence of self-identity across of range of behaviors as it may be the case that self-identity is only important in a restricted range of situations.

The influence of past on current behavior in SCMs has attracted much attention. It is argued that many health behaviors are determined by one's previous behavior rather than cognitions. The argument is based on the results of a number of studies showing past behavior to be the best predictor of future behavior. For example, Mullen, Hersey, and Iverson (1987) found initial behavior to be the strongest predictor of the consumption of sweet and fried foods, smoking, and exercise over an eight-month period. Similar results have been reported in relation to drug use (Bentler & Speckart, 1979; Huba, Wingard, & Bentler, 1981), exercise (Godin, Valois, & Lepage, 1993; Norman & Smith, 1995; Valois, Desharnais, & Godin, 1988), breast self-examination (Hodgkins & Orbell, 1998), attendance at health checks (Norman & Conner, 1993, 1996b) and seat belt use (Sutton & Hallett, 1989). Such results have led to calls for past behavior to be considered as an independent predictor of future behavior (Bentler & Speckart, 1979; Fredricks & Dossett, 1983). However, there are problems with this view. Ajzen (1988) argues that the effects of past behavior on future behavior should be mediated by the variables included in social cognition models; past behavior shapes individuals' beliefs about the behavior in question, and it is these cognitions that determine subsequent behavior. When past behavior is found to have a direct effect on future behavior it is because key social cognitive variables have not been considered (Ajzen, 1991).

There has also been focus on the concept of habit (i.e., behaving in a way you have acted before without thinking about it). Eagly and Chaiken (1993) review numerous studies where the addition of habit has added to the prediction of future behavior over and above the influence of variables such as intention, attitude, and subjective norm. Despite this evidence, few SCMs incorporate a measure of habit. A notable exception is Triandis (1977) who argues that it is possible to make a distinction between habitual and intentional behaviors. He argues that novel behaviors will be primarily determined by intention, while repeated behaviors will be primarily determined by habit. Ronis, Yates, and Kirscht (1989) make a similar distinction between habits and decisions, arguing that the performance of repeated behavior is determined by habit rather than social cognitive variables. For example, Dishman (1982) distinguished between the initiation and maintenance of behavior in relation to clinical exercise programs and found that only the initiation of exercise behavior was predicted by social cognitive variables. Sutton (1994) has proposed a further distinction between habits and routines. He argues that many health behaviors commonly considered habitual may be more appropriately considered routines. Sutton (1994) describes a routine as a sequence of behaviors which is repeated on a regular basis. However, what distinguishes them from habits is their need to be supported by self-reminders. It may be possible therefore to make the distinction between occasions when the influence of past behavior is mediated by social cognitive variables and those occasions when it is seen to have a direct influence via habitual responses. In particular, future work should develop measures of habit and routine that are discriminable from frequency of past behavior and outline the processes through which habit and routine determine behavior.

8.01.5.3 Processes by which Cognitions Influence Behavior

One important implication of Ronis et al.'s (1989) distinction between habits and decisions is the suggestion that social cognition models may only predict health behavior under certain conditions. This issue has been addressed by Fazio (1990) in the development of the MODE model of attitude–behavior relationships. He suggests that attitudes (and presumably other cognitions) influence behavior via two distinct processes: a deliberative (or controlled) process and a spontaneous process. Most social cognition models can be labeled as deliberative processing models as they assume that behavior results from a controlled process of conscious deliberation. However, Fazio (1990) argues that individuals may only make a behavioral decision in such a manner when they have the opportunity and motivation to do so. Under other conditions, attitudes which are highly accessible in memory may determine behavior in a spontaneous fashion (Fazio, Powell, & Williams, 1989). When the spontaneous process is operating, an attitude may be automatically activated from memory following the presentation of relevant cues, with the likelihood of activation determined by the accessibility of the attitude. Once activated, the attitude shapes the perception of the attitude object in an automatic, attitude-congruent, fashion. For example, if a positive attitude is activated then this will lead the individual to attend to and notice the positive qualities of the attitude object. This automatic process of selective perception will therefore shape the individual's definition of the event, and thus determine behavior. If the event is defined on the basis of positive perceptions of the attitude object, for example, then approach behaviors will follow. In addition, it is argued that normative guidelines (e.g., social norms or rules) may also influence the definition of the event and thus may help to determine behavior in some situations.

One important feature of the spontaneous processing model is that it outlines one way in which social cognitive variables (i.e., highly accessible attitudes) may determine behavior without systematic deliberation. To date, there has been little research on Fazio's model, although it has been successfully shown that the accessibility of relevant attitudes influences the strength of the relationship between attitudes and behavior (Fazio & Williams, 1986) and that highly accessible attitudes can lead to selective perception (Houston & Fazio, 1989; Fazio & Williams, 1986). Both findings are consistent with the spontaneous processing model. This above work has important implications for SCMs and health behaviors. The SCMs outlined here are deliberative processing models inasmuch as they focus on the conscious processing of information and fail to consider spontaneous or automatic influences on behavior. Hence, current SCMs may provide only a partial account of the social cognitive determinants of behavior, that is, they may only be applicable in situations where the individual has the ability and motivation to engage in deliberative processing of information (Conner, 1993). For many behavioral decisions, simplified or spontaneous decision-making rules may be employed instead (Norman & Conner, 1993). Fazio's (1990) spontaneous model has considerable potential in helping to provide a full account of the cognitive influences on behavior. However, it is clear that most of the empirical work to date has focused on issues surrounding the activation of attitudes and their influence on perception; later components of the model have received less attention.

8.01.5.4 Stage Models of Health Behavior

A number of researchers have suggested that there may be qualitatively different stages in the initiation and maintenance of health behavior, and that to obtain a full understanding of the determinants of health behavior it is necessary to conduct a detailed analysis of the nature of these stages. From a social cognitive perspective, an important implication of this position is that different cognitions may be important at different stages in promoting health behavior.

One of the first stage models was put forward by Prochaska and DiClemente (1984) in their transtheoretical model of change (TTM). Their model has been widely applied to analyze the process of change in alcoholism treatment (DiClemente & Hughes, 1990), smoking cessation (DiClemente et al., 1991), head injury rehabilitation (Lam, McMahon, Priddy, & Gehred-Schutlz, 1988), and psychotherapy (McConnaughly, DiClemente, Prochaska, & Velicer, 1989). In its most recent form, DiClemente et al. (1991) identify five stages of change: precontemplation, contemplation, preparation, action, and maintenance. Individuals are seen to progress through each stage to achieve successful maintenance of a new behavior. Taking the example of smoking cessation, it is argued that in the precontemplation stage smokers are unaware that their behavior constitutes a problem and have no intention to quit. In the contemplation stage, smokers start to think about changing their behavior, but are not committed to try to quit. In the preparation stage, the smoker has an intention to quit and starts to make plans about how to quit. The action stage is characterized by active attempts to quit, and after six months of successful abstinence the individual moves into the maintenance stage characterized by attempts to prevent relapse and to consolidate the newly acquired nonsmoking status. Whilst relative widely applied, the evidence in support of the model and the different stages is at present relatively weak (see Weinstein, Rothman, & Sutton, in press).

Heckhausen (1991) has similarly identified phases in the initiation and maintenance of behavior change; these being the predecisional, postdecisional, actional, and evaluative phases, which follow a similar progressive sequence as that outlined by Prochaska and DiClemente (1984). It is further suggested that different types of cognitions are important in each of these phases. So in the predecisional phase, cognitions about the desirability and feasibility of the behavior are believed to be important determinants of a desire to perform the behavior in question. This phase ends with the formation of an intention to change. In contrast, the decisional phase focuses on the development of plans and ends with the successful initiation of the behavior. In the actional phase the individual focuses on effectively achieving performance of the behavior and ends with the conlusion of the behavior. In the final, evaluative phase the individual compares achieved outcomes with initial goals in order to regulate and maintain behavior. While this four phase model of behavior was not developed for the prediction of health behavior, the potential for its application is clear (see Gollwitzer, 1993). Other stage models have been developed including the health action process approach (Schwarzer, 1992; Schwarzer & Fuchs, 1996), the precaution-adoption process (Weinstein, 1988, Weinstein & Sandman, 1992), and goal setting theory (Bagozzi, 1992, 1993; Bagozzi & Edwards, 1998).

There are two important themes in each of the stage models outlined above. First, they emphasize a temporal perspective with different stages of behavior change. While the models postulate different numbers of stages, they all follow the same pattern from a precontemplation stage through a motivation stage to the initiation and maintenance of behavior. The important point is that these models are dynamic in nature; people move from one stage to another over time. Second, these stage models imply that different cognitions are important at different stages (Sandman & Weinstein, 1993). For example, in the earlier stages information may be processed about the costs and benefits of performing a behavior, while in the later stages cognitions become more focused on the development of plans of action to initiate and support the maintenance of a behavior. This earlier motivational phase is assumed to end with the formation of an intention and only when the level of motivation or intention reaches a particular level is the individual assumed to be likely to move on to later stages.

The main SCMs of health behavior are primarily concerned with people's motivations to perform a health behavior and, as such, provide strong predictions of behavioral intentions (i.e., the end of a motivational stage). Ajzen (1991), for example, reports an average multiple correlation of 0.71 between variables in the TPB and behavioral intention. However, intentions do not always lead to corresponding actions. Studies examining the intention–behavior relationship have reported a wide range of correlations. For example, Sheppard et al. (1988) reported intention–behavior correlations ranging from 0.10 to 0.94. Clearly, many people who intend to perform a behavior fail to do so. However, the SCMs considered do not address the issue of translating intentions into action. They can be conceptualized as static models that stop at the formation of an intention without distinguishing between intenders who become actors and those who do not. As Bagozzi (1993) argues, the variables outlined in the main social cognition models are necessary but not sufficient determinants of behavior. Clearly, a detailed analysis of the social cognitive variables important in translating intentions into action is required to provide a full account of the determinants of health behavior.

Relatively little detailed attention has focused on the cognitive processes underlying the successful implementation of intentions. The main social cognition models contain few measures that account for the intention–behavior gap (Abraham & Sheeran, 1993). The TPB attempts to do this by proposing a direct link between perceived behavioral control and behavior. Thus, people's perceptions about the amount of control they have over a behavior influence the likely performance of behavior independently of their intentions, although an analysis of the volitional processes underlying performance of a health behavior is required, and a number of researchers have focused attention on this issue (e.g., Kuhl, 1985; Schwarzer, 1992; Weinstein, 1988). Here we focus on Gollwitzer's (1993) work on implementation intentions and Bagozzi's (1992) model of goal achievement to highlight the social cognitive variables important in the initiation and maintenance of behavior.

Gollwitzer (1993) made the distinction between goal intentions and implementation intentions. While the former are concerned with intentions to perform a behavior or achieve a goal (i.e., "I intend to achieve x"), the latter are concerned with plans as to when, where, and how the goal intention is to be translated into behavior (i.e., "I intend to initiate the goal-directed behavior x when situation y is encountered"). Goal intentions are most like the intention construct in the TPB, although in the TPB such intentions usually refer to actions or behaviors rather than goals. The important point about implementation intentions is that they commit the individual to a specific course of action when certain environmental conditions are met; in so doing they help translate goal intentions into action. Gollwitzer (1993) argues that by making implementation intentions individuals pass over control to the environment. The environment acts as a cue to action, such that when certain conditions are met, the performance of the intended behavior follows. These ideas have similarities with Weinstein's (1988) "messy desk" analogy, whereby people may have intentions to achieve a number of goals (i.e., "projects") which get "lost" on the "messy desk." Which project is actually worked upon is determined by environmental factors in a similar way as outlined by Gollwitzer (1993).

Gollwitzer (1993, 1996) has compiled a range of experimental evidence to support the view that the making of implementation intentions can aid the performance of intended behavior. To date, the only application of implementation intentions to the prediction of health behavior is a study by Orbell, Hodgkins, and Sheeran (1997) on breast self-examination. At the end of a questionnaire about breast self-examination, half the women were asked to indicate when and where in the next month they intended to perform breast self-examination. A one month follow-up found that 64% of these women had performed breast self-examination that month compared with only 16% of women who had not made an implementation intention, despite no difference in goal intentions. In addition, everyone in the implementation condition who actually performed the behavior reported doing so in response to the environmental cue in the implementation intention and not in response to any other environmental cue. This suggests that the making of an implementation intention can significantly increase the performance of a behavior. In a study on exercise behavior, Kendzierski (1990) found that respondents were more likely to implement their intentions to exercise when they had engaged in some prior planning. Further work needs to establish the utility of implementation intentions in predicting health behavior. However, initial findings are encouraging and suggest that those who make such plans of action are more likely to initiate and maintain behavior.

The work of Gollwitzer (1993) is important in that it identifies one way in which goal intentions may be translated into behavior. A similar but more comprehensive approach has been put forward by Bagozzi (Bagozzi, 1992, 1993; Bagozzi & Edwards, 1998; Bagozzi & Warshaw, 1990) in his model of goal achievement. He focuses on goal-directed behavior and argues that to initiate behavior individuals need to form an "intention to try" to achieve their desired goal. Once an intention to try has been formed, the individual focuses on the means, or instrumental acts, by which to attempt to achieve the desired goal. Considering the example of weight loss, a number of instrumental acts can be identified, including restricting between-meal foods, reducing overall calorie consumption, avoidance of high calorie foods, exercise, and so on. Bagozzi (1993) argues that for each of these instrumental acts, three appraisal tasks are performed. First, the individual considers the extent to which they are confident that they could perform the instrumental act (i.e., specific self-efficacies). Second, the likelihood that the instrumental act will help in achieving the desired goal is assessed (i.e., instrumental beliefs). Third, the individual considers an affective response towards the instrumental act (i.e., affect towards means). Once an individual initiates efforts to achieve a goal, there are a number of cognitive activities that support the successful initiation and maintenance of goal-directed behavior. First, the individual can develop plans in order to ensure that instrumental acts are performed. This involves identifying the situation or triggering conditions under which the instrumental act is performed (Bagozzi & Warshaw, 1990). This idea that certain environmental conditions may trigger behavior has a clear overlap with Gollwitzer's (1993) work on implementation intentions and Weinstein's (1988) "messy desk" analogy. One way in which plans are more likely to be acted upon is through the development of scripts or cognitive rehearsal, whereby the individuals imagine themselves performing the instrumental act (Anderson, 1983). Another is through the use of precommitting devices whereby the behavior is made more likely by precommitting oneself to it (e.g., avoiding eating butter at home by not having butter in the house). Bagozzi (1993) also proposes that ongoing behavior has to be monitored to ensure, for example, that the instrumental acts achieve their objectives. If any unforeseen impediments are encountered then these need to be taken into consideration and any future plans modified accordingly. These ideas overlap with Kuhl's (1985) theory of action control which identifies a number of processes by which individuals attempt to control their actions and achieve their goals. These processes may be particularly important in allowing individuals to overcome temptations to break their new behavior (see Loewenstein, 1996). Finally, goal-directed behavior is likely to be stronger and more persistent if the individual has a strong sense of commitment to the decision to try to achieve the goal and the means to achieve it.

8.01.5.5 Theoretical Integrations and Future Directions

The above research demonstrates some of the ways in which research into health behaviors based upon social cognitive approaches is developing. Here we outline possible directions for the future development of an integrative social cognition model of health behavior, outline some of the basic requirements for such a model, and indicate some fruitful avenues for future research.

It is clear that to fully explain health behavior it is necessary to develop a more dynamic model that examines different stages or phases in the contemplation, initiation, and maintenance of behavior. What is being proposed is an integration of current SCMs (such as the TPB) with stage models of health behavior (such as the TTM). Several authors have recommended such an integration (e.g., Courneya, Nigg, & Estabrooks, 1998; Godin, Desharnais, Valois, & Bradet, 1995; Marcus, Eaton, Rossi, & Harlow, 1994). Though the stage models considered in this chapter have suggested differing number of stages, it is likely that an integrative model should address at least four or five main stages: precontemplation, contemplation, planning, action, and maintenance (Norman & Conner, 1996a).

One implication of the identification of different stages is that different cognitive variables may be important in ensuring movement from one stage to the next. In the first stage, the individuals are not thinking about making a change to their behavior. However, this stage may be brought to an end by a range of cues to action, as outlined in the HBM, which may motivate the individual to start thinking about performing a health-related behavior. One such cue to action may be perceived threat (i.e., perceived susceptibility and perceived severity). While perceived susceptibility and perceived severity are seen to be important determinants of behavior in the HBM, research with these dimensions has tended to show that they are relatively weak predictors of behavior. However, as Schwarzer (1992) has argued, it may be more appropriate to consider these variables to have an indirect or more distal influence on behavior. Thus they may act as a cue to action, motivating the individual to start deliberating over performing a health-related behavior, and thus ensuring movement from the first to the second stage.

In the second stage, a decision-making or motivation stage, the individual is thinking about adopting a new behavior, and the stage ends when the individual forms an intention to perform the behavior. To date, most social cognition models have been primarily concerned with this stage. These models distinguish between three distinct determinants of individuals' intentions to perform a health behavior. First, are outcome expectancies, which focus on the perceived consequences of performing a behavior. These expectancies may also cover the notion of behavioral beliefs as considered in the TPB and include anticipated affective reactions. Second, are normative influences which are primarily tapped by the subjective norm and normative belief components of the TPB. This group of variables could also include descriptive norms, moral norms, and perceived social support. The third influence on individuals' intentions to perform a behavior is control beliefs (or self-efficacy expectancies) and may be based on a consideration of perceived barriers (HBM) and control beliefs (TPB). In the motivation stage, it is likely that other variables may have a more distal influence on behavioral intention via the variables outlined above. For example, health locus of control beliefs may help shape self-efficacy expectancies, self-identity and health values may influence the interpretation of the potential consequences of a behavior, and past behavior or experience may provide information which is used to determine the ease or difficulty of performing a behavior (i.e., self-efficacy).

Once a behavioral intention has been formed, it has to be translated into behavior. In the third stage, the individual is therefore concerned with planning; focusing on the specific actions or instrumental acts that need to be performed and the resources required to support them. Thus, a number of authors have highlighted the importance of action plans in this stage (Bagozzi & Warshaw, 1990; Schwarzer, 1992). Similarly, Gollwitzer (1993) focuses on implementation intentions which help ensure performance of the target behavior. Despite slight differences between definitions of these two concepts, both emphasize the need to construct fairly detailed plans of action in order to bridge the intention–behavior gap. As Schwarzer (1992) argues, self-efficacy may have an important role to play in the development and implementation of such plans, as might self-identity (Sparks & Shepherd, 1992) and a sense of commitment (Bagozzi, 1993). The planning stage is brought to an end when the individual initiates behavior.

In the fourth stage, the individual has to ensure that the behavior is successfully enacted. Various cognitive processes which are concerned with the monitoring and controlling of behavior may be important in this stage. Schwarzer (1992), for example, highlights the need for action control in which the behavior is re-evaluated against initial goals in order to regulate and maintain behavior (see also Bagozzi & Warshaw, 1990; Heckhausen, 1991). As with the development of action plans, self-efficacy, self-identity, and commitment may all be important variables in ensuring that behavior is maintained. In addition, Kuhl (1985) has argued that there may be individual differences in people's propensity to engage in monitoring activities that may account for why behavior is not always maintained. These monitoring activities help to ensure that the behavior is successfully completed. The same activities will be important in the fifth, maintenance stage for ongoing behaviors, where the individual is concerned with ensuring that the behavior is successfully repeated.

8.01.6 CONCLUSIONS

Health behaviors have been demonstrated to have important consequences for both the quality and length of life through influencing various disease outcomes. We have attempted to justify the interest in understanding health behaviors as a basis for attempting to change their occurrence in order to increase both length and quality of life. SCMs provide one approach to understanding health behavior in describing the important social cognitive variables underlying such behaviors. We believe that these models provide an important way of achieving these aims by providing a means for identifying useful targets for persuasion. It would seem that there is already sufficient literature to support the contention that important determinants of health behaviors are identified in these models. Further refinement and development of these models along the lines we have suggested and reviewed may lead to even better predictions of behavior and greater understanding of health behavior and how individuals may be encouraged to change.

However, persuasive messages targeted at relevant cognitions identified by SCMs may not be sufficient to produce the major behavior change necessary for health benefits to accrue. It may be that strategies which employ multiple level interventions which take account not only of the psychosocial factors influencing performance of the behavior (derived from SCMs) but also models of the process of persuasion of how people change and the context in which changes are made will be important (Glanz et al., 1990; Hockbaum & Lorig, 1992; Winett, 1985).

8.01.7 REFERENCES

Abraham, C., & Sheeran, P. (1993). Inferring cognitions, predicting behaviour: Two challenges for social cognition models. *Health Psychology Update, 14,* 18–23.

Abraham, C. S., Sheeran, P., Abrams, D., & Spears, R. (1994). Exploring teenagers' adaptive and maladaptive thinking in relation to the threat of HIV infection. *Psychology and Health, 9,* 253–272.

Adler, N., & Matthews, K. (1994). Health psychology: Why do some people get sick and some stay well? *Annual Review of Psychology, 45,* 229–259.

Aguirre-Molina, M., & Gorman, D. M. (1996). Community-based approaches for the prevention of alcohol, tobacco, and other drug use. *Annual Review of Public Health, 17,* 337–358.

Aho, W. R. (1979). Smoking, dieting and exercise: Age differences in attitudes and behavior relevant to selected health belief model variables. *Rhode Island Medical Journal, 62,* 95–102.

Aiken, L. S., West, S. G., Woodward, C. K., Reno, R. R., & Reynolds, K. D. (1994). Increasing screening mammography in asymptomatic women: Evaluation of a second-generation, theory-based program. *Health Psychology, 13,* 526–538.

Ajzen, I. (1988). *Attitudes, personality and behavior.* Milton Keynes, UK: Open University Press.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50,* 179–211.

Ajzen, I. (1996). The directive influence of attitudes on behavior. In P. Gollwitzer & J. A. Bargh (Eds.), *Psychology of action* (pp. 385–403). New York: Guilford Press.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior.* Englewood-Cliffs, NJ: Prentice-Hall.

Allied Dunbar Fitness Survey (1992). London: Health Education Authority.

Allison, K. R. (1987). *Perceived control as a determinant of preventive health behaviour for heart disease and lung cancer.* Unpublished doctoral dissertation, University of Toronto.

Amler, R. W., & Dull, H. B. (1987). *Closing the gap.* New York: Oxford University Press.

Anderson, C. A. (1983). Imagination and expectation: The effect of imagining behavioral scripts on personal intentions. *Journal of Personality and Social Psychology, 45,* 293–305.

Armitage, C. J., & Conner, M. (in press). Distinguishing perceptions of control from self-efficacy: Predicting consumption of a low fat diet using the theory of planned behavior. *Journal of Applied Social Psychology.*

Ashworth, H. W. (1963). An experiment in presymptomatic diagnosis. *Journal of the Royal College of General Practitioners, 6,* 71.

Aspinwall, L. G., Kemeny, M. E., Taylor, S. E., Schneider, S. G., & Dudley, J. P. (1991). Psychosocial predictors of gay men's AIDS risk-reduction behavior. *Health Psychology, 10,* 432–444.

Babrow, A. S., Black, D. R., & Tiffany, S. T. (1990). Beliefs, attitudes, intentions, and a smoking-cessation program: a planned behavior analysis of communication campaign development. *Health Communication, 2,* 145–163.

Baer, J. S., Marlatt, G. A., Kivlahan, D. R., Fromme, K., Larimer, M. E., & Williams, E. (1992). An experimental test of three methods of alcohol risk reduction with young adults. *Journal of Consulting and Clinical Psychology, 60,* 974–979.

Bagozzi, R. P. (1992). The self-regulation of attitudes, intentions and behavior. *Social Psychology Quarterly, 55,* 178–204.

Bagozzi, R. P. (1993). On the neglect of volition in consumer research: A critique and proposal. *Psychology and Marketing, 10,* 215–237.

Bagozzi, R. P., & Edwards, E. A. (1998). Goal setting and goal pursuit in the regulation of body weight. *Psychology and Health, 13.*

Bagozzi, R. P., & Warshaw, P. R. (1990). Trying to consume. *Journal of Consumer Research, 17,* 127–140.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84,* 191–215.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37,* 122–147.

Bandura, A. (1986). *Social foundations of thought and action: A cognitive social theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1991). Self-efficacy mechanism in physiological activation and health-promoting behavior. In J. Madden (Ed.), *Neurobiology of learning, emotion and affect* (pp. 229–270). New York: Raven Press.

Basen-Engquist, K., & Parcel, G. S. (1992). Attitudes, norms, and self-efficacy: A model of adolescents' HIV-related sexual risk behavior. *Health Education Quarterly, 19,* 263–277.

Beale, D. A., & Manstead, A. S. R. (1991). Predicting mothers' intentions to limit frequency of infants' sugar intake: Testing the theory of planned behavior. *Journal of Applied Social Psychology, 21,* 409–431.

Beck, K. H. & Lund, A. K. (1981). The effects of health threat seriousness and personal efficacy upon intentions and behavior. *Journal of Applied Social Psychology, 11,* 401–415.

Beck, K. H. (1981). Driving while under the influence of alcohol: Relationship to attitudes and beliefs in a college population. *American Journal of Drug and Alcohol Abuse, 8,* 377–388.

Beck, L., & Ajzen, I. (1991). Predicting dishonest actions using the theory of planned behavior. *Journal of Research in Personality, 25,* 285–301.

Becker, M. H. (1974). The health belief model and sick role behavior. *Health Education Monographs, 2,* 409–419.

Becker, M. H., Kaback, M. M., Rosenstock, I. M., & Ruth, M. (1975). Some influences on public participation in a genetic screening program. *Journal of Community Health, 1,* 3–14.

Becker, M. H., & Maiman, L. A. (1983). Models of health-related behavior. In D. Mechanic (Ed.), *Handbook of health, health care and the health professions* (pp. 539–568). New York: Free Press.

Becker, M. H., Radius, S. M., & Rosenstock, I. M. (1978). Compliance with a medical regimen for asthma: A test of the health belief model. *Public Health Reports, 93,* 268–277.

Becker, M. H., & Rosenstock, I. M. (1987). Comparing social learning theory and the health belief model. In W. B. Ward (Ed.), *Advances in health education and promotion* (Vol. 2, pp. 245–249). Greenwich, CT: JAI Press.

Belloc, N. B. (1973). Relationship of health practices to mortality. *Preventive Medicine, 2,* 67–81.

Belloc, N. B., & Breslow, L. (1972). Relationship of physical health status and health practices. *Preventive Medicine, 9,* 409–421.

Bengel, J., Belz-Merk, M., & Farin, E. (1996). The role of risk perception and efficacy cognitions in the prediction of HIV-related preventive behavior and condom use. *Psychology and Health, 11,* 505–525.

Bentler, P. M., & Speckart, G. (1979). Models of attitude-behavior relations. *Psychological Review, 86,* 452–464.

Bernier, M. & Avard, J. (1986). Self-efficacy, outcome and attrition in a weight reduction program. *Cognitive Therapy and Research, 10,* 319–338.

Best, J. A. (1975). Tailoring smoking withdrawal procedures to personality and motivational differences. *Journal of Consulting and Clinical Psychology, 43,* 1–8.

Biddle, B. J. Bank, B. J., & Slavings, R. L. (1987). Norms, preferences, identities and retention decisions. *Social Psychology Quarterly, 50,* 322–337.

Black Report (1988). *Inequalities in health: the Black report and the health divide.* London: DHSS.

Blair, S. N., Goodyear, N. N., Gibbons, L. W., & Cooper, K. H. (1984). Physical fitness and incidence of hypertension in healthy normotensive men and women. *Journal of the American Medical Association, 252,* 487–490.

Blane, D., Smith, G. D., & Bartley, M. (1990). Social class differences in years of potential life lost: Size trends, and principal causes. *British Medical Journal, 301,* 429–432.

Blaxter, M. (1990). *Health and lifestyles.* London: Tavistock.

Boer, I., & Seydel, E. R. (1996). Protection motivation theory. In M. Conner & P. Norman (Eds.), *Predicting health behaviour* (pp. 95–120). Buckingham, UK: Open University Press.

Boldero, J., Moore, S., & Rosenthal, D. (1992). Intention, context, and safe sex: Australian adolescents' responses to AIDS. *Journal of Applied Social Psychology, 22,* 1374–1396.

Borgida, E., Conner, C., & Manteufel, L. (1992). Understanding living kidney donation: a behavioral decision perspective. In S. Spacapan & S. Oskamp (Eds.), *Helping and being helped: Naturalistic studies* (pp. 183–212). Newbury Park, CA: Sage.

Boulton, M., Schramm Evans, Z., Fitzpatrick, R., & Hart, G. (1991). Bisexual men: Women, safer sex, and HIV infection. In P. Aggleton, P. M. Davies, & G. Hart (Eds.), *AIDS: Responses, policy and care* (pp. 55–67). London: Falmer Press.

Boyd, B., & Wandersman, A. (1991). Predicting undergraduate condom use with the Fishbein and Ajzen and the Triandis attitude-behavior models: Implications for public health interventions. *Journal of Applied Social Psychology, 21,* 1810–1830.

Bradley, C., Lewis, K. S., Jennings, A. M., & Ward, J. D. (1990). Scales to measure perceived control developed specifically for people with tablet-treated diabetes. *Diabetic Medicine, 7,* 685–694.

Brawley, L. R. (1993). The practicality of using social psychological theories for exercise and health research and intervention. *Journal of Applied Sport Psychology, 5,* 99–115.

Breslow, L., & Enstrom, J. E. (1980). Persistence of health habits and their relationship to mortality. *Preventive Medicine, 9,* 469–483.

British Medical Journal (1982). *Alcohol problems.* London: BMJ Publications.

Brock, B. M., Haefner D. P., & Noble, D. S. (1988). Alameda County Redux: Replication in Michigan. *Preventive Medicine, 17,* 483–495.

Brouwers, M. C., & Sorrentino, R. M. (1993). Uncertainty orientation and protection motivation theory: The role of individual differences in health compliance. *Journal of Personality and Social Psychology, 65,* 102–112.

Brown, J. D. (1991). Staying fit and staying well: Physical fitness as a moderator of life stress. *Journal of Personality and Social Psychology, 60,* 555–561.

Brown, N., Muhlenkamp, A., Fox, L., & Osborn, M. (1983). The relationship among health beliefs, health values, and health promotion activity. *Western Journal of Nursing Research, 5,* 155–163.

Brubaker, R. G., & Fowler, C. (1990). Encouraging College males to perform testicular self-examination: Evaluation of a persuasive message based on the revised theory of reasoned action. *Journal of Applied Social Psychology, 17,* 1411–1422.

Bundek, N. I., Marks, G., & Richardson, J. L. (1993). Role of health locus of control beliefs in cancer screening of elderly Hispanic women. *Health Psychology, 12,* 193–199.

Burk, C., & Kimiecik, J. (1994). Examining the relationship among locus of control, health value and exercise. *Health Values, 18,* 14–23.

Butts, S. V., & Chotlas, J. A. (1973). A comparison of alcoholics and nonalcoholics on perceived locus of control. *Quarterly Journal of Studies on Alcohol, 34,* 1327–1332.

Calnan, M. (1989). Control over health and patterns of health-related behaviour. *Social Science and Medicine, 29,* 131–136.

Carlson, B., & Petti, K. (1989). Health locus of control and participation in physical activity. *American Journal of Health Promotion, 3,* 32–37.

Chaiken, S., Lieberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212–252). New York: Guilford Press.

Chambliss, C. A., & Murray, E. J. (1979a). Efficacy attribution, locus of control, and weight loss. *Cognitive Therapy and Research, 3,* 349–353.

Chambliss, C. A., & Murray, E. J. (1979b). Cognitive procedures for smoking reduction: symptom attribution versus efficacy attribution. *Cognitive Therapy and Research, 3,* 91–95.

Champion, V. L. (1984). Instrument development of the health belief model constructs. *Advances in Nursing Science, 6,* 73–85.

Charng, H. W., Piliavin, J. A., & Callero, P. L. (1988). Role identity and reasoned action in the prediction of repeated behavior. *Social Psychology Quarterly, 51,* 303–317.

Chen, M. M., & Land, K. C. (1986). Testing the health belief model: LISREL analysis of alternative models of the causal relationships between health beliefs and preventive dental behavior. *Social Psychology Quarterly. 49,* 45–60.

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology, 58,* 1015–1026.

Colletti, G., Supnick, J. A., & Payne, T. J. (1985). The smoking self-efficacy questionnaire (SSEQ): Preliminary scale development and validation. *Behavioral Assessment, 7,* 249–260.

Colliver, J. D., & Mallin, H. (1986). State an national trends in alcohol-related mortality, 1975–1982. *Alcohol Health and Research World, 10,* 60–64.

Committee on Medical Aspects of Food Policy (1991). Dietary reference values for food energy and nutrients for the United Kingdom. *Report on Health and Social Subjects, 41.* London: HMSO.

Condiotte, M. M., & Lichtenstein, E. (1981). Self-efficacy and relapse in smoking cessation programs. *Journal of Consulting and Clinical Psychology, 49,* 648–658.

Connelly, C. E. (1984). Compliance with outpatient lithium therapy. *Perspectives in Psychiatric Care, 22,* 44–50.

Conner, M. (1993). Pros and cons of social cognition models in health behaviour. *Health Psychology Update, 14,* 24–31.

Conner, M., Martin, E., Silverdale, N., & Grogan, S. (1996). Dieting in adolescence: An application of the theory of planned behavior. *British Journal of Health Psychology, 1,* 315–325.

Conner, M., & Norman, P. (1994). Comparing the health belief model and the theory of planned behaviour in health screening. In D. R. Rutter & L. Quine (Eds.), *Social psychology and health: European perspectives* (pp. 1–24). Aldershot, UK: Avebury.

Conner, M., & Norman, P. (Eds.) (1996a). *Predicting health behaviour.* Buckingham, UK: Open University Press.

Conner, M., & Norman, P. (1996b). The role of social cognition in health behaviours. In M. Conner & P. Norman (Eds.), *Predicting health behaviour* (pp. 1–22). Buckingham, UK: Open University Press.

Conner, M., & Sparks, P. (1996). The theory of planned behaviour and health behaviours. In M. Conner & P. Norman (Eds.), *Predicting health behaviour* (pp. 121–162). Buckingham, UK: Open University Press.

Consensus Development Conference on Lowering Blood Cholesterol to Prevent Heart Disease (1985). Lowering blood cholesterol to prevent heart disease. *Journal of the Australian Medical Association, 253,* 2080–2086.

Corcoran, K. J. (1991). Efficacy, "skills," reinforcement, and choice behavior. *American Psychologist, 46,* 155–157.

Costello, R. M., & Manders, K. R. (1974). Locus of control and alcoholism. *British Journal of Addiction, 69,* 11–17.

Courneya, K. S., Nigg, C. R., & Estabrooks, P. A. (1998). Relationships among the theory of planned behavior, stages of change, and exercise behavior in older persons over a three year period. *Psychology & Health, 13.*

Cox, B. D, Huppert, F. A., & Whichelow, M. J. (1993). *The health and lifestyles survey: Seven years on.* Aldershot, UK: Dartmouth.

Dean, K. (1991). Relationships between knowledge and belief variables and health maintenance behaviours in a Danish population over 45 years of age. *Journal of Ageing and Health, 3,* 386–406.

Desharnais, R., Bouillon, J., & Godin, G. (1986). Self-efficacy and outcome expectations as determinants of exercise adherence. *Psychological Reports, 59,* 1155–1159.

DeVellis, B. M., Blalock, S. J., & Sandler, R. S. (1990). Predicting participation in cancer screening: the role of perceived behavioral control. *Journal of Applied Social Psychology, 20,* 639–660.

DeVito, A. J., Bogdanowicz, J., & Reznikoff, M. (1982). Actual and intended health-related information seeking and health locus of control. *Journal of Personality Assessment, 46,* 63–69.

DeVries, H., Backbier, E., Kok, G., & Dijkstra, M. (1995). The impact of social influences in the context of attitude, self-efficacy, intention, and previous behavior as predictors of smoking onset. *Journal of Applied Social Psychology, 25,* 237–257.

DeVries, H., Dijkstra, M., & Kuhlman, P. (1988). Self-efficacy: The third factor besides attitude and subjective norm as a predictor of behavioural intentions. *Health Education Research, 3,* 273–282.

DiClemente, C. C., & Hughes, S. O. (1990). Stages of change profiles in outpatient alcoholism treatment. *Journal of Substance Abuse, 2,* 217–235.

DiClemente, C. C., Fairhurst, S. K., Velasquez, M. M., Prochaska, J. O., Velicer, W. F., & Rossi, J. S. (1991). The process of smoking cessation: An analysis of precontemplation, contemplation, and preparation stages of change. *Journal of Consulting and Clinical Psychology, 59,* 295–304.

Dishman, R. K. (1982) Compliance/adherence in health-related exercise. *Health Psychology, 1,* 237–267.

Doll, R., Peto, R., Wheatley, K., Gray, R., & Sutherland, I. (1994). Mortality in relation to smoking: 40 years' observations on male British doctors. *British Medical Journal, 309,* 901–911.

Donovan, D. M., & O'Leary, M. R. (1978). The drinking-related locus of control scale. *Journal of Studies on Alcohol, 39,* 759–784.

Duffy, M. E. (1988). Determinants of health promotion in midlife women. *Nursing Research, 37,* 358–362.

Dzewaltowski, D. A., Noble, J. M., & Shaw, J. M. (1990). Physical activity participation—social cognitive theory versus the theories of reasoned action and planned behavior. *Journal of Sport and Exercise Psychology, 12,* 388–405.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 51,* 380–417.

Edwards, W. (1992). *Utility theories: measurements and applications*. Boston: Kluwer.

Expert Panel on Detection, Evaluation, and Treatment of High Cholesterol in Adults (1993). Summary of the second report of the National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high cholesterol in adults. *Journal of the American Medical Association, 269,* 3015–3023.

Farin, E. (1994). *Eine Metaanalyse Empirischer Studien zum Prädiktiven Wert Kognitiver Variablen der HIV-Bezogenen Risikowahrnehmung und-Verabeitung für das HIV-Risikoverhalten*. Dissertation. Universität Freiburg.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 75–109). San Diego, CA: Academic Press.

Fazio, R. H., & Williams, C. J. (1986). Attitude accessibility as a moderator of the attitude-perception and attitude-behavior relations: An investigation of the 1984 presidential election. *Journal of Personality and Social Psychology, 51,* 505–514.

Fazio, R. H., Powell, M. C., & Williams, C. J. (1989). The role of attitude accessibility in the attitude-to-behavior process. *Journal of Consumer Research, 16,* 280–288.

Feather, N. T. (1982). *Expectations and actions: Expectancy-value models in psychology*. Hillsdale, NJ: Erlbaum.

Fife-Schaw, C.R., & Breakwell, G. M. (1992). Estimating sexual behaviour parameters in the light of AIDS: A review of recent UK studies of young people. *AIDS Care, 4,* 187–201.

Fishbein, M. (1993). Foreword. In D. J. Terry, C. Gallois, & M. McCamish (Eds.). *The theory of reasoned action: Its application to AIDS-related behavior* (pp. xv–xxv). Oxford, UK: Pergamon.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior*. New York: Wiley.

Folkins, C. H., & Sime, W. E. (1981). Physical fitness training and mental health. *American Psychologist, 36,* 373–389.

Forrest, A.P.M. (1986). *Breast cancer screening*. London: HMSO.

Fredricks, A. J., & Dossett, D. L. (1983). Attitude–behavior relations: A comparison of the Fishbein–Ajzen and Bentler–Speckart models. *Journal of Personality and Social Psychology, 45,* 501–512.

Freeling, P. (1965). Candidates for exfoliative cytology of the cervix. *Journal of the Royal College of General Practitioners, 10,* 261.

Friedman, G. D., Dales, L. G., & Ury, H. K. (1979). Mortality in middle-aged smokers and non-smokers. *New England Journal of Medicine, 300,* 213–217.

Fries, J. F., Green, L. W., & Levine, S. (1989). Health promotion and the compression of mortality. *Lancet, i,* 481–483.

Frisch, D., & Clemen, R. T. (1994). Beyond expected utility: rethinking behavioral decision making. *Psychological Bulletin, 116,* 46–54.

General Household Survey (1989). London: OPCS.

General Household Survey (1992). London: OPCS.

General Household Survey (1993). London: OPCS.

General Household Survey (1994). London: OPCS.

Georgiou, A., & Bradley, C. (1992). The development of a smoking specific locus of control scale. *Psychology and Health, 6,* 227–240.

Gerrard, M., Gibbons, F. X., & Bushman, B. J. (1996). Relation between perceived vulnerability to HIV and precautionary sexual behavior. *Psychological Bulletin, 119,* 390–409.

Gianetti, V. J., Reynolds, J., & Rihen, T. (1985). Factors which differentiate smokers from ex-smokers among cardiovascular patients: A discriminant analysis. *Social Science and Medicine, 20,* 241–245.

Gibbons, F. X., & Gerrard, M. (1995). Predicting young adults' health risk behavior. *Journal of Personality and Social Psychology, 69,* 505–517.

Gibbons, F. X., & Gerrard, M. (1997). Health images and their effects on health behavior. In B. P. Buuuk & F. X. Gibbons (Eds.), *Health, coping, and social comparisons* (pp. 63–94). Hillsdale, NJ: Erlbaum.

Gibbons, F. X., Gerrard, M., Ouelette, J., & Burzette, R. (1998). Cognitive antecedents to adolescent health risk: Discriminating between behavioral intention and behavioral willingness. *Psychology and Health, 13.*

Glanz, K., Lewis, F. M., & Rimmer, B. K. (Eds.) (1990). *Health behavior and health education: Theory, research and practice*. San Francisco: Josey-Bass.

Godin, G., Desharnais, R., Valois, P., & Bradet, R. (1995). Combining behavioral and motivational dimensions to identify and characterize the stages in the process of adherence to exercise. *Psychology and Health, 10,* 333–344.

Godin, G., & Shepherd, R. J. (1987). Psychosocial factors influencing intentions to exercise in a group of individuals ranging from 45 to 74 years of age. In M. E. Berridge & G. R. Ward (Eds.), *International perspectives on adapted physical activity* (pp. 243–249). Champaign, IL: Human Kinetics Publishers.

Godin, G., Valois, P., & Lepage, L. (1993). The pattern of influence of perceived behavioral control upon exercise behavior: An application of Ajzen's theory of planned behavior. *Journal of Behavioral Medicine, 16,* 81–102.

Godin, G., Valois, P., Lepage, L., & Desharnais, R. (1992). Predictors of smoking behaviour—An application of Ajzen's theory of planned behaviour. *British Journal of Addiction, 87,* 1335–1343.

Gollwitzer, P. M. (1990). Action phases and mind-sets. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: foundations of social behaviour* (Vol. 2, pp. 53–92). New York: Guilford Press.

Gollwitzer, P. M. (1993). Goal achievement: The role of intentions. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 4, pp. 141–185). Chichester, UK: Wiley.

Gollwitzer, P. M. (1996). The volitional benefits of planning. In P. Gollwitzer & J. A. Bargh (Eds.), *Psychology of action* (pp. 287–312). New York: Guilford Press.

Gorsuch, R. L., & Ortberg, J. (1983). Moral obligation and attitudes: Their relation to behavioral intentions. *Journal of Personality and Social Psychology, 44,* 1025–1028.

Gottlieb, N. H., & Baker, J. A. (1986). The relative influence of health beliefs, parental and peer behaviours and exercise program participation on smoking, alcohol use and physical activity. *Social Science and Medicine, 22,* 915–927.

Gottlieb, N. H., & Green, L. W. (1984). Life events, social network, life-style and health: An analysis of the 1979 National survey of personal health practices and consequences. *Health Education Quarterly, 11,* 91–105.

Graham, H. (1987). Women's smoking and family health. *Social Science and Family Health, 25,* 47–56.

Gregg, I. (1966). The recognition of early chronic bronchitis. *Journal of the American Medical Association, 176,* 1114–1116.

Grube, J. W., Morgan, M., & McGee, S. T. (1986). Attitudes and normative beliefs as predictors of smoking intentions and behaviours: A test of three models. *British Journal of Social Psychology, 25,* 81–93.

Hallal, J. (1982). The relationship of health beliefs, health locus of control, and self-concept to the practice of breast self-examination in adult women. *Nursing Research, 31,* 137–142.

Harris, R., & Lynn, M. W. (1985). Health beliefs, compliance and control of diabetes mellitus. *Southern Medical Journal, 2,* 162–166.

Harrison, J. A., Mullen, P. D., & Green, L. W. (1992). A meta-analysis of studies of the Health Belief Model with adults. *Health Education Research, 7,* 107–116.

Hartman, P. E., & Becker, M. H. (1978). Non-compliance with prescribed regimen among chronic haemodialysis patients. *Journal of Dialysis and Transplantation, 7,* 978–985.

Hayes, D., & Ross, C. E. (1987). Concern with appearance, health beliefs and eating habits. *Journal of Health and Social Behaviour, 28,* 120–130.

Health Promotion Authority for Wales (1990) *Health in Wales*. Cardiff, UK: Author.

Heckhausen, H. (1991). *Motivation and action*. Berlin, Germany: Springer-Verlag.

Hennekens, C. H. (1983). Alcohol. In N. M. Kaplan & J. Stamler (Eds.) *Prevention of coronary heart disease: practical management of the risk factors* (pp. 130–138). Philadelphia: Saunders.

Hester, N. R., & Macrina, D. M. (1985). The health belief model and the contraceptive behavior of college women: Implications for health education. *Journal of American College Health, 33,* 245–252.

Hochbaum, G. M. (1958). *Public participation in medical screening programs: A socio-psychological study*. Public Health Service Publication No. 572. Washington, DC: US Government Printing Office.

Hockbaum, G. M., & Lorig, K. (1992). Roles and uses of theory in health education practice. *Health Education Quarterly, 19,* 289–412.

Hodgkins, S. & Orbell, S. (1998). Can protection motivation theory predict behaviour? A longitudinal study exploring the role of previous behaviour. *Psychology and Health, 13*.

Hofstetter, C. R., Sallis, J. F., & Howell, M. F. (1990). Some health dimensions of self-efficacy: analysis of theoretical specificity. *Social Science and Medicine, 31,* 1051–1056.

Honkanen, R., Ertama, L., Kuosmanen, P., Linnoila, M., Ahla, A., & Visuri, T. (1983). The role of alcohol in accidental falls. *Journal of Studies on Alcohol, 44,* 231–245.

Horwitz, M. B., Hindi-Alexander, M., & Wagner, T. J. (1985). Psychosocial mediators of abstinence, relapse, and continued smoking: A one-year follow-up of a minimal intervention. *Addictive Behaviors, 10,* 29–39.

Hounsa, A. M., Godin, G., Alihonou, E., Valois, P., & Girard, J. (1993). An application of Ajzen's theory of planned behaviour to predict mothers' intention to use oral rehydration therapy in a rural area of Benin. *Social Science and Medicine, 37,* 253–261.

Houston, D. A., & Fazio, R. H. (1989). Biased processing as a function of attitude accessibility: Making objective judgements subjectively. *Social Cognition, 7,* 51–66.

Huba, G. J., Wingard, J. A., & Bentler, P. M. (1981). Intentions to use drugs among adolescents: A longitudinal analysis. *Journal of the Additions, 16,* 331–339.

Janis, I. L. (1967). Effects of fear arousal on attitude change: recent developments in theory and experimental research. In L. Berkowitz (Ed.) *Advances in experimental social psychology* (Vol. 3, pp. 166–224). New York: Academic Press.

Janz, N. K., & Becker, M. H. (1984). The health belief model: A decade later. *Health Education Quarterly, 11,* 1–47.

Jeffrey, R. W., Bjornson-Benson, W. M., Rosenthal, B. S., Lindquist, R. A., Kurth, C. L., & Johnson, S. L. (1984). Correlates of weight loss and its maintenance over two years of follow-up among middle-aged men. *Preventive Medicine, 13,* 155–168.

Jonas, K. (1993). Expectancy-value models of health behaviour: An analysis by conjoint measurement. *European Journal of Social Psychology, 23,* 167–183.

Kaplan, G. D., & Cowles, A. (1978). Health locus of control and health value in the prediction of smoking behavior. *Health Education Monographs, 6,* 129–137.

Kavanagh, D. J., Pierce, J., Lo, S. K., & Shelley, J. (1993). Self-efficacy and social support as predictors of smoking after a quit attempt. *Psychology and Health, 8,* 231–242.

Kelley, J. A., St. Lawrence, J. S., Brasfield, T. L., Lemke, A., Amidei, T., Roffman, R. E., Hood, H. V., Smith, J. E., Kilgore, H., & McNeill, C. (1990). Psychological factors that predict AIDS high-risk versus AIDS precautionary behavior. *Journal of Consulting and Clinical Psychology, 58,* 117–120.

Kendzierski, D. (1990). Decision making versus decision implementation: An action control approach to exercise adoption and adherence. *Journal of Applied Social Psychology, 20,* 27–45.

Keys, A. (1970). Coronary heart disease in seven countries. *Circulation, 41*(Suppl. 1), entire issue.

King, A. C., Blair, S. N., Bild, D. E., Dishman, R. K., Dubbert, P. M., Marcus, B. H. Oldridge, N. B., Paffenbarger, R. S., Powell, K. E., & Yeager, K. K. (1992). Determinants of physical activity and interventions in adults. *Medicine and Science in Sports and Exercise, 24,* 1079–1092.

King, J. (1982). The impact of patients' perceptions of high blood pressure on attendance at screening. *Social Science and Medicine, 16,* 1079–1091.

Kinket, B., Paans, L., & Verplanken, B. (1992). Patient education by general practitioners: Application of the theory of planned behavior. *Gedrag and Gezondheid Tijdschrift voor Psychologie and Gezondheid, 20,* 187–195.

Kirsch, I. (1986). Early research on self-efficacy: What we already know without knowing we knew. *Journal of Social and Clinical Psychology, 4,* 339–358.

Kirscht, J. P., Becker, M. H., & Eveland, P. (1976). Psychological and social factors as predictors of medical behaviour. *Journal of Medical Care, 14,* 422–431.

Kok, G., de Vries, H., Mudde, A. N., & Strecher, V. J. (1991). Planned health education and the role of self-efficacy: Dutch research. *Health Education Research, 6,* 231–238.

Kristiansen, C. M. (1986). A two-value model of preventive health behaviour. *Basic and Applied Social Psychology, 7,* 173–184.

Kuhl, J. (1985). Volitional mediators of cognition-behavior consistency: Self-regulatory processes and action versus state orientation. In J. Kuhl & J. Beckman (Eds.), *Action control: From cognition to behavior* (pp. 101–128). New York: Springer.

Kurland, N. B. (1995). Ethical intentions and the theories of reasoned action and planned behavior. *Journal of Applied Social Psychology, 25,* 297–313.

Lam, C. S., McMahon, B. T., Priddy, D. A., & Gehred-Schutlz, A. (1988). Deficit awareness and treatment performance among traumatic head injury adults. *Brain Injury, 2,* 235–242.

Langlie, J. K. (1977). Social networks, health beliefs and preventive health behaviours. *Journal of Health and Social Behaviour, 18,* 244–260.

Lau, R. R. (1982). Origins of health locus of control beliefs. *Journal of Personality and Social Psychology, 42,* 322–334.

Lau, R. R., Hartman, K. A., & Ware, J. E. (1986). Health as value: methodological and theoretical considerations. *Health Psychology, 5,* 25–43.

Lazarus, R. S. & Launier, R. (1978). Stress related transactions between person and environment. In L. A. Pervin & M. Lewis (Eds.), *Perspectives in interactional psychology* (pp. 287–327). New York: Plenum.

Lefcourt, H. M. (1991). Locus of control. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 661–753). New York: Academic Press.

Leitner, M., Shapland, J., & Wiles, P. (1993). *Drug usage and drugs prevention: The views and habits of the general public.* London: HMSO.

Lennon, D. L. F., Stratman, F. W., Shrago, E., Nagle, F. J., Hanson, P. G., Maddon, M., & Spennetta, T. (1983). Total cholesterol and DHL cholesterol changes during acute, moderate intensity exercise in men and women. *Metabolism, 32,* 244–249.

Levenson, H. (1974). Multidimensional locus of control in psychiatric patients. *Journal of Consulting and Clinical Psychology, 41,* 397–404.

Leventhal, H. (1970). Findings and theory in the study of fear communications. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 5, pp. 119–186). New York: Academic Press.

Leventhal, H., Nerenz, D. R., & Steele, D. F. (1984). Illness representations and coping with health threats. In A. Baum & J. Singer (Eds.), *A handbook of psychology and health* (pp. 219–252). Hillsdale, NJ: Erlbaum.

Lewin, K. (1951). *Field theory in social science.* New York: Harper.

Liao, K. L., Hunter, M., & Weinman, J. (1995). Health-related behaviours and their correlates in a general population sample of 45-year old women. *Psychology and Health, 10,* 171–184.

Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes, 65,* 272–292.

Madden, T. J., Ellen, P. S., & Ajzen, I. (1992). A comparison of the theory of planned behavior and the theory of reasoned action. *Personality and Social Psychology Bulletin, 18,* 3–9.

Maddux, J. E. (1993). (Ed.) *Self-efficacy, adaptation and adjustment: theory, research and application.* New York: Plenum.

Maddux, J. E., Norton, L. W., & Stoltenberg, C. D. (1986). Self-efficacy expectancy, outcome expectancy, and outcome value: relative effects on behavioral intentions. *Journal of Personality and Social Psychology, 51,* 783–789.

Maddux, J. E., & Rogers, R. W. (1983). Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change. *Journal of Experimental Social Psychology, 19,* 469–479.

Maibach, E. W., Flora, J. A., & Nass, C. (1991). Changes in self-efficacy and health behavior in response to a minimal community health campaign. *Health Communication, 3,* 1–15.

Manning, M. M. & Wright, T. L. (1983). Self-efficacy expectancies, outcome expectancies, and the persistence of pain control in childbirth. *Journal of Personality and Social Psychology, 54,* 421–431.

Marcus, B. H., Eaton, C. A., Rossi, J. S., & Harlow, L. L. (1994). Self-efficacy, decision-making, and stages of change—An integrative model of physical exercise. *Journal of Applied Social Psychology, 24,* 489–508.

Marteau, T. M. (1989). Health beliefs and attributions. In A. K. Broome (Ed.), *Health psychology: processes and applications* (pp. 1–23). London: Chapman Hall.

McAuley, E. (1993). Self-efficacy and the maintenance of exercise participation in older adults. *Journal of Behavioral Medicine, 16,* 103–113.

McCaul, K. D., Glasgow, R. E., & O'Neill, H. K. (1992). The problem of creating habits: establishing health-protective dental behaviors. *Health Psychology, 11,* 101–110.

McCaul, K. D., Sandgren, A. K., O'Neill, H. K., & Hinsz, V. B. (1993). The value of the theory of planned behavior, perceived control, and self-efficacy expectations for predicting health-protective behaviors. *Basic and Applied Social Psychology, 14,* 231–252.

McCelland, D. C. (1985). How motives, skills, and values determine what people do. *American Psychologist, 40,* 812–825.

McConnaughly, E. A., DiClemente, C. C., Prochaska, J. O., & Velicer, W. F. (1989). Stages of change in psychotherapy: A follow-up report. *Psychotherapy, 26,* 494–503.

McCusker, J., & Morrow, G. (1979). The relationship of health locus of control to preventive health behaviors and health beliefs. *Patient Counselling and Health Education, 1,* 146–150.

McNeil, A. D., Jarvis, M. J., Stapleton, J. A., Russell, M. A. H., Eiser, J. R., Gammage, P., & Gray, E. M. (1988). Prospective study of factors predicting uptake of smoking in adolescents. *Journal of Epidemiology and Community Health, 43,* 72–78.

Mechanic, D., & Cleary, P. D. (1980). Factors associated with the maintenance of positive health behavior. *Preventive Medicine, 9,* 805–814.

Metzner, H. L., Carman W. J., & House, J. (1983). Health practices, risk factors and chronic disease in Tecumeseh. *Preventive Medicine, 12,* 491–507.

Mittag, W. & Schwarzer, R. (1993). Interaction of employment status and self-efficacy on alcohol consumption: A two-wave study on stressful life transitions. *Psychology and Health, 8,* 77–87.

Moore, L., Smith, C., & Catford, J. (1994). Binge drinking: prevalence, patterns and policy. *Health Education Research, 9,* 497–505.

Morrison, D. M., Gillmore, M. R., & Baker, S. A. (1995). Determinants of condom use among high-risk heterosexual adults: A test of the theory of reasoned action. *Journal of Applied Social Psychology, 25,* 651–676.

Muhlenkamp, A. F., Brown, N. J., & Sands, D. (1985). Determinants of health promotion activities in nursing clinic clients. *Nursing Research, 34,* 327–333.

Mullen, P. D., Hersey, J. C., & Iverson, D. C. (1987).

Health behavior models compared. *Social Science and Medicine, 24,* 973–981.

Nemeck, M. A. (1990). Health beliefs and breast self-examination among black women. *Health Values, 14,* 41–52.

Netemeyer, R. G., Burton, S., & Johnston, M. (1991). A comparison of two models for the prediction of volitional and goal-directed behaviors: A confirmatory analysis approach. *Social Psychology Quarterly, 54,* 87–100.

Nicassio, P., Wallston, K., Callahan, L., Herbert, M., & Pincus, T. (1985). The measurement of helplessness in rheumatoid arthritis: The development of the arthritis helplessness scale. *Journal of Rheumatology, 12,* 462–467.

Norman, P. (1990). Health locus of control beliefs and preventive health behaviours. *Journal of the Institute of Health Education, 28,* 113–117.

Norman, P. (1991). Social learning theory and the prediction of attendance at screening. *Psychology and Health, 5,* 231–239.

Norman, P. (1995). Heath locus of control and health behaviour: an investigation into the role of health value and behaviour-specific efficacy beliefs. *Personality and Individual Differences, 18,* 213–218.

Norman, P., & Bennett, P. (1996). Health locus of control. In M. Conner & P. Norman (Eds.), *Predicting health behaviour* (pp. 62–94). Buckingham, UK: Open University Press.

Norman, P., & Conner, M. (1993). The role of social cognition models in predicting attendance at health checks. *Psychology and Health, 8,* 447–462.

Norman, P., & Conner, M. (1996a). The role of social cognition models in predicting health behaviours: future directions. In M. Conner & P. Norman (Eds.), *Predicting health behaviour* (pp. 197–225). Buckingham, UK: Open University Press.

Norman, P., & Conner, M. (1996b). Predicting health-check attendance among prior attenders and nonattenders: The role of prior behavior in the theory of planned behavior. *Journal of Applied Social Psychology, 26,* 1010–1026.

Norman, P., & Smith, L. (1995). The theory of planned behaviour and exercise: An investigation into the role of prior behaviour, behavioural intentions and attitude variability. *European Journal of Social Psychology, 25,* 403–415.

Nucifora, J., Gallois, C., & Kashima, Y. (1993). Influences on condom use among undergraduates: Testing the theories of reasoned action and planned behaviour. In D. J. Terry, C. Gallois & M. McCamish (Eds.), *The theory of reasoned action: Its application to AIDS-preventive behaviour* (pp. 47–64). Oxford, UK: Pergamon.

Nutbeam, D., & Catford, J. (1990). Modifiable risks for cardiovascular disease among general practitioners in Wales. *Public Health, 104,* 353–361.

Oberman, A. (1985). Exercise and the primary prevention of cardiovascular disease. *American Journal of Cradiology, 55,* 10D–20D.

O'Connell, J. K., & Price, J. H. (1982). Health locus of control of physical fitness program participants. *Perceptual and Motor Skills, 55,* 925–926.

O'Leary, A. (1985). Self-efficacy and health. *Behavioral Research and Therapy, 23,* 437–451.

O'Leary, A., & Raffaelli, M. (1996). Preventing the sexual transmission of HIV: Current status and future directions. *Psychology and Health, 11,* 75–94.

Office of Population Census and Surveys (1968). *Drug use in the UK.* London: Government Statistical Services.

Oliver, R. L. & Berger, P. K. (1979). A path analysis of preventive health care decision models. *Journal of Consumer Research, 6,* 113–122.

Orbell, S., Crombie, I., & Johnson, G. (1996). Social cognition and social structure in the prediction of cervical screening uptake. *British Journal of Health Psychology, 1,* 35–50.

Orbell, S., Hodgkins, S., & Sheeran, P. (1997). Translating intentions into action. *Personality and Social Psychology Bulletin, 23,* 953–962.

Paisley, C. M., & Sparks, P. (1998). Expectations of reducing fat intake: The role of perceived need with the theory of planned behaviour. *Psychology and Health, 13.*

Parker, D., Manstead, A. S. R., & Stradling, S. G. (1995). Extending the theory of planned behaviour: The role of personal norm. *British Journal of Social Psychology, 34,* 127–137.

Parker, D., Manstead, A. S. R., & Stradling, S. G. (1996). Modifying beliefs and attitudes to exceed the speed limit: An intervention study based on the theory of planned behavior. *Journal of Applied Social Psychology, 26,* 1–19.

Parker, D., Manstead, A. S. R., Stradling, S. G., Reason, J. T., & Baxter, J. S. (1992). Intention to commit driving violations—An application of the theory of planned behavior. *Journal of Applied Psychology, 77,* 94–101.

Peak, H. (1955). Attitude and motivation. In M. R. Jones (Ed.), *Nebraska symposium on motivation* (Vol. 3, pp. 149–188). Lincoln, NE: University of Nebraska Press.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes of attitude change.* New York: Springer-Verlag.

Pomazal, R. P., & Jaccard, J. J. (1976). An informational approach to altruistic behavior. *Journal of Personality and Social Psychology, 33,* 317–326.

Prochaska, J. O., & DiClemente, C. C. (1984). *The transtheoretical approach: Crossing traditional boundaries of therapy.* Homewood, IL: Dow Jones Irwin.

Prwun, J., van der Borne, H., de Reuver, R., de Boer, M., Ter Pelkwijk, M., & de Jong, P. (1988). The locus of control scale for cancer patients. *Tijdscrift vour Sociale Gezondherdszong, 66,* 404–408.

Quadrel, M. J. & Lau, R. R. (1989). Health promotion, health locus of control, and health behavior: Two field experiments. *Journal of Applied Social Psychology, 18,* 1497–1521.

Quine, L., Rutter, D. R., & Arnold, L. (1998). Predicting safety helmet use among schoolboy cyclists: A comparison of the theory of planned behavior and the health belief model. *Psychology and Health, 13.*

Raats, M. M. (1992). *The role of beliefs and sensory to milk in determining the selection of milks of different fat content.* Unpublished doctoral thesis, University of Reading.

Redeker, N. S. (1989). Health beliefs, health locus of control, and the frequency of practice of breast self-examination in women. *Journal of Obstetrics, Gynaecology and Neonatal Nursing, 18,* 45–51.

Redhead, I. H. (1960). The incidence of glycosuria and diabetes mellitus in general practice. *British Medical Journal, 1,* 695.

Reiss, I. L., & Leik, R. K. (1989). Evaluating strategies to avoid AIDS: Numbers of partners versus use of condoms. *Journal of Sex Research, 26,* 411–433.

Richard, R., & van der Pligt, J. (1991). Factors affecting condom use among adolescents. *Journal of Community and Applied Social Psychology, 1,* 105–116.

Richard, R., van der Pligt, J., & de Vries, N. (1995). Anticipated affective reactions and prevention of AIDS. *British Journal of Social Psychology, 34,* 9–21.

Richard, R., van der Pligt, J., & de Vries, N. (1996a). Anticipated affect and behavioral choice. *Basic and Applied Social Psychology, 18,* 111–129.

Richard, R., Van der Pligt, J., & de Vries, N. (1996b). Anticipated regret and time perspective: Changing sexual risk-taking behavior. *Journal of Behavioral Decision Making, 9,* 185–199.

Rigotti, N. A. (1989). Cigarette smoking and body weight.

New England Journal of Medicine, 320, 931–933.

Rippetoe, P. A., & Rogers, R. W. (1987). Effects of components of protection motivation theory on adaptive and maladpative coping with a health threat. *Journal of Personality and Social Psychology, 52,* 596–604.

Rodin, J., & Salovey, P. (1989). Health psychology. *Annual Review of Psychology, 40,* 533–579.

Rogers, R. W. (1983). Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation. In J. T. Cacioppo & R. E. Petty (Eds.), *Social psychophysiology: a source book* (pp. 153–176). New York: Guilford Press.

Ronis, D. L., & Harel, Y. (1989). Health beliefs and breast self-examination behaviours: Analysis of linear structural relations. *Psychology and Health, 3,* 259–285.

Ronis, D. L., Yates, J. F., & Kirscht, J. P. (1989). Attitudes, decisions and habits as determinants of repeated behavior. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 213–239). Hillsdale, NJ: Erlbaum.

Rosen, T. J., & Shipley, R. H. (1983). A stage analysis of self-initiated smoking reductions. *Addictive Behaviors, 8,* 263–272.

Rosenstock, I. M. (1966). Why people use health services. *Millbank Memorial Fund Quarterly, 44,* 94–124.

Rosenstock, I. M. (1974). Historical origins of the health belief model. *Health Education Monographs, 2,* 1–8.

Rosenstock, I. M., Strecher, V. J., & Becker, H. M. (1988). Social learning theory and the health belief model. *Health Education Quarterly, 15,* 175–183.

Rosenthal, M., Haskell, W. L., Solomon, R., Widstrom, A., & Reavan, G. M. (1983). Demonstration of a relationship between level of physical training and insulin stimulated glucose utilization in normal humans. *Diabetes, 32,* 408–411.

Roskam, S. (1986). *Application of health locus of control typology toward predicting depression and medical adherence in rheumatoid arthritis.* Unpublished doctoral dissertation, Vanderbilt University, Nashville, TN.

Rotter, J. B. (1954). *Social learning and clinical psychology.* Englewood Cliffs, NJ: Prentice-Hall.

Rotter, J. B. (1966). Generalized expectancies for internal and external control of reinforcement. *Psychological Monographs: General and Applied, 80 (whole no. 609),* 1–28.

Royal College of Physicians (1983). *Health or smoking: Follow-up report.* London: Pitman.

Rutter, D. R., & Quine, L. (1996). Social psychological mediators of the relationship between demographic factors and health outcomes: A theoretical model and some preliminary data. *Psychology and Health, 11,* 5–22.

Sachs, M. L. (1984). Psychological well-being and vigorous activity. In J. M. Silva & R. S. Weinberg (Eds.), *Psychological foundations of sport* (pp. 435–444). Champaign, IL: Human Kinetics.

Sallis, J. F., Howell, M. F., Hofstetter, C. R., & Barrington, E. (1992). Explanation of vigorous physical activity during two years using social learning variables. *Social Science and Medicine, 34,* 25–32.

Saltzer, E. B. (1982). The weight locus of control (WLOC) scale: A specific measure for obesity research. *Journal of Personality Assessment, 46,* 620–628.

Sampos, C., Fulwood, R., Haines, C., Carroll, M., Anda, R., Williamson, D. F., Remington, P., & Cleeman, J. (1989). The prevalence of high blood cholesterol levels among adults in the United States. *Journal of the American Medical Association, 262,* 45–52.

Sandman, P. M., & Weinstein, N. D. (1993). Predictors of home radon testing and implications for testing promotion programs. *Health Education Quarterly, 20,* 471–487.

Scheier, M. F. & Carver, C. S. (1992). Effects of optimism on psychological and physical well-being: Theoretical overview and empirical update. *Cognitive Therapy and Research, 16,* 201–228.

Schifter, D. E., & Ajzen, I. (1985). Intention, perceived control and weight loss: An application of the theory of planned behaviour. *Journal of Personality and Social Psychology, 49,* 843–851.

Schlegel, R. P., D'Avernas, J. R., Zanna, M. P., DeCourville, N. H., & Mansue, S. R. (1992). Problem drinking: A problem for the theory of reasoned action? *Journal of Applied Social Psychology, 22,* 358–385.

Schwartz, S. H., & Tessler, R. C. (1972). A test of a model for reducing measured attitude-behavior discrepancies. *Journal of Personality and Social Psychology, 24,* 225–236.

Schwarzer, R. (1992). Self-efficacy in the adoption and maintenance of health behaviors: Theoretical approaches and a new model. In R. Schwarzer (Ed.), *Self-efficacy: thought control of action* (pp. 217–243). London: Hemisphere.

Schwarzer, R. (1994). Optimism, vulnerability, and self-beliefs as health-related cognitions: a systematic overview. *Psychology and Health, 9,* 161–180.

Schwarzer, R., & Fuchs, R. (1996). Self-efficacy and health behaviour. In M. Conner & P. Norman (Eds.), *Predicting health behaviour* (pp. 163–196). Buckingham, UK: Open University Press.

Seeman, M., & Seeman, T. E. (1983). Health behaviour and personal autonomy: A longitudinal study of the sense of control in illness. *Journal of Health and Social Behaviour, 24,* 144–160.

Segall, M. E., & Wynd, C. A. (1990). Health conception, health locus of control, and power as predictors of smoking behavior change. *American Journal of Health Promotion, 4,* 338–344.

Seydel, E., Taal, E., & Wiegman, O. (1990). Risk-appraisal, outcome and self-efficacy expectancies: Cognitive factors in preventive behaviour related to cancer. *Psychology and Health, 4,* 99–109.

Shaper, A. G., Pocock, S. J., Walker, M., Cohen, N. M., Wale, C. J., & Thomson, A. G. (1981). British Regional Heart Survey: Cardiovascular risk factors in middle-aged men in 24 towns. *British Medical Journal, 283,* 179–186.

Sheeran, P., & Abraham, C. (1996). The health belief model. In M. Conner & P. Norman (Eds.), *Predicting health behaviour* (pp. 23–61). Buckingham, UK: Open University Press.

Sheppard, B. H., Hartwick, J., & Warshaw, P. R. (1988). The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research, 15,* 325–343.

Sherlock, S. (1982). (Ed.) Alcohol and Disease. *British Medical Bulletin, 38* (whole issue).

Simonson, I. (1992). The influence of anticipating regret and responsibility on purchase decisions. *Journal of Consumer Research, 19,* 105–118.

Singer, R. N. (1992). Physical activity and psychological benefits: A position statement. *The Sport Psychologist, 6,* 199–203.

Slenker, S. E., Price, J. H., & O'Connell, J. K. (1985). Health locus of control of joggers and nonexercisers. *Perceptual and Motor Skills, 61,* 323–328.

Snyder, C. R., Harris, C., Anderson, J. R., Holleran, S. A., Irving, L. M., Sigmon, S. T., Yoshinobu, L., Gibb., J., Langelle, C., & Harney, P. (1991). The will and the ways: Development and validation of an individual-differences measure of hope. *Journal of Personality and Social Psychology, 60,* 570–585.

Sparks, P. (1994). Food choice and health: Applying, assessing, and extending the theory of planned behaviour. In D. R. Rutter & L. Quine (Eds.), *Social psychology and health: European perspectives* (pp. 25–46). Aldershot, UK: Avebury.

Sparks, P., & Shepherd, R. (1992). Self-identity and the theory of planned behaviour: Assessing the role of identification with "green consumerism". *Social Psychology Quarterly, 55,* 388–399.

Sparks, P., Shepherd, R., & Frewer, L. J. (1995). Assessing and structuring attitudes toward the use of gene technology in food production—The role of perceived ethical obligation. *Basic and Applied Social Psychology, 16,* 267–285.

Sparks, P., Shepherd, R., Wieringa, N., & Zimmermanns, N. (1994). *Barriers to healthy eating: An examination of perceived behavioural control and unrealistic optimism.* Paper presented at the BPS Social Psychology Section Annual Conference, University of Hertfordshire, UK, September 1994.

Sparks, P., Shepherd, R., Wieringa, N., & Zimmermanns, N. (1995). Perceived behavioural control, unrealistic optimism and dietary change: An exploratory study. *Appetite, 24,* 243–355.

Speake, D. L., Cowart, M. E., & Stephens, R. (1991). Healthy lifestyle practices of rural and urban elderly. *Health Values, 15,* 45–51.

Stacy, R. D., & Lloyd, B. H. (1990). An investigation of beliefs about smoking among diabetes patients: Information for improving cessation efforts. *Journal of Patient Education and Counselling, 15,* 181–189.

Stainback, R. D., & Rogers, R. W. (1983). Identifying effective components of alcohol abuse prevention programs: effects of fear appeals, message style and source expertise. *International Journal of Addictions, 18,* 393–405.

Stanley, M. A. & Maddux, J. E. (1986). Cognitive processes in health enhancement: Investigation of a combined protection motivation and self-efficacy model. *Basic and Applied Social Psychology, 7,* 101–113.

Stanton, A. (1987). Determinance of adherence to medical regimens by hypertensive patients. *Journal of Behavioural Medicine, 10,* 377–394.

Steptoe, A., & Wardle, J. (1992). Cognitive predictors of health behaviour in contrasting regions of Europe. *British Journal of Clinical Psychology, 31,* 485–502.

Steptoe, A., & Wardle, J. (1996). The European health and behaviour survey: The development of an international study in health psychology. *Psychology and Health, 11,* 49–73.

Steptoe, A., Wardle, J., Vinck, J., Tuomisto, M., Holte, A., & Wickstrom, L. (1994). Personality and attitudinal correlates of healthy and unhealthy lifestyles in young adults. *Psychology and Health, 9,* 331–343.

Stevens, V. J. & Hollis, J. F. (1989). Preventing smoking relapse, using an individually tailored skills-training technique. *Journal of Consulting and Clinical Psychology, 57,* 420–424.

Strax, P. (1984). Mass screening for control of breast cancer. *Cancer, 53,* 665–670.

Stroebe, W., & Stroebe, M. S. (1995). *Social psychology and health.* Buckingham, UK: Open University Press.

Sutton, S. (1982). Fear-arousing communications: A critical examination of theory and research. In J. R. Eiser (Ed.), *Social psychology and behavioural medicine* (pp. 303–337). London: Wiley.

Sutton, S. (1994). The past predicts the future: Interpreting behaviour–behaviour relationships in social psychological models of health behaviours. In D. R. Rutter & L. Quine (Eds.), *Social psychology and health: European perspectives* (pp. 71–88). Aldershot, UK: Avebury.

Sutton, S., & Hallett, R. (1989). Understanding seat-belt intentions and behavior: A decision-making approach. *Journal of Applied Social Psychology, 19,* 1310–1325.

Taylor, D. W. (1979). A test of the health belief model in hypertension. In R. B. Haynes, D. W. Taylor, & K. L. Sackett (Eds.), *Compliance in health care* (pp. 103–109). Baltimore: Johns Hopkins University Press.

Taylor, S. E. (1991). *Health psychology.* New York: McGraw-Hill.

Terry, D. J., & O'Leary, J. E. (1995). The theory of planned behaviour: The effects of perceived behavioural control and self-efficacy. *British Journal of Social Psychology, 34,* 199–220.

Terry, D. J., & Hogg, M. A. (1996). Group norms and the attitude-behavior relationship: A role for group identification. *Personality and Social Psychology Bulletin, 22,* 776–793.

Theodorakis, Y. (1994). Planned behavior, attitude strength, role identity, and the prediction of exercise behavior. *The Sport Psychologist, 8,* 149–165.

Triandis, H. C. (1977). *Interpersonal behavior.* Monterey, CA: Brooks-Cole.

US Department of Health and Human Services (1988). *The Surgeon General's report on nutrition and health.* Washington, DC: US Government Printing Office.

US Department of Health and Human Services (1990). *Health people 2000: National health promotion and disease prevention objectives.* Washington, DC: US Government Printing Office.

Valois, P., Desharnais, R., & Godin, G. (1988). A comparison of the Fishbein and the Triandis attitudinal models for the prediction of exercise intention and behavior. *Journal of Behavioral Medicine, 11,* 459–472.

Van den Putte, H. (1993). *On the theory of reasoned action.* Unpublished doctoral dissertation, University of Amsterdam.

Van der Pligt, J., & de Vries, N. (1998). Expectancy-value models of health behaviours: The role of salience and anticipated affect. *Psychology and Health, 13.*

Van der Pligt, J., & Eiser, J. R. (1984) Dimensional salience, judgement and attitudes. In J. R. Eiser (Ed.), *Attitudinal judgement* (pp. 43–63). New York: Springer-Verlag.

Van der Velde, W., & van der Pligt, J. (1991). AIDS-related behavior: Coping, protection motivation, and previous behavior. *Journal of Behavioral Medicine, 14,* 429–451.

Vernon, S. W., Laville, E. A., & Jackson, G. L. (1990). Participation in breast screening programs: a review. *Social Science and Medicine, 30,* 1107–1118.

Waldron, I. (1988). Why do women live longer than men? *Journal of Human Stress, 2,* 2–13.

Waller, K. V., & Bates, R. C. (1992). Health locus of control and self-efficacy beliefs in a healthy elderly sample. *American Journal of Health Promotion, 6,* 302–309.

Wallston, B. S., Wallston, K. A., Kaplan, G. D., & Maides, S. A. (1976). Development and validation of the health locus of control (HLC) scale. *Journal of Consulting and Clinical Psychology, 44,* 580–585.

Wallston, K. A. (1991). The importance of placing measures of health locus of control beliefs in a theoretical context. *Health Education Research: Theory and Practice, 6,* 251–252.

Wallston, K. A. (1992). Hocus-pocus, the focus isn't strictly on locus: Rotter's social learning theory modified for health. *Cognitive Therapy and Research, 16,* 183–199.

Wallston, K. A., Maiders, S., & Wallston, B. S. (1976). Health-related information seeking as a function of health-related locus of control and health value. *Journal of Research in Personality, 10,* 215–222.

Wallston, K. A., & Wallston, B. S. (1980). Health locus of control scales. In H. Lefcourt (Ed.), *Advances and innovations in locus of control research* (pp. 198–234). New York: Academic Press.

Wallston, K. A., & Wallston, B. S. (1981). Health locus of control scales. In H. Lefcourt (Ed.), *Research with the locus of control construct* (Vol. 1, pp. 189–243). New York: Academic Press.

Wallston, K. A., & Wallston, B. S. (1982). Who is responsible for your health? The construct of health locus of control. In G. Sanders & J. Suls (Eds.), *Social psychology of health and illness* (pp. 65–95). Hillsdale, NJ: Erlbaum.

Wallston, K. A., Wallston, B. S., & DeVellis, R. (1978). Development of multidimensional health locus of control (MHLC) scales. *Health Education Monographs, 6,* 160–170.

Wardle, J. (1995). Cholesterol and psychological well-being. *Journal of Psychosomatic Research, 39,* 549–562.

Warshaw, P. R., & Davis, F. D. (1985). Disentangling behavioral intention and behavioral expectation. *Journal of Experimental Social Psychology, 21,* 213–228.

Weatherburn, P., Hunt, A. J., Davies, P. M., Coxon, A. P. M., & McManus, T. J. (1991). Condom use in a large cohort of homosexually active men in England and Wales. *AIDS Care, 3,* 31–41.

Weinstein, N. D. (1988). The precaution adoption process. *Health Psychology, 7,* 355–386.

Weinstein, N. D. & Nicolich, M. (1983). Correct and incorrect interpretations of correlations between risk perceptions and risk behaviors. *Health Psychology, 12,* 235–245.

Weinstein, N. D., Rothman, A. J., & Sutton, S. R. (in press). Stage theories of health behavior. *Health Psychology.*

Weinstein, N. D., & Sandman, P. M. (1992). A model of the precaution adoption process: evidence from home radon testing. *Health Psychology, 11,* 170–180.

Weiss, G. L., & Larsen, D. L. (1990). Health value, health locus of control, and the prediction of health protective behaviours. *Social Behavior and Personality, 18,* 121–135.

Wilson, D., Zenda, A., McMaster, J., & Lavelle, S. (1992). Factors predicting Zimbabwean students' intentions to use condoms. *Psychology and Health, 7,* 99–114.

Winett, R. A. (1985). Ecobehavioral assessment in health life-styles: Concepts and methods. In P. Karoly (Ed.), *Measurement strategies in health psychology* (pp. 147–181). Chichester, UK: Wiley.

Wojcik, J. V. (1988). Social learning predictors of the avoidance of smoking relapse. *Addictive Behaviors, 13,* 177–180.

Wurtele, S. K. (1988). Increasing women's calcium intake: the role of health beliefs, intentions and health value. *Journal of Applied Social Psychology, 18,* 627–639.

Wurtele, S. K., Britcher, J. C., & Saslawsky, D. A. (1985). Relationships between locus of control, health value and preventive health behaviors among women. *Journal of Research in Personality, 19,* 271–278.

Wurtele, S. K., & Maddux, J. E. (1987). Relative contributions of protection motivation components in predicting exercise intentions and behavior. *Health Psychology, 6,* 453–466.

Young, H. M., Lierman, L., Powell-Cope, G., Kasprzyk, D., & Bonoliel, J. Q. (1991). Operationalizing the theory of planned behaviour. *Research in Nursing and Health, 14,* 137–144.

Zuckerman, M., & Reiss, H. T. (1978). Comparison of three models for predicting altruistic behavior. *Journal of Personality and Social Psychology, 36,* 498–510.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.02
Psychophysiological Bases of Disease

ANDREW STEPTOE
University of London, UK

8.02.1 INTRODUCTION 40

8.02.2 THE PSYCHOPHYSIOLOGICAL RESEARCH FRAMEWORK 41
- *8.02.2.1 Stress and Coping* 41
- *8.02.2.2 Research Strategies* 42
  - *8.02.2.2.1 Animal studies* 42
  - *8.02.2.2.2 Mental stress testing* 43
  - *8.02.2.2.3 Field studies* 43
  - *8.02.2.2.4 Survey studies* 44
  - *8.02.2.2.5 Convergence of methods* 44

8.02.3 PSYCHOPHYSIOLOGICAL RESPONSES 45
- *8.02.3.1 Neuroendocrine Parameters* 45
  - *8.02.3.1.1 Adrenal steroid stress response* 45
  - *8.02.3.1.2 Catecholamines and the sympathetic nervous system* 46
  - *8.02.3.1.3 Endogenous opioids* 47
  - *8.02.3.1.4 Other neuroendocrine and hormonal responses* 47
- *8.02.3.2 Cardiovascular and Autonomic Parameters* 48
  - *8.02.3.2.1 Blood pressure, heart rate, and regional blood flow* 48
  - *8.02.3.2.2 Cardiac function* 49
  - *8.02.3.2.3 Platelets and hemostasis* 50
  - *8.02.3.2.4 Other physiological systems* 50
- *8.02.3.3 Lipids and Glucose Metabolism* 51
  - *8.02.3.3.1 Insulin metabolism and abdominal obesity* 51
- *8.02.3.4 Immune Parameters* 52
  - *8.02.3.4.1 Regulation of immune function* 53
  - *8.02.3.4.2 Acute stress-induced immune responses* 53
  - *8.02.3.4.3 Chronic changes in immune function* 53
  - *8.02.3.4.4 Functional significance of stress-induced immune modulation* 54
- *8.02.3.5 Adaptive or Destructive?* 55

8.02.4 THE NATURE OF PSYCHOSOCIAL DEMANDS 55
- *8.02.4.1 Life Events, Chronic Stressors, and Daily Hassles* 55
- *8.02.4.2 Characteristics of Psychosocial Demands* 56
  - *8.02.4.2.1 Novelty, familiarity, and predictability* 56
  - *8.02.4.2.2 Control and lack of control* 57
- *8.02.4.3 Work Characteristics, Control, and Psychophysiological Responses* 57
- *8.02.4.4 Concurrent Stressors* 58

8.02.5 RESISTANCE AND VULNERABILITY FACTORS 58
- *8.02.5.1 Personality and Behavior Patterns* 58
  - *8.02.5.1.1 Disease-prone personality and negative affectivity* 59
  - *8.02.5.1.2 Anger and hostility* 59
  - *8.02.5.1.3 Depression, pessimism, and optimism* 60
  - *8.02.5.1.4 Emotional inhibition* 60

8.02.5.2 Social Isolation and Social Support 61
8.02.5.3 Prior Experience of Stressful Events 62

8.02.6 PATHWAYS TO DISEASE—PREDISPOSING FACTORS 63

8.02.6.1 Nutrition and Exercise 63
8.02.6.2 Advanced Age 64
8.02.6.3 Gender 64
8.02.6.4 Ethnicity 65
8.02.6.5 Socioeconomic Status and Education 65
8.02.6.6 Genetic Factors 65
8.02.6.7 Individual Patterns of Psychophysiological Responses 66

8.02.7 PSYCHOPHYSIOLOGICAL MODELS OF DISEASE 66

8.02.7.1 Physiological Reactivity as a Causal Factor 67
8.02.7.2 Physiological Reactivity as an Inhibitor of Host Resistance and Defense 68
8.02.7.3 Physiological Reactivity as a Disruptive Factor 69
8.02.7.4 Physiological Reactivity as a Trigger of Acute Clinical Events 70

8.02.8 CONCLUSION 70

8.02.9 REFERENCES 70

8.02.1 INTRODUCTION

All behavioral states and activities are underpinned by variations in physiological activity, and are associated with specific patterns of autonomic nervous system, neuroendocrine, and metabolic response. The relationship between emotional disturbances and physiological function has been recognized since the work of Sherrington, Cannon and their contemporaries was carried out nearly 100 years ago. One of the fundamental concepts underlying clinical health psychology is that psychosocial factors influence the etiology and maintenance of disease through perturbation of physiological functions. This chapter will evaluate the empirical evidence supporting this concept, and will outline the theoretical framework within which psychophysiological or psychobiological processes can best be understood.

The range of experimental and clinical research relevant to this topic is large, since it not only includes work on the regulation of peripheral bodily functions by the brain, but also the distinctive pathological processes implicated in individual medical disorders. The chapter begins with a discussion of the psychophysiological research framework, basic models of stress and coping, and methods of investigation and quantification. The nature and pattern of psychophysiological responses is then described, together with an outline of the psychosocial factors that influence the intensity and duration of these responses. Later sections of the chapter describe how the physiological responses elicited by psychosocial factors actually affect disease mechanisms and increase health risk.

In keeping with the remit of health psychology, the focus is on nonpsychiatric "medical" conditions and on peripheral rather than central nervous system psychophysiology. Thus the role of physiological factors in problems such as depression, panic, or hypochondriasis will not be discussed. This does not imply an endorsement of Cartesian dualism, or a belief that psychiatric and nonpsychiatric disorders are regulated by different processes. Rather it reflects the specific expertise that has developed within health psychology and behavioral medicine over recent years. Many of the important medical conditions that concern health psychologists are detailed in later chapters of this volume. The present chapter therefore concentrates on general models rather than specific conditions, and problems such as high blood pressure (hypertension), diabetes mellitus, and bronchial asthma are introduced only for illustrative purposes. The arguments here build on the models of psychophysiological processes in disease that have been presented in earlier publications (Steptoe, 1980, 1984a, 1991).

The major theme of this chapter is that there is now abundant evidence that psychophysiological processes have a role in a range of medical conditions. However, the nature of this role is varied, and depends on the stage of the development of the disorder, and in the ways psychophysiological processes interact with other factors and pathological mechanisms. Psychophysiological processes are implicated as cofactors in medical conditions, rather than exclusive etiological agents. As such, their importance varies across medical problems and between individuals. The era has passed in which certain diseases might be regarded as "psychosomatic," while others were purely biological. Instead, a more fluid concept needs to be understood, in which emotion, behavior, and other psychosocial factors play their parts to a varying extent right across the disease spectrum. Exciting possibilities for management and prevention in the new models of psychophysiological processes are emerging from the literature.

8.02.2 THE PSYCHOPHYSIOLOGICAL RESEARCH FRAMEWORK

Physiological adjustments are associated with the entire range of behavioral states from sleep to vigorous physical exercise. Patterns of blood flow, pulmonary function, gastrointestinal activity, and so on are adapted to sustain the motoric, energetic, and metabolic requirements of the organism. For example, acute exercise is accompanied by an increase in pulmonary ventilation, heart rate and blood pressure, preferential blood flow to skeletal muscle, and mobilization of energy supplies stored in the liver and adipose tissue. These responses maintain adequate glucose supplies to working muscle, and ensure prompt removal of waste products. In the context of disease, the physiological responses that are particularly relevant are those elicited by potential harm. Responses of this kind are typically described as psychophysiological stress responses.

8.02.2.1 Stress and Coping

The definition of psychological stress is an issue that has generated a great deal of discussion and dispute, reflected in the colorful title of the chapter by Engel (1985) *Stress is a noun! No, a verb!, No, an adjective!* It has been argued that stress is a term that is so misused and abused that it be jettisoned altogether as a construct of serious scientific value. Nevertheless, the term has proved difficult to discard, if only as a shorthand to help define a valid area of human experience and clinical importance.

Many of the problems surrounding the concept of stress have stemmed from the definition put forward by an early pioneer, Hans Selye. His definition of stress as the "nonspecific response of the body to any demand" gave rise to two major difficulties. First, the conditions eliciting stress responses were defined so widely (as "any demand") as to be virtually meaningless. The second difficulty is that he introduced the idea of an undifferentiated biological stress response, characterized principally by release of the group of steroid hormones known as glucocorticoids from the adrenal glands. He failed to appreciate that the activation of neuroendocrine, autonomic, and immune pathways is delicately patterned, depending on the precise demands on the organism and the behavioral and cognitive coping responses that are mobilized.

Weiner (1992) has eloquently deconstructed the background to Selye's model, and argues that it arose from the type of animal experiment that was carried out in the middle decades of the century. These studies utilized a variety of intense aversive experiences such as anoxia, limb fracture, extreme cold, and exposure to X rays. They involved potentially life-threatening damage, isolated the organism from its social context, and the stimuli were typically painful, unavoidable, and uncontrollable. These intense experiences, which lacked ecological relevance, masked variations in biological response produced by different contingencies and patterns of coping. They also led to the belief that only very profound and overwhelming experiences were likely to generate damaging physiological change. The remnants of this concept persist in today's literature with the emphasis placed on the study of major life events at the expense of milder everyday aversive experiences that may be cumulatively more significant.

It is now recognized that stress is a biopsychosocial construct, and cannot be defined purely in terms of physiological response. The physiological responses are only part of the profile that include changes at the affective, behavioral, and cognitive levels as well. Nor can stress simply be defined as the response to aversive environmental stimulation or disturbing events. The same situation may produce quite different responses in two people, or in one individual on different occasions. Most frameworks of contemporary research and clinical practice therefore recognize that stress responses arise through interactions between demands on the one hand, and psychosocial resources on the other. Cohen, Kessler, and Gordon (1995a), for example, have defined stress as a process in which "environmental demands tax or exceed the adaptive capacity of an organism, resulting in psychological and biological changes that may place persons at risk for disease" (p. 3).

This process, and the physiological responses that may be elicited, is outlined schematically in Figure 1. The psychosocial demands (or stressors) can be defined as potential or actual threats or challenges to the integrity, survival, and reproduction of the organism (Weiner, 1992). They may be anticipated, and may be real or imaginary. In the field of health psychology, the main categories of psychosocial demand are acute life events, chronic stressors, and day-to-day hassles and irritations. The characteristics of these demands that make them particularly potent elicitors of psychophysiological responses are described in Section 8.02.4.2. Adaptive capacity is operationalized in terms of resistance and vulnerability factors, and

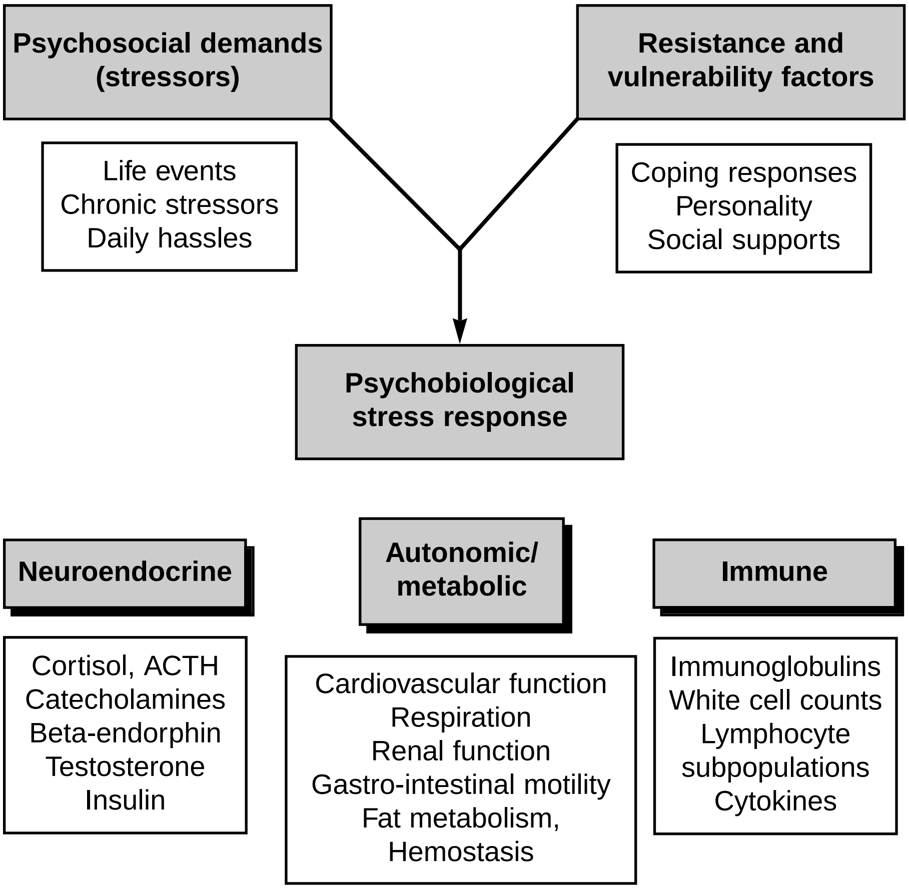
Psychosocial demands (stressors)
Life events
Chronic stressors
Daily hassles
Resistance and vulnerability factors
Coping responses
Personality
Social supports
Psychobiological stress response
Neuroendocrine
Cortisol, ACTH
Catecholamines
Beta-endorphin
Testosterone
Insulin
Autonomic/ metabolic
Cardiovascular function
Respiration
Renal function
Gastro-intestinal motility
Fat metabolism,
Hemostasis
Immune
Immunoglobulins
White cell counts
Lymphocyte subpopulations
Cytokines

Figure 1 Outline of the major physiological elements of the psychobiological stress response.

aspects of people's psychological makeup and social world that may render them more or less susceptible. Personality, behavioral coping style, and social resources are among the major resistance and vulnerability factors, and their role is examined in Section 8.02.5. Of course, it must be recognized that although, for convenience, demands and adaptive resources are viewed as distinct categories, they are difficult to disentangle in practice. For example, social isolation can be conceptualized both as the absence of a social resource or as a chronic stressor, while close social contact may be supportive or the source of major interpersonal conflict.

Psychophysiological responses are elicited when the organism's adaptive capacity fails to match demands. Typically, this will be under conditions of over-stimulation, but it is conceivable that lack of stimulation may also lead to adverse physiological change. The physiological responses themselves include adjustment in a number of biological systems, but the most important for health risk are neuroendocrine, immunological and autonomically-mediated changes in physiological function.

8.02.2.2 Research Strategies

Before the physiological systems perturbed by stressful transactions are detailed, a note about research strategies is needed. The clinical literature contains many sceptical comments about psychosocial factors in disease, typically based on negative results emerging from particular studies. It is important to recognize that evidence for a psychophysiological influence on disease does not rest with any one study or even with a single research paradigm. Rather, the evidence is assembled from a range of study types, each with their strengths and limitations.

8.02.2.2.1 Animal studies

Experimental studies with animals can document the ways in which pathological changes may emerge following exposure to different types of behavioral environment. Work with animals allows precise control over genetic background, rearing, and physical surroundings; factors that are difficult to regulate in human studies. Animals can be randomized to different conditions of behavioral stimulation for known periods of time, so true experimental designs are feasible. Detailed physiological assessments are practical, as is concurrent measurement of central neurochemistry.

As noted above, early work with animals tended to utilize physical insults; some of these methods, such as restraint stress, continue to be used (Glavin, Pare, Sandbak, Bakke, & Murison, 1994). Subsequently, experiments in which the behavioral contingencies surrounding physical insults are manipulated have helped to tease out the impact of psychological factors from the effects of pain or tissue damage. There have, for example, been numerous studies comparing controllable and uncontrollable electric shocks matched in intensity and duration, and these have demonstrated pronounced effects of behavioral control (Steptoe & Appels, 1989). However, stressors such as electric shock bear no resemblance to threats that animals encounter in their natural environments, so the value of studies on social challenges has been increasingly recognized. Social disruption and antagonism have been found to influence a range of physiological responses of potential pathological significance including blood pressure, coronary atherosclerosis (Kaplan, Manuck, Williams, & Strawn, 1993), cell-mediated and humorally-mediated immunity (Laudenslager & Fleshner, 1994), visceral fat accumulation (Jayo, Shively, Kaplan, & Manuck, 1993), and renal function (Von Holst, 1972).

Animal studies have clearly demonstrated that behavioral stimuli can elicit physiological responses of pathological significance, and have also underlined the importance of coping behaviors and social position. However, caution is required in their interpretation. Situations that promote pathophysiological responses in some species do not necessarily transfer to others, and there are important strain differences within the species as well (Gilad & Gilad, 1995). The dominant regulatory mechanisms for important target responses are not necessarily the same in different species. For example, glucocorticoids may play a more prominent role in regulating immune function under stress in rodents than in humans, in whom sympathetic pathways are salient (see Section 8.02.3.4). Finally, it is important not to be swayed unduly by superficial similarities between the types of challenge that provoke pathological responses in animal models and the demands to which humans are exposed.

8.02.2.2.2 Mental stress testing

Mental stress testing involves monitoring physiological responses in the laboratory or clinic during exposure to acute psychological challenge. A wide range of mental stress tests are employed, including cognitive and problem-solving tasks, emotionally charged interviews, public speaking tasks, upsetting movies and painful stimuli. Mental stress testing has been a dominant paradigm in cardiovascular research, and has also been utilized to investigate immune, musculoskeletal, respiratory and neuroendocrine responses. Both reactivity and recovery effects may be important (e.g., Seeman & Robbins, 1994).

The value of mental stress testing has been the subject of vigorous dispute. Briefly, the advantage of mental stress testing is that responses to standardized and uniform sets of stimuli can be monitored under environmentally controlled conditions, reducing many of the sources of bias and individual difference that are extraneous to the investigation. Experimental designs can be used with randomization to different conditions, and sophisticated measurement of physiological function is possible in clinical settings. The assertion that responses to mental stress tests show poor reproducibility seems not to be true provided that care is taken with procedures, although adaptation to repeated exposure may occur (Kamarck et al., 1992). However, there are other limitations. One is that use of standardized stimuli means that conditions may be somewhat divorced from everyday life. Few people spend much of their lives carrying out mental arithmetic under time pressure, or having ice packs placed on their foreheads. The monitoring of physiological activity during conversation and speech tasks goes some way towards evaluating ecologically more relevant situations. Second, physiological responses to standardized conditions may not always be representative of function outside the laboratory or clinic.

8.02.2.2.3 Field studies

Field studies are naturalistic investigations in which physiological measures are obtained while people go about their lives outside the laboratory or clinic. By taking repeated measures, it is possible to monitor the pattern of response to real-life stressors, or to assess associations with fluctuations in behavioral and emotional state. Three types of field study are relevant to psychophysiological aspects of disease. The first involves recording during challenging tasks such as parachuting, learning to fly aeroplanes, or speaking in public (e.g., Biselli, Farrace, Amelio, & Fattorossi, 1993; Van Doornen & Van Blokland, 1992). These investigations are closest to laboratory studies in that they evaluate responses to acute and relatively unusual events. The second type of study involves repeated assessments of physiological parameters as people go about their ordinary lives. For example, repeated samples of saliva can been obtained over the day and analyzed to measure the profile of cortisol release in people exposed to different levels of daily stress (Van Eck, Berkhof, Nicolson, & Sulon, 1996). The third type of field study involves using automated portable (ambulatory) apparatus to measure physiological activity in a relatively unobtrusive fashion from free living individuals. For instance, portable electrocardiograms (EKG) and ambulatory blood pressure monitoring equipment have proved valuable for assessing psychosocial factors related to cardiovascular disease, while electromyographic (EMG) assessments of muscle tension in the neck and forehead are feasible for the investigation of headache (Hatch et al., 1991 Pickering, 1991).

The main advantage of field studies is that they have ecological validity, in that physiological measures are obtained under naturalistic conditions of ongoing behavior and emotion. Studies of blood pressure in patients with hypertension suggest that ambulatory measures are even more predictive of pathological progression than are measures taken in the clinic (Pickering, 1991). The limits of field measures are essentially technological, but developments of instrumentation mean that even rather complex functions such as the dynamics of the filling and contraction of the heart can now be monitored.

There are, however, a number of factors that need to be borne in mind when interpreting these data. The first is that field measures are inevitably somewhat disturbing, since people are aware of the apparatus and may modify their behavioral responses. One study of ambulatory blood pressure demonstrated that in comparison with a control day, people were less physically active on the measurement day, and rested more (Blanchard, Cornish, Wittrock, & Jaccard, 1990). Second, there are numerous extrinsic influences on physiological function that need to be taken into account, including cigarette smoking, food and caffeine intake, and patterns of sleep and activity. Perhaps most problematic is distinguishing physiological responses related to physical activity and exertion from those elicited by psychosocial factors. For example, we have recently studied ambulatory blood pressure monitored from a group of men during an eight-hour work shift, and a comparable nonwork period (Steptoe, Roy, & Evans, 1996). Blood pressure was recorded every hour using portable apparatus, and ratings of activities were made in diaries. As might be expected, blood pressure was higher over the work than nonwork day, and the men also reported more frequent stress at work. However, the blood pressure readings were more likely to be associated with high physical activity at work than during nonwork periods, and this may have accounted in part for differences between the days. Sophisticated analytic procedures are required to take account of these factors (Schwartz, Warren, & Pickering, 1994).

8.02.2.2.4 Survey studies

The last type of study in which psychophysiological processes are evaluated involves measurement of physiological parameters in epidemiological surveys or in clinical samples. Typically, the physiological markers are recorded on a single occasion rather than repeatedly, with efforts being made to standardize recording conditions as far as possible. It is sampling of this type that established high blood pressure and elevated serum cholesterol concentration as standard risk factors for coronary heart disease, and recent years have witnessed extension of the methodology to the study of psychosocial factors. For instance, the relationship between blood pressure and psychological characteristics such as anger has been assessed in several surveys (Jorgensen, Johnson, Schreer, & Kolodziej, 1996). Immune measures derived from single blood samples have been shown to be disturbed in people under the chronic burden of caring for dementing relatives (Kiecolt-Glaser, Dura, Speicher, Trask, & Glaser, 1991). Concentration of fibrinogen, a substance involved in the process of blood clotting, has been related to social class, and may contribute to the socioeconomic gradient in coronary heart disease (Markowe et al., 1985).

Survey studies are valuable in identifying the biological parameters relevant to the influence of psychosocial disturbances on disease risk. Measures from large representative samples can be obtained at relatively low cost, prospective study designs can be employed, and potential confounders can be taken into account by statistical manipulation. However, survey studies provide limited information about the dynamics of physiological responses, their origins or consequences, and the causal sequences are often difficult to disentangle.

8.02.2.2.5 Convergence of methods

Efforts are increasingly being made to integrate the different methods outlined here. This trend has been particularly prominent in the cardiovascular area, with work on the interrelationship between responses to mental stress testing and ambulatory cardiovascular monitoring. The aim has been to discover whether acute blood pressure responses during behavioral tasks correlate with blood pressure levels recorded under naturalistic conditions (Turner et al., 1994). Another example is the extension of mental stress testing into the epidemiological framework, replacing the small selected samples typical of laboratory studies by large-scale evaluations (Carroll, Smith, Sheffield, Shipley, & Marmot, 1995). Developments in instrumentation technology are likely to promote this important trend, since at present the results of investigations carried out with different study types are often difficult to compare directly.

8.02.3 PSYCHOPHYSIOLOGICAL RESPONSES

The psychophysiological responses relevant to disease are many and varied. Each system has a complex organization involving the integration of autonomic, hormonal, and metabolic regulatory processes. The brain mechanisms responsible are also complicated, and several neurotransmitters are implicated, including serotonin, gamma-aminobutyric acid and dopamine (Gray, 1995; Williams, 1994). The aim here is to describe the major psychophysiological responses currently thought to be relevant to disease, and briefly to outline their regulation and effects. More detailed accounts will be cited where appropriate. Measurement issues will not be discussed, since there are texts directly devoted to these topics (Cohen, Kessler, & Gordon, 1995b).

8.02.3.1 Neuroendocrine Parameters

8.02.3.1.1 Adrenal steroid stress response

The adrenal steroids known as corticosteroids have held center stage in research on psychophysiological aspects of disease for many decades, largely as a result of Selye's work. There is, however, a growing feeling among investigators that this emphasis may be misplaced, originating as it does in rodent biology. In humans, the sympathetic nervous system and the catecholamines may be more important as mediators of peripheral physiological dysfunction. Nonetheless, there is no doubt about the extensive effects of corticosteroid release. Corticosteroids are produced by the cortex of the adrenal glands as a result of activity in the hypothalamic–pituitary–adrenal (HPA) axis. Neurosecretory neurons in the paraventricular nuclei of the hypothalamus contain a peptide called corticotropin releasing factor (CRF), which is transported through the hypothalamic hypophyseal portal system to the pituitary where it initiates the synthesis of adrenocorticotropin (ACTH) and beta-endorphin. The amygdala is in part responsible for stimulating CRF in response to stressors, while the hippocampus plays an inhibitory role (McEwen & Sapolsky, 1995). ACTH in turn travels through the circulatory system to the adrenal glands, where it stimulates production of steroid hormones such as androgens, estrogens, mineralocorticoids and glucocorticoids. The most important corticosteroid in the stress response is the glucocorticoid cortisol (corticosterone in rodents).

Corticosteroid production is regulated through negative feedback loops to the hippocampus, hypothalamus, and pituitary gland. Corticosteroid receptors in the hippocampus mediate a fast feedback loop responding to increases in plasma corticosteroids by inhibiting the HPA, while glucocorticoids inhibit CRF biosynthesis in the hypothalamus and ACTH release in the pituitary through a delayed feedback circuit that responds to tonic corticosteroid levels (Checkley, 1996). The dexamethasone suppression test is a method of assessing the feedback loop by measuring the suppression of HPA function with a synthetic glucocorticoid. Other important regulators of the HPA axis are cytokines, peptides released by activated immune cells that act on the brain structures involved in corticosteroid release, and these processes will be described in Section 8.02.3.4 (Besedovsky & Del Rey, 1996). Together, the regulatory mechanisms ensure that under normal circumstances, corticosteroid production is maintained within a narrow range.

Much of the cortisol secreted by the adrenals is inactive and bound to carriers. Free corticosteroids have a range of actions including stimulation of glucose production by the liver, the release of free fatty acids from adipose fat stores, the modification of water balance, alteration in vascular tone and the excretion of sodium by the kidney, the promotion of anti-inflammatory responses, and immunoregulatory effects. Disruption of corticosteroid output during stress responses may have a variety of adverse effects such as the elevation of lipid concentration in the bloodstream (which may in turn be incorporated into atherosclerotic plaque), deposition of fat in abdominal adipose tissue, the suppression of some immune functions, decalcification of bone, and impaired fertility (Sapolsky, 1992; Weiner, 1992). There is evidence that blood pressure increases with glucocorticoid stimulation, and cardiovascular reactivity to sympathetic nervous stimulation is enhanced (Whitworth, Gordon, Andrews, &

Scoggins, 1989). Within the central nervous system, sustained elevation of glucocorticoid concentration may lead to damage of the hippocampus, attenuating the hippocampal inhibition of CRF production. Increased levels of glucocorticoids are associated with clinical depression, and may promote hippocampal atrophy in humans (Holsboer & Barden, 1996).

Increased cortisol has been observed under a range of acutely challenging conditions in humans and animals. Early work in this area was reviewed by Mason (1968) who did much to dispel the notion that corticosteroid production bears a simple linear association with the emotional distress. Increased cortisol in the blood, urine, and saliva is found in response to mental arithmetic, speech tasks, simulated driving, and other challenges (e.g., Kirschbaum et al., 1995; Seeman et al., 1995). Chronic stress in animals has been shown to lead to a reduction in the number of corticosteroid receptors in the hippocampus, impairment of central nervous system inhibition of the HPA, and hypertrophy of the adrenal glands (Herman, Adams, & Prewitt, 1995). Results of studies of naturalistic stressors in humans have been somewhat variable, partly because measurement is complicated by circadian variations and the influence of factors such as physical exercise, smoking, caffeine, alcohol, and medication. A longitudinal investigation of an elderly population showed that plasma cortisol was elevated up to three months following severely threatening life events, and cortisol elevations are particularly marked among depressed people who experience long-term stressors (Willis, Thomas, Garry, & Goodwin, 1987). High cortisol has also been recorded from people exposed to major threats such as being a hostage of war (Rahe, Karson, Howard, Rubin, & Poland, 1990). On the other hand, ex-soldiers suffering from post-traumatic stress disorder show low cortisol levels, and an inverse relationship with intensity of combat exposure has been described (Boscarino, 1996).

Despite the opinion that many types of work are stressful, the results of investigations of corticosteroid levels in people at work have been inconsistent, and it appears that only a proportion of individuals respond even to very challenging occupations such as air traffic control (Rose & Fogg, 1993). Nonetheless, day-to-day stressors or hassles have been found to stimulate increased cortisol independently of factors such as smoking, food intake, and coffee consumption (Van Eck et al., 1996). There is sufficient evidence to infer that heightened cortisol secretion may be a mechanism through which psychophysiological factors influence risk of disease.

8.02.3.1.2 Catecholamines and the sympathetic nervous system

The autonomic nervous system is the branch of the nervous system that regulates the function of visceral organs, and has two divisions, the sympathetic and the parasympathetic. Many tissues are dually innervated, with stimulation of the sympathetic and parasympathetic nerves having opposite effects. For instance, heart rate accelerates under sympathetic stimulation and slows with parasympathetic (vagal) stimulation. The major neurotransmitter in the sympathetic nervous system is norepinephrine (noradrenaline) while acetylcholine is released from parasympathetic nerve terminals.

The sympathetic nervous system tends to induce activation in many tissues, so the second major component of the psychophysiological stress response is the sympathetic-adrenomedullary (or sympathoadrenal) system. Norepinephrine is stored in granules in the sympathetic nerve terminals, and is released when these nerves are stimulated. Much of the neurotransmitter binds to alpha-adrenergic and beta-adrenergic receptors on effector cells, while the remainder is released into the bloodstream. Sympathetic pathways also stimulate the synthesis of epinephrine (adrenaline) in the chromaffin cells of the adrenal medulla. A series of enzymes are involved in this process, including tyrosine hydroxylase, dopamine beta-hydroxylase, and phenylethanolamine *N*-methyltransferase, that are sometimes measured as indices of catecholamine production. The sympathetic nervous system and epinephrine from the adrenal glands typically operate in concert to stimulate a wide range of physiological responses that are summarized in Sections 8.02.3.2–8.02.3.4.

Catecholamine levels in the bloodstream are not easy to measure since responses are rapid and transient, and concentrations in venous blood may not be representative of the overall level of circulating epinephrine and norepinephrine. Other techniques such as the measurement of norepinephrine spill-over from nerve terminals, electrical recording from sympathetic nerves using microneurography, and measures of adrenergic receptor density provide more direct information. These techniques indicate that sympathetic activation is not a generalized phenomenon with mental stress, since the degree of sympathetic response differs substantially across tissues (e.g., Anderson, Wallin, & Mark, 1987). However, most human laboratory and field studies have measured catecholamines from venous blood samples or from urine. In the laboratory, acute increases in epinephrine and norepinephrine have been recorded in response to mental stress tests such as arithmetic, cold stimulation, and social conflict (Malarkey, Kiecolt-Glaser, Pearl, & Glaser, 1994). In field studies, elevations of catecholamines have been shown in diverse groups including the bereaved and victims of industrial disasters and sexual abuse (De Bellis, Lefter, Trickett, & Putnam, 1994; Gatchel, Schaeffer, & Baum, 1985).

A substantial amount of research has been carried out into sympathoadrenal responses to work. Studies from Sweden have shown that catecholamine excretion is elevated among people on piece work or those carrying out highly mechanized work (Johansson, Aronsson, & Lindstrom, 1978). Differences between sexes and people at higher and lower work grades have been reported, with some evidence for sustained elevations of norepinephrine in the evening among women in managerial positions (Frankenhaeuser et al., 1989). Frankenhaeuser, Lundberg, and Forsman (1980) have argued that behavioral performance efficiency under challenging conditions is positively associated with catecholamine secretion. However, one of the problems interpreting catecholamine responses is that excretion is positively correlated with physical effort and with motor activity; consequently, norepinephrine may increase when people are more actively engaged in tasks.

The sympathoadrenal system and the HPA axis are not independent of one another. The glucocorticoids have a role in regulating catecholamine biosynthesis in the adrenal medulla, while catecholamines stimulate ACTH release from the pituitary (Axelrod & Reisine, 1984). Under acute conditions, correlations have been observed between cortisol and sympathetically mediated responses in some but not all studies (Cacioppo, 1994). At the same time, there has been great interest in identifying situations that lead to differential release of corticosteroids and catecholamines. Extensive evidence (reviewed by Steptoe, 1983) suggests that HPA activation is associated with passivity in the face of aversive situations, with behavioral withdrawal, and with loss of control. Sympathoadrenal responses on the other hand are maximized under conditions of active efforts to cope and maintain control over the environment. This distinction may be significant when examining psychophysiological pathways to disease.

8.02.3.1.3 Endogenous opioids

Beta-endorphin and the enkephalins are endogenous opioid peptides, and are derived from the same precursor molecule as ACTH. The secretion of beta-endorphin from the anterior lobe of the pituitary gland is regulated by CRF, and occurs in response to a wide range of painful stimuli and stressors (Madden, Akil, Patrick, & Barchas, 1977). Inescapable electric shock in rats has been shown to produce stress-induced analgesia or increased pain tolerance. One form of this analgesia is dependent on endogenous opioids, in that it is blocked by antagonists such as naloxone, and shows cross-tolerance with morphine-induced analgesia (Terman, Shavit, Lewis, Cannon, & Liebeskind, 1984).

There is growing evidence that endogenous opioids are involved in the regulation of neuroendocrine stress responses (McCubbin, 1993). Serum concentration of beta-endorphin increases during mental stress tests, and assessments have also been made during naturalistic challenges such as anticipation of academic examinations (Malarkey, Pearl, Demers, Kiecolt-Glaser, & Glaser 1995; Sheps et al., 1995). When the opioid antagonist naloxone is administered under double-blind conditions, it leads to increases in stress-induced heart rate, epinephrine, ACTH and cortisol responses (Morris et al., 1990). This suggests that beta-endorphin normally operates to dampen excessive responses, and is part of the organism's self-regulatory mechanism, preventing overshoot of stress-induced physiological reactions. The pathway may operate less effectively in young people at elevated risk for high blood pressure, since their cardiovascular responses to mental stressors are unchanged by naloxone (McCubbin, Surwit, Williams, Nemeroff, & McNeilly, 1989). The increase in beta-endorphin with mental stress has been shown to correlate with raised pain thresholds in patients with coronary artery disease (Sheps et al., 1995). This may be one explanation for stress-induced episodes of reduced coronary blood flow (ischemia) occurring without concomitant chest pain in so-called "silent" ischemia.

8.02.3.1.4 Other neuroendocrine and hormonal responses

A range of other hormones are involved in psychophysiological responses. Dopamine in the brain increases in response to experimental stressors in animals, and peripheral dopamine may be particularly important in modulating gastrointestinal injury (Glavin, 1992). Reproductive function is inhibited by various components of the HPA system, with suppression of luteinizing hormone releasing hormone in the hypothalamus, and inhibition of testosterone and estrogen by glucocorticoids (Rabin, Gold, Margioris, & Chrousos, 1988). The classic studies of Christian (1970) showed that crowding in animals impaired reproductive efficiency. Social status has also been shown to affect ovarian function in cynomolgus monkeys, with more anovulatory cycles and progesterone-deficient luteal phases in subordinate animals (Adams, Kaplan, Clarkson, & Koritnik, 1985). Hypogonadal function in middle-aged men is associated with psychosocial risks such as work-related problems, living alone, lower social class, and high levels of physical symptoms (Nilsson, Moller, & Solstad, 1995). Vasopressin and prolactin have also been studied as components of psychophysiological responses, although their significance is not yet clear (Theorell, 1992).

8.02.3.2 Cardiovascular and Autonomic Parameters

Psychophysiological responses are organized hierarchically, and a series of physiological adjustments take place as a result of the neuroendocrine secretions described in Section 8.02.3.1 coupled with direct autonomic stimulation. They include changes in cardiovascular function, sweat gland activity, gastrointestinal motility and pulmonary ventilation. Not all of these responses are immediately relevant to disease risk. Studies using direct stimulation of brain regions in unanaesthetized animals indicate that emotional behaviors can be elicited that are associated with a range of damaging physiological changes such as cardiac arrhythmias, gastric erosions, and adrenal hyperplasia (Kojima et al., 1996). Cardiovascular and metabolic responses are particularly important when considering psychophysiological processes in disease.

8.02.3.2.1 Blood pressure, heart rate, and regional blood flow

Blood pressure and heart rate generally increase in response to acute challenges in the laboratory. During performance of tasks such as mental arithmetic, the increase in blood pressure and heart rate is associated with vasodilation in skeletal muscle and reduced blood flow to the kidneys, skin, and viscera (Tidgren & Hjemdahl, 1989). The high heart rate (tachycardia) is a result of sympathetic stimulation coupled with parasympathetic or vagal withdrawal. Blood flow in adipose tissue is increased, and is associated with enhanced lipolysis (Linde, Hjemdahl, Freyschuss, & Juhlin-Dannfelt, 1989). The sensitivity of the cardiac baroreceptor reflex, the mechanism through which increased blood pressure is offset by slower heart rates and vice versa, is inhibited during acute stress. Power spectrum analysis of heart rate and blood pressure signals show changes in spectral density at different frequencies indicative of these autonomic adjustments (Pagani et al., 1991). The capacity of the kidneys to excrete sodium and maintain fluid balance may also be impaired, and this may have implications for the development of hypertension (Light, Koepke, Obrist, & Willis, 1983).

Several factors need to be taken into account when evaluating acute stress responses in humans. First, response patterns depend critically on the nature of the demands imposed on the individual. Actively demanding situations such as information processing and problem solving tend to elicit cardiovascular responses that are mediated by increases in myocardial contractility leading to elevated stroke volume and cardiac output. In contrast, passive aversive conditions and situations over which people have little control elicit increases in peripheral resistance rather than cardiac responses (Sherwood, Dolan, & Light, 1990). Consequently, it is not appropriate to regard the cardiovascular response to acute stressors as a uniform phenomenon.

Second, individuals show characteristic response patterns that persist across different situations. Manuck (1994) identified groups of "cardiac" and "vascular" reactors, according to whether they showed increases in cardiac activity or total peripheral resistance during tasks. The magnitude of blood pressure response was similar in the two groups despite variation in the underlying pattern of haemodynamic adjustment. This means that individual differences as well as situational characteristics need to be taken into account (see also Section 8.02.6.7).

Third, there appears to be flexibility across hemodynamic mechanisms, so that certain aspects of cardiovascular responses are maintained during mental stress through alternative pathways when necessary. For example, when actively demanding mental stress tests are carried out during inhibition of sympathetic pathways by beta-adrenergic blockade, blood pressure elevations are maintained through other autonomic mechanisms (Julius, (1988). Similarly, adrenelectomized subjects show the same blood pressure response to mental arithmetic or other stimuli as healthy controls, despite the lack of epinephrine response and a reduced heart rate increase (Lenders, Peters, Pieters, Willemsen, & Thien, 1988). Julius (1988) had described this as a "blood pressure-seeking property" of the central nervous system, with integrative mechanisms in the brain operating to maintain sustained high blood pressure during challenge through a variety of physiological mechanisms. If this is true, it has important implications for interventions designed to reduce blood pressure stress responses.

8.02.3.2.2 Cardiac function

Several aspects of cardiac function may be disturbed as part of the psychophysiological stress response in vulnerable individuals (see Steptoe & Tavazzi, 1996). The two most important types of response are changes in the rhythm of cardiac contraction and alterations in blood flow in the coronary circulation.

Clinical disturbances of the generation and conduction of heart beats include various types of tachycardia and arrhythmia (Janse & Davies, 1996). Arrhythmias affecting the left ventricle of the heart (ventricular arrhythmias) can be provoked by sympathetic nervous stimulation, while the parasympathetic system is protective. Stimulation of the posterior hypothalamus in dogs reduces the threshold for potentially lethal ventricular fibrillation, and similar responses have been observed during avoidance conditioning and other types of psychological stress (Lown, Verrier, & Rabinowitz, 1977). In cardiac patients, ventricular arrhythmias have been associated with high levels of norepinephrine spill-over indicative of sympathetic activity (Meredith, Broughton, Jennings, & Esler, 1991). Tavazzi, Zotti, and Rondanelli (1986) evaluated the influence of mental stress on the electrophysiology of the heart using programed ventricular stimulation in postinfarction patients. Transient ventricular tachycardia and fibrillation were induced in patients during psychological stress, but not under resting control conditions. These studies suggest that the disturbances of cardiac rhythm can be induced by activation of psychophysiological pathways.

High levels of parasympathetic or vagal tone act in opposition to sympathetic traffic, and reduce vulnerability to certain arrhythmias. Clinically, it has been observed that high heart rate, which may be a reflection of low vagal activity, may be an independent risk factor for coronary heart disease (Dyer et al., 1980). The minimum heart rate achieved during 24-hour ambulatory monitoring is positively associated with coronary atherosclerosis as determined by angiography (Perski et al., 1992). Parasympathetic activity is also related to heart rate variability, with higher variability being indicative of greater parasympathetic tone. Mortality in postmyocardial infarction patients has been related to low heart rate variability, and this may be another important marker of the psychophysiological response (Klieger, Miller, Bigger, & Moss, 1987).

Blood flow abnormalities in the coronary circulation can provoke ischemia and abnormalities of contractility and left ventricular wall motion. In classical angina pectoris, cardiac ischemia due to increased myocardial oxygen demand during exercise is associated with chest pain, but episodes of "silent" ischemia (without pain) are also common. Mental stress may stimulate cardiac ischemia in a manner similar to exercise, with raised blood pressure and heart rate leading to heightened myocardial oxygen demand. Alternatively, there may be a direct influence of the central nervous system on the vascular tone of the coronary arteries and arterioles, leading to reductions in coronary blood flow.

Sustained infusions of norepinephrine have been shown to induce myocardial lesions indicative of ischemia (Schenk & Moss, 1966). Several new methods of assessing cardiac function in conscious people have allowed stress-induced silent ischemia to be evaluated in detail. In an early study, positron tomography was used in a study of 16 angina patients (Deanfield et al., 1984). Twelve patients showed abnormalities of perfusion of the heart during mental arithmetic, but pain was experienced by only four. Rozanski et al. (1988) documented a high incidence of left ventricular wall motion abnormalities during mental stress in patients with coronary artery disease using radionuclide ventriculography, with public speaking being a particularly provocative stress test. Similar results have emerged with other methods (Steptoe & Tavazzi, 1996).

Studies of ischemic responses under field conditions have largely been confined to examining abnormalities of the EKG rather than more sensitive measures, although the introduction of the ambulatory "nuclear vest" may be helpful in the future (Legault, Freeman, Langer, & Armstrong, 1995). Gabbay et al. (1996) assessed the frequency of ST-segment depression (an EKG change indicative of myocardial ischemia) in patients during ambulatory monitoring. It was found that both physical exertion and mental stress (anxiety and anger) were associated with ischemic episodes. The presence of ischemic responses to mental stress in the laboratory has recently been found to predict silent ischemia during ambulatory monitoring in a large sample of patients with coronary artery disease (Blumenthal et al., 1995). Interestingly, positive correlations have also been recorded between the number and duration of ischemic episodes in daily life, and heart rate responses to mental stress in the laboratory (Krittayaphong, Light, Biles, Ballenger, & Sheps, 1995). These results indicate that laboratory responses to mental stress may have genuine significance for cardiac function in patients with coronary artery disease.

8.02.3.2.3 Platelets and hemostasis

The ability of blood to coagulate rapidly is one of the primary defense mechanisms reducing the impact of injury. One of the early stages of the clotting process is the adhesion of blood platelets to the lips of a wound and the formation of a platelet plug. This is followed by reinforcement of the platelet plug with fibrin. Later the deposit is removed by fibrinolytic and other mechanisms. This process is also involved in the formation of the thrombi in the vascular system that contribute to clinical events such as myocardial infarction and stroke.

It has been known for many decades that catecholamines stimulate the activation of platelets and their aggregation (Ardlie, Glew, & Schwartz, 1966). Platelet activation and aggregation have been shown to increase in healthy subjects carrying out challenging mental stress tests (Larsson, Hjemdahl, Olsson, Egberg, & Hornstra, 1989). Rapid increases in platelet activation and aggregability have also been recorded in patients with coronary artery disease performing mental arithmetic (Grignani et al., 1991). The implication of such findings is that vulnerability to the formation of thrombi may be enhanced as part of the psychophysiological stress response. Unfortunately, results of naturalistic studies have been rather inconsistent, with both increases and decreases in platelet activation and aggregation being reported under challenging conditions (Haft & Arkel, 1976, Levine et al., 1985). One complication is that there are several markers of platelet function and they may not produce uniform results.

The later stages of the coagulation cascade culminate in the action of the enzyme thrombin, which converts fibrinogen into fibrin. There has been limited study of this process in the psychophysiological context, although one study showed a decrease in fibrinogen and other blood coagulation factors during a prolonged episode of work without sleep (Palmblad et al., 1977), while work stress has been associated with elevated fibrinogen levels (Davis, Matthews, Meilahn, & Kiss, 1995).

8.02.3.2.4 Other physiological systems

Respiration is regulated by brain stem mechanisms that are closely integrated with cardiovascular control systems. During challenges that involve active efforts to cope, increases in respiration rate are typically observed, while the depth of breathing (tidal volume) is augmented. Changes in the constriction of the airways can also be induced as part of the psychophysiological stress response. Restriction of the airways (or bronchoconstriction) is stimulated not by the sympathetic but by the parasympathetic nervous system, while epinephrine relaxes the smooth muscle of the bronchial tree. The airways therefore operate in the opposite fashion to cardiovascular parameters, in that problems are more likely to be induced by parasympathetic than sympathetic stimulation. Effects are found most prominently among people suffering from bronchial asthma, for whom challenges such as the suggestion that an agent which causes wheezing is being inhaled can stimulate pronounced bronchoconstriction (Steptoe, 1984b). These responses are blocked by atropine and other anticholinergic agents, confirming that the parasympathetic nervous system is involved. Passive aversive situations in which the person is not required to respond actively but rather to endure stressful conditions are especially provocative (Lehrer et al., 1996). There is some evidence that bronchoconstrictive responses can be elicited in nonasthmatics as well as asthmatic individuals, but the findings are controversial.

The gastrointestinal tract is innervated by the autonomic nervous system in such a way that stimulation of the sympathetic branch decreases digestive functions such as saliva production, bile secretion, and gut motility, while these processes are enhanced by ACTH and parasympathetic responses. Large individual differences have been recorded in studies of gastric acid secretion with mental stress (Holtmann, Kriebel, & Singer, 1990). Thompson, Richelson, and Malagelada (1982) have shown how a variety of acute stressors, such as the cold pressor test or ear irrigation with cold water, delay gastric emptying and perturb duodenal motility. Other measures of gastrointestinal function are disturbed with acute stressors in people suffering from disorders such as duodenal ulcer and irritable bowel syndrome (Kumar & Wingate, 1985).

Changes in skeletal muscle tension are important components of the psychobiological stress response. These reactions are difficult to quantify without elaborate instrumentation, since no one set of muscles is a reliable indicator of general tension. Levels of tension in various tissues are heightened among patients with anxiety disorders, and acute stressors may also affect both level and variability of muscle tension (Hoehn-Saric, McLeod, & Zimmerli, 1989). The role of these responses in headache and low back pain is discussed in further detail in Section 8.02.7.3.

8.02.3.3 Lipids and Glucose Metabolism

The level of blood cholesterol is well recognized as a risk factor for coronary artery disease, so the influence of psychosocial factors on cholesterol metabolism is of great interest. Cholesterol is a fat-like material synthesized mainly in the liver, and is transported through the circulation in the form of lipoprotein particles. These protein molecules are of several types. Low-density lipoproteins (LDL) carry much of the cholesterol, and are the main culprit in coronary artery disease. High-density lipoproteins (HDL) also carry cholesterol but are protective in that they clear lipid from blood vessel walls. In addition, triglycerides are fats that are not attached to proteins; they may be associated independently with cardiovascular disease risk, although the importance of this mechanism remains uncertain.

Indirect evidence for the involvement of psychophysiological stress responses in lipid metabolism comes from experiments showing that blood pressure, heart rate, and catecholamine responses are correlated with the concentration of cholesterol and lipid fractions (e.g., Fredrikson, Lundberg, & Tuomisto, 1991; Suarez, Williams, Kuhn, Zimmerman, & Schanberg, 1991). If people prone to large physiological stress responses show higher lipid levels, then a common pathway of sympathetic or HPA activation may be responsible. Psychological challenges can also lead to acute increases in total cholesterol and LDL-cholesterol concentration (McCann et al., 1995). However, it has been discovered that this response may be secondary to alterations in blood volume. Plasma volume shows small reductions during mental stress tests because of shifts between intra-cellular and extra-cellular compartments. This can lead to an increase in blood protein concentration that accounts for the apparent elevation in blood lipid levels.

The impact on lipids of longer-term stressors such as anticipation of academic examinations, stressful work, and threat of unemployment has also been evaluated (Brindley, McCann, Niaura, Stoney, & Suarez, 1993). Results have been quite mixed, with increases in total cholesterol and LDL-cholesterol in some studies but not others. The problem in understanding these responses is that under naturalistic conditions, changes may take place in dietary composition, smoking, and other factors that influence lipids independently of any psychophysiological processes. Clearer evidence derives from animal studies in which diet and other factors can be controlled. Infusion of epinephrine has been shown to increase cholesterol concentration in primates and rodents, while socially subordinate baboons have lower HDL-cholesterol than dominant animals, and this is coupled with a higher cortisol concentration (Sapolsky & Mott, 1987). However, it is striking that in the series of studies of social stress in cynomolgus monkeys carried out by Kaplan, Manuck and co-workers, increases in coronary atherosclerosis have been observed that were not due to elevations in lipid concentration (Kaplan et al., 1993; Manuck, Kaplan, Adams, & Clarkson, 1989). In this animal model, social stress promotes atherosclerosis independently of lipid concentration, through mechanisms such as damage to the endothelial surface of arterial walls, which may permit greater deposition of lipids in vessel walls (Strawn et al., 1991).

8.02.3.3.1 Insulin metabolism and abdominal obesity

One of the critical mechanisms linking neuroendocrine activation with lipid metabolism is insulin. Insulin is a hormone produced by the cells of the Islets of Langerhans in the pancreas, and facilitates the utilization of glucose by tissues. Destruction of the insulin-producing cells leads to insulin-dependent (or Type 1) diabetes mellitus, in which protein and fat are metabolized instead of glucose, resulting in accumulation of toxic substances in the blood and kidney damage. A more common disorder is noninsulin dependent (or Type II) diabetes, in which high levels of circulating glucose (hyperglycemia) arise either because of failure to produce adequate insulin, or because of defects in the ability of tissues to respond to insulin (insulin-resistance).

Both the sympathoadrenal and HPA axes influence insulin metabolism, affecting its secretion, the conversion of glycogen to glucose in the liver, and glucose uptake in tissues. A constellation of metabolic disturbances (sometimes known as syndrome X or the insulin resistance syndrome) coexist in people at raised risk for cardiovascular disease. Their problems include insulin resistance, hyperinsulinemia, glucose intolerance, low concentrations of HDL-cholesterol and high blood pressure (Reaven, 1988). Another element of this risk profile is obesity, due particularly to deposition of fat in the central or abdominal fat stores. Bjorntorp (1990) has argued that a disturbance of corticosteroid control over lipid metabolism may underlie the development of abdominal obesity. Cortisol responses to mental stress tests have been shown to correlate positively with the degree of abdominal obesity (Moyer et al., 1994), while animal studies have demonstrated that social stress promotes deposition of abdominal fat (Jayo et al., 1993). It appears, therefore, that lipid and glucose metabolism play important roles in psychophysiological responses, and may be relevant to the pathogenesis of coronary heart disease and diabetes.

8.02.3.4 Immune Parameters

One of the most exciting developments in the psychophysiological studies of disease over recent decades has been the discovery that immune function may alter as part of the stress response. Knowledge of these links increases understanding of the mechanisms through which psychosocial factors can influence vulnerability to infection, inflammation, autoimmune disease, and certain types of cancer. Sweeping assertions that stress suppresses immune function are not appropriate, since it is now recognized that the central nervous system and immune system engage in a complex interplay, and that immune responses can be both up-regulated and down-regulated by different behavioral and emotional states.

The immune system is the body's primary defense against infection and invading pathogens such as viruses, bacteria, and fungi. Some immunity is innate and is present from birth, acting nonspecifically to protect against foreign materials. The skin and mucus membranes are part of this innate immune defence, using autonomic and chemical methods to prevent invasion, while cells such as macrophages are also involved. Local invasion can also trigger a general or systemic "acute phase response," in which several general bodily processes are elicited, including fever and raised body temperature.

Although the innate system is highly effective, it cannot cope with all pathogens, so there is a second form of immunity which is acquired or specific. Acquired immunity involves the recognition of substances (antigens) as "nonself," and the destruction or elimination of these materials. The immune system has memory, reflected in the fact that secondary exposure to antigens elicits a more vigorous immune response than the initial reaction, although the mechanism is poorly understood (Ahmed & Gray, 1996). The main organs of the immune system include the bone marrow, thymus gland, spleen, and lymphatics, but there are also cells active in the skin, lungs and gut.

There are two major arms of the immune system. Humoral immunity is involved in defense against bacteria and viruses in body fluids, while cell-mediated immunity is relevant to intracellular viruses and fungi, and also to cancer cells and transplanted tissue. Humoral immunity is mediated by serum antibodies or immunoglobulins. These are proteins that derive from B-lymphocytes in the bone marrow, and react with specific antigens. There are five major classes of immunoglobulin, IgA, IgM, IgG, IgE, and IgD, and each has unique characteristics. For example, IgE has a primary role in defense against parasites, and elicits a range of responses designed to exclude these organisms including bronchoconstriction, vomiting, inflammation, itching, and coughing. It also has a key role in allergic responses which are generally caused by overproduction of IgE (Sutton & Gould, 1993). IgA is found in secretions such as saliva and tears, and is important in defense against local infections in the gut and respiratory system.

Cellular immunity involves T-lymphocytes that arise in the bone marrow and mature in the thymus before circulating in the blood and lymph. T-cells do not recognize antigens by themselves, so antigens are usually presented to them by macrophages. T-cells have to be activated, but once activated become memory cells, migrating to tissues in which they are most likely to re-encounter their specific activating antigen. They are differentiated into various types. Of particular importance to work on stress and immune function are cytotoxic T-lymphocytes that kill invading pathogens and are involved in defense against infections such as influenza and the herpes simplex virus. T-cells are classified by surface cell markers, and cytotoxic T-cells fall into the CD8+ class. Helper T-cells (CD4+) play a central regulatory role in immune defense, stimulating B-cells to produce antibody, activating macrophages to engage in antimicrobial activity, and cytotoxic T-cells. Some of these actions are mediated through immune products called cytokines that have very specific actions on other branches of the immune system. Another type of lymphocyte, the suppressor T-cell, largely falls into the CD8+ category, but is less well understood. Suppressor cells are thought to shut off the activity of helper cells, and the helper/cytotoxic-suppressor ratio is sometimes assessed as a marker of immune health. Natural killer cells are a further class of immune cell, and are involved in immune surveillance, acting nonspecifically to defend against certain tumor cells and virus-infected cells.

Many different assays of immune function have been used in the context of health psychology (Kiecolt-Glaser & Glaser, 1995). They include enumerative assays or counts of various cell types, measures of concentration of immunoglobulins in the blood and saliva, measures of cytokine activity, and functional assays. Functional assays include testing responses to mitogens, which are substances that stimulate replication or proliferation of T and B cells. Common mitogens include phytohemagglutinin (PHA) and concanavalin A (Con A), with larger responses indicating more effective function. Natural killer cell activity (or cytotoxicity) is assessed by measuring the ability of natural killer cells to lyse target cells from known tumor cell lines.

8.02.3.4.1 Regulation of immune function

Both the HPA and sympathoadrenal pathways are involved in immune regulation (Besedovsky & Del Rey, 1996). There are receptors in the cytoplasm of lymphocytes for cortisol, and glucocorticoids suppress the activation of circulating lymphocytes and inhibit production of cytokines (Chrousos, 1995). The HPA plays a major role in the regulation of inflammation, and corticosteroids have long been used as anti-inflammatory agents. In addition, stress-induced production of CRF by the hypothalamus has been shown to reduce T-cell proliferation in response to mitogens independently of adrenal mechanisms (Jain et al., 1991). The sympathetic nervous system innervates immune organs such as the spleen and lymph nodes, and there are adrenergic receptors on circulating lymphocytes. Many of the immune responses to acute stressors are similar to those elicited by injection of epinephrine. It has been argued that sympathetic stimulation may lead to contraction of smooth muscle in the spleen which results in extrusion of lymphocytes into the circulation. Alternatively, the sympathetic nervous system may inhibit production of the cytokine interleukin-2 by helper T cells. Catecholamines also modulate natural killer cell activity.

In many studies of the stress process, it is not possible to determine whether HPA or sympathoadrenal pathways (or both) are responsible for stress-induced changes in immune function. However, there are striking examples of chronic increases in corticosteroid activity being maintained with behavioral stressors in the absence of changes in immune activity (Klein et al., 1992). In addition, alterations of immune function may occur in acute settings before changes in circulating corticosteroid concentrations have evolved. In these cases, sympathoadrenal activity is almost certainly responsible for regulating immune function.

The brain and immune system engage in a two-way dialogue, and the flow of information and control is not simply from the central nervous system to the periphery. Cytokines such as interleukin-1 are released during immune activation and mediate the host's response to infection and inflammation. Cytokines also act centrally, affecting neural growth and repair, stimulating sickness behaviors such as reduced social exploration, food intake and weight loss, and in addition disturbing cognitive functioning (Aubert, Vega, Dantzer, & Goodall, 1995). Dantzer and Kelley (1989) have argued that many of the nonspecific symptoms of sickness and behavioral responses to infectious illness are mediated by immune products, helping to coordinate the central and peripheral responses to infection.

8.02.3.4.2 Acute stress-induced immune responses

A number of studies of cellular immune responses to acute challenges such as problem solving and public speaking have been reported. Among the most consistent effects are an increase in the number of circulating CD8+ and natural killer cells, and reductions in mitogen-induced lymphocyte proliferation (Herbert et al., 1994). Changes in natural killer cell activity have been variable, with increases in cytotocity in some studies and inhibition in others (Bachen et al., 1992; Gerritsen, Heijnen, Weigant, Bermond, & Frijda, 1996). In a particularly striking set of studies, Kiecolt-Glasser et al. (1993) assessed newly-wed couples during a discussion of martial problems, and analyzed neuroendocrine and immune functions from blood samples. They showed that those who manifest critical interpersonal behaviors produced greater decrements than did other couples in natural killer cell activity and mitogen-induced lymphocyte proliferation.

The evidence that these responses are related to sympathoadrenal pathways is strong. The magnitude of immune response has been shown to correlate with the magnitude of cardiovascular reactions (Herbert et al., 1994). Furthermore, many of these responses can be attenuated by the blockade of adrenergic pathways with pharmacological agents (Bachen et al., 1995; Benschop et al., 1994). In contrast with work on cell-mediated effects, studies of acute changes in immunoglobulins have been very limited, although some investigation of IgA concentration in the saliva have been carried out (Herbert & Cohen, 1993b). Immunoglobulins have long half-lives, and salivary IgA may be influenced by saliva flow rate, making interpretation difficult.

8.02.3.4.3 Chronic changes in immune function

There is an extensive literature in animals demonstrating the impact of chronic stressors on immune function such as mitogen-induced lymphocyte proliferation (Coe, 1993; Line et al., 1996). It appears that these effects are ameliorated by social affiliation, which may be an analogue of social support (Cohen, Kaplan, Cunnick, Manuck, & Rabin, 1992). In humans, some of the changes in immune function associated with long-term threats are similar to those observed with acute challenges. Thus a reduction in mitogen-induced lymphocyte proliferation to PHA has been recorded in caregivers for demented patients, bereaved people, and in couples experiencing severe disruption of marital relationships leading to separation or divorce (Herbert & Cohen, 1993b; Kiecolt-Glaser, Malarkey, Cacioppo, & Glaser, 1994). The effects are rather persistent, with deficits being recorded over several years, and responses may be associated with depression (Herbert & Cohen, 1993a). Changes in the number of helper T-cells and cytotoxic-suppressor cells have been inconsistent, with modifications being recorded in some studies but not others.

One important difference from acute reactions is that instead of an increase in natural killer cell number, a reduction has generally been observed with chronic stressors (Castle, Wilkins, Heck, Tanzy, & Fahey, 1995). In addition, natural killer cell cytotoxicity is suppressed. The reason for these differences from the responses to acute stressors is unclear, but immune parameters are influenced by lifestyle variables such as smoking, exercise, and intake of dietary fat and vitamins. These variables are not controlled in naturalistic studies, but may be modified among individuals enduring chronic threats, so affecting the immune responses that are recorded. In addition, the pattern of hormonal reaction may alter with long-term threat, and also has a bearing on immune competence.

One interesting study of innate immune responses was published by Bosch et al., (1996), who assessed the aggregation of streptococcal bacteria in the saliva of students under normal conditions and then prior to an academic examination. Bacterial aggregation was reduced in the pre-exam period, suggesting that the saliva was less effective in preventing bacterial adhesion to oral surfaces, and defending against the early stages of infection. As far as immunoglobulins are concerned, studies have been carried out relating nonspecific salivary IgA to mood fluctuations (Evans, Bristow, Hucklebridge, Clow, & Walters, 1993). No differences in the production of IgM or IgG in adults exposed to severe long-term threat and control were recorded by McKinnon, Weisse, Reynolds, Bowles, and Baum (1989), while another study recorded associations with daily hassles (Jabaaij, Benschop, et al., 1993a). Rather consistent changes in antigen-specific salivary IgA and serum antibody response have also been recorded in studies by Stone et al., (1994). Subjects consumed a capsule containing a novel oral antigen on a daily basis, so the IgA antibody response was elicited in a way that might be analogous to invasion by a novel virus. Over several weeks, desirable daily events were associated with elevated IgA antibody production, while undesirable events led to a suppressed antibody response.

8.02.3.4.4 Functional significance of stress-induced immune modulation

An issue that arises in the investigation of immune responses is whether the changes are of clinical significance, or are so small as to be solely of academic interest. Many alterations in immune function will only affect health if an invading pathogen is present, so a short-term change in function may have few consequences if it does not coincide with exposure. There are now studies that have addressed functional significance more directly. One of the first models to be used was the response to latent viruses that remain within the host undetected unless deficits in immune control allow replication to increase. Large antibody titers to latent herpes viruses such as herpes simplex or Epstein-Barr virus indicate poor control by the cellular immune system over these pathogens. Reports by Kiecolt-Glaser, Glaser and colleagues have shown that antibody titers are elevated as part of the stress response (Kiecolt-Glaser et al., 1994).

A second approach has been to assess responses to nonpathogenic antigens, since studying the immune response to experimental infection is rarely ethical. Vaccination provides a useful model, but results to date have been inconclusive. One study of responses to influence of vaccine showed that psychological distress was associated with poor lymphocyte proliferative responses, while a second study showed no significant effects. Poor immune responses to hepatitis B vaccines have been elicited by acute or chronic stressors in some investigations but not in others (Glaser et al., 1992, Jabaaij et al., 1996). Results evidently depend critically on the timing of measures of the immune response, and on the size of the vaccination dose.

A third method that shows considerable promise relates to wound healing. Cellular immunity plays an important role, since cytokines such as interleukin-1 and tumor necrosis factor help to protect against infection and repair the damaged tissue. An investigation of caregivers for Alzheimer victims and matched controls has been published (Kiecolt-Glaser, Marucha, Malarkey, Mercado, & Glaser, 1995). Participants were given a punch biopsy to create a standard small wound on the forearm, and the rate of healing was assessed in the two groups. Wound healing was delayed among care-givers, and cytokine responses were impaired. This approach provides direct evidence for stress-induced suppression of immune defense being significant for health, and may be relevant to studies of psychosocial influences over recovery from surgery.

8.02.3.5 Adaptive or Destructive?

It has been argued since the earliest scientific investigations that the physiological activation seen under conditions of behavioral stress is an adaptive response. Its "purpose" is to prepare the organism for vigorous physical work such as might be called upon with fighting or escaping from danger. Thus, stored energy supplies are mobilized to feed working muscle, clotting processes are stimulated to prevent blood loss from wounds, and vegetative functions that are not essential for short-term activity are suppressed. Unfortunately, this phylogenetically primitive psychophysiological response is stimulated in humans under circumstances that do not call for vigorous exercise, and may therefore become destructive.

There are, however, features of psychophysiological responses that are puzzling from this perspective. They include the inhibition of inflammatory responses, the suppression of cell-mediated immunity, and the promotion of fluid loss through inhibition of the antidiuretic hormone vasopressin. How can the hypothesis that the response pattern is adaptive be reconciled with the evidence that some of the body's major defense systems are disrupted? How could such an organization of physiological responses survive the pressures of evolution?

Various explanations of this apparent anomaly have been put forward. One possibility is that suppression of immune responses is adaptive in directing energy to more pressing needs. Maier, Watkin, and Fleshner (1994) have argued that the production of acute phase responses, which include substantial cellular proliferation plus increases in body temperature, requires an increment in caloric energy production of at least 10%. Such responses might be difficult to sustain in the presence of the simultaneous demands of vigorous exercise. A second possibility developed by Besedovksy and Del Rey (1996) is that it may be adaptive for a species if on occasion immune systems fail and lead to the death of an individual, since this may prevent transmission of pathogens to conspecifics. This active negative self-selection may sacrifice the individual to ensure survival of other members of the population.

Another view that has gained widespread acceptance is that the role of the HPA axis is not to stimulate defense reactions, but rather to act as a regulatory mechanism and a restraint on other defenses. Munck, Guyre, and Holbrook (1984) have proposed that glucocorticoids play a central role in preventing the overshoot of inflammatory, immune, and metabolic responses that might otherwise take place when the organism is challenged. They suggest, for example, that in the absence of glucocorticoid counter-control, high levels of insulin might cause dangerous hypoglycemia, and that the suppression of vasopressin prevents excessive fluid retention. Chrousos (1995) has extended this perspective by pointing out that corticosteroid release may protect against susceptibility to autoimmune diseases. These are diseases in which the body's own antibodies cause inflammation and tissue destruction, and include rheumatoid arthritis, pernicious anemia and various types of thyroid dysfunction. In the absence of sustained control over immune responses by corticosteroids, risk might be increased.

8.02.4 THE NATURE OF PSYCHOSOCIAL DEMANDS

It will be apparent from the discussion thus far that a wide range of physiological responses are mobilized as part of the stress response. The question arises of what types of psychosocial demand stimulate these responses, and what characteristics of stressors are particularly relevant.

8.02.4.1 Life Events, Chronic Stressors, and Daily Hassles

Adverse experiences that have an impact on psychophysiological responses are conveniently divided into acute events that can be pinpointed in time and cause substantial disruption to life, chronic adverse social factors that may continue for months or even years, and less severe day-to-day sources of irritation (daily hassles). All three types of phenomenon have been associated with psychophysiological responses and subsequent health risk. Among the acute events, earthquakes and other natural disasters may be followed by elevations in health problems and mortality (e.g., Adams & Adams, 1984, Trevisan et al., 1992). Bereavement through death of a partner leads to responses such as immune down-regulation and an increase in all-cause mortality risk (Jones, 1987), while other life events have been associated with problems such as premature delivery for pregnant women (Nordentoft, Lou, & Hanson, 1996). Chronic stressors, such as living in the presence of loud ambient noise, crowding, unemployment, or caring for dementing relatives, also have effects on psychophysiological processes, and in some cases on mortality (Evans, Hygge, & Bullinger, 1995; Kiecolt-Glaser et al., 1994; Morris, Cook, & Shaper, 1994). Daily hassles have been linked with fluctuations in salivary cortisol levels, blood pressure, and IgA (Steptoe et al., 1996; Stone et al., 1994; Van Eck et al., 1996).

The quantification of psychosocial demands has been discussed and criticized extensively elsewhere (Cohen et al., 1995b). One of the issues to emerge is that many standard measures of life events may misrepresent the impact of adverse experiences by failing to be comprehensive (Turner, Wheaton, & Lloyd, 1995). If certain types of events are more common in one sector of society than another, then failure to include such experiences in the assessment instrument would lead to underestimation of exposure. This may account for the observation that has repeatedly been made in the literature that young adults experience more stressful events than older people; this is because many life event measures target events such as pregnancy, marriage, and so forth at the expense of the infirmities and loss of capacity that are more common among the elderly. Aneshensel (1992) has argued that life event research may neglect the problems of ordinary social life and that the latter show strong gradients with socioeconomic status that may not be apparent for acute events. Although it is convenient to distinguish acute events from chronic stressors, in reality the classification is far from clear cut. Many important life events have long-term ramifications that either precede the actual event (such as discord leading up to a marital separation), or last well beyond the incident itself. Single acute events may show a sustained capacity to elicit physiological activation through repetitive intrusive memories and other cognitive transformations.

Rather little is known about the extent to which the impact of acute events or chronic severe stressors is mediated through daily hassles. An event such as the death of a spouse is distressing in itself, but may also lead to numerous petty difficulties in financial and legal matters, the practicalities of living, and in social life. An interesting analysis of this issue was described by Pillow, Zautra, and Sandler (1996), who investigated the links between major threats such as death of a spouse, divorce, or having a child with chronic illness, daily hassles, and psychological distress. Different links between hassles and the various types of major threat were observed. The impact of divorce on psychological distress was mediated almost entirely through daily hassles, while the impact of a serious childhood illness was independent of daily hassle levels. The associations between bereavement and distress were partly direct, and partly mediated through an increase in daily hassle levels.

These issues have implications for systematic investigations of psychosocial factors in health risk. One of the major problems encountered in the study of physical illness is the "independence" of events from the prodromata of clinical illness. Many cross-sectional or retrospective studies of psychosocial factors examine life events or chronic stressors in case-controlled designs, comparing, for example, women who show a recurrence of breast cancer with nonrecurrence (Ramirez et al., 1989), or insulin-dependent diabetics and their siblings (Robinson & Fuller, 1985). Inferences concerning the causal sequence depend critically on the timing of adverse life experiences. But many threatening events such as being sacked from a job, or a breakdown in important relationships, may be the culmination of a series of hassles, some of which may have preceded physical pathology, or have even been caused by early symptoms and limitations in capacity brought on by the medical condition itself.

8.02.4.2 Characteristics of Psychosocial Demands

The arguments outlined in earlier sections indicate that psychophysiological responses are not uniform across situations, but vary with the nature of demands. Some of the broad dimensions that appear relevant are described in this section.

8.02.4.2.1 Novelty, familiarity, and predictability

Psychophysiological responses tend to be more intense when animals or humans are placed in unfamiliar situations and confronted with novel demands. This was strikingly illustrated by Mason (1975) who found that the neuroendocrine reactions elicited in primates were as great when animals were placed in the experimental situation for the first time, as they were to any subsequent "stressors." The neuroendocrine and autonomic reactions in people learning new dangerous skills such as parachuting also typically diminish rapidly with repeated exposure.

The reduction in psychophysiological responsivity with familiarity may be due to several factors. Natelson et al. (1988) have argued that adaptation in physiological stress responses is similar to the process of sensory habituation. Many biological systems show reduced responsivity over time, with alterations in biosynthesis, storage, secretion and re-uptake capacity, modulation of receptor sensitivity, and the institution of counter-regulatory feedback mechanisms. Familiar conditions may also elicit smaller psychophysiological responses because they become predictable. Psychophysiological responses are more pronounced in unpredictable conditions, even if the duration and intensity of stimulation is the same as that in predictable conditions (Abbott, Schoen, & Badia, 1984). To take just a few examples, rat studies have shown that corticosteroid responses are enhanced and the proliferative capacity of lymphocytes is suppressed by unpredictable as opposed to predictable stressors (De Boer, Van der Gugten, & Slangen, 1989; Mormede, Dantzer, Michaud, Kelley, & Le Moal, 1988), while in humans, Zakowski, Hall, and Baum (1992) found that lymphocyte proliferative responses to Con A were impaired to a greater extent with unpredictable stressors.

8.02.4.2.2 Control and lack of control

Behavioral control over psychosocial demands can be defined as the possibility of being able to take actions that prevent, terminate, or modify aversive stimulation. Lack of control implies that such actions are not available to the individual because of the prevailing circumstances, or because the actions are not in the behavioral repertoire. In the context of psychophysiological stress responses, perceived control, or the perception that actions can be taken, may be as important as actual control.

There is ample evidence from acute studies in animals that simple forms of behavioral control such as pressing a lever to avoid or escape shock are associated with smaller psychophysiological responses than equivalent uncontrollable stimulation. Uncontrollable conditions elicit greater corticosterone and catecholamine responses, an increased tendency to gastric lesions, a decrease in natural killer cell cytotoxic activity and mitogen-induced lymphocyte proliferation, and greater weight loss (Steptoe & Appels, 1989). In humans, diminution in blood pressure, heart rate, and cortisol responses have been recorded with controllable demands as opposed to equivalent demands that are not controllable.

However, the beneficial effects of control are limited to circumstances in which the behavioral responses are relatively easy to carry out. When contingencies are more complex or when behavioral tasks are difficult, physiological responses may be enhanced. This was neatly demonstrated in the study of gastric lesions in rats administered escapable or yoked inescapable electric shocks (Tsuda, Tanaka, Nishikawa, & Hirai, 1983). When the behavioral response was simple, animals in the uncontrollable condition showed greater ulceration than those in the controllable group. But with complex contingencies, the situation reversed. In humans, it is known that increasing task difficulty is associated with greater blood pressure and sympathetic nervous system activity (Callister, Suwarno, & Seals, 1992). When the tasks that must be carried out to maintain control over aversive stimulation are difficult and require considerable effort, psychophysiological responses may be greater than in uncontrollable conditions (Bongard, Hodapp, Frisch, & Lennartz, 1994). Even the perception that the person has chosen difficult tasks to perform may elicit heightened cardiovascular responses (Sherwood, Royal, & Light, 1993). However, it should be emphasized that in humans, the negative consequences of effortful control have largely been recorded during acute laboratory studies. Whether effects are sufficient to offset the benefits of perceived control in clinical and naturalistic settings remains unknown.

8.02.4.3 Work Characteristics, Control, and Psychophysiological Responses

For people in employment, work occupies a considerable portion of waking life. It is scarcely surprising therefore that the work characteristics associated with ill-health have been extensively studied. Various types of pathology such as hypertension and peptic ulcer are more prevalent in people in certain demanding occupations, although it is often difficult to disentangle psychophysiological processes from variations in lifestyle and health behavior (Cobb & Rose, 1973). Work may also lead to increases in catecholamines and blood pressure, notwithstanding the points made in Section 8.02.2.2.3 about the role of physical activity that remain pertinent to these investigations.

Much research on the health risks of work has been oriented around the demand-control model of job strain put forward by Karasek (1979). This proposes that risks to health are particularly great when work is characterized by high demands or work load coupled with low control or autonomy. The model therefore points to control as being an important element of psychosocial demand. A substantial literature has accumulated suggesting that high demand/low control jobs are associated with increased risk of coronary artery disease and high blood pressure (Schnall, Landsbergis, & Baker, 1994). The data are not all consistent, and other factors such as social support at work and psychological traits also need to be taken into account. A study from the author's laboratory suggested that perceptions of job strain and a disposition to high physiological stress reactivity interact in determining elevated blood pressure at work (Steptoe, Roy, Evans, & Snashall, 1995). A group of healthy young men working as fire-fighters were classified on the basis of job strain and their blood pressure reactivity to standardized mental stress tests. Eighteen months later, blood pressure was monitored during a work shift and an equivalent period of leisure using automated ambulatory apparatus. Neither high job strain nor physiological reactivity in isolation predicted blood pressure at work, but men who displayed both characteristics had elevated blood pressure during the later hours of their work shifts.

8.02.4.4 Concurrent Stressors

Adverse life experiences do not occur in isolation, but are superimposed on a background of hassles or chronic stressors. An issue that is poorly understood is the role of concurrent stressors in determining levels of psychophysiological activity. It might be assumed that if people are already under pressure because of other problems in their lives, then the reactions to fresh threats would be heightened. Alternatively, it is possible that new demands might act as distractors from concurrent difficulties, dampening the intensity of responses.

Cardiovascular responses to mental stress were inversely related to recent life events in a study of adolescents (Boyce & Chesterman, 1990), and beta-adrenergic receptor density was lower among homeless people who had experienced severe life events over the past six months (Dimsdale, Mills, Paterson, Ziegler, & Dillon, 1994). Other studies have shown that high occupational demands are associated with reduced blood pressure responses to acute mental stress tests (Schaubroeck & Ganster, 1993). These reports are consistent with the notion that concurrent stressors dampen physiological responses to fresh challenges. But other work points to the opposite conclusion, with the chronic stress of crowding and high levels of recent hassles being associated with elevated cardiovascular reactions and reduced antibody formation to hepatitis B vaccination (Fleming, Baum, Davidson, Rectanus, & McArdle, 1987; Jabaaij, Grosheide et al., 1993b). It is not certain what factors determine whether concurrent stressors enhance or inhibit psychophysiological responsivity.

8.02.5 RESISTANCE AND VULNERABILITY FACTORS

The other side of the model outlined in Figure 1 concerns the personal characteristics, social resources, and experiential and biological factors that serve to increase or decrease the ability of people to adapt effectively to psychosocial demands. These factors will be outlined in this section, focusing particularly on the evidence for influence over psychophysiological processes. The inclusion of both resistance and vulnerability factors reflects the presumption that some characteristics are protective, while others are potentially damaging.

Perhaps the ideal type of evidence is from studies in which individuals or groups who differ on a putative resistance or vulnerability factor are exposed to identical demands. Any difference in psychophysiological response or health risk would indicate that the factor in question was indeed operative. Within this framework, the resistance or vulnerability factors might exert a "buffering" effect, being operative at high but not low levels of psychosocial demand. Alternatively, they might exert "direct" effects, influencing responses irrespective of the level of demand. Many resistance and vulnerability factors may influence psychophysiological variables through their associations with different types of behavioral and cognitive coping response.

The vulnerability factors discussed in this section are those related to psychophysiological responses in general. The factors that may predispose people to particular types of health problem are considered in Section 8.02.6.

8.02.5.1 Personality and Behavior Patterns

A number of personality factors have been postulated as relevant to stress vulnerability and resistance. Some have been linked to risk for specific disorders (the best known case being Type A coronary-prone behavior), while others are thought to affect general vulnerability to psychosocial demands. There are several processes through which personality might influence psychophysiological responses, and these have been described by Bolger and Zuckerman (1995). First, personality factors or habitual ways of behaving might affect exposure to demands, increasing or decreasing the experience of potential threats by influencing social, vocational, and recreational activities. For instance, a person high in extroversion may have a very different experience of life than one who is introverted. Second, personality factors may affect the appraisal of situations, and whether or not they are perceived as threatening, challenging, irrelevant, exciting, controllable, and so on. Neuroticism and hostility come to mind as exemplars of this process. Third, personality can influence the individual's choice of coping responses. For example, an individual with high internal locus of control may launch into efforts at active problem-solving more readily than someone with external beliefs.

Finally, personality may have an impact on the effectiveness of coping responses. A person with high hostility scores may mobilize social support in an effort to cope, but alienate his or her social contacts with abrasive behavior. Thus far, there are rather few studies that have been able to tease out these different modes of action. Intriguing findings have been described for a number of personality and behavioral patterns including locus of control (Parkes, 1984), sense of coherence and conscientiousness (Friedman et al., 1995). The concept of the hardy personality developed by Kobasa has attracted considerable attention, but has also been criticized on conceptual and empirical grounds (Funk, 1992; Kobasa, Maddi, Puccetti, & Zola, 1985). This review is therefore confined to four areas of research into personality and behavioral patterns that appear especially relevant to psychophysiological processes.

8.02.5.1.1 Disease-prone personality and negative affectivity

The notion that a constellation of personality characteristics may put people at higher risk for all manner of diseases was proposed by Friedman and Booth-Kewley (1987). They described meta-analyses in which high levels of anxiety, neuroticism, depression, anger, and hostility emerged as features linked with a range of pathologies including coronary heart disease, bronchial asthma, peptic ulcer, rheumatoid arthritis, and headache. It was postulated that a person with these characteristics was predisposed to developing physical illness, but that the nature of the illness would depend on biological factors such as genetic risk.

This position can be criticized on three main grounds. First, many of the studies used to support this proposition are cross-sectional or retrospective in nature. It cannot be assumed that the psychological characteristics displayed by someone with manifest disease were present prior to diagnosis, or are present among people with similar medical conditions who have not been diagnosed (the "clinical iceberg"). Many of the illnesses discussed by Friedman and Booth-Kewley lead to functional limitations, pain, and impaired quality of life. It is not surprising that individuals with these problems score higher on measures of anxiety, depression, and anger or irritation than healthy people. In some cases such as hypertension, it is known that diagnosed cases and awareness of the problem are associated with increased psychological distress and neuroendocrine activation independently of blood pressure levels (Rostrup, Mundal, Westheim, & Eide, 1991).

Second, the constellation of factors thought to reflect disease-prone personality overlaps strikingly with negative affectivity. Negative affectivity is a tendency to pervasive dysphoria and low mood, and has been found to influence reports of social support, life events, and symptoms, and to affect health service utilization (Watson & Pennebaker, 1989). It is frequently indexed by measures such as neuroticism, so is difficult to distinguish operationally from a personality trait in cross-sectional studies. Negative affectivity is a mediator of associations between stressors and affective responses, and may consequently contribute to the correlation between disease and personality as well.

The third reason for reservations about the disease-prone personality is that it agglomerates a range of characteristics, each of which may have distinct associations with health risk and psychophysiological responses. To group factors such as depression, anxiety, and hostility together loses the potential for refining the role of psychosocial factors.

8.02.5.1.2 Anger and hostility

The constructs of anger and hostility illustrate this point very well. There has been great interest in the role of these factors in physical health. However, there are many different types of measure and experience that fall within this area, and not all may be relevant to health risk. Thus anger as an emotional state or trait can be separated from hostility, which is generally regarded as a broader construct involving cognitive and behavioral factors. Cynical hostility, for example, is centered on beliefs that others are motivated by selfish concerns, and is a dimension tapped by the Cook and Medley Hostility Scale. These characteristics can in turn be distinguished from the expression of anger (a behavioral response) and from overt aggression.

Space prevents more than a brief summary of the literature relating anger and hostility to physical health and psychophysiological processes. Interest has chiefly focused on cardiovascular disorders. Although the literature is not all consistent, meta-analyses indicate that coronary heart disease and its manifestations such as myocardial infarction are positively associated with overt hostility (Miller, Smith, Turner, Guijarro, & Hallett, 1996). There is limited evidence that hostility also predicts all-cause mortality and by implication health in general. A different pattern is found for high blood pressure, which is associated with low levels of anger expression, particularly among people who are not aware of their blood pressure level (Jorgensen et al., 1996).

These effects are not of course necessarily mediated through psychophysiological pathways, since hostility has also been associated with poor health habits such as lack of exercise and drinking and driving (Leiker & Harley, 1988). However, a number of investigators have shown that hostility is positively related to cardiovascular reactions to mental stress, particularly under conditions of harassment and provocation (Suls & Wan, 1993). Hostility is also positively correlated with cortisol excretion during the day, and with blood platelet activation, an early marker of thrombotic processes (Markovitz, Matthews, Kriss, & Smitherman, 1996; Pope & Smith, 1991). It is reasonable to assume, therefore, that anger and hostility are vulnerability factors in psychophysiological processes and disease.

8.02.5.1.3 Depression, pessimism, and optimism

People with serious illnesses frequently experience depression. Clinically, it is commonly observed that people suffering from painful conditions such as chronic back pain or rheumatoid arthritis are at least mildly depressed for a great deal of the time. Depression is also associated with nonpainful but restrictive conditions such as bronchial asthma and various cardiomyopathies. Herbert and Cohen (1993a) have reported a meta-analysis of studies linking depression with immune function, and conclude that depressed people generally exhibit lower mitogen-induced lymphocyte proliferation, natural killer cell activity, and smaller numbers of circulating helper T-cells than do comparison groups. Studies of patients following myocardial infarction indicate that depressed individuals are at high risk for future complications and early mortality independent of cardiological variables (Frasure-Smith, Lespérance, & Talajic, 1993), and pessimism as a trait has been shown to predict future mortality in at least one longitudinal study (Peterson, Seligman, & Vaillant, 1988).

These associations raise the question of whether psychiatric disorder of an affective nature increases risk of physical illness. There has been speculation on this topic for many decades, and recent years have seen the publication of several convincing systematic studies. For example, Huppert and Whittington (1995) used data from the Health and Lifestyle Survey to show that elevated scores on the General Health Questionnaire predicted mortality over a seven-year follow-up period independently of age, diagnosed physical disorder at the start, social class, and smoking. A more rigorous interview-based measure of depression was used in a study of first-degree relatives of depressed patients and controls (Moldin, et al., 1993). People classified as having experienced clinical depression at some time in their lives were at increased risk for headache, skin infection, respiratory illness, diabetes, and low blood pressure, after controlling for age, sex, and alcohol use.

In a larger study of people with diagnosed anxiety and depressive disorder in Sweden, risk of death from coronary heart disease was increased among men but not women (Allgulander, 1994). Unfortunately, interpretation of studies of this kind is problematic, since utilization of healthcare services may be different among people who are distressed or depressed, leading in some circumstances to an increased likelihood of diagnosis of conditions that might otherwise go unnoticed. Psychological disturbance may be a product of early subclinical physical illness, and not predate health problems. Patterns of health behavior such as exercise and food choice also vary with mental health, and may contribute to increased risk of disease.

The other side of the coin is optimism, which has been defined by Scheler and Carver (1992) as a set of global expectancies in which people generally experience good as opposed to bad outcomes in life. Optimism as measured, for example, by the Life Orientation Test, has been associated with favorable adaptation to cardiac surgery, lower distress among women with breast cancer and HIV positive men, and fewer psychological problems following unsuccessful medical interventions (e.g., Carver et al., 1993; Taylor et al., 1992). An optimistic disposition appears to be associated with the use of adaptive coping strategies in the face of threat, such as planning and positive reinterpretation, with low levels of cognitive avoidance. In this respect, optimism would appear to fulfill the criteria for a general resistance factor.

8.02.5.1.4 Emotional inhibition

One of the key concepts underlying the psychodynamic tradition of psychosomatics is that physical disorders might emerge through repression of neurotic conflicts. From this, the idea developed that people suffering from various physical complaints might be emotionally repressed or inhibited. A more modern development of this theme is the construct of alexithymia, a difficulty in accessing feelings and bodily sensations and in relating to emotional experiences (Sifneos, 1973). The hypothesis that the inhibition of emotion is potentially damaging can be supported without accepting all the intellectual penumbra of psychodynamic thought. Openly discussing emotionally-charged events may facilitate emotional processing, may help give meaning to distressing experiences, and may allow for social comparison and input of information and advice from others.

Several different methods of operationalizing the inhibition of emotion have been developed in the research literature (Pennebaker, 1995). The method developed by Weinberger and coworkers involves measuring anxiety and social desirability so as to identify people who report low anxiety together with high social desirability scores (Weinberger, Schwartz, & Davidson, 1979). The rationale is that people with a "repressive coping style" fail to acknowledge emotional distress (hence having low anxiety) but also desire social approval. Weinberger et al., found that students with repressive coping styles showed greater autonomic responses to a task than did genuinely low anxious individuals. This finding has been replicated in other studies, although not entirely consistently. For example, antibody titers to Epstein-Barr virus antigen were found to be raised among students with high anxiety plus high social desirability scores by Esterling, Antoni, Kumar, and Schneiderman (1993), and not among repressors.

Simpler measures of emotional expression are also negatively associated with physiological responses to challenge, a pattern that has been identified even in children of three and four years old (Cole, Zahn-Waxler, Fox, Usher, & Welsh, 1996). Anger inhibition may be particularly relevant to the development of hypertension, and has been shown to predict heightened cardiovascular reactions to mental stress tests in people with a family history of high blood pressure (Vögele & Steptoe, 1993). Another approach has been to assess the psychophysiological and health consequences of disclosure of emotionally distressing information. In one study of this kind, students were randomized into writing about very traumatic episodes in their pasts, or essays on neutral topics (Pennebaker, Kiecolt-Glaser, & Glaser, 1988). Cellular immune function appeared to be favorably affected by confronting traumas, with some evidence for a reduction in episodes of minor illness over subsequent weeks. There are other intriguing preliminary findings suggesting that lack of disclosure may have important health consequences. Cole, Kemeny, Taylor, and Visscher (1996) followed a cohort of HIV-negative homosexual men with regular health checks over a five-year period. They found that men who concealed their homosexual identity experienced a high incidence of infectious illnesses such as sinusitis and bronchitis, independent of common health behaviors, negative affect, age, and education. Risk for these types of disorder might be mediated through suppression of immune defense.

8.02.5.2 Social Isolation and Social Support

The influence of social isolation, social integration, and social support on health risk is now well established. A substantial literature on social support and physiological processes has also emerged, and this has been amply reviewed by Uchino, Cacioppo, and Kiecolt-Glaser (1996). Social network refers to the person's web of social relationships and ties, and is typically assessed by measures such as marital status, number of people in the household, number of close friends and relatives, church attendance, and participation in group activities. Composite measures of social networks have been shown to predict mortality in the Alameda County study, Tecumseh, Scandinavian cohorts, and elsewhere (House, Landis, & Umberson, 1988). It is apparent that network measures may include a number of quantitatively different types of social contact, and it is unlikely that all are equally important. Church attendance may bring with it particular attitudes to problems and styles of coping that have little to do with group involvement *per se*, while marriage is much more than having someone close in the social network.

The function of social networks is thought to be to provide social resources and support, and several types of support can be distinguished. They include material or tangible support, such as the provision of assistance with tasks; emotional support, or the availability of someone with whom to share personal problems and from whom to receive comfort and concern; informational support which involves the provision of advice and suggestions about plans of action; and belongingness, or a sense that one is part of a group with similar interests and concerns. Again, not all these elements are equally important across situations. The mother with young children who has recently been left by an errant spouse might benefit particularly from practical support in terms of help with child-care and shopping, and informational support such as legal advice. By contrast, a person worried about being dismissed from their job might especially value emotional support and a feeling that they still belong to their social group.

Social support and strong social networks are not uniformly distributed through the population. Perceived social support from spouses, friends, relatives, and co-workers is greater among people of higher socioeconomic status (Turner & Marino, 1994). In the Whitehall II epidemiological study of British civil servants, satisfaction with practical support was greater in men of higher than lower employment grade (Marmot et al., 1991). Women tend to report higher social support than men from most sources, although ratings of relationships with spouses are frequently less positive. In their studies of the social origins of depression, Brown and Harris (1978) showed that poor spouse/partner support was more frequent among women of low social class with young children at home than it was in other groups. It should of course be recognized that in general, social ties may have many negative consequences, and some of the most potent sources of chronic distress arise from relationships in the family (Burg & Seeman, 1994).

The literature is rather consistent in showing that social networks and social support are associated with favorable cardiovascular function (Uchino et al., 1996). For example, a study of elderly Swedish men found that blood pressure was negatively correlated with degree of anchorage in formal and informal social groups after controlling for social class, marital status, alcohol, smoking, body weight, and physical activity (Hanson, Isacsson, Janzon, Lindell, & Rastam, 1988). A frequently cited follow-up study of Italian nuns from a secluded and close-knit religious order showed that their rise in blood pressure with age was less marked than in the general population, although being a nun has more to it than social integration (Timio et al., 1988). Dressler (1991a, 1991b) has performed an interesting set of studies showing an interaction between chronic life stress and support among black men in southern USA. A subgroup were characterized by "lifestyle incongruity," in that their lifestyle in terms of material goods and status displays exceeded their occupational position. Dressler argued that efforts to maintain such an ostentatious lifestyle in the absence of sufficient funds constituted a source of chronic life stress. He found that blood pressures were indeed higher among men with lifestyle incongruity, but only when they had low levels of social support.

Physiological variables have also been assessed in a series of acute laboratory studies of responses to standardized tasks carried out in the presence of supportive friends. Effects have been somewhat mixed, although there is some evidence that cardiovascular reactions are reduced in the presence of friends (Kamarck, Annunziato, & Amateau, 1995). Evidence concerning neuroendocrine function and social support is less consistent, with several studies showing little effect. Immune data are more compelling in showing that the down-regulation of function elicited by chronic stressors may be ameliorated by support (Baron, Cutrona, Hicklin, Russell, & Lubaroff, 1990; Kiecolt-Glaser et al., 1991).

The results of studies that have attempted to bolster social support but have been rather disappointing to date. For example, Arnetz et al. (1987) randomized unemployed women to control or intervention groups that provided information and engaged in alternative activities. No effect on immune or endocrine function was observed. A very substantial study providing social support to deprived young pregnant women in Latin America failed to find any protective effects on birth outcomes (Villar et. al., 1992). Helgeson and Cohen (1996) have argued that although emotional support groups are desired by patients with cancer, there is little evidence that they are of benefit in terms of psychological well-being or other aspects of adjustment. It is not yet clear that support provided externally will necessarily have comparable effects to support that emerges more naturally from the individual's evolving social world.

8.02.5.3 Prior Experience of Stressful Events

There are distinct schools of thought about whether prior adverse experience helps to strengthen the individual's resistance to later threats, or whether it increases vulnerability. On one side, there is evidence that earlier stressful experiences increase vulnerability to later adverse events. Thus experiences such as disruption of family relationships early in life increase risk for later emotional and behavioral problems (Wadsworth, Maclean, Kuh, & Rogers, 1990). In work on life events and depression, Brown and Harris (1978) found that women who had as children experienced the loss of their own mothers through death, ran a high risk of developing depression in adult life in the face of negative life events. Adverse effects are by no means inevitable, since the impact of early stressful experiences is ameliorated by temperamental factors and coping style (Garbarino, Kostelny, & Dubrow, 1991). Cortisol responses to stressful situations are smaller among infants with more secure attachments (Nachmias, Gunnar, Mangelsdorf, Parritz, & Buss, 1996). Animal studies have shown that early social isolation leads to increased HPA output, reduced lymphocyte activation by mitogens, and an elevation of central norepinephrine turnover (Suomi, 1991). The enhancement of psychophysiological responses arising from disruption of early development may increase vulnerability to later disease. Prenatal factors may also be significant. For example, the exposure of pregnant rats to behavioral stress has been shown to influence the immune function of their offspring (Klein & Rager, 1995), and influences on later neuroendocrine function have been described as well.

On the other side, prior experience may mitigate the effects of later aversive stimulation. Thus early handling in laboratory animals produces more rapid adaptation to avoidance conditioning, smaller adrenal responses to chronic stress, and enhanced antibody responses to experimental infections (Solomon, Levine, & Kraft, 1968). Dienstbier (1989) has argued that prior stressful experiences lead to "physiological toughening," and more effective adaptation to later challenges. A direct test of this model was described by Norris and Murrell (1988) who examined the impact of a natural disaster (serious flooding) on anxiety symptoms. It was found that older adults who had experienced similar disasters earlier in their lives showed better adaptation than did those who lacked prior experience.

These discrepant findings can probably be reconciled through consideration of the specific adverse events experienced in the past and current demands on the individual. Previous exposure to a particular threat (such as a natural disaster) might indeed improve adaptation, since the sequence of events will be more familiar and less unpredictable than before, and strategies for coping may already have evolved. Other types of threatening experience may disrupt developmental trajectories, impair self-esteem and reduce the individual's social resources, leading to heightened responsivity to later events.

8.02.6 PATHWAYS TO DISEASE—PREDISPOSING FACTORS

Thus far, factors have been described that influence the magnitude and duration of psychophysiological responses in a general fashion. However, as well as establishing credible biological pathways, models of psychophysiological processes in disease must account for individual differences. Why is it that groups who suffer from "stress-related" pathology often do not appear to have more stressful lives than those who remain healthy? How is it that two people may have similar life experiences, yet one becomes ill while the other remains healthy? Why, in the face of chronic stressors, do some people succumb to infectious illness while others show increases in blood pressure and the development of hypertension?

Faced with these issues, it has been traditional in psychophysiological models to invoke "biological vulnerability" as an explanation. People will succumb to the pathologies to which they are most vulnerable, depending on their *locus minoris resistentiae*. Thus an individual prone to coronary artery disease will show accelerated atherogenesis with adverse life experience, while another susceptible through biological factors to muscular contraction will be at increased risk for headache or muscular skeletal pain. Still another person may be physically robust, but prone to anxiety disorders.

This model may account superficially for individual differences, but unless it is properly characterized it has little explanatory power. The problem is that unless biological vulnerability is defined *a priori* with clear measurable markers, then there is the danger that vulnerabilities are identified after the event, and that reasoning becomes circular. Fortunately, there has been progress over recent years in delineating predisposing factors. Some of these help account for why one person becomes ill and another does not, while others explain why individuals succumb to particular forms of ill-health.

8.02.6.1 Nutrition and Exercise

Nutritional deficits in early life affect later functioning of the immune system even after proper nutrition has been reinstated. A persistent alteration in phagocytosis and an accelerated decline in cell-mediated immunity have been described (Dutz, Kohout, Rossipal, & Vessal, 1976). Serum immunoglobulins appear relatively unaffected by malnutrition through protein and energy deficits, but secretory IgA responses are reduced and may enhance the risk of succumbing to viral infection. Among elderly people, subtle and often undetected micronutrient deficiencies lead to reductions in lymphocyte function and natural killer cell activity (Chandra, 1989). Other aspects of nutrition such as the consumption of saturated fats, salt, and fruit and vegetables are of clear importance as risk factors in disease. Thus nutritional profiles may have a profound influence on the extent to which psychophysiological responses will have clinical consequences in terms of disease risk.

Regular physical activity and physical fitness may potentially be protective. A substantial number of studies have assessed associations between fitness or regular activity and psychophysiological stress responsivity. Although by no means consistent, investigators have shown that blood pressure and catecholamine stress responses are smaller among fit or more active individuals (Seraganian, 1993). However, cross-sectional differences may be due to selection factors or other aspects of the lifestyle of sedentary and active people. Efforts have therefore been made to assess psychophysiological responsivity to mental stress before and after exercise training. Again the results have been very mixed, with reductions in stress responsivity with physical training emerging in some studies but not others (De Geus, Van Doornen, & Orlebeke, 1993; Rogers, Probst, Gruber, Berger, & Boone, 1996). The explanation of these discrepancies is not clear, so the extent to which psychophysiological mechanisms are moderated by regular physical activity and fitness is uncertain. Field studies certainly suggest that fitness is protective. For example, Brown (1991) found that in a sample of students, adverse life events over the past year were associated with psychological distress and self-reported illness level. However, both illness rates and healthcare utilization were lower in those who had experienced life events but were also physically fit, as opposed to the unfit. Results of this kind indicate that further investigation of the influence of regular activity and fitness on psychophysiological processes and disease is warranted.

8.02.6.2 Advanced Age

Many of the medical disorders investigated in health psychology are more common in older than younger adults, including coronary artery disease, rheumatoid arthritis, high blood pressure, and various cancers. The issue of aging is becoming increasingly important to an understanding of psychosocial contributions to disease risk (Schaie, Blazer, & House, 1992). The nature of psychosocial demands changes with age. Heckhausen and Schulz (1995) have argued that later years are characterized by increasing exposure to events over which the individual has little control, including bereavement, health problems, the reduced economic power associated with retirement, and diminution in social networks due to death and changing roles. People who appear to age more successfully may adapt by lowering their evaluation of the importance of areas of life over which they lose control (Brandtstadter & Rothermund, 1994). However, flexibility and adaptation may be reduced at both the psychological and biological levels, contributing to increased disease risk.

The profile of biological stress responses changes with age. Older age is characterized by higher levels of circulating catecholamines but reduced responsiveness to sympathetic stimulation (Lakatta, 1993). Thus, heart rate responses to challenging tasks are reduced, while a decrease in the dilation of blood vessels supplying working muscle results in increased blood pressure responsivity. Animal studies suggest that the magnitude of acute responses in the HPA system is unchanged with age, but that cortisol elevations are more prolonged in older animals, suggesting disruption to regulatory feedback mechanisms (Seeman & Robbins, 1994). A limited amount of evidence suggests similar processes are at work in humans, although increased cortisol responses to challenging tasks with age have also been described (Gotthardt et al., 1995). The magnitude of ACTH and cortisol responses to a simulated real-life challenge (a driving test) were shown to be inversely associated with self-esteem in an elderly cohort, suggesting that maintenance of psychological well-being in old age may help reduce adverse psychophysiological response patterns (Seeman et al., 1995).

8.02.6.3 Gender

Many of the medical disorders discussed in other chapters of this volume have different rates of occurrence in men and women. There are also consistently higher reports of psychological distress on measures such as anxiety and depression among women than men, and these differences do not appear to be accounted for by disclosure or response biases (Mirowksy & Ross, 1995). The question therefore arises of whether psychophysiological processes contribute to differences in disease and distress between the sexes.

Faced with identical challenges, young men tend to show more sustained cortisol reactions than young women (Kirschbaum, Wust, & Hellhammer, 1992), together with higher epinephrine responses (Frankenhaeuser, Dunne, & Lundberg, 1976). Men also manifest greater blood pressure responses than women, and this pattern can be seen even in childhood (Allen, Stoney, Owens, & Matthews, 1993; Murphy, Stoney, Alpert, & Walker, 1995). By contrast, heart rate responses are typically greater in women. Studies comparing women at different phases of the menstrual cycle have not reached a consensus about effects on stress responsivity (Stoney, Owens, Matthews, Davis, & Caggiula, 1990). However, it is probable that the low responsivity of young women is mediated at least in part by circulating estrogen, and after menopause the blood pressure stress responsivity and ambulatory blood pressure of women approach that of men (Owens, Stoney, & Matthews, 1993).

8.02.6.4 Ethnicity

There are pronounced ethnic differences in the prevalence of a variety of diseases (Murray & Lopez, 1996). The contribution of psychosocial factors has been investigated extensively in relation to cardiovascular disorders and diabetes. Noninsulin dependent diabetes is particularly common among Micronesians, native Americans, Hispanic Americans, and in migrants from South Asia to the UK. It is accompanied by a high prevalence of insulin resistance, low concentrations of HDL-cholesterol and high levels of abdominal obesity. Although there have as yet been few studies of psychophysiological processes, it is plausible that these ethnic differences may relate to the stress-related pathways described in Section 8.02.3.3.1.

Hypertension is more prevalent among black than white people in urbanized societies, and research indicates that factors such as repressed hostility, socioeconomic status, and social integration may all contribute (Anderson, Myers, Pickering, & Jackson, 1989). There have been many studies of differences in the responses of black and white volunteers to laboratory mental stress tests. The data suggest that black people show greater increases in peripheral vascular resistance than whites, particularly when confronted by passive aversive tasks such as the cold pressor test (Saab et al., 1992). White people by contrast show increases in heart rate and myocardial contractility, indicative of greater sensitivity to beta-adrenergic sympathetic nervous stimulation. These findings indicate that when faced with the same psychosocial demands, ethnic differences in responsivity may contribute to differential disease risks.

8.02.6.5 Socioeconomic Status and Education

The major sources of morbidity and mortality in developed countries show strong socioeconomic gradients, with higher prevalence among people of lower socioeconomic status (Blane, Brunner, & Wilkinson, 1996; Kaplan & Keil, 1993). Socioeconomic status may be assessed by a variety of indices including occupation, income, type of housing and education. Differences are present not only in manifest disease but in subclinical pathology, and are present throughout the life course. As noted earlier, psychosocial demands are greater among people of the lower socioeconomic status, while perceived social support is poorer (Turner & Marino, 1994; Turner et al., 1995). Low socioeconomic status is associated with chronic stressors such as excessive work demands, inadequate rewards, economic hardship and uncertainty, increased exposure to crime and noise in everyday life, frustration in expectations associated with social roles, and greater barriers to achieving life goals (Aneshensel, 1992). A combination of biological factors, health behaviors and psychosocial influences such as poorer support appears to account for much of the socioeconomic gradient in mortality and cardiovascular disease risk factors (Lynch, Kaplan, Cohen, Tuomilehto, & Salonen, 1996). However, some social variations require further explanation. For example, Nordentoft et al. (1996) found that life stress and education were independent predictors of premature delivery in a large sample of pregnant women. There is limited evidence that psychophysiological responsivity is associated with socioeconomic markers such as education and work status, but further study of these issues is essential (Carroll, Smith, Sheffield, Shipley, & Marmot, 1997; Owens et al., 1993)

8.02.6.6 Genetic Factors

Genetic factors are probably responsible for many of the differences between people who are at risk of psychophysiological disease. They may operate in several different ways. First, twin adoption studies indicate that a substantial proportion of the variance in the experience of adult life events and in social support is accounted for by genetic influences (Bergeman, Plomin, Pedersen, McClearn, & Nesselroade, 1990; Plomin, Lichtenstein, Pedersen, McClearn, & Nesselroade, 1990). Psychological factors such as optimism, hostility, and Type A behavior are also heritable. Consequently, the likelihood that a person will manifest intense psychophysiological responses is determined not only by his or her experience, but by genetic makeup. Second, the magnitude of psychophysiological stress responses to standardized tasks in the laboratory shows significant heritability (Hewitt & Turner, 1995). Third, genetics contribute to other predisposing factors like obesity, and to activities such as alcohol consumption, thereby increasing risk of disease through behavioral pathways. Overall, there is a sizeable genetic contribution to risk of illnesses such as coronary heart disease, hypertension, rheumatoid arthritis, diabetes, and certain types of cancer, so it is likely that psychophysiological processes will have a greater impact on an individual who is genetically predisposed to a particular condition.

8.02.6.7 Individual Patterns of Psychophysiological Responses

The final predisposing factor that needs to be considered is each person's tendency to show a particular pattern of physiological responsivity to psychosocial demands. As noted earlier, physiological responsivity is a relatively stable characteristic for each individual, and it is one that emerges early in life in animals and humans (Brown-Borg, Klemcke, & Blecha, 1993). Lewis, Thomas, and Worobey (1990) showed that infants who were emotionally reactive to a painful stimulus (inoculation) at two months old were more prone to infection over the subsequent year, probably because of a tendency towards stress-induced immunosuppression. In adult life, many people show individual-response specificity, which is the tendency to respond maximally and consistently in a particularly physiological pathway. It has typically been evaluated by measuring a range of autonomic responses to a series of different tasks. A person manifests individual-response specificity if they are maximally reactive in one measure across situations. Individuals differ in the measures in which they are most responsive, so some may be blood pressure responders, while others are most sensitive in terms of the electrodermal system or cortisol (Steptoe, 1984a). Multivariate analyses suggest that about one-third of people show consistent individual-response specificity. It is appealing to speculate that these people may be prone to developing disorders associated with the psychophysiological parameter in which they are most responsive.

8.02.7 PSYCHOPHYSIOLOGICAL MODELS OF DISEASE

The previous sections have described the physiological responses that are activated when psychosocial demands outweigh the resources available to the individual, and have outlined the factors that increase vulnerability and predisposition to disease. There remain the actual mechanisms through which disease is initiated and exacerbated. However, before these psychophysiological processes are described, it must be emphasized that they are not the only pathways through which psychosocial factors affect disease and physical health. Psychosocial factors may influence disease not only through psychophysiological processes, but also through behavioral, affective, and social-cognitive pathways. These latter will not be discussed here, since they have been extensively described previously (Cohen & Rodriguez, 1995; Steptoe, 1991).

Cognitive processes such as the identification of symptoms or decisions to seek help from health professionals, and behaviors such as alcohol consumption, adherence to medical advice, and risky sexual practices, may all be affected by the psychosocial factors described in this chapter. They may in turn impact on diagnosis and disease. It is frequently difficult fully to apportion accountability to psychophysiological and behavioral mediators. For example, both poor social support and depressed mood have been shown to predict early mortality following myocardial infarction (Berkman, Leo-Summers, & Horwitz, 1992; Frasure-Smith et al., 1993). The known effects of these factors on physiological systems relevant to coronary disease may be responsible, but so might behavioral factors. For instance, there may be differences in adherence to medication, or small variations in the speed with which help is sought following the onset of symptoms, that might occur between the socially isolated or depressed patients and others. These might determine whether or not the patient survives a fresh cardiac event irrespective of psychophysiological disturbances.

Notwithstanding these considerations, there is ample evidence for psychophysiological involvement in disease pathology. In fact, several types of mechanism have now been identified, and these are summarized below. It is important to recognize the differences between them, since the role played by psychophysiological factors in disease pathology is somewhat varied. Moreover, the types of research design needed to evaluate the various mechanisms are also quite distinct.

(i) Psychosocially induced physiological reactivity as a causal factor.

(ii) Psychosocially induced physiological reactivity as an inhibitor of host resistance and host defenses.

(iii) Psychosocially induced physiological reactivity as a disruptive factor on existing pathology.

(iv) Psychosocially induced physiological reactivity as a trigger of acute clinical events.

8.02.7.1 Physiological Reactivity as a Causal Factor

The first possibility is that psychophysiological reactivity is directly responsible for disease. Particular individuals show heightened reactivity in specific physiological parameters such as blood pressure, gastrointestinal motility, or muscle tension in the head and neck. Through regular or sustained exposure to psychosocial demands that over-tax resources, these physiological responses may be repeatedly elicited and in due course may gravitate from acute reactions to sustained pathology. Of course, physiological reactivity does not act alone, but against the substrate of predispositions outlined in Section 8.02.6.

The type of study needed to evaluate this mechanism is a longitudinal investigation in which psychobiological predispositions and reactivity patterns are measured at the outset, and exposure to the stressors that trigger appropriate physiological responses is tracked. It would then be predicted that people with heightened reactivity in a particular physiological response system, and who also experience psychosocial demands that elicit these reactions over a sustained period, will be at increased risk for developing the disease.

This type of study is difficult to carry out in humans, although it forms the basis of some animal models of pathology (Manuck et al., 1989). More readily available in humans is evidence that people with a variety of disorders tend to react to cognitive and emotional challenges with heightened responses in the physiological systems relevant to their disorder. Thus people with hypertension typically show larger blood pressure responses to mental stress tests than do those with normal blood pressure. Asthmatics react to challenges with greater increases in bronchoconstriction, people with irritable bowel or irritable colon show greater disturbances in motility of the gut, patients suffering from tension headache or chronic low back pain show heightened reactivity in muscles in appropriate tissues, and so on (Flor & Turk, 1989; Kumar & Wingate, 1985; Steptoe, 1997).

These studies are consistent with psychophysiological reactivity being involved in disease etiology, but do not provide strong evidence. The reason is that disturbances in reactivity to behavioral challenges may be secondary to the pathological changes characteristic of the disease itself. Thus bronchial hyper-reactivity is a general process in asthma that triggers reactions to exercise, cold, housedust mite and other factors apart from emotional challenges. In other conditions as well, exaggerated reactivity may be an effect rather than cause of the medical condition.

A way of circumventing this cause–effect problem is to identify people who are at enhanced risk for a particular disorder prior to the onset of pathology. If aberrant reactivity is present among these individuals, then the psychophysiological disturbance clearly precedes the development of the disorder. Unfortunately, markers of enhanced risk are rare for many medical disorders, making this approach difficult to implement. The field in which this strategy has been developed most vigorously is hypertension, since there is a moderate but consistent hereditary contribution to blood pressure level. This means that young people whose parents have hypertension are at increased risk themselves. The blood pressure and other cardiovascular responses to mental stress of normotensive people with and without a family history of hypertension have therefore been examined. It is frustrating that despite a large number of studies, findings have been inconclusive, with heightened reactivity being recorded from high risk people in some studies but not in others (Steptoe, 1997). The explanation for the variability in results is not clear, although the nature of psychological demands and the co-existence of vulnerability factors such as emotional inhibition may be relevant. Perhaps more importantly, even this research strategy does not completely overcome the cause–effect problem in interpretation. A variety of pathophysiological processes are disturbed in young normotensives with a family history of hypertension, including sodium transport, vascular distensibility and left ventricular structure. Consequently, the possibility cannot be ruled out that heightened psychophysiological reactivity is secondary in these groups, just as it may be in established hypertensives.

Longitudinal studies have also been conducted, testing the hypothesis that normotensive people with heightened reactivity will be at increased risk of developing hypertension in the future. Here the evidence is rather more compelling, with an increasing number of studies showing that high reactivity does predict future hypertension (Light, Dolan, Davis, & Sherwood, 1992; Treiber et al., 1996). Blood pressure reactivity to mental stress is also associated with the pathological consequences of hypertension such as hypertrophy of the left cardiac ventricle (Hinderliter, Light, Girdler, Willis, & Sherwood, 1996). Yet it is also the case that some longitudinal studies have failed to show any prediction of future cardiovascular problems (Carroll et al., 1995). Two issues are relevant. The first is that heightened psychophysiological reactivity may only be significant in the presence of co-factors such as obesity or salt sensitivity. Second, it is necessary in longitudinal studies to include measures of the life experience of the cohort over the follow-up period. Unless this is done, one important element of the psychophysiological causal hypothesis is missing.

8.02.7.2 Physiological Reactivity as an Inhibitor of Host Resistance and Defense

The second mechanism through which psychosocially-induced physiological reactions may influence disease is by altering physical vulnerability in ways that render the person more susceptible to invasive organisms. The mechanism is presented diagrammatically in Figure 2, which shows four possible scenarios linking exposure to pathogens and bodily defenses. The mechanism is illustrated through the case of infectious agents and immune defense, although it is equally applicable to other pathogens and defense systems. Panel A represents the situation when the organism mounts an adequate immune response that tackles the infection effectively and prevents disease. In panel B, exposure takes place in the presence of a stress-induced impairment of immune response, allowing the infection to be acquired and illness to develop. This is the situation that pertains when physiological reactivity alters host resistance.

This psychophysiological mechanism is different from the process described in Section 8.02.7.1 in two important ways. First, the psychophysiological response is not presumed to cause illness, but rather to influence vulnerability. Second, it is not necessary to postulate that people are exposed to intense, repeated stressors in order for disease risk to be enhanced. Rather, heightened vulnerability for a single short time period may be sufficient, if it is contiguous with a critical state of invasion.

The ideal type of study of this mechanism is a longitudinal investigation in which psychosocial factors are monitored together with measures of bodily defenses, exposure to infectious agents, and illness. Such studies are difficult to carry out at a technical level and also require high levels of cooperation from participants. Interpretation of results is also complicated by the two additional scenarios shown in panels C and D of Figure 2. Panel C describes a situation when the infectious agent to which the individual is exposed is particularly virulent. Under these circumstances, illness may occur even though immune responses are intact. Another possibility is that bodily defenses are disrupted, but since there is no simultaneous exposure to a pathogen, no illness results (Panel D). Taking account of all these alternatives probably goes beyond the scope of any study in humans that has been conducted thus far.

Some of the most convincing evidence for psychosocially induced impairment of host defenses comes from studies in which infectious agents were administered experimentally. Cohen, Tyrrell, and Smith (1991) demonstrated that the acquisition of infection following administration of known doses of common cold viruses was related to a composite measure of psychological stress; subsequent analyses suggested that infection rate was predicted by negative affect and subjectively-rated stress, while illness was predicted by negative life events. These findings have been replicated in broad outline, albeit without assessment of the immune parameters that might have mediated clinical effects (Cohen, Boyle et al., 1995, Stone et al., 1992).

Psychosocial factors are associated with infectious illness in naturalistic settings as well (Cohen & Williamson, 1991). For example, the herpes virus responsible for cold sores and genital lesions may remain latent in the body after the first exposure, but can be reactivated to produce new episodes of illness. Several retrospective studies have shown that recurrence is more likely during episodes of psychosocial stress. One prospective investigation with daily ratings of mood indicated that the ten days prior to a recurrence of genital herpes were characterized by elevations in ratings of nervousness (Dalkvist, Wuahlin, Bartsch, & Forsbeck, 1995). Interestingly, Turner-Cobb and Steptoe (1996) found in another longitudinal experiment that symptoms of upper respiratory tract illness were preceded by high levels of perceived stress for two to three weeks. In the study of immune function in care-givers for dementing patients alluded to in Section 8.02.3.4.3, Kiecolt-Glaser et al., 1991) reported that care-givers experienced longer episodes of illness over a 12-month follow-up than did controls. This suggests that the immune suppression recorded among care-givers may have led to prolongation of illness.

Evidence for psychosocially-induced impairment of host resistance having a clinical role is therefore becoming increasingly convincing. In studies of pathophysiology, there is a new understanding that infectious processes are not confined to traditional self-limiting disorders, but may contribute to conditions such as peptic ulcer, gastric cancer, dementia, vasculitis, and insulin-dependent diabetes (Lorber, 1996; O'Connor, Buckley, & O'Morain, 1996). As these mechanisms are uncovered, the role of psychophysiological processes in reducing resistance may assume even greater importance.

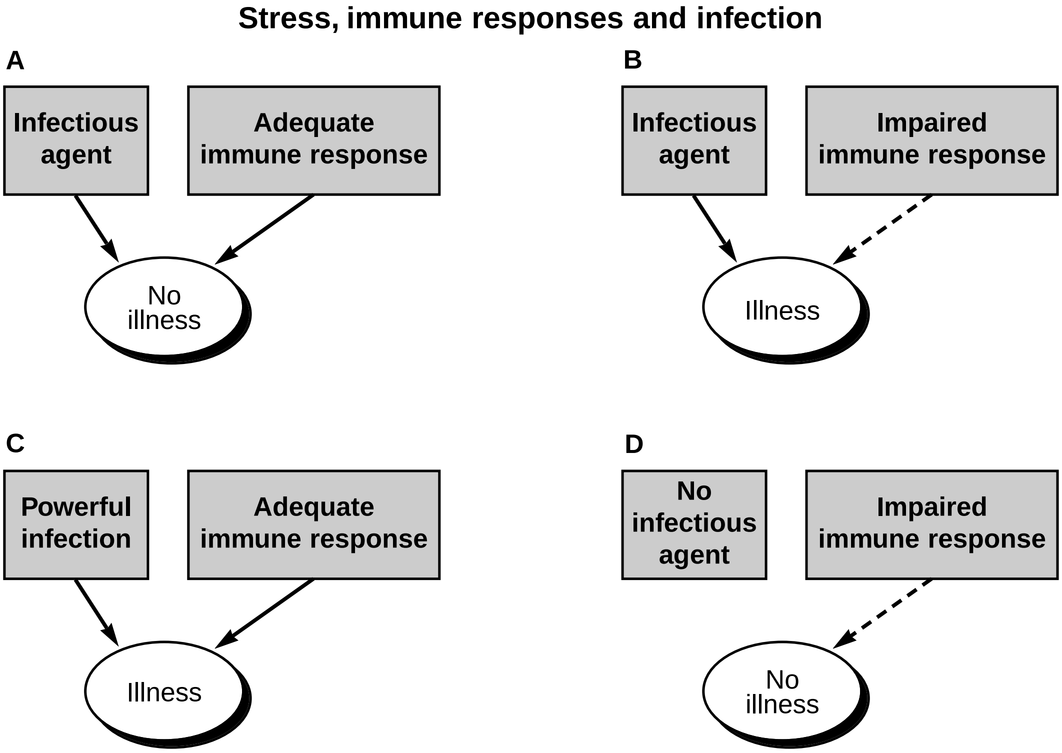

Figure 2 Schematic outline of associations between immune defenses and infection.

8.02.7.3 Physiological Reactivity as a Disruptive Factor

If psychosocially induced physiological reactions do not act as causal factors, how do they contribute to the disease process? An important possibility is that psychophysiological reactions act on existing pathology to disrupt regulatory mechanisms. The result may be exacerbation of the patient's clinical state, possibly leading to symptomatic episodes, or to more rapid degeneration in progressive conditions than would otherwise be the case.

The role of physiological reactivity as a disruptive factor differs from its involvement in the processes outlined in Section 8.02.7.1 and 8.02.7.2 in several ways. First, this mechanism is operative in people already suffering from disease, although the pathology may not necessarily have reached the stage of clinical diagnosis. It is particularly likely to be active in chronic or episodic conditions such as pain syndromes, bronchial asthma, and diabetes. Second, the influence of psychophysiological responses may be manifest at the level of day-to-day clinical status, and consequently relatively minor daily hassles may be influential. Third, the physiological reactions may either be directly involved in pathology (such as disturbances of bronchoconstriction in asthma) or may disrupt host resistance, as in the case of autoimmune conditions. There may be some problems such as insulin-dependent diabetes where effects could be mediated either directly through disturbances of insulin metabolism or indirectly through failures in resistance to enteroviruses.

There is a substantial amount of data showing that symptoms and other indicators of medical conditions covary with psychosocial factors, and the mechanism outlined here may be responsible for these associations. By way of illustration, factors such as negative mood and daily stressors have been shown to covary with disability and symptoms in rheumatoid arthritis (Affleck, Tennen, Urrows, & Higgins, 1994), irritable bowel syndrome (Suls, Wan, & Blanchard, 1994), the chronic inflammatory condition systemic lupus erythematosus (Adams, Dammers, Saia, Brantley, & Gaydos, 1994), headache, Crohn's disease (Brantley & Jones, 1993), insulin-dependent diabetes (Goldston, Kovacs, Obrosky, & Iyengar, 1995), and phantom limb pain (Arena, Sherman, Bruno, & Smith, 1990). However, the problem arises of whether these associations reflect genuine influences on pathology, or effects on symptom reporting and complaints. In order to show that psychophysiological processes are operative at the biological level, it is necessary to measure covariation of psychosocial factors with physiological regulatory parameters as well as symptoms. Here findings are more sparse, but nevertheless suggestive. Thus several studies of insulin-dependent diabetes have shown that stressors influence blood glucose levels and metabolic control independently of variations in diet, exercise, or compliance with medical advice (Goldston et al., 1995; Halford, Cuddihy, & Mortimer, 1990). In patients with rheumatoid arthritis, the experience of minor life events was associated with fluctuations in T and B lymphocyte counts (Zautra et al., 1989). As methods of physiological monitoring become more sophisticated and less obtrusive, it is likely that more influences of this kind will become apparent.

The role of psychosocial factors in the development of cancer and progression of HIV infection are controversial issues that are discussed in other chapters of this volume. It is probable that if psychophysiological processes are involved, they act through the dysregulation of immune surveillance in a fashion similar to the mechanisms outlined in this section.

8.02.7.4 Physiological Reactivity as a Trigger of Acute Clinical Events

The fourth distinct psychophysiological mechanism in disease concerns the elicitation of acute clinical events. It is possible that psychophysiological responses can be so intense as to trigger episodes of serious illness, and possibly even death. Such an effect is unlikely in the absence of severe underlying disease rendering the victim especially vulnerable. The mechanism is distinctive in that a single occurrence may have serious clinical consequences.

It has been proposed that such a process may be involved in the precipitation of epileptic seizures (Webster & Mawer, 1989). One study of civilians under threat of missile attack showed seizures to be triggered by raid alarms in a small group of patients (Neufeld, Sadeh, Cohn, & Korczyn, 1994). It is conceivable that the process is also relevant to the occurrence of stroke, although evidence is sparse (House, Dennis, Mogridge, Hawton, & Warlow, 1990). Probably the best illustration of the impact of psychophysiological processes on acute clinical events is in the cardiovascular field, with the triggering of myocardial infarction and sudden cardiac death. Recent studies of myocardial infarction victims suggest that the disturbances of cardiac rhythm and coronary blood flow described in Section 8.02.3.2.2 may precipitate clinical events in susceptible individuals. One study assessed anger during the hours prior to acute infarction in comparison with other time periods (Mittleman et al., 1995). A high incidence of severe anger was recorded in the two hours before symptom onset. The effect was not present among patients taking beta-blockers, suggesting that stimulation of coronary vasospasm or thrombotic processes by sympathoadrenal activation may have been involved. Emotional triggers of acute cardiac events have been reported in other studies as well (Tofler et al., 1990). Jain, Burg, Soufer, & Zaret (1995) showed that cardiac patients who responded to laboratory mental stress with disturbances of myocardial function were more likely over the next two years to have acute infarctions or develop unstable angina pectoris than were nonresponders to stress. Taken together, these findings suggest a plausible link between psychosocially-induced physiological reactions and acute episodes of clinical disease.

8.02.8 CONCLUSION

The psychophysiology of disease is a vigorous and exciting area of research. It involves the integration of several disciplines including neuroscience, pathophysiology, and health psychology, and in each of these areas new discoveries are constantly changing our levels of understanding. The complexity of links between the brain, peripheral physiological function, and disease risk is formidable, and linear models are of limited value. The broad processes linking psychophysiological factors with disease risk are now understood, although many of the biological mediators remain tantalizingly illusive. Important challenges for future research include delineation of the processes through which psychosocial factors such as social inequality and social isolation affect disease, understanding how emotional and behavioral coping responses can modify physiological reaction patterns and contribute to resistance and vulnerability, and defining the ways in which psychophysiological knowledge can be harnessed for prevention and disease management. Psychophysiology is one of the cornerstones of clinical health psychology, and is of prime importance in understanding how psychological and social experience can influence health and disease.

8.02.9 REFERENCES

Abbott, B. B., Schoen, L. S., & Badia, P. (1984). Predictable and unpredictable shock: Behavioral measures of aversion and physiological measures of stress. *Psychological Bulletin, 96,* 45–71.

Adams, M. R., Kaplan, J. R., Clarkson, T. B., & Koritnik, D. R. (1985). Ovariectomy, social status, and atherosclerosis in cynomolgus monkeys. *Arteriosclerosis, 5,* 192–200.

Adams, P. R., & Adams, G. R. (1984). Mount Saint Helen's ashfall: Evidence for a disaster stress reaction. *American Psychologist, 39,* 252–260.

Adams, S. G., Dammers, P. M., Saia, T. L., Brantley, P. J., & Gaydos, G. R. (1994). Stress, depression and anxiety predict average symptom severity and daily symptom fluctuation in systemic lupus erythematosus. *Journal of Behavioral Medicine, 17,* 459–477.

Affleck, G., Tennen, H., Urrows, S., & Higgins, P. (1994). Person and contextual features of daily stress reactivity: Individual differences in relation to undesirable daily events with mood disturbance and chronic pain intensity. *Journal of Personality and Social Psychology, 66,* 329–340.

Ahmed, R., & Gray, D. (1996). Immunological memory and protective immunity: Understanding their relation. *Science, 272,* 54–60.

Allen, M. T., Stoney, C. M., Owens, J. F., & Matthews, K. A. (1993). Hemodynamic adjustments to laboratory stress: The influence of gender and personality. *Psychosomatic Medicine, 55,* 505–517.

Allgulander, C. (1994). Suicide and mortality patterns in anxiety neurosis and depressive disorder. *Archives of General Psychiatry, 51,* 708–712.

Anderson, E. A., Wallin, B. G., & Mark, A. L. (1987). Dissociation of sympathetic nerve activity in arm and leg muscle during mental stress. *Hypertension, 9,* III-114–III-119.

Anderson, N. B., Myers, H. F., Pickering, T., & Jackson, J. S. (1989). Hypertension in blacks: Psychosocial and biological perspectives. *Journal of Hypertension, 7,* 161–172.

Aneshensel, C. S. (1992). Social stress: Theory and research. *Annual Review of Sociology, 18,* 15–38.

Ardlie, N. G., Glew, G., & Schwartz, C. J. (1966). Influence of catecholamines on nucleotide-induced platelet aggregation. *Nature, 212,* 415–417.

Arena, J. G., Sherman, R. A., Bruno, G. M., & Smith, J. D. (1990). The relationship between situational stress and phantom limb pain: Cross-lagged correlational data from 6 month pain logs. *Journal of Psychosomatic Research, 34,* 71–77.

Arnetz, B. B., Wasserman, J., Petrini, B., Brenner, S. O., Levi, L., Eneroth, P., Salovaara, H., Hjelm, R., Salovaara, L., Theorell, T., & Petterson, I-L. (1987). Immune function in unemployed women. *Psychosomatic Medicine, 49,* 3–12.

Aubert, A., Vega, C., Dantzer, R., & Goodall, G. (1995). Pyrogens specifically disrupt the acquisition of a task involving cognitive processing in the rat. *Brain, Behavior, and Immunity, 9,* 129–148.

Axelrod, J., & Reisine, T. D. (1984). Stress hormones: Their interaction and regulation. *Science, 224,* 452–459.

Bachen, E. A., Manuck, S. B., Cohen, S., Muldoon, M. F., Raible, R., Herbert, T. B., & Rabin, B. S. (1995). Adrenergic blockade ameliorates cellular immune responses to mental stress in humans. *Psychosomatic Medicine, 57,* 366–372.

Bachen, E. A., Manuck, S. B., Marsland, A. L., Cohen, S., Malkoff, S. B., Muldoon, M. F., & Rabin, B. S. (1992). Lymphocyte subset and cellular immune responses to a brief experimental stressor. *Psychosomatic Medicine, 54,* 673–679.

Baron, R. S., Cutrona, C. E., Hicklin, D., Russell, D. W., & Lubaroff, D. M. (1990). Social support and immune function among spouses of cancer patients. *Journal of Personality and Social Psychology, 59,* 344–352.

Benschop, R. J., Nieuwenhuis, E. E., Tromp, E. A., Godaert, G. L., Ballieux, R. E., & van Doornen, L. J. (1994). Effects of beta-adrenergic blockade on immunologic and cardiovascular changes induced by mental stress. *Circulation, 89,* 762–769.

Bergeman, C. S., Plomin, R., Pedersen, N. L., McClearn, G. E., & Nesselroade, J. R. (1990). Genetic and environmental influences on social support: The Swedish Adoption/Twin Study of Aging. *Journal of Gerontology, 45,* 101–106.

Berkman, L. F., Leo-Summers, L., & Horwitz, R. I. (1992). Emotional support and survival after myocardial infarction. *Annals of Internal Medicine, 117,* 1003–1009.

Besedovsky, H. O., & Del Rey, A. (1996). Immune-neuro-endocrine interactions: Facts and hypotheses. *Endocrine Reviews, 17,* 64–102.

Biselli, R., Farrace, S. D., Amelio, R., & Fattorossi, A. (1993). Influence of stress on lymphocyte subset distribution—a flow cytometric study in young student pilots. *Aviation Space and Environmental Medicine, 64,* 116–120.

Bjorntorp, P. (1990). "Portal" adipose tissue as a generator of risk factors for cardiovascular disease and diabetes. *Arteriosclerosis, 10,* 493–496.

Blanchard, E. B., Cornish, P. J., Wittrock, D. A., & Jaccard, J. (1990). Effects of 24-hour ambulatory blood pressure monitoring on daily activities. *Health Psychology, 9,* 647–652.

Blane, D., Brunner, E., & Wilkinson, R. G. (Eds.) (1996). *Health and social organization*. London: Routledge.

Blumenthal, J. A., Jiang, W., Waugh, R. A., Frid, D. J., Morris, J. J., Coleman, R. E., Hanson, M., Babyak, M., Thyrum, E. T., Krantz, D. S., & O'Connor, C. (1995). Mental stress-induced ischemia in the laboratory and ambulatory ischemia during daily life. *Circulation, 92,* 2102–2108.

Bolger, N., & Zuckerman, A. (1995). A framework for studying personality in the stress process. *Journal of Personality and Social Psychology, 69*, 890–902.

Bongard, S., Hodapp, V., Frisch, M., & Lennartz, K. (1994). Effects of active and passive coping on task performance and cardiovascular reactivity. *Journal of Psychophysiology, 8*, 219–230.

Boscarino, J. A. (1996). Post-traumatic stress disorder, exposure to combat, and lower plasma cortisol among Vietnam veterans: Findings and clinical implications. *Journal of Consulting and Clinical Psychology, 64,* 191–201.

Bosch, J. A., Brand, H. S., Ligtenberg, T. J. M., Bermond, B., Hoogstraten, J., & Nieuw-Amerongen, A. V. (1996). Psychological stress as a determinant of protein levels and salivary-induced aggregation of streptococcus gordonii in human whole saliva. *Psychosomatic Medicine, 58*, 374–382.

Boyce, W. T., & Chesterman, E. (1990). Life events, social support, and cardiovascular reactivity in adolescence. *Developmental and Behavioral Pediatrics, 11*, 105–111.

Brandtstadter, J., & Rothermund, K. (1994). Self-percepts of control in middle and later adulthood: Buffering losses by rescaling goals. *Psychology and Aging, 9*, 265–273.

Brantley, P. J., & Jones, G. N. (1993). Daily stress and stress-related disorders. *Annals of Behavioral Medicine, 15,* 17–25.

Brindley, D. N., McCann, B. S., Niaura, R., Stoney, C. M., & Suarez, E. C. (1993). Stress and lipoprotein metabolism: modulators and mechanisms. *Metabolism: Clinical and Experimental, 42* (9 Suppl. 1), 3–15.

Brown, J. D. (1991). Staying fit and staying well: Physical fitness as a moderator of life stress. *Journal of Personality and Social Psychology, 60,* 555–561.

Brown, G. W., & Harris, T. (1978). *Social origins of depression*. London: Tavistock.

Brown-Borg, H. M., Klemcke, H. G., & Blecha, F. (1993). Lymphocyte proliferative responses in neonatal pigs with high or low plasma cortisol concentration after stress induced by restraint. *American Journal of Veterinary Research, 54,* 2015–2020.

Burg, M. M., & Seeman, T. E. (1994). Families and health: The negative side of social ties. *Annals of Behavioral Medicine, 16,* 109–115.

Cacioppo, J. T. (1994). Social neuroscience: Autonomic, neuroendocrine, and immune responses to stress. *Psychophysiology, 31,* 113–128.

Callister, R., Suwarno, N. O., & Seals, D. R. (1992). Sympathetic activity is influenced by task difficulty and stress perception during mental challenge in humans. *Journal of Physiology, 454,* 373–387.

Carroll, D., Smith, G., Sheffield, D., Shipley, M. J., & Marmot, M. G. (1995). Pressor reactions to psychological stress and prediction of future blood pressure: Data from the Whitehall II study. *British Medical Journal, 310,* 771–776.

Carroll, D., Smith, G., Sheffield, D., Shipley, M. J., & Marmot, M. G. (1997). The relationship between socioeconomic status, hostility, and blood pressure reactions to mental stress in men: Data from the Whitehall II study. *Health Psychology, 16,* 131–136.

Carver, C. S., Pozo, C., Harris, S. D., Noriega, V., Scheier, M. F., Robinson, D. S., Ketcham, A. S., Moffat, F. L., Jr., & Clark, K. C. (1993). How coping mediates the effect of optimism on distress: A study of women with early stage breast cancer. *Journal of Personality and Social Psychology, 65,* 375–390.

Castle, S., Wilkins, S., Heck, E., Tanzy, K., & Fahey, J. (1995). Depression in caregivers of demented patients is associated with altered immunity. *Clinical and Experimental Immunology, 101,* 487–493.

Chandra, R. K. (1989). Nutritional regulation of immunity and risk of infection in old age. *Immunology, 67,* 141–147.

Checkley, S. (1996). The neuroendocrinology of depression and chronic stress. *British Medical Bulletin, 52,* 597–617.

Christian, J. J. (1970). Social subordination, population density, and mammalian evolution. *Science, 168,* 84–90.

Chrousos, G. P. (1995). The hypothalamic–pituitary–adrenal axis and immune-mediated inflammation. *New England Journal of Medicine, 332,* 1351–1362.

Cobb, S., & Rose, R. M. (1973). Hypertension, peptic ulcer and diabetes in air traffic controllers. *Journal of the American Medical Association, 224,* 489–492.

Coe, C. L. (1993). Psychosocial factors and immunity in nonhuman primates: a review. *Psychosomatic Medicine, 55,* 298–308.

Cohen, S., Boyle, W. J., Skoner, D. P., Fireman, P., Gwaltney, J. M., Jr., & Newsom, J. T. (1995). State and trait negative affect as predictors of objective and subjective symptoms of respiratory viral infections. *Journal of Personality and Social Psychology, 68,* 159–169.

Cohen, S., Kaplan, J. R., Cunnick, J. E., Manuck, S. B., & Rabin, B. S. (1992). Chronic social stress, affiliation, and cellular immune response in nonhuman primates. *Psychological Science, 3,* 301–304.

Cohen, S., Kessler, R. C., & Gordon, L. U. (1995a). Strategies for measuring stress in studies of psychiatric and physical disorders. In S. Cohen, R. C. Kessler, & L. U. Gordon (Eds.), *Measuring stress: A guide for health and social scientists* (pp. 3–26). New York: Oxford University Press.

Cohen, S., Kessler, R. O., & Gordon, L. U. (Eds.) (1995b). *Measuring stress: A guide for health and social scientists.* New York: Oxford University Press.

Cohen, S., & Rodriquez, M. S. (1995). Pathways linking affective disturbances and physical disorders. *Health Psychology, 14,* 374–380.

Cohen, S., Tyrrell, D. A. J., & Smith, A. P. (1991). Psychosocial stress and susceptibility to the common cold. *New England Journal of Medicine, 325,* 606–612.

Cohen, S., & Williamson, G. M. (1991). Stress and infectious disease in humans. *Psychological Bulletin, 109,* 5–24.

Cole, P. M., Zahn-Waxler, C., Fox, N. A., Usher, B. A., & Welsh, J. D. (1996). Individual differences in emotion regulation and behavior problems in preschool children. *Journal of Abnormal Psychology, 103,* 518–529.

Cole, S. W., Kemeny, M. E., Taylor, S. E., & Visscher, B. R. (1996). Elevated physical health risk among gay men who conceal their homosexual identity. *Health Psychology, 15,* 243–251.

Dalkvist, J. Wahlin, T.-B. R., Bartsch, E., & Forsbeck, M. (1995). Herpes simplex and mood: A prospective study. *Psychosomatic Medicine, 57,* 127–137.

Dantzer, R., & Kelley, K. W. (1989). Stress and immunity: An integrated view of relationships between the brain and the immune system. *Life Sciences, 44,* 1995–2008.

Davis, M. C., Matthews, K. A., Meilahn, E. N., & Kiss, J. E. (1995). Are job characteristics related to fibrinogen levels in middle-aged women? *Health Psychology, 14,* 310–318.

De Bellis, M. D., Lefter, L., Trickett, P. K., & Putnam, F. W., Jr. (1994). Urinary catecholamine excretion in sexually abused girls. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 320–327.

De Boer, S. F., Van der Gugten, J., & Slangen, J. L. (1989). Plasma catecholamine and corticosteroid responses to predictable and unpredictable noise stress in rats. *Physiology and Behavior, 45,* 789–795.

De Geus, E. J. C., Van Doornen, L. J. P., & Orlebeke, J. F. (1993). Regular exercise and aerobic fitness in relation to psychological make-up and physiological stress reactivity. *Psychosomatic Medicine, 55,* 347–363.

Deanfield, J. E., Shea, M, Kensett, M., Horlock, P., Wilson, R. A., De Landsheere, C. M., & Selwyn, A. P. (1984). Silent myocardial infarction due to mental stress. *Lancet, ii,* 1001–1005.

Dienstbier, R. A. (1989). Arousal and physiological toughness: implications for mental and physical health. *Psychological Review, 96,* 84–100.

Dimsdale, J. E., Mills, P., Patterson, T., Ziegler, M., & Dillon, E. (1994). Effects of chronic stress on beta-adrenergic receptors in the homeless. *Psychosomatic Medicine, 56,* 290–295.

Dressler, W. W. (1991a). Social support, lifestyle incongruity, and arterial blood pressure in a southern black community. *Psychosomatic Medicine, 53,* 608–620.

Dressler, W. W. (1991b). *Stress and Adaptation in the Context of Culture: Depression in a southern black community*. Albany, NY: SUNY Press.

Dutz, W., Kohout, E., Rossipal, E., & Vessal, K. (1976). Infantile stress, immune modulations, and disease patterns. *Pathology Annual, 11,* 415–454.

Dyer, A. R., Persky, V., Stamler, J., Paul, O., Shekelle, R. B., Berkson, D. M., Lepper, M., Schoenberger, J. A., & Lindberg, H. A. (1980). Heart rate as a prognostic factor for coronary heart disease and mortality: Findings in three Chicago epidemiologic studies. *American Journal of Epidemiology, 112,* 736–749.

Engel, B. T. (1985). Stress is a noun! No, a verb! No, an adjective! In T. Field, P. M. McCabe, & N. Schneiderman (Eds.), *Stress and coping* (pp. 3–14). Hillsdale, NJ: Erlbaum.

Esterling, B. A., Antoni, M. H., Kumar, M., & Schneiderman, N. (1993). Defensiveness, trait anxiety and Epstein-Barr viral capsid antigen antibody titers in healthy college students. *Health Psychology, 12,* 132–139.

Evans, G. W., Hygge, S., & Bullinger, M. (1995). Chronic noise and psychological stress. *Psychological Science, 6,* 334–338.

Evans, P., Bristow, M., Hucklebridge, F., Clow, A., & Walters, N. (1993). The relationship between secretory immunity, mood and life events. *British Journal of Clinical Psychology, 32,* 227–236.

Fleming, I., Baum, A., Davidson, L. M., Rectanus, E., & McArdle, S. (1987). Chronic stress as a factor in physiologic reactivity to challenge. *Health Psychology, 6,* 221–237.

Flor, H., & Turk, D. C. (1989). Psychophysiology of chronic pain: Do chronic pain patients exhibit symptom-specific psychophysiological responses? *Psychological Bulletin, 105,* 215–259.

Frankenhaeuser, M., Dunne, E., & Lundberg, U. (1976). Sex differences in sympathetic-adrenal medullary responses induced by different stressors. *Psychopharmacology, 47,* 1–5.

Frankenhaeuser, M., Lundberg, U., & Forsman, L. (1980). Dissociation between sympatho-adrenal and pituitary-adrenal responses to an achievement situation characterised by high controllability: Comparison between Type A and Type B males and females. *Biological Psychology, 10*, 79–91.

Frankenhaeuser, M., Lundberg, U., Fredrikson, M., Melin, B., Tuomisto, M., & Myrsten, A.-L. (1989). Stress on and off the job as related to sex and occupational stress in white-collar workers. *Journal of Organizational Behavior, 10*, 321–346.

Frasure-Smith, N., Lespérance, F., & Talajic, M. (1993). Depression following myocardial infarction: impact on 6 month survival. *Journal of the American Medical Association, 270*, 1819–1825.

Fredrikson, M., Lundberg, U., & Tuomisto, M. (1991). Serum lipid levels and cardiovascular reactivity. *Journal of Psychophysiology, 5*, 89–95.

Friedman, H. S., & Booth-Kewley, S. (1987). The "disease-prone personality": A meta-analytic view of the construct. *American Psychologist, 42*, 539–555.

Friedman, H. S., Tucker, J. S., Schwartz, J. E., Martin, L. R., Tomlinson-Keasey, C., Wingard, D. L., & Criqui, M. H. (1995). Childhood conscientiousness and longevity: Health behaviors and cause of death. *Journal of Personality and Social Psychology, 68*, 696–703.

Funk, S. C. (1992). Hardiness: A review of theory and research. *Health Psychology, 11*, 335–345.

Gabbay, F. H., Krantz, D. S., Kop, W. J., Hedges, S. M., Klein, J., Gottdiener, J. S., & Rozanski, A. (1996). Triggers of myocardial ischemia during daily life in patients with coronary artery disease: physical and mental activities, anger and smoking. *Journal of the American College of Cardiology, 27*, 585–592.

Garbarino, J., Kostelny, K., & Dubrow, N. (1991). What children can tell us about living in danger. *American Psychologist, 46*, 376–383.

Gatchel, R. J., Schaeffer, M. A., & Baum, A. (1985). A psychophysiological field study of stress at Three Mile Island. *Psychophysiology, 22*, 175–181.

Gerritsen, W., Heijnen, C. J., Weigant, V. M., Bermond, B., & Frijda, N. H. (1996). Experimental social fear: Immunological, hormonal, and autonomic concomitants. *Psychosomatic Medicine, 58*, 273–286.

Gilad, G. M., & Gilad, V. H. (1995). Strain, stress, neurodegeneration and longevity. *Mechanisms of Ageing and Development, 78*, 75–83.

Glaser, R., Kiecolt-Glaser, J. K., Bonneau, R. H., Malarkey, W., Kennedy, S., & Hughes, J. (1992). Stress-induced modulation of the immune response to recombinant hepatitis B vaccine. *Psychosomatic Medicine, 54*, 22–29.

Glavin, G. B. (1992). Dopamine: A stress modulator in the brain and gut. *General Pharmacology, 23*, 1023–1026.

Glavin, G. B., Pare, W. P., Sandbak, T., Bakke, H. K., & Murison, R. (1994). Restraint stress in biomedical research: An update. *Neuroscience and Biobehavioral Reviews, 18,* 223–249.

Goldston, D. B., Kovacs, M., Obrosky, D. S., & Iyengar, S. (1995). A longitudinal study of life events and metabolic control among youths with insulin-dependent diabetes mellitus. *Health Psychology, 14*, 409–414

Gotthardt, U., Schweiger, U., Fahrenberg, J., Lauer, C. J., Holsboer, F., & Heuser, I. (1995). Cortisol, ACTH, and cardiovascular response to a cognitive challenge paradigm in aging and depression. *American Journal of Physiology, 268*, R865–R873.

Gray, J. A. (1995). A model of the limbic system and basal ganglia: Applications to anxiety and schizophrenia. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1165–1176). Cambridge, MA: MIT Press.

Grignani, G., Soffiantino, F., Zucchella, M., Pacchiarini, L., Taccone, F., Bonomi, E., Pastoris, A., Sbaffi, A., Fratino, P., & Tavazzi, L. (1991). Platelet activation by emotional stress in patients with coronary artery disease. *Circulation, 83*(Suppl. 11), 128–136.

Haft, J. I., & Arkel, Y. S. (1976). Effect of emotional stress on platelet aggregation in humans. *Chest, 70*, 501–505.

Halford, W. K., Cuddihy, S., & Mortimer, R. H. (1990). Psychological stress and blood glucose regulation in Type 1 diabetic patients. *Health Psychology, 9*, 516–528.

Hanson, B. S., Isacsson, S.-O., Janzon, L., Lindell, S.-E., & Rastam, L. (1988). Social anchorage and blood pressure in elderly men—a population study. *Journal of Hypertension, 6*, 503–510.

Hatch, J. P., Prihoda, T. J., Moore, P. J., Cyr-Provost, M., Borcherding, S., Boutros, N. N., & Seleshi, E. (1991). A naturalistic study of the relationships among electromyographic activity, psychological stress, and pain in ambulatory tension-type headache patients and headache-free controls. *Psychosomatic Medicine, 53,* 576–584.

Heckhausen, J., & Schulz, R. (1995). A life-span theory of control. *Psychological Review, 102,* 284–304.

Helgeson, V. S., & Cohen, S. (1996). Social support and adjustment to cancer: Reconciling descriptive, correlational and intervention research. *Health Psychology, 15,* 135–148.

Herbert, T. B., & Cohen, S. (1993a). Depression and immunity: A meta-analytic review. *Psychological Bulletin, 113,* 472–486.

Herbert, T. B., & Cohen, S. (1993b). Stress and immunity in humans: A meta-analytic review. *Psychosomatic Medicine, 55,* 364–379.

Herbert, T. B., Cohen, S., Marsland, A. L., Bachen, E. A., Rabin, B. S., Muldoon, M. F., & Manuck, S. B. (1994). Cardiovascular reactivity and the course of immune response to an acute psychological stressor. *Psychosomatic Medicine, 56,* 337–344.

Herman, J. P., Adams, D., & Prewitt, C. (1995). Regulatory changes in neuroendocrine stress-integrative circuitry produced by a variable stress paradigm. *Neuroendocrinology, 61,* 180–190.

Hewitt, J. K., & Turner, J. R. (1995). Behavior genetic studies of cardiovascular responses to stress. In J. R. Turner, L. R. Cardon, & J. K. Hewitt (Eds.), *Behavior genetic approaches in behavioral medicine* (pp. 87–103). New York: Plenum.

Hinderliter, A. L., Light, K. C., Girdler, S. S., Willis, P. W., & Sherwood, A. (1996). Blood pressure responses to stress: Relation to left ventricular structure and function. *Annals of Behavioral Medicine, 18,* 61–66.

Hoehn-Saric, R., McLeod, D. R., & Zimmerli, W. D. (1989). Somatic manifestations in women with generalized anxiety disorder: Psychophysiological responses to psychological stress. *Archives of General Psychiatry, 46,* 1113–1119.

Holsboer, F., & Barden, N. (1996). Antidepressants and hypothalamic–pituitary–adrenocortical regulation. *Endocrine Reviews, 17,* 187–205.

Holtmann, G., Kriebel, R., & Singer, M. V. (1990). Mental stress and gastric acid secretion: Do personality traits influence the response? *Digestive Disease and Science, 35,* 998–1007.

House, A., Dennis, M., Mogridge, L., Hawton, K., & Warlow, C. (1990). Life events and difficulties preceding stroke. *Journal of Neurology, Neurosurgery and Psychiatry, 53,* 1024–1028.

House, J. S., Landis, A. R., & Umberson, D. (1988). Social relationships and health. *Science, 241,* 540–545.

Huppert, F. A., & Whittington, J. E. (1995). Symptoms of psychological distress predict 7-year mortality. *Psychological Medicine, 25,* 1073–1086.

Jabaaij, L., Benschop, R. J., Oostveen, F. G., Duivenvoorden, H. J., Ballieux, R. E., & Vingerhoets, A. J. J. M. (1993a). Immunologic, endocrine and psychological influences on cortisol-induced immunoglobulin synthesis in vitro. *Psychoneuroendocrinology, 18,* 591–605.

Jabaaij, L., Grosheide, P. M., Heijtink, R. A., Duivenvoorden, H. J., Ballieux, R. E., & Vingerhoets, A. J. J. M. (1993b). Influence of perceived psychological stress and distress on antibody response to low dose rDNA hepatitis B vaccine. *Journal of Psychosomatic Research, 37,* 361–369.

Jabaaij, L., van Hattum, J., Vingerhoets, A. J. J. M., Oostveen, F. G., Duivenvoorden, H. J., & Ballieux, R. E. (1996). Modulation of immune response to rDNA hepatitis B vaccination by psychological stress. *Journal of Psychosomatic Research, 41,* 129–137.

Jain, D., Burg, M., Soufer, A., & Zaret, B. L. (1995). Prognostic implications of mental stress-induced silent left ventricular dysfunction in patients with stable angina pectoris. *American Journal of Cardiology, 76,* 31–35.

Jain, R., Zwickler, D., Hollander, C., Brand, H., Saperstein, A., Hutchinson, B., Brown, C., & Audhya, T. (1991). Corticotropin-releasing factor modulates the immune response to stress in the rat. *Endocrinology, 128,* 1329–1336.

Janse, M. J., & Davies, M. J. (1996). Generation and conduction of the heartbeat and their disturbances. In D. G. Julian, A. J. Camm, K. M. Fox, R. J. C. Hall, & P. A. Poole-Wilson (Eds.), *Diseases of the heart* (pp. 88–114). London: Saunders.

Jayo, J. M., Shively, C. A., Kaplan, J. R., & Manuck, S. B. (1993). Effects of exercise and stress on body fat distribution in male cynomolgus monkeys. *International Journal of Obesity and Related Metabolic Disorders, 17,* 597–604.

Johansson, G., Aronsson, G., & Lindstrom, B. O. (1978). Social psychological and neuroendocrine stress reactions to highly mechanised work. *Ergonomics, 21,* 583–599.

Jones, D. R. (1987). Heart disease mortality following widowhood: Some results from the OPCS Longitudinal Study. *Journal of Psychosomatic Research, 31,* 325–333.

Jorgensen, R. S., Johnson, B. T., Schreer, G. E., & Kolodziej, M. E. (1996). Elevated blood pressure and personality: A meta-analytic review. *Psychological Bulletin, 120,* 293–320.

Julius, S. (1988). The blood pressure seeking properties of the central nervous system. *Journal of Hypertension, 6,* 177–185.

Kamarck, T. W., Annunziato, B., & Amateau, L. M. (1995). Affiliation moderates the effects of social threat on stress-related cardiovascular responses: Boundary conditions for a laboratory model of social support. *Psychosomatic Medicine, 57,* 183–194.

Kamarck, T. W., Jennings, J. R., Debski, T. T., Glickman-Weiss, E., Johnson, P. S., Eddy, M. J., & Manuck, S. B. (1992). Reliable measures of behaviorally-evoked cardiovascular reactivity from a PC-based test battery: Results from student and community samples. *Psychophysiology, 29,* 17–28.

Kaplan, G. A., & Keil, J. E. (1993). Socioeconomic factors and cardiovascular disease: A review of the literature. *Circulation, 88,* 1973–1998.

Kaplan, J. R., Manuck, S. B., Williams, J. K., & Strawn, W. (1993). Psychosocial influences on atherosclerosis: Evidence for effects and mechanisms in nonhuman primates. In J. Blascovich & E. S. Katkin (Eds.), *Cardiovascular reactivity to psychological stress and disease* (pp. 3–26). Washington, DC: American Psychological Association.

Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative Sciences Quarterly, 24,* 285–307.

Kiecolt-Glaser, J. K., Dura, J. R., Speicher, C. E., Trask, O. J., & Glaser, R. (1991). Spousal caregivers of dementia victims: Longitudinal changes in immunity and health. *Psychosomatic Medicine, 53,* 345–362.

Kiecolt-Glaser, J. K., & Glaser, R. (1995). Measurement of immune response. In S. Cohen, R. C. Kessler, & L. Underwood Gordon (Eds.), *Measuring stress: A guide for health and social scientists* (pp. 213–219). New York: Oxford University Press.

Kiecolt-Glaser, J. K., Malarkey, W. B., Cacioppo, J. T., & Glaser, R. (1994). Stressful personal relationships: Immune and endocrine function. In R. Glaser & J. K. Kiecolt-Glaser (Eds.), *Handbook of human stress and immunity* (pp. 321–339). San Diego, CA: Academic Press.

Kiecolt-Glaser, J. K., Malarkey, W. B., Chee, M., Newton, T., Cacioppo, J. T., Mao, H.-Y., & Glaser, R. (1993). Negative behavior during marital conflict is associated with immunological down-regulation. *Psychosomatic Medicine, 55,* 395–409.

Kiecolt-Glaser, J. K., Marucha, P. T., Malarkey, W. B., Mercado, A. M., & Glaser, R. (1995). Slowing of wound healing by psychological stress. *Lancet, 346,* 1194–1196.

Kirschbaum, C., Prussner, J. C., Stone, A. A., Federenko, I., Gaab, J., Lintz, D., Schommer, N., & Hellhammer, D. H. (1995). Persistent high cortisol responses to repeated psychological stress in a subpopulation of healthy men. *Psychosomatic Medicine, 57,* 468–474.

Kirschbaum, C., Wust, S., & Hellhammer, D. (1992). Consistent sex differences in cortisol responses to psychological stress. *Psychosomatic Medicine, 54,* 648–657.

Klein, F., Lemaire, V., Sandi, C., Vitiello, S., Van der Logt, J., Laurent, P. E., Neveu, P., Le Moal, M., & Mormede, P. (1992). Prolonged increase of corticosterone secretion by chronic social stress does not necessarily impair immune functions. *Life Sciences, 50,* 723–731.

Klein, S. L., & Rager, D. R. (1995). Prenatal stress alters immune function in the offspring of rats. *Developmental Psychobiology, 28,* 321–336.

Klieger, R. E., Miller, P., Bigger, J. T., Jr., & Moss, A. J. (1987). Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *American Journal of Cardiology, 59,* 256–262.

Kobasa, S. C., Maddi, S. R., Puccetti, M. C., & Zola, M. A. (1985). Effectiveness of hardiness, exercise and social support as resources against illness. *Journal of Psychosomatic Research, 29,* 525–533.

Kojima, K., Ogomori, K., Mori, Y., Hirata, K., Kinukawa, N., & Tashiro, N. (1996). Relationship of emotional behaviors induced by electrical stimulation of the hypothalamus to changes in EKG, heart, stomach, adrenal glands and thymus. *Psychosomatic Medicine, 58,* 383–391.

Krittayaphong, R., Light, K. C., Biles, P. L., Ballenger, M. N., & Sheps, D. S. (1995). Increased heart rate response to laboratory-induced mental stress predicts frequency and duration of daily life ambulatory myocardial ischemia in patients with coronary artery disease. *American Journal of Cardiology, 76,* 657–660.

Kumar, D., & Wingate, D. L. (1985). The irritable bowel syndrome: A paroxysmal motor disorder. *Lancet, 2,* 973–977.

Lakatta, E. G. (1993). Deficient neuroendocrine regulation of the cardiovascular system with advancing age in healthy humans. *Circulation, 87,* 631–636.

Larsson, P. T., Hjemdahl, P., Olsson, G., Egberg, N., & Hornstra, G. (1989). Altered platelet function during mental stress and adrenaline infusion in humans: Evidence for an increased aggregability in vivo as measured by filtragometry. *Clinical Science, 76,* 369–376.

Laudenslager, M. L., & Fleshner, M. (1994). Stress and immunity: Of mice, monkeys, models and mechanisms. In R. Glaser, & J. Kiecolt-Glaser (Eds.), *Handbook of human stress and immunity* (pp. 161–181). San Diego, CA: Academic Press.

Legault, S. E., Freeman, M. R., Langer, A., & Armstrong, P. W. (1995). Pathophysiology and time course of silent myocardial ischaemia during mental stress: Clinical, anatomical, and physiological correlates. *British Heart Journal, 73,* 242–249.

Lehrer, P. M., Hochron, S., Carr, R., Edelberg, R., Hamer, R., Jackson, A., & Porges, S. (1996). Behavioral task-induced bronchodilation in asthma during active and passive tasks: A possible cholinergic link to psychologically induced airway changes. *Psychosomatic Medicine, 58,* 413–422.

Leiker, M., & Harley, B. J. (1988). A link between hostility and disease: Poor health habits? *Behavioral Medicine, 14,* 129–133.

Lenders, J. W., Peters, J. H., Pieters, G. F., Willemsen, J. J., & Thien, T. (1988). Hemodynamic reactivity to sympathoadrenal stimulation in adrenalectomized women. *Journal of clinical Endocrinology and Metabolism, 67,* 139–143.

Levine, S. P., Towell, B. L., Suarez, A. M., Knieriem, L. K., Harris, M. M., & George, J. N. (1985). Platelet activation and secretion associated with emotional stress. *Circulation, 71,* 1129–1134.

Lewis, M., Thomas, D. A., & Worobey, J. (1990). Developmental organization, stress and illness. *Psychological Science, 1,* 316–318.

Light, K. C., Dolan, C. A., Davis, M. R., & Sherwood, A. (1992). Cardiovascular responses to an active coping challenge as predictors of blood pressure patterns 10 to 15 years later. *Psychosomatic Medicine, 54,* 217–230.

Light, K. C., Koepke, J. P., Obrist, P. A., & Willis, P. W. (1983). Psychological stress induces sodium and fluid retention in men at high risk for hypertension. *Science, 220,* 429–431.

Linde, B., Hjemdahl, P., Freyschuss, U., & Juhlin-Dannfelt, A. (1989). Adipose tissue and skeletal muscle blood flow during mental stress. *American Journal of Physiology, 256,* E12–E18.

Line, S. W., Kaplan, J. R., Heise, E. R., Hilliard, J. K., Cohen, S., Rabin, B. S., & Manuck, S. B. (1996). Effects of social reorganization on cellular immunity in male cynomolgus monkeys. *American Journal of Primatology, 39,* 235–249.

Lorber, B. (1996). Are all diseases infectious? *Annals of Internal Medicine, 125,* 844–851.

Lown, B., Verrier, R. L., & Rabinowitz, S. H. (1977). Neural and psychologic mechanisms and the problem of sudden cardiac death. *American Journal of Cardiology, 39,* 890–902.

Lynch, J. W., Kaplan, G. A., Cohen, R. D., Tuomilehto, J., & Salonen, J. (1996). Do cardiovascular risk factors explain the relation between socio-economic status, risk of all-cause mortality, cardiovascular mortality, and acute myocardial infarction? *American Journal of Epidemiology, 144,* 934–942.

Madden, J., Akil, H., Patrick, R. L., & Barchas, J. D. (1977). Stress-induced parallel changes in central opioid levels and pain responsiveness in the rat. *Nature, 265,* 358–360.

Maier, S. F., Watkins, L. R., & Fleshner, M. (1994). Psychoneuroimmunology: The interface between behavior, brain, and immunity. *American Psychologist, 49,* 1004–1017.

Malarkey, W. B., Kiecolt-Glaser, J. K., Pearl, D., & Glaser, R. (1994). Hostile behavior during marital conflict alters pituitary and adrenal hormones. *Psychosomatic Medicine, 56,* 41–51.

Malarkey, W. B., Pearl, D. K., Demers, L. M., Kiecolt-Glaser, J. K., & Glaser R. (1995). Influence of academic stress and season on 24-hour mean concentrations of ACTH, cortisol, and beta endorphin. *Psychoeneuroendocrinology, 20,* 499–508.

Manuck, S. B. (1994). Cardiovascular reactivity in cardiovascular disease: "Once more unto the breach." *International Journal of Behavioral Medicine, 1,* 4–31.

Manuck, S. B., Kaplan, J. R., Adams, M. R., & Clarkson, T. B. (1989). Behaviorally elicited heart rate reactivity and atherosclerosis in female cynomolgus monkeys (*Macaca fascicularis*). *Psychosomatic Medicine, 51,* 306–318.

Markovitz, J. H., Matthews, K. A., Kriss, J., & Smitherman, T. C. (1996). Effects of hostility on platelet reactivity to psychological stress in coronary heart disease patients and in healthy controls. *Psychosomatic Medicine, 58,* 143–149.

Markowe, H. L., Marmot, M. G., Shipley, M. J., Bulpitt, C. J., Meade, T. W., Stirling, Y., Vickers, M. V., & Semmence, A. (1985). Fibrinogen: A possible link between social class and coronary heart disease. *British Medical Journal, 291,* 1312–1314.

Marmot, M. G., Davey Smith, G., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E., & Feeney, A. (1991). Health inequalities among British civil servants: The Whitehall II study. *Lancet, 337,* 1387–1393.

Mason, J. W. (1968). A review of psychoendocrine research on the pituitary-adrenal cortical system. *Psychosomatic Medicine, 30,* 576–607.

Mason, J. W. (1975). Emotion as reflected in patterns of endocrine integration. In L. Levi (Ed.), *Emotions—their parameters and measurement* (pp. 143–181). New York: Raven Press.

McCann, B. S., Magee, M. S., Broyles, F. C., Vaughan, M., Albers, J. J., & Knopp, R. H. (1995). Acute psychological stress and epinephrine infusion in normolipidemic and hyperlipidemic men: Effects on plasma lipid and apoprotein concentrations. *Psychosomatic Medicine, 57,* 165–176.

McCubbin, J. A. (1993). Stress and endogenous opioids: Behavioral and circulatory interactions. *Biological Psychology, 35,* 91–22.

McCubbin, J. A., Surwit, R. S., Williams, R. B., Nemeroff, C. B., & McNeilly, M. (1989). Altered pituitary hormone response to naloxone in hypertension development. *Hypertension, 14,* 633–644.

McEwen, B. S., & Sapolsky, R. M. (1995). Stress and cognitive function. *Current Opinion in Neurobiology, 5,* 205–216.

McKinnon, A., Weisse, C. S., Reynolds, C. P., Bowles, C. A., & Baum, A. (1989). Chronic stress, leukocyte subpopulations, and humoral response to latent viruses. *Health Psychology, 8,* 389–402.

Meredith, I. T., Broughton, A., Jennings, G. L., & Esler, M. D. (1991). Evidence of selective increase in cardiac sympathetic activity in patients with sustained ventricular arrhythmias. *New England Journal of Medicine, 325,* 618–624.

Miller, T. Q., Smith, T. W., Turner, C. W., Guijarro, M. L., & Hallet, A. J. (1996). A meta-analytic review of research on hostility and physical health. *Psychological Bulletin, 119,* 322–348.

Mirowksy, J., & Ross, C. E. (1995). Sex differences in distress: Real or artefact? *American Sociological Review, 60,* 449–468.

Mittleman, M. A., Maclure, M., Sherwood, J. B., Mulry, R. P., Tofler, G. H., Jacobs, S. C. Friedman, R., Benson, H., & Muller, J. E. (1995). Triggering of acute myocardial infarction onset by episodes of anger. *Circulation, 92,* 1720–1725.

Moldin, S. O., Scheftner, W. A., Rice, J. P., Nelson, E., Knesevich, M. A., & Akiskal, H. (1993). Association between major depressive disorder and physical illness. *Psychological Medicine, 23,* 755–761.

Mormede, P., Dantzer, R., Michaud, B., Kelley, K. W., & Le Moal, M. (1988). Influence of stressor predictability and behavioral control on lymphocyte reactivity, antibody responses and neuroendocrine activation in rats. *Physiology and Behavior, 43,* 577–583.

Morris, J. K., Cook, D. G., & Shaper, A. G. (1994). Loss of employment and mortality. *British Medical Journal, 308,* 1135–1139.

Morris, M., Salmon, P., Steinberg, H, Sykes, E. A., Bouloux, P., Newbould, E., McLoughlin, L., Besser, G. M., & Grossman, A. (1990). Endogenous opioids modulate the cardiovascular response to mental stress. *Psychoneuroendocrinology, 15,* 185–192.

Moyer, A. E., Rodin, J., Grilo, C. M., Cummings, N., Larson, L. M., & Rebuffé-Scrive, M. (1994). Stress-induced cortisol response: A fat distribution in women. *Obesity Research, 2,* 255–261.

Munck, A. Guyre, P. M., & Holbrook, N. J. (1984). Physiological functions of glucocorticoids in stress and their relation to pharmacological actions. *Endocrine Reviews, 5,* 25–44.

Murphy, J. K., Stoney, C. M., Alpert, B. S., & Walker, S. S. (1995). Gender and ethnicity in children's reactivity: 7 years of study. *Health Psychology, 14,* 48–55.

Murray, C. J. L., & Lopez, A. D. (Eds.) (1996). *The global burden of disease*. Cambridge, MA: Harvard University Press.

Nachmias, M., Gunnar, M., Mangelsdorf, S., Parritz, R. H., & Buss, K. (1996). Behavioral inhibition and stress reactivity: The moderating role of attachment security. *Child Development, 67,* 508–522.

Natelson, B. H., Ottenweller, J. E., Cook, J. A., Pitman, D., McCartey, R., & Tapp, W. N. (1988). Effect of stressor intensity on habituation of the adrenocortical stress response. *Physiology and Behavior, 43,* 41–46.

Neufeld, M. Y., Sadeh, M., Cohn, D. F., & Korczyn, A. D. (1994). Stress and epilepsy: The Gulf War experience. *Seizure, 3,* 135–139.

Nilsson, P. M., Moller, L., & Solstad, K. (1995). Adverse effects of psychosocial stress on gonadal function and insulin levels in middle-aged males. *Journal of Internal Medicine, 237,* 479–486.

Nordentoft, M., Lou, H. C., & Hanson, D. E. (1996). Intrauterine growth retardation and premature delivery: The influence of maternal smoking and psychosocial factors. *American Journal of Public Health, 86,* 347–354.

Norris, F. H., & Murrell, S. A. (1988). Prior experience as a moderator of disaster impact on anxiety symptoms in older adults. *American Journal of Community Psychology, 16,* 665–683.

O'Connor, F., Buckley, M., & O'Morain, C. (1996). Helicobacter pylori: The cancer link. *Journal of the Royal Society of Medicine, 89,* 674–678.

Owens, J. F., Stoney, C. M., & Matthews, K. A. (1993). Menopausal status influences ambulatory blood pressure levels and blood pressure changes during mental stress. *Circulation, 88,* 2794–2802.

Pagani, M., Rimoldi, O., Pizzinelli, P. Furlan, R., Crivellaro, W., Liberati, D., Cerutti, S., & Malliani, A. (1991). Assessment of the neural control of the circulation during psychological stress. *Journal of the Autonomic Nervous System, 35,* 33–42.

Palmblad, J., Blomback, M., Egberg, N., Froberg, J., Karlsson, C.-G., & Levi, L. (1977). Experimentally induced stress in man: Effects on blood coagulation and fibrinolysis. *Journal of Psychosomatic Research, 21,* 87–92.

Parkes, K. R. (1984). Locus of control, cognitive appraisal, and coping in stressful episodes. *Journal of Personality and Social Psychology, 46,* 655–668.

Pennebaker, J. W. (Ed.) (1995). *Emotion, disclosure, and health*. Washington, DC: American Psychological Association.

Pennebaker, J. W., Kiecolt-Glaser, J. K., & Glaser, R. (1988). Disclosure of traumas and immune function: Health implications for psychotherapy. *Journal of Consulting and Clinical Psychology, 56,* 239–245.

Perski, A., Olsson, G., Landou, C., de Faire, U., Theorell, T., & Hamsten, A. (1992). Minimum heart rate and coronary atherosclerosis: Independent relations to global severity and rate of progression of angiographic lesions in men with myocardial infarction at a young age. *American Heart Journal, 123,* 609–616.

Peterson, C., Seligman, M. E., & Vaillant, G. E. (1988). Pessimistic explanatory style is a risk factor for physical illness: A thirty-five year longitudinal study. *Journal of Personality and Social Psychology, 55,* 23–27.

Pickering, T. G. (1991). *Ambulatory monitoring and blood pressure variability*. London: Science Press.

Pillow, D. R., Zautra, A. J., & Sandler, I. (1996). Major life events and minor stressors: Identifying mediational links in the stress process. *Journal of Personality and Social Psychology, 70,* 381–394.

Plomin, R., Lichtenstein, P., Pedersen, N. L., McClearn, G. E., & Nesselroade, J. R. (1990) Genetic influence on life events during the last half of the life span. *Psychology and Aging, 5,* 25–30.

Pope, M. K., & Smith, T. W. (1991). Cortisol excretion in high and low cynically hostile men. *Psychosomatic Medicine, 53,* 386–392.

Rabin, D., Gold, P. W., Margioris, A., & Chrousos, G. P. (1988). Stress and reproduction: interactions between the stress and reproductive axes. In G. P. Chrousos, D. L. Loriaux, & P. W. Gold (Eds), *Mechanisms of physical and emotional stress* (pp. 377–387). New York: Plenum.

Rahe, R. H., Karson, S., Howard, N. S., Rubin, R. T., & Poland, R. E. (1990). Psychological and physiological assessments on American hostages freed from captivity in Iran. *Psychosomatic Medicine, 52*, 1–16.

Ramirez, A. J., Craig, T. K., Watson, J. P., Fentiman, I. S., North, W. R., & Rubens, R. D. (1989). Stress and relapse of breast cancer. *British Medical Journal, 298*, 291–293.

Reaven, G. M. (1988). Role of insulin resistance in human disease. *Diabetes, 37*, 1595–1607.

Robinson, N., & Fuller, J. H. (1985). Role of life events and difficulties in the onset of diabetes mellitus. *Journal of Psychosomatic Research, 29*, 583–591.

Rogers, M. W., Probst, M. M., Gruber, J. J., Berger, R., & Boone, J. B. (1996). Differential effects of exercise training intensity on blood pressure and cardiovascular responses to stress in borderline hypertensive humans. *Journal of Hypertension, 14*, 1369–1375.

Rose, R. M., & Fogg, L. F. (1993). Definition of a responder: Analysis of behavior, cardiovascular and endocrine responses to varied workload in air traffic controllers. *Psychosomatic Medicine, 55*, 325–338.

Rostrup, M., Mundal, H. H., Westheim, A., & Eide, I. (1991). Awareness of high blood pressure increases arterial plasma catecholamines, platelet noradrenaline and adrenergic responses to mental stress. *Journal of Hypertension, 9*, 159–166.

Rozanski, A., Bairey, C. N., Krantz, D. S., Friedman, J., Resser, K. J., Morell, M., Hilton-Chalfen, S., Hestrin, L., Bietendorf, J., & Berman, D. S. (1988). Mental stress and the induction of silent myocardial ischemia in patients with coronary artery disease. *New England Journal of Medicine, 318*, 1005–1012.

Saab, P. G., Llabre, M. M., Hurwitz, B. E., Frame, C. A., Reineke, L. J., Fins, A. I., McCalla, J., Cieply, L. K., & Schneiderman, N. (1992). Myocardial and peripheral vascular responses to behavioral challenges and their stability in black and white Americans. *Psychophysiology, 29*, 384–397.

Sapolsky, R. M. (1992). *Stress, the aging brain, and the mechanisms of neuron death*. Cambridge, MA: MIT Press.

Sapolsky, R. M., & Mott, G. E. (1987). Social subordinance in wild baboons is associated with suppressed high density lipoprotein-cholesterol concentrations: The possible role of chronic social stress. *Endocrinology, 121*, 1605–1610.

Schaie, K. W., Blazer, D. G. & House, J. S. (Eds) (1992). *Aging, health behaviors and health outcomes*. Hillsdale, NJ: Erlbaum.

Schaubroeck, J., & Ganster, D. C. (1993). Chronic demands and responsivity to challenge. *Journal of Applied Psychology, 78*, 73–85.

Scheier, M. F., & Carver, C. S. (1992). Effects of optimism on psychological and physical well-being: Theoretical overview and empirical update. *Cognitive Therapy and Research, 16*, 201–228.

Schenk, E. A., & Moss, A. J. (1966). Cardiovascular effects of sustained norepinephrine infusions. 2: Morphology. *Circulation Research, 18*, 605–615.

Schnall, P. L., Landsbergis, P. A., & Baker, D. (1994). Job strain and cardiovascular disease. *Annual Review of Public Health, 15*, 381–411.

Schwartz, J. E., Warren, K., & Pickering, T. G. (1994). Mood, location and physical position as predictors of ambulatory blood pressure and heart rate: Application of a multi-level random effects model. *Annals of Behavioral Medicine, 16*, 210–220.

Seeman, T. E., Berkman, L. F., Gulanski, B. I., Robbins, R. J., Greenspan, S. L., Charpentier, P. A., & Rowe, J. W. (1995). Self-esteem and neuroendocrine response to challenge: MacArthur studies of successful aging. *Journal of Psychosomatic Research, 39*, 69–84.

Seeman, T. E., & Robbins, R. J. (1994). Aging and hypothalamic-pituitary-adrenal response to challenge in humans. *Endocrine Reviews, 15*, 233–260.

Seraganian, P. (Ed.) (1993). *Exercise Psychology*. New York: Wiley.

Sheps, D. S., Ballenger, M. N., De Gent, G. E., Krittayaphong, R., Dittman, E., Maixner, W., McCartney, W., Golden, R. N., Koch, G., & Light, K. C. (1995). Psychophysical responses to a speech stressor: Correlation of plasma beta-endorphin levels at rest and after psychological stress with thermally measured pain threshold in patients with coronary artery disease. *Journal of the American College of Cardiology, 25*, 1499–1503.

Sherwood, A., Dolan, C. A., & Light, K. C. (1990). Hemodynamics of blood pressure responses during active and passive coping. *Psychophysiology, 27*, 656–668.

Sherwood, A., Royal, S. A., & Light, K. C. (1993). Laboratory reactivity assessment: Effects of casual blood pressure status and choice of task difficulty. *International Journal of Psychophysiology, 14*, 81–95.

Sifneos, P. E. (1973). The prevalence of "alexithymic" characteristics in psychosomatic patients. *Psychotherapy and Psychosomatics, 22*, 55–62.

Solomon, G. F., Levine, S., & Kraft, J. K. (1968). Early experience and immunity. *Nature, 220*, 821–822.

Steptoe, A. (1980). Stress and medical disorders. In S. J. Rachman (Ed.), *Contributions to medical psychology* (Vol. 2, pp. 55–77) Oxford, UK: Pergamon.

Steptoe, A. (1983). Stress, helplessness and control: The implications of laboratory studies. *Journal of Psychosomatic Research, 27*, 361–367.

Steptoe, A. (1984a). Psychophysiological processes in disease. In A. Steptoe & A. Mathews (Eds.), *Health care and human behaviour* (pp. 77–112). London: Academic Press.

Steptoe, A. (1984b). Psychological aspects of bronchial asthma. In S. J. Rachman (Ed.), *Contributions to medical psychology* (Vol. 3, pp. 7–30). Oxford, UK: Pergamon.

Steptoe, A. (1991). Psychobiological processes in the etiology of disease. In P. R. Martin (Ed.), *Handbook of behavior therapy and psychological science* (pp. 325–347). New York: Pergamon.

Steptoe, A. (1997). Behavior and blood pressure: Implications for hypertension. In A. Zanchetti & G. Mancia (Eds.), *Pathophysiology of hypertension* (pp. 674–708). Amsterdam: Elsevier.

Steptoe, A., & Appels, A. (Eds.) (1989). *Stress, personal control and health*. Chichester, UK. Wiley.

Steptoe, A., Roy, M. P., & Evans, O. (1996). Psychosocial influences on ambulatory blood pressure over working and non-working days. *Journal of Psychophysiology, 10*, 218–227.

Steptoe, A., Roy, M. P., Evans, O., & Snashall, D. (1995). Cardiovascular stress reactivity and job strain as determinants of ambulatory blood pressure at work. *Journal of Hypertension, 13*, 201–210.

Steptoe, A., & Tavazzi, L. (1996). The mind and the heart. In D. G. Julian, A. J. Camm, K. M. Fox, R. J. C. Hall, & P. A. Poole-Wilson (Eds.), *Diseases of the heart* (pp. 1430–1441). London: Saunders.

Stone, A. A., Bovbjerg, D. N., Neale, J. M., Napoli, A., Valdimarsdottir, H., Cox, D., Hayden, F. G., & Gwaltney, J. M. (1992). Development of common cold symptoms following experimental rhinovirus infection is related to prior stressful life events. *Behavioral Medicine, 18*, 115–120.

Stone, A. A., Neale, J. M., Cox, D. S., Napoli, A., Valdimarsdottir, H., & Kennedy-Moore, E. (1994). Daily events are associated with a secretory immune response to an oral antigen in men. *Health Psychology, 13*, 440–446.

Stoney, C. M., Owens, J. F., Matthews, K. A., Davis, M. C., & Caggiula, A. (1990). Influences of the normal menstrual cycle on physiologic functioning during behavioral stress. *Psychophysiology, 27*, 125–135.

Strawn, W. B., Bondjers, G., Kaplan, J. R., Manuck, S. B., Schwenke, D. C., Hansson, G. K., Shively, C., A & Clarkson, T. B. (1991). Endothelial dysfunction in response to psychosocial stress in monkeys. *Circulation Research, 68*, 1270–1279.

Suarez, E. C., Williams, R. B., Jr., Kuhn, C. M., Zimmerman, E. H., & Schanberg, S. M. (1991). Biobehavioral basis of coronary-prone behavior in middle-aged men: II. Serum cholesterol, the Type A behavior pattern, and hostility as interactive modulators of physiological reactivity. *Psychosomatic Medicine, 53*, 528–537.

Suls, J., & Wan, C. K. (1993). The relationship between trait hostility and cardiovascular reactivity: A quantitative review and analysis. *Psychophysiology, 30*, 615–626.

Suls, J., Wan, C. K., & Blanchard, E. B. (1994). A multilevel data-analytic approach for evaluation of relationships between daily life stressors and symptomatology: Patients with irritable bowel syndrome. *Health Psychology, 13*, 103–113.

Suomi, S. J. (1991). Early stress and adult emotional reactivity in rhesus monkeys. In *The childhood environment and adult disease* (pp. 171–188). Ciba Foundation Symposium 156. Chichester, UK: Wiley.

Sutton, B. J., & Gould, H. J. (1993). The human IgE network. *Nature, 366*, 421–428.

Tavazzi, L., Zotti, A. M., & Rondanelli, R. (1986). The role of psychologic stress in the genesis of lethal arrhythmias in patients with coronary artery disease. *European Heart Journal, 7*(Suppl. A). 99–106.

Taylor, S. E., Kemeny, M. E., Aspinwall, L. E., Schneider, S. G., Rodriguez, R., & Herbert, M. (1992). Optimism, coping, psychological distress, and high-risk sexual behavior among men at risk for Acquired Immunodeficiency Syndrome (AIDS). *Journal of Personality and Social Psychology, 63,* 460–473.

Terman, G., Shavit, Y., Lewis, J. W., Cannon, J. T., & Liebeskind, J. C. (1984). Intrinsic mechanism of pain inhibition: activation by stress. *Science, 226,* 1270–1277.

Theorell, T. (1992). Prolactin: A hormone that mirrors passiveness in crisis situations. *Integrative Physiological and Behavioral Science, 27,* 32–38.

Thompson, D. G., Richelson, E., & Malagelada, J.-R. (1982). Perturbation of gastric emptying and duodenal motility through the central nervous system. *Gastroenterology, 83,* 1200–1206.

Tidgren, B., & Hjemdahl, P. (1989). Renal responses to mental stress and epinephrine in humans. *American Journal of Physiology, 257,* F682–F689.

Timio, M., Verdecchia, P., Venanzi, S., Gentili, S., Ronconi, M., Francucci, B., Monanari, M., & Bichisao, E. (1988). Age and blood pressure changes: A 20-year follow-up study in nuns in a secluded order. *Hypertension, 12,* 457–461.

Tofler, G. H., Stone, P. H., Maclure, M., Edelman, E., Davis, V. G., Robertson, T., Antman, E. M., & Muller, J. E. (1990). Analysis of possible triggers of acute myocardial infarction (the MILIS study). *American Journal of Cardiology, 66,* 22–27.

Treiber, F. A., Turner, J. R., Davis, S., Thompson, W., Levy, W., & Strong, W. B. (1996). Young children's cardiovascular stress responses predict resting cardiovascular functioning two and a half years later. *Journal of Cardiovascular Risk, 3,* 95–100.

Trevisan, M., Jossa, F., Farinaro, E., Krogh, V., Panico, S., Giumetti, D., & Mancini, M. (1992). Earthquake and coronary heart disease risk factors: A longitudinal study. *American Journal of Epidemiology, 135,* 632–637.

Tsuda, A., Tanaka, M., Nishikawa, T., & Hirai, H. (1983). Effects of coping behavior on gastric lesions in rats as a function of the complexity of coping tasks. *Physiology and Behavior, 30,* 805–808.

Turner-Cobb, J. M., & Steptoe, A. (1996). Psychosocial stress and susceptibility to upper respiratory tract illness in an adult population sample. *Psychosomatic Medicine, 58,* 404–412.

Turner, J., Ward, M., Gellman, M., Johnston, D., Light, K., & van Doornen, L. (1994). The relationship between laboratory and ambulatory cardiovascular activity: Current evidence and future directions. *Annals of Behavioral Medicine, 16,* 12–23.

Turner, R. J., & Marino, F. (1994). Social support and social structure: A descriptive epidemiology. *Journal of Health and Social Behavior, 35,* 193–212.

Turner, R. J., Wheaton, B., & Lloyd, D. A. (1995). The epidemiology of social stress. *American Sociological Review, 60,* 104–125.

Uchino, B. N., Cacioppo, J. T., & Kiecolt-Glaser, J. K. (1996). The relationship between social support and physiological processes: A review with emphasis on underlying mechanisms and implications for health. *Psychological Bulletin, 119,* 488–531.

Van Doornen, L. J. P., & Van Blokland, R. W. (1992). The relationship between cardiovascular and catecholamine reactions to laboratory and real-life stress. *Psychophysiology, 29,* 173–181.

Van Eck, M., Berkhof, H., Nicolson, N., & Sulon, J. (1996). The effects of perceived stress, traits, mood states, and stressful daily events on salivary cortisol. *Psychosomatic Medicine, 58,* 447–458.

Villar, J., Farnot, U., Barros, F., Victora, C., Langer, A., & Belizan, J. M. (1992). A randomized trial of psychosocial support during high-risk pregnancies. *New England Journal of Medicine, 327,* 1266–1271.

Vogele, C., & Steptoe, A. (1993). Anger inhibition and family history as modulators of cardiovascular responses to mental stress in adolescent boys. *Journal of Psychosomatic Research, 37,* 503–514

Von Holst, D. (1972). Renal failure as the cause of death in *Tupaia belangeri* exposed to persistent social stress. *Journal of Comparative Physiology, 78,* 236–273.

Wadsworth, M., Maclean, M., Kuh, D., & Rogers, B. (1990). Children of divorced and separated parents: Summary and review of findings from a long-term follow-up study in the UK. *Family Practice, 7,* 104–109.

Watson, D., & Pennebaker, J. W. (1989). Health complaints, stress, and distress: Exploring the central role of negative affectivity. *Psychological Review, 96,* 234–254.

Webster, A., & Mawer, G. (1989). Seizure frequency and major life events in epilepsy. *Epilepsia, 30,* 162–167.

Weinberger, D., A., Schwartz, G. E., & Davidson, R. J. (1979). Low-anxious, high-anxious, and repressive coping styles: Psychometric patterns and behavioral and physiological responses to stress. *Journal of Abnormal Psychology, 65,* 375–390.

Weiner, H. (1992). *Perturbing the organism: The biology of stressful experience*. Chicago: University of Chicago Press.

Whitworth, J. A., Gordon, D., Andrews, J., & Scoggins, B. A. (1989). The hypertensive effect of synthetic glucocorticoids in man: Role of sodium and volume. *Journal of Hypertension, 7,* 537–549.

Williams, R. B. (1994). Neurobiology, cellular and molecular biology, and psychosomatic medicine. *Psychosomatic Medicine, 56,* 308–315.

Willis, L., Thomas, P., Garry, P. J., & Goodwin, J. S. (1987). A prospective study of response to stressful life events in initially healthy elders. *Journal of Gerontology, 42,* 627–630.

Zakowski, S., Hall, M. H., & Baum, A. (1992). Stress, stress management, and the immune system. *Applied and Preventive Psychology, 1,* 1–3.

Zautra, A. J., Okun, M. A., Robinson, S. E., Lee, D., Roth, S. H., & Emmanual, J. (1989). Life stress and lymphocyte alterations among patients with rheumatoid arthritis. *Health Psychology, 8,* 1–14.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.03
Health Care

JOHN WEINMAN
United Medical and Dental Schools of Guy's and St. Thomas's Hospitals, London, UK

8.03.1 INTRODUCTION 79
8.03.2 THEORETICAL BACKGROUND 80
8.03.3 PERCEIVING SYMPTOMS AND ENTRY INTO THE HEALTH-CARE SYSTEM 81
  8.03.3.1 The Nature and Experience of Symptoms 81
  8.03.3.2 Factors Influencing Symptom Perception 82
    8.03.3.2.1 Attentional focus 82
    8.03.3.2.2 Knowledge and expectations 83
    8.03.3.2.3 Disposition and emotions 83
    8.03.3.2.4 Contextual factors 83
  8.03.3.3 Summary: Symptom Perception 84
  8.03.3.4 Symptom Perception and Help Seeking 84
8.03.4 HEALTH-CARE PROFESSIONAL–PATIENT COMMUNICATION 86
  8.03.4.1 Input Factors in Communication 86
    8.03.4.1.1 Patient input factors 87
    8.03.4.1.2 Health-care professional input factors 89
  8.03.4.2 The Consultation Process 91
  8.03.4.3 Outcomes of Health-care Communication 92
    8.03.4.3.1 Cognitive outcomes 92
    8.03.4.3.2 Affective outcome: patient satisfaction 92
    8.03.4.3.3 Behavioral outcome: adherence 94
    8.03.4.3.4 Consultation outcomes: an overview 98
  8.03.4.4 Improving Health-care Communication 98
8.03.5 HEALTH CARE IN HOSPITALS 100
  8.03.5.1 Psychological Effects of Hospitalization 100
    8.03.5.1.1 Physical and social environment 100
    8.03.5.1.2 Communication in hospital 101
  8.03.5.2 Children in Hospital 102
  8.03.5.3 Psychological Aspects of Specific Hospital Treatments 102
  8.03.5.4 Stressful Medical Procedures in Hospital 104
  8.03.5.5 Psychological Interventions for Stressful Medical Procedures 104
8.03.6 CONCLUSION 106
8.03.7 REFERENCES 107

8.03.1 INTRODUCTION

This chapter is concerned with a range of psychological processes involved in the entry into, experience of, and outcome from health care. It begins by considering how people make sense of and respond to symptoms, leading up to the decision to seek health care. Following this, a large part of the chapter is concerned with the process and outcome from medical consultations. Finally, it considers some very specific health-care settings such as hospital environments and surgical treatments where psychological factors can play an important role in adaptation and recovery.

Thus, the chapter covers the three following broad areas, which form its major subsections: (i) perceiving symptoms and entry into health care, (ii) health-care professional—patient communication, and (iii) health care in hospital.

It is important to note at the outset that this chapter does not deal with the nature of health-care systems and the ways in which these are organized and accessed. Health-care systems vary greatly from country to country and there can be considerable differences in the ease of access to health care, the relative roles of the primary and secondary care sectors, and the ways in which health care is funded and delivered. All these are extremely important contextual factors which can directly influence the experience of health care for the individual but a consideration of their influence is beyond the scope of this chapter. The focus here will be on those psychological processes involved in the entry into and use of health services, and which may be seen across different health-care systems.

8.03.2 THEORETICAL BACKGROUND

Although this chapter focuses on a range of psychological processes and issues, there is a general model which unifies many of the various themes, particularly from the patient's perspective. This is the self-regulatory approach (e.g., Carver & Scheier, 1990; Leventhal et al., 1997). The core concept here is that the individual attempts to make sense of threats to health using preexisting knowledge or schema which give rise to behavioral responses (e.g., taking nonprescribed medicines, deciding to seek medical help, following medical advice, etc.). Thus, many of the individual's behaviors within the health-care setting can be understood from the perspective of their own thoughts and ideas. These are not static but may well change and develop with the experience of new symptoms or information. Self-regulation is therefore a dynamic process in which the individual attempts to preserve the sense of self and to solve the problem of what is happening to their health.

The individual's initial representations of a health problem (e.g., "this stomach pain is a temporary problem brought on by something specific such as over-eating") give rise to specific coping procedures (e.g., not eating, taking antiacids) which are appraised for their effectiveness (see Figure 1). If the appraisal process results in the view that the mode of coping is not working, then another coping procedure may be selected or the individual may change their view about the nature of the problem (e.g., "this stomach pain has lasted for the whole day and has not responded to indigestion medication—it must be something more serious") and their response to it.

The core of the self-regulatory approach is the individual's own understanding or representation of their situation. Work by Leventhal and colleagues (Leventhal et al., 1997; Leventhal, Nerenz, & Steele, 1984) has shown that this representation is built around five distinct but interrelated themes, namely identity, cause, time-line, consequences, and cure/control. On experiencing a new symptom, the individual typically will provide a label or description and possibly search for or link this with other symptoms which they are experiencing. These aspects constitute their perceived identity of the problem, and typically this is linked with a causal explanation, as well as some expectations about how the problem will last (time-line), its likely effects (consequences) and the extent to which it is amenable to cure or control. These representations will influence how the individual responds to the problem in the short-term and the longer-term, if it persists. They will also provide the conceptual framework for providing and making sense of information within consultations with health-care professionals, as well as for evaluating the appropriateness and efficacy of recommended treatment or advice.

In some of the following sections (e.g., symptom perception), this self-regulatory, schema-based approach will be very obvious but in others it will be less clear or less pertinent.

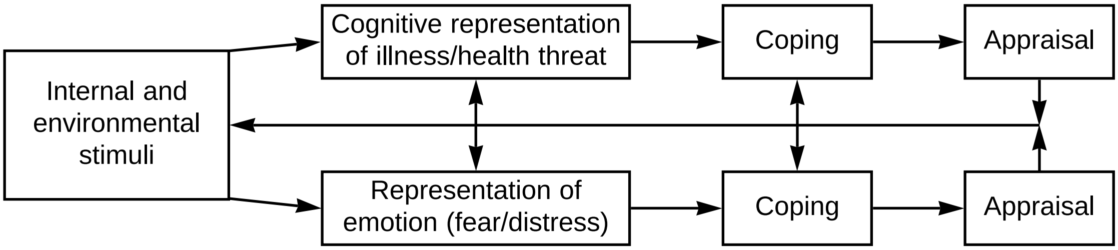
Internal and environmental stimuli
Cognitive representation of illness/health threat
Coping
Appraisal
Representation of emotion (fear/distress)
Coping
Appraisal

Figure 1 Leventhal's self-regulatory model (adapted from Leventhal, Diefenbach, & Leventhal, 1992).

Nevertheless, it is a central tenet in this chapter, just as it is in many other areas of psychological research. Just as individuals actively interpret and make sense of their physical (Eysenck & Keane, 1995) and their social environment (Fiske & Taylor, 1991), it will be shown that the psychological processes involved in experiencing symptoms and utilizing health care can be understood in a similar way (Weinman & Petrie, 1997).

8.03.3 PERCEIVING SYMPTOMS AND ENTRY INTO THE HEALTH-CARE SYSTEM

There are many reasons why people gain access to health care. Increasingly people are encouraged to attend on a regular basis for check-ups or for screening or preventive purposes, but by far the most common reason for seeking medical help is the experience of a symptom. At first sight, the relation between symptoms and entry into health care might appear straightforward. A simplistic model of the relation between symptom experience and health-care utilisation would probably involve the idea of a threshold of symptom severity, and that individuals would only seek care when this notional threshold is exceeded. As we shall see, this simple model is untenable and it is necessary to understand first how symptoms are perceived and then to examine how they influence behavior, including help-seeking behavior.

8.03.3.1 The Nature and Experience of Symptoms

There are continuous changes in bodily functions but individuals have limited awareness of these. In a series of studies, Pennebaker (1982) examined the awareness of a number of changing bodily processes (e.g., heart rate, finger temperature) and showed that this process is neither direct nor accurate The awareness and interpretation of bodily processes has been found to involve a range of factors, as will be outlined below, and an understanding of these is a basic first step in explaining why people seek help for health problems. A fundamental issue here concerns the point at which the individual decides that a particular change in bodily functioning might constitute a symptom and may therefore be indicative of a medical condition. Symptoms are experienced quite frequently. This has been found by asking people to recall symptoms experienced in the preceding days or weeks or by getting them to record their symptoms, by keeping a health or symptom diary. Typically people recall experiencing between two and three symptoms per week and very often symptom diaries reveal the presence of daily symptoms. Most of these are transient and, while they may make the individual wonder about their nature and cause, they may not give rise to any further behavior such as taking medication or seeking medical help. One study, of people keeping health diaries, found that while symptoms were recorded on 38% of the days in the study period, medical care was sought for only 5% of those symptoms (Verbrugge, 1985). This "clinical iceberg" has been reported elsewhere (e.g., Last, 1963) and it is clear that neither the experience nor the severity of a symptom provides an adequate explanation of why people seek medical help. There are many people who seek help for symptoms which doctors consider to be "minor" or "trivial" and there are also people who delay or do not seek help for symptoms which may reflect serious or life-threatening conditions, and some examples of this will be discussed in Section 8.03.3.4. In order to understand the relation between symptoms and help-seeking behavior, it is necessary first to examine how symptoms are perceived and then to examine their role, alongside other factors, in initiating access to health care.

The studies of Pennebaker (1982) and others show clearly that symptom perception shares many characteristics of other perceptual processes, such as those which are involved in the recognition of objects and sounds in the external world. Thus, there is considerable evidence that bodily changes are not perceived accurately and that there can be large variations in what is attended to and how it is perceived and interpreted. Just as attentional processes and "top-down" influences in the form of knowledge and prior experience can determine how objects are perceived (Eysenck & Keane, 1995), the same factors play a critical role in influencing our awareness of and interpretation of bodily changes.

Physical symptoms are often ambiguous and their interpretation involves a number of cognitive-perceptual processes which are subject to both psychological and social influences (Cioffi, 1991; Pennebaker, 1982). The experience of a bodily symptom initiates an active memory search in order to generate comparisons between current stimulus and the concrete and abstract knowledge contained in the person's illness schemas. These cognitive structures are the dominant source of information for the appraisal of health threats (Leventhal et al., 1997). According to Bishop (1991), the information contained in illness schemas forms a stable knowledge base or prototype which acts as a standard for the comparison and interpretation of bodily symptoms. Current symptoms are compared with prototypical patterns of illness retrieved from long-term memory. If there is a close enough resemblance between the illness prototype and the current symptoms, then a congruent interpretation will be made. However, H. Leventhal and Diefenbach (1991) propose that comparisons are made not with illness prototypes but with specific illness episodes. They suggest that matching to prototypes is likely to result in an ambiguous outcome and that the resulting uncertainty will initiate additional, more specific searches. Croyle and Barger (1993) comment that both types of matching may occur. Frequently experienced illnesses such as flu may promote the formation of prototypes, whereas unusual symptom clusters may give rise to matching with specific illness episodes. Cioffi (1991) has described a model of symptom interpretation which is compatible with the self-regulation model. Her model proposes that symptom interpretation is a function of interactions between "bottom-up," stimulus-driven processes, situational cues, and "top-down" processes which reflect the influence of higher level cognitive structures (e.g., knowledge, expectations) on other cognitive functions such as attention and memory. According to Cioffi (1991), ambiguous physiological changes compete with other events for a share of our fixed attentional capacity. If the stimuli are strong enough, or if there is little competition from other sources, the sensations are noticed and given a somatic label. An attribution is then made regarding the cause of the change and its consequences. Causal attributions reflect the relative influence of the procedures involved in the interpretative process and may be symptomological (e.g., muscle pain as a symptom of illness), or nonsymptomological (e.g., muscle pain due to exercise). Symptomological attributions can reflect preexisting hypotheses such as particular health concerns which may bias both attention to bodily stimuli and choice of attribution. Alternatively, in the absence of a prior hypothesis, a memory search is launched in order to generate plausible attributions. These processes, along with subsequent behavioral responses, are mediated by factors such as mood, coping repertoire, available choices, and both situation-specific and general goals. A schematic of this model is shown in Figure 2.

8.03.3.2 Factors Influencing Symptom Perception

The appraisal of health threats may be influenced by a range of personal, psychological, social, and cultural factors (Croyle, 1992). While many of these factors interact, they can be loosely categorized into those which affect or are a part of the self-system and those which stem from external, social, and cultural influences.

8.03.3.2.1 Attentional focus

Symptom perception is subject to the same limits and biases as other types of sensory processing, where there are well-demonstrated limits in attentional capacity which can be influenced by internal states, perceiver's intentions, and environmental stimuli and contexts (e.g., Kinchla, 1992). Similarly, a number of studies have shown that symptom perception depends on what the individual is attending to as well as what is happening in the perceiver's environment. For example, Pennebaker (1982) reports a study in which the symptom reports of fatigue were compared in two groups of joggers, one running around a track and the other on a cross-country run. The cross-country group report less fatigue, presumably because of the more varied and interesting surroundings. Similarly it has been reported that, while watching films, people are more likely to notice itchy/tickly throat sensations and then cough during boring parts of the film.

Attention to symptoms is to some extent under strategic control and a number of studies have examined the effects of attention to or distraction from physical sensations (Cioffi, 1991). Experimental studies examining the effects of various unpleasant stimuli (e.g., loud noise, pain induced by immersing one's hand into iced water) on tolerance and arousal, have shown that active distraction can have short-term advantages over more direct attention to the stimuli (Mullen & Suls, 1982). However, focused attention results in more favorable outcomes than distraction when the noxious stimulus is chronic and when the attentional strategy involves focusing on concrete aspects of the sensation (Suls & Fletcher, 1985). This raises the question as to whether focusing on aversive symptoms such as pain could have any positive or negative effects which could be utilized in the clinical situation. Work by Leventhal, Leventhal, Shacham, and Easterling (1989) indicates that focusing on sensory properties of sensations can result in reduced levels of distress. This type of sensory focusing or monitoring serves to direct attention to more objective, sensory properties of the stimulus and away from the more negative, emotionally arousing aspects, resulting in a more neutral interpretation and response. This can be used to good effect in providing sensory information to patients prior to stressful medical procedures and surgery (see Section 8.03.5.5).

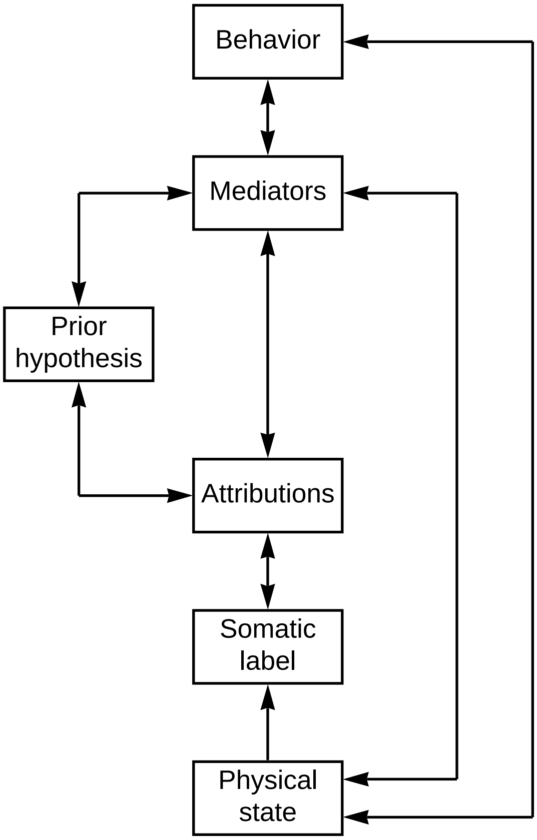

Figure 2 A model of symptom perception (adapted from Cioffi, 1991).

8.03.3.2.2 Knowledge and expectations

The knowledge contained within illness schemas imposes meaning on internal and external events (Leventhal et al., 1984). A powerful illustration of the influence of schema contents on the interpretation of somatic stimuli comes from studies of patients suffering from panic disorder and hypochondriasis whose catastrophic illness-related interpretations of benign bodily stimuli reflects their current health concerns (Warwick & Salkovskis, 1990). Similar influences of catastrophic interpretations on symptoms can be seen in patients with chronic fatigue syndrome (Moss-Morris, 1997) and with chronic pain disorders (e.g., Keefe, Brown, Wallston, & Caldwell, 1989). In all these situations the catastrophizing individual will be likely to interpret a bodily symptom in a very negative way (i.e., perceive it as indicative of a serious problem), which then will generate even higher levels of anxiety and an increased focus on and negative interpretation of related symptoms.

In addition to personal illness experience, illness schemas are derived from illness information current in the culture and from social communication with other people, particularly medical practitioners (Leventhal et al., 1984). The body or self schema also affects the interpretation of health threats as it provides a basis for comparing deviations from the norm and for discriminating normal age-related changes from illness (Keller, Leventhal, Prohasha, & Leventhal, 1989). Indeed, with older people there is evidence of a bias towards age-related attributions (i.e., "it's because I am old") which can serve to delay seeking help or prevent taking appropriate action (E. Leventhal & Crouch, 1997).

8.03.3.2.3 Disposition and emotions

The content and activation of illness schemas may also be influenced by dispositional factors such as psychological traits and emotions (H. Leventhal & Diefenbach, 1991; Leventhal, Diefenbach, & Leventhal, 1992). For example, Leventhal et al. (1992) propose that personality factors such as repression-sensitisation (Byrne, 1964) may affect illness representations by influencing attention to and interpretation of somatic states. Also, high levels of negative affectivity, which is highly correlated with neuroticism (Watson & Clark, 1984), have been found to predict symptom reporting and health complaints, but not actual health status (Watson & Pennebaker, 1989). Laboratory studies involving experimental induction of a negative mood state in volunteer subjects results in more negative judgements of health status and higher levels of symptom reporting than is reported by comparison subjects who had undergone positive mood induction (Croyle & Uretsky, 1987). Similarly, when individuals are asked to keep a diary of symptoms and moods, they report higher levels of symptoms during negative mood states (Verbrugge, 1985).

8.03.3.2.4 Contextual factors

Social information and the environmental and cultural context also affect symptom perception (Croyle & Barger, 1991). When intrapersonal memory searches aimed at establishing the meaning of symptoms fail to provide an answer, social information becomes an important factor. In particular, the social comparison process, in which illness information is shared within the social network, has a significant influence on symptom appraisal (Leventhal & Diefenbach, 1991). Medical sociologists describe the way people use "lay referral" networks in trying to make sense their symptoms. These reflect the use of advice given by family and friends, who may often refer the individual to others who are known to have knowledge or experience which could help in throwing some light on the nature of the problem (see Armstrong, 1989).

Situational and contextual cues can play an important role in interpreting symptoms and in the generation of illness representations. For example, Baumann, Cameron, Zimmermann, and Leventhal (1989) found that, in the presence of a stressful life event, participants were more likely to interpret ambiguous symptoms as signs of stress than signs of illness. However, in the absence of stress, an illness interpretation was more likely. While external factors affected judgements about ambiguous symptoms, the effect was constrained by prior knowledge since context did not influence judgements about symptoms representing familiar illnesses.

8.03.3.3 Summary: Symptom Perception

The interpretation and response to symptoms is guided by illness schemas and involves several cognitive procedures. Typically, it begins with a search of long-term memory for comparable illness episodes, which promotes comparisons between the current state and previous experience. The outcome of the comparison process influences judgements about illness status and the identity of symptom clusters. These judgements can be mediated by the self-system and by external factors. The influence of the self-system reflects personal illness experience, the cultural knowledge base, the body schema, and personality traits. Symptom appraisal may also be biased by motivational processes aimed at minimizing health threats (Croyle, Sun, & Hart, 1997). For example, in an experimental study, Croyle, Sun, and Louie (1993) found that students who had been told that they had a potentially risky cholesterol level actually rated high cholesterol as less serious than those who were informed that theirs was in the normal range. Contextual and situational cues may also modify symptom interpretation, by the provision of specific local information (e.g., other people with similar symptoms) or via selective attention to specific aspects of the symptom experience (Cioffi, 1991).

8.03.3.4 Symptom Perception and Help Seeking

From a population perspective, the experience of symptoms is widespread and it is clear that the majority of these are transient and benign. For the individual, each symptom is evaluated in terms of its potential threat and decisions will be made on how best to respond on the basis of past experience and the present context. Thus, the perceived identity and cause of the symptom and its anticipated timeline and consequences will give rise to particular health-related behaviors. The most common initial response is to anticipate that the symptoms will be short-lived and can be attributed to a specific situational factor but there are important individual and cultural differences in this respect (Robbins & Kirmayer, 1991; Kirmayer, Young, & Robbins, 1994). These authors have carried out studies in which they presented various imaginary symptoms (e.g., dry mouth) to participants and asked them to choose the most likely cause from three alternatives, of which one was situational ("I need to drink more liquids"), one of which was physical ("there is something wrong with my salivary glands"), and the other was psychological ("I must be scared or anxious about something"). Situational explanations were chosen most frequently but some individuals showed clear biases towards selecting more physical or psychological explanations.

A bias towards opting for physical explanations for symptoms could lead to potential conflicts in health care, when tests fail to reveal any organic cause for a problem. To some extent this can be observed in people with chronic fatigue syndrome since they commonly attribute their symptoms to a physical cause in the absence of any confirmatory evidence (Moss-Morris, 1997). This, in turn, may reduce the perceived relevance of any psychological treatments which may be offered and have adverse effects on treatment adherence (see Section 8.03.4.3.3).

Seeking medical help is a relatively uncommon response since the majority of individuals wait, do nothing, or self-medicate (Freer, 1980). However, there has been considerable interest in understanding the responses of two groups of individuals: those who persistently seek help for what appear to be quite minor symptoms and those who delay seeking help while experiencing serious or life-threatening symptoms. Studies of frequently attending patients reveal that they are typically high in trait anxiety (Banks, Beresford, Morrell, Waller, & Watkins, 1975). This is consistent with the finding that trait anxiety is associated with a stronger attentional focus on internal bodily states, as well as higher levels of symptom reporting (Watson & Pennebaker, 1989). Also, these individuals are more likely to make more negative interpretations and attributions for particular symptoms than their less anxious counterparts (Sensky, Macleod, & Rigby, 1996). Given the negative cognitions which accompany anxiety (Lucock, Morley, Peaks, & White, in press), it is therefore not surprising that there will be greater concern and an accompanying need for reassurance but, as is discussed more fully at the end of this section, the rather minimal reassurance which is very often provided by the doctor is not likely to be effective in reducing anxiety or preventing further help-seeking.

Delay in seeking help for serious symptoms is a complex process which depends critically on the individual's perception of their symptoms as well as contextual factors, such as barriers to health care. Three stages of delay can be identified and have been described by Safer, Tharps, Jackson, and Leventhal (1979) as appraisal delay, illness delay, and utilisation delay. Appraisal delay involves the decision as to whether one is ill, given the particular pattern of symptoms which are being experienced. This stage is very much a function of the sensory properties of the symptoms as well as the way in which these are interpreted either from past experience, from reading about one's symptoms, or from discussion with others. The appraisal stage can be understood easily in terms of the models of Leventhal et al. (1997) and Cioffi (1991) which were outlined earlier in the chapter, since symptoms which do not activate threat-related illness schema are not interpreted as serious or potentially serious and are very unlikely to instigate help-seeking. Illness delay describes the delay between recognising that there is an illness and deciding that professional help should be sought, whilst the time taken from that point to actually seeking help is referred to utilisation delay. Different factors contribute to the various stages of delay, including past experience of symptoms, their perceived consequences and concerns about the possible costs, and negative aspects of the treatment and outcome. One condition where delay can have very serious consequences is myocardial infarction (MI), since early utilization of medical care is associated with improved changes of survival. Death following MI typically occurs within a few hours of symptom onset, yet the evidence shows there is a large variation in utilization delay times and that some individuals put themselves at much greater risk as a result (Dracup et al., 1995). This work also shows that symptom severity and interpretation are important since those who perceive their symptoms as indicative of MI are less likely to delay. Problems caused by patient delay over a longer time period are also found with other conditions, including cancer (Facione, 1993), sexually transmitted diseases (Leenars, Rombouts, & Kok, 1993), and mental health problems (Clausen, Pfeffer, & Huffine, 1982).

From the health-care professional's (HCP) perspective it is therefore important to know what symptoms the individual has experienced and the way that these have been interpreted as a basis for understanding the decision to seek help. In view of this, it is noteworthy that guidelines for good patient interviewing in primary care include an encouragement to focus on why the patient has come (Pendleton, Schofield, Tate, & Havelock, 1984). Thus, in addition to eliciting signs and symptoms, doctors are expected to develop an interviewing approach which allows patients to express their own ideas or concerns about their condition. From the HCP's perspective, patients may appear to be seeking help for relatively minor symptoms (e.g., transient pain from indigestion) whereas specific contextual factors (e.g., a close relative with a heart problem) may cause the patient to hold a much more serious representation of their problem.

A further implication for HCP's lies in the issue of reassurance. As will be outlined in Section 8.03.4.1.1, one of the most common patient expectations for a primary care consultation is to be able to have a better understanding of a current complaint (Williams, Weinman, Dale, & Newman, 1995). Since the majority of symptoms will not be found to be indicative of any underlying problem, as the result of a physical examination or diagnostic test, it is not surprising that the HCP will expect this to be sufficient to reassure the patient. For some patients being told that there is no underlying or serious problem is effective in providing reassurance but there is consistent evidence that this is not always so. For a significant minority, there is continued concern about their health status (Channer, James, Papouchado, & Rees, 1987; Lucock et al., in press).

This continued anxiety following reassurance from the HCP may then result in further needless consultations or investigations, and is one of the factors involved in frequent attendance for health care. The key issue here seems to be the nature of the reassuring message which typically consists of reporting that no pathology can be detected and so there is no need to worry (McDonald, Daly, Jelinek, Panetta, & Gutman, 1996). Although apparently comforting from the HCP's perspective, this message only provides negative information and fails to contribute a positive explanation for the symptoms which the patient had or is experiencing. According to Lucock et al. (in press), reassurance in the form of negative information will only provide a short-term reduction of health concern. If symptoms persist or reoccur, then the health concern will also return since the patient still lacks a satisfactory explanation which would enable them to interpret the symptoms as benign. For reassurance to be effective, patients' concerns need to be elicited and appropriate information provided for explaining the symptoms. These findings can be understood in terms of Leventhal's self-regulatory model, which was outlined at the beginning of this chapter. If the patient's cognitive representation of a problem is that it is serious (e.g., heart disease) and the diagnostic test fails to provide any evidence for this, then an adequate alternative explanation and discussion of the symptoms will probably be necessary for the patient to generate a more benign representation and therefore be reassured by the test results.

8.03.4 HEALTH-CARE PROFESSIONAL–PATIENT COMMUNICATION

The consultation between the patient and the HCP lies at the centre of health care. The information which is transmitted during the consultation is very often critical in the formulation of diagnoses and in the organisation of treatment. Effective communication is necessary to ensure not only that the patients' problems and concerns are understood by the HCP but also that relevant information, advice, and treatment is received and acted upon by the patient. Since the late 1960s there has been considerable research on the medical consultation, prompted to a large degree by the fact that there has been consistent evidence that not only are the process and outcome often unsatisfactory for patients but also there is widespread noncompliance with subsequent treatment recommendations. Early research revealed quite high levels of patient dissatisfaction which was often associated with insufficient information, poor understanding of the medical advice, and subsequent reluctance or inability to follow recommended treatment or advice (e.g., Korsch & Negrete, 1972). These early findings showed that many patients complained that they were not given sufficient information about the nature of their problem, the treatment, and the likely outcome. For example, in the Korsch and Negrete (1972) study, which involved mothers taking their children to a paediatric outpatient clinic, about 20% of mothers were not informed clearly about the diagnosis and nearly 50% were uncertain afterwards as to the course of their child's illness. Similar findings have been produced in more recent studies in a variety of health-care settings and it is notable that these problems are unrelated to the length of time spent with the HCP (Korsch & Negrete, 1972).

A related problem involves the use of information, such as medical jargon, by the HCP which the patient misconstrues or cannot understand. Korsch and Negrete (1972) found that jargon was used in about half the consultations in their study. Other studies have shown how much this "communication gap" reflects differences in biomedical knowledge between patients and HCPs. A good example is seen in Figure 3 which shows some data from a study by Hawkes (1974) who investigated patients' and doctors' interpretations of anatomical terms. Not only did this study show that patients' understanding of anatomical terms is often quite different from that of their doctor but also that doctors may also differ in their use of some terms. Another complaint from patients is that the HCP does not appear to be interested in their presenting problem or in their concerns or worries about it and this is associated with lower levels of satisfaction with the consultation (Williams & Calnan, 1991). Moreover, patients often report feeling that they have not been sufficiently involved in the decision-making which may have occurred in the consultation. Hence, Roter and Hall's (1989) overview of research in this area revealed that patient satisfaction was higher following consultations in which the HCP engaged in more social conversation, positive verbal and nonverbal behavior, and partnership building.

The development of relatively unobtrusive audio and video-recording techniques has allowed researchers to investigate the consultation and many studies have analyzed the process of the consultation and attempted to relate process variables or characteristics to outcome. While these studies have identified important themes, they have not always been successful in making clear links between process and outcome (Stiles, 1989). In the absence of simple relations between consultation process variables (e.g., duration or style of consultation) and various outcomes (e.g., satisfaction, adherence, etc.), more recent studies have begun to examine what HCPs and patients bring to the consultation, as well as the importance of contextual factors. As a result, many current frameworks for understanding doctor-patient communication tend to be based on the relations between inputs (i.e., the attitudes, beliefs, expectations, etc. which patient and doctor bring to the consultation), process (the nature of the encounter), and outcome (the short- and longer-term effects on the patient). An example of one such framework is provided by Friederikson (1993) and is shown in Figure 4. An overview of input, process, and outcome factors follows.

8.03.4.1 Input Factors in Communication

Input factors which influence the consultation include not only aspects of the doctor and

In the picture below put a tick underneath the drawing that shows the correct position of the brain

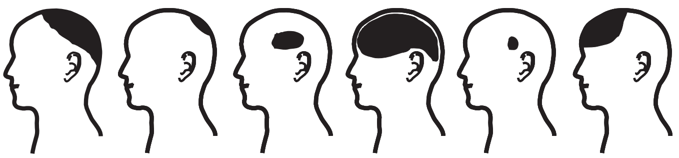

							Total	Rating
Doctors	0	0	0	53 (100%)	0	0	53	Good
Patients	11 (5.9%)	9 (4.8%)	17 (9.2%)	122 (66.3%)	4 (2.1%)	21 (11.4%)	184	Fair

In the picture below put a tick underneath the drawing that shows the correct position of the sciatic nerve

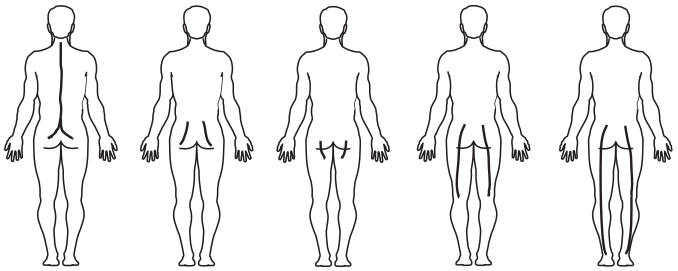

						Total	Rating
Doctors	0	1 (1.8%)	2 (3.7%)	29 (54.7%)	21 (39.6%)	53	Fair
Patients	37 (22.8%)	25 (15.4%)	8 (4.9%)	36 (22.2%)	56 (34.5%)	162	Poor

Figure 3 Examples of the differences in the use of anatomical terms by patients and doctors (from Hawkes, 1974).

patient but also the context and setting in which the consultation occurs. For example, for many primary care consultations in the UK, patients are booked in for 10-minute appointments whereas similar consultations in other countries may typically last two to three times longer. Although longer consultations do not inevitably result in better patient outcomes (Morrell, Evans, Morris, & Roland, 1986), the resulting process may well be different. Similarly, the layout of the consulting room, including the proximity and positioning of the HCP and patient, can also influence the communication process.

8.03.4.1.1 Patient input factors

When considering patient "input" factors which can influence the medical consultation, a number of studies have shown that patients cope with health threats in diverse ways and show consistent differences in how they want to be involved in the health-care process (Krantz, Baum, & Wideman, 1980). These studies also show that patients differ in the amount of information which they would like to receive about their health problem. Similarly, a distinction has been made by Miller, Brody, and Summerton (1988) between "monitors" and "blunters," with the former being more inclined to need and seek out information about their problem and treatment, whereas the latter group prefer consultations in which relatively limited information is provided. This is similar to the repression-sensitisation classification, described in Section 8.03.3.2.3, and reflects the extent to which patients choose to cope with their health problems in a problem-focused or an avoidant manner.

Patients come into the health-care setting with different levels of biomedical knowledge, based on their past experience. There is also consistent evidence that patients have differing expectations for specific consultations (Williams et al.,

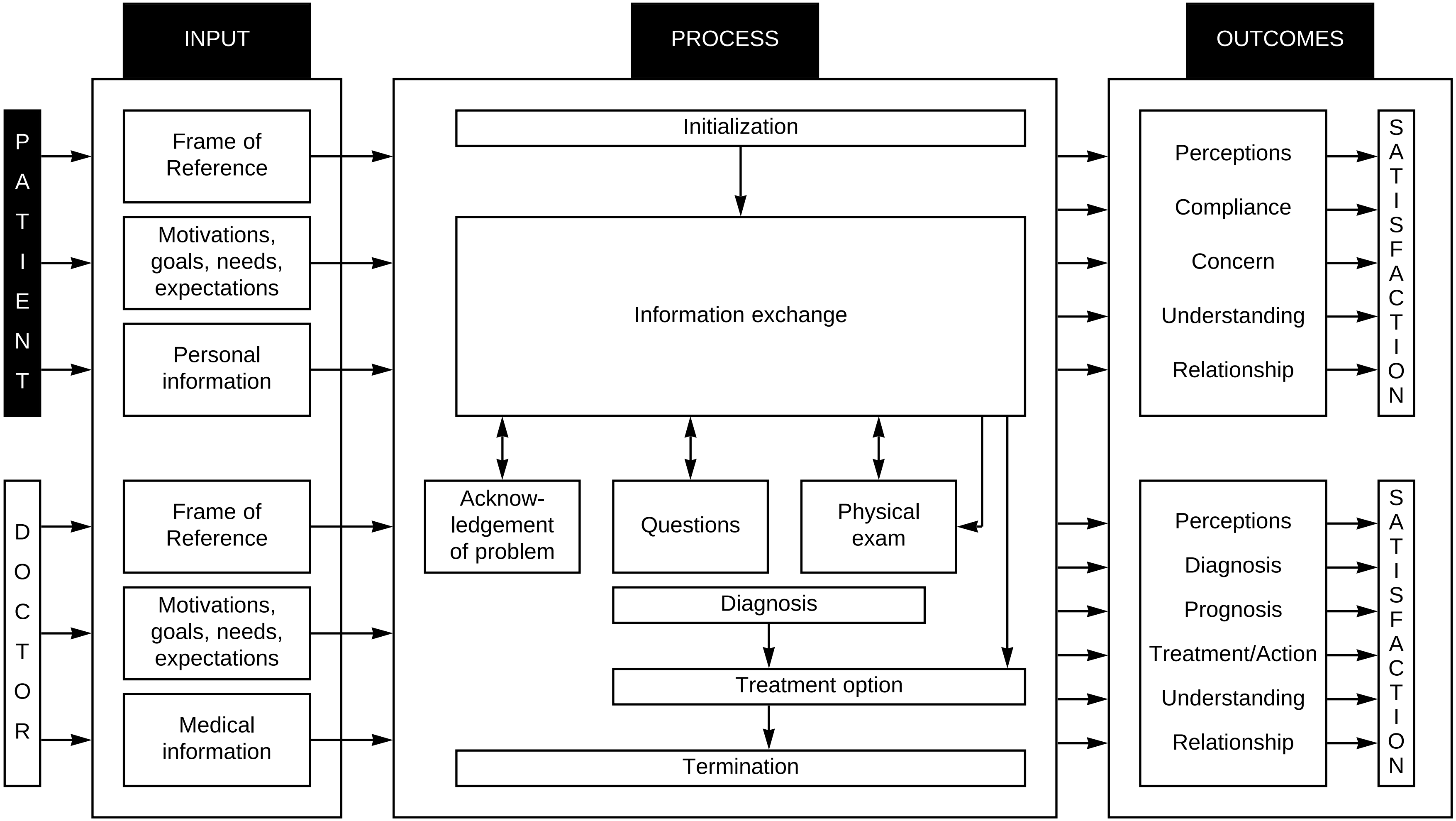
INPUT
PROCESS
OUTCOMES
PATIENT
Frame of Reference
Motivations, goals, needs, expectations
Personal information
DOCTOR
Frame of Reference
Motivations, goals, needs, expectations
Medical information
Initialization
Information exchange
Acknow-ledgement of problem
Questions
Physical exam
Diagnosis
Treatment option
Termination
Perceptions
Compliance
Concern
Understanding
Relationship
SATISFACTION
Perceptions
Diagnosis
Prognosis
Treatment/Action
Understanding
Relationship
SATISFACTION

Figure 4 Information processing model of medical consultation (adapted from Friedrikson, 1993).

1995), and an awareness of these can be helpful in understanding not only why they are seeking help at that time but also in being able to respond effectively to their needs. Following the self-regulatory model, described at the beginning of this chapter, it has been found that patients' illness representations will influence their decision to seek help (Cameron, Leventhal, & Leventhal, 1993) and hence their expectations of the consultation.

Patient expectations have been conceptualized and defined in a range of different ways. In their review of the literature, Thompson and Sunol (1995) have proposed working definitions for four types of expectations that are used frequently:

(i) *Ideal*. This may be referred to as an aspiration or desire;

(ii) *Predicted*. The realistic, practical, or anticipated outcome which therefore reflects what users actually believe will happen in an health care consultation;

(iii) *Normative*. What should or ought to happen and could be equated with what users are told, or led to believe, or personally deduce that they ought to receive from health services; and

(iv) *Unformed*. This state occurs when users are unable or unwilling to define their expectations, which may be because they do not have any, or find it too difficult to express, or do not wish to reveal their feelings, due to fear anxiety, conformity to social norms, etc.

The term "expectations" itself therefore needs to be clearly defined and one reason for the lack of consistency in the research in this area is that many studies have failed to do this. An example of this problem can be seen in a study carried out to investigate the relationship between patients' expectations and HCPs' actions (Webb & Lloyd, 1994). Primary care patients' expectations were measured using questionnaires which asked each patient "How do you think the doctor will be able to help you with your problem?" and instructed the subjects to tick as many of the following actions: "give you a prescription," "refer you to hospital," "give you advice," "help you in some other way." By this method it would not be clear to the patients whether they should respond according to what they themselves want (i.e., "ideal" expectations) or what they merely expect that they will be given by the doctor (i.e., "predicted" or "normative" expectations).

One method of measuring patients' expectations is to categorize their informational needs into distinct categories such as information in relation to aetiology, diagnosis, prognosis, treatment, and social effects (Kindelan & Kent, 1987). However, the measurement of patients' expectations by the simple categorization of needs according to the medical condition can be inadequate since patients may require varying levels of information about the presenting symptoms. The range of psychosocial and medical expectations of primary care patients has been investigated by Good, Good, and Nassi (1983). They have developed a scale (Primary Care Patient Request Scale) from the responses of North American primary care patients and it requires patients to rate the extent to which each item on the scale reflects their needs prior to the consultation. The results of a principal components analysis to investigate underlying components on the types of consultation requests of primary care patients revealed five distinct dimensions: treatment of psychosocial problems, medical explanation, supportive communication, test results, and ventilation and legitimation. A simpler version of this scale, developed by Salmon and Quine (1989), measures four components: explanation and understanding, support, medical treatment, and information-seeking. Using this scale, Williams et al. (1995) have shown that the most common expectation in primary care patients is for explanation and understanding of their problem, with much smaller numbers wanting tests, diagnoses, or support. Thus,patients do not always want or expect diagnosis or treatment since they may be looking to the consultation to gain more understanding of their health problem or may be hoping for support or understanding from their doctor. These prior expectations can be important in determining outcomes, since consultations in which patient expectations are met have been shown to result in greater satisfaction and an increased willingness to follow advice or treatment (Williams et al.).

8.03.4.1.2 Health-care professional input factors

It is important to acknowledge that health-care professionals can vary considerably in the attitudes and beliefs which they have not only about their own and the patient's role, but also about the function and conduct of the consultation. Doctors have been categorized in various ways according to their role perceptions and the extent to which they concentrate on the technical or more psychosocial aspects of patient care, as well as their beliefs about whether patients should be actively involved in the consultation and in decision-making about the management of the clinical problem (e.g., Grol, de Maeseneer, Whitfield, & Mokkunk, 1990). Inevitably, these broad attitudinal differences are reflected in differences in the way in which the consultation is conducted and in other aspects of professional behavior, including decision-making, prescribing, and management of clinical problems (McGee, 1997). Similarly, other HCPs have also been found to vary in their beliefs and attitudes in ways which impact on the way they manage their encounters with patients (Marteau & Johnston, 1990). In clinical psychology, one very obvious HCP input factor is the theoretical orientation of the practitioner, since this can have a major influence on the way in which clients' problems are conceptualized and managed. It has also been shown that personal characteristics of the practitioner can have significant effects on the process and outcome of psychological therapy (e.g., Strupp & Hadley, 1979).

A few studies have attempted to examine the interaction of clients' and therapists' values in psychological therapies (e.g., Kelly, 1990). There is evidence that strong initial differences between client and therapist may have significant negative effects on the early stages of therapy (Kelly & Strupp, 1992) and may be a contributory factor to client dropout (Vervaeke, Vertommen, & Storms, 1997).

Although HCPs undergo a common training and share a common body of knowledge and skills, wide variations in their clinical behaviour have been noted (e.g., Marteau & Johnston, 1990). Part of this variation is a by-product of personality differences and can be seen in differences in interpersonal aspects of clinical practice but part is a reflection of differences in the approach to clinical problems. Some of these are between-individual differences, whereas others reflect factors which can vary within the same individual, such as changing mood, time pressure, and various other contextual influences. In their overview of clinical decision-making, Schwartz and Griffin (1986) point out that there are often substantial disagreements between doctors when interpreting the same clinical information (e.g., X-rays). They also note that doctors may be inconsistent and disagree with their own previous judgements.

Clinical decision-making has been the subject of psychological research for a number of years and there is accumulating evidence of inter- and intra-individual differences in both the process and the outcome (Schwartz & Griffin, 1986). A detailed account of medical decision-making is beyond the scope of this chapter and so what follows is a brief outline of the process together with a focus on some factors which are associated with variation in decision-making performance.

In responding to a new patient, the doctor will first try to solve the problem of what is wrong, based on the various signs and symptoms, and then will have to make a decision about the best approach to treatment. Both these processes occur under conditions of uncertainty and, for much of the time, the doctor's thinking will be based on the use of probabilistic information. The signs and symptoms may be ambiguous and there is often a margin of error in test results. Also, treatment decisions, whether they involve waiting to see what will happen or the prescription of treatment, can involve risks of complications or side-effects.

During clinical decision-making, the doctor is often faced with having to process a large amount of information about the patient and the possible diagnoses and treatment options. In this they are subject to the same limits of information-processing capacity which affect all of us (see Eysenck & Keane, 1995). Thus, there are limits in what can be attended to and how much information can be held in working memory. In view of this it is not surprising to find that all doctors acquire heuristics, or general strategies, for processing diagnostic and treatment information and these are subject to a range of influences.

Kahnemann, Slovic, and Tversky (1982) have described a number of the heuristics which are used during clinical decision-making. These include the representativeness heuristic (decisions based on the similarity between the present situation and known previous ones) and the availability heuristic (judgments made on the basis of information which is most readily available). The influence of representativeness heuristics can be seen in the initial hypotheses which doctors generate for newly presented symptoms and availability heuristics can be inferred from the influence of such factors as recent or emotionally salient events which readily come to the doctor's mind during the clinical judgment process. Christensen-Szalanski, Beck, Christensen-Szalanski, and Keopsell (1983) have demonstrated an availability bias in doctors' estimates of the risk of different diseases since their overestimates of risks were found to be biased by encounters with people with the disease. Similarly, attitudes and beliefs about the doctor's and the patient's role can influence the weighting given to information from patients (e.g., the relative attention given to physical and psychosocial information). The doctor's mood can affect information-processing in a number of ways by directly influencing the speed and accuracy of clinical problem-solving (Isen, Rosenzweig, & Young, 1991) or possibly by constraining access to mood congruent semantic information, as has been demonstrated on more general studies of the effect of mood on cognition (e.g., Power & Dagleish, 1997).

8.03.4.2 The Consultation Process

There are a range of methods and frameworks for analyzing and describing the process of the consultation. One of the broadest distinctions made has been between consultations which are described as patient-centered and those which are HCP-centered, reflecting the extent to which the HCP or patient determines what is discussed (Grol et al., 1990). HCP-centered consultations are ones in which closed questions are used more often and the direction is determined by the HCP, typically with a primary focus on medical problems. In contrast, patient-centered encounters involve more open-ended questions with greater scope for patients to raise their own concerns and agendas. Related to this are consistent differences in the extent to which the HCP responds to the emotional agendas and the nonverbal cues of the patient (see Roter & Hall, 1989). Although there has been a tendency to consider the more patient-centered/emotion-focused approach as preferable, what appears to be more important is for HCP and patient to be in agreement over the nature of the problem and the best course of action (Starfield et al., 1981).

A number of specific methods have been developed for carrying out detailed analyses of the social interaction between HCP and patient based on audio or videotapes or transcripts of the consultations but, as yet, there is no real consensus as to the best method for this. One of the earliest of these was the Bales's (1950) process analysis system which distinguishes verbal statements into those which are task or emotion-focused and then into more specific categories and this approach has been adapted and extended by a number of other investigators (see Roter & Hall, 1989).

With Bales's system, the coding can be carried out using transcripts, audiotapes, or first-hand observations (Bales, 1950; Korsch & Negrete, 1972). Verbal behavior is usually studied, although nonverbal behavior may also be coded. A unit can be as short as a single word or as long as a lengthy sentence; compound sentences are usually divided at the conjunction and sentence clauses are scored as separate units when they convey a single item of thought or behavior. The rationale underlying this approach is that all interpersonal statements can be classified as falling into one of two domains, the task and the socioemotional. Interaction is described in terms of 12 mutually exclusive categories: six "effectively neutral" in the task dimension (e.g., doctor gives suggestion) and six equally divided into positive and negative affective categories in the socioemotional dimension (e.g., doctor agrees or disagrees).

Wasserman and Inui (1983) have noted that a strength of Bales's system is that it deals with both the content and relationship levels of the communication and that it is particularly strong on the relationship level, where it is sensitive to the feelings of each communicator. Moreover, it has been shown to be reliable and applicable in diverse situations. Its main weakness is the treatment of information transfer, that is, the transfer of information from the patient to the doctor as well as in the reverse direction. With this system there is a difficulty in categorising statements that involve both information transfer and affective components. As a result, a number of investigators, including Roter (1977), have modified the system to allow direct coding from audiotapes, assessment of the affective aspects of the clinician's behavior, and a more finely tailored categorisation of the clinical encounter.

An alternative approach, known as the Verbal Response Mode (VRM), has been developed by Stiles (1978) and which classifies each statement made by the HCP and patient into one of eight basic categories (e.g., questions, giving interpretations, etc.). Good overviews of these different approaches are available elsewhere (e.g., Roter & Hall, 1989) and attempts have been made to define a number of more general ways of classifying HCP-patient interactions. For example, one can distinguish between verbal and nonverbal information and, within the verbal domain, six broad categories can be defined: information-giving, information-seeking, social conversation, positive talk, negative talk, and partnership building. From a meta-analysis of these broad categories (Roter, 1989) it has been found that for the doctor, information-giving occurs most frequently (approximately 35% of the doctor's communication) followed by information-seeking (approximately 22%), positive talk (15%), partnership building (10%), social conversation (6%), and negative talk (1%). In contrast, the main type of patient communication consists of information giving (approximately 50%) with less than 10% involving information seeking.

A comparative study of process analysis methods was carried out by Inui, Carter, Kukull, and Haigh (1982). The Bales, Roter, and the VRM system were applied to 101 new patient consultations at a general medical clinic. The outcome measures were patient knowledge, satisfaction, recall of prescribed medications, and compliance. The findings revealed that the explanatory power of the three systems differed. Roter's system was found to explain 28% of the variance in adherence to prescribed medicines, compared to 19% for the Bales system and none for the VRM. The Roter and Bales systems were also better than VRM in explaining variation in satisfaction.

More recently, Roter et al. (1997) have used cluster analysis with the Roter Interactional Analysis System (RIAS) to describe doctor's communication patterns in a study of 127 doctors and 537 patients. Their results revealed five distinct patterns: (i) "narrowly biomedical," (ii) "expanded biomedical," (iii) "biopsychosocial," (iv) "psychosocial," and (v) "consumerist" (see Table 1). They found that "biomedical" approaches were used more often with more sick, older, and lower income patients by younger, male doctors. Patients and doctors completed a satisfaction questionnaire immediately after the consultation. The highest levels of doctor satisfaction were found in those using the consumerist approach and the lowest levels were found in those using the narrowly biomedical approach. In contrast the highest levels of patient satisfaction were found with those who had seen doctors using the psychosocial communication pattern whereas the lowest satisfaction scores were recorded in those who had experienced either of the two biomedical patterns.

A very different approach to process analysis can be found in the studies of Ley and colleagues (Ley, 1988), who have concentrated on the informational content of the consultation and the quality of information provided by the doctor. In particular they have analyzed the content in terms of its level of complexity, comprehensibility, and the extent to which the information is organized. They and others have found that medical information may be too detailed or complex with the result that important information may not be understood or retained by the patient. Ley's studies have shown patients' understanding and recall of the consultation can have a direct influence on treatment adherence or can influence adherence via their effects on satisfaction (see Figure 5).

There is even evidence that patients and doctors may interpret the same information in different ways and this communication gap can occur around anatomical information or other technical terms which are used to describe illness or treatment (see Figure 3).

These various ways of conceptualizing and analyzing the consultation process have given rise to a large number of indices or categories which have been related to outcomes, often in quite a limited fashion. Outcomes, such as patient satisfaction or adherence to treatment, are likely to be determined by a range of factors, reflecting a complex interaction of input, process, and situational variables.

8.03.4.3 Outcomes of Health-care Communication

The efficacy of any communication needs to be evaluated in terms of its effect on outcome. For health-care communication, the ultimate outcome is health but only relatively few studies have measured short or longer-term health outcomes following consultations and these will be mentioned briefly. However, a range of other outcomes have been assessed and these fall into three broad groups, namely cognitive, affective, and behavioral.

8.03.4.3.1 Cognitive outcomes

These can be assessed by evaluating changes in patients' knowledge, understanding, and recall of the relevant information provided in the consultation. Ley and colleagues have carried out a number of careful studies of patient recall and these have been summarized (Ley, 1997). Studies which have assessed how much patients are able to recall from consultations have shown that about half the information is retained but there is considerable range. This variation partly reflects the type of setting and sample used and partly the method which is used to test for recall. There are a number of other factors involved including the content of the information, the patients' prior knowledge, and their level of anxiety. With regard to the content and structure of the message, information which is presented early in the consultation is recalled better (the primacy effect) as are statements which are perceived as being important or relevant. Moreover, if the HCP makes an effort to present information in an organized way based around specific themes (e.g., the nature of the problem, the details of the treatment, etc.), then recall is improved. Not surprisingly, the more information which is presented, the smaller proportion is remembered.

8.03.4.3.2 Affective outcome: patient satisfaction

One of the broadest indicators of the patients' affective response to a health-care consultation is their overall level of satisfaction. Fitzpatrick (1997) has commented that the concept of patient satisfaction is important because it focuses on the need to understand how patients respond to health care. As a result, it is increasingly being assessed in surveys of health-care settings as a marker of quality of care, along with such other dimensions of quality as access, relevance to need, effectiveness, equity and efficiency.

Patient satisfaction is also considered important since it is associated with patient

Table 1 Communication patterns and their frequency of use in primary care consultations.

Type of pattern	*Characteristics*	*Consultations*	*% of doctors*[a]
Narrowly biomedical	Low % of talk on psychosocial topics; high % of biomedical information given by dr. & pt.; high % of question asking by dr.	32	68
Expanded biomedical	High % of question asking by dr.; moderate levels of biomedical and psychosocial exchange for dr. & pt.	33	61
Biopsycho-social	Balance between biomedical and psychosocial exchange in dr. & pt.; fewer questions than in the two biomedical patterns	20	42
Consumerist	High frequency of question asking by pt. and information giving from dr.; low levels of question asking by dr. and of psychosocial exchange	8	23
Psychosocial	Dominated by psychosocial exchange; dr. talk divided between psychosocial and biomedical; high % of pt. psychosocial talk; lowest levels of question asking by dr. & pt.; dr.- very positive and accepting of pt. input	7	19

[a] percentage of doctors with at least one of their consultations falling in the pattern Source: Roter et al. (1997).

cooperation (Hall, Roter, & Katz, 1988; Roter, Hall, & Katz, 1987), improved health status (Eraker, Kirscht, & Becker, 1984), and fewer malpractice suits (Vaccarino, 1977). Increased levels of satisfaction have also been found to be related to other important outcomes from the consultation, particularly adherence or compliance with treatment or advice.

Although the concept of satisfaction has good face validity, there is a singular lack of good theory or consensus about its nature or structure. Some approaches are based around the anxiety surrounding illness and equate satisfaction with the adequacy with which this is dealt with by HCPs. Stimson and Webb (1975) have suggested that patient satisfaction is related to the perceptions of the outcome of treatment and the extent to which treatment meets the patient's expectations. This can be related to the self-regulatory model, in that greater satisfaction would be predicted if the patient's own representation of the problem is taken into account by the HCP and if the information which is provided allows the patient to develop a representation of the problem and the treatment which reduces their uncertainty and facilitates their coping, including their adherence to treatment (Leventhal & Cameron, 1987).

More recently Fitzpatrick (1997) has promoted the idea of patient satisfaction as a multidimensional concept, since patients have been found to have differing views about different aspects of their health care, such as the HCP's behavior towards them, the information provided, the technical skills of the HCP, and the access to and quality of the health-care setting. Despite this, there is evidence that the behavior of the HCP is the critical determinant and one which can significantly influence ratings of all the other aspects of health care.

Although patient satisfaction appears to be a relatively straightforward concept, there remain a number of difficulties in measuring it. Results from most satisfaction surveys reveal very skewed data, apparently indicating very high levels of satisfaction across patient samples, particularly amongst older patients (Hall & Dornan, 1988). Part of this problem seems to lie in the reluctance of patients to criticize health-care services and part of it is due to the structure of the questionnaires which have been used to assess patient satisfaction. Comparative studies of different methods have shown variability in satisfaction across measurement approaches. Wide variation in satisfaction scores between three commonly used methods has been reported (Counte, 1979). For example, consistently lower rates of satisfaction are found when the same respondents use a five-point rating scale ranging from "poor" to "excellent" compared to a six-point rating scale ranging from "extremely satisfied" to "very dissatisfied" (Ware and Hays, 1988). In addition, the use of more in-depth methods which require patients to describe their experiences of health from their own perspective, typically give rise to a more critical view (Fitzpatrick & Hopkins, 1993).

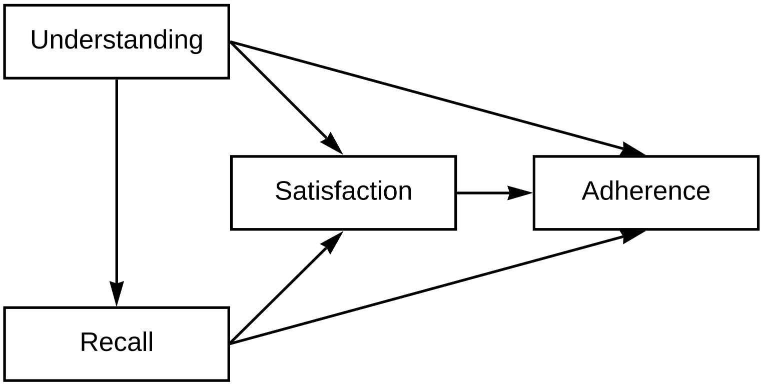

Figure 5 Ley's cognitive model, showing the relations between the patient's understanding, recall, and satisfaction with the consultation and subsequent treatment adherence.

Although these do not provide a quantitative measure, they can be used to highlight areas of dissatisfaction.

There are many patient satisfaction questionnaires that have been developed for use in different health-care settings. For assessing patient satisfaction with the consultation, one of the most widely used measures is the Medical Interview Satisfaction Scale (MISS) (Wolf, Putnam, James, & Stiles, 1978). This scale measures satisfaction with three aspects of the medical consultation. First, it assesses the extent to which the patient feels that the doctor listens, understands, and is interested (affective). Second, it assesses the patient's evaluation of the doctor's competence in the consultation (behavioral) and the third aspect is concerned with the amount and quality of information provided (cognitive).

Bowman, Herndon, Sharp, and Dignan (1992) have identified a number of important structural and measurement criteria for an adequate patient satisfaction scale and have reported that the MISS meets all of these. Moreover, it has good internal consistency and correlates highly with other equivalent measures (Kinnersley, Stott, Peters, Harvey, & Hackett, 1996).

8.03.4.3.3 Behavioral outcome: adherence

The most widely studied behavioral outcome from health-care consultations is reflected in the extent to which the patient adheres to the advice or treatment offered by the HCP. Many consultations result in the prescription of treatment or advice by the HCP and the appropriate adoption of self-care behaviors, including use of medicines, is a key aspect to the self-management of most chronic illnesses. However, many patients fail to do this and low rates of adherence to recommended treatment are seen as problematic in most chronic diseases including asthma (Yeung, O'Connor, Parry, & Cochrane, et al., 1994), diabetes (Glasgow, McCaul, & Schafer, 1986), heart disease (Horwitz & Horwitz, 1993; Monane, Bohn, Gurwitz, Glynn, & Avorn, 1994), cancer (Lilleyman & Lennard, 1996), and kidney disease (Cleary, Matzke, Alexander, & Joy, 1995), as well as in psychological treatments, such as relaxation training for anxiety-related disorders (Taylor, Agras, Schneider, & Allen, 1983).

The incidence of reported medication nonadherence varies greatly from 4–92% across studies, converging at 30–50% in chronic illness (Haynes, Taylor, & Sackett, 1979; Meichenbaum & Turk, 1987). In the area of primary prevention, it has been found that many participants drop out of lifestyle change programs designed to improve diet or reduce health risk behaviors (Dunbar & Agras, 1980). Even patients who have experienced major health problems, such as heart attacks, may show low levels of uptake of rehabilitation programs as well as considerable variation in the adoption of recommended lifestyle change (Petrie, Weinman, Sharpe, & Buckley, 1996). In the area of mental health, there is also evidence of significant rates of nonadherence to various recommendations from HCPs. For example, about half of those undergoing intake in a mental health clinic were found to fail to attend for the first arranged interview and something like three-quarters of psychotherapy patients have been found to drop out by the fifth session of treatment (Phillips, 1988).

There has been a major interest in patient adherence for many years since it has been found that many so patients fail to follow advice or treatment. In view of the importance of this topic, the following section will present an extended overview of it, focusing on the nature of the concept and its causes, as well as ways in which it can be influenced by specific interventions.

Although adherence may be simply defined as "the extent to which the patient's behavior coincides with the clinical prescription" (Sackett & Haynes, 1976), there are several conceptual and methodological issues (Gordis, 1976). It is neither usually possible nor desirable to define nonadherence as an "all or nothing" response in which the patient either completely follows the HCP's instruction (adherence) or deviates from it in some way (nonadherence). For most treatments, the need for total adherence is questionable and this has led Gordis (1976) to define nonadherence as "the point below which the desired preventative or desired therapeutic result is unlikely to be achieved." The percentage adherence necessary to achieve the desired effect varies between treatments and between and within individuals. Many individuals, particularly those with chronic health problems, are required to adhere to a variety of recommendations from their HCPs. For example, diabetic patients are required to take medication, control their diet, and check their feet and blood glucose levels on a regular basis but it has been found that these behaviors are not highly correlated (Glasgow, McCaul, & Schafer, 1987).

Nonadherence behaviors may be categorized as either active or passive. Active nonadherence arises when the patient makes a strategic decision not to take the treatment as instructed. An example of this type of behavior was found among hypertensive patients who believed that they could judge when their blood pressure was high by the presence of symptoms such as stress or headache and thus took antihypertensive medication only when these symptoms were experienced (Baumann et al., 1989; Meyer, Leventhal, & Guttman, 1985). Active nonadherence behavior has been noted among several chronic illness groups including those with asthma (Becker et al., 1978; Woller, Kruse, Winter, Mans, & Alberti, 1993), rheumatoid arthritis (Lorish, Richards, & Brown, 1990), epilepsy (Conrad, 1985), and hypertension (Morgan & Watkins, 1988), and may be particularly influenced by the patients' perception of their problem as well as their level of satisfaction with the HCP and the consultation. From a self-regulatory perspective, the level of treatment adherence may be indicative of a strategic coping response which is entirely consistent with the patient's view of their problem (Leventhal & Cameron, 1987). Thus, patients who believe that their problem will not last for long have been found to be less likely than those with a more chronic timeline representation to adhere to their medication over a long period of time (Meyer, Leventhal, & Guttman, 1985).

Passive nonadherence may be unintentional when the patient's intentions to follow treatment recommendations are thwarted by barriers such as forgetting, and inability to follow treatment instructions because of a lack of understanding or physical problems such as poor eyesight or impaired manual dexterity. Thus, if the quality of communication is poor and patients receive information which is difficult to understand or recall, as has been outlined above, then this makes it less likely that treatment will be adhered to (Ley, 1988).

(i) Measuring adherence

Adherence measures can be divided into two broad categories according to whether the assessment is direct or indirect. Direct measurement entails observing the required adherence behavior such as the ingestion of the drug or by detecting its presence in body fluids. Indirect measures assume ingestion based on proxy evidence such as the patient's report or the number of dosages removed from a container. The strengths and weaknesses of available direct and indirect methods have been reviewed extensively (e.g., Rudd, 1993). At first sight, direct methods might appear to be the best way to assess adherence as other methods do not directly confirm that the medication has been ingested (Caron, 1985). However, these techniques are problematic partly because they may be difficult to carry out and partly because there is not a one-to-one relationship between the amount of medication or tracer taken and the concentration found in body fluids. Furthermore, a major drawback of direct methods is that they are invasive, expensive, of questionable reliability, and provide no indication of the type or time-course of nonadherent behavior. A commonly used indirect method for assessing adherence is to count the number of dosage units left in the container and compare this to the number which would have been left had the patient followed the instructions. This method has the advantages of being technologically simple and inexpensive. However, some studies have shown that pill counts may under-estimate the true level of nonadherence (Kruse, Nikolaus, Rampmaier, Weber, & Schlierf, 1993; Rudd, Ahmed, Zachary, Barton, & Bonduelle, 1990). For example, patients may remove medication for reasons other than to use it. They might give the medication to someone else, or transfer it to another container, or may deliberately discard doses prior to monitoring to create an impression of high adherence.

In the 1990s, technological developments have allowed the incorporation of electronic devices into the medicine container to record the time and date of usage. The major advantage of these devices is that potentially they provide a profile of medication taking rather than simply detailing how much was taken (Kruse et al., 1993; Rudd et al., 1990). However, as with the pill count method, a dose removed is not necessarily a dose taken. Furthermore, if inclusion of the monitoring device changes the appearance of the medicinal product, this will alert the patient to the fact that they are being monitored and may change behavior. Similarly studies assessing adherence to behavioral treatments, such as the use of relaxation tapes, have required the users to keep diaries recording their daily use of the tapes (e.g. Bennett & Millard, 1985) and it is very likely that this self-monitoring process will influence rates of adherence.

One of the most widely used methods of assessing adherence is patient self-report which is held to be a valid indicator of treatment adherence that is practical and useful in a wide variety of research settings (Ley & Llewellyn, 1995; Meichenbaum & Turk, 1987). One problem with self-reported adherence is that questions about medicine taking are often presented at a time and place which is quite distant from the actual event and so reports are subject to recall effects. In particular, even when they admit to nonadherent behavior, people tend to overestimate the extent of their adherence (Ley & Llewellyn, 1995). The accuracy of self-report as a measure of adherence has been assessed by comparison with other, more objective, methods or evaluated on the basis of correlation between self-reported adherence and clinical outcome measures. In the main these studies show that the accuracy of self-report varies according to the type of adherence behavior (e.g., medicine taking vs. lifestyle change) as well as the type of self-report measure used. Nevertheless Morisky, Green, and Levine (1986) have shown that a four-item scale of adherence to antihypertensive medication had acceptable internal consistency and was moderately predictive of blood pressure control at two and five year follow-ups.

(ii) The determinants of nonadherence

The search for causal factors to explain patients' adherence or nonadherence to their recommended treatment or advice has progressed through different phases since the 1960s. Much of the early work focused on the possible contribution of demographic or personality factors, as well as particular characteristics of the treatment or the disease which might influence adherence levels. More recent studies have tended to concentrate on cognitive factors, particularly patients' beliefs and attitudes, since these are potentially modifiable.

In an early systematic review of 185 studies (Sackett & Haynes, 1976), no clear relationship emerged between race, gender, educational experience, intelligence, marital status, occupational status, income, and ethnic or cultural background and adherence behaviors. Moreover, there is little evidence that adherence behaviors can be explained in terms of personality characteristics (Becker, 1979; Bosley, Fosbury, & Cochrane, 1995; McKim, Stones, & Kozma, 1990). Also, the idea that stable sociodemographic or dispositional characteristics are the sole determinants of adherence is discredited by evidence that an individual's levels of adherence may vary over time and between different aspects of the treatment regimen (Cleary et al., 1995; Hilbrands, Hoitsma, & Koene, 1995; Rudd et al., 1990). This limitation also applies to the search for disease and treatment characteristics as antecedents of adherence since there are wide variations in adherence between and within patients with the same disease and treatment (e.g., Cleary et al.; Lilleyman & Lennard, 1996).

One very obvious explanation for nonadherence arises from the poor cognitive outcomes outlined above, particularly poor understanding and recall of information presented in the consultation. Many patients lack basic knowledge about their medication (Al Mahdy & Seymour, 1990; Cartwright, 1994; Eagleton, Walker, & Barber, 1993) but the relationship between this and their adherence is neither simple nor clear-cut. In a review of the adherence literature Haynes (1976) concluded that, although 12 studies had demonstrated a positive association between knowledge and adherence, there were more that had failed to demonstrate a link. Studies conducted since then generally indicate that associations between knowledge and adherence are at best small and inconsistent (Eagleton et al., 1993), and interventions which enhance knowledge do not necessarily improve adherence (George, Waters, & Nicholas, 1983; Haynes et al., 1978).

There is increasing interest in the role of patient satisfaction as a mediator between information provision, recall, and adherence, as would be predicted by Ley's model which was outlined earlier (see Figure 5). In a national UK survey of patients' satisfaction with medicines information, over 70% of respondents wanted more information than they were given (Gibbs & George, 1990). Dissatisfaction with attributes of the practitioner or the amount of information and explanation provided may act as a barrier to adherence by making the patient less motivated towards treatment (Hall et al., 1988).

The emphasis of adherence research since the late 1980s has moved away from attempts to identify stable trait factors which characterize the nonadherent patient to achieving a greater understanding of how and why patients decide to take some treatments and not others (Horne, 1993). Much of this research is informed by psychological theories which conceptualize behavior as the product of cognitions which occur within a social framework and these are described more fully in Chapter 8.01.

The application of social cognition models in research indicates that medication nonadherence may arise from a rational decision on the part of the patient and identifies some of the cognitions which are salient to these decisions. Although there is some variation in the specific type of beliefs which are associated with adherence across studies, the findings show that certain cognitive variables included in the Health Belief Model (HBM) (Janz & Becker, 1984) and Theory of Planned Behaviour (TPB) (Ajzen, 1988) appear to be prerequisites of adherence in certain situations. For example, beliefs that failure to take the treatment could result in adverse consequences and that one is personally susceptible to these effects tend to be associated with higher adherence (Cummings, Becker, Kirscht, & Levin, 1981; Kelly, Mamon, & Scott, 1987; Nelson, Stason, Neutra, Soloman, & McArdle, 1978). Perceived severity of anxiety has also been found to be related to adherence to recommended practice of relaxation training at home (Bennett & Millard, 1985). Additionally, adherence decisions may be influenced by a cost-benefit analysis in which the benefits of treatment are weighted against the perceived barriers (Brownlee-Duffeck et al., 1987; Cummings et al., 1981; Nelson et al., 1978). Other studies, based on the TPB, have shown that the perceived views of significant others such as family, friends, and doctors normative beliefs may also influence adherence (Cochrane & Gitlin, 1988; Reid & Christensen, 1988). Several studies have demonstrated the value of interventions based on the HBM in facilitating health-related behaviors, such as attending for medical check-ups (Haefner & Kirscht, 1970), or using emergency-care facilities in an acute asthma attack (Jones, Jones, & Katz, 1987).

Another cognitive approach which has been used to explain nonadherence is the self-regulatory model outlined at the beginning of this chapter. This model also acknowledges the importance of symptom perception in influencing illness representations and adherence as a coping behavior. Confirmatory evidence for this is provided by findings from a study of patients with diabetes who used perceived symptoms to indicate their blood glucose levels and to guide self-treatment (Gonder-Frederick & Cox, 1991). Unfortunately patients' beliefs about their symptoms, and estimations of their own blood glucose levels, were often erroneous and resulted in poor diabetic control. Further evidence of the importance of illness representations was obtained by Meyer et al. (1985) who noted a clear relationship between illness representations and behavior in their study of patients with hypertension. Patients who believed that their hypertension was an acute condition were more likely to cease taking antihypertensive medication than those who believed it to be a chronic condition. This study also showed that patients' representations of their illness often conflicted with the medical view and provided an insight into the effects of mismatch between the patients' representations and those of their doctor. In a group of 50 patients who had continued in treatment, 80% agreed with the statement that "people cannot tell when their blood pressure is up." However, 92% believed that they could tell when their own blood pressure was raised by monitoring symptoms such as tiredness, headache, and stress. Patients who believed their anti hypertensive medication improved symptoms were more likely to take it. A striking example of this was provided by five out of 17 patients who believed that their medication affected symptoms. These patients took their antihypertensive medication only when they judged their blood pressure to be raised. The patients had understood and accepted the abstract medical view of hypertension as an asymptomatic condition but their concrete experience or symptoms caused them to hold contrasting beliefs and to behave according to these.

Illness perceptions have been linked with a range of adherence-related behaviors other than the use of medicines. These include various self-management behaviors, such as dietary control and blood glucose testing in diabetes, attending rehabilitation, and the adoption of various lifestyle changes following myocardial infarction. A study of noninsulin dependent diabetic patients, has shown that personal models of diabetes are related to dietary self-management and to exercise adherence but not to the more medical aspects of control, such as blood glucose testing and taking medication (Hampson, 1997). Similarly, in a prospective study of patients following first-time myocardial infarctions (MI), Petrie et al. (1996) have found that specific illness perceptions are predictive of different post-MI behaviors. They found that attendance at rehabilitation, which is prescribed for all patients, was predicted by the strength of their belief in the cure/control of their MI whereas return to work depended more on the extent to which the patient saw their MI as having less serious consequences

In addition to studies of the role of illness beliefs, there is a small body of work which has examined people's beliefs about medicines and the ways in which these could influence adherence (Britten, 1994; Conrad, 1985; Donovan & Blake, 1992; Morgan & Watkins, 1988). The negative beliefs about medicines identified in these studies appear to be common across several illness and cultural groups and include worries about the potentially harmful effects of medicines and about long-term dependence on them. However, there are only a few studies which have assessed medication beliefs quantitatively and they have used different questionnaires or have investigated medication beliefs in the broader context of views about the practice of medicine (Echabe, Guillen, & Ozamiz, 1992; Marteau, 1990). Some studies have assessed peoples' ideas about medicines in general, whereas others have focused on specific medication prescribed for a particular illness.

Research by Horne (1997) indicates that four "core-themes" or factors underlie commonly held beliefs about medicines. Factor analysis of a pool of belief statements revealed two broad factors describing peoples' beliefs about their prescribed medicines: their perceived necessity for maintaining health (specific-necessity) and concerns based on beliefs about the potential for dependence or harmful long-term effects and that medication taking is disruptive (specific concerns). Two factors were also found to describe peoples' beliefs about medicines in general. The first relates to the intrinsic properties of medicines and the extent to which they are harmful, addictive substances (general-harm) and the second factor comprises views about whether medicines are overused by doctors (general-overuse).

Peoples' views about the specific medication regimen prescribed for them were found to be much more strongly related to adherence reports than are more general views about medicines as a whole. Moreover, an interplay was found between concerns and necessity beliefs which suggests that people engage in a risk-benefit analysis and consequently attempt to moderate the perceived potential for harm by taking less. Patients with stronger concerns based on beliefs about the potential for long-term effects and dependence reported lower adherence rates, whilst those with stronger beliefs in the necessity of their medication reported greater adherence to medication regimen (Horne, 1997).

8.03.4.3.4 Consultation outcomes: an overview

The cognitive, affective, and behavioral outcomes of the consultation are very closely linked and can influence each other. Patient satisfaction, understanding, and beliefs can play a major role in influencing adherence with treatment or advice, which is obviously important in situations where nonadherence results in adverse health consequences. Since there is evidence of high levels of nonadherence, this can clearly affect other outcomes including health and well-being. The latter have not often been studied as communication outcomes but there are a number of studies which demonstrate positive effects on patients' health and well-being arising from positive experiences in medical consultations (Stewart, 1995). These have focused on psychological states such as anxiety as well as changes in specific physical variables such as blood pressure and blood glucose control. Some of the most impressive findings here have been found in the patient-intervention studies, which are described below.

8.03.4.4 Improving Health-care Communication

In addition to increasing our understanding of health-care communication and its central role in health-care delivery, some research findings have also provided insights for developing interventions to improve the quality of communication. The majority of these have been aimed at improving the communication skills of health-care students or practitioners at various stages of training but a few have been targeted at patients, to enable them to get the most from a consultation. Both types of intervention approach will be outlined.

Communication skills training is now regarded as a fundamental part of the curriculum for medical, nursing, and other health-care students but this varies considerably in terms of the amount and type of teaching and the stage at which it is taught. Typically, students are provided with an overview of the basic skills of "active" listening which facilitate patient communication. At a basic level these include the importance of developing good rapport and the use of open-ended questions early in the consultation, appropriate eye-contact and other facilitatory responses to help the patient talk, together with the ability to summarize and arrive at a shared understanding of the patients' problem. These skills can be taught in a number of ways but the successful courses inevitably involve active learning, using role-plays with simulated patients, as well as real patient interviews (see Kendrick, 1997; McManus, Vincent, Thom, & Kidd, 1993). Feedback is important to identify problem areas as well as indicators for improvement, and increasing use is made of videotape for this purpose. There is consistent evidence that this type of training can result in clear improvements in basic communication skills which are maintained for a number of years (Maguire, Fairburn, & Fletcher, 1986).

In addition to these basic packages, it is also necessary for health-care students to learn how to communicate about sensitive or difficult areas of clinical practice, including dealing with distressed patients or relatives and giving "bad news." Research in this area, using tape-recorded consultations, has shown that doctors tend to give detailed information rather than find out and respond to patients' concerns and informational needs. As Maguire (1997) notes, this avoidance of patients' worries can have negative effects in the short and longer term. An immediate consequence is that patients may remain preoccupied with their own concerns and fail to take in information or advice. They may also selectively attend to negative phrases or messages and be unresponsive or misconstrue more positive or neutral information. Maguire (1997) provides the example of a surgeon who informs a recently diagnosed cancer patient that radiotherapy will be given "to mop up any residual cells," and then adds "I am sure that we will eradicate your cancer." When questioned afterwards, the patient only recalled the phrase "residual cells" and became very distressed since she thought that this meant that the cancer would spread through her body.

There were are various reasons why doctors and other health-care professionals may be poor at "bad news" communication and these include lack of training, fears about the effects of exploring patients concerns, lack of support, and a desire to protect their own emotional well-being. Some of the training packages which have been developed have taken these factors into account and provide sufficient time for the learners to explore and get feedback on different strategies, using role-play situations. Many of the component skills are those which are part of any communication training package, such as those outlined earlier in this section. The key first step is to ensure that this communication takes place in an appropriate environment and needs to start with the patient, exploring what they have been told and what they think and feel about their condition. If will often be necessary to take time to provide the patient with the information which they want, letting patients proceed at their own pace and allowing information to be assimilated. The other key skills involve developing the ability to recognise distress and allow this to be expressed without feeling discomfort or providing unrealistic reassurance (Maguire, 1997).

There is growing evidence that the way bad news is presented to patients can have major effects on the patient's perception of their condition and on subsequent coping and adjustment (Fallowfield, Hall, Maguire, & Baum, 1990). Effective training of health-care professionals in this area can have significant effects on the quality of health care for patients with serious or terminal conditions. In addition to improving the HCP's communication skills, there is also considerable scope for providing patients with more useful information, including tape-recordings of their consultations.

Since the 1980s, there have been a number of interesting interventions aimed at patients. Generally, these have involved interventions for patients prior to a consultation in order to increase their level of participation, particularly to ensure that their own concerns are dealt with and that information provided by the doctor is clearly understood. A successful development of this approach can be seen in the work of Greenfield, Kaplan, and Ware (1985) who used a preconsultation intervention lasting 20 minutes for hospital outpatients who were helped to identify their main questions and encouraged to ask these in the consultation. Compared with control patients, these patients participated more actively in the consultation and this was also associated with better long-term health outcomes, including lowered blood pressure in hypertensives and better glycaemic control in diabetic patients. These interventions can be quite time-consuming and a number of researchers have explored the possible efficacy of simpler patient-based interventions, such as the use of preparatory leaflets (e.g., McCann & Weinman, 1996), but these have not shown such significant effects in changing the process or outcome of consultations.

Finally, mention should be made of two specific patient-based approaches which have been very successful. The first by Ley and colleagues (see Ley, 1988) involved hospital patients and an additional short visit which allowed them to ask for any information to be clarified. Compared with control groups, these patients had a much higher level of satisfaction with communication, indicating that effective interventions need not be of complex or time-consuming. Hogbin and Fallowfield (1989) describe another simple intervention which consisted of making a tape-recording of the "bad news" consultation and allowing patients to take away and keep the tape. Since this type of consultation is often very distressing, patients may often find it very difficult to take in all the information. Thus, it was found that patients welcomed the use of these tapes as something which they could go back to and which others could also listen to.

The importance of providing patients with access to information in various forms, which can be discussed and shared with family and friends, is very pertinent in this area of health. In this respect it is interesting to note the development of patient support groups and information exchanges using the "internet" (Davison & Pennebaker, 1997). Providing patients with access to others with similar or related medical conditions may result in very important gains in communication. For example, patients may be able to locate and communicate with others who have very similar experiences and circumstances, which may not only have a supportive function but may also provide ways of sharing and coping with many of the problems which are experienced.

8.03.5 HEALTH CARE IN HOSPITALS

This section of the chapter will focus on a range of psychological aspects of health care in hospital settings. It begins with a general consideration of the psychological consequences of admission to hospital and is followed by a brief outline of some of the psychological problems associated with the hospitalization of younger children. Then there is a selective overview of particular hospital treatments such as intensive care and haemodialysis which can produce specific emotional reactions due to the limitations and demands they impose on patients. Similarly, many medical procedures are found to be painful or distressing and a number of psychological interventions, which have been devised to help patients cope with these, are outlined in the final part of this section.

8.03.5.1 Psychological Effects of Hospitalization

Hospital patients vary in many ways including their age, personality, and social circumstances, as well as in the severity of their health problems and the duration of their stay. Also, their experience in hospital will vary greatly and this inevitably means that there are difficulties in attempting a general discussion of the impact of hospital admission. Some specific issues concerning children are presented separately but the following account is intended to convey some of the more general factors which have been identified. The psychological reactions which are described may reflect not only a response to hospitalization but also to the illness itself.

8.03.5.1.1 Physical and social environment

One obvious feature of hospital life which involves a considerable change for the patient is the physical environment. Although a number of studies have demonstrated positive and negative effects of the built environment on psychological well-being and health (Spencer & Baum, 1997), there are relatively few studies which have directly investigated how the physical environment of the hospital affects the patient's condition. The hospital environment has been identified as one of a range of stressors which people experience in hospital (Koenig, George, Stangl, & Tweed, 1995). Even such factors as the nature of the view from the patient's bed have been found to affect recovery. For example, Ulrich (1984) found that patients with a window view of trees had a better postoperative recovery (i.e., fewer complications, less analgesia, better adaptation, shorter length of stay) than patients recovering from the same type of surgery but whose hospital window looked out on to a brick wall.

The hospital environment is frequently seen to be drab, clinical, impersonal, and cold by patients. Whereas in the outside world the individual operates with a strong sense of personal space, which is the amount of space necessary for optimal social behavior, in hospital this space is being constantly invaded as part of the daily routine. The daily routine itself is one which is very likely to be very different from life at home and therefore may require considerable adjustment. It is not therefore surprising to find that studies, which have compared home-treated and hospitalized patients with the same condition, have shown less psychosocial distress in those remaining at home (e.g., Oldenburg, Macdonald, & Perkins, 1988).

Enforced life-style changes have also been identified as a key hospital-based stressor in the Hospital Stress Index (Koenig et al., 1995). Admission to hospital removes individuals from a familiar, well-ordered world and places them in an environment which is different in every respect. In hospital they are likely to be totally dependent on others for most basic functions such as washing and feeding. Usually they will be restricted to one place, surrounded by totally new people, whose skills are now of vital importance to them, and it is significant that developing good relations with hospital staff is not only an important factor in adaptation to hospital life but can also be a potential source of stress for patients (Koenig et al.). Other hospital stressors which have been identified by an older measure, the Hospital Stress Rating Scale (Volicer & Bohannon, 1975), include worries about aspects of communication with staff as well as concerns about investigations and treatment, which are discussed in more detail below.

On the positive side, studies by Kulik and Mahler (1987) and Kulik, Mahler, and Moore (1996) have demonstrated the importance of social contact in the recovery of patients following surgery. These studies have focused on the effects of having a preoperative room-mate on the anxiety and recovery of surgical patients. In the early study they compared the effects of sharing a room either with a patient who was also about to undergo cardiac surgery or with one who had already had the same operation. The results showed clear beneficial effects of sharing a room with someone who was recovering from surgery. The patients who had postsurgical room-mates were less anxious prior to surgery, engaged in more postsurgical physical activity, and were discharged sooner (Kulik & Mahler, 1987). In a more recent variation on this study, two further variables were investigated, namely having a room-mate or being on ones own and having a room-mate who was about to undergo or had gone through same type of operation. In addition to replicating the earlier findings with respect to the advantage of sharing with a postsurgical patient, it was also found that it was advantageous to share with a postsurgical patient who had undergone the same type of surgery and that those who were in rooms on their own had the slowest recovery (Kulik et al., 1996).

8.03.5.1.2 Communication in hospital

In addition to the general stresses associated with hospital admission, some of the fear and anxiety which is found in many hospital patients may stem from the uncertainty and lack of information which they may have about the nature of their illness and its treatment and prognosis. There are a number of studies of hospital patients which show clearly that they are greatly dissatisfied with the communications aspect of hospital life (e.g., Ley, 1972). These studies indicate that about 40–50 % of hospital patients are critical of the communication aspects of their stay. Many of the communication problems which were discussed in the earlier part of this chapter (see Section 8.03.3) are also seen in hospital settings.

There is also evidence that hospital patients are often diffident in this respect and are unwilling to ask for information. The diffidence which seems to underlie patients' reluctance to complain or ask for information has been the subject of psychological study and it would appear that patients enter hospital with ideas about how they should behave. Lorber (1975) has identified three categories of patient which she labeled "good patients" (about 50% of her sample), "average patients" (about 25%), and "bad patients" (about 25%) on the basis of the ratings of their medical staff. Patients who were labeled "good" were ones who had straightforward medical problems, were uncomplaining and docile, who took up minimal staff time, and showed uncomplicated recovery. The "average" group were rather similar except they had some problems, but these were manageable. The "bad" group consisted either of patients with serious conditions or those without serious conditions but who complained and put extra demands on the staff as a result.

Interestingly, these patterns of behavior could be detected from interviews with patients about their views of the patient's role at the time of their admission. Thus, it would appear that many patients enter hospital with the expectation of being in a relatively passive role and the pattern of life in hospital actually reinforces this, sometimes to the point of helplessness. Taylor (1979) describes how helplessness can develop when questions go unanswered and desires for attention are not met. As a result "good" patient behavior may actually be detrimental to recovery since it prevents patients from taking an active role in their health care. Moreover "bad" patient behavior can be seen as a reaction against helplessness and an attempt to gain a degree of control, and hence could be thought of as a healthier response.

Fortunately, these problems do not appear to be insurmountable. For example, Ley, Bradshaw, Kincey, and Atherton (1976) conducted a study of hospital patients on three wards of a general hospital. In one ward there was an extra visit made to patients every 10 days, in which attempts were made to ensure that they had understood what they had been told. These visits were relatively brief and only concerned with clarifying existing information rather than raising new issues. The second ward essentially comprised a control group, who received an extra visit to discuss their welfare, food, and so on and not to clarify information. Patients in the third ward in the study received no extra visits. The results showed that 80% of patients who had received the informational visit were satisfied with the communication received, which was approximately twice the level of satisfaction with communication found in the other two groups.

In addition to these general psychological impacts of hospitalization, there may be specific problems or demands which occur either as the result of the particular health problems or the type of treatment which the patient has to undergo. An example of the way in which the patient's health problem may influence their experience of hospital care can be seen in some of the studies of patients with AIDS who may experience negative attitudes from staff or other patients. These negative attitudes are closely linked with the blaming attributions which staff or patients may have for the cause of AIDS. In a survey of 270 AIDS patients being treated in either special care units or integrated in more general hospital settings, specific stresses were reported by those in the integrated units (Van Servellen, Lewis, & Leake, 1990). Whereas both groups were bothered by the same range of general stressors identified above, particularly loss of independence, separation from loved ones, and problems with medicines, those who were in integrated settings more frequently experienced and were upset by factors associated with ambiguity about their care and their medical condition. Moreover, these patients also reported higher levels of stress associated with feelings of abandonment and impersonal or discriminatory treatment. A number of studies of HCPs have revealed evidence that their attributions of patient blame for their condition can have direct influences on their attitudes to the care of the patient as well as their approach to and involvement with them. Broadly, it has been shown that where staff see patients as instrumental in having brought about their own condition through their own behavior or neglect, then they may be less committed, motivated, and sympathetic towards them (Marteau & Riordan, 1992).

8.03.5.2 Children in Hospital

In many countries approaching 50% of children will have spent a period of time in hospital and for those with chronic conditions many have to spend longer periods or visit repeatedly over a number of years. There is evidence that children's reactions to illness and hospital treatment can be quite distinct and that the psychological and social consequences of separation from home can produce quite specific problems. Older studies of the emotional reactions of children to hospital (e.g., Illingworth & Holt, 1955) showed that admission to hospital was often very distressing for them, particularly the younger ones, but the situation has improved over the years. Since that time various recommendations have been put forward and, to some extent, implemented to allow greater access by parents to their hospitalized children. These have included increased flexibility in visiting arrangements as well as a greater provision of overnight accommodation for parents.

Adaptation to hospital life will depend greatly on the age and personality of the child as well as on the family and their reaction to the child, to disease, and to hospitals. When they occur, emotional problems in hospital are much more prevalent in younger children, particularly up to and around the age of four. Older children are more amenable to explanation and are usually less distressed by separation from home and by being surrounded by strangers. A more extensive discussion of children's responses to health issues and health care can be found in Chapter 8.27.

There are a number of specific psychological problems associated with the hospitalization of children. The problems would appear to stem from three sources:

(i) The social separations and disruptions incurred by admission to hospital;

(ii) The emotional response to the clinical problem; and

(iii) The lack of preparation for hospital life and the investigations and treatments to be carried out.

The more long-term effects of hospitalization on children suggest that some of the traumatic effects reported in early studies may have been overstated. Follow-up studies of hospitalized children indicate that serious problems are only found in children who had psychological problems prior to admission or who come from difficult families (La Greca & Stone, 1985). Studies of preparation of children for hospitalization also show that different types of preparation appear to be suitable for children of different ages and with different amounts of prior experience (see Schmidt, 1997 for an overview).

8.03.5.3 Psychological Aspects of Specific Hospital Treatments

In the same way that physical illness imposes physical and social limitations on the individual which can give rise to psychological reactions, some treatments are also very restricting and have been found to cause emotional and behavioral changes. In particular, a number of studies have been made of patients in specific treatment settings such as intensive care units (ICUs), coronary care units (CCUs), and haemodialysis. Although these treatment environments can give rise to specific stressors, they also offer challenging opportunities for psychological interventions.

Experimental work with healthy volunteers has shown that long periods of sensory deprivation or sensory overload will often give rise to a state characterized by increased wakefulness, disorientation, and visual hallucinations (Goldberger, 1982) These states have also been found in some patients who are being treated in ICUs since they may be exposed to long periods of sensory deprivation produced by monotony and immobilization, sometimes interrupted by periods of overstimulation, as well as sleep deprivation. It is claimed that the observed "ICU psychosis" is the result of all these factors on an individual who is already fearful of his or her life because of a serious physical illness. Fortunately, these dramatic psychological responses gradually disappear when normal amounts of sensory stimuli and sleep are restored by returning the patient to a normal hospital environment. More important, these changes can be avoided if the ICU contains a more varied sensory environment and with good contact with visitors and staff. For example, Keep, James, and Inman (1980) have compared patients in ICUs with and without windows, and found that those in the windowless units were less well oriented during their stay and had a less accurate recall of their length of stay afterwards. In addition to these general problems associated with the ICU, other studies have assessed the degree of stress experienced by patients, staff, and visitors. For example, in a study of a surgical ICU, patients rated their level of stress from a wide range of events which occurred (Pennock, Crawshaw, Maher, Prue, & Kaplan, 1994). The ratings were made shortly after their transfer from the ICU and indicated relatively minimal distress associated with many events with exception of being intubated and not being able to communicate. For patients' relatives there is evidence that they find the time spent by the patient on life support in the ICU particularly worrying. During this time they experience considerable fear and uncertainty but this can be resolved by seeking information and the use of other resources (Jamerson et al., 1996). For staff working in this climate, it is assumed that greater levels of work-related stress will be experienced, compared with other hospital personnel. However, the evidence on this is ambiguous and some studies have shown that ICU staff show a more positive attitude to their work and work environment than equivalent non-ICU staff (Boumans & Landerweerd, 1994).

In the context of coronary care, a number of studies have identified possible areas for psychological intervention in facilitating the recovery of patients in CCUs. Using a self-regulatory framework, one study has shown that patients with myocardial infarction have clear beliefs about the cause, timeline, consequences, and controllability of their heart condition during the acute phase of their hospital stay in the CCU. Moreover, these beliefs appear to be associated with subsequent attendance at cardiac rehabilitation and with later adaptive changes, such as return to work and social functioning (Petrie et al., 1996). In addition to these cognitive processes, there is also good evidence that patients' mood states at the early stages of recovery are related to longer-term physical and psychological well-being (Frasure-Smith, Lesperance, & Talajie, 1995) and there is now increased awareness of the ways in which appropriate care during the hospitalization phase of treatment can promote adaptive coping and recovery (see Bennett, 1994).

In contrast to the acute psychological restrictions and demands of intensive or coronary care, some patients are subject to much more chronic restrictions as part of their treatment. Many patients with renal failure are required to spend a considerable amount of time on renal dialysis machines, either in hospital or at home. These patients have a uniquely dependent relationship not only on their dialysis machine but also with the staff involved in their treatment. Dialysis can have major effects on an individual's psychological and social functioning, particularly giving rise to vocational impairment, reduced sexual activity, and mood changes (Oldenburg et al., 1988). In addition to the physical limitations and demands of dialysis, the patients are also faced with the need to adhere to strict recommendations regarding diet and fluid consumption, as well as complex medication regimens. Nonadherence can be a key issue in renal failure and the range of adherence-related factors outlined in an earlier section are very pertinent for patients on dialysis.

One finding which emerges from studies of patients on renal dialysis is that psychological dysfunction at an early stage is a predictor of longer-term adjustment. This means that it is not only vital that good psychological care is available during hospital treatment but also that "at risk" individuals can and should be identified at an early stage in order to anticipate later difficulties. It is noteworthy that comparative studies of patients on hospital or home-based dialysis show that the home-treated individuals show much less social dysfunction (Oldenburg et al., 1988). Part of this may well reflect the patient selection criteria for hospital and home dialysis, with more "problematic" patients tending to be treated in the hospital setting. However, part of this may also reflect the additional negative effects of hospitalization, outlined earlier, superimposed on the specific limitations and demands of dialysis.

The most commonly reported psychological difficulties found in patients on renal dialysis are depression and anxiety but the extent of this is unclear. There is a range in the reported prevalence of depression from 10–100% (Levy, 1994) and this variation is partly due to the different criteria and measures used (see Kaplan De-Nour, 1994). Another reason stems from the similarities between the physical effects of renal failure, such as fatigue, apathy, and sleep difficulties, and the symptoms of depression. Despite these methodological problems, there do appear to be a number of aspects of dialysis which can give rise to psychological distress. For example, Devins, Binck, Hollomby, Barre, and Guttman (1981) have identified the constant threat of death, dependence on the dialysis machine, medical staff, and the stringent dietary and liquid restrictions as key factors in the widespread feelings of helplessness and lack of control. Patients' perceptions of control over their treatment have been found to interact with treatment experiences and illness severity in determining mood outcomes (Christensen, Turner, Smith, Holman, & Gregory, 1991). A belief that one's health is subject to personal control has been found to be associated with lower levels of depression in people who had only been treated by dialysis whereas, for those who had returned to dialysis following a failed kidney transplant, higher control beliefs over health were associated with greater depression.

In line with self-regulatory theory, patients representations of their kidney disease, including their control perceptions, would be expected to influence coping and other outcomes, including treatment adherence (Horne & Weinman, 1994). Studies by Christensen, Benotsch, Wiebe, and Lawton (1995) have demonstrated that problem-focused types of coping were associated with better adherence to fluid intake restrictions when these coping strategies were used in response to stressors arising from a relatively controllable aspect of dialysis. For those stressors which patients perceived as less controllable, emotion-focused coping strategies were associated with better levels of adherence.

8.03.5.4 Stressful Medical Procedures in Hospital

In addition to the generally stressful effects of hospital admission, there is a range of medical procedures which can give rise to considerable discomfort and anxiety. These include specific investigative procedures such as barium X-rays (Allan & Armstrong, 1984), endoscopy (Johnson, Morrissey, & Leventhal, 1973), and cardiac catheterization (Kendall et al., 1979) which may not only be uncomfortable and sometimes physically distressing but which also carry the threat of uncovering a serious medical condition (Weinman & Johnston, 1988). The other obvious "event" in hospital which has been found to produce significant psychological effects is surgery. The psychological impact of surgery will partly depend on the procedure, the condition, and the likely outcome as well as on such psychological factors as the patient's expectations and coping style and the quality of communication.

The way in which a patient reacts to a medical procedure can also have a significant influence on the outcome, particularly in recovery from surgery. Patients who show the highest presurgical levels of stress response will also tend to experience adverse psychological reactions postsurgically and will be more likely to show poorer physical recovery (Johnston, 1986). These patients have been found to request more analgesia, show more postsurgical complications, and tend to recover more slowly with delays in discharge, as compared with less anxious or stressed patients.

8.03.5.5 Psychological Interventions for Stressful Medical Procedures

Since studies have shown a relation between patients' psychological state and their recovery, it has been recognized that there could be considerable gains from providing a psychological intervention designed to reduce or minimize the psychological impact of a medical procedure. There is a range of interventions which have been used to prepare patients for surgery or other stressful procedures in the hospital setting. In broad terms, they can help by providing the patient with information to reduce the uncertainty of the event, or with specific behavioral or cognitive skills to help with some of the discomfort or pain Mathews and Ridgeway (1984) have provided a clear overview of the various preparations and these are summarised in Table 2.

Procedural information is probably the most widely used approach and consists of providing information about the various procedures which will take place before and after the operation. In short, it involves a description of what will be done to the patient at different stages pre- and postoperatively. Sometimes this information is also accompanied by an explanation of the purpose of each of the procedures which are described. Thus, it provides the patient with a map of the events which will occur and, in doing so, can reduce the uncertainty of the whole process.

Sensory information describes what patients are likely to feel, particularly during the immediate postoperative pain period. The important point here is to provide matter of fact or benign interpretations of the sensations so that the patient can recognize them as part of the expected postoperative process. Thus, the patient who can recognize postoperative pain as an expected sensation caused by the incision and reflecting the healing process will be far less likely to be distressed than someone who has not been prepared for the pain and who may think of it as problem or a complication of the surgery.

Contrada, Leventhal, and Anderson (1994) have outlined an interpretation of the benefits of sensory and procedural information from the perspective of the self-regulatory model. They view the sensory preparation as providing a script which describes internal sensations and the procedural preparation as a script providing the objective external events involved in surgery. They maintain that it is the availability of the script which reduces uncertainty and worry for the patient. More specifically, Leventhal (1985) has argued that sensory information should be particularly helpful since it focuses on potentially threatening sensations (e.g., pain, discomfort) with the aim of ensuring that these are processed as nonthreatening or less threatening. However, evidence on the efficacy of both types of preparatory information indicates that procedural information is at least, if not more, effective in producing favorable outcomes (Johnston & Vogele, 1993, see below).

Behavioral instructions are also commonly provided and describe different behaviors which will help before, during, and after surgery. These include instructions about ways to cough and move in bed which will reduce the likelihood of pain associated with these movements. Other behavioral instructions such as deep breathing and ambulation exercises may also reduce the incidence of pain or complications as well as facilitating recovery.

Modeling is based on the use of filmed models who can be seen undergoing the same procedure as the patient. Following Bandura's social learning theory (Bandura, 1986), modeling or the observation of others completing a difficult or stressful task can serve to increase the individual's sense of self-efficacy for managing the same task. Two main types of model have been investigated: (i) mastery models who are shown dealing with the task with ease and ability, and (ii) coping models who are shown as having some anticipated concerns but who nevertheless are able to overcome these and cope with the procedure. The coping type of model has been found to be a more effective preparation for children undergoing surgery. For example, in a study of children about to be innoculated, Vernon (1974) has compared a group of children who saw a preparatory film which was realistic (the child in the film is seen to experience short-lived, moderate pain and emotion) with a group who saw an unrealistic film (no pain or emotional expression) and a group who saw no preparatory film. The realistically prepared group were found to experience least pain when receiving their injections. These methods have been more widely used with children than with adults, particularly since it may be difficult to provide children with sensory or procedural information or behavioral instructions in a meaningful way.

Relaxation-based interventions can involve a number of different techniques. These may involve the use of deep breathing, progressive muscle relaxation, or, less frequently, hypnosis. Relaxation can be used both to provide a general preparation involving anxiety-reduction and to give a specific skill which can be used postoperatively for dealing with pain or discomfort.

Cognitive coping procedures focus on patients' concerns and fears about the surgery and provide ways of dealing with them in one of two ways. First, they may make use of coping strategies which the patient has used successfully in the past for dealing with stressors, enabling the patient to rehearse and apply these in the surgical context (Langer, James, & Wolfer, 1975). The second cognitive approach involves dealing with negative thoughts by distracting attention from them and by focusing on positive aspects of the surgery and repeating positive self-statements (Ridgeway & Mathews, 1982).

The efficacy of these interventions has been evaluated by examining their effects on a range of postsurgical outcomes, including anxiety, pain, and use of pain medication, length of stay in hospital, and various indicators of recovery. All the interventions have been found to be successful in improving at least one aspect of outcome and the majority have a positive impact on many of the outcomes. The different interventions have been examined systematically in a meta-analysis by Johnston and Vogele (1993) and their findings will be outlined briefly. The largest recovery effects were obtained for pain, negative affect, and physiological indices

Table 2 Psychological preparations for stressful medical procedures.

Type of preparation	*Example*
Procedural information	"After the operation, you will be taken to the recovery room, where specialist staff will care for you"
Sensory information	"It is normal to feel a sharp, burning sensation along the line of the incision"
Behavioral instructions	"You should try to cough four times each morning and afternoon/evening to keep your chest clear"
Modeling	"This film shows someone like you coping with the procedures that you will be experiencing"
Relaxation	"Try to breathe deeply and concentrate on relaxing the muscles of your body whenever you feel tense"
Cognitive coping	"You say that you are worried about the anaesthesia. Try to think of ways in which you could make those thoughts more positive"

Source: Mathews & Ridgeway (1984).

of recovery but there was considerable variation in the magnitude of these effects. Smaller but more consistent advantages of psychological preparation were found on pain medication and length of hospital stay. The interventions which had the most widespread overall effects on all the outcomes were found to be procedural information provision and behavioral instructions. Relaxation was also found to have beneficial effects on the various outcomes. Whereas Mathews and Ridgeway (1984) had indicated that cognitive coping interventions were most likely to have the greatest efficacy, the meta-analysis results show that their effects appear to be restricted to specific outcomes. Thus, cognitive interventions have been shown to have positive effects on negative affect, pain and use of pain medication, and clinical recovery but do not appear to result in shorter lengths of stay or in improved physiological indices or behavioral recovery. Surprisingly, in view of the importance attached to patient evaluations of health care, only a few studies have examined the effects of these interventions on patient satisfaction but these show quite positive results, indicating that patients view them as acceptable and helpful.

In summary, there is considerable evidence to indicate that different types of psychological preparation can not only reduce the anxiety, stress, and pain involved in many medical procedures, but also that there are considerable related benefits (e.g., less analgesia, better recovery, faster discharge, etc.). Although the above outline of approaches provided separate descriptions of each, they can easily be used in conjunction and often are.

What is encouraging from the research and reviews of psychological preparation for surgery is that they show that it is possible to intervene effectively using relatively uncomplicated procedures. Moreover, there is now sufficient information about their efficacy to be confident in recommending that they should be included as routine components of standard medical and nursing care for all patients undergoing surgery.

8.03.6 CONCLUSION

The outline of psychological preparations for surgery provides a very positive endpoint for this chapter. The development and success of these procedures not only constitute very concrete evidence of the importance and contribution of psychological factors in health care but also serves to link together some of the other main areas covered in this chapter. For example, the preparations can have important effects on the way in which postsurgical symptoms are interpreted and responded to by patients. Indeed, the efficacy of the interventions provides additional evidence of the role of self-regulatory processes and expectations in symptom perception since the provision of sensory information can provide the patient with a benign interpretation of predictable postsurgical symptoms. Moreover, it is clear that whatever specific mode of intervention is used, each of them involves aspects of communication between the HCP and the patient.

At the beginning of this chapter a brief mention was made of the variety of systems which have evolved for the delivery of health care in different countries. The psychological processes and topics covered here are of equal relevance across all these systems. It is notable that Johnston and Vogele (1993) end their meta-analysis with a similar observation since they find that psychological preparations for surgery have produced similar results in countries with both free market and public health-care provisions. The ubiquitous role of psychological processes in health care is a reflection of the relevance and the potential of research and practice in health psychology.

ACKNOWLEDGMENTS

I am particularly grateful to Susan Pybus for her help in producing this manuscript. Special thanks are also due to Rob Horne, Mary Rome, and Sue Williams for their contribution to specific parts of the chapter. Finally, thanks are also due to the editors for their encouragement, critical feedback, and patience.

8.03.7 REFERENCES

Ajzen, I. (1988). *Attitudes, personality, and behaviour*. Buckingham, UK: Open University Press.

Allan, D., & Armstrong, D. (1984). Patient attitudes towards radiographic examinations involving contrast media. *Clinical Radiology, 35,* 457–459.

Al Mahdy, H., & Seymour, D. G. (1990). How much can elderly patients tell us about their medications? *Postgraduate Medical Journal, 66,* 116–121.

Armstrong, D. (1989). *An outline of sociology as applied to medicine*. Bristol, UK: J. Wright.

Bales, R. F. (1950). *Interaction process analysis: A method for the study of small groups*. Cambridge, MA: Addison-Wesley.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Banks, M. H., Beresford, S. A., Morell, D. C., Waller, J. J., & Watkins, C. J. (1975). Factors influencing demand for primary care in women aged 20–44 years. *International Journal of Epidemiology, 43,* 189–195.

Baumann, L., Cameron, L. D., Zimmermann, R., & Leventhal, H. (1989). Illness representations and matching labels with symptoms. *Health Psychology, 8,* 449–469.

Becker, M. H. (1979). Understanding patient compliance: The contributions of attitudes and other psychosocial factors. In S. J. Cohen (Ed.), *New directions in patient compliance*. Toronto, Canada: Lexington.

Becker, M. H., Radius, S. M., Rosenstock, I. M., Drackman, R. H., Schuberth, K. C., & Teets, R. C. (1978). Compliance with a medical regimen for asthma: A test of the health belief model. *Public Health Reports, 93,* 268–277.

Bennett, G., & Millard, M. (1985). Compliance with relaxation training: The effect of providing information. *Behavioural Psychotherapy, 13,* 110–119.

Bennett, P. (1994). Psychological care of the coronary patient. *Journal of Mental Health, 3,* 477–484.

Bishop, G. D. (1991). Understanding the understanding of illness: Lay disease representations. In J. A. Skeleton & R. T. Croyle (Eds.), *Mental representation in health and illness* (pp 32–59). New York: Springer-Verlag.

Bosley, C. M., Fosbury, J. A., & Cochrane, G. M. (1995). The psychological factors associated with poor compliance with treatment in asthma. *European Respiratory Journal, 8,* 899–904.

Boumans, N. P., & Landeweerd, J. A. (1994). Working in an intensive or non-intensive care unit: does it make any difference? *Heart and Lung, 23,* 71–79.

Bowman, M., Herndon, A., Sharp, P., & Dignan, M. (1992) Assessment of the patient–doctor satisfaction scale for measuring patient satisfaction. *Patient Education and Counselling, 19,* 75–80.

Britten, N. (1994). Patients' ideas about medicines: A qualitative study in a general practice population. *British Journal of General Practice, 44,* 465–468.

Brownlee-Duffeck, M., Peterson, L., Simonds, J. F., Goldstein, K., Kilo, C., & Hoette, S. (1987). The role of health beliefs in the regimen adherence and metabolic control of adolescents and adult with diabetes mellitus. *Journal of Consulting and Clinical Psychology, 55,* 139–144.

Byrne, D. (1961). The repression-sensitization scale: Rationale, reliability and validity. *Journal of Personality, 29,* 344–349.

Cameron, L. D., Leventhal, E., & Leventhal, H. (1993). Symptom representations and affect as determinants of care seeking in a community dwelling adult sample population. *Health Psychology, 12,* 171–179.

Caron, H. (1985). Compliance: The case for objective measurement. *Journal of Hypertension, 3,* 11–17.

Cartwright, A. (1994). The experience of patients and general practitioners. *Journal of the Royal Society of Medicine, 23* (Suppl.), 8–10.

Carver, C. S, & Scheier, M. F (1990). Principles of self-regulation: Action and emotion. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 3–52). New York: Guilford.

Channer, K. S., James, M. A., Papouchado, M., & Rees, J. R. (1987). Failure of a negative exercise test to reassure patients with chest pain. *Quarterly Journal of Medicine, 63,* 315–322.

Christensen, A. J., Benotsch, E. G., Wiebe, J. S., & Lawton, W. J. (1995). Coping with treatment-related stress: Effects on patient adherence in hemodialysis. *Journal of Consulting and Clinical Psychology, 63,* (3), 454–459.

Christensen, A. J., Turner, C. W., Smith, T. W., Holman, J. M., & Gregory, M. C. (1991). Health locus of control and depression in end-stage renal disease. *Journal of Consulting and Clinical Psychology, 59,* (3), 419–424.

Christensen-Szalanski, J. J. J., Beck, D. E., Christensen-Szalanski, C. M., & Koepsell, T. D. (1983). Effects of expertise and experience on risk judgements. *Journal of Applied Psychology, 68,* 278–284.

Cioffi, D. (1991). Beyond attentional strategies: A cognitive-perceptual model of somatic interpretations. *Psychological Bulletin, 109,* 25–41.

Clausen, J. A., Pfeffer, N. G., & Huffine, C. L. (1982). Help-seeking in severe mental illness. In D. Mechanic (Ed.), *Symptoms, illness behavior, and help-seeking*. New York: Prodist.

Cleary, D. J., Matzke, G. R., Aldexander, A. C., & Joy, M. S. (1995). Medication knowledge and compliance among patients receiving long-term dialysis. *American Journal of Health System Pharmacy, 52,* 1895–1900.

Cochrane, S. D., & Gitlin, M. J. (1988). Attitudinal correlates of lithium compliance in bipolar affective disorders. *The Journal of Nervous and Mental Journal Mental Disease, 176,* 457–464.

Conrad, P. (1985). The meaning of medications: Another look at compliance. *Social Science and Medicine, 20,* 29–37.

Contrada, R. J., Leventhal, E. A., & Anderson, J. R. (1994). Psychological preparation for surgery: Marshalling individual and social resources to optimise self-regulation. In S. Maes, H. Leventhal, & M. Johnston (Eds.), *International Review of Health Psychology* (Vol. 3, pp. 219–266). Chichester, UK: Wiley.

Counte, M. (1979). An examination of the convergent validity of three measures of patient satisfaction in an outpatient treatment center. *Journal of Chronic Disease, 32,* 583–588.

Croyle, R. T. (1992). Appraisal of health threats: Cognition, motivation, and social comparison. *Cognitive Therapy and Research, 16,* 165–182.

Croyle, R. T., & Barger, S. D. (1993). Illness cognition. In S. Maes, H. Leventhal & M. Johnston (Eds.), *International Review of Health Psychology* (pp. 29–47). Chichester, UK: Wiley.

Croyle, R. T., Sun, Y. C., & Hart, M. (1997). Processing risk factor information: Defensive biases in health-related judgements and memory. In K. J. Petrie & J. Weinman (Eds.), *Perceptions of health and illness: Current research and applications* (pp. 267–290). London: Harwood.

Croyle, R. T., Sun, Y. C., & Louie, D. H. (1993). Psychological minimisation of cholesterol test results: Moderators of appraisal in college students and community residents. *Health Psychology, 12,* 503–507.

Croyle, R. T., & Uretsky, M. B. (1987). Effects of mood on appraisal of health status. *Health Psychology, 6,* 239–253

Cummings, K. M., Becker, M. H., Kirscht, J. P., & Levin, N. W. (1981). Intervention strategies to improve compliance with medical regimens by ambulatory haemodialysis patients. *Journal of Behavioural Medicine, 4,* 111–127.

Davison, K. P., & Pennebaker, J. W. (1997). Virtual narratives: Illness representations in online support groups. In K. J. Petrie & J. Weinman (Eds.), *Perceptions of health and illness* (pp. 463–486). London: Harwood.

Devins, G. M., Binik, Y. M., Hollomby, D. J., Barre, P. E., & Guttmann, R. D. (1981). Helplessness and depression in end-stage renal disease. *Journal of Abnormal Psychology, 90(6),* 531–545.

Donovan, J. L., & Blake, D. R. (1992). Patient non-compliance: Deviance or reasoned decision making? *Social Science and Medicine, 34,* 507–513.

Dracup, K., Moser, D. K., Eisenberg, M., Meischke, H., Alonzo, A. A., & Braslow, A. (1995). Causes of delay in seeking treatment for health attack symptoms. *Social Science and Medicine, 40,* 379–392.

Dunbar, J. M., & Agras, W. S. (1980). Compliance with medical instructions. In J. M. Ferguson & C. B. Taylor (Eds), *Comprehensive handbook of behavioural medicine* (Vol. 3). New York: Spectrum.

Eagleton, J. M., Walker, F. S., & Barber, N. D. (1993, July). An investigation into patient compliance with hospital discharge medication in a local population. *International Journal of Pharmacy Practice,* 107–109.

Echabe, A. E., Guillen, C. S., & Ozamiz, J. A. (1992). Representations of health, illness and medicines: Coping strategies and health promoting behaviour. *British Journal of Clinical Psychology, 31,* 339–349.

Eraker, S. A., & Kirscht, J. P., & Becker, M. H. (1984). Understanding and improving patient compliance. *Annals of Internal Medicine, 100,* 258–268.

Eysenck, M. W., & Keane, M. T. (1995). *Cognitive psychology* (3rd ed.). Hove, UK: Erlbaum.

Facione, N. C. (1993). Delay versus help seeking for breast cancer symptoms: A critical review of the literature on patient and provider delay. *Social Science and Medicine, 36,* 1521–1534.

Fallowfield, L., Hall, A., Maguire, G. P., & Baum, M. (1990). Psychological outcomes of different treatment policies in women with early breast cancer outside a clinical trial. *British Medical Journal, 301,* 575–580

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition.* New York: McGraw-Hill.

Fitzpatrick, R. (1997). Patient satisfaction. In A. Baum, S. Newman, J. Weinman, R. West, & I. C. MacManus (Eds.), *Cambridge handbook of psychology, health and medicine.* Cambridge, UK: Cambridge University Press.

Fitzpatrick, R., & Hopkins, A. (Eds) (1993). *Measurement of patients' satisfaction with their care.* London: Royal College of Physicians of London.

Frasure-Smith, N., Lesperance, F., & Talajic, M. (1995). Depression and 18-month prognosis after myocardial infarction. *Circulation, 91,* 999–1005.

Freer, C. B. (1980). Self-care: A health diary study. *Medical Care, 18,* 853–861.

Friedrikson, L. G. (1993). Development of an integrative model for the medical consultation. *Health Communication, 5,* 225–237.

George, C. F., Waters, W. E., & Nicholas, J. A. (1983). Prescription information leaflets: A pilot study in general practice. *British Medical Journal, 287,* 1193–1196.

Gibbs, S. W., & George, C. F. (1990). Communicating information to patients about medicine. Prescription information leaflets: A national survey. *Journal of the Royal Society of Medicine, 83,* 292–297.

Glasgow, R. E., McCaul, K. D., & Schafer, L. C. (1986). Barriers to adherence among persons with insulin-dependent diabetes. *Journal of Behavioural Medicine, 9,* 65–77.

Glasgow, R. E., McCaul, K. D., & Schafer, L. C. (1987). Self care and glycemic control in Type 1 diabetes. *Journal of Chronic Diseases, 40,* 399–412.

Goldberger, L. (1982). Sensory deprivation and overload. In L. Goldberger & S. Brenitz (Eds.), *Handbook of stress: Theoretical and clinical aspects.* New York: Free Press.

Good, M.-J. D., Good, B. J., & Nassi, A. J. (1983). Patient requests in primary health care settings: development and validation of a research instrument. *Journal of Behavioural Medicine, 6,* 151–168.

Gonder-Frederick, L. A., & Cox, D. J. (1991). Symptom perception, symptom beliefs and blood glucose discrimination in the self-treatment of insulin dependent diabetes. In J. A. Skelton & R. T. Croyle (Eds.), *Mental representation in health and illness.* New York: Springer-Verlag.

Gordis, L. (1976). Methodological issues in the measurement of patient compliance. In D. L. Sackett & R. B. Haynes (Eds.), *Compliance with therapeutic regimens.* Baltimore: Johns Hopkins University Press.

Greenfield, S., Kaplan, S., & Ware, J. E. (1985). Expanding patient involvement in care: Effects on patient outcomes. *Annals of Internal Medicine, 102(4),* 520–528.

Grol, R., de Maeseneer, J., Whitfield, M., & Mokkink, H. (1990). Disease-centred versus patient-centred attitudes: Comparison of general practitioners in Belgium, Britain and The Netherlands. *Family Practice, 7,* 100–104.

Haefner, D. P., & Kirscht, J. P. (1970). Motivational and behavioural effects of modifying health beliefs. *Public Health Reports, 85,* 478–484.

Hall, J., & Dornan, M. (1988). What patients like about their medical care and how often they are asked: A meta-analysis of the satisfaction literature. *Social Science Medical, 27*(9), 935–939.

Hall, J., Roter, D., & Katz, N. R. (1988). Meta-analysis of correlates of provider behaviour in medical encounters. *Medical Care, 26,* 657–675.

Hampson, S. E. (1997), Illness representations and the self-management of diabetes. In K. J. Petrie & J. Weinman (Eds.), *Perceptions of health and illness* (pp. 323–347). London: Harwood.

Hawkes, C. (1974). Communicating with the patient—an example drawn from neurology. *British Journal of Medical Education, 8,* 57–63.

Haynes, R. B. (1976). A critical review of the "determinants" of patient compliance with therapeutic regimens. In D. L. Sackett & R. B. Haynes (Eds.), *Compliance with therapeutic regimens.* Baltimore: John Hopkins University Press.

Haynes, R. B., Sackett, D. L., Gibson, E. S., Taylor, D. W., Robers, R. S., & Johnson, A. L. (1978). Patient compliance with anti-hypertensive regimens. *Patient Counselling and Health Education, 1,* 18–21.

Haynes, R. B., Taylor, D. W., & Sackett, D. L. (Eds.) (1979). *Compliance in health care*. Baltimore: John Hopkins University Press

Hilbrands, L. B., Hoitsma, A. J., & Koene, R. (1995). Medication compliance after renal transplantation. *Transplantation, 60,* 914–920.

Hogbin, B., & Fallowfield, L. J. (1989). Getting it taped: The bad news consultation with cancer patients. *British Journal of Hospital Medicine, 41,* 330–333.

Horne, R. (1993). One to be taken as directed: Reflections on non-adherence (non-compliance). *Journal of Social and Administrative Pharmacy, 10,* 150–156.

Horne, R. (1997). Representations of medication and treatment: Advances in theory and measurement. In K. J. Petrie & J. Weinman (Eds.), *Perceptions of health and illness* (pp. 155–188). London: Harwood.

Horne, R., & Weinman, J. (1994). Illness cognitions: Implications for the treatment of renal disease. In H. McGee & C. Bradley (Eds.), *Quality of life following renal failure: Psychosocial challenges accompanying high technology medicine* (pp. 113–132). London: Harwood.

Horwitz, R. I., & Horwitz, S. M. (1993). Adherence to treatment and health outcomes. *Archives of Internal Medicine, 153,* 1863–1868.

Illingworth, R. S., & Holt, K. S. (1955). Children in hospital. *Lancet, 2,* 1257–1262.

Inui, T. S., Carter, W. B., Kukull, W., & Haigh, V. H. (1982). Outcome-based doctor–patient interaction analysis: I: Comparison of techniques. *Medical Care, 20,* 535–549.

Isen, A., Rosenzweig, A., & Young, M. J. (1991). The influence of positive affect on clinical problem solving. *Medical Decision Making, 11*(3), 221–227.

Jamerson, P. A., Scheibmeir, M., Bott, M. J., Crighton, F., Hinton, R. H., & Cobb, A. K. (1996). The experiences of families with a relative in the intensive care unit. *Heart and Lung, 25*(6), 467–474.

Janz, N. K., & Becker, M. H. (1984). The health belief model: A decade later. *Health Education Quarterly, 11,* 1–47.

Johnson, J. E., Morrissey, J. F., & Leventhal, H. (1973). Psychological preparation for an endoscopic examination. *Gastrointestinal Endoscopy, 19,* 180–182.

Johnston, M. (1986). Preoperative emotional states and postoperative recovery. *Advances in Psychosomatic Medicine, 15,* 1–22.

Johnston, M., & Vogele, C. (1993). Benefits of psychological preparation for surgery: A meta-analysis. *Annals of Behavioural Medicine, 15*(4), 245–256.

Jones, P. K., Jones, S. L., & Katz, J. (1987). Improving compliance for asthma patients visiting the emergency department using a health belief model intervention. *Journal of Asthma, 24,* 199–206.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.

Kaplan De-Nour, A. (1994). Psychological, social and vocational impact of renal failure: A review. In H. McGee & C. Bradley (Eds.), *Quality of life following renal failure: Psychosocial challenges accompanying high technology medicine* (pp. 33–42). London: Harwood.

Keefe, F. J., Brown, G. K., Wallston, K. A., & Caldwell, D. S. (1989). Coping with rheumatoid arthritis pain: Catastrophising as a maladaptive coping strategy. *Pain, 37,* 51–6.

Keep, P., James, J., & Imman, M. (1980). Windows in the intensive therapy unit. *Anaesthesia, 35,* 257–262.

Keller, M. L., Leventhal, H., Prohaska, T. R., & Leventhal, E. A. (1989). Beliefs about ageing and illness in a community sample. *Research in Nursing and Health, 12,* 247–255.

Kelly, G. R., Mamon, J. A., & Scott, J. E. (1987). Utility of the health belief model in examining medication compliance among psychiatric outpatients. *Social Science and Medicine, 25,* 1205–1211.

Kelly, T. A. (1990). The roles of values in psychotherapy: A critical review of process and outcome effects. *Clinical Psychology Review, 10,* 171–180.

Kelly, T. A., & Strupp, H. H. (1992). Client and therapist values in psychotherapy: Perceived changes, assimilation, similarity and outcome. *Journal of Consulting and Clinical Psychology, 60,* 34–40.

Kendall, P. C., Williams, L., Pechacek, T. F., Graham, L. E., Shisslak, C., & Herzoff, N. (1979). Cognitive-behavioral and patient education interventions in cardiac catheterization procedures: The Palo Alto Medical Psychology Project. *Journal of Consulting and Clinical Psychology, 47,* 49–58.

Kendrick, T. (1997). Teaching communication skills. In A. Baum, S. Newman, J. Weinman, R. West, & I. C. MacManus (Eds.), *Cambridge handbook of psychology, health and medicine* (pp. 322–325). Cambridge, UK: Cambridge University Press.

Kinchla, R. A. (1992). Attention. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual Review of Psychology* (Vol. 43). Palo Alto, CA: Annual Reviews Inc.

Kindelan, K., & Kent, G. (1987). Concordance between patients' information preferences and general practitioners' perceptions. *Psychology and Health, 1,* 399–409.

Kinnersley, P., Stott, N., Peters, T., Harvey, I., & Hackett, P. (1996). A comparison of methods for measuring patient satisfaction with consultations in primary care. *Family Practice, 13,* 41–49.

Kirmayer, L. J., Young, A., & Robbins, J. M. (1994). Symptom attribution in cultural perspective. *Canadian Journal of Psychiatry, 39,* 584–595.

Koenig, H. G., George, L. K., Stangl, D., & Tweed, D. L. (1995). Hospital stressors experienced by elderly medical in-patients: Developing a Hospital Stress Index. *International Journal of Psychiatry in Medicine, 25,* 103–122.

Korsch, B. M., & Negrete, V. F. (1972). Doctor–patient communication. *Scientific American, 227,* 66–74.

Krantz, D., Baum, A., & Wideman, M. (1980). Assessment of preferences for self-treatment and information in health care. *Journal of Personality and Social Psychology, 39,* 977–990.

Kruse, W., Nikolaus, T., Rampmaier, J., Weber, E., & Schlierf, G. (1993). Actual versus prescribed timing of lovastatin doses assessed by electronic compliance monitoring. *European Journal of Clinical Pharmacology, 45,* 211–215.

Kulik, J. A., & Mahler, H. I. (1987). Effects of preoperative room-mate assignment on preoperative anxiety and recovery from coronary by-pass surgery. *Health Psychology, 6,* 525–543.

Kulik, J. A., Mahler, H. I., & Moore, P. J. (1996). Social comparison and affiliation under threat: effects on recovery from major surgery. *Journal of Personality and Social Psychology, 71*(5), 967–979.

La Greca, A. M., & Stone, W. (1985). Behavioral pediatrics. In N. Scheiderman & J. T. Tapp (Eds.), *Behavioral medicine: The biopsychosocial approach*. Hillsdale, NJ: Erlbaum.

Langer, E. J., Janis, I. L., & Wolfer, J. A. (1975). Reduction of psychological stress in surgical patients. *Journal of Experimental Social Psychology, 11,* 155–165.

Last, J. M. (1963). The clinical iceberg: Completing the clinical picture in general practice. *Lancet, ii,* 28–30.

Leenars, P. E. M., Rombouts, R., & Kok, G. (1993). Seeking medical help for a sexually transmitted disease: Determinants of delay behaviour. *Psychology and Health, 8,* 17–32.

Leventhal, E., & Crouch, M. (1997). Are there differences in perceptions of illness across the lifespan? In K. Petrie

& J. Weinman (Eds.), *Perceptions of health and illness* (pp. 77–102). London: Harwood.

Leventhal, E., Leventhal, H., Shacham, S., & Easterling, D. V. (1989). Active coping reduces reports of pain in childbirth. *Journal of Consulting and Clinical Psychology, 57,* 365–371.

Leventhal, H. (1985). Role of theory in the study of adherence to treatment and doctor–patient interactions. *Medical Care, 23,* 556–563.

Leventhal, H., Benyamini, Y., Brownlee, S., Diefenbach, M., Levethal, E. A., Patrick-Miller, P. & Robitaille, C. (1997). Illness perceptions: Theoretical Foundations. In K. J. Petrie & J. Weinman (Eds.), *Perceptions of health and Illness* (pp. 19–46). London: Harwood.

Leventhal, H., & Cameron, L. D. (1987). Behavioural theories and the problem of compliance. *Patient Education and Compliance, 10,* 117–138.

Leventhal, H., & Diefenbach, M. (1991). The active side of illness cognition. In J. A. Skelton & R. T. Croyle (Eds.), *Mental representation in health and illness*. New York: Springer.

Leventhal, H., Diefenbach, M., & Leventhal, E. A. (1992). Illness cognition—using common sense to understand treatment adherence and affect cognition interactions. *Cognitive Therapy and Research, 16,* 143–163.

Leventhal, H., Nerenz, D. R., & Steele, D. J. (1984). Illness representations and coping with health threats. In A. Baum, S. E. Taylor, & J. E. Singer (Eds.), *Handbook of psychology and health, Vol. 4, Social psychological aspects of health*. Hillsdale, NJ: Erlbaum.

Levy, N. B. (1994). Psychological aspects of renal transplantation. *Psychosomatics, 35,* 427–433.

Ley, P. (1972). Complaints by hospital staff and patients: A review of the literature. *Bulletin of the British Psychological Society, 25,* 115–120.

Ley, P. (1988). *Communicating with patients*. Improving communication, satisfaction and compliance. London: Croom Helm.

Ley, P. (1997). Doctor–patient communication. In A. Baum, S. Newman, J. Weinman, R. West, & I. C. MacManus (Eds.), *Cambridge handbook of psychology, health and medicine*. Cambridge, UK: Cambridge University Press.

Ley, P., Bradshaw, P. W., Kincey, J. A., & Atherton, S. T. (1976). Increasing patients' satisfaction with communication. *British Journal of Social and Clinical Psychology, 15,* 403–413.

Ley, P. & Llewellyn, S. (1995). Improving patients' understanding, recall, satisfaction and compliance. In: A. Broome & S. Llewellyn (Eds.), *Health psychology: Process and applications* (2nd ed., pp 75–98). London: Chapman Hall.

Lilleyman, J. S., & Lennard, L. (1996). Non-compliance with oral chemotherapy in childhood leukaemia: An overlooked and costly cause of late relapse. *British Medical Journal, 313,* 1219–1220.

Lorber, J. (1975). Good patients and problem patients: Conformity and deviance in a general hospital. *Journal of Health and Social Behaviour, 16,* 213–225.

Lorish, C. D., Richards, B., & Brown, S. (1990). Perspective of the patient with rheumatoid arthritis on issues related to missed medication. *Arthritis Care and Research, 3,* 78–84.

Lucock, M. P., Morley, S., Peake, M. D., & White, C. (in press). Responses to reassurance: Time course and prediction of response. *British Medical Journal.*

McCann, S., & Weinman, J. (1996). Encouraging patient participation in general practice consultations: Effect on consultation length and content, patient satisfaction and health. *Psychology and Health, 11,* 857–869.

McDonald, I. G., Daly, J., Jelinek, V. M., Panetta, F., & Gutman, J. M. (1996). Opening Pandora's box: The unpredictability of reassurance by a normal test result. *British Medical Journal, 313,* 329–333.

McGee, H. (1997). Attitudes of health care professionals. In A. Baum, S. Newman, J. Weinman, R. West, & I. C. MacManus (Eds.), *Cambridge handbook of psychology, health and medicine*. Cambridge, UK: Cambridge University Press.

McManus, I. C., Vincent, C. A., Thom, S., & Kidd, J. (1993). Teaching communication skills to clinical students. *British Medical Journal, 306,* 1322–1327.

Maguire, P. (1997), Breaking bad news. In A. Baum, S. Newman, J. Weinman, R. West, & I. C. MacManus (Eds.) *Cambridge handbook of psychology, health and medicine*. Cambridge, UK: Cambridge University Press.

Maguire, G. P., Fairburn, S., & Fletcher, M. (1986). Consultation skills of young doctors: Benefits of feedback training in interviewing as students persist. *British Medical Journal, 292,* 1573–1576.

McKim, W. A., Stones, M. J., & Kozma, A. (1990). Factors predicting medicine use in institutionalized and noninstitutionalized elderly. *Canadian Journal on Aging, 9,* 23–34.

Marteau, T. M. (1990). Attitudes to doctors and medicine: The preliminary development of a new scale. *Psychology and Health, 4,* 351–356.

Marteau, T. M., & Johnston, M. (1990). Health professionals: A sources of variance in health outcomes. *Psychology and Health, 5(1),* 47–58.

Marteau, T. M., & Riordan, D. C. (1992). Staff attitudes towards patients: The influence of causal attributions for illness. *British Journal of Clinical Psychology, 31,* 107–110.

Mathews, A., & Ridgeway, V. (1984). Psychological preparation for surgery. In A. Steptoe & A. Mathews (Eds.), *Health care and human behaviour* (pp. 231–259). London: Academic Press.

Meichenbaum, D., & Turk, D. C. (1987). *Facilitating treatment adherence: A practitioner guidebook*. New York: Plenum.

Meyer, D., Leventhal, H., & Gutmann, M. (1985). Common-sense models of illness: The example of hypertension. *Health Psychology, 4,* 115–135.

Miller, S. M., Brody, D. S., & Summerton, J. (1988). Styles of coping with threat: Implications for health. *Journal of Personality and Social Psychology, 54,* 142–148.

Monane, M., Bohn, R. L., Gurwitz, J. H., Glynn, R. J., & Avorn, J. (1994). Non-compliance with congestive heart failure therapy in the elderly. *Archives of Internal Medicine, 154,* 433–437.

Morgan, M., & Watkins, C. J. (1988). Managing hypertension: Belief and responsed to medication among cultural groups. *Sociology of Health and Illness, 10*(4), 561–578.

Moriskey, E. E., Green, L. W., & Levine, D. M. (1986). Concurrent and predictive validity of a self-reported measure of medication adherence. *Medical Care, 24,* 67–74.

Morrell, D. C., Evans, M. E., Morris, R. W., & Roland, M. O. (1986). The five minute consultation: Effect of time constraint on clinical content and patient satisfaction. *British Medical Journal, 292,* 870–873.

Moss-Morris, R. (1997). The role of illness cognitions and coping in the aetiology and maintenance of the chronic fatigue syndrome. In K. J. Petrie & J. Weinman (Eds.), *Perceptions of health and illness: Current research and applications* (pp. 411–439). London: Harwood.

Mullen, B., & Suls, H. (1982). The effectiveness of attention and rejection of coping styles: A meta-analysis of temporal differences. *Journal of Psychosomatic Research, 26,* 43–49.

Nelson, E. C., Stason, W. B., Neutra, R. R., Soloman, H. S., & McArdle, P. J. (1978). Impact of patient perceptions on compliance with treatment for hypertension. *Medical Care, 16,* 893–906.

Oldenburg, B., Macdonald, G. J., & Perkins, R. J. (1988). Prediction of quality of life in a short of end-stage renal disease patients. *Journal of Clinical Epidemiology, 41*(6), 555–564.

Pendleton, D., Schofield, T., Tate, P., & Havelock, P. (1984). *The consultation: An approach to teaching and learning*. Oxford, UK: Oxford University Press.

Pennebaker, J. W. (1982). *The psychology of physical symptoms*. New York: Springer-Verlag.

Pennock, B. E., Crawshaw, L., Maher, T., Price, T., & Kaplan, P. D. (1994). Distressful events in the ICU as perceived by patients recovering from coronary artery bypass surgery. *Heart and Lung, 23*(4), 3323–3327.

Petrie, K. J., Weinman, J., Sharpe, N., & Buckley, J. (1996). Role of patient's view of their illness in predicting return to work and functioning following myocardial infarction. *British Medical Journal, 312,* 1191–1194.

Phillips, E. L. (1988). *Patient compliance: New light on health delivery systems in medicine and psychotherapy*. Lewiston, NY: Hans Huber.

Power, M., & Dagleish, T. (1997). *Cognition and emotion from order to disorder*. Hove, UK: Erlbaum.

Reid, L. D., & Christensen, D. B. (1988). A psychosocial perspective in the explanation of patients' drug-taking behavior. *Social Science and Medicine, 27,* 277–285.

Ridgeway, V., & Mathews, A. (1982). Psychological preparation for surgery: A comparison of methods. *British Journal of Clinical Psychology, 21,* 243–260.

Robbins, J. M., & Kirmayer, L. J. (1991). Attributions of common somatic symptoms. *Psychological Medicine, 21,* 1029–1045.

Roter, D. L. (1977). Patients participation in the patient-provider interaction: The effects of patients question asking on the quality of interaction, satisfaction and compliance. *Health Education Monographs, 5,* 281–315.

Roter, D. (1989). Which facets of communications have strong effects on outcome: A meta-analysis. In M. Stewart & D. Roter (Eds.), *Communicating with medical patients* (pp. 183–196). Newbury Park, CA: Sage.

Roter, D., & Hall, J. A. (1989). Studies of doctor patient interaction. *Annual Review of Public Health, 10,* 163–180.

Roter, D. L., Hall, J. A., & Katz, N. R. (1987). Relations between physicians' behaviors and analogue patients' satisfaction, recall and impressions. *Medical Care, 25,* 437–451.

Roter, D., Stewart M., Putnam, S. M., Lipkin, M., Stiles, W., & Inui, T. S. (1997). The patient-physician relationship: Communication patterns of primary care physicians. *Journal of the American Medical Association, 277*(4), 350–356.

Rudd, P. (1993). The measurement of compliance: Medication taking. In N. A. Krasnegor, L. H. Epstein, S. Bennett Johnson, & S. J. Yaffem (Eds.), *Developmental aspects of health compliance behavior* (pp. 185–213). Hillsdale, NJ: Erlbaum.

Rudd, P., Ahmed, S., Zachary, V., Barton, C., & Bonduelle, D. (1990). Improved compliance measures: Applications in an ambulatory hypertensive drug trial. *Clinical Pharmacology and Therapeutics, 48,* 676–685.

Sackett, D. L., & Haynes, R. B. (1976). *Compliance with therapeutic regimens*. Baltimore: John Hopkins University Press.

Safer, M. A., Tharps, Q., Jackson, T., & Leventhal, H. (1979) Determinants of three stages of delay in seeking care at a medical clinic. *Medical Care, 17,* 11–29.

Salmon, P., & Quine, J. (1989). Patients' intentions in primary care: Measurement and preliminary investigation. *Psychology & Health, 3,* 103–110.

Sanders, G. S. (1982). Social comparison and perceptions of health and illness. In G. S. Sanders & J. Suls (Eds.), *Social psychology of health and illness*. Hillsdale, NJ: Erlbaum.

Schmidt, L. (1997). The effects of hospitalisation on children. In A. Baum, S. Newman, J. Weinman, R. West, & C. McManus (Eds.), *The Cambridge handbook of psychology, health and medicine*. Cambridge, UK: Cambridge University Press.

Schwartz, S., & Griffin, T. (1986) *Medical thinking: The psychology of medical judgement and decision*. New York: Springer-Verlag.

Sensky, T., Macleod, A. K., & Rigby, A. F. (1996). Causal attributions about common somatic sensations among frequent general practice attenders. *Psychological Medicine, 26,* 641–646.

Spencer, S., & Baum, A. (1997). Architecture and health. In A. Baum, S. Newman, J. Weinman, R. West, & C. McManus (Eds.), *The Cambridge handbook of psychology, health and medicine* (pp. 70–72). Cambridge, UK: Cambridge University Press.

Starfield, B., Wray, C., Hess, K., Gross, R., Birk, P. S., & D'Lugoff, B. C. (1981). The influence of patient–practitioner agreement on outcome of care. *American Journal of Public Health, 71,* 127–131.

Stewart, M. A. (1995). Effective physician–patient communication and health outcomes: A review. *Canadian Medical Association Journal, 152,* 1423–1433.

Stiles, W. B. (1978). Verbal response modes and dimensions of interpersonal roles: A method of discourse analysis. *Journal of Personality and Social Psychology, 36,* 693–703.

Stiles, W. B. (1989). Evaluating medical interview process components: Null correlations with outcomes may be misleading. *Medical Care, 27,* 212–220.

Stimson, G. V., & Webb, B. (1975) *Going to see the doctor: The consultation process in general practice*. London: Routledge & Kegan Paul.

Strupp, H. H., & Hadley, W. S. (1979). Specific vs. nonspecific factors in psychotherapy. *Archives of General Psychiatry, 36,* 1125–1136.

Suls, J., & Fletcher, B. (1985). The relative efficacy of avoidant and non-avoidant coping strategies: A meta-analysis. *Health Psychology, 4,* 249–288.

Taylor, C. B., Agras, W. S., Schneider, J. A., & Allen, R. A. (1983). Adherence to instructions to practice relaxation exercises. *Journal of Consulting and Clinical Psychology, 51,* 952–953.

Taylor, S. E. (1979). Hospital patient behaviour: Reactance, helplessness or control. *Journal of Social Issues, 35,* 156–184.

Thompson, A., & Sunol, R. (1995). Expectations as determinants of patient satisfaction: Concepts, theory and evidence. *International Journal for Quality in Health Care, 7*(2), 127–141.

Ulrich, R. S. (1984). View through a window may influence recovery from surgery. *Science, 224,* 420–421.

Vacoarino, J. M. (1977). Malpractice: The problem in perspective. *Journal of the American Medical Association, 238,* 861–863.

Van Servellen, G., Lewis, C. E., & Leake, B. (1990). The stresses of hospitalisation amongst AIDS patients on integrated and special care units. *International Journal of Nursing Studies, 27*(3), 235–47.

Verbrugge, L. M. (1985). Triggers of symptoms and health care. *Social Science and Medicine, 20,* 855–870.

Vernon, D. T. A. (1974). Modeling and birth order in response to painful stimuli. *Journal of Personality and Social Psychology, 29,* 794–799.

Vervaeke, G. A. C., Vertommen, H., & Storms, G. (1997). Client and therapist values in relation to drop-out. *Clinical Psychology and Psychotherapy, 4,* 1–6.

Volicer, B. J., & Bohannon, M. W. (1975). A hospital stress rating scale. *Nursing Research, 24,* 352–359.

Ware, J., & Hays, R. (1988). Methods for measuring patient satisfaction with medical care. *Evaluation and Program Planning, 6,* 247.

Warwick, H. M., & Salkovskis, P. M. (1990). Hypochondriasis. *Behaviour Research and Therapy, 28,* 105–117.

Wasserman, R., & Inui, T. (1983). Systematic analysis of clinician–patient interactions: A critique of recent approaches with suggestions for future research. *Medical Care, 21,* 279–293.

Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin, 96,* 465–490.

Watson, D., & Pennebaker, J. W. (1989). Health complaints, stress and distress: Exploring the central role of negative affectivity. *Psychological Review, 96,* 234–254.

Webb, S. L. M., & Lloyd, M. (1994). Prescribing and referral in general practice: A study of patients expectations and doctors actions. *British Journal of General Practice, 44,* 165–169.

Weinman, J., & Johnston, M. (1988). Stressful medical procedures: An analysis of the effects of psychological interventions and of the stressfulness of the procedure. In S. Maes, P. Defares, I. G. Sarason, & C. D. Speilberger (Eds.), *Topics in health psychology* (pp. 205–217). Chichester, UK: Wiley.

Weinman, J., & Petrie, K. (1997). Illness perceptions: A new paradigm for psychosomatics? *Journal of Psychosomatic Research, 42,* 113–116.

Williams, S., Weinman, J., Dale, J., & Newman, S. (1995). Patient expectations: What do primary-care patients want from the GP and how far does meeting patient expectations affect patient satisfaction? *Family Practice, 12,* 193–201.

Williams, S., & Calnan, M. (1991). Key determinants of consumer satisfaction with general practice. *Family Practice, 8,* 237–242.

Wolf, M. H., Putnam, S. M., James, S. A., & Stiles W. B. (1978). The Medical Interview Satisfaction Scale: Development of a scale to measure patients perceptions of physicians behaviour. *Journal of Behavioural Medicine, 1,* 391–401.

Woller, W., Kruse, J., Winter, P., Mans, E. J., & Alberti, L. (1993). Cortisone image and emotional support by key figures in patients with bronchial asthma: An empirical study. *Psychotherapy and Psychosomatics, 59*(3–4), 190–196.

Yeung, M., O'Connor, S. A., Parry, D. T., & Cochrane, G. M. (1994). Compliance with prescribed drug therapy in asthma. *Respiratory Medicine, 88,* 31–35.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.04
Assessment and Measurement Issues

MARIE JOHNSTON and DEREK W. JOHNSTON
University of St. Andrews, UK

8.04.1 INTRODUCTION 114
  8.04.1.1 Theoretical Frameworks and Questions Investigated 114
  8.04.1.2 Measurement Approaches 115
    8.04.1.2.1 Interviews 115
    8.04.1.2.2 Questionnaires 116
    8.04.1.2.3 Observation of behavior 117
    8.04.1.2.4 Psychophysiological measurement 117
    8.04.1.2.5 Clinical and pathophysiological indices 119
    8.04.1.2.6 Health-care records 119
  8.04.1.3 Design of Investigations 119
8.04.2 MEASURES USED WITH HEALTHY PEOPLE 120
  8.04.2.1 Measuring Health 120
  8.04.2.2 Health-related Behavior 121
  8.04.2.3 Health Cognitions 121
  8.04.2.4 Predicting Disease 122
    8.04.2.4.1 Hypothesized link between behavior and disease 122
    8.04.2.4.2 Stability of behavior 122
    8.04.2.4.3 Independence of initial health status and predictor behavior 123
    8.04.2.4.4 Independence of predictor behavior and disease endpoint 123
    8.04.2.4.5 Costs and practicality of measures 123
8.04.3 MEASURES USED WITH ILL PEOPLE 123
  8.04.3.1 Measuring Illness 124
    8.04.3.1.1 Symptoms and pain 124
    8.04.3.1.2 Behavioral consequences of disease 124
    8.04.3.1.3 Emotional consequences of disease 125
    8.04.3.1.4 Quality of life 125
  8.04.3.2 Illness Cognitions 126
  8.04.3.3 Coping with Illness 127
8.04.4 MEASURES USED IN HEALTH-CARE CONTEXTS 127
  8.04.4.1 Measuring the Process of Health Care 127
    8.04.4.1.1 Communications between patients and providers 127
    8.04.4.1.2 Uptake and adherence 128
    8.04.4.1.3 Behavior of health professionals 128
    8.04.4.1.4 Psychological and behavioral interventions 128
  8.04.4.2 Health Care Cognitions 129
    8.04.4.2.1 Patients' health care cognitions 129
    8.04.4.2.2 Providers' health care cognitions 129
  8.04.4.3 Health Care Outcomes 129
8.04.5 MEASURES USED IN ALL CONTEXTS 130
  8.04.5.1 Stress, Life Events, and Emotional States 130
    8.04.5.1.1 Social deprivation 130

8.04.5.2 Resources 130
8.04.5.2.1 Social support 130
8.04.5.2.2 Individual differences 130
8.04.6 FUTURE DIRECTIONS 130
8.04.7 REFERENCES 131

8.04.1 INTRODUCTION

Most measurement issues in health psychology are similar to those in all other areas of psychology and are characterized by the problems of placing a metric on behavior and on mental states. Health psychology examines psychological and behavioral processes in health, illness, and health care and its measures are distinguished by the nature of the theoretical frameworks and questions addressed, by the diversity of measurement methods used, and by the context of the applications.

This introductory section deals with general measurement issues in health psychology. In the following sections, some of the measures used in the common areas of investigation are presented. We start with the assessment of healthy individuals, followed by measures used with people who are ill, then examine the measures adopted with both providers and users to investigate the process of health care and finally discuss measures that are used in all contexts.

8.04.1.1 Theoretical Frameworks and Questions Investigated

Health psychology measurement is undertaken in order to understand behavioral and psychological processes in health, illness, and health care. The purpose may be research or application, especially in the fields of public health or clinical medicine. Measures are used to address three main theoretically, based questions concerning:

(i) the assessment of psychological and behavioral indices of the status or amount of health, illness, or health care;

(ii) the assessment or evaluation of the psychological and behavioral consequences of health, illness, or health care;

(iii) the assessment of psychological and behavioral factors as predictors or explanations of health, illness, or health care.

In each case, the measurement of health, illness, and health care is essential.

Just as health is measured by psychological and behavioral processes, much of the assessment of illness and recovery from illness lies within this domain. Clinical signs and pathophysiological processes may be available in the assessment of patients with some conditions (such as respiratory function in chronic bronchitis, or immune function in AIDS). Frequently, however, the clinician assesses disease processes by ascertaining how the patients feel and what they are doing (as in rheumatoid arthritis or multiple sclerosis)—if patients are recovering well, they describe themselves as feeling better and doing more for themselves. Indeed, Kaplan (1990) has argued that these behavioral processes are the important outcome of illness and health care. Measurement of illness has therefore included assessment of the psychological and behavioral consequences of the disease processes.

Health care is largely a behavioral and interpersonal process. Indices of health care assess the use of services by patients, such as frequency of attendance at a clinic, as well as the performance of those involved in health care delivery (e.g., hand hygiene by surgical staff, provision of information by primary care physicians). These behavioral processes are likely to be a major influence on the effectiveness of health care as they determine the viability of delivery of biomedical treatments and advice. The behavior of patients in adhering to medical recommendations has long been recognized and methods of assessing adherence have been debated and developed, but the role of health professionals in determining patient outcomes has been less adequately assessed (Marteau & Johnston, 1991).

An important development in all industrialized countries has been the drive to develop evidence-based health care. Expansion of the need for medical care has resulted in financial pressures and the inevitable selectivity in the provision of care. It is clearly desirable to select those services which are both effective in achieving good health outcomes and efficient in their delivery, but this requires both evidence concerning which procedures or services are effective and efficient, and the successful implementation of these services. The assessment of effectiveness requires the measurement of health or health care outcomes, that is, scaling of the consequences of disease or health care. As indicated above, such measurement will frequently lie in a psychological or behavioral domain. Further, investigation of the implementation of interventions reveals that even those proven to be successful are not fully implemented; for example, appropriate thrombolysis for hospital patients has been shown to have the potential to reduce mortality, but there is evidence that it is not used appropriately in a large proportion of hospitals (Anderson et al., 1991). The assessment of health outcomes and the implementation of successful health care are key areas of health services research and are important areas of application for health psychology measures.

8.04.1.2 Measurement Approaches

As a result of the types of application, a wide range of measures is necessary. Health psychology investigations may include any of the following types of assessment to achieve measurement:

(i) interview (e.g., demographic information, self-report of behavior, beliefs, or feelings);

(ii) questionnaire, (e.g., standardized measures of individual differences such as optimism or coping);

(iii) observation of behavior (e.g., performance of self-care);

(iv) psychophysiology (e.g., cardiovascular reactivity to stressors);

(v) clinical indices (e.g., pulse rate, temperature, death);

(vi) pathophysiological indices (e.g., brain scan, blood cell counts);

(vii) information from health care records, (e.g., attendance at clinics, prescribed medication).

8.04.1.2.1 Interviews

Interviews or interviews incorporating questionnaires are the most common method of collecting data from people who are undergoing health care or who are ill. Interviews can help to standardize the presentation of the measurement materials in a way that could not be achieved in routine clinical assessments. They can also motivate and facilitate responses to long questionnaires, while at the same time being sensitive to fatigue, especially in people who are ill. Clearly, interview techniques can also be a source of error, if interviewers vary or if an interviewer's behavior is variable.

Interviews result in very diverse kinds of data. They may consist wholly or partially of closed precoded questions or they may contain open questions which require later coding. Some interviews may reproduce the format of a standardized interview, whereas others use newly designed questions. The most widely known example of a completely standardized interview in health psychology is the structured interview for the assessment of type A behavior (Friedman & Rosenman, 1964), where both the content and style of delivery of the questions is standardized as is the method of coding responses; both interviewers and coders require training and need to meet an acceptable standard of agreement with benchmark ratings. These ratings depend both on the content and style of response.

Open questions on interviews and questionnaires require subsequent work by the investigator in order to interpret the data and provide a basis for measurement. Some qualitative methods are used to explore new domains of investigation, by examining the predominant themes emerging from responses using a variety of techniques and theoretical frameworks (Michie, McDonald, & Marteau, 1996). These analyses can provide useful insights and may form a basis for the development of hypotheses or measures that can be satisfactorily validated. However, other methods are required if even the most basic level of measurement is to be achieved and this will be essential if evidence of the replicability of the methods and findings are desired. The minimal level of measurement enables the investigator to classify a response as belonging to a specified category or not. For example, in interviews with patients with motor neurone disease, a question about who was with them when they were told was reliably categorized as "told alone" or "with someone present" (Johnston, Earel, Mitchell, Morrison, & Wright, 1996). Without at least nominal or categorical measurement, it is impossible to establish the reliability, validity, frequency, associations, consequences, or causal status of the phenomena investigated.

Normally, investigators develop a coding frame from preliminary work with a subset of the data or from a preexisting theoretically or empirically-based framework. For example, the CAVE (Peterson, Buchanan, & Seligman, 1995) codes spontaneous utterances in terms of the attributions presented and classifies the material in terms of the respondent's explanatory style, whereas LIWC (Pennebaker, Mayne, & Francis, 1997) is a computer program that searches for key words reflecting the individual's attempts to find meaning. A group of four papers using different coding frames to analyze the same interview material (Folkman, 1997; Nolen-Hoeksema, McBride, & Larson, 1997; Pennebaker et al., 1997; Stein, Folkman, Trabasso, & Richards 1997) illustrates the diversity of measurement approaches that can be applied to spontaneous responses to open questions. This kind of coding frame guides the allocation of a code, usually a numerical code, to the responses. It is then possible to examine the reliability of the coding by comparing coders or the same coder on different occasions or on different parts of the data. Reliable data can then be investigated using appropriate statistical methods to examine the research questions (see Volume 3 of this series).

Where interview responses take a precoded closed format, then coding responses is not an issue. Scoring and measurement then depend on the same psychometric issues as apply to questionnaires.

8.04.1.2.2 Questionnaires

Questionnaires are widely used in health psychology to assess beliefs, attitudes, knowledge, mood, indices of individual differences, experience of health care, health-related behaviors, adherence to medical advice, etc. We have described or evaluated over 50 such measures in a user's portfolio of health psychology measures which provides a fuller account of the measures and their use (Johnston, Wright, & Weinman, 1995).

Such easily accessible measures can appear to the novice health psychology investigator as an attractive display of off-the-shelf techniques that can solve many of the problems in research design. However, there are a number of pitfalls. First, the measures vary in the quality of their psychometric properties. There are measures that have insufficient items to achieve satisfactory reliability and, for some, reliability may not even have been assessed. Some form of internal consistency should be available and, where appropriate, evidence of test–retest and interrater reliability should be provided. Measures should also have demonstrated validity. Validity is normally conceptualized within a theoretical framework that defines the content and the functioning of the construct being measured. A second pitfall for the naive researcher is the use of measures outside the theoretical framework of their development and validity. It can be perplexing when a series of quite different measures all seem to be relevant to the same domain, but this can often be explained in terms of the underlying theoretical framework. For example, many measures assess control cognitions; however, some of these are directly associated with Rotter's (1966) Social Learning Theory, others are related to Bandura's (1977, 1997) Social Cognitive Model, and others arise from Attribution Theory (Weiner, 1986). The investigator should ensure that the measures chosen are compatible with the theory guiding the research. A third pitfall lies in the use of measures with populations they were not designed for. For example, it may be tempting to use a test validated with adults in work with children or to use assessments shown to be valid with ill people in a study of a healthy community-based sample. Clearly it is important that measures are acceptable and applicable with the population being investigated.

There are even greater hurdles in designing new measures that have no previous development, evaluation, or normative data, but this may prove necessary for the exploration of new constructs, theoretical approaches, or applications. It is helpful if the theoretical formulation guides the development of measures as was done for the Theory of Reasoned Action (Fishbein & Azjen, 1975). However, the content and format of new questionnaires is likely to begin with pilot or exploratory work, often using qualitative methods of data collection (see Volume 3 of this series) to identify salient material. For example, in developing a measure of perceived control of recovery from disabling disease, we started with spontaneous comments by patients which were subsequently evaluated using standard reliability analyses and tested for construct validity (Partridge & Johnston, 1989).

The problems in using self-report interview or questionnaire data are widely recognized. Various response sets may affect responses and in health psychology, social desirability (the tendency to give responses attracting the approval of others) and negative affectivity (bias towards a negative view of self, circumstances, and events) may be particularly important. Social desirability is likely to be important where there are recommended or expected styles of behavior. For example, social desirability might lead to exaggeration of adherence to medical advice or health protective behaviors, but might result in under-reporting of health-damaging behaviors such as smoking. It is often valuable to assess and make allowance for social desirability by using an instrument such as the Marlowe–Crown Inventory as in Weinberger, Scwartz, and Davidson (1979). Watson and Pennebaker (1989) suggest that negative affectivity may be a source of error in examining the relationship between self-report measures where each measure may be influenced by the individual's tendency to focus on the negative aspect; for example, correlations between mood and health status might simply be due to individual differences in negative affectivity coloring perceptions of health status.

Although such considerations apply to all measures, they may prove particularly important where measures involve self-report. These measures are commonly used in health psychology and the sources of error in self-report measures are generally recognized (Abrahams & Hampson, 1996; Haaga, 1997). However, they may be a particularly serious source of bias if the investigator examines the relationship between two such measures. If both measures are subject to the same form of response set, such as social desirability or negative affectivity, then apparent relationships between constructs might be due entirely to the correlation in the biases in the measures.

8.04.1.2.3 Observation of behavior

Observations of behavior may be undertaken where self-report is impossible, or to supplement self-report measures. These observations may be made by health professionals, such as the clinical assessments frequently used in studies of preparation for surgery (Anderson, 1987), by members of the target person's household as in studies of smoking, or by the investigators as in the observation of limitations in the performance of daily activities following illness.

Such observations are limited by the observers' inability to see the full range of the behavior. Most health professionals will see the patient for quite short periods, and even nursing staff are typically limited to 8 or 12 hours per day. Lack of concordance between self-report and the observation of the professional may therefore be due to limitations in their observations as well as the noted problems in self-report. Nurses assess surgical patients as having more emotional problems than the patients report and are less accurate at pinpointing patients' emotional problems than other patients are (Johnston, 1982), perhaps because of the limited interactions they have with patients. Different health professionals perform different tasks and have very different styles of interaction with patients and this may result in systematic, rather than just random, variations in the assessments they make (M. Johnston et al., 1987). Thus, the choice of observer may critically affect the results obtained.

Members of the individual's household or social network may also be limited in the observations they can make. Very often they are involved in assessing the performance of undesirable, health-damaging behaviors such as smoking or consumption of fatty food, or adherence to a recommended medication, diet, or exercise regimen. The chosen observer may be unable to ascertain whether the person has participated in the target behavior when out of sight, or may be motivated to present a favorable picture of the friend or relative they are observing.

When investigators observe behavior, this is always restricted either in time or in the range of behaviors. For example, assessments of limitations of function typically require the performance of the activity, often in clinical rather than home settings. Thus, while observational methods are attractive as an adjunct to self-report methods used in interviews and questionnaires, they too have practical constraints, sources of unreliability, and factors which limit their validity.

8.04.1.2.4 Psychophysiological measurement

Many psychological processes implicated in the etiology, treatment, and assessment of disease and disease-related processes involve changes in physiological systems in response to psychological stimuli. This is particularly the case if stress, stress reduction, or emotion is involved. The measures used in particular investigations will often reflect the disease under study so, for example, heart rate and blood pressure are used in studies of heart disease (see Chapter 8.08, this volume), gastric measures in studies of gastrointestinal disorders (see Chapter 8.12, this volume), and measures of muscular tension in studies of headache (see Chapter 8.23, this volume). A general review of psychophysiological processes in disease is provided by Steptoe (Chapter 8.02, this volume).

(i) Stress testing

The actual use made of psychophysiological techniques obviously depends on the question under investigation. However, many studies involve contrasting a period when the subject is relaxed with when they are stressed. The choice of condition for both periods can be critical. Few subjects find the psychophysiological laboratory immediately relaxing and it may be difficult to control and standardize the temperature, humidity, and sound level. The choice of stressor and control tasks is also critical. The challenging video game that produces an enormous elevation in heart rate (HR) and systolic blood pressure (SBP) in a male teenager may be totally inappropriate in a late-middle-aged patient with heart disease where carefully selected role-play may be a more useful task (Ironson et al., 1992). It can also be difficult to find control tasks with similar metabolic requirements to the stressors to ensure that differences due to physical effort are not interpreted as being due to psychological stress.

(ii) Physiological measures

Some of the most commonly used psychophysiological measures are outlined below. Cacioppo and Tassiary (1990) describe most aspects of psychophysiological measurement in their excellent *Principles of psychophysiology*. The main bodily systems that are studied in health psychology using psychophysiological methods are the cardiovascular system, the respiratory system, aspects of sweating, the gastrointestinal system, and the skeletomuscular system. Hormonal measures such as adrenaline, noradrenaline, and cortisol are also often taken, as are a wide variety of measures of immune function.

(a) Cardiovascular. Cardiovascular measures are among the most widely used psychophysiological measures. This is because of their use in the study of the stress process and also because of the specific relationship between such responses and cardiovascular disease (see Chapters 8.02 and 8.08, this volume). Heart rate, the most common cardiovascular measure, is reliably recorded from the electrocardiogram (ECG), a record of the electrical activity of the heart.

There may be both intra- and interassessor unreliability in the manual assessment of blood pressure and therefore automatic assessments are better for most purposes. Blood pressure (BP) is determined by the amount of blood pumped by the heart (cardiac output) and the resistance of the blood vessels to the passage of blood (peripheral resistance). On each cardiac cycle, pressure fluctuates between a maximum value as the blood is ejected (SBP) and a minimal value just before the next heart beat (diastolic blood pressure, DBP). BP is measured in millimeters of mercury (mmHg). In healthy individuals the SBP is approximately 120 mmHg, and the DBP is 80 mmHg. BP can rise markedly during psychological stress and raised BP at rest defines hypertension.

The more specialist measures of cardiac output and peripheral blood flow are less commonly used. As there are considerable technical challenges in the direct measurement of cardiac output, the most common measurement approaches are indirect, and based on measuring changes in electrical impedance across the heart. Peripheral resistance cannot be determined directly, but can be calculated from BP and cardiac output. However, flow in specific parts of the vascular system can be measured. The most widely used method is photoelectric plethysmography, which relies on the fact that red blood cells scatter infrared light. When infrared light is passed through the skin or scattered by the blood in skin, variations in light between the light source and a detector can give an indication of blood flow. Such methods are used in some heart-rate detectors that do not use the ECG, and have been used in the study of abnormalities of blood flow in the temporal artery in headache (see Chapter 8.23, this volume). More demanding methods, based on direct measures of limb volume, can provide quantitative measures of peripheral blood flow. Freedman (1989) provides an application in Raynaud's disease. Further details on cardiovascular measurement can be found in Steptoe and Johnston (1991).

(b) Respiration. Respiration is studied in its own right in conditions with respiratory symptoms and also because variations in breathing have a profound effect on other bodily systems such as the cardiovascular system. Respiration rate can be determined from either belt-like devices round the chest that produce a electrical signal that varies with chest expansion or from thermistors placed near a nostril, used to detect the difference in temperature between inhaled and exhaled air. Other measures to indicate different types of breathing, such as in hyperventilation, require the use of several chest-movement detectors.

(c) Sweating. Sweating-related activity is a widely used psychophysiological measure. The number of sweat glands open determines variations in the resistance of the body to the passage of a small electric current between two sites. In certain areas of the skin, such as the fingertips, the main cause of sweating is arousal (rather than temperature), and therefore variations in electrodermal measures such as skin conductance provide a very useful measure of arousal. Since sweating is under purely sympathetic control, skin conductance is more readily interpreted than other measures, such as HR, which reflect the balance between the sympathetic and parasympathetic systems.

(d) Gastrointestinal. The gastrointestinal system is measured using the electrogastrogram, a record of the electrical signals associated with contractions of the gut.

(e) Musculoskeletal. Activity in the musculoskeletal system is recorded using the electromyogram, a record of the electrical impulses produced by contracting muscle fibers. This can give information on small specific muscles, such as the frontalis muscle often studied in relation to headache (see Chapter 8.23, this volume), or larger muscles relating to gross bodily movements (see Chapter 8.19, this volume).

(f) Stress hormones. The stress hormones epinephrine, norepinephrine, and cortisol can be measured in either blood, urine, or saliva. Although these measures can be illuminating, the assays involved and their interpretation is a complex matter. Difficulties can arise in the methods of collection of samples, in the storage prior to assay, and in the reliability of assays over laboratories, technicians, and time.

(g) Immune measures. The use of immunological measures, such as various types of immunoglobulin or T lymphocytes, is highly attractive in some areas of research such as infection, cancer, or HIV (see Chapters 8.10 and

8.23, this volume). These measures incur the same hazards as stress hormone assays but, additionally, this is a rapidly advancing field and the choice and interpretation of measures requires current, specialist expertise.

8.04.1.2.5 Clinical and pathophysiological indices

Medicine uses many clinical and pathophysiogical indices in screening populations, investigating disease, and charting recovery. These measures include: routine clinical measures such as temperature or pulse rate; indices of heart disease derived from the ECG or echocardiogram; measures of bacterial infection determined from samples of bodily fluids; the simple imaging techniques like the X-ray that have been in use for decades; and complex imaging of the functioning body now possible with techniques such as magnetic resonance imaging (MRI). In health psychology, one may wish to understand such measures either to have a better understanding of the medical condition of a patient in treatment or to use the measures in research. In either case it is important to ask oneself if the information is really necessary. Many psychologists enter this field with the belief that medicine has established valid and reliable measures of illness and recovery. This is not always so. Much of medicine relies on the patient's report of their symptoms. In many instances, treatments are given and terminated and patients are discharged from hospital when they indicate that they are better by their speech or actions. Other sections of this chapter indicate how such symptoms can be measured with much greater reliability and precision than is common in the medical assessment of symptoms.

If one is treating a patient with a condition like diabetes, hypertension, or asthma, one often wishes to know the severity of the condition, the prognosis, or the patient's medical progress. Information on blood sugar, BP, or peak respiratory flow may inform treatment or, at the very least, increase communication and rapport between psychologist and client. A considerable amount of such information is provided in the chapters of this volume on specific medical conditions. In addition, it is likely that one of the current editions of standard large medical textbooks designed for undergraduate use will provide more than enough information. Difficulties can arise if textbooks use language that is not appropriate to local or national diagnostic and treatment approaches, and locally available patient pamphlets may be equally valuable.

If medical test data is needed for research purposes, then the usual concerns over measurement standards apply. It is unwise to assume that medical measurement is either reliable or valid; for example, we have found that repeated measures of height in adults may not achieve high reliability. In overstretched medical environments, clinical testing is often carried out under far from ideal conditions. For example, sphygmomanometers (for measuring blood pressure) are often not serviced or even checked for long periods, and standard conditions for the taking of blood or urine samples are not always observed. The information obtained from such tests may be adequate for clinical purposes, such as preliminary screening or monitoring change, but not for research where accurate absolute values are required. It is therefore incumbent on the researcher to ensure that the measures they are using in research are of research quality, with known reliability and validity, and adequate to answer the research questions.

8.04.1.2.6 Health-care records

Health-care records may be used as sources of data on the behavior of a patient, such as in seeking health-care, the nature of medical recommendations or treatment, the results of clinical assessments, or the patient's health or illness status. As these data were typically not recorded for the purposes that the health psychology investigator has in mind, they are unlikely to be biased with respect to the question investigated. However, they may have other biases and sources of error, and are frequently incomplete and therefore an unreliable source of information. As a result, they are normally used to back up or validate other, more direct, methods, rather than as the main source of evidence.

8.04.1.3 Design of Investigations

Measurement can be improved if some thought is given to the selection and timing of measures, as well as to the measurement implications of the overall design of the investigation in the planning stage.

Investigators often use a series of measures, especially questionnaires, and then the order of presentation needs to be considered. If a measure of transitory state, such as anxiety or pain, is included, this should normally be presented first as otherwise the responses may be largely determined by the effects of completing the questionnaires. For example, if respondents are required to answer questions about their illness and its future management first, this might increase the level of anxiety reported.

Indeed, experimental approaches to mood manipulation may ask participants to recall just such events with reliable effects on mood (Fisher & Johnston, 1996a). There may still be order effects for other measures and some investigators present questionnaires in a randomized or counterbalanced order to ensure that results obtained are not simply due to order of presentation. Although this is desirable scientifically, it may sometimes reduce the acceptability to respondents if, for example, a clinical research interview does not appear to follow a logical sequence. Some error may be reduced if measures are not all completed at the same time. For example, mood might be measured on a different occasion from recall of illness events.

If a causal analysis is required, then a longitudinal design will be more effective than a cross-sectional design where it is not even possible to assert that the causal construct was assessed prior to the consequence. Experimental designs are even more effective and these may range from short-term laboratory manipulations to randomized controlled trials of therapeutic interventions. For example, the effect of perceived control on disability was examined experimentally in the laboratory by manipulating control cognitions and observing the subsequent functional limitations (Fisher & Johnston, 1996b). An example of the use of a randomized trial in investing causal explanations is provided in the recurrent coronary prevention project (Friedman et al., 1986) where cognitive-behavioral methods were used to reduce Type A behavior in the experimental group but not in the control group; finding a reduced rate of recurrent cardiac episodes in the experimental group furthered the hypothesis that Type A behavior caused coronary heart disease (CHD). Other designs have been developed specifically for research in clinical situations (see Chapter 8.05, this volume).

8.04.2 MEASURES USED WITH HEALTHY PEOPLE

This section addresses measures used in studying people who have not been selected as being ill or in need of medical care. Typically these are populations of students or community residents and may well contain many people who are ill and may even be receiving care.

Undoubtedly the two main theoretical questions addressed in these populations have been the understanding of factors influencing health-related behaviors, namely behaviors likely to affect health (see Chapter 8.01, this volume), and identifying psychological and behavioral characteristics predictive of disease, especially heart disease and cancer. A fundamental question that is sometimes addressed explicitly is how to measure health.

8.04.2.1 Measuring Health

It has proved extremely difficult to measure health (Bowling, 1991). This is hardly surprising given the lack of an agreed definition of health. The complex WHO (World Health Organisation) definition of health that is commonly adopted includes social, psychological, and physical aspects (Abelin, Brzezinski, & Carstairs, 1987). There is common agreement that health is more than the absence of illness and that it is characterized by the individual's vitality and functional capacity, their mood and outlook, and their participation in activities. Wright (1990) has conceptualized health as three dimensional: freedom from unpleasant symptoms and sensations; ability to perform conventional roles and duties; and possession of physiological, psychological, and social resources.

Assessment of absence of symptoms as an index of health has proved problematic as people who would be characterized as healthy are unlikely to be symptom-free. Even number of symptoms proves problematic, as symptom reporting is associated with mood, emotional disorder, and negative affectivity. For example, many measures of emotional disorder such as the GHQ (Goldberg & Williams, 1988) ask about the experience of somatic symptoms and have found that people with greater emotional disorder report more somatic symptoms. In a similar vein, use of health services may reflect individual patterns of responding to symptoms, rather than indexing health *per se*. Watson and Pennebaker (1989) suggest that there are individual differences in negative affectivity, the tendency to attend to and report negative experiences including sensations which might be interpreted as symptoms. They have developed a measure, the PILL, which assesses the tendency to endorse symptoms that can be used to allow for negative affectivity. Thus measures of symptoms or health-care usage may be misleading as an index of "health."

Simple self-report measures of health, such as ratings on four-point verbal rating scales from excellent to poor, have been found to be predictive of future mortality (Idler, 1992). Even when allowance is made for concurrent illness, people rating their health as poor are more likely to be dead at follow-up than those giving higher ratings. These results suggest that such ratings must have some validity as measures of health. Their main limitation is the lack of sensitivity inherent in such brief scales.

Measures of general well-being such as the Bradburn Affect Balance Scale (Bradburn, 1969) or even measures of satisfaction with life (Diener, Emmons, Larson, & Griffin, 1985) may come closer to the WHO concept of health as they allow the individual to evaluate all aspects of functioning. It is curious to note that the measurement of health-related quality of life has mainly been investigated in ill, rather than healthy, populations as discussed below. These measures are rarely evaluated with healthy populations and are therefore likely to be relatively insensitive to variations in their health.

Health may also be characterized as an expectation of continuing absence of illness, either due to engagement in health-protective behaviors such as taking exercise, or to having few risk factors for disease. Risk factors may be genetic (e.g., heart disease or diabetes), physiological (e.g., high blood pressure), or environmental (e.g., living in an area of high air pollution). However, many risk factors are behavioral, such as smoking or diet, and these may be assessed along with health-protective behaviors as part of the measurement of health.

8.04.2.2 Health-related Behavior

Two kinds of health-related behaviors can be identified. The first group of behaviors are those that have been demonstrated to confer health risk or disease protection. Evidence from the Alameda County study (Schoenborn, 1993) and from other epidemiological studies demonstrates that some behaviors, especially smoking (Doll, Peto, Wheatley, Gray, & Sutherland, 1994), increase risk, while others, such as engaging in exercise, appear protective. Clinical and public health interventions are frequently directed at changing behavior to reduce risk and enhance protective behaviors (Oxcheck Study Group, 1994; Family Heart Study Group, 1994; see Chapter 8.29, this volume). The behavioral assessments found to predict disease end-points in epidemiological studies have demonstrated predictive validity and should be used in order to assess risk in a population. For example, comparison with risk data from the British Regional Heart Study (Shaper et al., 1982; Shaper, Pocock, Phillips, & Wacker, 1987) is only valid if questioning about smoking and diet follow the methods used in that study.

The other group of behaviors are the health behaviors, defined as those behaviors that people undertake in the belief that they will prevent disease or ensure detection prior to symptoms (Kasl & Cobb, 1966). There is ample evidence that most people report engaging in a number of these behaviors (Amir, 1987; Cox et al., 1987; Harris & Guten, 1979). Several checklists of these behaviors have been published (Amir, 1987; Prochaska, Leventhal, Leventhal, & Keller, 1985), but unlike the risk/protective behaviors that can be validated against epidemiological findings, there is little evidence of validation of these checklists. Self-reports of behavior are likely to be subject to bias, especially social desirability bias, as discussed above. Until they are validated against some other measure of health behaviors, their value as assessments of health behaviors is restricted and their primary value is in assessing the rate of reporting behaviors.

8.04.2.3 Health Cognitions

A major group of theoretical models have been developed to predict health-related behaviours from health-relevant social cognitions (see Chapter 8.01, this volume; Conner & Norman, 1996). Models such as the Health Belief Model, Social Cognitive Theory, and the Theory of Planned Behavior propose that a range of health and health-behavior cognitions determine health-relevant behaviors. Health cognitions include outcome expectancy, perceived severity, perceived vulnerability, health locus of control, and health value, whereas health-behavior-relevant cognitions include response efficacy, self-efficacy, behavioral intention, implementation intention, and action plans.

These models raise some important measurement issues: the individual may not be aware of the cognitive processes influencing behavior; it may be difficult to operationalize all of the model's constructs; and the problems with self-report measures have been noted. Further, the models vary in the precision of specification of measurement of core constructs, ranging from the proscriptive Theory of Planned Behavior to the very poorly defined constructs of the Health Belief Model. Some constructs have standardised measures that have psychometric validation and published norms. The best known of these is the Multidimensional Health Locus of Control (MHLC) scale (Wallston, Wallston, & Devellis, 1978), which assess three components: internal, powerful others, and chance locus of control.

Although this measure has been criticized for its lack of specific relevance for people who are ill, it continues to be appropriate for healthy populations. Wallston (1992) has proposed that locus of control by itself should not be enough to predict health behavior and proposed that it should be combined with a new construct, perceived health competence, a construct akin to generalized health self-efficacy (Smith, Wallston, & Smith 1995). The social learning theory, from which locus of control derived, would suggest that the MHLC should only predict health behavior in combination with a measure of health value (Lau, Hartman, & Ware, 1986) and there has been criticism of investigations in this area for using the model inappropriately.

This relates to a more general methodological problem in this area of investigation—the selection of variables from models without addressing complete models. The investigator needs to have a clear objective. If the intention is simply to explain as much behavior as possible, then the investigator is justified in choosing the most likely combination of variables from whichever models seem appropriate. If the objective is to test a model, then all the constructs of the model, and none from other models, should be assessed. If the aim is to test the power of the model to explain behavior, then only the proximal determinants specified by the model need to be investigated; so, for example, in the Theory of Planned Behavior (see Chapter 8.01, this volume), only behavioral intention and perceived behavioral control would need to be assessed as all other variables act through these two.

8.04.2.4 Predicting Disease

The main issues of measurement in predicting disease from individual differences in behavior are:

(i) the specificity of the hypothesized link between behavior and disease, because this affects the choice of behavioral measure and disease endpoint;

(ii) the stability of the behavior over time or environment;

(iii) the independence of initial health status and the predictor measure;

(iv) the independent measurement of predictor and endpoint;

(v) the costs and practicality of the measures.

8.04.2.4.1 Hypothesized link between behavior and disease

For much of the twentieth century there have been two competing views on the psychological factors thought to predict disease. One camp, which was most clearly manifest in the writings of the psychoanalytically influenced psychosomaticists of the 1930s and 1940s, held that there was a very specific relationship between particular personality types and specific diseases. The alternative view, exemplified by the theories of Cannon and Selye, maintained that disease was the result of general stress-related processes. More sophisticated versions of the same positions can be seen today (see Chapter 8.02, this volume) and lead to very different measures. Type A, or coronary prone, behavior (Friedman & Rosenman, 1964), and the more recent emphasis on hostility (Hecker, Chesney, Black, & Frautschi, 1988), or even type of hostility (Siegman, Dembroski, & Ringel, 1987), is a good example of a specific personality/disease theory. Friedman and Booth-Kewley's (1987) concept of the disease-prone personality is a version of a very general theory, as is Seligman's theory of a pessimistic explanatory style (Peterson et al., 1985).

Friedman and Rosenman (1964) described Type A behavior (TAB) as a pattern of competitive, time-pressured, and aggressive behavior that led to CHD and developed a specific structured interview to assess it. They held that TAB was best measured by an observing behavior (during the interview) and not by questionnaire. In the landmark Western Collaborative Group study (WCGS), they showed that, in accord with their theory, TAB predicted a specific endpoint, CHD, independently of the classical risk factors for heart disease (Rosenman et al., 1975). Friedman and Rosenman's emphasis on specificity was extended by later researchers who showed the concept of TAB was too broad and that hostility (Hecker et al., 1988) or perhaps even a specific type of expressive hostility (Siegman et al., 1987) was the main psychological predictor of CHD. In marked contrast, Friedman and Booth-Kewley (1987) hypothesize that there is a general tendency to experience negative emotions which leads to enhanced risk of disease of many kinds. They are able to summarize an extensive literature describing many different negative psychological states, such as anxiety, depression, or anger, and relate them to many different diseases, such as CHD, asthma, or cancer. These two types of theory lead to markedly different measurement concerns with one requiring very specific behavioral measures, whereas the other is best served by very broad measures of negative emotion. Similarly, one theory requires careful measurement of a specific endpoint, CHD, whereas the other requires measures of all causes of morbidity or even total mortality.

8.04.2.4.2 Stability of behavior

Typically, theories in this field imply that an enduring trait or style exacerbates psychophysiological processes which result in disease.

Thus, both the predictive measure and the underlying trait being measured have to be stable over the prolonged time periods likely to be necessary for disease to develop. This may be plausible for the personality characteristics described by Friedman and Booth-Kewley, but is less certain for TAB. Friedman and Rosenman held that TAB was a product of the subject's environment and would change if the environment changed. It is the case that TAB in the WCGS was predictive of CHD at 8.5 years postmeasurement (Rosenman et al., 1975) but not at 22 years (Ragland & Brand, 1988) and there are some indications that most of the predictive power of TAB falls in the first few years after assessment. There are predictors that are thought to be specifically short term such as vital exhaustion that predicts myocardial infarction best in the year or two after assessment (Appels & Otten, 1992).

8.04.2.4.3 Independence of initial health status and predictor behavior

Predictive studies are seriously compromised if the subjects' current health status affects the predictive measure. So, for example, early reports of perceived heath predicting mortality in the Alameda County study might have been due to perceived health being affected by concurrent disease, but this is unlikely to be the explanation of prediction over decades (Schoenborn, 1993). This problem of health status affecting the predictor is rare in studies of healthy subjects, especially if care is taken to screen for disease. However, it is an obvious concern when subjects are selected for early signs of disease, as in the studies of vital exhaustion (Appels & Otten, 1992). The problem usually occurs when studying a short-term predictor of a comparatively rare event, such as a myocardial infarction; a convenient way of ensuring enough events in a short time is to select high-risk subjects but then their state of health may affect psychological predictors such as depression or exhaustion. The best solution is to test a substantial number of unselected subjects so that enough events occur in the critical time period.

8.04.2.4.4 Independence of predictor behavior and disease endpoint

In most predictive studies predictor and endpoint are, in measurement terms, independent, that is, the disease endpoint is determined without knowledge of the predictor. However, when self-report is used to define the predictor and the endpoint there is a real possibility of the relationship being inflated through common method variance. This is serious problem when concepts such as the disease-prone personality are used to predict conditions which depend on the patient's report of symptoms (e.g., angina pectoris), particularly as the disease-prone personality has an obvious similarity with Watson and Pennebaker's (1989) concept of negative affectivity which affects symptom reporting.

8.04.2.4.5 Costs and practicality of measures

Predictive studies often involve the testing and follow-up of many subjects. Such studies are expensive, are often multidisciplinary, and there is a great pressure to measure as many predictive factors as possible with minimal load on the respondents. As a result, the epidemiological literature is replete with studies in which inadequate psychological measures fail to predict disease. A striking example was the decision, understandable but probably unfortunate, of many researchers to replace Friedman and Rosenman's (1964) expensive and time-consuming TAB structured interview with much simpler inexpensive questionnaires (D. W. Johnston et al., 1987). It has been known for some time that, although different questionnaire measures of TAB correlate satisfactorily (Johnston & Shaper, 1983), they do not measure the same behaviors as the structured interview (Matthews, Krantz, Dembroski, & MacDougall, 1982) and that aspects of TAB which relate to CHD may not be assessed well by questionnaire. Adequate psychological measurement is necessary in predictive studies since the cost of a longitudinal study that fails through poor measurement is immense.

8.04.3 MEASURES USED WITH ILL PEOPLE

Under this heading, measures are considered which have been used with individuals identified by their illness status although they may not be concurrently experiencing illness. In addition to their relevance for ill people, such measures share a sensitivity to the respondent's status, for example, by being limited in length.

The central issues here are to measure how ill the person is, what the effects of illness are and how individuals respond to and cope with illness. (The word "illness" is used to refer to the subjective experience associated with disease processes as we believe it has universal meaning. In American English, the word "sickness" would be synonymous, but in other parts of the world this term is confined to illness associated with nausea.)

8.04.3.1 Measuring Illness

It can be a source of some confusion and frustration to discover that there are no biomedical indices of illness. Although the state seems subjectively recognizable and can be meaningfully replicated in animal studies, there is as yet no measure which reflects it. Recent animal research suggests that the state of illness may be associated with raised brain cytokines (Dantzer, 1997), but even this research offers little immediate prospect of a useful biomedical measure. The concept of illness, the subjective experience, is contrasted with the concept of disease, the pathophysiological process. For some diseases, it is possible to obtain a biomedical measure of the disease process such as degree of atherosclerotic occlusion in cardiovascular disease, degree of invasion of cancer cells, or loss of CD4 cells in AIDS. For others such as rheumatoid arthritis, there are indices of current inflammatory activity. However, patient "illness" may not relate directly to the disease process and, for many conditions, the main clinical and research assessments depend on patient observation or self-report. So, instead of reliable biomedical indices, "illness" is characterized by the subjective experience and associated behaviors.

The problem of measuring illness and disease is illustrated by Bowling (1995) in her book *Measuring disease*. She includes measures of symptoms, pain, emotional states, disability, and quality of life. However, none of these measures disease as such and each is influenced by psychological and/or social processes in addition to disease processes. As noted, the experience and report of symptoms may be influenced by individual differences in negative affectivity, but will also be affected by current attention, mood, cognitive, and social factors. As the behavioral, emotional, and quality of life consequences of disease are likely to be influenced by different factors, they are each considered separately.

8.04.3.1.1 Symptoms and pain

Problems in assessing the subjective experience of symptoms have been discussed above for healthy populations. Thus symptom checklists (Dunnell & Cartwright, 1972) need to be used with caution. In ill populations, there are disease-specific measures of symptoms such as measures included within composite measures. Examples, are the Rotterdam Symptom Checklist for cancer (de Haes, Van Knippenberg, & Neijt, 1990) or the Arthritis Impact Measurement Scale (Meenan, Gertman, & Mason, 1980). Such measures tend to have good internal consistency and have been validated using factor analytic methods and concurrent and predictive validation against other clinical or functional measures.

Pain is an important symptom of many conditions and an important motivator in seeking health care. Like other symptoms it is a subjective experience and therefore difficult to assess and validate measurement. The best known and most widely used measure is the McGill Pain Questionnaire (MPQ) (Melsack, 1975) which allows separate evaluation of sensory, affective, and evaluative components of pain. It records location, change with time, and factors influencing the pain and provides three main measures of intensity, the number of words chosen, the pain rating index, and the present pain intensity rating. A short form of the MPQ has been introduced which correlates well with the full MPQ and demonstrates sensitivity to change (Melsack, 1987). In clinical settings where, due to pain, patients may not have the concentration required for more extensive measures, a variety of single-item numerical and verbal pain rating scales have been used (see Chapter 8.24, this volume).

8.04.3.1.2 Behavioral consequences of disease

The WHO analysis of the consequences of disease identified three consequences: impairment or loss of structure/function, disability or loss of performance of activities, and handicap or loss of social roles (WHO, 1980). It was proposed that disability resulted from impairment and that handicap resulted from both impairment and disability. Although there has been much debate about this analysis, other (e.g. Nagi, 1991) and more recent WHO (1997) approaches retain the main characterizations of the elements involved.

Disability is defined in terms of behavior (Johnston, 1996) and is measured in terms of behavioral performance or capability. Thus, general disability measures assess limitations in the performance of normal functions; for example, the Barthel Index (Mahoney & Barthel, 1965) includes walking, bladder control, dressing, and using stairs, whereas more specific measures of Activities of Daily Living (ADL) assess the performance of everyday tasks such as self-care, shopping, and household tasks (Katz, Ford, Moskowitz, Jackson, & Jaffe, 1963; Lincoln & Edmans, 1990).

There is debate about whether measures should assess what the individual does or what they can do. Clearly, limitation in capacity is the construct most closely related to the direct consequence of disease, whereas limitation in performance may indicate the additional psychological, social, and environmental influences. Similar issues arise in other areas of performance psychology, where techniques such as dual-task performance have evolved to assess capacity; using the argument that, although performance may be maintained despite lower capability, the extra effort involved is assessed by examining the effects on a second task. To date, measures of disability have not taken account of this problem, and dual-task measures may enhance the measures in this field.

The behavioral consequences of disease may not simply be loss of activities, but may also involve the performance of new activities. The observable behaviors associated with pain have been assessed and they include additional behavior as well as loss of activities. Keefe and Block (1982) and Richards, Nepomuceno, Riles, and Su (1982) have developed measures of pain behaviors (e.g., grimacing, guarding movements, rubbing) which are assessed by observers. These measures are not necessarily closely related to self-report pain measures and may additionally reflect the individual's response to pain experience.

8.04.3.1.3 Emotional consequences of disease

Illness in oneself or in close family or friends is recognized as a significant source of stress in most methods of assessing life events and difficulties, including the original Social Readjustment Rating Scale (Holmes & Rahe, 1967). The emotional consequences can be assessed using any of the available measures of mood such as the STAI (Spielberger, 1983), the profile of Mood States (McNair, Lorr, & Droppelman, 1992) the CES-D (Radloff, 1977), or the GHQ (Goldberg & William, 1988). However, these can be confounding where somatic symptoms may be interpreted as emotional symptoms, such as where loss of appetite or slowing are indices of depression. Zigmond and Snaith (1983) introduced the Hospital Anxiety and Depression Scale (HADS) which they suggested minimized such confounding by reducing the number of such items. However, the success of this scale compared with other scales has not been assessed and, because it contains items such as "I feel as if I am slowed down," it may still contain contaminated items.

This problem does not arise in assessing the emotional consequences for people close to the patient and there is ample evidence that the emotional consequences may be as great for them as for the patient, for example, following myocardial infarction (MI). Where they are also caregivers, measures of caregiver burden are a combination of the emotional and behavioral burdens of caregivers (Orbell, 1996). The stress is observed not only in self-report measures but also in evidence of impaired immune function (Kennedy, Kiecolt-Glaser, & Glaser, 1988) with resulting vulnerability to disease. Some measures such as the GAIS and PAIS (Derogatis, 1975, 1986) assess adjustment to illness. Adjustment is frequently addressed as a form of coping or as illness cognitions (see Sections 8.04.3.2 and 8.04.3.3).

8.04.3.1.4 Quality of life

Although the construct of quality of life (QL) is widely used in assessing the consequences of disease, there are quite divergent schools of thought regarding definition and measurement. On the one hand, it is argued that the measures should assess the quality objectively, while on the other it is proposed that QL can only be evaluated according to the individual's expectations and values. Measures such as the Quality Adjusted Life Year (QALY) (Kind & Rosser, 1988; Rosser & Kind, 1978) or the SF-36 (Jenkinson, Coulter, & Loright, 1993) assume a universal value system, and so by assessing the individual's limitations it is possible to give an overall value to QL. By contrast, the Patient Generated Index (PGI) (Ruta, Garratt, Leng, Russell, & MacDonald, 1994) and the SEIQOL (O'Boyle, McGee, & Joyce, 1994) assess the individual's value system as well as the current levels of achievement in valued areas. The values are combined with the levels of achievement to generate an index reflecting quality of life according to the individual's own value system. Whereas the PGI requires the individual to respond in the context of their illness, the SEIQOL is one of the few measures of QL in current use in health psychology which could be used equally readily with healthy and ill individuals. Measures in other fields, such as the Satisfaction with Life Scale (Diener et al., 1985), assess QL without reference to illness and can be used successfully with individuals who are ill.

Some measures have been developed specifically for some diseases, including the AIMS (Meenan et al., 1980) for arthritis and the Rotterdam Symptom Check List (de Haes, Van Knippenberg, & Neijt, 1990) for cancer. The authors argue that a more sensitive and acceptable measure of QL can be obtained in this way. This reflects the design of disease-related QL measures, depending on assessment of symptoms and limitations in function as a major factor in QL. Where health-related QL measures assume a value system, this is usually based on assessment of emotional state, symptoms, and disability and the measure offers a method of combining these values. So, for example, the QALY gives values for each combination of distress and disability, and this can be done for individuals with diverse diseases.

It is clear that QL comes quite close to the WHO concept of handicap or to a combination of disability and handicap. There has been some attempt to measure handicap (e.g., the London Handicap Scale; Harwood, Gompertz, & Ebrahiam, 1994) and some measures clearly combine disability and handicap (e.g., the Sickness Impact Profile (de Bruin, de Wilte, Stevens, & Diederiks 1992) or the Functional Limitations Profile (Patrick & Peach, 1989). It is unlikely that there will be agreement on measurement while there continues to be debate about the constructs to measure. Although the construct of QL would appear to have value as a global measure of health outcome, it may mask the underlying processes as proposed by the WHO model. The WHO model proposes causal links between impairment, disability, and handicap rather than a homogeneous single construct and, further, allows the possibility that factors which may moderate the relationship between impairment and disability may be different from those that moderate the relationship between disability and handicap. Thus, the choice of measure, given the current stage of development of theory and measurement, is likely to be determined by whether one is seeking a simple health outcome measure or to understand the processes determining outcomes.

8.04.3.2 Illness Cognitions

The dominant models of response to illness, including stress and coping (Lazarus & Folkman, 1984), social-cognitive (Bandura, 1997) and self-regulation models (Leventhal et al., 1997), agree that individuals develop their own mental representations of the condition and it is these representations that guide the response. Models vary in the cognitions specified and in the level of measurement achieved.

Some address illness simply as a further example of a stressor and consider the degree of threat, outcome expectancies, perceived controllability, etc. as they would for any stressor. Leventhal et al. (1997) have attempted to characterize illness cognitions *per se*, and provide the most extensive work on this topic. They have identified five representations commonly occurring in individual representations:

(i) identity—the label and symptoms associated with the illness;

(ii) time-line—the pattern of change over time, especially whether it is acute, chronic or fluctuating;

(iii) cause—attributions for the illness;

(iv) consequences—the actual and expected consequences;

(v) control/cure—the extent to which the illness can be controlled or cured.

Their work was based initially on open-ended questions and qualitative analyses from which they developed guidelines for interviews (Leventhal & Cameron, 1987). More recently, Weinman, Petrie, Moss-Morris, and Horne (1996) have developed a questionnaire assessment of these five components of representations, the Illness Perceptions Questionnaire, using psychometric techniques to establish basic properties such as internal reliability of the scales. They have demonstrated that the scales have predictive and construct validity, finding that these cognitions predicted subsequent behaviors in patients post-MI as would be expected within the self-regulation model (Petrie, Weinman, Sharpe, & Buckley, 1996).

Weinman et al.'s analysis assumes but does not test the five-component structure found by Leventhal et al. It also assumes a dimensional structure, with each of the five types of representation being distributed along a numerical scale. There are other examples of assessments of illness cognitions which start from different premises.

Turk, Rudy, and Salovey (1986) used factor analytic methods implying an underlying dimensional structure, and found support for four dimensions: seriousness, personal responsibility, controllability, and changability. Although these factors have not proved entirely reproducible in subsequent studies, the degree of overlap between their factors, as characterized by the content rather than the labels for the factors, and the components proposed by Leventhal et al. is impressive. Thus, their controllability construct matches Leventhal et al., control/cure.

Other measures address single constructs other are still compatible with the Leventhal et al. model. Thus, Felton, Revenson, and Hinrichsen's (1984) measure of "Acceptance of Illness" appears to examine the consequences of the illness, containing items such as "my illness makes me a burden on my family and friends" and "I have a hard time adjusting to the limitations of my illness." Many measures address the control construct, including Wallston's recent form C adaptation of the Health Locus of Control scales for people with chronic disease (Wallston, Stein, & Smith 1994), scales for specific groups such as those with diabetes (Bradley, Brewin, Gamsu, & Moses, 1984), and those with the prospect of recovery from disabling conditions (Partridge & Johnston, 1989).

There are alternatives to the dimensional structured approach. Instead of building a representation of the illness from components, it may be useful to examine the prototypes used in evaluating new conditions (Bishop & Converse, 1986). So, for example, Lalljee, Lamb, and Carnibella (1993) have found seven clusters of illness prototypes. This type of analysis may prove useful in understanding responses to ambiguous sets of symptoms where patterns of response may be different to those where symptoms are more clear-cut.

8.04.3.3 Coping with Illness

Coping is defined as what people do to try to minimize stress and is commonly seen in health psychology as problem-focused, that is, directed at reducing the threats and losses of the illness, or emotion-focused, namely directed at reducing the negative emotional consequences. Perhaps the best known measure of coping is the ways of Coping Questionnaire, developed by Folkman and Lazarus (1980). They used factor-analytic procedures to establish the dimensional structure, but later analyses have produced different numbers and types of factors.

Many different coping styles have been identified and measured and contrasts are made between avoidant and attention coping, between active and passive coping, and between behavioral and cognitive coping. The COPE (Carver, Scheier, & Weintraub, 1989) attempts to represent all of these aspects of coping in a multidimensional inventory giving scores on 13 conceptually distinct types of coping. Further, the measure can be used to assess coping style, a dispositional measure of habitual ways of coping, or to assess coping strategy, the methods of coping used to address a specific stressor such as a current illness. The disadvantage of the comprehensiveness of the measure is that it is quite long, appears repetitive, and can seem burdensome to people who are ill. Carver (1997) has therefore published a shortened version which has proved more acceptable while retaining the multidimensional features of the measure.

Other measures address specific forms of coping, for example, Miller's (1987) measure of monitoring and blunting, or they deal with specific clinical conditions, such as Watson et al.'s (1988) assessment of coping styles employed by people with cancer.

Investigators choosing a measure of coping need to consider the compatibility of the measures with their theoretical perspective, the range of scores available, the acceptability of the measure to their clinical group and, possibly, the need to measure not only coping with the illness, but also coping with stressful forms of health-care.

8.04.4 MEASURES USED IN HEALTH-CARE CONTEXTS

People receive diverse forms of health-care, ranging from health checks, through consultations about minor symptoms, investigations of ambiguous symptoms, receiving diagnoses with threatening implications, enduring complex, painful, uncomfortable, or inconvenient forms of treatment, to being maintained on life-support systems. Investigations in health psychology examine factors resulting in individuals seeking health-care or accepting care that is offered, they explore the nature of the process of delivery of health-care, and they evaluate the psychological and behavioral outcomes of health-care.

For most forms of health-care, the recipients will be ill and therefore the questions and measures addressed in the previous section will also be pertinent. For those undergoing health checks or screening procedures, issues relevant to healthy individuals may have more relevance.

A major source of behavioral variation in health-care is due to the health professionals or providers as well as to the patients (Marteau & Johnston, 1991). As health services move towards implementation of evidence-based care, the assessment, explanation, and modification of their behavior is likely to be an increasing focus of investigations (Grol, 1992, 1997).

8.04.4.1 Measuring the Process of Health Care

The process of health-care delivery includes policy and management decisions and behaviors which occur at a level that tends to be outside the normal range of health psychology. The focus is primarily at the point of delivery, in the interaction between the patient and the health-care provider.

8.04.4.1.1 Communications between patients and providers

There is ample evidence that the nature and style of communications can determine the outcomes of health care (Ley, 1988; see Chapter 8.03, this volume). Various methods have been developed to describe and characterize the interpersonal communication process, using raters to assess records (audio, video, or typed transcripts) of the communication. Roter and Hall's Interaction Analysis (1989) based on the original work by Bales has been used to assess the components of interpersonal communication, such as giving information, asking questions, etc., whereas Stiles' Verbal Response Mode analysis (Putnam, Stiles, Jacob, & James, 1985; Stiles, 1989) identifies eight mutually exclusive modes of responding, including disclosure, interpretation, confirmation, and reflection. Other rating scale measures have been developed to assess the performance of health professionals against verbal and nonverbal performance criteria used in training (Maguire & Selby, 1989) or to assess the degree to which a doctor is patient- or doctor-centered (Byrne & Long, 1976). In each case, the methods are assessed in terms of inter-rater reliability and the predictive value of the assessments.

Methods of assessing written communications have also been developed (Ley, 1988). A major issue has been the assessment of intelligibility of written information such as leaflets for patients and some methods, such as the Flesch formula, can readily be applied to text using widely available word-processing software.

8.04.4.1.2 Uptake and adherence

When patients do not take up an offer of health care or do not adhere to recommended advice or treatment, the process of health care is interrupted. Although the assessment of uptake *per se* is fairly easy from health care records, it may be more difficult to ascertain whether a failure of uptake is due to not receiving the invitation, not understanding the invitation, not being able to accept the health care offered, or refusing the health care offered. These different patterns of interrupted health care have different causes and may have different effects on health outcomes; for example, Marteau et al. (1992) found different patterns for women refusing tests in pregnancy compared with those who simply omitted to have the test.

Similar contrasts have been drawn between intentional and nonintentional nonadherence to medication regimens and Horne (1997b) has developed a measure that allows separation of the two types of nonadherence. Although self-reports of adherence are widely criticized as being subject to social desirability biasing, they are correlated with other measures of adherence. In addition, all other measures of medication adherence have been criticized. Behavioral observation is limited as no observer can be permanently present, while pill counts and dispensers which record occasions of use can readily be manipulated by the patient who would wish to appear more compliant.

For adherence to advice on behavior change, some observational corroboration may be possible, but with the above caveats. For smoking advice, there are physiological indices such as serum cotinine which are less likely to be manipulated, whereas for dietary advice, measures of weight or serum cholesterol may be relevant. However, none of these measures has a one-to-one relationship with the behavior and they can therefore be misleading.

In sum, there appears to be no easy accepted solution to the problem of measuring adherence. It may be sufficient to use one, probably a self-report measure, for most investigations, but if this is a key variable, then more than one method should be used if possible.

8.04.4.1.3 Behavior of health professionals

Two main areas of behavior of health professionals have been investigated: their adherence to recommendations and the amount of stress they experience. As for patient adherence, it can be difficult to measure what cannot be observed and self-report may prove useful. However, some behaviors, such as hand hygiene, are observable and have been investigated. Other behaviors, such as drug dispensing, are likely to be assessable from patient or ward records.

Stress in health professionals has been investigated because of the recognition that working with people who are ill may be stressful and a concern that stressed professionals may be less able to deliver effective health care (Payne & Firth-Cozens, 1987). Perceived stress and the emotional impact can be investigated using conventional measures such as the Perceived Stress Scale (Cohen, Kamarck, & Mermelstein, 1983) or measures of mood or emotional state. In addition measures of occupational stress have been used, for example, the nursing stress index (Cooper & Mitchell, 1990), and the Maslach Burnout Inventory (Maslach, 1982) has been widely used to evaluate disabling levels of stress.

The effects of stress on performance of tasks has been assessed using self-reports of errors or by using cognitive tasks such as reaction time tasks as a proxy for the real task. The prevalence of major errors, such as in drug prescribing, can be investigated from health care or medicolegal records (Vincent, Ennis, & Audley, 1993).

8.04.4.1.4 Psychological and behavioral interventions

The recording of psychological interventions is not normally regarded as a measurement issue. However, a wide range of psychological and behavioral interventions is used in health care and it becomes increasingly important to be able to at least classify interventions. In the field of preparation for stressful medical procedures there are many interventions, for example, procedural information, sensory information, cognitive coping, and relaxation (Johnston & Vogele, 1993). For methods to be replicable and discriminable, they need to be characterized in sufficient detail to be able to at least classify which type of intervention is being used and perhaps to be able to rate interventions on the extent to which they use each mechanism (e.g., giving information, providing behavioral instruction, and training in relaxation).

This issue becomes increasingly urgent as high-profile reports of investigations of such interventions are published without any data, descriptive information, or methods of accessing such information about the interventions being provided (Johnston, 1995).

8.04.4.2 Health Care Cognitions

The beliefs, attitudes, and expectations of patients and health care providers to various aspects of health care have been investigated, often with a focus on the potential mismatch between patient and provider of health care or on beliefs likely to interfere with effective health care delivery.

8.04.4.2.1 Patients' health care cognitions

Some of patients' health care cognitions are those already considered with reference to health and illness; for example, perceived effectiveness of health care may be assessed in social cognition models of health behaviors, and perceived cure and controllability is a key element in illness cognitions.

Patients' expectations of health care have been investigated in terms of their general preference for managing their own treatment and the information relevant to their condition (Krantz, Baum, & Wideman, 1980) as well as preferences for care in a specific consultation with a doctor (Salmon & Quine, 1989). In each case, the scale has been validated using factor-analytic procedures and in both, items referring to provision of information have loaded on different factors from items dealing directly with treatment issues. Thus, there is some evidence that distinct elements of patients' preferences for medical care are being investigated.

In addition, there are more general measures of attitudes toward physicians and towards medicine (Horne, 1997a; Hulka, Zyzanski, Cassel, & Thompson, 1970; Marteau, 1990). Measures that evaluate a specific consultation are considered as health outcomes below.

Health care cognitions also include assessments of the stressfulness of receiving health care. For example, the Hospital Stress Rating Scale (Volicer & Bohannon, 1975) provides scaled values of the perceived stress associated with a large number of hospital experiences.

8.04.4.2.2 Providers' health care cognitions

As for patients' cognitions, health care providers' cognitions have also been investigated using the constructs of social cognition models. For example, their self-efficacy for providing dietary instructions was examined in the context of reducing serum cholesterol (Schucker et al., 1987). Investigators have also investigated their views of their working environment in the context of investigations of the stressfulness of caring for patients (Parkes, 1982). Further, measures of clinicians' attitudes to different patients groups have been developed; for example, St Lawrence, Kelly, Owen, Hogan, and Wilson's (1990) measure of attitudes to AIDS gives an index of favorableness of attitudes and one of willingness to interact with them in a clinical situation.

8.04.4.3 Health Care Outcomes

The main health care outcomes are health outcomes and these are measured as measures of illness as described in Section 8.04.3.1. In addition, health care outcomes include measures of health care utilization such as length of stay, number of primary care consultations, or use of medication; these measures can normally be obtained from health care records. For some investigations, these indices will be converted into health care costs, using the techniques of health economics. Delay in receiving health care, such as following MI, or deprivation of health care, especially preventive health care, may also constitute relevant outcomes. Where health care is directed at enhancing adherence with a prescribed regimen of medication or changing health-related behaviors such as diet, adherence to advice may also be an outcome, although it is more commonly seen as part of the process of health care. Where the objective is to impart information, measures of knowledge may constitute an outcome, although one might expect this improved knowledge to impinge on health or illness outcomes or on utilization of health care.

An important health care outcome is patient satisfaction. The most commonly used measure of satisfaction with a medical consultation is the Medical Interview Satisfaction Scale (MISS) (Wolf, Putman, James, & Stiles, 1978) which assesses affective, behavioral, and cognitive aspects of the consultation. Most measures of patient satisfaction have problems of validity and scaling as responses are skewed towards favorable ratings. Ware and Hays (1988) have found that measures of satisfaction are less skewed and therefore have improved psychometric qualities if respondents are asked to assess the excellence of the medical consultation rather than their satisfaction with the consultation.

8.04.5 MEASURES USED IN ALL CONTEXTS

Some measures are used in all contexts in health psychology.

8.04.5.1 Stress, Life Events, and Emotional States

Measures of stressors such as Life Events Scales (Holmes & Rahe, 1967), Impact of Events Scales (Horowitz, Wilner, & Alvarez, 1979), perceived stress scales (Cohen et al., 1983), or emotional states (Radloff, 1977; Zigmond & Snaith, 1983) are relevant in all contexts of health psychology. Life events, stress, or emotional states may influence health or the practice of health-related behaviors, may affect the likelihood of disease, may result from illness, and may be a consequence of health care for patients, their families, and health care providers alike. Thus measures that tap the stress process may be relevant for investigations in all areas of health psychology.

8.04.5.1.1 Social deprivation

Social deprivation has been shown to be related to the performance of health behaviors, to susceptibility to disease, to recovery from disease, and to receipt of health care (Adler et al., 1994; Carroll, Bennett, & Davey-Smith 1993; Carroll & Davey-Smith, 1997). It has been conceptualized in various ways: educational level, occupational status, financial status, socioeconomic status, residence in a deprived area, etc. Given the pervasiveness of this variable, it is important to ensure that it is measured as well as possible in all investigations.

8.04.5.2 Resources

Similarly, the resources that the individual brings to the health, illness, or health-care context are likely to determine how individuals respond. Coping has already been considered in the context of coping with illness, but clearly models of stress postulate that coping responses are likely to moderate the effects of stress and therefore it is appropriate to measure coping in contexts where stress may be relevant.

8.04.5.2.1 Social support

Social support has been conceptualized as directly beneficial, with lack of social support being akin to a stressor or as a resource that buffers the effects of stress (Cohen & Wills, 1985). Social support has been found to predict susceptibility to disease, to moderate the effects of stress, and to influence health and health care outcomes.

Social support has been measured in diverse ways and a major distinction is made between measures of the structure or network available to the individual and the quality or satisfaction with the support available; for example, the Social Support Questionnaire (Sarason et al., 1983; Sarason, Sarason, Shearin, & Pierce, 1987) gives measures of amount and satisfaction with social support. Similarly, there are different types of support with the most common distinction being made between emotional and practical support (Power, Champion, & Aris, 1988), but other forms such as esteem, informational, and companionship have also been distinguished (Cohen & Wills, 1985).

Such measures are difficult to validate, but correlations with relevant other self-report measures supports the construct validity of the scales.

8.04.5.2.2 Individual differences

Individual differences have already been referred to in the context of predicting disease (Section 8.04.2.4) and the confounding effects of negative affectivity in measurement especially in measuring illness (Section 8.04.3.1). Negative affectivity has been measured as trait anxiety (Spielberger, 1983), neuroticism (Eysenck & Eysenck, 1985), or using the scale developed by Watson, Clark, and Tellegen (1988), the Positive and Negative Affect Schedule (PANAS).

Positive affectivity has been found to be independent of negative affectivity and may act as a resource in stressful situations. Similarly, other attributes that may serve this protective function are Dispositional Optimism (Scheier & Carver, 1985, 1987) and self-esteem (Rosenberg, 1989).

8.04.6 FUTURE DIRECTIONS

This chapter has described the main measurement issues in health psychology. As this is a comparatively new area of research and practice, there are problems as well as solutions in many of these areas. However, there is increasing sophistication in measurement in most of the areas covered. In part this comes about simply as a result of increasing experience of the measures, so that gradually issues of reliability, validity, ambiguity, consumer resistance, etc. are identified and tackled. It is to be expected that this process will continue and increasing experience with measures and ever widening areas of application will lead to improved measurement. Measurement in health psychology will also benefit from the changes in technology that will affect most areas of psychology. Increasingly, testing will be computerized and as a result more relevant information will be gathered without overburdening the respondent. Advances in technology will also lead to more information, both behavioral and physiological, being gathered in the client's own environment, rather than the laboratory or clinic.

However, it is expected that improvements will increasingly come about because of more fundamental changes in how questions are asked and in the tools used to answer them. As indicated, many measurement problems arise because of a lack of specificity in how theoretical concepts should be operationalized. Although still far from perfect, theories are becoming more specific and directly address issues of operationalization (see Chapter 8.01, this volume). If this trend continues then there should be an increasingly good fit between theory and measure, and also increased comparability across different domains when the same theory is applied to different problems. A tool that is starting to increase our understanding of what we measure (particularly in questionnaire design) and how to improve it is the widening use of complex multivariate methods such as structural equation modeling and the related procedures of confirmatory factor analysis. A central problem in health psychology is the very substantial overlap between concepts; in particular, the real possibility that much of what is measured reflects primarily the pervasive effects of negative affectivity. The application of appropriate statistical techniques should see this issue resolved and is likely to lead to a reduced set of health-related individual difference measures and a better focused series of specific measures.

ACKNOWLEDGMENT

Some parts of this chapter have been influenced and informed by the work of John Weinman and Stephen Wright in co-authoring *Measures in health psychology: A user's portfolio* (Johnston, Wright, & Weinman, 1995).

8.04.7 REFERENCES

Abelin, T., Brzezinski, Z. J., & Carstairs, V. D. L. (Eds.) (1987). *Measurement in health promotion and protection.* European Series No. 22. Copenhagen: WHO Regional Office for Europe.

Abrahams, C. S., & Hampson, S. E. (1996). Special issue on: Controversy and method in the interpretation of verbal reports in health psychology research. *Psychology and Health, 11*(2).

Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Khan, R. L., & Syme, S. L. (1994). Socioeconomic status and health: The challenge of the gradient. *American Psychologist, 49,* 15–24.

Amir, D. (1987). Preventive behaviour and health status among the elderly. *Psychology and Health, 1,* 353–78.

Anderson, E. (1987). Preoperative preparation for cardiac surgery facilitates recovery, reduces psychological distress, and reduces the incidence of acute postoperative hypertension. *Journal of Consulting and Clinical Psychology, 55,* 513–520.

Anderson, F. A., Wheeler, H. B., Goldberg, R. J., Hosmer, D. W., Forcier, A. M., & Patwardhan, A. (1991). Physician practices in the prevention of venous thromboembolism. *Annals of Internal Medicine, 115,* 591–595.

Appels, A., & Otten, F. (1992). Exhaustion as precursor of cardiac death. *British Journal of Clinical Psychology, 31,* 351–356.

Bandura, A. (1977). Self-efficacy: Towards a unifying theory of behaviour change. *Psychological Review, 84,* 191–215.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: Freeman.

Bishop, G. D., & Converse, S. A. (1986). Illness representations: A prototype approach. *Health Psychology, 5,* 95–114.

Bowling, A. (1991). *Measuring health.* Buckingham, UK: Open University Press.

Bowling, A. (1995). *Measuring disease.* Buckingham, UK: Open University Press.

Bradburn, N. M. (1969). *The structure of psychological well-being.* Chicago: Aldine.

Bradley, C., Brewin, C. R., Gamsu, D. S., & Moses, J. L. (1984). Development of scales to measure perceived control of diabetes mellitus and diabetes-related health beliefs. *Diabetic Medicine, 1,* 213–218.

Byrne, P. S., & Long, B. E. L. (1976). *Doctors talking to patients.* London: HMSO.

Cacioppo, J. T., & Tassinary, L. G. (Eds.) (1990). *Principles of psychophysiology: Physical, social, and inferential elements.* Cambridge: Cambridge University Press.

Carroll, D., Bennett, P., & Davey-Smith, G. (1993) Socioeconomic health inequalities: their origins and implications. *Psychology and Health, 8,* 295–316.

Carroll, D., & Davey-Smith, G. (Eds.) (1997). Special Issue: Health Variations. *Journal of Health Psychology, 2*(3).

Carver, C. S. (1997). You want to measure coping but your protocol's too long: Consider the brief COPE. *International Journal of Behavioral Medicine, 4,* 92–100.

Carver, C. S., Scheier, M. F., & Weintraub, J. K. (1989). Assessing coping strategies: A theoretically-based approach. *Journal of Personality and Social Psychology, 56*(2), 267–83.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behaviour, 24,* 385–396.

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin, 98,* 310–357.

Conner, M., & Norman, P. (Eds.) (1996). *Predicting health behaviour.* Buckingham, UK: Open University Press.

Cooper, C. L., & Mitchell, S. (1990). Nurses under stress—a reliability and validity study of the NSI. *Stress Medicine, 16,* 21–24.

Cox, B. D., Blaxter, M., Buckle, A. I. J., Fenner, N. P., Golding, J. F., Gore, M., Huppert, F. A., Roth, M., Stark, J., Wadsworth, M., & Whichelow, M. (1987). *The health and lifestyle survey*. London: Health Promotion Research Trust.

Dantzer, R. (1997). Stress and immunity: What have we learned from psychoneuroimmunology? *Acta Physiologica Scandinavia, 161*(Suppl. 640), 43–46.

De Bruin, A. F., De Witte, L. P., Stevens, F., & Diederiks, J. P. M. (1992). Sickness impact profile: The state of the art of a generic functional status measure. *Social Science and Medicine, 35,* 1003–1014.

De Haes, J. C. J. M., Van Knippenberg, F. C. E., & Neijt, J. P. (1990). Measuring physical and psychological distress in cancer patients: Structure and application of the Rotterdam Symptom Checklist. *British Journal of Cancer, 62,* 1034–1038.

Derogatis, L. R. (1975). *The Global Adjustment to Illness Scale (GAIS)*. Baltimore: Clinical Psychometric Research.

Derogatis, L. R. (1986). The Psychosocial Adjustment to Illness Scale (GAIS). *Journal of Psychosomatic Research, 30,* 77–91.

Diener, E., Emmons, R. A., Larson, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49,* 71–75.

Doll, R., Peto, R., Wheatley, K., Gray, R., & Sutherland, I. (1994). Mortality in relation to smoking: 40 years, observations on male British doctors. *British Medical Journal, 309,* 901–911.

Dunnell, K., & Cartwright, A. (1972). *Medicine takers, prescribers and hoarders*. London: Routledge and Kegan Paul.

Eysenck, H. J., & Eysenck, S. B. G. (1985). *Manual for the Eysenck personality questionnaire*. London: Hodder and Stoughton.

Family Heart Study Group (1994). Randomised controlled trial evaluating cardiovascular screening and intervention in general practice: Principal results of the British Family Heart Study. *British Medical Journal, 308,* 313–320.

Felton, B. J., Revenson, T. A., & Hinrichsen, G. A. (1984). Stress and coping in the explanation of psychological adjustment among chronically ill adults. *Social Science and Medicine, 18,* 889–898.

Fishbein, M., & Ajzen, I. (1975). *Belief attitude, intention and behaviour: An introduction to theory and research.* Reading, MA: Addison-Wesley.

Fisher, K., & Johnston, M. (1996a). Emotional distress as a mediator of the relationship between pain and disability: An experimental study. *British Journal of Health Psychology, 1,* 207–218.

Fisher, K., & Johnston, M. (1996b). Experimental manipulation of perceived control and its effect on disability. *Psychology and Health, 11,* 657–669.

Folkman, S. (1997). Introduction to the special section: Use of bereavement narratives to predict well-being in gay men whose partners died of AIDS M four theoretical perspectives. *Journal of Personality and Social Psychology, 72,* 851–854.

Folkman, S., & Lazarus, R. S. (1980). An analysis of coping in a middle-aged community sample. *Journal of Health and Social Behaviour, 21,* 219–239.

Freedman, R. R. (1989). Quantitative measurements of finger blood flow during behavioral treatments for Raynaud's disease. *Psychophysiology, 26,* 437–441.

Friedman, H. S., & Booth-Kewley, S. (1987). The "disease-prone personality": A meta-analytic view of the construct. *American Psychologist, 42,* 539–555.

Friedman, M., & Rosenman, R. H. (1964). Procedures for assessment of behavior pattern. *Journal of the American Medical Association, 189,* 15–22.

Friedman, M., Thoresen, C. E., Gill, J. J., Ulmer, D., Powell, L. H., Price, V. A., Brown, B., Thompson, L., Rabin, D. D., Breall, W. S., Bourg, E., Levy, R., & Dixon, T. (1986). Alteration of type A behavior and its effects on cardiac recurrences in post myocardial infarction patients: Summary results of the recurrent coronary prevention project. *American Heart Journal, 112,* 653–665.

Goldberg, D., & Williams, P. (1988). *A user's guide to the General Health Questionnaire*. Windsor, UK: NFER-Nelson.

Grol, R. (1992). Implementing guidelines in general practice care. *Quality in Health Care, 1,* 184–191.

Grol, R. (1997). Beliefs and evidence in changing clinical practice. *British Medical Journal, 315,* 418–421.

Haaga, D. A. F. (Ed.) (1997). Introduction to the special section on measuring cognitive products in research and practice. *Journal of Consulting and Clinical Psychology, 65,* 907–1000.

Harris, D. M., & Guten, S. (1979). Health-protective behavior: An exploratory study. *Journal of Health and Social Behavior, 20,* 17–29.

Harwood, R. H., Gompertz, P., & Ebrahim, S. (1994). Handicap one year after a stroke: Validity of a new scale. *Journal of Neurology, Neurosurgery and Psychiatry, 57,* 825–829.

Hecker, M. H. L., Chesney, M. A., Black, G. W., & Frautschi, N. (1988). Coronary-prone behaviors in the Western Collaborative Group study. *Psychosomatic Medicine, 50,* 153–164.

Holmes, T. H., & Rahe, R. H. (1967). The Social Readjustment Rating Scale. *Journal of Psychosomatic Research, 11,* 213–218.

Horowitz, M., Wilner, N., & Alvarez, W. (1979). Impact of events scale: A measure of subjective stress. *Psychological Medicine, 41,* 209–218.

Horne, R. (1997a). Representations of medication and treatment: Advances in theory and measurement. In K. J. Petrie & J. A. Weinman (Eds.), *Perceptions of health and illness*. Amsterdam: Harwood.

Horne, R. (1997b). *The nature, determinants and effects of medication beliefs in chronic illness*. Unpublished Ph.D. thesis. University of London.

Hulka, B. S., Zyzanski, S. J., Cassel, J. L., & Thompson, S. J. (1970). Scale for the measurement of attitudes towards physicians and primary health care. *Medical Care, 8,* 429.

Idler, E. L. (1992). Self-assessed health and mortality: A review of studies. In S. Maes, H. Leventhal, & M. Johnston (Eds.), *International review of health psychology* (Vol. 1). Chichester, UK: Wiley.

Ironson, G., Taylor, C. B., Boltwood, M., Bartzokis, T., Dennis, C., Chesney, M., Spitzer, S., & Segall, G. M. (1992). Effects of anger on left ventricular ejection fraction in coronary artery disease. *American Journal of Cardiology, 70,* 281–285.

Jenkinson, C., Coulter, A., & Wright, L. (1993). The short form — 36 (SF-36) health survey questionnaire: Normative data for adults of working age. *British Medical Journal, 306,* 1437–1440.

Johnston, D. W., Cook, D. G., & Shaper, A. G. (1987). Type A behaviour and ischaemic heart disease in middle aged British men. *British Medical Journal, 295,* 86–89.

Johnston, D. W., & Shaper, A. G. (1983). Type A behaviour in British men: Reliability and intercorrelation of two measures. *Journal of Chronic Diseases, 36,* 203–207.

Johnston, M. (1982). Recognition of patients, worries by nurses and by other patients. *British Journal of Clinical Psychology, 21,* 255–261.

Johnston, M. (1995). Health related behaviour change. In I. Sharpe (Ed.), *Cardiovascular prevention in primary care:*

The way forward, National Forum for Coronary Disease Prevention. London: Kings Fund.

Johnston, M. (1996). Models of disability. *The Psychologist, 9,* 205–210.

Johnston, M., Bromley, I., Boothroyd-Brooks, M., Dobbs, W., Ilson, A., & Ridout, K. (1987). Behavioural assessments of physically disabled patients: Agreement between rehabilitation therapists and nurses. *International Journal of Rehabilitation Research, 10*(Suppl. 5), 204–213.

Johnston, M., Earll, L., Mitchell, E., Morrison, V., & Wright, S. (1996). Communicating the diagnosis of motor neurone disease. *Journal of Palliative Medicine, 10,* 23–34.

Johnston, M., & Vogele, C. (1993). What benefits can psychological preparation for surgery achieve? *Annals of Behavioral Medicine, 15,* 245–256.

Johnston, M., Wright, S., & Weinman, J. (1995). *Measures in health psychology: A User's portfolio*. Windsor, UK: NFER-Nelson

Kaplan, R. M. (1990). Behaviour as the central outcome in health care. *American Psychologist, 45,* 1211–1220.

Kasl, S. N., & Cobb, S. (1966). Health behavior, illness behavior, and sick role behavior. I: Health and illness behavior. *Archives of Environmental Health, 12,* 246–266.

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., & Jaffe, M. W. (1963). Studies of illness in the aged: The index of ADL—a standardised measure of biological and psychosocial function. *Journal of the American Medical Association, 185,* 914–919.

Keefe, F. J., & Block, A. R. (1982). Development of an observational method for assessing pain behaviour in chronic low-back pain patients. *Behaviour Therapy, 13,* 363–375.

Kennedy, S., Kiecolt-Glaser, J. K., & Glaser, R. (1988). Immunological consequences of acute and chronic stressors: Mediating role of interpersonal relationships. *British Journal of Medical Psychology, 61,* 77–86.

Kind, P., & Rosser, R. (1988). The quantification of health. *European Journal of Social Psychology, 18,* 63–77.

Krantz, D., Baum, A., & Wideman, M. (1980). Assessment of preferences for self-treatment and information in health care. *Journal of Personality and Social Psychology, 39,* 977–990.

Lalljee, M., Lamb, R., & Carnibella, G. (1993). Lay prototypes of illness—their content and use. *Psychology and Health, 8,* 33–49.

Lau, R. R., Hartman, K. A., & Ware, J. E. (1986). Health as a value: Methodological and theoretical considerations. *Health Psychology, 5*(1), 25–43.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.

Leventhal, H., Benyamini, Y., Brownlee, S., Diefenbach, M., Leventhal, E. A., Patrick-Miller, L., & Robitaille, C. (1997). Illness representations: Theoretical foundations. In K. J. Petrie & J. A. Weinman (Eds.), *Perceptions of health & illness* (pp. 19–46). Amsterdam: Harwood.

Leventhal, H., & Cameron, L. (1987). Behavioural theories and the problem of compliance. *Patient Education and Counselling, 10,* 117–138.

Ley, P. (1988). *Communicating with patients*. London: Croom Helme.

Lincoln, N. B., & Edmans, J. A. (1990). A re-validation of the Rivermead ADL Scale for elderly patients with stroke. *Age and Ageing, 19,* 19–24.

Maguire, P., & Selby, P. (1989). Assessing quality of life in cancer patients. *British Journal of Cancer, 60,* 437.

Mahoney, F. I., & Barthel, D. W. (1965). Functional evaluation: The Barthel Index. *Maryland State Medical Journal, 14,* 61–65.

Marteau, T. M. (1990). Attitudes towards doctors and medicine: the preliminary development of a new scale. *Psychology and Health, 4,* 351–356.

Marteau, T., & Johnston. M. (1991). Health professionals: A source of variance in health outcomes. *Psychology and Health, 5,* 47–58.

Marteau, T. M., Johnston, M., Kidd, J., Michie, S., Cook, R., Slack, J., & Shaw, R. W. (1992). Psychological models in predicting uptake of prenatal screening. *Psychology and Health, 6,* 13–22.

Maslach, C. (1982). *Burn-out: The cost of caring*. Englewood Cliffs, NJ: Prentice-Hall.

Matthews, K., Krantz, D. S., Dembroski, T. M., & MacDougall, J. M. (1982). Unique and common variance in structured interview and Jenkins activity measures of the Type A behavior pattern. *Journal of Personality and Social Psychology, 42,* 303–313.

McNair, D. M., Lorr, M., & Droppelman, L. F. (1992). EdITS manual for the Profile of Mood States (POMS). San Diego: EdITS/Editorial and Industrial Testing Service.

Meenan, R. F., Gertman, P. M., & Mason, J. R. (1980). Measuring health status in arthritis: The Arthritis Impact Measurement Scales. *Arthritis and Rheumatism, 23,* 146–152.

Melsack, R. (1975). The McGill Pain Questionnaire: Major properties and scoring methods. *Pain, 1,* 277–299.

Melsack, R. (1987). The short-form McGill Pain Questionnaire. *Pain, 30,* 191–197.

Michie, S., McDonald, V., & Marteau, T. (1996). Understanding responses to predictive genetic testing: A grounded theory approach. *Psychology and Health, 11,* 455–470.

Miller, S. M. (1987). Monitoring and blunting: Validation of a questionnaire to assess styles of information seeking under threat. *Journal of Personality and Social Psychology, 52,* 345–353.

Nagi, S. (1991). Disability concepts revisited: Implications for prevention. In A. Pope & A. Tarlow (Eds.), *Disability in America: Toward a national agenda for prevention.* Washington, DC: National Academy Press.

Nolen-Hoeksema, S., McBride, A., & Larson, J. (1997). Rumination and psychological distress among bereaved partners. *Journal of Personality and Social Psychology, 72,* 855–862.

O'Boyle, C., McGee, H. M., & Joyce, C. R. B. (1994). Quality of life: Assessing the individual. *Advances in Medical Sociology, 5,* 159–180.

Orbell, S. (1996). Informal care in social context: A social psychological analysis of participation, impact and intervention in care of the elderly. *Psychology and Health, 11,* 155–178.

Oxcheck Study Group (1994). Effectiveness of health checks conducted by nurses in primary care: results of the OXCHECK study after 1 year. *British Medical Journal, 308,* 308–312.

Parkes, K. R. (1982). Occupational stress among student nurses: A natural experiment. *Journal of Applied Psychology, 67,* 784–796.

Partridge, C., & Johnston, M. (1989). Perceived control of recovery from physical disability: measurement and prediction. *British Journal of Clinical Psychology, 28,* 53–59.

Patrick, D. L., & Peach, H. (Eds.) (1989). *Disablement in the community*. Oxford, UK: Oxford University Press.

Payne, R., & Firth-Cozens, J. (1987). *Stress in health professionals*. Chichester, UK: Wiley.

Pennebaker, J. W., Mayne, T. J., & Francis, M. E. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology, 72,* 863–871.

Peterson, C., Buchanan, G. M., & Seligman, M. E. P. (1995). Explanatory style: History and evolution of the field. In G. M. Buchanan & M. E. P. Seligman (Eds.), *Explanatory style*. Hillsdale, NJ: Erlbaum.

Petrie, K. J., Weinman, J. A., Sharpe, N., & Buckley, J. (1996). Role of patients' view of their illness in predicting return to work and functioning after myocardial infarction. *British Medical Journal, 312,* 1191–1194.

Power, M. J., Champion, L. A., & Aris, S. J. (1988). The development of a measure of social support: The significant others scale (SOS). *British Journal of Clinical Psychology, 27,* 349–358.

Prochaska, T. R., Leventhal, E. A., Leventhal, H., & Keller, M. L. (1985). Health practices and illness cognition in young, middle aged, and elderly adults. *Journal of Gerontology, 40,* 569–578.

Putnam, S. M., Stiles, W. B., Jacob, M. C., & James, S. A. (1985). Patient exposition and physician explanation in initial medical interviews and outcomes of clinic visits. *Medical Care, 23,* 74–83.

Radloff, L. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1,* 385–401.

Ragland, D. R., & Brand, R. J. (1988). Coronary heart disease mortality in the Western Collaborative Group study: Follow-up experience of 22 years. *American Journal of Epidemiology, 127,* 462–475.

Richards, J. S., Nepomuceno, C., Riles, M., & Suer, Z. (1982). Assessing pain behaviour: The UAB pain behaviour scale. *Pain, 14,* 393–398.

Rosenberg, M. (1989). *Society and the adolescent self-image* (reprint ed.). Middletown, CT: Wesleyan University Press.

Rosenman, R. H., Brand, R. J., Jenkins, C. D., Friedman, M., Straus, R., & Wurm, M. (1975). Coronary heart disease in the Western Collaborative Group study: Final follow-up experience of 8½ years. *Journal of American Medical Association, 233,* 872–877.

Rosser, R., & Kind, P. (1978). A scale of valuations of states of illness: Is there a social consensus? *International Journal of Epidemiology, 7,* 347–358.

Roter, D. L., & Hall, J. A. (1989). Studies of doctor–patient interaction. *Annual Review of Public Health, 10,* 163–200.

Rotter, J. B. (1966). Generalised expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*(1), 609.

Ruta, D. A., Garratt, A. M., Leng, M., Russell, I. T., & MacDonald, L. M. (1994). A new approach to the measurement of quality of life — The patient-generated index. *Medical Care, 32,* 1109–1126.

Salmon, P., & Quine, J. (1989). "Patients" intentions in primary care: Measurement and preliminary investigation. *Psychology and Health, 3,* 103–110.

Sarason, I. G., Levine, H. M., Basham, R. B., & Sarason, B. R. (1983). Assessing social support: The social Support Questionnaire. *Journal of Personality and Social Psychology, 44,* 127–139.

Sarason, I. G., Sarason, B. R., Shearin, E. N., & Pierce, G. R. (1987). A brief measure of social support: Practical and theoretical implications. *Journal of Social and Personal Relationships, 4,* 497–510.

Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology, 4,* 219–47.

Scheier, M. F., & Carver, C. S. (1987). Dispositional optimism and physical well-being: The influence of generalized outcome expectancies on health. *Journal of Personality, 55,* 169–210.

Schoenborn, C. A. (1993). The Alameda study—25 years later. In S. Maes, H. Leventhal, & M. Johnston (Eds.), *International Review of Health Psychology* (Vol. 2). Chichester, UK: Wiley.

Schucker, B., Wittes, J., Cutler, J. A., Bailey, K., Mackintosh, D. R., Gordon, D. J., Haines, C. M., Mattson, M. E., Goor, R. S., & Rifkind, B. M. (1987). Change in physician perspective on cholesterol and heart disease. *Journal of the American Medical Association, 258,* 3521–3526.

Shaper, A. G., Pocock, S. J., Phillips, A. N., & Walker, M. (1987). A scoring system to identify men at high risk of heart attack. *Health Trends, 19,* 37–39.

Shaper, A. G., Pocock, S. J., Walker, M., Cohen, N. M., Wale, C. J., & Thomson, A. G. (1982). British regional heart study: Cardiovascular risk factors in middle-aged men in 24 towns. *British Medical Journal, 283,* 179–186.

Siegman, A. W., Dembroski, T. M., & Ringel, N. (1987). Components of hostility and the severity of coronary artery disease. *Psychosomatic Medicine, 49,* 127–135.

Smith, M. S., Wallston, K. A., & Smith, C. S. (1995). The development and validation of the perceived health competence scale. *Health Education Research, 10,* 51–64.

Spielberger, C. D. (1983). *Manual for the State–Trait Personality Inventory*. Palo Alto, CA: Consulting Psychologists Press.

St Lawrence, J. S., Kelly, J. A., Owen, A. D., Hogan, I. G., & Wilson, R. A. (1990). Psychologists' attitudes towards AIDS. *Psychology and Health, 4,* 357–365.

Stein, N., Folkman, S., Trabasso, T., & Richards, T. A. (1997). Appraisal and goal processes as predictors of psychological well-being in bereaved caregivers. *Journal of Personality and Social Psychology, 72,* 872–884.

Steptoe, A., & Johnston, D. (1991). Clinical applications of cardiovascular assessment. *Psychological Assessment: Journal of Consulting and Clinical Psychology, 3,* 337–349.

Stiles, W. B. (1989). Evaluating medical interview process components: Null correlations with outcomes may be misleading. *Medical Care, 27,* 212–220.

Turk, D. C., Rudy, T. E., & Salovey, P. (1986). Implicit models of illness. *Journal of Behavioral Medicine, 9,* 453–474.

Vincent, C., Ennis, M., & Audley, R. (1993). *Medical accidents*. Oxford, UK: Oxford University Press.

Volicer, B. J., & Bohannon, M. W. (1975). A hospital stress rating scale. *Nursing Research, 24,* 352–359.

Wallston, K. A. (1992). Hocus-pocus, the focus isn't strictly locus: Rotter's social learning theory modified for health. *Cognitive Therapy and Research, 16,* 183–199.

Wallston, K. A., Stein, M. J., & Smith, C. A. (1994). Form C of MHLC Scales, a condition-specific measure of locus of control. *Journal of Personality and Social Assessment, 63,* 534–553.

Wallston, K. A., Wallston, B. S., & DeVellis, R. (1978). Development of the multidimensional health locus of control (MHLC) scales. *Health Education Monographs, 6,* 161–170.

Ware, J. E., & Hays, R. D. (1988). Methods for measuring patient satisfaction with specific medical encounters. *Medical Care, 26,* 393–402.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affects: the PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063–1070.

Watson, D., & Pennebaker, J. W. (1989). Health complaints, stress and distress: Exploring the central role of negative affectivity. *Psychological Review, 96,* 234–254.

Watson, M., Greer, S., Young, J., Inayat, G., Burgess, C., & Robertson, B. (1988). Development of a questionnaire measurement of adjustment: The MAC scale. *Psychological Medicine, 18,* 203–209.

Weinberger, D. A., Schwartz, G. E., & Davidson, R. J. (1979). Low-anxious, high-anxious, and repressive coping styles: Psychometric patterns and behavioral and physiological responses to stress. *Journal of Abnormal Psychology, 88,* 369–380.

Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer.

Weinman, J. A., Petrie, K. J., Moss-Morris, C., & Horne, R. (1996). The illness perception questionnaire: A new method for assessing the cognitive representation of illness. *Psychology and Health, 11,* 431–445.

Wolf, M. H., Putman, S. M., James, S. A., & Stiles, W. B. (1978). The medical interview satisfaction scale: Development of a scale to measure patient perceptions of physician behavior. *Journal of Behavioral Medicine, 1,* 391.

World Health Organization (1980). *International Classification of impairments, disabilities, and handicaps*. Geneva: Author.

World Health Organization (1997). *Towards a common language for disablements: ICIDH*. Geneva: Author.

Wright, S. J. (1990). Health status measurement: review and prospects. In P. Bennett, J. Weinman, & P. Spurgeon (Eds.), *Current developments in health psychology*. London: Harwood Academic.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression Scale. *Acta Psychiatrica Scandinavica, 67,* 361–370.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.05

Design Issues for Clinical Research in Health Psychology

CAROLYN E. SCHWARTZ
Frontier Science & Technology Research Foundation Inc., Chestnut Hill and Harvard Medical School, Boston, MA, USA

8.05.1 INTRODUCTION 137

8.05.2 IMPLEMENTING RESEARCH IN A CLINICAL SETTING 138

- *8.05.2.1 Key Considerations for Meaningful Research* 138
  - *8.05.2.1.1 Ask a meaningful research question* 138
  - *8.05.2.1.2 Have a strong theoretical foundation* 138
  - *8.05.2.1.3 Measure appropriate outcomes* 138
  - *8.05.2.1.4 Use an appropriate study design* 139
  - *8.05.2.1.5 Maintain a high response rate* 139
  - *8.05.2.1.6 Maintain complete data* 139
  - *8.05.2.1.7 Plan and implement appropriate statistical analysis* 140
  - *8.05.2.1.8 Plan study with adequate statistical power* 140
  - *8.05.2.1.9 Disseminate findings to appropriate audiences* 140
- *8.05.2.2 Study Designs to Evaluate Clinical Interventions* 140
  - *8.05.2.2.1 Needs assessment followed by randomized consent* 141
  - *8.05.2.2.2 Three-arm crossover design* 145
  - *8.05.2.2.3 Case–control* 146
  - *8.05.2.2.4 Solomon four-group pre–post design* 146
  - *8.05.2.2.5 Pre–post one-group design* 146
- *8.05.2.3 Selecting the Appropriate Study Design for Evaluating Treatment Intervention Effects: Balancing Research Questions and Assumptions* 146
- *8.05.2.4 Practical Challenges to Implementing Research in a Clinical Setting* 147
  - *8.05.2.4.1 Who pays for treatments being evaluated?* 147
  - *8.05.2.4.2 Incentives of capitated care* 148
  - *8.05.2.4.3 Legal considerations* 148
  - *8.05.2.4.4 Unavoidable costs of doing research* 149
  - *8.05.2.4.5 Ethical considerations* 149
- *8.05.2.5 An Emerging Theoretical Challenge for Outcomes Research* 149

8.05.3 SUMMARY 150

8.05.4 REFERENCES 150

8.05.1 INTRODUCTION

With the increasing focus on cost-containment and cost-effectiveness, clinical health psychologists are finding it increasingly necessary to document the value of their services. This documentation requires relying not only on empirical data to demonstrate the impact of treatment, but also on research designs which are rigorous yet feasible in active clinical settings where funding may not be available to support resource-intensive efforts. Given the emphasis of health psychology on the individual's well-being, it would seem incumbent on such research to link the patient's perspective with objective clinical outcomes as well as cost. Such linkage would also be consistent with the biopsychosocial model which drives much of health psychology theory and practice. The purpose of this chapter will be to discuss key designs, considerations, and challenges of doing outcomes studies of interventions in a clinical health psychology and/or medical setting. Although similar issues will be relevant for health psychology research which is correlative/predictive in nature, it is beyond the scope of the present chapter. Further information on some of the various concepts and designs discussed can be found in Volume 3.

8.05.2 IMPLEMENTING RESEARCH IN A CLINICAL SETTING

8.05.2.1 Key Considerations for Meaningful Research

Active clinical settings present an inviting paradox to the health psychologist. On the one hand, they have the ultimate resource for relevant outcomes research: an ample source of patients. Further, there is an obvious potential to link patient outcome data to clinical and cost data. The challenge of such settings is implementing data collection efforts from patients which are both externally and internally valid, often without financial and personnel resources dedicated to research endeavors. Thus, key considerations in study implementation need to be planned and integrated into the daily clinical routine so that the resulting data yield meaningful and useful glimpses into the patient's experience of illness and its management. The considerations which are crucial to successful study implementation (Table 1) will be discussed briefly.

8.05.2.1.1 Ask a meaningful research question

Since any well-done research endeavor is likely to require substantial effort and organization, it would be important that the research be motivated by research questions which are meaningful to patients, providers, payers, and other parties. A clinical setting provides the health psychology researcher with ample opportunities to hear about potentially interesting research issues and to develop hypotheses about what social and behavioral factors would be worthy of exploration in a formal study. By listening to patients, their caregivers, healthcare providers, payers, and others, researchers can develop questions which are responsive to current forces in healthcare delivery and health enhancement in general. Meaningful research questions can be developed in response to maintaining an open ear when interacting with patients and others in the clinical setting. Other more formal approaches might include focus groups (Morgan, 1988) and semistructured individual interviews with people with various relevant perspectives.

8.05.2.1.2 Have a strong theoretical foundation

A reasonable background of both knowledge and theory would be necessary conditions for a research study to be worth the effort and resources it will cost. Such a background can be gleaned by consulting with health psychologists who have a current and broad foundational understanding of models of change or growth from fields such as health education, public health, clinical and health psychology, and psychiatry. These models would provide a necessary foundation for hypothesizing what psychosocial and medical factors would be relevant and thus should be measured (Slife & Williams, 1997). This understanding might be supplemented with recent literature reviews as well as searches of computer databases of medical and psychological publications.

8.05.2.1.3 Measure appropriate outcomes

In beginning to plan a study with a strong theoretical foundation, one must select relevant outcomes. This selection process should also be informed by clinical acumen and patient experience. Focus groups (Morgan, 1988) with providers and patients can be helpful in identifying germane outcomes. This method usually identifies relevant groups whose perspective should be considered and begins with a series of open-ended questions. The underlying theme of focus groups is to expand the realm of possibilities rather than to contract them. One continues focus groups and various provider and patient subgroups until the information gleaned becomes redundant (Morgan, 1988). Focus groups can also be useful for pretesting questionnaire packets and for selecting one questionnaire among many which ostensibly measure the same thing. For example, there are numerous measures of functional status which might be used in an outcomes study. In the pilot-testing phase of a study evaluating the impact of two psychosocial interventions (Schwartz, 1994; Schwartz & Rogers, 1994; Schwartz & Sprangers, under review), Schwartz pretested a set of functional status measures, and asked chronically ill patients to track their mood and energy level before and after completing each measure.

Table 1 Key considerations for successful study implementation.

Consideration	*Method for achieving*
a. Meaningful research question	Listening to patients and clinicians, literature review, attenting to regulatory forces in health care delivery
b. Theory-based	Literature review; expert consultation
c. Appropriate outcomes	Focus groups
d. Appropriate study design	Match with research question and assumptions (see Table 2)
e. Response rate	Total Design Method
f. Complete data	Telephone follow-up
g. Appropriate analysis	Involvement of statistician when planning study
h. Statistical power	Involvement of statistician when planning study
i. Dissemination to appropriate audiences	Peer-review journals; patient education; managed care journals

The measure which was selected finally was the one which was the least depressing and was rated by patients as asking questions they deemed pertinent to their physical well-being.

8.05.2.1.4 Use an appropriate study design

Although collecting data from consecutive patients who come into a clinic might be considered research by some people, most professionals would agree that studies must be designed before they are implemented if they are to result in data worth analyzing. It is believed that study design is a crucial stumbling block for many aspiring researchers, and an entire section of this chapter has been devoted to describing five study design options which can be implemented feasibly in busy clinical settings (see below).

8.05.2.1.5 Maintain a high response rate

Ascertaining the generalizability of the study results will be an important consideration in assessing the impact that it may have. Studies which lose a large proportion of eligible participants are necessarily suspect of suffering from selection bias which may hamper the investigators' ability to make any valid conclusions about the treatments being evaluated. Further, substantial drop-out rates can result in more time required for accrual and more costly resources required to implement studies. Thus, planning carefully to maximize response rate would be a necessary condition to successful research.

In his Total Design Method, Dillman (1978) provides an accessible and feasible method which integrates the large literature on survey research to specific steps for assuring a high response rate. These steps can be easily implemented by the clinical researcher. For example, Dillman provides details regarding the letter which is sent to patients to inform them and invite them to participate in a given research study. He outlines the information which should be included in each paragraph, the letterhead parameters which have been documented to enhance response rate, the color ink which should be used to sign the letter, whose names should grace the letter, and the way in which the letter should be folded. Dillman (1978) provides a similar level of detail regarding preparation of the questionnaire booklets. Although this level of detail might seem unnecessary, he documents the increased percentage of response rate which results from their consideration. If one follows Dillman's protocol, one can be assured of a 90% response rate. Indeed, following these steps very carefully has allowed our group to maintain a high response rate in a number of studies by mail.

8.05.2.1.6 Maintain complete data

Missing data can pose significant problems to data analysis and interpretation. Most self-report questionnaires require data on at least 75% of the items within a subscale to allow imputation of a subscale score. Similarly, most statistical methods require complete data to include a participant's data in the analysis. Although some statistical methods can accommodate missing data (e.g., growth curve; Francis, Fletcher, Stuebing, Davidson, & Thompson, 1991), the meaning of the missingness can be an important consideration. For example, if data are "informed" missing, then their absence is indicative of some notable cause, such as dying during the course of the study. If they drop out due to health problems or access to care, then their missing data would also be informative. Thus, it is incumbent on the health psychology investigator to maintain complete data as much as possible, and to document reasons for drop-outs and missing data.

Developing a good rapport with study participants can be crucial for attaining complete data. This rapport may be developed initially in an intake interview, and maintained via telephone follow-up. For example, if a participant has not returned a questionnaire packet within two weeks of mailing it out, the investigator would be well-advised to inquire via telephone and to negotiate a date by which the participant is able to return the completed packet. Similarly, if the packet is returned but there are items which have not been answered, a respectful phone call can be beneficial. In general, it has been found that such calls are perceived by patients to be indicative of the seriousness and meaningfulness of the research endeavor, and they are cooperative in answering the missing items. If the patient purposefully left the item blank, then the investigator should respect the confidentiality of the patient, code the item as "refused," and not press the participant for an answer.

8.05.2.1.7 Plan and implement appropriate statistical analysis

A well-planned investigation would include planning the statistical analysis before the study is implemented. Such intention would ensure that the appropriate data (e.g., covariates, predictors, and outcomes) are collected, and that the research questions will be addressed adequately by the statistical analysis. It is recommended to include on the research team a professional who has expertise in statistical analysis. This person's expertise should be included at the planning stage of any study.

8.05.2.1.8 Plan study with adequate statistical power

By planning the statistical analysis when the study's research questions are being defined, the thoughtful investigator paves the way for adequate statistical power. Simply stated, an investigator's power to detect meaningful differences depends on three parameters: the sample size, the magnitude of the effect which is deemed clinically significant, and the inherent variability of the outcome measure. The power curve is exponential, requiring few subjects per group for large effects, and large numbers of subjects per group for moderate and small effects (Figure 1) (Cohen, 1988; Pocock, 1983). It is recommended that power calculation be used during the planning phases of a study by a professional who has expertise in statistical analysis, or with the aid of software designed for behavioral science studies (e.g., Bornstein & Cohen, 1988).

8.05.2.1.9 Disseminate findings to appropriate audiences

One of the most important products of a research study is developing vehicles for disseminating study results. Such dissemination should take place through peer-reviewed professional journals, a mechanism which can enhance the state of knowledge in the broader arena. Other mechanisms exist which can influence other highly relevant audiences. For example, there are private foundations which focus on a specific disease and which would provide conduits by which investigators might implement patient educational programs based on the results of their research. There might also be access to other healthcare providers and payers by publishing jargon-free articles in managed care journals and health management organizational newsletters. The more mechanisms by which access to study results can be provided, the more influential such results can be. Even negative (i.e., nonsignificant) findings can be informative, so investigators are to be encouraged to disseminate study results in all cases.

8.05.2.2 Study Designs to Evaluate Clinical Interventions

Thus far, important considerations in successful study implementation to enhance the external and internal validity of the research endeavor have been reviewed. Although the clinical setting does not necessarily pose challenges to assessing needs or concerns of patient populations, many important research questions may be more focused on how efficacious and cost-effective an intervention may be. Despite the above suggestions, the clinical setting may pose challenges to implementing the standard randomized controlled trial. In this design, eligible patients are approached and asked to participate in a trial in which they will be randomized to one of two treatment arms. Study designs which maximize statistical power are placebo-controlled randomized trials with longitudinal follow-up (Kraemer & Thiemann, 1989). Given the ethical, theoretical, and practical problems with placebo-controlled studies (Schwartz, Chesney, Irvine, & Keefe, 1997), the modern investigator often may be forced to consider statistical power issues in the context of intervention studies of two active treatments, such as standard care as compared to an experimental treatment or psychological adjunct (Figure 2(a)). The more one controls for nonspecific effects, the smaller the anticipated effect size and hence the larger the sample size requirements (Kazdin & Bass,

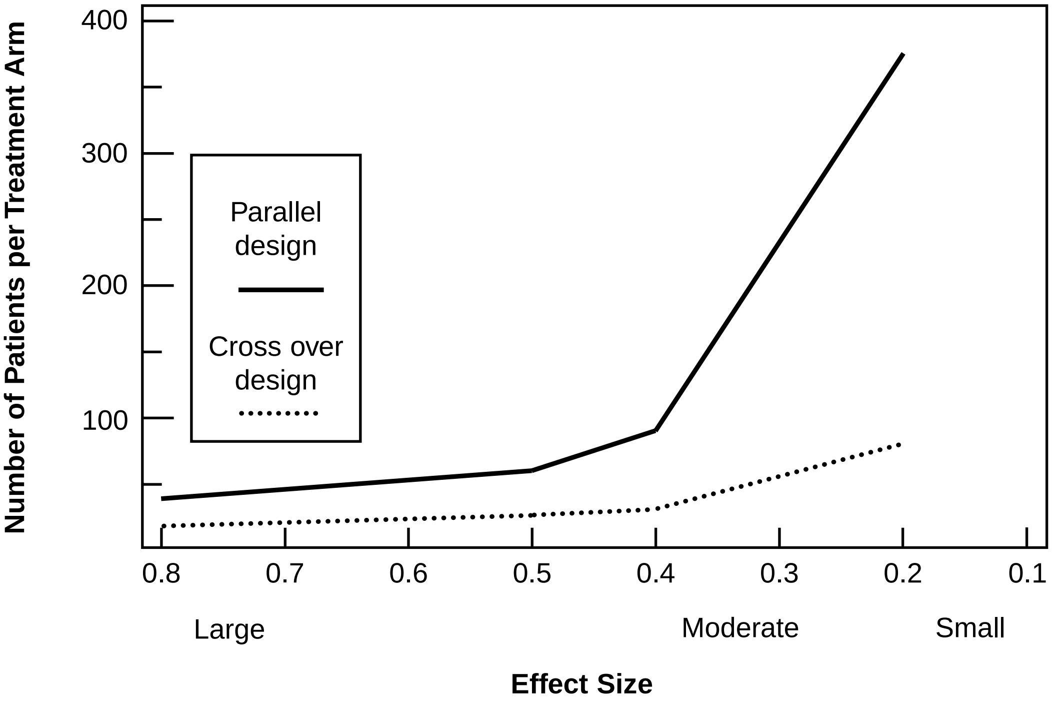

Figure 1 Statistical power curves for parallel design and crossover trial. This plot shows how the sample size requirements for a study increase with the decreasing magnitude of the anticipated effect size. Using a crossover design can reduce the required sample to approximately one-quarter the number of subjects as a parallel-arm study for the same power (Fleiss, 1986). This estimation assumes that the within-subject variability is equal to the between-subject variability.

1989). Similarly, the more the intervention is tailored to the individual patient, the more heterogeneous the resulting comparison groups. This heterogeneity also reduces the statistical power due to the increased variability in the outcomes. The increased cost and challenge of implementing such a large study might lead to considering the several design alternatives outlined below.

8.05.2.2.1 Needs assessment followed by randomized consent

An alternative to the standard randomized controlled trial is the randomized consent design (Zelen, 1979, 1990). In this design, eligible patients are randomized prior to obtaining consent, and consent is obtained for participation in that arm to which the patient has been randomized (Figure 2(b)). This design might be beneficial in a clinical setting where the efficacy or cost-effectiveness of an adjunctive treatment which is a reimbursable clinical service is evaluated. In this case, "reimbursable" refers to the idea that a third party (i.e., insurance company) would deem the service worthy of coverage by the individual's healthcare insurance. Such a concern is particularly relevant in countries which do not have universal healthcare coverage (e.g., the USA). By using a randomized design, the statistical comparability of the two treatment arms is ensured, thereby reducing the systematic error encountered in more quasi-experimental designs. The primary caveat of the randomized consent design is that it is feasible only if participants will be highly likely to accept the treatment to which they have been randomized. This caveat is due to the intention-to-treat principle of randomized trials: that is, that patient data are analyzed according to the intervention to which they were assigned, regardless of how much of the treatment they actually received. This principle ensures that the more consistent the efficacy (i.e., effectiveness and feasibility) of an intervention, the more ecologically valid the representation of the real-world impact it is likely to have.

To ensure that the participants will be highly likely to accept the treatment to which they have been randomized, implementing a needs assessment survey of clinic patients is suggested. This survey can be used to help determine the psychosocial needs and meaningful interventions for a given target population (Edwards & White, 1987). Data from such surveys can also be used to evaluate the sociodemographic and medical correlates of expressed needs (Kraft,

a. Standard Randomized Controlled Trial

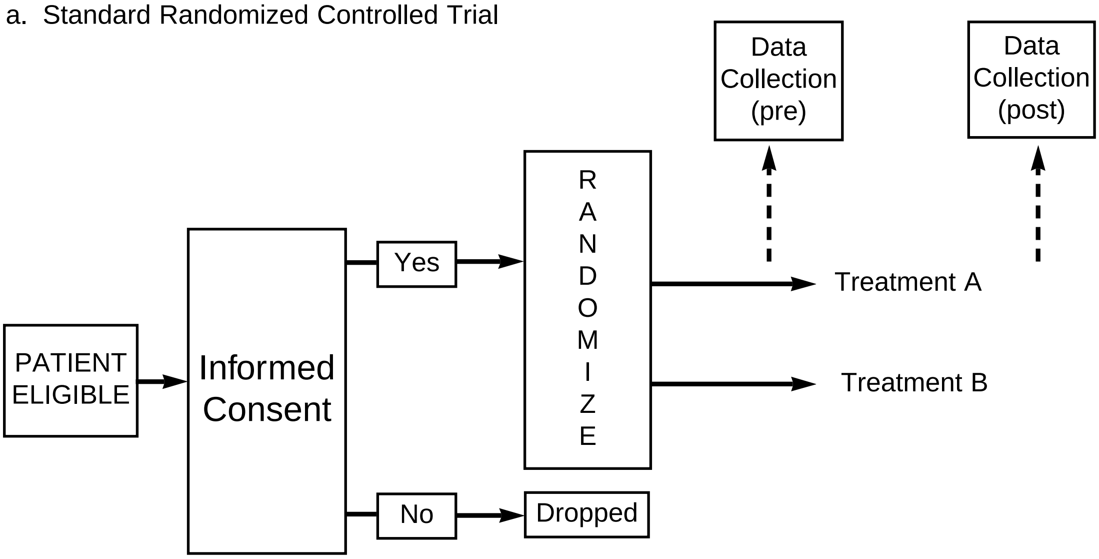
PATIENT ELIGIBLE
Informed Consent
Yes
No
RANDOMIZE
Dropped
Treatment A
Treatment B
Data Collection (pre)
Data Collection (post)

b. Randomized Consent Design

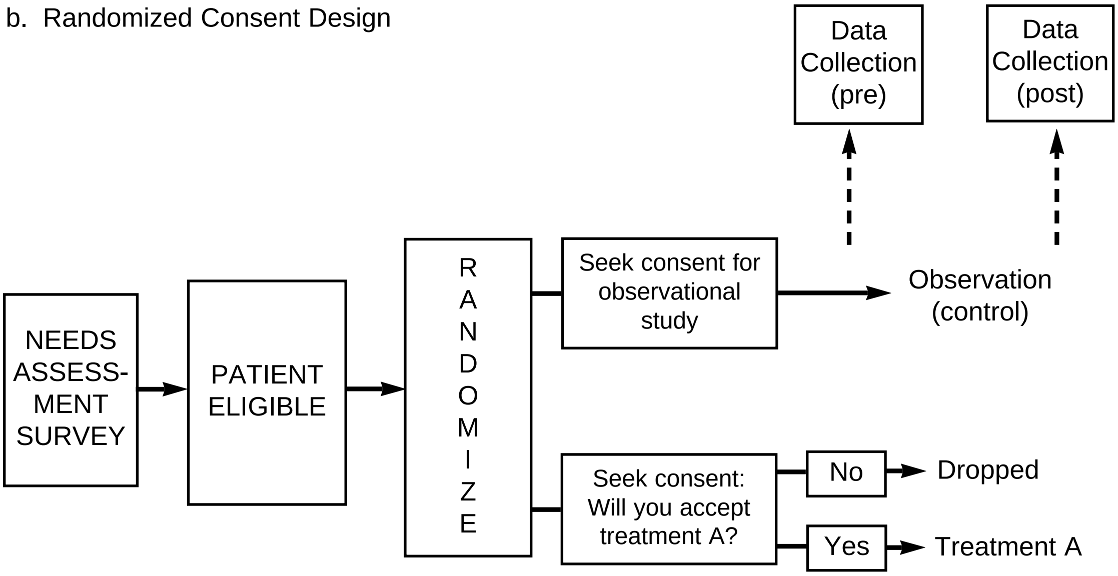
NEEDS ASSESS-MENT SURVEY
PATIENT ELIGIBLE
RANDOMIZE
Seek consent for observational study
Observation (control)
Seek consent: Will you accept treatment A?
No
Dropped
Yes
Treatment A
Data Collection (pre)
Data Collection (post)

c. Three-arm Crossover Design

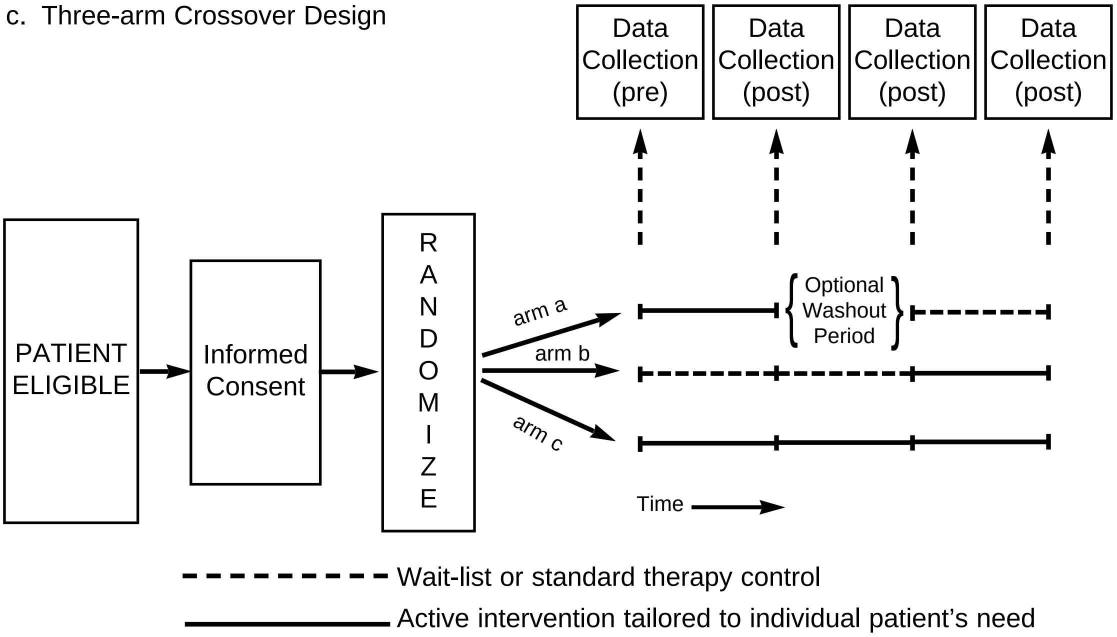
Data Collection (pre)
Data Collection (post)
Data Collection (post)
Data Collection (post)
PATIENT ELIGIBLE
Informed Consent
RANDOMIZE
arm a
arm b
arm c
Optional Washout Period
Time
Wait-list or standard therapy control
Active intervention tailored to individual patient's need

d. Case-control Design

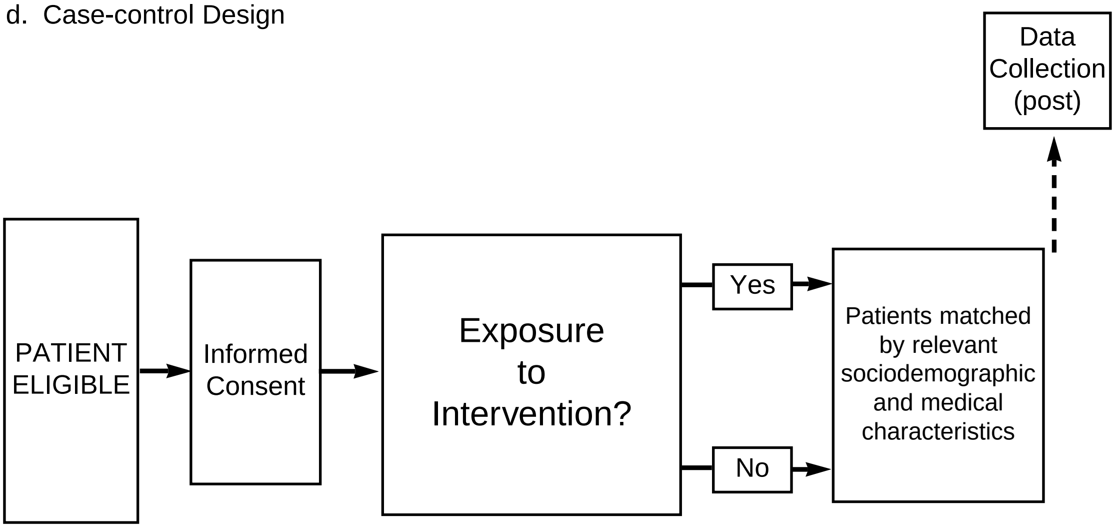
Data Collection (post)
PATIENT ELIGIBLE
Informed Consent
Exposure to Intervention?
Yes
No
Patients matched by relevant sociodemographic and medical characteristics

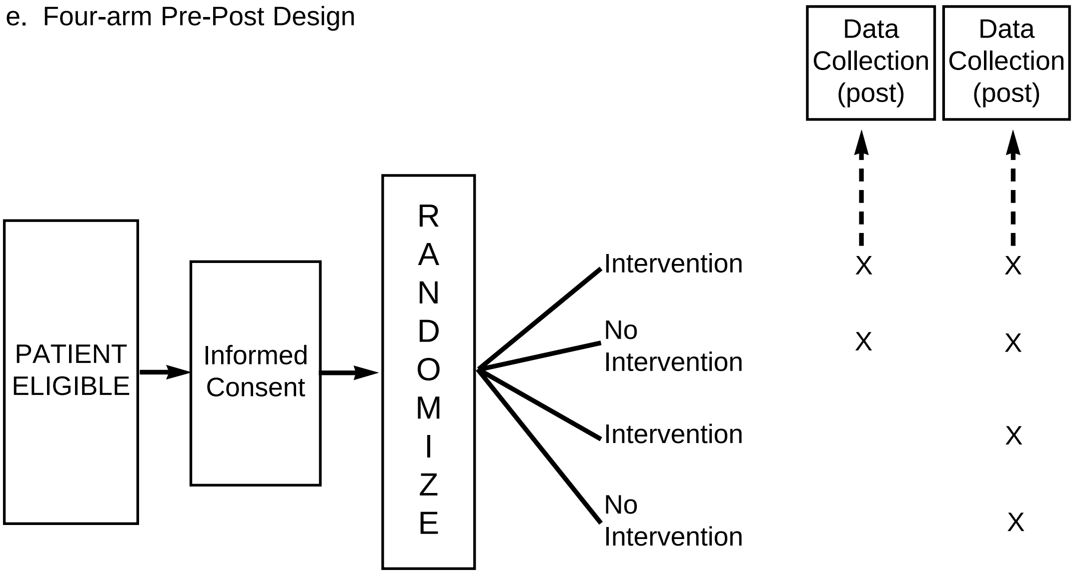
e. Four-arm Pre-Post Design
PATIENT ELIGIBLE
Informed Consent
R A N D O M I Z E
Intervention
No Intervention
Intervention
No Intervention
Data Collection (post)
Data Collection (post)
X
X
X
X
X
X

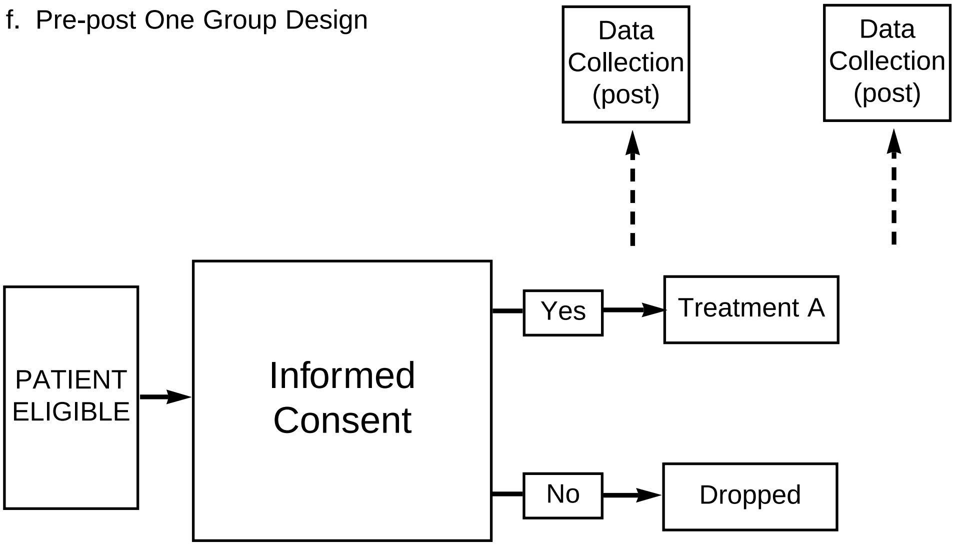
f. Pre-post One Group Design
PATIENT ELIGIBLE
Informed Consent
Yes
No
Treatment A
Dropped
Data Collection (post)
Data Collection (post)

Figure 2 Research designs which would be feasible to implement in an active clinical setting. (a) Schema for a standard randomized controlled trial for the sake of comparison with the subsequent designs. (b) Schema for a needs assessment survey followed by a randomized consent design. (c) Schema for the three-arm crossover design, which allows one to evaluate the impact of duration of treatment on outcomes. (d) Case–control design which would require matching patients on relevant sociodemographic and medical characteristics. (e) The Solomon four-arm pre–post design controls for maturation effects and investigates the impact of salience cuing on treatment outcomes. (f) Schema for a pre–post one-group design which is the simplest and least expensive study to implement. This study allows one to evaluate patient factors associated with better treatment outcomes.

Freal, & Coryell, 1986). By beginning with a needs assessment survey, a needy target population can be identified for a specific trial who would be likely to want the intervention(s) being evaluated (Figure 2(b)).

8.05.2.2.2 Three-arm crossover design

Given the difficulty in randomizing patients to a placebo–control condition, Schwartz et al. (1996) have proposed a randomized three-arm variation of a standard crossover trial. This design allows evaluation of treatment, order, and dose–response effects. In the first two arms, patients would receive active treatment followed by a control condition (arm "a" in Figure 2c), or vice versa (arm "b" in Figure 2(c)). The patients randomized to the third arm (arm "c" in Figure 2(c)) would receive the active treatment for the entire duration of the other treatment arms (i.e., treatment and control). This arm would allow the investigator to address the longer-term effect of the intervention, without dramatically increasing the sample size.

The control condition in the first two arms (i.e., arms "a" and "b" in Figure 2(c)) would be a wait-list which could involve standard care or no treatment. The decision about whether to use a standard care or no-treatment wait-list control would depend on the usual treatment approach for the study population and symptom complex. If there were a standard treatment for the symptom complex (e.g., pharmaceuticals, support groups, antidepressants, etc.), then the wait-list patients would have normal access to them. If the syndrome is one that does not interfere significantly with daily life and is not threatening to health then a no-treatment control would be appropriate. If the trial population is in crisis, then the control group would get standard care.

Since crossover trials compare the impact of an active treatment to standard care and/or wait list control, the designs avoid the need to create a placebo–control. Thus, the three-arm crossover similarly minimizes the ethical dilemma by providing treatment to all patients. The enhanced statistical efficiency of the crossover design would allow the use of individually-tailored interventions, and thus increase the clinical validity of the trial results. Such treatment protocol heterogeneity would normally reduce the statistical power of the trial and lead to an expensive increase in the number of patients required to detect a difference between treatments. However, crossover designs use patients as their own controls and are thereby statistically efficient: they require approximately one quarter the number of patients per study arm as those required for a standard parallel design (Figure 1, Fleiss, 1986). This estimation assumes that the within-subject is equal to the between-subject variability. This increased power would offset the reduced statistical power of the protocol heterogeneity.

The proposed three-arm crossover design allows for an examination of dose–response relationships, whereas the standard crossover trial limits the investigator's ability to learn about the effects of longer-term intervention. The three arms provide information about the effects of shorter-term interventions (i.e., arm "a" or "b") and about longer-term interventions (i.e., arm "c"). The investigator might seek to evaluate the maintenance of effects over time (i.e., carryover effects) by shorter-term interventions using data from arm "a" or design the study to include a washout period between treatment and follow-up so that the follow-up could be a clear no-treatment comparison condition. Consequently, treatment outcomes should be measured pre- and postintervention as well as before and after the washout and no-treatment periods. Thus, the proposed design responds to the standard problem of crossover designs, that of not allowing the investigator to examine whether continued adherence to an intervention produces benefits over a long period of time or follow-up. This use of the three-arm crossover trial data is similar to an approach used by Thoresen (1991) to evaluate the long-term impact of a behavioral intervention for postinfarction patients who had participated in a crossover trial. They were able to show that the larger the "dose" received of the behavioral intervention, the greater behavior change and health benefit.

Finally, the three-arm crossover design can allow the investigator to evaluate order effects, which may prove illuminating for hypothesis generation. For example, assume that condition "a" is nonspecific support (e.g., social support), condition "b" is coping effectiveness training (i.e., the experimental intervention), and condition "c" is two cycles of the coping effectiveness training. The design is: AB, BA, and BB. Suppose that analysis of the results revealed that AB was better than BA and BB. The interpretation, framed as hypothesis generation, might be that social support prepared patients for the coping skills training, whereas coping skills training given first inadvertently increased resistance in some patients. Therefore, this three-arm crossover design could identify possible timing or staging effects of the different therapies.

The appropriateness of the three-arm crossover trial will depend on the objectives of the research. If the intervention is not expected to have extended carryover effects (e.g., acupuncture, massage), then no washout period will be required. If the intervention is expected to have brief carryover effects (e.g., pharmacotherapy), then a brief washout period may be necessary. If the intervention is a cognitive-behavioral treatment for stress management, then a washout period may not be feasible because it may not be possible to remove the effects of the exposure. If the objective is to study carryover effects, then a standard parallel-arm design would be appropriate (i.e., treatment vs. control with no crossover). For example, the experimental treatment, best standard care, and a wait-list control or usual care (if the latter is not the same as the comparison condition) could be compared. The latter design will be expensive to implement due to the sample size requirements in parallel-arm trials comparing active treatments (i.e., small effects sizes, see Figure 1). Further, it may not be feasible among proactive patients who refuse to participate if randomized to a wait-list control. The three-arm crossover design is a feasible and appropriate design if carryover effects are not the primary endpoints. It represents one of several possible approaches to design.

8.05.2.2.3 Case–control

Borrowed from the field of epidemiology, the case–control design can be useful for health psychology research on interventions where some naturalistic assignment has occurred. In this design, patients are naturalistically "exposed" to one of the interventions under study and followed over time to determine the effect of the intervention (Figure 2(d)). This design might be used in cases where similar patient groups at difference sites receive distinct interventions. For example, if two sites of a managed care organization had similar patient populations in a diabetes clinic, and only one of the clinics implemented a patient education self-management program. Patients from both sites could be followed over time to evaluate the patient education program.

8.05.2.2.4 Solomon four-group pre–post design

This design allows the explicit consideration of external validity factors (Campbell, 1963). By paralleling a standard pretest–post-test control group design with experimental and control groups lacking the pretest (Figure 2(e)), the main effects of both the intervention and testing can be evaluated, as well as the interaction of the intervention and testing. This design consequently allows for an investigation of the impact of salience cuing on treatment outcomes. For example, measuring an outcome requires asking patients to self-report about that outcome. Thus the outcome's salience may be increased as a function of study participation, thereby leading to a shift in internal standards (i.e., response shift as described subsequently) on behalf of study participants. The Solomon four-group design would provide a vehicle for examining this aspect of response shift in clinical health psychology research.

8.05.2.2.5 Pre–post one-group design

Many investigators may not feel that they have the expertise to implement studies with more complex designs. Consequently, they rely on recruiting consecutive patients for treatment evaluation studies and asking patients to self-report on various quality-of-life outcomes before and after the intervention (Figure 2(f)). Because there is no comparison group, such data are not appropriate for testing the effectiveness of an intervention. However, this design can facilitate correlative or predictive studies in health psychology which seek to understand which factors are associated with better outcomes. One example of this use of the pre–post one-group design was a study done by Mohr et al. (1996) in which they examined how patient expectations were related to treatment compliance among a cohort of multiple sclerosis patients. They found that adverse events and unrealistic expectations were important predictors of treatment compliance in 86% of patients initiating an immunomodulating therapy. This use of the pre–post one-group design exemplifies how clinically relevant information can be gleaned with this simple and feasible research design.

8.05.2.3 Selecting the Appropriate Study Design for Evaluating Treatment Intervention Effects: Balancing Research Questions and Assumptions

The above five study designs are reasonable alternatives to randomized controlled trials for health psychology research. However, selecting the appropriate design will require balancing the research question being asked as well as the assumptions on which each design relies (Table 2). Although the first five study designs address the relative effectiveness of treatment A vs. treatment B, they differ in their assumptions about patient characteristics. For example, the standard randomized controlled trial assumes that patients will accept randomization to the control group, and will not seek alternative and similar treatments to the experimental group due to resentful demoralization (Cook & Campbell, 1978). In contrast, the randomized consent design only asks that patients accept the singular treatment option to which they have already been randomized. Thus, the former may take longer to accrue participants if the control group is obvious. On the other hand, the randomized consent design is only feasible if the researcher is fairly certain that patients will be likely to accept the treatment option to which they have been assigned.

A common assumption of randomized designs is that the random allocation method results in groups which are similar on sociodemographic and medical characteristics. As this is not always the case, stratification may be necessary for the crucial characteristics. Covariate adjustment can be done in the analysis phase for other pretreatment differences which occurred despite randomization. Although covariate adjustment can be done in both case–control and pre–post one-group designs, the former would benefit from matching study participants on relevant sociodemographic and medical factors to enhance the comparability of the two groups. This matching will prevent confounding due to differential representation of various prognoses by study arms, a design problem which cannot be resolved with covariate adjustment. One example of such a problem would be comparing the impact of a behavioral intervention on people with epilepsy, and not matching for location and focality of brain lesions. It is well-known that patients with focal temporal lesions are more likely to be considered as candidates for neurosurgery (Lechtenberg, 1990), and hence have a better psychosocial prognosis than those with diffuse lesions in other parts of the brain. Not matching on such important prognostic factors can lead to unnecessary confounding and perhaps even selection bias.

Given the similarities in research questions and assumptions across the several designs suggested, clinical investigators may need to base their research design selection on characteristics of the context in which they would like to initiate the research study. For example, randomized trials may not be feasible in settings in which there is no support staff to oversee the blind allocation of treatment or vigilant follow-up. In these contexts, the case–control or pre–post one-group designs may be optimal as they would allow one to dovetail data collection efforts to standard clinic visits and can address relevant, albeit limited, questions about treatment outcomes. The clinical investigator therefore has several design options which can facilitate outcomes research despite limited resources. Selecting an appropriate design will require the careful balance of the research question one seeks to ask, the assumptions one is comfortable making, and the constraints imposed by the clinical context.

8.05.2.4 Practical Challenges to Implementing Research in a Clinical Setting

Once an appropriate design has been selected, clinical investigators may wish to begin collecting data immediately. However, there are other practical challenges which they may confront. These challenges reflect the emerging financial and legal considerations which are playing an increasing role in healthcare. Highlighting a few of these considerations should make it easier for potential investigators to prepare for them.

8.05.2.4.1 Who pays for treatments being evaluated?

The appeal of implementing research in a clinical setting is partially founded on the assumption that the cost of such research can be reduced by having it dovetail with standard clinical work. Indeed the goal of outcomes research is to integrate an evaluation process into standard clinical practice and to provide relevant and timely feedback to clinicians so that they might improve the quality of care provided. However, many healthcare payers (e.g., insurance companies) may consider that such attempts to evaluate clinical practice indicate that the practice is "experimental" and therefore not reimbursable. Consequently, the investigator is faced with the following dilemma: if the payers think it is research, they may not cover it. Indeed studies of treatments which are already standard clinical practice might be more analogous to Phase IV (i.e., postmarketing) clinical trials than to Phase III (i.e., effectiveness) clinical trials. This tautological problem within outcomes research may be solved in various ways, depending on the resources available to the institution. For example, some institutions may opt for internal funding of clinical work being investigated, thereby not seeking reimbursement for clinical services. Others may opt for the pre–post one-group design which may only require collecting data from patients during standard clinic visits. Others may hire independent private organizations to manage the data collection effort so that the apparent investigators are not obviously linked to the clinical facility. Whatever the solution, it will be advantageous for the clinical investigator to consider this tautological problem before initiating any research study.

Table 2 Feasible clinical health psychology research study designs: research questions and assumptions.

Study design	*Research question*	*Assumption*
a. Standard randomized controlled trial	Is treatment A better than treatment B on x outcome(s)?	Patients will accept randomization to a control group; randomization takes care of pretreatment group differences
b. Randomized consent	Is treatment A better than treatment B on x outcome(s)?	Patients will be likely to accept the treatment to which they have been randomized
c. Three-arm crossover	*i.* Is treatment A better than treatment B on x outcome(s)? *ii.* Does duration of treatment influence outcome(s)?	Washout period duration is known
d. Case–control	Is treatment A better than treatment B on x outcome(s)?	Groups are matched on relevant sociodemographic and medical variables
e. Solomon four-group pre–post	Is treatment A better than treatment B on x outcome(s)? Does measurement cue patients about salience and thereby influence treatment outcomes?	All groups are similar at baseline on outcome variables
f. Pre–post one-group	What factors are associated with better outcome(s) of treatment A?	Relevant factors are measured

8.05.2.4.2 Incentives of capitated care

Since health care becomes increasingly controlled by managed care, the incentives of this system must be considered by clinical investigators. The idea behind capitated care is that the provider is paid in advance for all medical care on a per patient (i.e., capitated) basis. Unlike the fee-for-service arrangement where the incentive is to provide more medical services, capitated care presents an incentive to minimize the medical services provided so that the cost per patient is reduced.

Given this incentive structure, research which seeks to evaluate adjunctive treatments such as behavioral medicine interventions may face additional barriers. Such programs may reduce unnecessary medical care use in the future by attending to current psychological morbidities. These morbidities can lead to increased healthcare utilization (Browne, Aprin, Corey, Fitch, & Gafini, 1990), but treating them with adjunctive behavioral interventions might lead to a cost offset in the longer term (Friedman, Sobel, Myers, Candill, & Benson, 1995). However, managed care providers may be focused on short-term savings, while psychological interventions may appear to increase rather than reduce healthcare costs. Investigators may need to consider that short-term increases in utilization may appear to work against the incentives of the system. Consequently short- and long-term costs and outcomes should be addressed in clinical research studies.

8.05.2.4.3 Legal considerations

Although informed consent has become an assumed step in any research process, there are still legal issues which may arise due to an increasingly fearful and litigious culture. Legal issues may play a larger role as clinical investigators are required to seek consultations from medical attorneys rather than focusing on the best research design to ask the research question. A real-life example of such a situation was one where researchers sought to evaluate the predictive value of inability to answer a standard quality-of-life questionnaire, the SF-36 (Ware, Snow, Kosinski, & Gandek, 1993). Their original design was a standard case–control design which followed elderly patients in a primary care setting. Eligible participants were to be those patients who were unable to complete the questionnaire without missing data. Half of the patients were to be observed without additional intervention, and half were to be telephoned monthly by a clinical nurse to ensure that they adhered to treatment regimens and/or saw a physician if needed. After consulting with lawyers, however, the investigators were informed that such a design put the clinic at risk of malpractice: if one of the control patients died and a relative learned that the patient had participated in this study, the relative might sue the clinic for withholding treatment even though the efficacy of the treatment was exactly the focus of the investigation. The researchers were advised to change their design so that half the patients got a pre- and post-test along with the telephone intervention. The other half provided data only at the post-test. Clearly such a design limited the investigators" ability to adjust statistically for baseline health status. However, the fear of possible litigation over an unproven treatment dominated the design selection. Future clinical investigators might be wise to consider the legal perspective, and to ensure informed consent so that scientific rigor does not lose priority to litigious concerns.

8.05.2.4.4 Unavoidable costs of doing research

No matter how much a clinical research study attempts to dovetail with standard clinical care, there are some costs which cannot be ignored. For a study to be rigorous and to maintain a reasonable response rate, some staff time will need to be devoted to the initial training as well as for implementing the Total Design Method (Dillman, 1978). Further, it may be necessary to pay for some of the clinicians' time for aspects of the research which do not strictly involve delivering clinical care. For example, some time will be needed for occasional meetings to confirm that the study procedures are clear and implemented correctly, for interpreting results of the data analysis, and for completing manuscript preparation for peer review. Finally, for studies implemented by mail or by telephone there will be costs associated with pretesting instruments, printing, postage, and telephone follow-up. Although these costs are minimal, there are necessary components of the research endeavor and should be considered before beginning the study.

8.05.2.4.5 Ethical considerations

The integration into any research endeavor of review by the appropriate committee for the protection of human subjects and informed consent have become crucial steps for maintaining ethically sound research. Nonetheless, some ethical concerns may be raised which are beyond the scope of such processes, especially as research efforts seek to extend clinical service evaluations to lower income patient populations. One example would be research which seeks to evaluate differences in clinical effectiveness of a behavioral medicine intervention on higher vs. lower income patients. Although standard informed consent procedures might list the risks and benefits of such benign interventions, they would be unlikely to highlight the impact of treatment termination. When the referent patient population is unlikely to be able to afford continuing the treatment after the study ends, it is incumbent upon the investigator to make the implications of treatment termination explicit. Further, the investigator may need to provide access to social service providers who can direct outreach after study termination and identify resources which would be otherwise unavailable to patients. Such considerations may not be apparent in the initial review by the committee for the protection of human subjects and may arise in the course of co-investigator meetings to clarify study procedures.

8.05.2.5 An Emerging Theoretical Challenge for Outcomes Research

An emerging area of interest and work by researchers from a broad range of disciplines is focused on understanding "response shift phenomenon." This term refers to the idea that individuals facing a significant health challenge may experience a change in internal standards, values and the meaning of quality of life (Breetvelt and VanDam, 1991; Schwartz & Sprangers, under review; Sprangers 1988, 1996; Sprangers et al., 1995; Sprangers, Rogemuller, VanDen Berk, Bowen, & VanDam, 1994). As both Heraclitus and Alphonse Karr noted long ago, change is a constant in life; yet there is an underlying and undeniable structure to personality (Funder & Colvin, 1991). Quality-of-life investigators have documented a type of inconsistency in self-reported health outcomes which is likely to have important implications for evaluating the impact of cognitively-based interventions.

Acute health-state changes may have an impact on psychological morbidity, followed by accommodation and adaptation to the functional limitations imposed by the illness. This shift explains how an individual's life satisfaction may not be directly related to their functional status. Bach and Tilton (1994) found, for example, that individuals with tetraplegia who were dependent on a ventilator reported higher life satisfaction than tetraplegics who were able to breathe independently. Individuals facing a significant health challenge may scale down their expectations of health, may be more appreciative of the social resources which support their daily living activities, and may be making significant adjustments in the importance of life domains (Bach & Tilton, 1994).

These individual values may also play an important role in determining the complex inter-relationships underlying quality of life. For example, satisfaction with one's functional status has been found to be related to psychological well-being only among those individuals who viewed the abilities being evaluated as very important (Blalock, B. M. DeVellis, R. F. DeVellis, & Santer 1993). Interventions which improve social support may affect patient values, priorities, and appreciation of the resources they have (Norman & Parker, 1996). These social support interventions may allow them to maximize their quality of life despite important physical setbacks. Thus, it may be hard to differentiate change due to active interventions and truly improved functional status from change due to patient accommodation to level of function.

Response shift represents a challenge to health researchers. It lies under the surface of measurement, camouflaged by an apparent lack of change in treatment outcomes. Intra-individual shifts in referents and priorities may mediate both well-being and functional status (Sprangers & Schwartz, under review). Understanding the predictive significance of this intra-individual variability is at the heart of meaningful outcomes measurement, and has the potential to lead to a paradigm shift in the many fields of investigation that rely on patient self-report.

8.05.3 SUMMARY

This chapter reviewed various design issues which might facilitate implementing health psychology intervention research in an active clinical setting. Key considerations were described to improve the ability to detect treatment effects, and several research design options were presented to highlight how one's assumptions and research questions determine the optimal research design. Finally, response shift phenomenon was discussed to emphasize how internal standards, values, and an individual's concept of quality of life can be a dynamic process which is under the surface of current measurement techniques. This emerging construct should be considered in future clinical health psychology research, so that the full impact of interventions is more likely to be revealed in clinical research.

ACKNOWLEDGMENTS

The author would like to thank Elissa Laitin for her assistance in manuscript preparation. The project was supported by grant number R01 HS08582-01A1 from the agency for Health Care Policy and Research to Dr. Schwartz.

8.05.4 REFERENCES

Bach, J. R., & Tilton, M. C. (1994). Life satisfaction and well-being measures in ventilator assisted individuals with traumatic tetraplegia. *Archives of Physical Medicine & Rehabilitation, 75,* 626–632.

Blalock, S. J., DeVellis, B. M., DeVellis, R. F., & Santer, S. C. (1997). Self-evaluation processes and adjustment to rheumatoid arthritis. *Arthritis & Rheumatism, 31,* 1245–1251.

Bornstein, M., & Cohen, J. (1988). *Statistical Power Analysis: A Computer Program.* Hillsdale, NJ: LEA Software and Alternative Media.

Breetvelt, I. S., & VanDam, F. S. (1991) Underreporting by cancer patients: The case of response-shift. *Social Science and Medicine, 32,* 981–987.

Browne, G. B., Arpin, K., Corey. P., Fitch, M., & Gafni, A. (1990) Individual correlates of health service utilization and the cost of poor adjustment to chronic illness. *Medical Care, 28,* 43–58.

Campbell, D. T. (1963) *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Boston: Houghton Mifflin.

Dillman, D. A. (1978). *Mail and telephone surveys: The total design method.* New York: Wiley.

Edwards, M., & White, N. (1987). A cancer patient survey to help determine psychosocial needs, design, and implement meaningful interventions. *Advances in Cancer Control—15 Years of Progress, 248,* 153–180.

Francis, D. J., Fletcher, J. M., Stuebing, K. K., Davidson, K. C., & Thompson, N. M. (1991). Analysis of change: Modeling individual growth. *Journal of Consulting and Clinical Psychology, 59,* 27–37.

Fleiss, J. L. (1986). *The design and analysis of clinical experiments.* New York: Wiley.

Friedman, R., Sobel, D., Myers, P., Caudill, M., & Benson, H. (1995). Behavioral medicine, clinical health psychology, and cost offset. *Health Psychology, 14,* 509–518.

Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology, 60*(5), 773–794.

Kazdin, A. E., & Bass, D. (1989). Power to detect differences between alternative treatments in comparative psychotherapy outcome research. *Journal of Consulting and Clinical Psychology, 57,* 138–147.

Kraemer, H. C., & Thiemann, S. (1989) A strategy to use soft data effectively in randomized controlled clinical trials. *Journal of Consulting and Clinical Psychology, 57,* 148–154.

Kraft, G. H., Freal, J. E., & Coryell, J. K. (1986). Disability, disease duration, and rehabilitation service needs in multiple sclerosis: Patient perspectives. *Archives of Physical Medicine & Rehabilitation, 67,* 164–168.

Lechtenberg, R. (1990). *Seizure recognition and treatment.* New York: Churchill Livingstone.

Mohr, D. C., Goodkin, D. E., Gatto, N., Neilley, L. K., Griffen, C., Likosky, W., & Stiebling, B. (1996). Therapeutic expectations of patients with multiple sclerosis upon initiating interferon beta 1-b: Relationship to adherence to treatment. *Multiple Sclerosis, 2,* 222–226.

Morgan, D. L. (1988). *Focus groups as qualitative research.* Newbury Park, CA: Sage.

Norman, P., & Parker, S. (1996). The interpretation of change in verbal reports: Implications for health psychology. *Psychology and Health, 11,* 301–314.

Pocock, S. J. (1983). *Clinical trials: A practical approach* (p. 182). New York: Wiley.

Schwartz, C. E. (1994, July). *How do psychosocial interventions influence functional status in multiple sclerosis? Results of a randomized trial* [Abstract]. Proceedings of the Third International Congress of Behavioral Medicine, Amsterdam.

Schwartz, C. E. (under review). The psychosocial impact of two social support interventions: Results of a randomized trial.

Schwartz, C. E., Chesney, M. A., Irvine, M. J., & Keefe, F. J. (1997). The control group dilemma in clinical research: Applications for psychosocial and behavioral medicine trials. *Psychosomatic Medicine, 59,* 362–371.

Schwartz, C. E., & Rogers, M. (1994). Designing a psychosocial intervention to teach coping flexibility. *Rehabilitation Psychology, 39*(1), 57–72.

Schwartz, C. E., & Sprangers, M. (under review). Methodological approaches for assessing response shift in longitudinal quality of life research.

Slife, B. D., & Williams, R. N. (1997). Toward a theoretical psychology: Should a subdiscipline be formally recognized? *American Psychologist, 52,* 117–129.

Sprangers M. (1988). *Response shift and the retrospective pretest: On the usefulness of retrospective pretest–posttest designs in detecting training related response shifts*. Amsterdam: SVO.

Sprangers, M. (1996). Response-shift bias: A challenge to the assessment of patients" quality of life in cancer clinical trials. *Cancer Treatment Reviews, 22,* 55–62.

Sprangers, M., Broersen, J., Lodder, L., Wever, L., Smets, E., & VanDam, F. S. (1995). The need to control for response shift bias in longitudinal quality of life research [Abstract]. *Quality of Life Research, 4,* 488.

Sprangers, M., Rozemuller, N., Vanden Berk, M. B. P., Boven, S. V., & VanDam, F. S. (1994). Response shift bias in longitudinal quality of life research [Abstract]. *Quality of Life Research, 3,* 49.

Sprangers, M., & Schwartz, C. E., (under review). Integrating response shift into health-related quality-of-life research: A theoretical model.

Thoresen, C. E. (1991). Long-term results of recurrent coronary prevention project at eight years. Invited paper presentation at the First International Congress of Behavioral Medicine, Uppsala, Sweden.

Ware, J. E., Snow, K. K., Kosinski, M., & Gandek, B. (1993). *SF-36 Health survey: Manual and interpretation guide*. Boston: The Health Institute.

Zelen, M. (1979). A new design for randomized clinical trials. *New England Journal of Medicine, 300,* 1242–1245.

Zelen, M. (1990). Randomized consent designs for clinical trials: An update. *Statistics in Medicine, 9,* 645–656

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.06
Consultancy

SUSAN MICHIE
United Medical and Dental Schools of Guy's and St. Thomas's Hospitals, London, UK

8.06.1 INTRODUCTION 153
  8.06.1.1 Clinical Health Psychology 154
  8.06.1.2 What is Consultancy? 155
  8.06.1.3 Issues for Consultancy 155
  8.06.1.4 A Systemic Approach 155
  8.06.1.5 Who are the Consultees? 156
8.06.2 WHAT DOES CONSULTANCY INVOLVE 158
8.06.3 EVIDENCE-BASED CONSULTANCY 159
8.06.4 OPPORTUNITIES FOR PSYCHOLOGY CONSULTANCY 159
8.06.5 PROMOTING THE CONSULTANCY ROLE 160
8.06.6 EXAMPLES OF CONSULTANCY WORK 161
  8.06.6.1 Developing and Evaluating Written Material for Patients 161
    8.06.6.1.1 The breast clinic 162
    8.06.6.1.2 The colposcopy clinic 162
    8.06.6.1.3 The rehabilitation department 162
  8.06.6.2 The Health and Well-being of Health Professionals 162
    8.06.6.2.1 Stress management 162
    8.06.6.2.2 Health checks and feedback 163
    8.06.6.2.3 Stress in medical school 163
    8.06.6.2.4 Twelve-hour nursing shifts 164
  8.06.6.3 Health Promotion in Primary Care 164
    8.06.6.3.1 Individual lifestyle interventions in primary care 165
    8.06.6.3.2 Group-based health promotion interventions in primary care 165
    8.06.6.3.3 A consultancy framework for developing health promotion intervention programs 166
8.06.7 THE FUTURE: AVOIDING GAPS AND TRAPS 167
8.06.8 REFERENCES 167

8.06.1 INTRODUCTION

In the future,

> In the health care arena, psychologists will function less within an independent (autonomous) practice model and more within organised health care settings (Pion, 1991). The emphasis within these organisational frameworks will be on an interdisciplinary, collaborative model with other health care professionals in which psychologists will have shared responsibility and decision-making for patient care. (Siegel, 1995, p. 342)

The development of new health technologies and health care providers in the developed countries is leading to an enlarging "health industry." The way in which health services are organized is changing rapidly and profoundly on both sides of the Atlantic. These developments provide increased opportunities for professionals concerned with health issues, and also create a need for health professionals to examine their role and contribution. Several innovations demand expertise in both measuring and improving behavioral and psychological processes (Johnston & Weinman, 1995). For example, the growing emphasis in health care on the prevention of disease has led to a need for professional skills relevant to understanding and altering health-related cognitions and behaviors. The increased prevalence of chronic rather than acute disease has created a need for the understanding of emotional responses and problem-solving coping skills, in addition to existing biomedical knowledge. Changes in the structure of the health services in the 1990s has increased the need to evaluate the impact of these services on the users.

In the UK, the control and direction of health care services have moved from practitioner to purchaser, and from clinicians to general managers and accountants. This poses the challenge of whether health care professionals can adapt to these changes and whether they can influence the changes to achieve their objective of improving the health and well-being of the population or of sections of the population. This chapter describes the emergence and skills of clinical health psychologists, the consultancy role and the stages it involves, its possibilities and constraints and the responsibilities of psychologists in promoting themselves and their disciplines to existing and potential consultees. Whilst many of the examples are from the UK, the issues raised are of more general relevance.

8.06.1.1 Clinical Health Psychology

Health psychologists have much to contribute to the prevention and management of problems within organizations, and have the skills to function in a variety of settings (Altman & Cahn, 1983). The essence of training in psychology is learning about the factors that affect the causes, maintenance, and changes of human functioning at the individual, group, and organizational levels. Human functioning is considered in its component parts (behavior, thinking, and emotion) and an analysis is made of the relationships between these components, and of the relationships between these components and situational factors. Health psychology is a broad discipline: for convenience, content areas have been summarized into five broad areas (Johnston & Weinman, 1995).

(i) *Health risk behaviours*. The nature, causes, and health effects of such behaviors as smoking, substance abuse, unsafe sexual behavior, and harmful dietary choice; the role of stress and dispositional factors in the etiology and exacerbation of disease; psychological theories and interventions for primary prevention and health-related behavior change in community and workplace settings.

(ii) *Health protective or enhancing behaviors*. The uptake and maintenance of such behaviors as exercise, healthy dietary choice, teethbrushing, health checks/self-examination, and attendance for preventive medical screening; psychological approaches to health promotion.

(iii) *Health-related cognitions*. Cognitive processes that mediate and determine health and illness behaviors. These include general health beliefs/attitudes; attributions, control, and self-efficacy beliefs; symptom perception; and illness representations in children and adults. There are a range of theoretical models and frameworks that can be used, not only for explanation and prediction, but also for developing interventions (e.g. to modify health beliefs, increase internal control, or self-efficacy beliefs).

(iv) *Processes influencing health care delivery*. The nature and effects of communication between health care practitioners and patients; patient satisfaction and adherence to treatment and advice; the uptake and response to preventive services such as screening; stressful medical investigations and treatment; hospitalization; health care staff attitudes; decision-making, work stress; quality assurance; measurement issues in audit and outcome evaluation; psychological interventions to improve communication, facilitate adherence, and prepare for stressful medical procedures.

(v) *Psychological aspects of illness*. The psychological impact of acute and chronic illness on individuals and their families and carers across the lifespan; pain behavior and rehabilitation and health-related behavior change; dying, death, and bereavement. Psychological intervention to promote self-management, facilitate coping with pain or illness, to improve quality of life, and to reduce disability and handicap.

In all these areas, a wide range of skills is used, including assessment, evaluation, communication, teaching, and intervention. The work of a health psychologist is targeted at the problems of public health, the provision of health care and the response to illness. The problems may be identified by health care agencies, health professions, and organizations and employers outside the health care system. Psychologists can bring a systems approach to problems and locate them and the health care system within their wider contexts. By putting these skills together, psychologists are able to address “what if” questions by developing theoretical and quantitative causal models of behavioral predictors.

Health psychology is a new discipline. The first book including the term “health psychology” in its name was published in 1979 (Stone, Cohen, & Adler, 1979). Health psychology is defined as the study of psychological and behavioral processes in health, illness and health care (Johnston, 1994). Clinical health psychology has been used to refer to clinical, organizational, or health psychology in a health-service setting. There are two primary reasons for interest in the contribution of clinical health psychology to other health professionals and to health service managers (Jansen, Methorst, & Kerkhof, 1991). The first is the increasing recognition that solutions to health problems may be found by seeking behavioral solutions (e.g., risk-factor management in the case of heart disease). The second is the escalating cost of health care, which fosters a welcoming of prevention and of interventions that reduce reliance on the medical system (e.g., psychological interventions for chronic pain management).

Clinical health psychology aims to translate research into practical applications that can help people lead healthier lives. This may take several routes: teaching and training health professionals and patients; working with patients to adopt healthier lifestyles or to manage problems such as pain or stress; and consulting with professionals about patient, staff, or organizational issues. Many research findings are not applied and do not influence clinical practice or policy, a situation that is itself the subject of analysis and study by health psychologists (Marteau, Sowden, & Armstrong, submitted).

The knowledge bases required for a successful health psychology consultancy role are theoretical models, (see Chapter 8.01, this volume) measurement techniques (see Chapter 8.04, this volume) successful interventions, factors that influence behavior and behavior change (a portfolio of examples that is currently being collected by the British Psychology Society Special Group is cited here), and specific health psychology knowledge (e.g., psychophysiological processes, health behavior, stress, communication and decision-making in health care, consequences of and coping with illness) (see Chapters 8.02 and 8.07, this volume).

8.06.1.2 What is Consultancy?

The term “consultation” refers to one party (individual, committee, or organization) formally seeking the advice of another. The consultee poses the initial questions and the consultant shapes them, proposes a methodology, implements it, and reports back. Health psychology consultancy is the use of these psychological skills and knowledge to facilitate, develop, or enhance the effectiveness of others in the maintenance and improvement of human health.

Consultancy has the potential to influence large groups of people and to change processes that depend on many and varied people and organizational policies, procedures and practices. As with other health services, consultancy must be evidence based, and not just taken on trust. In this age of squeezed resources for health care, consultancy must be shown to be effective and cost-effective. This requires making the consultancy component of health psychology explicit and embedding it within a research framework.

In order to expand their role to include consultancy, psychologists need to apply the skills and knowledge acquired in working with individual problems to working with organizational and service problems (Michie, 1993). Areas in which psychologists have expertize relevant to consultancy work include:

(i) understanding the contextual nature of health problems;

(ii) the systematic analysis of problems and potential solutions;

(iii) the understanding of change processes that enable recommendations about the best ways of achieving the task or goal identified by the consultee; and

(iv) research skills that enable the evaluation of services, interventions, and changes.

8.06.1.3 Issues for Consultancy

The main issues for clinical health psychology consultancy are:

(i) Patients: improving care and decreasing patient anxiety.

(ii) Staff: risk assessment, reducing stress, reducing ill-health, and increasing job satisfaction.

(iii) Improving communication: between staff and between staff and patients.

(iv) Assessing need for services.

(v) Service evaluation, quality assurance, and audit.

8.06.1.4 A Systemic Approach

A systems framework for health is described well by Stone (1979, 1982). A system is a set of elements related to each other in such a way that the actions of one element affect the states of, and initiate or modify activity in, other elements of the system. A goal of health psychology is described as discerning the relationships between health activities and the entities that appear in the activities. The health system is seen as being much broader than the health care system, including entities such as environmental health hazards and health resources, public health measures and planning agencies, health education, and political behavior.

The health care system is composed of people, organizations, and machines that protect, enhance, or restore health. The people may be users, providers, family members, policy-makers, educators, or researchers. The organizations are any collections of individuals who act together to promote good health (e.g. community organizations, hospitals, or government departments). The machines are usually computers. The broader world of health hazards includes toxins, pathogens, stressors, and safety hazards such as defects in tools or materials. Health resources include knowledge and technical skills, leisure, recreation and health care facilities, material commodities such as housing and money, tools and machines, and regulations, laws, and procedures.

Health psychologists seek to understand the system in which, or with which, they are working, in order to develop ideas about how it could be improved and to intervene to change it, if necessary. Their contribution lies both in communication skills to gain the agreement of relevant people that change is desirable and in developing strategies to achieve the desired changes. Other skills that psychologists have are the ability to make predictions about the effects of change in particular areas, to test those predictions, and to interpret the results within a theoretical framework. These are all skills that are required in organizational consultancy.

The work of health psychologists has been described as lying on a dimension, extending from basic research to the design and development of new interventions, to the implementation of interventions in controlled trials, to the evaluation of programs of service delivery (Stone, 1982).

Organizations have certain tasks to perform, and patterns of behavior and beliefs about the organization develop around the performance of these tasks. Consultants describe and examine patterns of beliefs and behavior around specific tasks and the effect of these patterns on relationships and communication throughout the organization (Campbell, Draper, & Huffington, 1989). Observing, understanding, and developing hypotheses about the way in which these patterns or routines protect certain relationships are key to enabling organizations to change. Organizations change when people's perceptions, beliefs, and behavior change. This can be achieved by systemic interviewing, feedback within the organization, and targeted interventions. Systemic interviewing involves asking about the way in which people think, the way in which their beliefs affect their actions and relationships in such a way that challenges discrepancies between beliefs and actions.

8.06.1.5 Who are the Consultees?

Consultees are individuals, groups, or organizations that enter into a negotiated contract with the consultant, agreeing the objectives, process, and conditions of the organizational consultancy work. They may be those directly providing services, both medical and nonmedical, they may be those planning and managing these services, or they may be those purchasing services (see Figure 1). Most would not think of psychologists as the appropriate professionals to help them. However, the problems that they face are often ones of behavior and of behavior change. Management problems in audit and evaluation, in implementing policies and guidelines, and in developing and maintaining teams are examples of this.

In the UK, about 40% of health psychologists are employed by health services, either to help develop and evaluate services to patients or to deal with general problems of quality assurance, preventive health care, intervention at an organizational level, or to assist in setting and evaluating targets for health care. Examples of psychologists providing consultancy in the UK health service can be found in *The change directory: Key issues in organisational development and the management of change* (Brunning, Cole, & Huffington, 1990) and *Internal consultation in health care settings* (Bor & Miller, 1991). Similar work in other areas of the educational and social services is described in *Internal consultancy in the public sector: Case studies* (Huffington & Brunning, 1994).

In the USA, business and industry pay for more than 25% of the nation's health care costs, giving employers a direct financial interest in developing health promotion programs. In response to this, occupational health psychology has developed a "systems consultation" model (as described above) to help managers identify employees' needs, to meet these needs with appropriate forms of intervention, to communicate effectively these needs and expectations to the individuals responsible for the development of health promotion programs, and to ensure that managers' expectations of success are reasonable and in line with current capabilities and resources (Everly, 1986).

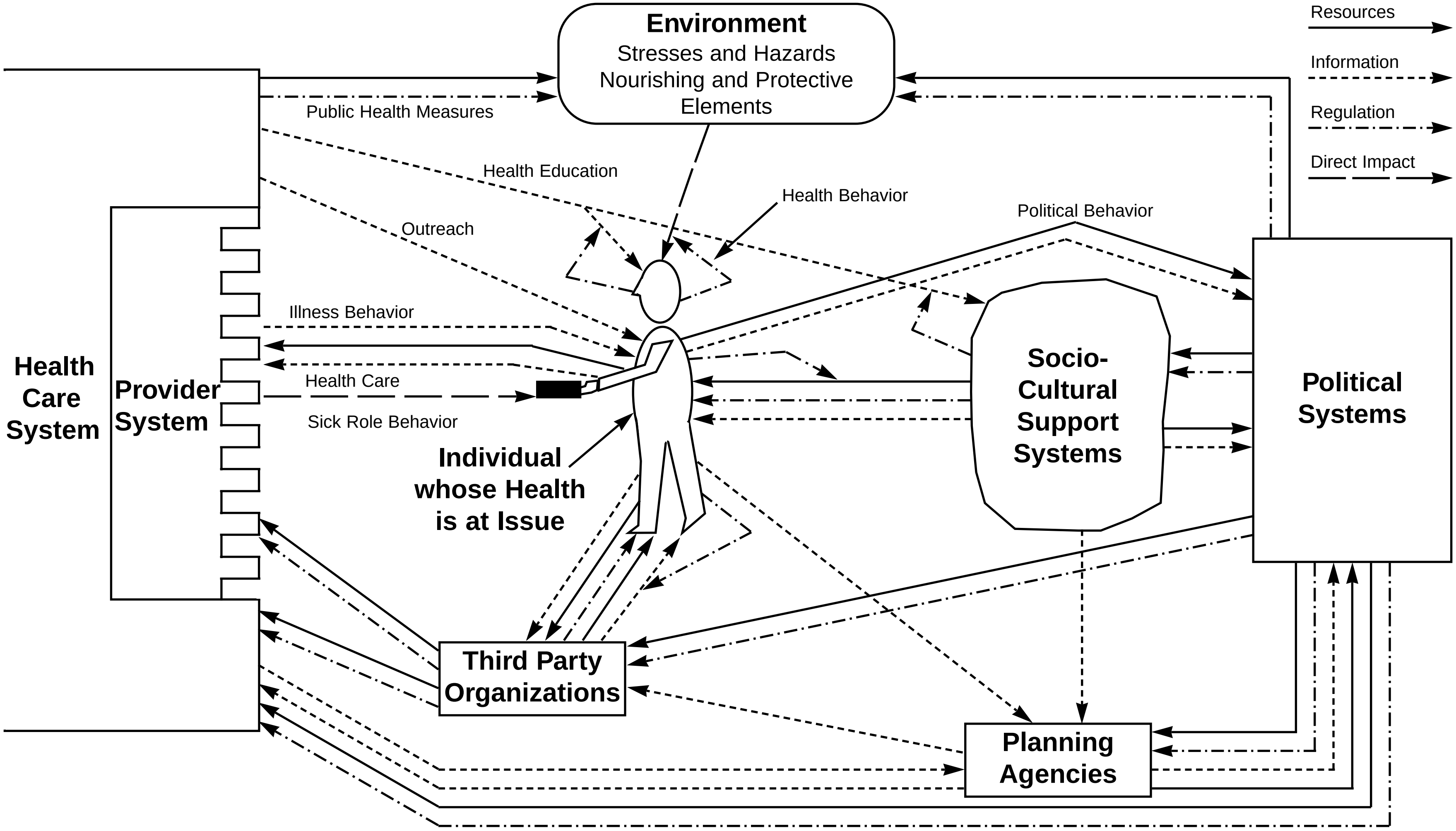
Environment
Stresses and Hazards
Nourishing and Protective Elements
Resources
Information
Regulation
Direct Impact
Public Health Measures
Health Education
Outreach
Health Behavior
Political Behavior
Illness Behavior
Health Care
Sick Role Behavior
Health Care System
Provider System
Individual whose Health is at Issue
Socio-Cultural Support Systems
Political Systems
Third Party Organizations
Planning Agencies

Figure 1 The health system (source: Stone 1982).

8.06.2 WHAT DOES CONSULTANCY INVOLVE

There are five stages to consultancy: the approach from the consultee, refining the question, developing the methodology, implementing the methods, and reporting back. Although these stages are sequential in principle, there is a flexibility in practice, with overlapping activities and revisiting of earlier stages as the work develops. Practical examples of these stages are given in Section 8.06.6.

The approach from the consultee involves the following:

(i) being clear about who is the consultee (e.g., manager, clinician), what the initial questions are, why the questions are being asked, what the problem is that needs solving, (e.g., lack of resources, interpersonal conflict, the consultee's anxiety);

(ii) discovering what answers the consultee is seeking (e.g. literature review, descriptive research, an intervention); and

(iii) establishing the nature of the approach (e.g. an informal contact, a formal research initiative, a direct approach for consultancy).

Refining the question requires:

(i) an analysis of the situation presented by the consultee and the formulation of problems such that they are specific and testable, with clear operationalized objectives;

(ii) acquiring relevant background information (e.g., understanding formal and informal systems of communication within the organization, knowing who are the key individuals and committees with power and influence, and anticipating potential barriers to consultancy work within the organization).

(iii) defining the proposed product of the consultation (e.g., a literature review, a psychological opinion, a new procedure, a descriptive study, a study of factors associated with, antecedent to or consequences of an event or phenomenon, an evaluation of the consultee's intervention, an evaluated psychological intervention, which may be new or an adaptation of an intervention from elsewhere or from research literature); and

(iv) the ability to win consent for evaluation at an early stage of the consultation, including developing a baseline.

Developing the methodology requires:

(i) A methodology that allows the generation of specific answers to specific questions. This may involve a systematic literature review (Chalmers & Altman, 1995), qualitative methodology (Smith, Harre, & Langenhove, 1995), or research designs for descriptive studies (see Chapter 5, this volume).

(ii) The use of valid and reliable measures to determine whether objectives have been achieved and to test any research questions within the consultation (see Chapter 8.04, this volume). This will require knowledge of the relevant research literature, and an ability to use databases and to design and pilot measures where appropriate ones do not already exist. A comprehensive range of current psychometric measures are given in *Measures in health psychology: A user's portfolio* (Johnston, Weinman, & Wright, 1995).

(iii) The ability to include a simple economic evaluation.

(iv) Strategic thinking (e.g., anticipating hurdles and barriers and developing plans for overcoming these, planning for any anticipated adverse consequences).

(v) Negotiating access to resources and winning agreement from all necessary individuals, groups, and departments.

(vi) Identifying the blocking and facilitating forces within the organization.

(vii) Negotiating access to all information sources, and identifying relevant gatekeepers and those who must give approval.

(viii) Setting agreed boundaries (e.g., confidentiality issues, time constraints, starting and finishing dates, review dates, location, report writing and publications).

Implementing the methods requires:

(i) Achieving the understanding and consent of the consultee and other relevant individuals, groups, and departments. This involves a sensitivity to the consultee's agenda, both explicit and hidden, and to potential threats posed by the consultancy process.

(ii) Involving the appropriate individuals and groups in the relevant aspects of the process (e.g., student health professionals in data collection).

(iii) Assigning clear roles and responsibilities to those involved, including the consultant and the consultee, and a clear statement of these to the relevant people.

(iv) Appropriate and accurate data collection, analysis, and interpretation.

(v) The maintenance of a professional role.

(vi) Drawing up a written contract of the objectives, process, and conditions of the consultancy work, and defining the resources required for all stages of the consultancy process.

(vii) Clear communication and effective meetings.

Reporting back requires:

(i) The ability to give feedback to the consultee and other relevant audiences in appropriate ways such that maximum effect is achieved.

(ii) Clear and realistic recommendations.

(iii) Negotiating where to publish and present the results and who the authors should be.

Outlets may include those relevant to the consultee's discipline, to health psychology, or to multidisciplinary audiences.

(iv) Dissemination of results, including both informal and formal written reports, and verbal presentations.

8.06.3 EVIDENCE-BASED CONSULTANCY

If the consultancy process is to be evaluated, its objectives must be defined and measurable. This is impossible in situations in which the consultancy role adopted by health psychologists is not made explicit. The outcome criteria used in evaluations may be very varied (e.g., staff or patient satisfaction, financial savings, and adherence to guidelines). Many consultees do not appreciate the importance of a scientific approach to service evaluation, including the evaluation of consultancy work. Psychologists need to introduce the importance of well-designed and conducted evaluation at every opportunity in their contact with other health professionals.

Psychologists have an important role in defining the outcomes used in evaluation. Instead of using a biomedical disease model, in which signs and symptoms are reflections of tissue pathology and patient reports are often discounted, psychologists use a behavioral model. The important indicators of health and well-being are seen as behavioral-overt behavior, cognitions, and emotions. Outcome measures in health and medicine should therefore be anchored in their relationships with behavior (Kaplan, 1994), and psychologists are the obvious professionals with expertise in this area.

8.06.4 OPPORTUNITIES FOR PSYCHOLOGY CONSULTANCY

There has been a fear among UK psychologists that they could become marginalized by health service changes (Ovretveit, Brunning, & Huffington, 1992). This situation has been mirrored in Australia:

> Lack of knowledge about the consequences of organisational change resulting from restructuring, and exclusion from a strategic influence in the process of change, makes allied health professions vulnerable to a substantial reversal of both power and prospects. (Boyce, 1991)

Alongside the possible intended benefits of health service re-organization are possible problems for staff and for health care organizations. Staff may become anxious or demoralized by uncertainty and by change, with the consequence that they work less effectively, with more absenteeism and less job satisfaction. Services may suffer if staff have less motivation, time, and energy to develop services and to improve their quality. The organization may become less effective and efficient if apathy, poor communication, and conflicts between individuals, teams, and departments develop. However, organizational changes are a potential opportunity, as well as a threat. The very problems that may be produced by the changes and the tasks faced by those trying to manage the changes give the potential for a new role for psychologists, that of organizational consultant. Psychologists are the group of health professionals seen as having skills and knowledge relevant to the process of change.

The importance of understanding the role of broader forces within society has been emphasized in a volume considering the international development of health psychology (Jansen & Weinman, 1991). The future shape of health psychology will be influenced both by the needs of society and developments within medicine, and by pressures within the discipline. In order for psychologists to understand the ways in which social conditions affect health and health-related behavior, they need to understand and collaborate with a variety of other disciplines (e.g., epidemiology and researchers of consumer behavior).

As health psychology is a relatively new discipline, jobs and consultancy contracts will often have to be created. It is not enough for health psychologists to have skills relevant to situations in which they have no widely recognized track record. They will have to define the job tasks, clarify the skills needed, and demonstrate to potential employers that their skills and orientation will make a contribution (Altman & Cahn, 1987). To a large extent, health psychologists will have to create their own job market. This will require an understanding of, and an ability to engage with, what potential employers want. A survey of potential employers of applied psychologists in the USA in 1979 found that the most important skills sought were the ability to communicate, interview, conduct impact evaluations, plan human resources, use statistics, and have knowledge of the particular area (Edwards & Holmgren, 1979). In a qualitative study of the future organization of six therapy services in the UK, Ovretveit (1992) summarized some of the elements needed for their consultancy role: to know how the health service market works, to attend to what purchasers want, and to judge how they may be persuaded to finance the service. He advised the development of strategic plans, which should include an analysis of the "business environment," a rationale for developing activities, the resources needed (people, capital, and revenue), and the end points to be reached (strategic objectives and a business plan).

Psychologists have a well-established role of providing psychological services directly to patients. They are also known for the research they conduct. The consultancy role is less associated with psychologists. In the UK, health psychologists are becoming increasingly aware that this is a role that is already being played, and one that could be developed with good effect for the quality of services, for multidisciplinary work and for the spread of psychological knowledge and skills. The potential for consultancy work by psychologists is much greater than is currently demonstrated and recognized. The realization of this potential will partly depend on the willingness of psychologists and the profession to develop their consulting role (Ovretveit et al., 1992).

A broader perspective for translating research findings into effective practical programs has been taken by Mary Jansen and colleagues in the USA (Jansen, Methorst, & Kerkhof, 1991). As well as having a knowledge of health care systems, psychologists should be familiar with media techniques to promote their messages and their work and thereby play a greater role in the political process. Cynthia Belar (1995) makes the point that health psychology is not just clinical psychology in a medical setting, citing examples of health psychologists being consulted by a government on noise abatement, designing advertising campaigns to increase the use of safety belts, and being consulted by schools about smoking prevention.

National and local governments are seen as ideal employers of health psychologists since they regulate health service delivery and costs, formulate laws to regulate health behavior (e.g., drink driving), and regulate public health activities affecting the environment (e.g., the Clean Air Act). Potential positions for health psychologists have been identified as:

> working with legislative committees specialising in health issues (e.g., House and Senate subcommittees), executive agencies and administrative departments (e.g., Health Care Financing Administration, National Institutes of Health, National Institute of Mental Health, and Environmental Protection Agency), government research organizations (e.g., Congressional Research Service, General Accounting Office, and Office of Technology Assessment), special bodies empowered to help control health costs and quality (e.g., health systems review bodies, professional review organizations), trade associations, and lobbyists who provide government officials with information. (Altman & Cahn, 1987, p. 238)

Psychology has the potential to be an influential partner in the development and management of health care resources and systems, but only if it widens its role. Elliott and Shewchuk (1995) state that:

> For this change to occur, psychologists must assume a more visible profile in lobbying at local, state and national tiers to ensure that behavioural health issues are included in policy formation, research agenda and the allocation of services. If professional psychology fails to assume this responsibility and meet this challenge, the profession risks being marginalised as a stake holder in the health care system. (p. 346)

In some cases, targeted outcomes and intervention may be more appropriate and effective at the level of society and community rather than at the clinic and organization level. There needs, however, to be clarity about the role being played by psychologists in different situations. Lobbying, campaigning, advocating, and influencing those with political power are political activities and may be very appropriate for the objectives of health psychologists. On the other hand, consultees need to be confident that consultancy work will be carried out within a neutral, rather than a political, perspective. Having said this, it is doubtful that neutrality can ever be achieved, and a better approach may be to be explicit about one's beliefs about issues relevant to the consultancy work.

8.06.5 PROMOTING THE CONSULTANCY ROLE

Much consultancy by applied psychologists goes on outside a formally contracted arrangement, which means that it may not be recognized, either by the consultee or the consultant. In order to help develop this area of work, psychologists should be encouraged to formalize and evaluate the consultancy work they carry out and to publish the results. Psychologists need first to establish what potential employers want, which can often be done by obtaining their annual reports, strategic plans, and other relevant documents. The next step is to convince them that they have the expertise to meet their needs, to achieve agreed tasks within the time and cost constraints, and that their track records are reasons for the employer to choose them. What hinders many psychologists is not their lack of skill but their lack of focus on what they can deliver, and their diffidence in persuading others that they have something to offer of a quality that cannot be obtained from others.

The point is well made if the term "health psychology" is substituted for "your company" in the following statement, cited in Altman and Cahn (1987):

> I don't know who you are.
> I don't know your company.
> I don't know your company's product.
> I don't know what your company stands for.
> I don't know your company's customers.
> I don't know your company's record.
> I don't know your company's reputation.
> Now—what is it you wanted to sell me?

One way to promote psychologists as consultants to potential purchasers would be for psychological professional associations to collect case studies of successful consultancy work. Producing an informative and attractive summary of the consultancy services on offer, in leaflet, video, or internet homepage form, is essential. Purchasers should be encouraged to write consultancy into the service contracts and job descriptions of psychologists. The training of consultancy skills should be included in postgraduate courses of applied psychology and as part of continuing professional development

8.06.6 EXAMPLES OF CONSULTANCY WORK

Examples of specific questions raised with a clinical health psychologist providing a consultancy service to UK health care organizations are listed below (taken from Michie, 1993, and Earl L., 1997, personal communication).

(i) Questions from managers:

(a) Hospital

What will the effects be on staff and patients of decentralizing our outpatient service into clinic-based teams?

How can we develop an integrated oncology service from the present divided service?

What are the reasons for the very high absenteeism on a particular ward and what recommendations would you make to improve the situation?

How can we best assess patient satisfaction with our services?

Is a new receptionist skills training program effective in changing staff behavior and in improving patient satisfaction with their reception within the hospital?

(b) Community

What are the problems for the various staff groups in introducing parent-held health records and how could they be overcome?

What are the effects of changing health visitors' work from individual to "corporate" caseloads?

What can be done to improve the communication and integration of work between community-based nurses and nurses working in doctor-led primary care teams?

(ii) Questions from clinicians:

Is the way in which we advise patients with diabetes about its management contributing to the worryingly high rate of eating disorders we are finding?

Can we do something to reduce the very high anxiety in women referred with suspected breast cancer at their first hospital appointment?

Is the arthroscopy patient information booklet and video helpful and, if so, how?

Is the oncology counseling service effective?

(iii) Questions from purchasers:

(a) Health authority

What are the dimensions of quality of health care valued by the local population and are they sufficiently represented in our contracts with provider units?

What are the priorities of maternity service users and how far do current services meet changing childbirth national guidelines?

(b) Multiagency:

What is the current service provision for children who have been sexually abused, the perpetrators, and nonabusing carers?

What are the gaps in service and recommendations for resource allocation?

Although there are many examples of work such as this, there is a paucity of consultancy work published in peer-reviewed journals. This limits the possibility of others learning from and building on the work and limits the promotion of psychologists as competent and experienced in carrying out this type of work. One example of published consultancy work comes from a health psychology unit based in the medical school of a London teaching hospital. This unit was set up in the late 1980s by Professor Marie Johnston to integrate clinical, organizational, and health services research work. In carrying out consultancy work, the team sought to include psychological questions within the projects undertaken and to ensure that the work was designed in such a way that it would be of publishable standard.

Consultancy projects were primarily in two areas: developing and evaluating written material for patients, and the health and well-being of health professionals. Some examples are described below. This is followed by a description of consultancy work carried out within a Health Promotion Unit by a UK psychologist, Barbara Wren.

8.06.6.1 Developing and Evaluating Written Material for Patients

One of the interests of this group was the effect of written information on patient outcome. Clinicians in several specialities were aware of the problem of high anxiety in patients referred for investigation and approached the health psychology unit for help. Several studies were carried out and the results were reported both to the appropriate parts of the commissioning organization and to the wider research community, the latter being in the form of publications and conference presentations.

8.06.6.1.1 The breast clinic

The consultant breast surgeon was concerned by his clinical experience of very high levels of anxiety shown by women attending a breast outpatient clinic for the first time, and approached the health psychology unit for advice as to how this might be prevented or reduced. The first contribution of the health psychologists was their knowledge of previous relevant research. This had suggested that written information may improve patients' feeling of well-being and make them feel better informed and less worried (Eardly, 1988; Hutchcroft, Snodgrass, Troyan, & Wares, 1984; Wallace, 1984). On the basis of this, a booklet aimed at addressing procedural worries by giving information and coping advice was designed and given to women within a randomized controlled trial (Madden, Johnston, & Parbhoo, 1994). This study confirmed the clinical observation that this group of patients was highly anxious, and also found that their most dominant worry was more often about the outcome of their appointment rather than about the procedure. Although the booklet increased women's knowledge, it did not reduce their anxiety. These results led to a subsequent study comparing three kinds of preparatory information: that dealing with clinic procedures, that dealing with the outcome of breast problems, and that designed to increase perceived control (Michie, Rosebert, Heaversedge, Madden, & Parbhoo, 1996). Women receiving information were found to be less anxious than the control group and perceived their problem as less serious. There was one difference between the types of information booklet: those receiving coping or outcome information were less worried about what the doctor might find than those receiving procedural or no information.

8.06.6.1.2 The colposcopy clinic

Another group of patients reported by clinicians at this hospital to be highly anxious was women referred for colposcopy following an abnormal cervical smear result. The effectiveness of two booklets, one produced nationally and the other designed locally to be simpler and shorter, was evaluated in a randomized controlled study (Marteau, Kidd, Cuddeford, & Walker, 1996). Although both booklets led to increased knowledge, only the simpler one reduced anxiety. In addition to using simpler language, the shorter booklet also contained information about coping strategies and was associated with the clinic that the women were attending. It was recommended to the clinic that they use this booklet.

8.06.6.1.3 The rehabilitation department

A different use of written information was used in response to a request by the hospital's physiotherapy department to help recovery during rehabilitation. A previous study had found that patients with a higher level of perceived control over recovery made more progress with recovery from disability than did those with a lower level of perceived control (Partridge & Johnston, 1989). A simple intervention was designed with the aim of increasing patients' sense of control over their recovery (Johnston, Gilbert, Partridge, & Collins, 1992). Using a randomized controlled design, patients were sent an appointment letter with additional paragraphs under the heading "Information sheet for patients receiving physiotherapy." Patients receiving this letter were found to have higher levels of perceived control and tended to be more satisfied with information than did those who received the standard letter. The extended letter was therefore used routinely by the physiotherapy department.

8.06.6.2 The Health and Well-being of Health Professionals

The same health psychology unit was also approached to carry out various projects concerning staff health, since this was recognized as being important both in its own right and as a factor in delivering high-quality patient care.

8.06.6.2.1 Stress management

We were asked by the Director of Nursing and the General Manager of the hospital to assess stress levels amongst nurses, and found they reported greater problems in dealing with stress than the norm (Michie, Ridout, & Johnston, 1996). This was tackled in three ways: introducing stress management training courses, offering a counseling service based at the hospital's Occupational Health Unit, and running a health screening and advice program for staff.

Although the first two initiatives were not evaluated within a randomized controlled trial, comparison of scores before and after the interventions showed that there was lower anxiety and increased job satisfaction after the stress management training (Michie & Ridout, 1990). Six months after counseling, staff were less anxious and depressed, more satisfied with their work and life outside work, and perceived themselves to be functioning better at work compared to before counseling (Michie, 1992a, 1992b). There were fewer days and episodes of sickness absence in the six months following counseling compared with the equivalent period before counseling (Michie, 1996).

8.06.6.2.2 Health checks and feedback

The Occupational Health Unit of the hospital was interested in the possibility of staff health screening as a way of detecting and reducing potential health risk factors, including stress, and of promoting healthier behaviors. They approached the Health Psychology Unit, who suggested an evaluation that included the impact of two methods of giving screening results on health behavior and beliefs (Michie, Johnston, Cockcroft, Ellinghouse, & Gooch, 1995). A free health screening program was promoted throughout the hospital for 10 days before under the slogan "Good Health Campaign" by posters, letters to departmental heads, and advertisements in the local press. Screening sessions were held in the Staff Recreation Club twice a day for one week. Questionnaires were completed before the physical measurements and staff were given written results of these. Those found to have raised blood pressure or very high scores on the General Health Questionnaire were given a standard letter to take to their general practitioner. All the 297 staff attenders were contacted by post six months later and invited for a repeat assessment.

Staff attending the program were compared with a stratified sample of non-attenders. Attenders were allocated randomly to getting simple informational feedback of their results or to receiving feedback plus verbal and written advice, personal target setting, and a written contract. Those receiving this more extensive feedback were found to have lost more weight and increased their exercise more than those receiving simple feedback. This project is an example of including a psychological research question within a piece of consultancy work.

8.06.6.2.3 Stress in medical school

This research group was also consulted about concern expressed by the medical school about stress in their students. The importance of medical schools recognizing the problem of stress and taking steps to reduce it was emphasized by the British Medical Association (BMA, 1992). This report concluded that medical students should be taught stress management techniques. A stress management course was designed and evaluated by a health psychologist for students in their first clinical year, which is known to be particularly stressful (Michie & Sandhu, 1994). Questionnaires assessing mood, attitudes, and the causes of stress were given to all students at the beginning of the year and one year later. The course was offered four times a year, and students were allocated to a waiting list of about three months. This allowed students to act as their own controls in order to assess whether the courses had an impact. The courses ran for three weekly sessions and students were asked to complete questionnaires before and after the course. Whilst on the waiting list, the students had completed the same questionnaire at two time points with a similar intervening interval. The uptake of this optional course (which coincided with sports afternoon) was 17%. Compared to nonattenders, attenders were more anxious, less satisfied with themselves and their life, and perceived both their work and outside functioning to be poorer. Over the year, nonattenders became significantly more dissatisfied at work and increasingly perceived their functioning at work and outside to be poorer, whilst attenders showed no such deterioration over the year. Those who attended the course showed more improvement in reported work functioning after the course than when on the waiting list. These results showed that stress management was providing long-term protective effects as well as short-term benefits. The results were reported to the medical school, and the school subsequently agreed to fund an extra half day per week to continue this work. Without data to demonstrate that the intervention had a positive impact, it is unlikely that this piece of consultancy work would have led to long-term funding for the consultant.

The BMA is concerned about the problems of stress and long hours of work at all stages of the medical career. After a junior doctor died in the UK following excessively long hours of work, a health psychologist and a medical consultant known for their work in occupational health were approached to write an editorial in the *British Medical Journal* on the question of whether overwork can kill (Michie & Cockcroft, 1996). This required a structured literature search and the resulting article attracted worldwide attention, with many invitations to participate in conferences and media events.

If those being consulted do not have the necessary data to answer the consultee's question, or the means to acquire such data, a literature review (preferably systematic) may be the best method of response. An example of this is a review commissioned by a local health authority of interventions to promote contraceptive use and sexual health among young people (McAlpine, Abraham, & Davies, 1995).

8.06.6.2.4 Twelve-hour nursing shifts

As cost pressures on the health service have increased, various ways of reorganizing work have been attempted. One such change was the change in the length of working day among nurses from 8-hour shifts to 12-hour shifts. The health psychology unit was approached to evaluate this change, as described below (Johnston & Pollard, 1991).

(i) The approach from the consultee

The issue of introducing and evaluating 12-hour nursing shifts arose in the context of increasing efficiency in health services and was under discussion by national nursing bodies in the UK. It was raised in a multidisciplinary hospital management meeting. After some discussion, the meeting agreed that the director of nursing should obtain the assistance of the psychology department in evaluating the new shift system.

(ii) Refining the question

The question was refined to be a comparison of the 12-hour shifts with the more usual 8-hour shifts. As the new shifts were only being introduced on a limited set of wards, a comparison was possible. Attendance at meetings of the nursing group implementing the shifts made it clear that the nurses were concerned about stress and the possibility of poorer performance. Therefore the question was reformulated to: Do 12-hour shifts result in greater stress and poorer performance than 8-hour shifts?

(iii) Developing the methodology

Wards that were comparable in function and structure to the 12-hour shift wards were identified; for example, surgical admission wards were matched with surgical admission wards. Stress was conceptualized as perceived stress to be measured by self-report measures of stress, tiredness, and mood. Performance was assessed in terms of speed and error rates using computer-presented cognitive tasks that were already available in the hospital. Computer assessment was used to minimize a possible confounding bias of nurses' responses reflecting what they wanted management to believe, as may be more likely in self-report measures. The design was a matched comparison of assessments every four hours and at the beginning and end of each shift. There might be an effect with duration of the shift, so that the nurses on 12-hour shifts would, after eight hours, be comparable to eight hour shift nurses, or there might be cumulative effects of working 12-hour shifts so that nurses would show effects earlier in the shift.

(iv) Implementing the methods

The proposed methods were discussed and agreed with the implementation group and they helped to introduce the study to the relevant ward staff. With the assistance of other departments in the hospital, a computer was installed in the nursing office in each of the participating wards. A medical student doing his project work in psychology played a major part in collecting data, but additional assistance was obtained from other members of the health psychology unit. Initially, a member of the psychology unit was present to ask the nurses on each ward to complete questionnaires and computer tasks at the start and end of each shift and every four hours (throughout the 24 hours). This was gradually reduced to telephone-call prompts and beginning/end of shift attendance. Data were analyzed by the medical student and a member of the health psychology unit.

(v) Reporting back

The results of the study were reported directly to the nursing implementation group. Of particular interest was the finding that self-report stress measures showed no differences between the two shift patterns and that the differences observed occurred at the beginning of shifts on the computer performance tasks. Nursing management decided not to implement 12-hour shifts, largely due to the retirement of the nurse leading this change. The results of the work were presented at conferences and published (e.g., Johnston & Pollard, 1991).

8.06.6.3 Health Promotion in Primary Care

A recent coronary heart disease (CHD) prevention project involved establishing a health psychology post within a health promotion unit to support primary care teams in their CHD prevention activities. Consultancy requests were often made as a result of the consultee attending a skill-based training course run by the consultant, seeking further support for skills development. The development of a health psychology consultancy role within a health promotion setting enabled the sharing of skills and knowledge about health psychology, and also began the process of establishing a culture in which the importance of self and peer review became acknowledged. The model of health psychology that was developed linked a training and a consultancy role, as described below.

8.06.6.3.1 Individual lifestyle interventions in primary care

(i) The approach from the consultee

The consultee was a general practice nurse who, following attendance at a training course, requested help with planning the health promotion consultations that she was offering.

(ii) Refining the question

The meeting with the consultee was used to explore the reason for her request and to help her consider the specific areas on which she would like to focus. From this discussion it became clear that her concern was about the need for a framework and skills to plan and carry out health promotion interventions in primary care. The consultee felt that a counseling approach was not always appropriate, but was unclear about how else to structure consultations. Discussion indicated the need to clarify the approach the consultee was using at present and to offer support in developing new strategies if appropriate. The initial request for help was reformulated into the question: What framework is being used at present in individual health promotion consultations?

(iii) Developing the methodology

The consultant and consultee decided to audit 30 consecutive health promotion interventions. An audit form was developed that recorded the following information: demographic information, health information, and a record of the consultation and of any health promotion materials that were used. Health information was recorded in terms of CHD risk factors (i.e., body mass index, smoking status, units of alcohol, physical activity levels, blood pressure) to provide a profile of the health status of clients who were being offered health promotion. The health promotion intervention was recorded in an open-ended way to allow for as full a description as possible of the interaction between the patient and the practice nurse. The practice nurse was simply asked to write a description of the intervention with the client. Finally, a record of the health promotion materials used allowed us to look at the appropriateness of the information being provided and to consider alternatives if necessary.

(iv) Implementing the methods

The form was discussed and agreed with the consultee. During the period of collecting the data the health psychologist met twice with the consultee to review the use of the form and discuss any issues that it raised. This process helped to clarify the approaches that the consultee was using in health promotion consultations. It allowed for a focus on individual consultations and practical examples of the difficulties that the consultee was experiencing. Once data on 30 consultations had been collected the psychologist collated the information, drawing out the main themes and issues identified in the descriptions of the consultations.

(v) Reporting back

The results were fed back to the consultee. The data helped to identify the range of skills that she already had and was using with this group. It also gave a profile of the patients with whom she was working and allowed her to explore the difference between primary and secondary prevention in health promotion work. It provided an opportunity to discuss the different types of interventions being used (e.g. counseling and advice giving), to consider others (e.g., motivational interviewing techniques) and to discuss when each approach might be appropriate or inappropriate. It also allowed for a focus on the importance of assessing the client's health beliefs in order to develop responsive health promotion interventions. Finally, the consultee was able to identify further training and development needs, which she later pursued through the health promotion unit.

8.06.6.3.2 Group-based health promotion interventions in primary care

(i) The approach from the consultee

Two health visitors contacted the health promotion unit to ask for help with setting up a smoking cessation group. The health visitors worked with the CHD health promotion officer to set up a display and develop materials and leaflets to advertise and plan the course, which was targeted at new mothers. However, they were unable to recruit enough numbers for a viable group. It was decided to use the time that had been set aside for the group to meet with the health visitors and discuss how the identified client group could be worked within one-to-one consultations, as the client group did not seem to be accessible through group work. An initial consultancy session was offered, and four health visitors said they would like to attend.

(ii) Refining the questions

The session was facilitated by the health psychologist and the health promotion officer and focused on considering why it had been so difficult to recruit to the group. It emerged from this discussion that one of the difficulties had been the health visitors' own anxieties about raising the issue of smoking with clients with whom they were already involved. Fears about jeopardizing their working relationship and not knowing how best to proceed had meant that, while much energy had been put into planning and advertising the group, the health visitors had not mentioned it to any of their clients. The question that the consultation needed to focus on was refined as: What skills can I use to raise the issue of smoking with my clients?

(iii) Developing the methodology

After this initial consultation the consultees and consultants agreed to meet for four further consultation sessions, which would focus on frameworks and skills for working with smokers. The health psychologist and health promotion officer planned these sessions to include using the Stages of Change model (Prochaska & DiClemente, 1982) in order to provide a framework with which to structure health promotion interventions, explore and practise some communication skills (e.g., open and closed questions), and choose and use health promotion materials appropriately.

(iv) Implementing the methods

The four sessions went ahead as planned. The health visitors were asked to do work between each session. This included using a case study from their own work to apply the framework presented, and reviewing a health promotion resource for its usefulness and appropriateness for their client group. The feedback from the health visitors on the sessions was very positive, and the sessions were slightly refocused to concentrate on basic communication skills and assessing clients' health beliefs.

(v) Reporting back

The final session was used to review what had been covered and discuss the training and development needs identified. Preliminary evaluation of this consultation indicated that this was a useful and valued way of working. In particular, consultees identified the benefits of being able to focus on skills in a small-group setting and of sharing their experiences of smoking cessation work in a structured way. The importance of offering this approach in a systematic way to other groups of health visitors was identified.

8.06.6.3.3 A consultancy framework for developing health promotion intervention programs

The examples given above demonstrate the similarities in requests for consultation. The three main issues focused on in health psychology consultancy were:

(i) offering support for, and an opportunity to review, skills development;

(ii) using health psychology models to provide frameworks to help consultees review their approaches to, and planning of, health promotion work; and

(iii) developing intervention programs.

It was decided to use this framework to develop a consultancy service that would allow the health promotion unit to develop health promotion programs in conjunction with primary care teams. An example of this work that led to the development of a weight management program is described below.

(i) The Approach from the consultee

The health promotion officer for exercise and nutrition was approached by two practice nurses who wanted to set up a group program in their practice for patients with weight problems. This officer was interested to work with the CHD health promotion officer to develop a program. The health promotion officer asked the health psychologist for support and help with this work and for advice on including a psychological perspective.

(ii) Refining the question

Initial meetings were held with the two health promotion officers to agree on the focus of the work. Their aim was to develop a joint working relationship with the practice nurses to allow for skill sharing in the development of a group program. The health promotion officers could offer health promotion skills, knowledge of the topic and its health implications, and knowledge of health promotion materials for running group interventions. The practice nurses had experience of working with this client group and facilitating groups. After discussions it was agreed that consultancy sessions would be used to review the development of this working relationship. The question was refined to become: How can we establish a working framework with primary care to develop an intervention program? A number of features of this relationship were identified that were to be reviewed in each consultancy session: the timing and setting of meetings between both groups; the issues raised, including any difficulties and how, or if, they were resolved; opportunities to discuss skills and approaches and to share concerns with regard to the group itself; and the use of health promotion materials. The final issue that was considered in consultancy sessions was how to capture the learning from this method of joint working.

(iii) Developing and implementing the methodology

A series of joint sessions with the health promotion officers was planned in which the above topics would be reviewed. These sessions took place on a weekly basis subsequent to the health promotion officers' meetings with the practice nurses. The sessions included feedback from the health promotion officers on the progress of the group. The health psychologist facilitated a review, using an acceptant consultancy style, of the issues that were raised, and the above points were considered each time.

(iv) Reporting back

The health psychologist kept a record of the consultancy sessions, and the final sessions were used to summarize the issues that had been raised, in order to identify further training and development needs and to consider how to capture the learning from this process. It was decided to write up the work in a way that would capture both the content and the process of learning for both the practice nurses and the health promotion officers. This led to the development of a weight management program pack, which was later piloted in another practice.

8.06.7 THE FUTURE: AVOIDING GAPS AND TRAPS

One general problem of consultancy is that health professionals may feel threatened by asking for psychological consultation (Belar, 1991). The request for help may reflect a breakdown in communication or relations between people, and may be a challenge to self-esteem, especially in health care settings where people are expected to have expertise in these areas. Empathy for the perspective of the individuals and the staff groups involved and tact to handle these situations are essential. Whilst these subjective factors may put limits on the consultancy role, the objective potential for the future is favorable.

In considering the future role of psychology in health care settings, Siegel (1995) wrote:

> An emphasis on cost containment, quality assurance, accountability, and documentation of health care programs' effectiveness will necessitate the use of psychologists' research and program evaluation skills. These developments will provide psychologists with the opportunity to have a significant role in shaping health-related programs for prevention and in the delivery of services. (p. 342)

and, as was confidently stated in 1987,

> there are few limits to the issues a health psychologist with vision, dedication, and persistence can address [but] the responsibility for forging new pathways lies on the shoulders of the entire health psychology community. (Altman & Cahn, 1987, pp. 240–241)

However, as psychologists know only too well, there is often a large gap between intention and behavior. We need to use the strategies identified by psychologists to bridge this gap, in the development of our own future, and we need to avoid the trap of the individualism that has sometimes been associated with psychologists. Instead, we should plan, organize, and act as a collective.

ACKNOWLEDGMENTS

Susan Michie is supported by the Wellcome Trust. The author thanks Louise Earll and Barbara Wren for comments on an earlier draft of this, and Barbara Wren for contributing the section on her consultancy work in primary care.

8.06.8 REFERENCES

Altman, D. G., & Cahn, J. (1983). The rest of the challenge: Position statement on employment opportunities. *Health Psychology, 2*(5 suppl): 119–122.

Altman, D. G., & Cahn, J. (1987). Employment options for health psychologists. In G. C. Stone, S. M. Weiss, J. D. Matarazzo, N. E. Miller, J. Rodin, C. D. Belar, M. J. Follick & J. E. Singer (Eds.), *Health psychology: A discipline and a profession* (pp. 231–244). Chicago: University of Chicago Press.

Belar, C. (1991). Professionalism in medical settings. In J. J. Sweet, R. H. Rozensky, & S. M. Tovian (Eds.), *Handbook of clinical psychology in medical settings* (pp. 81-92). New York: Plenum.

Belar, C. (1995). In B. Johnston, R. G. Frank, C. Belar, S. Berk, L. A. Bielianskas, E. D. Bigler, B. Caplan, T. R. Elliot, R. L. Glueckanf, R. M. Kaplan, J. S. Kreutzer, C. A. Mateer, D. Patterson, A. E. Puerle, J. S. Richards, M. Rosenthal, M. Sherer, R. Shewchuk, L. J. Siegel, & J. J. Sweet (Eds.), Psychology in health care: Future directions. *Professional Psychology: Research and Practice, 26*(4), 341–365.

Bor, R., & Miller, R. (1991). *Internal consultation in health care settings*. London: Karnac.

Boyce, R. (1991). Hospital restructuring—the implications for allied health professions. *Australian Health Review, 14,* 147–153.

British Medical Association (1992). *Stress and the medical profession.* London: Author.

Brunning, H., Cole, C., & Huffington, C. (1990). *The change directory: Key issues in organisational development and the management of change.* Leicester, UK: British Psychological Society.

Campbell, D., Draper, R., & Huffington, C. (1989). *A systemic approach to consultation.* London: DC Publishing.

Chalmers, I., & Altman, D. G. (1995). *Systematic reviews.* London: BMJ Publishing Group.

Eardly, A. (1988). Patient worries about radiotherapy. Evaluation of a preparatory booklet. *Psychology and Health, 2,* 79–89.

Edwards, J. D., & Holmgren, R. L. (1979). Some prerequisites for becoming a "really" applied, non-academic, social psychologist. *Personality and Social Psychology Bulletin, 5, 516–523.*

Elliott, T. R., & Shewchuk, R. (1995). In B. Johnston, R. G. Frank, C. Belar, S. Berk, L. A. Bielianskas, E. D. Bigler, B. Caplan, T. R. Elliot, R. L. Glueckanf, R. M. Kaplan, J. S. Kreutzer, C. A. Mateer, D. Patterson, A. E. Puerle, J. S. Richards, M. Rosenthal, M. Sherer, R. Shewchuk, L. J. Siegel, & J. J. Sweet (Eds.), Psychology in health care: Future directions. *Professional Psychology: Research and Practice, 26*(4), 341–365.

Everly, G. S. (1986). An introduction to occupational health psychology. In P. A. Keller & L. G. Ritt, (Eds.), *Innovations in clinical practice: A source book* (Vol. 5). Sarasota, FL: Professional Resource Exchange Inc.

Huffington, C., & Brunning, H. (1994). *Internal consultancy in the public sector: Case studies.* London: Karnak.

Hutchcroft, S., Snodgrass, T., Troyan, S., & Wares, C. (1984). Testing the effectiveness of an information booklet for cancer patients. *Journal of Psychology and Oncology, 2*(2), 73–83.

Jansen, M. A., Methorst, G. J., & Kerkhof, A. J. F. M. (1991). Health psychology in international perspective: Summary and thought for the future. In M. A. Jansen & J. Weinman (Eds.), *The international development of health psychology* (pp. 165–171). Reading, UK: Harwood.

Jansen, M. A., & Weinman, J. (Eds.) (1991). *Introduction. The international development of health psychology* (pp. vii–viii). Readiing, UK: Harwood.

Johnston, M. (1994). Current trends. *The Psychologist, 7*(3), 114–118.

Johnston, M., Gilbert, P., Partridge, C., & Collins, J. (1992). Changed perceived control in patients with physical disabilities: An intervention study with patients receiving rehabilitation. *British Journal of Clinical Psychology, 31*(1), 89–94.

Johnston, M., & Pollard, B. (1991). Length of nurses' working shift: Stress and information processing. In G. Biondi (Ed.), *Psychology in hospital: Models of intervention* (pp. 101–105). Rome: NES.

Johnston, M., & Weinman, J. (1995). Health psychology. In *Professional psychology handbook* (pp. 61–68). Leicester, UK: BPS Books.

Johnston, M., Weinman, J., & Wright, S. (1995). *Measures in health psychology: A user's portfolio.* Windsor, UK: NFER–Nelson.

Kaplan, R. M. (1994). The Ziggy theorem: toward an outcomes-focused health psychology. *Health Psychology, 13*(6), 451–460.

Madden, S., Johnston, M., & Parbhoo, S. (1994). Evaluation of women's worries and the effects of a preparatory booklet for patients attending a breast clinic. *The Breast, 3,* 169–172.

Marteau, T. M., Kidd, J., Cuddeford, L., & Walker, P. (1996). Reducing anxiety in women referred for colposcopy using an information booklet. *British Journal of Health Psychology, 1,* 181–189.

Marteau, T. M., Sowden, A. J., & Armstrong, D. (submitted). Implementing research findings into practice: Beyond the information deficit model.

McAlpine, S., Abraham, C., & Davies, H. (1995). *I. Promoting contraceptive use amonst young people: a brief review.* II. *Promoting sexual health interventions for young people: a brief revew.* Reports for Tayside Health Board, Dundee, UK.

Michie, S. (1992a). Evaluation of a staff stress management service. *Health Manpower Management, 18,* 15–17.

Michie, S. (1992b). Reducing stress in NHS staff. *Employee Counselling Today, 4,* 4–7.

Michie, S. (1993). The work of psychologists in the changing organisation of the NHS. *Clinical Psychology Forum, 52,* 27–29.

Michie, S. (1996). Reducing absenteeism by stress management: Valuation of a stress counselling service. *Work and Stress, 10*(4), 367–372.

Michie, S., & Cockcroft, A. (1996). Overwork can kill. *British Medical Journal, 312,* 921–922.

Michie, S., Johnston, M., Cockcroft, A., Ellinghouse, C., & Gooch, C. (1995). Methods and impact of health screening for hospital staff. *Journal of Organisational Behaviour, 16,* 85–92.

Michie, S., & Ridout, K. (1990). Stress management for nurses. *Clinical Psychology Forum, 26,* 16–19.

Michie, S., Ridout, K., & Johnston, M. (1996). Stress in nursing and patients' satisfaction with health care. *British Journal of Nursing, 5*(16), 1002–1006.

Michie, S., Rosebert, C., Heaversedge, J., Madden, S., & Parbhoo, S. (1996). The effects of different kinds of information on women attending an out-patient breast clinic. *Psychology, Health and Medicine, 1,* 285–296.

Michie, S., & Sandhu, S. (1994). Evaluating stress management in medical students. *Medical Education, 28, 528–533.*

Ovretveit, J. (1992). *Therapy services: Organisation, management and autonomy.* Reading, UK: Harwood.

Ovretveit, J., Brunning, H., & Huffington, C. (1992). Adapt or decay: Why clinical psychologists must develop the consulting role. *Clinical Psychology Forum, 46,* 27–29.

Partridge, C. J., & Johnston, M. (1989). Perceived control of recovery from physical disability: Measurement and prediction. *British Journal of Clinical Psychology, 28*(1), 53–59.

Pion, G. M. (1991). A rational human resources agenda for psychology: The need for a broader perspective. *Professional Psychology: Research and Practice, 22,* 449–455.

Prochaska, J. O., & DiClemente, C. C. D. (1982). Transtheoretical therapy: Toward a more integrative model of change. *Psychotherapy: Theory, Research and Practice, 19*(3), 276–88.

Siegel, L. J. (1995). In: B. Johnston, R. G. Frank, C. Belar, S. Berk, L. A. Bielianskas, E. D. Bigler, B. Caplan, T. R. Elliot, R. L. Glueckanf, R. M. Kaplan, J. S. Kreutzer, C. A. Mateer, D. Patterson, A. E. Puerle, J. S. Richards, M. Rosenthal, M. Sherer, R. Shewchuk, L. J. Siegel, & J. J. Sweet (Eds.). Psychology in health care: Future directions. *Professional Psychology: Research and Practice, 26*(4), 341–365.

Smith, J. A., Harre, R., & Langenhove, L. V. (1995). *Rethinking methods in psychology.* London: Sage.

Stone, G. C. (1979). Health and the health system: a historical overview and conceptual framework. In G. C. Stone, F. Cohen, & N. E. Adler (Eds.), *Health psychology: A handbook: theories, applications, and challenges of a psychological approach to the health care system.* San Francisco: Jossey-Bass.

Stone, G. C. (1982). *Health Psychology,* a new journal for a new field. *Health Psychology, 1*, 1–6.

Stone, G. C., Cohen, F., & Adler, N. E. (1979). *Health psychology: A handbook: theories, applications, and challenges of a psychological approach to the health care system*. San Francisco: Jossey-Bass.

Wallace, L. M. (1984). Psychological preparation as a method of reducing the stress of surgery. *Journal of Human Stress, 10*(2), 62–77.

Weinman, J. (1991). Determinants of health psychology's development. In M. A. Jansen & J. Weinman (Eds.), *The international development of health psychology* (pp. 159–164). Reading, UK: Harwood.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.07
Teaching and Training Other Health Disciplines

HELEN R. WINEFIELD
University of Adelaide, SA, Australia

8.07.1 INTRODUCTION 172

8.07.2 KEY HEALTH PROFESSIONS 172
- *8.07.2.1 Nursing* 173
- *8.07.2.2 Medicine* 174
- *8.07.2.3 Medical Social Work* 174
- *8.07.2.4 Physiotherapy/Physical Therapy* 174
- *8.07.2.5 Occupational Therapy* 174
- *8.07.2.6 Pharmacy* 174
- *8.07.2.7 Dentistry* 175
- *8.07.2.8 Speech Pathology/Therapy* 175
- *8.07.2.9 Other Health-related Occupations* 175

8.07.3 MAJOR GOALS OF HEALTH PSYCHOLOGY TEACHING 175
- *8.07.3.1 Improved Healthcare* 175
  - *8.07.3.1.1 Measurement of quality of care* 175
- *8.07.3.2 Health and Well-being of the Health Workers* 176
- *8.07.3.3 Knowledge and Skills from Health Psychology Input* 177
  - *8.07.3.3.1 Communication skills* 177
  - *8.07.3.3.2 Understanding of and respect for scientific enquiry* 178
  - *8.07.3.3.3 Appropriate patterns of practice* 179

8.07.4 FORMAT AND METHODS OF TEACHING 179
- *8.07.4.1 Lectures or Practice?* 179
  - *8.07.4.1.1 Lectures* 180
  - *8.07.4.1.2 Structured observations of practice* 180
  - *8.07.4.1.3 Small (tutorial) group learning* 180
  - *8.07.4.1.4 Learning via the Internet* 180
- *8.07.4.2 Integrated, Problem-based Learning* 180
- *8.07.4.3 Assessment* 181
- *8.07.4.4 Interdisciplinary Teaching* 181
- *8.07.4.5 Applying the Principles of Permanent Behavior Change* 181

8.07.5 CONTENT 182
- *8.07.5.1 Topics About Health and Illness* 183
- *8.07.5.2 Topics About the Interaction Between Health Professional and Patient* 183

8.07.6 IMPLICATIONS AND CONCLUSIONS 183
- *8.07.6.1 Needed Attributes of Health Psychologists who Teach Other Professions* 183
- *8.07.6.2 Needs of Health Psychologists who Teach Other Professions* 184
- *8.07.6.3 Consequences for Health Psychology of Teaching Other Professions* 184

8.07.7 REFERENCES 185

8.07.1 INTRODUCTION

In addition to the assessment, treatment, and prevention of health problems as described in other chapters of this volume, health psychologists have a valuable role to play in teaching and training members of other health-related disciplines. This chapter will outline the possibilities, some of which, such as communication skills teaching, are being applied widely in health education internationally, while others such as encouraging surveillance of the occupational health of health workers, remain underdeveloped.

This chapter begins by introducing the key health professions other than psychology, and briefly reviewing the psychological content of their curricula (Section 8.07.2). Section 8.07.3 then describes the desirable outcomes of health psychology within training for health professionals, in terms of three broad goals: improving the health of the community, improving the health and well-being of health workers, and imparting relevant knowledge and skills to health workers.

The chapter then reviews the format and methods which seem to be most effective in such teaching (Section 8.07.4). Traditional lecture and tutorial methods need extension for professional training, and health psychology teaching is characterized by multidisciplinary collaborations. Section 8.07.5 lists the health psychology topics likely to be widely useful to health professionals, and therefore recommended for inclusion in their educational preparation. Finally, Section 8.07.6 raises some wider issues about what qualities are needed in the health psychologists who do such teaching, what their own needs are, and what the consequences of sharing our knowledge may be for health psychology.

8.07.2 KEY HEALTH PROFESSIONS

This chapter will concentrate on the health disciplines which receive preregistration training in universities, including institutions of higher education which were formerly known as colleges of advanced education, polytechnics, or institutes of technology. Other disciplines which promise health care but for which training generally occurs outside universities include chiropractice, naturopathy, hypnotherapy, aromatherapy, and so on. While these "alternative" treatments sometimes win acceptance from doctors and nurses who recommend them to their patients, little information is publicly available about their training curricula or assessment methods. However, at least some of the material in this chapter seems likely to be applicable to them too.

Many health professions provide continuing education for their members after entry to the workforce, in the form of workshops, seminars, conferences, and journals, either on a voluntary basis or in a formal way overseen by registration or professional bodies. With recognition of the ever-faster expansion of knowledge, and acknowledgement that undergraduate instruction cannot suffice for a life-long practicing career, we may expect that some form of regulated continuing professional education will become more rather than less common. Thus postregistration health psychology teaching and training seems likely to increase in extent in the relatively near future. Everything said below about mutual collaboration between health psychologists and members of the profession concerned, applies with even greater force to postregistration teaching.

A great deal of variability exists in the methods and content of teaching health psychology to different sets of students. When we multiply the variations within different professions, undergraduate vs. postgraduate courses, teaching institutions, cities and countries, it is clearly unrealistic for a chapter such as this to give a detailed account of hours spent, methods used and course content in every case. Rather the goal here will be to convey a flavor of what aspect of health psychology are taught in the nonpsychology disciplines, largely based on a review of published textbooks. Learning via the Internet, which has been predicted to change academic methods radically, will be discussed later (see Section 8.07.4.1.4).

Published teaching materials are probably only a small proportion of those in use. The existence of a textbook indicates that somebody, perhaps an individual, probably a curriculum committee but certainly a publisher, perceived a need for that educational resource. However every teacher is well aware that published materials may not accurately represent the examinable content of the course, and even less do they reflect the methods used to teach it. Another major problem for the comprehensiveness of a review such as this is the fact that we have limited understanding of the relationship between the teaching provided, and what students learn. Doubtless the teachers' credibility, skills, knowledge, enthusiasm, and similar factors mediate how the teaching is received and therefore what impact it may have on the later professional practice of graduates. Apart from some comments on when and by whom teaching should occur (in Section 8.07.4), such issues are beyond the scope of this chapter, but do need systematic attention and research by educators.

The main health professions described below are not in order of status within the health care industry, nor in order of the competition for places and thus the prior academic qualifications of the students. Current numbers of Australian students have been chosen as an ordering principle. The health education system in Australia is very similar to the British, and Australian health professionals are found all over the world. While the number of students may not be proportional to the number of practitioners, it is reasonable to use student numbers in a chapter on training because the great bulk of formal training occurs before professional registration and practice.

Table 1 shows the numbers of health and social work students who were enrolled in higher education in Australia in 1995. An average of 6.5% of students across health disciplines came from overseas and were paying full fees, with the highest proportions being in pharmacy (14%), medicine (12%), and dentistry (11%). Overseas students mostly speak English as a second language, and language-rich subjects such as psychology need to provide extra support and language tuition to help them succeed. The most effective support includes discussion of cultural assumptions as well as language help (Ballard & Clanchy, 1991; Chur-Hansen & Barrett, 1996). Asian and European cultures have different norms and expectations about health-relevant behavior, challenging what has sometimes been an unself-conscious Anglocentrism in health psychology. In health and health care, cultural differences in beliefs and practices need to be understood by teachers and trainers as well as by practitioners. To a less marked extent the same point applies to students of different social class backgrounds within the same language group.

Table 1 Student enrollments in health by field of study, higher education, Australia, 1995.

Field of study	*1995 enrollments*
Basic nursing	23 228
Postbasic nursing	10 209
Medicine	10 471
Medical science	1973
Medical radiography	1549
Medical technology	1153
Dentistry	1408
Dental therapy	37
Nutrition and dietetics	671
Optometry	786
Pharmacy	2078
Podiatry	352
Rehabilitation	584
Other rehabilitation	512
Occupational therapy	2378
Physiotherapy	3037
Speech pathology/audiology	1235
Health support	1841
Health administration	2276
Health counseling	759
Health surveying	760
Health science and technology	1060
Health sciences — other	1280
Health support — other	660
Health — general	1840
Total health	72 137
Social work	6050

Source: Department of Employment, Education, Training & Youth Affairs, Canberra.

8.07.2.1 Nursing

In terms of student and practitioner numbers, the discipline of nursing is the dominant healthcare profession: there were three times as many nursing students as medical students (see Table 1).

Nursing education has established itself in Australian universities since the early 1980s in contrast with the previous apprentice-style, bedside teaching methods (McCue & White, 1983). A recent review of nurse education concluded:

> The need now is to move beyond the era of the transfer [of nursing to higher education] . . . [to] the development of professional practice, the growth of nursing as an academic discipline, and the future contribution of education and research to the quality of nursing care. (Reid, 1994, p. 341)

This review identified as core curriculum:

> basic strands in the areas of medical/surgical nursing, community health nursing, and mental health nursing, for individuals across the lifespan in institutional and noninstitutional settings [including] communication, primary health, behavioral sciences, biological sciences and broadbased, general health care subjects. (Reid, 1994, pp. 170–172)

Doctor–nurse relationships have been described by sociologists and anthropologists (e.g., Campbell-Heider & Pollock, 1987) but there is now an interest among nurses themselves in understanding the political dimensions of health care and in improving community health through active advocacy (Gray & Pratt, 1992). The establishment of doctoral training programs in nursing which include coursework and fieldwork in addition to research training is intended to assist nurses who seek careers in consultancy, management, and policy making (Pearson, 1996). In the USA there is a very wide choice of masters and doctoral programs for nurses.

At the undergraduate level there is longstanding and continuing interest in the psychological aspects of nursing (Byrne & Byrne, 1992; Hall, 1982; McGhie, 1973, Paton & Brown, 1991; Payne & Walker, 1996). One of the few texts to identify itself specifically as health psychology for nurses (Niven, 1994) includes in its preface an explicit articulation of a crucial issue in health psychology training for all nonpsychologists—namely what is the most desirable framework to adopt? Whereas early texts in nursing as in other health disciplines tended to adopt the framework of what aspects of psychology might be relevant to health care, Niven adopted instead the framework of a psychology of health. Taking health rather than psychology as the primary focus need not in fact result in vastly different contents from those of the more traditional approach (although they will be differently arranged and presented), but seems likely to have the important pedagogic advantage of being more attractive and credible to students (see Section 8.07.4).

8.07.2.2 Medicine

Medical courses have included psychological content, often under titles such as "behavioral science," "medical psychology," and "behavioral medicine," for many years now (DiMatteo, Friedman, & Robin 1982; Hetherington, Miller, & Neville, 1964; Pomerleau & Brady, 1979; Rachman, 1977; Stoudemire, 1994; Wedding, 1995; Weinman, 1987; Winefield & Peay, 1991). Health psychologists have seized the opportunity to teach medical students and thereby develop relationships with medical practitioners and engage in collaborative research (Winefield, 1991). While the technological advances of the post-World War II period have fostered a strong emphasis and high prestige within medicine for pharmaceutical, surgical, and diagnostic techniques—in general the "bio" element within the "biopsychosocial" model of health care (Engel, 1977)—health psychology with its focus on behaviors, beliefs, and feelings in relation to health and illness, adds the balance of a multivariate and holistic perspective on health care and health promotion.

Just as clinical psychology has experienced border disputes with the medical specialty psychiatry, health psychology seems most likely to compete with consultation–liaison psychiatry, the clinical derivative of psychosomatic medicine with its interests in cardiovascular disease, oncology, and chronic pain (Wise, 1986). Some clinical health psychologists spend their time in the delivery of services to individual patients, but others such as occupational health psychologists work in ways more congruent with public health (Quick, 1996).

8.07.2.3 Medical Social Work

The 1991 Australian census indicated that 35% of social workers will work in health settings (Martin, 1996), thus justifying the inclusion of this profession here. This profession shows large national differences in its organization; in the USA, for example, social workers much more often practice autonomously than in Britain and Australia.

Teaching materials tend to have a strong practice orientation (Danbury, 1994; James & Vinson, 1989; Shulman, 1993), although Thompson (1995) has argued in favor of integrated theory and practice. Developmental and social psychology seem to have been most valued, but social workers in medical settings may learn to welcome health psychology input.

8.07.2.4 Physiotherapy/Physical Therapy

This discipline, like others traditionally seen as "adjunct," "allied," or even "para" in relation to medical services, was once taught using an apprenticeship model within the work (usually hospital) setting. Physiotherapists use exercise, manipulation, and physical interventions such as heat and ultrasound to treat problems of pain, restricted movement, and the consequences of injury or aging. An early textbook by Dunkin (1981) focused on learning, individual differences, and interpersonal relations, and encouraged students to search the literature to follow up their interests.

8.07.2.5 Occupational Therapy

The practice of this discipline seeks to foster the health benefits of occupation, in the sense of purposeful activity, especially for persons with psychiatric, intellectual or physical disabilities. While virtually theory free in the past, the profession recognizes a need to identify and evaluate its skills and to prepare practitioners for work in a variety of settings (Fransella, 1982; Kielhofner, 1992).

8.07.2.6 Pharmacy

The Report of the Study Commission on Pharmacy (1975) found an imbalance in the curriculum of the day, with too much attention being given to drug products and their effects and too little to "human behavior, cultural determinants, health service systems and their economics. ... Needed and optimally effective drug therapy results only when both drugs and those who consume them are fully understood" (pp. 126–127). Wertheimer and Smith (1989) aimed to redress this imbalance, including readings on topics such as the patient's view of the illness, compliance, choices of prescribed and nonprescribed drugs, tranquillizers and social control, and ethics and informed consent. Public attitudes to medicines, which seem to be increasingly ambivalent (Vuckovic & Nichter, 1997), need to be understood by many health professionals in addition to pharmacists.

8.07.2.7 Dentistry

Textbooks on the psychological aspects of dental practice have been available for some time (e.g., Ayer & Hirschman, 1972; Cinotti, Grieder, & Springob, 1972). Obvious topics of interest include dental anxieties and phobias, bruxism, oro-facial pain and compliance with self-care or orthodontic instructions.

8.07.2.8 Speech Pathology/Therapy

Disorders of oral communication whether due to childhood lack of skill development or to loss of skills after stroke or head injury occupy speech pathologists, who treat them using graded exercises. Purser (1982) was an early contributor to their understanding of developmental and social psychology and behavior change.

8.07.2.9 Other Health-related Occupations

Various forms of medical scientist, and support professions such as radiography, dietetics, podiatry (chiropody), and health administration are also taught in universities. It appears that their curricula are task-focused and in some cases their direct service roles limited; where health psychological input is required it may well be taught using materials originally prepared for other disciplines, or specific to the course and not commercially available.

8.07.3 MAJOR GOALS OF HEALTH PSYCHOLOGY TEACHING

The health psychology input for health profession students does not need to be uniform in format or content across courses. Indeed health psychology courses need to be geared sensitively to the practice modes, traditions, and expectations of practitioners of each discipline. For example other professions vary in whether they see their discipline as a science or a craft, how independently they expect to practice, and where practice usually occurs, as well as in the academic achievement levels and backgrounds of incoming students. In this section of the chapter, however we will canvass some broad goals which could apply to all health psychology input. Section 8.07.6 will address the controversial issue of what the consequences for health psychology might be, of "giving psychology away" to other disciplines.

8.07.3.1 Improved Healthcare

Most teachers, trainers, and curriculum committees would probably agree that the main purpose of including health psychology is to improve the quality of healthcare which students are later able to deliver. In Section 8.07.3.3, I shall try to distinguish some superordinate skills of the well-taught and well-trained health professional, to which health psychology has contributed.

A secondary goal of course organizers might be for health psychology itself to gain recognition and respect amongst other health professions, and for health psychologists to develop useful professional networks and access to research opportunities by being visible and successful within the training institutions.

A large practical and conceptual obstacle to devising a successful health psychology course, in terms of knowing that it is contributing to improved healthcare, has been the lack of markers or indicators for quality of health care practice. When no data are available about how successful various practice interventions are, either in improving patients' quality of life or in reducing their symptoms and death rate, the role of health psychology knowledge and skills (as with every other component of the students' learning experiences) becomes difficult to judge. Presently there is increasing awareness of the need for health professions to develop appropriate outcome measures, to validate them, and to begin the process of exploring their antecedents (Kaplan, 1994a).

8.07.3.1.1 Measurement of quality of care

Life expectancy has improved in developed countries following the great public health reforms of clean water, sanitation, public education about hygiene, nutrition, and prevention; and accessibility of antenatal care, child health screening, and emergency services. However, mortality rates are insufficiently sensitive as outcome measures of many health care interventions, particularly given the chronic nature of much illness in developed countries. Ongoing measures of health are needed which will reflect changes over time in both physical and psychological dimensions. Health-related quality of life (QOL), a concept which embraces affective and social well-being, functional capacities, and lack of symptoms, is an outcome measure well-suited to the evaluation of the effectiveness and cost-effectiveness of health care (Kaplan, 1994a, 1994b; Ware, 1995). An example of the latter analysis is that some preventive and educational interventions have much greater payoff in terms of "cost per well year," than more specialized surgical treatments. Specifically, costs per well year in 1991 US dollars of seat belt laws = 0, of smoking cessation counseling = 6463, and of coronary artery bypass graft for heart disease = 662 835 (Kaplan, 1994a). The development and standardization of QOL measures is widely agreed to be a top priority project for health psychology (Sarafino, 1994; Weinman, 1990).

Patients are often forced to rely at least in the short term, on interpersonal features of the care-delivery interaction to assess its value. Some of these "process" features have been shown to correlate with beneficial health outcomes although the relationship is not straightforward (Kaplan, Greenfield, & Ware, 1989; Stewart, 1995).

"The patient is now ... to some degree an adjudicator of service quality ... [which] means that the doctor's perspective on what is valuable in this interaction ... is marginalized" (Fairhurst & May, 1995, pp. 389–390). Patient satisfaction with care is a proximal measure of quality of care but cannot be accepted as adequate by itself. Apart from the likely response-biases when seeking feedback from people who may need help again soon, there are problems in accepting patients' definitions of what they need. Some patients may have unrealistic expectations of what can be achieved, or of how much time the health worker can devote to their case. Others may feel most secure only after numbers of extra diagnostic tests, investigations, and treatments which are regarded as unnecessary and excessive by their health carers, and/or by the insurers and administrators who oversee health care costs (Aharony & Strasser, 1993; Pickering, 1993; Williams, 1994). Another difficulty is that patients' satisfaction with an episode of healthcare may change over time, and not be validly captured by a questionnaire after a single consultation (Avis, Bond, & Arthur, 1997).

Nonetheless, there are compelling arguments for seeking evaluative information from users of health care services. In industrialized countries the individual consultation between a patient and a general practitioner, for example, only rarely results in dramatic improvements in the health of the person concerned (Winefield, Murrell, & Clifford, 1995). Often, patient satisfaction with the experience of being cared for may be the only measurable outcome of a health care interaction, but this sense of being cared for and valued by credible experts can have powerful nonspecific benefits (Shepherd & Sartorius, 1989). An example is the finding by Fitzpatrick, Hopkins, and Harvard-Watts (1983) that headache sufferers were more likely to report fewer and less severe headaches one year after referral to a neurology clinic, regardless of what, if any, medication was prescribed, if they had felt satisfied immediately after the consultation. The popularity of alternative healthcare practitioners with the public has been attributed to their willingness to spend time, show interest, and develop a caring personal relationship with the "client"; these factors plus the shared rationale for illness and the confident prescription of healing activities are also common to nonmedical healers in other cultures (Frank, 1989).

8.07.3.2 Health and Well-being of the Health Workers

A secondary goal of including health psychology in courses for other professions might be seen as preparing them not only to deliver health care to patients/clients, but to maintain their own health and well-being while doing so. The high turnover in some health professions, particulary nursing where there are many back injuries, and the reports of high job stress, career dissatisfaction, burnout, and associated problems of alcoholism, drug abuse, and marital strains (Allen, 1994; Howell & Schroeder, 1984), provoke consideration of the need for some preventive educational preparation.

Students often enter health courses when quite young (many are teenagers direct from secondary school, in Australia and Britain), and with relatively little life experience. The higher the academic qualifications for entry, it seems the less opportunity students have had to mix with people outside their own peer group, people of different ages, social class and cultural background, and to explore their own talents and interests beyond study. It is not surprising therefore that the demands of tertiary study—often combined with financial pressures and personal development issues around independence, sexuality, and vocation—in combination with the exposure to illness, suffering, and death, can be stressful for students (Ashton & Kamali, 1995; Firth-Cozens, 1987; Simpson & Budd, 1996).

The role models available to students in clinical settings may or may not be positive examples of healthy coping. There is a delicate balance between maintaining a degree of professional detachment which facilitates action but still expresses care and respect for the patient, and an emotional distance which patients and professionals alike find unsatisfying and depersonalizing. Getting this balance right is possibly the hardest task in professional training, or certainly the most psychologically challenging aspect, and clinical health psychologists may have valuable roles as both teachers and therapists.

After graduation, other work-related pressures are added. Autonomy, a traditional reason for job satisfaction in medicine for example, is perceived as being eroded through the intrusion of third parties such as insurers, administrators, and managers (Allen, 1994; Burke, 1996; Siegler, 1993). An awareness of the psychological hazards of healthcare is the first requirement for prevention, so that affected individuals do not castigate themselves as weak and inadequate. This awareness can be raised during undergraduate years and may subsequently need reinforcement with stress management interventions (Bunce & West, 1996; Midence, Gregory, & Stanley, 1996; Quick Murphy, & Hurrell, 1992).

8.07.3.3 Knowledge and Skills from Health Psychology Input

In this section I shall review some general categories of health psychology knowledge and skills which, in the context of the restrictions described above on how to recognize quality care, could be expected to equip students for their future work in health professions.

First it is relevant to draw the distinction between "knowledge" which is acquired through teaching, and "skills" which are acquired through training. While the distinction cannot be rigidly applied, it usefully indicates different types of learning and the activities meant to promote them. Traditionally, health students have acquired knowledge in an abstract way through lectures and readings, assessed through exams, then proceeded to acquire clinical skills though supervised "hands on" practice, assessed to competency standards by experienced practitioners. The knowledge acts as a foundation for the skills but also goes further to give students a general appreciation of the theoretical and conceptual roots of their discipline and its methods of accumulating further knowledge. It is also meant to equip students to read critically, to continue their education life-long, and possibly to contribute to the knowledge base of the discipline themselves by engaging in research.

Skills have the more immediate practical value of equipping students to deal effectively with their work tasks. As well as being capable of giving injections, bandaging, and a host of other physical procedures, health professionals have to develop cognitive skills of clinical decision-making and time management. They must learn to respond calmly and automatically in distressing or life-threatening situations and to cope with behaviors, injuries, or disfigurements which untrained members of the public would find overwhelming. Skills can also become obsolete, and re-training or revision of them is then needed.

While each health profession has a more or less specific knowledge base and distinctive skills, health psychology can contribute important general elements which are identified below. Each of them could usefully become a strand of instruction, taking different forms with progression through the course (see Section 8.07.4), but with continuity throughout it. More specific topics are listed in Section 8.07.5.

8.07.3.3.1 Communication skills

Every health practitioner needs to develop effective communication skills. These include being able to access information from the patient which is relevant to diagnosis and treatment, to explain matters using comprehensible terms (including written materials) and to check patient understanding, to encourage patient questions, to listen attentively, to communicate respect and care, and to persuade service-recipients to follow recommended patterns of self-care after the consultation ends (whether taking medications, changing health habits, making a follow-up appointment, etc.). These skills were identified largely through research into the outcomes of psychotherapy (Greenberg & Pinsof, 1986; Stiles, Shapiro, & Elliott, 1986; Strong & Claiborn, 1982) and acquiring them is a clear case of behavioral training via graded steps. A great deal is now known about teaching such skills (Cushing, 1996; Byrne & Byrne, 1996; Dickson, 1989; Egan, 1994; Evans, Stanley, Mestrovic, & Rose, 1991; Froelich, Bishop, & Dworkin, 1976; Kimberlin, 1989; Winefield, 1992).

To be most effective, the communication process will often need to take into account how patients' responses and needs are affected by their psychosocial development, age, education, level of anxiety and cultural background. Thus flexible communication skills need sound psychological foundations.

Skilled communication increases patient satisfaction and compliance, reduces the rates of malpractice litigation, and achieves nonspecific (placebo) benefits, all of which may additionally contribute to the health worker's job satisfaction. Public health benefits can follow from targeted communications or information campaigns with health-promoting or educational messages. These require rather different sets of skills from the one-to-one interactions which are the present focus.

Health workers also need skills in communicating with their colleagues, superiors (clinicians, managers, and administrators), members of related disciplines, and often with the relatives or friends of the designated patient. Less commonly but just as critically, health professionals may need to communicate effectively with politicians, accountants, and other decision-makers. So far none of these skills has received as much attention as the skills for communicating with patients, yet they may prove vital in securing community support and resources for effective health care to proceed.

8.07.3.3.2 Understanding of and respect for scientific enquiry

It does not need a high-flying mathematical background (though some health students have that) to comprehend the basic concepts of experimental design such as matched control groups, random allocation to treatment or control groups, standardized outcome measures, and credible lengths of follow-up. Also practitioners of every discipline need to be aware of the nonspecific or placebo influences on healing which are independent of discipline-specific treatment techniques, whether pharmacotherapy, spinal manipulation, or craft activities. The psychological and therefore physical benefits of trust, hope, a sense of control and a sense of self-worth are relevant both to "feeling better" after contact with health professionals and also to the whole range of self-care (compliance) and preventive activities by the patient.

If students do not learn about the need for untreated and placebo control groups and about how to evaluate claims of treatment efficacy, there is a risk of inflated beliefs about the impact of their own discipline's specific effectiveness (an example perhaps being claims for the rehabilitation of torture victims by physiotherapy (Bloch & Moller, 1990). Evidence-based healthcare, which means using the best techniques available according to pooled high-quality information, will increasingly be demanded by health-management organizations as well as by health professionals concerned to offer the best care (Friedman, Sobel, Myers, Caudill, & Benson, 1995; Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996). The access to information offered by the electronic revolution in information technology will help, but the attitude of curiosity and willingness to learn remain vital. Confidence in one's own capacity to evaluate claims and remain an active self-educator, as well as to work with collaborators from a variety of disciplinary backgrounds, depends on this basic ability to assess the evidence—an ability which is not specific to health psychology but which psychologists have great experience and expertise in imparting (MacLeod & McCullough, 1994).

The psychologist's skills in research design and methodology, standardized measurement, and statistical analysis are increasingly valued by other health disciplines, as the sophistication of their research rises. Even where health students cannot become expert in these matters themselves, they can acquire a respect for the scientific approach to understanding human behavior, and the cognitive tools to ask critical questions of their own research literature and other sources of claims about new treatments. There is a permanent gain from having learned how to appraise the flood of information to which students are likely to be exposed during their working lives, and to develop tools for assessing its value and meaning.

Once data have been collected and analyzed, and even after a scientific paper has appeared in a prestigious journal, much remains to be done to translate the findings to improve community health and to ensure the sustainability of interventions after the researchers leave (Altman, 1995). Action research principles (Greene, 1988; Hart & Bond, 1995; Whyte, 1991) take seriously the obligation of researchers to ensure visibility of their work to decision-makers. Often this involves collaborating actively with key stakeholders from the community in order to gain access to subjects and measurements, as well as help in publicizing the results and ensuring their conversion to policy (Altman, 1995).

Another development within health psychology research which owes some impetus to health professions other than psychology, is the relatively recent interest in qualitative research methods (Patton, 1990; Waitzkin, 1990). Triangulation, the combination of quantitative and qualitative approaches in the same investigation can be extremely fruitful, as shown for example in recent studies of occupational stress and health in a factory by Baker, Israel, and Schurman (1996), and of the complex determinants and consequences of patient participation in consultations by O'Brien and Petrie (1996).

8.07.3.3.3 Appropriate patterns of practice

The ability to monitor one's own professional performance increases recognition of the limits of one's own disciplinary expertise, while failure to know one's limits is probably responsible for some of the worst cases of mismanaged health care efforts.

With the typically crowded curriculum of health disciplines, the health psychologist teacher will be able to give practical training in communication skills but not in performing psychological diagnosis or managing behavior change. Gaining an understanding of the principles of the latter two areas will allow nonpsychologist professionals to make effective referrals to specialist services including from clinical health psychologists (Belar, Deardorff & Kelly, 1987; Winefield, 1996).

The psychological awareness and the rejection of a narrow biomedical perspective which are the desirable outcomes of health psychology teaching, can be usefully applied by other professions not only to their interactions with patients but also to reflections on their own work practices. Health professionals need a sensitivity to the causes and effects of job stress in themselves and their colleagues, as the unproductiveness and danger inherent in dissatisfied, stressed, and impaired professionals is considerable (Jones et al., 1988; Keltner & Leung, 1995; Nuffield Provincial Hospitals Trust, 1996). All health workers should receive some educational preparation about the psychological as well as physical hazards to their safety and well-being, be capable of detecting signs of emotional exhaustion and other signs of strain in themselves and colleagues, and know how to proceed next.

Occupational health psychologists may be valuable to develop job stress management interventions for health workers and more broadly, to help health professions to select, train and monitor their members to prevent occupational stress and its damaging consequences for both practitioners and patients (Landy, Quick, & Kasl, 1994). Governments may not welcome such initiatives if they lead to calls for shorter working hours (e.g., for junior, salaried doctors in training), better working conditions (e.g., not being forcibly sent to rural locations), more support from psychologists, and policies to reduce environmental and social causes of illness such as air and water pollution, unsafe workplaces, poverty, racism, unemployment, and unequal access to education and health care.

8.07.4 FORMAT AND METHODS OF TEACHING

The following comments, about how health psychology should be included in courses for other professionals, need to be seen within the context described above of (i) different practice traditions and expectations in each profession, and (ii) difficulties in comparing the effectiveness of different teaching methods due to lack of definition of the desired outcome: high-quality health care.

Beginning teachers may find they have little autonomy to decide course content or teaching methods, and that crowded student timetables allow little flexibility. The health psychologist will then need to gain the trust of other faculty members as in any relationship between a service provider and a client.

There are no agreed formulae for ensuring student interest in and usage of health psychology material, as with other parts of their curriculum. In fact some students who enter their training with unrealistic ideas of the biological nature of health care, or who have limited tolerance for ambiguity, may be expected to find health psychology quite aversive. On the other hand some may find this holistic perspective a very welcome change from more molecular approaches. Evaluative feedback from students through anonymous ratings has many of the same problems as a measure of teaching quality as does patient satisfaction as a measure of health care quality. However, the very fact of staff showing concern about the students' response can raise their morale.

8.07.4.1 Lectures or Practice?

Swerissen and Foreman (1991) advocate utilizing different teaching methods for the "knowledge" and the skills or "application" strands of health professional courses. One strand uses lectures to teach health-focused psychological knowledge including psychological factors in illness and recovery, behavioral risk factors, stress and pain, and reactions to illness, hospitalization, disability, and death. The other strand consists of training vocationally-relevant skills such as communication, research and program evaluation, case management, and compliance strategies, and is taught to competency standards with supervised practice in applied settings. This was the procedure explicitly accepted by Niven (1994).

The second strand especially, is time- and labor-intensive and demands a cadre of trained health psychologists to act as tutors and supervisors. Yet this is likely to be the component of the course which ensures student respect for and retention of the more theoretical knowledge.

8.07.4.1.1 Lectures

Although listeners retain only a fraction of a lecture's content afterwards, the efficiency of one person presenting carefully selected material to a large class ensures the survival of this teaching technique in most establishments. On the whole, anything which increases the active participation of students, such as questions, short sessions of discussion ("buzz groups"), or segments of audiovisual material, is likely to improve student attentiveness (Newble & Cannon, 1994). Guest lecturers who use clinical examples to illuminate the connections between their daily work and the course material, can also stimulate student interest.

8.07.4.1.2 Structured observations of practice

Clinical settings can bewilder novice students, who often feel confused, superfluous, and embarrassed by their own lack of knowledge. Observing what a hospital is like, and how difficult it is to get information into and out of the system, can be valuable, but specific efforts to link observations to course material (for example, to report back on communication interchanges or behavior influence attempts) oblige students to process what they observe in the terms of psychological concepts. Debriefing may sometimes be required after the observation of negative models: entry-level students in particular have a capacity to identify with the patient and this furnishes powerful teaching opportunities.

8.07.4.1.3 Small (tutorial) group learning

A group of up to a dozen or so students which meets regularly with a tutor who adopts an egalitarian role and who facilitates active problem-solving, can provide both social and educational benefits. Again, adding structure to the meetings rather than always relying on spontaneous discussion, can increase motivation by making participation in group meetings assessable. Assignments to be carried out between meetings can include small fieldwork projects (ask new parents about early attachment experiences, find examples of different types of human conditioning, monitor own heart rate and emotional state, etc.) or involve critically reviewing a pertinent journal article. Demonstrations of diagnostic methods (including psychological ones), experiences of muscle relaxation training, and whole sessions devoted to independent group research on a topic for subsequent presentation are other useful activities.

In early stages of skill acquisition, group demonstrations and rehearsals in a nonthreatening environment are a good way to begin. One very natural application of this is in communication skills training, where feedback from standardized patients and group support can help to overcome self-consciousness. Assessable practice tasks and cumulative learning across course-years ("vertical integration") then build skills to the point of clinical competence.

The peer supportiveness which can emerge from a group such as this, and the experience of cooperative mutually helpful work, can be valuable learning experiences in themselves (Lucero, Jackson, & Galey, 1985; Newble & Cannon, 1994). Although work in the health professions is often collaborative and interdisciplinary, too often competition rather than cooperation is the norm between students, and skills in building constructive team dynamics are frequently neglected (Rowe, 1996).

8.07.4.1.4 Learning via the Internet

Computer Aided Learning (CAL) modules have been available for health students for some time, but the focus was initially on biological rather than psychosocial content. However, a consortium of British universities is now working to produce a set of CAL programs on behavioral science topics including pain, child development, clinical decision-making, changing health beliefs, managing stress, and quality of life measures.

The Internet, an exponentially expanding source of information, opinion, and marketing efforts, promises to revolutionize access to knowledge for both students and teachers. Methods not only of instruction but of assessment will consequently be modified. It will be all the more important for students to learn how to evaluate claims and develop critical thinking skills, in order for them to exploit this medium effectively and with discernment.

8.07.4.2 Integrated, Problem-based Learning

Just as education in nursing, physiotherapy, and some other disciplines is moving from the hospital to the classroom and increasing its scope and formality, medical education which has always been carried out in universities is in a sense moving in the reverse direction, with increased emphasis on "problem-based learning" (PBL) which blurs the old preclinical–clinical boundaries (Engel, 1992; Kaufman, 1985). In this approach a small group of students is presented by videotape, audiotape, or in writing with a carefully-constructed clinical case, and over the course of a series of meetings with their tutor, decide what information they need to solve the problems presented. Learning is thus thematically organized rather than discipline based, and is cumulative and learner directed. The goal is the development and enhancement of active reasoning and problem-solving skills rather than passive and superficial rote learning.

Norman and Schmidt (1992), although finding no evidence that PBL improves content-free problem-solving, reported that it enhances transfer of concepts to new problems and of basic science to clinical problems, and that it enhances intrinsic interest and self-directed learning skills; a gain which is maintained. Dolmans, Gijselaers, Schmidt, and van der Meer (1993) investigated whether students miss out on crucial curriculum material when staff are not in control of what they learn, which is one frequent concern with PBL for students and teaching staff. They reported a 64% overlap between faculty objectives and the learning issues generated by students, and that students also generated relevant learning issues not expected by the faculty. Low overlap of student- and staff-determined content could be used to screen out or reformulate unsuitable problems.

Part of the difficulty in evaluating the effect of a PBL curriculum in comparison to a more traditional format lies in the fact that students who are drawn to PBL may be different from the others. For example Cariaga-Lo, Richards, Hollingsworth, and Camp (1996) found that PBL entrants had higher admission test scores and were more self-reliant. Where the whole curriculum has become problem-based, Kaufman and Holmes (1996) found a favorable response to the change in students, tutors, and faculty staff, and identified four strategies which they believed had facilitated this acceptance, namely in-service training, resource support, feedback mechanisms, and participation in decision-making.

8.07.4.3 Assessment

To be congruent with the goals of PBL, reasoned decision-making rather than recall should be built into the assessment procedure (Engel, 1992). Taking this to its logical conclusion implies some reliance on self-assessment of progress towards self-selected goals (Kaufman, 1985; Lofgren & Lundahl, 1996; Ryan, Little, Smith, McMillan, & Hengstberger-Sims, 1992). However in practice a place for external assessments is likely to be maintained, due to the demands for credentials, public accountability and rank-ordered access to the most desirable training posts.

Written examinations in health psychology lend themselves to essay-type answers rather than to multiple choice questions. Although the marking is relatively more labor intensive, the teacher has the satisfaction of knowing that students are practicing a valuable and relevant skill (written communication), and that such methods are supposed to foster deep as opposed to superficial learning. Assignments in fieldwork or independent research can also be presented for assessment in written form, thus reducing the proportion of the assessment which is derived from examination papers, and providing students with a souvenir of the course which they may find instructive to retain. A real acceptance of responsibility for their own learning by students, will hopefully reduce the temptations of plagiarism which becomes more and more difficult for teachers to detect.

The assessment of practical skills needs other approaches. Segments of audio- or videotape can illustrate the student's level of competence in the early stages, for example, of communication skills acquisition, to be followed later by observed performance in standardized situations such as an objective structured clinical examination (Hodges, Turnbull, Cohen, Bienenstock, & Norman, 1996; Morrison, McNally, Wylie, McFaul, & Thompson, 1996). In general, feedback is the key to turning assessment experiences into learning experiences (Ende, 1983; Pendleton, Schofield, Tate, & Havelock, 1984; Prideaux, 1992).

8.07.4.4 Interdisciplinary Teaching

Much of the input of health psychologists to teaching in other health disciplines occurs in cooperation with multidisciplinary teams. Therefore flexibility about other perspectives and a capacity for constructive teamwork can be presented to students as part of showing what psychology understands and can offer to teach. Material on roles, group dynamics, and the social psychology of power and influence is all relevant.

8.07.4.5 Applying the Principles of Permanent Behavior Change

The psychosocial aspects of health and health care can seem irritatingly vague and complex to students who are more familiar with the clarities of the early stages of physical sciences. Health psychologists therefore need to make concrete for them how the psychosocial and physical are interdependent and how patient response to treatments is a subtle combination of both (as summarized for example by Cohen & Rodriguez, 1995). Because of its long association with education and its particular skills in understanding and producing behavior change, psychology should be well-placed to facilitate the permanent learning which is required.

The aim is to achieve long-term learning which will last students not just to the end-of-year exam but for the duration of their professional working lives. Hopefully they never develop undesirable habits such as adopting an exclusively biomedical focus or inattentive, unresponsive communication styles. However as with children beginning to smoke cigarettes, bad influences may emanate from peers and from prestigious models. To maintain learned changes over long periods of time, the natural agents of reinforcement must replace the psychologist; in other words other faculties than health psychology must encourage the application of health psychology knowledge and skills, and eventually the application of these in practice must become self-rewarding. Continuing on-the-job training or retraining ("booster sessions") are likely to be valuable, especially once gaps in knowledge and skills have become apparent to workers or their supervisors.

The "stages of change" model (Prochaska, DiClemente, & Norcross, 1992) from the behavior change literature shows that it is not worthwhile to attempt behavior change with people who have not committed themselves to make the change. Analogously perhaps, there is no point in advocating health psychology learning until students see its relevance. Thus the ground needs to be prepared by showing its usefulness to their future work. In early stages of education careful planning is needed to do this within a largely nonpracticing environment for the students. A long-term perspective is needed when health psychologists begin to try to influence other professions; fortunately other health professions have shown receptivity to psychological perspectives for some time now.

One way to facilitate the integration of health psychology knowledge and professional applications as early in the course as possible, is to engage in co-teaching with high-status practitioners of the discipline in question. Health psychologists working in clinical settings also present valuable teaching opportunities. Another way to encourage retention of undergraduate learning is to make explicit its role as a foundation for health psychology components of postgraduate masters and doctoral courses and entrance exams for specialist colleges.

8.07.5 CONTENT

Health psychology ranges in the size of its unit of analysis from the culture and norms of large community groups and institutions to individual neuroendocrines, and obviously different aspects of the subject will be of greatest interest to different disciplines. Also, the relevance of parts of health psychology knowledge may depend on the setting in which students of other disciplines will mostly practice, such as the large institution (e.g., hospital or corporation), in the community, or in a solo or small-group practice/business. Nonetheless certain common themes occur in the work of all health professionals, and these are centered on the whole person level of analysis. Some such common topics across disciplines, where health psychology has a large teaching contribution to offer, are listed below.

For reasons of space and duplication, this chapter cannot offer exhaustive references. Readers are referred to excellent reviews in other chapters in this volume, where experts provide comprehensive surveys of health psychology knowledge about 22 kinds of health problems. Resource books in health psychology include those by Jospe, Nieberding, and Cohen (1980), King and Remenyi (1986), Harvey (1988), Johnston and Marteau (1989), Bennett, Weinman, and Spurgeon (1990), Taylor (1995), Pitts and Phillips (1991), Brannon and Feist (1992), Sarafino (1994), Steptoe and Wardle (1994), and Broome and Llewelyn (1995). Key periodical sources of timely research reports and reviews include *Health Psychology*, *Psychology and Health*, *Psychology, Health and Medicine*, the *Journal of Health Psychology*, the *British Journal of Health Psychology*, the *Journal of Occupational Health Psychology*, *Social Science & Medicine*, behavioral medicine journals, and clinical, social, and organizational psychology journals. Papers with health psychologist authors also appear regularly in the psychiatric, general practice/primary care, public health, and health services administration journals.

Obtaining mastery of the Medline and other computerized data retrieval systems, and of the Internet, are highly desirable skills for educators. As libraries are forced to reduce their journal subscriptions, many seem to be more willing to obtain photocopied articles from sister institutions, sometimes by facsimile. Information technology improvements are of critical value for health psychologists as for other academics, since nobody can expect that their own colleagues or contacts will be able to provide a comprehensive picture of the state of current knowledge in this rapidly expanding field. National and international conferences abound, and there are numerous newsletters, seminars, and workshops sponsored by scientific and professional societies (some with multidisciplinary membership) with interests in health psychology.

The broad goals of introducing health psychology to the curricula of other health professions were outlined in Section 8.07.3, with a selection of references. The following checklist therefore summarizes the specific topics which health psychologist teachers and trainers in other disciplines may find useful to include in their courses.

8.07.5.1 Topics About Health and Illness

Health. The definition and measurement of health; how people assess their own state of health and decide how to proceed; comorbidity of psychological and physical distress; how health beliefs and practices can be changed by educational campaigns in order to promote health and prevent illness; genetic screening and counseling and their ethical implications.

Stress. How physical well-being is affected by chronic stress or anxiety, and how such effects can be prevented or reduced; the role of coping strategies and social supports; acute stress arising from invasive or painful procedures and how to reduce patients' suffering and hasten their recovery; psychological responses to chronic illnesses and disabilities, including in family caregivers.

Pain. This includes chronic pain and headaches; its determinants and management by psychological methods.

Aging. The stages of human cognitive and psychosocial development through life, including the challenges and vulnerabilities typical of each age group, and the influences of sex, social class, and cultural background; how people respond to loss and bereavement, including how they can be helped to face their own death or that of loved ones.

Risk factors. Sociocultural and behavioral risk factors for various illnesses; special needs of women, children, the elderly, handicapped, ethnic minorities, homosexuals, and the unemployed.

Referral. When and how to make referrals to local clinical health psychologists for problems outside the other discipline's own area of expertise.

8.07.5.2 Topics About the Interaction Between Health Professional and Patient

Roles. The roles of professional help-provider and help-seeker, and the dynamics of power and influence between the two parties; roles within health care teams and institutions; decision-making about health care services and resources.

Communication. The mechanisms of effective communication between the professional and patient with their differing amounts and kinds of knowledge and differing expectations; techniques for the specially demanding communication tasks such as giving bad news or dealing with "difficult" patients; informed consent and ethical relations with clients.

Adherence. The nature and reasons for patient nonadherence to professional recommendations, and how nonadherence can be reduced or at the least, explored and taken into account.

Behavior change. Attitude and behavior change and how to influence health-relevant behaviors such as smoking, alcohol and drug use, unsafe sex, weight, diet, and exercise; other behavioral interventions such as biofeedback, relaxation training, and postaddiction relapse immunization; worksite stress management. Other psychological treatments including counseling for psychosocial problems and preparation for surgery and hospitalization.

Placebo effects. The powerful nonspecific healing effects of hope and trust, and how these can be both maximized in practice and taken into account in evaluating treatments.

Burnout. The psychological hazards of work in human service professions, and how helping professionals may notice and deal flexibly with the particular stressors which arise (for them and for colleagues) from their work.

8.07.6 IMPLICATIONS AND CONCLUSIONS

This final section gives consideration to the implications for health psychology and health psychologists, of their participation in the teaching and training of other disciplines.

8.07.6.1 Needed Attributes of Health Psychologists who Teach Other Professions

The credible teacher of health psychology to students in other health disciplines should have a sound training in psychological research, ideally some experience in health care settings as researcher or clinician, and in a perfect world, pedagogic skills too. There may however be some delays until many such people are available. In the current economic climate of higher education there is likely to be a cyclic relationship between the availability of employment positions for health psychologists, and the diversion of resources to develop training programs in health psychology. Specific professional training in health psychology is becoming available, and opportunities for practice will increase with recognition of the health psychologist's contribution and corresponding financial support from hospitals and insurers (Diekstra, 1991; Friedman, Sobel, Myers, Caudill, & Benson, 1995).

Psychologists who accepted pioneering positions in health faculties from the 1960s were rewarded for venturing outside the mainstream of academic psychology, by opportunities for collaborative research with other health professionals and correspondingly, chances to advance the discipline of health psychology. A bootstrapping process then meant that opportunities also arose for clinical work with cardiologists, oncologists, pediatricians, community nurses, general practitioners, and many other health professionals, again with advantages for the development of the discipline. However questions are now being raised about the long-term consequences of health psychologists sharing their expertise with others, which will be discussed below.

8.07.6.2 Needs of Health Psychologists who Teach Other Professions

Information about the other professions' traditions, expectations, practice patterns, and needs should be readily available from teaching colleagues and associated practitioners. This will give a vital flavor of reality to the teaching examples and assignments used.

Health psychologists who teach nonpsychology students may also need peer support. Recognition and feedback from nonpsychologist teaching colleagues would be very welcome, but it may be valuable to use professional societies and their journals or newsletters and conferences to develop networks of people with similar jobs, in order to share information, practical advice, and understanding. Groups with a focus on health psychology within the professional societies include Division 38 in the American Psychological Association, a special group within the British Psychological Society, a section in the Canadian Psychological Association, a college within the Australian Psychological Society, and a division in the International Association of Applied Psychology. In addition there is an Association of Medical School Professors of Psychology, a European Health Psychology Society, an International Society for Health Psychology, and doubtless other national and international organizations. As noted above, the Internet, and the national societies for health informatics (information technologies applied to health-related knowledge) with their conferences, newsletters, and e-mail discussion groups, greatly increase the opportunities for connection with health psychology teachers around the world.

8.07.6.3 Consequences for Health Psychology of Teaching Other Professions

Some of the advantages to individual health psychologists of involvement in teaching other disciplines have been outlined above. Opportunities for joint research with a wider subject pool help with physiological measures such as blood assays, immune function tests, and cardiovascular health and access to funds for health and medical research can all advance individual careers as well as the discipline itself. The experience gained in other disciplines and settings could also very valuably supplement teaching to psychology students. However some doubts have been expressed about psychologists being prepared to share their knowledge and especially their skills, with other disciplines (Frank & Ross, 1995; James, 1994; Richards, 1994).

Often psychological content is not explicitly recognized as such, in teaching materials for other disciplines. For example in one health-oriented nursing text (Clark, 1986), there is considerable information about how to change behavior by collecting baseline data, deciding what reinforcers are relevant, goal-setting and behavioral contracts. Yet the only reference to psychology in the index is to an explanation of the field of psychoneuroimmunology (immediately after the section explaining reflexology). Similarly a popular textbook in occupational therapy (Kielhofner, 1992) refers to the theory of behavior modification as being a contribution from the related field of psychology which is sometimes useful in occupational therapy practice; yet does not mention the contribution of psychological research to our understanding of group dynamics and group therapy, a field where occupational therapists feel themselves to have a major role (Gibson, 1990; Howe & Schwartzberg, 1986).

A similar lack of recognition for the role of health psychology has been notable in health insurers and health management organizations. Unless health psychology can establish its unique role in providing cost-effective health care, and provide sufficient numbers of trained practitioners, there are fears that physicians (of whom for the first time there may soon be an oversupply), nurses, social workers, occupational therapists, and others may step in to fill the gap (Frank & Ross, 1995; Swerissen & Foreman, 1991).

At the 1996 Convention of the American Psychological Association, Cynthia Belar, the current President of Division 38 (Health Psychology) gave an address entitled "Clinical health psychology: A specialty for the twenty-first century" in which she encouraged health psychologists to place public welfare above "guild issues," and to continue to cooperate with other health professions. Mutual respect and recognition amongst health professions, with due credit for discipline-specific skills and lack of encroachment, are goals to which health psychologists may aspire. Health care and psychology's role within it will benefit from efforts to share psychological knowledge and some basic skills; future expansion of health psychology will follow from widening recognition of psychologists' value as consultants and collaborators in the processes of health-related diagnosis, treatment, prevention, and research. Health psychologists who teach in other health disciplines thus have a crucial role in demonstrating to those future practitioners, the value of our discipline and profession.

8.07.7 REFERENCES

Aharony, L., & Strasser, S. (1993). Patient satisfaction: What we know about and what we still need to explore. *Medical Care Review*, *50*, 49–79.

Allen, I. (1994). *Doctors and their careers: A new generation*. London: Policy Studies Unit.

Altman, D. G. (1995). Sustaining interventions in community systems: On the relationship between researchers and communities. *Health Psychology*, *14*, 526–536.

Ashton, C. H., & Kamali, F. (1995). Personality, lifestyles, alcohol and drug consumption in a sample of British medical students. *Medical Education*, *29*, 187–192.

Avis, M., Bond, M., & Arthur, A. (1997). Questioning patient satisfaction: An empirical investigation in two outpatient clinics. *Social Science & Medicine*, *44*, 85–92.

Ayer, W. A., & Hirschman, R. D. (Eds.) (1972). *Psychology and dentistry: Selected readings*. Springfield, IL: Charles C. Thomas.

Baker, E., Israel, B., & Schurman, S. (1996). Role of control and support in occupational stress: An integrated model. *Social Science & Medicine*, *43*, 1145–1159.

Ballard, B., & Clanchy, J. (1991). *Teaching students from overseas: A brief guide for lecturers and supervisors*. Melbourne, Australia: Longman Cheshire.

Belar, C. D., Deardorff, W. W., & Kelly, K. E. (1987). *The practice of clinical health psychology*. Oxford: Pergamon.

Bennett, P., Weinman, J., & Spurgeon, P. (Eds.) (1990). *Current developments in health psychology*. London: Harwood.

Bloch, I., & Moller, G. (1990). Rehabilitaton of torture victims: Physiotherapy as part of the treatment. In T. Hegna & M. Sveram (Eds.), *Psychological and psychosomatic problems* (pp. 121–144). Edinburgh, UK: Churchill Livingstone.

Brannon, L., & Feist, J. (1992). *Health psychology: An introduction to behavior and health* (2nd ed.). Belmont, CA: Wadsworth.

Broome, A. K., & Llewelyn, S. (Eds.) (1995). *Health psychology: Process and applications* (2nd ed.). London: Chapman & Hall.

Bunce, D., & West, M. A. (1996). Stress management and innovation interventions at work. *Human Relations*, *49*, 209–232.

Burke, R. J. (1996). Stress, satisfaction and militancy among Canadian physicians: A longitudinal investigation. *Social Science & Medicine*, *43*, 517–524.

Byrne, A., & Byrne, D. (1992). *Psychology for nurses: Theory and practice*. Melbourne, Australia: Macmillan.

Byrne, D., & Byrne, A. (1996). *Counselling skills for health professionals*. Melbourne, Australia: Macmillan.

Campbell-Heider, N., & Pollock, D. (1987). Barriers to physician–nurse collegiality: An anthropological perspective. *Social Science & Medicine*, *25*, 421–425.

Cariago-Lo, L. D., Richards, B. F., Hollingsworth, M. A., & Camp, D. L. (1996). Non-cognitive characteristics of medical students: Entry to problem-based and lecture-based curricula. *Medical Education*, *30*, 179–186.

Chur-Hansen, A., & Barrett, R. J. (1996). Teaching colloquial Australian English to medical students from non-English speaking backgrounds. *Medical Education*, *30*, 412–417.

Cinotti, W. R., Grieder, A., & Springob, H. K. (1972). *Applied psychology in dentistry* (2nd ed.). St. Louis, MO: Mosby.

Clark, C. C. (1986). *Wellness nursing: Concepts, theory, research, and practice*. New York: Springer.

Cohen, S., & Rodriquez, M. S. (1995). Pathways linking affective disturbances and physical disorders. *Health Psychology*, *14*, 374–380.

Cushing, A. (1996). Editorial: Communication skills. *Medical Education*, *30*, 316–318.

Danbury, H. (1994). *Teaching practical social work* (3rd ed.). Aldershot, UK: Arena.

Dickson, D. A. (1989). Interpersonal communication in the health professions: A focus on training. Special issue: Health psychology: Perspectives on theroy, research and practice. *Counselling Psychology Quarterly*, *2*, 345–366.

Diekstra, R. F. W. (1991). Psychology, health and health care. In M. A. Jansen & J. Weinman (Eds.), *The international development of health psychology* (pp. 19–31). Philadelphia: Harwood.

DiMatteo, M. R., Friedman, H. S., & Robin, M. (1982). *Social psychology and medicine*. Cambridge, MA: Oelgeschlager, Gunn & Hain.

Dolmans, D. H., Gijselaers, W. H., Schmidt, H. G., & van der Meer, S. B. (1993). Problem-effectiveness in a course using problem-based learning. *Academic Medicine*, *68*, 207–213.

Dunkin, E. N. (1981). *Psychology for physiotherapists*. London: Macmillan.

Egan, G. (1994). *The skilled helper: A problem-management approach to helping* (5th ed.). Pacific Grove, CA: Brooks/Cole.

Ende, J. (1983). Feedback in clinical medical education. *Journal of the American Medical Association*, *250*, 777–781.

Engel, C. E, (1992). Problem-based learning. *British Journal of Hospital Medicine*, *48*, 325–329.

Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. Science, *196*, 129–136.

Evans, B. J., Stanley, R. O., Mestrovic, R., & Rose, L. (1991). Effects of communication skills training on students' diagnostic efficiency. *Medical Education*. *25*, 517–526.

Fairhurst, K., & May, C. (1995). Consumerism and the consultation: The doctor's view. *Family Practice, 12*, 389–391.

Firth-Cozens, J. (1987). The stresses of medical training. In R. Payne & J. Firth-Cozens (Eds.), *Stress in health professionals* (pp. 3–22). Chichester, UK: Wiley.

Fitzpatrick, R. M., Hopkins, A. P., & Harvard-Watts, O. (1983). Social dimensions of healing: A longitudinal study of outcomes of medical management of headaches. *Social Science & Medicine*, *17*, 501–510.

Frank, J. D. (1989). Non-specific aspects of treatment: The views of a psychotherapist. In M. Shepherd & N. Sartorius (Eds.), *Non-specific aspects of treatment* (pp. 95–114). Toronto, ON: Huber.

Frank, R. G., & Ross, M. J. (1995). The changing workforce: The role of health psychology. *Health Psychology, 14*, 519–525.

Fransella, F. (1982). *Psychology for occupational therapists*. London: Macmillan.

Friedman, R., Sobel, D., Myers, P., Caudill, M., & Benson, H. (1995). Behavioral medicine, clinical health psychology, and cost offset. *Health Psychology, 14*, 509–518.

Froelich, R. E., Bishop, F. M., & Dworkin, S. F. (1976). *Communication in the dental office: A programmed manual for the dental professional*. St Louis, MO: Mosby.

Gibson, D. (1990). *Group protocols: A psychosocial compendium*. New York: Haworth.

Gray, G., & Pratt, R. (Eds.) (1992). *Issues in Australian nursing 3*. Melbourne, Australia: Churchill Livingstone.

Greenberg, L. S., & Pinsof, W. M. (Eds.) (1986). *The psychotherapeutic process: A research handbook*. New York: Guilford Press.

Greene, J. G. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review, 12*, 9–16.

Hall, J. (1982). *Psychology for nurses and health visitors*. London: Macmillan.

Hart, E., & Bond, M. (1995). *Action research for health and social care: A guide to practice*. Milton Keynes, UK: Open University Press.

Harvey, P. G. (1988). *Health psychology*. London: Longman.

Hetherington, R. R., Miller, D. H., & Neville, J. G. (1964). *Introduction to psychology for medical students*. London: Heinemann.

Hodges, B., Turnbull, J., Cohen, R., Bienenstock, A., & Norman, G. (1996). Evaluating communication skills in the objective structured clinical examination format: Reliability and generalizability. *Medical Education, 30*, 38–43.

Howe, M. C., & Schwartzberg, S. L. (1986). *A functional approach to group work in occupational therapy*. Philadelphia: Lippincott.

Howell, J. B., & Schroeder, D. P. (1984). *Physician stress: a handbook for coping*. Baltimore: University Park Press.

James, D., & Vinson, T. (Eds.) (1989). *Advances in social welfare education 1989*. Sydney, Australia: School of Social Work, University of New South Wales.

James, J. E. (1994). Health care, psychology, and the scientist-practitioner model. *Australian Psychologist, 29*, 5–11.

Johnston, M., & Marteau, T. (Eds.) (1989). *Applications in health psychology*. New Brunswick, NJ: Transaction.

Jones, J. W., Barge, B. N., Steffy, B. D., Fay, L. M., Kunz, L. K., & Wuebker, L. J. (1988). Stress and medical malpractice: Organizational risk assessment and intervention. *Journal of Applied Psychology, 73*, 727–735.

Jospe, M., Nieberding, J., & Cohen, B. D. (1980). *Psychological factors in health care: A practitioner's manual*. Lexington, KY: Heath.

Kaplan, R. M. (1994a). Using quality of life information to set priorities in health policy. *Social Indicators Research, 33*, 121–163.

Kaplan, R. M. (1994b). The Ziggy theorem: Toward an outcomes-focused health psychology. *Health Psychology, 13*, 451–460.

Kaplan, S. H., Greenfield, S., & Ware, J. E. (1989). Assessing the effects of physician–patient interactions on the outcomes of chronic disease. *Medical Care, 27*, S110–S127.

Kaufman, A. (Ed.) (1985). *Implementing problem-based medical education: Lessons from successful innovations*. New York: Springer.

Kaufman, D. M., & Holmes, D. B. (1996). Tutoring in problem-based learning: Perceptions of teachers and students. *Medical Education, 30*, 371–377.

Keltner, A., & Leung, R. (1995). Perceived stressors and stress responses among hospital-based nursing trainees. *International Journal of Stress Management, 2*, 97–111.

Kielhofner, G. (1992). *Conceptual foundations of occupational therapy*. Philadelphia: Davis.

Kimberlin, C. L. (1989). Communications. In A. I. Wertheimer and M. C. Smith (Eds.), *Pharmacy practice: Social and behavioral aspects* (pp. 159–177). Baltimore: Williams & Wilkins.

King, N. K., & Remenyi, A. (1986). *Health care: A behavioral approach*. Sydney, Australia: Grune & Stratton.

Landy, F., Quick, J. C., & Kasl, S. (1994). Work, stress, and well-being. *International Journal of Stress Management, 1*, 33–73.

Lofgren, M., & Lundahl, L. (1996). Self-marking in written examination: A way of feedback and learning. *Medical Education, 30*, 322–325.

Lucero, S. M., Jackson, R., & Galey, W. R. (1985). Tutorial groups in problem-based learning. In A. Kaufman (Ed.) *Implementing problem-based medical education: Lessons from successful innovations* (pp. 45–70). New York: Springer.

MacLeod, S. M., & McCullough, H. N. (1994). Social science as a component of medical training. *Social Science & Medicine, 39*, 1367–1373.

Martin, E. W. (1996). An update on census data: Good news for social work? *Australian Social Work, 49*, 29–36.

McCue, H., & White, R. (1983). *Innovation in nursing curricula: A survey of developments in Australia*. Kensington, Australia: Centre for Medical Education Research and Development, University of New South Wales.

McGhie, A. (1973). *Psychology as applied to nursing* (6th ed.). Edinburgh, UK: Churchill Livingstone.

Midence, K., Gregory, S., & Stanley, R. (1996). The effects of patient suicide on nursing staff. *Journal of Clinical Nursing, 5*, 115–120.

Morrison, H., McNally, H., Wylie, C., McFaul, P., & Thompson, W. (1996). The passing score in the Objective Structured Clinical Examination. *Medical Education, 30*, 345–348.

Newble, D. & Cannon, R. (1994). *A handbook for medical teachers* (3rd ed.). Dordrecht, The Netherlands: Kluwer.

Niven, N. (1994). *Health psychology: An introduction for nurses and other health care professionals* (2nd ed.). Edinburgh, UK: Churchill Livingstone.

Norman, G. R., & Schmidt, H. G. (1992). The psychological basis of problem-based learning: A review of the evidence. *Academic Medicine, 67*, 557–565.

Nuffield Provincial Hospitals Trust Report of a Working Party (1996). *Taking care of doctors' health: Reducing avoidable stress and improving services for doctors who fall ill*. London: Nuffield Provincial Hospitals Trust.

O'Brien, M., & Petrie, K. J. (1996) Examining patient participation in medical consultations: A combined quantitative and qualitative approach. *Psychology and Health, 11*, 871–890.

Paton, D., & Brown, R. (1991). *Lifespan health psychology: Nursing problems and interventions*. London: HarperCollins.

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage.

Payne, S., & Walker, J. (1996). *Psychology for nurses and the caring professions.* Buckingham, UK: Open University Press.

Pearson, A. (1996). *Doctor of nursing: Curriculum document.* Unpublished pro forma, University of Adelaide Faculty of Medicine, Department of Clinical Nursing.

Pendleton, D., Schofield, T., Tate, P., & Havelock, P. (1984). *The consultation: An approach to learning and teaching.* Oxford: Oxford University Press.

Pickering, W. G. (1993). Patient satisfaction: An imperfect measurement of quality medicine. *Journal of Medical Ethics, 19*, 121–122.

Pitts, M., & Phillips, K. (Eds.) (1991). *The psychology of health: An introduction.* London: Routledge.

Pomerleau, O. F., & Brady, J. P. (Eds.) (1979). *Behavioral medicine: Theory and practice.* Baltimore: Williams & Wilkins.

Prideaux, D. (1992). "We are only interested in the exam'. Assessment and self directed learning in medical education. In R. M. Harden, I. R. Hart, & H. Mulholland (Eds.), *Approaches to the assessment of clinical competence Part I* (pp. 90–95). Dundee, UK: Centre for Medical Education, University of Dundee.

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist, 47*, 1102–1114.

Purser, H. (1982). *Psychology for speech therapists.* London: Macmillan.

Quick, J. (1996, August) *Occupational health psychology: From the crossroads …* Invited address for Division 38 (Health Psychology), delivered at the American Psychological Association Annual Convention, Toronto.

Quick, J. C., Murphy, L. R., & Hurrell, J. J., Jr. (Eds.) (1992). *Stress and well-being at work: Assessments and interventions for occupational mental health.* Washington, DC. American Psychological Association.

Rachman, S. (Ed.) (1977). *Contributions to medical psychology* (Vol. 1). Oxford: Pergamon.

Reid, J. C. (1994). *Nursing education in Australian universities.* Canberra, Australian: Australian Government Publishing Service.

Report of the Study Commission on Pharmacy (1975). *Pharmacists for the future.* Ann Arbor, MI: Health Administration Press.

Richards, J. (1994). Giving psychology away? A comment on "Health care, psychology, and the scientist-practitioner model" by James (1994). *Australian Psychologist, 29*, 12–14.

Rowe, H. (1996). Multidisciplinary teamwork—myth or reality? *Journal of Nursing Management, 4*, 93–101.

Ryan, G., Little, P., Smith, G., McMillan, M., & Hengstberger-Sims, C. (1992). Implementing problem-based learning in nursing. In G. Gray, & R. Pratt (Eds.), *Issues in Australian nursing 3* (pp. 195–212). Melbourne, Australia: Churchill Livingstone.

Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal, 312*, 71–72.

Sarafino, E. P. (1994). *Health psychology: Biopsychosocial interactions* (2nd ed.). New York: Wiley.

Shepherd, M., & Sartorius, N. (Eds.) (1989). *Non-specific aspects of treatment.* Toronto, ON: Huber.

Shulman, L. (1993). *Teaching the helping skills: A field instructor's guide.* Alexandria, VA: Council on Social Work Education.

Siegler, M. (1993). Falling off the pedestal: What is happening to the traditional doctor–patient relationship? *Mayo Clinic Proceedings, 68*, 461–467.

Simpson, K. H., & Budd, K. (1996). Medical student attrition: A 10-year survey in one medical school. *Medical Education, 30*, 172–178.

Steptoe, A., & Wardle, J. (Eds.) (1994). *Psychosocial processes and health: A reader.* Cambridge: University of Cambridge Press.

Stewart, M. A. (1995). Effective physician–patient communication and health outcomes: A review. *Canadian Medical Association Journal, 152*, 1423–1433.

Stiles, W. B., Shapiro, D. A., & Elliott, R. (1986). "Are all psychotherapies equivalent?" *American Psychologist, 41*, 165–180.

Stoudemire, A. (1994). *Human behavior: An introduction for medical students* (2nd ed.). Philadelphia: Lippincott.

Strong, S. R., & Claiborn, C. D. (1982). *Change through interaction: Social psychological processes in counseling and psychotherapy.* New York: Wiley.

Swerissen, H., & Foreman, P. (1991). Training health professionals in health psychology. In M. A. Jansen, & J. Weinman (Eds.), *The international development of health psychology* (pp. 125–133). Chur, Switzerland: Harwood.

Taylor, S. E. (1995). *Health psychology* (3rd ed.). New York: McGraw-Hill.

Thompson, N. (1995). *Theory and practice in health and social welfare.* Buckingham, UK: Open University Press.

Vuckovic, N. & Nichter, M. (1997). Changing patterns of pharmaceutical practice in the United States. *Social Science & Medicine, 44*, 1285–1302.

Waitzkin, H. (1990). On studying the discourse of medical encounters. A critique of qualitative and quantitive methods and a proposal for reasonable compromise. *Medical Care, 28*, 473–488.

Ware, J. E., Jr. (1995). The status of health assessment 1994. *Annual Review of Public Health, 16*, 327–354.

Wedding, D. (1995). *Behavior and medicine* (2nd ed.). St. Louis, MO: Mosby.

Weinman, J. (1987). *An outline of psychology as applied to medicine* (2nd ed.) Bristol, UK: Wright.

Weinman, J. (1990). Health psychology: Progress, perspectives and prospects. In P. Bennett, J. Weinman, & P. Spurgeon (Eds.), *Current developments in health psychology* (pp. 9–33). London: Harwood.

Wertheimer, A. I., & Smith, M. C. (Eds.) (1989). *Pharmacy practice: Social and behavioral aspects.* Baltimore: Williams & Wilkins.

Whyte, W. F. (1991). *Participatory action research.* Newbury Park, CA: Sage.

Williams, B. (1994). Patient satisfaction: A valid concept? *Social Science & Medicine, 38*, 509–516.

Winefield, H. R. (1991). Health psychology for medical students. In M. A. Jansen, & J. Weinman (Eds.), *The international development of health psychology* (pp. 135–143). Philadelphia: Harwood.

Winefield, H. R. (1992). Doctor–patient communication: An interpersonal helping process. In S. Maes, H. Leventhal, & M. Johnston (Eds.), *International review of health psychology* (Vol. 1, pp. 167–187). Chichester UK: Wiley.

Winefield, H. R. (1996). Counselling work in normal general practice: An analysis of Australian consultations. *Journal of Health Psychology, 1*, 223–234.

Winefield, H. R., & Peay, M. Y. (1991). *Behavioural science in medicine* (2nd ed). Adelaide, Australia: University of Adelaide.

Winefield, H. R., Murrell, T. G., & Clifford, J. (1995). Process and outcomes in general practice consultations: Problems in defining high quality care. *Social Science & Medicine, 41*, 969–975.

Wise, T. N. (1986). Psychosomatic medicine: Integrating dynamic, behavioral and biological perspectives. *Psychotherapy and Psychosomatics, 46*, 85–95.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.08
Cardiovascular Disorders

DAVID S. KRANTZ and NICOLE R. LUNDGREN
Uniformed Services University of the Health Sciences, Bethesda, MD, USA

8.08.1 INTRODUCTION 190

8.08.2 CORONARY HEART DISEASE 190

- *8.08.2.1 Overview of Pathophysiology* 190
  - *8.08.2.1.1 Standard risk factors for CHD* 190
- *8.08.2.2 Psychosocial Risk Factors for CHD* 191
  - *8.08.2.2.1 Stress and CHD* 191
  - *8.08.2.2.2 Animal studies of stress and atherosclerosis* 191
  - *8.08.2.2.3 Occupational stress* 192
  - *8.08.2.2.4 Low levels of social support and lack of economic resources as CHD risk factors* 193
  - *8.08.2.2.5 Acute stress and anger as cardiac disease triggers* 193
  - *8.08.2.2.6 Stress and cardiac ischemia* 195
  - *8.08.2.2.7 Summary* 196
- *8.08.2.3 Individual Characteristics as CHD Risk Factors* 196
  - *8.08.2.3.1 Type A behavior: current status* 196
  - *8.08.2.3.2 Anger and hostility* 198
  - *8.08.2.3.3 Clinical depression as a CHD risk factor* 198
  - *8.08.2.3.4 Stress reactivity* 200
- *8.08.2.4 Treatment of CHD* 200
  - *8.08.2.4.1 Medical and surgical treatment options* 200
  - *8.08.2.4.2 Exercise and behavioral components of cardiac rehabilitation* 201
- *8.08.2.5 Psychosocial Treatment Approaches for the Cardiac Patient* 201
  - *8.08.2.5.1 Modifying hostility and type A behavior* 202
  - *8.08.2.5.2 Interventions to increase social support and reduce life stress* 202
  - *8.08.2.5.3 The lifestyle heart trial* 203
  - *8.08.2.5.4 Summary* 204

8.08.3 ESSENTIAL HYPERTENSION 204

- *8.08.3.1 Genetic–Environment Interactions* 206
- *8.08.3.2 Stress, Behavioral Factors, and Essential Hypertension* 206
  - *8.08.3.2.1 Personality and essential hypertension* 207
  - *8.08.3.2.2 Salt intake* 208
  - *8.08.3.2.3 Obesity* 208
- *8.08.3.3 Treatment of Essential Hypertension* 208
  - *8.08.3.3.1 Weight loss and dietary salt restriction in hypertension management* 209
  - *8.08.3.3.2 Effects of exercise training in hypertension* 209
  - *8.08.3.3.3 Effects of reducing alcohol intake* 209
  - *8.08.3.3.4 Effects of stress management, biofeedback, and other cognitive interventions in hypertension* 209
- *8.08.3.4 Adherence to Antihypertensive Treatments* 210
- *8.08.3.5 Summary* 211

8.08.4 CONCLUSION 211

8.08.5 REFERENCES 211

8.08.1 INTRODUCTION

Chronic diseases of the cardiovascular system, which include coronary heart disease (CHD) (also known as ischemic heart disease or atherosclerotic heart disease), high blood pressure, and stroke, constitute a major public health problem and the leading cause of death in Western countries (Higgins & Luepker, 1988; National Heart, Lung, and Blood Institute [NHLBI], 1994). Many physiological, environmental, and behavioral variables interact in the development of these disorders. For example, many of the causal agents for CHD can be modified, relate to habits of living, and are under the control of the individual. It can therefore be said that CHD is a disorder that is a result of the individual's lifestyle, and it is not surprising that cardiovascular diseases have been among the most widely studied topics in health psychology (e.g., Baum, Gatchel, & Krantz, 1997; Krantz, Grunberg, & Baum, 1985).

Over the past 30 years in the USA, there has been encouraging news that death rates from heart disease have been decreasing (NHLBI, 1994). Through the use of epidemiologic and statistical analyses these changes have been attributed partly to dramatic medical developments such as drugs and improvements in medical technology, but also to awareness by the public of behaviors that are risk factors, and attempts by individuals to modify components of lifestyle that constitute risk factors for heart disease (Higgins & Luepker, 1988). Goldman and Cook (1988), for example, used the epidemiologic literature to estimate that about 40% of the reduction in cardiovascular mortality was due to improvements in medical care, such as drug treatments for high blood pressure and improvements in cardiac care. However, more than half of the decline in death rates from CHD (54%) could be attributed to lifestyle changes, such as changes in fat consumption and reductions in cholesterol levels (30%) and cigarette smoking (24%). These data reinforce the important link between individual lifestyle and risk of coronary disease. Because of space limitations and the vastness of the available literature, this chapter gives a selective overview of behavioral science contributions to understanding the etiology and treatment of two of the major cardiovascular disorders: CHD and essential hypertension. For comprehensive reviews of various aspects of this vast literature, the reader is referred to Allan and Scheidt (1996), Dubbert (1995), Julius and Bassett (1987), Ockene and Ockene (1992), Shumaker and Czajkowski (1994), and Siegman and Smith (1994).

8.08.2 CORONARY HEART DISEASE

8.08.2.1 Overview of Pathophysiology

Coronary atherosclerosis is a condition that results when the coronary arteries, which deliver blood to the heart, become narrowed by fatty deposits called plaques. Ischemic heart disease refers to a set of conditions thought to result from coronary atherosclerosis. Myocardial ischemia—the inadequate flow of blood to cardiac tissue—results when arterial narrowing can cause insufficient blood supply to the heart. Ischemia, in turn, can sometimes be accompanied by chest pain—a condition called angina pectoris. Severe or repeated myocardial ischemia also predisposes the heart to life-threatening disturbances of cardiac rhythm, which can result in sudden cardiac death. A heart attack, or myocardial infarction (MI), results when cardiac tissue dies because arterial plaque becomes unstable and a complete blockage occurs, or when ischemia becomes prolonged or severe.

8.08.2.1.1 Standard risk factors for CHD

Based on a set of so-called "standard" risk factors, individuals who are likely to develop CHD can be identified with a moderate degree of accuracy. When multiple risk factors are present in the same individual, the CHD risk increases multiplicatively with the addition of each risk factor (Higgins & Luepker, 1988). Several of the coronary risk factors cannot be modified directly. These include chronological age, sex, race, and a strong family history of coronary disease.

With aging, there is a greater likelihood of developing heart disease, and nearly half of all coronary victims are aged over 65 years. Especially at younger ages, males are at greater risk of CHD than are females. The lower risk in women at younger ages is in large part attributable to a protective effect of estrogens (Lerner & Kannel, 1986). Thus, gender differences in coronary disease become much less pronounced as women reach menopause, and exogenously administered estrogens lower the risk of CHD in postmenopausal women. Also, in the USA, blacks are more prone to CHD than whites. This may be related to the fact that blacks are much more susceptible to high blood pressure, with high blood pressure being a major CHD risk factor. Another nonmodifiable risk factor is family history. Susceptibility to CHD at an early age appears to be transmitted genetically, and certain families are at higher risk than others (Glueck et al., 1985; Nicolosi & Schaefer, 1992). However, even if individuals fall into the higher risk groups because they possess the nonmodifiable risk factors, they can still minimize their risk by modifying other coronary risk factors that are amenable to change.

Modifiable risk factors for coronary disease include essential hypertension, cigarette smoking, high blood levels of low density lipoprotein (LDL) cholesterol and low levels of high density lipoprotein (HDL), diabetes, obesity, and sedentary lifestyle (Ockene & Ockene, 1992). (We discuss the so-called "psychosocial risk factors" in Section 8.08.2.2). For example, cigarette smoking is a preventable behavior, and the death rate from MI is markedly higher among smoers compared to nonsmokers. However, for those who give up the habit, the death rate begins to decrease almost to the level of those who have never smoked (Higgins & Luepker, 1988). Essential hypertension, or chronic high blood pressure, is also an important risk factor for the development of CHD that can be modified via pharmacological or behavioral means (see Section 8.08.3).

It should be noted that there is still some controversy as to the relative importance of such factors as diet and exercise in the development of coronary disease. For example, although large reductions in dietary intake of saturated fats and cholesterol have been shown to slow the progression of CHD (e.g., Ornish et al., 1990), several studies have not found that dietary patterns within the population are predictive of later development of heart disease (Mann, 1977). Generally, the least controversial and most widely accepted risk factors are considered to be smoking, high LDL and low HDL serum cholesterol levels, and high blood pressure (Kannel, 1979).

8.08.2.2 Psychosocial Risk Factors for CHD

Despite extensive research into the standard physiological risk factors, the best combinations of these factors still do not account for the occurrence of heart disease in many individuals (Jenkins, 1983). Thus, the search has been broadened to examine potential new risk factors, including social and psychological characteristics, and characteristics of the environment. In this chapter we discuss four areas of research on cardiovascular disorders:

(i) chronic and acute stress and factors such as social support that can moderate the effects of stress on cardiovascular pathology;

(ii) personality traits, such as hostility, anger, and type A behavior, which may be risk factors for CHD;

(iii) depression, which can lead to poor prognosis in cardiac patients; and

(iv) individual differences in physiologic reactivity (responsiveness) to behavioral stressors.

8.08.2.2.1 Stress and CHD

Among the lay public, it is widely believed that acute and chronic stress play a role in the onset of CHD manifestations such as MI and sudden cardiac death. Until recently, this view was corroborated primarily by anecdotal and case reports, some epidemiological research (Engel, 1971; Myers & Dewar, 1975), and human and animal studies demonstrating physiologic and pathophysiologic effects of mental stress on the cardiovascular system (Kamarck & Jennings, 1991; Manuck, Marsland, Kaplan, & Williams, 1995). An increased interest in the role of stress as a trigger of clinical cardiovascular events has been fostered by evidence suggesting that, in susceptible patients, acute cardiovascular events (e.g., MI, cardiac ischemia) are frequently triggered by activities such as physical or mental stress, rather than occurring spontaneously (Muller, Tofler, & Stone, 1989). In addition, studies making use of a variety of current techniques for assessing cardiac function in the laboratory and in the field provide convincing evidence for the pathophysiological mechanisms and effects of behavioral factors as triggers of myocardial ischemia (Deanfield et al., 1984; Gottdiener et al., 1994; Krantz et al., 1996; Rozanski et al., 1988).

The effects of chronic stress on coronary disease have been studied by assessing psychological and social conditions at work and in other life domains (e.g., home and family). In addition, useful animal behavioral models of coronary disease have been developed and utilized. Research on the role of behavior in the development of coronary artery disease has revealed that the effects of stressful situations on physiology and behavior depend on psychological and contextual factors. If situations (e.g., the workplace) are not viewed or interpreted as harmful, threatening, or challenging, they can produce smaller and even an absence of adverse physiologic responses (Lazarus & Folkman, 1984). Therefore, the relationship of various psychological and social stresses (e.g. occupational conditions) to CHD seems to depend on the meaning of the situation to the individual, and the way in which individuals perceive their life situations (Cohen, Evans, Stokols, & Krantz, 1986).

8.08.2.2.2 Animal studies of stress and atherosclerosis

In recent years, a particularly important series of studies in this area have been conducted at the Bowman–Gray School of Medicine (Kaplan & Bush, 1982). They used cynomolgus monkeys, whose coronary disease pathology closely resembles that of the human condition in these animals. In addition, many of the prominent aspects of their behavior, such as forming a social hierarchy, competition, and aggression, are salient aspects of human behavior analogous to those which have been implicated as potential contributors to coronary disease in humans.

The introduction of unfamiliar monkeys into an established social group is a powerful social stressor, leading to increased aggressive behavior as group members attempt to re-establish a social dominance hierarchy (Manuck, Kaplan, & Matthew, 1986). These stressful conditions were created by periodically reorganizing social groups (the "unstable" condition). Unstressed animals in "stable" social conditions were assigned to similarly sized groups having fixed memberships over the duration of the study (Kaplan & Bush, 1982). Based on the patterns of their behaviors in the groups, monkeys were categorized as either dominant or subordinate.

In one study of male monkeys, the dominant animals in the unstable social condition developed more extensive coronary atherosclerosis than their subordinate counterparts. However, differences were less evident between dominants and subordinates in the stable social condition—under stable social conditions, dominants were slightly less affected than subordinates. Providing a clue as to the possible behavioral reasons for the development of disease, dominant animals in the unstable condition showed the most aggression toward other animals and a disruption of positive social interactions. In another study conducted with female monkeys, subordinate animals developed greater coronary artery disease than dominants (Kaplan, Adams, Clarkson, & Koritnick, 1984), and as in the other studies with male animals, these effects could not be attributed to the physical risk factors. Reproductive function in many of the subordinate animals was also disrupted, and these behaviorally induced reproductive problems may have lessened these animals' "protection" against coronary disease.

In summary, experiments with nonhuman primates have demonstrated a causal effect of psychosocial stress in the development of coronary artery disease. Interestingly, the specific effects of stress depend on individual characteristics (e.g., level of dominance) that determine how objective conditions will affect each animal's behavior. Interestingly, the particular traits in animals (e.g., dominance and aggression) that relate to risk of atherosclerosis foreshadow similar findings obtained in human populations.

8.08.2.2.3 Occupational stress

Research on occupational stress and health has sought to determine which occupations are most stressful and the particular characteristics of occupations that lead to elevated risk of coronary disease (Karasek & Theorell, 1990). In this regard, several characteristics of settings and of individuals have been identified that potentially can be modified to reduce stress. Working conditions that have been associated with CHD risk include the psychological demands of the job, autonomy on the job (how much input a worker has in making decisions), and satisfaction on the job. Job demands refer to job conditions that tax or interfere with the worker's performance abilities, such as workload and work responsibilities. Level of job autonomy or control refers to the ability of the worker to control the speed, nature, and conditions of work. Job satisfaction includes gratification of the worker's needs and aspirations derived from employment (Wells, 1985).

Low levels of control over one's job and excessive workload seem to be a particularly important combination in heightening job-related stress. Karasek and colleagues (e.g., Karasek & Theorell, 1990) proposed that conditions of high work demands combined with few opportunities to control the job situation (low decisional latitude) are associated with increased coronary disease risk. Conditions of high demand and low control are called "high strain" situations. The Karasek "job demand/control" hypothesis has been tested in several populations by applying a "job characteristics scoring system" based on responses to several national surveys of workers. These job characteristic scores can distinguish between occupations along the dimensions comprising job strain. The job strain model has shown the ability to predict cardiovascular disease and mortality in studies of male Swedish workers, and in studies of men and women in the USA (Karasek & Theorell, 1990; Karasek et al., 1988; Schnall, Devereux, Pickering, & Schwartz, 1990). However, a recent longitudinal study of patients who underwent diagnostic testing for coronary disease did not find that occupational stress was related to extent of disease in this population or to subsequent cardiac morbidity or mortality (Hlatky et al., 1995). These negative findings may be attributable to the fact that the study group consisted of a selected group of largely symptomatic patients, and the effects of job strain may be obscured in such a population.

A study of occupational stress and CHD among women also suggests that the relationship of job conditions to coronary disease depends on the individual's work and family demands, and their control over these situations. Haynes and Feinleib (1980) analyzed data from the Framingham Heart Study, a major epidemiologic study of heart disease conducted by the National Institutes of Health. They asked the question of whether the increasing employment of women outside the home has adversely affected their cardiovascular health. In the mid-1960s, middle-aged women were examined for the development of CHD for the ensuing eight years. Working women, that is women who had been employed outside the home for more than half their adult years, were compared with housewives and with men. The results indicated that working women in general were not at significantly higher risk of subsequent coronary disease than housewives. However, clerical workers (who perhaps have low job control), and working women with children (who have high family demands) were more likely to develop CHD, as were women whose bosses were nonsupportive. Interestingly, the likelihood of CHD increased linearly with the number of children for working women, but not for housewives (Lacroix & Haynes, 1987).

8.08.2.2.4 Low levels of social support and lack of economic resources as CHD risk factors

Certain aspects of the social environment, particularly social isolation, low social support, and lack of social and economic resources can also increase an individual's risk of cardiovascular disease (Shumaker & Czajkowski, 1994). Social support refers to the instrumental, emotional, and informational aid obtained from an individual's social ties and community resources (Cohen & Wills, 1985). It has been hypothesized that social support may serve as a buffer or moderator between life stressors and cardiovascular disease outcomes. On the other hand, the absence of tangible aid from a supportive social network may also have direct adverse effects on health, for example, by impeding access to needed help or interfering with medical compliance. In studies in the USA and other countries, people with strong social networks are at a relatively lower risk of developing CHD (Shumaker & Czajkowski, 1994).

In industrialized societies, rates of cardiac morbidity and mortality are inversely related to socioeconomic status, with disease rates higher among poorer individuals. The inverse association of CHD mortality with socioeconomic level cannot be explained by different use of medical care services by individuals in lower socioeconomic groups (Marmot, 1983). Although it is known that the prevalence of standard CHD risk factors (e.g., high blood pressure and smoking) decreases with increasing socioeconomic status, this is not sufficient by itself to explain all the increase in coronary mortality attributable to low socioeconomic status.

Low social support and lack of economic resources appear to increase interactively the risk of adverse prognosis in CHD patients. Ruberman, Weinblatt, Goldberg, and Chaudhary (1984) (Figure 1) demonstrated that level of education and social support resources (elements associated with socioeconomic level) are significant factors affecting the survival of men after a first MI. In this study, men with low educational levels were likely to experience more isolation and life stress than were men with higher levels of education. Low education men had mortality rates twice as high as men with more education and social support. More recently, two studies demonstrated that living alone led to a near doubling of the risk of recurrent heart attack and death in post-MI patients, and that, independent of other prognostic medical factors, unmarried male and female coronary artery disease patients without a confidant and patients with low household incomes were at increased risk of cardiovascular death over a five-year period (Case, Moss, Case, McDermott, & Eberly, 1992; Williams et al., 1992). Social factors presumably influence prognosis by influencing patients" psychological reactions to MI. For example, the poorly educated may be at increased risk of job loss following an infarction, and for those who live alone the future may be even more bleak.

8.08.2.2.5 Acute stress and anger as cardiac disease triggers

As opposed to focusing on chronic social and environmental conditions, a relatively new research area has focused on the role of acute stress and of emotions as triggers of the onset of coronary artery disease manifestations in individuals with pre-existing disease (see Krantz et al., 1996; Muller et al., 1989). Earlier studies observed that stressful life events, such as death of a spouse, occurred in the 24 hours preceding death among patients who died suddenly from coronary disease (e.g., Cottington, Matthews, Talbott, & Kuller, 1980; Myers & Dewar, 1975) and that loss events (e.g., death of a loved one) occurred to sudden death victims more

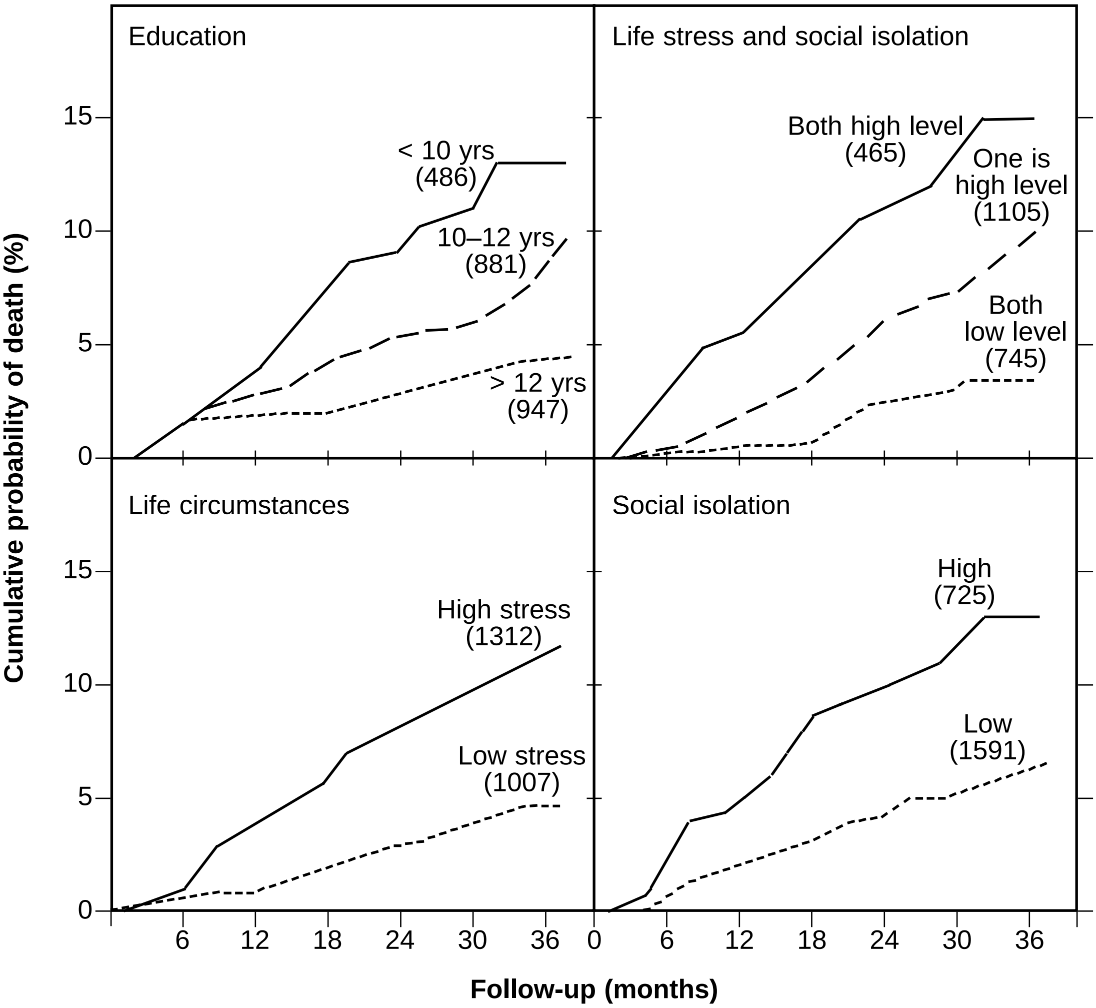

Figure 1 The effect of social factors on the survival of men after a first MI.

frequently than to controls. These studies corroborated earlier findings (e.g., Parkes et al., 1969) that observed a 40% increase in cardiovascular mortality among widowers. However, these and related studies are subject to the criticism of biased recall of stressful events by relatives or friends of sudden death victims, who served as informants in these studies.

The occurrence of disasters and personal traumas has also been related to an increase in cardiac deaths and increased rates of MI. For example, during the Iraqi missile attacks on Israel during the initial days of the 1991 Gulf War, there was a significant increase in fatal and nonfatal cardiac events among the population living close to Tel Aviv (Meisel et al., 1991). Cases of acute MI treated in the intensive care unit of a Tel Aviv medical center were elevated during the week following the missile attacks (January 17–25, 1991), compared to the week prior to the attacks and to an index period consisting of the same week a year earlier (Figure 2). In addition, there was an increase in the sudden death rate during January 1991 compared with the same period a year earlier. Mortality statistics among the entire Israeli population during period showed that on the day of the first missile strike, excess mortality observed was greater among women than among men (Kark, Goldman, & Epstein, 1995). This increase in cardiovascular mortality occurred largely in the Tel Aviv area where the missile attacks occurred.

Mittleman et al. (1995) used a novel epidemiologic methodology that compares each patient's activities before MI with his or her usual levels of activities, in order to assess the immediate physical and mental triggers of onset of heart attack. In a study of patients interviewed a median of four days after MI, 2.4% reported episodes of anger within the two hours prior to onset of MI. Following these

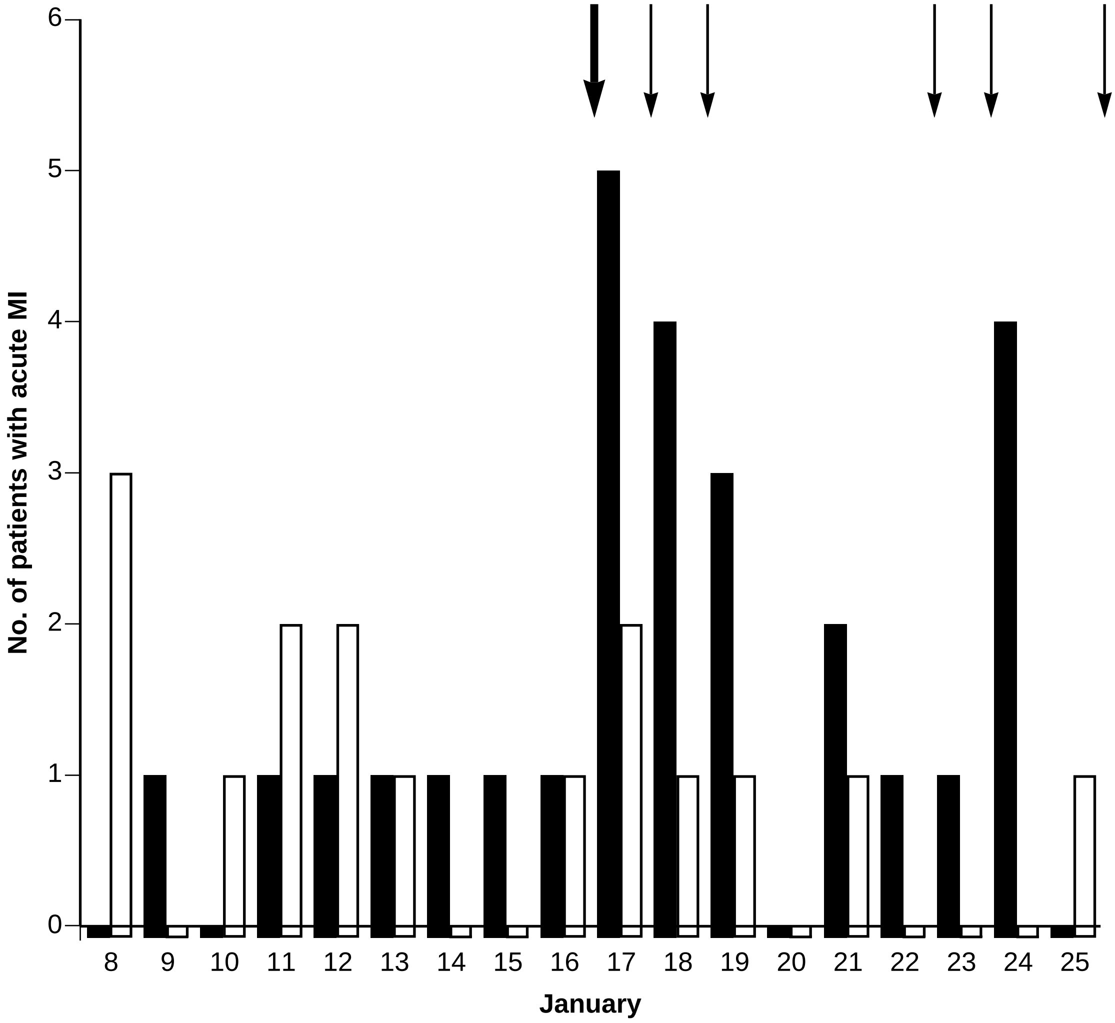

Figure 2 Daily incidence of acute myocardial infarction (MI) observed in Tel Aviv-area coronary care units during the period of the Iraqi missile attacks (January 8–25, 1991) (filled colomns) compared with the same time period in 1990 (unfilled columns) (large arrow = beginning of the Gulf War; small arrows = missile attacks on Israel).

episodes of anger, the risk of MI following episodes of anger was more than twice as high (Figure 3).

8.08.2.2.6 Stress and cardiac ischemia

It is possible to monitor continuously the electrocardiogram (ECG) during daily life in patients with coronary disease in order to detect the occurrence of cardiac ischemia—the inadequate supply of blood to the heart that is a clinical manifestation of coronary heart disease activity. In several studies, mental and physical activities of cardiac patients undergoing ambulatory monitoring of the ECG have been assessed using structured diaries. During daily life, ischemia in these patients is accompanied by strenuous physical exercise relatively infrequently, the majority of such episodes occurring during normal daily activities (e.g., Schang & Pepine, 1977). Mental stress and emotion also appear to be potent triggers of daily life ischemia in coronary patients (Barry et al., 1988; Gabbay et al., 1996). For example, Gabbay et al. (1996) assessed the effects of specific emotions on ischemia by using a more sophisticated diary in patients undergoing ambulatory monitoring. Ischemia occurred most frequently during moderate intensity physical and mental activities. Patients spent the largest proportion of time engaged in low intensity physical and mental activities, but the likelihood of ischemia was greatest during intense physical activities and during stressful mental activities. The percentage of time in ischemia (5% of the time) was approximately equivalent for high-intensity physical (Figure 4(a)) and high-intensity mental activities (Figure 4(b)), compared with 0.2% of

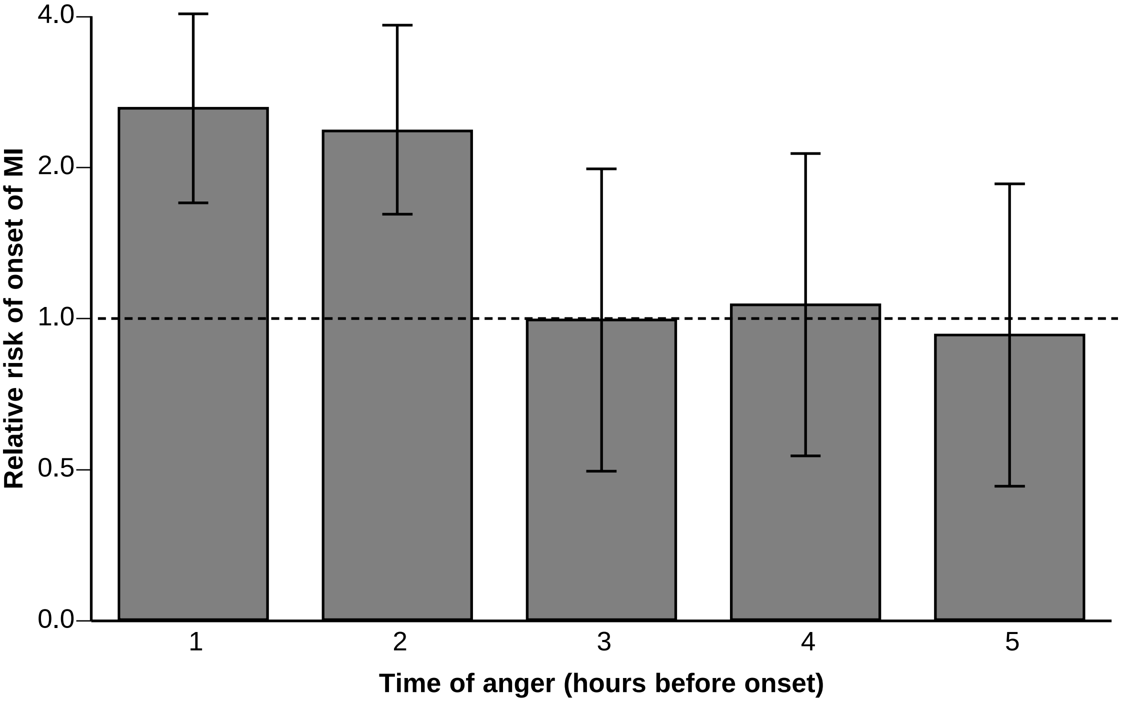

Figure 3 Episodes of anger and the relative risk of MI.

the time when patients were engaged in low intensity activities (see Figure 5).

Strenuous physical activity (e.g., walking) and the experience of intense anger were potent ischemic triggers.

Studies have also utilized a variety of sensitive means to evaluate the effects of mental stress on myocardial ischemia in the laboratory, enabling investigators to infer causal relationships between stress and ischemia and to assess physiologic and psychological mechanisms responsible for these effects. These studies have revealed that laboratory mental stressors can provoke ischemia in a substantial subset of patients with coronary artery disease (e.g., Blumenthal et al., 1995; Deanfield et al., 1984; Goldberg et al., 1996; Gottdiener et al., 1994; Rozanski et al., 1988). Ironson et al. (1992) further reported that an anger-inducing stressor was a particularly potent psychological stressor in its ability to trigger ischemia.

8.08.2.2.7 Summary

In summary, there is evidence that both acute and chronic stress may either promote the development of or trigger CHD events. Key aspects of chronic stress that are potential risk factors for cardiac events and that can potentially be modified include job strain and, particularly among cardiac patients, low social support and lack of economic resources. Converging evidence also suggests that anger is an emotion that is a potent trigger of cardiac events such as ischemia or infarction. In addition to these environmental and social factors, several individual behavioral traits are reviewed below that have been suggested as CHD risk factors.

8.08.2.3 Individual Characteristics as CHD Risk Factors

8.08.2.3.1 Type A behavior: current status

The concept of a coronary-prone personality—a set of emotions, behaviors, and personality attributes that characterize people who are likely to develop CHD—has origins as far back as a century or more ago. During the 1950s, cardiologists Friedman and Rosenman (1959) described the type A behavior pattern (TABP), characterized by excessive competitive drive, impatience, hostility, and vigorous speech characteristics. A contrasting behavior pattern, called type B, consists of the relative lack of these characteristics and a more easygoing style of coping. Friedman and Rosenman (1974) developed a structured interview to measure type A behavior based on the presence of observable behaviors such as speech characteristics and the manner in which subjects respond to questions, rather than relying on whether subjects describe themselves as impatient,

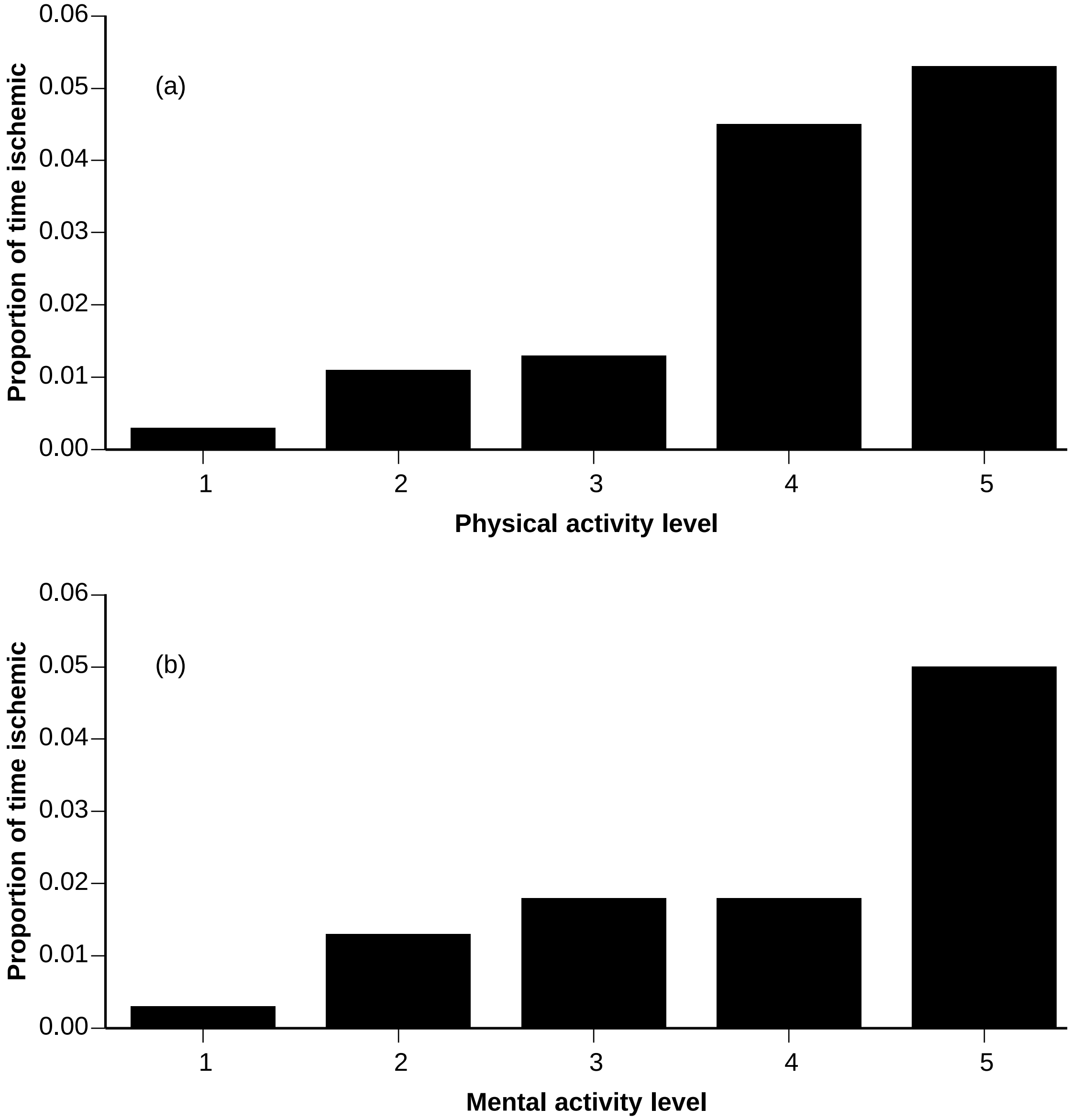

Figure 4 The percentage of time in ischemia at different physical and mental activity levels (adjusted for the time spent in the activity level).

competitive, etc. In contrast to the structured interview, several questionnaire measures developed to assess type A behavior, such as the Jenkins Activity Survey (JAS; Jenkins, Zyzanski, & Rosenman, 1971), relied solely on subjects" self-reports of their own behavior (Matthews & Haynes, 1986), and have generally shown a weaker relationship to CHD than interview-derived type A measures.

Most studies conducted in the 1960s and 1970s to examine the relationship of type A behavior to heart disease revealed a positive correlation between TABP and risk of CHD in men and women, which was both comparable to and independent of the effects of risk factors such as smoking and hypertension. For example, evidence supportive of type A as a risk factor included two major studies of initially healthy individuals. The Western Collaborative Group Study (WCGS) examined over 3000 men for 8.5 years. Those men assessed as type A by interview and questionnaire at study intake were more likely to have developed heart disease over the course of the study compared to type B men (Jenkins, Rosenman, & Zyzanski, 1974; Rosenman et al., 1975). In the Framingham Heart Study, type A behavior was a predictor of CHD among men in white collar occupations and in women working outside the home (Haynes, Feinleib, & Kannel, 1980).

However, since the 1980s, most studies have failed to find a relationship between type A behavior and coronary disease (see Matthews & Haynes, 1986). For example, the Multiple Risk Factor Intervention Trial (MRFIT) was a study conducted to determine whether interventions to modify coronary risk factors such as smoking, high cholesterol levels, and high blood pressure in high-risk men would lessen the likelihood of coronary disease in these individuals. Measures of type A behavior were obtained for over 3000 MRFIT subjects, who were followed for seven years. Type A behavior, as assessed by the structured interview and by the JAS in this high-risk group was unrelated to incidence of a first heart attack (Shekelle, Holley, & Neaton, 1985). This study, then, failed to support a link between type A behavior and heart disease. In addition, there were research reports indicating that, after a heart attack, type B patients (and not type A patients) were more likely to die (Ragland & Brand, 1988). Although this finding could be explained in terms of the healthier type A patients being the ones who initially survived their first heart attack, this result clearly presented problems for the notion of type A behavior as coronary-prone behavior.

The reasons for these inconsistent findings for type A behavior are not entirely clear, and some (e.g., Matthews & Haynes, 1986) have suggested that type A behavior might not be a risk factor for certain high-risk groups (as were tested in the MRFIT study). Others have suggested that type A behavior is more likely to be a risk factor in younger, as opposed to older, subgroups (Williams et al., 1988). Nevertheless, it appears that certain components of type A behavior—particularly anger, hostility, and vigorous speech characteristics—have remained correlated with coronary disease even in studies where overall, or "global," type A behavior has not been related to CHD.

8.08.2.3.2 Anger and hostility

As noted above, type A behavior consists of several behaviors, including competitiveness, time urgency, and hostility, yet it is possible that not all these behaviors contribute equally to coronary risk. What have consistently emerged as correlates of CHD in these studies are characteristics relating to hostility, anger, and certain speech characteristics derived from the structured interview, as well as the characteristic of not expressing anger or irritation or "anger-in." For example, a reanalysis of data from the WCGS described above showed that "potential-for-hostility," vigorous speech, and reports of frequent anger and irritation were the strongest predictors of CHD (Matthews, Glass, Rosenman, & Bortner, 1977). Even in the MRFIT study, which was not able to relate type A behavior to CHD, hostility characteristics in MRFIT subjects were associated with increased CHD risk (Dembroski, MacDougall, Costa, & Grandits, 1989).

The Cook and Medley Hostility Inventory (Cook & Medley, 1954), a scale derived from the Minnesota Multiphasic Personality Inventory (MMPI), has been shown in two studies to be related to occurrence of coronary disease. This scale appears to measure attitudes such as cynicism and mistrust of others (Barefoot & Lipkus, 1994). In one study involving a 25-year follow-up of physicians who completed the MMPI while in medical school, high Cook–Medley scores measured in college students predicted incidence of CHD as well as mortality from all causes, and the relationship was independent of the individual effects of smoking, age, and presence of high blood pressure (Barefoot, Dahlstrom, & Williams, 1983) (Figure 5). There is also evidence that low hostility scores are associated with decreased death rates during a subsequent 20-year follow-up of nearly 1900 participants in the Western Electric Study (Shekelle, Gale, Ostfeld, & Paul, 1983). In other studies it was shown that traits of hostility (e.g. assessed in terms of behaviors and attitudes indicative of hostility derived from the type A behavior structured interview) were related to the development of CHD in initially healthy men and in high-risk participants in the MRFIT study (for a review see Helmers, Posluszny, & Krantz, 1994). Subsequent research has further suggested that hostility scores on the Cook–Medley (1954) scale are higher in low socioeconomic status groups, higher in men and non-whites in the USA, and also positively related to the prevalence of smoking (Siegler, 1994). Thus, it is possible to hypothesize that hostility may account for some of the socioeconomic and gender differences in death rates from cardiovascular diseases (Stoney & Engbretson, 1994).

8.08.2.3.3 Clinical depression as a CHD risk factor

In addition to hostility, the presence of other psychological traits can confer increased risk of poor health outcomes among CHD patients, and hence are important to consider as elements of effective cardiac patient treatment and rehabilitation. Signs and symptoms of clinical depression, including fatigue, exhaustion and dysphoric mood, are prevalent among patients with CHD. Moreover, studies have reported that nearly 1 in 5 cardiac patients can be diagnosed as having the signs and symptoms of

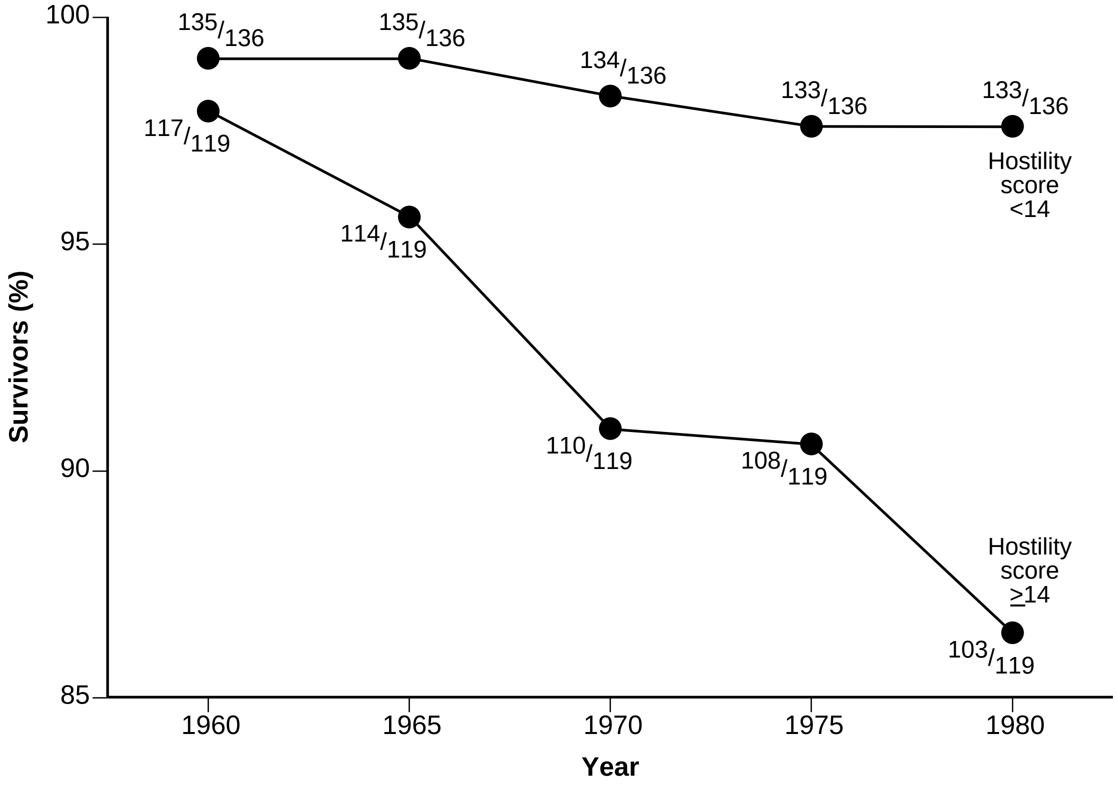

Figure 5 The relationship between hostility score and death rate in the Western Electric Study.

clinical depression (Carney et al., 1987; Frasure-Smith, Lesperance, & Talajic, 1993). However, despite its high prevalence, depression among post-MI patients is often untreated.

The presence of a major depressive episode in CHD patients is associated with poor psychosocial rehabilitation and increased medical morbidity (Carney, Freedland, Rich, & Jaffe, 1995). At least four studies have compared the clinical course of depressed versus nondepressed cardiac patients and noted an increase of events and lower mortality rates associated with depression. Diagnosis of major depressive disorder at the time of coronary angiography was the best predictor of a significant cardiac event (death, reinfarction, bypass, angioplasty) during a one-year follow-up period (Carney et al., 1987). Schleifer, Keller, Bond, Cohen, and Stein (1989) reported a trend toward a higher rate of rehospitalization and reinfarction in depressed patients. Kennedy, Hofer, Cohen, Schindledecker, and Fisher (1987) reported a significant correlation of cardiac arrhythmias, depression, and mortality. Frasure-Smith et al. (1993) prospectively followed 222 post-MI patients and observed that a diagnosis of major depression was a strong predictor of mortality in the 6 months following hospital discharge. In addition, two more recent studies of initially healthy populations have suggested that the presence of a major depressive episode (Pratt et al., 1996) or worsening of depressive symptoms (Wassertheil-Smoller et al., 1996) are prospectively predictive of the subequent development of cardiac events. A parallel set of studies conducted by Appels and colleagues (e.g., Appels, 1990; Kop, Appels, de Leon, de Swart, & Var, 1994) suggest that symptoms of fatigue or exhaustion—even in the apparent absence of clinical symptoms of dysphoria or depressive affect—are predictive of the subsequent development or worsening of cardiovascular symptoms and events. This syndrome of fatigue symptoms has been termed "vital exhaustion" (Appels, 1990). Moreover, evidence suggests that the predictive value of symptoms of fatigue or exhaustion cannot be explained by the effects of illness on mood or energy level (Kop et al., 1994).

To explain the link between depression and/or vital exhaustion, low social support, and mortality in coronary patients, studies have observed reduced compliance with medical therapeutic and exercise regimens in depressed cardiac patients (Carney et al., 1995), and a supportive social network is thought to be necessary to facilitate adequate compliance with medical regimens and reduction of cardiac risk factors (Amick & Ockene, 1994). It has also been suggested that increased susceptibility to cardiac arrhythmias and sudden cardiac death is a mechanism linking major depression to sudden cardiac death in coronary patients (Carney et al., 1995). In addition, recent studies have suggested that the presence of clinical depression can be associated with neural and neuroendocrine responses that predispose to sudden death, and changes in blood-clotting processes that may predispose to coronary thrombosis and to MI (Frasure-Smith, 1991). Research on mechanisms linking depression, social support, and/or lack of economic resources to increased cardiac morbidity and mortality is in its early stages, and this area promises to be a fruitful area for exploration.

8.08.2.3.4 Stress reactivity

The adverse effects of stress on cardiac pathology have been attributed to the neural, endocrine, and cardiovascular effects of stress (Krantz et al., 1996). It has long been known that there are wide individual differences that exist in physiologic reactions to stress, and research suggests that physiological responses (reactivity) to emotional stress may be involved in the development of CHD and/or high blood pressure (see Krantz & Manuck, 1984; Manuck, 1994). The measurement of "reactivity" involves the assessment of cardiovascular and/or hormonal changes in response to stress—not just resting levels of physiologic variables. There are wide individual differences in the magnitude of physiological responses shown during such a task, with some people (so-called "hot reactors") showing sizable increases to the challenging task, and others showing little or no increases. For example, there is some evidence that behaviors evidenced by hostile type A persons are accompanied by the same kinds of cardiovascular and neuroendocrine responses thought to link psychosocial stress to CHD (Contrada & Krantz, 1988; Krantz & Durel, 1983; Matthews, 1982).

Research has examined the possibility that excessive reactivity to stress may itself be a risk factor for coronary disease. In one study of initially healthy men followed for 23 years (Keys et al., 1971), the magnitude of their diastolic blood pressure reactions to a cold pressor test (which involves immersing the hand in cold water) predicted later heart disease. In fact, this physiologic response was a stronger predictor than many of the standard risk factors assessed in the study. However, a later study (Coresh, Klag, Mead, Liang, & Whelton, 1992) failed to replicate these findings.

Additional evidence of a relationship between cardiovascular reactions to stress and coronary disease was obtained in the Bowman–Gray monkeys described earlier. In studies of male and female cynomolgus monkeys fed on cholesterol-rich diet, animals were exposed to a standard laboratory stress (threat of capture) that produced large heart-rate elevations (Manuck, Kaplan, & Clarkson, 1983; Manuck et al., 1986). Animals were categorized as either high or low heart-rate reactive monkeys. At the end of the study, high heart-rate reactors had nearly twice the amount of coronary atherosclerosis as did low heart-rate reactors. Because of these animal and human studies, several behavioral methods for reducing reactivity have been studied. Among the behavioral techniques used for this purpose are cognitive techniques, relaxation training, biofeedback, and aerobic exercise (Jacob & Chesney, 1986; Matthews et al., 1986).

8.08.2.4 Treatment of CHD

Although the field of cardiac rehabilitation is comparatively young, it has made considerable inroads in the clinical management of coronary patients. Standard rehabilitation programs involve a combination of drug treatment, exercise, and dietary interventions (Blumenthal & Emery, 1988). The current status of interventions in this field focusing on medical and/or exercise interventions, has been documented in several reviews (Blumenthal & Emery, 1988; Lau et al., 1992; Oldridge, Guyatt, Fischer, & Rimm, 1988). Recently, more treatment plans for CHD have included psychological interventions, such as group and individual cognitive therapy, relaxation therapy, and meditation techniques. Because behavioral treatments for CHD must be implemented in the context of current medical practice, a brief overview of the medical and surgical treatment options for coronary disease is presented before describing psychosocial interventions in this area.

8.08.2.4.1 Medical and surgical treatment options

Medical and surgical treatment options for cardiac patients range from minimally invasive medical and drug treatments to coronary bypass surgery. Among the many pharmacological agents that have been proven useful for the treatment of CHD patients are: anti-ischemic agents, such as beta-blockers and calcium channel drugs, which reduce the heart's demand for blood and help the heart function more efficiently with a limited supply; nitrates, which are useful in the treatment of ischemia and anginal pain; various lipid-lowering drugs that reduce harmful cholesterol levels; aspirin and other drugs that reduce the clotting tendency of blood platelets; and agents that lower blood pressure. Other procedures, such as percutaneous transluminal angiography (PTCA) and coronary artery bypass grafting (CABG) are also commonly used in treatment. (PTCA and CABG are together referred to as revascularization procedures, as they aim to restore blood flow in areas where flow was previously restricted.) During PTCA, atherosclerotic points in selected arteries are widened by means of a balloon-inflation procedure. CABG involves the surgical bypassing of blocked sections of coronary arteries with sections of a saphenous vein, usually from the leg, or with portions of an artery from another part of the body. Studies indicate that CABG is effective in reducing anginal pain and improving functional status, but has only been shown to lengthen survival in specific subsets of CHD patients (Coronary Artery Surgery Study [CASS] Principal Investigators, 1984).

8.08.2.4.2 Exercise and behavioral components of cardiac rehabilitation

There are numerous behavioral components of successful comprehensive cardiac rehabilitation programs (e.g., Blumenthal & Emery, 1988; Krantz & Blumenthal, 1987). These include an aerobic exercise conditioning program, smoking cessation, weight loss, and maintenance of optimal weight, and modification of diet to emphasize foods low in saturated fat and cholesterol. Each of these interventions often involve significant and demanding behavioral changes, and issues of promoting behavioral change, ensuring compliance, and long-term maintenance of behavioral modifications are important in ensuring therapeutic effectiveness.

Exercise training is a core element of contemporary cardiac rehabilitation programs (for reviews see Blumenthal & Emery, 1988; Blumenthal & Wei, 1993). Physical inactivity is a significant and independent risk factor for CHD, and regular physical activity in the form of aerobic-type exercises increases functional capacity and plays an important role in both primary and secondary prevention. Active individuals have better cardiovascular risk factor profiles relative to their sedentary counterparts, including lower blood pressure, higher HDL cholesterol, lower LDL cholesterol, lower triglycerides, and very low density lipoprotein cholesterol levels. Physically fit individuals, in comparison to unfit persons, also have attenuated hemodynamic and neuroendocrine responses to behavioral stressors (Blumenthal & Wei, 1993).

Exercise improves functional capacity in most cardiac patients, including patients with limited exercise tolerance, and those who engage in regular physical activity may be better able to perform activities of daily living and enjoy many aspects of life. However, despite these positive effects, the precise effects of exercise training in prolonging life or reducing recurrence rates in post-MI patients is still the subject of some debate. Recent meta-analyses of controlled studies in rehabilitation have demonstrated significant reductions in mortality for patients assigned to rehabilitation in comparison with those assigned to control groups (Fletcher, 1996; Pate et al., 1995). However, these results have been difficult to demonstrate in individual studies, perhaps because large sample sizes are needed to reveal these effects.

Interventions to increase participation in physical activity generally indicate that, as with many lifestyle interventions, incentives and social support can increase exercise patterns in the short term, but that people have difficulty maintaining their exercise programs over the course of time (Blumenthal & Emery, 1988; Dubbert, 1995). Exercise maintenance is improved by integrating physical activity and exercise into everyday lifestyle. Because recent evidence suggests that stress management and psychosocial treatments have important beneficial effects on morbidity and on quality of life when implemented in conjunction with standard medical interventions, these interventions are reviewed in some detail in the following section.

8.08.2.5 Psychosocial Treatment Approaches for the Cardiac Patient

As noted previously, evidence suggests that psychosocial interventions may reduce morbidity and mortality, psychological distress, and some biological risk factors when used in addition to standard treatment regimens among cardiac patients. More specifically, Linden, Stossel and Maurice (1996) recently performed a meta-analysis of 23 controlled trials that evaluated the impact of psychosocial treatments directed specifically at reducing distress-related factors among cardiac patients. Results of this analysis indicated that patients receiving psychosocial treatments (e.g., stress management, individual or group cognitive therapy, and meditation) showed greater reductions, compared with control conditions, not only in psychological distress but also in blood pressure, heart rate, and serum cholesterol levels. Morbidity and mortality data, available in about half of these studies, indicated that patients not receiving psychosocial treatments showed greater mortality and cardiac recurrence rates during the first two years of follow-up (Table 1).

8.08.2.5.1 Modifying hostility and type A behavior

A variety of clinical intervention studies have attempted to decrease type A behavior either in persons with elevated levels of other CHD risk factors or in samples of coronary patients. Most of these early studies have demonstrated that elements of type A behavior can be decreased to some extent in subjects who are motivated to change (Allan & Scheidt, 1996; Suinn, 1982). Accompanying changes in type A behavior, some studies also measured changes in traditional CHD risk factors such as serum cholesterol levels or blood pressure. The findings from these studies are inconsistent (Allan & Scheidt, 1996; Suinn, 1982). These results are not surprising considering that the therapeutic regimens are not typically directed at decreasing risk factor levels other than type A behavior.

The most important and ambitious type A behavior intervention study done to date is the Recurrent Coronary Prevention Project (RCPP; Friedman et al., 1986). The major purpose of this project was to determine whether the constellation of type A behaviors—focusing on anger, impatience, aggressiveness, and irritability—could be modified in a large group of heart attack patients, and whether such behavior changes would lower the recurrence of heart attacks and deaths from CHD. Beginning in 1979, over 1000 patients were recruited for a five-year intervention study.

Patients were assigned to one of three groups: a cardiology counseling treatment group, a combined cardiology counseling and type A behavior modification group, or a no-treatment control group. The cardiology counseling included encouragement to comply with dietary, exercise, and drug regimens prescribed by the participants' personal physicians, education about CHD and its treatment, and counseling about psychological problems (other than type A behaviors) associated with the postcoronary experience. Type A behavior counseling included drills to change specific type A behaviors, focused discussions on beliefs and values underlying type A behavior, rearrangements of home and work demands, and relaxation training to decrease physiologic arousal.

In the final results of the study after 4.5 years, the rate of heart attack recurrence for the type A behavioral counseling group was significantly lower than for the cardiology counseling and control groups (Friedman et al., 1986). These results demonstrated, within a controlled experimental design, that an intervention directed at altering type A behavior reduces coronary disease recurrence in post-heart-attack patients. However, in light of the recent negative evidence regarding associations of type A behavior with CHD, it is possible that the beneficial effects of the RCPP study resulted from more general salutary effects of the interventions in reducing distress and increasing social support among patients.

8.08.2.5.2 Interventions to increase social support and reduce life stress

Another, perhaps more cost-effective, approach to preventing stress-related morbidity and mortality in cardiac patients was utilized by the Ischemic Heart Disease Life Stress Monitoring Program (Frasure-Smith & Prince, 1987,

Table 1 Benefit of psychosocial interventions for reductions in mortality and morbidity.

End-point	*No. of studies*	*No. of patients (control/experimental group)*	*Weighted log OR**	χ^2	*P*	*95% CI**
Morbidity, *y*:						
$\leqslant 2$	8	390/462	1.84	5.80	0.02	1.12–2.99
>2	3	298/297	1.64	5.06	0.02	1.06–2.54
Mortality, *y*:						
$\leqslant 2$	10	530/620	1.70	5.46	0.02	1.09–2.64
>2	3	298/297	1.35	2.24	0.13	0.83–1.53

OR, odds ratio; CI, confidence interval.
* Fully randomized, controlled trials only

1989), conducted in Montreal, Canada. This study made clinical use of the evidence that periods of increased life stress may precede recurrences of heart attack. Patients who had experienced a heart attack (or MI) were assigned to either a treatment group ($n = 229$) or to a control group ($n = 224$). Patients assigned to the control group did not receive life stress monitoring and intervention, but instead received routine medical follow-up care. In the treatment group, patients were contacted by phone on a monthly basis and asked about 20 symptoms of distress, such as insomnia, feelings of depression, and inability to concentrate. When stress levels exceeded a critical level (more than 4 of 20 possible symptoms), a project nurse made a home visit to assess the causes of increased distress, and attempted to help the patient with whatever life problems seemed to be producing the distress. Over the year of the program, about half of patients in the treatment group had high stress levels requiring intervention. On average, study nurses provided each high-stress patient with 5–6 hours of home nursing contact, including teaching patients about coronary disease, counseling, providing social and emotional support, and referring patients to cardiologists or other health professionals as needed.

The study hypothesized that by reducing stress the program would result in a reduction in cardiac mortality and in cardiac recurrences. Results showed that for the year of the project cardiac deaths were reduced by 50%, and that the reduced death rate persisted for six months beyond the end of the program (Figure 6). In addition, although there were similar rates of MI recurrences in the two groups, during the seven years following the program there were fewer MI recurrences among patients in the treatment group.

The most obvious explanation for the outcomes obtained in the Ischemic Heart Disease Life Stress Monitoring Program is that the program provided patients with a sense of social and emotional support that helped ameliorate depression and feelings of helplessness and distress, thereby reducing physiological arousal and its adverse effects on the cardiovascular system. The special clinical features of the program, including selection for treatment based on stress scores and individualized tailoring of interventions to patients" needs, may have been responsible for the program's success.

8.08.2.5.3 The lifestyle heart trial

One of the most important intervention studies conducted to date is the one carried out by Ornish and colleagues (Ornish et al., 1990). These researchers conducted a controlled study of 28 patients with significant coronary artery disease documented by coronary angiography (a procedure used to assess the severity of blockages in coronary arteries). Patients were randomly assigned to either an experimental group or to a usual care control group. Experimental group patients were given a lifestyle-modification program consisting of several components: (i) a very low fat vegetarian diet; (ii) stress management training and group support consisting of twice weekly yoga and meditation conducted in group sessions together with individual practice for one hour a day; (iii) a smoking cessation program; and (iv) a program of moderate levels of aerobic exercise. Patients assigned to the control group were not asked to make lifestyle changes other than those recommended in traditional cardiologic care. The intervention lasted for one year, and the extent of progression of coronary disease was assessed in all patients by comparing coronary angiograms obtained at the study onset and at one year.

The study results (Ornish et al., 1990) demonstrated that, for the experimental group, on average there was a slight reduction (regression) in the extent of stenosis (or blockage) in coronary arteries, whereas the controls showed a progression in severity of stenosis. Overall, 82% of patients in the experimental group had an average change toward regression of disease. Interestingly, the degree of measured regression of disease was related to the overall extent of adherence to the intervention recommendations among the study subjects, with the most compliant subjects showing the most improvement in disease status and the least compliant patients showing the least change (Figure 7).

Blumenthal et al. (1997) recently assessed the independent effects of stress management and exercise training on the prognosis of 107 cardiac patients with CHD and evidence of myocardial ischemia. In the Myocardial Ischemia Intervention Trial, patients were randomly assigned to a four-month program of stress management or exercise training. An additional control group of patients (those patients who lived too far from the intervention center to participate in the intensive and frequent treatment sessions) constituted a nonrandom usual-care group. To assess possible mechanisms associated with the effects of interventions, patients underwent mental stress and exercise testing for ischemia prior to and after the interventions. The results of the study indicated that, compared to controls, patients in the stress management group showed a lower risk of cardiac events over a

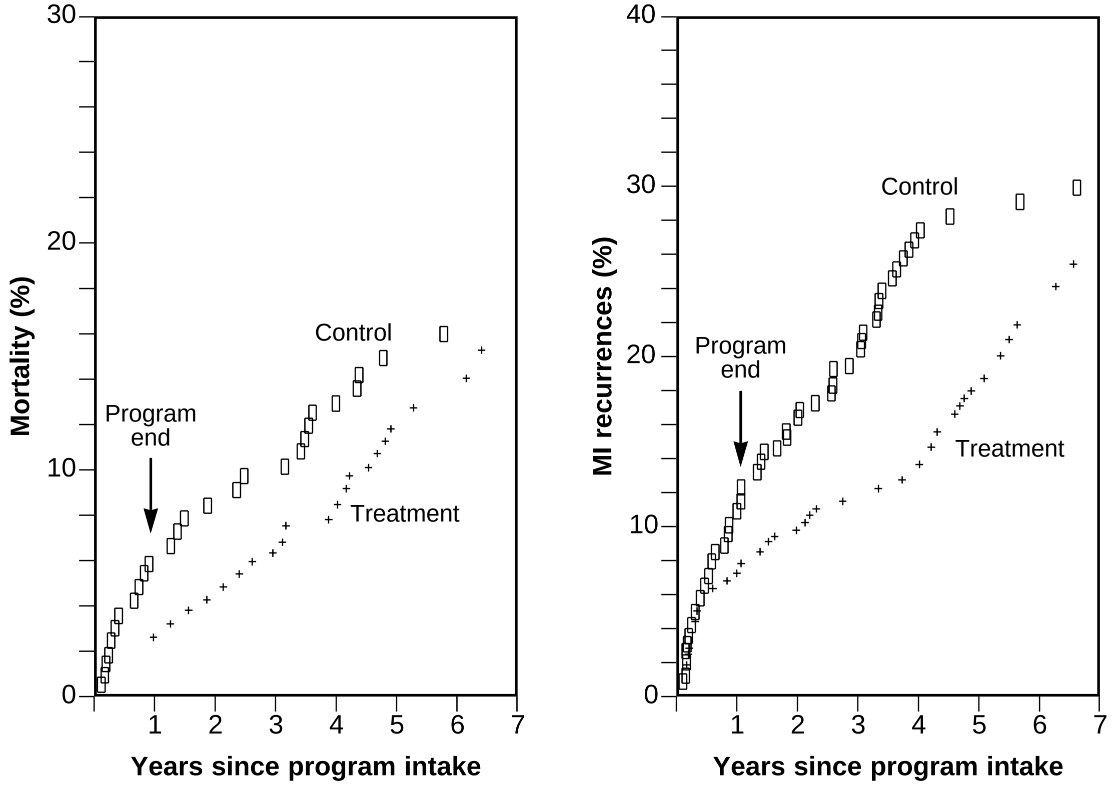

Figure 6 Long-term outcomes in the Ischemic Heart Disease Stress Monitoring Program.

mean follow-up period of 38 months. The risk of cardiac events in the exercise training group was lower than that of controls, but the effect did not reach statistically significant levels. Other beneficial effects of the stress management intervention included reduced occurrence of myocardial ischemia, lowered serum cholesterol levels, and reduced self-reported distress. One shortcoming of this study, however, was the failure to randomize patients to the usual-care control group. However, analyses failed to uncover any medical, demographic, or socioeconomic differences between control and treatment groups that could have accounted for the results obtained.

8.08.2.5.4 Summary

In summary, the identification of psychosocial risk factors for coronary disease has led to several promising behavioral and psychosocial interventions to aid in the treatment and prevention of coronary disease in high-risk individuals. Four such behavioral and psychosocial interventions with coronary patients were described: cognitive-behavioral interventions directed at lessening hostility and type A behavior (RCPP); a tailored social support and counseling intervention delivered by a nurse directed at reducing life stress (Ischemic Heart Disease Life Stress Monitoring Program); a lifestyle-modification program consisting of a very low fat vegetarian diet, a stress management training with group support and practiced individually, stopping smoking, and moderate levels of aerobic exercise (Lifestyle Heart Trial); and a controlled study of the effects of stress management interventions on cardiac ischemia and cardiac morbidity (Myocardial Ischemia Intervention Trial). As noted by Linden et al. (1996), the aggregate data provide a basis for the inclusion of psychosocial treatment components in cardiac rehabilitation. However, given the diversity of treatment approaches employed, these findings also suggest an important need to identify the most specific, most effective types of psychosocial interventions for particular patients.

8.08.3 ESSENTIAL HYPERTENSION

The term "essential hypertension" refers to an asymptomatic condition in which the blood pressure is chronically elevated, and for which no single cause can be identified. According to current diagnostic standards (Joint National

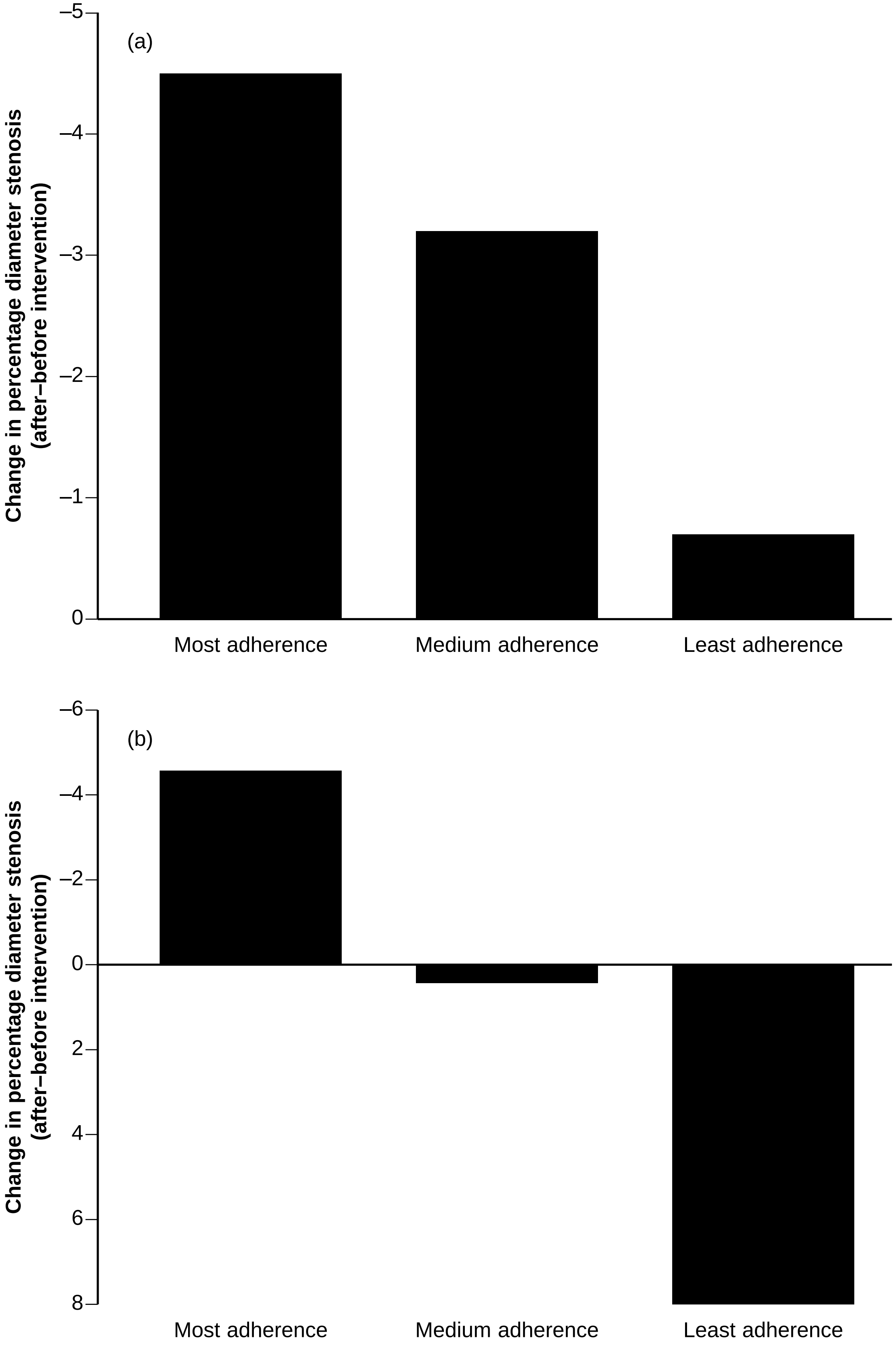

Figure 7 Disease regression measured in the study by Ornish et al. (1990): (a) experimental study group; (b) whole study group.

Committee on Detection, Evaluation, and Treatment of High Blood Pressure [JNC], 1993), hypertension is defined as a diastolic blood pressure consistently at or above 90 mmHg and/or a systolic blood pressure at or above 140 mmHg. Recent data further identify the optimal blood pressure as at or below 120/85 mmHg. The prevalence of essential hypertension in Western societies is high; for example, this condition is estimated to be present in 15% of the US population. Hypertension is a serious and potentially lethal disease, and untreated hypertension increases the risk of stroke, heart attack, kidney, and vascular disease. Progress in the control of hypertension over the past 30 years has been enormous, and although it is a chronic disease its control is now possible through a variety of pharmacologic as well as nonpharmacologic behavioral means.

Factors associated with risk of hypertension, include age, family history of hypertension, "borderline" high blood pressure, dietary intake of salt, and obesity. Like CHD, there are numerous social, environmental, and cultural factors that interact with genetic background in predisposing individuals to hypertension, and considerable attention has been paid to studying the role of stress and personality factors in the development of this disorder. In addition, there has been considerable attention devoted to cognitive and behavioral treatment approaches for essential hypertension, and these approaches have met with mixed results at best. It is likely that many findings in the area of the role of behavior in the development and treatment of hypertension are contradictory because of the complex interactions between behavioral, physiological, and genetic factors, and also because essential hypertension is probably not a single homogeneous disease. Rather, blood pressure is thought to progress over a period of years from moderately elevated or "borderline" levels to more appreciably elevated levels, the latter being called "established hypertension." The phenomenon of "white coat hypertension" has also been identified, in which individuals display elevated blood pressure when clinically examined in the doctors" office, but may display normal blood pressure when assessed in the patient's usual environment outside of the clinic (Pickering, James & Boddie, 1988; Siegel, Blomenthal, & Divine, 1990). It it likely that white coat hypertension results in an increase in "false-positive" diagnosis and treatment of high blood pressure, and the clinical significance of white coat hypertension remains a topic of current study (Dubbert, 1995).

Several different pathogenic mechanisms can bring about blood pressure elevations, and different physiologic and/or behavioral mechanisms are implicated at different stages of the disorder. For example, individuals with borderline and earlier stages of hypertension are commonly observed to have an elevated output of blood from the heart but little evidence of increased resistance to the flow of blood in the body's vasculature (Julius & Esler, 1975). This physiological pattern is consistent with an increased activation of the sympathetic nervous system, which is the body's initial reaction to psychological stress. Indeed, high levels of blood and tissue catecholamines (such as those produced by stress) have been found in some hypertensive humans and animals (Julius & Bassett, 1987; Julius & Esler, 1975). However, in older individuals with more established hypertension, the heart's output of blood is normal or even decreased and resistance to flow of blood is elevated.

8.08.3.1 Genetic–Environment Interactions

Evidence from animal research and from studies of human twins indicates that genetic factors are important in the development of hypertension (Pickering & Sleight, 1977). In humans, however, it is likely that sustained hypertension is produced by an interaction of environmental and genetic factors. Population studies reveal a difference in the prevalence of hypertension among various social and cultural groups, a difference that cannot be accounted for by genetic factors alone (Henry & Cassel, 1969). For example, even though African-Americans in the USA experience more hypertension than other groups, the prevalence of this disease is higher among poor than middle class black Americans (Harburg et al., 1973). Animal research similarly reveals examples where environmental factors, such as dietary salt intake or stress, can lead to sustained blood pressure elevations, but only in certain genetic strains (Dahl, Heine, & Tassinari, 1962).

8.08.3.2 Stress, Behavioral Factors, and Essential Hypertension

Sociocultural and psychological studies of humans, in conjunction with research on animals, have identified some factors related to behavior that might play a role in the development of hypertension. The increased risk for hypertension among blacks compared to whites in the USA and among persons of lower socioeconomic compared to higher socioeconomic status has been attributed to a variety of factors, including differences in dietary patterns, exercise habits, or the social and physical characteristics of the environments (including racism) (Krieger & Sidney, 1996) in which these individuals live and work (Julius & Bassett, 1987). Some studies have suggested that exposure to environments (e.g., urban high crime settings) that require sustained vigilance as well as recurrent mobilization of coping resources to ward off harm may raise blood pressure (Guttman & Benson, 1971; Henry & Cassel, 1969), and it is well documented that poorer people are overrepresented in such environments. For example, the role of stress in hypertension is suggested by a study of residents of Detroit (Harburg et al., 1973). Four areas were categorized as "high stress" or "low stress" based on socioeconomic status, crime rate, population density, residential mobility, and marital breakup rates. Blood pressure levels were highest among black high stress males, while white areas and black low stress areas did not differ in blood pressure levels. A recent report suggesting that racism may be linked to higher blood pressure in blacks (Krieger and Sidney, 1996) also supports this hypothesis. However, the simple notion that cultural change and social stress are causally linked to hypertension has been criticized as not being consistent with the aggregate data available (Syme & Torfs, 1978).

Research has also shown that people engaged in highly stressful occupations, such as workers in busy airport control towers, have more than four times the prevalence of hypertension when stress is high than individuals of similar age in other professions (Cobb & Rose, 1973). Current studies of urban bus drivers in several places in the world suggest that they also have higher rates of hypertension than demographically comparable groups of employed persons. The high rates of hypertension observed in these occupational groups is consistent with the job demand/control hypothesis advanced to explain the relationship between occupational stress and CHD (Karasek et al., 1982). Specifically, as with CHD, hypertension may occur more frequently in jobs that are demanding but in which there is little opportunity or flexibility to deal with these demands.

Experimental studies with animals further demonstrate that social or behavioral stress can produce hypertension in predisposed individuals. Classic studies by J. P. Henry and associates (Henry & Stephens, 1977) have shown that psychosocial stimuli can cause hypertension in mice. Other studies demonstrate that hypertension can be induced in animals exposed to environmental stressors such as fear or shock and experimentally produced conflict (Campbell & Henry, 1983). However, unlike the human condition, blood pressure tends to normalize when the stress is removed unless the animals are genetically predisposed to hypertension. Consistent with the notion that genetic–environment interactions are important in the development of hypertension, studies have demonstrated that strains of animals that are susceptible to hypertension are also likely to show stress-induced blood pressure elevations (Friedman & Iwai, 1976).

In humans, excessive stress and salt consumption might also interact in producing blood pressure elevations. Haythornthwaite, Pratley, and Anderson (1992) studied medical students during examination periods (high stress) and during vacation periods (low stress). Half the students were randomly assigned to consume salt tablets (sodium chloride) tablets or placebo tablets (a usual salt intake condition) during a two-week period of high or low stress. Results of the study indicated that high sodium intake during the high stress period was associated with greater elevations in resting systolic and mean arterial pressure than either the usual salt intake during the high stress period or the high salt intake during the low stress period. These results suggest that some forms of hypertension or blood pressure elevations may result from the interaction of psychosocial stress with periods of high salt intake.

8.08.3.2.1 Personality and essential hypertension

Beginning with the observation of early clinicians that patients with chronic hypertension showed certain personality features (e.g., Aymann, 1933), there has been a long-standing interest in personality correlates of essential hypertension (for reviews see Cottier, Perini, & Rauchfleisch, 1987; Johnson, Gentry, & Julius, 1992). Over time, several diverse personality characteristics have been identified or studied as being prevalent in hypertensives, with psychodynamic formulations driving early research in this area. Traits associated with hypertension in empirical studies have included, but are not limited to, suppressed hostility or anger (Dunbar, 1943; Harburg et al., 1973; Johnson et al., 1992), neuroticism and anxiety (Markovitz, Matthews, Kannel, Cobb, & D'Agonstino, 1993), and submissiveness (Elser et al., 1977). Although associations between these traits and hypertension have been widely reported, many existing studies suffer from methodological shortcomings, such as the use of selected or convenience samples, or the failure to distinguish in hypertensive patients characteristics (e.g., neuroticism) that may be prevalent in other diagnosed medical groups with chronic diseases. In addition, inconsistencies reported across studies in associations of certain traits with hypertension may have resulted not only from the aforementioned patient selection factors, but also from the fact that essential hypertension may not be a heterogeneous disease, progressing over a period of years from moderately elevated or "borderline" levels to more appreciably elevated levels (Weiner & Sapira, 1987).

Given the existence of a core of traits associated with blood pressure elevation, various investigators differ in their assumptions regarding the significance of personality trait differences in hypertension. Early psychosomatic investigators (e.g., Alexander, 1939; Dunbar, 1943) and more recent researchers (e.g., Jern, 1987) have postulated a causal role for personality traits in the development of hypertension. Other investigators (e.g., Esler et al., 1977; Weder & Julius, 1987) have suggested that personality traits associated with hypertension reflect excessive central and sympathetic nervous system arousal that may predispose to hypertension. Another perspective on personality traits in hypertensives has suggested that behavioral differences between hypertensive and normotensive populations that are reflected on personality tests result from the medical attention and labeling associated with the diagnosis of disease. For example, Irvine, Garner, Olmsted, and Logan (1989) noted that hypertensives who were aware of their diagnosis scored significantly higher than normotensives and than hypertensives unaware of their blood pressure elevation on neuroticism, trait and state anxiety, and self-reported type A behavior. However, an explanation in terms of labeling cannot account for the prospective association recently reported between anxiety in apparently normal individuals and the subsequent development of hypertension (Markovitz et al., 1993).

8.08.3.2.2 Salt intake

Salt intake is implicated as a factor causing or worsening essential hypertension, largely because excessive intake of sodium causes the kidneys to increase the volume of blood. However, studies indicate that high salt intake may only be related to high blood pressure levels in some cultures and population groups. For example, among people who live in underdeveloped tribal societies, sodium intake is often low, as is the prevalence of hypertension (Page, Danion, & Moellering, 1970). As individuals move or become acculturated to modern societies the prevalence of hypertension increases. This phenomenon has been discussed in terms of increasing stress with acculturation, but it has also been argued that acculturation is associated with increased salt intake and plays an equally important role in the development of hypertension (Page et al., 1970). Clearly, both the salt and stress explanations for the increased prevalence of hypertension in modern societies have some merit, and the research mentioned here suggests that salt consumption and exposure to stress might interact together to produce blood pressure elevations (Friedman & Iwai, 1976; Haythornthwaite et al., 1992).

8.08.3.2.3 Obesity

Obesity is another behavioral phenomenon that plays an important role in hypertension. There is an increased prevalence of hypertension in obese persons, although the precise reasons for this remain to be determined (Shapiro, 1983). Some have thought that it is merely that obese patients consume more sodium, but recent studies have demonstrated that weight loss without salt restriction can result in significant decreases in blood pressure (Reisen et al., 1978). For this reason, weight loss is an important behavioral method for managing high blood pressure.

8.08.3.3 Treatment of Essential Hypertension

Pharmacologic treatment by an increasingly broad variety of agents has revolutionized the treatment and control of hypertension. A landmark study in this area was the Hypertension Detection and Follow-up Program (HDFP, 1979). The HDFP, which is considered to have been a major breakthrough of modern medical care, studied a sample of 10 000 hypertensives. Half the patients were assigned to a special care group, involving a strict medication protocol, with the goal of normalizing blood pressure delivered via special community based clinics. A usual care group was returned to their usual sources of care available in the mid-1970s (the time of the study). Results indicated a significant decrease in overall mortality in the special care group. The beneficial effects of pharmacologic treatment were most evident in older patients and in black patients in the sample. This study convincingly demonstrated the benefits and efficacy of pharmacologic treatment of hypertension.

Because hypertension is a chronic, asymptomatic disease, to mitigate the life-long need to take medication, which often has side effects, a variety of nonpharmacologic behavioral interventions, including lifestyle interventions such as dietary modification, weight loss, exercise, and stress management, have been assessed as possible treatment approaches for essential hypertension. In this regard, a comprehensive set of behavioral changes, including weight loss, reduced salt intake, and exercise, are initial recommendations in the treatment of hypertension. If these approaches are not effective in normalizing blood pressure, medication is prescribed (JNC, 1993), and the aforementioned behavioral changes can lessen the amount of medication needed. In the following sections, we review the behavioral and lifestyle interventions used in the treatment of hypertension.

8.08.3.3.1 Weight loss and dietary salt restriction in hypertension management

Reduction of salt intake and weight loss have become important components of the nonpharmacological treatment of hypertension. Clinical trials completed in the last 10–15 years have shown that even relatively modest weight loss can have clinically significant effects on blood pressure over long periods, of up to 3–5 years in some studies (for a comprehensive recent review of the effects of dietary modification and weight loss interventions in hypertension see Dubbert, 1995). Dietary studies also indicate that decreasing sodium intake in the diet of hypertensives appears to have the salutary effect of lowering blood pressure, although these effects tend to be smaller than those obtained via weight loss (Dubbert, 1995). In accordance with the above discussion of genetic–environment interactions, we should note that high salt intake is correlated with increased prevalence of hypertension only in certain populations or groups, and that animal studies reveal that excessive salt intake results in sustained hypertension only in certain genetic strains (Dahl et al., 1962). Thus, we might expect salt restriction to be effective in lowering blood pressure only in certain susceptible individuals.

8.08.3.3.2 Effects of exercise training in hypertension

Dubbert (1995) has also reviewed the effects of exercise in the management of hypertension. Although several intervention studies have noted a beneficial effect of aerobic exercise training in hypertensives, the beneficial effects of exercise *per se* appear to be equivocal. Both positive and negative clinical trials of exercise have been reported in the literature (Kelley & McClellan, 1994; Martin, Dubbert, & Cushman, 1990) and, although meta-analyses (Kelley & McClellan, 1994) suggest significant beneficial effects of exercise training in hypertension, it is not clear to what extent these effects are due to the effects of weight loss. Thus, the role of exercise in the management of hypertension, independent of the effects of weight loss, remains an area of active research.

8.08.3.3.3 Effects of reducing alcohol intake

Evidence also suggests that a reduction in alcohol consumption is a behavioral change that can also have beneficial effects in the management of essential hypertension (Dubbert, 1995; MacMahon, 1987). In fact, because heavy drinking can both raise blood pressure and interfere with the efficacy of drug treatments for hypertension, the taking of a detailed drinking history and limiting daily alcohol intake to no more than two drinks per day has become standard recommended practice in hypertension treatment (JNC, 1993). The potentially beneficial long-term effects of reducing alcohol consumption among hypertensives who are heavy drinkers is currently being studied in a multicenter study in the USA (Cushman et al., 1994).

8.08.3.3.4 Effects of stress management, biofeedback, and other cognitive interventions in hypertension

Fostered by enthusiasm for exploring applications of biofeedback and stress management in medical conditions that occurred with the formation of the field of behavioral medicine, early studies suggested that techniques such as biofeedback and relaxation training could be used to modify the stress-induced components of high blood pressure, and thereby reduce blood pressure in hypertensive patients. Several studies have demonstrated that small but significant decreases in blood pressure (e.g., 15 mmHg systolic and 10 mmHg diastolic blood pressure) can be achieved in hypertensives after a series of training sessions with biofeedback or relaxation methods such as meditation, progressive relaxation, or yoga (for reviews see Dubbert, 1995; Eisenberg et al., 1993; Johnston et al., 1993). For example, one study that yielded positive effects was conducted by Patel and colleagues (Patel, Marmot, & Terry, 1981; Patel et al., 1985), who assessed the effects of eight weekly group sessions of training in breathing exercises, deep muscle relaxation, meditation, and a stress management program. Subjects were also instructed to practice relaxation, and meditation for 15–20 minutes at home daily and to try to relax during stressful daily activities. Compared with a control group, patients who received relaxation training showed decreases in both systolic and diastolic blood pressure over a follow-up period of up to four years after treatment. The magnitude of these changes were in the range of 7 mmHg in systolic and diastolic blood pressure. Early comparative studies of these various behavioral techniques indicated that no one of them is clearly superior to the others, with each producing modest declines in blood pressure (Shapiro et al., 1977). These reductions in blood pressure are achieved without any side effects or medical contraindications, thus enhancing the attractiveness and appeal of stress-reducing techniques as nonpharmacologic adjuncts to the treatment of hypertension.

Unfortunately, the results of more recent research and meta-analyses of an extensive body of controlled studies tend to reveal that most of the effects of stress management on hypertension appear to be relatively small and may be attributable to a mixture of nonspecific (i.e., placebo) effects and to habituation to repeated blood pressure measurements over the course of clinical trials (Dubbert, 1995; Eisenberg et al., 1993; Jacob, Chesney, Williams, & Ding 1991; Johnston et al., 1993). For example, Johnston et al. (1993) conducted a meticulously well-controlled study of the effects of stress management on resting and ambulatory blood pressure and on left ventricular mass (a measure of cardiac size that is a clinically important consequence of hypertension). Ninety-six mild hypertensives were studied, and underwent an extensive baseline evaluation to habituate them to the effects of blood pressure measurement. Subsequently, they were assigned to either 10 clinical sessions and practice at home of either stress management and relaxation training, or to a nonaerobic stretching condition control. Results indicated that patients' blood pressure fell during the habituation period, but neither resting nor ambulatory blood pressures were changed during the treatments. However, the magnitude of blood pressure increases produced by a stressful interview was reduced by stress managment. Thus, stress management appears not to be effective in lowering resting blood pressure among patients who are habituated to the measurement procedures. An interesting aspect of the study is that responses to acute stressors might be diminished by stress management. These general findings are corroborated by other recent studies of stress management conducted by Chesney, Black, Swan, & Ward (1987) at patients' worksites.

The only strong recent counterevidence for the efficacy of stress-management-like interventions comes from a recent set of studies by Schneider et al. (1995) and Alexander et al. (1996). These investigators report benefits of transcendental meditation on blood pressure in elderly African-American hypertensives. The discrepancies between these results and those of other recent trials may be due to the fact that blood pressure elevations among inner city blacks is stress related and therefore susceptible to stress reduction procedures. Alternatively, transcendental meditation may have unique effects not evident in other interventions, although prior analyses (e.g., Eisenberg et al., 1993; Jacob et al., 1991) appear not to corroborate this latter notion.

8.08.3.4 Adherence to Antihypertensive Treatments

As hypertension is a chronic, asymptomatic disorder, and because patients need to take medication that often has significant side effects (e.g., impotence, depression, and fatigue), nonadherence with antihypertensive medication regimens is a serious and prevalent problem. Descriptive studies indicate that fewer than 33–66% of patients comply with antihypertensive treatment regimens (Dunbar-Jacob, Wyer, & Dunning, 1991). Thus, the failure of patients to adhere to antihypertensive regimens is an important behavioral issue that has been investigated and clinically addressed by health psychologists. A comprehensive consideration of the problem of patient nonadherence to medical regimens used in the treatment of hypertension and other chronic diseases is beyond the scope of this chapter, and this issue will only be touched upon briefly here. For thorough considerations of this issue the reader is referred to Dunbar-Jacob et al. (1991) and Haynes, McKibbon, & Kanai (1996).

Research suggests that there are important differences between treatment settings in reported antihypertensive compliance rates, with family practice settings observing higher compliance and hospital clinics reporting poorer rates (Dunbar-Jacob et al., 1991). Interventions to increase adherence to regimens have demonstrated that factors such as increasing provider and family support, simplifying and making the regimen less complex, and introducing several behavioral self-management strategies may be effective in some situations. Nevertheless, the problem of antihypertensive compliance remains widespread.

Unfortunately, our inability to measure noncompliance precisely hinders interpretation of these data. The various ways in which compliance and noncompliance are assessed (e.g., patient self-report, physician report, and pill count) are not always accurate, and each one provides a separate set of problems. One of the more exciting developments in the area of compliance is the development of electronic pill monitors, which permit the unobtrusive and accurate assessment of medication compliance during daily life (Dunbar-Jacob, Sereika, Rohay, & Burke, in press).

8.08.3.5 Summary

In summary, behavioral influences on essential hypertension include excessive salt intake, obesity, and stress. There is evidence that inner city African Americans and those subject to racism are more prone to hypertension, and animal studies suggest that stress can result in hypertension in genetically susceptible strains. However, the effectiveness of behavioral intervention techniques, such as relaxation training, biofeedback, and meditation, in lowering blood pressure have been rather small and are generally of questionnable clinical significance. However, behavioral interventions such as weight loss, physical conditioning, and reduction of dietary salt intake have proven effective as adjuncts to pharmacologic interventions for hypertension. In addition, patient nonadherence to antihypertensive medication regimens is an important and prevalent problem that remains a topic of current study.

8.08.4 CONCLUSION

There are many environmental, behavioral, and physiological variables that interact in the development of cardiovascular disorders. Many of the standard CHD risk factors have important behavioral components, and increasing evidence suggests important psychosocial risk factors for CHD, including occupational stress, hostility, and physiologic reactivity to stress. In cardiac patients, the presence of acute stress, low social support, lack of economic resources, and psychological depression also appear to be important psychosocial risk factors. The identification of psychosocial risk factors for coronary disease have led to several promising behavioral and psychosocial interventions to aid in the treatment and prevention of coronary disease in high-risk individuals.

There also appear to be important biobehavioral influences in the development and treatment of essential hypertension. These include excessive salt intake, obesity, and stress. Evidence also indicates that genetic and environmental factors interact in the development of hypertension. However, the modest effects of behavioral stress-reducing techniques such as relaxation training, biofeedback, and meditation in lowering blood pressure have proven disappointing. Nevertheless, the important necessity for the involvement of health psychologists in the treatment of essential hypertension is underscored by the potential efficacy of weight loss, dietary modification, and exercise conditioning, as well as the need to ensure that patients adhere to medication regimens in order for pharmacologic interventions that have been shown to reduce subsequent morbidity to be effective.

ACKNOWLEDGMENTS

The opinions and assertions contained herein are those of the authors and are not to be construed as those of the USUHS or the US Department of Defense. Preparation of this chapter was supported by grants from the NIH (HL47337) and the USUHS (RO7233). Portions of this chapter were adapted from "Cardiovascular disorders and behavior" in A. Baum, R. Gatchel, & D. S. Krantz, *An introduction to health psychology* (3rd ed.). New York: McGraw-Hill, 1997.

8.08.5 REFERENCES

Alexander, C. N., Schneider, R. H., Staggers, F., Sheppard, W., Clayborne, B. M., Rainforth, M., Salerno, J., Kondwani, K., Smith, S., Walton, K. G., & Egan, B. (1996). Trial of stress reduction for hypertension in older African Americans. *Hypertension, 28,* 228–237.

Alexander F. (1939) Emotional factors in essential hypertension. *Psychosomatic Medicine, 1,* 173–179.

Allan, R., & Scheidt, S. (1996). Empirical basis for cardiac psychology. In R. Allan & S. Scheidt (Eds.), *Heart and mind: The practice of cardiac psychology* (pp. 63–124). Washington, DC: American Psychological Association.

Amick, T. L. & Ockene, J. K. (1994). The role of social support in the modification of risk factors for cardiovascular disease. In S. A. Shumaker & S. M. Czajkowski (Eds.), *Social support and cardiovascular disease* (pp. 259–280). (Plenum Series in Behavioral Psychophysiology and Medicine). New York: Plenum.

Appels, A. (1990). Mental precursors of myocardial infarction. *British Journal of Psychiatry, 156,* 465–471.

Aymann, D. (1933). The personality type of patients with arteriolar essential hypertension. *American Journal of Medical Science, 186,* 213.

Barefoot, J. C., Dahlstrom, W. G., & Williams, R. B. (1983). Hostility, CHD incidence and total mortality: A 25 year follow-up study of 255 physicians. *Psychosomatic Medicine, 45,* 59–63.

Barefoot, J. C., & Lipkus, I. M. (1994). The assessment of anger and hostility. In A. W. Siegman & T. W. Smith (Eds.), *Anger, hostility, and the heart* (pp. 43–66). Hillsdale, NJ: Erlbaum.

Barry, J., Selwyn, A. P., Nabel, E. G., Rocco, M. B., Mead, K., Campbell, S., & Rebecca, G. (1988). Frequency of ST segment depression produced by mental stress in stable angina pectoris from coronary artery disease. *American Journal of Cardiology, 61,* 989–93.

Baum, A., Gatchel, R., & Krantz, D. S. (1997). *An* *introduction to health psychology* (3rd ed.). New York: McGraw-Hill.

Blumenthal, J. A., & Emery, C. F. (1988). Rehabilitation of patients following myocardial infarction. *Journal of Consulting and Clinical Psychology, 56*(3) 374–381.

Blumenthal, J. A., Wei, J., Babyak, M., Krantz, D. S., Frid, D., Coleman, R. E., Waugh, R., Hanson, M., Appelbaum, M., O'Connor, C., & Morris, J. J. (1997). Stress management and exercise training in cardiac patients with myocardial ischemia: Effects on prognosis and on markers of myocardial ischemia. *Archives of Internal Medicine, 157,* 2213–2223.

Blumenthal, J. A., Jiang, W., Reese, J., Frid, D. J., Waugh, R., Morris, J. J., Coleman, R. E., Hanson, M., Thyrum, E. T., Krantz, D. S., O'Connor, C. (1995). Mental stress-induced ischemia in the laboratory and ambulatory ischemia during daily life: Association and hemodynamic features. *Circulation, 92,* 2102–2108.

Blumenthal, J. A., & We,i J. (1993). Psychobehavioral treatment in cardiac rehabilitation. *Cardiology Clinics, 11,* 323–331.

Campbell, R. J., & Henry, J. P. (1983). Animal models of hypertension. In D. S. Krantz, A. Baum, & J. E. Singer (Eds.), *Handbook of psychology and health: Vol. 3 Cardiovascular disorders and behavior* (pp. 155–197). Hillsdale, NJ: Erlbaum.

Carney, R. M., Freedland, K. C., Rich, M. W., & Jaffe, A. S. (1995). Depression as a risk factor for cardiac events in established coronary heart disease: A review of possible mechanisms. *Annals of Behavioral Medicine, 17*(2), 142–149.

Carney, R. M., Rich, M. W., Tevelde, A., Saini, J., Clark, K., & Jaffee, A. S. (1987). Major depressive disorder in coronary artery disease. *American Journal of Cardiology, 60,* 1273–1275.

Case, R. B., Moss, A. J., Case, N., McDermott, M., & Eberly, S. (1992). Living alone after myocardial infarction. *Journal of the American Medical Association, 267,* 515–519.

Chesney, M. A., Black, G. W., Swan, G. E., & Ward, M. M. (1987). Relaxation training for essential hypertension at the worksite. I. The untreated mild hypertensive. *Psychosomatic Medicine, 49*(3), 250–263.

Cobb, S., & Rose, R. M. (1973). Hypertension, peptic ulcer, and diabetes in air traffic controllers. *Journal of the American Medical Association, 224,* 489–492.

Cohen, S., Evans, G. W., Stokols, D., & Krantz, D. S. (1986). *Behavior, health and environmental stress.* New York: Plenum.

Cohen, S., & Wills, T. A. (1985). Stress, social support and the buffering hypothesis. *Psychological Bulletin, 98,* 310–357.

Contrada, R. J., & Krantz, D. S. (1988). Stress, reactivity and type A behavior: current status and future directions. *Annals of Behavioral Medicine, 10,* 64–70.

Cook, W. W., & Medley, D. M. (1954). Proposed hostility and pharisaic virtue scales for the MMPI. *Journal of Applied Psychology, 38,* 414–418.

Coresh, J., Klag, M. J., Mead, L. A., Liang, K. Y., & Whelton, P. K. (1992). Vascular reactivity in young adults and cardiovascular disease. A prospective study. *Hypertension, 19,* 218–223.

Coronary Artery Surgery Study Principal Investigators (1984). Myocardial infarction and mortality in the Coronary Artery Surgery Study (CASS) randomized trial. *New England Journal of Medicine, 310,* 750–758.

Cottier, C., Perini, C., & Rauchfleisch, U. (1987). Personality traits and hypertension: An overview. In S. Julius & D. R. Bassett (Eds.), *Handbook of hypertension: Vol 9. Behavioral factors in hypertension.* Amsterdam: Elsevier.

Cottington, E. M., Matthews, K. A., Talbott, E., & Kuller, L. H. (1980). Environmental events preceding sudden death in women. *Psychosomatic Medicine, 42,* 567–574.

Cushman, W. C., Cutler, J., Bingham, S. F., Harford, T., Hanna, E., Dubber, P., Collins, P., Dufour, M., Follman, D. & Allender, P. S. for the PATHS Group (1994). Prevention and treatment of hypertension study (PATHS): Rationale and design. *American Journal of Hypertension, 1,* 814–823.

Dahl, L. K., Heine, M., & Tassinari, L. (1962). Role of genetic factors in susceptibility to experimental hypertension due to chronic excess salt ingestion. *Nature, 194,* 480–482.

Deanfield, J. E., Shea, M., Kensett, M., Horlock, P., Wilson, R. A., de Landsheere, C. M., & Selwyn, A. P. (1984). Silent myocardial ischemia due to mental stress. *Lancet, ii,* 1001–1005.

Dembroski, T. M., MacDougall, J. M., Costa, P. T., & Grandits, G. A. (1989). Components of hostility as predictors of sudden death and myocardial infarction in the Multiple Risk Factor Intervention Trial. *Psychosomatic Medicine, 51,* 514–522.

Dubbert, P. M. (1995). Behavioral (life style) modification in the prevention and treatment of hypertension. *Clinical Psychology Review, 15*(3), 187–216.

Dunbar, H. F. (1943). *Psychosomatic diagnosis.* New York: Harper & Brothers.

Dunbar-Jacob, J., Sereika, S., Rohay, J., & Burke, L. E. (in press). Methods in ambulatory monitoring: Assessing adherence to medial regimens. In D. S. Krantz (Ed.), *Perspectives on behavioral medicine: innovations in technology and methodology.* Hillsdale, NJ: Erlbaum.

Dunbar-Jacob, J., Wyer, K., & Dunning, E. J. (1991). Compliance with antihypertensive regimen: A review of the research in the 1980s. *Annals of Behavioral Medicine, 13*(1), 31–39.

Eisenberg, D. M., Delbanco, T. L., Berkey, C. S., Kkaptchuk, T. J., Kupelnick, B., Kuhl, J., & Chalmers, T. C. (1993). Cognitive behavioral techniques for hypertension: Are they effective? *Annals of Internal Medicine, 188,* 964–972.

Engel G. L. (1971). Sudden and rapid death during psychological stress. *Annals of Internal Medicine, 74,* 771–782.

Esler, M., Julius, S., Zweifler, A., Randall, O., Harburg, E., Gardiner, H., & De Quattro, V. (1977). Mild high renin essential hypertension. A neurogenic human hypertension? *New England Journal of Medicine, 296,* 405–411.

Fletcher, G. F. (1996). The antiatherosclerotic effect of exercise and development of an exercise prescription. *Cardiology Clinics, 14,* 85–96.

Frasure-Smith, N. (1991). In-hospital symptoms of psychological stress as predictors of long-term outcome after acute myocardial infarction in men. *American Journal of Cardiology, 67,* 121–127.

Frasure-Smith, N., Lesperance, F., & Talajic, M. (1993). Depression following myocardial infarction: Impact on 6 month survival. *Journal of the American Medical Association, 270,* 1819–1825.

Frasure-Smith, N., & Prince, R. (1987). The Ischemic Heart Disease Life Stress Monitoring Program: Possible therapeutic mechanisms. *Psychology and Health, 1*(3), 273–285.

Frasure-Smith, N., & Prince, R. (1989). Long-term follow-up of the Ischemic Heart Disease Life Stress Monitoring Program. *Psychosomatic Medicine, 51*(5), 485–513.

Friedman, R., & Iwai, J. (1976). Genetic predisposition and stress-induced hypertension. *Science, 193,* 161–192.

Friedman, M., & Rosenman, R. H. (1959). Association of specific overt behavior pattern with blood and cardiovascular findings. *Journal of the American Medical Association, 169,* 1286–1296.

Friedman, M., & Rosenman, R. H. (1974). *Type A behavior and your heart.* New York: Knopf.

Freidman, M., Thorese, C., Gill, J., Ulmer, D., Powell, L., Price, V., Brown, B., Thompson, L., Rabin, D., Breall, W., Bourg, E., Levy, R., & Dixon, T. (1986). Alteration of type A behavior and its effects on cardiac recurrences in post myocardial infarction patients: Summary results of the recurrent coronary prevention project. *American Heart Journal, 112,* 653–665.

Gabbay, F. H., Krantz, D. S., Kop, W. J., Hedges, S. M., Klein, J., Gottdiener, J. S., & Rozanski, A. (1996). Triggers of myocardial ischemia during daily life in patients with coronary artery disease: Physical and mental activities, anger and smoking. *Journal of the American College of Cardiology, 27,* 585–592.

Glueck, C. J., Fallat, R. W., Millett, F., Gartside, P., Elston, R. C., & Go, R. C. (1975). Familial hyperalphaliporproteinemia: studies in eighteen kindreds. *Metabolism, 24,* 1243–1265.

Goldberg, A. D., Becker, L, C,, Bonsall, R., Cohen, J., Ketterer, M. W., Krantz, D. S., Light, K. C., McMahon, R. P., Noreuil, T., Pepine, C. J., Raczynski, J., Stone, P. H., Strother, D., Taylor, H., & Sheps, D. S. (1996). Ischemic, hemodynamic, and neurohormone responses to mental and exercise stress: Experience from the Psychophysiologic Investigations of Myocardial Ischemia. *Circulation, 94,* 2402–2409.

Goldman, L., & Cook, E. F. (1988). Reasons for the decline in coronary heart disease mortality: Medical interventions versus life style changes. In M. W. Higgins & R. V. Luepker (Eds.), *Trends in coronary heart disease mortality: The influence of medical care* (pp. 67–75). New York: Oxford University Press.

Gottdiener, J. S., Krantz, D. S., Howell, R. H., Hecht, G. M., Klein, J., Falconer, J. J., & Rozanski, A. (1994). Induction of silent myocardial ischemia with mental stress testing: Relationship to the triggers of ischemia during daily life activities and to ischemic functional severity. *Journal of the American College of Cardiology, 24,* 1645–1651.

Gutmann, M. C., Benson, H. (1971). Interaction of environmental factors and systemic arterial blood pressure: a review. *Medicine, 50*(6), 543–553.

Harburg, E., Erfurt, J. D., Haunsteinm, L. S., Chape, C., Schull, W. J., & Schork, M. A. (1973). Socioecologic stress, suppressed hostility, skin color and black-white male blood pressure: Detroit. *Psychosomatic Medicine, 35,* 276–296.

Haynes, S. B., & Feinleib, M. (1980). Women, work and coronary disease: Prospective findings from the Framingham Heart Study. *American Journal of Public Health, 700,* 133–141.

Haynes, S. G., Feinleib, M., & Kannel, W. B. (1980). The relationship of psychosocial factors to coronary heart disease in the Framingham Study. III. Eight-year incidence of coronary heart disease. *American Journal of Epidemiology, 3,* 37–58.

Haynes, R. B., McKibbon, K. A., & Kanai, R. (1996). Systematic review of randomized trials of interventions to assist patients to follow prescriptions for medications. *Lancet, 348,* 383–386.

Haythornthwaite, J. A., Pratley, R. E., & Anderson, D. E. (1992). Behavioral stress potentiates the blood pressure effects of a high sodium intake. *Psychosomatic Medicine, 54,* 231–239.

Helmers, K., Posluszny, D., & Krantz, D. (1994). Associations of hostility and coronary artery disease: A review of studies. In A. Siegman & T. Smith (Eds.), *Anger, hostility and the heart* (pp. 67–96). Hillsdale, NJ: Erlbaum.

Henry, J. P., & Cassel, J. C. (1969). Psychosocial factors in essential hypertension: Recent epidemiologic and animal experimental data. *American Journal of Epidemiology, 90,* 171–200.

Henry, J. P., & Stephens, P. M. (1977). *Stress, health and the social environment*. New York: Springer.

Higgins, M. W., & Luepker, R. V. (1988). *Trends in coronary heart disease mortality*. New York: Oxford University Press.

Hlatky, M. A., Lam, L. C., Lee, K. L., Clapp-Channing, N. E., Williams, R. B., Pryor, D. B., Califf, R. M., & Mark, D. B. (1995). Job strain and the prevalence and outcome of coronary artery disease. *Circulation, 92,* 327–333.

Hypertension Detection and Follow up Program (1979). Five year finding of the Hypertension Detection and Follow-up Program. I. Reduction in mortality of persons with high blood pressure, including mild hypertension. *Journal of the American Medical Association, 242,* 2562–2571.

Ironson, G., Taylor, C. B., Boltwood, M., Bartzokis, T., Dennis, C., Chesney, M., Spitzer, S., & Segall, G. M. (1992). Effects of anger on left ventricular ejection fraction in coronary artery disease. *American Journal of Cardiology, 70,* 281–285.

Irvine, M. J., Garner, D. M., Olmsted, M. P., & Logan, A. G.. (1989). Personality differences between hypertensive and normotensive individuals: Influence of knowledge of hypertension status. *Psychosomatic Medicine, 51,* 537–549.

Jacob, R. G., & Chesney, M. A. (1986). Psychological and behavioral methods to reduce cardiovascular reactivity. In K. A. Matthews (Ed.), *Handbook of stress, reactivity, and cardiovascular disease* (pp. 417–457). New York: Wiley.

Jacob, R. G., Chesney, M. A., Williams, D. M., & Ding, Y. (1991). Relaxation therapy for hypertension: Design effects and treatment effects. *Annals of Behavioral Medicine, 13,* 5–17.

Jenkins, C. D. (1983). Psychosocial and behavioral factors. In N. Kaplan & J. Stamler (Eds.), *Prevention of coronary heart disease: Practical management of risk factors* (pp. 98–112). Philadelphia: Saunders.

Jenkins, C. D., Rosenman, R. H., & Zyzanski, S. J. (1974). Prediction of clinical coronary heart disease by a test for the type A coronary-prone behavior pattern. *New England Journal of Medicine, 290,* 1271–1275.

Jenkins, C. D., Zyzanski, S. J., & Rosenman, R. H. (1971). Progress toward validation of a computer-scored test for the type A coronary-prone behavior pattern. *Psychosomatic Medicine, 33,* 193–202.

Jern, S. (1987). Specificity of personality factors found in hypertension. In S. Julius & D. R. Bassett (Eds.), *Handbook of hypertension: Vol 9. Behavioral factors in hypertension* (pp. 150–161). Amsterdam: Elsevier.

Johnson, E. H., Gentry, W. D., & Julius, S. (1992). *Personality, elevated blood pressure, and essential hypertension*. Washington, DC: Hemisphere.

Johnston, D. W., Gold, A., Kentish, J., Smith, D., Vallance, P., Shah, D., Leach, G., & Robinson, B. (1993). Effects of stress management on blood pressure in mild primary hypertension. *British Medical Journal, 306,* 963–966.

Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure. (1993). The fifth report of the Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure (JNC-V). *Archives of Internal Medicine, 153,* 154–183.

Julius,. S, & Bassett, D. R. (1987). *Handbook of hypertension: vol 9. Behavioral factors in hypertension*. Amsterdam: Elsevier.

Julius, S., & Esler, M. (1975). Autonomic nervous cardiovascular regulation in borderline hypertension. *American Journal of Cardiology, 36,* 685–696.

Kamarck, T., & Jennings, J. R. (1991). Biobehavioral factors in sudden cardiac death. *Psychologic Bulletin, 109,* 42–75.

Kannel, W. B. (1979). Cardiovascular disease: A multifactorial problem (insights from the Framingham study). In M. L. Pollack & D. H. Schmidt (Eds.), *Heart disease and rehabilitation* (pp. 15–31). Boston: Houghton Mifflin.

Kaplan, J. R., Adams, M. R., Clarkson, T. B., & Koritnick, D. R. (1984). Psychosocial influences on female "protection" among cynomolgus macaques. *Atherosclerosis, 53,* 283–295.

Kaplan, R. M., & Bush, J. W. (1982). Health related quality of life measurement for evaluation research and policy analysis. *Health Psychology, 1,* 61–80.

Karasek, R. A., Russell, R. S., & Theorell, T. (1982). Physiology of stress and regeneration in job related cardiovascular illness. *Journal of Human Stress, 8,* 29–42.

Karasek, R. A., & Theorell, T. G. (1990). *Healthy work, stress, productivity, and the reconstruction of working life.* New York: Basic Books.

Karasek, R. A., Theorell, T. G., Schwarz, J., Pieper, C., & Alfredsson, L. (1982). Job, psychological factors and coronary heart disease: Swedish prospective findings and US prevalence findings using a new occupational inference method. *Advances in Cardiology, 29,* 62–67.

Karasek, R. A., Theorell, T. G., Schwartz, J., Schnall, P., Pieper, C., & Michela, J. (1988). Job characteristics in relation to the prevalence of myocardial infarction in the US Health Examination Survey (HES) and the Health and Nutrition Examination Survey (HANES). *American Journal of Public Health, 78,* 910–918.

Kark, J. D., Goldman, S., & Epstein, L. (1995). Iraqi missile attacks on Israel. *Journal of the American Medical Association, 273,* 1208–1210.

Kelley, G., & McClellan, P. (1994). Antihypertensive effects of aerobic exercise. A brief meta-analytic review of randomized controlled trials. *American Journal of Hypertension, 7,* 115–119.

Kennedy, G. H., Hofer, M. A., Cohen, D., Shindledecker, R., & Fisher, J. D. (1987). Significance of depression and cognitive impairment in patients undergoing programmed stimulation of cardiac arrhythmias. *Psychosomatic Medicine, 49,* 410–421.

Keys, A., Taylor, H. L., Blackburn, H., Brozek, J., Anderson, J., & Simonson, E. (1971). Mortality and coronary heart disease among men studied for 23 years. *Archives of Internal Medicine, 128,* 201–214.

Kop, W. J., Appels, A. P. W. M., de Leon C. M., de Swart H. B., & Bar F. W. (1994). Vital exhaustion predicts new cardiac events after successful coronary angioplasty. *Psychosomatic Medicine, 56,* 281–287.

Krantz, D. S., & Blumenthal, J. (1987). *Behavioral assessment and management of cardiovascular disorders.* Sarasota, FL: Professional Resource Exchange.

Krantz, D. S., & Durel, L. A. (1983). Psychobiological substrates of the type A behavior pattern. *Health Psychology, 2,* 393–411.

Krantz, D. S., Grunberg, N. E., & Baum, A. (1985). Health psychology. *Annual Review of Psychology, 36,* 349–383.

Krantz, D. S., Kop, W. J., Gabbay, F. H., Rozanski, A., Barnard, M., Klein, J., Pardo, Y., & Gottdiener, J. S. (1996). Circadian variation of ambulatory myocardial ischemia: Triggering by daily activities and evidence for an endogenous circadian component. *Circulation, 93,* 1364–1371.

Krantz, D. S., & Manuck, S. B. (1984). Acute psychophysiologic reactivity and ride of cardiovascular disease: a review and methodologic critique. *Psychological Bulletin, 96,* 435–464.

Krieger, N., & Sidney, S. (1996). Racial discrimination and blood pressure: The CARDIA study of young black and white adults. *American Journal of Public Health, 86,* 1370–1378.

LaCroix, A. Z., & Haynes, S. G. (1987). Gender differences in the health effect of workplace roles. In R. C. Barnett, L. Biener & G. K. Baruch (Eds.), *Gender and stress* (pp. 96–121). New York: Free Press.

Lau, J., Antman, E. M., Jimenez-Silva, J., Kupelnick, B., Mosteller, F., & Chalmers, T. C. (1992). Cumulative meta analysis of therapeutic trials for myocardial infarction. *New England Journal of Medicine, 327,* 248–254.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping.* New York: Springer.

Lerner, D. J., & Kannel, W. B. (1986). Patterns of coronary heart disease morbidity and mortality in the sexes: A 26 year follow-up of the Framingham population. *American Heart Journal, 111,* 383–390.

Linden, W., Stossel, C., & Maurice, J. (1996). Psychosocial interventions for patients with coronary artery disease: A meta-analysis. *Archives of Internal Medicine, 156,* 745–752.

MacMahon, S. (1987). Alcohol consumption and hypertension. *Hypertension, 9,* 111–121.

Mann, G. V. (1977). Diet-heart: End of an era. *New England Journal of Medicine, 97,* 644–650.

Manuck, S. B. (1994). Cardiovascular reactivity in cardiovascular disease: "Once more unto the breach." *International Journal of Behavioral Medicine, 1,* 4–31.

Manuck, S. B., Kaplan, J. R., & Clarkson, T. B. (1983). Behaviorally induced heart rate reactivity and atherosclerosis in cynomolgus monkeys. *Psychosomatic Medicine, 45,* 95–108.

Manuck, S. B., Kaplan, J. R., & Matthew, K. A. (1986). Behavioral antecedents of coronary heart disease and atherosclerosis. *Atherosclerosis, 6*(1), 2–14.

Manuck, S. B., Marsland, A. L., Kaplan, J. R., & Williams, J. K. (1995). The pathogenicity of behavior and its neuroendocrine mediation. An example from coronary artery disease. *Psychosomatic Medicine, 57,* 275–283.

Markovitz, J. H., Matthews, K. A., Kannel, W. B., Cobb, J. L., & D'Agonstino, R. B. (1993). Psychological predictors of hypertension in the Framingham Study: Is there tension in hypertension. *Journal of the American Medical Association, 270,* 2439–2443.

Marmot, M. G. (1983). Stress, social and cultural variations in heart disease. *Journal of Psychosomatic Research, 27,* 377–384.

Martin, J. E., Dubbert, P. M., & Cushman, W. C. (1990). Controlled trial of aerobic exercise in hypertension. *Circulation, 81,* 1560–1567.

Matthews, K. A. (1982). Psychological perspectives on the type A behavior pattern. *Psychological Bulletin, 91,* 293–323.

Matthews, K. A., Glass, D. C., Rosenman, R. H., & Bortner, R. W. (1977). Competitive drive pattern A and coronary heart disease: A further analysis of some data from the Western Collaborative Group Study. *Journal of Chronic Diseases, 30,* 489–498.

Matthews, K. A., & Haynes, S. G. (1986). Type A behavior pattern and coronary disease risk: Update and critical evaluation, *American Journal of Epidemiology, 123,* 923–960.

Matthews, K. A., Weiss, S. M., Detre, T., Dembroski, T. M., Falkner, B., Manuck, S. B., & Williams, R. B., Jr. (Eds.) (1986). *Handbook of stress, reactivity, and cardiovascular disease.* New York: Wiley.

Meisel, S. R., Kutz, I., Dayan, K. I., Pauzner, H., Chetboun, I., Arbel, Y., & David, D. (1991). Effect of Iraqi missile war on incidence of acute myocardial infarction and sudden death in Israeli civilians. *Lancet, 338,* 660–661.

Mittleman, M. A., Maclure, M., Sherwood, J. B., Mulry, R. P., Tofler, G., Jacobs, S. C., Friedman, R., Benson, H., & Muller, J. E. (1995). Triggering of acute myocardial infarction onset by episodes of anger. *Circulation, 92,* 1720–1725.

Muller, J. E., Tofler, G. H., & Stone, P. H. (1989). Circadian variation and triggers of onset of acute cardiovascular disease. *Circulation, 79,* 733–743.

Myers, R., & Dewar, H. A. (1975). Circumstances surrounding sudden deaths from coronary artery disease with coroner's necropsies. *British Heart Journal, 37,* 1133–1143.

National Heart, Lung, and Blood Institute (1994). *Report of the task force on research in epidemiology and prevention of cardiovascular diseases*. Bethesda, MD: Author.

Nicolosi, R. J., & Schaefer, E. J. (1992). Pathobiology of hypercholesterolemia and atherosclerosis: Genetic and environmental determinants of elevated lipoprotein levels. In I. S. Ockene & J. K. Ockene (Eds.), *Prevention of coronary heart disease* (pp. 69–102). Boston: Little, Brown.

Ockene, I. S., & Ockene, J. K. (1992). *Prevention of coronary heart disease*. Boston: Little, Brown.

Oldridge, N. B., Guyatt, G. H., Fischer, M. E., & Rimm, A. A. (1988). Cardiac rehabilitation after myocardial infarction: Combined experience of randomized clinical trials. *Journal of the American Medical Association, 260,* 945–950.

Ornish, D., Brown, S. E., Scherwitz, L. W., Billings, J. H., Armstrong, W. T., Ports, T. A., McLanahan, S. M., Kirkeede, R. L., Brand, R. J., & Gould, K. L. (1990). Can lifestyle changes reverse coronary heart disease? *Lancet, 336,* 129–133.

Page, L., Danion, A., & Moellering, R. C. (1970). Antecedents of cardiovascular disease in six Solomon Island societies. *Circulation, 49,* 1132–1140.

Parkes, C. M., Benjamin, B., and Fitzgerald, R. G. (1969). Broken heart: A statistical study of increased mortality among widowers. *British Medical Journal, 1,* 740–743.

Pate, R. R., Pratt, M., Clair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C., Buchner, D., Ettinger, W., Heath, G. W., King, A. C., Kriska, A., Leon, A., Marcus, B., Morris, J., Paffenbarger, R., Patrick, K., Pollock M., Rippe, J., Sallis, J., & Wilmore, J. (1995). Physical activity and public health: A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *Journal of the American Medical Association, 273,* 402–407.

Patel, C., Marmot, M. G., Terry, D. J., (1981). Controlled trial of biofeedback-aided behavioral methods in reducing mild hypertension. *British Medical Journal, 282*(6281), 2005–2008.

Patel, C., Marmot, M. G., Terry, D. J., Carruthers, M., Hunt, B., & Patel, M. (1985). Trial of relaxation in reducing coronary risk: Four year follow up. *British Medical Journal, 290,* 1103–1106.

Pickering, T. G. (1969). Personal views on mechanisms of hypertension. In J. Genest, E. Koiw, & O. Kuchil (Eds.) Hypertension: Recent epidemiologic and animal experimental data. *Hypertension: Physiopathology and treatment* (pp. 598–606).

Pickering, T. G., James, G. D., & Boddie, C. (1988). How common is white coat hypertension? *Journal of the American Medical Association, 259,* 225–228.

Pickering, T. G., Sleight, P. (1977). Baroreceptors and hypertension. *Progress in Brain Research, 47,* 43–60.

Pratt, L. A., Ford, D. E., Crum, R. M., Armenian, H. K., Gallo, J. J., & Eaton, W. W. (1996). Depression, psychotropic medication, and risk of myocardial infarction. *Circulation, 94,* 3123–3129.

Ragland, D. R., & Brand, R. J. (1988). Type A behavior and mortality from coronary heart disease. *New England Journal of Medicine, 318,* 65–69.

Reisen, E., Abel, R., Modan, M., Silverberg, D., Eliahou, H., & Modan, B. (1978). Effect of weight loss without salt restriction on the reduction of blood pressure in overweight hypertensive patients. *New England Journal of Medicine, 298,* 1–6.

Rosenman, R. H., Brand, R. J., Jenkins, C. D., Friedman, M., Straus, R., & Wurm, M. (1975). Coronary heart disease in the Western Collaborative Group Study: Final follow-up experience of 872 years. *Journal of the American Medical Association, 233,* 872–877.

Rozanski, A., Bairey, C. N., Krantz, D. S., Friedman, J., Resser, K. J., Morell, M., Hilton-Chalfen, S., Hestrin, L., Bietendorf, J., & Berman, D. S. (1988). Mental stress and the induction of silent myocardial ischemia in patients with coronary artery disease. *New England Journal of Medicine, 318,* 1005–1012.

Ruberman, W., Weinblatt, E., Goldberg, J. D., & Chaudhary, B. S. (1984). Psychosocial influences on mortality after myocardial infarction. *New England Journal of Medicine, 311,* 552–559.

Schang, S. J., & Pepine, C. J. (1977). Transient asymptomatic ST depression during daily activity. *American Journal of Cardiology, 39,* 396–402.

Schleifer, S., Keller, S. E., Bond, R. N., Cohen, J., & Stein, M. (1989). Major depressive disorder and immunity. *Archives of General Psychiatry, 46,* 81–87.

Schnall, P. L., Peiper, C., Schwartz, J. E., Krasek, R. A., Devereux, R. B., Pickering, T. G., & Schwartz, J. E. (1990). The relationship between "job strain," workplace diastolic blood pressure, and left ventricular mass index. Results of a case-control study. *Journal* of the American Medical Association, 263(14), 1929–1935.

Schneider, R. H., Staggers, F., Alexander, C. N., Sheppard, W., Rainforth, M., Kondwani, K., Smith, S., & King, C. G. (1995). A randomized controlled trial of stress reduction for hypertension in older African Americans. *Hypertension, 26,* 820–827.

Shapiro, A. P. (1983). The non-pharmacologic treatment of hypertension. In D. S. Krantz, A. Baum & J. E. Singer (Eds.), *Handbook of psychology and health: Vol. 3. Cardiovascular disorders and behavior* (pp. 277–293). Hillsdale, NJ: Erlbaum.

Shapiro, A. P., Schwartz, G. E., Ferguson, D. C. E., Redmond, D. P., & Weiss, S. M. (1977). Behavioral methods in the treatment of hypertension. I. Review of their clinical status. *Annals of Internal Medicine, 86,* 626–636.

Shekelle, R. B., Hulley, S. B., & Neaton, J. (1985). The MRFIT behavior pattern study. II. Type A behavior and incidence of coronary heart disease. *American Journal of Epidemiology, 122,* 559–570.

Shekelle, R. B., Gale, M., Ostfeld, A. M., & Paul, O. (1983). Hostility, risk of coronary heart disease, and mortality. *Psychosomatic Medicine 45,* 109–114.

Shumaker, S. A., & Czajkowski, S. M. (Eds.) (1994). *Social support and cardiovascular disease* (Plenum Series in Behavioral Psychophysiology and Medicine). New York: Plenum.

Siegel, W. C., Blumenthal, J. A., & Divine, G. W. (1990). Physiological, psychological and behavioral factors and white coat factors and white coat hypertension. *Hypertension, 16,* 140–146.

Siegler, I. C. (1994). Hostility and risk: Demographic and lifestyle variables. In A. W. Siegman & T. W. Smith (Eds.), *Anger, hostility, and the heart* (pp. 199–214). Hillsdale, NJ: Erlbaum.

Siegman, A., & Smith, T. W. (1994). *Anger, hostility, and the heart*. Hillsdale, NJ: Erlbaum.

Stoney, C. M., & Engbretson, T. O. (1994). Anger and hostility: Potential mediators of the gender difference in coronary heart disease. In A. W. Siegman & T. W. Smith (Eds.), *Anger, hostility, and the heart*. Hillsdale, NJ: Erlbaum.

Suinn, R. M. (1982). Intervention with type A behaviors. *Journal of Consulting and Clinical Psychology, 50*(6), 933–949.

Syme, S. L., & Torfs, C. P. (1978). Epidemiologic research in hypertension: A critical appraisal. *Journal of Human Stress, 4,* 43–47.

Wassertheil-Smoller, S., Applegate, W. B., Berge, K., Chang, C. J., Davis, B. R., Grimm, R., Kostis, J., Pressel, S., & Schron, E. (1996). Change in depression as a precursor of cardiovascular events. *Archives of Internal Medicine, 156,* 553–561.

Weder, A. B., & Julius, S. (1987). Behavior, blood pressure variability and hypertension: a research agenda. In S. Julius & D. R. Basset (Eds.), *Handbook of hypertension: Vol. 9. Behavioural factors in hypertension.* Amsterdam: Elsevier.

Weiner, H., & Sapira, J. D. (1987) Hypertension: A challenge to behavioral research. In S. Julius & D. R. Basset (Eds.), *Handbook of hypertension: Vol. 9. Behavioral factors in hypertension.* Amsterdam: Elsevier.

Wells, J. A. (1985). Chronic life situations and life change events. In A. M. Ostfeld, E. D. Eaker, & T. J. Truss (Eds.), *Measuring psychosocial variables in epidemiologic studies of cardiovascular disease* (NIH Publication No. 85–2270). Bethesda, MD: NIH.

Williams, R. B., Barefoot, J. C., Haney, T. L., Harrell, F. E., Jr., Blumenthal, J. A., Pryor, M. A., & Peterson, B. (1988) Type A behavior and angiographically documented coronary atherosclerosis in a sample of 2289 patients. *Psychosomatic Medicine, 50,* 139–152.

Williams, R. B., Barefoot, J. C., Califf, R. M., Haney, T. L., Saunders, W. B., Pryor, D. B., Hlatky, M. A., Siegler, I. C., & Mark, D. B. (1992). Prognostic importance of social and economic resources among medically treated patients with angiographically documented coronary artery disease. *Journal of the American Medical Association, 267*(4), 559–560.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.09
Cancer

BARBARA L. ANDERSEN and DEANNA M. GOLDEN-KREUTZ
Ohio State University, Columbus, OH, USA

8.09.1 INTRODUCTION 217
8.09.2 MAGNITUDE OF THE PROBLEM INCIDENCE, DEATH RATES, AND GENDER DIFFERENCES 218
8.09.3 ETIOLOGY, RISK FACTORS, AND PREVENTION 218
8.09.4 PSYCHOLOGICAL, SOCIAL, AND BEHAVIORAL IMPACT OF CANCER 220
  8.09.4.1 Symptom Appearance 220
  8.09.4.2 Diagnosis 221
  8.09.4.3 Treatment 222
    8.09.4.3.1 Surgery 223
    8.09.4.3.2 Radiotherapy 223
    8.09.4.3.3 Chemotherapy 223
    8.09.4.3.4 Bone marrow transplantation 224
    8.09.4.3.5 Treatment side effects 225
    8.09.4.3.6 Choosing cancer treatments 226
  8.09.4.4 Recovery and Long-term Survival 227
  8.09.4.5 Recurrence and Death 229
8.09.5 CONCLUSION 231
8.09.6 REFERENCES 232

8.09.1 INTRODUCTION

The cost of human lives to cancer (defined as any malignant tumor; Rothenberg, 1992), is overwhelming. In many countries cancer is the second leading cause of death, only outnumbered by heart disease. In the USA alone, over 3500 people are diagnosed with cancer and another 1500 people die from cancer-related deaths *per day* (Parker, Tong, Bolden, & Wingo, 1997). While much of the increase in cancer incidence and mortality over the years appears related to advances in early detection as well as the general aging of the population (age is a risk factor; Garfinkel, 1994), Cunningham (1997) has noted cancer death rates may now be showing a downward trend in the USA. There is, however, variability across countries. For example, age-adjusted death rates per 100 000 population across sites and gender for 1990–1993 ranged from 394 in Hungary to 139 in Albania; the rate for the USA is 276 (average: Parker et al., 1997). In short, cancer is a significant medical problem that affects the health status of millions of people worldwide.

Research on the psychological, social, and behavioral aspects of oncology began in the early 1950s; however, our knowledge base has significantly expanded since the 1970s. In particular, research has clarified relationships between psychological responses (e.g., personality, mood, coping style, relationships), social factors (e.g., presence/absence of partner, degree of social support), and behavioral variables (e.g., compliance with treatment, diet, exercise). More recent research has incorporated biologic systems (e.g., immune and endocrine), examining the interaction of psychological, social, and behavioral variables and their relationship to disease course (see

Fawzy et al., 1993, for an example; see Andersen, Kiecolt-Glaser, & Glaser, 1994, for a discussion).

This chapter includes an overview of the central findings on the psychological, social, and behavioral impact of cancer. By way of introduction, we begin by addressing the magnitude of the cancer problem. Data on cancer incidence, death rates, and gender differences are highlighted. Where data from the USA are used, it is noted that they represent the same general trends found in other industrialized, Western countries. Second, we briefly discuss etiologic and risk factors and the role of prevention. The largest section of the chapter organizes research findings by disease relevant time points, from symptom discovery to recovery or death. Where appropriate, specific intervention studies will be detailed. Information about cancer's impact on a patient's social network (spouse, family, etc.) is also included. We conclude with a discussion of psychological interventions and techniques that appear to be effective in promoting improved quality of life and coping in cancer patients.

8.09.2 MAGNITUDE OF THE PROBLEM INCIDENCE, DEATH RATES, AND GENDER DIFFERENCES

Cancers vary in their prevalence and mortality. Tables 1 and 2 display data from the USA on the incidence and death rates by specific sites and genders. These data indicate, for example, that the most common diagnosis for women is breast cancer and for men, prostate cancer, but that lung cancer is the number one killer for both (Parker et al., 1997). Similarly, Tables 3 and 4 provide death rate data from selected countries (Parker et al., 1997). These data illustrate the geographic, nationality, and sex differences that exist (e.g., lung cancer death rates of 81.6 per 100 000 for Hungarian men, but a rate of 5.2 for French women). In examining the death rates worldwide, the top killers of women include breast, colon and rectum, lung, and stomach cancers. For men, lung, stomach, colon and rectum, and prostate cancers are the common killers worldwide.

8.09.3 ETIOLOGY, RISK FACTORS, AND PREVENTION

The specific events that turn a normal cell into a cancerous cell are still unknown (Rugo, 1993). However, we do know that it takes generations of mutations for cells to become cancerous. The development of cancer within an individual appears to depend on genetic or environmental factors, with exposure to carcinogens often playing a significant role in many cancers (Rugo, 1993). In fact, cancer incidence and concomitant premature death have been linked to behavioral habits that increase exposure. In the United States, for example, the National Cancer Institute set a goal of a 50% reduction in cancer mortality by the year 2000 through prevention and control efforts focused, in large part, on lifestyle (Greenwald & Cullen, 1985).

Primary prevention tactics have attempted to reduce the probability of cancer onset by decreasing exposure to risk factors. For women and men, unhealthy habits have been a behavioral research target, and smoking has been emphasized, as it is related to 30% of all cancer deaths (Rugo, 1993). Cigarette smoking is directly linked to lung cancer, one of the leading causes of death for men and women, and is implicated in cancers of the larynx, head and neck, esophagus, bladder, kidney, pancreas, and stomach.

Other lifestyle behaviors that have been linked to greater risk are dietary patterns and sun exposure (e.g., ultraviolet light). Diet research has suggested that diets high in fat and low in fiber may contribute to increased risk for particular cancers (e.g., breast and colon, respectively). Recommendations for modification—reductions in fat intake and increases in complex carbohydrates and fiber—and weight reduction *per se* have been suggested (Ashley, 1993; Wynder & Cohen, 1993). In breast cancer prevention, for example, interventions have been developed for at-risk women to reduce weight and fat intake and to increase physical activity (Heber, Bagga, Ashley, & Elashloff, 1993). Parallel interventions focused on increasing fiber in the diet may be important for decreasing one's risk for gastrointestinal cancers. Concerning sun exposure, the number of people diagnosed with skin cancers has been steadily increasing over recent years. In fact, close to 1 million diagnoses of basal and squamous cell skin cancers are expected in 1997 and this figure does not include malignant melanoma (Parker et al., 1997). It has been recommended that time spent exposed to the sun, especially during the hottest part of the day, be decreased and the use of lotions that block exposure to harmful ultraviolet rays be used. Another risk factor involves exposure to certain viruses. The human papilloma virus has been linked to gynecologic cancers and the human immunodeficiency virus has been associated to non-Hodgkin's lymphoma and Kaposi's sarcoma. Prevention efforts have focused on safer sexual practices with barrier methods (e.g., diaphram,

Table 1 1997 estimates of cancer incidence by site and gender for the five leading sites.

Male (total estimated 785 800)		*Female (total estimated 506 000)*	
Site	*Number (%)*	*Site*	*Number (%)*
Prostate	334 500 (43)	Breast	180 200 (30)
Lung	98 300 (13)	Colon/rectum	79 800 (13)
Colon/rectum	48 100 (8)	Lung	50 900 (11)
Bladder	39 500 (5)	Uterus	34 900 (6)
Lymphoma	30 300 (4)	Ovary	26 800 (4)

Adapted from *Cancer Statistics — 1997*, American Cancer Society.

Table 2 1997 estimates of cancer mortality by site and gender for the five leading sites.

Male (total estimated 294 100)		*Female (total estimated 265 900)*	
Site	*Number (%)*	*Site*	*Number (%)*
Lung	94 400 (32)	Lung	66 000 (25)
Prostate	41 800 (14)	Breast	43 900 (17)
Colon/rectum	27 000 (9)	Colon/rectum	27 900 (10)
Pancreas	13 500 (5)	Pancrease	14 600 (5)
Lymphoma	12 400 (4)	Ovary	14 200 (5)

Adapted from *Cancer Statistics — 1997,* American Cancer Society.

Table 3 Age-adjusted death rates in adult males per 100 000 for selected countries and disease sites for 1990–1993.

Country	*All sites*	*Colon/rectum*	*Lung*	*Prostate*	*Stomach*
USA	165.3 (27)	16.5	57.1	17.5	5.0
Hungary	258.7 (1)	38.0	81.6	15.8	23.4
Poland	204.2 (8)	15.4	71.3	9.9	21.6
Italy	189.2 (15)	15.3	57.0	11.6	16.9
New Zealand	170.6 (21)	26.4	42.6	19.0	9.0
Finland	149.9 (31)	12.4	46.1	18.3	11.5
Israel	116.9 (43)	15.3	25.7	9.2	8.7
Mexico	81.6 (48)	3.3	15.9	10.6	9.9

Figures in parentheses are order of rank based on data from 48 countries during the years 1990–1993. Adapted from Parker et al. (1997).

Table 4 Age-adjusted death rates in Adult females per 100 000 for selected countries and disease sites for 1990–1993.

Country	*All sites*	*Colon/rectum*	*Lung*	*Prostate*	*Stomach*
USA	111.1 (8)	22.0	11.2	25.6	2.3
Denmark	138.7 (1)	27.2	17.2	24.8	3.9
Austria	105.6 (16)	21.8	13.7	9.3	7.8
Russian Fed.	99.9 (21)	14.9	12.3	7.3	16.9
Argentina	96.4 (27)	21.3	9.2	6.0	4.8
France	86.5 (36)	19.7	10.1	5.2	3.3
Japan	75.2 (43)	6.6	9.8	8.1	14.2
Albania	42.7 (48)	5.2	2.2	5.0	6.3

Figures in parentheses are order of rank based on data from 48 countries during the years 1990–1993. Adapted from Parker et al. (1997).

use of condoms). Additionally, a family history of cancer may double a person's risk for that same cancer. For example, in breast cancer, the greater the number of affected relatives and the younger these relatives were at the time of diagnosis, the higher a woman's risk. Women with multiple first-degree relatives (e.g., mother, sisters) with premenopausal breast cancer are at highest risk. In such families, breast cancer appears to be transmitted as a gene, and the risk of breast cancer may approach 80%. New evidence is also emerging on a gene associated with the development of prostate cancer.

Secondary prevention efforts have been defined as those that identify the disease at the earliest of stages (e.g., when it is preinvasive or localized and asymptomatic) so that effective treatment can be administered sooner and mortality reduced. The chances of successful early cancer detection and treatment depend upon the clinical characteristics of the disease and the screening strategy, and it is unlikely that secondary prevention can proceed effectively unless certain conditions for both are met.

Behavioral scientists have focused mainly on the early detection of breast cancer, especially with mammography (the most reliable means of early detection). Investigators (e.g., Rimer, 1992) have considered the psychological, behavioral, sociodemographic, and healthcare factors that appear to be related to women's usage of mammography. Older women, poorer women, and those with weak linkages to the healthcare system, for example, are less frequent users of mammography. Factors such as these result in deadly consequences for women, as poorer or older women without a history of mammography are often diagnosed with more advanced disease. Thus, to the extent that behavioral scientists can affect women's acceptance and utilization of the technique, survival may be directly improved. Early detection efforts have also focused on prostate, skin, and colorectal cancers with behavioral scientists finally beginning to study prostate cancer.

8.09.4 PSYCHOLOGICAL, SOCIAL, AND BEHAVIORAL IMPACT OF CANCER

8.09.4.1 Symptom Appearance

In the study of delay in seeking treatment for medical conditions, several definitions of delay have been used, but an oft-chosen operationalization is the number of days from the detection of the first symptom to an end point. This latter variable is less consistently chosen, but has included seeing a physician for symptoms (e.g., Coates et al., 1992; Waters, Nichols, Wheeler, Fraser, & Hayes, 1983), being diagnosed with a medical condition (Marshall, Gregorio, & Walsh, 1982), or beginning treatment for the condition (e.g., Howson, 1950). According to this view, all individuals, and even some physicians, would be "delayers." Still another strategy has been to designate some delay as more "reasonable" than others (e.g., one month from symptom appearance to appearance before a physician; see early work in cancer by Pack & Gallo, 1938).

Andersen and Cacioppo (1995) have conceived of delay as comprising a series of stages (e.g., appraisal, illness, behavioral, scheduling) each governed by a conceptually distinct set of decisional and appraisal processes. Appraisal delay, delay surrounding symptom interpretation, has been found to account for most of the delay in seeking a cancer diagnosis (see Andersen & Cacioppo, 1995, for a model of delay). For example, the appraisal interval was approximately 80% of the total delay for women with gynecologic cancer, whereas for women with breast symptoms, a site that has a narrower range of symptom diversity, the interval accounted for 60% of the delay.

It is likely that there are several reasons for the finding of appraisal processes accounting for the bulk of the delay time. First, the development of malignancy and the appearance of cancer symptoms are protracted, and a complex and changing symptom picture can be typical, unlike the presentation of many other medical problems (e.g., myocardial infarction; see Matthews, Siegel, Kuller, Thompson, & Varat, 1983). Symptoms can also vary with the site and extent of the disease. For example, ovarian cancer has varied presentations—pelvic cramping, low back pain, pain or bleeding with intercourse, urinary frequency irregularities. Moreover, as cancers progress, they can involve other bodily systems (e.g., gastrointestinal, lymphatic) and the symptom picture tends to change from specific or localized complaints (e.g., vaginal discharge/bleeding) to diffuse ones (e.g., loss of appetite, nausea, flu symptoms). Finally, cancer is a life-threatening disease but a low probability disease for many, and thus the appraisal process may lengthen.

Briefly considering the other stages, illness delay is defined as the number of days elapsing from the time an individual concludes he or she is ill to the day he or she decides to seek medical help. At this time individuals must decide, for example, whether to seek assistance from others (e.g., physician, others with a similar condition) or to self-treat the illness. After this, the delay is spent in making two remaining decisions. One is the delay between the decision to seek medical attention and the person acting on this decision by making an appointment (behavioral delay). The other is the time that elapses between the person making an appointment and his or her first receipt of medical attention (scheduling delay). Economic factors such as affordability, social factors such as family influence, and cognitive factors such as the extent to which the decision to seek medical help is based on issue-relevant thinking are more likely to modulate the delay time between a decision to seek and an action to receive medical care (behavioral delay). In contrast, both patient characteristics (such as the manner in which the person describes his or her concerns and symptoms) and medical environment characteristics (such as a physician's appointment backlog), which are not under an individual's control, may modulate the delay incurred when scheduling an appointment. Therefore, shortening delay to seeking a diagnosis once symptom/sign awareness has occurred (tertiary prevention) would improve survival, as early detection is linked to improved survival rates for many sites (e.g., cervix, breast, colon, prostate).

8.09.4.2 Diagnosis

An early clinical study suggested that the diagnosis of cancer produces an "existential plight," meaning that the news brings shock, disbelief, and emotional turmoil (Weisman & Worden, 1976). Today, we know that individuals even become alarmed at the time of medical screening, long before a cancer diagnosis is suggested (Wardle & Pope, 1992). Perhaps because of these natural, difficult reactions, over the years there has been some variability in whether or not patients were told that their diagnosis was cancer.

In the USA as well as most Western countries, patients (including children) are now told that their diagnosis is cancer, since this is regarded as the moral, ethical, and legal procedure to follow (Woodard & Pamies, 1992). Support of this policy has documented the beneficial effects of coping successfully with an extreme stressor such as cancer (Taylor, 1983), and follow-up studies of childhood leukemia survivors, for example, have pointed toward more favorable long-term adjustment for children when they learn of their diagnosis early, whether directly, accidentally, or through their own efforts, in comparison with children who learn late (Slavin, O'Malley, Koocher, & Foster, 1982). Likewise, Mosconi, Meyerowitz, Liberati and Liberati (1991) found that of 1171 Italian breast cancer patients, those who were told their diagnosis was cancer were more satisfied with the information they had received from their physicians. Still other cross-cultural research has indicated that the majority of patients want to be told their diagnosis truthfully (e.g., Japan, Mizushima et al., 1990; UK, Sell et al., 1993). Further, the manner in which the information is disclosed is important. Physicians who communicate hope have patients who are more hopeful, more favorable towards their cancer diagnosis, and who report more favorable overall emotional adjustment (Sardell & Trierweiler, 1993). This points to the need of physicians and other health professionals to learn more effective and thorough means of communicating information (e.g., diagnosis, treatment, prognosis) to their patients (Sell et al., Woodard & Pamies, 1992).

Empirical reports clarify the specific emotions—sadness (depression), fear (anxiety), and anger—that characterize this period. It is not surprising that depression is the most common affective problem. Survey estimates for major depression are of the order of 5–6% (Derogatis et al., 1983; Lansky et al., 1985), similar to that of the general population (Locke & Regier, 1985). However, when major depression and depressed mood are considered, prevalence rates are higher (e.g., 16% in Derogatis et al., 1983; 25% in a review by Massie & Holland, 1990). Lower rates have been found when specified diagnostic criteria are used and/or the patients assessed are ambulatory with good physical functioning (Levin, Jones, & Sack, 1993; Massie & Holland, 1990). In general, depression is more common for those patients in active treatment rather than those on follow-up, receiving palliative rather than curative treatment, with pain or other disturbing symptoms rather than not, and/or with a history of affective disorder or alcoholism.

Study of the psychological responses that characterize the response to cancer diagnosis may shed light on the reactions of individuals to other life-threatening disease or treatment circumstances. In fact, when moods (Cassileth et al., 1985; Westbrook & Viney, 1982) or coping strategies (Felton, Revenson, & Hinrichsen, 1984) have been compared across disease groups (e.g., rheumatoid arthritis, diabetes, cancer, renal disease, hypertension), few differences have been found, suggesting that cancer patients may respond similarly to those with other life-threatening and/or chronic illnesses. On the other hand, patients in these early studies were heterogeneous on many disease variables (e.g., time since diagnosis, active treatment vs. follow-up status, and/or disability) that are important predictors of differential responses among cancer patients.

Differences in the patterns of emotional distress may exist. For example, it has been suggested that breast and gynecological cancer patients are significantly less depressed, anxious, and hostile (angry) possibly due to perceptions of their illness as less serious than women with other forms of cancer (e.g., lung, gastrointestinal, head and neck; Sneed, Edlund, & Dias, 1992). Similarly, a comparison of patients with limited vs. extensive lung cancer and equivalent levels of physical impairment found that significantly greater mood disturbance was reported by the patients with extensive disease and poorer prognoses (Cella et al., 1987). Related findings have emerged for individuals who perceive their illness as severe (Marks, Richardson, Graham, & Levine, 1986) or who have a sense of pessimism about their life (Carver et al., 1994). In fact, Carver et al. found that general pessimism (vs. optimism) at diagnosis predicted poorer well-being (mood and life satisfaction) at one day before surgery as well as at 3, 6, and 12 month follow-ups. In addition, when patients are followed longitudinally, the emotional responses at recurrence are characterized by significantly higher levels of depression and anger than had been reported at the time of initial diagnosis (Andersen, Anderson, & deProsse, 1989b).

It is encouraging that the clinical problem of diagnostic and treatment-related distress can be alleviated through psychological interventions (see Andersen, 1992, for a review). A comprehensive example of such an effort is the study of Fawzy, Cousins et al. (1990) and Fawzy, Kemeny et al. (1990). They attempted to reduce distress and enhance immune functioning in newly diagnosed melanoma patients via a structured group support intervention which included health education, illness-related problem solving, relaxation training, and group support. The format was weekly group treatment for six sessions. Eighty patients with early stage melanoma participated and were randomized to intervention or control conditions. The post-treatment analyses indicated that the intervention subjects reported significantly more vigor, but there were no other emotional distress differences. By six months, emotional distress had improved for the intervention subjects with significantly lower depression, confusion, and fatigue, and higher vigor. Coping data indicated that the intervention subjects reported significantly more use of active-behavioral strategies by the treatment's end, a pattern that continued with the addition of active-cognitive strategies by six months. Regarding the immunologic findings, at six months there was a significant difference in groups with better immunologic status for the intervention subjects. These data are impressive in that they suggest that improvements in psychological status and coping are associated with changes (upregulation) in immune responses.

Improvements such as these in mood and coping are all the more impressive because they are often achieved with brief, cost-effective interventions (e.g., 10 therapy hours with delivery in a group format; Andersen, 1992). Studies that have provided follow-up data also suggest some consolidation of intervention effects across time (upwards of six months post-treatment), with lowered emotional distress, enhanced coping, and/or improved sexual functioning (e.g., Capone, Good, Westie, & Jacobson, 1980; Christensen, 1983).

8.09.4.3 Treatment

A certain component of the emotional distress occurring at diagnosis is due to the anticipation of treatment. Current therapies include surgery, radiotherapy and radioactive substances, chemotherapy, hormonal therapy, immunotherapy, and combination regimens and procedures (e.g., bone marrow transplantation, intraoperative radiotherapy). Some patients also undergo difficult diagnostic or treatment monitoring procedures (e.g., bone marrow aspirations), and all treatments are preceded or followed by physical examinations, tumor surveys, and/or laboratory studies. Thus, the diagnostic process of selecting the appropriate therapy and subsequent treatment events can represent multiple medical stressors. As will be discussed below, the data are consistent in their portrayal of more distress (particularly fear and anxiety), slower rates of emotional recovery, and, perhaps, higher rates of other behavioral difficulties (e.g., food aversions, continued fatigue and malaise) in cancer patients compared with healthy individuals undergoing noncancer-related medical treatment. The emotional crisis that characterizes the diagnostic period lessens as time passes, and longitudinal studies find that as treatments end and recovery begins there is an emotional rebound (e.g., Andersen et al., 1989b; Bloom, 1987; Devlen, Maguire, Phillips, Crowther, & Chambers, 1987; Edgar, Rosberger, & Nowlis, 1992). This lowering of emotional distress over time is found even for patients undergoing radical treatment requiring major adjustments (e.g., radical neck dissection with laryngectomy; Manuel, Roth, Keefe, & Brantley, 1987). As cancer treatments (surgery, radiation, etc.) vary considerably in their intent, morbidity, and mortality, we will review each of the major modalities separately and also discuss three clinical problems—compliance, appetite and weight loss, and fatigue—that are common across therapies.

8.09.4.3.1 Surgery

There have been few investigations of cancer surgery, but there are numerous descriptive and intervention studies of the reactions of healthy individuals undergoing surgery for benign conditions. The latter studies are consistent in their portrayal of (i) high levels of self-reported preoperative anxiety predictive of higher postoperative anxiety, and (ii) higher postoperative anxiety predictive of longer recovery (e.g., decreased time out of bed, increased pain reports). What may distinguish cancer surgery patients are higher overall levels of distress and slower rates of emotional recovery. For example, Gottesman and Lewis (1982) found greater and more lasting feelings of crisis and helplessness among cancer patients in comparison to benign surgery patients for as long as two months following discharge.

Considering these data, findings on the interaction patterns of physicians and cancer patients on morning surgical rounds is disturbing. Blanchard, Ruckdeschel, Labrecque, Frisch, and Blanchard (1987) found attending physicians on a cancer unit less likely to be supportive and to address patients" needs than physicians treating general medical patients. The heavier volume and more seriously ill patients common to cancer units might be sources for this unfortunate relationship. Related findings indicate that oncology nurses might find their job significantly more stressful than other taxing assignments (e.g., cardiac, intensive care, or operating room; Stewart, Meyerowitz, Jackson, Yarkin, & Harvey, 1982), and may limit their communication with cancer patients, especially those experiencing a recurrence (Wilkinson, 1991).

As noted above, there has been considerable research on the psychological and behavioral aspects of response to surgery, and many effective interventions have been tested. Components of these interventions include procedural information (e.g., how the surgery is to be performed as well as pre- and postoperative events from the perspective of the patient), sensory information on the actual physical sensations of the surgery or preparatory events, behavioral coping instructions, cognitive coping interventions, relaxation, hypnosis, and emotion-focused interventions. In a meta-analysis of this literature, Johnston and Vogele (1993) reported that procedural information and behavioral instructions show consistent and strong positive effects on postoperative recovery. Effects are significant for a broad band of measures, including ratings of negative affect and pain, amount of pain medication, length of stay, behavioral recovery, and physiological indices.

8.09.4.3.2 Radiotherapy

At least 350 000 individuals in the USA receive radiation therapy each year. Clinical descriptions have noted patients" fears (e.g., being burned, hair loss, sterility); while such outcomes do occur, they are site- and dosage-dependent. To understand radiation fears, the surgical anxiety studies already described above have been used as a research paradigm. Here, again, anxiety (and sleeplessness) can often cause more overall distress than physical symptoms (Munro, Biruls, Griffin, Thomas, & Vallis, 1989), and are predictors of treatment response (Wallace, Priestman, Dunn, & Priestman, 1993). If interventions to reduce distress (especially anticipatory anxiety) are not conducted, heightened post-treatment anxiety is also found (Andersen, Karlsson, Anderson, & Tewfik, 1984; Andersen & Tewfik, 1985) and might be maintained for as long as three months post-therapy, particularly when treatment symptoms linger (e.g., diarrhea, fatigue; King, Nail, Kreamer, Strohl, & Johnson, 1985). When acute side effects resolve (usually by 12 months post-treatment), there appears to be no higher incidence of emotional difficulties for radiotherapy patients than for cancer surgery patients (Hughson, Cooper, McArdle, & Smith, 1987).

8.09.4.3.3 Chemotherapy

Of all the treatment modalities, the greatest progress has been in the understanding of psychological reactions to chemotherapy and the related toxicities, particularly nausea and vomiting. Over a decade of research has revealed that approximately 60% of cancer patients will develop nausea and another 50% will develop vomiting in response to cytotoxic treatments (Morrow & Hickok, 1993). Psychological research has focused on eliminating or reducing these problems through hypnosis (e.g., Redd, Andersen, & Minagawa, 1982), progressive muscle relaxation with guided imagery (e.g., Burish & Jenkins, 1992; Carey & Burish, 1988), systematic desensitization (e.g., Morrow et al., 1992), cognitive distraction (e.g., Greene, Seime, & Smith, 1991; Vasterling, Jenkins, Tope, & Burish, 1993), and biofeedback (e.g., Burish & Jenkins).

The routine use of antiemetic and antianxiety drugs and the development of newer chemotherapeutic agents has resulted in an overall lower incidence and clinical severity of nausea and vomiting. However, the above gains may be somewhat offset by the more current medical practice of using dose-intensive regimens as well as treating disease types or stages not previously given adjuvant chemotherapy (Andrykowski, 1993). It is clear that pretreatment with medications can control nausea and thus reduce vomiting, but pretreatment can also reduce the likelihood of anticipatory reactions. Anticipatory reactions are believed to be conditioned responses resulting from the learning process of classical conditioning in which side effects (e.g. nausea and vomiting) become associated with previously neutral stimuli (driving to the hospital, sitting in the waiting room, etc.). Once anticipatory nausea and vomiting develops, antiemetics are less effective, and it has been suggested that behavioral strategies should be used to aid patients (Andrykowski, 1993; Morrow & Hickok, 1993; Redd, 1993). Other research that may be helpful includes identifying the characteristics of individuals (e.g., high pretreatment anxiety or autonomic activity; severity of post-treatment vomiting in the early cycles; age; susceptibility to motion sickness or gastrointestinal distress), as well as characteristics of the agents (e.g., rated emetogenic potential of the regimen, dosages, rate of infusion) that place patients at risk of the development of nausea and vomiting (see Carey & Burish, 1988, or Morrow, 1986, for empirical tests of variables; see Morrow & Hickok, 1993, for a review of the potential variables).

8.09.4.3.4 Bone marrow transplantation

Over the past two decades, bone marrow transplantation (BMT) has evolved from an experimental procedure performed only when conventional therapies have failed (e.g., recurrent breast cancer) to an effective therapy considered as an alternative worthy of study in clinical trials as a first-line therapy (Baker, Curbow, & Wingard, 1991). It was initially used only for those with lymphohematopoietic malignancies or disorders characterized by bone marrow failure (e.g., leukemia, Hodgkin's disease, and non-Hodgkin's lymphoma; Wingard, Curbow, Baker, & Piantadosi, 1991), but is now being used with a variety of solid tumors (e.g., neuroblastomas, germ cell tumors, and cancers of the breast and ovary). The rapid acceptance and expanded use of this toxic modality is no less than dramatic. For example, a Phase II study has begun in the USA for BMT as an adjuvant treatment for breast cancer patients with 10 or more positive lymph nodes (Winer & Sutton, 1994). If beneficial, it is possible that thousands of breast cancer patients may undergo BMT as a standard treatment for metastatic as well as regional disease in the near future (Winer & Sutton, 1994). At present it is estimated that there are over 5000 transplants performed each year worldwide (Alby, 1991).

BMT is a complex and potentially fatal treatment. The target of the treatment, the bone marrow, is destroyed with high dose chemotherapy, with or without whole body radiation afterwards. The bone marrow to be transplanted comes from a donor (allogenic BMT) or from the patient (autologous BMT) after it has been removed and treated. Although allogenic BMT has a role, its expansion is limited by the need for a suitable donor and the subsequent risk of graft vs. host disease for the patient. In the case of autologous BMT, a variety of *ex vivo* procedures are used to destroy the malignant cells in the patient's marrow prior to reinfusion, including treatment with cytotoxic drugs, exposure to monoclonal antibodies that will attack tumor-associated antigens, or harvesting and introducing stem cells (cells from peripheral blood).

Patients (and their families) are faced with a number of stressors: a life-threatening illness, location of a suitable donor (for allogeneic transplants see Alby, 1991; Patenaude, 1990), a toxic treatment, and common and potentially fatal side effects (e.g., liver failure). There are many acute side effects as well (e.g., hair loss, mouth and gastrointestinal mucositis, infertility, skin breakdown, infection, pneumonia). Hospitalization is prolonged (often three to six weeks or longer) and it is generally spent in isolation (Alby, 1991; Baker et al., 1991; Winer & Sutton, 1994). Of course, in the end, the treatment can fail and the cancer persist or rapidly recur (Patenaude, 1990; Winer & Sutton, 1994).

The many difficulties—toxicity, uncertainty, illness and isolation, dependency, constant need for care, expense—all contribute to patients feeling out of control (helplessness), alone, anxious, and/or depressed (Alby, 1991; Altmaier, Gingrich, & Fyfe, 1991; Brown & Kelly, 1976). In attempting to cope, patients may be demanding, or the converse, withdrawn (Brown & Kelly, 1976; Patenaude, 1990). Psychological efforts have focused on providing support to patients, their families, and staff (Alby, 1991), and maximizing control for patients, such as making choices about the hospital environment whenever possible (Patenaude, 1990). For example, BMT rooms are often equipped with televisions/radios and weights and/or bicycles are provided for exercise as patients become stronger. Additionally, patients have been encouraged to bring personal items from home (e.g., pictures, photographs) to decorate their rooms. Increased medical complications, age, distance from home, and poorer pretransplant psychosocial adjustment are associated with the need for more intensive psychological consultation (Futterman, Wellisch, Bond, & Carr, 1991).

8.09.4.3.5 Treatment side effects

Important steps have also been made toward understanding the etiology and prevention of at least three common disease/treatment-related complications. First, appetite and weight loss are significant clinical problems for cancer patients susceptible to tumor-induced metabolism or taste changes, having tumor-related obstructions (often diagnosed as primary cachexia/anorexia), or receiving gastrointestinal-toxic chemotherapy or abdominal radiotherapy (secondary cachexia/anorexia). Malnutrition is associated with increased morbidity and mortality (Knox, 1991). Approximately 50% of the above patients will develop food aversions, and have two or three aversions to particular foods; common aversions involve protein sources such as meats, eggs, and dairy products (see Jacobson & Schwartz, 1993, for a review). In particular, learned food aversions due to chemotherapy appear to be robust, with rapid acquisition (usually after one to three treatments) and maintenance after long delays (e.g., 48 hours) between food intake and aversive reactions (nausea) from the drugs (Bernstein, 1986; Jacobson & Schwartz, 1993). Research in this area has pointed the way, for example, to interventions employing novel tastes or "scapegoat" foods (e.g., lemon-lime Kool-Aid, unusually flavored hard candies such as coconut; Jacobson & Schwart, 1993) to "block" conditioning to familiar diet items, reducing food and beverage intake prior to drug administration, and ingesting carbohydrate rather than protein source meals. While food aversions may not involve appetite or weight loss, patients may unknowingly develop aversions to their favorite foods and this can affect their daily routine and perceived quality of life (Jacobson & Schwartz, 1993; Knox, 1991).

Second, fatigue is a problem reported by the majority of patients receiving radio- or chemotherapy (Nail & King, 1987; Smets, Garssen, Schuster-Uitterhoeve, & deHaes, 1993). Fatigue, and related symptoms described by patients as tiredness, lack of energy, sleepiness, confusion, and poor concentration have been related to cancer morbidity and poor treatment compliance. Although a common experience, few studies have been conducted on the correlates of fatigue (see Irvine, Vincent, Bubela, Thompson, & Graydon, 1991, for a review; Pickard-Holley, 1991; see Smets et al., 1993, for a review). Irvine et al. (1991) note that research has not demonstrated consistent relationships between fatigue, anemia, sleeplessness, and/or psychological distress (e.g., negative mood/depression), but it appears that fatigue does reduce overall daily functioning. Psychological and behavioral interventions have focused on alleviating or increasing tolerance to fatigue through preparatory information on side effects and activity–rest cycle recommendations (e.g., naps in the afternoon; Nail & King, 1987; Smets et al., 1993), as well as other activity planning or coping efforts (e.g., planning/scheduling activities, decreasing nonessential activities, and relying on others for assistance as needed; Rhodes, Watson, & Hanson, 1988).

The expectation and/or experience of unpleasant side effects can compromise a patient's quality of life such that the patient may miss treatment appointments and be unwilling or unable to continue treatment, regardless of whether it is curative or palliative (Morrow & Hickok 1993; Richardson, Marks, & Levine, 1988). With noncompliance, the dosage of treatment may be reduced, which can, in turn, lower the cure rate. In short, noncompliance is a behavioral problem that can have a direct impact on the effectiveness of cancer therapy.

Noncompliance with treatment has been related to increased emotional distress (e.g., Gilbar & Kaplan De-Nour, 1989; Richardson et al., 1988), severity of treatment side effects (e.g., nausea and vomiting, Lewis, Linet, & Abeloff, 1983; Richardson et al., 1988), and lower income (Lebovits et al., 1990). Even when patients are responsible for self-administration of therapy, such as taking their chemotherapy at home in order to reduce the number of required, but inconvenient, hospital visits, their noncompliance may continue. One report of multidrug therapy with adults having hematologic malignancies indicated that self-reports (vs. sera reports) overestimated compliance by a factor of two (Richardson et al., 1987).

Psychological interventions have focused on a variety of methods to improve patient compliance, including appointment reminders, clearly written and specific treatment communications (Anderson & Kirk, 1982), home visits, and medication-taking shaping interventions (Richardson et al., 1987), as well as hospital interventions such as offering a tour of the oncology clinic, videotape presentations about the therapy, discussion/question sessions, and take-home information (Burish, Snyder, & Jenkins, 1991). Burish et al. and others (e.g., Nail & King, 1987; Rainey, 1985) suggest that preparatory information can improve coping with treatment.

Importantly, psychological interventions can reduce distress during and immediately following cancer treatments. We now present the findings from two illustrative studies. Cain, Kohorn, Quinlan, Latimer, and Schwartz (1986) compared individual and group therapy formats with 72 women with recently diagnosed gynecologic cancer or who were currently receiving therapy for such a cancer. The intervention had eight components including discussion of the causes of cancer at diagnosis, impact of the treatment(s) on body image and sexuality, relaxation training, emphasis on good dietary and exercise patterns, communication difficulties with medical staff and friends/family, and setting goals for the future to cope with uncertainty and fears of recurrence. The eight-session program was conducted during individual sessions conducted in the hospital or the woman's home or in weekly groups of four to six patients conducted at the hospital. Outcome measures were administered pre- and post-treatment and at a six month follow-up. Post-treatment analyses indicated all groups improved with time; however, anxiety was significantly lower for the individual therapy subjects only. Gains for the intervention subjects were more impressive with the six month follow-up data when there were no differences between the intervention formats, but both groups reported less depression and anxiety and better psychosocial adjustment (including health perspectives, sexual functioning, and use of leisure time) than the no treatment control group. Thus, the brief intervention, delivered either in individual or group format, appeared to be immediately effective, with gains enhanced during the early recovery months.

Edgar et al. (1992) provided a psychosocial intervention to 205 patients who were randomized into two groups: one group received the intervention soon after entering the study and the other group after a wait of four months. The majority ($n = 159$) of patients had been diagnosed with breast, colon, lung, uterine, or head and neck cancers. Both groups were assessed on depression, anxiety, illness worry, perceived personal control, and ego strength at entry into the study, and at 4, 8, and 12 month intervals. The intervention (comprising five one-hour sessions) focused on coping skills and included problem solving, goal setting, cognitive reappraisal, relaxation training, and, at four month intervals, workshops on healthcare information and available resources. While coping improved over the year for all patients, the later intervention group experienced lower levels of distress sooner than the group who had received the intervention earlier. Edgar et al. (1992) suggest that the delay in intervention may have been beneficial because it afforded patients the time to reduce feelings of being overwhelmed and they were, therefore, more ready to participate in the intervention. Additionally, data indicated that patients with lower ego strength and cancer diagnoses other than breast cancer reported greater distress and, therefore, particularly benefited from the intervention, while patients with higher ego strength tended to cope well regardless of the intervention. These and related data (e.g. Forester, Kornfield, & Fleiss, 1985; see Andersen, 1992, for a review) attest to the significant distress that occurs when patients are in the midst of coping with diagnosis and treatment. Importantly, impressive gains can be achieved by the end of the intervention and often these positive gains are stronger with continued follow-up. Again, these effects have been achieved with brief interventions.

8.09.4.3.6 Choosing cancer treatments

Psychological and behavioral data have been (and should continue to be) important to patients and physicians alike for making choices among comparable treatments. Treatments that result in less quality of life disruption often become standard treatment. The most obvious example of the importance of psychological data influencing cancer treatments was that documenting the more positive outcomes (e.g., feeling more attractive, fewer sexual difficulties) for women treated with breast-saving procedures (lumpectomy plus adjuvant radiation and/or chemotherapy) rather than modified radical mastectomy (e.g., Margolis, Goodman, & Rubin, 1990). Similarly, lower rates of erectile and ejaculation difficulties (e.g., 30% vs. 90%) are important reasons for some men to choose supervoltage irradiation rather than surgery for prostate cancer (radical prostatectomy; see Andersen & Lamb, 1995, for a review). Parallel treatment data are available for women with genital cancer including those with uterine disease (Andersen, Anderson, & deProsse, 1989a) or preinvasive vulvar disease (Andersen, Turnquist, LaPolla, & Turner, 1988). Additionally, Rathmell, Ash, Howes, and Nicholls (1991) suggest that unless a survival advantage is demonstrated, patients with advanced head and neck cancer may want radiotherapy alone vs. surgery plus radiotherapy, as the treatment combination is associated with poorer quality of life scores (e.g., psychological well-being, speech quality, ability to eat, levels of energy and activity).

8.09.4.4 Recovery and Long-term Survival

The most important cancer end points have been treatment response rates, length of disease-free interval, and survival. Yet, as the prognosis for some sites has improved, there has been increased attention to the quality of life, particularly for long-term survivors of cancer. The term "survivor" typically refers to individuals surviving at least five years, as the probability of late recurrence declines significantly after that time for most sites. As individuals recover and resume their life patterns, there may be residual emotional distress, with some difficulties which require continued coping, and even the appearance of still other problems (late sequelae). We will discuss each of these circumstances, and provide examples of both emotional and physical challenges which may confront the cancer survivor.

An investigation by Dunkel-Schetter, Feinstein, Taylor, and Falke (1992) sheds light on the strategies cancer patients use as they recover and resume their life activities. They studied coping patterns of 603 cancer patients. Patient diagnoses included breast cancer, gastrointestinal, circulatory or lymph, gynecological, and respiratory cancers, as well as others. The timing of the assessment ranged from initial diagnosis to more than five years post-treatment. Five coping patterns were identified: seeking or using social support, focusing on positive aspects, distancing, cognitive escape-avoidance (e.g., fantasizing or wishful thinking with fatalistic thoughts of poor outcomes), and behavioral escape-avoidance (e.g., social withdrawal, drug use). All patients used multiple coping strategies, but distancing was the most common. Unique patterns were found for subgroups. For example, individuals who viewed cancer as more stressful tended to use cognitive and behavioral escape-avoidance strategies, whereas patients reporting less distress relied on seeking or using social support, focusing on the positive, and distancing. Data such as these may have many uses, such as being used to tailor psychological interventions to patient subgroups (e.g., identifying those patients who are likely to use drugs/alcohol to cope).

Lingering emotional distress from the trauma of diagnosis, treatment, and, more generally, life threat, may occur for a small subset, 5–10%, of cancer patients. When pronounced, this long-term distress has been likened to post-traumatic stress disorder. In fact, having residual distress from the diagnosis and treatment of a life-threatening illness is included as one of the circumstances which may precipitate a *Diagnostic and statistical manual of mental disorders* (American Psychiatric Association, 1994) diagnosis of post-traumatic stress disorder. It is unlikely that such extreme distress occurs for the "average" cancer patient, and instead may only be a possibility for those who have undergone the most difficult of treatment regimens (e.g., BMT), multiple agent and/or lengthy chemotherapy regimens, as might be given to Stage III ovary patients, or those who undergo life-altering and/or disfiguring cancer treatments (e.g., limb amputations, pelvic exenteration, laryngectomies). As is the case with anxiety disorders, individuals with a prior history of traumatic stress might also be at heightened risk.

Secondly, some cancer survivors may need to cope with the expected, but nevertheless troubling, sequelae which may be consequences of the disease or treatment and be permanent. For example, coping with altered abilities (e.g., loss of natural speech following laryngectomy) or changes in organ functions (e.g., infertility) may require adjustment which demands new behaviors or emotions. Others may have to cope with losses (e.g., a sexual relationship that does not include intercourse). Thirdly, late side effects of cancer treatment, for example, a bowel dysfunction that is traced to pelvic radiotherapy, can occur and change health status as well as impact mood and coping.

Despite these possibilities, longitudinal data indicate that if the disease is controlled, by one year post-treatment the severe distress of diagnosis will have dissipated and emotions will have stabilized. The first longitudinal studies conducted in the UK for breast cancer patients indicated that by 12 (Maguire et al., 1978) and 24 months (Morris, Greer, & White, 1977) that approximately 20% of the patients had problems with moderate to severe depression in comparison to 8% of benign disease comparison subjects. However, controlled longitudinal studies of breast (Bloom, 1987; Vinokur, Threatt, Caplan, & Zimmerman, 1989) and gynecologic (Andersen et al., 1989b) patients conducted in the USA and replicated with data from The Netherlands (deHaes, van Oostrom, & Welvaart, 1986) have indicated no differences between the levels of emotional distress of women with cancer and either benign disease or healthy comparison subjects. Similar declines and lowered levels of distress have been found in retrospective (Cella & Tross, 1986) and longitudinal (Devlen et al., 1987) studies of Hodgkin's disease and non-Hodgkin's lymphoma patients. The consistency of findings for the studies conducted during the 1980s is important because it represents replications across site, and, to some degree, treatment toxicity. In sum, we preface the remaining discussion by noting that global adjustment problems do not occur for the majority of cancer survivors; a more likely scenario is the occurrence of problem areas.

Irvine, Brown, Crooks, Roberts, and Browne (1991) concluded in their review of survivorship following breast cancer that few patients experience long-term psychological distress, yet 20–30% will experience loss of roles (e.g., employment), decreased functional abilities, and problems with social relationships. Data on the adjustment of BMT survivors reveal a somewhat slower recovery (Syrajala, Chapko, Vitaliano, Cummings, & Sullivan, 1993). Comparison of BMT survivors and other cancer patients on maintenance chemotherapy indicated that psychological functioning was satisfactory for both groups when assessed three to four years postdiagnosis (Altmaier et al., 1991). However, when BMT patients were assessed sooner (two or four years post-treatment), they reported poorer physical functioning, greater impaired personal functioning (e.g., need for self-care assistance), and more relational problems (e.g., sexual difficulties) than individuals on maintainence chemotherapy (Altmaier et al., 1991). Also, there are data to suggest that there may be some risk of neuropsychological impairment from BMT procedures. Data from Andrykowski et al. (1992) suggest impairments in memory and higher cognitive processing, which may be a sequela of BMT, *per se*, or prior cancer treatments which precede BMT (e.g., cranial radiation, intrathecal chemotherapy). The many difficult aspects of BMT noted above (e.g., intensive chemotherapy, possible whole body irradiation, long recovery time, isolation) may account for this slowed and potentially more problematic recovery (see Andrykowski, 1994; or Winer & Sutton, 1994 for a review).

Data suggest that sexuality may be one dimension that is more likely to undergo disruption than other major life areas. All cancer patients with solid tumors (approximately 85% of adult patients) are vulnerable to sexual dysfunction. Across sites, estimates range from 10% (e.g., breast cancer patients treated with lumpectomy), 70–90% (e.g., women with vulva cancer treated with modified radical vulvectomy), to 100% (e.g., men with prostate cancer treated with radical prostatectomy), with the distribution skewed toward greater levels of disruption (see Andersen & Lamb, 1995, for a review). Among the hematologic malignancies, such as Hodgkin's disease, estimates are in the range of 20% (Andersen & Lamb, 1995). The data within treatment sites indicate that disease and treatment factors are the primary etiologies for sexual problems (e.g., gynecologic complications associated with long-term adjuvant tamoxifen therapy for breast cancer, Wolf & Jordan, 1992; female androgen deficiency syndrome resulting from cytotoxic agents and/or bilateral salpingo-oophorectomies, Kaplan & Owett, 1993). Controlled longitudinal studies of breast cancer patients (Maguire et al., 1978; Morris et al., 1977) and gynecologic cancer patients (Andersen et al., 1989a) have indicated that if sexual problems develop, they do so as soon as intercourse resumes, and, if untreated, they are unlikely to resolve (see Andersen & Elliott, 1994, for a discussion).

In addition to medical contributors, we have focused on identifying sexual relevant psychological factors in predicting risk for sexual morbidity. Specifically, sexual self-schema (self-concept) is a cognitive view about sexual aspects of oneself; it is derived from past experience, manifest in current experience, and it guides the processing of domain-relevant social information (Andersen & Cyranowski, 1994). The concept includes two positive aspects—an inclination to experience romantic/passionate emotions and a behavioral openness to sexual experiences and/or relationships; and a negative aspect—embarrassment and/or conservatism—which appears to be a deterrent to sexual expression. Using a measure developed for this purpose, the predictive power of sexual schema was examined in the context of risk for sexual morbidity following gynecologic cancer. Consistent with our definition of the construct, we anticipated that women with a negative sexual self-concept, in contrast to women with a positive sexual self-concept, would be at greatest risk for sexual difficulty. Negative sexual schema women would be expected to have more difficulties because they are, in general, less romantic/passionate in their emotions, less open to sexual experiences, and more likely to have negative feelings about their sexuality.

In a field test of sexual schema in the prediction of gynecologic cancer outcomes (Andersen, Woods, & Copeland, 1997), we assessed 62 women who were currently disease free but who received treatment from six months to five years previously for Stage I–II disease. Comparison subjects included 68 women seeking routine gynecologic care. Analysis of the quality of life data replicated earlier prospective longitudinal findings (Andersen, Andersen, & de Prosse, 1989a)—specifically, sexuality is the major life area of disruption for the survivor. For example, there were no differences between the groups in the areas of mental health (emotional distress, depression) or social functioning. In contrast, a comparison between the samples in terms of current sexual functioning found significant differences, with the cancer sample reporting lower levels of sexual behavior, sexual responsiveness, and global evaluations. We tested the utility of the schema construct with the cancer sample in the prediction of sexual responsiveness (e.g., desire, excitement, orgasm, and resolution) and sexual behavior (frequency of intercourse) with regression analyses. Sexual self-schema accounted for a significant and large portion of the variance (26%) in the prediction of current sexual responsiveness.

The high incidence of sexual and fertility disruption has, in part, been the reason for the concern over marital disruption among adult cancer patients. An early clinical study of women receiving radical mastectomy noted the realistic feelings of body disfigurement that both the women and spouses would feel—prompting sexual retreat, emotional estrangement, and, not surprisingly, marital disruption (Bard & Sutherland, 1952). Other concerns over the marriage originate from analyses of the interpersonal relationships, *per se*, (Wortman & Dunkel-Schetter, 1979). Despite the emotional distress and, for some, accompanying sexual disruption that couples experience, data from retrospective studies with comparison groups (Cella & Tross, 1986) and from the controlled longitudinal studies previously discussed indicate that the majority of marriages remain intact and satisfactory. These data are consistent with prospective studies showing that, when health problems arise for newly married couples, they are not among those problems precipitant to divorce (Bentler & Newcomb, 1978). However, multicenter studies have indicated that young survivors of childhood cancers are significantly less likely to marry, and once married they may be at greater risk for subfertility (Byrne et al., 1985, 1988; Teeter, Holmes, Holmes, & Baker, 1987).

Yet, the cancer experience is stressful for those closest to the patient (see Baider, Cooper, & Kaplan De-Nour, 1996, for a review), and the kin's distress may approach that of the patient's (Baider & Kaplan De-Nour, 1988; Cassileth et al., 1985). Family strain appears to be affected by illness variables (e.g., prognosis, stage/duration of illness, care-giving demands, patient's distress), family variables (e.g., age and gender of family members, socioeconomic status, other family stressors), and relational variables (e.g., quality of marriage, marital communication, family stage, and social support; see Sales, Schulz, & Biegel, 1992, for a review). Those couples in which the wife/mother has cancer and young children are still in the home may be at heightened risk for relationship difficulties (Vess, Moreland, & Schwebel, 1985). In addition, Ell, Nishimoto, Mantell, and Hamovitch (1988) found that those kin who were functioning poorly (e.g., lower perceived personal control, less adequate emotional support from close others, and greater stress unrelated to cancer) when the patient was diagnosed, or who lost personal and social resources during the patient's treatment and recovery, tended to function poorly at follow-up. In sum, a subset of partners and family members appear to be at psychosocial risk. For a review of strategies (focusing on information and support) that have been used to assist spouses of cancer patients cope with stress over the illness course, see Northouse and Peters-Golden (1993).

Finally, other, nonpsychological difficulties may occur following cancer, and these problems may have psychological concomitants (e.g., low self-esteem, perceived worthlessness). For example, it has been found that survivors of childhood or adolescent cancer are at risk of rejection from the armed services or college entrance and have difficulty obtaining health and life insurance (Teta et al., 1986). Job discrimination and insurance difficulties have also been reported by survivors of bone marrow transplantation (Wingard et al., 1991).

There have been few interventions targeted for cancer survivors *per se*. One such study was that by Telch and Telch (1986). They compared the effectiveness of coping skills instruction vs. supportive therapy for a heterogeneous sample of cancer patients on follow-up. An important research design aspect was the screening of the subjects, and only those with "clear evidence of psychological distress" were included in the trial. Both interventions were offered in a group format. Instruction in coping skills taught cognitive, behavioral, and affective strategies and included goal setting, self-monitoring, and role playing. Relaxation training and stress management skills were also included and patients provided ratings of their home practice. The group support intervention provided an environment for patients to discuss concerns but there was no specific agenda. Each group met for six weeks. Forty-one cancer patients completed the study. Analyses for the emotional distress data indicated that the coping skills group improved significantly across all measures, the support group improved on anxiety and depression only, but the no treatment control worsened and reported significantly more mood distress.

8.09.4.5 Recurrence and Death

Cancer recurrence is devastating; the magnitude of distress is even greater than that found with the initial diagnosis (Mahon, Cella, & Donovan, 1990; Thompson, Andersen, & DePetrillo, 1992), and studies contrasting cancer patients showing no evidence of disease with those receiving palliative treatment (e.g., Cassileth et al., 1985) have reported the greatest distress for those with disseminated disease (Bloom, 1987). Patients recently told of their cancer recurrence reported being less hopeful, more discouraged, having increased thoughts of death and dying as well as feelings of guilt and regret about their previous treatment decisions. In one study (Mahon et al., 1990), 45% of the patients indicated that they perceived their recurrence as "punishment." Therefore, difficult decisions (e.g., beginning a regimen that offers little chance for cure and has side effects vs. no treatment) are made in a context of extreme emotional distress and physical debilitation. However, the few studies of psychological interventions for adult patients have indicated that important emotional gains can be achieved during terminal stages (Linn, Linn, & Harris, 1982) and that children and adolescents, as well as adults, can make independent decisions about the continuation of therapy when death is imminent (Nitschke et al., 1982).

At this time of significant emotional turmoil and physical difficulty, psychological interventions appear to enhance the quality of life. One example is an investigation of a group support intervention for women with breast cancer conducted by Spiegel and colleagues (Spiegel & Bloom, 1983; Spiegel, Bloom, & Yalom, 1981). Women were randomized to no treatment or a group treatment intervention that included discussion of death and dying, family problems, communication problems with physicians, and living fully in the context of a terminal illness. The intervention subjects were also randomized a second time to two conditions: no additional treatment or self-hypnosis for pain problems (Spiegel & Bloom, 1983), which was incorporated into the support group format. All intervention groups met for weekly meetings for one year, for a total of 75 therapy hours. At the end of the first year the groups formally ended, but members could continue to meet as they wished or were able; some groups lasted for an additional two years. Eighty-six women, 50 intervention and 36 no treatment control, with metastatic breast cancer and referred to the intervention by their oncologist, participated. Following random assignment, there was subject loss (e.g., refusal, too weak, death) with the study beginning with 34 intervention and 24 control participants; however, the survival data are reported for the original sample of 86. Analyses indicated that the intervention group reported significantly fewer phobic responses and lower anxiety, fatigue, and confusion and higher vigor than the controls. These differences were evident at all assessments but the magnitude increased from 4 to 12 months. There was also a significant decrease in the use of maladaptive coping responses by the intervention group.

Regarding the findings from the hypnosis substudy, women receiving hypnosis within the group support intervention reported no change in their pain sensations during the year, while pain sensations significantly increased for the other women in group support who did not receive hypnosis. Similar findings were reported for pain suffering—a slight decrease for the women who also received hypnosis and a significant increase in suffering for the remaining intervention women. It is important to note that pain sensation scores for both groups were, however, significantly lower than those for the no intervention controls, suggesting that the hypnosis component provided an additive analgesic effect to other group treatment components. The most startling data from this project were reported in a survival analysis described at the beginning of the chapter. A variety of follow-up analyses, controlling for initial disease stage, days of radiotherapy, or use of androgen or steroid treatments were conducted and all indicated the same survival differences favoring the intervention participants. Unlike other studies, interventions for terminal patients tend to be intensive and lengthy, such as "several sessions" or "until death."

Finally, a frequent complication of disseminated disease is pain. Although it might be one of the first symptoms of cancer or be present when disease is localized, pain is more common and less controllable for those with metastatic disease (Ahles, Ruckdeschel, & Blanchard, 1984). Some 80% of patients with recurrent cancer report moderate to severe pain compared with 40% of earlier stage patients (see Ashburn & Lipman, 1993, for a review). The major cause of cancer pain, accounting for roughly 70% of the cases is due to direct tumor involvement (e.g., metastatic bone disease, nerve compression). Another 20–30% of cases are due to medical therapy (e.g., postoperative pain, radiation-induced pain). The remaining cases are individuals with pain problems unrelated to their cancer (e.g., osteoarthritis; Foley, 1985). Thus, the pain experienced by cancer patients can be any combination of acute malignant pain, chronic malignant pain, and/or chronic nonmalignant pain (Ashburn & Lipman, 1993).

As the pain experienced by each patient is variable (e.g., type, duration, resting vs. active, diagnosis/treatment factors), it is important that a pain assessment be completed before and during implementation of treatment for pain control (Ashburn & Lipman, 1993; Coyle, Adlehardt, Foley, & Portenoy, 1990). Besides pain-specific qualities and diagnostic/treatment considerations, other factors are also important in assessing the patient's experience of pain. For instance, pain is associated with depression, anxiety, and delirium (Massie & Holland, 1992), and may increase suicidal ideation (Coyle et al., 1990). If pain worsens or is difficult to control, quality of life deteriorates and emotional distress increases (Massie & Holland, 1992), physical mobility may lower (Ashburn & Lipman, 1993), and social interactions may decline (Strang & Qvarner, 1990).

The most difficult circumstance of pain control is when chronic pain accompanies disease progression. Treatment combinations of antitumor therapy, anesthetic blocks, and behavioral approaches are considered. Behavioral research has focused on assessment strategies (Daut, Cleeland, & Flanery, 1983; Keefe, Brantley, Manuel, & Crisson, 1985) and on pain reduction interventions, particularly hypnosis, to reduce pain (Spiegel & Bloom, 1983). When palliative therapy is of little use and/or brings further debilitation, psychological interventions may provide support and pain control and, secondarily, treat pain sequelae (e.g., sleep disturbances, reduced appetite, irritability).

8.09.5 CONCLUSION

Significant progress has been made in understanding the psychological, social, and behavioral aspects of cancer. There has been progress in prevention and control. Research on smoking prevention and cessation addresses the factor responsible for 30% of all cancer deaths. Similar efforts to encourage participation in screening and early detection programs for breast, cervix, skin, prostate, and colon and rectum cancers may also improve mortality rates.

More is known about the psychological processes and reactions to the diagnosis and treatment of cancer than is known about any other chronic illness. Breast cancer patients (see Glanz & Lerman, 1992, for a review) have been well studied, but other disease sites, men, and children, are becoming more commonly studied. Future research will probably test the generalizability of these descriptive data and formulate general principles of adjustment to illness. While providing estimates of the magnitude of quality of life problems, these data can be used for models that predict which patients might be at greatest risk for adjustment difficulties (see Andersen, 1996, for a discussion). The latter is an important step toward designing interventions tailored to the difficulties and circumstances of cancer patients.

The mental health community emphasizes the need to reduce stress and prevent deteriorations in quality of life for those with cancer. The importance of such efforts is underscored by three contextual factors. One, the stability of many cancer mortality rates, particularly those with the highest incidence such as lung and breast, makes it imperative that new, innovative treatments be developed to improve survival rates. Two, research has demonstrated that psychological interventions result in significant improvements in quality of life (see Andersen, 1992, for a review). Three, both qualitative (Maier, Watkins, & Fleshner, 1994) and quantitative (Herbert & Cohen, 1993a, 1993b) summaries of the psychoneuroimmunology literature conclude that psychological distress and stressors (e.g., negative life events, both acute and chronic) are reliably associated with changes—downregulation—in immunity. Thus, addressing the mental health needs of those with cancer will have important quality of life benefits and the possibility is raised of positive biologic (or health) consequences as well. Figure 1 provides a conceptual model of the psychological and behavioral factors and biologic mechanisms by which disease or health outcomes might be influenced. The majority of the paths move in one causal direction. We have discussed the importance of conceptual models in guiding future intervention research and in testing for any health benefits of psychological interventions (see Andersen et al., 1994).

A growing literature on the use of psychological interventions to improve the cancer patient's quality of life also exists (Andersen, 1992). The effectiveness of these interventions is robust, as they have reduced distress and enhanced the quality of life of many cancer patients differing on disease stage as well as disease site. Despite the challenges of studying these patients, well-controlled investigations have been conducted. Improvements in emotional distress are found at the end of the interventions with continued gains at follow-up. In addition, change in other areas—self-esteem/concept, death perceptions, life satisfaction, and/or locus of control—have been found. Important for quality of life, psychological interventions could also lower or stabilize pain reports. The positive outcomes for terminal patients are notable considering their worsening pain and/or increasing debilitation.

While there appear to be unique intervention components for different phases of the disease,

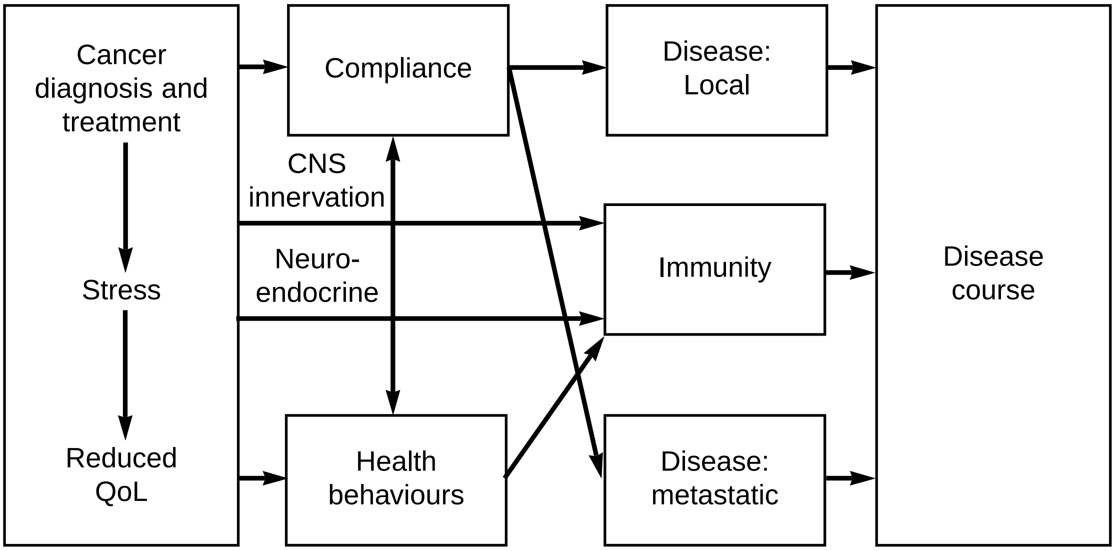

Figure 1 Theoretical model. A biobehavioral model of the psychological, behavioral, and biologic pathways from cancer stressors to disease course. CNS = central nervous system. From Andersen et al. (1994).

there are some commonalities. Therapy components have included: an emotionally supportive context to address fears and anxieties about the disease, information about the disease and treatment, behavioral coping strategies, cognitive coping strategies, and relaxation training to lower "arousal" and/or enhance one's sense of control. The descriptive data also highlight the need for focused interventions for sexual functioning, particularly those treated for gynecologic, breast, and prostate cancer, and the intervention studies attest to the effectiveness of these specific components. These various components appear more important to outcome than procedural variations. For example, therapy format, such as individual or group, appears to have little impact. Also, there were null findings for (group) interventions that included no structured content, suggesting that group support alone is insufficient to produce any measurable benefit (see Helgeson & Cohen, 1996, for a review).

How do psychological interventions achieve these effects? In large measure, the psychologic mechanisms may not be different from those operative in interventions designed for coping with other stressors. That is, confronting a traumatic stressor with positive cognitive states, active behavioral strategies and, eventually, lowered emotional distress may enhance one's sense of self-efficacy, feelings of control, and provide realistic stress appraisals of the disease or treatment process. Similarly for sexual interventions, information provides realistic expectations for sexuality and specific strategies to manage sexual activity when it is difficult or impossible. That the interventions produce more than situational improvement and may alter an individual's longer-term adjustment processes is suggested by data indicating that adjustment gains continue (and often increase) during the first post-treatment year. Immediate and longer-term psychologic changes may, in turn, increase the likelihood of changes in behavioral mechanisms, such as increasing the likelihood of adaptive health behaviors (complying with medical therapy; improving diet, exercise, etc.), to improving mental health directly, "adjustment" and, possibly, medical outcomes (e.g., disease-free interval). These data indicate that, increasingly, issues of quality of life are being raised, and positive results have been achieved by the behavioral scieces. But as with most issues, further commitment and action are needed.

8.09.6 REFERENCES

Ahles, T. A., Ruckdeschel, J. C., & Blanchard, E. B. (1984). Cancer related pain. I. Prevalence in an outpatient setting as a function of stage of disease and type of cancer. *Journal of Psychosomatic Research, 28,* 115–119.

Alby, N. (1991). Leukaemia: Bone marrow transplantation. In M. Watson (Ed.), *Cancer patient care: Psychosocial treatment methods* (pp. 281–297). Cambridge, UK: BPS Books.

Altmaier, E. M., Gingrich, R. D., & Fyfe, M. A. (1991). Two-year adjustment of bone marrow transplant survivors. *Bone Marrow Transplantation, 7,* 311–316.

American Cancer Society (1994). *Cancer facts and figures—1994.* Atlanta, CA: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.) Washington, DC: Author.

Andersen, B. L. (1992). Psychological interventions for cancer patients to enhance the quality of life. *Journal of Consulting and Clinical Psychology, 60,* 552–568.

Andersen, B. L. (1996). Predicting, understanding, and treating the sexual difficulties of gynecologic cancer survivors. *Cancer Control, 3,* 27–33.

Andersen, B. L., Anderson, B., & deProsse, C. (1989a). Controlled prospective longitudinal study of women with cancer: I. Sexual functioning outcomes. *Journal of Consulting and Clinical Psychology, 57,* 683–691.

Andersen, B. L., Anderson, B., & deProsse, C. (1989b). Controlled prospective longitudinal study of women with cancer: II. Psychological outcomes. *Journal of Consulting and Clinical Psychology, 57,* 692–697.

Andersen, B. L., & Cacioppo, J. T. (1995). Delay in seeking a cancer diagnosis: Delay stages and psychophysiological comparison processes. *British Journal of Social Psychology, 34,* 33–52.

Andersen, B. L., & Cyranowski, J. M. (1994). Women's sexual self schema. *Journal of Personality and Social Psychology, 67,* 1079–1100.

Andersen, B. L., & Elliott, M. L. (1994). Female cancer survivors: Appreciating their sexual concerns. *The Canadian Journal of Human Sexuality, 3,* 107–122.

Andersen, B. L., Karlsson, J. A., Anderson, B., & Tewfik, H. H. (1984). Anxiety and cancer treatment: Response to stressful radiotherapy. *Health Psychology, 3,* 535–551.

Andersen, B. L., Kiecolt-Glaser, J. K., & Glaser, R. (1994). A biobehavioral model of cancer stress and disease course. *American Psychologist, 49,* 389–404.

Andersen, B. L., & Lamb, M. A. (1995). Sexuality and cancer. In A. I. Holleb, D. Fink, & G. P. Murphy (Eds.), *American Cancer Society textbook of clinical oncology* (2nd ed., pp. 699–713). Atlanta, GA: American Cancer Society.

Andersen, B. L., & Tewfik, H. H. (1985). Psychological reactions to radiation therapy: Reconsideration of the adaptive aspects of anxiety. *Journal of Personality and Social Psychology, 48,* 1024–1032.

Andersen, B. L., Turnquist, D., LaPolla, J., & Turner, D. (1988). Sexual functioning after treatment of *in situ* vulvar cancer: Preliminary report. *Obstetrics & Gynecology, 71,* 15–19.

Andersen, B. L., Woods, X. A., and Copeland, L. J. (1997) Sexual self schema and sexual morbidity among gynecologic cancer survivors. *Journal of Consulting and Clinical Psychology, 65,* 221–229.

Anderson, R. J., & Kirk, L. M. (1982). Methods of improving patient compliance in chronic disease states. *Archives of Internal Medicine, 142,* 1673–1675.

Andrykowski, M. A., Schmitt, F. A., Gregg, M. E., Brady, M. J., Lamb, D. G., & Henslee-Downey, P. J. (1992). Neuropsychologic impairment in adult bone marrow transplant candidates. *Cancer, 70,* 2288–2297.

Andrykowski, M. A. (1993). The Morrow/Hickok article reviewed: Behavioral treatment of chemotherapy-induced nausea and vomiting. *Oncology, 7,* 93–94.

Andrykowski, M. A. (1994). Psychosocial factors in bone marrow transplantation: A review and recommendations for research. *Bone Marrow Transplantation, 13,* 357–375.

Ashburn, M. A., & Lipman, A. G. (1993). Management of pain in the cancer patient. *Anesthesia Analog, 76,* 402–416.

Ashley, J. M. (1993). Dietary guidelines for cancer prevention and control. *Oncology, 7*(suppl.) 27–31.

Baider, L., Cooper, C. L., & Kaplan De-Nour, A. (1996). *Cancer and the family*. Chichester, UK: Wiley.

Baider, L., & Kaplan De-Nour, A. (1988). Adjustment to cancer: Who is the patient—the husband or the wife? *Israel Journal of Medicine Sciences, 24,* 631–636.

Baker, F., Curbow, B., & Wingard, J. R. (1991). Role retention and quality of life of bone marrow transplant survivors. *Social Science and Medicine, 32,* 697–704.

Bard, M., & Sutherland, A. M. (1952). Adaptation to radical mastectomy. *Cancer, 8,* 656–671.

Bentler, P. M., & Newcomb, M. D. (1978). Longitudinal study of marital success and failure. *Journal of Consulting and Clinical Psychology, 46,* 1053–1070.

Bernstein, I. L. (1986). Etiology of anorexia in cancer. *Cancer, 58,* 1881–1886.

Blanchard, C. G., Ruckdeschel, J. C., Labrecque, M. S., Frisch, S., Blanchard, E. B. (1987). The impact of a designated cancer unit on house staff behaviors toward patients. *Cancer, 60,* 2348–2354.

Bloom, J. R. (1987). Psychological response to mastectomy. A prospective comparison study. *Cancer, 59,* 189–196.

Brown, H. N., & Kelly, M. J. (1976). Stages of bone marrow transplantation: A psychiatric perspective. *Psychosomatic Medicine, 38,* 439–446.

Burish, T. G., & Jenkins, R. A. (1992). Effectiveness of biofeedback and relaxation training in reducing the side effects of cancer chemotherapy. *Health Psychology, 11,* 17–23.

Burish, T. G., Snyder, S. L., & Jenkins, R. A. (1991). Preparing patients for cancer chemotherapy: Effect of coping preparation and relaxation interventions. *Journal of Consulting and Clinical Psychology, 59,* 518–525.

Byrne, J., Mulvihill, J. J., Connelly, R. R., Myers, M. H., Austin, D. A., Holmes, F. F., Holmes, G. E., Latourette, H. G., Meigs, J. W., & Strong, L. C. (1988). Reproductive problems and birth defects in survivors of Wilm's tumor and their relatives. *Medical and Pediatric Oncology, 16,* 233–240.

Byrne, J., Mulvihill, J. J., Myers, M. H., Abbott, S. C., Connelly, R. R., Hanson, M. R., Hassinger, D. D., Naughton, M. D., Austin, D. F., Gurgin, V. A., Holmes, F. F., Homes, G. F., Latourette, H. B., Weyer, P. J., Meigs, J. W., Teta, M. J., Strong, L. C., & Cook, J. A. (1985). Risk of infertility among survivors of childhood and adolescent cancer. *American Journal of Human Genetics, 37,* 24–48.

Cain, E. N., Kohorn, E. I., Quinlan, D. M., Latimer, K., & Schwartz, P. E. (1986). Psychosocial benefits of a cancer support group. *Cancer, 57,* 183–189.

Capone, M. A., Good, R. S., Westie, K. S., & Jacobson, A. F. (1980). Psychosocial rehabilitation of gynecologic oncology patients. *Archives of Physical Medicine and Rehabilitation, 61,* 128–132.

Carey, M. P., & Burish, T. G. (1988). Etiology and treatment of the psychological side effects associated with cancer chemotherapy: A Critical review and discussion *Psychological Bulletin, 104,* 307–325.

Carver, C. S., Pozo-Kaderman, C., Harris, S. D., Noriega, V., Scheier, M. F., Robinson, D. S., Ketcham, A. S., Moffat Jr., F. L., & Clark, K. C. (1994). Optimism vs. pessimism predicts the quality of women's adjustment to early stage breast cancer. *Cancer, 73,* 1213–1220.

Cassileth, B. R., Lusk, E. J., Strouse, T. B., Miller, D. S., Brown, L. L., & Cross, P. A. (1985). A psychological analysis of cancer patients and their next-of-kin. *Cancer, 55,* 72–76.

Cella, D. F., Orofiamma, B., Holland, J. C., Silberfarb, P. M., Tross, S., Feldstein, M., Perry, M., Maurer, L. H., Comis, R., & Orav, E. J. (1987). The relationship of psychological distress, extent of disease, and performance status in patients with lung cancer. *Cancer, 60,* 1661–1667.

Cella, D. F., & Tross, S. (1986). Psychological adjustment to survival from Hodgkin's disease. *Journal of Consulting and Clinical Psychology, 54,* 616–622.

Christensen, D. N. (1983). Postmastectomy couple counseling: An outcome study of a structured treatment protocol. *Journal of Sex & Marital Therapy, 9,* 266–274.

Coates, R. J., Bransfield, D. D., Wesley, M., Hankey, B., Eley, J. W., Greenberg, R. S., Flanders, D., Hunter, C. P., Edwards, B. K., Forman, M., Chen, V. W., Peynolds, P., Boyd, P., Austin, D., Muss, H., & Blacklow, R. S. (1992). Differences between black and white women with breast cancer in time from symptom recognition to medical consultation. *Journal of the National Cancer Institute, 84,* 938–950.

Coyle, N., Adelhardt, J., Foley, K. M., & Portenoy, R. K. (1990). Character of terminal illness in the advanced cancer patient: Pain and other symptoms during the last four weeks of life. *Journal of Pain and Symptom Management, 5,* 83–93.

Cunningham, M. P. (1997). Giving life to numbers. *CA—A Cancer Journal for Clinicians, 47,* 3–4.

Daut, R. L., Cleeland, C. S., & Flanery, R. C. (1983). Development of the Wisconsin Brief Pain Questionnaire to assess pain in cancer and other diseases. *Pain, 17,* 197–210.

deHaes, J. C., van Oostrom, M. A., & Welvaart, K. (1986). The effect of radical and conserving surgery on quality of life of early breast cancer patients, *European Journal of Surgical Oncology, 12,* 337–342.

Derogatis, L. R., Morrow, G. R., Fetting, J., Penman, D., Piasetsky, S., Schmale, A. M., Henrichs, M., & Carnicke, C. L. (1983). The prevalence of psychiatric disorders among cancer patients. *Journal of the American Medical Association, 249,* 751–757.

Devlen, J., Maguire, P., Phillips, P., Crowther, D., & Chambers, H. (1987). Psychological problems associated with diagnosis and treatment of lymphomas. I: Retrospective study and II: Prospective study. *British Medical Journal, 295,* 953–957.

Dunkel-Schetter, C., Feinstein, L. G., Taylor, S. E., & Falke, R. L. (1992). Patterns of coping with cancer. *Health Psychology, 11,* 79–87.

Edgar, L., Rosberger, Z., & Nowlis, D. (1992). Coping with cancer during the first year after diagnosis: Assessment and intervention. *Cancer, 69,* 817–828.

Ell, K., Nishimoto, R., Mantell, J., & Hamovitch, M. (1988). Longitudinal analysis of psychological adaptation among family members of patients with cancer. *Journal of Psychosomatic Research, 32,* 429–438.

Fawzy, F. I., Cousins, N., Fawzy, N. W., Kemeny, M. E., Elashoff, R., & Morton, D. (1990). A structured psychiatric intervention for cancer patients: I. Changes over time in methods of coping and affective disturbance. *Archives of General Psychiatry, 47,* 720–725.

Fawzy, F. I., Fawzy, N. W., Hyun, C. S., Gutherie, D., Fahey, J. L., Morton, D. L., & Elashoff, R. (1993). Maliganant melanoma: Effects of an early structured psychiatric intervention, coping, and effective state on recurrence and survival six years later. *Archives of General Psychiatry, 50,* 681–689.

Fawzy, F. I., Kemeny, M. E., Fawzy, N. W., Elashoff, R., Morton, D., Cousins, N., & Fahey, J. L. (1990). A structured psychiatric intervention for cancer patients: II. Changes over time in immunological measures. *Archives of General Psychiatry, 47,* 729–735.

Felton, B. J., Revenson, T. A., & Hinrichsen, G. A. (1984). Stress and coping in the explanation of psychological adjustment among chronically ill adults. *Social Science and Medicine, 18,* 889–898.

Foley, K. M. (1985). The treatment of cancer pain. *New England Journal of Medicine, 313,* 84–95.

Forester, B., Kornfeld, D. S., & Fleiss, J. (1985). Group psychotherapy during radiotherapy: Effects on emotional and physical distress. *American Journal of Psychiatry, 150,* 1700–1706.

Futterman, A. D., Wellisch, D. K., Bond, G., & Carr, C. R. (1991). The psychosocial levels system: A new rating scale to identify and assess emotional difficulties during bone marrow transplantation. *Psychosomatics, 32,* 177–186.

Garfinkel, L. (1994). Evaluating cancer statistics. *CA—A Cancer Journal for Clinicians, 44,* 5–6.

Gilbar, O., & Kaplan De-Nour, A. K. (1989) Adjustment to illness and dropout of chemotherapy. *Journal of Psychosomatic Research, 33,* 1–5.

Glanz, K., & Lerman, C. (1992). Psychosocial impact of breast cancer: A critical review. *Annals of Behavioral Medicine, 14,* 204–212.

Gottesman, D., & Lewis, M. S. (1982). Differences in crisis reactions among cancer and surgery patients. *Journal of Consulting and Clinical Psychology, 50,* 381–388.

Greene, P. G., Seime, R. J., & Smith, M. E. (1991). Distraction and relaxation training in the treatment of anticipatory vomiting: A single subject intervention. *Journal of Behavioral Therapy and Experimental Psychiatry,* 22, 285–290.

Greenwald, P., & Cullen, J. W. (1985). The new emphasis in cancer control. *Journal of the National Cancer Institute, 74,* 543–551.

Heber, D., Bagga, D., Ashley, J. M., & Elashloff, R. M. (1993). Nutritional strategies for breast cancer prevention. *Oncology, 7,* 43–46.

Helgeson, V. S., & Cohen, S. (1996). Social support and adjustment to cancer: Reconciling descriptive, correlational, and intervention. *Health Psychology, 15,* 135–148.

Herbert, T. B., & Cohen, S. (1993a). Depression and immunity: A meta-analytic review. *Psychological Bulletin, 113,* 472–486.

Herbert, T. B., & Cohen, S. (1993b). Stress and immunity in humans: A meta-analytic review. *Psychosomatic Medicine, 55,* 364–379.

Howson, J. Y. (1950). The procedures and results of the Philadelphia committee for the study of pelvic cancer. *Wisconsin Medical Journal, 49,* 215–219.

Hughson, A. V., Cooper, A. F., McArdle, C. S., & Smith, D. C. (1987). Psychosocial effects of radiotherapy after mastectomy. *British Medical Journal, 294,* 1515–1518.

Irvine, D., Brown, B., Crooks, D., Roberts, J., & Browne, G. (1991). Psychosocial adjustment in women with breast cancer. *Cancer, 67,* 1097–1117.

Irvine, D. M., Vincent, L., Bubela, N., Thompson, L., & Graydon, J. (1991). A critical appraisal of the research literature investigating fatigue in the individual with cancer. *Cancer Nursing, 14,* 88–199.

Jacobson, P. B., & Schwartz, M. D. (1993). Food aversions during cancer therapy: Incidence, etiology, and prevention. *Oncology, 7,* 139–143.

Johnston, M., & Vogele, C. (1993). Benefits of psychological preparation for surgery: A meta-analysis. *Annals of Behavioral Medicine, 15,* 245–256.

Kaplan, H. S., & Owett, T. (1993). The female androgen deficiency syndrome. *Journal of Sex & Marital Therapy, 19,* 3–24.

Keefe, F. J., Brantley, A., Manuel, G. M., Crisson, J. E. (1985). Behavioral assessment of head and neck cancer pain. *Pain, 23,* 327–336.

King, K. B., Nail, L. M., Kreamer, K., Strohl, R. A., & Johnson, J. E. (1985). Patients" descriptions of the experience of receiving radiation therapy. *Oncology Nursing Forum, 12,* 55–61.

Knox, L. S. (1991). Maintaining nutritional status in persons with cancer. In *Cancer nursing: Maintaining nutritional status in persons with cancer* (pp. 1–11). Atlanta, GA: American Cancer Society.

Lansky, S. B., List, M. A., Herrman, C. A., Ets-Hokin, E. G., Das Gupta, T. K., Wilbanks, G. D., & Hendrickson, F. R. (1985). Absence of major depressive disorders in female cancer patients. *Journal of Clinical Oncology, 2,* 1553–1560

Lebovits, A. H., Strain, J. J., Schleifer, S. J., Tanaka, J. S., Bhardwaj, S., & Messe, M. R. (1990). Patient noncompliance with self-administration chemotherapy. *Cancer, 65,* 17–22.

Levin, S. H., Jones, L. D., & Sack, D. A. (1993). Evaluation and treatment of depression, anxiety, and insomnia in patients with cancer. *Oncology, 7*(suppl.), 119–125.

Lewis, C., Linet, M. S., & Abeloff, M. D. (1983). Compliance with cancer therapy by patients and physicians. *The American Journal of Medicine, 74,* 673–678.

Linn, M. W., Linn, B. S., & Harris, R. (1982). Effects of counseling for late stage cancer patients. *Cancer, 49,* 1048–1055.

Locke, B. Z., & Regier, D. A. (1985). Prevalence of selected mental disorders. In C. A. Taube & S. A. Barrett (Eds.), *Mental Health United States* (pp. 1–6). Rockville, MD: National Institute of Mental Health.

Maguire, G. P., Lee, E. G., Bevington, D. J., Kuchemann, C. S., Crabtree, R. J., & Cornell, C. E. (1978). Psychiatric problems in the first year after mastectomy. *British Medical Journal, 1,* 963–965.

Mahon, S. M., Cella, D. F., & Donovan, M. I. (1990). Psychosocial adjustment to recurrent cancer. *Oncology Nursing Forum, 17,* 47–52.

Maier, S. F., Watkins, L. R., & Fleshner, M. (1994). Psychoneuroimmunology: The interface between behavior, brain, and immunity. *American Psychologist, 12,* 1004–1017.

Manuel, G. M., Roth, S., Keefe, F. J., & Brantley, B. A. (1987, Winter). Coping with cancer. *Journal of Human Stress, 13,* 149–158.

Margolis, G., Goodman, R. L., & Rubin, A. (1990). Psychological effects of breast-conserving cancer treatment and mastectomy. *Psychosomatics, 31,* 33–39.

Marks, G., Richardson, J. L., Graham, J. W., & Levine, A. (1986). Role of health locus of control beliefs and expectations of treatment efficacy in adjustment to cancer. *Journal of Personality and Social Psychology, 51,* 443–450.

Marshall, J. R., Gregorio, D. I., & Walsh, D. (1982). Sex differences in illness behavior: Care seeking among cancer patients. *Journal of Health and Social Behavior, 23,* 197–204.

Massie, M. J., & Holland, J. C. (1990). Depression and the cancer patient. *Journal of Clinical Psychiatry, 51,* 12–19.

Massie, M. J., & Holland, J. C. (1992). The cancer patient with pain: Psychiatric complications and their management. *Journal of Pain and Symptom Management, 7,* 99–109.

Matthews, K. A., Siegel, J. M., Kuller, L. H., Thompson, M., & Varat, M. (1983). Determinants of decision to seek medical treatment by patients with acute myocardial infarction symptoms. *Journal of Personality and Social Psychology, 44,* 1144–1156.

Mizushima, Y., Kashii, T., Hoshino, K., Morikage, T., Takashima, A., Hirata, H., Kawasaki, A., Konisha, K., & Yano, S. (1990). A survey regarding the disclosure of the diagnosis of cancer in Toyama prefecture, Japan. *Japanese Journal of Medicine, 29,* 146–155.

Morris, T., Greer, H. S., & White, P. (1977). Psychological and social adjustment to mastectomy: A two-year follow-up study. *Cancer, 40,* 2381–2387.

Morrow, G. R. (1986). Effect of the cognitive hierarchy in the systematic desensitization treatment of anticipatory nausea in cancer patients: A component comparison with relaxation only, counseling and no treatment. *Cognitive Therapy and Research, 10,* 421-446.

Morrow, G. R., Asbury, R., Hammon, S., Dobkin, P., Caruso, L., Pandya, K., & Rosenthal, S. (1992). Comparing the effectiveness of behavioral treatment for chemotherapy-induced nausea and vomiting when administered by oncologists, oncology nurses, and clinical psychologists. *Health Psychology, 11,* 250–256.

Morrow, G. R., & Hickok, J. T. (1993). Behavioral treatment of chemotherapy-induced nausea and vomiting. *Oncology, 7*(Suppl.), 83–89.

Mosconi, P., Meyerowitz, B. E., Liberati, M. C., & Liberati, A. (1991). Discolosure of breast cancer diagnosis: Patient and physician reports. *Annuals of Oncology, 2,* 273–280.

Munro, A. J., Biruls, R., Griffin, A. V., Thomas, H., & Vallis, K. A. (1989). Distress associated with radiotherapy for malignant disease: A quantitative analysis based on patients" perceptions. *British Journal of Cancer, 60,* 370–374.

Nail, L. M., & King, K. B. (1997). Symptom distress fatigue. *Seminars in Oncology Nursing, 3,* 257–262.

Nitschke, R., Humphrey, G. B., Sexauer, C. L., Catron, B., Wunder, S., & Jay, S. (1982). Therapeutic choices made by patients with end-stage cancer. *Journal of Pediatrics, 101,* 471–476.

Northouse, L. L., & Peters-Golden, H. (1993). Cancer and the family: Strategies to assist spouses. *Seminars in Oncology Nursing, 9,* 74–82.

Pack, G. T., & Gallo, J. S. (1938). Culpability for delay in treatment of cancer. *American Journal of Cancer, 33,* 443–462.

Parker, S. L., Tong, T., Bolden, S., & Wingo, P. P. (1997). Cancer Statistics, 1997. *CA—A Cancer Journal for Clinicians, 47,* 5–27.

Patenaude, A. F. (1990). Psychological impact of bone marrow transplantation: Current perspectives. *The Yale Journal of Biology and Medicine, 63,* 515–519.

Pickard-Holley, S. (1991). Fatigue in cancer patients. *Cancer Nursing, 14,* 13–19.

Rainey, L. C. (1985). Effects of preparatory patient education for radiation oncology patients. *Cancer, 56,* 1056–1061.

Rathmell, A. J., Ash, D. V., Howes, M., & Nicholls, J. (1991). Assessing quality of life in patients treated for advanced head and neck cancer. *Clinical Oncology, 3,* 10–16.

Redd, W. H. (1993). The Morrow/Hickok article reviewed: Behavioral treatment of chemotherapy-induced nausea and vomiting. *Oncology, 7,* 94–95.

Redd, W. H., Andersen, G. V., & Minagawa, R. Y. (1982). Hypnotic control of anticipatory emesis in patients receiving cancer chemotherapy. *Journal of Clinical and Consulting Psychology, 50,* 14–19.

Rhodes, V. A., Watson, P. M., & Hanson, B. M. (1988). Patients" descriptions of the influence of tiredness and weakness on self-care abilities. *Cancer Nursing, 11,* 186–194.

Richardson, J. L., Marks, G., Johnson, C. A., Graham, J. W., Chan, K. K., Selser, J. N., Kishbaugh, C., Barranday, Y., & Levine, A. M. (1987). Path model of multidimensional compliance with cancer therapy. *Health Psychology, 6,* 183–207.

Richardson, J. L., Marks, G., & Levine, A. (1988). The influence of symptoms of disease and side effects of treatment on compliance with cancer therapy. *Journal of Clinical Oncology, 6,* 1746–1752.

Rimer, B. K. (1992). Understanding the acceptance of mammography by women. *Annals of Behavioral Medicine, 14,* 197–203.

Rothenberg, R. (1992). *The new American medical dictionary and health manual* (6th ed.). New York: Penguin.

Rugo, H. (1993). Cancer. In L. Tierney, S. Mcphee, M. Papadakis, & S. Schroeder (Eds.), *CURRENT: Medical diagnosis & treatment.* Norwalk, CT: Appleton & Lange.

Sales, E., Schulz, R., & Biegel, D. (1992). Predictors of strain in families of cancer patients: A review of the literature. *Journal of Psychosocial Oncology, 10,* 1–26.

Sardell, A. N., & Trierweiler, S. J. (1993). Disclosing the cancer diagnosis. *Cancer, 72,* 3355–3365.

Sell, L., Devlin, B., Bourke, S. J., Munro, N. C., Corris, P. A., & Gibson, G. J. (1993). Communicating the diagnosis of lung cancer. *Respiratory Medicine, 87,* 61–63.

Slavin, L. A., O'Malley, J. E., Koocher, G. P., & Foster, D. J. (1982). Communication of the cancer diagnosis to pediatric patients: Impact on long-term adjustment. *American Journal of Psychiatry, 139,* 179–183.

Smets, E. M., Garssen, B., Schuster-Uitterhoeve, A. L., & deHaes, J. C. (1993). Fatigue in cancer patients. *British Journal of Cancer, 68,* 220–224.

Sneed, N. V., Edlund, B., & Dias, J. K. (1992). Adjustment of gynecological and breast cancer patients to the cancer diagnosis: Comparisons with males and females having other cancer sites. *Health Care for Women International, 13,* 11–22.

Speigel, D., & Bloom, J. R. (1983). Group therapy and hypnosis reduce metastatic breast carcinoma pain. *Psychosomatic Medicine, 45,* 333–339.

Speigel, D., Bloom, J. R., & Yalom, I. (1981). Group support for patients with metastatic cancer: A randomized outcome study. *Archives of General Psychiatry, 38,* 527–533.

Stewart, B. E., Meyerowitz, B. E., Jackson, L. E., Yarkin, K. L., & Harvey, J. H. (1982, October). Psychological stress associated with outpatient oncology nursing. *Cancer Nursing,* 383–387.

Strang, P., & Qvarner, H. (1990). Cancer-related pain and its influence on quality of life. *Anticancer Research, 10,* 109–112.

Syajala, K. L., Chapko, M. K., Vitaliano, P. P., Cummings, C., & Sullivan, K. M. (1993). Recovery after allogeneic marrow transplantation: Prospective study of predictors of long-term physical and psychosocial functioning. *Bone Marrow Transplantation, 11,* 319–327.

Taylor, S. E. (1983). Adjustment to threatening events: A theory of cognitive adaptation. *American Psychologist, 38,* 1161–1173.

Teeter, M. A., Holmes, G. E., Holmes, F. F., & Baker, A. B. (1987). Decisions about marriage and family among survivors of childhood cancer. *Journal of Psychosocial Oncology, 5,* 59–68.

Telch, C. F., & Telch, M. J. (1986). Group coping skills instruction and supportive group therapy for cancer patients: A comparison of strategies. *Journal of Consulting and Clinical Psychology, 54,* 802–808.

Teta, M. J., Del Po, M. C., Kasl, S. V., Meigs, J. W., Myers, M. H., & Mulvihill, J. J. (1986). Psychosocial consequences of childhood and adolescent cancer survival. *Journal of Chronic Disease, 39,* 751–759.

Thompson, L., Andersen, B. L., & DePetrillo, D. (1992). The psychological processes of recovery from gynecologic cancer. In M. Coppleson (Ed.), *Gynecologic oncology* (2nd ed.). Edinburgh, UK: Churchill Livingstone.

Vasterling, J., Jenkins, R. A., Tope, D. M., & Burish, T. G. (1993). Cognitive distraction and relaxation training for the control of side effects due to cancer chemotherapy. *Journal of Behavioral Medicine, 16,* 65–80.

Vess, J. D., Moreland, J. R., & Schwebel, A. I. (1985). A follow-up study of role functioning and the psychosocial environment of families of cancer patients. *Journal of Psychosocial Oncology, 3,* 1–14.

Vinokur, A. D., Threatt, B. A., Caplan, R. D., & Zimmerman, B. L. (1989). Physical and psychosocial functioning and adjustment to breast cancer: Long-term follow-up of a screening population. *Cancer, 63,* 394–405.

Wallace, L. M., Priestman, S. G., Dunn, J. A., & Priestman, T. J. (1993). The quality of life of early breast cancer patients treated by two different radiotherapy regimens. *Clinical Oncology, 5,* 228–233.

Wardle, J., & Pope, R. (1992). The psychological costs of screening for cancer. *Journal of Psychosomatic Research, 36,* 609–624.

Waters, W. E., Nichols, S., Wheeler, M. J., Fraser, J., & Hayes, A. J. (1983). Evaluation of a health education campaign to reduce the delay in women presenting with breast symptoms. *Community Medicine, 5,* 104–108.

Weisman, A. D., & Worden, J. W. (1976). The existential plight in cancer: Significance of the first 100 days. *International Journal of Psychiatry in Medicine, 7,* 1–15.

Westbrook, M. T., & Viney, L. L. (1982). Psychological reactions to the onset of chronic illness. *Social Science and Medicine, 16,* 899–905.

Wilkinson, S. (1991). Factors which influence how nurses communicate with cancer patients. *Journal of Advanced Nursing, 16,* 677–688.

Winer, E. P., & Sutton, L. M. (1994). Quality of life after bone marrow transplantation. *Oncology, 8,* 19–27.

Wingard, J. R., Curbow, B., Baker, F., & Piantadosi, S. (1991). Health, functional status, and employment of adult survivors of bone marrow transplantation. *Annals of Internal Medicine, 114,* 113–118.

Wolf, D. M., & Jordan, V. C. (1992). Gynecologic complications associated with long-term adjuvant tamoxifen therapy for breast cancer. *Gynecologic Oncology, 45,* 118–128.

Woodard, L. J., & Pamies, R. J. (1992). The disclosure of the diagnosis of cancer. *Primary Care, 19,* 657–663.

Wortman, C. B., & Dunkel-Schetter, C. (1979). Interpersonal relationships and cancer: A theoretical analysis. *Journal of Social Issues, 35,* 120–155.

Wynder, E. L., & Cohen, L. A. (1993). Nutritional opportunities and limitations in cancer prevention. *Oncology, 7*(Suppl.) 13–16.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.10
HIV and AIDS

MICHAEL H. ANTONI and NEIL SCHNEIDERMAN
University of Miami, Coral Gables, FL, USA

8.10.1 INTRODUCTION 238
- *8.10.1.1 Pathogenesis of HIV Infection* 238
- *8.10.1.2 Health Psychology and HIV Infection* 239
  - *8.10.1.2.1 Psychoneuroimmunology research and HIV infection* 240

8.10.2 PREDICTING ADJUSTMENT TO HIV INFECTION 242
- *8.10.2.1 Psychological Challenges in HIV Infection* 242
- *8.10.2.2 HIV-associated Psychosocial and Behavioral Sequelae* 244
  - *8.10.2.2.1 Distress and depression* 244
  - *8.10.2.2.2 Risk behavior and substance use* 246
- *8.10.2.3 Psychosocial Factors Predictive of Depression and Risk Behaviors in HIV Infection* 247
  - *8.10.2.3.1 Uncontrollable stressors* 247
  - *8.10.2.3.2 Stressor appraisals* 248
  - *8.10.2.3.3 Coping with HIV-related symptoms* 248
  - *8.10.2.3.4 Social support* 248

8.10.3 PREDICTING THE HEALTH COURSE OF HIV INFECTION 250
- *8.10.3.1 Biological Factors and Disease Progression* 250
  - *8.10.3.1.1 Constitutional factors* 250
  - *8.10.3.1.2 Other viral infections* 251
- *8.10.3.2 Behavioral and Psychosocial Factors and Disease Progression* 252
  - *8.10.3.2.1 Behavioral factors* 252
  - *8.10.3.2.2 Depressive symptoms* 253
  - *8.10.3.2.3 Stressful life events* 254
  - *8.10.3.2.4 Stressor appraisals* 255
  - *8.10.3.2.5 Coping strategies* 256
  - *8.10.3.2.6 Social support* 257
- *8.10.3.3 Psychosocial Factors in Long-term Survivors of AIDS* 258

8.10.4 INTERVENTIONS, ADJUSTMENT, AND HEALTH COURSE IN HIV INFECTION 259
- *8.10.4.1 Health Psychosocial Interventions with HIV-infected People* 259
- *8.10.4.2 Stress Management Intervention Effects at Different Critical Points in HIV Infection* 259
  - *8.10.4.2.1 The initial HIV seropositive diagnosis* 261
  - *8.10.4.2.2 The asymptomatic stage of HIV infection* 262
  - *8.10.4.2.3 The pre-AIDS symptomatic stage of HIV infection* 263
  - *8.10.4.2.4 Full-blown AIDS* 263
- *8.10.4.3 Common Features of Psychosocial Interventions in HIV Infection* 264
  - *8.10.4.3.1 Relaxation aspects* 264
  - *8.10.4.3.2 Delivery format* 264
  - *8.10.4.3.3 Treatment orientation* 264
- *8.10.4.4 Summary of Completed Studies of Psychosocial Interventions with HIV-infected Persons* 265

8.10.5 CONCLUSION 266

8.10.6 REFERENCES 267

8.10.1 INTRODUCTION

8.10.1.1 Pathogenesis of HIV Infection

The human immunodeficiency virus (HIV) is a retrovirus of the human T-cell leukemia/lymphoma line and is the causative agent of the acquired immunodeficiency syndrome (AIDS). (Although AIDS is believed to be caused by the human immunodeficiency virus-type 1 (HIV-1), we have for simplicity used the term HIV throughout this chapter.) A unique feature of HIV infection is that it can have a long asymptomatic phase lasting up to 10–15 years (Munoz et al., 1988) followed by the appearance of constitutional signs and symptoms of disease progression. In most people full-blown AIDS develops next (Kaplan, Wofsy, & Volberding, 1987), and even with aggressive and timely treatment, death often occurs a few years later (Lemp, Payne, Neal, Temelson, & Rutherford, 1990). A diagnosis of AIDS is made when a person shows laboratory-based signs of severe decline in the number of T-helper–inducer (CD4) cells ($< 200\,\text{cell}\,\text{mm}^{-3}$), or clinical signs consistent with the emergence of opportunistic infections and neoplasias. Some of these include: candidiasis of the esophagus, trachea, bronchi, or lungs; cryptoccosis; cryptosporidiosis with diarrhea persisting longer than one month; cytomegalovirus disease of an organ other than the liver, spleen, or lymph nodes; herpes simplex virus infection causing a mucocutaneous ulcer for longer than one month; Kaposi's sarcoma; primary lymphoma of the brain; pneumocystis carinii pneumonia; progressive multifocal leukoencephelopathy; toxoplasmosis of the brain; and a variety of other opportunistic infections and neoplasias. Since some infected people can remain free of clinical symptoms for a prolonged period of time and given that appropriate patient management can delay the onset of AIDS, it may be useful to view HIV as a chronic disease.

Many staging systems for HIV infection have been developed over the past decade (Walter Reed system, World Health Organization, and Centers for Disease Control), though the most widely accepted system used today is that developed by the Centers for Disease Control (CDC) in 1993. People with primary HIV infection often develop an acute mononucleosis-like syndrome about a month after initial infection (Tindall & Cooper, 1991). This period is associated with a high level of viral load or free virus released throughout the circulation (i.e., viremia) (Daar, Moudgil, Meyer, & Ho, 1991) which usually occurs 4–12 weeks after the onset of acute infection (Clerici, Berzofsky, Shearel, & Tacket, 1991). The virus becomes widely disseminated during this early stage of infection, suggesting that the course of infection may be influenced by the "seeding" of the virus throughout the lymph nodes and other immune system components (Tindall & Cooper, 1991). Even after this acute stage of the infection it appears that viral replication is not completely curtailed since it remains detectable in lymph nodes during the quiescent stages of the infection (Pantaleo, Poli, & Fauci, 1993).

A period of clinical, but not microbiological, latency lasting for a number of years follows the initial sequence of primary infection, viral dissemination, development of HIV specific immunity, and curtailment of extensive viral replication. Importantly, during the clinically latent period the disease is still progressing. This is observable in an increasing viral load, the depletion of CD4 cells in peripheral blood, and in the increasing proportion of HIV infected lymphoid cells (Pantaleo et al., 1993). It is generally accepted that the decline in CD4 cells leaves the infected person susceptible to a number of opportunistic infections and cancers characteristic of AIDS, though there is evidence that malfunctions in other aspects of immune functioning may also be associated with disease progression. Some of the more commonly observed infections include *pneumocystic carinii* pneumonia (PCP) cryptococcal meningitis, toxoplasmosis, and candida esophagitis (Kaplan et al., 1987). Many other diseases manifest in HIV-infected persons are caused by ubiquitous herpes viruses (e.g., cytomegalovirus (CMV) associated retinitis is a major cause of blindness in HIV-infected persons). HIV infected persons are also vulnerable to relatively rare cancers including Kaposi's sarcoma and Burkitt's lymphoma and, in women, cervical carcinoma. Interestingly these cancers are believed to be promoted by fairly common viruses as well (e.g., human papillomavirus types are associated with cervical neoplasia and squamous cell cervical carcinoma (Maiman & Fruchter, 1996). In contrast, it is rare that these viruses are successful in eluding the immune system for long enough to establish a clinical neoplastic process in healthy people, unless these people are undergoing pharmacologic treatments that suppress the immune system (Antoni, Esterling, Lutgendorf, Fletcher, & Schneiderman, 1995). HIV-infected persons are then extremely vulnerable to a wide range of pathogens normally controlled by the immune system and, with time, these individuals may contract a number of life-threatening diseases over the course of the infection. Because a person's awareness of their diagnosis and, more importantly, their symptom status may play an important role in determining their perceived psychosocial and physical burdens as well as their chosen coping strategies, this chapter shall later refer to different phenomenological stages of HIV infection tied to critical events as follows: the initial HIV seropositive diagnosis; the (post diagnosis) asymptomatic stage; the pre-AIDS symptomatic stage (formerly referred to as AIDS related Complex (ARC); and AIDS (often referred to as full-blown or frank AIDS).

Because there is no cure for AIDS, prevention is the major tool for limiting the spread of the disease. Primary prevention efforts use many behavioral change techniques, aimed at increasing availability and promoting use of condoms, and substance abuse management and treatment (Schneiderman, Antoni, Ironson, LaPierriere, & Fletcher, 1992). Many of these techniques are those that are commonly employed by health psychologists working in the area of risk management. Secondary prevention programs have also been undertaken to slow HIV spectrum disease progression. Because only a relatively small proportion of the one million people in the United States infected with HIV have AIDS (CDC, 1995) and because of the relatively prolonged period of time between the onset of infection and the development of AIDS (Munoz et al., 1989), there is a clear need to develop treatments to manage HIV spectrum disease and slow down disease progression. Much of this chapter will focus on such secondary prevention activities.

Many pharmacologic agents have been developed to manage symptoms and to slow HIV progression. These include such agents as pentamidine (to combat acute PCP and as a prophylaxic) and a large number of antiretrovirals, the most well-known being azidothymidine (AZT), to block the replication of HIV. To at least some extent advances in patient management have been successful in ameliorating symptoms and dealing with intercurrent illnesses (Longini, Clark, & Karon, 1993). In contrast, the benefits of antiviral therapies such as AZT, dideoxyinosine (ddI), and dideoxycytidine (ddC) appear to be only marginally effective and more controversial. In fact, in a letter to the *Lancet*, the Concorde Coordinating Committee (Aboulker & Swart, 1993) reported that although AZT modestly boosted the number of CD4 cells in the blood of asymptomatic patients, it did not delay progression to AIDS. Previous research had already shown that AZT only prolonged the life of HIV infected people for a few months (Fischl et al., 1987). The antiviral drugs also have undesirable physiological (e.g., bone marrow suppression, anemia, neutropenia; Richman, et al., 1987) and psychological (e.g., depression) side effects. We now know that using a single antiretroviral agent may lead to the development of AZT-resistant strains of HIV which are ultimately harder to treat; thus contemporary treatment approaches often focus on combination therapies (e.g., AZT and ddC). Newly emerging treatments such as triple combination therapy and protease inhibitors offer some hope that viral replication can be slowed more efficiently and with a lower likelihood of resistance or undue side effects (Lewin, 1996). However, we are now learning that the effectiveness of these newer regimens is critically dependent on the ability of patients to adhere to a rigid and demanding medication schedule. Failure to do so can substantially increase the risk of developing drug-resistant strains of the virus (Lewin, 1996; Schneiderman, Antoni, & Ironson, 1997).

8.10.1.2 Health Psychology and HIV Infection

Behavioral scientists focusing on HIV infection are usually involved in either primary prevention efforts (e.g., examining factors underlying risk behaviors that lead to infection in order to inform community-wide interventions and health policies) or secondary prevention efforts (e.g., examining factors contributing to the mental and physical health course of HIV-infected persons in order to develop clinical interventions). This chapter will focus on the role of health psychology in such secondary prevention efforts, though findings relevant to primary prevention efforts (e.g., predictors of sexual behavior and substance use) will also be addressed.

Managing HIV disease involves addressing psychosocial as well as biomedical issues (Antoni, Schneiderman et al., 1990; Schneiderman et al., 1994). The anticipation and the impact of HIV antibody test notification, for instance, is highly stressful (Ironson et al., 1990). Here, feelings of life-threat, doom, and anger are usually encountered as is the need to contemplate making major lifestyle changes (Christ & Wiener, 1985; Kaisch & Anton-Culver, 1989; Viney, Henry, Walker, & Crooks, 1989). Because of the large and accelerating number of people who are infected with HIV, there is a clear need to develop behavioral interventions to help them cope with the psychosocial aspects of their situation. There is also some evidence that improving psychological adaptation to HIV may have implications for physical health among infected individuals. Two areas that have been the focus of much health psychology research in HIV infection and AIDS examine how psychosocial factors such as distress states and depressive symptoms as well as certain risk behaviors might contribute to (i) the ways in which people adjust to the having HIV infection; and (ii) the actual health course of the infection.

Health psychological studies focusing on HIV and AIDS began in about 1987, coincident with the identification of HIV as the putative cause of many AIDS-related conditions that were emerging in clinics and hospitals around the world. At that time, behavioral scientists began studying the effects of psychosocial factors such as stressors and stress management interventions in people infected with HIV. The first major reason for this health psychological research concerned the issue of psychological adjustment. HIV infection presents multiple burdens at the psychological and physical levels, thereby creating a state of chronic stress that may overwhelm an individual's coping resources, thus greatly impairing their ability to emotionally adjust to what will likely be a long journey. Accordingly, it was reasoned that HIV infected individuals are among the groups who might benefit substantially from interventions that teach them to cope with the chronic demands that many people will never have to face. By learning how stress-reducing psychosocial interventions help such people adapt, adjust, and adhere to various lifestyle changes health psychologists gain insight into ways to facilitate the coping process in people facing this life-threatening condition as well as other chronic diseases (for reviews see Antoni, 1991; Ironson, Antoni, & Lutgendorf, 1995; Lutgendorf, Antoni, Schneiderman, Ironson et al. 1994; Schneiderman et al., 1994).

A second major reason for examining the effects of psychosocial factors in this population follows from a conceptualization of HIV infection as a chronic disease whose clinical course may be affected by multiple behavioral and biological factors (Antoni, Schneiderman et al., 1990). As in the case of other chronic diseases, HIV infection is characterized by a disorder in one or more bodily systems wherein signs and symptoms of the disorder are clinically manifest across long periods of time as a function of the degree of disorder in the bodily system(s) affected (e.g., immune system impairment). People with HIV infection become ill or die from the complications of their chronic disease not the disease itself. Just as diabetics who are unable to maintain their blood glucose levels within a certain range may develop kidney disease or suffer heart attacks, HIV-infected people who are unable to retain adequate immune system functioning necessary for surveillance of pathogens go on to develop the life-threatening infections or cancers that characterize full-blown AIDS.

8.10.1.2.1 Psychoneuroimmunology research and HIV infection

To the extent that HIV infection increases distress levels (Antoni, Schneiderman et al., 1990; Kaisch & Anton-Culver, 1989; Viney et al., 1989), which in turn have been shown to influence the immune system (Calabrese, Kling, & Gold, 1987; Ironson et al., 1990; Irwin, Daniels, Bloom et al., 1987; Kiecolt-Glaser et al., 1987), it is plausible that behavioral interventions that decrease distress may beneficially impact immune status and possibly health status in HIV-infected people. Because the degree to which the immune system becomes disordered seems to be the strongest predictor of the progression to HIV-related symptoms and death, much biomedical research has been focused upon ways to either slow the growth of the virus or slow the decline of the immune system by modulating the function of its cells. Psychoneuroimmunologic (PNI) research examines ways in which psychosocial influences such as stressful life events, distress states, and stress reducing techniques—all of which may contribute to successful adjustment to HIV infection—modulate immune system functioning and, ultimately, disease course. Applying these ideas from the point of view of a chronic disease model, to the extent that health psychology interventions modify psychosocial factors such as risk behaviors (unprotected sex, alcohol and drug abuse), emotional distress, maladaptive coping strategies, and social isolation, it might also modulate biological factors such as certain endocrine and immune system components. By diminishing the impact of psychosocial and behavioral factors on the immune system, the individual might retard the onset of disease complications by maintaining immunologic status (e.g., by preserving an adequate number of T-helper cells and/or maintaining the functioning of surviving immune cells) within a range necessary to defend against certain infections and cancers.

There are many studies relating psychosocial and behavioral factors to the immune system in a wide collection of species using an even wider number of different experimental paradigms. This area is marked by significant agreement and disagreement about the ways in which a wide range of factors impact the immune system in healthy animal and human populations. Clinical researchers have attempted to translate many of these findings into the clinical arena in order to classify how psychological phenomena such as self-efficacy and sense of control, coping strategies, social support, and emotional states such as depression—all reasonable targets of psychological interventions—contribute to impairments in the immune systems of HIV infected individuals (e.g., Antoni, 1990; Solomon, Kemeny, & Temoshok, 1991). Given existing evidence that experimentally induced and naturally occurring stressors, perceived loss of control, social isolation, and depression are related to decrements in the numbers and functions of immune cells known to be altered by HIV infection, it can be reasoned that these psychosocial and behavioral factors might influence immunologic status and, possibly, disease course in HIV-infected individuals. Similarly, stress management interventions that target these factors might provide both psychological and physical benefits for infected people, especially for those in the early stages of this chronic disease before the virus has established a stranglehold on the immune system and other physiological regulatory processes.

This line of reasoning can be operationalized in the context of HIV infection as follows: because CD4 cells are depleted in the advancing stages of HIV infection (Klimas et al., 1991), increases in qualitative aspects of HIV itself (e.g., mutation and replication rate) (Panteleo et al., 1993) and of lymphocytes (e.g., proliferation [growth and reproductive abilities], and cytotoxicity [recognition and killing abilities]) might be important in predicting those HIV seropositive individuals who do and do not develop opportunistic infections quickly. The ability of lymphocytes to multiply when challenged by antigens, often evaluated using tests of lymphocyte proliferative responses to plant mitogens such as phytohemaglutinin (PHA), concanavalin-A (conA), or pokeweed mitogen (PWM), may compensate to some degree for the smaller number of CD4 lymphocytes available as the disease progresses. Innate immune functions such as natural killer cell cytotoxicity (NKCC), a largely CD4-cell independent function, may also compensate in conferring protection against extant viral infections. For instance, herpes viruses such as Epstein Barr virus (EBV) or herpes simplex virus (HSV) are often poorly controlled in the context of HIV-induced defects in CD4 cell-directed cytotoxic function (Biron, Byron & Sullivan, 1989; Habu, Akamatsu, Tamaoki & Okumura, 1984, Bancroft, Shellam, & Chalmer, 1981; Bukowski, Warner, Denner & Welsh, 1985). It is plausible that NKCC, which may be partially preserved in HIV infection, may help to survey and control these and other infections in the HIV-infected host. Conversely, stress-induced impairments in NKCC may permit herpes virus reactivation to go unchecked in HIV-infected individuals (Glaser & Kiecolt-Glaser, 1987) with subsequent effects of HIV replication and progression to AIDS (Carbonari, Fiorilli, Mezzoroma, Cherchi, & Aiuti, 1989; Rosenberg & Fauci, 1991). Because several herpes viruses are known to have immunosuppressive effects, in and of themselves, reactivation of these viruses could have implications for compounding HIV-induced immune system decrements and, possibly, disease progression (Griffiths & Grundy, 1987).

Implicit in much PNI research is the notion that psychosocial factors relate to immune system changes by way of stress- or distress-induced changes in hormonal regulatory systems resulting in neuroendocrine elevations and imbalances (Maier, Watkins, & Fleshner, 1994). There are several neuroendocrine substances known to be altered as an function of an individual's appraisal (i.e., controllable vs. uncontrollable) of and coping response (i.e., active vs. passive) to stressful stimuli that have also been shown to impair certain components of the immune system. Some of these hormones, including those produced by the adrenal gland, are also known to be dysregulated in depressed individuals (Calabrese et al., 1987) and those reporting significant degrees of loneliness (Kiecolt-Glaser et al., 1984). Given that uncontrollable stressors, perceived loss of control (similar to low self-efficacy), and social losses such as divorce and bereavement have been related to alterations in some of these immunomodulatory hormones, it has been reasoned that some PNI relationships might be mediated, in part, by neurohormonal changes that are linked to an individual's appraisals of and coping responses to environmental burdens (for reviews see Antoni, Schneiderman et al., 1990 and Schneiderman et al., 1994). Although many different neurotransmitters, neurohormones, and other peptides have been studied as potentially important mediators of stressor-induced immunomodulation, two well-delineated "stress response pathways," the hypothalamic–pituitary–adrenal (HPA) axis and the sympathetic adrenomedullary (SAM) system, have received the most focus.

The HPA axis is a complex system that is regulated by several integrated central nervous system and peripheral mechanisms, a regulation that may be disrupted at several levels in depressive illness and other affective distress states (Amsterdam, Maislin, Gold, & Winokur, 1989). Several of these HPA axis abnormalities gradually normalize with clinical improvement in depression (Greden et al., 1983; Amsterdam, Lucki, & Winokur, 1985). Because immunologic function is impaired by cortisol (Plaut, 1987) in humans, immunologic correlates (especially blastogenesis and NKCC) of depressed states have been studied extensively. Despite the conclusive evidence for immunologic effects and neuroendocrine changes in depression, the exact mechanisms involved are lacking (Stein, Miller, & Trestman, 1991). However, it has been shown that glucocorticoids enhance HIV replication (Markham, Salahuddin, Veren, Orndorf, & Gallo, 1986), are immunosuppressive (Cupps & Fauci, 1982), and accompany affective disorder (Gold et al., 1986) and stressful experiences (Borysenko & Borysenko, 1982; Jacobs et al., 1986). On the other hand, one study found that 24 hour urinary-free cortisol levels were not related to CD4 (or CD8) cell counts in HIV positive individuals (Gorman et al., 1991). However, this does not preclude the possibility that distress or depression-associated cortisol elevations impact immune function (e.g., NKCC) in HIV-infected individuals and that such functional changes might relate directly to HIV disease progression by way of impaired surveillance of latent viruses and other pathogens.

A separate, but related neuroendocrine response system, the SAM system, is also activated during stressful circumstances (e.g., bereavement) and releases norepinephrine (NE) and epinephrine (E). Lymphocytes have beta-adrenergic receptors (Plaut, 1987), and behavioral stressors (Manuck, 1991; Antoni, Rodriguez, Starr et al., 1991; Naliboff et al., 1991) increase CD8 counts and may induce cyclic AMP-mediated suppression of lymphocyte and natural killer (NK) function (Plaut, 1987). Activation of both the HPA axis and the SAM system, which may interact (Axelrod & Reisine, 1984), during exposure to chronic uncontrollable stressors, could adversely affect immunologic status (Cupps & Fauci, 1982; Maier et al., 1994; Roszman et al., 1985). These neuroendocrine patterns may also be displayed when an individual is exposed to uncontrollable chronic stressors, has inadequate social support, and uses denial/avoidance coping strategies (Antoni et al., 1990; Schneiderman et al., 1992). If depression and uncontrollable stressors are associated with impaired immunologic function mediated by neuroendocrine changes in HIV-infected individuals, and if intervention-associated decreases in depressed mood help to normalize neuroendocrine functioning in this population, then such interventions may concurrently act to attenuate neuroendocrine-associated immunosuppression.

Putting all of this information together it can be hypothesized that stressful events or burdens that are interpreted by HIV-infected individuals as beyond their control might lead to social isolation, loneliness, anxiety, and depressed affect which might accompany alterations in some neurohormones (e.g., peripheral catecholamine and cortisol elevations) due to sympathetic nervous system (SNS) activation and dysregulation of the HPA axis. These neuroendocrine changes may also be accompanied by changes in the immune system (redistribution of lymphocytes and decrements in functions concerning lymphocyte proliferation and NKCC) via interactions among neural and neuroendocrine signals at the immune cell membrane, intracellular cyclic nucleotide activation, and the production of cytokines such as interleukin (IL)-I and II, and γ-interferon (γ-IFN). Since these structural (cell counts) and functional (proliferation and cytotoxicity) aspects of the immune system are known to decline progressively across the course of HIV infection (Panteleo et al., 1993), it can be further hypothesized that superimposed stressor-induced changes in the functioning of the immune system may increase the rate at which infected people develop clinical symptoms such as opportunistic infections and neoplasias. This model prescribes at least two sets of related targets for intervention: psychosocial experiences and pathophysiological processes (see Figure 1) (for details see Schneiderman et al., 1994).

8.10.2 PREDICTING ADJUSTMENT TO HIV INFECTION

8.10.2.1 Psychological Challenges in HIV Infection

A diagnosis of HIV seropositivity is the beginning of a long road of challenging life events and extraordinary personal changes that can overwhelm even the most psychologically well-adjusted individuals. One unique aspect of HIV infection is that it is a chronic physical disease embedded in an atmosphere of multiple chronic psychosocial and physical demands and burdens. From a health psychology perspective, it is plausible to hypothesize that disease-related and external "life stressors" operate interactively and that success at managing one can influence the course of the other. To test these types of hypotheses behavioral researchers have for years been examining associations between the occurrence of stressful life events and alterations in people's resiliency or susceptibility to a wide variety of pathogenic processes underlying chronic diseases such as coronary heart disease, diabetes mellitus, rheumatoid arthritis, and several types of cancer. In addition to studying the impact of stressful events, health psychology researchers, since the mid-1970s, have also focused on the role of stress "moderating" psychosocial variables (e.g., personality characteristics and social support) and

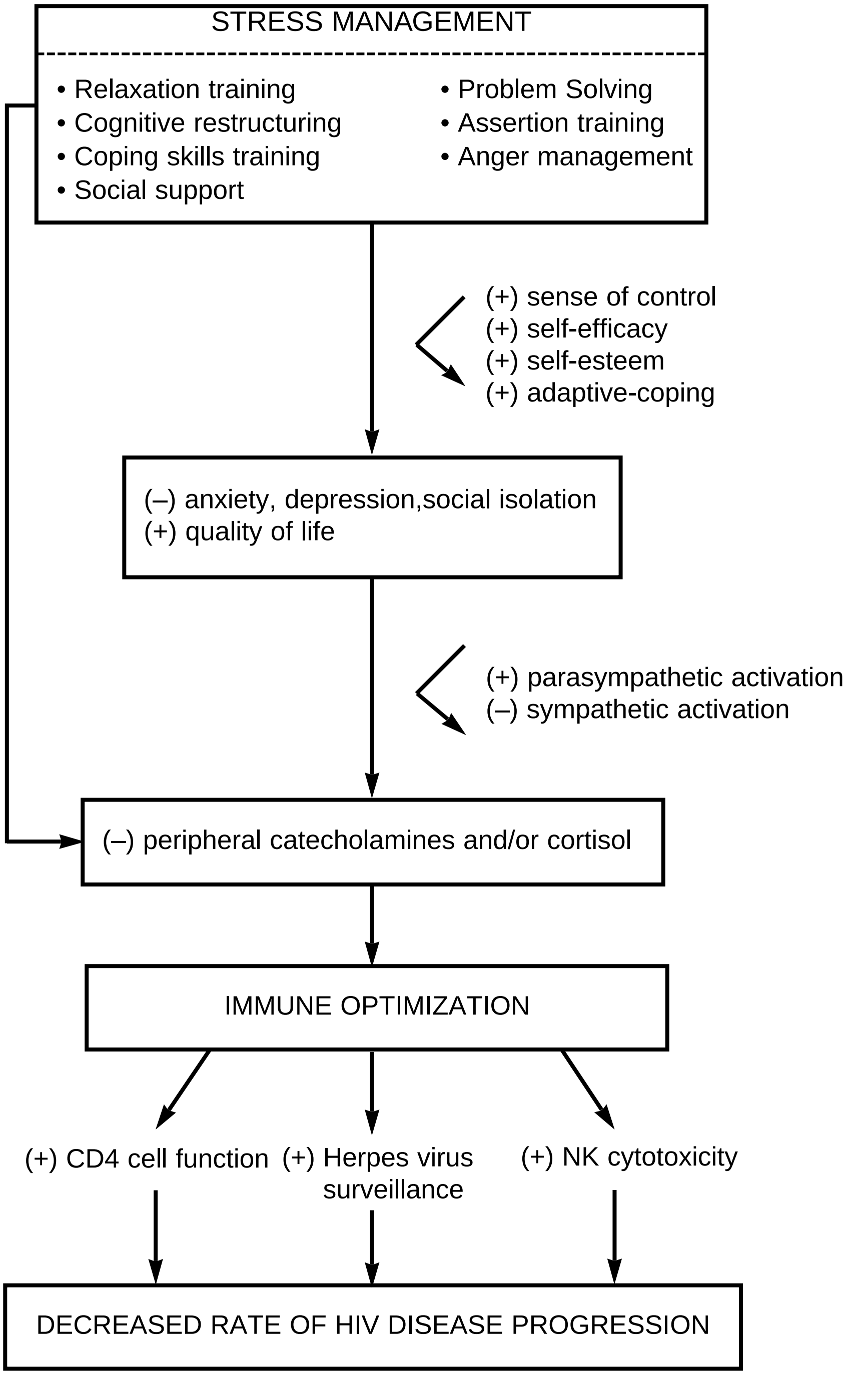
STRESS MANAGEMENT
• Relaxation training
• Cognitive restructuring
• Coping skills training
• Social support
• Problem Solving
• Assertion training
• Anger management
(+) sense of control
(+) self-efficacy
(+) self-esteem
(+) adaptive-coping
(–) anxiety, depression,social isolation
(+) quality of life
(+) parasympathetic activation
(–) sympathetic activation
(–) peripheral catecholamines and/or cortisol
IMMUNE OPTIMIZATION
(+) CD4 cell function
(+) Herpes virus surveillance
(+) NK cytotoxicity
DECREASED RATE OF HIV DISEASE PROGRESSION

Figure 1 Theoretical model for effects of stress management on disease progression in human immunodefiency virus (HIV) infection (CD4, T-helper-inducer cells; NK, natural killer cells).

psychosocial interventions (e.g., stress management) that might reduce the negative health influences of stressors in medically vulnerable populations such as those recovering from myocardial infarction or taxing treatments for some types of cancer.

Because more effective medical treatment for HIV infection is becoming available, there is now a fast-growing population of infected individuals who are coping with the complex and multiple psychosocial demands of a chronic life-threatening illness (Antoni, 1991). Homosexual men comprise a large majority of HIV-infected individuals in the United States. Today, however, an increasingly large percentage of HIV-infected people are heterosexual male drug users and women (Centers for Disease Control, 1995). Many of the people who are infected face lifestyle-associated social stigmas in addition to the direct burdens of the HIV infection. They endure many of the psychosocial phenomena previously associated with impairments in the immune system. Some of these changes include overt signs of progressive physical deterioration (Redfield & Burke, 1988), chronic legal problems (Blendon & Donelan, 1988; Ginzburg & Gostin, 1986; Walkey, Taylor, & Green, 1990), astronomical health-care costs (Bloom & Carliner, 1988), and multiple losses due to AIDS deaths (Martin, 1988)—all of which can be characterized as severe, chronic, uncontrollable, and unpredictable stressors. At the initial onset of HIV-related signs and symptoms these individuals may be overwhelmed and socially isolated, and may employ maladaptive coping strategies such as denial/avoidance possibly resulting in an increased likelihood of depression and distress, and perhaps sexual risk behaviors and substance use—all of these are potentially associated with changes in immunologic status (Antoni, 1991).

Since HIV infection is a chronic, progressive disease with an uncertain clinical course lasting up to 10–15 years, affected individuals must develop lifestyle changes, effective coping strategies, and, often, new social networks to meet their daily needs for hope, strength, and growth. Adjusting to a long-term progressive lethal illness entails addressing issues such as changes in self-expectations and preparation for death. These are, indeed, some of the central tasks in this stage of HIV infection (Hoffman, 1991; Lutgendorf, Antoni, Schneiderman, & Fletcher, 1994). A very tangible issue is that employment status often changes dramatically at this point in the disease (Chuang, Devins, Hunsley, & Gill, 1989; Dilley, Ochitil, Pert, & Volberding, 1985). One study reported that the average monthly income of patients with AIDS or ARC had fallen by an average of $1000 per month since their diagnosis (Ellerman, 1989). Intermittent episodes of PCP are highly associated with diminished abilities to maintain employment (Wachtel et al., 1992). Change in employment status can pervade many areas of life—loss of sources of social support, identity, income, and medical insurance—that, in turn, can deplete one's resources for dealing with the demands of the illness (Lutgendorf, Antoni, Schneiderman, Ironson et al., 1994).

Even for those who are successful at mustering the energy and resources for taking this journey, there still remain critical stressor events—emergence of new symptoms, development of resistance to a successful drug regime, loss of health insurance benefits, deaths of friends, rejection from family members, etc.—which can trigger lapses into depression, anxiety, and other distress reactions on the one hand and negative health behaviors on the other. Such lapses, if severe and unremitting, may contribute to a negative health course by way of increased mental health problems (Cleary et al., 1988; Jacobsen, Perry, & Hirsch, 1990) and acceleration of disease progression through PNI processes. If so, then psychosocial interventions may be capable of reducing the frequency and severity of these lapses with the accompanying effect of improving quality of life, reducing immunologic perturbations, and slowing the progression of the infection.

Some of the mechanisms that could mediate the effects of stress management intervention on the course of HIV infection have been proposed (Antoni, Schneiderman et al., 1990). According to this model, cognitive behavioral stress management (CBSM) interventions such as relaxation training and cognitive restructuring may reduce anxiety, depression, and social isolation by way of their ability to increase a sense of control, self-efficacy, and self-esteem. These psychological changes may also be accompanied by decreases in peripheral catecholamines and cortisol via decreases in sympathetic activation and improved regulation of the HPA axis, respectively. These neuroendocrine changes may, in turn, be associated with a partial normalization of immune system functions such as lymphocyte proliferation and NKCC that facilitate efficient surveillance of pathogens. To the degree that such a normalization of stress-associated immune system decrements can be achieved by stress management techniques, then any contribution of stressor and stress responses processes to acceleration of disease and expression of clinical symptoms might be forestalled or minimized.

It is arguable that such interventions can be targeted first to some of the more prevalent psychosocial and behavioral sequelae that HIV-infected people may face at various points in the disease process and some of the potentially changeable cognitive, affective, behavioral, and social factors that may aggravate or forestall these phenomena. Some of the most well-established sequelae include distress states, depressive symptoms, and negative health behaviors.

8.10.2.2 HIV-associated Psychosocial and Behavioral Sequelae

8.10.2.2.1 Distress and depression

HIV-infected persons appear to be at somewhat increased risk from *DSM-IV*-Axis I affective and adjustment disorders, and many suffer significant subclinical distress reactions following the initial news of an HIV seropositive diagnosis (Cleary et al., 1988; Jacobsen et al., 1990; Rabkin, Wagner, & Rabkin, 1996). Among diagnosable reactions, adjustment disorder with depressed mood may be the most common presenting complaint with major depressive disorder being far lower in prevalence (Perkins, Stern, Golden et al., 1994; Rabkin et al., 1996; Williams, Rabkin, Remien et al., 1991). One study found that the suicide rate among HIV-infected men may be up to 36 times that of age-matched uninfected men (Marzuk et al., 1988; Rundell, Paolucci, & Beatty, 1988), supporting the view that depressive symptoms are a common sequel to an HIV diagnosis and especially in those with a prior history of depression (Rabkin, Rabkin, Harrison, & Wagner, 1994). Pharmacologic treatments for depression such as imipramine and serotonin reuptake inhibitors have been shown to be safe and effective with HIV-infected samples (Rabkin et al., 1996). Newer work suggests that other biological approaches such as testosterone replacement therapy may also offer antidepressant effects (Rabkin et al., 1996). The coexistence of feelings of life-threat, doom, anger, and responsibility for making immediate lifestyle changes following a seropositive diagnosis (Christ & Wiener, 1985; Kaisch & Anton-Culver, 1989; Viney et al., 1989) suggests that psychosocial interventions that provide support, teach coping strategies, and offer the opportunity for mastery experiences may also be beneficial. Such interventions may also help these individuals to ventilate feelings such as anger (Miller, 1988) and manage anticipatory grief over the expectation of loss and death (Cohen, 1990; Pickrel, 1989). To the extent that HIV infection constitutes a chronic long-term disease, it is arguable that psychosocial intervention plays a key role in the treatment of HIV-infected persons. We have reasoned that all of these putative intervention effects are likely to be facilitated within a supportive group environment (Antoni, Schneiderman, & Ironson, in press).

Some of the evidence supporting the possible psychological benefits of a cognitive-behavioral treatment orientation for HIV-infected individuals comes from the patterns of results emerging from natural history studies conducted since the mid-1980s. For instance, there is growing evidence that the ways in which individuals cognitively and behaviorally cope with HIV-related stressors can influence their ability to establish psychological equilibrium at least within a given stage of the infection. In a community sample of 208 HIV-positive and HIV-negative, successful, and well-educated gay men, the coexistence of high levels of faith in the future and low levels of syndromal disorder or depressive symptoms (Rabkin, Williams, Neugebauer, & Goetz, 1990) suggested that interventions capable of instilling hope and self-efficacy may facilitate psychological adjustment. There is evidence that these "moderator" variables are not unique to gay men. Psychosocial characteristics of HIV-positive hemophiliacs related to elevated depression and distress included (i) a previous history of premorbid psychiatric disturbance/distress, (ii) experiencing recent life events involving loss, (iii) low social support from spouse, family, and friends, and (iv) a poor sense of mastery over life (Dew, Ragni, & Nimorwicz, 1990). Although based on a small heterosexual sample and a cross-sectional design, this work suggests the importance of uncontrollable loss-oriented life stressors, social isolation, and a lack of self-efficacy as deterrents of psychological adjustment to HIV infection.

While the psychosocial sequelae of HIV infection are reasonably well-documented for male samples, far less is known about what women experience during the course of this disease. Such information is imperative given the fact that the incidence of AIDS has increased at a faster rate for women than for any other subgroup in the population of the United States (Carpenter, Mayer, Fisher, Desai, & Durand, 1989; CDC, 1995a). AIDS incidence data through June 1996 indicates nearly 20% of all new AIDS cases are women (CDC, 1995b). AIDS cases and infected and at-risk women are disproportionately represented among minority populations, especially Blacks (Antoni, Schneiderman et al., 1992; CDC, 1995b). Young, Black American, low socioeconomic status (SES) women have been characterized as a group experiencing the triple stigmas of race, class, and gender, prior to the addition of the burden of HIV seropositivity (Quinn, 1993) as well as multiple chronic daily stressors (Freudenberg, Lee, & Silver, 1989; Mays & Cochran, 1988). These stressors include drug and alcohol dependency, poor social networks, medical problems, financial problems, overburdened public clinics, lack of transportation, and inaccessibility of child care, all potentially fueling a general feeling of helplessness and powerlessness (Quinn, 1993).

These stressors may be compounded even more by pregnancy (Antoni, Schneiderman et al., 1992). Relevant stressors here include the decision to abort or continue pregnancy to term, fears of becoming too ill to provide for offspring, and preparing older children for life as orphans. Seropositive mothers with unprotected partners are also likely to face rejection by the father of their children, resulting in emotional and financial isolation at a time when these women are most in need of support. Conversely, HIV seropositive women may react to an infected partner with self-isolation, both of which are likely to leave her and her offspring alone. If both mother and child are seropositive, the lack of emotional and financial support could make the situation even worse. Thus, the risk for significant emotional problems in HIV-infected mothers is obvious. It is critical that future work focus on factors contributing to depressive symptoms and mental adjustment in these growing populations of HIV-infected persons so that effective interventions can be developed.

8.10.2.2.2 Risk behavior and substance use

In addition to the influence of mood disturbances, psychosocial adjustment to HIV may be intimately related to negative health behaviors. In fact, certain negative health behaviors may show a bidirectional relationship with depression and distress levels. Two such negative health behaviors that are salient in the context of HIV infection are sexual risk behaviors and substance use. It is known for instance, that depressed people are often more likely to engage in risky (unprotected) sexual behaviors (Gold & Skinner, 1992; Kelly et al., 1993) and substance abuse (Rabkin et al., 1996). Beyond the fact that these two sets of behaviors are associated with transmission of the virus, it is also well known that they occur at very high rates in persons who are having difficulty dealing with being HIV-infected.

Despite reductions in sexual risk behaviors among urban gay men since 1987, a significant number continue to engage in unprotected anal intercourse (Martin, 1989), the sexual activity that carries the highest risk for HIV infection in gay men (Kingsley et al., 1987). Importantly, as many as 15% of those gay men who modified their sexual behaviors in the early years of the AIDS crisis have relapsed to unsafe sexual activities (Coates, 1991). One important factor that may influence such risk-taking behavior is the use of alcohol and drugs during sex (Martin, 1990b). The prevalence of alcohol abuse in the gay community is significantly higher (30%) than for the general population (5–10%) (Fifield, 1975; Lohrenz, Connelly, Coyne, & Spare, 1978; Martin, 1990b). Importantly, alcohol and drug use is strongly asociated with sexual behaviors and may determine adherence to safer sex guidelines in gay men (Martin, 1990b) as well as other populations who are at risk from HIV infection (Lollis, Antoni, Johnson, Chitwood, & Griffin, 1995; Lollis, Johnson, Antoni, & Hinkle, 1996). Finally, depressive symptoms may be more likely to occur in HIV-infected men and women who are substance abusers, especially those who are injection drug users (IDUs) (Rabkin et al., 1996). Thus, chronic depressive symptoms, substance abuse, and unprotected sexual behaviors may be highly interrelated and self-perpetuating.

Psychological factors related to the adoption and maintenance of reduced sexual risk behaviors include self-efficacy for safe sex, having alternative coping strategies for dealing with high-risk sexual impulses, and possessing skills for negotiating with interpersonal contacts (Kelly & St. Lawrence, 1988; McCusick, Hortsman, & Coates, 1985), while similar constructs such as condom efficacy and perceived vulnerability have been shown to predict sexual risk behaviors in injection drug users (Lollis et al., 1995). Denial of virulence and denial of personal responsibility, and frequent drug and alcohol use preceding sex (which are likely to facilitate one another) are also associated with continued high-risk behavior (Martin, 1990b). Specifically, Martin (1990b) found consistent associations between drug use (marijuana, inhaled nitrates, cocaine, hallucinogens, barbituates, and amphetamine) and unprotected insertive and receptive anal intercourse among 604 urban, middle-class gay men in the year prior to their development of AIDS. Cessation of drug use with sex predicted subsequent lower rates of unprotected anal intercourse (Martin, 1990). Thus, one way to intervene to reduce high-risk sexual behaviors may be to modulate drug and alcohol use perhaps by improving the ways individuals manage HIV-related stressors. Among 624 gay men, two categories of life stressors—deaths of loved ones due to AIDS and fear of illness and one's own death due to HIV infection—were associated with high-risk sexual behaviors and drug and alcohol use as well as psychological distress (Martin, 1989). Perceived self-efficacy and social support have also been shown to be critical for maintaining adherence to safer sex guidelines (Stall, Coates, & Hoff, 1988). Thus, among HIV-infected individuals, in addition to being associated with depressive symptoms, low self-efficacy, social support losses, and denial coping strategies may relate to the onset and maintenance of substance use and sexual risk behaviors and these may all be interrelated.

The role of ethnic, racial, cultural, and religious factors in influencing sexual behavior patterns is also an important consideration (Mays & Cochran, 1988); hence sociocultural influences on risk behavior reduction need to be understood. For poor ethnic minority women who are at risk of HIV infection or are already infected, the risk of AIDS is only one risk in a long list of multiple risks that are prevalent in daily existence. The key to behavior change is where in this list the women place the relatively longer-term consequences of AIDS-related risk behavior in relation to the short-term consequences of the risks of daily life (e.g., safety, food, shelter, human contact). From this vantage point, behavior change specialists need to be aware of more immediate survival needs which, if met, can provide the opportunity for professionals to educate and intervene effectively. Within this deeper understanding of the sociocultural milieu is the knowledge that the pervasive powerlessness experienced by disadvantaged women at risk may lead them to disengage from attempts at coping when informed about risk behaviors.

8.10.2.3 Psychosocial Factors Predictive of Depression and Risk Behaviors in HIV Infection

There is good evidence suggesting that the risk of depressive symptoms in HIV-infected persons is strongly tied to a prior history of depression (Rabkin et al., 1994). Other risk factors include a family history of depression, alcohol abuse, other substance abuse, and social isolation (Rabkin et al., 1996). There are also several psychosocial events that may follow the onset of initial HIV diagnosis and related symptoms (Antoni, Baggett et al., 1991; Schneiderman et al., 1992) and these may predict an increased likelihood of depression and distress, sexual risk behaviors, and substance use. These include: chronic, uncontrollable, and unpredictable stressors such as treatment nonresponsiveness and side effects, disease recurrence, stigmatizing behaviors, and medical costs (Blendon & Donelan, 1988; Bloom & Carliner, 1988; Ginzburg & Gostin, 1986; Redfield & Burke, 1988; Walkey, Taylor, & Green, 1990). Inability to cope with these mounting stressors may bring about a loss of self-esteem and self-efficacy, feelings of helplessness and hopelessness, depression, and an increase in maladaptive, potentially self-destructive negative health behaviors. The combination of uncontrollable life stressors and feelings of helplessness may increase cognitive and emotional burdens (Noh, Chandarana, Field, & Posthuma, 1990) that may overwhelm premorbid coping strategies such as active problem-solving and positive reframing resulting in the use of denial/avoidance coping strategies or giving up entirely. Loss of familiar sources of social support due to deaths of close friends (Martin, 1988), the infected person's own desire for social insularity and the avoidance by acquaintances and significant others following disclosure of homosexual orientation (Beckett & Rutan, 1990; Gambe & Getzel, 1989) may create a chronic lack of resources available to deal with stressors, thus compounding these adjustment problems. We now review some of these phenomena in greater detail to demonstrate how they interact.

8.10.2.3.1 Uncontrollable stressors

There are numerous psychosocial challenges following a diagnosis of HIV seropositivity. Among 240 gay and bisexual men, specific stressors—receiving an HIV-positive diagnosis, an AIDS diagnosis, an AIDS-related complex (ARC) diagnosis (the prior designation for having overt HIV-related symptoms but not yet developing AIDS), and learning of the death or AIDS diagnosis of a lover—were rated as the most emotionally distressing events that these people deal with (Rosser, Simon, & Michael, 1988). It is well-established that stressful life events shown to relate to depression and physical health consequences are those which are perceived as being beyond the individual's control (Justice, 1985). These events may also relate to increases in negative health behaviors. A variety of stressful events have been associated with alcohol consumption (Foy, Sipprelle, Rueger et al., 1984; Newcomb & Harlow, 1986) and low personal control is believed to be a mediator of this relationship (Beckman, 1980; Pearlin & Radabaugh, 1976). Similar findings are evident in gay men and may be more prevalent than in other populations given the reinforcement of heavy drinking in gay bars, a primary social gathering place (Nardi, 1982). Prevalence may similarly be enhanced in homosexual men as well as in ethnic minority group members due to the likelihood that alcohol (and drugs) are used to cope with burdens associated with being gay or otherwise disenfranchised in the society (e.g., guilt, anxiety, alienation, powerlessness) (Lohrenz et al., 1978). According to the identity disruption model (Lohrenz et al., 1978), negative life events that disturb one's self-concept or identity have the greatest deleterious effects. Self-conceptions determine the way personal information is appraised and processed, and the nature of social interactions. Identity disruptions, according to this model, may impair functioning by causing individuals to devote excess time and energy to the stressor, and undermine their sense of control and ability to predict other aspects of their lives. For those diagnosed with HIV infection, each HIV-associated stressor experienced may erode their previous self-concept giving way, possibly, to self-blame for their seropositivity (Hoffman, 1991).

8.10.2.3.2 Stressor appraisals

In addition to providing an overt sign of disease progression, physical changes such as skin lesions, muscle wasting, and fungal infections may compromise self-image in HIV-infected persons (Hoffman, 1991; Ostrow, Monjan et al., 1989). Infected individuals who see the onset of these changes as uncontrollable may attempt to cope by way of denial and increased substance use (Kaisen & Anton-Culver, 1989). These maladaptive strategies may affect psychological adjustment, future risk behaviors, and physical health (Burns et al., 1991; Ironson et al., 1992; Penkower et al., 1991). It is well known that a person's appraisal of stressful events may determine the individual's coping responses and emotional state (e.g., depression) (Lazarus & Folkman, 1984). Appraisals that are particularly relevant to the demands of an HIV diagnosis include those concerning harm/loss (e.g., loss of relationships, physical abilities, independence [Beckett & Rutan, 1990; Ferrara, 1984], threat (e.g., threat/anticipation of future symptoms [Hoffman, 1991]), and challenge (e.g., opportunities for personal growth)—three domains that have been studied extensively by Lazarus and Folkman. Thus, one important goal of psychosocial intervention is to help individuals enhance their sense of self-efficacy for dealing with the losses, threats, and challenges characterizing the physical and psychosocial sequelae of their HIV infection. Efforts to build self-efficacy may be most effective early in the course of the psychosocial intervention and can be enhanced by constructing a supportive group environment and teaching participants tangible strategies for managing daily hassles as well as major environmental stressors and their cognitive, behavioral, emotional, and interpersonal responses to these events.

8.10.2.3.3 Coping with HIV-related symptoms

Individuals previously diagnosed with HIV infection, who have not yet developed AIDS-defining symptoms, report an overwhelming sense of uncertainty (Dew et al., 1990). This uncertainty may actually contribute to the higher incidence of adjustment disorders reported for those with symptomatic pre-AIDS conditions vs. those with full-blown AIDS (Dilley et al., 1985). Uncertainty about one's immediate future creates critical demands for coping, the success of which may depend upon an individual's appraisals, attributions of blame, opportunities for social comparison, repertoire of adaptive coping strategies, and the sense of self-efficacy necessary for using these strategies.

Coping strategies classified as active-behavioral (planning, seeking social support) or active-cognitive (finding meaning in an illness, reframing) have been related to more positive affect and higher self-esteem in various chronic disease populations (Billings & Moos, 1981; Bloom, 1982; Fawzy et al., 1990; Felton, Revenson, & Hinrichsen, 1984). Conversely, denial and avoidance strategies are associated with greater depression and distress (Billings & Moos, 1981; Namir, Wolcott, Fawzy, & Alumbaugh, 1987). Individual differences in available coping strategies for dealing with stressors may be among the most significant predictors of depression and future adjustment to HIV seropositivity (Namir et al., 1987; Weimer, Nilsson-Schonnesson, & Clement, 1989) and may, possibly, play a role in future clinical disease progression (Ironson et al., 1992). One group prospectively related coping strategies and immunologic measures in HIV seropositive gay men over the one year period following notification of their antibody test results (Antoni et al., 1995a). Asymptomatic seropositive men's coping responses to antibody status notification were measured concurrently with phenotypic and functional immunologic markers, and these immunologic measures were taken again one year later. Seropositive men scoring above the median on postnotification disengagement coping strategies (denial, behavioral disengagement, mental disengagement) had significantly lower concurrently measured T-helper/suppressor (CD4/CD8) cell ratios, T-helper–inducer subset (CD4 + CD45RA +) % values, and proliferative responses to PHA than men scoring below the median on these scales. Disengagement coping responses also predicted one-year follow-up immune status; greater disengagement predicted poorer lymphocyte responsivity to PHA. Together these preliminary findings suggested that strategies that distract seropositive individuals from the stress of infectivity are related to greater levels of depressed affect and distress, as well as more impairment on some immunologic measures, and these relationships appeared to persist over as long as a 12-month period.

8.10.2.3.4 Social support

Social resources have been found to buffer the negative physical and psychological effects of major stressful events ranging from concentration camps to hurricanes (Lutgendorf et al., 1995; Sarason, Potter, Sarason, & Antoni, 1985; Schneiderman et al., 1992). Buffering effects of social support have been attributed to increased opportunities for intimacy, integration through shared concerns, reassurance of worth, nurturance, reliable alliance, and guidance—all potentially facilitating task-oriented thinking and active coping while decreasing helplessness and depression (Sarason, 1979). On the other hand, socially isolated individuals may be more likely to maintain maladaptive stressor appraisals such as catastrophizing, self-blame, and helplessness, and may employ denial as a chief coping strategy. The many losses in interpersonal and institutional support resources (Hoffman, 1991) may be responsible for the sense of isolation reported by some HIV-infected individuals (Dilley et al., 1985).

Issues regarding social support may be especially salient for HIV-infected women, many of whom are indigent minority group members who have been marginalized by their social groups leaving them with limited resources available to care for themselves or their children (Guinan, 1987; Nyamathi & Lewis, 1991). This threatens their roles as caregivers and mothers significantly (Buckingham & Rehm, 1987; Chung & Magraw, 1992). Trying to cope with being ill and the challenges of motherhood concurrently can be overwhelming. Having to disclose to children one's serostatus and medical condition are further challenges that these women face. Women with AIDS feel and experience stigma and shame by being presumed to be dangerous (Chung & Magraw, 1992). Community support efforts may appear to them to be primarily directed to gay men and they may respond to this with resentment and despair (Mackie, 1993).

It is important to remember, however, that gay men may also be at increased risk of social isolation. Given the number of HIV-infected gay men experiencing multiple bereavements (Martin, 1988), coupled with their reported low average support network size (comprised mainly of friends) (Namir, Alumbaugh, Fawzy, & Wolcott, 1989), any loss in such social support may be especially devastating. Other, less stable, social resources such as employment circles (Chung et al., 1989; Dilley, et al., 1985) may also be lost as the HIV infection progresses, with job loss having negative effects on both financial status, self-esteem, and emotional functioning.

In a recently published review, Zuckerman and Antoni (1995) suggest that certain social support criteria are related to optimal psychological adjustment in HIV-positive individuals. These include:

(i) a general sense of feeling supported (Namir et al., 1989);

(ii) satisfaction with received support (Hays, Turner, & Coates, 1992);

(iii) perceived helpfulness of peers (Hays, Chauncey, & Tobey, 1990);

(iv) total perceived availability of support as well as individual dimensions of support (O'Brien, Wortman, Kessler, & Joseph, 1990; Turner, Hays, & Coates, 1993; Zich & Temoshok, 1987);

(v) absence of social conflict (Lesserman et al. 1995; O'Brien et al., 1993);

(vi) greater involvement in AIDS-related community activities (Lesserman, Perkins, & Evans, 1992); and

(vii) a greater number of close friends (Hays et al., 1990).

Each of these aspects of support has been associated with lower levels of depression and, in some studies, with lower levels of hopelessness, anxiety, mood disturbance, risk behaviors, and greater overall sense of psychological well-being in a variety of populations (Christ & Wiener, 1985; Cutrona & Russell, 1987; Donlou, Wolcott, Gottlieb, & Landsverk, 1985; Friedman et al., 1991; Namir et al., 1989; Ostrow, Joseph et al., 1989; Zich & Temoshok, 1987). One study found that greater support satisfaction and greater involvement in AIDS-related community activities were associated with a greater use of coping strategies such as cognitive reframing and fighting spirit which were in turn associated with less dysphoria and greater self-esteem (Lesserman et al., 1992). Prior studies of HIV-positive men suggest that social support may determine the choice between active behavioral coping and less adaptive strategies. For instance, greater social support availability and satisfaction associated with active coping and lower support availability is associated with a greater use of denial/avoidance strategies (Martin, 1989). Also, lack of family support and other important support sources have been proposed as contributors to alcohol abuse (Nardi, 1982), though little is known about the ways in which ethnicity and socioeconomic status may moderate this relationship. In a large cohort of gay men in San Francisco, men who reported always using condoms had higher levels of social support from informal sources of help (friends, sex partners) (Catania et al., 1991). Thus social support may be viewed as a coping resource to the extent that it channels, facilitates, or perpetuates the use of coping strategies with stress-buffering properties.

Unfortunately, social support may actually decrease as the number of symptoms increases in HIV-infected individuals (Hays et. al., 1990). This may be due to caregivers and other support network members pulling back from persons with AIDS (PWA) to prevent burnout and emotional exhaustion (Turner et. al., 1993). Since family members may attempt to distance themselves emotionally from this traumatic situation, the importance of peers and fellow PWA's in the social support networks of HIV-infected individuals may increase. These people can be important sources of informational support and often provide the greatest opportunity for support (Hays et al., 1990). As such, peers often fill in the "support gaps" left behind by family and professional caregivers (Zuckerman & Antoni, 1995).

It is important to realize that some forms of social interaction may actually be associated with increased levels of psychological distress in HIV-infected persons. Among asymptomatic HIV-infected gay men, increases in social conflict, and dissatisfaction with social support in general predicted greater increases in distress and depression over a one-year period (Lesserman et al., 1995). These social conflicts can include rejection messages from family members, breaking up or declining quality of romantic relationships, and trouble with employers, among others. Another study revealed that the type of support deemed most useful depended largely on the situation confronting the individual and the context of the support provision: emotional support was generally seen as being helpful from all providers in almost any situation, while informational or problem-solving support was viewed as helpful only from medical experts and other PWA's (Zich & Temoshok, 1987). The relative effectiveness of problem-focused vs. emotional-focused support may also depend on factors such as disease stage (asymptomatic vs. AIDS) and stressor qualities such as controllability. Folkman and colleagues have suggested that problem-focused coping strategies may be most useful when directed toward controllable aspects of stressors, while emotion-focused strategies may be most effective when targeted toward uncontrollable aspects of stressors (Folkman et al., 1991). Extending the coping theory of Folkman and her colleagues to the context of matching social support resources (e.g., emotional vs. informational) to situational demands (e.g., uncontrollable vs. controllable) has been proposed as a reasonable strategy for providing the most useful support to individuals who are dealing with the different burdens of HIV infection (Zuckerman & Antoni, 1995).

In sum, inadequate social support or a sense of social isolation may exaggerate the deleterious effects of stressful events on depression and negative health behaviors in HIV-infected persons (Cutrona & Russell, 1987). A seropositive diagnosis may reduce both the availability and utilization of social support (Christ & Wiener, 1975; Donlou et al., 1985) which may increase the risk of depression, hopelessness, and less adaptive strategies such as substance use (Friedman et al., 1991; Nardi, 1982; Ostrow, Joseph et al., 1989). Similar phenomena may occur in those diagnosed with HIV symptoms or full-blown AIDS for the first time (Namir et al., 1989; Zich & Temoshok, 1987). Thus social support may be related to depression, substance use, and sexual risk behaviors at many stages HIV infection. Its influence may also vary as a function of ethnicity, among other factors such as SES and educational level. While available and adequate social supports might enhance HIV-infected people's ability to manage their infection and related stressors by functioning as a key coping resource, negative conflictual social interactions may actually act as a deterrent to successful adjustment.

8.10.3 PREDICTING THE HEALTH COURSE OF HIV INFECTION

With the emergence of data in the 1990s from long-term cohort studies, it is now well-established that some patients who are HIV positive develop symptoms rapidly while others remain free of AIDS symptoms over a decade. From a secondary prevention standpoint this means that it is important to isolate factors that account for the wide variability in the transit time for the clinical manifestations of HIV infection. Since 1990, new terms have been developed to reflect those who somehow outlive or outpace their prognosis following HIV infection. One grouping consists of nonprogressors with HIV—persons who are HIV-positive yet maintain near normal levels of CD4 numbers. Another grouping consists of those who are asymptomatic low CD4—people with very low CD4 counts ($< 50\,\text{cells}\,\text{mm}^{-3}$) who remain asymptomatic for extended periods of time. Finally, there are those termed long-term survivors of symptomatic AIDS. While the sources of individual variability in HIV progression remain largely unknown, a growing body of literature suggests that the key may lie in some combination of biological and psychological factors (Ironson et al., 1995).

8.10.3.1 Biological Factors and Disease Progression

8.10.3.1.1 Constitutional factors

Some biological factors influencing length of survival (Friedland et al., 1991; Hardy et al., 1991) include age (e.g., younger persons live more years once infected [Jacobson et al., 1991; Lemp et al., 1990]), gender (e.g., men live longer following diagnosis with AIDS [Brown, 1989]), ethnicity, and initial AIDS-related disease manifestation (having Kaposi's sarcoma vs. PCP as the presenting symptomatic criterion for AIDS (Friedland et al., 1991). Risk group membership (Detels et al., 1988; Shine et al., 1987) and presence of, or prior history of, other concomitant sexually transmitted diseases (STDs) also appear important. Disease progression markers, such as CD4 cell counts, may also be predictive of disease progression. For instance, prior rate of CD4 decline and body mass index predict the amount of time until the infected people's level of CD4 cells drops below 200 cells mm^{-3}, the point at which symptomatic AIDS often occurs (Hoover et al., 1995).

8.10.3.1.2 Other viral infections

Another set of biological factors influencing the course of HIV infection may be the degree to which other latent viruses (e.g., several types of herpes viruses) become reactivated (Rosenberg & Fauci, 1991). This point, taken in view of the psychosocial phenomena reviewed previously, is a critically important consideration for PNI research with HIV-infected populations given that (i) several psychosocial stressors have been related to immune system decrements and to signs of reactivation of one or more latent herpes viruses, and (ii) increases in sexual risk behaviors, which may ensue following some of these psychosocial phenomena, may place infected individuals at greater risk of contracting a new herpes virus infection. It has been reasoned that if stressors (Glaser & Kiecolt-Glaser, 1987) and behavioral interventions (Antoni, Esterling, Lutgendorf et al., 1995) affect the immune system's ability to control latent herpes viruses in a wide variety of populations, then stressors and interventions might also influence herpes virus-mediated progression of clinical disease in HIV-infected individuals.

Importantly, herpes viruses persist in the host indefinitely once they are infected (Rinaldo, 1990). Varicella and HSV establish latent infections in neurons, whereas CMV, Human Herpes Virus Type-6 (HHV-6), and EBV persist in lymphocytes. When an individual is infected with one of the herpes viruses, seroconversion and latency results, with or without the presence of clinical disease. In healthy individuals, herpes viruses can reactivate spontaneously and often in the absence of clinical symptoms. Under sustained states of cellular immunodeficiency, such as in persons infected with HIV or in those undergoing immunosuppressive therapy for renal transplant, reactivation of one or more of the latent herpes viruses results in severe morbidity (Ernberg, 1986; McVoy & Adler, 1989; Rinaldo, 1990; Yao, Rickinson, Gaston, & Epstein, 1985). Moreover, herpes viruses are known to suppress a wide number of immunologic functions including cytokine production and lymphocyte blastogenesis to mitogens (Rinaldo, 1990). These phenomena suggest that HIV and herpes viruses may interact to perpetuate disruptions in the biological control of each respective infection, possibly contributing to the severity and chronicity of related clinical signs and symptoms.

Because several herpes viruses prevalent in HIV-infected individuals are known to have independent immunosuppressive effects, reactivation of these viruses could have implications for progression to AIDS (Griffiths & Grundy, 1987). Subclinically these surveillance effects may be observed as rising antibody (IgG) titers to EBV, HHV-6, CMV, and HSV-2. One could speculate that clinically, the inability to survey these latent viruses may be manifest in EBV-associated Burkitt's lymphoma (Doll & List, 1982) and oral hairy leukoplakia (Eversole, Stone, & Beckman 1988; Greenspan et al., 1985); CMV-associated retinitis, encephalitis hepatitis, adrenalitis, and gastrointestinal pathologies (Blumberg, 1984; Henderly et al., 1987; Schooley, 1990; Tapper et al., 1984); and HSV-associated proctitis (Goodell et al., 1983), esophagitis (Levine, Woldenberg, Herlinger, & Laufer, 1987), and aseptic meningitis (Dahan, Haettich, LeParc, & Paolaggi, 1987).

There is also evidence that reactivation of several of these herpes viruses may act to further stimulate HIV replication and progression to AIDS (Carbonari et al., 1989; Hammarskjold & Heimer, 1988; Lusso et al., 1989; Rosenberg & Fauci, 1991; Sumaya, Boswell, & Ench, 1986). CMV, one of the herpes viruses, is an important opportunistic pathogen in HIV infection (Rinaldo, 1990) and has been shown to enhance susceptibility to secondary opportunistic pathogens in immunosuppressed populations (Perssan et al., 1987). It is among the most immunosuppressive of the herpes viruses impairing blastogenic responses and γ-interferon production by T-lymphocytes (Rinaldo, 1990) and it has also been shown to alter the growth and expression of HIV in culture, possibly accelerating progression to AIDS among infected persons (Biegalke & Geballe, 1991; Paya, Virelizier, & Michelson, 1991). Webster (1991) found that among HIV-infected individuals, the age-adjusted relative risk of developing AIDS was 2.5 times higher in those who were coinfected with CMV compared to those who were CMV seronegative.

In sum, although some mechanisms are still being explored, the fact that herpes viruses may be related to HIV associated sequelae is intriguing. For example, because HHV-6 may reactivate during EBV infections, and because both EBV and HHV-6 may reactivate following immunosuppression by HIV, these three viruses may potentiate one another in the progression of clinical disease. Through direct immunosuppression or transactivation of HIV (by CMV or HSV), latent herpes viruses may be considered as potential cofactors in the development of AIDS. It is quite plausible that HIV-associated effects on the cells of the immune system may perpetuate a downward spiral in the functional ability of the immune system to survey and control latent viruses which, when left unchecked, may act to increase replication of the primary virus.

The implication of these findings is that the host's ability to control such viruses may be a "downstream" immune system marker that has more meaning as a predictor of HIV clinical disease and possibly the acceleration of viral load than does the CD4 cell count alone. We know that changes in HIV viral load, viral character (appearance of syncytium-inducing variants), and CD8-directed cytotoxity against HIV are the key predictors of individual differences in clinical disease progression (Pantaleo et al., 1993). These "predictors" all speak to the importance of the host maintaining vigorous surveillance and control over HIV (through well-coordinated cytotoxic responses), and stabilized HIV replication (reflected in stable viral load and predominantly nonsyncytium-producing types). Given this understanding of the importance of qualitative aspects of the immune system (surveillance functions and signs of poorly surveyed viruses) in HIV disease progression, it is surprising that so little research has been published on the effects of biomedical and psychosocial interventions on qualitative immune measures including those reflecting surveillance of specific pathogens such as latent herpes viruses.

8.10.3.2 Behavioral and Psychosocial Factors and Disease Progression

8.10.3.2.1 Behavioral factors

Behavioral factors such as the use of alcohol, tobacco, caffeine, recreational drugs (cocaine, nitrites) during sex, unprotected anal intercourse, and poor nutrition have also been suggested as possible cofactors of accelerated HIV disease progression (Chiappelli, Ben-Eliyahu, Hanson & Wylie 1992; Griensven et al., 1990; Haverkos, Pinsky, Drotman, & Bregman, 1985). It is noteworthy that several of these behaviors have been associated with immunomodulation in other populations. Cigarette smoking has been associated with significant decreases in helper cells and increases in suppressor/cytotoxic cells (Miller, Goldstein, & Murphy, 1982). Alcohol use has been related to depressed cell-mediated immunity (Watson, Eskelson, & Hartman, 1984), decreased lymphocyte number and diminished lymphocyte blastogenesis (Glassman, Bennet, & Randal, 1985), bone marrow myelosuppression (Tisman et al., 1971), and abnormal granulocyte chemotaxis (Atkinson et al., 1977). Other work suggests that alcohol abuse may amplify the suppressive effects of mental depression on immunologic impairments (Irwin, Caldwell et al., 1990). Recreational drug usage (e.g., heroine) has been associated with depressed cellular immune functioning in one study (Lazzarin, Mella, & Trombini, 1984), and with suppressed NK cell killing in another (Katz, Zaytoun, & Faici, 1982). Some of this work has focused specifically on HIV-infected samples. For instance, there is growing evidence that injection drug use (Seage et al., 1993; Weber, Ledergerber, Opravil, Siegenthaler & Luthy, 1990) and even cigarette smoking (Burns et al., 1991; Nieman, Fleming, Coker, Harris, & Mitchell, 1993; Royce & Winkelstein, 1990) may be associated with faster disease progression in HIV-infected people. The specific mechanisms by which these substances affect immune functioning are still in the process of being elucidated. It is noteworthy that some of these substances are associated with alterations in those physiological stress response systems noted previously. For instance, nicotine in cigarette smoke is associated with catecholamine discharge (Blaney, 1985) and ethanol consumption is also linked with catecholamine elevations due to blockage of re-uptake (Davidson, 1985).

Nutritional factors and vitamin deficiencies may contribute to immunologic decrements in HIV infection. Protein-caloric malnutrition is associated with a wide range of immune impairments and is believed to explain the increased susceptibility to disease in undernourished populations of many Third World countries (Stites, Stobo, Fudenberg, & Wells, 1982). Some work also suggests that alterations in certain vitamin levels, for instance, vitamin B^6 deficiencies, may be associated with immune system decrements in HIV-infected persons (Baum et al., 1991).

A series of studies has presented evidence for immunomodulatory effects of sleep deprivation including decreases in lymphoproliferative responses and diminished granulocyte functioning (Palmblad et al., 1976; Palmblad, Petrini, Wasserman, & Akerstedt, 1979). One study, done in the mid-1990s, noted that the decrements in immune function (NKCC) evident in victims of a hurricane over the months following the storm were mediated, in part, by sleep disruptions (Ironson et al., 1997). Other work has indicated that human IL-1 levels peak at the onset of slow wave sleep (Krueger & Karnovsky, 1987), suggesting that cellular immunity may be affected by qualitative, as well as quantitative, aspects of sleep and accompanying cytokine changes. Important for HIV-infected populations is the fact that sleep disruptions are one of the hallmarks of clinical depression.

Numerous other biological factors may alter stress hormone levels and immune functioning such as aging, steroid use, sex hormone fluctuations, radiation treatment, as well as a multitude of medication regimens (Stites et al., 1982). Generally these variables are not likely to be responsive to health psychology interventions and thus will not be reviewed here. However, it is essential that these phenomena and others be measured and carefully controlled in any investigations of the influence of biological, behavioral, and psychosocial factors in HIV disease progression.

Another variable that may deserve further attention is heightened sympathetic nervous system responsiveness to stressful situations. The stress hormonal elevations known to accompany this hyper-responsiveness (Rose, 1980) may have immunomodulatory effects. Interestingly, a growing literature suggests that pharmacologic agents that attenuate cardiovascular hyper-responsiveness (Bonelli, 1982; Dimsdale, Hartley, Ruskin, Greenblatt, & Labrie, 1984) may also alter immune functioning (Benschop et al., 1994; Hatfield, Petersen, & DiMicco, 1986; Weisdorf & Jacob, 1987). There is also some evidence that HIV-infected persons may display altered cardiovascular, endocrine, and immunologic reactivity to laboratory stressors (Kumar, Morgan, Szapocznik, & Eisdorfer 1991; Starr et al., 1996), indicating that these individuals may not respond to and recover from stressful challenges in the same way that healthy, uninfected people do.

Although there is little known about the direct immunomodulatory effects of different forms of sexual behaviors, it is plausible that repeated exposure to novel strains of HIV and other sexually transmitted pathogens by way of unprotected intercourse (e.g., HSV and human papilloma viruses [HPVs], may contribute directly to the development of opportunistic infections (e.g., systemic HSV-2 infections) and neoplasias (e.g., HPV-associated cervical and anal intraepithelial neoplasias [Antoni & Goodkin, 1996]). In addition, exposure to these pathogens may indirectly contribute to an accelerated disease course by promoting HIV replication (for review see Antoni, Esterling, Lutgendorf et al., 1995). Therefore, efforts to modify risky sexual behaviors in already infected individuals can be seen as a secondary prevention strategy. Finally, PCP treatment prophylaxis, antiretroviral combination therapy, and the new breed of protease inhibitors may influence survival after AIDS onset (Saah et al., 1994). From a behavioral standpoint, an HIV-infected individual's decision and actions taken to pursue these treatments, at the earliest point following the emergence of symptoms or after dropping below landmark values for CD4 cell counts (e.g., $< 200\,\text{cell mm}^{-3}$), may play a large role in health maintenance.

We have reasoned that if psychological factors affect the immune systems of HIV-infected individuals, then even small immune changes might have clinical relevance since this population's immunologic functioning is already significantly impaired by the virus (Antoni, Schneiderman et al., 1990; Schneiderman et al., 1994). Healthy uninfected people in the normal population possess many checks, balances, and compensatory mechanisms built into their immune systems. Since HIV-infected individuals are just at, or above, the number of T-lymphocyte subsets necessary for responding to encountered pathogens (e.g., 200–500 CD4 cells mm^{-3}), then small changes, of unknown magnitude, could quite possibly have an impact on the incidence of a wide range of infections due to normally harmless bacterias, viruses, fungi, and other opportunistic pathogens.

In addition to behavioral factors, there is some evidence that several psychosocial variables may also contribute to individual differences in disease progression among asymptomatic HIV-infected individuals as well as contributing to differences in survival time in those who have already developed AIDS. Specifically, longitudinal studies have identified psychological factors such as depressive symptoms, stressful life events, maladaptive coping strategies, and social isolation that predict either survival time or CD4 changes over a period of years.

8.10.3.2.2 Depressive symptoms

There is considerable evidence that depressive symptoms are correlated with immunosuppression in otherwise healthy individuals (Herbert & Cohen, 1993; Irwin, Caldwell et al., 1990; Irwin, Patterson et al., 1990), though available evidence suggests that depression and affective disturbances may be associated with physical disorders in a bidirectional fashion (Cohen & Rodriquez, 1995). Much of the work relating clinical depression and depressed affect to immunologic changes is plagued with inconsistencies which may be attributable, in part, to the use of small samples, variability in control/comparison groups utilized, and variance in the depression diagnostic criteria employed (Stein et al., 1991). However, alterations in NKCC appear to be the most commonly observed immunologic correlate of depressed affect and depressive disorder employed (Stein et al., 1991). Among the more recent and well-controlled studies, Irwin, Caldwell et al. (1990) found significant NKCC decrements associated with depressive disorders.

Longitudinal studies relating depressive symptoms to rates of immunologic decline and disease progression in HIV-infected persons have produced mixed results (Burack et al., 1993; Lyketsos et al., 1993; Patterson et al., 1996). Burack and colleagues found that greater depressive symptoms predicted a faster decline in CD4 counts but were not associated with time to AIDS or survival time in early-stage HIV-infected gay men followed over a six-year period (Burack et al., 1993). Lyketsos used similar criteria for documenting the presence of depressive symptoms and found no prospective relationship with either CD4 count slopes of decline or clinical disease progression over a similar period (Lyketsos et al., 1993). However, there was a tendency for the men in the Lyketsos study to be at a more advanced stage of CD4 cell loss at start of the study than those men in the Burack study. Another study following HIV-infected men over a five year period observed that depressive symptoms predicted shorter survival time after controlling for initial symptoms and CD4 cell number (Patterson et al., 1996). Other psychosocial factors may interact with depressive symptoms or distress levels in predicting changes in immune status and health. Because depressed affect, substance use, and risky sexual behaviors may all contribute to health outcomes in HIV-infected individuals, it seems reasonable that factors which moderate the occurrence of these sequelae may also have an influence on immunity and health. Some of these include the occurrence of overwhelming stressful life events, stressor appraisals, individual differences in coping skills, and differences in social support.

8.10.3.2.3 Stressful life events

There is now ample literature (Herbert & Cohen, 1993) showing that stressors are associated with decrements in immune function. Some of these stressors may be particularly prevalent among HIV-infected people attempting to cope with their illness and include: bereavement (Bartrop, Lazarus, Luckhurst, Kiloh, & Penny, 1977; Irwin, Daniels, Bloom et al., 1987; Kemeny et al., 1995; Linn, Linn, & Jensen, 1981); unemployment (Arnetz et al., 1984); the stress of being a caregiver for a loved one with a terminal disease (Kiecolt-Glaser, Glaser et al., 1987; Kiecolt-Glaser et al., 1991); divorce (break-up) (Kiecolt-Glaser, Fisher et al., 1987); and marital distress (Kiecolt-Glaser, Fisher et al., 1987; Kiecolt-Glaser et al., 1988). Chronic uncontrollable stress has been operationalized in one form as people dealing with catastrophic environmental changes. These stressors are also associated with immune alteration—for instance, both at the Three Mile Island nuclear plant disaster (McKinnon, Weisse, Reynolds, Bowles, & Baum, 1989) and in Miami after Hurricane Andrew (Ironson et al., 1997; Lutgendorf et al., 1995).

Recently conducted investigations indicate that the most consistent immune decrements found in studies of stressors are functional measures associated with cellular immunity (Bartrop et al., 1977; Glaser, Rice, Speicher, Stout, & Kiecolt-Glaser, 1986; Glaser, Kiecolt-Glaser, Speicher, & Holliday, 1985; Glaser et al., 1987, 1990, 1991; Ironson et al., 1995; Irwin et al., 1990; Kemeny et al., 1995; Kiecolt-Glaser et al., 1985, 1987, 1988; Kiecolt-Glaser, Dura, Speicher, Trask, & Glaser, 1991; Levy, Herberman, Lippman, d'Angelo, 1987; Linn et al., 1981). These include, but are not limited to:

(i) decrements in blastogenic responses to phytohemaglutinin (PHA) and pokeweed mitogen (PWM);

(ii) decreased natural killer cell cytotoxicity (NKCC);

(iii) increased antibody titers to Epstein-Barr Virus (EBV) and Herpes Simplex Viruses (HSV);

(iv) poorer DNA repair; and

(v) disturbance in cytokine regulation including less γ-interferon and interleukin-2 (IL-2) production and IL-2 receptor gene expression.

There are some PNI studies that have related stress responses to people's ability to respond to viral pathogens. Cohen and colleagues (Cohen, Tyrrell, & Smith, 1991) innoculated volunteers with one of five rhinoviruses (cold viruses) in a restricted environment and then followed them over a period of weeks. Subjects with the highest levels of perceived stress were significantly more likely to display increases in antibody titers to the specific rhinovirus and clinical colds than those with the lowest levels of psychological stress. Stressors have also been related to alterations in primary immune responses to other viruses. Glaser et al. (1992) demonstrated that those medical students undergoing a stressor (exams) who were least stressed and anxious were also the most likely to seroconvert (produce a primary antibody response to) the Hepatitis B virus (HBV) after the first vaccine innoculation. In addition, participants with the highest levels of social support had a stronger immune response to the vaccine at the third innoculation. Thus, participants with low stress and anxiety, and high social support, appear to have been relatively well protected from infection by HBV. Together these studies suggest that stressors may influence mechanisms controlling both primary and secondary immune responses to viruses. This work is important for understanding the role of stressor–immune associations as they possibly relate to health outcomes in the context of HIV infection since an HIV-infected person's inability to control viral co-infections may act to aggravate the course of primary HIV infection (Antoni et al., 1995).

Despite some of the evidence associating stressful life events with changes in specific indices of immune system functioning in healthy populations that has accrued since the mid-1980s, there is less conclusive evidence linking stressors with immune and health changes in HIV-infected individuals. Briefly, two studies found that stressful life events were unrelated to CD4 counts or HIV-related symptoms (Kessler et al., 1991; Rabkin et al., 1991), while another found that elevated life events were predictive of decreases in CD4 counts (Goodkin, Fuchs, Feaster, Leeka, & Rishel, 1992). It should be noted that Goodkin and colleagues analyzed their data in line with a stressor–moderator model which simultaneously takes into account the direct effects of life events, plus their interaction with theoretically derived moderator variables such as coping and social support. Such more sophisticated statistical modeling may be necessary to delineate the often interactive nature of life event effects on the immune system and health.

One stressful event associated with disease progression in HIV-infected persons is bereavement (Coates, Stall et al., 1989; Kemeny et al., 1994). Coates, Stall et al. (1989) found that both the number of losses and accompanying distress levels were significantly related to faster HIV progression during a two-year follow-up. Similarly, Kemeny et al. (1994) found that a bereaved group of HIV-infected men who had lost a partner had significantly higher serum neopterin levels and lower proliferative responses to PHA than a matched group of nonbereaved HIV-positive men. This suggests that life events that are commonly experienced by HIV-infected individuals may have significant effects on the functioning of their immune system, but that gross changes in CD4 counts and clinical symptoms may take longer to develop. While the Rabkin et al. (1991) and Kessler et al. (1991) studies found no association between stressful life events and outcome measures, they did not assess individual differences in subjects" perceived control over stressful life events nor their coping responses for dealing with these stressors. While other life events have been associated with some immune measures in HIV-infected people (Evans, Pettito, & Leserman, 1992; Goodkin et al., 1992), stressful life events per se have not been shown to be predictive of a faster rate of disease progression (Kessler et al., 1991). Patterson and colleagues note that the interaction between mounting life events and distress levels may be a more reliable predictor of immunologic and health changes in HIV-infected persons (Patterson et al., 1996). Specifically, individuals who display the greatest distress levels during particularly stressful periods may show the poorest health outcomes. It is plausible that those individuals who are at greater risk of displaying such distress levels are those who utilize maladaptive stressor appraisals and coping strategies and who also have inadequate social resources.

8.10.3.2.4 Stressor appraisals

Some aspects underlying the cognitive processing of stressors have been associated with measures of physiological activation and immune functioning in healthy people. Davidson and Baum (1986) found that both intrusive imagery and avoidant cognitions, as measured by the Impact of Event Scale (IES) (Horowitz, Wilner, & Alvarez, 1979), were positively associated with increased cortisol levels in residents living near Three Mile Island (TMI) following the nuclear reactor accident. Many of these individuals displayed elevated antibody titers to certain viruses, suggesting impairments in immunologic surveillance years after the reactor accident (McKinnon et al., 1989). Workman and LaVia (1987) found that intrusive cognitions (measured by the IES) related to an impending medical school examination were associated with poorer blastogenic responses to PHA among young healthy medical students. Ironson and colleagues found that the intrusive thoughts and post-traumatic stress symptoms experienced by many victims of Hurricane Andrew were associated with decrements in NKCC (Ironson et al., 1997).

One study examined how the manner in which HIV-infected individuals processed the stress of seropositivity notification predicted their immunologic status during the early period of adjustment to this stressor. They found that changes in style of cognitive processing of this stressor were significantly correlated with changes in affective and immunological functioning over this 10 week period (Lutgendorf, Antoni, Ironson, Schneiderman et al., 1997). The IES and the Profile of Mood States (POMS) (McNair, Lorr, & Droppleman, 1981) were administered five weeks before (baseline) and five weeks after notification (week 10). A panel of immune assays were also conducted on blood samples collected at these time points. The IES instructions required subjects to report on intrusive thoughts that they had been experiencing in the past week regarding one of the most salient issues that infected people deal with on a daily basis: the threat of AIDS in their life. The IES also tapped the degree to which they had attempted to distract themselves or avoid thinking about this threat over this period. Participants were classified as those who had IES-Intrusion score increases vs. decreases and those who had IES-Avoidance score increases vs. decreases over the 10-week period. Results indicated that men with avoidance increases demonstrated greater anxiety, depression, and confusion at week 10 than those individuals who had avoidance decreases. Subjects with avoidance increases also displayed lower proliferative responses to PWM and lower NKCC at the end of the 10 weeks compared to those who decreased in avoidance scores over the study. These findings suggested that postnotification processing, colored by avoidance, was associated with poorer immune system functioning in the weeks after antibody testing among asymptomatic HIV-positive men.

Another stressor appraisal variable that has received some attention in HIV research concerns a construct called fatalism. In a series of studies, Kemeny (reviewed in Kemeny, 1994) found fatalistic appraisals to be related to poorer health outcomes in HIV-infected gay men. In the first study, fatalistic gay men with a diagnosis of AIDS had a significantly shorter survival time than men scoring low on fatalism (Reed, Kemeny, Taylor, Visscher et al., 1994). Interestingly, men who were both fatalistic and bereaved within the past year had the shortest survival time (Reed, Kemeny, Taylor, Visscher et al., 1994), the steepest declines in CD4 number and proliferative response to PHA, and a more rapid increase in disease progression markers, including neopterin and beta-2 microglobulin (Kemeny et al., 1995). Less is known about the ways in which other cognitive appraisal variables, such as optimism, different cognitive distortions, such as catastrophizing, and self-efficacy might relate to the course of HIV infection.

8.10.3.2.5 Coping strategies

There is some evidence that the use of certain coping strategies for dealing with HIV infection may be associated with the rate of disease progression. Ironson et al. (1994) found that the use of denial and behavioral disengagement to cope with an HIV-positive diagnosis, was predictive of lower CD4 counts at 1 year follow-up (controlling for CD4 counts at entry to the study) and greater likelihood of progression to HIV-related symptoms and AIDS at 2-year follow-up. A more fine-grained analysis exploring the immunologic changes that occurred in this sample prospectively related coping strategies and immunologic measures in the HIV seropositive gay men (and a seronegative comparison group of gay men) over the 3 week and 1-year period following notification of their antibody test results (Antoni et al., 1995). Seropositive men scoring above the median on postnotification disengagement coping strategies (denial, behavioral disengagement, mental disengagement) had significantly lower concurrently measured T-helper/suppressor (CD4/CD8) cell ratios, T-inducer subset (CD4 + CD45RA +) percent values, and proliferative responses to PHA than subjects scoring below the median on these coping scales. Greater disengagement coping responses also predicted poorer lymphocyte responsivity to PHA at 1-year follow-up.

Consistent with these findings, Goodkin et al. (1993) found that use of passive coping strategies (including denial and disengagement) was inversely related to long-term CD4 cell count in HIV-positive gay men. Solano et al. (1993) also found that denial/repression was associated prospectively with the emergence of symptoms in an HIV-positive sample. Mulder, Antoni, Duivenvoorden et al. (1995) found that active confrontational coping with HIV infection was predictive of decreased clinical progression over a 1-year period after taking into account baseline biomedical and behavioral variables. Finally, Blomkvist et al. (1994) followed HIV-infected hemophiliacs longitudinally and found that self-oriented active-optimistic coping behavior was related to prolonged survival after controlling for age and CD4 counts at study entry. Thus, there is some evidence that coping strategies may be reliably associated with certain immunologic and health measures in HIV-infected persons from a variety of risk groups.

It should be noted that while Reed, Kemeny, Taylor, Wang, & Visscher (1994) found that realistic acceptance was a significant predictor of decreased survival time in gay men diagnosed with AIDS, Ironson et al. (1994) observed that greater use of denial among asymptomatic HIV-infected gay men predicted a greater likelihood of developing symptoms or AIDS two years later. Thus it appears that being at the extreme end of either denial or realistic acceptance might be dysfunctional. Kemeny refers to the realistic acceptance factor as fatalism–acceptance, and notes, specifically, that "fatalistic men survived a significantly shorter period of time when compared to their less fatalistic counterparts." What is not clear is whether this variable represents acceptance as a healthy coping style or a predominantly negative view (pessimism) about the future. One study may shed some light on the notion of a balance between the extremes of denial and acceptance. Mulder et al. (in press) found that greater use of distraction (e.g., focusing on other things to take the mind off things) predicted a slower rate of decline in CD4 cells, less appearance of syncytium-inducing HIV variants, and less progression to immunologically defined AIDS (< 200 CD4 cells mm^{-3}) over a 7-year period. Ironson et al. (1995) notes that together these findings suggest that strategies which are predictive of longevity may involve using some distraction coupled with staying away from the extremes of either denial or acceptance to the point where fatalism and rumination occurs.

8.10.3.2.6 Social support

The health promoting effects of social support that have been most widely researched are usually in terms of its ability to buffer the effects of stressors (Cohen & Wills, 1985). A person's social support network may also influence their ability to manage, cope with, and recover from serious illness (Dimatteo & Hays, 1981; Taylor, Falke, Shoptaw, & Lichtman, 1986; Wortman & Conway, 1985). Since the 1970s there has been an enormous amount of research undertaken studying the ability of social networks to provide positive physical health benefits to people (Berkman & Syme, 1979; House, Landis & Umberson, 1988; Smith, Fernegel, Holcroft, Gerald, & Marien, 1994; Wortman & Dunkel-Schetter, 1987). Some findings have converged across investigations. For instance, it appears that being involved in a committed relationship (i.e., marriage) is associated with greater survival time among patients who have experienced a myocardial infarction (Wiklund et al., 1988; Williams et al., 1992) or a diagnosis of breast cancer (Neale, Tilley, & Vernon, 1986). However, these associations have not been replicated in other studies. Although a growing body of evidence supports the potential benefits of social ties in healthy people and several chronic disease populations (e.g., cancer patients), scientists are just beginning to examine the health correlates of social support in HIV-infected individuals.

A variety of models have been developed to explain the ways in which social support might provide direct influences on psychological and physical well-being (Cohen & Wills, 1985). Some of these direct psychological benefits may stem from the possibilities that social support helps people define certain stressors as being less overwhelming, provides opportunities for people to openly express fears, frustrations, and other emotions, thereby decreasing ruminations and obsessive thinking about stressors. Social support also provides key information and informational support provided by knowledgeable others, which may help individuals to assume a more realistic attitude toward life goals. In addition, social support may help people feel connected, which may increase their sense of self-worth, belongingness, meaning, purpose, and well-being. Finally, social ties may reduce initial feelings of burden caused by the onset of a stressor (e.g., a medical diagnosis), allowing the person to employ active coping strategies to deal with the stressor. Obviously, all of these benefits of social support may be highly salient for HIV-infected persons (Zuckerman & Antoni, 1995).

Some research suggests that the same mechanisms which provide these psychological benefits from social support also seem to buffer endocrine and immunologic changes associated with stress responses (e.g., Baron, Cutrona, Hicklin, Russell, & Lubaroff 1990; Cobb, 1974; Esterling, Kiecolt-Glaser, & Glaser, 1996; Levy et al., 1987). It has been hypothesized that HIV-infected people who maintain strong support networks may have less extreme and protracted biological responses to stressors because SAM and HPA-related stress hormone levels (e.g., peripheral catecholamines and cortisol) are better regulated (Antoni et al., 1990). Some work provides preliminary evidence that social support may relate to the rate of HIV disease progression as well. One study found that larger social network size and greater perceived informational social support predicted greater survival time in HIV-infected men, but only among those with AIDS symptoms (Patterson et al., 1996). Other prospective studies have found that greater social support availability (Theorell et al., 1995) and more frequent use of problem-focused support (Solomon, Temoshok, O'Leary, & Zich, 1987) were predictive of a slower decline in CD4 counts and longer survival, respectively. Another study may illuminate a factor underlying an HIV-infected person's decision to enlist social support. Cole, Kemeny, Taylor, Visscher, and Fahey (1996) followed 80 HIV seropositive (non-AIDS, normal levels of CD4 lymphocytes at entry) gay men for 9 years and found that those who concealed their sexual identity (being at least "half in the closet") had a faster decline in CD4 counts, shorter time to AIDS diagnosis, and shorter time to AIDS mortality than those who did not conceal their identity. Therefore, social support availability, decisions to utilize available support, and identity factors that influence decisions to make these connections may all contribute to the course of HIV disease.

In sum, it could be hypothesized that a mounting number of losses and grieving experiences, which constitute prevalent life stressors for HIV-infected individuals, might contribute to a faster rate of disease progression. Further work suggests that maintaining fatalistic stressor appraisals could predict shorter survival time among men with AIDS. Moreover, attempting to avoid dealing with intrusive thoughts concerning future AIDS risk was associated with greater immune system decrements in asymptomatic HIV-positive men. The use of coping strategies such as extreme denial or acceptance were both associated with an accelerated disease course, while a certain degree of distraction (and perhaps low rumination about the disease) was associated with a slower rate of progression. Finally, people who remain socially isolated (and who may be the least comfortable in approaching their support network) appear to show the fastest progression of disease.

8.10.3.3 Psychosocial Factors in Long-term Survivors of AIDS

Another research strategy for establishing salubrious psychosocial factors in HIV-infected individuals has been to document the characteristics of those who substantially "outlive" their prognosis after a diagnosis of AIDS. This research has been reviewed by Ironson and colleagues (Ironson et al., 1995). The CDC currently estimates median survival time after an AIDS diagnosis (by the 1987 symptomatic criteria) to be 18–20 months (CDC, personal communication 1994; Eickhoff, 1994), a number that increases slowly each year as better diagnostic and treatment options become available. The CDC defines "long-term survivors" of AIDS as persons who survive twice the median expected time. Only a few studies of the psychosocial characteristics of long-term survivors of AIDS (defined by clinical symptoms) have been done, and as Ironson and colleagues note, almost all of these have been cross-sectional (Ironson et al., 1995). Ironson and her colleagues did, however, find a good deal of consensus on the existence of several different psychosocial characteristics of these individuals, noting four psychosocial adjustment strategies (Ironson et al., 1995): (i) healthy self-care behaviors; (ii) a sense of connectedness; (iii) a sense of meaning and purpose; and (iv) maintaining perspective. The first factor included three components—medical care, a healthy lifestyle, and an awareness and ability to take action to meet personal needs. The second factor, maintaining connectedness, is closely related to the maintenance of social support. Having at least one confidant is part of this factor, as is being able to communicate openly about concerns and, for homosexuals, disclosing a gay lifestyle. This is interesting in light of the findings of Cole and colleagues (Cole et al., 1996). The third factor, maintaining a sense of meaning and purpose, ties into the notion of having something to live for and, thus, represents a cognitive appraisal factor. Ironson and colleagues note that having an optimistic attitude, (or conversely being low on negative expectancies) are all a part of this. Some long-term survivors actually report finding new meaning as a result of being HIV-positive (Schwartzberg, 1994). Still others find meaning in their relationship with a higher power, or spirituality/religion as they define it (Woods, Antoni, Ironson, & Kling, 1996). The final factor is maintaining perspective. A failure to maintain perspective is represented by depression, negative affect, and hopelessness. Ironson et al. (1995) note that accepting the reality of an AIDS diagnosis without seeing it as an imminent death sentence is part of this perspective, as is believing one can live with AIDS, realizing AIDS is episodic in nature, and remembering that AIDS usually includes periods of good health. Using selective distraction as a coping strategy and not allowing AIDS to become the sole focus of one's life captures the essence of this factor. Ironson et al. (1995) conclude that maintaining perspective is quite close to the idea that balance is very important. They cogently use the optimism/negative expectancies construct and the denial/acceptance construct to point out that the extreme at either end is probably the least functional. It remains to be determined which psychological and biological mechanisms might explain the ways in which these four long-term survivor strategies relate to health outcomes in HIV-infected people.

8.10.4 INTERVENTIONS, ADJUSTMENT, AND HEALTH COURSE IN HIV INFECTION

8.10.4.1 Health Psychosocial Interventions with HIV-infected People

Since 1990 a small collection of studies have examined the effects of psychosocial interventions in HIV-infected populations. These studies used both psychological and immune measures as dependent variables, although none of them have completed long enough follow-up periods to clearly establish a health/disease progression effect. Most of these interventions employ cognitive behavioral techniques and operate in a group-based format. Comprehensive reviews of the psychological effects (Chesney & Folkman, 1994), and immune and health effects (Ironson et al., 1995) of psychosocial interventions in HIV populations are now available.

The psychological, immune, and health effects of psychosocial interventions in HIV-infected individuals are summarized in Table 1. Studies included in this table are restricted to those that monitored intervention-related changes in both psychological and immunologic status. The first study to examine the effects of stress management intervention on psychological and immune status in HIV-infected individuals was conducted by Coates and colleagues at the University of California at San Francisco (Coates, McKusick, Stites, & Kuno, 1989). They randomized 64 HIV-positive gay men to either an 8-week stress management training program or a control group. The intervention consisted of systematic relaxation (with take-home tapes), health habit education (diet, rest, exercise, drug and alcohol use, smoking), and skills for managing stress. While the intervention appeared to change sexual behavior (significantly fewer sexual partners), there were no effects on enumerative or functional immune measures. The panel of immune assays included: CD4 cells CD4/CD8 ratio, NKCC, lymphocyte response to con A, candida antigen and CMV. Psychological changes in the intervention or control groups were not reported though the authors acknowledge that the protocol they used may have been ineffective in reducing stress.

Another study which had some positive effects on psychosocial variables, but no effect on CD4 cell number, was done by Auerbach and colleagues (Auerbach, Oleson, & Soloman, 1992). Twenty-six symptomatic seropositive gay men were randomized to an 8 week behavioral medicine intervention, thermal biofeedback, guided imagery, and hypnosis, or a wait-list control group. The intervention reduced self-reported symptoms of HIV (fever, fatigue, pain, headache, nausea, and insomnia) and increased vigor and hardiness relative to controls, but did not change anxiety or depression. The authors speculate that if they had selected a more distressed HIV-infected sample they may have obtained changes in immune status variables and anxiety and depression measures. It is also possible that stress management interventions may work best if the groups start out with high stress or encounter a stressor during the course of the intervention. The most appropriate way to evaluate the effects of these types of interventions might be to test them in the context of either high distress states or measurable recent or ongoing stressful events (e.g., bereavement, diagnosis notification) and to include a comparison group of HIV-negative controls.

8.10.4.2 Stress Management Intervention Effects at Different Critical Points in HIV Infection

Some of the most well-documented behavioral interventions applied to HIV-infected individuals have attempted to use procedures that reduce distress, promote self-efficacy, and encourage the use of active coping strategies. Operating from a chronic disease-PNI rationale we have hypothesized that if behavioral interventions can retard or normalize declines in immune status by bringing about these psychosocial and behavioral changes, then they may be able to forestall the onset of disease complications, such as opportunistic infections and neoplasias (see Figure 1). Our group developed a 10-week CBSM intervention for HIV infection that was based specifically upon the psychosocial sequelae that have been reviewed previously (Antoni, Baggett et al., 1991). For some HIV-infected people these sequelae include being faced with a wide variety of chronic, uncontrollable, and unpredictable stressors (e.g., such as changes in health, job status, health insurance, and medical costs) and a loss of social support resources which is triggered by their own withdrawal, the death of close friends, or the responses of friends and family. These are likely to persist across and grow in magnitude over the course of the infection (see Hoffman, 1991).

This model reasoned that the combination of multiple uncontrollable burdens and declining coping resources may overwhelm HIV-infected persons" previously adequate direct coping strategies such as active coping, positive appraisals of burdens, problem-solving, and seeking social support, resulting in the adoption of more indirect strategies such as avoidance

Table 1 Psychosocial intervention effects in HIV infection.

Team	*Intervention*	*Sample*	*Psychosocial effect*	*Immune effect*	*Health*
Coates, McKusick et al. (1989)	8 week stress management vs. no tx. ctrl.	64 HIV-positive gay men	↓ number sex partners, distress ND	no Δ CD4 no Δ PHA	ND
Auerbach et al. (1992)	8 week biofeedback imagery, hypnosis vs. WLC	26 HIV-positive gay men	↑ POMS vigor ↑ hardiness no Δ POMS anx. dep.	no Δ CD4	↓ HIV sx.
(i) Antoni, Baggett et al. (1991) (ii) Antoni, Ironson et al. (April, 1992) (iii) Esterling et al. (1992) (iv) Ironson et al. (1994) (v) Lutgendorf et al. (1997)	(i)–(iv) 10 week CBSM vs. no tx. ctrl. (v) 10 week CBSM vs. WLC	(i), (iii), & (iv) 47 HIV-positive asymptomatic men learning ab. status (ii) asymptomatic HIV-positive men (v) status known > 6 mo 40 symptomatic HIV-positive non-AIDS	(i) buffered notification POMS distress Δ (i) maintained social support (ii) ↓ behavioral and mental disengagement ↓ denial (v) ↓ BDI depression ↓ POMS distress ↑ social support, ↑ acceptance	(i) buffered CD4 Δ, CD56 Δ, & PHA Δ (ii) pre–post notification ↑ NKCC, ↑ PHA (iii) ↓ EBV ab. ↓ HHV-6 ab. (v) ↓ HSV-2 ab. (ii), (iii), & (v) no Δ CD4	(i), (ii), (iii), & (v) ND (iv) 2 year disease progression correlates ↓ adherence ↑ denial ↑ distress
Mulder et al. (1994, 1995)	12 week CBSM, 12 week EE vs. WLC	39 HIV-positive asymptomatic gay men, ab. status known	↓ POMS distress ↓ BDI depression no Δ coping, social support emo. exp.	distress Δ, CD4 Δ over 2 year	ND
Ironson, Field et al., 1995	4 week daily massage therapy vs. WLC	29 (20 HIV-positive, 9 HIV-negative) gay men	↓ POMS anxiety ↑ relaxation	↑ CD56, NKCC ↑ cytotoxic cells no Δ CD4	ND

tx. = treatment; ab. = antibody; ctrl. = control; WLC = wait-list control; dep. = depression; anx. = anxiety; ND = not done; EE = existential experiential; sx. = symptoms; CD4 = T-helper cells; PHA = lymphocyte proliferation to PHA; CD56 = natural killer cells; EBV Ab. = Epstein-Barr Virus antibodies; HHV-6 Ab. = human herpes virus type 6 antibodies; HSV-2 Ab. = herpes simplex virus type 2 antibodies; BDI = Beck Depression Inventory; Δ = change in.

and denial, on the one hand, and substance use and possibly risky sexual behaviors on the other. One self-defeating "loop" that this model highlights follows: the infected person's awareness of the ineffectiveness of their coping modes and strategies may increase their sense of low self-efficacy, helplessness, and hopelessness, which in turn perpetuates social isolation, depressive symptoms, continued substance use, and other negative health behaviors.

The intervention that was developed addresses each of these domains by enhancing self-efficacy and perceived control, teaching adaptive coping strategies, and improving social support availability, satisfaction, and utilization (Antoni, Schneiderman & Ironson, in press). This program teaches participants a variety of active coping strategies (e.g., cognitive restructuring, relaxation exercises, assertive responses) and may enhance their psychological functioning by increasing perceptions of self-efficacy, personal control, and mastery (Fishman & Loscalzo, 1987); changing maladaptive cognitive distortions (Simons, Garfield, & Murphy, 1984); modifying stressor appraisals and providing the person with an available coping response (Turk, Holzman & Kerns, 1986); decreasing a sense of hopelessness (Rush, Beck, Kovacs, Weisenberger, & Hollon, 1982); and increasing the availability and utilization of social support networks (Vachon, Lyall, Rogers, Freedman, & Freedman, 1980).

This CBSM intervention program uses 10 weekly modules to address these aims as follows: (i) to increase personal awareness by providing information on sources of stress, the nature of human stress responses, and different coping strategies used to deal with stressors; (ii) to teach anxiety reduction skills such as progressive muscle relaxation and relaxing imagery; (iii) to modify maladaptive cognitive appraisals using cognitive restructuring; (iv) to enhance cognitive, behavioral, and interpersonal coping skills through coping skills training, assertion training, and anger management techniques; and (v) to provide a supportive group environment and increase utilization of social support networks. The effectiveness of this intervention was systematically evaluated in different cohorts of HIV-infected persons who were dealing with different landmark challenges known to occur during the infection and focused on specific "critical points" that these HIV-infected people were passing through. The rationale underlying this approach is that HIV infection comprises not only a "spectrum" of diseases, but also a spectrum of psychosocial challenges that change over time. At least four, somewhat artificial, critical points can be used to characterize these challenges: (i) responding to the initial diagnosis of seropositivity; (ii) adjustment to being infected during the early asymptomatic period when individuals are still healthy; (iii) adjustment to the experience of HIV-related symptoms that are not life-threatening but do impact the quality of life; and (iv) adjustment to a diagnosis of AIDS (Lutgendorf, Antoni, Schneiderman, & Fletcher 1994).

8.10.4.2.1 The initial HIV seropositive diagnosis

In the first set of studies, 65 gay men who did not know their HIV serostatus were randomly assigned to either a CBSM intervention, an aerobic exercise intervention, or a control group. After 5 weeks of participating in one of these conditions, blood was drawn for antibody testing and the men received news of their serostatus 72 hours later. The intervention continued for another 5 weeks (10 weeks total) and the men were followed through the initial "adjustment" period. Across the notification period the HIV-positive controls showed significant increases in anxiety and depression, whereas the HIV-positive CBSM subjects showed no significant changes in anxiety, or depression scores (Antoni, Baggett et al., 1991). The same buffering effect was present for the HIV-positive men assigned to the aerobic exercise group (LaPerriere et al., 1990). With respect to immune findings, the HIV-positive controls showed slight decrements in responsivity to PHA, NKCC, and NK cell counts pre- to postnotification (with no change in CD4 counts). In contrast, the HIV-positive CBSM subjects had significant increases in CD4 and NK cells and slight increases in PHA responsivity and NKCC, thus showing a "buffering" effect (Antoni, Baggett et al., 1991). Exercise had a similar buffering effect on several immune measures, though these are not reviewed in detail here (see LaPerriere et al. [1990] or LaPerriere et al. [1994]).

A 2-year follow-up of 23 HIV-positive men recruited from these studies was conducted by Ironson and colleagues (Ironson et al., 1994) to determine psychological predictors of disease progression to symptoms and death from AIDS. They found that distress at diagnosis, HIV-specific denial coping (5 weeks postdiagnosis minus prediagnosis), and low treatment adherence (poorer attendance at either CBSM or exercise groups, a lower frequency of relaxation practice during the 10 weeks for those in the CBSM, and not doing homework for those in CBSM group) all predicted faster disease progression. Increases in denial and better treatment adherence remained significant even after controlling for initial disease severity at study entry (CD4 number). Furthermore, decreases in denial and a greater frequency of relaxation home practice during the 10-week intervention period were predictive of higher CD4 cell counts and greater blastogenic responses to PHA at 1 year follow-up (Ironson et al., 1994). These findings suggest that those men who attended intervention sessions regularly, and dealt with their denial during the intervention period, were most likely to show longer term immune and health benefits.

There were also several consistent immune changes noted over the total 10 week period of intervention. The CBSM (and exercise) interventions were associated with a significant decrease in antibody titers to EBV-VCA and to HHV-6 as compared to assessment-only controls, whose antibody titers remained constant over the 10-week intervention period and elevated relative to healthy male laboratory controls (Esterling et al., 1992). An analysis of other psychosocial changes revealed that men in the control group showed significant decrements in social support during the postnotification period, whereas those in the CBSM group maintained their social support levels (Friedman et al., 1991). We also found that pre–postintervention decrements in social support were associated with greater declines in NKCC over a similar period (Antoni, Ironson et al., 1992) and that the reductions in EBV antibody titers during this period appeared to be mediated by increases in social support (Antoni et al., 1996).

8.10.4.2.2 The asymptomatic stage of HIV infection

Even asymptomatic HIV disease is accompanied by wide-ranging psychosocial challenges including the many uncertainties with regard to future health course and available resources. The former arises due to the unpredictable physical course of the HIV infection and the latter as a function of uncontrollable stressors such as impaired occupational and social functioning, decreased earning power, high costs of medical care, complex medical treatments, difficulties with self-care, multiple bereavements, and declining social supports (Blendon & Donelan, 1988; Ginzburg & Gostin, 1986; Redfield & Burke, 1988; Walkey et al., 1990).

The effects of CBSM intervention have been examined with gay men who knew of their HIV-positive status for at least 6 months, but were still asymptomatic. Although their physical functioning had not yet been affected by the virus, knowledge of their positive serostatus, and its eventual implications, had affected their lifestyle, behaviors, and sense of well-being. Elsewhere the types of issues that these men were reporting have been described in greater detail (Lutgendorf, Antoni, Schneiderman, Ironson et al., 1994). Some issues included reluctance to terminate relationships, hesitating to have sex with someone who was uninfected for fear of infecting them, deciding to disclose HIV status to family members, and making job and career transitions. The predominant emotional experiences reported by these men included anger, fear, loss, and uncertainty.

The 10-week CBSM intervention developed for these asymptomatic men was specifically adapted to address these issues (Antoni, Ironson et al., 1992). The early part of the intervention maintained its psychoeducational focus, with sessions on HIV and the stress response, cognitive appraisals, refuting, and replacing cognitive distortions. For this cohort, the group sessions discussing impending notification and postnotification concerns were replaced with a session dealing with unfinished issues related to seropositivity notification and notification of others regarding one's seropositivity. Behavior change, assertiveness, and social support were addressed extensively in these group sessions, especially strategies for eliciting social support. From a sample of 23 men, 11 were randomized to CBSM and 12 to an aerobic exercise program, each lasting 10 weeks. Because this was a pilot study there was no control condition. Participants in CBSM significantly reduced their anxiety, depression, anger, fatigue, and confusion levels, as well as increasing feelings of vigor (Antoni, Ironson et al., 1992). They also reported increased use of adaptive coping strategies, active coping, planning, and acceptance, and decreases in several maladaptive coping strategies such as mental disengagement and denial. These individuals also revealed significant increases in blastogenic responses to PHA and NKCC.

Other groups have also reported significant reductions in distress (Chesney et al., 1996a, 1996b; Kelly et al., 1993; Mulder et al., 1994) and improvements in self-efficacy (Chesney et al., 1996a, 1996b) with cognitive behavioral interventions in asymptomatic HIV-infected men. Kelly and colleagues randomized participants to either cognitive-behavioral group intervention, social support group intervention, or no-treatment, and found significant reductions in depression, hostility, and somatization in the cognitive-behavioral groups, as well as the social support group, compared to controls. The cognitive-behavioral group was somewhat more effective at reducing drug use (Kelly et al., 1993). Mulder and colleagues compared cognitive-behavioral group intervention and existential experiential group intervention in

HIV-infected gay men and found that both groups reduced distress and depressive symptoms compared with a wait-list control condition (Mulder et al., 1994). Finally, Chesney et al, compared the effects of a CBSM group intervention (Coping Effectiveness Training (CET)) (Chesney et al., 1994; Folkman et al., 1991) with an information group and a wait-list control group, and found that CET significantly increased self-efficacy and decreased perceived stress and burnout compared to the other two conditions (Chesney et al., 1996a, 1996b).

8.10.4.2.3 The pre-AIDS symptomatic stage of HIV infection

The effects of a group-based CBSM intervention have also been evaluated in HIV-positive gay men who are dealing with the symptoms of the infection. Forty HIV-infected gay men who had mild symptoms (category B of the 1993 CDC definition) were randomly assigned to either the CBSM intervention or a modified wait-list control group in which they complete a 10 week waiting period before being reassessed and provided a one day CBSM seminar. Across the 10 weeks the CBSM group showed a significant decrease in depression, anxiety, and antibody titers to HSV-2 (Lutgendorf, Antoni, Ironson, Klimas et al., 1997). Moreover, the 10 week decreases in HSV-2 antibody titers were strongly associated with the degree to which participants lowered their depression scores on the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) over this period. The control group showed no such changes in depressive symptoms or antibodies to HSV-2. Subsequently, conducted path analyses revealed that the effects of CBSM on HSV-2 antibody reductions appeared to be mediated by the intervention-associated depression changes (Antoni et al., 1996). Since this structured intervention was designed to increase cognitive and behavioral coping skills and support among group members, we next examined the relative contribution of intervention-related changes in coping skills and social support to reductions in dysphoria, anxiety, and distress-related symptoms over the study period. Those randomized to CBSM showed significant improvement in cognitive coping strategies, such as positive reframing and acceptance, as well as in total perceived social support provisions compared to a wait-list control condition (Lutgendorf et al., 1996). Cognitive coping changes, specifically acceptance of the HIV infection, were strongly related to lower dysphoria, anxiety, and total mood disturbance in this sample. Path analyses suggested that changes in social support, and in cognitive coping strategies, appear to mediate the effects of the CBSM intervention on the changes in distress noted during the intervention. These results suggest that cognitive coping strategies and social support factors can be modified by psychosocial interventions. These changes may in turn be important determinants of psychological well-being and quality of life during symptomatic HIV infection.

8.10.4.2.4 Full-blown AIDS

To date, the bulk of the published empirical studies evaluating the efficacy of psychosocial interventions with HIV-infected people have focused on gay men at the earlier stages of the infection. While there is a growing literature documenting those psychosocial factors that are predictive of both psychological adjustment and physical health course in AIDS patients (e.g., long-term survivor literature reviewed in Ironson et al. [1995]), little or no work exists testing the effects of various interventions with patients who have progressed to AIDS. What we do know is that the onset of AIDS may be accompanied by a lessening of psychological distress from the levels experienced by symptomatic (pre-AIDS) HIV-infected men (Dilley et al., 1985; Tross & Hirsch, 1988). This may reflect a relief from the unpredictability of not knowing if and when the first signs of AIDS may occur.

With the steady progression of HIV infection to more and more serious symptoms, and eventually AIDS, individuals go through a continual process of emotional crises, recycling through the stages that Kubler-Ross (1969) originally noted as comprising the dying process (Lutgendorf, Antoni, Schneiderman, & Fletcher, 1994). Moreover, HIV-infected patients may have emotional reactions that are more intense and labile than those described by Kubler-Ross (Hoffman, 1991). The high percentage of HIV-infected individuals diagnosed with adjustment disorders is a testimony to the difficulties encountered as these people try to come to terms with their illness and its impact on their lives (Hoffman, 1991). While it is clear that these individuals are in need of psychosocial services, one must be cautious in applying what has been learned from asymptomatic and early-stage populations to those with AIDS. There is a lack of empirical studies evaluating the effectiveness of psychosocial interventions such as CBSM or other approaches such as Expressive Supportive therapy conducted in a group or individual-based format with individuals at this stage of the infection. It may be that those with AIDS may find structured CBSM techniques useful, but may also require additional emotional support. For these individuals a therapeutic plan that combines a group-based stress reduction intervention with individual-based therapy directed toward these existential issues may be very useful. An optimal therapy plan may also involve pharmacologic treatment blended with the psychosocial intervention.

8.10.4.3 Common Features of Psychosocial Interventions in HIV Infection

8.10.4.3.1 Relaxation aspects

It is noteworthy that many of the psychosocial intervention studies with HIV-infected persons have utilized some form of relaxation-based technique as part of the intervention protocol. This is relevant given the great deal of distress and anxiety that these people experience. Relaxation strategies are also salient given the reasonably large literature base relating relaxation to improvement in immune system functioning (e.g., Ironson et al., 1995; Kiecolt-Glaser et al., 1985, 1986). With regard to HIV infection, one of these studies found that frequency of home relaxation practice was related to better immunological functioning and slower disease progression in HIV-infected men (Ironson et al., 1994).

Another study examined alternate ways of inducing relaxation in HIV-infected individuals (Ironson et al., 1996)—massage therapy—because of its expected relaxing effects, and because it is easy to monitor objectively how much massage someone is getting. HIV-infected gay men showed a significant increase in NK cell number, NKCC, the soluble CD8 receptor, and the cytotoxic subset of CD8 cells during the weeks of daily massage intervention compared to the same men's values during a no-massage period. These men also showed significant decreases in 24 hour urinary cortisol output during the intervention period, suggesting that distress decreases and/or relaxation increases may have mediated these immunologic changes. Indeed, significant decreases in anxiety and increases in a sense of relaxation were also observed, which were both correlated with increases in NK cell number. This study suggests that the relaxation components of CBSM interventions noted previously may contribute directly to both the psychological and immunologic changes experienced by the participants. Psychosocial interventions used with HIV-infected persons use many other techniques in addition to relaxation and also vary considerably in terms of format and theoretical orientation. It is possible to examine some of the parameters on which these approaches converge and diverge.

8.10.4.3.2 Delivery format

Regarding delivery format, it is important to note that most of the psychosocial interventions evaluated with HIV-infected persons (including those involving aerobic exercise training [LaPerriere et al., 1990]) have been conducted in a group format (Ironson et al., 1995). These CBSM interventions typically employed groups of approximately 6–8 participants facilitated by two group leaders. The sessions were conducted in comfortable rooms, ran about 2–2.5 hours in duration, and ran on a weekly basis over a 10-week period. While there was some variability in terms of the duration of programs evaluated across different research groups (range = 8–12 weeks), the number of participants, group leaders, and session length and frequency were quite similar. Because of the fact that these sessions were conducted in a group format, little is known about the effects of these sorts of techniques with HIV-infected persons when delivered as individual-based psychotherapy sessions or self-help approaches.

8.10.4.3.3 Treatment orientation

Different treatment orientations to psychosocial interventions have been compared in HIV-infected persons by one group in the Netherlands (in collaboration with the University of Miami). They examined the influence of treatment orientation by comparing the psychological (Mulder et al., 1994) and immunologic (Mulder, Antoni, Emmelkamp et al., 1995) effects of one group-based intervention using a cognitive-behavioral orientation vs. one using an existential/experiential orientation in HIV-infected gay men. These men were recruited from a large cohort study conducted through the Amsterdam Municipal Health Center. The cognitive-behavioral therapy (CBT) intervention included training in cognitive restructuring, relaxation exercises, behavior change, assertiveness skills, coping fit for controllable vs. uncontrollable stressors, and information on stress responses. As such, it was quite similar to the CBSM interventions described previously. The experiential therapy (ET) condition was less structured and focused on increasing awareness of the here and now, of their inner experiential process, and restoring congruence between emotional, cognitive, and behavioral schemata. The intervention also attempted to help participants deal with a shortened life perspective and anxiety about future illness and death. Despite the differences in theoretical orientation the authors note that both interventions were designed to reduce stress, improve coping, build social support, and encourage emotional expression. They found that both (CBT and ET) conditions were associated with significant decreases in POMS total mood disturbance, and BDI scores compared to the controls, though there were no changes in coping style, social support, or emotional expression over the intervention period.

Mulder and colleagues compared the 26 men randomized ultimately to either CBT or ET with men from the Amsterdam cohort study who did not participate in the intervention study for differences in slope of immune changes that had occurred over the period after the intervention (Mulder, Antoni, Emmelkamp et al., 1995). Those intervention group participants with the largest decreases in distress, between the beginning of either intervention and 9 months later, showed the least decline in CD4 cell counts over a 2-year follow-up period from the beginning of the intervention. These findings suggest that while theoretical orientation may not determine the psychological improvements gained by HIV-infected men participating in group-based psychosocial interventions, the size of treatment-related reductions in distress may be predictive of longer-term health benefits. This interpretation seems to be in line with the findings reviewed previously (Antoni, Baggett et al. [1991] for short-term effects and Ironson et al. [1994] for longer-term effects).

8.10.4.4 Summary of Completed Studies of Psychosocial Interventions with HIV-infected Persons

Here we summarized the qualities characterizing the psychosocial intervention studies with HIV-infected individuals that were completed between 1989 and 1997. These studies recruited moderate sized samples that were approximately equally distributed across experimental conditions usually resulting in 10–30 people actually receiving the intervention in any single trial. While individuals participating in these studies were all gay or bisexual men, congruent with the predominant AIDS incidence patterns occurring in the 1980s, it should be noted that women now make up nearly 15% of all AIDS cases and current incidence patterns for HIV infection suggest that this proportion will increase. Therefore, these studies need to be replicated with other HIV-infected populations such as indigent women.

Of the interventions tested to date, most were conducted in groups, used trained facilitators or therapists, and met at least once weekly over periods of 8–12 weeks. Since the studies evaluating these interventions were designed to test their immediate efficacy, cohorts were typically assessed at single time points at the beginning and end of the intervention period. At least two studies followed subjects over a 2-year postintervention period (Ironson et al., 1994; Mulder et al., 1995).

While the results of this set of studies are not entirely consistent, they suggest the potential efficacy of psychosocial interventions for increasing psychological adjustment, immune functioning, and health status in HIV-infected people at early and middle stages of the disease who are dealing with mounting stressors. Less is known about the effects of these interventions with people at the more advanced stages of the disease. One study focused specifically on how a time-limited CBSM intervention could moderate the acute impact of learning one's HIV serostatus for the first time (Antoni, Baggett et al., 1991), while the remainder have examined how psychosocial interventions can improve psychological functioning in people dealing with the more chronic stress of having had HIV for a number of years (Auerbach et al., 1992; Coates, McKusick et al., 1989; Ironson et al., 1995; Mulder et al., 1994). The most consistent trends among the psychological effects of these interventions are: a decrease in distress and depressed affect; a decrease in the use of avoidance and denial as coping strategies; an increase in the use of acceptance and positive reframing strategies; and an increase or maintenance of social support provisions.

While CBSM intervention has been shown to reduce antibodies to herpes viruses and to modulate NKCC, no direct effects have been observed on HIV disease progression markers. Three group-based stress management interventions reported no significant impact on CD4 cell counts (Auerbach et al., 1992; Coates, McKusick et al., 1989; Mulder, Antoni, Emmelkamp et al., 1995), but none of these measured herpes virus antibodies or NKCC, which were the measures found to be most responsive to change in some of the other studies (Esterling et al., 1992; Ironson et al., 1996; Lutgendorf, Antoni, Schneiderman, & Fletcher, 1994). Ironson et al. (1995) note that no studies have tested the long-term effects of these interventions on 5- or 10-year survival, although one study with a small sample showed that intervention-related variables (denial coping reductions and greater treatment adherence) were predictive of 2-year survival (Ironson et al., 1994). Another study (Mulder, Antoni, Emmelkamp et al., 1995) showed distress changes (which can often be modified by psychosocial interventions) were related to changes in CD4 counts over a similar 2-year period. However, it should be noted that "assignment to a psychosocial intervention" has not been causally related to a slower rate of disease progression or greater survival time in HIV-infected persons. This lack of a causal effect may be due to the limited statistical power offered by the small samples that have been studied to date. Clinical trials that utilize hundreds of participants may be necessary to capture the effects of psychosocial interventions on the health course of this disease. It is also possible that measures of immune function might be more sensitive to the effects of stress-reducing psychosocial interventions than lymphocyte counts because stress hormone interactions with receptors that mediate cell activity and certain immune functional indices may be especially relevant to the pathophysiology of HIV infection in particular. Finally, it has been suggested that health psychology interventions possessing cognitive and behavioral elements, similar to those just noted, may be quite useful in addressing other aspects of HIV infection, such as nutritional disorders, pain, sleep problems, and medication adherence, all having a potentially huge impact on the quality of life of infected individuals (Sikkema & Kelly, 1996). It will be important to incorporate these additional variables into intervention protocols conducted in the future as more and more people live longer and longer with this disease.

8.10.5 CONCLUSION

This chapter has summarized the rationale and empirical support for the role of health psychology in the context of HIV infection and AIDS. Based on this work it is plausible to suggest that psychosocial targets of interventions should include reductions in distress and depression, as well as avoidance, denial, and rumination as methods of dealing with the chronic stress of the infection. Equally important intervention targets would involve increasing active coping and a sense of social attachment as important stress buffers. Several studies are underway which further examine the effects of psychosocial interventions on psychological adjustment, immune status, and disease course in HIV-infected individuals. There are several reasons to be encouraged by the trends that are apparent across ongoing investigations of psychosocial interventions in HIV-infected populations.

First, all of the studies use standardized time-limited interventions (10–17 weeks) that have been "manualized" with specific treatment modules and criteria for therapist training and implementation (e.g., Antoni, Schneiderman, & Ironson, in press). In some cases these interventions follow a CBSM orientation and in others a more experiential-existential approach is utilized while, in others, CBSM and experiential techniques have been integrated into the same intervention condition. The efficacy of these different treatment orientations may ultimately be shown to vary systematically as a function of the disease stage and demographic characteristics of the targeted population.

Second, these investigations are each recruiting people who meet specific inclusion and exclusion criteria that are similar across studies. Specifically, the samples are largely gay men at the early to middle stages of the infection (pre-AIDS symptoms and a restricted range of CD4 counts) who are dealing with similar psychosocial issues. Shifting demographic patterns of the HIV/AIDS epidemic make it imperative to develop techniques to help other HIV-infected populations, such as women of color, manage their disease. One study, led by Antoni, Greenwood, and Schneiderman in Miami, is examining the effects of a 10-week CBSM intervention in HIV-infected African-American women who are developing the first symptoms of the infection. Another trial is evaluating the effects of a similar intervention in women who have already developed AIDS. Clearly, a next step is to develop psychosocial interventions to address secondary prevention issues in the growing number of HIV-infected injection drug users.

Third, each study is assessing the psychosocial effects of these interventions with a collection of well-validated instruments designed to track changes in distress and depression as well as intermediary psychosocial variables such as coping, social support, and stressor appraisals. In most cases, the most reliable instruments have been employed in these batteries supplemented with more specific measures designed to tap the most salient areas of participants" lives.

Fourth, most of these studies currently being done reflect an excellent breadth of immunologic measures (including qualitative and quantitative indices). For instance, several studies are now assessing the effects of psychosocial interventions on NKCC and blastogenic responses to PHA. Several studies are also examining changes in IgG antibody titers to a panel of relevant herpes viruses (e.g., EBV, HHV-6, CMV, HSV) and assessing serum neopterin as a subclinical disease progression marker. In the future it is likely that assessments of HIV viral load will join this list of assays as a way to monitor treatment-related changes in disease status. It is also imperative that more disease- and antigen-specific immunologic measures be incorporated into assessment protocols to better understand specific pathophysiologic links between psychological changes and health course.

Fifth, all of the studies are tracking the effects of their interventions on clinical health changes using standardized staging procedures and criteria (Centers for Disease Control, 1993). Due to the relatively large sample sizes, these studies should be able to assess long-term effects on survival. All of these studies are also monitoring changes in several other behavioral measures such as sexual behaviors, substance use, physical activities, sleep, and medications. Information about changes in these additional measures, over the course of the interventions, may be useful for clarifying the mechanisms underlying treatment changes in affective, immunologic, and health endpoints, as well as representing important quality of life changes that may be likely to occur following participation in psychosocial interventions. Once completed these studies may provide efficacy data supporting the initiation of larger clinical trials as well as informing behavioral scientists about the psychosocial mechanisms underlying psychological adjustment, physiological functioning, and health preservation in this chronic disease.

Sixth, the arrival of protease inhibitors and triple combination therapies to the treatment arena has brought about important improvements in the health care of AIDS patients (Lewin, 1996; Schneiderman et al., 1997). The resulting treatment philosophy for HIV and AIDS increasingly views HIV infection as a chronic disease in which patient management is critical. It is quite reasonable to propose that psychosocial interventions can be an essential part of this management (Schneiderman et al., 1997). The key here is consistent adherence to a demanding medication schedule in the context of an already stressful daily existence. We now know, for example, that once patients are given protease inhibitors, medication protocols must be rigidly followed with respect to drug schedules and food intake or the virus may mutate and become drug resistant (Lewin, 1996). Failure to adhere to specific guidelines cannot only compromise the effects of the particular protease inhibitor being used, but can also reduce the efficacy of other related compounds to which cross-resistance has developed (Lewin, 1996). Psychosocial interventions that provide information, skills, and support to patients can facilitate adherence to difficult medication protocols. Since interventions such as CBSM can decrease denial and depressed affect, they may, in turn, decrease negative health behaviors and facilitate adherence to medical treatment regimens. It has been shown that better adherence to a stress management protocol is predictive of a slower rate of disease progression (Ironson et al., 1994). The improvements in immune function that have been associated with these psychosocial interventions raise the interesting question of whether CBSM can help reconstitute the compromised immune system once pharmacologic agents such as protease inhibitors have contained the virus. These latter questions constitute the focus of the ongoing work using health psychology strategies to address contemporary secondary prevention issues in the growing populations of HIV-infected persons who are attempting to manage their disease.

8.10.6 REFERENCES

Aboulker, J. P., & Swart, A. M. (1993). Preliminary analysis of the Concorde trial. *Lancet, 341,* 889–890.

Amsterdam, J. D., Lucki, I., & Winokur, A. (1985). The ACTH stimulation test in depression. *Psychiatric Medicine, 3,* 91–100.

Amsterdam, J. D., Maislin, G., Gold, P., & Winokur, A. (1989). The assessment of abnormalities in hormonal responsiveness at multiple levels of the hypothalamic–pituitary–adrenocortical axis in depressive illness. *Psychoneuroendocrinology, 14,* 43–62.

Antoni, M. H. (1991). Psychosocial stressors and stress management with HIV-1 seropositive and seronegative gay men, *International Reviews in Psychiatry, 3,* 385–402.

Antoni, M. H., August, S., LaPerriere, A., Baggett, H. L., Klimas, N., Ironson, G., Schneiderman, N., & Fletcher, M. A. (1990). Psychological and neuroendocrine measures related to functional immune changes in anticipation of HIV-1 serostatus notification. *Psychosomatic Medicine, 52,* 496–510.

Antoni, M. H., Baggett, L., Ironson, G., August, S., LaPerriere, A., Klimas, N., Schneiderman, N., & Fletcher, M. A. (1991). Cognitive behavioral stress management intervention buffers, distress responses, and immunologic changes following notification of HIV-1 seropositivity. *Journal of Consulting and Clinical Psychology, 59,* 906–915.

Antoni, M. H., Esterling, B., Lutgendorf, S., Fletcher, M. A., & Schneiderman, N. (1995). Psychosocial stressors, herpes virus reactivation and HIV-1 infection. In M. Stein & A. Baum (Eds.), *Chronic disease: Perspectives in behavioral medicine.* Hillsdale, NJ: Erlbaum.

Antoni, M. H., Goldstein, D., Ironson, G., LaPerriere, A., Fletcher, M. A., & Schneiderman, N. (1995a). Coping responses to HIV-1 serostatus notification predict concurrent and prospective immunologic status. *Clinical Psychology and Psychotherapy, 2*(4), 234–248.

Antoni, M. H., & Goodkin, K. (1996). Cervical neoplasia, human papilloma virus, and psychoneuroimmunology. In H. Friedman et al. (Eds.), *Psychoneuroimmunology and infectious disease* (pp. 243–262). Boca Raton, FL: CRC Press.

Antoni, M. H., Goodkin, K., Goldstein, V., LaPerriere, A., Ironson, G., Fletcher, M. & Schneiderman, N. (1991). Coping responses to HIV-1 serostatus notification predict short-term affective distress and one-year immunologic status in HIV-1 seronegative and seropositive gay men. *Psychosomatic Medicine, 53,* 227.

Antoni, M. H., Ironson, G., Helder, L., Lutgendorf, S., Friedman, A., LaPerriere, A., Fletcher, M. A. & Schneiderman, N. (1992, April). *Stress management intervention reduces social isolation and maladaptive coping behaviors in gay men adjusting to an HIV-1 seropositive diagnosis.* Paper presented at the scientific meeting of the Society of Behavioral Medicine, New York.

Antoni, M. H., Lutgendorf, S., Ironson, G., Fletcher, M. A., & Schneiderman, N. (1996). CBSM intervention effects on social support, coping, depression, and immune function in symptomatic HIV-infected men. *Psychosomatic Medicine, 58,* 86.

Antoni, M. H., Rodriguez, M., Starr, K., et al. (1991, August). *Neuroendocrine and immunologic reactivity to behavioral challenge in HIV-infected gay men.* Paper presented at the American Psychological Association, San Francisco, CA.

Antoni, M. H., Schneiderman, N., Fletcher, M. A., Goldstein, D., Ironson, G., & LaPerriere, A. (1990) Psychoneuroimmunology and HIV-1. *Journal of Consulting and Clinical Psychology, 58,* 38–49.

Antoni, M. H., Schneiderman, N., & Ironson, G. (in press). *Stress management for HIV-infection.* Society of Behavioral Medicine Clinical Research Guidebook Series.

Antoni, M. H., Schneiderman, N., LaPerriere, A., O'Sullivan, M. J., Marks, J., Efantis, J, Skyler, J., & Fletcher, M. A. (1992). Mothers with AIDS. In P. Ahmed (Ed.), *Living and dying with AIDS.* NJ: Plenum.

Arnetz, B. B., Wasserman, J., Petrini, B., Brenner, S. O., Levi, L., Eneroth, P., Salovaara, H. K., Hjelm, R., Salovaara, L., Theorell, T., et al. (1987). Immune function in unemployed women. *Psychosomatic Medicine, 49,* 3–12.

Atkinson, J., Sullivan, T., Kelly, J., et al. (1977). Stimulation by alcohols and cyclic AMP metabolism in human leukocytes. *Journal of Clinical Investigation, 6,* 284.

Auerbach, J. E., Oleson, T. D., & Solomon, G. F. (1992). A behavioral medicine intervention as an adjunctive treatment for HIV-related illness. *Psychology and Health, 6,* 325–334.

Axelrod, J., & Reisine, T. (1984). Stress hormones: their interaction and regulation. *Science, 224,* 452–459.

Bancroft, G., Shellam, G., & Chalmer, J. (1981). Genetic influences on the augmentation of natural killer cells (NK) during murine cytomegalovirus infection: correlation with patterns of resistance. *Journal of Immunology, 126,* 988–994.

Baron, R. S., Cutrona, C. E., Hicklin, D., Russell, D. W., & Lubaroff, D. M. (1990). Social support and immune function among spouses of cancer patients. *Journal of Personality and Social Psychology, 59,* 344–352.

Bartrop, R., Lazarus, L., Luckhurst, E., Kiloh, L., & Penny, R. (1977). Depressed lymphocyte function after bereavement. *Lancet, i,* 834–836.

Baum, M. K., Mantero-Atienza, E., Shor-Posner, G., Fletcher, M. A., Morgan, R., Eisdorfer, C., Sauberlich, H. E., Cornwell, P. E., & Beach, R. S. (1991). Association of vitamin B_6 status with parameters of immune function in early HIV-1 infection. *Journal of Acquired Immune Deficiency Syndrome, 4,* 1122–1132.

Beck, A. I., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561–571.

Beckett, A., & Rutan, J. S. (1990). Treating persons with ARC and AIDS in group psychotherapy. *International Journal of Group Psychotherapy, 40,* 19–29.

Beckman, L. (1980). Perceived antecedents and effects of alcohol consumption in women. *Journal of Studies of Alcohol, 41,* 518.

Benschop, R. J., Nieuwenhuis, E. E., Tromp, E. A., Godaert, G. L., Ballieux, R. E., & van Doornen, L. J. (1994). Effects of b-adrenergic blockade on immunologic and cardiovascular changes induced by mental stress. *Circulation, 89,* 762–769.

Berkman, L., & Syme, S. (1979). Social networks, host resistance, and mortality: A nine-year follow-up study of Alameda County residents. *American Journal of Epidemiology, 109,* 186–204.

Biegalke, B. J., & Geballe, A. P. (1991). Sequence requirements of reactivation of the HIV-1 LTR by human cytomegalovirus. *Virology, 183,* 381–385.

Billings, A., & Moos, R. (1981). The role of coping responses and social resources in attenuating the stress of life events. *Journal of Behavioral Medicine, 4,* 139–157.

Biron, C. A., Byron, K., & Sullivan, J. (1989). Severe herpes virus infections in an adolescent without natural killer cells. *New England Journal of Medicine, 320,* 1731–1735.

Blaney, N. (1985). Smoking: Psychophysiological causes and treatment. In N. Schneiderman & J. Tapp (Eds.), *Behavioral medicine: The biopsychosocial approach* (pp. 347–377). Hillsdale, NJ: Erlbaum.

Blendon, R. J., & Donelan, K. (1988). Discrimination against people with AIDS: The public's perspective. *New England Journal of Medicine, 319,* 1022–1026.

Blomkvist, V., Theorell, T., Jonsson, H., et al. (1994). Psychosocial self-prognosis in relation to mortality and morbidity in hemophiliacs with HIV infection. *Psychother Psychosom, 62,* 185–192.

Bloom, D. E. (1982). Social support, accommodation to stress and adjustment to breast cancer. *Social Science and Medicine, 16,* 1329–1338.

Bloom, D. E. & Carliner, G. (1988). The economic impact of AIDS in the United States. *Science, 239,* 604–610.

Blumberg, R., Kelsey, P., Perrone, T., et al. (1984). Cytomegalovirus-associated and cryptosporidium-associated acalculous gangrenous cholecystis. *American Journal of Medicine, 76,* 1118–1123.

Bonelli, J. (1982). Stress, catecholamines, and beta-blockade. *Acta Medical Scandinavia, 660,* 214–218.

Borysenko, M., & Borysenko, J. (1982). Stress, behavior, and immunity: Animal models and mediating mechanisms. *General Hospital Psychiatry, 4,* 59–67.

Buckingham, S., & Rehm, S. (1987, Winter). AIDS and women at risk. *Health and Social Work, 12,* 5–11.

Bukowski, J., Warner, J., Dennert, G. & Welsh, R. (1985). Adoptive transfer studies demonstrating the antiviral effect of natural killer cells in vivo. *Journal of Experimental Medicine, 161,* 40–52.

Burack, J. H., Barrett, D. C., Stall, R. D., Chesney, M. A., Ekstrand, M. L., & Coates, T. J. (1993). Depressive symptoms and CD4 lymphocyte decline among HIV-infected men. *Journal of the American Medical Association, 270,* 2568–2573.

Burns, D. N., Kramer, A., Yellin, F., Fuchs, D., Wachter, H., DiGioia, R. A., Sanchez, W. C., Grossman, R. J., Gordin, F. M., Biggar, R. J., & Goedert, J. J. (1991). Cigarette smoking: A modifier of Human Immunodeficiency Virus Type 1 infection? *Journal of Acquired Immune Deficiency Syndromes, 4,* 76–83.

Calabrese, J., Kling, M., & Gold, P. (1987). Alterations in immunocompetence during stress, bereavement, and depression: Focus on neuroendocrine regulation. *American Journal of Psychiatry, 14,* 1123–1134.

Carbonari, M., Fiorilli, M., Mezzaroma, I., Cherchi, M., & Aiuti, F. (1989). CD4 as the receptor for retroviruses of the HTLV family: Immunopathogenetic implications. *Advances in Experimental Medicine and Biology, 257,* 3–7.

Carpenter, C., Mayer, K., Fisher, A., Desai, M., & Durand, L. (1989). Natural history of acquired immunodeficiency syndrome in women in Rhode Island. *American Journal of Medicine, 86,* 771–775.

Catania, J., Coates, T., Stall, R., Bye, L., Kegeles, S., Capell, F., Henne, J., McKusick, L., Morin, S., Turner, H., & Pollack, L. (1991). Changes in condom use among homosexual men in San Francisco. *Health Psychology, 10,* 190–199.

Centers for Disease Control (1993). Revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *Morbidity and Mortality Weekly Report, RR17,* 1–19.

Centers for Disease Control (1995a). CDC: Update: AIDS among women—United States 1994. *Journal of the American Medical Association, 273*(10), 767.

Centers for Disease Control (1995b). *HIV/AIDS Surveillance Report, 7*(2).

Chesney, M. A., & Folkman, S. (1994). Psychological impact of HIV disease and implications for intervention. *Psychiatric Clinics of North America, 17,* 163–182.

Chesney, M. A., Folkman, S., & Chambers, D. (1996a). Coping effectiveness training for men living with HIV: Preliminary findings. *International Journal of STDs and AIDS, 7*(Suppl. 2), 75–82.

Chesney, M. A., Folkman, S., & Chambers, D. (1996b). The impact of a cognitive-behavioral intervention on coping with HIV disease. *Psychosomatic Medicine, 58,* 86.

Chiapelli, F., Ben-Eliyahu, S., Hanson, L., & Wylie, T. (1992). Modulation of immune surveillance of HIV-related cancers by alcohol and cocaine metabolites. *Alcologia, 4,* 219–227.

Christ, G., & Weiner, L. (1985). Psychosocial issues for AIDS. In V. Devita, Jr., S. Hellman, & S. Rosenberg (Eds.), *AIDS* (pp. 275–297). Philadelphia: Lippincott.

Chuang, H. T., Devins, G. M., Hunsley, J., & Gill, M. J. (1989). Psychosocial distress and well-being among gay and bisexual men with human immunodeficiency virus infection. *American Journal of Psychiatry, 146,* 876–880.

Chung, J., & Magraw, M. (1992). A group approach to psychosocial issues faced by HIV-positive women. *Hospital and Community Psychology, 43,* 891–894.

Cleary, P., Singer, E., Rogers, T., Avorn, J., VanDevanter, N., Soumerai, S., Perry, S., & Pindyck, J. (1988). Sociodemographic and behavioral characteristics of HIV antibody-positive blood donors. *American Journal of Public Health, 78,* 953–857.

Clerici, M., Berzofsky, J. A., Shearer, G. M., & Tacket, C. O. (1991). Exposure to human immunodeficiency virus (HIV) type 1 indicated by HIV-specific T-helper cell responses before detection of infection by polymerase chain reaction and serum antibodies. *Journal of Infectious Diseases, 164,* 178–184.

Coates, T. (August, 1991). *Who has been tested for HIV antibodies in the United States? The National AIDS Behavioral Survey (NABS).* Paper presented at the American Psychological Association, San Francisco, CA.

Coates, T., McKusick, L., Stites, D., & Kuno, R. (1989). Stress reduction training changed number of sexual partners but not immune function in men infected with HIV. *American Journal of Public Health, 79,* 885–887.

Coates, T., Stall, R., Ekstrand, M., & Solomon, G. (June, 1989). *Psychological predictors as cofactors for disease progression in men infected with HIV: The San Francisco men's health study.* Presented at the Vth International AIDS Conference, Montreal, Quebec, Canada.

Cobb, S. (1974). Physiological changes in men whose jobs are abolished. *Journal of Psychosomatic Research, 18,* 245–258.

Cohen, E. D. (1990). Confidentiality, counseling, and clients who have AIDS: Ethical foundations of a model rule. *Journal of Counseling and Development, 68,* 282–286.

Cohen, S., & Rodriquez, M. (1995). Pathways linking affective disturbances and physical disorders. *Health Psychology, 14,* 374–380.

Cohen, S., Tyrrell, D. A, & Smith, A. P. (1991). Psychological stress in humans and susceptibility to the common cold. *New England Journal of Medicine, 325,* 606–612.

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin, 98*(2), 310–357.

Cole, S., Kemeny, M., Taylor, S., Visscher, B., & Fahey, J. (1996). Accelerated course of HIV infection in gay men who conceal their homosexuality. *Psychosomatic Medicine, 58,* 219–231.

Cupps, T., & Fauci, A. (1982). Corticosteroid-mediated immunoregulation in man. *Immunology Review, 65,* 133–155.

Cutrona, C., & Russell, D. (1987). The provisions of social relationships and adaptation to stress. In Jones & Perlman (Eds.), *Advances in personal relationships.* Greenwich, CT: JAI Press.

Daar, E. S., Moudgil, T., Meyer, R. D., & Ho, D. D. (1991). Transient high levels of viremia in patients with primary human immunodeficiency virus type 1 infection. *New England Journal of Medicine, 324,* 961–964.

Dahan, P., Haettich, B., Le Parc, J., & Paolaggi, J. (1988). Meningoradiculitis due to herpes simplex virus disclosing HIV infection [letter]. *Annals of Rheumatic Disease, 47,* 440.

Davidson, L. M., & Baum, A. (1986). Chronic stress and posttraumatic stress disorders. *Journal of Consulting and Clinical Psychology, 54,* 303–308.

Davidson, R. (1985). Behavioral medicine and alcoholism. In N. Schneiderman & J. Tapp (Eds.), *Behavioral medicine: The biopsychosocial approach.* Hillsdale, NJ: Erlbaum.

Detels, R., English, P. A., Giorgi, J. V., Visscher, B. R., Fahey, J. L., Taylor, J. M. G., Dudley, J. P., Nishanian, P., Munoz, A., Phair, J. P., Polk, B. F., & Rinaldo, C. R. (1988). Patterns of CD4+ cell changes after HIV-1 infection indicate the existance of a co-determinant of AIDS. *Journal of the Acquired Immune Deficiency Syndrome, 1,* 390–395.

Dew, M., Ragni, M., & Nimorwicz, P. (1990). Infection with human immunodeficiency virus and vulnerability to psychiatric distress: A study of men with hemophilia. *Archives of General Psychiatry, 47,* 737–744.

Dilley, J. W., Ochitil, H. N., Perl, M., & Volberding, P. A. (1985). Findings in psychiatric consultations with patients with acquired immune deficiency syndrome. *American Journal of Psychiatry, 142,* 82–86.

DiMatteo, M. R., & Hays, R. (1981). Social support and serious illness. In B. H. Gottlieb (Ed.), *Social networks and social support* (pp. 117–148). Beverly Hills, CA: Sage.

Dimsdale, J., Hartley, L., Ruskin, J., Greenblatt, D., & Labrie, R. (1984). Effect of beta blockade on plasma catecholamine levels during psychological and exercise stress. *American Journal of Cardiology, 54,* 182–185.

Doll, D. C., & List, A. F. (1982). Burkitt's lymphoma in a homosexual. *Lancet, 1,* 1026–1027.

Donlou, J. N., Wolcott, D. L., Gottlieb, M. S., & Landsverk, J. (1985). Psychosocial aspects of AIDS and AIDS-related complex: A pilot study. *Journal of Psychosocial Oncology, 3,* 39–55.

Eickhoff, T. (1994). Doctors urged to offer more HIV tests. *American Medical News, 37*(9), 15.

Ellerman, D. (1989). *Quality of life of persons with AIDS/ARC.* Proceedings of the Vth International Conference on AIDS, Montreal, Quebec, Canada.

Ernberg, I. (1986). The role of Epstein-Barr virus in lymphomas of homosexual males. In E. Klein (Ed.), *Acquired immunodeficiency syndrome* (pp. 301–318). Basel, Switzerland: Karger.

Esterling, B., Antoni, M., Schneiderman, N., Ironson, G., LaPerriere, A., Klimas, N., Fletcher, M. A., & Carver, C. S. (1992). Psychosocial modulation of antibody to Epstein-Barr viral capsid antigen and herpes virus type-6 in HIV-1 infected and at-risk gay men. *Psychosomatic Medicine, 54,* 354–371.

Esterling, B., Kiecolt-Glaser, J., & Glaser, R. (1996). Psychosocial modulation of cytokine-induced natural killer cell activity in older adults. *Psychosomatic Medicine, 58,* 264–272.

Evans, D., Pettito, J., & Lesserman, J. (1992). Stress, depression and natural killer cells: Potential clinical relevance. *Clinical Neuropharmacology, 15*(Suppl. 1A), 656A–657A.

Eversole, L., Stone, C., & Beckman, A. (1988). Detection of EBV and HPV DNA sequences in oral "hairy" leukoplakia by *in situ* hybridization. *Journal of Medical Virology, 26,* 271.

Fawzy, F., Cousins, N., Fawzy, N., Kemeny, M., Elashoff, R., & Morton, D. (1990). A structured psychiatric intervention for cancer patients I. Changes over time in methods of coping and affective disturbance. *Archives of General Psychiatry, 47,* 720–725.

Felton, B., Revenson, T., & Hinrichsen, G. (1984). Stress and coping in the explanation of psychological adjustment among chronically ill adults. *Social Science and Medicine, 18,* 889–898.

Ferrara, A. J. (1984). My personal experience with AIDS. *American Psychologist, 39,* 1285–1287.

Fifield, L. (1975). *On my way to nowhere: An analysis of gay alcohol abuse and an evaluation of alcoholism rehabilitation services.* Los Angeles Gay Community Services Center and Department of Health Services.

Fischl, M. A., Richman, D., Grieco, M., Gottlieb, M., Volberding, P., Laskin, O., Leedom, J., Groopman, J., Moldvan, D., Schooley, R., Jackson, G., Durack, D., & King, D. (1987). The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. *New England Journal of Medicine, 317,* 185–191.

Fishman, B., & Loscalzo, M. (1987). Cognitive-behavioral interventions in management of cancer pain: Principles and application. *Medical Clinics of North America, 71*(2), 271–287.

Folkman, S., Chesney, M., McKusick, L., Ironson, G., Johnson, D., & Coates, T. J. (1991). Translating coping theory into intervention. In J. Eckenrode (Ed.), *The social context of coping* (pp. 239–260). New York: Plenum.

Foy, D., Sipprelle, P., Rueger, D., Carroll, E. M. et al. (1984). Etiology of post-traumatic stress disorder in Vietnam veterans: Analysis of premilitary, military, and combat exposure influences. *Journal of Consulting and Clinical Psychology, 52,* 79–87.

Freudenberg, N., Lee, J., & Silver, D. (1989). How Black and Latino community organizations respond to the AIDS epidemic: A case study in one New York city neighborhood. *AIDS Education and Prevention, 1* (1), 12–21.

Friedland, G. H., Saltzman, B., Vileno, J., Freeman, K., Schrager, L. K., & Klein, R. S. (1991). Survival differences in patients with AIDS. *Journal of Acquired Immune Deficiency Syndromes*, 144–153.

Friedman, A., Antoni, M. H., Ironson, G., LaPerriere, A., Schneiderman, N., & Fletcher, M. A. (1991). Behavioral interventions, changes in perceived social support and depression following notification of HIV-1 seropositivity. Presented at Society of Behavioral Medicine Annual Meeting, Washington, DC.

Gambe, R., & Getzel, G. S. (1989). Group work with gay men with AIDS. *Social Casework, 70,* 172–179.

Ginzburg, H. M., & Gostin, L. (1986). Legal and ethical issues associated with HTLV-III diseases. *Psychiatric Annals, 16,* 180–185.

Glaser, R., & Kiecolt-Glaser, J. (1987). Stress-associated depression in cellular immunity: implications for acquired immune deficiency syndrome (AIDS). *Brain, Behavior and Immunity 1,* 107–112.

Glaser, R., Kiecolt-Glaser, J. K., Bonneau, R., et al. (1992). Stress-induced modulation of the immune response to recombinant hepatitis B vaccine. *Psychosomatic Medicine, 54,* 22–29.

Glaser, R., Kiecolt-Glaser, J. K., Speicher, C. E., & Holliday, J. E. (1985). Stress, loneliness, and changes in herpesvirus latency. *Journal of Behavioral Medicine, 8,* 249–260.

Glaser, R., Kennedy, S., Lafuse, W. P., Bonneau, R. H., Speicher, C. E., Kiecolt-Glaser, J. K., & Hillhouse, J. (1990). Psychological stress-induced modulation of IL-2 receptor gene expression and IL-2 production in peripheral blood leukocytes. *Archives of General Psychiatry, 47,* 707–712.

Glaser, R., Pearson, G. P., Jones, J. F., Hillhouse, J., Kennedy, S., Mao, H., & Kiecolt-Glaser, J. K. (1991). Stress-related activation of Epstein-Barr virus. *Brain, Behavior and Immunity, 5,* 219–232.

Glaser, R., Rice, J., Sheridan, J., Fertal, R., Stout, J., Speicher, C., Pinshy, D., Kotur, M., Post, A., Beck, M., & Kiecolt-Glaser, J. (1987). Stress-related immune suppression: Health implications. *Brain, Behavior and Immunity, 1,* 7–20.

Glaser, R., Rice, J., Speicher, C. E., Stout, J. C., Kiecolt-Glaser, J. K. (1986). Stress depresses interferon production by leukocytes concomitant with a decrease in natural killer cell activity. *Behavioral Neuroscience, 100,* 675–678.

Glassman, A., Bennet, C., & Randal, C. (1985). Effects of ethyl alcohol on human peripheral lymphocytes. *Archives of Pathology and Laboratory Medicine, 109,* 540.

Gold, R., & Skinner, M. (1992). Situational factors and thought processes associated with unprotected intercourse in young gay men. *AIDS, 6,* 1021–1030.

Goodell, S., Quinn, T., Mkritichian, E., et al. (1983). Herpes simplex virus proctitis in homosexual men: Clinical, sigmoidoscopic, and histopathological features. *New England Journal of Medicine, 308,* 868–871.

Goodkin, K., Fuchs, I., Feaster, D., Leeka, J., & Rishel, D. D. (1993). Life stressors and coping style are associated with immune measures in HIV-1 infection—a preliminary report. *International Journal of Psychiatry in Medicine, 22,* 155–172.

Gorman, J., Kertzner, R., Cooper, T., Goetz, R., Lagomasino, I., Novacenko, H., Williams, J., Stern, Y., Mayeux, R., & Ehrhardt, A. (1991). Glucocorticoid level and neuropsychiatric symptoms in homosexual men with HIV infection. *American Journal of Psychiatry, 148,* 41–45.

Greden, J., Gardner, R., King, D., Grunhaus, L., Carroll, B., & Kronfol, A. (1983). Dexamethasone suppression tests in antidepressant treatment of melancholia: The process of normalization and test–retest reproducibility. *Archives of General Psychiatry, 40,* 493–500.

Greenspan, J., Greenspan, D., Lennette, E., et al. (1985). Replication of Epstein-Barr virus within the epithelial cells of oral "hairy" leukoplakia, an AIDS-associated lesion. *New England Journal of Medicine, 313,* 1564–1571.

Griensven, G., Vroome, E., Wolf, F., Goudsmit, J., Roos, M., & Coutinho, R. (1990). Risk factors for progression of human immunodeficiency virus (HIV) infection among seroconverted and seropositive homosexual men. *American Journal of Epidemiology, 132,* 203–210.

Griffiths, P. D., & Grundy, J. E. (1987). Molecular biology and immunology of cytomegalovirus. *Journal of Biochemistry, 241,* 313–324.

Guinan, M. (1987). Women, children and AIDS. *Journal of the American Medical Womens Association, 42,* 186.

Habu, S., Akamatsu, K., Tamaoki, N., Okumura, K. (1984). In vivo significance of the NK cell in resistance against (HSV-1) infections in mice. *Journal of Immunology, 133,* 2743–2747.

Hammarskjold, M. L., & Heimer, J. (1988). Activation of the HIV LTR by EBNA2 and LMP. *Journal of Experimental Cancer Research, 7,* 148.

Hardy, A. M., & the Long-term Survivor Collaborative Study Group (1991). Characterization of long-term survivors of acquired immunodeficiency syndrome. *Journal of Acquired Immune Deficiency Syndromes, 4,* 386–391.

Hatfield, S., Petersen, B., & DiMicco, J. (1986). Beta adrenoreceptor modulation of the generation of murine cytotoxic T lymphocytes in vitro. *Journal of Pharmacology and Experimental Therapeutics, 239*(2), 460–466.

Haverkos, H. W., Pinsky, P. F., Drotman, D. P., & Bregman, D. J. (1985). Disease manifestation among homosexual men with acquired immunodeficiency syndrome: A possible role of nitrates in Kaposi's Sarcoma. *Sexuality Transmitted Diseases, 12,* 203–208.

Hays, R. B., Chauncey, S., & Tobey, L. A. (1990). The social support networks of gay men with AIDS. *Journal of Community Psychology, 18,* 374–385.

Hays, R. B., Turner, H., & Coates, T. J. (1992). Social support, AIDS-related symptoms, and depression among gay men. *Journal of Consulting and Clinical Psychology, 60*(3), 463–469.

Henderly, D., Freeman, W., Smith, R. et al. (1987). Cytomegalovirus retinitis as the initial manifestation of acquired immune deficiency syndrome. *American Journal of Opthalmology, 103,* 316–320.

Herbert, T. B., & Cohen, S. (1993). Stress and immunity in humans: a meta-analytic review. *Psychosomatic Medicine, 55,* 364–379.

Hoffman, M. A. (1991). Counseling the HIV infected client: A psychosocial model for assessment and intervention. *The Counseling Psychologist, 19,* 467–542.

Hoover, D. R., Rinaldo, C., He, Y., Phair, J., Fahey, J., & Graham, N. M. H. (1995). Long-term survival without clinical AIDS after CD4+ cell counts fall below 200 × 10^6/1. *AIDS, 9,* 145–152.

Horowitz, M., Wilner, N., & Alvarez, (1979). Impact of event scale. *Psychosomatic Medicine, 41,* 209–218.

House, J. S., Landis, K. R., & Umberson, D. (1988). Social relationship and health. *Science, 241,* 540–545.

Ironson, G., Antoni, M., & Lutgendorf, S. (1995). Can psychological interventions affect immunity and survival? Present findings and suggested targets with a focus on cancer and human immunodeficiency virus. *Mind/Body Medicine, 1*(2), 85–110.

Ironson, G., Antoni, M., Simoneau, J., Starr, K., LaPerriere, A., Klimas, N., Schneiderman, N., & Fletcher, M. A. (1992). *Denial, distress, and treatment adherence predict disease progression in HIV-1 infected asymptomatic gay men.* Presented at American Psychosomatic Society, New York.

Ironson, G., Field, T., Scafidi, F., Hashimoto, M., Kumar, M., Kumar, A., Price, A., Goncalves, A., Burman, I., Tetenman, C., Patarca, R., & Fletcher, M. A. (1996). Massage therapy is associated with enhancement of the immune systems cytotoxic capacity. *International Journal of Neuroscience, 84,* 205–217.

Ironson, G., Friedman, A., Klimas, N., Antoni, M., Fletcher, M. A., LaPerriere, A., Simoneau, J., & Schneiderman, N. (1994). Distress, denial and low adherence to behavioral interventions predict faster disease progression in gay men infected with Human Immunodeficiency Virus. *International Journal of Behavioral Medicine, 1* (1), 90–105.

Ironson, G., LaPerriere, A., Antoni, M., O'Hearn, P., Klimas, N., Fletcher, M. A., & Schneiderman, N. (1990). Changes in immunologic and psychological measures as a function of anticipation and reaction to news of HIV-1 antibody status. *Psychosomatic Medicine, 52,* 247–270.

Ironson, G., Wynings, C., Schneiderman, N., Baum, A., Rodriguez, M., Greenwood, D., Benight, C. C., Antoni, M., LaPerriere, A., Huang, H. S., Klimas, N., & Fletcher, M. A. (1997). Post traumatic stress symptoms, intrusive thoughts, loss and immune function after Hurricane Andrew. *Psychosomatic Medicine, 59,* 128–141.

Irwin, M., Caldwell, C., Smith, T., Brown, S., Schuckit, M., & Gillin, C. (1990). Major depressive disorder, alcoholism, and reduced natural killer cell cytotoxicity. *Archives of General Psychiatry, 47,* 713–719.

Irwin, M., Daniels, M., Bloom, E., Smith, T., & Weiner, H. (1987). Life events, depressive symptoms, and immune function. *American Journal of Psychiatry, 144,* 437–441.

Irwin, M., Daniels, M., Smith, T., Bloom, E., & Weiner, H. (1987). Impaired natural killer cell activity during bereavement. *Brain, Behavior, and Immunity, 1*(1), 98–104.

Irwin, M., Daniels, M., & Weiner, H. (1987). Immune and neuroendocrine changes during bereavement. *Psychiatric Clinic of North America, 10*(3) 449–465.

Irwin, M., Patterson, T., Smith, T., Caldwell, C., Brown, S., Grant, I., & Gillin, J. C. (1990). Reduction of immune function by life stress and depression. *Biological Psychiatry, 27,* 22–30.

Jacobs, S., Mason, J., Kosten, T., Wahby, V., Kasl, S., & Ostfeld, A. (1986). Bereavement and catecholamines. *Journal of Psychosomatic Research, 30,* 489–496.

Jacobsen, P., Perry, S., & Hirsch, D. (1990). Behavioral and psychological responses to HIV antibody testing. *Journal of Consulting and Clinical Psychology, 58*(1), 31–37.

Jacobson, M. A., Bacchetti, P., Kolokathis, A., Chaisson, R. E., Szabo, S., Polsky, B., Valainis, G. T., Mildvan, D., Abrams, D., Wilber J., Winger, E., Sacks, H. S., Hendricksen, C., & Moss, A. (1991). Surrogate markers for survival in patients with AIDS and AIDS-related complex treated with zidovudine. *British Medical Journal, 302,* 73–78.

Justice, A. (1985). Review of the effects of stress on cancer in laboratory animals: Importance of time of stress application and type of tumor. *Psychological Bulletin, 98,* 108–138.

Kaisch, K., & Anton-Culver, H. (1989). Psychological and social consequences of HIV exposure: Homosexuals in Southern California. *Psychology and Health, 3,* 63–75.

Kaplan, L. D., Wofsy, C. B., & Volberding, P. A. (1987). Treatment of patients with acquired immunodeficiency syndrome and associated manifestations. *Journal of the American Medical Association, 257,* 1367–1374.

Katz, P., Zaytoun, A., & Fauci, A. (1982). Mechanisms of human cell-mediated cytotoxicity. Modulation of natural killer cell activity by cyclic nucleotides. *Journal of Immunology, 129,* 287–296.

Kelly, J. A. Murphy, D., Bahr, G., Koob, J., Morgan, M., Kalichman, S., Stevenson, L., Brashfield, T., Bernstein, B., & St. Lawrence, J. S. (1993). Factors associated with severity of depression and high-risk sexual behavior among persons diagnosed with human immunodeficiency virus (HIV) infection. *Health Psychology, 12,* 215–219.

Kelly, J. A., & St. Lawrence, J. S. (1988). AIDS prevention and treatment: Psychology's role in the health crisis. *Clinical Psychology Review, 8,* 255–284.

Kemeny M. E. (1994). Stressful events, psychological responses and progression of HIV infection. In R. Glaser & J. K. Kiecolt-Glaser (Eds.), *Handbook on human stress and immunity* (pp. 245–266). San Diego, CA: Academic Press.

Kemeny, M. E., Weiner, H., Duran, R., Taylor, S. E., Visscher, B., & Fahey, J. L. (1995). Immune system changes after the death of a partner in HIV-positive gay men. *Psychosomatic Medicine, 57,* 547–554.

Kemeny, M. E., Weiner, H., Taylor, S. E., Schneider, S., Visscher, B., & Fahey, J. L. (1994). Repeated bereavement, depressed mood, and immune parameters in HIV seropositive and seronegative gay men. *Health Psychology, 13,* 14–24.

Kessler, R. C., Foster, C., Joseph, J., Ostrow, D., Wortman, C., Phair, J., & Chmiel, J. (1991). Stressful life events and symptom onset in HIV infection. *American Journal of Psychiatry, 148,* 733–738.

Kiecolt-Glaser, J. K., Dura, J. R., Speicher, C. E., Trask, O. J., & Glaser, R. (1991). Spousal caregivers of dementia victims: Longitudinal changes in immunity and health. *Psychosomatic Medicine, 53,* 345–362.

Kiecolt-Glaser, J. K., Fisher, L. D., Ogrocki, P., Stout, J. C., Speicher, C. E., & Glaser, R. (1987). Marital quality, marital disruption, and immune function. *Psychosomatic Medicine, 49,* 13–34.

Kiecolt-Glaser, J. K., Garner, W., Speicher, C., Penn, G. M., Holiday, J., & Glaser, R. (1984). Psychosocial modifiers of immunocompetence in medical students. *Psychosomatic Medicine, 46,* 7–14.

Kiecolt-Glaser, J. K., Glaser, R., Shuttleworth, E. C., Dyer, C. S., Ogrocki, P., & Speicher, C. E. (1987). Chronic stress and immunity in family caregivers of Alzheimer's disease victims. *Psychosomatic Medicine, 49,* 523–535.

Kiecolt-Glaser, J. K., Glaser, R., Strain, E., Stout, J., Tarr, K., Holiday, J., & Speicher, C. E. (1986). Modulation of cellular immunity in medical students. *Journal of Behavioral Medicine, 9,* 5–21.

Kiecolt-Glaser, J. K., Glaser, R., Williger, G., Stout, J., Messick, G., Sheppard, S., Ricker, D., Romisher, S. C., Briner, W., Bonnell, G., Donnerberg, R. (1985). Psychosocial enhancement of immunocompetence in a geriatric population. *Health Psychology, 4*(1), 25–41.

Kiecolt-Glaser, J. K., Kennedy, S., Malkoff, S., Fisher, L., Speicher, C. E., & Glaser, R. (1988). Marital discord and immunity in males. *Psychosomatic Medicine, 50,* 213–229.

Kingsley, L., Kaslow, R., Rinaldo, C., Detre, K., Odaka, N., VanRaden, M., Detels, R., Polk, B., Chmiel, J., Kelsey, S., Ostrow, D., & Visscher, B. (1987). Risk factors for sero-conversion to human immunodeficiency virus among male homosexuals. *Lancet, 1,* 345–349.

Klimas, N., Caralis, P., LaPerriere, A., Antoni, M. H., Schneiderman, N., & Fletcher, M. A., (1991). Evaluation of immunological status in a cohort of HIV-1 positive and negative healthy gay men. *Journal of Clinical Microbiology, 29,* 1413–1421.

Krueger, J., & Karnovsky, M. (1987). Sleep and the immune response. *Annals of the New York Academy of Sciences, 496,* 510–516.

Kubler-Ross, E. (1969). *On death and dying*. New York: MacMillan.

Kumar, M., Morgan, R., Szapocznik, J., & Eisdorfer, C. (1991). Norepinephrine response in HIV+ subjects. *Journal of Acquired Immune Deficiency Syndrome, 4,* 782–785.

LaPerriere, A., Antoni, M. H., Schneiderman, N., Ironson, G., Klimas, N., Klimas, N., Caralis, P., & Fletcher, M. A. (1990). Exercise intervention attenuates emotional distress and natural killer cell decrements following notification of positive serologic status for HIV-1. *Biofeedback and Self-Regulation, 15,* 125–142.

LaPerriere, A., Ironson, G., Antoni, M. H., Schneiderman, N., Klimas, N., & Fletcher, M. A. (1994). Exercise and psychoneuroimmunology. *Medicine and Science in Sports and Exercise, 26,* 182–190.

Lazarus, R. S., & Folkman, S. (1984). *Stress appraisal and coping*. New York: Springer.

Lazzarin, A., Mella, L., & Trombini, M. (1984). Immunologic status in heroine addicts: Effects of methadone maintenance treatment. *Drug and Alcohol Dependency, 13,* 117–123.

Lemp, G. F., Payne, S. F., Neal, D., Temelso, T., & Rutherford, G. W. (1990). Survival trends for patients with AIDS. *JAMA, 263,* 402–406.

Lesserman, J., DiSantostefano, R., Perkins, D., Murphy, C., Golden, R., & Evans, D. (1995). Longitudinal study of social support and social conflict as predictors of depression and dysphoria among HIV-positive and HIV-negative gay men. *Depression, 2,* 189–199.

Lesserman, J., Perkins, D., & Evans, D. (1992). Coping with the threat of AIDS: The role of social support. *American Journal of Psychiatry, 149,* 1514–1520.

Levine, M., Woldenberg, R., Herlinger, H., & Laufer, I. (1987). Opportunistic esophagitis in AIDS: Radiographic diagnosis. *Radiology, 165,* 815–820.

Levy, S., Herberman, R., Lippman, M., & d'Angelo, T. (1987). Correlation of stress factors with sustained depression of natural killer cell activity and predicted prognosis in patients with breast cancer. *Journal of Clinical Oncology, 5,* 348–353.

Lewin, D. (1996). Protease inhibitors: HIV-1 summons a Darwinian defense. *Journal of NIH Research, 8,* 33–35;

Linn, B. S., Linn, M. W., & Jensen, J. (1981). Anxiety and immune responsiveness. *Psychological Reports, 49,* 969–970.

Lohrenz, L., Connelly, L., Coyne, L., & Spare, K. (1978). Alcohol problems in several midwestern homosexual communities. *Journal of Studies of Alcohol, 39,* 1959–1963.

Lollis, C., Antoni, M. H., Johnson, E., Chitwood, D., & Griffin, D. (1995). Does the health belief model predict risky sexual practices in injection drug users? *Clinical Psychology and Psychotherapy, 2*(4), 227–233.

Lollis, C., Johnson, E., Antoni, M. H., & Hinkle, Y. (1996). Characteristics of African Americans with multiple risk factors associated with HIV/AIDS. *Journal of Behavioral Medicine, 19,* 55–70.

Longini, I. M., Clark, W. S., & Karon, J. M. (1993). The effect of routine use of therapy in slowing the clinical course of HIV infection in a population-based cohort. *American Journal of Epidemiology, 137,* 1229–1240.

Lusso, P., Ensoli, B., Markham, P. D., Ablashi, D. V., Salahuddin, S. Z., Tschachler, E., Wong-Staal, F., & Gallo, R. C. (1989). Productive dual infection of human CD4+ T lymphocytes by HIV-1 and HHV-6. *Nature, 337,* 370–373.

Lutgendorf, S., Antoni, M. H., Ironson, G., Fletcher, M. A., Renedo, F., VanRiel, F., Baum, A., Schneiderman, N., & Klinas, N. (1995). Physical symptoms of chronic fatigue syndrome are exacerbated by the stress of Hurricane Andrew. *Psychosomatic Medicine, 57,* 310–323.

Lutgendorf, S., Antoni, M. H., Ironson, G., Klimas, N., Starr, K., Schneiderman, N., & Fletcher, M. A. (1996, March). *Coping and social support predict distress changes in symptomatic HIV-seropositive gay men following a cognitive behavioral stress management intervention*. Paper presented at the Society of Behavioral Medicine, Washington, DC.

Lutgendorf, S., Antoni, M. H., Ironson, G., Klimas, N., Starr, K., Schneiderman, N., McCabe, P., Cleven, K., & Fletcher, M. A. (1997). Cognitive behavioral stress management decreases dysphoric mood and Herpes Simplex Virus-Type 2 antibody titers in symptomatic HIV-seropositive gay men. *Journal of Consulting and Clinical Psychology, 65,* 31–43.

Lutgendorf, S. K., Antoni, M. H., Ironson, G., Schneiderman, N., & Fletcher, M. A. (1997). Cognitive processing style, affective status, and immune function following HIV seropositivity notification. *Cognitive Therapy and Research, 57,* 60–61.

Lutgendorf, S. K., Antoni, M. H., Kumar, M., & Schneiderman, N. (1994). Changes in cognitive coping strategies predict EBV antibody titre following a stressor disclosure induction. *Journal of Psychosomatic Research, 38,* 63–78.

Lutgendorf, S., Antoni, M. H, Schneiderman, N., & Fletcher, M. A. (1994). Psychosocial counseling to improve quality of life in HIV-infected infection. *Patient Education and Counseling, 24,* 217–235.

Lutgendorf, S., Antoni, M. H., Schneiderman, N., Ironson, G., & Fletcher, M. A. (1994). Psychosocial interventions and quality of life changes across the HIV spectrum. In A. Baum & J. Dimsdale (Eds.), *Perspectives in behavioral medicine* (pp. 205–239). Hillsdale, NJ: Erlbaum.

Lyketsos, C. G., Hoover, D. R., Guccione, M., Senterfitt, W., Dew, M. A., Wesch, J., VanRaden, M., Treisman, G. J., & Morganstem, H. (1993). Depressive symptoms as predictors of medical outcomes in HIV infection. *Journal of the American Medical Association, 270,* 2563–2567.

Mackie, I. (1993). AIDS, the female patient and the family physician. *Canadian Family Physician, 39,* 1600–1607.

Maier, S. F., Watkins, L. R., & Fleshner, M. (1994). Psychoneuroimmunology: The interface between behavior, brain, and immunity. *American Psychologist, 49,* 1004–1017.

Maiman, M., & Fruchter, R. (1996). Cervical neoplasia and the human immunodeficiency virus. In S. Rubino & W. Hoskins (Eds), *Cervical cancer and preinvasive neoplasia* (pp. 405–416). Philadelphia: Lippincott-Raven.

Manuck, S. (1991), Individual differences in cellular immune response to stress. *Psychological Science, 2,* 111–115.

Markham, P., Salahuddin, S., Veren, K., Orndorff, S., & Gallo, R. (1986). Hydrocortisone and some other hormones enhance the expression of HTLV-III. *International Journal of Cancer, 37,* 67–72.

Martin, D. (1989). Human immunodeficiency virus infection and the gay community: Counseling and clinical issues. Special Issue: Gay, lesbian, and bisexual issues in counseling. *Journal of Counseling and Development, 68,* 67–72.

Martin, J. (1988). Psychological consequences of AIDS-related bereavement among gay men. *Journal of Consulting and Clinical Psychology, 56,* 856–862.

Martin, J. (1990). Drug use and unprotected anal intercourse among gay men. *Health Psychology, 9,* 450–465.

Marzuk, P. M., Tierney, H., Tardiff, K., Gross, E. M., Morgan, E. B., Hsu, M. A., & Mann, J. (1988). Increased risk of suicide in persons with AIDS. *Journal of the American Medical Association, 259,* 1333–1337.

Mays, V., & Cochran, S. (1988). Issues in the perception of AIDS risk and risk reduction activities by Black and Hispanic/Latina women. *American Psychologist, 43*(11), 949–957.

McKinnon, W., Weisse, C. S., Reynolds, C. P., Bowles, C. A., & Baum, A. (1989). Chronic stress, leukocyte subpopulations, and humoral response to latent viruses. *Health Psychology, 8,* 399–402.

McKusick, L., Horstman, W., & Coates, T. (1985). AIDS and sexual behavior reported by gay men in San Francisco. *American Journal of Public Health, 75,* 493–496.

McNair, D. Lorr, M., & Droppleman, L. (1981). *EITS manual for the profile of mood states*. San Diego, CA: Educational and Industrial Testing Service.

McVoy, M. A., & Adler, S. P. (1989). Immunologic evidence for frequent age-related cytomegalovirus reactivation in seropositive immunocompetent individuals. *Journal of Infectious Disease, 160,* 1–10.

Miller, D. (1988). HIV and social psychiatry. *British Medical Bulletin, 44,* 130–148.

Miller, L., Goldstein, G., Murphy, M., & Ginns, L. C. (1982). Reversible alterations in immunoregulatory T cells in smoking. *Chest, 82,* 526–529.

Mulder, C. L., Antoni, M. H., Duivenvoorden, H., Kauffmann, R., & Goodkin, K. (1995). Active confrontational coping predicts decreased clinical progression over a one-year period in HIV-infected homosexual men. *Journal of Pyschosomatic Research, 39*(8), 957–965.

Mulder, C. L., Antoni, M. H., Emmelkamp, P., Veugelers, P., Sandfort, T., Vijver, F. vander, & Vries, M. de (1995). Psychosocial group intervention and the rate of decline of immunologic parameters in asymptomatic HIV-infected homosexual men. *Journal of Psychotherapy and Psychosomatics, 63,* 185–192.

Mulder, C. L., Emmelkamp, P., Antoni, M. H., Mulder, J., Sandfort, T., & de Vries, M. (1994). Cognitive-behavioral and experiential group psychotherapy for HIV-infected homosexual men: A comparative study. *Psychosomatic Medicine, 56,* 423–431.

Mulder, C. L., Vroome, E. de, Griensven, G. van, Antoni, M. H., & Sandfort. T. (in press). Avoidance as a predictor of the virological course of HIV-1 infection over a 7-year period in gay men. *Health Psychology*.

Munoz, A., Wang, M. C., Good, R., Detels, H., Ginsberg, L., Kingsley, J., Phair, J., & Polk, B. F. (1988, June). *Estimation of the AIDS-free times after HIV-1 seroconversion*. Paper presented at the IVth Annual Meeting of the International Conference on AIDS, Stockholm, Sweden.

Naliboff, B., Benton, D., Solomon, G., Morley, J., Fahey, J., Bloom, E., Makinodan, T., & Gilmore, S. (1991). Immunological changes in young and old adults during brief laboratory stress. *Psychosomatic Medicine, 53,* 121–132.

Namir, S., Alumbaugh, M. J., Fawzy, F. I., & Wolcott, D. L. (1989). The relationship of social support to physical and psychological aspects of AIDS. *Psychology and Health, 3,* 77–86.

Namir, S., Wolcott, D. L., Fawzy, F. I., & Alumbaugh, M. J. (1987). Coping with AIDS: Psychological and health implications. *Journal of Applied Social Psychology, 17,* 309–328.

Nardi, P. (1982). Alcoholism and homosexuality: A theoretical perspective. *Journal of Homosexuality, 7,* 9–25.

Neale, A., Tilley, B., & Vernon, S. (1986). Marital status, delay in seeking treatment and survival from breast cancer. *Social Science and Medicine, 23,* 305–312.

Newcomb, M., & Harlow, L. (1986). Life events and substance use in adolescents: Mediating effects of perceived loss of control and meaninglessness in life. *Journal of Personality and Social Psychology, 51,* 564–577.

Nieman, R., Fleming, J., Coker, R., Harris, J., & Mitchell, D. (1993). The effect of cigarette smoking on the development of AIDS in HIV-1-seropositive individuals. *AIDS, 7,* 705–710.

Noh, S., Chandarana, P., Field, V., & Posthuma, B. (1990). AIDS epidemic, emotional strain, coping and psychological distress in homosexual men. *AIDS Education and Prevention, 2,* 272–283.

Nyamathi, A., & Lewis, C. (1991). Coping of African American women at risk for AIDS. *Women's Health Issues, 1,* 53–61.

O'Brien, K. O., Wortman, C. B., Kessler, R. C., & Joseph, J. G. (1993). Social relationships of men at risk for AIDS. *Social Science and Medicine, 36*(9), 1161–1167.

Ostrow, D., Joseph, J., Kessler, R., Soucy, J., Tal, M., Eller, M., Chmiel, J., & Phair, J. (1989). Disclosure of HIV antibody status: Behavioral and mental health correlates. *AIDS Education and Prevention, 1,* 1–11.

Ostrow, D. G., Monjan, A., Joseph, J., VanRaden, M., Fox, R., Kingsley, L., Lawrence, K., Dudley, J., & Phair, J. (1989). HIV-related symptoms and psychological functioning in a cohort of homosexual men. *American Journal of Psychiatry, 146,* 737–742.

Palmblad, J., Cantrell, K., Strander, H., Froberg, J., Karlesson, C., Levi, L., Granstan, M., & Unger, P. (1976). Stressor exposure and immunological response in man: Interferon-producing capacity and phagocytosis. *Journal of Psychosomatic Research, 20,* 193–199.

Palmblad, J., Petrini, B., Wasserman, J., & Akerstedt, T. (1979). Lymphocyte and granulocyte reactions during sleep deprivation. *Psychosomatic Medicine, 41,* 273–278.

Pantaleo, G., Poli, G., & Fauci, A. S. (1993). The immunopathogenesis of human immunodeficiency virus infection. *Clinical and Infectious Diseases, 17*(Suppl. 1), S224–S229.

Patterson, T., Shaw, W., Semple, S., Cherner, M., McCutchan, J., Atkinson, J., Grant, I., & Nannis, E. (1996).

Relationship of psychosocial factors to HIV disease progression. *Annals of Behavioral Medicine, 18,* 30–39.

Paya, C. V., Virelizier, J. L., & Michelson, S. (1991). Modulation of T-cell activation through protein kinase C- or A-dependent signalling pathways synergistically increases human immunodeficiency virus long terminal repeat induction by cytomegalovirus immediate-early proteins. *Journal of Virology, 65,* 5477–5484.

Pearlin, L., & Radabaugh, C. (1976). Economic strains and coping functions of alcohol. *American Journal of Sociology, 82,* 652.

Penkower, L., Dew, M. A., Kingsley, L., Becker, J. T., Satz, P., Schaerf, F. W., & Sheridan, K. (1991). Behavioral, health and psychosocial factors and risk for HIV infection among sexually active homosexual men: The multicenter AIDS cohort study. *American Journal of Public Health, 81,* 194–196.

Perssan, U., Myrenfors, P., Ringden, O., Sundberg, B., Larsson, P., Gunnar, S., & Johansson, O. (1987). T-lymphocytes subpopulations in bone marrow-transplanted patients in relation to graft-versus-host disease and cytomegalovirus-induced infection. *Transplantation, 43,* 663–668.

Pickrel, J. (1989). "Tell me your story": Using life review in counseling the terminally ill. *Death Studies, 13,* 127–135.

Plaut, M. (1987). Lymphocyte hormone receptors. *Annual Review of Immunology, 5,* 621–669.

Quinn, S. (1993). AIDS and the African American woman. The triple burden of race, class and gender. *Health Education Quarterly, 20,* 305–320.

Rabkin, J., Rabkin, R., Harrison, W., & Wagner, G. (1994). Effect of imipramine on mood and enumerative measures of immune status in depressed patients with HIV illness. *American Journal of Psychiatry, 151,* 516–523.

Rabkin, J., Wagner, G. & Rabkin, R. (1996). Treatment of depression in HIV+ men: Literature review and report of an ongoing study of testosterone replacement therapy. *Annals of Behavioral Medicine, 18,* 24–29.

Rabkin, J., Williams, J., Neugebauer, R., & Goetz, R. (1990). Maintenance of hope in HIV-spectrum homosexual men. *American Journal of Psychiatry, 147,* 1322–1326.

Rabkin, J. G., Williams, J. B. W., Remien, R. H., Goetz, R. R., Kertzner, R., & Gorman, J. M. (1991). Depression, lymphocyte subsets, and human immunodeficiency virus symptoms on two occasions in HIV-positive homosexual men. *Archives of General Psychiatry, 48,* 111–119.

Redfield, R., & Burke, D. (1988). HIV infection: The clinical picture. *Scientific American, 259,* 90–98.

Reed, G. M., Kemeny, M. E., Taylor, S. E., Visscher, B. R., & Fahey, J. L. (1994, April). *Negative HIV-specific expectancies and health outcomes in HIV-related disease.* Invited symposium presented at the 52nd Annual Scientific Meeting of the American Psychosomatic Society, Boston, MA.

Reed, G. M., Kemeny, M. E., Taylor, S. E., Wang, H. J., & Visscher, B. (1994). Realistic acceptance as a predictor of decreased survival time in gay men with AIDS. *Health Psychology, 13,* 299–307.

Richman, D. D., Fischl, M., Grieco, M., Gottlieb, M., Volberding, P., Laskin, O., Leedom, J., Groopman, J., Mildvan, D., Hirsch, M., Jackson, G., Durack, D., & Nusinoff-Lehrman, S. (1987). The toxicity of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. *New England Journal of Medicine, 317,* 192–197.

Rinaldo, C. R. (1990). Immune suppression by herpes viruses. *Annual Reviews of Medicine, 41,* 331–338.

Rose, R. (1980). Endocrine responses to stressful psychological events. *Psychiatric Clinics of North America, 3,* 251–276.

Rosenberg, Z. F., & Fauci, A. S. (1991). Activation of latent HIV infection. *Journal of the National Institutes of Health Research, 2,* 41–45.

Rosser, B., Simon, R., & Michael, W. (1988). Perceived emotional and life change impact of AIDS on homosexual men in two countries. *Psychology and Health, 2,* 301–317.

Roszman, T., Jackson, J., Cross, R., Titus, M., Markesbery, W., & Brooks, N. H. (1985). Neuroanatomic and neurotransmitter influences on immune function. *Journal of Immunology, 135,* 769s–772s.

Royce, R., & Winkelstein, W. (1990). HIV infection, cigarette smoking and CD4+ T-L lymphocyte counts: Preliminary results from the San Francisco Men's Health Study. *AIDS, 4,* 327–333.

Rundell, J., Paolucci, S., Beatty, D., & Boswell, R. (1988). Psychiatric illness at all stages of human immunodeficiency virus infection [letter]. *American Journal of Psychiatry, 145,* 652–653.

Rush, A., Beck. A., Kovacs, M., Weisenberger, J., & Hollon, S. (1982). Comparison of the effects of cognitive therapy and pharmacotherapy on hopelessness and self-concept. *American Journal of Psychiatry, 139,* 862–866.

Saah, A. J., Hoover, D. R., He, Y., Kingsley, L. A., Phair, J. P., & the Multicenter AIDS Cohort Study (1994). Factors influencing survival after AIDS: Report from the Multicenter AIDS Cohort Study (MACS). *Journal of Acquired Immune Deficiency Syndromes, 7,* 287–295.

Sarason, I. (1979, August). *Life stress, self-preoccupation, and social supports.* Paper presented at the annual scientific meetings of the American Psychological Association.

Sarason, I., Potter, E., Sarason, B., & Antoni, M. H. (1985). Life events, social support, and illnesses. *Psychosomatic Medicine, 47*(2), 156–163.

Schneiderman, N., Antoni, M. H., & Ironson, G. (1997). Cognitive behavioral stress management and secondary prevention in HIV/AIDS. *Psychology & AIDS Exchange, 22,* 1–8.

Schneiderman, N., Antoni, M. H., Ironson, G., Fletcher, M. A., Klimas, N., & LaPerriere, A. (1994). HIV-1, immunity and behavior. In Glaser, R. (Ed.) *Handbook* of human stress and immunity. San Diego, CA: Academic Press.

Schneiderman, N., Antoni, M., Ironson, G., LaPerriere, A., & Fletcher, M. A. (1992). Applied psychosocial science and HIV-1 spectrum disease. *Journal of Applied and Preventative Psychology, 1,* 67–82.

Schooley, R. (1990). Cytomegalovirus infection in the seting of infection with the human immunodeficiency virus. *Reviews of Infectious Disease, 17,* S811.

Schwartzberg, S. (1994, August). *HIV as a catalyst for personal growth: Clinical and qualitative data.* Paper presented at the annual meeting of the American Psychological Association, Los Angeles, CA.

Seage, G., Oddleifson, S., Carr, E., Shea, B., Makarewicz-Robert, L., vanBeuzekom, M., & DeMaria, A. (1993). Survival with AIDS in Massachusetts, 1979 to 1989. *American Journal of Public Health, 83,* 72–78.

Shine, D., Moll, B., Emerson, E., Spigland, I., Harris, C., Small, C. B., Friedland, G., Weiss, S. H., & Bodner, A. J. (1987). Serologic, immunologic, and clinical features of parental drug users from contrasting populations. *American Journal of Drug Abuse, 13,* 401–412.

Sikkema, K., & Kelly, J. (1996). Behavioral medicine interventions can improve quality of life and health of persons with HIV disease. *Annals of Behavioral Medicine, 18,* 40–48.

Simons, A., Garfield, S., & Murphy, G. (1984). The process of change in cognitive therapy and pharmacotherapy: Changes in mood and cognition. *Archives of General Psychiatry, 41,* 45–51.

Smith, C., Fernengel, K., Holcroft, C., Gerald, K., & Marien, L. (1994). Meta-analysis of the associations between social support and health outcomes. *Annals of Behavioral Medicine, 16,* 352–362.

Solano, L., Costa, M., Salvati, S., Coda, R., Aiuta, F., Mezzaroma, I., & Bertini, M. (1993). Psychosocial factors and clinical evolution in HIV-1 infection: A longitudinal study. *Journal of Psychosomatic Research, 37*(1), 39–51.

Solomon, G., Kemeny, M., & Temoshok, L. (1991). Psychoneuroimmunologic aspects of Human Immunodeficiency Virus infection. In R. Ader, D. Felten, & N. Cohen (Eds.), *Psychoneuroimmunology* (2nd ed., pp. 1081–1114). San Diego, CA: Academic Press.

Solomon, G., Temoshok, L., O'Leary, A., & Zich, J. (1987). An intensive psychoimmunologic study of long-surviving persons with AIDS. *Annals of the New York Academy of Science, 496,* 647–655.

Stall, R., Coates, T., & Hoff, C. (1988). Behavioral risk reduction for HIV infection among gay and bisexual men: A review of results from the United States. *American Psychologist, 43,* 878–885.

Starr, K., Antoni, M. H., Hurwitz, B., Rodriquez, M., Ironson, G., Fletcher, M. A., Kumar, M., Patarca, R., Lutgendorf, S., Quillian, R., Klimas, N., & Schneiderman, N. (1996). Patterns of immune, neuroendocrine, and cardiovascular stress responses in asymptomatic HIV seropositive and seronegative men. *International Journal of Behavioral Medicine, 3,* 135–162.

Stein, M., Miller, A. H., & Trestman, R. L. (1991). Depression, the immune system and health and illness: Findings in search of meaning. *Archives of General Psychiatry, 48,* 171–177.

Stites, D., Stobo, J., Fudenberg, H., & Wells, J. (1982). *Basic and clinical immunology* (4th ed.). Los Altos, CA: Lange.

Sumaya, C. V., Boswell, R. N., Ench, Y., Kisner, D. L., Hersh, E. M., Ruben, J. M., & Mansell, P. W. (1986). Enhanced serological and virological findings of Epstein-Barr virus in patients with AIDS and AIDS-related complex. *Journal of Infectious Disease, 154,* 864–870.

Tapper, M., Rotterdam, H., Lerner, C. et al. (1984). Adrenal necrosis in the acquired immunodeficiency syndrome. *Annals of Internal Medicine 100,* 239–241.

Taylor, S. E., Falke, R. L., Shoptaw, S. J., & Lichtman, R. R. (1986). Social support, support groups and the cancer patient. *Journal of Consulting and Clinical Psychology, 54,* 608–615.

Theorell, T., Blomkvist, V., Jonsson, H., Schulman, S., Berntorp, E., & Stigendal, L. (1995). Social support and the development of immune function in human immunodeficiency virus infection. *Psychosomatic Medicine, 57,* 32–36.

Tindall, B., & Cooper, D. A. (1991). Primary HIV infection: Host responses and intervention strategies. *AIDS, 5,* 1–14.

Tisman, F., Herbert, V., Go, L. et al. (1971). In vitro demonstration and immunosuppression without bone marrow suppression by alcohol and bleomycin. *Clinical Research, 19,* 730.

Tross, S., & Hirsch, D. A. (1988). Psychological distress and neuropsychological complications of HIV infection and AIDS. *American Psychologist, 43,* 929–934.

Turk, D., Holzman, A., & Kerns, R. (1986). Chronic pain. In K. Holroyd, & T. Creer (Eds.), *Self-management of chronic disease: Handbook of clinical interventions and research.* Orlando, FL: Academic Press.

Turner, H. A., Hays, R. B., & Coates, T. J. (1993). Determinants of social support among gay men: The context of AIDS. *Journal of Health and Social Behavior, 34* (March), 37–53.

Vachon, M., Lyall, W., Rogers, K., Freedman, K., & Freeman, S. (1980). A controled study of self-help intervention with widows. *American Journal of Psychiatry, 137*(11), 1380–1384.

Viney, L., Henry, R., Walker, B., & Crooks, L. (1989). The emotional reactions of HIV antibody positive men. *British Journal Medical Psychology, 62,* 153–161.

Wachtel, T., Piette, M. S., Mor, V., Stein, M., Fleishman, J., & Carpenter, C. (1992). Quality of life in persons with Human Immunodeficiency Virus infection: Measurement by the Medical Outcomes Study instrument. *Annals of Internal Medicine, 116,* 129–137.

Walkey, F. H., Taylor, A. J., & Green, D. E. (1990). Attitudes to AIDS; A comparative analysis of a new and negative stereotype. *Social Science and Medicine, 30,* 549–552.

Watson, R., Eskelson, C., & Hartman, B. (1984). Severe alcohol abuse and cellular immune functions. *Arizona Medicine, 41,* 665–668.

Weber, R., Ledergerber, B., Opravil, M., Siegenthaler, W., & Luthy, R. (1990). Progression of HIV infection in misusers of injected drugs who stop injecting or follow a program of maintenance treatment with methadone. *British Medical Journal, 301,* 1362–1365.

Webster, A. (1991). Cytomegalovirus as a possible cofactor in HIV disease progression. *Journal of Acquired Immune Deficiency Syndrome, 4*(Suppl.) S47–S52.

Weimer, E., Nilsson-Schonnesson, L., & Clement, U. (1989). HIV-Infektion: Trauma und Traumaverarbeitung. *Psyche-Zeitschrift-fur-Psychoanalyse-und-inhre-Anwendungen, 43,* 720–735.

Wiklund, I., Oden, A., Sanne, H., Ulvenstam, G., Wilhelmsson, C., & Wilhemsen, L. (1988). Prognostic importance of somatic and psychosocial variables after a first myocardial infarction. *American Journal of Epidemiology, 128, 786–795.*

Williams, J., Rabkin, J., Remien, R. et al. (1991). Multidisciplinary baseline assessment of homosexual men with and without human immunodeficiency virus infection II. Standardized clinical assessment of current and lifetime psychopathology. *Archives of General Psychiatry, 48,* 124–130.

Williams, R., Barefoot, J., Califf, R., Haney, T., Saunders, W., Pryor, D., Hlatky, M., Siegler, I., & Mark, D. (1992). Prognostic importance of social and economic resources among medically treated patients with angiographically documented coronary artery disease. *Journal of the American Medical Association, 267,* 520–524.

Woods, T., Antoni, M. H., Ironson, G., & Kling, D. (1996). Religiosity and its relationship to affective distress, immune and health status in HIV-infected gay men. *Psychosomatic Medicine, 58,* 84.

Workman, E. A., & La Via, M. F. (1987). T-lymphocyte polyclonal proliferation: Effects of stress and stress response style on medical students taking national board examinations. *Clinical Immunology and Immunopathology, 43,* 308–313.

Wortman, C. B., & Conway, T. L. (1985). The role of social support in adaptation and recovery from physical illness. In S. Cohen & S. L. Syme (Eds.), *Social support and health* (pp. 281–302). New York: Academic Press.

Wortman, C. B., & Dunkel-Schetter, C. (1987). Conceptual and methodological issues in the study of social support. In A. Baum, S. Taylor, & J. E. Singer (Eds.), *Handbook of psychology and health* (vol. 5, pp. 63–108). Hillsdale, NJ: Erlbaum.

Yao, Q. Y., Rickingon, A. B., Gaston, J. S. H., & Epstein, M. A. (1985). In vitro analysis of the Epstein-Barr virus: Host balance in long-term renal allograft recipients. *International Journal of Cancer, 35,* 43–49.

Zich, J., & Temoshok, L. (1987). Perceptions of social support in men with AIDS and ARC: Relationships with distress and hardiness. *Journal of Applied Social Psychology, 17,*(3), 193–215.

Zuckerman, M., & Antoni, M. H. (1995). Social support and its relationship to psychological physical and immune variables in HIV infection. *Clinical Psychology and Psychotherapy, 2*(4), 210–219.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.11
Diabetes Mellitus

CLARE BRADLEY, AFSANE RIAZI, and SHALLEEN BARENDSE
Royal Holloway, University of London, UK
MARY B. PIERCE
Charing Cross and Westminster Medical School, London, UK
and
CHRISTEL HENDRIECKX
Vrije Universiteit Brussel, Brussels, Belgium

8.11.1 INTRODUCTION 278
  8.11.1.1 The Conditions 278
  8.11.1.2 Insulin Dependent Diabetes Mellitus 279
  8.11.1.3 Noninsulin Dependent Diabetes Mellitus 280
8.11.2 IDDM: ETIOLOGY, RISK FACTORS, AND PREVENTION 280
  8.11.2.1 Risk Factors in IDDM 281
  8.11.2.2 Prevention of IDDM 281
    8.11.2.2.1 Primary prevention 281
    8.11.2.2.2 Secondary prevention 282
    8.11.2.2.3 Tertiary prevention 282
8.11.3 NIDDM: ETIOLOGY, RISK FACTORS, AND PREVENTION 282
  8.11.3.1 Risk Factors in NIDDM 282
  8.11.3.2 Prevention of NIDDM 283
    8.11.3.2.1 Primary prevention 283
    8.11.3.2.2 Secondary prevention 283
    8.11.3.2.3 Tertiary prevention 284
8.11.4 IMPACT OF THE CONDITION ACROSS THE LIFE SPAN ON PERSON AND FAMILY 284
8.11.5 PSYCHOLOGICAL ASPECTS OF SYMPTOMS 285
8.11.6 SPECIFIC ISSUES OF ADHERENCE 286
8.11.7 PSYCHOLOGICAL INTERVENTIONS IN THE MANAGEMENT OF DIABETES 288
  8.11.7.1 Blood Glucose Awareness Training and Other Approaches to Regaining Hypoglycemia Awareness 288
  8.11.7.2 Identification of Stress Reactivity 288
    8.11.7.2.1 Mechanisms involved in stress and diabetes control 288
    8.11.7.2.2 Acute stress and blood glucose reactivity 289
    8.11.7.2.3 Life stress and blood glucose reactivity 289
  8.11.7.3 Stress Management Training 290
  8.11.7.4 Monitoring of Psychological Outcomes and Processes 291
    8.11.7.4.1 Measurement of psychological well-being and patient satisfaction 291
    8.11.7.4.2 Measurement of quality of life 292
    8.11.7.4.3 Measurement of knowledge, beliefs, and other cognitions 293
  8.11.7.5 Patient Empowerment 294
  8.11.7.6 Weight Management 295

8.11.7.6.1 Eating disorders and withholding of insulin 296
8.11.7.6.2 Weight reduction programs in NIDDM 297
8.11.7.7 Treatment of Sexual Dysfunction 298
8.11.8 CULTURAL ISSUES 298
8.11.9 PROFESSIONAL ISSUES 299
8.11.10 REFERENCES 299

8.11.1 INTRODUCTION

8.11.1.1 The Conditions

Diabetes mellitus is not a single disorder but a heterogeneous group of disorders. All forms are characterized by hyperglycemia and disturbances of carbohydrate, fat, and protein metabolism which are associated with absolute or relative deficiencies of insulin action and/or insulin secretion. The World Health Organization (WHO) developed a now widely accepted classification of the disorder, largely based on clinical characteristics (see Table 1, WHO, 1985).

Insulin is needed to allow glucose to pass from the blood into most of the body cells. Only the cells of the brain and central nervous system can use glucose from the blood in the absence of insulin. Without insulin, most body cells metabolize substances other than glucose for energy. However, fat metabolism in the absence of glucose metabolism, creates ketone bodies which are poisonous and their build up is associated with hyperglycemic coma. In the absence of sufficient insulin, unmetabolized glucose builds up in the blood. Water is drawn from body cells by osmosis to dilute the highly concentrated blood, and is then excreted along with much of the glucose, once the renal threshold for glucose (usually 10 mmol/L) is exceeded. Dehydration follows.

The characteristic presenting symptoms of diabetes are fatigue, extreme thirst, excessive urination, unexplained itch, recurrent infections (e.g., boils and thrush), and otherwise unexplained weight loss. These symptoms recur if treatment is not controling the condition adequately.

Complications of diabetes are responsible for considerable morbidity and mortality. The acute complications of diabetes are hypo- and hyperglycemic coma and infections. The chronic complications include microvascular complications such as retinopathy and nephropathy, and the macrovascular complications of heart disease and stroke. Diabetes mellitus is the commonest cause of blindness and renal failure in the UK and the USA. Other common complications include autonomic and peripheral neuropathy. A combination of vascular and neuropathic disturbances results in a high prevalence of impotence in men with diabetes. Peripheral neuropathy causes lack of sensation in the feet which can cause minor injuries to go unnoticed, become infected and, with circulatory problems obstructing healing, ulceration and gangrene are serious risks and amputation is not uncommon. Evidence from meta-analysis of studies of the relationship between glycemic control and microvascular complications (Wang, Lau, & Chalmers, 1993), and from the longitudinal multicenter Diabetes Control and Complications Trial (DCCT) in the USA (DCCT Research Group, 1993), have established a clear relationship between improved blood glucose control and reduction of risk of retinopathy and other microvascular complications in insulin-dependent diabetes mellitus (IDDM). It is likely that there would be similar findings for noninsulin-dependent diabetes mellitus (NIDDM) though the studies did not include NIDDM patients. However, the DCCT included highly selected, well-motivated, well-educated and well-supported patients, cared for by well-staffed diabetes care teams involving educators and psychologists as well as diabetologists and diabetes specialist nurses.

It is a considerable challenge to obtain the goals of the intensively treated patients in the DCCT with the vast majority of people with diabetes given the more limited health care resources typically available in routine practice. If diabetes control can be improved without significant damage to quality of life, the economic, health, and quality of life savings associated with a reduction in complications in later life will be vast. Although some people who have had poorly controlled diabetes over many years do not develop complications, complications commonly arise after 15–20 years of diabetes and individuals in their 40s or even 30s may develop several complications in rapid succession. However, up until the early 1980s, patients had no way of monitoring their own blood glucose levels at home. Urine glucose monitoring only told them when their blood glucose had exceeded the renal threshold of approximately 10 mmol/L (i.e., was far too high), without being able to discriminate between the too high levels of 7–10 mmol/L or the hypoglycemic levels below 4 mmol/L.

Table 1 WHO (1985) classification of diabetes mellitus.

Types	*Description*
1.	Insulin-dependent diabetes mellitus (IDDM)
2.	Noninsulin-dependent diabetes mellitus (NIDDM) a. Nonobese b. Obese
3.	Malnutrition-related disorders
4.	Other types of diabetes associated with certain conditions and syndromes including: pancreatic disease, diseases of hormonal etiology, drug-induced or chemical-induced conditions, abnormalities of insulin or its receptors, certain genetic syndromes
5.	Gestational diabetes mellitus

NIDDM and IDDM are the most common forms of diabetes in first-world countries, with gestational diabetes also being common.

Clinics relied on random blood glucose testing and there were no measures of average blood glucose over a longer period. Since the 1980s there have been measures of glycosylated hemoglobin (GHb, HbA_1, or HbA_{1c}) which indicate average blood glucose over a six to eight week period and measures of glycosylated protein, fructosamine, which indicates average blood glucose over a two-week period. Blood-glucose meters for patients were first introduced in the early 1980s and the accuracy and convenience of the meters and the reagent strips they use has improved dramatically since early models. By the late 1990s blood-glucose monitoring is part of the daily routine for most people using insulin in developed countries. Blood-glucose monitoring is less often prescribed for tablet- and diet-alone-treated patients, financial reasons probably being allowed to outweigh the educational value of accurate feedback in improving control long term. The reduced risk of hypoglycemia and diabetic ketoacidosis in NIDDM patients not using insulin means that acute crises rarely arise in these patients though their risk of long-term complications is at least as great as in IDDM and might be expected to be reduced if feedback from blood-glucose monitoring were provided.

8.11.1.2 Insulin Dependent Diabetes Mellitus

IDDM, also known as juvenile-onset diabetes, ketosis-prone diabetes, or Type 1 diabetes, is the commonest form of diabetes in children and young adults of European origin. Age at clinical onset of the condition is usually under the age of 30 years.

People who develop IDDM lose the ability to produce insulin altogether. Diagnosis and treatment with exogenous insulin is often followed by a "honeymoon" period when endogenous insulin production restarts and the exogenous insulin doses required are reduced and, occasionally, can be discontinued. This is, however, always only a temporary respite. Thereafter, exogenous insulin, usually delivered by subcutaneous injection, will be needed continuously throughout the lifetime of a person with IDDM. If insulin therapy is discontinued, even for a few days, ketotic coma will develop and, if not treated, will result in death.

The onset of IDDM is abrupt, with severe thirst, excessive urination, and dramatic weight loss. Individuals usually present to the doctor with one or more of these symptoms and an elevated blood glucose level. They will also have glucose, and possibly ketones, in their urine. This combination indicates diabetic ketoacidosis and is a potentially fatal medical emergency. Diagnosis is made clinically and confirmed on the basis of an elevated blood glucose concentration (> 6.7 mmol/L).

IDDM can be managed by balancing injected insulin, carbohydrate intake, and energy expenditure. The goal of treatment is to maintain blood glucose levels as close to the normal range as possible (i.e., 4–6 mmol/L) in order to reduce the risk of chronic complications, while also avoiding the dangers of blood glucose levels falling to hypoglycemic levels. To achieve this goal, people with IDDM will usually need to measure their blood glucose levels several times daily. In the 1980s and early 1990s, two injections a day were commonly used by IDDM patients. However, more recent evidence has shown that more intensive therapy, with up to five injections per day, is associated with reduced risk of chronic complications (DCCT Research Group, 1993; Wang et al., 1993) though at the price of threefold increases in occurrence of severe hypoglycemic episodes (DCCT). Intensive insulin regimens have become more commonly used in the 1990s and this trend has been facilitated by the development of insulin delivery devices such as pen-injectors which simplify the task of administering insulin, and continuing improvements in blood glucose monitoring methods (Bradley, Christie, & Home, 1991). The dangers of hypoglycemia, which tend to be increased with the intensity of the insulin therapy, give increasing cause for concern. Severe hypoglycemic episodes cause acute behavior changes with behavior often becoming aggressive, and unless action is taken to increase the blood glucose levels, collapse and coma follow rapidly. Insulin-dependent individuals need to carry with them at all times glucose tablets or some other form of rapidly absorbed glucose so that they can respond to early signs of hypoglycemia by increasing their blood glucose levels and thus prevent a severe hypoglycemic episode.

8.11.1.3 Noninsulin Dependent Diabetes Mellitus

NIDDM, also known as maturity-onset diabetes or Type 2 diabetes, is the commonest form of diabetes worldwide (85% of cases in the developed world). Age at clinical onset of the disorder is usually over the age of 40 years.

People with NIDDM continue to produce insulin, but it is inadequate either in amount or in action to maintain their blood glucose levels in the normal range. With the development of the pathological process involved, patients first have asymptomatic glucose intolerance and then go on to develop diabetes mellitus. Unlike IDDM, the process of development of NIDDM is insidious. Individuals may have NIDDM for years before it is diagnosed.

Diagnosis may result from patients presenting with classical symptoms, or with one or more of the complications of diabetes, or diabetes may be found on routine medical examination (confirmed by oral glucose tolerance testing). Dietary management is the first choice of treatment for NIDDM. A low-fat weight-reducing high fiber diet will be recommended to those who are overweight and if weight can be reduced and maintained close to ideal weight it may be the case that no further treatment is required. The diet recommended is usually high in complex carbohydrates and fiber and low in refined carbohydrates in order to avoid rapid absorption of glucose and to maintain blood glucose at a more even level without peaks of hyperglycemia. Reduction in fats may also be recommended (particularly if the individual has abnormal lipid levels) with a view to lowering and maintaining triglyceride and cholesterol levels and thereby reducing the risks of cardiovascular disease which accompany NIDDM.

If glucose levels cannot be controlled adequately with diet alone, oral hypoglycemic agents are added to treatment. These tablets do not contain insulin which, being a protein, would be digested if taken orally, but rather act to stimulate production of endogenous insulin or to enhance the effectiveness of existing levels of insulin. A variety of oral hypoglycemic drugs are available, such as sulfonylureas and biguanides. These differ in their modes of action and side effects, for example, sulfonylureas may increase the risk of hypoglycemia. If adequate control of blood glucose cannot be achieved with diet and tablets, insulin is likely to be recommended. It is important to recognize that NIDDM patients using insulin are just as prone to severe hypoglycemia as IDDM patients. However, because in NIDDM there is limited endogenous insulin secretion, sufficient to allow some glucose metabolism, ketoacidosis and hyperglycemic coma are unlikely to develop if exogenous insulin is reduced or discontinued.

Although NIDDM is often misguidedly referred to as "mild diabetes," the risk of chronic complications is no less than that associated with IDDM. People with NIDDM are two or three times as likely to develop cardiovascular disease than people without diabetes and 75% die of cardiovascular disease. The life expectancy of people with NIDDM is, on average, reduced by a decade as a result of the disease. Although the later onset of NIDDM might be expected to reduce the risk of complications which are usually associated with raised blood glucose levels over 15 or more years, the insidious onset of NIDDM delays diagnosis, and chronic hyperglycemia over many years probably accounts for the many cases which are only diagnosed after the onset of complications.

8.11.2 IDDM: ETIOLOGY, RISK FACTORS, AND PREVENTION

IDDM is caused by the autoimmune destruction of insulin-producing beta-cells in the pancreas. The pathological process is probably initiated several years, sometimes even decades, before clinical onset by an interaction of genetic predisposition and environmental triggers (Zimmet, 1995). The incidence of IDDM has been reported to change with age, rising to a peak around puberty (Laakso & Pyörälä, 1985) with a second peak later in life (Krolewski, Warram, Rand, & Kahn, 1987). Altogether, more cases of IDDM are diagnosed in adulthood rather than in childhood (Vandewalle et al, 1993) and, in some studies, a higher incidence has been found in males than in females (Blohmé et al., 1992; Bruno et al., 1993). Marked geographical variation in IDDM incidence has been observed, varying from 35.3 per 100 000 inhabitants in Finland to 1.9 per 100 000 in Japan (Karvonen, Tuomilehto, Libman, & LaPorte, 1993). Differences in incidence between neighboring countries has led to the assumption that diet and lifestyle are important determinants of the disease (Green, Gale, & Patterson, 1992), but differences in the frequency of "diabetogenes" may also play a role. Since the 1970s the incidence of IDDM has doubled in Europe (Bingley & Gale, 1989) and is steadily increasing worldwide (Krolewski et al., 1987). However, most of the incidence data are only available for those with early onset before the age of 20 years. Since most cases are diagnosed during adulthood, it may be that the apparent doubling in incidence is due to a shift towards earlier clinical onset of diabetes rather than to an increase in the number of people developing IDDM at some point in their lives (Bingley & Gale, 1989).

8.11.2.1 Risk Factors in IDDM

A genome-wide search has revealed that more than 12 gene loci contribute to genetic susceptibility for IDDM (Todd, 1995). The genetic component of the condition is indicated by the 5% prevalence of IDDM in first degree relatives, compared with less than 1% in the general population (Bingley, Bonifacio, & Gale, 1993; Gorus, Anselmo, & Keymeulen, 1993), and the higher concordance rate for IDDM in identical than in nonidentical twins (Leslie & Elliott, 1994). Nevertheless, 90% of the new cases occur in individuals without a family history of diabetes (Bingley et al., 1993). These individuals, however, also carry certain predisposing genotypes. Genetic susceptibility is not sufficient to cause diabetes (Karvonen et al., 1993) since a majority of people with disease-associated alleles do not develop IDDM. Environmental agents appear to have an important role. The way in which the environment affects the pathogenesis is still unclear. It has been hypothesized that viral infection (Krolewski et al., 1987; Zimmet, 1995) and proteins in the diet (Gerstein, 1994) are important factors. Psychological stress may also have a role in increasing vulnerability to viral infection or impairing defense mechanisms against such infection, thereby facilitating the progression of the hidden pathological process. Clear evidence for such an effect is difficult to obtain, though suggestive evidence is available to support the notion that stress may contribute to the development of IDDM (see Bradley, 1988 for review). There is stronger evidence to suggest that psychological stress can trigger the onset of symptoms in a person in whom islet beta-cell damage is at an advanced stage (Bradley, 1988; Clayer, Bookless-Pratz, & Harris, 1985). Blood glucose levels already raised as a consequence of diminishing insulin production can be elevated further by neuroendocrine responses to psychological stress and tipped over the renal threshold to precipitate the overt symptoms of diabetes which would have appeared in due course without the occurrence of the stress.

Knowledge about the incidence of IDDM has developed as a result of national and regional registries (Karvonen et al., 1993). Standardized data are collected on previous exposure to environmental factors and occurrence of metabolic and genetic markers for new IDDM cases (Gorus et al., 1993). The long-term follow-up of at-risk families allows comparative studies of the cause and pathogenesis of the disease and can help clarify temporal, regional, and environmental variations in the incidence of diabetes (Diabetes Epidemiology Research International Group, 1988).

8.11.2.2 Prevention of IDDM

There is no cure for IDDM. Despite evidence that improved blood glucose control reduces the risks of complications and despite the development of increasingly effective insulins and more convenient delivery systems and monitoring devices, the complications of diabetes continue to present major health problems with massive personal, social, and economic costs. The possibilities of preventive strategies are therefore attracting considerable attention.

8.11.2.2.1 Primary prevention

The aim of primary prevention is to avoid the initiation of the hidden pathological process leading to IDDM. This strategy requires identification of genetically predisposed individuals and protection against avoidable environmental factors (e.g., nutritional or viral). Because of the lack of specific knowledge about the nature of these triggering factors, large-scale initiatives could lead to treatment of many people who might never develop IDDM (Palmer & McCulloch, 1991). The role of the diabetes registries is crucial in improving understanding of the interaction between genetic and environmental factors. A first attempt at primary prevention has been initiated in Finland (Akerblom et al., 1993), where the role of cows" milk proteins in the diet of young children is under investigation.

8.11.2.2.2 Secondary prevention

The preclinical phase of IDDM is accompanied by humoral, immune, and metabolic abnormalities. Much knowledge has been accumulated since the 1980s about the diabetes-predictive value of several of these early markers (Dotta, Dionisi, Farilla, & DiMario; 1996). Refining the use of these marker combinations would allow early diagnosis of diabetes and offers the possibility of secondary prevention with the aim of halting or delaying the progression of further destruction of the beta-cell mass by pharmacological intervention at an early stage of the pathological process.

The presence of antibodies against various diabetes-associated islet-cell autoantigens is considered to be one of the best indicators of a latent autoimmune disease process in the islets (Dotta et al., 1996), although the relation of these antibodies to the pathological changes is poorly understood. Data obtained in a representative sample recruited through diabetes registries show that at least one of these immunological markers is present in approximately 90% of the individuals at clinical diabetes onset and also in the vast majority of relatives in the preclinical stage (Gorus et al., 1997; Verge et al., 1996). The presence of more than one type of autoantibody greatly increases the risk of IDDM (Eisenbarth, Vergen, Allen, & Rewers, 1993). Metabolic markers can be detected in a later phase of the preclinical stage. A decreased early-insulin response to an intravenous glucose challenge and increased proinsulin levels can indicate impending clinical onset of IDDM. There are indications that combined presence of these biological markers can also be helpful in identifying preclinical IDDM in the general population (Hagopian et al., 1995). The early diagnosis of a hidden disease process creates the theoretical possibility of influencing the process pharmacologically in such a way that the clinical disease manifestation is prevented. Pilot studies in animal models and on a small scale in humans suggest that such an approach is feasible. Large-scale intervention trials have been initiated with nicotinamide and with insulin. It is expected from preliminary studies that nicotinamide (vitamin B3) could delay the onset of IDDM (Elliott & Chase, 1991; Pozzilli & Andreani, 1993), but that the preventive effect could vary from individual to individual (Herskowitz, Jackson, Soeldner, & Eisenbarth, 1989). Multicenter programs were launched to investigate the effect of this vitamin on the pathological process and on the biological markers in first-degree relatives (Nordenfelt, 1996). A second approach has been the prophylactic use of insulin to allow the remaining beta-cell mass to rest or to induce tolerance against beta-cell autoantigens. A clinical trial has begun in the USA (Keller, Eisenbarth, & Jackson, 1993).

8.11.2.2.3 Tertiary prevention

Tertiary prevention aims to prevent long-term complications despite the onset of diabetes. One approach has been pancreatic organ transplantation which has been shown to restore normoglycemia and to halt the progression of chronic complications (Robertson, 1992). However, the intervention is by no means a cure, and the required immunotherapy can cause serious unwanted effects, including increased risk of cancers. For these reasons, solitary pancreas transplantation is not a suitable treatment for diabetes. Pancreas transplantation is, however, sometimes performed in combination with a kidney graft in individuals with severe chronic complications. Since the 1970s, transplantation of insulin-producing islets has been studied as an alternative for pancreatic organ transplantation (Pipeleers, Keymeulen, & Korbutt, 1994). Several clinical trials have been initiated worldwide (Sutherland et al., 1996). Although many obstacles have to be overcome, transplantation of allogeneic (including human to human) or xenogeneic (including pig to human) donor tissue has the potential to treat IDDM in the early stages with a view to avoiding complications (Pipeleers et al., 1994).

8.11.3 NIDDM: ETIOLOGY, RISK FACTORS, AND PREVENTION

8.11.3.1 Risk Factors in NIDDM

NIDDM appears to be a complex multigenic-environmental disorder, with a strong genetic component. Theories about its pathogenesis implicate intrauterine malnutrition (Hales & Barker, 1992) and a common (little understood) metabolic disorder, Syndrome X (Reaven, 1988), which leads in differing circumstances to NIDDM—obesity, hypertension, or hyperuricemia (gout).

Familial clustering is a characteristic feature of NIDDM. Monozygotic twin studies show a concordance rate of over 90% (Newman et al., 1987; Pyke, 1979). The US National Health and

Examination Survey (Harris, 1984) showed that in people aged 35–74, about 35% of those with a medical history of diabetes had one or both parents with diabetes. Other studies have shown an even greater concordance for NIDDM in sibling pairs than parents and children (Beaty, Neel, & Fajans, 1982). The literature on risk of NIDDM in the offspring of affected people has been reviewed (Pierce, Keen, & Bradley, 1995). The reviewers concluded that having a parent with NIDDM increases offspring risk by between two- and fourfold, but a family history of NIDDM does not increase risk of IDDM. NIDDM risk increases if the affected parent is the mother rather than the father, or if both parents are affected, and with increasing age, and in certain racial groups.

NIDDM is more common in those who are overweight, people who take little exercise, and in racial groups other than northern European whites. The association between obesity and risk of NIDDM is strong, graded, and consistent. The literature has been comprehensively reviewed by Felber, Acheson, and Tappy (1992). The distribution of obesity, central or peripheral, has been shown to be a better predictor of later diabetes than the degree of adiposity. People with "central" or "male" type obesity are at much higher risk than those with "peripheral" obesity.

Epidemiological studies have shown a relationship between inactivity and NIDDM (Helmrich, Ragland, Taylor, Leung, & Paffenbarger, 1991; Taylor, Ram, Zimmet, Raper, & Ringrose, 1984), although these studies have been cross-sectional rather than prospective and hence caution is needed in deducing causality.

People with impaired glucose tolerance progress to NIDDM at a rate of between 2–6% per year, largely dependent on race (Jarret, Keen, Fuller, & McCartney, 1979; Saad et al., 1988). However, it is not possible to predict who will progress to NIDDM and who will revert to metabolic normality. Gestational diabetes can be regarded as an extreme form of impaired glucose tolerance in some pregnant women in response to the metabolic strain of pregnancy. After the baby is born, the diabetes may disappear, but it is likely to reappear in later life.

Certain racial groups, including Micronesians (e.g., Nauruans), native Americans (e.g., Pimas) Asian Indians, Mexican Americans, and Australian Aborigines, seem to be particularly prone to NIDDM, a fact which may be due to their genotype (Neel, 1982). Their susceptibility manifests itself when they move from a "traditional" lifestyle to a more "modern" one with increased availability of food and with reduction in the demand for physical exertion (Zimmet, 1992). Possibly, the genetic make-up of these people is adapted to survival in the face of variable supplies of food. This "thrifty genotype" (Neel, 1982) lays down fat stores very quickly when food is available and so facilitates survival during periods of famine. Such a genetic make-up would have clear evolutionary advantages in a "traditional" lifestyle such as that of the Australian Aboriginal hunter-gatherers. However, when these people move to a lifestyle characterized by a continuously plentiful food supply, they readily become overweight and develop NIDDM. Alternatively Hales and Barker (1992) have hypothesized that infants who are malnourished *in utero* develop physical changes in their pancreas and other organs, which in later life cause a decreased ability to cope with excess calories and so these people become overweight and develop NIDDM. This "thrifty phenotype" hypothesis could also account for these racial groups being particularly prone to NIDDM in the face of westernization.

8.11.3.2 Prevention of NIDDM

8.11.3.2.1 Primary prevention

There have been several studies testing the possibility of primary prevention of NIDDM using either pharmaceutical agents or behavioral methods (Erikksom & Lindegarde, 1991; Jarret, et al, 1979; Keen, Jarret, & McCartney, 1982; Knowler, Sartor, & Schersten, 1987; Sartor, Schersten, Carlstrom, Melander, & Persson, 1980). These have given somewhat contradictory results. However, Pan et al. (1994) reported the results of a randomized controlled trial of diet and exercise alone, or in combination, suggesting that all of these interventions reduced the rate of progression to NIDDM in people with impaired glucose tolerance in mainland China. In the light of these findings, a large multicenter primary prevention trial is being planned in the USA. The interventions are likely to include hypoglycemic drugs and/or behavioral methods aimed at the reduction of obesity and the promotion of exercise. The challenges of designing such a trial are considerable as potential recruits are likely to have preferences for drug or lifestyle intervention which will make a simple randomized controlled trial design inappropriate (Bradley, 1993).

8.11.3.2.2 Secondary prevention

The secondary prevention of NIDDM is under investigation by the UK Prospective

Diabetes Study Group (UKPDS) who are investigating the effects of early treatment regimens for NIDDM (UKPDS VIII, 1991).

8.11.3.2.3 Tertiary prevention

Tertiary prevention of NIDDM involves the detection and treatment of complications of NIDDM, and the reduction of associated risk factors. Retinopathy (American Diabetes Association, 1992), neuropathy, and early signs of nephropathy (Viberti, Walker, & Pinto, 1992) can be detected by annual screening. Retinopathy can usually be treated effectively by laser therapy to prevent progression to blindness if detected in the early stages (Diabetic Retinopathy Study Group, 1981). Treatment of established neuropathy has been described by the St. Vincent Guidelines as "unsatisfactory" (Krans, Porta, Keen, & Staehr Johansen, 1995), though there is evidence from the DCCT that improved blood glucose control reduces risk of nerve damage (DCCT Research Group, 1993). There has been considerable interest in the possibilities of using drugs, particularly aldose reductase inhibitors, as neuropathy prevention therapy, though their value is far from clear (Krans et al., 1995). The St. Vincent Guidelines suggest that the best hope for preventing the development of neuropathic foot problems is through the education of patients and professionals and the organization of foot care services (Krans et al.). Incipient nephropathy can be detected by testing the urine for small amounts of albumin (microalbuminuria), and at this stage the process is potentially reversible by improved management of blood pressure and metabolic control. Avoidance of smoking (Mulhauser, 1990) and reduction of any modifiable risk factors for cardiovascular disease is particularly important for people with NIDDM, because of their elevated risk of myocardial infarction and stroke.

As with IDDM, psychological stress may trigger the manifestation of symptoms in NIDDM, probably by a similar mechanism (Bradley, 1988; Clayer et al., 1985). Evidence for a causal role of stress in initial onset of the pathological process is again less clear. Mechanisms by which stressful events may play a causal role will vary depending on the kind of NIDDM where, unlike IDDM, there is not always a reduction in insulin production. Indeed in individuals with NIDDM who are not overweight, insulin resistance may be involved where insulin production may increase as its effectiveness decreases. There is some evidence to suggest that stressful events may be causal in the onset of insulin resistance with stress-induced increases in adrenaline and cortisol secretion reducing the effectiveness of an individual's insulin (Rizza, Mandarino, & Gerich, 1982; Smith, 1984). If repeated or prolonged stressful events are experienced, insulin resistance over a period of months or years may result in excessive compensatory insulin production and, eventually, to exhaustion of pancreatic islet cells and manifestation of diabetes.

8.11.4 IMPACT OF THE CONDITION ACROSS THE LIFE SPAN ON PERSON AND FAMILY

With increasing knowledge of the familial risks of diabetes, diabetes may have a psychological impact on the outlook on life of an at-risk individual well before onset or without the actual onset necessarily occurring. People with diabetes may be concerned about the possibilities of onset of diabetes in their children. Indeed, the fact that a proportion of parents of at-risk children are prepared to enter them in clinical trials of the use of daily insulin injections in the hope of preventing diabetes onset (Keller et al, 1993) provides an indication of the level of concern and lengths to which some are prepared to go in order to reduce the risk of diabetes onset.

When diabetes is diagnosed, the demands of diabetes and its treatment vary considerably with the type of diabetes (IDDM usually being more demanding in terms of insulin treatment required and the increased need for monitoring of blood glucose levels). Within any one type of diabetes there is considerable individual variation in the effort needed and difficulty experienced in achieving blood glucose levels which approximate normoglycemia. The need to manage diabetes can add to the psychological demands of meeting the commonly experienced challenges throughout the lifecycle, such as leaving home, marriage, and pregnancy. Pregnancy, and preparing for pregnancy, are especially demanding times for women with diabetes when the need for normoglycemia is emphasized as essential for the health of the baby and recommended prior to conception through to birth. The success which many women achieve in obtaining near-normal blood glucose levels at this time is sometimes attributed by their physicians to the hormonal changes of pregnancy rather than to the efforts of the woman herself. Such attributions can be demotivating and depressing. The challenges of a first pregnancy are, however, dwarfed by the challenges of a second pregnancy where the needs of the first child are often incompatible with the needs of the developing second child.

Pregnant women with diabetes attempting to achieve normoglycemia are particularly prone to hypoglycemia and this risk carries new dangers when the pregnant woman is also in sole charge of a young child. Fear of hypoglycemia, a concern to most people with insulin-requiring diabetes (Irvine, Cox, & Gonder-Frederick, 1994), is likely to be heightened in pregnant women, especially those with young dependent children.

Many people with insulin-requiring diabetes are sufficiently fearful of hypoglycemic reactions to avoid tight control of blood glucose levels and the associated increase in risk of hypoglycemia, especially if their warning signs of hypoglycemia are unreliable (see Sections 8.11.5. and 8.11.7.1). Fear of hypoglycemia may lead individuals to aim for higher blood glucose levels which increase their risk of long-term complications. Others, anxious to reduce their risks of other complications may be experiencing numerous, potentially dangerous, hypoglycemic episodes.

When NIDDM patients are unable to achieve adequate diabetes control with tablets and diet, it is usually recommended that insulin be introduced. Indeed the prospect of insulin injections is often used by doctors as a stick to motivate greater dietary efforts and more reliable use of sulfonylurea tablets. As a result, the need for insulin may be taken as a sign of failure or a sign of progression from "mild" to increasingly "severe" diabetes. Many patients also fear the process of injection, prior to the experience. In practice, the injections are usually found to be an obstacle soon overcome and incorporated into daily life with little difficulty. Injections allow greater flexibility of timing and content of meals as well as improved blood glucose control. Improved satisfaction with treatment following changeover from tablets to insulin has been found in several studies (Jennings et al., 1991; Taylor, Foster, Kyne-Grzebalski, & Vanderpump, 1994). Such findings could be used by health professionals to inspire their patients who are reluctant to change. Interestingly, the aspect of quality of life widely felt by people with diabetes to be impaired by their diabetes and rated as very important, is enjoyment of food (Bradley, Todd, Gorton, Plowright, & Symonds, 1995). After decades of attempting to persuade patients to modify the timing and content of their food intake to fit in with fixed insulin doses, and to use food manipulation to compensate for changes in energy expenditure, there is growing recognition that a more rewarding approach is to manipulate insulin doses to fit in with food and other lifestyle factors (Howorka, 1996).

8.11.5 PSYCHOLOGICAL ASPECTS OF SYMPTOMS

Individual variation in the nature and extent of symptoms of hypoglycemia is considerable. Symptoms may include autonomic symptoms such as irritability or anxiety, shaking, or sweating associated with adrenaline released in response to the falling blood glucose level. Symptoms may also include neuroglycopenic symptoms such as loss of concentration or memory loss, due to the shortage of glucose in the brain (Deary, Hepburn, Macleod, & Frier, 1993). Hunger is also a common, though by no means universal symptom.

Individuals prone to hypoglycemia include anyone taking insulin and individuals taking certain kinds of oral hypoglycemic agent (particularly sulfonylureas). Individuals usually come to recognize the idiosyncratic but fairly reliable symptoms which for them herald the onset of hypoglycemia, and learn to take action to raise the blood glucose levels before they are rendered incapable of doing so by increasingly severe symptoms which impair rational action and precede the unconscious state of hypoglycemic coma. However, some individuals lose the ability to recognize the symptoms of hypoglycemia and become highly vulnerable to severe hypoglycemia as a result. Loss of warning of hypoglycemia, or hypoglycemia unawareness, may result from autonomic neuropathy, one of the microvascular complications of diabetes, which becomes increasingly common with duration of diabetes of 15 or more years. Damage to the autonomic nervous system impairs adrenal responses to low blood glucose levels. In the absence of adrenal symptoms, neuroglycopenic symptoms may not be noticed until the cognitive effects are too debilitating to take action to raise the blood glucose levels. However, onset of hypoglycemic unawareness is not always associated with autonomic neuropathy and there has been much concern among people with diabetes, clinicians, and researchers to identify the causes of loss of warning with a view to preventing this loss or restoring awareness.

It has commonly been assumed that hypoglycemic unawareness develops when symptoms of hypoglycemia occur below the level of blood glucose at which cognitive impairment begins. It is also widely assumed that the detection of hypoglycemia is primarily dependent on autonomic symptoms. However, Cox, Gonder-Frederick, Anderson et al. (1994) have argued that neuroglycopenic symptoms, such as difficulty in concentrating and lack of coordination, are as prevalent as autonomic symptoms and occur at similar levels of blood glucose. Furthermore, people with IDDM have reported using neuroglycopenic symptoms to recognize hypoglycemia, sometimes in preference to autonomic symptoms. Cox et al. also cited evidence to suggest that an episode of hypoglycemia leads to a reduction in strength of autonomic but not neuroglycopenic symptoms associated with any subsequent hypoglycemia within 24 hours. In line with these findings are reports from elsewhere (Cranston, Lomas, Maran, Macdonald, & Amiel, 1994) that neuroglycopenic symptoms provided earlier warning of impending hypoglycemia than autonomic symptoms and may indeed be more useful and more reliable cues to low blood glucose than has generally been assumed. A Blood Glucose Awareness Training program has been developed (Cox, Gonder-Frederick, Julian, & Clarke, 1992) which has been used successfully to regain lost awareness of hypoglycemia, and Amiel's research group at the University of London (Cranston et al., 1994) and others (Fanelli et al., 1993) have also reported success using more traditional medical approaches. These interventional approaches are reviewed in Section 8.11.7.1.

In the late 1980s there was a nationwide move in the UK to switch from animal insulins extracted from cattle and pigs to the genetically engineered "human" insulin. Most clinicians appear to have assumed that the changeover would have little effect on blood glucose levels. However, many people with insulin-treated diabetes were profoundly upset by the changeover as indicated by the influx of letters to the British Diabetic Association (BDA; Posner, 1992). Problems of rapid onset of severe hypoglycemia occurring without warning were reported, some with devastating and even fatal consequences. Although undoubtedly some of the problems were coincidental to the changeover to human insulin, it is by no means clear that all the problems would have occurred anyway (Alexander, 1993). The problems were further exacerbated by an unwillingness of some doctors to accept the possibility that human insulin may have caused the problems, and their refusal to accede to their patients" requests to change back to animal insulin. Many patients continue to suffer from the aftermath of the changeover to human insulin, since new developments in insulin delivery systems are usually only developed for use with human insulin and thus are not available to those who changed back to animal insulins. The BDA continues to campaign for the provision of animal-insulin alternatives to the now standard human-insulin cartridges used in pen-injector devices, so that patients using animal insulins may also benefit from the convenience of using these devices (North, 1991).

A major reason for anxiety about hypoglycemia is the increasing concern that it may lead to chronic forms of cognitive dysfunction. It has long been recognized that severe hypoglycemia involving coma can cause chronic impairments of cognitive function. Attention has been directed, however, to the chronic effects of less severe, but recurrent hypoglycemia (Gold et al., 1994). It is difficult to demonstrate a causal association but the available evidence (including uncontrolled case reports and retrospective studies) does suggest that repeated and severe hypoglycemia may lead to chronic cognitive dysfunction. Ryan (1994) has provided a useful overview of work assessing cognitive impairments related to hypo- and hyperglycemia which assesses the pros and cons of various measures of cognitive function in determining the nature and extent of cognitive impairments associated with different levels of blood glucose.

In contrast to the marked and troublesome nature of symptoms of hypoglycemia, hyperglycemia is quite often asymptomatic and this in itself creates challenges. Indeed, some individuals who have adapted to raised blood glucose levels feel uncomfortable when initially reducing their blood glucose levels. However, adaptation to the experience of blood glucose levels closer to normal levels soon follows. Resistance to tightening control of blood glucose levels may be reduced if such patients appreciate that any discomfort is likely to be temporary.

Most of the early signs of chronic complications of diabetes are also asymptomatic. High blood pressure, microalbuminuria indicative of early signs of kidney problems, and early signs of retinal damage are usually, if not always, without symptoms, and require medical screening to allow for early diagnosis and treatment. Peripheral neuropathy causes numbness in the extremities and a real risk that damage to the feet will go unnoticed. Infections and ulcers, once developed, can be difficult to clear up in the face of circulatory problems which usually accompany the neuropathy and may lead to gangrene and the subsequent need for amputation. These serious problems depend for their solution on patient education, vigilant self-care, and rapid access to skilled chiropody and medical attention when minor damage to the feet is first noticed.

8.11.6 SPECIFIC ISSUES OF ADHERENCE

Adherence, or compliance, has been defined as the extent to which a person's behavior coincides with health advice (Haynes, Wang, Gomes, 1987). The concept implies that specific recommendations can be made which, if followed, are likely to lead to health benefits. Certainly, the measurement of adherence requires that the recommendations have to be specified, and then the extent to which those recommendations have been implemented needs to be assessed. Such measurement may be practical if considering whether or not a patient had been adherent in taking a course of antibiotics. In this case, the recommendations can be specified in detail and the behavior monitored in various ways to establish the extent of match between recommendations and behavior. Where diabetes is concerned, potentially effective recommendations are likely to be complex, to vary from patient to patient and, for individual patients, to vary from day to day. "Simple" recommendations to take the same amount and types of insulin at the same time each day and eat the same amount and type of food at the same times to balance that insulin while somehow keeping energy expenditure constant are unlikely to be effective in achieving adequate diabetes control. If such recommendations did lead to improved metabolic control they would be likely to cause serious damage to patients" psychological well-being and quality of life. More complex regimens involve balancing the various manipulable components in the equation (e.g., timing and dose of the next injection of insulin, timing and content of the next meal or snack, and voluntary exercise) to deal with those components of the equation which are either not immediately controlable or cannot be controlled without undesirable costs (e.g., family or work commitments, action of the previous insulin injection). Such regimens require sufficient education and understanding to allow problem-solving appropriate to the occasions that arise and cannot be specified in detail in advance. "Adherence" under these circumstances is an unmeasureable and unsuitable concept.

Some authors have referred to measures of diabetes control as "proxy measures of adherence." Such a practice is highly misleading as it assumes that if the patient had followed whatever recommendations had been made then the diabetes control would be maximal. Use of HbA_1 or other measures of diabetes control as a measure of adherence denies the possibility that the recommendations made by the health professional may have been inadequate to the task of controling that individual's diabetes. There is no good reason to attribute all departures from optimal outcomes to patient "nonadherence" and there are many good reasons for not doing so. In particular, it is depressing and discouraging for patients who are trying to deal with the demands of diabetes management to have poor outcomes attributed to their actions (or lack of them), particularly when they have been struggling to follow recommendations which were not adequate to the task. The common tendency to take credit for successes and blame other factors for failure is one that affects health professionals as much as other groups (Gamsu & Bradley, 1987; Gillespie & Bradley, 1988) and it undoubtedly serves to protect their own self-esteem. However, such an ego-defensive stance is not only discouraging for patients and their families but also serves to reduce the perceived need for interventions to improve the organization and functioning of the health-care system (Bradley, 1989). It would be more supportive of patients" self-esteem to recognize patients" efforts and achievements and to adopt a collaborative problem-solving approach when outcomes need to be improved.

There are many different ways of managing diabetes. If the patient is not able to achieve satisfactory diabetes control with one approach, it may be that some aspect of the approach is deficient in some way. The education provided for the patient may not have resulted in the necessary knowledge and understanding required. Barriers to care may be obstructing application of adequate education and may include social pressures and competing demands, fatalistic or unrealistically optimistic beliefs about outcomes of diabetes and the risk of complications (Glasgow, 1994; Lewis & Bradley, 1994). Knowledge, understanding, and beliefs may all be compatible with self-care but the recommendations made may create conflicts with other priorities in the individual's life which may best be dealt with by changing the recommendations made without necessarily compromising the level of diabetes control that can be achieved.

Thus, in the complex field of diabetes management, the notion of self-care behavior is far more useful than the concept of "adherence." It can be useful and constructive to investigate self-care behavior and to explore the reasons for actions taken and consider the extent to which the actions were likely to result in the intended outcomes (and avoid unwanted outcomes) given the circumstances concerned. While it is neither necessary nor helpful to ask whether the actions were "adherent" it can be useful to assess self-care activities and efforts have been made to do so (Toobert & Glasgow, 1994).

A useful review of the complexities in attempting to apply the concept of "compliance" to diabetes management is provided by Glasgow (1991).

8.11.7 PSYCHOLOGICAL INTERVENTIONS IN THE MANAGEMENT OF DIABETES

8.11.7.1 Blood Glucose Awareness Training and Other Approaches to Regaining Hypoglycemia Awareness

A program has been developed at the University of Virginia of Blood Glucose Awareness Training (BGAT) to train individuals with diabetes to recognize their level of blood glucose. The BGAT program and associated manual (Cox, et al., 1992) is based on blood glucose monitoring together with monitoring of physical symptoms, mood, and environmental cues associated with each blood glucose reading in order to identify the symptoms and cues which predict high blood glucose levels and those which predict low blood glucose levels for each individual. An error grid analysis was developed to identify those errors made in estimating blood glucose which matter clinically. Efforts are focused on reducing the errors that would be clinically important, such as failing to detect extreme blood glucose levels, or mistaking an acceptable blood glucose level for one that is too high or too low.

BGAT has been shown to be effective in teaching individuals with IDDM to improve their accuracy in recognizing blood glucose levels. Associated improvements in diabetes control have also been found with improved glycosylated hemoglobin results in BGAT patients compared with untrained control patients (Cox et al., 1991). Long-term follow-up has shown that participants in the BGAT training program had had fewer episodes of severe hypoglycemia and fewer car crashes than controls at five-year follow-up (Cox, Gonder-Frederick, Julian, & Clarke, 1994). More recently, BGAT-2 has been developed to increase sensitivity to low blood glucose levels, and benefits have included fewer motoring offenses and car accidents, and fewer nocturnal episodes of severe hypoglycemia (Cox, Gonder-Frederick, Kovatchev, et al., 1995). A multicenter trial of BGAT-2 used by clinicians other than those in the original research team developing the program, suggested that BGAT-2 offers a useful strategy for treating the disabling problem of hypoglycemia unawareness. BGAT-2 improved detection of low blood glucose levels in individuals who had reported impaired hypoglycemia awareness on recruitment to the study (Cox, Gonder-Frederick, Polonsky, et al., 1995). There is increasing evidence that patients who have lost their warning symptoms of hypoglycemia can learn to attend to new signs of hypoglycemia which may be physical, cognitive, or mood symptoms, or may be circumstantial indicators that hypoglycemia is likely (e.g., running for a bus when late for a meal).

Another approach to treating loss of awareness of hypoglycemia has been to avoid hypoglycemia by diet review, advice about exercise, and redistribution of insulin in patients with hypoglycemic unawareness (Cranston, et al., 1994; Fanelli et al., 1993). After at least three clear weeks without hypoglycemia, patients became aware of symptoms of hypoglycemia. Adrenaline responses were more marked at higher blood glucose levels than was previously the case, and at levels higher than those at which cognitive impairment was indicated by a reaction-time task (Cranston et al.). It has been suggested (Bradley & Gamsu, 1995) that, in attempting to avoid hypoglycemia, participants in such studies may learn to become more sensitive to neuroglycopenic symptoms at higher levels of blood glucose than they were before. Thus, the intervention to avoid hypoglycemia may act as an indirect form of BGAT. It is also possible that the success of BGAT in increasing awareness of hypoglycemic symptoms may in part be due to a reduction in occurrence of hypoglycemia.

8.11.7.2 Identification of Stress Reactivity

8.11.7.2.1 Mechanisms involved in stress and diabetes control

Stress may affect metabolic control in diabetes in at least two ways (Barglow, Hatcher, Edidin, & Sloan-Rossiter, 1984).

(i) A direct psychophysiological effect via sympathetic and pituitary activity which results in the elevation of catabolic hormone levels and the suppression of anabolic hormones. In people with diabetes, this may result in increased blood glucose levels, although, for a small minority, less readily understood decreases in blood glucose levels result.

(ii) A behavioral mechanism whereby stress leads to behavioral changes capable of disrupting self-care behavior (Barglow et al., 1984). For example, time urgency may make blood glucose monitoring and well-balanced meals impractical leading to disruptions in metabolic control. Glycemic fluctuations themselves can contribute to the behavioral changes via impaired cognitive function (Holmes, Hayford, Gonzalez, & Weydert, 1983; Ryan, 1994; Wang et al., 1993), which, for example, may cause poor self-care responses to feedback from blood glucose monitoring. A two-way causal link between stress and diabetes control has been suggested (Bradley, 1988) whereby life events cause disruption in diabetes control which in turn causes increases in the number of life events.

8.11.7.2.2 Acute stress and blood glucose reactivity

Several seemingly well-controlled acute stress studies conducted in the 1980s reported no significant changes in blood glucose control in response to potential stressors such as mental arithmetic and public speaking (Edwards & Yates, 1985; Kemmer et al., 1986; Naliboff, Cohen, & Sowers, 1985). However, this work overlooked the possibility that the commonly observed individual differences in response to stress might be real and interesting, and not a reflection of methodological inadequacies (Bradley, 1988). Studies of experimental stress have examined physiological mechanisms hypothesized to mediate the relationship between psychological stress and blood glucose control. There is evidence that reduced blood flow to the insulin injection site (Hildebrandt, Mehlsen, Sestoft, & Nielson, 1985) and insulin resistance over several hours (Moberg, Kollind, Lins, & Adamson, 1994) may cause increased blood glucose levels in individuals with IDDM in response to acute laboratory stressors. It has been noted elsewhere (Bradley & Gamsu, 1995) that laboratory studies with a short time frame are likely to underestimate such long-lasting effects of stress on blood glucose.

The effects of a laboratory stressor (Stroop test) on changes in blood flow at the insulin injection site have been studied in IDDM patients (Greenhalgh, Jones, Jackson, Smith, & Yudkin, 1992). The results showed that mean levels of blood glucose rose in some patients and fell in others in response to acute stress, and that these changes were largely explicable in terms of changes in free insulin levels, which in turn were partially mediated by changes in injection-site blood flow. The authors proposed that vasodilation at the subcutaneous insulin injection site may in some cases lead to a paradoxical hypoglycemic effect during acute stress via an increased rate of absorption of insulin, an effect which will be counterbalanced to a greater or lesser degree by increases in counter-regulatory hormones. In other individuals, absorption of insulin may fall during stress and contribute to hyperglycemia. Again these findings underline the importance of looking at individual differences.

Idiosyncratic blood glucose responses which were reliable across a 12-week time period within individuals with IDDM have been reported in response to caffeine and to the competitive playing of a video game (Carter, Gonder-Frederick, Cox, Clarke, & Scott, 1985). A more recent study (Gonder-Frederick, Carter, Cox, & Clarke, 1990) has reported similar results. Such studies suggest that the inconsistent findings of previous research may have resulted because of differences in the stress-responsiveness of the individual patients recruited.

8.11.7.2.3 Life stress and blood glucose reactivity

Studies of major life events and diabetes control have suggested that increased life events are associated with raised blood glucose levels (Barglow et al., 1994; Chase & Jackson, 1981). However, within-subject variations in stress and in blood glucose values would have been obscured by the methods used which correlated aggregated retrospective reports of life events with GHb or HbA_1 measures of averaged blood glucose, thereby oversimplifying the picture. Research has begun to look at the relationship between minor daily events and diabetes control (using serial blood glucose measurements) in order to overcome the methodological limitations of the earlier studies of relationships between life events and blood glucose levels.

Investigations of daily stressful events are providing more interpretable results with studies reflecting individual differences in response to stress that were also seen in some of the laboratory based acute stress studies. Although the overall picture suggests that increased daily stress correlates with increased blood glucose levels in IDDM (Hanson & Pichert, 1986) and in NIDDM (Goetsch, Abel, & Pope, 1994), examination of individual's blood glucose response to stress indicates that some people display stress-reactivity and some people do not. Halford, Cuddihy, and Mortimer (1990) found that approximately half the sample of 15 IDDM participants had significant associations between stress and blood glucose levels, and that this association was independent of the effects of diet and exercise self-management.

Within those people who do display stress-reactivity, most show an increase in blood glucose while a few show a decrease in blood glucose (Aikens, Wallander, Bell & McNorton, 1994; Riazi, Pickup, & Bradley, 1996). Aikens et al. studied twenty-five women with IDDM who monitored daily stress and blood glucose for 30 consecutive days. Blood glucose was found to be higher on high-stress days than on low-stress days, with eight of the 25 subjects showing significant positive associations between stress and same-day blood glucose. Stress showed little relation to next-day blood glucose, although two subjects showed associations (one positive and one negative). In a study of 54 men and women with IDDM which measured daily stress and blood glucose levels over a three week period (Riazi et al., 1996), individual differences were marked in the magnitude, direction, and duration of blood glucose stress reactivity. Those with higher levels of HbA_1, indicating poorer control of diabetes, were more stress-reactive than those with lower levels of HbA_1.

These findings show the individual differences in response to stress that can be seen in laboratory-based acute stress studies which investigate individual differences but are not seen in studies which limited the analysis to group differences or in studies of life events using average measures of blood glucose. Although increases in blood glucose levels are the most common form of stress reactivity, decreases are not uncommon and there is a need for individuals to discover empirically how their own blood glucose levels respond to different kinds of stress in order to take appropriate action to prevent or correct disruptions in diabetes control.

8.11.7.3 Stress Management Training

The overly simple model of life stress causing raised blood glucose levels has inspired the use of stress management training, particularly relaxation training, as an aid to diabetes control. Relaxation is thought to decrease adrenocortical activity (Jevning, Wilson, & Davidson, 1978; DeGood & Redgate, 1982) as well as circulating levels of catecholamines (Mathew, Ho, Kralik, Taylor, & Claghorn, 1980; Mathew, Ho, Kralik, Taylor, Semchuk, et al., 1980) and hence is expected to prevent stress-induced increases in blood glucose levels.

Research supports the view that stress management techniques may be valuable to aid diabetes management for some people but not with others. It has been suggested (Bradley, 1994a) that relaxation techniques are unlikely to do harm except when blood glucose is already tightly controlled (Seeburg & DeBoer, 1980) and the insulin dosage is not reduced appropriately to balance the effects of relaxation on insulin requirements (which may well be reduced) and/or when used at a time when blood glucose is already low (< 4 mmol/L) and there is a risk of hypoglycemia. It is important to take the precaution of measuring blood glucose immediately before all relaxation training or practice sessions. Twenty minutes of relaxation can lead to a drop in blood glucose as great as 3 mmol/L. Such a substantial fall is only likely if blood glucose is well above the normal range of 4–6 mmol/L to start with and probably results from a suppression of catecholamine secretion. It is recommended that relaxation training should not be conducted at blood glucose levels below 4 mmol/L because of the risk of hypoglycemia. Relaxation training might be continued if the anticipated reduction in blood glucose were counteracted with pretraining intake of slow release carbohydrate (e.g., an apple). A further precaution of measuring the blood glucose after a relaxation session, before leaving the therapist's office, is also recommended to avoid the risk of hypoglycemia while traveling home.

Several studies have reported some success with improvements in diabetes control following relaxation training (usually involving electromyograph (EMG) biofeedback) in NIDDM and IDDM patients (McGrady, Bailey, & Good, 1991; McGrady & Gerstenmaier, 1990; Rosenbaum, 1983; Surwit, Feinglos, & Hasted, 1983). There is evidence to suggest that relaxation training is least useful for those subjects whose glycemic control was good to start with and most useful when used by subjects who not only had poor control of their diabetes but who felt that stress disrupted their diabetes control and who were currently experiencing stressful events (Bradley, Moses, Gamsu, Knight, & Ward, 1985; Lammers, Naliboff, & Straatmeyer, 1984).

In contrast, Feinglos, Hastedt, and Surwit (1987) found no significant effect of relaxation training in people with poorly controlled IDDM who reported stress-induced hyperglycemia. After six weeks in which the trial group practiced relaxation techniques at home, measures of glucose tolerance, GHb levels (which reflect blood glucose levels over 6–8 weeks), and insulin requirements did not differ between relaxation and control groups. Unfortunately, the authors did not provide any information about the variability of the individual patients.

Lane, McCaskill, Ross, Feinglos, and Surwit (1993) investigated 38 people with NIDDM who were treated with intensive conventional diabetes therapy (including diabetes education and dietary interventions) with half of the subjects assigned at random to receive biofeedback-assisted relaxation training. Both groups demonstrated significant improvements in GHb level after eight weeks, but not in glucose tolerance. No added benefit of relaxation training was seen on either measure though it was found that those subjects who showed improvements in glucose tolerance after pretreatment with the anxiolytic drug alprazolam also showed improvements in glucose tolerance after relaxation training. It was also found that subjects who responded best to relaxation training were those scoring higher on measures of trait anxiety. Overall, the authors concluded that relaxation may be useful for the treatment of neurotic, anxious, and autonomically reactive individuals but may be ineffective for other patients. Lack of selection of patients likely to benefit from relaxation training may have been responsible for the nonsignificant finding of relaxation of the group as a whole.

What is emerging from the studies of relaxation training on people with diabetes is that it is most useful for those who show stress-related disturbance of blood glucose control or for those who are anxious and autonomically reactive. The degree to which these groups overlap is unclear though there is some evidence to suggest that the groups may be separate. Self-reported anxiety did not correlate with blood-glucose reactivity to stress in the studies of Riazi et al. (1996). Assessment of subjects" stress-reactivity before embarking on relaxation training in future studies will help in the identification of patients most likely to benefit from relaxation training. Most studies have incorporated biofeedback into the relaxation training. Relaxation training on its own would be more practical and less costly. There is limited evidence that biofeedback is an unnecessary addition for most people (Bradley et al., 1985). Evaluation of low-technology forms of relaxation training with patients shown to be stress reactive would be a useful next step.

8.11.7.4 Monitoring of Psychological Outcomes and Processes

Just as measurement of blood glucose is essential to the task of achieving and maintaining tight control of diabetes, the measurement of psychological outcomes is central to protecting and improving the quality of life of people with diabetes. However, while measurement of blood glucose is standard practice in diabetes management and research, measurement of psychological outcomes is common only in research settings. Quality of life measurements are increasingly recognized as essential parameters in clinical trials evaluating new treatments for diabetes and improvements in aspects of quality of life are often among the primary outcomes to be investigated. Treatments, including insulins and insulin delivery systems, are being developed to be more convenient and flexible and compatible with quality of life. There is growing recognition that the desired biomedical outcomes are more likely to follow if treatment regimens are convenient and flexible and can be integrated into individuals" lives without impairing the quality of life. Although measurement of psychological outcomes in routine clinical practice is far from standard practice, several initiatives are encouraging such developments. The St. Vincent Declaration Action Programme for diabetes care in Europe has recommended the routine measurement of psychological well-being and patients" satisfaction with treatment in the monitoring of diabetes management (Bradley & Gamsu, 1994; Krans et al., 1995), and many national initiatives are underway, including a national study of monitoring of diabetes care in the UK (Wilson et al., 1993), and an initiative in The Netherlands to evaluate the effects of routine monitoring of the psychological outcomes recommended by the St. Vincent Guidelines (F.J. Snoek and colleagues, personal communication, 1996).

8.11.7.4.1 Measurement of psychological well-being and patient satisfaction

The measures recommended in the St. Vincent Guidelines are the Well-being Questionnaire (Bradley, 1994b; Bradley & Lewis, 1990) and the Diabetes Treatment Satisfaction Questionnaire (DTSQ; Bradley, 1994c) which were designed and developed specifically for people with diabetes. The Well-being Questionnaire was first designed in the early 1980s for diabetes clinical trials because it was recognized that generic instruments to measure depression or anxiety are unsuitable for people with diabetes. Measures such as the Beck Depression Inventory (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), developed for the general population, include items concerned with somatic symptoms which may well be indicative of depression in the population as a whole but in a diabetic sample are more likely to result from inadequate blood glucose control. For example, symptoms of fatigue, appetite disturbances, weight loss, sleep disturbances, irritability, and loss of sexual interest can all be symptoms of high blood glucose levels. At around the same time and for the same reasons, Zigmond and Snaith (1983) developed the Hospital Anxiety and Depression (HAD) scale to focus on cognitive symptoms and avoid somatic symptoms. The HAD was intended to be appropriate for use with hospital patients and, encouragingly, work with diabetic populations has confirmed the original factor structure of the HAD (McColl et al., 1995).

The Well-being Questionnaire has subscales to measure depression and anxiety which are similar to those of the HAD but, unlike the HAD, also has subscales to measure energy and positive well-being. These additional subscales are more sensitive to change than depression and anxiety (Bradley, 1994b) and the Well-being Questionnaire is likely to be more sensitive than the HAD to positive benefits of new treatments for diabetes. If only depression and anxiety scales are used, the optimal state measurable is an absence of depression or anxiety and floor effects occur if, as usually happens, in diabetes research, patients are not depressed or anxious when baseline measures are taken. WHO has recommended wider use of the Well-being Questionnaire with other patient groups in addition to those with diabetes, in a context of continuing development (WHO, 1993, 1995). As with the HAD scale, care is needed when using the instrument with a different population of patients as, despite efforts to reduce confounding of depression scores with symptoms of poorly controlled diabetes, different illnesses may influence responses in different ways. Efforts to improve and evaluate translations of the Well-being Questionnaire in European and other languages have been associated with the widespread use of the instrument in multinational clinical trials and in the St. Vincent European monitoring initiative and work is underway to meet the demand for a shortened version of the 22-item Well-being Questionnaire which will facilitate use of the instrument in routine clinical practice (Bradley, 1996a).

The DTSQ (Bradley, 1994c; Bradley & Gamsu, 1994; Bradley & Lewis, 1990; Krans et al., 1995) is a single-page instrument that has been used widely in clinical trials and is valued for its sensitivity to change in response to alterations to diabetes treatment regimens. A broader, though less detailed, single-page measure of satisfaction with the diabetes clinic service (The Diabetes Clinic Satisfaction Questionnaire [DCSQ]) has since been developed to meet the need to audit patients" satisfaction with the service as a whole as well as, more specifically, with their particular treatment regimen (Wilson et al., 1993). The DCSQ is designed to identify sources of dissatisfaction so that efforts to improve the service can be well informed. This instrument is being used in the UK national study of audit of diabetes care steered by the Diabetes Information Managment and Audit Committee of the British Diabetic Association (Wilson et al.).

The use of diabetes-specific instruments such as the Well-being Questionnaire, DTSQ, and DCSQ allows for instruments which focus on issues of particular relevance to the diabetic population. The instruments can be short because issues of little concern in diabetes can be excluded. Although generic instruments are less likely than diabetes-specific instruments to be sensitive to changes following changes in treatment for diabetes, there is some demand for generic instruments which comes particularly from health service managers and health economists who want to compare outcomes in different patient groups. The reader is referred elsewhere for a review which considers the use of generic measures with diabetic populations (Bradley, 1996b).

8.11.7.4.2 Measurement of quality of life

The phrase "quality of life measures" is often indiscriminately used to cover health status measures, well-being and satisfaction measures, and even measures of self-reported symptoms as well as measures of how good or bad life is felt to be. Focusing here on the latter group of measures, an important distinction can be made between individualized measures and those which are not tailored to the individual.

With most measures of quality of life, diabetes-specific or generic, the aspects of life deemed to be relevant to quality of life are specified by the designers of the questionnaire. Although the choice of aspects of life may have been informed by prior research with samples of people with diabetes and by input from health professionals involved in diabetes care, there is no scope for the individual respondent to decide if any one particular aspect of life asked about is applicable to them and, if it is, whether it is important to their quality of life. The Diabetes Quality of Life (DQOL) measure used in the Diabetes Control and Complications Trial includes 56 items which all receive equal weight regardless of their relative importance to the individual (Jacobson & the DCCT Research Group, 1994). It has been suggested that the lack of differences reported between DQOL scores of those who intensified their treatment in the trial and those who did not may be attributable to the limitations of the measure rather than to any real lack of impact of intensification of treatment on quality of life (Bradley, 1994d). Thus, an item concerned with immediate and pervasive worries about hypoglycemia which may markedly impair an individual's quality of life will have no more impact on the DQOL total score than an item concerned with satisfaction with time taken to determine blood glucose levels or an item about how often the individual has to explain what it means to have diabetes. The DCCT research group reported that the DQOL total scores showed no impact of intensified treatment on quality of life (DCCT, 1993). This was surprising given the considerable increase in severe hypoglycemia experienced by the intensified treatment group in the DCCT, but not surprising when the nature of the quality of life measure used is examined.

Individualized measures of quality of life ask the individual respondent which domains are important to their quality of life. An increasingly valued measure of individual quality of life is the Schedule for Evaluation of Individual Quality of Life (SEIQoL) interview method (Hickey et al., 1996; McGee, O'Boyle, Hickey, O'Malley, & Joyce, 1991; O'Boyle, McGee, Hickey, O'Malley, & Joyce, 1992). The SEIQoL establishes in interview the domains of life important to a particular individual's quality of life, elicits ratings for each domain from "as good as it could possibly be" to "as bad as it could possibly be," determines weights for each domain using a method of judgment analysis. The philosophy underpinning the SEIQoL has influenced the design of a questionnaire measure for Audit of Diabetes Dependent Quality of Life (ADDQoL) (Bradley, Todd, Gorton, Plowright, & Symonds, 1995; Todd, Bradley, & Symonds, 1993). The ADDQoL includes a series of domains identified in pilot work to be commonly affected by diabetes and its treatment. Working life and family life are two of the standard domains included together with worries about the future, enjoyment of food, and other items more specifically relevant to diabetes. Each domain is first rated for the impact of diabetes on that domain and then rated for the importance of that domain for quality of life. Domain ratings are then multiplied by importance ratings such that a domain which is unimportant will score zero regardless of how much it is affected by diabetes while ratings other than zero for an important domain will receive greater weight. Any domain that is said to be inapplicable is ignored. The weighted ratings of all applicable domains are then summed and divided by the number of applicable domains to give scores which vary from the same maximum (+9) to the same minimum (−9) for all individuals. In this way, the ADDQoL, like the SEIQoL, focuses on applicable domains that are important to the particular individual. Unlike the SEIQoL, the ADDQoL is concerned specifically with the effects of diabetes and its treatment on quality of life. The ADDQoL provides a framework which is readily modified to be suitable for use with other patient groups and has been adapted for use with renal patients (Bradley, 1997). There is considerable interest in patient-centered measures of quality of life and although other measures aim to be relevant by ensuring that item design is informed by qualitative work with patients, few instruments are patient-centered in providing opportunities for patients to indicate which items are relevant to them personally and which are not and provide their own weightings of importance for each relevant item (Bradley, 1996b; Greenhalgh, 1996).

8.11.7.4.3 Measurement of knowledge, beliefs, and other cognitions

Many different measures of knowledge of diabetes were designed and developed in the 1980s but more recently there has been widespread disenchantment with such measures (Glasgow, Toobert, Hampson, & Wilson, 1995). This has resulted in part from the abstract content of some knowledge measures and partly from the inappropriate expectation that knowledge of diabetes will correlate with measures of blood glucose control (Beeney, Dunn, & Welch, 1994). When no correlation is found the validity of the knowledge questionnaire has been questioned. However, there is no logical reason to expect that responses to many of the items on knowledge questionnaires will correlate with blood glucose levels (e.g., items concerned with foot care). In addition, validation of knowledge items is complicated by the fact that patients often acquire considerable knowledge after they develop a problem, thereby leading to counterintuitive associations between knowledge and outcomes in cross-sectional surveys. Care and ingenuity is needed to select appropriate methodologies to develop knowledge measures for particular purposes (Bradley, 1995). In the light of continuing evidence of inadequate levels of knowledge, with patients with NIDDM being particularly disadvantaged (Anderson, Hiss, Stepien, Fitzgerald, & Funnel, 1994), the measurement of knowledge and provision of remedial education as required is likely to result in marked improvements in biomedical and psychosocial outcomes of diabetes care.

Other measures of psychological processes which can be valuable in understanding difficulties with diabetes management include measures of health beliefs, and perceptions of control, or locus of control, of diabetes. A variety of reliable diabetes-specific measures have been validated for various purposes (Bradley, 1994e; Kohlmann, Schuler, Petrak, Kustner, 1993; Lewis & Bradley, 1994). Measures of health beliefs and, particularly, measures of patients" attributions or locus of control have proved useful in understanding patients preferences for injection treatments or continuous subcutaneous insulin infusion (CSII) pumps, and in understanding why some patients developed the dangerous acute complication, diabetic ketoacidosis (DKA), during pump use (Bradley, Gamsu, Knight, Boulton, & Ward, 1986; Bradley et al., 1987).

In a feasibility study involving all insulin treated patients at the Royal Hallamshire Hospital in Sheffield, UK (Knight et al., 1984), patients were offered the choice of CSII pumps, intensified injection treatment, or conventional injection treatment. An illustrated lecture informed patients about the treatments available and patients were told that "the evidence suggests that CSII controls blood glucose levels better than injections." Unfortunately, only with hindsight was it realized that patients needed to know, but had not been told, that the effectiveness of CSII depended on the actions of the individual user and that CSII required at least as much, if not more, thought, effort, and precautions to use it effectively and safely than was needed with injection therapies. Those attracted to CSII tended to be those who were seeking a medical solution for their condition and who had less sense of personal control over their diabetes (as measured by the perceived control of diabetes scales). Sadly, it was those with the strongest sense of medical control and least sense of personal control who did least well with CSII pumps both in terms of blood glucose control at 12 months (Bradley et al., 1987) and occurrence of DKA during the study (Bradley et al., 1986). A massive increase in rates of DKA occurred among CSII users in the first two years of the study, and the only variables to discriminate between those who developed DKA and a matched group of those with no DKA were the psychological variables, perceived control and health beliefs.

Although questionnaire measures of health beliefs and cognitions are somewhat cumbersome for clinical use, experience with the tools in a research context and in detailed assessment of individuals presenting with particular problems can be valuable. The concepts underlying the measures can usefully guide clinical interviews seeking to establish sources of difficulty with diabetes management with a view to tailoring the treatment to suit the individual patient and identifying inappropriate beliefs and expectations of treatments which need to be addressed.

The *Handbook of psychology and diabetes* (Bradley, 1994f) provides ready access to most of the questionnaire measures described above and provides reference to yet more measures. Despite the availability of such measures and increasing evidence for their value in clinical as well as research contexts, there is resistance among some clinicians to monitoring psychological outcomes (Barendse, 1995). The unfamiliarity of the tools leads many physicians to feel that they would need training in use of the instruments before they would feel able to use them. There is also concern about how they would deal with cases of depression and anxiety which may be brought to light as a result of using the questionnaires. These are both understandable concerns which may be dealt with by involvement of psychologists to support use and interpretation of the psychological measures and in helping to deal with the psychological problems identified. Interestingly, one source of resistance to using the instruments was the reported concern that patients would not want to have psychological outcomes considered as part of routine diabetes care though all the available evidence suggests that quite the reverse is the case: people with diabetes and parents of children with diabetes have long complained about the excessive emphasis on their blood glucose levels in diabetes clinics while little attention is paid to the effects of diabetes management on their lifestyle and quality of life. Clinical psychologists have an important role to play in encouraging and supporting diabetes care teams in measuring psychological outcomes and in taking action to protect or improve the psychological outcomes of people with diabetes.

8.11.7.5 Patient Empowerment

With diabetes, as much if not more than with other chronic disorders, optimal management can only be achieved by means of complex behavioral changes on the part of the patient. It is not enough for the clinician merely to give instructions. Health professionals are increasingly working towards holistic approaches to treatment, which include an emphasis on providing the education, support, and encouragement to enable patients, where possible, to manage their own diabetes. Indeed, the second edition of the St. Vincent Declaration Action Programme to improve diabetes care in Europe (Krans et al., 1995) now includes specific guidelines to facilitate patient empowerment.

The construct of empowerment has been in use in other domains of psychology for some time, and has at times been both ill-defined and misunderstood. In the context of diabetes care, a useful overview of patient empowerment has been provided by Anderson (1995).

The "empowerment" of individual patients is a complex goal. It requires first that patients have a level of authority and responsibility to be able to make decisions and manage their own treatment. Second, it requires that they be willing to accept that responsibility and to achieve the necessary skills and knowledge to make them capable of managing their diabetes. Third, the environment of care needs to be such that policies and practices facilitate self-management of diabetes care. Central to the achievement of each of these elements is the healthcare professional. They must accept patients as active and responsible self-carers, educate them both in knowledge of diabetes and self-care choices, and facilitate the setting of individual goals of self-care through an assessment of individual preferences. Patently, this becomes even more vital in an age of technologically advanced healthcare systems, where the tendency may be to alienate patients from active participation in their own medical treatment (Stabler, 1993).

Whilst numerous accounts of research into patient empowerment may be found, the literature specific to diabetes care is limited, but growing (Anderson, 1995; Anderson, Funnell, Barr, Dedrick, & Davis, 1991; Anderson et al., 1995; Doherty, Hall, James, & Roberts, 1996; Feste, 1992; Kinmonth et al., 1996). In looking to the future for a more widespread adoption of attitudes and practices which facilitate patient empowerment, it is clear that certain barriers need to be overcome. Some patients may be reluctant to accept responsibility for their own diabetes management, and need to be encouraged with the knowledge that they are best placed to make decisions which impact on the day to day running of their lives. They then need access to professional support, education in the knowledge of their condition, self-care skills, and choices available to them. Healthcare professionals may also misunderstand patient empowerment. They may see this as an unethical suggestion that they should abdicate responsibility for care and relinquish clinical control and decision making. At the other end of the spectrum, some may see patient empowerment as a license to divest themselves of the more intractable problems associated with diabetes care. In reality, however, patient empowerment involves active participation of the health professional in the role of facilitator, enabling patients to set individual goals for diabetes management and to achieve optimal outcomes in both medical and psychosocial health.

There is early evidence for the success of patient empowerment interventions. Anderson et al. (1995) reported a randomized waiting-list control group trial of patient empowerment education. Despite some problems associated with drop-out rates and the combination of data from randomized and nonrandomized participants for analysis, the results suggested a modest improvement in blood glucose control, along with gains in self-efficacy and attitudes toward diabetes.

Greenhalgh, Chowdhury, and Helman (1996) evaluated a patient empowerment model with a sample of British Bangladeshis with diabetes, and identified a variety of culturally related factors which may create obstacles to patient empowerment including ignorance about diabetes and available services, poverty, fatalism, and religious convictions. These findings underline the need for individual patient focus in empowerment programs if those programs are to be appropriate to patients needs. Quatromoni et al. (1994) described a study which used focus groups to explore culture-specific information that could be useful in empowerment interventions with Caribbean Latinos with NIDDM. Amongst the issues they identified were

> feelings of social isolation, little understanding of long-term consequences of diabetes, fatalism regarding the course of the disease, barriers to diet and exercise interventions, skepticism regarding the value of preventive health behaviors, and a clear need for culturally-sensitive health-care provision. (Quatromoni et al., p. 869)

Similarly, the Haida Gwaii Diabetes Project (Herbert, 1996) described an approach which uses community-based research to identify a culturally sensitive approach to diabetes prevention and management. Outcomes suggest an empowerment benefit to individuals with diabetes, their families, and community. Moreover, there were reported benefits to healthcare professionals, in that they also felt empowered to implement diabetes care in ways which were more pertinent to patients.

Evidence suggests that patient empowerment offers a promising way forward in diabetes care. It is perhaps worthwhile, however, to sound a cautionary note. Should empowerment interventions fail to show significant benefits in the future, it may be that the interventions themselves have not been adequate to the task, either in that health professionals did not receive adequate training to enable them to move to more participative care, or that the empowerment goals were not focused sufficiently on the cultural or individual needs of patients.

8.11.7.6 Weight Management

Weight management is a problem both for people with IDDM and NIDDM in rather different ways. People with IDDM need to balance their carbohydrate intake, insulin, and exercise in order to control their blood glucose levels. Dietary recommendations tend to be less restrictive than was the case in the 1980s though typically patients would be advised to avoid foods containing simple sugars (sweets, cakes, sugary drinks, etc.) and foods with a high fat content. Paradoxically, if they develop hypoglycemic symptoms they need to take instantly available glucose to avoid a severe hypoglycemic episode which can lead to coma. The consequent rise in blood glucose may overshoot and result in hyperglycemia. If their carbohydrate intake is too high, an increase in insulin dose will be needed to maintain metabolic control and, as with people who do not have diabetes, carbohydrate excess will lead to weight gain. However, in people with diabetes, crash diets to deal with unwanted weight may result in metabolic chaos and a careful weight reduction program is needed which balances reduced carbohydrate intake and/or increased exercise with reduced insulin doses.

8.11.7.6.1 Eating disorders and withholding of insulin

Adolescent girls with IDDM tend to be heavier than matched nondiabetic girls for reasons which are not understood, and are likely to be dieting more intensely in order to control their weight. In cross-sectional and case–control studies (Fairburn, Peveler, Davies, Mann, & Mayou, 1991; Peveler, Fairburn, Boller, & Dunger, 1992), eating disorders have been found to be as prevalent but not more common in young women with IDDM than those without. Nine percent of adolescent girls in the latter study met the diagnostic criteria used for "eating disorder not otherwise specified." Other studies have indicated that adolescents and young women with diabetes are at greater risk of eating disorders than those who do not have diabetes (Rodin, Daneman, Johnson, Kenshole, & Garfinkel, 1985; Rosmark et al., 1986). However, care needs to be taken over the methods used to identify eating disorders in people with diabetes. Generic eating disorder self-report assessments, such as the Eating Attitudes Test (EAT) or Eating Disorder Inventory (EDI), are likely to yield different scores from people with diabetes because of the emphasis on diet and dietary restraint in diabetes education. Items such as I "avoid foods with sugar in them" and "engage in dieting behavior" on the EAT and "I eat sweets and carbohydrates without feeling nervous" on the EDI may be useful indicators of eating disorders in the general population, but in people with diabetes, responses will be affected by the diabetes treatment regimens commonly recommended. Efforts have been made to control for this bias by excluding from the EAT and EDI those items which would be answered significantly differently by those with diabetes (Cantwell & Steel, 1996; Steel, Young, Lloyd, & Macintyre, 1989). However, further evaluation is needed before these modified measures can be used with confidence in studies involving people with diabetes. Rodin and Daneman (1992), in a review of this literature, suggested that the reliability of prevalence rates for "caseness" is improved when based on structured interviews and standardized diagnostic criteria such as *ICD-9*, rather than self-report questionnaires which are often used in prevalence studies.

However, disordered eating has been found to be common even in more methodologically rigorous studies in addition to more commonly recognized eating disorders and Peveler et al. (1992) also reported that 15% of their sample were omitting or reducing their insulin in order to reduce their weight. Insulin withholding has also been found to be associated with current or past symptoms of anorexia or bulimia nervosa (Biggs, Basco, Patterson, & Raskin, 1994). When insulin is withheld, blood glucose levels will rise until they reach the renal threshold (10 mmol/L) and then glucose will be excreted in the urine taking with it water which would otherwise have been retained. Thus, withholding of insulin is a means of effortless weight loss though not without potentially serious consequences. In particular, the chronic hyperglycemia of a person who habitually withholds insulin to lose weight, will put the individual at risk of DKA requiring hospitalization in the short term (if they omit insulin for too long a period) and, in the longer term, the risk of microvascular complications of diabetes is markedly increased. Pediatricians responsible for the care of adolescents with diabetes may well downplay concerns about future complications for fear of alarming young patients and their parents and it can, therefore, be quite possible for young people with diabetes to be ignorant of the well-established causal link between chronic hyperglycemia and the serious microvascular complications which are the reasons for concern that motivate the clinicians to attempt to improve diabetes control. Where young people with IDDM are well informed about the risks of microvascular complications and the importance of blood glucose control for reducing those risks, the attractions of quick and easy weight loss may nevertheless outweigh the more distant concerns of complications in 15 years or so and, by injecting some insulin occasionally, the dangers of DKA and associated coma may be avoided. Withholding of insulin thus presents a diabetes-specific form of eating disorder which can have much sought after immediate effects but with serious short- and longer-term consequences.

8.11.7.6.2 Weight reduction programs in NIDDM

In NIDDM, obesity is more prevalent than in the general population. Dietary therapy is the mainstay of treatment of NIDDM. A significant proportion of people who follow the dietary recommendations will regain metabolic control without having to take oral hypoglycemic tablets or insulin injections. In the overweight, calorie restriction is particularly important for metabolic control and weight loss is associated with improved glycemic control. Many studies have shown that it is possible to induce weight loss in NIDDM patients using a variety of pharmaceutical methods (for example, appetite suppressants), behavioral methods, exercise, conventional diets, or very low calorie diets. However, longer-term studies typically show that subjects lose weight for the first six months of a program and thereafter regain the weight (Wing & Anglin, 1996). The UK Prospective Diabetes Study Group (1990) examined the responses of more than 3000 newly diagnosed NIDDM patients to diet therapy and showed that 16% achieved near-normal fasting glucose after three months (with an average weight loss of 8 kg). However in the second three months the blood glucose increased a little in spite of continued but decelerated weight loss. Metabolic control appears to be related more to calorie intake than body weight *per se*. This study shows the value of diet therapy in the successful management of NIDDM, but the weight loss required for adequate diabetes control may be considerable and oral hypoglycemic agents may be needed as well.

The evidence suggests that even when dietary change is achievable in the short term, it is not sustainable at the same level over a longer period. Perhaps more attention needs to be given to developing diets which are more acceptable to the individual and, hence, more sustainable. Diets may be unacceptable for economic, psychological, or cultural reasons. A diet that may be affordable and tolerable in the short term, with the support of therapist and fellow patients on a weight loss program, may be unsustainable without such support. There is marked individual variation in weight loss achieved in weight-loss programs. In one intervention study, outcomes varied from a weight loss of 29.5 kg to a gain of 7.7 kg from pretreatment to follow-up (Wing, Epstein, Nowalk, Koeske, & Hagg, 1985). An understanding of the reasons for the variability may well be helpful in determining how to improve the success rate. In particular, consideration of patients" preferences for different weight loss programs may be of value in facilitating successful weight loss. It is likely that a group program will suit some but not others and different kinds of exercise program will suit different people. It is important to bear in mind that a conventional randomized clinical trial of particular weight-loss programs will have recruited only those individuals prepared to follow any of the programs under study and will be likely to have failed to recruit those patients with a strong preference for one program if the program (or a similar program) was available outside the trial. Where a new program is only available within a trial, the patients recruited are likely to have marked preferences for that program, and disappointment will follow if randomization allocates the nonpreferred treatment. Subsequent responses, which may include dropping out of the trial, will undermine the assumptions of a randomized trial. Such trials, attempting to answer the, perhaps inappropriate, question, “Which weight loss program works best?” may be less useful in informing routine clinical practice than preference trials or partially randomized preference trials, where patients with preferences are encouraged to follow the treatment they prefer and only those without preferences are invited to accept allocation of treatment by randomization. In such preference trials the question of interest is “Which weight loss program works best for which patients?” (Bradley, 1993).

Some studies have begun to look at reasons for success and failure in weight loss. National survey data in the US were used to investigate self-reported weight loss and it was found that intentional weight loss in people with diabetes was associated with having been hospitalized twice in the previous year, and having seen a dietitian (Will, 1995) suggesting that a strong motivation to lose weight is important for success. It has been shown that men lose weight faster than women but no association was found between weight loss and whether or not subjects were taking insulin, nor between weight loss and duration of diabetes (Wing, Shoemaker, Marcus, McDermott, & Gooding, 1990). Thus, use of insulin and diabetes of long duration do not appear to present particular obstacles to weight loss. In a later study, smaller weight losses were found in blacks than whites related to faster regaining of weight in blacks (Wing & Anglin, 1996). Pierce and Armstrong (1996) found that in Afro-Caribbean diabetic patients the prescribed diets were held in contempt and largely ignored as being so mean as to be incompatible with a healthy life. Moreover, the subjects disagreed profoundly with the medical notion of a healthy body weight. Large body size was perceived as healthy and attractive within this group and valued as such.

8.11.7.7 Treatment of Sexual Dysfunction

The vascular and neuropathic complications of diabetes are associated with a higher incidence of sexual problems among men with diabetes than among nondiabetic men. Women with diabetes may also experience more sexual difficulties than nondiabetic women though not all studies have found differences between women with and without diabetes (Jensen, 1981). It has been erectile failure in men with a prevalence of 35–59% (Bancroft & Gutierrez, 1996) which has attracted the attention of clinicians, researchers, and pharmaceutical companies. The potential contribution of psychologists in dealing with the sexual problems of men with diabetes has been little recognized. The increased likelihood of an organic component to erectile dysfunction associated with diabetes has led to a disproportionate emphasis on drug (self-administered papaverine injections to induce erection), mechanical (vacuum tumescence), and surgical interventions (penile implants) vigorously promoted by the manufacturers. Although diabetic men are at least, if not more, likely than other men to have impotence of psychogenic origins, the ready availability of funding for trials of drug and surgical interventions has encouraged and facilitated such approaches, while funding for trials of psychological therapy is far less readily available and such studies are few and far between. Nevertheless, there is evidence that good results can be achieved with psychological approaches. Veves, Webster, Chen, Payne, and Boulton (1995) reported that psychosexual counseling resulted in successful intercourse in 17 (60%) of the 24 men opting for this treatment. A further seven (25%), in whom the marital relationship had previously been unstable, went on to medical treatment after marital therapy. The success rate for psychosexual interventions compared favorably with the success rate (55%) for those opting for papaverine injections and did not have the unwanted consequences experienced by 15% of men using papaverine which included priapism, pain at injection site, and dizziness. However, the modest success rates of psychosexual counseling with unselected patients reported by McCulloch, Hosking, and Tobert (1986) have been widely quoted as reason for pursuing biomedical approaches to treatment. Increasing experience of the failure of surgical implants and complications following surgery have led to the view that surgery should be a last resort (Alexander, 1990; Dunsmuir & Holmes, 1996). Furthermore, there is increasing emphasis on the importance of psychosexual counseling prior to biomedical interventions (Dunsmuir & Holme, 1996).

Several authors have noted an unwillingness among their patients to view their erectile dysfunction as psychogenic in origin, and the limited enthusiasm for psychosexual counseling (Alexander, 1990; McCulloch et al., 1986; Veves et al., 1995). How much of this resistance is due to physicians" emphasis on biomedical approaches and limited appreciation of psychological approaches is unclear. An editorial in the *British Medical Journal* recommended that patients with erectile impotence "should be told that impotence in diabetes has a physical cause and will not improve without a physical treatment" (Price, 1993). Others have taken issue with this view (Bancroft & Gutierrez, 1996), but it is clear that this widespread exclusive focus on erectile dysfunction as an organic problem in diabetes will present a challenge to those wishing to encourage psychological approaches.

Meisler, Carey, Lantinga, and Krauss (1989) have offered guidelines for a multidisciplinary approach to assessment of sexual dysfunction. Further development and wider use of psychological interventions has the potential for improved psychosexual functioning and prevention of the often irreversible complications of biosurgical interventions.

8.11.8 CULTURAL ISSUES

Cultural issues have been referred to at various points throughout the chapter along with life span, age, and gender issues where each has been particularly relevant. Sections of the chapter concerned with risk factors for diabetes review evidence showing diabetes to be less common in white Western populations than in other populations. The high incidence of diabetes among racial minorities in Western society presents particular challenges to diabetes care services when language barriers pose particular problems of management for a disorder which requires complex self-management tasks and dietary change. Good communication between health professionals and their patients is central to the development of self-management skills and appropriate practical solutions need to be found to facilitate communication across languages with multilingual health professionals or use of translators familiar with diabetes and its management. Clashes between cultural beliefs, lifestyle, and recommendations for diabetes management, especially dietary recommendations, create particular difficulties where weight reduction is an important component of the recommended management plan for overweight people with NIDDM (see Section 8.11.7.6.2).

8.11.9 PROFESSIONAL ISSUES

In the USA involvement of psychologists in diabetes care teams is now standard practice. In the UK it is becoming increasingly common to find psychologists attached to diabetes care teams for two or more sessions per week although this is still the exception rather than the rule. A briefing paper from the British Psychological Society (1994) on services for people with diabetes mellitus recommended one whole-time equivalent post for each local diabetes service allocated between adult services, services to the elderly, and services to children and young people. It was recognized that clinical psychologists who work with adults will not normally have the skills and knowledge to work effectively with children and their families and that special expertize may also be required for some older people. Such recommendations are realistic rather than ideal. This level of input is likely to be mostly taken up with crisis intervention and problem solving rather than preventive work to improve communication, knowledge, and skills in diabetes management across the service as a whole which in the longer term would be of most benefit.

There is considerable scope for clinical psychologists to make valuable contributions to diabetes management, providing assessment and treatment, and conducting research and evaluation as well as teaching other health professionals and patients. Several specific psychological therapies which clinical psychologists may be able to offer include stress management and relaxation training, weight management interventions, and management of eating disorders and sexual dysfunction. More general psychological skills will be valuable in helping patients to identify and overcome barriers to self-care and in facilitating the process of helping patients and their advisors to negotiate manageable treatment regimens which are tailored to individuals" lifestyles and preferences.

For clinical psychologists to be most effective in diabetes care, detailed understanding is required of diabetes and the varieties of management that are possible. The management regimens recommended in a particular clinic are likely to be only a small minority of those which could be made available. There are many diabetologists who continue to recommend particular doses and timing of insulin which then dictate the timing of meals and snacks and require patients to manipulate their diet and exercise to fit in with the regimen prescribed. Initially, it may well be more demanding to teach patients how to manipulate their insulin appropriately to fit in with preferred meal times and varying levels of energy expenditure, but in the longer term it is a great deal less demanding and more convenient for the patient, their family, friends, and other associates if they adjust their insulin dose rather than their life (Howorka, 1996). Problems with poor diabetes control, which diabetologists may construe as "compliance" problems, may more productively be construed as inflexible treatment recommendations and/or inadequate education in diabetes management and self-care skills in adjusting insulin doses and timing to accommodate a flexible lifestyle. By developing an understanding of how diabetes can be managed physiologically and by appreciating the range of different medical approaches to diabetes management, clinical psychologists are better placed to apply their psychology to the task of improving diabetes management while at the same time protecting, and even improving, the psychological well-being and quality of life of people with diabetes.

8.11.10 REFERENCES

Aikens, J. E., Wallander, J. L., Bell, D. S. H., & McNorton, A. (1994). A nomothetic-idiographic study of daily psychological stress and blood glucose in women with Type 1 diabetes mellitus. *Journal of Behavioral Medicine, 17,* 535–548.

Akerblom, H. K., Savilahti, E., Saukkonen, T. T., Paganus, A., Virtanen, S. M., Terami, K., Knip, M., Ilonen, J., Reijonen, H., Karjalainen, J., Vaarala, O., & Reunanen, A. (1993). The case for elimination of cows" milk in early infancy in the prevention of Type 1 diabetes: the Finnish experience. *Diabetes/Metabolism Reviews, 9,* 269–278.

Alexander, W. D. (1990). The diabetes physician and an assessment and treatment program for male erectile impotence. *Diabetic Medicine, 7,* 540–543.

Alexander, W. D. (1993). Human insulin: lessons from the UK? *The Medical Journal of Australia, 159,* 75–76.

American Diabetes Association. (1992). Screening for diabetic retinopathy. *Diabetes Care, 15*(Suppl. 2), 16–18.

Anderson, R. M. (1995). Patient empowerment and the traditional medical model: A case of irreconcilable differences? *Diabetes Care, 18,* 412–415.

Anderson, R. M., Funnell, M. M., Barr, P. A., Dedrick, R. F., & Davis, W. K. (1991). Learning to empower patients: Results of professional education program for diabetes educators. *Diabetes Care, 14,* 584–590.

Anderson, R. M., Funnell, M. M., Butler, P. M., Arnold, M. S., Fitzgerald, J. T. & Feste, C. C. (1995). Patient empowerment: Results of a randomized controlled trial. *Diabetes Care, 18,* 943–949.

Anderson, R. M., Hiss, R. G., Stepien, C. J., Fitzgerald, J. T., & Funnel, M. M. (1994). The diabetes education experience of randomly selected patients under the care of community physicians. *Diabetes Educator, 20,* 399–405.

Bancroft, J., & Gutierrez, P. (1996). Erectile dysfunction in men with and without diabetes mellitus: a comparative study. *Diabetic Medicine, 13,* 84–89.

Barendse, S. (1995). *Doctors" intentions to use psychological measures in diabetes clinics: A field evaluation of the theory of reasoned action.* (Project Report). Department of Psychology, Royal Holloway, University of London.

Barglow, P., Hatcher, R., Edidin, D. V., & Sloan-Rossiter, M. (1984). Stress and metabolic control in diabetes: Psychosomatic evidence and evaluation of methods. *Psychosomatic Medicine, 46,* 127–144.

Beaty, T. H., Neel, J. V., & Fajans, S. S. (1982). Identifying risk factors for diabetes in first degree relatives of non-insulin-dependent diabetes. *American Journal of Epidemiology, 115,* 380–397.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561–571.

Beeney, L. J., Dunn, S. M., & Welch, G. (1994). Measurement of diabetes knowledge: the development of the DKN scales. In C. Bradley (Ed.) *Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice* (pp. 159–189). Chur, Switzerland: Harwood Academic Publishers.

Biggs, M. M., Basco, M. R., Patterson, G., & Raskin, P. (1994). Insulin withholding for weight control in women with diabetes. *Diabetes Care, 17,* 1186–1189.

Bingley, P. J., Bonifacio, E., & Gale, E. A. M. (1993). Can we really predict IDDM? *Diabetes, 42,* 213–220.

Bingley, P. J., & Gale, E. A. M. (1989). Rising incidence of IDDM in Europe. *Diabetes Care, 12,* 289–295.

Blohmé, G., Nyström, L., Arnqvist, H. J., Lithner, F., Littorin, B., Olsson, P. O., Schersten, B., Wibell, L., & Östman J. (1992). Male predominance of Type 1 (insulin-dependent) diabetes mellitus in young adults: Results from a 5-year prospective nationwide study of the 15–34 year age group in Sweden. *Diabetologia, 35,* 56–62.

Bradley, C. (1988). Stress and diabetes. In S. Fisher & J. Reason (Eds.), *Handbook of life stress, cognition and health* (pp. 383–401). Chichester, UK: Wiley.

Bradley, C. (1989). Psychosocial influences and adjustment: family and health care systems. In R. Larkins, P. Zimmet, & D. Chisholm (Eds.), *Diabetes 1988.* Amsterdam: Elsevier.

Bradley, C. (1993). Designing medical and educational intervention studies: a review of some alternatives to conventional randomized controlled trials. *Diabetes Care, 16,* 509–518.

Bradley, C. (1994a). The contribution of psychology to diabetes managment. *British Journal of Clinical Psychology, 33,* 11–21

Bradley, C. (1994b). The Well-being Questionnaire. In: C. Bradley (Ed.), *Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice* (pp. 89–109). Chur, Switzerland: Harwood Academic Publishers.

Bradley, C. (1994c). The Diabetes Treatment Satisfaction Questionnaire (DTSQ). In C. Bradley (Ed.), *Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice* (pp. 111–132). Chur, Switzerland: Harwood Academic Publishers.

Bradley, C. (1994d). What about quality of life? *Balance, 140,* Special issue on the implications of the Diabetes Control and Complications Trial. VIII-IX June/July.

Bradley, C. (1994e). Measures of perceived control of diabetes. In C. Bradley (Ed.), *Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice* (pp. 291–331). Chur, Switzerland: Harwood Academic Publishers.

Bradley, C. (Ed.) (1994f). *Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice.* Chur, Switzerland: Harwood Academic Publishers.

Bradley, C. (1995). Health beliefs and knowledge of patients and doctors in clinical practice and research. *Patient Education and Counseling, 26,* 99–106.

Bradley, C. (1996a). The WHO (Ten) Well-being Index: A critique. *Psychotherapy and Psychosomatics, 65,* 331–333.

Bradley, C. (1996b). Measuring quality of life in diabetes. In S. M. Marshall, P. D. Home, & R. A. Rizza (Eds.), *The Diabetes Annual 10* (pp. 207–224). Amsterdam: Elsevier.

Bradley, C. (1997). Design of a renal-dependant individualised quality of life questionnaire. *Advances in Peritoneal Dialysis, 13,* 116–120.

Bradley, C., Christie, M., & Home, P. (Eds.) (1991). *The technology of diabetes care: Converging medical and psychosocial perspectives.* Chur, Switzerland: Harwood Academic Publishers.

Bradley, C., & Gamsu, D. S. (1994). Guidelines for encouraging psychological well-being. Report of a working group of the World Health Organization regional office for Europe and International Diabetes Federation European St. Vincent Declaration action program for diabetes. *Diabetic Medicine, 11,* 510–516

Bradley, C., & Gamsu, D. S. (1995). Diabetes: psychosocial issues in adolescents and adults. *Current Opinion in Psychiatry, 6,* 400–406.

Bradley, C., Gamsu, D. S., Knight, G., Boulton, A. J., & Ward, J. D. (1986). Predicting risk of diabetic ketoacidosis in patients using continuous subcutaneous insulin infusion. *British Medical Journal, 293,* 242–243.

Bradley, C., Gamsu, D. S., Moses, J. L., Knight, G., Boulton, A. J. M., Drury, J., & Ward, J. D. (1987). The use of diabetes-specific perceived control and health belief measures to predict treatment choice and efficacy in a feasibility study of continuous subcutaneous insulin infusion pumps. *Psychology and Health, 1,* 133–146.

Bradley, C., & Lewis, K. S. (1990). Measures of psychological well-being and treatment satisfaction developed from the responses of people with tablet-treated diabetes. *Diabetic Medicine, 7,* 445–451.

Bradley, C., Moses, J. L., Gamsu, D. S., Knight, G., & Ward, J. D. (1985). The effects of relaxation on metabolic control of Type 1 diabetes; a matched controlled study. *Diabetes, 34,* 17A.

Bradley, C., Todd, C., Gorton, T., Plowright, R. E. A., & Symonds, E. (1995) *A patient-centered measure of the impact of a chronic disorder on quality of life: design and development of the ADDQoL.* International Conference and Exhibition: Scientific Basis of Health Services, London.

British Psychological Society, Division of Clinical Psychology (1994). *Purchasing clinical psychology services: services for people with diabetes mellitus.* Briefing paper No. 4, British Psychological Society.

Bruno, G., Merletti, F., Vuolo, A., Pisu, E., Giorio, M., & Pagano, G. (1993). Sex differences in incidence of IDDM in age group 15–29 yr. Higher risk in males in Province of Turin, Italy. *Diabetes Care, 16,* 133–136.

Cantwell, R., & Steel, J. M. (1996). Screening for eating disorders in diabetes mellitus. *Journal of Psychosomatic Research, 40,* 15–20.

Carter, W. R., Gonder-Frederick, L. A., Cox, D. J., Clarke, W. L., & Scott, P. (1985). Effects of stress on blood glucose in IDDM. *Diabetes Care, 8,* 411–412.

Chase, H. P., & Jackson, G. G. (1981). Stress and sugar control in children with insulin-dependent diabetes. *Journal of Pediatrics, 98,* 1011–1013.

Clayer, J. R., Bookless-Pratz, C., & Harris, R. L. (1985).

Some health consequences of a natural disaster. *The Medical Journal of Australia, 143,* 182–184.

Cox, D. J., Gonder-Frederick, L., Anderson, R., Schlundt, D., Jacobson, A., Marerro, D., & Clarke, W. L. (1994) Professionals" beliefs about useful symptoms of hypoglycemia. *Diabetes Care, 17,* 776–777.

Cox, D. J., Gonder-Frederick, L. A., Julian, D. M., & Clarke, W. L. (1992). *Blood glucose awareness training manual.* Charlottesville, VA: University of Virginia.

Cox, D. J., Gonder-Frederick, L., Julian, D. M., & Clarke, W. L. (1994). Long-term follow-up evaluation of blood glucose awareness training. *Diabetes Care, 17,* 1–5.

Cox, D. J., Gonder-Frederick, L., Julian, D. M., Cryer, P., Lee, J. A., Richards, F. E., & Clarke, W. L. (1991). Intensive versus standard blood glucose awareness training (BGAT) with insulin-dependent diabetes: Mechanisms and ancillary effects. *Psychosomatic Medicine, 53,* 453–462.

Cox, D. J., Gonder-Frederick, L. A., Kovatchev, B., Polonsky, W., Schlundt, D., Julian, D. M., & Clarke, W. L. (1995). Reduction of severe hypoglycemia (SH) with blood glucose awareness training (BGAT-2). *Diabetes, 44*(Suppl.), 27A.

Cox, D. J., Gonder-Frederick, L., Polonsky, W., Schlundt, D., Julian, D. M., & Clarke, W. L. (1995). A multi-center evaluation of blood glucose awareness training—II. *Diabetes Care, 18,* 523–528.

Cranston, I., Lomas, J., Maran, A., Macdonald, I., & Amiel S. A. (1994); Restoration of hypoglycaemia awareness in patients with long-duration insulin-dependent diabetes. *Lancet, 344,* 283–287.

DCCT Research Group (1993). The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *New England Journal of Medicine, 329,* 977–986.

Deary, I. J., Hepburn, D. A., Macleod, K. M., & Frier, B. M. (1993). Partitioning the symptoms of hypoglycemia using multisample confirmatory factor analysis. *Diabetologia, 36,* 771–777.

DeGood, D. E., & Redgate, E. S. (1982). Interrelationships of plasma cortisol and other activation indices during EMG biofeedback training. *Journal Behavioral Medicine, 5,* 213–223.

Diabetes Epidemiology Research International Group (1988). Geographic patterns of childhood insulin-dependent diabetes mellitus. *Diabetes, 37,* 1113–1119.

Diabetic Retinopathy Study Group (1981). Photocoagulation for proliferative diabetic retinopathy. Clinical application of Diabetic Retinopathy Study (DRS) findings. DRS report No. 8. *Opthalmology, 88,* 583–600.

Doherty, Y., Hall, D., James, P. T., & Roberts, S. H. (1996). Teaching change counseling skills to a multidisciplinary diabetes team: Lessons learnt. *Diabetic Medicine, 13*(Suppl. 7), S4.

Dotta, F., Dionisi, S., Farilla, L., & DiMario, U. (1996). The use of immunological markers in IDDM after more than 20 years from ICA description. *Diabetes/Metabolism Reviews, 12,* 121–126.

Dunsmuir, W. D., & Holmes, S. A. (1996). The aetiology and management of erectile, ejaculatory, and fertility problems in men with diabetes mellitus. *Diabetic Medicine, 13,* 700–708.

Edwards, C., & Yates, A. J. (1985). The effects of cognitive task demand on subjective stress and blood glucose levels in diabetic and non-diabetic subjects. *Journal of Psychosomatic Research, 29,* 59–69.

Eisenbarth, G. S., Verge, C. F., Allen, H., & Rewers, M. J. (1993). The design of trials for prevention of IDDM. *Diabetes, 42,* 941–947.

Elliott, R. B., & Chase, H. P. (1991). Prevention or delay of Type 1 (insulin-dependent) diabetes mellitus in children using nicotinamide. *Diabetologia, 34,* 362–365.

Erikksonn, K. F., & Lindegarde, F. (1991). Prevention of Type II (non-insulin-dependent) diabetes mellitus by diet and physical exercise. The 6-year Malmo feasibility study. *Diabetologia, 34,* 891–898.

Fairburn, G. C., Peveler, R. C., Davies, B., Mann, J. I., & Mayou, R. A. (1991). Eating disorders in young adults with insulin-dependent-diabetes mellitus: a controlled study. *British Medical Journal, 303,* 17–20.

Fanelli, C. G., Epifano, L., Rambotti, A. M., Pampanelli, S., Di Vincenzo, N., Modarelli, F., Lepore, M., Annibale, B., Ciofetta, M., Bottini, P., Porcellati, F., Scionti, L., Santensanio, F., Brunetti, P., & Bolli, G. B. (1993). Meticulous prevention of hypoglycemia normalises the glycaemic thresholds and magnitude of most neuroendocrine responses to, symptoms of, and cognitive function during hypoglycemia in intensively treated patients with short-term IDDM. *Diabetes, 42,* 1683–1689.

Feinglos, M. N., Hastedt, P., & Surwit, R. S. (1987). Effects of relaxation therapy on patients with Type I diabetes mellitus. *Diabetes Care, 10,* (1), 72–75.

Felber, J. P., Acheson, K. J., & Tappy, L. (1992). *From obesity to diabetes.* Chichester, UK: Wiley.

Feste, C. (1992). A practical look at patient empowerment. *Diabetes Care, 15,* 922–925.

Gamsu, D. S., & Bradley, C. (1987). Clinical staff's attributions about diabetes: scale development and staff versus patient comparisons. *Current Psychological Research and Reviews, 6,* 69–78.

Gerstein, H. C. (1994). Cow's milk exposure and Type I diabetes mellitus. A critical overview of the clinical literature. *Diabetes Care, 17,* 13–19.

Gillespie, C. R., & Bradley, C. (1988). Causal attributions of doctor and patients in a diabetes clinic. *British Journal of Clinical Psychology, 27,* 67–76.

Glasgow, R. E. (1991). Compliance to diabetes regimens: conceptualization, complexity and determinants. In J. A. Cramer & B. Spilker (Eds.), *Patient compliance in medical practice and clinical trials.* New York: Raven.

Glasgow, R. E. (1994). Social-environmental factors in diabetes: barriers to diabetes self-care. In C. Bradley (Ed.), *Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice* (pp. 335–349). Chur, Switzerland: Harwood Academic Publishers.

Glasgow, R. E., Toobert, D. J. Hampson, S. E., & Wilson, W. (1995). Examples of programmatic research in behavioral medicine. *Annals of Behavioral Medicine, 17,* 32–40.

Goetsch, V. L., Abel J. L., & Pope, M. K. (1994). The effects of stress, mood and coping on blood glucose in NIDDM: A prospective pilot evaluation. *Behavioral Research and Therapy, 32,* 503–510.

Gold, A. E., Deary, I. J., Jones, R. W., O'Hare, J. P., Reckless, J. P., & Frier, B. M. (1994). Severe deterioration in cognitive function and personality in five patients with long-standing diabetes: a complication of diabetes or a consequence of treatment? *Diabetic Medicine, 11,* 499–505.

Gonder-Frederick, L. A., Carter, W. R., Cox, D. J., & Clarke, W. L. (1990). Environmental stress and blood glucose change in IDDM. *Health Psychology, 9,* 503–515.

Gorus, F. K., Anselmo, J., & Keymeulen, B. (1993). Type-I (insulin-dependent) diabetes mellitus: an autoimmune, predictable and preventable disease? Lessons from national registries and new challenges to clinical biology. *Acta Clinica Belgica, 48,* 96–108.

Gorus, F. K., Goubert, P., Semakula, C., Vandewalle, C. L., De Schepper, J., Scheen, A., Christie, M. R., Pipeleers, D. G. et al. for the Belgian Diabetes Registry (1997). IA-2-autoantibodies complement GAD65-autoantibodies in new-onset IDDM patients and help predict impending diabetes in their siblings. *Diabetologia, 40,* 95–99.

Green, A., Gale, E. A. M., & Patterson C. C. (1992). Incidence of childhood-onset insulin-dependent diabetes mellitus: the Eurodiab Ace Study. *Lancet, 339,* 905–909.

Greenhalgh, J. (1996). (The role of) user outcomes within diabetes care. *Outcomes Briefings, 8,* 10–16.

Greenhalgh, P. M., Jones, J. R., Jackson, C. A., Smith C. C. T., & Yudkin, J. S. (1992). Changes in injection-site blood flow and plasma free insulin concentrations in response to stress in Type 1 diabetic patients. *Diabetic Medicine, 9,* 20–29.

Greenhalgh, T., Chowdhury, M., & Helman, C. (1996). Applying the patient empowerment model to non-white ethnic groups: a qualitative survey of health beliefs, illness experience, and barriers to change in British Bangladeshis with diabetes (Abstract). *Diabetic Medicine, 13*(Suppl. 7), S5.

Hagopian, W. A., Sanjeevi, C. B., Kockum, I., Landin-Olsson, M., Karlsen, A. E., Sundkvist, G., Dahlquist, G., Palmer, J., & Lernmark, Å. (1995). Glutamate decarboxylase-, insulin- and islet-cell antibodies and HLA typing to detect diabetes in a general population-based study of Swedish children. *Journal of Clinical Investigation, 95,* 1505–1511.

Hales, C. N., & Barker, D. J. (1992). Type 2 (non-insulin-dependent) diabetes mellitus: The thrifty phenotype hypothesis. *Diabetologia, 35,* 595–601.

Halford, W. K., Cuddihy, S., & Mortimer, R. M. (1990). Psychological stress and blood glucose regulation in Type 1 diabetic patients. *Health Psychology, 9,* 516–528.

Hanson, S. L., & Pichert, J. W. (1986). Perceived stress and diabetes control in adolescents. *Health Psychology, 5,* 439–452.

Harris, M. I. (1984). Unpublished data from the 1976–80 US National Health and Nutrition Examination Survey. National Center for Statistics.

Haynes, R. B., Wang, E., & Gomes, M. D. (1987). A critical review of interventions to improve compliance with prescribed medications. *Patient Education and Counseling, 10,* 155–166.

Helmrich, S. P., Ragland, D. R., Taylor, D. W., Leung, R. W., & Paffenbarger, R. S. (1991). Physical activity and reduced occurrence of non-insulin-dependent diabetes mellitus. *New England Journal of Medicine, 325,* 147–152.

Herbert, C. P. (1996). Community-based research as a tool for empowerment: The Haida Gwaii Diabetes Project example. *Canadian Journal of Public Health, 87,* 109–112.

Herskowitz, R. D., Jackson, R. A., Soeldner, J. S., & Eisenbarth, G. S. (1989). Pilot trial to prevent Type 1 diabetes: progression to overt IDDM despite oral nicotinamide. *Journal of Autoimmunity, 2,* 733–737.

Hickey, A. M., Bury, G., O'Boyle, C. A., Bradley, F., O'Kelly, F. D., & Shannon, W. (1996). A new short-form individual quality of life measure (SEIQoL-DW): application in a cohort of individuals with HIV/AIDS. *British Medical Journal, 313,* 29–33.

Hildebrandt, P., Mehlsen, J., Sestoft, L., & Nielson, S. L. (1985). Mild mental stress in diabetes: Changes in heart rate and subcutaneous blood flow. *Clinical Physiology, 5,* 371–376.

Holmes, C. S., Hayford, J. T., Gonzalez, J. L., & Weydert, J. A. (1983). A survey of cognitive-functioning at different glucose levels in diabetic persons. *Diabetes Care, 6,* 180–185.

Howorka, K. (1996). *Functional insulin therapy. Principles, teaching approach, and practice.* (2nd ed.). Berlin, Germany: Springer Verlag.

Irvine, A., Cox, D. J., & Gonder-Frederick, L. (1994). The Fear of Hypoglycaemia Scale. In C. Bradley (Ed.), *Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice* (pp. 133–155). Chur, Switzerland: Harwood Academic Publishers.

Jacobson, A. M., & the DCCT Research Group. (1994). The Diabetes Quality of Life measure. In C. Bradley (Ed.), *Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice* (pp. 65–87). Chur, Switzerland: Harwood Academic Publishers.

Jarret, R. J., Keen, H., Fuller, J. H., & McCartney, M. (1979) Worsening of diabetes in men with impaired glucose tolerance borderline diabetes. *Diabetologia,* 16, 25–30.

Jennings, A. M., Lewis, K. S., Murdoch, S., Talbot, J. F., Bradley, C., & Ward, J. D. (1991). Randomized trial comparing continuous subcutaneous insulin infusion and conventional insulin therapy in Type II diabetic patients poorly controlled with sulfonylureas. *Diabetes Care, 14,* 738–744.

Jensen, S. B. (1981). Diabetic sexual dysfunction: A comparative study of 160 insulin treated diabetic men and women and an age-matched control group. *Archives of Sexual Behavior, 10,* 493–504.

Jevning, R., Wilson, A. F., & Davidson, J. M. (1978). Adrenocortical activity during meditation. *Hormones and Behavior, 10,* 54–60.

Karvonen, M., Tuomilehto, J., Libman, I., & LaPorte, R. for the World Health Organisation DIAMOND project group (1993). A review of the recent epidemiological data on the worldwide incidence of Type I (insulin-dependent) diabetes mellitus. *Diabetologia, 36,* 883–892.

Keen, H., Jarret, R. J., & McCartney, M. (1982). The ten-year follow-up of the Bedford survey 1962–72: glucose tolerance and diabetes. *Diabetologia, 22,* 73–78.

Keller, R. J., Eisenbarth, G. S., & Jackson R. A. (1993). Insulin prophylaxis in individuals at high risk of Type I diabetes. *Lancet, 341,* 927–928.

Kemmer, F. W., Bisping, R., Steingruber, H. J., Baar, H., Hardtmann, F., Schlaghecke, R., & Berger, M. (1986). Psychological stress and metabolic control in patients with Type I diabetes mellitus. *New England Journal of Medicine, 314,* 1078–1084.

Kinmonth, A-L., Woodcock, A. J., Griffin, S., Campbell, M., Spiegal, N., Burgess, H. (1996). Diabetes care from diagnosis: Development of a patient-centered approach to primary care of newly diagnosed non-insulin dependent diabetes mellitus for evaluation in a randomized controlled trial. *Diabetic Medicine 13* (Suppl. 7), S13.

Knight, G., Boulton, A. J. M., Drury, J., Gamsu, D. S., Moses, J. L., Bradley, C., & Ward, J. D. (1984). A feasibility study of the use of continuous subcutaneous insulin infusion in a diabetic clinic: Patients" choice of treatment. *Diabetic Medicine, 1,* 267–272.

Knowler, W. C., Sartor, G., & Schersten, B. (1987). Effects of glucose tolerance and treatment of abnormal tolerance on mortality in Malmohus County, Sweden. *Diabetologia, 30,* 541.

Kohlmann, C. W., Schuler, M., Petrak, F., Kustner, E., Kronhne, H. W., & Beyer, J. (1993). Associations between type of treatment and illness-specific locus of control in Type 1 diabetes patients. *Psychology and Health, 8,* 383–391.

Krans, H. M. J., Porta, M., Keen, H., & Staehr Johansen, K. (1995). Diabetes care and research in Europe: The St Vincent Declaration action program. *Diabetologia, 15,* 1–84.

Krolewski, A. S., Warram, J. H., Rand, L. I., & Kahn, C. R. (1987). Epidemiologic approach to the etiology of Type 1 diabetes mellitus and its complications. *New England Journal of Medicine, 317,* 1390–1398.

Laakso, M., & Pyörälä, K. (1985). Age of onset and type of diabetes. *Diabetes Care, 8,* 114–117.

Lammers, C. A., Naliboff, B. D., & Straatmeyer, A. J. (1984). The effects of progressive relaxation on stress and diabetic control. *Behavior Research and Therapy, 22,* 641–650.

Lane, J. D., McCaskill, C. C., Ross, S. L., Feinglos, M. N., & Surwit, R. S. (1993). Relaxation Training for NIDDM: Predicting who may benefit. *Diabetes Care, 16,* 1087–1094.

Leslie, R. D., & Elliott, R. B. (1994). Early environmental events as a cause of IDDM. *Diabetes, 43,* 843–850.

Lewis, K. S., & Bradley, C. (1994). Measures of diabetes-specific health beliefs. In C. Bradley (Ed.), *Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice* (pp. 247–289). Chur, Switzerland: Harwood Academic Publishers.

Mathew, R. J., Ho, B. T., Kralik, P., Taylor, D., & Claghorn, J. L. (1980). Catecholamines and migraine: Evidence based on biofeedback induced changes. *Headache, 20,* 247–252.

Mathew, R. J., Ho, B. T., Kralik, P., Taylor, D., Semchuk, K., Weinman, M., & Claghorn, J. L. (1980). Catechol-O-methyltransferase and catecholamines in anxiety and relaxation. *Psychiatric Research, 3,* 85–91.

McColl, E., Steen, I. N., Meadows, K. A., Hutchinson, A., Eccles, M. P., Hewison, J., Fowler, P., and Blades, S. M. (1995). Developing outcome measures for ambulatory care: an application to asthma and diabetes. *Social Science and Medicine, 41,* 1339–1348.

McCulloch, D. K., Hosking, D. J., & Tobert, A. (1986) A pragmatic approach to sexual dysfunction in diabetic men: Psychosexual counseling. *Diabetic Medicine, 3,* 485–489.

McGee, H. M., O'Boyle, C. A., Hickey, A., O'Malley, K., & Joyce, C. R. B. (1991). Assessing the quality of life of the individual: The SEIQoL with a healthy and a gastroenterology unit population. *Psychological Medicine, 21,* 749–759.

McGrady, A., Bailey, B. K., & Good, M. P. (1991). Controlled study of biofeedback-assisted relaxation in Type 1 diabetes. *Diabetes Care, 14,* 360–365.

McGrady, A., & Gerstenmaier, L. (1990). Effect of biofeedback assisted relaxation training on blood glucose levels in a Type 1 insulin-dependent diabetic: A case report. *Journal of Behavior Therapy and Experimental Psychiatry, 21,* 69–75.

Meisler, A. W., Carey, M. P., Lantinga, L. J., & Krauss, D. J. (1989). Erectile dysfunction in diabetes mellitus: A biopsychosocial approach to etiology and assessment. *Annals of Behavioral Medicine, 11,* 18–27.

Moberg, E., Kollind, M., Lins, P. E., & Adamson, U. (1994). Acute mental stress impairs insulin sensitivity in IDDM patients. *Diabetologia, 37,* 247–251.

Mulhauser, I. (1990). Smoking and diabetes. *Diabetic Medicine, 7,* 10–15.

Naliboff, B. D., Cohen, H. J. & Sowers, J. D. (1985). Physiological and metabolic responses to brief stress in non-insulin dependent diabetic and control subjects. *Journal of Psychosomatic Research, 1985, 29,* 367–374.

Neel, J. V. (1982). The thrifty genotype revisited. In J. Kobberling & R. Tattershall (Eds.), *The genetics of diabetes mellitus.* London: Academic Press.

Newman, B., Selby, J. V., King, M. C., Slemenda, C., Fabsitz, R., & Friedman, G. D. (1987). Concordance for Type 2 (non-insulin-dependent) diabetes mellitus in male twins. *Diabetologia, 30,* 763–768.

Nordenfelt, L. (1996). Prevention and ethics in medicine: The case of diabetes prevention, *Journal of Pediatric Endocrinology and Metabolism, 9,* 381–386.

North, J. (1991) The pen-injector's impact on self-management. In C. Bradley, M. Christie, & P. Home (Eds.), *The technology of diabetes care: Converging medical and psychological perspectives* (pp. 84–88). Chur, Switzerland: Harwood Academic Publishers.

O'Boyle, C. A., McGee, H., Hickey, A. O'Malley, K., & Joyce, C. R. (1992). Individual quality of life in patients undergoing hip replacement surgery. *Lancet, i,* 1088–1091.

Palmer, J. P., & McCulloch, D. K. (1991). Prediction and prevention of IDDM—1991. *Diabetes, 40,* 943–947.

Pan, X., Guanwei, L., Zuoxin, A., Jixing, W., Bennett, P. H. & Howard, B. V. (1994). *Effect of dietary and/or exercise on the incidence of diabetes in subjects with IGT–The Da Qing IGT and Diabetes Study.* Paper given at the 15th International Diabetes Federation, Kobe, Japan.

Peveler, R. C., Fairburn, G. C., Boller, I., & Dunger, D. (1992). Eating disorders in adolescents with IDDM. *Diabetes Care, 15,* 1356–1360.

Pierce, M., & Armstrong, D. (1996). Afro-Caribbean lay beliefs about diabetes: An exploratory study. In K. Kelleher & S. Hillier (Eds.), *Researching cultural differences in health* (pp. 97–102). London: Routledge.

Pierce, M., Keen, H., & Bradley, C. (1995). Risk of diabetes in offspring of parents with non-insulin-dependent diabetes. *Diabetic Medicine, 12,* 6–13.

Pipeleers, D., Keymeulen, B., & Korbutt, G. (1994). Islet transplantation. In S. M. Marshall & P. D. Home (Eds.), *The Diabetes Annual/8* (pp. 299–330). Amsterdam: Elsevier.

Posner, T. R. (1992). *Draft report to BDA LOW taskforce on letters about the changeover to human insulin.* British Diabetic Association, London.

Pozzilli, P., & Andreani, D. (1993). The potential role of nicotinamide in the secondary prevention of IDDM. *Diabetes/Metabolism Reviews, 9,* 219–230.

Price, D. E. (1993). Managing impotence in diabetes. *British Medical Journal, 307,* 275–276.

Pyke, D. A. (1979). Diabetes: the genetic connections. *Diabetologia, 17,* 333–343.

Quatromoni, P. A., Milbauer, M., Posner, B. M., Carballeira, N. P., Brunt, M., & Chipkin, S. R. (1994). Use of focus groups to explore nutrition practices and health beliefs of urban Caribbean Latinos with diabetes. *Diabetes Care, 17,* 869–873.

Reaven, G. M. (1988). Banting Lecture 1988. The role of insulin resistance in human disease. *Diabetes, 37,* 1595–1607.

Riazi, A., Pickup, J., & Bradley, C. (1996). Blood glucose reactivity to stress: Individual differences in magnitude, timing, and direction of responses. *Diabetic Medicine, 13*(Suppl. 7), S5–6.

Rizza, R. A. Mandarino, L. J., & Gerich, J. E. (1982). Cortisol induced insulin resistance in man: Impaired suppression of glucose utilization due to a postreceptor defect of insulin action. *Clinical Endocrinology and Metabolism, 54,* 131–138.

Robertson, R. P. (1992). Pancreatic and islet transplantation for diabetes—cures or curiosities? *New England Journal of Medicine, 327,* 1861–1868.

Rodin, G. M., & Daneman, D. (1992). Eating disorders and IDDM: A problematic association. *Diabetes Care, 15,* 1402–1412.

Rodin, G. M., Daneman, D., Johnson, L. E., Kenshole, A., & Garfinkel, P. (1985). Anorexia nervosa and bulimia in female adolescents with insulin dependent diabetes mellitus: A systematic study. *Journal of Psychiatric Research, 19,* 381–384.

Rosenbaum, L. (1983). Biofeedback-assisted stress management for insulin-treated diabetes mellitus. *Biofeedback and Self-Regulation, 8,* 519–532.

Rosmark, B., Berne, C., Holmgren, S., Lago, C., Renholm, G., & Sohlberg, S. (1986). Eating disorders in patients with insulin dependent diabetes mellitus. *Journal of Clinical Psychiatry, 47,* 547–550.

Ryan, C. M. (1994). Measures of cognitive function. In C. Bradley (Ed.), *Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research* *and practice* (pp. 191–219). Chur, Switzerland: Harwood Academic Publishers.

Saad, M. F., Knowler, W. C., Pettitt, D. J., Nelson, R. G., Mott, D. M., & Bennett, P. H. (1988). The natural history of impaired glucose tolerance in the Pima Indians. *New England Journal of Medicine, 319,* 1500–1506.

Sartor, G., Schersten, B., Carlstrom, S., Melander, A., & Persson, G. (1980). Ten-year follow-up of subjects with impaired glucose tolerance: Prevention of diabetes by tolbutamide and diet regulation. *Diabetes, 29,* 154–157.

Seeburg, K. N., & DeBoer, K. F. (1980). Effects of EMG biofeedback on diabetes. *Biofeedback and Self-Regulation, 5,* 289–293.

Smith, U. (1984). Insulin resistance in obesity, Type II diabetes and stress. *Acta Endocrinologica, 105*(Suppl. 262), 67–69.

Stabler, B. (1993). On the role of patient. *Journal of Pediatric Psychology, 18,* 301–312.

Steel, J. M., Young, R. J., Lloyd, G. G., & Macintyre, C. C. A. (1989). Abnormal Eating attitudes in young insulin-dependent diabetics. *British Journal of Psychiatry, 155,* 515–521.

Surwit, R. S., & Feinglos, M. N. (1983). The effects of relaxation on glucose tolerance in non-insulin-dependent diabetes. *Diabetes Care, 6,* 176–179.

Sutherland, D. E. R., Gores, P. F., Hering, B. J., Wahoff, D., McKeehen, D. A., & Gruessner, W. G. (1996). Islet transplantation: An update. *Diabetes/Metabolism Reviews, 12,* 137–150.

Taylor, R., Foster B., Kyne-Grzebalski, D., & Vanderpump, M. (1994) Insulin regimens for the non-insulin dependent: Impact on diurnal metabolic state and quality of life. *Diabetic Medicine, 11,* 551–557.

Taylor, R., Ram, P., Zimmet, P., Raper, L. R., & Ringrose, H. (1984). Physical activity and prevalence of diabetes in Melanesian and Indian men in Fiji. *Diabetologia, 27,* 578–582.

Todd, J. A. (1995). Genetic analysis of Type 1 diabetes using whole-genome approaches. *Proceedings of the National Academy of Science USA, 92,* 8560–8065.

Todd, C., Bradley, C., & Symonds, L. (1993). Psychosocial measurement in the audit of diabetes services. *Audit Trends, 1,* 141–143.

Toobert, D. J., & Glasgow, R. E. (1994). Assessing diabetes self-management: The summary of Diabetes Self-Care Activities Questionnaire. In C. Bradley (Ed.), *Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice* (pp. 351–375). Chur, Switzerland: Harwood Academic Publishers.

UK Prospective Diabetes Study Group VII (1990). Response of fasting plasma glucose to diet therapy in newly presenting type II diabetic patients. *Metabolism, 39,* 309–312.

UK Prospective Diabetes Study Group VIII (1991). Study design, progress and performance *Diabetologia, 34,* 877–890.

Vandewalle, C. L., Decraene, T., Schult, F. C., De Leeuw, I. H., Pipeleers, D. G., & Gorus, F. K. (1993). Insulin antibodies and high titre islet-cell antibodies are preferentially associated with the HLA DQA1 *0301-DQB1*0302 haplotype at clinical onset of Type 1 (insulin-dependent) diabetes mellitus before age 10 years, but not at onset between age 10 and 40 years. *Diabetologia, 36,* 1155–1162.

Verge, C. F., Gianani, R., Kawasaki, E., Yu, L., Pietropaolo, H., Jackson, R. A., Chase, H. P., & Eisenbarth, G. S. (1996). Prediction of Type 1 diabetes in first-degree relatives using a combination of insulin-, GAD- and ICA512 bdc/IA-2 autoantibodies. *Diabetes, 45,* 926–933.

Veves, A., Webster, L., Chen, T. F., Payne, S., & Boulton, A. J. (1995). Aetiopathogenesis and management of impotence in diabetic males: 4 years experience from a combined clinic. *Diabetic Medicine, 12,* 77–82.

Viberti, G. C., Walker, J. D., & Pinto J. D. (1992). Diabetic nephropathy. In K. G. M. M. Alberti, R. A. De Fronzo, H. Keen, & P. Zimmet (Eds.), *International textbook of diabetes mellitus* (pp. 1267–1328). Chichester, UK: Wiley.

Wang, P. H., Lau, J., & Chalmers, T. C. (1993). Metaanalysis of effects of intensive blood-glucose control on late complications of Type I diabetes. *Lancet, 341,* 1306–1309.

Will, J. C. (1995). Self-reported weight loss in adults with diabetes: Results from a national health survey. *Diabetic Medicine, 12,* 974–978.

Wilson, A., Home, P. D., Bishop, A., Bradley, C., Brown, K. J. E., Hargreaves, B. & members of the working group of the Research Unit of the Royal College of Physicians and British Diabetic Association (1993). A dataset to allow exchange of information for monitoring continuing diabetes care. *Diabetic Medicine, 10,* 378–390.

Wing, R. R., & Anglin, K. (1996). Effectiveness of a behavioral weight control program for blacks and whites with NIDDM. *Diabetes Care, 19,* 409–413.

Wing, R. R., Epstein, L. H., Nowalk, M. P., Koeske, R., & Hagg, S. (1985). Behavior change, weight loss, and physiological improvements in type II diabetic patients. *Journal of Consulting and Clinical Psychology, 53,* 111–122.

Wing, R. R., Shoemaker, M., Marcus, M. D., McDermott, M., & Gooding, W. (1990). Variables associated with weight loss and improvements in glycemic control in type II diabetic patients in behavioral weight control programs. *International Journal of Obesity, 14,* 495–503.

World Health Organization (1985). *WHO study group on diabetes mellitus. Second report* (WHO Technical Report Series, No 727). Geneva, Switzerland: Author.

World Health Organisation Regional Office for Europe (1993). Consensus Meeting on Quality Assurance Indicators in Mental Health Care. Stockholm, August, 1993.

World Health Organisation Regional Office for Europe (1995). Consensus Meeting on Quality Assurance Indicators. Stockholm, November, 1995.

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 67,* 361–70.

Zimmet, P. Z. (1992). Kelly West Lecture 1991: Challenges in diabetes epidemiology—from West to the rest. *Diabetes Care, 15,* 232–252.

Zimmet, P. Z. (1995). The pathogenesis and prevention of diabetes in adults. *Diabetes Care, 18,* 1050–1064.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.12

Functional Bowel Disorders

SHANNON M. TURNER

Utah State University, Logan, UT, USA

8.12.1 INTRODUCTION 305

8.12.2 INFLAMMATORY BOWEL DISEASE AND PEPTIC ULCER DISEASE 306

8.12.3 IRRITABLE BOWEL SYNDROME AND FUNCTIONAL BOWEL DISORDER 306
- *8.12.3.1 Classification of IBS* 307
- *8.12.3.2 Classification of the Functional Bowel Disorders* 308
- *8.12.3.3 Epidemiology of IBS and the Functional Bowel Disorders* 308
- *8.12.3.4 Proposed Etiological/Assessment Issues: Physiological, Psychological and Psychosocial Determinants of IBS and FBD* 309
  - *8.12.3.4.1 Proposed physiological mechanisms for IBS and the functional bowel disorders* 309
  - *8.12.3.4.2 Proposed psychological/sociocultural explanations for IBS and the functional bowel disorders* 310

8.12.4 ASSESSMENT 312
- *8.12.4.1 Daily Symptom Diaries* 312
- *8.12.4.2 Assessing Psychological and Psychosocial Distress* 312
- *8.12.4.3 Assessing Psychiatric Disorder* 312
- *8.12.4.4 Assessing Health Impact of the Functional Bowel Disorders* 312
- *8.12.4.5 Summary* 313

8.12.5 TREATMENT AND/OR MANAGEMENT OF IBS AND THE FUNCTIONAL BOWEL DISORDERS 313
- *8.12.5.1 Medical Treatments for IBS* 313
  - *8.12.5.1.1 Summary* 313
- *8.12.5.2 Psychological and Psychophysiological Management of IBS and FBD* 314
  - *8.12.5.2.1 Brief psychodynamic psychotherapy* 314
  - *8.12.5.2.2 Hypnotherapy* 314
  - *8.12.5.2.3 Cognitive-behavioral approaches* 314
  - *8.12.5.2.4 Summary* 316

8.12.6 PROFESSIONAL ISSUES—IBS AND THE FUNCTIONAL BOWEL DISORDERS 316

8.12.7 SUMMARY 317

8.12.8 REFERENCES 317

8.12.1 INTRODUCTION

Although gastrointestinal (GI) disorders are highly prevalent, at least in the USA and in other Western countries, only a few of these have received attention from clinical psychologists. Among this limited group of GI disorders, those of a primarily functional nature, specifically irritable bowel syndrome, functional bowel disorders, nonulcer dyspepsia, gastroesophageal reflux, and (sometimes) fecal incontinence, are often refractory to usual medical care and thus cause the medical clinician great difficulty (Drossman, 1995). The economic costs of such disorders are large, particularly given sufferers' significant levels of healthcare consumption, quantities purchased of ineffective, over-the-counter treatments (Talley, Gabriel, Harmsen, Zinsmeister, & Evans 1995; Thompson, 1991), and absenteeism (Drossman et al., 1993). For these reasons, the present chapter will devote most of its coverage to these disorders. However, the psychological aspects of two organic conditions, peptic ulcer disease (PUD) and inflammatory bowel disease (Crohn's disease and ulcerative colitis), will be briefly addressed, given these disorders' deceptively similar presentation to irritable bowel syndrome and other functional bowel disorders, historical (albeit incorrect) conceptualization as "psychosomatic disorders," and the utility of psychiatric/behavioral techniques in their management.

8.12.2 INFLAMMATORY BOWEL DISEASE AND PEPTIC ULCER DISEASE

Inflammatory bowel disease (IBD) is a descriptive label which subsumes two painful and sometimes life-threatening conditions, Crohn's disease and ulcerative colitis. Crohn's disease is best characterized by inflammation that can penetrate the layers of the entire gastrointestinal tract, often resulting in bowel obstruction and/or infection of extraintestinal systems (Olden, 1992). Ulcerative colitis manifests similarly, but involves only the mucosa of the large bowel. Like IBS, both forms of IBD follow a chronic, episodic course, with symptoms including severe abdominal pain, loss of appetite, vomiting, and diarrhea. Complications of exacerbations can include electrolyte imbalance and anemia due to vomiting, anorexia, bloody diarrhea, and mineral malabsorption. Other extraintestinal effects include, but are not limited to, arthritis, carcinoma of the bile ducts, skin lesions, osteoporosis/osteomalacia, and increased risk for the development of cancers of the GI tract (see Olden, 1992).

Following the publication of a single case report in 1930 documenting an individual's development of ulcerative colitis during an interpersonally stressful period, a body of literature suggesting a psychosomatic etiology for this disease has developed (Murray, 1930; North & Alpers, 1994). A similar situation occurred regarding Crohn's disease (Blackburn, Hadfield, & Hunt, 1939). However, despite these uncontrolled reports, a consistent, causally directional relationship between "stress" or similar psychological variables, and IBD has not been effectively documented (Lewin & Lewis, 1995; North & Alpers, 1994). Although the exact etiology of both forms of IBD remains unclear, the clinical lore ascribing a causal role to psychiatric variables is based largely on clinical interpretation, rather than on statistically reliable and valid measurement (Olden, 1992).

However, while psychological variables and stress may not be etiologically related to IBD, certainly they may represent consequences of these diseases. Corticosteroids and other such medications as azulfidine, are used for long periods of time, and often have unpleasant side effects, making compliance issues relevant. Surgical intervention may be stressful and in the case of ostomy, may raise concerns about body image and intimacy (see Olden, 1992), as can skin conditions related to IBD. Psychological assessment and intervention may certainly be useful in such areas as quality of life (see Drossman, Patrick & Appelbaum, 1989; Drossman, Patrick, & Mitchell, 1989; Olden, 1992), and stress and pain management (Blanchard et al., 1992).

PUD includes ulceration of the gastric and duodenal areas, causing abdominal pain, loss of appetite, and altered bowel function. Recent research suggests that this problem is best attributable to a bacterium known as *Heliobacter pylori* (Warren & Marshall, 1983; Peura & Graham, 1994), which can be spread through oral–fecal contact (Delaney, 1995). Risk of contact with this organism is significantly related to childhood poverty (Webb, Knight, & Greaves, 1994). Pharmacological intervention is currently the most effective tool for PUD (see Delaney, 1995), which raises the possibility for behaviorally focused intervention.

Like IBS, PUD also follows a chronic, relapsing course, and until the discovery of *Heliobacter pylori*, psychiatry had long held that PUD was strongly related to prolonged stress (see Whitehead, 1992a), despite the lack of consistently supportive empirical data, and failure of psychological intervention to affect PUD symptoms. However, behavioral treatments focused on medication compliance may represent an effective intervention for PUD, given the often complicated pharmacological regimen involved in eradication of this bacterium (Delaney, 1995). Further empirical investigation is suggested.

In short, although psychological etiology in both IBD and PUD have not been supported empirically, this does not rule out a potential role for clinical psychologists in the assessment and intervention of these diseases. Assessment of stress, quality of life, sexual functioning, body image, and medication compliance, as well as behavioral intervention in all of these areas, seem appropriate avenues for psychologists in the future.

8.12.3 IRRITABLE BOWEL SYNDROME AND FUNCTIONAL BOWEL DISORDER

By far the most prevalent of the functional GI disorders (Talley, Weaver, Zinsmeister & Melton, 1992), and the one which has been the focus of the most attention since the late 1980s, is the irritable bowel syndrome (IBS). Functional bowel disorders (FBD), a target of increasing research attention, have only recently been delineated from IBS and therefore will also be covered in this section.

8.12.3.1 Classification of IBS

IBS is a common, sometime disabling (Corney & Stanton, 1990), and poorly understood functional condition of the lower GI tract. It is most often characterized by "abdominal pain and change in bowel habit (diarrhea and/or constipation) occurring in the absence of abnormalities on the appropriate physical and laboratory investigations" (Latimer, 1983, p. 205). However, individuals may often suffer other GI symptoms such as bloating, flatulence, nausea, belching, abdominal distention (Kellow & Langluddecke, 1989); mucous stools, abdominal rumbling and gurgling, and loss of appetite and/or weight (Latimer, 1983). Nongastrointestinal symptoms, including headache, palpitation, dyspnea (Davis, 1980), fatigue, weakness, muscular pains (Latimer, 1983; Lynn & Friedman, 1993), and disorders of anxiety (Blanchard, Scharff, Suls, & Barlow, 1990; Clouse, 1988; Liss, Alpers, & Woodruff, 1973; Lydiard, 1992; Tollefson, Tollefson, Pederson, Luxenberg, & Dunsmore, 1991; Walker, Roy-Byrne, & Katon, 1990; Walker, Roy-Byrne, Katon, Li et al., 1990) and mood (Blanchard et al., 1990; Liss et al., 1973; Lydiard, Fossey, Marsh, & Ballenger, 1993; Walker, Roy-Byrne, & Katon, 1990; Walker, Roy-Byrne, Katon, Li et al., 1990) are also often reported.

Clearly a heterogeneous problem, IBS may also follow a chronic and relapsing course (Latimer, 1983). Harvey, Salih, and Read (1983) found that in a prospective study of 2000 outpatients, 449 who had suffered from IBS at initial contact, continued to struggle with the disorder two to five years later. Similarly, Thompson, Creed, Drossman, Heaton, and Mazzacca (1992) state that IBS will often affect individuals for a large part of their lives. These two factors certainly complicate the management of this disorder.

Understandably, the complexity of IBS has led to numerous attempts to diagnose and/or classify it appropriately. Powell, in 1820, first described IBS as "painful affectations of the intestinal canal" (in Chandhary & Truelove, 1962) Since then, 30 definitions for this disorder have been coined, all reflecting the current etiological zeitgeist. However, given consistent failure to identify specific organic and/or physiologic causal variables, IBS has been defined descriptively and by exclusion. For instance, an attempt to standardize diagnostic practice was made by Kruis et al. (1984), who developed an actuarial approach that would initially separate IBS patients from those with organic GI disease by way of symptom presentation and lack of physical findings, and then further discriminate between the two groups by a weighted system utilizing historical, physical, and empirical variables, these data then being analyzed by logistic regression analysis. Although an important attempt, and one that attended to the necessity to differentiate IBS from more deadly conditions like IBD, the approach of Kruis et al. (1984) has been criticized for its lack of clinical utility, given the heterogeneity of IBS (Blanchard, 1993; Camilleri & Prather, 1992; Drossman, 1984), as well as the growing need to avoid expensive and unnecessary medical tests and procedures (Talley, Phillips, Melton et al., 1990).

In 1988, the Thirteenth International Congress of Gastroenterology met in Rome, Italy, to attempt a more positive diagnostic classification of IBS. Their decision was based in part on the work of Manning and his colleagues (Manning, Thompson, Heaton, & Morris, 1978), who had compared the prevalence of 15 symptoms in patients with IBS to that of individuals suffering from organic GI disease, and found that six primary symptoms (i) visible abdominal distention, (ii) pain relieved by defecation, (iii) more frequent stools with pain onset, (iv) looser stools with pain, (v) passage of mucus, and (vi) a subjective sensation of incomplete evacuation, differentiated IBS from bowel disease. The Rome Criteria are as follows: at least three months continuous or recurrent symptoms of:

(i) Abdominal pain or discomfort which is
(a) relieved by defecation,
(b) and/or associated with a change in frequency of stool,
(c) and/or associated with a change in consistency of stool; and
(ii) Two or more of the following, on at least a quarter of occasions or days:
(a) altered stool frequency (either more than three bowel movements per day or fewer than three bowel movements per week,
(b) altered stool form (lumpy/hard or loose/watery stool),
(c) altered stool passage (straining, urgency, or feeling of incomplete evacuation),
(d) passage of mucus,
(e) bloating or feeling of abdominal distention.

The "Rome Criteria," as they have come to be known, are currently the most commonly accepted and utilized for IBS. However, as noted by Talley, Phillips Melton et al. (1990), few studies have been conducted to test their validity. The aforementioned group assessed 361 gastroenterological outpatients presenting to the Mayo Clinic with both a validated bowel symptom questionnaire and a full clinical evaluation. These individuals were retested for reliability. Logistic regression analysis indicated that as the number of positive diagnostic criteria increased, so did the probability of IBS. However, the sensitivity was somewhat low. In short, a positive, sympomatic diagnostic approach based on the Rome Criteria seems quite useful and important.

8.12.3.2 Classification of the Functional Bowel Disorders

Since the classification of IBS in 1988, another team has constructed a similar system for all the functional gastrointestinal disorders, and has placed IBS in the category of the "functional bowel disorders." Thompson et al. (1992) define these disorders, propose diagnostic criteria, and discuss their differentiation from IBS.

Functional abdominal bloating (FAB) is a disorder characterized by a feeling of fullness, bloating and/or abdominal distention. Audible bowel sounds and expulsion of gas may also be present. Symptoms must have been present for at least three months, and must not be related to dietary factors. Individuals with FAB do not meet diagnostic criteria for IBS or other functional bowel disorders.

Functional constipation (FC) is defined as "persistent symptoms of difficult, infrequent, or seemingly incomplete defecation" (Thompson et al., 1992, p. 80), which has no other apparent etiology. Symptoms must be present again, for at least three months, individuals must report straining upon defecation at least 25% of the time, lumpy or hard stools, a sensation of incomplete evacuation, and two or fewer bowel movements in a week. Abdominal pain is not present, and individuals do not meet criteria for IBS.

Functional diarrhea (FD), defined by Thompson et al. (1992) as "the frequent and/or urgent passage of unformed stool" (p. 83), must include two or more of the following symptoms present for at least two months: unformed stool passage 75% of the time; three or more bowel movements per day, at least half of the time; and increased stool weight (for details, see Thompson et al., 1992). Patients do not report abdominal pain or hard or lumpy stools. They experience urgency and may occasionally soil themselves.

Unspecified bowel disorders include combinations of symptoms that neither fit into the aforementioned categories, nor suggest organic disease.

Lastly, functional abdominal pain (FAP), is described as six months of ongoing pain that is experienced in the abdomen, but bears no relationship to the function of the GI tract, or other physiological systems (e.g., reproductive system, etc.). As with all functional GI disorders, organic disease is conspicuously absent.

Thompson et al. acknowledge the significant overlap of reported GI symptoms in many of the aforementioned disorders; however, such differentiation is necessary for research into, and development of, effective and appropriate management strategies for all of these, given their prevalence.

8.12.3.3 Epidemiology of IBS and the Functional Bowel Disorders

Although epidemiological estimates have been made by way of surveying outpatients reporting to gastroenterology clinics, the contention that most individuals suffering from bowel symptoms do not seek medical attention for them has necessitated general population studies of both IBS and the functional bowel disorders. Talley, Zinsmeister, Van Dyke, and Melton (1991) conducted such a questionnaire study of the Caucasian, middle-aged population of Olmstead County, Minnesota. They found that the prevalence of IBS, identified by the Rome Criteria, was approximately 17%. Contrary to most published reports, this investigation did not uncover a gender difference in bowel symptom reporting. The prevalence of functional bowel disorders other than IBS was as follows: functional abdominal pain 26%, chronic constipation 17.4%, and chronic diarrhea 17.9%. Interestingly, only 9% of individuals reporting such symptoms had sought medical attention in the year prior to the study's inception, and neither pain severity, duration, or frequency accounted for such help-seeking.

Similarly, though on a somewhat larger scale, Drossman et al. (1993) examined the prevalence, sociodemography and health impact of functional GI disorders in a stratified probability random sample of US homeowners. Utilizing the Rome Criteria and the newly developed criteria for functional bowel disorders mentioned above, the authors found that 69% of the entire sample (5430 responders) reported having experienced at least one functional bowel syndrome in the previous three months, and based on the data received, estimated that 9% of the nation may suffer from IBS, 3.0% from functional constipation, 1.6% from functional diarrhea, 32.1% from functional abdominal bloating, and 2% from chronic functional abdominal pain. In contrast with the aforementioned report by Talley, the estimated prevalence rate of IBS was low, which the authors attribute to their use of stringent diagnostic criteria. Also, the present group found that females were indeed more likely to report IBS, in addition to functional constipation and functional abdominal pain, than were men, who more often reported functional bloating. Lastly, Drossman et al. (1993) found that females were more likely than males to seek medical assistance, whereas Talley et al. (1991) did not assess this variable.

Earlier epidemiological data collected in Europe and the Far East suggest a similar prevalence (Drossman & Thompson, 1992) to that estimated by Talley et al. (1991), but studies conducted in some areas of Southeast Asia and even rural North America report a lower rate, one more similar to that of Drossman et al. (1993). Such a discrepancy suggests that GI symptoms are very prevalent and many more suffer privately than seek treatment.

Regarding gender, in Western countries, the typical IBS patient appears to be female and under 35 years of age (Drossman & Thompson, 1992), but in India, for example, most individuals presenting to physicians for treatment of IBS are male, suggesting that IBS (or at least its report) may be in part, culturally determined, an issue that will be more broadly discussed below.

8.12.3.4 Proposed Etiological/Assessment Issues: Physiological, Psychological, and Psychosocial Determinants of IBS and FBD

The two epidemiological studies discussed above illustrate the improved status of current IBS and FBD research, as definitions/diagnoses of these syndrome(s) are more positive and symptom-determined, and no longer necessarily reflective of etiological viewpoints, which have yet to reflect a consensus in the field. The following sections will critically review the prominent etiological theories of IBS and, when appropriate, the functional bowel disorders (many of these are the same or very similar and fall under two main categories, physiological and psychological/psychosocial). Important assessment issues will then be addressed, including daily symptom tracking, the measurement of psychological distress and/or disorder, and the necessity to conduct a thorough psychosocial history.

8.12.3.4.1 Proposed physiological mechanisms for IBS and the functional bowel disorders

(i) Abnormal gut motility

According to Whitehead (1992b), increased motility in the intestinal tract is characteristic of most IBS patients. However, this motility is neither consistent, nor is it necessarily abnormal. Griffin (1985) summarized a number of different bowel motility patterns that have been observed in IBS. For example, cramping abdominal pain is often associated with large and small intestinal changes in the duration and frequency of increased intraluminal pressure and contractions; parasympathomimetic agents tend to provoke increased colonic muscle activity; increased resting colonic muscle activity is observed in diarrhea-predominant IBS, and the opposite pattern is often observed when constipation predominates. A representative study of motility dysfunction was conducted by Latimer et al. (1981), wherein they examined the basal motor activity (BMA) of the distal colon in diarrhea-predominant, constipation-predominant, and normal subjects. Findings were equivocal, in that although significant motility variation was found in the IBS groups, the normal group also exhibited motility disruption in the 15 cm segment of distal colon.

Similarly, motility dysfunction has been studied with slow-wave myoelectrical activity as the dependent variable. Snape, Carlson, and Cohen (1976) found that IBS patients' slow-wave activity in the distal colon was increased by three cycles per minute, compared to normals. Lynn and Friedman (1993) describe such activity as correlated with "nonpropulsive segmenting contractions of the colon [which] might result in constipation" (p. 1940). Again, though, in a study conducted by Madhu et al. (1988), such findings were not supported, as no significant differences in colonic motility were observed between IBS patients and normal controls.

Given that the functional bowel disorders have only recently been differentiated from IBS, research investigating functional diarrhea and/or functional constipation, as delineated here, has yet to be conducted. The few studies that do exist echo the results of the IBS research: support is mixed for dysfunctional motility.

(ii) Anxiety and the "brain–gut link"

Fossey and Lydiard (1990) have hypothesized that neuroanatomical connections between the hypothalamus, the "origin of information transmitted to the gut via parasympathetic and sympathetic pathways" (p. 178), and the GI tract may be indirectly responsible for IBS symptoms, through "primary pathology at the CNS level [which] may secondarily impact on the gastrointestinal system via neurohumoral processes, resulting in gastrointestinal distress" (p. 179). Unfortunately, however, it is unclear to what the authors refer by "primary pathology at the CNS level." A similar theory has been put forth by others, including Clouse (1988), who suggested that the "enteric nervous system (ENS)," made up of the nerves of the gut, is often considered the "third division" of the autonomic nervous system (ANS), and appears to function accordingly (p. 412). Walker, Roy-Byrne, Katon, Li et al. (1990) and Lydiard (1992) concur, and suggest that because high activity of the locus coeruleus, a pontine noradrenergic nucleus, has been correlated with "vigilance and attention to novel or fear-provoking stimuli, [and] less activation with ... behaviors such as sleep, grooming, and feeding " (Walker, Roy-Byrne, Katon, Li et al. 1990b, p. 569), this supports the notion that IBS is related to the physiological aspects of clinical anxiety. The locus coeruleus receives afferent input from the viscera (Lydiard, 1992) and is directly innervated by the medullary nucleus solitarius, which is known to communicate with the GI system (Walker, Roy-Byrne, Katon, Li et al. 1990b), potentially providing the "missing link" between IBS and the body's protective "fight or flight" response which, when firing inappropriately, can be experienced as anxiety (Barlow, 1988).

Albeit somewhat reductionistic and not sufficiently explanatory, given that not all IBS patients experience clinical anxiety, the "brain–gut link" theory represents a promising area for future research. What follows are some of the prominent psychological/sociocultural theories of the etiology of IBS.

8.12.3.4.2 Proposed psychological/sociocultural explanations for IBS and the functional bowel disorders

Psychological theories of IBS have been divided into five main sections: abnormal pain perception, specific historical variables, maladaptive psychological or behavioral characteristics, the presence of psychological disturbance or disorder, and stress.

(i) Abnormal pain perception

Early studies of lower intestinal function found that colonic distention produced pain (Bloomfield & Polland, 1931). Ritchie (1973) found that balloon insertion and inflation into the colons of both IBS patients and healthy but "constipated" control subjects produced significantly different pain perception in the IBS patients than in the controls. However, a study conducted by Latimer et al. (1979), which included a group of psychological controls without bowel symptoms, failed to find such between-group differences. Support for the colonic hyperalgesia hypothesis, therefore, is equivocal.

(ii) Historical variables: physical and sexual abuse history

Another potentially etiologically important issue in IBS and other functional gastrointestinal disorders is sexual and physical abuse history. A representative study was conducted by Walker, Katon, Roy-Byrne, Jemelka, and Russo (1993), who assessed both IBS and IBD patients with a structured psychiatric interview and a sexual trauma interview. The results indicated that the prevalence of severe lifetime sexual abuse was significantly higher in IBS patients (32% vs. 0%), as was any sexual abuse (54% vs. 5%). However, prevalence estimates vary significantly across studies and sites of investigation. For a recent review of this literature, see Drossman, Talley, Leserman, Olden, and Barreiro (1995).

(iii) Maladaptive psychological and/or behavioral characteristics

Research suggesting psychological and/or behavioral characteristics as explanatory for IBS and other psychophysiological disorders encompasses a very large area, including personality, specifically the traits of "neuroticism" (Latimer, 1983) and negative affectivity (NA) (Watson & Pennebaker, 1989); depressive and anxious symptomology (Blanchard et al., 1986); and abnormal illness behavior, a particularly interesting concept first defined by Whitehead, Winget, Fedoravicius, Wooley, and Blackwell (1982) as "frequent visits to physicians, multiple somatic complaints, and disability disproportionate to physical findings" (p. 202). Abnormal illness behavior has since been expanded upon and studied, due to its potential relevance to IBS. Gomborone, Dewsnap, Libby, and Farthing (1995) assessed IBS patients, individuals with organic GI disease, depressed individuals, and healthy controls with the Illness Attitudes Scale (IAS), a measure of abnormal illness attitudes, as well as the Beck Depression Inventory (BDI). They found that BDI scores of the IBS patients were significantly higher than those of patients with organic GI disease and healthy controls, and that IBS patients' IAS scores were also significantly elevated, particularly on such subscales as bodily preoccupation, hypochondriacal beliefs, and disease phobia. The authors concluded that IBS patients exhibit abnormal attitudes that are not related solely to depression or to the experience of physical symptoms, but reflect distress about illness disproportionate to reality.

(iv) Psychological disturbance or disorder

The above data document the apparent existence of abnormal psychological/behavioral characteristics in many IBS patients. Another area of research has implicated diagnosable psychological/psychiatric disorder in the etiology of IBS. It has been estimated that 50 to 100% of IBS patients meet diagnostic criteria for some psychiatric illness (Blanchard, 1993). However, for psychiatric illness to be etiologically connected to IBS, it would need to have predated or have coincided with GI symptom onset. Walker, Roy-Byrne, Katon, Li et al. (1990) investigated the prevalence of psychiatric illness as measured by structured interview in IBS patients and in those with inflammatory bowel disease (IBD). They found that IBS patients had suffered significantly more lifetime diagnoses of major depression, generalized anxiety disorder, panic disorder, somatization disorder, and phobia, than had the IBD patients. The majority of these individuals had experienced the psychiatric disorders prior to the onset of GI symptoms. Blanchard et al. (1990) found similar results, with IBS patients suffering significantly more anxiety disorders than individuals with IBD. However, not all investigations have found this to be the case. For example, Thornton, McIntyre, Murray-Lyon, and Gruzelier (1990) found that the majority of their sample did not suffer from psychological disorder. Again, although the bulk of the evidence suggests that psychological distress and/or disorder is etiologically related to IBS, it does not sufficiently explain its occurrence.

(v) Stress

The study of "stressful life events" and/or "stress" in the etiology of IBS is not new, and given that many IBS patients report that stressful events exacerbate their symptoms and may have even precipitated them (Chaudary & Truelove, 1962) it is an extremely important area of research. Arun, Kanwal, Vyas, and Sushil (1993) found that IBS patients recalled significantly more stressful life events than did healthy controls. The results of Craig and Brown (1984) indicated that in their sample, stressful life events predated the onset of symptoms in individuals with functional gastrointestinal disorders. A recent, well-designed study by Whitehead, Crowell, Robinson, Heller, and Schuster (1992) found that in a female sample unselected for bowel symptoms, individuals who met the Rome Criteria for IBS experienced significantly more stressful life events (as measured by a standardized inventory) than did subjects who experienced what can now be defined as functional bowel disorders other than IBS, or healthy controls. Whitehead's group also controlled for "neuroticism," which has also been implicated in IBS (Latimer, 1983; Blanchard et al. 1990), strengthening their findings.

Using a different conceptualization of "stress," subjects" cognitive appraisal of situational demands relative to their coping abilities, DeLongis, Coyne, Dakof, Folkman, and Lazarus (1982) assessed individuals with the "Hassles and Uplift Scale," a measure of the *distress* occasioned by daily events, and not their frequency, and found that this distress was more closely related to somatic health than were stressful life events. Similarly, Jandorf, Deblinger, Neale, and Stone (1987) found that the perceptions of daily events as negative and/or stressful strongly predicted physical symptoms in healthy subjects, even more so than in adults with pre-existing medical conditions, lending credence to the notion that one's *experience* of daily life as stressful affects one's physical health (or at least its perception).

These data appear supportive of a stress hypothesis in the etiology of IBS. However, as noted in an important editorial by Whitehead (1994) the distinction between "stressful life events" and "stress" reflects the field's conceptual differences, which have made it difficult to draw clear conclusions about the relationship between stress and IBS symptoms. A study conducted by Suls, Wan, and Blanchard (1994) attempted to address some of these differences by examining the relationship between stress and GI symptoms prospectively, using a daily diary format to assess both. Interestingly, their results indicated no relationship. However, Whitehead suggests that in their attempt to guard against a Type I error, Suls et al. (1994) may have committed a Type II error by rejecting the stress-symptom hypothesis prematurely, and that given the significant amount of clinical evidence in support of this hypothesis, as well as its theoretical basis for the treatment of IBS, further investigation should occur.

(vi) Summary

In summary, extensive empirical investigation has revealed that neither motility dysfunction nor individual difference/psychological variables have emerged solely responsible for IBS. Rather, the findings have revealed an apparently *interactive* relationship between colonic and external stimuli. Thus far, the body of findings regarding dysfunctional motility, the possible "brain–gut link," learned illness behavior, historical variables, psychological characteristics, stress and other biobehavioral variables, including eating (Harvey, Salih, & Read, 1983) and fiber intake (Payler, Pomare, Heaton, & Harvey, 1975), suggests that the clinician must be sensitive to the interaction of both the biological and behavioral systems, and treat IBS as a *psychophysiological* disorder. This implies that in order adequately to assess and treat this disorder, the approach should be comprehensive, assessing physiological and psychosocial factors.

As mentioned above, examination of the etiological role of these variables in the functional bowel disorders has not yet been undertaken, given their relatively recent distinction from IBS, and is therefore suggested.

8.12.4 ASSESSMENT

8.12.4.1 Daily Symptom Diaries

In their 1994 investigation, Suls et al. utilized daily symptom diaries, an immeasurably important part of the IBS assessment and treatment process. These diaries are organized such that once a day, from pretreatment through post-treatment and sometimes follow-up, patients are asked to rate on a five-point scale, the severity and distress occasioned by the following symptoms: abdominal pain, abdominal tenderness, diarrhea, constipation, bloating, flatulence, belching, and nausea. The scale ranges from zero to four, with the following accompanying descriptive headings: not a problem; mild severity and distress; moderate severity and distress; severe severity and distress; and debilitating severity and distress. Patients are also asked to document the daily number of successful bowel movements and unsuccessful attempts to defecate, whether or not abdominal problems caused them to avoid certain food or activities, and any medications taken for IBS symptoms. According to Blanchard (1993), such diaries serve an important function for both patient and clinician, in that they allow both to assess bowel patterns and progress throughout treatment. Both empirical and practical utilization of this assessment methodology is also encouraged with the functional bowel disorders.

8.12.4.2 Assessing Psychological and Psychosocial Distress

Given the aforementioned relationship between stress and GI symptoms, it is also suggested that patients keep a diary of daily stress similar to the GI symptom diary. However, the assessment burden must obviously be adjusted according to the patient's circumstances. As noted in Blanchard (1993), assessment of IBS patients with self-report inventories of depressive and anxious symptomatology, as well as daily hassles, life experiences, and perhaps learned illness behavior may be instructive, and may also serve as another way of assessing progress throughout treatment. Similarly, assessing these variables may aid in the growing understanding of the functional bowel disorders.

8.12.4.3 Assessing Psychiatric Disorder

Lastly, given the estimated prevalence of comorbid diagnosable psychiatric disorders (generally assessed by the fourth edition of the *Diagnostic and statistical manual of mental disorders* [*DSM-IV*; American Psychiatric Association, 1994]), in individuals with IBS (and possibly persons suffering from the functional bowel disorders), assessment for these (particularly the anxiety disorders) is also critical. Blanchard (1993) recommends the use of a structured psychiatric interview, particularly the Anxiety Disorders Interview Schedule-IV (ADIS-IV; DiNardo, Brown & Barlow, 1994).

8.12.4.4 Assessing Health Impact of the Functional Bowel Disorders

Although neither IBS nor the functional bowel disorders are life-threatening conditions, their impact on sufferers' lives can be significant. The epidemiological study conducted by Drossman et al. (1993) reported health impact data. The authors defined "health impact" operationally by work absenteeism and physician visits, both gastrointestinal and nongastrointestinal. Their results indicated that work absenteeism in persons with functional GI disorders ranged from 29.4% for fecal incontinence to 6.3% for functional abdominal bloating, and was significantly higher than absenteeism for those without functional GI diagnoses (4.2%). Similarly, the mean number of annual GI-related physician visits for individuals with functional GI disorders ranged from 3.53 for fecal incontinence to 0.49 for bloating. Non-GI-related visits were also high, ranging from 9.29 for those with fecal incontinence to 2.45 visits, made by those with functional diarrhea. Again, the mean number of physician visits for individuals with functional GI disorders was significantly higher than that of those without such diagnoses. Also, functional disorders associated with significant pain were generally associated with greater absenteeism and more physician visits than were disorders that did not include pain. Therefore, it seems that having a functional GI disorder impacts well-being in the areas of occupational functioning and perceived health.

A unique study conducted in 1987 by Guthrie, Creed, and Whorwell investigated sexual dysfunction in women with IBS. Their results indicated that 83% of the women with IBS reported sexual problems, compared to 30% of women with IBD, and only 16% of those with duodenal ulcers. The authors concluded that the mechanisms for this association are still not known, but in their sample, sexual dysfunction was significantly problematic.

8.12.4.5 Summary

Given the interaction of physiological, psychological, and psychosocial factors in IBS, it is necessary to conduct a comprehensive assessment of not only bowel symptoms, utilizing the most recent and standardized criteria, but psychological symptoms, learning history, historical variables, psychiatric disorder, and variables depicting impact on health and occupational/social functioning, in an effort to understand the patient most clearly, and thus formulate an appropriate treatment plan. This section briefly reviewed these issues in IBS and the functional bowel disorders. The following section reviews the treatment literature for IBS and the functional bowel disorders, focusing on both medical and psychological treatments.

8.12.5 TREATMENT AND/OR MANAGEMENT OF IBS AND THE FUNCTIONAL BOWEL DISORDERS

8.12.5.1 Medical Treatments for IBS

Given the enormous volume of studies conducted to find effective treatments for this common and refractory problem, examining the treatment literature can be quite daunting and seemingly fruitless. For the purposes of the present discussion, medical treatments include pharmacological and dietary interventions, both of which have been reviewed in the past 10 years. In an effort to assess the efficacy of pharmacological intervention for IBS, Klein, in 1988, reviewed all the controlled and randomized trials of pharmacological agents published in the English language. He summarized 43 studies, after having rejected over 50 due to inadequate experimental design, laboratory-induced dependent variables, and the unpublished status of some reports. He then categorized the remaining studies into the following nine treatment types: antispasmodic/anticholinergics; anticholinergic/barbituate combinations; antidepressants; bulking agents; carminatives (peppermint oil); dopamine antagonists; opiods; tranquilizers/tranquilizer combinations; and miscellaneous treatments, including such drugs as timolol, phenytoin, and diltiazem. Although Klein did not quantitatively review this literature, his qualitative review was based on expected standard of research and years of clinical experience. Klein determined that, based on the literature reviewed, not a single study provided sufficient evidence that any particular medical treatment effectively treats IBS. He then reasoned that the ethical clinician likely should not prescribe "unproven agents to treat IBS [given that] such treatment may result in needless side effects and expense to the patient without promise of significant benefit" (p. 240). Klein then suggested that treatment of this complicated condition should likely include such elements as education, support, psychotherapy, and appropriate symptom-reducing agents.

A review conducted in 1988 by Muller-Lissner deserves mention, because of its methodological rigor, and as an exemplar of the complexity of IBS and the consequent difficulty operationally defining relevant variables. Muller-Lissner examined the effect of wheat bran, a commonly used treatment for IBS, particularly the constipation-predominant variant, on stool weight and transit time, his chosen definitions of colonic function. Muller-Lissner gathered 35 papers, 27 of which were ultimately summarized and quantitatively reviewed. The results indicated that stool weight and transit time were decreased, regardless of wheat bran treatment, implying that wheat bran cannot be considered an effective treatment for IBS.

8.12.5.1.1 Summary

These reviews illustrate two critical issues in IBS treatment research. First, the aforementioned volume of treatment literature and the failure of both pharmacological and dietary interventions as assessed by fairly rigorous review necessitate attention to a more holistic, interdisciplinary approach to the treatment/management of this disorder. Klein's recommendation that treatment include support, psychotherapy, and education, as well as appropriate symptom management, is insightful and, as the next section reviewing psychological and psychophysiological interventions will reveal, particularly appropriate.

Second, given that IBS is a common and distressing *functional* problem, it is important to define clinically useful dependent variables when assessing treatment efficacy. While stool weight and transit time are of scientific interest in the study of constipation-predominant IBS, clinically, they provide neither clinician nor patient with relevant information regarding progress in symptom management. As mentioned above by Blanchard (1993), symptom-based measurement allows for more accurate assessment, provided clients faithfully track their symptoms on a consistent, preferably daily basis.

At present this writer is unaware of published literature regarding the management of FBDs. Accordingly, delineation of clinically useful dependent variables in the study of the functional bowel disorders will likely assist clinical researchers in understanding these problems. As with IBS, it is strongly suggested that comprehensive biopsychosocial assessment of individuals with FBDs, including daily symptom recording, be undertaken.

8.12.5.2 Psychological and Psychophysiological Management of IBS and FBD

In a recent qualitative review of this literature, Blanchard and Malamood (1996) identified four effective types of psychological/psychophysiological treatment for IBS and the functional bowel disorders that meet two very important criteria: first, controlled studies exist from at least two different research groups supporting the efficacy of a particular approach; and second, the treatment has been shown to be equivalent or superior to standard medical care and/or well-documented attention-placebo conditions. The approaches that meet these criteria, according to Blanchard and Malamood (1996), are brief psychodynamic psychotherapy, hypnotherapy, pure cognitive therapy, and combinations of cognitive-behavioral techniques. These approaches will be briefly discussed here, and then will be reviewed and discussed in the context of current important health-care issues, including managed care and the need for services research.

8.12.5.2.1 Brief psychodynamic psychotherapy

Svedlund, Sjodin, Ottosson, and Dotevall (1983) reported one of the first controlled trials of psychotherapy for IBS. In 101 patients recruited from a medical clinic, they compared standard medical care (bulking agents, antispasmodic and tranquilizing medications) to standard medical care plus brief (10 sessions) "psychodynamically oriented" psychotherapy, focused on the identification of stress and emotions contributing to bowel symptoms and then the development of effective coping strategies. The results indicated that the psychotherapy condition was more effective than that of standard medical care on reducing reports of abdominal pain, bowel dysfunction, and somatic complaints. Interestingly, symptoms of anxiety and depression improved similarly in both conditions.

Guthrie, Creed, Dawson, and Tomenson (1991) compared a similar dynamic psychotherapy condition which also included progressive muscle relaxation to standard medical care (bulking agents and antispasmodics) in 102 patients, also from a medical clinic. Their results indicated that the diarrhea and abdominal pain experienced by the psychotherapy/relaxation group had improved significantly more than that of the patients receiving standard medical care.

8.12.5.2.2 Hypnotherapy

In 1984, Whorwell, Prior, and Faragher conducted a study in which they compared home-based audiotaped hypnotherapy, defined as suggestions of general relaxation and mental imagery focused on warmth and relaxation of the gut and abdominal area, specifically, to a medication placebo and physician support. The hypnotherapy group experienced significantly larger reductions in diarrhea and abdominal pain than the control group, and these results were maintained at one-year follow-up.

Replicating and extending the work of Whorwell et al. (1984), Harvey, Hinton, Gunary and Barry (1989) reported that such a technique could be utilized as effectively in groups. Covino and Frankel (1993) document its use as well.

8.12.5.2.3 Cognitive-behavioral approaches

Approaches to the management of IBS that utilize both cognitive and behavioral elements have included biofeedback (for a detailed review of the use of biofeedback in gastrointestinal disorders, see Whitehead, 1992b) and multicomponent treatments, which include such elements as relaxation training, assertiveness training, and biofeedback (Bennett & Wilkinson, 1985; Lynch & Zamble, 1989; Neff & Blanchard, 1987). All of these studies have found the psychophysiological treatments superior to symptom monitoring controls, and/or standard medical care.

Various investigators have conducted uncontrolled examinations of behavioral and cognitive-behavioral treatment packages consisting of progressive muscle relaxation, thermal biofeedback, educational information and training in stress-coping strategies. Successful pilot studies are often followed by randomized, controlled investigations. One such study, based on preliminary data collected by Neff and Blanchard (1987), was conducted with 19 IBS patients, comparing such a multicomponent treatment ($n = 10$) with a symptom-monitoring condition ($n = 9$). The results indicated that 6 of the 10 treated subjects experienced clinically meaningful relief in their GI symptom severity. When the subjects in the symptom-monitoring condition were offered treatment, over half of them improved.

In an effort to improve the cost-effectiveness of service delivery, Blanchard and Schwarz (1987) adapted this multicomponent therapy to a small-group format following a three-month symptom-monitoring baseline. Seventeen subjects participated in the trial, with 14 indicating satisfaction with the format (three subjects dropped out). The multicomponent treatment led to significant reductions in abdominal pain and diarrhea, with 9 of 14 subjects experiencing clinically significant improvement.

The next step in this research program was to control for nonspecific treatment effects by including a credible attention-placebo condition. Blanchard et al. (1992) reported two randomized, controlled trials of the multicomponent treatment package compared to both a symptom monitoring control and a pseudo-meditation/EEG alpha suppression biofeedback control group. Study 1 included 10 subjects per cell, and was then replicated with approximately 30 subjects per condition (Study 2). In neither study was the multicomponent treatment found to be superior to the meditation/alpha suppression condition, although in Study 1, such trends were evident. Individuals in both therapeutic conditions experienced significant reductions in GI symptoms (as indicated by daily GI diaries) and in psychometric measures of trait anxiety and depression (STAI and BDI). Gastrointestinal symptom change was maintained at six-month follow-up. Blanchard et al. (1992) interpreted the high attention-placebo response in these two studies to indicate that such individuals actively "converted" this condition into an effective treatment. They report that many patients commented that they had utilized the placebo condition to "relax" or "calm themselves," despite being told not to do so (p. 187). Physiological measures of hand temperature in these subjects corroborate this interpretation, as these measurements were consistent with those of highly relaxed individuals.

Blanchard, Greene, Scharff, and Schwarz-McMorris (1993), sought potentially to isolate another active treatment component in the previously well-validated multicomponent strategy. They chose to evaluate the efficacy of relaxation training in a sample of eight IBS patients. These individuals received 10 sessions of progressive muscle relaxation training combined with home practice. They were compared to eight patients who merely monitored their GI symptoms. Results indicated that the patients in the relaxation condition exhibited significantly more improvement on a composite measure of GI symptom reduction than those who only monitored their symptoms. In addition, 50% of the relaxation group were clinically improved at the end of treatment. One important caveat to these results must be noted: Forty-three percent of subjects initially randomized to the relaxation condition dropped out of this study. Blanchard et al. (1993) reported that this was the highest rate ever reported at the Center for Stress and Anxiety Disorders. They speculated that the structured, relatively impersonal nature of progressive relaxation training may have impacted these individuals negatively, given that there was little individualized, personal attention devoted to their IBS problems. However, the speculative nature of this interpretation must be stressed.

Recently, another attempt was undertaken to identify an additional active component of the multifaceted treatment package. Greene and Blanchard (1994) investigated the efficacy of intensive, individualized cognitive therapy as compared to GI symptom monitoring. They randomly assigned 20 subjects to one of the two conditions, and then treated the intervention group with 10 sessions of cognitive therapy adapted from the work of Beck (1976), Meichenbaum (1985), and Persons (1989). The therapy was focused largely on cognitive change, and process measures were included in order to discover whether or not such cognitive change was associated with GI symptom change. The results indicated that the cognitive therapy group showed significantly larger reductions in GI symptoms than that of the monitoring group. Moreover, at post-treatment, 80% of the cognitive therapy group were improved to a clinically significant degree, whereas only 10% of the symptom monitoring group had reached this level. In addition, changes in cognitive measures, particularly automatic thoughts (both positive and negative changes were in the desired direction) were significantly associated with reductions in GI symptoms. The beneficial results were maintained over three months.

Payne and Blanchard (1995) then compared individual cognitive therapy to self-help support groups, in an effort to control for nonspecific treatment effects. Thirty-four subjects were randomly assigned to either individual cognitive therapy, identical to that described in Greene and Blanchard (1994), a self-help support group, or a symptom-monitoring waitlist control group. Each condition lasted approximately eight weeks. The results indicated significantly greater gastrointestinal symptom reduction in the cognitive therapy group than in either the support or waitlist conditions. In addition, the cognitive therapy group exhibited significant reductions in psychometric measures of depression and anxiety. Importantly, the support group and cognitive therapy were equivalent in initial credibility and patient expectations. As with Greene and Blanchard, these results were maintained at three-month follow-up.

Vollmer and Blanchard (1997) compared the effectiveness of individual cognitive therapy with the same intervention adapted for small groups. This study was undertaken in an effort to replicate the results of Greene and Blanchard with a third therapist; and to develop and evaluate a more cost-effective approach to the treatment of patients with IBS. Preliminary results suggest that neither treatment group is significantly more effective than the other, but at least 60% of each group has experienced GI symptom improvement (A. Vollmer, personal communication, September 15, 1995).

van Dulmen and colleagues (van Dulmen, Fennis, & Bleijenberg, 1996) reported that in a sample of 25 IBS patients, small-group cognitive therapy significantly reduced IBS symptoms, avoidance behavior, and increased successful coping strategies in the treated group, relative to a waitlist control group. Moreover, these changes were maintained throughout 2.25 year follow-up.

8.12.5.2.4 Summary

The results from these studies suggest that IBS is treatable by psychological means, particularly cognitive-behavioral methodology. It appears that cognitive therapy, whether administered individually or in groups is significantly effective for IBS. Given this strong pattern of results, process variables inherent in this treatment modality as applied to the treatment of IBS are now being investigated.

The utility of such strategies for the FBDs is unknown at the present time, and both theoretical and applied research is strongly encouraged. Behavioral/cognitive-behavioral strategies may be useful, given their wide applicability beyond management of bowel symptoms (restructuring negative cognitive schemata, general stress management, etc.), but until empirical data are collected, this remains speculative.

8.12.6 PROFESSIONAL ISSUES—IBS AND THE FUNCTIONAL BOWEL DISORDERS

One of the most important issues to consider when assessing and treating IBS patients, is that due to the psychophysiological nature of this disorder, these individuals are likely to have either sought or received services from providers in a variety of disciplines (e.g., internal medicine, obstetrics/gynecology, gastroenterology, dietetics, psychiatry). It is important for psychologists to develop and maintain positive, interdisciplinary relationships, as open communication and collaboration between professionals will best facilitate comprehensive and efficient patient care.

Cost-effectiveness in healthcare has also become an increasingly significant issue, at least in the United States. Maintaining high-quality patient care while addressing the need to keep costs down challenges administrators and providers alike. For this reason, continued research focused on ultimately establishing an empirically validated “first-line” intervention for IBS is suggested. Training providers in many disciplines to offer such interventions in group format (or even self-help format, provided these are empirically supported) may meet this challenge.

In the same vein, although the FBDs appear similar to IBS in presentation, further research is necessary to understand etiological and other variables involved in all of these disorders. It is as yet unclear why so few patients seek treatment for IBS, and such data are still minimally available for the FBDs (see Drossman et al., 1995). Continued investigation should clarify these issues.

Treatment of the relatively recently delineated FBDs may be difficult at the present time, due to the lack of empirical data supporting any particular intervention. Prudence should therefore be exercised while research continues on the development of effective treatments. Theoretically/empirically driven intervention development is encouraged, as not only should professionals attend to cost issues, but to ethical, high-quality practice, for which we can be accountable.

8.12.7 SUMMARY

Due largely to research conducted in a variety of disciplines, knowledge about such gastrointestinal disorders as IBS, PUD, IBD, and FBD, has increased tremendously in the past 100 years. The contribution of the medical, nutritional, and behavioral sciences has greatly improved the assessment, diagnosis, classification, and management of these. However, much remains to be done. Clarification of etiological and maintaining variables, as well as the establishment of effective primary and secondary intervention strategies for all of these disorders represent important goals for clinical psychological science in the twenty-first century.

8.12.8 REFERENCES

American Psychiatric Association (1994). *Diagnostic and satistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Arun, P., Kanwal, K., Vyas, J. N., & Sushil, C. S. (1993). Life events and irritable bowel syndrome. *Indian Journal of Clinical Psychology, 20,* 108–112.

Barlow, D. H. (1988). *Anxiety and its disorders: The nature and treatment of anxiety and panic*. New York: Guilford Press.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: New American Library .

Bennett, P., & Wilkinson, S. (1985). Comparison of psychological and medical treatment of the irritable bowel syndrome. *British Journal of Clinical Psychology, 24,* 215–216.

Blackburn, G., Hadfield, G., & Hunt, A. H. (1939). Regional ileitis. *St. Bartholomew's Hospital Report, 72,* 181–224.

Blanchard, E. B. (1993). Irritable bowel syndrome: Research and clinical application. In R. J. Gatchel & E. B. Blanchard (Eds.), *Psychophysiological disorders* (pp. 23–62) Washington, DC: American Psychological Association.

Blanchard, E. B., & Malamood, H. S. (1996). Psychological treatment of irritable bowel syndrome. *Professional Psychology: Research and Practice, 27*(3), 241–244.

Blanchard, E. B., Greene, B., Scharff, L., & Schwarz-McMorris, S. P. (1993). Relaxation training as a treatment for irritable bowel syndrome. *Biofeedback and Self-Regulation, 18,* 125–132.

Blanchard, E. B., Radnitz, C. L., Evans, D. D., Schwarz, S. P., Neff, D. F., & Gerardi, M. (1986). Psychological comparisons of irritable bowel syndrome to chronic tension and migraine headache and nonpatient controls. *Biofeedback and Self-regulation, 12,* 221–230.

Blanchard, E. B., Scharff, L., Schwarz, S. P., Suls, J. M., & Barlow, D. H. (1990). The role of anxiety and depression in the irritable bowel syndrome. *Behavior Research and Therapy, 28,* 401–405.

Blanchard, E. B., & Schwarz, S. P. (1987). Adaptation of a multicomponent treatment program for irritable bowel syndrome to a small group format. *Biofeedback and Self-regulation, 12,* 63–69.

Blanchard, E. B., Schwarz, S. P., Suls, J. M., Gerardi, M. A., Scharff, L., Greene, B., Taylor, A. E., Berreman, C., & Malamood, H. S. (1992). Two controlled evaluations of multicomponent psychological treatment of irritable bowel syndrome. *Behaviour Research and Therapy, 30,* 175–189.

Bloomfield, A. L. & Polland, W. S. (1931). Experimental referred pain from the gastrointestinal tract. Part II. Stomach, duodenum and colon. *Journal of Clinical Investigation, 10,* 453–473.

Camilleri, M., & Prather, C. M. (1992). The irritable bowel syndrome: Mechanisms and a practical approach to management. *Annals of Internal Medicine*, *116*(12), 1001–1008.

Chaudhary, N. A., & Truelove, S. C. (1962). The irritable colon syndrome: A study of the clinical features, predisposing causes, and prognosis in 130 cases. *Quarterly Journal of Medicine, New Series XXXI*(123), 307–322.

Clouse, R. E. (1988). Anxiety and gastrointestinal illness. *Psychiatric Clinics of North America*, *11*(2), 399–417.

Comey, R. H., & Stanton, R. (1990). Physical symptom severity, psychological and social dysfunction in a series of outpatients with irritable bowel syndrome. *Journal of Psychosomatic Research*, *34*(5), 483–491.

Covino, N. A., & Frankel, F. H. (1993). Hypnosis and relaxation in the medically ill. *Psychotherapy Psychosomatic*, *60*(2), 75–90.

Craig, T. K. J., & Brown, G. W. (1984). Goal frustration and life events in the aetiology of painful gastrointestinal disorder. *Journal of Psychosomatic Research*, *28*(5), 411–421.

Davis, W. D. (1980). Lower bowel disorders: 1. Irritable bowel syndrome. *Postgraduate Medicine: Bowel Disorders, Part 1*, *68*(4), 60–62, 64.

Delaney, B. C. (1995). Role of *Heliobacter pylori* in gastrointestinal disease: Implications for primary care of a revolution in management of dyspepsia. *British Journal of General Practice*, *45*, 489–494.

DeLongis, A., Coyne, J. C., Dakof, G., Folkman, S., & Lazarus, R. S. (1982). Relationship of daily hassles, uplifts, and major life events to health status. *Health Psychology*, *1*(2), 119–136.

DiNardo, P. A., Brown, T., & Barlow, D. H. (1994). *Anxiety disorders interview schedule (ADIS-IV)*. Albany, NY: Phobia and Anxiety Disorders Clinic, Center for Stress and Anxiety Disorders, SUNY at Albany.

Drossman, D. A. (1984). Diagnosis of the irritable bowel syndrome. A simple solution? *Gastroenterology*, *87*(1), 224–225.

Drossman, D. A. (1995). Diagnosing and treating patients with refractory functional gastrointestinal disorders. *Annals of Internal Medicine*, *123*(4), 688–697.

Drossman, D. A., Li, Z., Andruzzi, E., Temple, R. D., Talley, N. J., Thompson, W. G., Whitehead, W. E., Janssens, J., Funch-Jensen, P., Corazziari, E., Richter, J. E., & Koch, G. G. (1993). US Householder Survey of Functional Gastrointestinal Disorders: Prevalence, sociodemography, and health impact. *Digestive Diseases and Science*, *38*(9), 1569–1580.

Drossman, D. A., Patrick, D. L., & Appelbaum, M. I. (1989). Do IBD ostomates do better? A study of symptoms and health-related quality of life. *Gastroenterology*, *96*, A130. Abstract.

Drossman, D. A., Patrick, D. L., Mitchell, C. M., Zagami, E. A. & Applebaum, M. I. (1989). Health-related quality of life in inflammatry bowel disease: Functional status and patient worries and concerns. *Digestive Diseases and Science*, *34,* 1379–1386.

Drossman, D. A., Talley, N. J., Leserman, J., Olden, K. W., & Barreiro, M. A. (1995). Sexual and physical abuse and gastrointestinal illness. *Annals of Internal Medicine, 123,* (10), 782–794.

Drossman, D. A., & Thompson, W. G. (1992). The irritable bowel syndrome: Review and a graduated multicomponent treatment approach. *Annals of Internal Medicine, 116*(12), 1009–1016.

Folks, D. G., & Kinney, F. C. (1992). The role of psychological factors in gastrointestinal conditions. A review pertinent to DSM-IV *Psychosomatics, 33*(3), 257–270.

Fossey, M. D., & Lydiard, R. B. (1990). Anxiety and the gastrointestinal system. *Psychiatric Medicine, 8*(3), 175–186.

Giles, S. L. (1978). *Separate and combined effects of biofeedback training and brief individual psychotherapy in the treatment of gastrointestinal disorders*. Unpublished doctoral dissertation, Department of Psychology, University of Colorado.

Gomborone, J., Dewsnap, P., Libby, G., & Farthing, M. (1995). Abnormal illness attitudes in patients with irritable bowel syndrome. *Journal of Psychosomatic Research, 39*(2), 227–230.

Greene, B., & Blanchard, E. B. (1994). Cognitive therapy for irritable bowel syndrome. *Journal of Consulting and Clinical Psychology, 62*(3), 576–582.

Griffin, P. H. (1985). Irritable bowel syndrome: A final common response to stress? *Drug Therapy, 4,* 45–61.

Guthrie, E., Creed, F., Dawson, D., & Tomenson, B. (1991). A controlled trial of psychological treatment for the irritable bowel syndrome. *Gastroenterology, 100*(2), 450–457.

Guthrie, E., Creed, F. H., & Whorwell, P. J. (1987). Severe sexual dysfunction in women with the irritable bowel syndrome: Comparison of inflammatory bowel disease and duodenal ulceration. *British Medical Journal.*

Harvey, R. F., Gunary, R. N., Hinton, R. A., & Barry, R. E. (1989). Individuals and group hypnotherapy in treatment of refractory irritable bowel syndrome. *Lancet, 1,* 424–425.

Harvey, R. F., Salih, S. Y., & Read, A. E. (1983). Organic and functional disorders in 2000 gastroenterology outpatients. *Lancet, 1,* 632–634.

Jandorf, L., Deblinger, E., Neale, J. M., & Stone, A. A. (1986). Daily versus major life events as predictors of symptom frequency: A replication study. *The Journal of General Psychology, 113*(3), 205–218.

Kellow, J. E., & Langloddeck, P. M. (1989). Advances in the understanding and management of the irritable bowel syndrome. *Medical Journal of Australia, 151*(2), 92, 95–99.

Klein, K. B. (1988). Controlled treatment trials in the irritable bowel syndrome: A critique. *Gastroenterology, 95*(1), 232–241.

Kruis, W., Thieme, C., Weinzierl, M., Schussler, P., Holl, J., & Paulus, W. (1984). A diagnostic score for the irritable bowel syndrome. Its value in the exclusion of organic disease. *Gastroenterology, 87*(1), 1–7.

Latimer, P. R. (1983). Irritable bowel syndrome: Psychosomatic illness review: No. 7 in a series. *Psychosomatics, 24* (3), 205–218.

Latimer, P. R., Campbell, D., Latimer, M. R., Daniel, E. E., & Waterfall, W. E. (1979). Irritable bowel syndrome: A test of the colonic hyperalgesia hypotheses. *Journal of Behavioral Medicine, 2*(3), 285–293.

Latimer, P. R., Sarna, S. K., Campbell, D., Latimer, M. R., Waterfall, W. E., & Daniel, E. E. (1981). Colonic motor and myoelectrical activity: A comparative study of normal patients, psychoneurotic patients and patients with irritable bowel syndrome (IBS). *Gastroenterology, 80*(9), 893–901.

Lewin, J., & Lewis, S. (1995). Organic and psychosocial risk factors for duodenal ulcer. *Journal of Psychosomatic Research, 39*(5), 531–548.

Liss, J. L., Alpers, D., & Woodruff, R. A. (1973, April/May). The irritable colon syndrome and psychiatric illness. *Diseases of the Nervous System, 34,* 151–157.

Lydiard, R. B. (1992). Anxiety and the irritable bowel syndrome. *Psychiatric Annals, 22*(12), 612–618.

Lydiard, R. B., Fossey, M. D., Marsh, W., & Ballenger, J. C. (1993). Prevalence of psychiatric disorders in patients with irritable bowel syndrome. *Psychosomatics, 34*(3), 229–234.

Lynch, P. N., & Zamble, E. (1989). A controlled behavioral treatment of irritable bowel syndrome. *Behavior Therapy, 20,* 509–523.

Lynn, R. B., & Friedman, L. S. (1993). Irritable bowel syndrome. *New England Journal of Medicine, 239,* 1940–1945.

Madhu, S. V., Vij, J. C., Bhatnagar, O. P., Krishnamurthy, N., Anand, B. S., & Chuttani, H. K. (1988). Colonic myoelectrical activity in irritable bowel syndrome before and after treatment. *Indian Journal of Gastroenterology, 7*(1), 31–33.

Manning, A. P., Thompson, W. G., Heaton, K. W., & Morris, A. F. (1978). Towards positive diagnosis of the irritable bowel. *British Medical Journal, 2*(2), 653–654.

Meichenbaum, D. (1985). *Stress inoculation training*. Elmsford, NY: Pergamon.

Mitchell, C. M., & Drossman, D. A. (1987). The irritable bowel syndrome: Understanding and treating a biopsychosocial illness disorder. *Annals of Behavioral Medicine, 9*(3), 13–18.

Muller-Lissner, S. A. (1988). Effect of wheat bran on weight of stool and gastrointestinal transit time: A meta analysis. *British Medical Journal, 296,* 615–617.

Murray, C. D. (1930). A brief psychological analysis of a patient with ulcerative colitis. *Journal of Nervous and Mental Disease, 72*(6), 617–627.

Neff, D. F., & Blanchard, E. B. (1987). A multi-component treatment for irritable bowel syndrome. *Behavior Therapy, 18*(1), 70–83.

North, C. S., & Alpers, D. H. (1994). A review of studies of psychiatric factors in Crohn's Disease: Etiologic implication. *Annals of Clinical Psychiatry, 6*(2) 117–124.

Olden, K. W. (1992). Inflammatory bowel disease: A biopsychosocial perspective. *Psychiatric Annals, 22*(12), 619–623.

Payler, D. K., Pomare, E. W., Heaton, K. W., & Harvey, R. F. (1975). The effect of wheat bran on intestinal transit. *Gut, 16*(3), 209–213.

Payne, A., & Blanchard, E. B. (1995). A controlled comparison of cognitive therapy and self-help support groups in the treatment of irritable bowel syndrome. *Journal of Consulting and Clinical Psychology, 63*(5), 779–786.

Persons, J. B. (1989). *Cognitive therapy in practice: A case formulation approach.* New York: Norton.

Peura, D. A., & Graham, D. Y. (1994). *Heliobacter pylori*: Consensus reached: Peptic ulcer is on the way to becoming a historic disease. *American Journal of Gastroenterology, 89*(8), 1137–1139.

Radnitz, C. L., & Blanchard, E. B. (1988). Bowel sound biofeedback as a treatment for irritable bowel syndrome. *Biofeedback and Self-regulation, 13*(2), 169–179.

Ritchie, J. (1973). Pain from distention of the pelvic colon by inflating a balloon in the irritable colon syndrome. *Gut, 14,* 125–132.

Schwarz, S. P., & Blanchard, E. B. (1990). Inflammatory bowel disease: A review of the psychological assessment and treatment literature. *Annals of Behavioral Medicine, 12*(3), 95–105.

Schwarz, S. P., Taylor, A. E., Scharff, L., & Blanchard, E. B. (1990). A four-year follow-up of behaviorally treated irritable bowel syndrome patients. *Behaviour Research and Therapy, 28*(4), 331–335.

Snape, W. J., Carlson, G. M., & Cohen, S. (1976). Colonic myoelectrical activity in the irritable bowel syndrome. *Gastroenterology, 70*(3), 326–330.

Suls, J., Wan, C. K., & Blanchard, E. B. (1994). A multilevel data-analytic for approach for evaluation of relationships between daily life stressors and symptomatology: Patients with irritable bowel syndrome. *Health Psychology, 13,* 103–113.

Svedlund, J., Sjodin, I., Ottosson, J.-O., & Dotevall, G. (1983). Controlled study of psychotherapy in irritable bowel syndrome. *Lancet, 2,* 589–592.

Talley, N. J., Gabriel, S., Harmsen, W., Zinsmeister, A., & Evans, R. (1995). Medical costs in community subjects with irritable bowel syndrome. *Gastroenterology, 101,* 927–934.

Talley, N. J., Phillips, S., Bruce, B., Twomey, C., Zinsmeister, A., & Melton, J. (1990). Relation among personality and symptoms in non-ulcer dyspepsia and the irritable bowel syndrome. *Gastroenterology, 99*(2), 327–333.

Talley, N. J., Phillips, S., Melton, L., Mulvihill, C., Wiltgen, C., & Zinsmeister, A. (1990). Diagnostic value of the Manning Criteria in irritable bowel syndrome. *Gut, 31*(1), 77–81.

Talley, N. J., Weaver, A. L., Zinsmeister, A. R., & Melton, III, L. J. (1992). Onset and disappearance of gastrointestinal symptoms and functional gastrointestinal disorders. *American Journal of Epidemiology, 136*(2), 165–177.

Talley, N. J., Zinsmeister, A. R., Van Dyke, C., & Melton, III, L. J. (1991). Epidemiology of colonic symptoms and the irritable bowel syndrome. *Gastroenterology, 101,* 927–934.

Thompson, W. G. (1991). Symptomatic presentations of the irritable bowel syndrome. *Gastroenterology Clinics of North America, 20,* 235–247.

Thompson 1986.

Thompson, W. G., Creed, F., Drossman, D. A., Heaton, K. W., & Mazzacca, G. (1992). Functional bowel disease and functional abdominal pain. *Gastroenterology International, 5,* 75–91.

Thornton, S., McIntyre, P., Murray-Lyon, I., & Gruzelier, J. (1990). Psychological and psychophysiological characteristics in irritable bowel syndrome. *British Journal of Clinical Psychology, 29*(3), 343–345.

Tollefson, G. D., Tollefson, S. L., Pederson, M., Luxenburg, M., & Dunsmore, G. (1991). Comorbid irritable bowel syndrome in patients with generalized anxiety and major depression. *Annals of Clinical Psychiatry, 3*(3), 215–222.

van Dulmen, A. M., Fennis, J. F. M., & Bleijenberg, G. (1996). Cognitive-behavioral group therapy for irritable bowel syndrome: Effects and long-term follow-up. *Psychosomatic Medicine, 58*(5), 508–514.

Vollmer, A., & Blanchard, E. B. (1997). *A controlled comparison of individual and small-group cognitive therapy for the irritable bowel syndrome*. Manuscript under editorial review.

Walker, E. A., Katon, W. J., Roy-Byrne, P. P., Jemelka, R. P., & Russo, J. (1993). Histories of sexual victimization in patients with irritable bowel syndrome or inflammatory bowel disease. *American Journal of Psychiatry, 150*(10), 1502–1506.

Walker, E. A., Roy-Byrne, P. P., & Katon, W. J. (1990). Irritable bowel syndrome and psychiatric illness. *American Journal of Psychiatry, 147*(5), 565–572.

Walker, E. A., Roy-Byrne, P. P., Katon, W. J., Li, L., Amos, D., & Jiranek, G. (1990). Psychiatric illness and irritable bowel syndrome: A comparison with inflammatory bowel disease. *American Journal of Psychiatry, 147,* 1656–1661.

Warren, J. R., & Marshall, B. J. (1983). Unidentified curved bacillus on the gastric epithelium in active chronic gastritis. *Lancet, 1,* 1273–1275.

Watson, D., & Pennebaker, J. W. (1989). Health complaints, stress and distress: Exploring the central role of negative affectivity. *Psychological Review, 96*(2), 234–254.

Webb, P. M., Knight, T., Greaves, S. & Wilson, A. (1994). Relation between infection with *Heliobacter pylori* and living conditions in childhood: Evidence for person to person transmission in early life. *British Medical Journal, 308,* 750–753.

Whitehead, W. E. (1992a). Behavioral medicine approaches to gastrointestinal disorders. *Journal of Consulting and Clinical Psychology, 60*(4), 605–612.

Whitehead, W. E. (1992b). Biofeedback treatment of gastrointestinal disorders. *Biofeedback and Self-regulation, 17*(1), 59–76.

Whitehead, W. E. (1994). Assessing the effects of stress on physical symptoms. *Health Psychology, 13*(2), 99–102.

Whitehead, W. E., Crowell, M. D., Robinson, J. C., Heller, B. R., & Schuster, M. M. (1992). Effects of stressful life events on bowel symptoms: Subjects with irritable bowel syndrome compared with subjects without bowel dysfunction. *Gut, 33*(6), 825–830.

Whitehead, W. E., Winget, C., Fedoravicius, A. S., Wooley, S., & Blackwell, B. (1982). Learned illness behavior in patients with irritable bowel syndrome and peptic ulcer. *Digestive Diseases and Sciences, 27*(3), 202–208.

Whorwell, P. J., Prior, A., & Colgan, S. M. (1987). Hypnotherapy in severe irritable bowel syndrome: Further experience. *Gut, 28*(4), 423–425.

Whorwell, P. J., Prior, A., & Faragher, E. B. (1984). Controlled trial of hypnotherapy in the treatment of severe refractory irritable bowel syndrome. *Lancet, 2,* 1232–1234.

Wise, T. N., Cooper, J. N., & Ahmed, S. (1982). The efficacy of group therapy for patients with irritable bowel syndrome. *Psychosomatics, 23*(5), 465–469.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.13
Psychological Aspects of End-stage Renal Disease

ALAN J. CHRISTENSEN and PATRICIA J. MORAN
University of Iowa, Iowa City, IA, USA

8.13.1 INTRODUCTION 321
8.13.2 OVERVIEW OF ESRD AND ITS TREATMENT 322
8.13.3 PATIENT ADHERENCE TO THE RENAL TREATMENT REGIMEN 323
  8.13.3.1 Factors Influencing Patient Adherence 324
    8.13.3.1.1 Demographic factors 324
    8.13.3.1.2 Cognitive factors 324
    8.13.3.1.3 Social support 325
    8.13.3.1.4 Personality influences 326
    8.13.3.1.5 Patient coping and adherence 327
  8.13.3.2 Patient × Treatment Interactions in ESRD: Influence on Patient Adherence 328
  8.13.3.3 Improving Adherence Through Intervention 328
8.13.4 PATIENT DEPRESSION IN END-STAGE RENAL DISEASE 329
  8.13.4.1 Psychological Determinants of Depression 330
    8.13.4.1.1 Control, stress, and depression 330
    8.13.4.1.2 Illness intrusiveness 331
    8.13.4.1.3 Social support and depression 331
  8.13.4.2 Treatment of Depression in ESRD 332
8.13.5 PSYCHOSOCIAL INFLUENCES ON SURVIVAL 332
8.13.6 IMPLICATIONS FOR THE CONSULTING PSYCHOLOGIST 333
8.13.7 CONCLUDING OBSERVATIONS 335
8.13.8 REFERENCES 335

8.13.1 INTRODUCTION

The over 240 000 end-stage renal disease (ESRD) patients in the USA face an incurable, life-threatening, chronic disease. In many ways, the illness is unique among medical conditions. The extreme dependence on artificial means for survival and the substantial behavioral demands typically placed on the ESRD patient have few parallels in healthcare. However, despite the uniqueness of the disorder and its treatment, most of the central clinical issues in this population are common to all chronic medical conditions. The goals of this chapter are to review the major psychological problems faced by ESRD patients and to illustrate the unique features of ESRD and its treatment that impinge upon these problems. In doing so, we will emphasize the importance of considering the interaction between psychological qualities of the individual and aspects of the disease and treatment context. Finally, the chapter discusses the implications of these issues for the consulting psychologist assisting in the care and screening of the ESRD patient.

8.13.2 OVERVIEW OF ESRD AND ITS TREATMENT

The onset of ESRD is defined by an irreversible, total or near-total cessation of kidney function. ESRD afflicts individuals of all ages, ethnic groups, and socioeconomic strata. The disorder is slightly more common among males, and is substantially more prevalent among African Americans (US Renal Data System [USRDS], 1995). ESRD is most commonly due to the advanced complications of another medical condition (e.g., diabetes, hypertension). For other patients the underlying etiology is specific to the renal system (e.g., glomerulonephritis and polycystic kidney disease). Diabetic nephropathy is the most common etiological factor accounting for over one-third of new ESRD cases (USRDS, 1995).

From a healthcare policy standpoint, ESRD is unique among chronic medical conditions. The Medicare ESRD entitlement program enacted in 1973 makes all individuals with chronic renal failure eligible for benefits. In 1994 over six billion dollars in healthcare benefits were provided by Medicare for the treatment of ESRD (USRDS, 1995).

Only three decades ago, a diagnosis of ESRD meant near-certain death. Upon the cessation of renal function, excess fluid, metabolic toxins, and electrolytes rapidly accumulate in blood and bodily tissues. These substances must be removed by alternative means if the ESRD patient is to survive. Current renal replacement therapies include renal transplantation and several forms of renal dialysis. Renal grafts come from either a cadaveric (brain-dead) or living (typically a first-degree relative) donor. Despite recent advances in immunosuppressive therapy, activation of the patient's immune system resulting in organ rejection and failure remains an important limitation to the potential benefit of transplantation. The success rate for cadaveric donor organs is lowest, with a five-year graft survival rate of approximately 53%. For organs from living related donors, a five-year graft survival rate of 70% has recently been reported (USRDS, 1995). Despite the superior success rate for living related donation, a large majority (approximately 75%) of renal transplants actually performed involve cadaveric donors (USRDS, 1995).

Given a shortage of donor organs and a significant transplant rejection rate, most ESRD patients rely on some form of renal dialysis as treatment for their condition. Currently, approximately 75% of the ESRD patients in the USA are being treated with one of two forms of renal dialysis (USRDS, 1995). These primary dialysis modalities are center hemodialysis (82% of dialysis patients nationally) and continuous ambulatory peritoneal dialysis (CAPD) (13% of dialysis patients nationally).

There is an important difference in the role taken by the patient while undergoing the different forms of dialysis. The center hemodialysis patient is a passive recipient of treatment in the confines of a dialysis center. Hemodialysis achieves removal of toxins and excess fluid via the circulation of blood through an artificial kidney (the dialyzer). The dialysis procedure is performed three times a week by trained technicians in a hospital setting, requiring approximately four hours per session. Hemodialysis treatment commences by way of a vascular connection made between the dialysis machine and the patient, usually through an arteriovenous fistula permanently placed in the patient's forearm. Little participation is allowed or required of the patient while undergoing the procedure. During hemodialysis, patients may experience a number of physiological stressors including rapid drops in blood pressure, excess bleeding at the access site, nausea, muscle cramps, and increased fatigue. For a small minority of patients hemodialysis is carried out at home with assistance from a technician or caregiver. Although mechanically similar procedures are involved, home hemodialysis patients typically have the opportunity to be more actively involved in treatment delivery.

In contrast to hemodialysis, CAPD treatment requires the patient to take an active role to ensure treatment success. In CAPD a permanent catheter is surgically implanted in the abdomen. CAPD is carried out manually by the patient. A sterile tube is used to carefully connect the catheter to a bag of sterile dialysis solution (dialysate). The bag is elevated to allow flow of the dialysate into the peritoneal cavity. After this procedure is completed the bag is tucked away under the patient's clothing. Over the next four to six hours (eight hours while sleeping), the patient remains ambulatory as continuous dialysis ensues. During this "dwell time," blood filters through the peritoneal membrane leaving toxins behind in the dialysate. After the dwell time is complete, the bag is lowered and the used solution is allowed to drain back into the bag where it is discarded and the procedure begins again. Throughout each step of the procedure the patient must remain highly vigilant in keeping themselves and the dialysis connections sterile to help avoid serious infection (peritonitis) that can begin at the catheter access site. Because of the intricacies of the procedures involved, the successful administration of CAPD is clearly dependent on the patient taking a more active role in his or her treatment compared with patients undergoing hemodialysis. In contrast to the diminished use of home hemodialysis, the popularity of CAPD has steadily increased over the past 10 years (USRDS, 1995).

The availability of different treatment modalities, each with its own unique characteristics and patient demands, makes the management of ESRD unique. Hemodialysis, CAPD, and renal transplantation are all medically acceptable treatment alternatives for the large majority of ESRD patients. The different treatments have not generally been linked to differences in patient survival (Flechner, 1994). There is some indication that a successful renal transplant holds certain advantages in terms of patient quality of life (e.g., less emotional distress, greater mobility) (Christensen, Holman, Turner, Smith, & Grant, 1991; Evans et al., 1985). However, there is little evidence of quality of life differences among patients using the various types of renal dialysis (Evans et al., 1985; Simmons, Abress, & Anderson, 1988). In general, the choice of a particular ESRD treatment modality is largely a function of nonmedical factors, including patient and provider preferences and judgements about which modality would provide a particular patient with the highest quality of life.

8.13.3 PATIENT ADHERENCE TO THE RENAL TREATMENT REGIMEN

As with most chronic illnesses, patient nonadherence to the prescribed medical treatment regimen is a pervasive problem among individuals receiving chronic renal replacement intervention. Adherence to a multifaceted treatment regimen is required of both renal dialysis and transplantation patients. However, the vast majority of adherence studies have focused on individuals receiving some form of dialysis to compensate for a life-threatening loss of kidney function.

In addition to undergoing a regular schedule of time-consuming dialysis treatments, dialysis patients are required to follow a strict medication and dietary regimen. Both CAPD and hemodialysis patients are required to take regular doses of phosphate-binding medication as well as to reduce intake of phosphorus-rich foods (e.g., dairy products) due to the body's inability to excrete phosphorus while undergoing dialysis treatment. For many patients, phosphate-binding medication is poorly tolerated by the digestive tract leading to unpleasant gastrointestinal side effects. However, if the medication regimen is not followed, serum phosphorus (P) will rise above normal levels. Sustained elevations in serum P are associated with a variety of complications including renal osteodystrophy, serious decreases in calcium, and subsequent bone demineralization.

Hyperkalemia is also a clinically significant problem for all dialysis patients due to the body's inability to regulate serum potassium (K). Dialysis patients are required to follow strict dietary guidelines to maintain safe serum K levels. If dietary guidelines are not followed, serum K will rise and potentially life-threatening cardiovascular complications, such as cardiac arrythmia, can occur.

Because of their importance for clinical outcomes, serum P and serum K levels are assessed approximately monthly as part of the routine biochemical monitoring of dialysis patients. Both higher serum P and serum K levels are typically interpreted as reflecting poorer adherence. Serum K values above $6.0\,\mathrm{mEq\,L^{-1}}$ and serum P values greater than $5.5\,\mathrm{mg\,dL^{-1}}$ are generally considered indicative of problematic adherence (Wolcott, Maida, Diamond, & Nissenson, 1986).

In the case of hemodialysis treatment, there are also extreme limitations on the amount of fluid that can be safely consumed due to the intermittent nature of the fluid and waste clearance performed by the artificial kidney. Prolonged fluid overload is associated with congestive heart failure, hypertension, dizziness, shortness of breath, and severe muscle cramping, and is potentially life-threatening.

Interdialytic weight gain (IWG) is typically used to define adherence to the fluid-intake restrictions. IWG is determined by subtracting the postdialytic weight for the previous treatment session from the predialytic weight for the current session. The values resulting from this computation are believed to be a valid reflection of the amount of fluid that the patient ingests between dialysis sessions (Manley & Sweeney, 1986). Higher IWG values are interpreted as reflecting poorer patient adherence, with values over 2.5 kg (or over 4% of the patient's body weight) generally indicative of problematic adherence (Wolcott et al., 1986).

Studies examining the prevalence of nonadherence among renal dialysis patients have typically observed that between 30 and 50% of dialysis patients do not adhere to diet, fluid-intake, and medication regiments (Bame, Petersen, & Wray, 1993; Christensen et al., 1992; Schneider, Friend, Whitaker, & Nadhwa, 1991; Weed-Collins & Hogan, 1989; Wolcott et al., 1986). In general, these reports indicated that nonadherence is most common for fluid-intake guidelines and somewhat less common for dietary or medication restrictions. This pattern is consistent with reports that adhering to the fluid-intake restrictions may be the most challenging and stressful aspect of the hemodialysis regimen (Baldree, Murphy, & Powers, 1982; Rosenbaum & Ben-Ari Smira, 1986).

Less research is available examining rates of nonadherence among patients receiving a renal transplant. Post-transplantation patients are required to follow a strict immunosuppressive medication regimen as well as to remain vigilant about physical changes that may signal organ rejection, infection, or some other medical problem. Immunosuppressive medications can result in a range of side effects including mood changes, changes in appearance (e.g., hypertrichosis, cushnoid appearance), and nausea and vomiting. There is an emerging consensus that patient nonadherence is one of the leading causes of graft failure. One recent study indicated that 78% of cases of renal graft failure in the second year after transplant were due to nonadherence (Kiley, Lam, & Pollak, 1993). Thus, strict adherence to the immunosuppressive medication regimen is a central determinant of treatment success.

Because serum levels of many immunosuppressive medications are quite unstable over time and are influenced by a number of factors other than medication adherence (e.g., presence of an acute illness, other medications), reliably determining the rate of nonadherence is difficult (Lemaire, Fahr, & Maurer, 1990). Estimates of nonadherence among transplanted patients have varied greatly ranging from less than 5% to as many as 75% of patients (De Geest et al., 1995; Didlake, Dreyfus, Kerman, Van Buren, & Kahan, 1988; Kiley et al., 1993; Rovelli et al., 1989). This inconsistency across adherence studies is, at least in part, probably due to the lack of reliable and valid methods to assess adherence among transplant patients. To obtain more precise estimates of adherence in this population, researchers need to apply more direct methods of assessing actual medication-taking behavior.

8.13.3.1 Factors Influencing Patient Adherence

Given the prevalence and clinical importance of nonadherence in all forms of renal replacement therapy, it is not surprising that the identification of factors related to adherence has received considerable empirical attention. We review research examining several categories of factors that have been examined in relation to adherence in this population.

8.13.3.1.1 Demographic factors

Despite considerable research, few consistent patterns have emerged involving the relation of demographic variables to adherence among ESRD patients. In a number of studies, younger patients have exhibited consistently poorer adherence to various aspects of the prescribed dialysis regimen when compared to older patients (Bame et al., 1993; Boyer, Friend, Chlouverakis, & Kaloyanides, 1990; Christensen & Smith, 1995; Cummings, Becker, Kirscht, & Levin, 1982; Kimmel et al., 1995). A similar effect for age has been observed among renal transplant patients (Didlake et al., 1988). There is also evidence that males exhibit significantly poorer adherence than females when undergoing either dialysis or transplantation (Boyer et al., 1990; De Geest et al., 1995; Kiley et al., 1993; Morduchowicz et al., 1993). No other consistent patterns involving demographic variables have emerged in the ESRD adherence literature.

8.13.3.1.2 Cognitive factors

Patient beliefs or appraisals as determinants of medical regimen adherence have received considerable attention. A number of studies involving renal dialysis patients have examined components of the Health Belief Model (HBM; Rosenstock, 1966) as predictors of adherence. The HBM predicts that an individual will carry out a prescribed behavior when the perceived benefits of the action (i.e., a reduction in the threat of adverse health consequences) outweigh any perceived barriers associated with taking the action. Research involving both center hemodialysis patients (Cummings, Becker, Kirscht, & Levin, 1981; Weed-Collins & Hogan, 1989) and renal transplant patients (Kiley et al., 1993) have reported that greater perceived barriers (e.g., "being away from home," "medication cost") were associated with poorer medication adherence. However, with few exceptions other cognitive factors composing the HBM have been found to be unrelated to adherence among ESRD patients (Cummings et al., 1982; Hartmann & Becker, 1978; Rosenbaum & Ben-Ari Smira, 1986). Rosenbaum and Ben-Ari Smira (1986) have suggested that behavioral skills play a more central role in influencing adherence than do health-related beliefs. From this perspective, motivation or intention to adhere may not be closely linked to success in adhering to a regimen that is demanding or requires certain behavioral or self-control skills that some patients may not possess (e.g., the ability to delay gratification).

There is some evidence to suggest that self-efficacy expectations are related to fluid-intake adherence among hemodialysis patients as well as medication adherence among transplanted patients (De Geest et al., 1995; Rosenbaum & Ben-Ari Smira, 1986; Schneider et al., 1991). In general, patients who believe that they are capable of performing compliance behaviors are more likely to do so. De Geest et al. (1995) reported that patients perceiving stronger self-efficacy expectations regarding their ability to follow the immunosuppressive regimen exhibited both more favorable medication adherence and better renal graft survival at a five-year follow-up. In a study involving center hemodialysis patients, Rosenbaum and Ben Ari Smira (1986) suggested that an observed association between self-efficacy and adherence was largely due to patients" perceptions of past adherence success and attributions of that success to their own efforts. This suggests that patient perceptions of past behavior and the specific attributions they make regarding this behavior may play a central role in mediating adherence.

A cognitive construct that has often been applied to the prediction of adherence behavior among chronically ill individuals involves individual differences in perceived control over health. The constructs of generalized (Rotter, 1966) and health-specific locus of control (Wallston, Wallston, & DeVellis, 1978) have both received considerable attention in the adherence literature. Locus of control reflects the degree to which individuals believe attaining a desired outcome (e.g., a positive health outcome) is contingent upon their own behavior or some external influence. Despite considerable examination of the relation of locus of control (and health-specific locus of control) to adherence among ESRD patients, no clear consensus regarding the adaptive significance of control expectancies has emerged. There is some evidence that hemodialysis patients with an internal locus of control exhibit more favorable adherence to dietary and fluid-intake aspects of the treatment regimen (Kaplan-DeNour & Czaczkes, 1972; Oldenburg, MacDonald, & Perkins, 1988; Poll & Kaplan-DeNour, 1980). However, other research suggests that internal control expectancies are not significantly related to hemodialysis regimen adherence (Brown & Fitzpatrick, 1988; Schneider et al., 1991; Witenberg et al., 1983).

One recent study suggests that health locus of control and perceived self-efficacy may interact in influencing adherence behavior (Christensen, Wiebe, Benotsch, & Lawton, 1996). In this study involving a mixed sample of hemodialysis and CAPD patients, greater generalized self-efficacy expectations ("perceived health competence") were associated with more favorable medication regimen adherence only for those patients possessing a strong belief in provider control over health outcomes (i.e., high "powerful others health locus of control"). This pattern suggested that a dialysis patient's perceived ability to manage health effectively is related to adherence only when the patient believes positive health outcomes are contingent upon following the advice and actions of healthcare providers.

8.13.3.1.3 Social support

Considerable evidence suggests that the availability and perceived quality of social resources are important correlates of health-related behavior in both clinical and nonclinical populations (e.g., Kaplan & Hartwell, 1987; Ruggeiro, Spirito, Bond, Coustan, & McGarvey, 1990). However, limited research is available regarding adherence among ESRD patients. Christensen et al. (1992) examined the effects of social support in the family with regard to patient adherence in a sample of hemodialysis patients. Results indicated that patients holding perceptions of a more supportive family environment characterized by greater cohesion and expressiveness among family members and less intrafamily conflict, exhibited significantly more favorable adherence to fluid-intake restrictions than did patients reporting less family support. However, family support was not associated with adherence to dietary restrictions. Other evidence suggests that more favorable marital adjustment among center hemodialysis patients and their spouses is related to better fluid intake but not dietary adherence (Somer & Tucker, 1988, 1992). Similarly, involvment in a stable, dyadic relationship has been associated with more regular dialysis session attendance (Kimmel et al., 1995).

Studies involving family support or marital adjustment reflect a relatively consistent association of environmental factors and patient adherence. However, evidence for the role of extrafamilial social support is mixed. Several studies involving samples of center hemodialysis patients found that perceived support was not related to biochemical markers of fluid intake or medication adherence (Boyer et al., 1990; Cummings et al., 1982; Hitchcock, Brantley, Jones, & McKnight, 1992). Kimmel et al. (1995) have suggested the possibility of gender differences in the association of support with adherence. In this study of 149 center hemodialysis patients, a significant association between greater perceived support and more favorable medication adherence was obtained among males but not females. Finally, a study examining patient–provider interactions suggests that greater emotional support delivered by dialysis unit nurses to their patients is associated with more favorable fluid-intake adherence (Rorer, Tucker, & Blake, 1988).

8.13.3.1.4 Personality influences

Past reviews have concluded that there is no predictable or consistent association between stable, person factors and the adherence behavior of chronically ill patients (Haynes, 1979; Kaplan & Simon, 1990). However, more recent research involving renal dialysis has produced considerably more promising results (e.g., Christensen & Smith, 1995; Christensen, Wiebe, Edwards, Michels, & Lawton, 1996; Christensen, Wiebe, & Lawton, in press; Wiebe & Christensen, 1996).

In their review of the Five-Factor Model of personality, Wiebe and Christensen (1996) suggest that conscientiousness may be the personality trait most relevant to adherence behavior. The conscientiousness factor has been described as reflecting a highly purposeful, self-disciplined style, high in self-control and constraint (Digman, 1990; McCrae & John, 1992). Correspondingly, previous investigators have characterized nonadherence as a self-control problem citing low frustration tolerance, inability to delay gratification, and acting out as underlying causes of the problem (Kaplan-De Nour & Czaczkes, 1972; Rosenbaum and Ben-Ari Smira, 1986).

In a study of 72 renal dialysis patients, Christensen and Smith (1995) reported that higher conscientiousness scores from the NEO-Five Factor Inventory (Costa & McCrae, 1991) predicted a significant portion of the variance in medication adherence after the effects of relevant demographic variables were controlled. However, conscientiousness was not related to serum K-defined dietary adherence.

Empirically and conceptually related to the conscientiousness trait is the construct of learned resourcefulness. Learned resourcefulness as measured by Rosenbaum's (1980) Self-control Schedule (SCS) is believed to reflect one's tendency to apply self-control skills in solving behavioral problems (e.g., use of strategies to delay gratification or tolerate frustration). Rosenbaum and Ben-Ari Smira (1986) reported that lower SCS scores were related to poorer adherence with prescribed fluid-intake restrictions among 53 center hemodialysis patients. A path analysis suggested that learned resourcefulness affected fluid adherence indirectly through self-efficacy expectations and a tendency for high scorers on the SCS to attribute past adherence success to their own efforts. This suggests that personality characteristics such as learned resourcefulness or conscientiousness may promote adaptive health-related cognitions which are in turn related to more favorable adherence behavior.

Another personality factor that may have implications for adherence behavior is trait hostility. Considerable evidence links hostility with poorer physical health outcomes (Smith, 1992). An association between hostility and health may, in part, be mediated by differences in adherence behavior among chronically ill individuals. In the first study directly examining this possibility, Christensen, Wiebe, and Lawton (in press) examined hostility defined on the Cook-Medley (Cook & Medley, 1954) hostility (Ho) scale as a potential determinant of adherence. These authors argued that the generally suspicious, mistrusting, and cynical style of high scorers on the Ho scale may predispose them to reject or disregard the advice of healthcare providers.

Christensen, Wiebe, and Lawton (in press) obtained a predicted main effect between higher Ho scale scores and poorer adherence to the phosphorus control regimen. However, moderational analyses indicated that this effect was qualified by the interaction of hostility and powerful others health locus of control (PHLC) expectancies. As depicted in Figure 1, the deleterious effect of hostility on adherence (serum P levels) was most pronounced among patients possessing the expectancy that positive health outcomes are not strongly contingent on the actions or advice of powerful others (i.e., healthcare providers). This suggests that individual differences in trait hostility and control expectancies jointly contribute to patient dietary and medication regimen nonadherence. However, similar analyses failed to show significant effects for hostility or PHLC in the prediction of fluid-intake adherence. In addition to identifying a personality trait potentially related to some aspects of patient adherence, these findings provided some support for the notion that the association between hostility and health may be mediated by differences in patient adherence and other health-related behaviors.

Two recent studies suggest that individual differences in attentional style (i.e., self-focusing tendencies, dispositional monitoring) may be associated with medical regimen adherence among ESRD patients (Christensen et al., 1996; Christensen, Moran, & Lawton, in press). In the most recent study, Christensen, Moran et al., (in press) classified 51 center hemodialysis patients on the monitoring dimension from the Miller Behavioral Styles Scale (MBSS; Miller, 1987) and examined the relation of monitoring

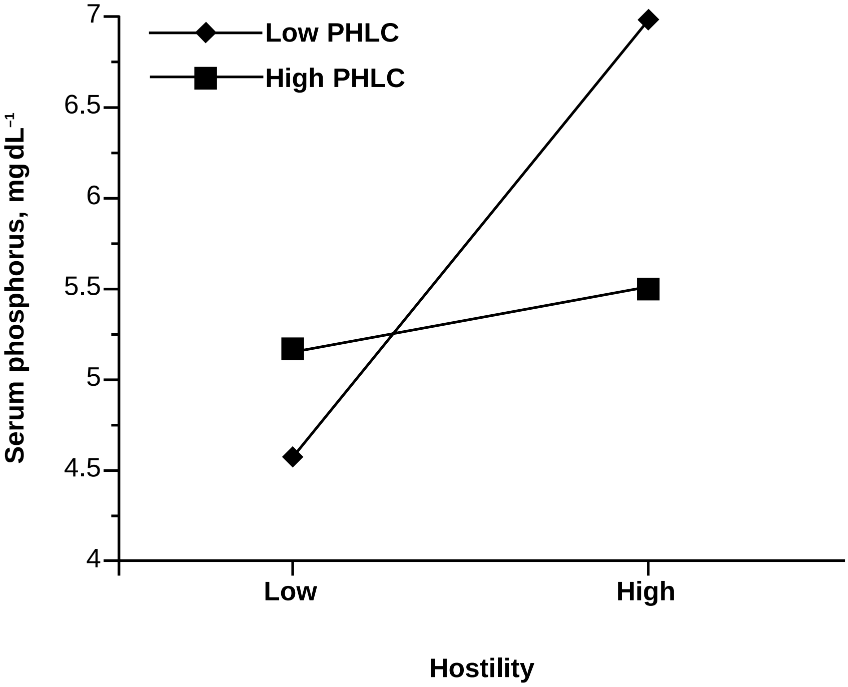

Figure 1 Effects of cynical hostility and powerful others health locus of control (PHLC) on patient adherence (from Christensen, Wiebe, & Lawton, in press).

scores to patient adherence. The MBSS was designed to assess the degree to which individuals are characteristically vigilant to threat-related cues or information. Patients scoring high on the monitoring dimension displayed poorer dietary and fluid-intake adherence relative to patients with less pronounced monitoring tendencies. Moreover, support was found for a mediational model suggesting that a lack of perceived control is largely responsible for the relationship between higher monitoring and poorer regimen adherence. This pattern suggests that among individuals facing severe, relatively uncontrollable chronic health stressors, a characteristic tendency to closely monitor threat-relevant information may exacerbate a perceived lack of control and contribute to patient nonadherence.

8.13.3.1.5 Patient coping and adherence

The general stress and coping paradigm (Lazarus & Folkman, 1984) is the foundation of much of the research literature on adaptation in chronic medical illness. Lazarus (1993) has suggested that the adaptational value of a particular type of coping may be a function of the characteristics of the stressful encounter being considered (e.g., controllability of the stressor). One recent study examined the potential interactive pattern involving coping, controllability of the stressor, and patient adherence, in a sample of 60 ESRD patients undergoing center hemodialysis (Christensen, Benotsch, Wiebe, & Lawton, 1995). Participants were asked to complete two forms of the Ways of Coping Checklist (WOC; Folkman, Lazarus, Dunkel-Schetter, De Longis, & Gruen, 1986). Instructions given for the measure were modified in order to obtain a situation-specific assessment of coping with two distinct stressful encounters differing in terms of how amenable each was to patient control. Directions for the first form asked subjects to recall and describe "the last time you experienced a problem with the dialysis procedure itself" (a relatively uncontrollable situation). Subjects were then asked to respond to the coping items based on how they dealt with the specific problem they had just described. Directions for the second form asked subjects to recall "the last time you experienced a problem with your fluid-intake or blood chemistries" (a relatively controllable situation). Subjects then completed another set of WOC items using this situation as the target incident.

Consistent with prediction, the relation of a particular coping strategy to adherence was dependent on the nature of the stressor ;being confronted. When confronting potentially controllable stressors, adherence with fluid-intake restrictions was best when patients employed planful problem solving. In contrast, when dealing with less controllable stressors, coping through the use of emotional self-control was associated with better adherence. Thus, the adaptive significance of specific coping strategies appeared to depend on the degree to which a specific stressful encounter was actually amenable to control.

8.13.3.2 Patient × Treatment Interactions in ESRD: Influence on Patient Adherence

Most studies attempting to identify determinants of patient adherence have examined only the main effects of psychosocial variables, while ignoring the tremendous heterogeneity of the ESRD and treatment context. Wiebe and Christensen (1996) have recently argued that relevant contextual or situational features should be explicitly assessed, and the interaction of these factors with patient variables tested directly. This general person-by-situation interaction perspective has a long history in clinical and health psychology (e.g., Dance & Neufeld, 1988; Schultheis, Peterson, & Selby, 1987). For example, across a variety of behavioral outcomes (e.g., smoking cessation, weight loss) individuals with highly active coping styles respond better to interventions that emphasize self-direction rather than direction by the therapist (Dance & Neufeld, 1988). However, the possibility of psychological attributes interacting with medical interventions to predict patient adaptation has rarely been explored.

The availability of a number of different treatments for ESRD each presenting a distinct set of patient demands offers a unique opportunity to apply the person × context framework to the understanding of patient adherence. From an interactional perspective, adherence should be best when the patient's characteristic or preferred style of coping is consistent with the contextual features or demands of the particular type of renal replacement treatment the patient is undergoing. Christensen, Smith, Turner, Holman, and Gregory (1990) tested this premise in a study of patients undergoing either center hemodialysis (a predominantly staff-directed treatment) or home hemodialysis (a treatment offering greater patient control). Patients were classified according to their characteristic or preferred style of coping using the "preference for behavioral involvement scale" from the Krantz Health Opinion Survey (Krantz, Baum, & Wideman, 1980). Adherence was predicted to be maximized in cases in which coping style classification matched the degree of actual patient control or involvement offered by the two types of dialysis treatment.

Consistent with prediction, among patients undergoing the more staff-directed center hemodialysis, a preference for active involvement in one's own healthcare delivery was associated with significantly worse dietary (i.e., potassium control) adherence. In contrast, for patients undergoing hemodialysis at home, where patient involvement and control is greater, patients with strong preferences for active involvement displayed better adherence. Moreover, among the patients dialyzing at home, those with low preferences for involvement in their own care had higher, clinically elevated serum K levels. Thus, adherence was best explained by the degree of congruency between patient coping style and the levels of involvement permitted or required by the particular treatment patients received.

In a related application of the person × treatment context hypothesis, Christensen, Smith, Turner, and Cundick (1994) examined medical regimen adherence among 52 staff-treated center hemodialysis patients and 34 self-treated CAPD patients. CAPD is a relatively new form of self-directed dialysis treatment that requires the patient to take an even more active role in treatment delivery to ensure success. Patients were classified on composite measures of coping style reflecting the tendency or motivation to exercise personal control in health contexts and to attend actively to threat-relevant information related to health and treatment.

As shown in Figure 2, the results of the cross-sectional analyses were consistent with the interactive model. Higher scores on a composite "information vigilance" dimension were associated with better dietary adherence (i.e., lower serum K levels) for self-treated CAPD patients, but poorer dietary adherence for center hemodialysis patients. In contrast, low information vigilance was associated with better dietary adherence for staff-treated hemodialysis patients, but poorer dietary adherence for CAPD patients. Analysis of prospective follow-up data revealed a highly similar, though only marginally significant pattern. This pattern again suggests that the degree of congruence between patient's coping style and the self-care demands of a prescribed medical treatment is an important determinant of adherence.

8.13.3.3 Improving Adherence Through Intervention

Given the clinical importance of adherence among ESRD patients it is not surprising that a number of studies have evaluated the effects of intervention techniques designed to improve adherence. A large majority of studies have evaluated the impact of behaviorally oriented intervention on adherence outcomes. There is modest evidence to suggest that behavioral

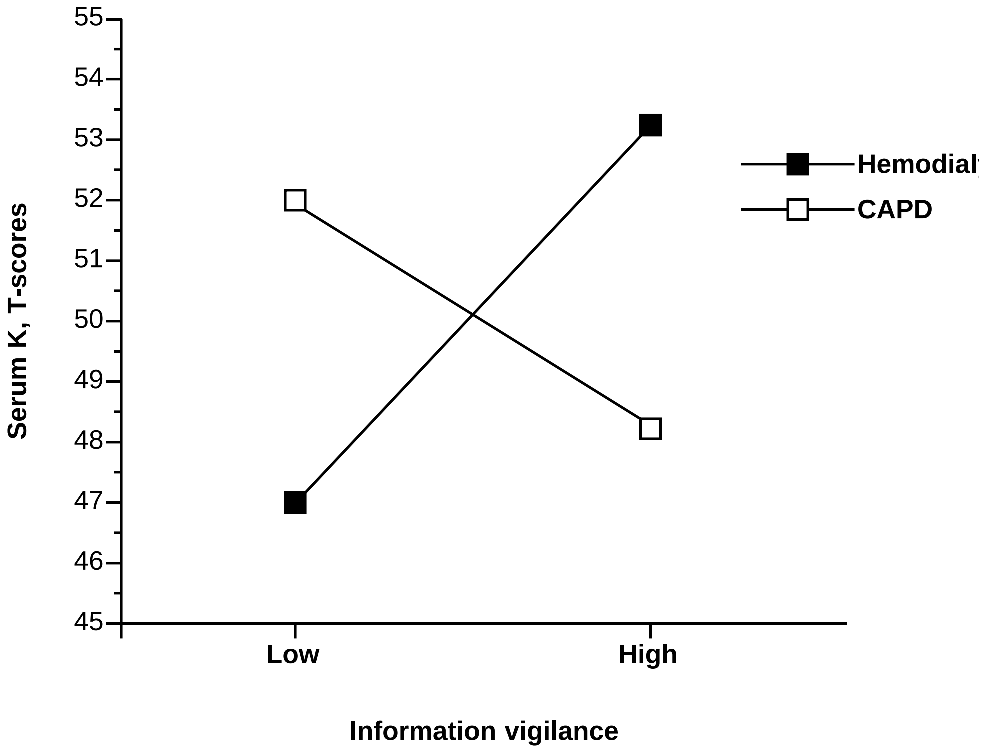

Figure 2 Effects of information vigilance and dialysis type (i.e., continuous ambulatory peritoneal dialysis (CAPD) vs. hemodialysis) on serum potassium (from Christensen, Smith et al., 1994).

strategies (e.g., self-monitoring, behavioral contracting, and positive reinforcement) are associated with improved adherence among hemodialysis patients (Barnes, 1976; Brantley, Mosley, Bruce, McKnight, & Jones, 1990; Hart, 1979; Hegel, Ayllon, Thiel, & Oulton, 1992; Keane, Prue, & Collins, 1981; Kirschenbaum, Sherman, & Penrod, 1987). However, most of these studies are limited to single-subject or small-group designs. A notable exception is a study evaluating several behavioral techniques to encourage adherence among 87 hemodialysis patients (Cummings et al., 1981). These authors reported that behavioral contracting and the use of various types of positive reinforcement (i.e., lottery tickets or verbal reinforcement) were all associated with significant improvement in both fluid-intake and medication adherence immediately following a six-week intervention period. However, these improvements did not persist at a three-month follow-up assessment.

One recent study used a small-group within-subject design to compare the efficacy of a cognitively oriented intervention (i.e., modification of health-related beliefs) to a behavioral intervention involving reinforcement and patient self-monitoring (Hegel et al., 1992). The cognitive approach was associated with little sustained change in fluid-intake adherence. However, the behavioral intervention did effect significant and sustained improvement in fluid-intake adherence over a two-month period. To date, no adherence intervention studies involving renal transplantation patients have been published.

The potential influence of modifying the patients" social environments or providing applied social support has yet to be examined in this population. However, research involving adherence in populations other than ESRD has provided evidence that interventions aimed at increasing social support are associated with more favorable patient adherence (Levy, 1983).

8.13.4 PATIENT DEPRESSION IN END-STAGE RENAL DISEASE

Individuals with ESRD face a variety of chronic recurrent stressors, significant lifestyle disruption, and threatened personal control. Thus, it is not surprising that high rates of depressed mood are commonly observed in ESRD patients (Christensen, Holman, Turner, & Slaughter, 1989; Craven, Rodin, & Littlefield, 1988; Hinrichsen, Lieberman, Pollack, & Sternberg, 1989; Rodin & Voshart, 1987; Smith, Hong, & Robson, 1985). Estimates of the prevalence of clinical depression in this population have varied substantially depending on differences in the method and criteria used to define a depressive disorder. Craven et al. (1988) reported that 45% of their ESRD patient sample was identified as depressed using the Beck Depression Inventory (BDI: i.e., BDI score greater than 10), but only 12% were diagnosed with major depressive disorder using more stringent criteria from the *Diagnostic and statistical manual of mental disorders* (3rd ed., *DSM-III*). Lowry and Atcherson (1980) reported an even higher rate (18%) of *DSM-III*-defined major depression among patients who were just beginning treatment. This finding suggests that psychological distress may be greatest at the outset of treatment. Other evidence suggests the rate of more chronic, moderate depression (i.e., dysthymia) is particularly high in this population (Craven, Rodin, Johnson, & Kennedy, 1987; Hinrichsen et al., 1989). In general, the available data suggest that 15–30% of ESRD patients meet criteria for a diagnosable mood disorder (either major depression or dysthymia). This high rate of depression is made even more alarming by epidemiological research reporting the suicide incidence among hemodialysis patients to be 10–100 times greater than the rate in the general population (Abram, Moore, & Westervelt, 1971; Haenel, Brunner, & Battegay, 1980).

A factor complicating the diagnosis of depression in this population is the confound between somatic symptoms of depression and physical symptoms of renal failure and side effects of treatment. Typical uremic symptoms seen in ESRD mimic classic somatic symptoms of depression. Like depression, uremia may produce irritability, cognitive deficits, decreased appetite, insomnia, apathy, and fatigue. Other conditions associated with ESRD such as anemia, electrolyte disturbances, and underlying systemic disease (e.g., diabetes) may also mimic depressive symptoms. In addition, ESRD patients may take medications such as antihypertensives, corticosteroids, and anti-inflammatory agents which can produce depressive side effects.

Given the overlap between physical changes associated with ESRD and some symptoms of depression, it is not surprising that nonsomatic or cognitive symptoms of depression appear to more accurately discriminate depressed from nondepressed ESRD patients (Craven et al., 1987; Hinrichsen et al., 1989; Smith et al., 1985). For example, Craven et al. (1987) reported that depressed mood, loss of interest, guilt, and concentration problems were significantly associated with a diagnosis of major depression, while only one somatic indicator, appetite and weight changes, was related to major depression in an ESRD sample. The *DSM-III-R* depression criteria of loss of energy, insomnia, and decreased sexual interest, were common in the entire sample and thus were not useful in distinguishing depressed from nondepressed patients.

8.13.4.1 Psychological Determinants of Depression

8.13.4.1.1 Control, stress, and depression

Much of the research involving determinants of depression has focused on the considerable loss of control encountered by ESRD patients. One of the most salient examples of a loss of control occurs when a renal transplant patient experiences organ rejection. Several studies suggest that depression is especially common among patients experiencing an unsuccessful transplant (Christensen, Holman et al., 1989; Johnson, McCauley, & Copley, 1982; Simmons et al., 1988). Christensen, Holman et al. (1989) compared levels of depression symptoms in patients with a functioning renal transplant, untransplanted patients on dialysis, and patients forced to return to dialysis after a transplant failure. After controlling for a variety of demographic factors and medical characteristics, patients composing the failed transplant group exhibited significantly greater depression than the other two groups. Patients with a functioning renal transplant had the lowest depression overall. This pattern is consistent with the view that depressive illness is often precipitated by the experience of uncontrollable, negative events (Seligman, 1975). However, this study did not consider potential cognitive moderators (e.g., attributional style) which have been proposed to play a central role in determining whether the experience of a stressor actually results in depressed mood (Abramson, Alloy, & Metalsky, 1989; Abramson, Seligman, & Teasdale, 1978).

Research involving renal transplant patients suggests that the experience of uncontrollable, aversive stressors is associated with greater levels of depression. Considerable chronic illness research has also examined the relation of individual differences in control appraisals to depression. Results of this research are mixed. The findings from a number of studies indicate that a more pronounced belief in personal control over health-related outcomes (an internal health locus of control) is associated with more favorable emotional adjustment including less depression (e.g., Devins et al., 1982; Poll & Kaplan-DeNour, 1980). However, other research has reported either no association between perceived control over health and adjustment (Fitzgerald, Tennen, Affleck, & Pransky, 1993) or has actually found greater perceived control to be associated with poorer adjustment (Burish et al., 1984).

Several recent studies have suggested that the adaptive significance of control appraisals may depend upon contextual or situational factors (Andrykowski & Brady, 1994; Christensen, Turner, Smith, Holman, & Gregory, 1991; Helgeson, 1992). This research has indicated that the degree of congruence between control beliefs and the constraints of the illness and treatment context may better explain emotional adjustment than either factor in isolation. Christensen et al. (1991) hypothesized that the association between health locus of control and depression in hemodialysis patients would vary as a function of the patient's treatment history (i.e., whether patients had experienced a renal transplant failure) and the severity of their disease. As seen in Figure 3 the results matched the interactive hypothesis. For patients with relatively low disease severity, depression (as measured by the BDI) was quite low overall. However, among severely ill patients, the belief that one's health is controllable (i.e., high internal health locus of control) was associated with less depression for those who had not previously experienced a failed renal transplant. However, high perceived control was associated with greater depression for patients who had returned to dialysis following an unsuccessful transplant. An analogous pattern was obtained when the locus of control measure reflected a belief in healthcare providers" ability to influence health outcomes (i.e., powerful others control). Thus, it appears that the extent to which patients" control beliefs are congruent with contextual factors (e.g., treatment outcomes) is an important determinant of the adaptiveness of those beliefs.

8.13.4.1.2 Illness intrusiveness

ESRD and its treatment are highly disruptive to an individual's lifestyle. Illness-related disruption is common in a variety of life domains including social and vocational functioning, marital, sexual, and family relationships and recreational activities (Devins, Mandins et al., 1990). Devins et al. (1984) suggested that the degree to which an illness interferes with central life domains is an important determinant of depression. Several studies have established that ESRD patients perceiving greater illness-related disruption of lifestyle and social activities report greater negative and less positive mood (Devins et al., 1984; Devins, Mandins et al., 1990; Sacks, Peterson, & Kimmel, 1990). The effect of perceived intrusiveness on depression has been found to persist even after controlling for physical disease severity (Sacks et al., 1990). Moreover, subjective appraisals of illness intrusiveness appear more important in terms of adjustment than do apparent objective differences (Devins et al., 1984; Devins, Mandins et al., 1990).

Eitel, Hatchett, Friend, Griffin, and Wadhwa (1995) recently examined the effects of perceived illness intrusiveness and clinically rated illness severity on depression in samples of staff-treated hemodialysis and self-treated CAPD patients. In this study, severely ill CAPD patients reported both greater perceived illness intrusiveness and higher levels of depression than did less physically ill CAPD patients. Moreover, CAPD patients with severe illness were more depressed and perceived greater intrusiveness than hemodialysis patients with similarly severe disease. Eitel et al. (1995) suggested that as medical illness becomes more severe, the greater control over treatment afforded by CAPD becomes a burden to very ill patients, resulting in greater perceived illness intrusiveness and ultimately more depression. These findings again underscore the importance of jointly considering both patient appraisals and contextual variables to understand depression in this population.

8.13.4.1.3 Social support and depression

Social support has long been associated with better psychological adjustment among the general population as well as a variety of medical populations (Cohen & Wills, 1985; Wortman & Conway, 1985). Similarly, among ESRD patients, various indices of social support have been associated with more favorable adjustment and less depression (Christensen, Turner, Slaughter, & Holman, 1989; Dimond, 1979; Kimmel et al., 1995; Moran & Christensen, 1996; Shulman, Pacey, Price, & Spinelli, 1987; Siegal, Calsyn, & Cuddihee, 1987). For example, in a study of 149 center hemodialysis patients, Kimmel et al. (1995) found higher scores on a generalized measure of perceived social support to be associated with lower BDI scores. Moreover, a lack of social support has been identified as a contributing factor to the high rates of suicide observed among ESRD patients (Abram et al., 1971; Foster & McKegney, 1978).

Both theory and empirical evidence suggest that the effect of support on well-being is most pronounced among individuals experiencing a high degree of stress (the "buffering model," Cohen & Wills, 1985). ESRD patients generally experience a high degree of stress relative to individuals not living with a chronic disease. However, there is considerable variability in

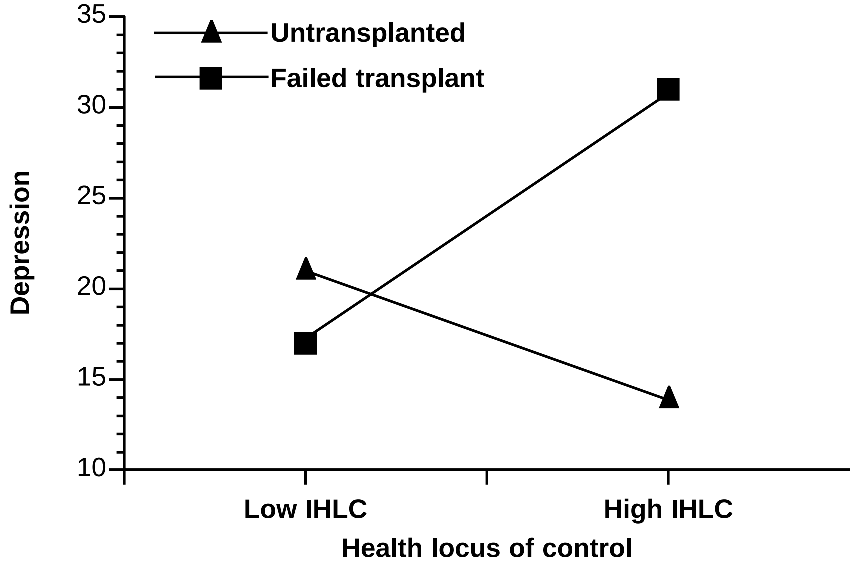

Figure 3 Effects of internal health locus of control (IHLC) and treatment status on adjusted Beck Depression Inventory scores for patients with severe disease (from Christensen et al., 1991).

the degree of impairment and stress different patients face. The buffering model would suggest that the salutary effect of support may be most pronounced among patients with relatively more severe illness. In a test of this model, Christensen, Turner et al. (1989) examined the effects of perceived family support and degree of illness-related impairment on emotional adjustment in a sample of 57 renal transplant patients. Consistent with the buffering model, severely ill patients perceiving a less supportive family environment (i.e., less cohesion and expressiveness and greater conflict) displayed significantly higher levels of depression and anxiety than patients with a more supportive family environment. Among patients experiencing relatively low illness-related impairment (and presumably less stress), the effect of family support was not significant.

8.13.4.2 Treatment of Depression in ESRD

The psychological treatment of depression in the ESRD population has received surprisingly little attention. Descriptive reports suggest that psychotherapeutic intervention for depression in this population often follows a behavioral or cognitive-behavioral orientation (Kirschenbaum, 1991). One recent study provides evidence that depression among ESRD patients can be successfully treated with psychotherapy (Hener, Weisenberg, & Har-Even, 1996). This controlled study compared the effectiveness of eight sessions of supportive therapy or cognitive-behavioral therapy (involving couples) to a no-intervention control group. Results indicated both treatment groups experienced significant reductions in depressive symptomatology compared to the control group. There were no significant differences in efficacy between the two active treatment groups.

Pharmacologic treatment of depression has also been shown to be effective in this population (Surman, 1987). However, antidepressant medications with pronounced anticholinergic effects (e.g., amitryptiline) are not well tolerated by ESRD patients while the efficacy and safety of some newer medications (e.g., sertraline) has not been well established in this population (see review by Kimmel, Weihs, & Peterson, 1993). These limitations to the use of antidepressant medications increases the need for careful evaluation of nonpharmacological alternatives.

8.13.5 PSYCHOSOCIAL INFLUENCES ON SURVIVAL

The high rate of depression typically observed among ESRD patients is an important clinical problem as it poses a clear detriment to patient quality of life. Growing evidence suggests that depression may also be related to earlier mortality among ESRD patients (Burton, Kline, Lindsay, & Heidenheim, 1986; Peterson et al., 1991; Shulman, Price, & Spinelli, 1989). In the most recent study, Peterson et al. (1991) reported that a cognitive item subset from the BDI (somatic items were eliminated to avoid a confound with disease severity) was a significant predictor of renal dialysis patient mortality over a two-year follow-up period. This effect persisted after controlling for several more traditional mortality risk factors (e.g., age, medical comorbidity). Moreover, Burton et al. (1986) reported item level analyses that suggest a subset of depression symptoms involving self-deprecating thoughts, pessimism, and helplessness may be particularly important determinants of earlier mortality. Several other studies involving similar methodologies have, however, failed to replicate a relationship between depression and survival (Christensen, Wiebe, Smith, & Turner, 1994; Devins, Mann et al., 1990; Husebye, Westlie, Styrvoky, & Kjellstrand, 1987).

A second psychosocial variable that has been examined as a predictor of earlier mortality in both nonclinical and medically ill populations is perceived social support (Berkman, 1985). However, evidence for the role of support in influencing survival among ERSD patients is limited. Recently, Christensen et al. (1994) reported that perceived family support is a significant predictor of survival among renal dialysis patients.

As seen in Figure 4, estimated five-year mortality rates among patients classified as having low family support were nearly three times higher than estimated mortality for high support patients (52% vs. 18%, respectively). The effect of family support was independent of the significant predictive influence of demographic and biochemical risk factors (i.e., age, blood urea nitrogen levels).

A study by Friend, Singletary, Mendell, and Nurse (1986) provided initial evidence that participation in a supportive group intervention may prolong survival among hemodialysis patients. Center hemodialysis patients participating in a largely unstructured patient support group were significantly more likely to be alive at the end of the 10-year study period compared to nonparticipants. Although this study was largely uncontrolled (e.g., no random assignment or control group comparison) the data underscore the need for further research examining the salutary influence of applied social support.

One of the most important unanswered questions from survival research involves a lack of evidence concerning the mediational mechanisms that may underlie the effect of depression or family support on survival. An often cited model linking a lack of social support to increased morbidity and mortality involves the influence of support on health or adherence behaviors (Cohen, 1988). Consistent with this view, an association between family support and patient adherence has been demonstrated (e.g., Christensen et al., 1992). Moreover, there is evidence that patient nonadherence is associated with earlier mortality in this population (Wolcott et al., 1986). Unfortunately, the only published study to date examining the potential mediating role of patient adherence failed to support the model (Christensen et al., 1994).

A second potential mechanism mediating the association of psychosocial factors to mortality involves diminished immunocompetence. Indirect evidence for the role of immune functioning comes from several sources. There is clear evidence of diminished cellular immunocompetence among ESRD patients (Keane & Maddy, 1989). In addition, systemic infection (septicemia) is a leading cause of death in this population (USRDS, 1995). Finally, both a lack of social support and elevated depression have been associated with diminished immunocompetence (O'Leary, 1990). However, no research has simultaneously evaluated the effect of immune functioning and support or depression on ESRD patient outcomes. Such evidence is needed to clarify the role that psychosocial factors may play in influencing patient mortality.

8.13.6 IMPLICATIONS FOR THE CONSULTING PSYCHOLOGIST

Mental health professionals have long served in a consultant role to renal dialysis and transplantation programs. However, until recently, little empirical knowledge existed for practitioners to draw from in approaching the range of clinical problems seen in this unique patient population. Several target problems or issues compose the core of the referral questions a consulting psychologist is likely to receive from a renal dialysis or transplantation program. In reviewing the activities of our own clinical health psychology consultation service over a 12-month period, the most common consultation request involved the assessment and treatment of depression (55% of ESRD patient referrals).

Assessing depression in this population presents a unique challenge given the confounding of physical changes secondary to a patient's medical condition with somatic or vegetative signs of depression. However, the modification of commonly used depression inventories and diagnostic schemes to emphasize cognitive symptoms of depression (e.g., guilt, anhedonia) rather than the less specific somatic symptoms (e.g., appetite changes, fatigue), has proven to be a useful strategy for valid depression assessment (Peterson et al., 1991; Sacks et al. 1990; Smith et al., 1985). For example, Sacks

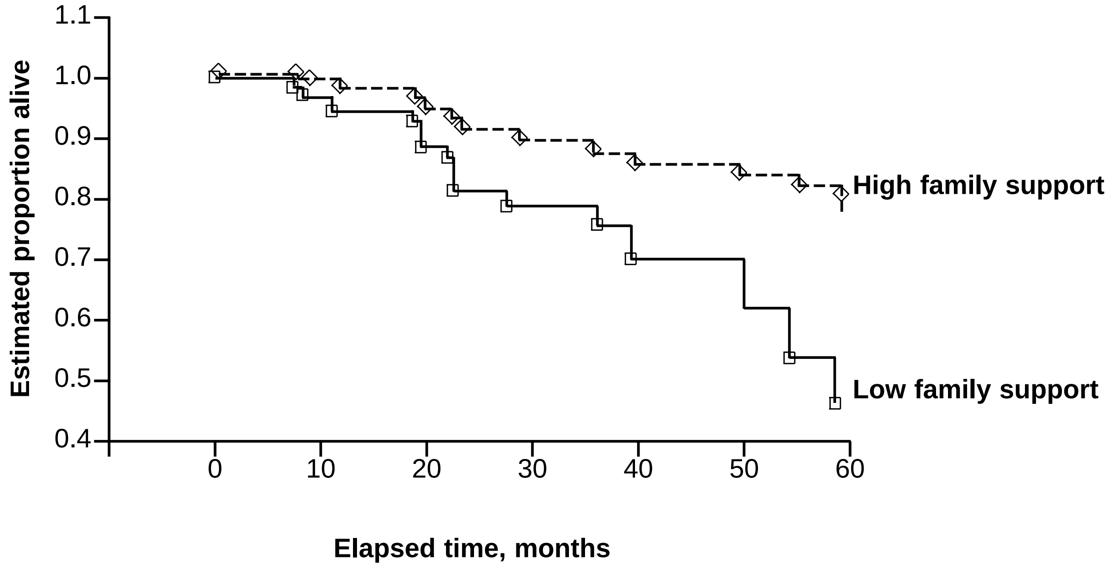

Figure 4 Estimated survival functions for high and low family support (from Christensen, Wiebe et al., 1994).

et al., (1990) have reported that items 1–14 of the widely used BDI compose a "cognitive depression index" that is less confounded by the severity of medical illness than the full-scale BDI. Use of this BDI item subset provides clinicians with a brief, valid assessment tool for depression screening.

Empirical evaluations of treatment efficacy for depressive illness in this population are extremely limited. Moreover, there is some evidence that depression among ESRD patients frequently goes untreated (Smith et al., 1985). There is insufficient data at this time to suggest which treatment approach may provide the greatest efficacy in this population. The results of comparative treatment studies (including both psychological and pharmacological intervention) involving depression in primary psychiatric patients (i.e., without comorbid medical problems) suggest few significant differences in outcome between treatments (Elkin et al., 1989). Accordingly, we would expect few between-treatment differences in outcome among depressed ESRD patients. Thus, we suggest that the specific choice of treatment strategy is less important than is simply initiating some form of active intervention for the depressed ESRD patient.

Patient nonadherence to the prescribed medical regimen reflects the second most common reason (30% of referrals to our service) for psychological consultation in this population. The available empirical studies suggest that behaviorally oriented intervention (e.g., self-monitoring, behavioral contracting) may provide an effective strategy for facilitating patient adherence (e.g., Cummings et al., 1981; Hegel et al., 1992). Studies of adherence determinants in this population have also produced a number of findings with implications for patient screening. For example, patients low in conscientiousness or high in hostility may be at risk for difficulties in adherence to the prescribed renal treatment regimen and may warrant early intervention.

An emerging challenge for the clinical health psychologist involves developing strategies for using psychosocial assessment data to assist in selecting the most beneficial medical or surgical treatment alternative for a given patient. Various types of renal dialysis (e.g., hemodialysis, CAPD) as well as renal transplantation are all medically acceptable treatment alternatives for the large majority of ESRD patients. Thus, medical considerations alone are not sufficient to identify the most appropriate treatment alternative.

The clinical literature suggests that psychosocial factors are often considered by healthcare providers when recommending a treatment alternative for a particular ESRD patient (Goodenough, Lutz, & Gregory, 1988; Gulledge, Bustza, & Montague, 1983; Knight, 1981). However, until recently there was little empirical basis for making such recommendations. The interactive framework discussed above has direct implications for choosing among medical treatment alternatives for ESRD patients in an effort to optimize outcomes. For example, patients possessing a highly active or vigilant style of coping with their healthcare should demonstrate better adherence and adjustment when undergoing more self-directed renal treatments (e.g., CAPD carried out at home; renal transplantation). In contrast, patients with a more passive coping orientation or those who choose to distract themselves from stressful stimuli are likely to show better adjustment when undergoing a staff-controlled treatment (e.g., hemodialysis treatment in a dialysis center).

8.13.7 CONCLUDING OBSERVATIONS

The medical management of ESRD has benefited from substantial technological advances over the past several decades. As the use of life-sustaining technology has become more routine, an increasing focus on psychosocial and behavioral issues has emerged. Accordingly, clinical psychology is in a position to make a significant contribution to the care of the ESRD patient. Knowledge from past research provides a basis for identifying patients at risk for behavioral (e.g., nonadherence) and emotional (e.g., depression) difficulties, both of which seem clearly linked to poorer patient outcomes. Moreover, a growing body of literature suggests that psychological assessment information can be useful in the selection of the most beneficial renal treatment modality for a particular patient. Future research can more clearly evaluate psychological intervention strategies that may be useful for enhancing patient adherence, improving emotional wellbeing, and ultimately prolonging patient survival. We believe clinical research with the ESRD population provides a unique opportunity to apply psychological concepts and procedures to improve patient outcomes as well as to contribute to health psychology theory and research more generally.

ACKNOWLEDGMENT

Preparation of this chapter was supported in part by National Institute of Diabetes and Digestive and Kidney Diseases grant DK49129–01, awarded to Alan Christensen.

8.13.8 REFERENCES

Abram, H. S., Moore, G. L., & Westervelt, F. B. (1971). Suicidal behavior in chronic dialysis patients. *American Journal of Psychiatry, 127,* 1199–1204.

Abramson, L. Y., Alloy, L. B., & Metalsky, G. I. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review, 96,* 358–372.

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87,* 49–74.

Andrykowski, M. A., & Brady, M. J. (1994). Health locus of control and psychological distress in cancer patients: Interactive effects of context. *Journal of Behavioral Medicine, 17,* 439–458.

Baldree, K. S., Murphy, S. P., & Powers, M. J. (1982). Stress identification and coping patterns in patients on hemodialysis. *Nursing Research, 31,* 107–112.

Bame, S. I., Petersen, N., & Wray, N. P. (1993). Variation in hemodialysis patient compliance according to demographic characteristics. *Social Science in Medicine, 37,* 1035–1043.

Barnes, M. R. (1976). Token economy control of fluid overload in a patient receiving hemodialysis. *Journal of Behavior Therapy and Experimental Psychiatry, 7,* 305–306.

Berkman, L. F. (1985). The relationship of social networks and social support to morbidity and mortality. In S. Cohen & S. L. Syme (Eds.), *Social support and health* (pp. 241–262). Orlando, FL: Academic Press.

Boyer, C. B., Friend, R., Chlouverakis, G., & Kaloyanides, G. (1990). Social support and demographic factors influencing compliance in hemodialysis patients. *Journal of Applied Social Psychology, 20,* 1902–1918.

Brantley, P. J., Mosley, T. H., Bruce, B. K., McKnight, G. T., & Jones, G. N. (1990). Efficacy of behavioral management and patient education on vascular access cleansing compliance in hemodialysis patients. *Health Psychology, 9,* 103–113.

Brown, J., & Fitzpatrick, R. (1988). Factors influencing compliance with dietary restrictions in dialysis patients. *Journal of Psychosomatic Research, 32,* 191–196.

Burish, T. G., Carey, M. P., Wallston, K. A., Stein, M. J., Jamison, R. N., & Lyles, J. N. (1984). Health locus of control and chronic disease: An external orientation may be advantageous. *Journal of Social and Clinical Psychology, 2,* 326–332.

Burton, H. J., Kline, S. A., Lindsay, R. M., & Heidenheim, A. P. (1986). The relationship of depression to survival in chronic renal failure. *Psychosomatic Medicine, 48,* 261–269.

Christensen, A. J., Benotsch, E., Wiebe, J., & Lawton, W. J. (1995). Coping with treatment-related stress: Effects on patient adherence in hemodialysis. *Journal of Consulting and Clinical Psychology, 63,* 454–459.

Christensen, A. J., Holman, J. M., Turner, C. W., & Slaughter, J. R. (1989). Quality of life in end-stage renal disease: Influence of renal transplantation. *Clinical Transplantations 3,* 46–53.

Christensen, A. J., Holman, J. M., Turner, C. W., Smith, T. W., & Grant, M. K. (1991). A prospective examination of quality of life in end-stage renal disease. *Clinical Transplantation, 5,* 40–47.

Christensen, A. J., Moran, P. J., & Lawton, W. J. (in press). Monitoring attentional style and medical regimen adherence. *Health Psychology.*

Christensen, A. J., & Smith, T. W. (1995). Personality and patient adherence: Correlates of the Five-Factor Model in renal dialysis. *Journal of Behavioral Medicine, 18,* 305–313.

Christensen, A. J., Smith, T. W., Turner, C. W., Holman, J. M., & Gregory, M. C. (1990). Type of hemodialysis and preference for behavioral involvement: Interactive effects on adherence in end-stage renal disease. *Health Psychology, 9,* 225–236.

Christensen, A. J., Smith, T. W., Turner, C. W., Holman, J. M., Gregory, M. C., & Rich, M. A. (1992). Family support, physical impairment, and adherence in hemodialysis: An investigation of main and buffering effects. *Journal of Behavioral Medicine, 15,* 313–325.

Christensen, A. J., Smith, T. W., Turner, C. W., & Cundick, K. E. (1994). Patient adherence and adjustment in renal dialysis: A person by treatment interactive approach. *Journal of Behavioral Medicine, 17,* 549–566.

Christensen, A. J., Turner, C. W., Slaughter, J. R., & Holman, J. M. (1989). Perceived family support as a moderator of psychological well-being in end-stage renal disease. *Journal of Behavioral Medicine, 12,* 249–265.

Christensen, A. J., Turner, C. W., Smith, T. W., Holman, J. M., & Gregory, M. C. (1991). Health locus of control and depression in end-stage renal disease. *Journal of Consulting and Clinical Psychology, 59,* 419–424.

Christensen, A. J., Wiebe, J. S., Benotsch, E. G., & Lawton, W. J. (1996). Perceived health competence, health locus of control, and patient adherence in renal dialysis. *Cognitive Therapy and Research, 20,* 411–421.

Christensen, A. J., Wiebe J. S., Edwards, D. L., Michels, J. D., & Lawton, W. J. (1996). Body consciousness, illness-related impairment, and patient regimen adherence in hemodialysis. *Journal of Consulting and Clinical Psychology, 64,* 147–152.

Christensen, A. J., Wiebe, J. S., & Lawton, W. J. (in press). Cynical hostility, expectancies about health care providers, and patient adherence in hemodialysis. *Psychosomatic Medicine.*

Christensen, A. J., Wiebe, J. S., Smith, T. W., & Turner, C. W. (1994). Predictors of survival among hemodialysis patients: Effect of perceived family support. *Health Psychology, 13,* 521–525.

Cohen, S. (1988). Psychosocial models of the role of social support in the etiology of physical disease. *Health Psychology, 7,* 269–297.

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin, 98,* 310–357.

Cook, W. W., & Medley, D. M. (1954). Proposed hostility and pharasaic-virtue scales for the MMPI. *Journal of Applied Social Psychology, 13,* 99–125.

Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R professional manual.* Odessa, FL: Psychological Assessment Resources.

Craven, J. L., Rodin, G. M., Johnson, L., & Kennedy, S.H. (1987). The diagnosis of major depression in renal dialysis patients. *Psychosomatic Medicine, 49,* 482–492.

Craven, J. L., Rodin, G. M., & Littlefield, C. (1988). The Beck Depression Inventory as a screening device for major depression in renal dialysis patients. *International Journal of Psychiatry in Medicine, 18,* 365–374.

Cummings, M. K., Becker, M. H., Kirscht, J. P., & Levin, N. W. (1981). Intervention strategies to improve compliance with medical regimens by ambulatory hemodialysis patients. *Journal of Behavioral Medicine, 4,* 111–127.

Cummings, M. K., Becker, M. H., Kirscht, J. P., & Levin, N. W. (1982). Psychosocial factors affecting adherence to medical regimens in a group of hemodialysis patients. *Medical Care, 20,* 567–580.

Dance, K. A., & Neufeld, R. W. J. (1988). Aptitude-treatment interaction research in the clinical setting: A review of attempts to dispel the "patient uniformity myth". *Psychological Bulletin, 104,* 192–213.

De Geest, S., Borgermans, L., Gemoets, H., Abraham, I., Vlaminck, H., Evers, G., & Vanrenterghem, Y. (1995). Incidence, determinants, and consequences of subclinical noncompliance with immunosuppressive therapy in renal transplant recipients. *Transplantation, 59,* 340–347.

Devins, G. M., Binik, Y. M., Gorman, P., Dattel, M., McCloskey, B., Oscar, G., & Briggs, J. (1982). Perceived self-efficacy, outcome expectancies, and negative mood states in end-stage renal disease. *Journal of Abnormal Psychology, 91,* 241–244.

Devins, G. M., Binik, Y. M., Hutchinson, T. A., Hollomby, D. J., Barre, P. E., & Guttman, R. D. (1984). The emotional impact of end-stage renal disease: Importance of patients" perceptions of intrusiveness and control. *International Journal of Psychiatry in Medicine, 13,* 327–343.

Devins, G. M., Mandin, H., Hons, R. B., Burgess, E. D., Klassen, J., Taub, K., Schorr, S., Letourneau, P. K. & Buckle, S. (1990). Illness intrusiveness and quality of life in end-stage renal disease: Comparison and stability across treatment modalities. *Health Psychology, 9,* 117–142.

Devins, G. M., Mann, J., Mandin, H., Paul, L. C., Hons, R. B., Burgess, E. D., Taub, K., Schorr, S., Letourneau, P. K., & Buckle, S. (1990). Psychosocial predictors of survival in end-stage renal disease. *Journal of Nervous and Mental Disease, 178,* 127–133.

Didlake, R. H., Dreyfus, K., Kerman, R. H., Van Buren, C. T., & Kahan, B. D. (1988). Patient noncompliance: A major cause of late graft failure in cyclosporine treated renal transplants. *Transplantation Proceedings, 20,* 63–69.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41,* 417–440.

Dimond, M. (1979). Social support and adaptation to chronic illness: The case of maintenance hemodialysis. *Research in Nursing and Health, 2,* 101–108.

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., Glass, D. R., Pilkonis, P. A., Leber, W. R., Docherty, J. P., Fiester, S. J., & Parloff, M. B. (1989). NIMH Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Archives of General Psychiatry, 46,* 971–982.

Evans, R. W., Manninen, D. L., Garrison, L. P., Jr., Hart, L. G., Blagg, C. R., Guttman, R. A., Hull, A. R., & Lowrie, E. G. (1985). The quality of life of patients with end-stage renal disease. *New England Journal of Medicine, 312,* 553–559.

Eitel, P., Hatchett, L., Friend, R., Griffin, K. W., & Wadhwa, N. K. (1995). Burden of self-care in seriously ill patients: Impact on adjustment. *Health Psychology, 14,* 457–463.

Fitzgerald, T. E., Tennen, H., Affleck, G., & Pransky, G. S. (1993). The relative importance of dispositional optimism and control appraisals in quality of life after coronary artery bypass surgery. *Journal of Behavioral Medicine, 16,* 25–43.

Flechner, S. M. (1994). Current status of renal transplantation: Patient selection, results, and immunosuppression. *Urologic Clinics of North America, 21,* 265–282.

Folkman, S., Lazarus, R. S., Dunkel-Schetter, DeLongis, A., & Gruen, R. J. (1986). Dynamics of a stressful encounter: Cognitive appraisal, coping, and encounter outcomes. *Journal of Personality and Social Psychology, 50,* 992–1003.

Foster, F. G., & McKegney, F. P. (1978). Small group dynamics and survival on chronic hemodialysis. *International Journal of Psychiatry in Medicine, 8,* 105–116.

Friend, R., Singletary, Y., Mendell, N. R., & Nurse, H. (1986). Group participation and survival among patients with end-stage renal disease. *American Journal of Public Health, 76,* 670–672.

Goodenough, G. K., Lutz, L. J., & Gregory, M. C. (1988). Home-based renal dialysis. *American Family Physician, 37,* 203–214.

Gulledge, A. D., Bustza, C., & Montague, D. K. (1983). Psychosocial aspects of renal transplantation. *Urologic Clinics of North America, 10,* 327–335.

Haenel, T. H., Brunner, F., & Battegay, R. (1980). Renal dialysis and suicide: Occurrence in Switzerland and in Europe. *Comprehensive Psychiatry, 21,* 140–145.

Hart, R. (1979). Utilization of token economy within a chronic dialysis unit. *Journal of Consulting and Clinical Psychology, 47,* 646–648.

Hartman, P. E., & Becker, M. H. (1978). Noncompliance with prescribed regimen among hemodialysis patients: A method of prediction and educational diagnosis. *Dialysis and Transplantation, 7,* 978–989.

Haynes, R. B. (1979). Determinants of compliance: The disease and the mechanics of treatment. In R. B. Haynes,

D. W. Taylor, & D. L. Sackett (Eds.), *Compliance in health care* (pp. 49–62). Baltimore: Johns Hopkins University Press.

Hegel, M. T., Ayllon, T., Thiel, G., & Oulton, B. (1992). Improving adherence to fluid-restrictions in male hemodialysis patients: A comparison of cognitive and behavioral approaches. *Health Psychology, 11,* 324–330.

Helgeson, V. S. (1992). Moderators of the relation between perceived control and adjustment to chronic illness. *Journal of Personality and Social Psychology, 63,* 656–666.

Hener, T., Weisenberg, M., & Har-Even, D. (1996). Supportive versus cognitive-behavioral intervention programs in achieving adjustment to home peritoneal kidney dialysis. *Journal of Consulting and Clinical Psychology, 64,* 731–741.

Hinrichsen, G. A., Lieberman, J. A., Pollack, S., & Sternberg, H. (1989). Depression in hemodialysis patients. *Psychosomatics, 30,* 284–289.

Hitchcock, P. B., Brantley, P. J., Jones, G. N., & McKnight, G. T. (1992). Stress and social support as predictors of dietary compliance in hemodialysis patients. *Behavioral Medicine, 18,* 13–20.

Husebye, D. G., Westlie, L., Styrvoky, T. J., & Kjellstrand, C. M. (1987). Psychological, social, and somatic prognostic indicators in old patients undergoing long-term dialysis. *Archives of Internal Medicine, 147,* 1921–1924.

Johnson, J. P., McCauley, C. R., & Copley, J. B. (1982). The quality of life of hemodialysis and transplant patients. *Kidney International, 22,* 286–291.

Kaplan, R. M., & Hartwell, S. L. (1987). Differential effects of social support and social network on physiological and social outcomes in men and women with Type II diabetes mellitus. *Health Psychology, 6,* 387–398.

Kaplan, R. M., & Simon, H. J. (1990). Compliance in medical care: Reconsideration of self-predictions. *Annals of Behavioral Medicine, 12,* 66–71.

Kaplan-DeNour, A., & Czaczkes, J. W. (1972). Personality factors in chronic hemodialysis patients causing noncompliance with medical regimen. *Psychosomatic Medicine, 34,* 333–344.

Keane, T. M., Prue, D. M., & Collins, F. L. (1981). Behavioral contracting to improve dietary compliance in chronic renal dialysis patients. *Journal of Behavior Therapy and Experimental Psychiatry, 12,* 63–67.

Keane, W. F., & Maddy, M. M. (1989). Host defenses and infectious complications in maintenance hemodialysis patients. In J. F. Maher (Ed.), *Replacement of renal function by dialysis* (pp. 856–880). Norwell, MA: Kluwer Academic.

Kiley, D. J., Lam, C. S., & Pollak, R. (1993). A study of treatment compliance following kidney transplantation. *Transplantation, 55,* 51–56.

Kimmel, P. L., Peterson, R. A., Weihs, K. L., Simmens, S. J., Boyle, D. H., Verme, D., Umana, W. O., Veis, J. H., Alleyne, S., & Cruz, I. (1995). Behavioral compliance with dialysis prescription in hemodialysis patients. *Journal of the American Society of Nephrology, 5,* 1826–1834.

Kimmel, P. L., Weihs, K., & Peterson, R. A. (1993). Survival in hemodialysis patients: The role of depression. *Journal of the American Society of Nephrology, 4,* 12–27.

Kirschenbaum, D. S. (1991). Integration of clinical psychology into hemodialysis programs. In J. Sweet, R. Rozensky, & S. Tovian (Eds.), *Handbook of clinical psychology in medical settings* (pp. 567–586). New York: Plenum.

Kirschenbaum, D. S., Sherman, J., & Penrod, J. D. (1987). Promoting self-directed hemodialysis: Measurement and cognitive-behavioral intervention. *Health Psychology, 6,* 373–385.

Knight, J. A. (1981). The liaison psychiatrist in kidney transplantation. *International Journal of Personality and Social Psychology, 39,* 977–990.

Krantz, D. S., Baum, A., & Wideman, M. V. (1980). Assessment of preferences for self-treatment and information in health care. *Journal of Personality and Social Psychology, 39,* 977–990.

Lazarus, R. S. (1993). Coping theory and research: Past, present, and future. *Psychosomatic Medicine, 55,* 234–247.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.

Lemaire, M., Fahr, A., & Maurer, G. (1990). Pharmacokinetics of cyclosporine: Inter- and intra-individual variations and metabolic pathways. *Transplantation Proceedings, 22,* 1110–1112.

Levy, R. L. (1983). Social support and compliance: A selective review and critique of treatment integrity and outcome measurement. *Social Science in Medicine, 17,* 1329–1338.

Lowry, M. R., & Atcherson, E. (1980). A short-term follow-up of patients with depressive disorder on entry into home dialysis training. *Journal of Affective Disorders, 2,* 219–227.

Manley, M., & Sweeney, J. (1986). Assessment of compliance in hemodialysis adaptation. *Journal of Psychosomatic Research, 30,* 153–161.

McCrae, R. R., & John, O. P. (1992). An introduction to the Five-Factor Model and its applications. *Journal of Personality, 60,* 175–215.

Miller, S. M. (1987). Monitoring and blunting: Validation of a questionnaire to assess styles of information-seeking under threat. *Journal of Personality and Social Psychology, 52,* 345–353.

Moran, P. J., & Christensen, A. J. (1996, March). *Social Support, Personality, and Patient Adaptation in Hemodialysis*. Paper presented at the Fourth International Congress of Behavioral Medicine, Washington, DC.

Morduchowicz, G., Sulkes, J., Aizic, S., Gabbay, U., Winkler, J., & Boner, G. (1993). Compliance in hemodialysis patients: A multivariate regression analysis. *Nephron, 64,* 365–368.

Oldenburg, B., MacDonald, G. J., & Perkins, R. J. (1988). Factors influencing excessive thirst and fluid intake in dialysis patients. *Dialysis and Transplantation, 17,* 21–40.

O'Leary A. (1990) Stress, emotion, and human immune function. *Psychological Bulletin, 108,* 363–382.

Peterson, R. A., Kimmel, P. L., Sacks, C. R., Mesquita, M. L., Simmens, S. J., & Reiss, D. (1991). Depression, perception of illness and mortality in patients with end-stage renal disease. *International Journal of Psychiatry in Medicine, 21,* 343–354.

Poll, I. B., & Kaplan-DeNour, A. (1980). Locus of control and adjustment to chronic hemodialysis. *Psychological Medicine, 10,* 153–157.

Rodin, G., & Voshart, K. (1987). Depressive symptoms and functional impairment in the medically ill. *General Hospital Psychiatry, 9,* 251–258.

Rorer, B., Tucker, C. M., & Blake, H. (1988). Long-term nurse–patient interactions: Factors in patient compliance or noncompliance to the dietary regimen. *Health Psychology, 7,* 35–46.

Rosenbaum, M. (1980). A schedule for assessing self-control behaviors: Preliminary findings. *Behavior Therapy, 11,* 109–121.

Rosenbaum, M., & Ben-Ari Smira, K. (1986). Cognitive and personality factors in the delay of gratification in hemodialysis patients. *Journal of Personality and Social Psychology, 51,* 357–364.

Rosenstock, I. M. (1966). Why people use health services. *Millband Memorial Fund Quarterly, 44,* 94–127.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80* (whole issue No. 609, 1).

Rovelli, M., Palmeri, D., Vossier, E., Bartus, S., Hull, D., & Schweizer, R. (1989). Noncompliance in organ transplant recipients. *Transplantation Proceedings, 21,* 833–834.

Ruggiero, L., Spirito, A., Bond, A., Coustan, D., & McGarvey, S. (1990). Impact of social support and stress on compliance in women with gestational diabetes. *Diabetes Care, 13,* 441–443.

Sacks, C. R., Peterson, R. A., & Kimmel, P. L. (1990). Perception of illness and depression in chronic renal disease. *American Journal of Kidney Diseases, 15,* 31–39.

Schneider, M. S., Friend, R., Whitaker, P., & Wadhwa, N. K. (1991). Fluid noncompliance and symptomatology in end-stage renal disease. Cognitive and emotional variables. *Health Psychology, 10,* 209–215.

Schultheis, K., Peterson, L., & Selby, V. (1987). Preparation for stressful medical procedures and person x treatment interactions. *Clinical Psychology Review, 7,* 329–352.

Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death.* San Francisco: Freeman.

Shulman, R., Pacey, I., Price, J. D. E., & Spinelli, J. (1987). Self-assessed social functioning on long-term hemodialysis. *Psychosomatics, 28,* 429–433.

Shulman, R., Price, J. D. E., & Spinelli, J. (1989). Biopsychosocial aspects of long-term survival on end-stage renal failure therapy. *Psychological Medicine, 19,* 945–954.

Siegal, B. R., Calsyn, R. J., & Cuddihee, R. M. (1987). The relationship of social support to psychological adjustment in end-stage renal disease patients. *Journal of Chronic Disease, 40,* 337–344.

Simmons, R. G., Abress, L., & Anderson, C. (1988). Rehabilitation after kidney transplantation. In G. J. Cerilli (Ed.), *Organ transplantation and replacement* (pp. 481–489). Philadelphia: Lippincott.

Smith, T. W. (1992). Hostility and health: Current status of a psychosomatic hypothesis. *Health Psychology, 11,* 139–150.

Smith, M. D., Hong, B. A., & Robson, A. M. (1985). Diagnosis of depression in patients with end-stage renal disease: Comparative analysis. *American Journal of Medicine, 79,* 160–166.

Somer, E., & Tucker, C. M. (1988). Patient life engagement, spouse marital adjustment, and dietary adherence of hemodialysis patients. *Journal of Compliance in Health Care, 3,* 57–65.

Somer, E., & Tucker, C. M. (1992). Spouse-marital adjustment and patient dietary adherence in chronic hemodialysis: A comparison of Afro-Americans and Caucasians. *Psychology and Health, 6,* 69–76.

Surman, O. S. (1987). Hemodialysis and transplantation. In T. Hackett & N. H. Cassem (Eds.), *Massachusetts General Hospital Handbook of Psychiatry* (pp. 380–402). Littleton, MA: PSG Publishing.

US Renal Data System (1995) *USRDS Annual Report.* Bethesda, MD: National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases.

Wallston, K. A., Wallston, B. S., & DeVellis, R. (1978). Development of the Multidimensional Health Locus of Control Scales. *Health Education Monographs, 6,* 160–170.

Weed-Collins, M., & Hogan, R. (1989). Knowledge and health beliefs regarding phosphate-binding medication in predicting compliance. *ANNA Journal, 16,* 278–283.

Wiebe, J. S., & Christensen, A. J. (1996). Patient adherence in chronic illness: Personality and coping in context. *Journal of Personality, 64,* 815–835.

Witenberg, S. H., Blanchard, E. B., Suls, J., Tennen, H., McCoy, G., & McGoldrick, M. D. (1983). Perceptions of control and causality as predictors of compliance and coping in hemodialysis. *Basic and Applied Social Psychology, 4,* 319–336.

Wolcott, D. W., Maida, C. A., Diamond, R., & Nissenson, A. R. (1986). Treatment compliance in end-stage renal disease patients on dialysis. *American Journal of Nephrology, 6,* 329–338.

Wortman, C. B., & Conway, T. L. (1985). The role of social support in adaptation and recovery from physical illness. In S. Cohen & S. L. Syme (Eds.), *Social support and health* (pp. 281–302). New York: Academic Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.14
Respiratory Disorders

THOMAS L. CREER and DEIRDRE A. LEVSTEK
Ohio University, Athens, OH, USA

8.14.1 INTRODUCTION 340

8.14.2 DISEASE STATE 340
  8.14.2.1 Definitions of Asthma and COPD 340
    8.14.2.1.1 Asthma 340
    8.14.2.1.2 Chronic obstructive pulmonary disease 341
  8.14.2.2 Epidemiology of Respiratory Disorders 342
  8.14.2.3 Morbidity Associated with COPD and Asthma 342
  8.14.2.4 Mortality Related to COPD and Asthma 343

8.14.3 FACTORS ASSOCIATED WITH COPD AND ASTHMA 343
  8.14.3.1 Environmental Risks for COPD and Asthma 343
  8.14.3.2 Individual Factors Associated with COPD and Asthma 343
  8.14.3.3 Genetic Contributions to Respiratory Disease 344
  8.14.3.4 Prevention of Respiratory Disease 344

8.14.4 MEDICAL MANAGEMENT OF COPD AND ASTHMA 344
  8.14.4.1 Evaluation of Respiratory Disorders 344
  8.14.4.2 Ratings of Illness Severity and Prognosis 345
  8.14.4.3 Medications used to Treat COPD and Asthma 345
    8.14.4.3.1 Maintenance or preventive medications 345
    8.14.4.3.2 Bronchodilators 346
    8.14.4.3.3 Oxygen therapy 346

8.14.5 PSYCHOLOGICAL ASPECTS OF COPD AND ASTHMA 346
  8.14.5.1 Impact on the Individual 347
    8.14.5.1.1 Emotional impact 347
    8.14.5.1.2 Cognitive impact 348
    8.14.5.1.3 Behavioral impact 349
  8.14.5.2 Impact on Family and Social Environment 349

8.14.6 ADHERENCE TO TREATMENT REGIMENS 350
  8.14.6.1 Medication Compliance 350
  8.14.6.2 Lifestyle Changes 351
    8.14.6.2.1 Activity restrictions 351
    8.14.6.2.2 Environmental prescriptions 351
    8.14.6.2.3 Smoking cessation 351

8.14.7 PSYCHOLOGICAL INTERVENTIONS FOR COPD AND ASTHMA 352
  8.14.7.1 Pulmonary Rehabilitation for COPD 352
  8.14.7.2 Asthma Management 353
    8.14.7.2.1 Stage 1: Psychotherapy, particularly psychoanalysis 353
    8.14.7.2.2 Stage 2: Behavioral approaches 353
    8.14.7.2.3 Stage 3: Self-management 353

8.14.8 IMPLICATIONS FOR HEALTH AND CLINICAL PSYCHOLOGY 354
  8.14.8.1 COPD 354
    8.14.8.1.1 Smoking cessation 354
    8.14.8.1.2 Assessment 355

8.14.8.1.3 Integration of behavioral skills 355
8.14.8.1.4 Extension of behavioral expertise 355
8.14.8.2 Asthma 355
8.14.8.2.1 Medication compliance 355
8.14.8.2.2 Avoidance of attack triggers 355
8.14.8.2.3 Environmental control 355
8.14.8.2.4 Refinement of self-management processes 356
8.14.8.2.5 Maintenance of self-management skills 356
8.14.8.2.6 Design of studies 356
8.14.9 CONCLUSIONS 356
8.14.10 REFERENCES 357

8.14.1 INTRODUCTION

Behavioral scientists are frequently asked to develop techniques to help treat chronic obstructive pulmonary disease (COPD) and asthma in an effective and efficient manner. Terms such as self-care, collaborative management, and self-management—denoting processes traditionally associated with behavioral science—have increasingly entered the vocabulary of medical and health care personnel. Knowledge about the pathology of asthma and COPD has led to conclusions that environmental and behavioral management are, in most cases, as important as medical treatment in the control of chronic respiratory disorders. Furthermore, risk prediction and prevention efforts, both involving behavioral change, are increasingly viewed as ways to lessen the impact of morbidity and mortality related to asthma and COPD. The range of respiratory disorders is expansive, and includes patients with airflow limitation who exhibit a variety of symptoms and responses to treatment. Two major types of chronic respiratory diseases, asthma and COPD, are the most likely types of chronic respiratory conditions a health or clinical psychologist is likely to encounter. Optimal application of behavioral interventions requires a basic knowledge of these respiratory disorders and their management.

8.14.2 DISEASE STATE

8.14.2.1 Definitions of Asthma and COPD

Asthma, emphysema, and chronic bronchitis are chronic diseases that involve airflow obstruction due to a narrowing of the airways. At one time, the three conditions were subsumed under the title of COPD. In 1987, however, COPD was defined as "a disorder characterized by abnormal tests of expiratory flow that do not change over several months observation" (American Thoracic Society, 1987, p. 225). The definition was written to emphasize that irreversibility of lung function is likely in COPD, consisting of emphysema and chronic bronchitis, and distinguishes the condition from asthma where reversibility is common. More recently, the American Thoracic Society (1995, p. 578) revised the definition of COPD, and described it as a "disease state characterized by the presence of airway obstruction due to chronic bronchitis or emphysema; the airflow obstruction is generally progressive, may be accompanied by airway hyperreactivity, and may be partially reversible." Symptoms likely to signal acute or chronic obstruction, such as wheeze, cough, and dyspnea, are not in themselves diagnostic as they occur in many types of pulmonary disease, including both COPD and asthma (Petty, 1990). In addition, differences between chronic bronchitis and late-onset asthma in adults are subtle and by no means clear-cut (Bernstein, 1988). The major difference between COPD and asthma is the salience of inflammation, involving complex cellular and chemical mediators, as a characteristic of asthma but not of COPD.

8.14.2.1.1 Asthma

The American Thoracic Society (1993) suggested that asthma could be distinguished from COPD in that a patient's clinical status is variable, and involves airway hyperreactivity associated with irritants such as dust, allergens, and smoke or fumes. The National Asthma Education Program (1991, p. 1) defined asthma as

> a lung disease with the following characteristics: (1) airway obstruction that is reversible (but not completely in some patients) either spontaneously or with treatment; (2) airway inflammation; and (3) airway hyperreactivity to a variety of stimuli.

Two slightly different definitions of asthma are found in the *International consensus report on the diagnosis of asthma* (National Heart, Lung, and Blood Institute, 1992) and the *Global initiative for asthma* (National Institutes for Health, 1995). Despite the fact that each successive definition has become longer and more complex, any of the three reports established a set of guidelines that can be followed to treat asthma. Indeed, there are indications that when guidelines presented by the National Asthma Education Program (1991) are followed, improvement in the treatment of patients with asthma occurs (Lantner & Ros, 1995). It is recommended that any set of guidelines be adopted by medical and behavioral scientists who work with asthma. Unfortunately, as Marwick (1995) pointed out, only one of 10 physicians surveyed in the US was familiar with or knew that the National Asthma Education Program (1991) guidelines had been issued. This finding not only has implications for the control of asthma, but it suggests problems in professional education that must be considered by clinical and health psychologists (Creer, Levstek, & Reynolds, in press).

Over the years, three characteristics of asthma have made the disorder especially challenging to treat: (i) the intermittency; (ii) variability; and (iii) reversibility of airway obstruction. These factors are a function of individual differences among patients, and variations in the number of asthma triggers, severity of disease, degree of patient and medical control, and other mediating variables (Creer, 1979, 1982; Renne & Creer, 1985). Each characteristic must be analyzed in order to understand and treat asthma.

Intermittency refers to the frequency of asthma episodes. The frequency of attacks can range from several episodes over a period of days to the passage of several months or even years between attacks. One patient may experience severe asthma episodes and then not have any symptoms for an extended period of time, whereas another patient may have symptoms most days of the year (what is called perennial asthma). Both medical and behavioral scientists must accept that intermittency is a common characteristic of asthma, otherwise a false sense of control over asthma may be assumed by both health care professionals and patients. Patients, in particular, may not expect future episodes and be unprepared to deal with them (Renne & Creer, 1985).

Variability is the severity of asthma across patients or within a given patient. Both the National Asthma Education Program (1991) and the *International consensus report* (National Heart, Lung, and Blood Institute, 1992) offered criteria for judging both asthma severity and asthma attacks as mild, moderate, and severe. According to a report edited by Spector and Nicklas (1995), severity should be rated by analyzing a combination of objective and subjective criteria; these include patient symptoms, activity restrictions, pulmonary function, hyperresponsiveness of airways, health care usage, and medication use. As noted, it is common for asthma severity to be rated as mild, moderate, or severe. However, the report by Spector and Nicklas (1995, p. 749) stressed that "asthma severity is a continuum across the population and often within a given individual, and that some characteristics of asthma may be more applicable in defining the severity in one patient, whereas different characteristics may be more applicable in another." Severity classification can be misleading to patients; it is again important to recognize that patient expectations are influenced by their perceptions of episode severity. If most of their attacks are mild, patients are often unprepared to help control more severe episodes (Renne & Creer, 1985).

Reversibility refers to the degree to which airway obstruction reverses either spontaneously or with appropriate treatment. Reversibility has long been a distinguishing feature of asthma, but thought regarding the characteristic is changing as researchers learn more about the nature of inflammation in asthma and its underlying pathology (Sanford, Weir, & Pare, 1996; Spector & Nicklas, 1995). Creer and Levstek (1997) noted that the concept of reversibility is complicated. They pointed out that airway reversibility, like severity, must be perceived as a continuum: whereas most patients show complete reversibility of airway obstruction when treated appropriately, others experience some degree of airway obstruction even with intensive medical treatment. The fact that reversibility of airway obstruction can occur either in response to therapy or spontaneously presents a challenge when attempting to show cause and effect relationships between treatment and remission of asthma symptoms (Creer, 1982; Creer & Levstek, 1997).

8.14.2.1.2 Chronic obstructive pulmonary disease

Until recently, there was no commonly accepted definition for COPD (Snider, 1995). The American Thoracic Society (1995), in their guidelines for the treatment of COPD, developed a working definition for COPD that includes three classifications: airflow obstruction, emphysema, and chronic bronchitis. These characteristics are as follows:

(i) Abnormal tests of expiratory flow or airway obstruction are prominent in COPD patients. They are indicative of airflow obstruction, which is progressive in nature. Airflow obstruction has been attributed to the alteration of the small airways and, to some degree, to bronchoconstriction.

(ii) Emphysema is a state of progressive airflow obstruction due to abnormal permanent enlargement of airspaces distal to the terminal bronchioles and alveoli. The destruction is described as a lack of uniformity in the pattern of airspace enlargement, and is a major cause of airflow obstruction for COPD patients with emphysema. That the definition of emphysema is in terms of pathological changes was emphasized by the American Thoracic Society (1995).

(iii) Chronic bronchitis is the presence of a persistent mucous cough of no known etiology. Specifically, chronic bronchitis is defined clinically by the persistence of a chronic productive cough for three months or more in each of two successive years in patients in whom other causes of chronic cough have been ruled out. The American Thoracic Society definition excludes patients with either chronic bronchitis or emphysema who do not exhibit persistent airflow obstruction, and includes a subset of asthma patients in whom airflow obstruction is not totally reversible (Snider, 1995). The symptom configuration for COPD patients and their treatment varies according to the severity and progression of the disease state. COPD normally follows a progressive course in that patients exhibit a rapid decline in lung function (measured by lung volume), symptoms of cough and dyspnea, more acute chest illness, and later in the disease, hypoxemia (American Thoracic Society, 1995).

8.14.2.2 Epidemiology of Respiratory Disorders

Prevalence rates for COPD and asthma continue to climb. The American Thoracic Society (1995) estimated that 14 million people, about 3–5% of the population in the USA, suffer from COPD; this represents a 42% increase in prevalence since 1982. The prevalence of asthma ranges between 10 and 15 million people in the USA (Centers for Disease Control, 1992, 1996a; Evans, 1993). However, this estimate is probably low, as Marwick (1995) suggested 25 million Americans experience asthma.

Rates of COPD and asthma vary by gender, age, and race. The American Thoracic Society (1995) reported that COPD prevalence was increasing at a significantly higher rate for women (increasing by 30% in 1985 alone) than for men (a slight increase). Despite this increase, Sherrill, Lebowitz, and Burrows (1990) found a higher prevalence of COPD in men (4–6%) than in women (1–3%). Data from the Centers for Disease Control (1996b) indicated that in people aged 5–34 years, considered the most stable population to study, the rate of asthma increased 52%, from 34.6 per 1000 to 52.6 per 1000 in the years between 1981 and 1992. The rate of asthma increased by 29% for males and 82% for females. The prevalence rate for pediatric asthma is estimated at 7% in the USA, a rate that is increasing more rapidly than for the total population (Centers for Disease Control, 1996a; Evans, 1993). Racial differences for adult asthma are notable: across age groups, both Black and Hispanic minorities in the USA have a higher prevalence of asthma than Whites (Evans, 1993; Coultas et al., 1993). The Centers for Disease Control (1992) reported that between 1981 and 1988, the annual prevalence rate of asthma among Black females increased by 45%; during the same period, the annual prevalence rate of asthma for White females increased by 63%.

8.14.2.3 Morbidity Associated with COPD and Asthma

Asthma morbidity has been closely studied by epidemiologists. Of greatest interest are data associated with physician contacts, hospitalization, disability, lost work and school days, and financial costs. Recent data has shown increasing morbidity cost due to asthma and COPD (American Thoracic Society, 1995; Centers for Disease Control, 1992, 1996a; Evans, 1993; Stempel, Hedblom, Duncanin-Robbins, & Sturm, 1996). Physician contacts, a primary source of data, have risen steadily since 1980 (Centers for Disease Control, 1992). Hospitalizations due to COPD and asthma are also increasing at a greater rate for females, minorities, and children. As of 1990, females were more likely to be hospitalized for asthma than males; Black adults with asthma were more than twice as likely to be hospitalized as Whites (Centers for Disease Control, 1992). The annual hospitalization rate for children and young adults with asthma (aged 0–24) years showed an increase of 28%, with rates consistently higher for Blacks than for Whites. Overall, asthma is the leading reason for hospitalization among children in the USA (Reed, 1986).

The financial burden associated with asthma can be separated by inpatient costs, outpatient costs, and outpatient medication costs. Weiss, Gergen, and Hodgson (1992) found the economic impact of asthma was an estimated US $6.2 billion in 1990. This figure included direct costs, such as hospitalization and medication costs, and indirect costs, such as lost productivity due to missed work and school. Stempel et al. (1996) reported that asthma patients belonging to one of four large HMOs had an annual expenditure of about US $467 per patient, with the greatest proportion of spending on medications (38%) and inpatient hospital charges (32%). Direct costs for asthma were estimated to be in excess of US $5 billion annually (Stempel, Sturm, Hedblom, & Durcainin-Robbins, 1995).

8.14.2.4 Mortality Related to COPD and Asthma

The annual death rate for COPD is currently estimated at 18.6 per 100 000, a 33% increase since 1979. Gender differences occur in older individuals with COPD (over age 55 years), with death rates for men by age 70 being twice that of women with COPD (American Thoracic Society, 1995). Racial differences for patients with COPD indicate that mortality due to COPD is lower for Blacks than for Whites, a finding that is partly due to the lower rates of COPD in minority populations. Death resulting from COPD occurs with greater frequency than for asthma, in part because of the commonly advanced age of diagnosis and of the progressive nature of COPD. Asthma is a treatable disorder, however. This makes the death rate from asthma, however small, a significant concern for health care professionals. Until 1980, asthma mortality was on the decline. Since that time, however, there has been a 31% increase in deaths from asthma. Death rates are generally higher for older patients with asthma, but there are a disproportionate number of deaths due to asthma in children (Evans, 1993). The most recent statistics on asthma mortality for children indicate that the rate nearly doubled from 1980 to 1993 for children aged 5–14 years, and increased an alarming 118% from 1980 to 1993 for patients aged 0–24 years (Centers for Disease Control, 1996a). Of this group, death rates have consistently been the highest for Blacks aged 15–24 years (from 2.5 per million in 1980 to 5.2 per million in 1993).

8.14.3 FACTORS ASSOCIATED WITH COPD AND ASTHMA

8.14.3.1 Environmental Risks for COPD and Asthma

The greatest risk factor for developing COPD is cigarette smoking. It is estimated that 80–90% of the risk for COPD can be accounted for by tobacco smoke (American Thoracic Society, 1995). Only about 15% of all smokers develop COPD; about 15% of one-pack-per-day and 25% of two-pack-per-day cigarette smokers will develop COPD if they continue smoking (Fletcher, Peto, Tinker, & Speiter, 1976). There is no documented direct link between smoking and asthma. Maternal and paternal smoking, however, have been correlated with severe asthma and, in some cases, with early onset asthma (Sherrill et al., 1990; Evans, 1993; Kaplan & Mascie-Taylor, 1989). Passive smoking has also been implicated as a potential risk factor for developing asthma, and there is often a worsening of asthma symptoms in the presence of tobacco smoke (Spector & Nicklas, 1995).

Air pollution is an irritant for exacerbations of COPD and asthma. However, it is unknown whether air pollutants are causative to the onset of respiratory disease. Air pollutants include smoke, particulate matter, acid rain, reactive gases, and ozone. The negative effects of these airborne pollutants can be damaging to the structure and functional capabilities of the lung, especially with prolonged exposure and exposure during exercise. It is thought that asthma patients are hypersensitive to pollutants. Additionally, exposure to pollutants increase the patients' sensitivity to other triggers for asthma. The relationship of asthma to air pollution is complicated by the fact that both asthma and pollution vary according to weather and outdoor temperature. However, temperature-adjusted rates for air pollution were still significantly related to increases in asthma attacks. Occupational factors are associated with increased prevalence of chronic airway obstruction, quicker rate of lung function decline, and higher mortality from COPD. One subset of emphysema, centricinar emphysema, is linked to long-term exposure to ambient particles such as coal dust; as expected, centricinar emphysema is widespread among coal workers (American Thoracic Society, 1995).

8.14.3.2 Individual Factors Associated with COPD and Asthma

The major demographic difference between COPD and asthma is the occurrence of the disorders at different ages. About 25% of childhood asthma cases persist into adulthood, and more adult women are likely to be diagnosed with asthma than men (Burrows, 1991; Spector & Nicklas, 1995). Gender differences in childhood asthma favor higher prevalence rates in boys and, as described earlier, racial differences are clear-cut for childhood asthma in that Black children are 2.5 times more likely than White children to be diagnosed with asthma. Racial differences are also noted in the COPD population, with a lower prevalence of COPD in Blacks (3.2%) and other minorities than in Whites (6.2%). The relationship of allergic disease to asthma has been well documented (Evans, 1993; Kaliner & Lemanske, 1992; Sherrill et al., 1990). Studies have repeatedly demonstrated that asthma prevalence is closely linked to immunoglobin levels associated with allergy (Sherrill et al.). Asthma has been shown to be, in many cases, allergy-based, induced by environmental agents and allergens, including pollens, fumes, smoke, weeds, grasses, and hay (Kaliner & Lemanske, 1992; Kaplan & Mascie-Taylor, 1989; Spector & Nicklas, 1995). Other risk factors found to predict childhood asthma include residential regions, home environment, parental smoking, low birthrate, young maternal age, low family income, and urban environments (Coultas et al., 1993; Kaplan & Mascie-Taylor, 1989).

8.14.3.3 Genetic Contributions to Respiratory Disease

There exists an hereditary predisposition for both asthma and COPD. As early as the 1800s, hereditary factors were implicated in the onset of COPD, especially emphysema (Sherrill et al., 1990). In the twentieth century, researchers found a known hereditary defect, α-1-antitrypsin deficiency, which occurs in about 1% of COPD cases. α-1-Antitripsin deficiency is suspected in early-onset cases of COPD where there is a family history of lung disease and the patient is a nonsmoker. Studies of monozygotic and dyzygotic twins have demonstrated some degree of family aggregation for COPD, pulmonary function, asthma, and allergic disease (American Thoracic Society, 1995; Sherrill et al.). In households where one parent has a history of an obstructive lung disease, there is a significantly higher prevalence for asthma and bronchitis, compared to households where no adult members have respiratory disease (Sherrill et al.). A report by Sanford, Weir, and Pare (1996) on the genetics of asthma outlined the heritability of asthma, and suggested that individuals are genetically susceptible for allergic airway inflammation, but that asthma will only occur if the individual is exposed to the correct allergens; important in this equation are the timing, intensity, and mode of exposure. Two common viral agents, respiratory syncytial virus (RSV) and the parinfluenza virus, are particularly prominent in the onset of childhood asthma (Ellis, 1993).

8.14.3.4 Prevention of Respiratory Disease

The Centers for Disease Control (1992) declared that prevention and control of known risk factors must be addressed as part of an overall plan to reduce the morbidity and mortality associated with respiratory disease. Prevention efforts for people with asthma have been twofold: (i) to control the symptoms of asthma, and (ii) to prevent the occurrence of asthma attacks (Spector & Nicklas, 1995). Factors identified as triggers for specific asthma attacks include: (i) viral respiratory infections; (ii) allergens, such as mold, animal dander, cockroaches, dust mites, and airborne pollens; (iii) environmental irritants such as fumes and strong odors, air pollutants, and tobacco smoke; (iv) ingested chemicals such as aspirin and food additives; (v) exercise or inhalation of cold air; (vi) weather changes; and (vii) emotional reactions, especially laughing, crying, and shouting. Despite this knowledge, patients and health care personnel are limited in their ability to identify the onset and precipitating factors of a specific asthma attack. There is not only great variation in the interactions produced by variables in a given individual, but a matrix of stimuli that impinge upon that patient at a given moment. In addition, immediate- or late-onset asthma that may follow exposure to a stimulus further limits the ability to determine what produced a given asthma episode.

Sherrill et al. (1990) derived a statistical model for risk prediction of COPD by looking at prediction of low lung function over time and how it related to certain risk factors. Their initial formulations have included initial lung function, age, gender, and smoking as the major determinants of COPD. However, these risk factors are likely to interact with the same variability as they do for asthma patients, leaving researchers and clinicians alike still searching for an understanding of the preclinical determinants of COPD (Burrows, 1991).

8.14.4 MEDICAL MANAGEMENT OF COPD AND ASTHMA

Medical treatment of COPD differs somewhat from that of asthma. It is based on the consideration of the state and severity of disease, and the configuration of presenting symptoms. Recognition of the underlying pathology of respiratory diseases has helped delineate specific treatments for different respiratory disorders.

8.14.4.1 Evaluation of Respiratory Disorders

It is imperative for practitioners to: (i) conduct a comprehensive evaluation of the individual patient, his or her history and symptoms, and the course of the disease; (ii) perform tests of pulmonary function; and (iii) establish a treatment regimen based both on the information obtained and the needs of the individual patient. For example, the realization that inflammation plays a significant role in the pathogenesis of asthma has led to the use of different medications to treat the disorder and prevent asthma exacerbations (Creer & Levstek, 1997).

Evaluation for the asthma patient consists of a complete medical and environmental history that includes exploration of the potential triggers for a patient's asthma attacks. Most asthma experts believe the patient's history is a significant part of the evaluation (Ellis, 1993; National Asthma Education Program, 1991). Pulmonary tests can aid in differentiating asthma from other respiratory disease. These include tests of pulmonary function, such as spirometry or peak expiratory flow rate (PEFR). One especially useful technique is to measure forced expiratory volume (FEV-1). This is the total amount of air a patient can exhale in one second. Pulmonary function data may help differentiate restrictive from obstructive airways. If given a bronchodilator, patients with asthma will usually exhibit complete reversibility of the obstruction, whereas lung function may remain impaired for patients with irreversible damage. In addition, measurements of pulmonary function are useful in determining the severity of the airflow difficulties; although they should not be ignored, subjective patient perceptions of their symptoms usually correlate poorly with the more objective pulmonary function tests (Spector & Nicklas, 1995). Patients may be presented with a standard form of bronchoprovocation challenge, such as inhaling methacholine, inhaling cold air, or exercising. This process helps the clinician diagnose and understand the specificity of the patient's symptoms. At one time, a 15–20% decrease in the pulmonary test data, when challenged, often resulted in a diagnosis of asthma. Provocation challenges are still used to determine precipitants of a patient's asthma precipitants, although in most cases, the actual diagnosis of asthma is based upon a 15–20% improvement in pulmonary function in response to a bronchoinhaler.

In recent years, peak flow rates, obtained with a simple, portable, and inexpensive peak flow meter, have been widely adapted in the management of asthma. When done correctly, PEFR values correlate highly with those obtained with the spirometer. PEFR values are useful in assessing the degree and severity of airway obstruction, for monitoring response to treatment, for diagnosing exercise-induced asthma, and for detecting asymptomatic deterioration. When patients learn to take PEFR values at home, the clinician's ability to provide effective treatment is enhanced (National Asthma Education Program, 1991).

More sophisticated testing is required for patients suspected of having COPD to determine the pathology and disease state. A complete history of the patient, particularly regarding his or her smoking history, is essential. A physical examination is helpful in revealing signs of dyspnea and airway hyperinflation; chest radiography provides the most direct evidence for COPD, because of the known anatomical changes that occur in the disease. Pulmonary function tests are also useful in determining the rate of decline in pulmonary function for patients with COPD. If chronic bronchitis is indicated, sputum examination is required (American Thoracic Society, 1995).

8.14.4.2 Ratings of Illness Severity and Prognosis

Rating systems for severity of COPD are unique from those used to gauge asthma severity. Severity of COPD is based on the inter-relationship between the degree of airflow obstruction, the subjective sensation of breathlessness (dyspnea), and the impairment of gas exchange (especially oxygen levels). Most patients are in a category of stage I where there is minimal interference of COPD with daily activities; medical care is provided on an outpatient basis and is modest in nature. At stage II, impairment on daily life is more significant, and patients are usually required to seek the ongoing support of a respiratory therapist. Stage III includes a minority of patients; it is the late stage of COPD where quality of life is greatly diminished. Constant care by a respiratory therapist is required. At each stage of the disease process, treatment is guided by the needs of the individual patient, and includes consideration of medical, functional, and psychological factors (American Thoracic Society, 1995). As noted, attempts to classify asthma severity have generally involved a formula based on the interaction between symptom frequency and intensity, activity impairment, pulmonary function, and response to treatment (International Consensus Report, 1992; National Asthma Education Program, 1991; Spector & Nicklas, 1995

8.14.4.3 Medications used to Treat COPD and Asthma

Despite the differences in diagnosis of asthma and COPD, many of the same treatment options are available to individuals with both disorders.

8.14.4.3.1 Maintenance or preventive medications

Recognition that inflammation was a major component in asthma led to the use of anti-inflammatory medications, particularly inhaled corticosteroids and cromolyn sodium, with an increasing number of patients. These drugs are taken on a prophylactic basis to help prevent asthma. The basis of their use is analogous to individuals who are susceptible to sun damage.

In this example, individuals can prevent the inflammation of a sunburn by applying a sun blocker. Similarly, asthma patients use the anti-inflammatory agents to protect against reactions to allergens and other irritants.

Though the exact mechanism of action is not fully understood, cromolyn sodium is effective with selected patients. It is also an excellent medication because it produces few side effects. Inhaled corticosteroids are effective anti-inflammatory agents, and have been widely prescribed in the treatment of asthma. Oral doses of corticosteroids, particularly when taken over an extended period of time, have potentially damaging side effects such as skin damage, cataracts, diabetes, osteoporosis, and infection.

The efficacy of anti-inflammatory drugs in COPD is unclear. Cromolyn sodium could only be of value to COPD patients who have associated respiratory allergies. There is no evidence that inhaled steroids have a role in the management of COPD. Side effects of oral steroids, along with a discouraging 20–30% positive response rate in COPD patients, make steroids an unlikely intervention choice for COPD treatment (American Thoracic Society, 1995).

8.14.4.3.2 Bronchodilators

Bronchodilators are used in the treatment of both asthma and COPD. In asthma therapy, bronchodilators are used to treat exacerbations of asthma because they dilate constricted bronchial airways. The effect is usually more pronounced in asthma patients than in patients with COPD because the latter are sometimes less tolerant of medication side effects such as tremor and nervousness (American Thoracic Society, 1995). β-Adrenergic agonists, anticholinergic agents, and methylxanthines are the three primary types of bronchodilators used to treat asthma and COPD; each is briefly discussed.

β-Adrenergic agonists relax the airway smooth muscle and may contribute to the biochemical process of inflammation reduction. Inhaled β-adrenergic agonists are the preferred medication for alleviating acute asthma exacerbations and for preventing exercise-induced asthma (Spector & Nicklas, 1995). For patients with COPD, β-adrenergic agonists do not alter the disease process, although they can alleviate some symptoms of the disorder. There is a concern that β-agonists are misused by some patients with asthma (Spector & Nicklas, 1995), and are overused by COPD patients to relieve dyspnea (American Thoracic Society, 1995).

Anticholinergic agents are not recommended as the sole treatment for asthma. However, their use may be indicated if taken with a β-agonist in asthma patients with chronic airflow obstruction (Spector & Nicklas, 1995). There is evidence that administration of an anticholinergic agent may be more effective than a β-agonist to relieve daily symptoms of COPD. The drugs have fewer side effects and are longer acting than β-agonists; they may also be used in combination with β-agonists to treat the exacerbation of symptoms (American Thoracic Society, 1995). Methylxanthines, particularly theophylline products, are useful in the treatment of mild, moderate, and, in particular, nocturnal asthma. They serve as a mild-to-moderate bronchodilator.

Methylxanthines are useful in the treatment of COPD where inhaled medication is contraindicated (American Thoracic Society, 1995). Although its bronchodilation and long-action effects are attractive, potential side effects associated with theophylline has often led to it not being prescribed by physicians (Creer & Levstek, 1997). However, harmful side effects can generally be avoided by appropriate dosing and monitoring (National Asthma Education Program, 1991).

8.14.4.3.3 Oxygen therapy

Oxygen therapy is rarely indicated for patients with asthma, except in the case of severe intractable asthma where the goal is to increase blood oxygen levels to at least 90% arterial saturation. It is likely that these patients will be treated in an intensive care setting where the risk for respiratory failure can be adequately addressed (Spector & Nicklas, 1995). For COPD patients, hypoxemia can be a common problem; one of the goals of therapy is to correct and prevent hypoxemia. Long-term oxygen therapy (LTOT) provides a variety of pulmonary, cardiac, quality of life, and functional benefits for COPD patients in later stages of disease.

8.14.5 PSYCHOLOGICAL ASPECTS OF COPD AND ASTHMA

Respiratory disorders have been perceived in various ways by professional groups who work with the condition. If one were to look at how asthma was perceived in the 1940s, one would discover that many thought it was a significant psychological disorder that centered around unresolved conflicts that patients, particularly children, had with their mothers. This view represented the last remnants of the psychoanalytic view of asthma espoused by French and Alexander (1941). In the 1990s, a wide spectrum of views concerning asthma are found; it is possible, for example, to compile a set of articles which suggest that patients with asthma are psychologically different from some comparative group. However, if one were to divide and stack articles into whether asthma patients are abnormal or normal, there would be far more articles in the latter pile. This does not mean that patients with asthma or COPD do not exhibit psychological or behavioral problems; what it illustrates is that it is inappropriate to classify all patients with asthma or COPD into a specific psychological category.

With this as an introduction, how should the psychological aspects of COPD and asthma be perceived? A prevailing view, described by a number of authors (Creer, Renne, & Christian, 1976; Desquin, Holt, & McCarthy, 1994; Russo & Varni, 1982), is that patients with chronic illness, including COPD and asthma, are healthy people with a unique set of problems. In short, the assumption is made that asthma and COPD patients are normal people in an abnormal situation. Patients are influenced, like other people, by a variety of social, environmental, developmental, and cognitive variables. One strength of this view is that it allows health and clinical psychologists to focus on behaviors they can change to permit COPD and asthma patients to function in a healthier manner.

8.14.5.1 Impact on the Individual

Any chronic illness can have an impact on the individual patient, his or her family, and society. The degree of influence of asthma and COPD varies from person to person and, in many cases, from situation to situation. It is often impossible to make general conclusions about the impact a respiratory disorder has on a given patient. You may work with a patient with severe COPD who maintains a positive outlook. Through a combination of medical and behavioral expertise, she appears to live a relatively normal life. She may take her oxygen tank with her, but she would not let her COPD disrupt her Wednesday bridge group or her Friday appointment with a hairdresser. In another instance, you may be asked to assess and consult on a case involving a boy with mild asthma, and discover that his condition is at the center of both his and his family's lives. The youngster may be physically in poor shape, yet avoid taking physical education courses because exercise once produced an attack.

As in any type of case referred to a clinical or health psychologist, professionals must consider how the patient perceives his or her disorder to determine when and how to intervene. The clinical emphasis should be, first, on assessing the individual patient to determine the specific types of problems he or she experiences. Many problems may be related to the COPD or asthma of the patient. The clinician should next determine how best to help the patient resolve any detected problems. In recent years, the aim of medical treatment, particularly for asthma, has been to tailor treatment regimens for the individual patient (National Asthma Education Program, 1991). The trend of individualizing or tailoring treatments has been expanded to all chronic respiratory disorders. A similar approach has long been taken in the psychological and behavioral treatment of asthma and COPD patients (Creer et al., 1976; Creer & Christian, 1976; Creer, 1979).

8.14.5.1.1 Emotional impact

Being unable to breathe has as much of an emotional impact as anything we are likely to experience. The impact is amplified when patients recognize that they don't know how to reverse their condition and breathe in a normal manner. Would we not expect that patients with asthma or COPD experience emotional distress under these circumstances? Would we not anticipate a patient's response to be contagious to those around him or her, including medical personnel? Imagine yourself a parent who suddenly realizes that your child is no longer breathing, yet the youngster looks to you for help. The emotional impact of both a child's condition and a parent's feelings of helplessness can be overwhelming. The fact that so many patients and members of their families remain calm when the patients cannot breathe—particularly under the conditions of uncertainty that earmark the intermittency and severity of breathing exacerbations—has always impressed those who work with asthma and COPD. Professionals, whether they be medical or behavioral personnel, often report that they are frightened and anxious when treating asthma and COPD patients. The awareness of their own feelings only strengthens the admiration clinicians have for the patients whose very existence is threatened. For this reason, teaching patients to remain calm is an important component for the management of COPD and asthma. Behavioral scientists possess techniques that are effective in reducing the fears and anxiety that can accompany COPD and asthma.

Depression is common in both COPD and asthma. Ries, Kaplan, Limberg, and Prewitt (1995) found that 29 out of 119 patients (24%) with COPD had significant scores on the Centers for Epidemiological Depression Scale (CES-D). A higher rate of depression was found in a group of COPD patients by Agle, Baum, Chester, and Wendt (1973), who found that 17 out of a group of 23 COPD patients experienced depression. However, Agle and Baum (1977) noted that a psychiatric diagnosis of depression for COPD patients is difficult to arrive at because many of the complaints that signal depression may be confused with disease symptoms. Similar comments were noted by Dudley and Sitzman (1988). They described four categories of depressive behavior in COPD patients: (i) physical symptoms; (ii) emotional reactions, such as feelings of hopelessness and helplessness; (iii) cognitive factors, such as indecisiveness and low self-esteem; and (iv) motivational factors, such as passivity and dependence.

It is also difficult to arrive at a psychiatric diagnosis of depression in asthma, although self-reports of the condition are common. In a study by Creer et al. (1988), for example, 37% of children with asthma and 32% of their parents, respectively, reported that the children were depressed because of asthma. Were the children depressed? Probably. Was their depression indicative of clinical depression? Probably not. When standardized depression inventories are used with asthma patients, higher scores can be obtained yet fail to be clinically significant. For example, elevated scores on the Beck Depression Inventory (Beck, 1967) were obtained in a study of 76 adults with asthma (Kotses et al., 1995). However, the average depression scores were within the normal range of values, indicting that most patients were not clinically depressed. This finding does not mean that depression should not be addressed by a clinical psychologist. What it does suggest is that both asthma and COPD patients may report depression that, depending upon the evaluation of the individual patient, may or may not require psychological intervention.

Panic–fear is another psychological dimension that has been investigated, particularly in patients with asthma. Two lines of research have been followed. First, Kinsman, Luparello, O'Banion, and Spector (1973) analyzed the subjective symptmatology of asthma within a group of 100 asthma patients. A set of 77 symptom adjectives were found to be associated with asthma attacks. Key cluster analysis of the 77 adjectives identified five symptom clusters: (i) panic–fear; (ii) irritability; (iii) hyperventilation–hypocapnia, for example, low levels of carbon dioxide in the blood; (iv) bronchoconstriction; and (v) fatigue. Panic–fear was reported by 42% of the sample investigated. A series of studies were conducted on the construct of fear–panic and asthma. Among findings were that the panic–fear influenced such variables as medication taking, length of hospitalization, symptom vigilance or disregard, medication compliance, and medical decision making (Dahlem, Kinsman, & Horton, 1979; Jones, Kinsman, Dirks, & Dahlem, 1979; Kinsman, Dirks, & Schraa, 1981; Staudenmayer, Kinsman, Dirks, Spector, & Wangaard, 1979).

A second line of research attempted to operationally define panic in behavioral terms (Creer, 1974, 1979). This proved more difficult than anticipated because the label "panic" was attached to a broad array of behaviors. At one end of the continuum, there were children who constantly complained that they required more of a prescribed drug or that the medication they were taking was ineffective. They appeared frightened and agitated, a behavior that was contagious to medical personnel. At the other end of the continuum, there were patients who remained frozen-like and silent and did not ask for help even if they lapsed into unconsciousness. After several attempts to define the behavior (Creer, 1992), it was decided to define panic as any behavior exhibited by children that frightened physicians, nurses, and other medical personnel. Defining panic in this manner led to the children being referred for psychological assessment and successful treatment via systematic desensitization (Creer, 1974, 1979).

Subsequent work has led to the inclusion of relaxation exercises, sometimes accompanied by systematic desensitization, in many treatment programs for asthma. Creer (1991) asserted that the use of relaxation with systematic desensitization to change panic behavior represented the single most effective application of behavioral procedures to manage asthma-related behaviors. In addition, relaxation during attacks can prevent the exacerbation of an ongoing episode and, perhaps more importantly, assist patients to think of and perform steps required to control their asthma.

8.14.5.1.2 Cognitive impact

Enough data was collected concerning the cognitive impact of COPD to warrant Grant, Heaton, McSweeny, Adams, and Timms (1980) to conclude the disease affects the brain as well as the lungs and heart. Prigatano and Grant (1988) suggested that COPD patients demonstrate signs of neuropsychological impairment, particularly including abstract reasoning, speed of information processing, and general psychomotor speed and accuracy. Their findings led Prigatano and Grant (1988) to suggest future research should: (i) investigate parameters that influence psychomotor deterioration and delay in speed of information processing; (ii) clearly describe the underlying nature of the neuropsychological deficits and the related medical parameters; (ii) understand changes in neuropsychological function that occur over time in COPD patients; (iv) determine the extent to which neuropsychological disturbance compromises the patient's ability to comply with rehabilitation; and (v) determine how neuropsychological deficits influence the quality of life in COPD patients. Attention has been focused on the impact of COPD on quality of life as assessed through such instruments as the Quality of Well-Being Scale and the Sickness Impact Profile (McSweeny, 1988; McSweeny & Creer, 1995; Prigatano & Grant, 1988).

A major concern in asthma has been on the effect asthma medications have on cognitive functions. A number of studies have been conducted on this subject by Bender and his colleagues, particularly on children with asthma. In a recent summary of the research (Creer & Bender, 1995), it was noted that the results of the studies are inconsistent. Asthma medications have been purported to impede memory or to improve it; to cause motor tremor; to cause hyperactivity in children; to improve attention; to cause anxiety; and to induce psychosis. Many reports of cognitive and neuropsychological side effects of asthma medications are prone to exaggeration, but side effects do exist that need to be considered in making treatment decisions regarding asthma, particularly in children. These involve all the major categories of asthma medications including corticosteroids, both oral and inhaled; theophylline; β-adrenergic agonists, and over-the-counter medications (Bender & Klinnert, in press; Creer & Bender, 1993; 1995).

8.14.5.1.3 Behavioral impact

Part of the behavioral impact of COPD is embedded in the morbidity and mortality data discussed earlier. Kaplan and Atkins (1988) succinctly summarized the four reasons why COPD has such a major impact: (i) it is among the top four leading causes of death; (ii) it has a major impact upon the activities of daily living; (iii) it affects a large number of people; and (iv) the occurrence of COPD is increasing. Many individuals who started smoking 30 or 40 years ago are only now beginning to experience COPD. The disorder, in turn, affects other psychological and behavioral patterns, such as anxiety and depression, cognitive and psychomotor functions, psychosocial activities, and sexual behavior.

The behavioral impact of asthma was also discussed within the context of epidemiological information. Patients with asthma miss more work and school than those who do not have the disorder. Asthma has long been the leading cause of school absenteeism among chronic health conditions (Schiffer & Hunt, 1963); this finding has not changed in the subsequent decades. Consequently, many children with asthma face the struggle of striving to keep up with their peers academically (Creer et al., 1976). Missing school also isolated these youngsters from the mainstream of social activities and interactions that help them develop into productive adults. Missing work is equally deleterious to adults with asthma. They are not only fearful that they may lose their jobs, but that co-workers believe they are passing their duties onto others. Many adults with asthma worry about taking medications at work, even if needed to abort an attack, because they believe their action could jeopardize their job, and their chances for advancement.

Finally, although it is bad enough to experience asthma, being labeled an asthmatic conjures up various negative images to different people. Educators may automatically assume that a child with asthma requires special educational treatment, such as skipping physical education classes, even though a child with asthma needs to be in good physical condition in the event he or she has a serious attack (Creer, 1979).

8.14.5.2 Impact on Family and Social Environment

COPD and asthma are chronic disorders that have the potential to dissolve the fabric of family unity and cohesiveness (Creer, 1979). There are a several reasons for this. First, there are the expenses incurred by the disorders. Vance and Taylor (1971) found that childhood asthma could consume anywhere from 2% to 30% of a family's income; similar data were reported a decade later by Marion, Creer, and Reynolds (1985). Not only may more intensive medical treatment be required to treat COPD, but the family could face the prospect of lost income because the patient with COPD is unable to work. Second, family activities are curtailed because of COPD or asthma. Vacations are a rarity not only because of their expense, but because, in the case of asthma, families are fearful of exposing a member to stimuli that could induce an attack; with COPD, patients and their families are often fearful of being apart from medical treatment, such as oxygen, if needed. With a progression in the severity of COPD, both the patient and members of their family begin to see themselves as prisoners with COPD as the jailer. Finally, as family tensions increase, there is the potential for families to break up and disintegrate. Creer, Ipacs, and Creer (1983) found that over a 10-year period, the divorce rate among parents of children admitted to a residential treatment facility for asthma increased from 15% to 61%. The rates of divorce among these families were far higher than found in the overall population at that time.

Social support is a strong predictor of prognosis, even more so than other predictors such as severity of illness (Jensen, 1983). This finding is applicable to patients with both COPD and asthma.

8.14.6 ADHERENCE TO TREATMENT REGIMENS

Adherence or compliance to treatment regimens is poor in most chronic conditions. Patients with COPD and asthma are no exceptions. With COPD, nonadherence occurs both in smoking cessation and in taking medications. Smoking cessation is difficult to attain, particularly in smokers who have smoked for 20 or more years. Adherence by patients to asthma regimens is also difficult. Compliance is especially low when both preventative and as-needed medications are required to help control asthma. This reflects the finding that more complex regimens lead to lowered rates of compliance. Equally problematic is having asthma patients avoid stimuli that could trigger or exacerbate their conditions. Many health care personnel, including psychologists, have failed when asking patients or families to rid their homes of a pet whose dander is provoking a patient's asthma.

8.14.6.1 Medication Compliance

Medication compliance has been studied extensively in asthma, particularly in children with the disorder. There are three themes to the research (Creer, 1993). First, interest in compliance was sparked by the widespread prescription of theophylline medications for asthma. These metaxanthine products are effective in the treatment of asthma, but only when drug levels fall within a therapeutic band in the blood. To determine safe doses of theophylline, a technology emerged to assess the drug in the blood stream.

A second theme of studies of compliance in asthma was the focus on methods of assessment. Studies investigated the use of direct measures—including blood serum assays, the use of tracers or markers, and observation—and indirect measures—including patient self-reports, pill and liquid medication assessment, treatment outcomes, and mechanical instruments for recording medication use.

A final theme was pointed out by Creer and Levstek (1996). They noted that there were two competing sets of research literature in asthma. On the one hand, there were the investigations showing the wide disparity of medication compliance rates among patients. For this reason, the tendency has been to blame patients for not adhering to their medication regimen. On the other hand, studies have indicated that not only were there many physicians who did not know how to use medication inhalers correctly, but that even more medical personnel did not teach proper inhalation techniques to their asthma patients. Based upon these disparate sets of data, Creer and Levstek (1996) concluded that it was time to share the blame for noncompliance in asthma between medical personnel who do not teach their patients to take their medications correctly and patients who do not adhere to their treatment regimens.

Compliance has not been directly studied in COPD patients (Windsor, Green, & Roseman, 1980). However, Parker (1988) surmized that COPD patients would be apt not to comply with a treatment regimen for several reasons: (i) COPD is a chronic disease, and adherence to medication regimens is not only low but tends to decrease over time; (ii) COPD patients are generally older, and may have several other ailments that require medical treatment; (iii) cognitive impairment in COPD patients may lead to forgetfulness; and (iv) as the dyspnea experienced by patients with COPD is a constant symptom, there are suggestions that increased symptom frequency may be associated with noncompliance. However, Parker (1988) concluded that the severity of a patient's disability may influence compliance in a positive direction and provide incentive for patients to comply with medical instructions. Two other factors may also contribute to noncompliance in COPD patients. First, compliance rates are low where major behavioral changes, such as quitting smoking and increasing exercise, are required. Second, Kopacz (1988) pointed out that compliance is apt to be more of an issue as the focus of long-term treatment of COPD shifts from hospital- to community-based care where patients are expected to perform self-management skills to help manage their disorder.

8.14.6.2 Lifestyle Changes

Life activities in asthma are often a function of severity of the disorder. Those with severe asthma, particularly those who experience asthma exacerbations on a perennial basis, may have more activity limitations than patients who have mild asthma on an intermittent basis. However, asthma is a disorder that, with proper treatment, can be controlled in almost all patients (National Asthma Education Program, 1991). For this reason, the degree of control established over asthma becomes of major significance. If asthma is treated on a regular basis by skilled medical personnel, there is no reason why asthma should dictate the lifestyle of patients.

There has been a strong effort to assess quality of life in patients with COPD (McSweeney, 1984; McSweeney & Creer, 1995). Four dimensions of life quality for COPD patients were suggested by McSweeny, Grant, Heaton, Adams, and Timms (1982). They define quality of life in terms of: (i) emotional functioning, including mood changes and other psychological symptoms; (ii) social role functioning, including employment, home management, and family or social relationships; (iii) daily living, such as activities of self-care skills and mobility; and (iv) the ability to engage in enjoyable activities, such as hobbies and recreational activities.

8.14.6.2.1 Activity restriction

Newacheck, Budetti, and Halfron (1986) reported that data from the National Health Interview Survey showed that over two million children under the age of 17 years experienced some degree of limitation in their school, play, or other recreational activities because of chronic conditions. Looking at the causes of the increase in activity limitation, Newacheck and his co-workers found that asthma counted for most of the net change (65%) in the prevalence of activity-limiting disease. More recently, Taylor and Newacheck (1992) examined activity restrictions in children with asthma. In comparison to their peers without asthma, it was found that youngsters with asthma missed an additional 10.1 million days of school, had 12.9 million more contacts with physicians, and were hospitalized 200 000 more times in 1988. Depending upon the severity of the disorder, patients with COPD may experience problems ranging from curtailing their favorite activity because of a shortness of breath to being restricted to travel only when accompanied by an omnipresent oxygen tank. Beyond these observations, there is a paucity of data concerning activity restriction and COPD.

8.14.6.2.2 Environmental prescriptions

Environmental control is becoming of increasing importance in the management and control of asthma. A prescription for environmental control was provided by Spector and Nicklas (1995, p. 760):

> Important steps in environmental control are (1) to minimize house dust mite exposure in mite exposure in mite-allergic patients with asthma, (2) to lessen exposure to domestic animals in appropriate patients, (3) to not allow smoking in the home, (4) to avoid strong odors and chemical fumes, (5) to install kitchen and bathroom exhaust fans, (6) to use humidifiers with caution in mite- and mold-sensitive patients, (7) to use air conditioners or electrostatic air purifiers, (9) to install a dehumidifier and reduce water entry in damp basements, and (10) to initiate other measures as indicted for specific allergies as appropriate.

Recent efforts have been made to establish environmental control for people with asthma, but overall it has been difficult to establish and maintain such control in the homes of patients. As pointed out by Creer and Levstek (1996), environmental management of asthma offers strong research and clinical opportunities for behavioral scientists.

Environmental risk factors for COPD include passive smoking, ambient air pollution, and occupational factors. As noted earlier, interactions between cigarette smoking and job exposure to hazardous airborne substances results in higher rates of COPD.

8.14.6.2.3 Smoking cessation

Most patients with asthma do not smoke. In fact, a common report by children involved in a self-management program was that they made every effort to avoid being around those who smoked, including other family members (Creer et al., 1988). Getting parents or other family members to quit smoking, however, is another story: most family members know their behavior harms the patient with asthma, yet they continue to smoke. As a clinician, it is a major victory to persuade family members to avoid smoking in the home.

A number of approaches have been taken to produce smoking cessation in COPD patients. One approach is the adaptation of the "four As" of smoking cessation outlined for physicians by the National Cancer Institute (Lewis & Fiore, 1995): (i) ask about smoking status at every opportunity; (ii) advise every smoker to quit smoking; (iii) assist your patient in stopping smoking, even including using nicotine patches or gum; and (iv) arrange follow-up visits. Others (Muzurek & Gross, 1995) suggest a three-stage system of smoking cessation that includes preparation, intervention, and maintenance. Other specific steps for smoking cessation, compiled from both medical and behavioral scientists, were suggested by the American Thoracic Society (1995).

8.14.7 PSYCHOLOGICAL INTERVENTIONS FOR COPD AND ASTHMA

Separate intervention approaches are taken towards COPD and asthma. The approach with COPD has been to introduce pulmonary rehabilitation to improve the quality of life of patients and their families. A number of behavioral techniques are involved in a pulmonary rehabilitation program. The approach with asthma is broader, in part because of the wide array of problems presented by patients with asthma.

8.14.7.1 Pulmonary Rehabilitation for COPD

Ries (1995) outlined a guide for the pulmonary rehabilitation of patients with moderate to severe disease. Rather than focusing solely on reversing the disease process, the goal of pulmonary rehabilitation is to minimize disability from the disease. The goal could be achieved by enhancing medical therapy, controlling and alleviating symptoms, improving pulmonary functions, and enhancing the quality of life of patients. There are three foci of a pulmonary rehabilitation program: (i) the patients; (ii) a multidisciplinarian approach; and (ii) attention to physiology and psychopathology. The selection of patients was based upon a study of rehabilitation programs conducted by Ries, Kaplan, Limberg, and Prewitt (1995). Six criteria for inclusion were used: (i) patients had symptomatic lung disease; (ii) patients were stable with standard therapy; (iii) patients had functional limitation because of chronic lung disease; (iv) patients were under the care of a primary care provider; (v) patients had no other interfering or unstable medical conditions; and (vi) patients were motivated to be involved in and responsible for their own healthcare. Components of the comprehensive pulmonary rehabilitation program required a multidisciplinarian team. There were two major classes of components of the program: patient evaluation and program content. Patient evaluation included the following five stages:

Stage 1: Interview. A screening interview was conducted to assess suitability for a pulmonary rehabilitation program. An additional aim of the interview was to establish teamwork between patients and their rehabilitation team.
Stage 2: Medical evaluation. Reviewing the medical history of patients, including laboratory data, helped confirm their disease and its severity.
Stage 3: Diagnostic testing. Pulmonary function evaluation, exercise testing, and the assessment of blood gases at rest and during exercise were performed during this stage.
Stage 4: Psychological assessment. Psychological, neuropsychological, social, or cognitive problems were detected during this stage. Other information concerning reactions to progressive dyspnea, sexual dysfunction, and social support was obtained.
Stage 5: Goals. After evaluating medical, physiological, and psychological data, realistic goals that were compatible with the patient's disease, needs, and expectations were established.
The second component of a pulmonary rehabilitation program consisted of four areas of program content:

(i) *General education*. Patients were provided education about their disorder. The general philosophy was to teach patients skills to permit them to become allies with their physicians in the management of the patient's respiratory condition.

(ii) *Respiratory and chest physiotherapy instruction*. Each patient's need for respiratory care techniques was assessed and instruction provided on various topics, including: (a) bronchial hygiene; (b) breathing retraining techniques, including pursed lips breathing; (c) proper use of respiratory equipment; and (d) appropriate use of oxygen.

(iii) *Exercise prescription*. Developing a standard program of exercise for each patient was another content area. The aim was to prescribe a specific regimen, consisting of simple, useful, and inexpensive exercises, to improve a patient's exercise tolerance and endurance.

(iv) *Psychosocial support*. The final content area consisted of teaching patients to combat feelings of hopelessness and the inability to cope with their COPD. Significant others of patients were involved in this component of treatment.

Significant changes in the condition of COPD patients were reported in several studies, including those conducted by Ries et al. (1995) and Niederman et al. (1991). Ries (1995) suggested that the following findings could be expected as a result of pulmonary rehabilitation of COPD patients: (i) decreases in respiratory and psychological symptoms, and the patient's use of medical resources; (ii) increases in quality of life, physical activity, exercise tolerance, activities of daily living, knowledge about COPD, and independence; and (iii) changes in lung function, possible return to work, and longer survival. As will be noted, these changes represent a gold standard for pulmonary rehabilitation because no single study has produced all of the findings.

8.14.7.2 Asthma Management

In tracing psychological approaches to asthma, three stages can be delineated.

8.14.7.2.1 Stage 1: Psychotherapy, particularly psychoanalysis

It was pointed out earlier that in the 1940s most psychiatric and psychological treatment was guided by the work of French and Alexander (1941). As the influence of psychoanalysis declined, the relationship of resolving underlying presumed psychological factors to help manage asthma also faded (Renne & Creer, 1985). The focus shifted towards intervention with treatable problems rather than a search for psychological underpinnings to what is a physical problem: asthma.

8.14.7.2.2 Stage 2: Behavioral approaches

The rise of behavioral therapy approaches in the 1960s and 1970s shifted the focus of asthma management to resolving asthma-related psychological and behavioral problems. The theme is ongoing as health psychologists investigate, for example, behavioral and psychological aspects of asthma medication (Bender & Milgrom, 1992), or fears generated by the disorder (Creer & Christian, 1976). In the meantime, health psychologists have been successful in developing approaches for treating panic, correct inhaler use, overuse of hospital services, phobic behaviors, and compliance behaviors (Creer, 1979, 1992). Many of the techniques developed by health psychologists, such as using negotiation and contracting, were introduced to effectively manage medication compliance in children with asthma in the 1970s. An example of the use of negotiation and contracting was summarized by Creer (1979):

> Over two decades ago, the staff at the National Asthma Center, a residential treatment facility for pediatric asthma, found that more and more children returned to single-parent homes when discharged. As parents often worked, the children were required to treat some attacks by themselves. Hence, it was decided by behavioral and medical scientists to teach children to take their own medications as directed. The procedure involved three steps. First, negotiation took place between psychologists, physicians, and children. Clinical psychologists mediated the discussions and wrote up the negotiated contract. The typical contract, signed by a physician, psychologist, and their patient, specified that when a child demonstrated he or she was taking medications correctly, the youngster would be discharged from the facility. Second, once given medications, psychologists established a procedure whereby children were closely monitored for a period of time to insure that they were taking their medications correctly. Monitoring was gradually faded from the procedure as the children showed they were properly taking their medications. Finally, upon demonstrating that he or she was taking his or her medications as prescribed over a period of time and in different settings, the child, his or her physician, and the youngster's psychologist set the discharge date for the child to return home. Monthly reports from the children and their parents indicated that the procedure was extremely effective in helping the children control their asthma.

These procedures are still state-of-the-art with respect to the management of pediatric asthma. This does not mean that the problem of medication compliance has been solved; however, it remains a major barrier to the proper treatment of both COPD and asthma.

8.14.7.2.3 Stage 3: Self-management

Processes involved in asthma self-management are summarized according to the model of self-management proposed by Creer (1997) and Creer and Holroyd (1997). It consists of the following processes.

(i) Goal setting

Setting goals involves patients reaching an accord with their physicians and other medical personnel how best to control a patient's asthma. Typical goals jointly established by physicians and patients include patient's taking prescribed medications as scheduled and avoiding known triggers of their asthma.

(ii) Information collection

Self-monitoring—the observation and recording of data on oneself—is the basis of information collection. In most studies of asthma self-management, a diary is employed. A typical asthma diary would ask patients to list the symptoms they experienced daily, their highest peak flow readings obtained in the morning and evening, and the medications they took for their asthma. Asthma diaries also serve to prompt patients to monitor their asthma. So as not to overwhelm patients in monitoring their behavior, Creer and Bender (1993) cautioned that target behaviors should be operationally defined, and that patients collect data for only specified periods of time.

(iii) Information processing and evaluation

Patients must learn to process and evaluate the information they collect about themselves and their asthma. Creer and Holroyd (1997) delineated five steps involved in information processing and evaluation: (a) patients must be able to detect changes that occur in the information they observe, record, and process about themselves; (b) a set of objective criteria must be established, such as their personal best peak flow reading, against which patients can evaluate any changes that have occurred in their breathing; (c) patients must evaluate and make judgments about the data they process; (d) patients must evaluate any changes that occur in their breathing in terms of the antecedents that may have produced the change, the behaviors they need to perform to alter any changes, and the consequences of their actions; and (e) asthma patients need to consider the contextual variables in processing and evaluating information about themselves.

(iv) Decision making

Making the correct decision is essential for successful self-management. A study by Creer (1990) analyzed decision making in gold standard physicians and patients. It was determined that both groups used decision-making strategies similar to those describe by Arkes (1981) in that they: (a) avoided preconceived notions; (b) were thoughtful and cautious; (c) generated a number of management alternatives; (d) did not misperceive the severity of the situation; (e) referred to a personal database; (f) thought in terms of probabilities; and (g) did not rely on memory.

(v) Action

Action involves the performance of self-management skills to help control asthma. A number of components have been tested in self-management programs for asthma, but self-instruction, including self-statements, is a common thread of these programs.

(vi) Self-reaction

Self-reaction refers to the attention patients direct towards evaluating their performance (Bandura, 1986). On the basis of self-reaction, patients can establish realistic expectations about whether they need more training and expertise. Self-efficacy, the belief that one can adequately perform specific skills in a given situation (Bandura, 1977), influences the performance of self-management skills. The development of self-efficacy scales have permitted this construct to be evaluated in asthma (Tobin, Wigal, Winder, Holroyd, & Creer, 1986).

A number of self-management programs have been developed and evaluated for both pediatric and adult asthma. Common outcomes of these programs include a decrease in related hospitalizations, ER visits, school and work absenteeism, asthma medication use, depression, and costs for asthma. At the same time, there have been reported increases in knowledge about asthma, performance of self-management skills, medication compliance, self-efficacy, and quality of life. The positive results have led expert panels on asthma (e.g., National Asthma Education Program, 1991) to recommend such training for all patients with the disorder.

8.14.8 IMPLICATIONS FOR HEALTH AND CLINICAL PSYCHOLOGY

There has never been such a need for health and clinical psychology in the management of COPD and asthma than at the present time. The increasing prevalence of people with the disorders, as well as morbidity and mortality markers, reinforce the need for active psychological and behavioral input. Yet, despite an increasing number of signals for behavioral change in patients with asthma and COPD, the urgent need for psychological and behavioral interventions has gone unheeded by both medical or behavioral scientists. The need is illustrated by suggesting ways that health and clinical psychologists could become involved with COPD and asthma.

8.14.8.1 COPD

A review of the literature suggests a myriad of ways that behavioral and psychological expertise could be applied to COPD. Four areas merit a brief comment: (i) smoking cessation; (ii) assessment; (iii) integration of behavioral skills; and (iv) extension of behavioral expertise in COPD.

8.14.8.1.1 Smoking cessation

It was estimated that three million people would die because of smoking in 1995; by 2025, the number of deaths from smoking is estimated to be 10 million (Peto, 1994). This is but one index of problems expected from cigarette smoking, a behavior that has hit a plateau in the USA (Centers for Disease Control, 1996b). Attempts by behavioral scientists to bring about smoking cessation have met with mixed results, particularly in the USA (Meltzer, 1994). A recent report from the Lung Health Study (Kanner, 1996) indicated that an extensive smoking cessation program could produce a 20% success rate in getting smokers at risk for COPD to give up smoking permanently.

8.14.8.1.2 Assessment

Most outcome variables used in the rehabilitation of COPD patients are medical. However, the most important outcome of rehabilitation is that patients live a longer, happier, and more productive live. Attempts to assess quality of life by the Quality of Well-Being Scale have been made in a number of studies of COPD patients by Kaplan and colleagues (Kaplan, Atkins, & Timms, 1984). The results have been mixed in that quality of life has been shown to improve in some studies with COPD but not in others (Ries et al., 1995). The latter finding suggests that more disorder-specific instruments, in addition to general instruments such as the Quality of Well-Being Scale, might be employed to determine the impact that COPD has on the quality of life experienced by COPD patients.

8.14.8.1.3 Integration of behavioral skills

In reviewing the literature by COPD, particularly the rehabilitation of patients with the disorder, it is invariably noted that success is a team effort. The skills of a number of medical and behavioral specialists were used to assist patients to live a more normal life. A synthesis of skills could provide the fuel for pulmonary rehabilitation programs for COPD. However, although each specialist brought their particular expertise to the rehabilitation process, there was no evidence that there had been any integration of skills by an interdisciplinary team to rehabilitate patients.

8.14.8.1.4 Extension of behavioral expertise

Behavioral techniques can be extended in new directions, and to help solve new problems in COPD. For example, there have been no systematic attempts to develop and evaluate self-management programs for COPD. This is unfortunate given the success of self-management with asthma. Second, new medical procedures are being developed and tested with COPD. An illustration is the use of lung reduction surgery to improve the breathing of COPD patients (Rogers, Sciurba, & Keenan, 1996). Use of these procedures will generate behavioral problems that will need to be addressed.

8.14.8.2 Asthma

Researchers and clinicians who have worked with asthma for a period often discuss the increasing complexity of the disorder and how it might be controlled. There are several areas that are deserving of behavioral research: (i) medication compliance; (ii) avoidance of attack triggers; (iii) environmental control; (iv) refinement of self-management processes; (v) maintenance of self-management performance; and (vi) introduction of new experimental designs.

8.14.8.2.1 Medication compliance

As long as compliance rates in asthma patients remain, on average, below 50%, there will be the need for more medication compliance investigations. Medication compliance has become even more of an issue with patients who take both preventive and as-needed medications; these patients must remember to take both medications, but on different schedules. It would be easy for a health psychologist to spend an entire career investigating medication compliance in asthma.

8.14.8.2.2 Avoidance of attack triggers

In many instances, patients can avoid exacerbations of their asthma by avoiding or escaping from stimuli known to trigger their asthma. When such triggers or precipitants for a patient are known, it is important to teach the patient to avoid them. A behavioral scientist should be involved in such training.

8.14.8.2.3 Environmental control

A major thrust in the control of asthma is to control environmental stimuli known to precipitate asthma (Spector & Nicklas, 1995). In most cases, attempts to control the environment require application of behavioral modification techniques. Thus far, limited success has been attained in teaching patients to alter their environment in order to manage their asthma (Creer, 1998). If there is a group of scientists who are likely to modify the patient's environments, it is clinical or health psychologists. The challenge of achieving such a goal should excite the most radical behaviorist!

8.14.8.2.4 Refinement of self-management processes

Despite the development and evaluation of a number of programs for both pediatric and adult asthma we do not know exactly how patients perform self-management skills to control their asthma. There is need for more research into the processes of goal setting, information collection and processing, evaluation of data, decision making, action, and self-reaction.

8.14.8.2.5 Maintenance of self-management skills

Although asthma self-management has reached the stage of refinement in development and evaluation, there is need to determine how to teach patients to perform self-management skills as long as is necessary to control their asthma (Creer, Levstek, & Reynolds, in press). There are intriguing data which indicate the maintenance of asthma self-management skills, over time and across settings, is stronger than anticipated (Levstek, 1996). However, more investigation is needed on this topic.

8.14.8.2.6 Design of studies

Most studies on the self-management of asthma have used traditional clinical trials designs employed in determining the effects of medications. Patients are randomly assigned to a treatment group; the control group is given a placebo condition. At the end of a period of time, usually a year, outcomes are assessed. Creer, Levstek, and Reynolds (in press) argued that these are not appropriate designs to use in behavioral research, particularly in self-management. First, they pointed out that intervention is educational and behavioral in nature. Although those in the intervention group learn to manage their asthma, self-management education is withheld from the control group for a year or more. The approach is simply unethical. Second, self-management programs differ from traditional clinical trials in that one is not simply investigating acquisition of self-management skills—an investigator wants to look at a more important outcome, namely the performance of learned skills. This cannot be determined from a traditional clinical trial because acquisition processes cannot be separated from performance processes.

Creer, Levstek, and Reynolds (in press) concluded that designs developed by behavioral scientists must be used. With these designs, patients can be randomly divided into treatment and waiting-list control groups. At the end of the intervention, differences between the two groups in knowledge of asthma and self-management can be assessed. Thereafter, whereas the waiting-list control receives the educational component, the patients taught the self-management skills should serve as their own control. This permits investigators to carefully study the main question in using rehabilitation and self-management approaches: how do patients perform and maintain their performance of self-management skills across settings and over time?

8.14.9 CONCLUSIONS

We would like to conclude with three comments. First, there is a pronounced need for health and clinical psychologists to be involved in the management of COPD and asthma. The need for behavioral expertise will only increase in the future. To actively help establish control over the disorders, however, psychologists must make a concerted effort to sell their skills to medical scientists who work with these conditions. This has not happened. Psychologists who have offered their expertise with COPD and asthma have sometimes been rebuffed by physicians in the past, but this reaction is fast fading. If you can help a patient with COPD to quit smoking or a child with asthma to comply with a medication regimen, your assistance will be welcomed.

Second, it behoves any clinical or health psychologist to learn about the medical and biological aspects of asthma or COPD if they are to work with the disorders. Not only will they be able to apply their expertise in a more effective manner, but they will not inadvertently harm patients. The latter event could have occurred in the past, for example, had psychologists increased compliance to asthma treatment regimens that were potentially harmful (Creer & Levstek, 1996).

Finally, one of the authors has worked with asthma for three decades. He used to hear medical scientists, particularly immunologists, proclaim that someday they would discover a cure for asthma and put everyone working with the disorder out of work. The response by the psychologist was that the day couldn't come soon enough. However, the medical scientists were always reminded that once a cure was developed, it would be the psychologist who would be called on to teach patients to properly use and comply with the medical regimen. As there are no cures in sight for either COPD or asthma, this is a moot comment. The exchange puts into perspective, however, the role we as clinical and health psychologists can play with respect to asthma and COPD in the future.

8.14.10 REFERENCES

Agle, D. P., Baum, G. L., Chester, E. H., & Wendt, M. (1973). Mulidiscipline treatment of pulmonary insufficiency: I. Psychological aspects of rehabilitation. *Psychosomatic Medicine, 35,* 41–49

Agle, D. P., & Baum, G. L. (1977). Psychosocial aspects of chronic obstructive pulmonary disease. *Medical Clinics of North America, 61,* 749–758.

American Thoracic Society (1987). Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. *American Review of Respiratory Diseases, 136,* 225–244.

American Thoracic Society (1993). Guidelines for the evaluation of impairment/disability in patients with asthma. *American Review of Respiratory Diseases, 147,* 1056–1061.

American Thoracic Society (1995). Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease. *American Journal of Respiratory and Critical Care Medicine, 152,* S77–S120.

Arkes, H. R. (1981). Impediments to accurate clinical judgment and possible ways to minimize their impact. *Psychological Bulletin, 110,* 486–498.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review, 84,* 191–215.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York: Harper & Row.

Bender, B. G., & Klinnert, M. D. (in press). Psychological correlates of asthma severity and asthma treatment outcome in children. In H. Kotses & H. Harver (Eds.), *Behavioral contributions to the management of asthma*. New York: Marcel Dekker.

Bender, B. G., & Milgrom, H. (1992). Theophylline-induced behavior change in children: An objective evaluation of parents' perception. *Journal of the American Medical Association, 267,* 2621–2624.

Bernstein, I. L. (1988). Asthma in adults: Diagnosis and treatment. In E. Middleton, Jr., O. E. Reed, & E. F. Ellis (Eds.), *Allergy: Principles and practice* (4th ed., pp. 901–934). St. Louis, MO: Mosby.

Burrows, B. (1991). Epidemiologic evidence for different types of chronic airflow obstruction. *American Review of Respiratory Diseases, 143,* 1452–1455.

Centers for Disease Control (1992). Asthma—United States, 1980–1990. *Morbidity and Mortality Weekly Reports, 45,* 733–735.

Centers for Disease Control (1996a). Asthma mortality and hospitalization among children and young adults—United States, 1980–1993. *Morbidity and Mortality Weekly Reports, 45,* 350–353.

Centers for Disease Control (1996b). Cigarette smoking among adults—United States, 1994. *Morbidity and Mortality Weekly Reports, 45,* 588–590.

Coultas, D. B., Gong, H., Grad, R., Handler, A., McCurdy, S. A., Player, R., Rhoades, E. R., Samet, J. M., Thomas, A., & Westley, M. (1993). Respiratory diseases in minorities of the United States. *American Journal of Respiratory and Critical Care Medicine, 149,* S93–S131.

Creer, T. L. (1974). Biofeedback and asthma. *Advances in Asthma and Allergy, 1,* 6–11.

Creer, T. L. (1979). *Asthma therapy. A behavioral health care system for respiratory disorders*. New York: Springer.

Creer, T. L. (1982). Asthma. *Journal of Consulting and Clinical Psychology, 72,* 912–921.

Creer, T. L. (1990). Strategies for judgment and decision-making in the management of childhood asthma. *Pediatric Asthma, Allergy, and Immunology, 4,* 253–264.

Creer, T. L. (1991). The application of behavioral procedures to childhood asthma: Current and future perspectives. *Patient Education and Counseling, 17,* 9–22.

Creer, T. L. (1992). Psychological and behavioral assessment of childhood asthma. Part II. Behavioral approaches. *Pediatric Asthma, Allergy, and Immunology, 6,* 21–34.

Creer, T. L. (1993). Medication compliance and childhood asthma. In N. A. Krasnegor, L. Epstein, S. B. Johnson, & S. J. Yaffe (Eds.), *Developmental aspects of health compliance behavior* (pp. 303–333). Hillsdale, NJ: Erlbaum.

Creer, T. L. (1997). *Psychology of adjustment. An applied approach*. Upper Saddle River, NJ: Prentice-Hall.

Creer, T. L. (1998). Childhood asthma. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of child psychopathology* (3rd ed., pp. 395–415). New York: Plenum.

Creer, T. L., Backial, M., Burns, K. L., Leung, P., Marion, R. J., Miklich, D. R., Morrill, C., Taplin, P. S., & Ullman, S. (1988). Living with asthma: Part I. Genesis and development of a self-management program for childhood asthma. *Journal of Asthma, 25,* 335–362.

Creer, T. L., & Bender, B. G. (1993). Asthma. In R. J. Gatchel & E. B. Blanchard (Eds.), *Psychophysiological disorders* (pp. 151–203). Washington, DC: American Psychological Association.

Creer, T. L., & Bender, B. G. (1995). Pediatric asthma. In M. C. Roberts (Ed.). *Handbook of pediatric psychology* (2nd ed., pp. 219–240). New York: Guilford.

Creer, T. L., & Christian, W. P. (1976). *Chronically-ill and handicapped children: Their management and rehabilitation*. Champaign, IL: Research Press.

Creer, T. L., & Holroyd, K. A. (1997). Self-management. In A. Baum, C. McManuis, S. Newman, J. Weinman, & R. West (Eds.), *Cambridge handbook of psychology, health, and medicine* (pp. 255–258). Cambridge, UK: Cambridge University Press.

Creer, T. L., Ipacs, J., & Creer, P. P. (1983). Changing behavioral and social variables at a residential treatment facility for childhood asthma. *Journal of Asthma, 20,* 11–15.

Creer, T. L., & Levstek, D. (1996). Medication compliance and asthma: overlooking the trees because of the forest. *Journal of Asthma, 33,* 203–211.

Creer, T. L., & Levstek, D. (1997). Adherence to asthma regimens. In, D. S. Gochman (Ed.), *Handbook of health behavior research* (Vol. II, pp. 131–148). New York: Plenum.

Creer, T. L., Levstek, D., & Reynolds, R. V. C. (in press). History and conclusions. In H. Kotses & A. Harver (Eds.), *The self-management of asthma*. New York: Marcel Dekker.

Creer, T. L., Renne, C. M., & Christian, W. P. (1976). Behavioral contributions to the rehabilitation of childhood asthma. *Rehabilitation Literature, 37,* 226–232; 247.

Dahlem, N. W., Kinsman, R. A., & Horton, D. J. (1979). Requests for as-needed medications by asthmatic patients. *Journal of Allergy and Clinical Immunology, 63,* 23–27.

Desquin, B. W., Holt, I. J., & McCarthy, S. M. (1994). Comprehensive care of the child with a chronic condition. Part 1. Understanding chronic conditions in childhood. *Current Problems in Pediatrics, 24,* 199–218.

Dudley, D. L., & Sitzman, J. (1988). Psychobiological evaluation and treatment of COPD. In A. J. McSweeney & I. Grant (Eds.), *Chronic obstructive pulmonary disease: A behavioral perspective* (pp. 183–135). New York: Marcel Dekker.

Ellis, E. F. (1993). Asthma in infancy and childhood. In W. W. Busse (Eds.), *Allergy: Principles and practice* (4th ed., pp. 1225–1262) St. Louis, MO: Mosby.

Evans, R. III. (1993). Epidemiology and natural history of asthma, allergic rhinitis, and atopic dermatitis. In

E. Middleton, Jr., C. E. Reed, E. F. Ellis, N. F. Adkinson Jr., J. W. Yunginger, & W. W. Busse (Eds.), *Allergy: Principles and practice* (4th ed., pp. 1109–1136). St. Louis, MO: Mosby.

Fletcher, C. M., Peto, R., Tinker, C. M., & Speizer, F. E. (1976). *The natural history of chronic bronchitis and emphysema.* Oxford, UK: Oxford University Press.

French, T. M., & Alexander, F. (1941). Psychogenic factors in bronchial asthma. *Psychosomatic Medicine Monographs, 4,* 1–92.

Grant, I., Heaton, R. K., McSweeny, A. J., Adams, K. M., & Timms, R. M. (1980). Brain dysfunction in COPD. *Chest, 77,* 308–309.

Jensen, P. S. (1983). Risk, protective factors, and supportive interventions in chronic airway obstruction. *Archives of General Psychiatry, 40,* 1203–1207.

Jones, N. F., Kinsman, R. A., Dirks, J. F., & Dahlem, N. W. (1979). Psychological contributions to chronicity in asthma: patient response styles influencing medical treatment and its outcome. *Medical Care, 17,* 1103–1118.

Kaliner, M., & Lemanske, R. (1992). Rhinitis and asthma. *Journal of the American Medical Association, 268,* 2851–2873.

Kanner, R. E. (1996). Early intervention in chronic obstructive pulmonary disease: A review of the Lung Health Study results. *Medical Clinical of North America, 80,* 523–544.

Kaplan, R. M., & Atkins, C. J. (1988). Behavioral interventions. In A. J. McSweeney & I. Grant (Eds.), *Chronic obstructive pulmonary disease. A behavioral perspective* (pp. 123–162). New York: Marcel Dekker.

Kaplan, R. M., & Atkins, C. J., & Timms, R. M. (1984). Validity of a quality of well-being scale as an outcome measure in chronic obstructive pulmonary disease. *Journal of Chronic Disease, 37,* 85–95.

Kaplan, B. A., & Mascie-Taylor, C. G. (1989). Biosocial correlates of asthma in a national sample of young adults. *Journal of Science, 21,* 475–482.

Kinsman, R. A., Dirks, J. F., & Schraa, J. C. (1981). Psychomaintenance in asthma. Personal styles affecting medical management. *Respiratory Therapy,* 39–46.

Kinsman, R. A., Luparello, T., O'Banion, K., & Spector, S. (1973). Multidimensional analysis of the subjective symptomatology of asthma. *Psychosomatic Medicine, 35,* 250–267.

Kopacz, M. A. (1988). Nursing care of patients with COPD: a sociobiological perspective. In A. J. McSweeney & I. Grant (Eds.), *Chronic obstructive pulmonary disease. A behavioral perspective* (pp. 237–278). New York: Marcel Dekker.

Kotses, H., Bernstein, I. L., Bernstein, D. I., Reynolds, R. V. C., Korbee, L., Wigal, J. K., Ganson, E., Stout, C., & Creer, T. L. (1995). A self-management program for adult asthma. Part I. Development and evaluation. *Journal of Allergy and Clinical Immunology, 95,* 529–540.

Lantner, R. R., & Ros, S. P. (1995). Emergency management of asthma in children: impact of NIH guidelines. *Annals of Allergy, Asthma, & Immunology, 74,* 188–190.

Levstek, D. A. (1996). Follow-up of patients receiving asthma self-management training. Unpublished raw data.

Lewis, S. F., & Fiore, M. C. (1995). Smoking cessation: What works? what doesn't? *The Journal of Respiratory Diseases, 16,* 497–510.

Marion, R. J., Creer, T. L., & Reynolds, R. V. C. (1985). Direct and indirect costs associated with the management of childhood asthma. *Annals of Allergy, 54,* 31–34.

Marwick, C. (1995). Inner-city asthma control campaign under way. *Journal of the American Medical Association, 274,* 1004.

Mazurek, D. A. & Gross, N. J. (1995). Smoking cessation lessons from the Lung Health Study. *The Journal of Respiratory Diseases, 16,* 1049–1060.

McSweeney, A. J. (1984). Assessing the quality of life in patients with chronic obstructive pulmonary disease. In N. K. Wenger, M. E. Mattson, C. D. Furberg, & J. Elinson (Eds.), *Assessment of quality of life in clinical trials of cardiovascular therapies* (pp. 185–217). New York: LeJacq.

McSweeney, A. J. (1988). Quality of life in relation to COPD. In A. J. McSweeney & I. Grant (Eds.), *Chronic obstructive pulmonary disease. A behavioral perspective* (pp. 59–85). New York: Marcel Dekker.

McSweeney, A. J., & Creer, T. L. (1995). Health-related quality of life assessment in medical care. *Disease-a-Month, 41,* 1–72.

McSweeney, A. J., Grant, I., Heaton, R. K., Adams, K. M., & Timms, R. M. (1982). Life quality of patients with chronic obstructive pulmonary disease. *Archives of Internal Medicine, 142,* 473–478.

Meltzer, E. O. (1994). Prevalence, economic, and medical impact of tobacco smoking. *Annals of Allergy, 73,* 381–388.

National Asthma Education Program (1991). *Executive summary: Guidelines for the diagnosis and management of asthma* (Publication No. 91–3042A). Washington, DC: US Department of Health and Human Services.

National Heart, Lung, and Blood Institute (1992). *International consensus report on diagnosis and treatment of asthma* (Publication No. 92–3091). Bethesda, MD: Author.

National Institutes of Health (1995). *Global initiative for asthma* (National Heart, Lung, and Blood Institute Publication No. 95–3659). Bethesda MD: Author.

Niederman, M. S., Clemente, P., Fein, A. M., Feinsilver, S. H., Robinson, D. A., Ilowite, J. S., & Bernstein, M. G. (1991). Benefits of a multidisciplinary pulmonary rehabilitation program. *Chest, 99,* 798–804.

Newacheck, P. W., Budetti, P. P., & Halfon, N. (1986). Trends in activity-limiting chronic conditions among children. *American Journal of Public Health, 76,* 178–184.

Parker, S. R. (1988). Behavioral science aspects of COPD: current status and future directions. In A. J. McSweeney & I. Grant (Eds.), *Chronic obstructive pulmonary disease. A behavioral perspective* (pp. 279–303). New York: Marcel Dekker.

Peto, R. (1994). Smoking and death: the past 40 years and the next 40. *British Medical Journal, 309,* 937–939.

Petty, T. L. (1990). Definitions in chronic obstructive pulmonary disease. *Clinics in Chest Medicine, 11,* 363–373.

Prigatano, G. P., & Grant, I. (1988). Neuropsychological correlates of COPD. In A. J. McSweeney & I. Grant (Eds.), *Chronic obstructive pulmonary disease. A behavioral perspective* (pp. 39–57). New York: Marcel Dekker.

Reed, C. E. (1986). New therapeutic approaches in asthma. *Journal of Allergy and Clinical Immunology, 77,* 537–543.

Renne, C. M., & Creer, T. L. (1985). Asthmatic children and their families. In M. L. Walraich & D. K. Routh (Eds.), *Advances in developmental and behavioral pediatrics,* (Vol. 6, pp. 41–81). Greenwich, CT: JAI Press.

Ries, A. L. (1995). What pulmonary rehab can do for your patients. *The Journal of Respiratory Diseases, 16,* 685–704.

Ries, A. L., Kaplan, R. M., Limberg, T. M., & Prewitt, L. M. (1995). Effects of pulmonary rehabilitation on physiologic and psychosocial outcomes in patients with chronic obstructive pulmonary disease. *Annals of Internal Medicine, 122,* 823–832.

Rogers, R. M., Sciurba, F. C., & Keenan, R. J. (1996). Lung reduction surgery in chronic obstructive lung disease. *Medical Clinics of North America, 80,* 623–644.

Russo, D. C., & Varni, J. W. (1982). Behavioral pediatrics. In D. C. Russo & J. W. Varni (Eds.), *Behavioral pediatrics: Research and practice* (pp. 3–24). New York: Plenum.

Sanford, A., Weir, T., & Pare, P. (1996). The genetics of asthma. *American Journal of Respiratory and Critical Care Medicine, 153,* 1749–1765.

Schiffer, C. G., & Hunt, E. P. (1963). *Illness among children* (Children's Bureau Publication No. 405). Washington, DC: US Government Printing Office.

Sherrill, D. L., Lebowitz, M. D., & Burrows, B. (1990). Epidemiology of chronic obstructive pulmonary disease. *Clinics in Chest Medicine, 11,* 375–387.

Snider, G. L. (1995). What's in a name? Names, definitions, descriptions, and diagnostic criteria of diseases with emphasis on chronic obstructive pulmonary disease. *Respiration, 62,* 297–301.

Spector, S. L., & Nicklas, R. A. (Eds.) (1995). Practice parameters for the diagnosis and treatment of asthma. *Journal of Allergy and Clinical Immunology, 96,* 707–870.

Staudenmayer, H., Kinsman, R. A., Dirks, J. F., Spector, S. L., & Wangaard, C. (1979). Medical outcome in asthmatic patients: Effects of airways hyperreactivity and symptom focused anxiety. *Psychosomatic Medicine, 41,* 109–118.

Stempel, D. A., Hedblom, E. C., Durcanin-Robbins, J. F., & Sturm, L. L. (1996). Use of a pharmacy and medical claims database to document cost centers for 1993 annual asthma expenditures. *Archives of Family Medicine, 5,* 36–40.

Stempel, D. A., Sturm, L. L., Hedblom, E. C., & Curcanin-Robbins, J. F. (1995). Total costs of asthma care. *Journal of Allergy and Clinical Immunology, 95,* 217.

Taylor, W. R., & Newacheck, P. W. (1992). Impact of childhood asthma on health. *Pediatrics, 90,* 657–662.

Tobin, D. L., Wigal, J. K., Winder, J. A., Holroyd, K. A., & Creer, T. L. (1987). A self-efficacy scale for asthma. *Annals of Asthma, 59,* 273–277.

Vance, V. J., & Taylor, W. F. (1971). Status and trends in residential asthma homes in the United States. *Annals of Allergy, 29,* 428–437.

Weiss, K. B., Gergen, P. J., & Hodgson, T. A. (1992). An economic evaluation of asthma in the United States. *New England Journal of Medicine, 326,* 862–866.

Windsor, R. A., Green, L. W., & Roseman, J. M. (1980). Health promotion and maintenance for patients with obstructive pulmonary disease: a review. *Journal of Chronic Disease, 33,* 5–12.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.15
Gynecological Problems

MYRA HUNTER
University College London, UK
and
ANNE WALKER
University of Aberdeen, UK

8.15.1 INTRODUCTION 362
  8.15.1.1 Gynecological Problems 362
  8.15.1.2 Theoretical Models and Practices in Gynecology 362
  8.15.1.3 Gender Issues in Gynecology 364
8.15.2 PSYCHOLOGICAL ASPECTS OF GYNECOLOGICAL PROCEDURES 364
  8.15.2.1 Investigating Complaints 364
  8.15.2.2 Treatment Procedures 365
  8.15.2.3 Psychological Interventions 367
8.15.3 MENSTRUAL DISORDERS 367
  8.15.3.1 Definition and Diagnosis 367
  8.15.3.2 Dysmenorrhea 368
  8.15.3.3 Menorrhagia 368
8.15.4 CHRONIC PELVIC PAIN 369
  8.15.4.1 Gynecological Approaches to Chronic Pelvic Pain 369
  8.15.4.2 Women's Experience of Chronic Pelvic Pain 370
  8.15.4.3 Psychological Interventions for Chronic Pelvic Pain 370
8.15.5 PREMENSTRUAL SYNDROME 371
  8.15.5.1 Definition and Diagnosis 371
  8.15.5.2 Epidemiology and Risk Factors 371
  8.15.5.3 Explanations for Premenstrual Syndrome 372
  8.15.5.4 Biomedical Approaches to Treatment 372
  8.15.5.5 Psychological Interventions for Premenstrual Syndrome 373
8.15.6 MENOPAUSE 373
  8.15.6.1 Definition 373
  8.15.6.2 Epidemiological Studies 374
  8.15.6.3 Hot Flushes/Flashes and Night Sweats 375
  8.15.6.4 Hormone Replacement Therapy 375
  8.15.6.5 Psychological Interventions 376
8.15.7 CLINICAL PSYCHOLOGIST'S ROLE 376
8.15.8 CONCLUSION 376
8.15.9 REFERENCES 377

8.15.1 INTRODUCTION

8.15.1.1 Gynecological Problems

Gynecology can be broadly defined as the science of physiological functions and disorders specific to women (Moscussi, 1993). In practice, this has come to mean functions and disorders of the female reproductive system, rather than of women *per se* (see below). Gynecological problems, then, are those complaints which result from a disorder within the female reproductive system (e.g., ovarian cancer, endometriosis), or are temporally linked with some aspect of this system (e.g., premenstrual syndrome, menopausal hot flushes), and for which help is sought from a gynecologist. The nature of these problems means that gynecology often overlaps with other medical and surgical specialties (e.g., oncology, psychiatry, and pain management), while their frequency means that many gynecological problems are investigated and treated in nonspecialist primary healthcare settings.

The female reproductive system is largely hidden from direct observation and, in most cultures, its functions are private and rarely discussed. Its visible aspects, menstruation and the menstrual cycle, vary from one woman to the next and across the lifespan (Vollman, 1977). As a result, it is often difficult for both women and clinicians to know what is normal. For example, when is menstruation considered to be too heavy, or menstrual pain so severe as to be abnormal? What is a normal frequency for menopausal hot flushes? So, women seek help from healthcare services for a variety of reasons, for information, advice and reassurance, for investigation of asymptomatic abnormalities found in routine screening (e.g., abnormal cervical smear test), as well as investigation and treatment of symptoms. The most common complaints for which women seek help are menstrual changes, pain, and vaginal discharge (Anderson & McPherson, 1984). Some symptomatic women experience a number of conditions, although they may not have been referred (and/or be seeking help) for all of them. For example, Bancroft, Williamson, Warner, Rennie, and Smith (1993) found considerable overlap in symptom experience between women referred to a gynecologist for either premenstrual syndrome (PMS), menorrhagia, or dysmenorrhea.

Most of the women referred to gynecologists are of reproductive age, that is, between adolescence and menopause. As the population ages, however, the number of postmenopausal women referred to a gynecologist for treatment of abnormal uterine bleeding is increasing (Jaluvka & Weitzel, 1995). Hence, women of all ages may experience gynecological problems.

Gynecological problems are associated with high levels of emotional distress, as measured by standard psychological instruments (e.g., General Health Questionnaire, Beck Depression Inventory, Hospital Anxiety and Depression Scale) in surveys of women attending gynecological clinics (Ballinger, 1997; Byrne, 1984; Worsley, Walters, & Wood, 1977). For example, between 40% and 50% of gynecology clinic attenders have been estimated to be at risk for clinical depression as measured by the General Health Questionnaire (Byrne, 1984; Goldberg, 1972). Gynecological problems in themselves may be a cause of distress. In addition, some problems such as premenstrual syndrome include a wide range of psychological as well as physical symptoms within the diagnosis. However, surveys of gynecology clinic attenders suggest that their distress is often associated with concurrent personal and social problems (Worsley et al., 1977), and occurs regardless of whether or not an organic cause can be found for the woman's symptoms (Pearce, 1987). These findings are consistent with evidence from clinic attenders in general that psychological and social factors are important predictors of helpseeking behaviour (Mechanic, 1980) and suggest that some women may, for complex reasons, be expressing emotional distress (which may relate to problems elsewhere in their lives) in a gynecological setting. However, the precise mechanisms underlying presentation of particular symptoms or problems is little understood. Whatever the cause of women's distress, the degree of it is ample evidence that psychological interventions have an important role to play alongside traditional gynecological treatments.

8.15.1.2 Theoretical Models and Practices in Gynecology

Modern gynecology is a medical and surgical specialty located firmly within a biomedical paradigm. However, many of the problems seen by gynecologists cannot easily be explained within this framework (Beard, 1984), and numerous critics have questioned its biologically reductionist approach to women's distress (Daley, 1978; Ussher, 1992). Hence, a broad range of theories have been developed to explain specific gynecological problems, both from within the medical profession and outside it. These theories in turn have had significant impact on both research and practice. Current theories can be broadly separated into biomedical, sociocultural, psychological, and biopsychosocial accounts (Hunter, 1994; Ussher, 1992; Walker, 1995).

The biomedical model proposes that biological factors are the major cause of gynecological complaints. Biological factors include overt organic pathology (e.g., carcinoma, endometriosis, ovarian cysts) and assumed dysfunctions in the hypothalamo–pituitary–ovarian (HPO) axis, which controls the reproductive system (Medvei, 1982). The HPO axis is characterized by a complex and intricate interplay of a variety of peripheral hormones and neurotransmitters, including estrogen, progesterone, luteinizing hormone, and follicle stimulating hormone (Oudshoorn, 1994). Theories which implicate the HPO axis, often described as hormonal theories, have been particularly influential in explanations of menstrual and menopausal problems. The biomedical model has achieved some success in explaining gynecological problems; however, as indicated above, physical symptoms often persist in the absence of evidence of organic pathology or HPO dysfunction (Worsley et al., 1997). In particular, there has been little empirical support for biomedical explanations of psychological symptoms or emotional distress associated with gynecological problems, despite its popularity in both lay and medical discourse (Ussher, 1992). The treatments that women receive from gynecologists are, however, generally derived from these theories. Surgical intervention, usually the removal or modification of parts of the reproductive system (e.g., hysterectomy, endometrial ablation, pelvic floor repair), is the treatment of choice in most cases of overt pathology, but is often also seen as a last resort solution to chronic, unexplained gynecological problems (O'Brien, 1994), raising concerns that it is being used inappropriately (Goudsmit, 1994; Smith, 1996). Medical treatments range from those targeted at specific symptoms (e.g., analgesics, anti-inflammatories, antidepressants) to those which alter aspects of HPO function (e.g., antigonadotropins, gonadotropin agonists, oral contraceptives, hormone replacement therapy, etc.), each of which is associated with its own risks and benefits (Mortola, 1994).

Sociocultural accounts of gynecological problems locate their cause in some aspect of a woman's social or cultural circumstances. These circumstances may expose her to hazards that increase the likelihood of pathology, for example, environmental pollution, aspects of working conditions, heterosexual activity, poverty, nutrition, and so on (Kuh & Stirling, 1995; Messing, Saurel-Cubizolles, Bourgine, & Kominski, 1993). In addition, there is considerable variation between cultures and across time in the definition of particular physical or emotional states as signs of illness (Helman, 1990). Hence, sociocultural forces may also act to "pathologize" experiences which are not inherently abnormal, by labeling them as problems or diseases (Dingwall, 1992). For example, it has been argued that the normal feelings which most women experience premenstrually have been constructed as problematic by modern Western cultures (and labeled PMS), because they result in behaviors which do not conform to societal expectations of femininity (Johnson, 1987; Laws, Hey, & Eagon, 1985; Martin, 1989). Hence, radical feminist writers have argued that women are controlled by the simultaneous medicalization of "normal" experiences, such as menstruation and menopause, and lack of medical interest in abnormal experiences, such as endometriosis or pelvic inflammatory disease (Daly, 1978). From this perspective, interventions are targeted at societies, rather than individuals, and may range from lobbying for safer working conditions to promoting safer sex or changing public perceptions of women.

Psychological models are predominantly concerned with the explanation of emotional distress in relation to gynecological problems or experiences. The earliest models were psychodynamic, attributing anxiety or depression to unconscious conflicts triggered by physiological states, for example, around menstruation (Daly, 1935; Horney, 1931). More recently, cognitive, social cognition, and biopsychological models have been developed (Asso, 1992; Ruble & Brooks-Gunn, 1979). Currently, the most influential psychological models relate to stress and coping, and are adapted largely from the work of Lazarus and Folkman (1984), among others. These models propose that the cognitive appraisal of physiological states or gynecological problems as stressful is the cause of emotional distress, and that a variety of coping mechanisms can be used to relieve the distress. Clinical interventions function either to alter the appraisal of the situation or to develop effective coping strategies. A variety of techniques have been used to do this, with cognitive-behavior therapy approaches being most popular (Hunter, 1994). Relaxation training, stress management, social support through self-help groups, and other strategies are often used additionally to help women to cope with their experiences.

Despite the proliferation of models and practices in gynecology, it is still not uncommon for distress in a gynecological context to be attributed to either a woman's body (Studd & Smith, 1993) or her mind (Goudsmit, 1994). This mind–body split can lead to inevitable biases in evaluation and treatment of problems, the neglect of psychosocial causes of problems, as well as communication difficulties and dissatisfaction with consultations. Biopsychosocial or multifactorial models have been developed to address the need to examine the complex relationships between psychological, social and biological influences (Hunter, 1994; Ussher, 1992; Walker, 1995). A biopsychosocial perspective attempts to reconcile the different etiological theories and to explain individual differences between women. A woman's psychosocial and cultural contexts are taken into account as well as her experience of physical changes, in determining the meaning and impact of gynecological problems. Cognitive factors such as symptom attributions or illness representations are often positioned centrally in biopsychosocial models, and the model can be expanded and tailored to individual women in clinical settings. From this perspective, single or multiple interventions may be needed, depending on the individual concerned, and the clinical health psychologist becomes a crucial member of the therapeutic team.

8.15.1.3 Gender Issues in Gynecology

Gynecology has been practiced since ancient times (McKay, 1901; Moscucci, 1993), and the theories and practices it has adopted reflect the social and cultural gendered discourses, specific to particular times (Ehrenreich & English, 1973; Moscucci, 1993; Smith-Rosenberg & Rosenberg, 1973). For most of written history, behavior which does not conform to a particular female stereotype has been regarded as a sign of either illness or madness, and attributed to women's reproductive body. For example, in Greco-Roman times emotional and health problems in women were attributed to the "wandering womb" (McKay, 1901). Similarly, nineteenth century expressions of distress or dissatisfaction were seen as sexual or "hysterical" in origin and gynecological surgery (e.g., clitoridectomy) was often used as a treatment (Showalter, 1987; Ussher, 1991). In the past 20–30 years, hormonal explanations of women's distress during the menstrual cycle, the postpartum period, and the menopause have become more pervasive (Dalton, 1977; Studd & Smith, 1993). No matter what form specific theories take, it remains the case that women's well-being and behavior is attributed to some aspect of reproductive function in a way that men's well-being and behavior is not (andrology is a far less developed discipline than gynecology, for example (Gannon, 1996). Hence, gender issues underlie the very existence of gynecology (Moscucci, 1993).

Gender is also an issue at a more interpersonal level in gynecological practice. Gynecology is an unusual medical specialty in that the client group is specifically female and the doctors predominantly male (Cockburn & Bewley, 1996). The problems under investigation, being related to women's bodies, sexuality, and reproduction, are inevitably influenced by social constructions of gender and gender-appropriate behavior. For example, a gynecological consultation, which frequently includes an internal examination, may be seen to recreate a stereotypical heterosexual relationship in which the female is powerless, passive, and dependent (Weiss & Meadow, 1979).

It has been argued, ironically, that even this branch of medicine derives its norm from the heterosexual male (Rosser, 1994), defining women's health, especially their sexual health, in terms of relationships with men. Thus gynecological practice focuses primarily upon heterosexual activity, and therefore may regulate aspects of women's sexuality and reproduction, for example, by means of access to abortion and fertility treatments, and by surgical treatments such as hysterectomy and oophorectomy. Scant attention has been paid to the gynecological health needs of lesbian women, who are assumed (erroneously) to experience fewer gynecological problems than heterosexual or bisexual women (Burns, 1992; Johnson, Smith, & Guenther, 1987). In addition, some reproductive issues for which both men and women are responsible, such as contraception and infertility, have become defined as solely or primarily women's health issue (Woollett, 1996).

8.15.2 PSYCHOLOGICAL ASPECTS OF GYNECOLOGICAL PROCEDURES

8.15.2.1 Investigating Complaints

The first stage after a woman has been referred to a gynecologist is investigation of her complaint. The range of problems seen in gynecological clinics means that this investigation may take a variety of forms, from invasive procedures, such as dilatation and curettage (D&C), to daily diary keeping and questionnaire completion. The increasing emphasis on screening and well-woman clinics in primary health care means that gynecological investigations (especially pelvic examinations and cervical smear tests) are not only conducted in gynecology clinics, although more invasive procedures are more likely to occur in hospital settings.

Most investigations require a vaginal examination which for many women is unpleasant and for some traumatic (Areskog-Wijma, 1987). Female genitalia are generally hidden from view, and in most cultures showing and touching of them is taboo outside intimate sexual relationship (Rafferty & Williams, 1996). Although the doctor may be concerned with safety and the routine nature of these examinations, many women feel uncomfortable about a stranger looking at and touching an intimate part of their anatomy (Wijma & Areskog-Wijma, 1987). Some women describe these examinations as frightening and degrading, and in surveys many express a preference for a female doctor (Areskog-Wijma, 1987; Ivins & Kent, 1993). For a subgroup of women, gynecological procedures may be so stressful that they experience symptoms of post-traumatic stress disorder (PTSD) afterwards. Menage (1993) surveyed the gynecological experiences of 500 women recruited through press advertisements, 30 of whom met the criteria for PTSD. Although this is not an epidemiological study and may overestimate the frequency of traumatic experiences, it does suggest that the psychological impact of these procedures can be considerable. Thus, from the woman's point of view the gynecological consultation, involving discussion of both personal and sexual issues in combination with an intimate physical examination, is rarely considered routine, and for many women can be a stress in itself (Rafferty & Williams, 1996).

Medical consultations in general are unlike ordinary social encounters and have been extensively studied by social scientists, especially with regard to characteristics which hinder communication between healthcare professionals and clients (Ley, 1988; Pitts, 1991; Stewart & Roter, 1989). Communication between health-care professionals and women seeking help for gynecological problems is often difficult. Many women express dissatisfaction with the amount or quality of information they receive, especially from general practitioners, about the interventions or investigations they undertake (Corney & Stanton, 1991; Grace, 1995a). Many women feel that their symptoms are not taken seriously by their doctors, or that they are seen as neurotic. This is particularly true of conditions such as PMS, menorrhagia, or chronic pelvic pain, which are difficult to treat effectively. In these situations, many women lose faith in doctors and develop their own strategies for coping (Scambler & Scambler, 1993; Zadinsky & Boyle, 1996.)

8.15.2.2 Treatment Procedures

Women experiencing gynecological procedures are a heterogeneous group in terms of age, and other factors—religion, ethnicity, sexuality, social class, education, physical and intellectual ability, for example. As a consequence, it is difficult to generalize about the psychological aspects of their experience or the meaning that particular procedures will have for all women. Epidemiological studies have shown that some groups of women are more likely to experience some types of procedure than others. For example, women with learning difficulties are more likely to experience therapeutic amenorrhea, and less likely to experience laparotomies or D&C procedures than other women (Houvinen, 1993). Similarly, a large nationally representative cohort study of hospital admissions in Britain found significant differences in the frequency of hysterectomy, D&C, and hospital admission rates according to social class and educational qualification, with all procedures being more frequent amongst the less advantaged groups (Kuh & Stirling, 1995). Studies like this suggest that, although there may be similarities across women, there will also be differences—which will be mediated by inequalities in healthcare and clinical assumptions about what is in a particular woman's interest. Further research is needed to investigate why gynecological procedure rates vary in this way.

In addition to the stresses associated with any gynecological consultation, particular procedures also arouse anxiety, because of concerns about what an investigation may show or the longer term consequences of treatment. It is difficult in these situations to disentangle the psychological effects of particular diagnoses (especially those which might imply cancer) and the stress of the procedure itself. For example, women who undergo laser treatment to prevent the development of cervical intraepithelial neoplasia (CIN) into cervical cancer experience considerable emotional upheaval and upset as they progress through diagnosis and treatment (Nugent & Tamlyn-Leamann, 1992; Posner & Vessey, 1988), although the treatment itself may not have measurable psychological costs over and above the effect of diagnosis (Palmer, Tucker, Warren, & Adams, 1993). Palmer et al. (1993) found that women who were diagnosed as having CIN had significantly higher scores on measures of intrusive thoughts, avoidance, and state anger than women who received a negative smear test result; however, there were no significant differences on these measures before and after laser treatment. Although these findings suggest that the treatment does not have additional psychological costs, they also imply that there is no immediate psychological benefit of the treatment. Interviews with the women suggested that fear of recurrence and progression of the CIN into cancer was a major worry, and although they were glad to have had the treatment, they were not fully reassured that it would prevent abnormal cells in the future. Many of the women attributed the cause of the disease to sexual activity, especially female promiscuity, and felt differently about their own bodies and their sexual relationships as a result. The feelings of defilement many of them described were an additional component of their distress. Studies like this suggest that women might benefit from both more information about their condition and greater opportunities to talk through their experiences. As yet, however, psychological interventions in this area are awaiting evaluation (Rafferty & Williams, 1996).

Hysterectomy, removal of the uterus, is one of the most common major surgical operations in many countries, although the rates of hysterectomy vary considerably both between countries and in different parts of the same country (Teo, 1990). In Britain, 10% of women have had a hysterectomy by the age of 43 (Kuh & Stirling, 1995), and 20% by the age of 65 (Chepple, 1995). In many cases, the woman's ovaries are removed (oophorectomy) at the same time to prevent ovarian cancer. In premenopausal women, hormone replacement therapy (HRT) is usually prescribed if the ovaries are removed to counteract the effect of a premature menopause. A hysterectomy can be conducted abdominally (in which an incision is made and the uterus removed), or vaginally. Both procedures require general anesthesia and hospital admission. Postoperative recovery usually takes about eight days in hospital and then six weeks or so to convalesce (Evans & Richardson, 1988).

The psychological impact of hysterectomy has been extensively studied. Early research was based on the hypothesis that the loss of the uterus would make a woman feel less feminine and have a negative psychological effect. A number of studies supported this idea, finding high rates of psychiatric disorder after the operation (Baker, 1968). More recent research has criticized these findings, primarily for failing to assess psychological state before hysterectomy. In carefully controlled studies, rates of psychological distress (e.g., depression or anxiety) after hysterectomy are similar to or lower than those prior to surgery (Gath, Cooper, & Day, 1982; Ryan, Dennerstein, & Pepperell, 1989). There is also some evidence to suggest that psychiatric morbidity both before and after hysterectomy varies across different cohorts of women. Gath et al. (1995) describe findings from three studies conducted between 1975 and 1990, all of which investigated the psychological state of women having a hysterectomy for menorrhagia before and six months after the operation, using the Present State Examination. In all three studies, rates of psychiatric morbidity were lower after the operation than before. However, the number of women experiencing distress at both times fell dramatically over the three studies. The proportion of women considered to be psychiatric cases before and after the operation was 58:26 in study one, 28:7 in study two, and 9:4 in study three. One of the biggest changes over this time period has been the development of medical treatments and less radical surgical strategies for managing menorrhagia. Hence, it may be that menstrual problems of the women in the later studies were better controlled before the operation than those in the earlier studies, causing them less distress. As with the studies of CIN and colposcopy (above), this research shows that it is difficult to disentangle the psychological effects of a woman's problem or diagnosis from the effects of the procedure itself. However, there is little evidence here to suggest that hysterectomy is necessarily traumatic.

As well as the effect of hysterectomy on psychological state, its effect on sexual activity and enjoyment has also been investigated. These studies suggest that effects differ between women, with some experiencing an improvement in their sex lives after surgery, and others (approximately one in five) a deterioration (Alexander et al., 1996; Helstrom, Lurdburg, Sorbon, & Backstrom, 1993). Women whose ovaries have been removed are more likely to be dissatisfied with sex after the operation, often describing difficulties with sexual arousal and vaginal lubrication, regardless of HRT use (Natharst-Boos, von Schoultz, & Carlstrom; 1993). The best predictors of satisfaction with sex after surgery are measures of satisfaction before the operation and feelings towards a sexual partner (Helstrom et al., 1993), suggesting that for some women relationship or sexual counseling may be helpful.

A major problem with most studies of hysterectomy (and other procedures) is variation in the length of follow-up time after the operation. Positive effects soon after the operation may be due to relief of the woman's menstrual or pain symptoms, whereas any longer-term impact of the operation cannot be assessed. Women interviewed two years after hysterectomy reported at least some negative symptoms which they attributed to the operation, despite being fairly positive about the outcomes three months after surgery (Bernhard, 1992), suggesting that more long-term research is needed to understand the impact of hysterectomy in women's lives.

Recently a variety of surgical techniques have been developed which remove or reduce the endometrium (lining of the uterus) without removing the uterus itself. These techniques have been evaluated largely in the treatment of menorrhagia, which accounts for approximately 30% of hysterectomies (Lalonde, 1994). Endometrial ablation (EA) or resection techniques require a general anesthetic, but are minor surgical procedures compared to hysterectomy, requiring less time in hospital (they can often be undertaken in day-care units) and less postoperative recovery time (Rafferty & Williams, 1996). These new techniques offer an alternative for women who do not want a hysterectomy or for whom hysterectomy is unsuitable. A proportion of women, however, experience a return of their symptoms at some point after the procedure, and may go on to have a hysterectomy. Hence, many women are faced with a difficult choice between endometrial resection/ablation and hysterectomy. There have been few studies so far which investigate the psychosocial consequences of EA. A recent randomized controlled trial of EA and hysterectomy found that both significantly reduced anxiety and depression, with no differences in mental health between the two groups 12 months after surgery (Alexander et al., 1996).

8.15.2.3 Psychological Interventions

Psychological interventions are usually presented as part of the preparation for either invasive or surgical procedures, with the purpose of reducing anxiety or enhancing coping abilities, and thereby promoting recovery. Their presentation is managed in a variety of ways, during a brief interview with a psychologist, nurse, or anesthetist, as an audiotape, or most frequently in the form of a booklet (Wallace, 1984). Ridgeway and Matthews (1982) describe five main types of psychoeducational intervention that have been investigated in relation to preparation for all types of surgery. Two of these are broadly educational, that is, information-giving interventions, which provide patients with either factual or sensory descriptions of what will happen before, during and after the surgery, and behavioral instruction techniques, in which patients are taught skills which may be helpful after surgery, for example, deep breathing or specific exercises. Interventions with a larger psychological component include relaxation training, modeling (in which patients observe someone else overcoming anxiety), and cognitive coping training (in which patients are encouraged to replace worries with more positive thoughts). Studies that have tested these interventions are difficult to evaluate because of the variety of outcome measures used, although early reviews concluded that all of the strategies have some potential benefits, especially if psychological and educational strategies are combined (Mumford, Schlesinger, & Glass, 1982; Weinman & Johnston, 1988). A more recent meta-analysis of 38 randomized controlled trials of psychological preparation of adult patients for a variety of types of surgery found that all of these approaches produce better-than-chance benefits on a variety of outcome measures (e.g., length of stay in hospital, negative affect, pain, satisfaction), with procedural information and behavioral instructions showing the most wide-ranging effects across all outcome measures, followed by relaxation training (Johnston & Vogele, 1993).

Studies of interventions such as these specifically in relation to gynecological procedures have found similar results (Rafferty & Williams, 1996; Wallace, 1984). For example, Wallace (1983) found that patients given a booklet containing procedural, sensory, temporal, and coping information as well as general reassurance had significantly better recovery after laparoscopy than patients who received routine care or a booklet containing global reassurance alone. When anxiety is the outcome measure, however, increased information alone may not be helpful. Marteau, Kidd, Cuddeford, and Walker (1996) sent women referred for colposcopy either a simple booklet, containing procedural information, behavioral instructions, and outcome information, or a more complex booklet containing more detailed procedural and outcome information, and assessed their anxiety at the time of receiving the booklet and when attending the hospital. Although both booklets increased women's knowledge about the procedure, the simple one was more effective in reducing anxiety. This finding suggests that information booklets are not universally effective in relieving anxiety, and it may be the case, as other studies have suggested, that detailed information is more suitable for some women than others, depending on their individual coping styles (Miller, 1988; Miller & Mangan, 1983; Shipley, Butt, Horwitz, & Farby, 1978).

8.15.3 MENSTRUAL DISORDERS

8.15.3.1 Definition and Diagnosis

The average woman experiences around 400 episodes of menstrual bleeding in her lifetime (Profet, 1993) The neurophysiological control of the menstrual cycle is complex and not yet completely understood, and its expression varies considerably between women (Profet, 1993). Hence, women vary in terms of the number of days between menstrual periods, (although 28 days is an average figure, only one in eight cycles is 28 days long) (Vollman, 1977), the number of days on which menstruation occurs, and the amount of blood and fluid lost during menstruation (Treloar, Boynton, Benn, & Brown, 1967; Snowden & Christian, 1983). Not surprisingly, then, menstrual disorders are difficult to define, and a cause cannot always be identified (Scambler & Scambler, 1993). Menstrual disorders are among the 10 most common complaints seen in primary care settings (Scambler & Scambler, 1993). Their epidemiology is difficult to assess, however, because not all women who experience them seek help from physicians. In a community-based study in London, Scambler and Scambler (1985) found that over 50% of the women who described severe menstrual disorders had not sought help from a doctor in the last year. The reasons for this are complex, but in this study, women described feeling that their doctor (or doctors in general) had little empathy with or understanding of menstrual problems and had little to offer in the way of acceptable remedies. Information about the epidemiology can be obtained from community-based questionnaire surveys: however, these are difficult to interpret because women's perceptions do not always match medically defined disorders (Scambler & Scambler, 1993). It is likely, however, that large numbers of women experience menstrual problems at some time in their life, and that most of these women develop their own ways of coping with them.

8.15.3.2 Dysmenorrhea

The most common menstrual problem is dysmenorrhea, or pain before and during menstruation. Painful periods are so common that many people dismiss them as normal. In studies of nonpatient populations as many as 70% of women report some pain associated with menstruation, and 5–10% regularly experience pain which is severe enough to be incapacitating for between an hour to three days each month (Golub, 1992). The pain is usually described as "primary dysmenorrhea" if it occurs in the absence of any pelvic pathology and "secondary dysmenorrhea" if there is pathology present (e.g., endometriosis, fibroids). Women describe a number of different types of pain which fall under the general heading of dysmenorrhea. Typically there may be cramps—intermittent sharp pains in the lower midabdomen which may radiate to the lower back or thighs or a steadier dull aching pain which may also extend into the legs. Some women experience nausea, vomiting, gastrointestinal disturbance, or headaches at the same time.

Although the etiology of dysmenorrhea is not completely understood, it is associated with high levels of prostaglandins in the uterus (Smith, 1986), and treatment with antiprostaglandins (nonsteroidal anti-inflammatory drugs [NSAIDs]) is effective in about 80% of cases (Benassi, Bertani, & Avanzini, 1996). The frequency of primary dysmenorrhea decreases with age and childbearing (Golub, 1992). As well as age and not having experienced pregnancy, there are suggestions in the literature that a diet low in fish oils (Deutch, 1995; Harel, Biro, Kottenhahn, & Rosenthal, 1996), occupational hazards (Messing et al., 1993), and smoking (Parazzini et al., 1994) may increase the risk of dysmenorrhea. Period pain is less severe in the absence of ovulation, so for some women oral contraceptives are a useful treatment (Robinson, Plichta, Weisman, Nathanson, & Ensminger, 1992). Nonpharmacological interventions can also be effective, for example, transcutaneous electrical nerve stimulation (Kaplan et al., 1994) or spinal manipulation (Korjohn, Schmid, Triano, & Brennan, 1992). For the majority of women, though, the NSAIDs are the preferred method of coping (Jarrett, Heitpempes, & Shaves, 1995).

Early studies suggested a link between dysmenorrhea and personality or attitudes towards femininity (Gannon, 1985). However, studies conducted after the discovery of uterine prostaglandins in the 1960s assume that it is pain which has an effect on psychological state, and not vice versa. Women who experience severe period pain are more likely to be depressed or anxious during menstruation than women who do not (Golub, 1992).

8.15.3.3 Menorrhagia

Although dysmenorrhea may be the most commonly experienced menstrual problem, menorrhagia (heavy bleeding) is the one most likely to be seen in gynecology clinics. In clinical terms, menorrhagia is usually defined as loss of 60–80 ml of blood or more in one menstruation (Smith, 1986) a level which may put women at risk for anemia. The amount of blood lost during menstruation is difficult to assess, and accurate diagnosis requires assessment of the quantity of blood in all of the tampons and/or pads used, a procedure which is rarely conducted outside research studies. Studies which have done this suggest that there is an imperfect relationship between perceived heaviness of menstruation, or the number of sanitary pads or tampons used, and excessive blood loss, partly because blood is only one component of menstrual fluid (Fraser, 1994). If women are asked to use a pictorial chart to rate the amount of blood loss, a much better relationship with actual blood loss is found (Janssen, Schoefen, & Heintz, 1995). There are many organic causes of menorrhagia, for example, uterine fibroids, thyroid dysfunction, or intrauterine contraceptive devices. For as many as half of the women complaining of heavy periods, however, no organic cause can be found (Anderson & McPherson, 1983). For both groups, the most likely medical investigation is a D&C, followed by hormonal and/or surgical (usually hysterectomy) treatments.

There has been little research which investigates women's experience of menorrhagia or its psychological concomitants. Greenberg (1983) found that 62% of women seeking help from a gynecologist for heavy periods were also suffering from mild to moderate depression. These women were also experiencing more life stresses and less severe gynecological symptoms than the women who were not depressed. Harris (1989) also found a positive relationship between recent life events and help seeking for menorrhagia. She hypothesizes that, rather than the life events having a direct effect on menstrual physiology, it may be that provoking events and depression increase the psychological impact of blood loss. It may simply be that the woman's usual ability to cope is exhausted by other pressures, and so she seeks help from a physician. Alternatively, it may be the case that heavy menstruation is itself difficult to cope with, for example, interfering with work and relationships, causing depressions and a negative perception of recent events. Further research is needed to investigate the intricate relationships between menstrual disorders, psychological states and help seeking, but this may be an area in which cognitive coping interventions prove beneficial.

For women who experience heavy periods, alternatives to surgery and medical treatments are limited. Most women would prefer a treatment that normalizes menstruation, with a reversible contraceptive effect, and are equally divided as to whether they would prefer tablets or a one-off operation (Warner, 1994). This ideal treatment for menstrual problems is some way in the future. In the meantime, as well as psychological interventions, development of new methods of sanitary protection, such as a menses cup (Cheng, Kung, Hannah, Wilansky, & Shime, 1995) which can effectively manage the menstrual flow, may enhance the quality of life of many women.

8.15.4 CHRONIC PELVIC PAIN

8.15.4.1 Gynecological Approaches to Chronic Pelvic Pain

Chronic pelvic pain (CPP) refers to persistent pelvic pain lasting for six months or more (Rapkin & Kames, 1987). The exact prevalence of CPP is unknown: however, it is one of the most common complaints amongst women seeking help from a gynecologist, accounting for as many as 10% of outpatient consultations (Reiter, 1990) and, in a recent study, 38% of women attending a gynecology clinic reported experiencing CPP at some time in their lives (Walker, Katon, Neraas, Jemelka, & Massoth, 1992). CPP is a frustrating condition for gynecologists because it often occurs without any apparent organic pathology. The rate of chronic pelvic pain without organic pathology (CPPWOP) varies between studies and according to the techniques used to identify pathology. In studies that rely on pelvic examination, up to two-thirds of CPP sufferers fall into the CPPWOP category (Gross, Doerr, Caldirola, Guzinski, & Ripley, 1980–81; Pearce & Beard, 1984). Studies that use laparoscopy are more variable, finding no demonstrable pathology in between 9% and 80% of CPP sufferers (Reiter, Gambone, & Johnson, 1991). However, in most of these studies between a quarter and a third of women complaining of pain obtain negative laparoscopic findings (Stout, Sleege, Dodson, & Hughes, 1991; Vercellini, 1989). When pathology is found it is usually endometriosis or adhesions, the extent of which is not related to women's self-reports of their pain intensity (Stout et al., 1991). Hence, the relationship between CPP and organic pathology is complex and variable.

In an attempt to understand why pain might occur, researchers have compared groups of women experiencing CPP without identifiable organic pathology with women whose CPP is associated with organic pathology on a variety of psychological and sociodemographic variables. Researchers have found higher rates of neuroticism (Rosenthal, Ling, Rosenthal, & Stovall, 1991), depression and anxiety (Magni, Saloni, Delco, & Ceola, 1984), sexual problems (Bak, Verhage, Drogendijk, Voitus van Hamme, & Duivenoorden, 1989) and somatization (Magni, Andreoli, de Leo, Martinotti, & Rossi, 1986) among women without apparent pathology. However, well-controlled studies (which use laparoscopy to diagnose organic pathology) and a recent meta-analysis have failed to support these findings (Hodgkiss & Watson, 1994; McGowan, Clark-Carter, & Pitts, 1994; Pearce, 1987). In general, women with CPP, regardless of whether or not an organic cause can be identified, are more depressed and anxious than women who are pain-free (Davies, Ganger, Drummond, Saunders, & Beard, 1992) and have higher rates of sexual problems. Women seeking help for CPP also report more experiences of sexual abuse or trauma than pain-free women (Reiter & Gambone, 1990; Schei, 1991) but not than women with chronic pain in other locations (Rapkin, Kames, Darke, Stampfle, & Naliboff, 1990). Although there is a relationship between sexual abuse and CPP, it is not the only factor, and many women who experience CPP have not been abused. Stout (1995) argues that CPP is a heterogeneous problem area rather than a single diagnosis, and that in order to understand it a multifactorial model is needed, incorporating biological, psychological, and social aspects, which act together to form a vicious cycle that maintains the pain.

Surgical (usually laparoscopic surgery or hysterectomy) and hormonal treatments are predominantly offered for CPP, although neither is universally successful, and many women experience side effects, inadequate pain relief, or a relapse of symptoms (Ivins & Kent, 1993; Stout, 1995).

8.15.4.2 Women's Experience of Chronic Pelvic Pain

The experience of CPP for many women is one of repeated investigations, a variety of medical or surgical interventions, and the development of self-care coping strategies (Grace, 1995b; Zadinsky & Boyle, 1996). The absence of a clear diagnosis for many women with CPP is a source of frustration. Many women feel that their complaints are being dismissed as imaginary or that there is nothing wrong with them. In a survey of 336 women using the health services in New Zealand for CPP, 43% reported that a GP had at some stage told them they had nothing wrong with them, and 24% claimed that a gynecologist had at some stage told them this (Grace, 1995a). Similarly, 25% claimed that a GP had told them the pain was due to a psychological problem, although very few were referred to a psychologist. As well as problems with the diagnosis, these women also described problems with communicating with doctors (over 40% felt that there were things about their pain which they had not been able to talk to the doctor about, and 58% felt that the doctor was not taking their pain seriously) and problems with side effects of medication and/or surgery. The frequency of these problems was greatest (> 50%) among women in lower income groups, but still high (25–35%) in the higher and middle income groups. These findings suggest that communication in relation to CPP could be improved. Zadinsky and Boyle (1996) interviewed CPP sufferers and found that in the absence of a cure being offered by their doctor, women developed their own self-care strategies to enable them to deal with the disruptive effects of their pain on their lives.

8.15.4.3 Psychological Interventions for Chronic Pelvic Pain

A growing body of research suggests that psychological interventions may be helpful in enhancing women's ability to cope with CPP (Bak et al., 1989; Kames, Rapkin, Naliboff, Afifi, & Ferrer-Brechner, 1990; Peters et al., 1991; Stout, 1995). However, since these are often offered in combination with other pain management strategies, specific effects are difficult to evaluate (Stout, 1995). A small number of studies have focused specifically on psychological interventions. For example, Pearce, Knight, and Beard (1982) compared relaxation training with behavioral counseling and nondirective psychotherapy, in a group of 32 patients with pain of at least six months' duration, and no apparent pathology after laparascopic investigations. The women were randomly allocated to one of the three treatment groups, and each was seen individually for four to six sessions at approximately fortnightly intervals. They rated their pain daily on a five-point scale continuously from a month before laparoscopy to three months after treatment ended. All three groups showed a significant increase in pain-free days at follow-up when compared to the initial assessment. These increases were continuing at one year follow-up in both the relaxation and nondirective counseling groups. Studies like this suggest that interventions that incorporate aspects of progressive muscle relaxation may be particularly helpful for CPP. Other forms of intervention may also be helpful, for example, sexual counseling or assertiveness training, but have yet to be evaluated. Given the experiences women describe, interventions aimed at health care professionals to enhance communication and limit psychologization might also have beneficial effects (Grace, 1995a).

8.15.5 PREMENSTRUAL SYNDROME

8.15.5.1 Definition and Diagnosis

Premenstrual tension first appeared in the medical and psychoanalytic literature in the 1930s (Frank, 1931; Johnson, 1987), to describe feelings of "indescribable tension" (Frank, 1931) in the days before a period. It was renamed premenstrual syndrome (PMS) by Greene and Dalton (1953) to encompass any physical or emotional complaint occurring regularly before menstruation, and became notorious in the 1980s when it was used as a mitigation in two murder trials (Dingwall, 1992; Rittenhouse, 1991).

There is no doubt that many women have physical and/or emotional experiences that indicate that their period is imminent. The nature and intensity of these experiences varies dramatically, however. Budieri, Li Wan Po, & Dornan (1974) identified 199 different potential signs and symptoms assessed in studies of PMS. Some women experience distressing moods or emotions and/or physical states around or before menstruation others feel at their most creative or energetic at this time (Logue & Moos, 1988; Stewart, 1989), and others still do not experience noticeable cyclical change (Ainslough, 1990; Laessle, Tuschl, Schweiger, & Dirke, 1990). As well as this variability between individuals, premenstrual experiences may also vary within individuals. Some women have different premenstrual experiences in consecutive cycles (Walker, 1994), or at different times of the year (Parry, Rosenthal, Tamarkin, & Wehr, 1987) and, as Gise, Levouits, Paddison, and Strain (1990) point out, events in women's lives (crises, holidays, physical illness, etc.) have effects on both emotional and physical well-being which may mask, exacerbate, or ease premenstrual changes. Given this range and variability of experience, it is not surprising that PMS has proved to be an elusive concept, difficult to define or diagnose.

Numerous attempts have been made within both gynecology and psychiatry to define PMS so that those women who "need" clinical intervention can be distinguished from those who do not (Gold & Severino, 1994; Walker, 1997). Current clinical definitions usually require that the symptoms interfere to some degree with a woman's life, and that they are relieved during or immediately after menstruation. This is usually assessed by completion of daily rating scales over two or more cycles in addition to a detailed history. For example, the *Diagnostic and statistical manual of mental disorders* definition of "premenstrual dysphoric disorder" (PMDD) requires women to experience at least five symptoms, one of which must be either depression, tension, emotional lability, or irritability, in most menstrual cycles for a year, of sufficient intensity to interfere with her life (American Psychiatric Association, 1994). This definition is controversial because it focuses on affective experience—excluding women whose experiences are entirely physical from a diagnosis, while simultaneously implying that all clinically significant premenstrual experiences constitute a psychiatric disorder (Caplan, McCurdy-Myers, & Gans, 1992; Stotland & Harwood, 1994). For some researchers, PMDD is only one manifestation of PMS, and the terms are not interchangeable (Goudsmit, 1994; Stotland & Harwood, 1994). The possibility of several "premenstrual syndromes" or "perimenstrual symptom patterns" has been suggested (Gise, 1988; Mitchell, Woods, & Lantz, 1991), although not always supported empirically (Jorgensen, Rossignol, & Bornlander, 1993). Others have argued that PMS is an "idiosyncratic syndrome," best approached from an individual perspective rather than through attempts at generalization (Gotts, Mose, & Donnerstein, 1995). Finally, there are those who argue that what matters is a woman's perception that she suffers from PMS, rather than whether or not she meets arbitrary scientific criteria (Corney & Stanton, 1991). As many as 80% of women who seek help from specialist PMS clinics do not confirm their self-diagnosis after completing daily ratings for two months (Metcalf, Livesey, & Braiden, 1989; Schnurr, Hurt, & Stout, 1994), whereas community studies show that many women who do meet scientific criteria for PMS do not identify themselves as having a problem or seek help for it (Dan & Monagle, 1994; Waner & Bancroft, 1990).

8.15.5.2 Epidemiology and Risk Factors

Controversy about the definition of PMS inevitably means that epidemiological studies are difficult to interpret. The majority of epidemiological studies have assessed the incidence of particular experiences around menstruation (e.g., irritability, swelling), rather than a syndrome. Crosscultural studies suggest that the physical and emotional experiences reported by women premenstrually in other cultures are similar to those reported in Britain or North America, but that the number of women reporting particular experiences varies (Dank & Monagle, 1994; Longue & Moos, 1986). Epidemiological studies in Scandinavia (Andersch, Wenderstam, Hahn, & Öhman, 1986) and North America (Woods, Most, & Dary,

1982) suggest that approximately 5–10% of women between age 15 and 50 regularly experience one or more moderate or severe premenstrual symptoms. However, Ramcharan, Love, Fick, and Goldfein (1992), in a large-scale cross-sectional study in North America, found no evidence to suggest that depression or anxiety are more frequent premenstrually than at any other time, and estimated a 1% rate of severe premenstrual dysphoria in the general population.

A large number of women identify themselves as PMS sufferers, however, and seek help with their symptoms. Warner and Bancroft (1990) investigated the factors related to self-diagnosis among 5457 respondents to a British magazine survey. Of these women, 61% described themselves as having PMS. The factors that best predicted self-diagnosis were the numbers of symptoms rated as severe (especially emotional symptoms), the number of years of natural cycles (since the last interruption by pregnancy, oral contraceptive use, etc.) and psychosocial factors (perceived stress, cohabitation, self-rated happiness of relationship, and type of employment). This supports the finding of Logue and Moos (1986) that women who report PMS are more likely to be over 30, have long and/or heavy periods, a history of affective illness, and high life stress. Other factors such as personality, social class, cycle length, contraceptive use, and general health are not consistently related to symptom reporting, although there is some evidence that women with high levels of trait negative affectivity may be more likely to seek help for PMS (Von den Akkov, Sharifian, Packer, & Eves, 1995).

8.15.5.3 Explanations for Premenstrual Syndrome

There are many theories about the causes of PMS, ranging from physiological to social constructionist (Bancroft, 1995; Walker, 1995). Researchers have investigated a range of possible physiological mechanisms—from ovarian hormones to neurotransmitters (Bancroft, 1993; Dinan & O'Keare, 1991; Severino, 1993; Walker, 1992), none of which has been well supported. For example, although premenstrual experiences occur in the luteal phase of ovulatory cycles, when progesterone is present, and not the follicular phase, when it is absent, suggesting a simple causal relationship, studies that have manipulated progesterone levels, either by suppressing ovulation (Bancroft, Boyle, Warner, & Fraser, 1987; Walker & Bancroft, 1990) or supplementing endogeneous hormones (Sampson, 1979) have failed to show a dose–response pattern. After reviewing the literature, Walker (1992, 1997) concludes that the timing of premenstrual emotional experiences is linked to the endocrinology of the menstrual cycle, but their intensity is not. Physical experiences, such as breast tenderness or swelling, may be more directly related to ovarian hormone levels, but even this relationship does not appear to be a simple one (Bancroft, 1993; Walker, 1997; Walker & Bancroft, 1990).

Similarly, studies of psychological factors produce inconsistent findings. Psychological theories have implicated personality (Coppen & Kessel, 1963), cyclical changes in arousal (Daly, 1935), cognitive processing (Reading, 1992) and stress (Woods, Mitchell, & Lentz, 1995), and social psychological approaches have demonstrated that social cognitions play a role in symptom perception and reporting (Asso, 1992). In addition, it is clear that PMS is a socially constructed phenomenon; that is, that whatever the experiences of women really are, the inference that large numbers of women are ill or unpredictable once a month serves a particular political purpose within postindustrialized cultures (Maddocks, Hahn, Moller, & Rerd, 1986; Martin, 1989). Integrated and biopsychosocial models of PMS are currently the most popular in both medicine and psychology. For example, Bancroft (1993, 1995) hypothesizes that PMS results from the interaction between a timing factor, related to the hormonal fluctuations of the cycle, a menstruation factor, involving processes leading up to menstruation, and a vulnerability factor, characteristics of the woman which make her vulnerable (such as personality, or a predisposition to depression or stress). Others see PMS as the result of interactions between physiological, psychological, socioeconomic, and cultural processes (Severino, 1993; Ussher, 1992). The interactions between these systems are proposed to be so complex and intricate that PMS is effectively an idiosyncratic syndrome with different causes and different symptoms in different women (Gotts et al., 1995).

8.15.5.4 Biomedical Approaches to Treatment

The various theories of PMS imply different approaches to treatment. The best known treatment is progesterone, which was popularized and marketed in the 1970s and 1980s. Although individual women may find progesterone helpful, well controlled trials have shown that it is no more effective than a placebo (Maddocks et al., 1986; Sampson, 1979). Recent reviews of PMS management in the medical literature emphasize the importance of thorough investigation and assessment of symptoms and then recommend three broad treatment strategies (Leigh, 1994; Severino & Moline, 1995): changes in diet and lifestyle ((e.g., dietary supplements (Carter & Verhoef, 1994), exercise)); pharmacological management of specific symptoms ((e.g., benzodiazepines or selective serotonin reuptake inhibitors for dysphoric symptoms (Severino & Moline, 1995; Steiner et al., 1995)); or disruption control of the menstrual cycle, either medically (Leigh, 1994; Magos, Briheat, & Studd, 1986) or surgically (Casson, Hahn, van Vugt, & Reid, 1990).

Rivera-Tovar, Rhodes, Pearlstein, and Frank (1994) have reviewed the efficacy of 31 different surgical, pharmacological, dietary, and psychosocial treatments for PMS, finding that many of these are effective for at least some premenstrual symptoms, but none can abolish PMS without simultaneously abolishing the menstrual cycle. In addition, all of the surgical and pharmacological treatments have associated risks and side effects (Mortola, 1994). Careful trials are essential in PMS research because of the high initial placebo effect, which is usually greater than 50%, and can be as high as 94% (Magos et al., 1986).

8.15.5.5 Psychological Interventions for Premenstrual Syndrome

A biopsychosocial or cognitive model of PMS suggests that psychological interventions may be useful. A number of studies have now been conducted of a variety of psychological interventions, including support groups (Walton & Youngkin, 1987), relaxation (Goodale, Domar, & Benson, 1990), and cognitive-behavior therapies with individuals (Corney, Stanton, Newell, & Clare, 1990; Slade, 1989) or groups (Blake, 1995; Christensen & Oei, 1995; Kirkby, 1974; Morse, Dennerstein, Farrell, & Varnavides, 1991; Staut, 1995). For example, Morse et al. (1991) randomly allocated 44 women with prospectively confirmed PMS to receive either hormone therapy (dydrogesterone), relaxation therapy (using audio tapes), or coping skills training. The coping skills training consisted of 10, weekly, 90 minute, group sessions, covering general coping skills, anxiety reduction techniques, stress innoculation, assertiveness training, and relaxation training. The women were assessed after treatment and monthly for three months, using the daily Moos Menstrual Distress Questionnaire, by someone who was unaware of the therapy they had received. Initially all three groups showed a good response to treatment; however, by three months, only the coping skills group were maintaining their treatment gains.

It is not clear whether the beneficial effects of psychological interventions relate specifically to premenstrual experiences or whether the woman is generally happier or more able to cope. Similarly, it is not clear whether some psychological interventions are more effective than others, and particularly whether the cognitive elements of therapy are necessary. Christensen and Oei (1995) for example, found that two-hour group sessions weekly for 13 consecutive weeks, of either cognitive-behavior therapy (cognitive restructuring and assertion training) or information-focused therapy (relaxation training, dietary and lifestyle advice, assertion training, and child management training), were equally effective in relieving premenstrual dysphoria (assessed by the BDI) at both six and 12 month follow-up. What is clear is that enabling women to develop their own coping strategies and skills can be as effective as the majority of current drug therapies for most women (Stout, 1995).

8.15.6 MENOPAUSE

8.15.6.1 Definition

The menopause, a woman's last menstrual period, occurs on average between the ages of 50 and 51 (typical age range 45–55) and has been defined as "permanent cessation of menstruation resulting from loss of ovarian activity" (World Health Organization [WHO], 1981). A period of gradual decreased output of estrogen by the ovaries and reduction in ovulation precedes the menopause. The most commonly used classification in medical practice and research is based on menstrual criteria, defining women as postmenopausal if no menstruation has occurred during the previous 12 months, and as perimenopausal if menstruation has become irregular but has occurred during the past 12 months (WHO, 1981). Hormonal measures (lowered estradiol $<100\,\mathrm{pm\,l^{-1}}$ and raised follicle stimulating hormone $>30\,\mathrm{iu\,l^{-1}}$) can also be used in order to confirm menopausal status. On average it takes four years from the first change in menstruation until 12 months after the last menstrual period (Kaufert & Gilbert, 1986). However, there is marked variation between women with respect to its timing and duration. Menopause can also result from surgery (oophorectomy), medical treatment, or disease and it is considered to be premature if it happens before the age of 40.

Hot flushes and night sweats, also called vasomotor symptoms, are the only definite symptoms of the menopause (Utian, 1972).

However, a wide variety of symptoms have been attributed to hormone changes during the menopause, such as irritability, headaches, depression, anxiety, weight gain, aches and pains, poor memory and concentration, and loss of libido—symptoms often included under the broad term "menopausal syndrome." With the development of hormonal treatments in the twentieth century the dominant medical model positions the menopause as an estrogen-deficiency disease, a cluster of physical and emotional symptoms that can be treated by HRT. Early studies were carried out on clinic samples, thus the link between symptomatology and the menopause was reinforced. The boundaries of the "disease" have shifted markedly over the past 20 years, HRT being first used for the treatment of hot flushes but more recently to prevent health problems in later life, such as osteoporosis and cardiovascular disease (Maxim, Ettinger, & Spitalny, 1995; Stephenson, Crook, Collins, Godsland & Whitehead, 1994).

From a woman's perspective the menopause has acquired complex social and cultural meanings, being inextricably linked to age and often coinciding with life changes. Thus, "the change of life" in popular discourse reflects the view that the menopause might be paralleled by role changes and emotional and social adaptations which may occur during midlife. Concepts such as "empty nest," physical decline, loss of femininity, loss of sexuality, involutional melancholia, and vaginal atrophy have contributed to varied, but generally negative, stereotypes of menopausal women in Western cultures. In contrast, anthropological studies provide examples of how menopause can be a positive event, particularly when it signifies a change in social status (Flint, 1975). In general, women living in non-Western societies appear to report fewer symptoms at the menopause than those living in the West (Payer, 1991). However, it is likely that material differences, such as dietary and reproductive practices and levels of exercise, contribute to cross-cultural differences, as well as cultural values and the social position of midaged and older women.

8.15.6.2 Epidemiological Studies

In the 1980s several prospective studies (Holte, 1992; Hunter, 1992; Kaufert, Gilbert, & Tate, 1992; Matthews, Wing, & Kuller, 1990; McKinlay, Brambilla, & Posner, 1992; Oldenhave, 1991) were carried out in North America and Europe, using general population samples in order to clarify what changes occurred and whether any changes were associated with stage of menopause, aging, or psychosocial factors. The studies included a cross-sectional phase, as well as following women for between three and five years across the menopause transition. Overall, the results of these studies suggest that the menopause is not associated with emotional symptoms. In four of the five studies no change in measures of depressed mood were evident. Similarly, there is no substantial evidence to support the view that psychiatric disorder is more prevalent during the menopause (Gath & Iles, 1990; Pearce, Hawton, & Blake, 1995).

Hot flushes and night sweats increased as expected, but the prevalence rates averaged around 50–60% of menopausal women, lower than the commonly quoted 70%. In one study there was a slight increase in psychological symptoms in perimenopausal women who had hot flushes and night sweats (Oldenhave, 1991), and an association between length of perimenopause and mood, which was explained by prolonged exposure to hot flushes and night sweats (Avis, Brambilla, McKinley, & Vass, 1994). Women who experienced chronic ill health, such as arthritis or thyroid problems, were also more likely to report depressed mood (Kaufert et al., 1992).

When asked, the majority of women reported relief from the cessation of menstruation and the possibility of pregnancy. However, a proportion expressed negative beliefs and expectations about the menopause beforehand. Such beliefs (that the menopause is associated with a host of physical and emotional problems) were found to predict depressed mood and other symptoms when these women reached the menopause (Hunter, 1992). Depressed mood was predicted mainly by past depression, together with low socioeconomic status, not being employed outside the home, as well as negative beliefs (Hunter, 1988). In general, psychological problems were associated with psychosocial factors to a greater extent than with stage of menopause. For example, life stress, particularly losses such as bereavements, have been found to be associated with emotional and physical symptoms, especially if they occurred in the context of additional stresses (Greene & Cooke, 1980). However, there is little evidence to suggest that the menopause is more stressful than other life stages, for example, early motherhood (Boulet, Lehert, & Riphagen, 1988). Similarly, there is neither strong support for the assumption that a child leaving home is necessarily stressful, nor for the belief that single women are more at risk during the menopause.

Overall, the epidemiological studies suggest that for the majority of women the menopause does not lead to changes in psychological or physical symptoms, apart from possible secondary effects resulting from problematic hot flushes and night sweats. Nor did it adversely affect ratings of general health or health-related behaviors such as medication intake or visits to the doctor (Hunter, 1992; McKinlay & McKinlay, 1986). Taken together with cross-cultural findings, the notion of an essentially biological cause of distress is certainly challenged by epidemiological research. When general population samples have been compared with those seeking help during the menopause a different picture emerges. Menopause clinic attenders report much higher rates of distress and more physical and emotional problems than non-attenders (Ballinger, 1985). Decisions to seek medical help may be more likely when emotional distress is experienced with hot flushes in the context of certain beliefs—in particular that the symptoms are due to a current disease-like process (Hunter, 1988).

8.15.6.3 Hot Flushes/Flashes and Night Sweats

Subjectively described in terms of sensations of heat in the face, neck, and chest, and frequently accompanied by perspiration and/or shivering, hot flushes or hot flashes are associated with increases in skin conductance and finger temperature (Tateryn et al., 1981), peripheral blood flow, and heart rate (Ginsberg, Hardman, & O'Reilly, 1989). The etiology of hot flushes is unknown but they are associated with rates of change of plasma estrogen. Surprisingly few psychological studies have been carried out given the prevalence of hot flushes.

There are considerable cultural differences in hot flush reporting, but even within Western cultures the experience of hot flushes is highly variable when site, frequency, duration, and intensity have been studied. A range of emotional reactions have been reported, including embarrassment, panic, feelings of suffocation, and precipitants have been identified such as hot drinks, alcohol, external heat, and stressful situations, but many flushing episodes are described as spontaneous (Kronenberg, 1990; Voda, 1981). Gannon, Hansel, and Goodwin (1987) monitored 10 women and found a positive correlation between hot flushes and stress (daily hassles) for half the sample.

Approximately 10–15% of women with hot flushes describe them as problematic, largely because of physical discomfort, disruption of sleep, and social embarrassment (Hunter & Liao, 1995), and hot flushes and night sweats are the main reason given for seeking medical help during the menopause. The frequency of hot flushes appears to be difficult to predict and was not associated with sociodemographic variables in a recent study. However, women who viewed their flushes as problematic tended to report higher levels of anxiety and depressed mood, and lower self-esteem (Hunter & Liao, 1995). HRT is an effective and commonly recommended treatment for hot flushes. However, in view of the possible relationship between hot flush reporting and stress, psychological interventions have been developed which include behavioral treatments (Hunter & Liao, 1995), particularly deep breathing and relaxation (Freedman & Woodward, 1992). In a recent study (Hunter & Liao, 1996) a cognitive–behavioral intervention, including relaxation, was compared to HRT and a no-treatment control condition, using a patient preference design. Of those wanting treatment, 59% of the sample ($N = 61$) chose the psychological treatment and 41% HRT. The psychological treatment includes four hourly sessions (discussion of hot flushes, monitoring precipitants, relaxation training, modification of precipitants, cognitive strategies to cope with hot flushes, and stress reduction) and was found to be as effective as HRT in reducing hot flush frequency. Those receiving the psychological treatment also reported lower anxiety ratings after treatment. Further research is needed to examine the potent factors within these treatments.

8.15.6.4 Hormone Replacement Therapy

HRT, which may be administered orally or via subcutaneous implants or skin patches, is being increasingly promoted for the prevention of osteoporosis and cardiovascular disease, as well as relief of hot flushes and, more recently, for improvements in quality of life. HRT has also been advocated as a treatment for depression. However, correlational studies do not support an association between estrogen levels and mood (Alder, Bancroft, & Livingstone, 1992), and treatment studies have failed to provide conclusive evidence that HRT alleviates depressed mood over and above placebo effects and the secondary relief from hot flushes (Hunter, 1990; Pearce et al., 1995).

In order to derive long-term health benefits from HRT, at least five years' use is generally recommended (Coope, 1989). In medical literature compliance with HRT is regarded as low in terms of both uptake and adherence rates, particularly in the UK where approximately 10% of 45–55-year-old women currently use HRT. Reasons for low compliance include not wanting to take medication, having menstrual periods, fears of risks, and side effects. It has been assumed that, with appropriate information and reassurance, uptake and adherence would increase (Sinclair, Bond, & Taylor, 1993). However, an educational model that neglects health beliefs is likely to be overly simplistic. In general, surveys of women's and doctors attitudes to the menopause have found that women are more likely than doctors to see the menopause as a normal developmental phase of life, whereas doctors are more likely to see it as a disease-like process (Leiblum & Swartzman, 1986). One study of women's attitudes found that promotion of HRT was considered of low priority compared to other health concerns (Griffiths, 1995). In a British study women were interviewed about the menopause and HRT and their responses analyzed qualitatively. They tended to prefer not to take medication for what they regarded as a natural process, unless they had troublesome current symptoms, that is, hot flushes, and thus appeared, in general, to be using different criteria from health professionals (who tend to view HRT as a long-term health measure for the menopause—a clinical problem) when making decisions about HRT use (Hunter, O'Dea, & Britten, 1997). This may partly explain why many women are dissatisfied with medical consultations about the menopause and HRT (Hibbard & Sampson, 1993). Many women express uncertainty about HRT, which points to a need for balanced information and the opportunity to discuss the relative costs and benefits and possible health risks, as well as the woman's values and preferences. Much of the existing knowledge base draws upon cohort studies of postmenopausal women electing to use HRT, resulting in selected samples and possible underestimates of the risks of breast cancer and overestimates of the prevention of fractures (Hemminki & Sihvo, 1993). Long term prospective randomized trials are currently in progress.

8.15.6.5 Psychological Interventions

Few psychological interventions have been developed for menopausal women. Apart from research on the treatment of hot flushes, described above, preliminary studies have been carried out focusing on prevention, ways of enhancing doctor–patient communication, and dealing with emotional problems reported by women who seek help. For example, the effects of providing 45-year-old women with balanced information, the opportunity to discuss beliefs and expectations, and health education is currently being evaluated (Liao & Hunter, 1994). Research is also being developed to evaluate ways to empower women in order to facilitate decision making during consultations, and to overcome barriers such as lack of information and stigma of the menopause, differing agendas, gender differences, and power relationships between women and health professionals (Hampson & Hibbard, 1996; Rothbert et al., 1994). Psychological therapies are likely to be helpful for women who attend clinics with emotional and physical problems. Individual (Green & Hart, 1987) or group (Hunter & Liao, 1995b) therapy can help women to clarify the causes of distress in their lives and seek appropriate solutions.

8.15.7 CLINICAL PSYCHOLOGIST'S ROLE

As outlined in the preceding sections, changes are needed at the level of training of health professionals, in describing and communicating about gynecological and reproductive problems, and in assessing and treating individuals, couples, and groups.

In practice the clinical health psychologist's role might include:

(i) Direct work with women, including assessments, clarification of problems, individual, couple, and group therapy.

(ii) Services to health professionals working in gynecology, such as training in communication skills, teaching about the application of a biopsychosocial model and psychological interventions, offering staff support and facilitating team development, consultation with other agencies on the management of gynecological problems, and carrying out clinically relevant research.

(iii) Services to purchasers and planners, such as advising about the psychological needs in gynecological settings, developing service evaluation and audit projects, evaluation of psychological interventions, and development of health promotion projects.

As an often scarce resource in these settings, clinical psychologists tend to work in a facilitating role, working with other health professionals (British Psychological Society, 1995). In order to provide such services clinical health psychologists should have access to supervision and research facilities and be supported by a clinical health psychology service.

8.15.8 CONCLUSION

Psychological research in gynecology clearly draws attention to the psychosocial and cultural influences upon women's experience of gynecological problems. Similar issues are faced by women seeking help across the lifespan, for example, how to make sense of complex biological and social changes that impact upon her experience of herself and her body. Cultural factors serve as a filter, influencing the meaning of bodily changes, as well as defining social traditions and practices. A biopsychosocial model that does not privilege biology over and above cultural and psychosocial factors is a useful starting point for clinical assessment, treatment, teaching, and research. Such a model might lessen the likelihood of dualistic thinking and polarized views about women's problems, and open up the possibility of developing and offering women a broader range of treatments. Clinical health psychologists working with gynecologists in clinics is one way to provide more integrated services and a more holistic approach (British Psychological Society, 1995).

Psychological research has also provided information about normal developmental processes and individual differences in the experience of reproductive events. Appreciation of these factors may help to counter stereotypes about women which have tended to be based upon clinic samples. Examples are provided in the above sections that challenge myths about women seeking help—as neurotic, hysterical, unfeminine, or ruled by their hormones. Stereotyping and labeling, as well as differing understandings or illness representations between doctor and patient, can lead to dissatisfaction with communication and inappropriate treatment decisions in clinical practice. Communication and decision making are increasingly important concerns for clinical health psychologists working in gynecological settings (Hunter, 1994), since women are required to make extremely complex choices about their health. Further research is needed to develop optimum procedures that facilitate informed decision making. An additional area for further study is the nature of help-seeking behavior, especially given the high levels of emotional problems reported by women attending gynecological clinics. For example, why do some women who are distressed seek help from a PMS clinic, when others with similar symptoms might seek help from a counselor or psychologist?

The value of psychological interventions has been demonstrated in many areas of gynecology, such as preparation for surgery, alleviation of pelvic pain, and is promising in the areas of PMS, menopause, and menstrual disorders. There appears to be a gap between research findings and clinical applications of these findings, particularly in the case of preparation for surgical interventions. Future research could usefully examine optimum methods of service delivery in hospital settings.

8.15.9 REFERENCES

Ainscough, C. E. (1990). Premenstrual emotional changes: a prospective study of symptomatology in normal women. *Journal of Psychosomatic Research, 34,* 35–45.

Alder, E. M., Bancroft, J., & Livingstone, J. (1992). Estradiol implants, hormone levels and reported symptoms. *Journal of Psychosomatic Obstetrics & Gynaecology, 13,* 223–235.

Alexander, D. A., Naji, A. A., Pinion, S. B., Mollison, J., Kitchener, H. C., Parkin, D. E., Abramovich, D. R., & Russell, I. T. (1996). Randomised trial comparing hysterectomy with endometrial ablation for dysfunctional uterine bleeding: psychiatric and psychosocial aspects. *British Medical Journal, 312,* 280–284.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Andersch, B., Wendestam, C., Hahn, L., & Öhman, R. (1986). Premenstrual complaints. I. Prevalence of premenstrual symptoms in a Swedish urban population. *Journal of Pychosomatic Obstetrics and Gynaecology, 5,* 39–50.

Anderson, A., & McPherson, A. (1983). Menstrual problems. In A. Anderson & A. McPherson (Eds.), *Women's problems in general practice* (pp. 19–40). Oxford, UK: Oxford University Press.

Areskog-Wijma, B. (1987). The gynaecological examination—women's experiences and preferences and the role of the gynaecologist. *Journal of Psychosomatic Obstetrics and Gynaecology, 6,* 59–69.

Avis, N. E., Brambilla, D., McKinlay, S. M., & Vass, K. (1994). A longitudinal analysis of the association between menopause and depression. Results from the Massachusetts Womens Health Study. *Annals of Epidemiology, 4,* 214–220.

Bak, A., Verhage, F., Drogendijk, A., Voitus van Hamme, J., & Duivenoorden, A. (1989). Sexual problems of women with chronic pelvic pain. *Journal of Psychosomatic Obstetrics and Gynaecology, 10,* 157–162.

Ballinger, C. B. (1977). Psychiatric morbidity and the menopause: survey of a gynaecological out-patient clinic. *British Journal of Psychiatry, 131,* 83–89.

Ballinger, S. E. (1985). Psychosocial stress and symptoms of menopause: a comparative study of menopause clinic patients and non-patients. *Maturitas, 7,* 315–327.

Bancroft, J. (1993). The premenstrual syndrome—a reappraisal of the concept and the evidence. *Psychological Medicine,* Suppl 24, 1–47.

Bancroft, J. (1995). The menstrual cycle and the well-being of women. *Social Science and Medicine, 41,* 785–791.

Bancroft, J., Boyle, H., Warner, P., & Fraser, H. M. (1987). The use of an LHRH agonist—buserilin, in the long term management of premenstrual syndromes. *Clinical Endocrinology, 27,* 171–182.

Bancroft, J., Williamson, L., Warner, P., Rennie, D., & Smith, S. K. (1993). Perimenstrual complaints in women complaining of PMS, menorrhagia and dysmenorrhea: Toward a dismantling of the premenstrual syndrome. *Psychosomatic Medicine, 55,* 133–145.

Barker, M. (1968). Psychiatric illness after hysterectomy. *British Medical Journal, ii,* 91–95.

Beard, R. (1984). Preface. In A. Broome & L. Wallace (Eds.), *Psychology and gynaecological problems* (pp. ix–xii). London: Tavistock.

Benassi, L., Bertani, D., & Avanzini, A. (1993). An attempt at real prophylaxis of primary dysmenorrhea: comparison between meclofenamate sodium and naproxen sodium. *Clinical and Experimental Obstetrics and Gynecology, 20,* 102–107.

Bernhard, L. (1992). Consequences of hysterectomy in the lives of women. *Health Care for Women International, 13,* 281–291.

Blake, F. (1995). Cognitive therapy for PMS. *Cognitive and Behavioural Practice, 2,* 167–185.

Boulet, M., Lehert, P. H., & Riphagen, F. E. (1988). The menopause viewed in relation to other life events—a study performed in Belgium. *Maturitas, 10,* 333–342.

British Psychological Society (1995). *Purchasing clinical psychology services in obstetrics & gynaecology: Briefing paper*. Leicester; UK: Author.

Budeiri, D. J., Li Wan Po, A., & Dornan, J. C. (1994). Clinical trials of treatments of premenstrual syndrome: Entry criteria and scales for measuring treatment outcomes. *British Journal of Obstetrics & Gynaecology, 101,* 689–695.

Burns, J. (1992). The psychology of lesbian health care. In P. Nicolson & J. Ussher (Eds.), *The psychology of women's health and health care* (pp. 225–248). London: Macmillan.

Byrne, P. (1984). Psychiatric morbidity in a gynaecology clinic: An epidemiological survey. *British Journal of Psychiatry, 144,* 28–34.

Caplan, P., McCurdy-Myers, J., & Gans, M. (1992). Should "premenstrual syndrome" be called a psychiatric abnormality? *Feminism and Psychology, 2,* 27–44.

Carter, J., & Verhoef, M. J. (1994). Efficacy of self-help and alternative treatments of premenstrual syndrome. *Women's Health Issues, 4,* 130–137.

Casson, P., Hahn, P. M., van Vugt, D. A., & Reid, R. (1990). Lasting response to ovariectomy in severe intractable premenstrual syndrome. *American Journal of Obstetrics and Gynecology, 162,* 99–105.

Chapple, A. (1995). Hysterectomy: British National Health Service and private patients have very different experiences. *Journal of Advanced Nursing, 22,* 900–906.

Cheng, M., Wilansky, D. Kung, R., Shime, J. K., & Hannah, M., (1995). Menses cup evaluation study. *Fertility and Sterility, 64,* 661–663.

Christensen, A., & Oei, T. (1995). The efficacy of cognitive behaviour therapy in treating premenstrual dysphoric changes. *Journal of Affective Disorders, 33,* 57–63.

Cockburn, J., & Bewley, S. (1996). Do patients prefer women doctors? *British Journal of Obstetrics and Gynaecology, 103,* 2–3.

Coope, J. (1989). *Hormone replacement therapy*. London: RCGP.

Coppen, A., & Kessel, N. (1963). Menstruation and personality. *British Journal of Psychiatry, 109,* 711–721.

Corney, R., & Stanton, R. (1991). A survey of 658 women who report symptoms of premenstrual syndrome. *Journal of Psychosomatic Research, 35,* 471–482.

Corney, R. H., Stanton, R., Newell, R., & Clare, A. W. (1990). Comparison of progesterone, placebo and behavioural psychotherapy in the treatment of premenstrual syndrome. *Journal of Psychosomatic Obstetrics and Gynaecology, 11,* 211–220.

Dalton, K. (1977). *The premenstrual syndrome and progesterone therapy*. New York: Heinemann.

Daly, C. D. (1935). The menstruation complex in literature. *Psychoanalytic Quarterly, 4,* 307–340.

Daly, M. (1978). *Gyn/Ecology: The metaethics of radical feminism*. Boston: Beacon Press.

Dan, A., & Monagle, L. (1994). Sociocultural influences on women's experiences of perimenstrual symptoms. In J. Gold & S. Severino (Eds.), *Premenstrual Dysphorias: Myths and Realities* (pp. 201–211). Washington, DC: American Psychiatric Press.

Davies, L., Gangar, K., Drummond, M., Saunders, D., & Beard, R. (1992). The economic burden of intractable gynaecological pain. *Journal of Obstetrics & Gynaecology, 12*(Suppl. 2), 554–556.

Deutch, B. (1995). Menstrual pain in Danish women correlated with low n-3 polyunsaturated fatty acid intake. *European Journal of Clinical Nutrition, 49,* 508–516.

Dinan, T. G., & O'Keane, V. (1991). The premenstrual syndrome: A psychoneuroendocrine perspective. *Balliéres Clinical Endocrinology and Metabolism, 5,* 143–165.

Dingwall, R. (1992). "Don't mind him—he's from Barcelona." Qualitative methods in health studies. In J. Daly, I. McDonald, & E. Willis (Eds.), *Researching health care: Designs, dilemmas, disciplines* (pp. 161–175). London: Tavistock/Routledge.

Ehrenreich, B., & English, D. (1973). *Complaints and disorders: The sexual politics of sickness*. London: Writers and Readers Publishing Co-operative.

Evans, C., & Richardson, P. (1988, August). Improved recovery and reduced post-operative stay after therapeutic suggestion during general anaesthesia. *Lancet, ii,* 491–493.

Flint, M. (1975). The menopause: Reward or punishment. *Psychosomatics, 16,* 161–163.

Frank, R. T. (1931). The hormonal causes of premenstrual tension. *Archives of Neurology and Psychiatry, 26,* 1053–1057.

Fraser, I. (1994). Menorrhagia — a pragmatic approach to the understanding of causes and the need for investigations. *British Journal of Obstetrics & Gynecology, 101* (Suppl. 11), 3–7.

Gannon, K. (1996). Andrology: Men and reproduction. In C. A. Niven & A. Walker (Eds.), *The psychology of reproduction. 1. Reproductive potential and fertility control* (pp. 92–108). Oxford, UK: Butterworth-Heinemann.

Gannon, L. (1985). *Menstrual disorders and menopause*. New York: Praeger.

Gath, D., Cooper, P., & Day, A. (1982). Hysterectomy and psychiatric disorder. I: levels of psychiatric morbidity before and after hysterectomy. *British Journal of Psychiatry, 140,* 335–342.

Gath, D., Rose, N., Bond, A., Day, A., Garrod, A., & Hodges, S. (1995). Hysterectomy and psychiatric disorders: are the levels of psychiatric morbidity falling? *Psychological Medicine, 25,* 277–283.

Gise, L. (1988). *The Premenstrual Syndromes*. New York: Churchill Livingstone.

Gise, L. H., Lebovits, A. H., Paddison, P. L., & Strain, J. J. (1990). Issues in the identification of premenstrual syndromes. *Journal of Nervous & Mental Disease, 178,* 228–234.

Gold, J., & Severino, S. (Eds.) (1994). *Premenstrual dysphorias: Myths and realities*. Washington, DC: American Psychiatric Press.

Goldberg, D. (1972). *The Detection of Psychiatric Illness by Questionnaire*. London: Oxford University Press.

Golub, S. (1992). *Periods: From Menarche to Menopause*. Newbury Park, CA: Sage.

Goodale, L. L., Domar, A. D., & Benson, H. (1990). Alleviation of premenstrual syndrome symptoms with the relaxation response. *Obstetrics & Gynaecology, 75,* 649–665.

Gotts, G., Morse, C. A., & Dennerstein, L. (1995). Premenstrual complaints: an idiosyncratic syndrome. *Journal of Psychosomatic Obstetrics and Gynaecology, 16,* 29–35.

Goudsmit, E. M. (1994). All in her mind! Stereotypic views and the psychologisation of women's illness. In S. Wilkinson & C. Kitzinger (Eds.), *Women and health: Feminist perspectives* (pp. 7–12). London: Taylor Francis.

Grace, V. (1995a). Problems of communication, diagnosis, and treatment experienced by women using the New Zealand health services for chronic pelvic pain: A quantitative analysis. *Health Care for Women International, 16,* 521–535.

Grace, V. (1995b). Problems women patients experience in the medical encounter for chronic pelvic pain: A New Zealand study. *Health Care for Women International, 15,* 509–519.

Greenberg, M. (1983). The meaning of menorrhagia: an investigation into the association between the complaint of menorrhagia and depression. *Journal of Psychosomatic Research, 27,* 209–214.

Freedman, R. R., & Woodward, S. (1992). Behavioural treatment of menopausal hot flushes: Evaluation by ambulatory monitoring. *American Journal of Obstetrics and Gynaecology, 167,* 436–439.

Gannon, L., Hansel, S., & Goodwin, J. (1987). Correlates of menopausal hot flushes. *Journal of Behavioural Medicine, 10,* 277–285.

Gath, D., & Iles, S. (1990). Depression and the menopause. *British Medical Journal, 300,* 1287–1288.

Ginsburg, J., Hardman, P., & O'Reilly, B. (1989). Peripheral blood flow in menopausal women who have hot flushes and those who do not. *British Medical Journal, 298,* 1488–1490.

Greene, J. G., & Cooke, D. J. (1980). Life stress and symptoms at the climacteric. *British Journal of Psychiatry, 136,* 2186–2191.

Greene, J. G., & Hart, D. M. (1987). The evaluation of a psychological treatment programme for climacteric women. *Maturitas, 9,* 41–48.

Greene, R., & Dalton, K. (1953). The premenstrual syndrome. *British Medical Journal, 1,* 1007–1014.

Griffiths, F. (1995). Women's health concerns. Is the promotion of hormone replacement therapy for prevention important to women? *Family Practice, 12,* 54–59.

Gross, R., Doerr, H., Caldirola, D., Guzinski, G., & Ripley, H. (1980–81). Borderline syndrome and incest in chronic pelvic pain patients. *International Journal of Psychiatry in Medicine, 10,* 79–96.

Hampson, S. E., & Hibbard, J. H. (1996). Cross-talk about the menopause: Enhancing provider—patient interactions about the menopause and hormone therapy. *Patient Education & counseling, 27,* 177–184.

Harel, Z., Biro, F. M., Kottenhahn, R. K., & Rosenthal, S. L. (1996). Supplementation with omega-3 polyunsaturated fatty acids in the management of dysmenorrhea in adolescents. *American Journal of Obstetrics & Gynecology, 174,* 1335–1338.

Harris, T. (1989). Disorders of menstruation. In G. Brown & T. Harris (Eds.), *Life events and illness* (pp. 261–294). London: Unwin Hyman.

Helman, C. (1990). *Culture, health and illness* (2nd ed.). London: Wright.

Helstrom, L., Lundberg, P. O., Sorbom, D., & Backstrom, T. (1993). Sexuality after hysterectomy: a factor analysis of women's sexual lives before and after subtotal hysterectomy. *Obstetrics & Gynecology, 81,* 357–362.

Hemminki, E., & Sihvo, S. (1993). A review of postmenopausal hormone therapy recommendations: Potential for selection bias. *Obstetrics & Gynaecology, 82,* 1021–1028.

Hibbard, J. H., & Sampson, S. E. (1993). Enhancing women's partnership with health providers in hormone replacement therapy decision making: Research and practice directions. *Journal of Women and Ageing, 5,* 17–29.

Hodgkiss, A. D., & Watson, J. P. (1994). Psychiatric morbidity and illness behaviour in women with chronic pelvic pain. *Journal of Psychosomatic Research, 38,* 3–9.

Holte, A. (1992). Influences of natural menopause on health complaints: A prospective study of healthy Norwegian women. *Maturitas, 14,* 127–141.

Horney, K. (1931). Die prämenstruellen Verstimmungen. *Zeitschrift für Psychoanalytische Pädagogik, 5,* 161–167.

Houvinen, K. (1993). Gynecological problems of mentally retarded women: A case-control study from southern Finland. *Acta Obstetrica et Gynecologica Scandinavica, 72,* 475–480.

Hunter, M. S. (1990). Emotional well-being, sexual behaviour and hormone replacement therapy. *Maturitas, 12,* 299–314.

Hunter, M. S. (1988). *Psychological and somatic experience of the climacteric and postmenopause: Predicting individual differences and helpseeking behaviour*. University of London, UK: unpublished PhD thesis.

Hunter, M. S. (1992). The S.E. England longitudinal study of the climacteric and postmenopause. *Maturitas, 14,* 117–126.

Hunter, M. S. (1994). *Counseling in obstetrics and gynaecology*. Leicester, UK: British Psychological Society.

Hunter, M. S., & Liao, K. L. M. (1995a). A psychological analysis of menopausal hot flushes. *British Journal of Clinical Psychology, 34,* 589–599.

Hunter, M. S., & Liao, K. L. M. (1995b). Problem-solving groups for midaged women in general practice: A pilot study. *Journal of Reproductive and Infant Psychology, 13,* 147–151.

Hunter, M. S., & Liao, K. L. M. (1996). Evaluation of a four-session cognitive-behavioural intervention for menopausal hot flushes. *British Journal of Health Psychology, 1,* 113–125.

Hunter, M. S., O'Dea, I., & Britten, N. (1997). Decision-making and hormone replacement therapy: A qualitative analysis. *Social Science and Medicine, 45*(10), 1541–1548.

Ivins, J., & Kent, G. (1993). Women's preferences for male or female gynaecologists. *Journal of Reproductive and Infant Psychology, 11,* 209–214.

Janssen, C. A. H., Scholten, P. C., & Heintz, A. P. M. (1995). A simple visual assessment technique to discriminate between menorrhagia and normal menstrual blood loss. *Obstetrics & Gynecology, 85,* 977–982.

Jarrett, M., Heitkemper, M. M., & Shaver, J. F. (1995). Symptoms and self-care strategies in women with and without dysmenorrhea. *Health Care for Women International, 16,* 167–178.

Jaszmann, N. D. (1973). Epidemiology of the climacteric and postmenopause. In P. van Keep & C. Lauritzen (Eds.), *Ageing and oestrogens* (pp. 22–25). Basel, Switzerland: Karger.

Johnson, T. (1987). Premenstrual syndrome as a western culture specific disorder. *Culture, Medicine and Psychiatry,* 11, 337–356.

Johnson, S., Smith, E., & Guenther, S. (1987). Comparisons of gynaecological health care problems between lesbians and bisexual women. *Journal of Reproductive Medicine, 32,* 805–811.

Johnston, M., & Vogele, C. (1993). Benefits of psychological preparation for surgery: A meta-analysis. *Annals of Behavioural Medicine, 15,* 245–256.

Jorgensen, J., Rossignol, A. M., & Bonnlander, H. (1993). Evidence against multiple premenstrual syndromes: results of a multivariate profile analysis of premenstrual symptomatology. *Journal of Psychosomatic Research, 37,* 257–263.

Kames, L. D., Rapkin, A. J., Naliboff, B. D., Afifi, S., & Ferrer-Brechner, T. (1990). Effectiveness of an interdisciplinary pain management program for the treatment of chronic pelvic pain. *Pain, 41,* 41–46.

Kaplan, B., Peled, Y., Pardo, J., Rabinerson, D., Hirsh, M., Ovadia, J., & Neri, A. (1994). Transcutaneous electrical nerve stimulation (TENS) as a relief for dysmenorrhea. *Clinical & Experimental Obstetrics & Gynecology, 21,* 87–90.

Kaufert, P. A., & Gilbert, P. (1986). Women, menopause and medicalisation. *Culture, Medicine and Psychiatry, 10,* 7–21.

Kaufert, P. A., Gilbert, P., & Tate, R. (1992). The Manitoba Project: A re-examination of the relationship between menopause and depression. *Maturitas, 14,* 143–155.

Kirkby, R. J. (1994). Changes in premenstrual symptoms and irrational thinking following cognitive-behavioural coping skills training. *Journal of Consulting and Clinical Psychology, 62,* 1026–1032.

Kokjohn, K., Schmid, D. M., Triano, J. J., & Brennan, P. C. (1992). The effect of spinal manipulation on pain and prostaglandin levels in women with primary dysmenorrhea. *Journal of Manipulative & Physiological Therapeutics, 15,* 279–285.

Kronenberg, F. (1990). Hot flashes: epidemiology and physiology. *Annals of New York Academy of Sciences, 592,* 52–86.

Kuh, D., & Stirling, S. (1995). Socioeconomic variation in admission for diseases of female genital system and breast in a national cohort aged 15–43. *British Medical Journal, 311,* 840–843.

Laessle, R. G., Tuschl, R. J., Schweiger, U., & Pirke, K. M. (1990). Mood changes and physical complaints during the normal menstrual cycle in healthy young women. *Psychoneuroendocrinology, 15,* 131–138.

Lalonde, A. (1994). Evaluation of surgical options in menorrhagia. *British Journal of Obstetrics & Gynaecology, 101*(Suppl. 11), 8–14.

Laws, S., Hey, V., & Eagan, A. (1985). *Seeing red: The politics of premenstrual tension.* London: Hutchinson.

Lazarus, R., & Folkman, S. (1984). *Stress, appraisal and coping.* New York: Springer.

Leiblum, S. R., & Swartzman, L. C. (1986). Women's attitudes toward the menopause: An update. *Maturitas, 8,* 47–56.

Ley, P. (1988). *Communicating with patients.* London: Chapman & Hall.

Liao, K. L. M., & Hunter, M. S. (1994). The women's midlife project: An evaluation of psychological services for midaged women in general practice. *Clinical Psychology Forum, 65,* 19–22.

Logue, C. M., & Moos, R. H. (1986). Perimenstrual symptoms: prevalence and risk factors. *Psychosomatic Medicine, 48,* 388–414.

Logue, C. M., & Moos, R. H. (1988). Positive premenstrual changes: Toward a new perspective on the menstrual cycle. *Journal of Psychosomatic Research, 32,* 31–40.

Maddocks, S., Hahn, P., Moller, F., & Reid, R. L. (1986). A double-blind placebo controlled trial of progesterone vaginal suppositories in the treatment of premenstrual syndrome. *American Journal of Obstetrics and Gynecology, 154,* 573–581.

Magni, G., Andreoli, C., de Leo, D., Martinotti, G., & Rossi, C. (1986). Psychological profile of women with chronic pelvic pain. *Archives of Gynecology, 237,* 165–168.

Magni, G., Salmi, A., Deleo, D., & Ceola, A. (1984). Chronic pelvic pain and depression. *Psychopathology, 17,* 132–136.

Magos, A. L., Brincat, M., & Studd, J. W. W. (1986). Treatment of the premenstrual syndrome by subcutaneous oestradiol implants and cyclical oral norethisterone: placebo controlled study. *British Medical Journal, 292,* 1629–1633.

Marteau, T., Kidd, J., Cuddeford, L., & Walker, P. (1996). Reducing anxiety in women referred for colposcopy using an information booklet. *British Journal of Health Psychology, 1,* 181–189.

Martin, E. (1989). *The woman in the body: A cultural analysis of reproduction.* Milton Keynes, UK: Open University Press.

Matthews, K. A., Wing, R. R., & Kuller, L. H. (1990). Influences of natural menopause on psychological characteristics and symptoms of middle-aged healthy women. *Journal of Consulting & Clinical Psychology, 58,* 345–363.

Maxim, D., Ettinger, B., & Spitalny, G. M. (1995). Fracture protection provided by long term oestrogen treatment. *Osteoporosis International, 5,* 23–29.

McGowan, L., Clark-Carter, D., & Pitts, M. (1994). *Chronic pelvic pain: A meta-analytic review.* Paper presented at British Psychological Society Special Group in Health Psychology Annual Conference, University of Sheffield, UK.

McKay, W. J. S. (1901). *The history of ancient gynaecology.* London: Balliere, Tindall & Cox.

McKinlay, M., & McKinlay, J. B. (1986). Health status and health care utilization by menopausal women. In M. Notelovitz & P. A. van Keep (Eds.), *The climacteric in perspective* (pp. 59–75). Lancaster, UK: MTP Press.

McKinlay, S. M., Brambilla, D. J., & Posner, J. G. (1992). The normal menopause transition. *Maturitas, 14,* 103–115.

Mechanic, D. (1980). The experience and reporting of common physical symptoms. *Journal of Health and Social Behaviour, 21,* 146–155.

Medvei, V. C. (1982). *A short history of endocrinology.* Lancaster, UK: MTP Press.

Menage, J. (1993). Post-traumatic stress disorder in women who have undergone obstetric and/or gynaecological procedures: A consecutive series of 30 cases of PTSD. *Journal of Reproductive and Infant Psychology, 11,* 221–228.

Messing, K., Saurel-Cubizolles, M. J., Bourgine, M., & Kaminski, M. (1993). Factors associated with dysmenorrhea among workers in French poultry slaughterhouses and canneries. *Journal of Occupational Medicine, 35,* 493–500.

Metcalf, M. G., Livesey, J. H., Wells, J. E., & Braiden, V. (1989). Mood cyclicity in women with and without the premenstrual syndrome. *Journal of Psychosomatic Research, 33,* 407–418.

Miller, S. (1988). The interacting effects of coping styles and situation variables in gynecologic settings: implications for research and treatment. *Journal of Psychosomatic Obstetrics & Gynaecology, 9,* 23–34.

Miller, S., & Mangan, C. (1983). Interacting effects of information and coping style in adapting to gynecologic stress: Should the doctor tell all? *Journal of Personality and Social Psychology, 45,* 233–236.

Mitchell, E., Woods, N., & Lentz, M. (1991). Recognizing PMS when you see it: Criteria for PMS sample selection. In D. Taylor & N. Woods (Eds.), *Menstruation, health and illness* (pp. 89–102). New York: Hemisphere.

Morse, C. A., Dennerstein, L., Farrell, E., & Varnavides, K. (1991). A comparison of hormone therapy, coping skills training and relaxation for relief of premenstrual syndrome. *Journal of Behavioural Medicine, 14,* 469–489.

Mortola, J. (1994). A risk-benefit appraisal of the drugs used in the management of premenstrual syndrome. *Drug Safety, 10,* 160–169.

Mumford, E., Schlesinger, H. J., & Glass, G. V. (1982). The effects of psychological intervention for surgery and heart attacks: an analysis of the literature. *American Journal of Public Health, 72,* 141–151.

Nathorst-Boos, J., von Schoultz, N., & Carlstrom, K. (1993). Elective ovarian removal and estrogen replacement therapy—effects on sexual life, psychological well-being and androgen status. *Journal of Psychosomatic Obstetrics & Gynecology, 14,* 283–293.

Nugent, L., & Tamlyn-Leamann, K. (1992). The colposcopy experience: What do women know? *Journal of Advanced Nursing, 17,* 514–520.

O'Brien, P. (1994). Helping women with premenstrual syndrome. *British Medical Journal,* 308, 205.

Oldenhave, A. (1991). *Well-being and sexuality in the climacteric.* Geneva, Switzerland: International Health Foundation.

Oudshoorn, N. (1994). *Beyond the natural body: An archaeology of sex hormones.* London: Routledge.

Palmer, A., Tucker, S., Warren, R., & Adams, M. (1993). Understanding women's responses to treatment for cervical intra-epithelial neoplasia. *British Journal of Clinical Psychology, 32,* 101–112.

Parazzini, F., Tozzi, L., Mezzopane, R., Luchini, L., Marchini, M., & Fedele, L. (1994). Cigarette smoking, alcohol consumption and risk of primary dysmenorrhea. *Epidemiology, 5,* 469–472.

Parry, B. L., Rosenthal, N. E., Tamarkin, L., & Wehr, T.

A. (1987). Treatment of a patient with seasonal premenstrual syndrome. *American Journal of Psychiatry, 144,* 762–766.

Payer, L. (1991). The menopause in various cultures. In H. Burger & M. Boulet (Eds.), *A portrait of the menopause* (pp. 3–21). Lancaster, UK: Parthenon.

Pearce, S. (1987). The concept of psychogenic pain: a psychological investigation of women with pelvic pain. *Current Psychological Research and Reviews, 6,* 219–228.

Pearce, J., Hawton, K., & Blake, F. (1995). Psychological and sexual symptoms associated with the menopause and the effects of hormone replacement therapy: A review. *British Journal of Psychiatry, 167,* 163–173.

Pearce, S., & Beard, R. (1984). Chronic pelvic pain. In A. Broome & L. Wallace (Eds.), *Psychology and Gynaecological Problems* (pp. 95–116). London: Tavistock.

Pearce, S., Knight, C., & Beard, R. W. (1982) Pelvic pain—a common gynaecological problem. *Journal of Psychosomatic Obstetrics & Gynaecology, 1,* 12–21.

Peters, A. A. W., van Dorst, E., Jellis, B., van Zuuren, E., Hermans, J., & Trimbos, J. B. (1991). A randomized clinical trial to compare two different approaches in women with chronic pelvic pain. *Obstetrics & Gynaecology, 77,* 740–744.

Pitts, M. (1991). The medical consultation. In M. Pitts & K. Phillips (Eds.), *The psychology of health: An introduction* (pp. 49–63). London: Routledge.

Posner, T., & Vessey, M. (1988). *Presentation of cervical cancer: The patients' view*. London: King's Fund.

Profet, M. (1993) Menstruation as a defense against pathogens transported by sperm. *Quarterly Review of Biology, 68,* 335–386.

Rafferty, P., & Williams, S. (1996). Psychological aspects of gynaecological procedures. In C. A. Niven and A. Walker (Eds.), *The psychology of reproduction. 1. Reproductive potential and fertility control* (pp. 77–91). Oxford, UK: Butterworth-Heinemann.

Ramcharan, S., Love, E. J., Fick, G. H., & Goldfien, A. (1992). The epidemiology of premenstrual symptoms in a population-based sample of 2650 urban women: Attributable risk and risk factors. *Journal of Clinical Epidemiology, 45,* 377–392.

Rapkin, A. J., & Kames, L. D. (1987). The pain management approach to chronic pelvic pain. *Journal of Reproductive Medicine, 32,* 323–327.

Rapkin, A. J., Kames, L. D., Darke, L. L., Stampfler, F. M., & Naliboff, B. D. (1990). History of physical and sexual abuse in women with chronic pelvic pain. *Obstetrics & Gynecology, 76,* 92–96.

Reading, A. (1992). Cognitive model of premenstrual syndrome. *Clinical Obstetrics & Gynecology, 35,* 693–701.

Reiter, R. C. (1990). A profile of women with chronic pelvic pain. *Clinical Obstetrics & Gynaecology, 33,* 130–136.

Reiter, R. C., & Gambone, J. C. (1990). Demographic and historic variables in women with idiopathic chronic pelvic pain. *Obstetrics & Gynaecology, 75,* 428–432.

Reiter, R. C., Gambone, J. C., & Johnson, S. R. (1991). Availability of a multidisciplinary pelvic pain clinic and frequency of hysterectomy for pelvic pain. *Journal of Psychosomatic Obstetrics and Gynaecology, 12,* 109–16.

Ridgeway, V. & Mathews, A. (1982). Psychological preparation for surgery: A comparison of methods. *British Journal of Clinical Psychology, 21,* 271–280.

Rittenhouse, C. A. (1991). The emergence of premenstrual syndrome as a social problem. *Social Problems, 38,* 412–425.

Rivera-Tovar, A., Rhodes, R., Pearlstein, T., & Frank, E. (1994). Treatment efficacy. In J. Gold & S. Severino (Eds.), *Premenstrual dysphorias: Myths and realities* (pp. 99–148). Washington, DC: American Psychiatric Press.

Robinson, J. C., Plichta, S., Weisman, C. S., Nathanson, C. A., & Ensminger, M. (1992). Dysmenorrhea and use of oral contraceptives in adolescent women attending a family planning clinic. *American Journal of Obstetrics & Gynecology, 166,* 578–583.

Rodin, M. (1992). The social construction of premenstrual syndrome. *Social Science & Medicine, 35,* 49–56.

Rosenthal, R. H., Ling, F. W., Rosenthal, T. L., & Stovall, T. G. (1991). Assessing chronic pelvic pain: clinical use of the Eysenck Personality Questionnaire. *Journal of Psychosomatic Obstetrics and Gynaecology, 12*(Suppl.), 31–38.

Rosser, S. V. (1994). *Women's health—missing from US medicine*. Bloomington, IN: Indiana University Press.

Rothert, M., Padonu, G., Holmes-Rovner, M., Kroll, J., Talarczyk, G., Rovner, D., Schmitt, N., & Breer, L. (1994). Menopausal women as decision makers in health care. *Experimental Gerontology, 29,* 463–468.

Ruble, D., & Brooks-Gunn, J. (1979). Menstrual symptoms: A social cognition analysis. *Journal of Behavioural Medicine, 2,* 171–194.

Ryan, M., Dennerstein, L., & Pepperell, R. (1989). Psychological aspects of hysterectomy: a prospective study. *British Journal of Psychiatry, 154,* 516–522.

Sampson, G. (1979). PMS: A double-blind controlled trial of progesterone and placebo. *British Journal of Psychiatry, 135,* 209–215.

Scambler, A., & Scambler, G. (1985). Menstrual symptoms, attitudes and consulting behaviour. *Social Science and Medicine, 20,* 1065–1068

Scambler, A., & Scambler, G. (1993). *Menstrual Disorders.* London: Routledge.

Schei, B. (1991). Sexual factors in pelvic pain—a study of women living in physically abusive relationships and of randomly selected controls. *Journal of Psychosomatic Obstetrics and Gynaecology, 12,* 99–108.

Schnurr, P., Hurt, S., & Stout, A. (1994). Consequences of methodological decisions in the diagnosis of late luteal phase dysphoric disorder. In J. Gold & S. Severino (Eds.), *Premenstrual Dysphorias: Myths and Realities* (pp. 19–40). Washington, DC: American Psychiatric Press.

Severino, S. (1993). Late luteal phase dysphoric disorder: A scientific puzzle. *Medical Hypotheses, 41,* 229–234.

Severino, S., & Moline, M. (1995). Premenstrual syndrome: Identification and management. *Drugs, 49,* 71–82.

Shipley, R., Butt, J., Horwitz, B., & Farby, J. (1978). Preparation for a stressful medical procedure: Effect of stimulus pre-exposure and coping style. *Journal of Consulting and Clinical Psychology, 46,* 499–507.

Showalter, E. (1987). *The female malady: Women, madness and English culture, 1830–1980*. London: Virago.

Sinclair, H. K., Bond, C. M., & Taylor, R. J. (1993). Hormone replacement therapy: a study of women's knowledge and attitudes. *British Journal of General Practice, 43,* 365–370.

Slade, P. (1989). Psychological therapy for premenstrual emotional symptoms. *Behavioural Psychotherapy, 17,* 135–150.

Smith, S. (1986). Aetiology of menstrual disorders. *Prostaglandin Perspectives, 2,* 8–9.

Smith, S. K. (1996). Gynaecology—medical or surgical. A more medical approach could reduce inappropriate surgical procedures. *British Medical Journal, 312,* 592–593.

Smith-Rosenberg, C., & Rosenberg, C. (1973). The female animal: Medical and biological views of woman and her role in the nineteenth century. *Journal of American History, 60,* 332–356.

Snowden, R., & Christian, B. (1983). *Patterns and Perceptions of Menstruation.* Beckenham, UK: Croom Helm.

Steiner, M., Steinberg, S., Stewart, D., Carter, D., Berger, C., Reid, R., Grover, D., & Streiner, D. (1995). Fluoxetine in the treatment of premenstrual dysphoria. *New England Journal of Medicine, 332,* 1529–1534.

Stephenson, J. C., Crook, D., Godsland, I. F., Collins, P., & Whitehead, M. I. (1994). Hormone replacement therapy and the cardiovascular system nonlipid effects. *Drugs, 47,* 35–41.

Stevenson, D. W., & Delprato, D. J. (1983). Multiple component self-control programme for menopausal hot flashes. *Journal of Behaviour Therapy and Experimental Psychiatry, 14,* 137–140.

Stewart, D. E. (1989). Positive changes in the premenstrual period. *Acta Psychiatrica Scandinavica, 79,* 400–405.

Stewart, M., & Roter, D. (1989). *Communicating with medical patients*. London: Sage.

Stotland, N. L., & Harwood, B. (1994). Social, political and legal considerations. In J. Gold & S. Severino (Eds.), *Premenstrual dysphorias: Myths and realities* (pp. 185–200). Washington, DC: American Psychiatric Press.

Stout, A. (1995). Cognitive-behavioural treatment of premenstrual syndrome and chronic gynecologic pain. *Depression, 3,* 60–65.

Stout, A. L., Steege, J. F., Dodson, W. C., & Hughes, C. L. (1991). Relationship of laparoscopic findings to self-report of pelvic pain. *American Journal of Obstetrics & Gynaecology, 164,* 73–79.

Studd, J. W. W., & Smith, R. N. J. (1993). Oestradiol and testosterone implants in menopause. In H. Burger (Ed.), *The menopause* (pp. 203–224). London: Balliere Tindall.

Tateryn, I. V., Lomax, P., Meldrum, D. R., Bajorek, J. G., Chesarek, W., & Hudd, H. L. (1981). Objective techniques for the assessment of postmenopausal hot flashes. *Obstetrics and Gynaecology, 57,* 340–344.

Teo, P. (1990). Hysterectomy: a change of trend or a change of heart? In H. Roberts (Ed.), *Women's health counts* (pp. 113–146). London: Routledge.

Treloar, A., Boynton, R., Benn, B., & Brown, B. (1967). Variation of the human menstrual cycle through reproductive life. *International Journal of Fertility, 12,* 77–126.

Ussher, J. (1991). *Women's madness: Misogyny or mental illness*? London: Routledge.

Ussher, J. (1992) Research and theory related to female reproduction: Implications for clinical psychology. *British Journal of Clinical Psychology, 31,* 129–151.

Utian, W. (1972). The mental tonic effect of oestrogens administered to oophorectomized females. *South African Medical Journal, 46,* 1079–1082.

van den Akker, O., Sharifian, N., Packer, A., & Eves, F. (1995). Contribution of generalized negative affect to elevated menstrual cycle symptom reporting. *Health Care for Women International, 16,* 263–272.

Vercellini, P., Fedele, L., Acaini, L., Bianchi, S., Rognomi, M. T., & Candiani, G. B. (1989). Laparoscopy in the diagnosis of chronic pelvic pain in adolescent women. *Journal of Reproductive Medicine, 34,* 827–830.

Voda, A. M. (1981) Climacteric hot flash. *Maturitas, 3,* 73–90.

Walker, A. (1992). Premenstrual symptoms and ovarian hormones: A review. *Journal of Reproductive and Infant Psychology, 10,* 67–82.

Walker, A. (1994). Mood and well-being in consecutive menstrual cycles: methodological and theoretical implications. *Psychology of Women Quarterly, 18,* 271–290.

Walker, A. (1995). Theory and methodology in premenstrual syndrome research. *Social Science and Medicine, 41,* 793–800.

Walker, A., & Bancroft, J. (1990). Relationship between premenstrual symptoms and oral contraceptive use: A controlled study. *Psychosomatic Medicine, 52,* 86–96.

Walker, A. E. (1997). *The menstrual cycle*. London: Routledge.

Walker, E. A., Katon, W. J., Neraas, K., Jemelka, R. D., & Massoth, D. (1992). Dissociation in women with chronic pelvic pain. *American Journal of Psychiatry, 149,* 534–537.

Wallace, L. (1983). *Psychological studies of the development and evaluation of preparatory procedures for women undergoing minor gynaecological surgery*. University of Birmingham, UK: unpublished PhD thesis.

Wallace, L. (1984). Psychological preparation for gynaecological surgery. In A. Broome & L. Wallace (Eds.), *Psychology and gynaecological problems* (pp. 161–188). London: Tavistock.

Walton, J., & Youngkin, E. (1987). The effect of a support group on self-esteem of women with premenstrual syndrome. *Journal of Obstetric, Gynaecological and Neonatal Nursing, 16,* 174–178.

Warner, P. (1994). Preferences regarding treatments for period problems: relationship to menstrual and demographic factors. *Journal of Psychosomatic Obstetrics & Gynaecology, 15,* 93–110.

Warner, P., & Bancroft, J. (1990). Factors related to self-reporting of the premenstrual syndrome. *British Journal of Psychiatry, 157,* 249–260.

Weinman, J., & Johnston, M. (1988). Stressful medical procedures: an analysis of the effects of psychological interventions and of the stressfulness of the procedures. In S. Maes, C. Spielberger, P. Dafares, & I. Sarason (Eds.), *Topics in Health Psychology* (pp. 205–217). London: Wiley.

Weiss, L., & Meadow, R. (1979). Women's attitudes towards gynaecologic practices. *Obstetrics and Gynecology, 54,* 110–114.

World Health Organisation (1981). *Research on the menopause*. Geneva, Switzerland: Author.

Wijma, K., & Areskog-Wijma, B. (1987). Women's experiences of their genitals as an important aspect of their meetings with gynaecologists. *Journal of Psychosomatic Obstetrics and Gynaecology, 6,* 133–141.

Woods, N., Mitchell, E., & Lentz, M. (1995). Social pathways to premenstrual symptoms. *Research in Nursing and Health, 18,* 225–237.

Woods, N. F., Most, A., & Dery, G. K. (1982). Prevalence of perimenstrual symptoms. *American Journal of Public Health,72,* 1257–1264.

Woollett, A. (1996). Reproductive decisions. In C. A. Niven & A. Walker (Eds.), *The psychology of reproduction. 2. Conception, pregnancy and birth* (pp. 1–13). Oxford, UK: Butterworth-Heinemann.

Worsley, A., Walters, W. A. W., & Wood, C. (1977). Screening for psychological disturbance amongst gynaecological patients. *Australian & New Zealand Journal of Obstetrics & Gynaecology, 17,* 214–218.

Zadinsky, J., & Boyle, J. (1996). Experiences of women with chronic pelvic pain. *Health Care for Women International, 17,* 223–232.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.16
Reproductive and Obstetric Issues

BETH ALDER
University of Dundee, UK

8.16.1 INTRODUCTION 384
  8.16.1.1 Reproductive Issues 384
  8.16.1.2 Gender Differences 384
  8.16.1.3 The Social Context 385
  8.16.1.4 Reproductive Issues and Health 385
8.16.2 CONTRACEPTION 385
  8.16.2.1 Current Contraceptive Use 385
  8.16.2.2 Oral Contraception 386
  8.16.2.3 Emergency Contraception 386
  8.16.2.4 Sterilization 387
    8.16.2.4.1 Vasectomy 387
    8.16.2.4.2 Laparoscopic sterilization 388
  8.16.2.5 Other Contraceptive Methods 388
  8.16.2.6 Models of Contraceptive Choice 388
8.16.3 INFERTILITY 389
  8.16.3.1 The Experience of Subfertility 389
  8.16.3.2 Unexplained Infertility 390
  8.16.3.3 Stress and Infertility 390
  8.16.3.4 Consequences of Diagnosis and Treatment 391
    8.16.3.4.1 Distress 391
    8.16.3.4.2 Grief 391
  8.16.3.5 Sexuality and Infertility 391
  8.16.3.6 Reproductive Technology 392
    8.16.3.6.1 Donor insemination 392
    8.16.3.6.2 In vitro *fertilization* 392
    8.16.3.6.3 Surrogacy 393
  8.16.3.7 Adoption 393
8.16.4 MISCARRIAGE AND ABORTION 393
  8.16.4.1 Miscarriage 393
    8.16.4.1.1 Spontaneous abortion 393
    8.16.4.1.2 Grief and loss 394
    8.16.4.1.3 Postmiscarriage care 394
  8.16.4.2 Abortion 395
    8.16.4.2.1 Psychological consequences 395
8.16.5 PREGNANCY 396
  8.16.5.1 The Impact of Pregnancy 396
  8.16.5.2 Transition to Parenthood 396
  8.16.5.3 Antenatal Classes 396
  8.16.5.4 Psychological and Social Changes 397
  8.16.5.5 Changes in Sexuality 397
  8.16.5.6 Teenage Pregnancy 398

8.16.6 THE POSTNATAL PERIOD 398
- *8.16.6.1 Changes after the Birth* 398
- *8.16.6.2 Postnatal Change in Mood* 399
- *8.16.6.3 Postnatal Depression and Development of the Infant* 400
- *8.16.6.4 Breast-feeding* 400
  - *8.16.6.4.1 Breast-feeding as a health behavior* 400
  - *8.16.6.4.2 Breast-feeding and the mother* 401
  - *8.16.6.4.3 Breast-feeding and sexuality* 402

8.16.7 THE OLDER MAN AND WOMAN 402

8.16.8 PROFESSIONAL ISSUES 402

8.16.9 REFERENCES 403

8.16.1 INTRODUCTION

8.16.1.1 Reproductive Issues

Reproductive issues in psychology cross many disciplinary boundaries. Medicine, sociology, anthropology, and social policy all have relevance to reproduction and are important in the lives of both men and women. Reproductive events such as menarche, the first sexual relationship, pregnancy, or the birth of the first child all have considerable psychological impact on individuals and their families.

The epidemiology of reproductive issues has been widely researched. Most of the general population comes into contact with the medical profession in the context of reproductive issues, and their health and behavior can be monitored. There have been studies on the incidence, prevalence, and etiological factors of infertility, sexual problems, and postnatal mental illness. Reproductive issues are relevant to all of us throughout our lives, and this chapter takes a lifespan approach. Life is sometimes taken to be cradle to the grave, but it has also been described as being "sperm to worm." For example, the mother's future maternal behavior may be influenced by the circumstance of the conception or events during pregnancy (hence, parent craft classes). The consequences of reproductive and obstetric events persist throughout life. Family members mainly care for elderly people and the implication of changes in family patterns will influence health care into the twenty-first century. Some reproductive events have psychological significance that go beyond the scope of this chapter. Sexual abuse in childhood may influence sexual relationships, sexual behavior, and parenting. A psychodynamic approach to sexual problems may contribute to understanding and therapy, although many couples may benefit from behavioral therapy approaches. A feminist perspective considers that femininity itself is closely tied to the regulation of sexuality of women; female sexual problems and sexual dysfunction may be defined by male views of female sexuality.

8.16.1.2 Gender Differences

Inequalities of health have probably always occurred even though equal health status between nations, rich and poor, urban and rural, and between ethnic populations has been a goal for many. The differences in health between men and women must be due in part to the biological differences in reproductive and obstetric behavior and to the cultural and social differences between men and women. An approach which recognizes that differences are not the same as undesirable inequalities is likely to be more fruitful in considering health in relation to reproductive and obstetric events.

Much of health psychology has concentrated on men's diseases and male health promotion, for example, heart disease. The first major study of psychological factors in heart disease was based on a prospective study of over 3000 middle-aged male nonmanual workers (Rosenman et al., 1975), and a prospective study of psychological factors and heart disease in Chicago looked at 2000 men only (Persky, 1987). The pattern of morbidity in heart disease and cancer in women in some countries is approaching that of men. Smoking is now increasing in young women in the UK, whereas the number of men smoking has decreased (Jarvis, 1994). Women live longer than men but have higher rates of mental and physical illness. They consult physicians more frequently, have more surgical procedures, have higher prescription rates, and take more over-the-counter medicines. These behavior patterns may be reflections of illness behavior rather than physical morbidity and there is some evidence that women consult for more minor illnesses than men (McCormick, Rosenbaum, & Fleming, 1990). Women are more likely than men to receive psychiatric care, whether measured by inpatient admissions, general practice consultations, prescription of psychotropic drugs, or rates of ECT (Ussher, 1991). In community studies, women are found to more likely to be depressed than men. The epidemiological evidence is based on diagnostic psychiatric interviews or screening scales and it has been argued that these are based on male-based criteria (Ussher, 1991). Interestingly, the gender differences in psychoses are less clear. Men are more likely to be admitted for alcohol-related illnesses than women. One explanation for this increased morbidity is the influence of gynecological events. Although the myth of the "wandering uterus" no longer pertains, there may be a belief that by virtue of their reproductive system, women are psychologically different: the biological determinants usually thought to be hormones and social influences. Two mechanisms can be contrasted. Hormones are held to be responsible for cyclical fluctuations in mood and behavior. This has been described as premenstrual syndrome (PMS), although there is little agreement on definition and the research evidence is full of methodological pitfalls (Choi, 1994). This is considered in more detail in Chapter 8.15, this volume.

Alternatively, women are seen as suffering from more material and social deprivation than men. Women are poorer than men, own less wealth, and have occupations of lower status. Deprivation is closely linked to morbidity, and the association is strong for mental illnesses. However, causal relationships have not been established. Depression may both cause and be caused by poverty. The links may be complex. Single parents may be more at risk for depression but also suffer material deprivation and may lack a confiding supportive relationship.

8.16.1.3 The Social Context

As more people live longer, more women will require medical care, but more young elderly women will themselves be carers.

"Care in the community" is being encouraged but women take up the burden of caring for the elderly. Social support comes from family and the community (Scambler, 1997). Mortality may be related to the degree of social support (House, Robbins, & Metzner, 1982). For example, mortality is related to the level of social support for men but not for women. The relationships are nonlinear, that is, low levels are disadvantageous but high levels give no advantage over middle levels of support. Social support could be related to economic factors related to health such as diet or housing. Some health behaviors may be enhanced if there is social support. Adherence to drug therapy may be greater if there is a family member to remind the patient. The buffer theory suggests that social support may protect against physiological stress (Pearlin, Mullen, Semple, & Skaff, 1990).

Differences in social support between the sexes may occur because of definitions of social networks. The extent of social networks may not reflect their quality. Carroll, Niven, and Sheffield (1993) suggest that, as social networks of women expand (in part because of increased work outside the home), the demands on them as expected providers of care and support may increase. This may impose extra strain incumbent on increased helping and giving. Women may more often be the providers of social support and there are increasing numbers of women living alone as a result of widowhood, single state, or divorce (Carroll, Niven, & Sheffield, 1993).

8.16.1.4 Reproductive Issues and Health

Psychological and physical health and their interaction are influenced by sexuality and fertility. An apparently unrelated disease syndrome or health problem may have been affected by previous reproductive events. Its management may depend on the stage of the reproductive career. Decisions about hysterectomy may depend on whether the couple have as many children as they want. Reaction to the diagnosis of HIV infection may be particularly problematic if it is made during pregnancy. The use of psychotropic drugs in the management of postnatal mental illness will be influenced by breast-feeding. Reproductive changes over the life span particularly affect women, and therefore this chapter focuses on women's health. Men's health may also be greatly influenced by reproductive issues and will be considered in context, that is, in contraceptive choice and infertility problems.

8.16.2 CONTRACEPTION

8.16.2.1 Current Contraceptive Use

The birth rate in the Western world has fallen dramatically since the 1940s (Scambler, 1997). It is not clear to what extent this is because of the increase in education of women, relative affluence, parental choice, or provision of contraceptive advice and supplies. In the European Union the fertility rate measured in births per 1000 women fell from 2.4 in 1970 to 1.5 in 1992. In 1993, 39% of women between 16 and 35 used the pill. Contraceptive use is shown in Table 1 (*Social Trends*, 1995).

The evidence for a contraceptive career (Lindemann, 1977) suggests that young people use condoms or no method at all in their first sexual encounters (Abraham & Sheeran, 1994). With increasing sexual and contraceptive experience they move to oral contraceptives. After the birth of their first child, they change to intrauterine device (IUD) and, after their family

Table 1 Contraceptive use in 1993 in the United Kingdom (*Social Trends*, 1995).

Method	*Percentage of users aged 16–49 years*
Oral contraceptive pill	25
Condom	17
Hysterectomy	12
Vasectomy/sterilization	12
IUD	5
Withdrawal	3
Other	3

is considered to be complete, they choose either male or female sterilization. Oddens (1996), in a population survey of contraceptive use in the UK and Germany, found that patterns of contraception use were closely related to variables such as age, parity, marital status, and intention to have children. Age and the desire to have children are the major factors associated with contraceptive choice. The decline in the use of oral contraceptives with age has been related to perceptions of health risk. In 1989 a "pill scare" in Germany associated certain pills with an increased risk of thromboembolism and this was followed by a fall in the number of oral contraceptive users. This happened again in the UK in October 1995 (Martin, Hilton, & Kerry, 1997). A comparison of the major reasons given for choosing different contraceptives is given in Table 2.

8.16.2.2 Oral Contraception

Guillebaud (1991) estimated that 70 million women worldwide take the oral contraceptive pill every day. Over 90% of women aged 35 in 1987/88 in the West of Scotland had used the oral contraceptive pill at one time (Hunt, 1990). Oral contraceptives tend to be the first choice of contraception, although there has been a fall in their popularity in the UK, accompanying a rise in the use of condoms (*Social Trends*, 1995). Oral contraceptives have been thought of as the gold standard of contraception. They are easy to use, very effective, safe, inexpensive, and readily available. They do not interfere with the act of sexual intercourse and give women control over their fertility. However, they are not without some disadvantages. Side effects reported in the early high oestrogen oral contraceptive pills included reduced sexual interest and depression (Guillebaud, 1991).

Condon, Need, Fitzsimmons, and Lucy (1995) carried out a retrospective study of 145 women and found that nearly half reported side effects and most changes in well-being were perceived as negative. It is difficult to carry out good prospective studies of pill users because of ethical difficulties and high attrition rates. Women who have side effects may stop using the pill, and those who continue and become long-term users are less likely to report side effects. Women who take the pill probably do so because it is very important to them to have a very safe form of contraception, and they may be more likely to tolerate side effects for the sake of a high level of protection (Guillebaud, 1991).

Women who use oral contraceptives may differ in their sexual behavior from those using other methods. They may wish to avoid interference with the act of sexual intercourse. The evidence for a direct psychological effect of the pill is weak. The most common psychological side effects are depression and loss of libido. Bancroft and Sartorius (1990) in a literature review conclude that oral contraceptives decrease premenstrual symptoms and suggest that a small group of women, number unknown, are adversely affected. In a study of readers of a monthly women's magazine, Warner and Bancroft (1990) asked about the level of well-being and sexual interest in "the week before the period," "during the period," "the week after the period," and "other times" or "never." Women taking the pill were less likely to report highs and lows in well-being overall than nonpill taking women. They were more likely to report high sexual interest "during the period" or "before," but this did not relate to changes in well-being. These findings suggest that there may be a direct effect of the pill on sexuality, independent of its effects on mood. Bancroft (1988) suggests that oral contraceptives may have an adverse effect on mood initially but some women stabilize, whereas other women will change their brand of pill or abandon oral contraceptives altogether.

8.16.2.3 Emergency Contraception

Hormonal postcoital contraception can prevent pregnancy in nearly 100% of cases, and yet few women presenting for termination of pregnancy have attempted to use this form of contraception (Smith, Gurney, Aboulela, & Templeton, 1996). George et al. (1994) found in a study of women attending a general practice surgery that the greater level of awareness of emergency contraception was associated with previous use of barrier contraception, young age group, parity, and previous emergency contraception enquiry. Smith et al. (1996) carried out a population-based survey of 1214 women in Grampian, Scotland (65% response rate). They found that although most women (94%) were aware of emergency contraception, only 39%

Table 2 Motives for choosing contraceptive method currently used (percentage). More than one answer could be given. Germany, $n = 1064$; UK, $n = 967$.

	Oral contraception		*Condom/ diaphragm*		*IUD*		*Sterilization (vasectomy and laparascopic)*	
Reason	*Germany*	*UK*	*Germany*	*UK*	*Germany*	*UK*	*Germany*	*UK*
Reliability	94	88	34	45	66	69	52	64
Nondisturbance of sex life	62	65	18	8	83	48	20	38
Easy to use	51	69	27	40	41	47	2	7
Menstrual problems	33	29	0	1	3	1	4	1
Medical advice	21	12	4	7	43	39	32	14
Infrequent intercourse	2	1	26	15	21	2	0	2
Health concerns	1	2	36	32	21	11	34	18
Partner involvement	2	0	23	11	0	0	15	12
Problems with former method	3	7	30	37	39	40	33	16

Source: Oddens et al. (1994a, 1994b).

knew the correct timing for its use. Two pills need to be taken 12 hours apart within 72 hours of unprotected sexual intercourse. The morning-after pill may have had a negative image in the media and there is debate about the advisability of deregulation (Glasier, 1993). Only 13% referred to it as being a form of abortion or gave a moral aspect as a reason for being opposed to increasing advertising. The so-called "once a month pill" (the antiprogesterone, RU486) is taken at about the time of the expected menses and induces menstrual-like bleeding and/or abortion. In a survey of attitudes in Scotland, Romania, and Slovenia it was found that 72%, 81%, and 94% of women, respectively, felt positive to the idea of a pill which inhibited ovulation. Over 50% thought that a pill which inhibited or interfered with implantation was an acceptable idea. A pill which was taken only if menstruation was delayed by one or two days was acceptable by only 24% in Scotland, but 80% in Romania, and 80% in Slovenia. This suggests that attitudes to abortion, availability of contraception, and religious beliefs are likely to be important (Rimmer et al., 1992).

8.16.2.4 Sterilization

Male or female sterilization is the most widely used method of contraception worldwide. Vasectomies in men and laparoscopic sterilization in women are safe and effective. However, sterilization operations are effectively irreversible and so the decision to be sterilized means the couple need to be very sure that they will not want more children.

8.16.2.4.1 Vasectomy

Vasectomy is a simple minor surgical procedure that takes about 15 minutes to perform (Hargreave, 1994). A small opening is made in the scrotum and the vas deferens that carries the sperm from the testes to the urethra is severed. The procedure is carried out under sterile conditions and under local anesthetic. It usually takes place in outpatient clinics. The patient is usually advised to rest for one or two hours in the clinic, and avoid hard physical work or strenuous exercise for the next few days. Infertility may take up to 10 weeks to achieve depending on the frequency of ejaculation. It is one of the most effective methods of contraception and there are very few side effects. Studies on the long-term effect of vasectomy have found no evidence of adverse psychological effects. Retrospective studies, for example, Cliquet, Thiery, Staelens, and Lambert (1981) report that less than 5% have problems such as decreased libido and depression. However, most studies did not include control groups, and studies carried out in the 1970s used nonstandardized assessments, heterogeneous groups, and were mostly retrospective. Regret may lead to requests for reversal, and these usually follow remarriage after divorce, death of one or more children, change in family circumstances, or psychological problems related to infertility. The operation can be reversed but may be successful in only 50% of cases. Some men are reluctant to undergo an operation on their genitals. Men are potentially capable of fathering children into their sixties, and some men do not want to end their fertility prematurely. The decision to be sterilized needs to be considered carefully and counseling is often given in order to help the couple decide which partner should undergo the operation (Hunter, 1994).

8.16.2.4.2 Laparoscopic sterilization

Laparoscopic sterilization in women involves tying, cutting, or more often clipping, the fallopian tubes that transport the eggs from the ovary to the uterus. It is a simple and effective surgical procedure, involves one post-operative check, has few side effects, and has a low failure rate. It leaves only two small scars and is a very safe surgical procedure. In a retrospective study of over 500 patients, only 2% complained about physical problems (Lawson, Cole, & Templeton, 1979).

It is very difficult to repair fallopian tubes successfully, but very few women regret that they have had the operation. Lambers, Trimbos-Kemprer, and van Hall (1984) identified the factors associated with regret following sterilization as being lack of freedom of choice in the decision-making process; a belief that the sterilization operation would solve other problems; an inability to cope with loss of fertility; and a strong adherence to a sex role identity that was linked to reproductive capacity. Regret rates of up to 15% were reported in early studies, but prospective studies have found much lower rates of nearer 5% (Alder, 1984). There is some evidence that young women may be more likely to regret sterilization, but this is probably because there is an increased risk of marital breakdown associated with early marriages. In a retrospective study (Alder et al., 1981), interviews were carried out with a sample of 45 women who had been sterilized and a matched group of women whose partners had had a vasectomy. It was found that women who had been sterilized reported more menstrual changes both positive and negative than those whose husbands had had a vasectomy. Six women expressed some regret. These consisted of two whose housing conditions had improved, one whose marriage had broken down, one because of a perceived loss of femininity, one who wanted more children, and one because of menstrual problems that began after she had stopped taking the oral contraceptive pill. A prospective study of 115 Chinese women followed up over a year (Tang & Chung, 1997) found that there was no adverse effect on sexual adjustment and that their mental health measured by the General Health Questionairre (GHQ) (Goldberg, 1978) significantly improved. The regret rate after one year was 3.4% (Tang & Chung, 1997).

Women requesting sterilization appear to have no higher rates of psychiatric disorder than the general population (Cooper, Gath, Rose, & Fieldsend, 1982). In a prospective study of 200 women who were interviewed preoperatively, and at six and 18 months postoperatively, the number of psychiatric cases assessed using the Present State Examination (PSE) fell from 10.5% preoperatively to 4.7% at six months and 9.3% at 18 months. The PSE is a very robust standardized psychiatric diagnostic interview. Only 6.7% of women became psychiatric cases for the first time, which was the same as in the general population. They found that the frequency of sexual intercourse had remained unchanged over the period of the study in nearly half, fallen in just over a quarter, and increased in just under a quarter. Sexual satisfaction improved, possibly because of the removal of fear of pregnancy.

There are cultural factors which may influence the reaction to sterilization in men and women, and postpartum sterilization may be followed by a higher rate of physical and psychological problems than sterilization at least three months after delivery (Oates & Gath, 1989).

8.16.2.5 Other Contraceptive Methods

The IUD is mainly used by parous women and confers good protection. Like the pill it does not interfere with sexual activity. No adverse reports on mood or sexuality have been reported with the use of IUDs (Bancroft, 1988). Since the 1980s there has been an increase in the use of condoms, but few new methods of contraception have been developed recently, possibly because of fear of potential litigation. Female condoms have been introduced and appear to be acceptable (McCabe, Golub, & Lee, 1997). The male pill is another possible development, but there is some doubt about whether women would trust men to take it regularly (Guillebaud, 1991). In a report based on surveys in 1992, Oddens (1996) suggests that attitudes towards the perceived medical nature of some contraceptive methods (oral contraceptives, sterilization, and IUDs) influence contraceptive choice.

8.16.2.6 Models of Contraceptive Choice

Lindemann (1977) described a three-stage model of a contraceptive career in women. In the first "natural" stage, sexual intercourse is relatively rare and unplanned. A woman does not perceive herself as a sexual being and she neither uses contraception nor takes responsibility for it. In the second "peer prescription" stage, sexual activity is more frequent. There is a moderate acceptance of sexuality, and she may seek information from friends. In the third "expert stage" she accepts her sexuality and becomes willing to seek contraceptive advice. The career model was supported for women but not for men in a meta-analysis of adolescent contraceptive use (Whitley & Schofield, 1986).

Men and women may differ in their relative influence on the choice of contraceptive method. It has been assumed that women primarily influence contraceptive behavior within marriage and much family planning advice is directed at women. One of the few research studies on couples' contraceptive choice studied 40 couples over a four-year period (Miller & Pasta, 1996). They distinguished between method-choice and method-use in decision-making. Method-choice is about continuing with the present method or changing to a new one. Method-use is the actual use of the selected method, for example, taking the oral contraceptive pill daily or inserting a diaphragm. They found that husbands and wives appeared to have equal influence on method-choice but intentions depend on their own preference. This suggests that the final choice is based on sharing information and preferences, perhaps on a cost-benefit basis.

Social cognition models have been used with varying degrees of success to predict contraceptive use (e.g., Galavotti et al., 1995). The transtheoretical model of behavior change (Prochaska & DiClemente, 1984) suggests that people go through a series of stages in adopting a new behavior: precontemplation, contemplation, preparation, action, and maintenance. Grimely, Prochaska, Velicer, and Prochaska (1995) tested this model on 248 American college students and found that it could predict patterns of condom use. Mahoney, Thombs, and Ford (1995) tested the health belief model in college students in relation to condom use and HIV infection. Sporadic users differed from both consistent and nonusers in the number of sex partners in the previous year and frequency of drunkenness during sexual intercourse, but only perceived susceptibility to infection and self-efficacy differed between the groups. In a meta-analysis of over 130 studies of contraceptive use (Whitely & Schofield, 1986), contraceptive use was strongly associated with perceived risk of pregnancy. Social cognition models have attempted to predict condom use (Scheeran & Abraham, 1996), but in a longitudinal study of condom use over one year among 16- and 18-year-olds, Health Belief Mode variables were not found to predict condom use (Scheeran & Abraham, 1996).

The choice of contraceptive method may depend on the perceived attributes of different contraceptive methods. Harvey, Beckman, and Murray (1991) compared 330 vaginal sponge users with 330 oral contraceptive or diaphragm users drawn at random from the same geographical areas and interviewed over the telephone. Effectiveness and safety were rated most highly. Each user group emphasised those attributes that characterised their chosen method. Convenience of use and interference with sexual activity were the least emphasised. The relative undesirability of other methods (e.g., side effects of the pill) was important.

This suggests that the choice of a method of contraception is a rational decision-making process. However, if the method requires an action such as putting on a condom (possibly in an emotionally charged sexual interaction), the actual behavior may not match the intention. The role of perceived behavioral control suggested in the theory of planned behavior is also likely to be important in negotiation of contraceptive use (Conner & Norman, 1996).

8.16.3 INFERTILITY

8.16.3.1 The Experience of Subfertility

Some couples choose to remain childless, and increasing numbers of couples delay the birth of their first child. The proportion of women who never have children appears to be increasing. It is expected that over a third of women born in 1967 will be childless at the age of 30 (*Social Trends*, 1995). A young couple may be expected to have children after a few years of marriage and their families and friends may well begin to ask them about their family plans or to assume that they are childless by choice and to isolate them for being selfish. Campbell (1985) suggested that in Western society, parenthood is perceived as having high status. Society may assume that for women, bearing children is a natural expression of femininity and that for men having children is a duty.

The research evidence for a link between intentions to have children and fertility is unclear (Miller & Pasta, 1995). In reported studies, intentions have been assumed to mean wanting any children at all, wanting more children, or wanting children within a specified time period. Pregnancy, not behavior, is often used as an outcome measure. The category of pregnant women may include women who have unplanned pregnancies and exclude those who would have become pregnant but are subfertile. In their study of nearly 400 couples they found that child-timing intentions followed by child-bearing intentions were the most important factors in predicting attempts to conceive. In a small qualitative study, Jones and Hunter (1996) used discourse analysis to describe the experience of infertility and question the assumption that it is necessarily stressful. However, men and women can have very intense desires for children, and this has been recognized in folklore. Many stories focus on the longing for a child which is accomplished in some magical way, such as *The Gingerbread Man* or *Tom Thumb*. There may be conditions attached as in the story of *Rapunzel*. In this story the couple must surrender their child to a witch who proceeds to lock her in a high tower.

Most couples use contraception when they first establish sexual relationships and make a conscious decision to stop using contraception in order to have children. It is then assumed that pregnancy will result from unprotected intercourse. After a year with regular intercourse but no conception, many couples become concerned and anxious about their fertility. About 10–15% of couples are involuntarily infertile and childlessness can occur in couples who have never had children (primary infertility), and in those who have already had children (secondary infertility). It is not known how many couples find that they do not conceive but do not seek medical help. Many causes lie within the reproductive system of the woman, but male factors account for about a quarter of couples with subfertility.

Sexual problems associated with conditions such as diabetes may well reduce the chances of conception as well as excessive use of alcohol or drug abuse. Psychiatric illnesses such as anorexia nervosa are associated with infertility. Recent suggestions that the sperm count may be falling in men have given rise to anxiety about future conception rates (Irvine & Cawood, 1996). In her novel *The children of men*, James (1972) describes a world in which there are no new births and people take desperate steps to try to cope with their lack of fertility.

8.16.3.2 Unexplained Infertility

In about one-quarter of couples seeking investigation, the infertility problem will be unexplained, although this depends on the length of time of investigation, and the extent to which the couple is prepared to undergo prolonged and invasive investigations. Estimates of unexplained infertility vary across clinics, but reported rates range from 6 to 60% (Templeton & Penney, 1982). In the absence of any anatomical, physiological, or endocrinological dysfunction it is tempting to attribute the failure to conceive to psychological causes. Sexual problems in men such as impotence and retrograde ejaculation (when the semen is not fully expelled) will reduce fertility. There have been some reports of sexual inhibitions leading to inability to ejaculate within the vagina, but these are uncommon (Irvine & Cawood, 1996).

A number of studies have attempted to compare the characteristics of patients with apparent organic infertility to those with unexplained infertility. These are difficult because the extent of unexplained infertility may reflect the sophistication of the infertility investigations. These can be very invasive and some couples may not wish to comply with requests for postcoital semen samples. Edelmann and Connolly (1989) reviewed studies between 1972 and 1983 and concluded that there were no differences in personality characteristics. However, O'Moore and Harrison (1991) found that stress-related factors distinguished couples with unexplained infertility and couples who were normally fertile. They found significantly higher scores of anxiety, measured by the State Trait Anxiety Inventory (STAI), in infertile women compared with controls. They suggest that stress may raise anxiety levels which influence hormone levels. Prolactin is the most likely candidate. Levels of prolactin can be raised by stressful stimuli but the picture is unclear. Harper, Lenton, and Cooke (1985) found a positive correlation between prolactin levels and stress measured by an adjective check list in 20 couples attending an infertility clinic. They also found a positive correlation between STAI state, but not trait anxiety, and plasma prolactin.

8.16.3.3 Stress and Infertility

Edelmann and Golombok (1989) argue for a link between stress and failure to conceive, but point out that single stressful events do not necessarily inhibit conception. Rape is highly stressful but may still result in conception. It seems unlikely that much infertility often has a "psychogenic" origin, but psychological factors may have some effect. Demyttenaere, Nifs, Steeno, Koninckx, and Evers-Kiebooms (1988) found that initial trait anxiety levels were related to the number of treatment cycles necessary for conception by donor insemination. Counseling may reduce the anxiety and guilt associated with failure to conceive, but even without treatment a proportion will conceive. Hull et al. (1985) found that 56% of women conceived after one year and 72% after two years. Bell (1981) suggests that the stress of infertility investigations themselves may contribute to failure to conceive. Connolly, Edelmann, and Cooke (1987) found that the emotional well-being of couples was affected by prolonged periods of clinical investigation and these were greater if the cause of the infertility lay with the man. In a study of 20 women undergoing attempts to conceive by *in vitro* fertilization in an early program, it was found that they were generally positive about the very stressful and invasive procedures (Alder & Templeton, 1985). Even though they were unsuccessful in conceiving, they thought that it was their only chance, and some felt that they could then begin to accept their childlessness because they had made an effort. Harrison, Callan, and Hennessey (1987) found in a sample of 500 male partners of women undergoing infertility treatment that there were lower sperm concentrations in pretreatment semen samples compared with those given on the day of treatment.

It may be assumed by both medical staff and lay people that women may be more upset if they cannot bear children, but this may be culturally and psychologically determined. Lampman and Dowling-Guyer (1995) gave six scenarios describing couples with different family situations to over 200 undergraduate students. They rated the scenarios on characteristics of the couples and their relationships. The results showed negative ratings of voluntary childlessness but less so towards involuntary childlessness.

8.16.3.4 Consequences of Diagnosis and Treatment

8.16.3.4.1 Distress

Berg and Wilson (1995) looked at the pattern of distress in 104 couples with primary infertility attending for infertility investigations. Forty percent of men and 49% of women met caseness criteria on the SCL-90-R scale (Derogatis, 1977). In a third of couples both were not distressed, in 18% the male only, in 22% the female only, and in 27% both partners were distressed. Edelmann and Connolly (1996) assessed sex-role type (Bem, 1974) on psychological functioning in a sample of 130 couples presenting with primary infertility but before diagnosis. Masculine men were the least anxious, but there were no relationships between anxiety and sex-role type in women. Depressive symptoms may be associated with infertility but the infertility may also be associated with the use of antidepressant drug therapy (Lapane, Zeirler, Lasater, Stein, et al., 1995). Recognition of the problem and the feeling that something is being done may be very reassuring. Couples who have difficulty in conceiving probably discuss their reasons for having children and think through their reproductive choices in a way that most couples do not.

Of course, facing a problem can bring couples together, and Raval, Slade, Buck, and Lieberman (1987) found a reduction in marital and sexual problems after attending discussions about sexual performance and the perceived association between virility and fertility offered by an infertility clinic. Burns (1995) reviews the extent of sexual problems among infertile couples and suggests that nonconsummation may be under-reported.

8.16.3.4.2 Grief

The discovery that conception does not occur on demand and that pregnancies cannot be planned to order may be a shock to modern young women who have been used to controlling their fertility by the use of contraception. Women's sense of identity and personal relationships may be affected. It has been argued that medical intervention for diagnosis and treatment can result in guilt and shame for women who seek treatment for infertility because they see themselves as defective (Whiteford & Gonzalez 1995). Pfeffer and Woollett (1983) describe the emotional reaction to infertility and the pressures of medical investigations. For those that want children but are unable to have them, grief can be very intense. Their pattern of grief may resemble that described for bereavement (Kubler Ross, 1969). Their initial reaction to the diagnosis may be shock and disbelief. Pfeffer and Woollett, who had themselves experienced infertility investigations, suggest that the couple should be told the results of the tests together and given time to absorb and return to ask questions later. Denial may result in rejection of the clinic's findings and a renewed search for medical advice.

8.16.3.5 Sexuality and Infertility

Some couples become so obsessed with monitoring their basal temperature, frequency, and timing of sexual intercourse that their sexual relationships suffer (Bell & Alder, 1994). They may feel very guilty about their feelings and may resent other women who are pregnant or who have as many children as they want. Their feelings of anger may be directed at themselves, perhaps for a previous termination, venereal disease, or premarital promiscuity. The angry feelings may also be directed at the infertile partner, so disrupting the sexual relationship. Paradoxically, some women report feeling guilty if their partners rather than themselves are responsible for their difficulties in conception (Alder, 1984). Anger may also be directed at the health professionals when treatment is unsuccessful. Infertility patients are extraordinarily compliant and clinicians find that they will often put up with the most invasive procedures and lengthy operations if there is even a small chance of success.

Sexuality may be adversely affected following the diagnosis of infertility, and women and men may resent the constraints of timing of intercourse to optimize the chances of conception. Women may express more grief and distress about the lack of ability to conceive, and may feel that their femininity is threatened. There may be a greater adverse effect on male sexuality if the infertility problem is of male origin, and men are more likely to deny that there is a problem (Alder, 1984).

8.16.3.6 Reproductive Technology

8.16.3.6.1 Donor insemination

The opportunities for treatment for infertile couples have been increased by advances in reproductive technology. Artificial insemination by a donor is used in couples where the man is infertile but the woman ovulates normally. Semen from an anonymous donor is inseminated into the vagina of the fertile woman. If she conceives she will have a child who carries half of her genes and half of those of the donor. The pregnancy proceeds normally and the couple need not tell anyone that the conception did not occur in the usual way. The couples may be counseled against keeping the circumstances of the conception secret from the child. In an interview study, it was found that out of 20 couples receiving donor insemination (DI), seven wives and 10 husbands would never tell a child that it was conceived by DI (E. M. Alder, 1984). Snowden and Snowden (1984) argue strongly for openness and suggest that children should be told the full facts of their conception when they are considered old enough to understand. In the past, donors have remained anonymous but a Swedish law in 1995 gave those born by DI the right to know the identity of the donor when they become mature. This legislation was followed by a decline in the number of DI births although it is not clear whether this was because of the reaction of health professionals to legislation, or to a decline in the number of donors. An open approach is also advocated in New Zealand, but by a voluntary code (Daniels & Lewis, 1996).

Follow-up studies of DI couples and their children have been difficult to carry out because of the secrecy involved. No doubt part of this is because of the stigma surrounding male infertility and its association with virility. Many couples prefer to raise the children as their own and without any further contact with the medical team that carried out the insemination. Some couples want to deny the existence of the donors but a study of the motivations and attitudes of semen donors in New Zealand suggests that donors themselves are interested in the future of their offspring (Daniels, 1989). Twenty-three semen donors completed an anonymous questionnaire about their motivation and attitudes to donor insemination. Almost all were interested in the outcome of their donations although most had not enquired. Most of them said that they thought about their offspring. Fifteen said they would be happy for the children to trace them at age 18.

8.16.3.6.2 **In vitro** *fertilization*

A very significant breakthrough came in 1978 with the birth of the first "test tube baby." Since then *in vitro* fertilization (IVF) has been used extensively, and more than 40 000 babies are estimated to have been born as a result of IVF worldwide. The procedure carries a low success rate (22% in women under age 38 in a London hospital between 1988–1995; Power, 1997), but it may be the only way in which women who have damaged fallopian tubes can bear children. Couples who conceive by IVF are not typical of the general population. They may have had intensive and prolonged infertility investigations and treatment and may be very anxious about the outcome of the pregnancy. Johnston, Shaw, and Bird (1987) found that couples entering an IVF program overestimated their chances of success. At the time the published success rate of having a live baby was as low as 8%, but over half the sample rated their chances as over 30%. They are likely to be older and highly motivated to become parents. They appear to have as good or better marital relationships than comparison groups. This may reflect a high level of commitment to marriage and the demands of infertility treatment, and IVF may foster communications. However, those that take part in psychological studies may feel that they need to assert their suitability as parents (McMahon, Ungerer, Beaurepaire, Pennant, & Saunders, 1995).

IVF treatment is expensive and Woollett, Dosanjh-Matawala, and Hadlow (1991) point out that couples may become more anxious because of the costs of private infertility treatment and its restricted availability. There are concerns about the future of reproductive technology and in the UK treatment is regulated by law. Shaw, Johnston, and Shaw (1988) showed that about half of 118 couples waiting for IVF or embryo transfer (ET) treatment in 1986 desired counseling and the offer of counseling is now mandatory. The general public may be sympathetic to the problems of those with infertility problems. In an early study, Alder, Cook, Davidson, West, and Bancroft (1986) surveyed nearly 2000 women attending a family planning clinic and found that the majority were in favor of research into *in vitro* fertilization and ET.

8.16.3.6.3 Surrogacy

A surrogacy arrangement is an undertaking given by a woman (the surrogate mother) to bear a child for another person or couple (the commissioning parents), the latter proposing to undertake the future care of the child (Blyth, 1995). Conception can occur by insemination of the surrogate mother by the commissioning father's sperm, or by sexual intercourse for the purposes of conception. "Host" surrogate mothers have a fertilized egg implanted into the uterus and the embryo will be the genetic offspring of the commissioning parents. There may be legal control over surrogacy arrangements, particularly financial arrangements. Currently, commercial agencies are not allowed to operate in the UK. There is still controversy and debate about the moral and ethical issues.

The case of "Baby M" in the 1980s in the USA involved a commissioning couple who contracted with a surrogate mother. She then gave birth to a baby which she refused to hand over. Eventually, after legal battles, custody of the child was given to the commissioning couple. This raised a number of questions including the motives of surrogate mothers, mother–child bonding relationships (she breast-fed for three months), and the rights of the grandmothers. The psychological and social issues are discussed by Schwartz (1991). In a qualitative study, Blyth (1995) interviewed 20 commissioning couples who were members of a self-help group (childlessness overcome through surrogacy, COTS). Most of them knew about surrogacy from the media, and the idea was usually initiated by the wife. A constant anxiety for the commissioning couples was that the surrogate might change her mind. Most commissioning parents expected to be present at the birth. All of them believed that the child should be told the full truth of its genetic and biological origins. Blyth describes the shift in attitudes of the medical profession and compares the openness now recognized in adoption with increasing openness in surrogacy.

8.16.3.7 Adoption

Adoption can be difficult because of the shortage of babies, and the selection procedures are very rigorous. The principle that the best interest of the child is paramount is held by most of the legal systems in most countries. Changes in approaches that have occurred since the 1970s have been towards more openness about the process. Open adoption is where the birth parents and the adopting parents know each other and maintain a relationship with each other. It has the advantages of increased feelings of control and open communication and maintenance of self-identity. However, there are risks of threatening the adopted parents' confidence in parenting and ambivalence about contact with birth parents (Demick & Wapner, 1988). The implications for the family are discussed by Humphrey and Humphrey (1988) and Schwartz (1991).

There may be many barriers to adoption such as marital status, age of the partner, previous medical history, or other social criteria. These are in spite of the evidence that single women do raise children successfully on their own, that older parents have no more problems than younger parents, and that disability does not preclude successful (however defined) parenting. Perhaps because of the shortage of babies offered for adoption together with societal norms, couples who have difficulty in conceiving and bearing their own children seek medical intervention.

8.16.4 MISCARRIAGE AND ABORTION

8.16.4.1 Miscarriage

8.16.4.1.1 Spontaneous abortion

Miscarriage is defined as the spontaneous loss of a pregnancy within the first 24 weeks of gestation. Most miscarriages are caused by genetic malformation of the fetus, but they may also be caused by incompetence of the cervix, hormone deficiencies, uterine abnormalities, or infections. About 20% of pregnancies are lost by 24 weeks. There is probably a much higher level of undiagnosed loss than recorded in hospital statistics, and this may remain unrecognized. In the UK there is no need to register the birth if the pregnancy is lost before 24 weeks, and the miscarrying woman may not be referred to a maternity ward.

Miscarriage may be treated entirely as a gynecological event and women are admitted for removal of the products of conception either medically or surgically. The psychological effects of the loss of early pregnancy may be unrecognized. Clinicians may assume that the pregnancy is not thought of as a baby and that there is no need to mourn the loss (Iles, 1989). Women probably only stay in hospital for a short while, and medical and nursing staff may not be aware of later distress. There may be a reluctance to talk about the event to others. It may be perceived as "nature's way" or as "being for the best." Miscarriage carries the double "social taboo" of being both a gynecological event and a death.

The introduction of home pregnancy kits may mean that women become aware that they are pregnant early and in the privacy of their own home. They do not need to acknowledge it to medical staff, and can choose those in whom they confide. This may mean that social terminations can take place early, but may also mean that more miscarriages are recognized as such. The widespread use of ultrasound scans early in pregnancy means that women may become aware that they are pregnant at an early stage in pregnancy and at a time when there is a relatively high risk of miscarriage. They may also identify themselves as carrying a baby rather than a fetus because they have seen the picture from the ultrasonic scan.

8.16.4.1.2 Grief and loss

In retrospective interview studies, grief reactions similar to those following stillbirth have been found (Friedman & Gath, 1989). The literature on bereavement suggests that in any loss followed by grief and mourning, resolution may take some time (Parkes, 1985). Leppert and Pahlka (1984) described grief in 22 women who had had a miscarriage. The grief did not resolve until three to four months after the loss. Prettyman, Cordle, and Cooke (1993) found that 41% of women were anxious one week after miscarriage and 32% after 12 weeks. Janssen, Cuisinier, de Graaw, and Hoogduin (1997), in a study of women who had had pregnancy loss in the Netherlands, found that grief intensity was higher in women who had been pregnant longer, had preloss neurotic personalities, and did not have other living children.

There can be a great sense of loss of a future life but women may also be mourning the loss self-esteem and their status as a future mother. Jackman, McFee, and Turner (1991) used semistructured home interviews to assess 27 women who had had a previous miscarriage and attended hospital. Half were found to be "cases" measured by the GHQ-30 scale (Goldberg, 1972). Grief can be considered normal but in a study of 67 women admitted to hospital for spontaneous abortion (Friedman & Gath, 1989), 84% were diagnosed as being "cases" using the PSE (Wing, Cooper, & Sartorius, 1974). Those who were more likely to show psychiatric morbidity were those who were single, had a previous psychiatric history, were high on scores of neuroticism, and had had previous obstetric problems.

8.16.4.1.3 Postmiscarriage care

Iles (1989) suggests that psychiatric morbidity could be reduced by advising women not to become pregnant too soon after miscarriage. A new pregnancy may inhibit the completion of mourning for the baby that was lost. Medical advice is usually to delay a further pregnancy for up to six months but many parents disregard this advice (Davis, Stewart, & Harmon, 1989). If the grief is unresolved the woman may remain emotionally detached in a subsequent pregnancy, so as to protect herself in case the pregnancy fails again. It is suggested that the mother may continue to be detached after the birth and remain unresponsive to the new baby, or alternatively over-react and become over protective. They suggest that in order to reduce the risks of adverse reactions, women

(i) should not necessarily be nursed apart from other women because this could lower still further their self-esteem;

(ii) should be assisted to comprehend the reality of the loss, particularly in the absence of a physical presence to mourn. Only 4/27 women in the study by Jackman et al. (1991) saw the physical remains, and only two were consulted about a burial. The reality of the loss also needs to be explained to the woman's family;

(iii) need to be given as much information as possible and to be reassured that they are unlikely to have been responsible for the loss of their pregnancy; and

(iv) need someone to listen to their experience and with whom they can share their sense of guilt and failure.

(i) Psychological consequences

Results based on retrospective accounts can be difficult to interpret, but of course prospective studies on miscarriage are difficult to do. Cuisinier et al. (1996) surveyed over 2000 women recruited from a Dutch magazine for expectant and new parents. 221 women lost a singleton through miscarriage, stillbirth, or neonatal death, and were followed up on four occasions at six-month intervals by a postal questionnaire. 193 miscarried within the first 16 weeks of pregnancy and the women were assessed on a number of standardized questionnaires. Grief was assessed by the perinatal grief scale (PGS; Potvin, Latvin, & Toedler, 1989), but feelings about future pregnancies, contraceptive use, and intentions to conceive were not reported. Most (86%) conceived again within 18 months. Of those who became pregnant within six months, only 3% reported that they saw the baby as a replacement child. The responses to the questionnaire were analyzed by regression analysis and it was found that a new pregnancy and birth of a new child significantly lessened grief, while grief significantly intensified following a new pregnancy loss. They also suggest that the speediness of the new pregnancy appears to reduce grief to some extent. They found little indication of severe maternal psychopathology likely to interfere with a healthy mother–child relationship. However, they did not use a standardized psychiatric interview, the data were collected by post, and the sample of readers may not be representative. These authors, and Davis et al. (1989), suggest that the topic of a new pregnancy should be discussed with the couples and they should be helped to make their own informed decisions and given subsequent support.

(ii) The male partner

Few studies have considered the effect of miscarriage on the male partner. Women may be more ready to cry and express their loss than men, and men may be seen as supporting the woman in her loss rather than grieving themselves. If father–infant "bonding" takes place after birth then it would be expected that men would not become attached to the fetus and that the duration of the pregnancy would not contribute to the extent of the grief. Johnson and Puddifoot (1996) described the response of male partners. They interviewed an opportunity sample of 126 stable partners of women who had recently miscarried. The scores on the PGS and the impact of events scale were similar to those found in women. They found a very similar reaction to the loss that has been described in women. The duration of the pregnancy and seeing the ultrasound scan were associated with an increasing level of grief in men on all subscales of the PGS.

(iii) Miscarriage after IVF

Loss of pregnancy following IVF may be particularly problematic. If the eggs are fertilized and the embryo is successfully transferred, women may feel that they are already pregnant. They may believe that they are pregnant early in gestation and be treated by others as being pregnant. Greenfield, Diamond, and DeCherney (1988) describe the grief response following failure to become pregnant during an IVF cycle as being analogous to the reaction that women experience after miscarriage.

8.16.4.2 Abortion

Abortion (social termination of pregnancy) was legalized in the UK by the 1967 UK Abortion Act, but abortion is still controversial and open to legal disputes in different countries. The Abortion Act legalized termination of pregnancy "if the continuation of the pregnancy would involve risk to the life of, or injury to the physical or mental health of the pregnant woman or to any existing children of her family, greater than if the pregnancy were terminated." In 1990, The Human Fertilisation and Embryo Act 1990 amended the 1967 Act to limit the gestation to 24 weeks, where formerly it had been limited to 26 weeks.

Adler et al. (1990) reviewed psychological responses to abortion in the USA. Most descriptive studies found a low incidence of severe negative responses after abortion. Two studies (Athanasiou, Oppel, Michaelson, Unger, & Yager, 1973; Zabin, Hirsch, & Emerson, 1989) compared those who had abortions with those who gave birth at term and neither found adverse effects of abortion. Distress is highest before abortion and decreases afterwards. Later levels of distress were related to difficulties in making the decision. However, Adler et al. point out that there are methodological problems of sample selection, no baseline data, short follow-up time, and lack of control for "wantedness" of the pregnancy.

8.16.4.2.1 Psychological consequences

The psychological effects of abortion have been much disputed. It has been suggested that women who have terminations before their first baby may resolve their feelings adequately at the time, but when they come to have a baby later they view the termination differently and may have a resurgence of guilt (Riley, 1995). It is difficult to carry out research into the psychological effects of abortions because it is not known how women would have felt if they had not had the abortion and continued with the pregnancy. Abortions that are carried out for medical reasons as a consequence of prenatal diagnosis or screening tests may be very different from those that are carried out because of lack of social support. Women who have tests in pregnancy such as alphafetoprotein (AFP) screening for neural defects or amniocentesis for Down's syndrome may be faced with the decision to have a relatively late termination. Attitudes towards abortion may influence the uptake of prenatal screening (Marteau et al., 1992). Green and Statham (1996) describe the experience of women undergoing an abortion for a fetal abnormality as a distressing event. Both parents may experience guilt about the decision and loss of self-esteem because of the conception of a handicapped child. The implications of genetic testing in pregnancy are complex and the rapidly expanding range of genetic testing raises many issues that need to be addressed (Marteau & Richards, 1996).

8.16.5 PREGNANCY

8.16.5.1 The Impact of Pregnancy

For women, pregnancy is a time of considerable biological and psychological changes. When a woman becomes pregnant and delivers her first baby she may be admitted into hospital for the first time in her life. It may be the most significant interaction with the medical profession that she ever has. From the medical point of view it is an opportunity for formal screening and health checks. Psychological problems may be identified at this stage, but of course there are normal emotional changes occurring during pregnancy. Both men and women undergo psychological and social changes and the transition to parenthood is a major life event.

8.16.5.2 Transition to Parenthood

Pregnancy is the beginning of the transition to parenthood, although it could be argued that if a couple stop using contraception in order to conceive, they have already begun the transition. As well as the emotional changes that occur during pregnancy, there are also considerable social changes. Relationships with the couples' own parents will change, and their place in the extended family changes (especially if their baby is the first of a new generation). During pregnancy women adjust to a new identity, and this may be harder for some than others. Breen (1975) described how women come to terms with themselves as new mothers. They have to consider how they will function as mothers. Women having their first baby may have a concept of ideal motherhood. Many women may have doubts about their ability to be "good" mothers. Dimitrovsky, David-Fuchs, and Itskowitz (1989) investigated the relationships of self-ideal discrepancies and attitudes towards pregnancy. The sample were 92 middle-class primiparae who were given a semantic differential scale (Osgood, 1952) to measure their level of self-acceptance; the maternal attitudes to pregnancy instrument (Blaw, Wekowitz, & Cohen, 1964) and a measure of attitudes toward femininity, the objective social perception inventory (Kipper, Zigler-Shani, Serr, & Insler, 1977). They found no difference in the scores until the last two months of pregnancy when there was an increase in negative attitudes to femininity. However, scores on the maternal attitudes scale did not change. The more negative the self-acceptance, the more negative the attitudes to femininity. The results confirmed the distinction between levels of general self-acceptance and self-acceptance as a new mother.

8.16.5.3 Antenatal Classes

As families get smaller, less cumulative wisdom may be provided by family members and more by health professionals and the media. Antenatal classes are routinely offered to new parents in the UK and elsewhere. They attempt to increase knowledge and decrease anxiety about labor and infant care. Evaluation of these is difficult because they are offered routinely. It would be unethical to allocate women randomly to classes or nonclasses, although this was done in an early study (Huttell, Mitchell, Fischer, & Meyer, 1972) and showed a positive effect on birth outcome. There are socioeconomic differences between attenders and nonattenders. In a small study, Michie, Marteau, and Kidd (1992a) attempted to increase attendance at antenatal clinics by changing women's perceptions of the clinics. They sent out a letter to every second woman who said that they did not intend to attend the classes, but found that there was no difference in the response compared with those who had not had the letter. They conclude that increasing attendance may need a more powerful intervention. Hillier and Slade (1989) compared women choosing to attend either hospital or community classes. They assessed the women at their first antenatal class and again after their last class. The antenatal classes, which are fairly typical of classes offered routinely in the UK, consisted of eight weekly sessions, each lasting approximately two hours. They included one hour of teaching relaxation, breathing, and physical exercises to prepare for labor, and one hour for teaching and group discussion. Topics included self-care during pregnancy, labor, the first few weeks postpartum, and a tour of the maternity unit and delivery suite. As expected, knowledge levels rose, and anxiety levels fell but the changes were not significantly correlated at either the first or last assessment. Knowledge levels were related to social class at the initial assessment but not at the final assessment. This suggests that the effect of the classes was to reduce the differences related to social background and this confirmed the finding by Michie, Marteau, and Kidd (1992b) who tested the ability of Fishbein and Ajzen's theory of reasoned action (see Chapter 8.15, this volume) to predict antenatal attendance in a study of over 500 women. They found that the women's intentions to attend predicted their behavior and that perceptions of the father's attitudes were important. Walker and Erdman (1984) found that antenatal class attendance eliminated the differences in confidence and knowledge between primiparous and multiparous women. Noncompleters were more likely to have poorer educational qualifications, and to have attended their first antenatal classes later in pregnancy.

Antenatal classes give a limited input into the educational support for pregnant women. A different approach has been taken by Oakley and colleagues (Oakley, 1992). The value of social support in promoting health has been well documented in other contexts. The Social Support and Pregnancy Outcome study (SSPO) was a randomized controlled trial of social support intervention in pregnancy conducted at four centers in England in 1986–1988. Five hundred and nine mothers with a history of low birthweight babies were recruited. Half were offered a social support intervention consisting of a program of home visits and 24-hour contact with a midwife, and half had routine care. The intervention group had fewer low birthweight babies, fewer antenatal hospital admissions, nonspontaneous labor, and deliveries.

8.16.5.4 Psychological and Social Changes

Couples expecting their first child may reflect on their own experience as children and seek out the history of their own infancy. In terms of reproductive life events, pregnancy may be equally problematic for fathers. The expectant father may be a figure of fun, as portrayed in the media, but many men become closely involved with the progress of the pregnancy. Young men tend to have had less experience of babies than young women, and are less likely to have been to classes on parenthood at school than women. In a study of fathers, Lewis (1986) found that few fathers accompanied their wives to the antenatal clinic or attended parent craft classes.

The parents of the expectant couple may take on a renewed significance in their lives. Anderson, Fleming, and Steiner (1994) surveyed 30 women during pregnancy and then followed them up at eight weeks postpartum. They were asked about issues of pain, body image, attitudes towards parenting and mood during the third trimester of pregnancy, and the early postpartum period. Depressed mood (measured by the BDI) was worse during pregnancy than after the birth. Body image and pain tolerance were positively correlated with high scores on the BDI and on depressed mood (MAACL-D, Zuckerman & Lubin, 1965). Later in pregnancy the woman may give up work and this may change her social status. She may miss the social contacts, the structure and routine of work, and feel isolated from her former colleagues. Changes in body image as she increases in size may reawaken fears of being fat (Strang & Sullivan, 1985). Towards the end of pregnancy both parents focus on the impending birth, and this will include the pain associated with childbirth. There are well-recognized changes in mood during pregnancy. "Three months dreary, three months cheery, and three months weary" is a traditional saying that probably reflects the experience of many mothers. During the first trimester some women are nauseous and feel tired. They may be anxious about the implications of being pregnant, and unsure of the future. Women who have had previous problems of infertility, previous miscarriages, or who have delayed their childbearing may be particularly anxious. By the second trimester most women will have had their pregnancy confirmed and become overtly "expecting." They begin to make plans about relationships, employment, and housing. The fetus takes on an identity, and may even be given a nickname. The issues of antenatal screening become pertinent. They will be aware that their own health behavior could affect the health of their baby and may give up or reduce their smoking and alcohol intake. There is probably only a small increase in risk of mental illness during pregnancy (Oates, 1989), and episodes are often related to previous psychiatric history. The risk of developing a major mental illness (manic depressive illness or schizophrenia) is much lower during pregnancy than at other times in a woman's life. Depressive illness in the last trimester of pregnancy is related to severe postnatal depression. Oates (1989) argues for close cooperation between obstetricians and psychiatrists in the management of major mental illness in pregnancy, and suggests specific approaches for women on medication and at risk.

8.16.5.5 Changes in Sexuality

At the beginning of pregnancy, physical and psychological changes start to affect sexuality. There are physical changes in muscle tone, size, and shape of the body. Weight gain may be considerable and back pain is not uncommon. Sleep is often disturbed, especially when there is pressure on the bladder in the night, and sometimes pregnant women experience overwhelming fatigue. The change in role in the family may affect sexual feelings and there may be conflicts between image of self as a lover and potential mother (Alder, 1994; Reamy & White, 1987).

Most studies of changes in sexual activity during pregnancy have found a gradual decline over successive months, and this is followed by slow return to prepregnancy levels over the first postnatal year. Frolich et al. (1990) found a significant decline in sexuality toward the end of pregnancy and this was not related to marital status, age, or parity. In a prospective study of middle class women in London, Robson, Brant, and Kumar (1981) found little change in sexual activity during pregnancy until the third trimester. Those reporting little or no pleasure from sex before pregnancy were much more likely to have stopped having intercourse in the first trimester. Only 40% said that they found sex enjoyable in the third trimester and only 26% had orgasm. The incidence of sexual problems is difficult to estimate but 8–15% may have difficulty in achieving orgasm (Bancroft, 1989). However, this is not necessarily related to satisfaction with the sexual relationship. Masters and Johnson (1966) found that although women had been advised to avoid sexual intercourse during pregnancy, only 21 out of 71 husbands understood, agreed, or honored the advice. Today, most women are more likely to be advised to continue with sexual activity as long as it remains comfortable.

8.16.5.6 Teenage Pregnancy

High conception rates in teenagers could be due to lack of knowledge about contraception, difficulties of access to contraception, unavailability of contraception at the time of sexual intercourse, barriers to the use of contraception, or even the desire to become pregnant.

It can be assumed that most teenagers know that unprotected intercourse carries a risk of becoming pregnant, but they may underestimate the extent of that risk. There is no single reason why teenagers become pregnant and it may not be simply a lack of information or access to contraception (Woodcock, Stenner, & Ingham, 1992). Contraception may not be used the first time they have intercourse. Johnson (1994) found that of those who had had intercourse before age 16, almost half the women and over half the men reported not using contraception the first time that they had intercourse. Most research studies have concentrated on access to and knowledge of contraception and on teenage sexual behavior in investigating factors associated with teenage pregnancy (MacAlpine, Abraham, & Davies, 1995). A number of studies have considered attitudes to sexuality, changing patterns of sexual behavior, use of contraception, effect of deprivation, provision of family planning services, and sex education (SNAP, 1994). All of these factors contribute to the chances of becoming pregnant. Pregnancies could also result from positive attitudes towards pregnancy and childbearing even though the pregnancy may be later reported as being unplanned at the time of sexual intercourse. Having babies when young may be seen as conforming to a desirable social norm (Phoenix, 1989). Sutherland (1997) suggests that young women may choose to become pregnant because it demonstrates their independence, feel that it shows devotion to their boyfriend, they desire something of their own to love, they have no employment prospects, or they are worried about later infertility. These feelings need to be interpreted in the context of the young person's cultural context, and for some young people a pregnancy in late teens is a positive event.

8.16.6 THE POSTNATAL PERIOD

8.16.6.1 Changes after the Birth

The biopsychosocial model can be well illustrated by the changes to parents during the first postnatal year (Holden, 1991; Kumar & Brockington, 1988; Oates, 1989). Physical changes in the mother following the pregnancy and delivery, the effects of breastfeeding, lack of sleep, interrupted sleep, and general fatigue are all physical factors that may affect mood and well-being. The psychological changes of role, the emotional experience of being parents for the first time, and the altered dynamics in the family all make demands on the new couple. The life events of childbirth, caring for a new member of the family, and change in work role may all act as stressors. The birth of the first child may be followed rapidly by a change in financial circumstances and moving house. Even though family sizes are often small, and more women are in employment, having children is still significant in economic, psychological, and health terms. It has been suggested that maternal morbidity after delivery is extensive and under-recognized. Glazener et al. (1995) carried out a questionnaire survey of a 20% random sample of women delivering in one week ($n=1249$) at two weeks, eight weeks, and 18 months after delivery, with a response rate of 90%. Psychological problems reported included tiredness (85%, 87%, and 76%) and tearfulness/depression (16%, 21%, and 17%). These were not related to parity or method of delivery.

8.16.6.2 Postnatal Change in Mood

In the first 10 days after birth a new mother may experience a greater range of emotions than at any other time of life. She may be ecstatically happy, worried, and anxious; feel loved and esteemed; tearful and confused; she may be very excited and terribly tired (Oates, 1989) The first three days are sometimes known as "the pinks" and about the fourth or fifth day she may experience "the blues." The blues are short episodes of emotional ups and downs, and are experienced by the majority of first-time mothers. The causes of the blues are not clear.

The blues should not be confused with postnatal depression which is a "morbid and persistent depressed mood, usually commencing six to twelve weeks after delivery" (Cox, 1984, p 840). It can be diagnosed using psychiatric diagnostic criteria such as *DSM-III* or *DSM-IV*, but the somatic symptoms sometimes associated with depression such as sleep disturbance have to be considered carefully. Postnatal depression probably occurs in about 9–13% of women (O' Hara & Zekowski, 1988), and the more severe postpartum psychosis occurs in about two in 1000 deliveries (Kumar, 1989). The causes and treatment of postnatal depression have been extensively described by Cox, Holden, and Sagovsky (1994) and Riley (1995), and reviewed in a meta-analysis by O'Hara and Swain (1996).

A meta-analysis of 59 studies (O'Hara & Swain, 1996) included studies which (i) reported statistical relationships between postpartum depression and psychosocial variables, (ii) assessed the variables during pregnancy or delivery, (iii) recruited by random or quasi-random sampling, and (iv) assessed depression at least two weeks postpartum. Fifty-nine studies were included although the method of searching was not described. They found an overall prevalence rate of 0.128 (95% CI 0.123–0.134). Postnatal depression is not therefore a major problem in numerical terms and the incidence may not be much greater than that found in the general population. However, the context is very different. The presence of a new baby makes depression particularly significant, not least because expectations of fulfillment and happiness are high.

There is thus some evidence of a slight rise in depression in the puerperium and some suggestion that it differs from depression at other times of life. There are likely to be several research approaches to understanding that are likely to develop in the next few years. There are considerable advances in understanding of the biological causes. There are changes in brain metabolism that take place around childbirth, possibly in response to the dramatic changes in hormone levels. The links are difficult to establish because all women experience physiological changes during childbirth but not all become depressed. We know that sex steroids have an effect on serotonin function. They alter the levels of 5-HT1a receptors and have specific effects on dopamine transmission (Wieck, 1996). There are also peripheral effects of changing estrogen levels. Lactating women have low levels of circulating estrogen and these may be related to vaginal dryness and dysparunia (painful intercourse) after childbirth. It is possible that estrogen levels may also affect mood but although many women and their doctors may believe that changes in mood and sexuality are all due to hormone changes, the evidence is unclear (Alder, 1994).

It might seem obvious that obstetric problems would be related to postnatal depression and there have been a large number of studies that have tried to relate obstetric intervention to later depression. Dimatteo et al. (1996) carried out a meta-analysis of 358 papers comparing women with either cesarean or vaginal delivery. They found relatively few differences, although they found that those who had cesarean deliveries had a less positive reaction to the infant at six weeks, had less tactile stimulation, caretaking and intimate play with the baby, and there was a nonsignificant trend in risk of postnatal depression.

The psychosocial argument for postnatal depression has focused on social and psychological risks and vulnerability factors. O'Hara and Swain (1996) in their meta-analysis found that postnatal depression was predicted by poor marital relationships, past psychiatric history, low levels of social support, and stressful life events. They found no association with age, marital status, level of education, or number of children. It is interesting that the significant predictors are probably already known to health visitors. In a recent review, Lee (1997) argues for a social construction of postnatal depression and suggests that for the majority of women their problems lie in the social demands of motherhood and not in any individual physiological or psychological deficit. However, she ignores the physical changes in the postnatal period as described above, and the discussion on the parents' relationship does not discuss changes in sexuality.

With so much media interest it might be expected that women's understanding of postnatal depression would be relatively high. Whitton, Warner, and Appleby (1996) recruited 78 mothers into a treatment trial for postnatal depression (Appleby, Warner, Whitton, & Faragher, 1997), and asked them about their current symptoms and attitudes to treatment. Nearly all (97%) reported that they had been feeling worse than usual but only 25 (32%) believed that they were suffering from postnatal depression, even though they had reached research diagnostic criteria (Spitzer, Endicott, & Robins, 1978) for depression. The most common reasons that they gave to explain how they were feeling were tiredness (33%), family problems (17%), the baby (9%), or other children (7%). Over 80% had not reported their symptoms to any health professional.

Counseling has been a major advance in treatment. In a randomized controlled trial, Holden, Sagovky, and Cox (1989) showed that nondirective counseling given in up to eight visits of 30 minutes was effective in alleviating postnatal depression. Recovery took place in 69% of those counseled compared with 38% of those receiving routine care. This has led to the advocacy of listening visits that can be carried out by health visitors. In a major study, Cooper and Murray (1995) found no difference between nondirective counseling, psychotherapy, and cognitive-behavioral therapy.

Appleby et al. (1997) carried out a randomized controlled trial of fluoxetine or placebo plus one or six sessions of cognitive-behavioral counseling. It could be argued that to offer cognitive-behavioral therapy in one session is a contradiction in terms but the results were very positive. There were highly significant improvements in all four treatment groups. Fluoxetine was significantly better than placebo and six sessions of counseling were better than one. There was no significant interaction between counseling and fluoxetine. They concluded that both drug therapy and counseling were effective but there was no advantage in receiving both.

8.16.6.3 Postnatal Depression and Development of the Infant

Gotlib and Whiffen (1989) reported a relationship between maternal depression and low scores on the Bayley Scales of Mental Development in infants of two months. The influence of depression on the development of the child could be mediated in a number of different ways (Murray & Cooper, 1996). Depression disrupts normal communication and interaction between the mother and her baby. If the mother does not communicate and stimulate the baby, she reduces its opportunity to learn early communication skills. Prolonged exposure to the symptoms of maternal depression results in learned dysfunctional behavior. The model presented to the child will be of a noncommunicating and withdrawn adult and children easily imitate adults' behavior (Cooper & Murray, 1997). Alternatively, social factors could explain variations in child development. The depression could be related to some other factor such as social deprivation or abuse in the family.

Field, Healy, Goldstein, and Gutherz (1990) assessed infant behavior in depressed, disadvantaged mothers and found that their infants tended to be withdrawn and lacked positive affect. Murray, Fiori-Cowley, Hooper, and Cooper (1996) compared healthy children of a sample of 29 mothers from a maternity hospital who were depressed (RDC) with 20 children of control mothers who were not depressed. They observed face-to-face interactions at two months and they found few differences between the two groups. However, there was some effect on older 18-month infants. Murray and Cooper (1996) suggest that the most likely mechanism is impaired pattern of communication occurring between the mother and child in the context of depression, rather than the child's exposure to depression *per se*. The effect is likely to be an interaction, because an uncommunicative child may affect the mother's behavior towards it. Alder and Thome (1994) found that those mothers scoring high on the Edinburgh Postnatal Depression Scale (Cox, Holden, & Sagovsky, 1987) were also more likely to report infant difficulties. Having a difficult infant could also increase the likelihood of becoming depressed, or at least feeling fatigued and low in maternal self-efficacy.

Pop et al. (1995) found an incidence of postpartum depression of 21% in their sample of 293 women using RDC (Spitzer, Endicott, & Robins, 1978). In their logistic regression analysis, only major depression contributed significantly to intelligence quotient (measured in five-year old children by three scales of the McCarthy Scales of Children's Abilities, MSCA), with an odds ratio of 7.0 (95% CI = 1.6–43, $P = 0.04$).

8.16.6.4 Breast-feeding

8.16.6.4.1 Breast-feeding as a health behavior

Breast-feeding can be regarded as a health behavior in the sense that it has a direct effect on health. Antibodies from the mother are passed through the breast milk to the infant, and the composition of human breast milk may be superior to artificial formula feed. Infant feeding behavior has been analyzed using social cognition models but breast-feeding is a complex behavior pattern involving relationships as well as health beliefs (Alder, 1989). Breast-feeding benefits the health of the baby (Howie, Forsyth, Ogston, Clark, & Florey, 1990), although it is unusual in that it may have positive benefits for the child but has costs for the mother. Breast-feeding can be painful, physically tiring, and emotionally demanding; it is also very satisfying and fulfilling.

Mothers are encouraged to breast-feed, and most first-time mothers attempt to breast-feed. In the UK, the continuation rates are very low and only about 30% of mothers are breast-feeding at three months postpartum. The incidence is strongly related to socioeconomic status (White, Freeth, & O'Brien, 1992), and in the UK mothers from lower socioeconomic groups are less likely to breast-feed. The situation is different in developing countries and urban mothers may be more likely to use formula feeds (Foster, Lader, & Cheeseborough, 1997).

The effect of the method of feeding on the development of the child could result from either the milk itself (breast milk or formula) or from the feeding behavior, or both. In early reports of an association between breast-feeding and intelligence, the association was confounded with socioeconomic factors. Florey, Leech, and Blackhall (1995) found that children who were breast-fed on discharge from hospital had an average value on the Bayley scales of between four and six points higher than formula-fed children. Experimental trials of infant feeding in healthy infants are not possible, but Morely, Cole, Powell, and Lucas (1988) investigated the effect of breast milk given to preterm children by bottle. Children of mothers who chose to provide breast milk for their preterm infants in addition to formula were found to have an eight-point increase on the Bayley scales compared with those receiving formula only. However, those mothers choosing to provide breast milk might be different in their social characteristics. After controlling for socioeconomic factors, the difference was still 4.3 points, and these differences were maintained at 7.5–8 years of age when assessed by the Weschler Intelligence Test. However, Doyle, Rickards, Kelly, Ford, and Callanan (1992) collected data in a prospective study of low birthweight babies and controlled for social factors including social class, maternal education, age of mother, and gender of baby. They found no significant differences in IQ scores of seven-year-old children between those who had been given expressed breast milk and those who had formula only. Children who had received direct breast-feeding had higher IQ scores, which suggests that breast-feeding may have a direct effect on the child or an indirect effect via the behavior of the mother.

The behavior of the infant when breast-fed is very different from that of the formula-fed baby not only in the sucking pattern. Wright (1981) found that breast-fed infants have different feeding patterns from formula-fed babies. Breast-fed babies at two months have the largest size of feed in the morning and the feed size decreases during the day. Bottle-fed babies tend to have constant-sized feeds and mothers may attempt to encourage the baby to finish the bottle. As night feeds are dropped, breast-fed babies go on to take larger feeds at the end of the day. Wright (1988) suggested that this is an example of anticipatory learning on the part of breast-fed infants. They have more control over the pace and duration of breast-feeding, and this may increase their expectations of mastery. The opportunity to learn (in the context of a biologically important behavior) is not available to formula-fed babies. They tend to be offered the same size of formula at each feed although appetite changes over the day, and mothers may well attempt to control the amount of milk consumed.

8.16.6.4.2 Breast-feeding and the mother

Breast-feeding mothers feed more often and more often in the night than formula-feeding mothers (Elias, Nicolson, Brora, & Johnston, 1986; Romito, 1988). Although breast-feeding mothers may be more fatigued, less sexually active, and report more pain with intercourse, there appears to be no relationship with ratings of clinical depression (Paykel, Emms, Fletcher, & Rassaby, 1980). It is possible that the influence of breast-feeding on depression, whether hormonal or social, is small, that psychiatric diagnoses of depression do not identify postpartum distress, that there are too many confounding variables that protect breast-feeding mothers from depression, or that negative associations have suffered from reporting bias.

Being depressed may affect the progress of breast-feeding. Tamminen (1989), in a cross-sectional study in Finland, surveyed four groups of mothers: (group 1 in late pregnancy ($n = 10$), group 2 just after delivery ($n = 17$), group 3 with babies 2–4 months ($n = 30$), and group 4 with babies 6–12 months ($n = 33$). Most of the delivered mothers were currently breast-feeding (group 2, 90%, group 3, 97%, group 4, 50%). Eight percent of mothers were depressed and they reported more breast-feeding difficulties than nondepressed women. Successful breast-feeding can have a positive influence on their mood, whereas nonsuccessful breast-feeding may have a negative influence on mood. If breast-feeding and depression are interrelated then this could be an indirect route for the effect of breast-feeding on infant development.

8.16.6.4.3 Breast-feeding and sexuality

Breast-feeding mothers may be more fatigued and are less interested in sexual activity (Alder & Bancroft, 1988). Breast-feeding mothers, whose babies persisted in waking up for a night feed, took longer to resume sexual intercourse with their partners than those who slept through the night (Alder & Bancroft, 1988). There may be psychological differences between those who persist in breast-feeding for several months, those who never attempt breast-feeding, and those who begin breast-feeding in hospital, but discontinue when they go home. Some women may give up breast-feeding in order to resume an active sex life, because they are more sexually interested or feel the need to resume sexual activity for the sake of the relationship. Some women who discontinue may be inhibited about their bodies and dislike physical contact. The relationship between breast-feeding and sexuality has been investigated in a number of studies (Alder & Bancroft 1988; Alder, Cook, Davidson, West, & Bancroft, 1986), but it is a very complex issue (Alder, 1989).

The effects of breast-feeding on sexuality could be a direct effect of endocrine changes on the central nervous system mediated in some way not yet understood. There has been little progress in establishing an endocrinological basis for changes in sexual interest, although Alder et al. (1986) found that lactating women with reduced sexual interest after delivery had lower testosterone levels than those with no reduction in interest. There could be a direct effect of the lowered level of circulating oestrogen on the vaginal epithelium, vaginal dryness, and possible effects of perineal pain could inhibit sexual intercourse. Alder et al. (1986) found that more perineal pain was reported in breast-feeding women. Fatigue may play a part. Breast-feeding women are more likely to have infants that wake in the night and they persist for longer. Glazener (1997) found that women who were breast-feeding were more likely to report a lack of interest in intercourse than those who were bottle feeding at eight weeks, but not at 12–18 months. This difference held even when controlled for tiredness and depression.

8.16.7 THE OLDER MAN AND WOMAN

It may seem that reproductive events are only relevant to the reproductive years, but the impact of reproduction lasts well into old age. The age at which women are capable of bearing children has now been extended beyond the menopause. The age of birth of the first child may have increased in Western countries but there is no evidence of a decline in parenting ability with age. Berryman (1991) concluded that older mothers were well adjusted and coped well with later childbearing. The menopause is discussed in Chapter 8.15, this volume. Men experience a gradual decline in androgen function in later life. There may be direct effects of endocrine changes on health conditions such as cardiovascular disease or prostrate hyperplasia, or indirectly on sexual function as part of the aging process. Osteoporosis in women is clearly related to decline in estrogen levels and there is a well-established relationship between bone mineral density and fractures. The social and psychological consequences of fractures in old age and consequent loss of mobility and independence pose a challenge to the health professions as the number of elderly increase as a proportion of the population. In men the increase in fractures rises in the late seventies, whereas for women there is a gradual rise following the menopause (Donaldson, Cook, & Thomson, 1990). There have been suggestions that estrogens and androgens, whose levels decline with age, influence cognitive functioning. Experimental studies show that estrogen and androgen therapy in women and testosterone treatment in men have cognitive effects (Christiansen, 1996; Sherwin, 1988).

8.16.8 PROFESSIONAL ISSUES

Psychologists may be directly involved with reproductive issues in, for example, abortion counseling, or indirectly involved when the reproductive status of the man or woman is relevant to the presentation of the problem or its management.

An understanding of basic reproductive physiology is fundamental to understanding the influence of reproductive and obstetric issues on psychological health. For instance, breast-feeding is controlled by a complex hormonal mechanism. In order to understand the changes that occur in mid-age it is necessary to understand the changes in the ovarian cycle that occur in the menopause.

The practice of medicine is closely involved with reproductive issues, and pregnancy and childbirth may be the first occasion on which there is any lengthy contact with the medical profession. Medical intervention moves into the most intimate and private areas of people's lives. Direct medical intervention takes place during childbirth. Medical control over prescribing oral contraception may influence sexual behavior. Cervical screening and breast screening take place in a medical context but they have sexual and psychological implications.

Much of the research interest in contraceptive use has come from clinicians concerned with infectious disease and the transmission of HIV infection (Chapter 8.15, this volume). However, the incidence of pregnancies in young teenagers has also focused attention on contraceptive use in young people.

Health psychologists may be involved in planning interventions. Preparation for childbirth is widely practiced and relaxation techniques are learned by many women in antenatal classes. The environment of special baby units has been shown to influence the behavior of the babies and the stress on medical staff and parents can be modified by using principles of health psychology (Padden & Glenn, 1997).

Social cognition models that are so widely used in health psychology have been applied to reproductive issues such as uptake of breast screening, contraceptive behavior, screening for genetic disorders, or prenatal diagnosis. General health issues may have special relevance in relation to obstetrics. Smoking and alcohol use in pregnancy may affect the health of the fetus. Health behaviors may be modified because of the reproductive context. Reasons why people do not exercise during pregnancy are different from reasons for not exercising in adolescence or old age. The concept of loss has been explored in studies of miscarriage and stillbirth. At a time when we expect to celebrate new life, it is particularly difficult to cope with loss and bereavement. The grief process will be influenced by the reproductive context. It has been suggested that couples with fertility problems may go through a grief process or mourning for the children they may never have.

Reproductive issues have attracted attention from health psychologists although often in clinical contexts or in attempts to explain and predict health behavior. Broome and Wallace (1984) made an early attempt to address the inter-relationships of psychology and reproductive issues and Niven and Walker (1996) have taken a broad approach to psychological aspects of reproduction.

The influence of reproduction continues throughout life and pervades all aspects of life. Psychologists, clinical and health, will benefit from considering its implications for health.

8.16.9 REFERENCES

Abraham, S. C. S., & Sheeran, P. (1994). Modelling and modifying young heterosexual's HIV-preventive behaviour; a review of theories, findings and educational implications. *Patient Education and Counselling, 23,* 173–186.

Adler, N. E., David, H. P., Major, B. N., Roth, S. H., Ruddo, N. F., & Wyatt, G. E. (1990). Psychological responses after abortion. *Science, 248,* 41–44.

Alder, B. (1993). Contraception. In C. Niven & D. Carroll (Eds.), *The health psychology of women.* Chur, Switzerland: Harwood Academic Publishers.

Alder, B., & Thome, M. (1994). Parental distress in Iceland. *British Psychological Society Scottish Branch Annual Conference*, Crieff.

Alder, E. (1984). Sterilisation. In A. Broome & L. Wallace (Eds.), *Psychology and gynaecological problems.* London: Tavistock.

Alder, E. M. (1984). Psychological aspects of AID. In A. Emery & I. Pullen (Eds.), *Psychological aspects of genetic counselling.* London: Academic Press.

Alder, E. M. (1989). Sexual behaviour in pregnancy, after childbirth and during breast-feeding. *Bailliere's Clinical Obstetrics and Gynaecology, 3,* 805–821.

Alder, E. M. (1994). Post natal sexuality. In P. Choi & N. Nicolson (Eds.), *Female sexuality: psychology, biology and social contact*: London: Harvester Wheatsheaf.

Alder, E. M., & Bancroft, J. (1988). The relationship between breast-feeding persistence, sexuality and mood in post partum women. *Psychological Medicine, 18,* 389–396.

Alder, E. M., Cook, A., Davidson, D., West, C., & Bancroft, J. (1986). Hormones, mood and sexuality in lactating women. *British Journal of Psychiatry, 148,* 74–79.

Alder, E. M., Cook, A., Gray, J., Tyrer, G., Warner, P., Bancroft, J.,& Loudon, N. B., & Loudon, J. (1981). The effects of sterilisation: A comparison of sterilised women with the wives of vasectomised men. *Contraception, 23,* 45–54.

Alder, E. M., & Templeton, A. A. (1985). Patient reaction to IVF treatment. (letter) *Lancet, i,* 8.

Anderson, V. N., Fleming, A. S., & Steiner, M. (1994). Mood and the transition to motherhood. *Journal of Reproductive and Infant Psychology, 12,* 69–78.

Appleby, L., Warner, R., Whitton, A., & Faragher, B. (1997). A controlled study of fluoxetine and cognitive behavioural counselling in the treatment of postnatal depression. *British Medical Journal, 14,* 932–936.

Athanasiou, R., Oppel, W., Michaelson, L., Unger, T., & Yager, M. (1973). Psychiatric sequalae to term birth and induced early and late abortion: A longitudinal study. *Family Planning Perspective, 5,* 227.

Ajzen, I., & Fishbein, B. (1980). *Understanding attitudes and predicting behavior.* Englewood Cliffs, NJ: Prentice-Hall.

Bancroft, J. (1989). *Human sexuality and its problems* (2nd ed.). Edinburgh: Churchill Livingstone.

Bancroft, J., & Sartorius, N. (1989). The effects of oral contraceptives in well-being and sexuality. *Oxford Review of Reproductive Biology, 12,* 57–92.

Bem, S. (1974). The measurement of psychological androgyny. *Journal of Clinical and Consulting Psychology, 42,* 155–162.

Bell, J. S. (1981). Psychological problems among patients attending an infertility clinic. *Journal of Psychosomatic Research, 25,* 1–3.

Bell, J. S., & Alder, E. (1994). Psychology of infertility and management. In T. M. Hargreaves (Ed.), *Male infertility* (2nd ed.). New York: Springer.

Berg, B. J., & Wilson, J. F. (1995). Patterns of psychological distress in infertile couples. *Journal of Psychosomatic Obstetrics and Gynaecology, 16,* 65–78.

Berryman, J. (1991). Perspectives on later motherhood. In A. Phoenix, A. Woollett, & E. Lloyd (Eds.), *Motherhood: Meanings, practices and ideologies.* London: Sage.

Blaw, A., Wekowitz, I., & Cohen, I. (1964). Maternal attitudes to pregnancy instrument. *Archives of General Psychiatry, 10,* 324–331.

Blyth, E. (1995). "Not a primrose path": Commissioning parents' experiences of surrogacy arrangements in Britain. *Journal of Reproductive and Infant Psychology, 13,* 185–196.

Breen, D. (1975). *The birth of a first child: Towards and understanding of femininity*. London: Tavistock.

Broome, A., & Wallace, L. (1984). *Psychology and gynaecological problems*. London: Tavistock.

Burns, L. H. (1995). An overview of sexual dysfunction in the infertile couple. *Journal of Family Psychotherapy, 6,* 25–46.

Campbell, E. (1985). *The childless marriage. An exploratory study of couples who do not want children*. London: Tavistock.

Carroll, D., Niven, C., & Sheffield, D. (1993). Gender, social circumstances and health. In C. Niven & D. Carroll (Eds.), *The health psychology of women*. Chur: Switzerland: Harwood Academic Publishers

Choi, P. Y. L. (1994). Women's raging hormones. In P. Y. L. Choi & P. Nicolson (Eds.), *Female sexuality: Psychology, biology and social context*. London: Harvester Wheatsheaf.

Christiansen, K. (1996). Androgens and sexual functioning and mood in men. In B. Oddens & A. Vermuelen (Eds.), *Androgens and the aging male*. London: Parthenon.

Cliquet, R. L., Thiery, M., Staelens, R., & Lambert, G. (1981). Voluntary sterilization in Flanders. *Journal of Biosocial Science, 13,* 47–61.

Cooper, P., & Murray, L. (1995). *The impact of postnatal depression on infant development: A treatment trial*. Proceedings of the Cape Town Conference on Infant Psychiatry.

Coleman, P. M. (1993). Depression during the female climacteric period. *Journal of Advanced Nursing, 18,* 1540–1546.

Condon, J. T., Need, J. A., Fitzsimmons, D., & Lucy, S. (1995). University students' subjective experiences of oral contraceptive use. *Journal of Psychosomatic Obstetrics and Gynecology, 16,* 37–43.

Conner, M., & Norman, P. (1996). *Predicting health behaviour*. Buckingham, UK: Open University Press.

Connolly, K. J., Edelmann, R. J., & Cooke, I. D. (1987). Distress and marital problems associated with infertility. *Journal of Reproductive and Infant Psychology, 5,* 49–57.

Cooper, P., Gath, D., Rose, N., & Fieldsend, R. (1982). Psychological sequalae to elective sterilisation in women: a prospective study. *British Medical Journal, 284,* 461–464.

Cooper, P. J., & Murray, L. (1997). Prediction, detection and treatment of postnatal depression. *Archives of Disease in Childhood, 77,* 97–101.

Cox, J. L. (1986). *Postnatal depression. A guide for health professionals*. Edinburgh: Churchill Livingstone.

Cox, J. L., Holden, J., & Sagovsky, R. (1987). Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal depression scale. *British Journal of Psychiatry, 150,* 782–786.

Cuisinier, M. K., Janssen, H., de Graauw, C., Bakker, S., & Hoogduin, C. (1996). Pregnancy following miscarriage: Course of grief and some determining factors. *Journal of Psychosomatic Obstetrics and Gynaecology, 17,* 168–174.

Daniels, K. R. (1989). Semen donors: Their motivations and attitudes towards their offspring. *Journal of Reproductive and Infant Psychology, 7,* 121–127.

Daniels, K., & Lewis, G. M. (1996). Openness of information in the use of donor gametes: Developments in New Zealand. *Journal of Reproductive and Infant Psychology, 14,* 57–68.

Davis, D. L., Stewart, M., & Harmon, R. J. (1989). Postponing pregnancy after perinatal death: Perspectives on doctors' advice. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 481–487.

Demick, J., & Wapner, S. (1988). Open and closed adoption: a developmental conceptualization. *Family Process, 27, 229–249.*

Demyttenaere, K., Nijs, P., Steeno, O., Koninckx, P., & Evers-Kiebooms, G. (1988). Anxiety and conception rates in donor insemination. *Journal of Psychosomatic Obstetrics and Gynaecology, 8,* 175–181.

Derogatis, L. R. (1977). The SCL-90-R manual. A step in the validation of a new self-report scale. *British Journal of Psychiatry, 128,* 280–290.

Dimatteo, M. R., Lepper, H. S., Damush, T. M., Morton, S. C., Carney, M. F., Pearson, M., & Kahn, K. L. (1996). Cesarean childbirth and psychosocial outcomes a meta analysis. *Health Psychology, 15,* 303–314.

Dimitrovsky, L., David-Fuchs, M., & Itskowitz, R. (1989). Self-acceptance and attitudes towards pregnancy and femininity among primiparae. *Journal of Reproductive and Infant Psychology 7,* 203–210.

Donaldson, L. J., Cook, A., & Thomson, R. G. (1990). Incidence of fractures in a geographically defined population. *Journal of Epidemiology and Community Health, 44,* 241–245.

Doyle, L. W., Rickards, A., Kelly, E. A., Ford, G. W., & Callanan, C. (1992). Breastfeeding and Intelligence. *Lancet, 339,* 744–745.

Edelmann, R. J., & Connolly, K. J. (1996). Sex role and emotional functioning in infertile couples: Some further evidence. *Journal of Reproductive and Infant Psychology, 14,* 113–119.

Edelmann, R. J. & Golombok, S. (1989). Stress and reproductive failure. *Journal of Reproductive and Infant Psychology, 7,* 79–93.

Elias, M., Nicolson, N., Brora, C., & Johnston, J. (1986). Sleep/wake patterns of breastfed infants in the first 2 years of life. *Pediatrics, 77,* 322–329.

Field, T., Healy, B., Goldstein, S., & Gutherz, M. (1990). Behavior-state matching and synchrony in mother infant interactions of non depressed versus depressed dyads. *Developmental Psychology, 26,* 7–14.

Florey, C. du V., Leech, A. M., & Blackhall, A. (1995). Infant feeding and mental and motor development at 18 months of age in first born singletons. *International Journal of Epidemiology, 24*(Suppl.1), S21–S26.

Foster, K., Lader, D., & Cheeseborough, S. (1997). *Infant feeding 1995*. London: The Stationery Office.

Frohlich, E. P., Herz, C., van der Merwe, F. J., van Tonder, D. M., Booysen, J. P. M., & Becker, P. J. (1990). Sexuality during pregnancy and early puerperium and its perception by the pregnant and puerperal woman. *Journal of Psychosomatic Obstetrics and Gynaecology, 11,* 73–80.

Friedman, T. (1989). Infertility and assisted reproduction. *Bailliere's Clinical Obstetrics and Gynaecology, 3,* 751–767.

Friedman, T. (1989). The psychiatric consequences of spontaneous abortion. *British Journal of Psychiatry, 146,* 55–61.

Galavotti, C., Cabral, R. J., Lansky, A., Grimely, D. M., Riley, G. E., & Prochaska, J. (1995). Validation of measures of condom use and other contraceptive use among women at high risk for HIV infection and unintended pregnancy. *Health Psychology, 14,* 570–578.

George, J., Turner, J., Cooke, E., Hennessy, E., Savage, W., Julian, P., & Cochrane, R. (1994). Women's knowledge of emergency contraception. *British Journal of General Practice, 44,* 451–454.

Glasier, A. (1993). Emergency contraception: Time for deregulation? *British Journal of Obstetrics and Gynaecology, 100,* 611–612.

Glazener, C. M. A. (1997). Sexual function after childbirth: Women's experiences, persistent morbidity and lack of professional recognition. *British Journal of Obstetrics and Gynaecology, 104,* 330–335.

Glazener, C. M. A., Abdalla, M., Stroud, P., Naji, S., Templeton, A., & Russell, I. T. (1995). Postnatal maternal morbidity; extent, causes, prevention and treatment. *British Journal of Obstetrics and Gynaecology, 102,* 282–287.

Goldberg, D. (1972). *The detection of psychiatric illness by questionnaire*. London, New York: Oxford University Press.

Gotlib, I. H., & Whiffen, V. E. (1989). Depression and marital functioning: An examination of specificity and gender difference. *Journal of Abnormal Psychology, 98,* 23–30.

Green, J., & Statham, H. (1996). Psychosocial aspects of prenatal screening and diagnosis. In T. Marteau & M. Richards (Eds.), *The troubled helix: Social and psychological implications of the new human genetics*. Cambridge: Cambridge University Press.

Greenfield, D. A., Diamond, M., & DeCherney, A. H. (1988). *Journal of Psychosomatic Obstetrics and Gynecology, 8,* 169–174.

Grimely, D. M., Prochaska, J. O., Velicer W. F., & Prochaska, G. E. (1995). Contraceptive and condom use adoption and maintenance: A stage paradigm approach. *Health Education Quarterly, 22,* 20–35.

Guillebaud, J. (1991). *The pill*. Oxford: Oxford Publishers.

Hargreave, T. M. (1994). Vasectomy. In T. M. Hargreave (Ed.), *Male infertility* (2nd ed.). London: Springer.

Harper, R., Lenton, E. A., & Cooke, I. D. (1985). Prolactin and subjective reports of stress in women attending infertility treatment: effect on semen quality relationships. *Journal of Reproductive and Infant Pschology,* 3–8.

Harrison, K. L., Callan, V. J., & Hennessey, J. F. (1987). Stress and semen quality in an *in vitro* fertilisation program. *Fertility and Sterility, 48,* 633–636.

Harvey, S. M., Beckman, L. J., & Murray, J. (1991). Perceived contraceptive attributes and method choice. *Journal of Applied Social Psychology, 21,* 774–790.

Hillier, C. A., & Slade, P. (1989). The impact of antenatal classes on knowledge, anxiety and confidence in primiparous women. *Journal of Reproductive and Infant Psychology, 7,* 3–13.

Holden, J. (1991). Postnatal depression: Its nature, effects and identification using the Edinburgh Postnatal depression scale. *Birth, 18,* 211–221.

Holden, J. M., Sagovsky, R., & Cox, J. L. (1989). Counselling in a general practice setting: Controlled study of health visitor intervention in treatment of postnatal depression. *British Medical Journal, 298,* 223–226.

House, J. S., Robbins, C., & Metzner, H. L. (1982). The association of social relationships and activities with mortality: Prospective evidence from the Tecumseh community health study. *American Journal of Epidemiology, 116,* 123–140.

Howie, P. W., Forsyth, J. G., Ogston, S. A., Clarke, A., & Florey, C. du V. (1990). Protective effect of breastfeeding against infection. *British Medical Journal, 300,* 11–16.

Humphrey, M., & Humphrey, H. (1988). *Families with a difference: Varieties of surrogate parenthood.* London: Routledge.

Hull, M. G., Glazener C. M., Kelly, N. J., Conway, D. I., Foster, P. A., Hinton, R. A., Coulson, C., Lambert, P. A., Watt, E. M., & Desai, K. M. (1985). Population study of causes, treatment and outcome of infertility. *British Medical Journal of Clinical Research Education, 291,* 1693–1697.

Hunt, K. (1990). The first pill-taking generation: past and present use of contraception amongst a cohort of women born in the early 1950s. *The British Journal of Family Planning, 16,* 3–15.

Hunter, M. (1994). *Counselling in obstetrics and gynaecology*. Leicester, UK: BPS Books.

Huttel, F. A., Mitchell, I., Fisher, M. W., & Meyer, A. E. (1972). A quantitative evaluation of prophylaxis in childbirth. *Journal of Psychosomatic Research, 16,* 81.

Iles, S. (1989). The loss of early pregnancy *Balliere's Clinical Obstetrics and Gynaecology, 3,* 769–790.

Irvine, S., & Cawood, E. (1996). Male infertility and its effect on male sexuality. *Sexual and Marital Therapy, 11,* 273–280.

Jackman, C., McGee, H. M., & Turner, M. (1991). The experience and psychological impact of early miscarriage. *The Irish Journal of Psychology, 12,* 108–120.

James, P. D. (1992). *The children of men*. Harmondsworth, UK: Penguin.

Janssen, H. J., Cuisinier M. C., de Graaw, K. P., & Hoogduin, K. A. (1997). A prospective study of risk factors predicting grief intensity following pregnancy loss. *Archives of General Psychiatry, 54,* 56–61.

Jarvis, M. J. (1994). A profile of tobacco smoking. *Addiction, 89,* 1371–1376.

Johnson, A. (1994). *Sexual attitudes and lifestyles*. Oxford: Blackwell.

Johnson, M. P., & Puddifoot, J. E. (1996). The grief response in the partners of women who miscarry. *British Journal of Medical Psychology, 69,* 313–327.

Johnson, M. P., & Puddifoot, J. E. (1996). The grief response of women who miscarry. *British Journal of Medical Psychology, 69,* 313–327.

Johnston, M., Shaw R., & Bird, D. (1987). Test tube Baby procedures: Stress and judgements under uncertainty. *Psychology and Health, 1,* 25–38.

Jones, S. C., & Hunter, M. (1996). The influence of context and discourse on infertility experience. *Journal of Reproductive and Infant Psychology, 14,* 93–111.

Kipper, D. A., Zigler-Shani, Z., Serr, D. M., & Insler, V. (1977). Psychogenic infertility, neuroticism and the feminine role: A methodological inquiry. *Journal of Psychosomatic Research, 21,* 353–358.

Kubler Ross, E. (1969). *On death and dying*. New York: Macmillan.

Kumar, R. (1989). Postpartum psychosis. *Bailliere's Clinical Obstetrics and Gynaecology, 3*, 823–838.

Kumar, R., & Brockington, I. F. (1988). *Motherhood and mental illness 2: Causes and consequences*. London: Wright.

Lampman, C., & Dowling-Guyer, S. (1995). Attitudes towards voluntary and involuntary childlessness. *Basic and Applied Social Psychology, 17,* 213–222.

Lapane, K. L., Zeirler, S., Lasater, T. M., Stein, M., Barbour, M. M., & Hume, A. L. (1995). Is a history of depressive symptoms associated with an increase of infertility in women? *Psychosomatic Medicine, 57,* 509–513.

Lambers, K. J., Trimbos-Kemprer, G. C. M, & van Hall, E. V. (1984). Regret and reversal of sterlization. In A. Broome & L. Wallace (Eds.), *Psychology and gynaecological problems*. London: Chapman Hall.

Lawson, S., Cole, R. A., & Templeton, A. A. (1979). The effects of laparascopic sterilisation by diathermy or silastic bands on post-operative pain menstrual symptoms and sexuality. *British Journal of Obstetrics and Gynaecology, 86,* 659–663.

Lee, C. (1997). Social context, depression and the transition to motherhood. *British Journal of Health Psychology, 2,* 93–108.

Leppert, P. C., & Pahlka, B. S. (1984). Grieving characteristics after spontaneous abortion: A management approach. *Obstetrics and Gynaecology, 64,* 119–122.

Lewis, C. (1986). *Becoming a father*. Milton Keynes, UK: Open University Press.

Lindemann, C. (1977). Factors affecting the use of contraceptives in the non marital context. In R. Gemme & C. Wheeler (Eds.), *Progress in sexology*. New York: Plenum.

Mahoney, C. A., Thombs, D. L., & Ford, O. J. (1995). Health belief and self efficacy models: Their utility in explaining college student condom use. *AIDS Education and Prevention, 7,* 32–49.

Marteau, T. M., Johnston, M., Kidd, J., Michie, S., Cook, R., Slack, J., & Shaw, R. W. (1992). Psychological models in predicting uptake of prenatal screening. *Psychology and Health, 6,* 13–22.

Marteau, T., & Richards, M. (Eds.) (1996). *The troubled helix. Social and psychological implications of the new human genetics*. Cambridge: Cambridge University Press.

Martin, R. M., Hilton, S. R., & Kerry, S. M. (1997). The impact of the October 1995 "pill scare" on oral contraceptive use in the United Kingdom: Analysis of a general practice automated database. *Family Practice, 14,* 279–284.

Masters, W. H., & Johnson, V. E. (1966). *Human sexual response*. Toronto, London: Little Brown.

McAlpine, S., Abraham, C., & Davies, H. (1995). Promoting contraceptive use amongst young people: a brief review. Report to Tayside Health Board. Dundee: University of Dundee.

McCabe, E., Golub, S., & Lee, A. C. (1997). Making the female condom a "reality" for female adolescents. *Journal of Pediatric Adolescent Gynaecology, 10,* 15–23.

McCormick, A., Rosenbaum, M., & Fleming, D. (1990). Socio-economic characteristics of people who consult their general practitioners. *Population Trends, 59,* 8–10.

McMahon, C. A., Ungerer, J. A., Beaurepaire, J., Pennant, C., & Saunders, D. (1995). Psychosocial outcomes for parents and children after *in vitro* fertilisation: A review. *Journal of Reproductive and Infant Psychology, 13,* 1–16.

Michie, S., Marteau, T. M., & Kidd, J. (1992a). Predicting antenatal class attendance. Attitudes of self and others. *Psychology and Health, 7,* 225–234.

Michie, S., Marteau, T. M., & Kidd, J. (1992b). An evaluation of an intervention to increase antenatal class attendance. *Journal of Reproductive and Infant Psychology, 10,* 183–185.

Miller, W. B., & Pasta, D. (1995). Behavioral intentions: Which ones predict infertility behavior in married couples. *Journal of Applied Social Psychology, 25,* 530–555.

Miller, W. B., & Pasta, D. (1996). The relative influence of husbands and wives on the choice and use of contraception, a diaphragm and condoms. *Journal of Applied Social Psychology, 26,* 1749–1774.

Morely, R., Cole, T. J., Powell, R., & Lucas, A. (1988). Mothers choice to provide breastmilk and developmental outcome. *Archives of Disease in Childhood, 63,* 1382–1385.

Murray, L., & Cooper, P. J. (1996). The impact of post partum depression on child development. *International Review of Psychiatry,* 55–63.

Murray, I. L., Fiori-Cowley, A., Hooper, R., & Cooper, P. J. (in press). The impact of post natal depression and associated adversity on early mother infant interactions and later infant outcome. *Child Development*.

Niven, C., & Carroll, D. (Eds.) (1993). *The health psychology of women*. Chur, Switzerland: Harwood Academic Publishers.

Niven, C., & Walker, A. (Eds.) (1996). *The Psychology of reproduction Vol 2. Conception, pregnancy and birth.* Oxford: Butterworth Heineman.

Oakley, A. (1992). *Social support and motherhood: The natural history of a research project*. Oxford, UK: Blackwell.

Oates, M. (1989). Management of major mental illness in pregnancy and the puerperium. *Bailliere's Clinical Obstetrics and Gynaecology, 3,* 905–920.

Oates, M., & Gath, D. (1989). Psychological aspects of gynaecological surgery. *Bailliere's Clinical Obstetrics and Gynaecology, 3,* 729–749.

Osgood, C. E. (1952). The nature of measurement and meaning. *Psychological Bulletin, 49,* 197–237.

Oddens, B. J. (1996). *Determinants of contraceptive use.* Netherlands: Euburon.

Oddens, B. J., Visser, A. P., Vemer, H. M., Everaerd, W. T. A. M., & Lehert, P. (1994a). Contraceptive use and attitudes in Great Britian. *Contraception, 49,* 73–86.

Oddens, B. J., Visser, A. P., Vemer, H. M. E., Everaerd, W. T. A. M. (1994b). Contraceptives and attitudes in reunified Germany. *European Journal of Obstetrics and Gynaecology and Reproductive Biology, 57,* 3201–3208.

O'Hara, M. W., & Swain, A. M. (1996). Rates and risks of postpartum depression: A meta-analysis. *International Review of Psychiatry, 8,* 37–54.

O'Hara, M. W., & Zekoski, E. M. (1988). Post partum depression: Causes and consequences, a comprehensive review. In R. Kumar & I. F. Brockington (Eds.), *Motherhood and mental illness 2*. London: Wright.

O'Moore, M. A., & Harrison, R. F. (1991). Anxiety and reproductive failure: Experiences from a Dublin infertility clinic. *The Irish Journal of Psychology, 12,* 276–285.

Padden, T., & Glenn, S. (1997). Maternal experiences of preterm birth and neonatal intensive care. *Journal of Reproductive and Infant Psychology, 15,* 121–139.

Parkes, C. M. (1985). Bereavement. *British Journal of Psychiatry, 146,* 11–17.

Paykel, E. S., Emms, E. M., Fletcher, J., & Rassaby, E. S. (1980). Life events and social support in puerperal depression. *British Journal of Psychiatry, 136,* 339–346.

Pearlin, L. I., Mullen, J. T., Semple, S., & Skaff, M. B. (1990). Caregiving and the stress process: an overview of concepts and their measurement. *The Gerontologist, 30,* 583–594.

Persky, V. W., Kempthornerawson, J., & Shekelle, R. B. (1987). Personality and risk of cancer 20-year follow up of the Western Electric study. *Psychosomatic Medicine, 49,* 435–449.

Pfeffer, N., & Woollett, A. (1983). *The experience of infertility*. London: Virago.

Pheonix, A. (1989). Influences on previous contraceptive use /non use in pregnant 16–19 year olds. *Journal of Reproductive and Infant Psychology, 7,* 211–225.

Pop, V. J., de Rooy, H. A., Vader, H. L., et al. (1991). Postpartum thyroid dysfunction and depression in an unselected population. *New England Journal of Medicine, 324,* 1815–1816.

Pop, V. J., de Vries, E., van Baar, A. L., Waelkens, J. J., de-Rooy, H. A., Horsten, M., Donkers, M. M., Komproe, I. H., van Son, M. M., & Vader, H. L. (1995). Maternal thyroid peroxidase antibodies during pregnancy: A marker of impaired child development? *Journal of Clinical Endocrinology and Metabolism,* 3561–3566.

Potvin, L., Lasker, J., & Toedter, L. (1989). Measuring grief: A short version of the Perinatal Grief scale. *Journal of Psychosomatic Behavioural Assessment, 11,* 29–45.

Power, M. (1997). Fertility problems. In G. Andrews (Ed.), *Women's sexual health*. London: Balliere Tindall.

Prettyman, R. J., Cordle, C. J., & Cook, G. D. (1993). A three-month follow up of psychological morbidity after early miscarriage. *British Journal of Medical Psychology, 66,* 363–372.

Prochaska, J. O., & DiClemente, C. C. (1984). *The transtheoretical approach: Crossing traditional boundaries of change*. Homewood, IL: Dow Jones Irwin.

Raval, H., Slade, P., Buck, P., & Lieberman, B. E. (1987). The impact of infertility on emotions and the marital and sexual relationships. *Journal of Reproductive and Infant Psychology, 5,* 221–234.

Reamy, K. J., & White, S. E. (1987). Sexuality in the puerperium: A review. *Archives of Sexual Behavior, 16,* 165–186.

Riley, D. (1995). *Perinatal mental health: A source book for health professionals.* Oxford: Radcliffe Medical Press.

Rimmer, C., Horga, M., Cerar, V., Alder, E. M., Baird, D. T., & Glasier A. (1992). Do women want a once-a-month pill. *Human Reproduction, 7,* 608–611.

Robson, K. M., Brant, H. A., & Kumar, R. (1981).

Maternal sexuality during first pregnancy and after childbirth. *British Journal of Obstetrics and Gynaecology, 88,* 882–889.

Romito, P. (1988). Mothers' experience of breastfeeding. *Journal of Reproductive and Infant Psychology, 6,* 89–99.

Rosenman, R. H., Brand, R. J., Jenkins, C. D., Friedman, M., Straus, R., & Wurm, M. (1975). Coronary disease in the Western Collaborative Group Study. Final follow up experience of 8½ years. *Journal of the American Medical Association, 233,* 872–877.

Scambler, G. (1997). *Sociology as applied to medicine* (4th ed.). London: WB Saunders.

Scheeran, P., & Abraham, C. (1996). The health belief model. In M. Conner & P. Norman (Eds.), *Predicting health behaviour*. Buckingham, UK: Open University Press.

Scheeran, P., White, D., & Phillips, K. (1991). Premarital contraceptive use: a review of the psychological literature. *Journal of Reproductive and Infant Psychology, 9,* 253–269.

Schwartz, L. L. (1991). *Alternatives to infertility: is surrogacy the answer*. New York: Brunner/Mazel.

Shaw, P., Johnston, M., & Shaw, R. (1988). Counselling needs, emotional and relationship problems in couples awaiting IVF. *Journal of Psychosomatic Obstetrics and Gynaecology, 9,* 171–180.

Sheeran, P., White, D., & Philips, K. (1992). Premarital contraceptive use: a review of the psychological literature. *Journal of Reproductive and Infant Psychology, 10,* 253–269.

Sherwin, B. B. (1988). Estrogen and/or androgen replacement therapy and cognitive functioning in surgically menopausal women. *Psychoneuroendocrinology, 13,* 345–357.

Smith, B. H., Gurney, E., Aboulela, & Templeton, A. (1996). Emergency contraception: A survey of women's knowledge and attitudes. *British Journal of Obstetrics and Gynaecology, 103,* 1109–1116.

SNAP (1994). Teenage pregnancy in Scotland. Scottish needs assessment programme. Glasgow Scottish Forum for Public Health Medicine.

Snowden, R., & Snowden, E. (1984). *The gift of a child.* London: Allen & Unwin.

Social Trends (1995). London: HMSO.

Spitzer, R. I., Endicott J., & Robins, E. (1978). Research diagnostic criteria: Rationale and reliability. *Archives of General Psychiatry, 35,* 773–782.

Strang, V. R., & Sullivan, P. L. (1985). Body image attitudes during pregnancy and the postpartum period. *Journal of Obstetric, Gynecologic and Neonatal Nursing, 14,* 332–337.

Sutherland, C. (1997). Young people and sex. In G. Andrews (Ed.), *Women's sexual health*. London: Balliere Tindall.

Tamminen, T. (1989). The impact of mother's depression on her breastfeeding attitudes and experiences. *Journal of Psychosomatic Obstetrics and Gynaecology, 10*(Suppl.1), 69–70.

Tang, C. S.-K., & Chung, T. K. H. (1997). Psychosexual adjustment following sterilization: A prospective study on Chinese women. *Journal of Psychosomatic Research, 42,* 187–196.

Templeton, A., & Penney, G. C. (1982). The incidence, characteristics and prognosis of patients whose infertility is unexplained. *Fertility Sterility, 31,* 175–179.

Ussher, J. M. (1991). *Women's madness: Misogyny or mental illness*. London: Harvester Wheatsheaf.

Walker, B., & Erdman, A. (1984). Childbirth education programs: the relationship between confidence and knowledge. *Birth, 11,* 103–108.

Warner, P., & Bancroft, J. (1990). Factors relating to self-reporting of the pre-menstrual syndrome. *British Journal of Psychiatry, 157,* 249–260.

White, A., Freeth, S., & O'Brien, M. O. (1992). *Infant feeding 1990*. OPCS.

Whiteford, L., & Gonzalez, L. (1995). Stigma: The hidden burden of infertility. *Social Science and Medicine, 40,* 27–36.

Whitley, B. E., & Schofield, J. W. (1986). A meta-analysis of research on adolescent contraceptive use. *Population and Environment, 8,* 173–203.

Whitton, A., Warner, R., & Appleby, L. (1996). The pathway to care in post-natal depression. Women's attitudes to postnatal depression and its treatment. *British Journal of General Practice, 46,* 427–428.

Wing, J. K., Cooper, J. E., & Sartorius, N. (1974). *The measurement and classification of psychiatric symptoms*. Cambridge, UK: Cambridge University Press.

Woodcock, A., Stenner, K., & Ingham, R. (1992). "All these contraceptives, videos and that..." Young people talking about school sex education. *Health Education Research,* 517–531.

Woollett, A., Dosanjh-Matawala, N., & Hadlow, J. (1991). The attitudes to contraception of Asian women in East London. *The British Journal of Family Planning, 17,* 72–77.

Wright, P. (1988). Learning experiences in feeding behaviour during infancy. *Journal of Psychosomatic Research, 32,* 613–619.

Zabin, L. S., Hirsch, M. B., & Emerson, M. R. (1989). When urban adolescents choose abortion: Effects on education, psychological status and subsequent pregnancy. *Family Planning Perspective, 21,* 248.

Zuckerman, M., & Lubin, B. (1965). *Manual for the Multiple Affect Adjective checklist*. Educational and Industrial Testing Service, San Diego.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.17
Genetic Disorders

HEIDI A. HAMANN and ROBERT T. CROYLE
University of Utah, Salt Lake City, UT, USA

8.17.1 OVERVIEW OF GENETIC DISORDERS 409
    8.17.1.1 Biological View of Genetic Disorders 410
    8.17.1.2 Age of Onset and Course of Genetic Disease 410
    8.17.1.3 Symptoms, Prognosis, and Treatment 410
8.17.2 IMPACT OF THE CONDITION ON PERSON, FAMILY, AND SOCIETY 411
    8.17.2.1 Intrapersonal, Family, and Public Reactions to the Individual with a Genetic Disease 411
        8.17.2.1.1 Person 411
        8.17.2.1.2 Family 411
        8.17.2.1.3 Society 412
    8.17.2.2 Reproductive Issues 412
8.17.3 PSYCHOLOGICAL ASPECTS OF GENETIC SCREENING AND TESTING 413
    8.17.3.1 Informed Consent and Protocol Procedures 413
    8.17.3.2 Factors Influencing Participation in Testing 414
    8.17.3.3 Genetic Screening 414
        8.17.3.3.1 Newborn genetic screening 414
        8.17.3.3.2 Prenatal genetic screening and testing 415
        8.17.3.3.3 Carrier testing 416
    8.17.3.4 Predictive Testing 416
        8.17.3.4.1 Huntington's disease 416
        8.17.3.4.2 Cancer susceptibility 418
8.17.4 PSYCHOLOGICAL INTERVENTIONS 420
    8.17.4.1 Support Groups 421
    8.17.4.2 Individual Therapy 421
    8.17.4.3 Family Therapy 422
    8.17.4.4 Psychological Intervention Within the Medical Setting 422
8.17.5 PROFESSIONAL ISSUES 422
    8.17.5.1 Role of the Genetic Counselor 422
    8.17.5.2 Role of the Psychologist 423
8.17.6 CONCLUSIONS 423
8.17.7 REFERENCES 423

8.17.1 OVERVIEW OF GENETIC DISORDERS

This chapter focuses on the psychological aspects of a specific type of chronic illness: genetic disorders. The label "genetic disorder" describes a heterogeneous group of conditions that are at least partially caused by changes in a person's hereditary information (Thompson, McInnes, & Willard, 1991). Although the term genetic disorder can be defined in a variety of ways, the labels genetic disorder and genetic disease are used here to refer to hereditary conditions that usually (but not always) pose unusual physical or psychological challenges for an individual's quality or duration of life. This definition intends to capture some common psychological themes associated with these disorders while respecting the fact that each person's interpretation of his or her genetic status can be extremely variable.

Genes are the specific units of hereditary information found in almost every cell of the body, and are carried on structures known as chromosomes (Mange & Mange, 1990). In a genetic disease, either specific genes are changed (mutated) or there is extra or missing chromosomal material (Thompson et al., 1991). Genetic disorders pose unique demands for the medical practitioner, including the need to assess genetic risk and to inform patients about genetic testing. However, the issue of hereditary transmission in genetic disorders not only raises medical concerns, it also raises important psychological concerns. This chapter focuses on the predominant psychological issues in genetic disease and on the very timely topic of psychological issues in genetic testing.

8.17.1.1 Biological View of Genetic Disorders

Before psychologists can understand the psychological sequelae of genetic disease, it is important that they have basic knowledge about the mechanism of transmission. In reproduction, genes are passed from parents to offspring; therefore, children inherit half of their genetic material from their mothers and half from their fathers (Nora, 1994). Accordingly, a person with a disease-causing gene mutation may pass that mutation to their children. Some gene mutations, however, occur in an individual without being found in either of the individual's parents; in this case, the disease-causing mutation usually occurred in either the sperm or egg of a parent. For example, in Down's syndrome, a mistake in chromosome replication in the sperm or egg cell causes an individual to have extra chromosomal material, even if both parents had normal numbers of chromosomes (Korf, 1996). Therefore, a disease does not necessarily need to be inherited from a parent in order to be classified as genetic. This means that there are two ways to contract a genetic disease: disease-contributing mutations or chromosome aberrations may either be passed from a parent (inherited) or formed newly in that individual (*de novo*). However, once a gene is mutated or chromosome material is added or subtracted, and this genetic change exists in the reproductive cells of an individual, there is the possibility the genetic change may be passed on to future offspring (Harper, 1993).

It is a popular, but misguided, perception that every genetic disorder is caused by a single gene acting in isolation. Although it is true that some genetic disorders are influenced by mutations in one gene (single-gene disorders), many others are influenced by mutations in many genes (polygenic disorders). Furthermore, the expression and severity of all genetic disorders are mediated by the environment in some manner, be it nutrition, toxins, the prenatal environment, or the gene's own cellular environment. Diseases that have a combination of genetic and environmental components are often termed multifactorial conditions (Harper, 1993).

8.17.1.2 Age of Onset and Course of Genetic Disease

Although many genetic conditions have similar biological etiologies and mechanisms of inheritance, the age of onset and course of the conditions are very diverse. In Down's syndrome, the physical manifestations of down-slanting eyes, epicanthal eye folds, and low-set ears are usually apparent at birth (Thompson et al., 1991). However, in neurofibromatosis type 1, the characteristic *cafe au lait* spots on the skin may not be present until early childhood (Viskochil, 1995). Adult onset is a feature of Huntington's disease, a disease in which many affected individuals first exhibit jerky, uncoordinated movements between the ages of 30 and 50 years (Huntington Society of Canada, 1993).

Considerable variability in age of onset and disease course characterizes many of the heritable adult cancer syndromes. For example, some women with a mutation in the *BRCA1* gene contract breast or ovarian cancer in their early 30s, whereas others are not affected until their 70s. Some *BRCA1* mutation carriers will never contract cancer. Even among those who carry the same cancer susceptibility mutation and contract cancer, the course of the disease can vary substantially (Gould, 1997).

8.17.1.3 Symptoms, Prognosis, and Treatment

As with the age of onset and course of disease, information about symptoms, prognosis, and treatments also differ depending upon the specific condition. Many genetic disorders are associated with varying degrees of mental retardation, although some are associated with normal levels of intelligence. Some genetic disorders carry physically distinguishing features, but others may not be apparent to the casual observer. Almost all genetic disorders shorten a person's life span, but they do so in varying degrees. Whereas individuals with achondroplasia (a form of dwarfism) may have long, relatively healthy lives into late adulthood, infants with a chromosomal disorder called trisomy 13 have such severe malformations that most die within the first year of life (Thompson et al., 1991). New treatments, however, promise to improve the prognosis and quality of life for many individuals with genetic disorders. For example, new respiratory drugs for cystic fibrosis have lengthened the average life span from less than 20 years to over 30 years (ASHG, 1992). Unfortunately, medical treatments can be costly, time-consuming, and sometimes painful. Because there are no cures for most genetic conditions, once a treatment or surveillance program is initiated it is usually continued for life. Therefore, compliance and adherence to medical follow-up and treatment are significant issues for an individual with a genetic disorder.

8.17.2 IMPACT OF THE CONDITION ON PERSON, FAMILY, AND SOCIETY

There is a wide range of psychological issues that may be associated with having a genetic predisposition or condition. Many of the issues are similar to those found in people with other chronic illnesses. However, the hereditary etiology of most genetic disorders creates unique psychological issues rarely encountered in persons with nongenetic chronic illnesses. This section focuses on the psychological issues that may be especially salient, or completely unique, among people affected by genetic conditions.

8.17.2.1 Intrapersonal, Family, and Public Reactions to the Individual with a Genetic Disease

8.17.2.1.1 Person

A genetic disorder can have critical effects on an individual's self-perception and self-esteem (Kessler, 1979). For most people, health is considered an integral part of the self. Therefore, when one has a chronic illness, it can severely threaten the integrity of the self. One way that individuals attempt to maintain the self concept in the face of disease is by projecting the illness externally (Shapiro, 1983). Distancing the self from the illness facilitates the maintenance of the self concept, because the disease is viewed as ego-alien, an enemy to be conquered (Kessler, 1979).

For individuals with genetic disorders, however, it may prove extremely difficult to project the genetic disorder outward and distance the self from it. Genetic disease is at least partially caused internally, and the disease-causing enemy is literally one's own genetic material in every cell of the body. As a result, the self concept of a person with a genetic disorder may be intrinsically bound with the view of the disorder (Kessler, 1979). This inability to distance the self from a negatively viewed disorder may affect self-esteem, leading to feelings of unworthiness. Individuals may feel that since their genes are defective, they are defective as well (Weiss, 1981). Furthermore, since most disease-causing genetic changes are present from birth and are permanent, this feeling of defectiveness may continue throughout a person's lifetime (Schild & Black, 1984).

8.17.2.1.2 Family

Genetic disorders are unique in the fact that every individual diagnosis becomes a family diagnosis. After the diagnosis of a genetic disorder in one individual, family members may discover that they themselves are mildly affected by the disorder, are at risk of developing it, and/or may potentially pass it to their offspring (Schild & Black, 1984). The family diagnosis can have a significant impact on family relationships and communication. One parent may blame the other for carrying a gene that has adversely affected their children. Siblings may compare their genetic status with each other and attach values of worth to each other as a result of their status. Traditional family roles may be radically altered as one or more individuals give up or take on new responsibilities to compensate for the presence of the genetic disorder in other family members (Weiss, 1981).

Another familial issue unique to genetic disease involves the transmission of information about the condition between family members. Not only is there biological inheritance of a genetic condition, there is also a kind of psychological inheritance attached to it. Just as a disease-related gene may be passed from parent to child, representations of the disease and its consequences may be inherited too. This type of psychological inheritance uniquely colors and shapes a person's experience with the condition (Cullinan, 1989). In one family, a member with achondroplasia may be shunned and considered someone of whom to be ashamed. Family members with the condition may be encouraged to integrate into the taller culture as much as possible and not share details of their condition with others. They are stigmatized and marginalized within their own families, and the condition, in effect, becomes a sort of family secret. In another family the condition may be seen as completely normal and worthy of frank discussion. Because family attitudes toward the genetic condition play a critical role in psychological adjustment, clinicians should always explore and discuss these beliefs with their clients.

Since more than one individual in a family is often affected with a given condition, the burden of care for the rest of the family becomes greater than usual. Some parents may find themselves caring for two sons with Duchenne muscular dystrophy, a progressive muscle weakening disease, at one time. A woman at risk of a predisposing breast cancer gene mutation may be the primary caretaker for her mother and sister, both stricken with breast cancer. As if being a caregiver to multiple persons within the family were not difficult enough, individuals who are not affected can experience survivor guilt. The sense of relief over being spared of the disease can be experienced as a secret not to be shared with affected relatives. This ambivalence has been reported, for example, by individuals who test negative for Huntington's disease when a sibling tests positive (Tibben et al., 1994), and a similar dilemma may occur among unaffected members of families with inherited cancer syndromes (Baty et al., 1997).

8.17.2.1.3 Society

Societal reactions to the person with a genetic disorder may be stigmatizing. Because many genetic conditions have physical or behavioral manifestations, the affected individual is often quite noticeable to others. Furthermore, since most genetic conditions are relatively rare, the lay public is often unfamiliar with and unaccustomed to seeing affected individuals. The combination of these two factors can lead people to stare at, laugh at, marginalize, and shun an individual with a genetic disorder. For individuals with neurofibromatosis type 1, the physical features of pigmented *cafe au lait* spots and raised neurofibromas seem unusual and unsightly to others when seen for the first time. Individuals with physically stigmatizing genetic disorders may already have issues of insecurity and feelings of unattractiveness with which to deal, so the public's tendency to turn away, laugh, or make cruel and hurtful remarks can exacerbate these issues substantially (Lynch, Lynch, & Lynch, 1979). When affected individuals enter the initial stages of their disease, public recognition of the changes can be especially hurtful. For example, a middle-aged person who is entering the initial stages of Huntington's disease often shows uncontrollable jerky movements. These symptoms can be troubling and embarrassing for the individual, not only because of their public impact, but also because each movement symbolizes to the individual that he or she has an incurable degenerative disease (Wexler, 1979). Unaffected family members also experience shame from public stigmatization. Some parents of newborns with Down's syndrome report that not only do they have to deal with conflicting emotions about their newborn baby, they must continuously deal with others' insensitive reactions (Kelly, 1977). Some people simply avoid public situations in order to avoid scrutiny.

8.17.2.2 Reproductive Issues

For many parents of a chronically ill child, guilt is a common emotion. In the face of a seemingly unexplainable event—an innocent child struck with a serious illness—parents often stop to wonder if their actions somehow caused the disease. In the context of genetic disorders, guilt and shame regarding inheritance are extremely common and sometimes debilitating emotions (Kessler, 1979). Some parents feel guilty for passing defective genes to their children, even after it is explained to them that the process was determined by chance. Others see the condition as a challenge from God to test their faith and perseverance (Kushner, 1981). An individual with a genetic condition may experience mixed emotions about the chance that his or her child will be affected. On the one hand, a person with a genetic condition does not want their children to experience a disease that will disadvantage them; on the other hand, the condition may be deeply integrated into the parent's concept of self and family. As a result, it is not uncommon for deaf parents to express some relief when they learn that their newborn is deaf as well (Harper, 1993).

Even when neither parent is affected and the disease is unexpected, a genetic disease can stimulate guilt and shame in parents. The tendency for ill individuals to search for the causes of their disease is magnified in genetic disease, parents will often blame themselves as the cause of the disorder (Kelly, 1977). Some women analyze their entire pregnancies, chastising themselves for a single glass of wine or a few cigarettes. Despite a clinician's careful explanation that *de novo* mutations and chromosome aberrations are almost always caused by chance, parents may continue to believe in self-causation.

Among expectant parents, some degree of fear concerning the possibility of health problems is common. For people who know that a child is at increased risk of a genetic disorder, this fear may be overwhelming (Kessler, 1979). Normally enjoyable events such as ultrasounds and fetal heartbeat checks become agonizing indicators of the baby's status. For some diseases it is possible to screen or test prenatally, and as explained further in Section 8.17.3 the decision-making process and the waiting periods can cause significant distress for some people (Marteau, 1992).

Many genetic conditions carry with them an increased risk of infertility (Harper, 1993), and this fact may also complicate attitudes about childbearing. When a person is infertile it can significantly affect that person's self-concept and self-esteem. For persons affected with a genetic condition who already have conflicting attitudes toward child bearing, they may see it as a message that they should not have children and think more negatively about themselves. Some individuals may never have been told by their health care professional about the possibility of infertility, because the health care professional made the erroneous assumption that childbearing would never be an issue. In and of itself, infertility can be a significant stressor to an individual. However, when the phenomenon of genetic disease is added, these effects can be exacerbated.

Even for parents who made a well-informed decision to proceed with a high-risk pregnancy, and are able to reconcile their own feelings of guilt, the responses of others may create shameful feelings (Weiss & Mackta, 1996). Family members, friends, or even complete strangers may make comments such as "How can you bring children into this world knowing they will be at such a disadvantage?" For persons themselves affected with a genetic disease, such public pressure to forego childbearing may also deeply undermine their own feelings of self-worth. To the person with an integrated sense of disease and person, a public denouncement of the condition is also perceived as a denouncement of the person, a very personal form of eugenics. In contrast, within some communities of people with a high prevalence of genetic conditions, there may be pressure to *have* children in order to keep the ranks strong. The religious and cultural context of reproductive decisions can further complicate the ability of parents to make decisions that balance needs for autonomy and social acceptance (Punales-Morejon & Penchaszadeh, 1992).

8.17.3 PSYCHOLOGICAL ASPECTS OF GENETIC SCREENING AND TESTING

Recent advances in human genetics and molecular biology have allowed the identification of a rapidly increasing number of disease-susceptibility genes and the development of tests to detect the presence of mutations in these genes. These developments allow genetic counselors and medical geneticists to provide genetic information to allow informed decision-making about reproductive issues, risk modification, and disease management. Different screening programs and testing protocols have been developed in an attempt to systematize the dissemination of genetic information. The emerging abundance of genetic screening and testing programs has prompted studies on the psychological impact of genetic screening and testing. Although this body of literature is growing, much of this research is in its infancy. This section focuses on research that explores the psychological impact of testing and the psychological sequelae associated with learning one's genetic risk of a given disease. In addition, psychological issues important for further research and clinical intervention are outlined.

8.17.3.1 Informed Consent and Protocol Procedures

Screening for some genetic conditions, such as phenylketonuria (PKU), is performed through a blood test in many newborns. Although sometimes refused, it is considered fairly routine, and often involves a short informed consent procedure (Natowicz & Alper, 1991). However, genetic testing for other genetic conditions often involves a lengthy informed consent procedure. A person following the accepted protocol for Huntington's disease testing, for example, will have an initial information appointment with a genetic counselor and/or medical geneticist, an appointment with a neurologist to rule out any signs of the disorder, an appointment with a mental health professional to discuss any psychological issues that may effect testing, an appointment for the blood draw, a follow-up appointment to learn the results, and periodic follow-ups with various physicians depending upon the test results (University of Michigan Medical Center, 1994). The informed consent standards for most genetic conditions are rigorous because there is a great deal for patients to consider before getting tested, including possible psychological effects and insurance risks (Geller et al., 1997). Newly developed tests are often performed under the umbrella of a research protocol. Even when tests are offered clinically, there should be a well-defined protocol and a thorough informed consent and patient education process. In both research and clinical situations, the informed consent and patient education protocol may include multiple visits before testing is actually performed. Despite the endorsement of these standards by many relevant professional organizations, recent evaluations of clinical genetic testing suggest that many physicians do not comply with them (Giardiello et al., 1997).

8.17.3.2 Factors Influencing Participation in Testing

How does an individual decide whether or not to get tested for a disease-susceptibility mutation? This question is one that psychologists have been interested in since genetic testing first became available. Initial surveys of people at risk of various genetic diseases and of the population at large showed that the majority of individuals in both groups indicated interest in genetic testing. However, as more genetic tests have become available, studies have shown that a much lower percentage actually utilize testing. One example of this phenomenon is illustrated by experiences with Huntington's disease (HD) testing. As predictive tests for HD became available in the late 1980s, researchers surveyed at-risk individuals to see whether they may be interested in testing. Various studies showed that on average, approximately 70% of respondents indicated that they would take a predictive HD test if it were available (Kessler, Field, Worth, & Mosbarger, 1987; Markel, Young, & Penney, 1987; Mastromauro, Myers, & Berkman, 1987). However, in the 5–7 years following the initial availability of predictive HD testing, only about 25% of the at-risk population had been tested (Quaid, 1997; Tyler, Ball, & Crawford, 1992). When asked why they did not seek testing, at-risk individuals indicated that the desire to avoid threatening information was the most important reason for not seeking testing (Quaid & Morris, 1993). Interestingly enough, in this study the burdensome nature of the test was listed as the least important reason not to be tested. Therefore, it is not test complexity barriers that prevents people from being tested; instead, the motivations to forego testing often involve the desire to avoid potentially negative news. It seems that once people are faced with the reality of actual testing, they may reevaluate their initial interest in the test.

Research on predicted vs. actual uptake for other genetic conditions have not shown such dramatic results as that for HD. For prenatal genetic testing, there appears to be a stronger link between attitudes and behaviors (Croyle & Lerman, 1995). This phenomenon can be explained in that pregnancy has a strict time course so the option of postponing prenatal testing is not available. In prenatal situations people do not have time to ruminate about the possibility of negative news. In addition, although HD has no effective treatment, for some people pregnancy termination is a viable option for dealing with a positive test result. These individuals may therefore be more motivated to seek testing, because they foresee a possible alternative should a test result indicate that a fetus has a known genetic mutation. In other arenas, such as breast and ovarian cancer susceptibility, tests have been in existence for a relatively short amount of time, so it is not possible to determine the rate of uptake in the general population. However, initial reports regarding known at-risk individuals have shown greater uptake of testing than in HD (Croyle, Smith, Botkin, Baty, & Nash, 1997; Lerman et al., 1996; Patenaude et al., 1996). Croyle et al. (1997) reported that of the first 213 study respondents, 84% expressed interest in participating in BRCA1 testing. Lerman et al. (1996) found that 60% of participants who completed the baseline interview were tested for BRCA1 and requested their test results. They also reported that the number of first-degree relatives with breast cancer, baseline knowledge, and perceived importance of the test were all positively related to utilization of testing. Qualitative data reported by many of the same investigators indicated that concerns about risks to children and concern about surveillance and prevention also were important reasons for being tested (Lynch et al., 1997). This suggests that the more people perceive that prevention and treatment for the disease is possible, the more motivated they will be to seek testing.

8.17.3.3 Genetic Screening

8.17.3.3.1 Newborn genetic screening

Newborn screening for such conditions as PKU is the most common and most established type of genetic screening (Thompson et al., 1991). In many states of the USA these screens are performed on every newborn. Very little has been written about the psychological effects of routine newborn screening, perhaps because it is perceived to cause few stress-provoking responses for parents. Many parents are not even aware that these screening procedures are performed and are relieved to know that there are no detected problems. Obviously the news can be traumatic for those who discover through routine newborn screening that their child has a rare genetic disorder. They may feel shocked, frightened, and surprised. However, it is thought that the test procedure itself is usually not associated with significant distress, and since most parents do not even realize what is being tested for, pretest anxiety is seldom an issue.

8.17.3.3.2 Prenatal genetic screening and testing

Other types of screening are performed by taking a maternal blood sample in pregnancy and checking the levels of maternal serum α-fetoprotein (MSAFP). Fetuses of women with high or low MSAFP results are considered at increased risk of chromosome abnormalities or neural tube defects, depending upon the MSAFP level (Harper, 1993; Merkatz, Nitowski, Macri, & Johnson, 1984). Women with abnormal MSAFP results are often offered ultrasound and/or amniocentesis (Green & Statham, 1996). Ultrasound is a noninvasive technique used to visualize fetal structures, whereas amniocentesis is an invasive technique used to directly test the fetal chromosomes and fetal AFP level (Harper, 1993). In amniocentesis, a needle is inserted into the abdomen and a small amount of the amniotic fluid is withdrawn. This procedure is usually done after 14 weeks gestation. An amniocentesis has a small risk of miscarriage, usually from 0.25% to 1%, with rates differing from center to center (Medical Research Council Working Party on Amniocentesis, 1976; Tabor, Madsen, Obel, Philip, Bang, & Norgard-Peterson, 1986; United States Institute of Child Health and Human Development Study Group, 1976). As women get older they have increased risks for having fetuses with chromosome abnormalities. Thus, many women over a certain age (usually 35) are offered amniocentesis or another type of prenatal testing called chorionic villus sampling (CVS). CVS is a procedure usually performed at 10–13 weeks gestation, in which, depending on the chorionic position, a catheter is inserted into the vagina or a needle is inserted into the abdomen and a small sample of the chorion is withdrawn. Unlike amniocentesis which detects both chromosome abnormalities and fetal AFP levels, CVS detects only chromosome abnormalities (Thompson et al., 1991). At 0.5–1%, the risk of miscarriage with CVS is also slightly higher than the risk of amniocentesis (Brambati, Lanzani, & Tului, 1990; Canadian Collaborative CVS–Amniocentesis Clinical Trial Group, 1989).

Studies have shown that women with abnormal MSAFP results tend to show increased anxiety after learning their results (Robinson, Hibbard, & Lawrence, 1984). Although some studies have shown that this anxiety was maintained even after amniocentesis showed a definitive normal result (Marteau et al., 1988), others have found that this level of anxiety tended to decrease when amniocentesis showed a definitive normal result (Tabor & Jonsson, 1987). Most studies following women over 35 who have amniocenteses show that, although anxiety was high before and immediately following the amniocentesis, it dropped significantly once normal test results were learned (Croyle & Lerman, 1995). Coping styles can also influence a woman's reaction to prenatal testing. Women who tend to cope by monitoring or seeking information have been shown to have higher levels of distress prior to screening, immediately after screening, and after receipt of results, than those who tend to avoid information (Phipps & Zinn, 1986). Thus, although the thoughts of having an amniocentesis can be troubling, for most women good news from it will significantly reduce their anxiety. Most studies have not shown evidence to support the notion that prenatal screening causes significant distress.

Many studies that explored the psychological effects of prenatal screening involved cases in which the fetus was at low risk and the majority of results were normal. For individuals who discover through prenatal testing that a fetus is affected with a genetic disease, the distress is often overwhelming. Once faced with this dilemma, they must decide whether to continue with the pregnancy or terminate it (Green & Statham, 1996). For women who decide to continue it, they go through the pregnancy knowing that the fetus is affected, and often view their pregnancy as different than most, because they must plan for a life with a disabled child. The stress, immediately acute, can turn into an underlying chronic strain of anticipation for this known, but unchangeable event. However, when the child is born these parents are more prepared for the event and may be able to cope with it more effectively. Those individuals who plan to terminate the pregnancy are also faced with stressors. Termination of a fetus because of a genetic disease is situationally and psychologically unique from other types of fetal losses. It is unlike an elective termination in the fact that the pregnancy was very much wanted. However, it is also distinct from a miscarriage or stillbirth because there was some element of choice involved in the process (Jorgensen, Uddenberg, & Ursing; 1985). A woman who discovers through amniocentesis that her fetus has a genetic disorder may be over 20 weeks gestation when she receives the news. If she decides to terminate the pregnancy she will most likely have to go through a lengthy, and possibly distressing, induced delivery. At this time she will also be showing her pregnancy and will have to explain her situation to others. Couples who terminate a pregnancy after finding out that it has a genetic condition may tell others it was a miscarriage rather than explain the true situation. This secrecy about the situation can cause or exacerbate intense feelings of shame and guilt. These couples also face mourning and grief over the loss of a wanted pregnancy (Blumberg, Golbus, & Hanson, 1975; Green & Statham, 1996). Furthermore, they may have fears about the possible recurrence of the event and face future childbearing with fear and indecision. In fact, studies have shown that of those women contemplating amniocenteses, women with previous chromosome abnormalities had significantly higher levels of distress than others (Beeson & Golbus, 1979).

8.17.3.3.3 Carrier testing

For some recessive conditions it is possible that an individual may not be affected with a condition but could be a carrier of it. In these cases, the child of two carrier parents is at risk of a given genetic disorder even though neither of the individuals has the disorder. Individuals from ethnic groups known to have higher susceptibilities to certain recessive conditions, such as sickle cell anemia, Tay–Sachs or cystic fibrosis (CF), or who have an affected relative, can have carrier testing (Thompson et al., 1991). Some carriers may deal with increased concern about childbearing, along with concerns about self-esteem and self-worth after they discover their genetic status (Massarick & Koback, 1981). Research focusing on carrier screening programs for CF in the UK has shown that carriers show small increases in anxiety after testing but do not appear to have concerns after 6 months (Watson, Mayall, Lamb, Chapple, & Williamson, 1992). Other studies have shown that most individuals appreciate finding out their carrier status and tend not to stigmatize themselves (Childs, Gordis, Kaback, & Kazazian, 1976). Therefore, although carrier testing may produce negative effects for some, for other individuals it seems to be a relief-provoking experience that creates the perception of increased control over reproductive decision-making.

8.17.3.4 Predictive Testing

Predictive testing is the newest addition to the realm of genetic testing and screening, and has many potential psychological issues associated with it. Predictive testing differs from other types of genetic testing in that the purpose is to find one's risk of developing a genetic disease later in life as opposed to discovering the status of one's progeny. Predictive testing can also indirectly provide information about one's children and one's parents. Thus, for some, conflict arises if some family members wish to be tested while others do not.

8.17.3.4.1 Huntington's disease

The first research that explored potential psychological effects of predictive testing came in the realm of HD testing. HD testing is discussed in some depth because it is one of the most thoroughly studied contexts for genetic testing. HD is a late-onset, progressive neurological disease with no effective treatment or cure. It is a rare condition, but those individuals with an affected parent are at 50% risk of having the gene mutation and developing the disease. Virtually everyone with the gene mutation will develop the disease at some point in their lives, although the age of onset is variable (Huntington Society of Canada, 1993). When testing for HD was first initiated, the HD gene had not been isolated, but the area around the gene on the chromosome had been delineated enough for the test to be done through a technique called linkage (Gusella et al., 1983). The important aspect of this type of test is that several family members are required to give blood samples in order to compare chromosome regions between affected and unaffected individuals, and these linkage methods are not 100% accurate. In the early 1990s the gene was isolated. (Huntington's Disease Collaborative Research Group, 1993). and a predictive test was developed that was able to directly determine whether an individual had a mutation without requiring the blood of the individual's family members. In addition, this test increased the accuracy of prediction to over 99% (Huntington Society of Canada, 1993).

When a linkage test for HD first became available, there was fear about the potentially devastating effects of discovering that one was a gene carrier. Many were concerned that when people found themselves to be gene carriers the risk of severe depression and suicide would be overwhelming. The concern was justified, as suicide rates in people diagnosed with HD are four times that of the general population (Farrer, 1986) and in one study 29% of the respondents, all of whom were at 50% risk of HD, indicated that they would have suicidal ideations should they test positive for the condition (Mastromauro et al., 1987).

In 1986 researchers in British Columbia initiated a pilot study designed to track those who decided to be tested for the HD gene mutation. In this protocol, and subsequent others, individuals who desired testing went through intensive information sessions about HD. In addition, these participants were given a battery of psychological measures before genetic testing, immediately following testing, and during subsequent follow-ups. The pretesting found that those who initially registered for testing were not psychologically different from the population at large (Bloch, Fahy, Fox, & Hayden, 1989). However, it was thought that these early self-selected participants may have better overall adjustment than others at risk of the disease. In a different study, Codori, Hanson, and Brandt (1994) found that a greater number of test-decliners than test-acceptors list the fear of emotional reactions as a significant factor in the decision-making process. They concluded that people who choose testing are psychologically self-selected for better adjustment. British Columbia researchers also found that 33% of their subjects indicated that they would consider suicide had they a positive result (Bloch et al., 1989). One of the 51 participants was found to have early signs of HD and had a severe suicidal reaction (Lam et al., 1988).

A more comprehensive report of the British Columbia data reported that in the 7–10 day follow-up, individuals who tested at increased risk for HD showed increased levels of distress, whereas individuals who tested at decreased risk had decreased levels of distress. However, these distress differences decreased after 6 months, and after 12 months the two groups did not differ significantly in their distress levels. Furthermore, after 12 months both groups showed significant drops in their distress levels (Wiggins et al., 1992). Therefore, it appears that in this pilot group, individuals who tested at increased risk for the gene mutation experienced initial distress immediately after finding out their test results, but within a year were able to deal with the news and integrate it into their lives. This finding has held in subsequent studies of individuals with an identified HD gene mutation, indicating that in general, the news of a positive carrier status, although upsetting in the short term, does not appear to have long-term distressing effects for most participants.

One important point is that in all of these studies, individuals who decide to be tested undergo intensive genetic information sessions, are given strategies to deal with potential negative emotional reactions, undergo initial psychological evaluation, and receive follow-up medical and psychological care. These services have been viewed as beneficial not only by the health care professionals, but also by the HD community in general. In fact, when HD testing became available on a wider basis, it was the Huntington's Disease Society of America, an organization made up of primarily HD carriers and family members, that recommended a structured protocol for those who desired to be tested (Huntington's Disease Society of America, 1994). There are certain sites in the USA that offer HD testing through this protocol, which includes initial genetic counseling, a neurological evaluation, and a psychological evaluation, before anyone can proceed with testing (University of Michigan Medical Center, 1994). Studies of people after testing indicate that most are glad they had the opportunity to learn what they did through the protocol, and studies of individuals who decline testing indicate that the burden of the procedure is an infrequently endorsed reason for not getting tested (Quaid & Morris, 1993).

In the British Columbia study, plus other subsequent studies, one surprising result was the long-term reaction of individuals who tested negative for a mutation in the HD gene. On average, these individuals experienced short-term relief, demonstrated by decreases in measures of distress and depression. However, after the initial euphoria, most of these individuals returned to normal levels of distress and some even experienced increased levels of distress in the long term. These individuals had to readjust their life plans and views to include life without the disease. Many had lived their lives under the assumption that they would develop HD, and now had to redefine themselves to include longer, healthier lives. For some of these individuals there was also the phenomenon of survivor guilt with which to deal. Some had seen their siblings test positive while they tested negative, and had to face the guilt of the "Why not me?" question. For some, too, they felt unable to express the joy that they felt to their families because other members had tested positive. In a study by Huggins et al. (1992), it was reported that approximately 10% of those individuals who received test results indicating a decreased HD risk needed additional counseling to help them cope with their status.

Another interesting feature of the British Columbia study was that individuals who did not find out their HD status, either because they chose not to take the test or had uninformative tests through linkage analysis, were the people who fared the worst psychologically (Wiggins et al., 1992). This result suggests that the uncertainty reduction provided by the test, regardless of the results themselves, lends some psychological benefit. People who choose not to get tested may have problems reconciling their decision with the life of uncertainty that they face, whereas people who follow through with testing and who are then unable to get results feel frustrated. These people committed to a hard decision, but were unfulfilled by any indication of certainty.

Since the initial studies, others have filled in some of the gaps about who does well, who does poorly, and why they do so after testing. Tibben et al. (1993) found that carriers showed more intrusive thoughts than noncarriers without showing increased levels of overt hopelessness. The researchers attributed this finding to possible minimization or denial of problems. This finding shows that on the surface, carriers may appear to be adjusting well but there is a possibility of underlying distress. They also found that certain individual differences were important in reactions to test results. The more intrusive HD-related feelings and avoidance of HD-related situations before testing, the greater the chance of these feelings being there 6 months after testing for individuals who tested positive.

They also found that individuals who avoided HD-related situations at pretesting were more likely to do so after testing, and that pessimistic individuals were more likely to avoid HD-related situations, be dissatisfied with their social support, and become depressive or suicidal after testing (Tibben et al., 1993). Bloch, Adam, Wiggins, Huggins, and Hayden (1992) found that psychological preparedness was an important component of psychological reactions. Those who had prepared themselves psychologically for potential bad news seemed more ready to cope with the information. In addition, those who sought out social support did better than those who isolated themselves. Other studies have explored the reactions of family members, including spouses, to the results. Quaid and Wesson (1995) found that prior to testing, spouses of at-risk individuals were significantly more depressed than their partners. Furthermore, couples in which one individual was found to be at increased risk of HD had higher measures of distress at 3 and 6 month follow-ups than their lower-risk counterparts (Quaid and Wesson, 1995).

The HD findings support the need for psychosocial resources for those who are contemplating HD testing. Psychological baseline measures and initial psychological interviews help to identify people who may not have the psychological resources at that time to deal with discovering their status. Some have argued that the use of these baseline measures and interviews are paternalistic and should not be used to screen out people (Wilke, 1995). However, as described previously, research has shown that they can be helpful indicators (Bloch et al., 1989; Huggins et al., 1992; Lam et al., 1988). It is our opinion that psychological measures and interviews should be used responsibly and not function to abandon unsuitable research subjects. If questionnaires or interviews indicate that an individual has high levels of distress or insufficient coping resources, then counseling should be offered to help the individual deal with the distress and possibly obtain the coping skills needed to deal with a test result. Furthermore, the presence of mental illness at a previous point in one's life should not necessarily preclude one from getting testing. In the past, individuals who have been previously diagnosed with psychiatric disorders, but have since been treated, have been allowed to proceed with testing (Quaid, 1992). Psychological resources are also important for the individuals after testing in order to help them cope with unexpected results, circumstances, and emotions. The need for this support is often underestimated by physicians. As indicated by research, even individuals with a negative result may require additional counseling (Huggins et al., 1992). Furthermore, psychological intervention is important for spouses, children, and other relatives of those tested. They may be the people most overlooked in these situations, but are the people left as caretakers and left to mourn after the affected person dies (Kessler, 1993).

8.17.3.4.2 Cancer susceptibility

One of the largest efforts in genetic research since the mid-1980s has focused on identifying gene mutations that predispose individuals to certain types of cancer. Many of these cancers, such as Li–Fraumeni syndrome, retinoblastomatous, and Wilms' tumor, are quite rare. However, in recent years, predisposing genes to more common types of cancer have been identified. Two broad categories of cancers that have drawn attention are colon cancer and breast cancer.

Hereditary colon cancer comes in two main forms: an early childhood-onset (familial adenomatous polyposis), and a more common teenage-to adult-onset hereditary nonpolyposis colon cancer (HNPCC). There are multiple genes and multiple mutations known to contribute to HNPCC, but there are also families with clearly hereditary colon cancer in which no known mutations have been found, so it is thought that other unknown genes and/or mutations play a role in colon cancer (Croyle & Lerman, 1993; Vernon et al., 1997).

For years, clinicians had noted that breast and/or ovarian cancers appeared to run in families. It has been estimated that 6–19% of all cases of breast cancer can be attributed to a family history of the disease (Colditz et al., 1993; Slatery & Kerber, 1993). In 1994, the *BRCA1* gene was identified; mutations in this gene predispose women to breast and ovarian cancer (Miki et al., 1994). Shortly after that, another predisposing gene, *BRCA2*, was identified (Wooster et al., 1995). It is estimated that females with an identified mutation in *BRCA1* have an 87% lifetime risk of developing breast cancer and a 44% lifetime risk of developing ovarian cancer (Ford, Easton, Bishop, Narod, Goldgar, & Breast Cancer Linkage Consortium, 1994). Mutations in the *BRCA2* gene give males a 6% risk of developing breast cancer by age 70 and females an 87% risk of breast cancer plus a moderately increased risk of ovarian cancer (Hoskins et al., 1995).

Recent research has found that over 2% of Ashkenazi Jewish individuals carry mutations in *BRCA1* or *BRCA2* genes. However, interestingly, the risks of developing cancer for Ashkenazi Jewish individuals who have certain *BRCA1* or *BRCA2* mutations is less than previous estimates of high-risk individuals (Struewing et al., 1997). Findings of this nature underscore the role of modifying factors in determining whether a *BRCA* mutation will cause cancer in a given individual (Healy, 1997).

The identification of these genes has made it possible to identify some individuals who will be at increased risk of developing cancer. However, these genes are not completely predictive of cancer, they are merely predisposing factors. Unlike HD in which a person with an identified mutation has over 99% chance of getting the disease sometime in the person's lifetime, a person with an identified cancer predisposing mutation has anywhere from 50% to 90% chance of developing cancer by age 70 (Croyle & Lerman, 1995). Therefore, unlike genetic conditions that have fairly straightforward inheritance patterns, most common cancers are true multifactorial diseases. Thus, from the procedural end, for the genetic counselor or other health professional explaining a person's genetic risk, cancer counseling can be quite complex. Most risk estimates are based on empirical data, so when a cancer predisposing mutation is just recently discovered, there is little data upon which to base a risk estimate. Therefore, cancer risk counseling is a dynamic process; as more research is done, risk estimates continue to change.

As detailed previously, studies have shown that the desire to reduce uncertainty is one of the main reasons individuals give for wanting to go through with genetic testing. A great deal of uncertainty remains, however, for individuals found not to have an identifiable cancer-related mutation. A negative test result may mean that the individual has the population risk for the disease; thus the chance of disease cannot be completely eliminated by the negative result. Furthermore, especially for newly identified genes, not every disease-susceptibility mutation has been identified. Therefore, a person without a known mutation in the family, who is told that an identifiable mutation was not found, may be tested at a later date and be found to have a newly identified mutation. Thus, test participants who do not experience uncertainty reduction may be disappointed. In fact, data from the HD model have shown that it was those individuals whose uncertainty was not reduced who fared the poorest psychologically.

Several large prospective studies of the psychological impact of cancer genetic testing were initiated in the 1990s. One such study showed that although *BRCA1* noncarriers showed significant reductions in depressive symptoms and functional impairment as opposed to carriers and decliners of testing at a 1 month follow-up, mutation carriers did not show increases in depression and functional impairment from baseline to follow-up (Lerman et al., 1996). In another study, mutation carriers were found at a 1–2 week follow-up to manifest more symptoms of test-related distress and general distress than noncarriers (Croyle et al., 1997). All participants, regardless of their carrier status, had significant declines in anxiety between the pretest and 1–2 week follow-up.

Preliminary studies show that, similar to HD, in the short term after testing noncarriers are relieved at their status, whereas carriers are somewhat distressed. However, the overall declines in distress levels pre- and post-testing show that even among carriers, the mere fact of testing decreases anxiety. One way to interpret this finding is by theorizing that the act of testing reduces perceived uncertainty, thus reducing anxiety, even if the results are less than certain. It is important to remember, however, that both of these studies took place under research protocols including extensive genetic education, counseling, and psychosocial support. What remains to be seen is whether these outcomes change as the test is offered in non-research settings to people from families without a previously identified gene mutation.

If one believes that the issue of certainty is a critical component in genetic testing, then one could hypothesize that those individuals with a greater need for certainty will be more inclined to seek testing when they perceive it to satisfy their needs. This has been demonstrated by Croyle et al. (1995), who showed that individuals with a high need for certainty were less interested in a *BRCA1* test when the information that breast cancer affects one in nine women regardless of genetic status was included than when this information was omitted.

Another cause of uncertainty in gene mutation carriers is that of deciding treatment prevention, and surveillance options. Unlike HD in which there are no effective treatments or cures, some cancers have proven surveillance, prevention, and treatments. Studies have shown that a low-fat, high-fiber diet can help prevent colon cancer, but it is unclear whether a low-fat diet can help prevent breast cancer as well. In colon cancer, periodic colonoscopies have been shown to be effective in detecting early-stage polyps which can then be removed (Newcomb, Norfleet, & Storer, 1992; Winawer et al., 1993). If an advanced tumor is detected in the colon, a colectomy (removal of the colon), although not always a comfortable choice, can effectively stop spread of the cancer. In breast cancer, the merits of surveillance techniques such as breast self-exam (BSE) and mammograms remain controversial, but they are still recommended by most physicians (Olufunmilayo et al., 1997). Women who discover through genetic testing that they are at increased risk of breast and ovarian cancer are faced with choices regarding BSE, mammographies, prophylactic mastectomies, and prophylactic oopherectomies (Schrag, Kuntz, Garber, & Weeks, 1997). All of these choices can create a feeling of confusion and uncertainty, and psychological support and counseling throughout the decision-making process can be helpful. Research suggests that addressing distress related to cancer risk may also reduce maladaptive health behaviors. A study by Kash, Holland, Halper, & Miller (1992) showed that increased cancer anxiety decreased regular clinical examinations in women with a significant family history of breast cancer. High anxiety also predicted poor adherence to monthly self-breast examinations.

In addition to the need for certainty, psychological preparedness has also been shown to be related to anxiety in cancer susceptibility genetic testing. Marteau et al. (1988) found that individuals who were previously unaware of their genetic risk were more likely to be in distress after testing. People who are not aware of their genetic risk are not able to prepare themselves psychologically, and as a result, have a harder time dealing with risk notification. Croyle et al. (1997) found that mutation carriers who had never had cancer or cancer-related surgery showed more distress than those who had had cancer or cancer-related surgery. The researchers explain the finding as that the women who already had cancer may have been less surprised at their result. In addition, because they had experienced cancer-related distress in the past they were able to put the genetic test into the context of this perspective. The concept that lack of psychological preparedness may correlate with high anxiety levels is an important one to consider as new protocols are developed for clinical cancer susceptibility testing. As testing becomes more widespread, there is a greater likelihood that people may be tested in clinical settings with insufficient informed consent and education. In these scenarios individuals are likely to be ill-prepared for bad news and may experience increased levels of anxiety. This possibility emphasizes the importance of developing protocols that provide sufficient genetic information about the risk of being a carrier, and the psychological resources to help the patient negotiate personal and family implications of test results.

It is important for clinicians to assess family and social support networks prior to genetic testing. The lack of social support has been shown to be associated with higher levels of distress after receiving test results (Kash et al., 1992) Furthermore, the prospect of testing can stimulate unresolved grief in members of families with inherited cancer syndromes. The psychological impact of previous disease in the family can shape family members' beliefs about biomedical science and what medicine has to offer. Because family members who died of cancer might have been subjected to outdated treatments, it is important to explore the extent to which family members' views of the disease and its treatment are shaped by these prior experiences.

As with HD testing, psychological intervention is proving to be an important need for some individuals who undergo cancer susceptibility testing. Validated screening tools to assess a person's level of anxiety and depression are used in many research protocols to help identify individuals who need more extensive counseling than that provided by the genetic counselor or medical geneticist.

8.17.4 PSYCHOLOGICAL INTERVENTIONS

It is quite common for people affected by genetic disease to have ongoing relationships with physicians who treat their physical symptoms and help them maintain their physical health. However, it is not as common for them to seek help for psychological issues concerning their genetic condition. In the USA, the specialized field of genetic counseling grew partly out of the recognition that psychological issues related to genetic disease were important to acknowledge (Kessler, 1992). Nevertheless, most genetic counseling training programs in the USA focus on the teaching of genetic science rather than psychological issues and treatment techniques. In Europe, medical or clinical psychologists are often relied on to provide psychological counseling, with genetics information being provided by a physician with specialized training in medical genetics.

In their meetings with clients, medical geneticists and genetic counselors are able to explore some of the psychosocial issues associated with genetic conditions and testing.

However, these clinicians may not have an ongoing relationship with their clients, so it may not be possible for them to explore psychological issues in depth. For individuals who are affected by a genetic condition, considering genetic testing, or dealing with genetic results, regular, ongoing psychological intervention can be very helpful (Schild & Black, 1984).

8.17.4.1 Support Groups

Support groups represent the first type of psychological intervention designed to meet the needs of individuals affected by genetic disease. Historically, many support groups were formed by individuals affected by particular genetic diseases to fill the void of psychosocial services that addressed the needs of this population. Before becoming involved with a specialized support group, individuals with rare genetic conditions may never have encountered someone else with the same condition. For individuals who feel self-conscious and stigmatized by their conditions in most public situations, participation in a support group can be a liberating and comforting experience (Weiss & Mackta, 1996).

Genetic disease support groups allow individuals to share their feelings about aspects of their genetic condition and help each other anticipate psychological challenges. Support groups can be especially helpful for individuals going through certain life tasks that are affected by their particular genetic disease (Weiss & Mackta, 1996). For example, teenagers with cystic fibrosis may find it helpful to talk with one another about what it is like to be limited in their physical activities, how they tell their friends about their condition, and how they deal with an emerging sexuality in the face of the disease.

Support groups can also prove helpful for individuals close to a person with a genetic condition. The spouse of an individual affected with HD has been referred to as a forgotten person (Kessler, 1993). They may spend so much of their time caring for their affected spouse that they often neglect their own emotional needs. A support group allows caregivers to voice their fears about their spouse's deterioration, their worry about their children's risk, and their frustration at the increasing medical costs, all in an environment that does not make them feel as if they are being selfish for having such concerns. Couples who have lost a child to a genetic condition or who have terminated a pregnancy after an abnormal prenatal test can also benefit by going through the grieving process with others who have been in similar situations. Especially for couples who have terminated a pregnancy because of a genetic condition and are fearful of telling others, a forum of honesty and openness without judgment is extremely welcome.

Although support groups are helpful for many, they are not appropriate forums for everyone affected by a genetic condition. Some groups emphasize political advocacy, a role that some might wish to avoid. For an individual struggling with the initial acceptance of her diagnosis, a support group with advanced-stage patients may cause this individual to distance herself from the others because she does not want to admit that she is similar to them. Peer-led groups can sometimes disseminate inaccurate medical information. Furthermore, although the ideal support group is a caring, supportive, and nurturing environment, not all of them function in this way. Support groups may not be beneficial if there is too wide a variety of conditions, beliefs, or goals represented (Weiss & Mackta, 1996). For example, a general pregnancy-loss support group may not be the best place for couples who have terminated a pregnancy because of a genetic condition. In a group such as this, they may be judged by others because they chose to terminate instead of miscarrying their pregnancy. Furthermore, a couple who decided to terminate a fetus because an amniocentesis showed Down's syndrome may be condemned by another group member who believes that only life-threatening conditions are justified terminations. Finally, if they are populated by individuals with the need to constantly talk about themselves, support groups can turn into competitions for floor time where nobody actually listens to each other. A person who attends an unrewarding support group may need to find another one more suited to their needs or pursue some other form of psychological intervention.

8.17.4.2 Individual Therapy

Individual therapy may be a more appropriate type of psychological intervention for individuals who do not find comfort in the support group environment. Others may use it in conjunction with a support group to talk about some of their more private or in-depth issues. For an individual coping with a genetic disease, individual therapy can allow the person to discuss the physical and psychological difficulties of the disease and work on self-esteem issues. For an individual considering genetic testing, individual therapy can serve as a forum for a discussion about the possible pros and cons of each decision, along with preparing the individual for each possible outcome. A person dealing with the results of a genetic test may pursue individual therapy in order to talk about feelings of depression, suicidal ideations, or guilt associated with the test result.

Individual therapy may not be the most appropriate mode for some, however. As described earlier in this chapter, families and society influence a person's interpretation of and coping strategy for a genetic disease. By focusing mainly on the intrapsychic processes, individual therapy may not be able to address many of the interpersonal factors involved in coping within the family context. Clinicians who provide individual counseling must also be responsible for educating themselves about the disease and its medical implications. Close collaboration with a medical geneticist or genetic counselor is recommended.

8.17.4.3 Family Therapy

As previously described, one of the unique features of genetic conditions is that they may be inherited biologically from parent to child. The psychological interpretation given to a genetic disease is also inherited through the family. Since family dynamics regarding the disease are crucial to how the disease is interpreted in a given family, family therapy may be especially helpful for families affected by a genetic disease. In therapy, family members can discuss ways in which each other affect the coping process. Persons in the family with the disease may feel stigmatized or infantized, so family therapy gives them a forum to discuss these feelings. Individuals in the family without the disease sometimes feel ignored or burdened with care. Family therapy allows these individuals to have a voice when they may not have necessarily felt they had one before.

When family members who have a differing genetic status become involved in family therapy there may, however, be a conflict of interest. It may be difficult to negotiate different goals and concerns of various family members with a genetic disease and without the disease. In families affected by HD, the psychological needs of those who have tested positive for a gene mutation, those who have tested negative, and those who have chosen not to be tested may be too diverse to be addressed in family therapy. Ideally, family members should be provided with opportunities to discuss psychosocial and decision-making issues both within and outside of the family context.

8.17.4.4 Psychological Intervention Within the Medical Setting

Recent research has focused on psychological intervention for high-risk individuals within the traditional medical care setting. A study by Lerman et al. (1995) evaluated a standardized protocol for individualized risk assessment and counseling for first-degree relatives of breast cancer patients. Previous studies had shown that the majority of women with family histories of breast cancer significantly overestimate their risk of cancer and have higher levels of anxiety. High levels of anxiety have been associated in some studies with decreased adherence to mammography (Lerman, Daly, Sands et al., 1993) and fewer clinical and self-breast exams (Kash et al., 1992). Lerman et al. (1995) found that the women who received risk counseling were more likely to have an accurate estimation of their cancer risk. However, among those women with high levels of anxiety at baseline, risk counseling did not improve their risk comprehension. Lerman and her colleagues concluded that, to be effective, cancer risk counseling should address the issues of cancer-related anxiety (Lerman et al., 1995).

Gagnon et al. (1996) found that women who attended structured surveillance programs showed significant improvement in their perceived risk comprehension and reductions in their anxiety levels, but found no change in performance of breast self-exam. Lloyd et al. (1996) conducted a retrospective study comparing women with family histories of breast cancer who had received genetic counseling with controls who had no family history of cancer. They found that after genetic counseling, women with family histories of cancer had more cancer-related distress than the control group, and concluded that the finding suggests that these women have significant anxieties that are not alleviated through the process of genetic counseling. Although this research is in the preliminary stages, it suggests that traditional risk counseling (genetic education) may not be effective in alleviating anxiety for women at increased cancer risk, thus highlighting the need to supplement these services with psychological counseling.

8.17.5 PROFESSIONAL ISSUES

8.17.5.1 Role of the Genetic Counselor

In the USA, genetic counselors are health professionals usually trained with a 2 year postgraduate masters degree (Thompson et al., 1991). They provide genetic education and psychosocial support for individuals affected by genetic disease or who are considering genetic testing (Fraser, 1974). Genetic counselors often juggle duties of giving genetic information and providing psychological support. They are trained to give information to help a person make an informed decision, but do so in a way that is as nondirective as possible (Harper, 1993). Physicians who are trained in medical genetics often use a similar approach. Because genetic counselors and medical geneticists focus on knowledge acquisition and informed decision-making as a goal for their patients, clinical psychologists can focus on the broader meaning and implications of genetic tests results or genetic disease.

8.17.5.2 Role of the Psychologist

There are many ways in which psychologists can complement the medical services received by individuals who are affected by a genetic condition. In academic settings a common role is the development of evaluation or research protocols concerning the psychological impact of genetic disease, genetic counseling, and genetic testing. When new genetic tests are introduced into clinical practice, psychologists are often called upon to conduct psychiatric screening or baseline assessments of test participants. Screening for depression is especially important, given research suggesting that depression is relatively common among individuals who attend specialized clinics for genetic risk evaluation (Robinson et al., 1997). For psychologists working in a hospital or other medical setting, individuals with genetic conditions may be referred for assessment or therapy, and psychologists can play a role in the development of referral criteria. Neuropsychological assessment is commonly utilized for genetic conditions that affect nervous system functioning, and psychologists can help educate genetic counselors in the appropriate interpretation of these findings. Psychologists also encounter grief and bereavement issues in parents who have lost a child to a genetic disorder, couples who terminated a pregnancy based upon genetic findings, or a person who has lost a family member to a hereditary disease. Finally, psychology professionals may be called upon to be facilitators for genetic support groups. In all of these roles, it is critical that psychologists have at least a basic familiarity with the medical features of the disease and the surveillance or treatment options available.

8.17.6 CONCLUSIONS

The rapid pace of advances in genetic science and clinical genetics suggests that clinical psychologists will encounter patients with genetics-related disorders and concerns more frequently. In many medical settings, the role of the clinical psychologist in genetics services has yet to be established. In Europe, where genetic testing tends to be centralized within academic medical research centers, it may be easier for psychologists to establish and maintain a role in providing support services to patients. In the USA, where competitive forces encourage biotechnology companies to market genetic testing laboratory services to community physicians, there are fewer incentives to provide the comprehensive quality of care that psychologists can help to provide. By informing themselves about the distinctive concerns and challenges faced by individuals at-risk of or affected by genetic disease, psychologists stand a better chance of filling an important role in application of the new genetics to patients and their families.

8.17.7 REFERENCES

ASHG Ad Hoc Committee on Cystic Fibrosis Carrier Screening (1992). Statement of the American Society of Human Genetics on Cystic Fibrosis carrier screening. *American Journal of Human Genetics, 51,* 1443–1444.

Baty, B. J., Venne, V. L., McDonald, J., Croyle, R. T., Halls, C., Nash, J. E., & Botkin, J. R. (1997). BRCA1 testing: Genetic counseling protocol development and counseling issues. *Journal of Genetic Counseling, 6,* 223–244.

Beeson, D., & Golbus, M. S. (1979). Anxiety engendered by amniocentesis. *Birth Defects: Orginal Article Series, 15,* 191–197.

Bloch, M. Adam, S., Wiggins, S., Huggins, M., & Hayden, M. R. (1992). Predictive testing for Huntington disease in Canada: the experience of those receiving an increased risk. *American Journal of Medical Genetics, 42*(4), 499–507.

Bloch, M., Fahy, M., Fox, S., & Hayden, M. R. (1989). Predictive testing for Huntington disease: II. Demographic characteristics, life-style patterns, attitudes, and psychosocial assessments of the first fifty-one test candidates. *American Journal of Medical Genetics, 32,* 217–224.

Blumberg, B. D., Golbus, M. S., & Hanson, K. H. (1975). The psychological sequelae of abortion performed for a genetic indication. *American Journal of Obstetrics and Gynecology, 122,* 799–808.

Brambati, B., Lanzani, A., & Tului, L. (1990) Transabdominal and transcervical chrorionic villuis sampling: Efficiency and risk evaluation of 2,411 cases. *American Journal of Medical Genetics, 35,* 160–164.

Canadian Collaborative CVS–Amniocentesis Clinical Trial Group (1989). Multicentre randomised clinical trial of chorionic villus sampling and amniocentesis: First report. *Lancet, 1,* 1–6.

Childs, B., Gordis, L., Kaback, M. M., & Kazazian, H. H., Jr. (1976). Tay-Sachs screening: Social and psychological impact. *American Journal of Human Genetics, 28,* 550–558.

Codori, A. M., Hanson, R., & Brandt, J. (1994). Self-selection in predictive testing for Huntington's disease (see comments). *American Journal of Medical Genetics, 54*(3), 167–173.

Colditz, G. A., Willett, W. C., Hunter, D. J., Stampfer, M. J., Manson, J. E., & Hennekens, C. H. (1993). Family history, age, and risk of breast cancer: Prospective data from the Nurses' Health Study. *Journal of the American Medical Association, 270,* 338–343.

Croyle, R. T., & Lerman, C. (1993). Interest in genetic testing for colon cancer susceptibility: cognitve and emotional correlates. *Preventive Medicine, 22,* 284–292.

Croyle, R. T., & Lerman, C. (1995). Psychological impacts of genetic testing. In R. T. Croyle (Ed.), *Psychosocial effects of screening for disease prevention and detection.* New York: Oxford University Press.

Croyle, R. T., Dutson, D. S., Tran, V. T., & Sun, Y.-C. (1995). Need for certainty and interest in genetic testing. *Women's Health: Research on Gender, Behavior, and Policy, 1,* 329–339.

Croyle, R. T., Smith, K. R., Botkin, J. R., Baty, B., & Nash, J. (1997). Psychological responses to BRCA1 mutation testing: Preliminary findings. *Health Psychology, 16*(1), 63–72.

Cullinan, A. L. (1989). Psychological impact of genetic disease on individual. family, and societal systems. *Genetic disease: The unwanted inheritance.* New York: Haworth Press.

Farrer, L. A. (1986). Suicide and attempted suicide in Huntington disease: Implications for preclinical testing of persons at risk. *American Journal of Medical Genetics, 24,* 305–311.

Ford, D., Easton, D. F., Bishop, D. T., Narod, S. A., Goldgar, D. E., & Breast Cancer Linkage Consortium. (1994). Risks of cancer ion BRCA1-mutation carriers. *Lancet, 343,* 692–695.

Fraser, F. C. (1974). Genetic counseling. *American Journal of Human Genetics, 26,* 636–659.

Gagnon, P., Massie, M. J., Kash, K. M., Gronert, M., Heerdt, A. S., Brown, K., Sullivan, M. D., & Borgen, P. (1996). Perception of breast cancer risk and psychological distress in women attending a surveillance program. *PsychoOncology, 5,* 259–269.

Geller, G., Botkin, J. R., Green, M. J., Press, N., Biesecker, B. B., Wilfond, B., Grana, G., Daly, M. B., Schneider, K., & Kahn, M. J. E. (1997). *Journal of the American Medical Association, 277*(18), 1467–1474.

Giardiello, F. M., Brensinger, J. D., Petersen, G. M., Luce, M. C., Hylind, L. M., Bacon, J. A., Booker, S. V., Parker, R. D., & Hamilton, S. R. (1997). The use and interpretation of commercial APC gene testing for Familial Adenomatous Polyposis. *The New England Journal of Medicine, 336*(12), 823–827.

Gould, R. L. (1997). *Cancer and genetics: Answering your patients' question.* Huntington, NY: PRR Inc.

Green, J., & Statham, H. (1996). Psychosocial aspects of prenatal screening and diagnosis. In T. Marteau & M. Richards (Eds.)., *The troubled helix: Social and psychological implications of the new human genetics.* Cambridge, UK: Cambridge University Press.

Gusella, J. F., Wexler, N. S., Conneally, P. M., Naylor, S. L., Anderson, M. A., Tanzi, R. E., Watkins, P. C., Ottina, K., Wallace, M. R., Sakaguchi, A. Y., Young, A. B., Shoulson, I., Bonilla, E., & Martin, J. B. (1983). A polymorphic DNA marker genetically linked to Huntington's disease. *Nature, 306,* 234–238.

Harper, P. S. (1993). *Practical genetic counseling.* Oxford, UK: Butterworth-Heinemann.

Healy, B. (1997). BRCA genes—Bookmaking, fortunetelling, and medical care. *The New England Journal of Medicine, 336*(20), 1448–1449.

Hoskins, K. F., Stopfer, J. E., Calzone, K. A., Merajver, S. D., Rebbeck, T. R., Garber, J. E., & Weber, B. L. (1995). Assessment and counseling for women with a family history of breast cancer: A guide for clinicians. *Journal of the American Medical Association, 273*(7), 577–585.

Huggins, M., Bloch, M., Wiggins, S., Adam, S., Suchowersky, O., Trew, M., Klimek, M., Greenberg, C. R., Eleff, M., Thompson, L. P., Knight, J., MacLeod, P., Girard, K., Theilman, J., Hedrick, A., & Hayden, M. R. (1992). Predictive testing for Huntington disease in Canada: adverse effects and unexpected results in those receiving a decreased risk. *American Journal of Medical Genetics, 42*(4), 508–515.

Huntington's Disease Collaberative Research Group (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell, 72,* 971–983.

Huntington's Disease Society of America (1994). *Huntington's disease (HD): Facts at a glance* (Brochure). New York: Author.

Huntington Society of Canada (1993). *Predictive testing for Huntington's disease* (Brochure). Cambridge, ON: Author.

Jorgensen, C., Uddenberg, N., & Ursing, Z. (1985). Ultrasound diagnosis of fetal malformation in the second trimester: the psychological reactions of the women. *Journal of Psychosomatic Obstetrics and Gynecology, 4,* 31–40.

Kash, K. M., Holland, J. C., Halper, M. S., & Miller, D. G. (1992). Psychological distress and surveillance behaviors of women with a family history of breast cancer. *Journal of the National Cancer Institute, 84*(1), 24–30.

Kelly, P. T. (1977). *Dealing with dilemma: A manual for genetic counselors.* New York: Springer.

Kessler, S. (1979). The psychological foundations of genetic counseling. In S. Kessler (Ed.), *Genetic counseling: psychological dimensions.* New York: Academic Press.

Kessler, S. (1992). Process issues in genetic counseling. In G. Evers-Kiebooms, J-P. Fryns, J-J. Cassiman, & H. Van den Berghe (Eds.), *Psychosocial aspects of genetic counseling.* New York: Wiley-Liss.

Kessler, S. (1993). Forgotten person in the Huntington disease family. *American Journal of Medical Genetics, 48,* 145–150.

Kessler, S., Field, T., Worth, L., & Mosbarger, H. (1987). Attitudes of persons at risk for Huntington disease toward predictive testing. *American Journal of Medical Genetics, 26*(2), 259–70.

Korf, B. R. (1996). *Human genetics: A problem-based approach.* Cambridge, MA: Blackwell Science.

Kushner, H. (1981). *When bad things happen to good people.* New York: Schocken Books.

Lam, R. W., Bloch, M., Jones, B. D., Marcus, A. M., Fox, S., Amman, W., & Hayden, M. R. (1988). Psychiatric morbidity associated with early clinical diagnosis of Huntington disease in a predictive testing program. *Journal of Clinical Psychiatry, 49*(11), 444–447.

Lerman, C., Daly, M., Sands, C., Balshem, A., Lustbader, E., Heggan, T., Goldstein, L., James, J., & Engstrom, P. (1993). Mammography adherence and psychological distress among women at risk for breast cancer. *Journal of the National Cancer Institute, 85,* 1074–1080.

Lerman, C., Lustbader, E., Rimer, B., Daly, M., Miller, S., Sands, C., & Balshem, A. (1995). Effects of individualized breast cancer risk counseling: A randomized trial. *Journal of the National Cancer Institute, 87*(4), 286–292.

Lerman, C., Narod, S., Schulman, K., Hughes, C., Gomez-Caminero, A., Bonney, G., Gold, K., Trock, B., Main, D., Lynch, J., Fulmore, C., Snyder, C., Lemon, S. J., Conway, T., Tonin, P., Lenoir, G., & Lynch, H. (1996). BRCA1 testing in families with hereditary breast-ovarian cancer: Prospective study of patient decision making and outcomes. *Journal of the American Medical Association, 275*(24), 1885–1892.

Lloyd, S., Watson, M., Waites, B., Meyer, L., Eeles, R., Ebbs, S., & Tylee, A. (1996). Familial breast cancer: a controlled study of risk perception, psychological morbidity and health beliefs in women attending for genetic counselling. *British Journal of Cancer, 74,* 482–487.

Lynch, H. T., Lemon, S. J., Durham, C., Tinley, S. T., Connolly, C., Lynch, J. F., Surdam, J., Orinion, E., Slominski-Caster, S., Watson, P., Lerman, C., Tonin, P., Lenoir, G., Serova, O., & Narod, S. (1997). A descriptive study of BRCA1 testing and reactions to disclosure of test results. *Cancer, 79*(11), 2219–2228.

Lynch, H. T., Lynch, P. M., & Lynch, J. F. (1979) Genetic counseling and cancer. In S. Kessler (Ed.), *Genetic counseling: psychological dimensions*. New York: Academic Press.

Mange, A. P. & Mange, E. J. (1990). *Genetics: Human aspects*. Sunderland, MA: Sinauer Associates.

Markel, D. S., Young, A. B., & Penney, J. B. (1987). At-risk person's attitudes toward presymptomatic and prenatal testing of Huntington's disease in Michigan. *American Journal of Medical Genetics, 26*(2), 295–305.

Marteau, T. M. (1992). Psychological implications of genetic screening. In G. Evers-Kiebooms, J-P. Fryns, J-J. Cassiman, & H. Van den Berghe (Eds.), *Psychosocial aspects of genetic counselling*. New York: Wiley-Liss.

Marteau, T. M., Kidd, J., Cook, R., Johnston, M., Michie, S., Shaw, R. W., & Slack, J. (1988). Screening for Down's syndrome (letter). *British Medical Journal, 297,* 1469.

Massarick, F., & Kaback, M. M. (1981). *Genetic disease control: A social psychological approach*. Beverly Hills, CA: Sage.

Mastromauro, C., Myers, R. H., & Berkman, B. (1987). Attitudes toward presymptomatic testing in Huntington's disease. *American Journal of Medical Genetics, 26*(2), 271–282.

Medical Research Council Working Party on Amniocentesis (1976). An assessment of the hazards of amniocentesis. *British Journal of Obstetrics and Gynaecology, 85,* 1–14.

Merkatz, I. R., Nitowski, H. M., Macri, J. N., & Johnson, W. E. (1984). An association between low maternal serum alpha feto-protein and fetal chromosome abnormalities. *American Journal of Obstetrics and Gynecology, 148,* 866–894.

Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L. M., Ding, W., Bell, R., Rosenthal, J., Hussey, C., Tanh, T., McClure, M., Frye, C., Hattier, T., Phelps, R., Haugen-Strano, A., Katcher, H., Yakumo, K., Gholami, Z., Shaffer, D., Stone, S., Bayer, S., Wray, C., Bogden, R., Dayanath, P., Ward, J., Tonin, P., Narod, S., Bristow, P. K., Norris, F. H., Helvering, L., Morrison, P., Rosteck, P., Lai, M., Barrett, J. C., Lewis, C., Neuhausen, S., Cannon-Albright, L., Goldgar, D., Wiseman, R., Kamb, A., & Skolnick, M. H. (1994). A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science, 266,* 66–71.

Natowicz, M. R. & Alper, J. S. (1991). Genetic screening: Triumphs, problems, and controversies. *Journal of Public Health Policy, Winter,* 475–491.

Newcomb, P. A., Norfleet, N. G., & Storer, B. E. (1992). Screening sigmoidoscopy and colorectal cancer mortality. *Journal of the National Cancer Institute, 84,* 1572–1575.

Nora, J. J. (1994). *Medical genetics: Principles and practice*. Malvern, PA: Lea & Febiger.

Olufunmilayo, O. I., Blackwood, A., Cummings, S., Daly, M. B., Davies, J. G., Dragon, L., Garber, J. E., Lichter, A. S., Lynch, P., Offit, K., Peterson, G. M., Thigpen, T. J., Tomlinson, G., Strauss-Tranin, A., Weber, B. L., & Wells, S. A. (1997). *Resource document for curriculum development in cancer genetics education*. Unpublished manuscript, American Society of Clinical Oncology.

Patenaude, A. F., Schneider, K. A., Kieffer, S. A., Calzone, K. A., Stopfer, J. E., Basili, L. A., Weber, B. L., & Garber, J. E. (1996). Acceptance of invitations for p53 and BRCA1 predisposition testing: factors influencing potential utilization of cancer genetic testing. *Psycho-Oncology, 5,* 241–250.

Phipps, S., & Zinn, A. B. (1986). Psychological response to amniocentesis: II. Effects of coping style. *American Journal of Medical Genetics, 25,* 143–148.

Punales-Morejon, D., & Penchaszadeh, V. B. (1992). Psychosocial aspects of genetic counseling: Cross-cultural issues. In G. Evers-Kiebooms, J-P. Fryns, J-J. Cassiman, & H. Van den Berghe (Eds.), *Psychosocial aspects of genetic counseling*. New York: Wiley-Liss.

Quaid, K. A. (1992). Presymptomatic testing for Huntington disease: Recommendations for counseling. *Journal of Genetic Counseling, 1,* 277–302.

Quaid, K. A., & Morris, M. (1993). Reluctance to undergo predictive testing: the case of Huntington disease. *American Journal of Medical Genetics, 45*(1), 41–45.

Quaid, K. A., & Wesson, M. K. (1995). Exploration of the effects of predictive testing for Huntington disease on intimate relationships. *American Journal of Medical Genetics, 57*(1), 46–51.

Robinson, J. O., Hibbard, B. M., & Laurence, K. M. (1984). Anxiety during a crisis: emotional effects of screening for neural tube defects. *Journal of Psychosomatic Research, 28*(2), 163–169.

Robinson, G. E., Rosen, B. P., Bradley, L. N., Rockert, W. G., Carr, M. L., Cole, D. E. C., & Murphy, K. J. (1997). Psychological impact of screening for familial ovarian cancer: Reactions to initial assessment. *Gynecologic Oncology, 65,* 197–205.

Rosenfeld, D. L. (1984). Infertility. In A. E. H. Emery & I. Pullen (Eds.), *Psychological aspects of genetic counselling*. London: Academic Press.

Schild, S. & Black, R. B. (1984). *Social work and genetics: A guide for practice*. New York: Haworth Press.

Schrag, D., Kuntz, K. M., Garber, J. E., & Weeks, J. C. (1997). Decision analysis—Effects of prophylactic mastectomy and oophorectomy on life expectancy among women with BRCA1 or BRCA2 mutations. *The New England Journal of Medicine, 336*(20), 1465–1471.

Shapiro, J. (1983). Family reactions and coping strategies in response to the physically ill or handicapped child: A review. *Social Science Medicine, 17,* 913–931.

Slattery, M. L. & Kerber, R. A. (1993). A comprehensive evaluation of family history and breast cancer risk: The Utah population database. *Journal of the American Medical Association, 270,* 1563–1568.

Struewing, J. P., Hartge, P., Wacholder, S., Baker, S. M., Berlin, M., McAdams, M., Timmerman, M. M., Brody, L. C., & Tucker, M. A. (1997). The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. *The New England Journal of Medicine, 336*(20), 1401–1408.

Tabor, A., & Jonsson, M. H. (1987). Psychological impact of amniocentesis in low risk women. *Prenatal Diagnosis, 7,* 443–449.

Tabor, A., Madsen, M., Obel, E. B., Philip, J., Bang, J., & Norgaard-Pedersen, B. (1986). Randomised controlled trial of genetic amniocentesis in 4606 low-risk women. *Lancet, 1,* 1287–1293.

Thompson, M. W., McInnes, R. R., & Willard, H. F. (1991). *Genetics in medicine*. Philadelphia: W. B. Saunders.

Tibben, A., Duivenvoorden, H. J., Niermeijer, M. F., Vegter-van der Vlis, M., Roos, R. A. C., & Verhage, F. (1994). Psychological effects of presymptomatic DNA testing for Huntington's disease in the Dutch Program. *Psychosomatic Medicine, 56,* 526–532.

Tibben, A., Duivenvoorden, H. J., Vegter van der Vlis, M., Niermeijer, M. F., Frets, P. G., van de Kamp, J. J., Roos, R. A., Rooijmans, H. G., & Verhage, F. (1993). Presymptomatic DNA testing for Huntington disease: identifying the need for psychological intervention. *American Journal of Medical Genetics, 48*(3), 137–144.

Tyler, A., Ball, D., & Crauford, D. (1992). Presymptomatic testing for Huntington's disease in the United Kingdom. *British Medical Journal, 304,* 1593–1596.

United States Institute of Child Health and Human Development Study Group (1976). *Journal of the*

Americal Medical Association, 236, 1471–1476.

University of Michigan Medical Center (1994). *Genetic counseling and testing for Huntington's disease* (Brochure). Ann Arbor, MI: Author.

Vernon, S. W., Gritz, E. R., Peterson, S. K., Amos, C. I., Perz, C. A., Baile, W. F., & Lynch, P. M. (1997). Correlates of psychologic distress in colorectal cancer patients undergoing genetic testing for hereditary colon cancer. *Health Psychology, 16*(1), 73–86.

Viskochil, D. (1995). Identification and characterization of the Neurofibromatosis 1 gene. In D. J. Shaw (Ed.), *Molecular genetics of human inherited disease*. Chichester, UK: Wiley.

Watson, E. K., Mayall, E. S., Lamb, J., Chapple, J., & Williamson, R. (1992). Psychological and social consequences of community carrier screening programme for cystic fibrosis. *Lancet, 340,* 217–220.

Weiss, J. O. (1981). Psychosocial stress in genetic disorders: A guide for social workers. *Social Work in Health Care, 6*(4), 17–31.

Weiss, J. O. & Mackta, J. S. (1996). *Starting and sustaining genetic support groups*. Baltimore: Johns Hopkins University Press.

Wexler, N. S. (1979). Genetic "Russian roulette." In S. Kessler (Ed.), *Genetic counseling: psychological dimensions*. New York: Academic Press.

Wiggins, S., Whyte, P., Huggins, M., Adam, S., Theilmann, J., Bloch, M., Sheps, S. B., Schechter, M. T., & Hayden, M. R. (1992). The psychological consequences of predictive testing for Huntington's disease. *New England Journal of Medicine, 327*(20), 1401–1405.

Wilke, J. (1995). From a survivor: The emotional experience of genetic testing. *Journal of Psychosocial Nursing, 33*(4), 28–37.

Winawer, S. J., Zauber, A. G., Ho, M. N., O'Brien, M. J., Gottlieb, L. S., Sternberg, S. S., Waye, J. D., Schapiro, M., Bond, J. H., Panish, J. F., Ackroyd, F., Shike, M., Kurtz, R. C., Hornsby-Lewis, L., Gerdes, H., Stewart, E. T., & The National Polyp Study Workgroup. (1993). Prevention of colorectal cancer by colonoscopic polypectomy. *The New England Journal of Medicine, 329,* 1977–1981.

Wooster, R., Neuhausen, S. L., Mangion J., Quirk, Y., Ford, D., Collins, N., Nguyen, K., Seal, S., Tran, T., Averill, D., Fields, P., Marshall, G., Narod, S., Lenior, G. M., Lynch, H., Feunteun, J., Devilee, P., Cornelisse, C. J., Menko, F. H., Daly, P. A., Ormiston, W., McManus, R., Pye, C., Lewis, C. M., Cannon-Albright, L. A., Peto, J., Ponder, B. A. J., Skolnick, M. H., Easton, D. F., Goldgar, D. E., & Stratton, M. R. (1995). Identification of the breast cancer susceptibility gene BRCA2. *Nature, 378,* 789–792.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.18
Arthritis

JULIE H. BARLOW
Coventry University, UK

8.18.1 INTRODUCTION 427

8.18.2 DISEASE CHARACTERISTICS AND TREATMENT 428

8.18.2.1 Osteoarthritis 428
8.18.2.2 Rheumatoid Arthritis 428
8.18.2.3 Ankylosing Spondylitis 429
8.18.2.4 Juvenile Chronic Arthritis/Juvenile Rheumatoid Arthritis 429

8.18.3 THE ROLE OF PSYCHOLOGICAL FACTORS IN ONSET AND COURSE OF DISEASE 429

8.18.4 PSYCHOLOGICAL ASPECTS OF SYMPTOMS, DIAGNOSIS, AND TREATMENT 430

8.18.5 IMPACT ON PERSON AND FAMILY 430

8.18.5.1 Psychological Impact 430
8.18.5.2 Children with JCA 432
8.18.5.3 Social Impact: Family 432
8.18.5.4 Social Impact: Society 433

8.18.6 ADHERENCE 434

8.18.7 PSYCHOLOGICAL INTERVENTIONS 434

8.18.7.1 Information and Instruction 434
8.18.7.2 Cognitive Behavioral Interventions 435
8.18.7.3 Self-management Interventions 436
8.18.7.4 Interventions in JCA 437
8.18.7.5 Effectiveness of Psychological Interventions 438

8.18.8 PROFESSIONAL ISSUES 439

8.18.9 CONCLUSIONS AND FUTURE ISSUES 439

8.18.10 REFERENCES 440

8.18.1 INTRODUCTION

Arthritis is one of the most common chronic conditions, affecting millions of people worldwide. In the USA, prevalence rates of self-reported arthritis are projected to increase from 15% (37.9 million) in 1990 to 18.2% (59.4 million) by 2020 (Helmick, Lawrence, Pollard, Lloyd, & Heyse, 1995) with older people, women, and those with less education or lower incomes being at greater risk. Although not considered a major public health problem, arthritis is the largest single cause of physical disability in the UK (Badley & Tennant, 1993) and is linked to increased rates of mortality (Pincus & Callahan, 1993). Most forms of arthritis follow an unpredictable course of exacerbations and remissions, resulting in varying degrees of physical deformity. Prognosis is uncertain, and, since there is no cure, treatment is ameliorative aiming to alleviate inflammation, reduce pain, and preserve or improve function. Arthritis comprises over 200 different chronic diseases involving the joints and/or connective tissues, the most prevalent forms being osteoarthritis, rheumatoid arthritis and, ankylosing spondylitis in adults (Taal, Seydel, Rasker, & Weigman, 1993) and juvenile chronic arthritis (JCA) in children. Other forms of arthritis (e.g., systemic lupus erythematosus, scleroderma, Sjogren's syndrome, psoriatic arthritis, gout) are less common.

Medical assessment and evaluation of the patient includes clinical history, a physical examination, and laboratory tests. The pain of arthritis is the most frequent factor causing patients to seek health care. Characteristics, such as location, quantity, intensity, nature, and the course of pain can assist diagnosis. Other symptoms of arthritis include stiffness, limited motion, fatigue, weakness, and swollen joints. The latter can be assessed by inspection or by direct palpation of the joint. A count of the number of swollen joints provides an indication of the amount of inflamed tissue. Use of diaries could prove a useful adjunct to traditional methods of pain assessment (e.g., visual analogue scales) and can incorporate ratings of stiffness, fatigue, and mood. General laboratory tests can help to determine whether arthritis, and its treatment, have affected major body systems. For example, a full blood count may reveal the presence of anemia, a common feature of arthritis. In rheumatoid arthritis (RA), 80% of patients will test positive for rheumatoid factor, an antibody directed against immunoglobulin. Rheumatoid factor is associated with increased disease severity, the development of erosions, extra-articular manifestation, and greater disability (Harris, 1993). Synovial fluid aspiration and examination can aid diagnosis of joint pain. Levels of acute phase reactants (e.g., C-reactive protein, erythrocyte sedimentation rate) are used to monitor disease activity. Radiography, or X rays, are helpful both diagnostically and in monitoring disease progression. Other imaging techniques include ultrasonography, computed tomography, and magnetic resonance imaging. The most widely used system of disease classification for prevalence and incidence studies is based on relevant sets of criteria published by the American Rheumatism Association (ARA). For example, the ARA (1987) criteria for RA include morning stiffness, arthritis of hand joints, serum rheumatoid factor, and radiographic changes.

8.18.2 DISEASE CHARACTERISTICS AND TREATMENT

8.18.2.1 Osteoarthritis

Osteoarthritis (OA) is generally acknowledged to be the most prevalent form of arthritis, particularly among older people. Prevalence rates vary according to the reporting methods used, the age and gender of study participants, and the number of joints studied. Early epidemiological studies in the UK, using radiographic change as the detection criterion, found that almost everyone aged 65 and over had OA in at least one joint (Lawrence, Bremmer, & Bier, 1966). Since the proportion of older people in the population is rising, increasing numbers of older adults will develop OA. The condition is characterized by progressive loss of articular cartilage and secondary reactions in bone causing joint pain and stiffness, particularly at the start of movement. Bony enlargement of affected joints can occur and mobility can become progressively restricted. Hutton (1995) suggests that since there is no treatment that can influence disease progression, clinical management should be based on principles of logic and not doing harm. Control of pain is through low-toxicity analgesia, and although considered a noninflammatory condition, anti-inflammatory drugs are often prescribed. Exercise, splinting, and use of aids (e.g., walking stick) can assist mobility. Outcome for people with severe hip or knee OA can be improved with joint replacement. Most people with OA remain under the care of a general practitioner, and are only referred to hospital-based clinics when problems become severe (e.g., need for hip replacement). Comorbidity is common among older people.

8.18.2.2 Rheumatoid Arthritis

RA, affects approximately 1% of the population (Schumaker, 1988), manifests between the ages of 20 and 50, and is more prevalent among women, with a ratio (women to men) of 3:1 (Anderson, Bradley, Young, McDanial, & Wise, 1985). The precise etiology of RA remains largely unknown, and although a hereditary predisposition (i.e., an association with the human leukocyte antigen, HLA-DR4), has been established, it is poorly understood (Hazes & Silman, 1990). In RA, the synovial membrane becomes thickened and inflamed, eventually resulting in degeneration of the cartilage and ultimately the joint. Primary symptoms are persistent pain, stiffness, swollen joints, and fatigue, anemia is common. RA is a systemic condition, therefore people can feel generally unwell in addition to experiencing problems with specific joints. Typically, joints in the hands and feet are affected, thus leading to impaired physical functioning and difficulty in performing activities of daily living. The economic impact of disability of resulting from RA in England for the year 1992 has been estimated at £1256 billion (McIntosh, 1996).

Most people with inflammatory arthritis, including those with RA, will be prescribed nonsteroidal anti-inflammatory drugs (NSAIDs) as first-line drug treatment. Second-line drugs may be used if NSAIDs fail to control symptoms. People with RA are advised to carry out regular, gentle exercise and to learn when to rest inflamed joints. Splints may be used in attempts to prevent permanent joint deformity.

Finally, where joints become damaged beyond repair, joint replacement may be necessary to relieve pain and to improve function. People with RA often require long-term monitoring and care by specialized rheumatology clinics.

8.18.2.3 Ankylosing Spondylitis

Ankylosing spondylitis (AS) is characterized by an early age of onset (under 40) and is one of the few rheumatic diseases to exhibit a male predominance, with reported ratios in the region of 3 men:1 woman (Kahn & van der Linden, 1990). The etiology of AS remains unknown, although an association with the antigen HLA-B27 was identified over 20 years ago (Brewerton, Caffrey, & Hart, 1973). The disease affects mainly the spine and sacroiliac joints, although other parts of the body (e.g. hips, shoulders, knees, eyes) can become involved. Gradual fusion of affected joints in the spine results in progressive functional impairment and the development of a typical hunched posture. Treatment is two-fold, consisting of medication and therapeutic exercise. The aim of medication is to reduce pain and inflammation. Performance of regular strengthening and stretching exercises is advised in order to maintain mobility of affected joints, to improve or maintain posture, and to achieve general fitness. After diagnosis, some patients continue to be monitored in hospital-based clinics, while others may be referred back to community-based medical practitioners for long-term healthcare. People with AS can be referred to a physiotherapy department for advice on home exercise activities and may also receive a short course of hydrotherapy, although such treatment is dependent upon the availability of local facilities.

8.18.2.4 Juvenile Chronic Arthritis/Juvenile Rheumatoid Arthritis

Contrary to the general belief that arthritis is a condition of "old age," some forms of arthritis begin in childhood. JCA is defined as an inflammatory arthritis with onset prior to a child's 16th birthday (Munthe, 1990). The condition is categorized into three subgroups: systemic onset disease, characterized by fever and a rash; pauciarticular disease, four or fewer joints affected in the absence of systemic features; and polyarticular disease where five or more joints are involved in the absence of systemic features. JCA is one of the most common chronic diseases of childhood with prevalence estimates ranging from 0.16 to 1.13 per 1000 children (Benjamin, 1990). JCA is associated with chronic pain, restricted joint mobility, and reduced physical function, placing children at risk of long-term physical disability. JCA is a complex condition requiring complex disease management best achieved by a pediatric rheumatology team (rheumatologists, nurses, physiotherapists, occupational therapists, and psychologists).

8.18.3 THE ROLE OF PSYCHOLOGICAL FACTORS IN ONSET AND COURSE OF DISEASE

Early studies implicated personality and stress as predisposing factors leading to the onset of RA (Booth, 1937; Jones, 1909). In a review of the literature, Moos (1964) concluded that people with RA are conforming, conservative, compliant, subservient, depressed, sensitive, inhibited, and anxious. However, evidence of causality is lacking given the methodological problems associated with retrospective studies and the identification of appropriate comparison groups. One attempt to compare the personality characteristics and stress at disease onset of patients with RA and OA failed to find evidence of a generalized, simple, and distinct RA personality (Latman & Walls, 1996), although there were indications that a subgroup of RA patients, characterized by high stress at disease onset, may experience a higher degree of disease severity.

Beliefs, deriving from clinical experience and anecdotal evidence, can evolve into stereotypes of "typical" patients and become associated with certain forms of arthritis. For example, compared with RA, people with AS are generally believed to be aggressive, active, less depressed, less emotional, more socially oriented, more highly motivated, and to possess a high pain threshold (Williams, 1989). It is important to note that RA is more predominant in women, with a typical age at onset between 40 and 50 years, whereas AS is more predominant in men, with a typical age at onset between 20 and 30 years. In addition, treatment for AS comprises regular performance of a relatively aggressive exercise regime, while treatment for RA places greater emphasis on the importance of rest during periods of disease exacerbation and traditionally positions the patients as passive recipients of care. One attempt to assess the validity of health professionals' beliefs compared patients with RA, AS, and chronic low back pain in terms of sthenia (noncomplaining, active), ambition, and education (Zant, Dekker-Saeys, Van den Burgh, Kolman, & Van der Stadt, 1982). Results showed that although AS patients had a significantly higher education level, they did not differ from the other patient groups or Dutch population norms on sthenia or ambition. Similarly, T. Stiles (personal communication, 1993) found no difference on personality traits between women with AS and either women receiving physical therapy for generalized aches and pains or healthy controls.

Many of the characteristics associated with arthritis (e.g., depression) may be reactions to a painful, chronic, and disabling disease, rather than causal factors. Nevertheless, there are indications that stress may trigger disease onset and may influence the course of established disease (Anderson et al., 1985; Lerman, 1987). Links have been established between stress and immune system functioning in people with RA (Bradley, 1989; Wegener, 1991). Using a combination of interviews and blood tests, an investigation of the relationship between life stress and immune parameters among 33 women with RA attending three routine, monthly, clinic checkups showed that stress resulted in negative immune system reactions (Zautra et al., 1989). However, the precise mechanisms involved remain unclear.

8.18.4 PSYCHOLOGICAL ASPECTS OF SYMPTOMS, DIAGNOSIS, AND TREATMENT

One difficulty in diagnosing many forms of arthritis is the potential overlap of disease symptomatology with psychological symptomatology. Intermittent pain, fatigue, and sleep disturbance can be symptoms of both arthritis and depression. This can delay the acquisition of a medical label for one's condition, particularly where blood tests are not clear diagnostic indicators. Equally, the overlap of symptoms can work in the opposite direction, masking psychological distress. Hence, some people with arthritis may not receive appropriate psychological care when needed. Poor detection of major depressive disorder occurs when symptoms (e.g., fatigue) are attributed to RA and the possibility of concurrent psychological disorder is not considered (Creed & Ash, 1992). The authors of a study of depression in AS concluded that such oversights are not limited to RA, but can occur in other forms of arthritis (Barlow, Macey, & Struthers, 1993).

A long delay between onset of symptoms and medical diagnosis is a characteristic of AS associated with greater adherence to therapeutic exercise (Barlow & Barefoot, 1996). Thus a potentially negative experience, a long period between onset and medical diagnosis, can have a positive outcome in terms of health behavior. However, this has to be offset against negative aspects such as the impact of being informed that pain, stiffness, and fatigue are "all in the mind" and do not have an organic cause. During this prediagnostic period people are often considered "neurotic" (Dekker-Saeys, 1976). The impact of being given such opinions during the delay between onset and diagnosis has not been fully examined.

One issue likely to become more salient in the field of arthritis concerns the implications of genetic screening. Hereditary factors have been identified and associated with certain forms of arthritis (e.g., RA, AS), a fact reflected in the concern expressed by adults regarding their decisions to have a family, and anxiety about "passing on" arthritis to children and grandchildren (Barlow & Cullen, 1996).

8.18.5 IMPACT ON PERSON AND FAMILY

The unpredictable course of disease, allied to an uncertain prognosis, combine with painful, disabling symptoms to pose considerable challenges for people with arthritis. The incurable nature of most forms of arthritis necessitates long-term disease management, including maintenance of complex treatment regimes. Not surprisingly, the impact of this chronic condition is wide ranging and pervades all areas of life, such as psychological well-being (Barlow, Macey, & Struthers, 1993; Creed & Ash, 1992), ability to work (Callahan, Bloch, & Pincus, 1992; Lubeck, 1995), social relationships (Fitzpatrick, Newman, Archer, & Shipley, 1991), and family life (Reisine, 1993). Paralleling the medical literature, psychological studies have tended to focus on people with RA. Thus psychosocial impact is less well documented among OA, AS, and JCA.

8.18.5.1 Psychological Impact

Although episodes of psychological distress may be experienced along the way, most people with arthritis appear to achieve a degree of adjustment to the challenge of their disease. Much of the documented work examining psychological impact has focused on depression among people with RA (see Anderson et al., 1985; Creed & Ash, 1992; Katz & Yelin, 1993, for reviews). The consensus is that depressive disorders and depressive symptoms are more common among people with RA, with prevalence rates equivalent to those in other clinic samples of people with chronic diseases. For example, a population-based study conducted in the USA reports a prevalence of 20% for depression in RA (Hawley & Wolf, 1993).Using a UK sample and the Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983), Pincus, Griffith, Pearce, and Isenberg (1996) report a prevalence of 23% compared with 6% among case controls (using a score of 8+ for indicating a possible clinical state), with prevalence decreasing to 15% using a more stringent cut-off point (11+) for determining "cases." Odds ratios for depression and anxiety showed that RA patients were twice as likely to be depressed and four times as likely to be anxious. The overlap between symptoms of depression and symptoms of arthritis can be problematic when the purpose of measurement is to identify cases or severity of depression, or to determine prevalence rates. Although the inclusion of somatic items on depression scales can inflate estimates of severity and prevalence in RA, the level of inflation is believed to be modest (Blalock, DeVellis, Brown, & Wallston, 1989).

Given the persistence of arthritis pain, it is not surprising that pain has been identified as a correlate of depression in RA (Creed & Ash, 1992) and AS (Barlow et al., 1993). A key issue yet to be resolved concerns the direction of causality. Results from longitudinal studies in RA have been inconsistent. For example, Brown (1990) suggests a causal model in which pain predicts subsequent depression. Nicassio and Wallston (1992) extended this work in a longitudinal study of 242 RA patients, showing that prior pain predicted adverse changes in sleep problems and that the interaction between high sleep problems and high pain were independently associated with depression. In contrast, Parker et al. (1992) found that depression had more influence over pain than vice versa.

Appraisals, expectations, and coping strategies have been investigated as potential mediators of pain, psychological well-being, and disability. Feelings of helplessness can arise owing to the uncertain disease course, lack of satisfactory pain control, and the interference of arthritis with daily activities. People with high levels of helplessness tend to report high pain and greater depression. A cross-sectional study of 92 patients with RA, revealed that perceived helplessness mediated the relationship between pain and depression in RA (Smith, Peck, & Ward, 1990). This finding has been replicated in a sample of people with OA (Wallston, 1993) and in a multiwave study of patients with RA (Smith & Wallston, 1992). Similarly, longitudinal analyses have demonstrated that both helplessness and pain are significant mediators between disease activity (indicated by joint count) and future physical and psychological disability (Schoenfeld-Smith et al., 1996). Affleck, Tennen, Pfeiffer, and Fifield (1987) found that positive mood and better psychological adjustment were associated with perceived control over symptoms in a cross-sectional analysis based on a combination of interview and questionnaire data.

There is consensus that passive, avoidant, emotion-focused coping strategies are related to depression, negative affect, and poor self-esteem, whereas active coping strategies are related to positive affect and decreased depression (Young, 1992). Wishful thinking has been associated with poorer psychological well-being, whereas cognitive restructuring and information seeking are associated with better psychological well-being (Felton & Revenson, 1984; Manne & Zautra, 1989). Furthermore, passive coping strategies have been identified as mediators between pain and subsequent depression (Brown, Wallston, & Nicassio, 1989). In three independent samples of RA patients, Zautra et al. (1995) confirmed that coping mediates the relationship between disease variables and positive and negative affect. Using cluster analysis, Newman, Fitzpatrick, Lamb, and Shipley (1990) classified 158 RA patients on the basis of their coping strategies. Interestingly, the largest group ($n = 105$) did not demonstrate a distinctive coping style, but was found to utilize a range of strategies to limited extents. Although the four groups identified did not differ on demographic, clinical, or laboratory measures, one group ($n = 14$), characterized by lower levels of pain, stiffness, and disability and greater psychological well-being, tended to use more open and active strategies to cope with the disease.

Depression may be a key factor influencing perceptions of disability among people with arthritis. Psychological factors have been shown to be better predictors of disability compared with traditional, medical indicators of disease activity in RA (Katz & Yelin, 1993; McFarlane & Brooks, 1988; Newman, Fitzpatrick, Lamb, & Shipley, 1989; Wright et al., 1996). In a study of knee OA examining the roles of sociodemographic factors, obesity, radiographic evidence, and pain on self-reports of disability (Jordan et al., 1996), results revealed that severity of pain was more important than radiographic evidence in determining disability.

Taken together, these findings are consistent with anecdotal evidence from health care professionals suggesting that the degree of psychological distress is not always directly related to disease severity or to objective indicators of disability. It is important to note that disability in arthritis is often measured by self-report questionnaires, both for research purposes and screening in clinical settings. Respondents are typically asked to rate their ability to perform a range of daily activities (e.g., dressing, eating). Thus ratings of disability may be influenced by psychological factors such as depression, self-efficacy, or degree of illness acceptance. The role of depression may be one of mediation, whereby higher levels of depression produce a downward bias in perceived ability to perform daily activities, even controlling for objective measures of disease activity. Clearly, longitudinal studies are needed to disentangle the precise nature of causal relationships between pain, depression, and disability.

8.18.5.2 Children with JCA

Investigations of psychosocial well-being among children with JCA have produced inconsistent findings. Compared with a control group, children with arthritis are reported to have a more negative self-image and lower than average achievement scores in school (McAnarney, Pless, Satterwhite, & Friedman, 1974). However, these findings have not been replicated in later studies (Billings, Moos, Miller, & Gottieb, 1987; Ungerer, Horgan, Chaitow, & Champion, 1988). In a review of psychosocial factors, Miller (1993) concludes that in general children with rheumatic diseases do not show evidence of psychological or social dysfunction. Similarly, Revenson (1993) maintains that JCA places some, but not all, children at risk of psychosocial problems. Family studies suggest that family cohesion and harmony, few stressful events, and better parental health status are associated with fewer adjustment or behavioral problems among children (Harris, Newcomb, & Gewanter, 1991; Ross et al., 1993; Vandik & Eckblad, 1991; Varni, Wicox, & Hanson, 1988; Wallander & Varni, 1989). The nature of arthritis (pain, fatigue, visible changes to the body) and adherence to complex treatment regimens combine to make children with JCA feel and appear "different" from their peers (Barlow, Shaw, & Harrison, 1997). In addition, many children meet substantial barriers when attempting to participate in school and leisure activities. The influence of these factors on psychological adjustment has rarely been examined.

8.18.5.3 Social Impact: Family

The impact on family dynamics has rarely been addressed, despite the relatively high levels of physical disability associated with arthritis and the nature of that disability, which is likely to interfere with performance of instrumental and nurturant activities. Over 50% of people with RA and OA experience loss in eight life domains including household chores, leisure activities, work, and social relationships (Yelin, Lubeck, Holman, & Epstein, 1987). In large samples of people with RA, the most affected areas of family role functioning were found to be shopping, cleaning, and maintaining family ties (Reisine, Goodenov, & Grady, 1987). The importance of instrumental tasks, such as cooking, has been highlighted by Katz and Yelin (1994), who found that inability to perform valued family activities was related to depression in women with RA. A survey of 240 people with AS has revealed that approximately one-third experience difficulties in parenting (Barlow & Cullen, 1996). The main problem areas centered on breast-feeding, lifting, holding, bathing, and parents' inability to participate in "rough and tumble" play and sporting activities. As expected, a gender difference emerged with women reporting more difficulties with babies and toddlers, while men tended to experience more problems in caring for older children over the age of five. Fatigue, pain, restricted mobility, and negative emotional reactions were the main factors detracting from perceptions of successful parenting. Consequently, many people felt deprived of the joys and pleasures of simply being a parent. Feelings of frustration, impatience, anger, irritability, and depression were typically expressed: "I became irritable due to discomfort and stiffness"; "Being unable to join in physical activities left me feeling very frustrated"; and "I let my pain control my anger instead of understanding youthful behavior."

The impact of arthritis permeates through other family members with variable effects: some families find they are closer, others report no difference and some report negative effects on family life (Pritchard, 1989). In RA, married people perceived better access to help with instrumental tasks, more physical affection, and expressed feelings of being needed by others, compared with those who were not married (Manne & Zautra, 1989). A comparison of income, social support, and psychological well-being among women with RA revealed that widowed or divorced women had a lower income, less potential support, and higher levels of depression and anxiety than women who had never married or those living with a spouse

(Kraaimaat, van Dam-Baggen, & Bijlsma, 1995). Interestingly, pain contributed to depression and anxiety only in women living with a spouse. The effects of social support may not be positive if the nature of the support offered fails to match the needs of the person with arthritis. A close relationship may be a double-edged sword in that positive and negative aspects of such relationships may both assist and detract from coping and adjustment. Nevertheless, social support from family and friends has been associated with less depression (Revenson & Majerovitz, 1991), more life satisfaction (Burckhardt, 1985) and greater self-esteem (Fitzpatrick, Neuman, Lamb, & Shipley, 1988). Revenson (1993) concludes that the relationship between social support and psychological adjustment is robust and consistent across studies.

Both the direct-effects and stress-buffering model have been examined in RA, in attempts to determine the mechanism through which social support influences psychological adjustment. Brown, Wallston, and Nicassio (1989) found evidence of both direct and stress-buffering effects on depression in cross-sectional analyses, but only a direct effect longitudinally. Fitzpatrick, Newman, Lamb, and Shipley (1988) revealed that social support was associated with greater self-esteem and lower depression, regardless of level of disability, thus providing support for the direct effects model. Similarly, a study of newly diagnosed RA patients revealed that greater support was associated with less depression, regardless of physician-rated disease severity (Revenson & Majerovitz, 1991).

Arthritis can impact on the wellbeing of other family members. For example, the presence of RA can evoke negative consequences among healthy spouses (Reisine, 1993) and negative marital interactions surrounding the wife's illness can determine psychological adjustment of both husband and wife (Manne & Zautra, 1990). The effect on children of having a parent with arthritis has received little attention. However, there are indications that adolescents of parents with RA may have difficulty in maintaining self-esteem (Hirsch, Moos, & Reischl, 1985). Similarly, having a child with arthritis is likely to place an additional strain on the burden of care among parents and has been found to have negative consequences (Timko, Stovel, & Moos, 1992). One cause of psychological distress may be parents' perceived need to monitor their child's condition constantly for symptoms and side effects, combined with fear for their child's future health, career, and happiness (Barlow, Shaw, & Harrison, 1998). The result of prolonged emotional stress on parents has not been examined systematically and remains to be fully elucidated. Similarly, the impact of a child's arthritis on siblings and grandparents, and vice versa, has received little attention but may be an important avenue of further inquiry in attempts to understand the factors that promote family coping ability.

8.18.5.4 Social Impact: Society

There is a tendency for health-related research to be framed within the medical model of disability, whereby "the problem" is considered to be a deficit in adjustment, motivation, or ability of the person with arthritis. This model construes professionals as the experts, who intervene with "solutions" for those lacking in knowledge and expertise. In contrast, the social model of disability suggests that problems derive from the attitudinal, environmental, and organizational barriers faced by people with arthritis (Barlow & Harrison, 1996).

Arthritis is generally viewed as a condition associated with old age. This attitude derives from a lack of understanding in society, and is a constant source of frustration for young people and children who are told that they "cannot have arthritis" because they are "too young." The effect of such attitudes serves to deny their personal experience and identity.

The bodily impact of arthritis can be visible or invisible, as well as permanent or temporary. Social expectations can vary according to the visibility or invisibility of disabilities (Stone, 1995). Hence, it is difficult to explain how one can be relatively free from pain one day, but in agony and unable to move the next. People with invisible or fluctuating arthritis can be accused of being lazy, malingering, or using their condition as an excuse for failing to meet the expectations of others in their social network. When arthritis is visible there is an overwhelming propensity for people to see the arthritis rather than the person, resulting in feelings of being different, marginalization, and exclusion.

Young people and children often miss long periods of schooling during times of an arthritis "flare." The choice between segregation, if steered down the special needs route, or isolation, if left trying to cope in unsupportive and disabling institutions, is a false dichotomy faced by many young people. Institutional regulations frequently dictate inflexible working or school hours causing difficulties for those who experience fatigue as an integral part of their condition. Fatigue can render the individual unable to muster the intellectual or physical capacity necessary to negotiate a day in education or work.

Environmental barriers frequently derive from the inappropriate design of physical infrastructures that effectively limit not only access to buildings and public transport, but also freedom of movement once inside. Buses with high steps, the lack of ramps at the entrance to buildings, and the scarcity of disabled toilets are examples of environmental barriers frequently cited by people with arthritis of all ages (Barlow et al., 1997; Williams, Cullen, & Barlow, 1996). Restricted access to buildings and transport have major implications for leisure pursuits and establishing a social life. Everyday activities, such as shopping or going to the cinema, become significant feats of negotiation, endurance, and even risk taking.

A positive outlook, self-esteem, and self-confidence become difficult to maintain. Feelings of isolation, anger, and being different are common (Barlow & Cullen, 1996; Barlow et al., 1997). Finding a relevant peer group, through voluntary organizations or support groups, can enable people with arthritis to feel part of a community, a group with shared experiences and and values (Barlow, Cullen, Davis, & Williams, 1997). Such feelings of belonging can also be engendered by attending psychological interventions based on group formats, which offer the opportunity for participants to meet with similar others and to discuss common concerns.

8.18.6 ADHERENCE

As in all chronic diseases, responsibility for the daily management of arthritis gradually shifts from the healthcare team to the individual. Management of home-care regimens is often complex, involving adherence to a range of therapies, including medication, prescribed exercise, splint wearing, clinic attendance, and the avoidance of health risk behaviors. For the person with arthritis, the benefits of behavioral strategies (e.g., exercise, use of splints), may not be evident in the short term, but are necessary to prevent further disability. Complex, long-term therapies are notoriously difficult to maintain, thus it is not surprising that adherence in arthritis is often poor and not only detracts from the individual benefits of various therapies, but on a wider scale, can result in increased health costs.

A review of treatment adherence in RA (Bradley, 1989) revealed that rates of cooperation varied from 33% to 78% for medication and from 33% to 66% for therapeutic exercises. Patients' beliefs concerning the efficacy of treatment may be crucial determinants of adherence. The demands placed on individuals are well illustrated in the context of exercise therapy. Exercising regularly for the remainder of one's life is a daunting prospect for even the most motivated of individuals whatever the age of disease onset, but can be particularly problematic in conditions that begin during adolescence or young adulthood. For example, a young man diagnosed with AS at the age of 20 can expect to carry out over 18 000 exercise sessions by the time he is 70. Regular exercise and monitoring of posture have to become an integral part of everyday life. Psychological interventions can offer a useful means of enhancing adherence and as yet have not been fully exploited in the field of arthritis.

8.18.7 PSYCHOLOGICAL INTERVENTIONS

The range of psychological interventions utilized in arthritis is summarized in Table 1; the most common being cognitive behavioral therapy and self-management programs. Interventions can be delivered in individual or group format, in clinic or community settings.

8.18.7.1 Information and Instruction

People with arthritis require the necessary knowledge and understanding to make informed decisions concerning treatment options, to adhere to complex regimens and to perform self-care activities. Hence, provision of information is an integral aspect of clinical encounters and often forms the opening section of many interventions. The successful transfer of information not only represents a useful first step in the process of empowering individuals to perceive themselves as capable of controlling their arthritis, but may also go some way towards satisfying patients' desire for information.

Kay and Punchak (1988) found that only 46% of patients with RA reported that they had received information about their condition from health professionals. Nevertheless, these patients were keen to find out more about their disease, its management, and drug treatment and believed information on these topics would help them to cope better with their everyday lives. A survey of rheumatology outpatients revealed that 55% had sought further information about their condition outside of the clinic setting (Barlow & Pennington, 1996). People wanted to know more about their condition (e.g., the cause, symptoms) and disease management, particularly self-care. Information-seeking behavior was associated with a diagnosis of RA, being female, and holding educational qualifications.

Table 1 Summary of psychological interventions.

Type of intervention	*Individual or group format*	*Location*
Instruction materials: leaflets, books, videos, computer multimedia programs	Individual	Home
Cognitive-behavioral therapy	Individual or group	Hospital or home
Self-management programs	Group	Community setting

Leaflets offer a means of bridging the gap between the amount of information patients desire and the amount they receive, and can be assimilated in a clinical setting. In addition, leaflets represent a relatively low-level, low-cost method of reaching large numbers of people, although the gatekeeping role adopted by some health professionals serves to limit their potential (Barlow, Bishop, & Pennington, 1996). For example, a survey found that only 41.8% of patients attending rheumatology outpatient clinics recalled receiving an information leaflet (Barlow & Pennington, 1996). Distribution of leaflets was controlled by health professionals, many of whom felt that allowing patients free access may cause more harm than good, despite documented evidence to the contrary (Ong, De Haes, Hoos, & Lammes, 1995; Weinman, 1990).

Leaflets are an effective means of improving knowledge among OA patients (Moll, 1986) and patients with mixed diagnoses of arthritis (Maggs, Jubb, & Kemm, 1996). Extending this work, a randomized, controlled study found that leaflets were effective in terms of increasing knowledge, decreasing pain and improving depressed mood among RA patients at three weeks' postintervention (Barlow, Pennington, & Bishop, 1997). Furthermore, there was no evidence that patients found the leaflets worrying. On the contrary, leaflets appeared to have a positive effect on patients' feelings and beliefs, generating a sense of reassurance, reducing isolation, and helping them to come to terms with their condition. A follow-up study has shown that the positive changes on knowledge and depression were maintained at six months. However, it is not clear whether maintenance of such improvements was influenced by participation in a telephone interview following the three week assessment.

The optimal time for presenting patients with disease-related information (i.e., at diagnosis, after several years, or at repeated intervals throughout the disease course) has yet to be identified, although there are indications that patients in the early stages of their disease can feel that they are given more information than they can cope with (Lineker, Hughes, & Badley, 1995). Patients' interest in learning more about RA does not decrease over time (Silvers, Hovel, Weisman, & Mueller, 1985) and increase in knowledge following receipt of written materials appears to be independent of disease duration (Barlow et al., 1997). Thus the common practice of restricting leaflet distribution to the point of diagnosis may be limiting the potential for positive change to those in the early stages of their disease. There is an unwritten assumption that the written word is a universally acceptable medium for communicating health-related information. The effectiveness of alternative educational media, such as audiotapes, videos, the Internet, or computer multimedia programs, remains largely unexplored in the context of arthritis.

8.18.7.2 Cognitive Behavioral Interventions

Cognitive behavioral therapy (CBT) for people with arthritis focuses on pain management and comprises three components: a rationale for treatment, training in coping skills, and training for maintenance of coping skills and setbacks (Keefe & Caldwell, 1996). A range of coping skills are introduced, including relaxation, imagery, activity–rest cycling, and cognitive restructuring. Interventions are usually conducted in a group format (i.e., 5–10 people) by clinical psychologists or other suitably qualified health professionals trained in CBT techniques. Sessions typically last for two hours and are delivered on a weekly basis for 6–10 weeks.

Cognitive behavioral interventions have been shown to be effective for people with RA. Compared with patients in a social support control group, patients with RA receiving CBT reported significant reductions in pain, pain behavior, anxiety, and depression (Bradley, 1996). Other controlled studies have revealed that CBT is effective in improving coping and function (Applebaum, Blanchar, Hickling, & Alfonso, 1988; Parker et al., 1988). Moreover, such improvements appear to be maintained in the longer term (Bradley et al., 1988), particularly among patients who continue to practice coping skills on a regular basis (Applebaum et al., 1988). CBT has been shown to enhance perceived self-efficacy among patients with RA (O'Leary, Shoor, Lorig, & Holman, 1988), the degree of self-efficacy enhancement correlating with the magnitude of improvement on pain, joint inflammation, and psychosocial functioning (e.g., depression). However, generalizability of findings is limited by the relatively small sample (i.e., 15 in the control and intervention groups respectively). Family support has been combined with CBT for people with RA (Radojevic, Nicasio, & Weisman, 1992). Four conditions were contrasted to assess potential benefits: CBT with family support, CBT without family support, family support only, and a no treatment group. The CBT interventions were found to provide more benefits compared with the other two conditions.

Few studies have examined the effects of CBT among people with other forms of arthritis. However, there is evidence that CBT is more effective than arthritis education and standard care in reducing pain and psychological distress among people with OA (Keefe et al., 1990). One application of CBT among people with AS revealed improvements in depression and self-efficacy, but failed to show evidence of reduced pain or decreased disease severity (Basler, 1993). Overall, CBT appears to offer benefits to people with arthritis, particularly in relation to improvements in pain, and psychological well-being. CBT is often used in combination with education and self-management skills training.

8.18.7.3 Self-management Interventions

Recognition of the need for self-management ability among people with arthritis has led to the development of interventions designed to enhance relevant skills and coping strategies. One such intervention, the Arthritis Self-Management Program (ASMP), was developed for people in North America with mild to moderate arthritis (Lorig & Gonzalez, 1992). The ASMP comprises six weekly sessions of approximately two hours in length delivered in community settings by pairs of lay leaders, most of whom have arthritis. Topics covered include information on the disease process, exercise, cognitive pain management, depression, communication with medical professionals, problem solving, and goal accomplishment. The program draws on the basic tenets of self-efficacy theory (Bandura, 1982) and aims to enhance perceived ability to control various aspects of arthritis through skills mastery, modeling, reinterpretation of symptoms, and persuasion. The ASMP encourages participants to become active agents in the care of their arthritis, able to select the self-management strategies that best suit their individual needs at a given time.

Most studies of self-efficacy in people with arthritis have followed Bandura's (1977, 1982) definition of behavior-specific beliefs tailored to fit narrowly defined situations. Thus self-efficacy is measured in terms of confidence in one's ability to control various aspects of arthritis, such as pain, fatigue, and physical functioning, on a daily basis (Lorig, Chastain, Ung, Shoor, & Holman, 1989). Given equal disease severity, perception of arthritis self-efficacy may differentiate between those who are incapacitated by their disease and those who continue to live full and active lives. Randomized, controlled studies have shown that at four month follow-up, ASMP participants report a reduction in pain, a reduction in depression, and a decreased number of visits to physicians (Lorig & Holman, 1993). Moreover, a longer-term follow-up after four years showed that despite a slight increase in disability, reductions in pain and visits to physicians remained evident and there was a marked increase in self-efficacy (i.e., 17%). Levels of depression remained fairly stable.

The ASMP has been successfully delivered and evaluated in controlled trials in Australia (Lindroth, Bauman, Brooks, & Priestly, 1995; Simeoni, Bauman, Stenmarh, & O'Brien, 1995), and also in Scandinavia and Canada, indicating that well-designed psychological interventions can transcend cultural differences. The first evaluation of the ASMP in a UK context (Barlow, Williams, & Wright, 1997) focused on older people with arthritis (i.e., over 55 years of age) using a pre–post-test design. Results showed that after four months, participants demonstrated significant increases in arthritis self-efficacy, positive affect, cognitive symptom management, communication with doctors, exercise, and relaxation. In addition, significant decreases were found in terms of pain, depression, and visits to primary care practitioners. This study extended previous work by including assessment of generalized self-efficacy beliefs in addition to arthritis-specific beliefs. Although generalized trait-like beliefs are relatively stable over time compared with state-like beliefs (Scheier & Carver, 1992), a small improvement on generalized self-efficacy was identified, suggesting that techniques, such as problem solving, may not be confined to managing arthritis. It seems likely that the benefits of attending an ASMP are of a wider-ranging nature than first anticipated and may influence perceived ability to cope with the environmental demands faced in everyday life as an older person with arthritis. Further randomized controlled studies are in progress to determine whether the ASMP holds similar benefits for people of a wider age range in the UK.

The effectiveness of a different form of group patient education, Self-Management Courses for People with Ankylosing Spondylitis (SMC-AS), was examined in a pre–post-test design with an intervention group and a matched control group receiving standard care (Barlow & Barefoot, 1996). This intervention differs from the ASMP, by its greater emphasis on the promotion of home exercises and self-monitoring, essential aspects of disease management in AS. The SMC-AS was found to benefit participants in terms of enhanced self-efficacy, reduced disease severity, and improved psychological well-being after six months. Although significant increases in the range and frequency of exercise were evident at three weeks, these changes were not maintained over time. A positive association between severity and exercise at baseline suggests that people with greater perceived severity are more likely to adhere to therapeutic exercise. Hence, as group participants began to "feel better" in terms of pain, stiffness, and psychological well-being, motivation to carry out therapeutic exercise may have declined. This intriguing finding warrants further, more detailed, exploration and has implications for the delivery of future interventions.

Failure to maintain change over time through self-management interventions has been reported in controlled studies in The Netherlands (Taal et al., 1993) and Australia (Lindroth et al., 1995), where the interventions were delivered by health professionals. In contrast, long-term improvements have been reported following the ASMP delivered by lay leaders in the USA (Lorig & Holman, 1993; summarized in Table 2). It should be noted that the majority of participants in the studies reviewed by Lorig and Holman were women with OA with a mean age of 65 years. Thus it is not clear whether long-term benefits derive from the characteristics of ASMP participants (e.g., OA, age, gender), cultural differences between North Americans and Europeans, or variation in the content and context of delivery of interventions (e.g., community or hospital setting).

8.18.7.4 Interventions in JCA

Evidence from the few published studies evaluating psychological interventions for children with JCA suggests this is a promising avenue worthy of greater attention in a field where treatment is ameliorative rather than curative, treatment adherence is poor, and long-term adaptation is essential. Interventions for children include CBT, behavioral techniques, summer camps, and family retreats. In the context of pain management, the effectiveness of CBT delivered in a series of individual sessions, was examined among 26 children (Walco, Varni, & Ilowite, 1992). Of the 13 children who completed the program, pain ratings were improved in the short term and at six month follow-up. Using similar methods but with the addition of thermal biofeedback and autogenic exercises (e.g., positive self-statements and relaxing imagery), Lavigne et al. (1992) found improved pain ratings among four of the eight child participants postintervention and at six month follow-up. These findings were confirmed by maternal reports of children's pain and pain behavior. While indicating the potential of CBT for influencing pain management in children, generalizability is limited by the small samples, high attrition rates, and reliance on parametric statistical procedures in small data sets.

The efficacy of a parent-managed behavioral (token) reinforcement program designed to improve the medication compliance of a 14-year-old boy with juvenile rheumatoid arthrits was examined in a withdrawal (single subject) design (Rapoff, Purviance, & Lindsley, 1988a). Results showed increased compliance, clinical improvements (assessed by a rheumatologist), and improvement on parental ratings of symptoms. In an attempt to reduce the complexity and time-consuming nature of token reinforcement programs, Rapoff, Purniance, and Lindsley (1988b) used more simple behavioral strategies (e.g., self-monitoring and positive verbal feedback) combined with education among three female patients, aged 3, 10, and 13 years. At baseline, medication compliance increased for two patients but showed no change for the third patient. Moreover, for all three patients, compliance had decreased at four month follow-up, suggesting that long-term maintenance may require either continuous intervention or regular monitoring. It should be noted that the high level of compliance achieved by the youngest child may reflect the parents' attitudes to medication more than the beliefs and behaviors of the child. A systematic replication of these techniques with three female patients, two with systemic lupus erythematosus and one with dermatomyositis, showed improved compliance during the intervention and at six month follow-up (Pieper, Rapoff, Purviance, & Lindsley, 1989).

The impact of a very different intervention, a summer camp designed to foster independence and to promote disease self-management, was examined among children and young people aged between 7 and 20 in an uncontrolled,

Table 2 Summary of long-term outcomes of controlled, self-management interventions.

	Taal et al. (1993)	*Lorig and Holman (1993)*	*Lindroth et al. (1995)*
Context	Hospital-based	Community-based	Hospital-based
Tutors	Health professionals	Lay people with arthritis	Health professionals
Follow-up time	14 months	4 years	5 years
Outcomes			
Pain	NS	+	+
Disability	NS	NS	NS
Depression	NS	+	NA
Visits to physicians	NA	+	NA
Self-efficacy: pain	NS	+	NA
Self-efficacy symptoms: other	NS	+	NA
Internal locus of control	NA	NA	NS

+ = positive change in intervention group, NS = no significant difference between groups, NA = not assessed.

pre–post test design (Stefl, Shear, & Levinson, 1989). Given these aims it is surprising to note that no measures of self-management were included in the assessment battery. Results showed that children's locus of control and self-esteem had significantly improved immediately following the one week camp, but these changes were not maintained at six months. Acknowledging the fact that management of JCA occurs within a family context, Hagglund et al. (1996) examined the efficacy of a three day family retreat held at a "resort" hotel among 39 families. The retreat included formal educational and therapeutic sessions (e.g., coping with chronic illness, enhancing self-esteem) and family-oriented recreational activities. Caregivers reported that their children's internalizing behavior problems were significantly reduced six months after the retreat and mothers reported reductions in strain on family leisure and occupational functioning at six month follow-up. Mothers' psychological distress remained stable. A number of important methodological limitations were noted, such as the lack of a control group and the difficulty in controling for intervening variables between initial and follow-up assessments.

There are indications that psychological interventions may offer benefits to children and their families. However, a number of methodological issues remain to be addressed, such as the small sample sizes, lack of controlled studies, dearth of validated measuring instruments for children, and scarcity of resources available for nonmedical interventions. Furthermore, it has been argued that pain is not always viewed as a prominent issue in JCA (Lavigne et al., 1992). Until such time as psychological well-being, including the subjective experience of pain, is accepted as of equal importance to physical well-being, psychological interventions for children with JCA will continue to lag behind those for adults. Descriptive studies employing qualitative methodologies may allow greater insight into the experience of JCA from a child's perspective and would provide a useful framework for developing successful interventions in the future.

8.18.7.5 Effectiveness of Psychological Interventions

The main benefits of psychological interventions appear to center on promoting positive mood, reducing depression, and fostering confidence in self-management (i.e. self-efficacy). For example, Basler (1993), showed that the main effects of a cognitive-behavioral program were on well-being (e.g., depression) and self-efficacy among people with AS. Similarly, evaluations of the ASMP report positive effects on depression and self-efficacy (Lorig & Holman, 1993). Hence, although levels of disability and pain may remain fairly stable, participants feel better able to cope.

The effectiveness of psychological interventions is usually assessed over and above the effects of medical interventions. Contrary to clinical trials, participants are not required to cease their standard treatment for a wash-out period prior to entering the intervention. Wash-out periods serve to increase the probability of finding a significant effect in clinical trials, while maintenance of medication during an education intervention limits the potential for improvement. A review of arthritis patient education studies (including psychologically based interventions) (Hirano, Laurent, & Lorig, 1994) suggests that clinical studies of the effects of medication alone demonstrate a 20–50% improvement on health status, whereas educational interventions provide an additional 15–30% improvement. A meta-analysis comparing the effects of educational interventions and NSAIDs on pain and functional disability in OA and RA concluded that since most patients in education trials are being treated with medications, the sizes of the effects of these trials represent additional, marginal effects of patient education beyond those achieved by medication alone (Superio-Cabuslay, Ward, & Lorig, 1996). The weighted average effect sizes for pain, disability, and tender joint count in educational interventions were 0.17, 0.03, and 0.34, respectively, compared with 0.66, 0.34, and 0.43 for NSAID trials. Interestingly, the authors found evidence of a publication bias: smaller NSAID trials reporting that little effect or negative effects were not represented in the meta-analysis. Hence, the averaged effect sizes for NSAID trials may be inflated.

8.18.8 PROFESSIONAL ISSUES

Few rheumatology units have a resident psychologist dedicated to the care of people with arthritis. Rather, psychological input tends to revolve around provision of remedial help for acute problems or "problem patients." Parker and Wright (1995) argue for an extension of the traditional rheumatological approach to embrace genuine rehabilitation strategies, designed to maximize psychological functioning. Indeed, the interactions between symptoms of arthritis, psychosocial well-being and disability suggest a broader, more holistic approach to healthcare may promote positive outcomes. For example, depression in RA may be alleviated by treatment interventions designed to reduce life stress, enhance self-efficacy, and minimize physical disability (Wright et al., 1996). As knowledge of the psychosocial impact and the effectiveness of psychological interventions continues to grow, the demand for suitably qualified psychologists will escalate accordingly. Availability of adequate resources may be the crucial factor determining whether psychologists are able to take their place in the search to facilitate positive adaptation to the challenge of arthritis.

8.18.9 CONCLUSIONS AND FUTURE ISSUES

Clearly, psychology and psychologists are key players in the theater of life with arthritis. The chronic, painful, disabling nature of the condition combined with an unpredictable disease course of exacerbations and remissions, necessitates a continuous process of adaptation. In the long term, the majority of people appear to adapt with some degree of success, although many experience feelings of anxiety, depression, loss of control, and helplessness along the way. There is evidence that psychological factors are more important predictors of perceived disability than disease-related variables, such as duration. Further empirical applications of psychological theories are needed to elucidate the complex relationships between disease parameters, behaviors, and emotional well-being, including coping ability.

Knowledge of psychological factors in RA is accumulating. Most studies of psychological impact have been based on convenience samples of RA patients recruited through rheumatology clinics, who may be those with more severe disease or those experiencing an acute phase (i.e., a flare up). Thus little is known about prevalence or correlates of depression among community-based samples or people with milder arthritis. We need to know more about the psychological well-being of people with other forms of arthritis, such a OA, JCA, or scleroderma. Further knowledge of the psychological profile of people with OA may be particularly important given the aging demographic profile of the population. Similarly, more information on psychological adjustment among ethnic minorities, carers, and other family members is needed. The challenge for psychologists is to discover methods of identifying those likely to experience difficulty, regardless of their arthritis type, and to determine the most effective intervention strategies to aid psychological well-being.

The role played by positive well-being has not been extensively examined. Factors promoting positive well-being among people with arthritis may not necessarily be the same as those that serve to reduce negative reactions. The role of humor and positive experiences require further exploration and may help to illuminate ways in which psychologists can help people with arthritis to move forward in attempts to resolve the challenge of arthritis.

Psychological assessment and research would be enhanced by the availability of a greater range of reliable measuring instruments, validated for use among children, adolescents, and people with forms of arthritis other than RA. Specific instruments, used in combination with generic psychological measures, would enable comparisons to be made with the general population, acute illness, and other chronic disease groups. Thus, psychologists working in the field of arthritis could apply techniques found to enhance well-being in other health-related arenas.

There is a growing body of evidence demonstrating that psychological interventions are effective in helping people with arthritis to manage and cope better with their condition. Clearly, there is a great need for more psychological interventions, for all people with arthritis, regardless of the degree of severity and regardless of age. Interventions for children and adolescents are rare but appear to have the potential for promoting well-being. Group interventions hold a number of benefits for participants, including the opportunity to share their experiences with similar others in a nonthreatening environment. The potential for information technology to be harnessed as a means of providing psychological support to greater numbers of people remains to be explored. Psychologists and psychological theories have an important contribution to make to the development, delivery, and evaluation of interventions. There is an overwhelming need to persuade the purchasers and providers of services that psychologists have a fundamental role to play in empowering people with arthritis to take control of their lives.

ACKNOWLEDGMENT

I would like to thank Trudy Brockway for her valuable assistance in preparing this manuscript.

8.18.10 REFERENCES

Affleck, G., Tennen, H., Pfeiffer, C., & Fifield, J. (1987). Appraisals of control and predictability in adapting to a chronic disease. *Journal of Personality and Social Psychology, 53*(2), 273–279.

Anderson, K. O., Bradley, L. A., Young, L. D., McDanial, L. K., & Wise, C. M. (1985). Rheumatoid arthritis: Review of psychological factors related to etiology, effects and treatment. *Psychological Bulletin*, *98*, 358–387.

Applebaum, K. A., Blanchar, E. B., Hickling, E. J., & Alfonso, M. (1988). Cognitive behavioral treatment of a veteran population with moderate to severe rheumatoid arthritis. *Behavior Therapy, 19*, 489–502.

Badley, E. M., & Tennant, A. (1993). Disablement associated with rheumatic disorders in a British population: Problems with activities of daily living and level of support. *British Journal of Rheumatology*, *32*, 601–608.

Bandura, A. (1977). Self efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122–147.

Barlow, J. H. & Barefoot, J. (1996). Group education for people with arthritis. *Patient Education and Counseling, 27*, 257–267.

Barlow, J. H., Bishop, P. & Pennington, D. C. (1996). How are printed patient educational materials used in outpatient clinics? Insight from rheumatology. *Health Education Journal, 55*, 275–284.

Barlow, J. H., & Cullen, L. (1996). Parenting and ankylosing spondylitis: "I can't see where my baby is feeding'. *Disability, Pregnancy and Parenthood International, 16*, 4–5.

Barlow, J. H., Cullen, L., Davis, S., & Williams, R. B. (1997). The power of sharing common experiences: The "hidden" benefit of group education for people with arthritis. *British Journal of Therapy and Rehabilitation, 4*(1), 38–41.

Barlow, J. H., & Harrison, K. (1996). Focusing on empowerment: Facilitating self-help in young people with arthritis through a disability organisation. *Disability and Society, 11*(4), 539–551.

Barlow, J. H., Macey, S. J., & Struthers, G. R. (1993). Gender, depression and ankylosing spondylitis. *Arthritis Care and Research*, *6*(1), 45–51.

Barlow, J. H., & Pennington, D. C. (1996). Evaluation of patient education materials. In *The Arthritis & Rheumatism Council for Research, Scientific Reports 1995* (p. 389). Chesterfield, UK: Arthritis & Rheumatism Council.

Barlow, J. H., Pennington, D. C., & Bishop, P. (1997). Patient education leaflets for people with rheumatoid arthritis: A controlled study. *Psychology, Health and Medicine: Special Issue, Communications in Health Care, 2.*

Barlow, J. H., Shaw, K., & Harrison, K. (1998). The psychosocial impact of juvenile chronic arthritis. *British Journal of Rheumatology, 36,* 4.

Barlow, J. H., Williams, R. B., & Wright, C. (1997). Evaluation of arthritis self-management in older adults: "Just what the doctor didn't order". *British Journal of Health Psychology*, *2*, 1–13.

Basler, H-D. (1993). Group treatment for pain and discomfort. Patient education & counseling. *Psychosocial Aspects of Rheumatic Diseases, 20*(2 & 3), 167–175.

Benjamin, C. M. (1990). Review of UK data on the rheumatic diseases—1, Juvenile chronic arthritis. *British Journal of Rheumatology*, *29*, 231–233.

Billings, A. G., Moos, R. H., Miller, J. J., & Gottlieb, J. E. (1987). Psychosocial adaptation in juvenile rheumatic disease: A controlled evaluation. *Health Psychology, 6*, 343–359.

Blalock, S. J., DeVellis, R. F., Brown, G. K., & Wallston, K. A. (1989). Validity of the Centre for Epidemiological Studies Depression Scale in arthritis populations. *Arthritis and Rheumatism, 32*, 991–997.

Booth, G. C. (1937). Personality and chronic arthritis. *Journal of Nervous Mental Disorders*, *85*, 637–662.

Bradley, L. A. (1989). Adherence with treatment regimens among adult rheumatoid arthritis patients: Current status and future directions. *Arthritis Care and Research*, *2*, 533–539.

Bradley, L. A. (1996). Pain management. In S. T. Wegener, B. L. Bleza, & E. P. Gall, (Eds.), *Clinical care in the rheumatic diseases*, (pp. 111–115). Atlanta, GA: ACR.

Bradley, L. A., Young, L. D., Anderson, K. O., Turner, R. A., Agudelo, C. A., McDaniel, L. K., & Semple, E. (1988). Effects of cognitive behavioral therapy on rheumatoid arthritis pain behavior: one year follow-up. In R. Dubiner, G. Gebhardt, & M. Bond (Eds.), *Pain research and clinical management* (Vol. 34, pp. 310–314). Proceedings of the fifth World Congress on Pain. Amsterdam: Elsevier.

Brewerton, D., Caffrey, M., & Hart, D. (1973). Ankylosing spondylitis and HLA-B27. *The Lancet*, *1*, 902.

Brown, G. K. (1990). A causal analysis of chronic pain and depression. *Journal of Abnormal Psychology, 99*(2), 127–137.

Brown, G. K., Wallston, K. A., & Nicassio, P. M. (1989). Social support and depression in rheumatoid arthritis: A one-year prospective study. *Journal of Applied Social Psychology, 19*, 1164–1181.

Burckhardt, C. S. (1985). The impact of arthritis on quality of life. *Nursing Research, 34*, 11–16.

Callahan, L. F., Bloch, D. A., & Pincus, T. (1992). Identification of work disability in rheumatoid arthritis:

Physical, radiographic and laboratory variables do not add explanatory power to demographic and functional variables. *Journal of Clinical Epidemiology, 45*(2), 127–138.

Creed, F., & Ash, G. (1992). Depression in rheumatoid arthritis: Aetiology and treatment. *International Review of Psychiatry*, *4*, 23–33.

Dekker-Saeys A. J. (1976) *Spondylitis Ankylopetica Syndrome*. Dissertation. Amsterdam: Academic Press.

Felton, B. J., & Revenson, T. A. (1984). Coping with chronic illness: A study of illness controllability and the influence of coping strategies on psychological adjustment. *Journal of Consulting and Clinical Psychology, 52*(3), 343–353.

Fitzpatrick, R., Newman, S., Lamb, R., & Shipley, M. (1988). Social relationships and psychological well-being in rheumatoid arthritis. *Social Science and Medicine, 27*, 399–403.

Fitzpatrick, R., Newman, S., Archer, R., & Shipley, M. (1991). Social support, disability and depression: A longitudinal study. *Social Science & Medicine, 33*, 605–611.

Hagglund, K. J., Doyle, N. M., Clay, D. L., Frank, R. G., Johnson, J. C., & Pressly, T. A. (1996). A family retreat as a comprehensive intervention for children with arthritis and their families. *Arthritis Care and Research*, *9*(1), 35–41.

Harris, E. (1993). Clinical features of rheumatoid arthritis. In W. N. Kelley, E. D. Harris, S. Ruddy, & C. B. Sledge (Eds.), *Textbook of rheumatology* (pp. 122–154). Philadelphia: Saunders.

Harris, J. A., Newcomb, A. F., & Gewanter, H. L. (1991). Psychosocial effects of JRD: The family and peer systems as a context for coping. *Arthritis Care and Research, 4*, 123–130.

Hawley, D. J., & Wolf, F. (1993). Depression is not more common in rheumatoid arthritis: 10 Year longitudinal study of 6,153 patients with rheumatic diseases. *Journal of Rheumatology, 20*, 2025–2031.

Hazes, J. M., & Silman, A. J. (1990). Review of UK data on the rheumatic diseases—2, Rheumatoid arthritis. *British Journal of Rheumatology*, *29*, 161–165.

Helmick, C. G., Lawrence, R. C., Pollard, R. A., Lloyd, E., & Heyse, S. P. (1995). Arthritis and other rheumatic conditions: Who is affected now, who will be affected later? *Arthritis Care and Research*, *8*(4), 203–211

Hirano, P. C., Laurent, D. D., & Lorig, K. (1994). Arthritis patient education studies, 1987–1991: A review of the literature. *Patient Education and Counseling*, *24*, 9–54.

Hirsch, B. J., Moos, R. H., & Reischl, T. M. (1985). Psychosocial adjustment of adolescent children of a depressed, arthritic or normal parent. *Journal of Abnormal Psychology, 94*, 154–164.

Hutton, C. W. (1995). Osteoarthritis—clinical features & management. In R. C. Butler & M. I. V. Jayson (Eds.), *Collected reports on the rheumatic diseases* (pp. 35–38). Chesterfield, UK: Arthritis & Rheumatism Council.

Jones, R. L. (1909). *Arthritis deformans*. New York: William Wood.

Jordan, J. M., Luta, G., Renner, J. B., Linder, G. F., Dragomir, A., Hochberg, M. C., & Fryer, J. G. (1996). Self-reported functional status in OA of the knee in a rural southern community; the role of sociodemographic factors, obesity and knee pain, *Arthritis Care and Research, 9*(4), 273–278.

Kahn, A. M., & van der Linden, S. (1990). Ankylosing spondylitis and other spondylarthropathies. *Rheumatic Disease Clinics of North America*, *16*(30), 551–579.

Katz, P. & Yelin, E. (1993). Prevalence and correlates of depressive symptoms among persons with rheumatoid arthritis. *Journal of Rheumatology, 20*, 790–796.

Katz, P. P., & Yelin, E. H. (1994). Life activities of persons with rheumatoid arthritis with and without depressive symptoms. *Arthritis Care and Research, 7*, 69–75.

Kay, E. A., & Punchak, S. S. (1988). Patient understanding of the causes and medical treatment of rheumatoid arthritis. *British Journal of Rheumatology, 27*, 396–398.

Keefe, F. J., & Caldwell, D. S. (1996). Cognitive behavioral interventions. In S. T. Wegener, B. L. Belza, & E. P. Gall (Eds.), *Clinical care in the rheumatic disease* (pp. 59–64). Atlanta, GA: ACR.

Keefe, F. J., Caldwell, D. S., Williams, D. A., Gil, K. M., Mitchell, D., Robertson, C., Martinez, S., Nunley, J., Beckham, J. C., & Helms, M. (1990). Pain coping skills and training in the management of osteoarthritic knee pain: A comparative study. *Behavior Therapy, 21*, 49–62.

Kraaimaat, F. W., van Dam-Baggen, C. M. J., & Bijlsma, J. W. J. (1995). Depression, anxiety and social support in rheumatoid arthritic women without and with a spouse. *Psychology and Health, 10*, 387–396.

Latman, N. S., & Walls, R. (1996). Personality and stress: An exploratory comparison of rheumatoid arthritis and osteoarthritis. *Archives of Physical Medicine and Medical Rehabilitation*, *77*, 796–800.

Lavigne, J. V., Ross, C. K., Berry, S. L., Hayford, J. R., & Pachman, L. M. (1992). Evaluation of a psychological treatment package for treating pain in juvenile rheumatoid arthritis. *Arthritis Care and Research*, *5*, 101–110.

Lawrence, J. S., Bremmer, J. M., & Bier, F. (1966). Prevalence in the population and relationship between symptoms and x-ray changes. *Annals in Rheumatic Diseases*, *25*, 1–24.

Lerman, C. E. (1987). Rheumatoid arthritis: Psychosocial factors in the etiology, course and treatment. *Clinical Psychology Review*, *7*, 413–425.

Lindroth, Y., Bauman, A., Brooks, P. M., & Priestley, D. (1995). A 5-year follow-up of a controlled trial of an arthritis education program. *British Journal of Rheumatology*, *34*, 647–652.

Lineker, S., Hughes, A., & Badley, E. M. (1995). *Educational needs of clients with rheumatoid arthritis: A focus group experience* (working paper 95-1). Canada: Arthritis Community Research & Evaluation Unit ACREU, The Wellesley Hospital Research Institute.

Lorig, K., Chastain, R. L., Ung, E., Shoor, S., & Holman, H. R. (1989). Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. *Arthritis and Rheumatism*, *32*(1), 37–44.

Lorig, K., & Gonzalez, V. (1992). The integration of theory with practice: A 12 year case study. *Health Education Quarterly, 19*(3), 355–368.

Lorig, K., & Holman, H. (1993). Arthritis self-management studies: A twelve year review. *Health Education Quarterly*, *20*(1), 17–28.

Lubeck, D. P. (1995). The economic impact of arthritis. *Arthritis Care and Research, 8*(4), 304–310.

Maggs, F. M., Jubb, R. W., & Kemm, J. R. (1996). Single-blind randomised controlled trial of an educational booklet for patients with chronic arthritis. *British Journal of Rheumatology, 35*(8), 775–777.

Manne, S. L., & Zautra, A. J. (1989). Spouse criticism and support: Their association with coping and psychological adjustment. *Journal of Personality and Social Psychology, 56*, 608–617.

Manne, S. L., & Zautra, A. J. (1990). Couples with chronic illness: Women with rheumatoid arthritis and their healthy husbands. *Journal of Behavioral Medicine, 13*(4), 327–342.

McAnarney, E. R., Pless, B., Satterwhite, B., & Friedman, S. B. (1974). Psychological problems of children with juvenile arthritis. *Pediatrics, 53*, 523–528.

McFarlane, A. C. & Brooks, P. M. (1988). Determinants of disability in rheumatoid arthritis. *British Journal of Rheumatology, 27*, 7–14.

McIntosh, E. (1996). The cost of rheumatoid arthritis. *British Journal of Rheumatology*, *35*(8), 781–790.

Miller III, J. J. (1993). Psychosocial factors related to rheumatic disease in childhood *Journal of Rheumatology, S38*, 1–11.

Moll, J. M. (1986). Doctor patient communication in rheumatology: Studies of visual and verbal perception using educational booklets and other graphic material. *Annals of Rheumatic Diseases, 45*, 198–209.

Moos, R. H. (1964). Personality factors associated with rheumatoid arthritis: A review. *Journal of Chronic Diseases*, *17*, 41–55.

Munthe, E. (Ed.) (1990). The care of the rheumatic child. *EULAR Bulletin*, 47–50.

Newman, S. P., Fitzpatrick, R., Lamb, R., & Shipley, M. (1989). The origins of depressed mood in rheumatoid arthritis. *Journal of Rheumatology, 16*(6), 740–744

Newman, S. P., Fitzpatrick, R., Lamb, R., & Shipley, M. (1990). Patterns of coping in rheumatoid arthritis. *Psychology and Health, 4*, 187–200.

Nicassio, P. M., & Wallston, K. A. (1992). Longitudinal relationships among pain, sleep problems, and depression in rheumatoid arthritis. *Journal of Abnormal Psychology, 101*(3), 514–520.

O'Leary, A., Shoor, S., Lorig, K., & Holman, H. R. (1988). A cognitive-behavioral treatment for rheumatoid arthritis, *Health Psychology, 7*(6), 527–544.

Ong, L. M. L., De Haes, J. C., Hoos, A. M., & Lammes, F. B. (1995). Doctor–patient communication: A review of the literature. *Social Science and Medicine, 4*(7), 903–918.

Parker, J. C., Frank, R. G., Beck, N. C., Finan, M., Walker, S., Hewett, J. E., Broster, C., Smarr, K. L., Smith, E. T., & Kay, D. (1988). Pain management in rheumatoid arthritis patients: A cognitive behavioral approach. *Arthritis & Rheumatism, 31*, 593–601.

Parker, J. C., Smarr, K. L., Angelone, E. O., Mothersead, P. K., Lee, B. S., Walker, S. E., Bridges, A. J., & Caldwell, C. W. (1992). Psychological factors, immunologic activation, and disease activity in rheumatoid arthritis. *Arthritis Care and Research, 5*, 196–201.

Parker, J. C., & Wright, G. E. (1995). The implications of depression for pain and disability in rheumatoid arthritis. *Arthritis Care and Research*, *8*(4), 279–283.

Pieper, K. B., Rapoff, M. A., Purviance, M. R., & Lindsley, C. B. (1989). Improving compliance with prednisone therapy in pediatric patients with rheumatic disease. *Arthritis Care and Research*, *2*(4), 132–135.

Pincus, T., & Callahan, L. F. (1993). What is the natural history of rheumatoid arthritis? *Rheumatic Disease Clinics of North America*, *19*(1), 123–151.

Pincus, T., Griffith, J., Pearce, S., & Isenberg, D. (1996). Prevalence of self-reported depression in patients with rheumatoid arthritis. *British Journal of Rheumatology, 35*(9), 879–883.

Pritchard, M. L. (1989). *Psychological aspects of RA*. New York: Springer-Verlag.

Radojevic, V., Nicassio, P. M., & Weisman, M. H. (1992). Behavioral intervention with and without family support for rheumatoid arthritis. *Behavior Therapy, 23*(1), 13–30.

Rapoff, M. A., Purviance, M. R., Lindsley, C. B. (1988a). Improving medication compliance for juvenile rheumatoid arthritis and its effect on clinical outcome: A single-subject analysis. *Arthritis Care and Research*, *1*, 12–16.

Rapoff, M. A., Purviance, M. R., & Lindsley, C. B. (1988b). Educational and behavioral strategies for improving medication compliance in juvenile rheumatoid arthritis. *Archives of Physical Medicine and Medical Rehabilitation*, *69*, 439–441

Reisine, S. (1993). Marital status and social support in rheumatoid arthritis, *Arthritis and Rheumatism, 36*, 589–592.

Reisine, S. T., Goodenow, C., & Grady, K. E. (1987). The impact of rheumatoid arthritis on the homemaker. *Social Science and Medicine, 25,* 89–95.

Revenson, T. A. (1993). The role of social support in rheumatic diseases *Bailliere's Clinical Rheumatology. International Practice and Research, 7*(2), 377–396.

Revenson, T. A., & Majerovitz, S. D. (1991). The effects of chronic illness on the spouse: Social resources as stress buffers. *Arthritis Care and Research, 4*, 63–72.

Ross, C. K. Lavigne, J. V., Hayford, J. R., Dyer, A. R., & Pachman, L. M. (1993). Psychological factors affecting reported pain in JRA. *Journal of Pediatric Psychology, 18*, 561–573.

Scheier, M. F., & Carver, C. S. (1992). Effects of optimism on psychological and physical well-being: Theoretical overview and empirical update. *Cognitive Therapy and Research*, *16*(2), 201–228.

Schoenfeld-Smith, K., Petrosici, S. F., Hewett, J. E., Johnson, J. C., Wright, G. E., Smarr, K. L., Walker, S. E., & Parker, J. C. (1996). A biopsychosocial model of disability in rheumatoid arthritis. *Arthritis Care and Research, 9*(5), 368–375.

Schumaker, H. R. (Ed.) (1988). *Primer on the rheumatic diseases* (9th ed.). Atlanta, GA: Arthritis Foundation.

Silvers, I. J., Hovel, M. F., Weisman, M. H., & Mueller, M. R. (1985). Assessing physician–patient perceptions in rheumatoid arthritis: A vital component in patient education. *Arthritis and Rheumatism, 28*, 300–307.

Simeoni, E., Bauman, A., Stenmark, J., & O'Brien, J. (1995). Evaluation of a community arthritis program in Australia: Dissemination of a developed program. *Arthritis Care and Research*, *8*(2), 102–107.

Smith, T. W., Peck, J. R., & Ward, J. R. (1990). Helplessness and depression in RA. *Health Psychology, 9*, 377–389.

Smith, C. A. & Wallston, K. A. (1992). Adaptation in patients with chronic rheumatoid arthritis: Application of a general model. *Health Psychology, 11*, 151–162.

Stefl, M. E., Shear, E. S., & Levinson, J. E. (1989). Summer camps for juveniles with rheumatic disease: Do they make a difference? *Arthritis Care and Research*, *2*, 10–15.

Stone, S. D. (1995). The myth of bodily perfection. *Disability and Society, 10*(4), 413–424.

Superio-Cabuslay, E., Ward, M. M., & Lorig, K. R. (1996). Patient education interventions in osteoarthritis and rheumatoid arthritis: A meta-analytic comparison with non-steroidal anti-inflammatory drug treatment. *Arthritis Care and Research*, *9*(4), 292–301.

Taal, E., Riemsma, R. P., Brus, H. L., Seydel, E. R., Rasker, J. J., & Wiegman, O. (1993). Group education for patients with rheumatoid arthritis. *Patient Education and Counseling*, *20*, 177–187.

Taal, E., Seydel, E. R., Rasker, J. J., & Weigman, O. (1993) Psychosocial aspects of rheumatic diseases: Introduction. *Patient Education & Counseling*, *20*, 55–61.

Timko, C., Stovel, K. W., & Moos, R. H. (1992). Functioning among mothers and fathers of children with juvenile rheumatic disease: A longitudinal study. *Journal of Pediatric Psychology, 17*, 705–724.

Ungerer, J. A., Horgan, B., Chaitow, J., & Champion, G. D. (1988). Psychosocial functioning in children and young adults with juvenile arthritis. *Pediatrics, 81,* 195–202.

Vandvik, I. H., & Eckblad, G. (1991). Mothers of children with recent onset rheumatic disease: associations between maternal distress, psychosocial variables and the disease of the children, *Journal of Developmental and Behavioral Pediatrics, 12*, 84–91.

Varni, J. W., Wilcox, K. T., Hanson, V. (1988). Mediating effects of family social support on child psychological adjustment in juvenile rheumatoid arthritis. *Health Psychology, 7*, 421–431

Walco, G. A., Varni, J. W., & Ilowite, N. T. (1992). Cognitive-behavioral pain management in children with juvenile rheumatoid arthritis. *Pediatrics*, *89*, 1075–1079

Wallander, J. L. & Varni, J. W. (1989). Social support and adjustment in chronically ill and handicapped children. *American Journal of Community Psychology, 17*, 185–201.

Wallston, K. A. (1993). Psychological control and its impact in the management of rheumatological disorders, *Bailliere's Clinical Rheumatology. International Practice and Research, 7*(2), 281–295.

Wegener, S. T. (1991). Psychosocial aspects of rheumatic disease: The developing biopsychosocial framework. *Current Opinion in Rheumatology*, *3*, 300–304.

Weinman, J. (1990). Providing written information for patients: Psychological considerations. *Journal of the Royal Society of Medicine, 83*, 303–305.

Williams, G. H. (1989). Hope for the humblest? The role of self-help in chronic illness: The case of ankylosing spondylitis. *Sociology of Health and Illness*, *11*(2), 135–159.

Williams, B., Cullen, L., & Barlow, J. (1996). "My batteries have been recharged": The experiences of older people with arthritis attending a personal independence course: *Florida Arthritis and Rheumatism, 39*, 1161.

Wright, G. E., Parker, J. C., Smarr, K. L., Schoenfeld-Smith, K., Buckelew, S. P., Slaughter, J. R., Johnson, J. C., & Hewett, J. E. (1996). Risk factors for depression in RA. *Arthritis Care and Research, 9*, 264–172.

Yelin, E., Lubeck, D., Holman, H., & Epstein, W. (1987). The impact of rheumatoid arthritis and osteoarthritis the activities of patients with rheumatoid arthritis and osteoarthritis compared to controls. *Journal of Rheumatology, 14*, 710–717.

Young, L. D. (1992). Psychological factors in rheumatoid arthritis. *Journal of Consulting & Clinical Psychology, 60*(4), 619–627.

Zant, J. L., Dekker-Saeys, A. J., Van den Burgh, I. C., Kolman, A., & Van der Stadt, R. J. (1982). Sthenia, ambition and educational level in patients suffering from ankylosing spondylitis: A controlled study of personality features as compared to rheumatoid arthritis and unspecified low back pain. *Clinical Rheumatology*, *4*, 243–250.

Zautra, A. J., Burleson, M. H., Smith, C. A., Blacock, S. J., Wallston K. A., DeVellis R. F., DeVellis, B. M., & Smith, T. W. (1995). Arthritis and perceptions of quality of life: An examination of positive and negative affect in rheumatoid arthritis patients. *Health Psychology, 14*(5), 399–408.

Zautra, A. J., Okun, M. A., Robinson, S. E., Lee, D., Roth, S. H., & Emmanual, J. (1989). Life stress and lymphocyte alterations among patients with rheumatoid arthritis. *Health Psychology*, *8*, 1–14.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatra Scandinavia, 67*, 361–370.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.19
Spinal Cord Injuries

PAUL KENNEDY
Stoke Mandeville Hospital, Aylesbury, UK

8.19.1 THE INJURY 445
- *8.19.1.1 History* 445
- *8.19.1.2 Spinal Cord* 446
- *8.19.1.3 Epidemiology* 446
- *8.19.1.4 Etiology* 446
- *8.19.1.5 Incidence and Prevalence* 446
- *8.19.1.6 Rehabilitation Management* 447

8.19.2 PSYCHOLOGICAL IMPACT 447
- *8.19.2.1 Emotional Reactions* 447
- *8.19.2.2 Suicide and Indirect Self-destructive Behavior* 449
- *8.19.2.3 Neuropsychological Concerns* 450
- *8.19.2.4 Chronic Pain* 450
- *8.19.2.5 Secondary Complications* 450

8.19.3 INTERPERSONAL AND SOCIAL ADJUSTMENT 450
- *8.19.3.1 Sexual Dysfunction* 450
- *8.19.3.2 Impact on Relationships* 451
- *8.19.3.3 Employment After Injury* 451

8.19.4 PSYCHOLOGICAL ADJUSTMENT AND QUALITY OF LIFE 451

8.19.5 COPING WITH SPINAL CORD INJURY 452

8.19.6 PSYCHOLOGICAL INTERVENTIONS 453
- *8.19.6.1 Initial Contact* 454
- *8.19.6.2 Management of Emotional Concerns* 454
  - *8.19.6.2.1 Intervention methods* 454
  - *8.19.6.2.2 Cognitive-behavioral strategies* 454
  - *8.19.6.2.3 Coping effectiveness training* 455
- *8.19.6.3 Skills Training* 456
- *8.19.6.4 Rehabilitation Planning and Adherence* 456
- *8.19.6.5 Sexual Counseling* 457
- *8.19.6.6 Chronic Pain Management* 457
- *8.19.6.7 Biofeedback* 457

8.19.7 PROFESSIONAL ISSUES 457

8.19.8 LIFE SPAN AND SOCIETAL RESPONSES 458

8.19.9 REFERENCES 459

8.19.1 THE INJURY

8.19.1.1 History

The first descriptions of spinal cord injuries are found in the *Edwin Smith Surgical Papyrus* which records 5000-year-old cases of spinal cord injury (Hughes, 1988). An unknown ancient Egyptian physician summed up one of the cases as follows: "Having a dislocation in a vertebra of his neck while he is unconscious of his two legs and his two arms, and his urine dribbles. An ailment not to be treated." The first text book devoted exclusively to the spinal cord was written by the Dutch anatomist Gerard Blasius (1625/1692), in *Anatomy Medullae Spinalis Et Nervorum*. The initial pessimism recorded 5000 years ago, that this was not an ailment to be treated, persisted until 1944 when the first rehabilitation programs were established in Stoke Mandeville Hospital in the UK by Sir Ludwig Guttmann, and in West Roxbury, MA in the USA by Donald Munro. Up until then, 80% of people with this condition died in the first few weeks (Guttmann, 1976). Since the 1940s, both acute and long term survival rates for persons with spinal cord injury have improved dramatically (DeVivo & Stover, 1995). Now, with life expectancy greatly enhanced, and the availability of rehabilitation centres in most developed health care systems, quality issues, rather than basic survival, have become the main concern.

8.19.1.2 Spinal Cord

The spinal cord is part of the central nervous system and is the main conveyance for sensory, motor and autonomic neural information. It is encased in the spinal vertebrae and extends from the brain stem to the cauda equina. An injury to the spinal cord occurs when sufficient force causes the cord to be compressed, lacerated or stretched and may be associated with a fracture or fracture-dislocation of the vertebral column and displacement of the intervertebral disks. These injuries are often the consequence of sudden impact as with car accidents and falls. Nontraumatic damage may be caused by pressure on the spinal cord, as in the case of tumors, or an interruption of the blood supply to the spinal cord. Neurological losses will depend upon the completeness and level of the injury.

Cervical (neck) injuries may result in tetraplegia (quadriplegia). Injuries to the thoracic, lumbar and sacral levels (upper, middle, and lower back) may result in paraplegia. Fifty-three percent of injuries are neck injuries, and 47% are back injuries (Bedbrook, 1981). A complete injury is one in which all motor and sensory functions are lost below the level of injury as well as the loss of control of visceral functions such as bladder, bowel and sexual functions. In paraplegia, the legs and the lower trunk muscles are impaired, whereas cervical injuries can result in the loss of function in the hands, arms, shoulder and diaphram, as well as the legs and lower trunk. Complete injuries above the seventh cervical segment may preclude the possibility of independent living. Sensory losses are similar to the area of motor loss and include the loss of touch, pressure, temperature sensation and position sense. The loss of sexual sensation and responsivity, and control of bladder and bowel function, are common to all complete lesions above the sacral roots. Incomplete injuries may result in partial neurological impairment. In the American National SCI data base, Go, DeVivo, and Richards (1995) recorded that the most common discharge neurological category was incomplete tetraplegia (31%) followed by complete paraplegia (26%), complete tetraplegia (22%), incomplete paraplegia (20%), and complete recovery (1%). The spinal cord is part of the central nervous system and, once damaged, nerve cells do not regenerate (Guttmann, 1976). Complications postdischarge include problems associated with pressure ulcers, urinary tract infections, respiratory disorders, muscular spasm, and chronic pain.

8.19.1.3 Epidemiology

From the American National SCI data base of over 66 000 people (Go, DeVivo, & Richards, 1995), 82% of the sample were male and 18% female. Kennedy (1995) in a British sample found 81% of admissions to be male and 19% female. Population-based studies have revealed that spinal cord injury occurs most frequently in teenaged persons and young adults aged between 16 and 30 years, the mean age being 30.7, and the most frequently occurring age at injury being 19. Fifty-three percent of people with spinal cord injury are single, 31% are married and 16% are either divorced, separated or widowed (Go, DeVivo, & Richards, 1995).

8.19.1.4 Etiology

In general, in most developed economies road traffic accidents account for approximately 50% of spinal cord injuries, approximately 25% are caused by falls and 20% are caused by sporting injuries (Kennedy, 1995; Stover & Fine 1986). Most sporting injuries are caused by diving into the shallow end of a swimming pool. In the USA, a significant proportion of injuries are caused by acts of violence (accounting for almost 17% of injuries). In developing economies, road traffic accidents are less frequent, with increases in agricultural accidents and falls.

8.19.1.5 Incidence and Prevalence

The annual incidence of spinal cord injury in the UK, as for most other European countries, is between 10 and 15 per million. In the USA it is thought to be nearer 30 per million and in Japan, 27 per million. There are an estimated 40 000 people in the UK and 200 000 people in the USA with spinal cord injuries. Life expectancy estimates vary: for a young person with an incomplete injury the relative survival rate is 69%, but for complete tetraplegics over 50 it is estimated to be 33%. Primary causes of mortality include pneumonia, septicemia, renal failure, and pulmonary emboli which are more common in the first few months post injury. Saturday and Sunday are the high risk days for having a spinal cord injury, particularly during the afternoon, evening and early hours of the morning. Most injuries occur during the months of July, August, and June and are least likely to occur in February and January.

8.19.1.6 Rehabilitation Management

Most people are transferred to a specialist spinal cord injury treatment and rehabilitation center shortly after the onset of the disorder. Once the fracture is stabilized either through internal or external fixation, or postural reduction (bedrest), the person begins a period of rehabilitation which can last between three months and one year, depending on the level of injury and the individual's personal circumstances. DeVivo, Whiteneck, and Charles (1995) estimate that rehabilitation costs US$1.9 million (1992 dollars) of direct expenses for a 25-year-old with a complete C4 spinal cord injury. In the American Model Spinal Cord Rehabilitation Centers the overall mean length of stay in acute and rehabilitative care was 3.8 months. An informal internal audit in Stoke Mandeville Hospital between 1995 and 1996 found the overall length of stay for acute management and rehabilitation to be 4.9 months.

8.19.2 PSYCHOLOGICAL IMPACT

Given the relatively young age at onset and in terms of potential years of life changed, the spinal cord injury presents a formidable challenge for the individual. It may require major adjustment to physical, psychological and social domains. Much has been said about its devastating nature, but more recently researchers (Kennedy, 1995; Menter & Hudson, 1995) using longitudinal approaches indicate that many survivors of spinal cord injury go on to live satisfying, meaningful and rewarding lives. The extent to which people manage the consequences of their disabilities will depend on internal factors, such as self-efficacy and coping styles, and external factors, such as prevailing social attitudes and social support. The social context and associated cultural beliefs underpinning these domains are important given that it is not the specific disability *per se* which causes most of the individual's difficulty, but existing within physical, psychological and socially disabling environments (Swain, Finkelstein, French, & Oliver, 1993). Following the spinal cord injury, the person has usually survived a life-threatening incident and then has to learn to deal with the changes in their physical needs and concurrently address domestic, social, and vocational issues. The immediate impact of the injury may be compounded by disturbed consciousness because of the brain injury and medication, severe pain, and sensory deprivation. Trieschmann (1988) pointed out that, in the acute phase, psychological disturbance may result from sensory responses to the injury, the hospital environment, and the procedures utilized to enable survival of the person, rather than consideration of the spinal cord injury itself.

8.19.2.1 Emotional Reactions

Emotional reactions to a spinal cord injury are typified by increased levels of depression, anxiety (and associated avoidance behaviors), with some indication of temporary increases in anger and hostility. However, negative emotional responses are neither universal or necessary for successful reintegration and rehabilitation. Earlier researchers utilized clinical impressions and empirical studies to construct stage models that account for impact and adjustment. Depression is considered to be a normal part of the adjustment process and is readily observed. Many have held the view that all spinal cord injured patients must at some point suffer from depression (Nagler, 1950; Siller, 1969). Indeed, they suggested that if depression was not observed then they considered the patient was in denial and that for healthy adjustment depression is a stage that must be passed through. These models could consist of two stages (Guttmann, 1976), three stages (Stewart, 1977), four stages (Rigoni, 1977), and five stages (Bracken & Shepard, 1980; Hohmann, 1975). Stages generally include shock, denial, depression, anger, and conflict resolution. These models may have had some utility in recognizing and normalizing emotional reactions. However, there is little empirical validation of the sequence, duration, or existence of the stages, or indeed that they are helpful in fostering adjustment (Morris, 1992; Wortman & Silver, 1989). Furthermore, such unrealistic assumptions held by healthcare professionals and the social network may also unnecessarily exacerbate feelings of distress among those who encounter loss and lead to the self-perception that their own responses are inappropriate and abnormal under the circumstances. A brief clinical anecdote may illustrate this point. In 1985 I was approached by a patient who had acquired a complete traumatic spinal cord injury of the fifth cervical segment. He enquired of the possibility of providing him with help to become depressed! He was an Olympic gold medallist and intended to complete his rehabilitation with distinction. He had been told by a counsellor that, as he had not been depressed, this meant he had not "come to terms" with his disability. He had spent a number of weeks trying to become depressed without success and my intervention was to endorse and confirm the appropriateness of his effective coping strategies and tell him that depression was not a component of our rehabilitation program.

There is no evidence that a pre-injury personality trait exists, or that there is a specific personality characteristic associated with the onset of the spinal cord injury (Fordyce, 1964; Trieschmann, 1988). The Minnesota Multiphasic Personality Inventory (MMPI) has been used in a number of early studies of psychological impact. Boureston and Howard (1965) compared MMPI profiles for persons with spinal cord injuries, multiple sclerosis, and rheumatoid arthritis. They found that those with spinal cord injury had the least emotional distress, the least anxiety, and the least concern with traumatic problems. Taylor (1967) administered the MMPI to a group of persons with spinal cord injury within a month of injury. The profiles were within normal limits with no evidence of marked depression or anxiety. In a small but detailed study of depression, Lawson (1978) followed 10 tetraplegics for their entire rehabilitation, and analyzed staff and patient ratings of depression. Over this period he found the level of depression decreased with time, and that high levels of depression were associated with highly personalized significant events. Fullerton, Harvey, Klein, and Howell (1981) investigated depression in the study of 30 patients. They found that 30% of those followed up developed a depressive disorder after injury. This is a higher incidence than that reported in the general population which is estimated to be about 6% using the same criteria (Weisman, Myers & Harding, 1978).

Judd, Stone, Webber, Brown, and Burrows (1989) assessed 71 patients with spinal cord injury in their acute and rehabilitation phases of treatment. Patients completed the Beck depression inventory (BDI) weekly and a clinician completed the Hamilton rating scale for depression. Fourteen patients (i.e. 20%) met the *Diagnostic and statistical manual of mental disorders* (3rd ed., *DSM-III*) criteria for major depressive disorder, a further 13 patients had transient periods of depressed mood, whereas the majority of patients showed no clear evidence of depression. Richards (1986) investigated 36 persons with spinal cord injury at three weeks, three months and at one year following discharge using the BDI and a self-report questionnaire. Twenty-nine able bodied persons completed the same at similar time intervals. Immediately following discharge, moderately increased depression was found in the spinal cord injury group, but after a year this difference had gone. These studies found that depression, although common, does not necessarily follow the spinal cord injury and tended to be resolved either by the time of discharge, or within a year of discharge. It would appear that the rates of depressive disorder have been found to be between 25% and 30% per sample, less than previous literature suggested but more than the general population.

Although most interest in immediate psychological impact in spinal cord injury has focused on depression, other emotional reactions have been investigated. Cook (1979) investigated 118 persons with spinal cord injury less than one year following their injury using the State–Trait Anxiety Inventory (STAI) and the Mini-mult (an abbreviated form of the MMPI). Comparisons with normative data revealed no significant anxious or depressive reactions in the spinal cord injury group. Bracken, Shepard, and Webb (1981) examined 190 patients on discharge from acute hospitalization using a number of specific measures of anger, anxiety, and depression. Although appreciating the limitations in comparing the results of this study with more standardized studies, their results suggested increased levels of anger, anxiety, and depression, which continued following discharge and regulated that these problems would interfere with community integration. Kennedy, Lowe, Grey, and Short (1995) compared a group of 41 patients six weeks post injury with a group of 30 respondents between four and seven years post injury. They found no differences between the group scores on the BDI, the Beck Hopelessness Inventory, the STA, the intrusion subscale of the impact of Events scale and scale 7 (psychological distress) of the Psychosocial Adjustment to Illness scale. In line with previous findings, 34% of the inpatient group and 37% of the post discharge group scored above the clinical cut-off point of 14 on the BDI.

Hancock, Craig, Dickson, Chang, and Martin (1993) implemented a longitudinal study exploring depression and anxiety as measured by the BDI and the STAI of three time points following injury. They compared 41 spinal cord injured patients with a matched able-body control group. Whereas 25% of the spinal cord injured group were anxious and depressed, only 5% of the control group were anxious and 3% depressed. Trieschmann (1988) critically reviewed the published literature and emphasized the need to adopt a more longitudinal method of enquiry, elaborating on not only aspects of the psychological impact following the injury, but exploring the relationship between psychological well-being, coping strategies, and adjustment. In response to these recommendations, Kennedy (1995) implemented a longitudinal study following up a cohort of traumatic spinal cord injured people from onset of injury through rehabilitation and up to two years post hospital discharge. Respondents were assessed on nine observational periods on measures of depression, hopelessness, anxiety, and general psychological distress. The measures showed no change over time, but some reduction in psychological morbidity was noted following discharge. The mean score for inpatients on depression fell within the mild/moderately depressed range on the BDI and two years post discharge was found to be within the normal range. However, this longitudinal study also highlighted the relative consistency of those respondents who were initially identified as having problems with anxiety and depression. The number of respondents going above the clinical cut-off scores for depression and anxiety (i.e., BDI greater than 14 or STAI greater than 48) remained fairly stable across the nine observational periods, demonstrating that for this sample a group between a quarter and a third can remain psychologically vulnerable and stressed from onset of injury to community placement two years post discharge. However, it is arguable that this may have overestimated psychological morbidity in that those who responded may represent a more vulnerable subsection of the sample.

8.19.2.2 Suicide and Indirect Self-destructive Behavior

In a study of 9135 persons injured between 1974 and 1984, DeVivo, Black, Richards, and Stover (1991) found that 50 people had committed suicide (6.3% of deaths) producing a standard mortality ratio (SMR) of 4.9 (i.e., ratio of observed to expected suicides). They concluded that although suicide occurred more often in the spinal cord injured population, in an absolute sense it is a rare event. Those with the highest SMR were those with neurologically complete paraplegia. Charlifue and Gerhart (1991) reviewed 489 deaths in a population of 5200 persons with spinal cord injury and found that 9% of the deaths were caused by suicide. These rates of suicide, as a percentage of total deaths, may underestimate the true scale of the problem in a spinal cord injured population. Persons with spinal cord injury were more likely to die than the general population from a number of medical problems. Charlifue and Gerhart (1991) produced a model predicting 81% of the variance associated with suicide and factors included age (with risk increasing with age), injury from a fall, a complete neurological lesion, history of family fragmentation, alcohol abuse, and social withdrawal. Twenty percent of suicides occur in the first year after injury, and 70% occur between one and five years. In the UK, Charlifue et al. (1996) found that of 909 deaths (between 1943 and 1990) 22 were caused by suicide, producing an SMR of 2.4.

Suicide accounts for deaths by direct self-destructive behavior, but in spinal cord injury self neglect can equally be fatal. Health maintenance following spinal cord injury needs to be disciplined and consistent. Failure to do so may result in problems such as pressure sores or contractures. Wilcox and Stauffer (1972) found that suicide or self-neglect accounted for 43% of the 50 recorded deaths in a group of 423 patients. Nehemkis and Groot (1980) used the term "indirect self-destructive behavior" to describe this type of noncompliance. They explored in detail the autopsy protocol of all spinal cord injured patients who died between 1971 and 1976. Of the 52, they found that 49% of the cases were certified as natural, whereas 33% were considered to have an indirect self-destructive behavioral contribution to the cause of death.

Therefore, in summary, the onset of spinal cord injury is associated with elevations in general emotional distress. The sudden onset of the injury is often associated with periods of shock, numbness, disbelief, and denial. Depression and anxiety may emerge as part of a normal and rational response but were neither necessary nor sufficient for effective adjustment. Depression may emerge as part of a normal response to the perceived losses. Anxiety occurs when people believe that their coping strategies are exhausted or are insufficient in dealing with a current or anticipated situation. Adjustment may be further complicated by self-destructive behavior, but suicide rates are only slightly higher than in the general population, although indirect self-destructive behavior is more common.

8.19.2.3 Neuropsychological Concerns

There is a growing awareness of the need to assess and identify neuropsychological impairment, which complicates adjustment in many chronic health problems (Kelly and Doty, 1995).

Reported prevalence figures for neuropsychological deficits in patients with spinal cord injury vary. Estimates of the frequency of cognitive deficits range between 10% and 60% (Davidoff, Roth, & Richards, 1992; Morris, 1992). A variety of factors contribute to this frequency. The most often cited reason is concomitant traumatic brain injury which may be as high as 50% (Davidoff et al., 1988). The long term cognitive effects of alcohol and substance abuse may also contribute to cognitive deficits in this population (Heinemann, Keen, Donohue, & Schnoll, 1988). Deficits are generally diffuse in nature but the specific deficits that hinder rehabilitation include difficulties with attention, speed of information processing, memory, concentration, producing problem solving per judgment, and impaired social appropriateness (Morris, Roth & Davidoff, 1986).

8.19.2.4 Chronic Pain

Chronic pain is a prevalent and significant complication of spinal cord injury, the prevalence of which is estimated to be between 25% and 45% (Britell & Mariono, 1991; Kennedy, Frankel, Gardner, & Nuseibeh, 1997.). In a survey completed by Rose, Robinson, Ells, and Cole (1988) 36% of 885 out-patients with spinal cord injury suggested that it was pain severity and not their physical disabilities that prevented them from returning to work. This is a major factor which complicates adjustment and compounds the individual's sense of burden.

8.19.2.5 Secondary Complications

People with spinal cord injury have a lifelong risk of developing a pressure ulcer (sore) and Fuher, Garber, Rintala, Clearman, and Hart (1993) estimated the prevalence of pressure sores to be 33%. Yarkony and Heinemann (1995) in their review of 66 000 people with spinal cord injury and managed under the Model System Care program found the prevalence of pressure sores to be 32%. Seekins, Clay, and Ravesloot (1994) investigated the secondary complications in a community sample of adults with physical disabilities, of whom 20% had a spinal cord injury. Significant complications included mobility difficulties, chronic pain, spasticity, bladder and bowel dysfunction, and pressure sores. Ivie and DeVivo (1994) attempted to identify the factors predicting unplanned hospitalizations post injury. They found that measures of functional assessment (the Functional Independence Measure) rather than neurological impairment were more predictive, and that a greater risk of readmission was found with those who were unemployed. Seekins et al. (1994) suggest three components aimed at prevention and management. The first involved wellness promotion and physical reconditioning, the second included behavioral management procedures to address chronic pain, sleep disturbance, and depression, and the third involved increasing physical access to the resources of the community.

8.19.3 INTERPERSONAL AND SOCIAL ADJUSTMENT

8.19.3.1 Sexual Dysfunction

Sexual dysfunction following spinal cord injury has received increased attention since the 1970s. Sipski and Alexander (1992) reviewed this literature and highlighted that between 70% and 93% of males with complete upper motor neuron lesions retain reflex erections, but none sustained psychogenic erections. Ejaculation was noted to occur in 4% of these patients. Of those with lower motor neuron lesions, 26% were able to sustain a psychogenic erection, but none experienced reflex erections. Eighteen percent of this population were able to ejaculate. Therefore, in general, as most are upper motor neuron lesions, men lose the ability to produce a psychogenic erection, while retaining a reflexogenic capacity. Few experience orgasm and most have impairment in the ability to move during intercourse. Despite the increased interest in research, little was known of the physiology of female sexual dysfunction following spinal cord injury other than a vague suggestion that female sexual function is analogous to that in males. However, Sipski, Alexander, and Rosen (1995) report on the first laboratory-controlled analysis of female sexual response and elaborated on the nature of the female reflex vasocongestion. Despite these physiological changes, most women returned to sexual involvement within 12 months of injury (Alexander, Sipski, & Findlay, 1993) and 57% rated post injury sexual relations as satisfying or at least as satisfying as before (Siosteen, Lundqvist, Blomstrand, Sullivan, & Sullivan, 1990). However, the frequency of penis–vagina intercourse decreases with an associated increase in sexual activities, related to areas above the level of the spinal cord injury. Siosteen et al. (1990) found that positive sexual adjustment post injury was most closely related to a higher frequency of intercourse, a willingness to experiment with alternative modes of sexual expression, and that young people tended to have a better sexual adjustment than older.

Fertility is also compromised in males as the result of the neurological impairment and the nerves involved in emission and ejaculation. It also results in diminished concentration and motility of sperm, Linsenmeyer and Perkash (1991). Currently, assistive techniques aimed at producing ejaculates are being developed. After an initial period of amenorrhoea (Comarr, 1966) women remain fertile.

8.19.3.2 Impact on Relationships

The buffering role of social support is demonstrated throughout this volume. It is not surprising that the impact of the injury on relationships has been the subject of investigation. However, to date there is much ambiguity and equivocation in the findings. Siller (1969) suggests that divorce rates in pre-injury marriages were higher than in the general population, whereas Abrams (1981) summarized 30 years of the literature relating to marital stability, sexual interaction and marital satisfaction, and concluded that paraplegia does not seriously impair marital functioning. Dijkers and Abela (1993) found that the divorce rate for the first five years post injury was higher than expected given the general population, age, sex, and specific divorce rates. In a British study, Guttmann (1964) concluded that post injury marriages had a better rate of success. However, DeVivo and Fine (1985) followed 276 patients for three years post injury and reported that those patients had significantly fewer marriages and fewer divorces than would be expected according to base rates in the US population. However, these studies cover only the first few years post injury and it is possible that they reflect the acute disruption resulting from the injury and that the trends might not hold over a longer period of time.

Crewe and Krause (1988) sent questionnaires to 122 married individuals with spinal cord injury and compared persons in pre-injury marriages and those in post injury marriages. The mean time since injury for the pre-injury marriage group was 13 years, and for the post-injury marriages, 18 years. They obtained basic demographic information and used a modified version of the Life Satisfaction Questionnaire (LSQ) to examine satisfaction and adjustment. They concluded that marriages which survived for years after one partner experiences a spinal cord injury are generally viewed as satisfying and comfortable. Yet when compared with marriages established after the onset of disability, differences do emerge in that couples in post injury marriages were more likely to be employed, go out socially, and report higher levels of satisfaction with life. These discrepancies held even when age differences were statistically controlled. In a later study, Crewe and Krause (1992) suggested that it may be that only individuals who are especially likeable, active and well adjusted, succeed in attracting partners, and as a consequence post injury marriages are happier. The US National SCI database contains the most up-to-date and systematically collected data on marital status (Dijkers, Abela, Gans & Gordan, 1995). They concluded that changes in marital status tended to be highest in the earlier years post injury and that over time there is an increase in both divorces and marriages. Marriage failures were most common in the younger age group (16–30). In this study of marital relationships in a cohort of 60 000 people with spinal cord injury, they found that new marriages were less stable than pre-injury marriages. This contradicts previous findings and it may be that those in pre-injury marriages were older than those in post injury marriages. Further analysis is required on this data; controls for age at injury and age at time of divorce are required.

8.19.3.3 Employment After Injury

When Guttman was asked to define rehabilitation in a BBC UK television documentary (BBC, 1966), he defined it as "making the individual a tax payer again." The return to full-time gainful employment is a useful definition of successful rehabilitation. Re-employment figures after spinal cord injury reported in the literature vary tremendously. Trieschmann (1988) reports a range of return to competitive employment varying between 13% and 48%. Dijkers et al. (1995) in a recent review of the literature gives the following generalization: paraplegics have a higher re-employment rate than tetraplegics; those who are younger are more likely to become employed; pre-injury vocational experiences predict post injury employment; minority groups are less likely to become employed; and involvement in a vocational rehabilitation program makes employment more likely. Creek, Moore, Oliver, Silver, and Zarb (1988) found in Britain that, although 97% of people were employed at the time of the injury, only 50% were in employment at a five year follow-up.

8.19.4 PSYCHOLOGICAL ADJUSTMENT AND QUALITY OF LIFE

Trieschmann (1988) described rehabilitation as a life long process of learning to live with a disability in one's own environment, suggesting that there is no end point for adjustment and that it should improve over time. Factors which influence adjustment, such as personality characteristics, level of injury, and social support, have been investigated. There is no evidence that psychological well-being and positive adjustment are predicted by the level of injury (Cook, 1979; Kennedy, Lowe, et al., 1995; Kennedy, Gorsuch, & Marsh, 1995). Krause and Crewe (1990) followed a cohort of people with spinal cord injury from 1977 until 1984 and used the LSQ to confirm research findings that quality of life appears to improve over time. The LSQ assesses activity, adjustment, interpersonal satisfaction and economic satisfaction. Positive levels of satisfaction in psychological status were maintained, with increases in marriage rates and hours spent working. Woodrich and Patterson (1983) studied a group of 251 people who ranged from six months to two years post injury and they found a strong positive correlation between time post injury and acceptance of disability (as measured by the Acceptance of Disability Scale (Linkowsky, 1971). They found that females were more accepting than males, younger people were more accepting than older people, and they did not find an association between acceptance of disability and level of injury or ethnic origin.

Whiteneck et al. (1992) reviewed the mortality, morbidity, health, functional and psychosocial outcomes of 834 individuals with long term spinal cord injuries. Not surprisingly, they found that functional status declined with the ageing process, and although they found some decline in measures of life satisfaction, 75% of those interviewed rated their current quality of life as either good or excellent on a five-point scale. Decker and Schulz (1985) found that adjustment as measured by the Life Satisfaction Index was positively correlated with their own social support scale. In a mailed survey of 125 persons with spinal cord injury using the Quality of Life Index (QOLI) and the Reciprocal Social Support Scale, Anson, Stanwyck, and Krause (1993) found that increased social support was associated with positive adjustment and decreased secondary complications (i.e., fewer pressure sores and less pain).

In general, the main themes from the literature suggest: younger people tend to adjust to it better than older people; an internal locus of control has been shown to predict better adjustment than an external locus (Bracken, Shepard, & Webb, 1981; Shadish, Hickman, & Arrick, 1981); those who believe they were in some way responsible for their injury adjust better than those who believe themselves to be innocent victims (Bulman & Wortman, 1977). There appears to be contradictory evidence about the role of educational background, Woodrich and Patterson (1983) concluded that the more educated adjusted better; in contrast, Hancock, Craig, Dickson, Chang, and Martin (1993) found no educational effect. In summary then, the evidence suggests that following spinal cord injury the majority of individuals experience a quality of life which is comparable with their able bodied peers and poor psychological adjustment is not an inevitable consequence. Now let us examine the factors associated with coping.

8.19.5 COPING WITH SPINAL CORD INJURY

Frank et al. (1987) identified, by cluster analysis, two subgroups differing in coping styles within a group of 53 persons with spinal cord injury. The cluster variables were coping, as measured by the Ways of Coping (WOC) scale (Folkman & Lazarus, 1988), and locus of control as measured by the Multidimensional Health Locus of Control scale (Wallston, Wallston, & DeVellis, 1978). Cluster one tended to use all the coping strategies in the WOC scale to a greater extent than cluster two. However, the general pattern was quite similar. Most used was wishful thinking, followed by problem-focused strategies, with self-blame being the least used. Kennedy, Lowe, et al. (1995) investigated the relationship between depression (using the BDI; Beck & Steer, 1987), anxiety (using the STAI; Speilberger, Gorsuch, Lushene, Vagg & Jacobs, 1983) and social support using the Social Support Questionnaire (Sarason & Sarason, 1986) and coping strategies as measured by Carver, Scheier, and Weintraub's (1989) COPE Questionnaire. They examined two groups of patients, 41 at six weeks post injury and 30 between four and seven years post injury. A series of regression analyses found that the coping strategies explained more of the variance in psychological impact measures than moderating variables such as age, level of injury or social support. The models produced demonstrated an association between increased levels of psychological distress and the use of the coping strategies of behavioral disengagement, alcohol and drug use ideation, and the underutilization of the strategy of acceptance. The post-discharge group used emotional and instrumental social support less often and scored higher on venting emotions and behavioral disengagement.

However, there are difficulties in making causal interpretations from this type of cross-sectional study. Using a longitudinal design, Buckelew and Hanson (1992) measured coping using the adapted WOC questionnaire (developed from Felton & Revenson, 1984) in 57 persons with spinal cord injury during acute rehabilitation and followed them up five years post discharge. Adjustment was measured by the Acceptance of Disability scale (Linkowski, 1971) and the Symptom Check List-90-R (Derogatis, 1977). This preliminary and methodologically weak study explored the issue of coping strategies and the prediction of adjustment. They found that cognitive restructuring was associated with high levels on acceptance of disability, whereas wish-fulfilling fantasy strategies were associated with low levels. Reidy, Caplan, and Shawaryn (1991) followed up patients between 18 and 24 months post discharge. Using the WOC questionnaire, the BDI and the Multiple Affective Adjective Checklist they found that depression had a strong positive correlation with the use of escape avoidant coping strategies. Positive affective states were positively associated with distancing, seeking social support, positive reappraisal, and planned problem solving.

Kennedy (1995) employed a longitudinal methodology to examine psychological impact and coping strategies (using the COPE) following traumatic spinal cord injury. This prospective longitudinal design employed repeated measures to examine coping strategies over time from six weeks post injury to two years post discharge. A remarkably consistent pattern of coping strategies emerged and no significant differences were noted across time. The most frequently used strategies were acceptance, positive interpretation and growth, active coping, planning, and social support; the least used were behavioral disengagement, denial, and alcohol and drug use ideation. Concurrent multiple regression analyses implicated similar coping strategies (acceptance, behavioral disengagement, alcohol and drug use ideation) in models predicting levels of anxiety and depression during hospitalization and after discharge. At six months post injury, using coping strategies (COPE), demographic data (age, sex), and measures of functional status (FIM) as independent variables and depression (as measured by the BDI) as the dependent variable, a model was produced which predicted 83% of the variance in depression. Predictive variables (in order of contribution) included: behavioral disengagement, low acceptance, focusing on and venting emotions, low emotional social support, alcohol and drug use ideation, and low functional independence. The longitudinal methodology employed by this study enabled the exploration of prediction across time. Coping strategies assessed at six weeks post injury were found to predict 67% of the variance in depression at one year post discharge. The model here included the use of focusing on and venting emotions, low acceptance and low functional independence. Fifty-seven percent of the anxiety at one year post discharge was predicted by the strategies of focusing on venting emotions, positive reinterpretation and growth (negative coefficient) and restraint coping, all collected at six weeks post injury. These results demonstrate that a significant predictive relationship exists between the coping strategies people use and the level of psychological distress following traumatic spinal cord injury. Again, mediating variables, such as coping, accounted for more variance in emotional adjustment than moderating variables such as age, sex, and level of injury.

8.19.6 PSYCHOLOGICAL INTERVENTIONS

Trieschmann (1988), in her review of the literature, commented on the paucity of longitudinal investigations into psychological impact, coping and adjustment. To some extent this need has been met by recent studies such as Kennedy (1995), Krause and Crewe (1990) and Whiteneck et al. (1992). Unfortunately the same cannot be said about the research on psychological interventions with this client group. McAweeney (1996) reviewed the literature of psychological interventions in the rehabilitation of persons with spinal cord injury over the past 20 years. She identified 221 studies and found that the majority (83%) were of quasi-experimental design, 16% employed a repeated measures design and only 1% reported the effect size. She did not dismiss the clinical relevance reported in the quasi-experimental studies, but highlighted the need for randomized controlled trials. However, it is appropriate to recognize the efficacy of established psychological techniques in this population group (Crewe & Krause, 1987; Kennedy, 1991; Trieschmann, 1988). In general, psychological interventions with spinal cord injured people aim to assist in the management of emotional concerns, maximizing the individual's rehabilitative potential, and minimizing general psychosocial disruption. The American Association of Spinal Cord Injury Psychologists and Social Workers (AASCIPSW; 1992) suggested that the psychosocial rehabilitation process should involve a continuity of services from the onset of spinal cord injury, through comprehensive rehabilitation, to follow-up in the community.

8.19.6.1 Initial Contact

Many people who are referred for psychological intervention in the acute rehabilitation phase are either on bed rest or are in the early stages of mobilization. Wanlass, Palmer, and Cook (1992) have described some helpful guidelines for improving the psychologists' bedside manner. These guidelines are particularly important to help manage initial therapeutic resistance. Many patients have an initial focus on medical treatment, physical therapy, and the idea of getting better (Crewe & Krause, 1987), whereas accepting psychological intervention may imply making plans for a future that could include a permanent disability. Wanlass et al. (1992) emphasized the importance of being familiar with the available information on the ward of both medical and social information, and before entering the patient's space, ask their permission to pull up a chair and spend time talking to them. Once you have identified yourself, it is important to inform the patient of your role and the purpose for making contact. It is helpful to emphasize that most people have a normal psychological reaction and that the psychologist is available to help with individual concerns and to foster coping and positive adjustment. Traditional empathic counselling skills are important in facilitating patients' engagement. Discussion should involve individual responses, acknowledgment of feelings and emotional reactions, and exploration of major relationships and the wider social context.

8.19.6.2 Management of Emotional Concerns

8.19.6.2.1 Intervention methods

Clinical experience suggests the early provision of supportive counselling, which includes emphasizing that the patient is now safe, normalizing emotional reactions, and listening to the patient's concerns, is beneficial. Taylor and Aspinwall (1990) concluded that such informal interventions can reduce the emotional distress of patients in medical settings. Initially the focus of attention is often the patient, but family support during this period is essential (Padrone, 1994). Once the initial rapport is established and the patient has been supported in expressing their concerns, one can then begin the processes of cognitive restructuring and fostering adjustment. It is important to acknowledge the catastrophic elements, but statements that reinforce self-efficacy and adaptation should be encouraged. At this stage it is important to remember that most people make a successful adjustment to the spinal cord injury (Kennedy, Gorsuch, & Marsh, 1995; Krause & Crewe, 1990; Trieschmann, 1988) and their quality of life improves over time. Telling patients and their families what to expect considerably alleviates anxiety (Taylor & Aspinwall, 1990) and at this stage they should be informed on the nature of the rehabilitation process.

Interventions to manage psychological distress and facilitate adjustment in coping can be applied at a number of levels (Wilson, Kennedy, & Glass, 1994; AASCIPSW, 1992). The first level involves individual assessment and the treatment of specific emotional disorders utilizing individual psychotherapy. The second level incorporates the provision of information regarding adjustment and emotional reactions and the training of staff to enhance psychological awareness and responsiveness. The third level indirectly addresses emotional issues using behavioral techniques in the form of goal planning to enhance engagement in rehabilitation and foster self-efficacy. Fourthly, the application of research findings that can clarify the process of adjustment, help identify maladaptive and adaptive coping strategies and challenge negative beliefs, form the basis of the evidence required to challenge negative beliefs.

8.19.6.2.2 Cognitive-behavioral strategies

Cognitive-behavioral approaches have had a remarkable influence in the management of emotional disorders since the mid-1970s (Kendall, Vitousek, & Kane, 1991). More recently, Sharpe et al. (1996) and Speckens et al. (1995) implemented a series of randomized control trials which utilized between six and 16 sessions of cognitive-behavioral therapy to demonstrate the effectiveness in managing problems of chronic fatigue and medical patients with unexplained physical symptoms, respectively. As mentioned, there is no need to replicate the efficacy of cognitive-behavioral therapy with this client group but it is appropriate to elaborate on the main themes and highlight some of the common cognitive distortions.

The cognitive-behavioral approach has much to offer in the management of spinal cord injury. There is common concensus on the efficacy of its application to this client group. Crewe and Krause (1987) and Kennedy (1991) describe the applications of cognitive-behavioural therapy with this population. During therapy, thoughts and beliefs about the consequences of the injury are explored. Specific periods of emotional distress are examined; they include the antecedent events, thoughts and beliefs, and emotional and behavioral consequences. These are separated and their interrelationships are discussed. Once specific negative thinking patterns and inferences are identified, related core assumptions are made explicit. By focusing on maladaptive and irrational patterns, negative assumptions about the consequences of the injury can be challenged and people are encouraged to respond to their injury in a rational and realistic fashion.

Turk and Salovey (1995) describe the following components of cognitive behavioral interventions for people with diverse chronic diseases and physical impairments. The first component involves reconceptualization which aims to shift a patient from automatic and ineffective responses toward systematic problem solving and planning. Problems and issues are recast in forms amenable to solutions, and therapy aims to foster hope, positive anticipation, and expectations of success. The second component in skills acquisition, fundamental to cognitive-behavioral therapy, is the need to build up the individual's awareness of the relationship between their thoughts, affect, and behavior. The initial focus is to make explicit the internal dialogue and irrational assumptions. The first stage of cognitive restructuring is to monitor these relationships and evaluate their validity and viability. Catastrophic statements such as "my life's over" and "I'm just going through the motions of life," can be identified and demonstrated as to how they can contribute to the sense of catastrophe and interfere with adaptive coping responses. With individuals who, as a result of motor dysfunction, are unable to physically record their thoughts, therapy focuses on key events in their recent past and the core assumptions that maintain negative thinking. The following case example may give an insight into the specific cognitions involved.

John had just finished university and was due to start work as a quantity surveyor when he was injured diving into the shallow end of a swimming pool. He injured his fifth cervical segment which resulted in a complete tetraplegia. He required a powered chair and help with most activities of daily living. Six weeks into his rehabilitation program he began to withdraw and complained of losing his appetite and having difficulty in sleeping. When interviewed it was clear that he was having great difficulty acknowledging the permanence of the paralysis. When we explored the reasons the injury was so difficult to accept, the following themes emerged. He believed he was unable to do anything worthwhile, that he would never be in control of his life again, and that he had now become a problem that needed to be managed. This, in turn, mobilized cognitions associated with previous vulnerabilities and negative self-perceptions. These negative predictions, global judgments and selective abstractions were challenged over eight sessions using a variety of cognitive-behavioral strategies. Decatastrophizing began by encouraging John to discuss his concerns and then exploring the basis of his cognitive distortions. We then examined realistically the activities that he was capable of engaging in, which was enhanced by discussing of comparable peers. This was introduced while exploring evidence for and against his negative predictions. We then challenged the belief that you have to be physically independent to be in control and encouraged his engagement in decision making. Throughout, John was able to discuss his emotions, such as his fears about the future and concerns about his losses, but these were balanced by general coping skills training which involved problem solving, reframing and active coping. During this time he increased his engagement, his level of affective distress decreased and assessment of coping strategies indicated an increase in level of acceptance and active coping and a decrease in behavioral disengagement.

In summary, this intervention aimed to communicate that this was a situation, that despite its severity could be managed, and that John possessed the capacity to do so. John has since returned to the community, is pursuing full-time education and lives in adapted accommodation with a personal care package funded by the local authority.

8.19.6.2.3 Coping effectiveness training

Many aspects of the acquired disability may not be manageable or controllable and, as an adjunct to cognitive-behavioral therapy, elements from cognitive-relational coping interventions are important (Folkman et al., 1991). This model suggests that a relationship exists between the person and the environment which may be cognitively appraised by the individual as personally significant and taxes or exceeds their resources (Lazarus & Folkman, 1984). The process of coping and adjustment is mediated by cognitive appraisal and coping. Coping has two major functions, namely to alter the problem that is causing stress or to regulate the emotional response to the problem. Training in coping effectiveness utilizes cognitive methods to appraise the stressor, and to identify the changeable and unchangeable aspects. When demands are perceived to be amenable to resolution, they are trained to utilize problem-focused coping strategies, whereas in situations where demands are appraised as not changeable, individuals utilize emotion-focused forms of coping.

Following spinal cord injury, individuals engage a range of coping responses and it is possible to identify those associated with good outcomes. Kennedy (1995) concluded that individuals who were not distressed two years post injury were more likely to employ the coping strategies of acceptance and positive reframing, and less likely to use behavioral disengagement, alcohol and drug use ideation and denial.

Therapy aimed at increasing adaptive and decreasing maladaptive strategies may be worth further investigation. A pilot study to explore these issues is currently being implemented by this author. Two other critical components of cognitive, behavioral and coping training include techniques to consolidate skills and to ensure generalization and maintenance post therapy. It is helpful to supplement prevention with patient information booklets and handouts that elaborate on the therapeutic techniques and coping strategies.

8.19.6.3 Skills Training

Relaxation training is an important aspect of skills training (Turk & Salovey, 1995). In spinal cord injury Jacobsonean relaxation (muscles tense/relaxed) training is contra-indicated and the focus should be on autogenic relaxation training (Schultz, 1932) and imaging techniques.

In addition, general problem-solving skills training, social skills training and assertiveness training have proved to be of benefit (Crewe & Krause, 1987; Dunn, Van Horn, & Herman, 1981; Kennedy, 1989). Dunn et al. (1981) effectively developed a program for enhancing social relationships and interpersonal skills. Components of this skills training program included general assertiveness training, disability and specific social skills training (such as reactions to wheelchairs, bladder and bowel accidents) and behavioral rehearsal of managing difficult social situations. Frank (1992) used a randomized control group design to demonstrate the effectiveness of a personal achievement skills (PAS) program. This was a structured group psychotherapeutic approach, the results of which indicated that the PAS group participants in comparison to the control group improved significantly on measures of self-esteem.

Moore (1989) investigated the effectiveness of including a family/significant other as an explicit component in the psychosocial rehabilitation of men with recent onset of spinal cord injuries. Using a two-way mixed analysis of covariance, significant interactions were found which confirmed that including family members or significant other in the psychosocial rehabilitation programme has a positive effect on personal adjustment as measured by the California Psychological Inventory and Personal Orientation Inventory.

8.19.6.4 Rehabilitation Planning and Adherence

Trieschmann and Willems (1980) conceptualize rehabilitation in behavioral terms and Norris-Baker, Stephens, Rintala, and Willems (1981) concluded that behavioral engagement in rehabilitation was the best protective of both medical and behavioral status post discharge. Kennedy, Fisher, and Pearson (1988) in an observational study of a spinal rehabilitation unit concluded that patients spent a considerable proportion of their time in solitary and disengaged behaviors. Patients who utilized behavioral disengagement as a coping strategy have been more likely to be depressed and anxious (Kennedy, 1995).

Goal planning is a behavioral intervention strategy which was developed by Houts and Scott (1975) to manage and minimize problems associated with institutionalization and dependency. The main objective of this intervention strategy is to increase patient involvement in decision making, identify needs and strengths rather than disabilities and problems, break down targets into manageable units and specify who will do what, where and when, and the degree of success expected. Locke and Latham (1990) found that strong positive intercorrelations existed between self-efficacy, goal planning, and performance. Kennedy, Walker, and White (1991) describe the effectiveness of goal planning in increasing patient engagement, specifically time spent in therapy and involvement in formal meetings.

Kennedy and Swalwall (1995) utilized the system which involves a key worker coordinating fortnightly goal planning meetings that involve members of the rehabilitation team and the patient in setting goals. They developed a Needs Assessment Checklist which comprises over 200 behavioral indicators of needs ranging from activities of daily living, bladder management, and discharge planning. Individual progress is compared with peers and a goal planning action plan is provided to patients. Complex needs are broken into specific behavioral objectives that can be managed within a short timespan (usually two weeks). This system provides an ongoing assessment of engagement and enables the more systematic and contingent delivery of positive social reinforcement. Many patients are initially overwhelmed by their needs and are unable to see progress in rehabilitation because of their focus on how they were before the injury. Rehabilitation programs are a powerful source of positive reinforcement and goal planning ensures that this occurs on a regular basis. For example, mobility needs are one of the major aspects of rehabilitation, independent transfer to and from a wheelchair is an important goal and weekly targets may begin by building up upper limb strength by attending weight training and lifting 5 kg 60 times. These objective milestones reflect to patients and staff goal attainment, serve as directional indicators of progress, and promote growth and self-efficacy.

8.19.6.5 Sexual Counseling

There is general concensus (Alexander, 1991; Kennedy, 1991) that the P-LI-SS-IT (permission, limited information, specific suggestions, intensive therapy) developed by Annon (1974) provides a useful framework for organizing sexual counseling. This model has four levels of intervention, each increasing in sophistication, which allows counselors to structure a program according to their level of competence. Permission giving highlights the legitimacy of continued sexual interest and the need to discuss sexual concerns. Limited information is provided dependant on the patient's ability and level of concern and should be made available to patients on request. Most spinal rehabilitation centers possess a number of handbooks and pamphlets on this issue. Specific suggestions include advice on maintaining an erection, positioning, developing relationships and possible problems.

People with more complex sexual difficulties may require more intensive therapy which may focus on sexual behavior, such as Masters and Johnson (1970), sensate focus or on communication. This needs to be provided by people trained in sexual counseling, but all staff should be encouraged to refer patients to suitable therapists. Sex education following spinal cord injury is essential and should be structured around the sexual response cycle and expected changes due to the spinal cord injury, as well as the effects of medication and management of problems such as bowel difficulties and spasm.

8.19.6.6 Chronic Pain Management

As mentioned, almost 40% of people have a problem with chronic pain following a spinal cord injury. Chronic pain behavior is a function of the interplay between organic and learning factors including avoidance, social discomfort, absolution of responsibilities, negative prediction, and expectation of pain. The chronic pain management approaches detailed in Chapter 8.24 of this volume are appropriately applied to this population. Programs should involve relaxation training, ability scheduling, challenging negative predictions and beliefs, distraction, and problem-solving techniques.

8.19.6.7 Biofeedback

Biofeedback techniques have been used in physical rehabilitation since the mid-1970s but their use in spinal cord injury has been extremely limited and has had mixed success (Brucker & Bulaeva, 1996). Biofeedback is an operant conditioning technique used to establish learned voluntary control of specific physiological responses. Until recently, the most significant evidence that individuals with complete spinal cord injuries can gain specific control of physiological responses was the work on postural hypotension and learned voluntary control of blood pressure (Brucker & Ince, 1977).

Autonomic disregulation occurs in high thoracic and cervical level lesions, and can cause blood to pool in the lower extremeties. In this study, blood pressure feedback over 11 sessions enabled a subject to increase his blood pressure voluntarily by 20 mmHg on command. Electromyograph (EMG) biofeedback has been applied to cervical patients to maximize upper limb function, Brudny et al. (1976). More recently Brucker and Bulaeva (1996) utilized EMG biofeedback procedures to increase voluntary EMG responses from the tricep muscles of over 100 persons with cervical spinal cord injuries C6 and above and greater than one year duration. These results demonstrated the efficacy of biofeedback for increasing voluntary EMG result responses in long term spinal cord injured patients (Figure 1).

Gallego, Perez de la Sota, Vardon, and Jaeger-Denavit (1993) used EMG biofeedback of the inspiratory muscles of the upper thorax with 10 tetraplegic patients. Their results supported the clinical use of the EMG biofeedback, increasing thoracic motion and reducing paradoxical breathing.

8.19.7 PROFESSIONAL ISSUES

Most spinal cord injury rehabilitation units in developed economies have resident clinical psychologists or sessional psychological input. The AASCIPSW (1992) established a taskforce to produce a set of standards for professional services. They recommend a minimum standard

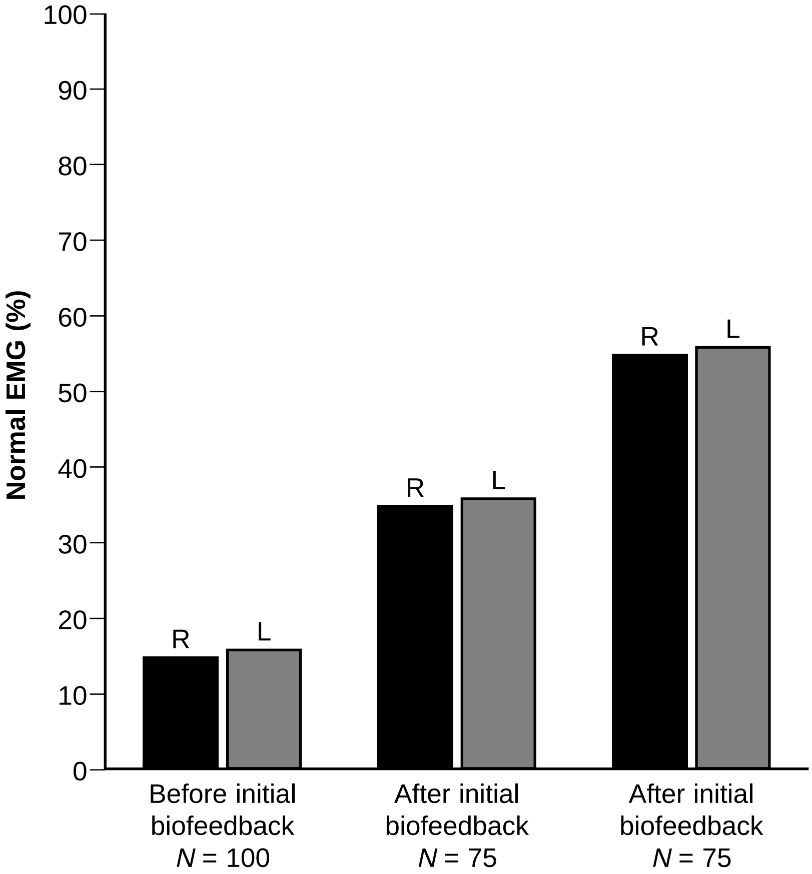

Figure 1 Percentage of normal EMG scores from the right and left triceps before initial biofeedback treatment and after initial and additional biofeedback treatments (Brucker & Bulaeva, 1996).

of 20 patients to one full-time psychologist. In this document they report on the psychologist's contribution to the psychosocial program that ranges from assessment, through treatment planning and follow-up. They also suggest psychological involvement in team consultation, research, resource dissemination, and in-service staff training. The author has been working with this population since 1984. Most of this time has been spent in direct clinical work (60%). The greatest demand for psychological services has come from patients in the rehabilitation phase but a small and significant proportion of referrals have been received in the acute phase and following discharge. Psychological issues are a core component of the rehabilitation phase. However, some people (estimated 10%) require referral to community psychological services for ongoing support. Adjustment disorder, anxiety and depression are the commonest needs requiring psychological support, but acquired cognitive impairment, psychiatric disorders and chronic pain are significant issues with a small proportion of clients.

8.19.8 LIFE SPAN AND SOCIETAL RESPONSES

In the 1960s, many people with severe spinal cord injuries were sent to residential homes following their rehabilitation. These were expensive in terms of resources, staff morale and individual psychological well-being. Now, with longevity only slightly compromised, quality of life issues have become paramount. Psychological interventions have helped individuals to manage emotional concerns, foster engagement in rehabilitation, and challenge some of the negative assumptions about acquired disability. The research previously discussed in this chapter, together with personal and clinical experience, have shown that most people make a positive adjustment and get on with their life goals. Nevertheless, a significant proportion continue to have difficulties and further research is required into identifying those at risk and exploring interventions to minimize their vulnerability. Whiteneck et al. (1992) in a study of over 20 years post injury, found that almost three-quarters of respondents rated their current quality of life as either good or excellent on a five-point scale, demonstrating a pattern of increasing quality of life up to the age of 50, and 30 years post injury. Clinical psychology services not only assist the individual in managing their adjustment concerns and acquired disability, but can enhance engagement in rehabilitation and contribute to the challenging of the myths and misconceptions that prevail in our communities about the impact of physical disability.

Once a decision has been made to prevent death of the severely injured person at the scene of a trauma, then it could be argued that healthcare providers have a responsibility to follow this decision through to enable access to a quality of life that is commensurate with previously held expectations. The challenge for clinical psychologists is to recognize the urgent need for new directions in research with an emphasis on treatment efficacy based on sound empirical methods. More attention needs to be given to health and rehabilitation outcomes in response to the growing competition for resources.

8.19.9 REFERENCES

Abrams, K. (1981). Impact on marriages of adult onset paraplegia. *Paraplegia, 19,* 253–259.

Alexander, C. J. (1991). Psychological assessment and treatment of sexual dysfunctions following spinal cord injury. *Journal of the American Paraplegia Society, 14,* 127–131.

Alexander, C. J., Sipski, M. L., & Findley, T. W. (1993). Sexual activities, desire and satisfaction in males pre- and post-spinal cord injury. *Archives of Sexual Behaviour, 22,* 217–228.

American Association of Spinal Cord Injury Psychologists and Social Workers (1992). *Standards (1992).* Jackson Heights, NY: Author.

Annon, J. S. (1974). *The behavioural treatment of sexual problems* (Vol. 1). Honolulu, HI: Enabling Systems.

Anson, C. A., Stanwyck, D. J., & Krause, J. S. (1993). Social support and health status in spinal cord injury. *Paraplegia, 31,* 632–638.

BBC (1966). *Man Alive Documentary: Paraplegics.* (Film) TX09-03-66 (Available from BBC, Wood Lane, London, UK).

Beck, A. T., & Steer, A. R. (1987). *Beck Depression Inventory manual.* San Antonio, TX: Psychological Corporation, & Harcourt, Brace and Jovanovich.

Bedbrook, G. (Ed.) (1981). *The care and management of spinal cord injuries.* Heidelberg, Germany: Springer-Verlag.

Boureston, N., & Howard, M. (1965). Personality characteristics of three disability groups. *Archives of Physical Medicine and Rehabilitation, 46,* 626–632.

Bracken, M., & Shepard, M. (1980). Coping and adaptation to acute spinal cord injury: A theoretical analysis. *Paraplegia, 18,* 74–85.

Bracken, M., Shepard, M., & Webb, S. (1981). Psychological response to acute spinal cord injury: An epidemiological study. *Paraplegia, 19,* 271–283.

Britell, W. W., & Mariano, A. J. (1991). Chronic pain in spinal cord injury. In N. E. Walsh (Ed.), *Physical medicine and rehabilitation; Rehabilitation of chronic pain* (pp. 71–82). Philadelphia: Hanley and Belfus.

Brucker, B. S., & Bulaeva, N. V. (1996). Biofeedback effect of electromyography responses in patients with spinal cord injury. *Archives of Physical Medicine and Rehabilitation, 77,* 133–137.

Brucker, B. S. & Ince, L. (1977). Biofeedback as an experimental treatment for postural hypotension in a patient with a spinal cord lesion. *Archives of Physical Medicine and Rehabilitation, 58,* 49–53.

Brudny, J., Korein, J., Grynbaum, B. B., Friedmann, L. W., Weinstein, S., Sachs-Frankel, G., & Belandres, P. (1976). EMG feedback therapy: A review of treatment of 114 patients. *Archives of Physical Medicine and Rehabilitation, 57,* 55–61.

Buckelew, S., & Hanson, S. (1992). Coping and adjustment following spinal cord injury. *Spinal Cord Injury Psychosocial Process, 5,* 99–103.

Bulman, R. J., & Wortman, C. B. (1977). Attributions of blame and coping in the "Real World'. *Journal of Personality and Social Psychology, 35,* 351–363.

Carver, C. S., Scheier, M. F., & Weintraub, J. K. (1989). Assessing coping strategies: A theoretically based approach. *Journal of Personality and Social Psychology, 56,* 267–283.

Charlifue, S. W., Coll, J., Frankell, H., Fraser, M., Gardner, B., Jamous, A., Krishnan, K., Menter, R. R., Nuseibeh, I., Stavic, G., & Whiteneck, G. (1996). *Changing patterns in causes of death in a population based sample of people with spinal cord injury.* Paper presented at the 35th Annual Scientific Meeting of the International Medical Society of Paraplegia, Atlanta, GA.

Charlifue, S. W., & Gerhart, K. A. (1991). Behavioural and demographic predictors of suicide after traumatic SCI. *Archives of Physical Medicine and Rehabilitation, 72,* 488–493.

Comarr, A. E. (1966). Observations of menstruation and pregnancy among female spinal cord injury patients. *Paraplegia, 3,* 263–272.

Cook, D. (1979). Psychological adjustment to spinal cord injury: Incidence of denial, depression and anxiety. *Rehabilitation Psychology, 26,* 97–104.

Creek, G., Moore N., Oliver, V., Silver, J., & Zarb, G. (1988). *Personal and social implications of spinal cord injury.* London: Thames Polytechnic.

Crewe, N. M., & Krause, J. S. (1992). Marital status and adjustment to spinal cord injury. *Journal of American Paraplegia Society, 15,* 14–18.

Crewe, N. M., & Krause, J. S. (1988). Marital relationships and spinal cord injury. *Archives of Physical Medicine and Rehabilitation, 69,* 435–438.

Crewe, N. M., & Krause, J. S. (1987). Spinal cord injury: Psychological aspects. In B. Kaplan (Ed.), *Rehabilitation psychology* (pp. 3–35). Baltimore: Aspen.

Davidoff, G. N., Roth, E. J., & Richards, J. S. (1992). Cognitive deficits in spinal cord injury: Epidemiology and outcome. *Archives of Physical Medicine and Rehabilitation, 73,* 275–284.

Davidoff, G., Thomas, P., Johnson, M. Berent, S., Dijkers, M., & Doljonac, R. (1988). Closed head injury in acute traumatic spinal cord injury: Incidents and risk factors. *Archives of Physical Medicine and Rehabilitation, 69,* 869–872.

Decker, S. D., & Schulz, R. (1985). Correlates of life satisfaction and depression in middle aged and elderly spinal cord injured persons. *American Journal of Occupational Therapy, 39,* 740–745.

Derogatis, L. R. (1977). *SCL-90-R administration scoring and procedures manual.* Towson, MD: Clinical Psychometric Research.

DeVivo, M. J., Black, K. J., Richards, J. S., & Stover, S. L. (1991). Suicide following spinal cord injury. *Paraplegia, 29,* 620–627.

DeVivo, M. J., & Fine, P. R. (1985). Spinal cord injury: Its short-term impact on marital status. *Archives of Physical Medicine and Rehabilitation, 66,* 501–504.

DeVivo, M. J., & Stover, S. L. (1995). Long-term survival and causes of death. In S. L. Stover, J. A. DeLisa, & G. G. Whiteneck (Eds.) *Spinal cord injury: Clinical outcomes from the model systems* (pp. 289–315). Baltimore: Aspen.

DeVivo, M. J., Whiteneck, G. G., & Charles, E. D. (1995). The economic impact of spinal cord injury. In S. L. Stover, J. A. DeLisa, & G. G. Whiteneck (Eds.) *Spinal cord injury: Clinical outcomes from the model systems* (pp. 234–270). Baltimore: Aspen.

Dijkers, M. P., Abela, N. B., Gans, B. M., & Gordon, W. A. (1995). The aftermath of spinal cord injury. In S. L. Stover, J. A. DeLisa, & G. G. Whiteneck (Eds.), *Spinal cord injury: Clinical outcomes from the model systems* (pp. 185–211). Baltimore: Aspen.

Dijkers, M., & Abela, M. (1993). Formal education after spinal cord injury. *Archives of Physical Medicine and Rehabilitation, 74,* 673.

Dunn, M., Van Horn, E., & Herman, S. H. (1981). Social skills and spinal cord injury: A comparison of three training procedures. *Behaviour Therapy, 12,* 153–164.

Felton, B. J., & Revenson, T. A. (1984). Coping with chronic illness: A study of illness controllability and the influence of coping strategies on psychological adjustment. *Journal of Consulting and Clinical Psychology, 52,* 343–353.

Folkman, S., & Lazarus, R. S. (1988). *Way of Coping Questionnaire—research edition*. Palo Alto, CA: Consulting Psychologists Press.

Folkman, S., Chesney, M., McKusick, L., Ironson, G., Johnson, D. S., & Coates, T. J. (1991). Translating coping theory into an intervention. In J. Eckenrode (Ed.) *The social context of coping* (pp. 239–260). New York: Plenum.

Fordyce, W. (1964). Personality characteristics of men with spinal cord injury as related to manner of onset of disability. *Archives of Physical Medicine and Rehabilitation, 45,* 321–325.

Frank, R. A. (1992). *Structured group psychotherapy for individuals with spinal cord injury*. Unpublished PhD thesis, Oregon State University.

Frank, R. G., Umlauf, R. L., Wonderlich, S. A., Askanazi, G. S., Buckelew, S., & Elliott, T. (1987). Difference in coping styles among persons with spinal cord injury: A cluster analytic approach. *Journal of Consulting and Clinical Psychology, 55,* 727–731.

Fuhrer, M. J., Garber, S. L., Rintala, D. H., Clearman, R., & Hart, K. A. (1993). Pressure ulcers in community resident persons with spinal cord injury: Prevalence and risk factors. *Archives of Physical Medicine and Rehabilitation, 74,* 1172–1177.

Fullerton, D. T., Harvey, R. F., Klein, M. H. & Howell, T. (1981). Psychiatric disorders in patients with spinal cord injuries. *Archives of General Psychiatry, 38,* 1369–1371.

Gallego, J., Perez de la Sota, A., Vardon, G., & Jaeger-Denavit, O. (1993) Learned activation of thoracic inspiratory muscles in tetraplegics. *American Journal of Physical Medicine and Rehabilitation, 72,* 312–317.

Go, B. K., DeVivo, N. J., & Richards, J. S. (1995). The epidemiology of spinal cord injury. In S. L. Stover, J. A. DeLisa, & G. G. Whiteneck, (Eds.), *Spinal cord injury: Clinical outcomes from the model systems.* Baltimore: Aspen.

Guttmann, L. (1964). Problems in the initial treatment of traumatic tetraplegia. *Excerpta Medica,* 593–596.

Guttmann, L. (1976). *Spinal cord injuries: Comprehensive management and Research* (2nd ed.). Oxford, UK: Blackwell Scientific.

Hancock, K. M., Craig, A. R., Dickson, H. G., Chang, E., & Martin, J. (1993). Anxiety and depression over the first year of spinal cord injury: A longitudinal study. *Paraplegia, 31,* 349–357.

Heinemann, A. W., Keen, M., Donohue, R., & Schnoll, S. (1988). Alcohol use by persons with recent spinal cord injuries. *Archives of Physical Medicine and Rehabilitation, 69,* 619–624.

Hohmann, G. (1975). Psychological aspects of treatment and rehabilitation of the spinal injured person. *Clinical Orthopaedics, 112,* 81–88.

Houts, P., & Scott, R. (1975). *Goal planning with mentally disabled persons: Procedures for developing an individualised client plan.* Pennsylvania: Milton Hershey Center.

Hughes, J. T. (1988). The Edwin Smith surgical papyrus: An analysis of the first case reports of spinal cord injuries. *Paraplegia, 26,* 71–82.

Ivie, C. S., & DeVivo, M. J. (1994). Predicting unplanned hospitalisations in persons with spinal cord injury. *Archives of Pysical Medicine and Rehabilitation, 75,* 1182–1188.

Judd, F. K., Brown, D. J., & Burrows, G. D. (1991). Depression, disease and disability: Application to patients with traumatic spinal cord injury. *Paraplegia, 29,* 91–96.

Judd, F. K., Stone, J., Webber, J. E., Brown, D. J., & Burrows, G. D. (1989). Depression following spinal cord injury: A prospective in-patient study. *British Journal of Psychiatry, 154,* 668–671.

Kelly, M. P., & Doty, R. E. (1995). Neuropsychological dysfunction: Research and evaluation. In P. M. Nicassio, & T. W. Smith (Eds.), *Managing chronic illness: A biopsychosocial prospective* (pp. 117–162). Washington, DC: American Psychological Association.

Kendall, P. C., Vitousek, K. B., & Kane, M. (1991). Thought and action in psychotherapy: Cognitive-behavioral approaches. In M. Hersen, A. Cazdin, & A. Bellack (Eds.). *The clinical psychology handbook* (pp. 596–626). New York: Pergamon.

Kennedy, P. (1989). Psychological approaches in the management of spinal cord injury. *British Journal of Holistic Medicine, 4,* 169–176.

Kennedy, P. (1991). Counselling with spinal cord injured people. In H. Davis & L. Fallowfield (Eds.). *Counselling and communication in healthcare* (pp. 129–145). Chichester, UK: Wiley.

Kennedy, P. (1995). *Psychological aspects of spinal cord injury: Behavioral approaches, emotion and impact in coping strategies*, Volume 2. DPhil Thesis, University of Ulster.

Kennedy, P., Fisher, K., & Pearson, E. (1988). Ecological evaluation of a rehabilitative environment for spinal cord injured people: Behavioural mapping and feedback. *British Journal of Clinical Psychology, 27,* 239–246.

Kennedy, P., Frankel, H., Gardner, B., & Nuseibeh, I. (1997). Factors associated with acute and chronic pain following traumatic spinal cord injuries. *Spinal Cord, 35,* 814–817.

Kennedy, P., Gorsuch, N., & Marsh, N. (1995). Childhood onset of spinal cord injury: Self-esteem and self-perception. *British Journal of Clinical Psychology, 34,* 581–588.

Kennedy, P., Lowe, R., Grey, N., & Short, E. (1995). Traumatic spinal cord injury and psychological impact: A cross sectional analysis of coping strategies. *British Journal of Clinical Psychology, 34,* 627–639.

Kennedy, P., & Swalwell, E. (1995). Clinical audit of physical rehabilitation. *Auditorium, 4,* 25–28.

Kennedy, P., Walker, L., & White, D. (1991). Ecological evaluation of goal planning and advocacy in a rehabilitative environment for spinal cord injured people. *Paraplegia, 29*, 197–202.

Krause, J. S., & Crewe, N. M. (1990). Long-term prediction of self-reported problems following spinal cord injury. *Paraplegia, 28,* 186–202.

Lawson, N. (1978). Significant events in the rehabilitation process: The spinal cord patients point of view. *Archives of Physical Medicine and Rehabilitation, 59,* 573–597.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.

Linkowski, D. (1971). A scale to measure the acceptance of disability. *Rehabilitation Counselling Bulletin, 14,* 236–244.

Linsenmeyer, T. A., & Perkash, I. (1991). Infertility in men with spinal cord injury. *Archives of Physical Medicine and Rehabilitation, 72,* 747–754.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Inglewood Cliffs, NJ: Prentice-Hall.

Masters, W. H., & Johnson, V. E. (1970). *Human sexual inadequacy*. Boston: Little, Brown and Company.

McAweeney, M. J. (1996). *Positive effects of psychosocial interventions in the rehabilitation of persons with spinal cord injury*. Paper presented at a 10th Annual Conference of the American Association of Spinal Cord Injury Psychologists and Social Workers, Las Vegas, NV.

Menter, R. R., & Hudson, L. M. (1995). Effects of age at injury and the aging process. In S. L. Stover, J. A. DeLisa, & G. G. Whiteneck (Eds.), *Spinal cord injury: Clinical outcomes from the model systems* (pp. 272–287). Baltimore: Aspen.

Moore, L. I. (1989). *Behavioural changes in male spinal cord injury following two types of psychological rehabilitation*. Unpublished PhD thesis, St Louis University.

Morris, J. (1992). Psychological and sociological aspects of patients with spinal cord injuries. In H. L. Frankel (Ed.), *Handbook of clinical neurology: Spinal cord trauma* (pp. 537–555). London: Elsevier.

Morris, J., Roth, E., & Davidoff, G. (1986). Mild closed head injury and cognitive deficits in spinal cord injured patients: Incidence and impact. *Journal of Head Trauma Rehabilitation, 1,* 31–42.

Nagler, B. (1950). Psychiatric aspects of spinal cord injury. *The American Journal of Psychiatry, 107,* 49–56.

Nehemkis, A., & Groot, H. (1980). Indirect self-destructive behavior in spinal cord injury. In N. Farberon (Ed.), *The many faces of suicide* (pp. 99–115). New York: McGraw-Hill.

Norris-Baker, C., Stephens, M. A. P., Rintala, D. H., & Willems, E. P. (1981). Patient behaviour as a predictor of outcomes in spinal cord injury. *Archives of Physical Medicine and Rehabilitation, 62,* 602–608.

Padrone, F. J. (1994). Psychotherapeutic issues with family members of persons with physical disabilities. *American Journal of Psychotherapy, 48,* 195–207.

Reidy, K., Caplan, B., & Shawaryn, M. (1991). *Coping strategies following spinal cord injury: Accommodation to trauma and disability*. Paper presented at the 68th Annual Meeting of the American Congress of Rehabilitation Medicine, Washington, DC.

Richards, J. S. (1986). Psychological adjustment to spinal cord injury during the first post-discharge year. *Archives of Physical Medicine and Rehabilitation, 67,* 362–365.

Rigoni, H. (1977). *Psychological coping in the patient with spinal cord injury*. Boston: Little and Brown.

Rose, M., Robinson, J. E., Ells, P., & Cole, J. D. (1988). Pain following spinal cord injury: Results from a postal survey. *Pain, 34,* 101–102.

Sarason, I. G., & Sarason, B. R. (1986). Experimentally provided social support. *Journal of Personality and Social Psychology, 50,* 1222–1225.

Schultz, I. H. (1932). *Das Autogene Training*. Stuttgart, Germany: Thieme.

Seekins, T., Clay, J., & Ravesloot, C. (1994). A descriptive study of secondary conditions reported by a population of adults with physical disabilities served by three independent living centers in a rural state. *Journal of Rehabilitation, 2,* 47–51.

Sharpe, M., Hawton, K., Simkin, S., Surawy, C., Hackmann, A., Klimes, I., Peto, T., Warrell, D., & Seagroagt, T. V. (1996). Cognitive behavior therapy for the chronic fatigue syndrome: A randomised controlled trial. *British Medical Journal, 312,* 22–25.

Shadish, W. R., Hickman, D., & Arrick, M. C. (1981). Psychological problems of spinal cord injury patients: Emotional distress as a function of time and locus of control. *Journal of Consulting and Clinical Psychology, 49,* 297.

Siller, J. (1969). Psychological situation of the disabled with spinal cord injuries. *Rehabilitation Literature, 30,* 290–296.

Siosteen, A., Lundqvist, C., Blomstrand, C., & Sullivan, L., & Sullivan, M. (1990). Sexual ability, activity, attitudes and satisfaction as part of adjustment in spinal cord injured subjects. *Paraplegia, 28,* 285–295.

Sipski, M. L., & Alexander, C. J. (1992). Sexual function and dysfunction after spinal cord injury. *Physical Medicine and Rehabilitation Clinics of North America, 3,* 811–828.

Sipski, M. L., Alexander, C. J., & Rosen, R. C. (1995). Physiological parameters associated with psychogenic sexual arousal in women with complete spinal cord injuries. *Archives of Physical Medicine and Rehabilitation, 76,* 811–818.

Speckens, A. E., Van-Hemert, A. M., Spinhoven, P., Hawton, K. E., Bolk, J. H., & Rooijmans, H. G. (1995). Cognitive behavioural therapy for medically unexplained physical symptoms: A randomised controlled trial. *British Medical Journal, 311,* 1328–1332.

Spielberger, C., Gorsuch, R., Lushene, R., Vagg, P., & Jacobs, G. (1983). *Manual for the State Trait Anxiety Inventory*. Palo Alto, CA: Consultant Psychologists Press.

Stewart, T. (1977). Coping behaviour and the moratorium following spinal cord injury. *Paraplegia, 15,* 338–342.

Stover, S., & Fine, R. (1986). *Spinal cord injury: The facts and figures*. Birmingham, AL: University of Alabama.

Swain, J., Finkelstein, V., French, S., & Oliver, M. (1993). *Disabling barriers—enabling environments*. London: Sage.

Taylor, G. (1967) *Predicted vs. actual responses to spinal cord injury: A psychological study*. Unpublished doctoral dissertation, University of Minnesota, Minneapolis.

Taylor, S. E., & Aspinwall, L. G. (1990). Psychosocial aspects of chronic illness. In P. T. Costa, G. R. Vandenbos, & G. H. Herek (Eds.), *Psychological aspects of serious-illness: Chronic conditions, fatal diseases, and clinical care* (pp. 3–60). Washington, DC: American Psychological Association.

Trieschmann, R. B. (1988) *Spinal cord injuries: Psychological, social and vocational rehabilitation* (2nd ed.). New York: Demos.

Trieschmann, R., & Willems, E. P. (1980). Behavioral programs for the physically disabled. In D. Glenwick & L. Jason (Eds.) *Behavioral community psychology: Progress and prospects* (pp. 45–57). New York: Praeger.

Turk, D. C., & Salovey, P. (1995). Cognitive-behavioral treatment of illness behavior. In P. M. Nicassio & T. W. Smith (Eds.) *Managing chronic illness—a biopsychosocial perspective* (pp. 245–285). Washington, DC: American Psychological Association.

Wallston, K. A., Wallston, B. S., & DeVellis, R. (1978). Development of the multidimensional-health locus of control scales. *Health Education Monographs, 6,* 160–170.

Wanlass, R. L., Palmer, D., & Cook, D. (1992). Bedside manner for rehabilitation psychologists. *SCI Psychosocial Process, 5,* 9–10.

Weisman, M. M., Myers, J. K., & Harding, P. S. (1978). Psychiatric disorders in a US urban community. *American Journal of Psychiatry, 135,* 459–462.

Whiteneck, G. G., Charlifue, S. W., Frankel, H. L., Fraser, B. M., Gardner, B. P., Gerhart, M., Krishnan, K., Menter, R. R., Nuseibeh, I., Short, D., & Silver, J. (1992). Mortality, morbidity and psychosocial outcomes of persons spinal cord injured more than 20 years ago. *Paraplegia, 30,* 617–630.

Wilcox, N., & Stauffer, E. (1972). Follow-up of 423 consecutive patients admitted to the spinal cord centre. *International Journal of Paraplegia. 10,* 115–122.

Wilson, C., Kennedy, P., & Glass, C. (1994) Meeting the psychological needs of people with traumatic spinal cord injury. *Clinical Psychology Forum, 71,* 17–21.

Woodrich, F., & Patterson, J. (1983). Variables related to acceptance of disability in persons with spinal cord injuries. *Journal of Rehabilitation, 49,* 26–30.

Wortman, C. B., & Silver, R. C. (1989). The myths of coping with loss. *Journal of Consulting and Clinical Psychology, 57,* 349–357.

Yarkony, G. M., & Heinemann, A. W. (1995). Pressure ulcers. In S. L. Stover, J. A. DeLisa, & G. G. Whiteneck (Eds.), *Spinal cord injury: Clinical outcomes from the modelled systems* (pp. 100–118). Baltimore: Aspen.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.20
Traumatic Brain Injury

BARBARA A. WILSON
Medical Research Council—Applied Psychology Unit, Cambridge, UK

8.20.1 INTRODUCTION 464

8.20.2 THE NATURE OF TBI 465
  8.20.2.1 Primary and Secondary Brain Damage 465
  8.20.2.2 How Common are Head Injuries? 465
  8.20.2.3 Recovery from TBI 465
  8.20.2.4 Who is at Risk of TBI? 466
  8.20.2.5 The Impact of TBI on the Family 467

8.20.3 PROBLEMS FACED BY PEOPLE WITH TBI 468
  8.20.3.1 Motor Problems 468
  8.20.3.2 Sensory Problems 469
  8.20.3.3 Emotional and Social Problems 469
  8.20.3.4 Behavior Problems 469
  8.20.3.5 Cognitive Problems 470

8.20.4 ASSESSMENT OF PEOPLE WITH TBI 470
  8.20.4.1 Models of Assessment 470
  8.20.4.2 Behavioral Assessment 471

8.20.5 PRINCIPLES AND PRACTICES OF REHABILITATION 473
  8.20.5.1 Restoration of Function 473
  8.20.5.2 Environmental Restructuring 474
  8.20.5.3 Anatomical Reorganization 474
  8.20.5.4 Using Residual Skills More Efficiently 474
  8.20.5.5 Alternative Solutions 474

8.20.6 COGNITIVE REHABILITATION 476
  8.20.6.1 Cognitive Retraining 476
  8.20.6.2 Cognitive Neuropsychology 477
  8.20.6.3 Combined Approaches 479
  8.20.6.4 The Holistic Approach 480

8.20.7 EFFECTIVENESS OF REHABILITATION 480
  8.20.7.1 Clinical Effectiveness 480
  8.20.7.2 Does Rehabilitation Improve Quality of Life? 481
  8.20.7.3 Overall View of Effectiveness 481

8.20.8 PROFESIONAL ISSUES 481

8.20.9 CONCLUSIONS: AN INTEGRATED APPROACH TO COGNITIVE REHABILITATION 481

8.20.10 REFERENCES 482

8.20.1 INTRODUCTION

Traumatic brain injury (TBI) has been defined as “an insult to the brain caused by an external force that may produce diminished or altered states of consciousness, which results in impaired cognitive abilities or physical functioning” (National Head Injury Foundation, 1988). The two main types of TBI are closed head injury and penetrating wounds. Closed head injury (CHI) is by far the most common. It is associated with acceleration and deceleration forces that may or may not involve skull fracture.

The earliest known scientific document about head injury is believed to be the Edwin Smith Surgical Papyrus found by Smith in Luxor in 1862 (Walsh, 1987). The papyrus is between 2500 and 3000 years old and contains 48 case descriptions of treatment. Eight of these cases refer to head injuries, and it is apparent that the physician recognized that head injury could affect other parts of the body. Evans (1981), discussing the same papyrus, said, “Bleeding from the nose or ears with shudderings (convulsions) are described, together with eye deformity (deviation), shuffling gait and one who does not release the shoulder fork (hemiplegic) posture of the arm” (p. 27).

Severity of head injury is usually judged by (i) the depth and duration of coma, and (ii) the length of post-traumatic amnesia (PTA). Coma has been defined as “not obeying commands, not uttering words, and not opening the eyes” (Jennett & Teasdale, 1981). PTA has been described as “a period of variable length following closed head trauma during which the patient is confused, disorientated, suffers from retrograde amnesia, and seems to lack the capacity to store and retrieve new information” (Schacter & Crovitz, 1977).

The Glasgow Coma Scale (GCS; Teasdale & Jennett, 1974) is widely used throughout the world to measure the depth and duration of coma. The GCS consists of three sections involving eye opening, verbal, and motor responses (see Table 1). The original 14-point scale of the GCS, later extended to 15 points, provides an uncomplicated and objective measure of coma that is easy to score. Within each section the best possible response is scored 4, 5, or 6, depending on the component being measured and the version of the test employed. The worst possible score for each section is 1, thus even patients who are brain dead score 3. The operational definition often used is that a combined score of 8 or less means the patient is in coma, and 9 or more means the patient is out of coma.

Table 1 Glasgow Coma Scale.

Eyes	4	=	Open Spontaneously
	3	=	To verbal command
	2	=	To pain
	1	=	No response
Motor	6	=	Obeys
	5	=	Localizes pain
	4	=	Flexion—withdrawal
	3	=	Flexion—abnormal
	2	=	Extension
	1	=	No response
Verbal	5	=	Oriented + converses
	4	=	Disoriented + converses
	3	=	Inappropriate words
	2	=	Incomprehensible sounds
	1	=	No response
Total 3–15			

Source: Teasdale and Jennett (1974).

Although there are some difficulties with this definition (see Watson, Horn, & Curl, 1992; Ponsford, Snow, & Sloan, 1995; Wilson, Shiel, Watson, Horn, & McLellan, 1994 for further discussion), the scale has proved useful in predicting outcome (Klonoff, Costa, & Snow, 1986; Ponsford, Olver, Curran, & Ng, 1995), and enables comparisons to be made from one center to another. The worse the GCS score, the greater the depth of coma and the worse the predicted outcome. Thus a patient with a GCS score of 4 on admission has a poorer prognosis than someone with a GCS of 7 on admission. Severe head injury is, however, often defined as a GCS of 8 or less for at least six hours and a very severe head injury is defined as a GCS score of 8 or less for at least 24 hours. Moderate head injury may be defined as a GCS of 9–12 or a score of 8 or less for less than 6 hours. Mild head injury is a GCS of 13–15 or a coma of less than half an hour. Snow and Ponsford (1995) point out, however, that GCS scores by themselves account for less than a third of the variance in outcome studies.

Assessment of PTA is believed to be an even better predictor of outcome than depth and duration of coma (Evans, 1981; Teasdale & Mendelow, 1984). Jennett (1990) classifies PTA of less than 5 minutes as a very mild head injury; 5–60 minutes as mild; 1–14 hours as moderate; 1–7 days as severe; 1–4 weeks as very severe; and more than 4 weeks as extremely severe. Although Jennett says, “there is no need to seek a very accurate figure for PTA; what matters is whether it lasted minutes, hours, days or weeks” (p. 4), there are occasions when it would appear useful to have a formal objective measure of PTA. For example, some patients are considered suitable for rehabilitation if they are "out of PTA." Other patients remain with persistent memory problems (one of the characteristics of PTA) even when they are no longer confused and disoriented (another characteristic of PTA). Wilson, Baddeley, Shiel, and Patton (1992) attempted to distinguish between PTA and other kinds of memory problems. They compared patients in PTA with (i) patients showing a pure amnesic syndrome, (ii) people with long-standing memory difficulties, and (iii) patients with orthopedic injuries. Although all three of the neurologically impaired groups had difficulty with memory tasks, those patients in PTA were slower on a reaction time task, had poorer verbal fluency, and made more errors on a true/false decision task. Wilson, Evans, Baddeley, Emslie, and Watson (1997) also showed that tests useful for differentiating between groups are not identical to those useful for monitoring recovery.

8.20.2 THE NATURE OF TBI

8.20.2.1 Primary and Secondary Brain Damage

Brain damage resulting from traumatic head injury can be due to both primary and secondary causes. Primary damage is that caused directly by the accident or assault, whereas secondary damage is due to complications arising from the first insult (Rose & Johnson, 1996). In the words of Miller, Pentland, and Berrol (1990), "the final outcome in any patient who suffers head injury is governed by three groups of factors: the preinjury status of the brain, the total amount of immediate damage done to the brain by the impact of the head injury (primary damage), and the cumulative effect of secondary pathological damage to the already injured brain" (p. 21).

Once an accident has occurred, the primary damage is there and has to be managed as well as possible with the resources available. The secondary damage, however, which may result in more permanent disability than the primary, is (at least potentially) preventable. Table 2 illustrates the possible causes of primary and secondary damage (Teasdale & Mendelow, 1984). For further discussion see Miller et al. (1990) and Ponsford (1995).

8.20.2.2 How Common are Head Injuries?

There is little doubt that head injuries are common. Jennett (1990) suggests that every year in the USA there are about 7 million head injuries with about 500 000 people admitted to hospital. The death rate in the USA is 22–25 per 100 000, more than twice that seen in Britain, Sweden, or Japan (Jennett, 1990). More than two-thirds of people sustaining TBI are under 30 years of age, with over twice as many males as females involved. McMillan and Greenwood (1993) suggest that of those admitted to hospital, 5–7% are considered to have a severe head injury and 10% a moderate head injury.

The prevalence of severe head injury is difficult to calculate. Jennett (1990) suggests a figure of 150 markedly disabled persons per 100 000 of the population in the UK, and approximately 400 per 100 000 in the USA. Wade and Langdon Hewer (1987) suggest a prevalence rate of 300 per 100 000 disabled survivors of head injury. This translates to some 174 000 people alive in the UK with a major disability caused by head injury.

8.20.2.3 Recovery from TBI

The process of recovery after TBI is not well understood and probably involves different biological processes (Ponsford, 1990, 1995). Changes in the first few minutes, as seen in those with mild head injury, presumably reflect the resolution of temporary dysfunction without accompanying structural damage. Recovery after several days is more likely to be due to temporary structural abnormalities such as oedema and vascular disruption (Jennett, 1990). Recovery after months or years is even less well understood. Finger and Stein (1982) suggest several ways this might be achieved including regeneration, diaschisis, and plasticity.

Regeneration in the central nervous system (CNS) may occur even in adults (Kolb, 1995), but there is great uncertainty as to the extent such regeneration leads to clinical and behavioral functional gains. Diaschisis is a term coined by von Monakow (1914, translated by Pribham, 1969). It assumes that damage to a specific area of the brain can result in neural shock or disruption elsewhere in the brain. The secondary neural shock can be adjacent to the site of the primary insult or much further away (Miller, 1984). In any case, the shock follows a particular neural route. Similar, although not identical, is Luria's (1963) theory of inhibition. In inhibition, however, the shock is more diffuse and affects the brain as a whole. Miller (1984) found the evidence for diaschisis unconvincing.

Plasticity implies anatomical reorganization based on the idea that undamaged areas of the brain can take on the skills or functions subserved by a damaged area (Wilson, 1989b). Although controversial as an explanation of recovery in adults, such plasticity appears to be widely accepted by the general

Table 2 Causes of primary and secondary brain damage after severe traumatic brain injury.

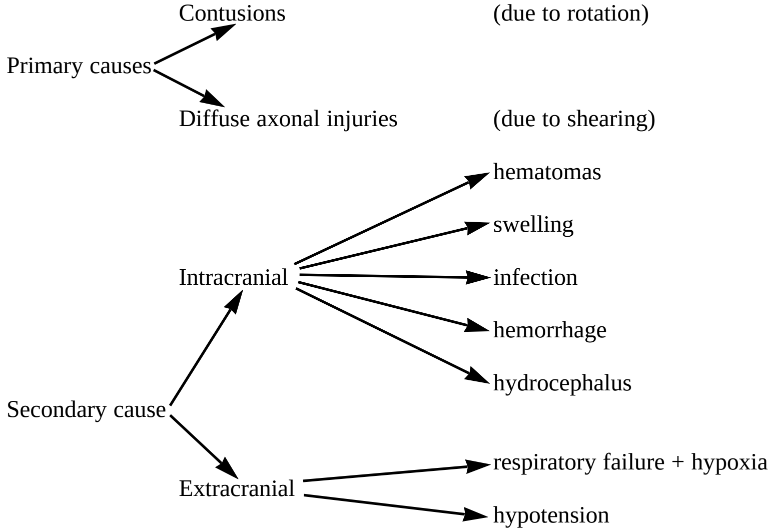

Source: Teasdale and Mendelow (1984)

public. MccGwire (1986), for example, writing of her partner who had survived a severe head injury, wrote, "We make use of only 10 per cent of our brains . . . " (p. 85). There is no doubt that cerebral plasticity occurs in infants but the extent to which it occurs in adults is unclear. Miller (1984) is somewhat skeptical, although Kolb (1995) feels the skepticism seen among neuroscientists is unfounded and argues that there is increasing evidence for such plasticity. Much of the late "recovery," however, is probably due to functional adaptation (Luria, Naydin, Tsvetkova, & Vinarskaya, 1969) whereby alternative solutions or compensatory strategies are employed to overcome problems and difficulties (Ponsford, Snow, & Sloan, 1995; Wilson, 1989b; Wilson & Patterson, 1990; Wilson & Watson, 1996).

Jennett (1990) says, "Whether recovery is due to restoration of activity in structures that are recovering or to diversion to other pathways, it seems likely that some of the restoration of function is essentially a learning process" (p. 6). The topic of learning and compensation will be discussed later when we consider rehabilitation.

8.20.2.4 Who is at Risk of TBI?

A famous quotation by Symonds (1937) states that "It's not only the kind of head injury that matters but the kind of head" (p. 1092). As noted earlier, young men between the ages of 15 and 24 years are most at risk of TBI. Approximately 70% of TBIs result from motor vehicle accidents (Ponsford, 1995). Cycling accidents, falls, and sports injuries account for most of the others. Of 1000 patients with severe head injury in a Scottish study (Bryden, 1989), 58% resulted from road accidents, 7% from assaults, 16% from falls, and 7% from work-related accidents. Falls account for a greater proportion of head injuries in children and elderly people than they do in young adults (Rimel, Jane, & Bond, 1990). A questionnaire sent to members of the National Head Injuries Association (Headway) reported by Talbott (1989) showed that car accidents accounted for the largest number of accidents and assaults the smallest number. These results are shown in Table 3.

Rimel et al. (1990) suggest that TBI from falls and assaults (but not from traffic accidents) is more likely to occur in lower socioeconomic groups, and a greater than average proportion have pre-existing problems such as a history of psychopathology, substance abuse, heavy alcohol consumption, and poor academic performance (Ponsford, 1995).

Two of the most successful measures of preventing TBI are using seat belts in cars and wearing a helmet if you are a motor cyclist. Rimel et al. (1990) reported that a study carried out in Virginia showed less than 14% of people involved in car accidents reported wearing seat belts at the time of injury, and only 58% of motor cyclists were wearing helmets. This was despite state law requiring the use of belts and helmets. The authors went on to show that "In Britain where there is a compliance rate of well

Table 3 A questionnaire study completed by 132 members of the National Head Injury Association. Causes of injury resulting in brain damage.

Cause of injury	*Number of patients* (%)
Car (driver or passenger)	54 (40.9)
Motorcycle (driver or passenger)	20 (15.15)
Cycle (rider)	5 (3.78)
Hit by vehicle	15 (11.36)
Assault	3 (2.27)
Industrial accident	5 (3.78)
Sport	4 (3.03)
Home	9 (6.81)
Other	17 (12.87)
Total	132 (100)

Source: Talbott (1989).

over 90% for several years ... (there) has been a significant decline in the number of head and facial injuries in car drivers and passengers and also in head injuries among motorcyclists" (p. 11).

Bryden (1989) also discusses the seat belt law (introduced into the UK in 1983) suggesting, however, that this has shifted the pattern of severity of injuries from (i) death to survival with major injury, (ii) major injury to minor injury, and (ii) minor injury to no injury.

Barnes (1996) confirms the value of seat belts and safety helmets, saying that there is good evidence these have reduced the number of fatalities and severe injuries. He goes even further when he argues for legislation enforcing the wearing of helmets for bicycle riders. He quotes two studies showing that bicycle riders had a significant reduction in the risk of head injury when wearing protective head gear. This reduction was as much as 63% in one study (Thomas et al., 1994) and 85% in another (Thompson, Rivara, & Thompson, 1989).

Drink driving regulations in the UK have significantly reduced head injuries and fatalities (Dunbar, Penttila, & Pikkarainen, 1987), although alcohol would still appear to be heavily involved in TBI. Kraus, Morgenstern, Fife, Conroy, and Nourjah (1989), for example, reported that over 60% of people blood tested for alcohol had a high level ($\geqslant 100$ mg%) of blood alcohol.

Barnes (1996) discusses other ways TBI can be reduced. He makes a strong argument for increased use of public transport, claiming that a journey made by car is three times more likely to injure a pedestrian than a journey made by bus. He also quotes a study by Allsop and Turner (1986), demonstrating that the removal of subsidies to public transport in London resulted in increased road casualties. Safety measures, air bags, car design, home, work and leisure environments, and education are other areas where changes could result in fewer head injuries.

8.20.2.5 The Impact of TBI on the Family

Families typically receive news of a TBI through a knock on the door by a police officer. They are given news of the road traffic accident and taken to the hospital where their son or daughter, wife or husband is in intensive care. If they are lucky there will be a member of staff to talk to them, to be kind, considerate, and able to provide explanations and information. Sometimes families will be left to their own devices in an intensive ward that will appear frenetic and confusing; the head injured family member will probably be attached to unfamiliar equipment with perhaps a catheter, a tracheostomy tube, and shaved head. Add to this, severe bruising of the face of the injured person, swollen, black, and tightly closed eyes, possibly bleeding and broken limbs encased in splints, it is not hard to see how terrifying the sight is to most families.

In the first few days relatives are typically concerned with life and death. Their main desire is to see the loved one survive. The greatest distress is seen at this stage (Oddy, 1995). Stress may, however, remain at a high level indefinitely (Oddy).

After the early stages, relatives are typically more concerned with the physical sequelae than with the cognitive or emotional sequelae. Full recovery may still be expected. It is only in the later stages, perhaps two, three, or even more years later that families realize the full extent of the cognitive and emotional problems, and are able to accept that the injured person will not get back to his or her old self. Oddy's (1995) chapter goes into this discussion in more detail.

Several studies have noted that it is the cognitive, social, and behavioral sequelae that ultimately cause families most distress and not the physical problems. Condeluci (1985) referred to the former group as the most disabling consequences of head injury. Oddy (1995) states that "Changes in cognition and emotional responsiveness may alter the core characteristics of the person. Many studies have suggested that it is these changes in personality that cause relatives most distress" (p. 167). An Israeli study by Rosenbaum and Najenson (1976) found that wives of head injured soldiers were more distressed than wives of soldiers with spinal injury. Brooks and McKinlay (1983) also found that distress in relatives was likely to increase over time. Wearing (1992) provides numerous examples of stress on families. She says that although the vast majority of brain injured people live with their families, the families themselves have to cope or fail to, largely unsupported. In an excellent chapter Wearing (1992) states:

> Families are often struck by the marked contrast between the frenetic activity surrounding the crisis of illness or head injury, when a hospital springs into action with emergency lights flashing, and then the long haul afterwards when the brain injured person and the family may be left largely to their own devices. Ironically, the family is typically excluded in the hospital: waiting in the corridor during the consultant's rounds, confined to visiting hours. The doctor is usually a neurologist, neurosurgeon, or general physician whose knowledge may not extend beyond the brain tissue to the consequences of damage and the possibilities of rehabilitation. ... Yet on discharge, the family may assume all responsibility overnight. (pp. 276–277)

8.20.3 PROBLEMS FACED BY PEOPLE WITH TBI

Brain injured people present with a variety of problems, the nature of which depends in part on the site, extent, and severity of the damage and the age at which it occurred. Inappropriate early treatment may also add to resulting problems. Behavioral, motor, sensory, emotional, social, and cognitive problems are common after many kinds of brain injury. Some people may demonstrate all these problems while others demonstrate just one or two. Neuropsychologists may be involved in the treatment of any and all of these different problems.

8.20.3.1 Motor Problems

Although motor problems are most likely to be treated by physiotherapists, who are trained to prevent deformities and contractures and to understand how to reteach motor skills, neuropsychologists can assist physiotherapists in encouraging success in rehabilitation in several ways. For example, knowledge of learning theory can be harnessed to improve motor learning as Robertson and Cashman (1991) showed when they improved the walking ability of a stroke patient through providing auditory feedback which increased the number of times the patient placed her heel properly. Their patient was a 29-year-old woman who had sustained several strokes which left her with unilateral neglect and frontal lobe problems. She caused difficulty in her physiotherapy sessions because despite being able to lower her heel to the floor on command, she nevertheless walked with her left heel off the floor and in an unstable position. Anxious to avoid injury her clinical psychologist and physiotherapist devised a pressure-sensitive switch attached to a buzzer on the patient's belt. The switch was inserted under her left heel, and a walking program implemented. The woman's ability to make contact with her heel on the floor was gradually increased through a process of charting progress and setting goals. Improvements generalized to the patient's everyday life.

Carr and Wilson (1983) also used a pressure switch to improve cooperation with a skin-care regime in a man with spinal cord injury and pressure sores.

Behavioral assessment and task analysis can be used to measure and understand motor disorders. Wilson and Powell (1994), for example, describe a behavioral treatment program which led to a young head injured woman tolerating physiotherapy exercises. The young woman, severely injured in a road traffic accident, was causing problems in physiotherapy because she refused to do her exercises. She had contractures, that is, shortening of the muscles, which meant she could not straighten her arms and legs. These contractures had developed during her three-month period of coma following the accident. She had also developed "myositis ossificans" as a result of the contractures. Myositis ossificans is the name given to a condition in which deposits of calcium accrue on the muscles. These caused the young woman considerable pain when she engaged in certain exercises. Nevertheless, she needed to do these exercises in order to gain sufficient flexibility for her family to manage to get her dressed, get her in and out of the car, and to the toilet.

Following a behavioral analysis, it was determined that certain exercises caused her more pain and distress than others. There was one exercise she positively enjoyed. This was head balancing performed to improve head control. The "Premack principle" (i.e., using a desired activity to reward an undesired activity) was employed together with shaping and positive reinforcement. The young woman was asked to increase the time spent on one "difficult" exercise by just a few seconds each day. If she succeeded she was praised and allowed to spend a few minutes on the head balancing exercise. Her success was also recorded on a wall chart. After five sessions on one difficult exercise, the procedure was applied to a second exercise and, later, to a third. The patient gradually tolerated longer at each exercise and her contractures slowly improved.

A number of head injured people show good motor recovery if this is measured by functional mobility (McKinlay, Brooks, Bond, Martinage,

& Marshall, 1981) although they may show more subtle motor deficits. Brain stem damage, for example, can lead to difficulty controlling fine movements. Damage to the cerebellum may lead to gross tremor and a staggering gait. Frontal lobe damage may cause difficulty in initiating movements or may result in the poor sequencing of motor movements involved in such tasks as making a cup of coffee or changing bed linen. Focal lesions following, for example, stroke or tumor can lead to permanent and intractable motor impairments.

8.20.3.2 Sensory Problems

Sensory impairments are also common after CNS damage. Hemiplegic stroke patients frequently lose sensation in their paralyzed limbs and may, as a consequence, damage themselves by trapping a leg in the spokes of a wheelchair or by burning an arm as a result of leaving it too close to a radiator. Visual sensory deficits are common following stroke, head injury, tumor, and other kinds of brain damage. Some studies have suggested that it is possible to reduce hemianopia in stroke patients (Zihl, 1990), although others have disputed this (Pommerenke & Markowitsch, 1989). Even if reduction of hemianopia is unachievable, hemianopic patients can be taught to compensate for their visual field loss through, for example, improved scanning (Kerkoff, Münßinger, Eberle-Strauss, & Stögerer, 1992). It is even possible to reduce other visual deficits such as myopia through treatment procedures. Collins, Ricci, and Burkett (1981) improved the short-sightedness of nonbrain damaged people and Thomas (reported in Wilson, 1991b) achieved the same effect with a young head injured woman whose poor visual acuity resulted from a motor cycle accident.

Auditory sensory problems are less frequently encountered, although occasionally head injured people become deaf as a result of the accident. It is sometimes possible to teach these people sign language or another alternative communication system. Typically these programs are implemented by speech and language therapists, although neuropsychologists may also take on such treatment, either alone or working together with the speech and language therapists. Once again, principles from learning theory and other branches of psychology can enhance learning.

8.20.3.3 Emotional and Social Problems

Emotional and social difficulties, including fear, anxiety, depression, and social isolation, are faced by perhaps the majority of brain injured people. Prigatano (1995) believes that dealing with the emotional effects of brain injury is crucial to rehabilitation success. Depression and anxiety can be expected in about two-thirds of patients with TBI (McKinlay et al., 1981). TBI patients are also likely to face social isolation (Talbott, 1989; Wilson, 1991a). Emotional changes are also seen frequently in stroke patients. Left and right hemisphere strokes tend to cause different emotional problems. Those with right hemisphere lesions are likely to show denial or indifference to their problems (Heilman, Watson, & Bowers, 1983), while those with left hemisphere lesions are more likely to show "a catastrophic reaction." Patients with brain stem strokes may well show extreme emotional lability (Ross, 1983). Of course emotional and social problems can also occur for nonorganic reasons such as problems with mobility, financial restrictions, or fear of what might happen in the future.

Gainotti (1993) distinguishes three main factors that cause emotional and psychological problems after brain injury. First are neurological factors that provoke disturbances by the disruption of the specific neural mechanisms subserving the regulation and control of emotional and social behavior. Second are psychological or psychodynamic factors involving attitudes towards the disability arising from awareness of the disability and its implications for the patient's quality of life. Third are the consequences of the functional impairment on the patient's social network and social activities.

The majority of clinical psychologists and some neuropsychologists are well placed to deal with the emotional and social consequences of brain injury. Anxiety management procedures, cognitive therapy, and other methods for dealing with depression, family therapy, and social skills training can all be applied to brain injured people as well as to those whose mental health problems arise from psychiatric causes.

8.20.3.4 Behavior Problems

Behavior problems, frequently seen at some stage in people with severe head injury, may worsen over time if not treated properly (BPS report, 1989). They certainly cause great stress for families and caregivers (Brooks, 1984). Livingston, Brooks, and Bond (1985) report that families of head injured people show twice the level of psychiatric dysfunction than can be expected in the general population.

A growing awareness of the importance of behavioral management programs is reflected in a number of publications (Alderman & Burgess, 1994; Alderman & Ward, 1991; Hanlon, Clontz, & Thomas, 1993; Wilson, 1991b; Wood, 1988). Common problems associated with severe head injury include yelling and swearing. Less common but more difficult to manage are physical violence and sexually offensive behaviors. Such problems may result in patients being admitted to long-term psychiatric care. Even very severe behavioral disturbance, however, may respond to behavior modification regimes (Alderman & Burgess, 1994; Alderman & Ward, 1991; Eames, 1989; Wood, 1984). Shiel, McLellan, Wilson, Evans, and Pickard (1996) provide some tentative evidence that violent behavior is more likely to occur in TBI patients who do not receive rehabilitation.

8.20.3.5 Cognitive Problems

Finally, we address cognitive problems, which almost always occur after severe TBI and frequently follow other causes of brain damage, and are probably of most interest to the majority of clinical neuropsychologists. Disorders of learning and memory, information processing, planning and organizational problems, slowness of intellectual activity, and communication are all common after TBI (Ponsford et al., 1995). Less common are the agnosias, apraxias, and aphasias although these are often found after stroke, hypoxic brain damage, and some tumors.

There has been an increasing interest in the management and remediation of cognitive deficits since the early 1980s (Diller & Gordon, 1981; Riddoch & Humphries, 1994; Seron & Deloche, 1989; Trexler, 1982; Uzzell & Gross, 1986; Wilson, 1987). Cognitive rehabilitation programs are now found in many rehabilitation centers. These may be part of the "holistic programs" of Ben-Yishay (New York) Prigatano (Phoenix), and Christensen (Copenhagen) (see Ben-Yishay & Prigatano, 1990; Christensen, 1994, and Prigatano, 1994 for more detailed descriptions), or specifically tailored to particular cognitive problems such as memory (Wilson, 1995), attention (Sohlberg & Mateer, 1989) executive deficits (Alderman, Fry, & Youngson, 1995; von Cramon & Matthes-von Cramon, 1992), and language (Byng & Coltheart, 1986). Whatever the approach, the aim of cognitive rehabilitation is to enable brain injured people and their families to live with, manage, by-pass, reduce, or come to terms with cognitive deficits resulting from an insult to the brain (Wilson, 1989b). This is achieved through a process in which brain injured people work together with health service professionals to remediate or alleviate cognitive deficits.

Cognitive, emotional, and behavior problems are far more likely to impede return to work or independent living than the more obvious physical problems. Although physical restrictions can indeed be traumatic for the individual concerned, physically handicapped people can and do achieve independence, hold down a job, and make rational decisions about their lives. People with cognitive handicaps, however, may have impaired judgment and be slow to learn new things; they may be unable to remember what they were doing a few minutes ago, and be very slow to process information. As a consequence, they may remain permanently dependent on their families. Many of these are unaware of the extent of their problems and may think they can do things just as well and just as quickly as they were able to do before the brain injury. Such lack of insight is likely to impede rehabilitation efforts. Behavior problems can exacerbate the situation as people with tantrums and poor self-control are not readily tolerated in society. Apathy and indifference reduce the chances of obtaining and holding down a job. Personality changes, mood swings, inappropriate social behavior, concrete thinking, and loss of sense of humor all reduce the chances of successful reintegration into society.

8.20.4 ASSESSMENT OF PEOPLE WITH TBI

Assessment of people with brain damage, like any other assessment, is conducted in order to answer questions. The nature of these questions determines the assessment procedure used. In the case of a brain-injured patient it is important to know a patient's cognitive strengths and weaknesses. Is there intellectual impairment? Is speed of processing information slower than one would expect for the patient's age? Is memory functioning affected and if so, which aspects of memory are intact and which impaired? What kind of language (or perceptual or reading) disorder is present?

8.20.4.1 Models of Assessment

Before treatment is planned, these questions must be addressed. Formal psychological and neuropsychological tests can usually answer these sorts of questions reasonably well. Formal, standardized tests used by neuropsychologists have been influenced by a number of theoretical approaches. The psychometric approach, for example, is based on statistical analysis and includes measures of reliability, validity, and performance of a selected sample of a given population. (See Anastasi, 1982, for a succinct account of the characteristics of psychological tests.)

The Wechsler Adult Intelligence Scale (WAIS) and the Wechsler Adult Intelligence Scale-Revised (WAIS-R) (Wechsler, 1955, 1981) are two examples of tests influenced by psychometry. Obviously, when a trained rehabilitation clinician uses a test of this kind, more than a mechanistic documentation of results will be produced. Such a clinician will understand the principles on which the test has been created and the relevance of data collected in support of the instrument's standardization, validity, and reliability. Without such an understanding, it is unlikely that good diagnosis will follow testing.

The identification of such neuropsychological syndromes as agnosia and apraxia requires a different form of assessment, in which the examiner must eliminate or exclude explanations for a particular problem. For example, to diagnose the recognition disorder known as visual object agnosia, the examiner must eliminate poor visual acuity and naming disorders as explanations of the failure to recognize objects. Lezak (1995) discusses both theoretical and practical considerations in her comprehensive account of the characteristics of neuropsychological assessment.

The development of theoretical models in the field of cognitive psychology has provided a rich source of theoretical support for the assessment of brain-injured people. For example, the working-memory model (Baddeley & Hitch, 1974) has influenced the assessment of neurologically impaired people, particularly those with Alzheimer's disease and other memory disorders. This model has enabled clinicians to assess separately the individual components of working memory, visual and verbal memory, and semantic and episodic memory. Furthermore, the model helps explain or predict such differences as those seen between people with short- and long-term memory deficits. Similarly, a model of reading proposed by Coltheart (1985), suggesting that readers can either use the whole word lexical route or the phonetic route, has led to a more careful and systematic analysis of acquired disorders of reading. (See Wilson, 1987, and Wilson & Patterson, 1990, for a more detailed discussion of these issues.)

Localization studies from neuropsychology provide another approach to assessment, whereby the examiner attempts to assess deficits in, for example, the functioning of the right and left hemispheres, the frontal lobes, the temporal lobes, and so forth. The Halstead–Reitan battery (Halstead, 1947; Reitan & Davison, 1974) is an example of this approach. The battery was originally used to discriminate among patients with frontal lobe lesions or between frontal patients and normal controls.

While these theoretical approaches to assessment offer relatively sophisticated diagnoses of individual patients, they do not, and indeed cannot, address other important issues arising from cognitive and neuropsychological deficits. They cannot, for example, address the effects on patients in terms of how their daily lives will be altered, and how changes after insult will affect the lives of relatives.

The enormous personal and social problems confronted by patients and their relatives in their home environments are a major concern for those working with brain injured people. Because these problems are not a primary concern for theorists working in such areas as localization, neuropsychological syndromes, or cognitive psychology, there is the danger of a split developing between theorists and practitioners. Indeed, I believe these different perspectives have led to division between theorists and practitioners in the past. But this division has very little substance and simply requires greater understanding of differences in role and purpose on the part of the protagonists to ensure that theory and practice are mutually supportive whenever this is possible and appropriate.

8.20.4.2 Behavioral Assessment

For those working in rehabilitation, behavioral assessment is an approach that is both theoretical and practically applicable to problems encountered in the everyday lives of patients. Behavioral assessment can bridge the gap between diagnosis and treatment in ways that other forms of assessment cannot. It is a major force in rehabilitation precisely because it can address a problem as it affects a patient's well-being. In behavioral assessment one can answer questions such as the following: What problems cause the most distress to a patient's family? How many times does this person ask the same question in the course of a day? Does this person bring a notebook to therapy sessions? Does this person remember to put on wheelchair brakes before transferring to the toilet? When questions like these are answered, a relevant treatment can be implemented.

Checklists, interviewing scales, and direct observation can all be used to identify and monitor the everyday implications of neuropsychological impairment (see Wilson, 1987, 1989a, for further discussion). One of the ways in which behavioral assessment differs from standardized testing is that it is usually part of the treatment process itself. Behavioral assessment identifies problems for treatment and can evaluate the effectiveness of that treatment. Consequently, the therapist or psychologist continues to assess the patient while treatment is progressing. Treatment can thus be modified or altered in response to information obtained. It is this inherent dynamism that is so useful to the rehabilitator: doors are always kept open, and alternative routes can be tried (i) in order to make life more bearable for the patient, and (ii) to make the behavior of the patient more conducive to tolerable living.

I want to emphasize here that there is no one right way to assess a brain-injured patient, and no one discipline can answer all questions arising during rehabilitation. When assessing a particular patient, the clinical neuropsychologist should draw from a number of disciplines and should keep in mind that theoretical principles and models are not necessarily static, but are there to be modified in the light of experience. Above all, good theory is dynamic and can grow as knowledge is extended. Neuropsychologists who continually relate practice to theory and keep up with current academic discussion are likely to find their rehabilitative work is indeed supported by a sturdy theoretical framework. An example of an assessment of an 18-year-old girl with TBI following a horse riding accident is provided to illustrate how psychometric, neuropsychological, and behavioral approaches contributed to (i) better understanding of her problems and (ii) planning for rehabilitation.

J was involved in a serious horse riding accident when she was 17 years old. She sustained a very severe head injury and was in coma for 3 months. A few months later, by which time she was 18, J was admitted to a rehabilitation center and underwent a neuropsychological assessment.

The first test administered was the WAIS. Results from the verbal subtests suggested that J's IQ was in the average range. It was not possible to estimate her performance IQ because of severe perceptual difficulties. Despite a gross tremor due to cerebellar damage, J had the motor ability to perform the tasks but she was not able to identify the visual stimuli. Furthermore, this difficulty extended to written words. She could not, therefore, do the National Adult Reading Test, a test often given to estimate premorbid ability (Nelson, 1982). We knew that J had been studying for examinations (O-levels) at the time of the accident that suggested that she was probably functioning in the average range before her head injury.

Standardized memory tests indicated normal immediate memory with severely impaired delayed memory for both visual and verbal material.

A summary of results from standardized tests (see Figure 1) was presented as a graph in the manner described by Lezak (1979). This was shown to the other staff who could see at a glance where J's main cognitive strengths and weaknesses lay.

Further assessment was carried out using other models described above. It became clear during administration of the WAIS that J could not recognize visual stimuli. For example, she thought a picture of a pig (in the picture completion subtest) was a hippopotamus. This difficulty could be due to (i) poor acuity, (ii) a naming disorder, or (iii) an object recognition disorder. Poor acuity was excluded as J could do a number of tests involving eyesight. She could also name to description well, provided the function of the object to be named was described rather than the physical appearance. Furthermore, her vocabulary was good, she just did not know what things looked like. When asked "What is an eagle?," she replied "A carnivorous bird," but when asked "What does an eagle look like?" she said, "it has four legs, lots of teeth and a long tail." J turned out to have a visual object agnosia (see Wilson & Davidoff, 1993, for further details). Thus, agnosia was diagnosed by excluding other explanations for her object recognition problem.

Cognitive neuropsychological models of reading were employed to try to understand the nature of J's reading deficit. In particular, Coltheart's (1985) dual-route model of reading helped explain that J's problem lay at a very early stage, namely identification of visual stimuli. Later when J had been taught to read again it was possible to see that she had permanent damage to the "whole word" route but was able to use the phonological route, that is, build up words from their sounds. Consequently, J changed from being alexic (totally unable to read) to being surface dyslexic (able to read regular but not irregular words). For a description of her reading program see Wilson (1987).

Finally, we needed to know how J's problems manifested themselves in real life. Direct observation was employed to note problems caused by her agnosia. It became clear that J had no major difficulty in structured situations when objects were in "context" and in their right place. She could manage at the toilet, she could feed and wash herself, and could dress herself provided clothes were given one at a time and J was told what they were.

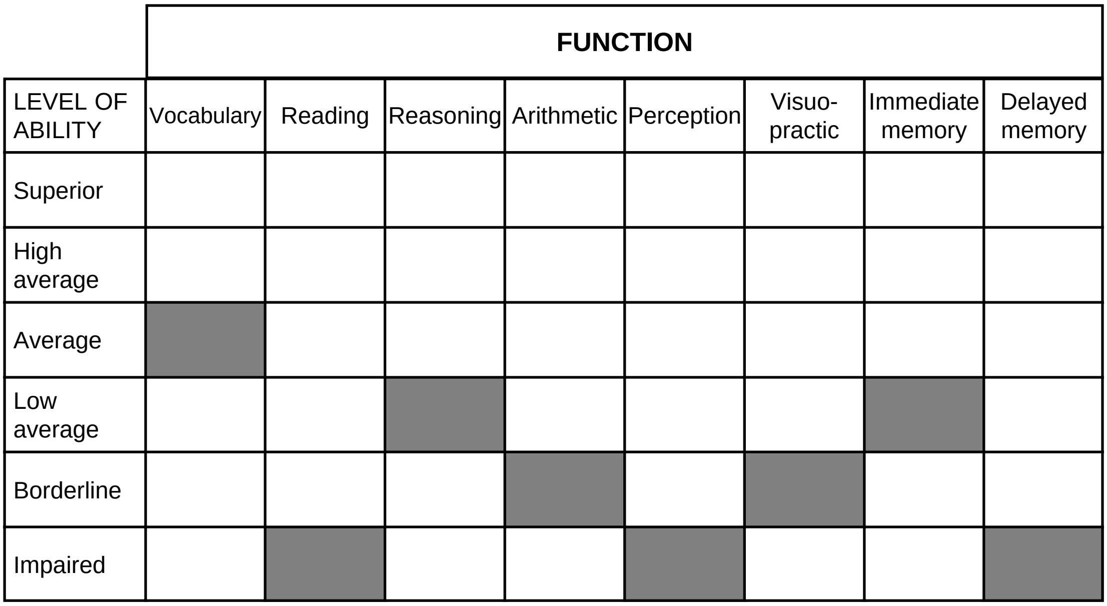

	FUNCTION							
LEVEL OF ABILITY	Vocabulary	Reading	Reasoning	Arithmetic	Perception	Visuo-practic	Immediate memory	Delayed memory
Superior								
High average								
Average	▨							
Low average			▨				▨	
Borderline				▨		▨		
Impaired		▨			▨			▨

Figure 1 J's cognitive strengths and weaknesses as identified by standardized tests (adapted from Lezak, 1979).

However, when things were out of context or when there were several choices, J became confused. She once said to a nurse on the ward "Is this toothpaste?" "Yes," said the nurse. "Good," replied J, "because yesterday I cleaned my teeth with Vaseline and the day before with hair cream."

J also had problems when out for a walk. She thought two dogs across the river were ducks, she could not distinguish cats, dogs, and sheep although she could distinguish people. Sometimes people thought she was stupid for her mistakes but more often they thought she had bad eyesight. Yet it was not her eyes at fault but the interpretation of what she saw.

As a result of these different approaches, it was possible to build up a detailed picture of J's cognitive strengths and weaknesses, provide a theoretical explanation of the observed phenomena and plan a coherent rehabilitation program.

8.20.5 PRINCIPLES AND PRACTICES OF REHABILITATION

Rehabilitation is a two-way process. Unlike treatment, which is given to a patient, rehabilitation is a process in which the patient, client, or disabled person takes an active part. Professional staff work together with the disabled person to achieve the optimum level of physical, social, psychological, and vocational functioning. The ultimate goal of rehabilitation is to enable the person with a disability to function as adequately as possible in his or her most appropriate environment.

8.20.5.1 Restoration of Function

In the early days and weeks following brain injury, attempts are usually made to restore lost functioning. Thus therapists working with stroke patients who have lost the ability to walk and talk will try to teach the patients to walk and talk again. After a while, however, if restoration of function has not occurred, therapists are likely to adjust their objectives and perhaps concentrate on the goal of wheelchair independence or teach an alternative communication system. Both of these are acceptable approaches to rehabilitation (Diller, 1994).

It is not always easy to decide when attempts towards restoration of function should be abandoned. In the early days, weeks, and even months after brain injury, recovery or partial recovery may be rapid. Even though the rate of recovery slows down later, it may nevertheless continue for several years (Robertson & Wilson, in press). It is probable that natural recovery can be enhanced or facilitated through rehabilitation. For example, lesioned animals placed in stimulating, enriched environments show greater recovery than animals left in unstimulating, barren environments (Gentile, Gren, Neiburgs, Schmelzer, & Stein, 1978; Kolb, 1992). Enriched environments also appear to lead to greater cortical recovery (Black, Sirevaag, & Greenough, 1987; Mayer, Brown, Dunnett, & Robbins, 1992) Furthermore, Mayer et al. (1992) showed that rats given striatal neural transplants only benefited from transplants when they were also given opportunities for learning. It is likely that enriched environments for humans also lead to greater recovery. Brain injured patients placed in unstimulating rooms are likely to do less well than those receiving rehabilitation.

Stimulation has been shown to change cortical activity in brain injured humans. Lindgren, Hagstadius, Åbjörnsson, and Ørbæk (1997) gave 14 patients with organic solvent induced encephalopathy a 10-week program of visual imagery training, following which a change in right hemisphere regional blood flow was observed during activation. Hemiplegic stroke patients also show greater recovery when forced to use their weak limb through immobilization of the nonaffected limb (Taub et al., 1993).

Attempting to restore lost functioning represents only one aspect of rehabilitation. Teaching or enabling people to compensate for their deficits and psychotherapy to address the emotional problems and enhance self-esteem are others. The provision of counseling, education, and support is yet another. Prigatano (1995), Jackson and Gouvier (1992), and Wilson (1995) address some of these issues. Our concern in this section is compensation, which is a widely used strategy in rehabilitation, particularly when the rate of natural recovery has slowed down or stopped altogether. Compensation can be achieved in several ways.

8.20.5.2 Environmental Restructuring

One can avoid or bypass certain problems through restructuring the environment. Compensation is thus achieved by removing the patient's need to function in a particular way and that function is provided by the environment. As an example, consider people who are totally paralyzed from the neck down. With the appropriate equipment they can control their physical environments through a voice-activated control mechanism or by using their mouths to control a stick. Through these mechanisms they can open and close doors and windows, turn the pages of a book, answer the telephone and so forth. The disabled person controls the structured environment through his or her mouth and no longer needs the use of limbs to do this. Similarly, a person with severe intellectual impairments can be enabled to function in a suitably structured environment with signposts, labeling of doors, reminders from staff, and alarms which alert carers if the person wanders off. "Smart houses" for people with cognitive deficits are already being designed for people with dementia (Slaven, personal communication) in an attempt to "disable the disabling environment." These environments are controlled by computers, video links, and telephones and they can remind people about toilets, baths, and medication, ensure that showers are the right temperature and turn electrical appliances on and off.

8.20.5.3 Anatomical Reorganization

It might be possible to achieve compensation through anatomical reorganization thereby enabling undamaged areas of the brain to take over the function of the damaged area. Although we know this can happen in babies and infants (Kohn & Dennis, 1978), it is often at the expense of other functions (Dennis & Kohn, 1975) and it is less clear to what extent this happens in adults. Robertson and Wilson (in press) consider this in some detail.

8.20.5.4 Using Residual Skills More Efficiently

Perhaps the two most important types of compensation in neuropsychological rehabilitation are (i) teaching people to use their residual skills more efficiently, and (ii) finding an alternative way to achieve the desired outcome.

In this approach it is assumed that some residual functioning remains. Thus people with amnesia do not lose all memory functioning and can possibly be helped to use what little remains more effectively. This might include allowing extra time to learn new information, making associations between the new information they are trying to remember and old information they already know. Mnemonics, for example, probably work because they allow previously isolated items to become integrated with one another (Bower, 1972). Many of the memory therapy techniques described by Wilson (1987) follow this approach. Rehearsal techniques such as Landauer and Bjork's (1978) method of expanding rehearsal have been employed by Moffat (1989) to teach new words to a woman with nominal aphasia, and by Camp and Schaller (1989) in their work with Alzheimer's patients. In both these studies people were helped to use their residual skills more efficiently.

8.20.5.5 Alternative Solutions

Finding an alternative solution to a problem or relying upon an intact cognitive skill to compensate for one that is damaged is one of the most long-standing procedures in rehabilitation. Zangwill (1947) believed it to be an important approach in rehabilitation and gave as an example of teaching people with a right hemiplegia to use their left hand. Later Luria et al. (1969) described the principle of "functional adaptation," that is, if you cannot do something one way, find another way to do it. Many of the aids provided by occupational therapists for activities of daily living can be described as functional adaptations. Equipment for disabled people provides an alternative solution to tasks such as dressing, eating, and cooking. This practice also works well in cognitive rehabilitation. Teaching dysphasic people to communicate with sign language (Skelly, 1979) or visual symbols (Gardner, Zurif, Berry, & Baker, 1976), "talking books" for people who are blind and the use of memory aids for people with amnesia are all examples of compensation through finding an alternative solution. Wilson and Watson (1996) discuss the development of compensatory behavior in people with severe memory problems. They also consider which characteristics predict good use of compensations. Age (younger than 30 years at the time of the brain injury), less severe memory deficits, and absence of additional cognitive problems appear to be the most important predictors.

These different approaches to rehabilitation are not mutually exclusive and can be used in combination. For example, a memory impaired person may be helped (i) by reorganizing the environment to reduce the load on memory, (ii) by being taught to use external memory aids, that is, use a functional adaptation approach, and (iii) by being encouraged to use residual skills more efficiently when learning new information through employing mnemonics and rehearsal strategies. It should be noted, however, that severely memory impaired people rarely use mnemonics spontaneously. Therapists or relatives typically supply the mnemonics, whose main value is that their use leads to faster learning. Wilson (1992) offers more detailed discussion of the use of mnemonics with memory impaired people. An example of a memory rehabilitation program for a head injured man is provided below.

Fred was injured in a road traffic accident when he was 54 years old. He sustained a severe head injury with a coma lasting 10 days. Three months later he was discharged from hospital having made a good physical recovery. His main problem at this stage was his memory. He had virtually no recall of what he had been doing a few minutes earlier. He asked the same questions so often he irritated his wife and son almost beyond endurance. He lost his belongings, could not remember where to meet his wife on shopping trips, and had difficulty learning any new information such as new telephone numbers or the name of his granddaughter born during his stay in hospital.

Fred was seen as an outpatient for three hours a week. He received individual therapy for the first hour, then after a short break he spent two hours attending a memory group with five other people whose memory problems also resulted from a neurological insult. The individual treatment lasted for 12 weeks and the group sessions for 10 weeks.

Fred's wife attended for part of each individual session. She completed a two-week "memory diary" in which she recorded Fred's everyday memory failures. Certain problems appeared very frequently and some of these were selected for treatment. Fred lost his belongings several times a day, particularly his keys, spectacles, and diary. He could not remember where his socks, shirts, and other clothes were kept in the bedroom and became frustrated when he could not find them. He never knew what he was supposed to do each day or where he was going, and he regularly forgot his new telephone number which had changed a few months earlier.

A number of strategies were implemented over the following weeks. Fred's wife was helped to arrange the house to make it easier for Fred to find his belongings and also to find his way around. She wrote labels for his chest of drawers and wardrobe so that he could see at a glance where his socks, shirts, and ties were. She tidied his shelves and made boxes for his keys and spectacles. She encouraged Fred to return his keys and spectacles to the same place every time he finished using them. When he took his keys out of the house he was prompted to clip his keyring to his belt so he would not lose his keys. She bought a smaller diary than the one Fred had been using so that Fred could keep it in his pocket. These small adjustments to the environment enabled Fred to be more independent. They also reduced his frustration. Furthermore, he gradually learned to put his keys and spectacles away without prompting.

The next step was to teach Fred to use his diary more efficiently. He wore a watch with an alarm. The plan was to set the alarm to sound every hour during waking hours and to teach Fred to check his diary every time the alarm sounded. A few practice sessions were needed first. The alarm sounded and Fred was prompted to look at his diary. It sounded a minute later and he was prompted again. We then set the alarm to sound two minutes later and he was prompted again, although by now he was initiating the action himself. In fact through prompting and expanding rehearsal (i.e., gradually increasing the test interval), Fred rapidly learned this easy "paired associate" task.

As well as checking his diary, it was also important for Fred to record information and to cross off each task or appointment once it was completed. This was also achieved through prompting and then fading out the prompts much in the way described by Kime, Lamb, and Wilson (1996). Fred learned to use his diary through the provision of very close supervision and structure initially and then slowly withdrawing the amount of help required. If Fred had been seen a year or two later we would have considered using NeuroPage, a simple to use but highly effective paging system developed in California (Hersh & Treadgold, 1994) and evaluated in the UK (Wilson, Evans, Emslie, & Malinek, 1997).

We then tackled the problem of the new telephone number and the new granddaughter's name. Fred was taught these without too much trouble by using a combination of errorless learning and expanding rehearsal. Errorless learning, originally used with pigeons (Terrace, 1963) and then with learning disabled people (Sidman & Stoddard, 1967), is a method of teaching which avoids trial and error and ensures success at the beginning of the learning process. Baddeley and Wilson (1994) and Wilson, Baddeley, Evans, and Shiel (1994) demonstrated that this method was effective for teaching memory impaired patients new information.

Fred was provided with the telephone number written on a card. He was "tested" with the number in front of him and in full view for several trials. He then had to repeat the number immediately after removal of the card. After several trials using this method, Fred was tested after two-second delay. After several more trials the delay was increased to four seconds and then eight and so on. At the first sign of hesitation the number was provided for him. We planned to return to an earlier and easier delay if Fred made an error, but fortunately he did not and the new number was learned within the first half hour of a one-hour session. Fred then had to demonstrate generalization of his new knowledge by telephoning his home several times. His new granddaughter's name was taught in a similar way.

Fred's wife was also given information and explanations about her husband's problem and what to expect. Her many questions were answered to the best of our ability and she was given some easy to digest literature to read to back up our information and explanations. Written advice was provided on the best way to manage some of Fred's problems.

The main purposes of the group sessions were to reduce Fred's anxiety and increase his opportunities for social interaction (Evans & Wilson, 1992; Wilson & Moffat, 1992). We also hoped he would imitate his peers" use of external aids and have an opportunity to try out his own strategies.

Thus, Fred's program incorporated: (i) reorganization of his environment to reduce the load on his memory, (ii) teaching him to use his diary more efficiently, (iii) helping him to use his residual skills through errorless learning and expanding rehearsal, and (iv) reducing anxiety through attendance at a memory group.

8.20.6 COGNITIVE REHABILITATION

"Cognitive rehabilitation is a process whereby brain injured people work together with health service professionals to remediate or alleviate cognitive deficits arising from a neurological insult" (Wilson, 1996, p. 637).

In the late, 1990s there would appear to be four main approaches to cognitive rehabilitation: (i) those focusing on cognitive retraining through drills and exercises; (ii) those based on models from cognitive neuropsychology; (iii) those combining theory and practice from neuropsychology, cognitive psychology, and behavioral psychology; and (iv) those using a holistic approach by addressing emotional, motivational, and other noncognitive aspects of functioning in addition to cognitive aspects.

8.20.6.1 Cognitive Retraining

The cognitive retraining approach seems to rely upon the assumption that it is possible to remediate underlying cognitive deficits (or at least to teach patients how to deal with their cognitive problems) by exercise, practice, and stimulation. It is therefore akin to the "mental muscle" approach (Harris & Sunderland, 1981) which assumes that appropriate cognitive exercises can improve cognition in the same way that physical exercise can improve physical well-being and muscle tone. Such programs typically involve the presentation of patients with a cognitive exercise regime to work through. The exercises are often computerized. Gianutsos (1980) described one such treatment for a woman in her forties who had memory, perceptual, and reading problems following a meningioma and a stroke. Gianutsos reported that in total 23 cognitive rehabilitation sessions were conducted with this patient, both before and after discharge from the hospital to her home. "The sessions lasted from 45 minutes to two hours and occurred within a two-month period, except for a follow-up session five months later. Cognitive retraining was organized around two tasks: (i) memory for short lists of three unrelated words with varying amounts of interference in the recall interval, and (ii) oral reading and recall of paragraphs" (p. 38)

Many American programs seem to follow the exercise, stimulation, and practice model, including the original Diller (1976) program in which patients were taken through a series of training models to improve, for example, visual information processing or spatial awareness. Others who use a retraining exercise approach include Parenté, Anderson-Parenté, and Shaw (1989), who developed an "iconic memory scanning" training to improve "iconic retention"; Adamovich, Henderson, and Auerbach (1985) and Sohlberg and Mateer (1989), who describe a range of exercises for improving various cognitive deficits; Lynch (1982), Bracy (1986), Gianutsos and Klitzner (1981) and others describe computerized cognitive retraining programs.

There is a distinct lack of evidence that retraining through exercise is effective. Miller (1984) found little evidence of success for the stimulation, exercise, and drilling approach. Wilson (1982) reported no effect of exercise on memory functioning, a finding supported by Glisky (1995). In a fairly comprehensive review of computerized cognitive rehabilitation, Robertson (1990) found no evidence of significant changes in memory, visuoperceptual, or visuospatial functioning as a result of computerized training. Language training programs for highly specific disorders fared a little better although there was no published evidence of more general effectiveness of computerized language training. Only in attention training were there some positive results, although even here the results were contradictory, with some studies reporting positively and others negatively.

Since Robertson's (1990) review paper, further computerized attentional training programs have appeared support of the tentative evidence available in the 1980s. Gray, Robertson, Pentland, and Anderson (1992) found that on a number of outcome measures a group of brain injured patients receiving computerized attentional retraining were marginally better than a control group of brain injured patients receiving recreational computerized training. However, six months later the experimental group did significantly better than the control group on two tests of attention, indicating benefits at follow-up that were not apparent at the time of training. Sturm and Willmes (1991) gave computerized training to right and left hemisphere damaged stroke patients. Right hemisphere damaged patients showed more pronounced impairments of sustained attention and vigilance, whereas left hemisphere damaged patients were worse on choice reaction tasks. Both groups improved with training on a number of attention functions but not for vigilance. Moreover, the training effects were less pronounced for the right brain damaged group. Just as Robertson reported in 1990, Sturm and Willmes (1991) also found no evidence of generalization to more general cognitive functions or to real life tasks.

Despite the lack of evidence, there are still those who believe in the exercise/drilling approach. Gianutsos (1992), for example, said, "restoration as a goal" should be pursued "even when that goal might be remote" because patients view it as appropriate (p. 29). Further problems with the retraining through exercising approach, however, include: (i) It fails to address the functional manifestations of cognitive problems as experienced by brain injured patients in real life, that is, it tackles impairments as measured by tests rather than disabilities as measured by problems encountered in everyday life. Learning lists of words or identifying particular stimuli on a computer screen is not, as a rule, of any particular concern to brain injured people. (ii) The approach fails to address generalization problems. The only point to "cognitive retraining" would be if improvements on the training tasks generalized to real life, but to date we have no convincing evidence of this. Furthermore, without such evidence it would seem wrong to build in unnecessary steps for brain injured clients, who typically have problems learning, especially when they could be taught directly the information or skills they require for real life functioning. (iii) Emotional, social, and behavioral sequelae of brain injury are ignored yet these can be just as handicapping as cognitive difficulties (McKinlay et al., 1981; Thomsen, 1985). Patients and families are sometimes in danger of exploitation through unscrupulous companies, selling computers and software to people desperately hoping for some improvement in their daily lives. (iv) This approach lacks any theoretical underpinning.

8.20.6.2 Cognitive Neuropsychology

Cognitive neuropsychological theoretical approaches use a cognitive model, typically of language or reading, to identify the specific deficit of an individual patient. The deficits are often very precisely identified by the models. The underlying rationale would appear to be that once a deficit can be identified it can be treated. In the words of Coltheart (1991), "One needs to have in mind a model of how naming is normally achieved before one can begin to seek to understand impairments of naming consequent upon brain damage; and one needs this understanding before one determines what kinds of treatment could be appropriate" (pp. 216–217).

In the UK and some other European countries this approach is held in high esteem. Eminent proponents are Coltheart (1991), Colheart, Bates, and Castles (1994), Riddoch and Humphries (1994), Seron, van der Linden, and de Partz (1991), and Mitchum and Berndt (1989, 1995). A typical rehabilitation program within this approach would involve a careful assessment (following one's particular model) to identify the component(s) of the model that are dysfunctional. Having carefully analyzed the deficit "the treatment is selected specifically to fit in with the theoretical interpretation of the functional impairment" (Mitchum & Berndt, 1995, p. 7). The treatment itself, however, typically includes practice at the damaged component of the model. Although we saw earlier that practice is not a good rehabilitation strategy, it would be argued by cognitive neuropsychologists engaged in rehabilitation that the drilling practice and exercising approach fails because it fails to identify the damaged component precisely enough.

Even among cognitive neuropsychologists there are critics of this approach. Caramazza (1989) remained unconvinced of the value of models for rehabilitation. Wilson and Patterson (1990) claimed that the recent interest in treatment on the part of neuropsychologists has been almost entirely for the benefit of the scientist rather than the patient. Baddeley (1993) suggests that cognitive neuropsychology has learned a great deal from the study of brain damaged patients over the past 20 years but it is by no means clear that brain damaged patients have benefited from cognitive neuropsychology.

One of the major problems with this approach is that while cognitive neuropsychological models are extremely useful in identifying the nature of the deficit, and for explaining observed phenomena, they provide little or no information on ways of treating the deficit (Wilson & Patterson, 1990). In other words, they tell us *what* to treat but not *how* to treat. Or, putting it in other terms, this approach can be extremely good at identifying the *impairment*, but has little or nothing to offer when it comes to treatment of the *disability*, which is the real business of rehabilitation.

These terms come from the conceptual framework developed by the World Health Organization (1980). The framework suggests classifying the sequelae of brain injury into impairments, disabilities, and handicaps. Impairments can be regarded as damage to physical or mental structures (for example, parietal lobe damage); disabilities refer to the particular problems caused by the impairment (for example, being unable to get from a wheelchair to the toilet because of unilateral neglect and spatial attention difficulties). Handicaps can be seen as problems imposed by society because of someone's disability (for example, a person in a wheelchair is handicapped in a building without wheelchair access but not in a building with wheelchair access).

A further limitation is that the theoretical cognitive neuropsychology approach is best suited to patients with relatively pure deficits. Indeed, rehabilitationists adopting models from cognitive neuropsychology typically seek out such patients, and typically report single cases. Of course single cases are invaluable for confirming or disconfirming theories (Shallice, 1979), and as Caramazza (1989) wrote, "Experimental research with brain damaged patients . . . has had as its explicit goal that of informing theories of normal cognitive functioning" (p. 384). For many cognitive neuropsychologists, the *raison d'être* for their interest in rehabilitation is to develop, extend, or modify their models. In clinical practice the majority of brain damaged patients have multiple deficits and for this group theoretical models have been less useful.

Because theoretical models tend to focus attention on test performance rather than ways in which cognitively impaired people dysfunction in everyday life, they are like retraining models focusing on impairment rather than disability. Patients rarely want or need to improve scores on tests in the face of far greater demands requiring them to function more efficiently in real life, in family, social, and vocational settings. If the subsequent cognitive rehabilitation offered by the theoretical cognitive neuropsychological approach was successful in terms of these real life needs of patients, then I would have one less objection to them, but so far there is little evidence of this happening. Typically, improvements on test scores are provided as evidence of success rather than some of the more functional measures such as reading for pleasure, greater confidence in using the telephone, or engaging more often in conversation.

My opposition is not aimed at the models *per se*, and indeed, regarded as tools that have enabled us to come to a greater understanding of the nature of a number of cognitive impairments I am deeply appreciative. However, like Caramazza and Hillis (1993), I would argue that understanding on a theoretical level, although essential, is not a sufficient requirement on its own to address rehabilitation as conceived of by the World Health Organization when they describe it as "enabling people to achieve optimum social integration." Cognitive neuropsychological approaches fail to address emotional, social, or behavioral sequelae of brain injury, yet, as we noted earlier, these are often more handicapping than the cognitive if not addressed properly.

8.20.6.3 Combined Approaches

The combined approach, involving learning theory, cognitive psychology, and neuropsychology, is perhaps practiced most in the UK, and probably reflects British clinical psychology training. I return to this later in Section 8.20.8.

A behavioral approach is usually incorporated into neuropsychological rehabilitation because it provides a structure, a way of analyzing a problem, of assessing everyday manifestations of difficulties, of evaluating the efficacy of treatment, and because there are numerous treatment strategies that can be modified or adapted for brain injured patients (Lincoln, 1978; Wilson, 1987, 1989a, 1991b). Although formal behavior modification programs are used in some centers (e.g., Wood, 1987; Wood & Eames, 1981), many psychologists engaged in cognitive rehabilitation are less inclined to use the more rigid behavioral approaches of Skinner (1953) or Kazdin (1975) and more inclined to use elements such as task analysis and behavioral assessment, together with other strategies such as anxiety management and systematic desensitization.

However, behavioral approaches are not in themselves sufficient for our purposes in cognitive rehabilitation. A better understanding of theory and the organization of the brain are required, hence the incorporation of cognitive and neuropsychology into the programs. Ponsford et al. (1995) also use this broad approach in their work in Australia.

Wilson (1984, 1987) puts forward the case for the combined approach, although Lincoln (1978) had earlier used behavioral elements in neuropsychological rehabilitation, as indeed had Ince (1976, 1980). Robertson (1988, 1990) also comes from this school although he perhaps emphasizes the cognitive and neuropsychological aspects rather more than the behavioral.

The basic principle upon which this approach is based is that the everyday problems of brain injured patients can be reduced. If the patient/client fails to learn or fails to change, it is the psychologist who is at fault for not finding the right strategy. The main strengths of this approach are perhaps its emphasis on monitoring, record keeping, and evaluation of success or failure.

One example of this combined approach is provided by Wilson (1988) in the treatment of a woman who sustained severe brain damage in an anesthetic accident. Despite being able to walk and talk, she was, nevertheless, unable to care for herself. She could not feed herself, drink from a cup unaided, dress herself, or manage herself at the toilet. Her problems were due to an extremely disabling apraxia, that is, a disorder of movement that is not due to paralysis, weakness, or failure to understand the task. The woman could manage the individual motor movements required but could not organize these movements into a coherent sequence.

Formal assessment demonstrated widespread cognitive impairments across a number of domains. It was hard to explain the pattern of deficits in terms of a theoretical model, so a behavioral approach was adopted in which one goal was selected for treatment—drinking from a cup alone. Detailed observations were carried out when she attempted to drink: she frequently misreached the cup, aiming too far to the right or left or too short or too far; sometimes she found the handle but could not grip and on other occasions she held the very top of the cup so that it was unstable. When she did get hold of the cup, she was often unable to let go of it.

The task was broken down into nine steps and each step scored from 1 to 4 depending on the amount of help required (1 = manages this step alone and 4 = needs maximum physical guidance to achieve this step). Taking the woman through these steps enabled her to learn the task within a few trials. The procedure was then employed to teach her other self-care tasks. Thus a neuropsychological syndrome—apraxia—was found to coexist with numerous other cognitive deficits. Although hard to find an adequate theoretical interpretation of her pattern of deficits, a behavioral approach enabled her to learn sufficient self care skills to remain outside institutional care. The behavioral approach included: (i) task analysis, (ii) careful behavioral assessment of each step in each task, and (iii) modification of a behavioral teaching strategy, namely chaining, to teach the task.

To the outside observer the main problem would appear to be that this approach fails to address the patient's emotional needs. Prigatano (personal communication) says that the approach does not consider the patient's feelings. In practice, this criticism is perhaps more apparent than real, as the patient's concerns and anxieties are frequently addressed.

Anxiety management strategies, particularly desensitization and relaxation, are commonly employed. In many programs patients are involved in setting their own rehabilitation goals, in monitoring their treatment, and in taking control of their own therapy. Nevertheless, it is true to say that patients" feelings, emotions, and their self-esteem are not as a rule formally seen as part of their cognitive rehabilitation. Furthermore, it is common practice to focus on one aspect of cognitive dysfunction, such as memory, attention, or reading, rather than address a broader spectrum.

8.20.6.4 The Holistic Approach

The "holistic" approach, pioneered by Diller (1976), Ben-Yishay (1977), and Prigatano et al. (1986), "consists of well-integrated interventions that exceed in scope, as well as in kind, those highly specific and circumscribed interventions which are usually subsumed under the term 'cognitive remediation'" (Ben-Yishay & Prigatano, 1990, p. 400). Proponents of the "holistic" approach regard it as futile to separate the cognitive, the psychiatric, and the functional from the affective sequelae of brain injury (Ben-Yishay et al., 1985). Ben-Yishay's model (1985) follows a hierarchy of stages through which the patient must work in rehabilitation and these are: engagement, awareness, mastery, control, acceptance, and identity.

Major themes of the daily programs include increased awareness, acceptance and understanding, cognitive remediation, development of compensatory skills, and vocational counseling. The rehabilitation day typically starts with a morning meeting of staff and patients, with the rest of the day structured around individual and small group therapy, together with a further meeting of all staff and clients. The individual sessions include cognitive training, psychotherapy, and other specific therapies as required (e.g., speech and language therapy or relaxation therapy). Group therapy also covers cognitive and psychotherapy, plus special groups as required (e.g., problem solving or communication). A relatives" group may be held once a week, and vocational counseling and work trials are usually included.

Clinically, there appears to be little wrong with this approach. It is hard to find irrefutable evidence of its success but it has probably been subjected to more research on efficacy than other approaches (Diller, 1994); and, as Cope (1994) says, "Although none of the outcome research provides definitive proof of rehabilitation efficacy, the sum of many studies ... provides reasonably convincing evidence that comprehensive rehabilitation does make a substantial difference in outcome of handicap for TBI (traumatic brain injury) patients" (p. 218).

Good as the holistic programs are, it is probably true to say they can be improved. Incorporating some of the positive aspects from the combined learning theory, cognitive psychology, and neuropsychology programs mentioned earlier might, for example, enhance their clinical effectiveness. Careful task analysis, baselines, monitoring, and the implementation of single case experimental designs should enable us to evaluate more easily the successes and failures. The implementation of theoretical models to (i) identify cognitive strengths and weaknesses, and (ii) explain observed phenomena should extend the boundaries of achievement for cognitive rehabilitation programs.

There is increasing evidence that these programs result in less emotional distress, increased self-esteem, and greater productivity for participants than is observed in those patients who do not undergo such treatment (Prigatano et al., 1986, 1994), as well as evidence that these programs are cost-effective (Mehlbye & Larsen, 1994).

8.20.7 EFFECTIVENESS OF REHABILITATION

Research into effectiveness can be divided into those studies looking at cost-effectiveness, those looking at clinical effectiveness, and those looking at improvement in the quality of life.

In one American study of 145 brain injured patients (Cope, Cole, Hall, & Barkan, 1991), the estimated savings in care costs following rehabilitation for severely brain injured people was over £27 000 per year. A Danish study (Mehlbye & Larsen, 1994) reported that spendings in health and social care for patients attending a nonresidential program were recouped in five years. The costs of not rehabilitating brain injured people are also considerable, given the fact that many are young with a relatively normal life expectancy (Greenwood & McMillan, 1993). In a review of rehabilitation effectiveness, Cope (1994) believed that postacute rehabilitation programs can produce sufficient savings to justify their support on a cost benefit basis. He also suggested that they could be run even more efficiently in the future.

8.20.7.1 Clinical Effectiveness

Numerous studies have shown clinical improvement or reduction in disability in handicap following rehabilitation. One study (Aronow, 1987) showed no difference between a group of brain injured patients receiving rehabilitation and a control group not receiving rehabilitation. However, the group receiving rehabilitation was more impaired to start with, so caught up with the others through rehabilitation. Two studies (Fryer & Haffey, 1987; Mills, Nesbeda, Katz, & Alexander, 1992) found that training in real life activities yielded more positive results than trying to address underlying impairment. Two other studies (Blackerby, 1990; Spivack, Spettell, Ellis, & Ross, 1992) found that high levels of treatment delivered in a relatively short time were associated with better outcome. A follow-up study of head injured patients (Shiel et al., 1996) found those who received rehabilitation were far less likely to show physical aggression than those who did not.

However, it is harder to find evidence directly comparing more than one of the four approaches described in this paper, although an interesting study by Rattok et al. (1992) looked at a comparison of different treatment mixes within a holistic program. They compared the different intensity of treatments for three groups of brain injured patients. All groups achieved the same success rate in terms of employment at six-months follow-up. The groups differed, however, in that those who spent most time in cognitive rehabilitation did better on cognitive tasks, those who spent more time on interpersonal skill training did better on tasks involving interpersonal skills, and those trained on both improved on both.

8.20.7.2 Does Rehabilitation Improve Quality of Life?

A British study (Johnson, 1987) found the employment rate of brain injured people could be tripled following individualized and intensive work re-entry programs. Prigatano and colleagues (Prigatano et al., 1986) found that patients undergoing treatment were less emotionally distressed than control group patients. More recently Prigatano and colleagues (Prigatano et al., 1994) reported greater productivity in those patients undergoing a holistic rehabilitation program.

8.20.7.3 Overall View of Effectiveness

Just a few of the many studies attempting to answer the question "Is rehabilitation effective?" have been selected. Despite some methodological problems, strong evidence of the overall effectiveness of rehabilitation on general measures of handicap has emerged in the past five years (Cope, 1994). There is mounting evidence that rehabilitation reduces the effect of cognitive, social, and emotional problems, leading to greater independence on the part of the patient, reduction in family stress, and eventual employability for many brain injured people.

8.20.8 PROFESSIONAL ISSUES

In order to achieve the skills necessary for the professional psychological tasks described here, neuropsychologists require a basic training in clinical psychology followed by specialization in neuropsychology as in the UK. The clinical psychology training should be generic, with all students spending part of their time working in adult mental health, child services, and learning disability. In addition to these compulsory placements, students choose three of four specialist placements which may include neuropsychology. The academic part of the training would typically include cognitive psychology, neuropsychology, learning theory (and behavior modification), developmental psychology, psychotherapy, abnormal psychology, statistics, and research design. In many courses the basic philosophy is one that recognizes that clinical psychologists are trained to solve health service problems: we are not allowed to say a patient is untestable or untreatable; instead we should use a problem-solving strategy to find a way to assess or treat or refer to the right service/department. Finally, within this tradition, clinical psychologists are expected to be research oriented and to evaluate the efficacy of their intervention programs.

8.20.9 CONCLUSIONS: AN INTEGRATED APPROACH TO COGNITIVE REHABILITATION

In the light of the discussion above, and the fact that feelings affect how we think and how we behave, we need to recognize that cognition should not be divorced from emotion, motivation, or other noncognitive functions. Consequently, an integrated program should address not only cognitive functions but also social, emotional, functional, and affective difficulties.

Programs should focus on *disability*, that is, the functional consequences of brain injury, rather than *impairment*. Effectiveness should not be evaluated by improvements on test results but on functional achievements in everyday life.

Cognitive neuropsychologists engaged in rehabilitation should encourage dialogue with other disciplines. We typically conduct diagnostic studies or experiments to elucidate and clarify the nature of impairment. Other rehabilitation staff are typically concerned with the remediation of real life problems and in helping brain injured people manage their disabilities. The challenge of those engaged in cognitive rehabilitation is to develop concepts, tools, and terminology to bridge these different interests.

We need a broader theoretical spectrum on which to base our rehabilitation programs. Wilson (1984, 1987) argued for the need to combine theoretical models and methodology from neuropsychology, cognitive psychology, and behavioral psychology. Diller (1987) said, "While current accounts of remediation have been criticized as lacking a theoretical base, it might be more accurate to state that remediation must take into account several theoretical bases" (p. 9). More recently, McMillan and Greenwood (1993) said, "rehabilitation should draw on fields of clinical neuropsychology, behavioral analysis, cognitive retraining and group and individual psychotherapy" (p. 352).

We need to strengthen the interaction between theory and clinical observation. Theory can sometimes predict which methods might work but clinical observation is often required to tell us how best to implement the theory. For example, work on errorless learning in the rehabilitation of amnesic patients (Baddeley & Wilson, 1994) grew out of theoretical work on implicit memory. The implementation of errorless learning principles in clinical rehabilitation, however, requires behavioral observation and analysis at an individual patient level (Wilson et al., 1994; Wilson & Evans, 1996).

We need to ensure proper evaluation of cognitive rehabilitation programs. Given that patients and therapists cannot be "blind" to treatment procedures, double-blind randomized control trials are always unlikely to be realized in neuropsychological rehabilitation. Single-blind trials are sometimes possible (Cope, 1994). Even more possible are small group or cohort studies, crossover designs, and single case experimental designs (Wilson, 1997, discusses in some detail research and evaluation in rehabilitation).

Research should be at a number of levels of specificity. At the most general level are questions such as "Is this approach effective?" and "What are the cost consequences of providing and not providing neuropsychological rehabilitation?" These questions are, perhaps, best answered through collaborative research maybe involving a health economist. At a more specific level we need to evaluate different rehabilitation strategies such as the comparison of two (or more) methods to improve attention or teach new information. This is where small group or cohort designs are potentially very valuable. At the greatest level of specificity is the evaluation of an individual client's response to treatment. Is this patient changing and is the change due to spontaneous recovery or to our intervention are two of the most important questions in rehabilitation and ones that every person engaged in clinical practice should attempt to answer.

8.20.10 REFERENCES

Adamovich, B. B., Henderson, J. A., & Auerbach, S. (1985). *Cognitive rehabilitation of closed head injured patients: A dynamic approach*. London: Taylor & Francis.

Alderman, N., & Burgess, P. (1994). A comparison of treatment methods for behaviour disorder following herpes simplex encephalitis. *Neuropsychological Rehabilitation, 4,* 31–48.

Alderman, N., Fry, R. K., & Youngson, H. A. (1995). Improvement of self-monitoring skills, reduction of behaviour disturbance and the dysexecutive syndrome: Comparison of response cost and a new program of self-monitoring training. *Neuropsychological Rehabilitation, 5,* 193–221.

Alderman, N., & Ward, A. (1991). Behavioral treatment of the dysexecutive syndrome: Reduction of repetitive speech using response cost and cognitive overlearning. *Neuropsychological Rehabilitation, 1,* 65–80.

Allsopp, R., & Turner, E. (1986). Road casualties and public transport fares in London. *Accident Analysis and Prevention, 18,* 147–156.

Anastasi, A. (1982). *Psychological testing* (5th ed.). New York: Collier/Macmillan.

Aronow, H. U. (1987). Rehabilitation effectiveness with severe brain injury: Translating research into policy. *Journal of Head Trauma Rehabilitation, 2,* 24–36.

Baddeley, A. D. (1993). A theory of rehabilitation without a model of learning is a vehicle without an engine: A comment on Caramazza and Hillis. *Neuropsychological Rehabilitation, 3,* 235–244.

Baddeley, A. D., & Hitch, G. (1974) Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York: Academic Press.

Baddeley, A. D., & Wilson, B. A. (1994). When implicit learning fails: Amnesia and the problem of error elimination. *Neuropsychologia, 32,* 53–68.

Barnes, M. (1996). Legislative and policy factors In F. D. Rose & D. A. Johnson (Eds.), *Brain injury and after: Towards improved outcome* (pp. 163–183). Chichester, UK: Wiley.

Ben-Yishay, Y. (Ed.) (1977). Working approaches to remediation of cognitive deficits in brain damaged persons. *Rehabilitation Monograph No. 59*. New York: New York University Medical Center, 1978.

Ben-Yishay, Y., & Prigatano, G. P. (1990). Cognitive remediation. In M. Rosenthal, E. R. Griffith, M. R. Bond, & J. D. Miller (Eds.), *Rehabilitation of the adult and child with traumatic brain injury,* (2nd ed., pp. 393–409). Philadelphia: F.A. Davis.

Ben-Yishay, Y., Rattok, J., Lakin, P., Piasetsky, E. B., Ross, B., Silver, S., Zide, E., & Ezrachi, P. (1985). Neuropsychologic rehabilitation: Quest for a holistic approach. *Seminars in Neurology, 5,* 252–259.

Black, J. E., Sirevaag, A. M., & Greenough, W. T. (1987). Complex experience promotes capillary formation in young rat visual cortex. *Neuroscience Letters, 83,* 351–355.

Blackerby, W. F. (1990). Intensity of rehabilitation and length of stay. *Brain Injury, 4,* 167–173.

Bower, G. H. (1972). Mental imagery and associative learning. In L. W. Gregg (Ed.), *Cognition in learning and memory* (pp. 51–88). New York: Wiley.

British Psychological Society Report (1989). *Services for young adult patients with acquired brain damage*. Leicester, UK: Author.

Bracy, O. (1986). *Programs for cognitive rehabilitation*. Indianapolis, IN: Psychological Software Services.

Brooks, D. N. (Ed.) (1984). *Closed head injury: Social, psychological and family consequences*. Oxford UK: Oxford University Press.

Brooks, D. N., & McKinley, W. W. (1983). Personality and behavioral change after severe blunt head injury—a relative's view. *Journal of Neurology, Neurosurgery, and Psychiatry, 46,* 336–344.

Bryden, J. (1989). How many head injured? The epidemiology of post head injury disability. In R. Ll. Wood & P. Eames (Eds.), *Models of brain injury rehabilitation* (pp. 17–27). London: Chapman & Hall.

Byng, S., & Coltheart, M. (1986). Aphasia therapy research: methodological requirements and illustrative results. In E. Hjelmquist & L. G. Nilsson (Eds.), *Communication and handicap* (pp. 383–398). Amsterdam: North-Holland.

Camp, C. J., & Schaller, J. R. (1989). Epilogue: Spaced retrieval memory training in an adult day-care center. *Educational Gerontology, 15,* 641–648.

Caramazza, A. (1989). Cognitive neuropsychology and rehabilitation: an unfulfilled promise? In X. Seron & G. Deloche (Eds.), *Cognitive approaches in neuropsychological rehabilitation* (pp. 383–398). Hillsdale, N.J.: Erlbaum.

Caramazza, A., & Hillis, A. E. (1993). For a theory of remediation of cognitive deficits. *Neuropsychological Rehabilitation, 3,* 217–234.

Carr, S., & Wilson, B. A. (1983). Promotion of pressure relief exercising in a spinal injury patient: A multiple baseline across settings design. *Behavioral Psychotherapy, 11,* 329–336.

Christensen, A-L. (1994). Visions for rehabilitation. In A-L. Christensen & B. P. Uzzell (Eds.), *Brain injury and neuropsychological rehabilitation* (pp. 293–299). Hillsdale, NJ: Erlbaum.

Collins, F. L., Ricci, J. A., & Burkett, P. A. (1981). Behavioral training for myopia: Long term maintenance of improved acuity. *Behavioral Research and Therapy, 9,* 265–268.

Coltheart, M. (1985). Cognitive neuropsychology and reading. In M. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 3–37). Hillsdale, NJ: Erlbaum.

Coltheart, M. (1991). Cognitive psychology applied to the treatment of acquired language disorders. In P. Martin (Ed.), *Handbook of behavior therapy and psychological science: An integrative approach* (pp. 216–226). New York: Pergamon.

Coltheart, M., Bates, A., & Castles, A. (1994). Cognitive neuropsychology and rehabilitation. In M. J. Riddoch & G. W. Humphreys (Eds.), *Cognitive neuropsychology and cognitive rehabilitation* (pp. 17–37). Hove, UK: Erlbaum.

Condeluci, A. (1985). *National Head Injury Survey*. Washington, DC: National Head Injury Foundation.

Cope, N. (1994). Traumatic brain injury rehabilitation outcome studies in the United States. In A.-L. Christensen & B. P. Uzzell (Eds.), *Brain injury and neuropsychological rehabilitation: International perspectives* (pp. 201–220). Hillsdale, NJ: Erlbaum.

Cope, D. N., Cole, J. R., Hall, K. M., & Barkan, H. (1991). Brain injury: Analysis of outcome in a post-acute rehabilitation system. *Brain Injury, 5,* 111–125.

Dennis, M., & Kohn, B. (1975). Comprehension of syntax in infantile hemiplegics after cerebral hemidecortication: Left hemisphere superiority. *Brain and Language, 2,* 472–482.

Diller, L. L. (1976). A model for cognitive retraining in rehabilitation. *The Clinical Psychologist, 29,* 13–15.

Diller, L. (1987). Neuropsychological rehabilitation. In M. J. Meier, A. L. Benton, & L. Diller (Eds.), *Neuropsychological rehabilitation*. Edinburgh: Churchill Livingstone.

Diller, L. (1994). Changes in rehabilitation over the past 5 years. In A.-L. Christensen & B. P. Uzzell (Eds.), *Brain injury and neuropsychological rehabilitation: International perspectives* (pp. 1–15). Hillsdale, NJ: Erlbaum.

Diller, L., & Gordon, W. A. (1981). Rehabilitation and clinical neuropsychology. In S. B. Filskov & T. J. Boll (Eds.), *Handbook of clinical neuropsychology* (pp. 702–733). New York: Wiley.

Dunbar, J. A., Penttila, A., & Pikkarainen, J. (1987). Drinking and driving: Choosing the legal limits. *British Medical Journal, 295,* 1458–1460.

Eames, P. (1989). Head injury rehabilitation: towards a "model" service. In R. Ll. Wood & P. Eames (Eds.), *Models of brain injury rehabilitation* (pp. 48–58). London: Chapman & Hall.

Evans, C. D. (1981). *Rehabilitation after severe head injury*. Edinburgh: Churchill Livingstone.

Evans, J. J. & Wilson, B. A. (1992). A memory group for individuals with brain injury. *Clinical Rehabilitation, 6,* 75–81.

Finger, S., & Stein, D. (1982). *Brain damage and recovery: Research and clinical perspectives*. New York: Academic Press.

Fryer, J., & Haffey, W. (1987). Cognitive rehabilitation and community readaptation: Outcomes from two program models. *Journal of Head Trauma Rehabilitation, 2,* 51–63.

Gainotti, G. (1993). Emotional and psychosocial problems after brain injury. *Neuropsychological Rehabilitation, 3,* 259–277.

Gardner, H., Zurif, E. B., Berry, T., & Baker, E. (1976). Visual communication in aphasia. *Neuropsychologia, 14,* 275–292.

Gentile, A. M., Green, S., Neiburgs, A., Schmelzer, W., & Stein, D. G. (1978). Disruption and recovery of locomotor and manipulatory behavior following cortical lesions in rats. *Behavioral Biology, 22,* 417–455.

Gianutsos, R. (1980). What is cognitive rehabilitation? *Journal of Rehabilitation, 1,* 37–40.

Gianutsos, R. (1992). The computer in cognitive rehabilitation: It's not just a tool any more. *Journal of Head Trauma Rehabilitation, 7,* 26–35.

Gianutsos, R., & Klitzner, C. (1981). *Computer programs for cognitive rehabilitation*. Bayport, NY: Life Science Associates.

Glisky, E. L. (1995). Computers in memory rehabilitation. In A. D. Baddeley, B. A. Wilson, & F. N. Watts (Eds.), *Handbook of memory disorders* (pp. 557–575). Chichester, UK: Wiley.

Gray, J. M., Robertson, I. H., Pentland, B., & Anderson, S. (1992). Microcomputer-based attentional training after brain damage: A randomized group controlled trial. *Neuropsychological Rehabilitation, 2,* 97–115.

Greenwood, R. J., & McMillan, T. M. (1993). Models of rehabilitation programmes for the brain injured adult. I: Current provision, efficacy and good practice. *Clinical Rehabilitation, 7,* 248–255.

Halstead, W. C. (1947). *Brain and intelligence*. Chicago: University of Chicago Press.

Hanlon, R., Clontz, B., & Thomas, M. (1993). Management of severe behavioral dyscontrol following subarachnoid haemorrhage. *Neuropsychological Rehabilitation, 3*, 63–76.

Harris, J. E., & Sunderland, A. (1981). A brief survey of the management of memory disorders in rehabilitation units in Britain. *International Rehabilitation Medicine, 3*, 206–209.

Heilman, K. M., Watson, R. T., & Bowers, D. (1983). Affective disorders associated with hemispheric disease. In K. M. Heilman & P. Satz (Eds.), *Neuropsychology of human emotion* (pp. 45–64). New York: Guilford Press.

Hersh, N., & Treadgold, L. (1994). NeuroPage: The rehabilitation of memory dysfunction by prosthetic memory and cueing. *Neuropsychological Rehabilitation, 4*, 187–197.

Ince, L. P. (1976). *Behavior modification in rehabilitation medicine*. Springfield, IL: Thomas.

Ince, L. P. (1980). *Behavior psychology in rehabilitation medicine: Clinical applications*. Baltimore: Williams & Wilkins.

Jackson, W. T., & Gouvier, W. D. (1992). Group psychotherapy with brain-damaged adults and their families. In C. J. Long & L. K. Ross (Eds.), *Handbook of head trauma: Acute care to recovery* (pp. 309–327). New York: Plenum.

Jennett, B. (1990). Scale and scope of the problems. In M. Rosenthal, E. R. Griffith, M. R. Bond, & J. D. Miller (Eds.), *Rehabilitation of the adult and child with traumatic brain injury* (pp. 3–7). Philadelphia: F. A. Davis

Jennett, B., & Teasdale, G. (1981). *Management of head injuries*. Philadelphia: F.A. Davis.

Johnson, R. (1987). Return to work after severe head injury. *International Disability Studies*, 9, 49–54.

Kazdin, A. E. (1975). *Behavior modification in applied settings*. Homewood, IL: Dorsey Press.

Kerkoff, G., Münβinger, U., Eberle-Strauss, G., & Stögerer, E. (1992). Rehabilitation of hemianopic alexia in patients with postgeniculate visual field disorders. *Neuropsychological Rehabilitation, 2*, 21–42.

Kime, S. K., Lamb, D. G., & Wilson, B. A. (1996). Use of a comprehensive program of external cuing to enhance procedural memory in a patient with dense amnesia. *Brain Injury, 10*, 17–25.

Klonoff, P. S., Costa, L. D., & Snow, W. G. (1986). Predictors and indicators of quality of life in patients with closed head injury. *Journal of Clinical and Experimental Neuropsychology, 8*, 469–485.

Kohn, B., & Dennis, M. (1978). Selective impairments of visuospatial abilities in infantile hemiplegics after right cerebral hemidecortication. *Neuropsychologia, 12*, 505–512.

Kolb, B. (1992). Mechanisms underlying recovery from cortical injury: reflections on progress and directions for the future. In F. D. Rose & D. A. Johnson (Eds.), *Recovery from brain damage: Reflections and directions* (pp. 169–186). New York: Plenum.

Kolb, B. (1995). *Brain plasticity and behavior*. Mahwah, NJ: Erlbaum.

Kraus, J., Morgenstern, H., Fife, D., Conroy, C., & Nourjah, P. (1989). Blood alcohol tests, prevalence of involvement, and outcomes following brain injury. *American Journal of Public Health, 79*, 294–299.

Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Gruneberg, P. E. Morris & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625–632). London: Academic Press.

Lezak (1979). *Neuropsychological assessment*. New York: Oxford University Press.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Lincoln, N. B. (1978). Behavior modification in physiotherapy. *Physiotherapy, 64*, 265–267.

Lindgren, M., Hagstadius, S., Åbjörnsson, G., & Ørbæk, P. (1997). Neuropsychological rehabilitation of patients with organic solvent-induced chronic toxic encephalopathy. A pilot study. *Neuropsychological Rehabilitation, 7*, 1–22.

Livingston, M. A., Brooks, D. N., & Bond, M. (1985). Three months after severe head injury: Psychiatric and social impact on relatives. *Journal of Neurology, Neurosurgery, and Psychiatry, 48*, 870–875.

Luria, A. R. (1963). *Recovery of function after brain injury*. New York: Macmillan.

Luria, A. R., Naydin, V. L., Tsvetkova, L. S., & Vinarskaya, E. N. (1969). Restoration of higher cortical functions following local brain damage. In P. J. Vinken & G. W. Bruyn (Eds.), *Handbook of clinical neurology* (Vol. 3, pp. 368–433). New York: Elsevier.

Lynch, W. (1982). The use of electronic games in cognitive rehabilitation. In L. E. Trexler (Ed.), *Cognitive rehabilitation: Conceptualization and intervention*. New York: Plenum.

Mayer, E., Brown, V. J., Dunnett, S. B., & Robbins, T. W. (1992). Striatal graft-associated recovery of a lesion-induced performance deficit in the rat requires learning to use the transplant. *European Journal of Neuroscience, 4*, 119–126.

MccGwire, S. (1986). *Kim's story. A fight for life*. London: Harrap.

McKinlay, W. W., Brooks, D. N., Bond, M. R., Martinage, D. P., & Marshall, M. M. (1981). The short-term outcome of severe blunt head injury as reported by relatives of the injured persons. *Journal of Neurology, Neurosurgery and Psychiatry, 44*, 527–533.

McMillan, T. M., & Greenwood, R. J. (1993). Model of rehabilitation programs for the brain-injured adult—II: Model services and suggestions for change in the UK. *Clinical Rehabilitation, 7*, 346–355.

Mehlbye, J., & Larsen, A. (1994). Social and economic consequences of brain damage in Denmark. In A.-L. Christensen & B. P. Uzzell (Eds.), *Brain injury and neuropsychological rehabilitation: International perspectives* (pp. 257–267). Hillsdale, NJ: Erlbaum.

Miller, E. (1984). *Recovery and management of neuropsychological impairments*. Chichester, UK: Wiley.

Miller, J. D., Pentland, B., & Berrol, S. (1990). Early evaluation and management. In M. Rosenthal, E. R. Griffith, M. R. Bond, & J. D. Miller (Eds.), *Rehabilitation of the adult and child with traumatic brain injury, second edition* (pp. 21–51). Philadelphia: Davis.

Mills, V. M., Nesbeda, T., Katz, D. I., & Alexander, M. P. (1992). Outcomes for traumatically brain injured patients following post acute rehabilitation programmes. *Brain Injury, 6*, 219–228.

Mitchum, C. C., & Berndt, R. S. (1989). Aphasia rehabilitation: An approach to diagnosis and treatment of disorders of language production. In M. G. Eisenberg (Ed.), *Advances in clinical rehabilitation*. New York: Academic Press.

Mitchum, C. C., & Berndt, R. S. (1995). The cognitive neuropsychological approach to treatment of language disorders. *Neuropsychological Rehabilitation, 5*, 1–16.

Moffat, N. (1989). Home based cognitive rehabilitation with the elderly. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 659–680). Cambridge, UK: Cambridge University Press.

National Head Injury Foundation (1988). Task Force on Special Education: *An educator's manual: What educators need to know about students with traumatic brain injury*. Southborough, MA: Author.

Nelson, H. E. (1982). *The National Adult Reading Test (NART)*. Windsor, UK: NFER-Nelson.

Oddy, M. (1995). He's no longer the same person: How families adjust to personality change after head injury. In M. Anne Chamberlain, V. Neumann & A. Tennant (Eds.), *Traumatic brain injury rehabilitation* (pp. 167–179). London: Chapman & Hall.

Parenté, R., Anderson-Parenté, J. K., & Shaw, B. (1989). Retraining the mind's eye. *Journal of Head Trauma Rehabilitation, 4*, 53–62.

Pommerenke, K., & Markowitsch, H. J. (1989). Rehabilitation training of homonymous visual field defects in patients with postgeniculate damage of the visual system. *Restorative Neurology and Neuroscience, 1,* 47–63.

Ponsford, J. (1990). Editorial: Psychological sequelae of closed head injury: Time to redress the imbalance. *Brain Injury, 4,* 111–114.

Ponsford, J. (1995). Mechanisms, recovery and sequelae of traumatic brain injury: A foundation for the REAL approach. In J. Ponsford, S. Sloan & P. Snow (Eds.), *Traumatic brain injury: Rehabilitation for everyday adaptive living* (pp. 1–31). Hove, UK: Erlbaum.

Ponsford, J. Olver, J. H., Curran, C., & Ng, K. (1995). Prediction of employment status two years after brain injury. *Brain Injury, 9,* 11–20.

Ponsford, J., Sloan, S., & Snow, P. (1995). *Traumatic brain injury: Rehabilitation for everyday adaptive living*. Hove, UK: Erlbaum.

Prigatano, G. P. (1994). Individuality, lesion location, and psychotherapy after brain injury. In A-L. Christensen & B. P. Uzzell (Eds.), *Brain injury and neuropsychological rehabilitation: International perspectives* (pp. 173–186). Hillsdale, NJ: Erlbaum.

Prigatano, G. P. (1995). Personality and social aspects of memory rehabilitation. In A. D. Baddeley, B. A. Wilson, & F. N. Watts (Eds.), *Handbook of memory disorders* (pp. 603–614). Chichester, UK: Wiley.

Prigatano, G. P., Klonoff, P. S., O'Brien, K. P., Altman, I. M., Amin, K., Chiapello, D., Shepherd, J., Cunningham, M., & Mora, M. (1994). Productivity after neuropsychologically oriented milieu rehabilitation. *Journal of Head Trauma Rehabilitation, 9,* 91–102.

Prigatano, G. P., Fordyce, D. J., Zeiner, H. K., Roveche, J. R., Pepping, M., & Wood, B. C. (Eds.) (1986). *Neuropsychological rehabilitation after brain injury*. Baltimore: Johns Hopkins University Press.

Rattok, J., Ben-Yishay, Y., Ezrachi, O., Lakin, P., Piasetsky, E., Ross, B., Silver, S., Vakil, E., Zide, E., & Diller, L. (1992). Outcome of different treatment mixes in a multidimensional neuropsychological rehabilitation program. *Neuropsychology, 6,* 395–416.

Reitan, R. M., & Davison, L. A. (1974). *Clinical neuropsychology: Current status and applications*. Washington, DC: Hemisphere.

Riddoch, M. J., & Humphreys, G. W. (Eds.) (1994). *Cognitive neuropsychology and cognitive rehabilitation*. Hove, UK: Erlbaum.

Rimel, R. W., Jane, J. A., & Bond, M. R. (1990). Characteristics of the head injured patient. In M. Rosenthal, E. R. Griffith, M. R. Bond, & J. D. Miller (Eds.), *Rehabilitation of the adult and child with traumatic brain injury* (2nd ed., pp. 8–16). Philadelphia: F. A. Davis.

Robertson, I. H. (1988). *Unilateral visual neglect*. Unpublished Ph.D. thesis, University of London.

Robertson, I. H. (1990). Does computerized cognitive rehabilitation work? A review. *Aphasiology, 4,* 381–405.

Robertson, I. H., & Cashman, E. (1991). Auditory feedback for walking difficulties in a case of unilateral neglect: A pilot study. *Neuropsychological Rehabilitation, 1,* 175–184.

Robertson, I. H., & Wilson, B. A. (in press). Neuropsychological rehabilitation. In J. W. Fawcett, M. V. Sofroniew & S. B. Dunnett (Eds.), *Brain damage, Brain repair*. Oxford, UK: Oxford University Press.

Rose, F. D., & Johnson, C. A. (1996). *Brain injury and after: Towards improved outcome*. Chichester, UK: Wiley.

Rosenbaum, M., & Najenson, T. (1976). Changes in life patterns and symptoms of low mood as reported by wives of severely brain-injured soldiers. *Journal of Consulting and Clinical Psychology, 44,* 881–888.

Ross, E. D. (1983). Right-hemisphere lesions in disorders of affective language. In A. Kertesz (Ed.), *Localization in neuropsychology* (pp. 493–508). New York: Academic Press.

Schacter, D., & Crovitz, H. (1977). Memory function after closed head injury: A review of the quantitative research. *Cortex, 13,* 105–176.

Seron, X., & Deloche, G. (Eds.) (1989). *Cognitive approaches in neuropsychological rehabilitation*. Hillsdale, NJ: Erlbaum.

Seron, X., van der Linden, M., & de Partz, M-P. (1991). In defence of cognitive approaches in neuropsychological therapy. *Neuropsychological Rehabilitation, 1,* 303–318.

Shallice, T. (1979). Case study approach in neuropsychological research. *Journal of Clinical Neuropsychology, 1,* 183–211.

Shiel, A., McLellan, D. L., Wilson, B. A., Evans, J. J., & Pickard, J. (1996). *The effect of rehabilitation on violent behaviour after severe head injury*. Manuscript submitted for publication.

Sidman, M., & Stoddard, L. T. (1967). The effectiveness of fading in programming simultaneous form discrimination for retarded children. *Journal of Experimental Analysis of Behavior, 10,* 3–15.

Skelly, M. (1979). *Ameri–Ind gestural code based on universal American Indian hand talk*. New York: Elsevier.

Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.

Snow, P., & Ponsford, J. (1995). Assessing and managing impairment of consciousness following traumatic brain injury. In J. Ponsford, S. Sloan, & P. Snow (Eds.), *Traumatic brain injury: Rehabilitation for everyday adaptive living* (pp. 33–64). Hove, UK: Erlbaum.

Sohlberg, M., & Mateer, C. (1989). *Introduction to cognitive rehabilitation*. New York: Guilford Press.

Spivack, G., Spettell, C. M., Ellis, D. W., & Ross, S. E. (1992). Effects of intensity of treatment and length of stay on rehabilitation outcomes. *Brain Injury, 6,* 419–434.

Sturm, W., & Willmes, K. (1991). Efficacy of a reaction training on various attentional and cognitive functions in stroke patients. *Neuropsychological Rehabilitation, 1,* 259–280.

Symonds, G. P. (1937). Mental disorder following head injury. *Proceedings of the Royal Society of Medicine, 30,* 1081–1094.

Talbott, R. (1989). The brain injured person and the family. In R. L. Wood & P. Eames (Eds.), *Models of brain injury rehabilitation* (pp. 3–16). London: Chapman and Hall.

Taub, E., Miller, N. E., Novack, T. A., Cook, E. W., Fleming, W. C., Nepomuceno, C. S., Connell, J. S., & Crago, J. E. (1993). Technique to improve chronic motor deficit after stroke. *Archives of Physical Medicine and Rehabilitation, 74,* 347–354.

Teasdale, G., & Jennett, B. (1974) Assessment of coma and impaired consciousness: a practical scale. *Lancet, 2,* 81–84.

Teasdale, G., & Mendelow, D. (1984). Pathophysiology of head injuries. In D. N. Brooks (Ed.) *Closed head injury: Social, psychological and family consequences* (pp. 4–36). Oxford, UK: Oxford University Press.

Terrace, H. S. (1963). Discrimination learning with and without "errors". *Journal of Experimental Analysis of Behavior, 6,* 1–27.

Thomas, S., Acton, C., Nixon, J., Battistutta, D., Pitt, W. R., & Clart, R. (1994). Effectiveness of bicycle helmets in preventing head injury in children: A case controlled study. *British Medical Journal, 308,* 173–176.

Thompson, R., Rivara, F., & Thompson, D. (1989). A case controlled study of the effectiveness of bicycle safety helmets. *The New England Journal of Medicine, 320,* 1361–1367.

Thomsen, I. V. (1985). Do young patients have worse outcomes after severe blunt head trauma? *Brain Injury, 3,* 157–162.

Trexler, L. E. (Ed.) (1982). *Cognitive rehabilitation: Conceptualization and intervention.* New York: Plenum.

Uzzell, B. P., & Gross, Y. (1986). *Clinical neuropsychology of intervention.* Boston: Martinus Nijhoff.

von Cramon, D., & Matthes-von Cramon, G. (1992). Reflections on the treatment of brain injured patients suffering from problem-solving disorders. *Neuropsychological Rehabilitation, 2,* 207–230.

von Monakow, C. (1914). Translated by Pribham (1969). *Brain and behaviour.* Harmondsworth, UK: Penguin.

Wade, D., & Langdon Hewer, R. (1987). Functional abilities after stroke: measurement, natural history and prognosis. *Journal of Neurology, Neurosurgery and Psychiatry, 50,* 177–182.

Walsh, K. (1987). *Neuropsychology: A clinical approach* (2nd ed.). Edinburgh, UK: Churchill Livingstone.

Watson, M., Horn, S., & Curl, J. (1992). Searching for signs of revival: Uses and abuses of the Glasgow Coma Scale. *Professional Nurse, 7,* 670–674.

Wearing, D. (1992). Self help groups. In B. A. Wilson & N. Moffat (Eds.), *Clinical management of memory problems second edition* (pp. 271–301). London: Chapman & Hall.

Wechsler, D. (1955) *Wechsler Adult Intelligence Scale. Manual.* New York: The Psychological Corporation.

Wechsler, D. (1981) *Manual for the Wechsler Adult Intelligence Scale-Revised.* New York: The Psychological Corporation.

Wilson, B. A. (1982). Success and failure in memory training following a cerebral vascular accident. *Cortex, 18,* 581–594.

Wilson, B. A. (1984). *Cognitive rehabilitation after brain damage.* Unpublished Ph.D. thesis, University of London.

Wilson, B. A. (1987). *Rehabilitation of memory.* New York: Guilford Press.

Wilson, B. A. (1988). Remediation of apraxia following an anaesthetic accident. In J. West & P. Spinks (Eds.), *Clinical psychology in action* (pp. 178–183). London: John Wright & Sons.

Wilson, B. A. (1989a). Management of problems resulting from damage to the central nervous system. In S. Pearce & J. Wardle (Eds.), *The practice of behavioral medicine* (pp. 51–81). Oxford, UK: Oxford University Press.

Wilson, B. A. (1989b). Models of cognitive rehabilitation. In R. Ll. Wood & P. Eames (Eds.), *Models of brain injury rehabilitation* (pp. 117–141). London: Chapman & Hall.

Wilson, B. A. (1991a). Long term prognosis of patients with severe memory disorders. *Neuropsychological Rehabilitation, 1,* 117–134.

Wilson, B. A. (1991b). Behavior therapy in the treatment of neurologically impaired adults. In P. R. Martin (Ed.), *Handbook of behavior therapy and psychological science: An integrative approach* (pp. 227–252). New York: Pergamon.

Wilson, B. A. (1992). Memory therapy in practice. In B. A. Wilson & N. Moffat (Eds.), *Clinical management of memory problems* (2nd ed., pp. 120–153). London: Chapman & Hall.

Wilson, B. A. (1995). Management and remediation of memory problems in brain-injured adults. In A. D. Baddeley, B. A. Wilson, & F. N. Watts (Eds.), *Handbook of memory disorders* (pp. 451–479). Chichester, UK: Wiley.

Wilson, B. A. (1996). Neuropsychological rehabilitation. In J. G. Beaumont, P. M. Kenealy & M. J. C. Rogers (Eds.), *The Blackwell dictionary of neuropsychology* (pp. 618–626). Oxford, UK: Blackwell.

Wilson, B. A. (1997). Research and evaluation in rehabilitation. In B. A. Wilson & D. L. McLellan (Eds.), *The handbook of rehabilitation studies* (pp. 161–187). Cambridge, UK: Cambridge University Press.

Wilson, B. A., Baddeley, A. D., Evans, J. J., & Shiel, A. (1994). Errorless learning in the rehabilitation of memory impaired people. *Neuropsychological Rehabilitation, 4,* 307–326.

Wilson, B. A., Baddeley, A. D., Shiel, A., & Patton, G. (1992). How does post traumatic amnesia differ from the amnesic syndrome and from chronic memory impairment. *Neuropsychological Rehabilitation, 2,* 231–243.

Wilson, B. A., & Davidoff, J. (1993). Partial recovery from visual object agnosia: A 10 year follow-up study. *Cortex, 29,* 529–542.

Wilson, B. A., & Evans, J. J. (1996). Error free learning in the rehabilitation of individuals with memory impairments. *Journal of Head Trauma Rehabilitation, 11*, 54–64.

Wilson, B. A., Evans, J., Baddeley, A. D., Emslie, H., & Watson, P. (1997, June). *Monitoring recovery from post traumatic amnesia.* Paper presented at the International Neuropsychological Society, Bergen, Norway.

Wilson, B. A., Evans, J. J., Emslie, H., & Malinek, V. (1977). Evaluation of NeuroPage: A new memory aid. *Journal of Neurology, Neurosurgery, and Psychiatry, 63,* 113–115.

Wilson, B. A., & Moffat, N. (1992). The development of group memory therapy. In B. A. Wilson & N. Moffat (Eds.), *Clinical management of memory problems (second edition)* (pp. 243–273). London: Chapman & Hall.

Wilson, B. A., & Patterson, K. E. (1990). Rehabilitation and cognitive neuropsychology: Does cognitive psychology apply? *Journal of Applied Cognitive Psychology, 4,* 247–260.

Wilson, B. A., & Powell, G. E. (1994). Neurological problems: treatment and rehabilitation. In S. J. E. Lindsay & G. E. Powell (Eds.), *The handbook of clinical adult psychology* (2nd ed., pp. 688–701). London: Routledge.

Wilson, B. A., Shiel, A., Watson, M., Horn, S., & McLellan, D. L. (1994). Monitoring behaviour during coma and post traumatic amnesia. In A. L. Christensen & B. Uzzell (Eds.), *Progress in the rehabilitation of brain-injured people* (pp. 85–98). Hillsdale, NJ: Erlbaum.

Wilson, B. A., & Watson, P. C. (1996). A practical framework for understanding compensatory behaviour in people with organic memory impairment. *Memory, 4*, 465–486.

Wood, R. Ll. (1984). Behaviour disorders following severe brain injury: Their presentation and psychological management. In D. N. Brooks (Ed.), *Closed head injury: Social, psychological and family consequences.* Oxford, UK: Oxford University Press.

Wood, R. Ll. (1987). *Brain injury rehabilitation: A neurobehavioural approach.* London: Croom Helm.

Wood, R. Ll. (1988). Management of behaviour disorders in a day treatment setting. *Journal of Head Trauma Rehabilitation, 3*, 53–62.

Wood, R. Ll., & Eames, P. G. (1981) Application of behavioral modification of traumatically brain injured adults. In G. Davey (Ed.), *Application of conditioning theory.* London: Methuen.

World Health Organization (1980). *International classification of impairments, disabilities and handicaps: A manual of classification relating to the consequences of disease.* Geneva, Switzerland: Author.

Zangwill, O. L. (1947). Psychological aspects of rehabilitation in cases of brain injury. *British Journal of Psychology, 37,* 60–69.

Zihl, J. (1990). Zur Behandlung von Patienten mit homonymen Gesichtsfeldstörungen. *Zeitschrift für Neuropsychologie, 1,* 91–98.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.21
Epilepsy

GUS A. BAKER
Walton Hospital, Liverpool, UK
and
ANN JACOBY
University of Newcastle upon Tyne, UK

8.21.1 INTRODUCTION 487
8.21.2 CLASSIFICATION OF SEIZURES AND EPILEPSY 488
8.21.3 PROGNOSIS OF EPILEPSY 489
8.21.4 MEDICAL AND SURGICAL TREATMENT OF EPILEPSY 489
  8.21.4.1 Adherence with Medical Treatment for Epilepsy 491
8.21.5 IMPACT OF EPILEPSY ON PERSON AND FAMILY 492
  8.21.5.1 Impact on Physical Function 492
  8.21.5.2 Impact on Psychological Function 493
  8.21.5.3 Impact on Social Function 494
  8.21.5.4 Impact of Epilepsy on Other Family Members 496
8.21.6 COPING WITH THE IMPACT OF EPILEPSY 497
8.21.7 LIFE SPAN ISSUES 498
8.21.8 GENDER ISSUES 499
8.21.9 PSYCHOLOGICAL INTERVENTIONS IN THE MANAGEMENT OF EPILEPSY 499
  8.21.9.1 Interventions Relating to Treatment of Seizures 499
  8.21.9.2 Interventions to Reduce the Impact of Epilepsy and Enhance Coping 499
8.21.10 CONCLUSION 501
8.21.11 REFERENCES 501

8.21.1 INTRODUCTION

Seizures are "an occasional, an excessive, and a disorderly discharge of nerve tissue" (Taylor, 1958), and epilepsy is "a chronic disorder characterized by recurrent unprovoked seizures" (Hauser & Annegery, 1993). The emphasis on the recurrent and unprovoked nature of the seizures manifested by epilepsy is important, since not all people who have seizures will be defined as having epilepsy. Febrile seizures (those occurring in young children and associated with a rise in body temperature), anoxic seizures, and seizures provoked by an acute cause such as illness or trauma are not considered to constitute epilepsy; single seizures do not constitute epilepsy; and seizures that occur sporadically probably do not constitute epilepsy.

Epilepsy is one of the most common neurological disorders, with an age-adjusted incidence of between 20 and 70 per 100 000 and a prevalence of 4–10 per 1000 (Chadwick, 1994). It has been estimated that around 50 million people worldwide have epilepsy. According to Porter (1993), in more than three-quarters of people with epilepsy, seizures begin before the age of 18 years. Though in the majority of cases there is no identifiable cause for epilepsy, it can be the result of virtually any major category of serious disease or disorder of humans, including congenital malformations, infections, tumours, vascular diseases, degenerative diseases, or injury. The main causes of seizures and epilepsy in children and adults are displayed in Table 1.

8.21.2 CLASSIFICATION OF SEIZURES AND EPILEPSY

Neurologists attempt to classify both seizure type and epilepsy syndrome for the purposes of treatment. Seizures are fundamentally divided into two groups—partial or generalized (International League Against Epilepsy, 1981). Partial seizures have clinical or electroencaphalographic evidence of a localized onset. Generalized seizures are as a group much more heterogeneous than partial seizures and have no evidence of a localized onset. Partial seizures can be further subdivided into two groups—simple and complex. Simple partial seizures are those attacks which show evidence of a localized onset but during which consciousness or responsiveness is preserved. The discharge is usually confined to a single hemisphere and the symptoms are specific to the affected region of the brain. Commonly patients may experience either a motor event such as jerking of an extremity, or a sensory event such as a bad odor or taste. Complex partial seizures are those attacks which show evidence of a localized onset and in which consciousness is altered. These attacks can be subdivided into frontal, parietal, and occipital complex partial seizures. Complex partial seizures of temporal lobe origin, the most common, often begin with disturbance of motor function followed by speech disturbance, and automatisms such as lip smacking frequently follow. The attacks usually last less than one minute and are followed by postictal confusion and gradual recovery.

Generalized seizures include tonic–clonic, clonic, tonic, atonic, myoclonic, and absence seizures. The most common form of generalized seizures, and that with which most people will be familiar, are generalized tonic–clonic seizures, which have certain well-defined characteristics. During the initial tonic phase, which usually lasts for less than a minute, the muscles of the body contract and the person becomes rigid. If the bladder is full, contraction of the muscles will result in urination. As the muscles in the lungs contract, air is forced out through the vocal cords, so that the person seems to cry out. Breathing may stop for a short time, and the person becomes cyanosed as a result. The tonic phase is followed by the clonic, or jerking, phase, with jerking caused by alternating phases of muscle relaxation and contraction. Tongue-biting sometimes occurs during this phase of the seizure. After a minute or two, the seizure ends, the person relaxes and may then be deeply unconscious and unrouseable, but consciousness normally returns within about five minutes. Arousal from the seizure may be followed by a brief period of confusion and irritability (Chadwick & Usiskin, 1987).

Absence seizures begin in childhood or early adolescence and are characterized by unresponsiveness and a variety of associated phenomena. The attacks are short, usually less than 10 seconds and rarely longer than 45 seconds. The onset is paroxysmal and without warning and the cessation is likewise sudden. Myoclonic seizures are a seizure type defined as a quick movement of the muscle which is normally stereotypical and presents as clonic jerking. The clonic seizure is usually a generalized seizure without a tonic phase and the reverse can be said

Table 1 Main causes of seizures in children and adults.

Infants and children	*Adults*
No definite cause determined	No definite cause determined
Birth and neonatal injuries	Vascular lesions
Vascular insults other than above	Head trauma
Congenital or metabolic disorders	Drug or alcohol abuse
Head injuries	Neoplasia
Infection	Infection
Neoplasia	Heredity
Heredity	

Source: Porter (1993).

of tonic seizures. Atonic seizures are characterized by a precipitous loss of tone which causes the patient to have a sudden nodding of the head or to fall to the floor. Some people with epilepsy experience partial seizures that subsequently evolve into generalized ones.

Epilepsy is more of a group of syndromes than a single disease entity, a syndrome being defined as a disorder characterized by a cluster of signs or symptoms which customarily occur together. Epileptic syndromes classify patients and emphasize the heterogeneity of these complex symptoms. Identification of syndromes is important in terms of predicting the prognosis, selecting the appropriate treatment, and defining the likelihood of identifying an underlying etiology. The most important division is between epilepsy with a recognizable cause (symptomatic) or epilepsy without a recognizable cause (cryptogenic or idiopathic). Other factors also important include: seizure type, age of onset, EEG abnormality, and associated neurological features. The syndromic classification currently used by neurologists was proposed by the International League Against Epilepsy.

8.21.3 PROGNOSIS OF EPILEPSY

Sander (1993) has defined the term prognosis, as applied to epilepsy, as meaning the chance of terminal remission once a patient has established a pattern of recurrent seizures. Obtaining evidence about the prognosis of epilepsy has been difficult because of the lack of definitional agreement and methodological inconsistencies between various studies. A recent review reported by Hauser and Headorffer (1990) that the probability of remission is anywhere between 41 and 75%. A community-based study by Annegers, Hauser, and Elveback (1979) following up 457 patients with a history of two or more nonfebrile seizures for a minimum of five years showed that the probability of being in remission lasting five years or more was 61% at 10 years and 70% at 20 years. It is now generally accepted that as many as 70–80% of people who develop epilepsy will achieve a remission, and that remission will in most cases occur shortly after antiepileptic drug treatment is initiated (Sander, 1993).

8.21.4 MEDICAL AND SURGICAL TREATMENT OF EPILEPSY

For over a century anticonvulsant drugs have been the mainstream of treatment for individuals with epilepsy. Early drug treatments included the use of extracts from human skulls, hippopotamus testicles, tortoise blood, and pigeon dung as well as a number of herbal remedies such as foxgloves or camomile flowers (Scott, 1993). Therapy reached a somewhat more sophisticated level after the use of bromides was introduced in 1857. Though first synthesized in 1864 barbituric acid, phenobarbitone, was not put into clinical practice until 1912. Phenobarbitone was followed in 1938 by the introduction of phenytoin, carbamazepine in 1954, and sodium valproate in 1974. Recently, a number of drugs have been marketed including vigabatrin, lamotrigine, and gabapentin and a number of others, topirimate, tiagabine, and levetiracetam, are in phased development.

The drug prescribed depends largely on the type of epilepsy—the aim being to control seizures with the simplest drug regime and the fewest side effects. The majority of patients can be successfully treated by monotherapy, although in some cases polytherapy is required in order to achieve satisfactory seizure control. Though the varied and unpredictable risk of side effects associated with polytherapy has made it unfashionable in the last decade (Brodie, 1992; Shorvon & Reynolds, 1979), the development of new antiepileptic drugs (AEDs) with better defined mechanisms of action may increase its acceptability among clinicians, as they identify combination regimens with synergistic properties. A list of the main anticonvulsants, indications for their use, and their common side effects is given in Table 2.

Though the side effects of AEDs are generally minor in patients with less severe epilepsy which is controlled by a single drug, they may impose significant impairments in patients with chronic epilepsy. The frequency with which patients perceive side effects associated with their medication has recently been documented (Baker, Jacoby, Buck, Spalgis, & Morret, 1997). Five thousand members of epilepsy support groups throughout Europe were sent questionnaires through the post which asked them abut their epilepsy, the treatment they had for it and its impact on their day-to-day functioning. They were asked to report any problems or side effects which they felt had been caused by their AED medication and whether they had "always or often," "sometimes," "rarely," or "never" been a problem in the last month. Only 12% of respondents currently on AEDs reported no side effects at all, the remainder reporting at least one. The most commonly experienced side effects were tiredness, memory problems, difficulty concentrating, sleepiness, difficulty thinking clearly, and nervousness or agitation. Table 3 shows, for

Table 2 Main antiepileptic drugs, indications for use, and common side effects.

Generic name	*UK trade name*	*Indications for use*	*Adverse effects*
Carbamazepine	Tegratol	Drug of choice: complex partial, tonic–clonic, and simple partial seizures	Dose-related: dizziness, double vision, unsteadiness, nausea, and vomiting
Clobazam	Frisium	Occasional use: tonic–clonic and partial seizures	Dose-related: drowsiness and sedation
Clonazepam	Rivitol	Drug of choice: status epilepticus	Dose-related: sedation, inflammation of veins if given intravenously
Diazepam	Valium	Drug of choice: status epilepticus	Dose-related: sedation
Ethosuximide	Zarontin	Drug of choice: simple absence	Dose-related: nausea, drowsiness, dizziness, unsteadiness, may exacerbate tonic–clonic seizures. Allergic: rashes
Phenobarbitone	Prominal	Effective in: tonic–clonic and partial seizures	Dose-related: drowsiness, unsteadiness. Allergic: rashes. Chronic toxicity: tolerance habituation, withdrawal seizures, adverse effects on cognitive functioning
Phenytoin	Epanutin	Drug of choice: tonic–clonic, simple, and complex partial seizures	Dose-related: drowsiness, unsteadiness, slurred speech, occasionally abnormal movement disorders. Allergic: rashes, swelling of lymph glands, hepatitis. Chronic toxicity: gum swelling, acne, coarsening of facial features, hirsuitism, folate deficiency
Primidone	Mysoline	Occasional use: tonic–clonic and partial seizures	Dose-related: drowsiness, unsteadiness. Allergic: rashes. Chronic toxicity: tolerance, habituation, withdrawal seizures, adverse effects on cognitive functioning
Sodium valproate	Epilim	Drug of choice: all idiopathic generalized epilepsies. Effective in partial and secondary generalized seizures	Dose-related: tremor, irritability, restlessness, occasional confusion. Allergic: gastric intolerance, liver damage. Chronic toxicity: weight gain, hair loss

Source: Chadirck and Usiskin (1987).

those respondents on monotherapy only, the percentage reporting each problem according to the antiepileptic medication they were taking. Forty-four percent of respondents said they worried "a lot" or "some" about the possible side effects of their medication. A third had changed their AED medication at least once in the last year because of associated side effects; 5% had changed medication at least three times.

Though AEDs continue to be the mainstay of epilepsy treatment, there is a small but significant proportion of patients with intractable epilepsy for whom surgery is now an available option. There are a number of different surgical procedures available for intractable epilepsy, including amygdalahippocampectomy (partial removal of the amygdala and the hippocampus), temporal lobectomy (removal of a significant amount of the temporal lobe), corpus callosotomy (partial separation of the two hemispheres), and hemispherectomy (removal of significant amount of one hemisphere of the brain). Around 20% of patients with difficult-to-control seizures will prove suitable candidates for epilepsy surgery and pooled studies of the outcome for those operated on show that

Table 3 Subjects' reports of AED* side effects by type of antiepileptic (percentage reporting side effects always, often, or sometimes in the last month; subjects on monotherapy only).

Reported side effect	*CBZ* *(n=944)* (%)	*VPA* *(n=556)* (%)	*PHT* *(n=387)* (%)	*PB* *(n=112)* (%)
Unsteadiness	24	20	29	21
Tiredness	54	49	55	45
Restlessness	25	24	33	31
Feelings of aggression	18	18	22	18
Nervousness/agitation	27	31	37	31
Headache	30	25	27	24
Hair loss	11	17	11	11
Problems with skin	19	15	20	20
Double/blurred vision	20	13	17	12
Upset stomach	14	20	15	19
Difficulty concentrating	41	38	45	33
Trouble with mouth/gums	15	14	33	16
Shaky hands	21	37	28	18
Weight gain	19	36	17	23
Dizziness	22	20	25	13
Sleepiness	40	35	38	31
Depression	23	26	28	19
Memory problems	47	37	48	30
Disturbed sleep	19	21	31	23
Difficulty thinking clearly	35	30	40	28
Slurred speech	16	15	24	12

Source: Baker, Jacoby, Buck, Stalgis, and Monnet (1997).

CBZ, carbamazepine; VPA, sodium valproate; PHT, phenytoin; PB, phenobarbitone.

Figures in brackets are the base numbers on which percentages were calculated, and exclude the small number of respondents who did not answer these questions.

around 70% will become seizure free (Engel, van Ness, Rasmussen, & Ojemann, 1993). This radical clinical improvement in their condition is not always accompanied by immediate psychosocial improvements, as a number of studies have shown (Bladin, 1992; Ferguson & Rayport, 1965; Rausch & Crandell, 1982). This is probably because of the difficulties inherent in adjusting to becoming seizure free after a significant period of time with very frequent and unpredictable seizures. Following surgery, both the patient's expectations and those of his or her family, that he or she will now lead a normal and fulfilling life, are often unattainable in the short term. Rehabilitating patients who have undergone surgery requires the input of a multidisciplinary team, whose aim should be to help the operated individual and their family set realistic goals and make a positive adjustment to their newly-acquired status.

8.21.4.1 Adherence with Medical Treatment for Epilepsy

The failure to adhere to prescribed AED regimes has been cited as a common reason for poor seizure control (Chadwick, 1993; Cramer & Mattson, 1991). Estimates of the level of nonadherence range from around a third of patients admitting to missing medication occasionally (Buck, Baker, Jacoby, Smith, & Chadwick, 1997) to almost two-thirds (Dawson & Jameson, 1971). Nonadherence may be a particular problem in adolescents (Buck et al., 1997; Takaki, Klinokaura, & Aoyama, 1985). It tends to be poorer with polytherapy and may be enhanced by simplifying treatment regimes. Cramer, Scheyer, and Mattson (1990) reported that adherence to AED therapy declined significantly between clinic visits, reflecting what has been dubbed "white-coat compliance," and indeed a number of studies show that adherence can be improved by increasing the frequency of clinic visits (Wannamaker, Morton, Gross, & Sanders, 1980) and introducing educational programs (Helgeson, Mittan, Tan, & Chayasirisobhon, 1990).

The extent to which treatment side effects contribute to nonadherence has been documented in the US Veterans Administration Study (Smith, Mattson, Cramer, Collins, Novelly, Craft, and Veterans Administration Epilepsy Cooperation Study Group, 1987). By the two-year end point for the study, 85 of the 421 patients originally enrolled had a drug failure because of intolerable side effects. Adhering to treatment regimes is, however, more complex than simply considering the issue of immediate side effects. Patients appear to perform their own cost-benefit analysis of adherence, based on their understanding of their condition and the importance of treatment, and the information they receive not only from their doctor but also from other informal sources such as friends, family, and the media (Trostle, 1988; Trostle, Hauser, & Susser, 1983). Schneider and Conrad (1983) found that patients altered their AED regimes either for the purpose of testing the progress of their epilepsy and whether in fact they still needed to take them, or because they resented being dependent on medication and wanted to reassert control over their condition. Scambler (1989) reported that subjects in his study who reduced or abandoned AED therapy did so either because they were concerned about possible immediate or longer-term side effects, because they recognized that AEDs could only control not cure their condition, and because continuing to take AEDs implied continuing epilepsy and the stigma associated with it.

8.21.5 IMPACT OF EPILEPSY ON PERSON AND FAMILY

The World Health Organization (WHO) identifies three distinct concepts as consequences of a chronic condition such as epilepsy: impairment, disability and handicap (WHO, 1980). Seizures, the manifestation of impairment in brain function associated with epilepsy, are only temporarily, even though dramatically, disabling. For the rest of the time, the majority of people with seizures are able to perform their normal activities without any resultant restriction or loss of ability. However, the stigma of epilepsy and associated discrimination can represent significant handicaps for them, for example, in the spheres of occupation and economic self-sufficiency. People with epilepsy are therefore often handicapped, even if not disabled by their condition, and the quality of their lives will be significantly reduced as a result. The impact of epilepsy on individuals who have it can be felt physically, psychologically, and socially. Epilepsy is also known to sometimes have a major impact upon the families of people who have it.

8.21.5.1 Impact on Physical Function

At its most extreme, epilepsy is associated with increased mortality (Cockerell et al., 1994; Hauser & Hesdorffer, 1990; Hauser & Kurland, 1975). Though patients who have infrequent seizures have a standardized mortality ratio (SMR) similar to the general population, it is higher in those whose seizures are frequent. Survival is much less for those individuals where a cause for their epilepsy can be identified, and higher mortality has been reported for some types of seizures than for others. Among people who have generalized tonic–clonic seizures, the mortality rate is 1.5–3 times greater than in the general population, whereas among people who have absences or complex partial seizures, the rate is the same (an unsurprising finding given the differing presentations of these different seizures). While increased mortality may be due in some cases to the underlying causes of epilepsy, it is also the case that deaths from accidents and trauma are higher. For example, people with seizures are twice as likely to die from drowning as people without them. There is some evidence that death from suicide is more common among people with epilepsy (Barraclough, 1981; Hawton, Fagg, & Marsack, 1980; Zielinski, 1994), though this is not confirmed by other studies (Cockerell et al., 1994; Hauser, Annegers, & Elveback, 1980). There is a significant literature on the risk of sudden unexpected death (SUD) in people with epilepsy (Annegers, Hauser, & Shirts, 1984; Eamest, Thomas, Eden, & Hossack, 1992; Leestma, Kalelker, Teas, Jay, & Hughes, 1984) and some evidence that the risk of SUD is increased around the time of unusually stressful life events, perhaps as a result of more erratic medication taking at such times. The risk of nonfatal accidents including fractures and burns is also higher for people with epilepsy than others (Buck et al., 1997; Nakken, & Lossius, 1993; Spitz, Towkin, Shantz, & Adler, 1994).

Less dramatically, people with epilepsy have to contend with the side effects of antiepileptic medications, which may be both cognitive and physical. Side effects may be more common with polytherapy and are apparently less common with the newer drugs—though it has been suggested this is no more than a reflection of the fact that they have been less well investigated (Shorvon, 1987). It has been argued that all major antiepileptic drugs have adverse cognitive effects, including reduced attentiveness, impoverished memory, and mental slowing (McQuire & Trimble, 1990), and that patients with more cognitive impairment also have more psychosocial problems (Dodrill, 1980). In addition, people with epilepsy, particularly those with a more severe disorder or on higher levels of drugs, may experience changes in their physical appearance including weight gain, hirsuitism, thinning of the hair, and gum hyperplasia.

8.21.5.2 Impact on Psychological Function

Reviewing studies of the psychiatric and psychological outcomes of epilepsy, Fenwick (1987) has estimated that the prevalence of psychiatric morbidity is high, with around a third of people with epilepsy affected. Hermann and Whitman (1986) hypothesize three alternative groups of variables which contribute to psychopathology in people with epilepsy: *psychosocial* variables such as fear of seizures, perceived stigma and discrimination, degree of adjustment to the diagnosis, other life events, and level of social support; *clinical* variables such as age of onset, duration of epilepsy, seizure type, and seizure severity; and *medication* variables including medication type and number. Empirical investigation has so far tended to focus on the role of clinical variables but there is evidence that both antiepileptic medication and psychosocial factors also play an important part in the development of psychopathology (Baker, Jacoby, & Chadwick, 1996; Jacoby, Baker, Steen, Potts, & Chadwick, 1996).

Anxiety and depression are the two most common manifestations of psychopathology reported among people with epilepsy. Betts (1981) lists a whole range of ways in which anxiety and depression may be related to epilepsy (Tables 4 and 5). Anxiety is the problem most commonly elicited from patients themselves (Arnston, Drodge, Norton, & Murray, 1986; Collings, 1990; Smith, Baker, Dewey, Jacoby, & Chadwick, 1991). In an unselected population of people with epilepsy ranging from those with complete seizure control to those with longstanding drug-resistant epilepsy, a quarter were identified, using the Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983), as "cases" of anxiety and a further 15% as "borderline," and there was a very clear relationship in this population between anxiety and seizure frequency, with the percentage of cases rising from 13% among people who had had no seizures in the last year to 44% among those who were having at least one a month (Jacoby et al., 1996).

It has been suggested that protracted anxiety may be the precursor of depression, and the two problems commonly co-exist in epilepsy (Robertson, Trimble, & Townsend, 1987). Depression was the most common reason for admission to hospital for psychiatric care in a study of people with epilepsy (Betts, 1974). In the community study cited above (Jacoby et al., 1996), 9% of subjects were classified as depressed and a further 15% were "borderline" and as for anxiety, the percentage of subjects thus classified rose with seizure frequency, from 4% of those currently seizure free to 21% of those with seizures more often than once a month on average. Betts (1981) suggests that depression can be self-reinforcing and its associated sequelae, such as loss of confidence and reduced self-esteem, can be more disabling and longer-lasting than the depression itself.

Table 4 Anxiety and epilepsy.

1. Anxiety reaction to acquiring the label of epilepsy
2. Anxiety reaction to the social and family problems of epilepsy
3. Prodromal anxiety before a seizure
4. Anxiety as an aura
5. Anxiety as an ictal experience
6. Anxiety occurring in association with an epileptic psychosis
7. Anxiety occurring in association with epilepsy-related brain damage
8. True phobic anxiety related to seizures

Source: Betts (1981).

Table 5 Depression and epilepsy.

1. Depressive reaction to acquiring the label of epilepsy
2. Depressive reaction to social or family problems of epilepsy
3. Prodromal depressive feelings before a seizure
4. Depressive feelings as an aura
5. Depressive feelings as an ictal state
6. Postictal depressive feelings
7. Depressive twilight state
8. Epileptic depressive delirium
9. Endogenous depression unrelated directly to seizures, but possibly to their decrease in frequency
10. Depressive symptoms occurring in association with other mental illnesses, particularly a paranoid or schizophrenic psychosis

Source: Betts (1981).

Two other common psychological problems for people with epilepsy are reduced self-esteem and poor sense of mastery. Potential sources of reduced self-esteem are parental overprotection and rejection, educational and occupational underachievement, and perceived stigma and discrimination. However, the relationship between epilepsy and self-esteem has been relatively little researched to date. Collings (1990) compared self-esteem in a UK population of people with epilepsy and age- and sex-matched controls, and found significant differences on 11 out of 20 self-esteem ratings, with people with epilepsy seeing themselves as *less* valuable, adaptable, dependable, mature, stable, able to cope with life, successful, well-adjusted, or happy, and *more* unpredictable and handicapped. An important finding from Collings" study was that the most important correlate of overall well-being among the epilepsy sample was the degree of discrepancy between their current self-perception and anticipated self-perception without epilepsy. Smith et al. (1991) found patient-perceived seizure severity to be an important predictor of self-esteem among adult patients whose seizures were intractable, raising the question of whether improved seizure control will be accompanied by increased self-esteem. In the same study, Baker, Smith, Dewey, Jacoby, and Chadwick (1993) reported that self-esteem was significantly related to patients" scores on other psychological dimensions, including anxiety, depression, and sense of mastery.

Perhaps because of the unpredictable nature of seizures, people with epilepsy tend to have a poorer sense of mastery than healthy people and a number of authors (Arnston et al., 1986; Baker et al., 1993; Matthews & Barabas, 1981) have shown that sense of mastery is significantly related to other psychological variables. For example, Matthews and Barabas (1981) found that poor sense of mastery in adults with epilepsy was associated with anxiety, poor self-esteem, feelings of helplessness, and an increased risk of suicide. From their study of 15 children with epilepsy, these authors concluded that their perceptions of control could affect their social functioning, their self-concept and their motivation to learn, and may predispose them to anxiety and depression. Similarly, Arnston et al. (1986) showed that a poor sense of mastery in adults with epilepsy was significantly and positively related to anxiety, depression, and a number of somatic symptoms: the more their respondents believed that their health was controlled by forces external to them, the more anxious, depressed, and ill they felt.

The role of the family in reducing sense of mastery in a person with epilepsy has been examined by a number of authors. For example, Ziegler (1982) considered the way in which seizures distorted family negotiations relating to a child's competence and autonomy and concluded that other family members, parents in particular, need a way to deal with the feelings of helplessness and loss of control they themselves experience when the affected person begins to have seizures. Since seizures rupture a person's control, Ziegler maintains that the other family members may become reluctant to grant him or her autonomy outside the seizures; their apparently transient effects thus acquire a much deeper significance in negotiations between the affected person and his/her family, continuing to repattern what has been called "the family dance" (Napier & Whitaker, 1978).

8.21.5.3 Impact on Social Function

Historically, the social meaning of epilepsy was that it was a disease to be feared and shunned and the treatment of people with epilepsy has throughout history been a reflection of this (Temkin 1971). In ancient and primitive societies, epilepsy was seen as a sign of sin, demonic possession, and madness, and although it is now recognized as a disorder of the brain rather than the soul, the old images of epilepsy to some extent still persist. Certainly many people with epilepsy are only too well aware of its potential for stigma, so that having epilepsy may be seen as worse than having seizures (Scambler, 1989; Schneider & Conrad, 1983). Scambler (1989) makes a distinction between "enacted" stigma, that is, episodes of discrimination against people with epilepsy and "felt" stigma, which he defines as an oppressive fear of enacted stigma. He showed that felt stigma was much more prevalent among people with epilepsy than enacted stigma—a finding confirmed in a recent large-scale quantitative study of people with well-controlled epilepsy (Jacoby, 1994). In the latter, only 2% of subjects could recall any actual event of discrimination in the workplace, but 32% felt that epilepsy made it more difficult for them than for people without epilepsy to find employment. The issue of stigma represents an important feature of the management and rehabilitation of people with epilepsy.

One way in which the potential for stigma may impact upon people with epilepsy is by reducing the level of their social contacts and activities. Social isolation is frequently the product of anxiety over the possibly hostile reactions of others if a seizure should occur in a public place. However, Mittan (1986) reports that fear for their physical safety also keeps many people isolated—more than a fifth of the people in his study said they were afraid to go out because of the possibility of seizures occurring which might lead to accidents and their most common coping strategy was therefore to stay at home. It has been suggested by Scambler and Hopkins (1980) that such fears lead to a self-denial of opportunities that can, in turn, result in a whole host of problems with personal relationships. These problems in personal relationships are reflected in the lower rates of marriage and fertility reported among people with epilepsy compared with the general population. Lechtenberg (1984) reported that 56% of the men he surveyed and 69% of the women were unmarried, and Collings (1990), comparing epilepsy and nonepilepsy samples in the UK, found that 71% of the nonepilepsy group were married, but only 42% of the epilepsy group.

There is some evidence that social withdrawal and isolation may also be the product of parental reactions to a diagnosis of epilepsy, which typically involve overprotectiveness (Ferrari, Matthews, & Barabas, 1986; Wolf, Thorbecke, & Even, 1986). Such overprotectiveness may lead parents to limit their child's activities, which in turn renders him or her more socially inept. Children with epilepsy have been shown to have increased levels of dependency and attachment to their parents (Mulder & Surmeijer, 1977; Stores & Piran, 1978). They have also been found to adopt a more passive role in family interactions and a reduced level of involvement in family decision-making (Ritchie, 1981). There is evidence that the parents of children with epilepsy, as well as being overprotective, have lowered expectations about their potential achievements. Ferrari (1989) investigated the expectations and attitudes of parents of children with epilepsy in 21 families and found that mothers had uniformly lower expectations of them than of their healthy siblings—their children with epilepsy were seen as less reliable, less trusting of others, less adept, more moody, more likely to have emotional problems, and more likely to perform poorly at school. Ferrari concludes that parents" responses to epilepsy can greatly influence their expectations for their child and play a significant role in their underachievment and adjustment problems.

Both under- and unemployment have been shown to be more common in people with epilepsy. One author (Floyd, 1986) has suggested that up to half of all people with epilepsy living in the UK may have experienced problems over their employability, and in the US, pooled data from a number of studies shows unemployment rates for people with epilepsy are twice that for the general population (Hauser & Hedorffer, 1990). Collings (1990) reported that whereas only 14% of a nonepilepsy sample were unemployed, 41% of those with epilepsy were, and long-term unemployment was more common in the sample of people with epilepsy—64% of them had been unemployed for more than two years, compared with only 13% of the nonepilepsy group. Elwes, Marshall, Beattie, and Newman (1991) also reported that people with epilepsy experienced significant difficulties in obtaining work, a finding which may in part be attributable to their educational underachievement. Elwes found that respondents in his study were less likely to leave school with any formal qualifications or to undergo any subsequent training apprenticeships. A number of clinical features of epilepsy appear related to the likelihood of being employed, including the failure to achieve total seizure control, seizure type, and the presence of multiple neurological or non-neurological handicaps.

UK community studies have generally shown that manual occupations are over-represented among people with epilepsy, suggesting that underemployment is also a problem for them. In the US, a study in Rochester, Minnesota, found a lower proportion of people with epilepsy than others employed in professional and managerial occupations and a higher proportion employed in lower-level occupations such as manual labor or machine operation (Trostle, 1987). It has been suggested that the under-representation of people with epilepsy in nonmanual occupations can be accounted for by the theory of "social drift," which proposes that people who are chronically ill or disabled are unable to maintain their original occupational status and drift downwards to jobs of lower status. This theory was examined by Scambler and Hopkins (1980) who report that among the community sample they followed, they were unable to find any evidence of significant intragenerational drift downwards as a result of their condition: they did find, however, a small but significant association between social class and current employment status which suggested that respondents from working class households were more vulnerable to unemployment.

It is important to note that from the point of view of their employability, people with epilepsy are not a homogeneous group. Those with severe epilepsy who may also have associated physical or mental handicaps and behavioral disorders are likely to be unemployable on the open market; those whose seizures are infrequent but not fully controlled should be employable given a suitable job and a sympathetic employer, and those whose epilepsy is controlled or in remission should be employable in any job except one requiring a completely clear history (Espir & Floyd, 1986).

One obvious implication of underemployment and unemployment is that it has an impact on a person's financial status. A study addressing this issue was done by Arnston et al. (1986) in the US. The authors noted a marked discrepancy between the educational achievement of the group they studied and their household income, the reported level of which was far below the national average. Clemmons (1983) reported that 50% of a US sample of people with epilepsy were dependent on family or federal subsidy. Other studies which have shown that people with epilepsy generally have lower than average incomes include those by Batzel, Dodrill, and Fraser (1980); Fraser, Clemmons, Trejo, and Temkin (1985); and Dodrill, Breyer, Diamond, Dubinsky, and Geary (1984). There are other broader implications of the reduced employability of people with epilepsy. The Reid Report on people with epilepsy (Central Health Services Council, 1969) highlighted the importance of work in determining not only financial status, but also social status and societal role, aspects of personal satisfaction and self-esteem, sense of purpose and social companionship, and recently, research has indicated that there is a significant relationship employment status and the ability of people with epilepsy to adjust to their condition (Collings, 1990; Hermann, Whitman, Wyler, Anton, & Vanderzwagg, 1990).

A number of the studies cited above of the psychosocial status of people with epilepsy focused on prevalent populations of individuals with resistant epilepsy drawn from hospital clinics who represent only between 20–30% of all people with epilepsy. This has biased our understanding of the impact of epilepsy and its treatment, presenting a somewhat gloomy picture. Recently, research has been published which helps to redress the balance. Chaplin, Yepez Lasso, Shorvon, and Floyd (1992) studied an unselected sample of 192 adults aged 17 years and over with a recent diagnosis of epilepsy. They investigated 14 different areas of psychosocial adjustment using a previously developed 42-item scale, and found that problems were experienced by almost all subjects in at least one area. However, in contrast with findings from studies of people with chronic epilepsy, problems in this group of newly diagnosed patients were generally mild (Table 6). There were only four areas where more than 10% of subjects reported severe problems: fear of seizures, fear of stigma in employment, adverse effects on leisure, and lack of energy. Chaplin suggests that the findings that the diagnosis of epilepsy did not of itself seem to lead to high levels of maladjustment calls into question the view that the label "epilepsy" has a powerful and constant stigmatizing effect, and argues that the fact that adjustment appears, from earlier studies, to be related to seizure frequency, weakens the case for the label's importance. We would suggest that an alternative interpretation is that infrequent seizures reduce the number of potentially discrediting situations but do not remove the ever-present threat of being discredited, and in those with newly diagnosed seizures, the full impact of this recently-acquired status on their identity and aspects of their daily life has yet to be realized.

A recent study of people with well-controlled epilepsy also presents a more positive picture of its psychosocial consequences. The Medical Research Council study of antiepileptic drug withdrawal was a multicenter clinical trial which compared clinical and psychosocial outcomes in groups of patients randomized to slow withdrawal from or continued treatment with antiepileptic medication (MRC Antiepileptic Drug Withdrawal Group, 1991). The study yielded a considerable quantity of information on a large cohort of individuals with epilepsy in remission. To be eligible for the trial, patients had to be seizure free for a minimum of two years and many had been so for considerably longer. In this group, levels of psychosocial distress were low. Respondents emerged as well-adjusted to their condition and experiencing few problems because of it. Their psychological profile compared to that of other epilepsy samples and normative populations was positive; rates of employment were comparable with those of the general population; rates of marriage were higher than in groups of people with severe epilepsy and the proportion living alone was the same as in the general population. There was little evidence of social isolation and few people appeared to feel stigmatized by their epilepsy. For the majority epilepsy did not appear to profoundly diminish their ability to function physically, psychologically, or socially (Jacoby, 1992, 1995). The findings from this study are important, because people with well-controlled epilepsy represent the largest group of people with epilepsy—between 70 and 80% of all those affected by the condition.

8.21.5.4 Impact of Epilepsy on Other Family Members

As well as affecting the physical, psychological, and social status of people with the condition, epilepsy also has implications for other family members. Families of children with epilepsy have been shown to be significantly less cohesive, have lower levels of esteem and communication, and have reduced levels of social support compared with those of children with other chronic conditions (Austin, 1988; Ferrari et al., 1983). The presence of a child with epilepsy in the family appears to have important repercussions for all members, and particularly for mothers and siblings. Rutter, Graham, and Yule (1970) found that having a child with epilepsy was a potent source of distress among their mothers, of whom a fifth in his study had experienced psychiatric problems. Sillanpaa (1973) reported higher rates of divorce among the parents of children with epilepsy. The siblings of children with epilepsy also appear

Table 6 Percentages of newly diagnosed patients reporting problems in psychosocial adjustment and their severity.

	Severity of problem			
Areas of adjustment	*None* (%)	*Mild* (%)	*Moderate* (%)	*Severe* (%)
Acceptance	36	36	20	7
Fear of seizures	20	8	57	15
Employment	31	28	27	13
Future	47	24	28	1
Travel	71	15	12	2
Social life	56	15	20	9
Leisure	57	12	21	10
Outlook on life	51	21	20	8
Family	65	19	14	2
Medication	50	19	27	4
Medical profession	63	16	16	4
Emotions	53	25	19	3
Isolation	68	13	17	2
Lack of energy	57	14	18	11

Source: Chaplin et al. (1992).

to be at greater risk of psychiatric disturbance (Hoare & Kerley, 1991). Finally, Betts (1988) has pointed out that there may be something of a "cycle of deprivation" in the functioning of families with a member with epilepsy since "the child that has had poor parenting as a result of his or her own epilepsy is likely to make a poor parent in turn."

8.21.6 COPING WITH THE IMPACT OF EPILEPSY

The ways in which people with epilepsy cope with its actual or perceived negative impact on their concepts of self and their lives was examined by Schneider and Conrad (1981) in interviews with 80 people, some of whom were experiencing enormous problems in their daily lives because of their diagnosis, others of whom saw it as "no big thing." It was interesting that these divergent perceptions of its impact were not a direct function of the medical severity of their condition: people who shared similar seizure experiences and histories nonetheless held very different views about their impact.

From the interviews, Schneider and Conrad (1981) were able to develop a typology of the characteristic ways in which their subjects handled and adapted to their condition. The typology distinguishes between two broad forms of adaptation, called by the authors the "adjusted" and the "unadjusted" within each of which there are a number of subtypes. One identifiable subtype of adjusted adaptation was the *pragmatic* type, wherein the predominant strategy for dealing with epilepsy was to minimize it as much as possible, both to one's self and to others. People who opted for this pragmatic approach, while recognizing the potential for negative reactions, were generally open about their epilepsy with others. In contrast, a second subtype of adjusted adaptation was the *secret* type, wherein epilepsy was managed by often elaborate procedures to control and conceal information about it. This type of adaptation rested on a belief in the stigmatizing quality of epilepsy and the extent of the secrecy was dependent on the intensity of the perceived stigma. A third adjusted subtype was termed by the authors the *quasiliberated*, sharing with the pragmatic a straightforward definition of oneself as having epilepsy and broadcasting the fact to others in an attempt to educate them and be free of the burdens of secrecy based on the fear of stigma. A fundamental difference between these adjusted adaptations to epilepsy and the unadjusted adaptations was that whereas the former was characterized by a sense of control, the latter was marked by a sense of being overcome or overwhelmed by epilepsy. The most extreme subtype of this unadjusted adaptation was what the authors called the *debilitated* subtype wherein epilepsy represented "an indelible and irrevocable threat to one's worth." People who were debilitated by their epilepsy were often, it appeared, the partial and unwitting engineers of their own fate, since they allowed their condition to dominate their identity and life, withdrawing into it and away from others.

Scambler (1984) applied the typology of adaptation outlined above to the population of people with epilepsy in his London study. He concluded that the subtype of adjusted adaptation referred to as the secret type represented the preferred strategy for the majority of his subjects, though it rarely brought them any lasting peace of mind. The second-choice strategy was that termed by Schneider and Conrad (1981) the pragmatic type, which several of Scambler's subjects reported they resorted to from time to time. Only a tiny percentage had employed the quasiliberated strategy, a finding Scambler attributes to cultural differences in the militancy and organization of the disability movement in the US and Britain. Finally, Scambler found that the debilitated subtype of unadjusted adaptation, as outlined by Schneider and Conrad, was, among the group he studied, a temporary rather than a permanent type of adaptation, enacted only at particular times and in particular circumstances when their status as "epileptic" became more than usually apparent and threatened the equilibrium of their lives.

The adaptation strategies employed by people with epilepsy as a means of reducing its impact are mirrored in other stigmatizing conditions. For example, Link, Mirotznik, and Cullen (1991) studied the way in which people with mental illness try to avoid its negative consequences and reported that secrecy, selective avoidance, or withdrawal from social interaction, and educating or "preventive telling" were common coping strategies. However, it appeared that none of these strategies was fully effective in diminishing the negative impact of their condition on employment or on levels of psychological distress. Scambler (1993) suggests that to fully understand how people cope with epilepsy, five dimensions of coping must be addressed: accommodation, that is, an individual's reaction to the physical properties of epilepsy and the iatrogenic effects of medication; rationalization, that is, the process by which people make sense of their epilepsy; their concept of self; sociability, that is, the quality and extent of the relationships negotiated by the person with epilepsy and significant others; and finally, fulfilment, an individual's sense of achievement which may be expressed in relation to any of his or her social roles and for which epilepsy may constitute a serious obstacle.

8.21.7 LIFE SPAN ISSUES

The incidence of epilepsy is highest in childhood and older age. It is the most common neurological problem of childhood, its incidence being highest in the first decade of life (Hauser & Hesdorffer, 1990). Most childhood epilepsy is benign and will remit in adolescence, but some severe syndromes persist into adulthood; these syndromes are often characterized by significant brain dysfunction and developmental delay (Dreifuss, 1995). There is also a group of syndromes that begin in adolescence, including juvenile absences and juvenile myoclonic epilepsy (Dravet, 1992). The effects of AED medication may be particularly significant in childhood because of their influence on the developing brain and many studies have examined the influence of AEDs on emotional and behavioral problems, cognitive functioning (including impairments of attention, concentration, and memory), and academic performance (Aldenkamp, 1995; Thompson, 1995; Trimble, 1988). It has been shown that children with epilepsy have a relatively more compromised life quality than those with other chronic conditions (Austin, Shelton, Risinger, & McNelis, 1994). The development of epilepsy can have a profound effect on a child's sense of self and this may be particularly significant at adolescence, in part because this is the point when the legal and statutory restrictions imposed by a diagnosis of epilepsy are first felt, and in part because adolescence is a time for consideration of one's identity, sexuality, and independence and epilepsy can severely disrupt this normal development.

The incidence of epilepsy is at its lowest between the ages of 20 and 60 and in the majority of adults in this age group it is well controlled. Despite the absence of seizures, such individuals still face the fear of further attacks and the psychosocial consequences of the label of epilepsy. As noted earlier in this chapter, adults with epilepsy have been shown to have poorer psychosocial functioning than the general population (Levin, Banks, & Berg, 1988). They also, like children, have to contend with the side effects of medication and the disruption to daily functioning that occurs as the result of such effects.

Epilepsy is the third most common neurological condition of old age after stroke and dementia and Tallis (1993) comments that the onset of the condition can represent a watershed in a person's life, with both immediate and longer-term impact on confidence, independence, and quality of life. Making a diagnosis of epilepsy in older people is more hazardous because seizures sometimes present differently in this age group and the range of potential causes is much wider. The after-effects of seizures may be more prolonged and may often be associated with more serious injury. The central nervous system-related effects of AEDs may further exacerbate any existing decline in cognitive function in older people, who are also more likely to be taking other non-AED medications and so face an increased risk of drug interactions (Jones, 1988; Troupin & Johannessen, 1990. Tallis (1993) has described the social effects of epilepsy in older age as embodying a process of "marginalization and disempowerment" which can profoundly undermine their concept of self.

8.21.8 GENDER ISSUES

Contraception, pregnancy, and parenthood represent a number of concerns for women with epilepsy. While oral contraception remains the most effective form of contraception, it is known to interact with enzyme-inducing AEDs (e.g., phenytoin and carbamazepine) and consequently a contraceptive with a higher dose of estrogen may be required to prevent contraceptive failure. A number of recommendations about the clinical management of women who are pregnant or contemplating pregnancy have been made. These include: consideration of AED withdrawal in women seizure free for at least two years; reduction of polytherapy and use of carbamazepine monotherapy where possible; early screening for neural tube defects and prescription of folic acid supplements; and reassurance about the acceptability of breast-feeding while taking AEDs (Appleton, 1994).

8.21.9 PSYCHOLOGICAL INTERVENTIONS IN THE MANAGEMENT OF EPILEPSY

Managing the various psychological and social problems associated with epilepsy is crucial to making a positive adjustment to the condition. Failure to resolve such problems may reduce the affected person's motivation to comply with treatment regimens and impair their ability to make a successful adaptation. Relatively little attention has so far been paid to the contribution of the nonmedical management of epilepsy to dealing with such issues, perhaps as a result of the small number of specialist centers worldwide providing a multidisciplinary approach to its assessment and treatment (Goldstein, 1990). Nonmedical approaches to the management of epilepsy can be divided into two broad categories: those aimed at reducing seizure frequency and those which target the psychological adjustment of the individual with epilepsy.

8.21.9.1 Interventions Relating to Treatment of Seizures

A variety of relaxation strategies have been advanced in the management of seizures because of the recognition that anxiety may act as a seizure precipitant and that people with epilepsy are more likely to have increased rates of anxiety when compared with the normal population. Progressive muscle relaxation techniques have been used, with some success, in patients who are able to identify heightened physiological arousal prior to seizure onset and who experience anxiety interictally (Rousseau, Hermann, & Witman, 1985). While there is substantial anecdotal evidence that people with epilepsy will benefit from anxiety management, there are few randomized studies to support the efficacy of anxiety management techniques.

Recently, a number of studies have focused on applying a behavioral analysis to the management of seizure activity. These approaches have generally focused upon the antecedants and consequences of the seizure (Goldstein, 1990). Dahl, Melin, and Leissner (1988) have shown that direct manipulation of stimuli which are known to elicit seizures may lead not only to a reduction in seizure behavior but also a reduction in EEG activity. The package of treatment proposed by Dahl and colleagues includes the use of contingent relaxation, symptom discrimination, countermeasures for arousal level, and positive reinforcement for appropriate behavior. Unfortunately, the generalizability of the findings from this study was hampered by the small number of subjects ($N = 3$) and the absence of a control group. Tan and Bruni (1986) in a later study examined the use of cognitive behavioral treatment for the reduction of seizure activity. Their study was conducted over eight weeks and consisted of 2 h weekly sessions involving self-monitoring of stressful events and associated cognitions. Standard cognitive-behavioral treatment, however, produced little change in seizure frequency. Other researchers (Dahl, Melin, & Lund, 1987; Pritchard, Holmstrom, & Giacinto, 1985; Brown & Fenwick, 1989) have used covert desensitization programs with relaxation training and cue-controlled arousal. One example of such an intervention is where the patient was taught to imagine an aura in a relaxed state and then used an arousal technique to stop it. However, this approach has yet to be evaluated in a group study (Goldstein, 1990).

8.21.9.2 Interventions to Reduce the Impact of Epilepsy and Enhance Coping

The role of the clinical psychologist in helping people with epilepsy to accept the diagnosis and to self manage their condition is well recognized. Given that people with epilepsy may develop psychological problems at any stage in their illness career, it is important that a psychological service be routinely available to them in epilepsy outpatient clinics. People with epilepsy should be fully investigated to determine what factors are likely to be important in the development and maintenance of psychological problems associated with their condition. Obviously some problems will be amenable to psychological treatment only, whereas other more severe problems may require pharmacological treatment in addition to psychological therapy. If the epilepsy is long-standing and psychological–behavioral problems have been in existence for a considerable time, individuals may require long-term psychotherapy rather than time-limited behavioral treatments. Cognitive-behavioral therapy involves shaping the way people think, feel and behave and clinical anecdotal evidence suggests that it may be the most appropriate treatment for people with epilepsy. Even though the research evidence for its efficacy in this condition is limited at present, there is evidence for its success in other chronic conditions (Illis, 1994) Some studies have shown that psychoeducational programs are effective in helping patients to self-manage their epilepsy (Helgeson & Mittan, 1990).

It is important for people with epilepsy and their families to have a proper understanding of their condition and to demystify any unhelpful beliefs they might have about it (Mittan, 1986). Hills and Baker (1992) demonstrated how possession of accurate information about epilepsy was significantly related to level of well-being; people with epilepsy who were more informed about the management and treatment of their condition perceived themselves as being in more control and subsequently had significantly better psychological profiles than those who were less informed. In this arena too the clinical psychologist has a key role to play, the aim being to help patients minimize the impact of their condition on their daily lives. Interventions aimed at the cognitive level emphasize the importance of providing new information about the illness or its management. Cognitive-behavioral techniques are also implemented to modify the underlying core beliefs that people have about their condition which may be at the root of their psychological impairments. Interestingly, use of group-based cognitive-behavioral therapy in the work of Tan and Bruni (1986) did not significantly reduce seizure frequency but *did* result in an increase in patients' ratings of their well-being.

As well as working with patients themselves, psychologists will be called upon to work with their families, particularly where the epilepsy develops in childhood and adolescence. Families have a significant part to play in the way in which the affected person adjusts to their condition. For example, parents may become overprotective of their child with epilepsy when fearful of the risks involved if seizures cannot be completely controlled. They may dominate and prevent the child from gaining normal independence. Conversely, epilepsy may lead to the rejection of the child which in turn may lead to psychological and psychiatric problems for both protagonists. For some families, education and counseling will be sufficient to ameliorate these difficulties; for others more formal family therapy may be necessary. In family therapy, the whole of the family are invited to explore and modify communication systems and behavioral functioning in order to minimize the negative aspects of epilepsy. While there is little evidence of the efficacy of this approach in families with epilepsy, there is substantial evidence of its usefulness in other conditions (Minuchin, 1974).

Clinical psychologists are only one part of the multidisciplinary team managing patients with epilepsy, and other team members can also contribute to a positive psychological adjustment. Clinicians need to be circumspect in the timing of communicating a diagnosis of epilepsy, since diagnosis marks the beginning of a protracted and sometimes lifelong process of adjustment. A number of authors have highlighted the role of the hospital clinician in the adjustment process, both through rational drug therapy and through awareness of the possible psychological or psychiatric sequelae of the condition (Hermann & Whitman, 1991; Scambler, 1993) The general practitioner, as the clinician with often the most regular contact with the patient, also has a role to play in providing relevant information regarding the nature and management of epilepsy. Such information should be provided on a regular basis and should be updated, as far as is possible, to keep pace with important developments in the field. The general practitioner is also well-placed to direct people with epilepsy and their families to local and national support groups and to organize practical support, particularly for patients who are severely affected by their condition, such as respite care, home help, and holiday opportunities.

Finally, the epilepsy support agencies also have a role in educating and supporting the individual with epilepsy and their family. Information about counseling services and the location of specialist epilepsy centers can be obtained through most of the organizations for people with epilepsy. There are both local and national organizations covering the whole of the UK.

8.21.10 CONCLUSION

In this chapter, we have tried to provide the reader with an overview of epilepsy and its management, with particular emphasis on the impact it can have on patients' physical, social, and psychological well-being. The ways in which people with epilepsy seek to accommodate its effects on their daily lives have been documented in research, and we have highlighted the means by which clinical psychologists can help to make this accommodation a positive one. The clinical aspects of epilepsy are generally easily managed by present-day treatments, but the psychological aspects may require a significant input from health professionals. We believe that the clinical psychologist can make a considerable contribution to the amelioration of both clinical and, more importantly, nonclinical issues confronting the patient with epilepsy.

8.21.11 REFERENCES

Aldenkamp, A. P. (1995). Cognitive side effects of antiepileptic drugs. In A. P. Aldenkamp, (Ed.), *Epilepsy in children and adolescents* (pp. 161–182). Boca Raton, FL: CRC Press.

Annegers, J. F., Hauser, W. A., & Elveback, L. R. (1979). Remission of seizures and relapse in patients with epilepsy. *Epilepsia, 20,* 729–737.

Annegers, J. F., Hauser, W. A., & Shirts, S. B. (1984). Heart disease mortality and morbidity in patients with epilepsy. *Epilepsia, 25,* 699–704.

Appleton, R., (1994). *Epilepsy* (3rd ed.). London: Martin Dunitz.

Arntson, P., Drodge, D., Norton, R., & Murray, E. (1986). The perceived psychosocial consequences of having epilepsy. In S. Whitman & B. Hermann (Eds.), *Psychopathology in epilepsy: Social dimensions.* New York: Oxford University Press.

Austin, J. K. (1988). Childhood epilepsy: child adaption and family resources. *Journal of Child and Adolescent Psychiatric Mental Health Nursing, 1,* 18–24.

Austin, J. K., Shelton, M. S., Risinger, M. W., & McNelis, A. M. (1994). Childhood epilepsy and asthma: comparison of quality of life. *Epilepsia, 35,* 608–615.

Baker, G. A., Smith, D. F., Dewey, M., Jacoby, A., & Chadwick, D. W. (1993). The initial development of a health-related quality of life model as an outcome measure in epilepsy. *Epilepsy Research, 16,* 65–81.

Baker, G. A., Jacoby, A., & Chadwick, D. W. (1996). The associations of psychopathology in epilepsy: a community study. *Epilepsy Research, 25,* 29–39.

Baker, G. A., Jacoby, A., Buck, D., Stalgis, C., & Monnet, D. (1997). Quality of life of people with epilepsy: a European study. *Epilepsia, 38,* 353–362.

Barraclough, B. (1981). Suicide and Epilepsy. In E. H. Reynolds & M. R. Trimble (Eds.), *Epilepsy and psychiatry.* Edinburgh: Churchill Livingstone.

Batzel, L. W., Dodrill, C. B., & Fraser, R. T. (1980). Further validation of the WPSI vocational score: comparisons with other correlates of employment in epilepsy. *Epilepsia, 21,* 235–242.

Betts, T. A. (1974). A follow-up study of a cohort of patients with epilepsy admitted to psychiatric care in an English city. In P. Harris, C. Mawdsley, & H. Berger (Eds.), *Epilepsy: Proceedings of the Hans Bergen Centenary Symposium.* Edinburgh: Churchill Livingstone.

Betts, T. A. (1981). Depression, anxiety and epilepsy. In E. H. Reynolds & M. R. Trimble (Eds.), *Epilepsy and psychiatry.* New York: Churchill Livingstone.

Betts, T. (1988). People with epilepsy as parents. In P. Hoare (Ed.), *Epilepsy and the family: a Medical Symposium on New Approaches to Family Care.* Manchester: Sanofi UK Ltd.

Bladin, P. F. (1992). Psychosocial difficulties and outcome after temporal lobectomy. *Epilepsia, 33,* 898–907.

Brodie, M. J. (1992). Drug interactions in epilepsy. *Epilepsia, 33,* S13–S22.

Brown, S. B., & Fenwick, P. B. (1989). Evoked and psychogenic epileptic seizures: 2. Inhibition. *Acta Neurologica Scandinavica, 80,* 541–547.

Buck, D., Baker, G. A., Jacoby, A., Smith, D. F., & Chadwick, D. W. (1997). Patients' experiences of injury as a result of epilepsy. *Epilepsia, 38,* 439–444.

Central Health Services Council. (1969) People with Epilepsy: Report of a Joint Sub-Committee of the Standing Medical Advisory Committee and the Advisory Committee on the Health and Welfare of Handicapped Persons. London: HMSO.

Chadwick, D. (1994). Epilepsy. *Journal of Neurology, Neurosurgery and Psychiatry, 57,* 264–277.

Chadwick, D., & Usiskin, S. (1987). *Living with epilepsy.* London: Macdonald Optima.

Chadwick, D. W. (1993). Seizures and epilepsy in adults. In J. Laidlaw, A. Richens, & D. Chadwick (Eds.), *A textbook of epilepsy.* Edinburgh: Churchill Livingstone.

Chaplin, J. E., Yepez Lasso, R., Shorvon, S. D., & Floyd, M. (1992). National general practice study of epilepsy: the social and psychological effects of a recent diagnosis of epilepsy. *British Medical Journal, 304,* 1416–1418.

Clemmons, D. (1983). Relationship of general aptitude test battery scores to successful employment for epileptics in a rehabilitation setting. *Epilepsia, 24,* 232–237.

Cockerell, O. C., Johnson, A. L., Sander, W. A. S., Hart, Y. M., Goodridge, D. M. G., & Shorvon, S. D. (1994). Mortality from epilepsy: results from a prospective population-based study. *Lancet, 344,* 918–921.

Collings, J. A. (1990). Psychosocial well-being and epilepsy: an empirical study. *Epilepsia, 31,* 418–426.

Cramer, J. A., Scheyer, R. D., & Mattson, R. H. (1990). Compliance declines between clinic visits. *Archives of Internal Medicine, 150,* 1509–1510.

Cramer, J. A., & Mattson, R. H. (1991). Monitoring compliance with antiepileptic drug therapy. In J. A. Cramer & B. Spilker (Eds.), *Patient compliance in medical practice and clinical trials.* New York: Raven Press.

Dahl, J., Melin, L., & Lund, L. (1987). Effects of a contingent relaxation treatment program on adults with refractory epileptic seizures. *Epilepsia, 28,* 125–132.

Dahl, J. A., Melin, L., & Leissner, P. (1988). Effects of a behavioural intervention on epileptic behaviour and paroxysmal activity: A systematic replication of three cases of children with intractable epilepsy. *Epilepsia, 29,* 172–83.

Dawson, K. P., & Jamieson, A. (1971). Value of phenytoin estimation in management of childhood epilepsy. *Archives of Diseases in Childhood, 46,* 386–389.

Dodrill, C. B. (1980). Inter-relations between neuropsychological data and social problems in epilepsy. In R. Canger, F. Angeleri, & J. K. Penry (Eds.), *Advances in epileptology: XIth Epilepsy International Symposium.* New York: Raven Press.

Dodrill, C. B., Breyer, D. N., Diamond, M. B., Dubinsky, B. L., & Geary, B. B. (1984). Psychosocial problems among adults with epilepsy. *Epilepsia, 25*(2), 168–175.

Dravet, C. (1992). The epilepsies beginning at adolescence. In Anonymous, *Epilepsy in Adolescence: Symposium* (pp. 15–22). Ruisbroek: Le Monde Medical Magazine.

Dreifuss, F. E. (1995). Classification of epilepsies in childhood and adolescence. In A. P. Aldenkamp (Ed.), *Epilepsy in children and adolescents* (pp. 1–16). Boca Raton, FL: CRC Press.

Eamest, M. P., Thomas, G. E., Eden, R. A., & Hossack, K. F. (1992). The sudden unexpected death syndrome in epilepsy: demographic, clinical and post-mortem features. *Epilepsia, 33,* 310–316.

Elwes, R. D., Marshall, J., Beattie, A., & Newman, P. K. (1991). Epilepsy and employment. A community based survey in an area of high unemployment. *Journal of Neurology, Neurosurgery and Psychiatry, 54,* 200–203.

Engel, J., Van Ness, P. C., Rasmussen, T. B., & Ojemann, L. M. (1993). Outcome with respect to epileptic seizures. In J. Engel (Ed.), *Surgical treatment of the epilepsies* (pp. 609–621). New York: Raven Press.

Espir, M., & Floyd, M. (1986). Epilepsy and recruitment. In F. Edwards, M. Espir, & J. Oxley (Eds.), *Epilepsy and Employment: a Medical Symposium on Current Problems and Best Practices*. London: Royal Society of Medicine Services.

Fenwick, P. (1987). Epilepsy and psychiatric disorders. In A. Hopkins (Ed.), *Epilepsy*. London: Chapman & Hall.

Ferguson, S. M., & Rayport, M. (1965). The adjustment to living without epilepsy. *Journal of Nervous and Mental Disease, 140,* 26–37.

Ferrari, M. (1989). Epilepsy and its effects on the family. In B. P. Hermann & M. Seidenberg (Eds.), *Childhood epilepsies: neuropsychological, psychosocial and intervention aspects*. Chichester: Wiley.

Ferrari, M., Matthews, W. S., & Barabas, G. (1983). The family and child with epilepsy. *Family Process, 22,* 53–59.

Floyd, M. (1986). A review of published studies in epilepsy and employment. In F. Edwards, M. Espir, & J. Oxley (Eds.), *Epilepsy and Employment: A Medical Symposium on Current Problems and Best Practices*. London: Royal Society of Medicine Services.

Fraser, R. T., Clemmons, D., Trejo, W., & Temkin, N. R. (1983). Program evaluation in epilepsy rehabilitation. *Epilepsia, 24,* 734–746.

Goldstein, L. H. (1990). Behavioural and cognitive-behavioural treatments for epilepsy: A progress review. *British Journal of Clinical Psychology, 29,* 257–269.

Hauser, W. A., & Hesdorffer, D. C. (1990). *Epilepsy: Frequency, causes and consequences*. Maryland: Epilepsy Foundation of America.

Hauser, W. A., & Annegers, J. F. (1993). Epidemiology of Epilepsy. In J. Laidlaw, A. Richens, & D. Chadwick (Eds.), *A textbook of epilepsy*. Edinburgh: Churchill Livingstone.

Hauser, W. A., Annegers, J. F., & Elveback, L. R. (1980). Mortality in patients with epilepsy. *Epilepsia, 21,* 339–412.

Hauser, W. A., & Kurland, L. T. (1975). The epidemiology of epilepsy in Rochester, Minnesota, 1935 through 1967. *Epilepsia, 16,* 1–66.

Hawton, K., Fagg, J., & Marsack, P. (1980). Association between epilepsy and attempted suicide. *Journal of Neurology, Neurosurgery and Psychiatry, 43,* 168–170.

Helgeson, D. C., Mittan, R., Tan, S. Y., & Chayasirisobhon, S. (1990). Sepulveda epilepsy education: The efficiency of a psychoeducational treatment program in treating medical and psychosocial aspects of epilepsy. *Epilepsia, 31,* 75–82.

Hermann, B. P., & Whitmann, S. (1986). Psychopathology in epilepsy: a multi-etiological model. In S. Whitman & B. P. Hermann (Eds.), *Psychopathology in epilepsy: Social dimensions*. New York: Oxford University Press.

Hermann, B. P., Whitman, S., Wyler, A. R., Anton, M. T., & Vanderzwagg, R. (1990). Psychosocial predictors of psychopathology in epilepsy. *British Journal of Psychiatry, 156,* 98–105.

Hermann, B. P., & Whitman, S. (1991). Neurobiological, psychosocial and pharmacological factors underlying interictal psychopathology in epilepsy. In D. B. Smith, D. M. Treiman, & M. R. Trimble (Eds.), *Advances in neurology Vol. 55. Neurobehavioural problems in epilepsy*. New York: Raven Press.

Hills, M. D., & Baker, P. G. (1992). Relationships among epilepsy, social stigma, self-esteem and social support. *Journal of Epilepsy, 5,* 231–238.

Hoare, P., & Kerley, S. (1991). Psychosocial adjustment of children with chronic epilepsy and their families. *Developmental Medicine and Child Neurology, 33,* 201–215.

International League Against Epilepsy (1981). Proposal for revised clinical and electroencephalographic classification of epileptic seizures. *Epilepsia, 22,* 489–501.

International League Against Epilepsy (1989). *International classification of epilepsies and epileptic syndromes*. New York: Raven Press.

Illis, L. S. (1994). *Neurological rehabilitation*. Oxford: Blackwell Scientific.

Jacoby, A. (1992). Epilepsy and the quality of everyday life. Findings from a study of people with well-controlled epilepsy. *Social Science and Medicine, 43*(6) 657–666.

Jacoby, A. (1994). Felt versus enacted stigma: a concept revisited. *Social Science and Medicine, 38,* 269–274.

Jacoby, A. (1995). *Psychosocial functioning in people with epilepsy in remission and the outcomes of antiepileptic drug withdrawal*. Unpublished doctoral dissertation, University of Newcastle upon Tyne, UK.

Jacoby, A., Baker, G. A., Steen, N., Potts, P., & Chadwick, D. W. (1996). The clinical course of epilepsy and its psychosocial correlates: findings from a UK community study. *Epilepsia, 37,* 148–161.

Jones, I. (1988). A pharmacist's view of drug compliance in the elderly. In R. Tallis (Ed.), *Epilepsy and the elderly* (pp. 42–51). Oxford: Alden Press.

Lechtenberg, R. (1984). *Epilepsy and the family*. Cambridge, MA: Harvard University Press.

Leestma, J. E., Kalelker, M. B., Teas, S., Jay, F. W., & Hughes, J. R. (1984). Sudden unexpected death associated with seizures: analysis of 66 cases. *Epilepsia, 25,* 84–88.

Levin, R., Banks, S., Berg, B. (1988). Psychosocial dimensions of epilepsy: a review of the literature. *Epilepsia, 29,* 805–816.

Link, B. G., Mirotznik, J., & Cullen, F. T. (1991). The effectiveness of stigma coping orientations: can negative consequences of mental illness labeling be avoided? *Journal of Health and Social Behavior, 32,* 302–320.

Matthews, W. S., & Barabas, G. (1981). Suicide and epilepsy: a review of the literature. *Psychosomatics, 22,* 515–524.

McQuire, A., & Trimble, M.R. (1990). Quality of life in patients with epilepsy: the role of cognitive factors. In D. Chadwick (Ed.), *Quality of life and quality of care in epilepsy*. London: Royal Society of Medicine Services.

Mittan, R. J. (1986). Fear of Seizures. In S. Whitman & B. Hermann (Eds.), *Psychopathology in epilepsy: Social dimensions*. New York: Oxford University Press.

Minuchin, S. (1974). *Families and family therapy*. London: Tavistock Publications.

MRC Antiepileptic Drug Withdrawal Group (1991). A randomised study of antiepileptic drug withdrawal in patients in remission of epilepsy: Medical Research Council Antiepileptic Drug Withdrawal Study Group. *Lancet, 337,* 1175–1180.

Mulder, H. C., & Surrmeijer, T. P. (1977). Families with a child with epilepsy: a sociological contribution. *Journal of Biosocial Science, 9,* 13–24.

Nakken, K. O., & Lossius, R. (1993). Seizure-related injuries in multi-handicapped patients with therapy-resistent epilepsy. *Epilepsia, 34,* 836–840.

Napier, A., & Whitaker, C. (1978). *The family crucible.* New York: Harper and Row.

Porter, R. J. (1993). Classification of epileptic seizures and epileptic syndromes. In J. Laidlaw, A. Richens, & D. Chadwick (Eds.), *A textbook of epilepsy.* Edinburgh: Churchill Livingstone.

Pritchard, P. B., III, Holmstrom, V. L., & Giacinto, J. (1985). Self-abatement of complex partial seizures. *Annals of Neurology, 18,* 265–267.

Rausch, R., & Crandall, P. H. (1982). Psychological status related to surgical control of temporal lobe seizures. *Epilepsia, 23,* 191–202.

Ritchie, K. (1981). Interaction in families of epileptic children. *Journal of Child Pyschology and Psychiatry and Applied Disciplines, 22,* 65–71.

Robertson, M. M., Trimble, M. R., & Townsend, H. R. A. (1987). Phenomenology of depression in epilepsy. *Epilepsia, 28,* 364–372.

Rousseau, A., Hermann, B., & Witman, S. (1985). Effects of progressive relaxation on epilepsy: Analysis of a series of cases. *Psychological Reports, 57,* 1203–1212.

Rutter, M., Graham, P., & Yule, W. (1970). A neuropsychiatric study in childhood. In Anonymous, *Clinics in developmental medicine Nos. 35/36.* London: Spastics International and Heineman Medical.

Sander, J. W. A. S. (1993). Some aspects of prognosis in the epilepsies: a review. *Epilepsia, 34,* 1007–1016.

Scambler, G. (1984). Perceiving and coping with stigmatising illness. In R. Fitzpatrick, (Ed.), *The experience of illness.* London: Tavistock.

Scambler, G. (1989). *Epilepsy.* London: Tavistock.

Scambler, G. (1993). Coping with epilepsy. In J. Laidlaw, A. Richens, & D. W. Chadwick (Eds.), *A textbook of epilepsy.* Edinburgh: Churchill Livingstone.

Scambler, G., & Hopkins, A. (1980). Social class, epileptic activity and disadvantage at work. *Journal of Epidemiology and Community Health, 34,* 129–133.

Schneider, J. W., & Conrad, P. (1981). Medical and sociological typologies: the case of epilepsy. *Social Science and Medicine, 15A,* 211–219.

Schneider, J. W., & Conrad, P. (1983). *Having epilepsy: the experience and control of illness.* Philadelphia: Temple University Press.

Scott, D. F. (1993). *The history of epileptic therapy: an account of how medication was developed.* Carnforth, Lancs: Parthenon.

Sillanpaa, M. (1973). Medico-social prognosis of children with epilepsy. *Acta Paediatrica Scandinavica, Suppl. 237,* 3–104.

Shorvon, S. D. & Reynolds, E. H. (1979). Reduction in polypharmacy for epilepsy. *British Medical Journal, 2,* 1023–1025.

Shorvon, S. (1987). The treatment of epilepsy by drugs. In A. Hopkins (Ed.), *Epilepsy.* London: Chapman & Hall.

Smith, D. B., Mattson, R. H., Cramer, J. A., Collins, J. F., Novelly, R. A., Craft, B., & Veterans Administration Epilepsy Cooperative Study Group. (1987). Results of a nationwide Veterans Administration Cooperative study comparing the efficacy and toxicity of carbamazepine, phenobarbitol, phenytoin and primidone. *Epilepsia, 28,* S50–S58.

Smith, D. F., Baker, G. A., Dewey, M., Jacoby, A., & Chadwick, D. W. (1991). Seizure frequency, patient perceived seizure severity and the psychosocial consequences of intractable epilepsy. *Epilepsy Research, 9,* 231–241.

Spitz, M. C., Towbin, J. A., Shantz, D., & Adler, L. E. (1994). Risk factors for burns as a consequence of seizures in people with epilepsy. *Epilepsia, 35,* 764–767.

Stores, G., & Piran, N. (1978). Dependency of different types of school-children with epilepsy. *Psychological Medicine, 8,* 441–445.

Takaki, S., Kurokawa, T., & Aoyama, T. (1985). Monitoring drug non-compliance in epileptic patients: assessing phenobarbitol plasma levels. *Therapeutic Drug Monitoring, 7,* 87–91.

Tallis, R. (1993). Through a glass darkly: assessing quality of care of elderly people with epilepsy using quality of life measures. In Anonymous, *Quality of life and quality of care in epilepsy: Update 1993* (pp. 79–96), London: Royal Society of Medicine Services.

Tan, S. Y., & Bruni, J. (1986). Cognitive-behaviour therapy with adult patients with epilepsy: a controlled outcome study. *Epilepsia, 27*(3) 225–233.

Taylor, J. (1958). *Selected writings of John Hughlings Jackson. Volume 1 On epilepsy and epileptiform convulsions.* New York: Basic Books.

Temkin, O. (1971). *The falling sickness.* Baltimore: Johns Hopkins University Press.

Thompson, P. J. (1995). The impact of epilepsy on behaviour and emotional development. In A. P. Aldenkamp (Ed.), *Epilepsy in children and adolescents* (pp. 239–250). Boca Raton, FL: CRC Press.

Trimble, M., R., & Cull, C. (1988) Children of school age: the influence of antiepileptic drugs on behavior and intellect. *Epilepsia, 29,* 15–19.

Trostle, J. A. (1987). *Managing epilepsy: A community study of chronic illness in Rochester, Minnesota.* Unpublished doctoral dissertation, University of California.

Trostle, J. A. (1988). Medical compliance as an ideology. *Social Science and Medicine, 27,* 1299–1308.

Trostle, J. A., Hauser, W. A., & Susser, I. S. (1983). The logic of non-compliance. *Culture, Medicine and Psychiatry, 7,* 35–36.

Troupin, A. S., & Johannessen, S. I. (1990). Epilepsy in the elderly: a pharmacological perspective. In D. B. Smith (Ed.), *Epilepsy: current approaches to diagnosis and treatment.* New York: Raven Press.

Wannamaker, B. B., Morton, W. A., Gross, A. J., & Sanders, S. (1980). Improvements in antiepileptic drug levels following reduction of intervals between clinic visits. *Epilepsia, 21,* 155–162.

Wolf, P., Thorbecke, R., & Even, W. (1986). Social aspects of psychosis in patients with epilepsy. In S. Whitman & B. P. Hermann (Eds.), *Psychopathology in epilepsy.* New York: Oxford University Press.

World Health Organization (WHO) (1980). *International classification of impairments, disabilities and handicaps. The consequences of disease.* Geneva: World Health Organization.

Ziegler, R. G. (1982). Epilepsy: Individual illness, human predicament and family dilemma. *Family Relations, 31,* 435–444.

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 67,* 361–370.

Zielinski, J. J. (1974). Epilepsy and mortality rates and cause of death. *Epilepsia, 15,* 191–201.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.22
Sleep Disorders

RACHEL MANBER, RICHARD R. BOOTZIN, and DEREK LOEWY
University of Arizona, Tucson, AZ, USA

8.22.1 INTRODUCTION 506
- *8.22.1.1 Basic Sleep Physiology* 506

8.22.2 WHAT CAUSES SLEEP DISTURBANCE? 507
- *8.22.2.1 Biological Rhythms* 507
- *8.22.2.2 Sleep Fragmentation* 508
  - *8.22.2.2.1 Sleep apnea* 508
  - *8.22.2.2.2 Periodic limb movement disorder* 508
- *8.22.2.3 Substances* 509
- *8.22.2.4 Hypnotics* 509
  - *8.22.2.4.1 Alcohol* 510
  - *8.22.2.4.2 Stimulants* 510
- *8.22.2.5 Behavioral, Cognitive, and Environmental Factors* 510
  - *8.22.2.5.1 Behavioral and cognitive factors* 510
  - *8.22.2.5.2 Lifestyle and environmental factors* 511

8.22.3 LIFESPAN ISSUES 511
- *8.22.3.1 Infancy and Childhood* 512
- *8.22.3.2 Adolescence* 512
- *8.22.3.3 The Elderly* 513

8.22.4 ASSESSMENT OF SLEEP DISORDERS 513
- *8.22.4.1 Assessment Interview* 513
- *8.22.4.2 Assessment Techniques* 514
  - *8.22.4.2.1 Polysomnography* 514
  - *8.22.4.2.2 Actigraphy* 515
  - *8.22.4.2.3 Sleep diaries* 515
  - *8.22.4.2.4 Multiple sleep latency test* 515

8.22.5 TREATMENT OF SPECIFIC SLEEP DISORDERS 515
- *8.22.5.1 Circadian Rhythm Disorders* 515
- *8.22.5.2 Primary Insomnia* 516
  - *8.22.5.2.1 Stimulus-control instructions* 516
  - *8.22.5.2.2 Bright light* 517
  - *8.22.5.2.3 Relaxation* 517
  - *8.22.5.2.4 Paradoxical intention* 517
  - *8.22.5.2.5 Sleep restriction* 518
  - *8.22.5.2.6 Cognitive components* 518
- *8.22.5.3 Treatment of Childhood Sleep Difficulties* 518
- *8.22.5.4 Parasomnias* 519
- *8.22.5.5 Sleep Apnea* 520
- *8.22.5.6 Restless Leg Syndrome and Periodic Limb Movements* 520
- *8.22.5.7 Narcolepsy* 520

8.22.6 INTERACTIONS WITH OTHER DISORDERS 521
- *8.22.6.1 Alzheimer's Disease* 521
- *8.22.6.2 Chronic Pain Disorders* 521
- *8.22.6.3 Acquired Immune Deficiency Disorder* 521
- *8.22.6.4 Psychiatric Disorders* 522

8.22.6.4.1 Mood disorders 522
8.22.6.4.2 Anxiety disorders 522
8.22.7 PROFESSIONAL ISSUES 523
8.22.8 CONCLUSION 523
8.22.9 REFERENCES 523

8.22.1 INTRODUCTION

Sleep disorders have significant impact on daytime functioning, health, and mood. Epidemiological data indicate that, in comparison to noninsomniacs, insomniacs more frequently report trouble remembering things, difficulty accomplishing needed tasks, and feeling less physically well (National Sleep Foundation, 1991; Silva, Chase, Sartorius, & Roch, 1996). The data also indicate that insomniacs more frequently report having a more negative mood, being less able to enjoy contact with family and friends, and involvement in vehicle accidents in which fatigue was a factor. Daytime sleepiness is a particularly troublesome consequence of poor sleep. A recent Gallop poll (National Sleep Foundation, 1995a) found that about half (52%) of the adults surveyed reported driving while feeling drowsy and about a third (31%) admitted that they have, at some time, dozed off while at the wheel of a vehicle. It is estimated that about 46% of total and 36% of fatal motor vehicle accidents are caused by sleepiness (Leger, 1994).

Symptoms of sleep disturbance are often reported by patients receiving psychotherapy. Although these symptoms are frequently viewed by mental health professionals as part of the presenting psychopathology, they appear to be resistant to change. In a study of 854 adult outpatients receiving individual psychodynamic and eclectic psychotherapy at five mental health centers, sleep disturbance symptoms were among the most resistant to change across 52 sessions (Kopta, Howard, Lowry, & Beutler, 1994). For example, the proportion of patients who showed clinically significant improvement in their sleep symptoms after eight psychotherapy sessions was 23%, after 26 sessions was 27%, and after 52 sessions was only 30%. In contrast, interventions that specifically address sleep problems produce clinically significant improvements in sleep. Two meta-analyses of 59 and 66 treatment studies focused on insomnia found that short-term brief therapies varying in length from 4 to 12 weeks produced substantial improvement (Morin, Culbert, & Schwartz, 1994; Murtagh & Greenwood, 1995). There are at least two possible reasons for the discrepancy in treatment response in the two settings. First, sleep disturbance symptoms are seldom the direct focus of general psychotherapy and specific empirically validated treatments for sleep disturbance are seldom used outside a sleep disorder center. Second, many outcome studies that focus on insomnia include subjects with only primary insomnia, excluding individuals with other sleep disturbances, such as those associated with psychopathology. Nevertheless, the clear implication from these findings is that to change sleep complaints, therapists must focus on sleep and use empirically supported treatments. Because sleep complaints do not change as the result of general psychotherapy, it is important to be well informed on the different sleep disorders and the best ways to assess and treat them. This chapter provides an overview of the most common sleep disorders with an emphasis on psychological aspects of etiology, assessment, and treatment. It is intended to provide information to help the clinical psychologist assess and treat sleep complaints and determine when a referral to a sleep disorder center may be indicated.

8.22.1.1 Basic Sleep Physiology

A knowledge of the physiology of normal sleep is essential for understanding sleep disorders. Electroencephalogam (EEG), eye movements (EOG), and chin muscle reactivity (EMG) are the basic elements from which internal organization of sleep is inferred. Normal sleep of an adult is organized into sleep cycles lasting about 90 minutes. Each sleep cycle consists of different sleep stages that are determined primarily by the dominant frequency in the EEG signal. The EEG signal gains amplitude and slows down as one becomes drowsy and progresses through the non-REM stages of sleep from stage one to stage two, to the slow wave (delta) sleep stages (three and four). This process is accompanied by a subjective sense of increase in depth of sleep as measured both by the lengthening of the arousal threshold and by the increase in the percentage of individuals who report being asleep when awakened from a given sleep stage.

Rapid eye movement (REM) sleep is the stage of sleep that is most likely to be associated with dreaming. REM sleep is markedly different from stages one through four. Unlike these four non-REM stages, REM sleep is characterized by rapid conjugate eye movements, similar to those occurring during wakefulness. It is also associated with a loss of postural muscle tone (atonia), that protects us from acting out our dreams. This loss of muscle tone is reflected in the flattening of all activity in the mentalis muscles (under the chin) during the transition from non-REM to REM sleep. In addition, REM sleep is associated with significant changes in important regulatory mechanisms including irregular respiration, increased heart rate, and disrupted thermoregulation. These normal regulatory changes are of clinical significance. For example, breathing cessation during sleep is more likely to manifest during REM sleep, when respiration is less regular.

The first sleep cycle typically progresses from stage one to stage two to slow wave sleep and might end with a brief REM episode. Later in the night the amount of slow wave sleep within each cycle decreases, and the amount of REM increases. The last sleep cycle typically contains no slow wave sleep and a long REM episode. This relative distribution of sleep stages across the night explains why some disorders, such as those associated with partial and confusional arousal from slow wave sleep, are observed at the first part of the night, whereas other disorders, such as those associated with REM abnormalities are observed during the latter part of the night.

8.22.2 WHAT CAUSES SLEEP DISTURBANCE?

Sleep is affected by many aspects of our existence. The biological clocks regulating sleep and wakefulness, our physical and psychological well-being, our lifestyle, and substances we consume all interact with the systems regulating our sleep. These factors and their effects the duration of sleep, its internal organization, and its quality, are discussed below.

8.22.2.1 Biological Rhythms

Sleep must be understood within the context of biological rhythms. There is substantial evidence that sleep–wake behavior has an endogenous circadian (daily) rhythm; that is, a daily rhythm that persists in the absence of environmental cues. Under normal conditions, the sleep–wake circadian rhythm is synchronized with other circadian rhythms including core body temperature, cortisol, and melatonin. Core body temperature is used as a means of measuring the underlying sleep–wake rhythm. Within the normal range, higher core body temperatures are associated with alertness and lower temperatures are associated with sleep and sleepiness.

Circadian rhythm disorders occur when individuals attempt to sleep at times that are inconsistent with their underlying biological clocks. In delayed sleep phase syndrome (DSPS), the sleep–wake circadian rhythm is delayed compared to when the individual attempts to sleep. Individuals with this problem report difficulty falling asleep at a desired bedtime, but have normal sleep if they attempt to sleep a few hours later. Delayed sleep phase individuals commonly identify themselves as "night people" and report being most alert during the late evening and night hours. This problem is often seen in adolescents and young adults. In contrast, in advanced sleep phase syndrome, sleep occurs at an earlier than desired time and the individual awakens earlier than desired. This problem is often seen in older adults.

Bright light significantly affects the circadian sleep–wake rhythm. In the absence of light and other cues regarding the time of day, the endogenous biological rhythm regulating sleep and wakefulness lasts about 25, rather than 24, hours (Moore-Ede, Martin, Sulman, & Fuller, 1982). Under normal circumstances, environmental influences, such as sunlight, sleep, meals, and social activity, entrain the endogenous rhythm to a 24-hour cycle. The combination of bright light during wake and total darkness during sleep is a particularly strong entraining influence and can facilitate adjustment to changes in sleep schedules (Czeisler et al., 1989).

Two common environmentally caused circadian rhythm problems are shift work and time-zone changes. Because of family and social demands, night shift workers usually attempt to live their days off work on a different schedule than their work days. A disrupted sleep–wake schedule often results in disturbed and shortened sleep, sleepiness on the job, reduced performance levels, and psychological distress due to disruptions in family and social life (Penn & Bootzin, 1990). Shift work coping problems are affected by individual factors as well as factors associated with the work system. For example, adjustment to shift work is more difficult after age 50, for individuals with a history of gastrointestinal complaints, and for workers on rotating rather than on permanent shifts schedules (Monk, 1994).

Longer biological rhythms also affect sleep. Seasonal and monthly biological clocks are called infradian rhythms. Seasonal sleep–wake rhythms can be seen in hibernating animals.

Humans who live in the extreme northern and southern latitudes experience disrupted sleep during seasonal periods of 24-hour darkness and 24-hour sunlight. Individuals with seasonal affective disorder (SAD) also report changes in sleep need and sleep quality with the changing seasons of the year. The biological clock that regulates menstruation is another infradian rhythm that interacts with the sleep–wake system. Both subjective and objective data indicate that sleep is more disturbed and more fragmented premenstrually compared to the more hormonally quiescent preovulatory phase (Manber & Bootzin, 1997; Parry, Mendelson, Duncan, Sack, & Wehr, 1989). Initial evidence suggests an interaction between the daily core body temperature and some of the longer biological rhythms. For example, the amplitude of the temperature rhythm is lower during the premenstrual phase (Lee, 1988), and its phase is advanced during the months of shorter photoperiods in individuals with SAD (Terman, 1994).

8.22.2.2 Sleep Fragmentation

Sleep is fragmented when it is disrupted by intermittent arousals. Even when repeated arousals are brief and do not significantly reduce total sleep time, the subjective perception upon awakening in the morning is that sleep was not refreshing. In fact, sleep fragmentation is a better predictor of daytime sleepiness than the amount of total sleep (Carskadon, Brown, & Dement, 1982; Martin, Engleman, Deary, & Douglas, 1996). Sleep experts distinguish between arousals and awakenings. The technical distinction between the two has to do with the length of time that the EEG signal is fast enough to be characterized as "nonsleep." The sleeper is usually not aware of arousals or short awakenings. Two common causes of sleep fragmentation are breathing cessation and periodic leg twitches. Both are symptoms of sleep disorders that are discussed next.

8.22.2.2.1 Sleep apnea

Sleep apnea is a sleep disorder in which an individual ceases breathing during sleep for brief periods (average duration 10–30 seconds) throughout the night. These events may be due to an obstruction of the upper airway, as is the case in obstructive sleep apnea, or to abnormal functioning of central respiratory mechanisms, as is the case with central sleep apnea. Apnea episodes are frequently associated with arterial oxygen desaturation (hypoxemia).

Sleep apnea, in its various forms, can cause significant sleep fragmentation. Although not all apnic events will result in full awakenings, almost all will be associated with some form of arousal. A patient may experience hundreds of these events each night and consequently might experience severe daytime sleepiness (Roehrs, Zorick, Witting, Conway, & Roth, 1989). Loud snoring is strongly associated with apnea and serves as a diagnostic symptom. Primary snoring, in itself, may also cause fragmentary arousals during sleep. A more recently documented sleep-related breathing disorder is upper airway resistance syndrome (Guilleminault, 1991). This disorder is associated with airway obstruction and breathing cessions that are shorter than 10 seconds but produce microarousals and result in nonrestorative sleep and daytime somnolence.

Sleep apnea is more common in men than in women and increases in men with age and in women after menopause. Obesity and the anatomy of the upper airways also contribute to sleep disordered breathing (Isono & Remmers, 1994). Ingestion of central nervous system depressants, such as narcotics, benzodiazepines, and alcohol, can exacerbate sleep apnea (Robinson & Zwillich, 1994).

8.22.2.2.2 Periodic limb movement disorder

Periodic limb movement disorder is another common physical cause of sleep fragmentation. In this disorder the limbs, most often the legs, will repeatedly "twitch" or "jerk" while the person is sleeping. Typically, the sufferer is unaware of this activity, and it is usually reported by the bed partner rather than the patient. Periodic limb movements (PLMs) usually cause microarousals, but may also lead to full awakenings. PLMs are commonly associated with complaints of excessive daytime sleepiness. A related phenomenon is restless legs syndrome (RLS) in which the patient may report odd sensations or discomfort in the legs during the day or, most notably, when lying down at bedtime. This discomfort is distinct from a leg cramp and is best described as an uneasy deep sensation. The patient may report having to move or rub the legs or get out of bed and walk about to alleviate the symptoms. Individuals experiencing symptoms of restless legs while awake frequently exhibit PLMs during sleep (Montplaisir, Godbout, Pelletier, & Warnes, 1994). Restless legs may be a symptom of iron deficiency and may also be seen in uremia, pregnancy, diabetes mellitus, rheumatoid arthritis, and polyneuropathy. However, both periodic limb movement disorder and RLS are sleep disorders that are often independent of other medical conditions.

8.22.2.3 Substances

Both prescription and nonprescription substances can cause sleep disturbances. The chronic use of drugs, alcohol, and caffeine to induce or to suppress sleepiness may cause adverse effects. Further, many substances taken for other health and lifestyle reasons can affect sleep.

8.22.2.4 Hypnotics

The prescription of sedative/hypnotics is the most frequently used treatment for insomnia. About 4.3% of the population use medically prescribed psychoactive medication to promote sleep (i.e., hypnotics, anxiolitics, and antidepressants; Mellinger, Balter, & Uhlenhuth, 1985). Epidemiological studies in the USA report a prevalence of hypnotics of approximately 2.5% (Balter & Uhlenhuth, 1992; Mellinger et al., 1985). The prevalence rates for hypnotic medication are higher in many other countries. For example, the figures of habitual use range from 3.8 to 6.2% in France (Quera-Salva, Orluc, Goldenberg & Guilleminault, 1991) and 6.5% in Israel (Matalon, Yinnon, & Hurwitz, 1990).

Of the sedative/hypnotics, sedative benzodiazepines are the most frequently prescribed (Morin & Kwentus, 1988). These drugs have a definite effect on sleep architecture including reduction in sleep latency, decrease in the number of nocturnal awakenings, increase in total sleep time, and decrease in amount of REM and delta sleep. All central nervous system depressants, including benzodiazepines, depress respiration and exacerbate sleep apnea. Hypnotics have a number of other deleterious side effects including drug hangover and, for some drugs, increased sleep disruptions in the latter part of the night. The pharmacological effects of hypnotics depend upon dose, absorption rate, and serum half-life (Greenblatt, 1992; Nicholson, 1994). Hypnotics with long half-lives (such as flurazepam) are likely to produce drug hangover and daytime sedation. Hypnotics with short half-lives (such as triazolam) may produce rebound insomnia the very same night the medication is taken. The same side effects and disrupted sleep architecture that Rare associated with hypnotics also apply to benzodiazepine tranquilizers such as diazepam (Valium).

The elderly are particularly vulnerable to deleterious side effects of hypnotics because they are more likely to have disorders that can be aggravated by hypnotics, such as respiratory, hepatic, renal, or cardiac disorders. In addition, the period of potential toxicity of hypnotics increases with the age-dependent decreases in protein-binding ability, circulation time, and kidney and liver metabolism of the drugs (Albert, 1981).

The use of sedative/hypnotics to induce sleep is effective for short-term use, with effects that last from two to four weeks. With continued use, tolerance develops rapidly and larger doses are required to achieve a soporific effect. As dosage is increased to offset tolerance, daytime carryover effects also increase and may include symptoms such as excessive sleepiness, poor motor coordination, visual–motor problems, and late afternoon restlessness (American Sleep Disorder Association [ASDA], 1990). Cessation of sedative/hypnotics results in withdrawal symptoms that include severe insomnia. Withdrawal symptoms promote psychological dependency as the individual becomes convinced that the hypnotics are the only thing preventing sleeplessness night after night. Withdrawal effects are usually stronger with the short half-life hypnotics. With the long half-life hypnotics, there is still medication in the blood stream for days after the insomniac stops taking the medication, producing the equivalent of a tapered withdrawal.

During the past few years, zolpidem, a nonbenzodiazepine hypnotic, has been available in the USA. Zolpidem is an imidazopyridine which is related to benzodiazepines but so far has not been shown to have rebound or daytime residual effects (Nicholson, 1994). Zolpidem is a very short half-life medication and is primarily useful for sleep onset problems. However, zolpidem, like all hypnotics, is a central nervous system depressant and may exacerbate other health problems.

Antidepressants with sedating properties are often used in the management of insomnia. However, not all antidepressants have beneficial effects on sleep. Tricyclic antidepressants suppress REM sleep and can exacerbate periodic limb movements. Monoamine oxidase (MAO) inhibitors have an even greater suppression of REM (Minot, Luthringer, & Macher, 1993). When REM suppressing agents are discontinued, there is a rebound in REM sleep that is frequently associated with nightmares and intense dreams. Many, but not all, of the new generation antidepressants produce increased physiological arousal during sleep resulting in reduced delta activity and decreased sleep efficiency (Armitage, Trivedi, Rush, & Hoffmann, 1995). Increased physiological arousal and REM suppression are particularly evident in some of the selective serotonin reuptake inhibitors. On the other hand, serotonin receptor modulators, such as Trazodone and Nefazodone, increase sleep continuity and do not reduce slow wave sleep (Reynold, Buysse, Nofzingor, & Kupfer, 1995).

Nonprescription substances that facilitate sleep include camomile tea, valerian root, tryptophan, and melatonin. None of these substances has been shown to be helpful for persistent insomnia. It should also be emphasized that because these substances are not regulated by the Food and Drug Administration, the consumer cannot be assured of purity. There is substantial current interest in the use of melatonin. While there are promising research results (e.g., Haimov et al., 1995; Zhdanova, Wurtman, Lynch, Morabito, & Matheson, 1995), important questions about dose size and long-term effects have yet to be answered.

8.22.2.4.1 Alcohol

The problems associated with hypnotics and tranquilizers are also prevalent with alcohol. Like other central nervous system depressants, alcohol is REM-sleep depriving. Continued use results in tolerance to the alcohol as a sleep-inducing agent and sleep fragmentation becomes more prominent. Patients often increase the amount of alcohol or add other sedatives in order to sleep. Some patients will report that they have no sleep disturbance as long as they continue to take substantial amounts of alcohol nightly. Withdrawal from heavy drinking produces a REM rebound effect which is accompanied by restless sleep and nightmares. Sleep fragmentation often persists long after alcohol use is discontinued (Gillin, 1994). Alcohol also exacerbates sleep apnea. An important additional danger is that alcohol potentiates the effects of hypnotics and other depressants. Thus, the combination of alcohol and hypnotics may intensify and prolong deleterious side effects.

8.22.2.4.2 Stimulants

The effects of stimulants on sleep include an increase in sleep latency, a decrease in total sleep time, and an increase in spontaneous awakenings (ASDA, 1990; Brown et al., 1995). Substances that have stimulating effects on the central nervous system include caffeine, some prescription mediations, nicotine, cocaine, and amphetamines.

Caffeine has a plasma half-life of approximately six hours and as a result its effects can be felt long after it has been ingested. Reducing or eliminating the intake of caffeine, particularly in the afternoon and evening, can reduce its negative effects on sleep. Some prescription medications, such as Ritalin, are prescribed primarily for increasing alertness in disorders of excessive daytime sleepiness. Other prescription medications, including analgesics, bronchodilators, decongestants, and appetite suppressants, often contain stimulants and disturb sleep architecture as a side-effect of long-term use.

Nicotine is another central nervous system stimulant that produces lighter and more fragmented sleep. Cocaine and amphetamines have euphoric and sleep suppressing effects that reinforce their continued use. As tolerance to the alerting and euphoric effects of cocaine and amphetamines increases, higher doses are necessary and greater disturbance in sleep are experienced. Cessation of these drugs, as well as nicotine, may be associated with withdrawal symptoms such as sleepiness, irritability, lassitude, and severe depression (ASDA, 1990; Wetter, Fiore, Baker, & Young, 1995).

8.22.2.5 Behavioral, Cognitive, and Environmental Factors

Insomnia may begin with a precipitating event. Stress, disruptive sleep environment, or an abrupt change in sleep schedule are three examples of events that could precipitate insomnia. The interaction between a precipitating event and predisposing and/or maintaining factors determines the extent and the course of insomnia. Predisposing factors explain individual differences in the response to a precipitating event. For example, two individuals with different underlying circadian rhythms might respond differently to a change from a night shift to a day shift. This change in schedule might precipitate insomnia in a "night person" yet be a welcome relief for a "morning person." In other words, physiological differences in the sleep–wake rhythm are predisposing factors that might interact with a precipitating event to produce sleep disruptions. Factors that maintain sleep problems could be behavioral, cognitive, or environmental. For example, the strategies one uses to solve or cope with sleep difficulties can sometimes exacerbate the sleep problem and maintain it long after a precipitating event is no longer relevant.

8.22.2.5.1 Behavioral and cognitive factors

Behaviors that are inconsistent with the maintenance of good quality sleep and full daytime alertness are referred to collectively as inadequate sleep hygiene. Inadequate sleep hygiene commonly develops when individuals adopt ineffective strategies for coping with their sleep problem. One common strategy for coping with lost sleep is to sleep in until the late morning hours following a rough night. This strategy could lead to irregular sleep–wake schedules that might prolong the sleep difficulties by weakening the underlying sleep–wake rhythm. Another common behavioral strategy to cope with poor sleep at night is to take an extended nap during the day. However, extended afternoon naps are often followed by difficulties initiating sleep at night. Extremely sedentary lifestyles also fall into the category of inadequate sleep hygiene. These behaviors are not likely to provide sufficient entrainment cues to the sleep–wake regulatory system and might consequently be associated with disturbed sleep.

Conditioning factors are often involved in the development of sleep problems. For example, operant conditioning might be involved in the development of sleep disturbance in the very young with parental soothing behaviors reinforcing night time calling behaviors by a toddler. Classical conditioning processes could also contribute to disturbed sleep. Many insomniacs engage in activities at bedtime that are incompatible with falling asleep, thus rendering the bed a cue for arousal rather than for sleep (Bootzin & Nicassio, 1978). For some, bedtime is the first available quiet time and it becomes the time to rehash the day's events and make plans for the next day. These activities promote arousal, rather than sleep. In addition, with time, for many insomniacs the bedroom becomes associated with anxiety and frustration about their inability to fall asleep, and worry about how lack of sleep might affect their performance the following day. These individuals often have less difficulty falling asleep while away from home.

A common cognitive characteristic of insomniacs is the tendency to overestimate latency to sleep onset and to underestimate total sleep time (Borkovec, Lane, & Vanoot, 1981; Carskadon et al., 1976). The causes of this misperception are not clear. It could be related to excessive worry in bed and increased cognitive arousal (Monroe, 1967) or to hypervigilance to external stimuli during sleep (Anch, Saskin, & Moldofsky, 1989). Alternately, poor time estimation could be related to a compromised ability to distinguish between sleep and wakefulness states (Engle-Friedman, Baker, & Bootzin, 1985). This appears to be the case in approximately 5% of insomniacs. These insomniacs are diagnosed with "sleep-state misperception" because objective sleep measures do not verify their subjective complaints.

8.22.2.5.2 Lifestyle and environmental factors

Stress is one of the most common precipitating cause of insomnia. It is usually associated with physiological, cognitive, and emotional arousal that, in themselves, contribute to sleep disruptions. Research has demonstrated that compared to normal controls, insomniacs have greater whole body metabolic rates, body temperature, and heart rate both during the day and at night (Bonnet & Aranda, 1995). It is not clear if this heightened arousal is the cause or consequence of poor sleep. In the last decade several investigators have raised the hypothesis that insomnia results from increased central nervous system activation that interferes with the sleep system (Stephanski, Zorick, Roehrs, Young, & Roth, 1988; Bonnet & Arand, 1996).

Environmental factors such as noise level, light intensity, room temperature, and safety are important to the quality and continuity of sleep. When these environmental factors are not optimal they often contribute to sleep disruptions though they are seldom the primary cause of insomnia (ASDA, 1990). Sleeping in an environment that is perceived to be unsafe leads to hypervigilance that results in lighter sleep. Noise, particularly unpredictable noise (Sanchez & Bootzin, 1985), decreases both the amount of slow wave sleep and the continuity of sleep. It also increases sleep fragmentation, body movement, and number of stage shifts. In contrast, continuous white noise is often useful for masking noisy sleep environments.

An uncomfortable temperature appears to have greater negative effect on measures of sleep continuity than on the ability to fall asleep at bedtime. Sleep is more likely to be disturbed if the temperature changes significantly during REM sleep than if the temperature change occurs during non-REM sleep (Satinoff, 1988).

Sleep problems also need to be considered within the family context. Choices about sleep schedules, sleep environments, and other sleep behaviors might depend on the family environment. For example, some insomniacs might be going to bed at times that are not congruent with their own biological clocks because it is important within their family context that they go to bed at the same time their spouse does. Family issues might also influence the sleep environment itself. Concerns about waking up spouses might prolong the time insomniacs might stay in bed worrying about their inability to fall asleep.

8.22.3 LIFESPAN ISSUES

The two extreme ends of the lifespan are associated with significant changes in the organization of sleep. These changes are accompanied by corresponding changes in the types and prevalence of sleep disorders.

8.22.3.1 Infancy and Childhood

During the first few months of life, the diurnal organization of sleep and wakefulness becomes synchronized with the light–dark cycle and with social cues. By six months of age, most infants have one long period of sleep at night and two naps during the day. During toddlerhood, the number of naps is reduced to a single nap a day and by the age of four or five most children do not nap during the day. The average total sleep time decreases from about 14 hours at the end of the first year to about 11 or 12 hours by age five.

The internal organization of each sleep episode undergoes significant changes during the first year. During the first few months sleep becomes organized into cycles lasting 50–60 minutes and the time spent in REM within each sleep cycle is reduced and becomes more prevalent in the latter part of the night. Within 10 minutes after sleep onset, the young child descends into slow wave sleep and seldom enters REM sleep before the second sleep cycle. Slow wave sleep is quantitatively and qualitatively different in children. It is characterized by greater amplitude and significantly higher arousal thresholds than in adults. The amount of time spent in slow wave sleep, the amplitude of the brain wave, and the arousal thresholds all decrease from about age 10 onwards (Feinberg, 1989).

The most common parental concern about their infant's sleep is the child's difficulty settling into sleep either at bedtime or in the middle of the night. During early childhood, insufficient sleep manifests as hyperactivity and irritability rather than sleepiness. From toddlerhood onward, sleep problems decrease from about 25% at age two to about 10% during the preschool years (Richman, 1987). Sleep disturbances in children have been found to be primarily transient in nature when prospective methodology was used (Strauch & Meier, 1988), but retrospective studies tend to find greater stability of childhood sleep disorders (Hauri & Olmstead, 1980; Richman, Stevenson, & Graham, 1982)

Disorders of partial arousal from slow wave sleep, such as sleep walking, night terrors, sleep talking, and nocturnal enuresis, are most prevalent during childhood when slow wave sleep is most abundant. These disorders tend to resolve with age as the amount of slow wave sleep decreases. Nocturnal enuresis is of particular relevance to this age group. It is estimated that 30% of children up to age four may wet their bed at night and that the prevalence declines to 3% by age 12. Although maturation decreases the prevalence of nocturnal enuresis, this decrease is slow. For example, when left untreated, nocturnal enuresis persists in 85% enuretic children (age eight) and in 84% of enuretic adolescents one year postinitial assessment (Forsythe & Redmond, 1974).

8.22.3.2 Adolescence

The internal organization of sleep of an adolescence resembles that of an adult with sleep cycles lasting about 90 minutes. Both the physiological and psychological processes that occur during this period of life have significant effects on sleep. The accumulation of research on the sleep of adolescents suggests that, contrary to societal beliefs, this is a period of high sleep need. Yet survey data indicate a decline in total sleep time from 10 hours in prepubertal children to seven in late puberty. School demands, strong pull for spending time with peers, extracurricular activity, and work for wages, all contribute to delayed bedtimes and to a reduction in total sleep time on school nights (Carskadon, Mancuso, & Rosokiad, 1989).

Not surprisingly, increased sleep deprivation during adolescence is accompanied by increased levels of daytime sleepiness. More than half of the adolescents surveyed both in France and in the United States complained of daytime sleepiness (Carskadon & Dement, 1987; Manber et al., 1995a, 1995b; Strauch & Meier, 1988). Objective measures of sleepiness (see Section 8.22.4.2.4) confirm this increase in daytime sleepiness during the second decade of life. Levels of daytime sleepiness in midadolescence, although high, are below the levels seen in patients with disorders of excessive daytime sleepiness. Nevertheless, the impact of chronic mild sleep deprivation on many aspects of adolescents" life is not trivial (Carskadon & Dement, 1981; Manber & Bootzin, 1991) and is amplified by alcohol and drug abuse and by emotional stress. Irritability, difficulty in sustaining attention, and vulnerability to minor illnesses and injuries could increase with increased sleepiness.

The second most common sleep complaint during adolescence is sleep onset insomnia. It is reported by about 12–14% of adolescents (Price, Coates, Thoresen, & Grinstead, 1978; White, Hahn, & Mitler, 1980). It has been suggested that sleep onset insomnia, particularly during adolescence, might actually be a mild form of delayed sleep phase insomnia (Henschel & Lack, 1987). To cope with insufficient sleep on school nights, adolescents typically try to catch up and sleep in during the weekend. This leads to a development of irregular sleep patterns that could weaken the underlying biological rhythm regulating sleep and produce a delayed sleep phase problem (Carskadon, Vieira, & Acebo, 1993).

Irregular sleep patterns appear to contribute not only to sleep difficulties but also to elevated levels of sleepiness (Billiard, Alperovitch, Perot, & Jammes, 1987; Strauch & Meier, 1988). Two recent studies of sleepy and irregular sleepers suggest that although sleep deprivation is a primary cause of daytime sleepiness, schedule irregularity further exacerbates the problem. In one study, sleepy and irregular sleepers regularized their sleep schedules but suffered sleep loss in the process. These subjects experienced an increase in daytime sleepiness and a concomitant deterioration in daily mood and concentration ratings (Manber & Bootzin, 1991). A second study examined the efficacy of regularizing sleep schedules while preventing sleep deprivation. This study found that a combination of adequate sleep durations with regular nocturnal sleep schedules produced greater improvement in daytime sleepiness than adequate sleep duration alone (Manber, Bootzin, Acebo, & Carskadon, 1996). Thus, to improve sleep and daytime functioning, adolescents should be encouraged to obtain sufficient sleep and to maintain regular sleep schedules.

8.22.3.3 The Elderly

The most frequent sleep complaints of older adults are increased frequency and duration of nocturnal awakenings and light sleep. These complaints parallel developmental changes in psychophysiological sleep parameters that occur with advancing age (Bliwise, 1994). One of the most consistent findings associated with aging, documented through polysomnography, is an increase in the frequency of awakenings, particularly during the latter half of the night (e.g., Karacan & Hursch, 1974; Reynolds 1985). Older adults also have more difficulty falling back to sleep than younger individuals (Webb & Campbell, 1980). As a consequence of increased time awake and difficulty falling back to sleep, older adults get less total sleep at night than younger adults. However, because daytime naps increase in frequency with age, the total sleep per 24 hours of the elderly is about equivalent to that of younger individuals.

Aging appears to affect both non-REM slow wave sleep and REM sleep. There is a marked reduction in the amount of stages three and four of sleep with little or no stage four sleep in about one-quarter of the population over the age of 60. The reduction in stages three and four results primarily from a reduction in amplitude of the EEG waves, so that they do not meet the amplitude criterion for scoring as slow wave sleep (e.g., Feinberg, Koresko, & Heller, 1967). Thus, slow EEG-frequency sleep still occurs in the elderly, but there is a marked loss in EEG amplitude.

There are also small decreases in REM sleep with aging. Unlike younger individuals, older adults have REM periods that remain uniform or decrease in duration during the night (Reynolds et al., 1985). Increased REM early in the night's sleep is associated with depression and with the normal advance in the phase of the circadian sleep rhythm with age. In addition, shortened latency to REM sleep and lower amounts of REM have been found to be associated with intellectual decline (Feinberg et al., 1967). It should be emphasized that while there are substantial changes in sleep associated with aging, these changes do not always result in a subjective complaint of sleep disturbance. Changes in slow wave sleep and REM sleep, for example, may not influence judgments about the quality of sleep, while frequent awakenings do (Bonnet & Johnson, 1978).

There are gender and age differences among the elderly, particularly related to sustaining sleep and the amount of slow wave sleep obtained. Although there is an age-related decrease in the slow wave sleep of the elderly, aging women appear to have more slow wave sleep than older men (Hoch et al., 1987). For older women, compared with older men, a higher and more stable relationship exists between self-reported soundness of sleep and the polysomnographic variable of slow wave sleep.

8.22.4 ASSESSMENT OF SLEEP DISORDERS

8.22.4.1 Assessment Interview

The clinical assessment of sleep disorders begins with a detailed diagnostic intake interview that determines the course of treatment as well as the potential need for further diagnostic procedures such as an all night sleep study (polysomnography). Because a patient's global statement is rarely accurate, it is important to quantify the severity and extent of a presenting sleep problem by systematically exploring the patient's sleep pattern including bedtime, time to sleep onset, time awake after sleep onset, wake–up time, and the degree of variability in these sleep parameters. A careful examination of the multiple causes of sleep disturbance is the central focus of the intake interview. Evaluating the contribution of causal factors such as sleep hygiene, medical and psychiatric histories, medications, and other substances, is relatively straightforward. Evaluation of circadian tendencies and of symptoms of sleep apnea or periodic limb movement disorder are less obvious. Circadian tendencies can be assessed by inquiring about the timing of optimal functioning as well as times of day when sleepiness is most pronounced.

The clinical assessment of potential sleep apnea is done by inquiring about loud snoring, and witnessed gasping for air or shortness of breath during the night. Sweating, unrefreshing sleep, morning headaches, morning dry mouth, and swelling in the extremities are also important clinical indicators of potential sleep apnea. Physiological features that are potential sources of breathing obstructions also need to be assessed. These include obesity and the structure of the chin, the uvula, and the tongue. A positive clinical indication for sleep disordered breathing needs to be followed up by a polysomnographic assessment of sleep that includes apnea monitoring.

The clinical assessment of RLS and PLMs relies on the symptom of an uneasy sensation in the leg while trying to fall asleep that typically triggers the patient to get out of bed and walk. Polysomnography is indicated when clinical symptoms are present. Polysomnography allows quantification of severity of RLS and is important in the diagnosis of periodic limb movement disorder.

The impact of poor sleep on daytime functioning, particularly on daytime sleepiness, is an important determinant of diagnosis and treatment course, and therefore needs to be an integral part of the assessment of sleep disorders. The quantification of daytime subjective sleepiness is accomplished by evaluating the tendency to fall asleep in different settings. Unwanted and inappropriately timed daytime sleep episodes often flag the clinician that disorders of excessive daytime sleepiness, such as sleep apnea, need to be ruled out. When reported levels of daytime sleepiness are severe and include frequent unintentional sleep episodes, there needs to be an assessment for narcolepsy. Narcolepsy is a neurological disorder characterized by excessive daytime sleepiness that is typically associated with cataplexy (a sudden spell of weakness, or muscle atonia), and other REM sleep phenomena such as sleep paralysis and hypnagogic hallucination (hallucination occurring at the interface between wakefulness and sleep). Narcolepsy can be assessed both through clinical symptoms and through a special polyomnographic procedure (see also Section 8.22.5.7). In contrast, an inability to nap, even when feeling very tired, is common in insomniacs and should not be interpreted as optimal alertness but rather as a reflection of the insomniac's difficulty in falling asleep.

The assessment of sleep disturbance during infancy and childhood requires attention to a unique set of relational and cultural factors. These include the temperament of the infant, bedtime interaction between parent and child, separation anxiety, difficulties in self soothing, and parental conflict regarding child rearing practices. The potential contribution of allergy to cow's milk to disturbed sleep in the infant needs to be assessed (Kahn et al., 1987), but other aspects of feeding, including amount and method, do not appear to be related to sleep difficulties for most infants.

The interview can be more efficient and thorough when it is preceded by a short battery of questionnaires addressing sleep habits, common sleep symptoms, daytime sleepiness, depression, and anxiety. A few existing sleep questionnaires can further illuminate the patients" attitudes and beliefs about their sleep problem (Morin, 1993), as well as level and type of anticipatory anxiety at bedtime (Bootzin, Shoham, & Kuo, 1994). Because one cannot observe oneself during sleep, input from a significant person in the patient's life is often valuable. This applies particularly to breathing behaviors during sleep and to odd sleep behaviors, such as sleep walking. In addition, information from a significant other is important when the patient does not remember, is unaware, or minimizes the presenting problem.

8.22.4.2 Assessment Techniques

8.22.4.2.1 Polysomnography

Polysomnography (PSG) is the most commonly used method for the objective assessment of sleep. In a clinical setting, it includes an overnight monitoring of sleep in the sleep laboratory that includes EEG, EOG, and chin EMG, as well as airflow, respiratory effort, and oximetry to screen for apnea, and leg or arm EMG to test for PLMs. An overnight sleep study can assess the severity of disorders such as sleep apnea, RLS, and PLMs and is part of the diagnostic procedure for narcolepsy. For these disorders, a single night of PSG is usually adequate because the defining characteristics of these disorders are stable (Bootzin et al., 1995). However, sleep parameters that are relevant to the assessment of insomnia are often altered on the first night of sleep in the unfamiliar laboratory environment. This "first night effect" and the great night-to-night variability of insomnia symptoms make it difficult to interpret the results from a single night of polysomnography. Consequently, it is generally agreed that polysomnography should not be used in the assessment of insomnia unless a careful history or physical examination suggest the presence of physiological disturbances such as sleep-related breathing disorder or PLM (Reite, Buysse, Reynolds, & Mendelson, 1995). Use of polysomnography is also recommended when the treatment of insomnia is unsuccessful (ASDA, 1995a). Polysomnography is also not an optimal tool in the clinical assessment of disorders that involve events that do not occur nightly, such as sleep walking and night terrors, because they are not likely to be "captured" during the particular night of the study. Recent advancements in sleep medicine include ambulatory or home-monitoring PSGs that are cheaper and reduce the impact of the first night effect. A disadvantage of ambulatory or home-monitoring PSGs is that they can monitor fewer sleep parameters than laboratory studies.

8.22.4.2.2 Actigraphy

A wrist actigraph is a movement-detecting device that is worn on the wrist. It provides a portable, longer-term means of assessing variables such as daily rest/activity, sleep–wake cycle, sleep latency, nocturnal arousals and awakenings, and total sleep time. The reliability and validity of actigraphs has been demonstrated in the assessment of sleep variables (Cole & Kripke, 1988) and in the assessment of sleep disorders (Sadeh, Alster, Urbach, & Lavie, 1989). Some actigraphs are equipped with light sensors to quantify the amount of light exposure a patient receives. Others are equipped with a behavioral response capacity to provide a continuous measure of reaction time performance for the assessment of daytime sleepiness. The primary use of actigraphy at present is in research settings. Clinical use of actigraphy is recommended primarily as an adjunct to diagnosis and to the assessment treatment compliance (ASDA, 1995b).

8.22.4.2.3 Sleep diaries

Daily sleep diaries are effective and inexpensive means of obtaining prospective information on sleep–wake behaviors. They are frequently used in the assessment of sleep disorders as well as in the assessment of treatment efficacy and compliance. When completed each morning, sleep diaries correlate highly with PSG variables (Carskadon et al., 1976). With time, patients become better at estimating their sleep behaviors, further increasing the validity of sleep diaries (Franklin, 1981).

8.22.4.2.4 Multiple sleep latency test

The multiple sleep latency test (MSLT) is the most widely used objective assessment procedure of daytime sleepiness. It consists of a series of four or five nap opportunities given every two hours in the sleep laboratory. The patient lays in a darkened, sound-attenuated room for up to 20 minutes and is instructed to fall asleep. The average latency to sleep onset over all nap opportunities provides an index of sleep propensity. A smaller number reflects greater level of sleepiness. Average latencies of less than five minutes indicate excessive daytime sleepiness. The MSLT is less accurate for assessment of sleepiness in insomniacs because the inability to fall asleep is part of the presenting pathology and is likely to carry over to the nap opportunities.

A related but different procedure is the wake maintenance test (WMT) which tests the ability to remain awake during soporific circumstances (Mitler, 1992). It is delivered in a manner similar to MSLT, with one important exception. The patient is seated and instructed to try to maintain wakefulness. It is used for assessment of degree of impairment and treatment efficacy in disorders of excessive day time somnolence.

8.22.5 TREATMENT OF SPECIFIC SLEEP DISORDERS

The treatment of sleep disorders is enhanced when a multidisciplinary approach is adopted. Psychological, behavioral, and medical aspects of treatment all contribute to treatment efficacy. Psychology has a central role in the treatment of primary insomnia, insomnia complicated by psychiatric illness, circadian rhythm disorders, nocturnal panic, and substance-induced sleep disorders. Psychological components can also enhance compliance with treatment of sleep apnea and augment treatment of certain parasomnias (see Sections 8.22.5.4 and 8.22.5.5).

8.22.5.1 Circadian Rhythm Disorders

Properly timed exposures to bright light (2500–10 000 lux) for a few days to a week is an effective treatment for circadian rhythm disorders that are associated with shift in the phase of the sleep rhythm (Terman, 1994). As a means of comparison, typical indoor room light is less than 500 lux; a few minutes after dawn, sunlight produces about 2500 lux; and, at noon, sunlight is about 100 000 lux. The direction of the shift depends on the timing of the exposure to light. A phase advance is achieved by light exposure in the morning, while a phase delay is achieved by light exposure in the evening. In addition to the proper timing of bright light, it is also important to have periods of darkness during which no bright light is allowed. Possible side effects of light therapy include eye irritation and headache. Light therapy should not be used in patients with retinopathies.

Individuals with DSPS (Section 8.22.2.1) have no problem with further delaying their rhythm but find it impossible to fall asleep at an earlier time. This observation is the basis for chronotherapy. This consists of delaying bedtime by two to three hours a day until the desired bedtime is reached, usually within 10 days. Best results are obtained when this procedure follows a careful assessment of the factors that might reinforce very late bedtimes and thus reduce motivation for compliance. This assessment is particularly relevant when treating adolescents with a delayed sleep phase who might miss school days because they cannot wake up on time. The procedure is also enhanced when familial factors are taken into account and when close follow-up is in place. It is also possible to combine chronotherapy with a simulation of morning light and evening darkness to strengthen the circadian rhythm during the transition period. To prevent relapse, individuals must maintain a regular sleep schedule upon completion of chronotherapy.

Bright light treatment is also effective for circadian rhythm disorders related to shift work. Bright light during night shift work was documented to increase the duration of daytime sleep and to improve alertness on the job. Other methods for enhancing alertness and performance among shift workers have been suggested. The optimal timing of work breaks, social activity during breaks, bright light and other sensory stimulation have the most potential for short-term alerting effects (Penn & Bootzin, 1990). Stress coping techniques, sleep hygiene information, and family counseling have the most potential for addressing the long-term effects of shift work (Penn & Bootzin, 1990).

8.22.5.2 Primary Insomnia

The *Diagnostic and statistical manual of mental disorders* (4th ed.) (*DSM-IV*) defines primary insomnia as a disorder characterized by difficulty initiating or maintaining sleep, or by a complaint of nonrestorative sleep. A diagnosis of primary insomnia is given only when these symptoms satisfy each of the following three conditions: (i) they are not attributable solely to extrinsic factors or to other sleep or psychiatric disorder; (ii) they cause clinically significant negative daytime sequelae; and (iii) they have lasted a month or longer. In other words, primary insomnia is defined as insomnia that is attributable primarily to internal mechanisms.

Sedative hypnotics are the most common treatment for insomnia in general practice settings and in psychiatric clinics. Even though they are recommended only for short term use, hypnotic medications are frequently prescribed and used long after they have lost their efficacy. A number of effective behavioral interventions for insomnia have been designed and tested alone or in combination. These include stimulus-control instructions, sleep restriction, relaxation training, paradoxical intention, cognitive therapy, light therapy, and sleep hygiene education. McClusky, Milby, Switzer, Williams, and Woolens (1991) found that while benzodiazepine hypnotics were more immediately effective, behavioral therapy showed significant improvement within two weeks and proved to be superior to pharmacotherapy at a five-week follow-up. A combination cognitive-behavioral therapy with pharmacotherapy (triazolam) was recently found be superior to triazolam alone on a small sample of insomniacs (Milby et al., 1993). Physiological and psychological dependency complicate the treatment of insomnia in individuals who are chronic users of hypnotic medications. Preliminary results demonstrate that gradual taper in combination with cognitive behavioral components is a promising approach for treatment of insomnia patients with long term use of hypnotic medications (Kirmil-Gray, Eagleston, Thorensen, & Zarcone, 1985; Morin et al., 1994). The important elements of such an intervention include, a very gradual taper, ample support, preparation for rebound insomnia, and an early introduction of effective cognitive behavioral skills for improving sleep.

8.22.5.2.1 Stimulus-control instructions

Stimulus-control instructions have become the gold standard of behavioral treatments for insomnia (Bootzin, Epstein, & Wood, 1991; Bootzin & Nicassio, 1978). The instructions are designed to strengthen the bed as a cue for sleep and weaken it as a cue for activities that are not consistent with sleep. Stimulus-control instructions are also designed to strengthen internal cues for sleep and to stabilize the sleep–wake rhythm. These are achieved by instructing patients (i) to go to bed only when tired, (ii) not to use the bedroom for any activity (excluding sexual) other than sleep, (iii) to leave the bed and bedroom when unable to fall asleep, (iv) to wake up at a regular time regardless of how well or how long they slept at night, and (v) to avoid taking naps. Patients do well when the rationale for each instruction is explained as they are presented and when the patient's progress and/or difficulties are carefully monitored in daily sleep logs.

The goal of the first rule is to help patients become more sensitive to internal cues of sleepiness, and to increase the likelihood that they will fall asleep quickly. The second rule is designed to remove activities that are associated with arousal away from the bedroom. Reflection on the day's activities should be done in a place other than the bedroom. In addition, although good sleepers often read or watch television in bed with no ill effects, this is not the case with insomniacs. This rule is designed to help break maladaptive bedtime routines and replace them by routines that are less likely to interfere with sleep. The third rule is intended to disassociate the bed from frustration about not being able to fall asleep. By using the time out of bed to engage in pleasant relaxing activities, a patient may get a sense of increased control, rather than being a victim to frustration. As a result, the problem becomes more manageable and less distressing, thus breaking the cycle of sleeplessness and frustration. Keeping consistent wake times serves two roles. In the short run it is likely to produce some sleep deprivation that will facilitate sleep onset the following night, thus strengthening the association between the bed and sleep. In the long run, consistent wake times affect the biological clock regulating sleep and wakefulness and help develop a more stable sleep rhythm. The goal of the last rule is to prevent irregular napping that might destabilize the sleep rhythm. A short (less than one hour) nap that takes place daily at around the same time is permissible.

Research has shown the stimulus control procedure to be effective both for sleep onset insomnia (e.g., Bootzin, 1984) and for sleep maintenance insomnia (Morin & Azrin, 1987). Improvements in sleep usually begin after a week or two of following stimulus control instructions. Compliance issues need to be carefully examined and addressed for this method to work. For example, some patients become anxious about following these instructions properly and become worried about when to leave the bedroom, and when to return to bed. These patients typically do well when the main ideas behind the intervention are clearly outlined and they have an opportunity to modify the instructions, while staying within the conceptual goals of the instructions.

8.22.5.2.2 Bright light

Encouraging results have been reported for the use of bright light with insomniacs. These positive results have been taken as evidence that some forms of insomnia are in fact related to small shifts in the phase of the sleep wake rhythm. Morning light has been found to reduce sleep onset latency in sleep onset insomnia (Lack, Wright, & Paynter, 1995). Evening light has been found to delay wake-up time, increase total sleep (Lack & Wright, 1993), and increase sleep efficiency (Campbell, Dawson, & Anderson, 1993) in sleep maintenance insomniacs.

8.22.5.2.3 Relaxation

Progressive muscle relaxation training was found to be effective for sleep onset difficulties both for patients with primary insomnia and for patients with sleep-state misperception (Van Oot, Lane, & Borkovec, 1984). This and other relaxation procedures are believed to exert their therapeutic effect by reducing arousal at bedtime. Using meta-analytic techniques, Morin et al. (1994) concluded that cognitive relaxation methods are slightly superior to those targeting physiological arousal.

8.22.5.2.4 Paradoxical intention

Paradoxical intention for insomnia instructs the patient to stay awake as long as possible. The instruction is intended to reduce anticipatory anxiety at bedtime (Ascher & Turner, 1980). Some controlled studies found paradoxical intention to be effective for sleep onset insomnia (e.g., Ascher & Turner, 1979; Espie, 1989) and others have not (Lacks, Bertelson, Gans, & Kunkel, 1983; Turner & Ascher, 1982). It has been suggested that symptom prescription, that is, instructing the patient to stay awake, might increase a sense of voluntary control over sleep (Bateson, Jaclson, Haley, & Weakland, 1950), because if one successfully follows the paradoxical intention, one will gain control over sleep by "not sleeping." This will subsequently generalize to a sense of control in the other direction, that is, "sleeping." Shoham, Bootzin, Rohrbaugh, and Urry (1995) found that a voice indicator of defiance predicted success with paradoxical intention for sleep.

8.22.5.2.5 Sleep restriction

Sleep restriction (Spielman, Saskin, & Thorpy, 1987) is most commonly used for treating insomnia in older adults with poor sleep efficiencies (Hoelscher & Edinger, 1988). Patients are instructed to restrict the time in bed to their estimated total sleep time. Since patients tend to underestimate total sleep time, they experience partial sleep deprivation. This initial sleep deprivation helps consolidate sleep and increase sleep efficiency. Patients are then instructed to gradually increase time spent in bed as long as sleep efficiency remains high. With time, increased sense of control and self-efficacy leads to more gains.

8.22.5.2.6 Cognitive components

Cognitive components typically target a patient's expectations and attribution styles and are designed to increase a sense of self-efficacy. It is usually used in combination with one or more of the behavioral interventions reviewed above. The basic elements of cognitive treatments of insomnia include setting realistic expectations, decatastrophizing the consequences of a bad night, highlighting small gains, examining and changing maladaptive coping styles, and shifting the patient away from attributions that are inconsistent with good sleep.

The efficacy of cognitive-behavioral treatments of insomnia was recently summarized in two meta-analyses. Both analyses concluded that nonpharmacological interventions for insomnia produce reliable and durable clinical benefits that include reduction in latency to sleep onset and increase in total sleep time (Morin et al., 1994; Murtagh & Greenwood, 1995). For example, patients with sleep onset difficulties were better off after treatment than 81% of untreated controls and patients with sleep maintenance problems were better off after treatment than 74% of untreated controls (Morin et al., 1994). The improvements were not only statistically significant but also clinically significant. For example, both studies report that average time to sleep onset was reduced from approximately 60 minutes at baseline to approximately 35 minutes post-treatment. The two studies, however, did not come to the same conclusions regarding the relative efficacy of different nonpharmacological interventions. Morin and his colleagues found stimulus control and sleep restriction to be the two most effective interventions, whereas Murtagh and Greenwood found no differential efficacy when comparing nonpharmacological interventions to each other. The primary reason for this discrepancy is that the authors used different criteria for what studies qualified for inclusion. Although there was a considerable overlap in studies included, there were differences as well. A secondary reason for the discrepancy is that the authors of the two meta-analyses performed somewhat different analyses.

8.22.5.3 Treatment of Childhood Sleep Difficulties

The treatments of sleep difficulties during childhood are based on the same basic principles as the treatment of adults, but they need to take into account developmental factors. Behavioral interventions for infant sleep difficulties are based on extinction and stimulus control principles. An important component of treatment is helping parents teach their children to fall asleep on their own so that the presence of the parent does not become a necessary distinguishing stimulus for sleep onset. The treatment typically involves establishing a consistent and relaxing bedtime and discontinuing the active involvement of the caretaker in soothing the child. Discontinuation of parental presence in the room can be accomplished abruptly, or by gradually reducing the length of the time interval in which the caretaker is in the room. Both methods are effective but the abrupt method is more difficult for parents to follow (Ferber, 1985). Efficacy of these treatments is high. An equally effective and more gentle approach is based on the premise that separation anxiety is more central to the problem of sleep initiation than inadequate cues. This method involves temporary "co-sleeping." That is, the parent might sleep in the child's bedroom each night for about a week without having any other interaction with the child during the night and gradually decrease his or her presence in the child's bedroom with time (Sadeh, 1994).

During childhood, sleep onset difficulties could be manifested as bedtime struggle related to parent–child conflict, fear of the dark, or difficulties settling down. Specific interventions for each of these manifestations need to be applied. Behaviors associated with bedtime struggle can be treated using positive reinforcement or paradoxical intentions, but when the parent–child conflict is present beyond the bedtime period, the treatment needs to focus on the most prominent oppositional domain. In these cases, the resolution of conflict in the most prominent domain commonly leads to improvement in the bedtime difficulty as well (Cashman & McCann, 1988). Fear of the dark can be handled using cognitive rehearsal of appropriate bedtime behavior during the day, followed by reinforcement of appropriate bedtime behavior the following morning (Cashman & McCann, 1988). Self-statements, self-assessment, and self-reinforcement form an effective treatment for night-time fears of older children (Graziano, Mooney, Huber, & Ignasiak, 1979). Children's difficulties settling down can be treated using the same principles that are effective in adults. Stimulus control instructions and relaxation procedures modified to suit the child's development (Oaklander, 1981) are most appropriate.

8.22.5.4 Parasomnias

This group of sleep disorders is characterized by unusual sleep-related behaviors that are often of great concern to the patient and family members. Some parasomnias, such as sleep walking, night terrors, and sleep talking, are disorders of partial arousal from slow wave sleep and are most common during childhood. This set of parasomnias is considered normal in the very young, in whom slow wave sleep is abundant, and pathological when persisting into adulthood. Other parasomnias occur during REM sleep. The parasomnias require careful differential diagnoses. For example, night terrors, like nightmares and nocturnal panic, are episodes of intense fear associated with a frightening experience. Night terrors can be differentiated from nightmares in several ways. First, night terrors are associated with a fearful situation but lack the elaborate story-like cognition typical of nightmares. Second, nightmares arise during REM sleep and typically occur in the latter half of the sleep period whereas night terrors occur mostly in the first half. Finally, unlike nightmares, night terrors are associated with disorientation and amnesia for the event upon awakening. It is more difficult to differentiate night terrors from nocturnal panic. Both are associated with increased physiological arousal and a feeling of panic. Night terrors, however, usually begin with a loud scream.

Another difficult diagnostic issue is the differential diagnosis between sleep walking (somnambulism) and REM behavior disorders (RBD). Both involve unusual sleep behaviors. However, sleep walking occurs early in the night, as a partial arousal from slow wave sleep, and is not associated with enacting thoughts, whereas REM behavior disorders occur during REM sleep and the sleeper reports acting out a dream. The pathology of RBD involves intermittent loss of muscle atonia during REM sleep. These episodes often involve aggression or violence, placing the patient and bed partner at risk of injury. RBD is most prevalent among elderly males and may involve a familial component (ASDA, 1990), whereas sleep walking is more common in childhood. Both disorders need to be differentiated from nocturnal epileptic seizures that are often associated with unusual motor behavior during sleep. The movements associated with sleep-related seizures may be characterized by unilateral twitching or jerking of the face, arm or leg, or appear as nonpurposeful, automatic behaviors such as fidgeting of the hands or lip smacking. A referral to a neurologist can assist in the differential diagnosis between a seizure disorder and RBD. When the unusual sleep behavior is frequent, an all night sleep study might provide a definitive answer because nocturnal seizures are most likely to occur during non-REM sleep and are identifiable in the EEG by well defined "spike-and-wave" patterns (ASDA, 1990).

There are multiple approaches to the treatment of parasomnias that can be combined and individually tailored. Prescription of benzodiazepines is the most common treatment for both REM and non-REM parasomnias (Keefauver & Guilleminault, 1994; Mahowald & Schenck, 1994). Although long-term nightly use of benzodiazepines is not recommended, pharmacotherapy is important when dangerous behaviors are involved and when there is a change in the sleep environment so that the parasomnia might be more dangerous or embarrassing. Injuries related to odd sleep behaviors can be prevented by making the sleep environment safe (e.g., covering glass windows, locking doors and windows, and placing an alarm on the bedroom door). The stress diathesis theory of parasomnia (Moldofsky, Gilbert, Lue, & MacLean, 1995) postulates that the disorder arises as an interaction between a predisposing "genetic or constitutional factor," indicated by childhood onset parasomnia or a family history of parasomnia, and a percipitating triggering event such as psychological distress, unusual sleep schedule, or substance abuse. This theory suggests that relaxation and stress management techniques could be used to reduce the frequency of the events. Hypnosis has been reported efficacious for night terror and sleep walking (Reid, Ahmed, & Levie, 1981). Interventions designed to consolidate and lengthen sleep are important in the treatment of parasomnias because both sleep fragmentation and sleep deprivation exacerbate these disorders (Guilleminaul, Moscovitch, & Leger, 1995). Because sleepers might be confused when awakened, it is best not to arouse an individual from these episodes. Instead, the patient can be gently led back to bed.

Treatment of nocturnal enuresis, a parasomnia most common during childhood, includes psychological and pharmacological approaches. A recent review of the outcome literature concluded that: (i) the use of a "urine alarm" was the most effective method of treatment, (ii) psychological treatment was more effective than pharmacological treatment, and (iii) overall, some form of treatment was better than no treatment (Houts, Berman, & Abramson, 1994).

Preliminary results indicate that chronic nightmares, considered a parasomnia in the sleep disorders nosology, can be treated with a cognitive-behavioral technique called imagery rehearsal (Krakow, Kellner, Rathak, & Lambert, 1995). The patient is instructed to recall the nightmare in detail, change its content so it is no longer frightening, and visualize the new set of images.

8.22.5.5 Sleep Apnea

Continuous positive airway pressure (CPAP) is the most common treatment for sleep apnea. The patient wears a nasal mask that provides air pressure to the upper airways at a level that is individually adjusted. Oxygenation is indicated in some cases. Other treatment approaches for obstructive sleep apnea include surgery, dental appliances, and medications that stimulate breathing. Many of the complex issues associated with the treatment of sleep apnea remain beyond the scope of this chapter and the reader is referred to Berry (1995) for a comprehensive review. Instead, we highlight the potential role of psychology.

A common problem with CPAP therapy is a fear reaction to the mask. Although effective treatment of simple phobias exist, we know little about their efficacy in treating the claustrophobic reaction to CPAP in adults. Gradual exposure has been successfully used to help children adjust to the CPAP mask (Rains, 1995). A simple learning-based approach is appropriate when sleep apnea occurs predominantly in a particular sleeping position. It involves structuring the sleep environment so that the undesirable position is avoided. For example, sewing tennis balls to the back of the nightgown can gradually teach the sleeper to avoid the supine position that is most commonly associated with apnea. Weight loss is also relevant in the treatment of sleep apnea, when obesity contributes to the presentation. In a recent review Strobel and Rosen (1996) concluded that all but one of the weight loss studies reviewed demonstrated positive treatment effects. The target symptoms that have improved with weight loss included sleep disordered breathing, sleep quality, cardiovascular and pulmonary functions, and alertness and performance.

8.22.5.6 Restless Leg Syndrome and Periodic Limb Movements

The treatment of PLM disorder and RLS is typically pharmacological (see review by Montplaisir et al., 1994). Opioids and benzodiazepines, in particular clonazepam, are effective in reducing arousals associated with PLM, thus reducing sleep fragmentation and improving sleep quality (Walters et al., 1993). The results concerning clonazepam in treating RLS are less conclusive. Dopaminergic drugs such as L-dopa are effective at reducing the occurrence of PLMs (Kaplan, Allen, Buchholz, & Walters, 1993).

Vibratory stimulation is a nonpharmacological treatment of PLMs and RLS that includes 30 minutes of vibratory stimulation to the dorsoflexors of the feet and toes before bedtime. Vibratory stimulation significantly reduces the number of leg movements during sleep (Kovacevic-Ristanovic, Cartwright, & Lloyd, 1991), but is less effective than clonazepam at relieving symptoms of RLS (Montagna, Sassoli de Bianchi, Zucconi, Cirignotta, & Lugaresi, 1984).

8.22.5.7 Narcolepsy

Narcolepsy is a neurological disorder of unknown origin that involves a pathology of REM sleep mechanisms. The primary daytime symptom of the disorder is irresistible bouts of sleep or "sleep attacks" that recur every 2–3 hours. Narcolepsy may be distinguished from other causes of excessive daytime sleepiness by a unique constellation of features which include sleep onset REM periods, cataplexy, and hypnogogic hallucinations. "Sleep paralysis," a waking state associated with an inability to move, may also occur at sleep onset or at sleep termination. The diagnosis of narcolepsy needs to be confirmed by nocturnal polysomnograph, followed by daytime MSLT. These procedures document both level of sleepiness and presence of a REM episode close to sleep onset. Although the sleep of a narcolepsy patient may be fragmented with frequent awakenings and increased time in stage one, daytime sleepiness is a feature of the disorder and not a consequence of nocturnal sleep fragmentation (Lamphere et al., 1989).

The treatment of narcolepsy includes stimulant medications for sleepiness and noradrenergic reuptake blockers for cataplexy, sleep paralysis, and hypnogogic hallucination. Treatment also includes behavioral management. The clinician and the patient are often faced with the challenge of achieving a careful balance between optimal alertness and increased irritability and other peripheral and central side effects of amphetamines. This is particularly relevant when tolerance levels increase and higher doses of medication are needed. An important behavioral treatment component is the prescription for strategically timed short naps. Recommended nap duration is 15–20 minutes with at least three repeated naps because the refractory sleep period after a nap lasts only 90–120 minutes (Guilleminault, 1994). In addition, the large majority of narcoleptic patients are interested in psychotherapy and support groups (Garma & Marchand, 1994) to help them cope with the effects of the disorder on their social and family relationships and other aspects of life.

8.22.6 INTERACTIONS WITH OTHER DISORDERS

Many medical and psychiatric disorders are associated with disturbed sleep. Far from being comprehensive, the discussion below focuses on conditions that are likely to be most relevant to clinical psychologists.

8.22.6.1 Alzheimer's Disease

There are striking sleep changes in persons with dementia, including frequent nocturnal awakenings and decreased REM sleep. These changes parallel but exceed in magnitude the changes of "normal aging" (Prinz et al., 1982). Loss of slow wave sleep and an increase in the amount and frequency of night time wakefulness have been reported at early stages of Alzheimer's disease with significant amounts of daytime sleep emerging with advanced severity (Vitiello, Bliwise, & Prinz, 1992). Controversy exists regarding whether there is a decrease in REM sleep (Bliwise, 1993).

The aspect of sleep disturbance most likely to be brought to the attention of the clinician is what has been called sundowning. This describes late afternoon or early evening exacerbation of behavioral disruption and agitation. Sundowning is a complex and poorly defined syndrome. Several treatment approaches have been suggested (Bliwise, 1993). Pharmacological treatments have not been promising since hypnotics have minimal effect and neuroleptics have strong risks for side-effects in older persons. Behavioral treatments include regularly scheduled activities both in the day and during the usual night time awakenings, exposure to bright outdoor light, bright light therapy, and sleep restriction (Bliwise, 1993; Okawa et al., 1991; Satlin, Volicer, Ross, Herz, & Campbell, 1992).

8.22.6.2 Chronic Pain Disorders

Chronic pain is a common cause of sleep complaints. In a recent Gallop poll, 53% of all adults surveyed reported they experienced night time pain in the last year and 60% of these night time pain sufferers stated it disturbed their sleep (National Sleep Foundation, 1995b). Rheumatological disorders such as osteoarthritis, a degenerative joint disease, are a common cause of insomnia. During the day the joints are kept mobile which helps to maintain flexibility and reduce pain. These joints become stiff at night and sleep may be disturbed by sudden pain when the individual moves or attempts to turn over in bed. Adequate dosage of anti-arthritic medication before bedtime can greatly alleviate the problem.

Fibromyalgia is a chronic pain syndrome in which the primary symptom is deep muscle tenderness in the neck and shoulders, forearms, small of the back, and/or the thighs. Along with muscle tenderness, fibromyalgic patients also complain of morning aching, poor appetite, chronic fatigue, sleep disturbance, and nonrestorative sleep. Polysomnographic sleep studies indicate that 60–70% of fibromyalgic patients exhibit alpha-sleep patterns (Moldofsky, Scarisbrick, England, & Smyth, 1975; Russell & Campos, 1986) in which an alpha EEG frequency wave occurs simultaneously with slower EEG waves. Alpha sleep is often found in patients who complain of nonrestorative sleep. It has been proposed that alpha sleep may cause the musculoskeletal symptoms of fibromyalgia (Moldofsky & Scarisbrick, 1976) and that it may be involved more generally in disorders that affect the immune system.

Other chronic pain conditions that may precipitate complaints of insomnia include burning foot pain from diabetes and chest and epigastric pain due to angina, reflux esophagitis, or peptic ulcer disease. Cancer patients often exhibit more severely disturbed sleep patterns than either healthy people or other patients with medical conditions.

8.22.6.3 Acquired Immune Deficiency Disorder

Approximately one-third of asymptomatic HIV-seropositive individuals report sleep onset or maintenance difficulties (Rothenberg, Zozula, Funesti, & McAuliffe, 1990). Polysomnographic studies indicate increased sleep fragmentation, increased amounts of slow wave sleep, slow wave sleep distributed throughout the night, and an alpha-sleep pattern (Norman, Demirozu, & Chediak, 1990; Norman et al., 1989). Abnormal sleep architecture was evident at a one year follow-up (Norman, Chediak, Kiel, Gazeroglu, & Mendez, 1990).

8.22.6.4 Psychiatric Disorders

Sleep is both affected by and affects psychopathology. The complex interaction between sleep and psychopathology is most evident in affective and anxiety disorders. Epidemiological data suggest that the incidence of psychiatric disorders in patients complaining of insomnia may be underdiagnosed (Walsh, Moss, & Sugarman, 1994). About 21% of persons reporting serious insomnia met the *DSM-III* criteria for major depression (Mellinger, Balter, & Uhlenhuth, 1985) and up to 57% suffered from some mental disorder or developed one within one year (Ford & Kramerow, 1989). The high prevalence of psychiatric comorbidity in insomniacs poses a challenge to mental health providers and sleep clinicians alike. Should one treat the psychopathology first and then treat residual sleep symptoms, or should one target the sleep difficulties without waiting to see if treatment of the comorbid psychiatric condition resolves the sleep problem? A recent report on the patterns of symptomatic recovery in psychotherapy indicates that sleep complaints do not change as the result of general psychotherapy (Kopta et al., 1994) and implies that to change sleep complaints, they need to be specifically targeted.

8.22.6.4.1 Mood disorders

Sleep onset and sleep maintenance difficulties, as well as early morning awakening, are commonly experienced by patients with major depression. About 90% of depressed patients have some abnormality in their sleep that is most pronounced in endogenous, rather than reactive, depression (Neylan, 1995). The main alteration in sleep architecture of depressed patients is a decrease in latency to first REM episode and increase in REM sleep in the beginning of the night (Kupfer & Reynolds, 1992). It has been suggested that both the increase in REM sleep earlier in the night and the presence of early morning awakenings result from an advance in the sleep–wake rhythm (Schulz & Lund, 1985). A reduction in slow wave sleep is another sleep abnormality associated with depression and it is more pronounced in men than in women (Reynolds et al., 1990). However, these alterations in EEG sleep parameters are not unique to affective disorders (Gillin et al., 1990; Mendelson et al., 1987).

Daytime sleepiness and hypersomnolence may also be associated features of depression. Reynolds and Shipley (1985) estimated that 15–20% of depressed patients suffer from hypersomnia rather than insomnia. In these cases, however, the diagnosis was often cyclothymia or bipolar depression. Individuals with major depression tend to spend a greater amount of time in bed, a fact that might contribute to report of daytime somnolence more than any actual increase in the amount of time spent sleeping.

Sleep abnormalities of patients with bipolar disorder are similar to those found in patients with major depression. In addition, for these patients, a manic episode is typically associated with short sleep durations, long sleep latencies, and decreased slow wave sleep (ASDA, 1990). In contrast, the sleep of patients with dysthymia is comparable to normal controls (Arriaga, Rosado, & Paiva, 1990).

8.22.6.4.2 Anxiety disorders

Although anxiety can be an important etiological factor in insomnia, the proportion of self-reported insomniacs meeting diagnostic criteria (*DSM-III*) for anxiety disorders is only about 15% (Tan, Kales, Kales, Soldatos, & Bixler, 1984). Nevertheless, insomniacs frequently (50%) obtain high scores on anxiety scales (Mellinger et al., 1985) and there is a high prevalence (70%) of insomnia among patients with generalized anxiety disorder (Hoehn-Saric & McLeod, 1990). Both patients with generalized anxiety disorders and patients with obsessive-compulsive disorder experience difficulties initiating or maintaining sleep and frequent nocturnal awakenings. Alterations in sleep architecture of both groups of patients include reduction of slow wave sleep and a shortening of latency to the first REM period (Insel et al., 1982; Reynolds, Shaw, Newton, Coble, & Kurfer, 1983). Sleep onset difficulties of patients with obsessive-compulsive disorder are often related to obsessive or compulsive behaviors that focus on sleep, such as checking rituals (Insel et al., 1982).

Recurrent disturbing nightmares are considered to be an integral part of post-traumatic stress disorder (PTSD). A closer look at sleep related symptomatology in PTSD reveals that the disorder is associated with a greater number of arousals and awakenings compared with normals and with combat veterans without PTSD (Mellman, Kulick-Bell, Ashlock, & Nolan, 1995). Recurrent awakening in PTSD patients is associated not only with disturbing dreams but, equally as frequent, with startles or panic features (Mellman et al., 1995) and, less frequently, with thrashing movements during sleep. These symptomatic awakenings tend to be associated with REM sleep. In addition to increased number of awakenings, patients with PTSD spend a longer time awake and have lower sleep efficiencies than controls (e.g., Ross, Ball, Sullivan, & Caroff, 1989).

About 30–45% of patients with panic disorder report repeated nocturnal panic episodes (Mellman & Uhde, 1990). Nocturnal panic attacks commonly arise out of non-REM sleep stage two and early stage three (Mellman

& Uhde, 1989). They can also arise from stage one sleep during transition to sleep onset (Hauri, Friedman, & Ravaris, 1989). Patients with predominantly nocturnal panic have relatively little social or occupational disability and exhibit few phobic features (Mellman & Uhde, 1990).

8.22.7 PROFESSIONAL ISSUES

The assessment and treatment of sleep disorders is not the exclusive province of any one discipline. Consequently, sleep disorder centers are multidisciplinary and may include professionals from respiratory medicine, neurology, pediatrics, psychiatry, and nursing, in addition to clinical psychology. There are important roles to be played by all professionals and the field of sleep disorders is characterized by a mutual respect of the contributions of other disciplines. The major professional sleep organizations, such as the American Sleep Disorders Association, the Sleep Research Society, and the World Federation of Sleep Research Societies, are all multidisciplinary. In the USA, the American Board of Sleep Medicine certifies clinical psychologists who wish to become Board Certified Sleep Specialists. To be eligible to take the Board examinations, a psychologist, as well as any other individual holding a doctoral degree in a health-related field, needs to have acquired sufficient experience in sleep medicine under the supervision of a diplomate of the American Board of Sleep Medicine, working in a Sleep Disorder Center. Eligibility requirements also include sufficient experience in interpreting polysomnograms.

As a practical matter, clinical psychologists dealing with sleep disorders must be prepared to consult with and refer to other professionals, when necessary. For example, the treatment of insomnia may require having the patient gradually withdraw from hypnotics while cognitive-behavioral interventions are being implemented. The consent and supervision provided by the prescribing physician are important to the treatment of the patient as well as to the relationship of the psychologist with the medical community. It is often necessary to consult with the patient's physician during both the assessment and treatment of the sleep problem regarding physical disorders and medications that might contribute to the presenting sleep disturbance.

8.22.8 CONCLUSION

This chapter is not, nor could it be, a comprehensive review of the fast growing field of clinical sleep research. In writing this chapter the goal was to inform the clinical psychologist on important diagnostic and treatment considerations relevant to patients who complain about difficulty sleeping or about difficulties staying optimally alert. Behavioral and cognitive approaches have made significant contributions to understanding the etiology, maintenance, and treatment of sleep disorders. Their efficacy has been demonstrated for the treatment of insomnia yet the efficacy of cognitive-behavioral approaches for treating other sleep disorders, although promising, remains open. In particular, further research is needed to asses the efficacy of cognitive–behavioral approaches for: (i) increasing adherence with CPAP, (ii) facilitating weight loss and weight control when excessive weight is exacerbating sleep apnea, (iii) reducing frequency of nonREM parasomnias, (iv) reducing chronic use of hypnotics, and (v) coping with chronic sleep disorders, such as narcolepsy.

More information on specific sleep disorders and their treatment can be found in the *International classification of sleep disorders: Diagnostic and coding manual* (ASDA, 1990) and in *Principles and practices of sleep medicine* (Kryger, Roth, & Dement, 1994).

8.22.9 REFERENCES

Albert, M. S. (1981). Geriatric neuropsychology. *Journal of Consulting and Clinical Psychology, 49,* 835–850.

American Sleep Disorders Association (1990). *The international classification of sleep disorders: Diagnostic and coding manual.* Rochester, MN: Author.

American Sleep Disorders Association (1995a). Practice parameters for the use of polysomnograph in the evaluation of insomnia. *Sleep, 18,* 55–57.

American Sleep Disorders Association (1995b). Practice parameters for the use of actigraphy in the clinical assessment of sleep disorders. *Sleep, 18,* 285–287.

Anch, A. M., Saskin, P., & Moldofsky, H. (1989). Behaviorally signaled awakenings in normals and fibrositis patients. *Sleep Research, 18,* 327.

Armitage, R., Trivedi, M., Rush, A. J., & Hoffmann, R. (1995). The effects of fluoxetine on period-analyzed sleep EEG in depression. *Sleep Research, 24,* 381.

Arriaga, F., Rosado, P., & Paiva, T. (1990). The sleep of dysthymic patients: A comparison with normal controls. *Biological Psychiatry, 27,* 649–656.

Ascher, L. M., & Turner, R. M. (1979). Paradoxical intention and insomnia: An experimental investigation. *Behavior Research and Therapy, 17,* 408–411.

Ascher, L. M., & Turner, R. M. (1980). A comparison of two methods for the administration of paradoxical intention. *Behavior Research and Therapy, 18,* 121–126.

Balter, M. B., & Uhlenhuth, E. H. (1992). New epidemiologic findings about insomnia and its treatment. *Journal of Clinical Psychiatry, 53*(Suppl.) 34–42.

Bateson, G., Jaclson, D., Haley, J., & Weakland, J. (1956). Toward a theory of schizophrenia. *Behavioral Science, 1,* 251–264.

Berry, R. B. (1995). Sleep related breathing disorders. In R. B. George (Ed.), *Chest medicine: Essentials of pulmonary and critical care medicine* (3rd ed., pp. 125–140). Baltimore: Williams & Wilkins.

Billiard, M., Alperovitch, A., Perot, C., & Jammes A.

(1987). Excessive daytime somnolence in young men. Prevalence and contributing factors. *Sleep, 10,* 297–305.

Bliwise, D. L. (1993). Sleep in normal aging and dementia. *Sleep, 16,* 40–81.

Bliwise, D. L. (1994). Normal aging. In M. H. Kryger, T. Roth, & W. G. Dement (Eds.), *Principles and practice of sleep medicine* (2nd ed., pp. 26–39). Philadelphia: Saunders.

Bonnet, M. H., & Arand, D. L. (1995). Twenty-four hour metabolic rate in insomniacs and matched normal sleepers. *Sleep, 18,* 581–588.

Bonnet, M. H., & Arand, D. L. (1996). The consequences of a week of insomnia. *Sleep, 19,* 453–461.

Bonnet, M. H., & Johnson, L. C. (1978). Relationship of arousal threshold to sleep stage distribution and subjective estimates of depth and quality of sleep. *Sleep, 1,* 161–168.

Bootzin, R. R. (1984). Evaluation of stimulus control instructions, progressive relaxation, and sleep hygiene as treatments for insomnia. In W. P. Koella, E. Ruther, & H. Schultz (Eds.), *Sleep* (pp. 142–144). Stuttgart, Germany: Fischer Verlag.

Bootzin R. R., Bell, I. R., Halbisch, R., Kuo, T. F., Wyatt, J. K., Rider, S. P., & Manber, R. (1995). Night-to-night variability in measures of sleep and sleep disorders: A six night PSG study. *Sleep Research, 24,* 121.

Bootzin, R. R., Epstein, D., & Wood, J. M. (1991). Stimulus control instructions. In P. J. Hauri (Ed.), *Case studies in insomnia.* New York: Plenum.

Bootzin, R. R., & Nicassio, P. (1978). Behavioral treatments for insomnia. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification.* New York: Academic Press.

Bootzin, R. R., Shoham, V., & Kuo, T. F. (1994). Sleep anticipatory anxiety questionnaire: A measure of anxiety about sleep. *Sleep Research, 23,* 188.

Borkovec, T. D., Lane T. W., & VanOot, P. H. (1981). Phenomenology of sleep among insomniacs and good sleepers: wakefulness experience when cortically asleep. *Journal of Abnormal Psychology, 90,* 607–609.

Brown, S. L., Salive, M. E., Pahor, M., Foley, D. J., Corti, M. C., Langlois, J. A., Wallace, R. B., & Harris, T. B. (1995). Occult caffeine as a source of sleep problems in an older population. *Journal of the American Geriatric Society, 43,* 860–864.

Campbell, S. S., Dawson, D., & Anderson, M. W. (1993). Alleviation of sleep maintenance insomnia with timed exposure to bright light. *Journal of the American Geriatrics Society, 41,* 829–836.

Carskadon, M. A., Brown, E. D., & Dement, W. C. (1982). Sleep fragmentation in the elderly: Relationship to daytime sleep tendency. *Neurobiology of Aging, 3,* 321–327.

Carskadon, M. A., & Dement, W. C. (1981). Cumulative effects of sleep restriction on daytime sleepiness. *Psychophysiology, 18,* 107–113.

Carskadon, M. A., & Dement, W. C. (1987). Sleepiness in the normal adolescent. In C. Guilleminault (Ed.), *Sleep and its disorders in children* (pp. 53–66). New York: Raven.

Carskadon, M. A., Dement, W. C., Mitler, M. M., Guilleminault, C., Zarcone, V. P., & Spiegel, R. (1976). Self reports versus sleep laboratory findings in 122 drug-free subjects with complaints of chronic insomnia. *American Journal of Psychiatry, 133,* 1382–1388.

Carskadon, M. A., Mancuso, J., & Rosekind, M. R. (1989). Impact of part-time employment on adolescent sleep patterns. *Sleep Research, 18,* 114.

Carskadon, M. A., Vieira, C., & Acebo, C. (1993). Association between puberty and delayed sleep phase preferences. *Sleep, 16,* 258–262.

Cashman, M. A., & McCann, B. S. (1988). Behavioral approaches to sleep/wake disorders in children and adolescents. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification* (pp. 215–283). Newbury Park, CA: Sage.

Cole, R. J., & Kripke, D. F. (1988). Progress in automatic sleep/wake scoring by wrist actigraph. *Sleep Research, 17,* 331.

Czeisler, C. A., Kronauer, R. E., Allen, J. S., Duffy, J. F., Jewett, M. E., Brown, E. N., & Ronda, J. M. (1989). Bright light induction of strong (type 0) resetting of the human circadian pacemaker. *Science, 244,* 1328–1333.

Engle-Friedman, M., Baker, E. A., & Bootzin, R. R. (1985). Reports of wakefulness during EEG identified stages of sleep. *Sleep Research, 14,* 152.

Espie, C. A., Lindsay, W. R., Brooks, D. N., Hood, E. M., & Turvey, T. (1989). A controlled comparative investigation of psychological treatments for chronic sleep-onset insomnia. *Behavior Research and Therapy, 27,* 79–88.

Feinberg, I. (1989). Effects of maturation and aging on slow wave sleep in man: Implications for neurobiology. In A. Wauquier, C. Dugovic, & M. Radulovacki (Eds.), *Slow wave sleep physiological, pathophysiological and functional aspects* (pp. 31–48). New York: Raven.

Feinberg, I., Koresko, R., & Heller, N. (1967). EEG sleep patterns as a function of normal and pathological aging in man. *Journal of Psychiatric Research, 5,* 107–144.

Ferber, R. (1985). *Solve your child's sleep problems.* New York: Simon & Schuster.

Ford, D. E., & Kramerow, D. B. (1989). Epidemiological study of sleep disturbances and psychiatric disorders: An opportunity for prevention? *Journal of the American Medical Association, 262,* 1479–1484.

Forsythe W. I., & Redmond A. (1974). Enuresis and spontaneous cure rate. Study of 1,129 enuretics. *Archives of Disease in Childhood, 49,* 259–263.

Franklin, J. (1981). The measurement of sleep onset latency in insomnia. *Behavioral Research and Therapy, 19,* 547–549.

Garma, L., & Marchand, F. (1994). Non-pharmacological approaches to the treatment of narcolepsy. *Sleep, 17,* S97–S102.

Gillin, J. C. (1994). Sleep and psychoactive drugs of abuse and dependence. In M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and practices of sleep medicine* (pp. 934–942). Philadelphia: Saunders.

Gillin, J. C., Smith, T. L. Irwin, M., Kripke, D. F. Brown, S., & Schuckit, M. (1990). Short REM latency in primary alcoholic patients with secondary depression. *American Journal of Psychiatry, 147,* 106–109.

Graziano, A. M., Mooney, K. C., Huber, C., & Ignasiak, D. (1979). Self-control instructions for children's fear-reduction. *Journal of Behavior Therapy and Experimental Psychiatry, 10,* 221–227.

Greenblatt, D. J. (1992). Pharmacology of benzodiazepine hypnotics. *Journal of Clinical Psychiatry, 53*(Suppl.), 7–13.

Guilleminault, C. (1994). Narcolepsy syndrome. In M. H. Kryger, T. Roth, & W. G. Dement (Eds.), *Principles and practices of sleep medicine* (pp. 549–561). Philadelphia: Saunders.

Guilleminault C., Moscovitch, A., Leger, D. (1995). Forensic sleep medicine: Nocturnal wandering and violence. *Sleep, 18,* 740–748.

Guilleminault, C., Stoohs, R., Nouriani, B., Cetel, M., Maistros, P., & Clerk, A. (1991). The upper airway resistance syndrome: Help for patients with idiopathic hypersomnia. *Sleep Research, 20A,* 494.

Haimov, I., Lavie, P., Laudon, M., Herer, P., Vigder, C., & Zisapel, N. (1995). Melatonin replacement therapy of elderly insomniacs. *Sleep, 18,* 598–603.

Hauri, P. J. Friedman, M., & Ravaris, C. L. (1989). Sleep in patients with spontaneous panic attacks. *Sleep, 12,* 323–337.

Hauri, P., & Olmstead, E. (1980). Childhood onset insomnia. *Sleep, 3,* 59–65.

Henschel, A., & Lack, L. (1987). Do many adolescents sleep poorly or just too late? *Sleep Research, 16,* 354.

Hoch, C. C., Reynolds, C. F. III, Kupfer, D. J., Houck, P. R., Berman, S. R., & Stack, J. A. (1987). The superior sleep of healthy elderly nuns. *International Journal of Aging and Human Development, 25,* 1–9.

Hoehn-Saric, R., & McLeod, D. R. (1990). Generalized anxiety disorder in adulthood. In M. Hersen and C. G. Last (Eds.), *Handbook of child and adult psychopathology: A longitudinal perspective* (pp. 247–260). New York: Pergamon.

Hoelscher, T. J., & Edinger, J. D. (1988). Treatment of sleep-maintenance insomnia in older adults: Sleep period reduction, sleep education, and modified stimulus control. *Psychology and Aging, 3,* 258–263.

Houts, A. C., Berman, J. S. & Abramson, H. (1994). Effectiveness of psychological and pharmacological treatments for nocturnal enuresis. *Journal of Consulting & Clinical Psychology, 62,* 737–745.

Insel, T. R., Gillin, C., Moore, A., Mendelson, W. B., Loewenstein, R. J., & Murphy, D. L. (1982). The sleep of patients with obsessive-compulsive disorder. *Archives of General Psychiatry, 39,* 1372–1377.

Isono, S., & Remmers, J. E. (1994). Anatomy and physiology of upper airway obstruction. In M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and practices of sleep medicine* (pp. 642–656). Philadelphia: Saunders.

Kahn, A., Rebuffat, E., Blum, D., Casimir, J., Duchateau, J., Mozin, M. J., & Jost, R. (1987). Difficulty in initiating and maintaining sleep associated with cow's milk allergy in infants. *Sleep, 10,* 116–121.

Kaplan, P. W., Allen, R. P., Buchholz, D. W., & Walters, J. K. (1993). A double blind, placebo-controlled study of the treatment of periodic limb movements in sleep using carbidopa/levidopa and propoxyphene. *Sleep, 16,* 717–723.

Keefauver, S. P., & Guilleminault, C. (1994). Sleep terrors and sleep walking. In M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and practices of sleep medicine* (pp. 567–573). Philadelphia: Saunders.

Kirmil-Gray, K., Eagleston, J. R., Thorensen, C. E., & Zarcone, V. P., Jr. (1985). Brief consultation and stress management treatment for drug-dependent insomnia: Effects on sleep quality, self-efficacy, and day-time stress. *Journal of Behavioral Medicine, 8,* 79–99.

Kopta, S. M., Howard, K. I., Lowry, J. L., & Beutler, L. E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology, 62,* 1009–1016.

Kovacevic-Ristanovic, R., Cartwright, R. D., & Lloyd, S. (1991). Nonpharmacologic treatment of periodic leg movements in sleep. *Archives of Physical Medicine & Rehabilitation, 72,* 385–389.

Krakow, B., Kellner, R., Pathak, D., & Lambert, L. (1995). Imagery rehearsal treatment for chronic nightmares. *Behavior Research and Therapy, 33,* 837–843.

Kryger, M. H., Roth, T., & Dement, W. C. (Eds.) (1994). *Principles and practices of sleep medicine*. Philadephia: Saunders.

Kupfer, D. J., & Reynolds, C. F., III. (1992). Sleep and affective disorders. In E. S. Paykel (Ed.), *Handbook of affective disorders* (2nd ed., pp. 311–323). New York: Guilford.

Lack, L., & Wright, H. (1993). The effect of evening bright light in delaying the circadian rhythms and lengthening the sleep of early morning awakening insomniacs. *Sleep, 16,* 436–443.

Lack, L., Wright, H., & Paynter, D. (1995). The treatment of sleep onset insomnia with morning bright light. *Sleep Research, 24A,* 338.

Lacks, P., Bertelson, A. D., Gans, L., & Kunkel, J. (1983). The effectiveness of three behavioral treatments for different degrees of sleep-onset insomnia. *Behavior Therapy, 14,* 593–605.

Lamphere, J., Young, D., Roehrs, T., Witting, R. M., Zorick, F., & Roth, T. (1989). Fragmented sleep, daytime somnolence, and age in narcolepsy. *Clinical Electroencephalography, 20,* 49–54.

Lee, K. A. (1988). Circadian temperature rhythms in relation to menstrual cycle phase. *Journal of Biological Rhythms, 3,* 255–263.

Leger, D. (1994). The cost of sleep-related accidents: A report for the National Commission of Sleep Disorders Research. *Sleep, 17,* 84–93.

Mahowald, M. W., & Schenck, C. H. (1994). REM sleep behavior disorder. In M. H. Kryger, T. Roth, & W. C. Dement (Eds.) *Principles and practices of sleep medicine* (pp. 574–588). Philadelphia: Saunders.

Manber, R., & Bootzin, R. R. (1991). The effect of regular wake-up schedule on daytime sleepiness in college students. *Sleep Research, 20,* 284.

Manber, R., & Bootzin, R. R. (1997). Sleep and the menstrual cycle. *Health Psychology, 16,* 209–214.

Manber, R., Bootzin, R. R., Acebo, C., & Carskadon, M. A. (1996). The effects of regularizing sleep–wake schedules on daytime sleepiness in college students. *Sleep, 19,* 432–441.

Manber, R., Pardee, E. R., Bootzin, R. R., Kuo, T., Rider, A. M., Rider, P. S., & Bergstrom, L. (1995a). Changing sleep patterns in adolescence. *Sleep Research, 24,* 106.

Manber R., Pardee, E. R., Bootzin, R. R., Kuo, T., Rider, A. M., Rider, P. S., & Bergstrom, L. (1995b). The relationship between sleepiness and sleep-patterns in three cohorts of students. *Sleep Research, 24,* 105.

Martin, S. E., Engleman, H. M., Deary, I. J., & Douglas, N. J. (1996). The effect of sleep fragmentation on daytime function. *American Journal of Respiratory and Critical Care Medicine, 153,* 1328–1332.

Matalon, A., Yinnon, A. M., & Hurwitz, A. (1990). Chronic use of hypnotics in a family practice—patient's reluctance to stop treatment. *Family Practice, 7,* 258–260.

McClusky, H. Y., Milby, J. B., Switzer, P. K., Williams, V., & Wooten, V. (1991). Efficacy of behavioral versus triazolam treatment in persistent sleep-onset insomnia. *American Journal of Psychiatry, 148,* 121–126.

Mellinger, G. D., Balter, M. B., & Uhlenhuth, E. H. (1985). Insomnia and its treatment: prevalence and correlates. *Archives of General Psychiatry, 42,* 225–232.

Mellman, T. A., Kulick-Bell, R., Ashlock, L. E., & Nolan, B. (1995). Sleep events among veterans with combat-related posttraumatic stress disorder. *American Journal of Psychiatry, 152,* 110–115.

Mellman, T. A., & Uhde, T. W. (1989). Electroencephalographic sleep in panic disorder: A focus on sleep related panic attack. *Archives of General Psychiatry, 46,* 178–184.

Mellman, T. A., & Uhde, T. W. (1990). Patients with frequent sleep panic: Clinical findings and response to medication treatment. *Journal of Clinical Psychiatry, 51,* 513–516.

Mendelson, W. B., Sack, D. A., James, S. P., Martin, J. V., Wagner, R., Garnett, D., Milton, J., & Wehr, T. A. (1987). Frequency analysis of the sleep EEG in depression. *Psychiatry Research, 21,* 89–84.

Milby, J. B., Williams, V., Hall, J. N. Khuder, S., McGill, T., & Wooten, V. (1993). Effectiveness of combined triazolam-behavioral therapy for primary insomnia. *American Journal of Psychiatry, 150,* 1259–1260.

Minot, R., Luthringer, R., & Macher, J. P. (1993). Effects of miclobemide on the psychophysiology of sleep/wake cycles: A neuroelectrophysiological study of depressed patients administered with miclobemide. *International*

Clinical Psychopharmaology, 7, 181–189.

Mitler, M. (1992). Maintenance of wakefulness test: A polysomnographic technique for evaluating treatment efficacy in patients with excessive somnolence. *Electroencephalography and Clinical Neurophysiology, 53,* 658–661.

Moldofsky, H., & Scarisbrick, P. (1976). Induction of neuroasthenic musculoskeletal pain syndrome by selective sleep stage deprivation. *Psychosomatic Medicine, 38,* 35–44.

Moldofsky, H., Scarisbrick, P., England, R., & Smythe, H. (1975). Musculoskeletal symptoms and non-REM sleep disturbance in patients with "fibrositis syndrome" and healthy subjects. *Psychosomatic Medicine, 37,* 341–351.

Moldofsky, H., Gilbert, R., Lue, F. A., & MacLean, A. W. (1995). Sleep-related violence. *Sleep, 18,* 731–739.

Monk, T. M. (1994). Shift Work. In M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and practices of sleep medicine* (pp. 471–481). Philadelphia: Saunders.

Monroe, L. J. (1967). Psychological and physiological differences between good and poor sleepers. *Journal of Abnormal Psychology, 2,* 255–264.

Montagna, P., Sassoli de Bianchi, L., Zucconi, M., Cirignotta, F., & Lugaresi, E. (1984). Clonazepam and vibration in restless legs syndrome. *Acta Neurologica Scandinavica, 69,* 428–430.

Montplaisir, J., Godbout, R., Pelletier, G., & Warnes, H. (1994). Restless legs syndrome and periodic limb movements during sleep. In M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and practices of sleep medicine* (pp. 589–597). Philadelphia: Saunders.

Moore-Ede, M. C., Martin C., Sulzman, F. M., & Fuller, C. A. (1982). *The clocks that times us: Physiology of the circadian timing system*. Cambridge, MA: Harvard University Press.

Morin, C. M. (1993). *Insomnia: Psychological assessment and management*. New York: Guilford.

Morin, C. M., & Azrin, N. H. (1987) Stimulus control and imagery training in treating sleep-maintenance insomnia. *Journal of Consulting and Clinical Psychology, 55,* 260–262.

Morin, C. M., Culbert, J. P., & Schwartz, S. M. (1994). Nonpharmacological interventions for insomnia: A meta-analysis of treatment efficacy. *American Journal of Psychiatry, 151,* 1172–1180.

Morin, C. M., & Kwentus, J. A. (1988). Behavioral and pharmacological treatments for insomnia. *Annals of Behavioral Medicine, 10,* 91–100.

Murtagh, D. R., & Greenwood, K. M. (1995). Identifying effective psychological treatments for insomnia: A meta-analysis. *Journal of Consulting and Clinical Psychology, 63,* 79–89.

National Sleep Foundation (1991). *Sleep in America: A national survey of US adults.* Princeton, NJ: Gallup Organization.

National Sleep Foundation (1995a). *Sleep in America.* Princeton, NJ: Gallup Organization.

National Sleep Foundation (1995b). *Adult public's experiences with nighttime pain.* Princeton, NJ: Gallup Organization.

Neylan, T. C. (1995). Treatment of sleep disturbances in depressed patients. *Journal of Clinical Psychiatry, 56* (Suppl 2), 56–61.

Nicholson, A. N. (1994). Hypnotics: Clinical pharmacology and therapeutics. In M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and practice of sleep medicine* (2nd ed., pp. 355–363). Philadelphia: Saunders.

Norman, S. E., Chediak, M., Kiel, H., Gazeroglu, H., & Mendez, A. (1990). HIV infection and sleep: Follow up studies. *Sleep Research, 19,* 339.

Norman, S. E., Demirozu, M. C., & Chediak, M. (1990). Distorted sleep architecture (NREM/REM sleep cycles) in HIV healthy infected men. *Sleep Research, 19,* 340.

Norman, S. E., Kiel, M., Nay, K. M., Demirozu, M. C., Dunbar, S., & Cohn, M. A. (1989). Alpha–delta sleep pattern and circadian rhythmicity in HIV seropositive patients. *Sleep Research, 18,* 353.

Okawa, A., Mishima, K., Hishikawa, Y., Hozumi, S., Hori, H., & Takahashi, K. (1991). Circadian rhythm disorders in sleep-waking and body temperature in elderly patients with dementia and their treatment. *Sleep, 14,* 478–485.

Oaklander V. (1988). *Windows to our children: a Gestalt therapy approach to children and adolescents.* Highland, NY: Center for Gestalt Development.

Parry, B. L., Mendelson, W. B., Duncan, W. C., Sack, D. A., & Wehr, T. A. (1989). Longitudinal sleep EEG, temperature, and activity measurements across the menstrual cycle in patients with premenstrual depression and age matched controls. *Psychiatry Research, 30,* 285–303.

Penn, P. E., & Bootzin, R. R. (1990). Behavioral techniques for enhancing alertness and performance in shift work. *Work & Stress, 4,* 213–226.

Price, V. A, Coates, T. J., Thoresen, C. E., & Grinstead, O. A. (1978). Prevalence and correlates of poor sleep among adolescents. *American Journal of the Disabled Child, 132,* 583–586.

Prinz, P., Peskind, E. R., Vitaliano, P. P., Raskind, M. A., Eisdorfer, C., Zemcuznikov, N., & Gerber, C. J. (1982). Changes in the sleep and waking EEGs of nondemented and demented elderly subjects. *Journal of the American Geriatrics Society, 30,* 86–93.

Quera-Salva, M. A., Orluc, A., Goldenberg, F., & Guilleminault, C. (1991). Insomnia and use of hypnotics: Study of a French population. *Sleep, 14,* 386–391.

Rains, J. C. (1995). Treatment of obstructive sleep apnea in pediatric patients, behavioral intervention for compliance with nasal continuous positive airway pressure. *Clinical Pediatric, 34,* 535–541.

Reid, W. H., Ahmed, I., & Levie, C. A. (1981). Treatment of sleep walking: A controlled study. *American Journal of Psychotherapy, 35,* 27–37.

Reite, M., Buysse, D., Reynolds, C., & Mendelson, W. (1995). The use of polysomnography in the evaluation of insomnia. *Sleep, 18,* 58–70.

Reynolds, C. F. III, Buysse, D. J., Nofzinger, E. A., & Kupfer, D. J. (1995). Treatment of persistent insomnia in psychiatric patients: effects of antidepressant therapies on sleep in mood disorders. In A. Schatzberg & C. B. Nemeroff (Eds.), *Textbook on psychopharmacology* Washington, DC: American Psychiatric Press.

Reynolds, C. F. III, Kupfer, D. J., Thase, M. E., Frank, E., Jarret, D. B., Coble, P. A., Hoch, C. C., Buysse, D. J., Simons, A. D., & Houck, P. R. (1990). Sleep, gender, and depression: An analysis of gender effects on the electroencephalographic sleep of 302 depressed outpatients. *Biological Psychiatry, 28,* 673–684.

Reynolds C. F. III, Kupfer, D. J., Taska, L. S., Hoch, C. C., Sewitch, D. E., & Spiker, D. G. (1985). Sleep of healthy seniors: A revisit. *Sleep, 8,* 20–29.

Reynolds III, C. F., Shaw, D. H., Newton, T. F., Coble, P. A., & Kupfer, D. J. (1983). EEG sleep in outpatients with generalized anxiety: *a preliminary comparison with depressed outpatients. Psychiatry Research, 8,* 81–89.

Reynolds, C. F. III, & Shipley, J. E. (1985). Sleep in depressive disorders. In: R. E. Hales & A. J. Frances (Eds.), *Psychiatry update: The American Psychiatric Association annual review* (Vol. 4., pp. 341–351). Washington, D.C: American Psychiatric Press.

Richman, N. (1987). Surveys of sleep disorders in children in a general population. In C. Guilleminault (Ed.), *Sleep and its disorders in children* (pp. 115–127). New York: Raven.

Richman, N., Stevenson, J., & Graham, P. J. (1982). *Preschool to school: a behavioral study.* London: Academic Press.

Robinson, R. W., & Zwillich, C. W. (1994). Drugs and sleep respiration. In M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and practices of sleep medicine* (pp. 603–620). Philadelphia: Saunders.

Roehrs, T., Zorick, F., Wittig, R., Conway, W., & Roth, T. (1989). Predictors of objective level of daytime sleepiness in patients with sleep-related breathing disorders. *Chest, 95,* 1202–1206.

Ross, R. J., Ball, W. A., Sullivan, K. A., & Caroff, S. N. (1989). Sleep disturbance as the hallmark of post-traumatic stress disorder. *American Journal of Psychiatry, 146,* 679–707.

Rothenberg, S., Zozula, R., Funesti, J., & McAuliffe, V. (1990). Sleep habits in asymptomatic HIV-seropositive individuals. *Sleep Research, 19,* 342.

Sadeh, A. (1994). Assessment of intervention for infant night waking: parental reports and activity-based home monitoring. *Journal of Consulting and Clinical Psychology, 62,* 63–68.

Sadeh, A., Alster, J., Urbach, D., & Lavie, P. (1989). Actigraphically based automatic bedtime sleep–wake scoring: Validity and clinical applications. *Journal of Ambulatory Monitoring, 2,* 209–216.

Sanchez, R., & Bootzin, R. R. (1985). A comparison of white noise and music: Effects of predictable and unpredictable sounds on sleep. *Sleep Research, 14,* 121.

Satinoff, E. (1988). Thermal influences on REM sleep. In R. Lydic, & J. Biebuyck (Eds.), *Clinical physiology of sleep* (pp. 135–144). Bethesda, MD: American Psychological Society.

Satlin, A., Volicer, L., Ross, V., Herz, L., & Campbell, S. (1992). Bright light treatment of behavioral and sleep disturbances in patients with Alzheimer's Disease. *American Journal of Psychiatry, 149,* 1028–1032.

Schulz, H., & Lund, R. (1985). On the origin of early REM episodes in the sleep of depressed patients: A comparison of three hypotheses. *Psychiatry Research, 16,* 65–77.

Shoham, V., Bootzin, R. R., Rohrbaugh, M., & Urry, H. (1995). Paradoxical versus relaxation treatment for insomnia: The moderating role of reactance. *Sleep Research, 24A,* 365.

Silva, J. A. C. E., Chase, M., Sartorius N., & Roth, T. (1996). Special report from a symposium held by the World Health Organization and the World Federation of Sleep Research Societies: An overview of insomnias and related disorders—recognition, epidemiology, and rational management. *Sleep 19,* 412–416.

Spielman, A. J., Saskin, P., & Thorpy, M. J. (1987). Treatment of chronic insomnia by restriction of time in bed. *Sleep, 10,* 45–56.

Stepanski, E., Zorick, F., Roehrs, T., Young, D., & Roth, T. (1988). Daytime alertness in patients with chronic insomnia compared with asymptomatic control subjects. *Sleep, 11,* 54–60.

Strauch, I., & Meier, B. (1988). Sleep need in adolescents: A longitudinal approach. *Sleep, 11,* 378–386.

Strobel, R. J., & Rosen R. C. (1996). Obesity and weight loss in obstructive sleep apnea: a critical review. *Sleep, 19,* 104–115.

Tan, T. L., Kales, J. D., Kales, A., Soldatos, C. R., & Bixler, E. O. (1984). Biopsychobehavioral correlates of insomnia, IV: diagnoses based on DSM-III. *American Journal of Psychiatry, 141,* 357–362.

Terman, M. (1994). Light therapy. In M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and practice of sleep medicine* (2nd ed., pp. 1012–1029). Philadelphia: Saunders.

Turner, R. M., & Ascher, L. M. (1982). Therapist factor in the treatment of insomnia. *Behavior Research and Therapy, 20,* 33–40.

VanOot, P. H., Lane, T. W., & Borkovec, T. D. (1984). Sleep disturbances. In H. E. Adams & P. B. Sutker (Eds.), *Comprehensive handbook of psychopathology* (pp. 683–723). New York: Plenum.

Vitiello, M. V., Bliwise, D. L., & Prinz, P. N. (1992). Sleep in Alzheimer's disease and the sundown syndrome. *Neurology, 42*(Suppl. 6), 83–93.

Walsh, J. K., Moss, K. L., & Sugarman, J. (1994). Insomnia in adult psychiatric disorders. In M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and practice of sleep medicine* (pp. 500–508). Philadelphia: Saunders.

Walters, A. S., Wagner, M. L., Hening, W. A., Grasing, K., Mills, R., Chokroverty, S., & Kavey, N. (1993). Successful treatment of the idiopathic restless legs syndrome in a randomized double-blind trial of oxycodone versus placebo. *Sleep, 16,* 327–332.

Ware, J., Russell, J., & Campos, E. (1986). Alpha intrusions into the sleep of depressed and fibromyalgia syndrome (fibrositis) patients. *Sleep Research, 15,* 210.

Webb, W. B., & Campbell, S. S. (1980). Awakenings and the return to sleep in an older population. *Sleep, 3,* 41–46.

Wetter, D. W., Fiore M. C., Baker, T. B., & Young, T. B. (1995). Tobacco withdrawal and nicotine replacement influence objective measures of sleep. *Journal of Consulting & Clinical Psychology, 63,* 658–667.

White, L., Hahn, P. M., & Mitler, M. M. (1980). Sleep questionnaire in adolescents. *Sleep Research, 9,* 108.

Williams, R. L., Karacan, I., & Hursch, C. J. (1974). *EEG of human sleep: Clinical applications*. New York: Wiley.

Zhdanova, I. V., Wurtman, R. J., Lynch, H. L., Morabito, C., & Matheson, J. K. (1995). Sleep inducing effects of low melatonin doses. *Sleep Research, 24,* 66.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.23
Headache

PAUL R. MARTIN
University of New England, Armidale, NSW, Australia

8.23.1 INTRODUCTION 530
8.23.2 DEFINITION, SYMPTOMS, AND CLASSIFICATION 531
8.23.3 FUNCTIONAL MODEL, RISK FACTORS, AND ETIOLOGY 532
  8.23.3.1 The Functional Model of Headaches 532
  8.23.3.2 Risk Factors for Developing a Headache Disorder 535
    8.23.3.2.1 Predisposing antecedents 535
    8.23.3.2.2 Onset antecedents 536
    8.23.3.2.3 Setting antecedents 536
    8.23.3.2.4 Summary of risk factors 537
  8.23.3.3 Etiology of Headaches 537
    8.23.3.3.1 Immediate antecedents 537
    8.23.3.3.2 Psychophysiological mechanisms 537
    8.23.3.3.3 Immediate reactions of sufferers and significant others 538
    8.23.3.3.4 Summary of etiological factors 539
8.23.4 IMPACT OF HEADACHES ON PERSON, FAMILY, AND SOCIETY 539
  8.23.4.1 Impact of Headaches on Individual and Family Functioning 539
  8.23.4.2 Impact of Headaches on Society 540
8.23.5 EPIDEMIOLOGY, LIFE SPAN, GENDER, AND CULTURAL ISSUES 541
8.23.6 MEDICAL DIAGNOSIS, TESTS, AND TREATMENT 542
  8.23.6.1 Medical Assessment 542
  8.23.6.2 Medical Management 542
8.23.7 PSYCHOLOGICAL ASSESSMENT AND TREATMENT 543
  8.23.7.1 Psychological Assessment 543
  8.23.7.2 Psychological Treatment 544
    8.23.7.2.1 Biofeedback training 544
    8.23.7.2.2 Relaxation training and cognitive coping training 546
    8.23.7.2.3 Group therapy 546
    8.23.7.2.4 Side effects of treatment 547
    8.23.7.2.5 Summary of treatment findings 547
8.23.8 MINIMAL INTERVENTIONS AND PREVENTION 547
  8.23.8.1 Minimal Interventions 548
  8.23.8.2 Prevention 548
8.23.9 ADHERENCE AND PROGNOSIS 549
  8.23.9.1 Adherence 549
  8.23.9.2 Prognosis 549
8.23.10 PROFESSIONAL ISSUES 550
8.23.11 REFERENCES 551

8.23.1 INTRODUCTION

> Headache roameth over the desert, blowing like the wind,
> Flashing like lightning, it is loosed above and below:
> It cutteth off like a reed him who feareth not his god . . .
> This man it hath struck and
> Like one with heart disease he staggereth,
> Like one bereft of reason he is broken.

Writings on headaches go back at least 6000 years. The above quote is from a Mesopotamian poem dating from 4000 to 3000 BC (Friedman, 1972). Descriptions of headaches have been found on the Ebers Papyrus which was transcribed around 1550 BC from earlier sources dating back as far as 3700 BC. It was not until Ancient Greece, however, that Hypocrates distinguished different types of headaches and described several of the well-known features of migraine. One of the earliest classification systems of headaches was developed by the Byzantine, Alexander Trallianus (AD 525–605), and subdivided headaches into three groups according to an ascending scale of severity—cephalalgia, cephalea, and hemicrania (Riley, 1932), which is an interesting historical footnote given that some modern writers have proposed severity models to replace taxonomies listing migraine and tension-type headache as distinct entities (Bakal & Kaganov, 1979).

The earliest description of headache trigger factors has been attributed to Cornelius Celsus (75 BC to AD 50) who listed drinking wine, "crudity" (i.e., dyspepsia), cold, heat of a fire, and the sun. Thomas Willis added other factors to this list in the seventeenth century including sexual activity, overeating, oversleeping, and hunger. Emotional factors did not appear in these lists, however, until Simone-André Tissot in the eighteenth century (Diamond & Maliszewski, 1992).

Early theories of headaches reflected the prevailing illness models of their time. In many cultures, headaches were viewed as a curse from the gods or a consequence of possession by evil spirits. The Ancient Egyptians, for example, wrote of Tiu, the evil spirit of headache (Diamond & Maliszewski, 1992). When humoral theories had currency in medicine, headaches were conceptualized as arising from an excess of body fluid or humor entering through the gastrointestinal system and irritating the brain. The vomiting of migraine was attributed to an excess of yellow bile (Rose, 1995).

Modern theories of headaches, such as the neurogenic and vascular theories, have their origins in the work of physicians from the seventeenth century onwards. Willis wrote a lecture in 1672 suggesting that the source of pain was not in the brain but rather the nerve fibers being pulled by the distended vessels (Cady & Everett, 1995). It was Harold Wolff and colleagues in the 1930s, however, who developed the vascular theory by arguing for an initial preheadache phase of intracranial vasoconstriction followed by extracranial vasodilatation occurring during the attack (Wolff, 1963). The vascular theory was challenged as early as 1873 by Edward Liveing who argued that migraine resulted from a brain seizure or "nervous storm." The suggestion that the origin of headaches was within the central nervous system was the basis for the neurogenic theory of migraine (Cady & Everett, 1995).

Theories of headaches have, of course, given rise to their own treatment modalities. Perhaps the earliest treatment for headaches was trepanning, in which demons were released from the head by cutting holes in the skull. Offerings to the gods, magic symbols, and incantations were part of the therapeutic armentarium. Remedies such as wearing a hangman's noose around the head, or moss scraped off the head of a statue on a red string tied around the neck, were offered (Cady & Everett, 1995). Humoral theories gave rise to treatments such as the letting of blood, application of leaches, and purging. Galen was an advocate of "counteriritation," suggesting the application of a live electric fish known as a torpedo to the forehead (Diamond & Maliszewski, 1992).

Pharmacological approaches to headaches have a very long history with their roots in the extraordinary concoctions suggested by early healers such as a ninth century English headache tonic consisting of elderberry juice, cow's brain, vinegar, and goat-dung (Cady & Everett, 1995). A significant step forward occurred toward the end of the nineteenth century with the introduction of ergot and ergot derivatives in the treatment of migraine, which constrict distended arteries. The history of psychological approaches is much shorter, although writers as early as Plato advocated comprehensive treatment of the whole person including their soul; and physicians such as William Gowers emphasized the importance of lifestyle modification in the late nineteenth century. Research on psychological treatment of headaches really began in the 1970s, however, stimulated by the evolving biofeedback literature (Martin, 1993). Following Neil Miller's ground-breaking work on learned control of autonomic functioning, researchers saw the potential of biofeedback for training headache sufferers to control physiological pain mechanisms.

8.23.2 DEFINITION, SYMPTOMS, AND CLASSIFICATION

Dictionaries usually define headaches as pain in the head and go on to define pain as suffering, anguish, or distress. More technical definitions refine the above by using phrases such as "pain located above the orbitomeatal line" (Headache Classification Committee of the International Headache Society, 1988). The pain of headache however, varies qualitatively, as it may, for example, be experienced as throbbing, pulsating, or a steady ache; and quantitatively, as indexed by measures of frequency, intensity, and duration. Headaches are often preceded, accompanied, or followed by other "symptoms" including visual auras, nausea, vomiting, photophobia, phonophobia, tinnitus, and vertigo.

A consistent assumption of the medical literature is that these pain characteristics and symptoms can be grouped into syndromes and that different syndromes have different physiological mechanisms. A further assumption is that the syndromes can be organized hierarchically into the categories and subcategories of a taxonomy. The previous section briefly mentioned some early attempts at classifying headache, but the system that dominated the field until recently was formulated by a committee of six eminent neurologists in 1962 (Ad Hoc Committee, 1962). This group divided headaches into 15 major categories of which the first two were "vascular headaches of the migraine type" and "muscle-contraction headache." These two types occurring on their own or in combination were believed to account for over 90% of all headaches.

The Ad Hoc Committee system merely provided brief descriptions of the various headache syndromes and was criticized on both conceptual and empirical grounds (e.g., Martin, 1985). In 1985, the International Headache Society (IHS) set up a committee to develop a new classification system which was published three years later (Headache Classification Committee of the IHS, 1988). The IHS system divides headaches into 13 major categories, the first four of which are considered as primary headache: "migraine," "tension-type headache," "cluster headache," and "miscellaneous headaches unassociated with structural lesions." The remaining nine categories are considered as secondary headaches and include headaches developing for the first time in close temporal relationship to a possibly causative organic factor. The categories "migraine" and "tension-type headache" parallel the first two categories of the Ad Hoc Committee's classification, and their subcategories are shown in Table 1. The subcategories of most relevance to psychologists are "migraine without aura," "migraine with typical aura," and all types of "tension-type headache."

In contrast to previous taxonomies, the IHS system provides operational diagnostic criteria for all types of headaches. The criteria are too extensive to list here but the descriptions of the most significant subcategories are provided in Table 2. The system classifies headaches rather than people so that an individual can be diagnosed as suffering from more than one type of headache.

The IHS system has gained wide acceptance among headache specialists and has been translated into more than 10 languages. It has been endorsed by all national headache societies represented by the IHS and the World Federation of Neurology. The major principles of the classification have been incorporated into the 10th revision of the *International classification of diseases* (World Health Organization, 1992). Specialist headache journals, such as *Headache* and *Cephalalgia*, expect contributors to use the system.

The Chairman of the IHS Classification Committee, Jes Olesen, has summarized the evidence accumulated in support of the system (Olesen & Lipton, 1994; Olesen & Rasmussen, 1995). He has argued that, first, the system has demonstrated applicability in diverse settings (i.e., headache clinics and the general population). Second, the system is exhaustive as the 13th category in the system ("headache not classifiable") is rarely used. Third, few studies have investigated the reliability of the system but those that have show substantial interobserver agreement. Validating the system is problematic as it is not clear what should constitute the critical comparison. One approach taken in some studies has been to compare the IHS system with the Ad Hoc Committee's system, and this has suggested a high correlation between the two classifications.

The positive acclaim for the IHS system has been accompanied by a series of criticisms, however, and suggestions for revisions. The most common criticism is that it is too complex for general practice and epidemiological research (e.g., Michel et al., 1993; Solomon, 1994). Parenthetically, it should be noted that the system is described in a volume supplement running to 96 pages in contrast to the Ad Hoc Committee's system which was described in two pages. Other authors have made a similar point in arguing that it is not widely used in everyday medical care (Gobel, 1994; Marcus, Nash, & Turk, 1994). Solomon (1994) has proposed simpler criteria for clinical practice and Michel et al. (1993) have proposed a variation of the IHS system for epidemiological research. Marcus

Table 1 First two categories of the classification system of the IHS.

- 1. Migraine
  - 1.1 Migraine without aura
  - 1.2 Migraine with aura
    - 1.2.1 Migraine with typical aura
    - 1.2.2 Migraine with prolonged aura
    - 1.2.3 Familial hemiplegic aura
    - 1.2.4 Basilar migraine
    - 1.2.5 Migraine aura without headache
    - 1.2.6 Migraine with acute onset aura
  - 1.3 Ophthalmoplegic migraine
  - 1.4 Retinal migraine
  - 1.5 Childhood periodic syndromes that may be precursors to or associated with migraine
    - 1.5.1 Benign paroxysmal vertigo of childhood
    - 1.5.2 Alternating hemiplegia of childhood
  - 1.6 Complications of migraine
    - 1.6.1 Status migrainous
    - 1.6.2 Migrainous infarction
  - 1.7 Migrainous disorder not fulfilling above criteria
- 2. Tension-type headache
  - 2.1 Episodic tension-type headache
    - 2.1.1 Episodic tension-type headache associated with disorder of pericranial muscles
    - 2.1.2 Episodic tension-type headache unassociated with disorder of pericranial muscles
  - 2.2 Chronic tension-type headache
    - 2.2.1 Chronic tension-type headache associated with disorder of pericranial muscles
    - 2.2.2 Chronic tension-type headache unassociated with disorder of pericranial muscles
  - 2.3 Headache of the tension type not fulfilling above criteria

et al. (1994) have emphasized the need for educating health care providers in the use of the system.

A second criticism that has been leveled at the IHS system by a number of authors is that the criteria are not well suited to diagnosing headaches in children (e.g., Gladstein, Holden, Peralta & Raven, 1993; Seshia & Wolstein, 1995). A third criticism evolved from a study by Michel et al. (1993) that led them to conclude the system had excellent specificity for diagnosing migraine but low sensitivity. The authors went on to argue that this results in the system being appropriate for research investigating the mechanisms or treatment of migraine but not for epidemiological research. A final criticism offered by some researchers is that the system does not adequately classify the daily and near daily headache disorders known as chronic daily headaches (Silberstein, 1994; Silberstein, Lipton, Solomon, & Mathew, 1994).

8.23.3 FUNCTIONAL MODEL, RISK FACTORS, AND ETIOLOGY

The committee that developed the IHS classification system recognized it would need periodic review and revision; and, in fact, encouraged researchers "to take an active part in the testing and further development of the classification" (Headache Classification Committee of the IHS, 1988, p. 19). Questions can be asked, however, about the validity of the assumptions on which classification systems are based, as outlined early in the previous section. Factor analytic studies, for example, have cast doubt on whether migraine and tension-type headache are discrete entities (e.g., Peck & Attfield, 1981; Ziegler, Hassanein, & Hassanein, 1972). Psychophysiological research has suggested that migraine and tension-type headache may have the same physiological mechanism (e.g., Lichstein et al., 1991; Martin, Marie, & Nathan, 1992). Finally, the two types of headaches do not respond differentially to the various psychological treatment modalities (e.g., Lichstein et al., 1991; Martin, Nathan, Milech, & van Keppel, 1989).

8.23.3.1 The Functional Model of Headaches

The above considerations have led to the advocating of exploring alternative ways of conceptualizing primary headaches. One alternative is to approach headaches from a

Table 2 Descriptions of the main categories and subcategories of headaches of relevance to psychologists.

Migraine without aura	Idiopathic, recurring headache disorder manifesting in attacks lasting 4–72 hours. Typical characteristics of headache are unilateral location, pulsating quality, moderate or severe intensity, aggravation by routine physical activity, and associated with nausea and photo- and phonophobia
Migraine with typical aura	Migraine with an aura consisting of homonymous visual disturbances, hemiparesis or dysphasia, or combinations thereof. Gradual development, duration under one hour and complete reversibility characterize the aura which is associated with headache
Episodic tension-type headache	Recurrent episodes of headaches lasting minutes to days. The pain is typically pressing/tightening in quality, of mild or moderate intensity, bilateral in location, and does not worsen with routine physical activity. Nausea is absent, but photo- or phonophobia may be present
Chronic tension-type headache	Headache present for at least 15 days a month during at least six months. The headache is usually pressing/tightening in quality, mild or moderate in severity, bilateral, and does not worsen with routine physical activity. Nausea, photo-, or phonophobia may occur
Headache of the tension-type not fulfilling above criteria	Headache that is believed to be a form of tension-type headache but does not quite meet the operational diagnostic criteria for any of the forms of tension-type headache

functional rather than a symptomatological perspective in which the emphasis is on understanding the controlling variables of headaches or the variance in headache activity. The key questions are the "why" questions: why does the person suffer from a headache one day rather than another; why does the person suffer from headaches more at this stage of their life than at other stages; why did the person develop a headache problem when they did; and why is the person predisposed to developing a headache disorder? Such questions seem more closely related to etiology than pain characteristics and associated symptoms, and appear to have more management implications, at least from a psychological perspective.

A functional model of chronic headaches is represented in Figure 1 (Martin, 1993). In the center of the model are the headaches and associated symptoms, and the underlying physiological mechanisms. The controlling variables are represented as antecedents and consequences. The antecedents are arranged along a temporal continuum grouped into four categories. The immediate antecedents are the stimuli that precipitate or aggravate headaches, and commonly reported triggers include: stress and negative emotions; perceptual phenomena such as flicker, glare, and loud noise; certain foods, alcohol, and hunger; and high temperature and humidity. Setting antecedents are the lifestyle factors that moderate current vulnerability and provide the psychosocial context in which headaches occur. If headaches are triggered by stress, for example, a full understanding of the problem requires investigation of the sources of stress and stress-moderating variables such as appraisal processes, coping skills, and social support. A dysfunctional marriage serving as a stressor could be viewed as a setting factor for stress-induced headaches as could a deficient social support system.

Onset antecedents are the processes or events that result in headaches developing initially or becoming significantly worse. Stressful life events and head or neck trauma are commonly reported events that sufferers associate with the start of their headache disorder. Predisposing antecedents are the constitutional and personality characteristics that account for individual differences in susceptibility. Headaches run in families and genetic factors are considered to play a role.

Consequences are included in the model because the reactions of headache sufferers and their significant others often exacerbate the headaches by setting up negative feedback loops with the antecedent factors; and similarly, the long-term consequences of headaches can contribute to a psychosocial context in which headaches thrive. Headaches and their associated symptoms often lead to sufferers feeling

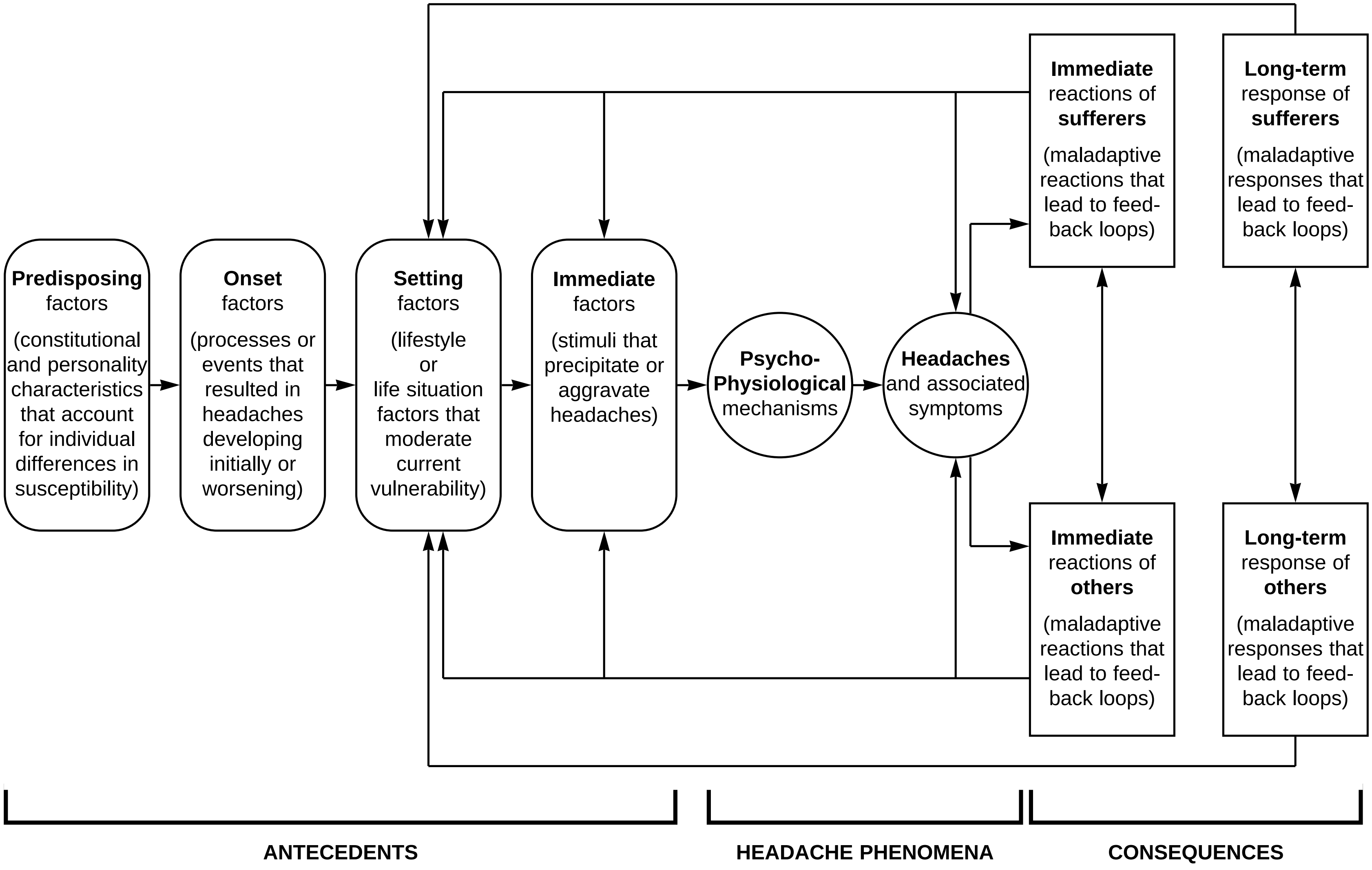
Predisposing factors
(constitutional and personality characteristics that account for individual differences in susceptibility)
Onset factors
(processes or events that resulted in headaches developing initially or worsening)
Setting factors
(lifestyle or life situation factors that moderate current vulnerability)
Immediate factors
(stimuli that precipitate or aggravate headaches)
Psycho-Physiological mechanisms
Headaches and associated symptoms
Immediate reactions of **sufferers**
(maladaptive reactions that lead to feed-back loops)
Long-term response of **sufferers**
(maladaptive responses that lead to feed-back loops)
Immediate reactions of **others**
(maladaptive reactions that lead to feed-back loops)
Long-term response of **others**
(maladaptive responses that lead to feed-back loops)
ANTECEDENTS
HEADACHE PHENOMENA
CONSEQUENCES

Figure 1 The functional model of chronic headaches (after Martin, 1993).

stressed, anxious, and depressed, for example, thus creating a vicious cycle whereby the reactions further exacerbate the problem. The tension resulting from headaches can lead to rows between sufferers and their partners, thereby causing further deterioration in the marital relationship which may be a setting factor for the headaches. Alternatively, partners may respond to headaches by providing attention, support, and sympathy, so that reinforcement may play a role in the maintenance of headaches. After headaches have adversely affected social and recreational activities on a number of occasions, sufferers may withdraw from such activities. Hence, a long-term consequence of headaches may be a reduction in the sufferer's social network, thereby resulting in reduced social support, a setting antecedent for stress-related headaches.

The functional model has been developed based on work with adults but has also been applied to children and adolescents (Holden, Gladstein, Trulsen, & Wall, 1994). The book contains examples of patients analyzed according to a functional perspective, and an intervention approach driven by the model (Martin, 1993).

The literature relevant to this model will now be reviewed briefly. Predisposing, onset, and setting antecedents will be discussed under the general heading of "risk factors for developing a headache disorder," and immediate antecedents, physiological mechanisms, and immediate reactions will be discussed under the general heading of "etiology of headaches." Long-term consequences will be reviewed in the following section on impact of headaches. Because of the feedback loops that characterize chronic headaches, it should be emphasized that the appropriate subsection for discussing research findings is often not clear. Inadequate social support can be viewed as both an antecedent and consequence of headaches, for example.

8.23.3.2 Risk Factors for Developing a Headache Disorder

8.23.3.2.1 Predisposing antecedents

Epidemiological studies consistently report that headaches are more prevalent in females than males (e.g., Nikiforow & Hokkanen, 1978). In one study, the gender ratio for migraine was estimated at 2.5:1, and for tension-type headache at 1.4:1 (Rasmussen, Jensen, Schroll, & Olesen, 1992b).

Migraine has long been accepted as a famial condition, and in fact many neurologists consider a family history to be a prerequisite for the diagnosis of migraine (Ziegler, 1978). Early studies reported concordance rates in excess of 90% (e.g., Allan, 1930) but later studies reported rates as low as 10% (e.g., Waters, 1971). The early findings led a number of authors to suggest that migraine was an inherited disorder. Twin studies have supported a genetic factor, although reported concordance rates have varied from 14% to 100% for monozygotic twins and 0% to 40% for dizygotic twins (Lucas, 1977). The mode of transmission has been a controversial topic with autosomal dominant, autosomal recessive, and multifactorial modes proposed. More recent and better controlled studies support the multifactorial mode (e.g., Kalfakis, Panas, Vassilopoulos, & Malliara-Loulakaki, 1996), and D'Amico, Leone, Macciardi, Valentini, and Bussone (1991) have suggested that it is "sex-limited."

One approach to inherited predisposition is to consider whether individuals with particular personality traits or profiles are more vulnerable to headaches, since personality is partially genetically determined. For many years it has been argued that headache sufferers tend to have a particular personality profile. Henryk-Gutt and Rees (1973), for example, characterized a migraine sufferer as "a tense, driving, obsessional perfectionist with an inflexible personality who maintains a store of bottled-up resentments which can neither be expressed nor resolved" (p. 142). Fewer descriptions of the personality traits associated with tension headache have appeared in the literature, but those that have share much in common with the migraine profile.

The literature on personality and headaches is extensive as indicated by a review which tracked down over one hundred studies (Blanchard, Andrasik, & Arena, 1984), but the findings are inconsistent. Some studies report positive findings such as increased neuroticism in migraineurs (Breslau & Andreski, 1995), greater rigidity in tension-type headache sufferers (Passchier, van der Helm-Hylkema, & Orlebeke, 1984a), and higher Type-A scores in headache sufferers (Martin, Nathan, & Milech, 1987); but other studies have found few abnormalities (e.g., Andrasik, Blanchard, Arena, Teders, & Rodichok, 1982; Kohler & Kosanic, 1992).

Methodological problems abound in this literature. Philips (1976) has pointed out, for example, that individuals who complain to their doctors have a particular type of personality characterized by extroversion and neuroticism, so that findings from clinical samples cannot be generalized to the headache population as a whole. Another problem relates to measurement artifact as elevated scores on some personality tests may result from measuring headache symptomatology rather than psychopathology (Dieter & Swerdlow, 1988).

Cluster analytic studies have identified four or five homogeneous subgroups of headache sufferers, some of which are characterized by quite abnormal profiles, such as conversion V (clinically significant elevations on hypochondriasis and hysteria but not depression), and some of which are relatively normal (Rappaport, McAnulty, Waggoner, & Brantley, 1987; Robinson, Geisser, Dieter, & Swerdlow, 1991). This line of research is more consistent with clinical experience which suggests that personality characteristics sometimes play a predisposing role in headaches and sometimes do not.

8.23.3.2.2 Onset antecedents

Events and processes that can initiate a headache disorder have received surprisingly little attention in the research literature. Some early studies have reported that the onset of headache problems is associated with stress on approximately 50% of occasions (Henryk-Gutt & Rees, 1973; Howarth, 1965). More recent studies have linked headache onset with particular types of stressful events such as physical and sexual abuse (Domino & Haber, 1987), personal loss (Kaiser & Primavera, 1993), and onset of low back pain (Duckro, Schultz, & Chibnall, 1994).

One study has associated headache onset with minor head or neck injury (Weiss, Stern, & Goldberg, 1991). The only other events linked with headache onset are hormonally mediated. Menarche is considered by some to be associated with onset or exacerbation of headaches but others have presented data that they argue do not support this contention (cf. Deubner, 1977; Kudrow, 1978). Oral contraceptives can increase the frequency of migraine headaches (Granella et al., 1993), and pregnancy can be associated with the onset of a headache disorder although, more commonly, pregnancy results in a decrease in headaches (Uknis & Silberstein, 1991).

8.23.3.2.3 Setting antecedents

Given that stress is commonly reported to be a precipitator of headaches, frequent exposure to stressors on the one hand, or inadequate personal or social resources for coping with stressors on the other hand, should result in vulnerability to headaches. Exposure to stressors has been investigated both at the levels of major and minor life events. Comparing headache sufferers and controls in terms of major life events has usually resulted in a finding of no significant difference (e.g., Martin & Theunissen, 1993), but comparisons in terms of minor life events or daily hassles have yielded significant differences (e.g., Spierings, Sorbi, Haimowitz, & Tellegen, 1996).

Holm, Holroyd, Hursey, and Penzien (1986) also found that headache sufferers appraised the stressful events they experienced more negatively than did controls. Specifically, headache sufferers perceived the event as having greater impact and themselves as having less control over the event than did controls. In addition, headache sufferers employed less effective coping strategics in their efforts to manage stressful events since they placed more reliance on the relatively ineffective coping strategies of avoidance and self-blame and made less use of social support than did controls. A follow-up study confirmed that sufferers of both migraine and tension-type headache appraised and coped with stressful events differently from headache-free controls (Ehde & Holm, 1992).

Two studies comparing headache sufferers and matched nonheadache controls have found headaches to be associated with deficits in perceived social support. Martin and Theunissen (1993) reported differences between headache and control groups in terms of availability and adequacy of both attachment and social integration. Martin and Soon (1993) found headache sufferers were significantly less satisfied with the support available to them and scored lower on all four types of functional support measured in the study.

If anxiety and depression can precipitate or aggravate headaches then emotional and mood disorders would provide a psychosocial context for headaches. A number of studies have documented associations between headaches, depression and anxiety. Featherstone (1985) found that depression or anxiety were significantly more commonly diagnosed in headache sufferers than in matched controls. Morrison and Price (1989) reported that 33% of referrals to a migraine clinic met the *Diagnostic and Statistical Manual of Mental Disorders*, third edition, criteria for affective disorder. A recent study by Marazziti et al. (1995) has gone further to argue that migraine, panic disorder, and some forms of depressive illness are part of the same spectrum.

Headaches have been linked to other disorders such as sleep disturbance (e.g., Paiva et al., 1994), eating disorders (e.g., Brewerton & George, 1993), and chronic low back pain (e.g., Duckro et al., 1994). Given that these disorders are likely to be associated with stress and negative emotions, they too could contribute to the psychosocial context of headaches.

8.23.3.2.4 Summary of risk factors

Before listing risk factors for headache disorders, two points need emphasizing. The evidence supporting the risk factors is tenuous as it is based on a small number of studies that have not always produced consistent findings, and it is open to alternative interpretations such as some "risk factors" being a consequence rather than a cause of headaches (e.g., perceptions of deficient social support). Also, the relationships that have been demonstrated are of a low order of magnitude compared with some risk factors for other disorders (e.g., smoking and lung cancer).

Having stated these limitations, the following summary is offered:

(i) Female gender and family history of headaches are risk factors for headaches, particular migraine.

(ii) Personality profiles such as the "neurotic triad" or "conversion V" on the Minesota Multiphasic Personality Inventory, and Type A personality, arguably, constitute risk factors for headaches.

(iii) High exposure to stressors, dysfunctional appraisal processes, deficient coping skills, and perceived inadequate social support are risk factors for developing a headache disorder.

(iv) Mood and emotional disorders, such as depression and anxiety, are headache risk factors.

8.23.3.3 Etiology of Headaches

8.23.3.3.1 Immediate antecedents

It has been suggested that "everything can produce a migraine" (Blau & Thavapalan, 1988, p. 433) and certainly the list of trigger factors reported in the literature is extensive. One study aimed to simplify the list via factor analysis (Martin, Milech, & Nathan, 1993). This approach resulted in 25 commonly reported precipitants being reduced to five readily interpretable patterns: negative affect (stress, anxiety, depression, and anger); visual disturbance (flicker, glare, and eyestrain); somatic disturbance (sneezing, coughing and pollen); environmental stress (humidity, high temperature, and opposite to relaxation); and consumatory stimuli (alcohol, food, and hunger).

Although the list of trigger factors is long, the empirical support for most of the factors is limited. A series of studies has been completed in which the main antecedent factors were manipulated in the laboratory to validate them experimentally as headache triggers. Negative affect, induced by requiring subjects to attempt to solve difficult anagrams accompanied by failure feedback, has been used in all three of the studies cited below and has consistently resulted in reported head pain in comparison to control procedures. Other antecedent factors that have been shown to induce headaches include: visual disturbance (subjects have to read small font words from a computer screen while the screen is illuminated by a bright stroboscopic light) (Martin & Teoh, 1997a); noise (exposure to 65 decibels of white noise) (Martin & Todd, 1996); and hunger (food deprivation for 19 hours) (Martin & Seneviratne, 1997).

The headache-inducing properties of stress have also been demonstrated in the laboratory in two studies (Gannon, Haynes, Cuevas, & Chavez, 1987; Haynes, Gannon, Bank, Shelton, & Goodwin, 1990). The only other trigger factor to be manipulated has been food. Moffett, Swash, and Scott (1974) investigated the effect of chocolate in migraine by carrying out two separate double-blind studies in migraineurs who had observed that headaches regularly occurred after the ingestion of small amounts of cocoa products. The results suggested that chocolate on its own is rarely a precipitant of migraine. Using a quite different methodology, Medina and Diamond (1978) randomly assigned migraineurs to three dietary regimes involving eating foods reported to be associated with headaches, eating foods not reported to be associated with headaches, and a normal diet control condition. Monitoring headaches over six week periods on each dietary regime revealed no differences in indices of migraine or headache medication.

8.23.3.3.2 Psychophysiological mechanisms

The peripheral physiological mechanism of tension-type headache was considered to be sustained contraction of skeletal muscles around the head, although the IHS classification system divides this group into tension-type headaches caused by pericranial muscle tension, and tension-type headaches caused by unknown factors. In contrast, the peripheral physiological mechanism of migraine was considered to be distended extracranial and intracranial arteries. The paradigm that has been used most extensively to investigate these mechanisms has been to compare headache sufferers and headache-free controls in terms of a number of physiological variables, under various experimental conditions such as rest/relaxation, physical or mental stress, and poststress recovery. The different conditions enable testing a variety of hypotheses relating to the physiological mechanisms of headaches such as whether headache sufferers differ from controls at rest, in response to stress (e.g., exaggerated response), or in terms of poststress recovery (e.g., slower recovery).

A large literature developed during the late 1970s and 1980s using this approach but a clear picture has not emerged. Considering first the hypothesized muscular mechanism of tension-type headache, the evidence is mixed, but on balance quite unconvincing. There is little support for the hypothesis that tension-type headache sufferers have generally elevated electromyographic (EMG) levels or are slow to recover following stress (e.g., Andrasik, Blanchard, Arena, Saunders, & Barron, 1982; Martin & Mathews, 1978). The evidence for headache sufferers responding to stress with more muscle tension than nonheadache controls is more promising (e.g., Cohen et al., 1983) but many negative findings have been reported (e.g., Martin & Mathews, 1978). In reviewing this literature, Flor and Turk (1989) reported that among studies they rated as more methodologically sound, compared with studies that were methodologically weaker, a lower proportion supported the muscular hypothesis.

Investigations of the vascular mechanism of migraine have also produced inconsistent findings but seem more promising. Comparisons of migraineurs and control subjects at rest suggest some temporal blood flow abnormalities such as increased variability and perhaps elevated cardiovascular arousal as indexed by heart rate and possibly other variables such as diastolic blood pressure and facial skin temperature (e.g., Drummond, 1982; Morley 1985). It is in the area of response to stress that the most impressive findings have occurred, however. Migraineurs show response stereotypy by responding maximally in terms of the cardiovascular system to a variety of stressors (e.g., Cohen, Rickles, & McArthur, 1978). In addition, relative to controls, migraineurs respond to stress with vasodilatation of the temporal artery, possibly preceded by vasoconstriction (e.g., Rojahn & Gerhards, 1986). There is also some tentative evidence that migraineurs are slower to recover from stress (e.g., Gannon, Haynes, Safranek, & Hamilton, 1981).

One other paradigm that has been used for investigating physiological mechanisms is to compare headache sufferers when experiencing a headache and when headache-free. Several studies have investigated muscular mechanisms with this approach and produced directly conflicting results. For example, Haynes, Griffin, Mooney, and Parise (1975) found significantly higher forehead EMG during headache periods while Martin and Mathews (1978) reported the reverse. Vascular mechanisms have received less attention but Passchier, van der Helm-Hylkema, and Orlebeke (1984b) reported that migraineurs experiencing a headache had significantly elevated heart rates relative to headache-free migraineurs.

Finally, it should be noted that vascular abnormalities have been found in tension-type headache sufferers (e.g., Lehrer & Murphy, 1991), and muscular abnormalities in migraineurs (e.g., Philips, 1977). Where migraineurs and tension-type headache sufferers have been included in the same study, it has often resulted in no significant differences emerging between them (e.g., Martin et al., 1992).

8.23.3.3.3 Immediate reactions of sufferers and significant others

There are a variety of reactions to headaches that are maladaptive in the sense that they are likely to aggravate the disorder. It is known, for example, that pain perception is affected by factors such as anxiety level, focus of attention, and sense of control (Melzack & Wall, 1988). Hence, reactions that are characterized by elevated anxiety, focusing on the headaches, and feelings of loss of control will enhance pain perception. Factor and cluster analytic studies of cognitive and emotional reactions to headaches demonstrate exactly these responses, as common responses include tension, worry, and concern; rumination about the headaches and their causes; and helplessness (e.g., Penzien, Holroyd, Holm & Hursey, 1985; Philips, 1989).

In similar vein, as stress and negative affect are the most common precipitators of headaches, any reactions to headaches that are characterized by negative affect will lead to a vicious cycle developing. The findings discussed above provide support for this feedback loop, and additional findings from these studies include reactions such as irritation over environmental difficulties and emotional reactivity (Penzien et al., 1985; Philips, 1989).

Philips (1987) has argued that avoidance behavior plays a role in maintaining chronic pain via increased sensitivity to pain-inducing stimuli in the short term and through adversely affecting self-efficacy beliefs pertaining to pain in the long term. This model is based on two types of evidence, one of which is the data showing that a characteristic response to pain is avoidance behavior (e.g., Philips & Hunter, 1981). The other type of evidence comes from a laboratory study in which it was shown that exposure to a pain-provoking stimulus (noise) led to increased tolerance of the stimulus, whereas avoidance of exposure to the stimulus resulted in decreased tolerance (Philips & Jahanshahi, 1985). One study also found that repeated exposure to a headache trigger (flicker, glare, and eyestrain) resulted in desensitization to the trigger (Martin & Teoh, 1997b).

A common response to a headache developing, or even to concern that a headache will develop, is to take medication, sometimes in large quantities. Excessive consumption of medication can exacerbate the headache disorder, however (Rapoport, 1988; Saper, 1987). The IHS classification includes a category "headache associated with substances or their withdrawal" and the subcategories "ergotamine induced headache" and "analgesic abuse headache." Baumgartner, Wessely, Bingol, Maly, and Holzner (1989) claimed that drug-induced headaches are common, accounting for 10%–15% of patients in headache clinics, and some individuals consume extraordinary quantities of medication. A case has been described, for example, of a woman who was taking on average 21 tablets a day and met the criteria for both IHS subcategories (Forsyth & Martin, 1997). Hospitalization for detoxification resulted in spectacularly beneficial results.

Finally, a number of authors have suggested that reinforcement or secondary gain play an important role in maintaining headaches in a proportion of cases (e.g., Martin, 1972; Stout, 1984). Unfortunately, this type of claim is extremely difficult to test empirically. Research on how sufferers and significant others react to headaches has indicated that the most common patterns of responses are positive (e.g., stop what you are doing, sit or lie down, take medication) rather than negative (e.g., ignoring, complaining) (Martin, Milech, & Nathan, 1993). Also, cases of headaches in children and adolescents treated successfully using a contingency management approach are consistent with consequences maintaining the disorder (Ramsden, Friedman & Williamson, 1983; Yen & McIntire, 1971).

8.23.3.3.4 Summary of etiological factors

The etiology of migraine and tension-type headache is not clear but the following factors seem to be involved.

(i) Some form of vascular mechanism probably plays a role in both migraine and tension-type headache and one feature of this mechanism is distended cranial arteries.

(ii) This mechanism can be triggered by a number of factors. The most commonly reported trigger and the one for which there is most supporting evidence is stress and negative affect. Additional trigger factors that have established experimental validity include visual disturbance (flicker, glare, and eyestrain), noise, and hunger. Other factors can probably trigger headaches, such as alcohol, certain foods, and particular weather conditions, but the supporting evidence is largely limited to retrospective self-report.

(iii) Maladaptive reactions to headaches by the sufferers and significant others can play a role in maintaining headaches via a number of different processes such as creating negative feedback loops and reinforcement.

8.23.4 IMPACT OF HEADACHES ON PERSON, FAMILY, AND SOCIETY

The impact of headaches on individuals, families, and the community has received far more attention in the 1990s than previous decades. A recent supplement to *Neurology* (1994, *44*, Supplement 4) was specifically devoted to the impact of migraine. The literature seems driven by two main trends. First, the shift in emphasis in medicine from a focus on reducing mortality to increasing quality of life has led to a series of studies looking at how headaches affect quality of life. Second, concerns about increasing health care costs and the need for developing a cost-effective approach to health has led to researchers looking at the costs of providing services to headache sufferers on the one hand (i.e., the direct costs) vs. the economic costs of headaches in terms of absenteeism and reduced work efficiency on the other (i.e., the indirect costs).

8.23.4.1 Impact of Headaches on Individual and Family Functioning

As discussed earlier, headaches are often accompanied by a number of symptoms. The majority of migraine sufferers report that pain is the most important symptom however, and a review by Stewart and Lipton (1994) found that across several studies, 57–85% of migraineurs rated their pain intensity as severe or very severe. Daily activities are severely disrupted at the time of a migraine attack as indicated by the number of migraineurs who endorsed the following statements (percentages agreeing with statements are included in parentheses): "when I have one of my migraines/severe headaches, I postpone household chores" (90%); "I always have to lie down when I have one of my migraines/severe headaches" (76%); "my migraines/severe headaches limit my ability to work as carefully as I usually do" (73%) (Clarke, MacMillan, Sondhi, & Wells, 1996). This study also showed that migraine had an impact between attacks as indicated by the following endorsed statements: "I'm afraid of letting others down because of my migraines/severe headaches" (66%); "my migraines/severe headaches interfere with how I get on with members of my family and friends" (54%).

A number of studies have administered quality of life scales to headache sufferers and various comparison groups. Osterhaus, Townsend, Gandek, and Ware (1994), for example, administered to migraineurs in the USA the Medical Outcomes Study (MOS) 36-item Short-Form Health Survey (SF-36) which provides measures on eight scales: physical functioning, role disability due to physical problem, bodily pain, general health, vitality, social functioning, role disability due to emotional problems, and mental health. Scores were adjusted for comorbid conditions. Migraineurs' adjusted scores were substantially and significantly lower on all eight scales than those for a general population sample of comparable age and gender who reported no chronic disease. Compared with other chronic conditions which had been adjusted for age, gender, and comorbidity, migraine appeared comparable to conditions that might otherwise be considered to be more debilitating than migraine. While similar to osteoarthritis and diabetes in physical functioning, scales measuring well-being were significantly lower for migraineurs. Migraine was significantly lower than hypertension on all measures except general health perceptions. Measures of bodily pain, role disability due to physical health and social functioning, discriminated best between migraineurs, disease-free individuals and patients with other chronic conditions.

Solomon, Skobieranda, and Gragg (1993) also administered the MOS Short-Form Health Survey to headache sufferers and concluded that patients with chronic headaches have a level of functioning worse than that of patients with such chronic diseases as arthritis, diabetes, depression, and back problems. The only chronic conditions that had similar levels of functional impairment to chronic headaches were myocardial infarction and congestive heart failure. The only disorder with worse levels of patient well-being and functioning was symptomatic HIV infection.

The SF-36 has been administered in other countries with similar results (e.g., Essink-Bot, Van Royen, Krabbe, Bonsel, & Rutten, 1995). Also, the impact of headaches on lifestyle, health status, and quality of life has been demonstrated on a number of other instruments such as the Comprehensive Pain Questionnaire (Lacroix & Barbaree, 1990), Nottingham Health Profile (Jenkinson, 1990), and Sickness Impact Profile (Carlsson, Augustinsson, Blomstrand, & Sullivan, 1990). A recent development in this literature is the construction of quality of life questionnaires specific to headache such as the Migraine-specific Quality of Life Questionnaire (Hartmaier, Santanello, Epstein, & Silberstein, 1995) and the Quality of Life Headache in Youth (Langeveld, Koot, Loonen, Hazebroek-Kampschreur, & Passchier, 1996). Impairment in psychosocial functioning has been demonstrated in children suffering from recurrent headaches as well as adults (Carlsson, Larsson, & Mark, 1996).

8.23.4.2 Impact of Headaches on Society

In a review of patterns of health care utilization, Lipton, Stewart, and von Korff (1995) reported that the percentages of migraineurs who had ever consulted a doctor ranged from 56% in Denmark to 66% in the USA and 81% in Canada. Although these proportions are high in every country studied, the majority of migraineurs were not being seen by doctors currently and most migraine sufferers reported that they had never been diagnosed by a doctor. Less information is available on tension-type headache but Rasmussen, Jensen, and Olesen (1992) reported that only 24% of tension-type headache sufferers had ever consulted their general practitioner. Hospitalization for headache is relatively rare. Rasmussen et al. (1992) reported that 2% of all headache sufferers were hospitalized in Denmark, and Stang and Osterhaus (1993) reported that 8% of migraineurs were hospitalized in the USA.

The indirect costs of headaches are enormous. Although the majority of headache sufferers claim that they do not take time off work because of their headaches (Clarke et al., 1996), the number of working days lost is still very high. Stang and Osterhaus (1993) estimated that migraineurs in the USA were bedridden for about three million days per month and had an estimated 74.2 million days per year of restricted activity due to migraine. Estimates of the cost of migraine in terms of lost productivity in the USA have ranged from $1.4 to $17 billion per year (Stewart & Lipton, 1994), and have been valued at greater than £600 million per year in the UK (Clarke et al., 1996).

Estimates of the direct costs of medical care for headaches are harder to find. Clarke et al. (1996) quoted a figure of £20–30 million per year as the cost of migraine to the National Health Service in the UK. This figure seems particularly low in the context of a study by Osterhaus, Gutterman, and Plachetica (1992) which calculated that the 648 migraineurs who responded to their survey in the USA incurred $529,199 in annual medical expenditure because of their condition. Of course, costs are likely to be lower than one would expect on the basis of the high prevalence rate of headache in the community (discussed in Section 8.23.5) since migraine is substantially underdiagnosed and undertreated. Certainly, reviewers typically conclude that the indirect costs of headaches far exceed the direct costs (e.g., Lipton et al., 1995) which is often used as an argument for the need for better management. Both types of cost are difficult to estimate but a number of authors have recently suggested improved methods of analysis (e.g., de Lissovoy & Lazarus, 1994).

8.23.5 EPIDEMIOLOGY, LIFE SPAN, GENDER, AND CULTURAL ISSUES

Epidemiological studies have produced wildly varying figures for headache prevalence. This is illustrated by a review that selected 24 out of 58 identified population-based studies and reported prevalence rates for migraine ranging from 1.7% to 46.3% in males and from 4.4% to 57.1% in females (Stewart, Simon, Shechter, & Lipton, 1995). This variability is usually attributed to unreliable or inaccurate data and the methodological problems inherent in population-based studies.

Studies using the IHS criteria have achieved greater consistency. Rasmussen, Jensen, Schroll, and Olesen (1991), investigating a Danish sample, reported lifetime prevalences for men of 93% for any kind of headache, 8% for migraine, and 69% for tension-type headache. For women, the lifetime prevalences were 99% for all headache, 25% for migraine, and 88% for tension-type headache. In the preceding year, migraine prevalences were 6% in men and 15% in women; the corresponding tension-type headache prevalences were 63% and 86%, respectively. The American Migraine Study reported very similar results for migraine prevalence (Stewart, Lipton, Celentano, & Reed, 1992) as did an epidemiological study conducted in France (Henry et al., 1992).

Epidemiological research has identified a number of factors that are related to prevalence. Gender is one such factor as revealed in the above figures, with both migraine and tension-type headache more common in females. Age is another factor. Headaches are common in childhood (Kristjansdottir & Wahlberg, 1993) and migraine prevalence is about equal in boys and girls prior to puberty (Goldstein & Chen, 1982). Prevalence increases more rapidly in girls than in boys as adolescence approaches. Prevalence increases in both sexes from 12 to 38 years of age before falling steadily to a low rate in later life (Stewart, Linet, Celentano, Van Natta, & Ziegler, 1993).

Physician and clinic-based studies suggest that migraine is associated with high intelligence and social class, but this is not supported by population-based studies. In the American Migraine Study, migraine prevalence was found to be lower in higher income groups (Lipton & Stewart, 1993), a finding replicated in the National Health Interview Study (Stang & Osterhaus, 1993). In the latter study, prevalence was lowest for middle-income groups, and began to rise in the high-income group. This study relied on self-reports of migraine, however, and that migraine awareness rises with income.

It is often suggested that migraine prevalence varies as a function of culture. Haimanot, Seraw, Forsgren, Ekbom, and Ekstedt (1995) reviewed population-based studies from nine countries and calculated one year migraine prevalence rates ranging from 0.7% in China to 29.1% in Thailand. It is not possible to say, however, the degree to which these figures represent differences in prevalence rates among countries or the problems of population-based studies referred to earlier. In the Stewart et al. (1995) review, for example, nine population-based studies were from a single country (the UK) and yet the reported rates varied from 17.0% to 46.3% in males, and from 23.1% to 57.1% in females.

Focusing on more recent studies using IHS criteria does suggest that while at least some countries in Europe and North America have similar rates, as discussed earlier, other countries differ. A study of Greek monks living in an all-male community, for example, reported prevalence rates for migraine of 3.6% and for tension-type headache of 3.3% (Mitsikostas, Thomas, Gatzonis, Ilias, & Papageorgiou, 1994). An Ethiopian study (Haimanot et al., 1995) reported prevalence rates for men of 1.7% for migraine and 1.0% for tension-type headache, and prevalence rates for women of 4.2% for migraine and 2.3% for tension-type headache. These figures are dramatically lower than those reported in other studies using IHS criteria, particularly the figures for tension-type headache. The implication of these findings is that the prevalence of migraine and tension-type headache is higher in developed, Westernized countries than in the developing nations and simpler societies, but given some of the earlier findings, such as the high prevalence of migraine in Thailand, such conclusions await further research.

One study suggested that migraine prevalence was increasing in the USA. The Center for Disease Control reported that migraine prevalence increased by 60% between 1981 and 1989 (Morbidity and Mortality Weekly Report, 1991). As the study relied on self-report, such a finding could reflect greater awareness of the disorder.

With respect to the contribution of epidemiological studies to our understanding of functional characteristics of headaches, there appears to be much overlap in trigger factors across individuals grouped in diverse ways. Migraine and tension-type headache have similar triggers, for example, although some differences exist such as a slightly greater vulnerability to visual stimuli (flicker, glare, eyestrain) in migraineurs (Martin, Milech, & Nathan, 1993). Trigger factors are much the same for males and females except, of course, for hormonally related precipitants (Rasmussen, 1993). Headache triggers reported in studies of childhood headache (e.g., Osterhaus & Passchier, 1992, Spierings & Schellekens, 1991) are similar to those reported in adult studies assuming that one translates predominantly childhood phenomena such as "doing homework" and "preparing for exams" into adult equivalents. Epidemiological studies carried out in different countries consistently report stress as a major factor but other factors appear more culturally bound. For example, alcohol was not listed as a precipitant of headaches in Saudi Arabia (Al-Rajeh, Bademosi, Ismaii, & Awada, 1990), witchcraft seemed a uniquely African precipitant (Levy, 1983), and sauna a rather Scandinavian phenomenon (Nikiforow & Hokkanen, 1978).

8.23.6 MEDICAL DIAGNOSIS, TESTS, AND TREATMENT

8.23.6.1 Medical Assessment

Headaches are one of the most common symptoms that neurologists evaluate and although most are caused by primary headache disorders, the differential diagnosis is one of the longest in medicine with over 300 different types and causes (Saper, Silberstein, Gordon, & Hamel, 1993). The causes or types of most headaches can be determined by taking a detailed history supplemented by a general and neurological examination. Experts differ in their opinion with respect to the value of diagnostic testing. Saper et al. (1993) state that "In general, we believe that most patients with headache require neurodiagnostic testing. A computed tomographic (CT) scan or magnetic resonance imaging (MRI) should be performed in most patients with headaches severe enough to prompt medical evaluation and treatment" (p. 8). In contrast, Lance (1993) argues that "only a small proportion of headache patients require investigation, other than a careful history and physical examination" (p. 268).

In a review of diagnostic testing, Evans (1996) listed clinical situations in which neurologists usually recommend diagnostic testing which include: the "first or worst" headache; headache with increasing frequency or severity; headache associated with symptoms and signs such as fever, stiff neck, nausea, and vomiting; new-onset headache after age 50; and headache associated with cognitive impairment. Neuroimaging is most commonly used to rule out brain tumors. MRI is generally more sensitive than CT scan but the yield for both types of test is low. Reviews show detection rates for tumors of less than one in a hundred scans (e.g., Frishberg, 1994).

The electroencephalogram (EEG) for many years was a standard test for the evaluation of headaches, especially in the pre-CT scan era. Experts now argue that EEG is not useful in the routine evaluation of headache (Evans, 1996). A lumbar puncture is recommended in certain circumstances such as where a neurologist is suspicious that a new-onset headache is caused by a subarachnoid hemorrhage. Similarly, cerebral arteriography and magnetic resonance angiography are recommended in specific circumstances associated with subarachnoid hemorrhage. For headaches following minor head injury, skull radiographs are frequently carried out but are probably not justified on a routine basis.

8.23.6.2 Medical Management

The medical management of primary headache will often incorporate some counseling but the mainstay is pharmacotherapy. Two approaches are differentiated, acute treatment aimed at aborting attacks and prevention or prophylaxis, although there is some overlap in the medications prescribed (Baumel, 1994; Leone, Grazzi, D'Amico, Moschiano, & Bussone, 1995; Schulman, & Silberstein, 1992). Four main types of drug are recommended for the acute treatment of migraine: analgesics, nonsteroidal anti-inflammatory drugs (NSAIDs), ergot derivatives, and sumatriptan. One of the most commonly used drugs for the treatment of acute migraine is aspirin which has both analgesic and anti-inflammatory properties. Migraine attacks can also be managed successfully using paracetamol or codeine. Nausea or vomiting is common during migraine attacks so an antiemetic may also be administered.

NSAIDs are probably the preferred treatment for mild to moderately severe attacks of migraine. As for aspirin, controlled trials have demonstrated the efficacy of this approach but both can lead to side effects such as gastrointestinal hemorrhaging. Ergotamine tartrate was the drug of choice for the acute treatment of migraine for many years but recent controlled trials have shown that it is effective in no more than 50% of patients. Dihydroergotamine (DHE) is a derivative of ergotamine whose action and side effects are similar to those of ergotamine tartrate but very much milder. It is particularly recommended for acute migraine management when given intravenously or intramuscularly. Both ergotamine and DHE are associated with a range of side effects including nausea, vomiting, abdominal pain, diarrhea, paresthesia, and leg cramps.

The newest and perhaps the most effective abortive drug available for migraine therapy is sumatriptan. Sumatriptan has been shown to be more effective than other commonly used antimigraine drugs including ergotamine and aspirin, and is also effective against the nausea and vomiting that accompany a migraine attack. The weakness of sumatriptan is that headache recurs more frequently than with other antimigraine medications. Another drawback is the high cost of the medication. Side effects of sumatriptan include sensation of hotness, chest symptoms, neck pain/stiffness, flushing, and tingling.

Five main classes of prophylactic drugs are currently used for migraine: beta-blockers, calcium antagonists, serotonin modulators, NSAIDs, and ergot compounds. A number of beta-blockers have been shown to be effective with migraine of which the most studied is propanolol. A review by Tfelt-Hansen (1986) concluded that propanolol led to an average reduction in migraine pain of 54%. The most common side effects are bradycardia, arterial hypertension, insomnia, asthma, gastrointestinal disturbances, vertigo, weight increase, depression, and bronchospasm. Flunarizine has been shown to be the most effective calcium antagonist (Leone, Grazzi, La Mantia, & Bussone, 1991) and is particularly effective in reducing headache frequency. Side effects include weight gain, depression, somnolence, and parkinsonism. Serotonin modulators include amitriptyline, methysergide, and pizotifen, the latter being the most useful and widely used. The most frequently reported side effects are somnolence, increased appetite, and weight gain, which can seriously limit the drug's usefulness.

Several NSAIDs have been shown to prevent migraine in controlled tests. These drugs impact on nausea and vomiting as well as pain. Prolonged prophylaxis with NSAIDs is contraindicated because of their well-known gastric and intestinal effects of ulceration and hemorrhage. DHE has been shown to be an effective prophylactic but is probably less so than beta-blockers, flunarizine, and pizotifen.

Many drugs used in the management of migraine have been discovered serendipitously. The introduction of beta-blockers, for example, followed the chance observation that angina patients suffered less migraine when given propanolol. The mechanism of most migraine medications is not known although multiple therapeutic actions are suspected. Combinations of these drugs are contraindicated in some cases, such as using ergotamine and sumatriptan simultaneously, or beta-blockers and calcium antagonists together.

The acute treatment of episodic tension-type headache can usually be accomplished effectively with analgesics and NSAIDs, but chronic tension-type headache is more difficult to treat (Biondi & Portuesi, 1994). The most frequently used drug treatments include three classes of drugs: analgesics, anxiolytics, and antidepressants. Treatment by analgesics is often practiced on the basis of self-prescription by the patients themselves. Benzodiazepines, such as diazepam and lorazepam, are the most commonly used anxiolytics. Long-term use can lead to analgesic and anxiolytic abuse and dependence. Caffeine frequently combined with analgesics may also cause dependence.

When headaches occur daily the medication of choice is a tricyclic antidepressant (Kunkel, 1991). These drugs are not used for their antidepressant properties but for multiple actions including an effect on serotonergic transmission. Only rarely do the common side effects of tricyclic antidepressants, such as somnolence, dry mouth, weight gain, and blurred vision, constitute a contraindication to therapy at the dosage usually employed in the treatment and prophylaxis of tension-type headache. Adding a beta-blocker such as propanolol is considered to be able to resolve cases in which the tricyclic antidepressant alone is shown to be insufficient (Speed, 1983).

Brennum et al. (1996) have argued that sumatriptan is not clinically useful in the management of episodic tension-type headache although at times it differs significantly from placebo.

8.23.7 PSYCHOLOGICAL ASSESSMENT AND TREATMENT

8.23.7.1 Psychological Assessment

Medical evaluation necessarily precedes psychological assessment as the latter does not discriminate between primary and secondary headaches. Psychological assessment of headaches usually involves three approaches: interviewing the patient and possibly significant others, self-monitoring, and administration of questionnaires (Martin, 1993). Interviewing is aimed at collecting a number of different types of information including: (i) background details that provide a psychosocial and developmental context for understanding the headaches, (ii) headache characteristics and symptoms that enable a diagnosis to be formulated, (iii) functional aspects of headaches that contribute to an understanding of the controlling variables of headaches, and (iv) assessment and treatment history. The specific questions that may be useful to ask at interview have been documented in sources such as the Psychological Assessment of Headache Questionnaire (Martin, 1993) and Headache History Form (Blanchard & Andrasik, 1985).

Self-monitoring usually takes two forms. First, headaches are typically recorded using some form of time-sampling approach whereby patients rate their headache intensity periodically and record the ratings on cards. Rating scales have varied from 3 points to 100 points and rating intervals from hourly to daily. Patients often record additional information on the cards, such as medication usage. Attempts to assess the reliability and validity of this approach have not generated very impressive data (Collins & Thompson, 1979; Blanchard, Andrasik, Neff, Jurish, & O'Keefe, 1981), but such records are likely to yield more accurate and detailed information than retrospective reports at interview.

An alternative form of self-monitoring is event-sampling whereby patients complete records when specified events occur, such as prior to, during, or immediately after headaches developing. Information required might include what patients are feeling, thinking, and doing, and the situation that they are in, at such times. A number of examples of this approach have been published (e.g., Diamond, 1979; Martin, 1983).

Many different types of questionnaire may prove helpful in assessing headaches. Questionnaires specific to headaches include the Headache Scale (Jahanshahi, Hunter, & Philips, 1986), the Headache Assessment Questionnaire (Bakal, 1982), the Headache Self-efficacy Scale (Martin, Holroyd, & Rokicki, 1993), the Headache-specific Locus of Control Scale (Vandecreek & O'Donnell, 1992), and the quality of life questionnaires referred to earlier. Of course, other types of questionnaire are also potentially helpful in assessing headache sufferers, such as questionnaires investigating commonly associated problems (e.g., anxiety and depression), stress-related phenomena (e.g., daily hassles, appraisal, coping, and social support), and personality characteristics.

8.23.7.2 Psychological Treatment

The literature on psychological treatment of headaches is very extensive and hence this review will only be able to summarize key findings. More detailed reviews are available and include Gauthier, Ivers, and Carrier (1996), Holroyd and Penzien (1994), and Martin (1993). Psychological treatments, for which the research literature is limited and no outstanding results have been reported, will be omitted from the review. This includes EEG biofeedback, skin conductance biofeedback, hypnosis, transcendental meditation, operant therapy, and covert positive reinforcement.

Before beginning the review, a few comments about terminology are necessary. Different names have been used by various writers for the same or substantially similar treatment techniques. In this review, consistent terminology will be used even when this results in departures from the original labels provided by the authors. The term thermal biofeedback will be used to cover all forms of thermal and temperature biofeedback; BVP (blood volume pulse) biofeedback will be used to include pulse amplitude biofeedback, vasomotor biofeedback, and vasoconstriction training; and cognitive coping training will be used to include cognitive therapy, cognitive-behavioral treatment, stress-coping training, and cognitive skills training.

This section is divided into five subsections. The first two cover the main forms of individually administered treatment, namely, biofeedback training; and relaxation training and cognitive coping training. These subsections are followed by consideration of group therapy, side effects of psychological treatment, and a summary of the treatment literature.

8.23.7.2.1 Biofeedback training

The main forms of biofeedback training for headaches are EMG biofeedback, thermal biofeedback and BVP biofeedback training. EMG biofeedback has usually been given from the forehead although other sites, such as the neck, have been used. The laboratory training is typically accompanied by instructions to practice at home daily. EMG biofeedback has been combined with other forms of biofeedback and with relaxation training. This type of biofeedback has usually been offered to tension-type headache sufferers but has also been used with migraineurs.

Numerous studies have shown EMG biofeedback training to be superior to headache monitoring alone (i.e., no treatment) and to a variety of attention-placebo control conditions such as pseudofeedback (e.g., Budzynski, Stoyva, Adler, & Mullaney, 1973) and pseudotherapy (Holroyd, Andrasik, & Noble, 1980). Comparisons with other forms of psychological treatment, such as relaxation training and other forms of biofeedback training, usually fail to reveal significant differences. An exception was a study by Holroyd, Andrasik, and Westbrook (1977) that demonstrated that cognitive coping training was significantly more effective than EMG biofeedback training. An interesting development is a study by Arena, Bruno, Hannah, and Meador (1995) which showed that EMG biofeedback from the trapezius muscle was superior to EMG biofeedback from the traditional forehead site. Meta-analytic reviews suggest that the combination of EMG biofeedback and relaxation training is no more effective than EMG biofeedback training alone (Holroyd & Penzien, 1986).

Paiva et al. (1982) compared the effectiveness of EMG biofeedback and diazepam. Improvement in headaches occurred more quickly with diazepam but relapse was a major problem with the pharmacological approach which was not the case for the psychological treatment.

A review of long-term follow-up studies by Blanchard (1987) revealed that while maintenance is good up to 12 months, progressive deterioration in treatment gains occurs beyond this point. The mechanism of change with EMG biofeedback training is not clear as numerous studies have failed to find correlations between reductions in EMG and headaches (e.g., Martin & Mathews, 1978). A study by Holroyd et al. (1984) suggested that cognitive mechanisms explained treatment success rather than physiological changes, and that improvement was predicted by changes in self-efficacy and locus of control.

Thermal biofeedback training usually involves training subjects to increase the temperature of their hands. Autogenic phrases are usually introduced to assist subjects, and subjects are typically encouraged to practice at home on a daily basis. Thermal biofeedback training has been combined with relaxation training and also with drug therapy. This type of biofeedback has usually been offered to migraineurs but has also been used with tension-type headache sufferers.

A number of studies have shown thermal biofeedback training to be superior to no treatment but comparisons with attention-placebo control conditions have often failed to reveal significant differences. Reading (1984) could not differentiate between thermal biofeedback and pseudofeedback, for example, and Blanchard, Appelbaum, Nicholson, et al. (1990) reported the same result with pseudomeditation. Comparisons with other forms of psychological treatment, such as relaxation training, have generally led to nonsignificant findings. One exception is a study by Elmore and Turksy (1981) which found that BVP biofeedback was superior to thermal biofeedback training. A meta-analytic review concluded that thermal biofeedback plus autogenic training was not significantly more effective than thermal biofeedback alone, although the average improvement associated with the combined treatment was 13% higher than thermal biofeedback alone (Blanchard, Andrasik, Ahles, Teders, & O'Keefe, 1980).

Thermal biofeedback training combined with relaxation training has been compared with pharmacological therapy. Mathew (1981) contrasted the psychological approach with ergotamine and found the former approach to be significantly more effective. Thermal biofeedback and relaxation training has been compared with propanolol in a number of trials and a meta-analytic review found that the two approaches produced almost identical results (Holroyd & Penzien, 1990). Mathew combined thermal biofeedback and relaxation training with either propanolol, amitriptyline, or propanolol plus amitriptyline. Amitriptyline did not enhance the effectiveness of the psychological treatment but adding propanolol yielded superior results.

The review by Blanchard (1987) suggested that treatment gains achieved with thermal biofeedback training deteriorated after 12 months but other studies have produced more promising results for maintenance (e.g., Lisspers & Ost, 1990b). The mechanism of action of thermal biofeedback is not clear as a number of studies have failed to find a relationship between temperature control and headache change (e.g., Gauthier, Bois, Allaire, & Drolet, 1981). The results of a study by Mizener, Thomas, and Billings (1988) suggested that the Holroyd model of cognitive change might apply to thermal biofeedback as well as EMG biofeedback training.

BVP biofeedback training involves providing subjects with feedback of pulse amplitude from the extracranial temporal artery. The aim is for a subject to learn to constrict the artery, thus achieving the same result as that from taking ergotamine. This form of biofeedback training has been used exclusively with migraineurs except for a few early case reports.

A number of studies have shown BVP biofeedback training to be significantly superior to headache monitoring. Friar and Beatty (1976) used BVP biofeedback from an irrelevant site as an attention-placebo control condition and demonstrated that BVP biofeedback training from the standard site was significantly more effective for headache relief. Comparisons with other forms of psychological treatment have generally failed to find differential effects except for the study by Elmore and Turksy (1981) discussed earlier.

Reich (1990) combined BVP biofeedback training with amitriptyline and found short-term enhancement of the psychological approach. BVP biofeedback alone was superior to the combination of BVP biofeedback and amitriptyline, however, after eight months.

Lisspers and Ost (1990b) published a five year follow-up of their BVP biofeedback study and reported good maintenance. The mechanism of action of BVP biofeedback is unclear as some studies have reported a relationship between BVP self-regulation skills and headache reduction (e.g., Lisspers & Ost, 1990a) and others have found no such relationship (Gauthier, Lacroix, Cote, Doyon, & Drolet, 1985).

8.23.7.2.2 Relaxation training and cognitive coping training

Relaxation training has usually been in the form of progressive relaxation or autogenic training, and is typically accompanied by instructions to practice at home. It has been used with both migraineurs and tension-type headache sufferers.

A number of studies have shown relaxation training to be superior to no treatment, and to a variety of attention-placebo control conditions, such as medication placebo (Cox, Freundlich, & Meyer, 1975) and pseudomeditation (Blanchard, Applebaum, Radnitz et al., 1990). Research has generally failed to find significant differences between relaxation training and various forms of biofeedback training or the combination of the two. One study found a trend for cognitive coping training to be superior to relaxation training (Sorbi & Tellegen, 1986), and another found a trend for enhancing the effectiveness of relaxation training by adding cognitive coping training (Blanchard, Applebaum, Radnitz et al., 1990).

Blanchard (1987) concluded that long-term maintenance associated with relaxation training was good for tension-type headache sufferers but not for migraineurs. Since this report, however, Sorbi, Tellegen, and Du Long (1989) have documented good long-term results with migraineurs. The mechanism of relaxation training is not certain, but a study suggested that Holroyd's model of cognitive change might apply to relaxation training as well as EMG biofeedback training (Blanchard, Kim, Hermann, & Steffek, 1993).

Two main variations of cognitive coping training have been developed. In the original formulation of this approach, Holroyd et al. (1977) focused on altering maladaptive cognitive responses that were assumed to lead to headaches. In contrast, Bakal, Demjen, and Kaganov (1981) developed a treatment approach that emphasized modifying stress reactions to headaches. Martin's (1993) treatment model combined these two approaches. Cognitive coping training was originally developed for tension-type headache but has also been used for migraine.

Studies have shown both variations to be superior to no treatment. Comparisons with attention-placebo control procedures have not been forthcoming but a number of studies have shown cognitive coping training to be equal to, or superior to, other treatments of proven efficacy. As noted earlier, one study reported cognitive coping training to be superior to EMG biofeedback (Holroyd et al., 1977) and another reported favorable results compared with relaxation training (Sorbi & Tellegin, 1986). The efficacy of cognitive coping training has also been demonstrated in a number of single group outcome studies (e.g. Bakal et al., 1981) and series of case studies (e.g., Anderson, Lawrence, & Olson, 1981).

Cognitive coping training is associated with good long-term results for both tension-type headache (Blanchard, 1987) and migraine (Sorbi et al., 1989). The mechanism of change is not clear, but Newton and Barbaree (1987) did demonstrate that cognitive coping training leads to positive changes in appraisal and coping, which were correlated with reduction in headache intensity.

8.23.7.2.3 Group therapy

The majority of studies on psychological treatment of headaches have used an individual format but a few studies have treated headache sufferers in groups of up to eight patients. Studies have shown group therapy using approaches such as relaxation training (Williamson et al., 1984) and cognitive coping training (Holroyd & Andrasik, 1978) to be superior to no treatment. Where comparisons have been made with attention-placebo control conditions, nonsignificant findings have often eventuated. However, Johnson and Thorn (1989) showed that cognitive coping training was as effective in the group format as the individual format.

Holroyd and Andrasik (1978) reported that adding relaxation training to cognitive coping training in a group format did not lead to greater headache reduction than did cognitive coping training alone. Murphy, Lehrer, and Jurish (1990) compared the efficacy of group relaxation training and group cognitive coping training and found the latter to be superior.

No long-term follow-up studies of group treatment have been reported nor studies of group treatment mechanisms.

8.23.7.2.4 Side effects of treatment

A paper by Szajnberg and Diamond (1980) raised the issue of symptom substitution with biofeedback training for headache. These authors interviewed a few individuals who had successfully completed biofeedback training for migraine and reported that 67% experienced some form of "symptom substitution," primarily psychosomatic symptoms or depression, accompanying or following the reduction of headache activity. This paper was criticized on methodological grounds by Haynes and Cuevas (1980), and other studies that have looked for symptom substitution have failed to find support for it (Budzynski et al., 1973). In fact there is much evidence to suggest that contrary to the findings of Szajnberg and Diamond, psychological treatment for headache seems to be associated with a reduction in psychosomatic symptoms generally and many beneficial behavioral, affective, and cognitive changes in addition to headache reduction.

Reduction in psychosomatic symptoms with headache treatment has been demonstrated with a number of types of treatment using different measures (e.g., Knapp, 1981). Decreased use of medication following treatment has been reported in investigations of all the major treatment approaches (e.g., Gauthier et al., 1981). Beneficial changes in pain behavior such as avoidance behavior, verbal complaint behavior, and nonverbal behavior has been demonstrated in a number of studies (e.g., Radnitz, Appelbaum, Blanchard, Elliott, & Andrasik, 1988). Improvement in mood states such as depression, anxiety, fatigue, irritability, and anger have been reported (e.g., Martin et al., 1989). Cognitive changes including locus of control, self-efficacy to cope with headaches, perceived control over pain, and increased ability to cope with headaches have been shown in several studies (e.g., Holroyd et al., 1984). A few studies have suggested that psychological treatment for headaches is associated with positive changes in personality such as decreases in hysteria, neuroticism, and somatization (e.g., Janssen, 1983).

8.23.7.2.5 Summary of treatment findings

(i) There is now strong evidence available that several psychological treatment approaches are effective in reducing migraine and tension-type headache, and that these gains are maintained for at least 12 months. The beneficial effects achieved with these treatments are more than attention-placebo effects. Meta-analytic reviews have calculated average percentage of improvement rates across studies, associated with biofeedback training and relaxation training, varying between 45% and 65%, in contrast to rates associated with placebo procedures varying between 12% and 38%.

(ii) In addition to reducing headaches, psychological treatment leads to decreased medication consumption and a range of positive effects in the behavioral, affective, and cognitive domains.

(iii) There is evidence to suggest that the treatment gains associated with thermal biofeedback training, BVP biofeedback training, relaxation training, and cognitive coping training are maintained for at least three to six years. Improvement achieved via EMG biofeedback training seems to deteriorate progressively after 12 months, although not back to pretreatment levels.

(iv) Comparisons among different forms of psychological treatment have not demonstrated clear superiority of one or more particular approach. Comparisons between psychological and pharmacological approaches have indicated that the former are equal or superior to the latter.

(v) There is little evidence available that combining psychological treatment approaches increases effectiveness, or that combining psychological and pharmacological treatments increases effectiveness, with the exception that the combination of thermal biofeedback, relaxation training, and propanolol seems particularly promising.

(vi) It is not clear how psychological treatments achieve their effects. EMG biofeedback training, thermal biofeedback training, relaxation training and cognitive coping training may all be mediated by cognitive mechanisms.

8.23.8 MINIMAL INTERVENTIONS AND PREVENTION

A series of studies and analyses have repeatedly made the point that current and future mental health and health needs far outstrip the capacities of present delivery systems (e.g., Winett, King, & Altman, 1991). This realization has led some researchers to explore new service delivery systems that are more cost-effective than individually administered treatment. Two such service delivery innovations are minimal interventions and prevention.

8.23.8.1 Minimal Interventions

Minimal interventions involve reducing the contact between therapists and patients by converting therapist-administered programs into largely self-administered programs. This can be accomplished by providing patients with manuals, tape recordings, and biofeedback devices that enable patients to pursue a treatment program at home. Such programs may be entirely self-administered or may be supported by limited contacts with therapists such as occasional visits and phone calls.

Most of the therapist-administered programs for headaches have been converted into minimal interventions so that the literature includes home-based versions of relaxation training, cognitive coping training, a combination of these two approaches, relaxation training plus thermal biofeedback training, and a combination of all three approaches.

The first published study of a minimal intervention for headaches compared clinic-based EMG biofeedback training and stress management training with a home relaxation program for tension-type headache (Steger & Harper, 1980). Only the clinic-based program was successful in significantly reducing headaches. Other early studies had more success, however. Sherman (1982) provided tension-type headache sufferers with relaxation tapes and all but one subject showed a substantial decrease in headache activity. Kohlenberg and Cahn (1981) contrasted the impact of providing migraineurs with an experimental book that contained instructions for thermal biofeedback training, relaxation and cognitive coping training, and a control book. The experimental book was associated with a 62% decrease in headaches compared with a 14% decrease for the control book.

Most of the home-based formats have been compared with no-treatment control conditions and have consistently been shown to be superior (e.g., Appelbaum et al., 1990; Blanchard, Applebaum, Nicholson et al., 1990). Comparisons with attention-placebo control conditions have generally been lacking except for the study by Kohlenberg and Cahn (1981) referred to above. Several studies have compared the home-based format with the clinic-based format and have consistently found no significant differences in effectiveness but large differences in cost-effectiveness in favor of the home-based format (e.g., Blanchard, Andrasik, Appelbaum et al., 1985; Teders et al., 1984).

Comparisons among the different approaches to home-based treatment programs have often suggested equal efficacy. One exception is a study by Tobin, Holroyd, Baker, Reynolds, and Holm (1988) that found the combination of relaxation plus cognitive coping training to be superior to relaxation training alone. A number of studies have compared home-based treatment with drug approaches. Relaxation plus cognitive coping training was shown to be superior to amitriptyline (Holroyd, Nash, Pingel, Cordingley, & Jerome, 1991), and thermal biofeedback plus relaxation training was shown to be equal to ergotamine tartrate (Holroyd et al., 1988). Penzien, Johnson, Carpenter, and Holroyd (1990) reported that the combination of thermal biofeedback, relaxation, and cognitive coping training was equal to propanolol, and Holroyd et al. (1995) found that adding propanolol to thermal biofeedback and relaxation training increased the effectiveness of the package.

One study showed good maintenance over two years with thermal biofeedback and relaxation training (Blanchard et al., 1988). No studies have looked at treatment mechanisms of minimal interventions.

8.23.8.2 Prevention

The public health field distinguishes three conceptually distinct levels of prevention: (i) primary, reduction of new cases of disorder or promotion of health competencies; (ii) secondary, efforts directed at individuals or groups determined at risk for or showing early signs of disorders; and (iii) tertiary, equivalent to traditional rehabilitation approaches (Burgoyne & Jason, 1991). Writers about the medical management of primary headache talk about prevention in the context of differentiating between medication prescribed for aborting attacks vs. medication prescribed for preventing attacks (see Section 8.23.6.2). Such an approach constitutes tertiary prevention as the disorder has already become established.

Writers about the psychological management of primary headaches rarely discuss whether treatment is aimed at coping with headaches or at preventing headaches occurring. A case can be made that most psychological approaches should be beneficial both for countering attacks and preventing attacks. It was suggested in Section 8.23.7.2 that EMG biofeedback, thermal biofeedback, and relaxation training might operate via cognitive changes such as increased self-efficacy and internal locus of control; and that cognitive coping training may operate through increased positive appraisal and coping thoughts. Such cognitive changes should help headache sufferers cope more adequately with headaches but should also help them cope more effectively with stressful situations that could trigger headaches. Even if the mechanism of EMG biofeedback, thermal biofeedback, and relaxation training is closer to their rationales, and involves reduced arousal/tension, the same argument could be made for both abortive and preventive beneficial effects.

BVP biofeedback training is somewhat different as it aims to promote constriction of the temporal arteries, which according to traditional models of migraine would be advantageous during attacks but not between headaches. Lisspers and Ost (1990a) recognized this fact in a study in which they trained subjects either to constrict the arteries and practice the skill during headaches, or to dilate the arteries and practice the skill during periods of stress (i.e., the preheadache stage). Parenthetically, the former approach proved more successful.

Should primary or secondary prevention approaches be developed for headaches? With respect to secondary prevention, adolescent and young adult females with a strong family history of headaches would be an obvious target group, particularly if they scored high on measures of anxiety, depression, and stress. Most psychologists see headache patients for the initial consultation around 15 years after headache onset, by which time the dynamics have become complex owing to reactions to the disorder and reactions to the reactions. Any innovations that prompted earlier referrals to psychologists would seem highly desirable.

Primary prevention specifically for headaches appears difficult to justify but any broad-based health enhancement program or stress management program should reduce the incidence of headache disorders.

8.23.9 ADHERENCE AND PROGNOSIS

8.23.9.1 Adherence

The issue of adherence has received surprisingly little attention in either the medical or psychological literature on headache. A study by Packard and O'Connell (1986) interviewed 100 patients and concluded that over 50% of headache sufferers failed to adhere properly to drug treatment regimes. Fitzpatrick, Hopkins, and Harvard-Watts (1983) found that one year after starting drug therapy, only 24% of patients reported that they had used headache medication exactly as instructed. In a study of compliance with ergotamine use, Holroyd et al. (1988) found that 70% of patients failed to make optimal use of this medication.

Holroyd et al. (1989) attempted to increase adherence with instructions for use of ergotamine by developing a brief adjunctive self-management training program administered by an allied health professional. The self-management intervention included one brief face-to-face session following a neurologist's prescription of abortive medication, and three telephone calls designed to identify and correct problems with ergotamine use. In a comparison between patients who received standard abortive therapy and patients who received abortive therapy plus the self-management intervention, the latter group attempted to abort a greater percentage of their migraine attacks (70% vs. 40%) and showed larger reductions in headaches (40% vs. 26%).

Another problem with medical management that has been addressed is compliance with attending follow-up visits (Spierings & Miree, 1993). This study found that 41% of patients attenting a specialist headache clinic had not complied with the recommended follow-up visit. The most frequently mentioned reasons for noncompliance were dislike of the clinician seen and seeking care elsewhere.

With respect to psychological management, drop-out rates are relatively low, averaging around 10–15% (e.g., Martin et al., 1989). More detailed studies pertaining to adherence are largely lacking, however. In one of the few studies in this area, Collins and Thompson (1979) investigated compliance with self-monitoring instructions. Using three objective procedures for detecting inaccurate recording, the results indicated that 40% of subjects were noncompliant. A more subjective measure suggested that this figure was a conservative estimate. Collins and Martin (1980) suggest enhancing recording compliance by reducing the demands of the procedures.

While several authors have emphasized the importance of monitoring compliance with instructions for practicing relaxation at home (e.g., Blanchard & Andrasik, 1982), a review by Hillenberg and Collins (1982) reported that only 19% of studies assessed compliance even with self-report measures. Gutkin, Holborn, Walker, and Anderson (1992) evaluated adherence to relaxation home practice instructions using a microcomputer-based method that required the patient to squeeze a hand control when instructed to tense a muscle. Of three patients in the study, the percentages of the time that they met two compliance criteria were 73%, 37%, and 8%. Improvement in headaches was greater with higher compliance.

8.23.9.2 Prognosis

The evidence with respect to the natural history of headache disorders is limited. Childhood headaches were once viewed as a transient phenomenon that the child would outgrow (Ryan, 1954), but there is increasing evidence from longitudinal studies to suggest that this is not so for the majority of children. Bille (1981), for example, found that 60% of children aged between 7 and 15 years who were diagnosed as suffering from migraine continued to report symptoms into adulthood. Similarly, Joffe, Bakal, and Kaganov (1983) stated that most adult headache sufferers report that their headaches began in childhood.

Many studies have reported predictors of success for psychological treatment but results have not always been consistent. A number of studies have reported age to be a significant predictor of treatment success (e.g., Blanchard, Andrasik, Evans, & Hillhouse, 1985). On the other hand, other studies have found a nonsignificant relationship (e.g., Martin et al., 1989), and studies of psychological treatment with the elderly have reported very positive results (Arena, Hannah, Bruno, & Meador, 1991). Diagnosis has failed to predict success for EMG biofeedback training (Bakal & Kaganov, 1977), cognitive coping training (Bakal et al., 1981), self-management training and cognitive therapy (Martin et al., 1989), and relaxation training (Williamson et al., 1984). In contrast, functional characteristics of headaches have emerged as significant predictors of treatment success although only weakly (Martin, Milech, & Nathan 1993). Hence, headaches that were precipitated by negative emotions, and headaches that were followed by the supportive partner reactions of encouraging the sufferer to stop what they were doing and sit or lie down, responded best to psychological treatment.

A number of studies have reported that patients who suffer from continuous or near continuous headaches respond less well to psychological treatments than do patients who suffer from episodic headaches (e.g., Blanchard, Appelbaum, Radnitz, Jaccard, & Dentinger, 1989). Similarly, poor response is associated with high levels of pretreatment headache activity (e.g., Jacob, Turner, Szekely, & Eidelman, 1983) and high levels of medication consumption (e.g., Michultka, Blanchard, Appelbaum, Jaccard, & Dentinger, 1989). Presence of depressive symptoms prior to treatment also suggests poor prognosis (e.g., Levine, 1984).

Although some studies have reported significant therapist effects, therapist variables such as warmth, empathy, genuineness, competence, and likeability do not predict treatment response (e.g., Martin et al., 1989). Several investigators have found regularity of home relaxation practice to be a significant predictor of treatment success (e.g., Blanchard et al., 1983), and Solbach, Sargent, and Coyne (1989) reported that quality of home practice was a better predictor than quantity of home practice.

8.23.10 PROFESSIONAL ISSUES

This final section will focus on a single issue, namely the relationship between medicine and psychology in the headache field, but will examine the issue as it manifests in a number of different domains. First to be considered will be the question of how headaches should be conceptualized. Medicine approaches the conceptualization of headaches as it does other disorders: a classification system has evolved and the task for practitioners is to formulate a diagnosis. The rationale of this approach is that it leads to the identification of headache mechanisms and consequently suggests the most appropriate treatment. Psychologists have a history of arguing for dimensional over categorical models, and for functional over symptomatological analysis. Although there is evidence pertaining to the problems associated with the current medical conceptualization of primary headaches, psychologists have been curiously reluctant to explore alternatives.

Medical researchers have recognized the importance of understanding trigger factors for headaches but have largely ignored other potentially important controlling variables. The reactions to headaches of the sufferers and significant others have received little attention in the medical literature, for example, despite the etiological significance of such reactions. The psychosocial and developmental context in which headaches occur has also received little attention. Integration between taxonomic and functional approaches should be explored.

Differential diagnosis between migraine and tension-type headache may be useful for medical practitioners in deciding, for example, between alternative symptomatic medications, such as whether to prescribe sumatriptan. The literature shows, however, that it is of limited value for psychologists. In contrast, research has demonstrated that whether headaches are accompanied by high stress levels has implications for choice of psychological treatments (Sorbi et al., 1989; Tobin et al., 1988), as does whether headaches are accompanied by depression (Martin et al., 1989).

Medical investigation of headaches is clearly a necessary part of headache management to differentiate primary and secondary headaches. Self-monitoring has been a major component of psychological assessment for many years, but has received less attention in the medical literature. This situation is changing, however, with medical researchers increasingly advocating the value of self-monitoring.

Medical practitioners predominantly treat headaches with drugs while psychologists treat headaches with psychological approaches. Evidence pertaining to the efficacy of combining a psychological approach (thermal biofeedback training plus relaxation) with a pharmacological approach (propanolol) suggests the importance of further research investigating combined technologies. Psychological therapy, perhaps involving cognitive coping training, and sumatriptan seems an obvious combination to investigate. Also, psychologists should bring their skills to bear more on the problem of increasing adherence with headache medication.

An area in which greater collaboration between psychologists and the medical profession is desirable is the referral of headache sufferers to psychologists. The average referral occurs after headache sufferers have endured headaches for around 15 years. The opportunity for earlier psychological intervention would be advantageous.

There are undoubtedly some settings in which psychologists and medical practitioners work together harmoniously and productively in the management of headaches. Overall, however, there seems to be a need for increasing interaction between researchers and practitioners from psychology and medicine. The author's own experience of attending conferences on headaches, for example, is that there are a handful of psychologists among hundreds of medical scientists and practitioners, predominantly neurologists. When medical practitioners look at alternative approaches to drug treatment, they often turn to approaches such as acupuncture and chiropractic despite the lack of research evidence that such approaches are of value in the management of primary headache. Psychologists need to collaborate with medical practitioners but psychology has much to offer medicine in understanding and managing headaches.

8.23.11 REFERENCES

Ad Hoc Committee (1962). Classification of headache. *Journal of American Medical Association, 179,* 717–718.

Allan, W. (1930). The inheritance of migraine. *Archives of Internal Medicine, 42,* 590–599.

Al-Rajeh, S., Bademosi, O., Ismaii, H., & Awada, A. (1990). Headache syndromes in the Eastern Province of Saudi Arabia. *Headache,* 30, 359–362.

Anderson, N.B., Lawrence, P.S., & Olson, T.W. (1981). Within-subject analysis of autogenic training and cognitive coping training in the treatment of tension headache pain. *Journal of Behavior Therapy and Experimental Psychiatry, 12,* 219–223.

Andrasik, F., Blanchard, E.B., Arena, J.E., Saunders, N.L., & Barron, K.D. (1982). Psychophysiology of recurrent headache: Methodological issues and new empirical findings. *Behavior Therapy, 13,* 407–429.

Andrasik, F., Blanchard, E.B., Arena, J.G., Teders, S.J., & Rodichok, L.D. (1982). Cross validation of the Kudrow–Sutkus MMPI classification system for diagnosing headache type. *Headache, 22,* 2–5.

Appelbaum, K.A., Blanchard, E.B., Nicholson, N.L., Radnitz, C., Kirsch, C., Michultka, D., Attanasio, V., Andrasik, F., & Dentinger, M.P. (1990). Controlled evaluation of the addition of cognitive strategies to a home-based relaxation protocol for tension headache. *Behavior Therapy, 21,* 293–303.

Arena, J.G., Bruno, G.M., Harmah, S.L., & Meador, K.J. (1995). A comparison of frontal electromyographic biofeedback training, trapezius electromyographic biofeedback training, and progressive muscle relaxation therapy in the treatment of tension headache. *Headache, 35,* 411–419.

Arena, J. G., Hannah, S. L., Bruno, G. M., & Meador, K. J. (1991). Electromyographic biofeedback training for tension headache in the elderly: A prospective study. *Biofeedback and Self-Regulation, 16,* 379–390.

Bakal, D.A. (1982). *The psychobiology of chronic headache.* New York: Springer.

Bakal, D.A., Demjen, S., & Kaganov, J.A. (1981). Cognitive behavioral treatment of chronic headache. *Headache, 21,* 81–86.

Bakal, D.A., & Kaganov, J.A. (1977). Muscle contraction and migraine headache: Psychophysiologic comparison. *Headache, 17,* 208–215.

Bakal, D.A., & Kaganov, J.A. (1979). Symptom characteristics of chronic and non-chronic headache sufferers. *Headache, 19,* 285–289.

Baumel, B. (1994). Migraine: A pharmacologic review with newer options and delivery modalities. *Neurology, 44* (Suppl. 3), S13–S17.

Baumgartner, C., Wessely, P., Bignol, C., Maly, J., & Holzner, F. (1989). Long-term prognosis of analgesic withdrawal in patients with drug-induced headaches. *Headache, 29,* 510–514.

Bille, B.S. (1981). Migraine in childhood and its prognosis. *Cephalalgia, 1,* 71–75.

Biondi, M., & Portuesi, G. (1994). Tension-type headache: Psychosomatic clinical assessment and treatment. *Psychotherapy and Psychosomatics, 61,* 4–64.

Blanchard, E.B. (1987). Long-term effects of behavioral treatment of chronic headache. *Behavior Therapy, 18,* 375–385.

Blanchard, E.B., & Andrasik, F. (1982). Psychological assessment and treatment of headache: Recent developments and emerging issues. *Journal of Consulting and Clinical Psychology, 50,* 859–879.

Blanchard, E.B., & Andrasik, F. (1985). *Management of chronic headaches: A psychological approach.* New York: Pergamon.

Blanchard, E.B., Andrasik, F., Ahles, T.A., Teders, S.J., & O'Keefe, D. (1980). Migraine and tension headache: A meta-analytic review. *Behavior Therapy, 11,* 613–631.

Blanchard, E.B., Andrasik, F., Appelbaum K.A., Evans, D.D., Jurish, S.E., Teders, S.J., Rodichok, L.D., & Barron, K.D. (1985). The efficacy and cost effectiveness of minimal therapist contact, nondrug treatments of chronic migraine and tension headache. *Headache, 25,* 214–220.

Blanchard, E.B., Andrasik, F., & Arena, J.G. (1984). Personality and chronic headache. Progress in Experimental Personality Research, 13, 303–360.

Blanchard, E.B., Andrasik, F., Evans, D.D., & Hillhouse, J. (1985). Biofeedback and relaxation treatments for headache in the elderly: A caution and a challenge. *Biofeedback and Self-regulation, 10,* 69–73.

Blanchard, E.B., Andrasik, F., Neff, D.F., Jurish, S.E., & O'Keefe, D.M. (1981). Social validation of the headache diary. *Behavior Therapy, 12,* 711–715.

Blanchard, E.B., Andrasik, F., Neff, D.F., Saunders, N.L., Arena, J.G., Pallmeyer, T.P., Teders, S.J., & Jurish, S.E. (1983). Four process studies in the behavioral treatment of chronic headache. *Behaviour Research and Therapy, 21,* 209–220.

Blanchard, E.B., Appelbaum, K.A., Guarnieri, P., Neff, D.F., Andrasik, F., Jaccard, J., & Barron, K.D. (1988). Two studies of the long-term follow-up of minimal therapist contact treatments of vascular and tension headache. *Journal of Consulting and Clinical Psychology, 56,* 427–432.

Blanchard, E.B., Appelbaum, K.A., Nicholson, N.L., Radnitz, C.L. Morrill, B., Michultka, D., Kirsch, C., Hillhouse, C., & Dentinger, M.P. (1990). A controlled evaluation of the addition of cognitive therapy to a home-based biofeedback and relaxation treatment of vascular headache. *Headache, 30,* 371–376.

Blanchard, E, B., Appelbaum, K. A., Radnitz, C. L., Jaccard, J., & Dentinger, M. P. (1989). The refractory headache patient-I: Chronic, daily, high intensity headache. *Behaviour Research and Therapy, 27,* 403–410.

Blanchard, E. B., Appelbaum, K. A., Radnitz, C. L., Morrill, B., Michultka, D., Kirsch, C., Guarnieri, P., Hillhouse, J., Evans, D. D., & Jaccard, J. (1990). A controlled evaluation of thermal biofeedback and thermal biofeedback combined with cognitive therapy in the treatment of vascular headache. *Journal of Consulting and Clinical Psychology, 58,* 216–224.

Blanchard, E. B., Kim, M., Hermann, C. U., & Steffek, B. D. (1993). Preliminary results of the effects on headache relief of perception of success among tension headache patients receiving relaxation. *Headache Quarterly, 4,* 249–253.

Blau, J. N., & Thavapalan, M. (1988). Preventing migraine: A study of precipitating factors. *Headache, 28,* 481–483.

Brennum, J., Brinck, T., Schriver, L., Wanscher, B., Sorensen, P. S., Tfelt-Hansen, P., & Olesen, J. (1996). Sumatriptan has no clinically relevant effect in the treatment of episodic tension-type headache. *European Journal of Neurology, 3,* 23–28.

Breslau, N., & Andreski, P. (1995). Migraine personality, and psychiatric comorbidity. *Headache, 35,* 382–386.

Brewerton, T. D., & George, M. S. (1993). Is migraine related to the eating disorders? *International Journal of Eating Disorders, 14,* 75–79.

Budzynski, T. H., Stoyva, J. M., Adler, C. S., & Mullaney, D. J. (1973). EMG biofeedback and tension headaches: A controlled outcome study. *Psychosomatic Medicine, 15,* 484–496.

Burgoyne, N. S., & Jason, L. A. (1991). Incorporating the ecological paradigm into behavioral preventive interventions. In P. R. Martin (Ed.), *Handbook of behavior therapy and psychological science* (pp. 457–472). New York: Pergamon.

Cady, R. K., & Everett, K. (1995). Introduction. In R. K. Cady & A. W. Fox (Eds.), *Treating the headache patient* (pp. 1–14). New York: Marcel Dekker.

Carlsson, J., Augustinsson, L. E., Blomstrand, C., & Sullivan, M. (1990). Health status in patients with tension headache treated with acupuncture or physiotherapy. *Headache, 30,* 593–599.

Carlsson, J., Larsson, B., & Mark, A. (1996) Psychosocial functioning in school children with recurrent headaches. *Headache, 36,* 77–82.

Clarke, C. E., MacMillan, L., Sondhi, S., & Wells, N. E. J. (1996). Economic and social impact of migraine. *Quarterly Journal of Medicine, 89,* 77–84.

Cohen, M. J., Rickles, W. H., & McArthur, D. L. (1978). Evidence for physiological response stereotypy in migraine headache. *Psychosomatic Medicine, 40,* 344–354.

Cohen, R. A., Williamson, D. P., Monguillot, J. E., Hutchinson, P. C., Gottlieb, J., & Waters, W. F. (1983). Psychophysiological response patterns in vascular and muscle contraction headaches. *Journal of Behavioral Medicine, 6,* 93–107.

Collins, F. L., & Martin, J. E. (1980). Assessing self-report of pain: A comparison of two recording procedures. *Journal of Behavioral Medicine, 2,* 55–63.

Collins, F. L., & Thompson, J. K. (1979). Reliability and standardization in the assessment of self-reported headache pain. *Journal of Behavioral Medicine, 1,* 73–86.

Cox, D. J., Freundlich, A., & Meyer, R. G. (1975). Differential effectiveness of electromyograph feedback, verbal relaxation instructions, and medication placebo with tension headaches. *Journal of Consulting and Clinical Psychology, 43,* 892–898.

D'Amico, D., Leone, M., Macciardi, F., Valentini, S., & Bussone, G. (1991). Genetic transmission of migraine without aura: A study of 68 families. *Italian Journal of Neurological Science, 12,* 581–584.

de Lissovoy, G., & Lazarus, S. S. (1994). The economic cost of migraine: Present state of knowledge. *Neurology, 44,* 556–562.

Deubner, D. C. (1977). An epidemiologic study of migraine and headache in 10–20 year olds. *Headache, 17,* 173–180.

Diamond, S. (1979). Headache. *American Academy of Family Physicians, October,* 6–18.

Diamond, S., & Maliszewski, M. (1992). A brief history of headache. In M. Maliszewski (Ed.), *Sexual aspects of headaches* (pp. 1–10). Madison, CT: International Universities Press.

Dieter, J. N., & Swerdlow, B. (1988). A replicative investigation of the reliability of the MMPI in the classification of chronic headaches. *Headache, 28,* 212–222.

Domino, J. V., & Haber, J. D. (1987). Prior physical and sexual abuse in women with chronic headache: Clinical correlates. *Headache, 27,* 310–314.

Drummond, P. D. (1982). Extracranial and cardiovascular reactivity in migrainous subjects. *Journal of Psychosomatic Research, 26,* 317–331.

Duckro, P. N., Schultz, K. T., & Chilbnall, J. T. (1994). Migraine as a sequela to chronic low back pain. *Headache, 34,* 279–281.

Ehde, D. M., & Holm, J. E. (1992). Stress and headache: Comparisons of migraine, tension and headache-free subjects. *Headache Quarterly, 3,* 54–60.

Elmore, A. M., & Turksy, B. (1981). A comparison of two psycho physiological approaches to the treatment of migraine. *Headache, 21,* 93–101.

Essink-Bot, M. L., Van Royen, L., Krabbe, P., Bonsel, G. J., & Rutten, F. F. (1995). The impact of migraine on health status. *Headache, 35,* 200–206.

Evans, R. W. (1996). Diagnostic testing for the evaluation of headaches. *Neurologic Clinics, 14,* 1–26.

Featherstone, H. J. (1985). Medical diagnosis and problems in individuals with recurrent idiopathic headaches. *Headache, 25,* 136–140.

Fitzpatrick, R. M., Hopkins, A. P., & Harvard-Watts, O. (1983). Social dimensions of healing: A longitudinal study of outcomes of medical management of headaches. *Social Science and Medicine, 17,* 501–510.

Flor, H., & Turk, D. C. (1989). Psychophysiology of chronic pain: Do chronic pain patients exhibit symptom-specific psychophysiological responses? *Psychological Bulletin, 105,* 215–259.

Forsyth, M. R., & Martin, P. R. (1997). The need for vigilance in identifying those at risk from drug-induced headache. *Behavior Change, 14,* 1–3.

Friar, L. R., & Beatty, J. (1976). Migraine: Management by trained control of vasoconstriction. *Journal of Consulting and Clinical Psychology, 44,* 46–53.

Friedman, A. P. (1972). The headache in history, literature, and legend. *Bulletin New York Academy of Medicine, 48,* 661–681.

Frishberg, B. M. (1994). The utility of neuroimaging in the evaluation of headache in patients with normal neurologic examination. *Neurology, 44,* 1191–1197.

Gannon, L. R., Haynes, S. N., Cuevas, J., & Chavez, R. (1987). Psychophysiological correlates of induced headaches. *Journal of Behavioral Medicine, 10,* 411–423.

Gannon, L. R., Haynes, S. N., Safranek, R., & Hamilton, J. (1981). A psychophysiological investigation of muscle-contraction and migraine headache. *Journal of Psychosomatic Research, 25,* 271–280.

Gauthier, J., Bois, R., Allaire, D., & Drolet, M. (1981). Evaluation of skin temperature biofeedback training at two different sites for migraine. *Journal of Behavioral Medicine, 4,* 407–419.

Gauthier, J., Lacroix, R., Cote, A., Doyon, J., & Drolet, M. (1985). Biofeedback control of migraine headaches: A comparison of two approaches. *Biofeedback and Self-regulation, 10,* 139–159.

Gauthier, J. G., Ivers, H., & Carrier, S. (1996). Non-pharmacological approaches in the management of recurrent headache disorders and their comparison and combination with pharmacotherapy. *Clinical Psychology Review, 16,* 542–571.

Gladstein, J., Holden, E. W., Peralta, L., & Raven, M. (1993). Diagnoses and symptom patterns in children presenting to a pediatric headache clinic. *Headache, 33,* 497–500.

Gobel, H. (1994). Paper-pencil tests for retrospective and prospective evaluation of primary headaches on the basis of the IHS criteria. *Headache, 34,* 564–568.

Goldstein, M., & Chen, T. C. (1982). The epidemiology of disabling headache. In M. Critchley, A. P. Friedman, & S. Gorini (Eds.), *Advances in Neurology* (Vol. 33, pp. 377–390). New York: Raven Press.

Granella, F., Sances, G., Zanferrari, C., Costa, A., Martignoni, E., & Manzoni, G. C. (1993). Migraine without aura and reproductive life events: A clinical epidemiological study in 1300 women. *Headache, 33,* 385–389.

Gutkin, A. J., Holborn, S. W., Walker, J. R., & Anderson, B. A. (1992). Treatment integrity of relaxation training for tension headaches. *Journal of Behavior Therapy and Experimental Psychiatry, 23,* 191–198.

Haimanot, R. T., Seraw, B., Forsgren, L., Ekbom, K., & Ekstedt, J. (1995). Migraine, chronic tension-type headache and cluster headache in an Ethiopian rural community. *Cephalalgia, 15,* 482–488.

Hartmaier, S. L., Santanello, N. C., Epstein, R. S., & Silberstein, S. D. (1995). Development of a brief 24-hour migraine-specific quality of life questionnaire. *Headache, 35,* 320–329.

Haynes, S. N., & Cuevas, J. L. (1980). Symptom substitution and biofeedback treatment of migraine headache: A comment. *Headache, 20,* 347–348.

Haynes, S. N., Gannon, L. R., Bank, J., Shelton, D., & Goodwin, J. (1990). Cephalic blood flow correlates of induced headaches. *Journal of Behavioral Medicine, 13,* 467–480.

Haynes, S. N., Griffin, P., Mooney, D., & Parise, M. (1975). Electromyographic biofeedback and relaxation instructions in the treatment of muscle contraction headache. *Behavior Therapy, 6,* 672–678.

Headache Classification Committee of the International Headache Society (1988). Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain. *Cephalalgia, 8*(Suppl. 7), 91–96.

Henry, P., Michel, P., Brochet, B., Dartigues, J. F., Tison, S., & Salamon, R. (1992). A nationwide survey of migraine in France: Prevalence and clinical features in adults. *Cephalalgia, 12,* 229–237.

Henryk-Gutt, R., & Rees, W. L. (1973). Psychological aspects of migraine. *Journal of Psychosomatic Research, 17,* 141–153.

Hillenberg, J. B., & Collins, F. L. (1982). A procedural analysis and review of relaxation training research. *Behaviour Research and Therapy, 20,* 251–260.

Holden, E. W., Gladstein, J., Trulsen, M., & Wall, B. (1994). Chronic daily headache in children and adolescents. *Headache, 34,* 508–514.

Holm, J. E., Holroyd, K. A., Hursey, K. G., & Penzien, D. B. (1986). The role of stress in recurrent tension headache. *Headache, 26,* 160–167.

Holroyd, K. A., & Andrasik, F. (1978). Coping and self-control of chronic tension headache. *Journal of Consulting and Clinical Psychology, 46,* 1036–1045.

Holroyd, K. A., Andrasik, F., & Noble, J. (1980). A comparison of EMG biofeedback and a credible pseudotherapy in treating tension headache. *Journal of Behavioral Medicine, 3,* 29–39.

Holroyd, K. A., Andrasik, F., & Westbrook, T. (1977). Cognitive control of tension headache. *Cognitive Therapy and Research, 1,* 121–133.

Holroyd, K. A., France, J. L., Cordingley, G. E., Rokicki, L. A., Kvaal, S. A., Lipchik, G. T., & McCool, M. R. (1995). Enhancing the effectiveness of relaxation–thermal biofeedback training with propanolol hydrochloride. *Journal of Consulting and Clinical Psychology, 63,* 327–330.

Holroyd, K. A., Holm, J. E., Hursey, K. G., Penzien, D. B., Cordingley, G. E., Theofanous, A. G., Richardson, S. C., & Tobin, D. L. (1988). Recurrent vascular headache: Home-based behavioral treatment vs. abortive pharmacological treatment. *Journal of Consulting and Clinical Psychology, 56,* 218–223.

Holroyd, K. A., Holm, J. F., Penzien, D. B., Cordingley, G. E., Hursey, K. G., Martin, N. J., & Theofanous, A. G. (1989). Long-term maintenance of improvements achieved with (abortive) pharmacological and nonpharmacological treatments for migraine: Preliminary findings. *Biofeedback and Self-regulation, 14,* 301–308.

Holroyd, K. A., Nash, J. M., Pingel, J. D., Cordingley, G. E., & Jerome, A. (1991). A comparison of pharmacological (amitriptyline HCL) and nonpharmacological (cognitive-behavioral) therapies for chronic tension headaches. *Journal of Consulting and Clinical Psychology, 59,* 387–393.

Holroyd, K. A., & Penzien, D. B. (1986). Client variables and the behavioral treatment of recurrent tension headache: A meta-analytic review. *Journal of Behavioral Medicine, 9,* 515–536.

Holroyd, K. A., & Penzien, D. B. (1990). Pharmacological vs. nonpharmacological prophylaxis of recurrent migraine headache: A meta-analytic review of clinical trials. *Pain, 42,* 1–13.

Holroyd, K. A., & Penzien, D. B. (1994). Psychosocial interventions in the management of recurrent headache disorders. 1: Overview and effectiveness. *Behavioral Medicine, 20,* 53–63.

Holroyd, K. A., Penzien, D. B., Hursey, K. G., Tobin, D. L., Rogers, L., Holm, J. F., Marcille, P. J., Hall, J. R., & Chila, A. G. (1984). Change mechanisms in EMG biofeedback training: Cognitive changes underlying improvements in tension headache. *Journal of Consulting and Clinical Psychology, 52,* 1039–1053.

Howarth, E. (1965). Headache, personality and stress. *British Journal of Psychiatry, 111,* 1193–1197.

Jacob, R. G., Turner, S. M., Szekely, B. C., & Eidelman, B. H. (1983). Predicting outcome of relaxation therapy in headaches: The role of "depression." *Behavior Therapy, 14,* 457–465.

Jahanshahi, M., Hunter, M., & Philips, C. (1986). The headache scale: An examination of its reliability and validity. *Headache, 26,* 76–82.

Janssen, K. (1983). Differential effectiveness of EMG feedback vs. combined EMG-feedback and relaxation instructions in the treatment of tension headache. *Journal of Psychosomatic Research, 27,* 243–253.

Jenkinson, C. (1990). Health status and mood state in a migraine sample. *International Journal of Social Psychiatry, 36,* 42–48.

Joffe, R., Bakal, D. A., & Kaganov, J. (1983). A self-observation study of headache symptoms in children. *Headache, 23,* 20–25.

Johnson, P. R., & Thorn, B. E. (1989). Cognitive-behavioral treatment of chronic headache: Group vs. individual treatment format. *Headache, 29,* 358–365.

Kaiser, R. S., & Primavera, J. P. (1993). Failure to mourn as a possible contributory factor to headache onset in adolescence. *Headache, 33,* 69–72.

Kalfakis, N., Panaos, M., Vassilopouilos, D., & Malliara-Loulakaki, S. (1996). Migraine with aura: Segregation analysis and heritability estimation. *Headache, 36,* 320–322.

Knapp, T. W. (1982). The treatment of migraine headache by training in vasoconstriction of the temporal artery and a cognitive stress-coping training. *Behavioral Analysis and Modification, 4,* 267–274.

Kohlenberg, R. J., & Cahn, T. (1981). Self-help treatment for migraine headaches: A controlled outcome study. *Headache, 21,* 196–200.

Kohler, T., & Kosanic, S. (1992). Are persons with migraine characterized by a high degree of ambition, orderliness and rigidity? *Pain, 48,* 321–323.

Kristjansdottir, G., & Wahlberg, V. (1993). Sociodemographic differences in the prevalence of self-reported headache in Icelandic school-children. *Headache, 33,* 376–380.

Kudrow, L. (1978). Current aspects of migraine headache. *Psychosomatics, 19,* 48–57.

Kundel, R. S. (1991). Diagnosis and treatment of muscle contraction (tension-type) headaches. *Medical Clinics of North America, 75,* 595–603.

Lacroix, R., & Barbaree, H. E. (1990). The impact of recurrent headaches on behaviour, lifestyle and health. *Behaviour Research and Therapy, 28,* 235–242.

Lance, J. W. (1993). *Mechanisms and management of headache.* Oxford, UK: Butterworth Heinemann.

Langeveld, J. H., Koot, H. M., Loonen, M. C. B., Hazebroek-Kampschreur, A. A. J. M., & Passchier, J. (1996). A quality of life instrument for adolescents with chronic headache. *Cephalalgia, 16,* 183–196.

Lehrer, P. M., & Murphy, A. I. (1991). Stress reactivity and perception of pain among tension headache sufferers. *Behaviour Research and Therapy, 29,* 61–69.

Leone, M., Grazzi, L., D'Amico, D., Moschiano, F., & Bussone, G. (1995). A review of the treatment of primary headaches. Part I: Migraine. *Italian Journal of Neurological Science, 16,* 577–586.

Leone, M., Grazzi, L., La Mantia, L., & Bussone, G. (1991). Flunarizine in migraine: A mini-review. *Headache, 31,* 388–391.

Levine, B. A. (1984). Effects of depression and headache type on biofeedback for muscle-contraction headaches. *Headache, 31,* 388–391.

Levy, L. M. (1983). An epidemiological study of headache in an urban population in Zimbabwe. *Headache, 23,* 2–9.

Lichstein, K. L., Fischer, S. M., Eakin, T. L., Amberson, J. L., Bertorini, T., & Hoon, P. W. (1991). Psychophysiological parameters of migraine and muscle-contraction headaches. *Headache, 31,* 27–34.

Lipton, R. B., & Stewart, W. F. (1993). Migraine in the United States: A review of epidemiology and health care use. *Neurology, 43*(Suppl. 3), 6–10.

Lipton, R. B., Stewart, W. F., & Von Korff, M. (1995). Migraine impact and functional disability. *Cephalalgia, 15,* 4–9.

Lisspers, J., & Ost, L. G. (1990a). BVP biofeedback in the treatment of migraine. The effects of constriction and dilatation during different phases of the migraine attack. *Behavior Modification, 14,* 200–221.

Lisspers, J., & Ost, L. G. (1990b). Long-term follow-up of migraine treatment: Do the effects remain up to six years? *Behaviour Research and Therapy, 28,* 313–322.

Lucas, R. N. (1977). Migraine in twins. *Journal of Psychosomatic Research, 20,* 147–156.

Marazziti, D., Toni, C., Pedri, S., Bonuccelli, U., Pavese, N., Nuti, A., Muratorio, A., Cassano, G. B., & Akiskal, H. S. (1995). Headache, panic disorder and depression: Comorbidity or a spectrum? *Neuropsychobiology, 31,* 125–129.

Marcus, D. A., Nash, J. M., & Turk, D. C. (1994). Diagnosing recurring headaches: IHS criteria and beyond. *Headache, 34,* 329–336.

Martin, M. J. (1972). Muscle-contraction headache. *Psychosomatics, 13,* 16–19.

Martin, N. J., Holroyd, K. A., & Rokicki, L. A. (1993). The Headache Self-efficacy Scale: Adaptation to recurrent headaches. *Headache, 33,* 244–248.

Martin, P. R. (1983). Behavioural research on headaches: Current status and future directions. In K. A. Holroyd, B. Schlote, & H. Zenz (Eds.), *Perspectives in research on headaches* (pp. 204–215). Lewiston, NY: Hogrefe.

Martin, P. R. (1985). Classification of headache: The need for a radical revision. *Cephalalgia, 5,* 1–4.

Martin, P. R. (1993). *Psychological management of chronic headaches.* New York: Guilford Press.

Martin, P. R., Marie, G. V., & Nathan, P. R. (1992). Psychophysiological mechanisms of chronic headaches: Investigation using pain induction and pain reduction procedures. *Journal of Psychosomatic Research, 36,* 132–148.

Martin, P. R., & Mathews, A. M. (1978). Tension headaches: Psychophysiological investigation and treatment. *Journal of Psychosomatic Research, 23,* 389–399.

Martin, P. R., Milech, D., & Nathan, P. R. (1993). Towards a functional model of chronic headaches: Investigation of antecedents and consequences. *Headache, 33,* 461–470.

Martin, P. R., Nathan, P. R., & Milech, D. (1987). The Type A behaviour pattern and chronic headaches. *Behaviour Change, 4,* 33–39.

Martin, P. R., Nathan, P. R., Milech, D., & van Keppel, M. (1989). Cognitive therapy vs. self-management training in the treatment of chronic headaches. *British Journal of Clinical Psychology, 28,* 347–361.

Martin, P. R., & Seneviratne, H. M. (1997). Trigger factors of benign recurring headaches: The role of negative affect and hunger. *Health Psychology, 16,* 1–9.

Martin, P. R., & Soon, K. (1993). The relationship between perceived stress, social support and chronic headaches. *Headache, 33,* 307–314.

Martin, P. R., & Teoh, H. J. (1997a). *Effect of visual stimuli and a stressor on head pain.* Manuscript submitted for publication.

Martin, P. R., & Teoh, H. J. (1997b). *Headache triggers: To avoid or not to avoid, that is the question.* Manuscript submitted for publication.

Martin, P. R., & Theunissen, C. (1993). The role of life event stress, coping and social support in chronic headaches. *Headache, 33,* 301–306.

Martin, P. R., & Todd, J. (1996). *Effect of noise and a stressor on head pain.* Unpublished manuscript.

Mathew, N. T. (1981). Prophylaxis of migraine and mixed headache: A randomized controlled study. *Headache, 21,* 105–109.

Medina, J. L., & Diamond, S. (1978). The role of diet in migraine. *Headache, 18,* 31–34.

Melzack, R., & Wall, P. D. (1988). *The challenge of pain.* Harmondsworth, UK: Penguin.

Michel, P., Dartigues, J. F., Henry, P., Tison, S., Auriacombe, S., Brochet, B., Vivares, C. K., Salamon, R., & the GRIM (1993). Validity of the International

Headache Society criteria for migraine. *Neuroepidemiology, 12,* 51–57.

Michultka, D. M., Blanchard, E. B., Appelbaum, K. A., Jaccard, J., & Dentinger, M. P. (1989). The refractory headache patient-II: High medication consumption (analgesic rebound) headache. *Behaviour Research and Therapy, 27,* 411–420.

Mitsikostas, D. D., Thomas, A., Gatzonis, S., Ilias, A., & Papageorgiou, C. (1994). An epidemiological study of headache among the monks of Athos (Greece). *Headache, 34,* 539–541.

Mizener, D., Thomas, M., & Billings, R. (1988). Cognitive changes of migraineurs receiving biofeedback training. *Headache, 28,* 339–343.

Moffett, A. M., Swash, M., & Scott, D. F. (1974). Effect of chocolate in migraine: A double-blind study. *Journal of Neurology, Neurosurgery and Psychiatry, 37,* 445–448.

Morbidity and Mortality Weekly Report (1991). Prevalence of chronic migraine headache—United States, 1980–1989. *MMWR Morbidity and Mortality Weekly Report, 40,* 331, 337–338.

Morley, S. (1985). An experimental investigation of some assumptions underpinning psychological treatments of migraine. *Behaviour Research and Therapy, 23,* 65–74.

Morrison, D. P., & Price, W. H. (1989). The prevalence of psychiatric disorder among female new referrals to a migraine clinic. *Psychological Medicine, 19,* 919–925.

Murphy, A. I., Lehrer, P. M., & Jurish, S. (1990). Cognitive coping skills and relaxation training as treatments for tension headaches. *Behavior Therapy, 21,* 89–98.

Newton, C. R., & Barbaree, H. E. (1987). Cognitive changes accompanying headache treatment: The use of a thought-sampling procedure. *Cognitive Therapy and Research, 11,* 635–651.

Nikiforow, R. M., & Hokkanen, E. (1978). An epidemiological study of headache in an urban and rural population in Northern Finland. *Headache, 18,* 137–145.

Olesen, J., & Lipton, R. B. (1994). Migraine classifications and diagnosis. *Neurology, 44,* 6–10.

Olesen, J., & Rasmussen, B. K. (1995). Classification of primary headaches. *Biomedicine and Pharmacotherapy, 49,* 446–451.

Osterhaus, J. T., Gutterman, D. G., & Plachetica, J. R. (1992). Healthcare resource use and lost labour costs of migraine headache in the US. *Pharmacoeconomics, 2,* 67–76.

Osterhaus, J. T., Townsend, R. J., Gandek, B., & Ware, J. E., Jr. (1994). Measuring the functional status and well-being of patients with migraine headache. *Headache, 34,* 337–343.

Osterhaus, S., & Passchier, J. (1992). Perception of triggers in young nonclinical school students with migrainous headaches and with tension headaches. *Perceptual and Motor Skills, 75,* 284–286.

Packard, R. C., & O'Connell, P. (1986). Medication compliance among headache patients. *Headache, 26,* 411–419.

Paiva, T., Martins, P., Batista, A., Esperanca, P., & Martins, I. (1994). Sleep disturbances in chronic headache patients: A comparison with healthy controls. *Headache Quarterly, 5,* 135–141

Paiva, T., Nunes, S., Moreira, A., Santos, J., Teixeira, J., & Barbosa, A. (1982). Effects of frontal EMG biofeedback and diazepam in the treatment of tension headache. *Headache, 22,* 216–220.

Passchier, J., van der Helm-Hylkema, H., & Orlebeke, J. F. (1984a). Personality and headache type: A controlled study. *Headache, 24,* 140–146.

Passchier, J., van der Helm-Hylkema, H., & Orlebeke, J. F. (1984b). Psychopsychological characteristics of migraine and tension headache patients. Differential effects of sex and pain state. *Headache, 24,* 131–139.

Peck, D. F., & Attfield, M. E. (1981). Migraine symptoms on the Waters Headache Questionnaire. A statistical analysis. *Journal of Psychosomatic Research, 25, 281–288.*

Penzien, D. B., Holroyd, K. A., Holm, J. E., & Hursey, K. G. (1985). Psychometric characteristics of the Bakal Headache Assessment Questionnaire. *Headache, 25,* 55–58.

Penzien, D. B., Johnson, C., Carpenter, D., & Holroyd, K. (1990). Drug vs. behavioral treatment of migraine: Long-acting propanolol vs. home-based self-management training. *Headache, 30,* 300.

Philips, C. (1976). Headache and personality. *Journal of Psychosomatic Research, 20,* 535–542.

Philips, H. C. (1977). A psychological analysis of tension headache. In S. Rachman (Ed.), *Contributions to medical psychology* (Vol. 1, pp. 91–151). Oxford, UK: Pergamon.

Philips, H. C. (1987). Avoidance behaviour and its role in sustaining chronic pain. *Behaviour Research and Therapy, 25,* 273–279.

Philips, H. C. (1989). Thoughts provoked by pain. *Behaviour Research and Therapy, 27,* 469–473.

Philips, H. C., & Hunter, M. (1981). Pain behavior in headache sufferers. *Behavior Analysis and Modification, 4,* 257–266.

Philips, H. C., & Jahanshahi, M. (1985). Chronic pain: An experimental analysis of the effects of exposure. *Behaviour Research and Therapy, 23,* 281–290.

Radnitz, C. L., Appelbaum, K. A., Blanchard, E. B., Elliott, L., & Andrasik, F. (1988). The effect of self-regulatory treatment on pain behavior in chronic headache. *Behaviour Research and Therapy, 26,* 253–260.

Ramsden, R., Friedman, B., & Williamson, D. (1983). Treatment of childhood headache reports with contingency management procedures. *Journal of Clinical Child Psychology, 12,* 202–206.

Rapoport, A. M. (1988). Analgesic rebound headache. *Headache, 28,* 662–665.

Rappaport, N. B., McAnulty, D. P., Waggoner, C. D., & Brantley, P. J. (1987). Cluster analysis of Minesota Multiphasic Personality Inventory (MMPI) profiles in a chronic headache population. *Journal of Behavioral Medicine, 10,* 49–60.

Rasmussen, B. K. (1993). Migraine and tension-type headache in a general population: Precipitating factors, female hormones, sleep pattern and relation to lifestyle. *Pain, 53,* 65–72.

Rasmussen, B. K., Jensen, R., & Olesen, J. (1992). Impact of headache on sickness, absence and utilisation of medical services: A Danish population study. *Journal of Epidemiology and Community Health, 46,* 443–446.

Rasmussen, B. K., Jensen, R., Schroll, M., & Olesen, J. (1991). Epidemiology of headache in a general population: A prevalence study. *Journal of Clinical Epidemiology, 44,* 1147–1157.

Rasmussen, B. K., Jensen, R., Schroll, M., & Olesen, J. (1992). Interrelations between migraine and tension-type headache in the general population. *Archives of Neurology, 49,* 914–918.

Reading, C. (1984). Psychophysiological reactivity in migraine following biofeedback. *Headache, 24,* 70–74.

Reich, B. A. (1990). Photoplethysmograph biofeedback combined with tricyclic medication: A two-year assessment. *Headache, 30,* 313.

Riley, M. A. (1932). Special article. Migraine. *Bulletin Neurological Institute of New York, 2,* 429–544.

Robinson, M. E., Geisser, M. E., Dieter, J. N., & Swerdlow, B. (1991). The relationship between MMPI cluster membership and diagnostic category in headache patients. *Headache, 31,* 111–115.

Rojahn, J., & Gerhards, F. (1986). Subjective stress sensitivity and physiological responses to an aversive auditory stimulus in migraine and control subjects. *Journal of Behavioral Medicine, 9,* 203–212.

Rose, F. C. (1995). The history of migraine from Mesopotamian to Medieval times. *Cephalalgia, 15*(Suppl. 15), 1–3.

Ryan, R. E. (1954). *Headache: Diagnosis and treatment*. St Louis, MI: Mosby.

Saper, J. R. (1987). Ergotamine dependency: A review. *Headache, 27,* 435–438.

Saper, J. R., Silberstein, S., Gordon, C. D., & Hamel, R. (1993). *Handbook of headache management: A practical guide to diagnosis and treatment of headache and facial pain*. Baltimore: Williams & Wilkins.

Schulman, E. A., & Silberstein, S. D. (1992). Symptomatic and prophylactic treatment of migraine and tension-type headache. *Neurology, 42*(Suppl. 2), 16–21.

Seshia, S. S., & Wolstein, J. R. (1995). International Headache Society classification and diagnostic criteria in children: A proposal for revision. *Developmental Medicine and Child Neurology, 37,* 879–882.

Sherman, R. A. (1982). Home use of tape-recorded relaxation exercises as initial treatment for stress related disorders. *Military Medicine, 147,* 1062–1066.

Silberstein, S. D. (1994). Tension-type headaches. *Headache, 34,* S2–S7.

Silberstein, S. D., Lipton, R. B., Solomon, S., & Mathew, N. T. (1994). Classification of daily and near-daily headaches: Proposed revisions to the IHS criteria. *Headache, 34,* 1–7.

Solbach, P., Sargent, J., & Coyne, L. (1989). An analysis of home practice patterns for non-drug headache treatments. *Headache, 29,* 528–531.

Solomon, G. D., Skobieranda, F. G., & Gragg, L. A. (1993). Quality of life and well-being of headache patients: Measurement by the Medical Outcomes Study Instrument. *Headache, 33,* 351–358.

Solomon, S. (1994). Migraine diagnosis and clinical symptomatology. *Headache, 34,* S8–S12.

Sorbi, M., & Tellegen, B. (1986). Differential effects of training in relaxation and stress-coping in patients with migraine. *Headache, 26,* 473–481.

Sorbi, M., Tellegen, B., & Du Long, A. (1989). Long-term effects of training in relaxation and stress-coping in patients with migraine: A 3-year follow-up. *Headache, 29,* 111–121.

Speed, W. G. (1983). Muscle contraction headache. In J. R. Saper (Ed.), *Headache disorders: Current concepts and treatment strategies* (pp. 115–24). Boston: J. Wright-PSG.

Spierings, E. L. H., & Miree, L. F. (1993). Non-compliance with follow-up and improvement after treatment at a Headache Centre. *Headache, 33,* 205–209.

Spierings, E. L. H., & Schellekens, J. A. (1991). Causes of headache in high school students. *Headache Quarterly, 2,* 225–228.

Spierings, E. L. H., Sorbi, M., Haimowitz, B. R., & Tellegen, B. (1996). Changes in daily hassles, mood and sleep in the 2 days before a migraine headache. *The Clinical Journal of Pain, 12,* 38–42.

Stang, P. E., & Osterhaus, J. T. (1993). Impact of migraine in the United States: Data from the National Health Interview Survey. *Headache, 33,* 29–35.

Steger, J. C., & Harper, R. G. (1980). Comprehensive biofeedback vs. self-monitored relaxation in the treatment of tension headache. *Headache, 20,* 137–142.

Stewart, W. F., Linet, M. S., Celentano, D. D., Van Natta, M., & Ziegler, D. (1993). Age- and sex-specific incidence rates of migraine with and without visual aura. *American Journal of Epidemiology, 134,* 1111–1120.

Stewart, W. F., & Lipton, R. B. (1994). The economic and social impact of migraine. *European Neurology, 34*(Suppl. 2), 12–17.

Stewart, W. F., Lipton, R. B., Celentano, D. D., & Reed, M. L. (1992). Prevalence of migraine headache in the United States: Relation to age, income, race and other sociodemographic factors. *Journal American Medical Association, 267,* 64–69.

Stewart, W. F., Simon, D., Shechter, A., & Lipton, R. B. (1995) Population variation in migraine prevalence: A meta analysis. *Journal of Clinical Epidemiology, 48,* 269–280.

Stout, M. A. (1984). A cognitive-behavioral study of self-reported stress factors in migraine headache. *Psychopathology, 17,* 290–296.

Szajnberg, N., & Diamond, S. (1980). Biofeedback, migraine headache and new symptom formation. *Headache, 20,* 29–31.

Teders, S. J., Blanchard, E. B., Andrasik, F., Jurish, S. E., Neff, D. F., & Arena, J. G. (1984). Relaxation training for tension headache: Comparative efficacy and cost-effectiveness of a minimal therapist contact vs. a therapist delivered procedure. *Behavior Therapy, 15,* 59–70.

Tfelt-Hansen, P. (1986). Efficacy of beta-blockers in migraine: A critical review. *Cephalalgia, 6* (Suppl. 5), 15–24.

Tobin, D. L., Holroyd, K. A., Baker, A., Reynolds, R. G. C., & Holm, J. E. (1988). Development and clinical trial of a minimal contact, cognitive-behavioral treatment for tension headache. *Cognitive Therapy and Research, 12,* 325–339.

Uknis, A., & Silberstein, S. D. (1991). Migraine and pregnancy. *Headache, 31,* 372–374.

Vandecreek, L., & O'Donnell, F. (1992). Psychometric characteristics of the Headache-specific Locus of Control Scale. *Headache, 32,* 239–241.

Waters, W. E. (1971). Migraine: Intelligence, social class, and familial prevalence. *British Medical Journal, 2,* 77–81.

Weiss, H. D., Stern, B. J., & Goldberg, J. (1991). Post-traumatic migraine: Chronic migraine precipitated by minor head or neck trauma. *Headache, 31, 451–456.*

Williamson, D. A., Monguillot, J. E., Jarrell, H. P., Cohen, R. A., Pratt, J. M., & Blouin. D. C. (1984). Relaxation for the treatment of headache: Controlled evaluation of two group programs. *Behavior Modification, 8,* 407–424.

Winett, R. A., King, A. C., & Altman, D. G. (1991). Extending applications of behavior therapy to large-scale intervention. In P. R. Martin (Ed.), *Handbook of behavior therapy and psychological science: An integrative approach* (pp. 473–490). New York: Pergamon.

Wolff, H. G. (1963). *Headache and other head pain*. New York: Oxford University Press.

World Health Organization (1992). *International classification of diseases and related health problems 10*. Geneva, Switzerland: Author.

Yen, S., & McIntire, R. W. (1971). Operant therapy for constant headache complaints: A simple response–cost approach. *Psychological Reports, 28,* 267–270.

Ziegler, D. K. (1978). The epidemiology and genetics of migraine. *Research and Clinical Studies in Headache, 9,* 21–33.

Ziegler, D. K., Hassanein, R., & Hassanein, K. (1972). Headache syndromes suggested by factor analysis of symptom variables in a headache prone population. *Journal of Chronic Diseases, 25,* 353–363.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.24
Chronic Pain

SHIRLEY PEARCE and ANNE-LISE McDONALD
University of East Anglia, Norwich, UK

8.24.1 INTRODUCTION 557

8.24.2 WHAT IS PAIN? 558

8.24.3 PSYCHOLOGICAL INFLUENCES ON PAIN 559
- *8.24.3.1 Perception of Control* 559
- *8.24.3.2 Mood* 559
- *8.24.3.3 Personality* 559
- *8.24.3.4 Social and Environmental Factors* 560

8.24.4 CURRENT MODELS OF PAIN 560
- *8.24.4.1 Psychogenic or Organic Pain* 561

8.24.5 PAIN ASSESSMENT 562

8.24.6 ASSESSMENT OF PAIN BEHAVIORS 562
- *8.24.6.1 Direct Observation of Pain Behavior* 562
- *8.24.6.2 Electromechanical Devices* 563
- *8.24.6.3 Self-observation* 563

8.24.7 MEASUREMENT OF THE SUBJECTIVE COMPONENT OF PAIN: MEASURING THE QUALITY AND INTENSITY OF PAIN 564
- *8.24.7.1 Rating Scales* 564
- *8.24.7.2 McGill Pain Questionnaire* 565
- *8.24.7.3 Assessment of Mood in Pain Patients* 565

8.24.8 PSYCHOLOGICAL INTERVENTIONS FOR PAIN 566
- *8.24.8.1 Contingency Management* 566
- *8.24.8.2 Multimethod: Developments Since Fordyce's Work* 567

8.24.9 COGNITIVE COPING METHODS 568
- *8.24.9.1 Altering Beliefs* 569

8.24.10 EARLY INTERVENTION 569

8.24.11 CONCLUSION 570

8.24.12 REFERENCES 571

8.24.1 INTRODUCTION

Pain is the primary symptom for which people seek medical treatment, and estimates suggest that over 70 million (80%) of all consultations to physicians in the USA occur because of pain (National Center for Health Statistics, 1986). In addition, in the USA there are over 50 million trauma injuries, many associated with high levels of pain (Chapman & Turner, 1986) and a further 15 million people being diagnosed with cancer with moderate to severe pain. Of these, 40–45% of patients experience such pain in the very early stages and 60–80% in the advanced stages of disease (Bonica, 1979). In addition to pain associated with disease or injury, medical and surgical procedures often produce pain. For example, in 1980 in USA, 23 million surgical procedures were conducted with more than 70% of the individuals undergoing these procedures reporting moderate to severe postoperative pain (Bonica & Benedetti, 1980). It is also important to note that any statistics such as those quoted above will only represent the smallest proportion of pain in the population, since most people endure quite large amounts of pain without seeking medical help (Von Korff, Dworkin, LeResche, & Kruger, 1988). For example, an epidemiological study of pain in adults in Seattle reported a prevalence of recurrent episodes of pain of 37%, with 8% reporting severe persistent pain. Pain is, therefore, a common symptom and experience, and its management presents a tremendous challenge to society, particularly to doctors and other health professionals. The impact of pain is not restricted to the health professions alone, however, since the size of the pain problem, particularly the chronic pain problem, affects all aspects of economic and professional activity. Most investigation of the economic impact of pain has focused on the back pain problem.

Back pain alone costs the UK National Health Service (NHS) £480 million per annum, and the cost has risen over recent years. In the USA $60 billion per year are spent on chronic pain problems (Bonica, 1980). The increasing size of the back pain problem led to the commissioning, in the UK, of the Clinical Standards Advisory Group (CSAG) report which attempts to identify the size of the problem and brings together some of the basic costs and existing evidence about effective approaches to management. The CSAG report (CSAG, 1994) estimates a population prevalence for back pain of 16.5 million, resulting in 2.4 million outpatient attendances and 12 million GP consultations. Much of the cost of back pain is contributed to by the relatively small proportion of back pain patients who continue to have back pain after about three months. A significant proportion of these patients will still be disabled by back pain at one year and the management of this long-term chronic group has attracted widespread interest and concern. It should be remembered that these figures represent the impact of back pain alone and, although this pain commonly represents approximately 60% of a pain clinic population, it is by no means the whole problem and the full cost of chronic pain is likely to be substantially greater.

Chronic pain has been distinguished from acute pain on the basis of two main variables, function and duration. Definitions based on function highlight the role of acute pain as a warning signal. The value of chronic pain is less clear, although it may serve to promote the inactivity which is sometimes necessary for successful recovery from serious injury (Bonica, 1977; Wall, 1979). There is little empirical support for a functional distinction between acute and chronic pain, and researchers and clinicians have usually based their working definitions of chronicity on the duration of pain. Pain is usually considered chronic if it persists for longer than six months (Sternbach, 1974). This cutoff of six months has, however, been rather arbitrarily determined, and recent studies looking at the prognosis for acute back pain suggest that three months may be a more appropriate cutoff (The Classification of Chronic Pain, 1986; Volinn, van Koevering, & Loeser, 1991).

8.24.2 WHAT IS PAIN?

There have been numerous attempts to define pain and a single, generally accepted, definition has emerged. Melzack and Wall (1988) defined it as an aversive, personal, subjective experience, influenced by cultural learning, the meaning of the situation, attention, and other psychological variables, which disrupts ongoing behavior and motivates the individual to attempt to stop the pain. This multidimensional view of pain is relatively recent and earlier views considered pain as a sensation. Early theories of pain perception, most particularly the specificity theory (Mountcastle, 1974) and the pattern theory (Crue & Carregal, 1975) view pain as primarily a sensory experience, resulting from activity in the peripheral pain receptors (nociceptors). These sensory theories view the experience of pain as a result of stimulation of specific nociceptors that send afferent stimulation to the brain via the pain pathways. This input to the brain, leading to stimulation of pain specific areas of the sensory cortex, was thought to lead to conscious awareness of pain. The intensity of the pain experienced was thought to be closely related to the degree of stimulation of the pain receptors and hence the extent of activity in the pain pathways. This direct line transmission model of pain is now thought to be oversimplistic and Melzack and Wall (1965) brought together a wealth of physiological, anatomical, and psychological evidence which shows that the experience of pain is not directly related to the extent of tissue damage or the level of activity in the pain receptors. As far as psychological factors are concerned, processes such as attention, or the individual's sense of control over the pain stimulus can intervene to determine the extent of the pain and distress associated with a given level of tissue damage. In order to accommodate these findings, Melzack and Wall proposed the "gate control theory" of pain in 1965. This essentially provided an anatomical basis for the view that sensory input could be modulated by certain physiological and psychological processes. Input from the peripheral nociceptors was thought to pass through a neural gate in the spinal cord before being transmitted to the brain. This gate may be opened or closed according to the relative activity in small and large afferent fibers, for example touch receptors, as well as the activity in fibers descending from central cortical areas, hence providing a mechanism by which psychological factors can exert an influence over pain perception. The theory is described at length in a number of places (Melzack, 1973; Melzack & Wall, 1988; Wall, 1976) and will not be discussed in detail here. The basic concept of input modulation is now widely accepted and the gate control theory is viewed as one of the most important theoretical advances in recent pain research. It is now clear that any effective model of pain must be multidimensional and provide for the possibility that physical and psychological aspects of pain may not be perfectly related. Examples of the range of psychological factors that influence pain perception are outlined below.

8.24.3 PSYCHOLOGICAL INFLUENCES ON PAIN

The range of psychological factors that have been shown to influence the level of pain intensity and distress associated with a given level of tissue damage include the perception of control, mood, personality, social, and environmental factors.

8.24.3.1 Perception of Control

Fiske and Taylor (1984) make a distinction between cognitive control, such as distraction or dissociation, and behavioral control where a subject is able to do something to reduce the pain or remove the pain stimulus. For example, evidence that subjects' control over the pain stimulus influenced their experience emerged when Bowers (1968) showed that subjects endure more pain when they are able to turn off the aversive stimulus, even though they often elected not to do so. Presumably the knowledge that they were able to remove themselves from the aversive stimulus enabled them to cope with the pain. Others have confirmed this link between pain tolerance and the knowledge that the pain can be controlled. An important practical example of behavioral control of pain is the development of patient controlled analgesia (PCA) in the management of postoperative pain. In PCA, patients self-administer postoperative analgesic via an easily controlled pump injection system. Mackie, Coda, and Hill (1991) showed that patients take less analgesic morphine from a PCA machine than they do if it is administered by a nurse or obtained through continuous infusion. A major report on pain after surgery produced in 1990 by the Royal Colleges of Surgeons and Anaesthetists concluded that PCA results in better pain management, reduced use of opiates, and earlier discharge compared with intramuscular therapy.

8.24.3.2 Mood

The interrelationship between depression and pain perception has been the subject of a variety of studies (Gupta, 1986; Pilowsky, 1988; Romano & Turner, 1985; Roy, Thomas, & Matas, 1984). There is still, however, considerable debate over the direction and nature of the relationship. A few studies have suggested that depression may be a cause of chronic pain, where pain is used as a method of expressing or masking of depression (Beutler, Engle, Oro-Beutler, Daldrup, & Meredith, 1986; Blumer & Heilbronn, 1982). Other studies suggest that pain is more likely to be the cause of the depression and that successful treatment of the pain leads to an improvement in mood (Atkinson, Slater, Patterson, Grant, & Garfin, 1991; Gupta, 1986; Romano & Turner, 1985; Roy, 1985; Widmer & Cadoret, 1978). Waddell and Turk (1992) suggest that it is more useful to regard both depression and anxiety as forms of distress. In their study of the relationship between anxiety and acute and chronic pain, Fordyce and Steger (1979) suggested the relationship for the two types of pain is quite different. Anxiety is experienced by acute pain sufferers but passes with the successful treatment of the pain, whereas in chronic pain, because the treatments often have little effect, patients become more anxious, which can in turn increase pain.

8.24.3.3 Personality

In 1959 Engel introduced the concept of a pain-prone personality. This suggested that certain personality types are more likely to have repeated episodes of chronic pain. He characterized the pain-prone personality by some specific traits, including: "excessive guilt, and a predisposition to use pain to atone for guilt, a strong unfulfilled aggressive drive, a history of suffering and defeat, intolerance of success, the use of pain as a replacement for loss when a relationship is threatened, and a tendency to masochistic sexual development" (Horn & Munafo, 1997, p. 82). Although there have been attempts to identify the pain-prone personality there is little empirical evidence to support the model (Horn & Munafo, 1997; Turk & Melzack, 1992). Indeed Sternbach and Timmermans (1975) suggest that so-called personality differences may be a result of the long term experience of pain as much as a cause and that personality difficulties are reduced by the successful treatment of pain. Studies of experimentally induced pain suggest that the personality dimensions of extroversion–introversion and neurotism may affect pain experience. Barnes (1975) showed that extroverts have higher levels of pain tolerance than introverts and Eysenck (1961), in a study of labor pain, found a relationship between extraversion and pain behavior. The practical implications of these relationships between pain and personality are unclear because the opportunities for changing personality traits are limited.

8.24.3.4 Social and Environmental Factors

Craig (1986) outlined a number of studies that demonstrated that the behavior of other individuals can influence both subjective and behavioral components of pain. They showed that the presence of a pain-tolerant model significantly increased the pain tolerance of naive subjects exposed to experimentally induced pain. Helman (1990) extended this to outline the role of social learning in pain and concluded that different social and cultural groups respond to pain in different ways depending on their cultural background and learning history. Perhaps the most marked examples of social influences on pain experience are provided by the Indian hook-swinging ceremonies, where in an annual ritual a person (celebrant) is selected to represent the power of the gods. The celebrant has steel hooks inserted under the skin and muscles on his back. Ropes are then used to attach him to a specially adapted cart. They travel to each village where the celebrant blesses the children and crops. After the blessing he swings free on the hooks. The chosen man shows no sign of pain and appears to be in a state of exaltation (Melzack & Wall, 1988).

In clinical practice the implications of the social environment of the ward or clinic setting on patient's experience of pain has received very little attention. For example, very little is known about the impact on a postsurgical patient of being placed in a ward where other patients may be demonstrating significant pain behavior. Craig's (1978) studies on experimentally induced pain would suggest that the patient's pain experience would be increased in the presence of other patients in pain.

8.24.4 CURRENT MODELS OF PAIN

For this reason recent models of pain use paradigms more akin to current models of emotion than sensation (Fordyce, 1986; Leventhal, 1984). Fordyce, for example, suggests that pain is a multidimensional experience involving affective and behavioral elements as well as the sensory experience. The relationship between these different components of pain is not linear and, depending on external environmental and internal factors, an individual may perceive a given level of tissue damage as either very painful or relatively painless. This is proposed for all pain, both acute and chronic, but Fordyce argues that, as pain persists, the relative impact of the nociceptive process becomes less important in influencing suffering and pain behaviors while other variables become more predominant. Fordyce, Brockway, Bergman, and Spengler (1986) has argued that this explains the failure of conventional medical and surgical interventions to reduce chronic pain.

Fordyce (1978) suggests that activity in the physiological system is perceived at a subjective level both as a sensation (e.g., burning or pricking) and as an unpleasant affect (e.g., distress or discomfort). He further suggests that this negative affect motivates pain behaviors (wincing, taking pills), which are aimed at communicating the pain experience to those around and at reducing the intensity of the pain experience. This view has become known as the three-systems (physiological, subjective, and behavioral) view of pain.

The further away from the level of nociception, the more other factors, besides the level of physiological activity, are likely to become implicated in the pain experience. For example, social and cultural variables are likely to influence pain behaviors in chronic pain, perhaps as much as activity in the nociceptive fibers. In Fordyce's view the behavioral consequences of pain will be influenced by environmental contingencies and maintained by operant learning processes, irrespective of the level of the underlying pathology, and he has developed a treatment program based on operant principles, which is discussed later.

Karoly (1985) has extended the three-systems model to include two additional factors: (i) lifestyle impact, for example, marital distress or vocational change; and (ii) information processing or central control, for example, coping styles, problem-solving skills, and health beliefs.

Philips (1987) outlines a model that considers the nature of the relationship between two of Karoly's response levels. It describes the influence of cognitions on pain behavior. Philips places emphasis on patients' expectations about the effects of their actions and their beliefs about their capacity to control pain episodes. She suggests that if self-efficacy is low, patients avoid activities that might provoke an increase in pain. In this way pain behaviors are viewed as avoidance behaviors that have the effect of maintaining the pain problem. The therapeutic implication of her model is that strategies that increase the patients self-efficacy will lead to increases in behavior and improved quality of life.

A model which possibly has greater predictive power is that of Leventhal (1984), which uses very different concepts drawn from an information-processing view of emotion. In common with more recent models of pain, he proposes that the notion of a linear relationship between noxious stimulation and pain intensity must be discarded. Instead he views pain processing as occurring at three hierarchically organized levels. The first involves the activation by a painful stimulus of an innate set of expressive motor reactions to environmental stimuli. Even very young children show specific patterns of facial expressions and it is possible to discriminate between different experiences such as pain, fear and anger with high inter-rater reliability. The second is the automatic encoding in memory of the emotional experience. This leads to the gradual development of a schematic system that which provides a record of the stimulus factors and their associated affective and expressive motor responses. The third level involves conscious beliefs and attitudes, comprising a set of abstract rules about emotional episodes and associated voluntary responses that arise over time, a consequence of self-observation.

A major strength of Leventhal's model is that it suggests that the experience of pain may be stimulated by physiological activation (bottom-up processing) or by top-down processing. Leventhal views the experience of pain as being stimulated by the activation of pain schema either as a result of activity at a conscious conceptual level (thoughts, attitudes, or beliefs) or as a result of activity at the expressive motor level (nociception) or both. The nature and process of the experience are therefore the same whether the pain is initiated by primarily top-down or bottom-up processes. This is in marked contrast to conventional medical and lay views of pain that perpetuate a psychogenic/organic distinction by considering pain to be either real and caused by organic factors, or not real or imaginary and caused by psychological factors. Having explanations for so-called organic and psychogenic pain within the same conceptual framework is important, and helps redress the view inherent in both lay and medical views of pain that there are two kinds of pain. The history and validity of this distinction are briefly considered because it has marked clinical and practical implications.

8.24.4.1 Psychogenic or Organic Pain

Pain, for which no clear physical cause can be identified, has often been labeled psychogenic in origin. It is assumed that patients have some emotional, motivational, or personality problem as the primary cause of their pain behaviors (Fordyce, 1978). For example, Engel (1959) identified a subgroup of patients without detectable lesions who, he suggested, were characterized by excessive guilt feelings, intolerant of personal success, and had family histories featuring pain and aggression. However, there is no clear evidence that these factors play a causal role and indeed there is some suggestion that psychological disturbance may be the result, rather than the cause of chronic pain. For example, although pain patients show higher scores on the hysteria, depression and hypochondriasis scales of the Minnesota Multiphasic Personality Inventory (MMPI), these scores decrease significantly after successful treatment of the chronic pain problem (Sternbach, 1974; Sternbach & Timmermans, 1975).

The distinction between organic and psychogenic pain is therefore probably of little value, since it is now recognized that psychological factors play a significant part in all pain. Nevertheless, this deeply entrenched concept does offer medical specialists a convenient label for those patients with unidentified pathology whose pain remains intractable despite the best efforts of practitioners to cure it. There is a concern that some patients labeled psychogenic, and for whom organic disease has been apparently excluded, may indeed have some underlying physical condition that is as yet undetected or unidentified. Either way, the term does little to serve the interests of patients or practitioners, and for the purposes of this chapter no further distinction is made, since the methods described are of equal relevance to the assessment of pain, whether or not organic pathology is identified.

8.24.5 PAIN ASSESSMENT

In view of the multidimensional nature of the experience of pain there are many aspects of both acute and chronic pain that clinicians and researchers may want to assess and the most appropriate methods chosen will be greatly influenced by the purpose for which the assessment is made. If the question to be addressed by the measurement of pain is "Does this drug X reduce the intensity of a particular aversive medical procedure, such as wisdom tooth extraction?" then the appropriate measures will be those that focus on assessing pain intensity or distress. If the question is "Is this multidisciplinary program for chronic pain effective?" then the range of dimensions that will need to be sampled may be much broader and include measures of pain behavior, lifestyle impact, social and marital functioning, as well as the more obvious intensity and affect measures. In practice there are three main issues that clinicians and researchers address in pain assessment:

(i) To determine the suitability of a patient for treatment.

(ii) To determine the individual patient's strengths and weaknesses so as to match or tailor a treatment program effectively.

(iii) To evaluate change during treatment and at follow-up periods.

In considering methods appropriate for each assessment question, the measures for the assessment of pain behaviors, pain intensity and distress, pain cognitions, and physiological aspects of pain are described separately. A full review of pain assessment methods can be found in Turk and Melzack (1992).

8.24.6 ASSESSMENT OF PAIN BEHAVIORS

The term "pain behavior" has been applied to a wide range of different activities. These include demands for medication, verbal complaints of pain, nonverbal responses such as gasping or grimacing, pain-relieving activities (e.g., rubbing or heating an affected location), and gross negative behaviors, such as not getting out of bed and avoiding activities that may exacerbate pain. The measures described in the following sections are aimed at measuring these specific pain behaviors, but it must be remembered that the term pain behaviors in the broadest sense includes Karoly's concept of lifestyle impact, which may require more complex assessment. Marital distress, social interactions, recreational activities, employment status, and domestic arrangements would be included in this category. Many aspects of such lifestyle impact are incorporated in the widely used Sickness Impact Profile (Bergner, Bobbitt, & Pollard, 1976), adapted for use with chronic back pain patients (Follick, Smith, & Ahern, 1985).

8.24.6.1 Direct Observation of Pain Behavior

Observer recordings of pain behavior may be performed in clinical settings as well as in more natural environments. They may be of particular benefit if the purpose of the assessment is to understand environmental influences on pain experience as part of a functional analysis of the pain problem. However, as in all direct observation the influence of the overall environment itself must be taken into account. Some studies use nurse observation in hospital settings. Ward assessments, however, raise the question of the generalizability of the measurements obtained. Aspects of the ward may exacerbate the demonstration of pain, either by means of social influences, such as other patients manifesting pain behaviors, or by nursing staff providing reinforcement of pain behaviors. Hence caution must be applied in generalizing findings from one setting to another. As a method of evaluating progress, however, during an inpatient program, direct observation on the ward may be very useful. Certain very specific pain behaviors are easily recorded by ward staff. For example, where nurses are the only source of drugs, medication intake can be easily recorded, as can requests for medication.

Well-behaviors, such as the number of sit-ups achieved in physiotherapy, can also be simply recorded. These are important measures of improvement and should be included among measures of outcome. Field assessment of the home and workplace is also currently being developed, not just as an aid to assessment but in conjunction with homebased management as part of treatment programs aimed to complement institutional management (Cott, Anchel, Goldberg, Fabich, & Parkinson, 1990).

Pain behaviors have been recorded by frequency counts during specified intervals. For example, Rybstein-Blinchik (1979) reports frequency counts of both verbal and nonverbal pain behaviors during 30 second observation periods of discussion initiated by the therapist. Inter-rater reliability was high, but the generalizability of the count was reduced by the therapist's presence. In addition, the period of time sampled was very short. This may have limited the value of the measure in terms of evaluating outcome. Repeated time-sampling over longer time periods is likely to provide more relevant information. The relative merits of different forms of observation, frequency counts, time sampling, and interval recording are described by Keefe and Williams (1992). No method is perfect, but Hartmann and Wood (1982) emphasize the need to train observers to a criterion level of reliability before starting data collection.

Other approaches to direct observation of pain behaviors have involved the use of rating scales. For example, ratings by nurses of the frequency of three different types of pain behavior observed during each shift. More detailed ratings of pain behavior are provided by Richards, who describes the development of a pain behavior scale. Nursing staff were asked to rate the frequency on three-point scales (none, occasional, frequent) of 10 verbal and nonverbal pain behaviors. The time to be sampled was a daily ward round, where the patient interacted with the therapy team in a semistructured interview lasting approximately five minutes. He reports high inter-rater reliability for this measure, and moderate correlations with other measures such as level of physical activity and behavioral improvement during the inpatient-program (Richards, Nepomuceno, Riles, & Suer, 1982). Again, the time period sampled limits the external validity of the measure. In addition, although the use of a three-point rating scale may make the instrument easy to apply, it is also likely to reduce its sensitivity to change.

Keefe et al. (1990) describes an observation method aimed specifically at evaluating change during treatment for patients with osteoarthritis in which trained observers keep detailed records of patients positions, level of movement and guarding, rubbing, shifting, rigidity, and joint flexing, which they show to have good reliability and validity and may be generalizable to other clinical populations.

Attempts at obtaining measures of pain behaviors in natural settings include asking spouses or peers of patients to record pain behavior (Fordyce, 1976). Unfortunately, a spouse may not be an entirely objective observer. A review of the problem of reactivity in observational measurement suggests that participant observers cannot be assumed to provide reliable data (Harris & Lahey, 1982).

8.24.6.2 Electromechanical Devices

The move towards objectivity in the assessment of pain behavior led to the development of a number of rather bizarre electromechanical devices, although methods favored in early studies are not much used today. Cairns, Thomas, Mooney, and Pace (1976) used an uptime clock placed above the patient's bed, which recorded the amount of time the patient spent lying down. Although they report this to be a useful measure of uptime, it is presumably dependent on the patients not being able to lie down anywhere else in the ward. Accuracy of recording also requires that patients can be trusted not to play the system by lying on their fellow ward mates' beds!

A slightly less problematic measure has been developed by Saunders. This is the pedometer and consists of a recording device worn on the patient's wrist, activated by a microswitch on the leg. It therefore assesses gross movements and gives estimates of walking distance (Saunders, Goldstein, & Stein, 1978).

8.24.6.3 Self-observation

Self-observation has the advantage of providing measurements obtained from the patient's natural environment. Furthermore, it gives responsibility for the evaluation of progress directly to the patient. However, it is dependent on patients recording their own behavior accurately, which is often difficult to ensure.

Care must be taken in designing self-recording schedules that maximize patient compliance (Collins & Thompson, 1979). Simple targets for self-recording, such as the number of specific exercises achieved and medication intake may provide more reliable information than very detailed measures (Melzack & Wall, 1989). It has been suggested that self-recordings should be used in conjunction with more objective behavioral measures, because discrepancies between the two types of measure may occur. Comparison of observer recordings of the activity of patients with self-monitoring data obtained from the patients themselves show considerable discrepancies. Similar discrepancies have been noted between self-monitoring data and uptime clock recordings (Sanders, 1983).

Such discrepancies may reflect some of the complex functional ramifications of pain and its meaning for the patient in a wider social context. For example, many patients, particularly at the start of treatment, may deny awareness of fluctuations in pain intensity in relation to environmental changes. Often the admission that social or interpersonal factors may exacerbate pain is seen as a statement that the pain is not real. Hence discrepancies between measures may provide important information for developing an understanding of the patients problem and should not necessarily be seen as inconvenient inconsistency.

8.24.7 MEASUREMENT OF THE SUBJECTIVE COMPONENT OF PAIN: MEASURING THE QUALITY AND INTENSITY OF PAIN

8.24.7.1 Rating Scales

Rating scales are the most commonly used measure of pain in clinical settings and pain research. They vary according to the number and nature of the anchor points supplied. Examples are numbers, words, or the presentation of visual analogue scales.

Verbal rating scales (VRSs) are usually 5 or 6 point scales, such as none–mild–moderate–severe–unbearable–intense (Gracely, McGrath, & Dubner, 1978). VRSs are normally scored by assigning each anchor a score as a function of its rank. The assumption, as with numerical rating scales, is that they can be considered interval scales, whereas in fact they are only ordinal scales. Cross-modality matching techniques have been used to demonstrate that considerable interindividual differences exist in the interpretation of verbal anchors, (mild to one person may mean considerably more pain than to his or her neighbor). Similarly, assumptions that patients within different diagnostic groups may use verbal rating scales in a similar fashion has been disputed using cross-modality matching tasks (Urban, Keefe, & France, 1984).

Visual analogue scales (VASs), such as 10 cm lines, suffer from some of the same difficulties although, as with VRSs, they have reasonable validity in that they show positive relationships with other measures of self-report (Jensen, Karoly, & Braver, 1986) and are sensitive to treatment effects (Turner, 1982).

Perhaps the easiest scale to administer is the numeral rating scale in which patients are asked to rate their pain from 0 to 10 or 0 to 100, where 0 indicates no pain and the maximum number represents pain as bad as it could be. These measures also demonstrate positive associations with other measures and are sensitive to change. Jensen et al. (1986) report higher levels of compliance than with VASs, probably because they are so simple to use. Jensen et al. also looked at the number of levels actually used on a rating scale. They looked at chronic pain patients' use of a 101 point rating scale in obtaining estimates of the patients least, most, current, and average pain. They also recoded the 101 point scales as first 11 and then 21 point scales and found that little information was lost if 101 scales were collapsed into these smaller number of categories. In addition, examination of the actual responses of patients on the 101 scale suggested that people treated the scale in terms of multiples of 5 or 10. They conclude that 21 or even 10 point scales provide sufficient levels of discrimination for chronic pain patients to describe their pain intensity.

Whatever rating scale is used, care must be taken to ensure that the anchors chosen do not confuse pain intensity descriptors with pain affect descriptors, because there is evidence that patient's ratings of pain intensity and pain affect are not always perfectly correlated (Gracely, Dubner, & McGrath, 1979). The importance of assessing both these components of pain separately is clear. If an outcome measure has anchors that could be ambiguous in their interpretation, unreliability of the measure may be introduced by the patient interpreting the question in one way on one occasion and in the other way at a later occasion of measurement. For example, if the scale asks "How bad is the pain?" with anchors from "not at all bad" to "extremely bad," the scale could be interpreted to mean either how bad is the pain intensity or how distressing is the pain. Used in different ways on different occasions this scale will become unreliable. Attempts to address this problem include the use of combinations of rating scales. Price, Harkins, and Baker (1987) used separate VAS ratings of the sensory and affective dimensions of pain. Previous research showed that VAS ratings approximated the properties of ratio scales (Price, McGrath, Raffi, & Buckingham, 1983) and were internally consistent (Price & Harkins, 1987). In a comparison of different patient groups, Price et al. (1987) found score profiles consistent with expectations, in that chronic pain patients produced higher affective ratings than either women in labor or volunteers exposed to experimental pain.

A second approach has involved using methods of cross-modality matching to validate the distinction between the two scales. Gracely (1980) describes the use of cross-modality matching techniques to develop separate ratio scales for the sensory (intensity) and the affective (distress) aspects. Pharmacological interventions were shown to have differential effects on the two scales. Although these findings concern acute experimentally induced pain rather than chronic pain, they nevertheless demonstrate that reliance on single scales may distort conclusions concerning changes in the subjective component of pain.

A further attempt to address this, and to develop related measures of intensity and affective aspects of the pain experience, is the Descriptor Differential scale of Gracely and Kwilosz (1988). This is a multiple item measure in which patients are asked to rate the intensity or affect of their pain as being more or less than the series of 12 descriptors, on a 21 point scale, (10 points either side of the verbal descriptor):

Intensity

(−) . . . . .strong . . . .(+)

Affect

(−) . . . . .distressing . . . .(+)

A strength of the scale is that the internal consistency of the scale can easily be assessed and it is quickly clear if a subject is using the scale in an unreliable/inconsistent fashion. Unfortunately, in most normal clinical practice, it may be rather time consuming to complete, and since there is, as yet, relatively little information about its properties, the relative sensitivity of the measure in relation to the standard VRS and VAS scales is not yet clear. However, it seems a valuable way forward in obtaining a more thorough measure of pain intensity and affect. A helpful development for the assessment of pain affect is the Pain Discomfort scale described by Jensen, Karoly, and Harris (1991). This is a 10 item questionnaire covering distress and anger responses to pain.

8.24.7.2 McGill Pain Questionnaire

The development of the McGill Pain Questionnaire (MPQ) by Melzack and Torgerson (1971) was an important move towards establishing a measure that assessed not just pain intensity but also the affective and motivational qualities of pain. Melzack and Torgerson (1971) started by asking subjects to classify pain adjectives and hence elicited classes of words describing the sensory qualities, the affective qualities, and the evaluative aspects of pain. They then asked groups of physicians, patients, and students to assign intensity values to each pain descriptor and thereby elicited scale values for individual descriptors in each category of pain.

The emerging MPQ consisted of 20 categories of words (10 sensory groups, five affective groups, one evaluative, and four miscellaneous). Details of the full questionnaire and its scoring procedures are described by Melzack (1975). It has been extremely widely used and a few examples of its application are provided by Melzack (1983), Reading (1989), and Wilke, Savedra, Holzener, Tesler, and Paul (1990). Evidence of its widespread use is the fact that it has been translated into 12 other languages, including Arabic and Chinese.

Three major indices of the MPQ are used to describe the overall pain problem:

(i) The pain rating index (PRI), based on the rank values of the words chosen by the patient to describe their pain.

(ii) The total number of words chosen.

(iii) The present pain intensity (PPI), based on an intensity rating 0–5.

The reliability of the adjective groupings has been confirmed by Prieto et al. (1980) and Reading, Everitt, and Sledmere (1982). In addition Turk, Rudy, and Saveloy (1985) have studied the internal structure of the questionnaire by using confirmatory factor analysis. Melzack and Katz (1992) and Torgerson (1986) discuss these findings in depth, together with the rationale for using factor analytic methods to delineate the dimensions of the MPQ.

The MPQ has been used to differentiate clinical pain and its value in this field is described by Melzack and Katz (1992). Although it is unlikely that any pain descriptor instrument would be used to make differential diagnoses, it is clear from comparisons across different clinical groups that patients with similar diseases or pain syndromes do tend to use similar words to describe their pain experience.

Melzack (1987) reported the development of the short form of the McGill (SFMPQ) which aims to be a quicker version, particularly relevant for research administration. It consists of 15 words (11 sensory, four affective) selected on the basis of their frequency of endorsement by patients with a variety of acute intermittent and chronic pains. Patients are asked to rate each adjective on an intensity scale 0 = more to 3 = severe. Melzack shows that the SFMPQ correlates very highly with the major PRI indices of the longer form and is sensitive to traditional therapies, suggesting it may be a valuable clinical and research tool.

8.24.7.3 Assessment of Mood in Pain Patients

Assessment of mood in pain patients can be problematic because many of the items of conventional anxiety and depression measures may be answered positively as a result of the pain itself, rather than the person's associated mood. The Beck Depression Inventory (BDI), for example, which is widely used in the assessment of depression, has a number of somatic items (sleep disturbance, fatigue) that may be caused by the pain itself rather than being a function of the associated mood. Hence scores on the BDI may be artificially inflated by this confounding. Indeed this has been borne out by a number of studies with both UK and US pain patients. Williams and Richardson (1993) conducted a factor analysis of the BDI scores in their UK chronic pain patient group and identified three factors: sadness about health, self-reproach, and somatic disturbance. The most frequently endorsed items were those loading on the somatic factor and Williams and Richardson suggest that the pattern of relationships between factor scores and measures of pain, mood and cognition, and physical functioning indicate that the use of the total BDI score may give a misleading impression of the supposed affective disturbance in this group of patients.

Similar conclusions are arrived at by the results of a factor analysis by Chibnall and Tait (1994) who conclude that the endorsement of somatic items means that the total BDI scores may not in fact indicate affective disturbance but rather be a consequence of the somatic aspects of pain.

Attempts to develop measures of mood that are less affected by contamination of endorsement of somatic items have included the Hospital Anxiety and Depression scale (Zigmund & Snaith, 1983) and more recently the Well-being Questionnaire (Pincus, Griffith, Isenberg, & Pearce, 1997).

8.24.8 PSYCHOLOGICAL INTERVENTIONS FOR PAIN

Since the late 1970s the increasing awareness of the multidimensional nature of pain has altered the way in which pain has been managed clinically and, particularly in the case of chronic pain, a range of different psychological interventions have been developed. These range from interventions that target a physiological system thought to play a causal role in the pain problem, (e.g., muscle biofeedback for tension headache, pulse volume biofeedback for migraine) through cognitive interventions aimed at teaching pain-coping skills, to operant behavioral measures that aim to modify reinforcing contingencies maintaining pain behavior. There is now considerable evidence that interventions that teach behavioral and cognitive pain coping strategies can be effective in reducing the life-style impact of chronic pain. Numerous reviews now exist (Flor, Fydrich, & Turk, 1992; Jensen, Turner, Romano, & Karoly; 1991; Linton, 1986; Pearce, 1983; Turner & Chapman, 1982a, 1982b; Williams et al., 1993). In clinical practice, there is an increasing number of psychology-led pain management programs on a range of different management strategies, including for example, relaxation training, cognitive skills training, and physical therapy. This chapter describes each of the interventions and considers the evidence for their efficacy separately. It then considers factors that may influence the overall success of the multimethod programs.

8.24.8.1 Contingency Management

The development of contingency management approaches for the treatment of chronic pain derives from Fordyce's view of chronic pain as primarily a behavioral problem. Fordyce considered that, in some cases, pain behaviors can be maintained largely by their reinforcing consequences, rather than by underlying physiological processes or anatomical change. The term pain behaviors is now commonly used to refer to both the learned and unlearned respondent and operant responses that contribute to and comprise chronic illness behaviors. Among the relevant and often discussed environmental variables that may strengthen illness behaviors are: (i) reinforcement of pain behaviors by significant others; (ii) avoidance behaviors in the form of restrictions in work, exercise, and even the most routine of daily activities; (iii) litigation and income protection plans (i.e., income contingent on disability); and (iv) unemployment or the failure to return to work following treatment or recovery. Pain behaviors may be either positively reinforced, for example by social factors, such as attention and concern from family members, or negatively reinforced, for example by avoidance of unwanted responsibilities. There is some experimental support for the underlying theoretical position that pain behaviors and activity levels in chronic pain patients can be controlled by altering reinforcement contingencies. McArthur, Cohen, Gottlieb, Naliboff, and Schandler (1987) review some of the literature dealing with the effect on activity levels of contingent social reinforcement. Others have looked at the influence of family members on chronic pain patients' behavior. For example, Romano et al. (1991) videotaped chronic pain patients and their spouses, as well as healthy controls, performing a series of routine household activities. They found that not only did chronic pain patients show higher levels of verbal and nonverbal pain behaviors, but their spouses showed more solicitous behavior and less facilitative behavior, for example, compliments and praise, than the spouses of control subjects. Although this does not prove that different patterns of spouse behavior play a central role in either causing or maintaining pain behaviors, it does provide support for the view that the environmental context in which chronic pain occurs may be different.

It is this environmental context that operant methods aim to change. The aim of operant programs is to increase the frequency of well behaviors and decrease that of pain behaviors. Reduction in pain intensity is not specified as an aim. Indeed, in clinical practice it is generally explained to the patient that the program is not aiming to remove their pain but rather to help them cope better with it and resume normal activities despite it. There is an underlying assumption that, once the patient is active again and engaged in more distracting activities, attention to the pain sensations will be reduced and hence pain intensity will diminish. The methods used to achieve these aims are described in detail by Fordyce (1976). Briefly, patients are admitted to specialized programs for periods of between two and six weeks. During this time staff give no attention to pain behaviors or requests for analgesics, but provide considerable social reinforcement for targeted well-behaviors. Physical therapy programs are developed to increase patient activity levels. Daily exercise quotas are determined on the basis of patients' initial tolerance levels. Each day the quota is increased so that the patient becomes able to manage significantly more than the initial level. Progress is charted graphically and staff attention, or some other desirable event, is made contingent on successful daily completion of the quota.

Reducing levels of medication is also an important aim of the operant programs. This is done by establishing initial drug requirements by putting the patient on a free operant schedule for few days immediately after admission. The amount of medication required is then provided in a strong-tasting masking vehicle called the pain cocktail. The patient is unable to tell from its appearance or taste how strong the concoction is. This pain cocktail is provided on a time contingent rather than pain contingent or PRN basis; the patient has to take it every four hours whether in pain or not. If it is refused, the patient must wait until the next time interval is up before medication is available again. In this way it is assumed that the proposed association between operant pain behavior and medication should be broken. The active ingredients of the pain cocktail are gradually reduced so that by discharge the patient should be consuming only the inert substance.

Typically, groups of patients are admitted to behavioral pain management programs, usually conducted in inpatient facilities and of short duration (e.g., four to six weeks). Such programs can produce dramatic results, but have been less successful in producing adequate maintenance on follow-up (Keefe & Gil, 1985). This may suggest that, although activity levels may be altered by environmental manipulation, any therapeutic approach based solely on this method is likely to have considerable problems in generalizing beyond the treatment environment.

8.24.8.2 Multimethod: Developments Since Fordyce's Work

Although Fordyce's original work revolutionized the approach to the management of longterm pain, relatively few recent psychological interventions have focused exclusively on operant methods. Although many use an operant framework they also include relaxation training, biofeedback, exercise therapy, and intensive education. These multiple ingredients make it difficult to identify the effective components of treatment and the process of change but there is now a clear body of evidence indicating that these multimethod interventions can produce significant improvements in quality of life. One of the early outcome studies is that of Seres and Newman (1976). They describe the progress of 100 patients in a treatment program incorporating operant conditioning along with physical therapy, body mechanics, biofeedback, relaxation training, and education. They reported positive results, showing increases in activity level as a result of the program. Later studies confirm these early positive findings. For example, McArthur et al. (1987) reviewed a large sample of chronic low back-pain patients from admission to a multidisciplinary program through to initial and long-term follow-up over a period of six months to five years following treatment. The therapeutic techniques used on the program included biofeedback, attribution-based psychotherapy, patient-controlled medication reduction, physical reconditioning, vocational counseling, educational presentations, pool and recreational therapy, and assertiveness training. Favorable outcomes were achieved by many of the respondents, and in several of the measures a good degree of stability was observed over the follow-up period, though obviously given the enormous number of components of treatment we have no idea of the effective ingredients

A total of 65 studies evaluating the efficacy of multidisciplinary, multimethod treatment centers were identified by Flor et al. (1992) and subjected to a meta-analysis. They report that both within- and between-group effect sizes show multidisciplinary treatments for chronic pain to be superior to no treatment, waiting list controls and conventional medical and surgical treatment or physiotherapy. Most of these studies have been conducted in the USA. One of the few UK studies that has drawn strongly from the operant method developed by Fordyce is that of Williams et al. (1993). This is an evaluation of a four week inpatient program delivered by a multidisciplinary team. During the admission, patients are managed in a largely operant culture, with reinforcement for well behaviors and pain-management behaviors, associated with relative lack of attention for pain behaviors. Graded exercise programs are taught but Williams et al. deviate from the strict Fordyce model by also teaching patients cognitive coping strategies. Although the study is not a randomized controlled trial, it does describe the progress of a large number of patients. Two hundred and twelve patients were assessed before and after treatment and at six months follow-up. Patients made significant improvements during treatment on measures of psychological and physical function and these improvements were largely maintained at six month follow-up. As with all treatment outcome studies they experienced problems in identifying all patients at follow-up. Patients who were unavailable for follow-up had higher pretreatment pain scores and poorer physical function, suggesting that the more severe patients may be vulnerable to dropping out and hence more difficult to keep in the treatment program. The approach is, therefore, probably of less value to them. Clearly in the absence of a control group one cannot be certain that the effects are due to treatment, but the results are in line with those reported by the USA studies (Linton, 1986; Malone & Strube, 1988; Phillips, 1987).

An Australian study also provides support for the efficacy of cognitive behavioral programs using an operant framework. Peters, Large, and Elkind (1992) compared inpatient pain management with an outpatient version of the treatment. The treatments are unfortunately not directly comparable because the inpatients had significantly greater contact time, whereas the outpatient program had less of an operant component and more emphasis on relaxation and coping skills. Therefore, it is not possible to draw conclusions of the relative benefits of inpatient vs. outpatient treatment. Nevertheless, both treatment groups showed significantly greater reductions in pain intensity ratings and pain behaviors than a no treatment control group, providing general support for the efficacy of such an operant-based approach.

The problems in comparing in- and outpatient treatment identified in the Peters et al. (1992) study raise the question of whether operant methods can realistically be applied on an outpatient basis. Controlling environmental contingencies may be viable on an inpatient basis, but is not easy in outpatient settings. Nevertheless, these basic environmental contingencies at home may be among the most important factors in maintaining the pain problem. There is increasing evidence that family members can influence pain experience and pain behavior and hence should be included in an operant program since they are powerful environmental reinforcing agents. Block, Kremer, and Gaylor (1980) showed that pain patient's ratings of their spouse as solicitous or nonsolicitous were correlated with patient's own pain ratings. Block suggested that the spouse was a discriminative cue for pain behavior and recommended spouse education to remove attention contingent upon pain display. Later studies have corroborated the importance of spouse reinforcement and social support (Flor, Kerns, & Turk, 1987; Paulsen & Altmaier, 1995; Romano et al., 1992). This would suggest that, to be effective, operant methods should more directly tackle the challenge of altering spouse behavior. Although some treatment programs involve significant others to some degree (Pearce & Erskine, 1989) there are very few studies that have extended operant methods into the home setting. Corey, Etlin, and Miller (1987) is an exception. They treated 72 patients by therapists going into the patients home for a six month period. Follow-up results at 18 months showed that significant lifestyle changes (including vocational status) were largely maintained, with nearly 70% identified as having achieved and maintained vocational goals. Although home based programs are likely to be difficult to organize and costly to implement, they do require further investigation because the longer-term cost effectiveness may outweigh the initial costs of the intervention.

8.24.9 COGNITIVE COPING METHODS

Attempts to reduce the impact of pain using cognitive manipulations have been used informally for generations. Adults attempt to distract children after accidents with magic kisses and encouraging them to imagine that they are brave little soldiers. In recent years psychologists have attempted to formalize and extend these phenomena. The term cognitive interventions refers to the growing number of techniques that attempt to influence pain through the medium of thoughts. These may include an individual's attentional processes, images, and self-statements. Fernadez and Turk (1989) summarized the literature and identified six categories of cognitive coping strategies:

(i) pleasant imaginings, for example, thinking of a bunch of flowers or a pleasant day in the country;

(ii) rhythmic cognitive activity, for example, counting backwards from 100 in threes;

(iii) external focus of attention, for example, counting ceiling tiles;

(iv) pain acknowledging, for example, reappraising the nociceptor stimuli objectively, such as concentrating on the dullness or numbness;

(v) dramatized coping, for example, imaging your are playing a heroic role in a football game or combatitive setting; and

(vi) neutral imaginings, for example, imagining being in a neither arousing nor unpleasant situation, such as watching television or attending a lecture.

In evaluating the efficacy of different cognitive coping strategies, Fernadez and Turk report a meta-analysis of 51 studies concluding that, overall, 85% of the investigations showed that cognitive strategies were effective in enhancing pain tolerance for experimental pain or attenuating clinical pain ratings. The relative efficacy of the different categories of coping is harder to elicit, since control subjects are not easily prevented from the spontaneous use of their own coping strategies. Fernadez and Turk conclude, however, that there is more evidence to support imagery strategies being more generally effective, and less evidence for the value of pain acknowledging. It should also be noted that many of the studies of specific coping strategies are conducted on normal subjects exposed to acute pain in laboratory conditions, hence care should be taken in extrapolating these findings to the clinical situation. In practice, because it is so difficult to predict which patients will find particular strategies most helpful, strategies are taught in a menu fashion and patients are encouraged to develop their own repertoire of chosen strategies.

8.24.9.1 Altering Beliefs

Chronic pain patients come to treatment with beliefs about their pain that have developed from their own experiences at attempting to manage it and their interactions with doctors and other health professionals. Not surprisingly, they tend to have negative views about the potential success of treatment and their ability to cope. Until the late 1980s there were few methods of assessing pain beliefs, and hence evaluating change in beliefs was difficult. We now have questionnaires by Williams and Thorn (1989), (the Pain Beliefs and Perception Inventory); by Jensen, Karoly, and Huger (1987), (Survey of Pain Attitudes); and by Edwards, Pearce, Turner-Stokes, and Jones (1992), (Pain Beliefs Questionnaire). Schwartz, Slater, and Birchler (1994) have also pioneered an assessment of pain beliefs that moves beyond questionnaire assessment and involves videotape presentation of common pain problems. The inclusion of such measures in the evaluation of treatment programs will provide important information on the process of change during interventions, as well as providing a method of assessing changes during the development of the pain problem from an acute to chronic condition.

Attention has been directed at people's self-efficacy beliefs and their role in predicting response to treatment. Council, Ahen, Follick, and Kline (1988) found that self-efficacy ratings predicted their performance on physical tasks (e.g., leg raises). Changes in self-efficacy may be an important process factor during multidisciplinary treatments and Jensen, Turner, and Romano (1991) report significant increases in self-efficacy as measured by the Pain Self-Efficacy Questionnaire during the treatment phase of their inpatient program.

8.24.10 EARLY INTERVENTION

Throughout the late 1980s and the 1990s, early intervention in the treatment of pain has been increasingly advocated. Although many patients with acute pain will naturally get better with time (Lindström et al., 1992), it is necessary to understand the development of chronicity to decide who would benefit from being treated at an early stage. Thus an important part of the research into the efficacy of early intervention is an understanding of the development and process of chronicity (Linton, Bradley, Jensen, Spangford, & Sundell, 1989). It is, however, difficult to identify individuals who are likely to become chronic pain sufferers. Psychosocial factors have been identified in some studies, although it is not clear if these factors are the cause of or are caused by the pain symptoms (Hellsing, Linton, & Kalvemark, 1994; Von Korff et al., 1990).

A few studies have attempted to predict who will become chronic pain sufferers, however, as yet the results of such studies are often conflicting. Philips, Grant, and Berkowitz (1991) in their randomized controlled study into the utility of behavioral counseling for back and neck pain patients ($n = 117$), found that risk factors for development of chronic pain problems included: main site being upper back or neck; acute pain from injury; and cases where there was a legal involvement. In the same study a model was developed to predict chronic pain using stepwise logistic regression. They used a series of measures as dependent variables including: pain intensity; pain quality (sensory, affective); pain behavior (avoidance, complaint); pain evaluation questionnaire; the BDI; the State Trait Anxiety Index; and the Sickness Impact Profile to predict patients likely to develop chronic pain. They measured pain at the acute stage, subchronic stage (three months), and chronic stage (six months). They were able to accurately predict subchronic from acute (75.5%), chronic from subchronic (80%), and chronic from acute (68.1%). This study, although useful, was limited in the number of patients sampled and the choice of dependent variables. Its authors suggested the need for further research.

In a pilot study of the evolution of chronic pain, Potter and Jones (1992) found that the presence of depression was significantly associated with the development of chronic pain. They then went on to suggest that if such depression was appropriately treated the likelihood of chronicity may be reduced. They also found a relationship between high pain intensity scores and chronicity. Hellsing et al. (1994) found, in patients with acute back and neck pain, that the development of chronic pain was not dependent on the severity or intensity of the acute pain. However, they found that those patients where active trunk movements provoked pain in several directions, or those with suspected nerve root involvement, had significantly higher risk of developing chronic pain. Lehmann, Spratt, & Lehmann (1993), in their study that attempted to predict long-term disability in back-injured workers, could not identify any predictive outcome measures for the development of chronic pain. They did, however, find a correlation with age and marital status and time until return to work. Coste et al. (1994) found that previous chronic low back pain resulted in a much lower chance of recovery. Other predictive factors included pain when standing or lying, initial disability, and employment and compensation status. This suggests that the development process of chronic pain is still not fully understood and further research is necessary to produce a consensus.

Several studies have been conducted to assess the advantages of early interventions (Hellsing et al., 1994; Jankus, Park, Van Keulen, & Weisensel, 1995; Linton, Hellsing, & Andersson, 1993; Mayer et al., 1987; Philips et al., 1991). This research has resulted in two key findings. First, the establishment of a critical time period within which treatment should be initiated, and second, that the advantages seem to be gained only with patients who have not been chronic pain suffers in the past.

The UK CSAG on back pain suggests that delayed access to treatment can contribute to create a chronic pain sufferer. Recommendations were given on the timescale for various treatments including, "all patients with acute back pain should be referred to and seen by a physical therapist before they are off work for six weeks" (CSAG, 1994, p. 36). This is because of clinical and epidemiological evidence that the duration of the initial episode of back pain is related to chronicity. Hackett, Bundred, Hutton, O'Brien, and Stanley (1993) found that open access to a physiotherapist, either at a health center or a hospital, results in fewer sick listed days than referral to a hospital consultant. Hellsing et al. (1994) also advocate early intervention for acute back and neck pain patients. They recommend that treatment should be completed within three months as this was a critical limit for developing new inactive habits.

In a study comparing standard treatment with an early active treatment program Linton et al. (1993) found that the advantages of early intervention were only evident in patients with no history of sick listing for musculoskeletal pain problems. The risk of developing chronic pain was eight times lower in the early activation no-history-of-pain group. This finding was reiterated by Hellsing et al. (1994) where the relative risk for developing chronic pain for patients with previous back/neck problems was six times higher than those not previously sick listed for musculoskeletal problems.

Philips et al. (1991) found that patients who received behavioral counseling early within the pain episode were more likely to be pain free at six months. The difference was not statistically significant but, as the intervention was limited to only one session of behavioral counseling on a relatively small sample, this might not be surprising. Philips et al. suggest that a larger study with a more significant intervention might produce more notable results.

8.24.11 CONCLUSION

This chapter provides an overview of some of the current theories of pain, methods for the assessment of pain, and psychological interventions for the management of chronic pain. It is clear that considerable advances have been made in the way chronic pain is conceptualized and treated. Whereas in the early 1980s patients with long term pain would only be able to expect physical interventions for their problem, now, especially in the USA but increasingly in other countries as well, local hospitals have some form of psychological or multidisciplinary service. These services are developing in the light of evidence that improvements in activity and quality of life can be effected by nonphysical, psychological, and multidisciplinary programs. However, although it is clear that these interventions produce change, the quality of the studies to date still leaves a number of questions about the process of change and the nature of the effective ingredients of psychological interventions. It is not yet known how to predict the outcome, or which kinds of patients do best under a particular form of psychological treatment.

Other areas that are still largely unknown concern the processes by which patients progress from acute to chronic pain. The next few years are likely to see interventions developing to prevent patients with acute pain moving along the continuum to chronic pain, and it is likely that the change in clinical services for chronic pain observed since the 1980s will be matched by increasing multidisciplinary interventions in primary care for the early chronic pain patient.

8.24.12 REFERENCES

Atkinson, J. H., Slater, M. A., Patterson, T. L., Grant, I., & Garfin, S. R. (1991). Prevalence, onset and risk of psychiatric disorder in men with chronic low back pan: A controlled study. *Pain, 45,* 111–122.

Barnes, G. E. (1975). Extraversion and pain, *British Journal of Social and Clinical Psychology, 14,* 303–308.

Bergner, M., Bobbitt, R. A., & Pollard, W. E. (1976). The Sickness Impact Profile: Validation of the health status measure. *Medical Care, 14,* 56–67.

Beutler, L. E., Engle, D., Oro-Beutler, M. E. Daldrup, R., & Meredith, K. (1986). Inability to express intense affect: A common link between depression and pain? *Journal of Consulting Clinical Psychology, 54,* 752–759.

Block, A. R., Kremer, E. F., & Gaylor, M. (1980). Behavioral treatment of chronic pain: Variables affecting treatment efficacy. *Pain, 8,* 367–375.

Blumer, D., & Heilbronn, M. (1982). Chronic pain as a variant of depressive disease: The pain prone disorder. *Journal of Nervous Mental Disorder, 170,* 381–406.

Bonica, J. J. (1977). Neurophysiologic and pathologic aspects of acute and chronic pain. *Archives of Surgery, 112,* 750–761.

Bonica, J. J. (1979). Cancer pain: Importance of the problem. In J. J. Bonica & V. Ventafridda (Eds.), *Advances in pain research and therapy* (Vol. 2, pp. 1–12). New York: Raven.

Bonica, J. J. (1980). Pain research and therapy: Past and current status and future needs. In L. K. Y. En & J. J. Bonica (Eds.), *Pain discomfort and humanitarian care* (pp. 1–46). Amsterdam: Elsevier.

Bonica, J. J., & Benedetti, C. (1980). Post-operative pain. In R. E. Condon & J. J. DeCosse (Eds.), *Surgical care: A physiological approach to clinical management* (pp. 394–415). Philadelphia: Febiger.

Bowers, K. S. (1968). Pain, anxiety and perceived control. *Journal of Consulting Clinical Psychology, 32,* 596–602.

Byrne et al (1980).

Cairns, D., Thomas, L., Mooney, V., & Pace, J. B. (1976). A comprehensive treatment approach to chronic low back pain. *Pain, 2,* 301–308.

Chapman, C. R., & Turner, J. A. (1986). Psychological control of acute pain in medical settings. *Journal of Pain and Symptom Management, 1,* 9–20.

Chibnall, J. T., & Tait, R. C. (1994). The short form of the Beck Depression Inventory: Validity issues with chronic pain patients. *Clinical Journal of Pain, 10,* 261–266.

Clinical Standards Advisory Group (1994). *Report of the Committee on Back Pain.* London: HMSO.

Collins, F. L., & Thompson, J. K. (1979). Reliability and standardization in the assessment of self-reported headache pain. *Journal of Behavioral Assessment, 1,* 73–86.

Corey, D. T., Etlin, D., & Miller, P. C. (1987). A home-based pain management and rehabilitation programme: An evaluation. *Pain, 29,* 219–230.

Coste, J., Delecoeuillerie, G., Cohen De Lara, A., Le Parc, J. M., & Paolaggi, J. B. (1994). Clinical course and prognostic factors in acute low back pain: An inception cohort study in primary care practice. *British Medical Journal, 308,* 577–580.

Cott, A., Anchel, H., Goldberg, W. M., Fabich, M., & Parkinson, W. (1990). Non-institutional treatment of chronic pain by field management: An outcome study with comparison group. *Pain, 40,* 183–194.

Council, J. R., Ahern, D. K., Follick, M. J., & Kline, C. L. (1988). Expectancies and functional impairment in chronic low back pain. *Pain, 33,* 323–331.

Craig, K. D. (1986). Social modeling influences on pain. In R. A. Sternbach (Ed.), *The psychology of pain* (pp. 67–96). New York: Raven.

Crue, B. L., & Carregal, E. J. A. (1975). Pain begins in the dorsal horn—with a proposed classification of the primary senses. In B. L. Crue (Ed.), *Pain: Research and treatment.* New York: Academic Press.

Edwards, L. C., Pearce, S. A., Turner-Stokes, L., & Jones, A. (1992). The Pain Beliefs Questionnaire: An investigation of beliefs in the causes and consequences of pain. *Pain, 51*(3), 267–272.

Engel, G. L. (1959). Psychogenic pain and the pain prone patient. *American Journal of Medicine, 26,* 899–918.

Eysenck, S. B. (1961). Personality and pain assessment in childbirth, of married and unmarried mothers. *Journal of Mental Sciences, 107,* 417–430.

Fernadez, E., & Turk, D. C. (1989). The utility of cognitive coping strategies for altering pain perception: A meta-analysis. *Pain, 38,* 123–136.

Fiske, S. T., & Taylor, S. E. (1984). *Social cognitions.* Wokingham, UK: Addison-Wesley.

Flor, H., Fydrich, T., & Turk, D. C. (1992). Efficacy of multidisciplinary pain treatment centers: A meta-analytic review. *Pain, 49,* 221–230.

Flor, H., Kerns, R. D., & Turk, D. C. (1987). The role of spouse reinforcement, perceived pain and activity levels of chronic pain patients. *Journal of Psychosomatic Research, 31,* 251–259.

Follick, M. J., Smith, T. W., & Ahern, D. K. (1985). The Sickness Impact Profile: A global measure of disability in chronic low back pain. *Pain, 21,* 67–76.

Fordyce, W. E. (1976). *Behavioural methods for chronic pain and illness.* St. Louis, MO: Mosby.

Fordyce, W. E. (1986). Learning processes in pain. In R. A. Sternbach (Ed.), *The psychology of pain* (pp. 49–72). New York: Raven.

Fordyce, W. E., Brockway, J. A., Bergman, J. A., & Spengler, D. (1986). Acute back pain: A control-group comparison of behavioural vs. traditional management methods. *Journal of Behavioural Medicine, 9,* 127–140.

Fordyce, W. E., & Steger, J. C. (1979). Chronic pain. In O. F. Pomerleau & J. P. Brady (Eds.), *Behavioral medicine: theory and practice.* Baltimore: Williams & Wilkins.

Gracely, R. H. (1980). Psychological assessment of human pain. In J. J. Bonica & D. Albe-Fessard (Eds.), *Advances in pain research and therapy* (Vol. 1, p. 818). New York: Raven.

Gracely, R. H., Dubner, R., & McGrath, P. A. (1979). Narcotic analgesia: Fentanyl reduces the intensity but not the unpleasantness of painful tooth pulp sensations. *Science, 203,* 1261–1263.

Gracely, R. H., & Kwilosz, D. M. (1988). The Descriptor Differential Scale: Applying psychophysical principles to clinical pain assessment. *Pain, 35,* 279–288.

Gracely, R. H., McGrath, P., & Dubner, R. (1978). Ratio scales of sensory and affective verbal pain descriptors. *Pain, 5,* 5–18.

Gupta, M. A. (1986). Is chronic pain a variant of depressive illness? *Canadian Journal of Psychiatry, 31,* 241–248.

Hackett, G., Bundred, P., Hutton, J., O'Brien, J., & Stanley, I. M. (1993). Management of joint and soft tissue injuries in three general practices: Value of on-site physiotherapy. *British Journal of General Practice, 43,* 61–64.

Harris, F. C., & Lahey, B. B. (1982). Recording system bias in direct observational methodology: A review and critical analysis of factors causing inaccurate coding behaviours. *Clinical Psychology Review, 2,* 539–557.

Hartmann, D. P., & Wood, D. D. (1982). Observational methods. In A. S. Bellack, M. Hersen, & A. E. Kazdin (Eds.), *International handbook of behavior modification and therapy* (pp. 109–138). New York: Plenum.

Hellsing, A. L., Linton, S. J., & Kalvemark, M. (1994). A prospective study of patients with acute back and neck pain in Sweden. *Physical Therapy, 74,* 116–124; discussion 125–128.

Helman, C. G. (1990). Pain and culture. In C. G. Helman (Ed.), *Culture, health and illness.* Oxford, UK: Butterworth-Heinemann.

Horn, S., & Munafo, M. (1997). *Pain, theory, research and intervention.* Buckingham, UK: Open University Press.

Jankus, W. R., Park, T. J., VanKeulen, M., & Weisensel, M. (1995). Interdisciplinary treatment of the injured worker with chronic pain: Long-term efficacy. *Wisconsin Medical Journal, 94,* 244–9.

Jensen, M. P., Karoly, P., & Braver, S. (1986). The measurement of clinical pain intensity: A comparison of six methods. *Pain, 27,* 117–126.

Jensen, M. P., Karoly, P., & Harris, P. (1991a). Assessing the affective component of chronic pain: Development of the Pain Discomfort Scale. *Journal of Psychosomatic Research, 35,* 149–154.

Jensen, M. P., Karoly, P., & Huger, R. (1987). The development and preliminary validation of an instrument to assess patients' attitude toward pain. *Journal of Psychomatic Research, 31,* 393–400.

Jensen, M. P., Turner, J. A., & Romano, J. M. (1991). Self efficacy and outcome expectancies relationships to chronic pain coping strategies and adjustment. *Pain, 44,* 263–269.

Jensen, M. P., Turner, J. A., Romano, J. M., & Karoly, P. (1991b). Coping with chronic pain: A review of the literature. *Pain, 47,* 249–283.

Karoly, P. (Ed.) (1985). *The assessment of pain: Concepts and procedures in measurement strategies in health psychology.* New York: Wiley.

Keefe, F. J., Caldwell, D. S., Williams, D. A. Gil, K. M., Mitchell, D., Robertson, C., Martinez, S., Nunley, J., Beckham, J. C., Crisson, J. E., & Helms, M. (1990). Pain coping skills training in the management of osteoarthritic knee pain: A comparative study. *Behaviour Therapy, 21,* 49–62.

Keefe, F. J., & Gil, K. M. (1985). Recent advances in the behavioural assessment and treatment of chronic pain. *Annals of Behavioural Medicine, 7,* 11–16.

Keefe, F. J., & Williams, D. A. (1992). Assessment of pain behaviors. In D. C. Turk & R. Melzack (Eds.), *Handbook of pain assessment* (pp. 277–292). New York: Guilford Press.

Lehmann, T. R., Spratt, K. F., & Lehmann, K. K. (1993). Predicting long-term disability in low back injured workers presenting to a spine consultant. *Spine, 18,* 1103–1112.

Leventhal, H. (1984). A perceptual motor theory of emotion. *Advances in Experimental Social Psychology, 17,* 117–183.

Lindström, I., Öhlund, C., Eek, C., Wallin, L., Peterson, L. E., Fordyce, W. E., & Nachemson, A. L. (1992). The effect of graded activity on patients with subacute low back pain: A randomised prospective clinical study with an operant-conditioning behavioural approach. *Physical Therapy, 72,* 279–293.

Linton, S. J. (1986). Behavioural remedication of chronic pain: A status report. *Pain, 24,* 125–141.

Linton, S., Bradley, L., Jensen, I., Spangford, E., & Sundell, L. (1989). The secondary prevention of low back pain: A controlled study with follow-up. *Pain, 36,* 197–207.

Linton, S. J., Hellsing, A. L., & Andersson, D. (1993). A controlled study of the effects of an early intervention on acute musculoskeletal pain problems. *Pain, 54,* 353–359.

Mackie, A. M., Coda, B. A., & Hill, H. F. (1991). Adolescents use patient-controlled analgesia effectively for relief from prolonged oropharyngeal mucositis pain. *Pain, 46,* 265–269.

Malone, M. D., & Strube, M. J. (1988). Meta-analysis of non-medical treatments for chronic pain. *Pain, 34,* 231–244.

Mayer, T. G., Gatchel, R. J., Mayer, H., Kishino, N., Keeley, J., & Mooney, V. (1987). A prospective 2 year study of functional restoration in industrial low back injury. An objective assessment procedure. *Journal of the American Medical Association, 258,* 1763–1767.

McArthur, D. L., Cohen, M. J., Gottlieb, H. J., Naliboff, B. D., & Schandler, S. L. (1987). Treating chronic low back pain: I. Admission to initial follow-up. II. Long-term follow-up. *Pain, 29,* 1–38.

Melzack, R. (1973). *The puzzle of pain.* New York: Basic Books.

Melzack, R. (1975). The McGill Pain Questionnaire: Major properties and scoring methods. *Pain, 1,* 277–299.

Melzack, R. (1983). *Pain measurement and assessment.* New York: Raven.

Melzack, R. (1987). The Short Form McGill Pain Questionnaire. *Pain, 30,* 191–197.

Melzack, R., & Katz, M. (1992). The McGill Pain Questionnaire: Appraisal and current status. In D. C. Turk & R. Melzack (Eds.), *The handbook of pain assessment* (pp. 152–168). New York: Guilford Press.

Melzack, R., & Torgerson, W. S. (1971). On the language of pain. *Anaesthesiology, 34,* 50–59

Melzack, R., & Wall, P. D. (1965). Pain mechanisms—a new theory. *Science, 150,* 971–979.

Melzack, R., & Wall, P. D. (1988). *The challenge of pain.* (2nd ed.). London: Penguin.

Melzack, R., & Wall, P. D. (1989). *Textbook of pain.* London: Churchill Livingstone.

Mountcastle, V. B. (1974). Pain and temperature sensibilities. In V. B. Mountcastle (Ed.), *Medical physiology.* St. Louis, MO: Mosby.

National Center for Health Statistics, & Koch, H. (1986). *The management of chronic pain in office-based ambulatory care: National Ambulatory Medical Care Survey. Advance data from Vital and Health Statistics,* No. 123 (DHHS Publication No. PHS 86-1250). Hyattsville, MD: Public Health Service.

Paulsen, J. S., & Altmaier, E. M. (1995). The effects of perceived verses enacted social support on the discriminative cue function of spouses for pain behaviors. *Pain, 60,* 103–110.

Pearce, S. (1983). A review of cognitive-behavioural methods for the treatment of chronic pain. *Journal of Psychosomatic Research, 27,* 431–440.

Pearce, S., & Erskine, A. (1989). Chronic pain. In S. Pearce & J. Wardle (Eds.), *The practice of behavioural medicine* (pp. 83–111). Oxford, UK: British Psychological Society/ Oxford University Press.

Peters, J., Large, R. G., & Elkind, G. (1992). Follow-up results from a randomised controlled trial evaluating in- and outpatient pain management programmes. *Pain, 50,* 41–50.

Philips, H. C. (1987). The effects of behavioural treatment on chronic pain. *Behaviour Research and Therapy, 25,* 365–377.

Philips, H. C., Grant, L., & Berkowitz, J. (1991). The prevention of chronic pain and disability: A preliminary investigation. *Behaviour Research and Therapy, 29,* 443–450.

Pilowsky, I. (1988). Affective disorders and pain. In R. Dubner, G. F. Gebhart, & M. R. Bond (Eds.), *Proceedings of the Vth World Congress on Pain* (Vol. 3, pp. 263–275). Amsterdam: Elsevier.

Pincus, T., Griffith, J., Isenberg, D., & Pearce, S., (1997). The Well-being Questionnaire: Testing the structure in groups with rheumatoid arthritis. *British Journal of Health Psychology, 2,* 167–174.

Potter, R. G., & Jones, J. M. (1992). The evolution of chronic pain among patients with musculoskeletal problems: A pilot study in primary care. *British Journal of General Practice, 42,* 462–446.

Price, D. D., & Harkins, S. W. (1987). The combined use of visual analogue scales and experimental pain in providing standardized assessment of clinical pain. *Clinical Journal of Pain, 3,* 1–8.

Price, D. D., Harkins, S. W., & Baker, C. (1987). Sensory affective relationship among different types of clinical and experimental pain. *Pain, 28,* 297–307.

Price, D. D, McGrath, P. N., Raffi, A., & Buckingham, B. (1983). Validation of visual analogue scales as ratio scales measures for chronic and experimental pain. *Pain, 17,* 45–56.

Prieto, E. J., Hopson, L., Bradley, L. A., Byrne, M., Geisinge, K. F., Midax, D., & Machisello, P. J. (1980). The language of low back pain: Factor structure of the McGill Pain Questionnaire. *Pain, 8,* 11–19.

Reading, A. E. (1989). Testing pain mechanisms in persons in pain. In R. Melzack & P. D. Wall (Eds.), *The Textbook of pain* (pp. 269–280). Edinburgh: Churchill Livingstone.

Reading, A. E., Everitt, B. S., & Sledmere, C. M. (1982). The McGill Pain Questionnaire: A replication of its construction. *British Journal of Clinical Psychology, 21,* 339–351.

Richards, J. S., Nepomuceno, C., Riles, M., & Suer, Z. (1982). Assessing pain behavior: The UAB Pain Behavior Scale. *Pain, 14,* 393–398.

Romano, J. M. & Turner, J. A. (1985). Chronic pain and depression: Does the evidence support a relationship. *Psychological Bulletin, 97,* 18–34.

Romano, J. M., Turner, J. A., Friedman, L. S., Bulcroft, R. A., Jensen, M. P., & Hops, H. (1991). Observational assessment of chronic pain patient–spouse behavioral interactions. *Behavior Therapy, 22,* 549–568.

Romano, J. M., Turner, J. A., Friedman, L. S., Bulcroft, R. A., Jensen, M. P., Hops, H., & Wright, S. F. (1992). Sequential analysis of chronic pain behaviours and spouse responses. *Journal of Consulting Clinical Psychology, 60,* 777–82.

Roy, R. (1985). Engel's pain-prone disorder patient: 25 years after. *Psychotherapy Psychosomatic, 43,* 126–135.

Roy, R., Thomas, M., & Matas, M. (1984). Chronic pain and depression: A review. *Comprehensive Psychiatry, 25,* 96–105.

Royal Colleges of Surgeons and Anaesthetists (1990, September). Commission on the provision of surgical services. Report of the Working Party on pain after surgery. London: Author.

Rybstein-Blinchik, E. (1979). Effects of different cognitive strategies on chronic pain experience. *Journal of Behavioural Medicine, 2,* 93–101.

Sanders, S. H. (1983). Component analysis of a behavioral treatment programme for chronic low-back pain. *Behavior Therapy, 14,* 697–705.

Saunders, K. J., Goldstein, M. K., & Stein, G. H. (1978). Automated measurement of patient activity on a hospital rehabilitation ward. *Archives of Physical Medicine and Rehabilitation, 59,* 255–257.

Schwartz, L., Slater, M. A., & Birchler, G. R. (1994). Interpersonal stress and pain behaviours in patients with chronic pain. *Journal of Consulting Clinical Psychology, 62,* 861–864.

Seres, J. L., & Newman, R. (1976). Results of treatment of chronic low back pain at the Portland Pain Centre. *Journal of Neurosurgery, 45,* 32–36.

Sternbach, R. A. (1974). *Pain patients: traits and treatments.* New York: Academic Press.

Sternbach, R. A., & Timmermans, G. (1975). Personality changes associated with reduction of Pain. *Pain, 1,* 177–181.

The classification of chronic pain. (1986). *Pain* (Suppl 3).

Torgerson, W. S. (1986). *Critical issues in verbal pain assessment: Multidimensional and multivariate issues.* Washington, DC: American Pain Society.

Turk, D. C. &, Melzack, R. (Eds.) (1992). *Handbook of pain assessment.* New York: Guilford Press.

Turk, D. C., Rudy, T. E., & Saveloy, P. (1985). The McGill Pain Questionnaire reconsidered: Confirming the factor structure and examining appropriate uses. *Pain, 21,* 385–397.

Turner, J. A. (1982). Comparison of group progressive-relaxation training and cognitive-behavioural group therapy for chronic low back pain. *Journal of Consulting and Clinical Psychology, 50,* 757–765.

Turner, J. A., & Chapman, C. R. (1982a). Psychological intervention for chronic pain: A critical review: I. Relaxation and training and biofeedback. *Pain, 12,* 1–21.

Turner, J. A., & Chapman, C. R. (1982b). Psychological intervention for chronic pain: A critical review: II. Operant conditioning, hypnosis and cognitive-behavioural therapy. *Pain, 12,* 23–46.

Urban, B. J., Keefe, F. J., & France, R. D. (1984). A study of psychophysical scaling in chronic pain patients. *Pain, 20,* 157–168.

Volinn, E., van Koevering, D., & Loeser, J. D. (1991). Back sprain in industry: The role of socioeconomic factors in chronicity. *Spine, 16,* 542–548.

Von Korff, M., Dworkin, S. F., & LeResche, L. (1990). Graded chronic pain status: An epidemiologic evaluation. *Pain, 40,* 279–291.

Von Korff, M., Dworkin, S. F., LeResche, L., & Kruger, A. (1988). An epidemiologic comparison of pain complaints. *Pain, 32,* 33–40.

Waddell, G., & Turk, D. C. (1992). Clinical assessment of low back pain. In D. C. Turk & R. Melzack (Eds.), *Handbook of pain assessment* (pp. 15–36). New York: Guilford Press.

Wall, P. D. (1976). Modulation of pain by non-painful events. In J. J. Bonica & D. Albe-Fessard (Eds.), *Advances in pain research and therapy* (Vol. 1.). New York: Raven.

Wall, P. W. (1979). On the relation of injury to pain. *Pain, 6,* 253–260.

Widmer, R., & Cadoret, R. (1978). Depression in primary care: Changes in patterns of patient visits and complaints during a developing depression. *Journal of Family Practice, 7,* 293–302.

Wilke, D. J., Savedra, M. C., Holzener, W. L., Tesler, M. D., & Paul, S. M. (1990). Use of the McGill Pain Questionnaire to measure pain: A meta-analysis. *Nursing Research, 39,* 36–41.

Williams, A. C., & Richardson, P. H. (1993). What does the BDI measure in chronic pain? *Pain, 55,* 259–266.

Williams, A. C., Nicholas, M. K., Richardson, P. H., Pither, C. E., Justins, D. M., Chamberlain, J. H., Harding, V. R., Ralphs, J. A., Jones, S. C., Dieudonne, I., Featherstone, J. D., Hodgson, D. R., Ridout, K. L., & Shannon, E. M. (1993). Evaluation of a cognitive behavioural programme for rehabilitating patients with chronic pain. *British Journal of General Practice, 43,* 513–518.

Williams, D. A., & Thorn, B. E. (1989). An empirical assessment of pain beliefs. *Pain, 36,* 351–358.

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scales. *Acta Psychiatrica Scandinavica, 67,* 361–370.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.25
Visible Disfigurement

NICHOLA RUMSEY
University of the West of England, Bristol, UK

8.25.1 INTRODUCTION 576
  8.25.1.1 Why is Physical Appearance Important? 576
  8.25.1.2 Psychological Research, Theory, and Practice in Disfigurement 577
8.25.2 THE INCIDENCE OF DISFIGUREMENT 577
8.25.3 THE CAUSES OF DISFIGUREMENT 577
  8.25.3.1 Congenital Malformations 577
  8.25.3.2 Visible Differences Resulting from Developmental Change 578
  8.25.3.3 Injuries Caused by Trauma and Burns 578
  8.25.3.4 Dermatological Conditions and Disease 578
  8.25.3.5 Disfigurements Resulting from Therapeutic Intervention 578
8.25.4 THE TREATMENT OF PEOPLE WITH DISFIGUREMENTS 579
  8.25.4.1 The Context of Treatment Provision 579
  8.25.4.2 Reconstructive and Aesthetic Plastic Surgery 579
  8.25.4.3 Prosthetic Devices 579
  8.25.4.4 Cosmetic Camouflage 580
8.25.5 PSYCHOLOGICAL ASPECTS OF TREATMENT 580
  8.25.5.1 Elective Surgery 580
    8.25.5.1.1 Motivation to seek treatment 580
    8.25.5.1.2 The outpatient experience 580
    8.25.5.1.3 The consultation 580
    8.25.5.1.4 Hospital admission and the postoperative experience 581
    8.25.5.1.5 Satisfaction with treatment outcome 581
  8.25.5.2 Emergency Treatment 581
8.25.6 THE IMPACT OF BEING VISIBLY DIFFERENT 581
  8.25.6.1 The Birth and Early Development of a Disfigured Child 582
  8.25.6.2 Childhood and Adolescence 582
  8.25.6.3 Common Problems Encountered by Adults 583
    8.25.6.3.1 Negative feelings about the self 583
    8.25.6.3.2 Problems with social encounters 583
    8.25.6.3.3 The behavior of the visibly different person 584
    8.25.6.3.4 Loss of identity through injury or treatment 584
8.25.7 INDIVIDUAL DIFFERENCES IN DISTRESS AND ADJUSTMENT 584
  8.25.7.1 The Etiology of the Disfigurement 584
  8.25.7.2 The Visibility and Extent of Disfiguration 585
  8.25.7.3 Age and Gender 585
  8.25.7.4 Socioeconomic Status and Educational Level 585
  8.25.7.5 Coping Style 585
  8.25.7.6 Social Support 586
  8.25.7.7 Levels of Social Skill 586
8.25.8 THE ROLE OF PSYCHOLOGISTS IN HEALTH CARE PROVISION 586
  8.25.8.1 Outpatient Care and Preoperative Assessment 586
    8.25.8.1.1 The clinic environment 587

8.25.8.1.2 Psychosocial assessment 587
8.25.8.1.3 Facilitating decision-making 587
8.25.8.1.4 Preparation for surgery 588
8.25.8.2 Emergency Admissions 588
8.25.8.3 The Postoperative Period 588
8.25.8.3.1 Discharge planning and preparation 588
8.25.8.3.2 Long-term follow-up 589
8.25.8.4 Emotional Support for Professional Carers 589
8.25.9 PSYCHOLOGICAL INTERVENTIONS AND THE DEVELOPMENTAL PROCESS 589
8.25.9.1 The Birth and Early Development of a Disfigured Child 589
8.25.9.2 The Family Environment 589
8.25.9.3 Social Interaction Skills 590
8.25.10 DISFIGUREMENT AND PSYCHOLOGY: PROFESSIONAL ISSUES 590
8.25.11 CONCLUSIONS 591
8.25.12 REFERENCES 591

8.25.1 INTRODUCTION

This chapter describes the current state of knowledge, research, and practice relating to the problems experienced by people with visible disfigurements.

Over the years, professionals working in this area have struggled to formulate a working definition which reflects appropriate inclusion and exclusion criteria. One particular area of debate centers on body dysmorphic disorder (BDD), in which the sufferer is considered to have grossly exaggerated negative feelings about an imagined or very minor physical defect. Although some of the problems reported by dysmorphophobics are similar to those with clearly visible disfigurements, other aspects of the condition (e.g., behavior patterns and response to treatment) set BDD apart. For the purpose of this chapter, findings related to research and practice concerning "people with one or more visible physical characteristic(s) perceived by the self and others as 'different' or 'abnormal' compared with societal and cultural norms" (adapted from Harris, 1997; Partridge, 1997) have been included.

8.25.1.1 Why is Physical Appearance Important?

Why do people who look "different" experience problems? Evidence for the strength and complexity of feelings associated with outward appearance comes from a rich variety of sources. These include mythology, legends, fairy tales, anecdotes from history, evidence from psychological research, and indications from contemporary society.

Whilst there is no compelling evidence for a relationship between personality and facial characteristics, the notion that character can be read from the face has an ancient and distinguished history. Subscribers to this view have included many great writers, including Cicero ("The face is the image of the soul"), Shakespeare ("There's no art to find the mind's construction in the face"), and Oscar Wilde ("It is only shallow people who do not judge by appearance!"). Folklore linking aspects of appearance to personality has existed throughout history and persists in current society, for example, linking red hair and fiery temper, a receding chin with weak character. In children's stories stereotypes concerning appearance abound. We are all familiar with the notion that the ugly frog cannot marry the princess until he is transformed into a handsome prince!

Evidence for the emphasis on physical appearance in society is plentiful. Millions of pounds are spent each year on cosmetic products and diet foods as people struggle to match up to the images promoted on television and in glossy magazines. Advertising constantly endorses the importance and benefits of physical attractiveness. Stereotypical type-casting is still rife in television and films. Characters with unprepossessing personalities and behavior are frequently portrayed as being ugly or in some way disfigured (Partridge, 1997). In Western society, language that reflects negative attitudes towards those who deviate from the perceived norm is used. Words such as "abnormality," "disfigured," and "flawed" convey the notion that there is something inherently "wrong" about the way a person looks. The care and treatment of people with disfigurements reinforces this notion by seeking to "normalize" appearance, thereby reinforcing the assumption that aesthetic improvement will enhance a person's quality of life (Rumsey, 1997).

The involvement of psychologists in the field of physical appearance would also seem to reflect the current biases in Western societies. In the USA and the UK, funding and interest in the effects of physical attractiveness abounded in the 1960s and 1970s, with the bulk of studies attesting to the advantages of being beautiful. The effects of ugliness and disfigurement have been much less fashionable, however, and this field has attracted very few devotees.

8.25.1.2 Psychological Research, Theory, and Practice in Disfigurement

As a consequence of the lack of involvement by researchers in this area, theory and research methodology are still at an early stage of development. The bulk of published research papers have emanated from the USA and the UK. They have focused on the prevalence of problems in relatively small samples of people with one type of disfigurement. The psychological effects of disfiguring burns and clefts of the lip and/or palate have received more attention than the consequences of port wine stains, various dermatological conditions, and disfigurements resulting from the removal of malignant tumors.

Methodological problems include small, often nonrepresentative samples, cross-sectional designs, the lack of standardized measuring instruments, and a lack of attention to individual differences in adjustment. Since the mid-1990s, a handful of researchers have begun to use models and concepts derived from health and social psychology. However, there is much capital still to be gained by the appropriate application of these frameworks to future research.

In view of the dearth of research studies and the lack of methodological sophistication, it is perhaps surprising that common themes emerge so strongly from the literature. Consensus exists amongst those in the field as to the most frequent problems experienced by people set apart from others by their visible difference, although knowledge of the factors underpinning individual differences in adjustment is still at the stage of informed speculation.

8.25.2 THE INCIDENCE OF DISFIGUREMENT

The precise incidence of disfigurement is impossible to calculate. There are many physical characteristics (such as height and weight) which can be objectively measured. Normality and abnormalities of appearance are more problematic, involving cultural norms, and social comparison (Harris, 1997). Although estimates of the incidence of various types of congenital deformities can be calculated with reasonable accuracy, figures for the total incidence of visibly disfigured people in Great Britain rely on statistics collected by the Office of Population Censuses and Surveys (OPCS). According to the OPCS, approximately 400 000 adults and children in Great Britain have a "severe disfigurement" defined as a "scar, blemish, or deformity which severely affects the ability to lead a normal life" (Martin, 1988; OPCS, 1988). The charity Changing Faces has used OPCS statistics to estimate that 10% of the population have some kind of disfigurement, with 2–3% having a difference which is clearly visible to others (Changing Faces, 1996).

8.25.3 THE CAUSES OF DISFIGUREMENT

Physical appearance is influenced by genetic inheritance, the physiological processes of growth, development and aging, accidental trauma, disease, and medical or surgical intervention. Thus, the causes of disfigurement are many and varied.

8.25.3.1 Congenital Malformations

Some disfigurements result from congenital malformations. The most common of these is a cleft (or gap) of the lip and/or palate, occurring about once in every 750 live births. Clefts of the lip can be unilateral or bilateral. Clefts occur twice as often in boys as in girls (Harris, 1997).

Other malformations of the head and neck are numerous, but rare. They can result from a failure of part of the face to develop (e.g., the cheek bones and upper jaw in Treacher Collins syndrome), the absence or malformation of an ear, eye, eyelid, or the nose, or the premature fusion of the suture lines which separate the bones of the skull (as in craniosynostoses such as Cruzon's and Apert's syndrome).

Some congenital syndromes involve additional difficulties, such as problems relating to feeding and drinking (as in cleft lip/palate), breathing, or sleeping (children with Apert's and Cruzon's syndromes have restricted nasal passages). In some infants, developmental delay and deficiencies in intellectual functioning are an added complication.

Birthmarks of various types are caused by malformation of blood vessels. Vascular malformations such as capillary "port wine stains," or venous haemangiomas, are present at birth and persist throughout life. They can occur in skin on any part of the body, but are most common on the face. Port wine stains are very noticeable due to their distinctive color, but they rarely lead to distortions of normal shape or contour. Cutaneous haemangiomata (strawberry nevi) result from delayed maturation of blood vessels. They usually appear about two weeks after birth, and initially resemble bright red patches. During the next six months, they grow rapidly into raised red, blue, and gray masses (frequently resembling bruising), gradually subsiding over subsequent years (Harris, 1997).

Malformations of limbs may be caused by a lack of development, either transversely (e.g., a missing hand or foot) or longitudinally (as in a "club" hand or foot). The most common malformation is syndactyly, or webbing, in which two or more fingers or toes fail to separate. Polydactyly is the presence of an extra digit.

8.25.3.2 Visible Differences Resulting from Developmental Change

There is considerable debate about whether or not changes resulting from the processes of growth and development, childbearing, the distribution of fat, and aging should be included in the field of disfigurement. Harris (1997) has argued that developmental changes may result in an appearance which is considered abnormal and disfiguring either by the person involved or by others. These changes in appearance may lead to stereotyping (or the fear of stereotyping), for example, drooping eyelids may lead a person to appear constantly tired, lowered corners of the mouth may suggest depression and misery, changes in skin texture and tone may lead to negative judgements of general health and well-being.

8.25.3.3 Injuries Caused by Trauma and Burns

Injuries may be caused by a wide variety of incidents, for example, industrial accidents, assaults, road traffic accidents, incidents in the home, and attacks by animals. Some injuries are self-inflicted, such as tattoos or piercing.

Scars associated with traumatic injury can be disfiguring when they result in alterations of shape, color, or contour, or if they are irregular or discolored (Harris, 1997). Facial disfigurements are particularly visible, and if they involve facial features or musculature, are likely to alter the person's repertoire of nonverbal behaviours.

In Great Britain each year there are about 40 000 injuries involving scalds and burns, with 25% of these requiring admission to a specialist burns unit. About half of these burns affect children under age 15 (Settle, 1994). Although fairly common, the majority of burn injuries are superficial, resulting only in some discoloration of the skin. More extensive burns result in considerable disfigurement. As part of the healing process, scar tissue contracts. These contractures can distort surrounding skin, leading to extensive deformity. Skin grafts, used to recover areas of skin loss, also contract, leaving an irregular surface in addition to a discolored donor site. As the appearance and texture of skin varies from one part of the body to another, grafted skin may not match the original surrounding skin, further contributing to the disfiguration (Changing Faces, 1996).

8.25.3.4 Dermatological Conditions and Disease

Dermatological conditions may result in changes in skin texture, unevenness, color, or coarseness of skin. Disfigurement can result from chronic skin conditions such as eczema, psoriasis (an incurable condition involving dry, flaking, and scaly patches of skin, occurring in 2–3% of the population), acne, rosacea (reddening of the skin), or from acne scarring. Neurofibromatosis results in patches of brown pigmented skin which can resemble freckles or can occur in large diffuse areas. In addition, sufferers can develop large numbers of very visible soft pinkish swellings or lumps. Vitiligo involves gradual depigmentation in the skin. The light patches become more extensive over time and can cover the whole body (Harris, 1997).

Facial paralysis occurs when the nerves that control sensation and movement in the face are damaged. This may be the result of Bell's palsy, the consequence of a stroke, the result of injury, or surgical intervention to the face (e.g., in the removal of a facial cyst or tumor).

8.25.3.5 Disfigurements Resulting from Therapeutic Intervention

Most operations leave a scar, the visibility and aesthetic quality of which depends on a number of factors. Harris (1997) outlined variables such as the direction of the scar in relation to lines of skin folds and skin tension, the part of the body operated on (scars on the face heal better than many other sites), the patient's characteristic disposition in relation to scar healing (scars may stretch in patients with overelastic skin, or stay thick and red in those with a tendency to form keloid scars) and the quality of wound healing (infections can contribute to poor-quality scarring). Scars will be more noticeable and disfiguring when they significantly alter the shape, tidiness, smoothness, and color of the affected area.

Extensive disfigurements may result from surgical removal of malignancies, for example, tumors in the head and neck region, lumpectomy or mastectomy in the case of breast cancer, or the removal of digits or limbs. The incidence of skin cancer is increasing rapidly (Changing

Faces, 1996), and the wide excision of skin followed by skin grafts (frequently leaving a dish-like deformity) is a growing cause of postoperative disfigurement. Harris (1997) has noted that patients with scars of the wrist or neck as the result of surgical intervention (as, for example, in a tracheotomy) may be misjudged as suicide attempts.

Other treatments may lead to temporary disfigurements. Pressure garments, worn to encourage the flattening of burns scars, are particularly visible. Chemotherapy may lead to the temporary loss of hair. Radiotherapy can cause discoloration of the skin and residual scarring.

8.25.4 THE TREATMENT OF PEOPLE WITH DISFIGUREMENTS

8.25.4.1 The Context of Treatment Provision

People with visible disfigurements may come into contact with the health services in a number of ways. Those born with congenital abnormalities will encounter health care professionals at an early stage following birth. Contacts may continue through a substantial part of the lifespan. Others acquire disfigurements as the result of trauma and will encounter the health service on an unplanned basis. In these cases, referral for plastic surgery is usually secondary to emergency treatment for life threatening injuries and functional deficits. Others will be referred as conditions develop (e.g., those with skin malignancies, tumors, or neurofibromatosis). Sufferers of intermittent conditions (e.g., eczema, psoriasis, acne) may be treated by their general practitioner, with occasional referral for specialist treatment when the symptoms flare up. Some, particularly in the private sector, will seek referral for plastic surgery procedures.

The services potential patients encounter will vary considerably, in terms of the professionals making up a care team, in the way care is delivered, and in relation to the funding of the units. Some patients may be referred to specialist, regional centres, staffed by multidisciplinary teams. Others may be treated in smaller, local, nonspecialist units.

Treatment possibilities vary on the type and location of the disfigurement. However, treatment is generally provided within the framework of the biomedical model. The aim of surgical intervention (generally reconstructive and aesthetic plastic surgery) is to normalize patients' actual (and perceived) abnormalities of appearance. It is assumed that improvements in appearance will result in decreases in psychological distress and associated emotional and behavioral dysfunction.

8.25.4.2 Reconstructive and Aesthetic Plastic Surgery

Reconstructive techniques frequently involve additional disfiguration during the course of treatment. Most reconstructive procedures employ the patient's own skin, fat, muscle, or bone taken from donor sites, which will in turn be scarred. In the past, many grafts were undertaken in stages to reduce the risk of rejection. One end of the "flap" was left *in situ*, while the other was attached to the donor site. The second end of the flap was repositioned only when the first stage of the graft had successfully taken. Although considerable advances have been made in grafting techniques, each graft carries a risk of failure.

Successful skin grafts may differ in color and texture from the original and surrounding skin. In order to overcome this problem, tissue expanders have been employed, in which a balloon-shaped expander has been inserted under the skin adjacent to the area to be reconstructed. Injections of saline are used to progressively expand the balloon and the surrounding skin. The expanded skin is then used to reconstruct the adjacent deficit (Harris, 1997).

With craniofacial defects and deformities, affected areas of the skull and facial skeleton are separated, then repositioned. Microplates are used to secure the repositioned sections of the skull and the affected facial structures.

Additional surgical techniques are employed in those who do not require reconstructive techniques. Facial features such as the nose, chin, jaw, and cheek bones can be altered through bone grafting and bone reduction. The surface of skin may be altered by dermabrasion, chemical peeling, or skin planing. Previous scars can be revised to be less conspicuous. Subcutaneous fat and superfluous skin can be removed. Laser therapy has become available and is used for the reduction and in some cases removal of port wine stains and other vascular disorders.

8.25.4.3 Prosthetic Devices

Prostheses are used for a range of purposes. A variety of implants have been developed to replace or augment breasts. In addition, prosthetic limbs are fitted following amputations or the loss of facial features such as eyes, ears, nose, cheek, and teeth. Developments in technology have led to great improvements in the appearance and tailoring of prostheses to an individual's requirements. The development of osseointegration now means that metal studs can be implanted into a bone allowing for much greater stability and security than was previously the case.

8.25.4.4 Cosmetic Camouflage

In many hospitals, cosmetic camouflage is provided by make-up specialists. They use cosmetic products which have been specially developed to provide adequate coverage of the affected area, to be waterproof, and resistant to accidental removal. These products are particularly useful for those with skin discoloration such as birthmarks and very visible scars.

Plastic surgeons employ a variety of techniques to treat people with disfigurements. Technological advances and the wider availability of plastic surgery are now widely advertised in magazines and other forms of the media. This increases the general acceptability of undergoing surgery to improve one's appearance and promotes the notion that visible variation is an abnormality requiring surgical intervention.

8.25.5 PSYCHOLOGICAL ASPECTS OF TREATMENT

Many factors will influence the impact that treatment has on a person with a visible difference. These include pretreatment variables and factors related to the eventual outcome.

8.25.5.1 Elective Surgery

8.25.5.1.1 Motivation to seek treatment

For the majority, decisions to seek surgical intervention to change appearance are influenced by a complex combination of variables, predominantly social and psychological in nature (Bull & Rumsey, 1988). A common motivator is the perception that a person's social life has been impaired by their appearance. As the result of problems in social interaction, many prospective patients experience high levels of self-consciousness and low self-esteem. The majority of people have been experiencing problems over a long period of time. Some have been encouraged or pressurized to seek treatment by family or friends, yet may personally feel ambivalent about the prospect.

Those with disfiguration resulting from congenital abnormalities (e.g., cleft lip/palate) will routinely be offered surgical and orthodontic intervention to improve their appearance. However, as particular stages of physical growth are critical to the successful outcome of many procedures, patients may feel frustrated at having to wait for a particular developmental stage before a procedure is undertaken.

In the UK, National Health Service waiting lists in plastic surgery can be extensive, especially for nonfunctional problems. Potential patients may experience distress at the low waiting list priority allocated to their case. In cost-cutting exercises, many health care trusts have excluded certain procedures (e.g., the removal of tattoos) from their routine provision of care.

8.25.5.1.2 The outpatient experience

Although "hard" evidence is lacking, indications from the personal accounts of people with disfigurements suggest that several factors conspire to make outpatient consultations less than ideal. Prospective patients may have to travel long distances to regional centers. In addition to the consultation with the surgeon and/or the treatment team, the patient may be required to undergo x rays, photographs, and preoperative tests, with each procedure involving additional waiting time. This can increase distress, particularly for children and their accompanying parents.

Many visibly disfigured patients feel uncomfortable waiting in busy, very public waiting areas. Pre-existing worries or fears may be exacerbated by seeing others who are at various stages of treatment.

8.25.5.1.3 The consultation

After months or years of waiting, consultations are likely to be brief. Aware of the time constraints faced by doctors, prospective patients (especially those in countries where health care is free at the point of delivery) feel under pressure to conclude the consultation quickly. Many patients feel that doctors are primarily concerned with any functional limitations associated with the disfigurement. They may underplay the extent of their concern and psychological distress.

At larger centers, multidisciplinary teams can comprise up to 20 people (Hearst & Middleton, 1997), including junior doctors, trainees, and other visitors. Patients can find the experience of a consultation with a team very daunting. Turner, Thomas, Dowell, Rumsey, and Sandy (1997) reported that the majority of patients and families attending multidisciplinary craniofacial clinics experienced difficulties expressing their opinions and asking questions about treatment. The feeling of being looked at and assessed by a large team of people is distressing for many. Patients may be confused by the necessary discussions between professionals and the jargon used in these discussions (Hearst & Middleton, 1997). Questions may be forgotten and sensitive issues avoided. Actual or perceived time pressures can act as further barriers to open discussion and effective decision-making concerning treatment.

8.25.5.1.4 Hospital admission and the postoperative experience

Following the decision to proceed with surgery, people with disfigurements will share the anxieties experienced by many patients surrounding admission to hospital and the immediate preoperative period. If their motivation for seeking treatment is largely psychosocial, they may fear negative evaluations by staff or fellow patients.

Coping with additional disfiguration in the immediate postoperative period may be particularly difficult for patients who have long awaited improvements in their appearance. Following surgical intervention, patients will initially look very much worse due to swelling and bruising before any long-term change can be detected. Pruzinsky (1988) has noted that during this time, patients often experience depression, anxiety, and doubts about the eventual outcome of the surgical procedure. Reconstructive surgery involving complicated (and often unsightly) grafting procedures can take many years to complete. Orthognathic surgery can necessitate the immobilization of the jaw postoperatively and the administration of a liquid diet. For those undergoing laser treatment for vascular disorders such as birthmarks, the treated area may be reddened and sensitive. Patients may be unable to apply normal camouflage make-up to the affected area for a number of months.

8.25.5.1.5 Satisfaction with treatment outcome

In line with treatment in many fields of health care delivery, the majority of patients report themselves to be satisfied following plastic surgery. Studies generally demonstrate decreases in depression, self-consciousness, and social anxiety, with increases in self-esteem (see, for example, Sheridan & Radmacher, 1992). However, correlations are not always observed between postoperative satisfaction and psychological improvement. Lovius et al. (1990) found patients who had undergone orthognathic surgery feared negative evaluation from others just as much after surgery as they had before, although improvements in body satisfaction were also reported. Rubinow, Peck, Squillace, and Gnatt (1987) wrote of continued problems with self-confidence and sensitivity to social rejection in patients whose acne had been substantially improved following treatment.

8.25.5.2 Emergency Treatment

Disfiguring injuries such as burns or head injuries will require emergency admission and treatment. The focus will initially be on stabilizing the patient, and subsequently on restoring function to damaged organs and limbs. In many cases, it will not be until after these issues have been addressed that residual disfigurement becomes an issue for the patient, the family, and the care team (Ambler, 1997). In the aftermath of dealing with the impact of the trauma, with unplanned admission to hospital and invasive, painful treatment, the realization that a person's former appearance will never be regained and the possibility of extensive permanent disfiguration result in an additional burden of stress for patients and their families.

8.25.6 THE IMPACT OF BEING VISIBLY DIFFERENT

In a society which places a high premium on physical attractiveness and "wholeness," it is not surprising that a considerable research literature reports the benefits of having a physically attractive appearance. When compared with the first impressions formed of those with an unattractive facial appearance, good-looking people are perceived to be more intelligent, popular, honest, and socially desirable (Bull & Rumsey, 1988). Research has suggested that in adults, attractiveness is positively related to expectations of future success, happiness, and satisfactory marital relationships.

Research evidence suggests that children from as young as five years of age stereotype on the basis of physical appearance. Studies have reported that subjects choose more attractive children as potential friends, expecting them to be more popular and more intelligent than their less attractive counterparts (see, for example, Tobiasen & Hiebert, 1988). Teachers have also been shown to rate attractive children as more intelligent and popular (for a review, see Bull & Rumsey, 1988).

Speculation about the origins of stereotypes and prejudices relating to physical appearance has focused on instinctive rejection of the less than perfect, a legacy of primitive beliefs, or processes of social conditioning and reinforcement based on ideals of attractiveness and the positive attributes that attractive people are assumed to possess and which, by implication, ugly or disfigured people supposedly lack. Research suggests that all of these explanations may play a part in responses to people on the basis of physical appearance (Rumsey, 1997).

Despite the lack of research specifically examining the effects of having an unattractive or visibly different facial appearance, there is considerable agreement concerning the common problems experienced by those with visible disfigurements (Macgregor, 1990).

8.25.6.1 The Birth and Early Development of a Disfigured Child

Advances in antenatal screening may mean that some parents are aware that their child will be born with a defect, but for the majority this comes as a complete surprise. The reaction to birth of a child with a defect has been likened to a bereavement reaction (Lansdown, Lloyd, & Hunter, 1991). Parents are likely to experience a variety of negative emotions, including grief, anxiety, guilt, and shock. They are faced with an immediate crisis, which may be compounded by associated problems such as difficulties with feeding (as in clefts of the lip and/or palate).

The handling of the birth is likely to have a major influence on the ability of the family to cope with the child's condition (Walters, 1997). The way the news is broken to parents has the potential to greatly affect their subsequent reaction. Staff may be as shocked as the parents, and may "freeze," feeling uncertain of how to deal with the situation. Subsequent discussion of the diagnosis, prognosis, and treatment possibilities is also crucial. Speculation about the causes of congenital malformations, for example, can cause many parents to relive the pregnancy and birth over and over again, agonizing about what they might have done to have caused the defect in their child (Walters, 1997). Uncertainties and delays are potentially damaging, as is the lack of opportunity for parents and the immediate family to ask questions.

The attachment of one or both parents to the infant may be adversely affected by the disfigurement. Less holding and less nurturant behavior have been observed by Barden, Ford, Jensen, Rogers-Sallyer, and Sallyer (1989) in mothers with craniofacially deformed children compared with those with normal appearance. A facial deformity may affect the normal movement of the baby's mouth and facial musculature, and may result in the caregiver misinterpreting the infant's emotional state. Smiling or other forms of interaction may also be affected, reducing the possibilities for reinforcement of caregivers (Walters, 1997).

Many children with congenital defects require early corrective surgery. Once an attachment has been formed, a common fear amongst parents is that the child will look, and therefore somehow be different postsurgery. The parents have become attached to their "disfigured" offspring and may fear losing the child they have learned to love. These anxieties are usually short-lived postoperatively.

8.25.6.2 Childhood and Adolescence

For children and adolescents, the majority of problems focus on teasing and bullying by others, on fear of going to new places, and problems associated with negative feelings about the self (Bradbury, 1996). Preferences shown by others for physically attractive friends may mean that social withdrawal is more frequent and avoidance of social situations can undermine confidence in future. Pertschuk and Whitaker (1982) have reported anxiety in disfigured children to be more common than in nondisfigured children. However, disfigurements in children do not result in any identifiable personality pattern, in reports of delinquency, or in raised incidence of bad behavior (Lansdown et al., 1991). The jury is still out in relation to lack of educational success. Some studies have reported that disfigured children achieve less academic and occupational status, whereas others have found no difference between visibly different children and their peers (see Nash, 1995, in relation to cleft affected children).

In adolescence, concerns about appearance can assume a new importance, and the onset of a disfigurement during adolescence is likely to be particularly distressing. Forming new relationships may be difficult if confidence in relationships has been steadily eroded (Walters, 1997). Adolescents with visible differences may lack the opportunities to try out new relationships and experiences (Sutherland, 1988). Leaving the familiar school environment can present difficulties. Disfigured adolescents and young adults may be disadvantaged in job interviews and prone to inaccurate stereotyping by potential employees.

The family environment is likely to be of crucial importance to the well-being of a child or adolescent who is visibly different. The attitudes, expectations, and support offered by the family will have an enormous influence on the offspring's feelings of self-worth and the way they cope with the disfigurement (Blakeney, Portman, & Rutan, 1990). In some families, parents and children may not discuss their difficulties. The parents may worry about upsetting the child, and may also be having difficulties coping with their own emotions of guilt, anger, anxiety, or helplessness (Bradbury, 1996). Children may feel protective of their parents and consequently will not want to upset them by openly admitting they are experiencing problems. Parents may become overprotective of the child. Ginsburg and Link (1989) wrote of a psoriasis patient who referred to his mother treating him as her "defective cub." The parents may find it hard to let go of their child, who in turn may find it difficult to leave the shelter of the family.

Siblings may also suffer as a result of the disfigurement. The family may have to direct financial and emotional resources towards the disfigured child, at the sibling's expense. Parental absence may occur during hospitalization of the disfigured child. Later on, siblings may feel ashamed of their disfigured brother or sister, and may be reluctant to invite friends home (Walters, 1997).

Despite the obvious problems encountered by children and adolescents with visible differences, some writers have questioned how they manage to be so "normal." Lansdown et al. (1991) and others, believe that problems in adolescence and early adulthood are underreported, as most data is collected in response to questions asked in medical settings, frequently in an atmosphere unconducive to open discussion of social and psychological issues. Many writers agree that life is likely to be appreciably harder for those with a face which is out of the ordinary. Interestingly, data from 20-year-olds who have completed treatment for clefts of the lip and/or palate suggest that for a variety of reasons, including not wishing to upset parents, disappoint clinicians, or even admit to themselves, they had downplayed and under-reported the difficulties they experienced during childhood and adolescence (Rumsey, Robinson, Turner, & Mercer, 1997).

8.25.6.3 Common Problems Encountered by Adults

8.25.6.3.1 Negative feelings about the self

Various negative emotions associated with visible disfigurements have been described in the literature, including negative self-image and low self-esteem. Many visibly different people feel shamed, embarrassed, or guilty about their appearance and report feeling in some way stigmatized or flawed. Salzer and Schallreuter (1995) reported that 75% of patients attending a vitiligo clinic found their disfigurement moderately or severely intolerable. Feelings of shame and embarrassment were reported in 89% of a sample of psoriasis patients (Ginsburg & Link, 1989).

8.25.6.3.2 Problems with social encounters

Situations involving encounters with others are frequently reported as problematic for visibly disfigured people (Macgregor, 1990). In studies of patients with vitiligo, Porter, Beuf, Lerner, and Norlund (1990) documented anxiety and embarrassment when meeting strangers. Difficulties are encountered in forming friendships (Lanigan & Cotterill, 1989) or beginning a new sexual relationship (Noar, 1991).

Problems are particularly apparent at times of transition or change, for example, moving to a new neighborhood, job seeking, or changing jobs. Anecdotal evidence suggests that people with visible differences may have more trouble than their nondisfigured counterparts obtaining the kind of employment they would like. Some studies of people affected by a cleft of the lip have suggested that they achieve a lower occupational status than their peers (Nash, 1995).

On the basis of her clinical experience, Bradbury (1996) suggested that lowered self-esteem and feelings of self-worth may result in disfigured people settling for inappropriate friendships and less-than-ideal partners in longer-term relationships. Other writers have suggested that fewer cleft affected people marry, and those that do have fewer children (see Nash, 1995 for a review).

Many allude to a lack of privacy resulting from staring, pointing, "double takes," hurtful comments, and intrusive, unsolicited questions about the disfigurement (Robinson, Rumsey, & Partridge, 1996). Revealing the disfigurement to others (e.g., in certain weather conditions such as wind and rain) or being seen without make-up can be stressful. Lanigan and Cotterill (1989) reported that 9% of a sample of women did not reveal their birthmarks to their husbands. Others with visible disfigurements describe their distress when posing for a photograph, laughing, smiling, eating, or drinking in public, or unexpectedly catching sight of their own disfigurement in a mirror or shop window (Robinson, 1996).

If the disfigurement appears to be contagious in some way (as in psoriasis and neurofibromatosis), sufferers are particularly sensitive to the fact that others may not wish to touch them. Dermatology patients often avoid activities where the skin might be "on show," such as swimming or holidaying in climates where minimal clothing is required (Jowett & Ryan, 1985).

Many disfigured people complain that they are avoided or rejected by the general public. Studies of the process of social interaction have shown that avoidance does indeed occur. In addition, those people who do engage in interaction tend to stand further away from a disfigured person than from a nondisfigured one, and make less frequent eye contact during the interaction (Rumsey, Bull, & Gahagen, 1986).

Why does this avoidance occur? A number of theories have been put forward, all of which may play a part in the avoidance experienced by visibly different people. Bernstein (1986) has suggested that because of the strong association between health and good looks, a person with a disfigurement may be avoided as they are assumed to have something which is contagious. In the context of the "just world" hypothesis, Lerner (1970) proposed that people are assumed to have deserved their misfortune and, as discredited people, are therefore avoided.

Langer, Fiske, Taylor, and Chanowitz (1976) suggested that uncertainty about how to behave can lead to a desire to escape or avoid an encounter. There is no doubt that a first encounter, particularly with a facially disfigured person, is potentially problematic, as the normal rules of social interaction will not apply. The disfigured person's capacity to communicate may be impaired or altered. The other party in the interaction may be concerned that if they use conventional patterns of nonverbal communication, they may be misconstrued (for example, eye contact directed at the face might be interpreted as staring). Avoidance may seem preferable to the risk of embarrassment for either or both parties.

However, despite evidence that avoidance does occur, studies of prosocial behavior, particularly in situations where social interaction is unavoidable, suggest that some people respond more favorably to a disfigured person than to a nondisfigured person (Bull & Rumsey, 1988).

8.25.6.3.3 The behavior of the visibly different person

During the 1990s, attention has been diverted towards the effects on social interaction of the behavior of disfigured people themselves. Many visibly different people become preoccupied with their appearance and the effect it may have on others. Anticipating negative reactions from others, they may adopt a shy, defensive, or aggressive style of interaction (Macgregor, 1990). They may actively seek out signs that other people have noticed the disfigurement and that they feel negatively about it. Strenta and Kleck (1985) used a make-up artiste to apply scars to subjects' faces. Unknown to some of the subjects, the scar was then removed (subjects were led to believe that a fixative was being applied). Those who believed they were disfigured maintained the scar was being scrutinized by people with whom they subsequently interacted, and perceived more negative reactions from them. Such beliefs and resulting patterns of behavior will increase the likelihood of negative reactions from others. The notion of unfavorable perceptions by members of the public is therefore reinforced, and a cycle of spiraling problems may continue, including avoidance and social withdrawal by the disfigured person (Rumsey, Bull, & Gahagen, 1986).

8.25.6.3.4 Loss of identity through injury or treatment

Following a traumatic injury, or rather more paradoxically, changes to appearance as the result of treatment, there can be an acute sense of loss of identity. In these cases, people complain that they no longer recognize themselves, nor look like the "real me" (Grealy, 1994). Distress may also result from feelings of loss of a family feature or a physical characteristic felt to reflect racial identity. The reaction to such changes in appearance has been likened to a bereavement reaction, particularly when it has been unanticipated (Bradbury, 1996). Old patterns of social behavior and styles of coping may no longer work so well, leaving the person vulnerable to problems in interactions and relationships with others.

8.25.7 INDIVIDUAL DIFFERENCES IN DISTRESS AND ADJUSTMENT

Despite some commonality in the types of problems encountered by those with visible differences, there is wide variation in the degree of distress experienced. Little attention has been paid to identifying the factors which might predict individual differences in distress and adjustment.

8.25.7.1 The Etiology of the Disfigurement

The cause of a disfigurement is an obvious variable when considering the potential impact on an individual. Complicating factors include feelings of blame or anger, whether any compensation claims are involved, and the degree of stigmatization felt in relation to the etiology. Adjustment is also likely to be linked to prognosis. Some disfigurements are temporary (e.g., acne, some trauma injuries), others are recurring (as in some dermatological conditions such as eczema), progressive (as in neurofibromatosis), or linked to the treatment of life-threatening conditions such as cancer.

In noncongenital conditions, additional variables are likely to include pre-existing levels of attractiveness, social confidence and social skill, an individual's existing coping repertoire, levels of self-esteem and self-efficacy. Although the passage of time has been proposed as a great healer, the length of time a person has had a disfigurement is not a reliable predictor of distress or adjustment (Robinson, 1997).

8.25.7.2 The Visibility and Extent of Disfiguration

Many assume a direct relationship between the severity and extent of a disfigurement and the degree of distress experienced. However, one of the most consistent findings in the literature is the lack of such a relationship (see Robinson, 1997 for a review). Some writers suggest that the responses of others to major disfiguration are more consistent and predictable, whereas reactions to minor disfigurements are less easily predicted. Uncertainty as to the likely reactions of others may add to the difficulties encountered in social interactions.

Disfigurements which are more readily concealed are associated with less distress. Porter, Beuf, Lerner, and Norlund (1986) found less distress in those suffering from vitiligo, which can be concealed with camouflage make-up, compared with patients with psoriasis. Williams and Griffiths (1991) found that burn injuries which could be covered were reported as less distressing than those in more visible locations. As mentioned above, disfigurements affecting the face can be particularly problematic due to the heavy involvement of the face in social encounters, and the role of the face in the process of impression formation.

8.25.7.3 Age and Gender

The period of adolescence (E. Wallace, 1993) and the early twenties (Long & DeVault, 1990) appear to be times when psychological problems relating to disfigurement may be particularly severe. In addition, times of transition or change (e.g., changing schools, moving to a new neighborhood, changing jobs, new relationships) are widely acknowledged to create or to exacerbate existing problems.

Little research has focused directly on the effects of gender in relation to adjustment to visible disfigurement. In more general terms, it has been pointed out that in our society, a high value is placed on personal appearance, with a disproportionate amount of this pressure directed at women (Liggett & Liggett, 1989). Consequently, it may be that women are more prone to distress associated with visible disfigurement. However, women may also find camouflage make-up more acceptable than their male counterparts. Men may find it more difficult to admit to concerns about their appearance than women.

Although few researchers have found main effects for gender and distress, Robinson (1997) has suggested that gender is likely to interact with other factors such as the use of particular coping techniques, and the use of social support.

8.25.7.4 Socioeconomic Status and Educational Level

Few studies have examined the effects of intelligence, educational level, and socioeconomic status on adjustment. However, Lefebvre and Arndt (1988) have suggested that in children, intelligence is an important protective factor against maladjustment to facial disfigurement, as it is related to academic achievement, a sense of control, and competence. In Robinson's (1996) sample of adults with a variety of disfigurements, higher levels of educational achievement were found to be related to lower levels of anxiety and depression, fewer difficulties in social situations, and more confidence in the company of strangers.

During the last decade, attention has turned to psychological factors which may be related to adjustment in people with disfigurements.

8.25.7.5 Coping Style

In an examination of the coping styles employed by a broad sample of people with visible disfigurements, Robinson (1996) found little evidence of consistent coping styles. Instead, a wide variety of cognitive and behavioral coping strategies were used which varied according to the situation, and to a lesser extent on the prevalent mood.

Common strategies included avoiding social situations, avoidance of reflective surfaces, escapism (e.g., drinking alcohol, smoking, using drugs), and distraction techniques (undertaking study, doing lots of exercise, doing alternative activities to promote distraction). A minority used strategies designed to gain some control in social encounters, for example, returning a stare, smiling, trying to be friendly, or initiating an interaction with a handshake. Additional techniques included self-talk and social comparison ("it could be worse, I could have been badly burned instead"). Camouflage techniques, for example, make-up, clothes, sunglasses, and a particular hairstyle were widely used. Others described seeking out social support (this was particularly the case for subjects of higher socioeconomic status) and seeking further treatment.

8.25.7.6 Social Support

The evidence for the role of social support in adjustment is mixed, but largely supportive of the beneficial effects in a variety of subject populations (Robinson, 1996). Supportive families have been widely acknowledged to positively influence the psychological adjustment of children (see, for example, Blakeney et al., 1990). Orr, Reznikoff, and Smith (1989) examined the effect of social support upon the self-esteem, body image, and levels of depression in adolescent and young adult burn-injured patients. They found that perceived social support from friends was the most crucial factor in determining adjustment to disfigurement.

In a study of head and neck cancer survivors, Baker (1992) found that perceived social support was positively related to patients' adjustment six months post-treatment. Davidson, Bowden, Tholen, James, and Feller (1981) reported in a long-term study of burned patients that social support from family, friends, and peers was significantly related to self-esteem, life satisfaction, and participation in social and recreational activities. Robinson et al. (1996) found that in a sample of people with a variety of facial disfigurements, decreases in depression scores following social interaction skills workshops were more notable for those living with another person, compared with those living alone.

Moss (1997) has discussed the possibility that low social support might be a consequence as well as a cause of poor adjustment. He noted that Brewin, McCarthy, and Furnham (1989) argued that stigma produces anxiety, which can lead to social withdrawal. This withdrawal is likely to decrease the availability of social support. In the context of attributional theory, Brewin et al. proposed an example which is highly relevant to the situations experienced by those with visible differences. They suggested that when a person believes an event is likely to recur, when other people do not have to cope with a similar situation, and when similar stresses are likely to occur in a variety of circumstances, the perception that lower amounts of social support are available is likely to prevail.

Although the precise mechanisms of the effects of social support remain unclear, important variables are likely to include the level of social support available, the type of support, the quality of support, and the person's appraisal of the availability of support. Future research will need to explore why some people seem able to make better use of support than others, and how to help people maximize the benefits (Robinson, 1996).

8.25.7.7 Levels of Social Skill

In addition to the individual differences in social skill which exist in the population as a whole, several factors may affect the level of social skill employed by visibly different people in their encounters with others. These factors include commonly reported experiences such as preoccupation with appearance, self-consciousness, social anxiety, a lack of confidence concerning social interactions, or the anticipation of negative reactions from others. In addition to effective self-presentation and communication, good social skills involve the accurate interpretation of information from others. This information is likely to affect both the self-image and feelings of self-worth. As discussed above, the selective attention employed by some visibly different people to the more negative aspects of this feedback is likely to compound pre-existing difficulties in social interaction.

8.25.8 THE ROLE OF PSYCHOLOGISTS IN HEALTH CARE PROVISION

Within the health services, the provision of care to meet the psychosocial needs of people with disfigurements is patchy at best. It is widely acknowledged, particularly in the UK, to be inadequate (L. Wallace, 1988). The lack of specialized support is partly the result of resource constraints, but also reflects the prevailing biomedical model of care. Although interventions designed to improve appearance are undoubtedly important, they form only part of the whole picture. Pruzinsky (1988) and others have argued that psychological factors play a part in all stages of elective, reconstructive, and emergency surgery. Psychological research, assessment, and interventions would therefore be appropriate at all of these stages.

8.25.8.1 Outpatient Care and Preoperative Assessment

Visits to outpatient clinics may occur for a number of reasons. Some people seek an appointment with a view to surgical intervention to improve their appearance. Others (e.g., those with congenital or chronic dermatological conditions) attend for regular outpatient consultations. Although rarely in evidence, psychologists have a potential role to play in all aspects of outpatient care.

8.25.8.1.1 The clinic environment

With minor alterations, the clinic environment can provide an atmosphere and ethos more conducive to open discussion and effective decision making concerning treatment.

The waiting area should be considered, as patients' pre-existing worries and fears may be exacerbated by the presence of others who are at various stages of treatment. Long waiting times provide the opportunity for a preliminary chat with a team member in order to allay fears and to establish an agenda for the subsequent consultation.

If the patient is to be assessed by a multidisciplinary team, a system of individual appointments could be implemented, as many patients and their families report difficulties openly discussing their motivation and their concerns about treatment in front of a large group of professionals. If a team appointment is unavoidable, arrangements could be made to put patients more at ease. In an outpatient craniofacial clinic in Bristol, UK, a system of advocacy has been used with some success by the author and the treatment team (details available on request). A team member meets with the patient and family before the consultation, with the aim of putting them more at ease and eliciting their main concerns. This team member then presents and reviews the case for the team, and subsequently acts as an advocate for the patient and family in the consultation, supporting and if necessary, prompting them if they wish.

8.25.8.1.2 Psychosocial assessment

Carr (1997) has pointed out that although the technical outcome of a treatment is a matter for judgement by a clinician, the effectiveness of the treatment depends on the extent to which it meets patients' needs. Critical to this process is a thorough and accurate assessment of the patient's initial problems and expectations. Carr suggests that assessment should explore the person's current state, previous and ongoing difficulties, current coping strategies, expectations of outcome, and the anticipated impact of the desired changes on current problems. Macgregor (1990) pointed out that the decision to seek surgery should be a personal one and not the result of pressure from family or friends. It is likely that the majority of patients will have been experiencing distress for some time. Factors precipitating the decision to seek surgery should therefore be explored.

In a preoperative assessment, the psychologist should decide whether another intervention, for example, counseling, psychotherapy, or social interaction skills training, would be appropriate either as a forerunner, an adjunct, or as an alternative to surgery. If the person is experiencing significant additional stressors, it may be appropriate to delay surgery in case the disfigurement has become a "hook" on which to hang other problems.

Specialist psychological interventions in the treatment of dermatological conditions have been reported in the literature. Although the mechanisms are not yet clear, psychological distress is thought to play a part in some skin diseases. In some cases, suggestion, with or without hypnosis, has been shown to produce palpable modifications of skin diseases. Grossbart and Sherman (1986) found the concept of a "cellular battle," in which healing forces are imagined to attack and overwhelm the skin disease, to produce beneficial results. Brown and Fromm (1987) have reviewed treatment methods for a range of skin complaints including acne, psoriasis, eczema, and virus-mediated diseases. Although the psychological interventions appear to have been useful in each of these, the majority of reports comprise case studies, or small samples lacking adequate controls.

A psychosocial assessment can be a useful screening device for those unlikely to derive benefit from surgical intervention. Concern has been expressed as to the advisability of treating those seeking cosmetic surgery for relatively minor, and in some cases invisible, disfigurements (see comments on BDD in Section 8.25.1).

8.25.8.1.3 Facilitating decision-making

Making decisions about treatment for visible disfigurements can be difficult for all parties involved—the potential patient, the family, and the professionals providing the care (Bradbury & Middleton, 1997). The priorities of those involved in the decision-making process are likely to vary. There may be treatment options to consider (e.g., which reconstruction technique to choose, whether or not to opt for a prosthesis).

The timing of surgical intervention can be particularly critical in childhood and adolescence. The balance between waiting for physical growth to be completed and the psychological problems linked with the disfigurement may be a precarious one. The level of involvement of young patients in decisions concerning treatment is often less than ideal.

Children and adolescents frequently report that treatment decisions are made by the clinicians and parents, and they feel they have little say. In such cases, there will be a need to elicit the views of the young person before treatment decisions are finalized and to provide an atmosphere in which it is possible to say "no" to treatment.

Psychologists clearly have a potential role in promoting effective decision-making, for example, by ensuring that comprehensible and user-friendly information (written, verbal, or taped) is available, and by checking the level of understanding of those involved in the decision-making process (Bradbury & Middleton, 1997).

8.25.8.1.4 Preparation for surgery

Once the decision has been made to proceed with treatment, psychologists can usefully play a part in the preparation of patients for surgery. In some cases there may be a need to address concerns relating to a diagnosis. Prospective patients are likely to benefit from accurate information about the procedures they will undergo. Realistic expectations about the likely results should be established in order to minimize any discrepancy between expectations and outcome. The postoperative period should be discussed, for example, the likelihood of bruising and swelling, and possible functional restrictions following surgery.

Presurgery, highly anxious patients may benefit from relaxation training or an exploration of other potentially beneficial coping skills. Outcome research in the areas of plastic and orthognathic surgery suggest that predictors of positive adaptation to surgery include a good repertoire of coping skills, good overall body image, good social support, and a positive relationship with the treatment team (Pruzinsky, 1988).

8.25.8.2 Emergency Admissions

Ambler (1997) has argued that psychological factors are involved at every stage of hospital care following a severe disfiguring injury. Accidents do not occur entirely at random. There is a greater tendency for accidents to be experienced by people who are for some reason vulnerable. Burn injuries, for example, are sustained more often by the very young or the elderly, or by people who are confused or unwell. Consequently, pre-existing problems need to be considered as part of the emotional response to the trauma.

Ambler (1997) maintains that distress in the acute phase frequently relates to the experience of pain, to sleep deprivation, to anxiety associated with unplanned hospitalization, and to emotions related to the cause of the injury. As this phase passes, intervention may be required in order to establish realistic expectations of the outcome of surgery (e.g., in relation to the person's original appearance and function). Careful management is needed of the first glimpse the patient has of his/her disfiguration. Ambler points out that this can be stage-managed by the deliberate introduction of a mirror, or that it may happen unexpectedly if the patient happens to glance at a reflective surface.

Clearly, a more proactive approach than is currently the norm is needed in the management of psychological symptoms during the acute and early recovery stages of trauma care. Ambler (1997) has pointed out that this approach to care is likely to pay dividends in terms of better long-term trust and cooperation from patients, in addition to better adjustment. Staff need to be made more aware of the causes and nature of the psychological problems experienced and advice should be given about appropriate intervention. As in all treatment teams involved in the care of people with disfigurements, staff will need to make decisions about who will be responsible for the psychological aspects of care.

8.25.8.3 The Postoperative Period

Pruzinsky (1988) has noted that in the immediate postoperative period, patients often experience depression, anxiety, and doubts about the eventual outcome of the surgical procedure. Reassurance and empathy are appropriate postoperatively and are likely to be well received from a psychologist or member of the health care team who has established a rapport with the patient in the preoperative stage.

8.25.8.3.1 Discharge planning and preparation

For those with extensive disfiguration, predischarge liaison with schools or the workplace has been shown to be beneficial (Bernstein & Robson, 1993). This can help to prepare the patient for a return to everyday life and will also be useful for peers, teachers, or colleagues who may be uncertain of how the patient will look and of how they should behave. Although staff in the unit, close family, and friends may have become accustomed to the disfigurement, facing the general public can be a major hurdle. A program of accompanied trips, first around the hospital and latterly to venues beyond the familiar surroundings, is likely to be helpful (Ambler, 1997).

8.25.8.3.2 Long-term follow-up

Psychological support may also be necessary in the longer term following planned or emergency surgical intervention. Many problems may be encountered in adjusting to scarring and altered appearance in the longer term, whether or not a "good" surgical outcome has been achieved. Some people experience unwelcome feelings of change to their "identity" (see Section 8.25.6.3.4). Some have difficulties in dealing with changes in the level of attention from others, whether this is an increase due to the onset of a visible disfigurement, or a decrease resulting from a now unremarkable appearance. Others express disappointment, as expectations of change in their lives are not fulfilled. A few may demonstrate an inappropriate desire for further surgery.

8.25.8.4 Emotional Support for Professional Carers

Ambler (1997) has discussed the emotional strain on those professional carers who are engaged in treating people with extensive disfiguration. Particular pressures result from caring for patients requiring intensive treatment (for example, as the result of severe trauma or burns), and those with life-threatening conditions (as in advanced cancer of the head or neck). Although widely recognized as desirable (see, for example, Bernstein & Robson, 1983), the routine provision of staff support is the exception rather than the rule. Psychologists are well qualified to initiate, participate in, and evaluate a variety of methods of formal support, and to encourage informal support amongst team members.

8.25.9 PSYCHOLOGICAL INTERVENTIONS AND THE DEVELOPMENTAL PROCESS

In addition to providing psychological services as an integral part of hospital-based care, information, support, and in some cases individual or group interventions are likely to be particularly appropriate at key stages of the development of the disfigured person and their family.

8.25.9.1 The Birth and Early Development of a Disfigured Child

Responses to birth defects are influenced by the cultural and social context in which they occur (Strauss, 1985). However, as described in Section 8.25.6.1, in most societies, the emotional response to the birth of a congenitally deformed child has been likened to a bereavement reaction. Staff may also experience shock following the birth of an abnormal baby. The sensitive handling of the birth itself is essential, as is the provision of appropriate support immediately after the birth and in the weeks following discharge from hospital. Ideally, this would require specialist training of midwives, community, and hospital staff, and the availability of a psychologist who could meet with parents and support staff at short notice.

In the case of craniofacial anomalies, for example, cleft palate, or craniosynastosis, early surgery will be necessary. At this stage, parents will need detailed information, help with decision-making, and emotional support to address fears relating to the outcome of surgery (Hearst & Middleton, 1997).

In conditions where the relationship between the care team, the patient, and the family is likely to continue over a long period of time, the early establishment of rapport and trust with a team member is likely to be of considerable benefit. The psychosocial development and well-being of the child should be checked on a regular basis, with appropriate advice, support, and referral for psychological intervention offered if problems are identified.

Bradbury (1996) has advocated particular psychosocial interventions designed to tackle the problems commonly faced by disfigured children. These may include play therapy in preparation for surgery, and the teaching of specific skills to deal with teasing or bullying by other children. A child can be encouraged to learn strategies for dealing with the behavior of others. Bradbury has suggested a "force-field" in which the child imagines she/he is protected by an invisible barrier, which deflects the insulting "arrows" back towards the offending person. The child can be encouraged to role play difficult situations, substituting and practising alternative responses.

8.25.9.2 The Family Environment

The family environment is of crucial importance to the well-being and support of disfigured people at all stages of development. Clinical experience suggests that disfigured children and adults have better prospects in a family that places little emphasis on the role of physical appearance in a successful and fulfilling life. Parents can be encouraged to develop positive coping strategies and to model adaptive ways of dealing with unkind comments and stares. They can be advised to encourage their children to focus on other, more positive attributes, and to foster their child's independence.

Interventions may be necessary to tackle communication difficulties within families. Children frequently report that they do not want to hurt their parents' feelings by talking openly about their problems, and parents may similarly feel that a cheerful exterior should be preserved at all costs. Families may need support in making treatment decisions, for example, in relation to the timing of hospitalization. Siblings may be the casualties of extra time, effort, and attention which may be devoted to the disfigured child. For many of these problems, supportive listening and empathy may suffice. Other family issues are more complex and will require professional intervention.

8.25.9.3 Social Interaction Skills

Since the mid-1970s there has been a huge amount of interest in "social therapy" and its application to a variety of problems concerning social functioning. Argyle's (1988) social skills model suggested that problems experienced during social interaction can be ameliorated through the development of a new repertoire of social behaviors, designed to elicit more positive response from others. Social skills training has been shown to improve the social problems experienced by various groups with problems similar to those experienced by disfigured people, for example, those with spinal cord injury (Dunn, van Horn, & Herman, 1981) and those with social phobia (Wlazlo, Schroeder, Hand, Kaiser, & Munchau, 1990).

As many of the problems associated with visible difference are associated with social encounters with others, attention in this field has been focused on the possible benefits of social interaction skills training. The results of a small-scale study by Rumsey et al. (1986) suggested that facially disfigured people could influence the impressions formed of them by others through the use of social interaction skills. Robinson et al. (1996) evaluated social interaction skills workshops developed and offered by the charity Changing Faces. Sixty facially disfigured attenders reported high levels of anxiety and social anxiety preworkshop. However, six weeks postworkshop and at six-month follow-up, participants reported significantly less anxiety both in general and social situations, and significantly more confidence in the company of strangers and when meeting new people.

Participants liked being able to discuss their problems with others who could empathize. They also reported that the workshop had made them more aware of the way strangers feel when meeting a person who is visibly different. This understanding in some cases led to a greater tolerance and a desire to put the other party at ease. However, despite generally positive results, there was also considerable individual variation in the improvements made following the workshops. Clearly, a great deal of effort is required to change long-established patterns of behavior. Long-term follow-up and support are likely to be required to maximize the gains achieved.

Additional strategies for tackling problems in social encounters have been recommended by others with clinical experience of disfigurement. For those suffering from staring or from a heightened vigilance about the reactions of others, Bradbury (1996) advocates behavioral strategies such as unfocusing the eye and not seeing others, or cognitive strategies such as distraction, or thinking of others as being unaware of what they are doing, rather than trying to do harm.

The further development of psychological interventions to ameliorate the problems associated with social interaction would seem to be worthwhile. There is a consensus amongst workers and researchers in the field that an abnormal appearance can even become a positive advantage for those at ease with their appearance, and confident in their ability to establish and maintain effective relationships.

8.25.10 DISFIGUREMENT AND PSYCHOLOGY: PROFESSIONAL ISSUES

There is considerable potential for greater involvement by psychologists in the field of visible disfigurement. Efforts to improve the lot of visibly different people are needed on several fronts, not all of which relate directly to health care.

In the past, disfigured people and researchers alike have discussed the desirability of changing society's bias towards physical attractiveness and "wholeness," with the aim of removing the stigma associated with visible difference. Experience in other areas of attitude change (e.g., in relation to racial prejudice) suggests this is an ambitious, laborious, and time-consuming task, but not one which should be shirked. There is a need, for example, to increase the general public's knowledge about the causes and consequences of disfigurement, and to increase awareness of the distress experienced by those who are set apart from others by virtue of their appearance. Psychologists have been, and should continue to be, involved in producing documentaries, books, and articles aimed at increasing levels of relevant knowledge and awareness.

A novel initiative aimed at changing attitudes in relation to appearance was launched in the UK in 1995 by the charities The Phoenix Appeal and Changing Faces. An educational pack has been developed which includes a teacher's guide, sample lesson plans, activity ideas, and videos. The aim of the pack is to encourage children and teachers to explore issues relating to appearance and the effects of disfigurement within the school curriculum, in the hope that a greater understanding of visible difference will be fostered at an early age. Clinical and health psychologists have been involved in the development and the ongoing evaluation of this pack.

In terms of changing the prevailing ethos of health care provision and delivery for visibly disfigured people, psychologists have an obvious role to play. Several suggestions are made in this chapter for how they might contribute, but what sort of psychologists (or other health care professionals) should fill the various roles? The choice of individuals will depend on availability and local circumstances, however, clinical and health psychologists would seem to be particularly well placed to provide an appropriate service. Although there will be some overlap between these subspecialties, clinical psychologists, by virtue of their training, are likely to have skills well suited to the assessment of patients, to the development and implementation of interventions, and appropriate outcome measures. Health psychologists have the potential to evaluate and advise on treatment provision, to audit patient needs, to devise and evaluate appropriate packages of care to meet those needs, and to further knowledge through the application of theories and models to research in the area.

The way in which psychological support and interventions should be provided is currently a subject of debate. A routine presence of a psychologist in those outpatient clinics and inpatient wards dealing with people with visible differences is desirable. However, these clinics are numerous, and a lack of appropriately trained psychologists coupled with the inevitable lack of funding means this is very much the exception rather than the rule.

A cost-effective alternative is being piloted in the UK in the 1990s. Outlook, a Disfigurement Support Unit, based at Frenchay Hospital in Bristol, is being funded on a three-year trial basis by the charity Changing Faces. The unit is staffed by two part-time clinical psychologists (one adult and one child specialist), and an "outreach" community link worker. The work of the unit is being evaluated by a health psychologist. The unit takes referrals from a wide variety of specialities in the South and West regions of the UK, including burns, plastic surgery, dermatology, orthodontics, laser therapy, and community agencies. Staff provide a specialist assessment and intervention service for patients experiencing problems related to a disfigurement. Social interaction skills and cognitive-behavioral interventions are offered to patients on an individual or group basis. In addition, the unit is also a source of information and support for families. Early indications are that the service is being well received by patients, families, and staff.

8.25.11 CONCLUSIONS

There are numerous stressors associated with being visibly different in a society which places considerable emphasis on physical attractiveness and "wholeness." These can have a significant impact on people disfigured and their families.

Current care provision focuses on improving function and aesthetics. It has not, in the most part, taken account of what is known from psychological theory and research. There is a lack of understanding, both about the psychosocial problems associated with disfigurement and of ways of addressing these.

Disfigurement is not a glamorous speciality. It does not easily attract funding for treatment and research. However, considerable benefit could be derived from the involvement of psychologists in raising awareness, changing attitudes, and in the provision of care. There is a wealth of theory and research in health psychology which could usefully be applied as the field of disfigurement grows in sophistication and significance.

8.25.12 REFERENCES

Ambler, N. (1997). Case studies. In R. Lansdown (Ed.), *Visibly different: Coping with disfigurement* (Chap. 25). Oxford: Butterworth-Heinemann.

Argyle, M. (1988). Social relationships. In M. Hewstone (Ed.), *Introduction to social psychology: A European perspective* (pp. 224–245). Oxford: Blackwell.

Baker, C. (1992). Factors associated with rehabilitation in head and neck cancer. *Cancer Nursing, 15,* 395–400.

Barden, R., Ford, M., Jensen, A., Rogers-Sallyer, M., & Sallyer, K. (1989). Effects of craniofacial deformity in infancy on the quality of mother–infant interactions. *Child Development, 60,* 819–824.

Bernstein, N. (1976). *Emotional care of the facially burned and disfigured.* Boston: Little, Brown.

Bernstein, N., & Robson, M. (1983). *Comprehensive approaches to the burned person.* New York: Medical Examination Publishing.

Blakeney, P., Portman, S., & Rutan, R. (1990). Familial values as factors influencing long-term psychological adjustment of children after severe burn injury. *Journal of Burn Care and Rehabilitation, 11,* 472–475.

Bradbury, E. (1996). *Counselling people with disfigurement.* Leicester, UK: BPS Books.

Bradbury, E., & Middleton, J. (1997). Decision-making about treatment. In R. Lansdown (Ed.), *Visibly different: Coping with disfigurement* (Chap. 23). Oxford: Butterworth-Heinemann.

Brewin, C., McCarthy, B., & Furnham, A. (1989). Social support in the face of adversity: The role of cognitive appraisal. *Journal of Research in Personality, 23,* 354–372.

Brown, D., & Fromm, E. (1987). *Hypnosis and behavioral medicine*. Hillsdale, NJ: Erlbaum.

Bull, R., & Rumsey, N. (1988). *The social psychology of facial appearance*. New York: Springer.

Carr, A. (1997). Assessment and measurement. In R. Lansdown (Ed.), *Visibly Different: Coping with disfigurement*. Oxford: Butterworth-Heinemann.

Changing Faces (1996). *Facial disfigurement: The full picture*. London: Author.

Davidson, T., Bowden, M., Tholen, D., James, M., & Feller, I. (1981). Social support and post-burn adjustment. *Archives of Physical Medicine and Rehabilitation, 62,* 274–278.

Dunn, M., van Horn, E., & Herman, S. (1981). Social skills and spinal cord injury: A comparison of three training procedures. *Behaviour Therapy, 12,* 153–164.

Ginsburg, I., & Link, B. (1989). Feelings of stigmatization in patients with psoriasis. *Journal of the American Academy of Dermatology, 20,* 53–63.

Grossbart, T., & Sherman, C. (1986). *Skin deep: A mind/body program for healthy skin*. New York: Morrow.

Grealy, L. (1994). *In the mind's eye*. London: Century.

Harris, D. (1997). Types, causes and physical treatment of visible differences. In R. Lansdown, N. Rumsey, E. Bradbury, A. Carr, & J. Partridge (Eds.), *Visibly different: Coping with disfigurement* (Chap. 14). London: Butterworth-Heinemann.

Hearst, D., & Middleton, J. (1997). Psychological intervention and models of current practice. In R. Lansdown (Ed.), *Visibly different: Coping with disfigurement* (Chap. 22). Oxford: Butterworth-Heinemann.

Jowett, S., & Ryan, T. (1985). Skin disease and handicap: An analysis of the impact of skin conditions. *Social Science & Medicine, 20*(4), 425–429.

Langer, E., Fiske, S., Taylor, S., & Chanowitz, B. (1976). Stigma, staring and discomfort: A novel-stimulus hypothesis. *Journal of Experimental Social Psychology, 12,* 451–463.

Lanigan, S., & Cotterill, J. (1989). Psychological disabilities amongst patients with port wine stains. *British Journal of Dermatology, 121,* 209–215.

Lansdown, R., Lloyd, J., & Hunter, J. (1991). Facial deformity in childhood: Severity and psychological adjustment. *Child: Care, Helath and Development, 17,* 165–171.

Lefebvre, A., & Arndt, E. (1988). Working with facially disfigured children: A challenge in prevention. *Canadian Journal of Psychiatry, 33,* 453–458.

Lerner, M. (1970). The desire for justice and reactions to victims. In J. Macaulay & L. Berkowitz (Eds.), *Altruism and helping behaviour*. New York: Academic Press.

Liggett, A., & Liggett, J. (1989). *The tyranny of beauty*. London: Gollancz.

Long, D., & DeVault, S. (1990). Disfigurement and adolescent development: Exacerbating factors in personal injury. *American Journal of Forensic Psychology, 8,* 3–14.

Lovius, B., Jones, R., Pospisil, O., Reid, D., Slade, P., & Wynne, T. (1990). The specific psychosocial effects of orthognathic surgery. *Journal of Cranio-Maxillo-Facial Surgery, 18,* 339–342.

Macgregor, F. (1990). Facial disfigurement: Problems and management of social interaction and implications for mental health. *Aesthetic Plastic Surgery, 14,* 249–257.

Martin, J. (1988). *Surveys of Disability in Great Britain. Report: The prevalence of disability among adults*. London: HMSO.

Moss, T. (1997). Individual differences in adjusting to visible differences. In R. Lansdown, N. Rumsey, E. Bradbury, A. Carr, & J. Partridge (Eds.), *Visibly different: Coping with disfigurement* (Chap. 18). London: Butterworth-Heinemann.

Nash, P. (1995). *Living with disfigurement: Psychosocial implications of being born with a cleft lip and palate*. Aldershot, UK: Avebury.

Noar, T. (1991). Questionnaire survey of attitudes and concerns of patients with cleft lip and palate and their parents. *Cleft Palate-Craniofacial Journal, 28,* 279–284.

Office of Population Censuses and Surveys (1988). *Surveys of Disability in Great Britain. Report*. London: HMSO.

Orr, D., Reznikoff, M., & Smith, G. (1989). Body image, self-esteem and depression in burn injured adolescents and young adults. *Journal of Burn Care and Rehabilitation, 10*(5), 454–461.

Partridge, J. (1997). Living visibly different. In R. Lansdown (Ed.), *Visibly different: Coping with disfigurement* (Chap. 12). Oxford: Butterworth-Heinemann.

Pertschuk, M., & Whitaker, L. (1982). Social and psychological effects of craniofacial deformity and surgical reconstruction. *Clinics in Plastic Surgery, 9*(3), 297–306.

The Phoenix Appeal & Changing Faces (1995). *Face Values: An education pack for Key Stages 1 & 2*. London: Changing Faces.

Porter, J., Beuf, A., Lerner, A., & Norlund, J. (1986). Psychological effects of vitiligo: A comparison of vitiligo patients with "normal" control subjects, with psoriasis patients and with patients with other pigmentary disorders. *Journal of the American Academy of Dermatology, 15,* 220–224.

Porter, J., Beuf, A., Lerner, A., & Norlund, J. (1990). The effect of vitiligo on sexual relationships. *Journal of the American Academy of Dermatology, 22,* 221–222.

Pruzinsky, T. (1988). Collaboration of plastic surgeon and medical psychotherapist. *Medical Psychotherapy, 1,* 1–13.

Robinson, E. (1996). *An analysis of stress, coping and adjustment in relation to facial disfigurement*. Unpublished MSc thesis, University of Exeter, UK.

Robinson, E. (1997). Psychological research on visible differences in adults. In R. Lansdown (Ed.), *Visibly different: Coping with disfigurement* (Chap. 16). Oxford: Butterworth-Heinemann.

Robinson, E., Rumsey, N., & Partridge, J. (1996). An evaluation of the impact of social interaction skills training for facially disfigured people. *British Journal of Plastic Surgery, 49,* 281–289.

Rubinow, D., Peck, G., Squillace, K., & Gnatt, G. (1987). Reduced anxiety and depression in cystic acne patients after successful treatment with oral isotretinoin. *Journal of the American Academy of Dermatology, 17,* 25–32.

Rumsey, N. (1997). Historical and anthropological perspectives on appearance. In R. Lansdown (Ed.), *Visibly different: Coping with disfigurement* (Chap. 15). Oxford: Butterworth-Heinemann.

Rumsey, N., Bull, R., & Gahagen, D. (1986). A preliminary study of the effects of social skills training for improving the quality of social interaction for the facially disfigured. *Social Behaviour, 1,* 143–145.

Rumsey, N., Robinson, E., Turner, S., & Mercer, N. (1997). Patient or provider centred care? *Transactions: 8th International Congress on Cleft Palate and Related Craniofacial Anomalies* (pp. 423–426). Stamford Press.

Salzer, B., & Schallreuter, K. (1995). Investigation of the personality structure in patients with vitiligo and a possible association with catecholamine metabolism. *Dermatology, 190,* 109–115.

Settle, J. (1994). *Burn care facilities in the United Kingdom.* London: British Burn Association.

Sheridan, C., & Radmacher, S. (1992). *Health Psychology: Challenging the biomedical model.* New York: Wiley.

Strauss, R. (1985). Culture, rehabilitation and facial birth defects: International case studies. *Cleft Palate Journal, 22*(1), 56–62.

Strenta, F., & Kleck, R. (1985). Physical disability and the attribution dilemma: Perceiving the causes of social behaviour. *Journal of Consulting & Clinical Psychology, 3,* 129–142.

Sutherland, S. (1988). Burned adolescents' description of their coping strategies. *Heart & Lung, 17*(2), 150–157.

Turner, S., Thomas, P., Dowell, T., Rumsey, N., & Sandy, J. (1997). Psychological outcomes amongst cleft patients and their families. *British Journal of Plastic Surgery.*

Wallace, E. (1993). Nursing a teenager with burns. *British Journal of Nursing, 2,* 278–281.

Wallace, L. (1988). Abandoned to a social death? *Nursing Times, 84,* 34–37.

Walters, E. (1997). Problems faced by children and families living with visible difference. In R. Lansdown (Ed.), *Visibly different: Coping with disfigurement* (Chap. 17). Oxford: Butterworth-Heinemann.

Williams, E., & Griffiths, T. (1991). Psychological consequences of burn injury. *Burns, 17*(6), 478–480.

Wlazlo, Z., Schroeder, H., Hand, I., Kaiser, G., & Munchau, N. (1990). Exposure *in vivo* vs. social skills training for social phobia: Long term outcome and differential effects. *Behavioral Research & Therapy, 28,* 181–193.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.26
Dental Health

GERRY KENT
University of Sheffield, UK

8.26.1 INTRODUCTION 595

8.26.2 DENTAL ANXIETY AND DENTAL PHOBIA 596
- *8.26.2.1 Description* 596
  - *8.26.2.1.1 Phobias* 597
- *8.26.2.2 Epidemiology* 598
- *8.26.2.3 Etiology* 598
  - *8.26.2.3.1 Negative experiences* 598
  - *8.26.2.3.2 The dentist–patient relationship* 600
  - *8.26.2.3.3 Modeling* 600
  - *8.26.2.3.4 Biological characteristics* 600
- *8.26.2.4 Maintenance of Anxiety* 601
- *8.26.2.5 Impact on Quality of Life* 602
- *8.26.2.6 Treatment* 603
  - *8.26.2.6.1 Aspects of the intervention* 603
  - *8.26.2.6.2 Patient characteristics* 604
  - *8.26.2.6.3 The dentist–patient relationship* 605

8.26.3 PREVENTIVE CARE 606
- *8.26.3.1 Epidemiological Studies* 606
- *8.26.3.2 Improving Preventive Care* 607
  - *8.26.3.2.1 The educational approach* 607
  - *8.26.3.2.2 The behavioral approach* 608
  - *8.26.3.2.3 Social cognition models* 609

8.26.4 CONCLUSIONS 610

8.26.5 REFERENCES 611

8.26.1 INTRODUCTION

Dental health is of interest to psychologists for both theoretical and therapeutic reasons. Theoretically, the study of dental care provides an opportunity to conduct research into a variety of areas relevant to psychology, including occupational stress (Cooper, Watts, Baglioni, & Kelly 1988; Hilliardlysen & Riemer, 1988; Kent, 1987b), the care of patients with handicaps (Brown, 1980; Clevenger, Wigal, Salvati, & Burchill, & Crinella 1993) or who are elderly (Ettinger, 1984; Strauss & Hunt, 1993), the effects of surgery for cosmetic or structural reasons (Kent & Johns, 1993; Kiyak, 1993), and psychological approaches to the control of temporomandibular pain dysfunction syndrome (Cohen & Hillis, 1979; Gessell & Alderman, 1871; Locker & Slade, 1988; Pomp, 1974; Rugh & Solberg, 1976; Rugh, 1977, see also Chapter 8.24, this volume). Because virtually everyone in Western societies has need of dental care and has contact with professional services, research in these areas does not need to be limited to small numbers of patients with particular problems. In many ways, the dental setting can be considered as a kind of natural laboratory where people have to cope with stressful events on a regular basis. Young children are required to meet and deal with strange adults, often for the first time (see Chapter 8.27, this volume), and dental health requires preventive care at home on a regular basis, as well as professional interventions, so that theoretical notions concerning health promotion and self care are relevant. The dental setting has also proved to be an ideal situation to test and develop theories of the cognitive aspects of anxiety and pain. Levels of anxiety state rise and fall and Pavlovian learning processes are mediated and affected by cognitive processes such as memory, perceptual biases, and catastrophic thinking patterns.

Much of this theoretical work has had a direct influence on clinical practice as provided by dentists and clinical psychologists. Increasingly, dentists are calling upon psychologists to help them help their patients. Dentists are aware that sociological and psychological factors play important roles in achieving dental well-being. Psychologists can help in the design of health promotion material and can provide training in aspects of managing pain and improving dentist–patient communication. Clinical psychologists can provide therapeutic interventions. Because of the extensive research on the etiology, nature, and correlates of dental anxiety, a number of efficacious interventions have been developed. Study of the efficacy of these approaches can then be fed back to elaborate theoretical considerations.

The purpose of this chapter is to describe some of the research on the applications of psychology to dental health. It covers two of the topics mentioned above. Because the bulk of research has been conducted on dental anxiety and dental phobia, this area is given the most attention. The other topic covered is the promotion of preventive care, which has been problematic in many areas other than dentistry (see Chapter 8.29, this volume).

8.26.2 DENTAL ANXIETY AND DENTAL PHOBIA

This section of the chapter is concerned with the measurement, epidemiology, etiology, maintenance, and treatment of dental anxiety and phobia.

8.26.2.1 Description

As outlined below, dental anxiety and dental phobia are common. Depending on how it is measured, 3–5% of the general population can be said to have debilitatingly high levels of fear of dentistry. Insofar as fear can result in avoidance of professional care, an understanding of this difficulty could have many beneficial effects for dental health.

When asked, most people in the community agree that regular professional care is an important aspect of dental health. Yet 40–45% attend only when they are experiencing symptoms (Todd, Walker, & Dodd, 1982). Although reasons for this discrepancy include financial considerations and family patterns of care, one significant factor is anxiety level. Longitudinal studies have indicated that highly anxious patients are less likely to make, and more likely to break, dental appointments (Curson & Coplans, 1970; Stewart, Marcus; Christenson & Lin, 1994). Many authors have concluded that there is an indisputable relationship between avoidance and poor oral health. There is evidence that patients with high levels of anxiety are more likely to have experienced extractions and to have more decayed surfaces and fewer fillings (Bedi, Sutcliffe, Barrett, & McConnachie, 1992; Hakeberg, Berggren, & Grondahl, 1993; Kaufman, Rand, Gordon, & Cohen, 1992; Klingberg, Lofqvist, Bjarnason, & Noren, 1994; Locker & Liddell, 1992). It is possible that a vicious circle develops, in which fear leads to avoidance of professional dental care, which in turn leads to greater perceived need and greater anxiety (Moore, Brodsgaard, & Birn, 1991).

But, in fact, the relationship between anxiety, attendance, and dental health is far from clear. Some patients attend regularly despite high levels of anxiety. It may be that some patients are anxious because of their perceptions of the poor state of their oral health, or they may have become anxious because of previous extractions rather than vice versa. Also, there are many influences on oral health, including diet and fluoridation of toothpaste and of the water supply. It is possible that these background factors may be responsible for high anxiety or outweigh the disadvantages of poor attendance *per se*.

Milgrom, Weinstein, and Getz (1995) use Miller's (1959) description of approach–avoidance conflicts to explain the relationship between perceived need for care on the one hand, and reluctance to attend on the other. They suggest that due to knowledge about the consequences of neglect and cultural pressures to attend the dentist, most people are motivated to attend. If anxiety is low, then there would be little conflict with this approach motivation and regular appointments would be made and kept. If, however, a patient were anxious, a conflict would develop as the time for an appointment approached. In keeping with research on increasing anxiety state and intrusiveness of cognitions over time (Jongh & Stoutland, 1993; Kent, 1987a; Kent & Gibbons, 1987 and see below), such a conflict could result in a "no show" or canceled appointment.

8.26.2.1.1 Phobias

DSM-IV (American Psychiatric Association, 1995) indicates that an individual with dental fears ought to meet seven criteria to qualify for diagnosis of a specific phobia:

(i) Marked and persistent fear that is excessive or unreasonable, cued by the presence or anticipation of a specific object or situation.

(ii) Exposure to the phobic stimulus almost invariably provokes an immediate anxiety response.

(iii) The person recognises that the fear is excessive or unreasonable.

(iv) The phobic situation(s) is avoided or else endured with intense anxiety.

(v) The avoidance, anxious anticipation, or distress in the feared situation interferes significantly with the person's normal routine, occupational functioning, social activities, or relationships, or there is marked distress about having the phobia.

(vi) In individuals under age 18 years, the duration is at least six months.

(vii) The anxiety, panic attacks, or phobic avoidance are not better accounted for by another mental disorder.

In the literature on dental phobia, some of these criteria are better researched and more adequately measured than others. Whereas there are several scales designed to measure the extent of anxiety according to criteria (i), (ii), and (iv), not until the 1990s have criteria (v) and (vii) received attention.

There are several self-report questionnaires. Perhaps the best-known and most commonly used scale is Corah's Dental Anxiety Scale or DAS (Corah, 1968), a four-item self-report questionnaire with scores ranging from 4 (low) to 20, with a score of 12 or greater indicating moderate anxiety and a score of 15 or greater indicating extreme anxiety (Corah, Gale, & Illig, 1978). It is quick to complete and provides a general measure of perceived emotional arousal and physiological reactions. A scale which inquires about fears of specific aspects of the dental setting (e.g., about drilling, injections) is the Dental Fear Survey or DFS (Kleinknecht, Klepac, & Alexander, 1973), which has somewhat more support in terms of validity and reliability (Moore, Berggren, & Carlsson, 1991; Schuurs & Hoogstraten, 1993) than the DAS. The Dental Belief Survey is concerned with patients' beliefs about the way they are treated interpersonally by dentists, including such items as "dentists do not really listen" and "I don't feel comfortable asking questions" (Milgrom, Weinstein et al., 1995; Smith, Weinstein, Milgrom, & Getz, 1984). There are several other scales (Jongh & Stouthard, 1993; Schuurs & Hoogstraten, 1993; Horst & Wit, 1993) which attempt to gauge various aspects of these social and emotional aspects of the dental situation.

There has been interest in assessing the cognitive component of dental anxiety. As in the case with other types of anxiety, attentional bias has been shown to occur in dental anxiety. Muris, Merckelbach, and Jongh (1995) used the Stoop test to examine the biases of patients with high and low dental anxiety, as measured by Corah's DAS. Highly anxious participants were significantly slower in naming words related to dental threats (e.g., nerve, blood, drilling, tooth) than those with low anxiety, but levels of trait anxiety and type of coping behavior (monitoring vs. blunting (Miller, 1987)) were not related to speed of naming. With the recognition that highly anxious dental patients experience many more negative cognitions of a catastrophic kind than low anxiety patients (Jongh, 1993; Jongh, Muris, Horst, Van-Zuuren, & Wit, 1994; Kent & Gibbons, 1987), Jongh, Muris, Schoenmakers, and Horst (1995) have developed the Dental Cognitions Questionnaire, which contains such items as "I will panic during treatment," "surely something will go wrong," and "anesthetics often do not work": patients are asked to indicate both the frequency and the believability of the statements. Not only does the frequency of such thoughts discriminate extremely well between phobic and nonphobic patients, but they contribute substantially to the variance in anxiety state.

There are also many scales which are appropriate for children, which can be completed by parents, such as the dental subscale of the Children's Fear Survey (Klingberg, 1994) or scales which are completed by children themselves. The latter include projective techniques (Klingberg & Hwang, 1994) and asking children to choose between a set of pictures which illustrate varying levels of distress (Venham, 1979).

Besides self-report questionnaires, attempts have also been made to assess the physiological aspects of dental anxiety and phobia, but as in other areas there are many conceptual difficulties in dealing with results. Several studies (Benjamins, Schuurs, Asscheman, & Hoogstraten, 1990; Berggren & Carlsson, 1985; Carlsson, Linde, Berggen, & Harrison, 1986; Harrison, Carlsson, & Berggren, 1985; Hugdhal & Ost, 1985) have shown that there is no clear correspondence between measures of physiological arousal as measured by heart rate, skin conductance, or blood pressure. Although Lang's (1971) three-factor model of anxiety does not require concordance between self-report, and behavioral and physiological reactions, one would expect that there would be fairly consistent differences between high and low anxious patients, or that there would be changes on physiological indices after successful therapeutic interventions. However, this has not been shown consistently (Carlsson et al., 1986; Hirschman, Revland, Hawk, & Young, 1969).

8.26.2.2 Epidemiology

The prevalence of dental anxiety and phobia is interesting in its own right, but by examining age and cultural patterns it might also be possible to identify factors which are relevant to etiology. If prevalence rates vary across cultures, for example, it is plausible that learning processes, cultural beliefs, or the way that dental care is delivered, may be crucial in the development of dental phobia. On the other hand, if prevalence is similar across cultures, then biological processes could be more significant.

Estimates of the prevalence of extreme levels of dental anxiety and phobia vary with the populations studied and the measures used. In several studies with a variety of populations, mean scores on the DAS range between 7 and 8 (Horst & Wit, 1993; Kaufman et al., 1992). Using the criterion of scores of 15 or greater, studies have indicated that between 3% and 5% of people in unselected adult populations experience severe dental anxiety (Hakeberg, Berggen, & Carlsson, 1992; Kleinknecht et al., 1973; Klingberg, Berggen, & Noren, 1994; Moore, Birn, Kirkegaard, Brodsgaard, & Scheutz, 1993; Vassend, 1993), but there are a few higher estimates, of up to 13% or higher (Horst & Wit, 1993; Mellor, 1992). For children and adolescents, using different scales, the estimates tend to be somewhat greater, from 6% to 7% (Klingberg, 1994; Moraes, Milgrom, Tay, & Costa, 1994). Using self-reported avoidance to define extreme anxiety, the prevalence seems to be higher. For example, in a large representative Dutch sample, 36% reported "dread" of visiting the dentist (cited in Lindsay & Jackson, 1993).

While it is possible to make some general observations concerning prevalence—that high levels of dental anxiety are very common in all countries studied, they are more likely in women than in men, and they peak during early adolescence and decline with advancing age (Liddell & Locker, 1993)—the wide range of instruments used means that it is not possible to provide a definitive answer with respect to prevalence. Prevalence studies using standard measures with representative samples are still required (Vassend, 1993). In order to understand etiologic processes, it is necessary to move away from epidemiological data towards more psychologically based research.

8.26.2.3 Etiology

In general, the evidence suggests that a Pavlovian model is helpful in understanding the onset of dental anxiety and dental phobias. The classical conditioning model applied to this situation is that an Unconditioned Stimulus (UCS) (e.g., pain during an invasive dental procedure), which results in the Unconditioned Responses of anxiety and flight, has been paired with a Conditioned Stimulus (CS) (e.g., a dentist, the dental office, the sounds made by dental instruments), so that presentation of the UCS results in anxiety and dental avoidance. This useful model, however, needs to be adapted to include the idea of latent inhibition, issues of perceived control, innate biological predispositions to learn anxiety responses, and the nature of the dentist–patient relationship. Unfortunately, the reliance on retrospective accounts of past traumas and experiences is exceeding problematic for many reasons, and there is a dearth of longitudinal research.

8.26.2.3.1 Negative experiences

The notion that dental anxiety is the result of a classically conditioned association between pain and dental care has been supported many times (Horst & Wit, 1993; Hugdhal & Ost, 1985; Ost, 1987; Milgrom, Mancl, King, & Weinstein, 1995; Moore, Berggren et al., 1991). It received early support from Lautch (1971), who interviewed phobic and nonphobic patients about their previous experiences. In his sample of 34 phobics, all reported at least one previous traumatic experience with dental care—usually the experience of intense pain—whereas only 10 of the 34 nonphobics could remember such an incident. Of the phobic patients who had returned to a dentist at least once after the trauma (30 patients), all reported further traumas but only one of the nonphobics did so, suggesting that more than one negative experience is often needed for the development of a phobia.

This Pavlovian model is consistent with the notion of preparedness, that humans are innately predisposed to become anxious about some objects and situations more readily than others (Silva, 1988). It is possible to see a link with dentistry in this regard. Lying on one's back with an adult placing sharp instruments in the mouth "should," in some sense, be frightening. Anxiety might be heightened particularly if the adult is a stranger (patients often find the prospect of joining a new practice distressing because they are concerned that the dentist will have a rough manner) or if the child is a patient (where there is a large disparity of strength). This attractive idea goes some way towards explaining the widespread prevalence of dental anxiety, why anxiety is most likely to develop before adulthood, and why it is often associated with fears about pain and mutilation.

However, there are a number of difficulties with this explanation which relates conditioning processes with preparedness. One difficulty is that many people who are not highly anxious or phobic have also experienced a situation where classical conditioning could be expected to occur. In one large scale survey, 30% of respondents reported that they had experienced "intense" or "violent" pain at some point in their dental care histories (Vassend, 1993). Conversely, some phobic patients cannot recall such an incident (Moore, Brodsgaard, & Birn, 1991). Another difficulty is that all of these studies have relied on retrospective accounts of experience, thus introducing all the cognitive and memory biases that retrospective accounts involve. For example, the timing of the request for information on previous experiences may be relevant. Many of the studies in this area have asked patients to recall incidents when they arrive for treatment, or at least when they are scheduled to attend, when levels of anxiety state are high. In a study designed to explore the effect of anxious mood on recall, Kent (1989) asked patients to describe a previous visit to a dentist, but some were asked to do this just before a dental appointment, whereas others were asked to do so just afterwards. Patients with higher levels of dental anxiety rated their previous experience as being more unpleasant than patients with lower anxiety. For females, but not for males, the experience was rated as being significantly more unpleasant before the appointment than afterwards. When judges were asked to rate the vividness of the descriptions, a similar pattern emerged. Thus, it is plausible that anxious patients—whose levels of anxiety state are high before an appointment—may recall previous incidents as being very unpleasant because of their anxious mood, and not vice versa.

A third difficulty with these studies is that phobic patients are typically asked to relate their current feelings to previous experiences. This raises issues of attribution, particularly the tendency to attribute negative events to external causes (Ross, 1964). It is possible that individuals who are already highly anxious, perhaps because of other factors or high levels of trait anxiety, will attribute their anxiety to a previous dental appointment, whereas others with low trait anxiety, who might have undergone a similar experience, will not do so. In the absence of a method for objectively rating the aversiveness of an experience, attribution theory can provide an equally plausible explanation for the association between "traumatic" events and the onset of a dental phobia.

Davey (1989) and Jongh et al. (1995) have added a degree of sophistication to this area by exploring the possibility that latent inhibition is relevant to this learning process. Latent inhibition refers to the hypothesis that if a CS (in this case a visit to the dentist) is presented alone on several occasions before it is paired with the UCS (a painful or traumatic experience such as pain), then it is more difficult to condition an association between the CS and UCS than if there were no CS alone trials prior to conditioning. Thus, according to this principle, dental patients might not become anxious if they have a history of previous positive (or at least not negative) visits to the dentist.

Although Davey and Jongh et al. again relied on retrospective accounts in their studies with students, they attempted to specify the conditions under which traumatic events failed to promote the development of anxiety. Davey asked his sample to indicate their current levels of anxiety, their experiences of pain or trauma in the dental setting, and, importantly, whether their levels of anxiety had changed over time. One group consisted of people who had always been anxious about dental treatment, a second group consisted of those who were once anxious but were now more relaxed, a third group comprised those who were once relaxed but who were now anxious, and the fourth group comprised those who had always been relaxed. For the present purposes the important finding was that, in both studies, latent inhibition seemed to have been operating. Subjects who reported the experience of a traumatic event but who did not become anxious were more likely to have had a number of previous nontraumatic encounters. Latent inhibition was attenuated after the experience of a particularly painful event. Like Lautch (1971), Davey found that those who had always been anxious had received two or more aversive experiences.

Clearly, prospective studies are required before such conclusions can be confirmed. One of the great weaknesses of research in dental anxiety and phobia is the almost complete lack of longitudinal data. An important exception to this is provided by Murray, Liddell, and Donohue (1989). They collected data from children aged 9 and then again at age 12, while monitoring their dental experiences from dental records. Children who did not receive invasive treatment (i.e., having an injection, having teeth filled or removed) during this time were significantly more anxious than those who had received treatment, while the anxiety of children who attended regularly and who had experienced invasive treatment did not change significantly. This result implies that exposure acts prophylactically and is inconsistent with the notion of latent inhibition.

8.26.2.3.2 The dentist–patient relationship

Most of the research relating dental anxiety to experience involves encounters with pain. However, there is another thread of research which examines encounters with the dentist him or herself. Bernstein, Kleinknecht, and Alexander (1979) first divided their sample of students into high- and low-fear groups. The students were then asked to write an essay describing their visits to the dentist as children. These essays were to include "the features and events associated with those visits which determine present attitudes." Of the high-fear group, 42% mentioned pain during early appointments as a factor in their present feelings about dental care, compared to 17.4% of the low-fear students. While these proportions discriminated between the groups to some extent, there was still the 58% of high-fear students who did not cite pain as a factor and a sizeable minority of low-fear students who did experience pain.

This overlap was clarified by considering another factor—the perceived manner of the dentist. Half of the high-fear group mentioned negative behavior on the part of the dentist as a factor in their present feelings, most of these not citing pain as a reason for their current anxiety. The dentists were considered "impersonal," "uncaring," "uninterested," or "cold." Furthermore, of the low-fear group who had experienced pain as children, many found the dentist "careful" "patient," and "friendly." The effects of any pain seemed to have been mitigated by a caring and concerned dentist. Thus this study suggested that the dentist had an independent effect on the students' feelings (i.e., cold or uninterested behavior was enough to make some students feel negatively about dentistry) and an interactive effect (i.e., caring and warmth could obviate the long-term effects of painful experiences).

8.26.2.3.3 Modeling

Besides the direct experience of aversive events, there is evidence that learning can take place in a more indirect fashion, via modeling or cultural beliefs about an association between dental care and pain. Social learning theory would suggest that children could learn that professional dental care is painful and distressing through observing parental reactions in the dentist's office or hearing about negative experiences. A number of studies have implicated the importance of parental—especially maternal—responses to dental care as an etiologic influence (Corkey & Freeman, 1994; Horst & Wit, 1993; Klingberg et al., 1994; Milgrom, Mancl, King, & Weinstein, 1995). However, such learning may be overridden by actual experience rather quickly (Koenigsberg & Johnson, 1972). In most cases these studies are again reliant on retrospective accounts (although actual observational studies support a link; Johnson & Baldwin, 1968), and may be more relevant to the development of moderate, rather than extreme, levels of anxiety. In any case, it is not possible to disentangle learning processes from biological ones in these studies: it is possible that children of anxious parents are fearful themselves not because they had observed distress but because both children and parents share a genetic propensity to be anxious. This possibility leads to another major explanation as to the etiology of dental fear—the notion that there are individual differences in propensity to learn anxiety responses.

8.26.2.3.4 Biological characteristics

If the conditioning approach places the cause of phobias within the experience of dental trauma, the biological approach suggests that there are individual differences in how patients react to dental care. It is possible that some people will be more liable to learning anxiety responses than others. This approach is particularly relevant to the finding that, in some patients at least, anxiety occurs in the absence of any specifiable negative experience. Some children are very anxious on their first visit, for example. There are three types of study which support this approach—studies on temperament, studies on neuroticism, and studies on other psychiatric difficulties reported by phobic patients.

The concept of temperament is based on the work of Thomas and Chess (Thomas, Birch, Chess, Hertzig, & Korn, 1963), whose studies describe the range of responses given by young children, especially when they encounter new situations. These reactions emerge at a young age and can persist for many years. Of particular relevance to this area is the notion of the "difficult child," one who tends to withdraw from and who has difficulty in adapting to novel situations. The relevance of temperament has been explored in five studies (Holst, Hallonsten, Schroder, Ek, & Edlund, 1993; Liddell, 1990; Lochary, Wilson, Griffen, & Coury, 1993; Radis, Wilson, Griffen, & Coury, 1994; Williams, Murray, Lund, Harkiss, & Franco, 1985), all of which suggest that children who have difficulty in meeting unfamiliar people and coping with new situations are less likely to be able to accept routine dental care, even under sedation (Lochary et al., 1993). Associated with temperament is Eysenck's notion of neuroticism. Again, several studies (Frazer & Hampson, 1988; Horst & Wit, 1993; Klepac, Dowling, & Hauge, 1982; Lautch, 1971) have indicated that patients with high levels of dental anxiety also score highly on the N scale of the Eysenck Personality Inventory (EPI) or Eysenck Personality Questionnaire (EPQ).

Taken together, the work on temperament and neuroticism would suggest that many people with high levels of dental anxiety also have a number of other unrelated anxieties and concerns. This is a central issue, since it bears upon the possibility that dental anxiety and phobia can be indicative of more general psychological or psychiatric difficulties, rather than as a specific difficulty. There is now growing evidence that this is often, but not always, the case (Alwin, Murray, & Niven, 1994; Brown, Wright, & McMurray, 1986; Corkey & Freeman, 1994; Horst & Wit, 1993; Klingberg et al., 1994; McNeil & Berryman, 1989; Milgrom, Weinstein, & Getz 1995; Moore, Brodsgaard, & Birn, 1991). A 1994 study (Roy-Bryne, Milgrom, Khoon-Mei, Weinstein, & Katon, 1994) indicated that in a sample of phobic patients, 40% could be assigned another concurrent psychiatric diagnosis, such as panic disorder.

8.26.2.4 Maintenance of Anxiety

In many respects, it is possible to see how high levels of dental anxiety are maintained, due to the defining characteristic of avoidance. A decision to put off visiting a dentist would itself be reinforcing since it results in a short term reduction of anxiety, as Mowrer's (Mowrer, 1960) two-stage theory would suggest, and temporarily reduces the approach–avoidance conflict (Milgrom, Weinstein et al., 1995). It would also explain how anxiety could increase over time. As perceptions of worsening dentition increase, so would the perception of the need for invasive treatment. However, this behavioral approach does not explain why many patients who continue to attend regularly still experience high levels of anxiety. Although anxious patients typically experience much less pain than they expect (Kent & Warren, 1985; Lindsay & Jackson, 1993), high levels of anxiety are often maintained. There are several possible reasons for this apparent paradox.

One reason is that dentistry is not consistently pain-free. One survey of dentists themselves indicated that local anesthetics can fail in about 13% of instances (Kaufman, Weinstein, & Milgrom, 1984), so that one negative experience could reconfirm an anxious patient's belief that dental treatment can be painful and thus anxiety is maintained. Lindsay (Lindsay & Jackson, 1993) argues that the anxious patient may believe that it is less distressing to expect pain—just in case it occurs—and be pleasantly surprised if it doesn't, than be unexpectedly hurt. The negative effects of unpredicted sudden pain have been demonstrated, under experimental conditions, supporting this idea. Arntz and colleagues (Arntz, Eck, & Jong, 1991; Arntz, Eck, Jong, & Hout, 1990) have demonstrated that underpredicted pain experiences increase anticipatory anxiety and avoidance, and also result in dishabituation (Arntz & Lousberg, 1990).

Even if all future dental appointments were pain free there are several cognitive mechanisms which can be used to explain the maintenance of anxiety. One concerns the notion of typicality. As noted above, latent inhibition theory suggests that dental anxiety is unlikely to develop if the patient had previously experienced many pain-free appointments. In the language of cognitive heuristics, this could be seen in terms of typicality: one painful experience could be interpreted as anomalous and unlikely to recur. The converse of this is that anxiety might be maintained if pain-free experiences were seen as atypical, perhaps as a result of chance, rather than because of the nonaversive nature of care itself. Kent (1986) explored this possibility by first ascertaining the discrepancy between expected pain (as ascertained before the appointment) and experienced pain (as ascertained at the end of treatment). Patients were also asked at the completion of the appointment to indicate the extent to which they believed the amount of pain they experienced was typical through the question "How confident are you that you will experience the same amount of pain (on the next appointment) as today?" When contacted three months later, decreases in dental anxiety were found for patients who (i) experienced a sizeable discrepancy between expected and experienced pain and who (ii) had a high level of confidence in the typicality of this discrepancy. A decrease in anxiety was less likely if the patient had low confidence. It is possible, then, that regular attendance, and the experience of pain-free dentistry, is slow to result in a reduction of anxiety because of the covariation biases which can be present in phobias more generally (Jong, Hout, & Merckelbach, 1995).

The common six month period between appointments may also contribute to the persistence of anxiety for patients who require little treatment. The "natural" exposure that regular attendance implies may not be sufficiently frequent. Studies of the effectiveness of behavioral therapies suggest that interventions for phobias and anxiety are more efficacious if they are applied frequently (e.g., a few times per week over a few weeks) rather than infrequently (e.g., one or two days per month) (Marshall, Gauthier, & Gordon, 1980). In cognitive terms, frequent interventions may be more efficacious in altering expectations than infrequent ones because there is less opportunity for positive experiences to be reconstructed into existing anxiety-dominated schemata.

That such schemata can serve to maintain anxiety was shown in two studies (Arntz, Eck, Jong et al., 1990; Kent, 1985b) where patients were asked to indicate the levels of expected pain (before the appointment), their levels of experienced pain (afterwards), but were then contacted again three months after the appointment. On this later occasion they were asked to recall the amount of pain they had experienced at the appointment. The findings implicated the relevance of cognitive structures, since the amount of pain remembered by the anxious patients had increased over time, becoming more similar to the original expected pain, while the nonanxious patients recalled their levels of experienced pain more accurately.

Clearly, there is considerable scope for future research into the reasons why anxiety is maintained, but equally importantly into the reasons why anxiety comes to be reduced without formal interventions. Although longitudinal studies are more expensive in time and effort, they are clearly needed here, since they could provide information which could serve to inform the choice and design of therapeutic interventions, an area which is considered below.

8.26.2.5 Impact on Quality of Life

While the scales and measures discussed earlier attempt to assess level of anxiety according to the (i), (ii), and (iv) *DSM-IV* criteria, since the mid-1980s there has been interest in criterion (v): the effect on social relationships and activities, or what has become known as quality of life. Not only might a phobia lead to avoidance of professional care, but also embarrassment, days off work, and a degree of social isolation. Research in these areas is begining to appear. Berggren and Carlsson (1986) found that successfully treated phobics reported several areas of psychosocial improvement, including reduced alcohol intake and reduced time on sick leave and, as shown below, treatment of dental phobia can improve functioning in these areas. Berggren (1993) sent questionnaires to 109 patients who were referred for assistance to a dental anxiety clinic. Although there was no comparison group, many patients scored highly on a modified Nottingham Health Profile, with 52% indicating that their dental fear caused problems with social activities (eating out, meeting friends), 46% with going on vacation, and 41% with family relationships. These types of situations, which demand interactions with other people, could be severely curtailed.

Another approach was taken by Kent, Rubin, Getz, and Humphries (1996), who designed a scale to measure the social and psychological effects of phobias. They compared the responses of patients who had been refered for anxiety difficulties with patients attending an emergency dental clinic for other difficulties. The phobic patients were more likely to endorse items which indicated social and cognitive consequences (e.g., "I feel that people will laugh at me if I tell them about my fears about dentistry" and "the need to see a dentist is constantly on my mind"), avoidance (e.g., "when walking or driving somewhere I take a route in order to avoid passing by a dentist's office" and "I become upset when I see adverts on TV about tooth decay and tooth loss"), and social inhibitions (e.g., "I stop myself from smiling or I cover my mouth when I laugh" and "I am reluctant to meet new people because of the state of my teeth").

Thus, it can be argued that an accurate and complete assessment of dental phobia needs to include a variety of areas. Stouthard, Hoogstraten, and Mellenberg (1995) contend that there are three facets to dental phobias; namely a situation facet (including the nature of the dentist–patient relationship as well as the specific fears of particular procedures, such as drilling and extractions, and preliminary aspects, such as sitting in the waiting room and making an appointment), a time facet (as the appointment approaches in time), and a reaction facet (including effect, physical reactions, and cognitive intrusions). The Dental Anxiety Inventory includes items on all three aspects of the phobia and, with development, could provide a more comprehensive assessment than other scales.

8.26.2.6 Treatment

A wide variety of interventions have been employed in the treatment of dental anxiety and phobia, with considerable success. With doubts about the appropriateness and safety of general anesthesia, there has been a growing interest in the use of conscious sedation as afforded by nitrous oxide (Roberts, Gibson, Porter et al., 1979; Veerkamp, Gruythuysen, Hoogstraten, & Amerongen, 1993) and other sedatives (Kupietzky & Houpt, 1993). Several behavioral therapies have also been used in dentistry, ranging from such minimal, yet effective, interventions as information leaflets (Jackson & Lindsay, 1995) and the playing of music (Anderson, Baron, & Logan, 1991), which are probably most effective for low or moderate levels of anxiety, to more intensive interventions which are appropriate for extreme levels of anxiety. Many of these interventions involve a traditional graded exposure approach to the dental setting, the encouragement of patient control while receiving dental treatment, and attention-focusing techniques such as hypnosis (Gokli, Wood, Mourino, Farrington, & Best, 1994; Horst & Wit, 1993; Kent, 1985a, 1991; Melamed, 1979; Robertson, Gatchel, & Fowler, 1991). As would be expected, they are successful in relieving anxiety for most patients (Milgrom, Weinstein et al., 1995; Moore, Brodsgaard, Berggren, & Carlsson, 1991b), although across many studies approximately 25–35% of patients do not benefit or cannot follow through with treatment.

One possible reason for apparent lack of success for some patients concerns the choice of outcome measure. The most usual measures concern reductions of anxiety—either anxiety state or scores on dental anxiety scales, and/or the achievement of more regular attendance patterns. However, there is no one-to-one relationship between a reduction of measured anxiety and subsequent improvement in attendance (Schuurs, Makkes, & Duivenvoorden, 1995). It can be argued that the range of outcome measures ought to be broadened to include some quality of life measures, such as reduction in embarrassment and dysphoric mood, and improvements in social functioning (Moore, Brodsgaard, Berggren et al., 1991b). For example, there is evidence that treatment for dental anxiety can have positive effects on a variety of psychosocial measures, such as use of drugs and days off work (Berggren & Carlsson, 1986). In order to examine the latter possibility, Hakeberg and Berggren (1992a) examined data from the local National Health Insurance Office in Sweden to ascertain the time spent on sick leave by successfully treated patients and by a comparison group of patients who were attending the dental hospital for difficulties not connected with anxiety. Before the intervention, the anxious patients spent a median of 15.0 days off work on sick leave in the previous year, but this was reduced to 3.0 days in the year following treatment. By contrast, there was no change in the number of days off work in the comparison group. These results suggest that a wider range of outcome measures would be appropriate and might provide a more sophisticated view of the effects of treatment.

There are other reasons for the variability in success rates, however, including aspects of the intervention, patient characteristics, and aspects of the relationship between patients and their dentists.

8.26.2.6.1 Aspects of the intervention

As in other areas of psychological care, there has been interest in combining techniques in multitreatment packages. Kleinknecht and Bernstein (1979) used this idea to help two patients who had avoided seeing a dentist for six and 13 years. The patients first saw a film of a model being treated, then they were given relaxation training and a course in systematic desensitization. Finally, the patients visited a dentist's office to sit in the waiting room for 30 minutes, and held and manipulated the dental instruments. One year after this program, both patients had made several visits involving extractions, a root canal, and many fillings. The total time required by the therapist was 3.5 hours for each patient.

One area which deserves attention is the control of anxiety-provoking cognitions. Since levels of anxiety are associated with both the content of cognitions and patients' ability to control their intrusiveness (Jongh et al., 1994; Kent & Gibbons, 1987), diminishing catastrophizing ideation and providing patients with means of controlling thoughts may be an important determinant of adjustment to dental treatment and reduction of distress (Jongh et al., 1995).

Another issue concerns the method of delivery of the intervention. In most studies, interventions have been provided on an individual basis. This may not be necessary, with group therapy being, potentially, at least as clinically effective and more cost-effective. Using systematic desensitization, Moore and Brodsgaard (1994) compared group vs. individual sessions for patients with extreme levels of dental anxiety. There were no significant differences in dropout rates between the two types of intervention, but group therapy required fewer therapist hours and, at the completion of treatment, the group approach resulted in greater declines in anxiety levels. At a one year follow-up, however, patients in group therapy were less likely to have continued to visit a dentist. Moore and Brodsgaard argue that a combination of group and individual therapy might provide the advantages of group support and allow the therapist to target the needs of individuals.

Bandura (1982) would argue that the treatments used to help patients with dental anxiety are successful because they increase patients' belief that they are able to control their behavior and affect outcomes. Thus, an intervention would be successful if perceptions of control and self-efficacy were enhanced. There is a large body of research in the stress literature which indicates that people are more able to cope with stressful events if they believe that they have some control over the nature and aversiveness of outcomes (Steptoe & Appels, 1989), particularly for those with a high desire for control (Clevenger et al., 1993). In the dental setting, patients often experience a lack of control (partly because it is difficult to communicate needs with a mouthful of instruments) and there are several studies which indicate that providing a measure of control through stop signals (Wardle, 1983), information (Thrash, Marr, & Box, 1982), and instructions to focus their attention on particular aspects of their treatment (Baron, Logan, & Hoppe, 1993) can have beneficial effects on anxiety levels.

Litt, Nye, and Shafer (1993) explored the importance of self-efficacy by assigning some patients to a relaxation condition and others to a self-efficacy enhancing condition. In the latter, the patients were given false feedback about their ability to relax. Not only was the biofeedback equipment set to indicate great success in relaxing, but also the experimenter verbally emphasised their ability. Measures of patients' belief in ability to cope, distress before a scheduled appointment to have a molar removed, and staff's blind ratings of distress during the appointment, all indicated that the self-efficacy enhancement strategy was successful and provided additional benefits to relaxation alone.

8.26.2.6.2 Patient characteristics

There are also characteristics of individual patients to consider. Although some researchers have suggested that patients with higher levels of pretreatment anxiety would be less likely to benefit, this has not been demonstrated (Berggren & Carlsson, 1985). There is evidence that many patients with dental phobia have more generalised difficulties (Roy-Bryne et al., 1994). Liddell and Locker (1994) found that elderly individuals with wide ranging problems were more likely to carry their dental anxiety into old age than those whose anxiety was more specific. We could expect that the more specific the dental anxiety, the more likely it would be that a behavioral treatment would be successful. Patients who do not benefit from traditional graduated exposure to the dental setting may "fail" because the dental anxiety would only be a manifestation of more deep-seated difficulties. This possibility was supported by Berggren and Carlsson (1985), who found that patients who were anxious about a variety of situations and objects were less likely to benefit from behavioral therapy than those whose anxiety was more specific. Liddell, Difazio, Blackwood, and Ackerman (1994) conducted a 1–4 year follow-up of patients who completed a group behavioral program. At that time, 70% of patients were paying regular dental visits. Those patients who were not regular attenders at follow-up reported higher scores (on the DAS) at the completion of treatment, but unfortunately measures of generalized anxiety were not taken.

A related possibility is that therapy is more likely to succeed if the intervention provided is consonant with the main component of the patients' anxiety. Within Lang's (1971) three-factor model, anxiety has cognitive, behavioral, and physiological components. These can covary or be largely independent, varying between individuals. Harrison et al. (1985) demonstrated that intraindividual correlations between various psychological assessments of dental anxiety, mood, and physiological recordings are largely idiosyncratic. Some have argued that if an individual's anxiety is primarily physiologically based then applied relaxation methods would be most effective, while if the anxiety is primarily cognitive in nature, then such methods as self-instructional training would be more appropriate. Jerremalm, Jansson, and Ost (1986) tested this hypothesis by first assessing the nature of their patients' anxiety and then randomly assigning them to a physiological or a cognitive program. Unfortunately for the concordance hypothesis, both interventions were about equally successful, and success was not associated with a match between individual response pattern and type of intervention, a conclusion supported elsewhere (Ning & Liddell, 1991).

Another related possibility is that when behavioral treatments do not succeed it is because of a misdiagnosis. There has been some interest in the possibility that the difficulties of some patients who present with dental phobia might be better understood within the context of panic disorder. That is, perhaps some patients are not simply fearful of aspects of dental care, but rather of their own anxiety reactions (Holst et al., 1993; Weiner & Sheehan, 1990). If so, then cognitive-behavioral interventions which aim to help patients cope with their catastrophic ideation (Clark, 1986) would be helpful.

There is support for the notion that attempting a match between coping style and intervention type is helpful. Of particular interest are designs in which individuals' desire for information or control is ascertained first, and then patients are provided with an intervention which is either consonant or not consonant with their coping style. There are a number of studies on this topic (Auerbach, Kendall, Cuttler et al., 1976; Auerbach, Martelli, & Mercuri, 1983; Baron, Logan, Keeley, Law, & Stein, 1991) which suggest that when a patient is presented with a nonpreferred coping mechanism (e.g., given more information than they require for their coping style) stress levels can increase. Law, Logan, and Baron (1994) first measured patients' desire for control and their perceived coping resources, and then assigned them to either a Stress Inoculation Training (SIT) program (which included a video which emphasized ways of controlling both emotions and the dentist's behavior) or to a placebo condition, in which patients watched a video on local areas of interest. As predicted, SIT significantly reduced the amount of pain experienced during a subsequent dental appointment only for patients who initially reported a high desire for control coupled with low perceived control.

Similarly, Litt, Nye, and Shafer (1995) divided a large group of patients who were due to have molar extractions into five groups, each receiving different levels of relaxation, self-efficacy, and control interventions. The results indicated that each type of intervention had significant, and roughly equivalent, effects and that interventions that provided more coping strategies were generally superior. Importantly for the hypothesis that treatments which are consonant with coping style will be most effective, those who scored high on Miller's (1987) blunting scale received most benefit from interventions which required least personal investment (the standard preparation of minimal desensitization), while those who scored low on the blunting scale reported less distress when given the most intensive interventions (which included self-efficacy enhancement plus needle desensitization). From these studies, it seems that an intervention is more likely to succeed when it is compatible with coping style.

8.26.2.6.3 The dentist–patient relationship

The above research suggests that aspects of the relationship between the patient and therapist, or patient and the dentist may be significant. When patients want to exercise control and are given an opportunity to do so, they are less likely to experience pain and heightened anxiety. Furthermore, some patients locate the reason for their anxiety in previous negative encounters with dentists themselves (Bernstein et al., 1979). The research on temperament indicates that some patients find it more difficult than others to adapt to the dental setting, suggesting that the same individual may have more or less anxiety depending on the dentist's ability to provide a secure setting (Lochary et al., 1993). Anxious patients often make great efforts to stay with a dentist they know (Kent, 1984). In the follow-up of patients treated for high levels of anxiety, Liddell, Di Fazio et al. (1994) found that the most successful discriminator between those who managed to attend regularly and those who didn't, was perception of the amount of communication and information given by the practice dentist. Such findings indicate that a more detailed examination of the dentist–patient relationship may provide important information about the nature and treatment of high levels of anxiety.

In this context, it is interesting to note studies on the relationship between levels of anxiety and behavioral disruption on the one hand, and the behavior of the dentist on the other. In observational studies (Melamed et al., 1983; Weinstein, Getz, Ratener et al., 1982), coercive and punitive behavior on the part of dentists has been shown to increase the likelihood of disruptive or anxiety-based behaviors in children. The suggestion that punitive statements and lack of warmth can exacerbate anxiety responses has been confirmed in some experimental studies (Melamed et al., 1983; Greenbaum, Lumley, Turner, & Melamed, 1993) where dentists have been coached to behave in supportive or nonsupportive ways. In one study (Greenbaum et al., 1993), children were assigned to either a touch condition (in which the dentist gently patted the patient on the arm and provided reassuring descriptions of the treatment) or the no-touch condition (in which reassurance alone was provided). Based on analysis of video tapes of the consultations, the touched children reported significantly fewer feelings of displeasure and of being dominated, and older children (7–10 years of age) were less likely to fidget in the chair.

It is also interesting to note the comments of patients who have been provided with behavioral treatments for their anxiety. From a psychologist's perspective, the implementation of graded exposure and relaxation training, or challenging catastrophic thinking patterns are important. From the patient's perspective, however, interventions can be conceptualized quite differently. Smith, Getz, Milgrom, and Weinstein (1987) asked a sample of 24 patients, who successfully completed a cognitive-behavioral program, about which aspects helped them to tolerate treatment. Patients indicated that a number of factors were important, including the provision of information, the time taken, perceptions of control, and the dentist understanding and listening to their concerns and not treating them as "silly."

Such factors are firmly grounded in an interpersonal model of anxiety and anxiety-reduction. Perhaps it would be helpful in understanding and treating high levels of anxiety and phobia to consider these difficulties not simply as residing within the individual or in the individual's perceptions of dental care, but more within the relationship with the dental care provider.

8.26.3 PREVENTIVE CARE

As is the case for work in dental anxiety, social scientists have been interested in conducting research on health promotion for both theoretical and practical reasons. Surveys have shown that the dental health of the general population has improved considerably over recent decades. By asking similar questions at different times, changes in the prevalence of disease can be monitored. In the 1986 Office of Population Census Survey (OPCS) report (OPCS, 1986) on dental health in the UK, the prevalence of complete tooth loss (edentulism) in 1968, 1978, and 1983 was compared. Across all ages, 37% of the population was edentulous in 1968, compared with 29% in 1978 and 25% in 1983. The decline was especially noteworthy for younger people. Nevertheless, there is certainly a need for improved preventive care, since the effects of dental disease are widespread. Cushing, Sheiham, and Maizels (1986) found that 26% of adults had experienced dental pain, 20% had difficulty in eating, and 15% had problems in communication during the year prior to the survey. Edentulism can have varied and severe effects. The wearing of dentures can result in a number of difficulties, not only including problems such as inability to chew certain foods, but also a greater likelihood of mandibular dysfunctions such as headaches (Magnusson, 1980), and social problems such as refusing invitations to others' homes and avoidance of speaking (Bloomberg & Linquist, 1983).

Much of the early research on prevention was heavily influenced by sociological concepts. Attempts were made to understand preventive care by examining large populations of people using epidemiological methods. There have also been several attempts to improve dental health through educational programs and other more psychologically based approaches. This part of the chapter considers research in these areas.

8.26.3.1 Epidemiological Studies

Large scale surveys have repeatedly shown that dental health is related to social class. In the 1983 OPCS survey (OPCS, 1986), 1000 people were given extensive interviews in their own homes about a variety of topics such as education, housing, and income, as part of a national census. There was a clear relationship between social class and tooth loss. While only 8% of people from professional classes had lost all of their teeth, 44% of people who had unskilled manual occupations had done so. An important question, though, is what sociological and psychological factors associated with social class are relevant. Is it income and education, availability of professional dental care, or attitudes and behavior? For instance, families with a higher income are more likely to have a higher level of education, to have dentists who stress the importance of preventive care, and to believe in the importance of dental care for their children (Kriesberg & Treiman, 1962).

Because large numbers of people can be studied in this way, it is possible to disentangle the various factors in sophisticated matching designs. For example, Metz and Richards (1967) were interested in identifying factors which could be used to predict children's visiting patterns. Previous research had implicated parental income, parental education, and parental patterns of preventive care, but the relative importance of these factors had not been specified. Metz and Richards were able to match parents on these three variables. By comparing families with similar incomes and levels of education, they could test the hypothesis that parental patterns of visiting a dentist act as an independent factor in predicting children's behavior. Similarly, they tested the effects of income and of education by matching patients on the other variables. When this was done, all three factors were found to have independent and significant effects, but the effect of parental practices was nearly as great as the income and educational factors combined. Thus, they were able to conclude that the best predictor of whether or not a child made preventive visits to the dentist was whether the parents did so, across all income and education levels.

8.26.3.2 Improving Preventive Care

The general improvement in oral health also applies to the incidence of caries—which many researchers have attributed to the addition of fluoride in toothpaste (Burt, 1985)—but a sizable problem remains. Gingivitis (or inflammation of the gums caused by plaque due to poor toothbrushing and flossing) is very prevalent and can lead to tooth loss if patients do not visit their dentist regularly for cleaning, and although the incidence of caries has declined overall, it is still a major cause of tooth loss amongst some deprived groups in society. This section of the chapter considers educational, behavioral, and social cognition approaches to these problems.

8.26.3.2.1 The educational approach

Put simply, the educational approach holds that people do not engage in preventive behaviors because they are either unaware of the importance of doing so or do not have the necessary ability and skills. Professional cleaning can reduce plaque and is effective in eliminating gingivitis (Axelsson & Lindhe, 1974), but several surveys have shown that there is considerable lack of knowledge about dental hygiene practices in the general population. In one sample, 63% had never heard of the word plaque, 29% could not remember having been instructed on how to brush their teeth, and none could remember having been told of the significance of the frequency of eating sugar (Linn, 1974). While it is important to realise that these results are based on patients' memories (perhaps they had been told but had forgotten), such studies indicate that there is a substantial lack of information about dental care. It may be that this is partially responsible for the problem.

There have been numerous attempts to confirm the hypothesis that behavior will change after exposure to an educational program. The usual method has been first to select a sample of people, randomly assign them into two or more groups, and then provide the intervention (such as films or reading material) to one group but not others. A good example of such an experiment on dental education is provided by Horowitz, Suomi, and Peterson (1976). One group of children was given 10, 30 minute sessions on plaque removal, being taught in small groups by a dental hygienist. They were told about the significance of plaque, how to identify it with a disclosing agent, and how to remove it. Each day for the following six months plaque removal was practised under supervision, and the children were asked to rebrush and refloss any remaining plaque. The comparison group of children was not given any of this information or practice. All of this added up to a considerable amount of time, effort, and education.

The program had some short-term but few long-term effects. At an assessment immediately after the intervention, there was a significant decline in the levels of gum inflammation for the education group compared to the comparison group, but this difference disappeared at the 12 month comparison. Nor were there any difference in the number of caries or amount of plaque. It seemed that participants reverted to their previous behavior once the intervention ceased. Unfortunately, this is a typical result. While many educational programs have shown significant increases in knowledge and changes in attitude, few have demonstrated changes in behavior and fewer still improvements in oral health (Sogaard & Holst, 1988).

Rather than dismiss the educational approach, however, it is more reasonable to see information and skill development as necessary, but not sufficient, conditions for improved oral health. That is, a patient may require knowledge about prevention, but this does not guarantee a change in behavior. Perhaps, with refinements, the educational approach would be more efficacious. There is increasing evidence that after many failed attempts to affect oral health with this approach, greater success is possible. Galgut, Waite, Todd-Pokropek, and Barnby (1987) explored the possibility that willingness to change behavior and to benefit from a preventive program would be related to locus of control. As predicted, participants in a preventive program who scored highly on the internal scale on the Multidimensional Health Locus of Control (Wallston, Wallston, & De Vellis, 1978) were more likely to benefit from an educational package than those who believed that susceptibility to oral disease was a matter of chance.

Craft, Croucher, and Dickinson (1981) employed a variety of approaches in their program with adolescents in the UK, including the use of teachers at school, the involvement of parents, stressing personal responsibility, and repetition and reinforcement of the oral hygiene message. They found that, with such refinements, oral health could be improved and maintained 6–9 months after completion of the intervention. It may well be the case that this program was successful because some behavioral principles were incorporated, an approach which is considered next.

8.26.3.2.2 The behavioral approach

From a psychological perspective, a functional analysis of the issue is most helpful. As for a number of other health-related behaviors, there is little contingency between neglect and the experience of symptoms. That is, it is difficult for patients to see the relationship between behavior and consequences. Plaque on the teeth cannot be seen without staining agents, and caries usually occur in relatively hidden parts of the mouth. Children are particularly at risk because they can be much more concerned with the short term advantages of eating sugary foods and find it difficult to appreciate that their current behavior could have negative consequences in middle age.

The use of positive reinforcement to bridge the gap between behavior and consequence has been validated in a large number of studies. Whether the target behavior has been improved attendance (Reiss, Piotrowski, & Bailey, 1976) or improved oral health (Iwata & Becksfort, 1981; Marten, Frazier, & Kirt, 1973), material rewards have been efficacious. In some studies these rewards have been substantial, but even inexpensive reinforcements such as placing stars on a calendar each time brushing or flossing takes place can be effective. Furthermore, changes in behavior tend to be maintained even after the incentives are withdrawn. Children's behavior is particularly amenable to behavioral change, but every attempt should be made to involve parents.

Weinstein and Getz (1978) provide a detailed analysis of behavioral change in this context. Although large and clinically significant changes in behavior and oral hygiene are unlikely to be achieved, they argue that the chances of success can be maximized by following a series of steps which are consistent with the behavioral approach. Below are the six steps they outline.

(i) *Formulating a clear and precise definition of the problem.* Many people are unaware of the basic facts and skills needed for adequate prevention. A rather detailed assessment may be needed so that the gaps in knowledge and skills may be identified and filled. However, the skills may not be put into practice, so that the problem may be seen as one of behavior and consequences. Specific behavioral definitions are required so that patients will know which aspects of their oral hygiene patterns need to be modified. For example, if the patient has severe gingivitis, it is likely that brushing and flossing need attention. The presence of new caries is indicative of the consumption of many cariogenic snacks and drinks, so diet must be improved.

(ii) *Monitoring the frequency of the behavior.* Once the broad nature of the specific difficulty is identified, information about current behavior is required. This can be accomplished by charting, which involves asking the patient to note each time he or she engages in the specified behavior. This helps to identify the scope of the problem, provides a baseline for assessing the effectiveness of the intervention, and provides feedback for the patient.

(iii) *Specifying the aims of the intervention.* Just as it is important to specify the nature of the problem, so too it is important to consider what the patient aims to achieve. The target behavior needs to be achievable within a specified time limit.

(iv) *Changing the behavior.* However, the target may not be easily achieved. Studies have shown that shaping can be effective in this context. If the aim were to improve diet, the patient is asked to cut down snack consumption from the baseline level gradually (from five to four snacks per day, for example) and, once this has been accomplished, a further reduction would be required. Children can be shown to brush competently by 3 or 4 years of age if the task is broken down into sufficiently small steps (Poch, McCubbrey, & Munn, 1982) as can people who have suffered brain injuries (Zasler, Devary, Jarman, & Friedman, 1993).

Shaping can be particularly useful in helping people with learning disabilities to look after their oral health, which could form a part of an overall program of self-care. Horner and Keilitz (1975) first ascertained what behaviors were necessary by conducting a task analysis, in this instance by videotaping a competent person brushing his teeth. From the tape they identified several small steps:

(a) pick up and hold the toothbrush,
(b) wet the toothbrush,
(c) remove cap from toothpaste,
(d) apply toothpaste to brush, and
(e) replace cap on the toothpaste.

Steps (f)–(k) involved brushing various parts of the mouth.

(l) rinse the toothbrush,
(m) rinse the sink,
(n) put equipment away, and
(o) discard any paper cups or tissues used.

When a step had been learned, it was chained to the next on the list of tasks.

(v) *Providing reinforcement.* It is important to provide reinforcement for each successful attempt at change. The nature of the reinforcement can often be left up to the patient: for one person it may be a trip to the theatre, for another the purchase of a small gift for themselves. Children can also choose their rewards, such as the choice of the dinner menu once a week.

(vi) *Failures in preventive programs.* There are several reasons why such a program might fail. Lack of support from family, inappropriate choice of reinforcement, and lack of motivation are some possibilities. Weinstein and Getz (1980) suggest that the use of a contract, which is a public commitment to change, can be helpful.

8.26.3.2.3 Social cognition models

Social cognition models share many characteristics of both the educational and behavioral models. Like educational approaches, they are concerned with beliefs, but, like the behavioral approaches, they are also centrally interested in predicting how people act. The Health Belief Model (HBM) (Kegeles, Lund, & Weisenberg, 1978) has been very influential. It postulates that the likelihood that someone will adopt a recommended health action is affected by perceived vulnerability (belief in the personal susceptibility to disease), seriousness (if the disease were to develop), preventability (of the disease if action were taken), salience (the relative importance of taking preventive actions compared to other behaviors), and readiness (the belief that action ought to be taken promptly rather than defered to a later date).

Numerous studies have found that people who have taken preventive actions are more likely to hold these beliefs that those who have not (i.e., the HBM describes behavior after it has occurred (Kegeles, 1963)). However, there is relatively little prospective work. Furthermore, a test of a model should include the prediction of a novel type of behavior, rather than one which is already established. This is necessary because beliefs and attitudes might follow behavior rather than precede it: a person may come to believe that preventive care is important because they have always brushed or flossed their teeth. This could have been due to parental encouragement when they were young, for example. In order for behavior and beliefs to be consistent, people might change their beliefs to correspond with their previous actions.

The few studies which have met the criteria of both prediction of behavior and an attempt to instigate novel behavior have not supported the HBM. Weinstein, Kegeles, and Lund (1980) reported a study in which children were asked to participate in a program which involved a novel preventive action, in this case topical fluoride application. After the treatment and its effects had been explained to all the children, their health beliefs were measured. Unfortunately for the HBM, the children who scored high on vulnerability, on seriousness, and on the effectiveness of treatment were no more likely to take part. In fact, where the results showed a relationship, they were opposite to the predicted direction. It seems, from this study, that general health beliefs are poor predictors of future preventative behavior.

Kegeles et al. (1978) combined the HBM with other approaches. They used the educational, HBM, and behavioral approaches in an attempt to encourage children to use a mouthrinse containing fluoride. First, all children in the study were given a slide show which explained the etiology of dental disease, the health and cosmetic consequences of decay, and the effectiveness of fluoride as a preventive measure. The children were then assigned to one of three groups. Those in one group were given no further assistance, except that they were allowed to ask questions about the slide show. Those in the second group were given two discussion periods on the HBM, exploring ideas of seriousness, vulnerability, and preventability. The children in the final group were offered some small cash incentives if they volunteered for and continued through to the end of the treatment program.

The fluoride treatment was then offered to all three groups. It involved using a mouthrinse at home twice a day from a bottle which contained a 14 day supply. The bottles were designed so that the children could not simply dispose of the contents all at once: a stopper ensured that they would be ready for use only at six-hourly intervals. Every two weeks the bottles were to be returned, when the incentive group children were given a small reward, and the next bottle picked up. The results indicated that the education plus incentive intervention was the most effective. Although initial uptake in the program was similar at the begining, with about 95% of the children initially volunteering to take part, participation declined over time. This decline was greatest for the discussion group and least for the incentive group. For example, 49% of the children given an incentive picked up the last bottle, 31% of the children in the education group did so, and only 18% of the children in the HBM group. That the HBM discussion group was the least effective is somewhat surprising, but a similar result was found in another study in which the treatment involved topical fluoride application (Lund, Kegeles, & Weinsenberg, 1977).

More recently there has been an increased interest in applying Fishbein and Ajzenn's early theory of reasoned action (TRA) (Ajzenn & Fishbein, 1980) and later theory of planned behavior (TPB) (Ajzenn, 1985), which incorporates beliefs about self-efficacy in affecting outcomes (Bandura, 1986) to preventive dental care (Tedesco, Keffer, & Fleck-Kandath, 1991).

McCaul, O'Neill, and Glasgow (1988) tested the ability of the TPB to explain brushing and flossing behavior in a sample of 77 college students. McCaul et al. took great care in taking baseline measures: the students' levels of oral health (plaque levels and gingival health) were checked by a dental hygienist, skills were assessed, and self-reports of brushing and flossing frequencies taken. The students completed several questionnaires designed to assess each component of the TPB (see Chapter 8.01, this volume). Over the next several weeks, participants were then asked to record the frequency with which they brushed and flossed in a self-monitoring booklet. The results provided some support for the TRA. About 30% of the variance in self-reported intentions to brush and floss could be accounted for by attitudes and norms, and 27% and 37% of the variances in self-reported brushing and flossing behavior, respectively, could be accounted for by intentions. The measure of self-efficacy did not add significantly to the prediction of behavior.

Tedesco, Keffer, Davis, and Christersson (1992) argued that self-efficacy could have a significant influence on preventive care if it were specifically targeted in an intervention. They randomly assigned participants (recruited through a variety of means) to either a traditional hygiene program (which included instruction on brushing and flossing techniques, and feedback on their oral health), or to an experimental group (which received the traditional program plus encouragement that their oral health was under their control and praise for their accomplishments). The participants completed several questionnaires designed to assess the seven basic dimensions of the TRA, self-efficacy, and self-reports of brushing and flossing frequencies. In addition, all were given an oral examination, when plaque and gingivitis indices were taken. Over the next several visits, the control group continued to be given feedback on oral health, while the successes of the intervention group were emphasized.

Both plaque and gingival indices showed improvement over the visits, with somewhat better oral health for the intervention group. At six months after completion of the intervention, the oral health of the experimental group was significantly better, but this difference disappeared at a final follow-up visit six months later. The authors concluded that the intervention delayed the relapse. However, there were no differences on the cognitive or self-report variables over time or between groups. Self-efficacy increased during intervention visits and decreased at the final follow-up.

In another report which focuses on the intervention group only, Tedesco, Keffer, Davis, and Christersson (1993) report a series of correlations between cognitive variables, self-reported behavior, and clinical indices. The correlations between self-reported behavior and social cognitions were higher, accounting for up to 51% of the variance in self-reported behavior on some occasions. When self-efficacy measures were included, the amount of variance in reported behavior accounted for increased further.

In summary, the educational, behavioral, and social cognition approaches to preventive care each show promise. Educational approaches, though intuitively attractive, do not appear sufficient in themselves for behavioral change. It seems that they are more effective when combined with behavioral methods. These latter approaches have been given the most support, particularly when they are applied on an individual basis. Two further issues should be raised. One is the reliance on self-reported behavior as an outcome measure. It is always possible that self-reports of behavior are not valid indicators of actual behavior. Although there is little work on the relationship between self-reports and actual preventive behavior in dentistry, there are indications that patients' reports of attendance, for example, do not correspond to visiting patterns (Eddie, 1984). Similarly, self-reports of brushing and flossing may not be accurate reflections of preventive care. Second, none of these approaches include the nature of the interpersonal relationship between dentist and patient, which may be of crucial importance (Sandell, Camner, & Sarhed, 1994).

8.26.4 CONCLUSIONS

This brief description of the research on dental anxiety and preventive dental care makes it clear that studies on these topics have had practical implications for the well-being of dental patients and have been very fruitful in theoretical terms. There is no doubt that, since the early 1970s, psychologists have helped to make significant improvements to the well-being of dental patients. This improvement is partly due to the application of psychological principles to dental care, and partly due to the rigor that psychologists have brought to the assessment of a variety of issues. Psychologists have helped to identify the difficulties faced by dental patients and have developed interventions which are effective in reducing the significant difficulties which many patients encounter. New treatment programs which are based on cognitive-behavioral approaches, and which target negative cognitions, have been of special benefit. Further studies have suggested that preventive programs can also benefit from the use of psychological assessments and interventions.

There are many other ways in which psychology can contribute. In particular, there is a need for assessments which include aspects of quality of life and an exploration of the ways in which dental disease can affect patients' lives outside of the dental setting. Some consensus on the applicability of particular assessments would also aid studies of prevalence. For the understanding of etiology, longitudinal studies are clearly required. While there are many studies based on retrospective accounts, following patients through their dental histories is likely to be the only method which could tease out issues of cause and effect. This may be particularly true when attempting to improve preventive care.

8.26.5 REFERENCES

Ajzenn, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckman (Eds.), *Action-control: From cognition to behavior*. Berlin, New York: Springer-Verlag.

Ajzenn, I. & Fishbein, M., (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Alwin, N., Murray, J. J., & Niven, N. (1994). The effect of children's dental anxiety on the behavior of the dentist. *International Journal of Pediatric Dentistry, 4,* 19–24.

American Psychiatric Association (1995). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Anderson, R. A., Baron, R., & Logan, H. (1991). Distraction, control, and dental stress. *Journal of Applied Social Psychology, 2,* 156–171.

Arntz, A., Eck, M. van, & Heijmans, M. (1990). Predictions of dental pain: the fear of any expected evil is worse than the evil itself. *Behavior Research and Therapy, 28,* 29–41.

Arntz, A., Eck, M. van, & Jong, P. de (1991). Avoidance of pain of unpredictable intensity. *Behavior Research and Therapy, 29,* 197–201.

Arntz, A., Eck, M. van, Jong, P. de, & Hout, M. van den (1990). The relationship between unpredicted pain and escape. *behavior Research and Therapy, 28,* 87–90.

Arntz, A., Hout, M. van den, Berg, G. van den, & Meijboom, A. (1991). The effects of incorrect pain expectations on acquired fear and pain responses. *Behavior Research and Therapy, 29,* 547–560.

Arntz, A., & Lousberg, R. (1990). The effects of underestimated pain and their relationship to dishabituation. *Behavior Research and Therapy, 28,* 15–28.

Auerbach, S., Kendall, P., Cuttler, H., et al. (1976). Anxiety, locus of control, type of preparatory information, and adjustment to dental surgery. *Journal of Consulting and Clinical Psychology, 44,* 809–818.

Auerbach, S., Martelli, M., & Mercuri, L. (1983). Anxiety, information, interpersonal impacts, and adjustment to a stressful health care situation. *Journal of Personality and Social Psychology, 44,* 1284–1296.

Axelsson, P., & Lindhe, J. (1974). The effect of a preventive programme on dental plaque, gingivitis, and caries in schoolchildren. Results after one and two years. *Journal of Clinical Periodontology, 1,* 126–138.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37,* 122–147.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Baron, R., Logan, H., & Hoppe, S. (1993). Emotional and sensory focus as mediators of dental pain among patients differing in desired and felt dental control. *Health Psychology, 12,* 381–389.

Baron, R. S., Logan, H., Keeley, K., Law, A., & Stein, S. (1991). Desired and felt control as mediators of stress in a dental setting. *Health Psychology, 10,* 352–359.

Bedi, R., Sutcliffe, P., Donnan, P., Barrett, N., & McConnachie, J. (1992). Dental caries experience and prevalence of children afraid of dental treatment. *Community Dentistry and Oral Epidemiology, 20,* 368–371.

Benjamins, C., Schuurs, A., Asscheman, H., & Hoogstraten, J. (1990). Anxiety and blood pressure prior to dental treatment. *Psychological Reports, 67,* 371–377.

Berggren, U. (1993). Psychosocial effects associated with dental fear in adult dental patients with avoidance behaviors. *Psychology and Health, 8,* 185–196.

Berggren, U., & Carlsson, S. (1985). Usefulness of two psychometric scales in Swedish patients with severe dental fear. *Community Dentistry and Oral Epidemiology, 13,* 70–74.

Berggren, U., & Carlsson, S. (1986). Qualitative and quantitative effects of treatment for dental fear and avoidance. *Anaesthesia Progress, 33,* 9–13.

Bernstein, D., Kleinknecht, R., & Alexander, L. (1979). Antecedents of dental fear. *Journal of Public Health Dentistry, 39,* 113–124.

Bloomberg, S., & Linquist, L. (1983). Psychological reactions to edentulousness and treatment with jawbone-anchored bridges. *Acta Psychiatrica Scandinavica, 68,* 251–262.

Brown, D., Wright, F., & McMurray, N. (1986). Psychological and behavioral factors associated with dental anxiety in children. *Journal of Behavioral Medicine, 9,* 213–218.

Brown, J. P. (1980). The efficacy and economy of comprehensive dental care for handicapped children. *International Dental Journal, 30,* 14–27.

Burt, B. (1985). The future of the caries decline. *Journal of Public Health Dentistry, 45,* 261–269.

Carlsson, S., Linde, A., Berggren, U., & Harrison, J. A. (1986). Reduction of dental fear: psychophysiological correlates. *Community Dentistry and Oral Epidemiology, 14,* 253–257.

Clark, D. M. (1986). A cognitive approach to panic. *Behavior Research and Therapy, 24,* 461–470.

Clevenger, W., Wigal, T., Salvati, N., Burchill, R., & Crinella, F. M. (1994). Dental needs of persons with developmental disabilities in Orange County. *Journal of Developmental and Physical Disabilities, 6,* 299–309.

Cohen, E. S., & Hillis, R. (1979). The use of hypnosis in treating the temporomandibular joint pain dysfunction syndrome. *Oral Surgery, 48,* 193–197.

Cooper, C., Watts, J., Baglioni, A., & Kelly, M. (1988). Occupational stress amongst general practice dentists. *Journal of Occupational Psychology, 61,* 163–174.

Corah, N. L. (1968). Development of a dental anxiety scale. *Journal of Dental Research, 48,* 596.

Corah, N. L., Gale, E., & Illig, S. (1978). Assessment of a dental anxiety scale. *Journal of the American Dental Association, 97,* 816–819.

Corkey, B., & Freeman, R. (1994). Predictors of dental anxiety in six-year-old children: Findings from a pilot study. *Journal of Dentistry for Children, 61,* 267–271.

Craft, M., Croucher, R., & Dickinson, J. (1981). Preventive dental health in adolescents: short and long term pupil response to trials of an integrated curriculum package. *Community Dentistry and Oral Epidemiology, 9,* 199–206.

Curson, I., & Coplans, M. (1970). The need for sedation in conservative dentistry. *British Dental Journal, 125,* 816–819.

Cushing, A., Sheiham, A., & Maizels, J. (1986). Developing socio-dental indicators—the social impact of dental disease. *Community Dental Health, 3,* 3–17.

Davey, G. (1989). Dental phobias and anxieties: evidence for conditioning processes in the acquisition and modulation of a learned fear. *Behavior Research and Therapy, 27,* 51–58.

Eddie, S. (1984). Frequency of attendance in the General Dental Service in Scotland—a comparison with claimed attendance. *British Dental Journal, 157,* 267–270.

Ettinger, R. (1984). Clinical decision making in the dental treatment of the elderly. *Gerodontology, 3,* 157–165.

Frazer, M., & Hampson, S. (1988). Some personality factors related to dental anxiety and fear of pain. *British Dental Journal, 165,* 436–439.

Galgut, P., Waite, I., Todd-Pokropek, A., & Barnby, G. (1987). The relationship between the multidimensional health locus of control and the performance of subjects on a preventive periodontal programme. *Journal of Clinical Periodontology, 14,* 171–175.

Gessell, A. H., & Alderman, M. (1971). Management of myofascial pain dysfunction syndrome of the temporomandibular joint by tension control training. *Psychosomatics, 12,* 302–309.

Gokli, M., Wood, A. J., Mourino, A., Farrington, F., & Best, A. (1994). Hypnosis as an adjunct to the administration of local anaesthetic in pediatric patients. *Journal of Dentistry for Children, 61,* 272–275.

Greenbaum, P., Lumley, M., Turner, C., & Melamed, B. G. (1993). Dentist's reassuring touch: effects on children's behavior. *Pediatric Dentistry, 15,* 20–24.

Hakeberg, M., & Berggren, U. (1992). Changes in sick leave among Swedish dental patients after treatment for dental fear. *Community Dental Health, 10,* 23–29.

Hakeberg, M., Berggren, U., & Carlsson, S. (1992). Prevalence of dental anxiety in an adult population in a major urban area in Sweden. *Community Dentistry and Oral Epidemiology, 20,* 97–101.

Hakeberg, M., Berggren, U., & Grondahl, H. (1993). A radiographic study of dental health in adult patients with dental anxiety. *Community Dentistry and Oral Epidemiology, 21,* 27–30.

Harrison, J. A., Carlsson, S., & Berggren, U. (1985). Research in clinical process and outcome methodology: psychophysiology, systematic desensitization, and dental fear. *Journal of Behavior Therapy and Experimental Psychiatry, 16,* 201–209.

Hilliardlysen, L., & Riemer, J. (1988). Occupational stress and suicide among dentists. *Deviant Behavior, 9,* 333–346.

Hirschman, R., Revland, P., Hawk, G., & Young, D. (1969). Effects of dental anxiety and phase of treatment on discomfort during dental stimulation. *Journal of Dental Research, 48,* 444–447.

Holst, A., Hallonsten, A., Schroder, U., Ek, L., & Edlund, K. (1993). Prediction of behavior-management problems in 3-year-old children. *Scandinavian Journal of Dental Research, 101,* 110–114.

Horner, R., & Keilitz, I. (1975). Training mentally retarded adolescents to brush their teeth. *Journal of Applied Behavioral Analysis, 8,* 301–309.

Horowitz, A., Suomi, J., & Peterson, J. (1976). Effects of daily supervised daily dental plaque removal by children. *Journal of Public Health Dentistry, 36,* 193–200.

Horst, G. ter, & Wit, C. de (1993). Review of behavioral research in dentistry 1987–1992: Dental anxiety, dentist–patient relationship, compliance, and attendance. *International Dental Journal, 43,* 265–276.

Hugdhal, K., & Ost, L. (1985). Subjectively rated physiological and cognitive symptoms in six different clinical phobias. *Personality and Individual Differences, 6,* 175–188.

Iwata, B., & Becksfort, C. (1981). Behavioral research in preventive dentistry. *Journal of Applied Behavioral Analysis, 14,* 111–120.

Jackson, C., & Lindsay, S. (1995). Reducing anxiety in new dental patients by means of leaflets. *British Dental Journal, 179,* 163–167.

Jerremalm, A., Jansson, L., & Ost, L. (1986). Individual response patterns and the effects of different behavioral methods in the treatment of dental phobia. *Behavior Research and Therapy, 24,* 587–596.

Johnson, R., & Baldwin, D. (1968). Relationship of maternal anxiety to the behavior of young children undergoing dental extraction. *Journal of Dental Research, 47,* 801–805.

Jong, P. de, Hout, M. van den, & Merckelbach, H. (1995). Covariation bias and the return of fear. *Behavior Research and Therapy, 33,* 211–213.

Jongh, A. de, & Horst, G. ter (1993). What do anxious patients think? An exploratory investigation of anxious dental patients' thoughts. *Community Dentistry and Oral Epidemiology, 21,* 221–223.

Jongh, A. de, Muris, P., Horst, G. ter, Van-Zuren, F., & Wit, C. de (1994). Cognitive correlates of dental anxiety. *Journal of Dental Research, 73,* 561–566.

Jongh, A. de, Muris, N., Schoenmakers, N., & Horst, G. ter (1995). Negative cognitions of dental phobics: reliability and validity of the dental cognitions questionnaire. *Behavior Research and Therapy, 33,* 507–515.

Jongh, A. de, & Stouthard, M. (1993). Anxiety about dental hygienist treatment. *Community Dentistry and Oral Epidemiology, 21,* 91–95.

Kaufman, E., Weinstein, P., & Milgrom, P. (1984). Difficulties in achieving local anesthesia. *Journal of the American Dental Association, 108,* 205–208.

Kaufman, E., Rand, R. S., Gordon, M., & Cohen, H. S. (1992). Dental anxiety and oral health in young Israeli male adults. *Community Dental Health, 9,* 125–132.

Kegeles, S. (1963). Some motives for seeking preventive health care. *Journal of the American Dental Association, 67,* 90–98.

Kegeles, S., Lund, A., & Weisenberg, M. (1978). Acccep-tance by children of a daily home mouthrinse program. *Social Science and Medicine, 12,* 199–210.

Kent, G. (1984). Satisfaction with dental care—its relationship to utilization and allegiance. *Medical Care, 22,* 583–585.

Kent, G. (1985a). Hypnosis in dentistry. *British Journal of Experimental and Clinical Hypnosis, 3,* 103–112.

Kent, G. (1985b). Memory of dental pain. *Pain, 21,* 187–194.

Kent, G. (1986). The typicality of therapeutic surprises. *Behavior Research and Therapy, 24,* 625–628.

Kent, G. (1987a). Self-efficacious control over reported physiological, cognitive, and behavioral symptoms of dental anxiety. *Behavior Research and Therapy, 25,* 341–347.

Kent, G. (1987b). Stress amongst dentists. In R. Payne & J. Firth-Cozens (Eds), *Stress in Health Professionals* (pp. 127–149). Chichester, UK: Wiley.

Kent, G. (1989). Memory of dental experiences as related to naturally occuring changes in state anxiety. *Cognition and Emotion, 3,* 45–53.

Kent, G., (1991). *The psychology of dental care.* Oxford: Wright and Sons.

Kent, G., & Gibbons, R. (1987). Self-efficacy and the control of anxious cognitions. *Journal of Behavior Therapy and Experimental Psychiatry, 18,* 33–40.

Kent, G., & Johns, R. (1993). Psychological effects of permanently implanted false teeth: a 2-year follow-up and comparison with dentate patients. *Psychology and Health, 8,* 213–222.

Kent, G., Rubin, G., Getz, T., & Humphris, G. (1996). The development of a scale to measure the social and psychological effects of severe dental anxiety: Social attributes of the Dental Anxiety Scale. *Community Dentistry and Oral Epidemiology, 24,* 394–397.

Kent, G., & Warren, P. (1985). A study of factors associated with changes in dental anxiety. *Journal of Dental Research, 64,* 1316–1318.

Kiyak, H. A. (1993). Psychological aspects of orthognathic surgery. *Psychology and Health, 8,* 197–212.

Kleinknecht, R., & Bernstein, D. (1979). Short term treatment of dental avoidance. *Behavior Research and Therapy, 10,* 311–315.

Kleinknecht, R., Klepac, R., & Alexander, L. (1973). Origins and characteristics of fear of dentistry. *Journal of the American Dental Association, 86,* 842–848.

Klepac, R., Dowling, J., & Hauge, G. (1982). Characteristics of clients seeking therapy for the reduction of dental avoidance: reactions to pain. *Journal of Behavior Therapy and Experimental Psychiatry, 13,* 293–300.

Klingberg, G. (1994). Reliability and validity of the Swedish version of the Dental Subscale of the Children's Fear Survey Schedule, CFSS-DS. *Acta Odontologica Scandinavica, 52,* 255–256.

Klingberg, G., Berggren, U., & Noren, J. G. (1994). Dental fear in an urban Swedish child population: prevalence and concomitant factors. *Community Dental Health, 11,* 208–214.

Klingberg, G., & Hwang, C. P. (1994). Children's dental fear picture test (CDFP): A projective test for the assessment of dental fear. *Journal of Dentistry for Children, 61,* 89–96.

Klingberg, G., Lofqvist, L., Bjarnason, S., & Noren, J. (1994). Dental behavior management problems in Swedish children. *Community Dentistry and Oral Epidemiology, 22,* 201–205.

Koenigsberg, S., & Johnson, R. (1972). Child behavior during sequential dental visits. *Journal of the American Dental Association, 85,* 128–132.

Kriesberg, L., & Treiman, B. (1962). Preventive utilisation of dentist services among teenagers. *Journal of the American College of Dentistry, 29,* 28–45.

Kupietzky, A., & Houpt, M. (1993). Midazolam: a review of its use for conscious sedation of children. *Pediatric Dentistry, 15,* 237–241.

Lang, P. (1971). The application of psychophysiological methods to the study of psychotherapy and behavior modification. In A. Bergin & S. Garfield (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis*. New York: J. Wiley and Sons.

Lautch, H. (1971). Dental phobia. *British Journal of Psychiatry, 119,* 151–158.

Law, A., Logan, H., & Baron, R. S. (1994). Desire for control, felt control, and Stress Inoculation Training during dental treatment. *Journal of Personality and Social Psychology, 67,* 926–936.

Liddell, A. (1990). Personality characteristics versus medical and dental experiences of dentally anxious children. *Journal of Behavioral Medicine, 13,* 183–194.

Liddell, A., Di Fazio, L., Blackwood, J., & Ackerman, C. (1994). Long-term follow-up of treated dental phobics. *Behavior Research and Therapy, 32,* 605–610.

Liddell, A., & Locker, D. (1993). Dental anxiety in the elderly. *Psychology and Health, 8,* 175–183.

Liddell, A., & Locker, D. (1994). Disproportionate dental anxiety in the over 50s. *Journal of Behavior Therapy and Experimental Psychiatry, 25,* 211–216.

Lindsay, S., & Jackson, C. (1993). Fear of routine dental treatment in adults: its nature and management. *Psychology and Health, 8,* 135–153.

Linn, E. (1974). What dental patients don't know about dental care. *Journal of Public Health Dentistry, 34,* 39–41.

Litt, M. D., Nye, C., & Shafer, D. (1993). Coping with oral surgery by self-efficacy enhancement and perceptions of control. *Journal of Dental Research, 72,* 1237–1243.

Litt, M. D., Nye, C., & Shafer, D. (1995). Preparation for oral surgery: evaluating elements of coping. *Journal of Behavioral Medicine, 18,* 435–459.

Lochary, M., Wilson, S., Griffen, A., & Coury, D. (1993). Temperament as a predictor of behavior for conscious sedation in dentistry. *Pediatric Dentistry, 15,* 348–352.

Locker, D., & Liddell, A. (1992). Correlates of dental anxiety among older adults. *Community Dentistry and Oral Epidemiology, 20,* 372–375.

Locker, D., & Slade, G. (1988). Prevalence of symptoms associated with temporomandibular disorders in a Canadian population. *Community Dentistry and Oral Epidemiology, 16,* 310–313.

Lund, A., Kegeles, S., & Weisenberg, M. (1977). Motivational techniques for increasing acceptance of preventive health measures. *Medical Care, 15,* 678–692.

Magnusson, T. (1980). Prevalence of recurrent headache and mandibular dysfunction in patients. *Community Dentistry and Oral Epidemiology, 8,* 159–164.

Marshall, W. L., Gauthier, J., & Gordon, A. (1980). The current status of flooding therapy. In M. Hersen, R. Eisler, & P. Miller (Eds.), *Progress in behavior modification 9*. New York: Academic Press.

Marten, L., Frazier, P., & Kirt, K. (1973). Developing brushing performance in second graders through behavior modification. *Health Services Reports, 88,* 818–823.

McCaul, K., O'Neill, H. K., & Glasgow, R. (1988). Predicting the performance of dental hygiene behaviors: An examination of the Fishbein and Ajzen model and self-efficacy expectations. *Journal of Applied Social Psychology, 18,* 114–128.

McNeil, D., & Berryman, M. (1989). Components of dental fear in adults. *Behavior Research and Therapy, 27,* 233–236.

Melamed, B. G. (1979). Behavioral approaches to fear in dental settings. *Progress in Behavior Modification, 7,* 171–203.

Melamed, B. G., Ross, S., Courts, F., Bennett, C., Bush, J., Ronk, S., Jerrell, G., & Hill, C. (1983). Dentists' behavior management as it affects compliance and fear in pediatric patients. *Journal of the American Dental Association, 106,* 324–330.

Mellor, A. C. (1992). Dental anxiety and attendance in the north west of England. *Journal of Dentistry, 20,* 207–210.

Metz, A., & Richards, L. (1967). Children's preventive dental visits. *Journal of the American College of Dentistry, 34,* 204–212.

Milgrom, P., Mancl, L., King, B., & Weinstein, P. (1995). Origins of childhood dental fear. *Behavior Research and Therapy, 33,* 313–319.

Milgrom, P., Weinstein, P., & Getz, T. (1995). *Treating fearful dental patients: A patient management handbook.* (2nd ed.). Seattle, WA: University of Washington.

Miller, N. E. (1948). Studies of fear as an acquired drive: 1. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *Journal of Experimental Psychology, 38,* 89–101.

Miller, N. E. (1959). Liberalization of basic S-R concepts: Extensions to conflict-behavior, motivation and social learning. In S. Koch (Ed.), Psychology: A study of a science (pp. 196–292). New York: McGraw-Hill.

Miller, S. M. (1987). Monitoring and blunting: validation of a questionnaire to assess styles of information seeking under threat. *Journal of Personality and Social Psychology, 52,* 345–353.

Moore, R., Berggren, U., & Carlsson, S. (1991). Reliability and clinical usefulness of psychometric measures in a self-refered population of odontophobics. *Community Dentistry and Oral Epidemiology, 19,* 347–351.

Moore, R., Birn, H., Kirkegaard, E., Brodsgaard, I., & Scheutz, F. (1993). Prevalence and characteristics of dental anxiety in Danish adults. *Community Dentistry and Oral Epidemiology, 21,* 292–296.

Moore, R., & Brodsgaard, I. (1994). Group therapy compared with individual desensitization for dental anxiety. *Community Dentistry and Oral Epidemiology, 22,* 258–262.

Moore, R., Brodsgaard, I., Berggren, U., & Carlsson, S. (1991). Generalization of effects of dental fear treatment in a self-refered population of odontophobics. *Journal of Behavior Therapy and Experimental Psychiatry, 22,* 243–253.

Moore, R., Brodsgaard, I., & Birn, H. (1991). Manifestations, acquisition, and diagnostic categories of dental fear in a self-refered population. *Behavior Research and Therapy, 29,* 51–60.

Moraes, A. de, Milgrom, P., Tay, K., & Costa, S. (1994). Prevalence of dental fear in Brazilian high school students in Sao Paulo state. *Community Dentistry and Oral Epidemiology, 22,* 114–115.

Mowrer, O. H. (1960). *Learning theory and behavior*. New York: J. Wiley and Sons.

Muris, P., Merckelbach, H., & Jongh, A. de (1995). Color-naming of dentist-related words: role of coping style, dental anxiety and trait anxiety. *Personality and Individual Differences, 18,* 685–688.

Murray, P., Liddell, A., & Donohue, J. (1989). A longitudinal study of the contribution of dental experience to dental anxiety in children between 9 and 12 years of age. *Journal of Behavioral Medicine, 12,* 309–320.

Ning, L., & Liddell, A. (1991). The effect of concordance in the treatment of clients with dental anxiety. *Behavior Research and Therapy, 29,* 315–322.

OPCS. (1986). The 1983 update on adult dental health from OPCS. *British Dental Journal, 160,* 246–253.

Ost, L. (1987). Age of onset in different phobias. *Journal of Abnormal Psychology, 96,* 223–229.

Poche, C., McCubbrey, H., & Munn, T. (1982). The development of correct toothbrushing technique in preschool children. *Journal of Applied Behavior Analysis, 15,* 315–320.

Pomp, A. (1974). Psychotherapy for the myofacial pain dysfunction syndrome: a study of factors coinciding with symptom remission. *Journal of the American Dental Association, 89,* 629–632.

Radis, F., Wilson, S., Griffen, A., & Coury, D. (1994). Temperament as a predictor of behavior during initial dental examination in children. *Pediatric Dentistry, 16,* 121–127.

Reiss, M., Piotrowski, W., & Bailey, J. (1976). Behavioral community psychology: encouraging low income patients to seek dental care for their children. *Journal of Applied Behavioral Analysis, 9,* 387–397.

Roberts, G., Gibson, A., Porter, J. et al. (1979). Relative analgesia: an evaluation of safety. *British Dental Journal, 146,* 177–182.

Robertson, C., Gatchel, R., & Fowler, C. (1991). Effectiveness of a videotaped behavioral intervention in reducing anxiety in emergency oral surgery patients. *Behavioral Medicine, 17,* 77–85.

Ross, L. (1964). The intuitive psychologist and his shortcomings. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10). New York: Academic Press.

Roy-Bryne, P., Milgrom, P., Khoon-Mei, T., Weinstein, P., & Katon, W. (1994). Psychopathology and psychiatric diagnosis in subjects with dental phobia. *Journal of Anxiety Disorders, 8,* 19–31.

Rugh, J. (1977). A behavioral approach to diagnosis and treatment of functional and oral disorders: biofeedback and self-control techniques. In J. Rugh, D. Perlis, & R. Disraeli (Eds.), *Biofeedback in dentistry: Research and clinical applications*. Phoenix, AZ: Semantodontics.

Rugh, J., & Solberg, W. (1976). Psychological implications in temporomandibular pain and dysfunction. *Oral Science Review, 7,* 3–30.

Sandell, R., Camner, L., & Sarhed, G. (1994). The dentists' attitudes and their interaction with patient involvement in oral hygiene compliance. *British Journal of Clinical Psychology, 33,* 549–558.

Schuurs, A. H., & Hoogstraten, J. (1993). Appraisal of dental anxiety and fear questionnaires: a review. *Community Dentistry and Oral Epidemiology, 21,* 329–339.

Schuurs, A. H. B., Makkes, P. C., & Duivenvoorden, H. (1995). Attendance pattern of anxiety-treated patients: a pilot study. *Community Dentistry and Oral Epidemiology, 23,* 221–223.

Silva, P. de (1988). Phobias and preparedness: replication and extension. *Behavior Research and Therapy, 26,* 97–98.

Smith, T., Getz, T., Milgrom, P., & Weinstein, P. (1987). Evaluation of treatment at a dental fears research clinic. *Special Care in Dentistry, 7,* 130–134.

Smith, T., Weinstein, P., Milgrom, P., & Getz, T. (1984). An initial evaluation of an institution-based dental fears clinic. *Journal of Dental Research, 63,* 272.

Sogaard, A., & Holst, D. (1988). The effect of different school based dental health education programmes in Norway. *Community Dental Health, 5,* 169–184.

Steptoe, A., & Appels, A. (Eds.) (1989). *Stress, personal control and health*. Chichester, UK: Wiley.

Stewart, J., Marcus, M., Christenson, P., & Lin, W. (1994). Comprehensive treatment among dental school patients with high and low dental anxiety. *Journal of Dental Education, 58,* 697–700.

Stouthard, M., Hoogstraten, J., & Mellenbergh, G. (1995). A study on the convergent and discriminant validity of the dental anxiety inventory. *Behavior Research and Therapy, 33,* 589–595.

Strauss, R., & Hunt, R. (1993). Understanding the value of teeth to older adults: Influences on the quality of life. *Journal of the American Dental Association, 124,* 105–110.

Tedesco, L., Keffer, M., Davis, E., & Christersson, L. (1992). Effect of a social cognitive intervention on oral health status, behavior reports, and cognitions. *Journal of Periodontology, 63,* 567–575.

Tedesco, L., Keffer, M., Davis, E., & Christersson, L. (1993). Self-efficacy and reasoned action: predicting oral health status and behavior at one, three, and six month intervals. *Psychology and Health, 8,* 105–121.

Tedesco, L., Keffer, M., & Fleck-Kandath, C. (1991). Self-efficacy, reasoned action, and oral health behavior reports: a social cognitive approach to compliance. *Journal of Behavioral Medicine, 14,* 341–355.

Thomas, A., Birch, H., Chess, S., Hertzig, M., & Korn, S. (1963). *Behavioral individuality in early childhood.* London: University of London Press.

Thrash, W., Marr, J., & Box, T. (1982). Effects of continuous patient information in the dental environment. *Journal of Dental Research, 61,* 1063–1065.

Todd, J., Walker, A., & Dodd, P. (1982). *Adult dental health* (Vol. 2). London: HMSO.

Vassend, O. (1993). Anxiety, pain and discomfort associated with dental treatment. *Behavior Research and Therapy, 31,* 659–666.

Veerkamp, J. S. J., Gruythuysen, R., Hoogstraten, J., & Amerongen, W. van (1993). Dental treatment of fearful children using nitrous oxide. Part 4: Anxiety after two years. *Journal of Dentistry for Children, 60,* 372–376.

Venham, L. (1979). The effect of mother's presence on child's response to dental treatment. *Journal of Dentistry* *for Children, 46,* 219–225.

Wallston, K., Wallston, B., & De Vellis, R. (1978). Development of multidimensional health locus of control (MHLC) scales. *Health Education Monographs, 6,* 160–170.

Wardle, J. (1983). Psychological management of anxiety and pain during dental treatment. *Journal of Psychosomatic Research, 27,* 399–402.

Weiner, A., & Sheehan, D. (1990). Etiology of dental anxiety: Psychological trauma or CNS chemical imbalance? *General Dentistry, 38,* 39–43.

Weinstein, P., & Getz, T. (1980). *Changing human behavior: strategies for preventive dentistry*. London: Mosby.

Weinstein, P., Getz, T., Ratener, P. et al. (1982). The effect of dentists' behaviors on fear-related behaviors in children. *Journal of the American Dental Association, 104,* 32–38.

Weinstein, P., Kegeles, S., & Lund, A. (1980). Children's health beliefs and acceptance of a dental prevention programme. *Journal of Health and Social Behavior, 21,* 59–74.

Williams, J. M. G., Murray, J. J., Lund, C., Harkiss, B., & Franco, A. de (1985). Anxiety in the child dental clinic. *Journal of Child Psychiatry and Psychology, 26,* 305–310.

Zasler, N., Devany, C., Jarman, A., Friedman, R., & Dinius, A. (1993). Oral hygiene following traumatic brain injury: A programme to promote dental health. *Brain Injury, 7,* 339–345.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.27
Child Health Psychology

CHRISTINE EISER and NICKY MAIN
University of Exeter, UK

8.27.1 INTRODUCTION 618
- *8.27.1.1 Terminology* 618
- *8.27.1.2 Do We Need a Special Child Health Psychology?* 618
  - *8.27.1.2.1 Children's illnesses are different from those of adults* 619
  - *8.27.1.2.2 Developmental changes in emotional and cognitive reactions to health and illness* 620

8.27.2 LIFESTYLES AND HEALTH PROMOTION 620
- *8.27.2.1 Substance Abuse* 620
- *8.27.2.2 Sunbathing* 622
- *8.27.2.3 Exercise* 623
- *8.27.2.4 Eating Behavior* 623
- *8.27.2.5 Accidental Injury* 624
- *8.27.2.6 Summary* 625

8.27.3 ILLNESS EXPERIENCE 625
- *8.27.3.1 Illness Experience in Healthy Children* 625
- *8.27.3.2 Hospitalization* 626

8.27.4 CHRONIC ILLNESS 626
- *8.27.4.1 Incidence and Definitions* 626
- *8.27.4.2 Theoretical Issues* 627
  - *8.27.4.2.1 Deficit-centered models* 627
  - *8.27.4.2.2 Conceptual models* 627
  - *8.27.4.2.3 Attachment theory* 628
  - *8.27.4.2.4 Social learning theory* 628
  - *8.27.4.2.5 Attribution theory* 629
  - *8.27.4.2.6 Difficulties with theories* 629
- *8.27.4.3 Methodological Issues* 629
  - *8.27.4.3.1 Sampling* 629
  - *8.27.4.3.2 Measures* 629

8.27.5 EMERGING ISSUES 631
- *8.27.5.1 Neuropsychology* 631
  - *8.27.5.1.1 Neuropsychological assessment* 632
- *8.27.5.2 Adherence* 632
- *8.27.5.3 Quality of Life* 633
  - *8.27.5.3.1 Why measure quality of life?* 634
  - *8.27.5.3.2 Domains of quality of life* 634
- *8.27.5.4 Family and Ethnic Issues* 635

8.27.6 COMMUNICATION 636
- *8.27.6.1 Understanding Health and Illness* 636
- *8.27.6.2 Genetic Testing* 637

8.27.7 PROFESSIONAL ISSUES 638

8.27.8 CONCLUSIONS 639

8.27.9 REFERENCES 639

8.27.1 INTRODUCTION

8.27.1.1 Terminology

Interest in children and health is shared by workers from a number of different backgrounds, including pediatricians, psychologists, and psychiatrists. Depending on the specific emphasis and interests of those involved, the area is variously described as "pediatric psychology," "child health psychology," or "clinical child psychology." Pediatric psychology is flourishing more in the USA than elsewhere (Wallander, 1996). This has been partly attributed to the fact that pediatricians in the UK tend to see themselves as interested in, and responsible for, the psychological well-being of children in their care. They are therefore less likely to refer children to other specialists. In contrast, pediatricians in the USA recognize the differing roles and expertise that can be brought by other professionals. Regardless of the specific terminology, the focus is on a commitment to the health and well-being of children and their families. Pediatric psychology is an "amalgamation of applied practice and basic research concerned with medical and psychological problems affecting children and families" (Roberts, McNeal, Randall, & Roberts, 1996, p. 107). The goal is to optimize the physical and psychological health of young people, recognizing the interdependence between health and developmental status. For this reason, it is important that child health psychology be firmly rooted in developmental psychology.

8.27.1.2 Do We Need a Special Child Health Psychology?

Traditionally, health psychology has been concerned with issues of greater relevance to adults than children. General textbooks devote much more attention to questions such as coping in adulthood with chronic or life-threatening illness (such as cancer or heart disease), the promotion of a healthy lifestyle, or to issues of health promotion in the workplace. Children, if mentioned, are often assigned to a few paragraphs. Work concerned with the impact of a diagnosis of cancer, for example, tends to center on issues of adult cancer, especially how disease impacts on work or social functioning. As far as children are concerned, information will often be based on parents' or doctors' reports about the impact. It is rare that much coverage is given to the issue of how children themselves consider that the illness affects their lives.

In many textbooks child issues are restricted to brief discussions of lifestyles (drug abuse, smoking, weight control), but even here the emphasis tends to remain on the implications for adult health rather than the child's current well-being. The impact of illness, the effects of hospitalization, pain, and adherence to treatment should perhaps all be considered separately for children and adults. Although whole chapters may be devoted to specific conditions which affect adults, it is much less likely that comparable coverage is given to illnesses affecting children (e.g., childhood cancer or diabetes). Other diseases, especially those partly or wholly prevalent during childhood (muscular dystrophy or cystic fibrosis), are often neglected. We may be urged to consider the age-appropriateness of information for children, but details of how this is achieved are rarely specified.

There are several reasons why health psychology has focused so much less on children compared with adults. First, there is a strong assumption that health is better during childhood than at any other time of life: we do not have to worry about it. Certainly we would like to hope that children have far fewer physical health problems compared with the middle aged or the elderly. Such an assumption is less well supported empirically than might be imagined. Recent work suggests a higher incidence of both physical and psychological complaints in children than might be expected (Dubow, Lovko, & Kausch, 1990). Almost all work suggests that these complaints are higher in girls compared with boys. From an adult perspective, some of these concerns might be considered trivial or not so important (e.g., health concerns include many anxieties about skin conditions which under many circumstances correct themselves with time). However, we know that a proportion of children suffer from more serious and chronic conditions (Cadman, Boyle, Szatmari, & Offord, 1987).

Largely because it has been assumed that children are healthy, interest in child health psychology has often been justified in terms of the implications for adult health status and behavior, rather than the implications for the child at the time. Much of the work concerned with legal or illegal drug use has therefore focused on prediction of adult drug use from child histories (Kandel & Yamaguchi, 1984). Similarly, poor eating habits are seen to be of interest as much in terms of their predictability for adult eating status and obesity as for the implications for child health (Simopoulos, 1987). The tendency to consider the implications of child health for adult health is also prevalent in work with children with chronic diseases.

Children with diabetes, for example, are encouraged to adhere to medical advice as much to reduce the likelihood of complications in the future as to result in any significant health benefit in the present. In all these examples, it is apparent that the interest in child health psychology for its own sake is seen to be less important than the implications for adult health.

Second, concerns about the reliability of data elicited from children can be a deterrent for many researchers. In the past, it was assumed that children were unable to make reliable judgments especially about the localization and extent of pain. Current work suggests that by three years of age children are able to make these judgments (McGrath & McAlpine, 1993). Children's relatively brief timespan and readiness to forget about injury if cuddled or bribed with goodies is also taken as evidence for the unreliability of children's responses. As our methodologies become more sophisticated this view is being challenged (Eder, 1990). This situation parallels the history of progress in developmental psychology. This was initially dominated by work which focused on what children did not know or understand (Piaget, 1929), and which highlighted difficulties experienced in abstract reasoning. Increasingly, it has been shown that these conclusions were partly attributable to methodological difficulties, in that when the experimental situation is less complex, children show greater sophistication (Nelson, 1986). A problem in this area has been the over-reliance on the semistructured clinical interview as a basis for data collection, with the inherent difficulties of coding. More experimental methods may overcome this problem to some extent (Hergenrather & Robinowitz, 1991). However, for the most part, a sound methodology has not been developed which is also child-friendly. Approaches to data collection are invariably extrapolated from adult methods at least to the point when it is clear that children are too young to be able to understand instructions.

Also, investigators point to cognitive skills, and emphasize the differences between child and adult abilities to process information. Differences between children and adults in their understanding of the causes of, and treatment for, different diseases have been widely researched. Undoubtedly, there are developmental changes in the ways in which children of different ages understand these issues (Bibace & Walsh, 1980). The notion that children proceed through a fairly invariant sequence of beliefs with rather naive implications for communication and explanations of illness (Perrin & Gerrity, 1981) has many-attractions. Even this view, however, is increasingly challenged (Carey, 1985; Hergenrather & Robinowitz, 1991).

Our argument then is that child health psychology is important in its own right and not merely in relation to the consequences for adult health. At the same time (i) children's illnesses are different from those of adults, and (ii) developmental changes in emotional and cognitive reactions to health and illness necessitate the development of a child health psychology which is unique and distinct from adult health psychology. This is not to deny that there are some issues which overlap with those in adult health psychology, at least at a superficial level. Even so, a different methodological and theoretical approach is often needed. For the most part, differences between the types of illness which affect children and adults, differences in cognitive and emotional response to illness, and the greater involvement of the family all necessitate a distinctive approach.

8.27.1.2.1 Children's illnesses are different from those of adults

Some diseases are unique to childhood. These include congenital conditions, or conditions which can be fatal during childhood, such as cystic fibrosis. Others, although going by the same name, have a different etiology when they occur in children compared with adults. For example, the kinds of cancer that affect children are different from those that affect adults, and the course of treatment differs (Boring, Squires, Tong, & Montgomery, 1994). Similarly, in adults with diabetes, residual insulin production means that patients can often be treated by oral medication and attention to diet. In these cases, the intrusiveness of the disease on daily life is minimal. In children with diabetes, the complete lack of insulin production means that treatment involves insulin injections. The degree of intrusiveness and the perception of severity is inevitably greater, and risk of long term complications higher than for adults with noninsulin dependent diabetes.

In addition to the medical differences between child and adult conditions, the social implications are different. In the case of young children, parents are responsible for care and disease management. Especially as children grow up, conflicts between child and parent can become common. Parents may feel guilty about the possible inherited nature of the disease, and children often resent the fact that parents are responsible for many routine aspects of daily care. The diagnosis of a chronic condition in a child is a diagnosis with implications for the whole family, to a larger extent than when an adult is the patient.

Pharmaceutically, it is inappropriate to simply reduce the quantity of drugs prescribed for a child compared with an adult. There are differences in metabolism and vulnerability of developing organs which can necessitate a qualitatively different kind of treatment. For the same reason simply scaled down adult psychology should not be used as a basis for understanding children's reactions to illness.

8.27.1.2.2 Developmental changes in emotional and cognitive reactions to health and illness

Children differ from adults in their understanding of the cause, treatment, and implications of illness, their recall of illness events and interpretations, and their perceptions of risk involved in different situations. They may be concerned about events which do not upset adults and vice versa.

Although we recognize that treatments can be painful, we have tended to make judgments about the extent of pain, or distress based on reports of adults, especially mothers. Mothers experience stress themselves (Hobfoll, 1991) and, as a consequence their accounts of the impact of the disease, may be influenced by their own emotional state. Increasingly it is recognized that it is important to seek the views of children directly rather than infer them from parents' reports.

8.27.2 LIFESTYLES AND HEALTH PROMOTION

Much of the attention of child health psychologists has focused on the treatment of existing physical and psychological problems so that children are able to regain normal developmental pathways. However, the discipline also maintains a basic orientation towards health promotion and the prevention of later onset problems in young children and adolescents (Roberts, 1986) (see Chapters 8.28 and 8.29, this volume). Some of the major challenges facing pediatric health psychology include the growing problems of adolescent drug and alcohol abuse, the issue of smoking in the under 18s, and expanding concern over nutritional issues such as the need for a healthy diet and the problem of eating disorders such as anorexia and bulimia. Exercise is also an ongoing health issue, as is prevention of accidental injuries.

8.27.2.1 Substance Abuse

One of the main reasons why the issues involved in the uptake of substances such as alcohol and drugs need to be examined in more depth is that it has always been assumed that patterns of pathological drinking, and perhaps also drug taking, are laid down in adolescence and early adulthood; the consequences only becoming apparent in later life. Many problem behaviors in youth are in fact interrelated (Jessor & Jessor, 1977). Analyses of college and high school longitudinal studies over a period from 1969 to 1973 found pervasive evidence for covariation among behaviors such as drinking, drug use, sexual behavior, and deviance levels. It was found that those who had engaged in a particular problem behavior had also engaged in various related behaviors at higher rates than those who had not engaged in that particular behavior. Jessor and Jessor (1977) regard their research as supportive of a "syndrome-like patterning" in which participation in a particular behavior is implicative of engaging in others. Clustering of problem behaviors in the same individuals, including alcohol and drug abuse, smoking, and sexual precocity, has now been widely reported (e.g., Aaro, Laberg, & Wold, 1995; Vingilis & Adlaf, 1990).

Regrettably, many young people choose to disbelieve or ignore the long term consequences of drinking or taking drugs. There is, therefore, an urgent need to identify the factors which are predictive of later substance abuse in order that interventions can be applied before dependence or addiction develops. Reducing the number of young people drinking or taking drugs to excess will also bring about benefits such as lowering the frequency of alcohol and drug related crime, accidents, aggressive behavior, and suicide. Many of these events result from altered perceptions and psychomotor reactions that the user experiences whilst under the influence of alcohol or drugs, and consequently could be avoided by lowering usage of these substances in young people.

Youth, defined by the World Health Organization as the period between 15 and 24 years, can be characterized by a period of experimentation for a variety of legal and illegal substances. Drug taking and excessive alcohol consumption appear to be related to a complex collection of factors. Longitudinal studies of childhood characteristics and later alcoholism or drug use (Block, Block, & Keyes, 1988; Zucher & Gomberg, 1986) suggest a combination of intrapersonal and interpersonal risk factors for the future development of substance abuse. Briefly, children at increased risk have been described as particularly temperamental or nonaffectionate at a young age, high in emotionality, hyperactivity and/or conduct disorder, and somewhat more controversially, sociability. Further research is required to determine what differentiates young people who outgrow such behaviors and those who go on to participate in more serious levels of antisocial behavior or drug abuse.

With respect to interpersonal variables, a number of central influences seem to be implicated in the uptake of alcohol and drugs among young people. Drug and alcohol use among teenagers and adolescents usually occurs in group situations, and the use of such substances has been found to be strongly motivated by self-presentational desires such as the need for social approval and peer acceptance (Leary, Tchividjian, & Kraxberger, 1994). For example, many adolescent boys believe that their peers admire the attributes that characterize someone who drinks alcohol and thus aspire to that image (Chassin, Tetzloff, & Hershey, 1985). Family factors, including parental modeling (often parents themselves are heavy users of the substance), and the family–child relationship, are key risk factors for subsequent substance use (Glynn & Haenlein, 1988). In addition, peer relationships appear to be strongly related to the extent of alcohol and illicit drug use during youth. Finally, both types of substance abuse seem to be more common among those who report academic difficulty or poor adjustment at school (Swain, 1991). Other reasons for the illicit use of alcohol and drugs by under age children include the wish to convey autonomy or rebelliousness to parents or other adults, and the reduction of anxiety in intergroup settings (Berglas, 1986). Some may use alcohol or drugs if they feel unable to cope with the pressures associated with growing up; as a sort of temporary escape from school problems, social anxieties, or home difficulties.

Tobacco use is one of the most damaging forms of behavior in the long term to the health of children and adolescents. Young people, particularly those in developed countries, are beginning to use tobacco at an early age (most start before the age of 19 years), and as a consequence find it more difficult to give up than those who start smoking later in life (Eiser, Morgan, & Gammage, 1987). Like alcohol and drug use, those whose parents or siblings smoke or whose friends have taken up the habit are particularly likely to begin smoking themselves (Smith, 1991). Early attempts to prevent the onset of drug and alcohol use failed for a variety of reasons, but often because they were based on faulty assumptions or limited theoretical bases (Botvin, & Wills, 1985; Falco, 1992). These programs relied on scare tactics, moral exhortations or the simple provision of factual information. There is little evidence regarding the efficacy of these approaches.

The most successful smoking and drug control programs have focused on encouraging the development of skills needed to resist peer pressure and media incitement to use tobacco. Programs based on social influence models (Hansen, 1992) suggest that positive effects occur in 63% of these studies, neutral results in a few (26%) and negative results in the smallest proportion (11%). In part, inconsistencies in outcome are likely to be attributable to the variety of strategies that are included under the rubric of social influence approaches.

The basic assumption underlying "social inoculation" theory (McGuire, 1964) is that young adolescents experience increased mobility, seek independence from adults, and experiment with different lifestyles. The decision to drink or use drugs is assumed to be dependent on the ability to resist the social pressures common in early adolescence. The basis of social innoculation theory is that by practising to avoid social pressure in a controlled environment, the adolescent will be more able to deal with the real situation if it arises. Adolescents are therefore trained to recognize social pressure from peers and the media. In addition they are trained to develop the social skills necessary to refuse offers of alcohol and drugs while at the same time coping with the negative social consequences that can result. The assumption is made that adolescents will be able to resist offers of drugs through learning refusal skills and maintaining self-esteem in the face of pressure to do so. Although there is good evidence that children learn resistance skills there is less evidence that participation in these programs is associated with reduced drug use.

Of the many criticisms that have been leveled at this work, the one that seems most likely to account for some of the discrepancies observed is the fact that school-based resistance models are directed uniformly at all children in the school involved. It is likely that the impact of any program will be dependent on individual differences in beliefs about the acceptability of drug use. While some young people may believe that drug use is unacceptable and benefit from learning refusal skills, an unpredicted effect of this training may be that adolescents come to believe that the likelihood of them being offered drugs, and the incidence of drug taking among their peers is actually higher than in reality. By increasing perceptions of the prevalence of drug taking among peers, this kind of teaching may inadvertantly increase the likelihood of experimentation (Donaldson, Graham, & Hansen, 1994).

In an attempt to unravel the processes operating in school based social resistance programs, Donaldson, Graham, Piccinin, and Hasen (1995) compared subsequent alcohol use as a consequence of four training programs. The information only group received four lessons about the health and social consequences of drug and alcohol use. The resistance training group was designed to teach adolescents how to refuse offers of drugs or alcohol in a socially acceptable manner. The normative education group received information to correct erroneous assumptions about the prevalence of alcohol use amongst peers. In the combined condition, students received elements from all three approaches. The programs were evaluated in terms of alcohol use, refusal skills, perceived acceptability of alcohol use, and perceived prevalence of offers.

For adolescents who believe it is not acceptable to drink, the authors concluded that social resistance training does delay the onset of alcohol use. For those who are more inclined to believe that it is acceptable to drink, resistance skills training was not successful in delaying alcohol use. A negative finding was that resistance skills training did result in overestimation of the frequency of alcohol use by others. This was, however, counterbalanced in the combined condition; when adolescents received resistance skills training in conjunction with normative education they reported lower prevalence rates compared with those in either the resistance skills training or normative information conditions.

This study underscores the importance of clarifying the processes through which intervention programs have effects (see Chapter 8.01, this volume). Resistance skills models may be useful but primarily for those who share the basic assumptions implicit within the model, that is, where adolescents accept that alcohol or drug use is unacceptable. It is also clear that evaluations of any intervention program need to include multiple measures of outcome. Simply relying on self-reported drug use is of very limited value, as this yields no information regarding the processes underlying decisions made prior to use. School-based strategies may need to become more sophisticated, and more attempts made to match the program with more specific information about adolescent beliefs about drug use and its prevalence.

8.27.2.2 Sunbathing

Environmental damage caused by the break-up of the ozone layer is generally acknowledged to be the critical factor in increased incidence of skin cancer, particularly among young people. In addition, increased leisure opportunities means that young people experience much greater exposure to the sun than previous generations. Since the cause of skin cancer is so clearly related to sun exposure, efforts to inform the public and reduce risky behavior is necessary and should have a direct impact on disease incidence (Cancer Research Campaign, 1995).

Reducing one's personal risk of developing skin cancer is clearly dependent on the adoption of appropriate behaviors, and on understanding how behavior aggravates risk. Much of what we know about people's current knowledge in this area and their preparedness to make changes in their behavior is based on research conducted in parts of the world where sun exposure is recognized as a major problem, notably Australia and the USA. Increasingly there are concerns about those who live in more temperate climates, where the challenge of education may be more demanding.

A number of barriers to appropriate behavior have been identified. Not least is the fact that within Western cultures, a dark tan is considered highly fashionable, and is seen as a measure of beauty and attractiveness (Weston, 1996). Surprizing though it may seem, being safe in the sun is not considered to be easy, for example, many parents complain that children dislike wearing protective clothing. As in almost all areas of health related behavior, females are more knowledgeable than males (Cody & Lee, 1990; Eiser, Eiser, & Pauwels, 1993; Keesling & Friedman, 1987). They are also more likely than males to use sunscreen regularly (Hill, Rassaby, & Gardner, 1984; Keesling & Friedman, 1987), and to perceive a lighter tan to be more attractive than a darker one (Broadstock, Borland, & Gason, 1992). These results suggest that young males are at higher risk of sun exposure, and may benefit from strategies targeting these behaviors. Young males who work outside, or play lots of sport, appear particularly vulnerable.

There is some evidence that interventions can be successful. A major health promotion in Victoria, Australia, was held over two summers. This included a focus on sun protection as a sensible and fashionable behavior and was aimed at the general population, although the primary target was adolescence and young adults. A key aim was to make tanned skin less fashionable.

Children received education on three separate occasions during schooling and knowledge improved. Following this promotion, virtually all Australian children were aware of the need for sun protection (Broadstock et al., 1992). They were also positive about the idea of taking appropriate action, except with regard to keeping out of the sun. Those at high risk of burning (defined in terms of skin and hair color) understood their own risk, had more positive attitudes to sunscreen, and were satisfied with achieving a lighter tan. However, those at slightly greater risk (somewhat more sensitive skin) tended to report similar beliefs to those at little risk. This cannot be attributed to ignorance as those with slightly sensitive skin were in fact more knowledgeable.

With regard to taking recommended action, girls were less in favor of wearing hats than boys irrespective of their knowledge or attitudes. Boys seem to be less concerned about fashion and accept hats for utility (although, maybe, contemporary fashion dictates that boys can wear baseball caps).

Girls were more likely than boys to remember having a school lesson about skin cancer, had more knowledge, had more positive sun protection attitudes, and were more likely to report using a sun protection factor 15/15 sunscreen. Girls expressed more favorable attitudes to sun protection in all aspects except wearing a hat and suntan preference. There were no sex differences concerning sun avoidance, beliefs about suntans or reported incidence of sunburn.

In this sample of 12–17 year olds, 15 year olds had attitudes least conducive to protection with more appropriate attitudes being held by both younger and older children. This may suggest that 15-year-olds take most risks, with this diminishing by 17 years of age. However, it may be that the focus on 17-year-olds in school introduces a bias (in Australia 98% of 15-year-olds are in school, but only 65% of 17-year-olds). The fact that this study was conducted during an Australian winter also introduces the possibility that people may have overestimated their protection behavior of the previous summer.

Since parents are usually strongly in favor of protection behavior (Bennetts, Borland, & Swerissen, 1991) it follows that young children are relatively well protected. Adolescents seem to be at special risk. For example, many adolescents have poor sun protection behaviors, in that they do not tend to use an adequate factor sunscreen, and do not apply it as frequently as they should when exposed to the sun (Wichstrom, 1994). Given their greater leisure time and more time spent on outdoor sports and activities, adolescents are likely to be an especially vulnerable group.

8.27.2.3 Exercise

Observational data indicate that age is inversely related to physical activity, with studies showing that the amount of time spent in physical activity declines by 50–75% between 6 and 18 years of age (Rowland, 1990; Sallis, Buono, Roby, Micale, & Nelson, 1993). There is also concern that the amount of physical exercise in schools is declining as a direct result of changes in the educational system since the mid-1980s (Woodruffe, Glickman, Barker, & Power, 1993). At all ages males are slightly more active than females, although according to the British Sports Council (1990) there is some evidence that this gender gap is closing. With regard to environmental variables, (and reflecting findings related to vulnerability for substance abuse) major influences on activity levels include both parental and peer activity. Other environmental factors include access to play areas/sports facilities, or to organizations and programs which aim to increase activity levels in young people. Because there is no single variable that may be considered the primary determinant of children's physical activity, no simple intervention is likely to be highly effective in promoting physical activity among children (Cheung & Richmond, 1995).

Although the end stages of chronic degenerative diseases become manifest primarily during adulthood, the processes underlying such diseases often begin during childhood and adolescence (Despres, Bouchard, & Malina, 1990). Health related fitness is thus a reasonable goal for children, particularly for preventing degenerative disease. This is because regular physical activity during childhood and youth may prevent or impede the development of several adult conditions, such as obesity and degenerative heart disease. It is often assumed that habits of engaging in regular physical activity developed during childhood and adolescence may persist into adulthood. Additional research is needed to resolve when, if at all, during childhood and adolescence regular physical activity is likely to exert a beneficial effect on subsequent adult health status, although the maximal effects of physical activity may be achieved if the habit of regular activity is begun in childhood and maintained throughout life (Bar-Or & Malina, 1995).

8.27.2.4 Eating Behavior

Eating problems can arise from a number of different sources. In developing countries inadequate intake of nutritious food often results in malnutrition, and in industrialized countries the diet of young people may be inadequately balanced or may be insufficient as a result of conditions such as anorexia or bulimia. Dietary habits in youth and adolescents can have important implications for both the short and longer term health of the individual. For example, iron deficiency can lead to fatigue and anemia in the short term, and calcium deficiency is one of the main causes of osteoporosis in later life, particularly for women. However, most children in the UK are adequately nourished, and although over the last decade there has been a trend towards healthier eating habits with diets including less saturated fat and sugar, fatty and sweet foods still make up a large part of the diet of young children (Golding, Haslum, & Morris, 1984; Nutbeam, Aar, & Catford, 1989). This is particularly true for those from the lower social classes (Woodruffe et al., 1993). Overnutrition leading to obesity or excessive weight gain is an increasing problem in many developed countries, with young people preferring widely available snack foods usually high in fat or sugar.

Excessive food intake at a young age will predispose people to a variety of late onset health problems such as heart disease, diabetes, arthritis, and some cancers (Gortmaker, Walker, Weitzman, & Sobol, 1990). Children become obese when their energy, or calorific, intake exceeds their energy expenditure. The resting metabolic rate comprises about 70% of total energy expenditure. Since the amount of weight being supported influences the resting metabolic rate, obese children have significantly higher rates than nonobese children. This suggests that they should also consume more calories than their nonobese counterparts, although due to the unreliability of self-report measures it is difficult to establish whether or not this is in fact the case. Another area requiring more research concerns the link between obesity and activity: there is some evidence that obese children have lower activity levels than nonobese children, although again more objective methods for measuring activity are needed. There are a number of individual factors to consider when investigating the etiology of obesity, in that some children may become obese by being very sedentary with normal eating, and others may eat large amounts of food and yet not become obese because their energy expenditure is so high (Epstein, 1992). A genetic predisposition to obesity has been well recognized, resemblances in fatness being greatest in monozygotic twins. Although a genetic component to obesity is supported by the familial associations of the disease, it is still not clear how this genetic predisposition is expressed. The link between parental obesity and obesity in offspring may be attributable to a shared environment as well as a shared genetic inheritance.

In general, concern has focused more on managing and treating excessive fatness in children than on preventing it. Because physical activity plays such an important role in regulating fatness, educators and the medical community should emphasize the value of regular exercise during childhood to prevent obesity. Although the most effective treatment for obesity involves simply decreasing calorific intake, weight loss may be accelerated by increasing exercise. Further research is needed into the long-term effects of exercise, the problem of encouraging children to adhere to exercise regimes, and on the role of sedentary activities in the development of obesity (Epstein, 1992).

The problem of obesity in young people is of further concern because of the associated impact on appearance and self-esteem. Concern over excessive weight gain has led to an increase in weight related disorders such as anorexia and bulimia among both school age children and adolescents in recent decades. Both bulimia and anorexia nervosa can be triggered by periods of dieting, although the weight loss associated with anorexia and the overeating reminiscent of bulimia can also be brought about by organic disease.

8.27.2.5 Accidental Injury

Deaths from accidents outnumber deaths from childhood cancer, heart disease, and genetic diseases combined (Baker, O'Neill, & Karpf, 1984). Recent laws enforcing the wearing of seat belts and safety helmets for motorcycles have helped to reduce the frequency of traffic injuries, although over half of accidental deaths in children aged 1–15 remain due to road traffic accidents (Wynne, 1991). Other accidents include leisure and sports pursuits, burns, poisonings, falls, and drownings. Homicide is also a major cause of accidental death in the age group 15–24 years. Behavioral factors that contribute to accidental injury include alcohol or drug intoxication, lack of experience, and risk-taking (Graham, 1996). Nonfatal accidents can result in a great deal of suffering both for the child and their family. In addition the incidence of serious injuries is higher in the lower social classes, thus adding disability to existing social disadvantages (Alwash & McCarthy, 1988). For this reason, and because accidental injury in children is so prevalent, those working in the area of child health have an important role to play in the prevention of accidents in the young (Jackson, 1988).

Injuries at home account for more than half of all fatalities in children under five years of age. For young children there is some evidence to suggest that parent behavior such as poor parenting skills and inadequate supervision may contribute to the occurrence of a child's injury

(Oldershaw, Walters, & Hall, 1986). Since many home hazards cannot simply be reduced or eliminated, it appears that active strategies requiring that safe behavior be taught and performed regularly and reliably are the most effective methods for injury prevention (Finney & Cataldo, 1991), and should be combined with passive strategies, such as blocking access to hazardous items wherever possible.

The more knowledgeable that parents are about a safety issue, the more competent they feel to intervene. In addition, the more they believe an intervention would prevent injury, the greater the reported teaching efforts. Thus, it is important that parents feel knowledgeable about safe behavior, and competent to teach their children these skills (Peterson, Oliver, Brazeal, & Bull, 1995).

8.27.2.6 Summary

A combination of instability in the family and excessive negative pressures from the surrounding environment (adults, peers, etc.) are more likely to lead to health damaging behaviors such as the use or abuse of tobacco, alcohol, drugs, poor eating, and excessive risk taking in adolescents (Buckstein, 1995). Such problems are of particular importance to the healthcare system because it is adolescents and young adults who are the least likely to come forward and admit that they need help, or come too late, at a stage when the problems have become more difficult to resolve. The focus on school-based education programs has resulted in an overemphasis on the roles of peers and a failure to acknowledge the contributions made by family relationships to onset of drug use. A more balanced view, reflecting the combined influence of peer and family relationships is required.

However extensive an education program, a proportion of young people will inevitably adopt unacceptable behavior. In these cases, failure to organize a referral service which is acceptable to young people, often from minority or deprived backgrounds, further hampers efforts to revert such socially unacceptable behaviors.

8.27.3 ILLNESS EXPERIENCE

8.27.3.1 Illness Experience in Healthy Children

All children come across illness during the course of growing up. This may be through the experience of acute illness themselves, thought to be a significant influence on child attitudes (Parmelee, 1986). Many children learn about illness following acute illness and hospitalization, or minor injury; others experience illness vicariously, through their friends or parents; a further group through the experiences of brothers or sisters.

There has been less work which focuses on the reactions of healthy peers to chronic disease or disability compared with work which focuses on healthy siblings' reactions. Early work which attempted to account for the experiences of healthy siblings paralleled work with sick children themselves; that is, from a "deficit-centered" perspective, it was assumed that healthy siblings would show compromised behavioral, social, and cognitive function in relation to children from healthy families (Drotar, 1981). There was some support for this view, although most studies were subject to the same criticism as work with sick children themselves. They were based on small samples, failed to use age-appropriate measures, and, most damning of all, did not give the children the opportunity to report anything other than negative emotions or behaviors. In one of the few studies to overcome this criticism, preschool siblings of a child with cancer were compared with children in families where there was no illness. Siblings of a child with cancer showed more empathy and altruistic behavior (Horwitz & Kazak, 1990). Such a study needs replication, both for the preschool age group and also with regard to older children.

Much of the early work about childrens' conceptions of health and illness centered around the Piagetian "stage model" of development, with the assumption that it was possible to learn as much from studying what children did not know as what they did. Systematic changes in how children perceive health and illness as they progress through childhood have been described (Bibace & Walsh, 1980; Perrin & Gerrity, 1981). Briefly, in the preoperational stage there is confusion between cause and effect, and no differentiation between illnesses. This is followed in the concrete operational stage by an understanding of the role of contagion and germs in the etiology of illness, and during the stage of formal operational thought a more sophisticated understanding of infection and health prevention. Some studies appear to support this analysis (Beales, Holt, Keem, & Mellor, 1983) in that younger children offer fairly simple explanations of illness whereas older children tend to offer more organized and detailed descriptions. However, much of this data is based on semistructured interviews with unnatural and repetitive language, and it is not clear that the same responses are always coded in the same way by different researchers (Eiser & Kopel, 1996). Consequently some of the more recent research has tended to focus more on what children *do know* as opposed to what they *do not know* (Nelson, 1986). Empirical work along this line suggests that both healthy and chronically ill children can in fact develop accurate and quite detailed accounts of illness events (Bearison & Pacifici, 1989), and have much greater insight about health and illness than was originally proposed by Piaget.

The study of the development of concepts of health and illness has been justified on a number of different practical and clinical grounds, one of which is that of providing appropriate health education. The underlying message here is that in order to be effective and promote changes in behavior, educational materials need to be tailored to the developmental level of the child. Developmental issues have been most clearly taken into account in recent programs aimed to increase knowledge and understanding about AIDS (Eiser & Kopel, 1996).

Most five-year-olds believed that simple proximity to an AIDS victim was sufficient to contract the disease, but older children understood that it could be transmitted through sexual contact or mixing blood (Osborne, Kistner, & Helgemo, 1993). Thus, it appears that young children would benefit from curricula that aimed to reduce fears about how the disease is transmitted, for example, not to touch cuts, and for the older children information about how to avoid contaminated blood and bodily fluid seems more appropriate.

8.27.3.2 Hospitalization

For the majority of children and young people, good health is taken for granted and it is clear that for the most part, today's young people are healthier than those in previous generations. Even so, illness is a fact of life. Although there have been many improvements in the care of children in hospital, it remains a potentially frightening experience.

There have been significant changes in the way children are treated and the whole ethos of child care in hospital settings. Much can be credited to the work of Robertson (1958). In contrast to practice during the 1950s (when parents were generally allowed brief formal visits to their child once or twice a week), it is now accepted that parents should be closely involved in the care of their child in hospital. Efforts are also made to ensure some continuity between home and hospital. Children are encouraged to bring toys from home; siblings and friends can visit; efforts are made to prepare children for the procedures to be experienced; special units have been established for adolescents. There is much evidence that these changes have resulted in a far less traumatic experience for children in hospital today (Cleary et al., 1986), but improvements still need to be made (Thornes, 1983).

Many changes can be identified but tend to be introduced without evaluation. As a consequence, it is difficult to be certain which changes contribute most to improved function. Evaluations are also necessary to determine the overall impact of an intervention. Although children may benefit when parents are able to live on the ward, the toll for parents and other children in the family needs to be considered (Meadow, 1992). Parents play an important role in determining the way in which children react both to hospitalization generally and treatments more specifically. For example, parental behavior, specifically in terms of how, and when, information about procedures is given to the child, is an important predictor of child distress (Jacobsen, Manne, Schorr, Rapkin, & Redd, 1990).

8.27.4 CHRONIC ILLNESS

8.27.4.1 Incidence and Definitions

Some 10–12% of children suffer from a chronic illness (Cadman et al., 1987). The numbers of children affected are relatively large and the implications are usually long term. For many children, the implications of their disease are relatively benign, but about 1–2% of the total child population have more severe conditions (Gortmaker & Sappenfield, 1984). Advances in biotechnology mean that it is now possible to treat many more children than in the past. However, the cost in financial terms to the health service is coming under greater scrutiny (Stevenson, McCabe, Pharoah, & Cooke, 1996) and the cost to families in social and psychological terms is increasingly recognized (Eisser, 1993; Wallender & Thompson, 1995). These include the added caretaking burden involved in care of a sick child (Quittner, 1992), increased psychological morbidity, especially among mothers (Hobfoll, 1991), and impact on other family members, especially healthy siblings (Lobato, Barbour, Hall, & Miller, 1987). For the child, there are likely to be interruptions to regular schooling, restricted mobility and activities, and sometimes long-term restrictions in adult life (Eiser, 1995). To determine the extent of maladjustment shown by children, Lavigne and Faier-Routman (1992) conducted a meta-analysis of the empirical literature. This included the results from 87 published articles. Taking all the results together, they suggested that children with chronic disease were significantly different from normal children. Specifically, there were more than twice as many children with chronic disease with problems in adjustment generally, internalizing and externalizing behavior problems and poor self-esteem compared with healthy children.

In the past, there has been a tendency for psychologists to work within a medical model in child health psychology. Thus, it is assumed that the consequences in terms of psychological health are disease specific. Such a view is perpetuated where psychologists are "attached" to clinical specialties, for example, in oncology or renal disease. Reservations about this approach have been put forward by a number of workers, while the disadvantages of a disease-specific model have been most clearly summarized by Stein, Bauman, Westbrook, Coupey, and Ireys (1993). They raise a number of objections. First, it is not possible to include every possible disorder to which children might be susceptible. Second, diagnoses may be made inconsistently by different physicians. This might especially apply to conditions such as asthma. Third, a diagnostic label is relatively uninformative with regard to the severity of the condition. Fourth, there may be a bias to include only those children who have access to the medical care system, and fifth, there may be a delay between the emergence of symptoms and diagnosis. In many instances, it may therefore be more useful to categorize diseases in terms of their consequences rather than their labels.

According to their definition, chronic diseases share the following characteristics: they have a biological, psychological or cognitive basis; they last, or are virtually certain to last, for more than one year, and they are associated with one or more of the following sequelae:

(i) Limitation of function, activities or social role in comparison with healthy same-age peers in the areas of physical, cognitive, emotional, and social functioning, as well as growth and development.

(ii) Dependency on one or more of the following to compensate or minimize limitation of function, activities or social role: medications, special diet, medical technology, assistive device, or personal assistance.

(iii) A need for medical care or related services, psychological services, or educational services over and above the usual for the child's age, or for special ongoing treatments, interventions or accommodations at home or school (Stein et al., 1993, p. 345).

While welcoming such a comprehensive definition, the question arises as to who should determine limitation of function, be it in physical, social, or emotional fields. Doctors, teachers, parents, siblings, and the children themselves may each see the limitations differently, partly as a function of their experiences (e.g., parents generally do not always know much about what goes on in school) and partly because of differences in personal values or personality. Doctors tend to see the limitations of a disease to be greater than parents, who put greater value on the implications (Cadman, Rosenbaum, Boyle, & Offord, 1991). Similarly, doctors rated the quality of life in children with cancer to be higher than ratings made by parents (Billson & Walker, 1994).

8.27.4.2 Theoretical Issues

8.27.4.2.1 Deficit-centered models

Given the obvious disadvantages associated with a chronic disease in childhood, early work adopted a "deficit-centered" approach (Drotar, 1981). This emphasized the extent to which children and their families experienced problems, with little attention given to coping efforts. Early reports tended to confirm the validity of this model, with work involving children with a variety of chronic diseases pointing to the existence of widespread deficits in psychological, behavioral, and emotional function. More recently, large scale epidemiologial studies have also concluded that children with chronic diseases are at greater risk of maladjustment than healthy children (Cadman et al., 1987). Based on a study involving 3294 Canadian children, Cadman et al. (1987) concluded that chronically sick children were twice as likely as healthy children to show psychiatric disorder. Those with both a chronic disease and physical disabilty were at greater risk (3.4 times as likely as healthy children).

8.27.4.2.2 Conceptual models

In an attempt to account for differences between children in their response to the stress of having a chronic illness, and guide interventions, recent theoretical explanations have attempted to include a wider range of variables than in the original deficit-centered models. Two models dominate much current empirical work (Thompson, Gustafson, Hamlett, & Spock, 1992; Varni & Wallandar, 1988).

According to the disability–stress-coping model (Varni & Wallandar, 1988), the factor primarily responsible for increasing the risk of adjustment problems in chronically sick children is stress. This may be associated with their physical condition (disease–disability parameters) and related behavioral and environmental circumstances (functional limitations associated with the condition). In addition to disease related stress, children inevitably experience everyday stress in common with others.

These may appear more difficult to manage because of the condition. For example, changing from junior to high school is stressful for all children (Simmons & Blyth, 1987) but more so for those who have a physical handicap which makes it difficult to get around a new environment. Such a change of school may also be stressful from a social point of view in that children may be obliged to explain their condition to new classmates and teachers. Adjustment is also thought to be mediated by individual differences in personal, social, and coping processes, and attenuated by various resistance factors such as family relationships. Although it is not possible that any single study can include all the variables identified in the model, a number of empirical studies have been reported which focus on elements of the model. For the most part, it seems that adjustment is not affected by the condition or its severity, but social and family variables are important. Family communication has been implicated in several studies as an important predictor of adjustment.

There are similarities between the model proposed by Varni and Wallander (1988) and that of Thompson and colleagues (1992). Both view chronic disease as a potential stressor. However, Thompson adopts a more ecological-systems perspective. In particular, the model is less generic and emphasizes the importance of different dimensions of illness in determining adjustment. However, both models draw initially on theoretical ideas basic to the coping model of Lazarus and Folkman (1984). Both, therefore, stress the importance of cognitive appraisal of the situation, family processes, and coping strategies.

These models are exemplary in defining a wider range of variables that interact to determine outcome in the child and are a major step forward from the earlier deficit centered models. Inevitably, they also share some problems. They are highly complex, with the result that it is only possible to test parts of the models. No overall assessment of the models has been attempted and none is likely to be made, given the complexity of the methods which would be involved. In addition, it is not always clear how the different variables identified contribute to the overall processes underlying adjustment; in many cases, they are simply identified and assumed to be relevant. Both make assumptions about the stress involved in living with a chronic disease, yet neither specify the nature of this stress with any precision. The very size and ambitiousness of these models is also their biggest disadvantage.

Several models in social psychology also have some potential applicability. Attachment theory, social learning theory, and attribution theory all have some potential, but until now have been used only in isolated studies.

8.27.4.2.3 Attachment theory

A number of studies involving young adults treated as children for chronic or life-threatening conditions suggest that it is in the area of intimate social relationships that most negative sequaelae are observed (Ireys, Werthamer-Larsson, Kolodner, & Gross, (1994). Such a finding points to the potential value of attachment theory, particularly as developed by Hazan and Shaver (1987). According to this model, attachment styles develop from experiences in regulating distress with attachment figures, and these styles guide responses to threatening situations. Secure attachment is thought to result from responsive caretaking and reflects situations in which individuals are allowed to acknowledge distress and request support. Anxious/ambivalent attachment develops from insensitive or inconsistent caretaking and results in hypervigilance to negative affect. It might be expected that prolonged illness in childhood affects the development of secure attachment, by limiting the availability of and accessibility to parenting. Despite the apparent relevance of attachment theory, surprisingly little work has investigated this model with sick children and their families. Feeney and Ryan (1994) present some empirical evidence in support of the model, although this was based on recall of early childhood experiences made by young adults. Their data support the predictions for anxious/ambivalent attachment but not for secure attachment. Secure attachment was more prevalent among females, although it was not associated with aspects of parenting as predicted.

8.27.4.2.4 Social learning theory

Both subjective and objective health have been linked with parental modeling of the sick role. Parental modeling refers both to the prevalence of chronic illness in a parent as well as parental attitudes which encourage sick role behavior (such as being given special privileges if unwell or allowed time away from school). This model has been used to understand the occurrence of unexplained pediatric pain and guide intervention work. Walker and Greene (1989) found that children with unexplained pain tend to come from homes where there is a high incidence of illness among family members and where parents are highly solicitous to the child's physical complaints. This suggests that interventions may need to target parental behavior. Family based interventions, based broadly on social learning models, appear successful both in the management of unexplained pain (Sanders et al., 1989) and sickle cell disease (Walco & Dampier, 1990).

8.27.4.2.5 Attribution theory

A study by Wright (1990) found that adolescents with renal disease who were poorly adjusted to their disease were more likely to be blamed by their parents compared with those in good control. For example, one father described his son, who was poorly adjusted, as follows: "It can be difficult getting him to stick to his diet because he can be awkward about what he eats."

In contrast, parents of well-adjusted adolescents suggested that any difficulties experienced were more a consequence of the environment or were associated with the disease: "Obviously she's missed out at school, but she's done incredibly well—she's the type of person who's well-motivated and will get on with something, and that's pulled her through."

8.27.4.2.6 Difficulties with theories

A major barrier in health psychology in trying to implement theories developed in social and developmental psychology is in the measurement of key concepts. This was clearly a difficulty in the work involving attachment theory. In adults, attachment tends to be assessed by lengthy interview or forced choice questionnaires. These methods are not always suitable when working with sick children. For example, in a study by Feeney and Ryan (1994), different results were obtained depending on whether subjects were given a series of descriptions of attachment styles and asked to select the one that best described them, or rated themselves in terms of applicability across the three styles. The selection of measuring instruments can, therefore, have implications for quite major differences in reported results. Subtle differences in method may be even more influencial when working with children.

8.27.4.3 Methodological Issues

8.27.4.3.1 Sampling

The appropriateness of the deficit-centered model has been challenged on a number of grounds (Lavigne & Faier-Routman, 1992), and initially explanations focused on methodological differences in the way in which studies were conducted. Certainly, much work in this area has not been above criticism from the point of view of methodological rigor. First, samples are often recruited from a single hospital or medical centre and may therefore lack comparability with samples elsewhere, in terms of characteristics of the population or in terms of the prevalent approach to care and management. The literature is dominated by work from prestigious medical centres. Inevitably such an approach also means that small sample sizes are the norm. In addition, it means that a hospital based sample includes children across a wide age-range, thus failing to take into account developmental changes in cognitive or social development which might modify the impact of illness.

Second, it is common practice to compare responses of the sick group with population norms if these are available, or more commonly with an age and gender matched control group of healthy children. Apart from age and gender, many other factors are likely to be important, but it is rarely possible to match with much sensitivity.

Third, there has been some disagreement about the best approach to data collection. Some have considered that the value of standardized instruments is paramount, even where this means that instruments are employed which were developed for different populations of children and may lack the sensitivity necessary to detect the impact of a chronic disease. Yet, given sample sizes, it is almost impossible to develop well standardized questionnaires with appropriate norms which tap the specific needs of the population of interest. Increasingly, workers are recognizing these difficulties and are turning to more qualitative methods of data collection (Lask, 1996).

8.27.4.3.2 Measures

Despite the number of measures developed for use with adults, it is crucial that new measures are developed for work with children. In almost all conditions, younger children show different and arguably more serious difficulties compared with older children. Many of these follow from the younger child's difficulties in understanding the reason for treatment. In sickle cell disease, younger children show an increased incidence of bacterial infections (e.g., meningitis) and consequently require more frequent follow-up than older children. They also need closer supervision to prevent pain episodes (e.g., reminders about the need to drink plenty of fluids), partly because of the lesser cognitive maturity and partly because of the greater risk of complications during this period (e.g., damage to the spleen). Younger children with cancer may have significant physical problems in the short and long term (poor nutrition, growth impairment, cognitive impairment following radiotherapy) as well as emotional problems resulting from interrupted care-taking and the need for extended hospitalization. Older children with cancer may experience different emotional problems resulting from greater embarrassment associated with changes in physical appearance, interrupted school attendance, and peer and family problems. Although they are more able to understand the reason for treatment, this in itself may be distressing information, raising questions about disability or long term survival. Further examples can be found in almost all pediatric conditions.

The challenge when working with children, however, is to take into account how concerns change with maturity. Central to many definitions of quality of life is the impact of disease on school or work and relationships with friends. Yet "getting on at school" may have more social meanings to the young child. The demands of the national examination system may create more academic concerns for the adolescent. There are qualitative shifts in patterns of friendship so that it is more common for older children to have fewer but more intimate friendships than younger children. Normative changes of this kind need to be taken into account. The speed of development means that concerns about illness can also change. A scale which was appropriate for three-year-old children was less adequate for five-year-olds (Cadman & Goldsmith, 1986). Thus, any measure of impact needs to be sensitive to normative developmental tasks and goals.

(i) Parent completed measures

In the past, there has been an assumption that parents are the most reliable sources of information about a child's well-being. In many situations, children may be too young, or too ill, to be able to answer for themselves. In these cases, medical staff have no choice but to rely on information from parents. Yet, it is surprizing how little evidence exists to suggest that parents are reliable informants about their child. Conclusions about the level of stress experienced by a child appear to be dependent on who is giving the information. Most agreement has been found between nurse ratings and behavioral observations, with lowest levels of agreement between parents and child self-report (Manne, Jacobsen, & Redd, 1992). Explanations about lower than expected correlations between parent and child report have focused on parents' own anxiety levels, but may also be dependent on other factors including age and gender. The real limitation is, however, that parents' reports reflect their own anxiety about child health or behavior over and above more objective indicators. At the same time, children may be reluctant to discuss their illness with parents, knowing that this causes distress, even if they are not sure exactly why.

We cannot, therefore, assume that parents' reports will inevitably match those of their child in all situations. Parents may be quite accurate reporters as far as some situations or behaviors are concerned. Parents appear well able to identify "externalizing" or acting out problems. They are less able to identify "internalizing" problems such as anxiety or sadness (Edelbrock, Costello, Dulcan, Conover, & Kala, 1986). In addition, they lack direct information which enables them to make competent ratings about difficulties the child experiences at school or in interactions with friends.

(ii) Self-ratings

Limitations in cognitive or linguistic skills raise unique methodological issues and have often been used as an argument against measuring symptoms or quality of life directly from children. First, it has often been assumed that children are less able, or even unable, to locate and identify pain with any reliability. In addition, they do not always use the same language as adults. For any child, treatment can be very painful. For the youngest, this may be aggravated by an inability to understand the reason for the pain. Second, the behavior of families appears important in how children express pain. Parents may influence children by modeling distress themselves or by differentially reinforcing inappropriate behavior. Parents who communicate anxiety, or repeatedly apologize about the treatment, reinforce distress behavior (Blount, Landolf-Fritsche, Powers, & Sturges, 1991). In contrast, parental use of distraction has been associated with less child distress. Although the magnitude of these relationships is generally small, the implication that child distress can be influenced by specific parental behavior has considerable implication for clinic staff.

Concern about both these issues means that children's distress has often gone unrecognized. However, several recent studies suggest that it is possible to quantify how young children experience pain. McGrath and McAlpine (1993) used structured play and story telling tasks and concluded that from 18 months of age children were able to say that a pain hurts, localize it, and make efforts to alleviate pain and recognize pain in someone else. Children at this age are aware that they can control the extent of pain through hugs and kisses or with asking for medicine. By three or four years children spontaneously use distraction and report that playing makes them feel better. The available evidence therefore suggests that the majority of children should be able to assess pain and other physical symptoms or quality of life for themselves.

If adults, parents, and staff are imperfect reporters of a child's health and well-being, how far can we rely on children to report their own illness experience? A number of objections have been raised to using children as their own reporters. Some of these are real and practical. Often children are very young, for example the mean age on diagnosis of leukaemia is four years of age (Boring et al., 1994). Children can be very ill, especially on diagnosis. As a consequence of their disease, they may have physical limitations which compromise their ability to make written or verbal responses. It is highly unlikely that any single measure could be developed which would tap the range of issues important throughout childhood, necessitating separate but related measures appropriate for different age groups.

(iii) Observational measures

A number of methods have been devised to assess the immediate pain of medical procedures. These include general measures, such as the Neonatal Facial Action Coding System (Grunau & Craig, 1987) and the Children's Hospital of Eastern Ontario Pain Scale (McGrath et al., 1985). Within pediatric oncology, the most widely used system (Jay & Elliot, 1984) includes provision for continuous behavioral recording in 15 second intervals and a weighted score of severity of distress for each of 11 behavioral categories assessed. Scores correlate well with physical parameters but not with child self-reports. Only one scale has been developed to assess longer lasting pain, although this is specifically for use with children with cancer (Gauvain-Piquard, Rodory, Rezvani, & Lemerle, 1987).

Recognition of the limitations of current methodology is now widespread and much has been written to suggest alternative approaches. Many argue that "within" group differences are more meaningful than "between" group differences. Thus, it is increasingly common to look for differences within a sample of children with a chronic disease, attempting to define subgroups who stand out in terms of particularly good or poor adjustment, rather than assume that sick children deviate from a normative sample. Also, many argue that it is important to pay careful attention to developmental issues involved. The kind of stressors experienced by child and family are age specific. Thus in studying the impact of hearing loss in children, measures should be developed which specifically tap the kind of difficulties experienced by parents of preschool children (Quittner, 1992). Such an approach is likely to be much more helpful from a practical point of view than the more global approaches adopted in the past, in that it allows for interventions to focus on the particular difficulties experienced by families.

While standardized, quantitative instruments are often considered desirable, they can be limited. Often the research question cannot be answered by use of a standardized instrument, since such an instrument is not available. In these circumstances, researchers are increasingly turning to qualitative methodologies, either alone or in conjunction with more standardized assessment procedures: "There are no easy answers to the kinds of assessment problems that pediatric psychologists commonly encounter" (LaGreca & Lemanek, 1996, p. 137). The nonavailability of instruments to measure psychological or social functioning in health psychology, and especially child health psychology, is often a main obstacle to the establishment of collaborative work with pediatricians. At the same time, decisions about use of alternative instruments which appear to measure the same concept, can generate much unproductive debate. In all of this, the inference is often that similar issues would not be a problem if psychology was a "science" in the true sense. In point of fact, many parallel dilemmas can be identified in medicine itself. It is not necessarily the case that clinicians would agree about the definition of asthma, or about the best method to assess severity. Debates about the most appropriate method to use are therefore not confined to psychology alone.

8.27.5 EMERGING ISSUES

8.27.5.1 Neuropsychology

Pediatric neuropsychology has a great deal in common with developmental and child clinical psychology, and because of these associations there is much to be gained by the sharing of methods and measures used by each discipline. More importantly, the relationship between behavior and the brain can only be addressed adequately by studying the clinical, developmental, and psychosocial dimensions of childhood disorders (Fletcher, Taylor, Levin, & Satz, 1995). Researchers in pediatric neuropsychology are becoming increasingly aware of avoiding or at least recognizing sample biases, identifying risks for adverse clinical outcomes, and exploring the implications of findings for the development of appropriate interventions. Such advances require both methodological sophistication and active interdisciplinary collaboration (Rourke, Bakker, Fisk, & Strang, 1983).

Since the 1980s there has been a substantial increase in the number of papers devoted to neurodevelopmental issues in behavior (Rourke et al., 1983; Taylor & Fletcher, 1995). As the central nervous system (CNS) is in a state of rapid change during childhood, dysfunction is typically reflected in a failure to develop certain cognitive and behavioral capacities rather than a loss of previously acquired skills. The numbers of children at risk are relatively high, some 100 000 children below the age of 15 years are hospitalized annually for traumatic brain injury (Snow & Hopper, 1994). A substantial number of these demonstrate injury-related impairments in cognitive functioning and behavior (Fletcher & Levin, 1988). Other conditions associated with learning and behavior problems include epilepsy, spina bifida, and very low birth weight. Some intellectual deterioration has been noted in children with leukemia treated by CNS irradiation, especially where treatment occurred in the first five years of life (Eiser, 1991). Others have noted some loss of function in children with insulin dependent diabetes (Kovacs, Ryan, & Obrosky, 1994).

Neurological impairment can have a significant impact on the overall functioning of a growing child or adolescent, whether the impairment is acquired or congenital. Although most children show at least some capacity for development or recovery of function, it is very difficult to predict the extent to which this will occur in any individual case. Studies have shown that the rate and degree of recovery and subsequent sparing of function appear to depend mainly on location, extent, and development of the lesion; exposure to environmental stimulation; and demographic factors such as sex and age (Rourke et al., 1983).

8.27.5.1.1 Neuropsychological assessment

One of the primary aims of neuropsychological assessment is to produce a reliable and valid description of the relationships between brain and behavior, by sampling those abilities that are thought to be subserved by the brain. This is followed by an understanding of those relationships, involving statements regarding the probability that the particular pattern of abilities and deficits exhibited by the child is due to compromised cerebral functioning as opposed to factors unrelated to brain functioning. This distinction may have important implications for the specification of the rehabilitative program that is recommended for the child. Neuropsychological assessment of children with brain injury is also important for making diagnoses, and for accurate specification of the child's adaptive abilities and deficits for the purpose of remedial planning (Rourke et al., 1983).

One of the most difficult aspects of the neuropsychological assessment of children is that of the formulation of satisfactory prognostic statements. Very little longitudinal research has been carried out in a number of crucial areas of concern to child clinical neuropsychologists. One possible exception is the area of learning disabilities (Rourke et al., 1983). Studies (Rourke & Orr, 1977; Staz, Taylor, Friel, & Fletcher, 1978) have yielded useful information regarding the sorts of developmental markers that indicate which children are at risk for academic learning disabilities over the long term. Such studies can also give important information about more transitory deficit patterns that are not associated with longer term consequences. Similar information may be useful for furthering understanding of neuropsychological assessment in other contexts.

Neuropsychological assessment is invariably limited to intellectual functioning. This is as much a consequence of the availability of standardized measures to assess intellectual function, as recognition of those behaviors which are most vulnerable. Paradoxically, social and emotional functioning appear to be sensitive following neuropsychological damage, at least for children with leukemia treated with CNS irradiation (Butler & Copeland, 1993) and in pediatric AIDS patients (Moss, Wolters, Brouwers, Hendricks, & Pizzo, 1996). A greater innovation in method is necessary to assess the social and behavioral consequences, although video recordings have been used to assess with some success behavior (Moss et al., 1996).

8.27.5.2 Adherence

For parents, encouraging young children to adhere to prescribed medication and behavioral treatment can be a major problem. For different reasons, adherence can continue to be a contentious issue among older children and adolescents. In addition, adherence has wide reaching implications for scientific progress. It is important that patients do what they are supposed to do so that treatments can be evaluated for their efficacy, safety, or possible side effects. Much attention therefore has been paid to questions of adherence.

The terms adherence and compliance are used almost interchangeably, though they have different implications. Compliance, perhaps the more traditional term, has connotations of strict adoption and execution of medical advice. As a concept, adherence is preferred by psychologists in that there is an implication that patients do not, and should not, be expected to follow advice blindly. Patients may be more adherent with regard to some aspects of their treatment compared with others. Especially with regard to long-term conditions, patients (and parents) may make modifications to their recommended protocol, depending on what they believe best suits them (Deaton, 1985).

Not surprisingly, pediatric adherence is often reported to be poor. This can include a range of behaviors, such as failure to keep appointments, or to take medication as prescribed. Nonadherence in the pediatric population varies from 20–80% (Litt & Cuskey, 1980), although much of this is restricted to children with acute, rather than chronic conditions. High rates have also been reported for children suffering from chronic or life-threatening conditions such as cancer (Teta et al., 1986), cystic fibrosis (Muszynski-Kwan, Perlman, & Rivington-Law, 1988) and renal disease (Korsch et al., 1978). Work on pediatric compliance has been biased toward a few diseases (asthma, diabetes, and juvenile arthritis) with the most common focus being on medication use and blood and urine testing in diabetes. Other aspects of compliance (appointment-keeping, diet) have been little studied (Dunbar-Jacob, Dunning, & Dwyer, 1993).

Part of the variability in results can be attributed to different methods to measure adherence. Patient reports, clinician estimates, and pill counts are often used, although all are subject to reporting bias. Physiological indicators such as blood serum assays to determine concentration of the prescribed drug in the blood are thought to be more objective methods. The difference is between attempting to determine adherence in children with diabetes by asking them what they have eaten compared with monitoring blood sugar levels.

Failure to adhere with treatment is understandable from the child's point of view. Some medications are associated with changes in physical appearance which are likely to be unacceptable to a young person. Immunosuppressants used following heart transplants for example are associated with cushingoid features and puffy faces. Steroids used to control cancer are associated with weight gain and a generally bloated look. Nonadherence increases with the length and complexity of treatment (Haynes, Taylor, & Sackett, 1979) and extent to which treatment disrupts everyday life (Friedman & Litt, 1987). Family variables may also influence adherence. Poor family communication, and negative mother–child relationships have been implicated in several studies (Chaney & Peterson, 1989). Although much attention has been paid to questions of how children understand illness, the issue of how far this understanding influences coping with illness, compliance with treatment, general behavior, and decisions about the future, are less well-addressed.

Although most work has been atheoretical, the "health belief model" (Becker, Maiman, Kirscht, Haefner, & Drachman, 1977) has been applied with some success (see Chapter 8.01, this volume). According to this model, the individual's readiness to comply with a prescribed regimen is dependent on perceptions of susceptibility to the specific illness, perceived severity of the condition, and the expected benefits of treatment against the barriers or difficulties of adherence. Using such a model, the importance of children's own cognitions and attitudes in determining their readiness to take prescribed medication has been emphasized (lannotti & Bush, 1993). It is likely, however, that other variables are involved when considering chronic, or life-threatening, rather than acute conditions.

8.27.5.3 Quality of Life

Nowhere does the link between psychology and medicine become as apparent as in the potential implications of chronic disease for quality of life. This applies to both the short- and longer-term. In the short term, children can experience lengthy hospitalization, school absence, broken friendships, and physical discomfort. In the longer term there may be physical problems such as growth impairment and infertility (Shalet, 1989), respiratory (Jenney, Faragher, Jones, & Woodcock, 1995) or cardiac damage (Lipschultz, Colan, & Gelber, 1991) as well as educational and psychological problems (Eiser & Havermans, 1994). While pronounced in children with cancer, longer term difficulties, particularly with growth and education, characterize many other chronic childhood conditions. On the surface then, there are many reasons to suppose that children with chronic illness might experience a compromised quality of life (QoL).

Health-related QoL refers to both the subjective and objective impact of dysfunction associated with illness, injury, or treatment. Investigators are agreed that QoL is a multifactorial concept, with implications for functioning along at least three dimensions:

physical, mental, and social (Mulhern et al., 1989). Any agreement about measurement, however, seems some way off.

Indeed, it has only recently been recognized that there is a need for a measure of patients' perceptions of the intensity of symptoms and limitations on daily life associated with different treatments. These patient perceptions are considered to be central for QoL, reflecting the uniqueness of individual experience. Although this has been addressed by several groups of researchers in adult patients (Aaronson, 1991) it is only recently being appreciated that there is a need for similar information specific to children.

8.27.5.3.1 Why measure quality of life?

(i) As a basis of interventions

More systematic attention to the determinants of QoL in children is needed as a basis for appropriate interventions. This may include interventions based on the social rather than the physical consequences of a disease. Work with young people with diabetes, for example, suggests that improvements in self-care can be achieved where friends are informed and involved in daily therapy (LaGreca, 1990). With few exceptions (Varni & Setoguchi, 1991), interventions have targeted physical symptoms at the expense of more social or behavioral consequences. As issues of importance to the child are increasingly understood, it should be possible to optimize outcomes by developing more appropriate interventions.

(ii) To compare clinical trials

As survival rates have improved there has been a recognition of the need for more sensitive and comprehensive measures of outcome. Such information may have implications for planning of future randomized studies. It is widely assumed that improvements in care and prognosis are best achieved in the context of controlled trials (Smyth et al., 1994). In cancer, for example, cranial irradiation is no longer routinely offered to standard risk children with acute lymphoblastic leukemia, following recognition of damage to long-term cognitive function or growth impairment identified in previous trials. In other diseases, it is important to know how far children perceive changes in medication to be beneficial or not. For children with asthma, for example, how are increases in steroid treatment reflected in psychological measures of QoL, regardless of changes in physical function? Accurate measures which reflect the impact of treatment from the child's perspective are urgently needed and should be seen to be an integral part of the evaluation of new treatments.

8.27.5.3.2 Domains of quality of life

A number of disease specific measures to assess QoL have been developed. These include QoL in children with asthma (West & Sammons, 1994), diabetes (Challen et al., 1988), and cancer (Eiser, Havermans, Craft, & Kernahan, 1995). Most are developed from interviews or focus groups involving children and their families, supplemented by information from medical and nursing staff. The general format involves asking patients to rate a series of statements on Likert-type scales. With the exception of West and Sammons (1991) who developed separate age-appropriate measures, they target a wide age range. To date there has been little attempt to develop a measure suitable for younger children and it has yet to be established whether or not this will prove possible.

As shown in Table 1, which summarizes a number of measures developed specifically for work with children with cancer, there is some overlap in the kinds of domains identified to be important in QoL, but not as much as might be expected. In part, this lack of agreement is likely to be attributable to the poorly developed theoretical understanding of the concept of QoL. To some, QoL is in danger of becoming an umbrella term, "under which are placed many different indexes dealing with whatever the user wants to focus on" (Feinstein, 1987, p. 18). In the absence of a theoretical model about how individuals make judgments about their quality of life, there will remain a confusion in terms of what is being measured and lack of consensus about what constitutes appropriate domains. We are left with a situation in which we are forced to recognize that QoL is what is measured by QoL instruments, where validation of new measures is made against older measures and where for any single measure, what is actually measured may relate very little to patient experience.

We recognize that QoL is an important issue for patients with chronic conditions. We know also that despite severe problems, many manage to maintain or even enhance their psychological well-being over preillness levels. How is this achieved? The question of how patients achieve satisfactory QoL is less discussed and remains an empirical rather than theoretical issue. Explanations based on discrepancies between real and ideal self may have some potential.

In Rosenberg's model (1986), it is assumed that there are multiple aspects of the self that

Table 1 Cancer-specific quality of life scales for children.

Scale	*Components*	*Respondents*	*Age range*	*Validity*
Play performance (Lansky et al., 1985)	None	Parents, physicians	1–16 years	Global function, research interviews
Quality of well-being (Bradlyn et al., 1993)	Mobility, physical function, social activity, symptoms	Parents	4–18 years	Play performance, treatment toxicity
Multiattribute health status (Feeny et al., 1992)	Mobility, cognition, sensation, pain, self-care, fertility, emotion	Physicians	8–25 years	Population norms
Quality of life (Goodwin et al., 1994)	Physical function, emotional distress, reaction to treatment	Parents	very wide	Play performance, child behavior, checklist, depression

develop through internal evaluations and interpersonal experiences. In response to external events, different aspects of the self can be enhanced or modified. Two of the hypothesized components of self are the ideal self and the actual self. Ideal self refers to an individual's aspirations about who they would like to be; actual self refers to concepts of who one really is. Individuals are motivated to achieve a match between ideal and actual self in order to reduce tension or discomfort, thus alleviating experienced depression or anxiety. Thus, when external events threaten a person's self-concept, different aspects of the self are altered in attempts to maintain actual-ideal self-comparability and reduce psychological distress. It is predicted that a good match between actual and ideal self is associated with good psychological adjustment. The theory is able to cope with the finding that people with a life threatening condition become adjusted over time, perhaps through a process of making adjustments to actual–ideal self discrepancies.

Some empirical work with adults supports this view. One study found that elderly women with cancer had lower actual and ideal self-ratings than women without cancer, but were no different in terms of self-discrepancies and adjustment (Heidrich & Ward, 1992). This may suggest that adjustment to a life-threatening situation can be achieved by reducing the discrepancy between real and ideal self.

Given the threatening nature of a diagnosis of a chronic condition, we might predict that greater actual–ideal self-discrepancies would occur around diagnosis which would subsequently be reduced over time. In this view, close correspondence between actual and ideal self at diagnosis may be nonadaptive in that appropriate adjusting mechanisms have not been brought into play. Discrepancies which persist after diagnosis might be expected to be associated with greater psychological dysfunction. The model also offers a means to test empirically the outcome of evaluations; we might predict that successful interventions would result in a closer match between actual and ideal self.

To what extent might this model be applied to children as well as adults? Given appropriate methods (Eder, 1990) it has been shown that children as young as three and a half years appear able to report consistent and meaningful concepts, suggesting an awareness of psychological self-understanding equivalent to actual self. There appears to be little empirical work regarding the development of the ideal or future self in children. However, we have some anecdotal evidence that children with chronic illness at least anticipate that some changes in their psychological make-up are attributable to illness and expect these to change with future (anticipated) improvements in health. A child with asthma reported that he would be less moody when he "outgrew" his asthma; a 12-year-old girl with a newly diagnosed bone tumor attributed her current emotional state of weepiness to the diagnosis, and suggested this was not the way she really was, or would be again when treatment was complete.

8.27.5.4 Family and Ethnic Issues

There is no doubt that the diagnosis of chronic illness has a major and potentially very damaging effect on family life. Conversely, the family may be the single most important factor in determining how well the child is able to adapt and cope with the situation. Family studies have most frequently been rooted in the circumplex model (Olson, Russell, & Sprenkle, 1983), which assesses two dimensions of family functioning: cohesion and adaptability. Cohesive families may be further described as disengaged, separated, connected, or enmeshed. In terms of adaptability, the defined levels include rigidity, structure, flexibility, or chaotic.

Family functioning appears to be related to the duration as well as characteristics of the disease. In diabetes, for example, the correlation between good metabolic control and family functioning which is often found following diagnosis, was substantially reduced with a longer duration of the disease (Hanson et al., 1989).

Community studies of children's mental health point to the importance of family structure. The moderating effect of both parents being at home appears as important for the well-being of children with chronic illness as the rest of the population (Gortmaker & Sappenfield, 1984). Behavioral consequences of a chronic condition were magnified where one of the biological parents was absent. Children with chronic conditions from single parent families have been reported to have poorer adherence to treatment (Christiaanse, Lavignes, & Lerner, 1989) and poorer school attendance (Anderson, Bailey, Gasper, Palmer, & West, 1983) compared with those with similar conditions from two parent families.

Most work concerned with the impact of illness on family coping and personal relationships has focused on responses from Caucasian, middle-class groups. However, variations in family structure are likely to be especially important. In addition, beliefs about the cause or etiology of disease are likely to be mediated by ethnic or cultural differences. The unique difficulties experienced by children with chronic conditions from non-Caucasian backgrounds has received virtually no attention. This is especially surprising given that, outside the realm of chronic illness, many ethnic minorities have problems which are specific to them, and often which are not well recognized by the healthcare system. For example, a number of risk factors associated with the use of illicit drugs were identified to be specific to African-American youth (Farrell, Danish, & Howard, 1992). This group have also been found to be at the highest risk for homicide (Yee, 1995), and are more likely than other racial or ethnic groups to be injured in violent episodes (Sheley, McGee, & Wright, 1992). Both African- and Latin-Americans and are at higher risk for HIV/AIDS than the rest of the population (Centers for Disease Control, 1995). "Unique social, cultural and ethnic antecedents" (Yee, 1995) must be incorporated into current models of risk behavior if progress is to be made in closing the gap between certain high-risk ethnic minorities and the rest of society.

Some recent studies have focused on conditions which exclusively, or predominantly affect certain ethnic groups (e.g., sickle cell disease; Lemanek, Moore, Gresham, Willianson, & Kelley, 1986). In these cases, greater care needs to be given to disentangling any confounds between ethnicity and social class (Lemanek et al., 1986). Piecemeal application of measures developed for work with Caucasian children may be inappropriate, given differences between ethnic groups in the meaning and implications of illness (Penn, Kar, Kramer, Skinner, & Zambrana, 1995). There is a need for more ethnic-specific services and care within the present healthcare systems, in the UK and in the USA.

8.27.6 COMMUNICATION

8.27.6.1 Understanding Health and Illness

As late as the 1950s and 1960s, it was considered appropriate that children should not be informed about their illness or involved in decision-making about alternative treatments. To some extent, this view is understandable when we consider the poor prognosis for most children who experience chronic illness or accidental injury. It was feared that information would be unnecessarily distressing for the child and would result in poorer adherence and cooperation with treatment. Over the years, attitudes have changed considerably. This can be attributable to the much improved prognosis for many conditions; it is no longer the case that information given is inevitably about the severity or life threatening nature of the condition. There are also many more legal and ethical issues which have been raised making it more standard practice to involve children in all aspects of health care. While the consensus is that children need to be involved, questions about what they should be told, and how or by whom, are much less resolved. A central issue in pediatric health psychology in recent years has been an increasing interest in how healthy children understand the concepts of health and illness, and how these beliefs are modified by personal and vicarious experience of illness.

With respect to the question of how sick children understand their illness, the research tool most often used has been the semistructured interview, with the general questions about minor childhood illnesses being replaced by questions about the the child's own condition. When the responses are categorized into one of Piaget's stages it appears that the developmental sequence described for healthy children also accounts for development in children with chronic conditions such as asthma, epilepsy or arthritis (Beales et al., 1983; Perrin, Stein, & Drotan, 1991). One area of concern however is how poorly informed many children appear to be about their illness. Part of this problem arises because often the child is diagnosed at a very young age when information is necessarily directed to the parents (Eiser & Kopel, 1996). There is also the issue of parental avoidance when the child mentions the illness, and in some families inappropriate blanket explanations are preferred regardless of the ages of the children concerned. Incongruities occur between what the parents want the child to know and what the child wants to know. Most children want more information, and that it is repeated (Canam, 1986). It is not clear how or when information is given to the child, and unfortunately many children learn about their disease through rather unsatisfactory sources such as friends, relations, or the media (Kendrick, Culling, Oakhill, & Mott, 1986). There is some evidence that the children who adjust well are those who are told about their illness on diagnosis or shortly after diagnosis (Claflin & Barbarin, 1991), but there has really been insufficient work in this area for definitive conclusions to be drawn.

Many chronic conditions can involve long and painful treatment, and in the case of young children parents make treatment decisions. With older children, however, the question arises as to how far they should be able to intervene in their own treatment. Despite the importance of this issue, there is little work concerned with how children and adolescents come to understand treatment alternatives and make reasoned decisions. One exception is the work by Weithorn and Campbell (1982), who developed a series of scenarios involving either psychological or medical treatment decisions for four different problems: diabetes, epilepsy, depression, and enuresis. The study compared the performance and decisions made by subjects aged 9, 14, 18, and 21 years. Results showed that minors (aged 14 years) demonstrated a level of competence equivalent to that of adults. Nine-year-olds appeared less competent on measures of understanding and rational reasons, but did not differ from adults on measures of treatment choices and reasonable outcome alternatives. The authors conclude that the majority of adolescents are able to make competent decisions about their own welfare before the currently accepted age of 18 years. However, the way in which children and adolescents make decisions in hypothetical situations may be very different from the choices they make when faced by real-life treatment dilemmas. Unfortunately moral and ethical objections mean that real-life decision making is particularly difficult to study, and evidence as to how well children are able to understand the concept of a clinical trial, or make rational decisions about their own treatment is scarce.

An exception is a study by Postlethwaite et al. (1995) in which the psychological consequences of involving children in decisions about treatment options was explored. As a consequence of renal disease, the children were offered growth hormone treatment to counteract retarded growth. Decisions about whether or not to agree to the treatment were difficult for parents and child. The majority of parents were able to understand the issues involved, but ease of decision-making was not related to understanding. Choices were influenced by the views of pediatricians, emotional factors and personal style. Decisions were not dependent on physical factors, such as the child's height deficit, concern about growth or understanding of the risks of treatment. The children (aged between 9–14 years) had much more difficulty understanding the information than their parents. Even so, most felt that they had been involved in decision-making. The authors highlight the need to provide children with age-appropriate information, but again, fail to specify what this might entail.

One of the main problems in this area is the failure to develop reliable methods of assessment for younger children with limited language and reading skills. To begin to understand more clearly what children know about their illness, child-centered instruments need to be developed which are less dependent on language than the traditional semistructured interview approach, and which can be used with sick or handicapped children as well as healthy children.

8.27.6.2 Genetic Testing

Recent developments in molecular genetics have enabled clinicians to test directly for the gene responsible for many inherited diseases. Where conditions do not manifest themselves until adulthood this poses considerable ethical problems when trying to decide whether to perform such tests on children and adolescents (see Chapter 8.17, this volume). The same dilemma occurs regarding children who have inherited genes for recessive disease, (such as cystic fibrosis) which, although having no impact on the individual's own health, may have important implications for the health of their own children at some later stage. The advent of molecular genetic screening programs for the general population may also create problems in that a family may be informed that their child has a disease for which no cure is available. In situations such as this it is difficult to see the benefits of such programs. Newborn and infant screening can be advantageous in that it provides the parents with useful information regarding future family planning and reproductive decisions, as well as sparing the parents a long period of anxiety when they realize that the child has a problem which has not yet been diagnosed (Michie, 1995). However, because of possible costs such as the marring of many potential years of happiness, and damaged family relationships, the balance between the pros and cons of across the board childhood genetic screening must be carefully balanced before being brought into general practice (Michie, 1995). This is also true with regard to the testing of children in families with histories of certain genetic diseases. Although in recent years there have been many requests to test minors presymptomatically, the general consensus among clinicians advises against testing due to fears that such knowledge could distort the child's future emotional or educational development, or blight their lives much earlier than the disease would given its natural course (Michie, 1995). Family problems may also arise in cases whereby one child in the family is identified to be at high risk and another at low risk or unaffected. The hope is that in these circumstances there may be some individuals who would come to terms with their result more successfully than they would if the test had been postponed until adulthood. In practice, however, it is very difficult either to support or refute this argument. Although the abilities of children to consider the issues of consent to genetic testing should not be underestimated, there are as yet no strategies to assess a child's understanding of all that is involved (Michie, 1995). Assumptions about children's cognitive functioning should not be used to exclude them from the decision-making process.

There are other ethical as well as practical problems when dealing with the issue of childhood testing. One of the main considerations is that testing removes the individual's future right to make their own decision to be tested as an adult. This is a particularly notable concern when adult predictive testing data is examined. For example, far fewer adults with Huntington's disease have actually chosen to have the test than was expected based on studies looking at expected uptake before the test was available (Crauford, Dodge, Kerzin-Storrar, & Harris, 1989). However if the decision had been made for them to have the test in childhood they would not as an adult have the option to "opt out" of having the test. Other concerns include the fact that testing children is not completely confidential in that the parents will know the results, and that knowledge of a child's genetic status may alter future upbringing and pattern of relationships with peers and other family members. It seems appropriate to advise caution in carrying out predictive genetic tests on children until more information is available about the effects of such tests. Costs such as damaged self-esteem and family relationships, and possible long-term health and life insurance difficulties, must be weighed against potential benefits, such as improved coping and the release of parental anxiety and concern over the child's future health.

8.27.7 PROFESSIONAL ISSUES

There is inevitably some friction between psychologists and pediatricians in the way in which they approach research and clinical issues concerned with chronically sick children. Although both share some ideals about what should be done to help children with chronic conditions and their families, their training and experience does not necessarily mean that they share the same research agenda (Stabler, 1988). Some compromises have to be made. Perhaps the major obstacle is in defining the question to be asked. Too much work which has started from a psychological perspective has proved to have little impact in a practical context. Many might argue that implications for practice should not be the sole criterion against which success should be considered. It might then be expected that contributions to theoretical development would be important. In fact, this does not always appear to be so. Some of the questions asked in psychological work, such as identifying "adjustment" in heterogeneous groups is of very limited value in a clinical context (Lavigne & Faier-Routman, 1992). Too often, in psychological work, a group of patients are used as a convenience sample to test some aspect of theory, regardless of specific relevance. In these situations, clinical implications are limited to a brief paragraph at the end of the paper, often highly speculative.

For many, the honeymoon period for pediatric psychology is over. A survey of published work (Roberts et al., 1996) suggests that the most common type of research is explicative, that is, studies attempt to show an association between medical and psychological variables. Some of this work is well-represented in the previous sections. In contrast, research aimed more directly at issues of clinical concern, including assessment, prevention, and intervention, is at a premium.

In contrast to psychological research, medical papers start very much from a practical question. The methods, however, are often not specified with sufficient detail to allow for replication, and the results are often discussed within a narrow clinical context. There is rarely any attempt to extrapolate from this context to more general practical or theoretical issues. The paper that takes a real practical issue and applies a satisfactory level of methodological expertise and statistical rigor to the analysis is rare indeed. Questions should be theory-driven, but sensitive to clinical needs (LaGreca & Lemanek, 1996). Such an ideal is dependent on the development of theories that lend themselves to clinical application. Disappointingly, many theories in mainstream psychology have been developed in relation to specific contexts and do not lend themselves readily to other situations. In practice, it is often difficult to find a theory which is appropriate for the clinical setting in question.

8.27.8 CONCLUSIONS

Pediatric psychology has been defined as an interdisciplinary field addressing physical, cognitive, social, and emotional functioning and development as they relate to health and illness issues in children, adolescents, and families. It aims to explore the relationship between the psychological and physical well-being of children, adolescents, and families, including: psychosocial and developmental factors contributing to the etiology, course, treatment, and outcome of pediatric conditions; assessment and treatment of behavioral and emotional concomitants of disease, illness, and developmental disorders; the role of psychology in health care settings; behavioral aspects of pediatric medicine; the promotion of health and health-related behaviors; the prevention of illness and injury among children and youth, and issues relating to training.

The goals of this branch of psychology are wide-ranging and ambitious. In the relatively brief history of the discipline, much has been achieved, though it is notable that attention to some issues is disproportionately high compared with others. Many more resources have been channeled toward exploring the relationship between psychological and physical well-being compared with assessment and treatment issues. There has generally been more of a focus on children with chronic, rather than acute conditions, and on the etiology of emotional and behavioral consequences of disease compared with lifestyle behaviors in the general population. There is a focus on illness rather than developmental disorders; a fleeting rather than comprehensive study of psychological issues in health care settings. Interest in injury prevention and training is small, but increasing.

There are strengths and weaknesses in a discipline which has such ambitious goals especially where many of these overlap with the interests of other professionals. Certainly some of the defined goals might be considered to be more the remit of clinical rather than pediatric or health psychology (e.g., developmental disorders). The nature of much of the work involves collaboration with other professionals; clinicians, teachers, and social services. Issues of training need particularly to address the potential obstacles to successful collaboration (Stabler, 1988). The rewards inherent in successful collaboration are nevertheless much higher than where collaboration is limited to *post hoc* attempts to speculate about the potential value of research for practice.

Despite the criticisms that can be made, there is no doubt that the establishment of a child health psychology in itself has done much to increase the profile of work with children. The difficulties involved may be considerable, but these also contribute to its attractiveness. That children have a unique perspective on life, despite the hardships of a serious illness, is without debate. From such a perspective, and for many people, work with children is therefore not only humbling, but intrinsically more satisfying, than work with other groups.

8.27.9 REFERENCES

Aaro, L. E., Laberg, J. C., & Wold, B. (1995). Health behaviors among adolescents: towards a hypothesis of two dimensions. *Special issue: Measured in health education research. Health Education Research, 10,* 83–93.

Aaronson, N. K. (1991). Methodological issues in assessing the quality of life of cancer patients. *Cancer, 67,* 844–850.

Alwash, R., & McCarthy, M. (1988). Measuring severity of injuries to children from home accidents. *Archives of Disease in Childhood, 63,* 635–638.

Anderson, H. R., Bailey, P. A., Cooper, J. S., Palmer, J. C., & West, S. (1983). Morbidity and school absence caused by asthma and wheezing illness. *Archives of Disease in Childhood, 58,* 777–784.

Baker, S. P., O'Neill, B., & Karpf, R. S. (1984). *The injury fact book.* Lexington, MA: D. C. Health.

Bar-Or, O., & Malina, R. M. (1995). Activity, fitness, and health of children and adolescents. In L. Cheung and J. Richmonds (Eds.), *Child health, nutrition and physical activity.* Champaign, IL: Human Kinetics.

Beales, J. G., Holt, P. J. L., Keem, J. H., & Mellor, V. P. (1983). Children with juvenile chronic arthritis: Their beliefs about their illness and therapy. *Annals of the Rheumatic Diseases, 42,* 481–486.

Bearison, D. J., & Pacifici, C. (1989). Children's event knowledge of cancer treatment. *Journal of Applied Developmental Psychology, 10,* 469–486.

Becker, M. H., Maiman, L. A., Kirscht, J. P., Haefner, D. P., & Drachman, R. H. (1977). The health belief model and prediction of dietary compliance: A field experiment. *Journal of Health and Social Behavior, 18,* 348–366.

Bennetts, K., Borland, R., & Swerissen, H. (1991). Sun protection behavior of children and their parents at the beach. *Psychology and Health, 5,* 279–287.

Berglas, S. (1986). A typology of self-handicapping alcohol abusers. In M. J. Saks & L. Saxe (Eds.), *Advances in applied social psychology* (Vol. 3, pp. 29–56). Hillsdale, NJ: Erlbaum.

Bibace, R., & Walsh, M. E. (1980). Development of children's concepts of illness. *Pediatrics, 66,* 912–917.

Billson, A. L., & Walker, D. A. (1994). Assessment of health status in survivors of cancer. *Archives of Disease in Childhood, 70,* 200–204.

Block, J., Block, J. H., & Keyes, S. (1988). Longitudinally foretelling drug usage in adolescence: Early childhood personality and environmental precursors. *Child Development, 59,* 336–355.

Blount, R. L., Landolf-Fritsche, B., & Powers, S. W. & Sturges, J. W. (1991). Difference between high and low coping children and between parent and staff behaviors during painful medical procedures. *Journal of Pediatric Psychology, 16,* 795–809.

Boring, C. C., Squires, T. S., Tong, T., & Montgomery, S. (1994). Cancer statistics. *CA Cancer Journal, 44,* 7–26.

Botvin, G. J., & Wills, T. A. (1985). Personal and social skills training: Cognitive-behavioral approaches to substance abuse prevention. In C. S. Bell & R. Battjes (Eds.), *Prevention research: Deterring drug abuse among children and adolescents.* National Institute on Drug Abuse Research Monograph 63. DHSS publication (AOM) 85-1334 (pp. 8–49). Washington, DC: Superintendent of Documents, US Government Printing Office.

Bradlyn, A. S., Harris, C. V., Warner, J. E., Ritchey, A. K., Zaboy, K. (1993). An investigation of the validity of well-being scale with pediatric oncology patients. *Health Psychology, 12,* 246–250.

Broadstock, M., Borland, R., & Gason, R. (1992). Effects of suntan on judgments of healthiness and attractiveness by adolescents. *Journal of Applied Social Psychology, 22,* 157–172.

Bukstein, O. G. (1995). *Adolescent substance abuse: Assessment, prevention and treatment.* New York: Wiley.

Butler, R. W., & Copeland, D. R. (1993). Neuropsycholgical effects of central nervous system prophylactic treatment in childhood leukemia: methodological considerations. *Journal of Pediatric Psychology, 18,* 319–338.

Cadman, D., & Goldsmith, C. (1986). Construction of social value or utility based indexes: The usefulness of factorial experimentation design plans. *Journal of Chronic Disease, 39,* 643–651.

Cadman, D. Boyle, M., Szatmari, P., & Offord, D. (1987). Chronic illness, disability, and mental and social well being: Findings of the Ontario child health study. *Pediatrics, 79,* 805–813.

Cadman, D., Rosenbaum, P. Boyle, M., & Offord, D. R. (1991). Children with chronic illness: Family and parent demographic characteristics and psychosocial adjustment. *Pediatrics, 87,* 884–889.

Canam, C. (1986). Talking about cystic fibrosis within the family: What parents need to know. *Issues in Comprehensive Pediatric Nursing, 9,* 167–178.

Cancer Research Campaign (1995). *Malignant melanoma—UK factsheet.* London, Author.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge; MA: MIT Press.

Centers for Disease Control (1995). *HIV/AIDS surveillance report.* Atlanta, GA: US Department of Health and Human Services.

Claflin, C. J., & Barbarin, O. A. (1991). Does telling less protect more? Relationships among age, information disclosure, and what children with cancer see and feel. *Journal of Pediatric Psychology, 16,* 169–191.

Challen, A. H., Davies, A. G., Williams, R. J., Haslum, M. N., & Baum, J. D. (1988) Measuring psychosocial adaptation to diabetes in adolescence. *Diabetic Medicine, 5,* 739–746.

Chaney, J. M., & Peterson, L. (1989). Family variables and disease management in juvenile rheumatoid arthritis. *Journal of Pediatric Psychology, 14,* 389–403.

Chassin, L., Tetzloff, C., & Hershey, M. (1985). Self-image and social-image factors in adolescent alcohol use. *Journal of Studies on Alcohol, 46,* 39–47.

Cheung, L., & Richmond, J. (1995). *Child health, nutrition and physical activity.* Champaign, IL: Human Kinetics.

Christiaanse, M. E., Lavigne, J. V., & Lerner, C. V. (1989). Psychosocial aspects of compliance in children and adolescents with asthma. *Developmental and Behavioral Pediatrics, 10,* 75–80.

Cleary, J., Gray, O. P., Hall, D. J., Rowlandson, P. H., Sainsbury, C. P., & Davies, M. M. (1986). Parental involvement in the lives of children in hospital. *Archives of Disease in Childhood, 61,* 779–787.

Cody, R., & Lee, C. (1990). Behaviours, beliefs and intentions in skin cancer prevention. *Journal of Behavioral Medicine, 13,* 373–389.

Crauford, D., Dodge, A., Kerzin-Storrar, L., & Harris, R. (1989). Uptake of presymptomatic predictive testing for Huntingtons Disease. *Lancet, 2,* 603–605.

Deaton, A. V. (1985). Adaptive non-compliance in pediatric asthma: The parent as expert. *Journal of Pediatric Psychology, 10,* 1–14.

Despres, J. P., Bouchard, C., & Malina, R. M. (1990). Physical activity and coronary heart disease risk factors during childhood and adolescents. *Exercise and Sports Sciences Reviews, 18,* 243–261.

Donaldson, S. I., Graham, J. W., & Hansen, W. B. (1994). Testing the generalizability of intervening mechanism theories: Understanding the effects of adolescent drug use prevention interventions. *Journal of Behavioral Medicine, 17,* 195–216.

Donaldson, S. I., Graham, J. W., Piccinin, A. M., & Hansen, W. B. (1995). Resistance-skills training and onset on alcohol use: Evidence for beneficial and potentially harmful effects in public schools and in private catholic schools. *Health Psychology, 14,* 291–300.

Drotar, D. (1981). Psychological perspectives in chronic childhood illness. *Journal of Pediatric Psychology, 6,* 211–228.

Dubow, E. F., Lovko, K. R., & Kausch, D. F. (1990) Demographic differences in adolescents' health concerns and perceptions of helping agents. *Journal of Clinical Child Psychology, 19,* 44–54.

Dunbar-Jacob, J., Dunning, E. J., & Dwyer, K. (1993) Compliance research in pediatric and adolescent populations: Two decades of research. In N. P. Krasnegor, L. Epstein, S. B. Johnson, & S. J. Jaffe (Eds.), *Developmental aspects of health compliance behavior.* Hillsdale, NJ: Erlbaum.

Edelbrock, C., Costello, A. J., Dulcan, M. K., Conover, M. C., & Kala, R. (1986). Parent–child agreement on child psychiatric symptoms assessed via structured interview. *Journal of Child Psychology and Psychiatry, 27,* 181–190.

Eder, R. A. (1990). Uncovering young children's psychological selves: Individual and developmental differences. *Child Development, 61,* 849–863.

Eiser, C. (1991). Cognitive deficits in children treated for leukamia. *Archives of Disease in Childhood, 66,* 164–168.

Eiser, C. (1993), *Growing up with a chronic disease: The* *impact on children and their families*. London: Jessica Kingsley.

Eiser, C., Eiser, J. R., & Pauwels, P. (1993). Skin cancer: Assessing perceived risk and behavioural attitudes. *Psychology and Health, 8,* 393–404.

Eiser, C., & Havermans, T. (1994). Long-term social adjustment after treatment for childhood cancer. *Archives of Disease in Childhood, 70,* 66–70.

Eiser, C., Havermans, T., Craft, A., & Kernahan, J. (1995). Development of a measure to assess the perceived illness experience after treatment for cancer. *Archives of Disease in Childhood, 72,* 302–307.

Eiser, C., & Kopel, S. (1996). Children's perceptions of health and illness. In K. Petrie & J. Weinman (Eds.) *The patients perceptions of illness and treatment, Current research and applications*. Harwood Academic Press.

Eiser, J. R., Morgan, M. J. & Gammage, P. (1987). Belief correlates of perceived addiction in young smokers. *Special issue: Juvenite substance use and human development: New perspectives in research and prevention. European Journal of Psychology of Education, 4,* 375–385.

Epstein, L. H. (1992). Exercise and obesity in children. *Journal of Applied Sport Psychology, 4,* 120–133.

Falco, M. (1992). *The making of a drug-free America: Programs that work*. New York: Times Books.

Farrell, A. D., Danish, S. J., & Howard, C. W. (1992). Risk factors for drug use in urban adolescents: Identification and cross-validations. *American Journal of Community Psychology, 20,* 263–286.

Feeney J. A., & Ryan S. M. (1994). Attachment style and affect regulation: Relationships with health behavior and family experiences of illness in a student sample. *Health Psychology, 13,* 334–345.

Feeny, D., Furlong, W., Barr, R. D. (1992). A comprehensive multi-attribute system for classifying health status of survivors of childhood cancer. *British Journal of Cancer, 10,* 923.

Feinstein, A. R. (1987), Clinimetric perspectives. *Journal of Chronic Disease, 40,* 635–640.

Finney, J. W., & Cataldo, M. F. (1991). Child and parent behavior and unintentional home injuries. In J. H Johnson & S. B. Johnson (Eds.), *Advances in child health psychology* (pp. 108–114). Gainsville, FL: University Press of Florida.

Fletcher, J. M., & Levin, H. S. (1988). Neurobehavioral effects of brain injury in children. In D. Routh (Ed.) *Handbook of pediatric psychology*. New York: Guilford Press.

Fletcher, J. M., Taylor, H. G., Levin, H. S., & Satz, P. (1995). Neuropsychological and intellectual assessment of children. In H. I. Kaplan & B. J. Sadock (Eds.), *Comprehensive textbook of psychiatry* (Vol. 6). Baltimore: Williams & Wilkins.

Friedman, I. M., & Litt, I. F. (1987). Adolescents' compliance with therapeutic regimens: Psychological and social aspects and intervention. *Special issue: Medication and drug use in adolescence. Journal of Adolescent Health Care, 8,* 52–67.

Gauvain-Piquard, A., Rodary, L., Rezvani, A., & Lemerie, J. (1987). Pain in children aged 2–6 years: A new observational rating scale elaborated in a pediatric oncology unit: Preliminary report. *Pain, 31,* 177–188.

Glynn, T. J., & Haenlein, M. (1988). Family theory and research on adolescent drug use: A review. *Journal of Chemical Dependency Treatment, 1,* 39–56.

Golding, J., Haslum, M., & Morris, A. C. (1984). What do our 10 year old children eat? *Health Visitor, 57,* 178–179.

Goodwin, D. A. J., Boggs, S. R., & Graham-Pole, J. (1994). Development and validation of the Pediatric Oncology Quality of Life scale. *Psychological Assessment, 6,* 321–328.

Gortmaker, S., & Sappenfield, W. (1984), Chronic childhood disorders: Prevalence and impact. *Pediatric Clinics of North America, 31,* 3–18.

Gortmaker, S. L., Walker, D. K., Weitzman, H., & Sobol, A. M. (1990) Chronic conditions, socioeconomic risks, and behavioral problems in children and adolescents. *Pediatrics, 85,* 267–276.

Graham, P. (1996). Alcohol and the young. *Archives of Disease in Childhood, 75,* 361–363.

Grunau, R. V., & Craig, K. D. (1987). Pain expression in neonates: Facial action and cry. *Pain, 28,* 395–410.

Hansen, W. B. (1992). School-based substance abuse prevention: A review of the state of the art in curriculum, 1980–1990. *Health Education Research, 7,* 403–430.

Hanson, C. L., Harris, M. A., Relyea, G., Cigrang, J. A., Carle, D. L., & Burghen, G. A. (1989). Coping styles in youths with insulin-dependent diabetes mellitus. *Journal of Consulting and Clinical Psychology, 57,* 644–651.

Haynes, R. B., Taylor, D. W., & Sackett, D. L. (Eds.) (1979) *Compliance in health care*. Baltimore: Johns Hopkins University Press.

Hazan, C., & Shaver, P. (1987). Romantic love conceptualised as an attachment process. *Journal of Personality and Social Psychology. 52,* 511–524.

Heidrich, S. M., & Ward, S. E. (1992). The role of the self in adjustment to cancer in elderly women. *Oncology Nursing Forum, 19,* 1491–1496.

Hergenrather, J. R., & Rabinowitz, M. (1991). Age-related differences in the organization of children's knowledge of illness. *Developmental Psychology, 27,* 952–959.

Hill, D., Rassaby, J., & Gardner, G. (1984). Determinants of intentions to take precautions against skin cancer. *Community Health Studies, 8,* 33–44.

Hobfoll, S. E. (1991). Gender differences in stress reactions: Women filling the gaps. *Psychology and Health, 5,* 95–109.

Horwitz, W. A., & Kazak, A. E. (1990). Family adaptation to childhood cancer: Sibling and family systems variables. *Journal of Clinical Child Psychology, 19,* 221–228.

Iannotti, R. J., & Bush, P. J. (1993). Towards a developmental theory of compliance. In N. A. Krasnegor, L. H. Epstein., S. B. Johnson, & S. J. Yaffe (Eds.), *Developmental aspects of health compliance behavior*. Hillsdale, NJ: Erlbaum.

Ireys, H., Werthamer-Larsson, L., Kolodner, K. B., & Gross, S. S. (1994). Mental health of young adults with chronic illness: The mediating effect of perceived Impact. *Journal of Pediatric Psychology. 19,* 205–222.

Jackson, R. H. (1988). The doctor's role in the prevention of accidents. *Archives of Disease in Childhood, 63,* 235–237.

Jacobsen, P. B., Manne, S. L., Schorr, O., Rapkin, B., & Redd, W. H. (1990). Analysis of child and parent behavior during painful medical procedures. *Health Psychology, 9,* 559–576.

Jay, S. M., & Elliot, C. (1984). Behavioral observation scales for measuring children's distress: The effects of increased methodological rigor. *Journal of Consulting and Clinical Psychology, 52,* 1106–1107.

Jenney, M. E., Faragher, E. B., Morris-Jones, P. H., & Woodcock, A. (1995). Lung function and exercise capacity in survivors of childhood leukaemia. *Medical and Pediatric Oncology, 24,* 222–230.

Jessor, R., & Jessor, S. L. (1977). *Problem behavior and psychosocial development. A longitudinal study of youth.* New York: Academic Press.

Kandel D. B., & Yamaguchi, K. (1984). Patterns of drug use from adolescence to young adulthood: 1. Periods of risk for initiation, continued use and discontinuation. *American Journal of Public Health, 74,* 673–681.

Keesling, B., & Friedman, H. S. (1987). Psychosocial factors in sunbathing and sunscreen use. *Health Psychology, 6,* 477–493.

Kendrick, C., Culling, J., Oakhill, T., & Mott, M. (1986).

Children's understanding of their illness and its treatment within a pediatric oncology unit. *Association for Child Psychology and Psychiatry (Newsletter), 8,* 16–20.

Korsch, B. M., Negrett, V. F., Gardner J. F., Weinstock, C. L., Mercer, A. S., Grushkin, A. S. et al. (1978) Kidney transplantation in children: Psychological follow-up study on child and family. *Journal of Pediatrics, 83,* 399–408.

Kovacs, M., Ryan, C., & Obrosky, D. S. (1994). Verbal intellectual and verbal memory performance of youths with childhood–onset insulin-dependent diabetes mellitus. *Journal of Pediatric Psychology, 19,* 475–483.

La Greca, A. M. (1990). Social consequences of pediatric conditions: Fertile area for future investigation and intervention. *Journal of Pediatric Psychology, 15,* 285–307.

La Greca, A. M., & Lemanek, K. L. (1996). Editorial: Assessment as a process in pediatric psychology. *Journal of Pediatric Psychology, 21,* 137–151.

Lansky, L. L., List, M. A., Lansky, S. B., Cohen, M. E., & Sinks, L. F. (1985). Toward the development of a play performance scale for children. *Medical Pediatric Oncology, 24,* 1837–1841.

Lask, B. (1996). Editorial: Reaching the parts other journals cannot reach. *Clinical Child Psychology and Psychiatry, 1,* 5–9.

Lavigne, J. V., & Faier-Routman, J. (1992). Psychological adjustment to pediatric physical disorders: A meta-analytic review. *Journal of Pediatric Psychology, 17,* 133–157.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.

Leary, M. R., Tchividjian, L. R., & Kraxberger, B. E. (1994). Self presentation can be hazardous to your health: Impression management and health risk. *Health Psychology, 13,* 461–470.

Lemanek, K. L., Moore, S. L., Gresham, F. M., Williamson, D. A., & Kelley, M. L. (1986). Psychological adjustment of children with sickle cell anemia. *Journal of Pediatric Psychology, 11,* 397–410.

Lipschultz, S. E., Colan, S. D., & Gelber, R. D. (1991). Late cardiac effects of doxorubicin therapy for acute lymphoblastic leukemia in childhood. *New England Journal of Medicine, 324,* 808–815.

Litt, I. F., & Cuskey, W. R. (1980) Compliance with medical regimens during adolescence. *Pediatric Clinics of North America, 27,* 3–15.

Lobato, D., Barbour, L., Hall, L. J., & Miller, C. T. (1987). Psychosocial characteristics of preschool siblings of handicapped and non-handicapped children. *Journal of Abnormal Child Psychology, 15,* 329–338.

Manne, S. L., Jacobsen, P. B., & Redd, W. H. (1992). Assessment of acute pediatric pain: Do child self-report, parent ratings and nurse ratings measure the same phenomenon? *Pain, 48,* 45–52.

McGrath, P. J., Johnson, G., Goodman, J. T., Schillinger, J., Dunn, J., & Chapman, J. (1985). A behavioral scale for rating postoperative pain in children. In H. L. Fields, R. Dubner, & F. Cervero (Eds.), *Proceedings of the Fourth World Congress on Pain, Seattle*. New York: Raven.

McGrath, P. J., & McAlpine, L. M. (1993). Psychological perspectives on pediatric pain. *Journal of Pediatrics, 122,* S2–58.

McGuire, W. J. (1964). Inducing resistence to persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 1). New York: Academic Press.

Meadow, R. (1992). Difficult and unlikeable parents. *Archives of Disease in Childhood, 67,* 697–702.

Michie, S. (1996), Predictive genetic testing in children: Paternalism or empiricism? In T. Marteau, & M. Richards (Eds.) *The troubled helix: Social and psychological implications of the new human genetics*. Cambridge: Cambridge University Press.

Moss, H. A., Wolters, P. L., Brouwers, P. J. M., Hendricks, M. L., & Pizzo, P. A. (1996). Impairment of expressive behavior in pediatric HIV-infected patients with evidence of CNS disease. *Journal of Pediatric Psychology, 21,* 379–400.

Mulhern, R. K., Horowitz, M. E., Ochs, J., Friedman, A. G., Armstrong, F. D., Copeland, D., & Kun, L. E. (1989). Assessment of quality of life among pediatric patients with cancer. *Journal of Consulting and Clinical Psychology, 1,* 130–138.

Muszynski-Kwan, A. T., Perlman, R., & Rivington-Law, B. A. (1988) Compliance with and effectiveness of chest physiotherapy: A review. *Physiotherapy Canada, 40,* 28–32.

Nelson, K. (1986). *Event knowledge: Structure and function in development*. Hillsdale, NJ: Erlbaum.

Nutbeam, D., Aar, L., & Catford, J. (1989), Understanding children's health behaviour. The implications for health promotion for young people. *Social Science and Medicine, 29,* 317–325.

Oldershaw, L., Walters, G. C., & Hall, D. K. (1986). Control strategies and noncompliance in abusive mother–child dyads: An observational study. *Child Development, 57,* 722–732.

Olson, D. H., Russell, C., & Sprenkle, D. H. (1983). Circumplex model of marital and family systems: VI Theoretical Update. *Family Process, 22,* 69–83.

Osborne, M. L., Kistner, J. A., & Helgemo, B. (1993). Developmental progression in children's knowledge of AIDS: Implications for education and attitudinal change. *Journal of Pediatric Psychology, 18,* 177–192.

Piaget, J. (1929). *The child's conception of the world.* New York: Harcourt Brace.

Parmelee, A. H. (1986). Children's illnesses: their beneficial effects on behavioural development. *Child Development, 57,* 1–10.

Penn, N. E., Kar, S., Kramer, J., Skinner, J., & Zambrana, R. E. (1995) Ethnic minorities, health care systems, and behavior. *Health Psychology, 14,* 641–646.

Perrin, E. C. & Gerrity, P. S. (1981). There's a demon in your belly: Children's understanding of illness. *Pediatrics, 67,* 841–849.

Perrin, E. C., Stein, R. E. K., & Drotar, D. (1991). Cautions in using the Child Behaviour Checklist: Observations based on research about children with a chronic illness. *Journal of Pediatric Psychology, 16,* 411–421.

Peterson, L. Oliver, K. K., Brazeal, T. J., & Bull, C. A. (1995). A developmental exploration of expectations for and beliefs about preventing bicycle collision injuries. *Journal of Pediatric Psychology, 20,* 13–22.

Postlethwaite, R. J., Reynolds, J. M., Wood, A. J., Evans, J. H., Lewis, M. A., & Eminson, D. M. (1995). Recruiting patients to clinical trials: lessons from studies of growth hormone treatment in renal failure. *Archives of Disease in Childhood, 73,* 30–35.

Quittner, A. L. (1992). Re-examining research on stress and social support: The importance of contextual factors. In A. M. Greca, L. J. Seigal, J. L. Wallander, & C. E. Walker (Eds.), *Stress and coping in child health* (pp. 85–115). New York: Guilford Press.

Roberts, M. C. (1986). Health promotion and problem prevention in pediatric psychology: An overview. *Journal of Pediatric Psychology, 11,* 147–161.

Roberts, M. C., McNeal, R. E., Randall C. J., & Roberts J. C. (1996). A necessary reemphasis on integrating explicative research with the pragmatics of pediatric psychology. *Journal of Pediatric Psychology, 21,* 107–114.

Robertson, J. (1958) *Young children in hospitals.* New York: Basic Books.

Rosenberg, M. (1986). *Conceiving the self*. Malabar, FL: Krieger.

Rourke, B. P., Bakker, D. J., Fisk, J. L., & Strang, J. D. (1983). *Child Neuropsychology: An introduction to theory, research, and clinical practise*. New York: Guilford Press.

Rourke, B. P., & Orr, R. R. (1977). Prediction of the reading and spelling performances of normal and retarded readers: A four-year follow-up. *Journal of Abnormal Child Psychology, 5,* 9–20.

Rowland, T. W. (1990). *Exercise and children's health*. Champaign, IL: Human Kinetics.

Sallis, J. F., Buono, M. J., Roby, J. J., Micale, F. G., & Nelson, J. A. (1993). Seven-day recall and other physical activity self-reports in children and adolescents. *Medicine and Science in Sports and Exercise, 25,* 99–108.

Sanders M. R., Rebgetz, M., Morrison, M., Bor, W., Gordon, R., Dadds, M., & Shepherd, R. (1989). Cognitive-behavioral treatment of recurrent nonspecific abdominal pain in children: An analysis of generalization, maintenance, and side effects. *Journal of Consulting and Clinical Psychology, 57,* 294–300.

Satz, P., Taylor, H. G., Friel, J., & Fletcher, J. M. (1978). Some developmental and predictive precursors of reading disabilities: A six-year follow-up. In A. L. Benton & D. Pearl (Eds.), *Dyslexia: An appraisal of current knowledge*. New York: Oxford University Press.

Shalet, S. M. (1989). Endocrine consequences of treatment of malignant disease. *Archives of Disease in Childhood, 64,* 1635–1641.

Sheley, J., McGee, Z., & Wright, J. (1992). Gun-related violence in and around inner city schools. *American Journal of Diseases of Childhood, 146,* 677–682.

Simmons, R. G., & Blyth, D. A. (1987). *Moving into adolescence: The impact of pubertal change and school context*. New York: Aldine de Gruyter.

Simopoulos, A. P. (1987). Diet and health: scientific concepts and principles. *American Journal of Clinical Nutrition, 45,* 1015–1410.

Smith, C. (1991). Smoking among young people: Some recent developments in Wales. *Health Education Journal, 50,* 8–11.

Smyth, J. F., Mossman, J., Hall, R., Hepburn, S., Pinkerton. R., Richards, M., Thatcher, N., & Box, J. (1994). Conducting clinical research in the new NHS: the model of cancer. UK Coordinating Committee on Cancer Research. *British Medical Journal, 309,* 457–461.

Snow, J. H., & Hooper, S. R. (1994). *Pediatric traumatic brain injury*. CA: Sage.

Sports Council (1990). *People in sport: Fact sheet*. London: Author.

Stabler, B. (1988). Pediatric consultation-liaison. In D. Routh (Ed.), *Handbook of pediatric psychology*. New York: Guilford Press.

Stein, R. E., Bauman, L. J., Westbrook, L. E., Coupey, S. M., & Ireys, H. T. (1993). Framework for identifying children who have chronic conditions: The case for a new definition. *Journal of Pediatrics, 122,* 342–347.

Stevenson, R. C., McCabe, C. J., Pharoah, P. O. D., & Cooke, R. W. I. (1996). Cost of care for a geographically determined population of low birthweight infants to age 8–9 years. I. Children without disability. *Archives of Disease in Childhood, 74,* 114–117.

Swain, R. (1991). Childhood risk factors and adolescent drug and alcohol abuse. *Educational Psychology Review, 3,* 363–398.

Taylor, H. G., & Fletcher, J. M. (1995). Editorial: Progress in pediatric neuropsychology. *Journal of Pediatric Psychology, 20,* 695–701.

Teta, M. J., Del Po, M. C., Kasl, S. V., Meigs, J. W., Myers, M. H., & Mulvihill, J. J. (1986) Psychosocial consequences of childhood and adolescent cancer survival. *Journal of Chronic Disease, 39,* 751–759.

Thompson, R. J., Jr., Gustafson, K. E., Hamlett, K. W., & Spock, A. (1992). Psychological adjustment of children with cystic fibrosis: The role of child cognitive processes and maternal adjustment. *Journal of Pediatric Psychology, 17,* 741–755.

Thornes, R. (1983). Parental access and family facilities in children's wards in England. *British Medical Journal, 287,* 190–192.

Varni, J. W., & Wallander, J. L. (1988). Pediatric chronic disabilities: Hemophilia and spina bifida as examples. In D. K. Routh (Ed.), *Handbook of pediatric psychology* (pp. 190–221). New York: Guildford Press.

Varni, J. W., & Setoguchi, Y. (1991). Correlates of perceived physical appearance in children with congenital/acquired limb deficiencies. *Journal of Developmental Behavioral Pediatrics, 12,* 171–176.

Vingilis, E., & Adlaf, E. (1990). The structure of problem behavior among Ontario high school students; a confirmatory factor analysis. *Special issue: Traffic injury prevention. Health Education Research, 5,* 151–160.

Walco, G. A., & Dampier, C. D. (1990). Pain in children and adolescents with sickle cell disease: A descriptive study. *Journal of Pediatric Psychology, 15,* 643–658.

Walker L. S., & Greene, J. W. (1989). Children with recurrent abdominal pain and their parents: More somatic complaints, anxiety and depression than other patient families? *Journal of Pedatric Psychology, 14,* 231–243.

Wallander, J. L. (1996). Editorial: International activities in pediatric psychology. *Journal of Pediatric Psychology, 21,* 299–305.

Wallander, J. L., & Thompson, R. J. (1995). Psychosocial adjustment of children with chronic physical conditions. In M. C. Roberts (Ed.), *Handbook of pediatric psychology* (Vol. 2, pp. 124–141). New York: Guilford Press.

West, A., & Sammons, P. (1994). The measurement of children's attitudes towards school. The use of the "Smiley" scale. In M. Christie and D. French (Eds.), *Assessment of quality of life in childhood asthma* (pp. 157–180). Switzerland: Harwood Academic Press.

Weston, R. (1996). Have fun in the sun; protect yourself from skin damage. In S. Clift and S. J. Page (Eds.), *Health and the international tourist*. London: Routledge.

Weithorn, L. A., & Campbell, S. B. (1982). The competency of children and adolescents to make informed treatment decisions. *Child Development, 53,* 1589–1598.

Wichstrom, L. (1994). Predictors of Norwegian adolescents' sunbathing and use of sunscreen. *Health Psychology, 13,* 412–420.

Woodruffe, C., Glickman, M., Barker, M., & Power, C. (1993) *Children, teenagers and health: The key data*. Buckingham, UK: Open University Press.

Wright, S. J. (1990). Family attributions and adjustment to dialysis in adolescent end-stage renal failure. *Psychology and Health, 1,* 77–88.

Wynne, J. (1991). Accidents. In D. Harvey & I. Kovar (Eds.), *Child health: A textbook for the D.C.H.* (2nd ed.). Edinburgh, UK: Churchill Livingstone.

Yee, B. W. (1995). Risk-taking and abusive behaviors among ethnic minorities. *Health Psychology, 14,* 622–631.

Zucher, R. A., & Gomberg, E.S. (1986). Etiology of alcoholism reconsidered: The case for a biopsychosocial process. *American Psychologist, 41,* 783–793.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.28
Tobacco Smoking

MARTIN J. JARVIS
University College London Medical School, UK
and
GAY SUTHERLAND
Institute of Psychiatry, London, UK

8.28.1 INTRODUCTION 646

8.28.2 PATTERNS OF SMOKING PREVALENCE WORLDWIDE 646

8.28.3 SMOKING AS A CAUSE OF DEATH 646
- *8.28.3.1 Diseases Caused by Smoking* 647
- *8.28.3.2 Smoking in Pregnancy* 648
- *8.28.3.3 Health Effects of Passive Smoking* 648
- *8.28.3.4 The Health Benefits of Smoking Cessation* 649

8.28.4 CIGARETTE SMOKING AND NICOTINE DEPENDENCE 650

8.28.5 THE SMOKING CAREER: PATTERNS OF UPTAKE, MAINTENANCE, AND CESSATION OF SMOKING 651
- *8.28.5.1 Recruitment of Young People to Cigarette Smoking* 651
- *8.28.5.2 Adult Smoking: Disadvantage and Dependence* 652
- *8.28.5.3 Factors Associated with Smoking Cessation* 652
- *8.28.5.4 Rates of Spontaneous Cessation in Smokers* 654

8.28.6 PSYCHOLOGICAL SYMPTOMS 654

8.28.7 ASSESSMENT 656
- *8.28.7.1 Assessing Motivation and Intention to Quit* 657
- *8.28.7.2 Assessing Nicotine Dependence* 658
- *8.28.7.3 Questionnaire Measures of Dependence* 658
- *8.28.7.4 Biochemical Measures of Dependence* 660

8.28.8 TREATMENT APPROACHES 661
- *8.28.8.1 Behavioral and Public Health Approaches* 661
- *8.28.8.2 Pharmacological Approaches to Smoking Cessation* 663
  - *8.28.8.2.1 Nicotine replacement therapy* 663
  - *8.28.8.2.2 Non-nicotine pharmacotherapy* 666

8.28.9 APPROACHES TO REDUCE SMOKING-RELATED DISABILITY 667

8.28.10 PROSPECTS FOR THE FUTURE 668

8.28.11 REFERENCES 668

8.28.1 INTRODUCTION

Tobacco smoking presents numerous paradoxes. Despite being possibly the most prevalent, and certainly the most lethal, form of drug dependence in the world, for years cigarette smoking was regarded as no more than a social habit and there was little awareness of its psychoactive component, nicotine. Nicotine is in many ways an invisible drug, which lacks obvious effects on users' cognitive performance, mood, or social functioning. Yet it is now regarded by many as the purest pharmacological dependence, the very paradigm of drug addiction. Nicotine is a stimulant drug which users say calms them down and sedates. Smoking is often viewed as a form of self-medication to cope with stress, but there is little evidence that nicotine possesses any anxiolytic or anti-depressant properties. Despite nicotine's lack of euphoriant or other obvious mood altering effects, smoking is remarkably recalcitrant to change. Most smokers want to give up smoking, and make numerous attempts to do so. But even those with severe smoking-related disease often fail to quit, and even in developed countries with a strong antismoking ethos, fewer than half of all those who have ever smoked regularly succeed in stopping before the age of 65. It is this feature above all which has made smoking a focus of interest for psychologists. It has long been an area of active research into methods of achieving behavior change, as well as for understanding the complex interplay between pharmacological, psychological, and social determinants of behavior. After years when progress was very limited, there have recently been genuine advances in developing effective, and cost-effective, treatment methods for smoking cessation. But failure to succeed in quit attempts, or relapse following initial success, are still the norm, and young people continue to be recruited to smoking in large numbers. The smoking problem shows no sign of going away.

8.28.2 PATTERNS OF SMOKING PREVALENCE WORLDWIDE

While tobacco use has a long history, with forms such as pipe and cigar smoking, chewing tobacco, and nasal snuff predominating in earlier periods, the manufactured cigarette, and with it the epidemic of smoking worldwide, is essentially a phenomenon of the twentieth century. Annual per capita consumption of cigarettes in adults in developed countries rose steadily from 600 in 1920 to a peak of over 3000 in the mid-1970s, since when there has been a modest decline of some 16% (Collishaw & Lopez, 1996). In developing countries, on the other hand, cigarette consumption increased sharply in the 1970s and 1980s, so that the gap in per capita cigarette consumption between developed and developing countries is narrowing. On current trends, per adult consumption in developing countries will exceed that of developed countries shortly after the turn of the century.

Table 1 shows WHO estimates of smoking prevalence in men and women in various regions in the world in the early 1990s, and also trends in per capita consumption since the late 1970s. Among men, worldwide prevalence, at 47%, approaches one-half of all adults, with particularly high levels being seen in China and in the former socialist countries of Eastern Europe. Globally, women's smoking is at much lower levels than that of men (12%), with relatively high prevalence being seen in developed countries and low rates in Africa, India, China, and other parts of Asia where cultural traditions still frown on female smoking. In terms of the total world market, the decline since 1970 of 1.5% per annum in per capita consumption in the American region and of 0.6% in Western European countries has been more than offset by rapid growth in China (4.8% per annum) and other less developed countries, with the result that overall the size of the market continues to increase.

8.28.3 SMOKING AS A CAUSE OF DEATH

Because the magnitude of the risk to health rises with increased duration of the habit, increases in deaths from smoking show a time lag of about 30–40 years from the onset of regular smoking. As a result, differing patterns of smoking related deaths are currently seen in men and women. In the UK, although male smoking declined from the 1960s onward, smoking deaths have only recently begun to decline; while in women death rates are rapidly increasing and will continue to do so for some years, despite prevalence peaking in the early 1970s (Peto et al., 1996).

Table 2 gives estimates of the total number of deaths caused by smoking in developed countries since 1955. In 1955, smoking was responsible for about 500 000 deaths per year, mostly among men. Since that time smoking-attributable deaths have risen dramatically and by 1995 almost 2 million people (1.5 million men and 500 000 women) in developed countries were dying each year from tobacco. The rate of increase is now slowing somewhat among men, but continues to increase rapidly in women. In the mid 1990s about 25% of all male deaths, and

Table 1 Daily smoking prevalence in men and women aged 15 and over, selected world regions, early 1990s: WHO estimates.

	Men (%)	*Women* (%)	*Annual % change 1970–1972 to 1990–1992 per capita cigarette consumption*
WHO regions			
African region	29	4	1.2
American region	35	22	−1.5
Eastern Mediterranean region	35	4	1.4
European region	46	26	0.0
Southeast Asia region	44	4	1.8
Western Pacific region	60	8	3.0
More developed countries	**42**	**24**	**−0.5**
Established market economies	37	23	−0.6
Formerly socialist economies of Europe	60	28	0.6
Less developed countries	**48**	**7**	**2.5**
China (1984)	61	7	4.8
India (1980s)	40	3	1.5
Other Asia and islands	54	7	2.1
Middle Eastern crescent	41	8	1.2
Sub-Saharan Africa	25	3	1.0
Latin America and the Caribbean	40	21	−0.4
World	**47**	**12**	**0.8**

Source: Collishaw and Lopez (1996).

9% of female deaths, in developed countries were due to smoking. The proportion of deaths in middle age (35–69) caused by smoking is considerably higher, at 36% in men and 13% in women. Estimates of smoking related deaths in the developing world are more uncertain, but it is already clear that the increases in smoking prevalence that have occurred in the developing world since the 1970s will lead inexorably to massive increases in deaths. The current worldwide annual figure for deaths caused by smoking of 3 million (made up of 2 million in developed and 1 million in developing countries) is estimated to increase to 10 million per year by about 2020 or 2030. The great bulk of these deaths will be in the developing world (Peto, Lopez, Boreham, Thun, & Heath, 1994).

8.28.3.1 Diseases Caused by Smoking

The link of smoking with lung cancer was first reported in prewar Germany (Smith, Strobele, & Egger, 1994), but this work was largely forgotten, and the findings of Doll and Hill (1950) and Wynder and Graham (1950) stimulated the setting up of a number of epidemiological studies which established the foundations of our understanding of smoking as a cause of disease. Based on the 20 year follow-up of the British doctors cohort, it was concluded that a young man who persists in smoking will run a 1 in 4 chance of being killed prematurely by tobacco (Royal College of Physicians, 1971). More recent studies, including the 40 year follow-up of the British doctors (Doll, Peto, Wheatley, Gray, & Sutherland, 1994) and the second large study of the American Cancer Society (CPS II) (Thun, Day-Lally, Calle, Flanders, & Heath, 1995) have extended this understanding and forced a re-evaluation of the extent of the risk. It is now apparent that persistent smokers run a 1 in 2 risk of being killed by cigarettes, losing on average eight years of life.

Table 3 summarizes data from the American Cancer Society study of over 1 million men and women aged 35 and over, giving mortality in cigarette smokers and life-long nonsmokers for the main fatal smoking-related diseases. The relative risk and absolute excess risk is given for each disease, together with the attributable proportion—the proportion of all deaths for each specified disease that is due to smoking. Smoking is recognized to cause 80% or more of all lung cancers with a relative risk in men of 22 and women of 12. In addition it is responsible for most cancers of the upper respiratory tract (lip, tongue, mouth, pharynx, and larynx) and for a smaller fraction of cancers of the bladder, pancreas, esophagus, and kidney. Among both men and women, smoking-attributable deaths

Table 2 Estimated deaths caused by smoking in developed countries, 1950–2000.

	Men			*Women*		
Mid-decade year	*No. of deaths per year*	*% of all deaths*	*% of deaths age 35–69*	*No. of deaths per year*	*% of all deaths*	*% of deaths age 35–69*
1955	447 000	10	20	26 000	<1	2
1965	793 000	17	28	70 000	2	4
1975	1 119 000	21	31	165 000	3	7
1985	1 369 000	24	35	317 000	6	11
1995	1 442 000	25	36	476 000	9	13
Total all deaths caused by smoking 1950–2000	52 million	20	30	10.5 million	4	7

Source: Peto et al. (1994).

from cardiovascular disease (ischemic heart disease, aortic aneurysm, and stroke) outnumber those from all other causes, including lung cancer. Over 70% of chronic obstructive lung disease is attributable to smoking, with a relative risk in both male and female smokers of about 10.

Findings from the 40 year follow-up of the British doctors' study are in general very similar, but additionally indicate a causal effect of smoking on stomach cancer and leukemia. Associations between smoking and cirrhosis of the liver, suicide, poisoning, and cancer of the liver have been regarded as being due to confounding (Doll et al., 1994). Similarly, there is uncertainty whether the association of smoking with cervical cancer is causal or attributable to confounding (Phillips & Smith, 1994).

As well as being the single largest cause of preventable premature death, cigarette smoking is a cause of a number of disabling but generally nonfatal conditions (Wald & Hackshaw, 1996). These include peripheral vascular disease, cataracts, Crohn's disease, gastric and duodenal ulcers, hip fracture in the elderly, and periodontitis, the major cause of teeth loss in adults.

8.28.3.2 Smoking in Pregnancy

Smoking is an important hazard in pregnancy, leading to an increased risk of spontaneous abortion and doubling the risk of ectopic pregnancy (Poswillo & Alberman, 1992). Babies of smoking mothers weigh on average 150–250 grams less at birth than babies of nonsmoking mothers. This association has been shown to be causal, since randomized trials of smoking cessation in pregnancy have shown that birth weight can be increased (Sexton & Habel, 1984). Smoking by mothers has consistently been found to be associated with sudden infant death syndrome, although it is uncertain whether prenatal exposure, or passive exposure to mother's and father's smoking postnatally, is more important (Blair et al., 1996; Haglund & Cnattingius, 1990; Klonoff-Cohen et al., 1995).

8.28.3.3 Health Effects of Passive Smoking

The adverse effects of cigarette smoking on health are not limited to active smoking. Since the early 1980s there has been an explosion of research into the effects on nonsmokers of breathing in other people's smoke, and this work has had a decisive impact on the debate about the social acceptability of smoking. Studies of the uptake of smoke constituents by nonsmokers have shown dose–response relationships with the extent of exposure to passive smoking (Jarvis, 1989). Cotinine concentrations in exposed individuals have been found to average about 0.7–1.0% of those in heavy smokers (Jarvis, Tunstall-Pedoe, Feyerabend, Vesey, & Saloojee, 1984; Jarvis et al., 1985). A number of reports by independent scientists have reviewed the evidence linking passive smoking with lung cancer in nonsmoking spouses exposed to their partners' smoke. All have concluded that passive smoking causes lung cancer in otherwise healthy nonsmokers, with an estimated number of deaths each year of 3000 in the USA (US Environmental Protection Agency [USEPA], 1992) and about 300 in the UK (Froggatt, 1988). The USEPA has classified environmental tobacco smoke as a known human carcinogen. The association of heart disease with passive smoking is of similar magnitude to that observed for lung cancer, with exposed spouses having about a 30% increase in risk. On a causal interpretation of this association, passive smoking may kill far more people through heart disease than through lung cancer—up to 60 000 deaths each year in

Table 3 Fatal diseases associated with smoking: American Cancer Society (CPS II) men and women aged 35+.

		Standardized mortality per 100 000 per year				
		Life-long nonsmoker	*Current cigarette smoker*	*Relative risk*	*Absolute excess risk per 100 000 per year*	*Attributable (%)*
Cancer						
Lung	M	24	537	22.4	513	87
	F	18	213	11.9	195	77
Upper respiratory	M	1	27	24.5	26	89
	F	2	10	5.6	8	58
Bladder	M	18	53	2.9	35	36
	F	8	21	2.6	13	32
Pancreas	M	18	38	2.1	20	25
	F	16	37	2.3	21	29
Esophagus	M	9	68	7.6	59	66
	F	4	41	10.3	37	74
Kidney	M	8	23	3.0	15	37
	F	6	8	1.4	2	11
Ischemic heart	M	500	970	1.9	470	22
	F	386	688	1.8	302	19
Aortic aneurysm	M	24	98	4.1	74	48
	F	11	52	4.6	41	52
Stroke	M	147	328	2.2	181	27
	F	236	434	1.8	198	20
Chronic obstructive	M	39	378	9.7	339	72
lung disease	F	21	216	10.5	195	74
All diseases	M	788	2520	3.2	1732	40
	F	708	1720	2.4	1012	30

the USA (Wells, 1994; Steenland, 1992). However, because of the difficulty in making appropriate allowance for confounders such as diet, uncertainty remains about how much of the observed association is truly causal.

Although the debate about the effects of passive smoking has largely centered on lung cancer in adults, the main public health burden is borne by children with smoking parents. Numerous studies have shown that babies from smoking households are at increased risk for admission to hospital with bronchitis and pneumonia (USEPA, 1992). In later childhood there is strong evidence of impairments in lung function (Strachan, Jarvis, & Feyerabend, 1990; Cook et al., 1993) and of increased risk of middle ear effusion when parents smoke (Strachan, Jarvis, & Feyerabend, 1989).

8.28.3.4 The Health Benefits of Smoking Cessation

Stopping smoking benefits health at any age, the more so the younger the smoker stops (US Department of Health and Human Services, 1989). In the British doctors' study, those who gave up by their mid-30s had a life expectancy indistinguishable from never smokers, and even those who gave up in their late 60s lived significantly longer than continuing smokers (Doll et al., 1994). Benefits have been observed for a variety of conditions. The Lung Health Study, a randomized trial of smoking cessation in people at increased risk of chronic respiratory disease, showed that on stopping smoking the rate of loss of ventilatory capacity reverts from the accelerated rate characteristic of smokers to the slower age-related decline of nonsmokers (Anthonisen et al., 1994; Jarvis, 1995). With lung cancer a somewhat different picture is seen. The extra risk incurred from a smoking career of a given duration is never lost, but remains at the level determined by the duration of smoking at the time of quitting (Halpern, Gillespie, & Warner, 1993). Since that risk would have escalated exponentially with continued smoking, a marked reduction in risk is observed by comparison with persisting smoking. For heart disease the risk in ex-smokers declines toward that of never-smokers, although there is some uncertainty as to how rapidly the excess risk is lost, whether rapidly (Dobson, Alexander, Heller, & Lloyd, 1991; Rosenberg, Kaufman, Helmrish, & Shapiro, 1985) or over many years (Cook, Shaper, Pocock, & Kussick, 1986).

8.28.4 CIGARETTE SMOKING AND NICOTINE DEPENDENCE

Drug taking aspects of cigarette smoking were largely ignored until the 1970s, but since then an accumulation of research findings from a variety of disciplines and from both human and animal work has led to a consensus that cigarette smoking is essentially a form of addiction to nicotine. This was enshrined in the landmark 1988 report of the US Surgeon General which concluded that the processes underlying addiction to nicotine are similar to those of other addictive drugs such as alcohol, heroin, and cocaine (US Department of Health and Human Services, 1988). Since then clear evidence has emerged that this view is shared, privately if not publicly, by the tobacco industry, which actually reached it considerably earlier (Slade, Bero, Hanauer, Barnes, & Glantz, 1995). The US Food and Drug Administration has based its assertion of jurisdiction over cigarettes and other tobacco products on the evidence of nicotine's addictive properties and the tobacco industry's exploitation of them (Kessler et al., 1996, 1997). Understanding of nicotine's role is fundamental to an appreciation of smoking uptake, maintenance, and cessation. It does not follow, of course, that cigarette smoking is explicable solely in terms of pharmacological factors. Nicotine effects provide a rich substrate for conditioning and social learning mechanisms, and for broad social, economic, and societal influences. In this respect it is no different from other drug dependencies.

The evidence relevant to nicotine's properties as an addictive drug comes from a number of areas (Henningfield & Keenan, 1993; Stolerman & Jarvis, 1995). Some of the main lines are briefly summarized here.

(i) *Brain neurochemistry*. Both animal and human studies have shown that chronic nicotine administration leads to an increase in the expression of specific nicotine receptors on neurons. Even prenatal exposure can produce such nicotine receptor up-regulation (Slotkin, Orband-Miller, & Queen, 1987). Postmortem studies have revealed raised concentrations of nicotine receptors in the brains of human smokers (Benwell, Balfour, & Anderson, 1988). Investigations of dopaminergic systems thought to be important in reward mechanisms have shown effects of nicotine similar to those of cocaine (Pontieri, Tanda, Orzi, & Dichiara, 1996; Pich et al., 1997).

(ii) *Animal self-administration*. Under appropriate schedules of reinforcement nicotine functions as a robust primary reinforcer (Corrigall & Coen, 1989; Spealman & Goldberg, 1982) in a variety of animal species. Early difficulties in demonstrating nicotine's reinforcing effect seem to have stemmed from the narrow range of rewarding blood concentrations and the ease of provoking aversive overdose.

(iii) *Pharmacokinetics*. Nicotine is readily absorbed through the skin, and through the lining of the mouth and nose, the rate of absorption being enhanced in an alkaline environment and reduced in an acidic environment. Because of the large surface area of the lungs the mildly acidic smoke of cigarettes is absorbed almost immediately and completely on inhalation, giving rise to high concentration arterial nicotine boli which reach the brain in less than 10 seconds (Henningfield, London, & Benowitz, 1990; Russell, 1980). Nicotine has a distributional half-life of about 15 minutes and a terminal half-life in blood of about two hours (Benowitz, 1988). This means that blood levels decline overnight to nonsmoking levels, and regular cigarettes are required over the course of the day to maintain elevated blood nicotine concentrations.

(iv) *Regulation of blood nicotine intake from different tobacco products and from cigarettes with differing deliveries (titration)*. Levels of blood nicotine maintained by cigarette smokers and by dependent users of either oral (Holm, Jarvis, Russell, & Feyerabend, 1992) or nasal (Russell, Jarvis, Devitt, & Feyerabend, 1981) snuff are remarkably similar, averaging in each case about 35 ng/ml. Given the very different sensory characteristics of burnt and noncombustible tobacco, the different routes of absorption, and the absence from snuff of components such as tar, this similarity points strongly to nicotine as the factor controlling the behavior.

Manufactured cigarettes contain about 10–14 mg of nicotine, an amount that does not vary greatly between different brands. But by techniques such as filtration and ventilation, yields of nicotine when smoked by a machine in a standardized way range from as little as 0.1 mg to over 1 mg, with concomitant tar yields being closely correlated and ranging from 1 to 15 mg or more. The relevance of these machine smoked yields to human smoking is open to serious question (US Department of Health and Human Services, 1996), since smokers extract a similar amount of nicotine (about 1 mg on average) from cigarettes of widely differing yields (Benowitz et al., 1983). Naturalistic studies of smokers smoking their own self-selected brands (Russell, Jarvis, Iyer, & Feyerabend, 1980; Gori & Lynch, 1985) and of smokers experimentally switched to lower yielding brands (Frost et al., 1995; Withey et al., 1992) have both shown that the slope relating machine-smoked yield and nicotine intake is either flat or very shallow, so that smokers can and do adjust the way they smoke so as to maintain similar nicotine intakes from cigarettes with widely differing deliveries. Similar upregulation of nicotine intake has been shown when smokers temporarily reduce the number of cigarettes smoked (Ho-Yen, Spence, Moody, & Walker, 1982). This pervasive phenomenon which is characteristic of human smoking is termed nicotine compensation or titration (Russell, 1980). A boundary model of nicotine regulation has been proposed, whereby chronic users seek to avoid the adverse effects of either too little (withdrawal) or too much (overdose) nicotine (Kozlowski, 1984).

(v) *Acquisition of nicotine inhalation in novice smokers*. It is well established that children take up smoking for mainly psychosocial reasons, but studies have shown that pharmacological motives take on importance very early in the smoking career (Lynch & Bonnie, 1994; McNeill, 1991). Already by the time they are smoking on a daily basis, children take in as much nicotine from each cigarette as dependent adult smokers (McNeill, Jarvis, Stapleton, West, & Bryant, 1989), and they report similar craving (Goddard, 1990) and withdrawal symptoms on trying to stop (McNeill, West, Jarvis, Jackson, & Bryant, 1986). The FDA has aptly described nicotine addiction as a "pediatric disease."

(vi) *Compulsive use*. Although long-term occasional smokers (so called "chippers") are found (Evans et al., 1992; Shiffman, 1989), they are rare, and the typical smoker smokes much the same number of cigarettes each day, averaging about 17 in the UK. Even those with severe smoking-related disease persist in smoking, whether it be heart disease (Bigelow, Rand, Gross, Burling, & Gottlieb, 1986), laryngectomy (Himbury & West, 1985), or lung cancer (Davison & Duffy, 1982). Dependent users of alcohol, heroin, and cocaine rate giving up smoking as at least as difficult as their problem drug (Kozlowski et al., 1989). Compulsive use is the key feature characterizing nicotine dependence.

(vii) *Nicotine withdrawal syndrome*. Cessation of smoking reliably leads to a variety of signs and symptoms, with an onset within 12 hours or less of last tobacco use, and a duration of three weeks or longer. The manifestations are predominantly affective in nature, with subjects reporting irritability, difficulty concentrating, anxiety, restlessness, increased hunger and depressed mood, as well as craving for tobacco. These findings have been observed in numerous experimental studies (Hughes et al., 1984; West, Jarvis, Russell, Carruthers, & Feyerabend, 1984) and have been thoroughly reviewed (Hughes, 1992; Hughes, Gust, Skoog, Keenan, & Fenwick, 1991; Hughes & Hatsukami, 1986). Similar effects are seen after cessation of nicotine chewing gum (West & Russell, 1985) or of smokeless tobacco (Hatsukami, Gust, & Keenan, 1987). That the tobacco withdrawal syndrome is due to loss of nicotine rather than behavioral aspects of use is shown by the consistent finding that it is relieved by nicotine replacement but not placebo (Gross & Stitzer, 1989; Jarvis, Raw, Russell, & Feyerabend, 1982; Russell et al., 1993; Sutherland, Russell, Stapleton, Feyerabend, & Ferno, 1992; West et al., 1984).

(viii) *Efficacy of nicotine replacement as an aid to smoking cessation*. Numerous randomized controlled trials have shown that nicotine replacement, whether by patch, gum, nasal spray, or inhaler, reliably enhances the chances of success in an attempt at cessation, roughly doubling success rates over placebo (Fiore et al., 1996; Silagy, Mant, Fowler, & Lodge, 1994).

8.28.5 THE SMOKING CAREER: PATTERNS OF UPTAKE, MAINTENANCE, AND CESSATION OF SMOKING

8.28.5.1 Recruitment of Young People to Cigarette Smoking

Uptake of smoking typically occurs in adolescence, with few people starting to smoke after the age of 20 (Kessler, 1995; Kessler et al., 1997). In the UK, from a prevalence of about 1% at age 11, there is a rapid increase over the following few years, so that by age 15 about 25% of teenagers smoke, with some evidence that at this age (although not among young adults of 16 and above) girls are more likely to smoke than boys (Diamond & Goddard, 1995). Rates of smoking in young people have shown no decline since the early 1980s in the UK, and there is evidence that they are currently increasing in the USA (Giovino, Henningfield, Tomar, Escobedo, & Slade, 1995; Glynn, Greenwald, Mills, & Manley, 1993).

Initiation of smoking is subject to a number of influences: environmental, behavioral, and personal factors all play a part (Goddard, 1990; Lynch & Bonnie, 1994). Environmental influences include parental smoking (approximately doubling the likelihood of a child starting to smoke), and smoking by siblings and friends. Tobacco advertising and promotions effectively target young people with images of smoking as trendy, sporty, and successful (Altman, Levine, Coeytaux, Slade, & Jaffe, 1996; DiFranza et al., 1991; Pollay et al., 1996; Wills, Pierce, & Evans, 1996). Young people from deprived backgrounds where smoking is the norm are more likely to become smokers, an association which becomes accentuated in adulthood, implying that smoking is a marker for social trajectory (Glendinning, Shucksmith, & Hendry, 1994).

Cigarette smoking is often an early manifestation of problem behavior. It is linked with poor school performance, truancy, low aspirations for future success, and early school leaving or drop out (Lynch & Bonnie, 1994). Cigarette smoking in adolescents is frequently associated with other problem behaviors, including alcohol and other drug use and other risk taking or rebellious behaviors.

Personal characteristics consistently linked with adolescent smoking include low self-esteem, low knowledge of smoking's adverse effects, and anxiety and depression (Breslau, Kilbey, & Andreski, 1993).

School-based interventions to reduce the uptake of smoking by teenagers have been extensively studied. Skills based approaches designed to equip children to resist social influences to smoke have shown some initial success, with a number of randomized trials reporting lower rates of recruitment in intervention than control groups (Bruvold, 1993; Flay, 1985). Unfortunately, longer term follow-up has found that these effects dissipate (Flay et al., 1989; Nutbeam, Macaskill, Smith, Simpson, & Catford, 1993), leading researchers to advocate approaches involving the creation of a wider social environment supportive of nonsmoking (Flay, 1985). Both in the USA and the UK, where rates of teenage smoking are on the increase, there is a recognition that preventing the uptake of smoking by young people is an area where up until now little has been achieved. If progress is to be made, it will require a far better research understanding than there is at present of the factors influencing experimental smoking and the development of nicotine dependence.

8.28.5.2 Adult Smoking: Disadvantage and Dependence

In developed Western countries rates of cigarette smoking have peaked and are now declining, albeit slowly. In the USA and UK about 25–30% of adults smoke, with little difference in prevalence between men and women, except in the oldest age groups containing a cohort of women for whom never-smoking was the norm. The most striking feature of the evolution of smoking since the late 1970s has been the increasing association of cigarettes with markers of disadvantage, whether it be socioeconomic position, or a range of factors indicating stressful living circumstances (Jarvis, 1994b). High rates of smoking are seen in the unemployed (Lee, Crombie, Smith, & Tunstall-Pedoe, 1991), lone parents (Marsh & McKay, 1994), people who are divorced or separated, the homeless (Gill, Meltzer, Hinds, & Pettigrew, 1996), heavy drinkers (Jarvis, 1994b), drug users (Meltzer, Gill, Pettigrew, & Hinds, 1995), and prisoners (Bridgwood & Malbon, 1995). Cigarette smoking is strongly associated with psychiatric illness, whether it be schizophrenia (DeLeon et al., 1995; Hughes et al., 1991), depressive illness (Glassman, 1993), or a variety of other neurotic disorders (Meltzer et al., 1995). The association of cigarettes with lowered levels of psychological well-being is not confined to those with a formal psychiatric diagnosis, but also extends into the general population of smokers (Anda et al., 1990; Schoenborn & Horm, 1993). In the general population there is an increasingly strong association between cigarette smoking prevalence and indicators of deprivation (Pierce, Fiore, Novotny, Hatziandreu, & Davis, 1989a, 1989b). Between 1973 and 1994, rates of smoking among affluent people halved in the UK, but among the poorest groups prevalence remained unchanged at close to 70% (Bennett, Jarvis, Rowlands, Singleton, & Haselden, 1996; Jarvis, 1997).

Cigarette smoking is unusual among drug dependencies in that high levels of dependence are typical of most users, rather than being found only in a minority (Stolerman & Jarvis, 1995). At least 80% of cigarette smokers meet the *Diagnostic and statistical manual of mental disorders* (4th ed.; *DSM-IV*) criteria for dependence (Cottler et al., 1995). Smokers' dependence is manifested in a variety of ways, which all point to a pattern of behavior that has become highly stereotyped. The average smoker smokes every day, at a rate of over one cigarette for every waking hour. Survey data indicate that almost one-third of smokers have never stopped for as long as a week (US Department of Health and Human Services, 1990). There is increasing consensus that time to first cigarette of the day is the strongest behavioral indicator of cigarette dependence. Figure 1 shows the distribution of time to first cigarette in a general population sample of smokers, and also the corresponding levels of nicotine intake as indexed by saliva cotinine concentrations. Some 17% of smokers light up within five minutes of waking, and over 50% within half an hour. Overall levels of nicotine intake are closely correlated.

8.28.5.3 Factors Associated with Smoking Cessation

At any one time, about two-thirds of cigarette smokers say that they would like to give up, but most lack confidence that they would succeed if

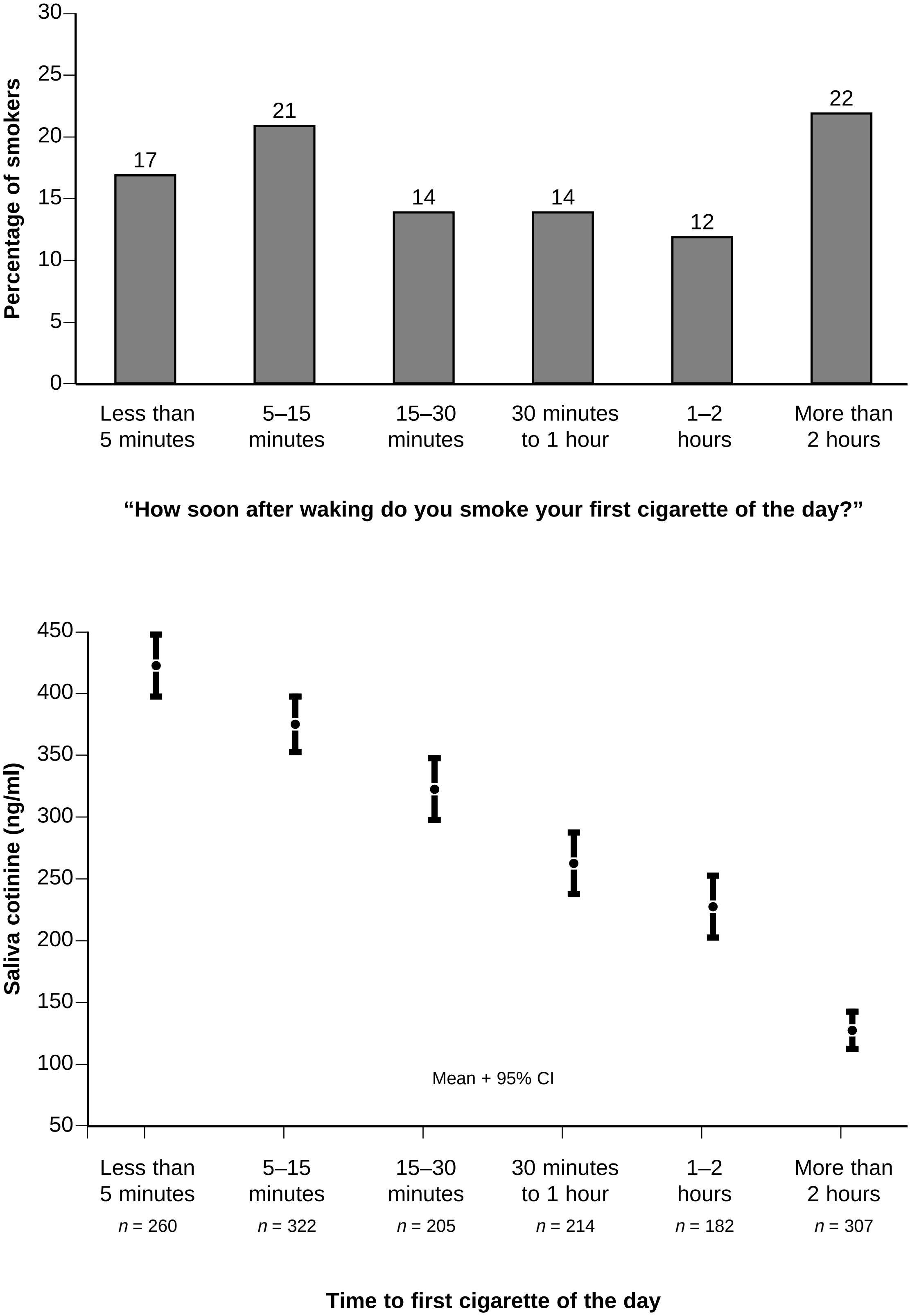

Figure 1 Distribution of time to first cigarette of the day among smokers surveyed in primary care, and corresponding saliva cotinine concentrations. Time to first cigarette is the strongest available predictor of nicotine dependence.

they tried (Bennett et al., 1996). Rates of cessation (defined as the proportion of ex-smokers among ever-regular cigarette smokers) rise steadily with age, from about 18% in young people in their early 20s to 60% among those aged 60 and above (Bennett et al., 1996). Across all age groups combined less than half of ever smokers in the UK and the USA have quit (Bennett et al., 1996; US Department of Health and Human Services, 1990). Overall, rates of cessation show little difference by gender, with young women significantly more likely to quit than men, and middle-aged women significantly less likely (Jarvis, 1994a).

Smoking cessation is influenced by a number of factors, including immediate family circumstances, the broader socioeconomic setting, psychological well-being, and pharmacological dependence (Jarvis, 1997). Influences within the family are shown both by the association of cessation in parents with number of children (Jarvis, 1996) and by the strong concordance in smoking habits among spouses and their partners. Spouse smoking has consistently emerged as one of the most powerful predictors of cessation (Murray, Johnston, Dolce, Wondra Wong, & Ohara, 1995; Nafstad, Botten, & Hagen, 1996; Severson, Andrews, Lichtenstein, Wall, & Zoref, 1995; Wakefield, Gillies, Graham, Madeley, & Symonds, 1993). Poverty has become increasingly associated with low rates of cessation, indicating that those who can least afford to smoke are the least likely to quit (Jarvis, 1997; Marsh & Mackay, 1994). Figure 2 shows rates of cessation by level of deprivation over the past 20 years in the UK. Among both affluent men and women cessation rates have more than doubled, but among the poorest only 10% or less of ever-smokers have quit, and there has been no improvement since the 1970s.

Depressive illness (Glassman et al., 1990b) and stressful life circumstances are associated with low rates of cessation, although interestingly there is evidence that successful cessation leads to lower, rather than higher, levels of perceived stress (Cohen & Lichtenstein, 1990).

8.28.5.4 Rates of Spontaneous Cessation in Smokers

The concept of "spontaneous" or unaided quitting is a difficult one. There are a multiplicity of influences, both those that are explicitly recognized by smokers, and others which, while unrecognized, may nevertheless be present and influence behavior. In a US survey of ex-smokers, respondents acknowledged little assistance in giving up smoking (Fiore et al., 1990). Over 90% said that they gave up without any help at all, either from family, physicians, or more formal services. Only 4% said that they had had any involvement with specialized treatment agencies.

One way to estimate rates of successful unaided quitting in the population is to look at the rate at which prevalence declines year by year, or, more specifically, the increase over time in the cessation rate. By this yardstick, there has been an approximate 1.5% increase in cessation each year in the UK since the 1970s. However, given that no information is available on the number of cessation attempts smokers make in any given year (even if it were possible to reach a satisfactory definition of what constitutes a serious attempt at cessation), this cannot easily be translated into what the chances of successfully quitting are from any one attempt. The figure of 1.5% represents the cumulative success rate from all quit attempts made each year, averaged across smokers. If smokers on average make several attempts to quit each year, it implies a success rate per quit attempt of under 1%; alternatively, if serious quit attempts occur at a rate of less than once per year, the implied success rate per quit attempt may be higher.

Two studies have explicitly recruited smokers planning to make an unaided attempt at quitting and have followed them up to establish success rates. In the first, only one-third abstained successfully for two days, one-quarter for one week, and at six months 3% remained abstinent (Hughes et al., 1992). Similar findings emerged in the second study, with 4.9% not smoking at six months (Cohen et al., 1989). Since some further relapse to smoking would be anticipated after six months, these percentages give an overestimate of the long-term success rate.

There is a reasonable degree of agreement between estimates from population cessation rates and from the studies of self-quitters. Both indicate a low rate of successful long-term quitting from any given quit attempt. It would appear that the chances of successfully quitting on a particular attempt range between about 0.5 and 3%. Such low figures point once again to the difficulty of the task faced by smokers who wish to give up, and suggest that treatments achieving modest success rates of only 5–10% would have a valuable part to play in lowering smoking prevalence, if they could be delivered cost-effectively to the bulk of the smoking population.

8.28.6 PSYCHOLOGICAL SYMPTOMS

Although the act of smoking does not in itself cause any discernible mental or functional impairment, this is not to say that there are no negative psychological sequelae. For many

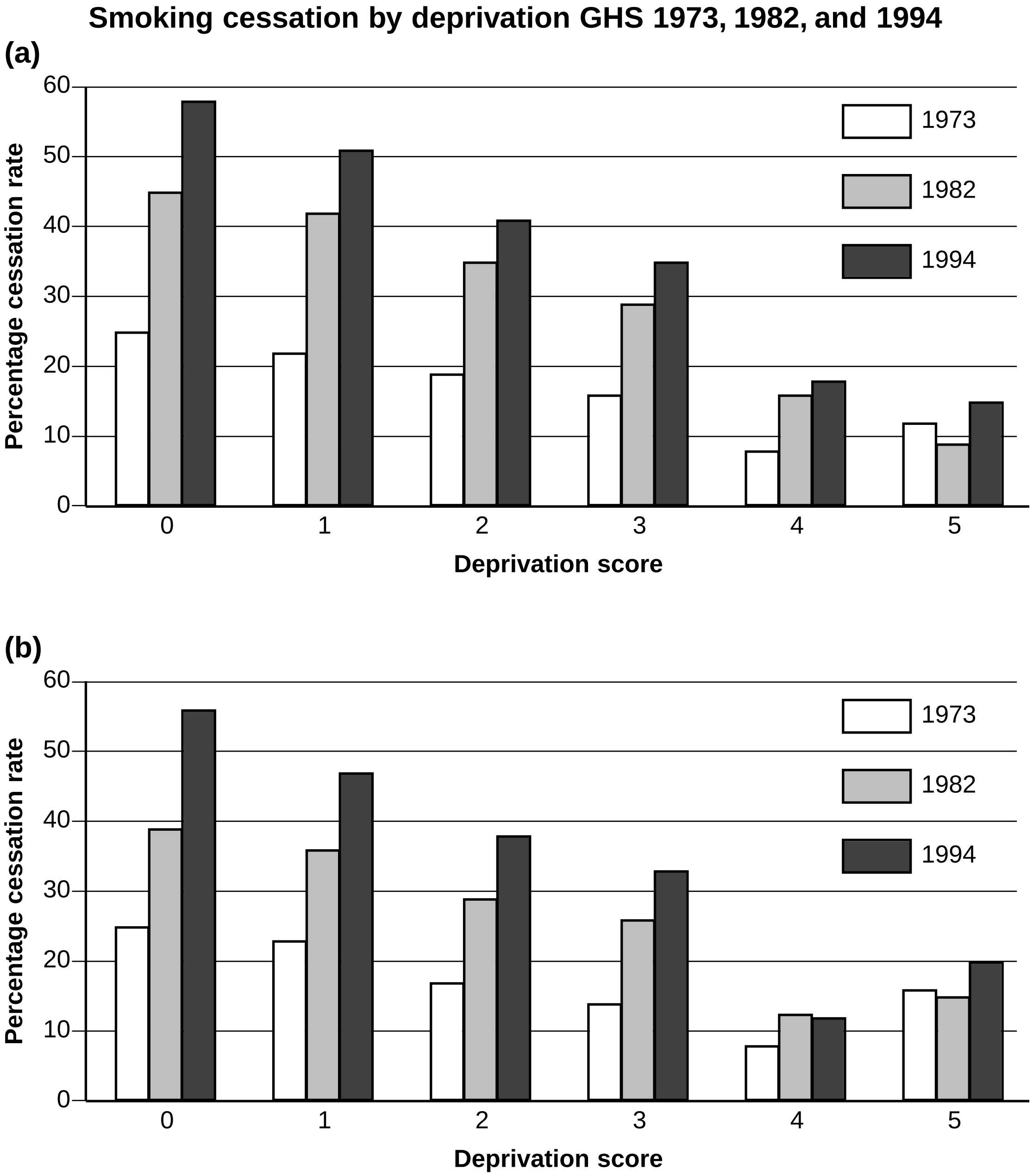

Figure 2 Trends in rates of smoking cessation by deprivation in (a) men and (b) women in the UK, 1973, 1982, and 1994. Deprivation was measured on a five-point scale, where respondents scored 1 for each of the following: manual occupation; rented housing; no access to a car; crowded accommodation; unemployment. From Jarvis (1997).

individuals there will be concern about current or future illness, the effect of passive smoking on children, and the financial burden. There is also likely to be pressure to abstain from non-smoking family and friends, and a dislike of being dependent on a drug. However, the clearest psychological impact of smoking arises when an individual tries to stop. Most, but not all, smokers experience some psychological and physical signs and symptoms when they stop smoking or significantly reduce the amount they smoke. There has been much research both prospectively and retrospectively on the nicotine withdrawal syndrome, most studies using simple self-report questionnaires scored on likert or visual analogue scales. The three most widely reported questionnaires are the Shiffman–Jarvik Smoking Withdrawal Questionnaire (Shiffman & Jarvik, 1976), the Smoker Complaints Scale (Schneider, Jarvik, & Forsythe, 1984), and the Minnesota Nicotine Withdrawal Questionnaire (Hughes & Hatsukami, 1986).

The most commonly reported subjective nicotine withdrawal symptoms are craving for cigarettes, depressed mood, irritability, restlessness, difficulty concentrating, anxiety/tension, sleep disturbance, and increased appetite, particularly for sweet, high calorific foods. Although the number and intensity varies considerably from person to person, withdrawal symptoms are experienced by about 80% of ex-smokers, three-quarters of whom may exhibit a pattern that could be described as severe (Gritz, Carr, & Marcus, 1991). Craving cigarettes, which can be intermittent or more or less continuous, is probably the most troublesome and fundamental symptom and is a major factor in triggering relapse in the first few weeks of an attempt to quit. Self-reports of withdrawal symptoms have been verified by observer ratings (Hughes & Hatsukami, 1986).

Withdrawal symptoms occur abruptly on cessation and changes in mood and ability to concentrate can be detected within two hours of the last cigarette. Symptoms peak in the first few days of cessation and, with the exception of craving, hunger, and weight gain which have a longer time course, gradually subside to baseline levels within a few days or weeks, with an average duration of about 3–4 weeks (Snyder, Davis, & Henningfield, 1989; West, Hajek, & Belcher, 1987). Bouts of craving can persist for months, however, and it is not uncommon for ex-smokers to experience a strong urge to smoke in certain situations or mood states associated with prior smoking even some years after giving-up. Increased appetite and weight gain continue for at least 10 weeks after quitting and ex-smokers typically gain about 6–10 lb in the first year of cessation, though there is great variation from one person to another, with the heaviest smokers tending to gain the most weight.

Other physical and objective changes that accompany tobacco withdrawal include decreases in catecholamine excretion, metabolic rate, heart rate (12–15 bpm) (West & Russell, 1987) blood pressure, and tremor, increases in skin temperature, changes in REM sleep, decreases in peak alpha EEG frequency indicating decreased cortical arousal (Knott & Venables, 1977) and gastrointestinal disturbances (Hughes & Hatsukami, 1986). Performance on tasks requiring vigilance or sustained attention and memory is also impaired (Snyder et al., 1989). Many of these subjective and objective withdrawal effects are the opposite of the effects that occur following nicotine use and can be reversed by smoking or giving nicotine by another route (e.g., nicotine gum, skin patch, or injection) but not by a placebo or nicotine-free cigarettes. Severity of withdrawal symptoms have been found to correlate with baseline smoking nicotine levels and also to predict treatment outcome (West, Hajek, & Belcher, 1989; West & Russell, 1985b).

The *DSM-IV* (American Psychiatric Association, 1994) diagnostic criteria for nicotine withdrawal are shown in Table 4. Of particular interest is the inclusion of depressed mood which had not been listed in *DSM-III-R*. A number of epidemiological studies have recently focused attention on the relationship between depression and smoking. Glassman et al. (1990a) found that smoking was more prevalent in individuals who had had a major depressive disorder at some time in their lives and also that they were less likely to succeed in quitting smoking than those without such a history. Mood changes, such as anxiety and depression have been shown to predict relapse.

8.28.7 ASSESSMENT

Although there are a wide range of assessment procedures which have been, or could be, applied to smoking behavior it is unfortunately the case that while many have face validity they have limited clinical utility. Clinically assessments are most useful when the results can be used to determine, for instance, which of a number of treatments to apply in order to maximize efficacy. This is rarely the case in smoking cessation, however, since much of the information gathered at assessment will at best have prognostic value in predicting a smoker's likelihood of succeeding but will have little to offer in the way of indicating specific procedures to improve the outcome. Empirical research into matching smokers to particular treatments is still in its infancy and has only recently started to highlight some useful leads. The thoroughness and depth of assessment will be determined by the treatment setting. In a specialized smokers clinic offering intensive treatment, for instance, a systematic assessment of individual patient characteristics is feasible and is likely to include measures of the following:

(i) motivation, intention, or desire to stop smoking;

(ii) demographics and smoking history including number and duration of previous attempts to quit and reasons for relapse;

(iii) self-report and objective measures of current tobacco consumption;

(iv) severity of nicotine dependence;

(v) suitability for nicotine replacement or other pharmacological treatment;

(vi) psychosocial factors such as degree of social support, stress, coping skills, and presence of psychiatric comorbidity.

Table 4 *DSM-IV* diagnostic criteria for nicotine withdrawal.

A. Daily use of nicotine for at least several weeks

B. Abrupt cessation of nicotine use, or reduction in the amount of nicotine used, followed within 24 hours by four (or more) of the following signs:

(1) dysphoric or depressed mood;
(2) insomnia;
(3) irritability, frustration, or anger;
(4) anxiety;
(5) difficulty concentrating;
(6) restlessness;
(7) decreased heart rate;
(8) increased appetite or weight gain

C. The symptoms in criterion B cause clinically significant distress or impairment in social, occupational, or other important areas of functioning

D. The symptoms are not due to a general medical condition and are not better accounted for by another mental disorder

Assessments in briefer community interventions and primary care are likely to be considerably less detailed and time consuming and might only include questions about motivation to stop smoking and severity of nicotine dependence since these are the two patient characteristics most likely to influence the course of subsequent treatment and outcome.

8.28.7.1 Assessing Motivation and Intention to Quit

There are very few smokers who are able to quit smoking permanently without first being highly motivated and believing that they can do so. Regardless of the method employed to stop smoking, a major factor will be the psychological resources available to an individual when it comes to changing a strongly conditioned behavior and resisting withdrawal symptoms and craving. A number of studies have shown that without strong motivation to quit the likelihood of succeeding is extremely low in all but the most minimally dependent smokers (Jackson, Stapleton, Russell, & Merriman, 1986). In brief intervention studies, motivation to quit has been assessed by simple questionnaire items such as: "How much do you want to stop smoking altogether?" (not at all, slightly, moderately, quite strongly, very strongly), and "Would you give up smoking altogether, if you could do so easily?" (yes definitely, yes probably, possibly, probably not, definitely not). Only those scoring in the highest categories are likely to be sufficiently motivated to persist in a serious quit attempt. The number of previous attempts to stop smoking is also a useful indicator of motivation.

The concepts of motivation and intention to change behavior have been extended into a theoretical model by Prochaska, DiClemente and colleagues (Prochaska & DiClemente, 1986; Prochaska, DiClemente, & Norcross, 1992). Their "Stages of Change" model is based on recognition of the fact that giving up smoking is a dynamic process, rather than a discrete event, often extending over many years, and characterized by repeated cycles of attempts to stop and relapses back to smoking. The model identifies five main stages through which smokers cycle:

(i) *Precontemplation*. Not seriously thinking about quitting.

(ii) *Contemplation*. Staring to think seriously about stopping within the next 6 months but not within the next 30 days.

(iii) *Preparation*. Intending to quit within the next 30 days.

(iv) *Action*. Stopped smoking within past 6 months.

(v) *Maintenance*. Stopped smoking for over 6 months.

Relapse leads to re-entry into the cycle at an earlier stage. Three simple questions have been identified to assess a smoker's current stage (Prochaska & Goldstein, 1991):

(i) *Are you intending to quit within the next 6 months?* If no, the smoker is in the precontemplation stage.

(ii) *Are you intending to quit smoking in the next month?* If no, but answered yes to question (i), they are in the contemplation stage.

(iii) *Did you try and quit smoking in the past year?* If yes, and answered yes to question (ii) they are in the preparation stage.

In addition to these stages, Prochaska and DiClemente have identified 10 different positive behavior change processes or activities (e.g., counterconditioning, reinforcement management, stimulus control) which facilitate an individual's progression from one stage to the next. They suggest that interventions which encourage the use of the behavior change processes most relevant to a smoker's current stage, will be most effective. So, for instance, precontemplators would be targeted with information about the health risks of smoking and benefits of stopping in the hope of raising their motivation and thus moving them to the contemplator stage, whereas individuals in the action stage might receive training in strategies aimed to prevent relapse. Although there has been some empirical support for the notion of stage-matched interventions, critics have exposed inconsistencies and theoretical problems with the model (Sutton, 1996). Perhaps the most important contribution made by the Stages of Change to date has been to encourage therapists to assist smokers to stop over repeated attempts, rather than using a "one shot" approach. The model has also focused attention on alternative measures of outcome besides total abstinence, that is, in terms of moving smokers along a stage (see also Chapter 8.01, this volume).

8.28.7.2 Assessing Nicotine Dependence

Nicotine dependence is now recognized as a psychoactive substance dependence disorder (*DSM-IV*, *ICD-10*). Under the *DSM-IV* criteria smokers are classified as dependent if they exhibit impaired control and continue to smoke despite awareness that it exacerbates a medical condition, and if they experience withdrawal symptoms when abstaining from smoking, or if they have a history of unsuccessful attempts at quitting. However the practical value of using this all-or-nothing dichotomous classification when assessing an individual smoker is limited since it has been estimated that about 90% of smokers are dependent (Woody, Cottler, & Cacciola, 1993). It is therefore, more useful to view smokers as lying on a continuum of nicotine dependence. When measured in this way dependence has proved to be the strongest and most consistent predictor of the success or failure of an attempt to quit smoking (see Figure 3) such that the likelihood of stopping smoking is several times greater at low, rather than high, levels of dependence (Stapleton et al., 1995).

Although this would seem to suggest considerable potential for tailoring cessation treatment according to dependence level in order to maximize overall cessation rates, little progress has been made in this area. There are not as yet any established behavioral treatments based on the differing dependence level of the smoker. One fundamental problem is that cessation rates are modest even among smokers at the lower end of the dependence range and special attention and effort is only likely to be invested in the more highly dependent smokers when treatment programs are more successful generally. Not surprisingly, given that tobacco dependence is largely a dependence on nicotine, certain forms of nicotine replacement therapy have been shown to be more effective at higher levels of dependence and it is in this area that most progress towards matching treatments is likely to be made in the future (Sutherland & Stapleton, 1994). This does not mean, however, that an assessment of a smoker's level of dependence has no clinical value, since it serves to alert therapists to cases where withdrawal symptoms are likely to be particularly severe and the prognosis poor, indicating that more intensive help may be required.

8.28.7.3 Questionnaire Measures of Dependence

Measures of the severity of nicotine dependence fall into two categories: (i) self-reported behavioral and subjective symptoms, and (ii) objective biochemical markers of nicotine intake. The simplest indicator of degree of dependence is daily cigarette consumption which can be obtained by self-report or by self-monitoring diary records over a period of time. Since smokers tend to report consumption in multiples of 5 or 10, self-monitoring diaries tend to be more accurate than self-report, especially if the cigarette is recorded at the time it is smoked, although the monitoring itself can be a reactive measure and result in a decrease in smoking (Frederiksen, Epstein, & Kosevsky, 1975; McFall, 1978). Consumption is only a rough indicator of dependence, however, as smoke intake per cigarette can vary a great deal from one individual to another. The most widely used paper and pencil test of dependence is the Fagerström Tolerance Questionnaire (FTQ), consisting of an eight-item inventory designed to assess: daily nicotine intake, smoking in the morning (when nicotine levels are low); and perceived difficulty in refraining from smoking (Fagerström, 1978). More recently a shorter, six-item, version of the FTQ, the Fagerström Test for Nicotine Dependence (FTND), has been developed to address some of the psychometric weaknesses of the original and has improved internal consistency and predictive validity (Heatherton, Kozlowski, Frecker, & Fagerström, 1991). Scores range from 0 (low dependence) to 10 (high dependence).

An even shorter version, the Heaviness of Smoking Inventory, is based on the two items which together account for 60% of the FTND score, namely, daily cigarette consumption, and

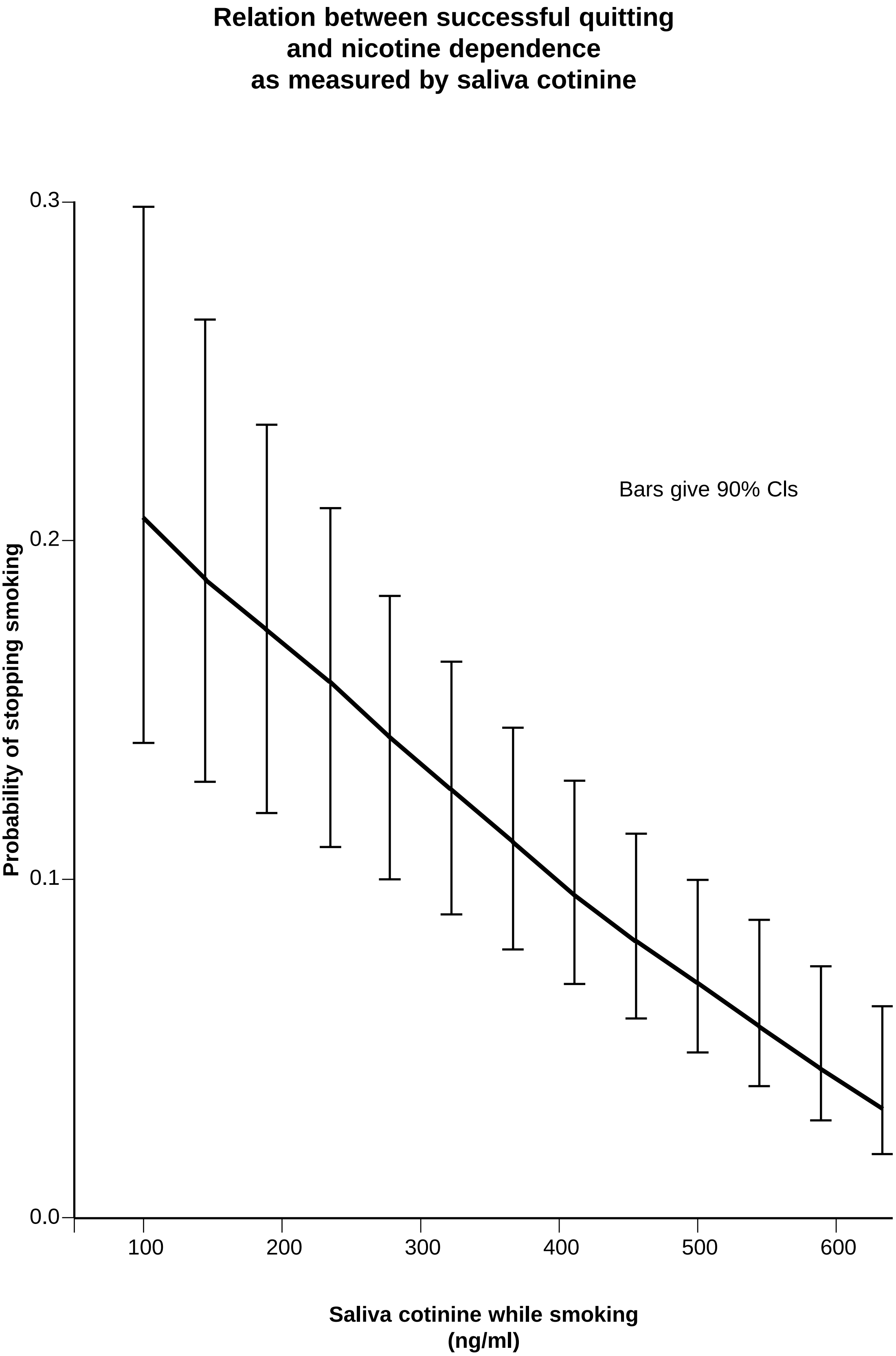

Figure 3 Relationship between nicotine dependence (indexed by saliva cotinine concentrations while smoking) and the probability of succeeding in a cessation attempt. Adapted from Stapleton et al. (1995).

how soon after waking the first cigarette is smoked (Heatherton, Kozlowski, Frecker, Rickert, & Robinson, 1989). Cut-off points have been proposed for the FTND score (e.g., score greater than seven equals "severe" dependence) in order to provide a simple classification. Although theoretically there may be some circumstances where such a classification is desirable the rationale for the precise levels of cut-off are weak and all the evidence points to there being a continuum of dependence in the population with no well defined clusters of smokers at the different levels.

A further questionnaire measure in use is the Overall Dependence Scale (ODS), a subset of nine items taken from the Horn–Russell Smoking Motivation Questionnaire. The ODS (score 0–27) assesses perceived and experienced difficulty in refraining from or quitting smoking, craving for cigarettes when unable to smoke, and degree of habitual or automatic smoking (West & Russell, 1985b).

8.28.7.4 Biochemical Measures of Dependence

In addition to the questionnaire measures, there are now also a number of objective measures which can be used to provide quantitative measures of smoke inhalation and nicotine intake (Jarvis, Tunstall-Pedoe, Feyerabend, Vesey, & Saloojee, 1987). These measures provide a more accurate estimate of the extent of smoking, and hence degree of dependence, than can be gained by knowing simply how many cigarettes an individual smokes. They also have an invaluable role in validating self-reports of abstinence. There are four main biochemical indicators: (i) carbon monoxide in expired air or blood; (ii) cotinine in saliva, blood, or urine; (iii) nicotine in blood; and (iv) thiocyanate (SCN) in saliva, blood, or urine.

In clinical settings the most useful and widely applied method is a measure of the concentration of carbon monoxide (CO) in the expired-air of the smoker which correlates highly with blood (plasma) nicotine and with carboxyhaemoglobin (COHb) (Jarvis, Russell, & Saloojee, 1980; Wald, Idle, Boreham, & Bailey, 1981). Obtaining a breath sample and a reading takes less than a minute, requires minimal training of both therapist and patient, and has the advantage over plasma nicotine of being noninvasive. The main disadvantage of this method is the relatively short half-life of COHb (4–5 hours), less in very physically active individuals, which means it is only possible to detect smoking which has occurred over the past few hours. Ideally, testing should be in the afternoon of a normal smoking day by which time CO levels have reached a stable level. Though rarely a problem in practice, low level exposure to CO from car exhausts and fires, for instance, can produce CO levels in nonsmokers equivalent to those found in very light smokers.

Inhaled nicotine from cigarettes is rapidly distributed throughout the body and about 90% is subsequently metabolized to cotinine (Benowitz, Kuyt, Jacob, Jones, & Osman, 1983). As measures of smoke intake, nicotine, and cotinine have the advantage over CO of being highly specific to tobacco users. However, unlike nicotine, cotinine can be measured reliably in saliva which is considerably easier to collect than a blood sample. Cotinine also has the advantage over nicotine in having a longer half-life of 15–20 hours (compared to nicotine's 1–2 hours) which means sampling time is less critical. Cotinine assays can reliably detect regular smokers who typically have levels between 200 and 400 ng/ml, as well as light smokers. Nonsmokers have levels below 10 ng/ml. The drawback of both nicotine and cotinine measures over CO is that they require a laboratory assay which is expensive, time consuming to collect, and cannot provide immediate feedback of results. Very recently a simple "dipstick" method for assessing total urinary nicotine metabolites has become available (Cope, Nayyar, Holde, Gibbons, & Bunce, 1996). Although it does not produce as accurate a measure as a laboratory assay, it will enable health professionals to assess nicotine dependence quickly and conveniently.

Thiocyanate (SCN) has also been used to determine smoke intake and to corroborate patients' self-reports but has a number of drawbacks and in recent years has increasingly been overtaken by cotinine as the marker of choice. Its use stems from the fact that hydrogen cyanide gas is present in high concentrations in tobacco smoke and is metabolized to SCN by the liver. It has a long half-life of between 10 and 14 days which is useful for validating smoking abstinence but in comparison to cotinine its sensitivity and specificity are low and it is affected by naturally occurring SCN in the diet (e.g., almonds, beer, and broccoli) resulting in some overlap between smokers and nonsmokers.

The questionnaire measures of dependence correlate moderately with the biochemical markers of intake (West & Russell, 1985), and with severity of withdrawal symptoms, but no single measure has been shown to be consistently superior to the others in predicting the success of a quit attempt. The choice of which measure to use will depend on both the purpose for which the measure is taken and the setting. For instance, the paper-and-pencil measures have the advantage of low cost and simplicity whereas salivary cotinine provides the most stable measure of the actual nicotine intake from cigarettes.

8.28.8 TREATMENT APPROACHES

There are four main approaches for reducing the health consequences of tobacco smoking: (i) treatment interventions aimed at encouraging and assisting adult smokers to stop; (ii) preventive interventions to reduce the uptake of smoking by children; (iii) modifications to cigarettes to make smoking less harmful; and (iv) fiscal measures, legislation, and restrictions on smoking in public places designed to reduce active smoking as well as nonsmokers' exposure to environmental tobacco smoke. This chapter will focus primarily on treatment approaches.

8.28.8.1 Behavioral and Public Health Approaches

Since the 1950s a wide variety of interventions, both pharmacological and behavioral, have been developed to help smokers stop, although the efficacy of many of these approaches has not been adequately evaluated in controlled clinical trials. This has been the case particularly with behavioral treatments, where inadequate statistical power through insufficient sample sizes has been a particular problem. Other methodological limitations include: very short-term or unspecified length of follow-up; absence of biochemical validation of abstinence; lack of a control group; and failure to present outcome on a strict "intent to treat" basis whereby nonattenders are classified as smokers. As a result there is little evidence that one behavioral technique is more effective than any other (Lichtenstein & Glasgow, 1992), although there is consensus that multiple component programs tend to be the most successful (Schwartz, 1987). Little attention has been paid to teasing out the specific therapeutic ingredients of the multimodal psychological packages which have been used with smokers. A broad array of behavioral/psychological/cognitive techniques have been tried either individually or in combination, in many settings, from brief self-help programs to inpatient treatment units (Hurt et al., 1992). The interventions have included: rapid smoking and other aversion therapies; smoke dilution; nicotine fading; self-management approaches; for example, contingency contracting and stimulus control, skills training, relapse prevention, cognitive, motivational, and educational interventions, cue exposure, hypnosis and individual or group counselling (Hajek, 1996).

In the 1950s the field was dominated by educational approaches, followed in the 1960s by conditioning procedures such as rapid smoking and satiation. The 1970s saw the introduction of intensive multicomponent cognitive self-management approaches which have promising success rates but are severely limited by the fact that most smokers will not attend formal clinic programs (Lichtenstein, 1992). This factor, together with a recognition of the sheer scale of the problem (approximately 12 million smokers in the UK), prompted a move away from these intensive clinical approaches towards a broader community wide public health perspective involving self-help and advice in a wide variety of settings from health care to the workplace. The current consensus is that a stepped-care treatment model is likely to be the most cost-effective approach in smoking cessation. Similar models have been advocated for other preventive medical interventions. Stepped-care involves initially offering the least intensive and least costly approaches to the greatest number of smokers, and saving the intensive and expensive second-line approaches for those who fail to respond. Though the absolute success rates of the brief interventions are lower they tend to be more cost effective as they have the potential to reach a far greater number of smokers. Typical one year success rates range from about 5% for brief advice in primary care to 10% when NRT is added, and up to 20–30% for specialized smokers clinics offering intensive multisession behavioral approaches combined with NRT.

Recently, after an extensive literature review, the US Agency for Health Care Policy and Research (AHCPR) published their influential guidelines on smoking cessation (Fiore, Bailey, et al., 1996). This targets three main audiences: primary care clinicians; smoking cessation specialists; and health care administrators. The strategies they recommend for those working in primary care, or for whom smoking cessation is just one of their many clinical activities, are shown in Table 5. As part of the assistance offered, the guidelines advise clinicians to provide basic information on the nature and time course of withdrawal and the addictiveness of smoking, for example, that any smoking increases the likelihood of full relapse. Smokers should be helped to develop problem-solving and coping skills by identifying events, mood states and activities which increase the risk of relapse, and identifying and practising skills intended to cope with these high-risk situations.

The AHCPR guidelines for smoking cessation specialists overlap with those for health professionals in primary care but involve collecting more information at assessment on

Table 5 Recommendations for smoking intervention from AHCPR Clinical Practice Guidelines.

Action	*Strategies for implementation*
ASK about smoking status at *every* visit and document	Include in vital signs, use stickers on patients notes, or implement computerized reminder system
ADVISE all smokers to quit, in a *clear, strong*, and *personalized* manner	for example, "Quitting smoking is the most important thing you can do to protect your health" Tie smoking to current illness or impact on children and family
IDENTIFY smokers willing to make a quit attempt at this time	If willing—provide assistance (see below). If patient prefers more intensive help or if it is indicated, refer to specialist. If unwilling—provide motivational intervention
A. ASSIST smoker with a quit plan	Set a quit date Help prepare smoker for quitting: inform family and friends; prepare the environment (discard cigarettes); review previous quit attempts; anticipate challenges
B. NICOTINE REPLACEMENT (NRT)	Encourage use of NRT unless contraindicated
C. GIVE KEY ADVICE	Total abstinence: "Not a single puff" Avoid alcohol Other smokers in household: consider quitting with others or make specific plans for maintaining abstinence among smokers
D. PROVIDE SUPPLEMENTARY MATERIALS	Booklets, etc., culturally, educationally, and age appropriate for the smoker.
ARRANGE follow-up in person or by telephone	First follow-up within 1 week of quit date, second within a month At follow-up, congratulate success, or use lapses for constructive learning, elicit recommitment to total abstinence, identify problems, anticipate challenges, assess use of NRT

Adapted from AHCPR Smoking Cessation Guideline. Source: Fiore, Bailey et al. (1996).

stress and psychiatric comorbidity, for instance, as these variables are known to negatively affect outcome. As there is a strong relationship between the intensity of counseling or contact with the patient and treatment success, the recommendation is for treatment programs, whether group or individual, to span a period of at least two weeks, though preferably more than eight weeks. This factor appears more important than the specific behavioral techniques employed. The optimal number of sessions is 4–7, with each session lasting for 20–30 minutes. Training in relapse prevention, thought by many in the addictions field to be critical to long-term success, has not been shown to be cost-effective in smoking cessation. This has led some to conclude that it is more beneficial to encourage relapsers to wait and have another attempt at a later time rather than offering time-consuming training in relapse prevention (Lichtenstein & Glasgow, 1992).

The AHCPR guidelines for health care administrators and purchasers focus on the importance of developing supportive systems within the clinical setting to facilitate systematic and repeated interventions with smokers. To achieve this, environmental prompts (e.g., computerized systems) to remind staff to ask about smoking are likely to be required, in addition to policy changes to include smoking assessment and cessation in the performance expectations of clinicians. The basic thrust of the recommendations is to greatly increase the number of clinicians who intervene with patients who smoke and to ensure that clinicians accept smoking cessation as a vital part of their role in health care provision.

In addition to brief physician-delivered and intensive multicomponent behavioral smoking interventions, workplace programs have also been investigated. These provide the opportunity to reach a large number of smokers, in a convenient location and may therefore be more acceptable to smokers than attending specialist clinics. A meta-analysis of 20 controlled workplace studies concluded that smokers were 58% more likely to quit in the intervention condition than in the control group (Fisher, Glasgow, & Terborg, 1990).

8.28.8.2 Pharmacological Approaches to Smoking Cessation

8.28.8.2.1 Nicotine replacement therapy

A number of nicotine and non-nicotine pharmacological aids to smoking cessation have been investigated. By far the most well researched is nicotine replacement therapy (NRT) which was designed to be used in conjunction with a behavioral program, however brief. The rationale for nicotine replacement is to temporarily provide smokers with a safer, more slowly delivered and lower dose of nicotine via a different route, to reduce the severity of the withdrawal symptoms and desire to smoke. The process of quitting is effectively broken down into two-stages. Initially the patient gives up the behavioral act of smoking (i.e., it is decoupled from the rewarding effects of nicotine) while starting unlearning the many nondrug elements of the habit and developing strategies to cope without cigarettes. Once the ex-smoker has gained confidence in their ability to stay off cigarettes they are weaned off nicotine completely, usually over a period of 3–4 months. NRT is the most rigorously and thoroughly evaluated treatment for smokers, and though success rates are modest, varying according to the setting and psychological support with which it is combined, it marked a breakthrough in smoking cessation. There are currently four different forms of NRT which have been approved for clinical use on prescription or are available over-the-counter (OTC), although not all have yet been licensed in every country. One very important difference between NRT products and inhaled tobacco smoke is the speed with which nicotine is absorbed. None can mimic the extremely rapid, high but transient nicotine peaks in arterial blood which characterize inhalation, so they do not provide the same positive satisfaction as smoking and therefore have less potential for sustaining dependence.

(i) Nicotine gum

The first formulation to be developed and thus the most extensively researched was nicotine gum, which is available in two strengths (2 or 4 mg per piece). Nicotine from the gum is absorbed slowly through the lining of the mouth (buccal mucosa) and blood levels reach a flat peak after about 30 minutes chewing. Typically, smokers obtain blood nicotine levels around 10–15 ng/mg, about one-third of usual smoking levels, when using the 2 mg gum *ad libitum* (West et al., 1984) and about two-thirds smoking levels with the 4 mg gum (Fagerström, 1988). At these levels, adverse mood effects and difficulty concentrating associated with withdrawal are partially relieved, though craving for cigarettes is not reliably decreased. This is probably because satisfaction from nicotine depends at least in part on a rapid boost in nicotine levels. If absorption rate is crucial nothing is gained by using more gum. This may explain why in routine clinical use smokers tend to chew only seven to eight pieces per day rather than the recommended 12–15 pieces (Russell, 1988). Despite relatively modest effects on withdrawal, the efficacy of the gum in smoking cessation is well established. Nearly 50 randomized controlled trials have now been conducted, and a recent metal-analysis concluded that the gum was more effective than placebo (odds ratio 1.6), regardless of the setting (Silagy et al., 1994).

Although the gum undoubtedly marked a breakthrough in the treatment of smokers, success rates have nevertheless been disappointing and it has a number of other limitations. It is contraindicated in patients with peptic ulcers and poses problems for denture wearers, tastes unpleasant and gives rise to a number of minor side effects and requires some effort and perseverance in order to learn the correct chewing technique.

(ii) Nicotine patches

The patch has the advantage over gum of providing therapeutic levels of nicotine from the first day of treatment with minimal effort and instruction. They are available in three doses (sizes) and most smokers start on the highest dose for 4–8 weeks, followed by a systematic weaning period of 2–8 weeks using progressively smaller size patches. Two different types are available; one worn during waking hours only, with the highest dose patch delivering 15 mg nicotine over 16 hours, and the other worn throughout the day and night delivering 21 mg over 24 hours. Both appear to be equally effective (Fiore, Smith, Jorenby, & Baker, 1994). After applying a patch, venous blood nicotine levels build up slowly over several hours to a plateau of around 10–20 ng/ml, roughly equivalent to 50% of mean plasma nicotine levels produced by smoking. Levels remain reasonably constant for much of the waking day before declining during the evening. Since nicotine absorption is very slow, patch users are unlikely to experience any of the positive subjective effects associated with rapidly administered nicotine which suggests they may be more helpful to individuals who smoke to relieve or avoid withdrawal (negative reinforcement) than to smokers seeking positive reinforcement. Other potentially important differences between the patch and the other forms of NRT include the fact that the patch does not allow patients to self-titrate their dose according to need, and does not provide a behavioral or sensory component to substitute for lighting-up. However, these factors together with the lack of positive effects suggest the dependence potential and abuse liability of the patch will be very low, which is supported by clinical experience. The most common side effect is mild skin irritation but only about 5% of patients have to stop treatment as a result of adverse effects (Hughes & Glaser, 1993).

Numerous controlled trials of the patch in a variety of settings and populations have demonstrated its efficacy with the highest absolute success rates (20–25% at 6 month follow-up) tending to occur when the patch is combined with an intensive behavioral treatment program. The broad consensus from several reviews and meta-analyses is that compared to a placebo, the nicotine patch increases success rates two to threefold irrespective of the intensity of the adjunctive behavioral treatment (Fiore et al., 1994; Silagy et al., 1994). It reduces the severity of withdrawal, and perhaps surprisingly, appears to reduce craving for cigarettes more consistently than the nicotine gum, but its main advantage over the latter is undoubtably the improved patient compliance.

(iii) Nicotine nasal spray

The development of nicotine nasal spray (NNS) stemmed from work showing the rapid rate of nicotine absorption from nasal snuff (Russell et al., 1981). Nicotine taken intranasally is very swiftly absorbed into the systemic circulation and peak venous levels are reached within 5–10 minutes of taking a dose of NNS, which is considerably faster than with the patch or gum (Sutherland et al., 1992), although not as rapid as inhaled cigarette smoke. Unlike smoking however, NNS does not deliver arterial bolus doses of nicotine, although the rate of nicotine absorption is sufficiently rapid to produce the positive stimulant subjective effects (often described as a "buzz") which many smokers experience while smoking their first cigarette of the day. Such effects are usually considered rewarding and satisfying by smokers and are not produced when nicotine is absorbed more slowly. The spray delivers 1 mg of nicotine per dose (0.5 mg to each nostril), and with repeated dosing has the potential to provide blood nicotine levels similar to those obtained while smoking, although results from clinical trials indicate smokers attain levels only about 30–50% of baseline smoking levels when using *ad libitum*.

The efficacy of NNS has been demonstrated in three randomized placebo-controlled trials in which it was combined with behavioral support and produced remarkably similar success rates at one year, averaging 26% vs. 11% for placebo. (Hjalmarson, Franzon, Westin, & Wiklund, 1994; Schneider et al., 1995; Sutherland et al., 1992). As with other forms of NRT, NNS reduces the severity of the affective symptoms of withdrawal compared to placebo, but also has the advantage of substantially relieving craving for cigarettes. At times of acute craving not only can the user obtain pharmacological relief in minutes but there is also a behavioral response (i.e., spraying) with a marked sensory impact. However, perhaps the most significant finding, both from a clinical and theoretical point of view, is that the spray appears to be of greatest benefit to the most highly dependent smokers, in contrast to findings with the patch (see Figure 4) which is equally effective across the full range of dependence.

Studies are currently underway to determine whether these initial but promising findings will be replicated when the spray is used with less intensive adjunctive treatment, such as in general practice. It takes a few days to learn the correct spraying technique and to acclimatize to the initial irritant side effects (to nose/throat), which may pose problems in settings where less time is available to support and encourage new users.

Another issue concerns the greater dependence potential of the spray compared to the slower-acting nicotine delivery systems (Hughes, 1988). In the two smoking cessation studies which allowed the spray to be used for up to a year, 43% and 29% of those who succeeded in stopping smoking used it for this time, representing approximately 12% of all those originally assigned active spray. This figure is somewhat higher than for the gum. However,

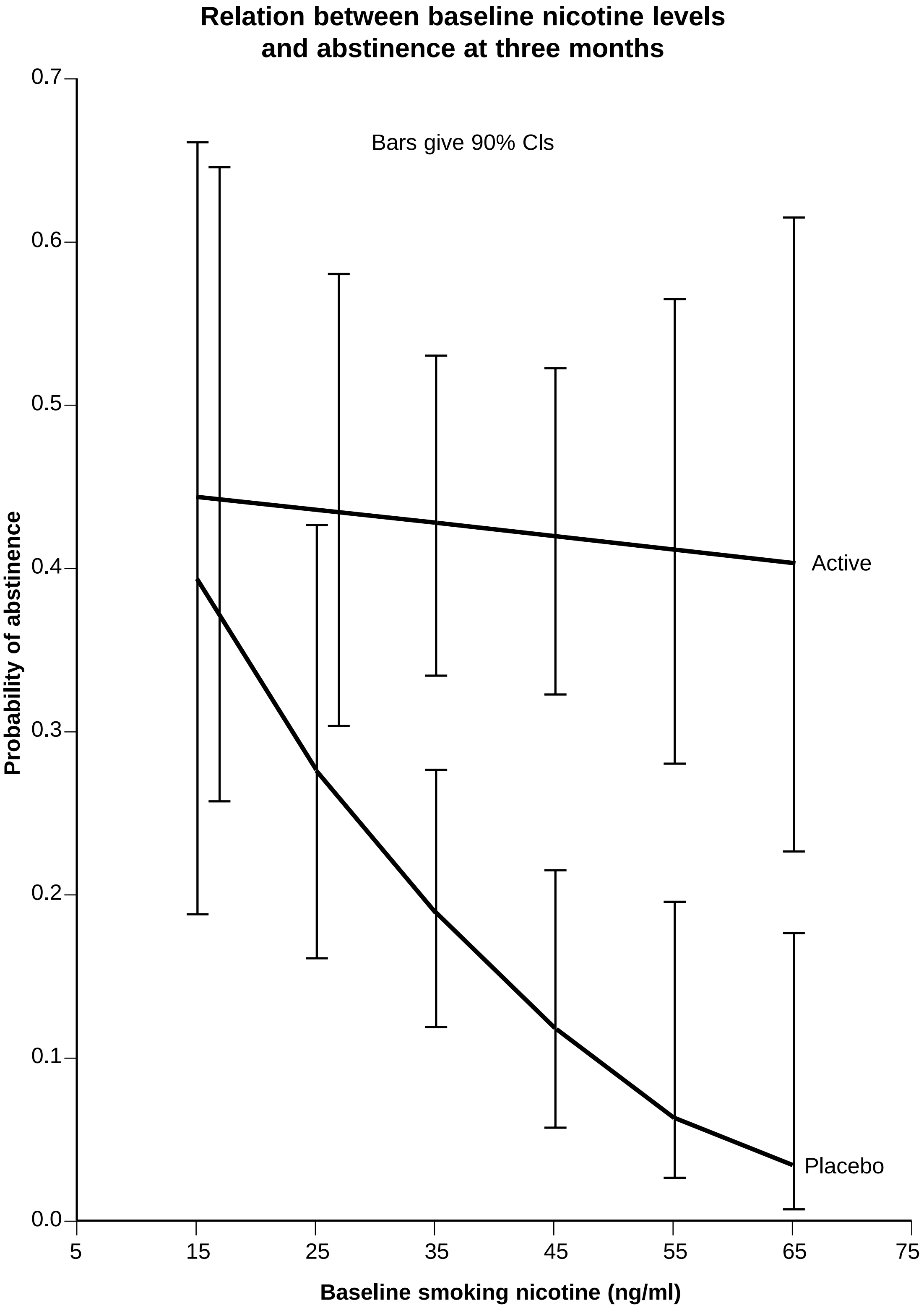

Figure 4 Relationship between nicotine dependence (indexed by plasma nicotine concentration from smoking) and the probability of successfully quitting on active or placebo nasal nicotine spray. Treatment with active spray is especially helpful at high levels of dependence. From Sutherland et al. (1992).

the fact that the spray was provided free of charge is likely to have encouraged longer term and higher usage since cost has been shown to influence these parameters for nicotine gum (Hughes, Wadland, Fenwick, Lewis, & Bickel, 1991). Not surprisingly, the long-term users had higher pretreatment nicotine levels (more dependent) than those who stopped the spray earlier on.

(iv) Nicotine inhaler

The inhaler, the newest form of NRT, is an oral puffing device that consists of a hollow plastic mouthpiece, roughly the size of a cigarette, which contains a nicotine impregnated porous plug. It evolved from an earlier smoke-free product called Favor®, which in preliminary studies was shown to have some therapeutic potential (Hajek, Jarvis, Belcher, Sutherland, & Feyerabend, 1989; Russell, Jarvis, Sutherland, & Feyerabend, 1987). Despite its name, absorption of nicotine vapor from the inhaler is not very rapid and does not give an arterial peak, since most of it is deposited in the mouth and upper airways rather than the lungs. The pharmacokinetic profile resembles the gum rather than cigarettes (Bergstrom, Nordberg, Lunell, Antoni, & Langstrom, 1995). Per puff, the nicotine delivered by the inhaler is about one-tenth of that from a cigarette and is both temperature and effort related, but even with very frequent and intensive puffing, the levels achieved are modest. However, the regular hand–mouth activity involved offers more habit replacement than other forms of NRT and it provides sensory characteristics that resemble smoking, which may have a role in alleviating cigarette cravings (Rose, Behm, & Levin, 1993).

The results of two randomized controlled trials of the inhaler have been published and report one year success rates of 15% active inhaler vs. 5% placebo, and 13% vs. 8%, respectively (Schneider et al., 1996b; Tønnesen, Nørregaard, Mikkelsen, Jørgensen, & Nilsson, 1993).

Research is being undertaken to see whether combining different types of NRT is more effective than using each individually and whether tailoring NRT to achieve 100% replacement (i.e., up to smoking levels) is warranted. Preliminary reports indicate that higher nicotine replacement is associated with better withdrawal relief and higher short- and long-term success rates (Dale et al., 1995; Fagerström, Schneider, & Lunell, 1993; Kornitzer, Boutsen, Dramaix, Thijs, & Gustavsson, 1995). From a theoretical point of view, the best combinations are likely to involve the patch, which provides a steady background level of nicotine, supplemented by one of the other faster-acting delivery systems which enable the patient to self-titrate dose as needed and provide an alternative coping response. Despite the vast literature on NRT many questions remain to be answered, such as the optimal duration of treatment and dosage, how it should be discontinued (abruptly or gradually), and with which psychological treatment it should be combined for maximum efficacy. What does seem clear at the present time is that NRT combined with supportive treatment, however brief, produces higher success rates than NRT or behavioral treatment alone. This might be because the two treatments affect different aspects of nicotine dependence, for example, NRT reduces withdrawal discomfort and behavioral treatment enhances coping skills (Hughes, 1995) or because each treatment increases adherence to the other (Klesges, Ward, & Debon, 1996).

8.28.8.2.2 Non-nicotine pharmacotherapy

A variety of non-nicotine pharmacological treatments or agents have also been examined including lobeline (Schneider, Mione, Raheman, Phillips, & Quiring, 1996a), naltrexone (Sutherland, Stapleton, Russell, Feyerabend, 1995), antidepressants (Berlin et al., 1995; Edwards, Murphy, Downs, Ackerman, & Rosenthal, 1989), mecamylamine (Rose et al., 1994), glucose (West, Hajek, & Burrows, 1990), the antihypertensive clonidine (Covey & Glassman, 1991; Gourlay, Stead, & Benowitz, 1997), buspirone (Cinciripini et al., 1995; Hilleman, Mohiuddin, Core, & Sketch, 1992; Schneider et al., 1996c; West, Hajek, & McNeill, 1991), anorectics (Spring et al., 1995), citric acid inhaler (Behm et al., 1993; Rose & Hickman, 1987), black pepper vapor inhaler (Rose & Behm, 1994), benzodiazepines, beta-blockers, anticholinergics, stimulants, ACTH, silver acetate, and sodium bicarbonate (Hughes, 1993).

Although some show promise, most have not been found to be effective, to have unacceptable side effects, or have not yet been evaluated thoroughly enough with adequate sample sizes to recommend their use at present. The only exception to this concerns the antidepressant, bupropion, which was approved for use in smoking cessation by the US Food and Drug Administration in 1997, though it has not yet been licensed in the UK. Bupropion (Zyban®) is an aminoketone antidepressant which is thought to act primarily via a noradrenergic mechanism but which also has some dopaminergic activity. Anecdotal reports of depressed patients quitting smoking spontaneously while taking bupropion, its generally favorable side effect profile, and its low dependence potential, stimulated interest in the drug for smoking cessation. Although the results of the controlled clinical trials have not yet been published the data are convincing and provided that it is licensed more widely, bupropion could, like NRT, make a significant contribution to future cessation rates.

8.28.9 APPROACHES TO REDUCE SMOKING-RELATED DISABILITY

Although in developed countries studies consistently show the majority of smokers are aware that smoking carries significant health risks, albeit grossly underestimating the real risks to their own health, reducing smoking prevalence has been a slow and hard fought process. Despite their best efforts a number of smokers cannot manage to give up and some others simply do not want to stop, and unlike other drug dependencies smokers do not seem to "mature" out of it. Until more effective interventions have been developed there is a strong case for exploring approaches designed to make smoking less dangerous. One way of achieving this is by modifying the constituents or design of cigarettes to reduce their toxicity. Such strategies are akin to those adopted in the drug dependence field under the umbrella of "harm minimization." This is an area that has highlighted a number of ethical issues in addition to having political implications, and which has divided opinion among researchers in the smoking cessation field.

In 1973 the Independent Scientific Committee on Smoking and Health (ISCSH) was established in the UK to formulate national policies on smoking and health. The main thrust of the program has concerned product modification. Initially the Committee were involved in a radical move examining the potential advantages of substituting tobacco in cigarettes with a less toxic synthetic substance based on modified cellulose. Although 11 new cigarette brands based on substitute tobacco (New Smoking Material and Cytrel) were launched in the UK in 1977 they were not adequately acceptable to smokers and were soon withdrawn from sale. Following this disappointing outcome the ISCSH turned its attention instead to the gradual reduction of tar yields of cigarettes which has been a far more successful strategy. Since 1972 the sales-weighted average tar yield of cigarettes has declined from 20.8 mg to around 11.0 mg per cigarette. This reduction has been achieved by improving cigarette filters, introducing ventilation holes, and increasing the porosity of the paper in order to dilute the smoke, as well as through changes to the tobacco itself. Determination of the yields of cigarettes are made using smoking machines using internationally agreed puff parameters (volume of smoke in puff, frequency of puffing, etc.). Since tar and nicotine yields are highly correlated (about 0.9) reductions in tar have also been accompanied by reductions in nicotine. Unfortunately when smokers switch to cigarettes which have lower nicotine levels, they tend to alter, often unconsciously, the way in which they smoke in an attempt to maintain adequate levels of nicotine. They may for instance take more puffs, hold the smoke in their lungs for longer, take larger puffs, smoke to a shorter butt length, or even hold the cigarette in a way that causes the ventilation holes to be covered. This "compensatory" smoking behavior, or upregulation of smoking, has meant that the percentage reduction in tar intake that would be expected when smokers transfer from a medium or high tar cigarette to a low tar brand is not as great as predicted by the smoking machine figures. Mainly as a result of the lower nicotine levels, low-tar cigarettes tend to be less acceptable to smokers which explains why, despite the health advantages and the selective advertising encouraging smokers to switch, such brands still only account for 30% of the market (Bennett et al., 1996).

Since some of the health advantage of switching to lower-tar cigarettes may be offset by the tendency to compensate by increasing inhalation, and since consumer acceptability is limited, there have been recommendations for the development of low-tar cigarettes with slightly enhanced nicotine yields (up to 1 mg). Although the case for slightly higher nicotine cigarettes is compelling, one must not underestimate the significance of the reductions in the tar intake per cigarette over the last 25 years which have contributed to a decline in mortality from lung cancer (Peto, 1986) and also been beneficial in chronic obstructive lung disease (Darby, Doll, & Stratton, 1989). No matter how intensively modern cigarettes are smoked it is virtually impossible to obtain tar levels as high as those delivered by the old high-tar brands.

The flip-side of this strategy, and one which is far more radical, is the approach being considered in the USA linked to the unprecedented proposed settlement with the tobacco industry. In addition to the financial penalties and other controls being imposed on tobacco companies, the settlement stipulates that the FDA will in future have the power to regulate nicotine as a drug and gives them authority to remove it from cigarettes in 12 years, subject to a number of provisions such as showing that this would not lead to a significant demand for black market cigarettes. The thinking behind this plan is that nicotine yields in cigarettes could be lowered slowly over a number of years until the maximum level a smoker could obtain from cigarettes would be insufficient to create or sustain nicotine dependence (Benowitz & Henningfield, 1994; Kessler et al., 1996, 1997; Kozlowski & Henningfield, 1995; Slade, 1995). In future children might still experiment with cigarettes but could not get a large enough dose of nicotine to become pharmacologically dependent and so would be unlikely to progress to regular daily smoking. Although intuitively appealing and plausible, a number of critical questions remain unanswered. What will happen to existing smokers? It would be reasonable to expect them to smoke more aggressively in an attempt to obtain more nicotine and maintain satisfaction, and in the process become exposed to greater amounts of the more dangerous components (CO and tar) than would otherwise have been the case. What is the critical threshold for creating dependence and what role does tolerance play? By reducing nicotine yields very slowly over several years, might neuronal readaptation and sensitization occur so that current smokers remain dependent albeit on lower nicotine levels?

Another innovative but highly controversial approach to safer smoking has come from tobacco companies themselves. In 1987 the R. J. Reynolds Tobacco Company spent $1 billion developing and market-testing a novel virtually tar free cigarette-like device (brand name Premier) which heated rather than burned tobacco (R. J. Reynolds Tobacco Co., 1988). The smoke particles in which the nicotine was transported were comprised mainly of glycerol and water rather than tar. Glycerol is harmless and metabolized as a source of energy and was added simply to produce "smoke." The yields of all major carcinogens and many other potentially harmful components were far lower than those of conventional cigarettes and the biological activity of its condensate was greatly reduced. The tar yield was about 0.7 mg per cigarette and the nicotine and carbon monoxide yields about 0.3 and 12.0 mg, respectively. It was estimated that the risks of tobacco-related cancers and chronic obstructive lung disease would be substantially reduced, possibly by as much as 90%, for an average smoker switching from a regular brand to Premier (Sutherland, Russell, Stapleton, & Feyerabend, 1993). Premier was never marketed, however, due to lack of acceptability (especially poor taste), and strong protests from public health organizations in the USA who petitioned the FDA to classify and therefore regulate Premier as a drug as opposed to a cigarette. In the last year R. J. Reynolds have launched another new smoke-free cigarette called Eclipse which like its predecessor, is virtually tar free but which is claimed to have a more acceptable taste. It is likely that in the USA this product will receive near unanimous opposition once again on the grounds that it might encourage smokers who would otherwise stop to switch to the new "safe" cigarette instead, and that it may encourage children to take up smoking.

8.28.10 PROSPECTS FOR THE FUTURE

Like all drug dependencies, cigarette smoking is a multifaceted problem, whose understanding requires attention to a range of factors, from drug pharmacokinetics and pharmacodynamics, genetics, individuals and their social milieu, to broader economic and political influences. No single policy response can hope to address the range of issues which determine smoking uptake, maintenance, and cessation. Future progress will depend on developing a better understanding of the processes operating at each stage of the smoking career, and clinical psychology will continue to have an important contribution to make in furthering understanding of individual liability to dependence, the process of recruitment, the role of nicotine in maintaining the habit, and in developing and disseminating effective cessation interventions. If there is one lesson which can be taken from the past, it is that future progress is likely to be equally slow and hard won.

8.28.11 REFERENCES

Altman, D. G., Levine, D. W., Coeytaux, R., Slade, J., & Jaffe, R. (1996). Tobacco promotion and susceptibility to tobacco use among adolescents aged 12 through 17 years in a nationally representative sample. *American Journal of Public Health, 86*(11), 1590–1593.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anda, R. F., Williamson, D. F., Escobedo, L. G., Mast, E. E., Giovino, G. A., & Remington, P. L. (1990). Depression and the dynamics of smoking: A national perspective. *Journal of the American Medical Association 264*(12), 1541–1545.

Anthonisen, N. R., Connett, J. E., Kiley, J. P., Altose, M. D., Bailey, W. C., Buist, A. S., Conway W., Jr., Enright, P. L., Kanner, R. E., Ohara, P., Owens, G. R., Scanlon, P. D., Tashkin, D. P., & Wise, R. A. (1994). Effects of smoking intervention and the use of an inhaled anticholinergic bronchodilator on the rate of decline of FEV 1: The lung health study. *Journal of the American Medical Association 272*(19), 1497–1505.

Behm, F. M., Schur, C., Levin, E. D., Tashkin, D. P., & Rose, J. E. (1993). Clinical evaluation of a citric acid inhaler for smoking cessation. *Drug & Alcohol Dependence, 31*(2), 131–138.

Bennett, N., Jarvis, L., Rowlands, O., Singleton, N., &

Haselden, L. (1996). *Living in Britain: Results from the 1994 General Household Survey*. London: HMSO.

Benowitz, N. L. (1988). Pharmacologic aspects of cigarette smoking and nicotine addiction. *New England Journal of Medicine 319*(20), 1318–1330.

Benowitz, N. L., Hall, S. M., Herning, R. I., Jacob, P., Jones, R. T., & Osman, A. L. (1983). Smokers of low-yield cigarettes do not consume less nicotine. *New England Journal of Medicine 309*(3), 139–142.

Benowitz, N. L., & Henningfield, J. E. (1994). Establishing a nicotine threshold for addiction: The implications for tobacco regulation. *New England Journal of Medicine 331*(2), 123–125.

Benowitz, N. L., Kuyt, F., Jacob, P. I., Jones, R. T., & Osman, A. L. (1983). Cotinine disposition and effects. *Clinical Pharmacology & Therapeutics, 34*(5), 604–611.

Benwell, M. E. M., Balfour, D. J. K., & Anderson, J. M. (1988). Evidence that tobacco smoking increases the density of (−)-(3H)nicotine binding sites in human brain. *Journal of Neurochemistry, 50*(4), 1243–1247.

Bergstrom, M., Nordberg, A., Lunell, E., Antoni, G., & Langstrom, B. (1995). Regional deposition of inhaled c-11 nicotine vapor in the human airway as visualized by positron emission tomography. *Clinical Pharmacology & Therapeutics, 57*(3), 309–317.

Berlin, I., Said, S., Spreux-Varoquaux, O., Launay, J. M., Olivares, R., Millet, V., Lecrubier, Y., & Puech, A. J. (1995). A reversible monoamine-oxidase-a inhibitor (moclobemide) facilitates smoking cessation and abstinence in heavy, dependent smokers. *Clinical Pharmacology & Therapeutics, 58*(4), 444–452.

Bigelow, G. E., Rand, C. S., Gross, J., Burling, T. A., & Gottlieb, S. H. (1986). *Smoking cessation and relapse among cardiac patients*. (NIDA Research Monograph Series 72, pp. 167–171). Rockville, MD: US Department of Health and Human Services.

Blair, P. S., Fleming, P. J., Bensley, D., Smith, I., Bacon, C., Taylor, E., Berry, J., & Tripp, J. (1996). Smoking and the sudden infant death syndrome: results from 1993–5 case-control study for confidential inquiry into stillbirths and deaths in infancy. *British Medical Journal, 313*, 195–198.

Breslau, N., Kilbey, M. M., & Andreski, P. (1993). Nicotine dependence and major depression: New evidence from a prospective investigation. *Archives of General Psychiatry, 50*(1), 31–35.

Bridgwood, A., & Malbon, G. (1995). *Survey of the physical health of prisoners*. London: HMSO.

Bruvold, W. H. (1993). A meta-analysis of adolescent smoking prevention programs. *American Journal of Public Health, 83*(6), 872–880.

Cinciripini, P. M., Lapitsky, L., Seay, S., Wallfisch, A., Meyer, W. J., & Van Vunakis, H. (1995). A placebo-controlled evaluation of the effects of buspirone on smoking cessation: Differences between high-anxiety and low-anxiety smokers. *Journal of Clinical Psychopharmacology, 15*(3), 182–191.

Cohen, S., & Lichtenstein, E. (1990). Perceived stress, quitting smoking, and smoking relapse. *Health Psychology, 9*(4), 466–478.

Cohen, S., Lichtenstein, E., Prochaska, J. O., Rossi, J. S., Gritz, E. R., Carr, C. R., Orleans, C. T., Schoenbach, V. J., Biener, L., Abrams, D., DiClemente, C., Curry, S., Marlatt, G. A., Cummings, K. M., Emont, S. L., Giovino, G., & Ossipktein, D. (1989). Debunking myths about self-quitting. Evidence from 10 prospective studies of persons who attempt to quit smoking by themselves. *American Psychologist, 44*(11), 1355–1365.

Collishaw, N. E., & Lopez, A. D. (1996, May). The tobacco epidemic: a public health emergency. *WHO Tobacco Alert*. Geneva, Switzerland: World Health Organization.

Cook, D. G., Shaper, A. G., Pocock, S. J., & Kussick, S. J. (1986). Giving up smoking and the risk of heart attack: A report from the British Regional Heart Study. *Lancet, 2*, 1376–1379.

Cook, D. G., Whincup, P. H., Papacosta, O., Strachan, D. P., Jarvis, M. J., & Bryant, A. (1993). Relation of passive smoking as assessed by salivary cotinine concentration and questionnaire to spirometric indices in children. *Thorax, 48*(1), 14–20.

Cope, G., Nayyar, P., Holder, R., Gibbons, J., & Bunce, R. (1996). A simple near-patient test for nicotine and its metabolites in urine to assess smoking habit. *Clinica Chimica Acta, 256*(2), 135–149.

Corrigall, W. A., & Coen, K. M. (1989). Nicotine maintains robust self-administration in rats on a limited-access schedule. *Psychopharmacology, 99*(4), 473–478.

Cottler, L. B., Schuckit, M. A., Helzer, J. E., Crowley, T., Woody, G., Nathan, P., & Hughes, J. (1995). The *DSM-IV* field trial for substance use disorders: Major results. *Drug and Alcohol Dependence, 38*(1), 59–69.

Covey, L. S., & Glassman, A. H. (1991). A meta-analysis of double-blind placebo-trials of clonidine for smoking cessation. *British Journal of Addiction, 86*(8), 991–998.

Dale, L. C., Hurt, R. D., Offord, K. P., Lawson, G. M., Croghan, I. T., & Schroeder, D. R. (1995). High-dose nicotine patch therapy: Percentage of replacement and smoking cessation. *Journal of the American Medical Association, 274*(17), 1353–1358.

Darby, S. C., Doll, R., & Stratton, I. M. (1989). Trends in mortality from smoking related diseases in England and Wales. In N. Wald & P. Froggatt (Eds.), *Nicotine, smoking, & the low tar program* (pp. 70–82). Oxford, UK: Oxford University Press.

Davison, G., & Duffy, M. (1982). Smoking habits of long-term survivors of surgery for lung cancer. *Thorax, 37*(5), 331–333.

DeLeon, J., Dadvand, M., Canuso, C., White, A. O., Stanilla, J. K., & Simpson, G. M. (1995). Schizophrenia and smoking: An epidemiological survey in a state hospital. *American Journal of Psychiatry, 152*(3), 453–455.

Diamond, A., & Goddard, E. (1995). *Smoking among secondary schoolchildren in 1994*. London: HMSO.

DiFranza, J. R., Richards J. Jr., Paulman, P. M., WolfGillespie, N., Fletcher, C., Jaffe, R. D., & Murray, D. (1991). RJR Nabisco's cartoon camel promotes Camel cigarettes to children. *Journal of the American Medical Association, 266*(22), 3149–3153.

Dobson, A. J., Alexander, H. M., Heller, R. F., & Lloyd, D. M. (1991). How soon after quitting smoking does risk of heart attack decline? *Journal of Clinical Epidemiology, 44*(11), 1247–1253.

Doll, R., & Hill, A. B. (1950). Smoking and carcinoma of the lung. Preliminary report. *British Medical Journal, ii*, 739–748.

Doll, R., Peto, R., Wheatley, K., Gray, R., & Sutherland, I. (1994). Mortality in relation to smoking: 40 years' observations on male British doctors. *British Medical Journal, 309*(6959), 901–911.

Edwards, N. B., Murphy, J. K., Downs, A. D., Ackerman, B. J., & Rosenthal, T. L. (1989). Doxepin as an adjunct to smoking cessation: A double-blind pilot study. *American Journal of Psychiatry, 146*(3), 373–376.

Evans, N. J., Gilpin, E., Pierce, J. P., Burns, D. M., Borland, R., Johnson, M., & Bal, D. (1992). Occasional smoking among adults: evidence from the California Tobacco Survey. *Tobacco Control, 1*, 169–175.

Fagerström, K. O. (1978). Measuring degree of physical dependence to tobacco smoking with reference to individualization of treatment. *Addictive Behaviors, 3*, 235–241.

Fagerström, K. O. (1988). Efficacy of nicotine chewing gum: A review. In O. F. Pomerleau & C. S. Pomerleau (Eds.), *Nicotine replacement: A critical evaluation* (pp. 109–128). New York: Alan R. Liss.

Fagerström, K. O., Schneider, N. G., & Lunell, E. (1993). Effectiveness of nicotine patch and nicotine gum as individual versus combined treatments for tobacco withdrawal symptoms. *Psychopharmacology, 111*(3), 271–277.

Fiore, M. C., Bailey, W. C., Cohen, S. J., Dorfman, S. F., Goldstein, M. G., Gritz, E. R., Heyman, R. B., Hobrook, J., Jaen, C. R., Kottbe, T. E., Lando, H. A., Mecklenburg, R., Mullen, P. D., Nett, L. M., Robinson, L., Stitzer, M. L., Tommasello, A. C., Nillejo, L., & Newers, M. E. (1996). *Smoking cessation: Clinical Practice Guideline No. 18* (AHCPR Publication No. 96-0692). Rockville, MD: US Department of Health and Human Services, Agency for Health Care Policy and Research.

Fiore, M. C., Novotny, T. E., Pierce, J. P., Giovino, G. A., Hatziandreu, E. J., Newcomb, P. A., Surawicz T. S., & Davis, R. M. (1990). Methods used to quit smoking in the United States. Do cessation programs help? *Journal of the American Medical Association, 263*(20), 2760–2765.

Fiore, M. C., Smith, S. S., Jorenby, D. E., & Baker, T. B. (1994). The effectiveness of the nicotine patch for smoking cessation: A meta-analysis. *Journal of the American Medical Association, 271*(24), 1940–1947.

Fiore, M. C., Wetter, D. W., Bailey, W. C., Bennett, G., Cohen, S. J., Dorfman, S. F., Goldstein, M. G., Gritz, E. R., Hasselblad, V., Henningfield, J. E., Heyman, R. B., Holbrook, J., Husten, C., Jaen, C. R., Kohler, C., Kottke, T. E., Lando, H. A., Manley, M., Mecklenburg, R., Melvin, C., Mullen, P. D., Nett, L. M., Piasecki, T. M., Robinson, L., Rothstein, D., Schriger, D. L., Stitzer, M. L., Stachenko, S., Tommasello, A., Villejo, L., Wewers, M. E., & Baker, T. B. (1996). The agency for health-care policy and research smoking cessation clinical-practice guideline. *Journal of the American Medical Association, 275*(16), 1270–1280.

Fisher, K. J., Glasgow, R. E., & Terborg, J. R. (1990). Work site smoking cessation: A meta-analysis of long-term quit rates from controlled studies. *Journal of Occupational Medicine, 32*(5), 429–439.

Flay, B. R. (1985). What we know about the social influences approach to smoking prevention: Review and recommendations. *NIDA Research Monograph Series* (63), 67–112.

Flay, B. R., Koepke, D., Thomson, S. J., Santi, S., Best, J. A., & Brown, K. S. (1989). Six-year follow-up of the first Waterloo school smoking prevention trial. *American Journal of Public Health, 79*(10), 1371–1376.

Frederiksen, L. W., Epstein, L. H., & Kosevsky, B. P. (1975). Reliability and controlling effects of three procedures for self-monitoring smoking. *Psychological Record, 25,* 255–264.

Froggatt, P. (1988). *Fourth report of the Independent Scientific Committee on Smoking and Health*. London: HMSO.

Frost, C., Fullerton, F. M., Stephen, A. M., Stone, R., Nicolaides-Bouman, A., Densem, J., Wald, N. J., & Semmence, A. (1995). The tar reduction study: Randomised trial of the effect of cigarette tar yield reduction on compensatory smoking. *Thorax, 50*(10), 1038–1043.

Gill, B., Meltzer, H., Hinds, K., & Pettigrew, M. (1996). *Psychiatric morbidity among homeless people*. London: HMSO.

Giovino, G. A., Henningfield, J. E., Tomar, S. L., Escobedo, L. G., & Slade, J. (1995). Epidemiology of tobacco use and dependence. *Epidemiologic Reviews, 17*(1), 48–65.

Glassman, A. H. (1993). Cigarette smoking: Implications for psychiatric illness. *American Journal of Psychiatry, 150*(4), 546–553.

Glassman, A. H., Helzer, J. E., Covey, L. S., Cottler, L. B., Stetner, F., Tipp, J. E., & Johnson, J. (1990a). Smoking, smoking cessation, and major depression. *Journal of the American Medical Association, 264,* 1546–1549.

Glassman, A. H., Helzer, J. E., Covey, L. S., Cottler, L. B., Stetner, F., Tipp, J. E., & Johnson, J. (1990b). Smoking, smoking cessation, and major depression. *Journal of the American Medical Association, 264*(12), 1546–1549.

Glendinning, A., Shucksmith, J., & Hendry, L. (1994). Social class and adolescent smoking behavior. *Social Science and Medicine, 38,* 1449–1460.

Glynn, T. J., Greenwald, P., Mills, S. M., & Manley, M. W. (1993). Youth tobacco use in the United States—Problem, progress, goals, and potential solutions. *Preventive Medicine, 22*(4), 568–575.

Goddard, E. (1990). *Why children start smoking*. London: HMSO.

Gori, G. B., & Lynch, C. J. (1985). Analytical cigarette yields as predictors of smoke bioavailability. *Regulatory Toxicology and Pharmacology, 5*(3), 314–326.

Gourlay, S. G., Stead, L. F., & Benowitz, N. L. (1997). A meta-analysis of clonidine for smoking cessation. In T. Lancaster, C. Silagy, & D. Fullerton (Eds.), *Tobacco addiction module of the Cochrane database of systematic reviews* (Vol. Issue 2). Oxford, UK: The Cochrane Library, The Cochrane Collaboration, Update Software.

Gross, J., & Stitzer, M. L. (1989). Nicotine replacement: ten-week effects on tobacco withdrawal symptoms. *Psychopharmacology, 98*(3), 334–341.

Gritz, E. R., Carr, C. R., & Marcus, A. C. (1991). The tobacco withdrawal syndrome in unaided quitters. *British Journal of Addiction, 86*(1), 57–69.

Haglund, B., & Cnattingius, S. (1990). Cigarette smoking as a risk factor for sudden infant death syndrome: A population-based study. *American Journal of Public Health, 80*(1), 29–32.

Hajek, P. (1996). Current issues in behavioral and pharmacological approaches to smoking cessation. *Addictive Behaviors, 21*(6), 699–707.

Hajek, P., Jarvis, M. J., Belcher, M., Sutherland, G., & Feyerabend, C. (1989). Effect of smoke-free cigarettes on 24h cigarettes withdrawal: A double-blind placebo-controlled study. *Psychopharmacology, 97*(1), 99–102.

Halpern, M. T., Gillespie, B. W., & Warner, K. E. (1993). Patterns of absolute risk of lung cancer mortality in former smokers. *Journal of the National Cancer Institute, 85*(6), 457–464.

Hatsukami, D. K., Gust, W. G., & Keenan, R. M. (1987). Physiologic and subjective changes from smokeless tobacco withdrawal. *Clinical Pharmacology and Therapeutics, 41*(1), 103–107.

Heatherton, T. F., Kozlowski, L. T., Frecker, R. C., & Fagerström, K. O. (1991). The Fagerström Test for Nicotine Dependence: A revision of the Fagerström Tolerance Questionnaire. *British Journal of Addiction, 86*(9), 1119–1127.

Heatherton, T. F., Kozlowski, L. T., Frecker, R. C., Rickert, W. S., & Robinson, J. (1989). Measuring the heaviness of smoking: Using self-reported time to the first cigarette of the day and number of cigarettes smoked per day. *British Journal of Addiction, 84,* 791–799.

Hilleman, D. E., Mohiuddin, S. M., Core, M. G., & Sketch, M. H. (1992). Effect of buspirone on withdrawal symptoms associated with smoking cessation. *Archives of Internal Medicine, 152*(2), 350–352.

Hjalmarson, A., Franzon, M., Westin, A., & Wiklund, O. (1994). Effect of nicotine nasal spray on smoking cessation: A randomized, placebo-controlled, double blind study. *Archives of Internal Medicine, 154*(22), 2567–2572.

Henningfield, J. E., & Keenan, R. M. (1993). The anatomy of nicotine addiction. *Health Values, 17,* 12–19.

Henningfield, J. E., London, E. D., & Benowitz, N. L.

(1990). Arterial-venous differences in plasma concentrations of nicotine after cigarette smoking. *Journal of the American Medical Association, 263*(15), 2049–2050.

Himbury, S., & West, R. (1985). Smoking habits after laryngectomy. *British Medical Journal, 291*(6494), 514–515.

Holm, H., Jarvis, M. J., Russell, M. A. H., & Feyerabend, C. (1992). Nicotine intake and dependence in Swedish snuff takers. *Psychopharmacology, 108*(4), 507–511.

Ho-Yen, D. O., Spence, V. A., Moody, J. P., & Walker, W. F. (1982). Why smoke fewer cigarettes? *British Medical Journal, 284*(6333), 1905–1907.

Hughes, J. R. (1988). Dependence potential and abuse liability of nicotine replacement therapies. In O. F. Pomerleau, & C. S. Pomerleau (Eds.), *Nicotine replacement: A critical evaluation* (1st ed., pp. 261–277). New York: Alan R. Liss.

Hughes, J. R. (1992). Tobacco withdrawal in self-quitters. *Journal of Consulting and Clinical Psychology, 60*(5), 689–697.

Hughes, J. R. (1993). Non-nicotine pharmacotherapies for smoking cessation. *Journal of Drug Development, 6*(4), 197–203.

Hughes, J. R. (1995). Combining behavioral therapy and pharmacotherapy for smoking cessation: An update. *NIDA Research Monograph Series* (150), 92–109.

Hughes, J. R., & Glaser, M. (1993). Transdermal nicotine for smoking cessation. *Health Values 17*(2), 25–31.

Hughes, J. R., Gulliver, S. B., Fenwick, J. W., Valliere, W. A., Cruser, K., Pepper, S., Shea, P., Solomon, L. J., & Flynn, B. S. (1992). Smoking cessation among self-quitters. *Health Psychology, 11*(5), 331–334.

Hughes, J. R., Gust, S. W., Skoog, K., Keenan, R. M., & Fenwick, J. W. (1991). Symptoms of tobacco withdrawal. A replication and extension. *Archives of General Psychiatry, 48*(1), 52–59.

Hughes, J. R., & Hatsukami, D. K. (1986). Signs and symptoms of tobacco withdrawal. *Archives of General Psychiatry 43*(3), 289–294.

Hughes, J. R., Hatsukami, D. K., Pickens, R. W., Krahn, D., Malin, S., & Luknic, A. (1984). Effect of nicotine on the tobacco withdrawal syndrome. *Psychopharmacology 83*(1), 82–87.

Hugnes, J. R., Wadland, W. C., Fenwick, J. W., Lewis, J., & Bickel, W. K. (1991). Effect of cost on the self-administration and efficacy of nicotine gum: A preliminary study. *Preventive Medicine, 20*(4), 486–496.

Hurt, R. D., Dale, L. C., Offord, K. P., Bruce, B. K., McClain, F. L., & Eberman, K. M. (1992). Inpatient treatment of severe nicotine dependence. *Mayo Clinic Proceedings, 67*(9), 823–828.

Jackson, P. H., Stapleton, J. A., Russell, M. A. H., & Merriman, R. J. (1986). Predictors of outcome in a general practitioner intervention against smoking. *Preventive Medicine, 15*(3), 244–253.

Jarvis, M. J. (1989). Application of biochemical intake markers to passive smoking measurement and risk estimation. *Mutation Research, 222*(2), 101–110.

Jarvis, M. J. (1994a). Gender differences in smoking cessation: real or myth? *Tobacco Control, 3,* 324–328.

Jarvis, M. J. (1994b). A profile of tobacco smoking. *Addiction, 89*(11), 1371–1376.

Jarvis, M. J. (1995). Smoking cessation: Time for action. *Thorax, 50*(S1), S22–S24.

Jarvis, M. J. (1996). The association between having children, family size and smoking cessation in adults. *Addiction, 91*(3), 427–434.

Jarvis, M. J. (1997). Patterns and predictors of smoking cessation in the general population. In C. T. Bolliger, & K. O. Fagerström, (Eds.), *The tobacco epidemic. Progress in respiratory research* (Vol. 28, pp. 151–164). Basel, Switzerland: Karger.

Jarvis, M. J., Raw, M., Russell, M. A. H., & Feyerabend, C. (1982). Randomised controlled trial of nicotine chewing-gum. *British Medical Journal, 285*(6341), 537–540.

Jarvis, M. J., Russell, M. A. H., Feyerabend, C., Eiser, J. R., Morgan, M., Gammage, P., & Gray, E. M. (1985). Passive exposure to tobacco smoke: Saliva cotinine concentrations in a representative population sample of non-smoking schoolchildren. *British Medical Journal, 291*(6500), 927–929.

Jarvis, M. J., Russell, M. A. H., & Saloojee, Y. (1980). Expired air carbon monoxide: A simple breath test of tobacco smoke intake. *British Medical Journal, 281*(6238), 484–485.

Jarvis, M. J., Tunstall-Pedoe, H., Feyerabend, C., Vesey, C. J., & Saloojee, Y. (1984). Biochemical markers of smoke absorption and self reported exposure to passive smoking. *Journal of Epidemiology and Community Health, 38*(4), 335–339.

Jarvis, M. J., Tunstall-Pedoe, H., Feyerabend, C., Vesey, C., & Saloojee, Y. (1987). Comparison of tests used to distinguish smokers from nonsmokers. *American Journal of Public Health, 77*(11), 1435–1438.

Kessler, D. A. (1995). Nicotine addiction in young people. *New England Journal of Medicine, 333*(3), 186–189.

Kessler, D. A., Barnett, P. S., Witt, A., Zeller, M. R., Mande, J. R., & Schultz, W. B. (1997). The legal and scientific basis for FDA's assertion of jurisdiction over cigarettes and smokeless tobacco. *Journal of the American Medical Association, 277*(5), 405–409.

Kessler, D. A., Witt, A. M., Barnett, P. S., Zeller, M. R., Natanblut, S. L., Wilkenfeld, J. P., Lorraine, C. C., Thompson, L. J., & Schultz, W. B. (1996). The Food and Drug Administration's regulation of tobacco products. *New England Journal of Medicine, 335*(13), 988–994.

Klesges, R. C., Ward, K. D., & Debon, M. (1996). Smoking cessation: A successful behavioral/pharmacologic interface. *Clinical Psychology Review, 16*(6), 479–496.

Klonoff-Cohen, H. S., Edelstein, S. L., Lefkowitz, E. S., Srinivasan, I. P., Kaegi, D., Jae Chun, C., & Wiley, K. J. (1995). The effect of passive smoking and tobacco exposure through breast milk on sudden infant death syndrome. *Journal of the American Medical Association, 273*(10), 796–798.

Knott, V. J., & Venables, P. H. (1977). EEG alpha correlates of non-smokers, smokers, smoking, and smoking deprivation. *Psychophysiology, 14*(2), 150–156.

Kornitzer, M., Boutsen, M., Dramaix, M., Thijs, J., & Gustavsson, G. (1995). Combined use of nicotine patch and gum in smoking cessation: A placebo-controlled clinical trial. *Preventive Medicine, 24*(1), 41–47.

Kozlowski, L. T., & Henningfield, J. E. (1995). Thinking the unthinkable: The prospect of regulation of nicotine in cigarettes by the United States government. *Addiction, 90*(2), 165–167.

Kozlowski, L. T., Wilkinson, A., Skinner, W., Kent, C., Franklin, T., & Pope, M. (1989). Comparing tobacco cigarette dependence with other drug dependencies. Greater or equal "difficulty quitting" and "urges to use," but less "pleasure" from cigarettes. *Journal of the American Medical Association, 261*(6), 898–901.

Kozlowski, L. T. (1984). The interaction of psychosocial and biological determinants of tobacco use: More on the boundary model. *Journal of Applied Social Psychology, 14*(3) 244–256.

Lee, A. J., Crombie, I. K., Smith, W. C. S., & Tunstall-Pedoe, H. D. (1991). Cigarette smoking and employment status. *Social Science and Medicine, 33*(11), 1309–1312.

Lichtenstein, E., & Glasgow, R. E. (1992). Smoking cessation: What have we learned over the past decade? *Journal of Consulting & Clinical Psychology, 60*(4), 518–527.

Lichtenstein, E., & Hollis, J. (1992). Patient referral to a smoking cessation program: Who follows through? *Journal of Family Practice, 34*(6), 739–744.

Lynch, B. S., & Bonnie, R. J. (Eds.) (1994). *Growing up tobacco free: Preventing nicotine addiction in children and youths*. Washington, DC: National Academy Press.

Marsh, A., & McKay, S. (1994). *Poor smokers*. London: Policy Studies Institute.

McFall, R. M. (1978). Smoking cessation research. *Journal of Consulting & Clinical Psychology, 46*(4), 703–712.

McNeill, A. D. (1991). The development of dependence on smoking in children. *British Journal of Addiction, 86*(5), 589–592.

McNeill, A. D., Jarvis, M. J., Stapleton, J. A., West, R. J., & Bryant, A. (1989). Nicotine intake in young smokers: Longitudinal study of saliva cotinine concentrations. *American Journal of Public Health, 79*(2), 172–175.

McNeill, A. D., West, R. J., Jarvis, M., Jackson, P., & Bryant, A. (1986). Cigarette withdrawal symptoms in adolescent smokers. *Psychopharmacology, 90*(4), 533–536.

Meltzer, H., Gill, B., Pettigrew, M., & Hinds, K. (1995). *The prevalence of psychiatric morbidity among adults living in private households* (Vol. Report 1). London: HMSO.

Murray, R. P., Johnston, J. J., Dolce, J. J., Lee, W. W., & Ohara, P. (1995). Social support for smoking cessation and abstinence: The Lung Health Study. *Addictive Behaviors, 20*(2), 159–170.

Nafstad, P., Botten, G., & Hagen, J. (1996). Partner's smoking: A major determinant for changes in women's smoking behavior during and after pregnancy. *Public Health, 110,* 379–385.

Nutbeam, D., Macaskill, P., Smith, C., Simpson, J. M., & Catford, J. (1993). Evaluation of two school smoking education programs under normal classroom conditions. *British Medical Journal, 306*(6870), 102–107.

Peto, R. (1986). Overview of cancer time-trend studies in relation to changes in cigarette manufacture. In D. G. Zaridze, & R. Peto (Eds.), *Tobacco: A major international health hazard* (IARC Scientific Publications No. 74, pp. 211–226). Lyon, France: IARC.

Peto, R., Lopez, A. D., Boreham, J., Thun, M., & Heath, C. (1994). *Mortality from smoking in developed countries 1950–2000: Indirect estimates from national vital statistics*. Oxford, UK: Oxford University Press.

Peto, R., Lopez, A. D., Boreham, J., Thun, M., Heath, C., & Doll, R. (1996). Mortality from smoking worldwide. *British Medical Bulletin, 52*(1), 12–21.

Phillips, A. N., & Smith, G. D. (1994). Cigarette smoking as a potential cause of cervical cancer: Has confounding been controlled? *International Journal of Epidemiology, 23*(1), 42–49.

Pich, E. M., Pagliusi, S. R., Tessari, M., Talabot-Ayer, D., Hooft van Huisduijnen, R., & Chiamulera, C. (1997). Common neural substrates for the addictive properties of nicotine and cocaine. *Science, 275*(5296), 83–86.

Pierce, J. P., Fiore, M. C., Novotny, T. E., Hatziandreu, E. J., & Davis, R. M. (1989a). Trends in cigarette smoking in the United States. Educational differences are increasing. *Journal of the American Medical Association, 261*(1), 56–60.

Pierce, J. P., Fiore, M. C., Novotny, T. E., Hatziandreu, E. J., & Davis, R. M. (1989b). Trends in cigarette smoking in the United States. Projections to the year 2000. *Journal of the American Medical Association, 261*(1), 61–65.

Pollay, R. W., Siddarth, S., Siegel, M., Haddix, A., Merritt, R. K., Giovino, G. A., & Eriksen, M. P. (1996). The last straw—cigarette advertising and realized market shares among youths and adults, 1979–1993. *Journal of Marketing, 60*(2), 1–16.

Pontieri, F. E., Tanda, G., Orzi, F., & Dichiara, G. (1996). Effects of nicotine on the nucleus accumbens and similarity to those of addictive drugs. *Nature, 382,* 255–257.

Poswillo, D., & Alberman, E. (1992). *Effects of smoking on the fetus, neonate, and child*. Oxford, UK: Oxford University Press.

Prochaska, J. O., & DiClemente, C. C. (1986). Towards a comprehensive model of change. In W. E. Miller, & N. Heather (Eds.), *Treating addictive behaviors: Processes of change* (pp. 3–27). New York: Plenum.

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist, 47*(9), 1102–1114.

Prochaska, J. O., & Goldstein, M. G. (1991). Process of smoking cessation: Implications for clinician. *Clinics In Chest Medicine, 12*(4), 727–735.

R. J. Reynolds Tobacco Company (1988). *Chemical and biological studies on new cigarette prototypes that heat instead of burn tobacco*. Winston-Salem, NC: Author.

Rose, J. E., & Behm, F. M. (1994). Inhalation of vapor from black pepper extract reduces smoking withdrawal symptoms. *Drug & Alcohol Dependence, 34*(3), 225–229.

Rose, J. E., Behm, F. M., & Levin, E. D. (1993). Role of nicotine dose and sensory cues in the regulation of smoke intake. *Pharmacology Biochemistry & Behavior, 44*(4), 891–900.

Rose, J. E., Behm, F. M., Westman, E. C., Levin, E. D., Stein, R. M., & Ripka, G. V. (1994). Mecamylamine combined with nicotine skin patch facilitates smoking cessation beyond nicotine skin patch treatment alone. *Clinical Pharmacology & Therapeutics, 56*(1), 86–99.

Rose, J. E., & Hickman, C. S. (1987). Citric acid aerosol as a potential cessation aid. *Chest, 92*(6), 1005–1008.

Rosenberg, L., Kaufman, D. W., Helmrish, S. P., & Shapiro, S. (1985). The risk of myocardial infarction after quitting smoking in men under 55 years of age. *New England Journal of Medicine, 313*(24), 1511–1514.

Royal College of Physicians (1971). *Smoking and Health Now*. London: Pitman Medical and Scientific Publishing.

Russell, M. A. H. (1980). Nicotine intake and its regulation. *Journal of Psychosomatic Research, 24*(5), 253–264.

Russell, M. A. H. (1988). Nicotine replacement: The role of blood nicotine levels, their rate of change, and nicotine tolerance. In O. F. Pomerleau, & C. S. Pomerleau (Eds.), *Nicotine replacement: A critical evaluation* (1st ed., pp. 63–94). New York: Alan R. Liss.

Russell, M. A. H., Jarvis, M. J., Devitt, G., & Feyerabend, C. (1981). Nicotine intake by snuff users. *British Medical Journal, 283*(6295), 814–817.

Russell, M. A. H., Jarvis, M., Iyer, R., & Feyerabend, C. (1980). Relation of nicotine yield of cigarettes to blood nicotine concentrations in smokers. *British Medical Journal, 280*(6219), 972–976.

Russell, M. A. H., Jarvis, M. J., Sutherland, G., & Feyerabend, C. (1987). Nicotine replacement in smoking cessation: Absorption of nicotine vapor from smoke-free cigarettes. *Journal of the American Medical Association, 257*(23), 3262–3265.

Russell, M. A. H., Stapleton, J. A., Feyerabend, C., Wiseman, S. M., Gustavsson, G., Sawe, U., & Connor, P. (1993). Targeting heavy smokers in general practice: Randomised controlled trial of transdermal nicotine patches. *British Medical Journal, 306*(6888), 1308–1312.

Schneider, F. H., Mione, P. J., Raheman, F. S., Phillips, B. M., & Quiring, J. N. (1996a). Reduction of tobacco withdrawal symptoms by sublingual lobeline sulfate. *American Journal of Health Behavior, 20*(5), 346–363.

Schneider, N. G., Jarvik, M. E., & Forsythe, A. B. (1984). Nicotine vs. placebo gum in the alleviation of withdrawal during smoking cessation. *Addictive Behaviors, 9*(2), 149–156.

Schneider, N. G., Olmstead, R., Mody, F. V., Doan, K., Franzon, M., Jarvik, M. E., & Steinberg, C. (1995). Efficacy of a nicotine nasal spray in smoking cessation: A placebo-controlled, double-blind trial. *Addiction, 90*(12), 1671–1682.

Schneider, N. G., Olmstead, R., Nilsson, F., Mody, F. V., Franzon, M., & Doan, K. (1996b). Efficacy of a nicotine inhaler in smoking cessation: A double-blind, placebo-controlled trial. *Addiction, 91*(9), 1293–1306.

Schneider, N. G., Olmstead, R. E., Steinberg, C., Sloan, K., Daims, R. M., & Brown, H. V. (1996c). Efficacy of buspirone in smoking cessation: A placebo-controlled trial. *Clinical Pharmacology & Therapeutics, 60*(5), 568–575.

Schoenborn, C. A., & Horm, J. (1993). *Negative moods as correlates of smoking and heavier drinking: Implications for health promotion.* Hyattsville, MD: National Center for Health Statistics.

Schwartz, J. (1987). *Review and evaluation of smoking cessation methods: The United States and Canada, 1978–1985* (National Institute Health Publication No. 87–2940). Bethesda, MD: National Cancer Institute.

Severson, H. H., Andrews, J. A., Lichtenstein, E., Wall, M., & Zoref, L. (1995). Predictors of smoking during and after pregnancy: A survey of mothers of newborns. *Preventive Medicine, 24*(1), 23–28.

Sexton, M., & Hebel, J. R. (1984). A clinical trial of change in maternal smoking and its effect on birth weight. *Journal of the American Medical Association, 251*(7), 911–915.

Shiffman, S. (1989). Tobacco "chippers." Individual differences in tobacco dependence. *Psychopharmacology, 97*(4), 539–547.

Shiffman, S. M., & Jarvik, M. E. (1976). Smoking withdrawal symptoms in two weeks of abstinence. *Psychopharmacology, 50*(1), 35–39.

Silagy, C., Mant, D., Fowler, G., & Lodge, M. (1994). Meta-analysis on efficacy of nicotine replacement therapies in smoking cessation. *Lancet, 343*(8890), 139–142.

Slade, J. (1995). Are tobacco products drugs? Evidence from US tobacco (Editorial). *Tobacco Control, 4,* 1–2.

Slade, J., Bero, L. A., Hanauer, P., Barnes, D. E., & Glantz, S. A. (1995). Nicotine and addiction: The Brown and Williamson documents. *Journal of the American Medical Association, 274*(3), 225–233.

Slotkin, T. A., Orband-Miller, L., & Queen, K. L. (1987). Development of (3H)nicotine binding sites in brain regions of rats exposed to nicotine prenatally via maternal injections or infusions. *Journal of Pharmacology and Experimental Therapeutics, 242*(1), 232–237.

Smith, G. D., Strobele, S. A., & Egger, M. (1994). Smoking and health promotion in Nazi Germany. *Journal of Epidemiology and Community Health, 48*(3), 220–223.

Snyder, F., Davis, F., & Henningfield, J, (1989). The tobacco withdrawal syndrome: Performance decrements assessed on a computerised test battery. *Drug & Alcohol Dependence, 23*(3), 259–266.

Spealman, R. D., & Goldberg, S. R. (1982). Maintenance of schedule-controlled behavior by intravenous injections of nicotine in squirrel monkeys. *Journal of Pharmacology and Experimental Therapeutics, 223*(2), 402–408.

Spring, B., Wurtman, J., Wurtman, R., El-Khoury, A., Goldberg, H., McDermott, J., & Pingitore, R. (1995). Efficacies of dexfenfluramine and fluoxetine in preventing weight gain after smoking cessation. *American Journal of Clinical Nutrition, 62*(6), 1181–1187.

Stapleton, J. A., Russell, M. A. H., Feyerabend, C., Wiseman, S. M., Gustavsson, G., & Säwe, U. (1995). Dose effects and predictors of outcome in a randomized trial of transdermal nicotine patches in general practice. *Addiction, 90*(1), 31–42.

Steenland, K. (1992). Passive smoking and the risk of heart disease. *Journal of the American Medical Association, 267*(1), 94–99.

Stolerman, I. P., & Jarvis, M. J. (1995). The scientific case that nicotine is addictive. *Psychopharmacology, 117*(1), 2–10.

Strachan, D. P., Jarvis, M. J., & Feyerabend, C. (1989). Passive smoking, salivary cotinine concentrations, and middle ear effusion in 7 year old children. *British Medical Journal, 298*(6687), 1549–1552.

Strachan, D. P., Jarvis, M. J., & Feyerabend, C. (1990). The relationship of salivary cotinine to respiratory symptoms, spirometry, and exercise-induced bronchospasm in seven-year-old children. *American Review of Respiratory Disease, 142*(1), 147–151.

Sutherland, G., Russell, M. A. H., Stapleton, J. A., & Feyerabend, C. (1993). Glycerol particle cigarettes: A less harmful option for chronic smokers. *Thorax, 48*(4), 385–387.

Sutherland, G., Russell, M. A. H., Stapleton, J., Feyerabend, C., & Ferno, O. (1992). Nasal nicotine spray: A rapid nicotine delivery system. *Psychopharmacology, 108*(4), 512–518.

Sutherland, G., & Stapleton, J. A. (1994). Nasal nicotine spray for dependent smokers. *Journal of Smoking Related Disorders, 5*(Suppl. 1), 195–201.

Sutherland, G., Stapleton, J., Russell, M. A. H., & Feyerabend, C. (1995). Naltrexone, smoking behavior and cigarette withdrawal. *Psychopharmacology, 120*(4), 418–425.

Sutherland, G., Stapleton, J. A., Russell, M. A. H., Jarvis, M. J., Hajek, P., Belcher, M., & Feyerabend, C. (1992). Randomised controlled trial of nasal nicotine spray in smoking cessation. *Lancet, 340*(8815), 324–329.

Sutton, S. R. (1996). Can "stages of change" provide guidance in the treatment of addictions? A critical examination of Prochaska and DiClemente's model. In G. Edwards & C. Dare (Eds.), *Psychotherapy, psychological treatments, and the addictions* (pp. 189–205). Cambridge, UK: Cambridge University Press.

Thun, M. J., Day-Lally, C. A., Calle, E. E., Flanders, W. D., & Heath, C. W. (1995). Excess mortality among cigarette smokers: Changes in a 20-year interval. *American Journal of Public Health, 85*(9), 1223–1230.

Tønnesen, P., Nørregaard, J., Mikkelsen, K., Jørgensen, S., & Nilsson, F. (1993). A double-blind trial of a nicotine inhaler for smoking cessation. *Journal of the American Medical Association, 269*(10), 1268–1271.

US Department of Health and Human Services (1988). *The health consequences of smoking: Nicotine addiction. A report of the Surgeon General* (DHSS Publ. No. (CDC) 88-8406). Washington, DC: US Government Printing Office.

US Department of Health and Human Services (1989). *Reducing the health consequences of smoking: 25 years of progress. A report of the Surgeon General.* Washington, DC: US Government Printing Office.

US Department of Health and Human Services (1990). *The health benefits of smoking cessation: a report of the Surgeon General.* Washington, DC: US Department of Health and Human Services, Public Health Service, Centers for Disease Control, Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

US Department of Health and Human Services (1996). *The FTC cigarette test method for determining tar, nicotine and carbon monoxide yields of US cigarettes* (NIH Publication No. 96-4028). Bethesda, MD: National Cancer Institute.

US Environmental Protection Agency (1992). Respiratory health effects of passive smoking: Lung cancer and other disorders. Washington, DC: Author.

Wakefield, M., Gillies, P., Graham, H., Madeley, R., & Symonds, M. (1993). Characteristics associated with smoking cessation during pregnancy among working-class women. *Addiction, 88*(10), 1423–1430.

Wald, N. J., & Hacksaw, A. K. (1996). Cigarette smoking: An epidemiological overview. *British Medical Bulletin, 52*(1), 3–11.

Wald, N. J., Idle, M., Boreham, J., & Bailey, A. (1981). Carbon monoxide in breath in relation to smoking and carboxyhaemoglobin levels. *Thorax, 36*(5), 366–369.

Wells, A. J. (1994). Passive smoking as a cause of heart disease. *Journal of the American College of Cardiology, 24*(2), 546–554.

West, R. J., Hajek, P., & Belcher, M. (1987). Time course of cigarette withdrawal symptoms during four weeks of treatment with nicotine chewing gum. *Addictive Behaviors, 12*(2), 199–203.

West, R. J., Hajek, P., & Belcher, M. (1989). Severity of withdrawal symptoms as a predictor of outcome of an attempt to quit smoking. *Psychological Medicine, 19*(4), 981–985.

West, R. J., Hajek, P., & Burrows, S. (1990). Effect of glucose tablets on craving for cigarettes. *Psychopharmacology, 101*(4), 555–559.

West, R. J., Hajek, P., & McNeill, A. (1991). Effect of buspirone on cigarette withdrawal symptoms and short-term abstinence rates in a smokers clinic. *Psychopharmacology, 104*(1), 91–96.

West, R. J., Jarvis, M. J., Russell, M. A. H., Carruthers, M. E., & Feyerabend, C. (1984). Effect of nicotine replacement on the cigarette withdrawal syndrome. *British Journal of Addiction, 79*(2), 215–219.

West, R. J., & Russell, M. A. H. (1985a). Effects of withdrawal from long-term nicotine gum use. *Psychological Medicine, 15*(4), 891–893.

West, R. J., & Russell, M. A. H. (1985b). Pre-abstinence smoke intake and smoking motivation as predictors of severity of cigarette withdrawal symptoms. *Psychopharmacology, 87*(3), 334–336.

West, R. J., & Russell, M. A. H. (1987). Cardiovascular and subjective effects of smoking before and after 24 h of abstinence from cigarettes. *Psychopharmacology, 92*(1), 118–121.

Wills, T. A., Pierce, J. P., & Evans, R. I. (1996). Large-scale environmental risk-factors for substance use. *American Behavioral Scientist, 39*(7), 808–822.

Withey, C. H., Papacosta, A. O., Swan, A. V., Fitzsimons, B. A., Ellard, G. A., Burney, P. G. J., Colley, J. R. T., & Holland, W. W. (1992). Respiratory effects of lowering tar and nicotine levels of cigarettes smoked by young male middle tar smokers: II. Results of a randomized controlled trial. *Journal of Epidemiology and Community Health, 46*(3), 281–285.

Woody, G. E., Cottler, L. B., & Cacciola, J. (1993). Severity of dependence: Data from the *DSM-IV* field trials. *Addiction, 88*(11), 1573–1579.

World Health Organization (1992). *The ICD-10 classification of mental and behavioral disorders. Clinical descriptions and diagnostic guidelines*. Geneva, Switzerland: Author.

Wynder, E. L., & Graham, E. A. (1950). Tobacco smoking as a possible etiologic factor in bronchogenic carcinoma: A study of 684 proved cases. *Journal of the American Medical Association, 143,* 329–336.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

8.29
Health Promotion: Perspectives on Physical Activity and Weight Control

NEVILLE OWEN and DAVID CRAWFORD
Deakin University, Melbourne, Vic, Australia

8.29.1 INTRODUCTION 675
8.29.1.1 Health Promotion 675
8.29.2 INDIVIDUAL VS. PUBLIC HEALTH APPROACHES IN HEALTH PROMOTION 676
8.29.3 PHYSICAL ACTIVITY AS A PUBLIC HEALTH CONCERN 677
8.29.4 WEIGHT CONTROL AS A PUBLIC HEALTH CONCERN 678
8.29.5 BEHAVIORAL EPIDEMIOLOGY 678
8.29.5.1 The Descriptive Epidemiology of Physical Activity 679
8.29.5.2 Environmental and Public Policy Approaches: Physical Activity Promotion 679
8.29.5.3 The Descriptive Epidemiology of Overweight and Obesity in Australia 680
8.29.5.4 Environmental and Public Policy Approaches: Dealing with Overweight and Obesity 681
8.29.6 UNDERSTANDING AND INFLUENCING HEALTH BEHAVIORS: THE EXAMPLE OF THE TRANSTHEORETICAL MODEL 682
8.29.6.1 Applying the Transtheoretical Model to Physical Activity Promotion 684
8.29.6.2 Applying the Transtheoretical Model to Weight Control Behavior 685
8.29.7 CONCLUSIONS 686
8.29.8 REFERENCES 687

8.29.1 INTRODUCTION

Clinical health psychology, with its roots in the understanding of human behavior and experience, its strong applied focus and the theoretical developments which it has spawned, has had a significant influence on the theory and practice of health promotion. Other chapters address clinical or individually focused health promotion and behavior change strategies, and some of the contemporary theoretical frameworks which are helpful in understanding how to change the health-related behaviors of individuals. In this chapter, our emphasis is on how research from health psychology and the related field of behavioral epidemiology can provide some useful components of the conceptual and informational bases of large-scale efforts to influence health risk. We illustrate how a widely used model from health psychology (the transtheoretical model) may be used to better understand physical activity and weight control behaviors.

8.29.1.1 Health Promotion

Health promotion is a broad field of endeavour, and it is beyond the scope of this chapter to explore fully all of its elements. The breadth of the field is made clear by the following definition of health promotion as: "any planned combination of educational, political, regulatory and organizational supports for actions and conditions of living conducive to the health of individuals, groups or communities" (Green & Kreuter, 1991, p. 432). This definition makes it quite clear that health promotion encompasses much more than education about how to be more healthy, or personal health-risk reduction through self-initiated behavioral change. A public health perspective on health promotion involves attention to such matters as the effectiveness of mass-reach communication campaigns, planned changes to community environments, and attention to broader patterns of social or fiscal incentives for health-related behaviors.

Such large-scale health promotion initiatives can be improved through stronger connections between health psychology theory and practical intervention procedures. This can be achieved, for example, by facilitating extrapolations from laboratory research to community intervention methods; by using consistent theoretical frameworks to guide the health promotion planning process; and by providing disciplined accounts of the data that are relevant to decisions about the appropriateness and likely impact of public health programs and strategies.

The type of regulatory and social changes implied by Green and Kreuter's above definition can have a broad impact on large numbers of people. The legitimacy of such large-scale social change initiatives have been challenged on ethical grounds, focusing on concerns about deliberately engineered social controls and the manipulation of the behaviors of large numbers of people for the purpose of promoting health (Skrabanek, 1986). Concerns raised include, for example, the extent to which participation in, or exposure to, large-scale health promotion programs is truly voluntary; the harm that might be done to those who may be strongly coerced into changing health-related behaviors when they may not wish to do so; and, the extent to which large numbers of people might be being deprived of personal liberties, for example, through bans on cigarette smoking in workplaces, being coerced to be more physically active, or having their choices of food restricted.

Any health promotion initiative, particularly when people are encouraged or constrained to make pervasive changes to lifestyle patterns, should be justified by sound epidemiological evidence. There must be a plausible biological and epidemiological rationale (see Chapter 8.01 this volume and also Kaplan, 1990) for why the behavior of concern should be regarded as a health risk of sufficient magnitude to be addressed by large-scale initiatives which will at least influence the knowledge and attitudes or will impact on the ways of life of large numbers of people.

Not all ethical concerns about health promotion interventions may be addressed by sound epidemiological rationales, by employing theoretical frameworks and methodologically rigorous data collection techniques to evaluate the efficacy of interventions, or by applying theory and methods from health psychology in as rigorous a fashion as possible. Nevertheless, rigorously applying soundly based theory and research may make it less likely that ineffective, wasteful or potentially deleterious approaches to health promotion will be developed and applied.

8.29.2 INDIVIDUAL VS. PUBLIC HEALTH APPROACHES IN HEALTH PROMOTION

Public-health approaches to changing health-related behaviors require the development of intervention methods which may be used on a community-wide basis. Rose (1992) describes some of the major differences between a clinical and a broader population perspective on understanding and influencing risks to health. This different style of thinking has been a significant influence on the practice of prevention, but the relevant knowledge base is far from complete. Although there are now results available from, for example, a number of community-wide cardiovascular risk intervention trials (e.g., Young, Haskell, Jatulius, & Fortmann, 1993), much of what is known about health-related behavior change is based on intensive research with high-risk populations, or on laboratory, clinic, and small community-based studies which often involve intensive programs delivered to small numbers of people.

Often, clinical health psychology activities deal with individuals who already are involved in some form of action to change. In contrast to this, in the health promotion and public-health context, a key issue is encouraging much larger groups in the community environment to contemplate change, and to choose to take action (Jeffery, 1989). A challenge for psychologists who may wish to promote health-related behaviors is to come to terms with a population perspective on behavior and health (Kaplan, 1985, 1990).

During the late 1980s, psychologists became more strongly interested in the public health issues which arise in relation to health-related behaviors (Best & Proctor, 1988; Ewart, 1990;

Gochman, 1988; Jeffery, 1989; Lichtenstein & Glasgow, 1992; Winett, King, & Altman, 1989). The aim of this approach is to provide as many individuals as possible with assistance at a low cost. This interest has created a demand for effective ways of integrating the approaches of the disciplines of psychology and public health (Oldenburg & Owen, 1990; Winett et al. 1989).

It is informative to consider the practical advantages and disadvantages of clinical compared to mass-reach health promotion (see Table 1). Clinical (or face-to-face) methods involve individual counseling, work with small groups, and other forms of personal contact. Public health or mass-reach approaches often use print or audiovisual media, and are usually made available through broadcast or mass-distribution channels (Flay, 1987).

It is not being suggested here that a clinical perspective is ineffective or irrelevant to large-scale health community and population health promotion efforts. The public health approach can incorporate a whole range of more focused programs and services, as well as community-wide and environmental strategies, which can act to strengthen, support and maintain the health gains that health professionals can achieve with individuals in clinical and other settings.

8.29.3 PHYSICAL ACTIVITY AS A PUBLIC HEALTH CONCERN

It is now clear from a number of large, cross-sectional studies that there is a wide range of health benefits associated with habitually high levels of energy expenditure, and with increases from relatively low habitual levels (Blair, Kohl, Gordon, & Paffenbarger, 1992). Regular, vigorous exercise results in a number of beneficial changes to cardiovascular risk factors and decreased risk of heart disease may be achieved at relatively low intensities of exercise levels (Pate et al., 1995). Physical activity also has a significant protective effect against some cancers (Lee, 1994).

The basal metabolic rate, which accounts for some 70% of energy expenditure, is increased by exercise. This, as well as the direct expenditure of energy during the activity, leads to weight loss if regular exercise is combined with moderate dietary restraint (Grilo, 1995). Other physical benefits include improved muscle strength and joint flexibility (DiPietro, 1996). Among noninsulin-dependent diabetics, exercise appears to have a number of beneficial metabolic effects: it can improve carbohydrate metabolism, can reduce blood glucose levels, and can increase the effectiveness of insulin and insulin receptors. It is estimated that physical inactivity accounts for about 35% of deaths from adult-onset diabetes (Powell & Blair, 1994). The likelihood of osteoporosis and risk of fracture may be reduced by exercising: higher levels of weight-bearing physical activity are associated with increased bone density levels (Drinkwater, 1994).

Regular physical activity may also have mental health benefits: alleviating mild-to-moderate depression and anxiety, and improving confidence and self-concept. While exercise has the potential to bring about such benefits, it is not yet possible to determine whether these

Table 1 Some relative advantages and disadvantages of clinical and public health approaches to health promotion.

	Advantages	*Disadvantages*
Clinical or face-to-face programs	Interactive Reliable Provide social support Allow for personalizing Allow modeling Appropriate sequencing easy Follow-up easy	Expensive Market penetration weak May encourage dependency May not be acceptable to many people
Large-scale public health programs	Cheap per contact Large numbers reached More acceptable Can use social modeling techniques effectively May stimulate self-initiated change Potential for further development through modern technology	Weak engagement of users Unreliable Dilution of content Distortion Sequencing difficult Follow-up difficult

Source: Adapted from Owen and Lee (1986, p. 84).

result from exercising itself. It seems equally likely that they could result from influences such as social support and confidence building, which are difficult to disentangle from the conduct of physical training programs. Improved sleeping patterns and feelings of relaxation can be produced by moderate levels of exercise, and exercise may be associated with other psychological benefits (Martinsen & Stephens, 1994).

As well as there being a wide range of evidence from cross-sectional studies, recent prospective studies have shown associations between change in physical activity levels and physical health benefits (e.g., Young et al., 1993). There is now a general scientific consensus that physical activity promotion efforts directed to whole populations are justified by a range of well-documented health benefits. (e.g., Bouchard, Shephard, Stephens, Sutton & McPherson, 1990; Gloag, 1992; US Department of Health and Human Services, 1996).

8.29.4 WEIGHT CONTROL AS A PUBLIC HEALTH CONCERN

Health authorities in many industrialized countries have for some years recommended that the prevalence of obesity should be reduced in the adult population. These recommendations are based on substantial evidence that demonstrates that higher relative weights are associated with greater risk of mortality and morbidity for a range of common conditions. This evidence was reviewed in the 1980s by expert groups in the UK (Royal College of Physicians, 1983), the USA (Burton, Hirsch, & Van Itallie, 1985), and Canada (Health and Welfare Canada, 1988), and as a consequence reports or statements were released identifying associations between excess weight and mortality and/or morbidity from conditions including noninsulin-dependent diabetes mellitus, coronary heart disease, hypertension, hyperlipidemia, and certain types of cancers.

In a more recent review of the literature, Pi-Sunyer (1991) concluded that there was " ample evidence to suggest that obesity increases both morbidity and mortality" (p.1600S), although debate continues concerning the strength of the independent risk of obesity vs. its effect on other risk factors, and where in the excess-weight spectrum specific risks begin. Similarly, an expert group in the UK recently concluded that "obesity is an important health hazard both independently and in association with other diseases," and they argued that "the true adverse consequences of obesity are underestimated, particularly for those who are only moderately overweight" (Kopelman, Finer, Fox, Hill, & MacDonald, 1994, p. 189). Research has also shown that increased upper body or abdominal fat is associated with greater risk of premature mortality and morbidity, although the epidemiological evidence relating differences in body fat distribution to health is far less extensive than the data concerning excess weight (Pi-Sunyer, 1991).

Two major sources of data are available on weight in the Australian adult population. These data are derived from the National Heart Foundation's Risk Factor Prevalence Study, conducted between 1980 and 1989 (Bennett & Magnus, 1994), and from the 1989–1990 National Health Survey (Australian Bureau of Statistics, 1992). The method used for characterizing body fatness in these studies was Body Mass Index (BMI = weight in kilograms/height in meters^2). Although it is recognized that BMI may give an inaccurate indication of obesity in special groups, such as elite athletes or body builders (Garrow & Webster, 1985; Pi-Sunyer, 1988), it is nonetheless regarded as a practical and cost-effective measure of body fatness for population studies, as it correlates well with other more sophisticated measures (Garrow & Webster, 1985; Willett, 1990).

A classification system derived from research findings on the health risks associated with varying levels of BMI has been adopted by Australian health authorities (National Health and Medical Research Council, 1985), and has been used to classify adults in the population studies described below. Under this classification system, "underweight" is defined as a BMI of less than 20 kg/m^2; a BMI in the range of 20–25 kg/m^2 inclusive is classified as an "acceptable weight"; "overweight" is defined as a BMI of greater than 25 and up to and including 30 kg/m^2; and a BMI of greater than 30 kg/m^2 is defined as "obese." These categories apply to adults regardless of age (National Health and Medical Research Council, 1992). The use of a BMI of 25 kg/m^2 to define the upper limit of the healthy weight range, regardless of age, is consistent with recommendations from the United States (Dietary Guidelines Advisory Committee, 1995). In practice, these two terms are often used interchangeably, some writers favoring the generic term "overweight"; others (particularly in the USA) tend to favor use of the term "obese."

8.29.5 BEHAVIORAL EPIDEMIOLOGY

The most fundamental level of behavioral epidemiology is concerned with the identification of behaviors that are causally linked to disease. As stated above, epidemiological research has identified strong and probably causal associations between physical inactivity and obesity, and the risk of several chronic diseases and other health problems. One step removed from the examination of the causal relationships between behavior and disease is the epidemiological study of the behavior itself.

With regard to physical activity, for example, behavioral epidemiology concerns itself with the study of who is active, who is inactive, why people are inactive, and how we might help them become more active. For weight control also, similar issues apply. The focus is on describing and understanding the distribution and determinants of the behavior, and not the disease itself (Raymond, 1989).

Large-scale public health initiatives to influence health-related behaviors rely heavily on the availability of valid and reliable data on the population prevalence, trends, and sociodemographic distribution and determinants of the behaviors about which we are concerned. This is the level of descriptive epidemiology, within which the aim is to be able to identify whether or not the behavior of concern is widespread in the population and the extent to which the prevalence of this behavior does constitute a substantial risk for the whole population (Mason & Powell, 1985).

8.29.5.1 The Descriptive Epidemiology of Physical Activity

Population prevalence data on physical activity participation are available from a number of developed countries, but until recently have tended to focus mainly on the 15–20% of the adult populations of the USA and Canada who exercise at a level sufficient strongly to enhance cardiorespiratory fitness (Caspersen, Merritt, & Stephens, 1994). However, it is more useful to focus on the physically inactive, given that the public health challenge is to activate the sedentary, rather than encourage those who are already active to do more. Australian studies (Bauman, Owen, & Rushworth, 1990) have found about one-third of adults to be sedentary in their leisure time, which is substantially similar to what has been found in North American surveys (Caspersen et al., 1994).

Representative population data are helpful for identifying which groups should be targeted, what the content of an intervention should be, or what the characteristics are of those likely to succeed or fail in continuing to be physically active. Owen and Bauman (1992) found that 30% of adults who could be classified as sedentary in their leisure time were more likely to be older, less well-educated, and to have lower incomes. The most commonly reported reasons for inactivity were as follows: "no time" by 35%, "physically unable" by 24%, "don't want to exercise" by 13%. These reasons were more likely to be reported by those who were older, and who had lower income. Women were more likely to report being physically unable to exercise. Those on lower incomes were more likely to report being physically unable and not wanting to exercise.

These Australian findings are likely to have relevance to other countries which are broadly similar in their economic and social characteristics (Owen, 1994). Women, older people, and the socially disadvantaged may benefit most from less strenuous forms of physical activity such as walking (to overcome "physically unable"). Some people perceive time pressure as a central reason for inactivity: these may include those with less discretionary time in which to exercise, such as blue-collar workers or adult members of families with several children. Strategies to increase the convenience of physical activity (such as worksite programs or community facilities providing child care), may help such people to integrate physical activity more readily into the demands of their everyday life. Low-cost and easily accessible activities such as walking may have a greater chance of being taken up by those who are physically inactive (Morris, 1990; Owen & Lee, 1989).

8.29.5.2 Environmental and Public Policy Approaches: Physical Activity Promotion

Stokols' (1992) ecological model has provided some useful insights into the options for a more comprehensive approach to large-scale health behavior change. Stokols makes the important point that most health promotion programs implemented in community, corporate, and clinical settings have been designed primarily to influence change in individuals rather than the environment. It has become increasingly clear that the wider environment and the context in which change is being addressed are also very important. It is important to understand the many social, cultural, and attitudinal factors that are determinants of some of the major behavioral risk factors.

Health psychologists have traditionally given considerable attention to models of individual health behavior such as the Health Belief Model and the Theory of Reasoned Action (see Glanz, Lewis, & Rimer, 1990). More recently, broader models of interpersonal behavior and social influence incorporating social learning theory (Bandura, 1986; Ewart, 1990) have been developed, but less attention has been given to models that attempt to understand health behavior change within whole communities (Glanz et al., 1990).

Many of the elements of a broader social-ecological approach may be seen in public health policy initiatives in Australia, which largely mirror approaches being taken up in the UK (Killoran, Fentem, & Caspersen, 1994) and the USA (US Department of Health and Human Services, 1996). A focus on environmental and social determinants of physical activity is strongly integrated within the Australian National Health Goals and Targets for Cardiovascular Disease (Commonwealth Department of Human Services and Health, 1994; Owen & Lowe, 1994).

The national target for Australia is a reduction in the proportion of adult Australians who are very inactive in their leisure time from 36% (at the time of the 1989 National Health Survey) to 25% by the year 2000. The Australian Goals and Targets identify strategies for reducing the proportion of the Australian population who are sedentary in their leisure time:

(i) Governments should develop strategies to encourage employers to offer their staff physical activity and fitness programs or incentives to participate in outside programs.

(ii) States and territories should work with the Australian Council for Health, Physical Exercise and Recreation (ACHPER) and the National Heart Foundation (NHF) to ensure that school curricula involve the regular participation of all children and adolescents in primary and high school in physical education programs several times a week.

(iii) States and territories should work with ACHPER and the NHF to ensure that school physical education programs consist of physical activities which may be continued and maintained through later life.

(iv) Tertiary education facilities should be encouraged to develop programs which increase the proportion of students who participate in sport or active recreation.

(v) Governments and nongovernment organizations such as the NHF should develop programs to further reinforce the convenience of regular, moderate physical activity for the maintenance of cardiovascular health.

(vi) States and local governments should be encouraged to include walking and cycle paths and low-traffic areas as part of community recreation and transport planning activities. The NHF should continue to work with local governments to increase community exercise facilities and programs.

(vii) States and local governments should investigate structural changes which will encourage people to walk or cycle instead of using motor transport wherever possible. Possibilities are car-free areas in cities (except for the disabled or elderly), and strategies to encourage cycling instead of driving.

To design programs to reach populations which is, after all, a fundamental aim of the public health approach, requires an understanding of how social systems operate, how change occurs within and among systems, and how large-scale environmental and public policy changes influence health behaviors (Sallis & Owen, 1996; Stokols, 1992). Diffusion of innovations theory and a number of other models address in detail the understanding of ways to translate effective interventions into more widespread change, but to date there has been relatively little research which has addressed these issues comprehensively, other than in a few key settings such as schools and the workplace (US Department of Health and Human Services, 1996, Chapter 6).

The example of the Australian national physical activity strategies show the extent to which an integrated and balanced approach to the promotion of physical activity must focus not only on mass-media awareness campaigns, but also on the provision of appropriate settings, facilities and resources that provide accessible and attractive opportunities for people to be more physically active (Donovan & Owen, 1994). What is emphasized in these strategies is that substantial and permanent opportunities to be more active in local communities need to be developed (King, 1991).

8.29.5.3 The Descriptive Epidemiology of Overweight and Obesity in Australia

While the National Health Survey provides extensive information on the health of Australians, it did not include any physical or biomedical measurements (Australian Bureau of Statistics, 1992). Participants in the survey were interviewed and were asked to report their height and weight. These self-reported heights and weights were used to calculate BMI for persons aged 18 years and over, and to classify them into one of the four BMI categories described above. There are, however, problems associated with the validity of such self-report data, which are likely to result in the misclassification of subjects into BMI categories (Nieto-Garcia, Bush, & Keyl, 1990). In spite of the limitations of the National Health Survey data, they have been used as the baseline for Australia's National Health Goals and Targets for overweight (Commonwealth Department of Human Services and Health, 1994), and they will be used to monitor trends in BMI and to evaluate the effectiveness of future initiatives.

The National Health Survey data suggest that, based on BMI, only half of the adult population have an acceptable weight. The proportion of men classified as overweight or obese (defined as a BMI $> 25\,\mathrm{kg/m^2}$) increased until the age of 55, rising from 26% of men aged 18–24 to 55% of the men aged 45–54, before falling in the older age groups. Among women, the prevalence of overweight or obesity increased from 15% in the youngest age group, to 45% in women aged 55–64, before beginning to fall. The prevalence of underweight was highest for men aged under 25 and for men 75 years and over, while for women underweight decreased as age increased, until the age of 55. Almost three times as many women as men were underweight, and slightly more women than men were in the obese category. However, men were more likely than women to be classified as overweight, and overall more men had a BMI greater than that which is currently considered acceptable.

During the 1980s the National Heart Foundation of Australia conducted three independent cross-sectional risk factor prevalence surveys, each of which employed essentially the same methods (Bennett & Magnus, 1994). In these surveys, a systematic probability sample of adults was selected from the federal electoral rolls of major urban centers (registration on the electoral rolls is compulsory for all citizens aged 18 years and over), and those selected were invited to attend a free cardiovascular risk factor screening at a local survey center. The data from the three Heart Foundation surveys are based on objectively measured height and weight.

The Heart Foundation data show that only about a half of the adults surveyed in 1989 were an "acceptable" weight (National Heart Foundation of Australia, 1990), that is, had a BMI which falls within the range that is currently considered to be associated with lowest risk of mortality and morbidity. Although obesity was also more common among women than men, a greater proportion of the men were overweight or obese, with almost one in two men affected, compared to approximately one in three women. The prevalence of overweight or obesity was found to increase from 25% in men aged 20–24, leveling out at approximately 60% in men aged 45 years or more. In women, the proportion who were overweight or obese increased from 17% in the youngest group, to slightly less than 60% of women aged 65–69.

The prevalence of overweight or obesity also varied by other sociodemographic characteristics. An analysis of the 1989 Heart Foundation survey data revealed that the odds of being overweight or obese were higher among persons born in Southern Europe compared to Australia; for persons with a low level of education compared to those with tertiary qualification; for married men compared to men who have never married; for plant/machinery operators compared to professional men; for laborers and related workers compared to professional women; and for women in the two lowest socioeconomic categories compared to those in the highest category (Waters & Bennett, 1995). The differences by social status are generally consistent with those reported for other industrialized countries (Sobal & Stunkard, 1989).

The Australian Risk Factor Prevalence Study has been analyzed to examine trends in BMI in the adult population over time (Bennett & Magnus, 1994). These analyses showed there has been an increase in the average weight of adults, with no significant changes in height during the 1980s. In 1989, women were an average of three kilograms heavier than their counterparts in 1980, while among men average weight increased by just over one and a half kilograms. Although these increases in weight do not appear dramatic, they in fact translate to a 23% increase in the proportion of men overweight or obese over the period 1980–1989, and a 58% increase in the proportion of women who carry excess weight. Other data derived from the 1989–1990 National Health Survey and the 1994 Australian Bureau of Statistics Population Survey Monitor, although relying on self-report height and weight data, suggest that among men and women aged 25–64 years the prevalence of overweight and obesity has continued to increase (Abraham, d'Espaignet, & Stevenson, 1995).

While definitions of obesity vary from country to country, the trend of increasing average weights is not unique to Australia, having also been observed in other industrialized countries, including the US (Kuczmarski, Flegal, Campbell, & Johnson, 1994), the UK (Gregory, Foster, & Tyler, 1990), Sweden (Kuskowska-Wolk & Bergstrom, 1993a; Kuskowska-Wolk & Bergstrom, 1993b), and Germany (Hoffmeister, Mensink, & Stolzenberg, 1994).

8.29.5.4 Environmental and Public Policy Approaches: Dealing with Overweight and Obesity

Many of the same ecological frameworks and theoretical models and issues described above in relation to physical activity apply to addressing overweight and obesity. For many years Australian health authorities and other health and professional organizations, like those in other developed countries, have promoted the virtue of an acceptable weight, focusing the benefits of weight loss. In addition to these activities and initiatives, commercial weight-loss centers, special diet foods, and diet plans proliferate, and information on weight control is common in the print media. Despite the activities of public- and private-sector organizations over many years, excess body weight is more common in the late 1990s than it was at the beginning of the 1980s.

In response to the high prevalence of overweight and obesity, and the disturbing trend of increasing average weights. Australian health authorities, like those of other countries (Health and Welfare Canada, 1988; US Department of Health and Human Services, 1991), have also identified health goals relating to overweight and obesity (Table 2). These targets are based on a recognition that the problem is a population phenomenon, and that a small reduction in the population's mean weight can lead to significant change in the prevalence of overweight (Rose, 1992).

While it is acknowledged that there is a need to promote behavioral and environmental change to improve dietary habits and increase levels of physical activity in the population, a comprehensive strategy to deal with overweight is only now, in the late 1990s, being developed in Australia. Given this, the Australian year 2000 targets for reducing overall levels of overweight and obesity must be viewed as optimistic, especially in light of the recent trends of increasing average weight.

8.29.6 UNDERSTANDING AND INFLUENCING HEALTH BEHAVIORS: THE EXAMPLE OF THE TRANSTHEORETICAL MODEL

There is a need to develop and apply appropriate theory and to generate innovative ways of delivering health promotion programs to larger numbers of individuals, not only through traditional clinical settings, but through the media, the workplace, and elsewhere in the community (Winett et al., 1989). A theoretical framework which now has considerable currency in public health analyses of health risk behaviors is the transtheoretical model (TTM). While the research supporting this model originated in clinical psychology (Prochaska, DiClemente, & Norcross, 1992), the TTM has been applied to health-related behavior change in large-scale studies and in a range of populations and settings (Prochaska et al., 1994).

The TTM is not the only model from health psychology with obvious relevance to understanding and influencing the practice of health promotion (see Chapter 8.01, this volume). Other theories give very helpful accounts of key issues such as the importance of observational learning, or the role of risk perceptions, health beliefs, and attitudes in health behavior (see

Table 2 Australia's health goals and targets relating to overweight.

Health goals for adults relating to overweight	*Year 2000 target* (%)	*Baseline level* (%)
Increase proportion with an acceptable body weight		
Men	60	48.8[b]
Women	60	51.0[b]
Reduce the prevalence of overweight[a]		
Men	40	44.2[b]
Women	25	30.6[b]
Reduce the prevalence of abdominal obesity		
Men	38	41.5[c]
Women	18	22.9[c]

Source: Adapted from Commonwealth Department of Community Services and Health (1994).
[a] Although this goal uses the term "overweight" it actually refers to the combined overweight and obese BMI categories. [b] Based on self-report data from the 1989–1990 National Health Survey. [c] Based on the 1989 National Heart Foundation Risk Factor Prevalence Study which included objective waist and hip measures.

Glanz et al., 1990). Here, we use the TTM as an example of a theoretical model which has performed well in the physical activity area, and which shows promise in understanding weight control.

The TTM proposes that there is a series of specific stages in the acquisition and maintenance of a behavior (DiClemente et al., 1991; Prochaska & DiClemente, 1983; see also Chapter 8.01 this volume). The model focuses on both current behavior and future behavioral intentions, and provides a framework that reflects motivational readiness for behavior change (Abrams, 1993; Marcus, Eaton, Rossi, & Harlow, 1994). Within this framework, behavior change is not considered to be an all-or-nothing phenomenon change, or a state of action vs. inaction (Abrams, 1993). This dynamic quality is important from a public health perspective, since it allows the targeting of specific segments of the population, the aim being to accelerate movement through the stages of behavior change toward action.

The five stages of behavior change (DiClemente et al., 1991) identified under the TTM are:

(i) precontemplation, in which change is not being considered;

(ii) contemplation, in which change is considered;

(iii) preparation, in which an individual takes small steps towards change;

(iv) action, involving the initial modification of the habit; and

(v) maintenance, when the change is sustained over six months or more.

The concept that there are discrete identifiable stages of change, reflecting varying degrees of readiness to modify behavior, has been proposed by others (e.g., Brownell, Marlatt, Lichtenstein, & Wilson, 1986). However, a distinctive feature of the TTM is its recognition of the importance of a broad set of stages, including the early stages of change (Rossi, Rossi, Velicer, & Prochaska, 1995).

Within the TTM framework, the change process is viewed as a continuum, with individuals leaving and re-entering at different points. Relapse (Marlatt, 1985; Brownell et al., 1986) should be considered a key element of any stage-base model of behavioral change. In relapse, individuals may lose resolve and abandon the new behavior pattern, feel guilty, lose confidence and revert to their previous habits. This terminates the action or maintenance stages, often prompting recycling through the earlier stages of precontemplation or contemplation.

The TTM was initially applied to smoking cessation (DiClemente & Prochaska, 1982), and has been widely used in that domain (Biener & Abrams, 1991, Owen, Wakefield, Roberts, & Esterman, 1992; Prochaska, Velicer, DiClemente, & Fava, 1988; Velicer, DiClemente, Prochaska, & Brandenburg, 1985). However, the TTM has also been employed to examine a range of other health behaviors such as quitting cocaine, safe sex, sun screen use, mammography screening, and delinquent behavior. These include behaviors that are addictive and non-addictive, that differ in terms of their frequency of occurrence, that are legal and illegal, socially acceptable and unacceptable, and that involve the cessation of negative behaviors as well as the acquisition of positive behaviors (Prochaska et al., 1994) Although few studies have used this model to explain weight control behaviors (discussed in Section 8.29.6.2), it has been extensively used to examine physical activity and dietary behaviors.

As well as identifying those who are ready to modify behavior, a key feature of the TTM is its emphasis on examining transitions between the various stages, by integrating components of other theories. While the TTM is one of a number of theoretical models that can be used to explain and predict health behavior change, it now has broad currency in the health promotion field, not only because of the empirical research supporting it, but also because of its practical face validity and appeal to practitioners and policymakers. Applications of the TTM to a wide range of behavioral domains are described by Prochaska et al. (1994).

Two of the key theoretical constructs within the TTM are of self-efficacy and decisional balance. According to self-efficacy theory, confidence in the ability to perform a particular behavior is strongly associated with the actual ability to perform that behavior (Bandura, 1977). Self-efficacy beliefs have been shown to be associated with the performance of a diverse range of health-related behaviors (O'Leary, 1985; Strecher, McEvoy-DeVellis, Becker, & Rosenstock, 1986). The notion of decisional balance is based on the conflict theory of decision making (Janis & Mann, 1977). Janis and Mann argued that, in making a sound decision, all relevant considerations are entered into a decisional "balance sheet" where the potential gains (pros) and losses (cons) resulting from the behavior to the individual and to others are evaluated, along with other factors such as self-esteem and the approval of others.

A number of studies of health-related behaviors have shown that, compared to those in the later stages, individuals in the early stages of change had little confidence in their ability to exercise or to change their diets (Glanz et al., 1994; Marcus & Owen, 1992; Marcus, Selby, Niaura, & Rossi, 1992; Sporny & Contento, 1995). Other data show that in moving from the precontemplation stage to the action stage, the perceived benefits of the change (the pros) increase, the perceived costs of the change (the cons) decrease, and overall decisional balance (pros–cons) increases (Marcus et al., 1994; Marcus & Owen, 1992; Marcus, Rakowski, & Rossi, 1992). These relationships between self-efficacy and decisional balance with stage of change for physical activity and dietary behaviors are generally consistent with those observed for other behaviors (DiClemente, Prochaska, & Gibertini, 1985; Prochaska et al., 1994; Schorling, 1995).

The following sections deal with the research findings and practical applications relating to stage-based approaches to physical activity and weight control. As will become apparent, the TTM has been applied more extensively in physical activity research and related practical initiatives than has so far been the case for weight control. These differences illustrate the scope that still exists for unique applications of health psychology theory and research methods within the health promotion field.

8.29.6.1 Applying the Transtheoretical Model to Physical Activity Promotion

Theory from health psychology may be used to provide a more in-depth understanding of physical activity or any other health-related behavior than is possible when only descriptive epidemiology studies like those described above are employed. For example, Australian studies have used stage-of-change measures from the TTM in behavioral epidemiology studies with representative population samples, and have found that about half of the sedentary Australian adults interviewed did not have being more active on their agenda (Booth, Bauman, Oldenburg, Owen, & Magnus, 1992; Booth et al., 1993).

The TTM's five main stages may be used to identify what is likely to be involved in taking up a more physically active lifestyle:

(i) In the precontemplation stage, individuals are not considering change; they state that they do not exercise and that they do not wish to do so.

(ii) In the contemplation stage, individuals are evaluating the benefits of being more active, and perhaps intend to change, but have not acted on this intention; for example, they have decided to lose weight by being more active but have not begun to do so.

(iii) In the preparation stage, individuals are taking concrete steps to become more active, for example joining a community activity centre, buying clothing and equipment or making commitments to be involved in active leisure pursuits with friends.

(iv) In the action stage, individuals have initiated changes, are consolidating those changes, and integrating them into their lifestyles.

(v) In the maintenance stage, a pattern of regular physical activity has been consolidated for at least six months, and individuals are persisting with a relatively stable new pattern of behavior change.

Assuming that these different stages are involved in the acquisition of physical activity habits, it follows that characteristics associated with the adoption of exercise may not be the same as those that relate to the successful maintenance of behavior change. For example, Marcus and Owen (1992) and Marcus, Rakowski, and Rossi (1992) examined the relationship of self-efficacy, a key component of Bandura's social-cognitive theory of behavior change (Bandura, 1986), with different stages of exercise adoption and maintenance. Self-efficacy reliably differentiated between the different stages. There was a significant trend of increased self-efficacy for participants in the different stages of the exercise-adoption sequence from precontemplation to maintenance.

Studies of the impact of large scale public health campaigns to influence exercise participation have used both descriptive behavioral epidemiology data and theoretically derived measures. Some informative findings come from studies of nationwide physical activity campaigns that were conducted by the National Heart Foundation of Australia in 1990 and 1991. The campaigns were informed by social, learning and social marketing models and each emphasized walking as the main activity, and also focused on the precursors and early stages of physical activity adoption, drawing on the Transtheoretical Model's stages of behavioral change (Booth et al., 1993; Prochaska et al., 1992).

The first campaign focused on the adoption of physical activity (in the framework of the TTM, shifting precontemplators and those in contemplation to taking action) and used the slogan "Exercise: make it part of your day" (Booth et al., 1992). The second campaign focused on maintenance (in the framework of the TTM, influencing those in preparation or action to persist with activity in the longer term) and used the slogan "Exercise: take another step" (Owen, Bauman, Booth, Oldenburg, & Magnus, 1995).

The two campaigns were promoted through paid television advertisements, prerecorded public service announcements for radio, the distribution of a professional paper, posters, leaflets and stickers, T-shirts, sweat shirts, publicity tours by two heart-health experts, a variety of magazine articles, and the scripting of one episode of each of two nationally broadcast television drama series. There was also opportunistic promotion through electronic and print media news coverage, editorials, and feature articles. In addition to national media-based strategies, local initiatives were pursued at the National Heart Foundation State Division level. These more community-based promotional activities emphasized special activity days and competitions.

These Australian campaigns provided an opportunity to examine systematically the influence of serial national mass-media campaigns directed at increasing levels of participation in physical activity. Face-to-face, home-based interviews with a representative national sample were carried out two weeks prior to and three to four weeks following each campaign. The evaluation of the 1990 campaign (Booth et al., 1992) found increases in the prevalence of reports of walking following the campaign. Changes were most marked in older people, and occurred across most socioeconomic groups. Awareness of the campaign message increased significantly from before the first campaign to after that campaign (46–71%). The level of physical activity–message awareness was still quite high prior to the 1991 campaign (63%), with postcampaign awareness increasing to 74%. Changes in reported walking for exercise and in readiness to exercise found in 1990 were not found in 1991 (Owen et al., 1995).

The measures of stages of change derived from the TTM were informative in interpreting elements of the evaluation. In 1990, the prevalence of reporting no intention to do more exercise showed the greatest pre- to postcampaign decreases, suggesting that the campaign may have been most effective in reaching those who were the least active and most resistant to change. The 1991 campaign message may have reinforced those who had previously adopted activity but failed to motivate change in others. Or most of those who were able to, or who intended to, change had already done so in response to the 1990 campaign.

8.29.6.2 Applying the Transtheoretical Model to Weight Control Behavior

In the weight control field, the research concerns of health psychologists and the current practical public health imperatives faced by practitioners and policy makers raise some issues in common with those for physical activity, but also different, and in some ways more basic, concerns. At the level of developing and using health psychology theory in the weight control field, applications of the TTM are promising, but at an earlier stage than is the case for physical activity.

An understanding of population weight-control behaviors is important because there is some evidence that, in trying to control their weights, individuals may put their health at risk (National Institutes of Health Technology Assessment Panel, 1992). The challenge facing health authorities in developed countries is to reduce the prevalence of overweight and obesity without exacerbating the problems of low body weight or unhealthy weight control. In order to achieve this, there is a growing recognition of the need to prevent weight gain in the population. In the United States, for example, health authorities have recommended that weight guidelines for the public be revised, so that there is an increased focus on weight maintenance, regardless of current weight, with less emphasis placed on weight loss. This shift in emphasis recognizes the importance of weight control as "the essential first step toward a reduction in the prevalence of obesity in the population" (Dietary Guidelines Advisory Committee, 1995, p. 23).

The use of the TTM in describing and understanding physical activity in a public health context has already been considered. Because these behaviors are directly relevant to the issue of weight control, the TTM is likely to be potentially useful in understanding population-wide weight-control behaviors, and its use may also facilitate the development of better targeted public health intervention strategies.

The few studies that have used the TTM to examine weight-related behaviors have shown that, as well as those actually taking action to lose weight, many individuals were trying to maintain their weight, and many were contemplating taking action. For example, in a group of college students only 18% were taking action for their weight; however, a further 15% were trying to maintain their weight, and just over half were contemplating weight loss (O'Connell & Velicer, 1988). Similarly, among participants in a smoking cessation study, 51% were taking action for their weight, with 27% attempting to maintain their weight, and another 19% contemplating weight loss (Ross et al., 1995). Consistent with the findings of previous research (Prochaska et al., 1994), the study of college students also found that the perceived costs (cons) of behavior change were significantly greater than the perceived benefits (pros) for those in the precontemplation stage, while individuals in the action stage showed the greatest imbalance in favor of the pros.

The TTM provides a theoretical framework within which to examine weight control using the broader definition proposed earlier, by encompassing both weight-loss and weight-maintenance behaviors (Rossi et al., 1995). However, the two studies which have applied the TTM were concerned with weight-loss behavior, not with weight-control behavior. In classifying the college students into one of the stages of change, for example, a minimum weight loss criterion of at least 10 pounds was set (O'Connell & Velicer, 1988). Therefore, individuals classified into the maintenance stage were those who had a minimum of 10 pounds weight loss in the past and were successfully taking action to avoid regaining the weight (but with no goal of further weight loss). Individuals who were actively trying to avoid weight gain, but who had not previously lost more than 10 pounds, were not classified into one of the stages of change.

Health authorities are now encouraging people to take steps to maintain their weight, even if they do not have a weight problem. Therefore, in considering weight-control behaviors from a public health perspective, it is important that we have data on population-wide weight-control behaviors, not just on weight loss among those who already have a weight problem. However, determining stage of change for weight control is potentially problematic. One key issue is that weight control is not a discrete behavior (Rossi et al., 1995). Unlike cigarette smoking, the health behavior to which the TTM was initially applied, weight control potentially involves a number of different strategies and it would not be feasible to examine each of the specific strategies employed.

Another potential problem arises from the fact that weight control includes weight loss and weight-maintenance behaviors, and can be practiced regardless of weight status. It would not be sensible from a public health perspective to use a single arbitrary weight standard to define stage of change for weight control and apply this to every individual. To do so would be to ignore that there are many individuals who are not overweight but who actively try to lose or maintain weight.

The public health challenge is to assess stage of change for the range of behaviors that are employed to control weight across the whole population. In this regard, the application of the TTM to weight control poses similar challenges to an examination of dietary behavior change. In that domain, classifying individuals in terms of their stage of change for each component of the diet is considered to be impractical, and there are also problems in trying to apply a strict criterion to define stage of change. Several of the studies that have examined dietary fat behavior have, therefore, used a "global" measure to define stage of change. For example, in a recent worksite study, individuals were asked to report whether they had deliberately changed the way they eat to reduce the amount of fat and fatty foods (Sporny & Contento, 1995). Subjects were classified into the precontemplation stage if they were not eating a reduced-fat diet and not considering doing so, contemplators were those who were not eating a reduced-fat diet but were considering doing so, subjects in action were those who reported eating a reduced-fat diet for less than two years, and those who had eaten such a diet for more than two years were considered to be in maintenance.

It may therefore be possible to determine stage of change for weight control using a similar measure. If so, this would allow researchers to distinguish those who are not taking any form of action to lose or maintain weight, those who are at least considering taking action, and those who are actively trying to lose weight or to maintain their weight. Such information is likely to be important in developing public health strategies to reduce the prevalence of overweight and obesity.

The TTM of behavior change is a framework through which more fine-grained analysis of the different stages in people's weight control efforts may be conducted. The TTM has been applied more extensively in physical activity research and related practical initiatives than has so far been the case for weight control. The more definitive application of such an understanding of weight-control behavior to public health intervention has yet to be developed and used in practice.

8.29.7 CONCLUSIONS

In health promotion research on physical activity and weight control, there remain gaps in the knowledge base needed to inform large-scale health promotion initiatives. Through behavioral epidemiology research, the links between health psychology theory and practical large-scale public health intervention strategies in these two areas will be strengthened into the twenty-first century. This will be the case not only for physical activity and weight control, but also for a number of the other rapidly developing areas of health promotion research and practice such as skin cancer prevention, and participation in mass screenings to detect treatable early signs or precursors of disease.

Health promotion research and practice in physical activity and weight control are relatively new areas of public health endeavor. There are potentially fruitful directions for research in making use of representative population data on physical activity and weight control and in understanding more clearly the stages of involvement in physical activity and weight control. Such research is needed to inform public health initiatives to encourage more physical activity and to prevent further weight gain in the populations of industrialized countries.

ACKNOWLEDGMENTS

We thank Jane Mudge for assistance in the preparation of this chapter.

8.29.8 REFERENCES

Abraham, B., d'Espaignet, E. T., & Stevenson, C. (1995). *Australian health trends 1995*. Canberra, Australia: AIHW and AGPS.

Abrams, D. B. (1993), Treatment issues: Towards a stepped-care model. *Tobacco Control, 2*(Suppl.), S17–S37.

Australian Bureau of Statistics. (1992). *1989–90 National Health Survey. Health Risk Factors—Australia. Catalogue No. 4380.0*. Canberra, Australia: Australian Bureau of Statistics.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior change. *Psychological Review, 84*(2), 191–215.

Bandura, A. (1986). *Social foundations of thought and action: A social-cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bauman, A., Owen, N., & Rushworth, R. L. (1990). Recent trends and socio-demographic determinants of exercise participation in Australia. *Community Health Studies, 14*, 19–26.

Bennett, S. A., & Magnus, P. (1994). Trends in cardiovascular risk factors in Australia. Results from the National Heart Foundation's Risk Factor Prevalence Study, 1980–1989. *Medical Journal of Australia, 161*, 519–527.

Best, J. A., & Proctor, S. (1988). Behavioral medicine training from a public health perspective. *Annals of Behavioral Medicine, 10,* 19–22.

Biener, L., & Abrams, D. B. (1991). The contemplation ladder: Validation of a measure of readiness to consider smoking cessation. *Health Psychology, 10*(5), 360–365.

Blair S. N., Kohl H. W., Gordon N. F., & Paffenbarger R. S. (1992). How much physical activity is good for health? *Annual Review of Public Health, 13,* 99–126.

Booth, M., Bauman, A., Oldenburg, B., Owen. N. & Magnus, P. (1992). Effects of a national mass-media campaign on physical activity participation. *Health Promotion International, 7,* 241–247.

Booth, M., Macaskill, P., Owen. N., Oldenburg, B., Marcus, B., & Bauman, A. (1993). The descriptive epidemiology of stages of change in physical activity. *Health Education Quarterly, 20,* 431–440.

Bouchard C., Shephard R. J., Stephens T., Sutton J. R., & McPherson B. D. (Eds.) (1990). *Exercise, fitness and health. A consensus of current knowledge*. Champaign, IL: Human Kinetics Press.

Brownell, K. D., Marlatt, G. A., Lichtenstein, E., & Wilson, G. T. (1986). Understanding and preventing relapse. *American Psychologist, 41*, 765–782.

Burton, B. T., Hirsch, F. J., & Van Itallie, T. B. (1985). Health implications of obesity: An NIH consensus development conference. *International Journal of Obesity, 9,* 155–169.

Caspersen C. J., Merritt R. K., & Stephens, T. (1994). International activity patterns: A methodological perspective. In R. K. Dishman (Ed.), *Advances in exercise adherence* (pp. 73–110). Champaign, IL: Human Kinetics Press.

Commonwealth Department of Human Services and Health. (1994). *Better health outcomes for Australians. National goals, targets and strategies for better health outcomes in the next century*. Canberra, Australia: Australian Government Publishing Service.

DiClemente, C. C., & Prochaska, J. O. (1982). Self-change and therapy change of smoking behavior: A comparison of processes of change in cessation and maintenance. *Addictive Behaviors*, *7*, 133–142.

DiClemente, C. C., Prochaska, J. O., Fairhurst, S. K., Velicer, W. F., Velasquez, M. M., & Rossi, J. S. (1991). The process of smoking cessation: An analysis of the precontemplation, contemplation and preparation stages of change. *Journal of Consulting and Clinical Psychology*, *59*(2), 295–304.

DiClemente, C. C., Prochaska, J. O., & Gibertini, M. (1985). Self-efficacy and the stages of self-change of smoking. *Cognitive Therapy and Research*, *9*(2), 181–200.

Dietary Guidelines Advisory Committee. (1995). *Report of the dietary guidelines advisory committee on the dietary guidelines for Americans, 1995*. Washington, DC: US Department of Agriculture.

DiPietro, L. (1996). The epidemiology of physical activity and physical function in older people. *Medicine and Science in Sports and Exercise*, *28*, 596–600.

Donovan, R. J., & Owen, N. (1994). Social marketing and population interventions. In R. K. Dishman (Ed.), *Advances in exercise adherence* (pp. 249–290). Champaign, IL: Human Kinetics Press.

Drinkwater, B. L. (1994). Does physical activity play a role in preventing osteoporosis? *Research Quarterly for Exercise and Sport*, *65*, 197–206.

Ewart, C. K. (1990). Social action theory for a public health psychology. *American Psychologist*, *46*, 931–946.

Flay, B. R. (1987). Mass media and smoking cessation: A critical review. *American Journal of Public Health*, *77*, 153–160.

Garrow, J. S., & Webster, J. (1985). Quetelet's Index (W/H^2) as a measure of fatness. *International Journal of Obesity*, *9*, 147–153.

Glanz, K, Lewis, F. M., & Rimer, B. K. (Eds.) (1990). *Health behavior and health education: Theory, research and practice*. San Francisco: Jossey-Bass.

Glanz, K., Patterson, R. E., Kristal, A. R., DiClemente, C. C., Heimendinger, J., & Linnan, L. (1994). Stages of change in adopting healthy diets: Fat, fiber and correlates of nutrient intake. *Health Education Research*, *21*, 499–519.

Gloag D. (1992). Exercise, fitness and health. *British Medical Journal*, *305*, 377–378.

Gochman, D. S. (1988). *Health behavior: Emerging perspectives*. New York: Plenum.

Green, L. W., & Kreuter, M. W. (1991). *Health promotion planning: An educational and environmental approach.* Mountain View, CA: Mayfield.

Gregory, J., Foster, K., & Tyler, H. (1990). *The dietary and nutritional survey of British adults. A survey of the dietary behaviour, nutritional status, and blood pressure of adults aged 16 to 64 living in Great Britain.* London: HMSO, Office of Population Censuses and Surveys, Social Survey Division.

Grilo, C. M. (1995). The role of physical activity in weight loss and weight loss management. *Medicine, Exercise, Nutrition, and Health*, *4*, 60–76.

Health and Welfare Canada. (1988). *Canadian guidelines for healthy weights. Report of an expert group convened by Health Promotion Directorate, Health Services and Promotion Branch*. Ottowa, Canada: Health Services and Promotion Branch, Health and Welfare Canada.

Hoffmeister, H., Mensink, G. B. M., & Stolzenberg, H. (1994). National trends in risk factors for cardiovascular disease in Germany. *Preventive Medicine, 23,* 197–205.

Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment*. New York: Free Press.

Jeffery, R. W. (1989). Risk behaviors and health: Contrasting individual and population perspectives. *American Psychologist, 44,* 1194–1202.

Kaplan, R. M. (1985). Behavioral epidemiology, health promotion, and health services. *Medical Care, 23,* 564–583.

Kaplan, R. M. (1990). Behavior as the central outcome in health care. *American Psychologist, 45,* 1211–1220.

Killoran, A., Fentem, P., & Caspersen, C. (Eds.) (1994). *Moving on international perspectives on promoting physical activity*. London: Health Education Authority.

King, A. C. (1991). Community intervention for promotion of physical activity and fitness. *Exercise and Sport Science Reviews, 19,* 211–259.

Kopelman, P. G., Finer, N., Fox, K. R., Hill, A., & MacDonald, L. A. (1994). ASO consensus statement on obesity. *International Journal of Obesity, 18,* 189–191.

Kuczmarski, R. J., Flegal, K. M., Campbell, S. M., & Johnson, C. L. (1994). Increasing prevalence of overweight among US adults. *Journal of the American Medical Association, 272*(3), 205–211.

Kuskowska-Wolk, A., & Bergstrom, R. (1993a). Trends in body mass index and prevalence of obesity in Swedish men 1980–89. *Journal of Epidemiology and Community Health, 47,* 103–108.

Kuskowska-Wolk, A., & Bergstrom, R. (1993b). Trends in body mass index and prevalence of obesity in Swedish women 1980–89. *Journal of Epidemiology and Community Health, 47,* 195–199.

Lee, I.-M. (1994). Physical activity, fitness, and cancer. In C. Bouchard, R. J. Shephard, & T. Stephens (Eds.), *Physical activity, fitness, and health* (pp. 814–831). Champaign, IL: Human Kinetics Press.

Lichtenstein, E., & Glasgow, R. E. (1992). Smoking cessation: What have we learned over the last decade. *Journal of Consulting and Clinical Psychology, 60,* 518–572.

Marcus, B. H., Eaton, C. A., Rossi, J. S., & Harlow, L. L. (1994). Self-efficacy, decision making, and stages of change: An integrative model of physical exercise. *Journal of Applied Social Psychology, 24,* 489–508.

Marcus, B. H., & Owen, N. (1992). Motivational readiness, self-efficacy and decision making for exercise. *Journal of Applied Social Psychology, 22,* 3–16.

Marcus, B. H., Rakowski, W., & Rossi, J. S. (1992). Assessing motivational readiness and decision making for exercise. *Health Psychology, 11,* 257–261.

Marcus, B. H., Selby, V. C., Niaura, R. S., & Rossi, J. S. (1992). Self-efficacy and the stages of exercise behavior change. *Research Quarterly for Exercise and Sport, 63,* 60–66.

Marlatt, G. A. (Eds.) (1985). *Relapse prevention*. Everett, WA: Audio Productions.

Martinsen, E. W., & Stephens, T. (1994). Exercise and mental health in clinical and free-living populations. In R. K. Dishman (Ed.), *Advances in exercise adherence* (pp. 55–72). Champaign, IL: Human Kinetics Press.

Mason, J. O., & Powell, K. E. (1985). Physical activity, behavioral epidemiology, and public health. *Public Health Reports, 100,* 113–115.

Morris J. N. (1990). Inequalities in health: Ten years and a little further on. *The Lancet, 336,* 491–493.

National Health and Medical Research Council. (1985). *Report of the one hundredth session*. Canberra, Australia: Australian Government Publishing Service.

National Health and Medical Research Council. (1992). *Dietary guidelines for Australians*. Canberra, Australia: Australian Government Publishing Service.

National Heart Foundation of Australia. (1990). *Risk factor prevalence study. Survey No. 3—1989*. Canberra, Australia: National Heart Foundation of Australia and Australian Institute of Health and Welfare.

National Institutes of Health Technology Assessment Panel. (1992). Methods for voluntary weight loss and control: Technology Assessment Conference statement. *Annals of Internal Medicine, 116,* 942–949.

New Zealand Department of Health. (1991).

Nieto-Garcia, F. J., Bush, T. L., & Keyl, P. M. (1990). Body mass definitions of obesity: Sensitivity and specificity using self-reported weight and height. *Epidemiology, 1,* 146–152.

O'Connell, D., & Velicer, W. F. (1988). A decisional balance measure for the Stages of Change Model for weight loss. *The International Journal of the Addictions, 23*(7), 729–750.

O'Leary, A. (1985). Self-efficacy and health. *Behaviour Reseach and Therapy, 23,* 437–451.

Oldenburg, B., & Owen, N. (1990). Health psychology in Australia. *Psychology & Health, 4,* 73–81.

Owen, N. (1994). Shaping public policies and programs to promote physical activity. In A. Killoran, P. Fentem, & C. Caspersen (Eds.), *International perspectives on promoting physical activity* (pp. 194–212) London: Health Education Authority.

Owen, N., & Bauman, A. (1992). The descriptive epidemiology of sedentary lifestyle in adult Australians. *International Journal of Epidemiology, 21,* 305–310.

Owen, N., Bauman, A., Booth, M., Oldenburg, B., & Magnus, P. (1995). Serial mass-media campaigns to promote physical activity: Reinforcing or redundant? *American Journal of Public Health, 85,* 244–248.

Owen, N., & Lee, C. (1986). Towards more rigorous evaluation of health promotion programmes. *Australian Psychologist, 21,* 79–91.

Owen, N., & Lee, C. (1989). Development of behaviorally-based policy guidelines for the promotion of exercise. *Journal of Public Health Policy, 10,* 43–61.

Owen, N., & Lowe, J. B. (1994). A perspective on the National Health Goals and Targets Cardiovascular Disease and Cancer Implementation Working Groups. *Australian Journal of Health Promotion, 4,* 24–27

Owen, N., Wakefield, M., Roberts, L., & Esterman, A. (1992). Stages of readiness to quit smoking: Population prevalence and correlates. *Health Psychology, 11,* 413–417.

Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C., Buchner, D., Ettinger, W., Heath, G. W., King, A. C., Kriska, A., Leon, A. S., Marcus, B. H., Morris, J., Paffenbarger, R. S., Patrick, K., Pollock, M. L., Rippe, J. M., Sallis, J., & Wilmore, J. H. (1995). Physical activity and public health: A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *Journal of the American Medical Association, 273,* 402–407.

Pi-Sunyer, F. X. (1988). Obesity. In M. E. Shils & V. R. Young (Eds.), *Modern nutrition in health and disease* (7th ed., pp. 795–816). Philadelphia: Lea & Febiger.

Pi-Sunyer, F. X. (1991). Health implications of obesity. *American Journal of Clinical Nutrition, 53,* 1595S–1603S.

Powell, K. E., & Blair, S. N. (1994). The public health burdens of sedentary living habits: Theoretical but realistic estimates. *Medicine and Science in Sports and Exercise, 26*, 851–856.

Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology, 51*, 390–395.

Prochaska, J. O., DiClemente, C. E., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist, 47*, 1102–1114.

Prochaska, J. O., Velicer, W. F., DiClemente, C. C., & Fava, J. (1988). Measuring processes of change: Applications to the cessation of smoking. *Journal of Consulting and Clinical Psychology, 56*, 520–528.

Prochaska, J. O., Velicer, W. F., Rossi, J. S., Goldstein, M. G., Marcus, B. H., Rakowski, W., Fiore, C., Harlow, L. L., Redding, C. A., Rosenbloom, D., & Rossi, S. R. (1994). Stages of change and decisional balance for 12 problem behaviors. *Health Psychology, 13*(1), 39–46.

Raymond, J. S. (1989). Behavioral epidemiology: The science of health promotion. *Health Promotion International, 4,* 281–286.

Rose, G. (1992). *The strategy of preventive medicine*. Oxford, UK: Oxford University Press.

Rossi, J. S., Rossi, S. R., Velicer, W. F., & Prochaska, J. O. (1995). Motivational readiness to control weight. In D. B. Allison (Ed.), *Handbook of assessment methods for eating behaviors and weight-related problems: Measures, theory, and research* (pp. 387–430). Thousand Oaks, CA: Sage.

Royal College of Physicians. (1983). Obesity: A report of the Royal College of Physicians. *Journal of the Royal College of Physicians, 17*(1), 6–65.

Sallis, J. F., & Owen, N. (1996). Ecological models. In K. Glanz, F. M. Lewis, & B. K. Rimer (Eds.), *Health behavior and health education: Theory, research, and practice* (2nd ed., pp. 403–424) San Francisco: Jossey-Bass.

Schorling, J. B. (1995). The stages of change of rural African–American smokers. *American Journal of Preventive Medicine, 11*, 170–177.

Skrabanek, P. (1986). Preventive medicine and morality. *Lancet, i*, 143–144.

Sobal, J., & Stunkard, A. J. (1989). Socioeconomic status and obesity: A review of the literature. *Psychological Bulletin, 105,* 260–275.

Sporny, L. A., & Contento, I. R. (1995). Stages of change in dietary fat reduction: Social psychological correlates. *Journal of Nutrition Education, 27,* 191–199.

Stokols, D. (1992) Establishing and maintaining healthy environments: Toward a social ecology of health promotion. *American Psychologist, 47,* 6–22.

Strecher, V. J., McEvoy-DeVellis, B., Becker, M. H., & Rosenstock, I. M. (1986). The role of self-efficacy in achieving health behavior change. *Health Education Quarterly, 13*(1), 73–91.

US Department of Health and Human Services. (1996) *The Surgeon General's report on physical activity and health.* Washington, DC: US Government Printing Office.

US Department of Health and Human Services. (1991). *Healthy people 2000. National health promotion and disease prevention objectives*. Washington DC: Author.

Velicer, W. F., DiClemente, C. C., Prochaska, J. O., & Brandenburg, N. (1985). A decisional balance measure for assessing and predicting smoking status. *Journal of Personality and Social Psychology, 48,* 1279–1289.

Waters, A. M., & Bennett, S. (1995). *Risk factors for cardiovascular disease: A summary of Australian data. (Cardiovascular Disease Series; no. 1).* Canberra, Australia: Australian Institute of Health and Welfare.

Willett, W. (1990). *Nutritional Epidemiology*. (Vol. 15). New York: Oxford University Press.

Winett, R. A., King, A. C., & Altman, D. G. (1989). *Health psychology and public health: An integrative approach.* New York: Pergamon.

Young D. R., Haskell W. L., Jatulius D. E., & Fortmann S. P. (1993). Associations between changes in physical activity and risk factors for coronary heart disease in a community-based sample of men and women: The Stanford Five-City Project. *American Journal of Epidemiology, 138,* 205–216.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.01
Mental Retardation

SARAH W. BISCONER
Eastern State Hospital, Williamsburg, VA, USA
and
JANENE N. SUTTIE
University of Hawaii at Manoa, Honolulu, HI, USA

9.01.1 INTRODUCTION 1
9.01.2 PHENOMENOLOGY 2
9.01.3 EPIDEMIOLOGY 3
9.01.4 ETIOLOGY 4
9.01.5 ASSESSMENT AND DIAGNOSIS 4
  9.01.5.1 Intelligence 4
  9.01.5.2 Adaptive Behavior 6
  9.01.5.3 Psychopathology 7
9.01.6 TREATMENT 8
  9.01.6.1 Behavioral Interventions 8
  9.01.6.2 Psychopharmacology 10
9.01.7 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 11
9.01.8 REFERENCES 16

9.01.1 INTRODUCTION

Mental retardation is a developmental disability characterized by subaverage intelligence and impairments in adaptive daily life skills. The disorder occurs before the age of 18 years, and is associated with numerous biological, environmental, and sociocultural causes. The concept of mental retardation and the language used to refer to mental retardation varies from culture to culture; therefore, policies and practices related to service eligibility, assessment and treatment modalities, and preventive measures tend to vary from culture to culture. Further, different terms may be used in some cultures to describe mental retardation. For example, in Australia, the term intellectual disability is used, and in the UK, the term learning disability is used.

Definitions of mental retardation are based on formulations developed by international leaders in the field (e.g., American Association on Mental Retardation (AAMR), American Psychiatric Association). According to the AAMR definition,

> *Mental retardation* refers to substantial limitations in present functioning. It is characterized by significantly subaverage intellectual functioning, existing concurrently with related limitations in two or more of the following applicable adaptive skill areas: communication, self-care, home living, social skills, community use, self-direction, health and safety, functional academics, leisure, and

work. Mental retardation manifests before age 18. (Luckasson et al., 1992, p. 1)

In addition to the definition, four assumptions are considered essential to its application (Luckasson et al., 1992, p. 1):

(i) valid assessment considers cultural and linguistic diversity as well as differences in communication and behavioral factors;

(ii) the existence of limitations in adaptive skills occurs within the context of community environments typical of the individual's age peers and is indexed to the person's individualized needs for supports;

(iii) specific adaptive limitations often coexist with strengths in other adaptive skills or other personal capabilities; and

(iv) with appropriate supports over a sustained period, the life functioning of the person with mental retardation will generally improve.

The AAMR classification system utilizes four levels of habilitation or support that parallels the individual's limitations (intermittent, limited, extensive, and pervasive). Intermittent supports typically are provided on a short-term, as-needed basis to facilitate life transitions (e.g., job loss, medical crisis). Limited supports typically are provided over longer periods of time in comparison to intermittent supports, but are still considered time limited (e.g., employment training, mobility training). Extensive supports are characterized by regular, long-term involvement in one or more life areas (e.g., long-term residential care, supported employment). Pervasive supports are characterized as constant, intensive, and often life sustaining supports provided long term across all life environments.

The current *Diagnostic and statistical manual of mental disorders* (*DSM-IV*; American Psychiatric Association, 1994) defines mental retardation as follows:

(i) significantly subaverage intellectual functioning: an IQ of approximately 70 or below on an individually administered IQ test (for infants, a clinical judgment of significantly subaverage intellectual functioning);

(ii) concurrent deficits or impairments in present adaptive functioning (i.e., the person's effectiveness in meeting the standards expected for his or her age by his or her cultural group) in at least two of the following areas: communication, self-care, home living, social/interpersonal skills, use of community resources, self-direction, functional academic skills, work, leisure, health and safety); and

(iii) onset before the age of 18 (p. 46).

The *DSM-IV* definition utilizes four degrees of severity that reflect the level of intellectual impairment: IQ levels between 50–55 to approximately 70 characterize mild mental retardation, 35–40 to 50–55 characterize moderate mental retardation, 20–25 to 35–40 characterize severe mental retardation, and IQ levels below 20–25 characterize profound mental retardation.

The AAMR classification and *DSM-IV* diagnostic systems include the same three criteria for diagnosis: significantly subaverage intellectual functioning, concurrent limitations in adaptive skills, and onset prior to age 18 years. Both systems operationally define "significantly subaverage intellectual functioning" as a cut-off IQ score approximately two standard deviation units below a mean of 100 (between 70 and 75 or below), and "concurrent limitations in adaptive skills" as two or more deficits in functional skill areas necessary to meet life's demands, as defined by the individual's age group, community setting, and sociocultural background. The two classification systems differ in that AAMR utilizes four levels of habilitation or support, which are not directly comparable to the levels of severity in *DSM-IV* (American Psychiatric Association, 1994).

9.01.2 PHENOMENOLOGY

Persons with mental retardation have been noted in the literature from early Greek and Roman times, and have been labeled with a variety of descriptors: innocents, incapables, idiots, morons, cretins, dunces, and imbeciles. The term "feeble mindedness" was used as a generic term in the UK and the USA during the eighteenth and nineteenth centuries to describe a broad range of persons unable, for one reason or another, to function according to the perceived expectancies of normal participation in society (Scheerenberger, 1983). In 1910, the Committee on Classification of Feeble-Minded (under the auspices of the American Association on Mental Deficiency), defined three levels of feeble-mindedness: idiots, or those whose intellectual development did not exceed that of a normal two-year-old child; imbeciles, or those whose intellectual development did not exceed that of a normal 7-year-old child; and morons, or those whose intellectual development did not exceed that of a normal 12-year-old child (Scheerenberger, 1983).

Internationally, general descriptors have varied. Scandinavia and Russia have used the term oligophrenia; the UK has used the term learning disabilities; The Netherlands and Hong Kong have used the term mental handicap; Australia and Singapore have used the term intellectual disability; and the USA and Japan have used the term mental retardation. Until the advent of consumer participation and self-advocacy in the 1990s, groups of persons with mental retardation were referred to as the mentally handicapped, the mentally deficient, or the mentally retarded. The "People First" movement, however, redefined such terminology to emphasize the individuality of people over the disability; hence, the current use of such terms as a person with mental retardation, or people with a developmental disability.

9.01.3 EPIDEMIOLOGY

The prevalence of mental retardation is estimated to be between 1% and 3% of the general population. Mental retardation is estimated to be between 1 and 1.6 times more common among males than females. Estimates vary depending on sampling procedures, the definition of mental retardation employed, and the assessment methods utilized. For example, prevalence rates are generally higher when a diagnosis of mental retardation is based only on intelligence, and lower when a diagnosis is based on intelligence and concurrent deficits in adaptive behavior. Prevalence rates vary by age, with the highest rate occurring between the ages of six and 16 years, or the years of formal education. The more severely impaired children and children with clearly identified conditions (e.g., Down syndrome) are typically diagnosed before the school years. Persons with mild levels of mental retardation often are not diagnosed until deficits are identified during the school years. Many of these individuals then lose their mental retardation diagnosis following their school years. This is most often due to an improvement in adaptive skill levels and a demonstrated ability to function in society (e.g., maintain employment). The prevalence of mental retardation also varies depending on the level of severity. An estimated 80% of mental retardation is classified as mild, 12% as moderate, 7% as severe, and 1% as profound. Severe and profound mental retardation are thought to be distributed uniformly across socioeconomic classes. Mild mental retardation is thought to be more common in lower socioeconomic classes or in underdeveloped regions of the world where poor prenatal care, immunization, nutrition, hygiene, and sanitation are common.

The prevalence of comorbid psychiatric and behavioral disorders is estimated to be three to four times greater among persons with mental retardation than in the general population (American Psychiatric Association, 1994). Numerous epidemiological surveys since the mid-1980s indicate that between one-third and two-thirds of persons with mental retardation also have a mental disorder. The prevalence of psychopathology tends to increase with the degree of mental retardation, suggesting that persons with more severe levels of mental retardation have a greater risk for mental illness (Kaplan, Sadock, & Grebb, 1994). An estimated 3% of persons with mental retardation have schizophrenia and up to 50% have a mood disorder (e.g., depression, mania). There is also a high rate of pervasive developmental disorders (e.g., autistic disorder, Rett's disorder, childhood disintegrative disorder, Asperger's disorder), particularly among persons with severe levels of mental retardation.

The prevalence of severe behavior problems among persons with mental retardation is also well documented. Generally, persons with mild and moderate mental retardation and challenging behaviors tend to exhibit behaviors related to hyperactivity, low frustration tolerance, physical and verbal aggression, and other antisocial or socially inappropriate behaviors. Persons with severe levels of mental retardation often are not aware of or interested in their immediate environment and tend to display self-stimulatory, self-injurious, aggressive, and destructive behaviors. Severe behavior constitutes the single most important variable influencing placement in an institution and is a major factor in community placement failure (Aman & Singh, 1991).

Use of psychotropic medication to address psychopathology and behavior problems among persons with mental retardation is well documented. Persons with mental retardation are among the most medicated populations in society (Aman & Singh, 1991). An estimated 30–70% of individuals living in institutional settings receive psychotropic medication, while an estimated 14–45% of individuals living in community settings receive psychotropic medication (Singh, Guernsey, & Ellis, 1992). Singh et al. (1992) conducted a comprehensive review of legislation and litigation that has arisen from the misuse and mismanagement of psychotropic medication with persons with mental retardation. They identified three areas of malpractice involving psychotropic medication: (i) negligence in diagnosis, (ii) negligence in administering the medication and monitoring for medication side effects, and (iii) failure to obtain informed consent by the client or an appropriate representative of the client. Litigation has occurred in cases where physicians have failed to: (i) review the drug regimen in a timely manner; (ii) justify the drug and dosage level; (iii) document the rationale for changes in medication (dose increases, change in therapeutic class); (iv) ensure that psychotropic medication does not interfere with the individual's treatment program unless medically indicated; and (v) ensure that psychotropic medication is not used as a substitute for active treatment, as a punishment, or for staff convenience.

The prevalence of medical or health-related conditions among persons with mental retardation may be significantly higher than in the general population, although some investigators dispute this view (Anderson, 1989). The medical conditions often are associated with the etiology of the mental retardation, and include neurological disorders, visual and hearing impairments, orthopedic conditions that affect mobility, and disorders that affect the heart, kidneys, lungs, or spine. Epilepsy is very common among persons with mental retardation, particularly among those with more severe levels of mental retardation. Individuals with mental retardation who are considered medically fragile often are vulnerable to malnutrition, metabolic disturbances, and infection (Luckasson et al., 1992).

In the past, persons with mental retardation did not live as long as their peers without mental retardation, although mortality rates have gradually improved due to technological advances in medical sciences and nutrition, and improved life-style options. Mortality demographics from a 10-year cohort (1984–1993) of persons with mental retardation in New York showed that the average age of death for persons with mental retardation was 66.1 years, as compared with 70.4 years for people in the general population, and 55.8 years for people with Down syndrome (Janicki, Dalton, Henderson, & Davidson, 1996).

9.01.4 ETIOLOGY

There are hundreds of known causes of mental retardation, and medical researchers continue to identify new entities that may be associated with mental retardation. Causes of mental retardation are typically classified into three broad categories: prenatal causes, perinatal causes, and postnatal causes (see Table 1) (Luckasson et al., 1992). A combination of biological, environmental, and psychosocial factors interact to cause mild levels of mental retardation, while a biological feature is typically identified in severe levels of mental retardation. In approximately 50% of persons with mild mental retardation, and in approximately 30% of persons with severe mental retardation, no definitive causative factors can be identified. A probable cause is found in approximately 65% of all persons with mental retardation. Multiple causes are identified in approximately 50% of all persons with mental retardation.

9.01.5 ASSESSMENT AND DIAGNOSIS

The diagnosis of mental retardation has relied, in large part, on the administration and interpretation of standardized tests of intelligence and adaptive behavior. However, a diagnosis of mental retardation incorporates more than just standardized scores of intelligence and adaptive behavior. Typically, the assessment process includes interviews with people acquainted with the individual and observations of the individual in his/her natural settings. This section focuses on a review of standardized instruments used to diagnose mental retardation and comorbid psychopathology. Numerous assumptions must be addressed when utilizing standardized assessment measures (Anastasi, 1988; Luckasson et al., 1992; Morgenstern & Klass, 1991):

(i) The examiner is appropriately trained, certified, or licensed to administer the tests, is familiar with a wide variety of tests, has significant experience working with persons with mental retardation at all levels of severity, at various ages, and from the represented cultural and linguistic backgrounds.

(ii) The tests are norm-referenced for the population being tested. The tests are reliable and internally consistent over time. Estimates of reliability can be determined by the standard error of measurement, which establishes a range around the obtained score. Tests are valid for the population and culturally and linguistically appropriate for the individual being tested.

(iii) The examiner determines that the test results reflect the individual's best performance effort. Factors such as anxiety, fatigue, poor motivation, poor concentration, physical illness, hearing or visual impairments, or inability to understand instructions because of cultural or language barriers reduce the validity of the test results and must be accounted for.

(iv) The individual or his/her parent(s) or legal guardian must give informed consent for the assessment, can refuse all or any part of the assessment, can stop the assessment process at any time, can appeal any determinations made as a result of the assessment process (diagnoses, placement or treatment recommendations), and can request a reassessment by another examiner.

9.01.5.1 Intelligence

The Wechsler scales include three individually administered scales of intelligence: Wechsler Preschool and Primary Scale of Intelligence-Revised (WIPPSI-R; Wechsler, 1989), Wechsler Intelligence Scale for Children-Third Edition (WISC-III; Wechsler, 1991), and Wechsler

Table 1 Etiology of mental retardation.

I. Prenatal causes

A. Chromosomal disorders

1. Autosome
2. X-Linked mental retardation
3. Other X chromosome disorders

B. Syndrome disorders

1. Neurocutaneous disorders
2. Muscular disorders
3. Ocular disorders
4. Craniofacial disorders
5. Skeletal disorders
6. Other syndromes

C. Inborn errors of metabolism

1. Amino acid disorders
2. Carbohydrate disorders
3. Mucopolysaccharide disorders
4. Mucolipid disorders
5. Urea cycle disorders
6. Nucleic acid disorders
7. Copper metabolism disorders
8. Mitochondrial disorders
9. Peroxisomal disorders

D. Developmental disorders of brain formation

1. Neural tube closure defects
2. Brain formation defects
3. Cellular migration defects
4. Intraneuronal defects
5. Acquired brain defects
6. Primary (idiopathic) microcephaly

E. Environmental influences

1. Intrauterine malnutrition
2. Drugs, toxins, and teratogens
3. Maternal diseases
4. Irradiation during pregnancy

F. Other (nonspecific, unknown)

II. Perinatal causes

A. Intrauterine disorders

1. Acute placental insufficiency
2. Chronic placental insufficiency (marginal reserve)
3. Abnormal labor and delivery
4. Multiple gestation (smaller, later, or male infant)

B. Neonatal disorders

1. Hypoxic-ischemic encephalopathy
2. Intracranial hemorrhage
3. Posthemorrhagic hydrocephalus
4. Periventricular leukomalacia
5. Neonatal seizures
6. Respiratory disorders
7. Infections
8. Head trauma at birth
9. Metabolic disorders
10. Nutritional disorders

III. Postnatal causes

A. Head injuries

1. Cerebral concussion (diffuse axonal injury)
2. Cerebral contusion or laceration
3. Intracranial hemorrhage
4. Subarachnoid (with diffuse injury)
5. Parenchymal

B. Infections

1. Encephalitis
2. Meningitis
3. Fungal infections
4. Parasitic infestations
5. Slow or persistent virus infections

C. Demyelinating disorders

1. Postinfectious disorders
2. Postimmunization disorders
3. Schilder disease

D. Degenerative disorders

1. Syndromic disorders
2. Poliodystrophies
3. Basal ganglia disorders
4. Leukodystrophies
5. Sphingolipid disorders
6. Other lipid disorders

E. Seizure disorders

1. Infantile spasms
2. Myoclonic epilepsy
3. Lennox–Gastaut syndrome
4. Progressive focal epilepsy (Rasmussen)
5. Status epileptious-induced brain injury

F. Toxic-metabolic disorders

1. Acute toxic encephalopathy
2. Reye syndrome
3. Intoxication
4. Metabolic disorders

G. Malnutrition

1. Protein-calorie (PCM)
2. Prolonged intravenous alimentation

H. Environmental deprivation

1. Psychosocial disadvantage
2. Child abuse and neglect
3. Chronic social/sensory deprivation

I. Hypoconnection syndrome

Source: Luckasson et al. (1992, pp. 81–91).

Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981). The Wechsler scales are widely used in the USA; however, they have been criticized for their low reliability with individuals with IQ scores less than 50, for being culturally biased, and for yielding lower scores with children who are culturally deprived (Kaplan et al., 1994). The Wechsler scales yield a standard score with a mean of 100 and a standard deviation of 15. The WIPPSI-R is appropriate for children ages 3 years to 7 years 3 months, the WISC-III is appropriate for children ages 6 years to 16 years 11 months, and the WAIS is appropriate for persons ages 16–74 years.

The Stanford–Binet, Fourth Edition (Thorndike, Hagen, & Sattler, 1986), is an individually administered measure of general intelligence for persons ages 2 years to 23 years 11 months. The scale yields a standard score with a mean of 100 and a standard deviation of 16. As such, a cutoff score reflecting 2 standard deviation units below the mean is lower (IQ of 68) when compared with the Wechsler scales. Like the Wechsler scales, the Stanford–Binet has been criticized for its low reliability with individuals with IQ scores less than 50, for being culturally biased, and for yielding lower scores with children who are culturally deprived (Kaplan et al., 1994). The Stanford–Binet does not assess for mental retardation in younger children, as the lower-limit composite standard age score is 95 for children 2 years 0 months.

The Gesell Developmental Schedules (Gesell, 1949), and Bayley Scales of Infant Development (Bayley, 1969) can be used when assessing for intellectual limitations in children younger than three years. The Gesell was originally published in 1940, and at the time, represented a pioneering attempt to provide a systematic, empirically based assessment of behavior development in young children. The test is appropriate for children ages four weeks to five years. Five major fields of behavior are assessed, including adaptive, gross motor, fine motor, language, and personal-social behavior. The data are obtained through observation of the child's responses to standard toys and stimulus objects, and by gathering information from the child's caretakers.

The Bayley Scales of Infant Development are appropriate for assessing the developmental status of children ages 2 months to 2 years 6 months. The test consists of three scales: mental scale, motor scale, and infant behavior record. Standard scores are derived for the mental and motor scales with norms classified by half-month steps from 2 to 6 months and by one-month steps from 6 to 30 months. The mental scale measures perception, memory, learning, problem solving, vocalization, and initial verbal communication and abstract thinking. The motor scale measures gross motor abilities (e.g., sitting, standing, walking, climbing stairs), and fine motor skills of the hands and fingers. The infant behavior record assesses emotional and social behavior, attention span, persistence, and goal directedness.

The Gesell, Bayley, and Cattell Measurement of Intelligence in Infants and Young Children (Cattell, 1950) also have been recommended to assess the intellectual levels of older persons who fall in the severe and profound levels of mental retardation and, consequently, below the floors of the Wechsler and Stanford–Binet scales. All three measures provide a chronological assessment of social, adaptive, language, and motor areas of growth and development (Morgenstern & Klass, 1991). The Peabody Picture Vocabulary Test-Revised (PPVT-R; Dunn & Dunn, 1981) is an alternative measure for persons with physical or other disabilities that make oral and written responses difficult or impossible. The PPVT-R is appropriate for children ages 2 years 6 months to adulthood. Scores on the PPVT-R can be standardized to reflect an equivalent Wechsler full scale IQ score ($M = 100$, $SD = 15$). When compared with the Wechsler scales and Stanford–Binet, the PPVT-R measures a similar yet narrower range of attributes. Therefore, it should not be used as the only or primary source for classifying intellectual functioning (Morgenstern & Klass, 1991).

The Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983) is appropriate for children ages 2 years 6 months to 12 years 6 months. The battery yields a mental processing composite score ($M = 100$, $SD = 15$) that is roughly equivalent to the Wechsler full scale IQ score, and subtest scores in four areas: sequential processing, simultaneous processing, nonverbal, and sociocultural. The K-ABC is considered a better predictor of intellectual ability in children with mental retardation, in comparison to the WISC-R. This is because the K-ABC measures intellectual ability without emphasizing academic and verbal skills, which is characteristic of the Wechsler scales (Morgenstern & Klass, 1991).

9.01.5.2 Adaptive Behavior

Diagnostic standards for mental retardation emphasize both the individual's intellectual capability and his/her adaptive daily living skills in the areas of communication, self-care, home living, social skills, community use, self-direction, health and safety, functional academics, leisure, and work (American Psychiatric Association, 1994; Luckasson et al., 1992). Historically, measures of intellectual capability have dominated the diagnosis of mental retardation even though previous definitions have specified that intelligence and adaptive behavior should have equal weight in diagnosis. The diagnostic standards implemented in the 1980s and 1990s specify that limitations must occur in two or more applicable adaptive skills areas in order to diagnosis mental retardation. Unlike measures of intelligence, however, there are no consistently used cutoff scores or precise age equivalent scores that structure the assessment process. Although there are numerous measures of adaptive behavior available (see Reschly, 1987, 1990 for a review), only two well-established instruments used in the diagnostic process are reviewed here.

The Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cicchetti, 1984) include a survey form, expanded form, and classroom form. The survey form is primarily used as a diagnostic and placement instrument for persons between the ages of birth and 18 years 11 months, and for low-functioning adults. The expanded form is used for preparing individualized educational or habilitative programs, and is appropriate for persons between the ages of birth and 18 years 11 months, and for low-functioning adults. The classroom form is used to assess adaptive behavior skills relevant to the classroom, and is appropriate for persons between the ages of 3 years and 12 years 11 months. The Vineland scales provide an overall level of functioning, or Adaptive Behavior Composite (ABC) score, and domain (subdomain) scores in the areas of communication (receptive, expressive, written); daily living skills (personal, domestic, community); socialization (interpersonal, play/leisure time, coping skills); and motor skills (gross motor, fine motor). The survey and expanded forms contain a maladaptive domain designed to assess maladaptive behavior.

AAMD Adaptive Behavior Scale, Residential and Community Edition (Nihira, Foster, Shellhaas, & Leland, 1975) is designed to provide objective descriptions and evaluations of an individual's ability to cope with the natural and social demands of his/her environment. The instrument was standardized on over 4000 individuals, ages 3 to 69 years, living in institutional settings. There are no normative data for individuals living in community settings. The instrument is divided into two sections. Part I is organized developmentally and evaluates skills and habits in 10 behavioral domains: independent functioning, physical development, economic activity, language development, numbers and time, domestic activity, vocational activity, self-direction, responsibility, and socialization. Part II consists of 14 domains designed to assess aberrant behavior: violent and destructive behavior, antisocial behavior, rebellious behavior, untrustworthy behavior, withdrawal, stereotyped behavior and odd mannerisms, inappropriate interpersonal manners, unacceptable vocal habits, unacceptable or eccentric habits, self-abusive behavior, hyperactive tendencies, sexually aberrant behavior, psychological disturbances, and use of medication. While the instrument offers valuable information about a variety of adaptive behavior domains, the results should be validated through other observation and assessment procedures, especially if it is being used as a diagnostic tool.

9.01.5.3 Psychopathology

The diagnosis of mental disorders is particularly difficult with persons with mental retardation because they may have a limited vocabulary to describe their symptoms or express their feelings or concerns. Most commonly used psychometric evaluations have not been validated for use with persons with mental retardation. Many instruments require reading or vocabulary comprehension that is beyond the ability of most persons with mental retardation. A structured clinical assessment conducted by a psychologist trained in mental retardation has its own set of challenges due to individual differences in the client's verbal ability, attention span, and memory. Also, characteristics of the individual's mental retardation might be misinterpreted as psychopathology, resulting in an inappropriate psychiatric diagnosis. The challenge of obtaining a valid psychiatric diagnosis increases with the level of severity of mental retardation. A comprehensive assessment of mental illness in persons with mental retardation should include the following sources of information (Luckasson et al., 1992, p. 53): "(a) interviews; (b) behavioral observation in everyday environments; (c) client and staff interviews; (d) formal psychometric evaluations, including social skills development, personality, and maladaptive behavior; and (e) medical and biological evaluations, including positron emission tomography and other types of neuroimaging assessment." Since the mid-1980s, numerous measures have been developed to assess psychopathology and behavior problems in persons with mental retardation (see Aman, 1991 for a review of these instruments). Three instruments with established psychometric properties are described here.

There are two versions of the Aberrant Behavior Checklists (ABC): ABC-Residential (Aman & Singh, 1986) and ABC-Community (Aman & Singh, 1994). The instruments are appropriate for use with individuals ages five years through adulthood. The instruments were designed to assess inappropriate and maladaptive behavior, and for use as a monitoring tool to evaluate treatment effects. The ABC-Residential was designed for persons with mental retardation living in institutional settings, and the community version was designed to provide a parallel measure for persons with mental retardation living in community settings. The instruments consist of five subscales: (i) irritability, agitation, crying; (ii) lethargy, social withdrawal; (iii) stereotypy; (iv) hyperactivity, noncompliance; and (v) inappropriate speech. Normative data for each subscale are available by age, gender, and level of mental retardation. The authors provide clear operational definitions of each item. The instrument can be administered by any trained individual who has regular contact with the client and knows him/her well.

The Reiss Screen for Maladaptive Behavior (Reiss, 1988) was designed to assess for significant mental health problems in persons with mental retardation living in institutional and community settings. The instrument is appropriate for use with persons with mild through profound mental retardation, ages 12 years and older. The instrument consists of eight subscales: aggressive behavior, psychosis, paranoia, depression (behavioral signs), depression (physical signs), dependent personality disorder, avoidant disorder, and autism. There also are items designed to evaluate serious behavior problems: drug/alcohol abuse, self-injury, stealing, overactivity, sexual problems, and suicidal tendencies. The instrument provides cutoff scores that indicate a need for further assessment; these cutoff scores are available for the total scale, each subscale, and for the items assessing serious behavior problems. The instrument is administered by at least two trained raters who have regular contact with the client and know him/her well.

The Diagnostic Assessment for the Severely Handicapped (DASH) Scale (Matson, Gardner, Coe, & Sovner, 1990) is appropriate for use with adolescents and adults with severe or profound mental retardation. The instrument consists of 13 subscales that are derived from the *DSM-III-R* classification system: anxiety, mood disorder–depression, mood disorder–mania, pervasive developmental disorder/autism, schizophrenia, stereotypies/tics, self-injurious behaviors, elimination disorders, eating disorders, sleep disorders, sexual disorders, organic syndromes, and impulse control and miscellaneous behavior problems. The instruments consists of two sections: a section seeking background information about the client; and a behavior-rating component designed to determine the frequency, duration, and severity of behavior problems or psychiatric symptoms. The instrument is administered by a mental health professional who conducts structured interviews with individuals who have regular contact with the clients and know them well.

9.01.6 TREATMENT

9.01.6.1 Behavioral Interventions

The development of appropriate adaptive skills involves the acquisition and maintenance of appropriate social and personal behavior and the elimination of behavioral excesses. Learning of these skills occurs in an ecological context. This ecological context is a system in which the biology and behavior of the individual, the physical environment, the social environment, and the behavior of other people function as a mutually interactive system in which each component influences the functioning of the entire system (Singh, 1997a).

Understanding the rationale of effective behavioral interventions requires some discussion of the development of adaptive behavior in an ecological context for an individual, and how that development differs for persons with mental retardation. Initially, infants have very little direct control over the environment or the behavior of other people. As children develop biologically and interact with the social and physical environment, they begin to learn the relationship between their own behavior, the behavior of others, and reinforcing and aversive environmental consequences. As this process continues in a balanced ecological system, the individual learns to increase those socially adaptive functional behaviors that promote pleasant social interactions, and help meet basic needs in a socially acceptable manner. At the same time, inappropriate behavior that does not promote social interaction or meet basic needs begins to extinguish due to a lack of reinforcement or decreases in frequency and intensity as a result of aversive consequences. This process occurs most efficiently and will most likely enhance the person's quality of life, when the individual's health is adequate, when the physical environment is appropriately stimulating but not overwhelming, when other people are appropriately responsive, and when the social environment is inclusive; that is, when the ecological system is balanced.

The process of acquiring adaptive skills is greatly enhanced by the development of language and cognitive capacity. The development of receptive and expressive language allows social learning to proceed more rapidly when contingencies can be represented symbolically, rather than through repetitive, direct experience. Increasing cognitive abilities allows the learner to be aware of contingencies that are remote in time or occur on infrequent schedules. Self-talk provides salient antecedents and immediate consequences not always available from the external environment.

For many individuals, the biological environment of good health, and social and physical environment of a family, community, society, and educational system creates a balanced ecological system that is likely to lead to the learning of a set of relatively adaptive behaviors. The learner moves through the environment and encounters naturally occurring contingencies, and contingencies structured by caretakers and a community that shapes the appropriate adaptive skills. The more effective these learned social and adaptive skills are, the more positive interactions the individual has with the social and physical environment, and the higher the quality of life.

For persons with mental retardation, interaction in a system that promotes and strengthens adaptive behavior may be limited. The more severe the deficits, the less likely appropriate learning experiences will occur in the natural environment. Deficits in expressive and receptive language skills and cognitive capacity reduce functional interactions with caretakers and peers in the social environment and are severe impediments to efficient adaptive learning. When socially appropriate skills are not learned, other people in the community are less likely to interact with the individual, and opportunities for further learning is restricted. If the individual also has inappropriate behavioral excesses, further restrictions in the social and physical environment are likely to be imposed. In this ecologically unbalanced system, the individual's quality of life diminishes as restrictions in the social and physical environment increase.

Thus, the limitations in cognitive functioning of the individual with mental retardation is only one part of the disordered system that promotes behavioral excesses, prevents development of adaptive behavior, and reduces the quality of life. Behavioral interventions that target only a narrowly defined behavior in a restricted environment are unlikely to promote more adaptive behavior. The ecology of the system in which the individual functions needs to be understood, and a plan for altering the system in ways that bring a better balance between the individual and the environment needs to be developed.

An awareness of the role of the ecological context in determining the adaptive functioning and quality of life for individuals with mental retardation has entered the community consciousness as the principle of the least restrictive environment, the goal of deinstitutionalization and community placement, and increased sensitivity to the civil rights of institutionalized persons (Singh, 1995). An awareness of the role of ecological context also has influenced the nature of education and training. Interdisciplinary assessment, full inclusion in the least restrictive environment, due process, and the use of written individualized education plans or individualized family service plans have helped to improve the breadth of treatment planning so that it includes plans for altering the environment in ways that are more likely to support positive change and increase quality of life. Services required include the development of formal educational environments to enhance academic and adaptive skill development, treatment for behavioral excesses, skills training for behavioral deficits, psychiatric services for emotional disorders, and regular medical and dental services.

Behavioral programs are based on the principles of learning. Thus, behavioral excesses that result from faulty or incomplete learning, and skill deficits that result from lack of opportunity to learn the skills are most amenable to behavioral programs. To be consistent with an approach that attempts to enhance the quality of life through greater ecological balance, behavioral interventions should target both behavior excesses to be reduced and skills to be enhanced. These interventions will be most effective if they take into account the context of the social, physical, and biological environment, and attempt to maximize the balance of the ecological system.

The assessment and analysis of the ecological context of problem behavior, and the treatment of that behavior, involve a process that includes:

(i) structural analysis of the problem behavior;

(ii) descriptive analysis of the individual's behavior in its ecological context;

(iii) hypothesis formation;

(iv) functional analysis done informally in the natural environment, or formally in an analogue setting;

(v) planning, implementing, and evaluating of treatment; and

(vi) maintenance and generalization of treatment (Singh, 1997b).

A structural analysis is designed to assess the impact of the physical, social, and biological environments on the problem behavior. Changes in these environments may lead to significant improvement in the individual's quality of life without the need for time-consuming and labor-intensive behavioral interventions. If the analysis indicates that the problem behavior is maintained by the individual's social (e.g., caretakers, peers) or physical (e.g., workplace structure) environment, one focus of treatment will be on the manipulation of these environments. If the analysis indicates that the problem is maintained by the individual's biological environment (e.g., earache, urinary tract infection, toothache), collaboration with a physician/dentist specializing in developmental disabilities will focus treatment on medical intervention and management (Feldman & Griffiths, 1997).

A descriptive analysis provides as comprehensive a picture as possible of the problem behavior in its context. A detailed description of the target behavior and its antecedents and consequences can be determined through the use of structured or unstructured observations in the natural setting, information from caretakers who know the individual and the situations in which the target behavior occurs, and systematic data collection using frequency or interval recording procedures. This descriptive analysis provides the information needed to form tentative hypotheses regarding the functional relationship between the target behavior and the environmental events that maintain it.

Testing the tentative hypotheses can be done formally in an experimental analysis of the problem behavior. This is done by establishing a controlled environment with a limited number of contingencies as an analogue to the natural environment. Variables are manipulated one at a time to determine their effect on the behavior. This can be time consuming and labor intensive. It requires that the problem behavior be allowed to continue and even increase in order to determine the effective contingencies. For this reason, an informal functional analysis carried out in the environment in which the problem behavior usually occurs, and in which it will be treated, is usually preferable. In these situations where the informal analysis does not lead to successful treatment, a formal assessment can be conducted under analog conditions.

The resulting treatment plan will have, as its primary goal, the enhancement of the quality of life of the individual. An effective treatment plan will focus on the creation of a balanced ecological system. The components of the plan will include guidelines for making changes in the biological, physical, and social environments, reducing excessive behavior, acquiring new skills, and improving interaction with the social and physical environment. Although no one of these components individually is an adequate plan, by systematically changing one or more components of the system, an effective treatment plan can be implemented.

9.01.6.2 Psychopharmacology

Psychotropic medication is any medication prescribed to stabilize or improve mood, mental status, or behavior (Kalachnik et al., 1995). It is used as one part of a comprehensive treatment plan to address an individual's psychiatric symptoms or aberrant behavior. Ideally, the medication acts as a setting event for the occurrence of appropriate behavior when maladaptive behavior is reduced. Thus, psychotropic medication can allow the individual to function more fully or appropriately at school, work, or home. Psychotropic medication should be used to enhance the individual's quality of life rather than to strictly reduce aberrant or undesirable behavior (Ellis, Singh, & Singh, 1997).

Kalachnik et al. (1995), as part of the International Consensus Conference on Psychopharmacology, delineated guidelines for the use of psychotropic medication with persons with mental retardation and other developmental disabilities (see Table 2). The following sources were utilized in the formation of the guidelines: (i) regulatory (e.g., Health Care Financing Administration, 1992); (ii) accreditation (e.g., Joint Commission on Accreditation of Healthcare Organizations, 1995); (iii) professional (e.g., American Psychiatric Association Committee on Research on Psychiatric Treatments, 1992); (iv) litigation (e.g., *Wyatt v. Stickney*, 1972); (v) legislation (e.g., Civil Rights of Institutionalized Persons Act, 1981); and (vi) proclamations and declarations (e.g., Assembly of the United Nations, "Declaration on the Rights of Mentally Retarded Persons"; see Beyer, 1988; Kalachnik et al., 1995; Singh et al., 1992). The guidelines are endorsed by professionals in the field of mental retardation and serve as a model for appropriate use.

Administration and monitoring of psychotropic medication should be conducted by a psychiatrist in cooperation with all individuals who interact with the client on a day-to-day basis (e.g., psychologist, teachers, caretakers). A multidisciplinary team approach enhances the reliable assessment as to whether the medication is successfully addressing the behaviors or psychiatric symptoms it was prescribed to address, and whether the medication is causing undesirable side effects (e.g., sedation, agitation) that interfere with the client's day-to-day activity. Medication monitoring processes are enhanced by the ongoing collection of systematic data that reflect the frequency, duration, and intensity of target behavior or psychiatric symptoms. This type of data is best collected by individuals who interact with the client for extended periods of time each day or night. These individuals can be trained to reliably assess and record the frequency, duration, and intensity of occurrence of aberrant behavior or psychiatric symptoms.

A thorough psychotropic medication review should take place every 3–6 months unless more frequent reviews are indicated (e.g., change in medication or dose, increase in behavior or psychiatric symptoms). The process should include a review of the client's psychiatric/behavioral history; a careful evaluation of the client's medical diagnoses and current medical status; behavioral changes since the last review (frequency, duration, intensity); observed medication side effects; a report of any environmental changes that may impact the client's

Table 2 Guidelines for the use of psychotropic medication.

1. *Psychotropic medication definition*. A psychotropic medication is any drug prescribed to stabilize or improve mood, mental status, or behavior.
2. *Inappropriate use*. Psychotropic medication shall not be used excessively, as punishment, for staff convenience, as a substitute for meaningful psychosocial services, or in quantities that interfere with an individual's quality of life.
3. *Multidisciplinary care plan*. Psychotropic medication must be used within a coordinated multidisciplinary care plan designed to improve the individual's quality of life.
4. *Diagnostic and functional assessment*. The use of psychotropic medication must be based upon a psychiatric diagnosis or a specific behavioral-pharmacological hypothesis resulting from a full diagnostic and functional assessment.
5. *Informed consent*. Written informed consent must be obtained from the individual, if competent, or the individual's guardian before the use of any psychotropic medication and must be periodically renewed.
6. *Index behaviors and empirical measurement*. Specific index behaviors and quality of life outcomes must be objectively defined, quantified, and tracked using recognized empirical measurement methods in order to monitor psychotropic medication efficacy.
7. *Side effects monitoring*. The individual must be monitored for side effects on a regular and systematic basis using an accepted methodology which includes a standardized assessment instrument.
8. *Tardive dyskinesia monitoring*. If antipsychotic medication or other dopamine blocking drugs (e.g., amoxapine or metoclopramide) are prescribed, the individual must be monitored for tardive dyskinesia on a regular and systematic basis using a standardized assessment instrument.
9. *Regular and systematic review*. Psychotropic medication must be reviewed on a regular and systematic basis. Clinical reviews must be conducted on a regular and systematic basis by the prescriber. Data reviews must be conducted on a regular and systematic basis by appropriate members of the multidisciplinary team.
10. *Lowest "optimal effective dose."* Psychotropic medication must be reviewed on a periodic and systematic basis to determine if it is still necessary or, if it is, if the lowest "optimal effective dose" is prescribed.
11. *Frequent changes*. Frequent drug and dose changes should be avoided.
12. *Polypharmacy*. Keep psychotropic medication regimens as simple as possible in order to enhance compliance and minimize side effects.
13. *Practices to minimize*. Minimize the following practices to the degree possible: (i) long-term use of PRN (as needed) orders; (ii) long-term use of benzodiazepine antianxiety medications such as diazepam; (iii) use of long-acting sedative-hypnotic medications such as chloral hydrate; (iv) long-term use of shorter-acting sedative-hypnotics such as temazepam (Restoril); (v) anticholinergic use such as benztropine without signs of EPSE; (vi) long-term anticholinergic use; (vii) antipsychotic medication at high doses; (viii) use of phenytoin, phenobarbital, and primidone as psychotropic medication.
14. *Peer or external review*. Establish a system of peer or external review of psychotropic medication prescribing which incorporates a system of flagging up cases of greatest concern.

Source: Bisconer, Sine, and Zhang (1996).

behavior or psychiatric symptoms; a progress report on skills training designed to decrease targeted behavior; and any changes to long-term plans for treatment and education (Bisconer, Zhang, & Sine, 1995).

Medication side effects should be systematically monitored by the psychiatrist via appropriate laboratory tests and objective screening procedures (see Gadow & Poling, 1988 for a review of screening instruments). The psychiatrist can rely on trained care staff and other professionals to conduct periodic screenings for side effects. These results can then be reviewed and validated when the psychiatrist meets with the client.

Table 3 provides psychotropic and antiepileptic medications by class, generic name, and recommended dosage by age for persons with mental retardation. Table 4 provides the most common indications and side effects for different classes of medications. A comprehensive review of psychotropic medication use with persons with mental retardation can be found in the following sources (Aman & Singh, 1991; Ellis et al., 1996; Singh, Ellis, & Singh, 1994).

9.01.7 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

It is becoming increasingly common for persons with mental retardation to live and interact in inclusive community settings rather than in segregated settings (e.g., institutions, training centers, group homes, sheltered workshops). With this movement toward full com-

Table 3 Recommended doses for the various classes of psychotropic drugs.

	Average daily dose		
Drug	*Children*	*Adolescents*	*Adults*
A. Antipsychotics			
chlorpromazine	30–200 mg	40–400	100–800 mg
(> 6 mos of age)[a]	(2.5–6 mg/kg/day)	(3–6 mg/kg/day)	(max 2000 mg/day)
thioridazine	75–200 mg	10–200+ mg	150–800 mg
(> 2 yrs of age)[a]	(0.5–3 mg/kg/day)		(max 800 mg/day)
trifluperazine	1–15 mg	1–20 mg	15–40 mg
(> 6 yrs of age)[a]			
thiothixene	2–10 mg	5–30 mg	20–60 mg
(> 12 yrs of age)[a]			
haloperidol	0.5–4 mg	2–16 mg	2–16+ mg
(> 3 yrs of age)[a]	(0.05–0.15 mg/kg/day)		
reserpine	0.02–0.25 mg	0.1–1.0 mg	0.1–1.0 mg
clozapine		50–200 mg	300–450 mg
(> 16 yrs of age)[a]		(3–5 mg/kg/day)	(max 900 mg/day)
loxapine	5–50 mg	20–100 mg	60–250 mg
(> 16 yrs of age)[a]			
B. Antidepressants			
amitriptyline	30–100 mg	50–100 mg	75–300 mg
(> 12 yrs of age)[a]	(1–5 mg/kg/day)	(1–5 mg/kg/day)	
bupropion	25–150 mg	75–300 mg	200–450 mg
(> 18 yrs of age)[a]		(3–6 mg/kg/day)	
clomipramine	25–100 mg	50–150 mg	100–250 mg
(> 10 yrs of age)[a]		(2–3 mg/kg/day)	
desipramine			
10–150 mg	50–150 mg	100–200 mg	
(> 12 yrs of age)[a]	(1–5 mg/kg/day)	(1–5 mg/kg/day)	
fluoxetine	5–20 mg	10–60 mg	20–80 mg
(> 18 yrs of age)[a]		(0.5–1 mg/kg/day)	
imipramine	10–150 mg	50–200 mg	75–200 mg
(> 6 yrs of age)[a]	(1–5 mg/kg/day)	(1–5 mg/kg/day)	
nortriptyline	10–100 mg	50–100 mg	75–200 mg
(> 12 yrs of age)[a]		(1–3 mg/kg/day)	
phenelzine		15–45 mg	45–90 mg
(> 16 yrs of age)[a]		(0.5–1 mg/kg/day)	
sertraline	25–100 mg	50–200 mg	50–200 mg
(not in children)[a]		(1.5–3 mg/kg/day)	
C. Antimatics			
lithium carbonate[b]	300–900 mg	900–1200 mg	900–1200 mg
(> 12 yrs of age)[a]		(10–30 mg/kg/day)	
D. Anxiolytics			
alprazolam	0.25–2 mg	0.75–5 mg	1–8 mg
(> 18 yrs of age)[a]		(0.02–0.06 mg/kg/day)	
chlordiazepoxide	10–30 mg	20–60 mg	20–100 mg
(> 6 yrs of age)[a]			
diazepam	1–10 mg	2–20 mg	4–40 mg
(> 6 mos of age)[a]		(max 0.8 mg/kg)	
lorazepam	0.25–3 mg	0.05–6 mg	1–10 mg
(> 12 yrs of age)[a]		(0.04–0.09 mg/kg/day)	
diphenhydramine	25–200 mg	50–300 mg	50–400 mg
		(1–5 mg/kg/day)	
hydroxyzine	25–100 mg	40–150 mg	75–400 mg
		(2 mg/kg/day)	
buspirone	2.5–15 mg	5–30 mg	15–60 mg
(> 18 yrs of age)[a]		(0.2–0.6 mg/kg/day)	

Table 3 (continued).

	Average daily dose		
Drug	*Children*	*Adolescents*	*Adults*
E. Stimulants			
dextroamphetamine	2.5–15 mg	5–40 mg	10–40 mg
(> 3 yrs of age)^a	(0.15–0.5 mg/kg/dose)	(0.15–0.5 mg/kg/dose)	(0.15–0.5 mg/kg/dose)
methylphenidate	2.5–30 mg	10–60 mg	20–60 mg
(> 6 yrs of age)^a	(0.3–1 mg/kg/dose)	(0.3–1 mg/kg/dose)	(0.3–1 mg/kg/dose)
pemoline	18.75–75 mg	37.5–112.5 mg	37.5–112.5 mg
(>6 yrs of age)^a	(1–3 mg/kg/day)	(1–3 mg/kg/day)	(1–3 mg/kg/day)
F. Antiepileptics			
carbamazepine^b	200–800 mg	400–1000 mg	600–1200 mg
(>6 yrs of age)	(5–20 mg/kg/day)	(10–30 mg/kg/day; max dose 1000 mg/day)	(max dose 1200 mg/day)
ethosuximide^b	250–800 mg	500–1500 mg	750–1500 mg
	(20–30 mg/kg/day)	(20–40 mg/kg/day)	(max dose 1500 mg/day)
phenobarbital^b	<250 mg	75–250 mg	150–250 mg
	(4–8 mg/kg/day)	(1–3 mg/kg/day)	
phenytoin^b	<300 mg	300–500 mg	300–400 mg
	(7.5–9 mg/kg/day)	(6–7 mg/kg/day)	
primidone	150–750 mg	750–1500 mg	750–2000 mg
sodium valproate	250–1000 mg	500–2000 mg	500–2500 mg
			(15–60 mg/kg/day)
G. Others			
propranolol	2–80 mg	20–140 mg	80–480 mg
		(max 2 mg/kg/day)	
clonidine	0.25–0.3 mg	0.3–0.4 mg	0.3–0.5 mg
(not in children)^a	(3–6 μg/kg/day)	(3–6 μg/kg/day)	
naltrexone	10–50 mg	40–120 mg	50–150 mg
(> 18 yrs of age)^a	(0.5–1.5 mg/kg/day)	(1–2 mg/kg/day)	(1.2 mg/kg/day)
			fenfluramine
30–60 mg	40–100 mg	60–120 mg	
(> 12 yrs of age)^a		(1–2 mg/kg/day)	
benzotropine	0.5–4 mg	0.5–6 mg	2–6 mg
(> 3 yrs of age)^a		(43–86 μg/kg/day)	

[a] Recommended FDA guidelines. [b] Dosage titrated using serum levels.
Source: Ellis et al. (1996); reprinted with permission from the authors.

munity inclusion comes the challenge of successfully integrating persons with mental retardation into neighborhoods, schools, employment settings, recreation and leisure settings, and retirement settings. Persons with mental retardation will continue to have a broad range of specialized needs, including a need for appropriate educational services and employment training, ongoing training in daily living and social skills, and oftentimes, long-term treatment for behavioral excesses and psychiatric disorders. Many persons with mental retardation have a life-long need for specialized medical care. Addressing the needs of persons with mental retardation typically requires the expertise of a variety of professionals working as part of a multidisciplinary team (e.g., physicians, dentists, nurses, psychologists, social workers, related services professionals, educators). Successful community inclusion also requires the ongoing support of the individual's family members and relatives, friends, neighbors, and peers with and without disabilities.

Clinicians working with persons with mental retardation need specialized training in the assessment and diagnosis of mental retardation and coexisting psychiatric disorders, intervention and treatment methods, and experience in multidisciplinary team models. Clinicians often become a liaison with medical specialists including psychiatrists, neurologists, dentists, and others who have less specialized training with this population. They also provide ongoing training and assistance to family members and care providers in behavioral support strategies.

Table 4 Psychiatric and behavioral indications and side effects of various classes of drugs.

A. Antipsychotic	
Indications:	Psychotic states; schizophrenia (exacerbations and maintenance); mania (in conjunction with lithium); behavior disorders with severe agitation, aggressivity and self-injury; and dyskinetic movement disorders (e.g., Tourette's disorder and juvenile Huntington's disease)
Side effects:	Anticholinergic effects, including dry mouth, constipation, blurred vision, and urinary retention (most common with low potency phenothiazines); extrapyramidal reactions, including acute dystonia, akathesia, tremor (particularly with high potency phenothiazines); neuroleptic malignant syndrome; tardive dyskinesia (lower risk with clozapine); other central nervous system effects, including sedation, fatigue, cognitive blunting, psychotic symptoms, confusion, and excitement; orthostatic hypotension and cardiac conduction abnormalities; endocrine disturbances (e.g., menstrual irregularities and weight gain); gastrointestinal distress; skin photosensitivity; granulocytopenia and agranulocytosis (clozapine); and allergic reactions
B. Antidepressants	
Indications:	Enuresis; Attention-Deficit/Hyperactivity Disorder; Major Depressive Disorder; and anxiety disorders (including school phobia, separation anxiety disorder, panic disorder, and obsessive-compulsive disorder)
Side effects:	*Tricyclics:* Anticholinergic effects, including dry mouth, constipation, blurred vision, and urinary retention; cardiac conduction slowing (treatment requires EKG monitoring), mild increases and/or irregularity in pulse rate and mild decreases or increases in blood pressure; confusion or the induction of psychosis; seizures; rash; and endocrine abnormalities *Monoamine oxidase inhibitors:* mild decreases or increases in blood pressure; drowsiness; weight gain; insomnia; hypertensive crisis with nonadherance to dietary restrictions (necessary to eliminate high tyramine foods from diet) or with certain drugs *Selective serotonin reuptake inhibitors:* irritability; gastrointestinal distress; headaches; insomnia *Other antidepressants:* irritability (bupropion, venlafaxine); insomnia (bupropion, venlafaxine); drug-induced seizures (bupropion, with high doses); changes in blood pressure (trazodone, venlafaxine); priapism (trazodone); sedation, sleepiness (trazodone, venlafaxine); gastrointestinal distress (venlafaxine); and headache (venlafaxine)
C. Antimanics	
Indications:	Manic episodes of Bipolar Disorder; unipolar depression/adjunct treatment in Major Depressive Disorder; behavior disorders with extreme aggression
Side effects:	Kidney abnormalities leading to increased urination and thirst; gastrointestinal distress; fine hand tremor; weakness and ataxia; possible thyroid abnormalities (with long-term use), weight gain, and electrolyte imbalances; sedation, confusion, slurred speech, irritability, headache, and subtle cogwheel rigidity; skin abnormalities; orthostatic hypotension and pulse rate irregularities; and allergic reactions
D. Anxiolytics	
Indications:	Anxiety disorders; seizure control; night terrors; sleepwalking; insomnia and acute management of severe agitation; adjunct treatment of mania and refractory psychosis; Tourette's Disorder
Side effects:	Headache, sedation, and decreased cognitive performance; behavioral disinhibition, including overexcitement, hyperactivity, increased aggressivity, and irritability; gastrointestinal distress, central nervous system disinhibition resulting in hallucinations, psychotic-like behavior, and depression; physical and psychological dependence (particularly with long-acting benzodiazepines); rebound or withdrawal reactions (particular with short-acting benzodiazepines); blood abnormalities; anticholinergic effects, including dry mouth, constipation, and blurred vision (antihistamines); and allergic reactions
E. Stimulants	
Indications:	Attention-Deficit/Hyperactivity Disorder (including those with mental retardation, fragile X syndrome, Tourette's disorder, head trauma, pervasive developmental disorders, or other comorbid disorder); narcolepsy; adjunctive treatment in refractory depression

Table 4 (continued).

Side effects:	Decreased appetite; weight loss; abdominal pain; headache, insomnia, irritability; sadness and depression; mild increases in pulse rate and blood pressure; possible temporary suppression of growth (with long-term use); choreoathetosis (pemoline) and rarely, tic disorders; and elevated liver function tests (pemoline)
F. Antiepileptics	
Indications:	Seizure control; Bipolar Disorder; adjunct treatment in Major Depressive Disorder; severe behavior problems (e.g., aggression, self-injury)
Side effects:	Sedation, weakness, dizziness, disturbances of coordination and vision, hallucinations, confusion, abnormal movements, nystagmus, slurred speech, and depression; blood abnormalities; gastrointestinal distress; skin rashes, alterations in pigmentation and photosensitivity reactions; increased or decreased blood pressure and congestive heart failure; abnormalities of liver functions (sodium valproate, carbamezapine–rate); genitourinary tract dysfunction; coarsening of facial features, enlargement of the lips, gingival hyperplasia, and excessive hair growth (phenytoin); and bone marrow suppression (carbamezapine, sodium valproate)
C. Others	
Propranolol	
Indications:	Behavior disorders with severe aggression, self-injury, or agitation; Tourette's disorder, akathesia
Side effects:	Decreased heart rate, peripheral circulation and blood pressure; fatigue, weakness, insomnia, nightmares, dizziness, hallucinations, and mild symptoms of depression; shortness of breath and wheezing (especially in patients with asthma); gastrointestinal distress; and rebound hypertension on abrupt withdrawal
Clonidine	
Indications:	Attention-Deficit/Hyperactivity Disorder; Tourette's Disorder; behavior disorders with severe aggression, self-injury, or agitation; adjunct treatment of schizophrenia and mania; possible use in anxiety disorders
Side effects:	Sedation; decrease in blood pressure; rebound hypertension; dry mouth; confusion (with high doses); depression
Guanfacine	
Indications:	Attention-Deficit/Hyperactivity Disorder; Tourette's Disorder
Side effects:	Sedation (less than with clonidine); decrease in blood pressure (less than with clonidine); rebound hypertension; dry mouth; confusion (with high doses); depression
Opioid antagonists	
Indications:	Reversal of narcotic depression; self-injury
Side effects:	Drowsiness, dizziness, dry mouth, sweating, nausea, abdominal pain, and loss of energy
Fenfluramine	
Indications:	Management of obesity; possible use in the control of some behavior problems in pervasive developmental disorders
Side effects:	Anorexia, weight loss; drowsiness, dizziness, confusion, headache, incoordination; mood alterations, anxiety, insomnia, weakness, agitation, and slurred speech; gastrointestinal distress; increased or decreased blood pressure and palpitations; skin rashes; dry mouth; eye irritation; and muscle aches

Source: Ellis et al. (1996); reprinted with permission from the authors.

A practitioner-as-researcher approach is recommended for all clinicians working with persons with mental retardation. There is an ongoing need to identify: (i) successful clinical intervention and treatment models, (ii) valid and reliable assessment and diagnostic procedures for psychopathology, and (iii) appropriate psychotropic medication treatments for various psychiatric conditions and challenging behaviors. Single-subject research designs can be integrated into a clinical practice and the outcomes of these studies provide a valuable source of information. Controlled, multisubject treatment-outcome designs are a more powerful study approach; they can validate the outcomes of single-subject research and increase generalizability (see Sulsky & Steiner, 1991 for a review of clinical research issues). Research conducted in cooperation with a psychiatrist helps determine the impact of a given medication on a targeted behavior or psychiatric symptom. Research conducted in cooperation with family members or care providers helps determine the impact of a given behavioral intervention or skills training strategy on a targeted behavior or psychiatric symptom. A critical aspect of the research process includes the development and validation of assessment and data-gathering tools. Assessing treatment or drug effects on a targeted behavior or psychiatric symptom requires trained professionals administering valid and reliable psychological assessments or standardized tests, as well as trained observers using valid and reliable observational tools (e.g., checklists, rating scales). The ongoing and careful study of therapeutic treatments and outcomes will enhance the scholarly literature in mental retardation, and ultimately improve the quality of life of persons with mental retardation, their families, friends, and care providers.

9.01.8 REFERENCES

Aman, M. G. (1991). *Assessing psychopathology and behavior problems in persons with mental retardation: A review of available instruments.* Rockville, MD: US Department of Health and Human Services.

Aman, M. G., & Singh, N. N. (1986). *Aberrant Behavior Checklist—Residential: Manual.* East Aurora, NY: Slosson Educational Publications.

Aman, M. G., & Singh, N. N. (1991). Pharmacological intervention. In J. L. Matson & J. A. Mulick (Eds.), *Handbook of mental retardation* (2nd ed., pp. 347–372). New York: Pergamon.

Aman, M. G., & Singh, N. N. (1994). *Aberrant Behavior Checklist—Community: Manual.* East Aurora, NY: Slosson Educational Publications.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association Committee on Research on Psychiatric Treatments (1992). Psychopharmacological screening criteria. *Journal of Clinical Psychiatry, 52,* 184–196.

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.

Anderson, D. J. (1989). Health care needs and residential settings: A national survey and policy perspectives. *Australia and New Zealand Journal of Developmental Disabilities, 15,* 289–300.

Bayley, N. (1969). *Bayley Scales of Infant Development.* New York: Psychological Corporation.

Beyer, H. A. (1988). Litigation and use of psychoactive drugs in developmental disabilities. In M. G. Aman & N. N. Singh (Eds.), *Psychopharmacology and the developmental disabilities* (pp. 29–57). New York: Springer.

Bisconer, S. W., Sine, L. F., & Zhang, X. Q. (1996). Prevalence and patterns of psychotropic medication use by adults with mental retardation living in community settings. *Journal of Developmental and Physical Disabilities, 8,* 291–311.

Bisconer, S. W., Zhang, X. Q., & Sine, L. F. (1995). Impact of a psychotropic medication and physical restraint review process on adults with mental retardation, psychiatric diagnoses, and challenging behaviors. *Journal of Developmental and Physical Disabilities, 7,* 123–135.

Cattell, P. (1950). *The measurement of intelligence of infants and young children.* New York: Psychological Corporation.

Civil Rights of Institutionalized Persons Act, 42 U.S.C. ξ 1997 (1981).

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test—Revised.* Circle Pines, MN: American Guidance Service.

Ellis, C. R., Singh, Y. N., & Singh, N. N. (1997). Use of behavior modifying drugs. In N. N. Singh (Ed.), *Prevention and treatment of severe behavior problems: Models and methods in developmental disabilities.* Pacific Grove, CA: Brooks/Cole.

Feldman, M. A., & Griffiths, D. (1997). Comprehensive assessment of severe behavior problems. In N. N. Singh (Ed.), *Prevention and treatment of severe behavior problems: Models and methods in developmental disabilities* (pp. 23–48). Pacific Grove, CA: Brooks/Cole.

Gadow, K. D., & Poling, A. G. (1988). *Pharmacotherapy and mental retardation.* Boston, MA: College-Hill.

Gesell, A. (1949). *Gesell Developmental Schedules.* New York: Psychological Corporation.

Health Care Financing Administration (1992). Medicare and medicaid programs: Omnibus nursing home requirements. Proposed rule. *Federal Register, 57,* 4516–4534.

Janicki, M. P., Dalton, A. J., Henderson, M., & Davidson, P. (1996). *Mortality and morbidity demographics of a 10 year cohort of older adults with intellectual disability in New York.* Paper presented at the 10th World Congress of International Association for the Scientific Study of Intellectual Disabilities, Helsinki, Finland.

Joint Commission on Accreditation of Healthcare Organizations (1995). *Accreditation manual for mental health, chemical dependency, and mental retardation developmental disabilities services.* Oakbrook Terrace, IL: Author.

Kalachnik, J. E., Leventhal, B. L., James, D. H., Sovner, R., Kastner, T. A., Walsh, K., & Klitzke, M. G. (1995). *Guidelines for the use of psychotropic medication.* Paper presented at the International Consensus Conference on Psychopharmacology, The Ohio State University, Columbus, OH.

Kaplan, H. I., Sadock, B. J., & Grebb, J. A. (1994). *Kaplan and Sadock's synopsis of psychiatry: Behavioral sciences, clinical psychiatry* (7th ed.). Baltimore, MD: Williams & Wilkins.

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children.* Circle Pines, MN: American Guidance Service.

Luckasson, R., Coulter, D. L., Polloway, E. A., Reiss, S., Schalock, R. S., Snell, M. E., Spitalnik, D. M., & Stark, J. A. (1992). *Mental retardation: Definition, classification, and systems of support.* Washington, DC: American Association on Mental Retardation.

Matson, J. L., Gardner, W. I., Coe, D. A., & Sovner, R. (1990). *Diagnostic Assessment for the Severely Handicapped (DASH) Scale (User manual).* Unpublished manuscript, Louisiana State University, Baton Rouge, LA.

Morgenstern, M., & Klass, E. (1991). Standard intelligence tests and related assessment techniques. In J. L. Matson & J. A. Mulick (Eds.), *Handbook of mental retardation* (2nd ed., pp. 195–210). New York: Pergamon.

Nihira, K., Foster, R., Shellhaas, M., & Leland, H. (1975). *AAMD Adaptive Behavior Scale Manual* (Rev. ed.). Washington, DC: American Association on Mental Deficiency.

Reiss, S. (1988). *Test manual for the Reiss Screen for Maladaptive Behavior.* Orland Park, IL: International Diagnostic Systems.

Reschly, D. J. (1987). *Adaptive behavior*. Tallahassee, FL: Florida Department of Education.

Reschly, D. J. (1990). Adaptive behavior. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (2nd ed., pp. 29–42). Washington, DC: National Association of School Psychologists.

Scheerenberger, R. C. (1983). *A history of mental retardation*. Baltimore, MD: Brookes.

Singh, N. N. (1995). Moving beyond institutional care for individuals with developmental disabilities. *Journal of Child and Family Studies, 4,* 129–145.

Singh, N. N. (1997a). *Prevention and treatment of severe behavior problems: Models and methods in developmental disabilities*. Pacific Grove, CA: Brooks/Cole.

Singh, N. N. (1997b). Enhancing quality of life through teaching and habilitation. In N. N. Singh (Ed.), *Prevention and treatment of severe behavior problems: Models and methods in developmental disabilities* (pp. 1–20). Pacific Grove, CA: Brooks/Cole.

Singh, N. N., Ellis, C. R., & Singh, Y. N. (1994). Medication management. In E. Cipani & F. Spooner (Eds.), *Curricular and instructional approaches for persons with severe handicaps* (pp. 404–423). Boston, MA: Allyn & Bacon.

Singh, N. N., Guernsey, T. F., & Ellis, C. R. (1992). Drug therapy for persons with developmental disabilities: Legislation and litigation. *Clinical Psychology Review, 12,* 665–679.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *The Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Service.

Sulsky, L. M., & Steiner, D. D. (1991). Design, analysis, and measurement issues in mental retardation research: An integrated approach. In J. L. Matson & J. A. Mulick (Eds.), *Handbook of mental retardation* (2nd ed., pp. 541–551). New York: Pergamon.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Guide for administering and scoring the Stanford–Binet Intelligence Scale: 4th Edition*. Chicago: Riverside Publishing.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale—Revised*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1989). *Manual for the Wechsler Preschool and Primary Scale of Intelligence—Revised*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children—Third Edition*. San Antonio, TX: The Psychological Corporation.

Wyatt v. Stickney, 344 F. Supp. 380 (M. D. Ala. 1972).

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.02

Pervasive Developmental Disorders

DONALD P. OSWALD
Virginia Commonwealth University, Richmond, VA, USA

9.02.1 INTRODUCTION 20

9.02.2 PHENOMENOLOGY 20
  9.02.2.1 Autistic Disorder 21
  9.02.2.2 Asperger's Disorder 21
  9.02.2.3 Childhood Disintegrative Disorder 22
  9.02.2.4 Rett's Disorder 22
  9.02.2.5 Pervasive Developmental Disorder Not Otherwise Specified 23
  9.02.2.6 Comorbid Conditions 23
    9.02.2.6.1 Mental retardation 23
    9.02.2.6.2 Seizure disorder 23
    9.02.2.6.3 Obsessive-compulsive disorder 23
    9.02.2.6.4 Schizophrenia 24
    9.02.2.6.5 Other behavior problems 24
  9.02.2.7 Adolescents and Adults with Pervasive Developmental Disorders 24

9.02.3 EPIDEMIOLOGY 25

9.02.4 ETIOLOGY 25
  9.02.4.1 Genetic Influences 25
  9.02.4.2 Prenatal/Perinatal Complications 25
  9.02.4.3 Immunologic Factors 26
  9.02.4.4 Microbial Infections 26
  9.02.4.5 Neuropsychological Features 26
  9.02.4.6 Neuroanatomic Features 26
  9.02.4.7 Neurochemical Findings 26

9.02.5 ASSESSMENT AND DIAGNOSIS 26

9.02.6 TREATMENT 28
  9.02.6.1 Structured Teaching and Behavior Modification Approaches 28
  9.02.6.2 Social Intervention Strategies 29
  9.02.6.3 Language Intervention Strategies 29
  9.02.6.4 Other Behavioral Interventions 29
  9.02.6.5 Psychopharmacology 29
    9.02.6.5.1 Serotonergic drugs 29
    9.02.6.5.2 Opiate antagonists 29
    9.02.6.5.3 Dopaminergic drugs 30
    9.02.6.5.4 Noradrenergic drugs 30
    9.02.6.5.5 Stimulants 30
    9.02.6.5.6 Nutritional supplements 30

9.02.7 SUMMARY 31

9.02.8 REFERENCES 31

9.02.1 INTRODUCTION

The pervasive developmental disorders (PDDs) are a class of neuropsychiatric syndromes that are typically identified in infancy or early childhood and persist throughout the lifespan. Characterized by marked developmental anomalies, particularly in the language and social domains, the PDDs represent a severe form of developmental psychopathology.

Autism or Kanner's syndrome (Kanner, 1943) is generally considered the prototypical PDD and has received the lion's share of attention over the past five decades. Recent developments in the field, however, have led to the conceptualization of a continuum of PDDs, or autism spectrum disorders, with a set of common core features. Current diagnostic systems generally include Asperger's disorder, Rett's disorder, childhood disintegrative disorder, and PDD not otherwise specified, along with autistic disorder, in the PDD continuum (American Psychiatric Association [APA], 1994).

A brief scan of the evolution of psychiatric nosology shows that, over the years, children with PDD have received a variety of diagnostic labels. Prior to Kanner's (1943) paper, the most commonly used terms were childhood psychosis, childhood schizophrenia, and borderline pathology of childhood. Children with characteristics that fit the PDD spectrum have also received a variety of other diagnoses, including infantile psychosis, infantile autism, symbiotic psychosis, autistic psychopathy, and schizoid and schizotypal personality disorders.

9.02.2 PHENOMENOLOGY

The pathognomonic behavioral feature of the PDDs is a marked impairment in social interaction skills. The social impairment is generally understood to be qualitative in nature; that is, individuals with PDD are not merely delayed in their social development but they display social interaction characteristics that are different from those seen in typically-developing persons of any age (van Engeland, Boldnár, & Bolhuis, 1985; Wenar, Ruttenberg, Kalish-Weiss, & Wolf, 1986). Persons in each of the subclasses of the PDD continuum manifest some form of social interaction abnormality, although the exact nature and extent of the impairment varies, in part, as a function of general intellectual ability.

The social interaction impairment displayed by individuals with a PDD is generally characterized as a failure to develop social reciprocity (APA, 1994). Social reciprocity is the ability to regulate one's behavior in social interactions as a function of the perceived verbal and nonverbal behavior and the attributed thoughts and feelings of the interaction partner. Deficits in social reciprocity have received considerable attention, particularly in persons with autism, and have sometimes been conceptualized as a failure of social cognitive development. One extensive line of related social cognition research has focused on the ability to perceive and interpret others' nonverbal expressions of emotion. Many studies have reported that persons with autism or other PDDs, even those whose general level of functioning is relatively high, show marked impairment in this ability, compared with nonautistic individuals with similar intellectual levels (e.g., Fein, Lucci, Braverman, & Waterhouse, 1992; Hobson, 1986; Macdonald et al., 1989).

Other researchers have described the affected social cognitive skill as a deficit in role-taking or perspective-taking ability (Dawson & Fernald, 1987; Oswald & Ollendick, 1989). Role-taking ability is the capacity to accurately predict what others are likely to be thinking or feeling, even when those thoughts and feelings are different from one's own. In recent years, a similar line of research has focused on the ability to conceive of mental states in oneself or others, commonly referred to as "a theory of mind" (Premack & Woodruff, 1978). Baron-Cohen and colleagues have produced a body of work that seeks to demonstrate the importance of theory of mind in understanding the social impairment of persons with autism. Persons with autism are said to display unique deficits in tasks that require a theory of mind, deficits that have been found to be associated with social interaction impairment (Baron-Cohen, Leslie, & Frith, 1985, 1986).

The theory-of-mind studies have generated substantial discussion in the literature. The key question that remains unanswered is whether theory-of-mind deficits are important in themselves in determining the profile of social abnormalities present in autism, or whether they represent artifacts that are determined by more fundamental social or language problems. Some recent data support the latter conclusion in that individuals with autism have been found to display a typical developmental sequence in the performance of theory-of-mind tasks and progress along that sequence is associated with verbal ability (Sparrevohn & Howie, 1995). In addition, evidence of social abnormalities is commonly seen in children well before even the earliest precursors of a theory of mind would be expected in the normal developmental progression (Klin & Volkmar, 1992). Finally, children with Down's syndrome have been found to display theory-of-mind deficits similar to those seen in persons with autism, calling into question the specificity of the deficit (Zelazo, Burack, Benedetto, & Frye, 1996).

Virtually all persons with PDD also manifest some form of communication impairment. The impairment may range from an apparent lack of all receptive and expressive language abilities to subtle peculiarities in the use or understanding of language that are difficult to characterize. One PDD subtype, Asperger's disorder, includes a diagnostic criterion emphasizing more or less age-appropriate speech development, although the language of persons with Asperger's disorder is nearly always marked by some unusual features, particularly with regard to social communication.

With the exception of Asperger's disorder, individuals with PDD also frequently present with intellectual functioning in the mentally retarded range (APA, 1994), and adaptive behavior deficits are even more common (Volkmar et al., 1987). Some authors have concluded that the empirical studies of adaptive behavior in persons with PDD support the presence of a specific pattern of adaptive skills that is common, if not unique. This pattern is characterized by relative deficits in adaptive skills in the socialization domain and relative strengths in the daily living skills domain (Carter, Gillham, Sparrow, & Volkmar, 1996). Such an adaptive skill pattern, however, may be primarily a reflection of the communication and social interaction problems inherent in the PDDs (Vig & Jedrysek, 1995).

The age at which symptoms associated with PDD first appear varies somewhat by subcategory but most typically falls within the first 36 months of life (Rogers & DiLalla, 1990). Autistic disorder is generally thought to be present from very early in infancy, if not from birth, and in most cases these individuals will display a disability that comes to the attention of their parents in the first year of life (Volkmar, Klin, Marans, & McDougle, 1996). The profoundly debilitating behavioral features of Rett's syndrome generally appear between 6 and 48 months of age, after 6–18 months of apparently normal psychomotor development (Tsai, 1994). Symptoms of childhood disintegrative disorder appear after a longer period of normal development (usually two to three years), but are generally apparent by age four (Volkmar, 1994). Persons with Asperger's disorder are generally not diagnosed until after age two, frequently much later, although their symptoms can usually be traced to unusual behaviors in early childhood (Klin, 1994).

Beyond the core feature of impaired social interaction, each of the PDDs manifests a slightly different presentation, although researchers and clinicians continue to debate the evidence regarding the distinctions among them. Nonetheless, the most widely accepted psychiatric taxonomies have converged on a set of four specific syndromes within the category (autistic disorder, Asperger's disorder, childhood disintegrative disorder, Rett's disorder) and one general designation for individuals who do not fit in any currently identified subcategory (pervasive developmental disorder not otherwise specified).

9.02.2.1 Autistic Disorder

Autistic disorder (AD) is the PDD subgroup that most closely matches Kanner's (1943) original description of autism. The defining characteristic of AD is deviant social development which may take the form of a lack of interest in other people, preferential interest in nonsocial aspects of the environment, failure to display differential attachment, and failure to develop typical prelinguistic communicative behaviors (Volkmar & Cohen, 1994).

Profound communication deficits are also an essential feature of AD, including failure to develop speech, echolalia, persistent pronoun reversal, extreme literalness, and unusual prosody (Paul, 1987). Even in those individuals with relatively intact intellectual function, language is a significant deficit area, particularly the understanding and use of social language.

The third major group of symptoms of AD has been designated "insistence on sameness" (APA, 1994). This term includes a variety of behavioral features associated with autism ranging from stereotyped motor movements to difficulty with transitions and preferential (even obsessive) interest in particular topics of conversation.

Kanner (1943) originally believed that most children with autism did not manifest mental retardation, a conclusion that was endorsed by others and led to the assertion that these individuals were merely unmotivated to engage in standardized intellectual assessment or that these instruments were somehow inappropriate for the population. Subsequent investigations, however, demonstrated that most individuals with AD also display general intellectual impairment (APA, 1994) and deficits in adaptive behavior (Schatz & Hamdan-Allen, 1995).

9.02.2.2 Asperger's Disorder

Asperger's disorder was originally described by Hans Asperger in 1944 (Tantum, 1988) and over the next 50 years appeared infrequently in the literature under the designation Asperger's syndrome. Individuals in the contemporary category of Asperger's disorder have been described in the literature under other designations as well, including mild autistic, high-functioning autistic, near-normal autistic, and schizotypal personality disorder. Recent work (Klin, Volkmar, Sparrow, Cicchetti, & Rourke, 1995) suggests substantial overlap between Asperger's disorder and two other diagnostic designations in contemporary literature: nonverbal learning disabilities syndrome (Rourke, 1989) and semantic-pragmatic disorder (Bishop, 1989).

The key features of Asperger's disorder, as described in the *Diagnostic and statistical manual of mental disorders* (4th ed., *DSM-IV*, APA, 1994) are (i) a qualitative impairment in social interaction; and (ii) restricted repetitive and stereotyped patterns of behavior, interests, and activities. Significant speech deficits are not part of the presentation, and persons with Asperger's disorder are less likely to display deficits in nonverbal concept formation, auditory perception, articulation, vocabulary, and verbal output than are high-functioning persons with autism (Klin et al., 1995). However, several unusual features in the communication domain are common in individuals with Asperger's disorder, including poor prosody, tangential and circumstantial speech, and marked verbosity on specific topics of interest (Klin & Volkmar, 1995).

In his original case descriptions, Asperger appears to have viewed clumsiness, or some other form of motor impairment, as a defining feature of the syndrome. Using high-functioning persons with autism as a comparison group, Asperger's disorder individuals have been reported to have greater difficulties with some motor tasks (Gillberg & Gillberg, 1989; Szatmari, Tuff, Finlayson, & Bartolucci, 1990). Reports of clumsiness in Asperger's disorder were sufficiently convincing that the feature was included among the diagnostic criteria for the syndrome in the *International classification of diseases* (10th ed., *ICD-10*), the most recent World Health Organization (1993) classification of diseases and health problems. Recent research findings differ on the question of whether motor skills deficits reliably distinguish individuals with Asperger's disorder from high-functioning individuals with autism (Klin et al., 1995; Manjivionna & Prior, 1995).

Other studies of the distinction between high-functioning autism and Asperger's disorder indicated that individuals with Asperger's disorder were more likely to engage in more pro-social behavior and less likely to use echolalic speech when young; their speech content was more likely to be repetitive and pedantic, without appropriate turn-taking; and they were likely to have collected facts about preferred subjects and to have asked questions or made statements on repetitive themes. Developmental features which distinguished individuals with Asperger's disorder included a higher verbal mental age and a greater mean age at diagnosis; they were also less likely to have experienced delayed language onset (Eisenmajer et al., 1996).

9.02.2.3 Childhood Disintegrative Disorder

The first description in the literature of the disorder currently designated as childhood disintegrative disorder (CDD) is commonly attributed to Heller, a Viennese special educator, who described six cases in 1908 (Volkmar, 1994). While there has been some variation in descriptions of CDD over the years, the condition is generally thought to be distinguished by several features including: (i) onset after at least two or three years of normal development, (ii) progressive deterioration in cognitive and behavioral functioning, and (iii) normal appearance and absence of signs of neurologic dysfunction (Volkmar, 1994). Some investigators have maintained that the criteria of at least two years of apparently normal development is not sufficiently documented in the literature on CDD (Perry, Cohen, & DeCarlo, 1995).

Once the disorder is established, it is virtually indistinguishable from autism in terms of behavioral characteristics. Individuals with CDD typically manifest speech loss or deterioration, social impairment, stereotypies and resistance to change, overactivity, affective symptoms and anxiety, and deterioration in self-help skills. In most cases of CDD, no specific neuropathological conditions associated with the onset of the disorder can be identified. In individuals who have a particularly late onset of CDD, however, such conditions are more frequently identified (Volkmar, 1992).

9.02.2.4 Rett's Disorder

The condition currently designated as Rett's disorder (RD) was first described by Rett in 1966 (Tsai, 1994). While aspects of the disorder are similar to autism (i.e., severe social and communication deficits), there are several key differences. RD has been reported only in females, while the other PDDs are substantially more common in males, and more individuals with RD fall into the severe range of mental retardation. Other key features of RD include: normal growth and development in the first months, followed by regression and decelerated head growth; loss of purposeful hand movement; and presence of characteristic stereotypies and gait abnormalities. There has been some controversy regarding whether RD should be considered one of the PDDs, but an emerging consensus supports this position (Volkmar, Klin, Marans, & Cohen, 1996).

The clinical course of individuals with RD has been described in detail by Hagberg and Witt-Engerström (1986). Four common stages of the development of the disorder include: (i) early-onset stagnation (6–18 months), (ii) rapid developmental regression beginning at 12–24 months, (iii) pseudostationary stage beginning at about three to four years, and (iv) late motor deterioration during the school-age years or adolescence.

Autistic-like features are most common in the period of rapid regression and are reported to diminish during the pseudostationary stage. There are indications that some individuals with RD maintain social interest and may develop social interaction skills as they grow older.

9.02.2.5 Pervasive Developmental Disorder Not Otherwise Specified

There has long been general acknowledgment that the PDDs represent a spectrum of conditions that include prototypical autism and several variations. *DSM-III* (APA, 1980) introduced the term "atypical PDD" as a means of classifying those patients that shared some of the features of autism but did not meet all of the diagnostic criteria. In subsequent editions of the *DSM*, the term has evolved to pervasive developmental disorder not otherwise specified (PDD NOS), and the category has been understood to include a variety of individuals. For example, children whose primary disability is severe pragmatic language dysfunction have sometimes been seen as fitting into the PDD NOS category (Towbin, 1994). Diagnostic criteria for PDD NOS are relatively nonspecific, as is typical of the NOS disorders.

One group of children in the PDD NOS category has been singled out for closer study in recent years. Designated as individuals with multiple complex developmental disorder (Cohen, Paul, & Volkmar, 1987), these children have been labeled with a variety of terms including childhood schizophrenia, childhood schizotypal disorder, schizotypal personality disorder, and borderline syndrome of childhood. The key features of multiple complex developmental disorder (MCDD) are defects in thinking, impaired social relatedness, and affect disregulation, all appearing by early childhood. Exploration of this proposed diagnostic category has yielded evidence supporting the view that the disorder is indeed best conceptualized as a developmental disorder, distinguishable from affective and disruptive behavior disorders of childhood (Towbin, Dykens, Pearson, & Cohen, 1993). The term is designed to avoid some of the semantic and conceptual problems raised by alternate labels such as "childhood schizophrenia."

9.02.2.6 Comorbid Conditions

9.02.2.6.1 Mental retardation

Samples of persons with autism typically show that about 80% have full-scale IQs in the mental retardation range (APA, 1994). Contrary to earlier notions that IQ was relatively meaningless in persons with autism, their scores on intelligence tests appear to be about as stable, reliable, and predictive as those for the general population (Volkmar, Klin, Marans, & McDougle, 1996).

Mental retardation is an exclusionary criterion for persons with Asperger's disorder, according to the *DSM-IV* definition (APA, 1994). Some reports, however, have suggested that persons with Asperger's disorder may present with full-scale IQs in the mild range of mental retardation (Gillberg & Gillberg, 1989). Persons with RD and CDD generally function in the moderate to profound range of mental retardation. PDD NOS is a heterogeneous category with respect to intellectual function but includes many individuals whose measured IQ falls in the mental retardation range.

9.02.2.6.2 Seizure disorder

Seizure disorders of all types are quite common in persons with a PDD. Seizures may develop during adolescence in up to 25% of individuals with AD, a pattern of onset that is somewhat unusual in the general population. The risk for seizures is increased for all PDD subgroups, but evidence suggests that individuals with prototypical autism are more likely to experience seizures than those with only some features of the disorder (Deykin & MacMahon, 1979).

9.02.2.6.3 Obsessive-compulsive disorder

The presence of repetitive behaviors is a very common feature in persons with autism that has long been included in descriptions of their behavioral features (Simons, 1974). Rumsey, Rapoport, and Sceery (1985) reported that 86% of their adult male sample demonstrated some form of compulsive behavior, including hand washing, arranging objects, repetitive tapping, phonic tics, and stereotyped touching of clothing or other objects. Half of the sample was also characterized by chronic, generalized anxiety. Many younger, lower-functioning children with autism display repetitive motor movements such as: rocking; toe-walking; arm, hand, or finger flapping; and whirling (Volkmar, Cohen, & Paul, 1986). These behaviors have a compulsive quality and may be associated with anxiety but their relationship to obsessive-compulsive disorder (OCD), if any, is unclear.

Obsessional thinking is also common in persons in the PDD spectrum. Such thinking is most common, or at least most readily detected, in highly verbal individuals. The content of obsessional thinking in PDD is likely to be idiosyncratic, though it may be similar in other respects to that described for OCD. Obsessive thoughts and compulsive behaviors in adults with autism and OCD have been found to differ in some respects from those seen in OCD patients who are not autistic; individuals with autism and OCD were more likely to display hoarding, touching, tapping, rubbing, and self-damaging behavior and were less likely to report repetitive thoughts involving aggression or symmetry or to engage in checking and counting (McDougle et al., 1995). Other anxiety symptoms are also frequently observed in individuals across the entire spectrum of PDDs; such symptoms include diffuse externalizing behavior problems in lower-functioning individuals, while persons with more typical intellectual skills may manifest obsessive worrying, psychomotor agitation, and excessive and rapid speech.

9.02.2.6.4 Schizophrenia

While not common, a number of case reports document the late-childhood or adolescent onset of schizophrenia in individuals with a clear-cut diagnosis of autism in childhood (Petty, Ornitz, Michelman, & Zimmerman, 1984). Most adults with autism are distinguishable from adults with schizophrenia, however, by the absence of positive symptoms of thought disorder (e.g., derailment and illogicality) although they manifest similar affective flattening (Rumsey, Andreasen, & Rapoport, 1986).

9.02.2.6.5 Other behavior problems

Hyperactivity is a commonly reported behavioral feature of children with PDD. Cross-sectional studies suggest that hyperactivity tends to decline from middle childhood to adolescence, although other behavior problems may persist or increase (Ando & Yoshimura, 1979). Stereotypy, self-injury, and pica are common associated behavior problems, particularly in children with marked intellectual impairment.

Many forms of behavior problems in persons with PDD have been hypothesized to serve communicative functions (Donnellan, Mirenda, Mesaros, & Fassbender, 1984). From this perspective, describing the exact form of the maladaptive behavior is less important than determining the behavior's function. The resulting functional analytic approach has been productive in expanding the nature and range of specific intervention strategies for problem behaviors in persons with PDD.

9.02.2.7 Adolescents and Adults with Pervasive Developmental Disorders

Some individuals with autism show improvement with the onset of puberty (Kanner, Rodriguez, & Ashenden, 1972), but adolescence is more commonly marked by increased difficulties (Gillberg, 1984). It is not uncommon for adolescents with no history of seizures to experience an onset of epilepsy during puberty, especially those with more severe intellectual impairment (Rutter, 1977). Problems with sexual behavior are quite common during adolescence, including public exposure and masturbation, and inappropriate sexual advances. Adolescents with autism, particularly those with higher IQs, may be vulnerable to depression as they become aware of the extent to which they differ from their peers. Behavioral problems may be exacerbated, especially in youth who experience marked physical changes as a result of puberty (Gillberg & Schumann, 1981).

The social impairment of the PDDs persists throughout the life-span. Although those affected individuals without significant intellectual impairment are generally able to improve their social functioning through therapeutic support and the investment of considerable effort, difficulties in negotiating social interactions are unlikely to disappear entirely. A substantial majority of persons with PDD require significant supports throughout their lives. Adaptive functioning is generally quite impaired, even relative to IQ, and most adults with autism require support and supervision in their employment and living arrangements (Rumsey et al., 1985). Historically, adults with autism have been at considerable risk for long-term institutionalization. One state-wide survey suggested that nearly 60% of the identified adults with autism were living in an institutional setting (Janicki & Jacobson, 1983).

In a description of adults with autism who had relatively good outcomes and were said to be "capable of functioning in society," Kanner et al. (1972, p. 27) reported that these adults continued to be characterized by social awkwardness and an inability to form personal friendships. One common feature across these good outcome cases was that all were said to have become aware, during their teen years, that they were different from their peers and made a conscious effort to address these differences. Most reports agree that even high-functioning adults with autism engage in social improprieties, would be perceived as unusual by a layperson, are generally characterized as loners, and lack social competencies and, in some cases, social motivation (Rumsey et al., 1985).

While long-range prediction of outcome for persons with PDD is an uncertain prospect, there are a few studies that offer some clues. The most powerful predictor of outcome appears to be measured IQ. Intellectual and language tests scores are quite stable in the population and are predictive of future educational placement (Freeman, Ritvo, Needleman, & Yokota, 1985). IQ has also been found to correlate with subsequent academic achievement and social behavior at home (Rutter & Bartak, 1973).

9.02.3 EPIDEMIOLOGY

AD occurs in approximately 5–10 cases per 10 000 children in the population (Bryson, 1996), and some researchers have indicated that prevalence of the disorder appears to be increasing (Volkmar, Klin, Marans, & McDougle, 1996) although this may be a function of increasingly sophisticated diagnostic procedures. Despite the vagaries of definition and diagnostic procedure, sex ratio figures of children with autism have remained remarkably stable over time: approximately four times as many males as females are diagnosed with autism (Wing, 1976). Evidence suggests, however, that females with autism are likely to have more severe intellectual impairment than males (Lord, Schopler, & Revicki, 1982).

Prevalence of persons with Asperger's disorder has been estimated at between 10 and 26 per 10 000, counting only individuals with at least average intelligence (Gillberg & Gillberg, 1989). Available evidence suggests that few persons with IQs in the mentally retarded range display the features of Asperger's disorder and, in fact, current diagnostic criteria preclude mental retardation as a comorbid condition (APA, 1994). Persons with Asperger's disorder are more likely to be males, but epidemiological data on the gender ratio are not yet available.

RD is observed in about one child out of 15 000. It is found in all races, and more than 1500 cases have been identified around the world. Generally considered to be a female-only disorder, there are a few cases in which RD was suspected in males. The disorder appears to be genetically determined, with spontaneous mutation the most likely cause (Hagberg & Witt-Engerström, 1986). The prevalence rate for CDD is thought to be only about 0.11 in 10 000, and the male:female ratio of reported cases is about eight to one (Volkmar, 1996).

In Kanner's (1943) original discussion of the case histories of his patients with autism, he described their families in terms that suggested they all enjoyed upper socioeconomic status. Although lacking any clear conceptual basis, reports of the association between autism and social class persisted in the literature for many years. Epidemiologic and clinical studies, however, have demonstrated that the relationship is illusory and have concluded that the early findings were likely the result of selection factors favoring the inclusion of the children of higher socioeconomic status families in published reports (Schopler, Andrews, & Strupp, 1979).

9.02.4 ETIOLOGY

9.02.4.1 Genetic Influences

Studies have shown that monozygotic twins show increased concordance for AD, compared with dizygotic twins, indicating that genetic factors play a role in the determination of the disorder. Concordance rates are not, however, 100% (Folstein & Rutter, 1977). Other relatives, particularly siblings, of persons with AD are also at increased risk for autism and other less severe developmental difficulties (Volkmar, Klin, Marans, & McDougle, 1996). The increased rate of occurrence of minor physical anomalies in persons with autism suggests a delay or disruption in early development that may contribute to the pathogenesis of the disorder; such a delay or disruption may result from a genetic predisposition, but no data exist at present to document such a pathway (Smalley & Collins, 1996).

One genetic abnormality associated with PDD is the fragile-X syndrome (Brøndum-Neilsen, 1983). Estimates of the prevalence of fragile-X in males with autism have ranged from 5% to 20%, and some researchers have maintained that all fragile-X individuals show some autistic features (Levitas et al., 1983). Other features apparently associated with the combination of autism and fragile-X syndrome are epilepsy, brainstem abnormalities, and a variety of neuropsychiatric symptoms, including elaborate repetitive routines and an aversion to being touched (Gillberg, Persson, & Wahlström, 1986).

9.02.4.2 Prenatal/Perinatal Complications

Folstein and Rutter (1977) reported that in a discordant twin pair the twin with autism was likely to be the one at greater obstetrical risk. Obstetric complications that are found with increased frequency in the histories of persons with autism and other developmental disabilities include prematurity, low birth weight, respiratory distress syndrome, rhesus incompatibility, emergency caesarean section, resuscitation, severe fetal/neonatal infection, hemolytic anemia, transfusion for anemia, gross physical abnormality in the fetus, and severe trauma during birth. Recent work, however, suggests that these complications probably represent epiphenomena of the condition or are the result of some shared risk factor, rather than playing a causal role (Bolton et al., 1997).

9.02.4.3 Immunologic Factors

There is some evidence for the role of immune system dysfunction in autism (Rimland & Baker, 1996). T-cell abnormalities have been found in children with autism (Warren, Yonk, Burger, Odell, & Warren, 1995), and immune booster treatment has been reported to have benefitted some such children (Reap & Lawson, 1990).

9.02.4.4 Microbial Infections

Anecdotal accounts of an onset of autistic-like features in infants following a course of high doses of antibiotics have led to speculation that PDDs may result from microbial infections. The antibiotics, by destroying the natural microorganisms in the intestinal tract, allow for the proliferation of harmful microbes such as *Candida albicans*. These organisms excrete neurotoxins that may cause symptoms like those seen in autism (Rimland & Baker, 1996). Empirical evidence for this etiological pathway to autism is, however, very limited.

9.02.4.5 Neuropsychological Features

Neuropsychological studies have documented impaired function in persons with autism in virtually every domain of cognition, including general intelligence, attention, memory, language, and executive function. In each domain, however, except for executive function, there are also functions that do not appear to be universally affected in persons with autism. Examples of commonly spared functions include visuospatial organization, sustained attention, paired associate learning, and others. Proposed neural substrates for some of the early neuropsychological symptoms of autism include the amygdala, the hippocampus, the cerebellum, the frontal lobe, and the basal ganglia (Dawson, 1996).

A study of nine, high-functioning autistic adult males indicated that, compared with normal controls, the autistic subjects showed significant weaknesses in conceptual problem-solving involving deficiencies in conceptual-level responding and a tendency to perseverate (Rumsey, 1985). A hypothesis focusing on weakness of frontal-subcortical functions is suggested by these data, but the wide variability in normative groups with demonstrable brain-damage precludes unequivocal interpretation of the findings.

9.02.4.6 Neuroanatomic Features

Morphological observations of the brains of persons with autism have revealed abnormalities in the brainstem and in the cerebellum (Hashimoto et al., 1995). Imaging studies have indicated enlargement of the parietal, temporal, and occipital lobes of the brains of individuals with autism (Piven, Arndt, Bailey, & Andreasen, 1996). Children with autism show frontal cortex abnormalities suggesting delayed frontal maturation (Zilbovicius et al., 1995).

Abnormal limbic system and cerebellum findings have also been reported for Asperger's disorder, although the abnormality is less pronounced; lesions in these areas of abnormality have been associated with symptoms similar to those seen in autism (Bauman, 1996). Animal models indicate that early medial temporal lobe damage involving the amygdala and the hippocampus produces abnormal social behaviors and stereotypies analogous to those seen in autism (Bachevalier, 1966).

RD appears to be associated with diffuse neuroanatomic abnormalities throughout the brain that are consistent with maturational failure (Bauman, 1996). A number of researchers have reported that macrocephaly is found in an unusually high proportion of persons with autism (e.g., Bailey, Luthert, Bolton, LeCouteur, & Rutter, 1993). However, this feature appears to be generally unrelated to most clinical characteristics (Lainhart et al., 1997).

9.02.4.7 Neurochemical Findings

Investigation of the neurochemistry of autism is a relatively nascent field. One finding that has been replicated is an increased level of serotonin in the platelets of children with autism (e.g., Anderson et al., 1987). Other promising methods of neurochemical investigation are under development but have yielded little in the area of PDD as yet (Cook, 1996).

9.02.5 ASSESSMENT AND DIAGNOSIS

Criteria for the diagnosis of autism have evolved over the past several decades, and this evolution has contributed to confusion and disagreement in the differential diagnosis of individuals with developmental disabilities. The use of latent class models in recent years, however, provides support for the assertion that diagnostic criteria have improved. PDD diagnostic criteria found in *ICD-10* and in *DSM-IV* have shown the best agreement with a "latent standard" compared with the criteria found in other recent definitions (Szatmari, Volkmar, & Walter, 1995).

The past four decades have seen the development of a wide range of diagnostic instruments and procedures that have contributed to improved diagnostic reliability. These tools include checklists, structured and semistructured interviews, and observation schedules.

The Autism Behavior Checklist (ABC) is a diagnostic screening instrument designed to examine behavioral characteristics common in autism (Krug, Arick, & Almond, 1980). The 57-item checklist was subsequently incorporated into the *Autism screening instrument for educational planning* (Krug, Arick, & Almond, 1993). Initial psychometric studies of the ABC were generally supportive and the instrument has received wide use, particularly in educational settings. A recent addition to the field, the Gilliam Autism Rating Scale (GARS; Gilliam, 1995), is a relatively brief, caregiver-report instrument designed to contribute to the diagnostic process for persons for whom autism is suspected. The GARS possesses reasonable psychometric properties and was normed on a sizeable sample ($N = 1092$) of individuals with autism.

One of the more promising directions in diagnostic procedures is the development of semi-structured interviews for caregivers of children and adults for whom PDD is a suspected diagnosis. The Autism Diagnostic Interview-Revised (ADI-R; Le Couteur et al., 1989) takes about 90 minutes to administer and includes items related to communication, social development and play, repetitive and restricted behaviors, and general behavior problems. The diagnostic algorithm for the ADI-R is designed to match the clinical descriptions and diagnostic criteria found in *DSM-IV* and *ICD-10* (APA, 1993; WHO, 1993). Reliability and validity data support the use of the ADI-R in distinguishing children with autism from nonautistic, mentally handicapped/language-impaired children (Lord, Rutter, & Le Couteur, 1994).

Many clinicians interested in the diagnosis of PDDs are reluctant to rely solely on parent report checklists and interview data as their primary sources of data. As a result, structured observation protocols have emerged as a way of standardizing clinician–patient interaction and eliciting diagnostically relevant behaviors. Among the first of such instruments for autism was the Childhood Autism Rating Scale (CARS; Schopler, Reichler, DeVellis, & Daly, 1980). The CARS consists of clinician ratings on 15 scales based on observation of the child during a diagnostic interview session. Substantial research has supported the validity of the CARS, and it continues to be one of the most widely used instruments of its kind.

Structured observation instruments have varied in terms of discriminant validity and typically performed best when distinguishing children with autism and severe mental retardation from normally developing or nonautistic children with mental retardation. The capacity of these instruments to identify higher-functioning persons with autism, however, has generally been found to be diminished.

Observation of social and communicative features of children and adults with mild or no intellectual impairment was the express goal of a structured observation system designated the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1989). The procedure provides a standardized set of contexts within which the observer can assess social and communicative behavior. The ADOS consists of eight tasks requiring about 20–30 minutes to complete. The tasks are intended to provide opportunities for specific diagnostically-relevant target behaviors, including asking for help, symbolic play, and reciprocal play. A downward extension and revision of the ADOS format has also been developed to assess children under six years of age who are suspected of having autism (DiLavore, Lord, & Rutter, 1995).

Even with the best diagnostic instruments and procedures, nothing can replace clinical experience with the population in producing an accurate, comprehensive evaluation. Further, an adequate assessment requires the knowledge and skills of parents, as well as professionals from a variety of disciplines, and an ability to coordinate those resources in a working team. In addition to the careful evaluation of diagnostic criteria, a comprehensive assessment of individuals with PDD will include at least three other major components: assessment of communication skills, psychological evaluation, and psychiatric assessment (Volkmar, Klin, Marans, & McDougle, 1996).

The assessment of communication in persons with PDD should not only include the traditional areas of articulation, vocabulary, and syntax but also emphasize those aspects of communication that are particularly affected by the condition. Specifically, attention is given to nonverbal communication (e.g., gaze, gestures), pragmatic intent and pragmatics in speech, and the use of language in social contexts. In individuals with more highly developed language, an evaluation of communication also incorporates assessment of subtle aspects of language such as the use of metaphors, humor, and irony and an understanding of ambiguity in communication (Klin & Shepard, 1994).

Key components of a psychological evaluation of individuals with PDD include assessment of intelligence, adaptive behavior, and personality as well as neuropsychological assessment (Klin & Shepard, 1994). There is abundant evidence that the use of standard intellectual instruments with individuals with PDD yields meaningful estimates of general intellectual ability and information regarding specific cognitive strengths and weaknesses. The fact that mental retardation is common in the population, however, frequently limits the choice of instruments, and the marked deficits in language often render portions of intelligence tests inaccessible to persons with PDD. Nonverbal instruments such as the Leiter International Performance Scale (Leiter, 1948) and tests of early development such as the Bayley Scales of Infant Development (Bayley, 1993) have often been used to derive estimates of general intellectual functioning when other scales do not include an adequate range of skills on the lower end of the developmental continuum.

Evaluation of adaptive behavior is less problematic, primarily because the most commonly used adaptive behavior assessment procedures rely on caretaker informants. The Vineland Adaptive Behavior Scale (Sparrow, Balla, & Cicchetti, 1984) and the AAMR Adaptive Behavior Scale (Lambert, Nihira, & Leland, 1993) are among the most commonly used instruments for the assessment of adaptive behavior in persons with PDD.

The evaluation of social/emotional or personality functioning of individuals with PDD often requires some adaptation of typical procedure. Evaluation procedures that rely on language are frequently of limited value because of the communication deficits inherent to the disorders. However, informant ratings of problem behaviors can be supplemented by clinical observation and structured or unstructured play assessment.

One of the more fruitful approaches to the assessment of problem behaviors in persons with PDD has been the functional analysis of behavior. Functional analysis is a conceptual framework for the assessment and treatment of challenging behaviors. The approach seeks to identify environmental stimuli and consequences that affect the frequency and intensity of such behaviors, with the goal of altering the environment or providing additional support to the individual. Functional analysis has a long history in the field of applied behavior analysis and has served to stimulate and support efforts to provide treatment that is nonpunitive and nonrestrictive with a focus on supporting adaptive behavior rather than eliminating problem behaviors (Singh, 1997). The functional analytic approach has yielded a variety of intervention strategies focusing on the hypothesized function of the maladaptive behavior. These strategies include intensive instruction in basic communication, extinction through the elimination of specific naturally-occurring reinforcers in the environment, and providing noncontingent attention.

9.02.6 TREATMENT

9.02.6.1 Structured Teaching and Behavior Modification Approaches

One of the most intensive and carefully studied approaches to structured teaching with children with autism was developed and originally described by Lovaas and colleagues (1980). The program was based on principles of operant reinforcement and involved intensive, highly structured instruction and behavior modification. Subjects, who were under four years of age at enrollment, were provided with one-on-one therapists for an average of 40 hours per week over a period of two or more years. Outcomes studies have documented clinically significant effects, including the finding that nearly half of the experimental group achieved normal intellectual and educational functioning at the post-treatment assessment (Lovaas, 1987). A long-term follow-up study of these children indicated that they had maintained their behavioral and intellectual progress, and continued to function much better than the original control group (McEachin, Smith, & Lovaas, 1993). A case report has also documented the success of an intervention based on Lovaas' work, with two children who showed significant deterioration similar to that seen in CDD (Perry et al., 1995).

Structured teaching also constitutes the core of the treatment approach provided by division TEACCH, the statewide program of services to individuals with autism in North Carolina (Mesibov, 1997). Division TEACCH has pioneered a range of teaching methods, particularly in the area of communication strategies, and TEACCH-related programs have yielded extensive documentation of the efficacy of these methods (e.g., Ozonoff & Cathcart, in press; Schopler, Brehm, Kinsbourne, & Reichler, 1971; Short, 1984). A hallmark of TEACCH's work has been the active and fruitful involvement of parents in the treatment and education of children with autism. TEACCH programs have demonstrated the importance and the impact of training parents in structured teaching methods; such parent training has been found to yield improved child behavior, parent–child interaction, and parental sense of control (Marcus, Lansing, Andrews, & Schopler, 1978). Outcome studies of TEACCH programs have documented their efficacy in terms of long-term placement in the community (Schopler, Mesibov, Devellis, & Short, 1981), vocational placement in supported employment (Mesibov, 1997), and parental satisfaction (Bristol & Schopler, 1983).

9.02.6.2 Social Intervention Strategies

Another approach to the treatment of children with autism has emphasized the importance of early social intervention, particularly with regard to peer influence in the development of interaction skills. Investigators at the Allegheny–Singer Research Institute have pursued a program of research that has served to clarify the type and level of support required for typically developing preschool-age children to elicit improved social interaction behaviors from peers with autism. Research in this area has explored differential responses to peer overtures (Kohler, Strain, & Shearer, 1992), the importance of group contingencies, and the role of specific instruction in corollary supportive peer behaviors (Kohler et al., 1995), and the role of teacher prompts (Odom & Strain, 1986). Some data suggest that a similar strategy can be applied in the home to increase positive interaction between young children with PDD and their siblings (Strain & Danko, 1995).

9.02.6.3 Language Intervention Strategies

While communication problems have always been an important component of the syndrome of autism, the notion that these problems could be explained in terms of semantic and conceptual deficits has been replaced by an emphasis on the pragmatics of language (Tager-Flusberg, 1981). The communication problems of children and adults with PDD have led researchers to propose a variety of language-related intervention strategies. Examples of such strategies include visual cuing (Quill, 1995) and direct instruction in pragmatic skills (Ogletree, Fischer, & Sprouse, 1995). Other developments in the field of language intervention include investigation of methods to overcome the effects of motor limitations on speech and communication development and the effects of a loss of speech or speech regression common in the history of many children with PDD (Prizant, 1996).

9.02.6.4 Other Behavioral Interventions

A variety of other types of interventions have emerged from the extensive use of behavior modification with persons with PDD. Behavioral interventions have shown promise in areas such as functional communication training (Carr & Durand, 1985), self-management in the treatment of social deficits and disruptive behavior (Koegel, Koegel, Hurley, & Frea, 1992), and the treatment of pica (Myles, Simpson, & Hirsch, 1997).

9.02.6.5 Psychopharmacology

Pharmacotherapy research related to PDDs has expanded dramatically in recent years. While it is unlikely that any drug will be equally effective for the entire PDD spectrum, researchers have begun to target key symptoms and subgroups for pharmacological intervention (McDougle, Price, & Volkmar, 1994).

9.02.6.5.1 Serotonergic drugs

Medications that affect levels of the neurotransmitter serotonin have been explored for individuals with autism in response to a finding that up to one-half of this population has blood platelet levels of serotonin at the 95th percentile or above. An initial report of the effectiveness of fenfluramine in three young children with autism (Geller, Ritvo, Freeman, & Yuwiler, 1982) prompted a flurry of related research activity. Controlled studies, however, have failed to replicate positive changes in core symptoms of PDD and suggested that fenfluramine may have negative effects on discrimination learning (August, Raz, & Baird, 1987; Campbell et al., 1988). In general, the evidence suggests that drugs that release serotonin or block its reuptake at the neuronal level may be useful in reducing aggression and repetitive behavior that interferes with functioning and in improving social relatedness in persons with autism (McDougle et al., 1994). However, a recent study of clomipramine, a tricyclic antidepressant and serotonin reuptake inhibitor, found that the drug was not therapeutic and was associated with significant side effects (Sanchez et al., 1996).

9.02.6.5.2 Opiate antagonists

In some persons with developmental disabilities who display self-injurious behavior, such behavior has been associated with elevated levels of endogenous opioids that are known to decrease pain sensitivity. This finding has led to trials of opiate antagonist drugs in the treatment of self-injury (Sandman et al., 1983). One such drug, naltrexone has been subjected to controlled studies which documented positive change in parent and teacher ratings of global improvement, restlessness, and impulsivity/hyperactivity (Kolmen, Feldman, Handen, & Janosky, 1995)

While opiate antagonists have shown promise in the treatment of self-injury in some persons with PDD, the findings are equivocal. Case reports have documented increases in self-injury when these medications are administered, the increases being commonly explained as extinction burst phenomena (e.g., Benjamin, Seek, Tresise, Price, & Gagnon, 1995). The extinction burst explanation is based on the conceptualization that when self-injurious behavior, previously reinforced by the natural secretion of endogenous opioids, is no longer reinforced, because of the suppression of opioid production, the behaviors may show the temporary increase in rate frequently associated with the removal of reinforcers.

9.02.6.5.3 Dopaminergic drugs

Drugs that affect the neurotransmitter dopamine have been found to decrease stereotyped behavior in persons with autism (Mikkelsen, 1982). Dopamine antagonists (e.g., thioridazine) have also been investigated in the treatment of pica. A controlled study has indicated that such drugs can actually worsen pica in persons with mental retardation (Singh, Ellis, Crews, & Singh, 1994).

Research on the effects of the dopaminergic drug risperidone in children and youth with PDD is encouraging, although the clinical trials are at a fairly early stage of development. An open-label trial of risperidone yielded clinically-significant global improvement and improvement in behavior symptoms (including repetitive behavior, aggression, and maladaptive behavior) in two-thirds of the subjects (McDougle et al., 1997).

Other dopamine antagonist drugs have long been used to good effect in the treatment of children and adults with PDD. The medication most systematically investigated in the treatment of autism has been haloperidol. This drug has been shown to facilitate learning and to reduce core symptoms and associated behavior of autism in some children (Campbell, Schopler, Cueva, & Hallin, 1996). One of the key drawbacks to the use of neuroleptics such as haloperidol has always been the occurrence of tardive dyskinesia (TD) associated with long-term use. TD has been reported to occur in about one-third of patients with autism. Susceptibility to TD may be associated with female gender and pre- and perinatal complications, and risk is increased by greater cumulative doses and longer exposure (Campbell et al., 1997).

There is considerable evidence that dopamine antagonists are effective in reducing maladaptive behavior in many children with autism. However, the older drugs in this class frequently produce dyskinesias, rendering them relatively less desirable. Newer drugs that affect dopamine transmission have not been sufficiently tested but may offer similar benefits without the substantial side effects (McDougle et al., 1994).

9.02.6.5.4 Noradrenergic drugs

Preliminary data indicate that clonidine, an alpha noradrenergic receptor agonist, may be useful in treating hyperactivity and irritability and that beta blockers may help control aggressive and self-injurious behaviors (Lewis, 1996). Few controlled studies have been conducted, however, and the existing data regarding the effects of this class of drugs on the core symptoms of autism are not particularly encouraging (McDougle et al., 1997).

9.02.6.5.5 Stimulants

Early clinical trials of the use of stimulants with children with autism were generally discouraging, indicating increased irritability and stereotypic movements (Campbell et al., 1972). Criticism of this work, however, on the basis of a heterogeneous sample and possible interactions with neuroleptics, has prompted additional investigation in the area. More recent studies have supported the use of methylphenidate (Ritalin) in the treatment of hyperactivity in children with autism and have not found the untoward side effects previously reported (Birmaher, Quintana, & Greenhill, 1988; Quintana et al., 1995).

9.02.6.5.6 Nutritional supplements

The most extensively investigated nutrition-related treatment of autism is the use of supplements of vitamin B_6 and magnesium. An early, double-blind crossover study (Rimland, Callaway, & Dreyfus, 1978) provided the impetus for considerable research in the area, leading to claims that more than 40% of the children with autism receiving this "megavitamin therapy" showed improvement (Rimland, 1988).

In general, the efficacy of nutritional supplements in the treatment of core symptoms of PDD remains in question. While testimonials and anecdotal evidence abound, documented responses in controlled studies have been relatively scarce, suggesting that the improvement in some individuals within the PDD spectrum may be based on idiosyncratic factors rather than reliable response mechanisms inherently related to the disability.

Nutritional supplements have been used to some effect in the treatment of specific symptoms common in autism. A serum zinc supplement and a general dietary supplement, for example, have been found to be effective in reducing pica in studies of individuals with mental retardation (Bugle & Rubin, 1993; Lofts, Schroeder, & Maier, 1990).

9.02.7 SUMMARY

Investigation related to the diagnosis, assessment, and treatment of persons with PDDs has been a tremendously fruitful area since the 1960s. Diagnostic differentiation is becoming increasingly sophisticated, instruments that can contribute to a comprehensive assessment are more numerous and more appropriate to the population, and information regarding treatment alternatives with solid empirical support is slowly being developed and disseminated.

Some promising emerging areas of research are in the burgeoning field of psychopharmacology and in the maturing study of ecological factors in the determination of behavior. The investigation of genetic, prenatal, and perinatal factors related to PDDs has also yielded many fruitful avenues for new research. Neuroanatomic studies continue to tantalize the field and represent a slow, steady accretion of knowledge that holds hope for tangible results. Finally, an increasing emphasis on early identification of persons in the PDD spectrum offers an opportunity for the development of more effective early intervention strategies.

PDDs are among the most puzzling and rewarding challenges to contemporary clinical psychology. As the knowledge base expands, the prospects for accurate diagnosis, meaningful assessment, and effective intervention steadily improve and, with them, the possibility that individuals with PDDs will suffer less from the effects of their disabilities and will enjoy an improved quality of life.

9.02.8 REFERENCES

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anderson, G. M., Freedman, D. X., Cohen, D. J., Volkmar, F. R., Hoder, E. L., McPhedran, P., Minderaa, R. B., Hansen, C. R., & Young, J. R. (1987). Whole blood serotonin in autistic and normal subjects. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 28,* 885–900.

Ando, H., & Yoshimura, I. (1979). Effects of age on communication skill levels and prevalence of maladaptive behaviors in autistic and mentally retarded children. *Journal of Autism and Developmental Disorders, 9,* 83–93.

August, G. J., Raz, N., & Baird, T. D. (1987). Fenfluramine response in high and low functioning autistic children. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 342–346.

Bachevalier, J. (1966). Medial temporal lobe and autism: A putative animal model in primates. *Journal of Autism and Developmental Disorders, 26,* 217–220.

Bailey, A., Luthert, P., Bolton, P., LeCouteur, A., & Rutter, M. (1993). Autism and megalencephaly. *Lancet, 341,* 1225–1226.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21,* 37–46.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1986). Mechanical, behavioural and intentional understanding of picture stories in autistic children. *British Journal of Developmental Psychology, 4,* 113–125.

Bauman, M. L. (1996). Neuroanatomic observations of the brain in pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 26,* 199–203.

Bayley, N. (1993). *Bayley Scales of Infant Development* (2nd ed.). San Antonio, TX: Psychological Corporation.

Benjamin, S., Seek, A., Tresise, L., Price, E., & Gagnon, M. (1995). Paradoxical response to naltrexone treatment of self-injurious behavior. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 238–242.

Birmaher, B., Quintana, H., & Greenhill, L. (1988). Methylphenidate treatment of hyperactive autistic children. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 248–251.

Bishop, D. V. M. (1989). Autism, Asperger's syndrome, and semantic-pragmatic disorder: Where are the boundaries? *British Journal of Disorders of Communication, 24,* 107–121.

Bolton, P. F., Murphy, M., Macdonald, H., Whitlock, B., Pickles, A., & Rutter, M. (1997). Obstetric complications in autism: Consequences or causes of the condition. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 272–281.

Bristol, M. M., & Schopler, E. (1983). Stress and coping in families of autistic adolescents. In E. Schopler & G. B. Mesibov (Eds.) *Autism in adolescents and adults* (pp. 251–278). New York: Plenum.

Brøndum-Neilsen, K. (1983). Diagnosis of the fragile-X syndrome (Martin-Bell syndrome): Clinical findings in 27 males with the fragile site at Xq28. *Journal of Mental Deficiency Research, 27,* 211–226.

Bryson, S. E. (1996). Epidemiology of autism. *Journal of Autism and Developmental Disorders, 26*, 165–167.

Bugle, C., & Rubin, H. (1993). Effects of a nutritional supplement on coprophagia: A study of three cases. *Research in Developmental Disabilities, 14*, 445–456.

Campbell, M., Adams, P., Small, A. M., Curren, E. L., Overall, J. E., Anderson, L. T., Lynch, S., & Perry, R. (1988). Efficacy and safety of fenfluramine in autistic children. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 434–439.

Campbell, M., Armenteros, J. L., Malone, R. P., Adams, P. B., Eisenberg, Z. W., & Overall, J. E. (1997). Neuroleptic-related dyskinesias in autistic children: A prospective, longitudinal study. *Journal of the American Academy of Child Psychiatry, 36*, 835–843.

Campbell, M., Fish, B., David, R., Shapiro, T., Collins, P., & Koh, C. (1972). Response to triiodothyronione and dextroamphetamine: A study of preschool schizophrenic children. *Journal of Autism and Childhood Schizophrenia, 2*, 343–358.

Campbell, M., Schopler, E., Cueva, J. E., & Hallin, A. (1996). Treatment of autistic disorder. *Journal of the American Academy of Child Psychiatry, 35*, 134–143.

Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis, 18*, 111–126.

Carter, A. S., Gillham, J. E., Sparrow, S. S., & Volkmar, F. R. (1996). Adaptive behavior in autism. *Child and Adolescent Psychiatric Clinics of North America, 4*, 945–961.

Cohen, D. J., Paul, R., & Volkmar, F. R. (1987). Issues in the classification of pervasive developmental disorders and associated conditions. In D. J. Cohen & A. Donellan (Eds.), *Handbook of autism and pervasive developmental disorders* (pp. 20–40). New York: Wiley.

Cook, E. H. (1996). Pathophysiology of autism: Neurochemistry. *Journal of Autism and Developmental Disorder, 26*, 221–225.

Dawson, G. (1996). Neuropsychology of autism: A report on the state of the science. *Journal of Autism and Developmental Disorders, 26*, 179–184.

Dawson, G., & Fernald, M. (1987). Perspective-taking ability and its relationship to the social behavior of autistic children. *Journal of Autism and Developmental Disorders, 17*, 487–498.

Deykin, E. Y, & MacMahon, B. (1979). The incidence of seizures among children with autistic symptoms. *American Journal of Psychiatry, 126*, 1310–1312.

DiLavore, P. C., Lord, C., & Rutter, M. (1995). The Prelinguistic Autism Diagnostic Observation Schedule. *Journal of Autism and Developmental Disorders, 25*, 355–379.

Donnellan, A. M., Mirenda, P. L., Mesaros, R. A., & Fassbender, L. L. (1984). Analyzing the communicative functions of aberrant behavior. *Journal of the Association for Persons with Severe Handicaps, 9*, 201–212.

Eisenmajer, R., Prior, M., Leekam, S., Wing, L., Gould, J., Welham, M., & Ong, B. (1996). Comparison of clinical symptoms in autism and Asperger's Disorder. *Journal of the American Academy of Child Psychiatry, 35*, 1523–1531.

Fein, D., Lucci, D., Braverman, M., & Waterhouse, L. (1992). Comprehension of affect in context in children with pervasive developmental disorders. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 33*, 1157–1167.

Freeman, B. J., Ritvo, E. R., Needleman, R., & Yokota, A. (1985). The stability of cognitive and linguistic parameters in autism: A five-year prospective study. *Journal of the American Academy of Child Psychiatry, 24*, 459–464.

Folstein, S., & Rutter, M. (1977). Infantile autism: A genetic study of 21 twin pairs. *Journal of Child Psychology and Psychiatry, 18*, 297–321.

Geller, E., Ritvo, E., Freeman, B., & Yuwiler, A. (1982). Preliminary observations on the effect on fenfluramine on blood serotonin and symptoms in three autistic boys. *New England Journal of Medicine, 307*, 165–169.

Gillberg, C. (1984). Autistic children growing up: Problems during puberty and adolescence. *Developmental Medicine and Child Neurology, 26*, 122–129.

Gillberg, C., & Schaumann, H. (1981). Infantile autism and puberty. *Journal of Autism and Developmental Disorders, 11*, 365–371.

Gillberg, C., Persson, E., & Wahlström, J. (1986). The autism–fragile-X syndrome (AFRAX): A population-based study of ten boys. *Journal of Mental Deficiency Research, 30*, 27–39.

Gillberg, I. C. & Gillberg, C. (1989). Asperger syndrome—some epidemiological considerations: A research note. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 30*, 631–638.

Gilliam, J. E. (1995). *Gilliam Autism Rating Scale: Examiner's manual.* Austin, TX: Pro-Ed.

Hagberg, B. A., & Witt-Engerström, I. (1986). Rett syndrome: A suggested staging system for describing impairment profile with increasing age toward adolescence. *American Journal of Medical Genetics, 24*, 47–59.

Hashimoto, T., Tayama, M., Murakawa, K., Yoshimoto, T., Miyazaki, M., Harada, M., & Kuroda, Y. (1995). Development of the brainstem and cerebellum in autistic patients. *Journal of Autism and Developmental Disorders, 25*, 1–18.

Hobson, R. P. (1986). The autistic child's appraisal of expressions of emotion: A further study. *Journal of Child Psychology and Psychiatry, 27*, 671–680.

Janicki, M. P., & Jacobson, J. W. (1983). Selected clinical features and service characteristics of autistic adults. *Psychological Reports, 52*, 387–390.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*, 217–250.

Kanner, L., Rodriguez, A., & Ashenden, B. (1972). How far can autistic children go in matters of social adaptation? *Journal of Autism and Childhood Schizophrenia, 2*, 9–33.

Klin, A. (1994). Asperger syndrome. *Child and Adolescent Psychiatric Clinics of North America, 3*, 131–148.

Klin, A. & Shepard, B. A. (1994). Psychological assessment of autistic children. *Child and Adolescent Psychiatric Clinics of North America, 3*, 53–70.

Klin, A., & Volkmar, F. R. (1992). Autistic social dysfunction: Some limitations of the theory of mind hypothesis. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 33*, 861–876.

Klin, A., & Volkmar, F. R. (1995). *Asperger syndrome: Some guidelines for assessment, diagnosis, and intervention.* Pittsburgh, PA: Learning Disabilities Association of America.

Klin, A., Volkmar, F. R., Sparrow, S. S., Cicchetti, D. V., & Rourke, B. P. (1995). Validity and neuropsychological characterization of Asperger Syndrome: Convergence with nonverbal learning disabilities syndrome. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 36*, 1127–1139.

Koegel, L. K., Koegel, R. L., Hurley, C., & Frea, W. D. (1992). Improving social skills and disruptive behavior in children with autism through self-management. *Journal of Applied Behavior Analysis, 25*, 341–353.

Kohler, F. W., Strain, P. S., Hoyson, M., Davis, L., Donina, W. M., & Rapp, N. (1995). Using a group-oriented contingency to increase social interactions between children with autism and their peers. *Behavior Modification, 19*, 10–32.

Kohler, F. W., Strain, P. S., & Shearer, D. D. (1992). The overtures of preschool social skill intervention agents: Differential rates, forms, and functions. *Behavior Modification, 16*, 525–542.

Kolmen, B. K., Feldman, H. M., Handen, B. L., & Janosky, J. E. (1995). Naltrexone in young autistic children: A double-blind, placebo-controlled crossover study. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*, 223–231.

Krug, D. A., Arick, J., & Almond, P. (1980). Behavior checklist for identifying severely handicapped individuals with high levels of autistic behavior. *Journal of Child Psychology and Psychiatry, 21*, 221–229.

Krug, D. A., Arick, J. R., & Almond, P. J. (1993). *Autism screening instrument for educational planning (2nd ed.): Examiner's manual.* Austin, TX: Pro-Ed.

Lainhart, J. E., Piven, J., Wzorek, M., Landa, R., Santangelo, S. L., Coon, H., & Folstein, S. E. (1997). Macrocephaly in children and adults with autism. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 282–290.

Lambert, N., Nihira, K., & Leland, H. (1993). *Adaptive Behavior Scale—School* (2nd ed.). Austin, TX: Pro-Ed.

Le Couteur, A., Rutter, M., Lord, C., Rios, P., Robertson, S., Hodgrafer, M., & McLennan, J. D. (1989). Autism Diagnostic Interview: A semi-structured interview for parents and caregivers of autistic persons. *Journal of Autism and Developmental Disorders, 19,* 363–387.

Leiter, R. G. (1948). *Leiter International Performance Scale.* Chicago: Stoelting.

Levitas, A. L., Hagerman, R. J., Braden, M., Rimland, B., McBogg, P., & Matus, I. (1983). Autism and the fragile-X syndrome. *Journal of Developmental and Behavioral Pediatrics, 4,* 151–158.

Lewis, M. H. (1996). Psychopharmacology of autism spectrum disorders. *Journal of Autism and Developmental Disorders, 26,* 231–235.

Lofts, R. Schroeder, S., & Maier, R. (1990). The effects of serum zinc supplementation on the pica behavior of a person with mental retardation. *American Journal of Mental Retardation, 95,* 103–109.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., & Schopler, E. (1989). Autism Diagnostic Observation Schedule: A standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders, 19,* 185–212.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24,* 659–685.

Lord, C., Schopler, E., & Revicki, D. (1982). Sex differences in autism. *Journal of Autism and Developmental Disorder, 12,* 317–330.

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55,* 3–9.

Lovaas, O. I., Ackerman, A. B., Alexander, D., Firestone, P., Perkins, J., & Young, D. (1980). *Teaching developmentally disabled children: The me book.* Austin, TX: Pro-Ed.

Macdonald, H., Rutter, M., Howlin, P., Rios, P., Le Conteur, A., Evered, C., & Folstein, S. (1989). Recognition and expression of emotional cues by autistic and normal adults. *Journal of Child Psychology and Psychiatry, 30,* 865–877.

Manjivionna, J., & Prior, M. (1995). Comparison of Asperger Syndrome and high-functioning autistic children on a test of motor impairment. *Journal of Autism and Developmental Disorders, 25,* 23–39.

Marcus, L., Lansing, M., Andrews, C., & Schopler, E. (1978). Improvement of teaching effectiveness in parents of autistic children. *Journal of the American Academy of Child and Adolescent Psychiatry, 17,* 625–639.

McDougle, C. J., Holmes, J. P., Bronson, M. R., Anderson, G. M., Volkmar, F. R., Price, L. H., & Cohen, D. J. (1997). Risperidone treatment of children and adolescents with Pervasive Developmental Disorders: A prospective, open-label study. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 685–693.

McDougle, C. J., Kresch, L. E., Goodman, W. K., Naylor, S. T., Volkmar, F. R., Cohen, D. J., & Price, L. H. (1995). A case-controlled study of repetitive thoughts and behavior in adults with autistic disorder and obsessive-compulsive disorder. *American Journal of Psychiatry, 152,* 772–777.

McDougle, C. J., Price, L. H., & Volkmar, F. R. (1994). Recent advances in the pharmacotherapy of autism and related conditions. *Child and Adolescent Psychiatric Clinics of North America, 3,* 71–89.

McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation, 97,* 359–372.

Mesibov, G. (1997). Formal and informal measures on the effectiveness of the TEACCH programme. *Autism: The International Journal of Research and Practice, 1,* 25–35.

Mikkelsen, E. (1982). Efficacy of neuroleptic medication in pervasive developmental disorders of childhood. *Schizophrenia Bulletin, 8,* 320–328.

Myles, B. S., Simpson, R. L., & Hirsch, N. C. (1997). A review of literature on interventions to reduce pica in individuals with developmental disabilities. *Autism: The International Journal of Research and Practice, 1,* 77–95.

Odom, S. L. & Strain, P. S. (1986). A comparison of peer-initiation and teacher-antecedent interventions for promoting reciprocal social interaction of autistic preschoolers. *Journal of Applied Behavior Analysis, 19,* 59–71.

Ogletree, B. T., Fischer, M. A., & Sprouse, J. (1995). An innovative language treatment for a child with high-functioning autism. *Focus on Autistic Behavior, 10,* 1–10.

Oswald, D. P. & Ollendick, T. H. (1989). Role taking and social competence in autism and mental retardation. *Journal of Autism and Developmental Disorders, 19,* 119–127.

Ozonoff, S., & Cathcart, K. (in press). Effectiveness of a home program intervention for young children with autism. *Journal of Autism and Developmental Disorders.*

Paul, R. (1987). Communication in autism. In D. Cohen & A. Donnellan (Eds.), *Handbook of autism and pervasive developmental disorders* (pp. 121–132). New York: Wiley.

Perry, R., Cohen, I., & DeCarlo, R. (1995). Deterioration, autism, and recovery in two siblings. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 232–237.

Petty, L. K., Ornitz, E. M., Michelman, J. D., & Zimmerman, E. G. (1984). Autistic children who become schizophrenic. *Archives of General Psychiatry, 41,* 129–135.

Piven, J., Arndt, S., Bailey, J., & Andreasen, N. (1996). Regional brain enlargement in autism: A magnetic resonance imaging study. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 530–536.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behavioural and Brain Sciences, 4,* 515–526.

Prizant, B. M. (1996). Communication, language, social, and emotional development. *Journal of Autism and Developmental Disorders, 26,* 173–178.

Quill, K. A. (1995). Visually cued instruction for children with autism and pervasive developmental disorders. *Focus on Autistic Behavior, 10,* 10–20.

Quintana, H., Birmaher, B., Stedge, D., Lennon, S., Freed, J., Bridge, J., & Greenhill, L. (1995). Use of methylphenidate in the treatment of children with autistic disorder. *Journal of Autism and Developmental Disorders, 25,* 283–294.

Reap, E. A., & Lawson, J. W. (1990). Stimulation of the immune response by dimethylglycine, a nontoxic metabolite. *Journal of Laboratory Clinical Medicine, 115,* 481–486.

Rimland, B. (1964). *Infantile autism.* Englewood Cliffs, NJ: Prentice-Hall.

Rimland, B. (1988). Controversies in the treatment of autistic children: Vitamin and drug therapy. *Journal of Child Neurology, 3*(Suppl.), S68–S72.

Rimland, B., & Baker, S. M. (1996). Alternative approaches to the development of effective treatments for autism. *Journal of Autism and Developmental Disorders, 26,* 237–241.

Rimland, B., Callaway, E., & Dreyfus, P. (1978). The effect of high doses of Vitamin B-sub-6 on autistic children: A double-blind crossover study. *American Journal of Psychiatry, 135,* 472–475.

Rogers, S. J., & DiLalla, D. L. (1990). Age of symptom onset in young children with pervasive developmental disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 863–872.

Rourke, B. P. (1989). *Nonverbal learning disabilities: The syndrome and the model*. New York: Guilford Press.

Rumsey, J. M. (1985). Conceptual problem-solving in highly verbal, nonretarded autistic men. *Journal of Autism and Developmental Disorders, 15,* 23–36.

Rumsey, J. M., Andreasen, N. C., & Rapoport, J. L. (1986). Thought, language, communication, and affective flattening in autistic adults. *Archives of General Psychiatry, 43,* 771–777.

Rumsey, J. M., Rapoport, J. L., & Sceery, W. R. (1985). Autistic children as adults: Psychiatric, social, and behavioral outcomes. *Journal of the American Academy of Child Psychiatry, 24,* 465–473.

Rutter, M. (1977). Infantile autism and other child psychoses. In M. Rutter and L. Hersov (Eds.), *Child psychiatry: Modern approaches* (pp. 717–747). Oxford, UK: Blackwell.

Rutter, M., & Bartak, L. (1973). Special educational treatment of autistic children: A comparative study. II: Follow-up findings and implications for services. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 14,* 241–270.

Sanchez, L. E., Campbell, M., Small, A. M., Cueva, J. E., Armenteros, J. L., & Adams, P. B. (1996). A pilot study of clomipramine in young autistic children. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 537–544.

Sandman, C. A., Data, P. C., Barron, J., Hoehler, F. K., Williams, C., & Swanson, J. M. (1983). Naloxone attenuates self-abusive behavior in developmentally disabled clients. *Applied Research in Mental Retardation, 4*, 5–11.

Schatz, J., & Hamdan-Allen, G. (1995). Effects of age and IQ on adaptive behavior domains for children with autism. *Journal of Autism and Developmental Disorders, 25*, 51–60.

Schopler, E., Andrews, C. E., & Strupp, K. (1979). Do autistic children come from upper-middle class parents? *Journal of Autism and Developmental Disorders, 9*, 139–152.

Schopler, E., Brehm, S., Kinsbourne, M., & Reichler, R. J. (1971). The effect of treatment structure on development in autistic children. *Archives of General Psychiatry, 24*, 415–421.

Schopler, E., Mesibov, G. B., Devellis, R. F., & Short, A. (1981). Treatment outcome for autistic children and their families. In P. Mittler (Ed.), *Frontiers of knowledge in mental retardation: Social, educational, and behavioral aspects* (pp. 293–301). Baltimore: University Park Press.

Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders, 10*, 91–103.

Short, A. B. (1984). Short-term treatment outcome using parents as cotherapists for their own autistic children. *Journal of Child Psychiatry and Allied Disciplines, 25*, 443–458.

Simons, J. M. (1974). Observations on compulsive behavior in autism. *Journal of Autism and Childhood Schizophrenia, 4*, 1–10.

Singh, N. N. (1997). *Prevention and treatment of severe behavior problems: Models and methods in developmental disabilities*. Pacific Grove, CA: Brooks/Cole.

Singh, N. N., Ellis, C. R., Crews, W., & Singh, Y. (1994). Does diminished dopaminergic neurotransmission increase pica? *Journal of Child and Adolescent Psychopharmacology, 4*, 93–99.

Smalley, S. L., & Collins, F. (1996). Genetic, prenatal, and immunologic factors. *Journal of Autism and Developmental Disorders, 26*, 195–198.

Sparrevohn, R. & Howie, P. M. (1995). Theory of mind in children with autistic disorder: Evidence of developmental progression and the role of verbal ability. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 36*, 249–263.

Sparrow, S., Balla, D., & Cicchetti, D. (1984). *Vineland Adaptive Behavior Scales Expanded Form*. Circle Pines, MN: American Guidance Service.

Strain, P. S., & Danko, C. D. (1995). Caregivers' encouragement of positive interaction between preschoolers with autism and their siblings. *Journal of Emotional and Behavioral Disorders, 3*, 2–12.

Szatmari, P., Tuff, L., Finlayson, A. J., & Bartolucci, G. (1990). Asperger's syndrome and autism: Neurocognitive aspects. *Journal of the American Academy of Child & Adolescent Psychiatry, 29*, 130–136.

Szatmari, P., Volkmar, F., & Walter, S. (1995). Evaluation of diagnostic criteria for autism using latent class models. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*, 216–222.

Tager-Flusberg, H. (1981). On the nature of linguistic functioning in early infantile autism. *Journal of Autism and Developmental Disorders, 11,* 45–56.

Tantum, D. (1988). Asperger's syndrome. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 29,* 245–255.

Towbin, K. E. (1994). Pervasive developmental disorder not otherwise specified: A review and guidelines for clinical care. *Child and Adolescent Psychiatric Clinics of North America, 3,* 149–160.

Towbin, K. E., Dykens, E. M., Pearson, G. S., & Cohen, D. J. (1993). Conceptualizing "borderline syndrome of childhood" and "childhood schizophrenia" as a developmental disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 775–782.

Tsai, L. Y. (1994). Rett syndrome. *Child and Adolescent Psychiatric Clinics of North America, 3,* 105–118.

van Engeland, H., Boldnàr, F. A., & Bolhuis, G. (1985). Some qualitative aspects of the social behaviour of autistic children: An ethological approach. *Journal of Child Psychology and Psychiatry, 26,* 879–893.

Vig, S., & Jedrysek, E. (1995). Adaptive behavior of young urban children with developmental disabilities. *Mental Retardation, 33,* 90–98.

Volkmar, F. R. (1992). Childhood disintegrative disorders: Issues for DSM-IV. *Journal of Autism and Developmental Disorders, 22,* 625–642.

Volkmar, F. R. (1994). Childhood disintegrative disorder. *Child and Adolescent Psychiatric Clinics of North America, 3,* 119–129.

Volkmar, F. R. (1996). The disintegrative disorders: Childhood disintegrative disorder and Rett's disorder. In F. R. Volkmar (Ed.), *Psychoses and pervasive developmental disorders in childhood and adolescence* (pp. 223–248). Washington, DC: American Psychiatric Press.

Volkmar, F. R., & Cohen, D. J. (1994). Autism: Current concepts. *Child and Adolescent Psychiatric Clinics of North America, 3,* 43–52.

Volkmar, F. R., Cohen, D. J., & Paul, R. (1986). An evaluation of DSM-III criteria for infantile autism. *Journal of the American Academy of Child and Adolescent Psychiatry, 25,* 190–197.

Volkmar, F. R., Klin, A., Marans, W., & Cohen, D. J. (1996). The pervasive developmental disorders. *Child and Adolescent Psychiatric Clinics of North America, 4,* 963–977.

Volkmar, F. R., Klin, A., Marans, W., & McDougle, C. J. (1996). Autistic disorder. In F. R. Volkmar (Ed.), *Psychoses and pervasive developmental disorders in childhood and adolescence* (pp. 129–190). Washington, DC: American Psychiatric Press.

Volkmar, F. R., Sparrow, S. S., Goudreau, D., Cicchetti, D. V., Paul, R., & Cohen, D. J. (1987). Social deficits in autism: An operational approach using the Vineland Adaptive Behavior Scales. *Journal of the American*

Academy of Child and Adolescent Psychiatry, 26, 156–161.

Warren, R. P., Yonk, J., Burger, R. W., Odell, D., & Warren, W. L. (1995). DR-positive T-cells in autism: Association with decreased plasma levels of the complement C4B protein. *Neuropsychobiology, 2,* 53–57.

Wenar, C., Ruttenberg, B. A., Kalish-Weiss, B., & Wolf, E. G. (1986). The development of normal and autistic children: A comparative study. *Journal of Autism and Developmental Disorders, 16,* 317–333.

Wing, L. K. (1976). *Early childhood autism.* Oxford, UK: Pergamon.

World Health Organization (1993). *The ICD-10 classification of mental and behavioural disorders, Diagnostic criteria for research.* Geneva, Switzerland: Author.

Zelazo, P. D., Burack, J. A., Benedetto, E., & Frye, D. (1996). Theory of mind and rule use in individuals with Down's syndrome: A test of the uniqueness and specificity claims. *Journal of Child Psychology, Psychiatry, and Allied Disciplines, 37,* 479–484.

Zilbovicius, M., Garreau, B., Samson, Y., Remy, P., Barthélémy, C., Syrota, A., & Lelord, G. (1995). Delayed maturation of the frontal cortex in childhood autism. *American Journal of Psychiatry, 152,* 248–252.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.03
Learning Disabilities

IVAN L. BEALE
University of Auckland, New Zealand

9.03.1 INTRODUCTION 38
9.03.2 PHENOMENOLOGY 38
  9.03.2.1 General 38
  9.03.2.2 A Typical Case of Learning Disabilities 38
  9.03.2.3 Definition 39
  9.03.2.4 Classification 40
    9.03.2.4.1 Learning disabilities subtypes 40
    9.03.2.4.2 Reading disability subtypes 40
    9.03.2.4.3 Mathematics disability subtypes 41
  9.03.2.5 Prognosis 41
  9.03.2.6 Other Problems Associated with Learning Disabilities 41
9.03.3 EPIDEMIOLOGY 42
  9.03.3.1 Prevalence 42
  9.03.3.2 Prevalence in Different Populations 42
9.03.4 ETIOLOGY 43
  9.03.4.1 Neurological Factors 43
  9.03.4.2 Psychological Factors 43
  9.03.4.3 Genetic Factors 43
  9.03.4.4 Environmental Factors 43
9.03.5 ASSESSMENT AND DIAGNOSIS 43
  9.03.5.1 The Behavioral Assessment Model 43
  9.03.5.2 The Issue of Diagnosis 44
  9.03.5.3 The Issue of Etiology 44
  9.03.5.4 Classification 45
  9.03.5.5 Goals of Assessment 45
    9.03.5.5.1 Identification of the focus of treatment 45
    9.03.5.5.2 Defining and measuring key behaviors 46
    9.03.5.5.3 Functional assessment (functional analysis) 46
    9.03.5.5.4 Selection of a treatment strategy 48
  9.03.5.6 The Assessment Procedure 49
9.03.6 TREATMENT 49
  9.03.6.1 General Treatment Issues 49
  9.03.6.2 Treatment Programs 49
  9.03.6.3 Treatment Implementation 50
  9.03.6.4 Monitoring of Treatment 51
9.03.7 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 51
9.03.8 SUMMARY 52
9.03.9 REFERENCES 53

9.03.1 INTRODUCTION

Learning disabilities (LDs) are learning problems that specifically affect one or more skill areas, such as reading, writing, or social skills, in persons who appear to be otherwise normally skilled and educated. The best-known, and most common, type of LD is dyslexia, which is characterized by a singular inability to read in an otherwise competent person. Although learning disabilities may be acquired by adults or children as a result of brain trauma, they originate typically as a developmental disorder of childhood, where they are as common as they are seemingly intractable to treatment.

There are many different psychological disorders that might afflict us at some time during our childhood. Some are known and feared by parents mainly because of the extensive media coverage they receive, featuring their dramatic nature, attendant family trauma, and poor prognosis. Examples include severe mental retardation, autism, fragile-X, and Tourette's and Rett's disorders. But many parents are less aware of the two psychological disorders of childhood that are by far the most likely to intrude into their lives. These are LD and Attention deficit hyperactivity disorder (ADHD). Each of these disorders affects about one child in every 20 and their combined prevalence is probably higher than that of all the other childhood psychological disorders (Pennington, 1991). Moreover, because LD and ADHD have a significant comorbidity and because each can lead to a secondary expression of some symptoms of the other, they are often confused with one another.

This chapter primarily is concerned with learning disabilities of developmental origin, as expressed in school-aged children. However, some reference is made to life-span implications, comorbid conditions, and the effects of brain trauma. As is appropriate in a volume on clinical psychology, emphasis is placed on the main concerns of the clinician: identification, assessment, and treatment. As far as possible, in the interests of providing a consistent framework, these are placed in the context of a broad conceptualisation of the behavioral assessment model (Nelson & Hayes, 1986).

9.03.2 PHENOMENOLOGY

9.03.2.1 General

The study of learning disabilities has a long history of controversy. Topics in dispute include not only definition, etiology, classification, and treatment but also fundamental questions such as whether learning disabilities even exist as identifiable disorders (Algozzine & Ysseldyke, 1983; Kavale & Forness, 1985; Radencich, 1984). For example, it is debated whether learning disabilities are reliably discriminable from cases of general learning backwardness or minor neurological damage, and it has long been argued whether medical, psychological, or educational expertise is most appropriate for dealing with the problem. At another level completely, there has been ongoing disagreement concerning the nature of some learning disabilities. For example, is dyslexia primarily a problem of visual processing, or does it result from a lack of phonological awareness (Hulme, 1988; Vellutino, 1979)?

The extent and persistence of the debate has led to a range of negative consequences. An example of political fallout is seen in New Zealand, where an otherwise developed country with a serious child education program does not recognize a need to provide assessment and remedial services explicitly for LDs (Chapman, 1992). Another undesirable consequence of the lack of professional unity on LD has been the persistence of a variety of speculative theories and their related interventions which have no basis in science or any other systematic knowledge. Examples include special diets, inappropriate medication, tinted glasses, gross motor exercises, and cranial manipulation (Beale, 1995; Silver, 1987). Perhaps the most troubling consequence of the LDs controversy is that its existence has been used by administrators, educators, and other professionals as an excuse for sidelining the LDs problem in decision-making about funding and other resource priorities.

Certainly the debate has importance within the framework of research and theory but, from a clinician's perspective, it is relevant mainly to the extent that it affects what is good practice in management of the clinical problems presented by the clients. Certainly, much of the controversy about etiology, definition, and classification can be regarded as only marginally relevant to the management of an individual case. This is because case presentation is primarily driven by concern about symptoms and the need for treatment rather than by the need for diagnosis, classification, and opinion regarding etiology.

9.03.2.2 A Typical Case of Learning Disabilities

This point of view can best be illustrated by considering briefly an example of a child with learning disabilities. The point here is to show that the major clinical considerations in the individual case are independent of the major issues under debate, including diagnosis, classification, and etiology. This case is revisited throughout this chapter to illustrate other issues.

A typical case presenting in a clinical setting is Hank, a nine-year-old boy referred by his parents, who cite persistent failure to learn to read or spell at anything like the normal rate. In other respects he has always seemed capable and is actually good at computational aspects of mathematics. He has some difficulty, however, working out the computational requirements of problems that are presented verbally. The problem was noticed in Hank's first year of elementary school instruction and has continued since, with Hank falling progressively further behind his peers in his level of performance in reading and writing. Concurrently, his initial positive attitude to schoolwork has reversed to the point of general distress and negativism about most school-related activities, even in those areas in which he can perform at an age-appropriate level. He is now anxious about school and has generally low expectations about his academic achievements. In particular, Hank hates any activities involving reading and writing. He considers himself to be "dumb," largely because of his difficulties with written language and the reaction to this, over the years, by some teachers and other students. He has suffered ridicule and teasing, his self-esteem generally has been eroded, and this is reflected in the deterioration of his social behavior. Hank now has no friends among his peers and few activities that he seems to enjoy. Of late, he has been inattentive and surly in the classroom, and his teacher has suggested to his parents that they seek expert advice.

Prior to this time, his teachers have not considered Hank's problem sufficiently severe to require expert intervention outside the remedial resources provided within the school. However, additional help provided in earlier years did not result in lasting gains in language skills. These remedial efforts consisted mainly of periods of one-to-one tutoring conducted by teacher aides who had no particular knowledge about learning disabilities nor insights into the nature of Hank's problem. The persistence of the reading problem in the face of remedial efforts has not been especially concerning to Hank's teachers, although it has been to Hank and his parents. His father recalls having reading problems at school and describes himself as still a "reluctant reader and hopeless speller." Hank's older brother had problems similar to Hank, although less severe. He seemed to get by after some remedial work was done with him. The impetus for the current expert assessment for Hank came not from concern about the reading problem but rather from the intrusive nature of the behavior problems that have emerged more recently.

The clinical issues are clear enough. There are specific academic-skills deficits overlaid by secondary behavioral and attentional problems. In addition, socio-emotional and motivational problems have begun to develop. The assessment and treatment of these problems can proceed regardless of the uncertainties or debate around issues of definition or etiology. However, an awareness of these matters may be helpful in discussing the problem with Hank and his family, and in securing the cooperation of teachers and other professionals who might be involved. Funding of treatment may also be dependent on diagnosis and classification of the problem, so there may be formal requirements to be met which require the clinician to grasp the nettle of LD definition.

9.03.2.3 Definition

The field of learning disabilities has a complex history. Kavale and Forness (1992) have traced the evolution of the definition of LD through the successive frameworks of brain injury (Strauss & Lehtinen, 1947), minimal brain dysfunction (Clements, 1966), and educational discrepancy (Kass & Myklebust, 1969) eventually to a series of consensual definitions required for legislative purposes (National Advisory Committee on Handicapped Children, 1968; National Joint Committee on Learning Disabilities, 1981). An example of a consensual definition of LD is that approved by the Interagency Committee on Learning Disabilities (ICLD) in 1988:

> Learning disabilities is a generic term that refers to a heterogeneous group of disorders manifested by significant difficulties in the acquisition and use of listening, speaking, reading, writing, reasoning or mathematical abilities, or of social skills. These disorders are intrinsic to the individual and presumed to be due to central nervous system dysfunction. Even though a learning disability may occur concomitantly with other handicapping conditions (e.g., sensory impairment, mental retardation, social and emotional disturbance), with socio-environmental influence (e.g., cultural differences, insufficient or inappropriate instruction, psychogenic factors), and especially with attention deficit disorder, all of which may cause learning problems, a learning disability is not the direct result of those conditions or influences. (Kavanagh & Truss, 1988, pp. 550–551)

Of course, such definitions are couched in broad descriptive terms that do not in themselves lead to a sure diagnosis. Concepts such as "significant difficulties" are useful only in as far as they can be translated into actual assessment protocols and clear interpretative criteria. However, past attempts to produce generally acceptable operational definitions for these descriptive terms usually have not been considered to be successful (Chalfant & King, 1976; Schere, Richardson, & Bailer, 1980). In the meantime, the ICLD definition serves as a useful guideline for identification of LD, to be operationalized where necessary by the requirements of whatever local authority sets the rules for definition of LD for the purpose of controlling access to treatment resources.

In the USA, requirements vary from state to state but are based on the federal requirement embodied in Public Law 101–476, the Individuals with Disabilities Education Act (1990). This includes a regulation stating that a student has a specific learning disability if:

(i) The student does not achieve at the proper age and ability levels in one or more of several specific areas when provided with appropriate learning experiences, and;

(ii) The student has a severe discrepancy between achievement and intellectual potential in one or more of these seven areas: oral expression, listening comprehension, written expression, basic reading skill, reading comprehension, mathematics calculation, and mathematics reasoning.

Again, these criteria can serve only as a general guide to identification of LD. In particular, attempts to develop discrepancy criteria that are valid across a wide range of ages and symptomatology have not been successful (Siegel & Metsala, 1992). Clinicians also find difficulty with the exclusionary clause in the ICLD definition because factors such as inadequate educational experiences, emotional disturbance, and neurological dysfunction, may well contribute to or modify the expression of a child's learning disabilities, even when they are not the root cause of the LD symptoms (Silver & Hagan, 1990).

The American Psychiatric Association's (1994) Diagnostic and Statistical Manual of Mental Disorders prefers the term "Learning Disorders." It identifies the categories of Reading Disorder, Mathematics Disorder, Disorder of Written Expression, and Learning Disorder Not Otherwise Specified. Subtypes are not specified within these disorders.

9.03.2.4 Classification

9.03.2.4.1 Learning disabilities subtypes

Early in its evolution as a clinical entity, LD was often regarded as a unitary concept, manifested in numerous research studies that contrasted "learning disabled vs. normal learners" as a means of revealing the essential factor underlying LD. Since the 1960s, this has given way to a general acceptance that LD is a heterogeneous disorder both at the level of symptomatology and etiology. The problem has been to find classificatory systems that are both enlightening theoretically and helpful in the process of choosing an effective intervention. Useful approaches might be based on areas of academic skill deficits, some aspect of the presumed etiology (Silver & Hagin, 1990), patterns of academic performance (Siegel & Metsala; 1992), or statistically identified clustering of performance variables (Rourke, 1985).

The skill areas of difficulty usually distinguished in the LD literature are: reading disability (oral reading and comprehension); oral language disability (listening and speaking); written language disability (spelling, handwriting, and written composition); mathematics disability (numerical concepts, computation); social disability (social skills); and motor disability (fine and gross movement, balance and kinaesthesis, orientation). Classification based on underlying processes generally makes reference to the following: memory (several categories); attention (several categories); phonological awareness; and cognitive and metacognitive strategies.

There has been an extensive research effort to identify useful classification systems for LD problems, driven by the need to deal rationally with the apparent heterogeneity of the disorder. The empirical approach to subtyping is based on statistical procedures, such as Q-factor analysis, and cluster analysis of scores on performance batteries mainly comprised by academic attainment tests and neuropsychological tests. Studies of this type have identified three (Petrauskas & Rourke, 1979) to five subtypes (Satz & Morris, 1981). However, the limitations of this approach have been pointed out repeatedly (Kavale, 1990a; Siegel & Metsala, 1992), and there has yet to emerge a consensual classification scheme that could simplify rather than complicate the conceptualization of LD heterogeneity.

9.03.2.4.2 Reading disability subtypes

Better understood are the predominant types within particular areas of academic skills deficit. The more specific category of reading disability, considered in less-enlightened times to be a unitary syndrome (Gaddes, 1985), has been classified into subgroups by several researchers (e.g., Doehring, Hoshko, & Bryans, 1979; Mattis, 1978). There is strong evidence that by far the most common type of reading disability is characterized by an impairment in language skills, especially the phonological processing required to discriminate the sequence of sounds in spoken words. The spelling to dictation of children and adults with LD is difficult to decipher, because the sequence of letters often does not correspond to the sequence of sounds in the words; they do not write the word as it sounds.

The least common is the visuo-spatial type of reading disability. Children in this category may have good phonological skills but poor visual memory. This is manifested in poor sight vocabulary in "flash" tests. Their spelling of irregular words is often phonetically regular but visually incorrect.

A third identified subtype is of intermediate prevalence relative to the other two types. The spelling problems in this group seem to be a mixture of the two other types. It has been suggested that this type of problem is characterized by difficulty with processing of sequential information (Mattis, 1978; Rourke, 1982).

It should not be thought that these three groups form a complete typology that is well accepted. Some investigators (e.g., Bakker, 1992) have preferred two categories, based essentially on theories about inappropriate balance of control between the two hemispheres of the brain. In fact, many cases of reading disability do not readily fit these classification systems (e.g., Harness, Epstein, & Gordon, 1984), indicating that the degree of heterogeneity among reading disabled children may be incompatible with any simple system of typology.

9.03.2.4.3 Mathematics disability subtypes

Mathematics disability has also attracted classificatory research. For example, Rourke and Strang (1983) and Rourke, Young, and Leenaars (1989) found useful subtypes based on patterns of children's performance on the reading, spelling, and arithmetic subtests of the Wide range achievement test (WRAT). Among those with mathematics problems, one type seemed primarily verbal (left hemisphere) and the other primarily visual-spatial (right hemisphere). This finding is partly supported by other studies (Share, Moffit, & Silva, 1988).

9.03.2.5 Prognosis

Longitudinal studies have addressed the issue of which is the most appropriate model: deficit, lag, delay, or some amalgam of these. In a review of these studies, Reiff and Gerber (1992) concluded that although LDs may be best characterized by the deficit concept, favorable outcomes are nevertheless possible, depending on a number of factors such as effective interventions and favorable social ecologies. LDs are viewed as a persisting condition which is expressed differently at different life stages. In adulthood, the former focus on academic skills development is replaced by social and vocational issues. The problems of adults with LDs are unlikely to be confined to the stigma and inconvenience of illiteracy, although it is illiteracy that usually is the focus of whatever services are readily available for adults.

Factors predicting poorer prognosis include greater severity of disability, lower intelligence (Minskoff, Hawks, Steidle, & Hoffman, 1989), failing to access intervention services (Bruck, 1987), and lack of family support (Reiff & Gerber, 1992). Notably, severity alone should not be equated with poor prognosis, given the many individuals with severe childhood disabilities who have led very successful adult lives. Much has been made of a possible link between learning disabilities and sociopathy in adolescence and adulthood, the so-called LD/JD (juvenile delinquency) connection. Several studies have shown that the prevalence of learning disabilities is unusually high among adjudicated delinquents (Brier, 1989; Broder, Dunivant, Smith, & Sutton, 1981) and that educational remediation can reduce recidivism significantly in LD delinquents (Brier, 1994). However, it is unclear what other factors may need to be present in a child with learning disabilities to precipitate the development of delinquency. A number of predisposing factors has been suggested by Reiff and Gerber (1992), including early conduct problems, association with a deviant peer group, and concurrent diagnosis of ADHD. The type of disorder may be an important predictor of adult outcomes. There is some evidence that a type of disorder with a nonverbal character may be associated with greater risk of depression and suicide (Rourke, Young, & Leenars, 1989).

9.03.2.6 Other Problems Associated with Learning Disabilities

As already noted, LD and ADHD have substantial co-occurrence or comorbidity, although estimates of the extent of this (3–80%) vary considerably with alternative definitional criteria, especially for ADHD (Anderson, Williams, McGee, & Silva, 1987). Because LD often leads to secondary ADHD symptoms and vice versa, there is a need for comorbidity estimates to be based more than usual on careful diagnosis. Studies indicate that similar heritable influences are in part responsible for both disorders (Narhi & Ahonen, 1995), and that comorbidity results from a different developmental pathway than that leading to LD or ADHD alone (Pennington, Grossier, & Welsh, 1993; Sanson, Prior, & Smart, 1996).

Although clinical samples have generally indicated that left-handedness is associated with LDs, this is not found in epidemiological studies, suggesting that children with learning problems are simply more likely to be referred if they are left handed (Corballis & Beale, 1983). According to teacher ratings, problems that are more common in children than in those without LDs include: social maladjustment (Cullinan, Epstein, & Lloyd, 1981); self-consciousness, inferiority, withdrawal, tension, and anxiety (Stone & LaGreca, 1984). Parent ratings indicate lower levels of participation, social interaction, and social competence (McConaughy & Ritter, 1986). A meta-analysis of social skills deficits in LDs found that about 75% of students with LD problems had significantly marked social skills deficits (Kavale & Forness, 1996). Moreover, the data did not support the idea that social skills problems themselves represent a specific learning disability, but rather that they are simply associated with LD.

Preschool or early-school characteristics that are predictive of LD diagnosis in later years include poorer ability to provide letter names and sounds or to perform rhyming tasks (Bryant & Bradley, 1985). In addition to such phonological skills, poor word attack and word identification skills have also been shown to predict later diagnosis of reading disability (Hurford, Schauf, Bunce, Blaich, & Moore, 1994). There are important implications for early intervention in these findings, at least to the extent that such skills are trainable in those children found to be deficient in them at the outset (Torgesen, Wagner, & Rashotte, 1994).

9.03.3 EPIDEMIOLOGY

9.03.3.1 Prevalence

The uncertainties about definition have led to doubts about the prevalence of LDs because different studies have tended to use different criteria for identifying the presence of LD (Epps, Ysseldyke, & Algozzine, 1985). Moreover, methodologies used in surveys of prevalence are not uniform, including surveys of school policy (Silverman & Metz, 1973) and studies of normal population cohorts (Rutter, Graham, & Yule, 1970). Not surprisingly, estimates of prevalence vary from 2 to 20%. A detailed analysis of prevalence studies is provided by Silver and Hagin (1990) and they reported that school administrators' estimates of the prevalence of LDs in US elementary schools ranged from 5 to 30%, the majority of estimates being between 10 and 15%.

The US Department of Education (1991) reported that 3.6% of children between six and 21 years were classified as LD in 1990, compared to 1.2% in 1977. Presumably the increase in children classified as LD was a direct result of the first introduction in 1977 of legislation requiring identification and adequate resourcing of LD cases [Public Law 94–142: The Education for All Handicapped Children Act (1975)]. What this trend makes clear is the impossibility of using official figures on children served under the provisions of the Act as a basis for estimating the prevalence of LD. Kavale and Forness (1992) have observed that over a similar period (1977–1986), the number of children classified as mentally retarded for the purpose of resourcing under the Act apparently decreased by about 29%, suggesting that the change in LD numbers may have had more to do with changes in classification procedures than anything else.

If only reading disability is considered, estimates again vary between studies. For example, epidemiological studies have reported ratios of about 2.5:1 (Berger, Yule, & Rutter, 1975) and about 1:1 (Shaywitz, Shaywitz, Fletcher, & Escobar, 1990). Studies of the familial incidence of reading disability show ratios of about 2:1 (deFries, 1989).

9.03.3.2 Prevalence in Different Populations

Given the problems of definition discussed above, there seems to be no valid basis for comparing real prevalence rates in different countries. There is little doubt that LD occurs in many countries, cultures, and languages (Junkala & Paul, 1987) although it may well be expressed in different ways, especially where the form of the language is very different from English. For example, Rozin, Poritsky, and Sotsky (1971) reported that US children with reading disabilities in a phonetic-based language (English) were able to learn vocabulary in an idiographic language (Chinese) without difficulty. Another report (Motluk, 1996) described a boy raised by English-speaking parents in Japan who had reading disability in English but not in either Kanji or Kana, the written languages used in Japan.

It is generally recognized that the prevalence of LD is higher among males than females, with estimates of the factor varying between 2:1

(Nichols & Chen, 1981) and 2.5:1 (US General Accounting Office, 1981). Vogel (1990) has reported evidence that the symptoms expressed by boys and girls tend to be different, with boys having more problems of a visuo-spatial character.

9.03.4 ETIOLOGY

9.03.4.1 Neurological Factors

In keeping with the origins of the LD concept in the framework of brain damage, theories of etiology primarily have had a neurological or neuropsychological character. These have dealt mainly with the predominant type of LD, reading disability, with relatively little attention being paid to other types of disability such as mathematics and social skills. Theories have been proposed at the level of basic neurological organization, basic psychological processes, and genetics. An early example of the first type, that has persisted in modified form even to the present day, is that proposed by the neurologist Orton (1937) who thought that reading disability was a phenomenon of visual memory and arose in individuals whose brains failed to develop the usual left-hemispheric dominance for language. The illogicality of some aspects of Orton's theory have been pointed out by Corballis and Beale (1983, 1993) who have proposed a modified version more in keeping with current knowledge of brain function, while retaining Orton's idea that lack of hemispheric asymmetry may be a critical factor. Other approaches to the understanding of the origins of reading disability at the level of brain anatomy or neurological organization have been presented by Geshwind and Galaburda (1987) and Hynd and Hynd (1984). The idea that reading disability has its origin in some aspect of lateralization of brain function also persists in theories that propose different rates of development of lateralization of function in reading disability (Satz & Sparrow, 1970; Witelson, 1977).

9.03.4.2 Psychological Factors

Turning now to theories about anomalous psychological processes in reading disability, examples have featured different aspects of visual processing such as visual memory (Willows, Kruk, & Corcos, 1993), visual perception (Lovegrove, 1992), eye movements (Pirozzolo, 1979), and different aspects of auditory processing, such as auditory sequential perception (Tallal, 1988) or linguistic skills generally (Shankweiler, Liberman, Mark, Fowler, & Fisher, 1979; Vellutino, 1979).

Many studies have pointed to ineffective use of cognitive and metacognitive strategies by children with learning disabilities (Brown & Palincsar, 1982; Wong, 1979). While these clearly may play an important role in the maintenance of reading problems in older children, it is not clear that they are a primary cause of reading disability.

In their review of neuropsychological theories of learning disabilities, Willis, Hooper, and Stone (1992) concluded that single-factor theories have not been successful in explaining the diversity of symptomatology in reading disability. The broader perspective of neuropsychological theories such as that of Rourke (1985) may have more promise as a guide to both research and clinical practice.

9.03.4.3 Genetic Factors

The apparent tendency of reading disability to run in families has prompted research into possible genetic factors that might underlie the nervous-system anomalies associated with the problem. Pennington (1991) has noted that the evidence shows both familiarity (35–40% of first-degree relatives are affected) and heritability (50%). Reading disability is apparently heterogenous genetically, although this does not mean that the physiological or neurophysiological expression will also be heterogeneous (Pennington, 1991).

9.03.4.4 Environmental Factors

It should be understood that environmental factors also have a role in the etiology of learning disabilities, although little is known about them. Badian (1984), for example, showed that reading problems are associated with large family size and low socioeconomic status. Of course, these factors also increase the risk of exposure to nervous-system toxins and poor nutrition that might affect brain development and learning directly or alternatively might interact in a negative way with a preexisting neurodevelopmental factor. In either case, the net effect still amounts to an "intrinsic factor" within the standard definition of a learning disability.

9.03.5 ASSESSMENT AND DIAGNOSIS

9.03.5.1 The Behavioral Assessment Model

Assessment of LDs takes many forms, corresponding to the many purposes to which it may be put (Salvia & Ysseldyke, 1991). This chapter uses the framework of a behavioral assessment model on the grounds that it is the most appropriate model for assessment of a problem that includes the intrinsic factors within the individual client, family, school, and other social factors, educational and other learning factors, and treatment considerations. Others making recommendations for the assessment of LDs have also proposed assessment approaches based on a behavioral assessment model. For example, Barkley (1981) advocated what he called a behavioral-neuropsychological assessment as being an approach that truly considers all of the aspects that might be relevant to the understanding and management of an LD problem.

The behavioral assessment approach has a clear advantage over other, more traditional approaches to assessment in that it considers the referral problem within a broad environmental context in which that problem occurs and in which the treatment will also be implemented. As such, it does not attempt to isolate the problem either within the person central to it or within particular aspects or subsets of the person's environment (Nelson & Hayes, 1986). In particular, choosing and monitoring the treatment are considered a part of the assessment process, ensuring an assessment that is truly driven by treatment considerations. Thus, it is recognized that LD problems have multifactorial causes and require multifactorial solutions. This is a significant departure from a type of approach that has been more common in the LD field, that is, choosing a treatment on the basis of either theoretical preconceptions of the etiology of the LD problem (e.g., Delacato, 1966) or on the basis that only the deficient academic behaviors need to be considered (e.g., Treiber & Lahey, 1983).

The assessment process essentially is one of gathering information to assist in reaching an understanding of the problem in order to design and test a suitable treatment. Texts on assessment of learning problems will often systematize formal aspects of the process of information gathering by organizing it on dimensions such as temporal (current vs. historical information) and methodological (direct observation, performance tests, rating scales, and checklists) (Salvia & Ysseldyke, 1991). There is heuristic value in this but, to be cost-effective, the information-gathering process needs to be driven by hypotheses about what sort of information is likely to be needed. Such hypotheses are derived from the general body of theory and knowledge about etiology, phenomenology, and treatment efficacy, as well as from the assessor's clinical experience or personal "case-lore."

Before proceeding to outline the formal structure of assessment, a return is made to Hank, the hypothetical boy with LD problems, to consider how an assessment might proceed in the less-than-perfect light of the referral information and the information on the nature of learning disabilities already garnered from the literature.

9.03.5.2 The Issue of Diagnosis

The funding of resources that may be needed to help Hank, including the cost of an assessment, are likely to be contingent on an LD diagnosis that is within the parameters specified by the local education administration. The information required is:

(i) Are his symptoms consistent with the inclusionary criteria for skills deficits? Spelling and reading are described as below his age level. Tests will need to be given to show his current level of these skills. There may be a requirement to use approved tests. Either norm- or criterion-referenced tests would serve.

(ii) Can an argument be made that central nervous system dysfunction is a contributing factor? No problem was apparent before he entered school, so evidence must be sought of some deficit in basic tests of psychological function. Such a deficit would not be definitive, but at least would be consistent with a hypothesis of central nervous system dysfunction. Norm-referenced tests would be required. The fact that his father and brother may have had similar problems indicates a familial and heritable disorder. This needs to be explored further by questionnaire or interview.

(iii) The exclusionary factors will need to be addressed. On current evidence, inappropriate instruction is an improbable cause, because remedial efforts have been unsuccessful. Judging from referral information, mental retardation and sensory impairment are probably not present, but intellectual status can be obtained from a total IQ score, and evidence of subtle sensory impairment can be looked for during formal testing. The attentional and behavioral problems, as well as current emotional and social problems, are apparently secondary to LD, rather than causal, but further information is required from school and home.

9.03.5.3 The Issue of Etiology

Although not necessarily helpful to the process of designing an effective treatment, etiology needs to be considered. Not only is it relevant to diagnosis but it can assist the family in coming to terms with the problem. Consideration of etiology also guides the assessment process in a way that may throw light on possible causal factors not indicated in the referral information. It cannot be assumed, on the basis of information about the possible occurrence of LD in other family members, that Hank's problem is totally developmental in origin. A medical history is required to check for the possibility of acquired neurological factors, such as head injury or meningitis, which may have contributed to a greater or lesser degree to the later emergence of the learning disability. Referral for neurological examination may need to be considered.

9.03.5.4 Classification

Classification might be seen as helpful for several reasons: it may assist the choice of an effective treatment; it may indicate one or another prognosis; it may increase the usefulness of the case for research purposes; it may clarify the essential nature of the problem; or, it may assist interdisciplinary assessment. The extent to which classification serves these purposes depends very much on the type of classification that is envisaged. The simplest type of classification, based just on the type of skill deficit, serves at least the last two of these aims. Hank's problem would probably be classed as a case of reading and spelling disability, with an overlay of secondary attentional, social, and behavioral problems. This would identify the case sufficiently for some research purposes, at least, and would provide a clear handle on the case for the purpose of a fuller assessment. This initial classification would be subject to modification on the basis of further information, of course. In any case, it would need to be confirmed by formal testing or reading and spelling skills.

Other types of classification, based on etiology or psychological processes, may contribute little in the immediate practical sense. They would assist prognosis and treatment choice only to the extent that the literature shows clear relationships between these variables and subtypes. But reviews of the research on subtypes are not hopeful in this respect (Forness, 1990; Siegel & Metsala, 1992). If there existed consensual guidelines on subtyping (e.g., Kavale, 1990a), these could at least be applied to a case to make it more accessible to retrospective research. Until that time, classification is likely to be driven by the theoretical framework in which the assessment is conducted, which may have more to do with the history of the assessor than with the nature of the problem being assessed. Nevertheless, the information gathered in the assessment process, particularly the formal psychometric tests that are used, will be determined jointly by the nature of the problem and the assessor's theoretical stance. Beyond the simple requirements of defining the nature of the problem, then, there can be no overall recommendation regarding classification. However, the behavioral assessment model provides clear guidance that is based on other considerations.

In proceeding with the assessment, there are two dimensions to consider that are simultaneously orthogonal and overlapping. The first is the goals of the assessment, including identification of the problem behaviors, identification of measurement methods, functional assessment (functional analysis) of the problem, and selection and monitoring of treatment (Nelson & Hayes; 1986). The second is the assessment procedure, including all aspects of data collection (interviewing, rating scales, tests, observation), data interpretation, and planning, implementation, and monitoring of treatment. The first dimension can be considered a guide to the second.

9.03.5.5 Goals of Assessment

9.03.5.5.1 Identification of the focus of treatment

This phase requires a consideration of which behaviors need to be increased, and which reduced. The context for this includes both social comparison and subjective evaluation. First, the fact that an LD problem is the referral issue does not mean that this is the best context for dealing with the problem. Parent referrals, in particular, may be mistaken in supposing that there is a significant academic skills problem, their judgement influenced perhaps by the greater success of older children in the family. Information from the school is essential for checking skills status at the earliest opportunity. Behavior problems identified by parents also need to be considered in a wider social context than just their judgement about the child's behavior at home. In the case of both academic skills and other behavior, appropriate norms should be consulted in making a decision about the generality and severity of the problem. With regard to socio-emotional issues in LD, subjective evaluation is also important. What are the attitudes, feelings, and beliefs held by the client and others about the referral problem? The point here is that there are subjective as well as objective contexts that are important; there is no single "correct" way of viewing the problem.

In Hank's case, the identification of target behaviors probably would be based primarily on teacher opinions, performance records such as reading and spelling tests given in school, and parent and teacher rating scale evaluations of social and other problem behaviors. Initially, it would appear that reading and spelling skills need to be strengthened and inappropriate attitudes, attention, and social behaviors need to be reduced. Anxiety may be treated as a collateral issue rather than being targeted directly, on the assumption that it is secondary to other problems, but it may need to be directly addressed if it turns out that it impedes either the assessment or compliance with treatment.

9.03.5.5.2 Defining and measuring key behaviors

When academic behaviors are targeted, this phase is achieved by the selection of formal tests, and care must be taken that the tests provide valid and reliable measurement of the behaviors at issue. The aim is to provide definitions that are objective, clear, and complete.

What precisely are Hank's reading and spelling problems? This will not necessarily be obvious from the data available from school. Reading includes whole-text oral reading skills such as pronunciation, punctuation, and fluency. It also includes comprehension or material dictated or self-read. It includes single-word recognition skills, with or without the opportunity for word analysis. Spelling includes oral naming of letters in words, written spelling to dictation, and proof-reading. Words spelled may be phonetic or nonphonetic. Both for spelling and reading, specific tests will be required to delineate the areas of difficulty. The primary criteria to be considered in choosing these tests are validity and reliability. Assessment of the degree of deficit can be accomplished using either norm- or criterion-referenced tests, and a good case can be made for using both (Salvia & Ysseldyke, 1991).

In Hank's case, attentional and behavioral problems are an issue. Regardless of whether it is intended to treat these directly or relegate them to the list of collateral behaviors, decisions are required about definition and measurement at the outset. This will be based, in the first instance, on rating scale and questionnaire information. Consideration might be given, however, to the possibility of a primary attentional problem, in which case tests will be required that differentiate between aspects such as sustained attention, focused attention, distractibility, and attention span (Lam & Beale, 1991).

In Hank's case, anxiety and social skills scales might be considered both for the purposes of gauging the severity of such problems and for measuring changes accompanying treatment, although teacher and parent ratings would also be required to check their perceptions of functioning in these domains.

Often it is advocated that systematic direct observation be included in the assessment process (Barkley, 1981), especially when referral issues emphasize behaviors that are public and easily observable. An advantage is that direct observation can be carried out by an "independent" person without a stake in Hank's problems or their management. Thus, direct observation may be seen as less subject to the influences that are regarded as possibly biasing informal information-gathering or the global ratings of interested persons such as teachers or family members. On the other hand, direct observation is only as good as the observation system is valid and reliable in the hands of the particular observers using it. This can be difficult to achieve, because of factors such as reactivity of observation, observer drift, and nonrepresentative sampling (Mudford, Beale, & Singh, 1990), not to speak of the resources required to train observers and to carry out the observations. In the clinical context, as opposed to the research context, use of systematic direct observation protocols is unlikely to be considered cost-effective. However, consideration would be given to using it in instances where more convenient forms of measurement are likely to be unreliable. In Hank's case, social activity in the school playground might be difficult to measure reliably by any other method.

Self-monitoring has been used extensively for the recording of semipublic or private behaviors, as well as providing a basis for the treatment of such behaviors (see Reid, 1996). For Hank, this method would be used to measure the proportion of time spent on-task in classroom instruction, both for the purpose of quantifying his initial level of distractibility and for treating this directly by reinforcing increases in time spent on-task. Variations of this method have been used for the measurement of accuracy (Marshall, Lloyd, & Hallahan, 1993) and productivity (Lloyd, Bateman, Landrum, & Hallahan, 1989) in classroom performance.

9.03.5.5.3 Functional assessment (functional analysis)

This phase of the assessment process is at the core of interpretation of the problem, leading to the choice of a suitable treatment. It includes the assessment of: (i) variability in the target behaviors; (ii) the settings or other conditions in which the problems are better or worse; (iii) environmental consequences of the problem behaviors; and (iv) intrinsic or organismic factors that may be relevant.

The first two of these, variability and settings, may be closely related. According to the behavioral assessment model, variability in a behavior is indicative of the joint control exercised over that behavior by setting events (discriminative stimuli) and consequences for that behavior in particular settings. LD problems may vary according to several variables, including who is the teacher, classroom dynamics, seating arrangements, time of day, day of the week, the general nature of the current classroom task, and especially, the specific subtask in which the child is engaged. A similar analysis is possible of behavior at home.

Hank's reading, spelling, and attentional problems can be measured within this framework. A common strategy is to supply teachers with a checklist that requests global ratings on all academic skill areas as well as on general behavior and motivation. For each skill area, the teacher is also asked to respond to questions about subskills and about how they vary in the context of different conditions: before vs. after breaks; in group vs. alone; morning vs. afternoon, and so on. Analysis of such information may show that distractibility and wandering is worse on tasks of a particular type, or during subtasks in which accuracy is low, or just before a break. Afternoon performance may, in general, be worse than morning.

Such observations not only indicate a causal relationship between academic skills problems and behavior and attentional problems, but they provide information about how the problems can be eased simply by environmental management. Examples are seating changes and rescheduling of academic lessons within the day or within the week. Some children have memory and organizational problems that can be circumvented greatly by changes of this sort.

Variations in consequences for behaviors are of course related both to the probability of those behaviors and the occurrence of competing and collateral behaviors, including some that are private. Criticism or ridicule of a child's errors will affect not only the behavior specifically criticized but also the child's willingness to participate generally, and self-esteem. Conditioned anxiety may develop that impedes further learning. Different teachers have different styles in this regard, and it is important that the assessor discovers what they are and how typically they are deployed in relation to the areas of difficulty. Hank's anxiety may have originated in the reactions of previous teachers and peers, but it is nevertheless maintained by prevailing conditions, and these need to be understood so that they can be addressed as required by the treatment program.

The next phase of the functional assessment involves consideration of intrinsic factors that might be relevant to the LD problem, including possible sensory impairment, neurological or neuropsychological problems, other physical and psychological conditions such as depression or anxiety, and psychopharmacological factors. Past history of treatment may also be included here, because there presumably is some legacy of this in the client's present constitution. Notably, this phase is not included in the assessment model often advocated by those trained in the educational discipline, who prefer to view learning disabilities in the context of the simpler ABC (Antecedents, Behavior, Consequences) model (e.g., Treiber & Lahey, 1983).

Sensory impairment can itself cause learning problems, or it can make worse problems that have a different cause. Visual and auditory perceptual problems clearly can have profound effects on the learning of academic skills. In most children referred to LD specialists, routine testing for hearing loss or for visual refractory errors will already have been conducted, but when this is not the case, such problems usually are apparent during psychological testing and a referral can be made for an expert assessment. More subtle problems, however, may have gone undetected. Rare individuals may have visual tracking problems that prevent fluent reading and are only identifiable on specialist optometric examination. Untreated chronic middle-ear infection is common in some countries, where it is a common cause of retardation of the development of language skills in preschool children. Visual perceptual factors such as scotopic sensitivity syndrome, and auditory problems such as central auditory processing deficit, have been regarded as causes of learning problems, and it is arguable whether they should be seen as simple sensory impairment for the purposes of exclusion under the standard definition of LD.

The assessment of neuropsychological factors is central to many theoretical models of LD, even though the role of these factors in treatment selection remains contentious (Kavale, 1990b). The use of neuropsychological tests in the assessment of LDs is usually primarily driven by the assessor's belief that such tests will assist in characterization of the basic processing problem(s) underlying the LDs. Because there are several classificatory systems that are based on different conceptualizations of neuropsychological processes relevant to LDs, assessment practices vary considerably. LD assessments in the context of research programs commonly use the same battery of tests in all cases (e.g., Rourke, 1985), while clinically-oriented assessment services are more likely to select tests on the basis of hypotheses being tested about the individual cases (Sattler, 1992). A compromise approach that serves both purposes is to use a limited set of tests for all clients, supplemented by additional tests chosen to fit the individual case.

An illustrative example is provided by the practice followed at the University of Auckland's Learning Assessment Centre. The core testing program includes the WISC-III, norm-referenced tests of academic skills attainment (reading, spelling, and math) and tests encompassing basic neuropsychological domains of attention, memory and information processing (see Table 1). Additional tests are used to test hypotheses about weaknesses indicated by the initial tests, and these vary from case to case.

Pharmacological factors are increasingly a concern in LD assessment (Aman & Rojahn, 1992). Many children referred for LD problems are currently on medication prescribed either for this or some related behavioral problem. This is especially the case for those with ADHD symptoms or diagnosis. In the USA, over 1% of schoolchildren are on stimulant medication, and the percentage has been as high as 10% in some states. Any psychoactive medication can be expected to influence the assessment process, including the judgement of severity and the choice of treatment. The issues are encapsulated in these questions: (i) Is the medication acceptable to the client and family? (ii) Is it perceived to be effective in controlling behavioral or attentional problems? (iii) Will the medication mask aspects of the learning disabilities during assessment, potentially invalidating some assessment conclusions? (iv) Is it proposed to continue the medication concurrently with another recommended treatment? and (v) Have adverse effects of the medication been reported?

The functional assessment concludes with an integrated synthesis of all the information collected to provide a coherent interpretation of the LD problem. This includes consideration of all probable contributing causal factors, intrinsic and environmental, an account of how these have been expressed in the development of the problem, and a description of the current problem that includes a discussion of the interactions between the LD problems and the current school and home environments.

9.03.5.5.4 Selection of a treatment strategy

Selection, implementation, and monitoring of treatment normally are included as part of the behavioral assessment model but, in this chapter, they are considered under a separate heading.

Table 1 Instruments (tests and scales) suitable for the assessment of LD.

Test	*Scales*
General intelligence	Weschler intelligence scale for children-III (WISC-III)
	Weschler preschool and primary scale of intelligence (WPPSI)
	Peabody picture vocabulary test-revised (PPVT-R)
Abilities	McCarthy scales of children's abilities
	Kaufman assessment battery for children (K-ABC)
	Peabody individual achievement test (PIAT)
	Detroit test of learning aptitude (DTLA-2)
	Lindamood auditory conceptualization test (LACT)
Academic achievement	Wide range achievement test-revised (WRAT-R)
	Progressive achievement test (PAT)
	Gray oral reading test-revised (GORT-R)
	The border test of reading-spelling patterns
	Neale analysis of reading ability-revised
	Gates-MacGinitie reading tests
	KeyMath diagnostic mathematics test-revised
	Woodcock-Johnson tests of achievement
	Proof-reading tests of spelling
Behavior rating scales	Conners teacher rating scale (CTRS)
	Swanson, Nolan, and Pelham rating scale (SNAP)
	Revised behavior problem checklist (RPBC)
Neuropsychological tests	Visual retention test
	Learning efficiency test
	Wisconsin card sorting test
	Trail making test for children
	Symbol-digit modalities test
	Selective reminding task
	Rey–Osterreith complex figure

9.03.5.6 The Assessment Procedure

The second dimension of assessment is the actual physical process by which the assessment proceeds. Typically, referral is followed by a contract for services being sent to the client (in the case on a child with LD, the client is a parent or guardian). Usually this stipulates the nature and cost of the service being offered, and requests written approval for other agencies (e.g., schools) to be contacted for information. It may also include a request to use anonymous assessment data for research purposes.

Receipt of approval may be followed by a mailing of questionnaires to parents and schools. Parents may be asked to complete medical and family history records, rating scales for their child's behavior problems, treatment acceptability scales, and so on. Teachers are asked to complete behavior rating scales and checklists relating to academic behaviors. When all the information has been returned, an appointment is made for the initial interview with the child and parents.

The purpose of the initial interview is threefold: to impart information about the assessment process; to review the information already provided in questionnaires; and to talk to the parents and child about the LD problem. It is advisable that the interview be structured, at least on paper, so that everything gets covered. On the other hand, it is important to allow the clients to talk freely and offer information without excessive prompting. It is usually advisable to speak to the parents both with and without the child being present, as this provides valuable information about their interactions with the child around the LD problem. One to two hours is usually sufficient for the interview.

Testing sessions should be conducted without a parent being present, but not if there is reason to expect a child to perform significantly worse with no parent present. Anxiety conditions, especially parent separation anxiety, can be exacerbated by the novel situation, and some children are untestable even with a parent close at hand in the waiting room. Recommended practice should be followed with regard to general factors such as rapport, sequence of tests, duration of test sessions, and so on (Sattler, 1992).

School records, including reports and results of school administered tests, can be a valuable source of information, as can children's exercise books. This type of information gives a much wider sample of a child's current skills and their developmental history than can be obtained in the context of brief testing sessions. A useful practice is to ask parents to gather this material and bring it to the initial interview.

The assessment report is unquestionably an important part of the assessment process, and plays a critical role in whether an effective treatment will be implemented. Reports are most useful when they follow the recommendations given in good assessment texts (Kamphaus, 1993; Sattler, 1992). The sovereign principle is that they should be written for the consumer, usually parents, so as to maximize understanding of the nature of the problem, prognosis, and preferred course of remedial action. A report that is overly expansive on details of testing and interpretation may be set aside as patronizing, technical, jargonistic, and indecisive. In any case, the client should be debriefed on the report and its implications, at least by telephone, but preferably in a recorded follow-up interview where all interested parties are invited to attend. The assessor should be prepared to address any questions and to facilitate plans for implementation of the recommended treatment.

9.03.6 TREATMENT

9.03.6.1 General Treatment Issues

Treatment issues can be divided into selection, implementation, and monitoring. Treatment selection follows logically from the functional assessment, which should provide an understanding of strengths and weaknesses in the client's repertoire of academic skills, social skills, and relevant psychological processes. It should also indicate the context, especially advantages and constraints, for possible alternative remedial approaches. A preferred initial treatment approach can then be chosen on the basis of the professional literature on treatment efficacy, together with local issues such as resources, treatment credibility, and acceptability. Reviews of alternative treatment strategies and specific treatment programs are provided elsewhere (Beale, 1995; Frostig & Horn, 1964; Hartlage & Telzrow, 1983; Kavale 1990b). It is emphasized here that treatment choice should not be based on theoretical preconceptions of the nature of LD, current fads, or anything other than functional assessment of the individual case.

9.03.6.2 Treatment Programs

Treatment programs for LDs fall into two categories. One, sometimes called training the brain or psychological process training (Gaddes, 1985; Hammill & Larsen, 1978; Kavale, 1981, 1990b; Torgesen, 1979), directs treatment at processes, the supposed weakness of which are the underlying cause of the LD problem. It has been argued that processes cannot be trained directly, only skills that use those processes (Beale & Tippett, 1992), and that the validity of this approach is heavily dependent on unproven assumptions about generality of behavior change across skills. The literature on the effectiveness of process training for LD is mostly unfavorable (Hoehn & Baumeister, 1994). At the same time, it can be argued that evaluations of process-oriented treatments have relied primarily on between-groups research designs which are unsuitable for evaluating the effectiveness of treatments on individuals (Hersen & Barlow, 1976). Studies using single-case designs are needed to provide a basis for deciding whether there might be some cases for which this approach can be useful (Beale, 1995).

The other category is referred to as direct remediation and consists of directly training the absent or poorly mastered academic or social skills, using established behavioral technology. There are abundant studies using good research designs which show that direct remediation programs can be applied effectively to a range of academic and social skills deficits in LD persons. Examples of useful programs and their areas of application are shown in Table 2.

The choice of an appropriate program is determined by the areas of skill deficit. Where there is more than one area of deficit, as in the illustrative case of Hank, a decision is needed about prioritizing areas for remediation. Sometimes it is possible to proceed with an integrated program for the concurrent remediation of different skills. For example, Hank might receive a phonological skills training program that addresses his reading and spelling skills in the context of a self-monitoring program aimed at increasing his attention to task. Clinical experience indicates that such concurrent programing places a great burden on those responsible for its implementation, and it should probably be restricted to those cases where trained professionals are involved in the actual day-to-day administration of the program. Usually, programs are delivered sequentially and are prioritized on the basis that the more fundamental skills should be taught first. In Hank's case, attention to task is considered a prerequisite to effective participation in phonological-skills training, so it has first priority. Another useful principle for prioritization is the importance of securing client cooperation and motivation to continue with the treatment. Programs that produce desirable changes quickly will establish motivation and treatment credibility, so are well-placed at the front of the queue.

9.03.6.3 Treatment Implementation

Implementation of a treatment is no less important than selection of the right treatment. Of the wide variety of treatment approaches described in the LD literature, there are a few approaches that are commonly considered among professionals to be effective, and not all of these will be effective in every case. The number of variables that influence the effectiveness of a particular treatment in a given case is much greater than there is space to describe here. It is recognized that treatment implementation is part science and part art (Talbott, Lloyd, & Tankersley, 1994). Having the right procedures is one thing but getting good results with a particular child in a particular setting is another thing altogether, about which very little is to be found in the formal literature on LD.

The identification of weaknesses and strengths at the level of neuropsychological testing may be useful in formulating strategies for effective implementation of the chosen treatment program. The relevant remedial principles are called circumventive and capitalization of strengths (Chronbach & Snow, 1977; Hartlage & Telzrow, 1983). The program can be arranged to minimize the impact of weaknesses on the learning process. For example, distractibility problems can be minimized by initially removing stimuli that distract from the teaching setting, and only progressively reintroducing these when the trained skill has been learned. Problems of sustained attention are addressed by using short training sessions, especially during difficult tasks. Divided attention problems indicate a need to arrange teaching procedures so that tasks are attempted sequentially rather than simultaneously. Memory problems can be minimized in the same way. For example, if visual sequential memory is weak, the learning task can be designed to minimize the need for this type of processing.

Capitalization of strengths essentially is the reverse of this strategy. The remediation program can take advantage of absolute or even relative strengths in say, visual memory by using visual representations of material or procedures to be remembered. The visual prompts can eventually be faded once the task has been mastered (Smeets, 1992).

It is a sovereign principle that standard remedial programs will be most effective if they are modified to suit the requirements of the individual case. This is most obvious at the level of practicalities of program administration (who does what, where, and for how long) and in the designing of an appropriate incentive

Table 2 Examples of remedial programs for LD.

Methods	*Comments*
1. *Reading and spelling*	
A. Multisensory phonological training	
Reading and spelling made simple[a]	A detailed published package, including materials.
	General description of two alternative methods; includes rationale
B. Other systematic approaches	
SRA spelling (McGraw-Hill school division)	Mainstream published program, includes all materials
Distar (McGraw-Hill school division)	Mainstream published program, includes materials
2. *Mathematics*	
Behavioral approaches to skills training[b]	Review of behavioral approaches to math remediation
Spatial skills training[c]	Method for remediating spatial skills problems in math
3. *Attention*	
Self-monitoring[d]	Review of use of self-monitoring with LD students

Source:[a] Andrew (1994) and Bryant and Bradley (1985); [b]Pereira and Winton (1991); [c]Lohman and Nichols (1990); [d](1991).

system. It is less obvious at the level of choice of incentives and choice of specific task materials and least obvious when it comes to incorporating the findings of neuropsychological assessment in the development of circumventive or capitalization-of-strengths strategies.

9.03.6.4 Monitoring of Treatment

The need to monitor the behavior changes that accompany treatment implementation is predicated by the limited knowledge about treatment effectiveness in any individual case. The knowledge base and resources are always insufficient for a perfect assessment, so treatment is best seen as an experiment based on a best-guess as to what will work. Under these circumstances, there is a clear professional duty to monitor progress. Effective monitoring presupposes the prior selection of valid and reliable measures of change of target behaviors. If these are collected and graphed on a regular basis, say weekly, they can be reviewed by the treatment planner for evidence of satisfactory progress and also used to illustrate progress to the client. Thus, monitoring allows for dynamic assessment of the client's response to treatment, so providing a basis for identifying and changing an unsuccessful treatment. Further, it also reinforces treatment credibility to the extent that it shows desirable changes under treatment.

Texts on applied behavior analysis generally recommend that treatment is monitored in the context of some single-subject research design capable of addressing the issue of internal validity, especially the ruling out of interpretations other than that the change in behavior was caused by the specific treatment being given (Hersen & Barlow, 1976; Kazdin, 1982). This is desirable if the case is intended to contribute to the research literature on treatment effectiveness. However, the requirements of a suitable research design are not readily achievable in the context of most clinical work. From the standpoint of clinical practice, it is not essential that a functional relationship between treatment and behavior change be demonstrated with every case in which a particular treatment is used, provided that the relationship has already been established in prior research. This is because the main aim of treatment in the clinical context is to produce change and not to show that a particular manipulation was responsible for the change. Nevertheless, single-subject designs should be used when this is practicable, if only to increase the clinician's understanding of what is happening when the treatment is given. In particular, variations of the multiple-baseline designs are well-suited to the clinical context where they can be used to good effect for monitoring the generality of behavior change under treatment (Hayes, 1981).

9.03.7 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

The term "learning disabilities" was first coined in the mid-1990s in its present sense (Lloyd, 1992), and relevant phenomena were studied long before (Kavale & Forness, 1992).

Much of the ensuing time has been devoted to struggles over matters such as disciplinary ownership, appropriate theoretical models for etiology, and, of course, definition and classification. This period has also seen a staggering diversity of treatment approaches, almost always originating in some narrow (usually contentious) view of the cause of learning disabilities. From this welter of contradictions between different ideas and practices has arisen a theoretical framework that may just be broad enough to encompass most of the phenomena of learning disabilities. There have been some necessary casualties along the way. One is the concept that learning disabilities refer to a homogeneous population; happily, it is rare now to read a paper comparing an LD with a non-LD group on some psychological variable. Of course, the admission of heterogeneity has resulted in a small industry devoted to taxonomy but there is some hope now that this will eventually yield a classificatory strategy that both improves understanding of the individual LD problem and how best to treat it.

Another casualty has been the idea that a medical, psychological, or an educational model alone might be sufficient for LD. There is a growing awareness that the field is truly interdisciplinary and that learning disabilities must be understood in a broad context that includes knowledge, ideas, and practices from neurology, neuropsychology, behavioral pharmacology, applied behavior analysis, cognitive psychology, education, social psychology, clinical psychology, and counseling.

An important casualty has been the persistent belief that there is a one-size-fits-all treatment that will be at least partly effective in all cases. Outside the scientific community, of course, the fads and fallacies department continues to live on and be featured in the media on a regular basis. But few professionals can now be found who give credence to unvalidated therapies for LD. The now-sizable treatment literature seems to show that even the best theoretical pedigree does not ensure that a treatment will be effective even for an appropriately chosen problem. Kavale (1990b, p. 18) has concluded from meta-analysis of published intervention studies that "LD interventions are more variable in their effects than they are beneficial." By way of explanation, he points to evidence of the existence of many factors, outside those usually specified in intervention programs, that likely play an important role in whether or not an intervention succeeds.

Researchers have got as far as recognizing that the effectiveness of LD interventions depends on the identification of variables that have not been considered routinely in theoretical accounts of LD assessment and remediation. These are variables relating to the individual client, the learning environment, and the treatment program itself. The view taken here is that the behavioral assessment model provides a framework that is sufficiently broad to encompass all the classes of variables that might be relevant in the assessment and treatment of learning disabilities. Moreover, the actual procedures specified in the model encourage the assessor to identify and define those variables in the context of the assessment. What is needed now is intervention research based on the behavioral assessment model, employing single-subject research designs that permit adequate control of threats to internal and external validity. Just as classificatory systems for learning disabilities can guide the selection of relevant variables for inclusion in the assessment process, LD research using the behavioral assessment model will provide validation for those classification variables that truly predict treatment outcome at the individual level.

9.03.8 SUMMARY

Once referred to as the "hidden handicap" because of the absence of distinctive diagnostic symptomatology, LD are now recognized widely both by professional and lay communities. Essentially, LD are learning problems that specifically affect one or more skill areas such as reading writing, or social skills in persons who appear to be otherwise normally skilled and educated. Despite continuing controversy about how they should best be defined and diagnosed, there has been little change to widely-accepted consensual definitions in the 1980s and 1990s. Alternative classificatory models have been proposed, based both on conceptual and statistical analyses of patterns of performance in persons with learning disabilities, but no consensual system has emerged. The prevalence of LD in school-aged children is estimated at about 2–20% in the few countries for which figures are available. The most common type, reading disability (developmental dyslexia), appears to have a prevalence of about 5%. Both neurological and environmental factors are implicated in the etiology of LD and, in the case of reading disability, a genetic factor has also been identified. Of the large variety of treatments that have been advocated for LD problems, few have been found to be reliably effective. On the basis of evidence regarding both the varied and complex nature of LD and the high variability in response to any particular treatment, the behavioral assessment model is suggested as being the most appropriate framework for the assessment and treatment of LD problems.

9.03.9 REFERENCES

Algozzine, B., & Ysseldyke, J. (1983). Learning disabilities as a subset of school failure: The over sophistication of a concept. *Exceptional Children, 50,* 242–246.

Aman, M. G., & Rojahn, J. (1992). Pharmacological intervention. In N. N. Singh & I. L. Beale (Eds.), *Learning disabilities: Nature, theory and treatment* (pp. 478–525). New York: Springer.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). New York: Author.

Anderson, J. C., Williams, S., McGee, R., & Silva, P. A. (1987). DSM-III disorders in preadolescent children. *Archives of General Psychiatry, 44,* 69–76.

Andrew, M. (1994). *Reading and spelling made simple* (4th ed.). Morrinsville, New Zealand: Gamlen Press.

Badian, N. A. (1984). Reading disability in an epidemiological context: Incidence and environmental correlates. *Journal of Learning Disabilities, 17,* 129–136.

Bakker, D. J. (1992). Neuropsychological classification and treatment of dyslexia. *Journal of Learning Disabilities, 25,* 102–109.

Barkley, R. (1981). Learning disabilities. In E. J.Mash & L. G. Terdal (Eds.), *Behavioural assessment of childhood disorders* (pp. 441–482). Chichester, UK: Wiley.

Beale, I. L. (1995). Learning disabilities: Current status and future prospects. *Journal of Child and Family Studies, 4,* 267–277.

Beale, I. L., & Tippett, L. J. (1992). Remediation of psychological process deficits in learning disabilities. In N. N. Singh & I. L. Beale (Eds.), *Learning disabilities: Nature, theory and treatment* (pp. 526–568). New York: Springer.

Berger, M., Yule, W., & Rutter, M. (1975). Attainment and adjustment in two geographical areas: The prevalence of specific reading retardation. *British Journal of Psychiatry, 126,* 510–519.

Brier, N. (1989). The relationship between learning disability and delinquency: A review and reappraisal. *Journal of Learning Disabilities, 22,* 546–553.

Brier, N. (1994). Targeted treatment for adjudicated youth with learning disabilities: Effects on recidivism. *Journal of Learning Disabilities, 27,* 215–222.

Broder, P. K., Dunivant, N., Smith, E. C., & Sutton, L. P. (1981). Further observations on the link between learning disabilities and juvenile delinquency. *Journal of Educational Psychology, 73,* 838–850.

Brown, A. L., & Palincsar, A. S. (1982). Inducing strategic learning from texts by means of informed, self-controlled training. *Topics in Learning and Learning Disabilities, 2,* 1–17.

Bruck, M. (1987). The adult outcomes of children with learning disabilities. *Annals of Dyslexia, 31,* 252–263.

Bryant, P., & Bradley, L. (1985). *Children's reading problems.* Oxford, UK: Blackwell.

Chalfant, J. C., & King, F. S. (1976). An approach to operationalizing the definition of learning disabilities. *Journal of Learning Disabilities, 9,* 228–243.

Chapman, J. W. (1992). Learning disabilities in New Zealand: Where kiwis and kids with LD can't fly. *Journal of Learning Disabilities, 25,* 362–370.

Chronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods.* New York: Irvington.

Clements, S. D. (1966). *Minimal brain dysfunction in children. Terminology and identification.* NINDS Monograph No. 3, U.S. Public Health Service Publication No. 1415. Washington, DC: US Department of Health, Education, and Welfare.

Corballis, M. C., & Beale, I. L. (1983). *The ambivalent mind: The neuropsychology of left and right.* Chicago: Nelson-Hall.

Corballis, M. C., & Beale, I. L. (1993). Orton revisited: Dyslexia, laterality and left-right confusion. In D. M. Willows, R. S. Kruk, & E. Corcos (Eds.), *Visual processes in reading and reading disabilities* (pp. 57–74). Hillsdale, NJ: Erlbaum.

Cullinan, D., Epstein, M., & Lloyd, J. W. (1981). School behaviour problems of learning disabled and normal girls and boys. *Learning Disability Quarterly, 4,* 163–169.

deFries, J. C. (1989). Gender ratios in reading-disabled children and their affected relatives. *Journal of Learning Disabilities, 22,* 544–555.

Delacato, C. H. (1966). *The treatment and prevention of reading problems.* Springfield, IL: Charles C. Thomas.

Doehring, D. G., Hoshko, I. M., & Bryans, B. N. (1979). Statistical classification of children with reading problems. *Journal of Clinical Neuropsychology, 1,* 5–16.

Epps, S., Ysseldyke, J. E., & Algozzine, B. (1985). Impact of different definitions of learning disabilities on the number of students identified. *Journal of Psychoeducational Assessment, 1,* 341–352.

Forness, S. R. (1990). Subtyping in learning disabilities: Introduction to the issues. In H. L. Swanson & B. Keogh (Eds.), *Learning disabilities: Theoretical and research issues* (pp. 195–200). Hillsdale, NJ: Erlbaum.

Frostig, M., & Horne, D. (1964). *The Frostig program for the development of visual perception.* Chicago: Follett.

Gaddes, W. H. (1985). *Learning disabilities and brain function.* New York: Springer.

Geshwind, N., & Galaburda, A. M. (1987). *Cerebral lateralization: Biological mechanisms, associations and pathology.* Cambridge, MA: Bradford/MIT Press.

Hammill, D. D., & Larsen, S. C. (1978). The effectiveness of psycholinguistic training: A reaffirmation of position. *Exceptional Children, 44,* 402–414.

Harness, B. Z., Epstein, R., & Gordon, H. W. (1984). Cognitive profile of children referred to a clinic for reading disabilities. *Journal of Learning Disabilities, 17,* 346–351.

Hartlage, L. C., & Telzrow, C. F. (1983). The neuropsychological basis of educational intervention. *Journal of Learning Disabilities, 16,* 521–528.

Hayes, S. C. (1981). Single case experimental design and empirical clinical practice. *Journal of Consulting and Clinical Psychology, 49,* 193–211.

Hersen, M., & Barlow, D. H. (1976). *Single-case experimental designs: Strategies for studying behaviour change.* New York: Pergamon.

Hoehn, T. P., & Baumeister, A. A. (1994). A critique of the application of sensory integration therapy to children with learning disabilities. *Journal of Learning Disabilities, 27,* 338–351.

Hulme, C. (1988). The implausibility of low-level visual deficits as a cause of children's reading difficulties. *Cognitive Neuropsychology, 5,* 369–374.

Hurford, D. P., Schauf, J. D., Bunce, L., Blaich, T., & Moore, K. (1994). Early identification of children at risk for reading disabilities. *Journal of Learning Disabilities, 27,* 371–382.

Hynd, G. W., & Hynd, C. R. (1984). Dyslexia: Neuroanatomical/neurolinguistic perspectives. *Reading Research Quarterly, 19,* 482–498.

Individuals with Disabilities Act (1990). *Public Law 101–476, Section 602(a)(19), October, 1990.* Washington, DC: US Office of Education.

Junkala, J., & Paul, J. (1987). Special monograph issue on international perspectives on learning disabilities. *Paedoperisse: An International Journal on Comparative Special Education, 1,* 1–107.

Kamphaus, R. (1993). *Clinical assessment of children's intelligence*. Boston: Allyn and Bacon.

Kass, C. E., & Myklebust, H. R. (1969). Learning disability: An educational definition. *Journal of Learning Disabilities, 2,* 377–379.

Kavale, K. A. (1981). Functions of the Illinois Test of Psycholinguistic Abilities (ITPA): are they trainable? *Exceptional Children, 47,* 496–510.

Kavale, K. A. (1990a). A critical appraisal of empirical subtyping research in learning disabilities. In H. L. Swanson & B. Keogh (Eds.), *Learning disabilities: Theoretical and research issues* (pp. 215–232). Hillsdale, NJ: Erlbaum.

Kavale, K. A. (1990b). Variances and verities in learning disabilities interventions. In T. E. Scruggs & B. Y. L. Wong (Eds.), *Intervention research in learning disabilities* (pp. 3–33). New York: Springer.

Kavale, K. A., & Forness, S. R. (1985). Learning disability and the history of science: Paradigm or paradox? *RASE, 6,* 12–23.

Kavale, K. A., & Forness, S. R. (1992). History, definition, and diagnosis. In N. N. Singh & I. L. Beale (Eds.), *Learning disabilities: Nature, theory, and treatment* (pp. 3–43). New York: Springer.

Kavale, K. A., & Forness, S. R. (1996). Social skills deficits and learning disabilities: A meta-analysis. *Journal of Learning Disabilities, 29,* 226–237.

Kavanagh, J. F., & Truss, T. J. (1988). Revised definition of learning disabilities. In J. F. Kavanagh & T. J. Truss (Eds.), *Learning disabilities: Proceedings of the national conference* (pp. 549–551). Parkton, MD: York Press.

Kazdin, A. (1982). *Single-case research designs: Methods for clinical and applied settings*. New York: Oxford University Press.

Lam, C. M., & Beale, I. L. (1991). Relation between sustained attention, reading performance and teachers' ratings of behaviour problems. *Remedial and Special Education, 12,* 40–48.

Lloyd, J. W. (1992). A commentary on learning disabilities. In N. N. Singh & I. L. Beale (Eds.), *Learning disabilities: Nature, theory and treatment* (pp. 569–586). New York: Springer.

Lloyd, J. W., Bateman, D. F., Landrum, T. J., & Hallahan, D. P. (1989). Self-recording of attention versus productivity. *Journal of Applied Behaviour Analysis, 22,* 315–323.

Lohman, D. F., & Nichols, P. D. (1990). Training spatial abilities: Effects of practice on rotation and synthesis tasks. *Learning and Individual Differences, 2,* 67–93.

Lovegrove, W. (1992). The visual deficit hypothesis. In N. N. Singh and I. L. Beale (Eds.), *Learning disabilities: Nature, theory and treatment* (pp. 246–269). New York: Springer.

Marshall, K. J., Lloyd, J. W., & Hallahan, D. P. (1993). The effects of training to increase self-monitoring accuracy on the attention-to-task of learning-disabled children. *Journal of Behavioural Education, 3,* 445–459.

Mattis, S. (1978). Dyslexia syndromes: A working hypothesis that works. In A. L. Benton & D. Pearl (Eds.), *Dyslexia: An appraisal of current knowledge* (pp. 43–60). New York: Oxford University Press.

McConaughy, S. H., & Ritter, D. R. (1986). Social competence and behavioural problems of learning disabled boys aged 6–11. *Journal of Learning Disabilities, 19,* 39–45.

Minskoff, E. H., Hawks, R., Steidle, E. F., & Hoffman, F. J. (1989). A homogeneous group of persons with learning disabilities: Adults with severe learning disabilities in vocational rehabilitation. *Journal of Learning Disabilities, 22,* 521–528.

Motluk, A. (1996). Why English is hard on the brain. *New Scientist, 20 January,* 14.

Mudford, O. C., Beale, I. L., & Singh, N. N. (1990). The representativeness of observational samples of different durations. *Journal of Applied Behaviour Analysis, 23,* 323–331.

National Advisory Committee on Handicapped Children (1968). *The first annual report: Special education for handicapped children*. Washington, DC: US Office of Education, Department of Health, Education, and Welfare.

National Joint Commission on Learning Disabilities (1981). *Learning disabilities: Issues on definition*. A position paper of the National Joint Committee on Learning Disabilities.

Narhi, V., & Ahonen, T. (1995). Reading disability with or without attention deficit hyperactivity disorder: Do attentional problems make a difference? *Developmental Neuropsychology, 11,* 337–349.

Nelson, R. O., & Hayes, S. C. (1986). *Conceptual foundations of behavioral assessment*. New York: Guilford.

Nichols, P., & Chen, T. C. (1981). *Minimal brain dysfunction: A prospective study*. Hillsdale, NJ: Erlbaum.

Orton, S. T. (1937). *Reading, writing and speech problems in children*. New York: Norton.

Pennington, B. F. (1991). *Diagnosing learning disorders*. New York: Guilford.

Pennington, B. F., Groisser, D., & Welsh, M. C. (1993). Contrasting cognitive deficits in attention deficit hyperactivity disorder versus reading disability. *Developmental Psychology, 29,* 511–523.

Pereira, J. A., & Winton, A. S. W. (1991). Teaching and remediation of mathematics: A review of behavioral research. *Journal of Behavioral Education, 1,* 5–36.

Petrauskas, R. J., & Rourke, B. P. (1979). Identification of subtypes of retarded readers: A neuropsychological, multivariate approach. *Journal of Clinical Neuropsychology, 1,* 17–37.

Pirozzolo, F. J. (1979). *The neuropsychology of developmental reading disorders*. New York: Praeger.

Public Law 94–142 (1975), 20 U.S.L. 1401 et seq. 1976.

Radencich, M. C. (1984). The status of learning disabilities: The emergence of a paradigm or a paradigm shift? *Learning Disabilities, 3,* 79–89.

Reid, R. (1996). Research in self-monitoring with students with learning disabilities: The present, the prospects, the pitfalls. *Journal of Learning Disabilities, 29,* 317–331.

Reiff, H. B., & Gerber, P. J. (1992). Adults with learning disabilities. In N. N. Singh & I. L. Beale (Eds.), *Learning disabilities: Nature, theory, and treatment* (pp. 170–200). New York: Springer.

Rourke, B. P. (1982). Central processing deficiencies in children: Towards a developmental neuropsychological model. *Journal of Clinical Neuropsychology, 4,* 1–18.

Rourke, B. P. (1985). *Neuropsychology of learning disabilities: Essentials of subtype analysis*. New York: Guilford.

Rourke, B. P., & Strang, J. D. (1983). Subtypes of reading and arithmetical disabilities: A neuropsychological analysis. In M. Rutter (Ed.), *Developmental neuropsychiatry* (pp. 473–488). New York: Guilford.

Rourke, B. P., Young, G. C., & Leenaars, A. A. (1989). A childhood learning disability that predisposes those afflicted to adolescent and adult depression and suicide risk. *Journal of Learning Disabilities, 22,* 169–175.

Rozin, P., Poritsky, S., & Sotsky, R. (1971). American children with reading problems can easily learn to read English represented by Chinese characters. *Science, 171,* 1264–1267.

Rutter, M., Graham, P., & Yule, W. (1970). *A neuropsychiatric study in childhood.* London: Spastics International Publications.

Salvia, J., & Ysseldyke, J. E. (1991). *Assessment*. Boston: Houghton Mifflin.

Sanson, A., Prior, M., & Smart, D. (1996). Reading disabilities with and without behaviour problems at 7–8 years: Prediction from longitudinal data from infancy to 6 years. *Journal of Child Psychology and Psychiatry, 37,* 529–541.

Sattler, J. M. (1992). *Assessment of children*. San Diego, CA: Sattler.

Satz, P., & Morris, R. (1981). Learning disability subtypes: A review. In F. J. Pirozzolo & M. C. Wittrock (Eds.), *Neuropsychological and cognitive processes in reading*. New York: Academic Press.

Satz, P., & Sparrow, S. S. (1970). Specific developmental dyslexia: A theoretical formulation. In D. J. Bakker & P. Satz (Eds.), *Specific reading disability: Advances in theory and method* (pp. 17–40). Rotterdam, The Netherlands: Rotterdam University Press.

Schere, R. A., Richardson, E., & Bailer, I. (1980). Towards operationalizing a psychoeducational definition of learning disabilities. *Journal of Abnormal Child Psychology, 8,* 5–20.

Shankweiler, D., Liberman, I., Mark, L., Fowler, C., & Fisher, F. (1979). The speech code and learning to read. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 531–545.

Share, D. L., Moffitt, T. E., & Silva, P. A. (1988). Factors associated with arithmetic-and-reading and specific arithmetic disability. *Journal of Learning Disabilities, 21,* 313–320.

Shaywitz, S. E., Shaywitz, B. A., Fletcher, J. M., & Escobar, M. D. (1990). Prevalence of reading disabilities in boys and girls: Results of the Connecticut Longitudinal Study. *Journal of the American Medical Association, 264,* 998–1002.

Siegel, L. S., & Metsala, J. (1992). An alternative to the food processor approach to subtypes of learning disabilities. In N. N. Singh & I. L. Beale (Eds.), *Learning disabilities: Nature, theory, and treatment* (pp. 44–60). New York: Springer.

Silver, A. A., & Hagin, R. A. (1990). *Disorders of learning in childhood.* New York: Wiley.

Silver, L. B. (1987). The "magic cure": A review of the current controversial approaches for treating learning disabilities. *Journal of Learning Disabilities, 20,* 498–512.

Silverman, L. A., & Metz, A. (1973). Numbers of pupils with specific learning disabilities in local public schools in the United States. *Annals of the New York Academy of Sciences, 205,* 146–157.

Smeets, P. M. (1992). Prompt-aided instruction of mirror-image discriminations: Abrupt prompt removal, time delay, and self-evaluation. *Journal of Behavioural Education, 2,* 87–104.

Stone, W. L., & LaGreca, A. M. (1984). Comprehension of nonverbal communication: A reexamination of the social competencies of learning disabled children. *Journal of Abnormal Child Psychology, 12,* 205.

Strauss, A. A., & Lehtinen, L. E. (1947). *Psychopathology and education of the brain-injured child.* New York: Grune & Stratton.

Talbott, E., Lloyd, J. W., & Tankersley, M. (1994). Effects of reading comprehension interventions for students with learning disabilities. *Learning Disability Quarterly, 17,* 223–233.

Tallal, P. (1988). Developmental language disorders. In J. F. Kavanagh & T. J. Truss (Eds.), *Learning disabilities: Proceedings of the national conference* (pp. 181–272). Parkton, MD: York Press.

Torgesen, J. K. (1979). What shall we do with psychological processes? *Journal of Learning Disabilities, 12,* 16–23.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1994). Longitudinal studies of phonological processing and reading. *Journal of Learning Disabilities, 27,* 276–286.

Treiber, F. A., & Lahey, B. B. (1983). Toward a behavioural model of academic remediation with learning disabled children. *Journal of Learning Disabilities, 16,* 111–116.

US Department of Education (1991). *To assure free appropriate public education of all children with disabilities*. Thirteenth annual report to Congress on the implementation of the individuals with disabilities. Washington, DC: US Government Printing Office.

US General Accounting Office (1981). *Disparities still exist in who gets special education.* Washington, DC: Comptroller General of the United States.

Vellutino, F. R. (1979). The validity of perceptual deficit explanations of reading disability: A reply to Fletcher and Satz. *Journal of Learning Disabilities, 12,* 160–167.

Vogel, S. (1990). Gender differences in intelligence, language, visual-motor abilities, and academic achievement in students with learning disabilities. *Journal of Learning Disabilities, 23,* 44–52.

Willis, W. G., Hooper, S. R., & Stone, B. H. (1992). Neuropsychological theories of learning disabilities. In N. N. Singh & I. L. Beale (Eds.), *Learning disabilities: Nature, theory and treatment* (pp. 201–245). New York: Springer.

Willows, D. M., Kruk, R., & Corcos, E. (1993). Are there differences between disabled and normal readers in their processing of visual information? In D. M. Willows, R. S. Kruk, & E. Corcos (Eds.), *Visual processing in reading and reading disabilities* (pp. 265–285). Hillsdale, NJ: Erlbaum.

Witelson, S. (1977). Developmental dyslexia: Two right hemispheres and none left. *Science, 195,* 309–311.

Wong, B. Y. L. (1979). Increasing retention of main ideas through questioning strategies. *Learning Disability Quarterly, 2,* 42–47.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.04
Visual Disabilities

STEVEN J. LaGROW and STEVEN HUMPHRIES
Massey University, Palmerston North, New Zealand

9.04.1 INTRODUCTION 57
9.04.1.1 Type 57
9.04.1.2 Severity 58
9.04.1.3 Age of Onset 58
9.04.2 NATURE 59
9.04.2.1 Reaction to Loss 59
9.04.2.2 Attitude Toward Disability 59
9.04.2.3 Sudden and Traumatic Onset 60
9.04.2.4 Age of Onset 60
9.04.2.5 Other Health Issues 61
9.04.3 PREVALENCE 61
9.04.3.1 Severe Visual Impairment 61
9.04.3.2 Legal and Total Blindness 61
9.04.3.3 Leading Causes 61
9.04.4 ASSESSMENT 62
9.04.4.1 Psychometric Testing 62
9.04.4.2 Measures of Adjustment 63
9.04.4.3 Measures of Functional Ability 63
9.04.5 TREATMENT 63
9.04.5.1 Skill Acquisition 63
9.04.5.2 Counseling and Support 64
9.04.5.3 Delivery Systems 65
9.04.5.4 Other Services 65
9.04.5.5 Services for Children 65
9.04.6 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 65
9.04.7 CONCLUSIONS 66
9.04.8 REFERENCES 66

9.04.1 INTRODUCTION

Visual disabilities result from a loss of function due to a visual impairment (Colenbrander, 1976). A visual disability may impose a number of limitations on the activities of an individual and restrict their level of independence. The number and degree of limitations experienced is dependent upon the type and severity of the disability, as well as the age and health of the individual affected (L. M. Teitelbaum, Davidson, Gravetter, Taub, & C. S. Teitelbaum, 1994).

9.04.1.1 Type

Visual disabilities may be of the progressive or nonprogressive type and categorized according to location. The location refers to the field of view affected. Demonstrable field defects may or may not occur. When they do, they are defined as occurring in either the central or peripheral field of view (Faye, 1976). Persons with a central field defect have difficulty with reading and other close work but experience few problems with mobility, while just the opposite is true for those with a peripheral field defect. Those with no demonstrable field defect have difficulty performing most tasks since they experience a reduction of visual ability across the entire field of view (Faye, 1976).

The degree of difficulty experienced with any of these types of disability is dependent on both its severity and age of onset. Both of these categories may be treated as being dichotomous. As a result, a visual disability is often seen as being either congenital (present at birth) or adventitious (not present at birth) and as producing either a total (blind) or partial loss of sight (low vision). A partial loss of sight is classified further by severity.

9.04.1.2 Severity

The severity of a visual loss can range anywhere along a continuum from minimal involvement to total blindness. The severity of that loss may be measured in terms of both visual acuity and visual field. An acuity rating is a standard measure of visual reception, is reported as a ratio, and provides a direct comparison to normative sight. The standard acuity measure for normal vision is 20/20 when measured in feet or 6/6 when reported in meters. Generally, the greater the denominator (e.g., 20/200 vs. 20/20) the worse are visual functions (Jose, 1984).

Persons with visual acuities of 20/70 or worse, for example, have difficulty reading regular newsprint and may be classified as being severely visually impaired for statistical purposes (Kirchner, 1985). Those with visual acuities of 20/200 or worse in the better eye after the best possible correction are considered to be legally blind. Persons may also be considered legally blind if their field of view does not exceed 20 degrees at its widest angle, regardless of their acuity ratings. To be considered totally blind one must have no usable or functional vision.

9.04.1.3 Age of Onset

Age of onset also plays a major role in determining the effect the visual loss will have on functioning. The earlier a significant loss of sight occurs in a child's life, for example, the more likely it is that the developmental process will be affected as well (Scholl, 1986). These children generally experience some delays or unevenness in their sensory perceptual, motor, concept, language, and social development (Ferrell, 1986; Scholl, 1986; Warren, 1984). These delays may result in deleterious effects in a number of areas of functioning including interpersonal and social skill development (Huebner, 1986). As a result, visually impaired children are thought to be at risk of social and emotional maladjustment (Ammerman, Van Hasselt, & Hersen, 1986; Van Hasselt, 1983). Infants often display deficits in those behaviors required for the development of secure mother–infant attachment (e.g., appropriate eye contact, responsive smiling, responsive motor responding) (Ammmerman, Van Hasselt, Hersen, & Moore, 1989). They also have difficulty modeling appropriate social behaviors and spontaneously and purposefully producing smiles and other facial expressions required in social interaction with their sighted peers (Ammerman et al., 1986; Van Hasselt, Hersen, Kazdin, Simon, & Mastantuono, 1985). Yet, results from research on the social and emotional adjustment of visually impaired children have generally been mixed, with a number of studies suggesting that these children have low self-concept, subnormal social functioning, poorer social experiences, poorer psychological and emotional adjustment, and a higher rate of social dysfunction within their family unit than their sighted peers (Beaty, 1991; Warren, 1984), while others have found no difference or even better adjustment in the visually impaired sample (Alexander, 1996; Ammerman et al., 1989; Beaty, 1994; Head, 1980; Van Hasselt et al., 1985). In general, it can be concluded that although there is a full range of social, emotional, and psychological adjustment exhibited among visually impaired children there is little evidence to suggest that there are any particular patterns of dysfunction among congenitally blind and visually impaired persons (Beaty, 1994; Warren, 1984). The only real exception to this conclusion is the relatively high incidence of stereotypic responding exhibited among visually impaired children especially those who are multiply disabled (Warren, 1984). However, these behaviors are not unique to visually impaired children nor are the interventions used in their treatment (Miner, 1991; LaGrow & Repp, 1984).

Adventitiously blinded or visually impaired adults, on the other hand, tend to exhibit a pattern of maladjustment involving a high rate of depressive states and other social, emotional, and psychological problems (Branch, Horowitz, & Carr, 1989; Van Hasselt, 1983). Yet, these problems are not an inevitable consequence of reduced visual function (Teitelbaum et al., 1994). Rather, they appear to be more in response to the loss of function or role associated with the onset of visual impairment (Dale, 1992; Dodds, Bailey, Pearson, & Yates, 1991; Hudson, 1994; Van Zandt, Van Zandt, & Wang, 1994). This is specially evident immediately following the sudden onset of a visual impairment (Carroll, 1961; Dale, 1992; Dodds et al., 1991) and seems to be exacerbated by old age (Hersen et al., 1995; Robbins & McMurray, 1988) and poor health (Bernbaum, Albert, & Duckro, 1988; Zarit, 1992).

9.04.2 NATURE

9.04.2.1 Reaction to Loss

The most common psychological responses exhibited following the onset of a significant visual impairment include feelings of shock, regression, denial, grief, depression, devaluation, and anxiety (Schulz, 1980). These responses are common reactions to any significant loss in life and not unique to the loss of sight. They are also not invariant, although most people will experience some of these reactions and some will experience all of them in response to a significant loss. The importance of the loss to the individual will have a bearing on the way the individual responds, as will its suddenness and completeness (Schulz, 1980).

These responses may appear to be particularly common following the onset of a significant visual impairment because the loss of vision results in numerous other losses to the individual. These are generally thought to include a loss of visual function for social communication, written communication, and various independent living skills including mobility which variously results in a loss of personal independence, social and familial role, psychological and financial security, and ultimately self-esteem (Carroll, 1961; Hersen et al., 1995; Teitelbaum et al., 1994; Van Hasselt, 1983; Van Zandt et al., 1994).

Carroll (1961) delineated no less than 20 separate losses which may be associated with the onset of a significant visual impairment. Some of these are temporary and occur immediately following a loss of sight but do not persist long after the individual has grown accustomed to the resulting change in function (e.g., the loss of reality contact with the environment, light security, and confidence in the remaining senses). Others are a direct result of the loss of sight and may persist a lifetime (e.g., the loss of physical integrity, visual background, the visual perception of the pleasurable, and the visual perception of the beautiful). A number are related to the functional limitations associated with the loss of sight and can often be overcome or minimized with appropriate intervention (e.g., the loss of mobility, the various techniques of daily living, written communication, spoken communication, recreational activities, and job). The remainder are associated with those disabilities and may change as the disabilities are minimized and skill in functioning is regained (e.g., the loss of personal independence, autonomy, social adequacy, self-esteem, meaningful interpersonal relationships, career, vocational goal, financial security, and various social and familial roles).

The loss of functioning associated with the onset of a significant visual impairment (e.g., mobility, various daily living skills, communication skills, recreation activities, job, and ultimately personal independence, autonomy, social adequacy, career prospects, financial security, and myriad social roles) may be shattering to one's sense of self (e.g., self-confidence, self-esteem, self-worth) (Delafield, 1976). Part of the depression so commonly seen in persons with recent sight loss may result from the shock to one's sense of self that comes from realizing that even the simplest tasks (e.g., dressing, eating, cooking, using a telephone, handling money, telling time, moving from room to room and place to place) may no longer be completed without considerable difficulty (Dodds et al., 1991).

9.04.2.2 Attitude Toward Disability

The degree to which one's sense of self is affected by the onset of a disability is obviously associated with the severity of the disability (Delafield, 1976; Dodds et al., 1991). The permanence of that effect, however, often has more to do with one's perception of the abilities and capabilities of disabled people in general, and blind and visually impaired people in particular, than the severity of the disability itself (Harrell & Strauss, 1986; Hersen et al., 1995). Some persons, for example, may be so hampered by their prejudices that they cannot recognize the difference between real limitations and perceived losses. They truly believe that they cannot learn to adjust. They believe that blindness is the worst thing that can befall them. They believe that they are helpless and to be pitied. The usually temporary loss of self-esteem, self-confidence, and/or self-worth resulting from the onset of a disability may be prolonged by their perception of what it means to be blind.

On the other hand, individuals who are not overwhelmed by a sense of helplessness, inadequacy, dependence, and self-pity will eventually become more comfortable, feel less inadequate, and begin to experience success in their ability to accommodate the change in their physical status (Dodds et al., 1991). As a result, the losses experienced in mobility, daily living skills, communication, recreation, job, career, and even financial security may be lessened or disappear. Thus, personal independence and/or autonomy may be re-established, as could the various social roles one fills, including those related to job and career. This success will build a more positive self-concept and a renewed feeling of social adequacy (Tuttle, 1984).

9.04.2.3 Sudden and Traumatic Onset

The suddenness of the onset of the visual impairment will also affect one's response to it, with the more sudden and unexpected causes of visual impairment resulting in the greatest difficulties with adjustment (Delafield, 1976; Dodds et al., 1991). This is especially true in the case of traumatic blindness. Traumatic blindness is defined as a sudden loss of sight, either partial or total, resulting from an external or internal somatic event such as a head injury, anoxia, toxic ingestion, a stroke, a severe adverse drug reaction, chemical burns, or surgical failure (Dale, 1992). The traumatic event itself may be more shocking than the resultant vision loss. Thus, the event may dominate the individual's attention, at least during the early stages of adjustment.

The process of adjustment may be delayed by a period of psychic numbness following the event. This may also have a negative effect on one's existing social support system because the normal and expected reactions to the trauma are not exhibited. Friends and families may find it difficult to relate to the individual who does not appear to be reacting normally. They may feel excluded, unwanted, useless, or even guilty. They may be traumatized by the event as well and, therefore, not be of much help to the affected individual.

Family members may have to take on new roles or cope with other members whose roles are changed as a result of the event or the resultant visual impairment. The family's financial security may be jeopardized by the cost of the medical care, the loss of income, or both. As a result, the individual's social support network may also be in need of support. Problems with interactions caused by this will not aid the adjustment process for either the family or the visually impaired family member. Post-traumatic stress syndrome may accompany or follow the event and, therefore, further complicate and prolong the process of adjustment (Dale, 1992).

In addition, the cause of the impairment is usually known and clearly identifiable in cases of traumatic blindness. As a result, the individual may spend a great deal of time reliving the circumstances leading up to it. This may be especially true if the circumstances were avoidable and/or caused by one's own (e.g., suicide attempt, substance abuse, careless or drunken driving) or another's actions (e.g., accident, criminal assault, careless or drunken driving, medical misadventure). In this case, the individual's thoughts may be dominated by a sense of guilt, remorse, blame, or wish for revenge (Dale, 1992).

Persons who blame others or themselves for events which affect their lives, or strongly hold to a sense of justice in the world, tend to have the greatest difficulty adjusting to the sudden and traumatic onset of visual impairment (Heinemann, Bulka, & Smetak, 1988; Malec, 1985; Rohe & Basford, 1989). Older persons also tend to have more difficulty adjusting to traumatic blindness than their younger peers (Dale, 1992).

9.04.2.4 Age of Onset

The older the person is the more likely he or she is to have trouble adjusting to a significant visual impairment whether it is the result of trauma or a disease process (Teitelbaum et al., 1994). This appears to reflect a cumulative effect among the losses associated with both aging and the onset of a visual impairment (Branch et al., 1989; Hersen et al., 1995; Van Zandt et al., 1994). Older persons also tend to be less tolerant of stress and more likely to experience other symptoms of ill health than their younger peers (Fagerstrom, 1992; Zarit, 1992).

Elderly persons who experience a loss in vision are often demoralized by this additional loss to their physical, social, and/or financial integrity (Branch et al., 1989; Vaughan & Hobson, 1990; Zarit, 1992). As a result, they tend to withdraw and become passive. They may complain a lot but do little to try to improve the situation they are in. They often refuse to try to learn to do new things or routine things in a new way. They also tend to reject the aids and devices which may be used to accommodate but not eliminate their vision loss. Often, their only goal is complete recovery. Anything less is dismissed as being unhelpful (Zarit, 1992). These behaviors may ultimately undermine their social support network and leave them isolated.

As a result, older persons often lose their place in their existing social support system or volintarily withdraw due to a perception that they are no longer able to reciprocate in social situations (Hersen et al., 1995). This is especially common if persons in the network tend to be paternalistic or overprotective in relation to the individual concerned (Van Zandt et al., 1994). Unfortunately, severe depression is highly correlated with the loss of social support in this population (Vaughan & Hobson, 1990).

9.04.2.5 Other Health Issues

Many of the conditions that result in visual impairment cause other physiological difficulties as well. The rate of chronic health conditions, such as hearing loss, arthritis, high blood pressure, diabetes, heart conditions, and orthopedic conditions is high among elderly blind persons (Kirchner & Peterson, 1988). When severe, these difficulties may compound the situation and intensify both the physical and emotional problems experienced in the adjustment process (Needham, Elridge, Harabedian, & Crawford, 1993; Rosenbloom, 1992).

Persons who experience severe visual impairments as a result of diabetes, for example, generally experience numerous other health and functional problems, including neuropathy, nephropathy, hypertension, arteriosclerosis, and myocardial infarction (Berkow, 1982). They have also been found to have significant levels of depression, poor self-esteem, and decreased psychological well-being (Bernbaum, Albert, & Duckro, 1993). They experience very high divorce rates compared to both the general population and the population of diabetics who are not visually impaired (Drury, Danchick, & Harris, 1985).

Disruptions to marriages and committed relationships appear to be directly related to the severity of the vision loss and occur relatively soon after its onset (i.e., within a year and a half) (Bernbaum et al., 1993). The stressors that cause divorce in this population are apparent across all groups of newly blinded persons. In fact, blindness in one person inevitably becomes a family affair (Jackson & Lawson, 1995). Roles of family members are invariably affected and financial security may be threatened due to lost wages, medical bills, or both. This is especially apparent when one's physical health is threatened, which is commonly the case among elderly, visually impaired persons who make up the bulk of the newly blinded population (Orr, 1992).

9.04.3 PREVALENCE

9.04.3.1 Severe Visual Impairment

Approximately 11 million people in the USA have some kind of visual impairment (Kirchner & Peterson, 1988). About three million of these are severely visually impaired (i.e., have difficulty reading newsprint). Nearly 70% of those with severe visual impairments are age 65 and over (Davis, Lovie-Kitchen, & Thompson, 1995). About 25% of the visually impaired population are 18–65 years of age and only about 5% are under 18 (Rosenbloom, 1992).

9.04.3.2 Legal and Total Blindness

Just over 20% of the severely visually impaired population are legally blind and 10–15% of that population are totally blind (Rosenbloom, 1992). Only about 6% of all cases of legal blindness are due to trauma (National Society for the Prevention of Blindness, 1988). Most are due to progressive disease and are often associated with conditions which cause other health or physical/sensory problems as well (Needham et al., 1993).

Although rare among children, adolescents, and young adults, the rate of severe visual impairment increases dramatically with age (Yeadon, 1984). In primary school age children, for example, the rate is around 0.5 per 100 and in all persons under 24 years it is just over 0.5 per 1000. However, it increases nearly 100-fold for those who are aged 65 to 74 to a rate of 47 per 1000. For those aged 75 to 84, it is 99 per 1000; and beyond 85, it jumps to 250 per 1000, or 1 in 4 (Van Zandt et al., 1994).

9.04.3.3 Leading Causes

As would be expected all the leading causes of visual impairment are age related. The most common of these are macular degeneration, cataracts, glaucoma, and diabetic retinopathy (Davis et al., 1995; Orr, 1992; Rosenbloom, 1992). All are adventitious, progressive, and usually result in a partial loss of sight, although both glaucoma and diabetic retinopathy may result in total blindness. In addition, a person may become totally blind due to senile cataracts. Yet, partial vision can usually be regained through surgery. Macular degeneration, by far the most common cause of new blindness (i.e., legal blindness), results in a central field loss but only rarely progresses to total blindness. Glaucoma results in a peripheral field loss, while cataracts and diabetic retinopathy generally result in an overall or, as Faye (1976) describes it, no demonstrable field loss.

9.04.4 ASSESSMENT

A thorough assessment is necessary before it can be assumed that the person to be served is known or understood (Vander Kolk, 1981). In the case of visually impaired persons, this would usually include ophthalmological, optometric, and functional evaluations as well as more general medical and psychological assessments.

Both ophthalmological and general medical examinations are required to establish the extent and nature of the visual disability as well as any secondary disabilities or health problems which may affect the individual's functional ability and physical well-being. Optometric assessments identify the individual's acuity, refractive error, and field of view. Functional assessments may be used to determine the way the individual functions on any number of daily living tasks. Psychological assessments may be conducted to identify the individual's aptitudes, interests, and abilities as well as their acceptance of and adjustment to vision loss. As a result, a number of instruments and a variety of methods including clinical examination, psychometric testing, behavioral observations, and self-report may be used.

9.04.4.1 Psychometric Testing

Psychometric assessment is widely used with visually impaired individuals (Swallow, 1981). However, there are only a limited number and range of tests designed specifically for use with this population (Johnson, 1989). Those which are available generally concentrate on either cognitive factors (i.e., IQ tests) or on aptitudes and abilities (e.g., tests of body and spatial concepts) (Dodds, Craig, & Flannigan, 1996). The latter are of limited applicability, designed for use with children, and used for planning specific educational interventions.

Most tests used with visually impaired persons have been designed for use with normally sighted persons and therefore must be modified in some way for use with this population. This is usually done in terms of content and/or administrative style (Needham & Eldridge, 1990). Administration of the test may be modified by reading the test aloud to the individual or into a tape recorder, reproducing the test in enlarged print or Braille, and/or allowing the individual to respond orally or by typing (Fuchs, Fuchs, Benowitz, & Barringer, 1987; Johnson, 1989). The content of tests is often modified by dropping items seen to be inappropriate for visually impaired persons. The entire performance sections of some tests (e.g., Stanford Binet and WISC-R) are dropped, while individual items considered to be visually loaded may be omitted from others (Ammerman et al, 1986; Jan, Freeman, & Scott, 1977).

Any modification from the original format, however, may alter the results of the test or otherwise disadvantage those for whom the modifications are made (Ammerman et al., 1989; Johnson, 1989). The use of Braille, for example, can be problematic for some because only a small percentage of the blind population are proficient with its use while emphasis and intonation accompanying an oral presentation of a written question may change its meaning (Johnson, 1989). Furthermore, modifications in presentation and response format often greatly increase the time and energy required to complete the test (Barker, Wright, Meyerson, & Gonick, 1953). This last factor has especially strong implications for the administration of timed tests to people with visual impairments.

There are a number of problems associated with dropping "invalid" test items. First, it is not always easy to detect which test items are visually loaded (Jan et al., 1977). Indeed, attempts to develop items whose solution is not dependent on sight, and that could be solved equally well by blind and sighted subjects, is extremely difficult to do (Brambring & Troster, 1994). Second, dropping items causes the reliability and the scope of a test to be reduced due to the shortened test length. Such truncated tests can give only a partial measurement of functioning when compared with the original instrument. For example, the information and vocabulary subscales of WISC correlate with the full-scale WISC-R at 0.71 and 0.74, respectively. That is, only half of the variance in the full-scale test can be accounted for by these two subscale measures.

Other problems may occur when using tests designed for one population with another. For example, care must be taken in assuming that any given item has the same interpretative significance for a person with a visual impairment as it does for sighted peers (Barker et al., 1953). One aspect of this is the need to distinguish a normal reaction to a disability from dysfunction (Ammerman et al., 1986). For example, some items on the MMPI that ostensibly measure hysteria may evoke positive responses in people with a disability, purely on the basis of the physical symptoms associated with their disability (Vash, 1981). Furthermore, many of the tests commonly used with subjects with visual impairment are norm-referenced tests based on sighted populations (Adrian, Miller, & De l'Aune, 1982; Scholl & Schnur, 1976). These tests include the Stanford Binet and WISC measures of intellectual functioning, the Benreuter Personality Inventory and California Personality Inventory, and the Maxfield–Buchholz Social Maturity Scale that is derived from the Vineland Social Maturity Scale (Ammerman et al., 1986; Greenberg & Jordan, 1957; Maxfield & Buchholz, 1957; Petrucci, 1953). The use of a norm-referenced test on a subgroup that it was not standardized on makes test interpretation questionable at best (Anastasi, 1988), and it may easily discriminate against that subgroup (Fuchs et al., 1987).

Some tests have been standardized on visually impaired populations (e.g., the Perkins Binet; the Anxiety Scale for the Blind; the Nottingham Ajustment Scale) (Jan et al., 1977). Yet, even these tests must be used and interpreted with caution because the existing norms do not adequately account for the heterogeneity of the visually impaired population (Ammerman et al., 1986). There is, for example, a notable lack of norms for very young and very old visually impaired people (Hersen et al., 1995), and little reason to believe that a test standardized on a congenitally visually impaired population would have validity for an adventitiously blinded person or vice versa.

9.04.4.2 Measures of Adjustment

Following the loss of sight, a period of adjustment is usually necessary in order for the individuals to prepare themselves mentally for the task ahead (Dodds et al., 1991). Assessments are often conducted to determine if the individual is ready to begin the rehabilitation process and if other treatment or support services may be needed to aid in this process.

Adjustment is a complex process which may be best reflected in multidimensional assessments. The Reactions to Disability and Impairment Inventory (RIDI) is one example of a scale that attempts to measure reaction to the onset of disability in terms of a multidimensional assessment, using eight different psychosocial scales (Antonak & Livneh, 1991). The Nottingham Adjustment Scale (NAS) is another scale designed specifically for use with and standardized on visually impaired persons (Dodds et al., 1996). The NAS consists of seven subsections that measure emotional and cognitive states, including anxiety, depression, self-esteem, acceptance of blindness, attitudes toward blind people, self-efficacy, locus of control, and attributional style. These seven factors are seen as constituting adjustment. As such, a low score on any one of the seven subsections is thought to be indicative of poor adjustment.

9.04.4.3 Measures of Functional Ability

Functional capacity is the product of a great many factors, including: the extent, etiology, and nature of vision loss; personality and cognitive factors; and environmental influences such as support networks and the attitudes of significant others (Van Hasselt, Hersen, & Kazdin, 1985). Self-report and direct observation may be used to gain a richness of information not possible with standard psychometric and clinical testing (Smith, De l'Aune, & Geruschat, 1992). While self-reports may not always be accurate (Wylie, 1961), they can provide good prediction of functioning, and high levels of agreement have been found between self-report and other measures of functional capacity, such as behavioral observation (Elliot & Kuyk, 1994; Szlyk, Arditi, Coffey Bucci, & Laderman, 1990). Behavioral observation tends to have good internal and ecological validity and has proven useful for determining a person's current level of functioning on a variety of tasks and evaluating the overall effectiveness of treatment programs, as well as the relative effectiveness of various treatment options (LaGrow & Murray, 1992; Van Hasselt & Hersen, 1981).

9.04.5 TREATMENT

Adjustment to the onset of a significant visual impairment generally involves both the acquisition of new skills and the restoration of psychosocial health (Elliott & Kuyk, 1994). Adjustment may be characterized by low levels of anxiety, an absence of depression, and high self-esteem (Dodds et al., 1991). The latter requires a high sense of self-efficacy and positive acceptance of one's visual state (Dodds et al.).

9.04.5.1 Skill Acquisition

Rehabilitation program for visually impaired persons are designed to provide the individual with the skills necessary to regain independence in functioning following a significant loss of sight. The re-establishment of functional capacity is a prerequisite for vocational rehabilitation for working age persons. Functional capacity has also been shown to be one of the most important factors involved in satisfaction ratings of a person's quality of life among older severely disabled persons (Osberg, 1987). In fact, restoration of the opportunity to participate in meaningful and enjoyable activities may be the best antidote for the sadness caused by the losses associated with aging and ill health (Zarit, 1992). As such, it is often the goal of rehabilitation for older visually impaired persons, while the restoration of independence, both personal and financial, to the level previously enjoyed is aimed for with younger persons. In order to achieve these goals, individualized instructional programs are provided in the areas of orientation and mobility, adaptive communication skills for both reading and writing, leisure time activities, and the techniques of daily living. The latter includes: eating etiquette, food preparation, grooming, house cleaning and maintenance, money identification and management, personal hygiene, shopping, sewing, telephone use, time-telling, and most other daily needs including those required to manage particular health or medical conditions (e.g., self-monitoring of glucose levels and self-administration of insulin by diabetics) (Huebner, 1986). Vocational rehabilitation may follow if appropriate.

The instructional system used is success based and relies upon both a multisensory approach and a logical sequence for the presentation of the skills and subskills to be mastered (LaGrow, 1992). These programs are designed as much to promote the restoration of a sense of self-efficacy and self-esteem through the development of skill and a history of success, as they are to establish the skills necessary to regain independence in functioning (Elliott & Kuyk, 1994). However, one's psychological state may mediate the amount of success experienced in these programs (Davis et al., 1995). Therefore, specific interventions to promote psychological adjustment may also be needed.

9.04.5.2 Counseling and Support

Psychological intervention may be required to alleviate extreme anxiety, emotional stress, and other depressive states (Dodds et al., 1991). Promotion of the use of positive coping strategies and mature defense mechanisms for dealing with the losses associated with the onset of significant visual impairment is often the focus of such intervention (Fagerstrom, 1992; Zarit, 1992). Both professional counseling and participation in specifically constructed support groups have proven beneficial for these purposes (Teitelbaum et al., 1994). Because "an individual's adjustment to sight loss is closely related to the attitudes, perceptions, and behaviors of his or her family" (Dumas & Sadowsky, 1984, p. 473), counseling services are often recommended for the family as well (Dumas & Sadowsky, 1984; Moore, 1984; Ponchillia, 1984).

The strategies and therapies (e.g., traditional psychotherapy, relaxation training, biofeedback, counseling, hypnosis, systematic desensitization, cognitive therapy for irrational thinking) developed for use with sighted persons can benefit visually impaired persons as they come to to terms with their loss (Needham et al., 1993). These interventions need no particular modification for use with visually impaired persons, yet their focus may vary according to the needs and experiences of the affected individual.

The objectives of therapeutic intervention following traumatic blindness, for example, are to help these individuals understand what has happened to them and to facilitate a sense of self-acceptance and forgiveness for themselves and for those who may be responsible for the event leading to their blindness (Dale, 1992). It is important to help those who are demoralized by the onset of a significant vision impairment verbalize their beliefs (i.e., those related to a sense of helplessness) and identify ways in which their lives can continue to be meaningful and full. Working age persons often require vocational or guidance counseling in addition to personal counseling (Needham et al., 1993). More severely depressed persons may need psychiatric treatment, including the use of antidepressant medications and psychotherapy (Zarit, 1992). Developing a thorough understanding of one's visual condition and an acceptance of the disabilities associated with it appears to be a requisite to adjustment for all persons with a vision impairment (Dale, 1992; Dodds et al., 1991; Teitelbaum et al., 1994; Zarit, 1992).

Although individual counseling is often provided as a means of crisis intervention, ongoing attention is generally provided in the context of a group. Group counseling procedures allow people to resolve their feelings about the loss of sight, vent their emotions to others, and discover that their feelings, attitudes, and experiences may not be completely unique (Schulz, 1980). Group counseling sessions may also provide the basis for the development of peer support.

Peer support may come naturally from interaction with others experiencing similar situations or through formal groups formed to provide support. These groups may prove to be especially beneficial for older persons with visual impairments who often find themselves being overprotected or treated paternalistically by those who do not understand vision impairment and/or aging. Peer support groups tend to balance such counterproductive support and therefore serve to reduce their members" feelings of isolation and the sense that their problems and experiences are unique (Van Zandt et al., 1994). As a result, peer support proves to be beneficial to spouses and other members of the affected person's family as well (Jackson & Lawson, 1995).

9.04.5.3 Delivery Systems

Counselling and support services may be provided by either general or specialist agencies, while the instructional programs provided for adapting to visual loss are generally only available from agencies specifically serving individuals who are blind or visually impaired (Hilton, Ellis, Genwright, Miller, & Stern, 1992). These services are traditionally delivered in the person's place of residence by an itinerant teacher, or at a rehabilitation center from a multidisciplinary team of specialists (Ponchillia & Kaarlela, 1986). Center-based programming may be either day or residential.

Residential rehabilitation programs provide the most therapeutic environment of the options available and have proven to be effective even for those who have had a history of psychological illness. In fact, those illnesses have been seen to become much less evident upon entry to such a program where structure, support and the promise of increased personal adequacy is available (Needham, Acton, & Eldridge, 1992). Residential rehabilitation programs have been shown to result in significant improvement in independent function, self-image, self-esteem (Elliott & Kuyk, 1994), and a significant decrease in irrational thinking about visual impairment (Needham, Ehmer, De l'Aune, & Carr, 1984). One of the greatest advantages of residential rehabilitation over home-based and even day programs is the opportunity for interaction with the other persons in attendance (Hilton et al., 1992).

9.04.5.4 Other Services

Rehabilitation services are not always the most appropriate treatment option for visually impaired persons. Such services assume that the loss of vision is the root cause of the problems the individual is facing. This is certainly not always the case among visually impaired persons. Congenitally blind persons may, in fact, experience no loss in visual function associated with their visual disability. Others, who have experienced the onset of a significant visual impairment later in life but have adapted for it and adjusted to it, may have other problems which require professional counseling or support which are not related directly to their loss in visual function. These problems may best be provided by those who specialize in the problem area (e.g., marriage or family counseling, alcohol and drug counseling, vocational or guidance counseling) rather than those who specialize in the adjustment process *per se*. However, referral and consultation with someone who has experience working with visually impaired persons may be helpful at times.

9.04.5.5 Services for Children

This chapter has concentrated on the needs of adventitiously visually impaired adults. The goal of treatment with this population is to restore the level of functioning to that obtained prior to the loss of sight. This goal and the strategies used to obtain it are not necessarily appropriate to meet the needs of visually impaired children. Their needs are much more general than those of adults who experienced a loss in function. In fact, many of these children experience no loss at all, and indeed have no previous level of function to return to.

Visually impaired children, like all other children, need to grow and develop in order to meet their potential as adults; receive an education which prepares them to participate fully in life; become acculturated in the norms, values, and mores of the society in which they live; and be prepared for their future as a productive member of that society. These children will need some special attention to aid them in meeting these goals (LaGrow, 1992). This assistance is generally available through the provision of special education and related services and is prescribed in an individualized education or development plan (Ward, 1986).

9.04.6 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

The number of elderly persons in the population is expected to increase dramatically in the future (e.g., in the USA the number of persons over 65 is expected to double from 1990 to 2030). As a result, both the number and proportion of persons in the population experiencing a significant visual impairment will also increase (Orr, 1992). Thus, we will need to know more about the effects of the onset of visual impairment on the elderly population and the most effective means of treatment. This research needs to be more theoretically driven than it is at present. Therefore, testable models of adjustment to adventitious blindness need to be developed that provide an explanation for the effects of the onset of a disability, a rationale for intervention, and predictions to guide future research. The loss model in use in the 1990s has been criticized as being merely descriptive and defies testing (Dodds et al., 1991).

New models of services delivery need to be developed and tested as well. The majority of older persons receive home-based rehabilitation programs without the benefit of comprehensive counseling services or ready access to formal or informal peer groups (Hilton et al., 1992). Yet, it appears that this population is most at risk of depression, devaluation, and demoralization from the onset of a visual impairment and the most likely to benefit from both counseling and peer support. Thus, the best way to provide services to this population and alternative ways to provide peer support need to be investigated. More specific assessment devices need be developed for this population which are standardized on this population (Hersen et al., 1995). Finally, a clear distinction between congenital and adventitious, and early and late onset, need to be made in future research.

More needs to be known about the effect of adventitious onset of visual impairment on children and the most effective means of treatment (Dale, 1992). Clear distinctions need to be made between rehabilitation and habilitation in this population, and the effects of each.

9.04.7 CONCLUSIONS

The psychological and emotional distress accompanying severe visual impairment and blindness has been well documented. This distress appears to be primarily a reaction to the loss of sight and related losses in functional adequacy and social role (Hudson, 1994). The more suddenly this occurs and the more severe the impairment, the more likely it is that the person will have significant problems with adjustment (Dale, 1992; Dodds et al., 1991).

These problems often manifest themselves in terms of anxiety, devaluation, demoralization, and depression (Schulz, 1980). Older persons and those in ill health tend to find it harder to adjust to the changes associated with the onset of a significant visual impairment than their younger and healthier peers (Bernbaum et al., 1988; Hersen et al., 1995: Robbins & McMurray, 1988; Teitelbaum et al., 1994; Zarit, 1992).

Rehabilitation services are designed to help with the adjustment process by restoring a person's functional ability through instruction and the use of adaptive equipment. This, in turn, should lead to a restoration of role, self-esteem, and security (LaGrow, 1992). Counseling and other psychological interventions may be most useful in helping this process unfold (Dodds et al., 1991). Peer support, either formal or informal, appears to be a critical factor for success in rehabilitation, especially for those who are older (Teitelbaum et al., 1994).

Rehabilitation may be provided either at home by one or more itinerant instructors or in a center by a multidisciplinary team of specialists. Center-based programs may be either day or residential. Residential programs appear to offer the most therapeutic environment, the most intensive instructional programing, and the greatest opportunity for peer support. The latter may, in fact, prove to be the greatest single advantage of this particular service delivery system (Hilton et al., 1992). The therapeutic and treatment programs available in residential rehabilitation tend to promote both the learning and overall adjustment of people with various degrees of psychological health, despite the anxiety involved in living away from home and learning new skills (Needham et al., 1992).

Unfortunately, such programs are more commonly provided for younger, working-age people than for the older population. These people are usually served by home-based programs (Hilton et al., 1992). As a result, they often have limited access to specialized counseling services and little opportunity for peer support, as well as less intensive instructional programs than those attending residential rehabilitation programs (Ponchillia & Kaarlela, 1986). It is assumed, however, that this population, who tend to have greater difficulty in their adjustment to the onset of a visual impairment than their younger and healthier peers, would need more of all of these things. Thus, if these services are not provided through specialized blindness services, they may have to be gained through other providers in the community.

Services for visually impaired children are primarily provided through the schools. Most of these children are congenitally visually impaired and tend to be about as well adjusted as their sighted peers (Beaty, 1994). Special services for adjustment to blindness are generally only needed if they experience the onset of a visual impairment and thus experience a loss in function.

9.04.8 REFERENCES

Adrian, R. J., Miller, L. R., De l'Aune, W. R. (1982). Personality assessment of early visually impaired persons using the CPI and MMPI. *Journal of Visual Impairment and Blindness, 76,* 172–178.

Alexander, F. E. (1996). Self-concepts of children with visual impairments. *RE: View, 28*(1), 35–43.

Ammerman, R. T., Van Hasselt, V. B., & Hersen, M. (1986). Psychological adjustment of visually handicapped children and youth. *Clinical Psychology Review, 6,* 67–85.

Ammerman, R. T., Van Hasselt, V. B., Hersen, M., & Moore, L. E. (1989). Assessment of social skills in visually impaired adolescents and their parents. *Behavioral Assessment, 11,* 327–351.

Anastasi, A. (1988). *Psychological testing* (6th. ed.). New York: Macmillan.

Antonak, R. F., & Livneh, H. (1991). A hierarchy of reactions to disability. *International Journal of Rehabilitation Research, 14,* 13–24.

Barker, R. G., Wright, B. A., Meyerson, L., & Gonick, M. R. (1953). *Adjustment to physical handicap and illness: A survey of the social psychology of physique and disability*. New York: Social Science Research Council.

Baumeister, A. A., & Forehand, R. (1973). Stereotyped acts. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 6). New York: Academic Press.

Beaty, L. A. (1991). The effects of visual impairment on adolescents" self-concept. *Journal of Visual Impairment and Blindness, 85,* 129–132.

Beaty, L. A. (1994). Psychological factors and academic success of visually impaired college students. *RE: View, 26,* 131–139.

Berkow, R. (1982). *The Merck Manual: Diagnosis and Therapy*. Rahway, NJ: Merck & Co, Ltd.

Bernbaum, M., Albert, S. G., & Duckro, P. N. (1988). Psycho-social profiles in patients with visual impairment due to diabetic retinopathy. *Diabetes Care, 11,* 551–557.

Bernbaum, M., Albert, S. G., & Duckro, P. N. (1993). Personal and family stress in individuals with diabetes and vision loss. *Journal of Clinical Psychology, 49,* 670–677.

Brambring, M., & Troster, H. (1994). The assessment of cognitive development in blind infants and preschoolers. *Journal of Visual Impairment and Blindness, 88,* 9–18.

Branch, L. G., Horowitz, A., & Carr, C. (1989). The implications for everyday life of self-reported visual decline among people over age 65 living in the community. *The Gerontologist, 29,* 359–365.

Carroll, T. J. (1961). *Blindness: What it is, what it does, and how to live with it*. Boston: Little, Brown & Co.

Colenbrander, A. (1976). Low vision: Definition and classification. In E. E. Faye (Ed.), *Clinical low vision* (pp. 3–6). Boston: Little, Brown & Co.

Dale, B. (1992). Issues in traumatic blindness. *Journal of Visual Impairment and Blindness, 86,* 140–143.

Davis, C., Lovie-Kitchen, J., & Thompson, B. (1995). Psychosocial adjustment to age-related macular degeneration. *Journal of Visual Impairment and Blindness, 89,* 16–27.

Delafield, G. L. (1976). Adjustment to blindness. *New Outlook for the Blind, 70,* 64–68.

Dodds, A. G., Bailey, P., Pearson, A., & Yates, L. (1991). Psychological factors in acquired visual impairment: The development of a scale of adjustment. *Journal of Visual Impairment and Blindness, 85,* 306–310.

Dodds, A. G., Craig, D., & Flannigan, H. (1996). The Nottingham Adjustment Scale: An estimation of response bias. *Journal of Visual Impairment and Blindness, 90,* 51–57.

Drury, T. F., Danchik, K. M., & Harris, M. I. (1985). Socio-demographic characteristics of adult diabetes. In *Diabetes in America*. National Diabetes Data Group, NIH Publication No. 85–1468 VII, 15–17. Washington, DC: Department of Health and Human Services.

Dumas, A., & Sadowsky, A. (1984). A family training program for adventitiously blinded and low vision veterans. *Journal of Visual Impairment and Blindness, 78,* 473–478.

Elliott, J. L., & Kuyk, T. K. (1994). Self-reported functional and psychosocial outcomes of blind rehabilitation. *Journal of Visual Impairment and Blindness, 88,* 206–212.

Fagerstrom, R. (1992). Defense mechanisms of elderly persons before and after a cataract operation. *Journal of Visual Impairment and Blindness, 86,* 361–363.

Faye, E. E. (1976). *Clinical low vision*. Boston: Little, Brown & Co.

Ferrell, K. (1986). Infancy and early childhood. In G. T. Scholl (Ed.), *Foundations of education for blind and visually handicapped children and youth: Theory and practice* (pp. 119–136). New York: American Foundation for the Blind.

Fuchs, D., Fuchs, L. S., Benowitz, S., & Barringer, K. (1987). Norm-referenced tests: Are they valid for use with handicapped students? *Exceptional Children, 54,* 263–271.

Greenberg, H. M., & Jordan, S. (1957). Differential effects of total blindness and partial sight. *Exceptional Children, 24,* 123–124.

Harrell, R. L., & Strauss, F. A. (1986). Approaches to increasing assertive behavior and communication skills in blind and visually impaired persons. *Journal of Visual Impairment and Blindness, 80,* 794–798.

Head, D. N. (1980). The stability of self-concept scores in visually impaired adolescents. *Education of the Visually Handicapped, 12,* 66–74.

Heinemann, A., Bulka, M., & Smetak, S. (1988). Attribution and disability acceptance following traumatic injury. *Rehabilitation Psychology, 33,* 195–206.

Hersen, M., Kabacoff, R. I., Van Hasselt, V. B., Null, J. A., Ryan, C. F., Melton, M. A., & Segal, D. L. (1995). Assertiveness, depression and social support in older visually impaired adults. *Journal of Visual Impairment and Blindness, 89,* 524–530.

Hilton, J., Ellis, F. J., Genwright, R., Miller, L. R., & Stern, H. (1992). The state of the blindness system today: Report of system of service delivery work group B. *Journal of Visual Impairment and Blindness, 86,* 323–330.

Hudson, D. (1994). Causes of emotional and psychological reactions to adventitious blindness. *Journal of Visual Impairment and Blindness, 88,* 498–503.

Huebner, K. (1986). Social skills. In G. T. Scholl (Ed.), *Foundations of education for blind and visually handicapped children and youth: Theory and practice* (pp. 341–362). New York: American Foundation for the Blind.

Jackson, R., & Lawson, G. (1995). Family environment and psychological distress in persons who are visually impaired. *Journal of Visual Impairment & Blindness, 89,* 157–160.

Jan, J. E., Freeman, R. D., & Scott, E. P. (1977). *Visual impairment in children and adolescents*. New York: Grune & Stratton.

Johnson, D. G. (1989). An unassisted method of psychological testing of visually impaired individuals. *Journal of Visual Impairment and Blindness, 83,* 114–118.

Jose, R. (1984). *Understanding low vision*. New York: American Foundation for the Blind.

Kirchner, C. (1985). *Data on blindness and visual impairment in the US: A resource manual on characteristics, education employment, and service delivery*. New York: American Foundation for the Blind.

Kirchner, C., & Peterson, R. (1988). Data on visual disability from NCHS, 1977. In C. Kirchner (Ed.), *Data on blindness and visual impairment in the US* (2nd ed., pp. 19–24). New York: American Foundation for the Blind.

LaGrow, S. (1992). *The rehabilitation of visually impaired people*. Auckland, New Zealand: Royal New Zealand Foundation for the Blind.

LaGrow, S., & Murray, S. (1992). Use of alternating treatment design to evaluate intervention in low vision rehabilitation. *Journal of Visual Impairment and Blindness, 86,* 435–439.

LaGrow, S., & Repp, A. (1984). Stereotypic responding. A review of intervention research. *American Journal of Mental Deficiency, 88,* 595–609.

Malec, D. (1985). Personality factors associated with severe traumatic disability. *Rehabilitation Psychology, 30,* 165–172.

Maxfield, K. E., & Buchholz, S. (1957). *A social maturity scale for blind preschool children: A guide to its use*. New York: American Foundation of the Blind.

Miner, D. (1991). Using nonaversive techniques to reduce self-stimulatory hand-mouthing in a visually impaired and severely retarded student. *RE: View, 22,* 185–194.

Moore, J. (1984). Impact of family attitudes toward blindness/visual impairment on the rehabilitation process. *Journal of Visual Impairment and Blindness, 78,* 100–105.

National Society for the Prevention of Blindness. (1988). *Vision loss in the United States*. Washington, DC: Author.

Needham, W. E., Acton, J. J., & Eldridge, L. S. (1992). Psychological disorders of blind persons and success in residential rehabilitation. *Journal of Visual Impairment and Blindness, 86,* 144–147.

Needham, W. E., Ehmer, M. N., De l'Aune, W. R., & Carr, R. B. (1984, August). *Irrational thinking, adjustment and change in the blind during rehabilitation*. Paper presented at the annual meeting of the American Psychological Association, Toronto, Canada.

Needham, W. E., & Eldridge, L. S. (1990). Performance of blind vocational rehabilitation clients on the Minnesota Rate of Manipulation Tests. *Journal of Visual Impairment and Blindness, 84,* 182–185.

Needham, W. E., Eldridge, L. S., Harabedian, B., & Crawford, D. G. (1993). Blindness, diabetes, amputations: Alleviation of depression and pain through thermal biofeedback therapy. *Journal of Visual Impairment and Blindness, 87,* 368–371.

Orr, A. L. (1992). Aging and blindness: Toward a systems approach to service delivery. In A. L. Orr (Ed.), *Vision and aging: Crossroads for service delivery* (pp. 3–31). New York: American Foundation for the Blind.

Osberg, J. S. (1987). Life satisfaction and quality of life among disabled elderly adults. *Journal of Gerontology, 42,* 228–230.

Petrucci, D. (1953). The blind child and his outlook. *New Outlook for the Blind, 47,* 240–246.

Ponchillia, P. (1984). Family services: Role of the center-based teaching professional. *Journal of Visual Impairment and Blindness, 78,* 97–100.

Ponchillia, P., & Kaarlela, R. (1986). Post-rehabilitative use of adaptive skills. *Journal of Visual Impairment and Blindness, 80,* 665–669.

Robbins, H. G., & McMurray, N. (1988). Psychosocial and visual factors in low vision rehabilitation of patients with age-related maculopathy. *Journal of Vision Rehabilitation, 2,* 11–21.

Rohe, D., & Basford, B. (1989). Traumatic spinal cord injury, alcohol and the MMPL. *Rehabilitation Psychology, 34,* 25–32.

Rosenbloom, A. A. (1992). Physiological and functional aspects of aging, vision, and visual impairment. In A. L. Orr (Ed.), *Vision and aging: Crossroads for service delivery* (pp. 47–68). New York: American Foundation for the Blind.

Scholl, G. T. (1986). Growth and development. In G. T. Scholl (Ed.), *Foundation of education for blind and visually handicapped children and youth: Theory and practice* (pp. 65–81). New York: American Foundation for the Blind.

Scholl, G., & Schnur, R. (1976). *Measures of psychological, vocational and educational functioning in the visually handicapped*. New York: American Foundation for the Blind.

Schulz, P. J. (1980). *How does it feel to be blind: The psychodynamics of visual impairment*. Los Angeles: Muse-Ed.

Smith, A. J., De l'Aune, W., & Geruschat, D. R. (1992). Low vision and mobility problems: Perceptions of O&M specialists and persons with low vision. *Journal of Visual Impairment and Blindness, 86,* 58–62.

Swallow, R. (1981). Fifty assessment instruments commonly used with blind and partially seeing individuals. *Journal of Visual Impairment and Blindness, 75,* 65–72.

Szlyk, J., Arditi, A., Coffey Bucci, P., & Laderman, D. (1990). Self-report in functional assessment of low vision. *Journal of Visual Impairment and Blindness, 84,* 61–66.

Teitelbaum, L. M., Davidson, P. W., Gravetter, F. J., Taub, H. A., & Teitelbaum, C. S. (1994). The relation of vision loss to depression in older veterans. *Journal of Visual Impairment and Blindness, 88,* 253–257.

Tuttle, D. W. (1984). *Self-esteem and adjusting with blindness*. Springfield, IL: Charles C. Thomas.

Van Hasselt, V. B. (1983). Visual impairment. In M. Hersen, V. B. Van Hasselt, & J. L. Matson (Eds.), *Behavior therapy for the developmentally and physically disabled* (pp. 109–128). New York: Academic Press.

Van Hasselt, V. B., & Hersen, M. (1981). Applications of single-case designs to research with visually impaired individuals. *Journal of Visual Impairment and Blindness, 75,* 359–362.

Van Hasselt, V. B., Hersen, M., & Kazdin, A. E. (1985). Assessment of social skills in visually handicapped adolescents. *Behavior Research and Therapy, 25,* 53–63.

Van Hasselt, V. B., Hersen, M., Kazdin, A. E., Simon, J., & Mastantuono, A. K. (1985). A behavioral-analytic model for assessing social skills in blind adolescents. *Behavior Research and Therapy, 23,* 395–405.

Van Zandt, P. L., Van Zandt, S. L., & Wang, A. (1994). The role of support groups in adjusting to visual impairment in old age. *Journal of Visual Impairment and Blindness, 88,* 244–252.

Vander Kolk, C. (1982). Rehabilitation and visual impairment: A human system. *Journal of Visual Impairment and Blindness, 76,* 346–350.

Vash, C. L. (1981). *The psychology of disability*. New York: Springer.

Vaughan, C. E., & Hobson, S. (1990). Reducing late-life dependence resulting from declining visual acuity. *Journal of Visual Impairment and Blindness, 84,* 370–372.

Ward, M. E. (1986). Planning the individualized education program. In G. T. Scholl (Ed.), *Foundations of education for blind and visually handicapped children and youth: Theory and practice* (pp. 215–238). New York: American Foundation for the Blind.

Warren, D. (1984). *Blindness and early childhood development* (2nd ed.). New York: American Foundation for the Blind.

Wylie, R. C. (1961). *The self-concept*. Lincoln, NE: University of Nebraska Press.

Yeadon, A. (1984). The informal care group: Problem or potential. *Journal of Visual Impairment and Blindness, 78,* 149–154.

Zarit, S. H. (1992). Psychosocial aspects of aging and visual impairment. In A. L. Orr (Ed.), *Vision and aging: Crossroads for service delivery* (pp. 93–117). New York: American Foundation for the Blind.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.05

Hearing Disabilities

LAURENCE McKENNA
Guy's Hospital and the Royal National Throat, Nose and Ear Hospital, London, UK
and
GERHARD ANDERSSON
Uppsala University, Sweden

9.05.1 INTRODUCTION 69
9.05.2 PREVALENCE 70
9.05.3 NATURE 71
  9.05.3.1 Detection 71
  9.05.3.2 Problems Experienced 71
  9.05.3.3 Psychological Factors and Theoretical Views 73
9.05.4 ASSESSMENT 74
  9.05.4.1 Audiological Assessments 74
  9.05.4.2 Self-report Measures 74
  9.05.4.3 Psychological Assessments 75
9.05.5 TREATMENT 75
  9.05.5.1 Hearing Aids and Other Prosthetic Devices 75
  9.05.5.2 Counseling and Self-help Approaches 76
  9.05.5.3 Psychological Treatments for Hearing Loss and Tinnitus 77
9.05.6 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 78
9.05.7 SUMMARY 79
9.05.8 REFERENCES 80

9.05.1 INTRODUCTION

The application of clinical psychology within the fields of audiological medicine and its associated science otolaryngology is comparatively new. Most of the clinical psychology and research efforts have taken place since the early 1980s. This chapter focuses on psychological issues relating to hearing loss and tinnitus, both of which are disorders of the ear.

Sensory restrictions are an almost universal part of human existence and are particularly prevalent among elderly people. All people experience problems with listening (Noble, 1983), especially in demanding auditory situations, and many people have experienced tinnitus after being exposed to loud sound. To review the physiological underpinnings of hearing loss and tinnitus is beyond the scope of this chapter and the reader is referred to textbooks on audiology (Katz, 1994). Here it suffices to say that hearing disorders can occur as a result of damage to or interference with any part of the auditory pathway from the outer ear to the brain.

Hearing disorders may arise from a wide range of etiologies. A basic distinction may be made between hearing loss that is present at birth or prelingually and that which is acquired postlingually. This chapter is concerned with only the latter type. Distinctions can also be made between conductive hearing loss and sensori-neural hearing loss. When the structures of the outer or middle ear are dysfunctional, the hearing loss is described as conductive in nature. Sensori-neural hearing loss arises as a result of cochlear (inner ear) dysfunction. The majority of hearing loss of 25 dB hearing level (HL) or greater is sensori-neural in nature. The causes of acquired hearing loss range from common influenza type infections to cerebral tumors and to iatrogenic causes.

Most commonly, hearing loss is attributed to exposure to loud noises, generally over a period of years, or to the process of aging (Katz, 1994). The term presbyacusis was coined to describe the almost ubiquitous presence of hearing loss in elderly people (Katz, 1994). It is most likely that hearing loss in elderly people involves changes in central auditory processes as well as the loss of sensory hair cells in the cochlea (Katz, 1994). Central presbyacusis is the component related to the aging central auditory nervous system. It is used to describe changes in speech understanding, especially in noise, that many older people experience.

Although sometimes observed, functional or nonorganic hearing difficulties should not be classified as hearing loss, but rather should be regarded a sign of psychological disturbance or abuse (Reidner & Efros, 1995). Malingering and faked hearing loss may also be found in the clinic, especially when monetary compensation is involved, but these are easily detected during testing.

Tinnitus is the perception of noise in the absence of any external stimulation to produce that noise (Hazell & Jastreboff, 1990). It is not a disease but a symptom that can be produced by an enormous variety of change in the auditory pathway. Tinnitus has for long been regarded as the result of cochlear dysfunction. Since the early 1990s, however, an emphasis has been placed on the central processing of cochlear pathology in the perception of tinnitus(Jastreboff & Hazell, 1993). Tinnitus is almost always a phantom auditory perception, perceived only by the patient. It is widely experienced in the general population and more particularly by people with hearing loss, although hearing loss and tinnitus can occur independently of one another. For many people seen in audiology clinics, tinnitus is the primary complaint.

Again, like hearing loss, the symptom can have many and diverse etiologies. In a clinical setting, the causes of many cases of tinnitus may go undiscovered. In seeking to account for tinnitus it is common clinical practice to tell the patient that "whatever caused your hearing loss also caused your tinnitus." Tinnitus can be distressing, and for the majority of people there are no practical medical or surgical solutions to the symptom (Stouffer, Tyler, Kileny, & Dalzell, 1991).

In some rare cases "objective tinnitus" has been documented where an outside observer can hear the tinnitus (Norton, Schmidt, & Stover, 1990). In such cases, the noises are somato sounds resulting from factors such as the patient's pulse or muscle twitches. Technically such noises should be not be included in the same classification as subjective tinnitus. The idea that temporomandibular dysfunction can be a cause of tinnitus has been debated and the evidence is not overwhelming (Chan & Reade, 1994).

Popular images of people with hearing impairment and of those with tinnitus are not flattering. Hearing disorders are associated with old age and as a consequence with frailty of body and mind. In word association the word "deaf" is very likely to be followed by "dumb" and to have the connotation of stupidity. It has been suggested that most physical disorders cut people off from "things," but hearing loss cuts the sufferer off from *people*. There are not only communication problems to overcome but also prejudice. For example, hearing-impaired people are commonly regarded as suffering from psychological disturbances such as depression and paranoia. A "hearing aid effect" has also been observed, that is, hearing aid users are rated less favorably than controls without a hearing aid (Danhauer, Johnson, Kasten, & Brimacombe, 1985). The popular view of the tinnitus patient is of a neurotic, complaining individual whom many physicians and surgeons would wish to avoid. Again, in the popular mind there is a link between tinnitus and psychiatric disorder including suicidal behavior. This is reflected in patients' views of tinnitus, and, although cases of tinnitus-related suicide have been reported, it is not associated with a greater increased risk than say unemployment (Lewis, Stephens, & McKenna, 1994).

9.05.2 PREVALENCE

Hearing loss and tinnitus affect large numbers of people. In the UK the National Study of Hearing conducted by the MRC Institute of Hearing Research (Davis, 1989) estimated that 16.1% of the adult population (aged 18–80 years) report a hearing disorder of 25 dB HL or greater. This means that there are about 7.7

million people in the UK with a hearing loss. As the severity of hearing impairment increases, the prevalence decreases. The prevalence of very profound hearing loss (95 dB HL or greater) is 0.2% of the population. The prevalence of hearing impairment at all levels of severity is very dependent on age. Below the age of 45 years even mild hearing loss is relatively rare (< 5% of the population), but close to 50% of people aged 70–74 years have a hearing loss of 25 dB HL or more.

Some 30% of the population experience some degree of tinnitus, although usually very minor in nature. About 10% of the population report persistent spontaneous tinnitus, that is, not just after exposure to noise, and nearly 7% of the population have consulted a doctor about tinnitus (Davis, 1993). Tinnitus causes severe annoyance in 1% of the population and brings about a severely reduced ability to live a normal life in 0.5%. Davis (1993) noted that tinnitus correlates highly with the individual's hearing impairment and that the best predictor of tinnitus is the severity of hearing impairment. Axelsson and Ringdahl (1989) studied 3600 randomly selected adults in Gothenburg, Sweden and reported that 14.2% of them suffered from tinnitus "often" or "always" and that 2.4% described tinnitus as "plaguing me all day."

9.05.3 NATURE

9.05.3.1 Detection

There are differences between tinnitus and hearing impairment regarding how and when they are noticed. Hearing loss may be more noticeable in noisy surroundings, whereas tinnitus may be more intrusive in a quiet environment. Tinnitus is, however, a complex phenomenon; in some cases it is aggravated by noise, sometimes the task in hand may distract a person from the tinnitus, while at other times tinnitus may be the distracter. There is no inevitability about when it will be detected.

Both tinnitus and hearing loss can come on gradually and may also be exacerbated at times of high stress such as when changes in lifestyle occur, for example, retirement. It is also possible for both symptoms to go unnoticed or at least unreported. It is well recognized that many older people never seek assistance for their hearing impairment. According to Brooks (1989), older people tolerate a greater degree of impairment before taking action than younger people, and they might also think that a hearing aid alone may be of little use. Commonly, it is a close relative that encourages the person to seek audiological help. Reasons for seeking help have been investigated. Swan and Gatehouse (1990) compared persons seeking help for hearing loss with those who did not. The results showed that help-seekers had more asymmetrical hearing, greater disability, and were more handicapped.

There is little work specifically investigating the process of tinnitus detection and the reasons for seeking help. Attias et al. (1995) did report that help-seekers displayed more psychiatric symptoms than nonhelp-seekers. Nonetheless, the tinnitus patients who did not seek help had more psychological problems than a control group without tinnitus. Curiously, the perceived loudness of tinnitus was significantly lower in the help-seeking group. Differences between tinnitus complainers and noncomplainers have, however, been found in terms of psychological differences (Hallberg & Erlandsson, 1993; Kirsch, Blanchard, & Parnes, 1989). To our knowledge, there are no prospective studies of the onset of tinnitus, although retrospective estimates of psychological changes exist (Stouffer & Tyler, 1992).

9.05.3.2 Problems Experienced

Hearing loss and tinnitus are associated with different kinds of problems. Tinnitus is more of a private, and hearing loss more of a communicative handicap. Since the comorbidity between the two is high (Lindberg, Lyttkens, Melin, & Scott, 1984), many patients are likely to experience both sets of problems.

Negative consequences of hearing loss as described by patients and in the literature include communication problems, noise sensitivity, technical problems with the hearing aid or assistive device, occupational difficulties, reduced performance on cognitive tests, and poor uptake of rehabilitation services. Moreover, there is an extensive literature on the negative consequences of hearing loss (e.g., Jakes, 1988; McKenna, 1993). Emotional problems such as depression, neurotic symptoms, paranoia, social stress, insecurity, loneliness, social isolation, and cognitive deficits have all been reported.

Attempts have been made to record the psychological impact of simulated hearing loss (Zimbardo, Andersen, & Kabat, 1981). While some negative emotional consequences have been described, the results are inconclusive; these studies have been restricted by the difficulty of achieving meaningful levels of hearing loss and by their temporary and artificial nature. Improvements in psychological functioning after partial restoration of hearing have been documented (Mulrow et al., 1990); more recently this has been studied after cochlear implantation (McKenna, 1991).

In terms of emotional consequences, the empirical picture is mixed with some studies showing an association between hearing impairment and psychiatric complications (Singerman, Riedner, & Folstein, 1980), whereas others have found less or conflicting support for these findings (Eastwood, Corbin, Reed, Nobbs, & Kedward, 1985). When other health problems are controlled, the correlations between hearing loss and anxiety and depression are found to be weaker (Jones, Victor, & Vetter, 1984). Andersson, Melin, Lindberg, and Scott 1995b), in studying older people with hearing impairment, found that health factors, hearing problems, and psychological factors clustered together, again pointing to the importance of other health problems. A lack of a clear relationship between the extent of hearing loss and the extent of psychological disturbance has thus been a frequent observation. However, Thomas and Gilholme-Herbst (1980) did identify psychological disturbance in a greater number of a subsample of subjects with more severe hearing loss and poor speech discrimination.

Obviously, for some patients hearing impairment is accompanied by psychological problems (McKenna, Hallam, & Hinchcliffe, 1991). There is only a small number of studies (Clark, Hirsch, Smith, Furman, & Jacob, 1994) that have examined the prevalence of emotional disturbances in the hearing-impaired population using *DSM-IV* (American Psychological Association, 1994) criteria and structured clinical interviews. McKenna (1993) reported that anxiety problems were common in his clinical practice. This observation is in line with a study by Eriksson-Mangold and Carlsson (1991) and with a study by Andersson and Green (1995) that showed a correlation between self-reported hearing handicap and anxiety. In other studies, the subjective experience of hearing impairment has been linked with depression (Knutson & Lansing, 1990), but the link between audiological measures and depression is less clear (Gilhome-Herbst & Humphrey, 1980).

The cognitive status of elderly hearing-impaired people has been a focus for study. Ohta, Carlin, and Harmon (1981) concluded that cognitive abilities of elderly persons can be underestimated because of their reduced sensory acuity. Hearing impairment has been associated with dementia (Uhlmann, Larson, Rees, Koepsell, & Duckert, 1989), but conflicting results have been reported (Gilhome-Herbst & Humphrey, 1980).

The social consequences of hearing impairment have been investigated. Meadow-Orlans (1985) cited research showing that elderly hearing-impaired people are less willing to guess, less flexible, and more cautious in social settings than younger hearing-impaired individuals. Weinstein and Ventry (1982) reported significant correlations between subjective and objective measures of isolation and hearing loss in a group of elderly people. In a population survey, Stephens, Lewis, Charny, Farrow, and Francis (1990) found that the most common problems associated with hearing impairment was trouble hearing television, radio, and general conversation. Other studies have also documented adverse effects on social interaction and communication (Tyler, Baker, & Armstrong-Bednall, 1983). The role of the spouse in the adjustment to hearing impairment has been investigated and has been found to be important in the sense that the spouse influences acceptance of the loss (Hallberg & Barrenäs, 1993).

There are numerous descriptions of the negative consequences of tinnitus (Hallam, Rachman, & Hinchcliffe, 1984; Tyler & Baker, 1983). Among the consequences described by patients and in the literature are sleep disturbances, concentration difficulties, muscular tension, noise sensitivity, and emotional consequences (Hallam, 1989; Hallam, Jakes, & Hinchcliffe, 1988). It is important to note that there are large individual differences in the extent to which patients experience these problems. The link between tinnitus and emotional distress has been investigated in several studies (Collet, Moussu, Disant, Ahami, & Morgon, 1990; Halford & Anderson, 1991). Most studies on the emotional consequences of tinnitus have been conducted on highly selected samples of patients with severe tinnitus distress (Briner, Risey, Guth, & Noris, 1990). Some studies have found relatively low (but significant) correlations between tinnitus distress and psychological complaints (Halford & Anderson, 1991; Hiller, Goebel, & Rief, 1994). McKenna et al. (1991), however, reported that 45% of those whose main complaint was tinnitus showed signs of significant psychological disturbance. Simpson, Nedzelski, Barber, and Thomas, (1988) found that 63% of tinnitus sufferers could be classified as psychiatrically disturbed and 46% had mood disorder as assessed by the structured interview for the *DSM-III-R* (SCID). Russo, Katon, Sullivan, Clark, and Buchwald (1994) used the *DSM-IV* to estimate the prevalence of psychiatric disorder among tinnitus patients. A number of studies have used the Beck depression inventory (BDI); (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) in the assessment of tinnitus patients. Kirsch et al. (1989) reported that the mean BDI score for a group of tinnitus patients was within the normal range. Wilson, Henry, Bowen, and Haralambous (1991) reported mean scores within the range of mild mood disturbance. In a recent review by

Hinchcliffe and King (1992), depression was found to be the principal distinguishing feature between tinnitus complainers and noncomplainers. As with hearing impairment, psychoacoustic measures (e.g., matching of tinnitus loudness) have not been found to be good predictors of tinnitus discomfort (Hinchcliffe & King, 1992).

Compared with the literature on chronic pain (Weiss & Kerns, 1995), little work has been done on the interpersonal aspects of tinnitus. One exception is a study by Sullivan, Katon, Russo, Dobie, and Sakai (1994) where role dysfunction in marital interaction was found to be associated with appraisal of tinnitus as salient. Punishing responses to illness behavior were associated with tinnitus-related dysfunction.

9.05.3.3 Psychological Factors and Theoretical Views

Psychological factors are involved in the illness behavior of hearing-impaired people and may disrupt the communication strategies such as lip-reading used by hearing-impaired people. A link between personality and vulnerability to noise-induced hearing loss has been postulated (Jakes, 1988). Although links between Type A personality and noise-induced peripheral vasoconstriction, and between temporary threshold shift and vasoconstriction, have been demonstrated, no clear link has been established between personality and causes of hearing loss. It seems unlikely that there is a psychological cause for most hearing loss.

A number of studies have examined the role of personality in coping with hearing impairment (Andersson, 1995). Stephens (1980) used the Eysenck personality inventory and found elevated levels of neuroticism and introversion. Coren and Harland (1995) found that reduced hearing acuity was associated with increased neuroticism.

The personality of the tinnitus patient has also been the subject of some interest (Gerber, Nehemkis, Charter, & Jones, 1985). For example, Collet et al. (1990) used the Minnesota multiphasic personality inventory (MMPI) and found normal scores overall, with the exception of elevated depression scores in men.

Dispositional optimism, as assessed by the life orientation scale (LOT); (Scheier & Carver, 1985), has been found to be positively related to coping with hearing impairment (Scott, Lindberg, Melin, & Lyttkens, 1994), and negatively related with tinnitus complaints (Andersson, 1996).

Personal control was found to be an important aspect of hearing-impaired peoples' communication success in a study by Scott et al. (1994). It is also an important predictor of tinnitus discomfort and adaptation (Scott, Lindberg, Melin, & Lyttkens, 1990). Budd and Pugh (1995) found significant relationships between locus of control, tinnitus severity, and emotional distress in tinnitus sufferers.

Kyle, Jones, and Wood (1985) described the notion that individuals usually attempt to control the access features of the information they receive by personal and social adjustment and that hearing loss disturbs the control the individual exerts. According to Kyle et al. (1985), there are at least three solutions available: (i) increase the level of control at all costs, (ii) accept or expect a reduced level of control and flow of information, and (iii) reject or avoid situations in which the level of control is threatened.

A different view of hearing originates from the work of Gibson (1986) who emphasized that our perceptions are rich and elaborate because the stimuli in our environment are rich with information rather than because our thought processes or experiences provide the richness. This is an ecological model. In the field of audiology, Noble (1983) has presented an ecological model that views hearing in relation to the audible features and characteristics of the real, day-to-day world. Noble especially pointed to the active role of the perceiver. In other words, hearing is more than just perceiving sounds, it also has to do with how a person acts in an environment. Noble and Hétu (1994) further described an ecological approach to hearing impairment which takes special notice of the interactions between people, environments, and their interfaces. The ecological approach has been important in diverting attention away from technological and towards psychological and environmental factors involved in rehabilitation. To our knowledge no research has yet dealt with an ecological approach to tinnitus.

The ideas of Skinner (1957) allow another view of hearing. One appealing aspect of Skinner's work was a focus on communication. According to this view, hearing could be viewed as an operant, that is, as a behavior classified on the basis of its effects and also being under the influence of contingencies of reinforcement. During his later years, Skinner extended his theories into the field of aging in the book *Enjoy old age* (Skinner & Vaughan, 1983), with some behavioral advice on how to handle hearing loss. Skinner advocates an assertive approach while still acknowledging when it is impossible to hear: "You do best to stop trying to hear things when you are having trouble. You are probably not enjoying what is said in a television program if you are straining to hear it" (Skinner & Vaughan, 1983, p. 44).

A related perspective is the functional approach. Applied to hearing impairment and communication difficulties, this can be viewed as an interaction between the capabilities of the individual and the particular task demand at hand in the communication situation (Andersson & Melin, 1993). Functional analysis deals with the identification of important, causal, and ideally controllable variables that are applicable to a specified set of target behaviors for the patient in question (Haynes & O'Brien, 1990). Emphasis is placed on viewing each patient as unique and finding strengths as well as needs.

In the field of tinnitus theory, recent neurophysiological approaches hold that tinnitus is better conceptualized as a problem in which central processes are involved (Jastreboff, Hazell, & Graham, 1994). This is in contrast to earlier models which emphasized the acoustic characteristics and peripheral causes of tinnitus. According to Jastreboff et al. (1994) and Jastreboff (1990), tinnitus is a phantom auditory sensation and they have developed an animal model to test the implications of their theory. Advances in neuroscience research, for example, by Flor et al. (1995), on phantom-limb pain and cortical reorganization, may improve the understanding of tinnitus from a neuropsychological point of view. Interestingly, a case study by Salah, DeQuardo, Jibson, Carli, and Tandon (1995) showed that tinnitus was alleviated by electroconvulsive therapy in a depressed patient. Although promising, research on evoked potentials in tinnitus patients (Attias, Urbach, Gold, & Sheemesh, 1993; Colding-Jorgensen, Lauritzen, Johnsen, Mikkelsen, & Saermark, 1992) has not yet given consistent results overall; a similar picture has emerged from work on auditory brainstem responses (Rosenhall & Axelsson, 1995).

The ideas of Jastreboff and co-workers about tinnitus are in keeping with a psychological habituation model for tinnitus annoyance described by Hallam et al. (1984) in which they theorized that the natural reaction to tinnitus is to habituate. This original model has been challenged (Carlsson & Erlandsson, 1991), but little research has been done. Dishabituation may be another way to describe the process of developing tinnitus-related distress, or that the emotional coloring of the tinnitus sound is interpreted as a warning signal that short-circuits the habituation process.

9.05.4 ASSESSMENT

9.05.4.1 Audiological Assessments

Hearing impairment can be identified from observations by the patient or by their family and friends, by questionnaires, and by audiometric evaluations (Brooks, 1989). After being referred by the general practitioner, the patient is screened by audiometric tests that serve as a guideline for further rehabilitation. The procedures used vary, but most include pure-tone air and bone-conduction threshold determination. All of these are based on the patient's self-report of detecting or not detecting the stimuli presented. Speech audiometry is also commonly used; this involves the presentation of speech material to the patient through earphones, again requesting their evaluation of hearing. The presentation of subjective methods may vary, and masking procedures are often needed to determine mono-aural functioning. Sometimes, more objective methods such as otoscopy, impedance audiometry, stapedius reflex thresholds, brainstem audiometry, and a number of other specialized tests are needed (Katz, 1994).

Tinnitus cannot be measured without the subject's active participation. The audiological assessment of tinnitus, that is, tinnitus matching, usually consists of obtaining an estimate of loudness, pitch, maskability, and sometimes residual inhibition (short period of absence of tinnitus after exposure to a loud tone) (Tyler & Stouffer, 1989). Tinnitus matchings have a weak correlation with measures of annoyance and also a doubtful prognostic value. The reliability and validity studies of tinnitus matchings show large intra individual variations (Penner & Bilger, 1992). Recent thoughts hold that tinnitus maskability could be used as an outcome measure, and that reduced levels of masking required to mask the tinnitus could be regarded as a sign of treatment success (Jastreboff et al., 1994).

9.05.4.2 Self-report Measures

Self-assessment scales for the hearing impaired have been used in audiological research and practice (Schow & Gatehouse, 1990). Audiological measures are insufficient as descriptors of the problems experienced. Self-assessments help clarify issues concerning the experience of hearing loss, hearing handicap, rehabilitation effectiveness, and communication strategies used by the hearing-impaired patient. This is reflected in the number of scales available (Andersson, Melin, Lindberg, & Scott, 1995a; Schow & Gatehouse, 1990). Although Pearson correlations between actual hearing impairment and experienced disability/handicap approaching $r = 0.60$ have been found (Brainerd & Frankel, 1985; Lutman, Brown, & Coles, 1987), self-reported hearing problems and audiometric tests generally show relatively low correspondence. Daily assessments of hearing aid use and hearing problems are relatively rare, but have been applied (Andersson, Palmkvist, Melin, & Arlinger, 1996).

Similarly, a number of self-report scales for the assessment of tinnitus-related distress have been developed (Hallam et al., 1988; Kuk, Tyler, Russell & Jordan, 1990; Wilson et al., 1991). Most of these scales have adequate psychometric properties. Further, daily measures of tinnitus distress and loudness on visual analogue scales have been applied (Andersson & Hägnebo, 1996), even though the use of repeated measures is relatively rare. Self-report measures such as like the BDI (Beck et al., 1961), the Beck anxiety inventory (Beck, Epstein, Brown, & Steer, 1988), and other measures of psychological problems, personality, and coping, may well be used in audiological populations, given some consideration of the group being assessed. Self-report is becoming an important tool in audiological research and rehabilitation. It has a potential value for research on rehabilitation effectiveness and in the treatment itself. Daily measures have not been widely used.

9.05.4.3 Psychological Assessments

Little has been written on the psychological assessment of hearing-impaired adults. Functional and behavioral analysis have been applied as mentioned earlier. Cognitive behavioral assessment methods are well suited, as long as the communication is secured and the client can understand what is being said (Kirk, 1989). As the therapist/clinical psychologist may serve as a model later on, it is important that they are familiar with communicating with hearing-impaired people.

For the assessment of tinnitus patients, a structured interview has been developed in Uppsala, Sweden. In this interview, cognitive-behavioral assessment methods are used. All patients referred to the psychologist for tinnitus treatment are assessed in this interview. The interview consists of questions about the following: tinnitus characteristics (including a metaphor for tinnitus): attention given to tinnitus (what time during the day, in which situations, and so on); variations in annoyance and loudness; coping ability, avoidance, ability to influence the tinnitus; psychological consequences and prior psychiatric history (e.g., depression, anxiety, irritation, concentration difficulties); sleep problems; influence of environmental sounds (masking), influence of tiredness and stress; influence of medication (if there is any), caffeine, alcohol, and tobacco. A positive response to any of the questions leads to further questioning.

A grading of tinnitus severity is also made. The availability of social support for the patient is investigated. Information is also gathered about changes in tinnitus since it first appeared, habituation, and the patient's beliefs about the cause of tinnitus. The use of other treatments for tinnitus and any other conditions are enquired about. The patient is asked about hearing problems, hearing rehabilitation, the effect of the hearing aid on tinnitus, noise sensitivity, headaches, dizziness, muscle tension, and jaw problems. Finally, a cognitive-behavioral model of how tinnitus becomes annoying is described. The patient's acceptance of a psychological model can influence the outcome of therapy for somatic complaints (Salkovskis, 1989) and it is therefore important to determine the tinnitus patient's beliefs in this respect.

9.05.5 TREATMENT

When possible, medical intervention seeks to treat the underlying cause of hearing loss. Surgical techniques are available for conditions that give rise to conductive hearing loss. Surgery is also possible for some causes of sensory-neural hearing loss such as acoustic neuroma. Cochlear implantation offers some restoration of hearing for some profoundly deafened people. It is important to note that for the majority of hearing losses (e.g., presbyacusis), no curative medical or surgical treatment exists.

Many medical treatments for tinnitus have been investigated, including pharmacological (Murai, Tyler, Harker, & Stouffer, 1992), electrical stimulation with cochlear implants (Hazell, 1991), transcutaneous nerve stimulation (Kaada, Hognestad, & Havstad, 1989), ultrasound (Carrick, Davies, Fielder, & Bihari, 1986), dental procedures (stomatognatic); (Erlandsson, Rubenstein, & Carlsson, 1991), and surgery (Hazell, 1990). A range of alternative therapies have also been attempted such as *Ginkgo biloba* (Holgers, Axelsson, & Pringle, 1994) and acupuncture (Andersson & Lyttkens, 1996). In other words, practically everything has been tried in the treatment of tinnitus. To sum up a large body of literature, the results of these attempts to alleviate tinnitus have been modest or unsuccessful. The exception may be treatment with antidepressants (Sullivan, Katon, Russo, Dobie, & Sakai, 1993), but in our experience it is not likely that this will become the treatment of choice for all tinnitus patients.

9.05.5.1 Hearing Aids and Other Prosthetic Devices

Since there exists no curative treatment for most cases of hearing loss (especially among the elderly), the hearing aid is the treatment of first choice for many individuals seeking audiological assistance. There are few controlled studies showing the benefits of hearing-aid fitting. One exception was a study by Mulrow et al. (1990) who randomly assigned elderly subjects to hearing-aid fitting or to a waiting-list control. Results showed significant improvements for social and emotional function, communication function, cognitive function, and depression. The authors concluded that hearing impairment is associated with adverse effects and that these are reversible with hearing aids.

The amount of use of hearing aids has been the subject of several studies (for a review see Andersson, 1995). The amount of use varies between those studies, but it is well recognized that many hearing aids remain in the *file-drawer* (Brooks, 1989). In the hearing-aid field, the theoretical framework is basically a technical approach. It could be argued that this focuses primarily on the provision of hearing aids and devotes little attention to the psychological aspects of adjustment to the hearing aid. Consequently, it is mostly the technical causes that are used to explain the limited hearing aid benefit often observed. There are many possible reasons why people may obtain limited benefit from hearing aids. A proper assessment of the issue must go beyond the technical considerations and include attitudes towards the hearing aid and other social and psychological factors (Andersson, 1995; Brooks, 1989; Tyler et al., 1983).

Hearing aids have been proposed as a treatment of tinnitus (Surr, Montgomery, & Mueller, 1985). The notion is that environmental sounds amplified by the hearing aid mask tinnitus. Although a controlled study by Melin, Scott, Lindberg, and Lyttkens (1987) failed to find any evidence of this, some patients report benefit from using a hearing aid.

Much hope and research effort has been invested in the fitting of tinnitus maskers and combined hearing aid/tinnitus maskers (Hazell et al., 1985). The purpose of tinnitus maskers is to obscure the tinnitus with a sound that is supposed to be more tolerable. Although early experience with maskers was positive, controlled studies and later experiences have been less encouraging (Erlandsson, Ringdahl, Hutchins, & Carlsson, 1987). Jastreboff et al. (1994) have suggested that masking should not involve totally obscuring the tinnitus. Instead, the tinnitus should be audible and be a part of the background sound. This is accomplished by a so-called white-noise generator.

A cochlear implant can be of assistance to people with profound hearing impairment. Parts of this device are surgically implanted into the ear. The cochlear implant receives sound, converts this into an electrical signal that is then carried past the damaged parts of the ear and used to stimulate the preserved parts of the cochlea or the auditory nerve directly (Clark, Tong, & Patrick, 1990). Cochlear implants therefore straddle the surgical and prosthetic approaches. They have been at the leading edge of bionic technology since the early 1980s. It has received some interest from a psychological point of view (McKenna, 1993). McKenna and Denman (1993) found that implant users reported larger perceived improvements in psychological well-being than might be expected from relatively modest acoustical benefits. Knutson et al. (1991) reported psychological improvements that were uncorrelated with acoustical abilities. The use of cochlear implants for the treatment of tinnitus is restricted to profoundly hearing-impaired people where tinnitus is a main primary complaint (Hazell, 1991).

Hearing aids can be very helpful in the rehabilitation of hearing handicap. There are, however, a number of problems relating to hearing-aid use yet to be resolved. Among the most important ones are motivating people to seek help and then to use the aid in appropriate settings. Tinnitus maskers have not been as helpful as expected, but the more recent approach of noise generators may be more effective.

9.05.5.2 Counseling and Self-help Approaches

The importance of counseling in the management of hearing loss, especially when fitting hearing aids, was pointed out by Brooks (1989). There is some empirical support for counseling programs for improving hearing-aid benefit (Brooks, 1989), but others have found less evidence of benefit (Parving & Sibelle, 1995). It is difficult to differentiate counseling from general clinical management. The time spent on each individual patient may be one distinction; counseling programs usually devote more time to each patient. A program described by Brooks (1989) includes home visits. The need for counseling tinnitus patients has also been pointed out (Stouffer et al., 1991), but with more emphasis on formal psychological treatment.

Auditory training is a technique that involves taking advantage of acoustic cues still available for the hearing-impaired person. Some support for auditory training exists (Tye-Murray, 1991), but there are conflicting results (Tye-Murray, 1992). Hearing tactics were defined by van der Lieth (1973) as "those methods used by someone suffering from a hearing impairment to solve the problems of his daily life—the practical, technical, and psychological problems caused by the handicap" (p. 209). Hearing tactics often deal with different ways of facilitating communication, such as optimizing signals and using conversational strategies. The use of hearing tactics involves a broader set of techniques than auditory training. In clinical practice, these techniques are usually taught by audiological scientists or hearing therapists; however, it is useful for clinical psychologists to be aware of these ideas. Few empirical investigations of the benefits of hearing tactics exist (Andersson, 1995).

Self-help approaches have been suggested and applied for both hearing loss and tinnitus (Hallam, 1989; Shimon, 1992; Slater, 1987), but only a few empirical investigations exist (Axelsson, Nilsson, & Coles, 1995; Loumidis, Hallam, & Cadge, 1991). These approaches usually amounts to patient education, which can sometimes be sufficient, although giving the patient adequate information should be normal practice.

9.05.5.3 Psychological Treatments for Hearing Loss and Tinnitus

A functional approach involving goal-setting activities in audiological rehabilitation has been proposed by McKenna (1987) and its use described in the care of an individual patient. A cognitive-behavioral adaptation of hearing tactics has been developed. The group treatment package includes applied relaxation, video self-modeling, exposure, information, and teaching of hearing tactics. Treatment is based on a cognitive-behavioral analysis of each patient. Increasing hearing ability in the acoustic sense has not been a goal, except for the benefits of behavioral adjustments when listening. Instead, the aim is to develop coping strategies for handling the consequences of hearing impairment.

Lindberg, Scott, Andersson, and Melin (1993) did a controlled study with elderly patients randomly allocated to treatment or to a control condition. The treatment group received a general information session before the actual treatment. Cognitive–behavioral treatment began according to the individualized functional analyses in which the subjects" strengths, needs, and treatment goals were established. This was followed by three sessions in which hearing tactics and coping strategies were taught and practised. The choice of these was based on the result of the functional analysis. Treatment was given in three one-hour sessions over three consecutive weeks, in either an individual, a group, or a combined format. The overall aim was to supply the subjects with a sense of control and an ability to control their hearing problems. The role of their spouses/relatives and cognitive aspects was highlighted during treatment. The results of self-report questionnaires showed some effects suggesting that treatment had improved subjects' ability to cope with their hearing loss. In a 15-month follow-up study, the effects were sustained and even more in favor of the treatment group (Andersson, Melin, Scott, & Lindberg, 1994).

In another controlled study, the effects of a comprehensive group behavioral treatment program were evaluated (Andersson, Melin, Scott, & Lindberg, 1995b). This consisted of four two-hour sessions after an individual behavioral analysis. The first treatment session provided general information about the importance of behavior in coping with hearing impairment, and individualized treatment goals were determined for each subject. The session also included a short version of applied relaxation (Öst, 1987). Relaxation skills were associated with difficult hearing situations that became the targets for homework assignments. In the second session, relaxation skills were further rehearsed and linked to actual behavior in difficult situations. A large part of the session was devoted to video self-modeling (Meharg & Woltersdorf, 1990). A difficult situation was created with the help of an accompanying background noise played on a tape-recorder. Subjects were instructed on how to handle conversation by means of hearing tactics. During this training the therapist prompted the subject until good performance was achieved. Then, the situation was video-recorded and later the group watched the video and saw themselves communicate successfully. During this training the use of hearing tactics was reinforced and the ability to concentrate on one person at a time without being tense and disturbed by the background noise was supported. The third session began with rehearsal of the relaxation skills and a refinement of these. The main theme of the third session was on how to cope with loud noises. In a relaxed state, subjects were gradually exposed to tones (440 Hz, 1 kHz, and white noise) and a tape of people talking that slowly increased in loudness. Subjects were never exposed to loud or potentially dangerous volumes, but rather the natural volume of sounds in daily life. The purpose of the exercise was to teach the subjects to be relaxed in a noisy environment. Coping skills including the use of hearing aid(s) were practiced. A leaflet with advice on communication with hearing-impaired people was provided, and discussion of this with relatives was set as a homework assignment. Earlier treatment ingredients were rehearsed in the final session. Attention was given to cognitive skills and motivation to engage in activities while having impaired hearing. Outcome measures included a structured video interview measuring coping behavior, daily ratings of hearing problems on visual analogue scales, a questionnaire, and a follow-up telephone interview. The structured interview included tasks with minor provocations, making communicating more difficult with the aim of eliciting coping behavior. Subjects were informed that the purpose of filming was to investigate communication during the interview. Significant beneficial effects in favor of the treatment package were found in terms of behavioral changes and self-reported problems. A two-year follow-up showed positive effects of the treatment in terms of self-assessed hearing problems (Andersson, Melin, Scott, & Lindberg, 1995a). Unfortunately, these studies have not been replicated but, given the large number of older people with hearing loss, clinical psychology methods could be an asset to audiology.

Far more empirical work has been done on the psychological treatment of tinnitus (Andersson, Melin, Hägnebo, Scott, & Lindberg, 1995g). The psychological treatment studies can be divided into hypnosis, biofeedback, relaxation, and cognitive-behavioral therapy (CBT) with or without applied relaxation. A cognitive-behavioral approach to tinnitus was formulated by Hallam et al. (1984) and by Scott, Lindberg, Melin, & Lyttkens (1985), but behavioral applications had been described earlier (Malesta, Sutker, & Adams, 1980). Andersson, Melin, Hägnebo et al. (1995) drew the conclusion that CBT in combination with applied relaxation was the treatment approach that had received the most empirical support. In one of the earlier controlled studies, Scott et al. (1985) investigated the effects of combining relaxation and distraction techniques with a waiting-list control condition. Treatment consisted of one-hour sessions during a three-week period and included training in self-control by distraction exercises. Visual analogue measures of loudness and annoyance were used as dependent measures. Significant decreases in ratings of annoyance but not loudness were observed. Lindberg, Scott, Melin, and Lyttkens (1987) carried out a nine-month follow-up and reported positive results. More recent studies corroborate these findings (Henry & Wilson, 1996; Kröner-Herwig et al., 1995; Lindberg, Scott, Melin, & Lyttkens, 1989). The use of relaxation training only (Ireland, Wilson, Tonkin, & Platt-Hepworth, 1985) or of strictly cognitive methods (Jakes, Hallam, McKenna, & Hinchcliffe, 1992) seem to lead to poorer results.

9.05.6 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

The application of clinical psychology within the field of hearing disorders is comparatively new and the possibilities for further developments in research and practice are considerable. We will touch on only some of the more obvious ones here.

Although there is evidence for the effectiveness of psychological treatment in the management of patients with hearing disorders, there is a need for this body of knowledge to be expanded. The development of effective management techniques, particularly for tinnitus, remains a priority. The relatively modest benefits reported in tinnitus treatment outcome studies are at odds with much clinical experience which suggests that psychological therapy leads to considerable reductions in distress. This disparity may be due to factors such as subject selection techniques used in formal outcome studies or to the use of inappropriate measures. These issues merit investigation.

The study of central factors in tinnitus perception is still in its infancy. Complaints of concentration problems and other minor neuropsychological difficulties are common among tinnitus patients. These complaints have traditionally been regarded as part of the emotional distress associated with tinnitus. Whether this is so or whether these cognitive difficulties represent a separate aspect of tinnitus complaint is open to question. To our knowledge there has been only one study (McKenna, Shurlock, & Hallam, 1996) of the cognitive effects of tinnitus. Neurophysiological studies (Attias et al., 1993) have pointed (although inconsistently) to the involvement of brain areas in tinnitus perception that can also be implicated in neuropsychological functioning. Investigations of cognitive functioning in tinnitus patients will offer a complementary line of study of this issue and may assist in the development of models of tinnitus perception.

There is some controversy about whether the natural history of tinnitus is characterized by habituation or by some other process. Clarification of this central issue would guide future management strategies. Surprisingly little longitudinal data are available on tinnitus and such data could make a valuable contribution to tinnitus research and management.

Much of the research carried out on emotional factors in hearing disorders has been psychiatric in nature. Historically, there has been an interest in classifying patients in psychiatric terms. Although there are areas that require further clarification, such as the extent to which hearing-impaired people suffer from anxiety disorders, there is a need now for further work on the psychological processes that underlie the changes which accompany hearing disorders.

The cognitive–behavioral approach to hearing tactics is a field in which little controlled treatment research has been done. The approach could benefit a very large group of people should the principles be applied in everyday audiological management. Little is known about what tactics are the most effective ones, and here some experimental work is needed (Andersson, Melin, Lindberg, & Scott, 1996). Assessment of social skills in hearing-impaired people would be an important research task and one that is especially relevant for planning of intervention strategies. Also, the benefits of hearing tactics training before the provision of hearing aids should be investigated. So far, this has not been done. The development of self-help treatments is another important research goal. In a controlled study, Andersson, Green, and Melin (1997) found positive effects of a self-help manual supported by telephone contacts in teaching hearing tactics.

Cochlear implants are now issued on a routine service basis; however, there is still considerable scope for psychological research in this field. In particular, there is a need to develop predictors of outcome. There are large individual differences in the success of implant use that can be only poorly predicted. Intuitively it might be supposed that differences in information processing abilities might be involved in determining the success of implant use, but this is an area that is largely unexplored. Clinically it is also apparent that a small number of implant users, who are successful in terms of the acoustic benefit derived, experience adverse psychological consequences. Although this matter has received no formal research attention, it appears that this is a result of the restoration of hearing leading to unfavorable changes in the dynamics of everyday life.

Few clinical psychologists work within audiology clinics. This chapter has focused on the problems of hearing loss and tinnitus; other audiological disorders, such as balance disorders (Yardley, 1994) and hypersensitivity to noise (Vernon, 1987), also have major psychological elements to them. The need for more clinical psychologists to be involved in this field is clear. The task here is twofold: to educate other health-care professionals and patients of the benefits of clinical psychology, and also to persuade clinical psychologists that this is an area that merits their attention. It may be that lessons can be learnt from the field of chronic pain where clinical psychologists are more firmly established. The development of multidisciplinary teams, including psychologists, for the management of hearing disorders are likely to be of benefit as they have been in the field of pain management. Wilson, Henry, and Nicholas (1993) are beginning to elaborate on the parallels between pain and tinnitus. It might be argued that because of the sacristy of resources, psychologists should not work with people who are only moderately distressed by their medical condition (although this cannot be said about hearing loss in general, many do adjust to their problems without the need for counseling). Preventive work, however, can be very cost-effective and could prevent the development of psychological problems in association with hearing disorders. To our knowledge nothing has yet been written about prevention in this field, in contrast to the field of chronic pain (Linton, 1987).

9.05.7 SUMMARY

Hearing impairment and tinnitus are among the most prevalent of symptoms that affect people and particularly elderly people. There are many prejudices associated with these symptoms. In reality, the experience of these symptoms can be very different from the widely held beliefs about them. An understanding of hearing disabilities may assist a clinical psychologist to help large numbers of people with these potentially very distressing symptoms.

From the evidence available it seems that hearing impairment does not directly cause severe psychological disturbances, although an elevated risk for depression might exist, mainly by its association with isolation (Gilhome-Herbst, 1983). Instead, it is likely that a vulnerability for psychological disturbance or premorbid/comorbid psychological or physical problems interact with hearing impairment to produce adverse consequences, that is, a stress–diathesis conception. There is more widespread agreement on the link between tinnitus and emotional disturbance, but no study has established a causal link between tinnitus and psychological complaints. In fact, Stouffer and Tyler (1992) have suggested that psychological complaints may precede tinnitus complaints. Again, a stress–diathesis conception is appealing; certainly, tinnitus may be an important cause of, for example, depression.

Psychological concepts have been used in the study of hearing impairment and tinnitus and individual differences are evident. The concepts of coping and personal control have been applied to describe the process of adjusting to hearing loss and tinnitus. Ecological, cognitive–behavioral, and functional approaches have been presented as alternatives to the predominating medical and technical models in audiological rehabilitation. In the field of tinnitus, a habituation model describing the natural history of tinnitus has been predominant since the late 1980s. A neurophysiological model that has some parallels to the habituation model has begun to receive empirical support.

Assessment of patients with hearing disorders must include audiological tests and medical screening. Self-assessment is an important but not a sufficient tool. Discussing the results of the self-assessment helps clarify issues surrounding the severity and situation characteristics of the problems. Cognitive-behavioral assessment methods are usually applicable with hearing-impaired patients, but the communication situation requires certain precautions. Summarizing and clarifying the content of the interview and asking the patient for their questions and views is all the more important when working with hearing-impaired people.

Psychological treatment methods have been successfully applied in controlled studies. Studies of cognitive-behavioral hearing tactics suggest that clinical psychology can be applied in the rehabilitation of elderly hearing-impaired people. More studies are needed in this field and the techniques need to be applied with other patient groups (for example, adults with noise-induced hearing loss). Controlled studies suggest that the psychological management of tinnitus patients leads to positive benefits. Few psychologists are involved in tinnitus treatment. In a survey of tinnitus management in Britain, Coles (1992) claimed that there were few clinical psychologists working with tinnitus patients and that few patients were referred to psychiatrists or psychologists. The field is much in need of more psychologists to help in the development of research and in the provision of clinical services.

ACKNOWLEDGMENTS

This work was supported by grants to Dr. Gerhard Andersson from the Swedish Council for Research in the Humanities and Social Sciences and The Sasakawa Young Leaders' Fellowship Fund.

9.05.8 REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Andersson, G. (1995). Hearing as behaviour: Psychological aspects of acquired hearing impairment in the elderly (Doctoral dissertation). *Comprehensive Summaries of Uppsala Dissertations from the faculty of Social Sciences 53.* Uppsala, Sweden: Acta Universitas Upsaliensis.

Andersson, G. (1996). The role of optimism in patients with tinnitus and in patients with hearing impairment. *Psychology and Health, 11,* 697–707.

Andersson, G., & Green, M. (1995). Anxiety in elderly hearing impaired persons. *Perceptual and Motor Skills, 81,* 552–554.

Andersson, G., Green, M., & Melin, L. (1997). Behavioural hearing tactics: a controlled trial of a short treatment programme. *Behaviour Research and Therapy, 35,* 523–530.

Andersson, G., & Hägnebo, C. (1996). Dysphoria, optimism, confidence in activities and daily symptoms of Ménière's disease. *Journal of Audiological Medicine, 5,* 83–91.

Andersson, G., & Lyttkens, L. (1996). Acupuncture for tinnitus: Time to stop? *Scandinavian Audiology, 25,* 273–275.

Andersson, G., & Melin, L. (1993). Functional analysis of hearing handicap. *Scandinavian Journal of Behaviour Therapy, 22,* 147–157.

Andersson, G., Melin, L., Hägnebo, C., Scott, B., & Lindberg, P. (1995). A review of psychological treatment approaches for patients suffering from tinnitus. *Annals of Behavioral Medicine, 17,* 357–366.

Andersson, G., Melin, L., Lindberg, P., & Scott, B. (1995a). Development of a short scale for self-assessment of experiences of hearing impairment: The Hearing Coping Assessment. *Scandinavian Audiology, 24,* 147–154.

Andersson, G., Melin, L., Lindberg, P., & Scott, B. (1995b). Dispositional optimism, dysphoria, health, and coping with hearing impairment in elderly adults. *Audiology, 34,* 76–84.

Andersson, G., Melin, L., Lindberg, P., & Scott, B. (1996). Elderly hearing impaired persons' coping behavior. *International Journal of Behavioral Medicine, 3,* 303–320.

Andersson, G., Melin, L., Scott, B., & Lindberg, P. (1994). Behavioural counselling for subjects with acquired hearing loss. A new approach to hearing tactics. *Scandinavian Audiology, 23,* 249–256.

Andersson, G., Melin, L., Scott, B., & Lindberg, P. (1995a). A two-year follow-up examination of a behavioural treatment approach to hearing tactics. *British Journal of Audiology, 29,* 347–354.

Andersson, G., Melin, L., Scott, B., & Lindberg, P. (1995b). An evaluation of a behavioural treatment approach to hearing impairment. *Behaviour Research and Therapy, 33,* 283–292.

Andersson, G., Palmkvist, A., Melin, L., & Arlinger, S. (1996). Predictors of daily assessed hearing aid use and hearing capability using visual analogue scales. *British Journal of Audiology, 30,* 27–35.

Attias, J., Shemsh, Z., Bleich, A., Solomon, Z., Bar-Or, G., Alster, J., & Sohmer, H. (1995). Psychological profile of help-seeking and non-help-seeking tinnitus patients. *Scandinavian Audiology, 24,* 13–18.

Attias, J., Urbach, D., Gold, S., & Sheemesh, Z. (1993). Auditory event related potentials in chronic tinnitus patients with noise induced hearing loss. *Hearing Research, 71,* 106–113.

Axelsson, A., Nilsson, S., & Coles, R. (1995). Tinnitus information: A study by questionnaire. *Audiology, 34,* 301–310.

Axelsson, A., & Ringdahl, A. (1989). Tinnitus—a study of its prevalence and characteristics. *British Journal of Audiology, 23,* 53–62.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. (1988). An Inventory for measuring clinical anxiety. Psychometric properties. *Journal of Consulting and Clinical Psychology, 56,* 893–897.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561–571.

Brainerd, S. H., & Frankel, B. G. (1985). The relationship between audiometric and self-report measures of hearing handicap. *Ear and Hearing, 6,* 89–92.

Briner, W., Risey, J., Guth, P., & Noris, C. (1990). Use of the million clinical multiaxial inventory in evaluating patients with severe tinnitus. *American Journal of Otolaryngology, 11,* 334–337.

Brooks, D. N. (1989). *Adult aural rehabilitation.* London: Chapman and Hall.

Budd, R. J., & Pugh, R. (1995). The relationship between locus of control, tinnitus severity, and emotional distress in a group of tinnitus sufferers. *Journal of Psychosomatic Research, 39,* 1015–1918.

Carlsson, S. G., & Erlandsson, S. I. (1991). Habituation and tinnitus: An experimental study. *Journal of Psychosomatic Research, 35,* 509–514.

Carrick, D. G., Davies, W. M., Fielder, C. P., & Bihari, J. (1986). Low-powered ultrasound in the treatment of tinnitus: A pilot study. *British Journal of Audiology, 20,* 153–155.

Chan, S. W. Y., & Reade, P. C. (1994). Tinnitus and temporomandibular pain-dysfunction disorder. *Clinical Otolaryngology, 19,* 370–380.

Clark, D. B., Hirsch, B. E., Smith, M. G., Furman, J. M. R., & Jacob, R. G. (1994). Panic in otolaryngology patients presenting with dizziness or hearing loss. *American Journal of Psychiatry, 151,* 1223–1225.

Clark, G. M., Tong, Y. C., & Patrick, J. (1990). Cochlear protheses: Introduction. In G. M. Clark, Y. C. Tong, & J. Patrick (Eds.), *Cochlear protheses* (pp. 1–14). New York: Churchill Livingstone.

Colding-Jorgensen, E., Lauritzen, M., Johnsen, N. J., Mikkelsen, K. B., & Saermark, K. (1992). On the evidence of auditory evoked magnetic fields as an objective measure of tinnitus. *Electroencephalography and Clinical Neurophysiology, 83,* 322–327.

Coles, R. R. A. (1992). A survey of tinnitus management in national health service hospitals. *Clinical Otolaryngology, 17,* 313–316.

Collet, L., Moussu, M. F., Disant, F., Ahami, T., & Morgon, A. (1990). Minnesota multiphasic personality inventory in tinnitus disorders. *Audiology, 29,* 101–106.

Coren, S., & Harland, R. E. (1995). Personality correlates of variations in visual and auditory abilities. *Personality and Individual Differences, 18,* 15–25.

Danhauer, J. L., Johnson, C. E., Kasten, R. N., & Brimacombe, J. A. (1985). The hearing aid effect: Summary, conclusions & recommendations. *Hearing Journal, 38,* 12–14.

Davis, A. C. (1989). The prevalence of hearing impairment and reported hearing disability among adults in Great Britain. *International Journal of Epidemiology, 18,* 911–917.

Davis, A. (1993). The prevalence of deafness. In J. Ballantyne, M. C. Martin, & A. Martin (Eds.), *Deafness* (5th ed., pp. 1–11). London: Whurr.

Eastwood, M. R., Corbin, S. L., Reed, M., Noobs, H., & Kedward, H. B. (1985). Acquired hearing loss and psychiatric illness: An estimate of prevalence and comorbidity in a geriatric setting. *British Journal of Psychiatry, 147,* 552–556.

Eriksson-Mangold, M., & Carlsson, S. G. (1991). Psychological and somatic distress in relation to perceived hearing disability, hearing handicap, and hearing measurements. *Journal of Psychosomatic Research, 35,* 729–740.

Erlandsson, S., Ringdahl, A., Hutchins, T., & Carlsson, S. G. (1987). Treatment of tinnitus: a controlled comparison of masking and placebo. *British Journal of Audiology, 21,* 37–44.

Erlandsson, S., Rubenstein, B., & Carlsson, S. G. (1991). Tinnitus: evaluation of biofeedback and stomatognatic treatment. *British Journal of Audiology, 25,* 151–161.

Flor, H., Elbert, T., Knecht, S., Weinbruch, C., Pantev, C., Birbaumer, N., Larbig, W., & Taub, E. (1995). Phantom-limb pain as a perceptual correlate of cortical reorganization following arm amputation. *Nature, 375,* 482–484.

Gerber, K. E., Nehemkis, A. M., Charter, R. A., & Jones, H. C. (1985). Is tinnitus a psychological disorder? *International Journal of Psychiatry in Medicine, 15,* 81–87.

Gibson, J. J. (1986). *The ecological approach to visual perception.* Hillsdale, NJ: Erlbaum.

Gilhome-Herbst, K. (1983). Psychosocial consequences of disorders of hearing in the elderly. In R. Hinchcliffe (Ed.), *Hearing and balance in the elderly* (pp. 174–200). New York: Churchill Livingstone.

Gilhome-Herbst, K., & Humphrey, C. (1980). Hearing impairment and mental state in the elderly living at home. *British Medical Journal, 281,* 903–905.

Halford, J. B. S., & Anderson, S. D. (1991). Anxiety and depression in tinnitus sufferers. *Journal of Psychosomatic Research, 35,* 383–390.

Hallam, R. S. (1989). *Living with tinnitus: Dealing with the ringing in your ears.* Wellingborough, UK: Thorsons.

Hallam, R. S., Jakes, S. C., & Hinchcliffe, R. (1988). Cognitive variables in tinnitus annoyance. *British Journal of Clinical Psychology, 27,* 213–222.

Hallam, R. S., Rachman, S., & Hinchcliffe, R. (1984). Psychological aspects of tinnitus. In S. Rachman (Ed.), *Contributions to medical psychology* (pp. 31–53). Oxford, UK: Pergamon.

Hallberg, L. R. M., & Barrenäs, M. L. (1993). Living with a male with noise-induced hearing loss: experiences from the perspective of spouses. *British Journal of Audiology, 27,* 255–261.

Hallberg, L. R. M., & Erlandsson, S. I. (1993). Tinnitus characteristics in tinnitus complainers and noncomplainers. *British Journal of Audiology, 27,* 19–27.

Haynes, S. N., & O'Brien, W. H. (1990). Functional analysis in behavior therapy. *Clinical Psychology Review, 10,* 649–668.

Hazell, J. W. P. (1990). Tinnitus II: Surgical management of conditions associated with tinnitus and somatosounds. *Journal of Otolaryngology, 19,* 6–10.

Hazell, J. W. P. (1991). Electrical tinnitus suppression. In H. Cooper (Ed.), *Cochlear implants. A practical guide* (pp. 355–369). London: Whurr.

Hazell, J. W. P., & Jastreboff, P. J. (1990). Tinnitus I: Auditory mechanisms: A model for tinnitus and hearing impairment. *Journal of Otolaryngology, 19,* 1–5.

Hazell, J. W. P., Wood, S. M., Cooper, H. R., Stephens, S. D. G., Corcoran, A. L., Coles, R. R. A., Baskill, J. L., & Sheldrake, J. B. (1985). A clinical study of tinnitus maskers. *British Journal of Audiology, 19,* 65–146.

Henry, J. L., & Wilson, P. H. (1996). Psychological management of tinnitus: Comparison of a combined cognitive educational program, education alone and a waiting-list control. *International Tinnitus Journal, 2,* 1–12.

Hiller, W., Goebel, G., & Rief, W. (1994). Reliability of self-rated tinnitus distress and association with psychological symptom patterns. *British Journal of Psychology, 33,* 231–239.

Hinchcliffe, R., & King, P. F. (1992). Medicolegal aspects of tinnitus. 1: Medicolegal position and current state of knowledge. *Journal of Audiological Medicine, 1,* 38–58.

Holgers, K.- M., Axelsson, A., & Pringle, I. (1994). Ginko biloba for the treatment of tinnitus. *Audiology, 33,* 85–92.

Ireland, C. E., Wilson, P. H., Tonkin, J. P., & Platt-Hepworth, S. (1985). An evaluation of relaxation training in the treatment of tinnitus. *Behaviour Research and Therapy, 23,* 423–430.

Jakes, S. (1988). Otological symptoms and emotion. A review of the literature. *Advances in Behaviour Research and Therapy, 10,* 53–103.

Jakes, S. C., Hallam, R. S., McKenna, L., & Hinchcliffe, R. (1992). Group therapy for medical patients: An application to tinnitus. *Cognitive Therapy and Research, 16,* 67–82.

Jastreboff, P. J. (1990). Phantom auditory perception (tinnitus): Mechanisms of generation and perception. *Neuroscience Research, 8,* 221–524.

Jastreboff, P. J., & Hazell, J. W. P. (1993). A neurophysiological approach to tinnitus: Clinical implications. *British Journal of Audiology, 27,* 7–17.

Jastreboff, P. J., Hazell, J. W. P., & Graham, R. L. (1994). Neurophysiological model of tinnitus: Dependence of the minimal masking level on treatment outcome. *Hearing Research, 80,* 216–232.

Jones, D. A., Victor, C. R., & Vetter, N. J. (1984). Hearing difficulty and its psychological implications for the elderly. *Journal of Epidemiology and Community Health, 38,* 75–78.

Kaada, B., Hognestad, S., & Havstad, J. (1989). Transcutaneous nerve stimulation (TNS) in tinnitus. *Scandinavian Audiology, 18,* 211–217.

Katz, J. (1994). *Handbook of clinical audiology* (4th ed.). Baltimore: Williams & Wilkins.

Kirk, J. (1989). Cognitive-behavioural assessment. In K. Hawton, P. M. Salkovskis, J. Kirk, & D. M. Clark (Eds.), *Cognitive behaviour therapy for psychiatric problems* (pp. 13–51). Oxford, UK: Oxford University Press.

Kirsch, C. A., Blanchard, E. B., & Parnes, S. M. (1989). Psychological characteristics of individuals high and low in their ability to cope with tinnitus. *Psychosomatic Medicine, 51,* 209–217.

Knutson, J. F., & Lansing, C. R. (1990). The relationship between communication problems and psychological difficulties in persons with profound acquired hearing loss. *Journal of Speech and Hearing Disorders, 55,* 656–664.

Knutson, J. F., Schartz, H. A., Gantz, B. J., Tyler, R. S., Hinrichs, J. V., & Woodworth, G. (1991). Psychological change following 18 months of cochlear implant use. *Annals of Otology, Rhinology and Laryngology, 100,* 877–882.

Kröner-Herwig, B., Hebing, G., Van Rijn-Kalkmann, U., Frenzel, A., Schilkowsky, G., & Esser, G. (1995). The management of chronic tinnitus: Comparison of a cognitive-behavioural group training with yoga. *Journal of Psychosomatic Research, 39,* 153–165.

Kuk, F. K., Tyler, R. S., Russell, D., & Jordan, H. (1990). The psychometric properties of a tinnitus handicap questionnaire. *Ear and Hearing, 11,* 434–445.

Kyle, J. G., Jones, L. G., & Wood, P. L. (1985). Adjustment to axquired hearing loss: A working model. In H. Orlans (Ed.), *Adjustment to adult hearing loss* (pp. 119–138). London: Taylor & Francis.

Lewis, J. E., Stephens, S. D. G., & McKenna, L. (1994). Tinnitus and suicide. *Clinical Otolaryngology, 19,* 50–54.

Lieth van der, L. (1973). Hearing tactics II. *Scandinavian Audiology, 2,* 209–213.

Lindberg, P., Lyttkens, L., Melin, L., & Scott, B. (1984). Tinnitus: Incidence and handicap. *Scandinavian Audiology, 13,* 287–291.

Lindberg, P., Scott, B., Andersson, G., & Melin, L. (1993). A behavioural approach to individually designed hearing tactics training. *British Journal of Audiology, 27,* 299–301.

Lindberg, P., Scott, B., Melin, L., & Lyttkens, L. (1987). Long-term effects of psychological treatment of tinnitus. *Scandinavian Audiology, 16,* 167–172.

Lindberg, P., Scott, B., Melin, L., & Lyttkens, L. (1989). The psychological treatment of tinnitus: An experimental evaluation. *Behaviour Research and Therapy, 27,* 593–603.

Loumidis, K. S., Hallam, R. S., & Cadge, B. (1991). The effect of written reassuring information on out-patients complaining of tinnitus. *British Journal of Audiology, 25,* 105–109.

Linton, S. (1987). Chronic pain: The case for prevention. *Behaviour Research and Therapy, 25,* 313–317.

Lutman, M. E., Brown, E. J., & Coles, R. R. A. (1987). Self-reported disability and handicap in the population in relation to pure-tone threshold, age, sex, and type of hearing loss. *British Journal of Audiology, 21,* 45–58.

Malesta, V. J., Sutker, P. B., & Adams, H. E. (1980). Experimental assessment of tinnitus aurium. *Journal of Behavioral Assessment, 2,* 309–317.

McKenna, L. (1987). Goal planning in audiological rehabilitation. *British Journal of Audiology, 21,* 5–11.

McKenna, L. (1991). The assessment of psychological variables in cochlear implant patients. In H. Cooper (Ed.), *Cochlear implants: A practical guide* (pp. 125–145). London: Whurr.

McKenna, L. (1993). The psychological consequences of acquired hearing loss. In J. Ballantyne, M. C. Martin, & A. Martin (Eds.), *Deafness* (5th ed., pp. 239–242). London: Whurr.

McKenna, L., & Denman, C. (1993). Repertory grid technique in the assessment of cochlear implant patients. *Journal of Audiological Medicine, 2,* 75–84.

McKenna, L., Hallam, R. S., & Hinchcliffe, R. (1991). The prevalence of psychological disturbance in neuro-otology outpatients. *Clinical Otolaryngology, 16,* 452–456.

McKenna, L., Shurlock, & Hallam, R. S. (1996). Cognitive functioning in tinnitus patients. In G. E. Reich & J. A. Vernon (Eds)., *Proceedings of the Fifth International Tinnitus Seminar 1995* (pp. 589–595). Portland, OR: American Tinnitus Association.

Meadow-Orlans, K. P. (1985). Social and psychological effects of hearing loss in adulthood: A literature review. In H. Orlans (Ed.), *Adjustment to adult hearing loss* (pp. 35–57). London: Taylor & Francis.

Meharg, S. S., & Woltersdorf, M. A. (1990). Therapeutic use of videotape self-modeling: A review. *Advances in Behaviour Research and Therapy, 12,* 85–99.

Melin, L., Scott, B., Lindberg, P., & Lyttkens, L. (1987). Hearing aids and tinnitus—an experimental group study. *British Journal of Audiology, 21,* 91–97.

Mulrow, C. D., Aguilar, C., Endicott, J. E., Tuley, M. R., Velez, R., Charlip, W. S., Rhodes, M. C., Hill, J. A., & DeNino, L. A. (1990). Quality-of-life changes and hearing impairment. A randomized trial. *Annals of Internal Medicine, 113,* 188–194.

Murai, K., Tyler, R. S., Harker, L. A., & Stouffer, J. L. (1992). Review of pharmacological treatment of tinnitus. *American Journal of Otology, 13,* 454–464.

Noble, W. (1983). Hearing, hearing impairment, and the audible world: A theoretical essay. *Audiology, 22,* 325–338.

Noble, W., & Hétu, R. (1994). An ecological approach to disability and handicap in relation to impaired hearing. *Audiology, 33,* 117–126.

Norton, S. J., Schmidt, A. R., & Stover, L. J. (1990). Tinnitus and otoacustic emissions: Is there a link? *Ear and Hearing, 11,* 159–166.

Ohta, R. J., Carlin, M. F., & Harmon, B. M. (1981). Auditory acuity and performance on the mental status questionnaire in the elderly. *Journal of the American Geriatrics Society, 29,* 476–478.

Öst, L. G. (1987). Applied relaxation: Description of a coping technique and review of controlled studies. *Behaviour Research and Therapy, 25,* 379–409.

Parving, A., & Sibelle, P. (1995). An audit of hearing rehabilitation within the health service—past and present. *Scandinavian Audiology, 24,* 33–38.

Penner, M. J., & Bilger, R. C. (1992). Consistent within-session measures of tinnitus. *Journal of Speech and Hearing Research, 35,* 694–700.

Reidner, E. D., & Efros, P. L. (1995). Nonorganic hearing loss and child abuse: Beyond the sound booth. *British Journal of Audiology, 29,* 195–197.

Rosenhall, U., & Axelsson, A. (1995). Auditory brainstem response latencies in patients with tinnitus. *Scandinavian Audiology, 24,* 97–100.

Russo, J. E., Katon, W. J., Sullivan, M. D., Clark, M. R., & Buchwald, D. (1994). Severity of somatization and its relationship to psychiatric disorders and personality. *Psychosomatics, 35,* 546–556.

Salah, R. S., DeQuardo, J. R., Jibson, M. D., Carli, T., & Tandon, R. (1995). Tinnitus and ECT. *Convulsive Therapy, 11,* 122–125.

Salkovskis, P. M. (1989). Somatic problems. In K. Hawton, P. M. Salkovskis, J. Kirk, & D. M. Clark (Eds.), *Cognitive behaviour therapy for psychiatric problems* (pp. 235–276). Oxford, UK: Oxford University Press.

Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology, 4,* 219–247.

Schow, R. L., & Gatehouse, S. (1990). Fundamental issues in self-assessment of hearing. *Ear and Hearing, 11,* 6–16.

Scott, B., Lindberg, P., Melin, L., & Lyttkens, L. (1985). Psychological treatment of tinnitus. An experimental group study. *Scandinavian Audiology, 14,* 223–230.

Scott, B., Lindberg, P., Melin, L., & Lyttkens, L. (1990). Predictors of tinnitus discomfort, adaption and subjective loudness. *British Journal of Audiology, 24,* 51–62.

Scott, B., Lindberg, P., Melin, L., & Lyttkens, L. (1994). Control and dispositional style among the hearing-impaired in communication situations. *Audiology, 33,* 177–184.

Shimon, D. A. (1992). *Coping with hearing loss and hearing aids.* San Diego, CA: Singular Publishing Group.

Simpson, R. B., Nedzelski, J. M., Barber, H. O., & Thomas, M. R. (1988). Psychiatric diagnoses in patients with psychogenic dizziness or severe tinnitus. *Journal of Otolaryngology, 17,* 325–330.

Singerman, B., Riedner, E., & Folstein, M. (1980). Emotional disturbance in hearing clinic patients. *British Journal of Psychiatry, 137,* 58–62.

Skinner, B. F. (1957). *Verbal behavior.* New York: Appleton-Century-Crofts.

Skinner, B. F., & Vaughan, M. E. (1983). *Enjoy old age. A program for self-management.* London: Hutchinson.

Slater, R. (1987). On helping people with tinnitus to help themselves. *British Journal of Audiology, 21,* 87–90.

Stephens, S. D. G. (1980). Evaluating the problems of the hearing impaired. *Audiology, 19,* 205–220.

Stephens, S. D. G., Lewis, P. A., Charny, M. C., Farrow, S. C., & Francis, M. (1990). Characteristics of self-reported hearing problems in a community survey. *Audiology, 29,* 93–100.

Stouffer, J. L., & Tyler, R. S. (1992). Ratings of psychological changes pre- and post-tinnitus onset. In J. M. Aran & R. Dauman (Eds.), *Tinnitus 91; Proceedings of the fourth international tinnitus seminar.* (pp. 449–452). Amsterdam: Kugler Publications.

Stouffer, J. L., Tyler, R. S., Kileny, P. R., & Dalzell, L. E. (1991). Tinnitus as a function of duration and etiology: Counselling implications. *American Journal of Otolaryngology, 12,* 188–194.

Sullivan, M. D., Katon, W. J., Russo, J. E., Dobie, R. A., & Sakai, C. (1993). A randomized trial of nortriptyline for severe chronic tinnitus. *Archives of Internal Medicine, 153,* 2251–2259.

Sullivan, M., Katon, W., Russo, J., Dobie, R., & Sakai, C. (1994). Coping and marital support as correlates of tinnitus disability. *General Hospital Psychiatry, 16,* 259–266.

Surr, R. K., Montgomery, A. A., & Mueller, H. G. (1985). Effect of amplification of tinnitus among new hearing aid users. *Ear and Hearine, 6,* 71–75.

Swan, I. R. C., & Gatehouse, S. (1990). Factors influencing consultation for management of hearing disability. *British Journal of Audiology, 24,* 155–160.

Thomas, A., & Gilhome-Herbst, K. (1980). Social and psychological implications of acquired deafness for adults of employment age. *British Journal of Audiology, 14,* 76–85.

Tye-Murray, N. (1991). Repair strategy usage by hearing-impaired adults and changes following communication therapy. *Journal of Speech and Hearing Research, 34,* 921–928.

Tye-Murray, N. (1992). Preparing for communication interactions: The value of anticipatory strategies for adults with hearing impairment. *Journal of Speech and Hearing Research, 35,* 430–435.

Tyler, R. S., & Baker, L. J. (1983). Difficulties experienced by tinnitus sufferers. *Journal of Speech and Hearing Disorders, 48,* 150–154.

Tyler, R. S., Baker, L. J., & Armstrong-Bednall, G. (1983). Difficulties experienced by hearing-aid candidates and hearing-aid users. *British Journal of Audiology, 17,* 191–201.

Tyler, R. S., & Stouffer, J. L. (1989). A review of tinnitus loudness. *Hearing Journal, 42,* 52–57.

Uhlmann, R. F., Larson, E. B., Rees, T. S., Koepsell, T. D., & Duckert, L. G. (1989). Relationship of hearing impairment to dementia and cogntive dysfunction in older adults. *Journal of the American Medical Association, 261,* 1916–1919.

Vernon, J. A. (1987). Pathophysiology of tinnitus: A special case—Hyperacusis and a proposed treatment. *American Journal of Otology, 8,* 201–202.

Weinstein, B. E., & Ventry, I. M. (1982). Hearing impairment and social isolation in the elderly. *Journal of Speech and Hearing Research, 25,* 593–599.

Weiss, L. H., & Kerns, R. D. (1995). Patterns of pain-relevant social interactions. *International Journal of Behavioral Medicine, 2,* 157–171.

Wilson, P. H., Henry, J., Bowen, M., & Haralambous, G. (1991). Tinnitus reaction questionnaire: Psychometric properties of a measure of distress associated with tinnitus. *Journal of Speech and Hearing Research, 34,* 197–201.

Wilson, P. H., Henry, J. L., & Nicholas, M. K. (1993). Cognitive methods in the management of chronic pain and tinnitus. *Australian Psychologist, 28,* 172–180.

Yardley, L. (1994). *Vertigo and dizziness.* London: Routledge.

Zimbardo, P. G., Andersen, S. M., & Kabat, L. G. (1981). Induced hearing deficit generates experimental paranoia. *Science, 212,* 1529.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.06
Jail and Prison Inmates

RONALD ROESCH, JAMES R. P. OGLOFF, PATRICIA A. ZAPF,
and STEPHEN D. HART
Simon Fraser University, Burnaby, BC, Canada
and
RANDY OTTO
University of South Florida, Tampa, FL, USA

9.06.1 INTRODUCTION 85

9.06.2 THE PREVALENCE OF MENTAL ILLNESS AND SUBSTANCE ABUSE AMONG INMATES IN JAILS AND PRISONS 86
- *9.06.2.1 Substance Abuse* 88
- *9.06.2.2 Prevalence and Assessment of Women in Jails and Prisons* 89

9.06.3 THE DIVERSION OF MENTALLY ILL OFFENDERS 91

9.06.4 PRETRIAL ASSESSMENT 93
- *9.06.4.1 Competency to Stand Trial* 93
  - *9.06.4.1.1 Community-based screening* 94
  - *9.06.4.1.2 Competency to be executed* 94
- *9.06.4.2 Criminal Responsibility* 95
- *9.06.4.3 Mental Disorder and Violence* 97

9.06.5 THE TREATMENT OF MENTALLY ILL OFFENDERS IN JAILS AND PRISONS 97
- *9.06.5.1 The Surrey Pretrial Mental Health Project* 100
- *9.06.5.2 Unique Concerns for the Treatment of Female Offenders* 100

9.06.6 CONCLUSION 101

9.06.7 REFERENCES 101

9.06.1 INTRODUCTION

Although greater attention has been paid to the issue of mentally ill offenders in the criminal justice system since the early 1980s (e.g., Monahan & Steadman, 1983; Steadman & Monahan, 1984; Teplin, 1983, 1984), this certainly is not a new issue. Historically, people with mental illness and mental retardation were often incarcerated with offenders and "treated" by cruel methods similar to those used to punish offenders (Ives, 1914). Monahan and Geis (1976) reported that Benjamin Franklin was instrumental in establishing the first mental hospital in the American colonies because of his belief that people with mental illness should be confined given their proclivity toward dangerousness. In many ways, then, the current attention that is being paid to mentally ill offenders is long overdue.

This chapter provides an overview of many of the issues regarding mentally ill inmates in jails and prisons, focusing on the role of psychologists in assessing and treating mental disorder within the criminal justice system. It begins with a review of the prevalence of mental disorder and substance abuse in jail and prison inmate populations. While males comprise the majority in these facilities, these issues for female offenders are also considered. The literature on diversion of mentally ill offenders out of the criminal justice system is then reviewed. Pretrial assessment of several major mental health/law issues, including competency to stand trial, criminal responsibility, and risk, are reviewed. Finally, the issue of treatment of mentally disordered offenders in jails and prisons is considered, and a current project in a pretrial jail facility is described in some detail.

9.06.2 THE PREVALENCE OF MENTAL ILLNESS AND SUBSTANCE ABUSE AMONG INMATES IN JAILS AND PRISONS

Perhaps at the very heart of the issue of people with mental illness in the criminal justice system is the question of how many inmates in jails and prisons are mentally ill. Not surprisingly, there is no simple answer to this question. The number of inmates with mental illness varies considerably across institutions and studies depending on such factors as whether the institution is a jail or prison; whether the institution draws inmates from urban, suburban, or rural areas; the definition of "mental illness" that is employed; and how the assessment of mental illness is made.

The terms *jail* and *prison* are often used interchangeably; however, jails and prisons are very different types of facilities. Prisons typically house inmates who have been convicted and sentenced to serve one or more years in a correctional facility. By contrast, jail inmates generally are detained for relatively short periods of time (often just one or two days) prior to arraignment or trial, or for short sentences (less than a year) if found guilty of a misdemeanor and sentenced to jail (Ogloff, Tien, Roesch, & Eaves, 1991). Due to the very different mandates of jails and prisons, the number of inmates with mental illnesses that they house likely varies considerably. Indeed, as the jail is the point of entry in the criminal justice system, and serves as a short-term holding facility, one might expect the prevalence of mental illness among offenders to be higher than it is in prisons. Presumably, the criminal justice system serves—at least to some extent—as a filtering system to identify those offenders with serious mental illness and to divert them to the mental health system (i.e., those who are mentally ill and commit minor offenses, those who are incompetent to stand trial, or are not guilty by reason of insanity).

Several studies show that a small but significant number of inmates in prisons are mentally ill (see Daniel, Robins, Reid, & Wilfey, 1988; Dvoskin, & Steadman, 1989; Hodgins & Cote, 1990; Roth, 1980; Steadman, Fabisiak, Dvoskin, & Holohean, 1987; Teplin, 1983, 1984, 1990). Steadman, Dvoskin and their colleagues (Dvoskin & Steadman, 1989; Steadman et al., 1987) conducted a survey of 9.4% of the then 36 144 inmates in the New York state prison system to determine the extent of psychiatric disabilities among inmates. The results showed that 5% of inmates were "severely psychiatrically disabled," demonstrating psychopathology similar to that found in state psychiatric center acute inpatients. Another 10% were "significantly psychiatrically disabled," similar to patients in crisis beds in the community.

Similar results have been obtained in other jurisdictions with other assessment procedures. In one study, interviewers administered a structured diagnostic interview to 1240 inmates of the Michigan state prison system (Neighbors, 1987). The interview yielded diagnoses of mental disorder based on the American Psychiatric Association's *Diagnostic and statistical manual of mental disorder* (*DSM-III*) criteria (American Psychiatric Association, 1987). The prevalence of severe mental disorders (i.e., any active psychotic disorder in the sample was 10.4%, and a further 38.4% had a less serious mood or anxiety disorder. Mental health professionals conducted follow-up interviews with 379 inmates and confirmed the high prevalence of severe mental disorder. They also rated 19.7% of the inmates as having "severely impaired" social functioning and judged that 4.4% were in need of immediate psychiatric hospitalization.

As with prisons, a number of studies examining the prevalence of mental disorder in jail populations have been conducted. Rates of mental illness reported for jail inmates vary considerably among studies. Prevalence rates of major mental illness among random samples of inmates vary between 3% (Schuckit, Herrman, & Schuckit, 1977) and 59% (Petrich, 1976), depending upon the criteria and diagnostic system used. Not surprisingly, higher prevalence rates have been obtained when samples have consisted only of those inmates referred by jail staff for mental health evaluations. In these cases, prevalence rates ranging from 24% (Nielsen, 1979) to more than 75% (Lamb & Grant, 1983) have been reported. Similar findings were reported by Allodi, Kedward, and Robertson (1977) in a Canadian study.

In one study that examined mental disorder among detainees in a large jail in Chicago, lay-interviewers assessed the detainees using a structured diagnostic interview (Teplin & Swartz, 1989). The prevalence of severe mental disorder (i.e., schizophrenia, bipolar affective disorder, major depression) was 8.5%. This rate was significantly higher than that of community residents with similar demographic characteristics (Teplin, 1991). Corrections staff failed to detect many cases of severe mental disorder during routine intake procedures.

Research in a pretrial facility in British Columbia, Canada, demonstrated that a substantial number of jail detainees have mental health problems. Although Canada's per capita prison population is considerably lower than in the USA, many similarities exist. Data was collected from a random sample of English-speaking admissions to the Vancouver Pretrial Services Centre, a large pretrial facility in the province of British Columbia, for a one-year period (Roesch, 1995). The total sample of 861 represented approximately 50% of all English-speaking males admitted to the jail during the study period. The estimate of the rate of major mental disorder (primarily schizophrenic and major affective disorders) in the total pretrial population was 15.6%. In addition, the prevalence of substance use disorders was exceptionally high, with over 77% considered to have alcohol use or dependence disorders, and over 63% with drug use disorders. Many of these individuals had co-occurring disorders (e.g., schizophrenia and substance abuse), further compounding the mental health and social problems these individuals experience.

When considering the research on the prevalence of mental illness in jails and prisons, it is important to keep in mind that the extant research on prevalence and assessment cannot be easily compared across studies as no consistent methodology has been utilized. The diagnostic systems used vary across studies as do the assessment procedures and protocols. Teplin, Abram, and McClelland (1996), in a review of the research on women in jail, argued that there has been little epidemiological research conducted on women in jail and that there are methodological problems with the studies that have been done. Three common methodological problems outlined by Teplin et al. (1996) would apply equally to the research on men in jails. First, most studies use only selected samples of jail inmates such as those referred for evaluation and/or treatment. Second, sample sizes are often too small to generate reliable rates of rare psychiatric disorders, and many of the studies have been conducted using nonstandardized instruments to assess psychiatric disorder. Despite the methodological flaws of some of the research, it appears that mental disorder is at least as common, and probably far more common in jails than in the community (Teplin, 1983).

A number of observations can be made upon reviewing studies investigating the prevalence of mental illness in jails and prisons. Foremost, findings from studies examining randomly selected inmates suggest that there are a large number of persons incarcerated in jails with significant mental disorders. Substantial numbers of inmates (perhaps as many as 10%) suffer from psychotic disorders and an even greater number suffer from anxiety or personality disorders. These disorders significantly impair the inmates' social functioning and require immediate treatment or management, and some of the inmates are so acutely ill that they require inpatient psychiatric care. The rate of serious mental disorder among inmates is significantly higher than the rate of serious mental disorder among community residents from similar socioeconomic conditions. Although the data are not clear on the point, the rate of mental illness in jails and prisons may increase over time due to deinstitutionalization and other factors (e.g., Pogrebin & Poole, 1987; Rabkin, 1979; Roesch & Golding, 1985). Also, in most cases, the onset of mental disorder precedes the onset of criminal behavior (Hodgins & Cote, 1990).

There is a general perception shared by jail healthcare administrators and jail mental heath professionals that the number of persons with mental illness entering jails has increased over the years. Sixty-nine percent of jail administrators responding to a survey prepared by Torrey et al. (1992) reported that the number of persons with mental illness who were entering jail had increased over the course of 10 years. Moreover, others have claimed that the proportion of mentally disordered jail inmates is increasing (Teplin, 1983; Torrey et al., 1992). For example, Teplin (1983) concluded that the "research literature, albeit methodologically flawed, offers at least modest support for the contention that the mentally ill are (now) being processed through the criminal justice system" (p. 54).

The contention that the mentally ill are entering jails in increasing numbers is not accepted by all, however (Monahan, Caldiera, & Friedlander, 1979). For example, some investigators have proposed that it is simply heightened awareness among professionals and the public of the problem of mentally ill in the jails that has resulted in the perception that they are entering in increasing numbers (Morrisey, Steadman, & Kilburn, 1983; Steadman & Ribner, 1980). If indeed the number of mentally disordered inmates entering the jails and prisons is increasing, it may be attributed to several societal changes that have occurred since the 1960s. Though separate, each of these factors is related to the general trend in which the methods used by the state to control the mentally ill have gradually eroded.

Adopting a sociological perspective, those with mental illness may be considered deviant persons for whom society has little tolerance or concern (Goffman, 1961; Heller & Monahan, 1977; Kiesler, 1980). Persons with mental disorder have been traditionally controlled and removed from society by the state mental healthcare system. Financial and legal factors have converged, however, so that the state mental health system has lost much of its ability to treat and control persons with mental illness, as mandated by society.

The issue of inmates with mental disorders is a major concern among jail administrators and mental health professionals who work in jails. Gibbs (1983) surveyed 39 jail managers and social service providers working in jails around the nation and found that mentally disordered inmates were their second most significant concern, with jail overcrowding being the first. Jail administrators believe persons with mental illness to be at significantly increased risk for suicide, violence, and abuse at the hands of other inmates than inmates without mental disorder. Jail administrators also see inmates with mental illness as requiring more attention from jail staff and being much more likely to disrupt jail activities than their nondisordered counterparts.

In addition, it appears that correctional officers are most likely to identify inmates with psychotic disorders. This is suggested by the large percentage of inmates when referred for mental health evaluations by correctional staff (who receive diagnoses of schizophrenia or affective disorder). This is of significance in terms of training correctional officers, who are front line personnel and, therefore, are in the best position to identify inmates with mental disorders.

A number of authors have expressed a concern that depressed inmates are especially likely to go unnoticed by correctional officers and, thus, are not referred to psychologists for evaluation and treatment. Depressed persons, unfortunately, may be considered "model inmates" by correctional officers. As a function of their depression, these inmates are unlikely to make demands upon staff or become involved in altercations with other inmates. Thus, they may not get referred for evaluation. Given the relationship between depression and self-injurious behavior, however, the referral of these inmates is crucial for their well-being and the successful operation of the institution. Thus, psychologists should take special care to alert correctional officers to this potential hazard.

Regardless of whether their number is increasing, extant research indicates that there are considerable numbers of mentally ill individuals in jails and prisons. Given the history of the lack of adequate mental health services in correctional facilities, it is clear that this group has been underserved and neglected. Thus, steps should be taken to increase the availability of mental health services in jails and prisons (Ogloff, Roesch, & Hart, 1993; Steadman, McCarty, & Morrisey, 1989).

9.06.2.1 Substance Abuse

Unfortunately, more is known about the number of mentally ill in jails than those who abuse or are addicted to alcohol and other drugs. The few studies that have examined the prevalence rates of mental illness in jails generally have often ignored or paid minimal attention to the problems of incarcerated substance abusers. A number of studies examining mental disorder in jails ignore or fail to report prevalence estimates of alcohol and drug abuse addiction (e.g., Lamb & Grant, 1983).

This cursory treatment of substance abuse problems in jails may be due to two factors. Although it is a significant mental health problem, substance abuse among jail inmates is unlikely to be a major concern for those administering or working in jails. The limited access to alcohol and drugs in the jail setting makes the problem of abuse during incarceration a minor issue for caretakers. When not under the influence or withdrawing, substance abusers are not likely to pose special problems to jail staff, although the problem of individuals being placed in jail while intoxicated or under the influence of other drugs remains.

Addicted inmates might be expected to pose a significant problem to the jail operation. Those withdrawing from alcohol and other drugs undergo considerable distress and in some cases may be in a life-threatening situation. In addition to the risk and discomfort to the individual, withdrawal from alcohol and drugs poses a clear management problem in terms of its effects on other inmates. Thus, the problem of substance abuse and addiction among jail inmates is an important issue with respect to the provision of mental health services in jails.

The few studies examining the prevalence of substance abuse and addiction suggest that the problem is considerable. In its 1983 census, the Department of Justice (1984) reported that 75% of the 5785 inmates surveyed reported using substances of some kind. Twenty-six percent indicated that they were using drugs of some type, and 46% indicated that they used alcohol immediately prior to committing the offense for which they were arrested. At the very least, these figures suggest that many individuals enter confinement under the influence of alcohol or drugs, a large number of whom may be serious abusers or addicted.

Findings from more thorough studies are consistent with the above information. Of 445 jail inmates self-referred or referred by correctional staff for psychiatric evaluation, Swank and Winer (1976) reported alcoholism and drug addiction rates of 11.9 and 8.5%, respectively. In their sample of 100 randomly selected inmates who were interviewed, the authors reported alcoholism and drug addiction rates of 18.0 and 6.0%. Somewhat higher rates were reported by Guy, Platt, Zwerling, and Bullock (1985). Upon interviewing 96 randomly selected inmates from the Philadelphia County Prison (county jail), the authors reported alcohol dependency and drug dependency rates of 25.0 and 11.5%, respectively. Moreover, 28% of a larger sample surveyed by the authors expressed a self-assessed need for alcohol or drug treatment.

As with prevalence studies of mental disorder, the differences between these studies may be explained by different sampling procedures, small sample sizes, different sampling periods, geographic differences, and varying diagnostic criteria. But even if the lowest estimates of substance abuse and dependence are assumed to be accurate, it remains clear that this is a significant problem among jail inmates. Unfortunately, substance abuse treatment may be the most neglected aspect of mental healthcare in jails.

In addition to the research investigating the prevalence of mentally disordered offenders in prisons, research shows that correctional officers perceive mentally disordered offenders less favorably than other inmates, and the officers feel the need for training in the area of identifying and handling mentally disordered offenders (Kropp, Cox, Roesch, & Eaves, 1989). Finally, mentally disordered offenders have higher rates of institutional infractions and incidents of misconduct than other offenders (Adams, 1986).

9.06.2.2 Prevalence and Assessment of Women in Jails and Prisons

Prevalence and assessment of mental disorder among female inmates will be reviewed simultaneously as much of the research on prevalence has been based on intake screening and assessment procedures in jails and prisons. The majority of the empirical research on the prevalence, assessment, and treatment of mental disorder in offender populations has focused on male offenders and has virtually ignored female offenders. Teplin et al. (1996) argued that the findings from studies of men cannot be generalized to women for three reasons. First, service utilization and diagnostic profiles are different for men and women. Second, men and women are managed differently in correctional settings, and third, the resources that are available in male correctional facilities are different from those available in female facilities. Thus, it is important to review studies conducted specifically with female offenders.

In recent years there has been a demonstrable increase in the number of incarcerated female offenders, although the number of women in prisons and jails in the USA is still relatively small when compared to men. A report by the US Department of Justice (1995) indicated that in 1994 there were almost 50 000 female jail inmates, representing approximately 10% of the total jail population. The report noted that this figure has more than tripled between 1983 and 1994.

A number of studies have examined the rates of mental disorder in women who have been screened and referred for mental health services within jails or prisons. A high rate of mental disorder would be expected in these samples as these individuals have already been flagged as having possible mental health problems. These studies do not examine prevalence rates *per se* but rather only the rates of certain disorders found in a referred population of women.

Lamb and Grant (1983) examined a sample of 101 women who had been referred for psychiatric evaluation to the Forensic Mental Health Unit in Sybil Brand Institute, Los Angeles county jail for women. This mental health unit receives approximately 4% of all the women who are booked into the jail. In this referred sample, 86% of the women had a history of psychiatric hospitalization and more than half met criteria for involuntary hospitalization. Birecree, Bloom, Leverette, and Williams (1994) assessed the need for mental health care in a sample of 91 women screened and referred for clinical interview upon entry in Oregon's prison system. Over half (55%) of these women had prior contact with the mental health system.

The first study to examine the rates of mental disorder in jails using a random of the entire jail population rather than a sample of previously identified patients was conducted by Washington and Diamond (1985) in California. They screened a random sample of inmates from five county jails. There were 115 women in their sample, 41.7% of whom had a *DSM-II* diagnosis. Washington and Diamond (1985) expressed concern at the high numbers of women in need of mental health treatment and questioned whether these women should be in the mental health system rather than the criminal justice system given that jails have limited and often inadequate mental health services.

Daniel et al. (1988) administered a structured diagnostic interview to 100 women admitted to the Missouri Correctional Classification Centre to determine lifetime and six-month prevalence rates for psychiatric disorders. They reported that the offender sample of women had significantly higher rates of schizophrenia (7%), major depression (21%), substance abuse (36%), and antisocial personality disorder (29%) compared to women from the general population. Daniel et al. concluded that mental health treatment and systematic screening needs to be available for all prisoners. Ingram-Fogel (1991) interviewed 135 women during the first week of incarceration at a major correctional facility in a southern state and concluded that substance abuse and depression were major health problems among these incarcerated women and that the incidence of alcohol and drug abuse among these women was far greater than that reported in the general population.

Teplin et al. (1996) administered a structured diagnostic interview to 1274 female arrestees awaiting trial. The results indicated that over 80% met criteria for one or more lifetime psychiatric disorders. The lifetime prevalence rates for the specific disorders were as follows: 2.5% met criteria for schizophrenia, 2.6% for manic disorder, 16.9% for major depressive episode, 33.5% for post-traumatic stress disorder, 63.5% for drug abuse/dependence, and 32.3% for alcohol abuse/dependence. With the exception of schizophrenia, the prevalence rates for all these disorders were significantly higher than in the general population. Teplin, Abram, and McClelland (1997) also investigated the percentage of incarcerated women that required mental health services, the percentage that received mental health services, and the variables that predicted which females would receive these services. The results indicated that 10.7% of the participants needed mental health services and, of those, 23.5% actually received services. As far as variables that predicted who would receive services, a history of psychiatric treatment greatly increased the odds of receiving services, whether or not the individual currently required services. Women with major depression were less likely to receive services unless it was comorbid with drug abuse or dependence and then they were more likely to receive services. Those women with schizophrenia or manic disorder were 65 times more likely to receive services than those without these disorders (except when these disorders were comorbid with drug abuse or dependence and then they were less likely to receive services). Finally, severely depressed white high school graduates had higher odds of receiving services than was predicted.

As in the USA, females make up only a small proportion of the total population of prisoners in other countries. In Canada, females comprise approximately 9% of the total jail population and 2% of the total prison population (Shaw, 1994). Turner and Tofler (1986) estimated that women constitute about 3% of the total prisoner population in the UK and the results of a study by Maden, Swinton, and Gunn (1994) indicated that women make up about 3.5% of the total prisoner population in England and Wales. The prevalence rates of mental disorder among women in jails and prisons in the USA is typical of that found in other countries. The UK data suggest much lower rates, although it should be remembered that it is unclear whether the differences are real or a function of differences in the manner in which mental illness is defined and assessed.

Robertson, Bankier, and Schwartz (1987) interviewed 100 consecutive female admissions to a jail in Winnipeg, Manitoba, Canada. They reported that 12% of these women had a psychiatric history, 34% had alcohol use disorders, 6% had drug use disorders, and 4% had psychotic disorders. Hurley and Dunne (1991) screened the entire population of 92 prisoners at the Brisbane (Australia) Women's Prison for psychological distress and psychiatric morbidity. This institution receives remanded and sentenced prisoners from the courts of the entire state of Queensland and comprises about 85% of Queensland's population of women prisoners. The results indicated that high levels of psychological distress were evident in this sample of women prisoners and that 53.3% were diagnosed with a current psychiatric disorder.

Turner and Tofler (1986) screened a sample of 708 women admitted to Holloway prison in the UK. Three variables—drug abuse, self-harm, and a history of psychiatric treatment—were used as indicators of psychiatric disorder. Although this research can be criticized because of the lack of standardized instruments to assess psychiatric disorder, these authors reported a high rate of drug abuse (14%), self-harm (27.5%), psychiatric history (17.7%), and a moderate amount of overlap between these variables. Turner and Tofler concluded that psychiatric units are needed in women's prisons and that changes to policies for admission to psychiatric hospitals should be vigorously pursued. Finally, Maden et al. (1994) collected interview and file data for 25% of the population of women serving a prison sentence in England and Wales to determine the prevalence rates of psychiatric disorder. The specific prevalence rates found were: 1.6% for psychotic disorder, 18% for personality disorder, 26% for drug abuse or dependence, and 9% for alcohol abuse or dependence. Prior psychiatric contact was reported by 45% of the women.

In summary, the available research indicates that the prevalence rates of mental disorder among women in jails and prisons throughout the USA and other countries are similar. In general, it appears that the highest prevalence rates occur for alcohol and drug use disorders, depression, and personality disorders.

9.06.3 THE DIVERSION OF MENTALLY ILL OFFENDERS

A substantial number of inmates in both jails and prisons are mentally ill and, for some of these offenders, it may be appropriate to divert them out of the jail or prison system into the mental health system where they can receive the treatment they need to remain offense free. The term diversion has many applications. For example, already incarcerated offenders, or offenders awaiting sentence, may be said to be diverted from prison through various means (probation, parole, electronic monitoring, and house arrest). In addition, prisoners may be deemed mentally unfit and diverted to a psychiatric facility. Pretrial diversion programs are used as an alternative to formally processing people charged with committing offenses in the criminal justice system. Rather than proceeding to trial in the usual manner, pretrial diversion involves redirecting offenders from the courts to other agencies for disposition. It was hoped that diversion programs would result in keeping the diverted accused out of the criminal justice system, making them less susceptible to committing other offenses (see Austin & Krisberg, 1981).

The movement toward diversion in the criminal justice system began in the USA in the 1960s, and in Canada in the 1970s, as a result of the increasing number of minor cases being processed by the clogged courts. Also instrumental in the genesis of pretrial diversion programs was an increasing dissatisfaction with the criminal justice system coupled with the emerging rehabilitative philosophy (Lawrence, 1991). Initially, diversion occurred when prosecutors and courts made informal arrangements with offenders and community agencies in an effort to rehabilitate offenders in order to prevent them from recidivating. Federal commissions in the 1960s in the USA sought reorientation of the correctional system away from punishment and incarceration and toward formal community correctional models, including diversion programs (Gottheil & Ghosh, 1983). Gottheil and Ghosh noted that by 1993 the number of diversion programs in the USA had skyrocketed. Similarly, the movement toward pretrial diversion gained considerable acceptance in other countries (e.g., Australia, Canada, and the UK; see Cooke, 1992; Davis, Boucherat, & Watson, 1989; Kraus & Hasleton, 1982; LeBlanc & Beaumont, 1992; Pratt, 1986).

Underlying any diversion program is the assumption that formal contact with the criminal justice system has inherently negative effects (Roesch & Corrado, 1983). Another assumption is that offenders need treatment or some other form of intervention to transform them into stable, law-abiding, community members (Lawrence, 1991; Roesch & Corrado, 1983). Finally, it is assumed that pretrial diversion is a more effective means of reintegrating offenders into the community than incarceration (Curran, 1988).

Five specific goals of pretrial diversion programs flow from the rationale behind pretrial diversion, and the assumptions that underlie it:

(i) to reduce recidivism, thereby lowering the crime rate (Gottheil & Ghosh, 1983);

(ii) to decongest the criminal justice system, thereby improving cost-effectiveness and allowing prosecutors and judges to attend to more serious offenders who pose greater threats to the community (Gottheil & Ghosh, 1983);

(iii) to provide necessary services (e.g., therapy, counseling, job training) to individuals to better prepare them for the demands of society (Covey & Menard, 1984);

(iv) to reduce the coercive, punitive social control of the criminal justice system by removing many less-serious offenders from the system (Lipsey, Cordray, & Berger, 1981); and

(v) to avoid the negative stigma and labeling (as a criminal) that occurs in the formal criminal justice system (Decker, 1985).

Regardless of the goal of the diversion program, pretrial diversion programs have certain eligibility criteria and program requirements. Accused who successfully meet the eligibility criteria and also satisfactorily complete the program requirements have the charges against them dismissed. Referral to a diversion program is typically at the prosecutor's discretion. Eligibility for pretrial diversion programs is usually restricted to first-time offenders charged with minor offenses. Selected individuals preferably should pose little threat to the community and require little supervision (Lawrence, 1991). Roesch, (1978a) reviewed 30 diversion programs and found that all excluded people accused of serious felonies (e.g., rape, murder, armed robbery), and seven excluded individuals accused of any felony. Roesch and Corrado (1983) noted that although many diversion projects purport to accept those accused of certain, less serious felonies, in practice this is rare since programs predominantly admit those accused of misdemeanors such as shoplifting.

There also is variability in eligibility criteria regarding the defendant's prior criminal record. Some programs exclude individuals with any prior record (see Royse & Buck, 1991) while others exclude only those who have committed a felony within the last five years, or who have ever committed a major felony (see Gottheil & Ghosh, 1983). Yet others exclude only those with records of major felonies (e.g., Fische & Jeune 1987). Also, there are many programs with selection criteria that fall between the extremes presented here, with differing combinations of offense history, time since last offense, and severity of present offense factored in. Although there are many varying eligibility criteria, the majority of programs accept only those accused of minor offenses with no prior felony convictions and few, if any, misdemeanor convictions.

Psychologically treatment-based pretrial diversion programs very often involve group therapy designed to help offenders gain insight into their behavior and to develop strategies for living in the community without committing further crimes. Specific therapeutic strategies are extremely broad ranging from psychoeducational treatment, stress management, and anger management to very specific treatment approaches for specific types of offenses (e.g., assaultive husbands programs). Given that many offenders are not particularly well educated and have few job skills, diversion programs have been developed to provide offenders with practical, hands-on, skills that will enable them to become gainfully employed and to decrease their need to rely on criminal behavior as a means of existence.

The number of diversion programs increased rapidly and gained widespread acceptance by the 1970s; however, little attention was paid to systematically evaluating the effectiveness of the programs (Roesch & Corrado, 1983). As Roesch (1978a) pointed out, although pretrial diversion programs are very expensive and require major resource commitments, the agencies that provide funding for the programs very often did not require "evaluation procedures to be included, and ... did not present any guidelines, beyond requesting certain demographic information on project participants, for such evaluation" (p. 80). By 1993, however, an extensive body of literature on the effectiveness of diversion programs had been published. Most of the studies investigated whether the goals of diversion were met. In general, although there have been some successful programs, the results are equivocal and diversion is not the panacea it was initially imagined to be.

A major issue is whether diversion programs actually reduce the number of individuals processed by the criminal justice system. Critics argue that many individuals referred to diversion programs would, in the absence of diversion programs, simply have been released from custody due to the minor nature of the offense or the lack of evidence for prosecution (see Roesch & Corrado, 1983). Thus, diversion programs may function to increase, not decrease, the extent to which the state is involved in people's affairs; the number of individuals who actually come into contact with the system increases because of diversion, a concept known as "widening the net of social control."

Several studies have indeed found evidence of this "widening of the net" phenomenon (e.g., Decker, 1985; Laberge & Morin, 1995; Lipsey et al., 1981; Sarri & Bradley, 1980) and it is widely considered a common, though undesirable, result of diversion programs (Decker, 1985). A corollary to this finding is that if diversion functions to increase the number of individuals in contact with the criminal justice system, then it is also failing to provide cost-effective service. However, there have also been studies which have failed to find evidence of net widening (e.g., Fischer & Jeune, 1987), and others have argued that it is the manner in which programs are implemented that determines whether they succeed, not any inherent weakness of the concept of diversion itself (Curran, 1988; Roesch & Corrado, 1983; Roesch & Foisy, 1986).

Another contested issue is whether diversion programs succeed in reducing recidivism, and consequently the crime rate. Again, there are mixed results. Many researchers have found that diversion programs do not decrease recidivism or prevent crime (Roesch & Corrado, 1983; Severy & Whitaker, 1982), others have found that diversion does decrease recidivism (Binder & Binder, 1983; Davidson, Redner, Blakely, Mitchell, & Emshoff, 1987; Land, McCall, & Williams, 1990; Lipsey et al., 1981; Quay & Love, 1979; Royse & Buck, 1991) while others have criticized the methodology of studies (Mrad, 1979; Roesch, 1978a; Roesch

& Corrado, 1979). Moreover, Osgood (1983) found no relationship between program effectiveness and the seriousness of offenders' criminal history. Again, while it is clear that diversion has not been as successful as was initially hoped, its failure may be due to poor implementation strategies such as applying the same counseling program to individuals with different needs (Roesch & Foisy, 1986).

Diversion also aims to be less coercive than formal processing, although critics maintain that the right to due process is violated because individuals are indeed coerced into diversion programs, because if they do not participate and successfully complete the program, they will be prosecuted (O'Brien, 1984). Participants must also admit responsibility without legal proof, which has been deemed coercive. However, there have been some attempts to reduce any coercive elements of diversion. Several programs provide legal counsel to individuals before the decision to participate is made (O'Brien, 1984) while others do not require that the program be successfully completed in order for charges to be dismissed (Gottheil & Ghosh, 1983).

Diversion can only work if it is applied to people who otherwise would have been prosecuted for their offense. Roesch and Corrado (1983) argued that diversion programs should include those charged with more serious offenses, and that dismissal of charges should not be contingent upon completion of a program, but should be automatic upon the decision to divert. The authors also suggested that program participation be voluntary, and that the type of program employed should have strong theoretical support and justification.

Some mentally ill inmates could benefit from psychological intervention which, in turn, might reduce the likelihood of them committing future offenses. Therefore, diversion programs may be particularly effective with this population. As with any program, care and attention will be required to identify and divert those persons who pose a low risk to society and demonstrate some likelihood of reducing their chance of reoffending. Of course, with mentally ill inmates who rely upon psychotropic medication to control their mental illness, particular care must be taken to ensure that they will remain compliant with the medication.

9.06.4 PRETRIAL ASSESSMENT

Pretrial assessment perhaps accounts for the greatest proportion of time spent by forensic clinicians working in the criminal justice system. There are three primary areas of pretrial assessment that we will review in this section: competency to stand trial, criminal responsibility, and risk assessment.

9.06.4.1 Competency to Stand Trial

Competency to stand trial is a legal provision that allows the postponement of judicial proceedings for defendants who do not have the capacity to assist in their defense. The competency to stand trial criteria specified in state statutes generally refer to variations of the standard set forth by the US Supreme Court in *Dusky v. USA* (1960). In *Dusky*, the Court held that for a defendant to be considered competent to stand trial, the defendant must have "sufficient present ability to consult with his attorney with a reasonable degree of rational understanding [and] a rational as well as factual understanding of the proceedings against him" (p. 402).

Psychologists have long been involved in providing information to the courts to assist in decisions about competency. Competency evaluations have not been uniform, varying widely from clinician to clinician (Roesch, Ogloff, & Golding, 1993), but a number of interview guides and rating scales have been developed to aid evaluators in assessing competency to stand trial. These include the competency screening test (Lipsitt, Lelos, & McGarry, 1971) the Competency Assessment Instrument (McGarry & Curran, 1973), the Interdisciplinary Fitness Interview (Golding, Roesch, & Schreiber, 1984; McDonald, Nussbaum, & Bagby, 1991), the Fitness interview Test-Revised (FIT-R, Roesch, Zapf, Eaves, & Webster 1998), the Competence Assessment for Standing Trial for Defendants with Mental Retardation (Everington, 1990), and the Georgia Court Competency Test (Wildman, White, & Brandenburg, 1990; see also Nicholson & Johnson, 1991). More recently, the MacArthur Research Network on Mental Health and the Law has been developing assessment tools for evaluating adjudicative competency (Bonnie, 1992; Hoge et al., 1997).

One criticism of the manner in which competency is assessed is that the evaluations focus too heavily on a defendant's mental status with little attention paid to the legal issues involved in a particular case. The recommendations resulting from these evaluations are typically global in nature, usually resulting in statements offering a conclusion about whether or not a defendant is competent to stand trial. These recommendations are based largely on a finding of psychosis or mental retardation, with virtually no attempt to relate these factors to the specific requirements of the instant criminal case (see Grisso, 1986; Roesch & Golding, 1980). This is unfortunate because the legal concept of competency to stand trial is quite specific and focuses on questions surrounding a defendant's functional abilities that are directly relevant to the criminal process in the law. Thus, information derived from traditional clinical assessments indicating the defendant's IQ score, personality profile, and mental status generally, is often not relevant to the legal question of the defendant's competence.

It has been suggested elsewhere that incompetency should be evaluated and treated in a context-dependent manner, taking into account the demands of a defendant's specific legal situation and what is likely to be required of that defendant in a given legal procedure (Golding et al., 1984; Roesch, Hart, & Zapf, 1996). It is quite possible for defendants to be competent for certain types of legal proceedings but not for others (Whittemore, Ogloff, & Roesch, 1997), so it is necessary that competency evaluations and the treatment of incompetent defendants consider the specific demands and contextual factors related to a particular case.

9.06.4.1.1 Community-based screening

In the USA and Canada, there has been a tradition of remanding large numbers of individuals to inpatient facilities for assessments of competency to stand trial. Grisso, Cocozza, Steadman, Fisher, and Greer (1994) conducted a survey to determine the organization of pretrial forensic evaluation services in the USA. The results of this research indicated that the traditional approach, though the dominant approach in only a minority of states, was still widely used throughout the USA. However, a number of states have implemented models that employ various types of outpatient approaches. One approach is to use screening evaluations, defined as "brief evaluations at jail or courthouse to determine whether there was a need for a full evaluation for competence to stand trial" (Grisso et al., 1994, p. 389).

In the late 1970s, Roesch (1978b) concluded that a brief, immediate screening interview could be used to evaluate competence to stand trial and that this method resulted in a reduced cost and an increase in the protection of individual rights over the traditional method of remanding an individual to an inpatient facility. He reported a high agreement between judgments of competence based upon a brief screening interview and lengthy institutional evaluations of competence and concluded that screening interviews are an effective alternative to inpatient evaluations holds true today, and more states are adopting brief screening interviews as the preferred model of evaluation (Grisso et al., 1994). Instruments for conducting these assessments are available. For example, research in 1997 conducted using the FIT-R has indicated that this measure demonstrates excellent utility as a screening instrument and that it can reliably screen out those individuals who are clearly fit to stand trial before they are remanded to an inpatient facility for evaluation (Zapf & Roesch, 1997). Zapf and Roesch found that about four of every five defendants could have been screened out before being remanded to an inpatient facility for assessment.

9.06.4.1.2 Competency to be executed

A specific form of criminal competency may arise with prison inmates sentenced to death, and that is the issue of competency to be executed. This is an infrequently raised issue, affecting only a small minority of inmates on death row, but the assessment and treatment issues, particularly ethical dilemmas for psychologists, warrant consideration in this chapter.

The American Bar Association (1989) proposed the following definition of the legal principle, "competency to be executed":

> A convict is incompetent to be executed if, as a result of mental illness or mental retardation, the convict cannot understand the nature of the pending proceedings, what he or she was tried for, the reason for the punishment, or the nature of the punishment ... [or] lacks sufficient capacity to recognize or understand any fact which might exist which would make the punishment unjust or unlawful, or lacks the ability to convey such information to counsel or to the court. (Standard 7-5.6, p. 290)

The US Supreme Court affirmed common law prohibitions against the execution of incompetent defendants in the case of *Ford v. Wainright* (1986), holding that such executions are prohibited by the Eighth Amendment provision against cruel and unusual punishment. In essence, the Court held that prisoners sentenced to death who subsequently become incompetent have a constitutional right not to be executed. There are several rationales that form the basis of common law and state statutes that prohibit the execution of incompetent prisoners. Perhaps the most compelling rationale is that incompetent individuals cannot assist counsel in the appeal process and other challenges of the original sentence. Other reasons include the belief that the goals of deterrence or retribution cannot be achieved by executing an incompetent person, it is inhumane and cruel, and it prevents the individual from making a final religious peace (Heilbrun, 1987; Melton, Petrila, Poythress, & Slobogin, 1987; Radelet & Barnard, 1986).

The issue of competency to be executed raises a number of ethical concerns for clinicians involved in the assessment and treatment of incompetent defendants. Is it possible to conduct such evaluations given that the major implication of a finding that an offender is competent to be executed, is that he or she will be executed? Would clinicians opposed to the death penalty be more inclined to find a defendant incompetent? Is it reasonable to ask clinicians to treat these defendants given that successful treatment will result in the execution of the defendant?

Many clinicians have refused to participate in assessment or treatment, citing ethical objections (Deitchman, Kennedy, & Beckham, 1991). Appelbaum (1986) makes the distinction that other competency evaluations (e.g., competency to stand trial) are based on the administration of justice whereas competency to be executed is an aspect of the administration of punishment. Bonnie (1990) took issue with this distinction and argued that clinicians should not rely on ethical objections as a basis for refusing to participate in such evaluations. He considered whether participation in competency for execution evaluations forces the clinician to cross the boundary between participation in the administration of justice to participation in the administration of punishment, since the latter may be considered ethically objectionable (see Appelbaum, 1986, for a discussion of this issue). Bonnie contended that clinicians have traditionally been involved in many aspects of the sentencing process and other postsentence evaluations, but he recognized that clinicians may make a personal choice that their "personal scruples against the death penalty are so strong that an objective evaluation is not possible" (p. 81), and on this ground, may decide against participation in the evaluation of competency to be executed.

Bonnie did not extend his arguments about the ethics of participation in evaluations to the treatment of offenders found incompetent. This is because treatment would not be beneficial to such defendants but rather would serve the interests of the state to carry out the execution. In the case of treatment, Bonnie favored refusal based on ethical objections, and added that one way out of the dilemma would be to automatically commute the sentence of any prisoner found incompetent to be executed. Bonnie recognized that this would open up the competency process as the last line of defense against execution, making the ethical and moral dilemmas even more pronounced.

Appelbaum (1986) commented on three positions that could be taken by psychologists regarding participation in competency for execution evaluations. The first is that since participation in these evaluations is no different from participation in any other forensic evaluation, there are no valid ethical objections. The second position allows participation, although it places conditions intended to serve as protections for prisoners, such as clarifying standards for determining competency, ensuring that prisoners be allowed to contest the finding and call their own experts. The third position is that psychologists should never participate in these evaluations. This position rests on the assumption that to do otherwise would require evaluators to cross the boundary between participation in the administration of justice and the administration of punishment. While the former is ethically defensible, the latter would not be considered so.

Forensic clinicians engaging in these evaluations will find that the evaluation of competency to be executed is similar to evaluations of other forms of criminal competency. Evaluations should be based on a functional assessment of an offender's competency. If communication between an attorney and an offender is an issue, the clinician may want to observe sessions between them. Because much is at stake for the offender, there may be considerable motivation for faking incompetency, so clinicians will likely want to obtain third party information and also conduct some assessment of the likelihood that the defendant is feigning incompetence (Rogers, 1988; Small & Otto, 1991).

9.06.4.2 Criminal Responsibility

Although the issues surrounding the insanity defense are complex and difficult, psychologists working in jails and prisons sometimes are asked to make determinations concerning an inmate's mental state at the time of the offense. Generally speaking, states employ one of two standards for insanity: the M'Naghten standard or the American Law Institute (ALI) standard (Ogloff, Roberts, & Roesch, 1993a). The M'Naghten standard comes from English common law and comprises three substantive elements. First, it must be determined that the defendant was suffering from a defect of reason and/or from disease of the mind. Next, evidence must show that, as a result of the mental disease or defect, the defendant did not know the nature and quality of the act. Finally, the M'Naghten standard also requires an inquiry to determine whether the defendant knew what he or she was doing was wrong. Therefore, the defendant who has a cognitive understanding or knowledge of his or her act, but who does not demonstrate knowledge that the act was morally wrong, may also be acquitted under the M'Naghten test. The M'Naghten test is referred to as a cognitive test of insanity because of its focus on the quality of the defendant's thought processes at the time of the crime (e.g., Low, Jeffries, & Bonnie, 1986). The M'Naghten standard was criticized for focusing exclusively on a defendant's cognitive capacity (Golding & Roesch 1987; Goldstein, 1967; Ogloff, Roberts, & Roesch, 1993).

The second common standard in place in many states is the ALI standard. In drafting the new insanity defense standard, the ALI subcommittee attempted to redress the shortcomings of earlier formulations, especially the requirement of total incapacity which seemed implicit in the M'Naghten test (American Law Institute, 1962, s. 4.01). The ALI standard introduced several important modifications. First, the ALI test recognizes that impairments in cognitive and volitional functioning which are caused by mental disease or defect may provide a basis for exculpation by reason of insanity. Second, the ALI test is founded upon an assessment of substantial rather than total incapacity, a standard which was seen as comporting more adequately with both medical facts and actual jury behavior. Third, the term *appreciate* rather than *know* was employed. Appreciate connoted affective as well as cognitive understanding and was perceived as fitting more consonantly with modern mental health perspectives on the unified nature of personality. Use of the word *conform* provided continuity with the older irresistible impulse notion of impairments in capacity for volitional control.

At the federal level, Congress passed the Insanity Defense Reform Act (1984) following John Hinckley, Jr.'s acquittal by reason of insanity for the attempted assassination of President Reagan, which removed the volitional prong of the ALI rule—leaving it substantively like M'Naghten, with its focus on the defendant's cognitive appreciation for purposes of determining whether he or she should be found not guilty by reason of insanity (Perlin, 1989). The modifier *substantial* was also removed, so that the substantive federal insanity test is whether the defendant lacks capacity to appreciate the wrongfulness of his or her conduct; thus federal legislators successfully re-introduced the historically earlier absolute or total deprivation of capacity standard. In addition, the new standard specified that the mental disease or defect must be severe (Finkel, 1989, Perlin, 1989).

In addition to more traditional not guilty by reason of insanity (NGRI) standards, the guilty but mentally ill (GBMI) verdict has been adopted by some states as a supplement to the insanity defense standards (Ogloff et al., 1993a). The GBMI verdict holds defendants criminally responsible for their acts, but recognizes that the defendant is mentally ill. The GBMI plea has been termed an in-between classification because defendants are neither acquitted nor found guilty in the traditional sense (Ogloff et al., 1993). Although a review of scholarly analyses of the GBMI verdict is beyond the scope of this article, readers should be aware that the verdict has stirred a great deal of controversy (e.g., Blunt & Stock, 1985; Golding & Roesch, 1987; Roberts & Golding, 1991; Slobogin, 1983).

Ogloff et al. (1993) described a clinical assessment procedure for helping to determine whether a defendant is NGRI. Such an assessment involves an investigation of the defendant's mental state at the time of the offense. However, a retrospective evaluation is difficult even under ideal circumstance where the clinician has access to perfectly valid information about the defendant's mental state immediately prior to and during the offense. As Ogloff et al. (1993) noted,

> it is important for clinicians to be cognizant of the limitations inherent in the mental nonresponsibility evaluation—that is, the retrospective nature of the evaluation always require some degree of speculation and the necessity of placing the data in a psychological conceptual framework that is relevant to the legal issues. (p. 168)

The assessment procedure described by Ogloff, Roesch, and Hart (1993) involves five phases. In Phase 1, the clinician establishes rapport and reviews with the defendant the nature of the proposed evaluation, along with the defendant's rights during the process, the limitations on psychosocial history from multiple sources of information (e.g., by reviewing the defendant's educational, mental health, and correctional records). To explore missing data and to continue the rapport-building process, defendants should be encouraged to provide their account of their psycho-social history as well and, where relevant, standardized tests or structured interviews are employed. Phase 3, the clinician assesses the defendant's present mental state, much as would be done in a more traditional assessment. Phase 4 is the most important phase insofar as the clinician must attempt to reconstruct the cognitive, conative, emotional, motivational, and psychopathological concomitants and determinants of the defendant's behavior at the time of the offense. Phase 5, the examiner must formulate the reconstructed elements in terms of their legal relevance while taking care to avoid conclusory pronouncements of the ultimate legal issues. At all times, clinicians are constrained by their general competence to render judgments about an individual's mental state at some time in the past and must acknowledge the limitations in their evaluation.

9.06.4.3 Mental Disorder and Violence

Another form of assessment frequently conducted by psychologists working in prisons and jails is risk assessment. Most often, the professional is asked to determine whether, as a consequence of any mental disorder from which they may suffer, offenders are likely to cause harm to themselves or to others. The focus of the evaluation may be institutional, that is, risk of violence during incarceration. Common concerns here are the likelihood of self-injurious behavior, for violence in the community after release from an institution. The primary concerns here are risk for violent and sexually violent recidivism, although risk for general recidivism (e.g., new property or drug offenses) may also be an issue.

Risk assessment differs from traditional mental health assessment in several ways. Traditional mental health assessments determine the presence, nature, and severity of symptoms of mental disorder, and then suggest appropriate strategies for alleviating those symptoms via psychotherapy and/or pharmacotherapy. Risk assessments determine the presence of risk factors, including symptoms of mental disorder, that are associated with the likelihood that an individual will be involved in some kind of critical incident, and then suggest strategies for reducing or managing the risks posed by the individual. The primary goal of traditional assessments is to improve the well-being of offenders, whereas the primary goal of risk assessments is the protection of institutional security or public safety. A second difference is that risk assessment requires the evaluator to have some expertise in crime and violence. These areas are outside the usual domain of training for psychologists, with the exception of those who completed specialized graduate programs or clinical internships in forensic psychology. A third difference is that whereas traditional mental health assessments are based principally on a clinical interview with the individual, risk assessments place much greater emphasis on historical factors that can be assessed through case records or collateral informants. Thus, risk assessments can, and often must, be done without the cooperation of the person who is being evaluated.

There are three primary methods of conducting risk assessments. They can be differentiated on the basis of the structure or systematization they impose on the assessment process: gathering information, collating it, and using it to reach a judgment. The first method, commonly referred to as the clinical method, has no structure. All the parameters for the assessment of the individual are determined by the evaluator. The second method relies on professional guidelines which provide recommendations concerning what information should be obtained, as well as specific risk factors that should be considered. The third method is highly structured. The defining feature of the actuarial method is that a decision-making process is specified. Typically, assessment information is reduced to a set of numbers, which are then combined using an arithmetic algorithm to yield a decision.

Increased structure in the assessment process has both positive and negative effects. On the negative side, it decreases flexibility and limits the evaluator's ability to consider case-specific information. On the positive side, it helps to ensure a reasonably comprehensive assessment, increases the accuracy of decisions, and reduces the risk of liability. In the view of the authors the professional and research literatures are sufficiently large that evaluators should not rely solely on clinical methods to assess risk in the vast majority of cases. Professional guidelines and actuarial instruments have been developed specifically to assist in the assessment of risk for institutional self-harm (e.g., Polvi, 1997), institutional misconduct (e.g., Cooke, 1996), postrelease general recidivism (e.g., Andrews, 1983), postrelease general violence (e.g., Violence Prediction Scheme, Webster, Harris, Rice, Cormier, & Quinsey, 1994); 2HCR-20 (Webster, Eaves, Douglas, & Wintrup, 1995), and postrelease spousal and sexual assault (e.g., *Spousal Assault Risk Assessment Guide,* Kropp, Hart, Webster, & Eaves, 1995); SVR-20 (Sexual Violence Risk Guide, Boer, Wilson, Gauthier, & Hart, 1997).

9.06.5 THE TREATMENT OF MENTALLY ILL OFFENDERS IN JAILS AND PRISONS

Being arrested is typically a traumatic experience. Individuals entering jail are often in a state of crisis, with high levels of stress and anxiety being quite common. In addition, many who are arrested have a history of mental health problems. Consequently, pretrial jail facilities may be a highly appropriate place for mental health interventions (Ogloff & Roesch, 1992). A jail can provide an excellent opportunity to deliver mental health services to those who need and are interested in obtaining such services. While many jail inmates do not need or desire mental health services, it is likely that a substantial number would benefit from such services, if they were available. It is important to note that inmates retain the right to refuse to participate in treatment so, with few exceptions such as an emergency situation to prevent harm to self or others, the provision of treatment is on a voluntary basis.

Since most detainees remain in jail for only a short time, mental health interventions in jails are fairly limited in scope, but inmates may wish to continue involvement after release. For this reason, it is important to ensure that treatment services can be continued outside jail. In this way, interventions in jails can be seen as a starting point, an opportunity to evaluate individuals and provide some short-term crisis intervention if needed, and then provide follow-up services in the community. A model of treatment that is integrated with community services is more likely to be successful (Ogloff & Roesch, 1992).

Jails, however, have traditionally provided little in the way of mental health services. Teplin (1990) noted that most mental health problems go undetected in jails so treatment is unlikely. Even if detected, there is considerable variability in the availability of mental health services in jails. For example, Steadman et al. (1989) conducted a national survey of jail mental health services and found that the emphases in most were on problem identification, including screening and evaluation, and dispensing of medication. Some type of drug or alcohol service was available in the majority of programs, but psychological counseling was available in less than one-half of the jails surveyed. Size of the jail seemed to have some impact on the type of services available. Therapies were unavailable in most small jails but medium-sized jails (51–250 inmates) were more likely than large jails to employ professional staff to deliver therapy and referral services.

What mental health services should be available in jail? The first step in determining services is to identify the needs of the jail population. An effective jail mental health program must begin with a brief mental health evaluation of all inmates soon after their admission. Ideally, this would be done by a mental health professional, but this is not always feasible. In some jails, correctional officers have been trained to conduct these interviews. The evaluation focuses on immediate mental health needs and, in particular, the risk for suicide or violent behavior. Research has consistently reported that suicide is the leading cause of death among jail detainees, often occurring within the first day or two after arrest. Self-mutilation is also a common problem. Early identification and subsequent crisis or suicide intervention programs, therefore, are of paramount importance.

Cox, Landsberg, and Paravotti (1989) suggested that suicides in a jail can be reduced if correctional officers are trained to identify inmates at risk, provided that clinical services are available to these referrals. Their model for crisis intervention incorporates the following procedures to identify suicidal and seriously mentally ill inmates, and provide for their safety and timely access to appropriate levels of mental health care:

(i) policy and procedural guidelines that outline administrative and direct service actions that enable staff to identify, manage, and serve high-risk mentally ill and suicidal inmates;

(ii) suicide prevention intake screening methods that can be administered by jail officers during intake to identify high-risk inmates;

(iii) a training program for jail officers in suicide prevention; and

(iv) the use of a mental health practitioner's manual designed to provide basic information to staff providing clinical care in the jail.

It was further suggested that emergency outpatient services, inpatient services, and nonemergency mental health services are also essential components in effectively meeting the needs of suicidal and other high risk inmates. Implementation of this program in one jail resulted in a 33% decline in suicides during the first year of their study.

Although identification and treatment services are becoming increasingly available in local jails, few jail programs are coordinated with outpatient services for the continuation of treatment after release. Steadman et al. (1989) proposed a comprehensive approach where identification, treatment, and referrals for ongoing care with inpatient mental health facilities or community service agencies are essential components. The key to a well-functioning system of ongoing treatment and integration into the community is interagency cooperation between the criminal justice and mental health systems. Adequate communication between these two systems permits the essential continuity of care necessary for mentally disordered offenders to "develop appropriate coping behaviors for situational catalysts that may contribute to recidivism' (Ogloff & Otto, 1989, p. 366).

Unfortunately, such continuity of treatment and intra-agency cooperation occurs all too infrequently. Based on their jail study, Steadman et al. (1989) found that one consistent factor seemed to be associated with more effective jail/community programs, and that was the existence of a core position that was responsible for managing the interactions of mental health, jail, and judicial personnel. Steadman (1992) used the term *boundary spanners* to characterize these individuals because they were able to cross over the boundaries of the separate systems and ensure that the needed services were being provided. For those interested in establishing a viable program, it may be essential to have an individual who is responsible for ensuring that the two systems interact effectively and efficiently.

In contrast to jails, inmates in prisons typically are detained for longer periods of time. This increased detention period provides an opportunity to provide more intensive and longer-term interventions with inmates (Ogloff, Roesch, & Hart, 1994; Webster, Hucker, & Grossman, 1993). Once an inmate is assessed, he or she should be referred to an appropriate treatment program within the correctional facility. If the inmate has a serious mental illness, and there is no existing treatment program available, the law would require that the prison provide individual psychological or psychiatric treatment to the inmate (Ogloff et al., 1994). Despite the legal requirements for providing mental health services and treatment in prisons, one of the primary impediments to prison treatment is a lack of motivation among many prison inmates to participate in the treatment programs. Consequently, inmates who might otherwise benefit from treatment do not take advantage of such programs.

The most common form of treatment in prisons is psychotropic drug therapy (Bloom & Bradford, 1988; Gudjonnson, 1990). Various forms of behavior therapy, particularly cognitive behavior therapy, are among the more popular types of psychotherapy (Gudjonnson, 1990). Some prisons have developed specific treatment programs for subgroups of inmates, such as sexual offenders (Bloom & Bradford, 1988), suicidal inmates (Ramsey & Tanney, 1987), or psychopaths (Ogloff, Wong, & Greenwood, 1990; Rice, Harris, & Cormier, 1992).

There are several subpopulations in jail that could be identified for treatment. One group would comprise those who have been arrested for domestic violence. Spousal assault is especially common among offenders. For example, Dutton and Hart (1992) surveyed male prison inmates and found that nearly 30% had a history of family violence, with the most common type of violence being wife assault. In recognition of the prevalence rates, some jails and prisons are developing institution- and community-based programs for treating offenders with a history of wife assault, as well as establishing training programs to help staff better identify and manage such offenders.

One of the most comprehensive treatment programs for mentally disordered offenders was developed in the New York Correctional System (see Steadman et al., 1989). The New York City Board of Corrections proposed minimum standards for mental health services in jails and prisons with the following service goals: crisis intervention and the management of acute episodes; suicide prevention; stabilization of the mentally ill and the prevention of deterioration in the prison setting; and elective therapy services and preventive treatment where resources exist. The standards further include criteria for identification and detection, training of staff, suicide prevention, diagnosis and referral, 24 hour access to nonemergency mental health services, and inpatient mental healthcare. Treatment options in the program are based on a needs assessment that determines the type and prevalence of problems requiring treatment, such as depression, suicide, or disruptive behavior (Dvoskin & Steadman, 1989; Steadman & Holohean, 1991; Toch & Adams, 1986). This more specific, problem-oriented approach could lead to the development of more effective treatment programs.

One subpopulation that already receives considerable attention are those whose competency to stand trial is being evaluated. Some studies have found that evaluations of competency are increasing, perhaps due to the changes in mental health commitment laws noted above (Roesch & Golding, 1985). Defendants are often transferred to a forensic facility for a period of up to 30 days for an assessment. The vast majority of these referrals are found competent, a finding that is at least in part due to the fact that many defendants are referred as a means to get them out of jail and into a mental health treatment facility rather than the result of a real concern for their fitness to stand trial. Again, this supports the need for mental health services in jails in order to minimize the inappropriate use of other mechanisms for obtaining treatment.

Inmates with alcohol and substance abuse comprise another identifiable group. Abram and Teplin (1991) found considerable overlap of mental disorder and drug or alcohol abuse in jail populations. This has been referred to as co-occurring disorders. The high rate of individuals with co-occurring disorders suggests the need for treatment programs both within jail and subsequent to release. Abram and Teplin suggested that these individuals might be particularly appropriate for alternative to prosecution, such as pretrial diversion programs.

9.06.5.1 The Surrey Pretrial Mental Health Project

The authors have been involved in the development of a comprehensive mental health program in a pretrial jail (Ogloff & Roesch, 1992). This project will be described briefly as an example of the type of program that forensic clinicians might help develop in local jails. The Mental Health Project at the Surrey Pretrial Services Centre (SPSC) began when SPSC opened in 1991. The facility houses approximately 160 inmates. Services offered by the project include screening for mental illness, assessment, treatment, institutional and community case planning and case management, liaison with other correctional institutions, as well as the training and utilization of correctional staff in the identification and management of mentally disordered inmates. The participation of community mental health services and forensic outpatient services makes community linkage possible.

A key aspect of the project is the initial interview conducted by senior doctoral students in a clinical psychology program at a local university. The screening interviews last approximately 10–15 minutes each and consist of a brief semistructured mental status interview covering seven content areas: personal/demographic information, suicide risk, orientation to time and space, criminal history, social adjustment during the past month, mental status in the past month, and mental health history. Particular attention is also paid during the intake assessment to identify inmates who require crisis intervention services, and who may be at risk for self-harm or suicide.

Such inmates are reported immediately to the correctional officer in charge so that necessary intervention/precautions may be taken. Similarly, attempts are made during the screening interview to identify those individuals who may be a danger to others. In 1995, 4412 inmates (85% of all admissions) were screened, with 1377 (31%) referred to the jail mental health team for further assessment and possibly intervention in the jail and following release. Approximately one-fifth of the referred inmates were considered to have a serious mental disorder.

Because of the limited time that most inmates spend in the SPSC, there is little opportunity for ongoing individual psychotherapy. Thus, much of the mental health treatment is limited to focused treatment groups as well as limited counseling and/or psychotherapy. With its community focus, one of the most important goals of the project is to identify the need for mental health intervention for inmates and to ensure that they are linked to appropriate mental health services agencies in the community, or in another correctional facility to which they may ultimately be sentenced.

One problem in developing an integrated community-based treatment program is the manner in which treatment and other services are divided among many different agencies. Often, there are disagreements between agencies over which one is responsible for providing treatment. For example, many mentally disordered offenders have two or more *DSM* Axis I disorders, usually substance abuse along with some other diagnosis. The community mental health system may refuse to treat such individuals, arguing that they fall under the jurisdiction of the alcohol/drug programs, but the latter programs feel that these dual diagnosis cases are not appropriate for treatment by them because the mental disorder must be treated before treatment for the addiction can take place. In short, there is no coordinated treatment effort, no attempt to focus on the individual who may have a range of problems, and may require interventions from a variety of perspectives. Clearly, for a project like this to be successful, there is a need to develop an integrated service delivery system in which the individual, and not the type of problem, is seen as the primary focus.

9.06.5.2 Unique Concerns for the Treatment of Female Offenders

It has been documented that women experience a variety of adjustment problems in jail and prison (e.g., Ingram-Fogel, 1991; MacKenzie, Robinson, & Campbell, 1989; Sultan et al., 1984). There have been some studies conducted on the effectiveness of different treatment approaches to women's adjustment and coping in jails and prisons (e.g., MacKenzie et al., 1989; Sultan, Kiefe, & Long, 1986; Sultan et al., 1984). There have also been studies conducted on treatment approaches to specialized problems, such as anger (Wilfley, Rodon, & Anderson, 1986) as well as sexual and physical abuse (Sultan & Long, 1988). However, there are no published studies on the treatment of mental disorder among women in jails and prisons. Steadman et al. (1989) surveyed 43 jail mental health programs in the USA and concluded that 100% of these institutions offer medication as a means of treating mental disorders, whereas only 30% offer any form of therapy or counseling. It may be that mentally disordered female inmates are either being diverted to the mental health system from the criminal justice system or are being treated only with medications while incarcerated. Research is needed to clarify the issue of the treatment of mental disorder among women in jails and prisons.

9.06.6 CONCLUSION

The literature on the prevalence, assessment, and treatment of mentally disordered offenders in jails and prisons has been reviewed. The review of the jail literature suggests that the prevalence of mental disorder in jails is high and probably on the increase; consequently, it is believed that it is important that mental health services be made available in jails. The emphasis of these services should be on brief screening assessments, crisis intervention, and community follow-up after release from jail. Increased use of diversion program for some mentally ill pretrial defendants should also be considered. With respect to prison mental health issues, it is recognized that the population of mentally ill inmates is smaller than in the jails, but nevertheless a range of psychological services, including assessment and treatment, should be considered an integral part of prison programs. Finally, some unique issues related to women with mental health problems in prisons and jails were reviewed.

9.06.7 REFERENCES

Abram, K. M., & Teplin, L. A. (1991). Co-occurring disorders among mentally ill jail detainees: Implications for public policy. *American Psychologist, 46,* 1036–1045.

Adams, K. (1986). The disciplinary experiences of mentally disordered inmates. *Criminal Justice & Behavior, 13,* 297–316.

Allodi, F., Kedward, H., & Robertson, M. (1977). Insane but guilty: Psychiatric patients in jail. *Canada's Mental Health, 25,* 3–7.

American Bar Association (1989). *ABA criminal justice mental health standards.* Washington, DC: Author.

American Law Institute (1962). *Model Penal Code.* Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

Andrews, D. A. (1983). *The Level of Supervision Inventory (LSI): The first follow-up.* Toronto, ON: Ontario Ministry of Correctional Services.

Appelbaum, P. S. (1986). Competence to be executed: Another conundrum for mental health professionals. *Hospital and Community Psychiatry, 37,* 682–684.

Austin, J., & Krisberg, B. (1981). Wider, stronger, and different nets: The dialectics of criminal justice reform. *Journal of Research in Crime and Delinquency, 18* 165–196.

Binder, A., & Binder, V. L. (1983). Juvenile diversion. *Counseling Psychologist, 11,* 69–77.

Birecree, E. A., Bloom, J. D., Leverette, M. D., & Williams, M. (1994). Diagnostic efforts regarding women in Oregon's prison system: A preliminary report. *International Journal of Offender Therapy and Comparative Criminology, 38,* 217–230.

Bloom, J. D., & Bradford, J. M. (1988). An overview of psychiatric treatment approaches to three offender groups. *Hospital and Community Psychiatry, 39,* 151–158.

Blunt, L. W., & Stock, H. V. (1985). Guilty but mentally ill: An alternative verdict. *Behavioral Sciences and the Law, 8,* 49–67.

Boer, D. P., Wilson, R. J., Gauthier, C. M., & Hart, S. D. (1997). Assessing risk for sexual violence: Guidelines for clinical practice. In C. D. Webster & M. A. Jackson (Eds.), *Impulsivity: New directions in research and clinical practice* (pp. 326–342). New York: Guilford Press.

Bonnie, R. J. (1990). Dilemmas in administering the death penalty: Conscientious abstention, professional ethics, and the needs of the legal system. *Law and Human Behavior, 14,* 67–90.

Bonnie, R. J. (1992). The competence of criminal defendants: A theoretical reformulation. *Behavioral Sciences and the Law, 10,* 291–316.

Cooke, D. J. (1992). Treatment as an alternative to prosecution: Offenders diverted for treatment. *British Journal of Psychiatry, 158,* 785–791.

Cooke, D. J. (1996). Predicting offending in prison: The predictive validity of the Prison Behavior Rating Scales. *Legal and Criminological Psychology, 65–82.*

Covey, H. C., & Menard, S. (1984). Community corrections diversion in Colorado. *Journal of Criminal Justice, 12,* 1–10.

Cox, J. F., Landsberg, G., & Paravotti, M. P. (1989). The essential components of a crisis intervention program for local jails: The New York local forensic suicide prevention crisis service model. *Psychiatric Quarterly, 60,* 103–117.

Curran, D. J. (1988). Destructuring, privatization, and the promise of juvenile diversion: Compromising community-based corrections. *Crime and Delinquency, 34,* 363–378.

Daniel, A. E., Robins, A. J., Reid, J. C., & Wilfley, D. E. (1988). Lifetime and six-month prevalence of psychiatric disorders among sentenced female offenders. *Bulletin of the American Academy of Psychiatry and Law, 16,* 333–342.

Davidson, W. S., Redner, R., Blakely, C. H., Mitchell, C. M., & Emshoff, J. G. (1987). Diversion of juvenile offenders: An experimental comparison. *Journal of Consulting and Clinical Psychology, 55,* 68–75.

Davis, G., Boucherat, J., & Watson, D. (1989). Pre-Court decision-making in juvenile justice. *British Journal of Criminology, 29,* 219–235.

Decker, S. H. (1985). A systemic analysis of diversion: Net widening and beyond. *Journal of Criminal Justice, 13,* 207–216.

Deitchman, M. A., Kennedy, W. A., & Beckham, J. C. (1990). Self-selection factors in the participation of mental health professionals in competency for execution evaluations. Law and Human Behavior, 15, 287–303.

Department of Justice (1984). *1983 Census.* Ottawa, ON: Statistics Canada.

Dusky v. United States, 362 U.S. 402 (1960).

Dutton, D. G., & Hart, S. D. (1992). Risk markers for family violence in a federally incarcerated population. *International Journal of Law and Psychiatry, 15,* 101–112.

Dvoskin, J. A., & Steadman, H. J. (1989). Chronically mentally ill inmates. The wrong concept for the right services. *International Journal of Law and Psychiatry, 12,* 203–210.

Everington, C. T. (1990). The Competence Assessment for Standing Trial for Defendants with Mental Retardation (CAST-MR): A validation study. *Criminal Justice and Behavior, 17,* 147–168.

Finkel, N. J. (1989). The Insanity Defense Reform Act of 1984: Much ado about nothing. *Behavioral Sciences and the Law, 7,* 403–419.

Fischer, D. G., & Jeune, R. (1987). Juvenile diversion: A process analysis. *Canadian Psychology, 28,* 60–70.

Ford v. Wainright, 477 U.S. 399 (1986).

Gibbs, J. J. (1983). Problems and priorities: Perceptions of jail custodians and social service providers. *Journal of Criminal Justice, 11,* 327–338.

Goffman, E. (1961). *Asylums.* Garden City, NY: Doubleday.

Golding, S. L., & Roesch, R. (1987). The assessment of criminal responsibility: An historical approach to a current controversy. In I. Weiner & A. Hess (Eds.), *Handbook of forensic psychology* (pp. 396–436). New York: Wiley.

Golding, S. L., & Roesch, R., & Schreiber, J. (1984). Assessment and conceptualization of competency to stand trial: Preliminary data on the Interdisciplinary Fitness Interview. *Law and Human Behavior, 8,* 321–334.

Goldstein, A. (1967). *The insanity defense.* New Haven, CT: Yale University Press.

Gottheil, D., & Ghosh, S. A. (1983). Pretrial intervention: Counseling in adult diversion programs. *Counseling Psychologist, 11,* 17–25.

Grisso, T. (1986). *Evaluating competencies: Forensic assessments and instruments.* New York: Plenum.

Grisso, T., Cocozza, J. J., Steadman, H. J., Fisher, W. H., & Greer, A. (1994). Organization of pretrial forensic evaluation services: A national profile. *Law and Human Behavior, 18,* 377–393.

Gudjonnson, G. H. (1990). Psychological treatment for the mentally ill offender. *Issues in Criminal Legal Psychology, 16,* 15–21.

Guy, E., Platt, J. J., Zwerling, I. & Bullock, S. (1985). Mental health status of prisoners in and urban jail. *Criminal Justice and Behavior, 12,* 29–53.

Heilbrun, K. S. (1987). The assessment of competency for execution: An overview. *Behavioral Sciences and the Law, 5,* 383–396.

Heller, K., & Monahan, J. (1977). *Psychology and community change.* Homewood, IL: Dorsey.

Hodgins, S., & Cote, G. (1990). The prevalence of mental disorders among penitentiary inmates. *Canada's Mental Health, 38,* 1–5.

Hoge, S. K., Bonnie, R. J., Poythress, N., Monahan, J., Eisenberg, M., & Feucht-Haviar, T. (1997). The MacArthur Adjudicative Competence Study: Development and validation of a research instrument. *Law and Human Behavior, 21,* 141–179.

Hurley, W., & Dunne, M. P. (1991). Psychological distress and psychiatric morbidity in women prisoners. *Australian and New Zealand Journal of Psychiatry, 25,* 461–470.

Insanity Defense Reform Act, 18 U.S.C. s. 20 (1986 Supp.).

Ingram-Fogel, C. (1991). Health problems and needs of incarcerated women. *Journal of Prison and Jail Health, 10,* 43–57.

Ives, G. (1914). *A history of penal methods.* London: S. Paul.

Kiesler, C. A. (1980). Mental hospitals and alternative care: Noninstitutionalization as potential public policy for mental patients. *American Psychologist, 35,* 349–360.

Kraus, J., & Hasleton, S. (1982). Juvenile offenders' diversion potential as a function of police perceptions. *American Journal of Community Psychology, 10,* 171–182.

Kropp, P. R., Cox, D. N., Roesch, R., & Eaves, D. (1989). The perceptions of correctional officers toward mentally disordered offenders. *International Journal of Law and Psychiatry, 12,* 181–188.

Kropp, P. R., Hart, S. D., Webster, C. W., & Eaves, D. (1995). *Manual for the Spousal Assault Risk Assessment Guide* (2nd ed.). Vancouver, BC: British Columbia Institute on Family Violence.

Laberge, D., & Morin, D. (1995). The overuse of criminal justice dispositions: Failure of diversionary policies in the management of mental health problems. *International Journal of Law and Psychiatry, 18,* 389–414.

Lamb, H. R., & Grant, R. W. (1983). Mentally ill women in a county jail. *Archives of General Psychiatry, 40,* 363–368.

Land, K. C., McCall, P. L., & Williams, J. R. (1990). Something that works in juvenile justice: An evaluation of the North Carolina Court Counselors' Intensive Protective Supervision Randomized Experimental Project, 1987–1989. *Evaluation Review, 14,* 574–606.

Lawrence, R. (1991). Reexamining community corrections models. *Crime and Delinquency, 37,* 449–464.

LeBlanc, M., & Beaumont, H. (1992). The effectiveness of juvenile justice in Quebec: A natural experiment in implementing formal diversion and a justice model. In R. R. Corrado, N. Bala, R. Linden, & M. LeBlanc (Eds.), *Juvenile justice in Canada: A theoretical and analytical assessment* (pp. 283–312). Toronto, ON: Butterworths.

Lipsey, M. W., Cordray, D. S., & Berger, D. E. (1981). Evaluation of a juvenile diversion program: Using multiple lines of evidence. *Evaluation Review, 5,* 283–306.

Lipsitt, P. D., Lelos, D., & McGarry, A. L. (1971). Competency for trial: A screening instrument. *American Journal of Psychiatry, 128,* 105–109.

Low, P. W., Jeffries, J. C., & Bonnie, R. J. (1986). *The trial of John W. Hinckley, Jr: A case study in the insanity defense.* Mineola, NY: The Foundation Press.

MacKenzie, D. L., Robinson, J. W., & Campbell, C. S. (1989). Long-term incarnation of female offenders: Prison coping and adjustment. *Criminal Justice and Behavior, 16,* 223–238.

Maden, T., Swinton, M., & Gunn, J. (1994). Psychiatric disorder in women serving a prison sentence. *British Journal of Psychiatry, 164,* 44–54.

McDonald, A., Nussbaum, D., & Bagby, R. M. (1991). Reliability, validity, and utility of the Fitness Interview Test. *Canadian Journal of Psychiatry, 36,* 1–5.

McGarry, A. L., & Curran, W. J. (1973). *Competency to stand trial and mental illness.* Rockville, MD: National Institute of Mental Health.

Melton, G., Petrila, J., Poythress, N., & Slobogin, C. (1987). *Psychological evaluation for the courts: A handbook for mental health professionals and lawyers.* New York: Guilford.

Monahan, J., Caldiera, C., & Friedlander, H. D. (1979). Police and the mentally ill: A comparison of committed and arrested persons. *International Journal of Law and Psychiatry, 2,* 509–518.

Monahan, J., & Geis, G. (1976). Controlling "dangerous" people. *Annals of the American Academy of Political and Social Science, 423,* 142–151.

Monahan, J., & Steadman, H. J. (1983). *Mentally disordered offenders: Perspectives from law and social science.* New York: Plenum.

Morrissey, J., Steadman, H. J., & Kilburn, H. (1983). Organizational issues in the delivery of jail mental health services. In J. Greenley (Ed.), *Social problems and public policy* (pp. 291–317). Greenwich, CT: JAI Press.

Mrad, D. F. (1979). The effect of differential follow-up on re-arrests: A critique of Quay and Love. *Criminal Justice and Behavior, 6,* 23–30.

Neighbors, H. (1987). The prevalence of mental disorder in Michigan prisons. *DIS Newsletter, 4,* 8–10.

Nicholson, R. A., & Johnson, W. G. (1991). Prediction of competency to stand trial: Contribution of demographics, type of offense, clinical characteristics, and psycholegal ability. *International Journal of Law and Psychiatry, 14,* 287–297.

Nielsen, E. D. (1979). Community mental health services in the community jail. *Community Mental Health Journal, 15,* 27–32.

O'Brien, D. (1984). Juvenile diversion: An issues perspective from the Atlantic Provinces. *Canadian Journal of Criminology, 26,* 217–230.

Ogloff, J. R. P., & Otto, R. K. (1989). Mental health intervention in jails. In P. A. Keller & S. R. Heyman (Eds.), *Innovations in clinical practice: A source book* (Vol. 8, pp. 357–369). Sarasota, FL: Professional Resources.

Ogloff, J. R. P., Roberts, C. F., & Roesch, R. (1993). The insanity defense: Legal standards and clinical assessment. *Applied and Preventive Psychology, 2,* 163–178.

Ogloff, J. R. P., & Roesch, R. (1992). Using community mental health centers to provide comprehensive mental health services to jails. In J. R. P. Ogloff (Ed.), *Psychology and law: The broadening of the discipline* (pp. 241–260). Durham, NC: Carolina Academic Press.

Ogloff, J. R. P., Roesch, R., & Hart, S. D. (1993). Screening, assessment, and identification of services for mentally disordered offenders in prisons. In H. J. Steadman & J. J. Cocozza (Eds.), *Providing services for offenders with mental illness and related disorders in prison* (pp. 61–90). Seattle, WA: National Coalition for the Mentally Ill in the Criminal Justice System.

Ogloff, J. R. P., Roesch, R., & Hart, S. D. (1994). Mental health services in jails and prisons: Legal, clinical, and policy issues. *Law and Psychology Review, 18,* 109–136.

Ogloff, J. R. P., Tien, G., Roesch, R., & Eaves, D. (1991). A model for the provision of jail mental health services: An integrative, community-based approach. *Journal of Mental Health Administration, 18,* 209–222.

Ogloff, J. R. P., Wong, S., & Greenwood, A. (1990). Treating criminal psychopaths in a therapeutic community program. *Behavioral Sciences and Law, 8,* 181–190.

Osgood, D. (1983). Offense history and juvenile diversion. *Evaluation Review, 7,* 793–806.

Perlin, M. L. (1989). *Mental disability law: Civil and criminal.* Charlottesville, VA: The Mitchie Company.

Petrich, J. (1976). Rate of psychiatric morbidity in a metropolitan county jail population. *American Journal of Psychiatry, 133,* 1439–1444.

Pogrebin, M., & Poole, E. (1987). Deinstitutionalization and increased arrest rates among the mentally ill. *Journal of Psychiatry and Law, 15,* 117–127.

Polvi, N. H. (1997). Assessing risk of suicide in correctional settings. In C. D. Webster & M. A. Jackson (Eds.), *Impulsivity: New directions in research and clinical practice* (pp. 278–301). New York: Guilford Press.

Pratt, J. (1986). Diversion from the juvenile court: A history of inflation and a critique of progress. *British Journal of Criminology, 26* 212–232.

Quay, H. C., & Love, C. T. (1979). Effects of a juvenile diversion program on rearrests: A reply to Mrad. *Criminal Justice and Behavior, 6,* 31–33.

Rabkin, J. (1979). Criminal behavior of discharged mental patients: A critical appraisal of the research. *Psychological Bulletin, 86,* 1–27.

Radelet, M. L., & Barnard, G. W. (1986). Ethics and the psychiatric determination of competency to be executed. *Bulletin of the American Academy of Psychiatry and Law, 14,* 37–53.

Ramsey, R. F., & Tanney, B. L. (1987). Suicide prevention in high-risk prison populations. *Canadian Journal of Criminology, 29,* 295–307.

Rice, M. E., Harris, G. T., & Cormier, C. A. (1992). An evaluation of a maximum security therapeutic community for psychopaths and other mentally disordered offenders. *Law and Human Behavior, 16,* 399–412.

Roberts, C. F., & Golding, S. L. (1991). The social construction of criminal responsibility and insanity. *Law and Human Behavior, 15,* 349–376.

Robertson, R. G., Bankier, R. G., & Schwartz, L. (1987). The female offender: A Canadian study. *Canadian Journal of Psychiatry, 32,* 749–755.

Roesch, R. (1978a). Does adult diversion work? The failure of research in criminal justice. *Crime and Delinquency, 24,* 72–80.

Roesch, R. (1978b). A brief, immediate screening interview to determine competency to stand trial: A feasibility study. *Criminal Justice and Behavior, 5,* 241–248.

Roesch, R. (1995). Mental health interventions in pretrial jails. In G. M. Davies, S. Lloyd-Bostock, M. McMurran, & C. Wilson (Eds.), *Psychology, law and criminal justice: International developments in research and practice* (pp. 520–531). Berlin, Germany: De Greuter.

Roesch, R., & Corrado, R. R. (1979). The policy implications of evaluation research: Some issues raised by the Fishman study of rehabilitation and diversion services. *The Journal of Criminal Law & Criminology, 70,* 530–541.

Roesch, R., & Corrado, R. R. (1983). Criminal justice system interventions. In E. Seidman (Ed.), *Handbook of social intervention* (pp. 385–407). Beverly Hills, CA: Sage.

Roesch, R., & Foisy, D. J. (1986). Evaluating criminal justice interventions: The myth of reform. In E. Seidman & J. Rappaport (Eds.), *Redefining social problems* (pp. 201–216). New York: Plenum.

Roesch, R., & Golding, S. L. (1980). *Competency to stand trial.* Urbana, IL: University of Illinois Press.

Roesch, R., & Golding, S. L. (1985). The impact of deinstitutionalization. In D. P. Farrington & J. Gunn (Eds.), *Current research in forensic psychiatry and psychology* (pp. 209–239). New York: Wiley.

Roesch, R., & Hart, S. D., & Zapf, P. A. (1996). Conceptualizing and assessing competency to stand trial: Implications and applications of the MacArthur Treatment Competence Model. *Psychology, Public Policy, and Law, 2,* 96–113.

Roesch, R., Ogloff, J. R. P., & Golding, S. L. (1993). Competency to stand trial: Legal and clinical issues. *Applied and Preventive Psychology, 2,* 43–51.

Roesch, R., Zapf, P. A., Eaves, D., & Webster, C. D. (1998). *The Fitness Interview Test.* Unpublished manuscript, Simon Fraser University, Burnaby, BC, Canada. (Available from Ronald Roesch, Department of Psychology, Simon Fraser University, Burnaby, BC, V5A 1S6 Canada).

Rogers, R. (1988). *Clinical assessment of malingering and deception.* New York: Guilford Press.

Roth, L. (1980). Correctional psychiatry. In W. Curran, A. McCory, & C. Petty (Eds.), *Modern legal medicine, psychiatry, and forensic science* (pp. 667–719). Philadelphia: Davis.

Royse, D., & Buck, S. A. (1991). Evaluation a diversion program for first-time shoplifters. *Journal of Offender Rehabilitation, 17,* 147–158.

Sarri, R., & Bradley, P. (1980). Juvenile aid panels: An alternative to juvenile court processing in South Australia. *Crime and Delinquency, 26,* 42–62.

Schuckit, M., Herrman, G., & Schuckit, J. (1977). The importance of psychiatric illness in newly arrested prisoners. *Journal of Nervous and Mental Disease, 165,* 118–125.

Severy, L. J, & Whitaker, M. J. (1982). Juvenile diversion: An experimental analysis of effectiveness. *Evaluation Review, 6,* 753–774.

Shaw, M. (1994). Women in prison: A literature review. *Forum on Corrections Research 6,* 13–18.

Slobogin, C. (1983). The guilty but mentally ill verdict: An idea whose time should not have come. *George Washington Law Review, 53,* 494–527.

Small, M. A., & Otto, R. K. (1991). Evaluations of competency to be executed: Legal contours and implications for assessment. *Criminal Justice and Behavior, 18,* 146–158.

Steadman, H. J. (1992). Boundary spanners: A key component for the effective interactions of the justice and mental health systems. *Law and Human Behavior, 16,* 75–87.

Steadman, H. J., Fabisiak, M. A., Dvoskin, J., & Holohean, E. J. (1987). A survey of mental disability among state prison inmates. *Hospital and Community Psychiatry, 10,* 1086–1090.

Steadman, H. J., & Holohean, E. J. (1991). Estimating mental health needs and service utilization among prison inmates. *Bulletin of the Academy of Psychiatry and Law, 19,* 297–307.

Steadman, H. J., McCarty, D. W., & Morrissey, J. P. (1989). *The mentally ill in jail: Planning for essential services*. New York: Guilford.

Steadman, H. J., & Monahan, J. (1984). *Crime and mental disorder*. Washington, DC: US Department of Justice.

Steadman, H. J., & Ribner, S. (1980). Changing perceptions of the mental health needs of inmates in local jails. *American Journal of Psychiatry, 137,* 1115–1116.

Sultan, F. E., Kiefer, A. A., & Long, G. T. (1986). The application of psychodidactic support groups to facilitate female offenders' adjustment to prison. *Psychotherapy, 23, 469–471.*

Sultan, F. E., & Long, G. T. (1988). Treatment of the sexually/physically abused female inmate: Evaluation of an intensive short-term intervention program. *Journal of Offender Counseling, Services, and Rehabilitation, 12,* 131–143.

Sultan, F. E., Long, G. T., Kiefer, A. A., Schrum, D. M., Selby, J. W., & Calhoun, L. G. (1984). The female offender's adjustment to prison life: A comparison of psychodidactic and traditional supportive approaches to treatment. *Journal of Offender Counseling, Services, and Rehabilitation, 9,* 49–56.

Swank, G. E., & Winer, D. (1976). Occurrence of psychiatric disorder in a county jail population. *American Journal of Psychiatry, 133,* 1331–1333.

Teplin, L. A. (1983). The criminalization of the mentally ill: Speculation in search of data. *Psychological Bulletin, 94,* 54–67.

Teplin, L. A. (1984). Criminalizing mental disorder: The comparative arrest rate of the mentally ill. *American Psychologist, 39,* 794–803.

Teplin, L. A. (1990). Detecting disorder: The treatment of mental illness among jail detainees. *Journal of Consulting and Clinical Psychology, 58,* 233–236.

Teplin, L. A. (1991). The criminalization hypothesis: Myth, misnomer, or management strategy. In S. A. Shah & B. D. Sales (Eds.), *Law and mental health: Major developments and research needs* (pp. 149–183). Rockville, MD: US Department of Health and Human Services.

Teplin, L. A., Abram, K. M., McClelland, G. M. (1996). Prevalence of psychiatric disorders among incarcerated women, I: Pretrial jail detainees. *Archives of General Psychiatry, 53,* 505–512.

Teplin, L. A., Abram, K. M., & McClelland, G. M. (1997). Mentally disordered women in jail: Who receives services? *American Journal of Public Health, 87,* 604–609.

Teplin, L. A., & Swartz, J. (1989). Screening for severe mental disorder in jails: The development of the Referral Decision Scale. *Law and Human Behavior, 13,* 1–18.

Toch, H., & Adams, K. (1988). Pathology and disruptiveness among prison inmates. *Journal of Research in Crime and Delinquency, 23,* 7–21.

Torrey, E. F., Stieber, J., Ezekiel, J., Wolfe, S. M., Sharfstein, J., Noble, J. H., & Flynn, L. M. (1992). Criminializing the seriously mentally ill: The abuse of jails as mental hospitals. Washington, DC: Public Citizen's Health Research Group and the National Alliance for the Mentally Ill.

Turner, T. H., & Tofler, D. S. (1986). Indicators of psychiatric disorder among women admitted to prison. *British Medical Journal, 292,* 651–653.

US Department of Justice (1995). *Jails and jail inmates 1993–94*. Washington, DC: Bureau of Justice Statistics.

Washington, P., & Diamond, R. J. (1985). Prevalence of mental illness among women incarcerated in five California county jails. *Research in Community and Mental Health, 5,* 33–41.

Webster, C. D., Eaves, D., Douglas, K., & Wintrup, A. (1995). *The HCR-20 scheme: The assessment of dangerousness and risk*. Burnaby, BC: Mental Health, Law, and Policy Institute, Simon Fraser University and Forensic Psychiatric Services Commission of British Columbia.

Webster, C. D., Harris, G., Rice, M. E., Cormier, C., & Quinsey, V. L. (1994). *The violence prediction scheme: Assessing dangerousness in high risk men*. Toronto, ON: Centre of Criminology, University of Toronto.

Webster, C. D., Hucker, S. J., & Grossman, M. G. (1993). Treatment programmes for mentally ill offenders. In K. Howells & C. R. Hollin (Eds.), *Clinical approaches to the mentally disordered offender* (pp. 87–109). Chichester, UK: Wiley.

Whittemore, K. E., Ogloff. J. R. P., & Roesch, R. (1997). An investigation of competency to participate in legal proceedings in Canada. *Canadian Journal of Psychiatry, 42,* 869–875.

Wildman, R. W., II, White, P. A., & Brandenburg, C. A. (1990). The Georgia Court Competency Test: The baserate problem. *Perceptual and Motor Skills, 70,* 1055–1058.

Wilfley, D. E., Rodon, C. J., & Anderson, W. P. (1986). Angry women offenders: Case study of a group. *International Journal of Offender Therapy and Comparative Criminology, 30,* 41–51.

Zapf, P. A., & Roesch, R. (1997). Assessing fitness to stand trial: A comparison of institution-based evaluations and a brief screening interview. *Canadian Journal of Community Mental Health, 16,* 53–66.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.07
Violent and Chronic Offenders

ISAAC K. WOOD
Virginia Commonwealth University, Richmond, VA, USA

9.07.1 INTRODUCTION 105
9.07.2 PHENOMENOLOGY 106
9.07.3 EPIDEMIOLOGY AND ETIOLOGY 108
9.07.4 INTERVENTION 113
9.07.5 CLINICAL APPLICATION 114
9.07.5.1 Cognitive Problem-solving Skills Training 115
9.07.5.2 Parent Management Training 115
9.07.5.3 Functional Family Therapy 115
9.07.5.4 Multisystemic Therapy 115
9.07.6 SUMMARY 116
9.07.7 REFERENCES 116

9.07.1 INTRODUCTION

The crime rate in the United States, particularly among youthful offenders, is growing at an astronomical rate. Of most concern is the increase in the occurrence of youthful violent crime. From 1992 to 1993, the total juvenile arrest rate increased by 5% and juvenile arrests for violent crimes increased 6%. Arrests for weapons violations by youthful offenders rose by 12%, and the arrest rate for homicide increased 14%. At the same time, adult changes in these categories were negligible (Howell, 1995).

Evidence continues to mount that a small proportion of offenders commit most of the serious and violent crimes. It is this small group that captures the public interest in making decisions about policies related to interventions with criminals as a whole. Furthermore, it is known that the onset of serious violent careers begins to increase at age 12, doubles between ages 13 and 14, peaks at ages 16 to 17, drops 50% by age 18, and continues to decrease through age 27. More than half of all violent offenders initiate their violence between ages 14 and 17 and their careers last only one year, with only 4% with a career of five years or more (Elliot, 1994; Elliot, Huizinga, & Morse, 1986). In fact, aggression is a very stable behavioral trait and involvement in violent behavior at an early age is the best predictor of a youth becoming a chronic violent offender. Therefore, to understand violent and chronic offenders, an analysis of this small group of youth must be undertaken.

The purpose of this chapter is to explore the characteristics of children and adolescents who become violent and chronic offenders. It is hoped that by understanding the antecedents to aggressive behavior, a more rational, comprehensive, and exhaustive plan of intervention may be developed. To best discern the attributes of youthful violent and chronic offenders, it is necessary to explore the various systems that contribute to the environments of these juveniles. This will require an analysis of the systems most intimate and internal to the child, namely the biology of chronic, violent offenders, and working outward to the most external spheres of influence on these youth.

9.07.2 PHENOMENOLOGY

There is no uniform definition of a violent and chronic offender. In fact the term "violent and chronic offender" is highly controversial and political. According to the Office of Juvenile Justice and Delinquency Prevention (Howell, 1995), states and local jurisdictions are responsible for establishing the criteria for designating an individual as a chronic and violent offender. Many factors, therefore, influence the specific criteria used to identify this group of criminals. For example, if a community has a particularly low threshold of tolerance for violent crime, the types of aggressive behaviors that would lead to the use of this label will be quite broad. However, in large metropolitan areas, the criminal justice system is often overwhelmed by the numbers of offenders. Designating a group as "chronic and violent" usually carries with it the expectation that these individuals will be dealt with in the most restrictive settings. The most restrictive settings are also usually the most expensive. Therefore, in systems with limited financial resources and a high rate of crime, the criteria for a violent and chronic offender are likely to be quite limited.

Regardless, there does seem to be some universal agreement that certain offenses are considered violent. For example, murder, manslaughter, and rape or sexual assault are usually included under the designation of "violent offenses." Armed robbery, malicious wounding, and aggravated assault may also be included. However, beyond this, there is a great deal of variation. The chronicity of offenses may be determined based on police contacts or convictions. A conservative definition of criminal chronicity would be conviction of a least three felonies or three serious misdemeanors.

In many ways, these designations are arbitrary and meaningless. Legal definitions do little to explain the trajectory of development of criminal behavior or the most effective methods of intervention. To arrive at this information, clinical definitions related to violent and chronic offenders must be explored. The vast majority of violent and chronic offenders first exhibit these behaviors in late childhood and early adolescence (Loeber & Stouthamer-Loeber, 1998). Therefore, it would seem prudent to begin an exploration of clinical definitions by looking at designations for juvenile offenders.

The *Diagnostic and statistical manual of mental disorders* (*DSM-III;* American Psychiatric Association, 1980) has no designation for "juvenile offenders." However, in place of this, there is a category for youth who demonstrate a basic disregard for rules, law, and authority. This would be the diagnosis of "conduct disorder." According to *DSM-III* and *DSM-IV* (American Psychiatric Association, 1994), a youth under the age of 18 may be diagnosed as having a conduct disorder if he has demonstrated "a repetitive and persistent pattern of behavior in which the basic rights of others or major age-appropriate societal rules or norms are violated" (p. 85). Three such behaviors must have been present for the last 12 months with at least one being present in the last six months, and the disturbance must significantly affect social, academic, or occupational functioning. The behaviors that may lead to a diagnosis of conduct disorder are grouped under four designations:

- (i) Aggression to people and animals:
  - (a) often bullies, threatens, or intimidates others,
  - (b) often initiates physical fights,
  - (c) has used a weapon that can cause serious physical harm to others (e.g., a bat, brick, broken bottle, knife, gun),
  - (d) has been physically cruel to people,
  - (e) has been physically cruel to animals,
  - (f) has stolen while confronting a victim (e.g., mugging, purse snatching, extortion, armed robbery), or
  - (g) has forced someone into sexual activity.
- (ii) Destruction of property:
  - (a) has deliberately engaged in fire setting with the intention of causing serious damage, or
  - (b) has deliberately destroyed others' property (other than by fire setting).
- (iii) Deceitfulness or theft:
  - (a) has broken into someone else's house, building, or car,
  - (b) often lies to obtain goods or favors or to avoid obligations (i.e., "cons" others), or
  - (c) has stolen items of nontrivial value without confronting a victim (e.g., shoplifting, but without breaking and entering; forgery).
- (iv) Serious violations of rules:
  - (a) often stays out at night despite parental prohibitions, beginning before age 13 years,
  - (b) has run away from home overnight at least twice while living in parental or parental surrogate home (or once without returning for a lengthy period), or
  - (c) is often truant from school, beginning before age 13 years.

Based on these criteria, a youth could stay out past curfew, skip school, and shoplift, and be diagnosed with a conduct disorder. By the same token, a juvenile could commit rape, armed robbery, and assault with a deadly weapon, and also be diagnosed with a conduct disorder. Just from the standpoint of the levels of aggression involved, common sense suggests that these two youth are most likely very different from each other in regard to the development of these behaviors and contributing circumstances.

In an attempt to better understand youth with a conduct disorder, several attempts have been made to subclassify these juveniles. Historically, these classification systems are based on descriptive criteria which imply little about the underlying etiology or natural course of the disorder. In *DSM-III*, conduct disordered youth are placed in four categories based upon two dimensions of behavior; the presence or absence of aggression and whether behaviors are socialized or undersocialized. The designation of "aggressive" or "nonaggressive" is self-evident. An aggressive act would be one in which there is direct or indirect confrontation of the victim with harm to the victim. This would include physical and sexual assaults, rape, and murder. Nonaggressive acts would include those offenses whereby there is either no direct contact with the victim (e.g., shoplifting, breaking and entering), there is contact with the victim but no harm is done, or those crimes that are considered "victimless." This would include criminal offenses and status offenses. Criminal offenses would be any act that is considered illegal regardless of the age of the perpetrator. Whether you are 14, 44, or 84 years old, assault, destruction of property, and vandalism are considered crimes. Status offenses would be any act that is considered illegal for a minor but not so for an adult. Examples of status offenses would include under age use of alcohol, truancy, staying out past curfew, and running away from home.

The concepts of "socialized" vs. "undersocialized" need explanation. The "socialized" types of conduct disordered youth "show evidence of social attachment to others, but may be similarly callous or manipulative toward persons to whom they are not attached and lack guilt when these "outsiders' are made to suffer." In contrast, the "undersocialized" types of conduct disordered youths are "characterized by a failure to establish a normal degree of affection, empathy, or bond with others. Peer relationships are generally lacking, although the youngster may have superficial relationships with other youngsters. Characteristically, the child does not extend himself or herself for others unless there is an obvious immediate advantage. Egocentrism is shown by readiness to manipulate others for favors without any effort to reciprocate. There is generally a lack of concern for the feelings, wishes, and well-being of others, as shown by callous behavior. Appropriate feelings of guilt or remorse are generally absent. Such a child may readily inform on his or her companions and try to place blame on them" (American Psychiatric Association, 1980, p. 45).

With this two-dimensional scheme, four subcategories of conduct disorder are designated. A youth may be socialized and aggressive. This suggests that the youth's disruptive behavior tends to be violent in nature and usually occurs within the context of a group of similar juveniles to whom the offender expresses loyalty. In colloquial terms, these youth are often considered members of "gangs." By the same token, a juvenile may be undersocialized and aggressive. This means that the juvenile's conduct disordered behaviors are violent, and that the youth is a loner who fails to demonstrate remorse for his or her behavior or empathy or affection towards others. In the vernacular, these offenders are often referred to as "sociopaths." A third designation is socialized and nonaggressive. These youth exhibit nonviolent behavior consisting primarily of rules violations, status offenses, and minor illegal activities such as property crimes within the context of a group. The overwhelming majority of youth who meet the criteria for a conduct disorder fall into this category. In fact, it is estimated that over 90% of adults have engaged in similar, delinquent behaviors at some point during adolescence. Therefore, the severity and chronicity of these behaviors, plus the degree to which the behaviors impact on the functioning of the youth and his or her family seem to be the most critical factors in this group as far as designating the presence or absence of pathology. Last, a conduct disordered youth may be labeled as nonaggressive and undersocialized. This category of youth are practically nonexistent as it seems that juveniles who are undersocialized are invariably violent in their conduct disordered behaviors. In addition, with youth who meet these behavioral criteria, a more appropriate diagnosis suggesting either pathology in the schizophrenia spectrum of illnesses or Cluster A personality disorders (i.e., paranoid, schizoid, and schizotypal) often is more appropriate in describing the underlying disorder.

DSM-III-R (American Psychiatric Association, 1987) attempts to eradicate the problems of the previous nomenclature by suggesting three new subcategories for conduct disorder. The essential feature for youth designated as having a "group type" of conduct disorder is a "predominance of conduct problems occurring mainly as a group activity with peers. Aggressive physical behavior may or may not be present" (p. 56). This tends to encompass the prior categories of socialized conduct disorder, whether the behavior be aggressive or nonaggressive. A juvenile described as having a "solitary aggressive type" of conduct disorder demonstrates a "predominance of aggressive physical behavior, usually toward both adults and peers, initiated by the person (not as a group activity)" (p. 56). This, in essence, is a mere renaming of the "undersocialized, aggressive" conduct disorder. Last, a youth may be designated as having an "undifferentiated" conduct disorder, which represents the designation for those juveniles who do not fit into one of the other two categories.

DSM-III-R nomenclature presents a new set of problems. First, the clustering of youth under the description of "group type" clearly creates an extremely heterogenous association. Included in this designation would be the extremes of violent gang members who may commit heinous crimes to those youth who fraternize with antisocial or counterculture peers and engage in status offenses, but who are not serious offenders. Such a radically heterogenous grouping presents insurmountable obstacles to research about the etiology of such behaviors. Second, this classification runs counter to epidemiological evidence suggesting a developmental schema for conduct disorder related to the age of onset rather than clustering based on group or nongroup affiliation.

In order to address these problems, *DSM-IV* presents a new designation for the subtypes of conduct disorder based on age of onset. If the disruptive behaviors start before the age of 10 years, the child is given the diagnosis of "childhood-onset conduct disorder." Signs of the disorder beginning at the age of 10 or thereafter are designated as "adolescent-onset conduct disorder." This subclassification suggests that there are differences in the developmental trajectory for conduct disorder related to age of onset. It also lends itself well to an understanding of serious and violent offenders as chronic; habitual, violent offenders are greatly over-represented in the childhood-onset group.

The vast majority of serious and violent offenders demonstrate the onset of aggressive behavior at a very early age with a decrement in violence in late adolescence and early adulthood (Elliot et al., 1986). Therefore, a focus on the epidemiology and etiology of the subtypes of conduct disorder that best encompasses the characteristics of serious and violent offenders is most prudent. Based on *DSM-III*, *DSM-III-R*, and *DSM-IV*, this includes a review of undersocialized-aggressive, solitary aggressive, and childhood-onset conduct disorders.

9.07.3 EPIDEMIOLOGY AND ETIOLOGY

Considering that multiple systems are involved in the development of serious and violent offenders, one approach to understanding these youth is to start with the most internal structures, namely the psychobiology of these juveniles, and then work outward to the most external structures.

Is it possible that a child may be "born bad?" Are some individuals genetically predisposed to be violent? In the animal literature, two types of aggressive behavior are described. "Predatory aggression" involves such activities as stalking other animals for a kill, the violence that ensues when one animal infringes upon the marked territory of another, or the posturing and attacks exhibited by nursing mothers when any animal, even the mate, approaches the nest. In contrast, "affective aggression" refers to the violent response of an animal to an emotional state. Affective aggression is characterized by "autonomic arousal, threatening or defensive postures and vocalizations, and, often, frenzied attacks" (Flynn, Vanegas, & Foote, 1970). *In vitro* measures of affective aggression are obtained by shocking the floor of a mouse or rat's cage while at the same time introducing a foreign object, such as a brush, a pencil, or another animal. In response to the shock, the animal will viciously attack the foreign object. This is more clearly the classic fight-or-flight response.

The animal literature clearly indicates that mammals may be bred to have higher and lower thresholds for both predatory and affective aggression. This results in a broad spectrum of behaviors to include animals who will attack with minimal stimulation to those who freeze and cower even in the presence of a clearly threatening situation (Petitto, Lysle, Gariety, & Lewis, 1994). Furthermore, it is clear that environmental circumstances may alter the degree and types of violence exhibited. For example, rodents, who at the age of weaning are raised in solitary isolation exhibit much higher levels of predatory and affective aggression than their litter mates who are raised in a group (Schneider-Stock & Epplen, 1995). Are there similar findings in humans?

Quite clearly, humans demonstrate both predatory and affective aggression. Predatory aggression is most evident in the violent activities of gang members who often mark their territories by spray painting buildings, fences, and walls in lieu of using urine or other scents deployed by "lower" mammals. Also, humans demonstrate affective aggression with outbursts of rage in response to frustration or the lack of immediate gratification of needs.

However, is there a genetic basis for such behaviors in man?

The simplest genetic study that may be done on aggression and conduct disorder is to examine the incidence and prevalence of these behaviors in boys and girls. Aggressive conduct disorder is overwhelmingly and universally manifested in boys at a significantly higher rate than in girls (Hinshaw, Lahey, & Hart, 1993). This suggests some type of genetic vulnerability to conduct disorder in boys that is not evident in girls, highly differential environmental experiences based on gender, or a significant influence of hormones.

To further explore the contribution of genetics, it seems expedient to examine data regarding the incidence of conduct disorder in the first-degree relatives of children who are chronically violent. Family studies of children with conduct disorder demonstrate a significantly higher rate of this diagnosis in siblings, particularly brothers, and a higher rate of antisocial personality disorder in parents, fathers much more so than mothers. In addition, there is an extremely high prevalence of attention deficit hyperactivity disorder, alcoholism, aggressive behavior, and criminality in fathers and brothers, and depression, somatization disorder, and substance abuse in mothers and sisters of male probands with an aggressive conduct disorder (August & Stewart, 1983; Biederman, Munir, & Knee, 1987; Hamdan-Allen, Stewart, & Breghly, 1989; Lahey et al., 1987). This suggests a clustering of these diagnoses within families of boys who have an aggressive, early-onset conduct disorder. Yet, these data still do not clearly answer the question of the relative contribution of genetic and environmental factors. This is best addressed by examining twin and adoption studies.

Twin studies involve the use of monozygotic and dizygotic twin pairs. Monozygotic twins are formed from the splitting of one ovum that has been fertilized by one sperm. Subsequently, these twins share 100% of their genetic material. Dizygotic twins result from two ova being fertilized by two separate sperm. The genetic material of these twins is no more related than that of one of the twins and a brother or sister and would be estimated as 50%. Thus, if a disorder is genetically based, when it appears in one twin, the concordance rate (i.e., the likelihood of the disorder occurring in the other twin) should be significantly higher in monozygotic than dizygotic sibling pairs. Several studies have substantiated this finding.

The relative concordance rate for criminality in monozygotic twins is 51%, whereas the rate in dizygotic twins is 22% (Christiansen, 1976; Cloniger & Gottesman, 1987; Dalgaard & Kringlen, 1976). This again suggests that there is a genetic contribution to the development of this disorder. However, it does not rule out the effect of environment as it has been argued that monozygotic twins, by virtue of being identical, are much more likely to elicit the same reactions from caretakers than dizygotic twins. Therefore, it would be anticipated that the environmental experiences of monozygotic twins would be more homogenous than that of dizygotics. To assess this issue, it becomes necessary to examine adoption studies.

The adopted-away offspring of criminal fathers are twice as likely to be antisocial as the adopted-away offspring of noncriminal fathers. Furthermore, the highest criminality rate has been noted in adoptees of criminal fathers who are adopted away to antisocial fathers (Mednick, Gabrielli, & Hutchings, 1984). Studies of Swedish adoptees who are criminals reveal a much higher rate of criminality in their biological fathers than in adoptees who were not antisocial. Given this information, what is inherited that could lead to conduct disordered behavior? Studies on the neurobiological basis of conduct disorder provide some clues.

There are a number of studies on the role of the catecholamines in conduct disorder. Those involving the enzyme dopamine-beta-hydroxylase (DBH) are of most interest considering the evidence to suggest a genetic basis or predisposition for this disorder. DBH is the enzyme responsible for converting dopamine to norepinephrine in the catecholamine synthetic pathway. The activity of this enzyme is almost entirely determined by genetic factors.

A number of studies demonstrate that boys with an undersocialized-aggressive conduct disorder have significantly lower DBH activity than girls with this type of conduct disorder, children of either gender who have a socialized conduct disorder, and unaffected children. Furthermore, the vast majority of boys with an undersocialized-aggressive conduct disorder have near-zero activity of the enzyme. The anticipated increase in dopamine and decrease in norepinephrine are not found (Rogeness, Jovros, Maas, & Macedo, 1990). This is understandable to a certain degree considering the biochemistry of the catecholamine cascade. All of the catecholamines (i.e., dopamine, norepinephrine, and epinephrine) are formed from the amino acid tyrosine. The rate-limiting step in the formation of the catecholamines is the conversion of tyrosine to dihydroxyphenylalanine (DOPA) by the enzyme tyrosine hydroxylase. The activity of tyrosine hydroxylase is determined by three factors: the amount of the precursor, tyrosine, that is available for metabolism; the availability of a cofactor, tetrahydrobiopterin; and feedback inhibition by dopamine. Therefore, when a certain amount of dopamine has been formed, it will shut down further conversion of the precursors to this catecholamine. As a result, it would not be anticipated that there would be excessive dopamine if there is low or near-zero activity of DBH.

The failure to find a decrease in norepinephrine is more difficult to explain except it is known that in most individuals, sufficient stores of norepinephrine are synthesized regardless of the activity of DBH. Regardless, this information suggests that the low or near-zero activity of DBH in undersocialized, aggressive, conduct disordered boys does not explain the etiology of this disorder. Rather, it is more likely that low DBH activity is a genetic marker for the type of conduct disorder which yields serious and violent offenders. This also makes sense in that for all of the psychiatric disorders for which a strong genetic predilection is being determined, multiple genetic factors, that is polygenetic inheritance, is being found in association with them rather than a single genetic locus.

Several studies also indicate that boys with an undersocialized-aggressive conduct disorder have lower levels of autonomic arousal as per cardiovascular measures and Galvanic skin conductance than unaffected boys or boys with a socialized conduct disorder (Zahn & Kruesi, 1993). The reasons for this underarousal are unknown. In the normal population, studies have indicated the presence of a "sensation seeking gene" in a limited number of individuals. These individuals do not seem to get the normal autonomic or endorphin rush that most people experience from usual events such as a scary movie, a ride on a roller coaster, or a highly emotionally charged situation. In order to feel aroused, individuals with the sensation seeking gene seem to require more extreme forms of stimulation such as bungee jumping or parachuting. The presence or absence of the sensation seeking gene has not been studied in chronic, violently offending youth.

From a theoretical standpoint, it has been postulated that autonomic underarousal may be the basis for the development of a defective conscience. In the typical developing toddler and preschooler, activation of the sympathetic nervous system serves as an internal stimulus signaling the fight or flight response. It is accompanied by heightened emotional arousal and physical changes which are normally associated with anxiety including an increase in heart and respiratory rate and blood pressure, pupillary dilatation, a shunting of the blood away from internal structures such as the gut and to the muscles, and a general state of "disease" or "displeasure." In the typical preschooler and toddler, this autonomic response serves to teach the youth about dangers in the environment. When a harmful situation is encountered, the unpleasant sympathetic arousal allows for learning to avoid similar circumstances in the future. As parents intervene to teach young children about what is dangerous or safe and what is "right" or "wrong," again, arousal of the sympathetic nervous system is called upon to serve as the reinforcer for these ideas. Eventually, the goal is to have a healthy conscience which will elicit sympathetic arousal and displeasure at the thought of doing something wrong before the act even commences.

As boys with an undersocialed, aggressive conduct disorder do not experience the normal types of autonomic, physical reactions to anxiety-provoking situations, it is hypothesized that they do not adequately internalize norms related to what is "right" or "wrong." Furthermore, this lack of activation of the autonomic nervous system seems to play an integral role in the baseline emotional states of these boys and their subsequent behaviors (Quay, 1993). The majority of the boys, particularly once they are adolescents, when asked to describe their mood will respond with "bored." It is not unusual to hear them recount an act of criminal behavior as being exciting with an accompanying increase in emotional arousal. This often explains the crimes these youth engage in which to the outside observer seems pointless, such as when one or several of these youth rob a home just to give or throw away the articles that are stolen.

The monoamine serotonin is also implicated as a correlate of aggression. Studies of peripheral and central serotonin activity demonstrate correlations between levels of serotonin and aggression, with the most violent youth having the lowest levels of serotonin. These results have been reinforced by studies examining the platelet binding of imipramine which indirectly reflects serotonin levels. Again, in aggressive, conduct disordered boys, there is decreased binding suggesting low peripheral concentrations of the neurotransmitter serotonin. Within the central nervous system, serotonin is found in high concentrations in the raphe nucleus within the brainstem. From there, serotonergic neurons emanate to all areas of the cortex, diencephalon, brainstem, cerebellum, and spinal cord where they are known to play a neuromodulatory role. Specifically, serotonin tends to play a role as a behavioral suppressant. It is known that adequate levels of serotonin serve an integral role in the regulation of mood, prevention of unnecessary anxiety, and in maintaining baseline attention (as opposed to phasic shifts of attention in response to internal or external stimuli) (Bear, Connors, & Paradiso, 1996). Regardless, current studies suggest a genetic predilection to the development of a childhood-onset conduct disorder, and accompanying abnormalities in neurotransmitter levels.

Social epidemiologic data on children who will become chronic, violent offenders suggest several defining characteristics. First, these children typically come from poor, urban centers characterized by overcrowded, insufficient public housing. This is not to say that chronic, violent offenders do not exist in the rural environment. In fact, there are data to support the appearance of such behaviors, often in the context of a gang with racial or religious ties, in many rural environments. However, the typical youth with a childhood-onset conduct disorder will evolve within a poor, urban center. Prenatal care, if there is any, is complicated by improper nutrition and substance abuse including the abuse of crack cocaine and alcohol. As would be expected, there is an increased risk of perinatal difficulties in these youth including premature birth, low birth weight for gestational age, and low APGAR scores at the time of birth (Kazdin, 1995).

Families of origin tend to be large relative to the general population. As there are more individuals of ethnic minority represented in the poor, urban population, it is not unusual to see that youth who will be labeled as having a childhood-onset conduct disorder will also be from the racially under-represented parts of the population. Here, the debate on nature vs. nurture becomes heated. Overall, the consensus seems to suggest that these life circumstances are integral in the development of conduct disorder as opposed to there being a higher incidence of chronic and violent behavior in children from ethnic minorities (Kazdin, 1995).

There are often different birth fathers for the children born to the mothers of these children, and a lack of availability of a paternal figure for emotional or financial support of the mother or children. Mothers of these children often begin child bearing at an early age and the maternal and paternal grandparents, grandmothers in particular, are often the nucleus of these families. The propensity towards violence and substance abuse in the fathers of these youth results in a high rate of domestic violence in the families of chronic and violent offenders (Kazdin, 1995).

There have been variable results of studies about the temperaments of these babies. Chess and Thomas (1991) described three basic types of temperaments in infants. Easy babies tend to be extremely flexible in response to the internal and external environments. They settle and accommodate very readily to new experiences. On the opposite extreme, difficult babies are overly responsive to change in their worlds. They are frequently described as colicky, cranky babies that are not easily soothed and are overwhelmed with minimal fluctuations in the internal and external environment. In between are the slow-to-warm-up infants. These babies often respond like difficult infants during novel situations, but settle much more readily in adjustment to the change, appearing like their counterparts, the easy babies. Some studies have suggested that youth who will evolve into chronic, violent offenders are more likely to have a difficult temperament, however, this has been disputed by other researchers (Lewis, 1991).

In summary, youth who tend to be chronic and violent offenders are most often boys who demonstrate a genetic propensity to these types of behaviors. With this, there are abnormalities in neurochemical functioning associated with an increase in aggressive behavior. The intrauterine environment is often subjected to toxins such as illicit drugs and the absence of appropriate nutrition in the mothers. Perinatal insults are common and further compromise the integrity of the developing nervous system. These infants are raised in poor, urban, single parent families where domestic violence is common. There is some evidence to suggest that these infants tends to have a difficult temperament. Their mothers are ill-prepared to nurture them. There are usually multiple siblings and the absence of a father figure to provide support. The mothers tend to be burdened by problems with substance abuse, depression, and somatization disorder. The foundation for appropriate development is absent.

As preschoolers and toddlers, these youth typically present with extremes of oppositionality and defiance, far beyond what would be considered as a normal component of development (Achenbach, 1993). Limited skills at parenting, previous experiences of abuse, and the presence of pathology in the parents often lead to neglect and corporal punishment as primary means of dealing with behavior problems in these children. Physical abuse is rampant. The head is the most common part of the body struck, usually with the hand, fist, or a blunt object. Although a history of loss of consciousness secondary to this is unusual, the impact of repeated trauma on the developing central nervous system must be considered (Wood, 1996).

From an early age, these boys exhibit aggressive behavior. They tend to push, hit, bite, scratch, etc. in preschool environments. In fact, the appearance and persistence of aggressive behavior is the most distinguishing characteristic of this population of youth (Loeber & Stouthamer-Loeber, 1998). It is not unusual to get a history of dismissal from preschools as a result of aggressive behavior. By the same token, these boys are often referred to their primary care physician or a mental health professional at an early age due to these difficulties.

Typically, these boys begin to exhibit problems with inattention, distractibility, impulsivity, and hyperactivity at a very early age. However, it is not unusual that these go unnoticed or untreated until the child is in a public school setting. There, the inability to sit still, participate in group activities, follow the rules, and play cooperatively will usually result in the child being referred for medical/psychiatric evaluation.

A number of studies have demonstrated the appearance of attention deficit hyperactivity disorder (ADHD) as a comorbid condition in these boys. Studies have indicated that three groups of boys may be identified with conduct disorder and ADHD. The first two groups have only conduct disorder or ADHD, but not both. There is no increase in aggressive or chronic, violent behaviors in either of these groups. A third group has comorbid conduct disorder and ADHD (Biederman, Newcorn, & Sprich, 1990). This group has been identified as having the neurochemical disorders delineated earlier in the chapter. They are also the group that develops into the childhood-onset conduct disorder with eventual evolution into the chronic and violent offender.

In elementary school, these boys are often started on stimulant medication to assist them with behavioral control. This will often lead to an improvement in the ADHD, but no real change in the conduct disorder symptoms (Lewis, 1991). The only difference here may be in terms of youth who tend to lash out aggressively in response to the environment. However, there is little change in predatory aggression. School is further compounded by specific developmental disorders, namely problems related to learning reading and math (Moffitt, 1993). As a result, these youth are identified early on as requiring special education services. The combination of ADHD, conduct disorder, learning problems, baseline sense of boredom, and lack of respect for rules and authority make these boys a "teacher's nightmare." They are difficult to keep in the classroom, easily distracted and bored, resistant to the teacher's authority, and perpetually involved in physical struggles with their peers. Academic failure and school drop out are common.

Overall, these boys tend to score in the low average to borderline range on tests of intelligence. Of interest is the difference in scores on the verbal and performance sections of intelligence tests (Moffitt, 1993). The verbal part of the intelligence test tends to measure "crystallized intelligence," that is, what a person has learned, primarily through school. The performance section of the intelligence test measures "fluid intelligence" or the ability of an individual to think through a novel situation and arrive at an appropriate conclusion. Typically, these boys score well within the normal range in tests of fluid intelligence, but fall below the average on the verbal section (Pennington & Bennetto, 1993). This is often attributed to poor school attendance (truancy is common), learning disabilities that interfere with the means of traditionally acquiring scholastic material, and cultural bias.

Although these boys usually do not come to the attention of the law at an early age, it is common to find that actual criminal behavior begins as a preschooler or in early elementary school. During this time, these boys will frequently report that they began shoplifting or engaged in other petty offenses that are illegal. As they grow older, the frequency and the severity of offenses worsens. Often there will be a progression from "victimless" crimes to those involving an actual human target. Shoplifting and petty offenses give way to destruction of property, breaking and entering, and assault with eventual evolution into the most violent types of infractions (Wood, 1996).

Concomitant with the onset of adolescence is usually the beginnings of sexual promiscuity and substance abuse. Overall, there seems to be little regard for the human condition or the rights of others. The cessation of boredom and the instant gratification of needs seems to be the basic rule for all behaviors. Although youth with a childhood-onset conduct disorder only make up about 3–5% of all minors with a conduct disorder, it is estimated that they are responsible for well over half of the violent offenses committed by youthful offenders (Elliot, Huizinga, & Ageton, 1985). Concomitantly, a disproportionate amount of public funding is spent on these youth from the time they are very young. This includes funding for educational and mental health services, and later, for legal interventions including detainment.

The likelihood of violent behavior occurring has also been shown to be related to additional environmental factors. The availability of firearms, and community laws and norms favorable toward drug use, firearms and crime have been associated with a higher incidence of violence.

Media portrayal of violence, low neighborhood attachment and community organization, and favorable parental attitudes and involvement in problematic behaviors negatively influences these youth (Hawkins, & Catalano, 1995; Howell, 1995).

9.07.4 INTERVENTION

The Coordinating Council on Juvenile Justice and Delinquency Prevention of the Office of Juvenile Justice has developed a national juvenile justice action plan to combat violence and delinquency. The plan consists of eight objectives aimed at a combative effort to address violent and chronic juvenile offenders. The following is a summary of the objectives derived directly from Combating Violence and Delinquency: The National Juvenile Justice Action Plan (1996).

(i) *Provide immediate intervention and appropriate sanctions and treatment for delinquent juveniles.* This is seen as a two-step measure. First, those factors most prevalent in communities which are associated with violent and delinquent behavior must be identified in order to establish effective programs to prevent or control delinquent behavior. "Risk factors, such as community disorganization, availability of drugs and firearms, and persistent poverty, make children more prone to involvement in delinquent behavior that if those factors were not present. Additionally, when a child's family life is filled with violence, problem behaviors, poor parental monitoring, and inconsistent disciplinary practices or maltreatment, a child's risk of delinquency increases" (Combating Violence and Delinquency, 1996, p. 7). Second, identifying programs that help youth, families, and communities protect themselves from these risk are paramount. "The most effective programs are those that address key areas of risk in the youth's life, those that seek to strengthen the personal and institutional factors that contribute to healthy adolescent development, those that provide adequate support and supervision, and those that offer youth long-term stake in the community" (Howell, 1995, p. 25).

Graduated sanctions that are swift, certain, and consistent are thought to be most effective as a deterrent. It is suggested that immediate intervention for first-time delinquent offenders (misdemeanors and nonviolent felonies) and many nonviolent repeat offenders be offered. Intermediate sanctions should be provided for many first-time serious and repeat offenders and some violent offenders. Finally, secure corrections should be the alternative for many serious, violent, and chronic offenders.

(ii) *Prosecute certain serious, violent, and chronic juvenile offenders in criminal court.* It is important to note that the vast majority of juvenile cases can be appropriately handled in the juvenile justice system. However, a slow trend since the late 1980s to remove more serious, violent, and chronic juvenile offenders from the juvenile justice system and turn them over to criminal courts has escalated. It is believed that this is a very effective deterrent to violent crimes and also a method to assure the safety of the public in dealing with this habitually offending youth who seem to have little regard for rules, law, and authority (Hunziker, 1995).

(iii) *Reduce youth involvement with guns, drugs, and gangs.* A trend analysis of juvenile homicide offenses shows that since the mid-1970s, the number of homicides in which no firearm was involved has remained fairly constant. However, homicides by juveniles involving a firearm have increased nearly threefold (Snyder & Sickmund, 1995). Drug activity appears to exacerbate juvenile violence in two ways. First, firearms are more prevalent around drug activity. Second, drugs impact the young person's perceptions allowing for an accentuation of the kind of "high" many of these youth experience during the commission of a crime. In regard to gangs, this is a significant connection among gang involvement, gang violence, and firearms. Researchers have identified a number of factors that put youth at risk for gang involvement: poverty, school failure, substance abuse, family dysfunction, and domestic and societal violence (Spergel, 1995). Easy access to illicit drugs and the perceived financial rewards of drug dealing pose attractive alternatives for youth with inadequate education and limited educational opportunities, leading them into high-risk behaviors and potential gang involvement.

(iv) *Provide opportunities for children and youth.* Providing children with opportunities to develop positive behaviors is the foundation of most efforts to prevent youth crime and violence. In recent years, the capacity of America's low-income rural and urban communities to provide critical positive activities or environments has declined. Public schools in many areas have deteriorated, and the quality of public education has been compromised. City parks and recreation centers are in disrepair, and financial support for youth facilities and programs has decreased, leaving high-risk environments for youth. Some youth who experience child abuse, neglect, poverty, poor health, or other risk factors do not become juvenile delinquents, school drop outs, or teenage parents. These youth have the benefit of a combination of protective factors that help guide them in making healthy choices. A resilient temperament and the development of close relationships with parents and other role models who provide encouragement, healthy beliefs, and clear standards of behavior offer protection from negative environmental influences (Hawkins & Catalano, 1993).

(v) *Break the cycle of violence by addressing youth victimization, abuse, and neglect.* Each year, abuse and neglect leave 1800 children permanently—often severely—disabled and inflict lifelong psychological damage on thousands of victims, siblings, and family members. Each year an estimated 2000 children—most under age 4—die at the hands of parents or caretakers, about five children per day (Combating Violence and Delinquency, 1996). Numerous studies demonstrate the connection between child abuse and neglect and later violent delinquent behavior (Thornberry, 1994). A National Institute of Justice funded study found that experiencing childhood abuse and neglect increases the likelihood of arrest as a juvenile by 50%, of arrest as an adult by 38%, and of committing a violent crime by 38% (Combating Violence and Delinquency, 1996). The system of child protective services and dependency courts, which addresses family violence and child abuse and neglect, has reached a crisis point. It is imperative that adequate resources be provided to such agencies to allow for effective and immediate intervention in cases of abuse and neglect with one goal being the cessation of the inherited cycle of violence.

(vi) *Strengthen and mobilize communities.* Communities play the primary role in preventing juvenile delinquency and the criminal victimization of juveniles. All community members—business leaders, media representatives, teachers, parents and grandparents, young people, policymakers, clergy, elected officials, and law enforcement—are responsible for ensuring the health and well-being of children. Community planning teams that include a partnership of agency and lay participants can help create a consensus of priorities and services to be provided.

(vii) *Support the development of innovative approaches to research and evaluation.* A solid knowledge base is necessary to establish effective public policy, laws, and intervention strategies to prevent and reduce juvenile crime and violence. Focused long-term research improves our understanding of the causes and correlates of juvenile crime and strengthen our ability to develop successful prevention, early intervention, and graduated sanction programs.

(viii) *Implement an aggressive public outreach campaign on effective strategies to combat juvenile violence.* Researchers have verified that long-term public education campaigns on violence prevention, family education, alcohol and other drug prevention, and gun safety curriculums in school are effective strategies to help prevent delinquency (DeJong, 1994). They also concur that involving youth and developing community consensus are essential to an effective public education process. A well-designed public education campaign can make a positive impact on public opinion, target specific audiences, and be a cost-effective way of providing critical information to a large number of people.

In summary, the provision of services to this population of youth remains multifaceted, impacting all of the spheres that influence these children and their families. The National Juvenile Justice Action Plan for combating violence and delinquency is to be commended for being comprehensive. However, it must also be commended in that it underscores the need to address the foundations of the problems (i.e., families, communities, schools, etc.) in addition to examining the specific needs of these problematic youth. Finally, it is important to note that the mental health concerns of these children and their parents must be recognized and adequately treated. Providing education, community support, and classes on improved parenting often miss the mark in a family plagued by poverty, substance abuse, and major psychopathology.

9.07.5 CLINICAL APPLICATION

Kazdin (1997) has reviewed and identified the various psychosocial approaches that are considered to be promising treatments involved in the intervention with conduct disordered youth. In order for a treatment to be considered as promising, four criteria must have been met. First, there must be clear conceptualization of theoretical basis for the treatment. In other words, there should be some rationale about how the dysfunction leading to conduct disorder develops and then how the intervention will alter that dysfunction. Second, the treatment conceptualization must be supported by basic research. In this regard, basic research refers to the collection of objective data that identify the factors associated with conduct disorder. Third, there should be preliminary outcome evidence indicating that the treatment was effective. Finally, there needs to be a process–outcome connection, namely there is a relationship established between the processes hypothesized to be necessary for therapeutic change and the actual objective evidence of a difference in clinical outcome. Based on these criteria, four treatments were identified as being promising.

9.07.5.1 Cognitive Problem-solving Skills Training

It has been demonstrated that youth who engage in conduct disordered behaviors often have a distorted or deficient perception of the world around them. This is not to say that their view of reality is skewed. However, they manifest errors of thinking that influence how they react and respond to their environments. This cognitive error results in these youth processing information incorrectly. For example, they may see a malevolent basis for an individual's behavior when nothing of this nature was intended. Disorders of cognitive processing may make it difficult for them to problem solve from an empathetic standpoint or be unable to view realistically the consequences of their behaviors. These thinking errors also serve as a means of justifying underlying antisocial behavior. For example, the woman who got her purse stolen was "asking for it" because she left her purse out, unattended. Therefore, she is at fault for the crime.

Problem-solving skills training involves the development of interpersonal cognitive problem-solving skills. By virtue of this training, children are taught novel ways to approach situations by exploring the thinking processes involved. Step-by-step approaches to problem-solving are emphasized with reinforcement of prosocial responses. The treatment is rendered through structured tasks such as games, stories, and groups with increasing emphasis on real-life situations. Modeling, role-playing, and reinforcement with mild punishment are the media for examining problem-solving skills.

There is evidence to suggest that older children fare better with this treatment than younger children, most likely due to their advanced abilities to cognitively process. Children who come from less disturbed family situations and who lack comorbid diagnoses, learning problems, and school delays benefit the least from this treatment modality. Furthermore, there is objective research which indicates that up to a year after treatment, children who receive problem-solving skills training continue to retain the benefits.

9.07.5.2 Parent Management Training

This involves procedures geared toward teaching parents to manage their child's behavior in the home. A therapist meets with the parents to teach them techniques to increase prosocial and decrease delinquent behaviors. There is also emphasis on helping parents to recognize when they may be inadvertently reinforcing negative behaviors in their child. The overall purpose of parent management training is to change the cyclical pattern of interaction between the youth and parent. Over the source of sessions, parents are taught increasingly more advanced and complicated techniques which build on the successful implementation of more basic interventions.

Outcome of parent management training has been well evaluated and these techniques are noted to be extremely effective overall. There is also evidence for generalization of the success of the intervention to the siblings of the identified, conduct disordered youth. It appears that this approach is most successful if carried out over a protracted course of 50–60 hours. Fewest gains are evidences in those families with the highest numbers of risk factors associated with these behavioral disorders (e.g., poor, overcrowded, parental psychopathology). Another advantage of this technique is the ready availability of training manuals that systematically approach how to work with the parents.

9.07.5.3 Functional Family Therapy

This therapeutic technique views the family as a system and the behavior of individual members, such as a youth that displays conduct disordered behaviors, as serving a means for some interpersonal functions (e.g., enmeshment, distancing). The goal of treatment then becomes to alter the way in which family members interact such that the cycle of behavior is permanently altered and interpersonal needs are met in such a way as to eliminate problematic behavior. Communication patterns and their different meaning to family members are explored from a social-learning perspective. The primary focus is within the treatment sessions, where family communication patterns are altered directly.

There have been few objection outcome studies of functional family therapy. These studies have illustrated the superiority of functional family therapy over psychodynamically-oriented family therapy and no treatment control groups. Gains have been noted for up to two and a half years after the completion of treatment. However, a major drawback to this technique is the requirement of a certain level of sophistication on behalf of the family and willingness to objectively examine the comments made by the therapist.

9.07.5.4 Multisystemic Therapy

There are multiple systems involved in the development of conduct disordered behaviors. One of the faults of the previous techniques described is that they tend to focus on only one aspect of the complicated systems impacting the child (e.g., the child's thinking processes, the ability of the parents to effectively intervene). Multisystemic therapy is a family systems approach which recognizes that the child is influenced by a multiplicity of systems (e.g., immediate and extended family, peers, schools, neighborhoods). Therefore, in order to affect change in the child, the role of these various systems in creating and reinforcing the maladaptive behavior must be understood and altered. For example, in an impoverished neighborhood where the majority of peers for a youth are delinquent, emphasis may need to be placed on development of aspects of the neighborhood that would encourage association with more prosocial peer groups. Therapy is rendered within a family preservation model of service delivery (Henggeler, Schoenwald, & Pickrel, 1995). Services are delivered in the natural environment (e.g., home, school, community). There is a great deal of collaboration between family members and the treatment managers. The strengths of the family and enveloping systems are emphasized and reinforced. Initially, treatment focuses on eliminating the barriers to effective parenting.

There have been several outcome studies of multisystemic therapy. It should be noted that the target audience for treatment and study has been a group of extremely antisocial and aggressive youth. Results have shown multisystemic therapy to be superior to "traditional" services offered to these youth in their communities. Guidelines are available to direct the therapist, but they tend to be very general. Thus, in working with a conduct disordered youth who is likely to have multiple risk factors involved in the creation and maintenance of delinquent behaviors, which interventions to use and when may be ambiguous.

In summary, there are several promising psychosocial techniques that have been implemented in the treatment of conduct disordered youth. In terms of the dimensions of conduct disorder outlined in this chapter, overall, it has been shown that these interventions are most effective in adolescent-onset rather than childhood-onset conduct disorder (Kazdin, 1997). As a result, the violent and chronic offender, the youth known to have the worst outcome, to a large degree remains an enigma. However, it is hopeful that the combination of intervention techniques may offer a better prognosis for these youth than any single method described.

9.07.6 SUMMARY

The numbers of violent crimes committed by youthful offenders is on the rise. The majority of these crimes may be attributed to a small percentage of the population of conduct disordered youth. Epidemiologic information suggests a developmental pathway for these children and adolescents who will go on to be chronic and violent offenders. There is compelling evidence that there is some type of genetic predilection to the development of these behaviors which occurs predominantly in males. Psychosocial variables such as poverty, overcrowding, and inadequate parenting contribute to the matrix created by nature. As preschoolers, these children exhibit the extremes in oppositional/defiant behavior and manifest the symptoms of ADHD. They also begin to behave aggressively which remains a pervasive theme for them throughout their lives. In school, they show developmental disorders related primarily to reading and arithmetic. Within the structured environment of school these youth's chronic disregard for rules, law, and authority usually first becomes convincingly apparent. Minor illegal activity with no confrontation of a victim gives way to felonious behavior with direct harm to individuals as the child ages. Substance abuse is a frequent comorbid condition in these adolescents. Without intervention, these youth continue on this pathway into adulthood.

The US Department of Justice has outlined the objectives for intervention with these youth. There are several psychosocial techniques that show promise. These involve cognitive-behavioral methods and interventions with the multiple systems (e.g., family, school, neighborhood) that impact and contribute to this type of functioning. National strategies such as limiting accessibility to firearms and changing community expectations regarding acceptable and unacceptable behaviors are also important in the prevention and elimination of these types of behaviors. Even with individually and community oriented programs, these youth remain difficult to treat with a relatively low success rate which has typically been measured by looking for the rate of recidivism after intervention. A better understanding of the relative contributions of nature and nurture to this disorder, with an equal assessment of variables allowing for resiliency, opens avenues for continued research.

9.07.7 REFERENCES

Achenbach, T. M. (1993). Taxonomy and comorbidity of conduct problems: Evidence from empirically based approaches. *Development and Psychopathology, 5*,51–64.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd. ed.). Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders-Revised* (3rd. ed., Rev.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th. ed.). Washington, DC: Author.

August, G. J., & Stewart, M. A. (1983). Familial subtypes of childhood hyperactivity disorder. *The Journal of Nervous and Mental Disease, 171,* 362–368.

Bear, M. F., Connors, B. W., & Paradiso, M. W. (1996). *Neuroscience: Exploring the brain.* Baltimore: Williams & Wilkins.

Biederman, J., Munir, K., & Knee, D. (1987). Conduct and oppositional disorder in clinically referred children with attention deficit disorder: A controlled study. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 724–727.

Biederman, J., Newcorn, J., & Sprich, S. (1991). Comorbidity of attention deficit hyperactivity disorder with conduct, depressive, anxiety and other disorders. *American Journal of Psychiatry, 148,* 564–577.

Chess, S., & Thomas, A. (1991). Temperament. In M. Lewis (Ed.), *Child and adolescent psychiatry: A comprehensive textbook* (pp. 145–160). Baltimore: Williams and Wilkins.

Christiansen, K. O. (1977). A review of studies of criminality among twins. In S. A. Mednick & K. O. Christiansen (Eds.), *Biological bases of criminal behavior* (pp. 45–88). New York: Gardner Press.

Cloniger, C. R., & Gottesman, I. I. (1987). Genetic and environmental factors in antisocial behavior disorders. In S. A. Mednick, T. E. Moffitt, & S. A. Stack (Eds.), *The causes of crime: New biological approaches* (pp. 92–109). Cambridge, UK: Cambridge University Press.

Combating violence and delinquency: The national juvenile justice action plan (1996). Washington, DC: US Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Dalgaard, O. S., & Kringlen, E. (1976). A Norwegian twin study of criminality. *British Journal of Criminality, 16,* 213–232.

DeJong, W. (1994). *Preventing interpersonal violence among youth: An introduction to school, community and mass media strategies.* Washington, DC: National Institute of Justice.

Elliot, D. S. (1994). Serious violent offenders: Onset, developmental course, and termination. *Criminology, 32,* 1–21.

Elliot, D. S., Huizinga, D., & Ageton, S.S. (1985). *Explaining delinquency and drug use.* Newbury Park, CA: Sage.

Elliot, D. S., Huizinga, D., & Morse, B. (1986). Self-reported violent offending: A descriptive analysis of juvenile violent offenders and their offending careers. *Journal of Interpersonal Violence, 1,* 472–514.

Flynn, J. P., Vanegas, H., & Foote, W. (1970). Neural mechanisms involved in a cat's attack on a rat. In R. E. W. Whatlen, R. F. Thompson, M. Verzeams, & N. M. Weinberger (Eds.), *The neural control of behavior* (pp. 35–41). New York: Academic Press.

Hamdan-Allen, G., Stewart, M. A., & Breghly, J. H. (1989). Subgrouping conduct disorder by psychiatric family history. *Journal of Child Psychology and Psychiatry, 6,* 889–897.

Hawkins, J. D.,& Catalano, R. F. (1993). *Communities that care: Risk-focused approach using the social development strategy.* Seattle, WA: Developmental Research and Programs.

Hawkins, J. D., & Catalano, R. F. (1995). *Risk-focused prevention: Using the social development strategy.* Seattle, WA: Developmental Research and Programs.

Henggeler, S. W., Schoenwald, S. K., & Pickrel, S. G. (1995). Multisystemic therapy: Bridging the gap between university- and community-based treatment. *Journal of Consulting and Clinical Psychology, 63,* 709–717.

Hinshaw, S. P., Lahey B. B., & Hart, E. L. (1993). Issues of taxonomy and comorbidity in the development of conduct disorder. *Development and Psychopathology, 5,* 31–49.

Howell, J. C. (1995). *Guide for implementing the comprehensive strategy for serious, violent, and chronic juvenile offenders.* Washington, DC: The Office of Juvenile Justice and Delinquency Prevention.

Hunziker, D. (1995). Juvenile crime, grown-up time. *State Legislatures, 21*(5),14–19.

Kazdin, A. E. (1995). *Conduct disorder in childhood and adolescence* (2nd ed.). Thousand Oaks, CA: Sage.

Kazdin, A. E. (1997). Psychosocial treatments for conduct disorder in children. *Journal of Child Psychology and Psychiatry, 38,* 161–178.

Lahey, B. B., Piacentini, J. C., McBurnett, K., Stone, P., Hartdagen, S., & Hynd, G. (1988). Psychopathology in the parents of children with conduct disorder and hyperactivity. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 163–170.

Lewis, D. O. (1991). Conduct disorder. In M. Lewis (Ed.), *Child and adolescent psychiatry: A comprehensive textbook* (pp. 561–573). Baltimore: Williams and Wilkins.

Loeber, R., & Stouthamer-Loeber, M. (1998). Development of juvenile aggression and violence: Some common misconceptions and controversies. *American Psychologist, 53,* 242–259.

Mednick, S. A., Gabrielli, W. F., & Hutchings, B. (1984). Genetic influences in criminal convictions: Evidence from an adoption cohort. *Science, 224,* 891–894.

Moffitt, T. (1993). The neuropsychology of conduct disorder. *Development and Psychopathology, 5,* 135–151.

Pennington, B. F., & Bennetto, L. (1993). Main effects of transactions in the neuropsychology of conduct disorder? *Development and Psychopathology, 5,* 153–164.

Pettito, J. M., Lysle, D. T., Gariety, J. L., & Lewis, M. H. (1994). Association of genetic differences in social behavior and cellular immune responsiveness: Effects of social experience. *Brain, Behavior and Immunity, 8,* 111–122.

Quay, H. C. (1993). The psychobiology of undersocialized aggressive conduct disorder: A theoretical perspective. *Development and Psychopathology, 5,* 165–180.

Rogeness, G. A., Javors, M. A., Maas, J. W., & Macedo, C. A. (1990). Catecholamines and diagnoses in children. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 234–241.

Schneider-Stock, R., & Epplen, J. T. (1995). Congenic AB mice: A novel means for studying the (molecular) genetics of aggression. *Behavior Genetics, 25,* 475–482.

Snyder, H., & Sickmund, M. (1995). *Juvenile offenders and victims: A national report.* Washington, DC: Office of Juvenile Justice and Delinquency Prevention, US Department of Justice.

Spergel, I. (1995). *The youth gang problem: A community approach.* New York: Oxford University Press.

Thornberry, T. P. (1994). *Violent families and youth violence.* Washington, DC: Office of Juvenile Justice and Delinquency Prevention, US Department of Justice.

Wood, I. (1996). Conduct disorder and oppositional defiant disorder. In D. X. Parmelee (Ed.), *Child and adolescent psychiatry for the clinician* (pp. 83–96). St. Louis, MO: Mosby Press.

Zahn, T. P., & Kruesi, M. J. P. (1993). Autonomic activity in boys with disruptive behavior disorders. *Psychophysiology, 30,* 605–614.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.08
Homelessness

PAUL A. TORO
Wayne State University, Detroit, MI, USA

9.08.1 INTRODUCTION 119

9.08.2 RESEARCH ON HOMELESSNESS: RECENT FINDINGS AND NEW FRONTIERS 120

- *9.08.2.1 Public Opinion and the Prevalence of Homelessness* 120
- *9.08.2.2 Methodological Issues* 120
  - *9.08.2.2.1 Definition* 121
  - *9.08.2.2.2 Measurement and sampling* 121
  - *9.08.2.2.3 Appropriate comparison groups* 121
- *9.08.2.3 Describing the Homeless Population* 122
  - *9.08.2.3.1 Psychiatric disorders* 122
  - *9.08.2.3.2 Demographic and other basic characteristics* 124
- *9.08.2.4 Homeless Families* 125
- *9.08.2.5 Homeless Adolescents* 126
- *9.08.2.6 Longitudinal Studies* 127
- *9.08.2.7 Causes and Consequences* 127
  - *9.08.2.7.1 Deinstitutionalization* 127
  - *9.08.2.7.2 Substance abuse* 128
  - *9.08.2.7.3 Social and demographic trends* 128
  - *9.08.2.7.4 Growing poverty and other economic changes* 128

9.08.3 INTERVENTIONS 128

- *9.08.3.1 The Problem with Emergency Services* 128
- *9.08.3.2 An Ecological Perspective* 129
- *9.08.3.3 Preventing Homelessness* 129
- *9.08.3.4 Intensive Case Management and Related Interventions* 129
- *9.08.3.5 Tapping Existing Community Resources to Help the Homeless* 129
- *9.08.3.6 Clinical Issues in Working with the Homeless* 129
  - *9.08.3.6.1 Immediate and long-term needs* 129
  - *9.08.3.6.2 The heterogeneity of the population* 130
  - *9.08.3.6.3 Substance abuse* 130
  - *9.08.3.6.4 Depression* 130
  - *9.08.3.6.5 Compliance with treatment* 130
  - *9.08.3.6.6 Family and other social characteristics* 130
  - *9.08.3.6.7 Competencies and deficits* 130
  - *9.08.3.6.8 Advocacy and traditional services* 130

9.08.4 SUMMARY 131

9.08.5 REFERENCES 131

9.08.1 INTRODUCTION

Since the early 1980s, the problem of homelessness has captured the attention of the public, the media, and political leaders in the USA and most other nations of the developed world (Fournier & Mercier, 1996; Gore, 1990; Lee, Link, & Toro, 1991; Shinn & Weitzman, 1990; Toro & Rojansky, 1990). Responding to this general concern, homelessness has become a major research and clinical focus of many psychologists and other professionals. The number of publications on homelessness in psychological and other professional literatures has been steadily rising (Shinn, Burke, & Bedford, 1990; Toro & Warren, 1991) and a wide range of books on the topic have recently appeared (e.g., Baumohl, 1996; Blau, 1992; Burt, 1992; Fournier & Mercier, 1996; Jahiel, 1992; Jencks, 1994; Liebow, 1993; Rossi, 1989; Wright, 1989). A variety of interventions have been initiated and evaluated to cope with the problem (Argeriou & McCarty, 1990; Levine & Rog, 1990; NIAAA, 1991; Shinn, 1992; Shlay & Rossi, 1992). After a brief review on what has been learned about the problem from recent research and discussing areas of additional research that are developing, the chapter will focus on providing suggestions for preventive and treatment interventions and considering a variety of clinical issues that arise in working with homeless populations. The discussion on interventions and clinical issues will be grounded in an ecological perspective on homelessness (Toro, Trickett, Wall, & Salem, 1991).

9.08.2 RESEARCH ON HOMELESSNESS: RECENT FINDINGS AND NEW FRONTIERS

9.08.2.1 Public Opinion and the Prevalence of Homelessness

Over 20 opinion surveys and polls on the topic of homelessness have been conducted, mostly in the USA, since the late 1980s (Link et al., 1995). Though some have involved only a handful of questions and/or nonrepresentative samples (e.g., Barbanel, 1989; Benedict, Shaw, & Rivlin, 1988; Media General Research, 1989; "Poll shows," 1991), other surveys have used sophisticated sampling methods and have assessed a wide range of attitudes, behaviors, and political beliefs (e.g., Fournier, Toro, Link, Manrique, & Glasser, 1996; Lee, Jones, & Lewis, 1990; Link et al., 1995; Manrique & Toro, 1994, 1995; Toro & McDonell, 1992). As a group, these polls and surveys paint a consistent picture of public opinion on homelessness, one that has remained surprisingly stable throughout the political changes and events since the late 1980s. For example, across the surveys, about two-thirds of the respondents have indicated a willingness to pay more taxes to help people who are homeless. Despite some media reports in recent years on growing "compassion fatigue" and the putative backlash against aggressive panhandling and other practices of homeless persons (e.g., Furillo, 1990; "Many cities," 1994; "Shift in feelings," 1991), public support for solving the problem of homelessness has remained strong (Link et al., 1995; Manrique & Toro, 1994). The average citizen is not only supportive but well-informed about the characteristics of the homeless population and possible remedies, such as more low-cost housing and job-training programs (Toro & McDonell, 1992). Analyses on the predictors of attitudes and knowledge about homelessness show support across the political spectrum, with only age and gender as consistent predictors (women and young people showing generally more positive attitudes).

Various estimates on the prevalence of homelessness in the USA have been produced since the mid-1980s by researchers, government officials, and advocates for the homeless. Estimated prevalence rates for the US have varied from as low as 200 000 (HUD, 1984; Barringer, 1991), to 2 million (Hombs & Synder, 1986), to as high as 12–15 million (Link et al., 1994; Manrique & Toro, 1995; Toro & McDonell, 1992). During the 1980s, a heated controversy developed on which were the most accurate rates, with more conservative politicians supporting low rates and advocates for the homeless supporting higher rates (Appelbaum, 1990; Foscarinis, 1991; Kondratas, 1991; Rossi, 1989). In addition to the variation in rates based on political agenda, estimates have also generally increased over time. Though it has been suggested that such increases reflect actual growth in the size of the homeless population over this period (Burt, 1992, 1994), they may also be due to a greater willingness on the part of politicians and policy analysts to accept the problem of homelessness and the use of different estimation methods. Even more important, the rates presented have varied by the time-frame used, that is, whether point, annual, or lifetime prevalence is estimated. The highest estimates come from recent national surveys that have provided lifetime rates suggesting that 7–8% of US citizens have already experienced an episode of homelessness (Link et al., 1994; Manrique & Toro, 1995). Regardless of which rate is chosen, estimation of the prevalence of homelessness is fraught with problems and, no matter which method is used, it is likely to be criticized by someone. It is the author's belief that devoting tremendous resources to further estimate the size of the homeless population may be unnecessary: the problem is clearly of unacceptable proportions by almost anyone's standards and, even with improving economic conditions in the USA and some other nations, the rates do not appear to be dropping.

9.08.2.2 Methodological Issues

Perhaps more than anything else, researchers since the 1980s have learned how to study the problem of homelessness. An improved understanding of methodological issues has occurred in three basic areas: defining the homeless; sampling and assessment techniques; and selecting appropriate comparison groups.

9.08.2.2.1 Definition

Defining who is homeless may seem fairly straightforward but, in fact, the issues involved are as complex as they are in defining mental disorder and other important constructs in clinical psychology. How long must one be without housing to be considered homeless? Should persons living in substandard housing (e.g., shacks in rural areas, rat-infested housing projects in the cities) be considered homeless? What about those living "doubled up" with family or friends? Most researchers in the USA and other developed nations have settled these issues by studying the "literal homeless," people staying in shelters for the homeless or sleeping on the streets or other similar settings (e.g., abandoned buildings, underneath bridges). In defining the homeless, it is also important to understand that there are three basic subgroups among the homeless: homeless adults without children (mostly men); homeless families (mostly young women with young children); and homeless adolescents (runaways and throwaways). Largely distinct sets of services and distinct research literatures have developed for each of these three groups. All three of these groups should be considered homeless and of concern to clinical psychologists.

9.08.2.2.2 Measurement and sampling

Since the 1980s, researchers have developed and adapted measures appropriate for use with homeless populations, including structured diagnostic tools, measures of social networks and social support, and inventories documenting housing history (e.g., Bates, 1993; Toro & Wall, 1991). In the mid-1980s, toward the beginning of the current new wave of research on homelessness, assessing mental illness became controversial (e.g., Snow, Baker, Anderson, & Martin, 1986; Susser, Conover, & Struening, 1989). Commonly used indicators of mental illness in studies of the homeless have included: (i) a diagnosis generated by a standardized clinical interview; (ii) a score on a symptom checklist; (iii) past history of psychiatric hospitalization; and (iv) global observer assessments (Robertson, 1992). Applied to the homeless, each of these methods has its merits in terms of reliability, validity, and/or practicality. The standardized clinical interview method tends to yield some of the most conservative estimates while the other, less rigorous, methods and combinations of methods yield higher estimates. Given the high rates of substance abuse among the homeless, the inclusion of substance abuse in the definition of "mental illness" yields especially high rates of disorder. Of the different methods, only the standardized clinical interview appears to clearly separate substance abuse from mental disorder (Bellavia & Toro, 1998; Toro & Wall, 1991).

Researchers have also investigated the impact of different sampling approaches on the nature of the homeless people obtained. For example, relatively few differences have been found in the composition of homeless samples obtained across the seasons and through different sampling methods (Hannappel, Calsyn, & Morse, 1989; Toro & Wall, 1991). Several recent large-scale studies have refined probability sampling procedures for selecting representative groups across large geographical areas from a variety of shelters, food programs, and other types of sites (Burnam & Koegel, 1988; Koegel, Burnam, & Morton, 1996; Robertson, Zlotnick, & Westerfelt, 1994, 1997; Toro et al., 1998).

9.08.2.2.3 Appropriate comparison groups

Many of the studies on the homeless done in the early to mid-1980s painted a rather dismal picture of the homeless population. They seemed to have a myriad of deficits, including mental illness, substance abuse, domestic violence, and poor educational attainment. However, it was also known that poor people generally also showed some of these same deficits. A number of more recent studies have selected appropriate comparison groups of persons who are like the homeless in most respects, aside from the condition of homelessness, in order to distinguish the unique characteristics of the homeless from those of the poor (e.g., Bukowski, Wolfe, & Toro, 1998; Goodman, 1991; Laurin, Fournier, Toupin, & Ostoj, 1996; Shinn, Knickman, & Weitzman, 1991; Sosin, 1992; Toro & Bukowski, 1995; Toro, Bellavia et al., 1995; Wood, Valdez, Hayashi, & Shen,1990). Such comparison group designs have now been applied to all three of the major groups of homeless people mentioned earlier (i.e., single adults, adolescents, and families). These studies suggest a number of specific areas that differentiate the poor from the homeless, as well as other areas that are less sensitive in distinguishing these groups. For example, substance abuse has emerged as a frequent predictor of being homeless (rather than poor, but housed), while severe mental illness has not.

9.08.2.3 Describing the Homeless Population

Research since the mid-1980s, primarily conducted in the USA, has now carefully documented the diagnostic and other basic characteristics of homeless adults based on large representative samples from relatively broad geographical areas within various cities. An overview of these characteristics, which are generally consistent from city to city and from study to study, are presented below. Note that most of the relevant studies have not included homeless families and homeless youths but have instead focused on single adults who form the largest segment of the homeless population. In later sections, the findings from the largely separate and more limited research literatures on homeless families and adolescents will be discussed.

9.08.2.3.1 Psychiatric disorders

Because it provides the most detailed data on the disorders observed among the homeless and because it separates substance abuse from other disorders, studies using a standardized clinical interview such as the Diagnostic Interview Schedule (DIS) will be emphasized (Eaton & Kessler, 1985; Regier et al., 1988; Robins, Helzer, Croughan, & Ratcliff, 1981). In particular, a set of nine samples of homeless adults from studies that all used the DIS will be focused on (see Fisher, Shapiro, Breakey, Anthony, & Kramer, 1986; Fournier, 1988; Fournier, Mercier, Raynault, Ohayon, & Caulet, 1993; Koegel, Burnam, & Farr, 1988; Koegel, Sullivan, Burnam, Morton, & Wenzel, 1997; North & Smith, 1993; Roberston et al., 1994, Toro et al., 1997; Toro et al., 1998).

These studies involved representative samples of homeless adults, generally obtained using similar probability sampling methods, and they were from seven different metropolitan areas in the US and Canada. The studies have produced a fairly consistent diagnostic portrait of the homeless based on *DSM-III* and *DSM-III-R* criteria (see Table 1). Because these studies used a common method to operationalize mental disorder and fairly similar sampling approaches, the consistency of diagnostic rates across them suggests that the variability in rates of disorder across various earlier studies on the homeless noted by some reviewers (e.g., Robertson, 1992; Toro & Wall, 1991) was largely due to variability in sampling and measurement methods, rather than to differences in the homeless populations across cities.

The nine studies using the DIS have yielded high rates for the major affective disorders among homeless adults (14–30%), with depression generally accounting for the majority of these disorders (bipolar disorders and mania being much less frequent, usually < 5% overall). Less severe depression, not meeting all diagnostic criteria, is even more prevalent among homeless adults (Bellavia & Toro, 1998; Toro & Wall, 1991) and post-traumatic stress disorder (PTSD) also appears very common (one recent study using the DIS identified a rate of 40%; North & Smith, 1992). Such high rates of depression and PTSD are consistent with recent views suggesting that homelessness often involves a series of traumatic life events (Goodman, Saxe, & Harvey, 1991). The stresses associated with being homeless have also now been well-documented (e.g., Koegel, Burnam, & Farr, 1990; Toro, Bellavia et al., 1995; Toro et al., 1998). Among the few studies that have included the DIS section on anxiety disorders (including panic and generalized forms), combined lifetime prevalence rates have ranged from 10 to 39% (Fischer et al., 1986; Koegel et al., 1988; North & Smith, 1992).

Although the rates for schizophrenic disorders among homeless adults have generally exceeded normative rates by a factor of two to six, such persons are still relatively rare. With the notable exception of the study by Koegel et al. (1988), which was done entirely in the large skid row area of Los Angeles, studies using the DIS have consistently found that under 10% of homeless adults have a schizophrenic disorder meeting diagnostic criteria. These relatively low rates for schizophrenia do not conform with expectations of many professionals and average citizens nor with the extensive research attention given the homeless mentally ill. This inconsistency may be a result of limitations of the DIS and other structured diagnostic measures, which have been criticized as systematically underestimating schizophrenic symptoms, in large part because they involve only a time-limited exposure between interviewer and subject (Arce & Vergare, 1984; Susser et al., 1989). However, the discrepancy between rates obtained in research and public perceptions could also the result of extensive media coverage of the homeless mentally ill and to the fact that such persons, even though they may be relatively rare, are generally very noticeable to the average citizen when observed on the streets.

The media and policy analysts often quote the general figure of one-third for mental illness among the homeless. This figure seems to be based on earlier studies that obtained their estimates through self-reported psychiatric hospitalization (see Roberston, 1992, for a review of such studies). It has been suggested that a fair number of such self-reported hospitalizations were likely for substance abuse

Table 1 Lifetime prevalence for severe mental disorders and substance abuse based on representative samples of homeless adults and the DIS.

			Mental disorders			*Substance abuse*			
	Sample size	*City*	*Affective*	*Schizophrenic*	*Either*	*Alcohol*	*Drug*	*Either*	*Dual diagnosis*
Fischer et al. (1986)	51	Baltimore, MD	13.7	2.0				70.6	
Fournier (1988)	299	Montreal, Canada	28.1	8.7		66.9	40.4	74.2	31.8
Koegel et al. (1988)	328	Los Angeles, CA	29.5	13.7		62.9	30.8	69.2	
Koegel et al. (1997)	1563	Los Angeles, CA	21.0	7.0		60.0	49.0	71.0	
North and Smith (1993)	900	St. Louis, MO	23.0	5.4		47.5	34.2	60.1	
Robertson et al. (1997)	564	Oakland, CA	21.6	4.1	22.3	52.6	52.2	69.1	19.0
Toro et al. (1997)	213	Buffalo, NY	18.8	3.3	20.3	45.5	37.6	57.6	14.2
Toro et al. (1998)	420	Buffalo, NY	20.2	4.1	22.6	55.6	48.9	70.0	17.4
Toro et al. (1998)	297	Detroit, MI	23.6	6.4	27.3	55.1	53.9	75.3	23.9
Mean rate for homeless			21.4	6.1	23.1	55.8	43.4	68.6	21.3
Normative rate			8.3	1.5		13.3	5.9	16.4	

All disorders are based on *DSM-III* or *DSM-III-R* criteria. Major affective disorders generally include bipolar and depressive disorders and dysthymia. Schizophrenic disorders generally include schizophrenia and schizophreniform disorder. Substance abuse disorders include both abuse and dependence. The dually diagnosed have both a severe mental disorder and a substance abuse disorder. The mean rate weights each of the nine samples equally. The national normative rates are based on the five-site Epidemiologic Catchment Area study ($N = 18\,571$; source: Regier et al., 1988).

(Bellavia & Toro, 1998). Furthermore, having a prior hospitalization, even if for mental illness, does not necessarily assure that all *DSM-III-R* criteria for a serious mental disorder were met at admission (since rigorous diagnostic methods are not universally used) and prior hospitalization is a poor indicator of current diagnosable mental illness requiring intervention. Thus, the often cited one-third figure is probably an overestimate and planning services for the homeless based on the assumption that a large proportion are actively psychotic is misguided. It is also unclear whether the high rates of mental disorders found can be attributed to homelessness *per se*, since the rates for the poor are also similarly high (Toro, Bellavia et al., 1995) and, especially in the case of depression, it is not clear to what extent the disorder is a response to the stresses of homelessness and how much it is a precursor to homelessness (Goodman et al., 1991).

Alcohol and drug abuse and dependence are very common among homeless adults, with studies using the DIS generally finding 60–75% having a lifetime diagnosis of a substance abuse disorder. A large number of these persons (generally about half) abuse both alcohol and drugs and, among those abusing drugs, often multiple drugs are used (most commonly marijuana, cocaine, and opioids). Though these rates of substance abuse are obviously high and many homeless adults are chronic abusers, a sizeable number (generally 30–50% of those with a lifetime diagnosis of substance abuse) report little or no recent symptoms of substance abuse. Even if such persons "in recovery" are being truthful about current use, given their past problems they are still obviously at risk of resuming use and abuse. Most homeless adults with a severe mental disorder have a co-occurring substance abuse disorder. Such "dually diagnosed" persons have been the focus of extensive recent attention by researchers and program developers and pose special problems for service providers (Blankertz & White, 1990; Dende & Kline, 1995; Dexter, 1990; Drake, Osher, & Wallach, 1991). Several studies have attempted to document the prevalence of organic brain disorders that are often associated with chronic substance abuse (Fischer et al., 1986; Koegel, 1988; North & Smith, 1992; Toro & Wall, 1991). Estimates of cognitive impairment based on the DIS have been as high as 15%.

Antisocial personality disorder and criminal histories in general are common among the homeless. Studies that have assessed antisocial personality based on the DIS and *DSM-III* or *DSM-III-R* criteria have found lifetime prevalence rates ranging from 16 to 37% (Fischer et al., 1986; Fournier, 1988; Koegel et al., 1988; North & Smith, 1992; Toro & Wall, 1991). About 25–30% of the homeless report having been convicted of a felony or spent some time in prison and many others (another 35–50%) report having been arrested. However, it has been noted that many arrests of the homeless are for "victimless" crimes such as loitering, disturbing the peace, or public drunkenness (Fischer, 1988).

9.08.2.3.2 Demographic and other basic characteristics

Based on recent studies involving representative samples of homeless adults, we now know a great deal about a range of other characteristics of the homeless population. Most homeless adults continue to be men (usually approximately 70% of the total population) and the majority (about two-thirds) are under 40 years of age. Though many homeless appear old, very few are over 60 (usually considerably <10%). This lack of old people among the homeless may be due to early mortality that reduces their presence in the population or to relatively generous benefits available to the elderly in our society that may prevent homelessness (Rossi, 1989).

African-Americans are heavily over-represented among all homeless persons in the USA, as they are in the poverty population more generally. In cities with relatively large numbers of African-Americans in the general population (e.g., 20% in metropolitan Detroit), African-Americans often constitute a clear majority of the homeless population (85% in Detroit; see Toro et al., 1998). Other ethnic groups that appear to be over-represented among the homeless, depending the city involved, include Hispanics and Native Americans.

Not only are homeless adults currently in severe poverty, but they also generally come from poor socioeconomic backgrounds. For example, in a recent two-city study involving representative samples of homeless adults, the parents of approximately 80% of the homeless in each city held unskilled or blue collar positions and only 3% had parents with professional backgrounds (Toro et al., 1998). The majority of homeless adults have never been married (usually approximately 50%) or are separated or divorced (usually another 40%). Though only a small minority of homeless adults have their children with them while homeless, many more (including both women and men) report having children for whom they feel financially responsible (a recent two-city study found rates of 39–43%; Toro et al., 1998). Often such children are being temporarily cared for by relatives or friends or are in foster care after having been removed from the parent's care by protective service agencies.

A great deal of recent attention has been paid to other social characteristics of the homeless. It is now well documented that most homeless persons have regular contact with family members (Toro & Wall, 1991; Toro et al., 1998). However, the nature of such contact may not always be positive. In fact, the existence of conflict and violence in social relationships have been identified as risk factors for homelessness among both adults, families, and adolescents. In assessing the childhood histories of both homeless adults and adolescents, out-of-home foster placements, sexual and/or physical abuse, family transience, punitive parenting practices, parental substance abuse and/or mental illness, and other indicators of noxious home environments have been commonly found (Boesky, Toro, & Wright, 1995; Koegel, Melamid, & Burnam, 1995; Susser et al., 1987; Toro, Bellavia et al., 1995; Wolfe, 1994). Violence in romantic relationships has also been identified as a significant risk factor for men as well as women (Toro, Bellavia et al., 1995). For women, such violence often leads to flight from an abusive situation and seeking refuge in a battered women's shelter (women in such shelters are sometimes included in representative samples of homeless adults; e.g., Toro et al., 1998). For men, whether they are perpetrators or victims of the violence (or both), relationships in which violence is prevalent may frequently dissolve and lead to a lack of stable housing.

As might be expected, the health status of homeless persons is generally poor (Gelberg, Linn, Usatine, & Smith, 1990; Institute of Medicine, 1988; Lozowski, Rowland, & Toro, 1994; Robertson & Cousineau, 1986; Wright & Weber, 1987). Particularly common among homeless adults are respiratory problems, skin disorders, head and other injuries, and hypertension. Rates of smoking are also high ($>80\%$; see Lozowski et al., 1994), which can help partially explain the prevalence of some of these health problems. The homeless also appear to be at heightened risk of HIV and tuberculosis infection (Gelberg, Panarites, Andersen, & Koegel, 1994; Stayanoff, Weber, Wiley, Ford, & Kerndt, 1994; Valencia, Saez, Susser, Conover, & Colson, 1994; Zolopa et al., 1994).

Finally, it should be noted that veterans, mainly those who served during the Vietnam era or later, may be another over-represented group, with studies generally finding approximately 25–30% of the total adult homeless population having military experience (Goldstein & Toro, 1995; Rosenheck & Koegel, 1993). Why so many veterans are homeless is less clear, although explanations have focused on the stresses associated with military life (including frequent moves and taxing work schedules, as well as combat experience) that may contribute to social and family problems and to high rates of substance abuse. A number of special interventions for homeless veterans, some targeted on a national scale, have also been attempted (e.g., Rosenheck, Gallup, & Frisman, 1993; Seidner, Burling, Fisher, & Blair, 1990).

9.08.2.4 Homeless Families

It is widely believed that homeless families are the fastest growing segment of the homeless population (e.g., US Conference of Mayors, 1989). In recent years, the media as well as researchers and service providers have been devoting increasing attention to this important subgroup among the homeless (Lee et al., 1991). Such families consist mostly of young women (generally in their twenties and thirties) who are homeless with their young children. Few families involving men are encountered, perhaps due in part to policies that prohibit men in many shelters for homeless families. Unlike homeless adults as a group (in which males clearly predominate, even when women homeless with their children are included in the samples; e.g., Toro et al., 1997; Toro et al., 1998), homeless children seem equally divided between girls and boys. Among children homeless with their mothers, those of school age (especially those 10 years or older) are relatively rare. Much more common are infants and toddlers. This lack of school-aged children among the homeless could be due to the eventual intervention of protective service workers (leading to the placement of such children) or the willingness of family members to care for such children, who may be seen as less demanding than infants or toddlers.

As in the general population, depressive symptoms are even more common among homeless women than homeless men, though women homeless along with their children seem to show levels of depression that are roughly comparable to those seen in homeless single men, suggesting a possible protective effect of motherhood (North & Smith, 1993; Roll, Toro, & Ortola, 1995; Smith & North, 1994). Again paralleling gender differences in the general population, both substance abuse and antisocial behavior among homeless women, especially those homeless with their children, appear to be much less common than among homeless men (Roll et al., 1995; Shinn & Weitzman, 1990; Smith & North, 1994).

Children homeless with their mothers have been found to be at elevated risk for a variety of problems, including poor health, educational delays, poor school performance, stigmatization by peers, behavioral problems (e.g., aggression, anxiety, and depression), and poor coordination (Bassuk & Gallagher, 1990; Bassuk & Rubin, 1987; Bassuk, Rubin, & Lauriat, 1986; Dail, 1990; Fox & Roth, 1989; Gewirtzman & Fodor, 1987; Molnar, Rath, & Klein, 1990; Passero Rabideau & Toro, 1997; Rafferty & Shinn, 1991; Wagner & Menke, 1990; Wright, 1990). Unfortunately, few of the studies documenting these problems have utilized comparison groups of stably-housed poor children in an attempt to differentiate the effects of homelessness from poverty. Those that have matched homeless children to poor, housed children (e.g., welfare recipients) have generally found that the two groups do not differ from each other, though both differ from the general population (Rafferty & Shinn, 1991).

Witnessing domestic violence has been consistently found to relate to a variety of poor outcomes in children (Jaffe, Wolfe, & Wilson, 1990). Given the high rates of domestic violence observed among homeless adults (including those with children), even when compared to comparable housed persons (Bassuk & Rosenberg, 1990; D'Ercole & Struening, 1990; Shinn et al., 1991; Toro, Bellavia et al., 1995), it seems likely that homeless children will often witness violence and be victimized themselves. Such exposure may require special attention from service providers.

9.08.2.5 Homeless Adolescents

Most investigators have found that the majority of runaway behavior occurs during mid-adolescence and studies suggest that 9–15% of adolescents have at some time run away from home or been thrown out by their parents (Nye & Edelbrock, 1980; Ringwalt, Greene, & Iachan, 1994; Smart, Adlaf, Walsh, & Zdanowicz, 1994; Windle, 1989). Recent research on homelessness has considered both runaways and throwaways. Some studies have also focused on "street youth" and others have included persons as old as 21 in their samples (e.g., Cauce et al., 1993, 1994). Probability sampling methods are not commonly used and existing studies on homeless adolescents have often obtained their samples from a single youth shelter, mental health or other service agency, or specific inner-city street area where homeless adolescents "hang out." The review of the literature suggests that studies that obtain "street" samples (e.g., Cauce et al., 1993, 1994; Robertson, 1990) often find a different profile for the homeless adolescent compared to studies that obtain their samples from youth shelters or other service settings (e.g., Bukowski et al., 1998; Mundy, Robertson, Greenblatt, & Robertson, 1989; Shaffer & Caton, 1984; Yates, MacKenzie, Pennbridge, & Cohen, 1988). The street samples appear to have higher levels of dysfunction. Studies that identify runaways from among those in much larger school-based or household survey samples (Ringwalt et al., 1994; Smart et al., 1994; Windle, 1989) are likely to be biased toward the less dysfunctional because homeless youth are less likely to be at home or in school.

Though the research literature on homeless adolescents is still rudimentary at this point in time (especially when compared to the comparable literature on homeless adults; see Robertson, 1991), a rough profile of the population can still be given. Unlike homeless adults, homeless adolescents seem more equally divided between girls and boys, though studies focusing on street youth and/or including persons aged 18–21 years show higher percentages of males (e.g., Cauce et al., 1993, 1994; Robertson, Koegel, & Ferguson, 1989). Combining this gender profile observed among adolescents with the well-documented over-representation of men among homeless adults, it has been suggested that the risk of homelessness may drop for women as they get older from adolescence through young adulthood, but rise for men during this period of life (Boesky, Toro, & Bukowski, 1997).

Two recent studies have collected probability samples of homeless adolescents and have used the DIS or DIS for Children (DISC; Fisher, Wicks, Shaffer, Piacentini, & Lapkin, 1992; Piacentini et al., 1993; Shaffer et al., 1993), providing diagnostic information based on *DSM-III* or *DSM-III-R* criteria. One of these sampled 122 adolescents (ages 12–17) from all six major shelters for homeless youth in the seven-county Detroit metropolitan area (see Boesky et al., 1997; Bukowski et al., 1998; Toro & Bukowski, 1995). The other sampled 97 street youth (ages 13–17) in Hollywood (see Robertson, 1990; Robertson et al., 1989). Both studies found high lifetime rates for affective disorders (44% for depression and mania combined in the Detroit study; 26% for depression only in the Hollywood study), substance abuse/dependence (including both alcohol and drugs; 75% in Hollywood; 34% in Detroit), and disruptive behavior disorders (93% for conduct disorder alone in Hollywood; 48% for conduct, attention deficit, and/or oppositional defiant disorders in Detroit). The Hollywood study also found a high rate for PTSD (38%). The Detroit study included a carefully matched comparison group of housed adolescents and found that both alcohol abuse/dependence and disruptive behavior disorders were significantly more prevalent among the homeless (Bukowski et al., 1998).

9.08.2.6 Longitudinal Studies

Longitudinal research with the homeless represents one of the exciting new frontiers now being explored. Preliminary findings from several ongoing studies, mostly involving homeless adults, are just beginning to emerge (e.g., Cauce et al., 1993, 1994; Cohen et al., 1993; Goldstein & Toro, 1996; Koegel & Burnam, 1994; Robertson et al., 1994; Rowland & Toro, 1996; Shinn, 1996; Toro et al., 1997; Wright, Allen, & Devine, 1993). At this point in time, based on follow-up data of 18 months or more, investigators have identified different patterns in the longitudinal course of homelessness. A sizable chronic group (approximately 25–35% of adults) remains homeless for much of the follow-up period and, at the other extreme, an escapee group (depending on the definition, 15–35%) appears to have made a permanent exit from their homelessness. Others fall in between these two extremes, perhaps experiencing an exit from homelessness followed by another brief return at some later point (such a group could be called the episodic homeless). A few recent studies that followed homeless adults for more than a year have found that, when tracked this long, the homeless show improvements on several outcomes, including physical health, homelessness, and perceived life stress (Rowland & Toro, 1996; Shinn, 1996; Toro et al., 1997). Such results suggest that, when sampled, homeless people may be at a particularly difficult point in their lives and can be expected to show some improvements over time.

9.08.2.7 Causes and Consequences

Psychologists are drawn to questions of causality. Unfortunately, they cannot use experimental methods to assess either the causes or consequences of homelessness: They cannot randomly assign people to homelessness nor to the sorts of factors (e.g., poverty, substance abuse, or mental disorder) that might cause it. They must generally rely on various correlational methodologies if they are to understand the important questions they have about homelessness. Creative use of such methodologies represents another research frontier.

Longitudinal studies on the homeless can be used to approach some causal questions. Psychologists can explore who among the heterogeneous homeless population has better and worse outcomes. Using structural equation modeling and other sophisticated new statistical techniques, they can even begin to understand more complex temporal ordering between various predictors and outcomes, including future homelessness. Such analyses are currently underway for a number of large-scale longitudinal studies of homeless adults (e.g., Wolfe, 1993). Studies following poor people or others at risk of homelessness might suggest answers to other causal questions. Such studies could identify factors that might be responsible for initial episodes of homelessness. They could also investigate the possible causes of homelessness using experimental methods by randomly assigning persons at risk of homelessness to interventions designed to mitigate the impact of risk factors. Unfortunately, such studies would be extremely costly to undertake, since large numbers would need to be tracked over several years in order to observe significant amounts of homelessness prospectively.

Designs that compare homeless groups to carefully constructed comparison groups of similar (but not homeless) groups can also help refine our understanding on the causes of homelessness. For example, such studies have been finding that substance abuse differentiates the housed from the homeless, while mental illness is a less sensitive predictor of group differences (Bukowski et al., 1998; Toro, Bellavia et al., 1995). Though by no means conclusive, such findings suggest that substance abuse may be the more important causal factor.

International comparisons, a few of which have already been made, could also be useful in the search for the causes of homelessness (e.g., Adams, 1986; Daly, 1990; Fournier et al., 1996; Toro, Alan et al., 1995; Toro & Rojansky, 1990). If comparable rates of homelessness could be produced across various nations, then economic, cultural, and other factors possibly associated with varying rates could be identified.

Based on the existing research, a number of plausible causes of homelessness will be briefly presented. These causes involve a variety of socioeconomic factors as well as conditions with which clinical psychologists have expertise.

9.08.2.7.1 Deinstitutionalization

The first thought that comes to mind for many when attempting to come to grips with explaining why a wealthy nation such as the USA would have so much homelessness involves the deinstitutionalization of the mentally ill that took place in the 1960s and 1970s. Unfortunately, this explanation does not go very far since most homeless people are not seriously mentally ill and have never been hospitalized. Even using a reasonably loose definition of mental illness that includes severe depression in addition to schizophrenia, no more than a third of all homeless people can be considered mentally ill. The deinstitutionalization argument ignores the majority of the homeless population who are not mentally ill. Casting further suspicion on this argument, recent comparison group studies find that rates of severe mental illness are nearly as high among the poor as they are among the homeless (Bukowski et al., 1998; Toro, Bellavia et al., 1995).

9.08.2.7.2 Substance abuse

The combination of very high rates observed (see Table 1), the results of comparison group designs which identify greater substance abuse among the homeless vs. the housed, and other research data all point to substance abuse as an important factor in homelessness. One international comparison suggested that the low rates of homelessness observed in certain developed nations (e.g., Japan) may be due to lower rates of substance abuse (Toro & Rojansky, 1990).

9.08.2.7.3 Social and demographic trends

The pervasiveness of family conflict and violence seen in the histories of homeless people supports the idea that growing family breakdown might be responsible for the increasing prevalence of homelessness seen in recent years. Nations with apparently lower rates of homelessness also seem to show fewer such trends. Related sociodemographic trends can also be implicated in homelessness. Delayed marriage, more single-parent families, and the associated greater demand for housing units all seem to be reasonable explanations for the increase in homelessness observed in the past decade or two. The impact of the baby boom generation can also be considered. With such a large number of middle-aged persons monopolizing the job and housing markets, many in this age group, as well as younger adults (in their 20s), have recently been competing for scarce employment and housing.

9.08.2.7.4 Growing poverty and other economic changes

The loss of well-paying but unskilled jobs, such as those in manufacturing industries, is one obvious explanation for the growing homelessness observed in recent years. The growth in income disparity between the rich and poor can also be suggested as a possible cause of increasing homeless. Homelessness can be seen as a by-product of more and more persons entering the ranks of the poor and the wealth of the USA and other developed nations becoming increasingly concentrated in the hands of a smaller percentage of people (Phillips, 1990). Reductions in real welfare benefits observed in recent years, in large part a function of the "welfare reform" that has been so politically popular, has also been suggested as a major cause of increased homelessness (Jencks, 1994; Rossi, 1989). Urban revitalization, a well-meant set of social policies that began in the 1960s and has destroyed many viable though poor and "blighted" neighborhoods having extensive low-income housing in the 1960s, has also been suggested as one factor that may be partially responsible for the recent rise in homelessness (Adams, 1986; Wright & Lam, 1987).

In summary, it seems likely that a number of alternative explanations for the recent rise in homelessness have some merit. While we will probably never have definitive answers to the key causal questions on homelessness, we can still proceed with interventions based on a research understanding of the problem, rather than responding solely on the basis of impressionistic beliefs.

9.08.3 INTERVENTIONS

Perhaps the most important new frontier in research on the homeless involves the identification and evaluation of effective interventions. Clinical psychologists have much to offer in promoting this research direction. In the rush to deal with the problem of homelessness in the past decade, a great many interventions for the homeless have been implemented, often haphazardly. Few of these interventions have been carefully evaluated. Not only do more interventions need to be developed to prevent and treat homelessness, but such interventions must be carefully evaluated to determine if they are effective.

9.08.3.1 The Problem with Emergency Services

The growth in services for the homeless has been occurring at a fast pace in the USA and UK since the late 1980s and has been since the early 1990s in many other developed nations. Most of the services being developed have been oriented toward meeting the immediate needs that homeless people have (e.g., food, shelter, health care) and are provided only on a temporary basis. However, based on existing research, it is known that the homeless have long-term needs. Emergency services also do little to stem the tide of the large numbers of people who continually enter the ranks of the homeless.

9.08.3.2 An Ecological Perspective

The inadequacy of the sorts of services being developed to address homelessness suggests the need for a broader, ecological perspective. Such a perspective highlights the need to consider both social and other contextual factors, in addition to individual vulnerabilities, when developing interventions (Toro et al., 1991). It also points out that interventions need to be considered at multiple levels of analysis, from individually-focused treatment to large-scale social policy.

9.08.3.3 Preventing Homelessness

Since the 1980s, there has been considerable progress in developing effective preventive interventions (see Price, Cowen, Lorion, & Ramos-McKay, 1988). Preventive interventions have now addressed most areas of traditional concern to clinical psychologists, including mental disorders, substance abuse, delinquency, and school drop-out. Interventions oriented toward preventing homelessness have yet to be developed, though some promising approaches for preventing the closely related problem of poverty have begun to emerge (e.g., Head Start; see Zigler & Styfco, 1990). Based on an ecological perspective, a number of potentially effective approaches to preventing homelessness can be proposed. Some involve changing large-scale social policies. For instance, rather than eliminating welfare benefits, such benefits could be expanded and programs designed to reduce domestic violence and substance abuse could be more widely implemented. Of course, such approaches would require large amounts of funds and would need to be politically viable. Given the many risk factors for abnormal development that have been observed among homeless children (Rafferty & Shinn, 1991), preventive interventions, such as Head Start, may be especially warranted for this vulnerable group.

9.08.3.4 Intensive Case Management and Related Interventions

Though not oriented toward prevention, a number of other innovative interventions, consistent with an ecological perspective, have been developed in the 1990s to help currently homeless persons escape homelessness. One general category of such interventions involves intensive case management. Such interventions work closely with homeless clients, attempting to meet all of their long- and short-term needs, including permanent housing, job training and placement, and linkages to services in mental health, substance abuse, health care, and other relevant areas. Interventions of this type have been adapted and shown some effectiveness in careful research evaluations for a variety of homeless populations, including the homeless mentally ill (Morse, Calsyn, Allen, Tempelhoff, & Smith, 1992), homeless street youth (Cauce et al., 1993, 1994), and more general samples of homeless adults and families (Toro et al., 1997).

9.08.3.5 Tapping Existing Community Resources to Help the Homeless

Another promising general approach for assisting those who are currently homeless involves tapping existing community resources in creative ways in order to reduce the incidence of homelessness. The Habitat for Humanity program is an example of this approach. This program taps interested volunteers in the community to develop low-income housing. Interventions can also consider homeless people themselves as valuable resources. Thus, the homeless can be taught to help rehabilitate housing for use by the poor and serve in other public works capacities. In this way, the homeless receive valuable training that can help them reintegrate into the community at the same time that they reduce the potential future risk of homelessness for others and improve the general quality of life in the community. Unfortunately, such approaches have only been developed on a small scale and there have been no systematic attempts to rigorously evaluate them.

9.08.3.6 Clinical Issues in Working with the Homeless

In the sorts of innovative programs described above, as well as in more traditional clinical interventions, there are a number of special issues that should be considered when working with the homeless.

9.08.3.6.1 Immediate and long-term needs

Most services targeted specifically to homeless persons have focused on their immediate needs. While it is obviously important that these needs be addressed, the long-term needs of the homeless require even more attention. These needs include substance abuse treatment, job training and placement, health care, suitable permanent housing, mental health services, and assistance to repair damaged family and other social relationships. Generally, any particular homeless person or family will show a variety of such needs. Unfortunately, the human service system is not designed to address such a multiplicity of long-term needs. Rather, that system is fragmented, with separate agencies able to handle only one or two areas. Intensive case management services are one way to improve on such a system, since these services insert someone who knows the system into the life of the homeless person or family so that the multiplicity of client needs can be readily met. A one stop shopping approach to service delivery for the homeless is recommended, in which all needs can be handled at a single site.

9.08.3.6.2 The heterogeneity of the population

In addressing both the short- and long-term needs of the homeless, service providers and planners need to consider the heterogeneous nature of the homeless population. Homeless people come from all age and ethnic groups, include both men and women, and show a variety of other background characteristics that are generally similar to the cities in which they are found. Such heterogeneity suggests that effective interventions need to be individually tailored.

9.08.3.6.3 Substance abuse

One of the factors most clearly distinguishing the homeless from other groups in our society is a history of drug and/or alcohol abuse. While many of those with such a history may not currently be abusing, all need long-term help to prevent future resumption of destructive substance abusing behaviors. In serving the homeless, the role of substance abuse cannot be forgotten. As in substance abuse treatment generally, research evaluations should be undertaken to assess both short- and long-term program effectiveness. A number of such evaluations have begun (e.g., Argeriou & McCarty, 1991; NIAAA, 1991).

9.08.3.6.4 Depression

The presence of depressive symptoms, of varying severity across homeless individuals, also needs to be considered. Whether such depression is seen as a cause or effect of homelessness, service providers must deal with the sense of hopelessness and despair that is so commonly seen among homeless people. Such negative affect can stymie well-intentioned intervention efforts. It may sometimes be necessary to address the affect itself in order to motivate the homeless to work on their other problems.

9.08.3.6.5 Compliance with treatment

Depression, substance abuse, physical and other disabilities, complicated social lives, the lack of financial resources, criminal histories, and the need to cope with immediate crises can all conspire to make serving the homeless less than rewarding. Flexibility of approach and a willingness to deal with whatever problem is of concern to the client can help engage homeless clients in both traditional and more innovative services.

9.08.3.6.6 Family and other social characteristics

While a small minority of the homeless fit the stereotype of the social isolate, considerable research has documented that most homeless people have regular contact with family. Unfortunately, this contact is not always positive. Among all segments of the homeless population, domestic violence and conflict are prevalent and often serve to precipitate and prolong episodes of homelessness. If homeless people are to make long-term adjustments in the community, attention needs to be paid to improving family functioning. Conflict resolution strategies, family therapy, and other intervention models can address these needs.

9.08.3.6.7 Competencies and deficits

While their deficits clearly need to be addressed if homeless people are to make positive long-term adjustments in the community, their competencies also need to be considered and strengthened. For example, many young homeless men have some employable skills and many women homeless with their children have good parenting abilities. These skills can be honed and put to use. For example, men with appropriate skills could be placed in positions to assist in the construction of more low-income housing and women with effective parenting skills could serve as paraprofessional parent or teacher aides.

9.08.3.6.8 Advocacy and traditional services

With their expertise in assessment and behavioral management, and their understanding of the many problems faced by homeless people (e.g., substance abuse, depression), clinical psychologists are in a unique position to provide direct services and supervise the services provided by others to this population. While such services are becoming increasingly funded through special grants and contracts and some psychologists have been providing their services to the homeless on a *pro bono* basis (Freiberg, 1991), fiscal pressures in the health and mental health care fields also operate to limit access to services for the poor and homeless. In addition to providing and developing traditional services, it is recommended that clinical psychologists interested in helping the homeless also consider other, less traditional roles. With their conceptual and research skills, clinical psychologists are well-equipped to assist in effective program evaluation and design and in policy analysis and planning. Also, since they are often respected members of their communities, psychologists can help address the needs of homeless people by educating policy makers and through political advocacy. If the incidence of homelessness is to be reduced, such advocacy as well as many more persons willing to assist homeless people directly are needed.

9.08.4 SUMMARY

Homelessness has become a troubling social problem in the USA and most other developed nations. Psychologists and other professionals have become involved with the problem as researchers and service providers. Recent research has shed light on many aspects of homelessness. Public opinion surveys have identified firm support among Americans for solving the problem and the diagnostic and other basic characteristics of the homeless in American cities have been well documented. Researchers have refined sampling and measurement methods and are now increasingly engaged in longitudinal and intervention research. In response to the great perceived need, services for the homeless have been developing at a rapid pace since the early 1980s. Unfortunately, most of the services are targeted only to the immediate needs of homeless people rather than their long-term problems and few of the services are being carefully evaluated. In working to help the homeless, a number of special issues are often encountered. These include dealing with the substance abuse and depression that are so common among these people and expanding traditional clinical approaches to consider the person's entire array of skills and resources and to include program design and evaluation activities and political advocacy.

9.08.5 REFERENCES

Adams, C. T. (1986). Homelessness in the post-industrial city: views from London and Philadelphia. *Urban Affairs Quarterly, 21,* 527–549.

Appelbaum, R. P. (1990). Counting the homeless. In J. Momeni (Ed.), *Homelessness in the United States: data and issues* (pp. 1–16). New York: Praeger.

Arce, A. A., & Vergare, M. J. (1984). Identifying and characterizing the mentally ill among the homeless. In H. R. Lamb (Ed.), *The homeless mentally ill: A task force report of the American Psychiatric Association* (pp. 75–90). Washington, DC: American Psychiatric Association.

Argeriou, M., & McCarty, D. (Eds.) (1990). *Treating alcoholism and drug abuse among homeless men and women: Nine community demonstration grants.* New York: Haworth Press.

Barbanel, J. (1989, June 29). New Yorkers fault city help for the homeless, poll shows. *New York Times*, pp. A1, B4.

Barringer, F. (1991, April 12). US homeless count is far below estimates. *New York Times,* p. A11.

Bassuk, E., & Gallagher, E. M. (1990). The impact of homelessness on children. *Child & Youth Services, 14,* 19–33.

Bassuk, E. L., & Rosenberg, L. (1990). Psychosocial characteristics of homeless children and children with homes. *Pediatrics, 85,* 257–261.

Bassuk, E. L., & Rubin, L. (1987). Homeless children: A neglected population. *American Journal of Orthopsychiatry, 57,* 279–286.

Bassuk, E. L., Rubin, L., & Lauriat, A. (1986). Characteristics of sheltered homeless families. *American Journal of Public Health, 76,* 1097–1101.

Bates, D. (1993, August). *Cross-sectional and longitudinal analyses of social support among the homeless.* Paper presented at the 101st Annual Convention, American Psychological Association, Toronto, Canada.

Baumohl, J. (Ed.) (1996). *Homelessness in America.* Phoenix, AZ: Oryx Press.

Bellavia, C., & Toro, P. A. (1998). Mental illness among homeless and poor people: A comparison of assessment methods. *Community Mental Health Journal.*

Benedict, A., Shaw, J., & Rivlin, L. (1988). Attitudes toward the homeless in two New York City metropolitan samples. *Journal of Voluntary Action Research, 17*(3–4), 90–98.

Blankertz, L., & White, K. K. (1990). Implementation of rehabilitation program for dually diagnosed homeless. *Alcoholism Treatment Quarterly, 7,* 149–164.

Blau, J. (1992). *The visible poor: Homelessness in the United States.* New York: Oxford University Press.

Boesky, L. M., Toro, P. A., & Bukowski, P. A. (1997). Differences in psychosocial factors among older and younger homeless adolescents found in youth shelters. *Journal of Prevention and Intervention in the Community, 15*(2), 19–36.

Boesky, L. M., Toro, P. A., & Wright, K. (1995, October). *Maltreatment in a probability sample of homeless adolescents: A subgroup comparison.* Paper presented at the Annual Meeting of the American Public Health Association, San Diego, CA.

Bukowski, P. A., Toro, P. A., & Wolfe, S. M. (1998). Homeless and matched housed adolescents: A comparative study of psychopathology. *Journal of Clinical Child Psychology, 27*(2).

Burnam, A., & Koegel, P. (1988). Methodology for obtaining a representative sample of homeless persons: The Los Angeles Skid Row Study. *Evaluation Review, 12,* 117–152.

Burt, M. R. (1992). *Over the edge: The growth of homelessness in the 1980s.* New York: Russell Sage Foundation.

Cauce, A. M., Morgan, C. J., Shantinath, S. D., Wagner, V., Wurzbacher, K., Tomlin, S., & Blanchard, T. (1993). Intensive case management for homeless youth: A description of a study in progress. *The Community Psychologist, 26*(2), 32–34.

Cauce, A. M., Morgan, C. J., Wagner, V., Moore, E., Sy, J., Wurzbacher, K., Weeden, K., Tomlin, S., & Blanchard, T. (1994). Effectiveness of intensive case management for homeless adolescents: Results of a 3-month follow-up. *Journal of Emotional and Behavioral Disorders, 2,* 219–227.

Cohen, E. H., Mowbray, C., Bybee, D., Yeich, S., Ribisl, K., & Freddolino, P. P. (1993). Tracking and follow-up methods for research on homelessness. *Evaluation Review, 17,* 331–352.

Dail, P. W. (1990). The psychosocial context of homeless mothers with young children: Program and policy implications. *Child Welfare, 69,* 291–308.

Daly, G. (1990). Programs dealing with homelessness in the United States, Canada, and Britain. In J. Momeni (Ed.), *Homelessness in the United States: Data and issues* (pp. 133–152). New York: Praeger.

Dende, J. D., & Kline, J. D. (1995). Overcoming crack, schizophrenia, and homelessness: A comprehensive case management approach. *New Directions for Mental Health Services, 65,* 53–69.

D'Ercole, A., & Struening, E. (1990). Victimization among homeless women: Implications for service delivery. *Journal of Community Psychology, 18,* 141–152.

Dexter, R. A. (1990). Treating homeless and mentally ill substance abusers in Alaska. *Alcoholism Treatment Quarterly, 7,* 25–30.

Drake, R. E., Osher, F. C., & Wallach, M. A. (1991). Homelessness and dual diagnosis. *American Psychologist, 46,* 1149–1158.

Eaton, W. W., & Kessler, L. G. (1985). *Epidemiologic field methods in psychiatry: The NIMH Epidemiologic Catchment Area Program.* Orlando, FL: Academic Press.

Fischer, P. (1988). Criminal activity among the homeless: A study of arrests in Baltimore. *Hospital and Community Psychiatry, 39,* 46–61.

Fischer, P. J., Shapiro, S., Breakey, W. R., Anthony, J. C., & Kramer, M. (1986). Mental health and social characteristics of the homeless: A survey of mission users. *American Journal of Public Health, 76,* 519–524.

Fisher, P., Wicks, J., Shaffer, D., Piacentini, J., & Lapkin, J. (1992). *Diagnostic interview schedule for children users' manual.* New York: Division of Child and Adolescent Psychiatry, New York State Psychiatric Institute.

Foscarinis, M. (1991). The politics of homelessness: A call to action. *American Psychologist, 41,* 1232–1238.

Fournier, L. (1988). *Rapport synthèse de l'operation d'énumération de la clientèle des missions et refuges de Montréal [Summary report on a census of the clientele of shelters and soup kitchens in Montreal].* Unité de recherche psychosociale, Centre de recherche de l'hopital Douglas, Verdun, Canada.

Fournier, L., & Mercier, C. (Eds.) (1996). *Sans domicile fixe: Au-delà du stéréotype [Homelessness: Beyond the stereotype].* Montreal, Canada: Méridien.

Fournier, L., Mercier, C., Raynault, M., Ohayon, M., & Caulet, M. (1993). Reaching the most destitute of the homeless: When success turns to failure. *Contemporary Drug Problems, 19,* 415–431.

Fournier, L., Toro, P. A., Link, B., Manrique, M., & Glasser, I. (1996, August). *A cross-cultural analysis of homelessness: Prevalence and public opinion.* Paper presented at the 104th Annual Convention, American Psychological Association, Toronto, Canada.

Fox, E. R., & Roth, L. (1989). Homeless children: Philadelphia as a case study. *Annals of the American Academy of Political and Social Sciences, 506,* 141–151.

Freiberg, P. (1991, February). APA *pro bono* program serves homeless people. *APA Monitor, 22*(2), 34.

Furillo, A. (1990, July 19). Homeless face growing hostility in nation's cities. *San Francisco Examiner/Chronicle,* pp. A1, A12.

Gelberg, L., Linn, L. S., Usatine, R. P., & Smith, M. H. (1990). Health, homelessness, and poverty: A study of clinic users. *Archives of Internal Medicine, 150,* 2325–2330.

Gelberg, L., Panarites, C. Andersen, R., & Koegel, P. (1994, November). *Tuberculosis among the homeless in Los Angeles County.* Paper presented at the Annual Meeting, American Public Health Association, Washington, DC.

Gewirtzman, R., & Fodor, I. (1987). The homeless child at school: From welfare hotel to classroom. *Child Welfare, 66,* 237–245.

Goldstein, M. N., & Toro, P. A. (1995, November). *A comparison of veterans and nonveterans in a representative sample of homeless adults.* Presented at the Annual Meeting, American Public Health Association, San Diego, CA.

Goldstein, M., & Toro, P. (1996, August). *Predicting outcomes among homeless adults: A longitudinal analysis.* Paper presented at the 104th Annual Convention, American Psychological Association, Toronto, Canada.

Goodman, L. A. (1991). The prevalence of abuse in the lives of homeless and housed poor mothers: A comparison study. *American Journal of Orthopsychiatry, 61,* 489–500.

Goodman, L., Saxe, L., & Harvey, M. (1991). Homelessness as psychological trauma: Broadening perspectives. *American Psychologist, 46,* 1219–1225.

Gore, A. (1990). Public policy and the homeless. *American Psychologist, 45,* 960–962.

Hannappel, M., Calsyn, R. J., & Morse, G. A. (1989). Mental illness in homeless men: A comparison of shelter and street samples. *Journal of Community Psychology, 17,* 304–310.

Hombs, M. E., & Snyder, M. (1986). *Homelessness in America: A forced march to nowhere.* Washington, DC: The Community for Creative Non-Violence.

HUD (1984). *A report to the Secretary on the homeless and emergency shelters.* Washington, DC: US Department of Housing and Urban Development.

Institute of Medicine (1988). *Homelessness, health, and human needs.* Washington, DC: National Academy Press.

Jaffe, P. G., Wolfe, D. A., & Wilson, S. K. (1990). *Children of battered women.* Newbury Park, CA: Sage.

Jahiel, R. I. (Ed.) (1992). *Homelessness: A prevention-oriented approach.* Baltimore: Johns Hopkins University Press.

Jencks, C. (1994). *The homeless.* Cambridge, MA: Harvard University Press.

Koegel, P., & Burnam, A. (1994, November). *The course of homelessness among adults in Los Angeles.* Paper presented at the Annual Meeting, American Public Health Association, Washington, DC.

Koegel, P., Burnam, A., & Farr, R. K. (1988). The prevalence of specific psychiatric disorders among homeless individuals in the inner city of Los Angeles. *Archives of General Psychiatry, 45,* 1085–1092.

Koegel, P., Burnam, A., & Farr, R. K. (1990). Subsistence adaptation among homeless adults in the inner city of Los Angeles. *Journal of Social Issues, 46,* 83–108.

Koegel, P., Burnam, A., & Morton, S. C. (1996). Enumerating homeless people: Alternative strategies and their consequences. *Evaluation Review, 20,* 378–403.

Koegel, P., Melamid, E., & Burnam, A. (1995). Childhood risk factors for homelessness among homeless adults. *American Journal of Public Health, 85,* 1642–1649.

Koegel, P., Sullivan, G., Burnam, A., Morton, S. C., & Wenzel, S. (1997). *Utilization of mental health and substance abuse services among homeless adults.* Unpublished manuscript, RAND Corporation, Santa Monica, CA.

Kondratas, A. (1991). Housing policy and homelessness in the United States. *American Psychologist, 46,* 1226–1231.

Laurin, I., Fournier, L., Toupin, J., & Ostoj, M. (1996, August). *Homeless and housed men: A study of risk factors.* Paper presented at the 104th Annual Convention, American Psychological Association, Toronto, Canada.

Lee, B., Jones, S. H. & Lewis, D. W. (1990). Public beliefs about the causes of homelessness. *Social Forces, 69,* 253–265.

Lee, B. A., Link, B. G., Toro, P. A. (1991). Images of the homeless: Public views and media messages. *Housing Policy Debate, 2,* 649–682.

Levine, I. S., & Rog, D. J. (1990). Mental health services for homeless mentally ill persons: Federal initiatives and current service trends. *American Psychologist, 45,* 963–968.

Liebow, E. (1993). *Tell them who I am: The lives of homeless women.* New York: Penguin.

Link, B. G., Schwarz, S., Moore, R., Phelan, J., Struening, E., Stueve, A., & Colten, M. E. (1995). Public knowledge, attitudes, and beliefs about homeless people. Evidence for compassion fatigue? *American Journal of Community Psychology, 23,* 533–555.

Link, B. G., Susser, E., Stueve, A., Phelan, J., Moore, R., Struening, E., & Colten, M. E. (1994). Lifetime and five-year prevalence of homelessness in the United States. *American Journal of Public Health, 84,* 1907–1912.

Lozowski, S., Rowland, L. L., & Toro, P. A. (1994, November). *Self-reported health in a probability sample of homeless adults.* Presented at the Annual Meeting, American Public Health Association, Washington, DC.

Manrique, M. A., & Toro, P. A. (1994, November). *National public opinion on homelessness: Is there compassion fatigue?* Paper presented at the Annual Meeting, American Public Health Association, Washington, DC.

Manrique, M. A., & Toro, P. A. (1995, August). *Life-time prevalence of homelessness in the United States.* Paper presented at the 103rd Annual Convention of the American Psychological Association, New York.

Many cities in crackdown on homeless. (1994, December 14). *New York Times,* p. 8.

Media General Research (1989). Codebook: Media General/Associated Press National Poll 23 (November, 1988). Richmond, VA: Author.

Molnar, J. M., Rath, W. R., & Klein, T. P. (1990). Constantly compromised: The impact of homelessness on children. *Journal of Social Issues, 46,* 109–124.

Morse, G., Calsyn, R. J., Allen, G., Tempelhoff, B., & Smith, R. (1992). Experimental comparison of the effects of three treatment programs for homeless mentally ill people. *Hospital and Community Psychiatry, 43,* 1005–1010.

Mundy, P., Robertson, J., Greenblatt, M., & Robertson, M. (1989). Residential instability in adolescent inpatients. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 176–181.

NIAAA (1991). *Synopses of cooperative agreements for research demonstration projects on alcohol and other drug abuse treatment for homeless persons.* Rockville, MD: US Department of Health and Human Services, National Institute on Alcohol Abuse and Alcoholism.

North, C. S., & Smith, E. M. (1992). Posttraumatic stress disorder among homeless men and women. *Hospital and Community Psychiatry, 43,* 1010–1016.

North, C. S., & Smith, E. M. (1993). A comparison of homeless men and women: Different populations, different needs. *Community Mental Health Journal, 29,* 423–431.

Nye, I. F., & Edelbrock, C. (1980). Some social characteristics of runaways. *Journal of Family Issues, 1,* 147–150.

Passero Rabideau, J. M., & Toro, P. A. (1997). Social and environmental predictors of adjustment in homeless children. *Journal of Prevention and Intervention in the Community, 15*(2), 1–17.

Phillips, K. (1990). *The politics of rich and poor: Wealth and the American electorate in the Reagan aftermath.* New York: Random House.

Piacentini, J., Shaffer, D., Fisher, P., Schwab-Stone, M., Davies, M., & Gioia, P. (1993). The Diagnostic Interview Schedule for Children-Revised Version (DISC-R): III. Concurrent Validity. *Journal of The American Academy of Child and Adolescent Psychiatry, 32,* 658–665.

Poll shows majority don't oppose shelters. (1991, March 6). *Washington Post,* p. B2.

Price, R. M., Cowen, E. L., Lorion, R. P., & Ramos-McKay, J. (1988). *Fourteen ounces of prevention: A casebook for practitioners.* Washington, DC: American Psychological Association.

Rafferty, Y., & Shinn, M. (1991). The impact of homelessness of children. *American Psychologist, 46,* 1170–1179.

Regier, D. A, Boyd, J. H., Burke, J. D., Jr., Rae, D. S., Myers, J. K., Kramer, M., Robins, L. N., George, L. K., Karno, M., & Locke, B. Z. (1988). One-month prevalence of mental disorders in the United States: Based on five epidemiologic catchment area sites. *Archives of General Psychiatry, 45,* 977–986.

Ringwalt, C. L., Greene, J. M., & Iachan, R. (1994, November). *Prevalence and characteristics of youth in households with runaway and homeless experience.* Paper presented at the meeting of the American Public Health Association, Washington, DC.

Robertson, M. J. (1990, August). *Characteristics and circumstances of homeless adolescents in Hollywood.* Paper presented at the Annual Convention of the American Psychological Association, Boston, MA.

Robertson, M. J. (1991). Homeless youth: An overview of the recent literature. In J. H. Dryder-Coe, L. M. Salamon, & J. M. Molnar (Eds.), *Homeless children and youth: A new American dilemna* (pp. 33–68). Baltimore: Johns Hopkins University Press.

Robertson, M. J. (1991). The prevalence of mental disorder among homeless people. In R. I. Jahiel (Ed.), *Homelessness: A prevention-oriented approach* (pp. 57–86). Baltimore: Johns Hopkins University Press.

Robertson, M. J., & Cousineau, M. R. (1986). Health status and access to health services among the urban homeless. *American Journal of Public Health, 76,* 561–563.

Robertson, M. J., Koegel, P., & Ferguson, L. (1989). Alcohol use and abuse among homeless adolescents in Hollywood. *Contemporary Drug Problems, 16,* 415–452.

Robertson, M. J., Zlotnick, C., & Westerfelt, A. (1994, November). *The course of homelessness among adults and families in Alameda County.* Paper presented at the Annual Meeting, American Public Health Association, Washington, DC.

Robertson, M. J., Zlotnick, C., & Westerfelt, A. (1997). Drug disorders and treatment contact among homeless men and women in Alemeda County. *American Journal of Public Health, 87,* 221–228.

Robins, L. N., Helzer, J. E., Croughan, J., & Ratcliff, K. S. (1981). National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics, and validity. *Archives of General Psychiatry, 38,* 381–389.

Roll, C. N., Toro, P. A., & Ortola, G. L. (1995, November). *Characteristics and experiences of homeless adults: A comparison of single men, single women, and women with children.* Paper presented at the Annual Meeting, American Public Health Association, San Diego, CA.

Rosenheck, R., Gallup, P., & Frisman, L. K. (1993). Health care utilization and costs after entry into an outreach program for homeless mentally ill veterans. *Hospital and Community Psychiatry, 44,* 1166–1171.

Rosenheck, R., & Koegel, P. (1993). Characteristics of veterans and nonveterans in three samples of homeless men. *Hospital and Community Psychiatry, 44,* 858–863.

Rossi, P. H. (1989). *Down and out in America: The origins of homelessness.* Chicago: University of Chicago Press.

Rowland, L. L., & Toro, P. A. (1996, August). *Predicting health outcomes among the homeless: An 18-month* *longitudinal study*. Paper presented at the 104th Annual Convention, American Psychological Association, Toronto, Canada.

Seidner, A. L., Burling, T. A., Fisher, L. M., & Blair, T. R. (1990). Characteristics of telephone applicants to a residential rehabilitation program for homeless veterans. *Journal of Consulting and Clinical Psychology, 58,* 825–831.

Shaffer, D., & Caton, C. L. M. (1984). *Runaway and homeless youth in New York City: A report to the Ittleson Foundation*. New York: New York State Psychiatric Institute and Columbia University.

Shaffer, D., Schwab-Stone, M., Fisher, P. C., Piacentini, J., Davies, M., Conners, C. K., & Regier, D. (1993). The Diagnostic Interview Schedule for Children-Revised version (DISC-R): Preparation, field testing, interrater reliability, and acceptability. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 643–650.

Shift in feelings on the homeless: Empathy turns to frustration. (1991, September 2) *New York Times,* p. 1.

Shinn, M. (1992). Homelessness: What is a psychologist to do? *American Journal of Community Psychology, 20,* 1–24.

Shinn, M. (1996, August). *Family homelessness: State or trait*. Paper presented at the 104th Annual Convention, American Psychological Association, Toronto, Canada.

Shinn, M., Burke, P. D., & Bedford, S. (1990). Homelessness: Abstracts of the psychological and behavioral literature, 1967–1990. *Bibliographies in Psychology Series, 7.* Washington, DC: American Psychological Association.

Shinn, M., Knickman, J., & Weitzman, B. C. (1991). Social relationships and vulnerability to becoming homeless among poor families. *American Psychologist, 46,* 1180–1187.

Shinn, M., & Weitzman, B. C. (1990). Research on homelessness: An introduction. *Journal of Social Issues, 46,* 1–11.

Shlay, A. B., & Rossi, P. H. (1992). Social science research and contemporary studies of homelessness. *Annual Review of Sociology, 18,* 129–160.

Smart, R. G., Adlaf, E. M., Walsh, G. W., & Zdanowicz, Y. (1994). Similarities in drug use and depression among runaway students and street youth. *Canadian Journal of Public Health, 85,* 17–18.

Smith, E. M., & North, C. S. (1994). Not all homeless women are alike: Effects of motherhood and the presence of children. *Community Mental Health Journal, 30,* 601–610.

Snow, D. A., Baker, S. G., Anderson, L., & Martin, M. (1986). The myth of pervasive mental illness among the homeless. *Social Problems, 33,* 407–423.

Sosin, M. R. (1992). Homeless and vulnerable meal program users: A comparison study. *Social Problems, 39,* 170–188.

Stayanoff, S. R., Weber, M. D., Wiley, D., Ford, W. L., & Kerndt, P. R. (1994, November). *HIV prevalence and associated risk factors among homeless patients in an inner city clinic in Los Angeles*. Paper presented at the Annual Meeting, American Public Health Association, Washington, DC.

Susser, E., Conover, S., & Struening, E. L. (1989). Problems of epidemiologic method in assessing the type and extent of mental illness among homeless adults. *Hospital and Community Psychiatry, 40,* 261–265.

Susser, E., Struening, E. L., & Conover, S. (1987). Childhood experiences of homeless men. *American Journal of Psychiatry, 144,* 1599–1601.

Toro, P. A., Alan, K., Saccarelli, T., Saccarelli, E. G., Lombardo, S. A., & Pienkowski, J. (1995, June). *A cross-cultural comparison of homelessness across developed nations*. Presented at the Fifth Biennial Conference on Community Research & Action, Chicago.

Toro, P. A., Bellavia, C., Daeschler, C., Owens, B., Wall, D. D., Passero, J. M., & Thomas, D. M. (1995). Distinguishing homelessness from poverty: A comparative study. *Journal of Consulting and Clinical Psychology, 63,* 280–289.

Toro, P. A., & Bukowski, P. A. (1995). Homeless adolescents: What we know and what can be done. *NMHA Prevention Update, 6*(1), 6–7.

Toro, P. A., & McDonell, D. M. (1992). Beliefs, attitudes, and knowledge about homelessness: A survey of the general public. *American Journal of Community Psychology, 20,* 53–80.

Toro, P. A., Rabideau, J. M. P., Bellavia, C., Daeschler, C., Wall, D. D., Thomas, D. M., & Smith, S. J. (1997). Evaluating an intervention for homeless persons: Results of a field experiment. *Journal of Consulting and Clinical Psychology, 65,* 476–484.

Toro, P. A., & Rojansky, A. (1990). Homelessness: Some thoughts from an international perspective. *Community Psychologist, 23*(4), 8–11.

Toro, P. A., Trickett, E. J., Wall, D. D., & Salem, D. A. (1991). Homelessness in the United States: An ecological perspective. *American Psychologist, 46,* 1208–1218.

Toro, P. A., & Wall, D. D. (1991). Research on homeless persons: Diagnostic comparisons and practice implications. *Professional Psychology: Research and Practice, 22,* 479–488.

Toro, P. A., & Warren, M. G. (1991). Homelessness, psychology, and public policy: Introduction to section three of the special issue on homelessness. *American Psychologist, 46,* 1205–1207.

Toro, P. A., Wolfe, S. M., Bellavia, C., Thomas, D. M., Rowland, L. L., Daeschler, C. V., & Bukowski, P. A. (1998). *Obtaining representative samples of homeless persons: A two-city study*. Unpublished manuscript, Department of Psychology, Wayne State University.

US Conference of Mayors (1989). *A status report on hunger and homelessness in America's cities: A 27-city survey*. Washington, DC: US Conference of Mayors.

Valencia, E., Saez, H., Susser, E., Conover, S., & Colson, P. (1994, November). *HIV and TB prevalence among mentally ill men in NYC shelter*. Paper presented at the Annual Meeting, American Public Health Association, Washington, DC.

Wagner, J., & Menke, E. (1990, November). *The mental health of homeless children*. Paper presented at the meeting of the American Public Health Association, New York.

Washington Post (1991). Poll shows majority don't oppose shelters. *Washington Post, 6 March,* B2.

Windle, M. (1989). Substance use and abuse among adolescent runaways: A four-year follow-up study. *Journal of Youth and Adolescence, 18,* 331–344.

Wood, D., Valdez, R. B., Hayashi, T., & Shen, A. (1990). Families in Los Angeles: A study comparing demographic, economic, and family function characteristics. *American Journal of Public Health, 80,* 1049–1052.

Wolfe, S. M. (1993, August). *Structural equation modeling among homeless adults: Stress, support, and self-esteem*. Paper presented at the 101st Annual Convention, American Psychological Association, Toronto, Canada.

Wolfe, S. M. (1994, August). *A comparative study of housed and homeless adolescents*. Paper presented at the 102nd Annual Convention, American Psychological Association, Los Angeles.

Wright, J. D. (1989). *Address unknown: The homeless in America*. New York: Aldine de Gruyter.

Wright, J. D. (1990). Homelessness is not healthy for children and other living things. *Child & Youth Services, 14,* 65–88.

Wright, J. D., Allen, T. L., & Devine, J. A. (1993, August).

Tracking non-traditional populations in longitudinal studies. Paper presented at the 101st Annual Convention, American Psychological Association, Toronto, Canada.

Wright, J. D., & Lam, J. A. (1987). Homelessness and the low-income housing supply. *Social Policy, 17(4),* 49–53.

Wright, J. D., & Weber, E. (1987). *Homelessness and health*. New York: McGraw-Hill.

Yates, G., MacKenzie, R., Pennbridge, J., & Cohen, E. (1988). A risk profile comparison of runaway and non-runaway youth. *American Journal of Public Health, 78,* 820–821.

Zigier, E., & Styfco, S. J. (1990). Head Start: Criticism in a constructive context. *American Psychologist, 49,* 127–132.

Zolopa, A. R., Hahn, J. A., Gorter, R., Miranda, J., Wlodarczyk, D., Peterson, J., Pilote, L., & Moss, A. R. (1994). HIV and tuberculosis infection in San Francisco's homeless adults: Prevalence and risk factors in a representative sample. *Journal of the American Medical Association, 272,* 455–461.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.09

Individuals with Traumatic Brain Injury

STEPHEN R. HOOPER
University of North Carolina, Chapel Hill, NC, USA

9.09.1 INTRODUCTION 138

9.09.2 DEFINITIONAL ISSUES 138
- *9.09.2.1 Alteration in Level of Consciousness* 138
- *9.09.2.2 Degree of Post-traumatic Amnesia* 138
- *9.09.2.3 Presence of Physical Damage* 139

9.09.3 EPIDEMIOLOGICAL CHARACTERISTICS 139
- *9.09.3.1 Incidence Rates and Related Factors* 139
- *9.09.3.2 Major Causes of Traumatic Brain Injury* 140
- *9.09.3.3 Related Factors* 140

9.09.4 NEUROPATHOLOGICAL MECHANISMS 140
- *9.09.4.1 Acceleration-dependent Injuries* 141

9.09.5 TYPES OF INJURIES 141
- *9.09.5.1 Penetrating Head Injuries* 141
- *9.09.5.2 Nonpenetrating Head Injuries* 141
- *9.09.5.3 Medical Complications* 142
  - *9.09.5.3.1 Hypoxia/ischemia* 142
  - *9.09.5.3.2 Cerebral edema* 142
  - *9.09.5.3.3 Hemorrhages* 142
  - *9.09.5.3.4 Cerebral atrophy* 143
  - *9.09.5.3.5 Post-traumatic epilepsy* 143

9.09.6 NEUROCOGNITIVE AND PSYCHOSOCIAL OUTCOMES 143
- *9.09.6.1 Neurocognitive Outcomes in Penetrating Head Injuries* 143
- *9.09.6.2 Neurocognitive Outcomes in Nonpenetrating Head Injuries* 144
  - *9.09.6.2.1 Intellectual* 144
  - *9.09.6.2.2 Motor/visual–motor* 145
  - *9.09.6.2.3 Attention* 145
  - *9.09.6.2.4 Language* 145
  - *9.09.6.2.5 Memory* 145
  - *9.09.6.2.6 Academic* 146
- *9.09.6.3 Psychosocial Outcomes* 146
  - *9.09.6.3.1 Impact on the family* 147

9.09.7 ISSUES RELATED TO OUTCOMES 147
- *9.09.7.1 Age at Time of Injury* 147
- *9.09.7.2 Recovery Rates* 148
- *9.09.7.3 Premorbid Functioning* 148

9.09.8 INTERVENTION 148
- *9.09.8.1 Cognitive Rehabilitation* 149
- *9.09.8.2 Educational Programming* 149

9.09.9 SUMMARY 150
9.09.10 REFERENCES 150

9.09.1 INTRODUCTION

Neurological involvement of any kind can have a significant impact on the overall functioning of an individual. This is especially possible in the case of traumatic brain injury (TBI) where the sudden nature of the injury and its severity can alter the developmental progression, expectations, and environmental arena of an individual forever. Recognizing the possible impacts of a TBI can assist the clinical psychologist and other practitioners in developing assessment, treatment, and tracking programs for these individuals; however, most professionals, particularly clinical child psychologists, are ill-prepared to manage such cases (Cooley & Singer, 1991).

This chapter discusses issues pertinent to TBI, including definitional issues, epidemiology, specific causes of TBI, associated risk factors, and outcomes. Selected intervention issues are also presented. Although a lifespan approach to TBI is put forth, this chapter focuses much of its discussion on children and adolescents.

9.09.2 DEFINITIONAL ISSUES

The National Head Injury Foundation (NHIF, 1985) defines a TBI as "an insult to the brain, not of a degenerative or congenital nature, but caused by an external force, that may produce a diminished or altered state of consciousness." This definition addresses most of the core issues pertinent to TBI. First, the definition states that this is an acquired problem, and consideration should be given to the many dynamics that can surround such an injury, as opposed to one of a degenerative or congenital nature (e.g., a specific learning disability). Second, it states that a TBI must be caused by an external force. Although a brain insult can occur via any number of neurological mechanisms (e.g., a cerebrovascular accident, tumors, epileptogenic brain activity), it is important to recognize that a TBI arises only in the case where an external force is involved. This external force does not necessarily have to be a direct blow to the head. The effects of acceleration–deceleration forces, as can be the case in a shaken infant or in a motor vehicle accident, may be the external force which creates the brain injury without actual physical contact (Jennett, 1986). Finally, the definition addresses the issue of change in functioning when compared to preinjury status. This is an extremely important issue with respect to the assessment and treatment course following a TBI.

Despite the inclusion of these core issues, the NHIF definition remains simplistic with respect to addressing the myriad of issues invoked in defining TBI. Indeed, Levin, Benton, and Grossman (1982) reviewed several systematic attempts to classify different types of traumatic injuries, but each one has been challenged on its ability to reflect actual brain injury, particularly in mild cases. At present, there are at least three key variables that appear pertinent to the definition of TBI: an alteration in the level of consciousness, the degree of post-traumatic amnesia (PTA), and the presence of neurophysiological, neuroanatomical damage, and/or some other type of physical damage (Bigler, 1990).

9.09.2.1 Alteration in Level of Consciousness

As noted in the NHIF definition, an altered state of consciousness is one of the key features of TBI. Although it is recognized that such changes can be common sequelae of most TBIs, variants to this corollary are quite common. For example, such alterations do not always appear with some brain injuries (Jennett, 1986; Levin et al., 1982), as can be the case with selected types of focal injuries. In other cases, the appearance of the alterations may be temporally delayed (e.g., edema, hematoma). Conversely, a brief alteration in the state of consciousness with quick and total recovery of functioning, as might be seen in a concussion (Bakay & Glasauer, 1980), may not be sufficient to invoke the inference of a TBI. Consequently, care should be taken when making a direct linkage between the presence/absence of an altered state of consciousness and a possible brain injury.

9.09.2.2 Degree of Post-traumatic Amnesia

The degree of PTA has also been invoked as a useful variable for defining the severity of a TBI. PTA is the amount of time following a brain injury that an individual experiences difficulties in learning and retaining new information. It is the time following an injury when the individual is alert and functioning (i.e., not comatose), and capable of responding in a relatively reliable manner. PTA ends when the individual is oriented and continuous memories are restored (Rosen & Gerring, 1986). Although there may also be significant retrograde amnesia (i.e., lack of recall for events prior to the accident), PTA has been deemed useful for defining the severity of a TBI. Brooks (1983) noted that a severe PTA of at least one week or more has been related to poorer cognitive and psychosocial outcomes.

9.09.2.3 Presence of Physical Damage

A final variable that must be included in conceptualizing a definition of TBI is whether there is any actual physical damage that can be identified in the brain. This physical damage can be in the form of: neurophysiological evidence, as can be uncovered with electroencephalographic assessment strategies; neuroanatomical evidence, as depicted with a variety of neuroimaging procedures such as computerized axial tomography (CT scan), magnetic resonance imaging (MRI), or functional magnetic resonance imaging (fMRI); and/or in the form of direct physical examination (e.g., specific neuropsychological deficits, speech/language deficits, hemiparesis, cortical blindness). Findings from these procedures, when used in tandem with the presence of PTA and the degree of consciousness, offer a powerful multidimensional approach to defining TBI.

9.09.3 EPIDEMIOLOGICAL CHARACTERISTICS

The complexity of the definitional issues relevant to TBI also contributes to what is known about the epidemiology of TBI. In fact, Kraus (1987) cited the lack of consistency across studies in the definition of a head injury and how this hindered the ascertainment of an accurate estimate of TBI cases. Further, Goldstein and Levin (1987) noted that many potential brain injuries have been diagnosed by the cause (e.g., fall, motor vehicle accident) rather than by the nature of the injury. These investigators also cited the pitfalls attached to developing incidence estimates based on hospital admission records and telephone surveys. These methodological issues contribute to the difficulties in gaining an accurate estimate of the incidence and prevalence of traumatic brain injury across the life span.

Despite these obstacles, it has been estimated that TBIs represent the most frequent neurological conditions that result in hospitalization of children and adolescents under 19 years of age, and are a major contribution to mortality and morbidity in children and adolescents (Frankowski, 1985). In general, TBI represents a major health problem because it is the leading cause of death or permanent disability in children and adolescents (Guyer & Ellers, 1990).

9.09.3.1 Incidence Rates and Related Factors

There have been a number of well-conducted epidemiological studies to examine issues pertinent to TBI (Annegers, 1983; Guyer & Ellers, 1990; Hahn et al., 1988; Ivan, Choo, & Ventureya, 1983; Kraus, Fife, Cox, Ramstein, & Conroy, 1986). Annegers, Grabow, Kurland, and Laws (1980) conducted one of the first epidemiological studies which employed clear diagnostic criteria to ascertain incidence rates in Olmsted County, MN. They obtained an overall incidence rate of 386 per 100 000, with males (270/100 000) showing over a 2:1 ratio when compared to females (116/100 000). In addition to the gender differences, specific age trends also were observed in the data. Specifically, Annegers et al. found that the incidence of TBI increased significantly in the 15–24 years age range, particularly for males. More specifically, males showed a heightened rate from ages 5–25 years, whereas females demonstrated a decline from age 3 years through to age 15 years. There were also second peaks which occurred during infancy and at the age of 70 years and older. Similar trends were obtained by Kraus et al. (1984) in their study of medical records of patients who died or who were admitted to San Diego County Hospital in 1981.

When taken together, Goldstein and Levin (1987) and Frankowski (1985) reported incidence rates of: 150/100 000 for ages birth to 4 years; 550/100 000 for ages 15–19; 160/100 000 around age 50; and 200/100 000 for age 65 and beyond. Kraus, Rock, and Hemyari (1990) also reported different incidence rates for different degrees of TBI severity for individuals ranging in age from birth to age 15 years. These rates indicated that about 5% of TBIs are fatal, whereas 6%, 8%, and 81% were described as severe, moderate, and mild, respectively. In addition to differences reported in the age, gender, and severity of the TBI, information has begun to emerge with respect to the impact of race and socioeconomic status on the incidence rates of TBI. Studies conducted by Kraus (1987), Rivara and Mueller (1986), and Cooper et al. (1983) revealed a higher rate of TBI in minority populations when compared to Whites. As Fennell and Mickle (1992) noted, however, these studies generally failed to take into account the effects of socioeconomic status on the incidence rates of TBI. In fact, Kraus (1987) reported that the highest rates of TBI were observed in sociocultural groups with the lowest median income, suggesting that socioeconomic factors should be considered when examining incidence rates of TBI.

Selected risk factors have also been discussed which might increase the chance of a TBI in selected populations. For example, Rutter, Chadwick, Shaffer, and Brown (1980) noted that these individuals may not represent a random cut of the general population. They noted that their mildly injured group of children showed a higher rate of psychosocial problems, that probably preceded the TBI. Similarly, Craft, Shaw, and Cartlidge (1972) found an increased rate of teacher-reported problems including hyperactivity, depression, and antisocial behaviors that predated the brain injury. Premorbid developmental difficulties and specific learning problems (Klonoff & Paris, 1974), language problems (Mahoney et al., 1983), and lower academic achievement (Chadwick, Rutter, Brown, Shaffer, & Traub, 1981) have also been reported for pediatric TBI samples. Finally, recent evidence does suggest that an individual who sustains a TBI is at increased risk, perhaps because of the sequelae incurred from the first TBI, for sustaining a second brain injury. Annegers et al. (1980) reported that this risk was age-related, with risk being increased twofold under age 14 years, threefold between the ages of 15 and 24 years, and fivefold after age 25 years. Males also had twice the likelihood of sustaining a second TBI as females.

9.09.3.2 Major Causes of Traumatic Brain Injury

Along with factors related to gender, race, socioeconomic status, and chronological age, specific causes of TBIs have been charted. For adults, motor vehicle accidents account for nearly half of all injuries, with motorcycle accidents tending to be about four times more common than other types of motor vehicle accidents. Falls are the second leading cause of TBI (Kalsbeek, McLaurin, Harris, & Miller, 1980). For children and adolescents, accidents in the home account for the major number of head injuries in preschool children, although child abuse is rapidly gaining in this age band (Christoffel, 1990). Rivara (1984) observed that infants and preschoolers may be at risk for head vs. trunk and extremity injuries, due to the disproportionate size of the child's head at these ages and the associated higher center of gravity. Falls, pedestrian–motor vehicle accidents, bicycle–motor vehicle accidents, and sporting activities comprise the majority of mild TBIs in school-age children (Frankowski, 1985). Similar to adult epidemiological findings, motor vehicle accidents constitute the majority of severe TBI cases for older adolescents (Rutter et al., 1980).

9.09.3.3 Related Factors

In general, the causes of TBI tend to change with age and may contribute to the actual pathophysiology of the TBI. For example, more severe brain injuries tend to be incurred by older adolescents and adults secondary to high-speed motor vehicle accidents. In contrast, Kalsbeek et al. (1980) reported that a similar level of severity in a TBI incurred from a fall may be created by compression of cortical and subcortical regions secondary to intracerebral hematoma, as opposed to frank brain damage. Similarly, a TBI resulting from a fall may develop intracranial hematomas (DiRocco & Velardi, 1986), whereas a TBI resulting from a motor vehicle accident usually results in concussion. Temporal fluctuations have also been reported with respect to the specific causes of TBI. These fluctuations have been seen with respect to specific times of the day (Levin et al., 1982), specific days (Edna, 1987), and specific seasons of the year depending on region of the country (Klauber et al., 1981). For example, the likelihood of a child or adolescent incurring a TBI increases from the time that school ends to the time they go to bed (Levin et al., 1982). Falls are most likely to occur between noon and 6.00 p.m., and motor vehicle accidents between 6.00 p.m. and 9.00 p.m. (Jagger, Levine, Jane, & Rimel, 1984). TBIs tend to occur with a higher frequency on weekends, with peak days being Friday and Saturday (Edna, 1987). Geographically, where weather patterns are more stable, stable rates of TBI are observed (Klauber et al., 1981); however, in regions of the country where more variable weather patterns are observed, an increased rate of TBIs are observed during the spring and summer months (Edna, 1987).

9.09.4 NEUROPATHOLOGICAL MECHANISMS

The mechanisms involved in a TBI occur as a result of three physical forces: tension (i.e., the tearing apart of tissues), compression (i.e., the pushing together of brain tissue), and shearing (i.e., the sliding of one portion of tissue over another). These physical forces combine to describe injuries from acceleration-dependent (i.e., acceleration, deceleration, rotation) and nonacceleration-dependent factors (e.g., many crushing types of injuries). The latter types of injuries are relatively rare, with much of the energy being evenly distributed and typically causing skull fractures, but relatively little brain damage in many instances. In these cases, the amount of damage is largely dependent upon the degree of skull deformation. Given the relative rarity of these types of injuries, acceleration-dependent injuries are the focus of this discussion.

9.09.4.1 Acceleration-dependent Injuries

Denny-Brown and Russell (1941) demonstrated that a moderate blow to a movable head could produce severe brain damage, but a blow with 20 times the strength could be sustained by a rigidly fixed head with relatively little brain damage being incurred. Although clinical situations involving a rigidly fixed head are infrequent, acceleration forces typically have been viewed as the dominant factors involved in TBI, particularly closed head injuries (Pang, 1985).

Acceleration injuries occur when the slower moving brain tissue is damaged or rendered dysfunctional by a sudden external force (e.g., from a punch, baseball, or rock). The brain tissue is of a different density than bone, and consequently it will accelerate at a slower rate than bone. As the external force hits the head, the skull moves in a direction opposite the impact, whereas the brain tissue initially moves in the direction of the impact and then follows the direction of the skull. In contrast to pure accelerating forces, where an object strikes a stationary head, deceleration forces occur when the moving head strikes a stationary object. In these instances, such as in motor vehicle accidents, the head is suddenly stopped with resultant damage occurring at the impact site (i.e., coup) as well as at the opposite site (i.e., contrecoup).

Rotational injuries occur as a result of hyperflexion, hyperextension, lateral flexion, and turning movements of the head on the neck and brainstem. These rotational or torquing forces can produce shearing and tearing of cerebral tissue, with some investigators attributing most contrecoup injuries to these forces (Ommaya, Grubbs, & Naumann, 1971). All of these mechanisms can result in the rupturing of superficial and deep blood vessels, and the subsequent development of different kinds of hemorrhages.

9.09.5 TYPES OF INJURIES

A trauma to the head does not always lead to a brain injury. To be of significance to the neurologist, psychologist, or other child or adult practitioner, a head trauma sufficient to produce a disruption of the central nervous system must have been incurred. There are several types of brain injuries that can be seen, with two broad classes generally being accepted: penetrating wounds and nonpenetrating injuries. A possible third category, skull fractures, will be subsumed under these two broad classes.

9.09.5.1 Penetrating Head Injuries

These kinds of injuries occur when an external object strikes the head with sufficient force to cause a depressed skull fracture, whereby skull fragments either tear the dura or cause a brain laceration. About 30% of all skull fractures reportedly are of a depressed nature, with compound fractures being present in about 20% of individuals with severe head trauma (Reitan & Wolfson, 1985)

A second kind of penetrating head injury occurs when a missile penetrates the skull. Such a penetration can lead to the missile being lodged in the brain tissue or passing through it. Some individuals may show an immediate loss of consciousness, whereas others may not lose consciousness at the time of injury, but gradually progress into a coma over several hours. If the injury is severe enough, then increased intracranial pressure and decreased blood pressure may be seen which may create a medical emergency. Brainstem dysfunctions (e.g., poor arousal modulation) also can be observed, and are intimately related to coma states.

In penetrating types of brain injuries, both focal and diffuse effects can be observed. Tearing of the dura, and subsequent damage to the underlying vasculature of the dura and meninges, along with focal contusions, can be observed in such cases. Further, the depressed skull fragments can cause tearing of the brain tissue directly underlying the depression (Dacey, Alves, Rimel, Winn, & Jane, 1986; Shapiro, 1987). Diffuse effects, such as stretching and shearing of the brain tissue, as well as contrecoup contusion, may also occur with penetrating types of brain injuries (Fennell & Mickle, 1992). Brain contusions (i.e., bruising), are most often seen in depressed skull fractures (Shapiro, 1987).

9.09.5.2 Nonpenetrating Head Injuries

In contrast to the penetrating types of injuries, nonpenetrating head injuries do not involve any actual penetration of foreign material into the brain. These types of injuries can occur when the trauma to the head does not produce a skull fracture, or it generates a nondepressed or linear skull fracture. This type of involvement includes the closed head injuries and concussions and, in general, accounts for over 90% major pediatric head injuries (Menkes, 1990).

Although a head injury can occur at any time in life, infants and young children are at particular risk for nonpenetrating closed head injuries because of specific skull–brain interface factors. For example, at birth there are six spaces, or fontanelles, between the neonatal skull bones. These remain unfused in order to accommodate the increasing size of the brain. Five of these fontanelles fuse by the age of 6 weeks, but the most anterior one does not fuse until about 9–16 months of age. As one might suspect, the infant brain is particularly susceptible to damage during this time period, although the flexibility of the immature skull can absorb a greater degree of deformation before resulting in a fracture (Menkes, 1990). This absorption factor, however, also can contribute to an increased amount of generalized shearing damage to the brain.

The actual interface between the skull and the brain also contributes to the different kinds of brain damage that can be seen. The brain tends to be more tightly packed in the posterior regions of the brain and more loosely packed in anterior regions. In addition, because the brain is more tightly packed in posterior regions, the internal surface of the skull tends to be smoother, whereas an increased number of bony protuberances can be found in the anterior regions. Consequently, the frontal and temporal lobes, particularly anterior temporal regions, are the cortical areas most likely to be damaged, regardless of the site or direction of the initial impact. Given this skull–brain interface, frontotemporal lacerations and contusions are frequently seen in all types of brain injuries, particularly nonpenetrating head injuries (Bigler, 1990). These injuries can result in memory disorders, executive functioning deficits, and a variety of emotional changes (Levin et al., 1982; Salazar et al., 1986).

Another primary finding in head trauma of all types is diffuse axonal injuries. Much neuropathological evidence suggests that the predominant impact of TBI is of a diffuse, nonspecific nature (Adams, Graham, & Gennarelli, 1985; Jennett, 1986). This is especially evident in cases of closed head injury. Even when focal deficits are evidenced, they are typically present in the face of more diffuse damage or dysfunction. These diffuse findings have been related to the effects of neuronal shearing and tearing (Auerbach, 1986). The diffuse damage that occurs results from the twisting, tearing, or breaking of the axonal fibers, damage to the connective supportive cells of the brain, or neuronal degeneration, these processes are referred to as diffuse axonal injury. Diffuse axonal injury can occur anywhere in the brain, but typically occurs in the brainstem, deep white matter structures, and in the frontotemporal regions.

9.09.5.3 Medical Complications

In addition to the above types of TBI, there are a number of medical complications that can ensue. Although many of these concerns are addressed prior to an individual leaving the hospital, it is important to be familiar with these complications as they may affect an individual's physical, neurocognitive, and general adaptive behavior functioning. These include ischemia/hypoxia, cerebral edema, hemorrhage, cerebral atrophy, ventricular enlargement, and post-traumatic epilepsy.

9.09.5.3.1 Hypoxia/ischemia

In addition to diffuse axonal injury, neuronal damage can also occur via obstructed blood flow (i.e., ischemia) and/or poorly oxygenated blood (i.e., hypoxia). Frequently, a loss of blood pressure to the brain can be seen, perhaps secondary to a significant loss of blood, and when this pressure falls to a certain critical point, ischemic hypoxia can occur. This process can contribute to brain tissue damage or destruction, particularly in the small vessels of the brain. In addition to the diffuse effects that can be seen with ischemic hypoxia, anoxic-type damage may disrupt specific structures as well (e.g., the hippocampus).

9.09.5.3.2 Cerebral edema

This is the most common secondary effect of TBI. Cerebral edema, or brain swelling, can create focal findings, such as with a lateralized contusion, or it can produce more generalized findings (Yoshino, Yamaki, Higuchi, Horikawa, & Hirakawa, 1985). Prolonged cerebral edema can compress blood vessels feeding the brain and lead to an infarction. It also can compress actual brain tissue which may lead to cell dysfunction or destruction (Ito, Tomita, Yamazaki, Takada, & Inaba, 1986). Children are at a greater risk than adults to develop diffuse cerebral swelling following a TBI. Shapiro (1987) estimated that about 30% of conscious head-injured children, and about 40% of comatose children, show diffuse cerebral edema CT Scan. With the possibility of tissue damage and destruction, concomitant neuropsychological sequelae typically can be seen.

9.09.5.3.3 Hemorrhages

Hemorrhages are bleeds that can occur in nearly any place. An extradural hematoma is a collection of blood between the skull and the dura which develops at the site of the trauma. It occurs in about 1% of children hospitalized for TBI and typically is unilateral in its appearance. The severity of the symptoms is associated, in part, with the size of the hematoma and the speed of its evolution. In children, there is typically a period of disorientation or loss of consciousness, but neurologic signs appear and a progressive loss of consciousness ensues minutes to days after evolution of the hematoma (Mealey, 1987). In general, the younger the child, the longer the latency.

A subdural hematoma is a localized collection of blood between the dura and the cerebral mantle, and typically has neurosurgical implications due to increased intracranial pressure. It typically is unilateral in its appearance (Menkes, 1990) and is responsible for death and physical injury in a significant number of battered or shaken babies (Caffey, 1974). It is important to note that a hematoma may form at the time of the injury, or sometime thereafter, so it might not be an immediate complication from TBI (Bucci, Phillips, & McGillicuddy, 1986; Soloniuk, Pitts, Lovely, & Bartkowski, 1986). In addition, the presence of any type of hematoma, regardless of location, may exacerbate underlying cerebral dysfunction or damage, and might be related to increased risk for neurocognitive deficits (Cullum & Bigler, 1986).

9.09.5.3.4 Cerebral atrophy

A TBI can cause cerebral atrophy, which typically corresponds to enlargement of the ventricular system and subsequent hydrocephalus (Cullum & Bigler, 1986; Lipper et al., 1985). Cullum and Bigler (1986) noted that this atrophy tends to be seen more in the frontotemporal regions, although it can also present in a diffuse fashion, with the greatest atrophy being associated with areas of contusion and/or prior hematoma. White matter degeneration has been described in a variety of age groups, and typically results from retraction of axonal pathways following injury (Graham, Adams, & Gennarelli, 1987). Results from neuroimaging procedures have suggested that changes in ventricular size may be more sensitive to structural deficits than the actual presence of cortical atrophy (Bigler, 1994).

9.09.5.3.5 Post-traumatic epilepsy

Seizures that are associated with TBI have been classified as to their time of onset. Immediate types of seizures occur within a few seconds of the head trauma. Menkes (1990) suggested that these types of seizures are the result of direct mechanical stimulation of brain tissue having a low seizure threshold. Early types of seizures typically appear during the first 24–48 hours following a brain injury. In general, these are due to cerebral edema, intracranial hemorrhage, contusion, laceration, or actual tissue necrosis. Children are more prone to developing seizures in the immediate and early postinjury period than adults, even after a minor TBI, although the overall incidence of early post-traumatic seizures is about 7% (Jennett, 1975).

Late post-traumatic seizures tend to emerge within the first two years following the injury, with about 50% of these occurring within the first year. These seizures typically have an origin in cerebromeningeal scarring. Further, there was a 5–10% incidence of post-traumatic epilepsy in children who experienced a loss of consciousness for at least one hour following their TBI, compared to a 2% occurrence when there was not a significant loss of consciousness. This incidence rate increased to 30% when a brain laceration was sustained, and doubled when the child experienced a post-traumatic amnesia of more than 24 hours. The topography of these seizures can take any form but, as a rule, generally is not manifested in absence form. It has been estimated that these seizures become less frequent after the third year following the injury in about 20–50% of patients, but that the individual will remain prone to a seizure disorder (Jennett, 1975). Finally, it must be remembered that many of the anticonvulsant medications also maintain their own liabilities with respect to their impact on neurocognitive functioning (e.g., phenytoin and phenobarbital can both cause and/or contribute to deficits on neuropsychological testing).

9.09.6 NEUROCOGNITIVE AND PSYCHOSOCIAL OUTCOMES

The range of neurocognitive and psychosocial outcomes for any particular TBI will be highly varied. In addition, there may be any number of medical outcomes that will interfere with the accurate assessment of an individual's functioning (e.g., post-traumatic seizures, spasticity, ataxia). This section focuses on the neurocognitive and psychosocial outcomes that might be seen in children and adolescents who have sustained a TBI.

9.09.6.1 Neurocognitive Outcomes in Penetrating Head Injuries

When the brain is damaged, deficits stem from injury to those areas of the brain that govern the involved functions. In penetrating head injuries, such as in an open head injury from a missile, discrete brain areas in the path of injury typically will be affected with largely predictable deficits. For example, if a bullet passes through the right motor strip, then disruption of functions on the left side of the body will result (e.g., hemiparesis, paralysis). In general, although the incidence of penetrating head injuries is difficult to ascertain, it has been estimated that only about 2–5% of all head injuries brought to medical attention are penetrating types of injuries (Kampen & Grafman, 1989), and this figure is usually less for children and adolescents. Penetrating head injuries typically occur as a result of military combat, suicide attempts, crime, or miscellaneous accidents that occur during work or leisure activities.

The amount and kind of tissue damage incurred from a penetrating head injury depends on a number of factors including the type of object, the location of entry to the brain, and the velocity of the object. For example, studies investigating the effects of high-velocity missile injuries have evidenced microscopic changes in brain regions far from the actual missile track, whereas the changes incurred from low-velocity objects (e.g., a knife) typically show focal damage. Further, a penetrating object can enter the brain at any particular point and then perform a number of actions. It can come to rest along the original missile trajectory, ricochet within the skull cavity, break apart leaving shrapnel and bone fragments scattered along the path and in adjacent and distal brain regions, or it can exit the skull. The penetrating type of injury also can produce hemorrhaging, although contrecoup injures are not typically observed (Kaufman, Makela, Lee, Haid, & Gildenberg, 1986).

Given the relatively high variability in these types of injuries, systematic neurobehavioral studies are few. A recent study of four adults and two children with penetrating wounds from gun shots showed focal deficits at one year postinjury (e.g., the patients with left hemisphere injuries exhibited greater verbal difficulties) as opposed to any generalized lowering of neurocognitive abilities. In addition, however, all of the patients experienced some form of memory dysfunction (Kampen & Grafman, 1989).

Deficits from a penetrating head injury should directly relate, in large part, to the areas of the brain that were damaged by the object. Other regions of the brain may also be affected, however, given the actions of the object and the response to the object by the brain. Nonetheless, if the entry point of an object can be identified, and perhaps its exit point, then initial functional hypotheses can be established based on what is known about brain–behavior relationships.

9.09.6.2 Neurocognitive Outcomes in Nonpenetrating Head Injuries

Nonpenetrating head injuries can result in more diffuse brain damage. Primary damage occurs to structures at the point of impact. Damage may also occur at the point opposite the impact point (i.e., contrecoup injury) as the brain is pushed against the interior of the skull. Secondary injury results from processes of vasodilation, edema, and increased intracranial pressure that can occur as a result of the primary injury. For example, if the right parietal motor strip is at the point of impact, then an associated left-sided paralysis may also be seen. A right-sided or bilateral paralysis may result as a consequence of the contrecoup injury of the more generalized secondary injury.

In its least severe form, a nonpenetrating head injury, or closed head injury, can produce a concussion. A concussion is a brief alteration/loss of consciousness. This is usually followed by a period of confusion, disorientation, and other neurological signs such as bradycardia, lowering of the blood pressure, and amnesia (Gennarelli, 1987). In children, these symptoms can be delayed for minutes or hours after a TBI, with full recovery typically occurring within 24 hours (Shapiro, 1987). Postconcussion syndrome has also been described and can include headaches, irritability, physical lethargy, emotional lability, memory disruption, and associated academic performance problems (Casey, Ludwig, & McCormick, 1986; Lanser, Jennekens-Schinkel, & Peters, 1988; Levin, Magnusson, Rafto, & Zimmerman, 1989). In addition to the range of symptoms that have been associated with postconcussion syndrome, a variety of other neurocognitive outcomes have been identified in nonpenetrating head injuries.

9.09.6.2.1 Intellectual

A review of numerous studies exploring the possibility of a decline in the level of intellectual performance following a TBI reveals a clear trend in that direction. Klonoff, Law, and Clark (1977) showed that initial estimates of intellectual function in a large group of children and adolescents were significantly below a matched group of normal controls. Levin and Eisenberg (1979) reported a low average level of intellectual functioning for many children who were comatose for at least 24 hours. For the most severely involved children (i.e., a PTA $>$ 7 days), scores on the Wechsler Intelligence Scale for Children-Revised have typically shown lower performance IQs compared to verbal IQs (Chadwick, Rutter, Brown, Shaffer, & Traub, 1981; Levin & Eisenberg, 1979; Winogron, Knights, & Bawden, 1984). This pattern of results is probably due to the dependence of many of the subtests on the verbal scale of old learning, with the exceptions being arithmetic and digit span, whereas subtests from the performance scale are more dependent upon speeded accuracy, problem-solving, and more novel learning (Boll, 1983).

It is unlikely that IQ scores return to preinjury levels for severely brain-injured children, with evidence indicating that only a partial recovery of intellectual abilities is generally possible (Mayes, Pelco, & Campbell, 1989). This finding has been buttressed by studies showing persistent IQ deficits at one (Berger-Gross & Shackelford, 1985; Chadwick, Rutter, Brown et al., 1981), two and a half (Chadwick, Rutter, Shaffer et al., 1991), three (Filley, Cranbert, Alexander, & Hart, 1987), four (Mahoney et al., 1983), and five (Klonoff et al., 1977) years postinjury. Although these findings appear to be related directly to the severity of the TBI, they seem to be robust across the age span.

9.09.6.2.2 Motor/visual–motor

Shapiro (1987) noted that the assessment of motor functions following a TBI has depended upon the procedures used in the neurological examination. Focal signs of motor dysfunction (e.g., hemiparesis, facial weakness) and problems with balance and gait typically have been described. Several studies have attempted to document motor functioning in a more explicit fashion. Levin et al. (1982) described motor slowing during the immediate time period following a mild, moderate, or severe head injury, although the severity of the injury eventually will exert a significant negative influence on the chronicity of these deficits. Gulbrandsen (1984) reported that mildly head-injured children did not differ from normal controls on simple motor speed at six months postinjury. For a severely injured group, however, simple and complex motor speed deficits have been documented (Chaplin, Deitz, & Jaffe, 1993), with these deficits persisting at one-year (Bawden, Knights, & Winogron, 1985) and two-year follow-up points (Klonoff et al., 1977). Visual–motor deficits also have been reported, particularly in copying and visual construction abilities (Bawden et al., 1985).

9.09.6.2.3 Attention

Attentional problems for children and adolescents following a TBI have been described across several studies, with these problems surfacing largely for more severely injured groups (Bruce, Schut, Bruno, Woods, & Sutton, 1978). Hyperactivity and poor attention have been reported by parents at follow-up intervals, with some evidence suggesting that these deficits can persist in a younger age group for up to five years postinjury (Klonoff et al., 1977). Deficits in concentration and speeded performance have been documented up through one year postinjury for individuals at all levels of severity (Bawden et al., 1985; Winogron et al., 1984), although these findings have not been universally supported (Chadwick, Rutter, Shaffer, & Shrout, 1981). Obviously, how attention and its various components are operationally defined and measured will influence these types of findings, and these issues continue to require refinement for study in TBI populations (Fennell & Mickle, 1992).

9.09.6.2.4 Language

In general, the incidence of speech and language deficits following a TBI increases proportionately with the degree of severity (Levin & Eisenberg, 1979). More pervasive deficits, such as mutism and frank aphasias, can be seen in many individuals, particularly those with severe head injuries (Ylvisaker, 1986). These types of global deficits tend to occur in younger children (i.e., less than age five years), although many of these deficits can resolve over the process of recovery. Estimates of aphasic-type disorders secondary to a TBI have ranged from about one-third of pediatric closed head-injury patients (Kaiser & Pfenninger, 1984) to nearly two-thirds of severely brain-injured individuals (Gilchrist & Wilkinson, 1979). From a developmental perspective, Ewing-Cobbs, Fletcher, Levin, and Landry (1985) speculated that the type of speech/language impairment incurred from a TBI would be related to the language abilities that are in primary ascendancy at the time of the injury.

In addition to these global speech/language deficits, specific deficits have also been uncovered. Specific deficits have been reported in object naming (Jordan, Ozanne, & Murdoch, 1990), verbal fluency (Slater & Bassett, 1988), word and sentence repetition (Levin & Eisenberg, 1979), speech (Filley et al., 1987), narrative discourse (Biddle, McCabe, & Bliss, 1996), and written output (Ewing-Cobbs & Fletcher, 1987).

9.09.6.2.5 Memory

Despite the high degree of interaction between memory functions and attentional efficiency, a variety of memory deficits typically can be seen following TBI, particularly in children suffering closed head injuries (Fennell & Mickle, 1992), with deficits continuing to be present through one year postinjury. Similar to other functional domains, the presence of memory deficits secondary to mild and moderate brain injuries is less clear, especially over the long term. Bassett and Slater (1990) showed deficits in immediate and delayed recall of story passages and in visual reproduction in their sample of severely involved adolescents; however, these deficits were not manifested in their mildly injured or age-matched normal controls. Deficits in visual recall (Berger-Gross & Shackelford, 1985) and spatial memory (Winogron et al., 1984) have been reported in severe TBI groups at one year postinjury, but these deficits tended not to be present in milder injuries beyond a 6–12 month period of time. Verbally based memory deficits have also been reported (Chadwick, Rutter, Shaffer, & Shrout, 1981), particularly for verbal retrieval of newly learned materials, although some investigations have uncovered these findings regardless of injury severity (Donders, 1993).

9.09.6.2.6 Academic

Few systematic studies have been conducted examining academic morbidity following a TBI. This is surprising given that a child spends nearly the equivalent of a work week in the school setting. One crucial distinction needs to be made with respect to understanding the academic functioning of children and adolescents with a TBI. This is the distinction between skill-based deficits and performance-based deficits. Skill-based deficits reflect a lack of knowledge regarding a specific academic domain (i.e., the skills have never been present and/or have deteriorated secondary to a TBI), whereas performance-based deficits reflect the poor deployment of skills and abilities that may be present. It is speculated that clinicians should expect to see skill and, perhaps, performance-based deficits at younger ages, but more performance-based deficiencies at older ages.

Academic problems following TBI at all severity levels have covered the gamut, although these findings have typically been reported for children and adolescents suffering moderate to severe injuries. In addition to specific problems in reading (Shaffer, Bijur, Chadwick, & Rutter, 1980), writing (Berger-Gross & Shackleford, 1985), and arithmetic (Levin & Benton, 1986), several investigators have documented the inreased need for special education programs, failure to return to school, and a tendency to return to a lower grade placement (Brink, Garrett, Hale, Woo-Sam, & Nickel, 1970; Fuld & Fisher, 1977; Klonoff et al., 1977). For example, Massagli, Michaud, and Rivara (1996) and Donders (1994) demonstrated that about 50–75% of their samples required some form of special education services postinjury. These rates reflect significant academic morbidity in this population.

The variety of deficits that can be manifested in individuals with TBI truly is impressive, and the impact of these deficits will undoubtedly affect school and vocational performance—even in cases where mild injuries have been sustained. Difficulties learning new or novel materials, problems with higher-order cognition (e.g., generalization, abstraction, organization, planning, strategy generation), slowed information processing speed, and overall independent work efforts all can impact upon an individual's classroom or vocational performance in a negative fashion. These difficulties should be taken into account when evaluating a TBI survivor's return to the formal academic or vocational setting.

9.09.6.3 Psychosocial Outcomes

From all available evidence, it is clear that a TBI can be associated with increased vulnerability for emergence of a psychiatric disorder, and it may contribute to some of the most devastating outcomes following an injury (Voogt & Voogt, 1994). In the 1970s, Klonoff and Paris (1974) found that for their younger age group (i.e., ages 2–8 years), denial of the injury and lack of concern for the injury were the most common reactions. For their older group (i.e., ages 9 through 15), however, they noted a significant deterioration in self-concept. Some examinations have suggested that behavior problems may emerge in the initial time period following even a mild TBI (Casey et al., 1986). In this regard, although studies of children with mild to moderate brain injuries have shown changes in temperament and other transient behavioral symptoms (e.g., increased irritability) immediately following an injury, mild TBI does not appear to be associated with an increased risk for psychiatric disturbance. The same cannot be said with respect to more severe TBI.

Probably one of the best controlled examination of this topic comes from the prospective studies of head-injured children by Rutter and his colleagues (Brown, Chadwick, Shaffer, Rutter, & Traub, 1981; Chadwick, Rutter, Brown et al. 1981; Chadwick, Rutter, Shaffer, & Shrout, 1981; Rutter et al., 1980). Children ranging from 5 to 14 years of age who had experienced closed head injuries of sufficient severity to result in a PTA of seven days or more were compared with a group of children having less severe head injuries (i.e., those with a PTA of less than seven days, but a duration of at least one hour). In addition, these groups were compared with a matched control group of hospital-treated children also suffering severe accidents, but with orthopedic rather than cranial injuries. All children were studied prospectively at 4 months, 1 year, and 2.5 years after their injuries. An important feature of this study was the care taken to determine the children's behavior before their accidents. This was done in an unbiased fashion by interviewing parents immediately after their child's injury, but before the child's postinjury psychiatric condition could have been known.

The children with severe head injuries did not differ from controls in their preinjury behavior, but they showed more than double the rate of psychiatric disorder at four months postinjury, and at each subsequent follow-up point. This was true even when children who had psychiatric disorders before their accidents were eliminated from the study, thereby focusing specifically on the comparative rate of new psychiatric disorders arising over the course of the follow-up period. There was a rather high threshold for an effect, however, because definite cognitive or psychiatric sequelae were found only in head-injured children having a PTA of at least one week. Whereas persistent psychiatric sequelae were quite common once this range of severity was reached, cognitive impairment lasting for more than two years generally required a PTA of at least three weeks. Head-injured children tended to show greater impairment on timed visual–spatial and visual–motor tests than on verbal tests but, apart from this, no pattern of cognitive deficits specific to head injury was identified. Likewise, the types of psychiatric disorder among the head-injured children were very similar to those found in controls. The only exception to this was in the case of grossly disinhibited social behavior, which was present only in children with very severe head injuries and may have been linked directly to frontal lobe damage or dysfunction.

In general, children with severe TBI showed an increased risk for psychiatric disorder regardless of the age, sex, or social class of the child—factors that ordinarily show a striking mediating effect in the general population. The risk was greater among children with histories of preaccident behavior disorders as well as those experiencing various psychosocial adversities within their homes, but the effects were additive rather than interactive. Thus, although psychiatric disorders in childhood have a multifactorial etiology, evidence from this series of studies indicated that severe brain injury can play a major independent role in the development and/or exacerbation of psychosocial difficulties, whereas this does not seem to be the case with milder injuries.

9.09.6.3.1 Impact on the family

Given the psychosocial as well as the neurocognitive changes that can be observed in individuals who have sustained a TBI, it stands to reason that the larger family system will be vulnerable as well. Martin (1990) reported that variables such as family support systems, family communication, extracurricular involvement of the family system (e.g., church), and availability of appropriate services and trained personnel were crucial to assisting families in their coping with the injured individual. Issues pertinent to increased dependency, financial difficulties secondary to costs of rehabilitation and medical needs, and increased family stress, perhaps related to parental or spousal guilt following an injury, all serve to disrupt the family system. Bragg, Klockars, and Berninger (1992) found that families with adolescents who had sustained a mild or moderate head injury experienced less autonomy and more dysfunction. Further, the adolescents themselves also described an increased amount of stress and more maladaptive behaviors.

Families may respond to this stress in a variety of ways. For example, Livingston and McCabe (1990) observed that families may try to cope with the injured individual by becoming overprotective and overindulgent at the expense of other family needs (e.g., siblings, marriage). In fact, Martin (1990) reported that these families seem to progress through the same stages as those who have experienced a significant loss and are grieving: shock, denial, anger, sorrow, and acceptance. Preinjury family status on such variables as communication, problem-solving, role flexibility, and stress also have been deemed predictive of family functioning postinjury (Rivara et al., 1997).

9.09.7 ISSUES RELATED TO OUTCOMES

In addition to rehabilitation efforts, there are a variety of factors that influence rate of recovery of functions and ultimate outcome. Many of these factors have been mentioned already, such as types of injuries, medical complications, and injury severity. In particular, the severity of the injury appears to hold the most weight with respect to prognosis for recovery. Several additional factors also require mention: chronological age at the time of injury, rate of recovery, and premorbid functioning of the child and/or family.

9.09.7.1 Age at Time of Injury

It is clear that a TBI may disrupt new learning, along with a variety of related neurocognitive functions. Given this, it has been speculated that younger children may be at greater risk for learning difficulties in that most of their learning will be new. In general, these difficulties are likely to be related to the physiological maturity of the developing brain and to the functional status of the brain at the time of the injury. Consequently, age-dependent differences might be seen not only in the loss or change in functional status, but in the recovery of functions and subsequent developmental trajectories of these functions (Fletcher, Miner, & Ewing-Cobbs, 1987; Thompson et al., 1994). For example, studies employing children aged five years or older have noted that age at the time of injury was unrelated to either the severity of the neurocognitive sequelae or the rate of recovery (Chadwick, Rutter, Brown et al., 1981; Levin & Eisenberg, 1979). Studies including infants and preschoolers, however, have reported more severe long-term neurocognitive deficiencies in these children (Lange-Cosack, Wider, Schlesner, Grumme, & Kubicki, 1979; Shaffer et al., 1980).

9.09.7.2 Recovery Rates

The rate of recovery appears to be dependent, at least in part, upon the severity of the TBI. In general, individuals incurring more severe injuries tend to show a slower rate of recovery than individuals who sustained mild or, in some instances, moderate TBI. Further, it has been suggested that the more rapidly that a specific function emerges following an injury, the better the prognosis for that function.

For adults and children, it seems that the major portion of recovery occurs during the first year following injury (Jaffe, Polissar, Fay, & Liao, 1995; Klonoff et al., 1977), perhaps due to spontaneous remission and/or the specific interventions employed. Chadwick, Rutter, Shaffer, and Shrout (1981) demonstrated that most of the recovery of intellectual functions tended to occur for more severely injured children. There was relatively little improvement in IQ for the controls and mildly injured individuals at one-year and 30 month follow-up points. In contrast, however, there was significant improvement in IQ by the severely injured group at the one-year follow-up point, with some modest improvements continuing at the 30 month point. Klonoff et al. (1977) have even shown that neuropsychological improvement can be demonstrated over a five-year period.

Fletcher and Levin (1988) and Taylor (1984) offered three different positions for age effects in a TBI and subsequent recovery of function: (i) children show an increased vulnerability to damage, but less recovery of function than adults; (ii) that children show greater behavioral sparing and recovery of function than adults; and (iii) the effects of a TBI will depend upon the age of the child/adolescent, and recovery will be dependent upon the capacity of the developing brain to evolve alternate coping/behavioral strategies. Although the field has not yet arrived at clear answers with respect to these positions, it is clear that the increased precision in how TBIs are evaluated will contribute to increasing our understanding of how age at the time of injury will affect ultimate outcome.

9.09.7.3 Premorbid Functioning

Despite the epidemiological features discussed earlier (e.g., more boys than girls sustaining TBI; more injuries occurring in lower socioeconomic status homes) and, in contrast to the similar gender differences typically found in clinical samples of learning and psychiatric disorders, the gender of a child does not appear to affect the ultimate outcome (Chadwick, Rutter, Brown et al., 1981; Shaffer et al., 1980). Likewise, social class seems to exert little influence on outcome (Chadwick, Rutter, Brown et al., 1981), although clinical intuition persists in suggesting that children from more affluent backgrounds tend to show better improvement following an injury. Despite these findings, preinjury difficulties, particularly in the psychiatric domain, have proved to be predictive of later problems (Brown et al., 1981). These findings suggest the importance of gaining a comprehensive preinjury history when working with individuals who have sustained a TBI.

It remains unclear, however, how these factors interact to influence outcome from a TBI. Although individuals with severe injuries generally tend to have poorer prognoses, those sustaining milder types of head injury can show neurocognitive residuals—even if only for several months after the injury. This is important in that these transient changes in neurobehavioral functioning can affect an individual's daily functioning within the home, school, and vocational settings. Further, given the timing of the injury (e.g., most injuries occur during the spring and summer months in seasonal regions), it is likely that these few months could stretch over two school years and disrupt an individual's academic progress quite significantly.

9.09.8 INTERVENTION

Interventions for individuals with TBI should be tailored to their specific strengths and weaknesses. Although this may seem a bit trite, this adage becomes extremely critical when working with these individuals. In particular, individuals in an acute phase of recovery may be showing frequent changes, and assessment–treatment linkages will need to be dynamic and reviewed on a regular basis. The range of interventions for individuals with TBI is quite wide, with many different kinds of strategies being applicable to this population. These have included pharmacological treatments, psychotherapies, behavioral strategies, family interventions, cognitive rehabilitation, and educational programming. This section will provide a brief overview of two treatments, cognitive rehabilitation and educational programming. The interested reader is referred to other, more comprehensive sources to gain more details pertaining to treatment strategies for individuals with TBI (e.g., Ylvisaker, Szekeres, & Hartwick, 1992).

9.09.8.1 Cognitive Rehabilitation

The term cognitive rehabilitation is extremely broad and, consequently, can include many different kinds of intervention delivered by many different disciplines. In general, the primary goal of cognitive rehabilitation is to promote recovery and to help individuals overcome barriers posed by their cognitive impediments (Ylvisaker et al., 1992). Over two decades ago the work of Diller and Weinberg (1977) illustrated the possible benefits of training individuals to increase their hemifield attention. Using individuals who sustained right hemisphere cerebral vascular accidents, these investigators employed specific training procedures to facilitate attentional adaptation. In a similar fashion, Ben-Yishay, Diller, and Rattok (1978) described the effectiveness of a cognitive retraining program, orientation remedial module, for improving basic deficits in attention. At the same time, these investigators were careful to note that these cognitive retraining strategies facilitated the use of spared abilities, perhaps more so than retraining damaged neurologic systems. The work of Ylvisaker and colleagues has clearly exemplified the importance of developing spared strengths to facilitate the improvement of other neurocognitive systems (e.g., memory, executive functions) following a TBI (Ylvisaker et al., 1992).

The specific strategy or strategies chosen for intervention and the extent of its deployment will depend on the severity of the injury, resulting spared and impaired abilities, the surrounding milieu and general psychosocial supports, and the skills of the interventionist. These variables not withstanding, however, Prigatano, O'Brien, and Klonoff (1993) have provided several pragmatic guidelines when considering any type of cognitive retraining:

(i) Never underestimate the severity of the cognitive deficits.

(ii) Assume initially that an individual does not fully recognize the severity or impact of a cognitive and/or personality deficit.

(iii) Approach any cognitive retraining task with a collaborative learning and problem solving attitude.

(iv) Allow the individual to record and observe their own behaviors, and then encourage the individual to examine the interventionist in the same fashion.

(v) Do not lecture the individual as to the deficits or the need for change.

(vi) Carefully record training activities and measure behavioral outcome on a regular basis.

(vii) Do not limit cognitive retraining activities to one-on-one sessions—*in vivo* situations are likely to prove more ecologically valid.

(viii) Watch the individual for an emotional reaction during the retraining activities.

(ix) Provide ample opportunities for practice in and out of the therapy setting.

(x) Focus on cognitive retraining as well as social–behavioral considerations.

These 10 general axioms should serve any interventionist regardless of the specific cognitive retraining strategies being utilized.

9.09.8.2 Educational Programming

Telzrow (1991) asserted that the process of educational consultation and programming should begin prior to the return of the student to the formal educational setting. In fact, Pieper (1991) stated that this process should begin at the time of the injury so as to promote communications between the family and the school, and to facilitate the planning for the student's reentry. The initial consultation might not only entail performing a comprehensive assessment to determine current neurobehavioral residuals from the injury, but it will also require understanding the nature of the individual's injuries and the current medical status of the student as the school reentry preparations begin. This will increase the need for clear communications between the hospital or rehabilitation facility, the family, and the school (Blosser & DePompei, 1991; Savage, 1991). Furthermore, placement decisions should consider the full range of services available in a school system in order to address the student's pending educational needs. This planning should also extend into the summer months in order to address the issues of maintenance and generalization of information (Ylvisaker, Hartwick, & Stevens, 1991).

Relatedly, it will be important for the practitioner to become familiar with specific teaching strategies that might be helpful with these students. Although many of these strategies may not appear to be very different from what is used in regular and special education settings, their use with individuals with TBI may be crucial with respect to learning, prevocational adaptation, and general school readjustment. For example, Cohen (1991) stated that direct instruction techniques, repetition and practice, cueing, modeling, instructional pacing, decreased use of time limits, and providing immediate feedback regarding the individual's performance will be important components of instruction. Further, Savage (1991) suggested that individual education plans should be written for shorter periods of time due to the rapid changes that can be seen during recovery, and parents should be encouraged to request regular meetings regarding progress and changes in functioning.

9.09.9 SUMMARY

There is much information to be learned and incorporated into the role of the professional working with individuals who have sustained a TBI. This chapter endeavors to provide an overview of this information including a discussion of definitional issues, epidemiological characteristics, and neuropathological mechanisms. Recognizing the differences in types of TBIs and the various kinds of outcomes that can occur are also essential knowledge for professionals working with this population. The recognition of factors that can influence outcomes is also important. Although not covered in this overview, some foundation in the neurological bases of behavior will also be necessary. Treatment strategies represent a wide array of activities, including the use of traditional intervention strategies (i.e., teaching, pharmacology, strategy development); however, their application to individuals with TBI will require a recognition of the spared and impaired abilities as well as the nature of the presenting brain injury. Finally, it is critical to note that one strategy for addressing TBI is via prevention. By advocating for primary prevention venues, such as those addressing bicycle and motorcycle helmets, child restraint systems, child abuse, and drinking-and-driving laws, perhaps some of these injuries could be prevented.

9.09.10 REFERENCES

Adams, J. H., Graham, D. L., & Gennarelli, T. A. (1985). Contemporary neuropathological considerations regarding brain damage in head injury. In D. P. Becker & J. T. Polviishack (Eds.), *Central nervous system trauma status report—1985*. Washington, DC: National Institute for Neurological and Communicative Disorders and Stroke, National Institutes of Health.

Annegers, J. F. (1983). The epidemiology of head trauma in children. In K. Shapiro (Ed.), *Pediatric head trauma* (pp. 1–10). Mount Kisco, NY: Futura.

Annegers, J. F., Grabow, J. D., Kurland, L. T., & Laws, E. R., Jr. (1980). The incidence, causes, and secular trends of head trauma in Olmsted County, Minnesota, 1935–1974. *Neurology, 30,* 910–919.

Auerbach, S. H. (1986). Neuroanatomical correlates of attention and memory disorders in traumatic brain injury: An application of neurobehavioral subtypes. *Journal of Head Trauma Rehabilitation, 1,* 1–12.

Bakay, L., & Glasauer, F. E. (1980). *Head injury*. Boston: Little, Brown.

Bassett, S. S., & Slater, E. J. (1990). Neuropsychological function in adolescents sustaining mild closed head injury. *Journal of Pediatric Psychology, 15,* 225–236.

Bawden, H. N., Knights, R. M., & Winogron, H. W. (1985). Speeded performance following head injury in children. *Journal of Clinical and Experimental Neuropsychology, 7,* 39–54.

Ben-Yishay, Y., Diller, L., & Rattok, J. (1978). A modular approach to optimizing orientation, psychomotor alertness, and purposive behavior in severe head trauma patients. In *Working approaches to cognitive deficits in brain damage* (Rehabilitation Monograph No. 59). New York: New York University Medical Center, Institute of Rehabilitation Medicine.

Berger-Gross, P., & Shackelford, M. (1985). Closed head injury in children: Neuropsychological and scholastic outcomes. *Perceptual and Motor Skills, 61,* 254.

Biddle, K. R., McCabe, A., & Bliss, L. S. (1996). Narrative skills following traumatic brain injury in children and adults. *Journal of Communication Disorders, 29,* 446–469.

Bigler, E. D. (1990). Neuropathology of traumatic brain injury. In E. D. Bigler (Ed.), *Traumatic brain injury: Mechanisms of damage, assessment, intervention and outcome* (pp. 13–49). Austin, TX: Pro-ed.

Bigler, E. D. (1994). Advances in brain imaging with children and adolescents. In M. G. Tramontana & S. R. Hooper (Eds.), *Advances in child neuropsychology* (pp. 48–83). New York: Springer.

Blosser, R., & DePompei, J. (1991). Preparing education professionals for meeting the needs of students with traumatic brain injury. *Journal of Head Trauma Rehabilitation, 6,* 73–82.

Boll, T. J. (1983). Minor head injury in children—Out of sight but not out of mind. *Journal of Clinical Child Psychology, 12,* 74–80.

Bragg, R. M., Klockars, A. J., Berninger, V. W. (1992). Comparison of families with and without adolescents with traumatic brain injury. *Journal of Head Trauma, 6,* 94–108.

Brink, J. D., Garrett, A. L., Hale, W. R., Woo-Sam, J., & Nickel, V. L. (1970). Recovery of motor and intellectual function in children sustaining severe head injuries. *Developmental Medicine and Child Neurology, 12,* 565–571.

Brooks, D. N. (1983). Disorders of memory. In M. Rosenthal, E. R. Griffith, M. R. Bond, & J. D. Miller (Eds.), *Rehabilitation of head injured adults* (pp. 185–196). Philadelphia: F. A. Davis.

Brown, G., Chadwick, O., Shaffer, D., Rutter, M., & Traub, M. (1981). A prospective study of children with head injuries: III. Psychiatric sequelae. *Psychological Medicine, 11,* 63–78.

Bruce, P., Schut, L., Bruno, L. A., Woods, J. H., & Sutton, L. N. (1978). Outcome following severe head injury in children. *Journal of Neurosurgery, 48,* 679–688.

Bucci, M. N., Phillips, T. W., & McGillicuddy, J. E. (1986). Delayed epidural hemorrhage in hypotensive multiple trauma patients. *Neurosurgery, 19,* 65–68.

Caffey, J. (1974). The whiplash shaken infant syndrome: Manual shaking by the extremities with whiplash-induced intracranial and intraocular bleedings, linked with residual permanent brain damage and mental retardation. *Pediatrics, 54,* 396–403.

Casey, R., Ludwig, S., & McCormick, M. (1986). Morbidity following minor head trauma in children. *Pediatrics, 73,* 497–502.

Chadwick, O., Rutter, M., Brown, G., Shaffer, D., & Traub, M. (1981). A prospective study of children with head injuries: II. Cognitive sequelae. *Psychological Medicine, 11,* 49–61.

Chadwick, O., Rutter, M., Shaffer, D., & Shrout, P. E. (1981). A prospective study of children with head injuries: IV. Specific cognitive deficits. *Journal of Clinical Neuropsychology, 3,* 101–120.

Chaplin, D., Deitz, J., & Jaffe, K. M. (1993). Motor performance in children after traumatic brain injury. *Archives of Physical Medicine and Rehabilitation, 74,* 161–164.

Christoffel, K. K. (1990). Violent death and injury in U.S. children and adolescents. *American Journal of Diseases of Children, 144,* 697–706.

Cohen, S. B. (1991). Adapting educational programs for students with head injuries. *Journal of Head Trauma Rehabilitation, 6,* 56–63.

Cooley, E., & Singer, G. (1991). On serving students with head injuries: Are we reinventing a wheel that doesn't roll? *Journal of Head Trauma Rehabilitation, 6,* 47–55.

Cooper, K. D., Tabaddor, K., Hauser, W. A., Schulman, K., Feiner, C., & Factor, P. R. (1983). The epidemiology of head injury in the Bronx. *Neuroepidemiology, 2,* 70–88.

Craft, W., Shaw, D. A., & Cartlidge, N. E. (1972). Head injuries in children. *British Medical Journal, 4,* 200–203.

Cullum, C. M., & Bigler, E. D. (1986). Ventricle size, cortical atrophy and the relationship with neuropsychological status in closed head injury: A quantitative analysis. *Journal of Clinical and Experimental Neuropsychology, 8,* 437–452.

Dacey, R. G., Alves, W. M., Rimel, R. W., Winn, R., & Jane, J. A. (1986). Neurosurgical complications after apparently minor head injury: Assessment of risk in a series of 610 patients. *Journal of Neurosurgery, 65,* 203–210.

Denny-Brown, D., & Russell, W. R. (1941). Experimental cerebral concussion. *Brain, 64,* 93–164.

Diller, L. V., & Weinberg, J. (1977). Hemiinattention in rehabilitation: The evolution of a rational remediation program. *Advances in Neuropsychology, 18,* 63–82.

DiRocco, C., & Velardi, F. (1986). Epidemiology and etiology of craniocerebral trauma in the first two years of life. In A. J. Raimondi, M. Choux, & C. DiRocco (Eds.), *Head injuries in the newborn and infant* (pp. 125–139). New York: Springer.

Donders, J. (1993). Memory functioning after traumatic brain injury in children. *Brain Injury, 7,* 431–437.

Donders, J. (1994). Academic placement after traumatic brain injury. *Journal of School Psychology, 32,* 53–65.

Edna, T. H. (1987). Head injuries admitted to hospital: Epidemiology, risk factors and long term outcome. *Journal of the Oslo City Hospitals, 37,* 101–116.

Ewing-Cobbs, L., & Fletcher, J. M. (1987). Neuropsychological assessment of head injury in children. *Journal of Learning Disabilities, 20,* 526–535.

Ewing-Cobbs, L., Fletcher, J. M., Levin, H. S., & Landry, S. H. (1985). Language disorders after pediatric head injury. In J. K. Darby (Ed.), *Speech and language evaluation in neurology: Childhood disorders* (pp. 97–112). Orlando, FL: Grune & Stratton.

Fennell, E. B., & Mickle, J. P. (1992). Behavioral effects of head trauma in children and adolescents. In M. G. Tramontana, & S. R. Hooper (Eds.), *Advances in child neuropsychology* (Vol. 1, pp. 24–49). New York: Springer.

Filley, C. M., Cranbert, L. D., Alexander, M. P., & Hart, E. J. (1987). Childhood and adolescence. *Archives of Neurology, 44,* 194–198.

Fletcher, J. M., & Levin, H. S. (1988). Neurobehavioral effects of brain injury in children. In D. K. Routh (Ed.), *Handbook of pediatric psychology* (pp. 258–295). New York: Guilford Press.

Fletcher, J. M., Miner, M. E., & Ewing-Cobbs, L. (1987). Age and recovery from head injury in children: Developmental issues. In H. S. Levin, J. Grafman, & H. M. Eisenberg (Eds.), *Neurobehavioral recovery from head injury* (pp. 279–291). New York: Oxford University Press.

Frankowski, R. F. (1985). Head injury mortality in urban populations and its relation to the injured child. In B. F. Brooks (Ed.), *The injured child* (pp. 20–29). Austin, TX: University of Texas Press.

Fuld, P. A., & Fisher, P. (1977). Recovery of intellectual ability after closed head injury. *Developmental Medicine and Child Neurology, 19,* 495–502.

Gennarelli, T. A. (1987). Mechanisms and pathophysiology of cerebral concussion. *Journal of Head Trauma Rehabilitation, 1,* 23–30.

Gilchrist, E., & Wilkinson, M. (1979). Some factors determining prognosis in young people with severe head injuries. *Archives of Neurology, 36,* 355–359.

Goldstein, F. C., & Levin, H. S. (1987). Epidemiology of pediatric closed head injury: Incidence, clinical characteristics, and risk factors. *Journal of Learning Disabilities, 20,* 518–525.

Graham, D. I., Adams, J. H., & Gennarelli, T. A. (1987). Pathology of brain damage in head Injury. In P. R. Cooper (Ed.), *Head injury* (pp. 72–88). Baltimore, MD: Williams & Wilkins.

Gulbrandsen, G. B. (1984). Neuropsychological sequelae of light head injuries in older children six months after trauma. *Journal of Clinical Neuropsychology, 6,* 257–268.

Guyer, B., & Ellers, B. (1990). Childhood injuries in the United States. *American Journal of Diseases of Children, 144,* 649–652.

Hahn, Y. S., Chyung, C., Barthel, M. J., Bailes, J., Flannery, A., & McLone, D. G. (1988). Head injuries in children under 36 months of age. *Child's Nervous System, 4,* 34–40.

Ito, U., Tomita, H., Yamazaki, S., Takada, Y., & Inaba, Y. (1986). Brain swelling and brain edema in acute head injury. *Acta Neurochiro, 79,* 120–124.

Ivan, L. F., Choo, S. H., & Ventureya, E. C. (1983). Head injuries in childhood: A 2-year survey. *Canadian Medical Association Journal, 128,* 281–284.

Jaffe, K. M., Polisar, N. L., Fay, G. C., & Liao, S. (1995). Recovery trends over three years following pediatric traumatic brain injury. *Archives of Physical Medicine and Rehabilitation, 76,* 17–26.

Jagger, J., Levine, J. L., Jane, J., & Rimel, R. W. (1984). Epidemiologic features of head injury in a predominantly rural population. *Journal of Trauma, 24,* 40–44.

Jennett, B. (1975). *Epilepsy after non-missile head injuries* (2nd ed.). London: William Heinemann Medical Books.

Jennett, B. (1986). Head trauma. In A. K. Asbury, G. McKhann, & W. I. McDonald (Eds.), *Diseases of the nervous system* (pp. 1282–1291). Philadelphia: W. B. Saunders.

Jordan, F. M., Ozanne, A. E., & Murdoch, B. E. (1990). Performance of closed head injured children on a naming task. *Brain Injury, 4,* 27–32.

Kaiser, G., & Pfenninger, J. (1984). Effects of neurointensive care upon outcome following severe head injuries in childhood: A preliminary report. *Neuropediatrics, 15,* 68–75.

Kalsbeek, W. D., McLaurin, R. L., Harris, B. S., & Miller, J. D. (1980). The national head and spinal cord injury survey: Major findings. *Journal of Neurosurgery, 53,* 19–31.

Kampen, D. L., & Grafman, J. (1989). Neuropsychological evaluation of penetrating head injury; In M. D. Lezak (Ed.), *Assessment of the behavioral consequences of head trauma* (pp. 49–60). New York: Alan R. Liss.

Kaufman, H. H., Makela, M. E., Lee, K. F., Haid, R. W., Jr., & Gildenberg, P. L. (1986). Gunshot wounds to the head: A perspective. *Neurosurgery, 18,* 689–695.

Klauber, M. R., Barrett-Connor, E., Marshall, L. F., & Bowers, S. A. (1981). The epidemiology of head injury: A prospective study of an entire community—San Diego County, California, 1978. *American Journal of Epidemiology, 113,* 500–509.

Klonoff, H., Low, M. D., & Clark, D. (1977). Head injuries in children: A prospective five-year follow-up. *Journal of Neurology, Neurosurgery, and Psychiatry, 40,* 1211–1219.

Klonoff, H., & Paris, R. (1974). Immediate, short-term and residual effects of acute head injuries in children: Neuropsychological and neurological correlates. In R. M. Reitan & L. A. Davison (Eds.), *Clinical neuropsychology: Current status and applications* (pp. 179–219). New York: Wiley.

Kraus, J. F. (1987). Epidemiology of head injury. In P. R. Cooper (Ed.), *Head injury* (2nd ed., pp. 1–19). Baltimore: Williams & Wilkins.

Kraus, J. F., Black, M. A., Hessol, N., Ley, P., Rokaw W., Sullivan, C., Bowers, S., Knowlton, S., & Marshall, L. (1984). The incidence of acute brain injury and serious impairment in a defined population. *American Journal of Epidemiology, 119,* 186–201.

Kraus, J. F., Fife, D., Cox, P., Ramstein, K., & Conroy, C. (1986). Incidence, severity, and external causes of pediatric brain injury. *American Journal of Diseases of Children, 140,* 687–693.

Kraus, J. F., Rock, A., & Hemyari, P. (1990). Brain injuries among infants, children, adolescents, and young adults. *American Journal of Diseases of Children, 144,* 684–691.

Lange-Cosack, H., Wider, B., Schlesner, H. J., Grumme, T., & Kubicki, S. (1979). Prognosis of brain injuries in young children (one until five years of age). *Neutropadiatrie, 10,* 105–127.

Lanser, J. B., Jennekens-Schinkel, A., & Peters, A. C. (1988). Headache after closed head injury in children. *Headache, 28,* 176–179.

Levin, A. V., Magnusson, M. R., Rafto, S. E., & Zimmerman, R. A. (1989). Shaken baby syndrome diagnosed by magnetic resonance imaging. *Pediatric Emergency Care, 5,* 181–186.

Levin, H. S., & Benton, A. L. (1986). Developmental and acquired dyscalculia in children. In I. Flehmig, & L. Stern (Eds.), *Child development and learning behavior* (pp. 317–322). Stuttgart: Guustav Fisher.

Levin, H. S., Benton, A. L., & Grossman, R. G. (Eds.) (1982). *Neurobehavioral consequences of closed head injury*. New York: Oxford University Press.

Levin, H. S., & Eisenberg, H. M. (1979). Neuropsychological impairment after closed head injury in children and adolescents. *Journal of Pediatric Psychology, 4,* 389–402.

Lipper, M. H. J., Kishore, P. R., Enas, G. G., Domingues-da Silva, A. A., Choi, S. C., & Becker, D. P. (1985). Computed tomography in the prediction of outcome in head injury. *American Journal of Radiology, 144,* 483–486.

Livingston, M. G., & McCabe, R. J. (1990). Psychosocial consequences of head injury in children and adolescents: Implications for rehabilitation. *Pediatrician, 17,* 255–261.

Mahoney, W. J., D'Souza, B. J., Haller, A., Rogers, M. D., Epstein, M. H., & Freeman, J. M. (1983). Long-term outcome of children with severe head trauma and prolonged coma. *Pediatrics, 71,* 756–762.

Martin, D. A. (1990). Family issues in traumatic brain injury. In E. D. Bigler (Ed.), *Traumatic brain injury: Mechanisms of damage, assessment, intervention, and outcome* (pp. 381–394). Austin, TX: Pro-ed.

Massagli, T. L., Michaud, L. J., & Rivara, F. P. (1996). Association between injury indices and outcome after severe traumatic brain injury in children. *Archives of Physical Medicine and Rehabilitation, 77,* 125–132.

Mayes, S. D., Pelco, L. E., & Campbell, C. J. (1989). Relationships among pre- and post injury intelligence, length of coma, and age in individuals with severe closed head injuries. *Brain Injury, 3,* 301–313.

Mealey, J. (1968). *Pediatric head injuries*. Springfield, IL: Charles C. Thomas.

Menkes, J. H. (1990). *Textbook of child neurology* (4th ed.). Philadelphia: Lea & Febiger.

National Head Injury Foundation (1985). *An educator's manual: What educators need to know about students with traumatic brain injury*. Framingham, MA: Author.

Ommaya, A. K., Grubbs, R. L., & Naumann, R. (1971). Coup and contrecoup injury: Observations on the mechanics of visible injuries in the rhesus monkey. *Journal of Neurosurgery, 35,* 503–516.

Pang, D. (1985). Pathophysiologic correlates of neurobehavioral syndromes following closed head injury. In M. Ylvisaker (Ed.), *Head injury rehabilitation: Children and adolescents* (pp. 3–70). San Diego, CA: College-Hill Press.

Pieper, B. (1991). *Traumatic brain injury: What the teacher needs to know*. Washington, DC: National Institute of Disability and Rehabilitation Research.

Prigatano, G. P., O'Brien, K. P., & Klonoff, P. S. (1993). Neuropsychological rehabilitation of young adults who suffer brain injury in childhood: Clinical observations. *Neuropsychological Rehabilitation, 3,* 411–421.

Reitan, R. M., & Wolfson, D. (1985). *Neuroanatomy and neuropathology. A clinical guide for neuropsychologists.* Tucson, AZ: Neuropsychology Press.

Rivara, F. P. (1984). Childhood injuries. III. Epidemiology of non-motor vehicle head trauma. *Developmental Medicine and Child Neurology, 26,* 81–87.

Rivara, J. M., Jaffe, K. M., Polissar, N. L., Fay, G. C., Liao, S., & Martin, K. M. (1997). Predictors of family functioning and change three years after traumatic brain injury in children. *Archives of Physical Medicine and Rehabilitation, 77,* 754–764.

Rivara, F. P., & Mueller, B. A. (1986). The epidemiology and prevention of pediatric head injury. *Journal of Head Trauma Rehabilitation, 1,* 7–15.

Rosen, C. D., & Gerring, J. P. (1986). *Head trauma: Educational reintegration*. San Diego, CA: College-Hill Press.

Rutter, M., Chadwick, O., Shaffer, D., & Brown, G. (1980). A prospective study of children with head injuries: 1. Design and methods. *Psychological Medicine, 10,* 633–645.

Salazar, A. M., Grafman, J., Schlesselman, S., Vance, S. C., Mohr, J. P., Carpenter, M., Pevsner, P., Ludlow, C., & Weingartner, H. (1986). Penetrating war injuries of the basal forebrain: Neurology and cognition. *Neurology, 36,* 459–465.

Savage, R. C. (1991). Identification, classification, and placement issues for students with traumatic brain injuries. *Journal of Head Trauma Rehabilitation, 6,* 1–9.

Shaffer, D., Bijur, P., Chadwick, O. F. D., & Rutter, M. L. (1980). Head injury and later reading disability. *Journal of the American Academy of Child Psychiatry, 19,* 592–610.

Shapiro, K. (1987). Special considerations for the pediatric age group. In P. R. Cooper (Ed.), *Head injury* (pp. 367–389). Baltimore: Williams & Wilkins.

Slater, E. J., & Bassett, S. S. (1988). Adolescents with closed head injuries. *American Journal of Diseases of Children, 142,* 1048–1051.

Soloniuk, D., Pitts, L. H., Lovely, M., & Bartkowski, H.

(1986). Traumatic intracerebral hematomas: Timing of appearance and indications for operative removal. *Journal of Trauma, 26,* 787–794.

Taylor, H. G. (1984). Early brain injury and cognitive development. In C. R. Almli & S. Finger (Eds.), *Early brain injury* (pp. 325–341). New York: Academic Press.

Telzrow, C. (1991). The school psychologist's perspective on testing students with traumatic brain injury. *Journal of Head Trauma Rehabilitation, 6,* 23–34.

Thompson, N. M., Francis, D. J., Stuebing, K. K., Fletcher, J. M., Ewing-Cobbs, L., Miner, M. E., Levin, H. S., & Eisenberg, H. M. (1994). Motor, visual-spatial, and somatosensory skills after closed head injury in children and adolescents: A study of change. *Neuropsychology, 8,* 333–342.

Voogt, R. D., & Voogt, K. R. (1994). Emotions and traumatic brain injury. In C. N. Simkins (Ed.), *Analysis, understanding and presentation of cases involving traumatic brain injury* (pp. 25–35). Washington, DC: National Head Injury Foundation.

Winogron, H. W., Knights, R. M., & Bawden, H. N. (1984). Neuropsychological deficits following head injury in children. *Journal of Clinical Neuropsychology, 6,* 269–286.

Ylvisaker, M. (1986). Language and communication disorders following pediatric head injury. *Journal of Head Trauma Rehabilitation, 1,* 48–56.

Ylvisaker, M., Hartwick, P., & Stevens, M. (1991). School reentry following head injury: Managing the transition from hospital to school. *Journal of Head Trauma Rehabilitation, 6,* 10–22.

Ylvisaker, M., Szekeres, S. F., & Hartwick, P. (1992). Cognitive rehabilitation following traumatic brain injury. In M. G. Tramontana & S. R. Hooper (Eds.), *Advances in child neuropsychology* (Vol. 1, pp. 168–218). New York: Springer.

Yoshino, E., Yamaki, T., Higuchi, T., Horikawa, Y., & Hirakawa, K. (1985). Acute brain edema in fatal head injury: Analysis by dynamic CT scanning. *Journal of Neurosurgery, 63,* 830–839.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.10
Street Gangs

DAVID W. THOMPSON
Clinical Psychology Associates, Burlington, WI, USA

9.10.1 INTRODUCTION 155
9.10.2 NATURE 155
  9.10.2.1 Definition 155
  9.10.2.2 Theories of Youth Gang Involvement 156
    9.10.2.2.1 Subcultural theories 156
    9.10.2.2.2 Opportunity theory 157
    9.10.2.2.3 Social development theory 157
9.10.3 PREVALENCE 160
  9.10.3.1 Methodological Issues 160
  9.10.3.2 Estimates 160
9.10.4 ASSESSMENT 160
  9.10.4.1 Macrolevel 160
  9.10.4.2 Microlevel 161
9.10.5 TREATMENT 161
  9.10.5.1 Street Worker Approaches 161
  9.10.5.2 Group Cohesiveness Approaches 162
  9.10.5.3 Prevention Approaches 162
  9.10.5.4 Gang Suppression Approaches 163
  9.10.5.5 Comprehensive Approaches 164
9.10.6 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 165
9.10.7 SUMMARY 166
9.10.8 REFERENCES 166

9.10.1 INTRODUCTION

Youth street gangs have long been a serious problem in the USA and elsewhere. Gang behavior is costly to the community. Not only are there direct costs to individual victims of assaults, vandalism, and robberies, but the entire community pays significantly for gang-related law enforcement, judicial proceedings, and imprisonment of gang offenders. Indirect costs including loss of life, physical and emotional suffering, and the loss of productivity are inestimable.

Given the severe impact of street gangs on society, effective and timely intervention and prevention efforts are crucial. This chapter seeks to identify the special issues which must be considered when working with youth street gangs. By providing a critical overview of the theoretical, practical, and empirical considerations related to youth street gangs, this chapter will equip others to more effectively intervene with this special population.

9.10.2 NATURE

9.10.2.1 Definition

Simply defining a gang is difficult. Thrasher (1927) described a gang as

> an interstitial group originally formed spontaneously, and then integrated through conflict. It is characterized by the following types of behavior: meeting face to face, milling, movement through space as a unit, conflict, and planning. The result of this collective behavior is the development of tradition, unreflective internal structure, esprit de corps, solidarity, morale, group awareness, and attachment to a local territory. (p. 57)

Thrasher also defined various subtypes of gangs, including the diffuse type (i.e., those without lasting solidarity), the solidified type (the opposite of the diffuse type), the conventional type (a social or athletic club), and the criminal type (those which habitually engage in delinquent and criminal behavior). Fagan (1989) proposed four types of gangs, including party gangs (i.e., those involved in drinking, drug use, and drug sales), social gangs (those which use drugs and commit crimes), delinquent gangs (those which seek to control "turf"), and young organizations (those which are well organized and exist for the purpose of making money).

Differences exist between gangs and delinquent groups. For example, Cohen (1969) compared the behavior of known gang groups with that of nongang groups, and concluded that (i) gang offenders were more homogenous with respect to age, race, sex, and residence patterns than their nongang counterparts; (ii) the victims of gang offenders were more homogenous in their characteristics than the victims of nongang offenders; and (iii) the gang offenders had more extensive prior records and were charged with more serious offenses.

Klein (1971) proposed a useful definition of youth street gangs which will be used in this chapter. He used the term gang to refer to groups of adolescents that "(a) are generally perceived as a distinct aggregation by others in their neighborhood; (b) recognize themselves as a denoteable group (almost invariably with a group name); and (c) have been involved in a sufficient number of delinquent incidents to call forth a consistent negative response from neighborhood residents or law enforcement agencies" (p. 13). The gangs described by Thrasher (1927) were usually formed along racial or ethnic lines, comprised of younger members, and used fists and sticks to settle arguments. Today, youth street gangs transcend racial and ethnic boundaries and routinely use lethal weapons such as knives and guns to settle disputes.

This chapter is limited to a discussion of street gangs. Other types of gangs exist (e.g., prison-based extensions of street gangs or organized crime), but these are fundamentally different from street gangs and will not be addressed.

9.10.2.2 Theories of Youth Gang Involvement

One of the more puzzling and controversial questions facing researchers investigating youth street gangs is the reason for gang existence. Why do street gangs form in the first place, and what factors contribute to their longevity?

9.10.2.2.1 Subcultural theories

Cohen (1955) suggested that gangs arise from a delinquent subculture, namely ... a way of life that has somehow become traditional among certain groups in American society (p. 13). He proposed that the important difference between delinquent and nondelinquent children is their degree of exposure to delinquent conduct and values. Subcultural theory essentially suggests that gang delinquency is the result of learned behavior based on subcultural values which differ from those of society at large or which are only perceived by a narrow segment of the middle class to be different values but are in reality not deviant but normal.

Although subcultural theories of street gang involvement have been examined empirically by a number of researchers, little evidence has been found to support them. For example, Rivera and Short (1967) conducted interviews of 458 lower-class Chicago youth, and found neither a rejection of middle-class concerns with occupational mobility nor the formulation of subcultural delinquency in reaction to actual or anticipated conventional failure.

Spergel (1967) found some support for Cohen's (1955) subcultural theory, although the data were not conclusive. He interviewed 131 boys aged 8–12 years from three Chicago neighborhoods and rated their responses on a 72-item interview schedule. About two-thirds of his sample were classified as deviant, while the remainder were categorized as conformist. He concluded that the delinquent behavior of his sample tended to be relatively independent of behavior patterns developed by older teenagers and adults.

Short, Tennyson, and Howard (1963) attempted to assess behavior patterns characteristic of the gangs studied to determine their significance to the subcultural theory of delinquency. Street workers rated 598 members of 16 delinquent gangs in Chicago on 37 different behavioral categories (e.g., joy riding, truancy, assault, arson, auto theft, homicide). Using factor analysis, the authors derived five factors: (i) conflict (e.g., fighting, use of weapons), (ii) stable corner activities (e.g., team sports, gambling), (iii) stable sex (e.g., statutory rape, petting), (iv) retreatist (e.g., narcotics use, deviant sexual behavior, suicide), and (v) authority-protest (e.g., auto theft, truancy, public nuisance). They reported that criminal behaviors were scattered across all five categories, and that no clearly criminal subculture pattern could be identified.

9.10.2.2.2 Opportunity theory

Opportunity theory is based both on the work of Emile Durkheim and Robert Merton which focused on the sources of social pressure which may lead to deviance, and on the work of Clifford Shaw, Henry McKay, and Edwin Sutherland which examined the ways in which features of social structure regulate the selection and evolution of deviant solutions (Cloward & Ohlin, 1960). Klein (1971) summarized five concepts central to Cloward and Ohlin's theory. Society has generally accepted cultural goals (such as material gain). Not all persons have equal access to these goals or to the accepted means of achieving them, and individuals may therefore turn to illegitimate means of attaining these goals. Communities differ in the availability of both legitimate and illegitimate means to access these goals, and the nature of the community's integration of legitimate and illegitimate means will determine the nature of the subcultural (delinquent) accommodation to goal achievement. Cloward and Ohlin made three predictions based on these concepts:

(i) A criminal subculture, and criminal gangs, will develop where there is cross-age integration of offenders plus close relations between the "carriers of criminal and conventional values."

(ii) A conflict subculture, and conflict gangs, will develop in neighborhoods not integrated as above, where social controls are relatively absent, and there is an absence of available systematic illegitimate means to material goals. The use of violence becomes in this setting an alternate means to an alternate goal (status rather than material goals).

(iii) A retreatist subculture, and retreatist gangs, will develop among boys locked out of the previous two subcultural adaptations because of the lack of means integration and because of "internalized prohibitions" or "socially structured barriers" to the use of violence. This "double failure" leaves only retreat, through drugs and alcohol most specifically (Klein, 1971, pp. 29–30).

Rivera and Short (1967) reported that gang membership seems to be related to lower occupational goals. This would appear to argue against Cloward and Ohlin's theory, which predicts that delinquents would develop higher aspirations to achieve status than nondelinquents, turning to delinquent behavior when they become frustrated by the lack of opportunity to achieve their goals.

However, Spergel (1964) found that delinquents did indeed aspire to higher status than nondelinquent youth when indicators such as anticipated income were used, providing some support for opportunity theory.

Kobrin, Puntil, and Peluso (1967) tested the opportunity theory postulate that styles of delinquency in lower-class communities are a function of the relative availability of illegitimate opportunity. They investigated a network of six street groups, four of which were delinquent. Their data indicated that the delinquent groups could be individually categorized as sophisticated, conspicuous, occupational, and versatile. They concluded that their data supported the implications of opportunity theory that adolescent street gangs' prestige would be highest for groups whose norms were closest to those of adult criminals.

Other aspects of opportunity theory have not received overwhelming empirical support. Elliott (1962) found that delinquents and nondelinquents evaluated the importance for success of school positions, education, and occupational positions similarly. Nondelinquents perceived greater opportunity to become representatives on the student council, but no significant differences existed between the two groups with respect to their perceived opportunity to achieve other school offices, although delinquents perceived lower opportunities to achieve successful educational and occupational positions than nondelinquents. Elliot concluded that the relationship between educational or occupational opportunity and social class was significant, but both middle- and lower-class delinquents perceived lower opportunities than nondelinquents. Although opportunity theory focuses on the discrepancy between success goals and the legitimate attainment of these goals as a source of delinquent behavior, Elliott's data suggest that this perceived discrepancy may be a result of the delinquent behavior itself, that is, the delinquents realize that their delinquent behavior may result in the loss of various opportunities.

9.10.2.2.3 Social development theory

In a comprehensive and systematic review of delinquency theories and research, several US government-funded studies have attempted to provide an empirically based model of delinquency which integrates the most valid parts of various theoretical perspectives (Weis & Hawkins, 1981; Weis, Sederstrom, Worsley, & Zeiss, 1980). The researchers drew upon the two major theoretical perspectives of delinquency: control theory (Briar & Piliavin, 1965; Hirschi, 1969) and cultural deviance theory (Akers, 1977; Akers, Krohn, Lanza-Kaduce, & Radosevich, 1979).

Control theory states that juvenile delinquency is the result of weak, absent, or ineffective social controls (Weis & Hawkins, 1981). Although numerous variations of control theory are found in the literature, Hirschi (1969) offered a version which specifies that a youngster's bond to society will be composed of four elements (i.e., attachment, commitment, involvement, and belief), and that significant units of social control include the family, school, and the law. Delinquency becomes possible when there is inadequate attachment to parents or school, inadequate commitment to educational or occupational success, and inadequate belief in the legitimacy and validity of the law. Thus, a delinquent child is the unsocialized product of his or her environment.

Hirschi's theory does not address the issue of why youth become delinquent, but instead questions why persons refrain from crime. He proposed that the four elements of a youngster's bond to society are important because of the following reasons. Attachment is important because it contains a moral dimension that dissuades persons from engaging in delinquency. He argued that although various subcultures do not exist, variance in a person's belief in society's norms does exist. He proposed that the less one's belief, the more prone one is to engage in delinquent behavior. Commitment to conventional activity dissuades persons from delinquency because they have invested time and energy into a conventional activity which therefore causes the person to weigh the risk of losing their investment against the benefits of engaging in delinquent behavior. Finally, he indicated that involvement reduces delinquency by limiting one's time to contemplate and commit delinquent acts.

Cultural deviance theory states that delinquency results from youths conforming to deviant cultural norms. In essence, cultural deviance theory is identical to the subcultural theories discussed above. Akers (1977) used a social learning model and proposed that the delinquent behavior is learned when the consequences of that behavior are rewarding to the youth engaging in delinquent acts. Thus, to prevent delinquency, delinquent behavior should be extinguished or punished.

Perhaps the most widely cited and investigated cultural deviance theory is Sutherland's theory of differential association (Sutherland & Cressey, 1970). This theory includes seven propositions which address specific causal factors of unlawful behavior. The concept of association with lawbreakers is seen as central to the origin of delinquency, and organized crime and criminal behavior could be overcome if we, as a society, organized ourselves better to fight it. Sutherland's theory of differential association continues to be quite influential.

Control and cultural deviance theories are easily integrated with each other. Although they are both socialization theories, they also have individual inadequacies. In combination, the two theories complement each other by compensating for each others' deficiencies—control theory specifies units and elements of socialization which result in a well-socialized individual, whereas cultural deviance theory focuses on the learning of deviant behavior, suggesting how individuals might be socialized to conventional norms.

Weis and Hawkins (1981) offered a model of delinquency, relevant to prevention and intervention efforts with street gangs, based upon this integration of control and cultural deviance theories. They further suggested that a dynamic multivariate causal model, which is responsive to the direct and interaction effects among variables over time, is appropriate. They postulated that the different causes of delinquency have different effects at different points in a youngster's life. Further, they suggested the existence of four stages of a youngster's life during which different units of socialization exert primary influence. They described these stages as parallel to those of the educational system—preschool, primary school, junior high school, and high school—and as having as their most salient units the family, the school, and peers, respectively. Based on this integrated social development model of delinquency, they described the general process and strategies of delinquency prevention as follows.

Opportunities for involvement in conventional activities and for interaction with appropriate role models are necessary for nondelinquent socialization. In order for these experiences to produce social bonds, participants must have certain basic skills, the application of which makes participation rewarding. It should be emphasized that skills must be possessed by both youthful participants and by others (such as parents and teachers) with whom youths are involved. For example, for involvement in school to be rewarding, students must develop cognitive skills, but teachers must also be skilled in recognizing and reinforcing students' progress. Furthermore, different actors in youths' social environments must be consistent in their expectations for and responses to behavior if conforming behavior is to be continually reinforced and deviant behavior prevented or extinguished.

Based on this model, gang intervention programs must strive to facilitate the formation of bonds between the youth and the appropriate institution (family, school, or peers, depending upon the youth's age). To accomplish this, programs must teach requisite social and educational skills when they are absent, and then facilitate their use and subsequent reinforcement in appropriate manners.

Despite the fact that social development theory is a relatively young theory, research evidence exists both for and against it. Jensen (1972) reanalyzed Hirschi's (1969) data in an attempt to answer two questions. First, he asked if parental control affected delinquency directly (as predicted by control theory) or indirectly by influencing the youth's exposure to delinquent behavior patterns (as asserted by differential association theory). Second, he asked if delinquent peers fostered delinquency directly (control theory) or indirectly through exposure to delinquent behavior patterns (differential association theory).

Jensen used school delinquency rates, perceptions of trouble in the neighborhood, and the number of delinquent friends reported by youth as variables for his analyses. In addition, he utilized self-reported delinquency (SRD) as his outcome variable. He found that when differential association variables were held constant, parental control variables (father's supervision and father's support) still decreased delinquency. He concluded that control theory was supported over differential association theory, and that the number of delinquent friends affected youths' SRD scores independently of the effect of delinquent definitions, and determined that group-process and situational-motive theories of delinquency were empirically superior to differential association theory.

Matsueda (1982) compared Hirschi's control theory against Sutherland's theory of differential association. Using the data upon which Hirschi's (1969) and Jensen's (1972) work was based, he applied mathematical modeling to test specific hypotheses derived from both the theories. Based on both control and differential association theories, Matsueda constructed structural models which specified the causal pathways between background variables (such as age, socioeconomic status, and broken vs. intact families), and youths' perceptions of neighborhood disturbance and other variables (such as parental control, peer relationships, and definitions of delinquent behavior patterns). He concluded that delinquent behavior was reduced in boys by parental supervision which had the effect of reducing their exposure to delinquent definitions of behavior. He further concluded that youths' association with delinquent friends mediated delinquency only by regulating exposure to delinquent definitions of behavior. He concluded that the data supported Sutherland's differential association theory.

Roff and Wirt (1984) followed 2453 three- to six-grade children into adulthood. They identified 127 low peer-status subjects and examined these youth in relation to subsequent delinquency. Using a multivariate design, they assessed the effects of social class, family disturbance, and childhood behavior problems as antecedents of delinquency. They found that childhood aggression was the best predictor of delinquency for males and, that although social class and family disturbance were associated with aggression, these variables exerted no direct effects on delinquency itself.

Suzuki, Nishimura, and Takahashi (1982), however, offered evidence in support of social development theory. Japanese high school students completed a self-report delinquency questionnaire as well as one examining their attitudes towards home and school. The delinquent group of students reported more feelings of maladjustment and discontent about families and schools, and tended to not believe in or adhere to school rules than the nondelinquent students. In addition, this group reported less positive perceptions of families and teachers, studied less, and felt more alienated than the nondelinquent youth. Suzuki et al. (1982) concluded that delinquent students tended to have weaker family and school bonds than nondelinquent students.

In conclusion, little empirical support exists for subcultural theory. Cloward and Ohlin's opportunity theory has received considerable research attention but, although the results of these investigations have been mixed, sufficient support for their theory has been reported to justify further investigation. Social development theory seems promising, but more data are needed before an adequate assessment of that model can be made. As Short (1967) has noted, gang behavior appears to be composed of mixtures of the various patterns as opposed to one pattern exclusively. It is clear that modifications in theory will be necessary in order to explain the current data. Indeed, most of the studies cited above provide partial evidence for practically any of the theories, while questioning portions of others.

The reasons for the current paucity of knowledge in this area stem in large part from methodological difficulties. It is clear from our review of the literature that gang behavior is more complex than Cohen or Cloward and Ohlin postulated. The formulation of more complex (e.g., multivariate) models of gang behavior requires experimental design and data collection techniques significantly more sophisticated than those reflected by the current literature.

There are those, however, who disagree with this formulation. Covey, Menard, and Franzese (1997), for example, assert that we currently possess adequate theoretical knowledge for the formulation of a comprehensive, integrated theory of juvenile gang behavior. They offer a series of 30 theoretical propositions which may account for juvenile gang formation, gang persistence or disintegration, gang structure, gang behavior, and the development of various types of juvenile gangs (Covey et al., 1997, pp. 212–250). Although available space does not allow for a comprehensive description of this theoretical formulation, these authors have attempted in a unique and provocative manner to integrate elements of numerous theories of gang formation and delinquent behavior to explain the origin and behavior of street gangs. Their 30 propositions derived from this approach appear to lend themselves readily to empirical study, perhaps providing a formulation for the multivariate study of street gang phenomena. To the extent that yet-to-be-collected data support these propositions, this comprehensive theoretical approach may suggest specific intervention strategies as well.

9.10.3 PREVALENCE

9.10.3.1 Methodological Issues

A variety of methodological difficulties preclude an accurate estimate of the magnitude of the gang problem. For example, few youth street gangs willingly share membership information with researchers or law enforcement agencies, and many youth who are not gang members may inaccurately identify themselves to peers and adults as gang members. It is clear, however, that the problem of youth street gangs continues to grow tremendously. Although law enforcement estimates of gang crime in specific locales are available, reliance on such data does not allow for the systematic assessment of the scope of gang problems nationwide. Uniform Crime Report (UCR) data are not available on gang crimes, and the National Incident Based Reporting System (NIBRS) does not include information on gang-related crimes. Difficulties in defining gangs and gang-related crime continue to be problematic, particularly since various jurisdictions define and report data differently (Ball & Curry, 1995). The development of methods for the systematic and consistent reporting of gang crimes is a fundamental requirement for accurate research in this area.

9.10.3.2 Estimates

Curry, Ball, and Decker (1996) described efforts to accurately estimate the scope of gang activity in the USA. They provided two sets of estimates of the nation's gang problem, one described as extremely conservative and the other as reasonably larger. Based on police reports, they conservatively estimated gang problems in approximately 57% of all US cities with populations over 25 000, with gang problems in 87.8% of cities with populations over 150 000. Using statistical procedures to estimate data from jurisdictions which did not formally produce annual statistics, they offered a larger estimate of the nation's gang problem based on 1994 National Institute of Justice extended survey data. Using this approach, they estimated the number of gang members as 555 181 and the number of gangs as 16 643. Their estimate of the national total gang crimes for 1993 was 580 331 (Curry et al., 1996).

If the prevalence of gangs, gang membership, and gang crimes is disturbing, the proliferation of gangs and gang-related crime in the USA is alarming. Klein (1995) described the spread of street gangs as "a seriously accelerating problem," citing a 74% growth in gang-involved cities between 1961 and 1970, 83% by 1980, and 345% by 1992.

9.10.4 ASSESSMENT

The assessment of street gangs is a complex endeavor fraught with problems. Given the problems associated with even the definition of a street gang, how does one approach the empirical assessment of gangs? Unlike many other groups or phenomena, street gangs often have no desire to be assessed, and have the potential for becoming uncooperative with or even dangerous toward those who attempt to quantify gangs or their activities. These issues pose significant challenges for the researcher or clinician seeking to assess the effectiveness of a given street gang intervention.

9.10.4.1 Macrolevel

One approach to the assessment of street gangs involves approaching them on a macrolevel. For example, one could collect data on the number of cities reporting street gang activity (Klein, 1995; Spergel & Curry, 1993). Data regarding emerging trends in gang membership and behavior also contribute to the large-scale understanding of street gangs (Chesney-Lind, Shelden, & Joe, 1996; Chin, 1996).

Although these data may prove useful in describing and assessing changes in the national scope of the street gang problem, a variety of political factors and the general lack of standardized reporting of gang crimes continue to decrease the reliability of these data. Curry et al. (1996) document many of these factors, and call for a more organized and comprehensive federal effort to collect national data on the US gang crime problem.

9.10.4.2 Microlevel

An alternative approach to the assessment of street gangs involves approaching them on a smaller, microlevel. Studying the individual members of a specific street gang, for example, can provide useful data relating to trends in gang membership (Thompson & Jason, 1988), the specific activities of individual members, and contribute to the assessment of gang cohesiveness.

Difficulties are also present with this microlevel approach to street gang assessment. How many gangs maintain formal membership rosters which they will openly share with interested researchers? If one adopts a more behavioral approach to assessing gang membership (e.g., classifying members based on their manner of dress, self-report, and peer affiliations), how does one classify "wannabe" or copy-cat individuals? At what point do we consider them real street gang members?

Although there are no definitive answers to these complicated assessment issues, several solutions seem reasonable. Macrolevel assessment of street gangs may be useful for tracking trends and changes in street gang behavior as a whole (e.g., increases in violent behavior). Such data may also prove beneficial when comparing gang problems in the same cities over time or when conducting larger-scale comparisons of various intervention approaches. These data must be tempered, however, with the realization that they have significant limitations. Caution must be exercised when utilizing data which may, on the surface, appear to be uniform but upon closer examination may reflect vastly diverse origins or regional idiosyncrasies. In a similar vein, microlevel assessment of street gangs may prove beneficial when evaluating the effects of specific intervention or gang suppression efforts.

The development of uniform definitions, data collection, and reporting for gang crimes and related behaviors is fundamental to the systematic study of street gangs and effective gang intervention programs. Until coordinated efforts result in improvements in these areas, the development and implementation of effective street gang intervention programs will remain severely handicapped.

9.10.5 TREATMENT

Treatment approaches to street gangs can generally be grouped into one of several categories. Traditional street worker model programs have attempted to intervene directly with street gangs on their turf through the use of gang intervention workers. Other programs have targeted the variables which allow street gangs to continue, such as group cohesiveness. Some treatment approaches have adopted a prevention approach, seeking to intervene with at-risk youth before they become gang members, whereas other programs are aimed at gang suppression. Finally, some intervention programs have attempted to adopt a more comprehensive approach to gang intervention by selectively drawing upon various promising components of all intervention categories.

9.10.5.1 Street Worker Approaches

Gang intervention programs utilizing street workers (also referred to as detached workers) are based on the proposition that because gang members do not respond well to agency-based (in-house) programs, the programs must be taken to the youth in order to be effective (Klein, 1971). Spergel (1996) defined street work as

> a systematic effort by an agency worker to help a group of young people, described as delinquent or potentially delinquent, to achieve a conventional adaptation through use of social work or social treatment techniques within distinct neighborhood contexts. Street work further requires work with, or manipulation of, those persons or agency representatives who interact critically with the members of delinquent groups (p. 44).

Unfortunately, few street worker programs have been carefully evaluated. For example, the New York City Youth Board established a project using street workers known as the Council of Social and Athletic Clubs. Gannon (1967) reported the results of questionnaires and interviews with agency street workers, but the results did not evaluate the overall program effectiveness, and instead were largely confined to general characteristics of the youth gangs serviced by the project.

When detached worker programs have been systematically evaluated, the results have been less than encouraging. Miller (1962) reported the results of the Roxbury project, a three-year detached worker program targeting about 400 members of 21 street gangs. Evaluation using a variety of data found no significant effects, and matched control-group gangs evidenced almost identical trends in court appearance rates.

The effectiveness of a joint detached worker project of the Chicago Boys Clubs and the Institute for Social Research of the University of Michigan was evaluated in a series of studies (Caplan, Deshaies, Suttles, & Mattick, 1967). Although the program successfully established relationships with various gangs and street youth, overall delinquency was not reduced (Klein, 1971). Indeed, Caplan (1968) reported that the project youth who received the most intensive services actually reported an increase in delinquency rates.

Klein (1969) described the Los Angeles Group Guidance Project, a program targeting lower- and lower-middle-class urban youth. Using a pre–post design, with gangs as their unit of observation, the intervention "had the overall effect of increasing the number of recorded offenses committed by the gangs included in its program" (Klein, 1969, p. 157). The evaluation suggested that the intervention actually tended to strengthen gangs, particularly those with initially lower levels of group cohesiveness.

9.10.5.2 Group Cohesiveness Approaches

The Ladino Hills Project attempted to reduce delinquency and direct gang members into more positive channels by breaking the bonds of gang cohesiveness (Klein, 1971). Rather than viewing group cohesiveness in gangs as a positive dynamic, the youth were conceptualized as being driven to gang membership by environmental influences. Intervention included arranging individual relationships with adults for targeted youth, de-emphasizing external threats, and varying the assignment of workers to the gang so that the continuity of a given worker's presence did not foster cohesiveness. Klein reported a reduction in gang cohesiveness ranging from 11% to 40%, depending upon the indicator used. Individual gang member offense rates did not change, but the entry of new members to the gang was completely halted, resulting in an overall 35% decrease in the number of offenses committed. Klein concluded that higher levels of group programming for gangs leads to increased gang cohesiveness, and a focus on younger gang members increases cohesiveness through recruitment. He also concluded that the transfer of street workers between gangs eliminates a major contributing factor to gang cohesiveness.

9.10.5.3 Prevention Approaches

Few approaches to street gangs have sought to prevent youth identified as at-risk from joining gangs. This may stem from the extensive past emphasis on detached worker programs, political climates which affect the availability of funding, or the methodological difficulties inherent in measuring something which has not occurred (Jason, Thompson, & Rose, 1986). Given the proliferation of street gangs, prevention must be considered. The costs of an effective prevention program are likely to be substantially less than the direct and indirect costs of gang membership, and the work of Klein (1971) strongly suggests that preventing new members from joining gangs is a highly effective strategy for the elimination of a gang altogether.

A sound theoretical basis for prevention efforts exists. Weis (1977) suggested several specific implications for delinquency prevention based on control theory and directed toward the family, the schools, and the law. First, he suggested that delinquency prevention efforts work toward enhancing the family's ability to develop self-control among its children. Second, prevention efforts should assist youth in developing and maintaining an attachment to school and education. Third, prevention efforts should be directed at desocialization of the juvenile court system so as to strengthen the youths' belief in the law. Finally, he suggested that all delinquency prevention efforts directed at these institutions should also attempt to enhance the self-concept of youth.

Weis (1977) conceptualized a community organizational approach to delinquency prevention under cultural deviance theory, and enumerated several points to be considered when implementing such interventions. First, a community must organize itself against delinquent behavior. Second, community control of delinquency prevention and other youth services should be encouraged. Third, both youth and adult participation in delinquency prevention should be effected. Finally, delinquent groups should be coopted or disbanded, and ties to conventional groups should be encouraged and developed.

Broader Urban Involvement and Leadership Development (BUILD) is a Chicago nonprofit agency which focuses entirely on the problem of youth street gangs. A school-based gang membership prevention program operated by BUILD provided a series of 12 classroom sessions to eighth-grade public school students in areas where active gang recruiting in the schools was reported. Following conclusion of the classroom sessions, participants who had been identified as being at risk for gang recruitment were invited to participate in an after-school athletic program conducted in their neighborhood by BUILD staff. This program consisted of organized sports clinics and competition and encouraged intragroup cooperation. The after-school program also facilitated the travel of these youth out of their neighborhood to participate in events and activities with similar groups from other locations. Finally, targeted youth were provided with job skills training workshops, social/recreational opportunities, and educational assistance programs sponsored by BUILD. The effectiveness of the prevention program was evaluated using a quasi-experimental design, and the results were reported by Thompson and Jason (1988). The dependent variable of gang membership was determined pre- and postintervention by comparing targeted youth names with gang rosters provided by gang members involved with other BUILD programs.

The BUILD program evaluation showed that one experimental group youth and four control group youth joined gangs during the year. None of the youth who participated in the after-school program had subsequently joined street gangs during the year. Although these results were not statistically significant, this strong trend suggested that the prevention program may have been beneficial despite methodological difficulties found in the evaluation of a multifaceted program's effects on a transient and often poorly identified population. The BUILD program data also suggested that viable alternatives to detached worker interventions may exist, and that they should certainly receive additional research and evaluation attention.

Cahill (1996) has explored empirical factors related to gang membership and serious delinquent behavior. She reported a relationship between psychopathy ratings and active gang membership, again suggesting that factors relevant to the identification of youth at risk for street gang membership can be isolated. This in turn suggests that further efforts aimed at preventing youth from joining street gangs may be developed with some degree of success.

9.10.5.4 Gang Suppression Approaches

Usually operated by police departments or law enforcement agencies, gang suppression efforts are designed to apply pressure to gang members. Efforts in this arena may include street sweeps where the police focus on a given area and attempt to arrest as many gang members as possible during the sweep. Other gang suppression efforts include harassment of gang members by police, increased police presence, and generally any get-tough approach tolerated by the community or the legal system.

Klein (1995) described the Los Angeles Police Department's Operation Hammer as a "massive, antigang police sweep." A special force of 1000 police officers swept through the targeted area arresting everyone available on the basis of outstanding warrants, observed criminal activities, and gang-related behavior. Klein indicated that of a total of 1453 arrests during the two day sweep, 1350 were released without charges, almost half were not gang members, and there were only 60 felony arrests, with 32 of these leading to charges (Klein, 1995).

Project HICOP (High Intensity Community Oriented Policing) developed out of the Aurora, Colorado volunteer Aurora Gang Task Force. Atkinson (1996) described the project as including police sweeps through targeted areas, mass arrests in an attempt to control gang-related drug activity, and assigning police officers to assist citizens in keeping crime out of their neighborhoods. Project HICOP was only one facet of an often-discussed and praised community antigang effort in Aurora. Nonetheless, Atkinson (1996, p. 262) reported that, despite six years of intervention efforts, "gang-related statistics from the Aurora police department portray mixed results."

Gang suppression efforts are generally based on deterrence theory. Klein (1995) capsulized this approach as follows:

(i) low priority is given to prevention and treatment efforts;

(ii) high priority is given to visible crime which frightens the public;

(iii) a policy of deterrence of gang crime through the use of harassments, surveillance, and penalties is adopted;

(iv) it is assumed that gang members will respond to these efforts by engaging in less crime.

Special prosecution efforts have been effective in addressing some areas of organized crime (e.g., federal RICO statutes and special prosecution teams). Similar efforts have been directed towards street gangs in an attempt to deter gang-related crime. For example, Klein (1995) described the efforts of the Los Angeles District Attorney to hold all gang members responsible for the actions of other members. As part of a graffiti-removal effort, the district attorney obtained a court order declaring gangs to be a form of corporate structure, allowing the city to force gang members to participate in clean-up efforts even if they had not placed the graffiti themselves. The Institute for Law and Justice (1993) reported other statutory mechanisms for prosecution-based gang-suppression efforts, including provisions for waiver from juvenile court to adult court, various penalty enhancers for gang-related crime, efforts to hold parents financially responsible for the behavior of their children, and enhanced provisions for the confiscation of property or money related to gang crimes.

Are gang suppression efforts effective? Klein (1995) has argued that many facets of these efforts may actually increase gang cohesiveness, thus inadvertently strengthening the very gangs the programs seek to destroy. By increasing opposing forces, gangs become stronger. When efforts at gang suppression are shown to be ineffective, communities do not abandon them but instead intensify efforts, thus intensifying gang cohesiveness.

Some antigang units can be effective, however. Klein (1995) contrasted the efforts of two Los Angeles area gang suppression programs, the Los Angeles Police Department's Community Resources Against Street Hoodlums (CRASH) units and the Los Angeles Sheriff Department's Operation Safe Streets (OSS). The CRASH units consist of uniformed officers and standard patrol cars which cruise gang areas and harass gang members. They are not assigned to the project over the long term, and are rotated out of the program after several years. The OSS officers wear plain clothes (not uniforms) and drive unmarked cars, and often remain assigned to a particular community for many years. Klein reported that the less suppressive approach of the OSS allows them to be more effective in their communities. He cited news reports (Los Angeles Times, 1986) claiming that the OSS program is more effective with fewer resources than the CRASH program.

Other less publicized gang suppression efforts are described by Klein (1995). These include techniques such as the cul-de-sac operation, where police blockade an area of known drug distribution and control access to it for a long period of time. Although no formal evaluations of this technique have been reported, it is likely that the drug dealing operations simply change locations, and although some relief to the immediate neighborhood may be obtained, no true impact on crime results from this technique. Shaming, or the use of public apologies or public appearances, has been attempted in some communities, as have local crackdowns on activities by specific individuals. No outcome data have been reported for these techniques. Education efforts directed toward citizens, groups, and law enforcement coalitions have been provided by metropolitan gang crime units and even the federal authorities, such as the FBI, the Drug Enforcement Agency (DEA), and the Bureau of Alcohol, Tobacco, and Firearms (ATF) have become involved in suppression efforts. Unfortunately, these agencies do not appear to effectively generalize their expertise in the area of organized crime to the dissimilar area of street gangs.

Finally, school-based gang suppression efforts seem to be increasing. Zero-tolerance discipline policies, dress codes, and prohibitions against gang haircuts and other accouterments of gang behavior are tools implemented in many school districts. Again, few data are available regarding the impact of these efforts on gang crime.

In general, gang suppression efforts do not appear promising. Not only have gang-related offenses and gang problems continued to escalate even in the presence of gang suppression projects, but the direct cost to communities for the funding of these efforts is high. Although sweeps, enhanced policing and prosecution of gang crimes, and community- and school-based suppression efforts may have temporary and transient effects on the activities or visibility of street gangs, they also clearly create an adversarial situation wherein the street gangs must combat an external threat to their survival. This may well have the unintended effect of strengthening street gangs through an increase in their cohesiveness. It is vital that communities, school boards, and law enforcement agencies alike implement carefully planned and effective gang intervention programs rather than simply responding to the perceived need to do something which does not work, and to do more of the same when efforts are seen to fail. Yablonsky (1997) noted that vigorous law enforcement in the absence of social programs is ineffective. Substantial changes in timing and certainty of consequences, as well as other aspects of the interaction between law enforcement agencies and individual gang members, are needed if gang suppression efforts are to be effective (Covey et al., 1997).

9.10.5.5 Comprehensive Approaches

Goldstein and Glick (1994) described an integrated gang intervention approach which consists of more than deterrence and incarceration. This model "seeks to prescriptively apply major features of detached worker, opportunities provision, and social control programming" (p. 43). Comprehensive gang intervention programming consists of a three-dimensional approach to street gangs. Individual-oriented interventions include the use of interpersonal skills training, anger control techniques, moral reasoning training, contingency management techniques, and cognitive-behavioral interventions. System-oriented interventions include family- and school-based components, as well as changes in the areas of vocational and employment, recreation, and the community. Criminal justice interventions include suppression efforts described above and encompass efforts of police, prosecutors, and the corrections system. This combined programming seeks to balance the punishing effects of deterrence efforts to decrease behavior with the establishment of new behavioral repertoires which seek to reward the use of prosocial adaptive behavior.

Comprehensive programming uses Aggression Replacement Training (ART) (Goldstein, Glick, Reiner, Zimmerman, & Coultry, 1986) as a pivotal component for gang intervention and instruction in nonviolent and prosocial behavior. ART incorporates the use of modeling and guided practice in the acquisition of new skills, systematic training in the inhibition of aggressive behavior, and enhancement of higher cognitive processes in the areas of moral reasoning and prosocial problem solving.

ART has been demonstrated to be effective in improving prosocial skills and behavior, reducing impulsive and acting out behavior, and in enhancing moral reasoning in incarcerated juvenile delinquents. ART programming was applied in a multigang delinquency intervention program in New York, where outcome measures included quantitative appraisal of changes in participants' interpersonal skills, anger reduction, community functioning, and rearrest. Goldstein and Glick (1994) reported improved interpersonal skills and community functioning ratings for youth who participated in ART programming and, although anger control measures did not show statistically significant differences between ART participants and control group members, ART participants demonstrated greater gains in anger control when compared to control group youth. Perhaps most importantly, ART group participants were rearrested at a significantly lower rate (13%) than the control group youth (52%) during the eight-month tracking period, suggesting that the behavioral gains made by ART group youth were effectively transferred to the community.

9.10.6 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

There is no question that the longstanding problems with street gangs continue to exist. The proliferation of gangs, the increase in violence associated with them, and the significant direct and indirect costs to individuals and society emphasize the pressing need to formulate effective and well-coordinated interventions. Decades of research, high levels of funding, and the efforts of a virtual army of law enforcement personnel have had little demonstrated effect on gang behavior, and may have even exacerbated the problem.

A plethora of data in many forms tell us what not to do, yet little reliable information is available to tell us what steps to take to effectively impact gang crimes in a lasting manner. This chapter has examined the conceptual and methodological difficulties inherent in addressing street gangs, and has described selected intervention programs utilizing popular approaches.

It is vital that individuals and groups seeking to have an impact on street gangs develop a full awareness of the pitfalls and potential boomerang effect of plunging into gang intervention activities which increase gang cohesiveness. Although it may be personally and politically difficult to tolerate a more constrained and less publicized approach to street gangs, it must be recognized that ill-conceived programs may strengthen street gangs and further exacerbate the entire range of problems associated with them.

At a federal policy level, a system for uniform data collection related to street gangs must be established. Conceptual and definitional difficulties in the field of gang research are so widespread that valid efforts to assess even the scope of the nation's gang problem must rely on statistical estimates of nonexistent data. Policies which create consistent definitions of street gangs and street gang membership may not immediately impact intervention attempts, but such definitions will clearly contribute to the accurate and reliable assessment of intervention techniques. Reliably defining gang crimes or gang-related episodes will allow improvement in our abilities to assess the magnitude and impact of street gangs and intervention attempts alike.

With recognition of the extreme economic impact of street gangs and related crime on America, federal policymakers should also consider the establishment of a national street gang data bank to promote effective research with gangs. By serving as a conduit for information and funding, such an effort could improve the effectiveness of gang prevention and intervention research at all levels. Although the framework for such efforts probably already exists within various federal departments, a national commitment to a concerted effort to effectively impact street gangs seems lacking. Without such a commitment, efforts are likely to remain reactive, ineffective, and poorly coordinated.

At the state or regional levels, similar policies coordinating the collection and sharing of accurate data must be developed. Just as importantly, resources and policies which effectively address the factors which contribute to the longevity of street gangs, such as increased group cohesiveness or the lack of viable alternative socialization resources, must be further developed.

At the individual project intervention level, further exploration is needed of the factors which contribute to or are associated with gang membership and active gang behavior. Rather than devoting extensive resources to large-scale intervention and gang suppression efforts in a poorly focused and unscientific approach, projects must first empirically identify relevant predictor variables for gang membership, determine what techniques are effective in preventing membership, and ascertain effective methods for efficiently implementing such programming. No projects should be undertaken in the absence of rigorous evaluation designs.

Although it is unsettling to review the ineffective results of decades of so many seemingly promising projects, it is clear that, as a whole, there are a number of principles that may serve to guide the effective and efficient use of gang intervention resources. First, without improvements in conceptualization and data collection, researchers are doomed to repeat the mistakes of the past. In the absence of the ability to reliably and validly measure the gang problem, intervention efforts cannot be assessed and directed. Second, ill-conceived interventions are likely to make the problem worse. This concern must be acknowledged and effective steps taken to address it. Finally, it is known that street gangs and their associated problems can be eliminated by preventing the influx of new gang members. Research that shapes and refines the technology for accomplishing this goal, however, is sorely needed.

Decades of gang intervention programs and research have not been wasted, but the lesson has been costly. Clearly the time has come to reexamine and refocus gang intervention efforts in effective directions.

9.10.7 SUMMARY

Street gangs, historically a serious problem in the USA, have proliferated markedly in past decades. Years of traditional intervention programs notwithstanding, estimates of gang membership, gang numbers, and gang-related crimes have increased exponentially. Basic research issues, including reliable and valid definitions of gang variables, have not been consistently addressed, and uniform reporting of gang-related crimes is not being done. Despite research to the contrary, intervention programs which do not work and which may actually amplify and strengthen street gangs continue to be implemented.

9.10.8 REFERENCES

Akers, R. L., Krohn, M. D., Lanza-Kaduce, L., & Radosevich, M. (1979). Social learning and deviant behavior: A specific test of a general theory. *American Sociological Review, 44*, 636–655.

Akers, R. L. (1977). *Deviant behavior: A social learning approach*. Belmont, CA: Wadsworth.

Atkinson, W. K. I. (1996). The community response to gangs. Section A: Organizing the community response in Aurora, Colorado. In C. R. Huff (Ed.), *Gangs in America* (2nd ed., pp. 257–262). Thousand Oaks, CA: Sage.

Ball, R. A., & Curry, G. D. (1995). The logic of definition in criminology: Purposes and methods for defining "gangs." *Criminology, 33*, 225–245.

Briar, S., & Piliavin, I. (1965). Delinquency, situational inducements and commitment to conformity. *Social Problems, 13*, 25–45.

Cahill, M. A. (1996). A comparison of gang member and non-gang member male juvenile delinquents. *Dissertation Abstracts International, 57*, 4023B.

Caplan, N. S. (1968). Treatment intervention and reciprocal interaction effects. *Journal of Social Issues, 24*, 63–88.

Caplan, N. S., Deshaies, D. J., Suttles, G. D., & Mattick, H. W. (1967). The nature, variety, and patterning of street club work in an urban setting. In M. W. Klein & B. G. Myerhoff (Eds.), *Juvenile gangs in context: Theory, research, and action* (pp. 194–202). Englewood Cliffs, NJ: Prentice-Hall.

Chesney-Lind, M., Shelden, R. G., & Joe, K. A. (1996). Girls, delinquency, and gang membership. In C. R. Huff (Ed.), *Gangs in America* (2nd ed., pp. 185–204). Thousand Oaks, CA: Sage.

Chin, K. I. (1996). Gang violence in Chinatown. In C. R. Huff (Ed.), *Gangs in America* (2nd ed., pp. 157–184). Thousand Oaks, CA: Sage.

Cloward, R. A., & Ohlin, L. E. (1960). *Delinquency and opportunity*. Glencoe, IL: Free Press.

Cohen, A. K. (1955). *Delinquent boys*. Glencoe, IL: Free Press.

Cohen, B. (1969). The delinquency of gangs and spontaneous groups. In T. Sellin & M. E. Wolfgang (Eds.), *Delinquency: Selected studies* (pp. 61–111). New York: Wiley.

Covey, H. C., Menard, S., & Franzese, R. J. (1997). *Juvenile gangs*. Springfield, IL: Charles C. Thomas.

Curry, G. D., Ball, R. A., & Decker, S. H. (1996). Estimating the national scope of gang crime from law enforcement data. In C. R. Huff (Ed.), *Gangs in America* (2nd ed., pp. 21–36). Thousand Oaks, CA: Sage.

Elliott, D. S. (1962). Delinquency and perceived opportunity. *Sociological Inquiry, 32*, 216–227.

Fagan, J. (1989). The social organization of drug use and drug dealing among urban gangs. *Criminology, 27*, 633–669.

Gannon, T. M. (1967). Dimensions of current gang delinquency. *Journal of Research in Crime and Delinquency, 4*, 119–131.

Goldstein, A. P., & Glick, B. (1994). *The prosocial gang: Implementing aggression replacement training*. Thousand Oaks, CA: Sage.

Goldstein, A. P., Glick, B., Reiner, S., Zimmerman, D., & Coultry, T. M. (1986). *Aggression replacement training: A comprehensive program for aggressive youth*. Champaign, IL: Research Press.

Hirschi, T. (1969). *Causes of delinquency*. Berkeley: University of California Press.

Jason, L. A., Thompson, D. W., & Rose, F. T. (1986). Methodological issues in prevention. In B. A. Edelstein & L. Michaelson (Eds.), *Handbook of prevention* (pp. 1–19). New York: Plenum.

Jensen, G. F. (1972). Parents, peers, and delinquent action: A test of the differential association perspective. *Amer-*

ican Journal of Sociology, 78, 562–575.

Institute for Law and Justice (1993). *Gang prosecution: Prosecutor survey*. Alexandria, VA: Author.

Klein, M. W. (1969). Gang cohesiveness, delinquency and a street-work program. *Journal of Research in Crime and Delinquency, 6,* 135–166.

Klein, M. W. (1971). *Street gangs and street workers*. Englewood Cliffs, NJ: Prentice-Hall.

Klein, M. W. (1995). *The American street gang: Its nature, prevalence, and control*. New York: Oxford University Press.

Kobrin, S., Puntil, J., & Peluso, E. (1967). Criteria of status among street groups. *Journal of Research in Crime and Delinquency, 4,* 98–118.

Los Angeles Times (1986, January 19). Policing gangs: Case of contrasting styles. *Los Angeles Times*.

Matsueda, R. L. (1982). Testing control theory and differential association: A causal modeling approach. *American Sociological Review, 47,* 489–504.

Miller, W. B. (1962). The impact of a "Total Community" delinquency control project. *Social Problems, 10,* 168–190.

Rivera, R. J., & Short, J. F. (1967). Occupational goals: A comparative analysis. In M. W. Klein & B. G. Myerhoff (Eds.), *Juvenile gangs in context: Theory, research, and action* (pp. 70–90). Englewood Cliffs, NJ: Prentice-Hall.

Roff, J. D., & Wirt, R. D. (1984). Childhood aggression and social adjustment as antecedents of delinquency. *Journal of Abnormal Child Psychology, 12,* 111–126.

Short, J. F., (1967). Introduction. *Journal of Research in Crime and Delinquency, 4,* 1–4.

Short, J. F., Tennyson, R. A., & Howard, K. I. (1963). Behavior dimensions of gang delinquency. *Sociological Review, 28,* 411–428.

Spergel, I. (1964). *Racketville, Slumtown, Haulberg: An exploratory study of delinquent subcultures*. Chicago: University of Chicago Press.

Spergel, I. (1966). *Street gang work: Theory and practice*. Reading, MA: Addison-Wesley.

Spergel, I. (1967). Deviant patterns and opportunities of preadolescent negro boys in three Chicago neighborhoods. In M. W. Klein & B. G. Myerhoff (Eds.), *Juvenile gangs in context: Theory, research, and action* (pp. 38–54). Englewood Cliffs, NJ: Prentice-Hall.

Spergel, I. A., & Curry, G. D. (1993). The national youth gang survey: A research and development process. In A. Goldstein & C. R. Huff (Eds.), *Gang intervention handbook* (pp. 359–400). Champaign-Urbana, IL: Research Press.

Sutherland, E. H., & Cressey, D. R. (1970). *Principles of criminology*. New York: J. B. Lippincott.

Suzuki, S., Nishimura, H., & Takahashi, Y. (1982). An analysis of the conditions conducive to delinquency involvement among high school students: II. On their weakened bonds to family and school. *Reports of the National Research Institute of Police Science, 23,* 147–165.

Thompson, D. W., & Jason, L. A. (1988). Street gangs and preventive interventions. *Criminal Justice and Behavior, 15,* 323–333.

Thrasher, F. (1927). *The Gang: A study of 1,313 gangs in Chicago*. Chicago: University of Chicago Press.

Weis, J. G. (1977). *Comparative analysis of social control theories of delinquency—the breakdown of social controls*. Washington, DC: US Government Printing Office.

Weis, J. G., & Hawkins, J. D. (1981). *Preventing delinquency*. Washington, DC: US Government Printing Office.

Weis, J. G., Sederstrom, J., Worsley, K., & Zeiss, C. (1980). *Family and delinquency*. Washington, DC: US Government Printing Office.

Yablonsky, L. (1997). *Gangsters*. New York: New York University Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.11
Families of Individuals with HIV Infection/AIDS

MARC J. TASSÉ
Université du Québec à Montréal, PQ, Canada
and
SUSAN M. HAVERCAMP
The Ohio State University Nisonger Center, Columbus, OH, USA

9.11.1 INTRODUCTION 170

9.11.2 NATURE OF THE DISEASE 170
- *9.11.2.1 Description of the Disease* 170
- *9.11.2.2 Transmission* 171
  - *9.11.2.2.1 During anal or vaginal sexual intercourse* 171
  - *9.11.2.2.2 Intravenous drug use* 171
  - *9.11.2.2.3 From infected mother to fetus* 171
  - *9.11.2.2.4 Breast-feeding* 171
  - *9.11.2.2.5 Through receipt of contaminated blood or blood-products* 171
- *9.11.2.3 Assessment* 171
- *9.11.2.4 Course and Prognosis* 172
- *9.11.2.5 Opportunistic Infections* 172
- *9.11.2.6 Treatment* 173

9.11.3 IMPACT OF THE DISEASE ON THE FAMILY 173
- *9.11.3.1 Family Stress* 173
- *9.11.3.2 Emotional Impact* 174
- *9.11.3.3 The Well-guarded Secret* 174
- *9.11.3.4 Social Support* 175
- *9.11.3.5 Triangulation* 175
- *9.11.3.6 Healthcare* 176

9.11.4 FAMILY NEEDS 176
- *9.11.4.1 Information and Education* 176
- *9.11.4.2 Medical* 177
- *9.11.4.3 Psychological Needs* 177
  - *9.11.4.3.1 Grieving* 178
  - *9.11.4.3.2 Social support* 178
  - *9.11.4.3.3 Role of the therapist* 179

9.11.5 FUTURE DIRECTIONS 179

9.11.6 REFERENCES 181

9.11.1 INTRODUCTION

Nearly two decades after the first documented case of human immunodeficiency virus (HIV) infection/acquired immunodeficiency syndrome (AIDS), we are just beginning to focus some clinical attention on its impact on families of individuals with HIV infection/AIDS, and on their needs. HIV infection and AIDS represent one of the most serious health threats to humankind and carry with them enormous financial, social, and psychological burdens for those infected, their friends, and family. It is an understatement to say that the effect of learning that one has HIV/AIDS can be traumatic for the person infected. Certain populations have been particularly hard hit by the epidemic including male homosexuals, intravenous drug users, and hemophiliacs or others who have disorders that require the frequent use of donated blood products or organs (Tsiantis et al., 1990; Winiarski, 1997). Families are especially burdened when the news of the HIV diagnosis is coupled with the first disclosure of homosexuality or intravenous drug use of their loved one (Kelly & Sykes, 1989). In the case of hemophilia, infected male hemophiliacs may unwittingly transmit the HIV/AIDS virus to their spouse, who may later transmit the virus prenatally or perinatally to their children.

This chapter will briefly describe the medical aspects of HIV/AIDS and the course of the disease, then we will focus on the impact of and needs created by HIV/AIDS with regard to the family of the person with AIDS. We will present the impact of HIV/AIDS in terms of issues related to medical, emotional, and social adjustments. Following the discussion of the different areas impacted by HIV/AIDS with respect to the family, we will present and discuss the needs incurred by the family of a person with HIV/AIDS. The presentation of these needs will center around information and education, and medical, psychological, and social supports. Finally, we will explore future trends for research in the area of families of individuals with HIV/AIDS, what we know and what we have yet to learn. We identify the areas of needed research, mostly applied, that can aid families of persons with HIV/AIDS to care for and support their loved ones with HIV/AIDS, as well as help them deal with the emotional burden of having a loved one with HIV/AIDS. This chapter will not include issues pertaining to pediatric AIDS. Interested readers should consult Armistead, Forehand, Steele, and Kotchick (1998) for a review.

9.11.2 NATURE OF THE DISEASE

9.11.2.1 Description and Definition of AIDS/HIV

Over 500 000 Americans have been reported with AIDS since the epidemic began (Centers for Disease Control and Prevention, 1997). At least 300 000 of them have died. Researchers estimate that between 650 000 and 900 000 Americans currently carry the human immunodeficiency virus that causes AIDS (Karon et al., 1996).

HIV belongs to a group of viruses, called retroviruses, which integrate their genetic material with a cell's genetic material through the complicated process of reverse transcription (Kalichman, 1995). Retroviruses are composed of two strands of ribonucleic acid (RNA) surrounded by a fatty envelope. The HIV envelope and its surface proteins help it enter other cells. Although HIV is believed to attack many kinds of cells by attaching itself at receptors, its infection of T-helper (T4) cells has received the most attention (Winiarski, 1997). This T4 cell helps orchestrate the immune response and serves as a biomedical marker of immune functioning. After attaching itself, HIV fuses with a protein on the T-helper cell which allows the contents of the virus to enter the T-helper cell. One of the contents, reverse transcriptase, produces deoxyribonucleic acid (DNA) from the viral RNA. This DNA, called provirus, is integrated in the host cell's genetic code and new viral RNA is transcribed. Viral proteins follow, and new complete virus particles begin to bud off from the host cell. Ultimately, the T cell playing host to such viral replications is destroyed. This reduction in the number of T4 cells is a characteristic indicator of the onset of AIDS. Depletion of T4 cells happens gradually. Vague symptoms of illness in the early stages of AIDS reflect the initial depression of the immune system. As greater numbers of T4 cells are overwhelmed, the immune response network is thrown into full decline. The individual is then vulnerable to the opportunistic infections that induce the fatal diseases associated with AIDS.

AIDS is a specific group of diseases or conditions which are indicative of severe immunosuppression related to infection with the human immunodeficiency virus (Centers for Disease Control and Prevention, 1997). AIDS is diagnosed when the patient both tests positive for the HIV infection and also exhibits severe HIV-related immunosuppression as defined by any of the following three conditions: (i) a T-helper count below 200 per cubic millimeter of blood, (ii) the percentage of T-helper cells relative to total T cells is less than 14%, or (iii) the patient is diagnosed with one of 26 "AIDS-defining illnesses" (Centers for Disease Control, 1992). More information about these 26 conditions is provided in section 9.11.2.6

9.11.2.2 Transmission

HIV exists in bodily fluids. HIV and HIV-infected cells have only been isolated in substantial quantities in human blood, vaginal secretions, semen, and breast milk. Insignificantly low quantities of the virus have been found in other body fluids such as saliva and urine (Glasner & Kaslow, 1990). Infection can occur when the bodily fluid of an HIV-infected person enters the bloodstream of a noninfected person. There is no evidence that HIV can be transmitted through casual contacts with, or even daily household exposure to, infected persons (Kalichman, 1995). According to the Centers for Disease Control and Prevention (1994a), even cases of transmission through direct household exposure to blood or other bodily fluids are rare. Transmission occurs in the following ways.

9.11.2.2.1 During anal or vaginal sexual intercourse

Sexual contact appears to be the most common way that HIV is transmitted. Semen or vaginal fluid containing white blood cells infected with HIV comes into contact with genital or rectal tissue. The virus can then enter the bloodstream of the host through micro-openings in the tissue surface. The risk is greatest for individuals engaging in receptive anal sex as there is a high probability of rectal tissue damage. However, penile–vaginal intercourse accounts for the greatest number of cases of HIV infection in the world (Mann, Tarantola, & Netter, 1992). Although transmission between men and women is bidirectional, the rates of transmission are 2.3 times higher for male-to-female infection than female-to-male transmission. There have been reported cases of HIV transmitted by artificial insemination with infected semen (Winiarski, 1997). Oral–genital contact appears to carry lower risk for HIV transmission than does anal and vaginal intercourse; however, it is not possible to say that oral sex with semen or vaginal fluid exchange constitutes a low-risk sexual behavior (Kalichman, 1995).

9.11.2.2.2 Intravenous drug use

Risk occurs when an HIV-infected person uses a syringe to inject drugs and leaves some infected blood in it, and another person then uses the same syringe, thereby injecting infected blood along with the drugs.

9.11.2.2.3 From infected mother to fetus

The mother may pass the virus in her blood across the placenta to the fetus or HIV may be transmitted through blood contamination of the fetus at delivery. Because of the presence of the mother's HIV antibodies, all babies born of HIV-infected mothers test HIV-positive immediately after birth. The mother's antibodies gradually disappear and from 13 to 30% of the babies subsequently develop their own antibodies and are found to be truly HIV-infected. A major study has shown that if pregnant women take zidovudine during the pregnancy and childbirth, and if the baby takes it during its first weeks of life, the risk of the baby's becoming HIV-positive is substantially reduced (Centers for Disease Control and Prevention, 1994a).

9.11.2.2.4 Breast-feeding

HIV exists in the breast milk of infected women and can be directly passed to the infant during breast-feeding.

9.11.2.2.5 Through receipt of contaminated blood or blood-products

In the USA, donated blood has been screened for the presence of HIV since 1985. In other countries, however, such screening is not mandated and does not always occur (Kalichman, 1995).

9.11.2.3 Assessment

Several blood tests exist that detect antibodies to an HIV infection. Because the tests screen for antibodies, and not for the virus itself, those who test positive are sometimes referred to as HIV antibody-positive. When routine testing is done, blood is taken from the individual. The most commonly used procedure is the enzyme-linked immunosorbent assay (ELISA), which employs a chemical method to indicate the presence of the antibodies. Another test, the Western Blot, is frequently used to confirm initial ELISA findings and provide an added safeguard against potentially inaccurate results. When done competently, these tests are very accurate. In a study testing 630 190 units of blood from 290 110 donors, the rate of obtaining a result that falsely indicated someone was HIV-positive was 0.0006 (MacDonald et al., 1989).

Polymerase chain reaction (PCR) is a relatively new test which detects HIV in macrophages before antibodies to the virus have formed. In this test, actual genetic material from the virus is magnified from blood and tissue samples. It is hoped that this screening procedure will provide earlier and more accurate detection because HIV has been known to "hide" in macrophages for variable lengths of time without triggering the production of HIV antibodies.

9.11.2.4 Course and Prognosis

HIV kills T-helper cells. With the death of these cells comes a decrease in immune functioning. The T-helper count is a widely accepted marker of immune function, and the diminishing of T-helper cell numbers is viewed as an important sign of disease progression. Normally, a person has in excess of 800 T-helper cells per cubic millimeter of blood. Over the course of infection, the number of T-helper cells decreases as the cells are killed. During the first three years of HIV infection, the T-helper cell count usually remains above 500 per cubic millimeter of blood and the patient is asymptomatic aside from a possible acute viral reaction at the time of infection. As the T-helper cell count drops to 300, typically during the fourth or fifth year of infection, the patient becomes susceptible to opportunistic infections caused by viruses, bacteria, fungi, and protozoa. When the count drops to fewer than 200 cells per cubic millimeter or less than 14% or AIDS-related conditions develop, the HIV-infected person is classified as having AIDS. The first occurrence of AIDS-related conditions occurs between 8 and 11 years after infection on average, as the T-cell count drops below 100 (Kalichman, 1995).

The time from infection with HIV until the emergence of AIDS is called the incubation period. Many factors may affect an individual's incubation period. Scientists have suggested that the following factors may be important: other viruses, such as cytomegalovirus and Epstein–Barr virus, bacterial infections, and severe malnutrition. Others have suggested that psychological factors such as emotional welfare and hope can play a part. AIDS often led to death in less than one year early in the epidemic. Now, due to the availability of zidovudine, combination therapies, and pneumocystic prophylaxis, longevity has increased dramatically. Osmond (1994), writing before protease inhibitors were licensed, noted that after the first opportunistic infection and with toleration of zidovudine therapy, the median survival time was from 15 months to two years.

9.11.2.5 Opportunistic Infections

Without adequate immune defenses, people suffering with AIDS are susceptible to opportunistic infections and the diseases they engender. The infections connected with AIDS represent a diverse group of pathogens: viruses, bacteria, fungi, and protozoa. Even though everyone is exposed to them daily, these infectious agents do not normally cause illnesses in those whose immune systems are healthy. In patients with AIDS, they are devastating and lethal.

The clinical courses of opportunistic infections vary in severity. Some may recede after an initial manifestation, only to reappear soon after, perhaps in a more virulent form. Certain maladies respond to drug therapies and may be brought under control temporarily. However, patients with AIDS are ordinarily afflicted with multiple infections at the same time. Thus, even if one is successfully treated, another will continue to ravage the body. The areas of the body most often targeted by opportunistic infections include the lungs, skin, gastrointestinal tract, eyes, lymph nodes, and brain (Flanders & Flanders, 1991).

Twenty-six conditions are considered AIDS-defining. The following conditions are among the most frequently occurring: *Pneumocystis carinii* pneumonia (PCP), the most common life-threatening condition in AIDS patients, caused by an organism that is genetically a fungus but is regarded as a parasite; retinitis, caused by cytomegalovirus; candidiasis, caused by fungus; and toxoplasmosis, caused by a protozoan parasite. *Candida albicans*, a yeast-like fungus, initially manifests itself in whitish sores along the mouth and later affects the throat and brain. This can be one of the first signs of an immune system weakened by HIV.

Kaposi's sarcoma (KS), a previously rare cancer affecting older men, manifests clinically as painless purple-to-brown skin lesions on the legs and torso. Later, swollen lymph nodes and toxoplasmosis of the brain develops (Winiarski, 1997). The cause of KS is hotly debated; although it is a cancer, there seems to be a viral contribution to its causes (Winiarski, 1997). Persons with AIDS also get malignancies at a rate that far exceeds that in the normal population. Invasive cervical cancer and lymphoma are two malignancies frequently seen in persons with AIDS (Winiarski, 1997). Particularly difficult for family members to cope with is the neurological condition of AIDS-related dementia. The latter is a collection of mental and behavioral disorders related to the toxic effects of HIV infection of the brain. It occurs in as many as 90% of patients in the later stages of

AIDS (Kalichman, 1995). This condition can cause alterations of consciousness, attention span, short- and long-term memory, motor disturbances, and mood stability and lability (Kelly & Sykes, 1989).

9.11.2.6 Treatment

Zidovudine (referred to as AZT, ZDV, or Retrovir) was the first major drug discovered to have an effect against HIV. It belongs to a class of drugs called reverse transcriptase inhibitors that interfere with the replication of HIV. Three additional reverse transcriptase-inhibiting drugs have been developed and tested: dideoxycytidine (ddC), dideoxyinosine (ddI), and 3′-deoxythymidine-2′-ene (d4T). Each functions similarly to zidovudine by inhibiting the action of reverse transcriptase and therefore interfering with HIV replication (Kalichman, 1995). Zidovudine is credited with extending the lifespan of many persons with AIDS. It is now studied and more widely used in combination with other drugs. Common side-effects include nausea, fatigue, malaise, and insomnia. Some individuals develop more severe side-effects such as bone marrow suppression and a reduction in white and/or red blood cell counts. These patients are taken off the medicine immediately. Most, if not all, patients treated with zidovudine for a long time develop some resistance to it, although the clinical significance of this is not always clear.

Protease inhibitors are a newer class of drugs that attack HIV at a different point in the reproduction process than does zidovudine. Typically, a protease inhibitor is prescribed along with other drugs to create a many-pronged attack on the virus. This results in drug regimens requiring 15 or more pills daily. At least initially, protease inhibitors were priced at thousands of dollars for a year's supply.

The typical treatment strategy for an asymptomatic HIV-infected person has two components. The first is an attack on the virus's replication process; the second is prevention of opportunistic infections. The attack on viral replication is typically accomplished through combination therapies: the simultaneous use of different types of drugs to disrupt replication at various stages of the process. The US Public Health Service and the Infectious Diseases Society of America have issued extensive guidelines for the prevention of opportunistic infections. These guidelines appeared in *Clinical infectious diseases*, and reprints are available from the National AIDS Clearinghouse (USPHS/IDSA Prevention of Opportunistic Infections Working Group, 1995). Standard-of-care recommendations include medications to prevent the most common opportunistic infections, including those caused by *Pneumocystis carinii*, or *Toxoplasma gondii*, and tuberculosis, and often *Mycobacterium avium* complex infections as well.

9.11.3 IMPACT OF THE DISEASE ON THE FAMILY

9.11.3.1 Family Stress

The impact of HIV/AIDS extends beyond the infected individual to that person's social support network. Referred to by some as the "secondary epidemic" (Turner & Catania, 1997), the effects of AIDS on the infected individuals' social support network may include family conflicts, shifts in social roles, financial hardships, deteriorating physical and mental health, and feelings of helplessness, grief, and loss. Research in the area of informal caregiving for HIV/AIDS is increasing (e.g., Pearlin, Mullan, Aneshensel, Wardlaw, & Harrington, 1994; Turner & Catania, 1997; Wardlaw, 1994). Informal caregivers have been defined as friends and family members (Turner & Catania, 1997). Many individuals with HIV/AIDS, who may have long been alienated by their parents and siblings, have come to refer to their friends and partner as their "family of choice" (Britton, Zarski, & Hobfoll, 1993; Nord, 1996; Wardlaw, 1994). The network of informal caregivers is more often comprised of the family of choice than of the family of origin (Nord, 1996). In surveys of informal caregivers, the person with HIV/AIDS was identified as a friend in 79% (Turner & Catania, 1997) and friend or lover/partner in 90% (Wardlaw, 1994) of cases. In contrast to research findings on informal caregivers with respect to other illnesses (e.g., cancer, Alzheimer), a large proportion of informal caregivers for HIV/AIDS patients are men (Pearlin et al., 1994; Turner & Catania, 1997; Wardlaw, 1994). HIV/AIDS has been evaluated as the most demanding of chronic illnesses with regard to the care provided, in terms of medical, psychological, emotional, and financial support (Tehan, 1991).

Families of choice also seem to have an important place in the lives of intravenous drug users, whose networks of friends are considered to provide more social support that their biological family members (Stowe, Ross, Wodak, Thomas, & Larson, 1993). Like people who are gay, intravenous drug users feel it necessary to constitute families of choice due to rejection or abandonment by the family of origin. When dealing with HIV/AIDS, the family has been defined by the Global Program on AIDS of the

World Health Organization (1994) as a group of persons linked by feelings of trust, mutual support, and common destiny. In this chapter, when discussing the impact on families, and their needs we will refer to both family of origin and family of choice, unless otherwise indicated.

Britton et al. (1993) described the difficulties many families have in dealing with the fact that their adult child is gay and has HIV/AIDS. Many families may not have yet come to grips with their child's homosexuality, denying to themselves and refusing to talk about it with others. These families are poorly prepared to deal with the devastating news that their child now is infected with HIV/AIDS, and often become immobilized by this information or, worse, constitute a source of negative support for the person with HIV/AIDS. Many families experience enormous stress when learning of AIDS and perhaps learning for the first time about their adult child's homosexuality or drug use (Caputo, 1985). Greif and Porembski (1989) reported a less severe impact of the AIDS diagnosis on families where the homosexuality of the person with HIV/AIDS had been previously accepted.

9.11.3.2 Emotional Impact

Nichols (1985) referred to the affective reaction in the case of AIDS as the "AIDS adjustment process" and identified four distinct phases: (i) crisis, (ii) transition, (iii) acceptance, and (iv) preparation for death and bereavement. Many people react with an extreme sense of rage and anger when they discover the HIV/AIDS diagnosis of a loved one. This anger can be diffuse and directed towards federal agencies for not moving quickly enough towards a cure or towards society in general for its stigmatization of persons with HIV/AIDS (Kelly & Sykes, 1989). As previously mentioned, learning of a loved one's HIV/AIDS diagnosis is often the first time the family of origin learns of their homosexuality or intravenous drug use. In families that have known or suspected but did not accept the sexual orientation or drug use, learning of the infection challenges their denial of their loved one's lifestyle. The announcement of the HIV/AIDS diagnosis can in some cases exacerbate family conflicts and tensions, rather than bring the family closer together in this time of need for support (Nord, 1996). In a review of the research literature on social support and HIV, Green (1993) concluded that although there is still much research needed in this area, there is evidence of a link between the psychological well-being of individuals with HIV/AIDS and their social support. For family members to provide this much needed support, they must first come to grips with the news of their loved one's homosexuality or intravenous drug use and HIV/AIDS diagnosis.

One demanding ailment often associated with AIDS is AIDS-related dementia. Its effect on the neurocognitive functioning of the person with HIV/AIDS may result in memory loss, mood disorder, thought disorders, attention problems, and behavior problems. Greif and Porembski (1989) noted the emotional reactions, including denial, guilt, anger, sadness, depression, and shame, frequently experienced by family members of the person with HIV/AIDS. Greif and Porembski also noted three common behavioral patterns within the families of persons with HIV/AIDS: (i) secrecy regarding the HIV/AIDS diagnosis, (ii) triangulation, and (iii) lingering feelings of hurt and anger.

9.11.3.3 The Well-guarded Secret

The proportion of people who say they personally know someone who has HIV/AIDS may in fact be higher than the 15% cited by Turner and Catania (1997), especially considering that many individuals with HIV/AIDS fear letting others know of their illness. People with HIV/AIDS must painfully grapple with the learning of their diagnosis and then determine who they want to or must inform.

Family acceptance or awareness of the sexual orientation or drug use of the person with HIV/AIDS also emotionally prepares the latter to reveal their HIV/AIDS status (Leask, Elford, Bor, Miller, & Johnson, 1997). Because of the strong stigma, keeping the HIV/AIDS diagnosis a secret from friends, families, and others is one of the first issues the person with HIV/AIDS grapples with. The stigma associated with AIDS and homosexuality compounds the already stressful situation (Crawford, 1996; Takigiku, Brubaker, & Hennon, 1993). Even among seronegative gay adolescents, social stigma has been associated with an increased risk for suicide (Radkowsky & Siegel, 1997). Some have referred to the AIDS stigma as a modern-day leprosy (Geis, Fuller, & Rush, 1986). This situation can be so painful that some persons with HIV/AIDS prefer to commit suicide rather than face telling their parents (Kubler-Ross, 1987). Failure on the part of the person with HIV/AIDS to disclose their HIV/AIDS diagnosis to members of their family of origin is even more widespread amongst certain cultural groups such as Latinos (Mason, Marks, Simoni, Ruiz, & Richardson, 1995).

Many family members dread the stigma associated with AIDS and battle with issues of disclosure long after the death of their loved one. Four of nine families interviewed by Greif and Porembski (1989) continued attempts at secrecy even after the death of the person with HIV/AIDS. Despite efforts to educate and inform, social stigma associated with people with HIV/AIDS is prevalent in all societies and is cited as the main reason why the friends and families of many persons with HIV/AIDS find it difficult to disclose (Greif & Porembski, 1988; Harek & Capitano, 1993). The secrecy issues for many family members of the person with HIV/AIDS may be related to shame, or fears of job loss or ostracism, and occur even when the transmission of AIDS was through a blood transfusion (Kelly & Sykes, 1989). Group support may be an effective means to helping families and persons with HIV/AIDS deal with secrecy issues and stigma associated with the HIV/AIDS diagnosis (Beckerman, 1994; Kelly & Sykes, 1989). Such support is especially important because this refusal to disclose to others leaves them isolated and cut off from natural social supports (Pomeroy, Rubin, & Walker, 1995, 1996).

In a study that included 34 gay men seeking treatment in a London clinic, the majority reported a positive change in their relationship with their family of choice following disclosure of their HIV status. A negative change was reported in the relationship with their lover or partner by 33% of men (Leask et al., 1997). When disclosing HIV diagnosis to their family of origin, Leask et al. reported that the gay men in their sample were more likely to disclose to their mother or sister than to their father or brother. Disclosing to their mothers rather than to their fathers may not be surprising, considering that mothers tend to have more positive attitudes than fathers towards homosexuality (Takigiku et al., 1993). However, similar disclosure patterns have been reported in heterosexual couples where one partner is seropositive (Kennedy, Skurnick, Foley, & Louria, 1995). Thus, even within the same family, some members may know of the HIV/AIDS diagnosis while others are not told.

9.11.3.4 Social Support

Greif and Porembski (1989) interviewed 11 significant others after the death of their loved ones. They reported that a common coping strategy used by family members of persons with HIV/AIDS was denial. Denial can be defined according to Billings and Moos (1981) as an avoidant coping measure. The use of avoidant coping strategies has been found to be a risk factor for increased psychological distress in persons with HIV/AIDS and their significant others (Klein, Forehand, Armistead, & Wierson, 1996). The lack of information regarding HIV/AID was also noted as an element that hindered their ability to better cope with the situation. The emotional impact of learning of a loved one's HIV/AIDS diagnosis was emphasized by almost all persons interviewed. Finally, the stigma of HIV/AIDS and the irrational fears of contagion prevalent in the general public frequently isolates family members from their own usual sources of social support.

Beckerman (1994) suggested that the parents and the person with HIV/AIDS be seen in a therapy session together in order to explore their respective fears and concerns. Acceptance of the diagnosis on the part of the parents may take some time. According to Beckerman (1994), if a family is to become supportive of their loved one with HIV/AIDS they will need to: (i) accept the diagnosis, (ii) accept their child's homosexuality, (iii) disclose (diagnosis and lifestyle), (iv) reintegrate their child into the family, and (v) let go, prepare for the death of the person with HIV/AIDS.

Social stigma associated with AIDS remains undeniably very present in all societies and contributes to the hardships attributed to this disease. Much of the negative attitude directed towards people infected with HIV/AIDS and to AIDS in general, may be largely a homophobic reaction (Crawford, 1996; Morin, 1991; Pryor, Reeder, Vinacco, & Kott, 1989; Trezza, 1994). In addition to society's negative reaction to homosexuality, irrational fear of contracting the HIV/AIDS virus is also fueling the shunning of persons with HIV/AIDS and anyone associated to them (Bishop, Alva, Cantu, & Rittiman, 1991). The severe stigma associated with HIV infection and AIDS tops the list of contributing factors to increased risk of suicide amongst persons with HIV/AIDS and their family members (see Wadland & Gleeson, 1991). For example, in a meta-analysis of 21 studies of AIDS stigma, Crawford (1996) found a greater stigma associated with AIDS than with other serious illnesses (e.g., leukemia, Alzheimer, lung cancer). Crawford also found that amongst healthcare professionals a greater magnitude of stigma was associated with AIDS.

9.11.3.5 Triangulation

One frequent target for the anger of families of origin of gay men with HIV/AIDS is the lover or partner (Dworkin & Pincu, 1993). Tensions between the parents of gay persons with HIV/AIDS and their child's lover or partner are common (Greif & Porembski, 1988; Lovejoy, 1989). Frequently, parents of a gay person with HIV/AIDS may accuse the lover, partner, or family of choice of being responsible for their son's HIV infection (Lovejoy, 1989). Even after the death of the person with HIV/AIDS, tensions and arguments erupt between the family of origin and the family of choice over funeral arrangements and disposition of belongings (Wardlaw, 1994). These reactions obviously cause additional strain and tension in an already painful situation.

9.11.3.6 Healthcare

With the ever-increasing cost of healthcare, more and more families of origin and families of choice are being called upon to provide the bulk of the caregiving to persons with HIV/AIDS. Many of these informal caregivers routinely help with such tasks as medical procedures, personal care, and managing the financial and legal affairs of the person with HIV/AIDS (Wardlaw, 1994). The discovery and refinement of effective drug regimens for suppressing symptoms and forestalling AIDS progression changes the nature of caregiving for persons with HIV/AIDS from acute to chronic. In their sample of 642 caregivers, Wardlaw (1994) reported an average caregiving duration of 1.5 years, and some participants reported providing caregiving to a person with HIV/AIDS for more than four years. While the family cares for their loved one, vacations, leisure activities, and long-term plans are put on hold. The burden of having to provide such long-term healthcare can be devastating for family members, especially since most of them receive no caregiving training of any kind (Wardlaw, 1994).

One concern of family members who provide home healthcare to loved ones with AIDS is the risk of being infected. Informal caregivers help with dressings on wounds and sores, injections, changing of soiled diapers and bed sheets, feedings, changing of catheters, and the starting of intravenous lines. The medical needs of persons with HIV/AIDS, and hence the demands upon their family, vary enormously with the course of the illness, at times involving intense medical care and at other times primarily social services (Milanese et al., 1997). Aside from the concerns of possible infection, these tasks executed by family members demand considerable time and physical strength and the impact on the family can be overwhelming (Nord, 1996). We will later discuss the needs of family members with respect to education concerning the risks of HIV transmission and the precautions to be taken in order to avoid being infected.

In the USA, the financial burden can be great for families with no health insurance or with a limited health insurance that may cap out quickly. Even for families living in a country that provides universal healthcare (e.g., Canada, Sweden) financial burdens may exist regarding unsubsidized types of care or medications. Persons with HIV/AIDS or their family members often lose their jobs because of inability to work, prolonged absences or an inability to continue working while providing care to a loved one. The economic strain on society attributed to AIDS has at times been alarmingly sounded, such as Lynch's (1989) prediction of bankruptcy of hospital and city administrations in the hard-hit areas (such as New York City and San Francisco). Such alarmist statements must be viewed with skepticism. The greatest financial burden created by the HIV infection and AIDS will be borne by persons with HIV/AIDS and their families.

9.11.4 FAMILY NEEDS

9.11.4.1 Information and Education

Information and education regarding the nature of HIV/AIDS are crucial in reassuring family members who may harbor irrational fears of contagion (Pomeroy et al., 1995). We only need to look at the historical perspectives of HIV/AIDS information to realize the degree of misinformation and ignorance that has circulated over the years. Initial misleading information announced that the AIDS virus could be transmitted by all bodily fluids, including saliva and tears. Public concern escalated with the speculation that HIV could be transported by mosquitos and that possible mutations could occur that would allow the AIDS virus to become airborne (Lynch, 1989). Stigma and social rejection has certainly been fueled by fear and inaccurate information.

Information and education needs to center on basic information regarding the transmission of HIV. As previously mentioned, transmission of HIV is possible through the use of infected needles or syringes during intravenous drug use, by having unprotected sexual intercourse (oral, anal, or vaginal) with an infected person (male or female), or by receiving a blood transfusion or organ/human tissue transplant from an infected donor. Infected mothers can also transmit the HIV virus to their fetus through the placenta or to their babies perinatally or through breast milk.

The HIV virus is not transmitted during casual contact with an infected person. It is important to educate family members in order to alleviate the fears they may have of contracting the HIV virus through activities of daily living or caregiving with a loved one who has AIDS. Below is a list published by the Centers for Disease Control and Prevention (1994b) of ways through which you cannot "catch" the HIV virus.

You will not be infected:

(i) through everyday contact with an infected person at home, work, school, etc.;

(ii) from clothes, phones, or toilet seats. HIV cannot be passed on by things like forks, cups, or other objects that an infected person has used;

(iv) from a mosquito bite. HIV cannot survive in a mosquito, and is not transmitted through a mosquito's bite like malaria. HIV is not transmittable through bedbugs, lice, fleas, or other insects;

(v) from an infected person's sweat, saliva or tears;

(vi) from a simple kiss from an infected person. Although, HIV transmission may be possible through prolonged or deep kissing, due to potential blood contact, most scientists agree that this would be most unlikely (Centers for Disease Control and Prevention, 1994b, p. 2).

If there is a need for education regarding the basics of HIV and AIDS amongst mental health professionals (Knox, Dow, & Cotton, 1989), it is even more pressing for family members called upon to care for a loved one with AIDS. Amongst the family members caring for a person with HIV/AIDS in Wardlaw's (1994) sample of 642 family member/caregivers, less than a third had gotten any type of training on how to care for a person with HIV/AIDS. Knox and Gaies (1990) proposed a model for the provision of effective AIDS education to community mental health professionals. This model can be adapted to educate family members and would consist of the following elements: (i) small group format (6–8 people); (ii) a didactic component to provide basic information about HIV/AIDS; (iii) an experiential component, such as behavioral rehearsal and role playing (e.g., practice safety precautions to guard against infections); and (4) group discussions and establishing social network/support.

According to Trezza (1994), HIV prevention programs geared towards college students should aim at reducing homophobia, which is strongly correlated with stigmatization of persons with HIV/AIDS, and address irrational fears pertaining to HIV/AIDS, education concerning high- versus low-risk behaviors, and safe sex techniques. Despite harboring their own negative attitudes towards gays and AIDS, psychologists, having been trained in behavior and attitude change, can become key figures in the elaboration of HIV/AIDS prevention and education programs (Trezza, 1994).

The American Psychological Association's HIV Office for Psychology Education (HOPE) program offers training workshops for psychologists to better prepare them to do therapy with persons with HIV/AIDS and their family. These continuing education workshops aim at providing current training with regard to case management, psychotherapy, and client education skills, to enable psychologists to work effectively with individuals with HIV infection/AIDS and their families. The HOPE program covers the following areas: (i) HIV disease information related to psychotherapy, (ii) one's own attitudes and values about HIV and individuals with HIV infection, (iii) psychosocial issues confronting clients, (iv) HIV-related counseling skills; (v) barriers to effective clinical practice, (vi) networking opportunities, (vii) awareness of local HIV/AIDS resources; and (viii) AIDS-related grief, loss, and bereavement (Eversole, 1995, p. 206).

9.11.4.2 Medical

Finding out that one's child or sibling has been diagnosed with HIV/AIDS can have a devastating effect on the entire family. Many family members may have only limited knowledge of the disease and require immediate information and education regarding the course of the illness, symptoms, opportunistic infections associated with AIDS, necessary precautions regarding contagion, and the types of care needed by a person with HIV/AIDS. Many family members may be called upon to provide care to their child or sibling who has been diagnosed with HIV/AIDS and will require additional information.

As the medical treatment of HIV/AIDS extends the lives of individuals infected, the once acute type of illness becomes a more chronic and prolonged one that requires extended care from family and friends. As the life expectancy of persons with HIV/AIDS is increased so are the needs of the family members who are called upon to provide care.

Education is needed regarding medication compliance, medication side-effects and interaction effects, signs of dementia and how to intervene, symptoms and course of HIV/AIDS (e.g., thrush, opportunistic infections, etc.), precautions to take to prevent infection, nutritional needs of persons with HIV/AIDS, and medical resources available for home healthcare.

9.11.4.3 Psychological Needs

Counseling and support to help family members cope with the HIV/AIDS infection status of their loved one is an important and essential step in providing services for families of individuals with HIV/AIDS (Gurney, 1995; Nord, 1996). Greif and Porembski (1989) pointed out the importance of providing therapeutic services to the members of the families of origin and choice, due to their role as primary caregivers to persons with HIV/AIDS and the toll this takes with respect to their own emotional health.

Support groups for family members of persons with HIV/AIDS seem helpful in providing a forum for open discussion and exchange with others about feelings, problems, and fears without being ostracized or stigmatized (Kelly & Sykes, 1989). Considering the societal stigma associated with HIV/AIDS, group-based interventions may be helpful in breaking the barriers of isolation and provide support to these family members.

An example of a more structured group intervention program was published by Pomeroy et al. (1995, 1996) where they presented their eight-session psychoeducational group for family members of persons with HIV/AIDS. During these eight sessions, participants are provided with information regarding topics such as dispelling myths about HIV/AIDS, HIV transmission, home healthcare and precautions to take in the home, the course of disease from HIV infection to AIDS, medical and nutritional needs of persons with HIV/AIDS, current treatment modalities and medication, opportunistic infections, and dementia associated with later stages of AIDS. Group members are also provided with a listing of existing resources in their communities that can provide them with a variety of medical or social services that are specialized in the care of persons with HIV/AIDS. Counseling and information are provided regarding insurance, legal aid, and legal aspects related to death (e.g., living wills) as well as concerning grief and the grieving process.

In their group intervention program, Pomeroy et al. (1995, 1996) also provided minimal skills training to the family members of persons with HIV/AIDS in the areas of anger management, relaxation training, assertiveness training, problem solving, and conflict resolution techniques. In addition to the education and support fostered during the group meetings, they added homework tasks that members were asked to complete and report upon at each meeting. These tasks involved phoning another group member and initiating a conversation in an attempt to foster social support amongst group members, finding and practicing a constructive outlet for their anger, and engaging in one leisure activity.

This psychoeducational group intervention program provides a range of needed skills and information for family members of persons with HIV/AIDS and an important component of peer social support for these family members. Their pre- and post-support group assessment indicated a significant improvement on the part of the participants, when compared to a wait-list group, on measures of perceived stigma associated with HIV/AIDS, depression, anxiety, and perceived stress. Surprisingly, the measures of perceived social support showed no appreciable change between the start and completion of the psychoeducational group. It may be that eight 90-minute sessions are insufficient to significantly ameliorate social support deficits in this population. The authors also acknowledged problems in outcome measurement sensitivity.

9.11.4.3.1 Grieving

Often, biological parents who learn of their child's HIV/AIDS diagnosis feel a sense of grief at the announcement of the diagnosis and may experience prolonged grief as the condition of the child deteriorates (Greif & Porembski, 1989; Wadland & Gleeson, 1991). Some families experience multiple losses when a loved one has HIV/AIDS. In some cases, such as the family of choice of gay men or the biological family of people with hemophilia, they may have experienced the death of more than one loved one as a result of HIV/AIDS. In some gay communities, entire groups of friends have lost their lives to HIV/AIDS. Nord (1996) presented the clinical case of a mother whose two hemophiliac sons contracted HIV and died of AIDS.

Discussing death is often avoided because of the immense emotional pain associated with it. However, being able to openly discuss the impending death of a loved one is critical for the family to begin healing from the loss (Millan & India Ivory, 1994). Honest discussion about death and dying is the penultimate indication of the level of trust between the client/group and the clinician. Dealing with grieving can only be effectively accomplished within a relationship of trust and emotional comfort.

9.11.4.3.2 Social support

The major stumbling block to families getting the much needed social support, from friends, family members, or professionals, is secrecy. Fears that fuel the need to keep the HIV/AIDS diagnosis of a loved one a secret isolate the family from everyone who is excluded from the circle of people who know. Group meetings between family members of persons with HIV/AIDS and therapist are an effective medium during which secrecy issues can be explored and discussed. Social support is crucially important in helping family members cope with the loss of a loved one and support groups are a valuable opportunity for family members to realize that others care and understand how they feel (Nord, 1996).

Kelly and Sykes (1989) studied the composition and structure of a support group intervention for family members of persons with HIV/AIDS. These authors ran a bi-monthly support group composed of a mix of family of origin and family of choice members. The support group provided a forum in which members could safely disclose that their loved one had AIDS, share with others their difficulties related to caring for a person with HIV/AIDS, and ventilate their rage and anger, as well as deal with issues of acceptance of their loved one's homosexuality or intravenous drug use.

Pomeroy et al. (1995, 1996) addressed issues of lack of social support in addition to providing information, education, and some coping skills training. Social support was provided by the group members within and between the group meetings. Members exchanged stories regarding their experiences in caring for a loved one with HIV/AIDS, the anger and guilt they at times felt towards the person with HIV/AIDS, their feelings of helplessness and depression, and how they dealt with these different issues. The fact, however, that group members did not report significantly higher perceptions of social support may indicate a need for more pervasive supports or the inclusion of natural supports such as friends and family members.

9.11.4.3.3 Role of the therapist

A certain cautionary note is needed with respect to cultural differences and the clinician's attempt to progress through different areas of need. Millan and India Ivory (1994) pointed out the important need for Latinos to feel a strong personal and intimate rapport with the therapist before dealing with issues such as completing a living will. The clinician must recognize the importance of ensuring that persons with HIV/AIDS and their family members are aware of and prepared for the important financial issues tied to their disease and that they are aware that their financial needs will certainly fluctuate with the course of the disease (Wadland & Gleeson, 1991).

Secrecy is a critical element in the client–therapist relationship. The therapist must abide by the desire for secrecy of a person with HIV/AIDS and must support the client who has HIV/AIDS in making progress towards dealing with the diagnosis and the difficult issues of disclosure (Greif & Porembski, 1989). Yet, Dworkin and Pincu (1993) recommended that clinicians ensure that their client is not using nondisclosure as a means to "avoid" dealing with the disease. This avoidance coping can contribute to more distress and emotional problems (Fukunishi et al., 1997). The clinician can play a crucial role in coaching and helping the person plan and rehearse how to tell people (e.g., role-play). Persons with HIV/AIDS, in an attempt to take charge of the situation and project an image of self-assurance, conversely, may decide to disclose their HIV/AIDS status to everyone they know. Dworkin and Pincu (1993) recommended clinicians prepare these clients for possible negative reactions from some individuals. Clinicians must support the person with HIV/AIDS in making rational and careful disclosure of their diagnosis. Psychologists must be aware of their biases and negative attitudes towards homosexuals and HIV/AIDS. Trezza (1994) reported that in his sample of nearly 600 clinical and counseling psychologists, many lacked the knowledge regarding HIV/AIDS that is necessary for effective counseling of persons with HIV/AIDS and their families.

9.11.5 FUTURE DIRECTIONS

With the advent of more effective means for slowing down the progression of the HIV virus, AIDS care is moving from a relatively short-term care plan to more chronic long-term managed care needs. Wardlaw (1994) advanced the concern regarding the resilience of families with respect to the emotional and physical strain involved in caring for a loved one with a lengthening interval between the AIDS diagnosis and death.

Given the ever-increasing desire to reduce healthcare costs, supports will be needed to provide home-based healthcare and help maintain HIV/AIDS infected people in their homes versus more expensive hospital settings. This emphasis on home-based healthcare creates an increased burden on family members, who will be expected to assume a greater caregiver role and heavier responsibilities. Service providers will need to plan the provision of needed family supports in educational, financial, medical, social, and psychological areas. Pomeroy et al. (1995) suggested additional research, to study the contribution of psychoeducational interventions that support informal caregivers in providing in-home healthcare to persons with HIV/AIDS, and the effectiveness of these education programs in reducing hospitalizations of persons with HIV/AIDS.

Clearly, social supports are a buffer, protecting the person with HIV/AIDS and their family from psychological distress associated with AIDS (Green, 1993; Klein et al., 1996). Further research is needed to evaluate and improve the effectiveness of support groups in significantly augmenting the level of perceived social support of family members and persons with HIV/AIDS. This need may be most keenly felt by family members who resist disclosure and are therefore cut off from their pre-established social support networks.

Given the over-representation of African-American and Hispanic persons with HIV/AIDS, more attention should be given to cross-cultural differences in counseling, support groups, and service delivery with respect to HIV/AIDS. Kelly and Sykes (1989) observe the under-representation of minority members in their support group sample, and emphasize the importance of offering support groups in a manner that ethnic minorities can and will accept. Studies examining caregiving have raised interesting questions regarding gender and cultural differences in the perceived burden on caregivers (see Takigiku et al., 1993; Turner & Catania, 1997). Given the unusually high percentage of male caregivers of persons with HIV/AIDS, research is needed to investigate possible gender differences in the disruptions caregiving creates within caregivers' jobs and social relationships. Further research is needed to study the relationship between individual caregiver characteristics and resilience to the caregiving burden, psychological distress, and complicated grieving. A detailed examination of risk factors for poor caregiving outcome is needed (Turner & Catania, 1997). Factors that may differentiate high- versus low-stressed parent caregivers include the family ethos of affection versus obligation, and attitudes toward homosexuality need futher investigation (Takigiku et al., 1993).

Additional research and clinical work is needed to provide a better understanding of the relationship between psychological distress experienced by family members and their perceived social support. Drotar, Agle, Eckl, and Thompson (1997) recommended the study of family relationships and family supports, including any fluctuation in these supports or changes in relationships over the course of the disease. They have suggested the need to study in greater depth the relationship between family support, stressful events, and psychological distress amongst the family members, both within the HIV/AIDS families and healthy families. Further, they suggested that research should identify specific elements of family relationships that parents experience as most supportive and hence may be most influential in reducing the psychological distress associated with AIDS. Future research is needed to better understand the coping behaviors of persons with HIV/AIDS and their families. As pointed out by Pomeroy et al. (1995), few studies have attempted to follow on a longitudinal basis the impact of HIV/AIDS on the family, and their needs as the disease progresses.

Wardlaw (1994) has pointed out the dearth of studies regarding the study of the psychological strain from AIDS mortality and morbidity on family members who have cared for and survived several other family members. Further studies are needed to better understand the repercussions of HIV infection/AIDS on families, even after the death of the person with HIV/AIDS.

Future studies will also need to address the issue of social stigma. Social stigma is one of the major variables contributing to increased stress and isolation for persons with HIV/AIDS and their families (Crawford, 1996; Greif & Porembski, 1988; Harek & Capitano, 1993; Trezza, 1994). Following her meta-analysis of 21 studies, Crawford (1996) concluded that there is a pressing need for further theory-driven empirical investigation of AIDS stigma, in order to better understand the underlying and mediating variables fueling the AIDS stigma in our societies, and to identify possible strategies to reduce this stigma. The elevated social stigma associated with AIDS is present in both the lay and professional populations (Crawford, 1996). Along with social stigma, future research also needs to examine the different patterns of AIDS disclosure and its impact within the family (Levine, 1995).

Research is needed to evaluate the impact on children who lose family members to AIDS and to assess their needs (Carten & Fennoy, 1997). Children and adolescents suffer the loss of a family member and disruption of the family without knowing what is really happening. Youngsters are kept in the dark in a misguided attempt to protect them. The discovery of the HIV infection and possible homosexuality of a family member may be especially difficult for the young, especially for adolescents who are still confronting their own sexuality. Research is needed to investigate the effects of information and support groups targeted at children and adolescents (Aronson, 1995; Kelly & Sykes, 1989; Levine, 1995).

We need to further study and better understand the role of social support in buffering against psychological distress, depression, stress, and other negative effects HIV/AIDS has upon the person with HIV/AIDS and the family. More empirical work is needed to study the effects of actual social supports (i.e., inventory) versus perceived social support on these psychological variables, and how the buffering effects of social support fluctuate over the course of the disease. Britton et al. (1993) found that support systems may or may not be perceived as helpful depending on the perceived distress of the person with HIV/AIDS. Further research towards understanding this phenomenon may help to plan interventions that will be perceived as beneficial and timely.

More work is needed in educating family members about this disease, and it is crucial to tailor this information or education to the culture and level of comprehension of the targeted audience. Improvements in AIDS awareness and education should focus on reducing high-risk behaviors associated with HIV infection, but also on promoting a better understanding of homosexuality, and of HIV infection and contagion, and on attacking the existing and pervasive irrational fear of AIDS, sometimes refer to as "AFRAIDS" (Doubleday, 1986).

9.11.6 REFERENCES

Armistead, L., Forehand, R., Steele, R., & Kotchick, B. (1998). Pediatric AIDS. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of child psychopathology* (pp. 463–481). New York: Plenum.

Aronson, S. (1995). Five girls in search of a group: A group experience for adolescents of parents with AIDS. *International Journal of Group Psychotherapy, 45,* 223–235.

Beckerman, N. L. (1994). Psychosocial tasks facing parents whose adult child has AIDS. *Family Therapy, 21,* 209–216.

Billings, A., & Moos, R. (1981). The role of coping responses and social resources in attenuating the stress of life events. *Journal of Behavioral Medicine, 4,* 139–157.

Bishop, G. D., Alva, A. L., Cantu, L., & Rittiman, T. K. (1991). Responses to persons with AIDS: Fear of contagion of stigma? *Journal of Applied Social Psychology, 21,* 1877–1888.

Britton, P. J., Zarski, J. J., & Hobfoll, S. E. (1993). Psychological distress and the role of significant others in the population of gay/bisexual men in the era of HIV. *AIDS Care, 5,* 43–54.

Caputo, L. (1985). Dual diagnosis: AIDS and addiction. *Social Work, 30,* 361–364.

Carten, A. J., & Fennoy, I. (1997). African American families and HIV/AIDS: Caring for surviving children. *Child Welfare, 76,* 107–125.

Centers for Disease Control (1992). 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *Morbidity and Mortality Weekly Report, 41*(RR-17), 1–19.

Centers for Disease Control and Prevention (1994a). Recommendations of the US Public Health Service Task Force on the use of zidovudine to reduce perinatal transmission of human immunodeficiency virus. *Morbidity and Mortality Weekly Report, 43*(RR-11), 1–20.

Centers for Disease Control and Prevention (1994b). *HIV infection and AIDS: Are you at risk?* Atlanta, GA: US Department of Health and Human Services.

Center for Disease Control and Prevention. (1997). *HIV/AIDS surveillance report* (Vol. 9 No. 1, pp. 1–37). Atlanta, GA: US Department of Health and Human Services.

Crawford, A. M. (1996). Stigma associated with AIDS: A meta-analysis. *Journal of Applied Social Psychology, 26,* 398–416.

Doubleday, W. A. (1986). Death, dying, and AIDS. In V. Gong & N. Rudnick (Eds.), *AIDS: Fact and issues* (pp. 291–302). New Brunswick, NJ: Rutgers University Press.

Drotar, D., Agle, D. P., Eckl, C. L., & Thompson, P. A. (1997). Correlates of psychological distress among mothers of children and adolescents with hemophilia and HIV infection. *Journal of Pediatric Psychology, 22,* 1–14.

Dworkin, S. H., & Pincu, L. (1993). Counseling in the era of AIDS. *Journal of Counseling and Development, 71,* 275–281.

Eversole, T. (1995). Family therapists' response to AIDS: An examination of attitudes, knowledge and contacts. *Journal of Marital and Family Therapy, 21,* 205–206.

Flanders, S. A., & Flanders, C. N. (1991). *Library in a book: AIDS*. New York: Facts On File.

Fukunishi, I., Hosaka, T., Negishi, M., Moriya, H., Hayashi, M., & Matsumoto, T. (1997). Avoidance coping behaviors and low social support are related to depressive symptoms in HIV-positive patients in Japan. *Psychosomatics, 38,* 113–118.

Geis, S. B., Fuller, R. L., & Rush, J. (1986). Lovers of AIDS victims: Psychological stresses and counseling needs. *Death Studies, 10,* 43–53.

Glasner, P. D., & Kaslow, R. A. (1990). The epidemiology of human immunodeficiency virus infection. *Journal of Consulting and Clinical Psychology, 58,* 13–21.

Global Program on AIDS of the World Health Organization (1994). *World AIDS Day Newsletter, 2,* 1.

Green, G. (1993). Social support and HIV. *AIDS Care, 5,* 87–104.

Greif, G. L., & Porembski, E. (1988). AIDS and significant others: Findings from preliminary exploration of needs. *Health and Social Work, 13,* 259–265.

Greif, G. L., & Porembski, E. (1989). Implications for therapy with significant others of persons with AIDS. *Journal of Gay and Lesbian Psychotherapy, 1,* 79–86.

Gurney, S. (1995). Counseling the HIV affected individual: A case study. *Counseling Psychology Quarterly, 8,* 17–25.

Harek, G. M., & Capitano, J. P. (1993). Public reactions to AIDS in the United States: A second decade of stigma. *American Journal of Public Health, 83,* 574–577.

Kalichman, S. C. (1995). *Understanding AIDS: A guide for mental health professionals*. Washington, DC: American Psychological Association.

Karon, J. M., Rosenberg, P. S., McQuillan, G., Khare, M., Gwinn, M., & Petersen, L. R. (1996). Prevalence of HIV infection in the United States, 1984 to 1992. *Journal of the American Medical Association, 276*, 126–131.

Kelly, J., & Sykes, P. (1989). Helping the helpers: A support group for family members of persons with AIDS. *Social Work, 34,* 239–242.

Kennedy, C. A., Skurnick, J. H., Foley, M., & Louria, D. B. (1995). Gender differences in HIV-related psychological distress in heterosexual couples. *AIDS Care, 7* (Suppl. 1), S33–S37.

Klein, K., Forehand, R., Armistead, L., & Wierson, M. (1996). The contributions of social support and coping methods to stress resiliency in couples facing hemophilia and HIV. *Advancement in Behaviour Research and Therapy, 16,* 253–275.

Knox, M. D., Dow, M. G., & Cotton, D. A. (1989). Mental health care providers: The need for AIDS education. *AIDS Education and Prevention, 1*, 285–290.

Knox, M. D., & Gaies, J. S. (1990). The HIV clinical tutorial for community mental health professionals. *Community Mental Health Journal, 26*, 559–566.

Kubler-Ross, E. (1987). *AIDS: The ultimate challenge*. New York: Macmillan.

Leask, C., Elford, J., Bor, R., Miller, R., & Johnson, M. (1997). Selective disclosure: A pilot investigation into changes in family relationships since HIV diagnosis. *Journal of Family Therapy, 19*, 59–69.

Levine, C. (1995). Orphans of the HIV epidemic: Unmet needs in six US cities. *AIDS Care, 7*(Suppl. 1), S57–S62.

Lovejoy, N. C. (1989). AIDS: Impact on the gay man's homosexual and heterosexual families. *Marriage and Family Review, 14*, 285–316.

Lynch, R. D. (1989). Psychological impact of AIDS on individual, family, community, nation, and world in a historical perspective. *Family and Community Health, 12*, 52–59.

MacDonald, K. L., Jackson, J. B., Bowman, R. J., Polesy, H. F., Rhame, F. S., Balfour, H. H., Jr., & Osterholm, M. T. (1989). Performance characteristics of serologic tests for Human Immunodeficiency Virus type I (HIV-I) antibody among Minnesota blood donors. Public health and clinical implications. *Annals of Internal Medicine, 110*, 617–621.

Mann, J., Tarantola, D. J. M., & Netter, T. W. (1992). *A global report: AIDS in the world*. New York: Oxford University Press.

Mason, H. R. C., Marks, G., Simoni, J. M., Ruiz, M. S., & Richardson, J. L. (1995). Culturally sanctioned secrets? Latino men's nondisclosure of HIV infection to family, friends, and lovers. *Health Psychology, 14*, 6–12.

Milanese, G., Abeni, D., Cancelli, A., Brancato, G., Fabrizi, E., Borgia, P., Perucci, C. A., & Rocchi, G. (1997). Heterogeneity of home care assistance needs of people with AIDS. *AIDS Care, 9*, 27–33.

Millan F., & India Ivory, L. (1994). Group therapy with the multiply oppressed: Treating Latino, HIV-infected injecting drug users. *GROUP, 18*, 154–166.

Morin, S. F. (1991). Removing the stigma: Lesbian and gay affirmative counseling. *The Counseling Psychologist, 19*, 245–247.

Nichols, S. E. (1985). Psychosocial reactions of persons with the acquired immunodeficiency syndrome. *Annals of Internal Medicine, 103*, 765–767.

Nord, D. (1996). The impact of multiple AIDS-related loss on families of origin and families of choice. *The American Journal of Family Therapy, 24*, 129–144.

Osmond, D. H. (1994). Trends in HIV disease survival time. In P. T. Cohen, M. A. Sande, & P. A. Volberding (Eds.), *The AIDS knowledge base* (2nd ed., pp. 1.3.1–1.3.7). Boston: Little, Brown.

Pearlin, L. I., Mullan, J. T., Aneshensel, C. S., Wardlaw, L., & Harrington, C. (1994). The structure and functions of AIDS caregiving relationships. *Psychosocial Rehabilitation Journal, 17*, 51–67.

Pomeroy, E. C., Rubin, A., & Walker, R. J. (1995). Effectiveness of a psychoeducational and task-centered group intervention for family members of people with AIDS. *Social Work Research, 19*, 142–152.

Pomeroy, E. C., Rubin, A., & Walker, R. J. (1996). A psychoeducational group intervention for family members of persons with HIV/AIDS. *Family Process, 35*, 299–312.

Pryor, J. B., Reeder, G. D., Vinacco, R., & Kott, T. L. (1989). The instrumental and symbolic functions of attitudes towards persons with AIDS. *Journal of Applied Social Psychology, 19*, 377–404.

Radkowsky, M., & Siegel, L. J. (1997). The gay adolescent: Stressors, adaptations, and psychosocial interventions. *Clinical Psychology Review, 17*, 191–216.

Stowe, A., Ross, W., Wodak, A., Thomas, G. V., & Larson, S. A. (1993). Significant relationships and social supports of injecting drug users and their implications for HIV/AIDS services. *AIDS Care, 5*, 23–33.

Takigiku, S. K., Brubaker, T. H., & Hennon, C. B. (1993). A contextual model of stress among parent caregivers of gay sons with AIDS. *AIDS Education and Prevention, 5*(1), 25–42.

Tehan, C. (1991). The cost of caring for patients with HIV infection in hospice. *AIDS in the Hospice Community, 7*, 41–59.

Trezza, G. R. (1994). HIV knowledge and stigmatization of persons with AIDS: Implications for the development of HIV education for young adults. *Professional Psychology: Research and Practice, 25*, 141–148.

Tsiantis, J., Anastasopoulos, D., Meyer, M., Panitz, D., Ladis, V., Platokouki, N., Aroni, S., & Kattamis, C. (1990). A multi-level intervention approach for care of HIV-positive haemophiliac and thalassaemic patients and their families. *AIDS Care, 2*, 253–266.

Turner, H. A., & Catania, J. A. (1997). Informal caregiving to persons with AIDS in the United States: Caregiver burden among central cities residents eighteen to forty-nine years old. *American Journal of Community Psychology, 25*, 35–59.

USPHS/IDSA Prevention of Opportunistic Infections Working Group (1995). USPHS/IDSA guidelines for the prevention of opportunistic infections in persons infected with the Human Immunodeficiency Virus: Disease-specific recommendations. *Clinical Infectious Diseases, 21*(Suppl. 1), S32–S43.

Wadland, W. C., & Gleeson, C. J. (1991). A model for psychosocial issues in HIV disease. *Journal of Family Practice, 33*, 82–86.

Wardlaw, L. A. (1994). Sustaining informal caregivers for persons with AIDS. *Families in Society, 75*, 373–384.

Winiarski, M. G. (1997). *HIV mental health for the 21st century*. New York: University Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.12
Families of Individuals with Developmental Disabilities

OLIVER C. MUDFORD and GILL McGRANE
University of Keele, UK

9.12.1 INTRODUCTION 183

9.12.2 PSYCHOLOGICAL THERAPIES FOR FAMILY MEMBERS 185
- *9.12.2.1 Guilt and Self-blame* 185
- *9.12.2.2 Stress Management and Coping Skills* 185
- *9.12.2.3 Inclusion of Psychological Interventions in Family Support Programs* 186

9.12.3 BEHAVIORAL INTERVENTIONS WITH FAMILY MEMBERS AS THERAPISTS 186
- *9.12.3.1 Characteristics of Developmental Disorders with Mental Retardation* 186
- *9.12.3.2 Importance of Behavioral Deficits and Problems for Families* 186
- *9.12.3.3 Characteristics of the Behavior Analytic Approach* 186
- *9.12.3.4 Outcomes of the Behavioral Approach* 187
- *9.12.3.5 Characteristics of Families Who Have Benefited Most from Behavioral Training* 188
- *9.12.3.6 Basic Behavioral Skills Taught to Families: Target Behaviors for Families* 188
- *9.12.3.7 Which Family Members are Taught?* 188
- *9.12.3.8 Which Teaching Formats are Families Taught to Use?* 189
- *9.12.3.9 How are Basic Behavioral Skills Taught by Professionals to Families?* 190
- *9.12.3.10 Teaching Parents to Manage Problem Behaviors* 191
  - *9.12.3.10.1 Timeout as a standard component in parent training programs* 191
  - *9.12.3.10.2 The least intrusive treatments imperative* 192
  - *9.12.3.10.3 One model for evaluating the least intrusive treatment* 192
  - *9.12.3.10.4 An even less intrusive treatment?* 192
  - *9.12.3.10.5 Treatment recommendations should be based on functional assessment* 193
  - *9.12.3.10.6 Should errorless compliance training replace CMT in parent training?* 193
- *9.12.3.11 Families and Insufficiently Validated Treatment Recommendations* 194

9.12.4 CONCLUSIONS 194

9.12.5 REFERENCES 195

9.12.1 INTRODUCTION

It is generally recognized that the demands of caring for a child with disabilities can place families under stress which is likely to have an adverse effect on relationships within and outside the family. This stress exceeds that reported by parents of nondisabled children and may be chronic (Dyson, 1993; Flynt, Wood, & Scott, 1992). Chronic stress on families can result in their seeking out-of-home placement for their child with disabilities and also places children with developmental disabilities at risk of abuse and neglect (Lutzker, Campbell, Harrold, & Huynen, 1992). For example, some research suggests that such stressed families have higher rates of divorce and lower rates of remarriage (Singer & Irvin, 1990), and less marital satisfaction, psychological well-being, and social support than families of children without disabilities (Friedrich & Friedrich, 1981). Other problems which have been identified are social isolation related to the negative reactions of others, the unavailability of social support and respite care, and limited financial resources as a result of the cost of caregiving, and the restrictions that caregiving places on parents' ability to work. Families may experience considerable fatigue, strain, and distress as a result of their child's behavior problems or the level of physical care that the child may need (Tausig, 1985). The need to arrange family activities and schedules to accommodate the medical, physical, and emotional needs of a disabled child can place further restrictions on the ability of the family to function normally.

Research on family stress has tended to give a negative perception of family functioning by concentrating on dysfunctions (Singer & Irvin, 1990). However, professionals who interact and advise families may discover that there are additional problems beyond the original reason for referral of the parent or, more likely, the child.

Family responsibility for a member with disabilities can extend to a lifelong role involving the whole family (Singer, Irvin, Irvine, Hawkins, & Cooley, 1989; Smith, Tobin, & Fullmer, 1995). Due to decreased mortality, improved healthcare, and the changes from institutional to community-based services, older parents increasingly are having to be involved in lifetime planning, even when the family member with disabilities is not living at home (Griffiths & Unger, 1994). For example, a recent survey of 140 adult siblings having a brother or sister with disabilities found that 36% expressed an intention to remain involved to the extent that they would eventually live together (Krauss, Seltzer, Gordon, & Friedman, 1996).

The presence of a child with developmental disabilities in the family may affect individual family members differentially. For example, mothers and fathers of a child with disabilities can significantly differ on psychological variables, with mothers often reporting significantly more negative effects (Beckman, 1991; Kravetz, Nativitz, & Katz, 1993). Dyson (1993) reported that family relationships were the most consistent factor that influenced parental stress which was not found to be the case for families of children without disabilities. This suggests that primary caregivers of children with disabilities expect greater levels of familial support than those of children without disabilities which may reflect the absence of extended networks of informal social support that are associated with decreased stress (Beckman, 1991). A comparison of needs between mothers and fathers of children with disabilities showed mothers as expressing significantly more needs than fathers (Bailey, Blasco, & Simeonsson, 1992). In particular, a need to have more time to themselves, for help in explaining their child's disability to others, and help with child care. Furthermore, they expressed more interest in having opportunities to meet and interact with other parents (Bailey et al., 1992).

As with most of the research on family characteristics, the results with regard to siblings are conflicting (e.g., Lobato, Faust, & Spirito, 1988). The effects of a family member with disabilities on their siblings have been the subject of recent research. For example, on a variety of ratings of psychological adjustment by the siblings of 19 children with autism, 20 children with Down's syndrome, and 20 children without disabilities, only on the Child Behavior Checklist (completed by mothers) were significant differences found (Rodrigue, Geffken, & Morgan, 1993). However, none of the grouped results were outside the normal range. Rodrigue et al. (1993) noted that these results are contrary to commonly held beliefs. However, they cited other research supporting the view that siblings may well not experience poor psychological adjustment. These findings are consistent with those reviewed by Lobato et al. (1988) that included chronically ill as well as siblings with developmental disabilities.

The effect on and of other family members has also received research attention. For example, Hornby and Ashworth (1994) assessed the amount of help and emotional support given by grandparents to 25 parents who had children with severe disabilities. Although 54% of grandparents provided little or no support, only 22% of parents reported that they would have liked more support. This survey was conducted in England and, compared with North American studies, levels of grandparental support were relatively low. Recent American research found that support from grandparents was positively correlated with family adjustment (Sandler, Warren, & Raver, 1995).

The recognition of the highly complex interactions between child characteristics, intrafamilial relationships, social and financial resources, environmental demands, and life-cycle transitions, has encouraged researchers to take an ecological perspective. In this model, the family is viewed as a unit consisting of multiple interacting elements in contrast to focusing on one element, such as the target child or mother (e.g., Lutzker et al., 1992; Sanders, 1992; Singer & Irvin, 1990). Many of the elements of an ecological approach do not constitute psychological interventions but may have psychological sequelae. An example of this concerns the costs of caring for a family member who is disabled (Fujiura, Roccoforte, & Braddock, 1994). For example, Herman (1994) and

Herman and Thompson (1995) reported that provision of a monthly cash subsidy increased parents' access to services such as respite care, professional support services, and parent support services. These supports in turn might be expected to have an impact upon psychological functioning within the family.

9.12.2 PSYCHOLOGICAL THERAPIES FOR FAMILY MEMBERS

In this section, we discuss interventions that were specifically targeted to the psychological distress experienced by parents. These interventions were not designed to change the behaviors of the family member with developmental disabilities.

9.12.2.1 Guilt and Self-blame

Families of children with disabilities often suffer from excessive amounts of guilt and self-blame (Nixon & Singer, 1993). In the absence of a clear causal explanation for their child's disability, as is often the case in diagnoses of mental retardation, parents search for an explanation which may result in self-blame. We suspect that some of the causes of guilt may stem from psychiatric opinion in the 1950s that attributed the etiology of some developmental disorders to parental personality characteristics. While these views no longer have credence in modern psychiatry, there may be residual negative attitudes among the general public towards parents resulting from these earlier professional views.

Nixon and Singer (1993) evaluated the use of cognitive restructuring as a means of changing self-blaming attributions. Their study included parents of children having developmental disabilities in the moderate to profound range. Parents were randomly assigned to either a treatment or a control group. One two-hour treatment session of cognitive therapy per week took place over a period of five weeks, containing a lecture, group discussion, and small-group exercises. The first treatment session introduced a cognitive model of emotions based on the assertion that people can change their feelings by changing their thinking. The second session introduced strategies for changing cognitive distortions and automatic thoughts. Parents' awareness of cognitive distortions and automatic thinking was facilitated by daily monitoring of their thoughts, and the links between self-blame, guilt, and depression with cognitive distortions and automatic thoughts were demonstrated. "Should" statements are believed to be particularly guilt-producing and parents were taught should-removal strategies, while reattribution taught parents to accept some responsibility for negative events but not all the blame. Finally, faulty schemas were addressed by reinterpreting and restructuring them into realistic and attainable goals. Later sessions concerned issues such as control, misattributions concerning explanation of events, and parental schemas in order to provide parents with adaptive strategies that enabled them to maintain a sense of perceived control and reduce their need to self-blame. Results of the study showed that self-blame, guilt, internal negative attributions, automatic thoughts, and depression were significantly reduced.

9.12.2.2 Stress Management and Coping Skills

Stress management training was evaluated by Singer, Irvin, and Hawkins (1988) in a study with 36 parents of children with severe handicaps. They were randomly assigned to either a treatment group or a waiting-list control group. The treatment consisted of one, two-hour class per week for eight weeks. The classes took the form of lecture, demonstration, and discussion during which three techniques were taught: (i) self-monitoring of stressful events and physiological reactions to them; (ii) muscle relaxation skills; and (iii) modification of cognitions associated with distress. Self-monitoring involved teaching the parents to recognize their own symptoms of stress and to identify the events which triggered those symptoms. Homework assignments required parents to assess their tension levels three times a day and to note any specific event which seemed to precipitate increased tension. Physiological symptoms of stress, such as headaches, insomnia, and indigestion, were also noted. Various stages of muscle relaxation were taught, progressing from a long form (25–30 min) to a very short form that could be used at work and in domestic situations where it was not possible to avoid the stressful event.

To learn cognitive modification skills, parents were required to note their thoughts at times when they were feeling particularly tense. Parents were taught to use structured diaries to recognize thoughts which were exaggerated or distorted, and to think in more realistic terms. Homework assignments and class demonstrations provided opportunities for parents to practice each newly acquired skill and encouraged generalization to the home environment. The results of the study indicated that stress management training did prove to be effective.

9.12.2.3 Inclusion of Psychological Interventions in Family Support Programs

We do not know whether guilt reduction has been included in most parent support programs, however, stress management and coping skills training has been incorporated into a comprehensive set of support services, including provision of case management, respite care, and behavioral parenting skills training (Singer et al., 1989). Mothers who received stress management training in addition to other services were less depressed and anxious up to a year following the multicomponent intervention. Given that the majority of parents do not experience excessive stress, we would have to question the automatic inclusion of stress management training in programs for helping parents.

In the Lutzker et al. (1992) model for family support systems, parents who have particular need of cognitive or cognitive-behavioral interventions have the opportunity to receive such therapies. Included are stress reduction, including relaxation training and/or anger control management, and problem-solving skills training. These options are in addition to training in providing routine care, behavioral pediatrics, and behavioral parenting skills. This model may be more economical and less intrusive for parents because only the components considered necessary are provided. Similarly, in another comprehensive family intervention model, Sanders and Dadds (1993) described what can be viewed as a least intrusive and most economical approach to helping families whose children have problem behaviors. Psychological intervention for parents is provided only if necessary, and usually in the context of helping to increase the likelihood of success of a child management program.

9.12.3 BEHAVIORAL INTERVENTIONS WITH FAMILY MEMBERS AS THERAPISTS

9.12.3.1 Characteristics of Developmental Disorders with Mental Retardation

Children and adults with mental retardation have deficits in two or more of the following skill domains: communication, self-care, self-direction, social skills, domestic (home living), functional academics, work, leisure, health, safety, and use of community resources. These areas of deficit define mental retardation, provided the deficits occur with an IQ of less than 70, in the *DSM-IV* classification system (American Psychiatric Association, 1994) or an IQ of less than 75 as in the American Association on Mental Retardation definition (Luckasson et al., 1992). In addition, significant behavioral problems (e.g., self-injury, aggression to others, repetitive stereotypies, and property destruction) may be exhibited by people with mental retardation. Current estimates of the prevalence of such problems range between 20 and 40% (Einfeld & Aman, 1995).

9.12.3.2 Importance of Behavioral Deficits and Problems for Families

The significance of behavioral deficits and problems for families are highlighted by Bromley and Blacher (1991) in a study into the reasons for parents seeking out-of-home residential services for their children with severe mental retardation. The families of 63 children rated the degree of influence of 21 variables in their decision for out-of-home residential services. The three variables which were rated highest in being a strong or very strong influence on the placement decision were day-to-day stress (81%), level of functioning of their child (75%), and behavior problems (60%). It seems likely that significant deficits and problems contributed to the reported levels of daily stress in families. Tausig's (1985) study included 64 families that had sought residential placement for their young family member with disabilities. Problem behaviors were the strongest predictor for seeking placement, and they were more important predictors than the level of retardation. Therefore, families generally welcome interventions that change behaviors of their children with developmental disabilities, for the benefit of the children and their families.

9.12.3.3 Characteristics of the Behavior Analytic Approach

Interventions to change behaviors either increase the rate of appropriate behavior which already occurs, teach new desirable behaviors which do not occur, or reduce the frequency and/or intensity of problem behaviors. Interventions which successfully change behaviors by altering the naturally occurring environmental antecedents and consequences are the procedures of applied behavior analysis (ABA). This approach was formerly known as behavior modification, and was based on operant laboratory findings in the experimental analysis of behavior. There is more to ABA than behavior modification procedures because it is not merely a successful technology for changing behavior. Analysis of the maintaining variables in baseline and of the intervention (independent) variables responsible for behavior change in individuals is a defining characteristic of ABA. These analytic skills include accurate measurement of behaviors of the individual and of that individual's social and physical environment and single-subject experimental designs.

Regarding family members' involvement in changing behavior, Graziano and Diament (1992) highlight the differences between the traditional child psychotherapy approach and that of psychologists guided by the ABA paradigm. Whereas the traditional approach involved psychologist–child interactions almost exclusive of family involvement, the ABA approach is for the psychologist to teach family members to act as therapists, that is, to become the agents of change for the behaviors of their family member with developmental disabilities. A review of all parent training research published between 1975 and 1990 identified the behavioral approach as most common, outnumbering all other approaches combined by a ratio of 9:1 (Rogers Wiese, 1992).

The child behaviors targeted for change through family-mediated behavioral interventions have been numerous. Some common examples are provided. Skill deficits addressed include language, especially with autistic disorder (e.g., Laski, Charlop, & Schreibman, 1988), social skills (e.g., Strain & Danko, 1995), self-care (e.g., Kashima, Baker, & Landen, 1988), play (e.g., Moran & Whitman, 1991), and independent walking (Robbins & Dunlap, 1992). Behavioral problems targeted have included aggression (e.g., Niemeyer & Fox, 1990) and oppositional behaviors, that is, tantrums, screaming, and refusal (Sanders & Plant, 1989). A common difficulty reported by parents has been failure of their children to follow their instructions, which can be viewed as either a deficit in compliance or a problem with noncompliance. This has been one of the primary reasons for referral by families for behavior change services for children with and without developmental disorders.

The potential advantages of family-mediated behavior therapy include: behavior is changed in a person's natural environment (settings and people), leading to the likelihood of improved maintenance and generalization of behavior change; generalized use by families of behavior change procedures beyond the initial reason for learning them; increased satisfaction for family members from successful outcome through their own efforts; reduction in stress associated with the increase in skills and/or reduction of problem behaviors; and increased cost-effectiveness (Graziano & Diament; 1992).

9.12.3.4 Outcomes of the Behavioral Approach

Concerning the effectiveness of parent training in ABA procedures, two outcome measures are most pertinent: (i) change in parent behavior, that is, the extent to which parents learned to perform the procedures taught; and (ii) change in child behavior, that is, whether parent training affected the child's behavior. Comprehensive reviews of the outcomes of parent behavioral training have concluded that the training of family members to act as therapists/trainers for their children provides benefits to both more often than not (see Graziano & Diament, 1992), even when the parents trained have developmental disabilities themselves (see Feldman, 1994, for a review). On the other hand, reviews of the methodology in parent training research have suggested that caution should be exercised in the interpretation of, and extrapolation from, the majority of such research (e.g., Rogers Wiese, 1992). In general, when the children of parents-in-training have developmental disabilities, desired parent behavior change is reported more often than desired change in their children (Graziano & Diament, 1992). A recent example was provided by Paisey, Fox, Curran, Hooper, and Whitney (1991) in which the mother of an 11-year-old girl with autism was unsuccessful in her attempts to reduce aggression using differential reinforcement and required relaxation. The combination of procedures was observed to be applied correctly by the mother, but was significantly effective only when conducted by trained staff in the family home. That is, the mother learned to perform the procedures and was consistent in their application, but they did not work for her. A recent counterexample, reporting two successes at reducing aggressive behavior with parental implementation of differential reinforcement, was provided by Niemeyer and Fox (1990).

Considering that production of truly socially significant improvement in individuals with profound disabilities (Reid, Phillips, & Green, 1991) or autism (McEachin, Smith, & Lovaas, 1993) has not been demonstrated consistently by behavior analytic researchers, yet it is not surprising that parental efforts are not always totally successful even if they have been trained adequately to use state-of-the-art behavioral procedures. This suggests that a realistic, but not pessimistic, view of potential behavioral gains should be communicated to families based on a thorough assessment of the child and of the intervention research literature. Maurice (1996) stated that parents can value (i.e., be reinforced for their training) the attainment of even small gains in their disabled children's skills. That is, parents need not, and should not, be promised unrealistically that behavioral methods will cure their child to become involved in training.

Generalization of parents' teaching skills from trained to novel tasks sometimes occurs without the addition of specific programming for generalization by the trainer (e.g., Bruder, 1987; Moran & Whitman, 1991). More often, generalization to other behaviors, settings, and siblings has not been spontaneous (Sanders & James, 1983) and requires deliberate promotion by the trainer. For example, Cordisco, Strain, and Depew (1988) found that one of three trained parents of children with autism generalized their learned behavioral skills. The data on maintenance of skills by parents are also equivocal. For instance, Forehand et al. (1979) followed up seven parents one year after their training had been completed and found that five continued to employ all learned skills.

Potential disadvantages have been suggested when behavior change agents are family members rather than a professional behavior analyst. One is that family stress may be increased by demands of carrying out behavioral teaching procedures. This question was addressed by Baker, Landen, and Kashima (1991) with 39 of 49 families who had completed a training program that focused on teaching self-care skills to their children. Questionnaires measuring depression, stress, and family adaptation and cohesion were administered before training and one year after training. There were small but significant decreases in grouped scores for depressive symptoms and stress, and an increase in family adaptation score. At present, there are insufficient data to suggest that parent training increases the burden on parents with children who have developmental disabilities.

9.12.3.5 Characteristics of Families Who Have Benefited Most from Behavioral Training

Reviewers of parent behavioral training research have indicated that desired behavior change in both parents and children is more likely when families have particular advantages (Frankel & Simmons, 1992; Graziano & Diament, 1992; Griest & Forehand, 1982; Webster-Stratton & Hammond, 1990). Higher socioeconomic status, a wide range of social contact and support, absence of psychopathology (particularly depression and psychosis), membership of the ethnic majority, absence of negative life crises and environmental stress, and an intact family low on marital distress have been found to be predictors of success. Families without some of these advantages may drop out of training (Frankel & Simmons, 1992), or they may not access training that may be available, or they may complete training but not benefit.

9.12.3.6 Basic Behavioral Skills Taught to Families: Target Behaviors for Families

What particular behavioral skills are taught to families with children who have, or are at risk of, developmental disabilities? For increasing behavioral repertoires, basic behavioral procedures are taught. Providing clear verbal prompts, hierarchies of prompting, contingent reinforcement, and shaping continue to be common elements across parent training programs (e.g., Harrold, Lutzker, Campbell, & Touchette, 1992; Moran & Whitman, 1991; Robbins & Dunlap, 1992; Sanders & Dadds, 1993). Basic procedures for reducing problem behaviors commonly include timeout (e.g., Harrold et al., 1992; Sanders & Dadds, 1993), physical guidance (Robbins & Dunlap, 1992), extinction, physical restraint, and differential reinforcement (Moran & Whitman, 1991).

It should be noted that the basic procedures listed are examples of the technology of ABA. Generalizable procedures are taught, but the analytic skills component seems not to be included generally. The teaching and/or training of skills to families may be viewed as a cookbook approach. At least some parents who have been exposed to the analytic nature of ABA advocate that the analytic aspect should be emphasized (e.g., Maurice, 1996).

9.12.3.7 Which Family Members are Taught?

The majority of studies of behavioral training in intact families with children with developmental disorders have focused on mothers as primary carers (Singer & Irvin, 1990). Although some reports do not differentiate between parents, others can be examined to illustrate the relatively low involvement of fathers (e.g., Kashima et al., 1988; Schreibman, Kaneko, & Koegel, 1991). However, fathers have been taught successfully in generalized behavioral skills, and there is some evidence that fathers may learn skills incidentally through a modeling effect from a trained mother (Horton, 1984). Horton's review of fathers' involvement in parent training was mainly of studies in which children with problem behaviors did not have additional disabilities. Although there may be an assumption that fathers ought to (or even must) be included in training, this may not be necessary for successful outcome for the child. Horton suggested that a husband's support of his wife's skill acquisition may be a more important determinant of success than being directly trained himself. Also, children have been taught to teach their sibling with disabilities (e.g., Coe, Matson, Craigie, & Gossen, 1991).

9.12.3.8 Which Teaching Formats are Families Taught to Use?

The discrete trials format involves repeated presentation of learning trials in discrete teaching sessions (Robbins & Dunlap, 1992). This format has been the one most often used to teach parents until relatively recently (Petronko, Harris, & Kormann, 1994). Some dissatisfaction has been expressed by parents and professionals (e.g., Laski et al., 1988; Singer & Irvin, 1990) with the artificiality and demands of the discrete trials approach. In contrast, the incidental teaching format requires therapists to contrive or identify teaching opportunities that occur more or less naturally during the course of regular activities and enable the use of more natural reinforcers (Hart & Risley, 1975).

Parents may well prefer the more naturalistic format which has been incorporated into several comprehensive programs variously described as the natural language paradigm (e.g., Laski et al., 1988), pivotal response training (e.g., Pierce & Schreibman, 1995), milieu teaching (e.g., Hemmeter & Kaiser, 1994), and planned activities training (e.g., Harrold et al., 1992; Sanders & Dadds, 1993). In studies that have compared the two approaches, autistic children (Koegel, Koegel, & Surratt, 1992; Mirandalinne & Melin, 1992) and adults with severe/profound mental retardation (Elliot, Hall, & Soper, 1991) have gained language skills at least as well with the naturalistic format. Parents have been trained successfully to use the naturalistic methods (e.g., Alpert & Kaiser, 1992; Hemmeter & Kaiser, 1994; Koegel, Bimbella, & Schreibman, 1996; Schreibman et al., 1991). The more naturalistic approaches appear to have significant advantages. For example, both children and parents have been found to show more behavioral expressions of positive affect while training is conducted (Schreibman et al., 1991) and outside specific training times (Koegel et al., 1996); children exhibit less disruptive behavior (Koegel et al., 1992); and there is some evidence of greater generalization (Elliot et al., 1991; Mirandalinne & Melin, 1992). However, evidence of comprehensive and maintained gains comparable to those demonstrated for the more traditional approach (McEachin et al., 1993) has not been reported.

Although the majority of research in the 1990s with parents as trainers for their children with developmental disabilities has emphasized and found benefits in the more naturalistic approach, we cannot recommend that the teaching of the discrete trials format to families be discarded yet for three main reasons. First, some well-validated and replicable child teaching programs which include extensive family participation prescribe this format, at least for initial training (e.g., Lovaas, 1987; Lovaas & Buch, 1997; McEachin et al., 1993). Second, given that the basic behavioral procedures are common across teaching formats, parents may find it easier to learn these within a discrete trials format before their incorporation into regular activities. Third, acquisition of new skills by a child may be faster when massed trials are employed, which could be expected to provide early positive encouraging feedback to families on their successful teaching (Mirandalinne & Melin, 1992). However, it may be advisable to move from discrete trials to a more naturalistic format once the basic behavioral skills are demonstrated by parents and some child learning has been shown to occur.

A further potential disadvantage of training parents to teach their children using behavioral methods has been mooted. Family members may come to view themselves and behave primarily as therapists or trainers rather than maintain their special role as mother, father, sister, and brother (Singer & Irvin, 1990). It does not seem to us that these roles are necessarily mutually exclusive, although there may be some formats for parents to be therapists that capitalize on a combined role better than others. For example, Schreibman et al. (1991) showed five-minute videotape segments of 24 parents conducting behavioral training for increasing language skills in their children with autism. Viewers were 120 undergraduates who were naive with respect to the study. The students rated levels of parental affect, specifically enthusiasm, interest, and happiness exhibited by the parent while training. The students judged that positive affect on all these dimensions was exhibited by parents when a more naturalistic child training method (i.e., pivotal response training) was being used by the parents. Ratings of affect were significantly lower when parents were shown using a discrete trial training format.

In another study, Koegel et al. (1996) found that ratings of parental affect from videotapes of extratraining family interactions were positive if families had been taught to use the more naturalistic teaching paradigm. In contrast, families that had received training to use the discrete trial format were rated as showing negative affect in the same setting. Studies of this type can be seen as approaching the question about parental behavior while interacting with their child according to a behavioral training protocol. The results imply that naturalistic formats for training children appear more parent-like. Koegel et al. suggest that the more naturalistic the format, the more likely the program may reduce family stress in everyday life.

9.12.3.9 How are Basic Behavioral Skills Taught by Professionals to Families?

Methods that have been used, either alone or in various combinations, include parent training manuals (Bernal & North, 1978), group training (Hornby & Singh, 1983), and individualized training (e.g., Ducharme & Popynick, 1993; Robbins & Dunlap, 1992). All of these teaching methods have been shown to be effective in some dimensions with some parents in one or more studies. Written material alone may be most useful when validated manuals address specific deficits or problems, although some parents have learned how to teach self-help skills generalizably from manuals alone (Bernal & North, 1978).

There seems to be conflicting data on this particular issue. A study required 19 parents to read sections of the same publication as well as attend four workshops on those readings presented by experienced presenters (Kashima et al., 1988). Parents, as a group, did not increase their scores on an analogue teaching simulation test more than a control group who had received no training. The conclusion must be that some parents learn (i.e., change their teaching behaviors with respect to their children's deficits in self-care skills) from books/manuals, and some do not. The study also assessed the effects of teaching parents in groups using videotapes of experienced presenters' workshops. Group exercises and discussions facilitated by inexperienced group leaders were interspersed. Little difference was found between the two groups. Although knowledge of behavioral procedures increased for both groups, improvements in demonstrable application did not exceed gains in the control group. Nevertheless, we expect and encourage further studies of the use of videotape as a means to increase parents' knowledge and skills. If training tapes can be produced that are successful, both for parents and their children, this could become the presentation format of choice as ongoing professional involvement would be minimized.

Reviews of studies comparing of the effects of group and individual instruction to parents suggest (Hornby & Singh, 1983) or recommend (e.g., Baker, 1989; Graziano & Diament, 1992) that reported outcomes favor the group teaching format. The results are similar across the formats so recommendations are based on other considerations. Group teaching is more cost-effective, and the opportunity for social contact with parents in similar situations may be beneficial. Hornby and Singh (1983) caution that some parents may prefer not to be part of a group. These parents may not accept the opportunity to receive training if the group format is the only available option. Alternatively, they may not benefit through dropping out before successful completion of training, especially when they have depression or low socioeconomic status (Forehand, Middlebrook, Rogers, & Steffe, 1983).

A further important consideration is the ethnic and linguistic composition of groups for parent training. For example, Rueda and Martinez (1992) described a group education and training forum for parents of children with developmental disabilities in the US for whom Spanish was their first language. The authors mentioned that groups for ethnic Korean and Chinese parents have also been established. It is not clear whether the content of training included ABA methods, however, the availability of culturally sensitive presentation would be expected to increase the acceptability of group training to parents who may not otherwise access services (Rueda & Martinez, 1992).

The effective components for teaching parents to employ basic behavioral skills have been identified, whether parents are taught in groups or individually. The usual order in which these components are provided is: brief instructions in a skill; modeling of the skill by the trainer; behavioral rehearsal (i.e., role play with the parent demonstrating the skill and differential feedback on role play performance by trainer); and differential feedback by trainer on parents' performance of the skill with their children, either in a clinical setting or preferably in the families' natural environments. Thus, active participation by parents in training, rather than passive receipt of information, may be a necessary component required to obtain behavior change. Bearing this in mind, designers of future video-based training efforts may have to include effective prompting strategies to promote active learning.

An innovative approach to delivery of parent training was reported recently by Neef (1995), which replicated and extended earlier work by Bruder (1987). A pyramidal parent training method for teaching functional skills to children with autism was evaluated. First, the procedure involved professional training of five parents in the use of generalizable behavioral skill-training methods, that is, task analysis, defining target behaviors, selecting opportunities for teaching during regular household routines, selecting materials, presenting instructions to children, providing prompts, delivering consequences, and recording data. These parents were described as tier 1 in the pyramid. Second, when the tier 1 parents had achieved mastery, they taught the methods to another parent (tier 2). Third, four tier 2 parents subsequently taught the methods to one tier 3 parent each. Parents who were trained by other parents demonstrated acquisition of child-skill training methods to levels of mastery comparable with those shown by 12 parents who had been trained by a master's level professional. Although all parents' teaching skills increased, seven of 26 children did not show improvements on the skills targeted for increase. Future replications should attempt to assess the following: the extent to which trained teaching skills are used by parents in their homes; the acceptability of peer training, both by trainer/parents and trainee/parents; and, why some children do not apparently benefit.

9.12.3.10 Teaching Parents to Manage Problem Behaviors

So far we have discussed the often successful training of families to increase their use of potentially generalizable skills to reduce the behavioral deficits that define mental retardation and to reduce or possibly avert the development of some problem behaviors by the use of basic behavioral techniques. Methods for reducing deficits have not changed markedly in clinical or family settings over the last three decades, although there is more emphasis now in research and in parent training on adapting operant techniques so that antecedent stimuli and reinforcers become more naturalistic, as discussed earlier. The situation regarding reduction of problem behaviors is quite different.

There has been considerable controversy regarding the inclusion of punishment in programming for reduction of problem behaviors (Repp & Singh, 1990; Singh, 1997). There are strong beliefs in psychology concerning the inclusion in treatment recommendations of any practices that can be considered coercive (e.g., Sidman, 1989). Coercive therapeutic interventions can be viewed to include the use of punishment and negative reinforcement procedures and physical guidance that is resisted by the client.

9.12.3.10.1 Timeout as a standard component in parent training programs

Timeout continues to be taught routinely as one component in training programs for parents to enable them to reduce common problem behaviors (e.g., Lutzker, 1992; Sanders & Plant, 1989). All these studies used a version of child contingency management training (CMT), a component of a comprehensive behavioral family intervention system developed by Sanders and Dadds (1993). It should be emphasized that behavioral family intervention has other components in which parents are trained, for example, planned activities training (Sanders & Dadds, 1982). Although the procedures were developed and researched with children without developmental disabilities with behavior problems, these studies have demonstrated the generalizability of the parent training methods and content to children with disabilities. The CMT prescription for managing noncompliance or disruptive child behavior when it occurs during an activity is, briefly: (i) obtain child's attention; (ii) briefly explain what child was doing wrong; (iii) prompt appropriate behavior; (iv) praise correct behavior if it occurs; (v) if disruptive behavior continues, place child in timeout chair or area; (vi) after two minutes of quiet in timeout, remove child from timeout; (vii) return to activity; and (viii) praise correct behavior when it occurs.

What happens if the child resists the timeout procedure? Unfortunately procedures for dealing with this situation have not been fully described in the CMT research. However, Roberts and Powers (1990) describe studies with 36 preschool children who resisted chair timeout for continued noncompliance in a parent implemented program rather similar to CMT. Four backup procedures were assessed: (i) spank twice on the buttocks; (ii) hold, that is, physical restraint; (iii) barrier, that is, seclusion timeout; and (iv) release, that is, child was told that since he or she had decided to leave the chair he or she was obviously ready to behave appropriately. Only spank and barrier were effective with the majority of timeout-resistant children. When the procedures were used at home, 64% of mothers elected to use spank as a backup to timeout. Although the authors cautioned that careful training must be used to prevent misuse, that potential clearly exists. Even though a hierarchy of contingencies, whether it includes spanking or not, may be effective at increasing compliance, it represents escalating coercion, which is best avoided if behavioral research suggests the availability of alternative procedures.

9.12.3.10.2 The least intrusive treatments imperative

Parents have reported high levels of satisfaction with their training to use timeout (McMahon & Forehand, 1983). Some mothers of oppositional children clearly believe that being taught a particular way by a professional to spank their children is acceptable (Roberts & Powers, 1990). However, ABA practitioners and researchers generally subscribe to the view that it is an ethical imperative to evaluate the effects of the least intrusive but likely-effective methods for reducing problem behaviors (Van Houten et al., 1988). The question is to what extent should professionals' recommendations to parents regarding the treatment of their children and training provided to parents be coherent with their own professional practice?

9.12.3.10.3 One model for evaluating the least intrusive treatment

The methods demonstrated by Harding, Wacker, Cooper, Millard, and Jensen-Kovalan (1994) with seven children with problem behaviors may provide a model for parental and professional involvement in developing effective treatment. The purpose of the assessment was to determine the least intrusive intervention that was likely to be effective for referred problem behaviors including aggression, destructive behavior, noncompliance, and tantrums. During a 90-minute session at an outpatient clinic, parents were first instructed and then verbally prompted to provide each potentially effective treatment for five minutes in a hierarchy of intrusiveness starting from encouraging their children to choose an activity, a low-intrusion procedure that has been shown to reduce problem behavior in children with developmental disabilities (e.g., Dunlap et al., 1994). Other levels of treatment were added cumulatively until an effective set of components was obtained. These were differential reinforcement with attention for appropriate behavior, differential reinforcement of communication, and, finally, either guided compliance (manual guidance response prompt) or one-minute timeout (standing facing a blank wall). Only one child required guided compliance and none required timeout. Telephone follow-up on six families at six months indicated that parents were highly satisfied with the assessment procedures, had continued to use the treatments identified in the assessment, and they were satisfied with the changes in their children's behaviors. As the authors acknowledged, direct observational measurement of child and parent behavior in the home would have provided more valid data concerning outcome.

9.12.3.10.4 An even less intrusive treatment?

An example of taking a new approach that avoids coercive interactions to reduce a common problem behavior, with mothers as change agents, has been demonstrated by Ducharme and his colleagues (Ducharme et al., 1994; Ducharme & Popynick, 1993; Ducharme & Worling, 1994). Ducharme and Popynick assessed a procedure for errorless learning of compliance with four children with developmental disorders. First, parents rated the frequency with which their children complied with common requests on a four-point scale from "almost always complies" (level 1) to "rarely complies" (level 4). Second, the frequencies of compliance for 12 requests at each rated level were empirically evaluated in the families' homes. Following assessment of baseline compliance at each level, the third phase involved training mothers to provide level 1 requests only and to reinforce their child's compliance with praise and/or physical contact (e.g., hugs and kisses). Mothers were taught not to repeat the same request or to provide any consequences for noncompliance. When high ($> 80\%$) and stable levels of compliance had been achieved with level 1 requests, mothers were instructed to introduce level 2 requests gradually, for example, 14 level 1 requests interspersed with seven level 2 requests. In the following session, the proportions of requests at each level were reversed and, in subsequent sessions, even greater densities of the level 2 requests were included until high and stable compliance to these requests were achieved. Likewise, level 3 and level 4 requests were introduced. Problem behaviors reduced markedly when compliance increased. For example, level 4 requests which had been associated with tantrums and aggression on 51% of trials in baseline reduced to 10% following the intervention. Considering that these behaviors were not targeted directly, their reduction showed a valuable side effect of the mother-mediated compliance training.

Ducharme et al. (1994) replicated their previous study with four children with developmental disabilities and a nondisabled sibling of two of those, all of whom were referred for lack of compliance. The therapist's input with each mother was reduced from 34 hours in Ducharme and Popynick (1993) to 17 hours, on average, by abbreviating the initial empirical assessment procedure and assessment of intervention. There was no decline in the clinical significance of outcome with these refinements.

Both studies assessed consumer satisfaction with the intervention. The mean score was 6.3 out of a maximum possible of 7.0, indicating that mothers found the intervention highly satisfactory. The authors note that therapist involvement in errorless compliance training can be implemented with fewer (five or six) sessions either at a clinic or in the family home when a case is not intended to provide research data for publication.

9.12.3.10.5 Treatment recommendations should be based on functional assessment

Research on treatment of aberrant behavior that has been based on thorough analysis of the environmental function of individual's problem behavior(s) has increased the range of effective treatments that do not include a punishment or other coercive components (see Singh, 1997). Briefly, reinforcers for aberrant behavior can be classified on two dimensions. One dimension is positive vs. negative reinforcement, that is, occurrence or termination (respectively) of a stimulus following the behavior. The other dimension concerns whether the reinforcing stimulus is external or internal. Behaviors reinforced by internal stimuli are variously named as self-stimulatory behaviors or automatically reinforced behaviors. In the present context we will not consider internal reinforcers further. External positive reinforcers for individuals in some settings (e.g., persons with disabilities in family environments) are usually dependent on other persons for their provision, and these reinforcers can be classified as social (e.g., parental attention, praise, physical contact, etc.) or, at least, socially mediated (e.g., parent provides food or a preferred activity). Identification of the reinforcers for the aberrant behavior leads to treatment recommendations. If an aberrant behavior is positively reinforced, the individual may be taught a more appropriate behavior to access the reinforcer. If negatively reinforced, either the stimulus that terminates following the behavior is removed from the environment or the individual may be taught appropriate behavior to terminate it.

The treatments that follow from such individualized assessment of antecedents and reinforcers for problem behavior are themselves individualized. This represents a shift from treatment recommendations based on the form of the behavior, for example, if any child does behavior X, apply treatment component Y. It also increases the complexity of assessment and treatment design. The emphasis on the social and physical environment of the individual with problem behaviors may provoke reassessment of general parent training programs to the extent that they are intended to help parents reduce problem behaviors. In other words, parent trainers may view the analytic components of ABA as a necessary inclusion in their programs.

There are many considerations for professionals in ABA with respect to educating and/or teaching parents to manage problem behaviors with increasing complexity of professional "best practice." Should, for example, parents be taught to perform functional assessment of problem behaviors and reinforcer assessment? It might be that retrospective reports based on informal parental observation may render these detailed and time-consuming assessments unnecessary. That is, parents can tell professionals accurately what are their children's motivations for problem behaviors and their reinforcers. We are unaware of research into the validity of parental reports on these issues, however, studies have shown that caregiving staff opinions are insufficiently reliable for designing effective and least intrusive interventions. Lack of reliability of staff opinions has been found in the areas of function of problem behavior (e.g., Zarcone, Rodgers, Iwata, Rourke, & Dorsey, 1991) and reinforcers that can be used to increase alternative behaviors (e.g., Green et al., 1988).

It is probably unnecessary to teach parents about the whole process of functional assessment, although some parents may find this information useful (e.g., Maurice, 1996). It is unrealistic to expect them to undertake such analyses. However, the parents' role in such assessment has been expanded from providing historical information about problem behaviors to becoming the therapist during such analyses (e.g., Derby et al., 1992). Given that socially-mediated reinforcement is the controlling function for many behaviors, and that parents are likely to be the providers of this reinforcement when the child is at home, it appears vital to include parents in the role of a therapist. In the case of aggressive or noncompliant behaviors, social reinforcers are likely to account for a greater percentage of function, so parents' involvement in assessment of function is even more essential.

9.12.3.10.6 Should errorless compliance training replace CMT in parent training?

Given that Ducharme and Popynick (1993) have demonstrated an alternative treatment for parents to use to increase compliance that appears not to include any coercive elements, should we advocate for the termination of parent training programs that teach parents how to use timeout? Given the ethical imperative to use the least intrusive treatment, and that timeout and manual guidance are judged as more intrusive than reinforcement, should not errorless compliance training replace CMT as the method of choice for managing disruptive behavior by their children, whether or not the children have disabilities?

Despite the apparent success, acceptability, and efficiency of the errorless compliance training programs, its proponents do not recommend the immediate introduction of such innovative procedures into general parent training curricula. Further research is required. First, replication by other researchers is required to determine whether professionals not involved with the origination of the procedures can obtain similar results. Second, further generalizability across children with developmental disabilities, levels of severity of noncompliance, and other parents needs to be established beyond two children in Ducharme and Worling (1994), four in Ducharme et al. (1994), and four in Ducharme and Popynick (1993). Third, the procedures were evaluated in homes. Cross-setting generalizability needs to be established. Parents may find it extremely difficult to avoid higher level requests than currently programmed for training or maintenance in other settings where noncompliance can be even more problematic. In community settings, parents may have difficulty in following the program component requiring that they ignore noncompliance. Fourth, Ducharme and Worling (1994) acknowledge that further research is required to establish further the efficacy of the procedures for "don't" and "stop" requests. Fifth, maintenance of parent and child behavior beyond four months needs to be assessed. Sixth, comparison of the acceptability of CMT and similar programs with that of errorless compliance training needs to be assessed by parents of children who need such interventions.

9.12.3.11 Families and Insufficiently Validated Treatment Recommendations

Psychologists and other professionals should be well aware of an extensive history of strong advocacy for unvalidated methods for remediating deficits and/or reducing problems for people with developmental disorders and their families. Jacobson, Mulick, and Shwartz (1995) discuss how families can be particularly vulnerable to the false promises of insufficiently researched treatments for their children with disabilities. Examples they cite are patterning, gentle teaching, and especially facilitated communication. To these, we would add holding therapy (see Maurice, 1993, for a personal account of her experience with this treatment) and auditory integration training (e.g., Berkel, Malgeri, & Streit, 1996, for a review of current controversy). Silver (1995) has suggested some other controversial therapies that may be oversold to families, including those based on optometric visual training, cerebellar-vestibular dysfunction, applied kinesiology, tinted lenses, megavitamins, trace elements, and allergic sensitivities.

Space does not permit discussion of the individual therapies listed above. Ongoing research may clarify whether any of these treatments have benefit and whether families should be encouraged to learn or purchase these treatments. Until these or any other new procedures can be properly recommended by professionals, families may need advice to protect them from false hopes. One approach to protecting families from ill-founded claims for successful treatment has been to address parent groups on this very issue (Favell, 1996). Parents often welcome presentations on how to distinguish scientifically validated treatments from those for which the research had not been sufficient to justify widespread adoption.

9.12.4 CONCLUSIONS

Review of the effects on families of a member with developmental disabilities indicates that parents may have a multitude of problem, some psychological, but others of a social or economic nature. Research on these issues has produced equivocal results, although it has been sufficient to prompt a more ecological approach to providing services to families. That is, more comprehensive services are being developed that facilitate multiple interventions for families, where necessary. These interventions may be targeted at the person with disabilities, or at mothers or fathers or other family members. With respect to psychological interventions, we have reviewed parent behavioral training in more detail as this type of intervention has attracted most research, and it continues to be a cornerstone of service provision. Recent research in this area suggests that continuing efforts to make training by parents of their child with developmental disabilities more naturalistic, while still preserving methods derived from basic operant studies, will be welcomed by families. Early results from studies incorporating more natural antecedents and consequences for desired behaviors suggest that increased benefits for the child and family may result. Such training may reduce parental psychological distress, as well as increase childrens' skills.

9.12.5 REFERENCES

Alpert, C. L., & Kaiser, A. P. (1992). Training parents as milieu language teachers. *Journal of Early Intervention, 16,* 31–52.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Bailey, D. B., Blasco, P. M., & Simeonsson, R. J. (1992). Needs expressed by mothers and fathers of young children with disabilities. *American Journal on Mental Retardation, 97,* 1–10.

Baker, B. L. (1989). *Parent training and developmental disabilities.* Washington, DC: American Association on Mental Retardation.

Baker, B. L., Landen, S. J., & Kashima, K. J. (1991). Effects of parent training on families of children with mental retardation: Increased burden or generalized benefit? *American Journal on Mental Retardation, 96,* 127–136.

Beckman, P. J. (1991). Comparison of mothers' and fathers' perceptions of the effect of young children with and without disabilities. *American Journal on Mental Retardation, 95,* 585–595.

Berkel, D. E., Malgeri, S. E., & Streit, M. K. (1996). Auditory integration training for individuals with autism. *Education and Training in Mental Retardation and Developmental Disabilities, 31,* 66–70.

Bernal, M. E., & North, J. A. (1978). A survey of parent training manuals. *Journal of Applied Behavior Analysis, 11,* 533–544.

Bromley, B. E., & Blacher, J. (1991). Parental reasons for out-of-home placement of children with severe handicaps. *Mental Retardation, 29,* 275–280.

Bruder, M. B. (1987). Parent-to-parent training. *American Journal of Mental Deficiency, 91,* 435–438

Coe, D. A., Matson, J. L., Craigie, C. J., & Gossen, M. A. (1991). Play skills of autistic children: Assessment and instruction. *Child and Family Behavior Therapy, 13,* 13–40.

Cordisco, L. K., Strain, P. S., & Depew, N. (1988). Assessment for generalization of parenting skills in home settings. *Journal of the Association for Persons with Severe Handicaps, 13,* 202–212.

Derby, K. M., Wacker, D. P., Sasso, G., Steege, M., Northup, J., Cigrand, K., & Asmus, J. (1992). Brief functional assessment techniques to evaluate aberrant behavior in an outpatient setting: A summary of 79 cases. *Journal of Applied Behavior Analysis, 25,* 713–721.

Ducharme, J. M., Pontes, E., Guger, S., Crozier, K., Lucas, H., & Popynick, M. (1994). Errorless compliance to parental requests II: Increasing clinical practicality through abbreviation of treatment parameters. *Behavior Therapy, 25,* 467–487.

Ducharme, J. M., & Popynick, M. (1993). Errorless compliance to parental requests: Treatment effects and generalization. *Behavior Therapy, 24,* 209–226.

Ducharme, J. M., & Worling, D. E. (1994). Behavioral momentum and stimulus fading in the acquisition and maintenance of child compliance in the home. *Journal of Applied Behavior Analysis, 27,* 639–647.

Dunlap, G., DePerczel, M., Clarke, S., Wilson, D., Wright, S., White, R., & Gomez, A. (1994). Choice making to promote adaptive behavior for students with emotional and behavioral challenges. *Journal of Applied Behavior Analysis, 27,* 505–518.

Dyson, L. L. (1993). Response to the presence of a child with disabilities: Parental stress and family functioning over time. *American Journal on Mental Retardation, 98,* 207–218.

Einfeld, S. L., & Aman, M. (1995). Issues in the taxonomy of psychopathology in mental retardation. *Journal of Autism and Developmental Disorders, 25,* 143–167.

Elliot, R. O., Hall, K., & Soper, H. V. (1991). Analog language teaching versus natural-language teaching: Generalization and retention of language-learning for adults with autism and mental retardation. *Journal of Autism and Developmental Disorders, 21,* 433–447.

Favell, J. E. (1996, May). *Developmental disabilities and behavior analysis: Can this relationship be saved?* Paper presented at the 22nd Annual Convention of the Association for Behavior Analysis, San Francisco, CA.

Feldman, M. A. (1994). Parenting education for parents with intellectual disabilities: A review of outcome studies. *Research in Developmental Disabilities, 15,* 299–332.

Flynt, S. W., Wood, T. A., & Scott, R. L. (1992). Social support of mothers of children with mental retardation. *Mental Retardation, 30,* 233–236.

Forehand, R., Sturgis, E. T., McMahon, R. J., Aguar, D., Green, K., Wells, K. C., & Breiner, J. (1979). Parent behavioral training to modify child noncompliance: Treatment generalization across time and from home to school. *Behavior Modification, 3,* 3–26.

Forehand, R., Middlebrook, J., Rogers, T., & Steffe, M. (1983). Dropping out of parent training. *Behavior Research and Therapy, 6,* 663–668.

Frankel, F., & Simmons, J. Q. (1992). Parent behavioral training: Why and when some parents drop out. *Journal of Clinical Child Psychology, 21,* 322–330.

Friedrich, W. N., & Friedrich, W. L. (1981). Psychosocial assets of parents of handicapped and nonhandicapped children. *American Journal of Mental Deficiency, 85,* 551–553.

Fujiura, G. T., Roccoforte, J. A., & Braddock, D. (1994). Costs of family care for adults with mental retardation and related developmental disabilities. *American Journal on Mental Retardation, 99,* 250–261.

Graziano, A. M., & Diament, D. M. (1992). Parent behavioral training: An examination of the paradigm. *Behavior Modification, 16,* 3–38.

Green, C. W., Reid, D. H., White, L. K., Halford, R. C., Brittain, D. P., & Gardner, S. M. (1988). Identifying reinforcers for persons with profound handicaps: Staff opinion versus systematic assessment of preferences. *Journal of Applied Behavior Analysis, 21,* 31–43.

Griest, D. L., & Forehand, R. (1982). How can I get any parent training done with all these other problems going on?: The role of family variables in child behavior therapy. *Child and Family Behavior Therapy, 4,* 73–80.

Griffiths, D. L., & Unger, D. G. (1994). Views about planning for the future among parents and siblings of adults with mental retardation. *Family Relations, 43,* 221–227.

Harding, J., Wacker, D. P., Cooper, L. J., Millard, T., & Jensen-Kovalan, P. (1994). Brief hierarchical assessment of potential treatment components with children in an outpatient clinic. *Journal of Applied Behavior Analysis, 27,* 291–300.

Harrold, M., Lutzker, J. R., Campbell, R. V., & Touchette, P. E. (1992). Improving parent–child interactions for families of children with developmental disabilities. *Journal of Behaviour Therapy and Experimental Psychiatry, 23,* 89–100.

Hart, B., & Risley, T. R. (1975). Incidental teaching of language in the preschool. *Journal of Applied Behavior Analysis, 8,* 411–420.

Hemmeter, M. L., & Kaiser, A. P. (1994). Enhanced milieu teaching: Effects of parent-implemented language intervention. *Journal of Early Intervention, 18,* 269–289.

Herman, S. E. (1994). Cash subsidy program: Family satisfaction and need. *Mental Retardation, 32,* 416–421.

Herman, S. E., & Thompson, L. (1995). Families' perceptions of their resources for caring for children with developmental disabilities. *Mental Retardation, 33,* 73–83.

Hornby, G., & Ashworth, T. (1994). Grandparents' support for families who have children with disabilities. *Journal of Child and Family Studies, 3,* 403–412.

Hornby, G., & Singh, N. N. (1983). Group training for parents of mentally retarded children: A review and methodological analysis of behavioural studies. *Child: Care Health and Development, 9,* 199–213.

Horton, L. (1984). The father's role in behavioral parent training: A review. *Journal of Clinical Psychology, 13,* 274–279.

Jacobson, J. W., Mulick, J. A., & Schwartz, A. A. (1995). A history of facilitated communication: Science, pseudoscience, and antiscience. *American Psychologist, 50,* 750–765.

Kashima, K. J., Baker, B. L., & Landen, S. J. (1988). Media-based versus professionally led training for parents of mentally retarded children. *American Journal on Mental Retardation, 93,* 209–217.

Koegel, R. L., Bimbella, A., & Schreibman, L. (1996). Collateral effects of parent training on family interactions. *Journal of Autism and Developmental Disorders, 26,* 347–359.

Koegel, R. L., Koegel, L. K., & Surratt, A. (1992). Language intervention and disruptive behavior in preschool children with autism. *Journal of Autism and Developmental Disorders, 22,* 141–153.

Krauss, M. W., Seltzer, M. M., Gordon, R., & Friedman, D. H. (1996). Binding ties: The roles of adult siblings of persons with mental retardation. *Mental Retardation, 34,* 83–93.

Kravetz, S., Nativitz, R., & Katz, S. (1993). Parental coping styles and the school adjustment of children who are mentally retarded. *The British Journal of Developmental Disabilities, 39,* 51–59.

Laski, K. E., Charlop, M. H., & Schreibman, L. (1988). Training parents to use the natural language paradigm to increase their autistic children's speech. *Journal of Applied Behavior Analysis, 21,* 391–400.

Lobato, D., Faust, D., & Spirito, A. (1988). Examining the effects of chronic disease and disability on children's sibling relationships. *Journal of Pediatric Psychology, 13,* 389–407.

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55,* 3–9.

Lovaas, O.I., & Buch, G. (1997). Intensive behavioral intervention with young children with autism. In N. N. Singh (Ed.), *Prevention and treatment of severe behavior problems: Models and methods in developmental disabilities* (pp. 61–86). Pacific Grove, CA: Brooks/Cole.

Luckasson, R., Coulter, D. L., Polloway, E. A., Reiss, S., Schalock, R. L., Snell, M. E., Spitalnik, D. M., & Stark, J. A. (1992). *Mental retardation: Definition, classification, and systems of supports*. Washington, DC: American Association on Mental Retardation.

Lutzker, J. R. (1992). Developmental disabilities and child abuse and neglect: The ecobehavioural imperative. *Behaviour Change, 9,* 149–156.

Lutzker, J. R., Campbell, R. V., Harrold, M. E., & Huynen, K. B. (1992). Project Ecosystems: An ecobehavioral approach to families with children with developmental disabilities. *Journal of Developmental and Physical Disabilities, 4,* 1–14.

McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation, 97,* 359–372.

McMahon, R. J., & Forehand, R. (1983). Consumer evaluation in behavioral treatment of children: Types, issues, and recommendations. *Behavior Therapy, 14,* 209–225.

Maurice, C. (1993). *Let me hear your voice: A family's triumph over autism*. New York: Fawcett Columbine.

Maurice, C. (1996, May). *One parent's reflections on persuasion and partnership*. Paper presented at the 22nd Annual Convention of the Association for Behavior Analysis, San Francisco, CA.

Mirandalinne, F., & Melin, L. (1992). Acquisition, generalization, and spontaneous use of color adjectives: A comparison of incidental teaching and discrete-trial procedures for children with autism. *Research in Developmental Disabilities, 13,* 191–210.

Moran, D. R., & Whitman, T. L. (1991). Developing generalized teaching skills in mothers of autistic children. *Child and Family Behavior Therapy, 13,* 13–37.

Neef, N. A. (1995). Pyramidal parent training by peers. *Journal of Applied Behavior Analysis, 28,* 333–337.

Niemeyer, J. A., & Fox, J. (1990). Reducing aggressive behavior during car riding through parent-implemented DRO and fading procedures. *Education and Treatment of Children, 13,* 21–35.

Nixon, C. D., & Singer, G. H. S. (1993). Group cognitive-behavioral treatment for excessive parental self-blame and guilt. *American Journal on Mental Retardation, 97,* 665–672.

Paisey, T. J., Fox, S., Curran, C., Hooper, K., & Whitney, R. (1991). Case study: Reinforcement control of severe aggression exhibited by a child with autism in a family home. *Behavioral Residential Treatment, 6,* 289–302.

Petronko, M. R., Harris, S. L., & Kormann, R. J. (1994). Community-based behavioral training approaches for people with mental retardation and mental illness. *Journal of Consulting and Clinical Psychology, 62,* 49–54.

Pierce, K., & Schreibman, L. (1995). Increasing complex social behaviors in children with autism: Effects of peer-implemented pivotal response training. *Journal of Applied Behavior Analysis, 28,* 285–295.

Reid, D. H., Phillips, J. W., & Green, C. W. (1991). Teaching persons with profound and multiple handicaps: A review of the effects of behavioral research. *Journal of Applied Behavior Analysis, 24,* 319–336.

Repp, A. C., & Singh, N. N. (1990). *Perspectives on the use of nonaversive and aversive interventions for persons with developmental disabilities*. Sycamore, IL: Sycamore.

Robbins, F. R., & Dunlap, G. (1992). Effects of task difficulty on parent teaching skills and behavior problems of young children with autism. *American Journal on Mental Retardation, 96,* 631–643.

Roberts, M. W., & Powers, S. W. (1990). Adjusting chair timeout enforcement procedures for oppositional children. *Behavior Therapy, 21,* 257–271.

Rodrigue, J. R., Geffken, G. R., & Morgan, S. B. (1993). Perceived competence and behavioral adjustment of siblings of children with autism. *Journal of Autism and Developmental Disorders, 23,* 665–674.

Rogers Wiese, M. R. (1992). A critical view of parent training research. *Psychology in the Schools, 29,* 229–236.

Rueda, R., & Martinez, I. (1992). Fiesta educativa: One community's approach to parent training in developmental disabilities for Latino families. *Journal of the Association for Persons with Severe Handicaps, 17,* 95–103.

Sanders, M. R. (1992). Enhancing the impact of behavioural family intervention with children: Emerging perspectives. *Behaviour Change, 9,* 115–119.

Sanders, M. R., & Dadds, M. R. (1982). The effects of planned activity and child management procedures in parent training: An analysis of setting generality. *Behavior Therapy, 13,* 452–461.

Sanders, M. R., & Dadds, M. R. (1993). *Behavioral family intervention*. Needham Heights, MA: Allyn and Bacon.

Sanders, M. R., & James, J. E. (1983). The modification of parent behavior: A review of generalization and maintenance. *Behavior Modification, 7,* 3–27.

Sanders, M. R., & Plant, K. (1989). Programming for generalization to high and low risk parenting situations in families with oppositional developmentally disabled preschoolers. *Behavior Modification, 13,* 283–305.

Sandler, A. G., Warren, S. H., & Raver, S. A. (1995). Grandparents as a source of support for parents of children with disabilities: A brief report. *Mental Retardation, 33,* 248–250.

Schreibman, L., Kaneko, W. M., & Koegel, R. L. (1991). Positive affect of parents of autistic children: A comparison across two teaching techniques. *Behavior Therapy, 22,* 479–490.

Sidman, M. (1989). *Coercion and its fallout*. Boston: Authors Cooperative.

Silver, L. B. (1995). Controversial therapies. *Journal of Child Neurology, 10,* S96–S100.

Singer, G. H. S., & Irvin, L. K. (1990). Supporting families of persons with severe disabilities: Emerging findings, practices, and questions. In L. H. Meyer, C. A. Peck, & L. Brown (Eds.), *Critical issues in the lives of people with severe disabilities* (pp. 271–312). Baltimore: Paul H. Brookes.

Singer, G. H. S., Irvin, L. K., & Hawkins, N. (1988). Stress management training for parents of children with severe handicaps. *Mental Retardation, 26,* 269–277.

Singer, G. H. S., Irvin, L. K., Irvine, B., Hawkins, N. & Cooley, E. (1989). Evaluation of community-based support services for families of persons with developmental disabilities. *Journal of the Association for Persons with Severe Handicaps, 14,* 312–323.

Singh, N. N. (1997). *Prevention and treatment of severe behavior problems: Models and methods in developmental disabilities*. Pacific Grove, CA: Brooks/Cole.

Smith, G. C., Tobin, S. S., & Fullmer, E. M. (1995). Elderly mothers caring at home for offspring with mental retardation: A model of permanency planning. *American Journal on Mental Retardation, 99,* 487–499.

Strain, P. S., & Danko, C. D. (1995). Caregivers' encouragement of positive interaction between preschoolers with autism and their siblings. *Journal of Emotional and Behavioral Disorders, 3,* 2–12.

Tausig, M. (1985). Factors in family decision-making about placement for developmentally disabled individuals. *American Journal of Mental Deficiency, 89,* 352–361.

Van Houten, R., Axelrod, S., Bailey, J. S., Favell, J. E., Foxx, R. M., Iwata, B. A., & Lovaas, O. I. (1988). The right to effective behavioral treatment. *Journal of Applied Behavior Analysis, 21,* 381–384.

Webster-Stratton, C., & Hammond, M. (1990). Predictors of treatment outcome in parent training for families with conduct problem children. *Behavior Therapy, 21,* 319–337.

Zarcone, J. R., Rodgers, T. A., Iwata, B. A., Rourke, D., & Dorsey, M. F. (1991). Reliability of the Motivation Assessment Scale: A failure to replicate. *Research in Developmental Disabilities, 12,* 349–360.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.13
Families of Alcoholics

BARBARA S. McCRADY, ELIZABETH E. EPSTEIN, and
CHRISTOPHER W. KAHLER
Rutgers-The State University of New Jersey, Piscataway, NJ, USA

9.13.1 INTRODUCTION 199
9.13.2 PREVALENCE 200
9.13.2.1 Alcohol Use Disorders and Families with Alcohol Problems 200
9.13.2.2 Consequences of Alcoholism to the Individual and Family 200
9.13.2.3 Gender, Ethnic, and Cultural Factors 202
9.13.3 EARLY MODELS OF FAMILIES OF ALCOHOLICS 203
9.13.4 CURRENT MODELS: THEORY, ASSESSMENT, AND TREATMENT 204
9.13.4.1 Cognitive-behavioral Models 204
9.13.4.1.1 Cognitive-behavioral theory 204
9.13.4.1.2 Cognitive-behavioral assessment 205
9.13.4.1.3 Cognitive-behavioral treatment 206
9.13.4.2 Family Systems Models 210
9.13.4.2.1 Family systems theory 210
9.13.4.2.2 Family systems assessment 210
9.13.4.2.3 Family systems treatment 211
9.13.4.3 Family Disease Models 212
9.13.4.3.1 Family disease model theory 212
9.13.4.3.2 Family disease model assessment 212
9.13.4.3.3 Family disease model treatment 213
9.13.5 SUMMARY AND CONCLUSIONS 214
9.13.5.1 Common Elements of the Models 214
9.13.5.2 Strengths of Each Model 215
9.13.6 REFERENCES 215

9.13.1 INTRODUCTION

The effects of alcohol use on family and marital functioning are complex, and several models have evolved to describe inter-relationships between alcoholic drinking and family life. In addition, because alcoholism has profound effects on the family, and because intrafamilial relationships and family patterns affect the course, severity, and pattern of alcohol misuse, several models of family-focused treatments for the alcoholic and his or her family members are in use. In fact, it is unusual for a contemporary alcohol treatment facility to exclude family treatment from its program.

There are several advantages to a family-based approach to conceptualizing and treating substance-use disorders. Family involvement is associated with better compliance with treatment (e.g., Berger, 1981) and with better treatment outcome. In addition, substance use has a negative impact on the functioning of the rest of the family, and involving the family in treatment may ease their distress as well. Family treatment models provide a framework for conceptualizing the inter-relationships between substance use and family functioning, can reduce isolation of the alcoholic, can facilitate establishment of common family goals, can help the family identify and deal with other problems such as poor communication, and can be used as a guide for treatment with any part of the family that is available for treatment (Collins, 1990; McCrady & Epstein, 1996).

This chapter is designed to acquaint the reader with theoretical and treatment models related to families of alcoholics. The first section provides a brief review of prevalence of alcohol use disorders and of alcohol-related consequences to spouses and children of alcoholics. Gender, ethnic, and cultural factors are considered. Early models of functioning in families with an alcoholic member are then reviewed. These include psychodynamic models and sociological models of stress and coping. The primary focus of the rest of the chapter is on three family models of alcoholism: cognitive-behavioral family, family systems, and family disease models. Theory, assessment, and treatment techniques for each model are described. Treatment of the alcoholic using a family approach, as well as treatment of family members of alcoholics are discussed.

9.13.2 PREVALENCE

9.13.2.1 Alcohol Use Disorders and Families with Alcohol Problems

The *Diagnostic and statistical manual of mental disorders*—Fourth Edition (*DSM-IV*; American Psychiatric Association [APA], 1994) defines an individual as having alcohol dependence if he or she meets three of seven criteria within one year. Criteria relate to impaired control over alcohol consumption, unsuccessful attempts or a persistent desire to quit drinking, excessive time spent drinking or recovering from the effects of drinking, curtailment of occupational, social, or recreational functioning due to drinking, continued drinking despite known physical and psychological problems related to drinking, and physical tolerance and withdrawal. Alcohol abuse is diagnosed if drinking continues despite any one of four symptoms: recurrent failure to fulfill obligations, recurrent legal problems, recurrent social or occupational problems due to substance use, or repeated use of alcohol in physically hazardous situations.

In the USA, alcohol abuse and dependence are "among the most prevalent mental disorders in the general population" (APA, 1994, p. 203). Based on the Epidemiologic Catchment Area (ECA) study begun in 1980 using *DSM-III* (APA, 1980) criteria for alcohol abuse and dependence, the estimated lifetime prevalence rate of alcoholism in the USA is 13.8% and the one-year prevalence is about 6.8% (Bucholz, 1992). Data from the National Health Interview Survey (NHIS) 1988 Supplement put the 1-year prevalence of *DSM-III-R* (APA, 1987) defined alcohol abuse and dependence at 8.6%. An even larger number of people may have problems related to alcohol use but do not meet diagnostic criteria. The prevalence of alcohol abuse and dependence may be increasing in the overall population and may be increasing faster among females than males (National Institute on Alcohol Abuse and Alcoholism, 1993).

Given the large number of individuals with alcohol use disorders, the number of family members potentially affected by the drinking of a relative is considerable. Because alcoholism is most prevalent from ages 18–44, many individuals are actively alcoholic at the time in their lives when they are most likely be married and raising children (Bucholz, 1992). Although lifetime prevalence of alcoholism is lowest among those stably married, the one-year prevalence of alcohol abuse or dependence among married people is 4.2%. Based on NHIS data, it is estimated that 22.6% of people in the USA have a first-degree blood relative who is alcoholic (Harford, 1992).

9.13.2.2 Consequences of Alcoholism to the Individual and Family

The consequences of alcoholism and heavy alcohol use are well documented. Five percent of all deaths in the USA in 1988 were attributable to alcohol (NIAAA, 1993). Heavy alcohol use has been associated with a number of types of liver injury, various cardiac conditions, immune system depression, damage to endocrine system and reproductive function, and multiple adverse neurological effects. Alcoholism is also associated with a high incidence of serious and fatal injuries, high-risk behavior, and suicide. About 48% of people with alcohol abuse/dependence meet criteria for another psychiatric disorder, a rate two times greater than that of nonalcoholics (Helzer, Bucholz, & Robins, 1992).

The consequences of living with an alcoholic family member are also significant. The divorce rate of alcoholics is estimated to be about four times that of the general population, and family disruption is probably much more likely with alcoholism than with other mental disorders (Paolino, McCrady, & Diamond, 1978). Families with alcoholic parents experience more marital conflict and more parent–child conflict than control families (Reich, Earls, & Powell, 1988). Sixth-grade (11–12 year old) children of alcoholics have been found to experience more negative events within the family, more family conflict, less positive family events, and less family cohesion than sixth-grade children of nonalcoholic parents (Heavey & Dodd, 1995). Families of relapsed alcoholics have been found to have significantly less cohesion and expressiveness than control families and families in which the alcoholic has relapsed to heavy drinking have been found to have more conflict and less familial organization (Moos & Moos, 1984).

As well as being linked to divorce and family disruption, alcoholism has often been linked to family violence. Although problems with research design, methodology, and the sampling have not allowed a causal relationship between alcohol use and violence to be established (Leonard & Jacob, 1988; Martin, 1992; Sher, 1991), the literature on spouse abuse has found consistently that many abused wives view their husbands as alcoholic or problem drinkers, with estimates of alcoholism in the husband ranging from 35 to 93% (Leonard & Jacob, 1988). Studies have found alcohol present in 25–50% of all incidents of wife beating (Martin, 1992) and have found violence levels in couples with alcoholic husbands to be five to six times greater than national prevalence estimates (O'Farrell & Choquette, 1991).

Research suggests that a substantial number of child abusers are also excessive drinkers and that alcohol often is involved in the abuse when the abuser is alcoholic (Leonard & Jacob, 1988). However, the relationship between alcoholism and child abuse may not be as strong as with spouse abuse (Sher, 1991) as major portions of the variance in child abuse are not accounted for by the use of alcohol (Leonard & Jacob, 1988). Some studies have found that alcoholic or highly intoxicated parents are more likely to abuse their children than other parents and that alcoholism is four times greater in parents of abused children than in parents of sick children (Martin, 1992). However, other studies have found no differences in alcoholism rates between abusive and nonabusive parents. Rates of parental alcoholism in families with reported child abuse have ranged from 0–92% (Sher, 1991). Estimates of the incidence of incest and alcoholism also vary with two reviews of the literature finding alcohol problems in 20–50% (Miller, 1990) and 20–75% (Hurley, 1991) of incestuous fathers.

Research indicates that family members of alcoholics experience higher levels of psychological distress than individuals without alcoholic family members. Most empirical work has studied wives and children of alcoholic men. For instance, in an early study, Kogan and Jackson (1965) found that wives of alcoholics showed elevations on a number of MMPI clinical subscales when compared to community controls, but less than half of these women were in the disturbed range on MMPI scale. Moos, Finney, and Gamble (1982) reported that spouses of relapsed alcoholics were functioning more poorly than community controls on a number of dimensions including depression, anxiety, alcohol consumption, physical health, and family cohesion. Hill (1993) found that nonalcoholic wives of alcoholics did not differ from female community controls on any of the scales of the Minnesota Personality Questionnaire and differed on only the Psychopathic Deviate (PD) scale of the MMPI. In a study comparing wives of incestuous sexual offenders, wives of alcoholics, and wives whose spouses were neither alcoholics nor sexual offenders, wives of alcoholics were found to differ from controls on a few subscales of the MMPI that dealt with transient states and on the PD scale (Pitman & Taylor, 1992). This elevation was due to an elevation in self-alienation, which reflects problems in concentration, feeling that life is unrewarding, feelings of guilt, and possible alcohol abuse.

Children of alcoholics also appear to function more poorly when compared to children from nonalcoholic families, although elevations in distress reported remain on average below clinical levels (Heavey & Dodd, 1995; Johnson & Jacob, 1995). Empirical studies consistently have found an increased risk of alcohol problems in children of alcoholics, and alcohol use during pregnancy has been the only factor associated with Fetal Alcohol Syndrome, which is now "the leading known environmental cause of mental retardation in the Western World" (NIAAA, 1993, p. 221). Children of alcoholics have been found to show less social competency, more internalizing and externalizing behaviors, and more total problem behaviors than children of control families (Johnson & Jacob, 1995). Adolescent children of alcoholics also exhibit more internalizing and externalizing behavior and are at increased risk for using alcohol (Chassin, Rogosch, & Barrera, 1991). Among college freshmen, children of alcoholics manifest more alcohol and drug problems, more behavioral under control and negative affectivity, and more psychiatric distress relative to peers without alcoholic parents, although differences found were small to moderate (Sher, Walitzer, Wood, & Brent, 1991).

Although alcoholism has been associated with numerous severe consequences such as domestic violence and divorce, it appears that many families with alcoholic members are financially stable, avoid violence, and experience no higher levels of depression or anxiety than the general population (Steinglass, Bennett, Wolin, & Reiss, 1987). However, even if more dramatic consequences are avoided, alcoholism in a family member can still have detrimental effects. For instance, a review of the literature suggests that not only are alcoholics susceptible to a wide range of traumas and diseases, but their families also utilize healthcare services at rates higher than the population as a whole (Langenbucher, McCrady, Brick, & Esterly, 1994). One study (Holder & Hallan, 1986) reports that members of alcoholic families used more than four times the amount of healthcare resources used by matched nonalcoholic families.

9.13.2.3 Gender, Ethnic, and Cultural Factors

In studying and treating alcoholics and their families, it is important to recognize that there may be systematic differences in the expression of alcoholism and its effects on the family according to the ethnicity or race of the family and the gender of the alcoholic involved. Such differences are not well understood, as research on ethnic minority populations often has used biased sampling procedures, failed to recognize heterogeneity within minority groups, and has been flawed by inconsistencies in the definitions of constructs used across studies (Collins, 1992). Alcoholism in women and their families also has been understudied (Nathan & Skinstad, 1987).

In the ECA study, no significant differences were found in lifetime alcoholism prevalence rates of African-Americans, Caucasian, and Hispanics in the USA (Helzer, Burnam, & McEvoy, 1991). Native Americans appear to have substantially higher rates of alcoholism and heavy drinking than other groups in the USA, although drinking varies greatly from tribe to tribe (Lex, 1987; Manson, Shore, Baron, Ackerson, & Neligh, 1992). Among adolescents, Native American youth also have been found to have particularly high rates of alcohol use and alcohol-related problems (Welte & Barnes, 1987). The expression of alcoholic symptoms across cultures and countries is remarkably similar (Helzer & Canino, 1992). However, some ethnic or racial groups are more susceptible to certain consequences of heavy alcohol use (Herd, 1994; Lex, 1987; Welte & Barnes, 1987).

In both ECA and NHIS studies, the prevalence rates of alcohol abuse and dependence were at least twice as high among men as compared to women (Bucholz, 1992). Men have been found to drink more and to manifest more alcoholism than women in almost all cultures studied (Helzer & Canino, 1992). However, this difference is particularly great in Hispanic and Asian cultures as compared to Western and Anglo-Saxon cultures. Among adolescents, males report drinking more than females in all ethnic groups, but this difference is largest among Asians and smallest among Caucasians (Welte & Barnes, 1987). One of the few studies conducted that included families of female alcoholics found no differences in family functioning between families with alcoholic husbands and those with alcoholic wives (Moos & Moos, 1984). However, the amount of drinking by women in a given subculture could affect how a family reacts to alcoholism in a female family member and either heighten or lessen the stigma associated with being alcoholic. For instance, it has been reported that alcoholic African-American women have their drinking criticized by their husbands less often than their Caucasian counterparts (Lex, 1987).

Data from the NHIS study reveal differences by ethnic group and gender in the rates of having at least one first, second, or third degree blood relative who was alcoholic (Harford, 1992). The lowest rate of familial alcoholism is in Asian-Americans, around 14%, and the highest rate is in Native Americans, over 50%. Females more often reported having alcoholic family members than males, suggesting possible differences by gender in conceptions of or concern with alcoholism.

Cultural differences also exist in the conceptualization of alcoholism (Babor, 1992), which could influence how families react to and interpret alcohol problems in family members. Studies in other countries reveal variations in how families respond to alcoholism and what concerns they express about the consequences of alcohol use. For example, 76% of alcoholics in Puerto Rico reported that their families had objected to their drinking (Canino, Burnam, & Caetano, 1992), whereas only 36% of alcoholics in the former West Germany reported familial objections (Wittchen & Bronisch, 1992). A study of Indo-Pakistani, Latin American, and Chinese groups in British Columbia, Canada, found that subjects from the Indo-Pakistani subsample more often disapproved of alcohol use by women and placed greater emphasis on a full range of family related consequences of heavy drinking than the other groups; the Latin American subsample expressed the least concern about wife and child abuse related to excessive alcohol use (Legge & Sherlock, 1991). Finally, differences have been found between African-American, Caucasian, and Asian-American families in family management practices, involvement in family activities, sibling deviance, family structure, and parental attitudes toward children's drinking (Catalano et al., 1992) and evidence has suggested that "there may well be cultural differences in family processes leading to early variety of substance initiation" (p. 216).

Alcohol use problems are widespread and affect a large number of families. Although gender, ethnicity, and culture generally are assumed to play a role in alcohol use and family functioning, research on the relationships among these factors is in its infancy. The models described below outline broad principles about the relationships between alcohol use and family and individual functioning that are hypothesized to operate in most families. None of these models explicitly addresses the role of gender, ethnicity, or culture, but clinicians and researchers should be aware that not all families subscribe to the same set of values and beliefs and that ways of coping with alcohol problems that seem unusual or intractable may be influenced by factors unaccounted for by present models.

9.13.3 EARLY MODELS OF FAMILIES OF ALCOHOLICS

In the 1930s, many alcoholics received treatment in state mental hospitals. Social workers in those facilities who interviewed the wives of alcoholic men observed the significant distress, particularly anxiety, depression, and psychosomatic symptoms that these women experienced (e.g., Lewis, 1937). Theoretical models were developed to explain these observations. The earliest model, the Disturbed Personality Hypothesis, was derived from the psychodynamic models that predominated at the time, and postulated that wives of alcoholics were disturbed women who resolved neurotic conflicts through their marriage to an alcoholic man. Although some authors postulated one primary underlying conflict, Whalen (1953) hypothesized four different types of conflicts that could be resolved through marriage to an alcoholic, including conflicts with aggression, control, masochism, and ambivalence.

A corollary to the Disturbed Personality Hypothesis was the Decompensation Hypothesis. Central to psychodynamic models is the notion that neurotic conflicts serve as a defense against more basic or primitive conflicts. If defenses are removed, an individual would be expected to exhibit these more primitive conflicts and decompensate. Thus, it was hypothesized that if a married alcoholic successfully stopped drinking, his wife would decompensate and exhibit more severe psychopathology. MacDonald (1956) reported, among 18 women hospitalized in a state mental hospital who were married to alcoholics, that 11 of the women's husbands had decreased their drinking recently, suggesting that these observations supported the Decompensation Hypothesis. No other study, however, has supported this hypothesis.

In the 1950s, alternative models began to be developed. Jackson (1954), after interviewing women who were attending Al-Anon meetings, suggested that a stress and coping perspective would better explain the experiences of these women than the prevailing psychodynamic models. She suggested that living with an alcoholic is stressful, and that most of the symptoms that wives of alcoholics experienced were common to family members living with long-term, familial stressors. She also suggested that families progress through "stages" in coping with alcoholism, and that each stage is characterized by different psychological phenomena. Family stages included family denial of the problem in the earliest stage, followed by attempts to control the problem, feeling hopeless and chaotic, attempting to maintain stable family functioning with the alcoholic present, attempting to escape from the problem through marital separation, organizing and maintaining the family without the alcoholic present, and a final adjustment phase if the alcoholic stopped drinking. Cross-sectional and prospective tests supported Jackson's model (Kogan & Jackson, 1965; Moos et al., 1982), finding that women whose husbands were actively drinking showed substantial psychological distress, but that wives of men who were not drinking were indistinguishable from wives of nonalcoholic men.

Concern about children of alcoholics dates to a compassionate book that reported clinical interviews and observations of children of alcoholics, and underscored the difficulties they experienced growing up in an alcoholic household (Cork, 1969). Concern was also stimulated by early genetic research suggesting a strong heritable component to more severe forms of alcohol dependence (Goodwin, Schulsinger, Hermansen, Guze, & Winokur, 1973). Popular clinical models for understanding and treating children of alcoholics evolved, and scientific research has progressed as well. Stress and coping models in particular have been supported in studies of children of alcoholics, with results showing that adolescent children with an alcoholic parent who has had no drinking problems for a period of three years function better than children of nonrecovered alcoholics but not as well as children with no alcoholism in the family (Chassin et al., 1991).

Beginning in the 1970s, three contemporary models for conceptualizing alcohol and family functioning began to evolve. These models, based in social learning theory, general systems theory, and the disease model of alcoholism, predominate theory, research, and clinical models of alcoholism and family functioning in the 1990s. All three models provide frameworks for understanding the relationships between drinking and family functioning, and also provide frameworks for understanding the functioning of individual family members.

9.13.4 CURRENT MODELS: THEORY, ASSESSMENT, AND TREATMENT

9.13.4.1 Cognitive-behavioral Models

9.13.4.1.1 Cognitive-behavioral theory

Cognitive-behavioral models have been used to describe a range of human problems. These models use social learning theory to explain the relationships among problem behaviors, environmental events, and private events (cognitions and emotions). Cognitive-behavioral models analyze controlling events for each problem behavior, and assume that controlling events will differ across individuals, rather than assuming the same controlling variables for all individuals with the same problem. As a consequence, cognitive-behavioral models of alcohol problems and family functioning examine classes of variables hypothesized to control the functioning of each family member as well as the interactions among family members. In this section, the cognitive-behavioral models developed to understand the interactions between drinking and family functioning are described and then the cognitive-behavioral applications to conceptualizing the problems of spouses and of children of alcoholics are detailed. Subsequent sections describe approaches to the assessment and treatment of the interactions between alcohol and family functioning, and then describe assessment and treatment of spouses and children.

(i) Drinking and family functioning

Cognitive-behavioral models in use in the 1990s describe current factors maintaining drinking, and include detailed consideration of the role of the family in the maintenance of drinking and the process of change. These models emphasize the drinker and spouse's roles as agents of change, and the paradigm has empirical support for its efficacy in treating alcoholic couples (McCrady & Epstein, 1995; McCrady, Stout, Noel, Abrams, & Nelson, 1991). Several basic assumptions are central: (i) reciprocal interactions between the drinker and his or her environment determine repetitive, stressful intra- and interpersonal behavioral patterns; (ii) external antecedents to drinking have a lawful relationship to subsequent alcohol consumption, through repeated associations with reinforcement or anticipation of reinforcement; (iii) cognitive and affective events mediate the relationship between external antecedents and drinking; and (iv) drinking is maintained by physiological, psychological, and/or interpersonal consequences (McCrady & Epstein, 1996).

Cognitive-behavioral models assume that substance use occurs in response to antecedent stimuli or cues. These stimuli or cues precede the substance use, and their presence increases the probability that substance use will occur. External antecedents may relate to the individual, familial, or other social systems. Familial antecedents might include family members" attempts to control the alcoholic's drinking, or negative interactions around specific problems resulting from the drinking, such as problems in communication, sexual functioning, or finances.

Aspects of the environment become cues for substance use through repeated sequences of events that include the antecedent stimulus, substance use, the occurrence of stimuli that reinforce use, and the lack of occurrence of punishing stimuli. The spouse and other family members unwittingly may supply positive consequences of drinking, for instance, by taking care of the intoxicated family member, pampering him or her when sick from drinking too much, or may protect the drinker from naturally occurring negative consequences, for instance, by calling in sick when the drinker cannot go to work. For certain couples, alcohol also may be associated with an increase in positive interactions.

The relationship between an antecedent stimulus and substance use is believed to be mediated by organismic events, including thoughts, feelings, and physiological reactions. Substance abusers may have a variety of positive expectancies about the reinforcing consequences of substance use, or may use in response to negative affect or physiological cues such as decreasing blood alcohol level. Substance abusers may also have cognitive and affective reactions to interpersonal situations such as an argument with a spouse, and may drink in response.

A number of cognitive constructs are important to conceptualizing couples" relationships. Couples make certain attributions about the motivations underlying each other's behavior, assuming that they know the reasons for certain actions. They also have certain expectancies about each other's future behavior, based partly in past experience and partly in their distorted attributions about the partner's motivations.

The repertoire of behavioral responses available to the substance user and to the couple or family is important. Alcoholic families often have skills deficits. Couples may have difficulty in coping with alcohol-related situations, expressing affect, disagreeing, making requests for change, listening to and understanding the partner's communication, providing positive support, or solving problems productively as a couple. Spouses may lack coping skills to respond effectively to the drinker and to drinking situations, and may have difficulty balancing attention to their own needs, the responsibilities they may have to maintaining the integrity and functioning of their families, and the stresses presented by the alcoholic's behavior.

The effect of alcohol on family interactions is complex, and negative consequences are often quite prominent. Family members may avoid or criticize the drinker. Verbal or physical aggression may be prominent. Typically, these negative consequences exacerbate the problems; the spouse and children may become fearful and angry; the drinker may leave the house and drink elsewhere, or may learn to drink secretly. In either case, a pattern of lying on the alcoholic's part, and hypervigilance and mistrust by the family may result, severely damaging family functioning. Finally, since alcoholics with a moderate to severe problem spend an inordinate amount of time seeking, consuming, or recovering from the effects of alcohol, they are typically unable to contribute fully to ensure daily functioning of the household. The nonalcoholic spouse and children must then take on more responsibilities around the house, leading to increased resentment and stress. All of this may serve to deepen the unhappiness in the home, cue the alcoholic to spend more time out of the home and drinking, and create further stress on the rest of the family.

(ii) Functioning of family members

Cognitive-behavioral models of the functioning of spouses and children in alcoholic families utilize a stress and coping perspective. Multiple factors are hypothesized to impact on the functioning of family members, including their own coping repertoire, other psychological problems, the types of stressors in their environment created by the drinking, and the quality of the marital or parent-child relationship (McCrady, Kahler, Shraedley, Epstein, & Hirsch, 1995). Family members utilize a variety of maladaptive coping strategies to deal with the chronic stressor of living with an actively drinking alcoholic family member. Spouses may engage in a variety of ineffective behaviors intended to change the drinking, such as nagging the drinker to change, or attempting to control the drinking or the drinker's behavior. These behaviors are conceived as maladaptive coping strategies, rather than indicators of underlying psychopathology. Over time, spouses assume extra role responsibilities, and often decrease the time they devote to activities that they experience as pleasurable. Depression, anxiety, and social isolation are understandable consequences.

Children in these families may be exposed to inconsistent parenting, high levels of conflict, and low levels of support from the alcoholic parent. Children may lack coping strategies to deal with common developmental stressors and may have learned to use alcohol as a primary coping strategy. They also may lack other social supports within or outside of the family that could buffer the negative effects of the stressors within the alcoholic family (e.g., Begun & Zweben, 1990).

9.13.4.1.2 Cognitive-behavioral assessment

The focus of a cognitive-behavioral assessment varies with the presenting problems and with who presents for treatment. Ideally, the clinician would see the entire family when alcohol problems are present, and the clinician would assess the functioning of each individual family member, the interrelationships between drinking and family functioning, and the quality of the marital and family relationships. However, if individual family members present without the drinker, the assessment addresses their individual functioning as well as how they are coping with the drinking.

The goals of a comprehensive cognitive-behavioral assessment are to assess the interrelationships between drinking and family behaviors, evaluate the current functioning of each member of the family unit, including strengths, problem areas, and coping skills, and assess the functioning of the family as a unit. The results of the assessment are used to develop a specific treatment plan to impact on the individual's drinking, enhance positive coping for members of the family, and to enhance the quality of marital or family relationships.

Assessment of the spouse focuses on behaviors related to coping with the partner's alcohol use as well as a broader assessment of functioning. A number of measures assess different aspects of the spouse's coping with drinking:

(i) *The Spouse Enabling Inventory*. This assesses spouse behaviors that may facilitate or maintain the partner's drinking (Thomas, Yosioka, & Ager, 1994a). Behaviors are divided into two types: those that increase events that facilitate the partner's drinking (e.g., buying liquor), and those that decrease events that could inhibit the partner's drinking (e.g., lying about the drinking to others).

(ii) *The Spouse Sobriety Influence Inventory*. This assesses partner behaviors intended to control the drinking, and behaviors intended to support the partner in efforts not to drink (Thomas, Yosioka, & Ager, 1994b).

(iii) *The Spouse Behavior Questionnaire*. This assesses spouse behaviors that may serve as cues for drinking, reinforce drinking, punish drinking, support sobriety, or that represent ways for the spouse to take care of his or her own physical and psychological needs (McCrady, Miller, Epstein, & Van Horn, 1993).

(iv) *The significant other behavior questionnaire*. This is an adaptation of the Spouse Behavior Questionnaire that assesses coping of any significant other, not just the spouse (Love, Longabaugh, Clifford, Beattie, & Peaslee, 1993).

(v) *The Spouse Situation Inventory*. This is a set of role play stimuli with scoring guidelines that allows the clinician to assess the actual behaviors that spouses use to cope with alcohol-related situations (Rychtarik, 1990).

Although there are a number of very good measures to assess the ways that spouses cope with the drinking itself, a broader assessment framework is necessary. A comprehensive assessment of the spouse of an alcoholic should evaluate: level of distress; behavioral repertoire for coping with alcohol related situations; social supports; social and economic resources and problems; and individual strengths and sources of positive reinforcement.

Standardized behavioral assessments specifically related to the effects of parental alcoholism on children are not in widespread use. The literature suggests that specific domains to target for assessment in children of alcoholics would include mood (i.e., depressive) disorders, hyperactivity, conduct disorder, and early (i.e., young adolescence) development of heavy alcohol and drug use. In addition, assessment of coping strategies is indicated. Assessment must also take into account several complexities, including comorbid psychiatric problems of the parent, for example, depression or antisocial personality, socioeconomic status of the family, genetic loading for substance-use disorders, psychosocial functioning of the parents, the marital relationship of the parents, degree of conflict or cohesion in the family, whether one or both parents have an alcohol problem, and the quality of the parent–child relationship (see Sher, 1991, for a complete review).

The preceding discussion of assessment of family members focuses on their individual behavior. It is also important to assess the degree to which there is a contingent relationship between family members" actions and the individual's drinking. Several strategies facilitate this assessment: (i) completing a functional analysis of specific drinking episodes (e.g., McCrady, 1993) to identify antecedent cues for the drinking, the drinker's cognitive and affective responses to family members" actions, and family behaviors consequent to drinking that might either reinforce or punish the drinking; (ii) using behavioral self-recording of drinking episodes and drinking urges, in conjunction with a functional analysis of these episodes; (iii) using drinker self-report questionnaires that identify high risk situations for drinking, such as the Inventory of Drinking Situations (Annis, 1982).

Assessment of marital and family functioning uses a wide range of assessment approaches from the broader behavioral marital and family field. For example, marital satisfaction and marital problems may be assessed with questionnaires such as the Dyadic Adjustment Scale (Spanier, 1976), or the Areas of Change Questionnaire (Margolin, Talovic, & Weinstein, 1983). Communication and problem solving skills may be assessed using self-report measures such as the Primary Communication Inventory (Navran, 1967), or by structured observations of interactions coded using any of a variety of standardized coding systems such as the Specific Affect Coding System (Gottman, 1996), or the Rapid Couples Interaction Scoring System (Gottman, 1996).

Domestic violence should also be assessed, using such instruments as the Conflict Tactics Scale (Strauss, 1979). Family communication and problem solving, role responsibilities, and parenting skills can be assessed using measures from the broader cognitive-behavioral family therapy literature.

9.13.4.1.3 Cognitive-behavioral treatment

At the completion of the assessment the clinician develops a treatment plan tailored to the individual case. Preferably treatment would begin with the couple. Drinking, spouse coping, and relationship functioning would be addressed. The functioning of the children would be evaluated, and appropriate treatment would be provided if necessary. If the drinker is not willing to come into treatment, the spouse would be engaged first, and treatment would address the spouse's individual coping and would also be directed at getting the drinker into treatment. Again, the functioning of the children would be evaluated.

(i) Cognitive-behavioral treatment of couples

The treatment model described here follows logically from cognitive-behavioral theory (McCrady & Epstein, 1995). This model for conceptualizing drinking problems necessitates intervention at multiple levels: with the individual, the spouse, the marital relationship, and other social systems in which the drinker is involved. To implement the therapeutic model, couples are seen together to educate both partners about the nature of drinking problems and the model of change, to decrease the mistrust and hypervigilance that are common in such relationships, and to increase the information available to the therapist. Alcohol is typically the primary focus of the early treatment sessions, and the relationship is addressed more as therapy progresses. By focusing on alcohol at the beginning of treatment, the therapist allies with both partners, who are presenting with drinking as their major concern.

At the individual level, the treatment helps the client assess reinforcers for continued drinking and for abstinence, as well as assessing negative consequences of drinking and abstinence. This provides an incentive framework for the rest of the therapy. Other strategies are also used to enhance motivation to change. Clients are taught individual coping skills to deal with alcohol-related situations such as self-management planning, stimulus control, drink refusal, and self-monitoring of drinking and drinking impulses. Behavioral and cognitive coping skills, individually tailored to the types of situations that are the most common antecedents to drinking ("high-risk situations") are also taught and include assertiveness, cognitive restructuring, relaxation, lifestyle balance, and recreational activities.

A second set of interventions revolves around the coping behaviors of the partner. The partner's motivation for entering and continuing in treatment, and the partner's perceptions of the positive and negative consequences of changes in drinking and the marital relationship are important factors contributing to the partner's willingness to engage in new behaviors and be an active participant in the therapy. If the client is successful in changing his or her drinking, the spouse may be more amenable to examining and changing his or her own behavior. Spouse coping skills might include learning new ways to discuss drinking and drinking situations, learning new responses to the partner's drinking and alcohol-related behavior, or individual skills to enhance functioning.

The third set of interventions focuses on the interactions between the two partners, around both alcohol and other issues. Alcohol-focused couples interventions use alcohol-related topics as a vehicle to introduce communication and problem solving skills. Such topics as how the couple could manage in a situation where alcohol is present, whether or not they will keep alcohol in the house, how the partner could assist the drinker in dealing with impulses to drink, or what the couple will tell family and friends about the alcoholic's treatment are all relevant topics that the couple must face in the process of dealing with the drinking. By using such topics as vehicles for discussion, the couple is taught basic communication skills. In addition, the model suggests that many alcoholic couples need to learn general communication and problem-solving skills, to decrease marital conflicts that may cue drinking, and to increase the rate of positive exchanges. When appropriate, the treatment also incorporates general reciprocity enhancement interventions to increase the overall reward value of the relationship.

As the alcoholic becomes sober and both marital partners adjust their behavior to accommodate this change, the marital relationship typically improves and communication between partners increases and becomes more productive. In this way, positive benefits of the therapy filter to other family members and to family functioning as a whole. Specific domains of family functioning may be targeted in treatment, for example, both spouses may be encouraged to discuss and improve parenting skills. In therapy, partners may be encouraged to discuss initiation of nonalcohol focused family rituals or events. As couples learn to problem-solve more effectively together, they are better able to set the tone for effective and healthier family functioning.

The fourth set of interventions focuses on other social systems important to the drinker and partner. Clients are helped to identify interpersonal situations and persons who are associated with heavy drinking, and are also helped to identify potential social situations and people who would be supportive of abstinence or decreased drinking. Social skills such as refusing drinks or general assertiveness may be taught.

Finally, the model includes techniques to increase generalization to the natural environment and maintenance of new behaviors.

Homework assignments, teaching clients how to anticipate high-risk situations, and planned follow-up treatment sessions all are designed to contribute to maintenance of change.

Research studies evaluating the effectiveness of conjoint behavioral alcohol treatment compared to individually-focused treatment generally have shown an incremental benefit for spouse involvement, that is, such treatment results in significant reductions in alcohol consumption and improvements in marital functioning (O'Farrell, 1993), and lead to reductions in marital violence (O'Farrell & Murphy, 1995).

(ii) Cognitive-behavioral treatment of family members of the alcoholic

Research has indicated that including behavioral marital therapy in the treatment of alcohol problems can enhance treatment success and improve marital functioning. However, it has not been established whether behavioral treatment involving family members is more beneficial to the family members than treatment that is conducted only with the alcoholic or is conducted non-behaviorally. Research has indicated that, regardless of the manner in which an alcohol problem was addressed, families of alcoholics often function as well as nonalcoholic families if the alcoholic is able to achieve stable abstinence. If disturbances in family members persist after the alcoholic has become abstinent, these disturbance can be treated using cognitive-behavioral techniques specific to the observed problems; the model does not assume that families of alcoholics are unique in their functioning. If the alcoholic continues to drink and is unwilling to seek treatment, a number of cognitive-behavioral interventions can be utilized to improve the coping of family members and to enhance the likelihood of the alcoholic seeking treatment. These interventions are described below.

(a) Unilateral Family Therapy. Unilateral Family Therapy (UFT) (Thomas & Ager, 1993) has been developed to treat families in which an alcoholic family member is unwilling to reduce or quit drinking or seek treatment. The main goal of UFT is to use family members, usually spouses, as a positive rehabilitative influence to motivate the alcoholic family member to change drinking behavior or seek treatment for alcohol problems. Therapy focuses on three areas of intervention: the coping of the spouse; the interactions of the couple; and the individual alcoholic or identified patient. The cooperative spouse first is educated about alcohol problems and their treatment and an assessment of the alcoholic's drinking is done based on an interview with the spouse.

Steps are taken to help the spouse assume a rehabilitative role. Spouses are asked to monitor the alcoholic's drinking, both to reinforce their own recognition of the drinking problem and to provide the therapist with more information about the alcoholic. Unilateral relationship enhancement is used to reduce conflict and increase positive reinforcement from the relationship. Spouses are encouraged to increase the rates of pleasing behaviors, such as demonstrating affection, when the alcoholic is sober. Spouses are taught to reduce their use of coercive techniques for controlling their partners" drinking such as nagging and threatening. Also, the therapist assesses and tries to prevent enabling behaviors by the spouse that protect the alcoholic partner from the consequences of heavy drinking.

After the coping of the cooperative spouse and the interactions of the couple have been addressed, efforts are aimed at helping the spouse encourage the alcoholic partner to acknowledge the existence of an alcohol problem and to take steps towards moderating or discontinuing alcohol use. Efforts begin with nonconfrontive interventions such as increased support for sobriety or recommendation from a physician to change alcohol use. If there is some indication that the alcoholic family member may respond favorably, more confrontive interventions may be used such as bringing the alcoholic to session with the therapist and making a direct request for change or having family and friends gather to state their support for reduced drinking or abstinence. Assessment of the alcoholic's drinking and decisions about the most appropriate treatment setting for the alcoholic are completed prior to the programed request so that the alcoholic can be moved into treatment as soon as he or she agrees to seek help. Once the alcoholic seeks treatment, spouse-mediated relapse prevention training is provided and efforts at relationship enhancement continue.

A pilot study and an experimental evaluation of UFT found that spouses in UFT decreased control attempts and enabling behaviors and showed improvements in psychological distress. Alcoholics in UFT decreased drinking and were more likely to seek treatment than a wait-list control group. UFT appears to be effective particularly if the spouse is compliant with treatment and has a lower level of psychopathology and if the couple has lower levels of marital conflict.

(b) The Community Reinforcement Approach. Models of family treatment of uncooperative alcoholics based on the Community

Reinforcement Approach (CRA) (Sisson & Azrin, 1993) have been developed and are being investigated. Like UFT, CRA also uses family members, generally spouses or parents, as a positive rehabilitative influence. In the first few sessions, family members provide information on the alcoholic's drinking pattern, discuss the problems caused by the alcoholic's drinking, and are given the opportunity to vent emotions related to these problems. The therapist describes CRA and the importance of Antabuse and instructs family members to monitor alcohol consumption. Family members are then taught to help reduce the alcoholic's drinking and enhance motivation for treatment by providing the alcoholic positive consequences of sobriety and negative consequences of drinking, appropriately discussing drinking with the alcoholic, and scheduling events incompatible with drinking. The family members are taught how to assess potentially violent or dangerous situations and to remove themselves before violence occurs. They are also encouraged to become more independent from the alcoholic by developing outside interests.

A plan for getting the alcoholic into treatment is developed with the family member that can be implemented quickly at any time. The request or demand for the alcoholic to seek treatment is made at a time when the alcoholic's motivation for treatment might be high, such as following a particularly embarrassing episode of drunkenness. If the alcoholic agrees to seek treatment, he or she is seen immediately and prescribed Antabuse so that alcohol consumption will not take place for at least a few days. Treatment then follows the basic program of CRA as practiced with alcoholics seeking treatment on their own.

Early evaluations of the effectiveness of CRA for families are promising and this approach appears to be particularly effective in families in which parents are trying to encourage their children to seek treatment (Meyers, 1996). CRA resulted in more identified patients seeking treatment when compared with families randomly assigned to a Johnson Institute Intervention. Dropout was particularly high in the Johnson Institute Intervention condition, suggesting that CRA is more broadly accepted by significant others.

(i) Treatment of children

Behavioral treatments have been developed to treat young children of alcoholic parents regardless of whether the parents change their own drinking behavior. For instance, Nye, Zucker, and Fitzgerald (1995) reported that three to six-year-old sons of alcoholic fathers showed increases in prosocial and decreases in negative behaviors following a behavioral intervention. The treatment consisted of social learning based parent training and marital conflict problem-solving components. Outcome was most positive when the mother was involved and invested in the program. Mothers" psychopathology was associated with outcome, highlighting the importance of addressing psychological distress in the nonalcoholic parent. No controlled trials of behavioral interventions have been reported on daughters of alcoholic fathers or children of alcoholic mothers.

No studies are known to have reported on treatments aimed at improving functioning of elementary school-age children or adolescent children of alcoholics. The stress and coping model suggests that children of alcoholics will manifest symptoms of distress because they are forced to cope with a chronic and unpredictable stressor. Research has shown that increased levels of environmental stress and parental psychopathology are related to the poorer functioning of children of alcoholics (Chassin et al., 1991; Johnson & Jacob, 1995) and that children with difficult temperaments may be particularly likely to evidence behavioral problems when living with parents with high levels of antisocial behavior, low socioeconomic status, and frequent alcohol-related problems (Jansen, Fitzgerald, Ham, & Zucker, 1995). Furthermore, because of the disruptions caused by the alcoholic parent's drinking, children of alcoholics may lack appropriate role models and therefore show generalized social and problem-solving deficits (Begun & Zweben, 1990).

These research results suggest possible targets of interventions for children of alcoholics. As with spouses of alcoholics, changing the drinking of the alcoholic is likely to be the most powerful intervention for the child. If drinking continues, general training in problem-solving skills and stress management techniques may enhance functioning of children. The importance of the nonalcoholic parent to the functioning of children suggests that interventions aimed at improving the psychosocial functioning, communication skills, and parenting skills of the nonalcoholic parent could ameliorate problems related to the alcoholic parent's drinking. Interventions in the school system as well as in the peer network should also be considered. Because children of alcoholics are at increased risk for alcohol problems, both alcohol and drug use problems in adolescents should be assessed and treated as early as possible to prevent further dysfunction.

9.13.4.2 Family Systems Models

9.13.4.2.1 Family systems theory

Whereas cognitive-behavioral models of alcoholism and the family grew out of social learning theory, family systems models grew out of General Systems Theory, which was developed as an alternative to reductionistic/mechanistic, linear models of the functioning of biological systems. Although several different systems theory models have been developed, all focus on patterns of interrelationships between parts of larger systems, rather than focusing on the individual as the sole unit of analysis. Systems theory focuses on pattern recognition rather than deductive reasoning, and explains individual behavior within the context of larger systems such as the marriage, family, or community. Different systems theory models place differential emphasis on three concept areas: wholeness, organization, and relationships (Steinglass, 1978).

Family systems models began to influence the alcohol field in the 1970s. In a series of studies, Steinglass and colleagues observed the behavior of alcoholics and family members hospitalized on an experimental unit, noting repetitive, patterned interactions (Steinglass, Davis, & Berenson, 1977; Steinglass, Weiner, & Mendelson, 1971) as well as significant differences in their patterns of interaction when sober than when intoxicated. These observations led to the hypothesis that alcohol performed certain positive functions in a family, by stabilizing family roles, allowing for the expression of affection, allowing for greater intimacy among family members, or allowing for the exploration of topics that the family might avoid when sober. These positive functions were called the "adaptive consequences of alcoholism" (Davis, Berenson, Steinglass, & Davis, 1974). After these initial observational studies, alcoholic families were studied later in their homes as well as in the laboratory. It was reported that there were observable differences among families with an alcoholic family member who was drinking, sober, or in transition from one drinking status to another. Families with a sober alcoholic were most flexible in their functioning, having a balance between time together and time apart when at home (Steinglass, 1981) and showed more flexibility in solving structured tasks in the laboratory (Steinglass, 1979). Drinking families showed the most rigidity of family roles and interactions, while transitional families were intermediate in their functioning.

Wolin, Bennett, Noonan, and Teitelbaum (1980) provided a second family systems perspective on alcoholism, focusing primarily on the intergenerational transmission of alcoholism. Their work examined the family "rituals" that characterize all families: how families vacation, have dinner, celebrate holidays, etc. They found that those families whose rituals were maintained intact were least likely to have offspring who became alcoholic; those families whose rituals were most disrupted were most likely to have alcoholic offspring.

Family systems theory has contributed a number of constructs to the conceptualization of family functioning. Families are assumed to be governed by the law of homeostasis. All systems are assumed to operate to try to maintain balance, stability, equilibrium. Factors that threaten to change the functioning of a family threaten that homeostasis, and the family system is assumed to function to try to avoid change. Thus, family systems models would assume, if an alcoholic family has functioned as a stable family unit with a drinking member, that introducing sobriety into the system would threaten homeostasis. A variety of aspects of the family's structure and functioning (organization) all serve to maintain homeostasis. Family members have defined roles that guide their actions. Roles may range from caretaker to bad child, but these roles are well defined and are assumed to be difficult to modify over time. Families also have unwritten family rules that govern the functioning of the family. These rules are not necessarily explicit or deliberate, but are more likely to be implicit and unspoken. Families have a variety of boundaries that define relationships within the family and between the family and the rest of the world. Boundaries define alliances between family members and the degree to which information is available to different family members, family members have influence in the decision-making of the family, outsiders are welcomed into the family, and the family seeks interaction with outside persons and social institutions.

Family systems models (e.g., Steinglass et al., 1987) stress the reciprocal interactions between substance use and family functioning. Steinglass et al. posit that all families obey the general laws of all systems, including organization, homeostasis, circular causality of events, and feedback. For some families, alcoholism or drug abuse is an organizing principle for the family, and the presence or absence of the drug is the most important variable defining interactional behavior in the family.

9.13.4.2.2 Family systems assessment

Assessment in family systems models varies depending on the school of family therapy. Family therapy assessment focuses on assessment of the structure and functioning of the family as a unit, with more limited focus on assessment of the functioning of individual family members. There may be structured initial assessment, but assessment is an ongoing process that occurs throughout therapy, particularly when the clinician assesses the family's response to specific interventions.

The initial interview provides crucial assessment data for the family therapist. The family therapist asks specific questions about the family's view of their presenting problems and goals for therapy (e.g., Wetchler, McCollum, Nelson, Trepper, & Lewis, 1993). Some family therapists will also conduct a traditional diagnostic interview and mental status examination, and will assess for violence, suicidality, individual health, alcohol and drug use, and other aspects of individual functioning (Wetchler et al., 1993). The family therapist observes family interactions such as communication patterns, mixed messages, and approaches to conflict resolution skills, and observes structural aspects of the family such as boundaries, coalitions, shifting alliances, family rules and roles, splits, and triangulations (Kaufman, 1985). An important aspect of the family assessment involves assessment of behavioral sequences that involve alcohol use.

A variety of structured family assessment devices may be used with alcoholic families. Killorin and Olson (1984), who use Olson's Circumplex Model to conceptualize functioning in alcoholic families, use a Clinical Rating Scale to assess cohesion and adaptability in the family. Family therapists may also use a genogram interview to assess multigenerational alcohol and drug use, domestic violence, family rituals, and the role of alcohol in family rituals (Wetchler et al., 1993).

9.13.4.2.3 Family systems treatment

(i) Treatment of the alcoholic using family systems models

Using a family systems approach to treatment, the therapist must first differentiate "families with alcoholic members," from the "Alcoholic Family" (Steinglass et al., 1987). In the Alcoholic Family, family regulatory behaviors have been reorganized around the alcohol such that alcohol-related behaviors are central to family life. In the Alcoholic Family, change in the alcoholic's drinking necessitates a change in the family's basic organization and regulatory system. Interventions are rooted in the developmental perspective of the family systems approach, the Family Life History Model (FLH) (Steinglass et al.) which divides the family life cycle into early-, middle-, and late-phase developmental stages. Each phase presents particular systemic tasks, and the role of alcohol in the Alcoholic Family varies depending on life cycle phase. This introduces heterogeneity into the treatment approach, since treatment must be geared toward the appropriate family life cycle phase issues. For example, in an early-phase couple, the primary developmental task is differentiation of the family unit from the two families of origin, and the meaning of alcohol in the new unit will be influenced strongly by the view or use of alcohol in both families of origin. In the middle-phase Alcoholic Family, the alcoholic identity has been established, but the alcoholic behavior is embedded in short-term problem-solving behavior and, thus, will probably not be seen by the family members as the primary problem. In late-phase Alcoholic Families, it is often the children in the family who are struggling, either with drinking themselves or with establishing a separate identity for themselves.

In the first, diagnostic stage of treatment, the therapist takes a detailed history from the family and pays particular attention to three regulatory behaviors of the family: short-term problem-solving strategies, daily routines, and family rituals. Preferably, all individuals in the nuclear family are included in the family therapy sessions, especially in the beginning of treatment. Family members are encouraged to express their views of the problem. In later sessions, therapy may focus on the parental unit, and other family members may discontinue attendance at the sessions.

During the early sessions, the therapist determines which areas of family life have been "invaded" by the alcohol use. The therapist diagnoses alcoholism, labels it a family problem rather than a problem of the identified patient, and helps to identify what functions the alcohol plays in maintaining homeostasis for the family. Using reframing techniques, the therapist explains to the family the functions that alcohol serves (McKay, 1996). In this way, the focus is taken off the alcoholic and placed more squarely on the family system as a whole. At this point implications of sobriety for family functioning and homeostasis become more clear.

The therapist then focuses on removal of alcohol from the family system by helping the alcoholic to become abstinent. This is predicated on clarification of the problem and commitment of family members to common goals. Once drinking has stopped, the family faces reorganization of problem-solving strategies, family rituals, and interpersonal distances and boundaries, all of which have been structured around the alcohol. This stage of treatment is difficult; Steinglass et al. (1987) call this the "Emotional Desert," a period of dysphoria and disorganization. The therapist helps the family through to a resolution phase, during which a new homeostatic level is achieved, independent of a sobriety-intoxication cycle that defined the earlier homeostasis.

(ii) Family systems treatment of family members of the alcoholic

Individual family members are seen as components of a primary organizational unit, that is, the family. The alcoholic carries the family symptom, but the object of treatment is considered to be the interrelationships among family members, family behavior patterns, and the balance of the family system. Thus, family systems treatment of family members of the alcoholic is the same treatment as described above for the alcoholic. Even without the family member with the alcohol problem being involved in the treatment, family systems treatment focuses on all individuals in the family and the imbalances in the system as a whole.

9.13.4.3 Family Disease Models

9.13.4.3.1 Family disease model theory

The two previous models, cognitive-behavioral and family systems, developed outside of the alcohol field and have been applied to a wide range of human problems, including alcoholism. In contrast, family disease models initially were developed to explain drinking and family problems, and only later have been applied to a broader range of human problems.

The family has been a focus of the disease model almost from the beginning of Alcoholics Anonymous (AA), a self-help group intended to provide support for alcoholics. The primary publication of AA, often called the "Big Book" (Alcoholics Anonymous, 1976), includes a chapter with advice to family members. Al-Anon began in 1949 as a separate organization to assist the family and friends of alcoholics. The contemporary family focus of disease model approaches began with Cork's (1969) book on children of alcoholics, followed by Wegsheider's (1981) and other books about children in alcoholic families and the adult sequelae of their early experiences. Later authors began to focus on the partners of alcoholics (e.g., Beattie, 1987; Cermak, 1986). Although controlled research related to disease model conceptualizations is limited, these models have had a substantial impact on treatment and on popular thinking.

In the 1990s, disease model approaches describe alcoholism as a "family disease." Family members are seen as suffering from a disease, just as is the alcoholic. The "disease" of the family member is called "codependence," which has been described as a "recognizable pattern of personality traits, predictably found within most members of chemically dependent families" (Cermak, 1986, p. 1). Cermak proposed several specific symptoms of codependence: investing self-esteem in controlling self and others in the face of serious adverse circumstances, assuming responsibility for meeting the needs of others before one's own, anxiety and distortions of boundaries around issues of intimacy and separation, enmeshment in relationships with persons with personality disorders or alcohol or drug problems, and having at least three of ten other signs and symptoms: (i) denial as a primary coping strategy; (ii) constricted emotions; (iii) depression; (iv) hypervigilance; (v) compulsive behavior; (vi) anxiety; (vii) being a substance abuser; (viii) being a victim of physical or sexual abuse; (ix) having stress-related illnesses; and (x) being in a relationship with a substance abuser for more than two years without seeking help. Family members who are codependent are assumed to engage in a variety of behaviors that enable the substance abuser. Enabling refers to patterns of behavior that perpetuate the substance use, either by making it easier for the person to use, or by providing positive responses to use and avoiding negative or limit-setting responses.

9.13.4.3.2 Family disease model assessment

Assessment from a family disease perspective focuses on the individual functioning of spouses and children of the alcoholic, and does not address assessment of the functioning of the family or the marital dyad as a unit. Similar to family systems approaches, assessment may be embedded within the therapy, with clients self-identifying their own codependent characteristics as they listen to lectures, read program materials, and participate in group discussions. In addition, a number of self-report measures of codependency have been developed. Potter-Efron and Potter-Efron (1989) have developed a definition of codependency that focuses on the effects of involvement with an alcoholic, and identified eight major types of effects: fear, shame and guilt, "prolonged despair," anger, denial, rigidity, "impaired identity development" (p. 39), and confusion. They developed an instrument to assess each of these areas of functioning, and defined criteria for diagnosing a respondent as codependent. The Webb-Fabean Codependence Assessment Guide

(Webb, Fabean, & Whalen, 1993a, 1993b) takes a broader view of codependence, drawing upon family systems perspectives as well as the family disease model, and emphasizes dysfunctional patterns of interaction, loss of autonomy, and "a distortion of reality" (1993a, p. 41).

Another aspect of family disease model assessment focuses on the identification of children of alcoholics. The Children of Alcoholics Screening Test (CAST) (Jones, 1983; Staley & el-Guebaly, 1991) assesses latency age, adolescent and adult children of alcoholics" perceptions of the drinking of their parents, and also assesses psychological characteristics believed to be characteristic of children of alcoholics. A short version of the CAST focuses only on the actual drinking of the parents (Hodgins, Maticka-Tyndale, el-Guebaly, & West, 1993).

9.13.4.3.3 Family disease model treatment

(i) Family disease model treatment of the alcoholic

Since proponents of the disease model consider alcoholism a family disease, treatment focuses on the alcoholic and family members. In this way, family disease models of treatment are somewhat similar to family systems models. In fact, family disease model treatment is based partially on the work of early system theorists, such as Bowen (1978), and some family disease model programs incorporate systems-based interventions in their treatment agenda (Laundergan & Williams, 1993) A major difference, however, is the disease model focus on self-help groups and on concepts such as enabling, detachment, and codependency, which are not associated with systems approaches. Through separate but parallel involvement with aspects of the 12-step recovery program, the alcoholic works on becoming and remaining abstinent while family members learn about chemical dependency and are encouraged to focus on their own needs rather than continue to organize their lives around alcoholism. A basic tenet of the family disease model is that family members should recognize the effect that alcoholism has had on their lives but allow the alcoholic to take responsibility for his or her own recovery.

Treatment of the alcoholic family member focuses on facilitation of participation in AA. This can be done on an inpatient or outpatient basis. Also, many people with an alcohol problem do not seek formal counseling, but become involved with AA on their own. The focus of AA on the alcoholic's recovery, especially in the early stages of recovery, precludes formal marital or family therapy until the alcoholic has attained sobriety on their own. Some disease model treatment programs (Nowinski, Baker, & Carroll, 1992) include conjoint sessions towards the end of treatment, to specifically address issues such as enabling and detachment.

(ii) Family disease model treatment of family members of the alcoholic

As mentioned above, families of alcoholics are presumed to suffer from the disease themselves, though it takes a different form in family members than in the alcoholic. Family members are treated not for a drinking problem, but for codependence. Treatment is through attendance at disease model self-help groups, such as Al-Anon, Al-Ateen (groups for teenaged family members or alcoholics), ACOA (Adult Children of Alcoholics) groups, and COA (Children of Alcoholics) groups. Family members learn about the concept of codependence, and about the effect alcohol has had on their lives and their behavior. They learn that a primary component of overcoming codependence is to stop taking responsibility for the alcoholic's drinking. Family members learn to do this by stopping their enabling behaviors, which minimize the negative consequences of the drinking (Nowinski et al., 1992). Family members are also encouraged to work towards "detachment," which is a functional response to the alcoholism that allows the alcoholic to take responsibility for their alcoholic behavior and its consequences. In detaching, the family members are encouraged to focus on their own well being, and to neutralize alcohol and the alcoholic as the focus of their lives.

Treatment of family members can occur on an inpatient basis. Most inpatient rehabilitation programs include a family component in which family members participate in family psychoeducational sessions. Families can also opt for more intensive involvement in some treatment sites, and spend several days living at the treatment facility and participating in a family program. The Hazelden Residential Family Program is a good example of this intensive type of treatment for family members of alcoholics (Laundergan & Williams, 1993). Family members live on site for five days, and participate in 12-step Al-Anon focused treatment in a multifamily group format. Treatment objectives include educating family members about chemical dependency and recovery and their effects on the family, learning about the universality of experience among family members of alcoholics, developing new strategies for coping with family relationships and anxiety, and learning about Al-Anon concepts such as enabling and detachment.

Finally, in instances where the alcoholic is unwilling to seek treatment, therapists following the disease model may utilize a particular type of family-based technique, called "Intervention," which is designed to motivate alcoholics to enter treatment. Intervention assumes that the alcoholic is in denial of their alcohol problem, that the family may have significant influence over the alcoholic's decision to seek treatment, and that the family can begin to recover along with the alcoholic (Liepman, 1993). This technique is also based on the family systems perspective. A family member (or "emissary") contacts a treatment professional who determines if the family is a good candidate for Intervention. Particular family members are then chosen to attend Intervention team training sessions. Teams may include both family members and close friends who have been affected by the alcoholism and who have influence over the alcoholic. Training sessions are used to prepare the team for the process of Intervention, to generate a list of concerns about the alcoholic's drinking and desired outcomes of the Intervention, to teach appropriate communication skills, and to rehearse the presentation. Contingency plans are made for possible resistance. A site is selected, and the alcoholic is brought to the site on the specified day, without prior knowledge, for the "confrontation." The alcoholic is asked to listen quietly while each team member presents their concerns about the drinking. The team emphasizes the effect the alcoholism has had on the alcoholic and on themselves, and the notion that the Intervention is being done in the context of love and concern for the alcoholic. The team asks the alcoholic to accept treatment. If the Intervention is successful, the alcoholic typically is taken immediately to a preselected treatment site, where preparations for admission have been made in advance. In order for an Intervention to be successful, the Interventionist must be highly skilled, the family must be quite motivated, and the preparations extensive (see Liepman, 1993).

9.13.5 SUMMARY AND CONCLUSIONS

Alcoholism affects the family profoundly, and family members are in a unique position to exert influence to help the alcoholic stop drinking. This chapter provides a broad overview of historical and modern theory, research, and treatment related to families of alcoholics, focusing, in particular, on three contemporary models of family treatment of alcoholism: cognitive-behavioral, family systems, and family disease models.

9.13.5.1 Common Elements of the Models

Although each model is distinct and has evolved from a different school of thought, several elements are common to the three approaches. All three approaches share the common goal of helping the alcoholic to eliminate alcohol use and to reduce the number and severity of alcohol-related problems. Each model seeks to enhance well being, to relieve alcohol-related distress of the alcoholic and family members, and to improve psychosocial and interactional functioning of the family in general.

For all three treatments, the issue of motivation to seek treatment is especially important, and each model has developed techniques to employ the spouse and/or family's influence to motivate the alcoholic, or to help the family despite lack of motivation on the alcoholic's part. The most extreme example is the Intervention, which uses both family disease and family systems interventions to reduce denial of the alcoholic and move the alcoholic toward acceptance of treatment. Unilateral family therapy has been developed to treat spouses and other family members of treatment refusing alcoholics. In each case, the family is involved in the issue of treatment seeking behavior of the alcoholic. Future research might focus on continuing to develop and evaluate methods by which family members can help motivate the alcoholic to get help.

Another issue common to all three models is the unit of target for treatment, that is, who is being treated, and who is included in the treatment sessions. Family therapy can include the spousal dyad, the entire nuclear family, the family of origin, and, in some systems and disease models, even other relatives and close friends. Cognitive-behavioral models, up to 1997, have focused on the parent dyad; techniques which include the entire family have yet to be developed. Family systems therapists view family interactions as the target of treatment, while family disease approaches emphasize personal responsibility and separate but parallel treatments. Researchers have focused on integration of disease model concepts into a cognitive-behavioral marital therapy format, with good results (McCrady, Epstein, & Hirsch, 1996). Further work might address the efficacy of cognitive-behavioral family therapy, and incorporation of family disease techniques in the context of couples or family therapy sessions.

9.13.5.2 Strengths of Each Model

Each of the three models of families of alcoholics has particular strengths. The cognitive-behavioral family model is grounded in a broader social learning theory, and is based on empirical studies of efficacy. Detailed manuals are available to help with training and delivery of cognitive-behavioral treatments. Cognitive-behavioral models are also very strong in assessment of alcohol consumption patterns, including behavioral and cognitive antecedents and consequences, and of coping patterns of the alcoholic and spouse.

The family systems model for alcoholism also has several strengths. For instance, this approach is rooted in a broader, theory-based systems model which is well-known and widely accessible. Family systems treatment of alcoholism allows for extensive family involvement and, thus, facilitates comprehensive assessment of the impact of alcohol on family functioning. This approach also facilitates identification and treatment of associated family problems and focuses attention on changes in family patterning and readjustment to a sober lifestyle.

Several strengths of the family disease model are notable. Family disease model treatment is quite prevalent, and AA and Al-Anon meetings are available almost daily in most areas. AA and other 12-step meetings are free of charge, making them even more accessible for attendance whenever necessary. Family disease model treatment thus provides free and accessible fellowship, psychoeducational components, and support on an ongoing basis, for years if necessary.

In summary, the three contemporary models of family treatment of alcoholism each have particular strengths, a defined theoretical approach, a set of treatment techniques, and varying degrees of empirical support. The choice of which approach to use typically is based not on particular patient characteristics, but on the training and theoretical framework of the therapist. Further work on patient–treatment matching is warranted.

The notion of cross-fertilization of treatment techniques is being explored on a limited basis. It is quite difficult to become equally facile with all three treatment approaches, since each requires years of training and practice to master. However, therapists who are solidly grounded in one school of thought might master techniques from other models to employ with certain patients. For instance, a therapist trained in cognitive-behavioral marital therapy could use family systems theory to help conceptualize cases in terms of family symptoms and to identify patterns of dysfunction. A therapist who is rooted in family disease theory might want to use particular cognitive-behavioral techniques such as functional analysis to assess drinking patterns and consequences. Further work in this area, testing whether therapists can integrate treatment approaches successfully and whether such integration increases efficacy of the treatment, is encouraged.

9.13.6 REFERENCES

Alcoholics Anonymous (1976). *Alcoholics Anonymous: The story of how many thousands of men and women have recovered from alcoholism* (3rd ed.). New York: Alcoholics Anonymous World Services.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. Rev.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Annis, H. M. (1982). *Inventory of Drinking Situations (IDS-100)*. Toronto, Canada: Addiction Research Foundation.

Babor, T. F. (1992). Cross-cultural research on alcohol: A quoi bon? In J. E. Helzer & G. J. Canino (Eds.), *Alcoholism in North America, Europe, and Asia* (pp. 33–52). New York: Oxford University Press.

Beattie, M. (1987). *Co-dependent no more*. Minneapolis, MN: Hazelden.

Begun, A., & Zweben, A. (1990). Assessment and treatment of adjustment and coping capacities in children living with alcoholic parents. *Alcoholism Treatment Quarterly, 7*(2), 23–40.

Berger, A. (1981). Family involvement and alcoholics" completion of a multiphase treatment program. *Journal of Studies on Alcohol, 42*, 517–520.

Bowen, M. (1978). *Family therapy in clinical practice*. New York: Aronson.

Bucholz, K. K. (1992). Alcohol abuse and dependence from a psychiatric epidemiologic perspective. *Alcohol Health and Research World, 16*, 197–208.

Canino, G. J., Burnam, A., & Caetano, R. (1992). The prevalence of alcohol abuse and/or dependence in two Hispanic communities. In J. E. Helzer & G. J. Canino (Eds.), *Alcoholism in North America, Europe, and Asia* (pp. 131–155). New York: Oxford University Press.

Catalano, R. F., Morrison, D. M., Wells, E. A., Gillmore, M. R., Iritani, B., & Hawkins, J. D. (1992). Ethnic differences in family factors related to early drug initiation. *Journal of Studies on Alcohol, 53*, 208–217.

Cermak, T. (1986). *Diagnosing and treating co-dependence*. Minneapolis, MN: Johnson Institute Books.

Chassin, L., Rogosch, F., & Barrera, M. (1991). Substance use and symptomatology among adolescent children of alcoholics. *Journal of Abnormal Psychology, 100*, 449–463.

Collins, R. L. (1990). Family treatment of alcohol abuse: Behavioral and systems perspectives. In R. L. Collins, K. E. Leonard, & J. S. Searles (Eds.), *Alcohol and the family: Research and clinical perspectives* (pp. 285–308). New York: Guilford Press.

Collins, R. L. (1992). Methodological issues in conducting substance abuse research on ethnic minority populations. *Drugs and Society, 6*, 59–77.

Cork, M. (1969). *The forgotten children*. Toronto, Canada: Addiction Research Foundation.

Davis, D. I., Berenson, D., Steinglass, P., & Davis, S. (1974). The adaptive consequences of drinking. *Psychiatry, 37*, 209–215.

Goodwin, D. W., Schulsinger, F., Hermansen, L., Guze, S. B., & Winokur, G. (1973). Alcohol problems in adoptees raised apart from alcoholic biological parents. *Archives of General Psychiatry, 28*, 238–243.

Gottman, J. M. (1996). *What predicts divorce? The measures*. Mahwah, NJ: Erlbaum.

Harford, T. C. (1992). Family history of alcoholism in the United States: Prevalence and demographic characteristics. *British Journal of Addictions, 87*, 931–935.

Heavey, J. M., & Dodd, D. K. (1995). Children of alcoholics, negative life events, and early experimentation with drugs. *Journal of School Psychology, 33*, 305–317.

Helzer, J. E., Bucholz, K., & Robins, L. N. (1992). Five communities in the United States: Results of the Epidemiologic Catchment Area survey. In J. E. Helzer & G. J. Canino (Eds.), *Alcoholism in North America, Europe, and Asia* (pp. 71–96). New York: Oxford University Press.

Helzer, J. E., Burnam, A., & McEvoy, L. T. (1991). Alcohol abuse and dependence. In L. N. Robins & D. A. Regier (Eds.), *Psychiatric disorders in America: The Epidemiologic Catchment Area Study* (pp. 81–115). New York: Free Press.

Helzer, J. E., & Canino, G. J. (1992). Comparative analysis of alcoholism in 10 cultural regions. In J. E. Helzer & G. J. Canino (Eds.), *Alcoholism in North America, Europe, and Asia* (pp. 289–308). New York: Oxford University Press.

Herd, D. (1994). Predicting drinking problems among Black and White men: Results from a national survey. *Journal of Studies on Alcohol, 55*, 61–71.

Hill, S. Y. (1993). Personality characteristics of sisters and spouses of male alcoholics. *Alcoholism: Clinical and Experimental Research, 17*, 733–739.

Hodgins, D. C., Matick-Tyndale, E. M., el-Guebaly, N., & West, M. (1993). The CAST-6: Development of a short-form of the Children of Alcoholics Screening Test. *Addictive Behaviors, 18*, 337–345.

Holder, H. D., & Hallan, J. B. (1986). Impact of alcoholism treatment on total health care costs: A six-year study. *Advances in Alcoholism and Substance Abuse, 6*, 1–15.

Hurley, D. L. (1991). Women, alcohol and incest: An analytic review. *Journal of Studies on Alcohol, 52*, 253–268.

Jackson, J. (1954). The adjustment of the family to the crisis of alcoholism. *Quarterly Journal of Studies on Alcohol, 15*, 562–586.

Jansen, R. E., Fitzgerald, H. E., Ham, H. P., & Zucker, R. A. (1995). Pathways into risk: Temperament and behavior problems in three- to five-year-old sons of alcoholics. *Alcoholism: Clinical and Experimental Research, 19*, 501–509.

Johnson, S. L. & Jacob, T. (1995). Psychosocial functioning in children of alcoholic fathers. *Psychology of Addictive Behaviors, 9*, 101–113.

Jones, J. W. (1983). The Children of Alcoholics Screening Test: A validity study. *Bulletin of the Society of Psychologists in Addictive Behaviors, 2*, 155–163.

Kaufman, E. (1985). *Substance abuse and family therapy*. Orlando, FL: Grune & Stratton.

Killorin, E. & Olson, D. H. (1984). The chaotic flippers in treatment. In E. Kaufman (Ed.), *Power to change. Family case studies in the treatment of alcoholism* (pp. 99–130). New York: Gardner.

Kogan, K. L., & Jackson, J. K. (1965). Stress, personality, and emotional disturbance in the wives of alcoholics. *Quarterly Journal of Studies on Alcohol, 26*, 486–495.

Langenbucher, J. W., McCrady, B. S., Brick, J., & Esterly, R. E. (1994). *Socioeconomic evaluations of addictions treatment*. Washington, DC: The White House.

Laundergan, J. C., & Williams, T. (1993). The Hazelden residential family program: A combined systems and disease model approach. In T. J. O'Farrell (Ed.), *Treating alcohol problems: Marital and family Interventions (pp. 145–169). New York: Guilford Press.*

Legge, C., & Sherlock, L. (1991). Perception of alcohol use and misuse in three ethnic communities: Implications for prevention programming. *The International Journal of Addictions, 25*, 629–653.

Leonard, K. E., & Jacob, T. (1988). Alcohol, alcoholism, and family violence. In V. B. Van Hasselt, R. L. Morrison, A. S. Bellack, & M. Hersen (Eds.), *Handbook of family violence* (pp. 383–406). New York: Plenum.

Lewis, M. L. (1937). Alcoholism and family casework. *Social Casework, 35*, 8–14.

Lex, B. W. (1987). Review of alcohol problems in ethnic minority groups. *Journal of Consulting and Clinical Psychology, 55*, 293–300.

Liepman, M. R. (1993). Using family influence to motivate alcoholics to enter treatment: The Johnson Institute Intervention Approach. In T. J. O'Farrell (Ed.), *Treating alcohol problems: Marital and family interventions* (pp. 54–77). New York: Guilford Press.

Love, C. T., Longabaugh, R., Clifford, P. R., Beattie, M., & Peaslee, C. F. (1993). The Significant-other Behavior Questionnaire (SBQ): An instrument for measuring the behavior of significant others towards a person's drinking and abstinence. *Addiction, 88*, 1267–1279.

MacDonald, D. E. (1956). Mental disorders in wives of alcoholics. *Quarterly Journal of Studies on Alcohol, 17*, 282–287.

Manson, S. P., Shore, J. H., Baron, A. E., Ackerson, L., & Neligh, G. (1992). Alcohol abuse and dependence among American Indians. In J. E. Helzer & G. J. Canino (Eds.), *Alcoholism in North America, Europe, and Asia* (pp. 113–130). New York: Oxford University Press.

Margolin, G., Talovic, S., & Weinstein, C. D. (1983). Areas of Change Questionnaire: A practical approach to marital assessment. *Journal of Consulting and Clinical Psychology, 51*, 920–931.

Martin, S. E. (1992). The epidemiology of alcohol-related interpersonal violence. *Alcohol Health and Research World, 16*, 230–237.

McCrady, B. S. (1993). Alcoholism. In D. H. Barlow (Ed.), *Clinical handbook of psychological disorders*. (2nd. ed., pp. 362–395). New York: Guilford Press.

McCrady, B. S., & Epstein, E. E. (1995). Marital therapy in the treatment of alcohol problems. In N. S. Jacobson & A. S. Gurman (Eds.), *Clinical handbook of couple therapy* (pp. 369–393). New York: Guilford Press.

McCrady, B. S., & Epstein, E. E. (1996). Theoretical bases of family approaches to substance abuse treatment. In F. Rotgers, D. S. Keller, & J. Morgenstern (Eds.), *Treating substance abuse: Theory and technique* (pp. 117–142). New York: Guilford Press.

McCrady, B. S., Epstein, E. E., & Hirsch, L. S. (1996). Conducting research on Alcoholics Anonymous: A model for using the randomized clinical trial. *Journal of Studies on Alcohol, 57*, 604–612.

McCrady, B. S., Kahler, C. W., Shraedley, P., Epstein, E. E., & Hirsch, L. (1995, June). *Coping and psychopathology: Models of wives of alcoholics*. Presented at the Annual Meeting of the Research Society on Alcoholism, Steamboat Springs, CO.

McCrady, B. S., Miller, K., Epstein, E. E., & Van Horn, D. (1993, November). *Spouses of alcoholics: Measurement of coping behaviors and self efficacy for coping*. Presented at the Annual Meeting of the Association for Advancement of Behavior Therapy, Atlanta, GA.

McCrady, B. S., Stout, R., Noel, N., Abrams, D., & Nelson, H. F. (1991). Effectiveness of three types of spouse-involved behavioral alcoholism treatment.

British Journal of Addiction, 86, 1415–1424.

McKay, J. R. (1996). Family therapy techniques. In F. Rotgers, D. S. Keller, & J. Morgenstern (Eds.), *Treating substance abuse: Theory and technique* (pp. 143–173). New York: Guilford Press.

Meyers, R. J. (1996, June). *A CRA strategy for engaging unmotivated alcohol abusers in treatment*. Paper presented at the Research Society on Alcoholism and the International Society for Biomedical Research on Alcoholism Joint Scientific Meeting, Washington, DC.

Miller, B. A. (1990). The interrelationships between alcohol and drugs and family violence. In M. De Las Rosa, E. Y. Lambert, & B. Gropper (Eds.), *National Institute on Drug Abuse Research Monograph Series 103: Drugs and violence: Causes, correlates, and consequences* (DDHS Publication No. ADM 90–1721, pp. 177–207). Washington, DC: US Government Printing Office.

Moos, R. H., Finney, J. W., & Gamble, W. (1982). The process of recovery from alcoholism II: Comparing spouse of alcoholic patients and matched community controls. *Journal of Studies on Alcohol, 43*, 888–909.

Moos, R. H., & Moos, B. S. (1984). The process of recovery from alcoholism: III. Comparing functioning in families of alcoholics and matched control families. *Journal of Studies on Alcohol*, *45*, 111–118.

Nathan, P. E., & Skinstad, A. H. (1987). Outcomes of treatment for alcohol problems: Current methods, problems, and results. *Journal of Consulting and Clinical Psychology*, *55*, 332–340.

National Institute on Alcohol Abuse and Alcoholism. (1993). *Eighth special report to the U.S. Congress on alcohol and health*. Washington, DC: EEI.

Navran, L. (1967). Communication and adjustment in marriage. *Family Process*, *6*, 173–184.

Nowinski, J., Baker, S., & Carroll, K. (1992). *Twelve Step facilitation therapy manual*. National Institute on Alcohol Abuse and Alcoholism, Project MATCH Monograph Series, Vol. 1. Rockville, MD.

Nye, C. L., Zucker, R. A., & Fitzgerald, H. E. (1995). Early intervention in the path to alcohol problems through conduct problems: Treatment involvement and child behavior change. *Journal of Consulting and Clinical Psychology*, *63*, 831–840.

O'Farrell, T. J. (1993). *Treating alcohol problems: Marital and family interventions*. New York: Guilford Press.

O'Farrell, T. J., & Choquette, K. (1991). Marital violence in the year before and after spouse-involved alcoholism treatment. *Family Dynamics of Addiction Quarterly*, *1*, 32–40.

O'Farrell, T. J., & Murphy, C. M. (1995). Marital violence before and after alcoholism treatment. *Journal of Consulting and Clinical Psychology*, *63*, 256–262.

Paolino, T. J. Jr., McCrady, B. S., & Diamond, S. (1978). Statistics on alcoholic marriages: An overview. *The International Journal of Addictions*, *13*, 1285–1293.

Pitman, N. E., & Taylor, R. G. (1992). MMPI profiles of partners of incestuous sexual offenders and partners of alcoholics. *Family Dynamics of Addiction Quarterly*, *2*, 52–59.

Potter-Efron, R. T., & Potter-Efron, P. S. (1989). Assessment of codependency with individuals from alcoholic and chemically dependent families. *Alcoholism Treatment Quarterly*, *6*, 37–57.

Reich, W., Earls, F., & Powell, J. (1988). A comparison of the home and social environments of children of alcoholics and nonalcoholic parents. *British Journal of Addiction*, *83*, 831–839.

Rychtarik, R. G. (1990). Alcohol-related coping skills in spouses of alcoholics: Assessment and implications for treatment. In R. L. Collins, K. E. Leonard, & J. S. Searles (Eds.), *Alcohol and the family. Research and clinical perspectives* (pp. 356–379). New York: Guilford Press.

Sher, K. J. (1991). *Children of alcoholics: A critical appraisal of theory and research*. Chicago: University of Chicago Press.

Sher, K. J., Walitzer, K. S., Wood, P. K., & Brent, E. E. (1991). Characteristics of children of alcoholics: Putative risk factors, substance use and abuse, and psychopathology. *Journal of Abnormal Psychology*, *100*, 427–448.

Sisson, R. W., & Azrin, N. H. (1993). Community reinforcement training for families: A method to get alcoholics into treatment. In T. J. O'Farrell (Ed.), *Treating alcohol problems: Marital and family interventions* (pp. 34–53). New York: Guilford Press.

Spanier, G. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family*, *38*, 15–28.

Staley, D., & el-Guebaly, N. (1991). Psychometric evaluation of the Children of Alcoholics Screening Test (CAST) in a psychiatric sample. *International Journal of the Addictions, 26*, 657–668.

Steinglass, P. (1978). The conceptualization of marriage from a systems theory perspective. In T. J. Paolino, Jr. & B. S. McCrady (Eds.), *Marriage and marital therapy: Psychoanalytic, behavioral and systems theory perspectives* (pp. 298–365). New York: Brunner/Mazel.

Steinglass, P. (1979). The alcoholic family in the interaction laboratory. *Journal of Nervous and Mental Disease, 167*, 428–436.

Steinglass, P. (1981). The alcoholic family at home. Patterns of interaction in dry, wet, and transitional stage of alcoholism. *Archives of General Psychiatry, 38*, 578–584.

Steinglass, P., Bennett, L. A., Wolin, S. J., & Reiss, D. (1987). *The alcoholic family*. New York: Basic Books.

Steinglass, P., Davis, D. I., & Berenson, D. (1977). Observations of conjointly hospitalized "alcoholic couples" during sobriety and intoxication: Implications for theory and therapy. *Family Process, 16*, 1–16.

Steinglass, P., Weiner, S., & Mendelson, J. H. (1971). Interactional issues as determinants of alcoholism. *American Journal of Psychiatry, 128*, 275–280.

Strauss, M. (1979). Measuring intrafamily conflict and violence: The Conflict Tactics (CT) Scales. *Journal of Marriage and the Family, 41*, 75–88.

Thomas, E. J., & Ager, R. D. (1993). Unilateral family therapy with spouses of uncooperative alcohol abusers. In T. J. O'Farrell (Ed.), *Treating alcohol problems: Marital and family interventions* (pp. 3–33). New York: Guilford Press.

Thomas, E. J., Yosioka, M. R., & Ager, R. D. (1994a). Spouse Enabling Inventory (SEI). In J. Fisher & K. Corcoran (Eds.), *Measures for clinical practice: A sourcebook. Vol. 1. Couples, families and children* (pp. 177–182). New York: Free Press.

Thomas, E. J., Yosioka, M. R., & Ager, R. D. (1994b). Spouse Sobriety Influence Inventory (SSII). In J. Fisher & K. Corcoran (Eds.), *Measures for clinical practice: A sourcebook. Volume I. Couples, families and children* (pp. 183–195). New York: Free Press.

Webb, M. T., Fabean, A., & Whalen, S. (1993a). A new means for measuring codependence. The Webb-Fabean assessment guide. *Professional Counselor, 7*, 40–42.

Webb, M. T., Fabean, A., & Whalen, S. (1993b). A new means for measuring codependence. The Webb-Fabean assessment guide. *Professional Counselor, 7*, 32–36.

Wegsheider, S. (1981). *Another chance: Hope and health for the alcoholic family*. Palo Alto, CA: Science and Behavior Books.

Welte, J. W., & Barnes, G. M. (1987). Alcohol use among adolescent minority groups. *Journal of Studies on Alcohol, 48*, 329–336.

Wetchler, J. L., McCollum, E. E., Nelson, T. S., Trepper, T. S., & Lewis, R. A. (1993). Systemic couples therapy for alcohol-abusing women. In T. J. O'Farrell (Ed.), *Treating alcohol problems. Marital and family interventions* (pp. 236–260). New York: Guilford Press.

Whalen, T. (1953). Wives of alcoholics: Four types observed in a family service agency. *Quarterly Journal of Studies on Alcohol, 14*, 632–641.

Wittchen, H. U., & Bronisch, T. (1992). Alcohol use, abuse, and dependency in West Germany: Lifetime and six-month prevalence in the Munich follow-up study. In J. E. Helzer & G. J. Canino (Eds.), *Alcoholism in North America, Europe, and Asia* (pp. 150–181). New York: Oxford University Press.

Wolin, S. J., Bennett, L. A., Noonan, D. L., & Teitelbaum, M. A. (1980). Disrupted family rituals: A factor in the intergenerational transmission of alcoholism. *Journal of Studies on Alcohol, 41*, 199–214.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.14
Families of Children with Chronic Medical Conditions

ROBERT J. THOMPSON, JR.
Duke University Medical Center, Durham, NC, USA
and
LAUREN B. RAEZER
Duke University, Durham, NC, USA

9.14.1 INTRODUCTION 219

9.14.2 CHRONIC MEDICAL CONDITIONS 220
- *9.14.2.1 Nature* 220
- *9.14.2.2 Prevalence* 221
- *9.14.2.3 Assessment* 222
- *9.14.2.4 Conceptual Models* 222
  - *9.14.2.4.1 Psychological adjustment of children with chronic illness* 223
- *9.14.2.5 Psychological Adjustment of Siblings* 224
- *9.14.2.6 Psychological Adjustment of Parents* 224
- *9.14.2.7 Treatment* 225
  - *9.14.2.7.1 Psychosocial adjustment* 226
  - *9.14.2.7.2 Adherence* 226
- *9.14.2.8 Pain Management* 226

9.14.3 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 227
- *9.14.3.1 Parenting* 227
- *9.14.3.2 Family Systems Functioning* 228
- *9.14.3.3 The Psychosocial Demands of Illness* 228
  - *9.14.3.3.1 The psychosocial demands of the family life cycle* 229
  - *9.14.3.3.2 The interface between the psychosocial demands of the illness and the family life cycle* 230

9.14.4 SUMMARY 230

9.14.5 REFERENCES 231

9.14.1 INTRODUCTION

Advances in biomedical science and health care have prolonged the lives of children who previously might have died in early childhood. As a consequence, once acutely lethal illnesses have now become chronic conditions. Correspondingly, many children and their families are engaged in the continuous process of adapting to the stresses and tasks associated with chronic medical conditions over substantial periods of their lives. This has led to a focus on enhancing quality of life in general, and psychosocial adjustment in particular, by fostering adaptation (Thompson & Gustafson, 1996).

Concurrent with these advances in biomedical science, there have been advances in understanding about how biological and psychosocial processes act together in human development across the life course. The social–ecological systems theory perspective on human development has provided a basis for investigating the process of adaptation to chronic medical conditions (Bronfenbrenner, 1977, 1979). There are two major tenets of this theory: there is a reciprocal influence between the child and his or her environmental settings; and transitions, that is, successive shifts in role and setting across the life span have developmental significance. The setting of immediate importance to the child is the family system.

Research on adaptation to chronic childhood medical conditions has moved to the analytic level in which various models have been proposed to delineate individual and family adaptational processes associated with quality of life. In general, these models have adopted a stress and coping framework based on cognitive-behavioral and social learning theories (Thompson & Gustafson, 1996; Wallander & Thompson, 1995). The development of these models has enabled application of the considerable body of knowledge regarding the role of cognitive processes, coping methods, and social support in managing stress to the specific situation of adapting to chronic childhood illness. In turn, new knowledge has been generated regarding adaptational processes, particularly in terms of developmental aspects. Dimensions of family functioning have been among the adaptational processes investigated. One of the purposes of this chapter is to summarize these findings. However, research in this area is just beginning to include more comprehensive family systems theory perspectives. Another purpose of this chapter is to integrate family system theories with cognitive-behavioral–social learning theories to formulate a conceptual framework to guide the next stage of research. This will provide the necessary basis for the development of intervention programs to enhance adaptation by fostering the goodness of fit of the special needs and characteristics of the child with a chronic medical condition with his or her family setting.

9.14.2 CHRONIC MEDICAL CONDITIONS

9.14.2.1 Nature

Chronic conditions of childhood include developmental disabilities and chronic diseases. The primary focus of this chapter is on chronic childhood illness. A generally accepted definition of a chronic illness has been provided by Mattsson (1972):

> a disorder with a protracted course which can be progressive and fatal, or associated with a relatively normal life span despite impaired physical or mental functioning. Such a disease frequently shows periods of acute exacerbation requiring medical attention. (p. 801)

There are a large number of specific chronic childhood illnesses. It is essential to understand the impact of disordered biological processes that result in specific symptoms and functional impairments and treatment regimens. It is also essential to recognize that the impact of chronic childhood illness is dynamic because it occurs within the context of ongoing developmental processes reflecting the mutual and continuous influence of the child on his or her environment (Thompson & Gustafson, 1996). Thus, an appreciation of multiple dimensions of chronic childhood medical conditions requires a biopsychosocial framework (Thompson & Gustafson, 1996).

In terms of biomedical dimensions, there is considerable diversity among the chronic illnesses of childhood in etiology and manifestation. Many conditions have a genetic component that affects the functioning of key biological processes, for example, hemophilia and muscular dystrophy, which are sex-linked recessive disorders, and cystic fibrosis and sickle cell disease, which are autosomal recessive disorders. Other conditions, such as cancer, have less clear etiologies and may reflect a combination of environmental factors and genetic predispositions.

Chronic conditions also vary in structural and functional manifestations. This variability has been the basis for classification systems that group conditions according to the body system affected such as endocrine, metabolic, hematologic, respiratory, digestive, and musculoskeletal systems (Newacheck, Budetti, & Haflon, 1986). Illnesses have also been categorized according to functional impact which includes both biomedical and psychosocial dimensions such as activity limitations and social roles. For example, Stein and Jessop (1982) developed a measure of functional status which assesses behavioral responses to illness, including communication, mobility, mood, sleeping, and eating, that interfere with the performance of normal social roles for infants, toddlers, preschoolers, and school-age children.

In terms of psychosocial dimensions, chronic childhood illness is associated with an increased risk of psychosocial adjustment difficulties for the child and his or her parents and siblings.

However, good adjustment is the norm. This variability in psychosocial adjustment has prompted the investigation of processes associated with risk and resiliency in the face of the adversity of chronic childhood illness (Thompson & Gustafson, 1996).

Not only does chronic childhood illness impact on biomedical and psychosocial functions, but it also brings a number of challenges that have been conceptualized by Moos and Tsu (1977) in their life crises model as adaptive tasks. Three adaptive tasks are illness-related: (i) dealing with the symptoms; (ii) dealing with the treatment procedures and environments; and (iii) developing and maintaining adequate relationships with the health care team. Four adaptive tasks are related to general life crises: (i) preserving reasonable emotional balance by managing feelings and maintaining hope; (ii) preserving self-image, sense of competence, and mastery and readjusting goals and expectations; (iii) preserving relationships with family and friends; and (iv) preparing for an uncertain future.

This delineation of illness tasks and general adaptive tasks has fostered a biopsychosocial approach to psychological adjustment that has included psychosocial perspectives without losing the advantages of the biomedical perspective. Although each chronic illness is characterized by distinct biological processes, symptoms, and specific therapies, the early history of the study of adaptation to chronic childhood illness did not reveal illness-specific patterns of adjustment. There was as much variability in adjustment patterns within illness subgroups as there was between illness subgroups. This led to a focus on commonalities confronting children and their families who are dealing with different chronic illnesses. For example, Pless and Pinkerton (1975) identified the chronicity of the illness and the impact on the family as factors likely to be more important for adjustment than specific characteristics of their illness. This emphasis on similarities across chronic illnesses became known as the noncategorical approach. Subsequently, Stein and Jessop (1982) suggested dimensions of particular importance to psychological adjustment, including whether the condition is visible or invisible, life-threatening or not, stable or unpredictable, involves the central nervous, sensory, or motor systems, or requires an intrusive or demanding routine of care.

In the efforts to foster adaptation, there are advantages to both the categorical, that is, illness specific, and noncategorical approaches to chronic childhood illness. What is needed are conceptual frameworks that delineate: (i) meaningful biomedical parameters within and across specific illness that elicit "significantly distinct psychosocial demands for the ill individual and his or her family" (Rolland, 1987, p. 204); and (ii) the specific pattern of individual and family adaptational processes that match or best fit these demands. The advent of systems theory perspectives in general, and family systems theories in particular, is an important tool in the effort to achieve this "goodness of fit."

9.14.2.2 Prevalence

Because of the large number and relative rarity of chronic illnesses in childhood, it is difficult to estimate prevalence. Estimates have been based on several epidemiological studies on children's health and development conducted in the USA and the UK (see Cadman, Boyle, Szatmari, & Offord, 1987; Gortmaker & Sappenfield, 1984; Pless & Douglas, 1971; Pless & Roghmann, 1971; Roghmann & Haggerty, 1970). Based on these data, it is estimated that one million children in the USA have a severe chronic illness that may impact their daily functioning (Perrin & MacLean, 1988a) and an additional 10 million children suffer from a less serious chronic disorder (Haggerty, 1984).

Overall, 10–20% of children have a chronic disorder (Gortmaker & Sappenfield, 1984) and these children comprise approximately 50% of pediatric practice (Magrab & Calcagno, 1978). In terms of specific illnesses, asthma and other allergic disorders ranked first, sensory disorders ranked second, and nervous systems disorders ranked third (Pless & Roghmann, 1971).

Whereas there is little evidence for significant change over time in the incidence of chronic childhood illness, the prevalence has increased. It has been estimated that from the 1960s to the 1980s there was a doubling from 1% to 2% of children with a severe chronic health condition (Gortmaker, 1985). There are three factors contributing to this increase in prevalence. First, advances in health care have increased the life expectancy and survival rates for many illnesses. For example, the median survival age for children with cystic fibrosis increased 2.5 fold, from 11 years in 1966 to 29.4 years in 1993 (Cystic Fibrosis Foundation, 1994). A second factor contributing to increased prevalence are new survivors. With advances in neonatal intensive care since the 1950s, there has been a progressive decline in mortality for infants born prematurely at very low birth weight ($<1\,500$ g) from 72% to 27% (McCormick, 1989). However, these new survivors have increased morbidity for respiratory, neurodevelopmental, behavioral, and learning problems. A third factor is new conditions, reflected in infants with prenatal drug exposure and children with AIDS who also have increased neurodevelopmental morbidity. Thus, children and their families are, in increasing numbers, confronted with the task of adapting to chronic illness over a substantial period of their lives.

9.14.2.3 Assessment

Assessment issues regarding adaptation to chronic childhood illness have two foci: psychosocial adjustment and the correlates, that is, processes or factors, that contribute to or account for differences in psychosocial adjustment. The term adjustment implies a broad range of functioning that includes the clinical/maladaptive level, and temporal and situational variability (Wallander & Thompson, 1995). Psychosocial adjustment is an umbrella term that encompasses three dimensions: psychological adjustment as reflected in emotional/behavioral problems/disorders and self-esteem; social adjustment and peer relationships; and school performance. It is useful to adopt a developmental-normative definition of adjustment in which behavior is assessed in terms of age appropriateness and trajectory toward positive adult functioning (Wallander & Thompson, 1995).

Because psychological adjustment has been conceptualized and operationalized in numerous ways, methods for assessing children's adjustment have been diverse and include checklists, questionnaires, and diagnostic clinical interviews completed by children or by parents, teachers, or clinicians about children. Classification systems for children's adjustment problems also vary by assessment methods, such as broad-band (e.g., internalizing and externalizing) and narrow-band (e.g., somatization or anxiety) behavior dimensions of the Child Behavior Checklist (Achenbach & Edelbrock, 1983) and the psychiatric diagnoses of the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 1994). Research indicates that there is a range in level of agreement among different informants and alternative methods as a function of age of the child, symptom type, and mental state of the informant (see Thompson & Gustafson, 1996). Evidence indicates that comprehensive assessment of the psychological adjustment of children requires multiple informants and measures (Thompson & Gustafson, 1996).

9.14.2.4 Conceptual Models

Research on adaptation to chronic childhood illness has reached the analytic, theory-driven stage and several conceptual models have emerged. These models have in common the conceptualization of chronic illness as a potential stressor and the conceptualization of adaptation in terms of social-ecological systems. Each model presents variables and processes that are hypothesized to affect adaptation to the stress of chronic illness. Models serve the heuristic value of integrating findings across diverse studies and guiding subsequent investigative studies and initial intervention efforts to enhance adaptation.

An early conceptual model, formulated by Pless and Pinkerton (1975), viewed adjustment as multiply determined by the individual's transactions with his or her environment. In addition to conceptualizing chronic illness as a stressor, seminal contributions of this model to subsequent efforts included viewing adaptation as a continuous process over time and the emphasis on the importance of coping methods in the process of adaptation.

The life crisis model developed by Moos and Tsu (1977) emphasized the role of individual cognitive processes in appraisal of the significance of the crisis, perceptions of tasks involved, and the selection of coping skills. This model postulates that these processes are influenced by three factors: background and personal characteristics; illness-related factors; and features of the physical and sociocultural environment.

In the disability–stress coping model developed by Wallander, Varni, Babani, Banis, and Wilcox (1989), factors hypothesized to play a role in adjustment to chronic illness are organized into a risk-resistance framework. The risk of psychosocial problems is elevated by stress emanating from disease parameters and associated functional limitations and from major life events, daily hassles, or transitions in the course of children's lives. Resistance factors include: intrapersonal factors of temperament, competence, effectance motivation, and problem-solving ability; stress processing factors of cognitive appraisal and coping behaviors; and social–ecological factors such as the family's psychosocial and practical resources.

The transactional stress and coping model, developed by Thompson and Gustafson (1996), hypothesizes that the illness–outcome relationship is a function of the transactions among biomedical parameters that reflect the type and severity of illness; demographic parameters of age, sex, and socioeconomic status (SES); and psychosocial processes including appraisal of illness-related and daily stress, expectations of efficacy and locus of control, coping methods, and family functioning. Psychological adjustment is considered in terms of maternal and child adjustment and their interrelationship.

In the following sections, what is known about the psychological and social adjustment of children with chronic illness and their siblings, and the psychological adjustment of their parents is summarized. Within each section, information on the correlates of adjustment is summarized in terms of the illness condition, individual, and social-ecological dimensions of concurrent conceptual models.

9.14.2.4.1 Psychological adjustment of children with chronic illness

Several hundred studies have addressed the psychological adjustment of children with chronic illness. Reviews available include Wallander and Thompson (1995), Thompson and Gustafson (1996), and a meta-analysis (Lavigne & Faier-Routman, 1993). Two research strategies have been employed: epidemiological surveys and clinical studies.

A review of three epidemiological surveys conducted in the late 1960s revealed that children with chronic medical conditions demonstrate a higher proportion of emotional and social problems than their healthy peers, with approximately 30% expected to be impaired by secondary psychological maladjustment by 15 years of age (Pless & Roghmann, 1971). A later epidemiological study reported similar findings, with 33% of children with a chronic physical condition diagnosed with at least one *DSM-III* disorder (Cadman et al., 1987). Children with chronic illness with and without disability had a risk of psychiatric disorder that was 3.4 times and 2.1 times that of healthy children, respectively. Children with chronic illness and disability were also at increased risk of social adjustment and school difficulties.

Studies of adjustment of children with chronic illness, using samples obtained in clinical settings, have suffered from methodological problems. However, studies have been conducted across a number of specific chronic illnesses that have used psychometrically sound assessment methods with samples of patients with a single illness or pooled illnesses compared with healthy or psychiatrically impaired groups or with normative data. Several conclusions can be reached: children with chronic illness are at increased risk of behavioral and emotional problems; internalizing, anxiety-based disorders are most frequent, although milder forms of externalizing problems such as oppositional disorder are evident in some illness and age groups; children with disorders that affect the brain may be at particular risk of adjustment difficulties; and stability in adjustment is only moderate with considerable change across time at the individual level in the presence or absence of a behavioral problem or diagnosis, particularly in terms of specific diagnoses or behavioral patterns.

Overall, based on epidemiological and clinical studies, it is estimated that children with chronic illness have a risk of adjustment problems that is 1.5–3 times as high as their healthy peers (Pless, 1984). However, major psychiatric disturbance is not common and good adjustment is possible. Thus, there is no simple or direct relationship between chronic childhood illness and psychosocial adjustment (Wallander & Thompson, 1995). This has led to a search for the processes and factors associated with good and poor adjustment.

There is substantial literature on the processes and factors that mediate or moderate the psychosocial adjustment of children to the stress of chronic childhood illness. The search for correlates of psychological adjustment is being guided by conceptual models. An organizational framework comprised of condition, child, and social-ecological perspectives has provided a useful framework for integrating findings across studies (Lavigne & Faier-Routman, 1993; Thompson & Gustafson, 1996; Wallander & Thompson, 1995).

Condition parameters have included condition type, severity, functional status, prognosis, and duration. Children with conditions involving the brain have more behavior problems and poorer social functioning than children with conditions not involving the brain among which no significant differences have generally been found. The findings of the meta-analysis (Lavigne & Faier-Routman, 1993) provide support for the contribution of condition severity, but numerous studies do not. There is emerging evidence for the relationship of functional status and illness duration with adjustment.

Child parameters have included age, gender, temperament, coping methods, and cognitive processes. Whereas most studies report no significant difference in parent and/or teacher reported behavior problems as a function of gender, where differences were found, boys were reported to have more problems. In terms of self-report, girls report more symptoms of distress than boys. Age effects on behavior/emotional problems and self-esteem have generally not been found but the evidence regarding age of onset is mixed. There is some support for the relationship of temperamental difficulty and mother-reported behavior problems. Children's pain coping strategies characterized by negative thinking have been associated with child-reported psychological symptoms, as have been an array of cognitive processes including high levels of perceived stress and negative perceptions of physical appearance and self-esteem. However, the findings regarding perceptions of health locus of control have been mixed.

Social-ecological parameters have included life stress, SES, parental adjustment, marital adjustment, and family functioning. SES has been measured inadequately and used infrequently and the findings are inconsistent. Child adjustment is positively related to maternal adjustment. Marital adjustment has been investigated infrequently but there is strong support for the role of family functioning in child psychological adjustment. In general, high levels of cohesion and supportiveness and low levels of conflict are associated with better adjustment.

In spite of efforts, the knowledge base is limited regarding the correlates of psychological adjustment of children with chronic illness. The focal points have been condition parameters and family functioning. The interrelationship among condition, child, and social-ecological parameters needs to be investigated longitudinally. To delineate salient intervention targets, there is a particular need for conceptually- and theoretically-driven studies.

9.14.2.5 Psychological Adjustment of Siblings

Research on the siblings of children with medical conditions has focused on the negative impact of the child with a handicap on the healthy siblings (Senapati & Hayes, 1988). This is consistent with the commonly held belief that children with a chronic condition have a harmful effect on healthy siblings that is reflected in healthy siblings having a high rate of adjustment problems. This belief in harmful effects is based on considerations of the impact of increased parental stress associated with caring for the disabled child, decreased parental attention to the healthy child, decreased family resources, and pathological identification with the child with a disability (Dyson, 1989). However, the number of empirical studies has been relatively small (see review by Drotar & Crawford, 1985; Lobato, Faust, & Spirito, 1988).

In general, there is wide variability in the findings on the frequency of sibling adjustment problems, with some studies reporting adverse psychological outcomes, no significant difference between siblings of children with disabilities and siblings of healthy children, and benefits to siblings of children with disabilities. The one extant longitudinal study reported an increase in depressive affect and social isolation over a five-year period in siblings of children with disabilities compared to healthy controls (Breslau & Prabucki, 1987). These diverse findings have stimulated efforts to delineate the correlates of good and poor sibling adjustment (Thompson & Gustafson, 1996).

Studies have not demonstrated a strong relationship between illness or disability and adverse psychological outcome among siblings (Thompson & Gustafson, 1996). Thus, investigations of the correlates of sibling adjustment have focused on sibling parameters including gender, birth order, and age spacing. In general, increased risk of sibling adjustment problems is associated with closer age spacing and older sisters and younger brothers in relation to the child with a disability (Lobato et al., 1988). Studies are investigating more dynamic processes such as the quality of sibling relations, especially sibling conflict (Hanson et al., 1992) and the impact of chronic illness on siblings' early functioning and interactions (McHale & Gamble, 1989). Again, to delineate specific processes likely to mediate or moderate sibling adjustment, theoretically- and conceptually-driven studies are necessary.

9.14.2.6 Psychological Adjustment of Parents

Psychological adjustment of parents is of concern as one of the impacts of chronic childhood illness because of the previously discussed role of parental adjustment in children's adjustment (Thompson & Gustafson, 1996). Assessment methods have been predominately self-report measures of psychological distress in terms of overall adjustment and specific symptom dimensions, particularly depression and anxiety. Studies have predominately been of mothers. The primary question addressed concerns the frequencies and types of adjustment problems. The search for correlates of parental adjustment is in its early stage.

One epidemiological survey, the Ontario Child Health Study, included an assessment of parental adjustment in terms of affective balance in families with children with health problems and families of well children (Cadman, Rosenbaum, Boyle, & Offord, 1991). The findings indicated significantly higher maternal negative affect scores among parents of children with a chronic problem and rates of parent-reported mental health treatment for nerves among mothers (29%) and fathers (8%) that was 2–3 times higher than that reported by parents of well children. However, there were no significant differences in other measures including social isolation and alcohol problems. It was concluded that "contrary to some clinical opinion, we found no evidence in this general community sample to support the belief that families of children with chronic health problems function any differently than families in the general population" (Cadman et al., 1991, p. 886).

Clinical studies of single or pooled illnesses are the basis for most of what is known about parental adjustment. The results of parent studies that compare self-reported adjustment of chronically ill children with that of healthy children in terms of mean scores on a number of different assessment measures have yielded mixed results. Some studies have reported increased, but subclinical, levels of symptoms, compared to normative data.

A research program at Duke University has addressed the questions of type and frequency of adjustment problems in mothers of children with chronic illness and changes in adjustment over time (see Thompson & Gustafson, 1996). Psychological adjustment has been assessed with the SCL-90-R (Derogates, 1977), which is a 90-item self-report measure of psychological distress along symptom dimensions including depression, anxiety, and hostility. The Global Severity Index combines information on numbers of symptoms and intensity of distress and is used as a measure of the overall level of psychological distress. Using nonpatient norms for females, T-scores greater than or equal to 63 (i.e., above the ninetieth percentile) are considered in the clinical range. Poor psychological adjustment is based on criteria for "caseness," T-scores > 63 on the Global Severity Index or any two of the nine symptoms dimensions. Across a number of separate studies of specific illness subgroups, the rates of poor maternal adjustment were relatively high: Duchenne muscular dystrophy (57%), spina bifida (44%), cystic fibrosis (34%), and sickle cell disease (36%). There was moderate stability in the individual classification of adjustment of mothers of children with cystic fibrosis and sickle cell disease across an 8–19 month period. A change in classification (between good and poor adjustment) was found in 21% and 25% of the mothers of children with cystic fibrosis and sickle cell disease, respectively.

Although there have been few studies, the biopsychosocial framework is also useful for integrating findings regarding the correlates of parental psychological adjustment. In terms of illness parameters, maternal adjustment is generally found not to be related significantly to the child's medical status (Kupst & Schulman, 1988) or to the degree of the child's physical impairment (Wallander et al., 1989). However, strong support has been provided for the relationship of intrapersonal parameters, in terms of maternal psychological processes, and social–ecological parameters to maternal adjustment.

Research guided by the transactional stress and coping model in the Duke University studies has demonstrated that maternal adjustment is associated with lower levels of appraisals of daily stress, less use of palliative methods of coping in general and in relation to adaptive methods of coping, and higher levels of family supportiveness and lower levels of family conflict (Thompson & Gustafson, 1996).

Research guided by the disability–stress coping model (Wallander et al., 1989) has provided support for the role of utilitarian family resources (i.e., structural characteristics of the family) and psychosocial family resources (i.e., the family's characteristic mode of perceiving and interacting with the social world) in maternal adjustment. Support has also been provided for the contribution of family life stress to maternal distress (Mullins et al., 1991).

9.14.2.7 Treatment

The biopsychosocial framework for adaptation to chronic childhood illness has implications for intervention. Health care services are directed at two outcomes: extending the duration of life and improving the quality of life (Kaplan & Anderson, 1988). With regard to quality of life, the focus of intervention is the adaptation of all children with chronic illness and their families, not just the relative few who experience psychological crises (Drotar, 1981). Thus, intervention includes treatment and prevention.

Stein and Jessop (1984) have articulated that the goals of care are to

> confine the consequences of the biologic disorder to its minimum manifestations, to encourage normal growth and development, to assist the child in maximizing potential in all possible areas, and to prevent or diminish the behavioral and social consequences. (pp. 193–194)

In the situation of chronic childhood illness, both biological and psychosocial processes are mediators of quality of life. One way to improve outcome is to modify health behaviors to enhance functioning. In this regard, adherence to medical regimens and pain management are important. Another way to improve outcome is to limit the disruptive impact of the illness on the normal processes of child development and family functioning (Perrin & MacLean, 1988b). Important in this regard are stress reduction and coping skills and social support enhancement. In general, the research on intervention is based on cognitive-behavior–social learning theory approaches to enhancing adaptation and have focused on three targets:

psychosocial adjustment, adherence, and pain management.

9.14.2.7.1 Psychosocial adjustment

Because the normal development of children is influenced by the social and academic progression that occurs in school (Sexson & Madan-Swain, 1993), preventing problems in these areas is integral to optimal adaptation to illness. With the onset of chronic illness, some children experience cognitive impairment, health-related absenteeism, and disruption of peer relationships. In response, multicomponent school reintegration programs have been developed and have been shown to enhance social competence and decrease behavioral adjustment problems (Katz, Rubenstein, Hubert, & Blew, 1988; Katz, Varni, Rubenstein, Blew, & Hubert, 1992).

Research has demonstrated the importance of peer support over and above family support in the adjustment of children with chronic illness (Varni, Katz, Colegrove, & Dolgin, 1994; Wallander & Varni, 1989). Consequently, efforts have been directed to developing methods to improve children's social skills and peer relationships through social skills and cognitive problem-solving training. Representative of these efforts is the behavioral group approach for social skills training of elementary school children developed by La Greca and Santogrossi (1980) and the social skills training modules developed by Varni, Katz, Colegrove, and Dolgin (1993).

Parenting processes have also been identified as salient intervention targets because of the firmly established contributions of parenting to behavioral adjustment and cognitive and social development (Thompson & Gustafson, 1996). A number of specific parent training programs have been developed and have been shown to be effective in teaching parents procedures for interacting with their children. With regard to behavioral adjustment, the program developed by Forehand (1993) is exemplary. With regard to cognitive and social development, The Partners for Learning Curriculum developed by Sparling and Lewis (1984) has been used successfully in early intervention efforts, such as the Infant Health and Development Program for low birth weight infants (Gross, 1990). The efforts to foster developmentally conducive parent–child interactions are being applied to the management of illness tasks, such as the multifamily group intervention program for children with diabetes and their families, that focuses on improving metabolic status and psychosocial functioning (Satin, La Greca, Zigo, & Skyler, 1989).

9.14.2.7.2 Adherence

Noncompliance with medical regimens is relatively common. The median adherence rate for pediatric populations has been estimated to be 50% (Dunbar-Jacob, Dunning, & Dwyer, 1993). The biopsychosocial framework identifies factors affecting adherence (Thompson & Gustafson, 1996). Illness parameters are reflected primarily in treatment regimens. Treatments that are more complex produce negative side effects, interfere with normal activities, are oriented toward the future, and have a negative impact on adherence. In terms of individual processes, adherence is related positively to cognitive and social/emotional developmental level and coping skills. In terms of social-ecological parameters, parenting, maternal functioning, and family functioning are related positively to adherence.

Given the less than one-to-one relationship of treatment to outcome and the complexity of the interrelationships of illness tasks with developmental life course issues in quality of family life, views regarding compliance are changing. There has been a movement away from viewing noncompliance as an indicator of irresponsibility or forgetfulness to a more considered view in which noncompliance can be a well-reasoned, adaptive choice (Deaton, 1985). Haynes (1976) has classified strategies for improving compliance into three broad categories: educational, behavioral, and combined. Educational strategies provide information. Behavioral strategies focus on changing behavior involved in compliance through stimulus control techniques, self-monitoring of symptoms and self-regulation of medications, and reinforcement of symptom reduction, medication use, and health contacts. The evidence for the effectiveness of behavioral strategies is strong (see Epstein & Cluss, 1982). Corresponding to the change in the view of adherence, there is an increasing emphasis on the process of tailoring a consensual regimen to the specific needs and life situation priorities and realities of the family (Thompson & Gustafson, 1996).

9.14.2.8 Pain Management

Pain is a multidimensional experience. It involves a sensory component, nociception, and the psychologic, physiologic, emotional, and behavioral responses to the nociception (McGrath et al., 1990). Approaches to pain management involve analgesics and behavioral intervention focused on cognition and socio-environmental processes that influence the perception, interpretation, and behavioral responses to pain.

Children with chronic illness may face two types of pain: acute pain associated with invasive medical procedures and chronic and recurrent pain associated with illnesses such as sickle cell disease and arthritis. The efficacy of the cognitive-behavioral approach for both types of pain is well established (Thompson & Gustafson, 1996). These approaches are multicomponent and focus on pain perception regulation and pain behavior modification. Pain perception regulation includes the self-regulatory techniques of deep breathing, muscle relaxation, imagery and behavioral rehearsal; distraction; and active coping attempts. Pain behavior techniques identify and modify socio-environmental factors that influence behavioral expressions of pain (for details regarding these techniques see Dahlquist, 1992; Jay, 1988; and Varni, Walco, & Wilcox, 1990). The role of the family environment in the reinforcement of pain-coping behaviors has been recognized and these multicomponent pain management programs have parent-training components as well as components to teach children pain-coping skills. For example, the efficacy of cognitive-behavioral family intervention has been demonstrated for children with recurrent abdominal pain (Sanders et al., 1989; Sanders, Shepherd, Cleghorn, & Woolford, 1994).

9.14.3 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

The next steps in the process of understanding and enhancing adaptation to chronic childhood illness involve further refinements at the analytic level of model development and movement to the experimental level of controlled interventions that would serve to increase quality of life and confirm relationships of processes to outcome. Although much remains to be learned, the next steps in the process must intentionally build on, refine, and extend the current knowledge base. In particular, "there is a compelling need to determine how psychosocial adjustment, mediating and moderating adaptational processes, and their interrelationship change over time relative to individual and family developmental phases and illness course" (Thompson & Gustafson, 1996, p. 20).

The experience gained from current models is that adaptation is a multicomponent process in which a large array of factors and processes account for relatively small portions of variance in psychological adjustment. Thus, one way in which advances could occur is by increasing the array. It is also clear that the interrelationship of processes and factors is important as reflected in the findings regarding goodness of fit, for example, between methods of coping and situational and adaptational processes. Thus, another way in which advances could occur is by delineating goodness of fit patterns among biomedical, psychosocial, and developmental processes. Goodness of fit patterns offer the potential for identifying at-risk and resilient families and salient intervention targets for treatment and prevention efforts.

In considering both extending the array and goodness of fit, two processes are ripe for inclusion within a stress and coping framework: parenting and family system functioning. Both of these processes are within the social-ecological domain and are areas of theoretical and empirical activity in and of themselves. Inclusion of these processes would serve to enhance the sophistication of constructs and measures within the social-ecological domain of biopsychosocial models of adaptation to chronic childhood illness and expand the targets for intervention.

9.14.3.1 Parenting

The impact of parents' child rearing strategies and behavior, that is parenting, on the child's cognitive and psychosocial development has been clearly established. "Parenting that is *sensitively* attuned to children's capabilities and to the developmental tasks they face promotes a variety of highly valued developmental outcomes" (Belsky, Lerner, & Spanier, 1984, p. 85). Thus, there has been a search for factors that influence parenting.

Belsky (1984) has formulated a process model of parenting that includes three domains that influence parenting: parent personality, child characteristics, and contextual sources of stress and support. Other lines of research have delineated other parent factors that affect parenting. Parental cognitions, in terms of beliefs and causal attributions regarding child development and behavior, have been shown to influence parents' proactive and reactive responses to their children and their children's development (see Dix, Ruble, & Zambarano, 1989; Miller, 1988). Linkages among parental stress and distress, child-rearing practices, and children's cognitive and social development have also been established. "Greater stress is associated with less optimal parent and family functioning, less optimal parent–child interactions, and lower child developmental competence" (Crnic & Greenberg, 1990, p. 1628). Maternal psychological distress has been shown to influence maternal perceptions of, and behavioral interactions with, her children (Webster-Stratton, 1990). For example, maternal depression has been linked to increased levels of vague or interrupting commands to which children cannot comply (Forehand, Lautenschlager, Faust, & Graziano, 1986).

An integration of these lines of research into a cognitive and stress processing model of parenting has been suggested previously (Thompson & Gustafson, 1996). The model is based on the premise that if parenting is the primary mechanism for reflecting the social-ecological effects on children's development, then enhancing the goodness of fit of parenting with the child's special needs and characteristics would promote development and adaptation. In turn, parents' psychological functioning, in terms of stress and distress, and parent child-rearing knowledge, beliefs, and attributions, are major influences on parenting. These processes become salient targets for intervention programs that combine educational and stress-processing components. In these efforts, basic programs to enhance parent effectiveness can be modified to address the illness-specific and general adaptative tasks associated with chronic childhood illness.

9.14.3.2 Family Systems Functioning

Family functioning has been one of the most frequently investigated correlates of psychological adjustment of children with chronic illness and their parents. Conceptual models afford family functioning a prominent role. However, for the most part, family functioning has been conceptualized and measured in terms of various dimensions such as cohesion, experiences, and conflict. The opportunity exists to go beyond such conceptualizations to employ the richer conceptualizations of family dynamics captured by family system theories. Although rich, these theories have been difficult to operationalize and evaluate empirically. However, with the formulation of the Family System Health Model, a basis for bringing family system theory approaches to bear in adaptation to chronic childhood illness is provided (Rolland, 1984, 1987, 1993).

Rolland's (1984, 1987, 1993) conceptual model of the adaptation of families confronting chronic and life-threatening disorders is based on the premise that illness and family processes interact in a constantly evolving fashion. Categorizing illnesses according to their psychosocial demands on the family system provides a common link between illness and family processes. The goodness of fit between the psychosocial demands of the disorder and the family's style of functioning and resources is viewed as a prime determinant of successful vs. problematic family coping and adaptation (Rolland, 1993). Within this model, goodness of fit is determined on the basis of a three-way congruence between the pull from the psychosocial demands of the illness, the pull from the demands of the family's life cycle stage, and the style of family functioning. Pull is conceptualized as either the need for family members to focus their energies inward and work together to perform family developmental tasks, termed centripetal pull, or the need for the family to work to support the development of individual members outside of the family system, termed centrifugal pull. Better family member adjustment is expected if the style of family functioning required (or pulled) by the family's current life cycle stage also fits that required by the illness.

Rolland's (1993) model provides a detailed explanation of how to delineate the psychosocial demands of the illness and the family life cycle and how to relate those demands to the family's life structure and style of functioning. One advantage of Rolland's (1993) model is its comprehensiveness, but his model and its components have not yet been operationalized. Thus, an effort will be made to illustrate potential ways to operationalize key constructs so that these can be investigated empirically.

9.14.3.3 The Psychosocial Demands of Illness

Rolland (1993) suggested that all illnesses elicit some degree of centripetal pull on the family system. However, the degree of pull varies as a function of the illness' psychosocial demands which arise out of the interaction between the biomedical parameters and the illness phase of the specific condition.

Biomedical parameters which place psychosocial demands on the family system include onset (acute or chronic), course (constant, progressive, or relapsing), outcome (fatal, nonfatal, or potentially fatal), and degree of incapacitation (incapacitating or nonincapacitating). These parameters have been combined to form a grid which yields a psychosocial typology comprising 32 illness types (Rolland, 1993). Illness types range from those that are acute in onset, constant, incapacitating, but nonfatal, such as congenital malformations, to those that are acute in onset, relapsing, nonincapacitating, and possibly fatal, such as sickle cell disease, to those that are gradual in onset, progressive, increasingly incapacitating, and fatal, such as cystic fibrosis. The tendency for an illness to elicit centripetal pull increases with the level of incapacitation or risk of progression and/or death (Rolland, 1993). For example, progressive illnesses elicit a stronger centripetal pull over time than illnesses with a constant course. The continual addition of new demands associated with progressive conditions keep the family's energy perpetually focused inward, often hindering the family's advancement through the life course developmental process (Rolland, 1993). In contrast, illnesses with a constant course (that are not severely incapacitating) require an initial period of readjustment, but then allow the family to regain its normal style of functioning.

There are two disease phases of importance to psychosocial demands and the corresponding degree of centripetal pull on the family system: crisis and chronic. Crisis phases are periods marked by the onset or significant exacerbation of disease symptoms. The diagnostic and terminal periods of the illness process are considered crisis phases. The diagnostic period ranges from the period in which the family first notices illness symptoms to the initial period of readjustment after the diagnosis. Family psychosocial tasks associated with this period include beginning the process of coping with the illness symptoms, dealing with the treatment regimens, and becoming familiar with the hospital staff and environment.

The terminal period of the illness process is the period during which the inevitability of the patient's death becomes apparent. During this period, the family must cope with issues of separation, death, grief, and resumption of normal life. Family members must shift their focus from trying to control the illness to coping with the process of letting go. Further, the family must remain open emotionally in order to deal with unfinished business, cope with practical tasks, and say their final goodbyes. Crisis phases, whether they are the diagnostic or terminal periods of the illness process or simply periods of symptom exacerbation during the chronic phase, typically place demands on the family system which require family integration and diffuse boundaries between family members. These phases are therefore viewed as eliciting a centripetal pull on the family system (Rolland, 1993).

The chronic phase is the asymptomatic period or period of baseline illness symptoms that occurs between the family's initial readjustment after the diagnosis and the period in which symptoms of terminal disease first appear. This phase is variable in length and, depending on the illness type, is marked either by constancy, progressive decline, or episodic change. Further, this period can be disrupted by crisis phases whenever there is an acute increase in illness symptoms. Within this period, the family must learn to live day-to-day with the chronic condition, while maintaining as normal a life as possible. It is no longer critical that the family focuses all of its energy on the illness process. Family members can, instead, begin to concentrate on their own individual development as well as the family's development as a system. The demands associated with the chronic phase are therefore thought to elicit a weak centripetal pull on the family system (Rolland, 1993).

The psychosocial demands associated with biomedical parameters and illness phases are combined to determine the degree of centripetal pull on the family system. For example, a severely incapacitating and/or fatal disease in the crisis phase of the illness process is thought to elicit a strong centripetal pull on the family system, whereas a nonincapacitating and nonfatal condition in the chronic phase of the illness process is thought to allow more individuation among family members.

It is hypothesized that better family member adjustment is expected if the family's style of functioning (e.g., high or low cohesion) fits the demands of the illness (i.e., strong or weak centripetal pull). For example, since the demands associated with the onset or crisis phase of an illness elicit a strong centripetal pull, it is hypothesized that high family cohesion during these periods leads to better family member adjustment. This hypothesis could be tested empirically by comparing the relationship between high family cohesion and child adjustment during the crisis phase of an illness to the relationship between high family cohesion and child adjustment during the chronic phase of an illness that is not progressive, incapacitating, or fatal.

9.14.3.3.1 The psychosocial demands of the family life cycle

The life cycle is a key determinant of the psychosocial demands on the family system. The family life cycle refers to the basic sequence and development of the life course within which family differences occur. The life cycle can be divided into a series of stages that progress from the stage during which the family bears children through the stage during which the family has adolescents to the stage during which the family launches young adults. The terms centripetal and centrifugal pull are also useful in conceptualizing the psychosocial demands of different family life cycle stages. The developmental tasks of the family life cycle shift from demanding interdependence and family teamwork (centripetal pull) during early stages of the life cycle, such as the childbearing stage, to emphasizing individuation and personal autonomy (centrifugal pull) among family members during later stages of the life cycle, such as the stage in which the family launches young adults (Rolland, 1993). In this way, stages of the life cycle can be defined in terms of the degree to which psychosocial demands reflect the relative need for family members to channel their energies and work together to perform those tasks.

It is hypothesized that better family adjustment is expected if the family's style of functioning fits the demands of the family's current life cycle stage (i.e., centripetal or centrifugal). For example, since the demands of the early stages of the life cycle require interdependence among family members, it is hypothesized that high family cohesion during these periods leads to better family member outcomes. This hypothesis could be tested empirically by comparing the relationship between high family cohesion, which can be assessed using measures such as the Family Environment Scale (Moos & Moos, 1981) or the Family Adaptability and Cohesion Evaluation Scales III (Olson, Portner, & Lavee, 1985), and child adjustment during an early stage within the life cycle to the relationship between high family cohesion and child adjustment during a late stage within the life cycle.

9.14.3.3.2 The interface between the psychosocial demands of the illness and the family life cycle

Combination of the key concepts regarding the family life cycle and the illness process provides a foundation for understanding family adaptation to chronic illness. Rolland (1993) proposed that the illness process and family life cycle interact, each placing their own demands on the family system. The goodness of fit between the pull from psychosocial demands of the illness and the pull from the developmental demands of the family's life cycle stage moderates family adaptation (Rolland, 1993). Better family member adjustment is expected if the pull of the demands associated with the family's life cycle stage is complementary, rather than conflicting, with the pull of the demands related to illness phase (Rolland, 1993). For example, if the crisis phase of an illness (which elicits a strong centripetal pull on the family) interfaces with the early childrearing years of the family life cycle (which also elicits a centripetal pull), better family member adjustment would be expected than if the crisis phase occurred during a centrifugal stage of family development (such as the period during which the family launches young adults). Specifically, illnesses that elicit a strong centripetal pull, those that are progressive, relapsing, severely incapacitating, and/or fatal, are thought to be particularly problematic to the family system if their onset coincides with a later stage of the family life cycle since the pull of the demands of the illness does not fit the pull of the demands associated with the family's life cycle stage (Rolland, 1993).

These hypotheses can be tested empirically by comparing child adjustment during a period in which a fit is thought to exist between the pull of psychosocial demands of the illness and that of the family life cycle demands (e.g., when the onset of a progressive, incapacitating, and/or fatal illness occurs during a stage early in the family life cycle) to child adjustment during a period in which a fit between demands is not thought to exist (e.g., when the onset of a progressive, incapacitating, and/or fatal illness occurs during a stage late in the family life cycle). It is hypothesized that child adjustment would be better in the fit rather than the nonfit condition.

Based on the interaction between the pull of illness and family life cycle demands, it is also hypothesized that a goodness of fit exists if the style of family functioning required for the family to deal appropriately with the psychosocial demands of the illness is congruent with the style of family functioning needed to meet the demands of the family's current life cycle stage. Specifically, high family cohesion would be hypothesized to fit both illness and family demands during periods in which both the family's life cycle stage and illness process elicit a strong centripetal pull on the family system. Since a fit exists between these variables within this situation, better family member adjustment would be expected than if high family cohesion did not fit illness and family demands. The model is not limited to cohesion as the measure of family style. Other aspects of family functioning such as control, expressiveness, or support could be hypothesized as a fit between illness and family processes during different periods of illness and family development.

It is the eventual hope that the information yielded from the operationalization of Rolland's (1993) model will enable investigators and clinicians to categorize families facing pediatric chronic illness as high or low risk for adjustment problems based on this three-way goodness of fit schema. Intervention efforts, appropriate to illness and family developmental demands, could then be developed and provided to at-risk families to enhance adaptation and prevent adjustment problems.

9.14.4 SUMMARY

Chronic illness impacts on an increasing number of children and their families and is a significant stressor with an increased risk for psychological adjustment difficulties. However, most children and families adapt well to the stresses and tasks associated with chronic childhood illness and good adjustment is the norm. Research on adaptation to chronic childhood illness has progressed to the analytic level of conceptually- and theoretically-driven studies that seek to delineate the biomedical, intrapersonal, and social-ecological processes associated with varying degrees of adjustment. The findings from these analytic level of studies have provided a foundation for moving to the experimental level of controlled intervention studies to enhance adaptation and quality of life. More specifically, improved stress processing is a salient target for interventions designed to enhance coping skills and foster social support.

In these efforts, it is recognized that adaptation is a multidimensional process. There is no one right way and the emphasis is upon achieving a good fit among illness demands and intrapersonal and social-ecological processes to enhance quality of life in general, and psychological adjustment, in particular, of all family members. While these intervention studies are underway, the proposed expansion of models to incorporate parenting processes and family system functioning will provide the foundation for the next wave of intervention targets. This expansion within the social-ecological domain of adaptative processes is particularly necessary to enable movement to another level of goodness of fit. That is, intervention programs can be tailored to diverse cultural beliefs and practices regarding health, illness, child development, and family functioning.

9.14.5 REFERENCES

Achenbach, T., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, VT: Queen City Printers.

American Psychiatric Association (APA) (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: APA.

Belsky, J. (1984). The determinants of parenting: A process model. *Child Development*, *55*, 83–96.

Belsky J., Lerner, R., & Spanier, G. (1984). *The child in the family*. Reading, MA: Addison-Wesley.

Breslau, N., & Prabucki, K. (1987). Siblings of disabled children. *Archives of General Psychiatry*, *44*, 1040–1046.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, *32*, 513–531.

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.

Cadman, D., Boyle, M., Szatmari, P., & Offord, D. R. (1987). Chronic illness, disability, and mental and social well-being: Findings of the Ontario Child Health Study. *Pediatrics*, *79*, 805–813.

Cadman, D., Rosenbaum, P., Boyle, M., & Offord, D. R. (1991). Children with chronic illness: Family and parent demographic characteristics and psychosocial adjustment. *Pediatrics*, *87*, 884–889.

Crnic, K. A., & Greenberg, M. T. (1990). Minor parenting stress with young children. *Child Development*, *61*, 1628–1637.

Cystic Fibrosis Foundation (CFC) (1994). *Patient Registry 1993 annual data report*. Bethesda, MD: CFC.

Dahlquist, L. M. (1992). Coping with aversive medical treatments. In A. M. La Greca, L. J. Siegel, J. L. Wallander, & C. E. Walker (Eds.), *Stress and coping in child health* (pp. 345–376). New York: Guilford Press.

Deaton, A. V. (1985). Adaptive noncompliance in pediatric asthma: The parent as expert. *Journal of Pediatric Psychology*, *10*, 1–14.

Derogatis, L. R. (1977). *SCL-90-R manual*. Baltimore, MD: Johns Hopkins University, Clinical Psychometrics Research Unit.

Dix, T. H., Ruble, D. N., & Zambarano, R. J. (1989). Mothers' implicit theories of discipline: Child effects, parents effects, and the attribution process. *Child Development*, *60*, 1373–1391.

Drotar, D. (1981). Psychological perspectives in chronic childhood illness. *Journal of Pediatric Psychology*, *6*, 211–228.

Drotar, D., & Crawford P. (1985). Psychological adaptation of siblings of chronically ill children. Research and practice implications. *Journal of Developmental and Behavioral Pediatrics*, *6*, 355–362.

Dunbar-Jacob, J., Dunning, E. J., & Dwyer, K. (1993). Compliance research in pediatric and adolescent populations: Two decades of research. In N. P. Krasnegor, L. Epstein, S. B. Johnson, & S. J. Yaffe (Eds.), *Developmental aspects of health compliance behavior* (pp. 29–51). Hillside, NJ: Erlbaum.

Dyson, L. L. (1989). Adjustment of siblings of handicapped children: A comparison, *Journal of Pediatric Psychology*, *14*, 215–229.

Epstein, L. H., & Cluss, P. A. (1982). A behavioral medicine perspective on adherence to long-term medical regimens. *Journal of Consulting and Clinical Psychology*, *50*, 950–971.

Forehand, R. (1993). Twenty years of research on parenting: Does it have practical implications for clinicians working with parents and children? *The Clinical Psychologist*, *46*, 169–176.

Forehand, R., Lautenschlager, G. L., Faust, J., & Graziano, W. G. (1986). Parent perceptions and parent-child interaction in clinic-referred children: A preliminary investigation of the effects of maternal depressive mood. *Behavior Research and Therapy*, *24*, 73–75.

Gortmaker, S. L. (1985). Demography of chronic childhood diseases. In N. Hobbs & J. M. Perrin (Eds.), *Issues in the care of children with chronic illness: A sourcebook of problems, services, and policies* (pp. 135–154). San Francisco: Jossey-Bass.

Gortmaker, S. L., & Sappenfield, W. (1984). Chronic childhood disorders: Prevalence and impact. *Pediatric Clinics of North America*, *31*, 3–18.

Gross, R. T. (1990). Enhancing the outcomes of low birth weight, premature infants: A multisite, randomized trial. *Journal of the American Medical Association, 263*, 3035–3042.

Haggerty, R. J. (1984). Symposium on chronic disease in children. *Pediatric Clinics of North America, 31*, 1–2.

Hanson, C. L., Henggeler, S. W., Harris, M. A., Cigrang, J. A., Schinkel, A. M., Rodrigue, J. R., & Klesges, R. C. (1992). Contributions of sibling relations to the adaptation of youths with insulin-dependent diabetes mellitus. *Journal of Consulting and Clinical Psychology, 60*, 104–112.

Haynes, R. B. (1976). Strategies for improving compliance: A methodologic analysis and review. In D. L. Sackett & R. B. Haynes (Eds.), *Compliance with therapeutic regi-*

mens (pp. 69–82). Baltimore, MD: Johns Hopkins University Press.

Jay, S. M. (1988). Invasive medical procedures: Psychological intervention and assessment. In D. K. Routh (Ed.), *Handbook of pediatric psychology* (pp. 401–425). New York: Guilford Press.

Kaplan, R. M., & Anderson, J. P. (1988). The general health policy model: Update and application. *Health Services Research, 23*, 203–235.

Katz, E. R., Rubenstein, C. L., Hubert, N. C., & Blew, A. (1988). School and social reintegration of children with cancer. *Journal of Psychosocial Oncology, 6*, 123–140.

Katz, E. R., Varni, J. W., Rubenstein, C. L., Blew, A., & Hubert, N. (1992). Teacher, parent, and child evaluative ratings of school reintegration intervention for children with newly diagnosed cancer. *Children's Health Care, 21*, 69–75.

Kupst, M. J., & Schulman, J. L. (1988). Long-term coping with pediatric leukemia: A six-year follow-up study. *Journal of Pediatric Psychology, 13*, 7–23.

La Greca, A. M., & Santogrossi, D. A. (1980). Social skills training with elementary school students: A behavioral group approach. *Journal of Consulting and Clinical Psychology, 48*, 220–227.

Lavigne, J. V., & Faier-Routman, J. (1993). Correlates of psychological adjustment to pediatric physical disorders: A meta-analytic review and comparison with existing models. *Journal of Developmental and Behavioral Pediatrics, 14*, 117–123.

Lobato, D., Faust, D., & Spirito, A. (1988). Examining the effects of chronic disease and disability in children's sibling relationships. *Journal of Pediatric Psychology, 13*, 389–407.

Magrab, P. R., & Calcagno, P. L. (1978). Psychological impact of chronic pediatric conditions. In P. R. Magrab (Ed.), *Psychological management of pediatric problems* (Vol. 1, pp. 3–14). Baltimore, MD: University Park Press.

Mattsson, A. (1972). Long-term physical illness in childhood: A challenge to psychosocial adaptation. *Pediatrics, 50*, 801–811.

McCormick, M. C. (1989). Long-term follow-up of infants discharged from neonatal intensive care units. *Journal of the American Medical Association, 261*, 1767–1772.

McGrath, P. J., Beyer, J., Cleeland, C., Eland, J., McGrath, P. A., & Portenoy, R. (1990). Report of the subcommittee on assessment and methodologic issues in the management of pain in childhood cancer. *Pediatrics, 86* (Suppl.), 814–817.

McHale, S. M., & Gamble, W. C. (1989). Sibling relationships of children with disabled and nondisabled brothers and sisters. *Developmental Psychology, 25*, 421–429.

Miller, S. A. (1988). Parents' beliefs about children's cognitive development. *Child Development, 59*, 259–285.

Moos, R. H., & Moos, B. S. (1981). *Family Environment Scale manual.* Palo Alto, CA: Consulting Psychologists Press.

Moos, R. H., & Tsu, U. D. (1977). The crisis of physical illness: An overview. In R. H. Moos (Ed.), *Coping with physical illness* (pp. 3–21). New York: Plenum.

Mullins, L. L., Olson, R. A., Reyer, S., Bernardy, N., Husztic, H. C., & Volk, R. J. (1991). Risk and resistance factors in the adaptation of mothers of children with cystic fibrosis. *Journal of Pediatric Psychology, 16*, 701–715.

Newacheck, P. W., Budetti, P. P., & Haflon N. (1986). Trend in activity-limiting chronic conditions among children. *American Journal of Public Health, 76*, 178–184.

Olson, D. H., Portner, J., & Lavee, Y. (1985). *FACES III manual.* Unpublished manual. (Available from D. H. Olson, Family Social Science, University of Minnesota, 290 McNeal Hall, St. Paul, MN, 55108).

Perrin, J. M., & MacLean, W. E., Jr. (1988a). Biomedical and psychosocial dimensions of chronic illness in childhood. In P. Karoly (Ed.), *Handbook of child health assessment: Biopsychosocial perspectives* (pp. 11–28). New York: Wiley.

Perrin, J. M., & MacLean, W. E., Jr. (1988b). Children with chronic illness: The prevention of dysfunction. *Pediatric Clinics of North America*, *35*, 1325–1337.

Pless, I. B. (1984). Clinical assessment: Physical and psychological functioning. *Pediatric Clinics of North America*, *31*, 33–45.

Pless, I. B., & Douglas, J. W. B. (1971). Chronic illness in childhood: Epidemiological and clinical characteristics. *Pediatrics*, *47*, 405–414.

Pless, I. B., & Pinkerton, P. (1975). *Chronic childhood disorders: Promoting patterns of adjustment*. Chicago: Year-Book Medical Publishers.

Pless, I. B., & Roghmann, K. J. (1971). Chronic illness and its consequences: Observations based on three epidemiological surveys. *Journal of Pediatrics*, *79*, 351–359.

Roghmann, K. J., & Haggerty, R. J. (1970). Rochester Child Health Surveys: I. Objectives, organization and methods. *Medical Care*, *8*, 47–59.

Rolland, J. S. (1984). Toward a psychosocial typology of chronic and life-threatening illness. *Family Systems Medicine*, *2*, 245–263.

Rolland, J. (1987). Chronic illness and the life cycle: A conceptual framework. *Family Process*, *26*, 203–221.

Rolland, J. S. (1993). Mastering family challenges in serious illness and disability. In F. Walsh (Ed.), *Normal family processes* (2nd ed.). New York: Guilford Press.

Sanders, M. R., Rebgetz, M., Morrison, M., Bor, W., Gordon, A., Dadds, M., & Shepherd, R. (1989). Cognitive-behavioral treatment of recurrent nonspecific abdominal pain in children: An analysis of generalization, maintenance, and side effects. *Journal of Consulting Psychology*, *57*, 294–300.

Sanders, M. R., Shepherd, R. W., Cleghorn, G., & Woolford, H. (1994). The treatment of recurrent abdominal pain in children: A controlled comparison of cognitive-behavioral family intervention and standard pediatric care. *Journal of Consulting and Clinical Psychology*, *62*, 306–314.

Satin, W. M., La Greca, A. M., Zigo, M. A., & Skyler, J. S. (1989). Diabetes in adolescence: Effects of multifamily group intervention and parent stimulation of diabetes. *Journal of Pediatric Psychology*, *14*, 259–275.

Senapati, R., & Hayes, A. (1988). Sibling relationships of handicapped children: A review of conceptual and methodological issues. *International Journal of Behavioral Development*, *11*, 89–115.

Sexson, S. B., & Madan-Swain, A. (1993). School reentry for the child with chronic illness. *Journal of Learning Disabilities*, *26*, 115–125.

Sparling, J., & Lewis, I. (1984). *Partners for learning*. Lewisville, NC: Kaplan Press.

Stein, R. E. K., & Jessop, D. J. (1982). A noncategorical approach to chronic childhood illness. *Public Health Reports*, *97*, 354–362.

Thompson, R. J., Jr., & Gustafson, K. E. (1996). *Adaptation to chronic childhood illness*. Washington, DC: American Psychological Association.

Varni, J. W., Katz, E. R., Colegrove, R., Jr., & Dolgin, M. (1993). The impact of social skills training on the adjustment of children with newly diagnosed cancer. *Journal of Pediatric Psychology*, *18*, 751–767.

Varni, J. W., Katz, E. R., Colegrove, R., & Dolgin, M. (1994). Perceived social support and adjustment of children with newly diagnosed cancer. *Journal of Developmental and Behavioral Pediatrics*, *15*, 20–26.

Varni, J. W., Walco, G. A., & Wilcox, K. T. (1990). Cognitive-biobehavioral assessment and treatment of pediatric pain. In A. M. Gross & R. S. Drabman

(Eds.), *Handbook of clinical behavioral pediatrics* (pp. 83–97). New York: Plenum.

Wallander, J. L., & Thompson, R. J., Jr. (1995). Psychosocial adjustment of children with chronic physical conditions. In M. C. Roberts (Ed.), *Handbook of pediatric psychology* (2nd ed., pp. 124–141). New York: Guilford Press.

Wallander, J. L., & Varni, J. W. (1989). Social support and adjustment in chronically ill and handicapped children. *American Journal of Community Psychology*, *17*, 185–201.

Wallander, J. L., Varni, J. W., Babani, L., Banis, H. T., & Wilcox, K. T. (1989). Family resources as resistance factors for psychological maladjustment in chronically ill and handicapped children. *Journal of Pediatric Psychology*, *14*, 157–173.

Wallander, J. L., Varni, J. W., Babani, L., DeHaan, C. B., Wilcox, K. T., & Banis, H. T. (1989). The social environment and the adaptation of mothers of physically handicapped children. *Journal of Pediatric Psychology*, *14*, 371–387.

Webster-Stratton, C. (1990). Stress: A potential description of parent perceptions and family interactions. *Journal of Clinical Child Psychology*, *19*, 302–312.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.15
Families of Children with Serious Emotional Disturbance

CYNTHIA R. ELLIS and NIRBHAY N. SINGH
Virginia Commonwealth University, Richmond, VA, USA

9.15.1 INTRODUCTION 235
9.15.2 CHARACTERISTICS OF CHILDREN AND THEIR FAMILIES 236
9.15.2.1 Children in Restrictive Settings 237
9.15.2.2 Children in Community Settings 237
9.15.2.3 Families 239
9.15.3 FAMILY IMPACT 239
9.15.3.1 Definition and Conceptualization 239
9.15.3.2 Research Findings 239
9.15.4 FAMILY SUPPORT 240
9.15.4.1 Definition and Conceptualization 240
9.15.4.2 Research Findings 241
9.15.5 FAMILY EMPOWERMENT 242
9.15.5.1 Definition and Conceptualization 242
9.15.5.2 Research Findings 242
9.15.6 FAMILY INVOLVEMENT 243
9.15.6.1 Definition and Conceptualization 244
9.15.6.2 Research Findings 244
9.15.7 FAMILY–PROFESSIONAL PARTNERSHIP 246
9.15.7.1 Definition and Conceptualization 246
9.15.7.2 Research Findings 246
9.15.8 FAMILY AND CULTURAL DIVERSITY 247
9.15.8.1 Definition and Conceptualization 247
9.15.8.2 Research Findings 248
9.15.9 SUMMARY 249
9.15.10 REFERENCES 249

9.15.1 INTRODUCTION

The seminal work of Knitzer (1982) documented that at least two of every three children with serious emotional disturbance (SED) in the USA did not receive the services that they needed and, when they did receive services, the services were inappropriate for the needs of many children. Given a population of three million children with SED, this meant that over two million of them were either not receiving needed services or the services that they were receiving were inappropriate. Spurred by these findings, as well as by demands of advocacy groups for better services for these children, Congress appropriated funds that enabled the creation of the Child and Adolescent Service System Program (CASSP) through the National Institute of Mental Health in 1984. This initiative enabled states and communities to begin developing comprehensive systems of care for children with SED and their families. Within 10 years, the CASSP had assisted all 50 states to improve their services for these children (Davis, Yelton, Katz-Leavy, & Lourie, 1995).

Most of the service systems developed by the states were based on a system-of-care concept that was advocated by CASSP (Stroul & Friedman, 1986). Although service delivery systems have taken new forms since the late 1980s, the major tenets of the system-of-care philosophy are still much in evidence today, both at local and state levels. The core values of this philosophy require that the system of care be child-centered and family focused, community-based, and culturally competent. In addition, Stroul and Friedman (1986, p. xxiv) provided a set of 10 philosophical guiding principles that is being used by communities and states in developing their systems of care, including:

(i) Children with emotional disturbances should have access to a comprehensive array of services that address the child's physical, emotional, social, and educational needs.

(ii) Children with emotional disturbances should receive individualized services in accordance with the unique needs and potentials of each child, and be guided by an individualized service plan.

(iii) Children with emotional disturbances should receive services within the least restrictive, most normative environment that is clinically appropriate.

(iv) The families and surrogate families of children with emotional disturbances should be full participants in all aspects of the planning and delivery of services.

(v) Children with emotional disturbances should receive services that are integrated, with linkages between child-serving agencies and programs and mechanisms for planning, developing, and coordinating services.

(vi) Children with emotional disturbances should be provided with case management or similar mechanisms to ensure that multiple services are delivered in a coordinated and therapeutic manner so that they can move through the system of services in accordance with their changing needs.

(vii) Early identification and intervention for children with emotional disturbances should be promoted by the system of care in order to enhance the likelihood of positive outcomes.

(viii) Children with emotional disturbances should be ensured smooth transitions to the adult service system as they reach maturity.

(ix) The rights of children with emotional disturbances should be protected, and effective advocacy efforts for children and youth with emotional disturbances should be promoted.

(x) Children with emotional disturbances should receive services without regard to race, religion, national origin, sex, physical disability, or other characteristics, and services should be sensitive and responsive to cultural differences and special needs.

Many communities and states have used these principles to develop local systems of care aimed at improving the outcomes for children and their families. Regardless of the details, all of the systems of care focus equally on the children and their families. Further, evaluation of these systems of care include family outcomes as an integral part of overall systems outcome.

The focus of this chapter is on issues that families of children with SED face in their daily lives. We begin with a discussion of the characteristics of the children with SED and their families, and note that, because of the heterogeneity of the children with emotional and behavioral disorders, these families face multiple problems not only in the direct care of their children but also in accessing multiple services for them. Next, we consider the issues of family impact and family support. Obviously, caring for a child with SED affects the lifestyle of the children's families and, although not all of it is negative, some families will need social, financial, and mental health support to cope with having a child with SED. Issues of family empowerment and involvement with respect to their importance in accessing services are discussed, and family–professional partnerships in terms of having parents as full partners with professionals in determining the needs and nature of the services provided to their children are considered. Finally, some of the issues related to cultural diversity and how these may affect families, professionals, and service delivery systems are discussed.

9.15.2 CHARACTERISTICS OF CHILDREN AND THEIR FAMILIES

Depending on the methodology and sample size used in epidemiological studies, the prevalence of children with SED ranges between 11% and 26%, with about 3–6% having serious psychopathology in need of treatment (Bird et al., 1988; Brandenberg, Friedman, & Silver, 1987; Costello et al., 1988). The impact of emotional disorders on these children's lives, as well as on their families and their communities, is immense. For example, because of their emotional and behavioral problems, these children are very likely to have severe adjustment and psychological problems as they grow older (Wagner, Blackorby, Cameto, Hebbeler, & Newman, 1993). The school dropout rate of students educationally labeled as SED exceeds 50% (US Department of Education, 1994), which means that the majority of them will face major obstacles in obtaining and maintaining a well-paid job. Of those who complete high school, up to 40% of them do not seek further education, vocational training, or regular employment (Wagner et al., 1993). Further, children with SED are often at risk for out-of-home and out-of-community placements because of their behavior and psychiatric problems.

9.15.2.1 Children in Restrictive Settings

Although many communities provide a single, community-based system of care, families often find that their children with SED continue to be served in restrictive settings, such as psychiatric hospitals, day treatment centers, juvenile justice learning centers, and residential treatment centers (Institute of Medicine, 1989; Nelson & Pearson, 1991; Singh, Landrum, Donatelli, Hampton, & Ellis, 1994). Even though inpatient psychiatric services have been criticized as an overused and unnecessarily restrictive treatment option, it is an integral part of a responsive system of care for certain children with SED (Cohen et al., 1990; Lundy & Pumariega, 1993). However, very little is known about the characteristics of the children who need such services.

In a study examining the characteristics of 250 children receiving partial hospitalization and inpatient psychiatric treatment, Singh et al. (1994) reported that, according to their educational status prior to admission, 46% had not been identified as handicapped for educational services, 36% were identified as SED, and another 18% had some other disabilities, such as mental retardation and learning disabilities. In the group of children with SED, boys outnumbered girls by a ratio of four to one. The mean age of the children with SED was about 12 years. About 77% of the children were Caucasian and the rest were African-American; 69% had lived at home prior to admission at a psychiatric hospital and 70% returned home at discharge from the hospital. A disproportionately higher percentage of children with SED had a prior history of psychiatric hospitalization. While only about 46% of children not having SED had previously been admitted to a psychiatric hospital, 73% of children with SED had a prior admission. This supports the findings of Mattison, Morales, and Bauer (1992) who reported significantly more children with SED receiving therapy for psychiatric problems than a control group of nondisabled children.

In Singh et al. (1994), the majority of the children with SED had no history of alcohol (83%) or other drug (87%) abuse. Similarly, the majority of the children with SED had no criminal history (91%). However, it must be remembered that the average age of these children was about 12 years, an age when children are just beginning to experiment with drugs and alcohol, as well as engage in criminal activity. Indeed, research shows that up to 40% of children with SED in our schools will become involved with the criminal justice system soon after leaving school (Jay & Padilla, 1987). Referrals for psychiatric services in the Singh et al. (1994) study came from the courts, mental health professionals, family, or self, with the major reasons for referral being threat to self and others and behavior problems. The three most frequently diagnosed psychiatric disorders for children with SED included, in order of prevalence, disruptive behavior disorders, attention-deficit hyperactivity disorder (ADHD), and mood disorders. Further, 64% of these children were on psychotropic medication, with 41% of those on medication receiving two or more medications.

In general, the findings of this study suggest that children and adolescents with SED and those without SED but who receive inpatient psychiatric treatment are similar in many respects, namely age, race, residential status at admission and discharge, type of referral (voluntary vs. involuntary) for psychiatric services, source and reason for psychiatric referral, self-reported amount of substance abuse, and criminal record. However, they differ in some significant ways as well. For example, more children with SED have a history of receiving inpatient and outpatient psychiatric services; need psychiatric services for ADHD; are prescribed medication, both at admission and at discharge; and are prescribed two or more medications, both at admission and at discharge. Singh et al. (1994) reported that although there were some significant differences between children with and without SED in their sample, the most important general finding was that the differences were more of *degree* than of *kind.*

9.15.2.2 Children in Community Settings

There is a recent but growing literature on characteristics of children with SED who receive community-based services (e.g., Barber, Rosenblatt, Harris, & Attkisson, 1992; Duchnowski, Johnson, Hall, Kutash, & Friedman, 1993; Epstein, Cullinan, Quinn, & Cumbald, 1994; Evans, Dollard, & McNulty, 1992; Landrum, Singh, Nemil, Ellis, & Best, 1995; Rosenblatt & Attkisson, 1992). These studies provide a fairly comprehensive picture of children with SED that families and service agencies deal with on a daily basis. These children are typically between 10 and 17 years, with studies from different geographical locations and sample sizes showing average ages ranging from 12 to 15 years (Epstein et al., 1994; Landrum et al., 1995; Rosenblatt & Attkisson, 1992). These findings are comparable with those reported for children treated in a residential mental health program (Wells & Whittington, 1993), as well as those identified as SED by their respective schools (Silver et al., 1992).

As with children seeking inpatient treatment, boys substantially outnumber girls in community-based services. For example, male prevalence rates between 75% and 80% have been reported in three studies (Epstein et al., 1994; Landrum et al., 1995; Quinn & Epstein, 1998), although slightly lower male prevalence rates (57–64%) were reported in other studies (Duchnowski et al., 1993; Rosenblatt & Attkisson, 1992). However, the higher male prevalence figures are in accord with national surveys of school children in special education programs who are identified as SED (e.g., Cullinan, Epstein, & Sabornie, 1992; Valdes, Williamson, & Wagner, 1990).

The data on racial prevalence is somewhat mixed. In the majority of studies, whites outnumber children from all other races. For example, prevalence of white children ranged from 63% to 77% in four studies (Duchnowski et al., 1993; Epstein et al., 1994; Quinn & Epstein, 1998; Rosenblatt & Attkisson, 1992), although in one of three communities surveyed by Rosenblatt and Attkisson (1992), the prevalence was only 45%. Further, Landrum et al. (1995) reported a figure of 54% African-American in a state-wide sample in Virginia, suggesting an overrepresentation of minorities among children and adolescents with SED, a finding noted by others as well (e.g., Kauffman, 1997; Singh, Ellis, Oswald, Wechsler, & Curtis, 1997; US Department of Education, 1994).

In terms of intellectual functioning, children and adolescents with SED score within the normal to low-normal range on standardized measures of intelligence (Duchnowski et al., 1993; Epstein et al., 1994; Quinn & Epstein, 1998). As for educational attainment, children and adolescents with SED have significant academic deficits in most content areas, are typically underachievers even when they have high IQs, are often in self-contained classrooms or alternative schools, and miss more days of school than their peers without SED because of truancy, suspension, expulsion, or dropout (Duchnowski et al., 1993; Epstein et al., 1994; Landrum et al., 1995). Further, the majority of these children exhibit behavior problems at some time during their school career. The combination of behavioral/psychiatric problems and academic deficits is a major risk factor for their continued educational achievements.

Children and adolescents with SED have a number of behavioral and psychiatric problems that can be generally characterized as externalizing disorders. Studies show that these children have significant disruptive behaviors or adjustment disorders, extreme noncompliance, physical aggression, and extreme verbal abuse, as well as mood disorders (Bruns, Burchard, & Yoe, 1995; Landrum et al., 1995; Rosenblatt & Attkisson, 1992). In one study, however, internalizing disorders were found to be twice as prevalent as externalizing disorders, a finding attributed by the investigators to a number of methodological differences between their study and the others that reported opposite findings (Quinn & Epstein, 1998). The use of prescribed medication as a treatment modality for these children ranges from 30% (Quinn & Epstein, 1998) to 56% (Landrum et al., 1995). This is a surprisingly high prevalence rate because the use of psychotropic medication by students with SED in the US school system is only about 11% (Cullinan, Gadow, & Epstein, 1987; Cullinan et al., 1992).

A high proportion of children and adolescents with SED have interactions with law enforcement agencies. Between 35% and 59% of them have a history of involvement with law enforcement agencies, typically because of their own unlawful behaviors (Duchnowski et al., 1993; Landrum et al., 1995; Quinn & Epstein, 1998). Further, many are referred to family courts because of abuse and neglect. Quinn and Epstein (1998) found that 20% of their sample had histories of abuse and neglect, and Zeigler-Dendy (1989) found that 33% had a history of abuse and 65% had experienced neglect. That children who are abused or neglected also have SED has been reported in other studies as well (e.g., Trupin, Tarico, Low, Jemelka, & McCellan, 1993). In addition, a majority of children and adolescents with SED have had out-of-home placements, typically in foster placements, psychiatric hospitals, or residential treatment centers (Epstein et al., 1994; Hawkins, Almedia, & Samet, 1989; Landrum et al., 1995; Wells & Whittington, 1993).

The picture that emerges is of children with multiple problems, requiring multiagency involvement. Thus, these children place immense burdens on their families, often so much so that they tear their families apart.

9.15.2.3 Families

Much less has been written about the characteristics of families of children with SED than about the children themselves. Less than a third of the nuclear families are intact, with 25–46% of parents being divorced (Landrum et al., 1995; Quinn & Epstein, 1998). Landrum et al. (1995) reported that 24% of the parents were single and another 21% of their sample of parents were either separated or widowed. Further, they found that African-American children were significantly more likely than white children to be living with a parent who was not married. Quinn and Epstein (1998) reported that in their sample, over 50% of the families had their parental rights qualified through either the child welfare system or the juvenile courts. They also reported that only 60% of the fathers and less than 50% of the mothers had full-time employment, and about 40% of the families received public assistance. Further, a significant percentage of the families had a history of mental illness, drug or alcohol abuse, and criminal activity. More than one half of these families had a history of family violence. Similar findings were reported by Epstein et al. (1994). Given the multiple problems of the children, the majority of these families had contact with multiple service agencies, including special education, juvenile justice, mental health, and child welfare. Accessing the needed services adds much strain on families that are already burdened by the daily grind of containing and helping their children to survive under difficult circumstances.

9.15.3 FAMILY IMPACT

9.15.3.1 Definition and Conceptualization

Having a child with mental health problems invariably affects the functioning of the family members (Donner et al., 1995; Friesen & Huff, 1996). This impact, which has traditionally been known as "family burden," can be defined as "the presence of problems, difficulties, or adverse events which affect the life of the ... patient's significant other(s)" (Platt, 1985, p. 383). Family members experience a high level of "objective" impact (e.g., disruption in family routine, providing transportation) and "subjective" impact (e.g., stigma, health problems, stress, reduced family cohesiveness) (Hoenig & Hamilton, 1967; Montgomery, Gonyea, & Hooyman, 1985). Often, these effects are in addition to other family stressors, such as poverty and poor housing. Friesen and Huff (1996) have noted that having a child with SED affects the family's lifestyle and routines, increases parental stress, decreases family time, and may negatively affect nuclear and extended family relationships. Of course, the level of family impact is dependent on a number of factors, including the severity of the child's emotional and behavioral problems, family strengths, cultural acceptance of the child's difficulties, and the parents' social support systems.

9.15.3.2 Research Findings

Although research on family impact has a long history, it has focused almost exclusively on assessing the effects that an adult family member with mental illness has on the rest of the family (Biegel, Sales, & Schulz, 1991; Clausen & Yarrow, 1955). In much of the research with adults, family impact was assessed using semistructured interviews or questionnaires; few psychometrically valid rating scales were used. Four new instruments have been developed to measure family impact of caring for a child with SED: the Impact on the Family Scale (IOFS; Stein & Riessman, 1980); the Impact of the Situation checklist (IOS; Regional Research Institute for Human Services, 1992); the Caregiver Strain Questionnaire (CGSQ; Brannan, Heflinger, & Bickman, 1997), and the Child and Adolescent Impact Assessment (CAIA; Patrick, Angold, Burns, & Costello, 1992). Some psychometric data are available for the IOFS, CGSQ, and the CAIA although much more work needs to be done to fully establish their psychometric properties.

Brannan et al. (1997) reported the development and psychometric properties of the CGSQ based on a sample of 984 families from the Fort Bragg Evaluation Project. The CGSQ is a 21-item self-report instrument that was developed specifically for assessing the impact of caring for a child with SED. Factor analyses revealed three related but distinct factors: objective burden, subjective internalized burden, and subjective externalized burden. The CSGQ has high internal consistency and convergent validity. The scale has been used in preliminary studies of family impact and the use of mental health services by children with SED (Bickman, Foster, & Lambert, 1996; Foster, Saunders, & Summerfelt, 1996). These studies found that high family impact is associated with the families using inpatient hospitalization and intermediate nonresidential services for their children, and lower family impact is associated with the families using traditional outpatient services alone. The suggestion is that family impact can be used as a predictor of the type of services that will be sought by the families of children with SED.

The link between family impact and the receipt of services has also been investigated. For example, it appears that the greater the family impact of caring for a child with SED, the more likely it is that the child will receive services (Angold et al., in press; Messer et al., 1996). Farmer, Burns, Angold, and Costello (1997) used the CAIA to assess the relationship between family impact and the families' use of services for their children. The CAIA measures the degree to which their children's emotional and behavioral problems impact on 17 areas of family functioning (e.g., the parent's relationship with his/her spouse/partner, restricted personal and family activities, parent well-being). The CAIA has good internal consistency and substantial discriminant validity between general population and clinical samples. Farmer et al. found that there was a strong association between family impact and initiation into and continued use of services for children with SED. Further, severity of the children's problems and degree of family impact were strongly associated with families accessing services. In addition, families accessed services when the care of the children impacted parental well being (e.g., depression resulting from the children's problems) and when parents felt incompetent in handling their children's problems. These findings suggest that parents access services for their children when the impact of caring for them gets beyond the parent's capability, and that they may not seek services solely on the basis of the mental health needs of the children. Given that children themselves do not seek services, and that some parents may not seek services for their children until the family impact is very severe, there may be children in our communities with unmet mental health needs.

9.15.4 FAMILY SUPPORT

9.15.4.1 Definition and Conceptualization

As the family impact literature shows, if families are to care for their children with SED at home and provide the context in which their children can grow and thrive in our society, we need to lessen their subjective and objective burdens. Caring for children exacts a toll on all parents, but more so on those parents who have children with disabilities because of the inadequacy of societal assistance available to them. Of course, caring for a child with SED is not inevitably a negative experience for most parents; however, current research does suggest that when family impact increases, parents of children with SED are at greater risk for psychological distress than those who do not have children with SED. Parental stresses may include economic stress, social isolation, and disruption of family life, including high rates of divorce and lower rates of remarriage (Singer & Irvin, 1989).

Families can overcome many of life's hurdles if they have access to a service system that supports their social, emotional, and tangible needs. Family support can be defined as,

> a constellation of formal and informal services and tangible goods that are defined and determined by families. It is "whatever it takes" for a family to care for and live with a child or adolescent who has an emotional, behavioral, or mental disorder. It also includes support needed to assist families to maintain close involvement with their children who are in an out-of-home placement, and to help families when their children are ready to return home. (Federation of Families for Children's Mental Health, 1992, p. 1)

It is important to note that this definition stresses the fact that the family members determine what their needs are and the nature of the assistance they seek to overcome these needs. Further, the "whatever it takes" approach means that family supports will vary among families and over time.

The Federation of Families for Children's Mental Health (1992, p. 1) has provided examples of family supports for those with a child with SED:

(i) Family self-help, support, and advocacy groups and organizations.

(ii) Information and referral.

(iii) Education that will support families in becoming active, informed decision-makers on behalf of their family and the child.

(iv) Advocacy with and on behalf of the family, if needed.

(v) The capacity to individualize, provide flexible support services, and meet unplanned needs quickly and responsively.

(vi) In-home and out-of-home respite care, with an emphasis on neighborhood and community participation for the child and conceptualized not as a clinical service but as a support for the whole family.

(vii) Cash assistance.

(viii) Assistance with family survival needs (housing, food, transportation, home maintenance).

(ix) Other supports, as determined by the family.

These examples suggest that family supports are society's responsibility, because it must be provided not only by the state and local agencies but also by families, friends, churches, and other social networks, such as parent-groups.

9.15.4.2 Research Findings

There are many studies showing that joining a support group reduces psychological and social distress in families. For example, Telleen, Herzog, and Kilbane (1989) enrolled mothers from low-income families in a family support program and found that, after three months, these mothers experienced less social isolation and parenting stress than mothers in a control group. This finding is in accord with others that show families who care for children with SED find family support programs to be supportive and rewarding (Rowe, 1991). Further, joining parent support groups enables parents to network with other parents with similar needs, identify other needs, and access services for these needs (Koroloff & Friesen, 1991). Respite care is another source of family support because it enables families to be less socially isolated and to undertake family activities that may not always be possible because of the emotional and behavioral problems of a family member. While the research literature on the impact of respite care is limited, the vast research base on families with children with developmental disabilities shows that respite care produces positive outcomes for the parents as well as the siblings of the child with disabilities (Singer & Irvin, 1991).

With the movement to providing family-centered services, service providers have begun including clusters of supports in their programs. For example, Armstrong and Evans (1992) included intensive case management, access to respite care, family support groups and other family support services, the services of a paid family advocate, and recreational activities in their family-centered intensive case management services. While it is difficult to tease out the contributions of the various aspects of the support services, their program showed that a cluster of support services was more beneficial than therapeutic foster care services.

In the 1990s, there has been a paradigm shift in the provision of services to children and adolescents with SED. A substantial number of states and local communities have begun using what is loosely called wraparound services. One of the eight key elements of this type of service states that "Services and supports must be individualized to meet the needs of the children and families ... " (VanDenBerg & Grealish, 1996, p. 9). This key element has been included a number of research-based wraparound services in the USA (e.g., Clark, Lee, Prange, & McDonald, 1996; Eber, Osuch, & Redditt, 1996; Hyde, Burchard, & Woodworth, 1996; Yoe, Santarcangelo, Atkins, & Burchard, 1996). Program evaluation data from these service programs indicate that family support is vital to positive outcomes for children with SED. However, cross-program comparisons are not possible as each included different combinations of family supports.

VanDenBerg and Grealish (1996, pp. 14–15) have suggested that the following domains should be considered when providing clusters of family support in community-based services:

(i) *Residence*. Do the current living arrangements meet the child's and family's needs?

(ii) *Family*. Who is in the family, by their definition? Do all family members have appropriate access to each other? What do the members of the family need to stay together or in touch with each other? Are there serious, unmet needs for any family members that impair family functioning?

(iii) *Social*. Do family members have friends and access to their friends? Does this family have the opportunity to socialize with each other?

(iv) *Emotional/Psychological*. Does the referred individual or family have any unmet needs in these areas? Are there unresolved issues which impede normal interactions within the family or in the community?

(v) *Educational/Vocational*. What will it take to ensure a viable education for the children, particularly the identified client? Do older children have access to employment opportunities? Are their rights to an appropriate education being enforced?

(vi) *Safety*. Is everybody in the family safe? Are there dangers to individual family members? Are clients potentially dangerous to themselves or to the community?

(vii) *Legal*. Are any family members involved in the judicial system, on probation, or parole? Do they have representation? Are there issues around custody of the children?

(viii) *Medical*. Are health care needs met? Does the family have access to any specialist services they may need?

(ix) *Spiritual*. Does this family identify unmet spiritual needs? Do all family members have access to spiritual leadership if they desire it?

(x) *Cultural*. Are cultural differences present and being considered?

(xi) *Behavioral*. Are any family members exhibiting behaviors which must be addressed if the quality of life for the family is to improve?

(xii) *Financial*. Does the family have pressing financial needs which are not being met through employment or public support?

Clearly, as this list shows, the development of a solid support system is a critical aspect of the care of children and adolescents with SED. Singer and Irvin (1991) have noted that respite care, child-focused education, parent training, and fiscal support are critical variables in reducing parental stressors. Supportive interventions that focus on helping families to use cognitive and behavioral coping strategies, such as cognitive reframing, have been found to be very helpful to families. Other supports that have been found to be effective include strengthening intrafamily support (Hawkins & Singer, 1989), strengthening informal social ties outside the family (Lehman, Ellard, & Wortman, 1986), and strengthening formal ties between parents and professionals (Singh, 1995). There is a growing literature on the need for family support services which shows that supportive interventions produce effective and lasting change in a family's ability to care for a child with SED.

9.15.5 FAMILY EMPOWERMENT

9.15.5.1 Definition and Conceptualization

Empowerment has been defined in a number of different ways and, in the broadest sense, it implies a process which enables individuals to gain control over their lives by influencing their interpersonal and social environments (Hasenfeld, 1987; Parsons, 1991; Rappaport, 1981; Zimmerman & Rappaport, 1988). The results of this process for the individual may be many, depending on the specific context in which the process occurs. For example, psychological empowerment has been shown to be associated with greater participation in community activities and organizations (Zimmerman & Rappaport, 1988). In general, however, the goal of individuals in becoming empowered is to bring about a positive change that improves their quality of life or achieves a lifestyle change that they desire (Curtis & Singh, 1996).

More recently, empowerment has been conceptualized within the framework of mental health service delivery systems. Although children and individuals with disabilities are frequently the recipients of mental health services, they are typically unable to advocate for themselves. Thus, increasing attention is being paid to the empowerment of their families needing access to the services on their children's behalf. The traditional paradigm of service providers and professionals "knowing all" and being directive in their prescription of treatments, regardless of the unique nature, expectations, and outcomes desired by a family, is no longer acceptable (Singh, 1995). Furthermore, the concept that empowerment involves a professional actively directing or compelling individuals or families to become a meaningful and influential participant in their own care has similarly fallen out of vogue because it is now generally accepted that one person cannot empower another. While professionals can encourage families to take more active roles in the care provided to their children, only the families can empower themselves to do so.

In the context of mental health service delivery systems, family empowerment has been conceptualized as "a process by which the families access knowledge, skills and resources that enable them to gain positive control of their own lives as well as improve the quality of their life-styles" (Singh, 1995, p. 13). Although mental health service providers and professionals cannot empower a family, there is growing support for the notion that they serve a vital role in facilitating or assisting families to empower themselves within this system. For example, professionals can assist families to increase their level of empowerment by making structural changes in the way they deliver services to these families and their children (Singh et al., 1995). It is now recognized that, along with the more traditional goals, providers of mental health services to families and children with disabilities should facilitate the empowerment of the recipients of such services (Freund, 1993; Parsons, 1991; Singh et al., 1995; Sluyter, 1994; Staples, 1990). There is ample evidence that formalized efforts by service providers to facilitate the empowerment of families involved in the mental health service delivery system are generally successful and lead to positive perceptions and outcomes (Battaglino, 1987; Dunst, Trivette, Davis, & Cornwell, 1988; Hasenfeld & Chesler 1989; Kopp, 1989).

9.15.5.2 Research Findings

Helping families to increase the level of their empowerment is a priority activity for service providers and professionals in the mental health system. Furthermore, in order to evaluate the impact of interventions designed to facilitate family empowerment, it is important that we assess families' perceptions of their empowerment status and empirically derive and validate the dimensions of the empowerment construct. To this end, Koren, DeChillo, and Friesen (1992) developed the Family Empowerment Scale (FES), a rating scale to assess family empowerment in the context of mental health service delivery systems. This scale was developed within a conceptual framework of empowerment that accounts for three levels of empowerment (Family, Service System, and

Community/Political) and three modalities for expressing empowerment (Attitudes, Knowledge, and Behaviors). A factor analysis by Koren et al. (1992) confirmed the validity of the three levels of empowerment, while only minimally confirming the three modalities for expressing empowerment. Another factor analysis of the FES resulted in the identification of four components of empowerment measured by the scale: Systems Advocacy, Knowledge, Competence, and Self-efficacy (Singh et al., 1995). These four components corresponded to four major areas of empowerment identified in the literature and represented a more atheoretical approach to the validation of the FES than that taken by Koren et al. (1992).

Simply defining the dimensions of the empowerment construct, however, is not useful on a practical level for service providers and professionals who want the families that they serve to increase their ability to empower themselves. In addition, an understanding of the nature of families with various levels of empowerment is important in the implementation of strategies designed to increase the effectiveness of mental health service delivery systems. As part of their validation study of the FES, Koren et al. (1992) assessed whether or not the FES was able to differentiate between those parents who participated in activities indicative of empowerment and those who did not. Koren et al. grouped these activities into six types, advisory, political, legal, assistance, organizing, and participant. Participation in these activities was postulated to be a behavioral indicator of empowerment, primarily at the Community/Political level. The results indicated that parents who participated in these activities did, in fact, score higher on the three subscales of the FES than parents who were not involved in these activities.

Other than for the work of Koren et al. (1992), there have been no further investigations of the relationship between family empowerment and theoretically proposed correlates of this construct. Singh, Curtis, Ellis et al. (1997) examined some of the demographic correlates of family empowerment in a sample of parents of children with SED alone and parents of children with SED as well as ADHD. Utilizing the FES as a measure of family empowerment and the four factors derived by Singh et al. (1995) as subscales, the empowerment status of this sample of parents across several family and demographic characteristics was investigated. The results indicated statistically significant correlations between family empowerment and membership in a parent support group, mental health status of the child, gender of the respondent (mother or father), and the parent's education level.

Case studies and anecdotal family reports have suggested that empowerment of families through the use of support groups yields positive results for their children (Battaglino, 1987; Friesen & Wahlers, 1993; Kopp, 1989). However, Singh, Curtis, Ellis et al. (1997) provided the first set of empirical data confirming the overriding positive impact of parent groups on the empowerment status of families. In this study, membership in the parent support group was the variable that most influenced a family's level of empowerment.

In a study discussed in the next section, Curtis and Singh (1996) found a correlational link between family empowerment and family involvement with services for their children with SED. However, the study did not deal with the issue of causality; that is, it did not determine whether greater involvement with the service system increased parental empowerment or if greater empowerment resulted in increased family involvement. Curtis and Singh suggested that the relationship between involvement and empowerment is reciprocal and inseparable. They noted that

> family involvement is strongly influenced by the level of parental empowerment and, at the same time, parental empowerment is strongly influenced by parental involvement. Our view, which is analogous to a constructivist position, focuses on the active role of parents in creating their cognitive, social, and political worlds in an effort to access the best services for their children. (p. 515)

9.15.6 FAMILY INVOLVEMENT

Families are the primary socialization and change agents for their children. They play a critical role in enhancing the quality of life of their children through direct care as well as through advocacy. Thus, it is to be expected that families will be closely involved in every aspect of the care of their children, including the formulation, planning, and evaluation of services. Regardless of whether they are heavily involved or not, families are the consumers of mental health services and, as such, they need to have a voice in determining what and how the services should be provided. Family members understand the context and nature of their children, so they are in a better position than most professionals in identifying their children's service needs. Further, because family members are often the ones who carry out interventions with their children, they need to be involved both in choosing and in determining how the interventions will be implemented.

9.15.6.1 Definition and Conceptualization

There is no single, empirically derived definition of family involvement in the child mental health services literature. However, it is important that the term be defined in terms of how the families perceive their own involvement in the care of their children, as well as their interactions with the mental health services system generally. Clearly, the concept of family involvement is multidimensional and includes a variety of activities that facilitate the process and ultimate achievement of involvement. Of course, it is not a state that can be achieved but a process along a continuum of involvement.

The components of family involvement have been identified through research, focus groups, and interviews with families receiving mental health services. These include, (i) involvement of families in all aspects of services at all times, with professionals genuinely encouraging this involvement; (ii) education of the family about the nature of services and how the service system works; (iii) involvement of the family in all decision-making; and (iv) keeping the family informed on the process and progress of treatment (Brotherson & Goldstein, 1992; Collins & Collins, 1990; DeChillo, 1993; Tarico, Low, Trupin, & Forsyth-Stevens, 1989).

Generally, family involvement is thought to occur at multiple levels within the service delivery system, and may take different forms (Fullan, 1982). At the *person* level, family involvement can be used to enhance the quality of care provided to the child with SED. At this level, family members may act as the child's therapist, collaborate with professionals in developing the child's educational and treatment plans, advocate for additional services that can be purchased in the community, evaluate the services being provided, and educate professionals about the family's world view of systems and services. At the *program* level, family involvement can be used to enhance the nature, quality, and quantity of services provided to the child with SED. Family members work closely with professionals in assessing the needs of their child and matching the services to these needs.

At the *community* level, family involvement is focussed at more general outcomes, such as working with service delivery agencies and making them more responsive to the needs of children with SED generally rather than just to their own child. At this level, family members are involved in educating the community about children with SED and in holding the community responsible for serving all children in their catchment area. Finally, at the *society* level, family involvement takes on the most global level of commitment. At this level, family members encourage change in the larger society that they live in and work with volunteer agencies, such as with local and national organizations, that advocate for and serve children with SED. Family members form alliances with other advocacy groups, educate legislators and other key government officials, and encourage greater advocacy for their children with SED at a national and international level. Clearly, families can be involved, and should be encouraged by professionals to be involved, at the level that best suits them.

9.15.6.2 Research Findings

The philosophical concept of family involvement was developed primarily within the context of human service delivery systems, particularly in the general area of services for children with developmental disabilities. Family involvement encompasses a conceptually narrower territory than family empowerment, relates specifically to involvement in service delivery systems, and does not carry the broader political and social ramifications that empowerment does. It reflects an ongoing process in mental health service delivery that is facilitated by service providers, and is characterized by a partnership or collaboration among families and service providers, with power and responsibility being shared among them (DeChillo, 1993; DeChillo, Koren, & Schultze, 1994; Friesen & Koroloff, 1990).

There is a paucity of studies on family involvement in child and adolescent mental health services. One of the major reasons for this state of affairs has been the lack of a well-validated instrument for assessing the construct of family involvement. Two versions of a self-report rating scale have been developed recently to describe empirically and quantify family involvement (Curtis, Singh, & Cohen, 1996). The Family Involvement Scale-Family Version (FIS-F) assesses the families' perceptions of their involvement in mental health services for their children with SED (Singh et al., 1996a) and the Family Involvement Scale-Professional Version (FIS-P) assesses professionals' perceptions of family involvement in mental health services (Singh et al., 1996b). These instruments can be used to investigate the relationship between parent-rated family involvement and selected demographic variables.

Curtis and Singh (1996) investigated the sociodemographic correlates of family involvement in mental health services for children with SED, as well as the relationship between family involvement and family empowerment. The

FIS-F was used to measure involvement of families in the provision of mental health services for their children, and the FES was used to measure family empowerment. The results showed that mothers reported greater involvement in services for their children than fathers, and those respondents (mothers and fathers) who had less formal education were more involved than those who had more formal education. Further, the knowledge subscale of the FES was moderately correlated with all three subscales of the FIS-F, suggesting that parents who perceive themselves as being knowledgeable about their local mental health delivery systems also perceive themselves as being involved in the services provided to their children within this system. This finding reinforces our intuitive belief that knowledge of mental health service delivery systems is a moderator of parental participation.

There are a number of barriers to a family's involvement in their child's mental health care, including resources, knowledge, and skills. Resource barriers that hinder family involvement include poverty; lack of adequate transportation, time, and opportunity; and limited family supports. Families that live in poverty, work long hours for minimum pay, or are in jobs that will withhold pay for time spent taking a child for a mental health appointment cannot be expected to be fully involved with the mental health system. Families that cannot find adequate child care for siblings of the child with SED do not have the luxury of becoming involved even if they so desired. Further, some families live in areas that are not served by public transportation or live at such a distance from the mental health facility that travel is a major barrier to closer involvement. Service delivery agencies can make accommodations in their planning for greater family involvement by scheduling meetings after work hours, providing child care services during the meetings, and arranging transportation for those who need it. Indeed, with some thought, service agencies can vastly increase family involvement at a personal and program level through skillful use of social supports and friendliness of the services they provide (Singh, Curtis, Ellis et al., 1997).

Knowledge barriers include those variables that limit the family's involvement either because they do not understand the service system or the skills needed to deal with professionals in an assertive manner. For example, families often do not have adequate knowledge of the agency's individualized planning process (e.g., individualized education programs in schools) or the political structure of the mental health system in their locality to become useful advocates for their children. Given the cultural diversity of our society, some ethnic minority families may need assistance in learning how to assert themselves with professionals in conflictual situations, gather necessary information so that they can fully participate in the educational and treatment planning process, and access the necessary community supports for their children. In such cases, family involvement can be increased by providing comprehensive learning opportunities to both the families as well as the professionals.

It is crucial that professionals learn how to be sensitive to the families' cultural nuances regarding their preferred level of involvement. Further, the Family Barriers Checklist (Elliott, Koroloff, Koren, & Friesen, 1998), a scale that families can use to rate the degree to which each of 13 variables pose a barrier to service involvement, can be used by professionals to assess barriers to involvement by individual families. In many cases, by carefully assessing family needs and actively listening to the family's views regarding their role in their child's care, professionals will be able to negotiate with the family regarding their level of involvement. Further, because of the vast tapestry of families that professionals deal with, it is important that they carefully examine their own and the families' assumptions, understanding, expectations, and level of involvement (Adams et al., 1997).

Many advocates have suggested that families take an active role in shaping the policies and practices of their local and national mental health agencies. In this regard, Adams et al. (1997, pp. 35–36) have provided the following guidelines for families that should make them more effective in bringing about meaningful policy change at the local level: (i) learn about the agency, (ii) examine the agency's written and stated goals, (iii) examine the agency's motivation, (iv) examine the agency's mission statement, (v) identify a common vision for yourself and the agency, (vi) develop a formal agreement with the agency, (vii) develop a reporting mechanism, (viii) define communication links, (ix) define the scope of work and time line in writing, and (x) secure financial support for family participation. These are very practical suggestions that provide the framework for initiating family involvement both at the personal and program levels. However, we need further research to show whether these tactics do indeed work, and how well they work in current mental health service delivery systems.

Family involvement is critical in the overall quality of services provided to children and adolescents with SED. Although conceptual work on family involvement has been available and empirical work on the topic has been abundant in the field of developmental disabilities, little has been done on it with regard to child mental health services. We suspect that with the development of appropriate assessment instruments, such as the FIS-F, FIS-P, and the Family Barriers Checklist, we will see much more research on the construct of family involvement.

9.15.7 FAMILY–PROFESSIONAL PARTNERSHIP

Parent–professional relationships have received considerable attention in the general area of childhood disabilities, especially in developmental disabilities (e.g., Mittler, Mittler, & McCanachie, 1987). The literature on parent–professional partnership in child mental health, however, is much more limited.

9.15.7.1 Definition and Conceptualization

There are a number of good reasons why child mental health professionals should involve parents as partners in the care of their children with SED. For example, it has been suggested that involving parents has intrinsic value (Friesen, 1989) and that it leads to improved outcomes for both the children and their parents (Burns & Friedman, 1989). Parent–professional partnership can be defined as parents and professionals having an active and equal role in making decisions about the needs and services for the child. It assumes that parents and professionals will have a collaborative working relationship, with each educating the other about the needs and services of the child, and both assuming responsibility for the process and outcome of the services.

Parent–professional partnerships have been conceptualized in several ways. For example, Hobbs et al. (1984) proposed a needs-based framework for providing services in which both the child and the family benefit; the child is the direct recipient of services and the family members increase their knowledge and competence. Parent–professional partnership in this model assumes that professionals will assist the families (i) to learn how to care better for their children, (ii) to access more of the available services, (iii) to develop competencies which will enable them to reduce their psychological and other stressors, and (iv) to network their social supports in the community. Other models include the "reflection in action" model (Schön, 1983), "family consultation" model (Bernheim, 1989), "enablement model of helping" (Dunst & Paget, 1993), "family alliance" model (Lefley, 1989), and "person-centered planning" model (Singh, 1997). At the core of these models is the premise that families are critical to the process of service delivery because they provide the context in which the child and family function. As noted by numerous professionals, the parents are the experts on their children and no professional can ever rival their knowledge of their children and the contexts in which their children need to enhance their functioning.

9.15.7.2 Research Findings

There is little research on parents' perceptions of professional behaviors and on parent–professional partnerships. In one study, Friesen, Koren, and Koroloff (1992) investigated parents views of the importance and frequency of professional behaviors in mental health services. Regardless of the professionals that the parents sought services from, they considered the following as very important to them: parent–professional relationship, honesty, nonblaming attitude, supportiveness, and inclusion in decision-making. Further, parents indicated that professionals only partly met their expectations. These findings are not surprising given that parents have been telling professionals that they wish to assume a more active role in decision-making regarding the services provided to their children (Collins & Collins, 1990). They strongly object to being excluded from decision-making or, worse, being blamed for their children's problems (Bernheim, 1989; Collins, & Collins, 1990; Lefley, 1989).

There is a growing feeling in the mental health field that the traditional service delivery system which assumes the primacy of the professional is not only ineffective in the long-term but also dehumanizing to children with SED and their families. The way services have been provided in the past are not in the best interests of the families and their children. The traditional model assumes that professionals, by their training and experience, know best how to assess, diagnose, and treat all manner of problems in children, regardless of the diversity of contexts that these children may come from. However, there is little data to support the efficacy of a service delivery system that revolves around professionals and excludes the families from full partnership in the treatment process. The tradition model creates a duality, a distinction between the professional as the knower and the provider of services and the family as the receiver of the services. There is little oneness between the professional and the family, and no oneness of vision of what may be best for the family.

In response, new service delivery models are emerging, such as wraparound (VanDenBerg & Grealish, 1996) and person-centered planning (Singh, 1997), and these need to be tested for their alignment with family values, expectations, and outcomes (Singh, Curtis, Wechsler, Cohen, & Ellis, 1997). Clearly, the process of paradigm shift is underway in child mental health service delivery systems. Professionals cannot equivocate on the changes that they must make if they are to be successful participants in the emerging paradigm. Singh (1995) has discussed what professionals can do to produce service delivery systems that are compassionate and sensitive to the needs of the families and their children. For example, above everything else, they need to learn to be one with the families and their children. They need to liberate themselves from the bondage of outdated theories, and from judging families in terms of their own mental constructs of a "healthy" or normative family. Professionals cannot facilitate family growth, empowerment, and involvement if they cannot see the world from the eyes of the family. They need to take a posture of mutual respect and reciprocity with families that will bring them closer to the world view of the families.

9.15.8 FAMILY AND CULTURAL DIVERSITY

A tapestry of families, rich in linguistic, racial, ethnic, and socioeconomic diversity, have children with SED who need mental health services. The cultural heritage and diversity of these families almost guarantees that none of them will fully fit the service provider's mental model of a healthy family, a model typically based on white, middle-class families. However, all families share some characteristics that are universal, regardless of diversity. For example, they may have a sense of loss of self-respect when they visit a professional because of the feeling that they have failed as parents. Nontraditional and culturally diverse families may have added problems because of their feeling that they are not understood by mental health professionals who are unaware of the structural aspects of their families and therefore, unable to fully comprehend the nature of their problems. While they believe that professionals may have a family's best interests at heart, they also know that families and professionals often view the same problem and its solutions differently.

The changing cultural demographics of our society is reflected in the demographics of children with SED and their families in systems of care. As the numbers of children from culturally diverse families increases in systems of care, it means that the context in which services are being delivered is also changing. This change has immense implications for the nature and effectiveness of the services delivered in systems of care.

9.15.8.1 Definition and Conceptualization

Differences in the culture of the families and mental health professionals can have a profound impact on the outcome of services. Cultural differences in values, approaches to child rearing and education, views of mental illness, and the seeking of services for their children undoubtedly influence many aspects of the involvement of families in systems of care (Harry, 1992; Julian, McKenry, & McKelvey, 1994; Kalyanpur & Harry, 1997; Lee, 1997). Further, the cultural competence and sensitivity of the professionals will impact not only the nature of the services sought and received by these families, but also the level of their involvement with different sectors of the service system.

Although no single definition of culture is likely to be universally accepted, it can be broadly defined as the shared values, traditions, arts, history, folklore, and institutions of a group of people that are unified by race, ethnicity, nationality, language, religious beliefs, spirituality, socioeconomic status, social class, sexual orientation, politics, gender, age, disability, or any other cohesive group variable (Singh, 1996). Clearly, this definition recognizes that all of us simultaneously belong to more than one cultural group and that each person is in a complex dynamic relationship with others from overlapping cultural traditions.

As with the term culture, there are many definitions of cultural awareness, sensitivity, and competency. Cultural *awareness* is a general term used to indicate that a person is conscious of the similarities and differences within, between, and among cultures. Awareness is a necessary but not sufficient condition for a person to behave in a culturally appropriate manner towards others from different cultures. Cultural *sensitivity* is a more specific term used to indicate that a person not only has an awareness of the nuances of one's own culture as well as those of other cultures, but also that he or she does not assign a negative or positive value to the differences within, between, and among cultures. Sensitivity means that the person accepts cultural differences nonjudgmentally. Cultural *competency* is a term used to indicate that a person has "knowledge and skills that enable him or her to appreciate, value and celebrate similarities and differences within, between, and among culturally diverse groups of people" (Singh, 1996, p. 124). Competency is a dynamic concept and does not imply that a person ever reaches a state of being universally culturally competent; rather, it implies that the person has the knowledge and skills for displaying and increasing his or her understanding and appreciation of the changing nature and nuances of his or her own culture as well as those of others.

9.15.8.2 Research Findings

According to Kauffman (1997), 3–6% of children in our schools probably have emotional and behavioral disorders that require some form of intervention. Although the data are limited, there is a good general indication that children with SED form a culturally diverse group. For example, an analysis of nine recent studies on the characteristics of children in systems of care shows that, on average, the age of the children ranges from about 8 to 16 years, about 70–75% of them are boys, and between 50% and 75% are Caucasians, followed by about 25–30% African-Americans, 10% Hispanic Americans, and others, including native Americans, Asians, Pacific Islanders, and those of mixed races (Barber et al., 1992; Cullinan, Epstein, & Quinn, 1996; Epstein, Cullinan, Quinn, & Cumbald, 1994, 1995; Landrum et al., 1995; Quinn et al., 1996; Quinn, Newman, & Cumbald, 1995; Silver et al., 1992; Singh et al., 1994).

These data indicate that even when only three indicators of cultural diversity (age, gender, and ethnicity/race) are taken into account, there is a broad spectrum of children with SED who are being served in various systems of care in this country. Of course, virtually no data are available on some of the other aspects of cultural diversity in these children and their families, such as their nationality, language, religious beliefs, spirituality, socioeconomic status, social class, sexual preference, politics, and disability, among others. The reporting of these data are crucial in future research because it may well be that some of these variables will be found to be highly correlated not only with the nature of the services needed or sought by children with SED and their families, but also with the types of services that are most appropriate.

A related issue is the lack of similar data reported in outcome studies from various systems of care. Such data provide an indication of cultural diversity of the sample that is being provided with services. In addition, the reporting of these data indirectly indicates that mental health professionals are aware of the unique needs of children and families from different cultural groups. For example, it is well established that clinical psychologists may need to use assessment instruments and diagnostic methods that are culturally appropriate for children from different racial and ethnic groups. Further, children from diverse cultural backgrounds may need additional interventions that are different from the generic programs used with white, middle-class children and their families (Singh, Ellis, Oswald et al., 1997; Singh, Williams, & Spears, in press).

Current research suggests that professionals have culturally laden interactions with families and their children throughout the process of providing psychological, educational, and mental health services (Harry, 1992; Kalyanpur & Harry, 1997; Singh, 1995). Singh (1998) has detailed the nature and extent of these interactions and the pitfalls that professionals should be aware of. These interactions range from collecting sociodemographic and intake information, assessment, cultural formulation of the presenting problems, and interventions to the evaluation of treatment outcome, cultural influences on data analysis and interpretation, and family satisfaction. Space precludes a detailed discussion and interested readers are referred to this source for further information.

Finally, there is the issue of whether mental health professionals, such as clinical psychologists, can ever be competent enough to deal with the cultural nuances of the myriad of culturally diverse families who will seek their services. Although this question has not been subjected to rigorous research, intuitively we feel that the answer is no. However, we believe that sensitivity, understanding, and compassion will bridge the gap between culturally sensitive professionals and families who seek their services. Above all, good listening skills and understanding the issues from the perspective of the families are critical skills that all professionals should have. Fortunately, Kleinman, Eisenberg, and Good (1978) have provided us with a list of eight simple questions that professionals can ask, and the family's answers provide the professionals with the family's explanatory model of their problems. When professionals do not appreciate the cultural underpinnings of their own actions and that of the families, at best, their efforts will be only partially effective and, at worst, they will have to assume some of the responsibility for the tragic outcomes that can occur (see Fadiman, 1997).

Given the pervasive effects of culture on human behavior, mental health professionals will find it difficult to be cognizant of and respond appropriately to all cultural variables in every facet of service provision to families with a child with SED. Obviously, the effects of even the major cultural variables can neither be assessed fully nor controlled with any one family. Our research efforts and clinical practice should focus on finding assessment and intervention principles and methods that are universal and would apply across cultures, as well as those that are specific to different cultural groups.

9.15.9 SUMMARY

Families play a critical role in the care of children and adolescents with SED; they, not professionals, are the experts who know the children best. They determine if and when their children enter mental health services, and what services are accessed. In addition, recent research is making clear that caring for a child with SED has its own set of joys and rewards, above and beyond raising children in general. This has brought to light the cultural, political, and practical issues that these families have to face on a daily basis. These issues have spawned conceptual, philosophical, and research papers and books since the 1980s; obviously, enough material that cannot be summarized in one chapter. This chapter covers a selected set of issues that are still in their formative stages, both in terms of concepts and research, and presents some ideas regarding areas of future research and intervention with these families.

9.15.10 REFERENCES

Adams, J., Gora, K., Huff, B., Burrell-Muhammad, V., Rivera, M., & Slaton, E. (1997). Defining family involvement. *Reaching Today's Youth, 1*(4), 35–36.

Angold, A., Messer, S. C., Stangl, D., Farmer, E. M. Z., Costello, E. J., & Burns, B. J., (in press). Perceived parental burden and service use for child and adolescent psychiatric disorders. *American Journal of Public Health.*

Armstrong, M. I., & Evans, M. E. (1992). Three intensive community-based programs for children and youth with serious emotional disturbance and their families. *Journal of Child and Family Studies, 1,* 61–74.

Barber, C. C., Rosenblatt, A., Harris, L. M., & Attkisson, C. C. (1992). Use of mental health services among severely emotionally disturbed children and adolescents in San Francisco. *Journal of Child and Family Studies, 1,* 183–207.

Battaglino, L. (1987). Family empowerment through self-help groups. In A. B. Hatfield (Ed.), *Families of the mentally ill: Meeting the challenges* (pp. 43–51). San Francisco: Jossey-Bass.

Bernheim, K. F. (1989). Psychologists and families of the severely mentally ill: The role of family consultation. *American Psychologist, 44,* 561–564.

Bickman, L., Foster, E. M., & Lambert, E. W. (1996). Who gets hospitalized in a continuum of care? *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 74–80.

Biegel, D., Sales, E., & Schulz, R. (1991). *Family caregiving in chronic illness.* Newbury Park, CA: Sage.

Bird, H., Canino, G., Rubio-Stipec, M., Gould, M. S., Ribera, J. C., Sesman, M., Woodbery, M., Heutos-Goldman, S., Pagan, A., Sanchez-Lakay, A., & Moscosco, M. (1988). Estimates of the prevalence of childhood maladjustment in a community survey in Puerto Rico. *Archives of General Psychiatry, 45,* 1120–1126.

Brandenberg, N. A., Friedman, R. M., & Silver, S. E. (1987). The epidemiology of childhood psychiatric disorders: Prevalence findings from recent studies. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 76–83.

Brannan, A. M., Heflinger, C. A., & Bickman, L. (1997). The Caregiver Strain Questionnaire: Measuring the impact on the family of living with a child with serious emotional disturbance. *Journal of Emotional and Behavioral Disorders, 5,* 121–222.

Brotherson, M. J., & Goldstein, B. L. (1992). Time as a resource and constraint for parents of young children with disabilities: Implications for early intervention services. *Topics in Early Childhood Special Education, 12,* 508–527.

Bruns, E. J., Burchard J. D., & Yoe, J. T. (1995). Evaluating the Vermont system of care: Outcomes associated with community-based wraparound services. *Journal of Child and Family Services, 4,* 321–340.

Burns, B. J., & Friedman, R. M. (1989). The research base for child mental health services and policy: How solid is the foundation? In P. Greenbaum, R. Friedman, A. Duchnowski, K. Kutash, & S. Silver (Eds.), *Conference proceedings: Children's mental health services and policy: Building a research base* (pp. 7–13). Tampa, FL: Research and Training Center for Children's Mental Health.

Clark, H. B., Lee, B., Prange, M. E., & McDonald, B. A. (1996). Children lost within the fostercare system: Can wraparound service strategies improve placement outcomes? *Journal of Child and Family Studies, 5,* 39–54.

Clausen, J., & Yarrow, M. (1955). Mental illness and the family. *Journal of Social Issues, 11,* 3–5.

Cohen, R., Parmelee, D. X., Irwin, L., Weisz, J. R., Howard, P., Purcell, P., & Best, A. M. (1990). Characteristics of children and adolescents in a psychiatric hospital and a corrections facility. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 909–913.

Collins, B., & Collins, T. (1990). Parent–professional relationships in the treatment of seriously emotionally disturbed children and adolescents. *Social Work, 35,* 522–527.

Costello, E. J., Edelbrock, C. S., Costello, A. J., Dulcan, M. K., Burns, B., & Brent, D. (1988). Psychiatric disorders in pediatric primary care: The new hidden morbidity. *Pediatrics, 82,* 415–424.

Cullinan, D., Epstein, M. H., & Sabornie, E. J. (1992). Selected characteristics of a national sample of seriously emotionally disturbed adolescents. *Behavioral Disorders, 17,* 273–280.

Cullinan, D., Epstein, M. H., & Quinn, K. P. (1996). Patterns and correlates of personal, family, and prior placement variables in an interagency community based system of care. *Journal of Child and Family Studies, 5,* 299–321.

Cullinan, D., Gadow, K. D., & Epstein, M. H. (1987). Psychotropic drug treatment among learning-disabled, educable mentally retarded, and seriously emotionally disturbed students. *Journal of Abnormal Child Psychology, 15,* 469–477.

Curtis, W. J., & Singh, N. N. (1996). Family involvement and empowerment in mental health service provision for children with emotional and behavioral disorders. *Journal of Child and Family Studies, 5,* 503–517.

Curtis, W. J., Singh, N. N., & Cohen, R. (1996). Measuring perceptions of family involvement in service provision for youth with serious emotional disturbance. In C. Liberton, K. Kutash, & R. Friedman (Eds.), *The eighth annual research conference proceedings, A system of care for children's mental health: Expanding the research base*. (pp. 105–109). Tampa, FL: University of South Florida, Florida Mental Health Institute.

Davis, M., Yelton, S., Katz-Leavy, J., & Lourie, I. S. (1995). Unclaimed children revisited: The status of state mental health service systems. *Journal of Mental Health Administration, 22,* 147–166.

DeChillo, N. (1993). Collaboration between social workers and families of patients with mental illness. *Families in Society, 74,* 104–115.

DeChillo, N., Koren, P. E., & Schultze, K. H. (1994). From paternalism to partnership: Family and professional collaboration in children's mental health. *American Journal of Orthopsychiatry, 64,* 564–576.

Donner, R. Huff, B., Gentry, M., McKinney, D., Duncan, J., Thompson, S., & Silver, P. (1995). Expectations of case management for children with emotional problems. In B. J. Friesen & J. Poertner (Eds.), *From case management to service coordination for children with emotional, behavioral, or mental disorders: Building on family strengths* (pp. 27–36). Baltimore, MD: Paul Brookes.

Duchnowski, A. J., Johnson, M. K., Hall, K. S., Kutash, K., & Friedman, R. M. (1993). The alternatives to residential treatment study: Initial findings. *Journal of Emotional and Behavioral Disorders, 1,* 17–25.

Dunst, C. J., & Paget, K. D. (1993). Parent–professional partnerships and family empowerment. In M. J. Fine (Ed.), *Collaboration with parents of exceptional children* (pp. 25–44). Brandon, VT: Clinical Psychology Publishing.

Dunst, C. J., Trivette, C. M., Davis, M., & Cornwell, J. (1988). Enabling and empowering families of children with health impairments. *Children Health Care, 17,* 71–81.

Eber, L., Osuch, R., & Reditt, C. A. (1996). School-based applications of the wraparound process: Early results on service provision and student outcomes. *Journal of Child and Family Studies, 5,* 83–99.

Elliott, D. J., Koroloff, N. M., Koren, P. E., & Friesen, B. J. (1998). Improving access to children's mental health services: The family associate approach. In M. H. Epstein, K. Kutash, & A. Duchnowski (Eds.), *Outcomes for children and youth with behavioral and emotional disorders and their families: Programs and evaluation best practices* (pp. 581–605). Austin, TX: Pro-Ed.

Epstein, M. H., Cullinan, D., Quinn, K. P., & Cumbald, C. (1994). Characteristics of children with emotional and behavioral disorders in community-based programs designed to prevent placement in residential facilities. *Journal of Emotional and Behavioral Disorders, 2,* 51–57.

Epstein, M. H., Cullinan, D., Quinn, K. P., & Cumbald, C. (1995). Personal, family, and service use characteristics of young people served by an interagency community-based system of care. *Journal of Emotional and Behavioral Disorders, 3,* 55–64.

Evans, M. E., Dollard, N., & McNulty, T. L. (1992). Characteristics of seriously emotionally disturbed youth with and without substance abuse in intensive case management. *Journal of Child and Family Studies, 1,* 305–314.

Fadiman, A. (1997). *The spirit catches you and you fall down*. New York: Farrar, Straus and Giroux.

Farmer, E. M. Z., Burns, B. J., Angold, A., & Costello, E. J. (1997). Impact of children's mental health problems on families: Relationships with service use. *Journal of Emotional and Behavioral Disorders, 5,* 230–238.

Federation of Families for Children's Mental Health. (1992). *Family support statement*. Alexandria, VA: Author.

Foster, E. M., Saunders, R. C., & Summerfelt, W. T. (1996). A comparison of multiple methods of predicting level of care in mental health services. *Evaluation and Program Planning, 19,* 143–153.

Freund, P. D. (1993). Professional role(s) in the empowerment process: Working with mental health consumers. *Psychosocial Rehabilitation Journal, 16,* 65–73.

Friesen, B. J. (1989). Parents as advocates for children and adolescents with serious emotional handicaps: Issues and directions. In R. M. Friedman, A. J. Duchnowski, & E. L. Henderson (Eds.), *Advocacy on behalf of children with serious emotional problems* (pp. 68–78). Springfield, IL: Charles C. Thomas.

Friesen, B. J., & Huff, B. (1996). Family perspectives on systems of care. In B. A. Stroul (Ed.), *Children's mental health: Creating systems of care in a changing society* (pp. 41–67). Baltimore, MD: Paul Brookes.

Friesen, B. J. & Koroloff, N. M. (1990). Family-centered services: Implications for mental health administration and research. *Journal of Mental Health Administration, 17,* 13–25.

Friesen, B. J., Koren, P. E., & Koroloff, N. M. (1992). How parents view professional behaviors: A cross-professional analysis. *Journal of Child and Family Studies, 1,* 209–231.

Friesen, B. J., & Wahlers, D. (1993). Respect and real help: Family support and children's mental health. *Journal of Emotional and Behavioral Problems, 2,* 12–14.

Fullan, M. (1982). *The meaning of educational change*. New York: Teachers College Press.

Harry, B. (1992). *Cultural diversity, families and the special education system: Communication and empowerment*. New York: Teachers College Press.

Hasenfeld, Y. (1987). Power in social work practice. *Social Service Review, 61,* 469–483.

Hasenfeld, Y., & Chesler, M. A. (1989). Client empowerment in the human services: Personal and professional agenda. *Journal of Applied Behavioral Sciences, 25,* 499–521.

Hawkins, N. E., & Singer, G. H. S. (1989). A skills training approach for assisting parents to cope with stress. In G. H. S. Singer & L. K. Irvin (Eds.), *Support for caregiving families: Enabling positive adaptation to disability* (pp. 71–83). Baltimore, MD: Paul Brookes.

Hawkins, R. P., Almedia, C., & Samet, M. (1989). Comparative evaluation of foster-family-based treatment and five other placement choices: A preliminary report. In A. Algarin, R. M. Friedman, A. J. Duchnowski, K. M. Kutash, S. E. Silver, & M. K. Johnson (Eds.), *Second annual conference proceedings of the Research and Training Center for Children's Mental Health: Building a research base* (pp. 98–119). Tampa, FL: Florida Mental Health Institute.

Hobbs, N., Dokecki, P. R., Hoover-Dempsey, K. V., Moroney, R. M., Shayne, M. W., & Weeks, K. H. (1984). *Strengthening families*. San Francisco: Jossey-Bass.

Hoenig, J., & Hamilton, M. W. (1967). The burden on the household in extramural psychiatric service. In J. Farndale & H. Freeman (Eds.), *New aspects of the mental health services* (pp. 612–635). London: Pergamon.

Hyde, K. L., Burchard, J. D., & Woodworth, K. (1996). Wrapping services in an urban setting. *Journal of Child and Family Studies, 5,* 67–82.

Institute of Medicine. (1989). *Research on children and adolescents with mental, behavioral, and developmental disorders: Mobilizing national initiative*. Washington, DC: National Academy Press.

Jay, D. E., & Padilla, C. L. (1987). *Special education dropouts*. Menlo Park, CA: SRI International.

Julian, T. W., McKenry, P. C., & McKelvey, M. W. (1994).

Cultural variations in parenting: Perceptions of Caucasian, African American, Hispanic, and Asian American parents. *Family Relations, 43,* 30–37.

Kalyanpur, M., & Harry, B. (1997). A posture of reciprocity: A practical approach to collaboration between professionals and parents of culturally diverse backgrounds. *Journal of Child and Family Studies, 6,* 487–509.

Kauffman, J. M. (1997). *Characteristics of emotional and behavioral disorders of children and youth* (6th ed.). Upper Saddle River, NJ: Merrill.

Kleinman, A., Eisenberg, L., & Good, B. (1978). Culture, illness, and care: Clinical lessons from anthropologic and cross-cultural research. *Annals of Internal Medicine, 88,* 251–258.

Knitzer, J. (1982). *Unclaimed children: The failure of public responsibility to children and adolescents in need of mental health services.* Washington, DC: Children's Defense Fund.

Kopp, J. (1989). Self-observation: An empowerment strategy in assessment. *Social Casework, 70,* 276–284.

Koren, P. E., DeChillo, N., & Friesen, B. J. (1992). Measuring empowerment in families whose children have emotional disabilities: A brief questionnaire. *Rehabilitation Psychology, 37,* 305–321.

Koroloff, N. M., & Friesen, B. J. (1991). Support groups for parents of children with emotional disorders: A comparison of members and nonmembers. *Community Mental Health Journal, 27,* 265–279.

Landrum, T. J., Singh, N. N., Nemil, M. S., Ellis, C. R., & Best, A. M. (1995). Characteristics of children and adolescents with serious emotional disturbance in systems of care. Part II: Community-based services. *Journal of Emotional and Behavioral Disorders, 3,* 141–148.

Lee, E. (1997). *Working with Asian Americans: A guide for clinicians.* New York: Guilford.

Lefley, H. P. (1989). Family burden and family stigma in major mental illness. *American Psychologist, 44,* 556–560.

Lehman, D. R., Ellard, J. H., & Wortman, C. B. (1986). Social support for the bereaved: Recipients' and providers' perspectives on what is helpful. *Journal of Consulting and Clinical Psychology, 54,* 438–446.

Lundy, M., & Pumariega, A. J., (1993). Psychiatric hospitalization of children and adolescents: Treatment in search of a rationale. *Journal of Child and Family Studies, 2,* 1–4.

Mattison, R. E., Morales, J., & Bauer, M. A. (1992). Distinguishing characteristics of elementary school boys recommended for SED placement. *Behavioral Disorders, 17,* 107–114.

Messer, S. C., Angold, A., Costello, E. J., Burns, B. J., Farmer, E. M. Z., & Patrick, M. K. S. (1996). The Child and Adolescent Burden Assessment (CABA): Measuring the family impact of emotional and behavioral problems. *International Journal of Methods in Psychiatric Research, 6,* 261–284.

Mittler, P., Mittler, H., & McCanachie, H. (1987). Working together: Guidelines for partnership between professionals and parents of children and young people with disabilities. In D. K. Lipsky (Ed.), *Family supports for families with a disabled member* (pp. 15–35). New York: International Exchange of Experts and Information in Rehabilitation.

Montgomery, R. J. V., Gonyea, J. G., & Hooyman, N. R. (1985). Caregiving and the experience of subjective and objective burden. *Family Relations, 34,* 19–26.

Nelson, C. M., & Pearson, C. A. (1991). *Integrating services for children and youth with emotional and behavioral disorders.* Reston, VA: Council for Exceptional Children.

Patrick, M., Angold, A., Burns, B., & Costello, E. J. (1992). *The Child and Adolescent Burden Assessment (CABA).* Durham, NC: Duke University Medical Center, Department of Psychiatry and Behavioral Sciences, Developmental Epidemiology Program.

Parsons, R. J. (1991). Empowerment: Purpose and practice principles in social work. *Social Work with Groups, 14,* 7–21.

Platt, S. (1985). Measuring the burden of psychiatric illness on the family: An evaluation of some rating scales. *Psychological Medicine, 15,* 383–393.

Quinn, K. P., & Epstein, M. H. (1998). Characteristics of children, youth, and families served by local interagency systems of care. In M. H. Epstein, K. Kutash, & A. Duchnowski (Eds.), *Outcomes for children and youth with behavioral and emotional disorders and their families: Programs and evaluation best practices* (pp. 81–114). Austin, TX: Pro-Ed.

Quinn, K. P., Epstein, M. H., Dennis, K., Potter, K., Sharma, J., McKelvey, J., & Cumbald, C. (1996). Personal, family, and service utilization characteristics of children served in an urban family preservation environment. *Journal of Child and Family Studies, 5,* 469–486.

Quinn, K. P., Newman, D. L., & Cumbald, C. (1995). Behavioral characteristics of children and youth at risk for out-of-home placements. *Journal of Emotional and Behavioral Disorders, 3,* 166–173.

Rappaport, J. (1981). In praise of paradox: A social policy of empowerment over prevention. *American Journal of Community Psychology, 9,* 1–25.

Regional Research Institute for Human Services. (1992). *Effects of situation.* Portland, OR: Oregon Partners Project Evaluation, Portland State University.

Rosenblatt, A., & Attkisson, C. C. (1992). Integrating systems of care in California for youth with severe emotional disturbance. I: A descriptive overview of the California AB377 evaluation project. *Journal of Child and Family Studies, 1,* 93–113.

Rowe, K. (1991). *Study of groups for parents of children with serious emotional disorders in Virginia.* Richmond, VA: Department of Mental Health, Mental Retardation and Substance Abuse Services.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action.* New York: Basic Books.

Silver, S. E., Duchnowski, A. J., Kutash, K., Friedman, R. M., Eisen, M., Prange, M. E., Brandenburg, N. A., & Greenbaum, P. E. (1992). A comparison of children with serious emotional disturbances served in residential and school settings. *Journal of Child and Family Studies, 1,* 43–59.

Singer, G. H. S., & Irvin, L. K. (1989). *Support for caregiving families: Enabling positive adaptation to disability.* Baltimore, MD: Paul Brookes.

Singer, G. H. S., & Irvin, L. K. (1991). Supporting families of persons with severe disabilities: Emerging findings, practices, and questions. In L. H. Meyer, C. A. Peck, & L. Brown (Eds.), *Critical issues in the lives of people with severe disabilities* (pp. 271–312). Baltimore, MD: Paul Brookes.

Singh, N. N. (1995). In search of unity: Some thoughts on family–professional relationships in service delivery systems. *Journal of Child and Family Studies, 4,* 3–18.

Singh, N. N. (1996). Cultural diversity in the 21st century: Beyond *E Pluribus Unum. Journal of Child and Family Studies, 5,* 121–136.

Singh, N. N. (1997, September). *Person-centered planning in mental health: A holistic approach to service delivery.* Invited keynote address presented at the 7th Annual Virginia Beach Conference on Children and Adolescents with Emotional and Behavioral Disorders, Virginia Beach, Virginia.

Singh, N. N. (1998). Cultural diversity: A challenge for evaluating systems of care. In M. H. Epstein, K. Kutash,

& A. Duchnowski(Eds.), *Outcomes for children and youth with behavioral and emotional disorders and their families: Programs and evaluation best practices* (pp. 425–454). Austin, TX: Pro-Ed.

Singh, N. N., Curtis, W. J., Cohen, R., Nicholson, M. W., Villani, T. M., & Wechsler, H. A. (1996a). *Measuring perceptions of family involvement. I: The Family Involvement Scale-Family Version*. Richmond, VA: Commonwealth Institute for Child and Family Studies, Virginia Commonwealth University.

Singh, N. N., Curtis, W. J., Cohen, R., Nicholson, M. W., Villani, T. M., & Wechsler, H. A. (1996b). *Measuring perceptions of family involvement. II: The Family Involvement Scale-Professional Version*. Richmond, VA: Commonwealth Institute for Child and Family Studies, Virginia Commonwealth University.

Singh, N. N., Curtis, W. J., Ellis, C. R., Villani, T., Nicholson, M. W., & Wechsler, H. A. (1995). Psychometric analysis of the Family Empowerment Scale. *Journal of Emotional and Behavioral Disorders, 3,* 85–91.

Singh, N. N., Curtis, W. J., Ellis, C. R., Wechsler, A. H., Best, A. M., & Cohen, R. (1997). Empowerment status of families whose children have serious emotional disturbance and attention-deficit hyperactivity disorder. *Journal of Emotional and Behavioral Disorders, 5,* 223–229.

Singh, N. N., Curtis, W. J., Wechsler, H. A., Cohen, R., & Ellis, C. R. (1997). Family friendliness of community-based services for children and adolescents with emotional and behavioral disorders: An observational study. *Journal of Emotional and Behavioral Disorders, 5,* 82–92.

Singh, N. N., Ellis, C. R., Oswald, D. R., Wechsler, H. A., & Curtis, W. J. (1977). Addressing and valuing diversity. *Journal of Emotional and Behavioral Disorders, 5,* 24–35.

Singh, N. N., Landrum, T. J., Donatelli, L. S., Hampton, C., & Ellis, C. R. (1994). Characteristics of children and adolescents with serious emotional disturbance in systems of care. Part I: Partial hospitalization and inpatient psychiatric services. *Journal of Emotional and Behavioral Disorders, 2,* 13–20.

Singh, N. N., Williams, E., & Spears, N. (in press). Value and address diversity: From policy to practice. In US Department of Education, OSEP (Ed.), *Improving results for children and youth with serious emotional disturbance*. Washington, DC: Chesapeake Institute.

Sluyter, G. V. (1994). Creating a vision for mental health services: A survey of states. *Administration and Policy in Mental Health, 21,* 247–250.

Staples, L. H. (1990). Powerful ideas about empowerment. *Administration in Social Work, 14,* 29–42.

Stein, R. E. K., & Riessman, C. K. (1980). The development of an Impact-on-Family Scale: Preliminary findings. *Medical Care, 18,* 465–472.

Stroul, B., & Friedman, R. M. (1986). *A system of care for children and adolescents with severe emotional disturbances*. Washington, DC: Georgetown University Child Development Center, National Technical Assistance Center for Children's Mental Health.

Tarico, V. S., Low, B. P., Trupin, E. & Forsyth-Stevens, A. (1989). Children's mental health services: A parent perspective. *Community Mental Health Journal, 25,* 313–326.

Telleen, S., Herzog, A., & Kilbane, T. L. (1989). Impact of a family support program on mother's social support and parenting stress. *American Journal of Orthopsychiatry, 59,* 410–419.

Trupin, E. W., Tarico, V. S., Low, B. P., Jemelka, R., & McCellan, J. (1993). Children on child protective service caseloads: Prevalence and nature of serious emotional disturbance. *Child Abuse and Neglect, 17,* 345–355.

US Department of Education. (1994). *Sixteenth annual report to Congress on the implementation of Individuals With Disabilities Education Act*. Washington, DC: US Government Printing Office.

Valdes, K. A., Williamson, C. L., & Wagner, M. M. (1990). *The national longitudinal transition study of special education students: Statistical almanac. Vol. 3: Youth categorized as emotionally disturbed*. Menlo Park, CA: SRI International.

VanDenBerg, J. E., & Grealish, E. M. (1996). Individualized services and supports through the wraparound process: Philosophy and procedures. *Journal of Child and Family Studies, 5,* 7–21.

Wagner, M., Blackorby, J., Cameto, R., Hebbeler, K., & Newman, L. (1993). *The transition experiences of youth with disabilities: A summary of findings from the National Longitudinal Transition Study*. Palo Alto, CA: SRI.

Wells, K., & Whittington, D. (1993). Characteristics of youths referred to residential treatment: Implications for program design. *Children and Youth Review, 15,* 195–217.

Yoe, J. T., Santarcangelo, S., Atkins, M., & Burchard, J. D. (1996). Wraparound care in Vermont: Program development, implementation, and evaluation of a statewide system of individualized services. *Journal of Child and Family Studies, 5,* 23–37.

Zeigler-Dendy, C. A. (1989). The invisible children project: Methods and findings. In A. Algarin, R. M. Friedman, A. J. Duchnowski, K. M. Kutash, S. E. Silver, & M. K. Johnson (Eds.), *Second annual conference proceedings of the Research and Training Center for Children's Mental Health: Children's mental health services and policy: Building a research base* (pp. 360–366). Tampa, FL: Florida Mental Health Institute.

Zimmerman, M. A., & Rappaport, J. (1988). Citizen participation, perceived control, and psychological empowerment. *American Journal of Community Psychology, 16,* 725–750.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.16
Family Lives of Lesbians and Gay Men

CHARLOTTE J. PATTERSON
University of Virginia, Charlottesville, VA, USA

9.16.1 INTRODUCTION 253
9.16.2 FAMILIES OF LESBIANS AND GAY MEN TODAY 254
  9.16.2.1 Estimates of the Lesbian and Gay Population 254
  9.16.2.2 Diversity Among Lesbians, Gay Men, and Their Families 255
  9.16.2.3 Legal and Public Policy Issues 256
9.16.3 RESEARCH ON LESBIAN AND GAY COUPLE RELATIONSHIPS 256
  9.16.3.1 Descriptions of Lesbian and Gay Couples 257
  9.16.3.2 Love and Commitment 257
  9.16.3.3 Power and Division of Labor 258
  9.16.3.4 Sexual Behavior 258
  9.16.3.5 Problems and Conflict in Couples 259
9.16.4 RESEARCH ON LESBIAN MOTHERS AND GAY FATHERS 260
  9.16.4.1 Lesbians and Gay Men Who Became Parents in the Context of Heterosexual Relationships 260
    9.16.4.1.1 Divorced lesbian mothers 260
    9.16.4.1.2 Divorced gay fathers 261
  9.16.4.2 Lesbians and Gay Men Choosing to Become Parents 262
9.16.5 RESEARCH ON CHILDREN OF LESBIAN AND GAY PARENTS 264
  9.16.5.1 Research on Children Born in the Context of Heterosexual Relationships 264
    9.16.5.1.1 Sexual identity 264
    9.16.5.1.2 Other aspects of personal development 265
    9.16.5.1.3 Social relationships 265
  9.16.5.2 Diversity Among Children With Divorced Lesbian or Gay Parents 266
  9.16.5.3 Research on Children Born to or Adopted by Lesbian Mothers 267
9.16.6 RESEARCH ON OTHER FAMILY RELATIONSHIPS 267
9.16.7 IMPLICATIONS OF RESEARCH FINDINGS 268
9.16.8 DIRECTIONS FOR RESEARCH 269
9.16.9 REFERENCES 270

9.16.1 INTRODUCTION

> In 1990, Ninia Baehr went on a blind date with another young woman, Genora Dancel. It was love at first sight. Later that year, they applied for a marriage license in Honolulu, where they were living. They were denied. Joining with a gay couple and another lesbian couple who had been similarly denied, they filed a lawsuit asking the State of Hawaii to grant them a marriage license. (Adapted from Bull & Gallagher, 1996, pp. 197–200)

Thus began the case that has become known as *Baehr v. Miike*, which is often called the Hawaii same-sex marriage case. In 1993, the Hawaii Supreme Court declared that the denial of marriage licenses to the three couples appeared to violate the state's prohibition on gender-based discrimination, and remanded the case to a lower court. In order to maintain the ban, the state would be required to prove that it served a compelling state interest. When the case was heard in 1996, the state's attorney argued that if same-sex couples could form legal marriages, then they would be likely to have more children, and this would harm children. Same-sex marriages would thus have an adverse impact on the welfare of children, the state's attorney argued, and this justified the ban. The trial judge's decision rejected the state's argument, holding that sexual orientation and parental fitness are not related, and struck down the prohibition on same-sex marriage. At the time of this writing, in 1997, the case is on appeal to the Hawaii Supreme Court.

The Baehr case has generated enormous controversy, not only in Hawaii, but also across the USA and around the world. Conservatives argued that legal marriages between same-sex couples would undermine the institution of marriage, and introduced into the US Congress a "Defense of Marriage Act" (DOMA) which said that states were not required to recognize as legal same-sex marriages contracted in another state. The bill passed, and was quickly signed into law. Meanwhile, a 1994 survey of a large group of gay men found that 85% said they would marry a same-sex partner if they were legally able to do so (Bull & Gallagher, 1996, p. 214). Religious groups also entered the debate. On one side, the Vatican declared that same-sex marriages would "undermine the foundation of the family model upon which human civilization was built" and would legitimize "moral disordertates." At the same time, other religious groups, such as the Central Conference of American Rabbis, voted to support legalization of same-sex marriages (Dunlap, 1996).

What is it that makes a lawsuit like *Baehr v. Miike* so controversial? This chapter addresses three inter-related reasons. First, the Baehr case involves controversies about the legitimacy of families formed by lesbians and gay men. Second, although widespread prejudices suggest that lesbians and gay men live outside the reach of family pleasures and responsibilities, the actual reality as illuminated by research is quite different. Third, if the family lives of lesbians and gay men are deemed acceptable under the law, the challenge to deeply held ideas about families and about child development is undoubtedly significant. Thus, when considering whether marriages like the one sought by Ninia Baehr and Genora Dancel are legitimate, much is at stake.

What are the findings of social science research about lesbian and gay family life? In this chapter, information about the prevalence and diversity of lesbian and gay family life, as well as about the legal context in which lesbian and gay families currently live, are first reviewed. The results of research on lesbian and gay couples, parents and their children, and other family members are then described and the implications of research findings for psychological theory as well as public policy are discussed. A discussion of directions for further research concludes the chapter.

9.16.2 FAMILIES OF LESBIANS AND GAY MEN TODAY

How many lesbian and gay people are there in the USA, and how many people are included in their families? What are important sources of diversity among these families? What is the nature of the legal context in which lesbians, gay men, and their families are living today?

9.16.2.1 Estimates of the Lesbian and Gay Population

For many reasons, no accurate count of lesbian and gay adults, much less of their families, is available. The numbers of lesbian and gay adults in the USA cannot be estimated with confidence. Because of fear of discrimination, many take pains to conceal their sexual orientation (Blumenfeld & Raymond, 1988). It is especially difficult to locate gay and lesbian parents. Fearing that they would lose child custody and/or visitation rights if their sexual orientation were to be known, many lesbian and gay parents attempt to conceal their gay or lesbian identities (Pagelow, 1980). It can also be difficult to locate other relatives of lesbians and gay men (e.g., siblings), because some may not be aware of having a lesbian or gay relative, and because others may not wish to acknowledge this aspect of their family life.

One approach to making estimates of this kind is to extrapolate from what is known or believed about base rates in the population. For example, there are approximately 260 million people in the USA (US Bureau of the Census, 1995). Based on data collected by the National Health and Social Life Survey (Laumann, Gagnon, Michael, & Michaels, 1994), 1.4% of women and 2.8% of men identified themselves to an interviewer as lesbian or gay. Using these numbers, one can estimate that there are approximately 5.5 million people who identify as lesbian or gay in the USA. If, on average, each one has five relatives (e.g., two parents, one sibling, and two grandparents), then there are at least 27 million heterosexual people with a lesbian or gay relative. That would mean that more than 32 million Americans either identify as lesbian or gay, or have a close relative who identifies in this way.

The accuracy of such estimates is, of course, no better than the accuracy of the numbers on which they are based. Laumann et al. (1994) also reported that 4.1% of women and 4.9% of men reported at least one same-sex sexual partner since 18 years of age. If estimates are based not on self-identification as lesbian or gay, but on numbers of adults who reported at least one same-sex sexual partner, then population estimates for lesbians, gay men, and their families will of course become larger. Whatever the precise numbers, it is clear that, when one considers the families of lesbians and gay men, very substantial numbers of people are under discussion.

9.16.2.2 Diversity Among Lesbians, Gay Men, and Their Families

Although lesbians, gay men, and their family members are often stereotyped in very specific ways, there is tremendous diversity among individuals and families. Just as other families embrace many variations, so the families of lesbians and gay men vary in many different ways from one another. Some of the variation is similar to that among heterosexual families, and some is more specific to the families of lesbians and gay men.

As is true for heterosexual families, the families of lesbians and gay men vary considerably on demographic characteristics. Lesbians, gay men, and their families may be of any racial or ethnic group, any cultural background, and they may or may not have had the benefit of extensive educational opportunities. Lesbians and gay men can be found in all economic strata (Badgett, in press), and in many different occupations. Although evidence suggests that men who identify as gay are more likely to live in urban than in suburban or rural environments, the geographic dispersal of lesbian women seems to be less pronounced (Laumann et al., 1994). Family members may live near their lesbian or gay relatives, or may live at some distance. Much of this type of variation is held in common with the families of heterosexual individuals.

Among lesbian and gay families with children, one important distinction concerns the sexual identity of parents at the time of a child's birth or adoption. Probably the largest group of children with lesbian and gay parents today are those who were born in the context of heterosexual relationships between the biological parents, and whose parent or parents subsequently identified as gay or lesbian. These include families in which the parents divorce when the husband comes out as gay, the wife comes out as lesbian, both parents come out, and one or both of the parents comes out and the parents decide not to divorce. Gay or lesbian parents may be single, or they may have same-sex partners. A gay or lesbian parent's same-sex partner may or may not assume step-parenting relationships with the children. If the partner also has children, the youngsters may assume step-sibling relationships with one another. In short, gay and lesbian families with children born in the context of heterosexual relationships are themselves a relatively diverse group.

In addition to children born in the context of heterosexual relationships between parents, lesbians and gay men are believed increasingly to be choosing parenthood (Martin, 1993; Patterson, 1994a, 1994b; Pies, 1985, 1990). The majority of such children are conceived by means of donor insemination (DI). Lesbians who wish to bear children may choose a friend, relative, or acquaintance to be the sperm donor, or they may choose to use sperm from an unknown donor. When sperm donors are known, they may take parental or avuncular roles relative to children conceived via DI, or they may not (Martin, 1993; Patterson, 1994a, 1994b; Pies, 1985, 1990). Gay men may also become biological parents of children whom they intend to parent, whether with a single woman (who may be lesbian or heterosexual), with a lesbian couple, or with a gay male partner. Options pursued by gay men and lesbians also include both adoption and foster care (Ricketts, 1991). Thus, children are today being brought up in a diverse array of lesbian and gay families.

In addition to differences in parents' sexual identities at the time of a child's birth, another distinction among the families of lesbians and gay men is the extent to which family members are related biologically to one another (Pollack & Vaughn, 1987; Riley, 1988; Weston, 1991). Although biological relatedness of family members to one another in heterosexual families is probably less often taken for granted than it once was, it is often even more prominent as an issue in lesbian and gay families with children. When children are born via DI into lesbian families, they are generally related biologically only to the birthmother, not to her partner. Similarly, when children are born via surrogacy to a gay couple, only the father who served as a sperm donor is biologically related to the child. In adoption and foster care, of course, the child will probably have no biological relation to any adoptive or foster parent.

Another distinction of particular importance for lesbian and gay families concerns custodial arrangements for minor children following the break-up of heterosexual marriages. As in heterosexual families, children generally live with one of their biological parents, or they spend part of their time in one biological parent's household, and part of their time in the other's home. Many lesbian mothers and gay fathers have, however, lost custody of their children to heterosexual spouses following divorce, and the threat of custody litigation almost certainly looms larger in the lives of most divorced lesbian mothers than it does in the lives of divorced heterosexual ones (Lyons, 1983; Pagelow, 1980). Although no authoritative figures are available, it seems likely that a greater proportion of gay and lesbian than heterosexual parents lost custody of children against their will. Probably for this reason, more lesbians and gay men are noncustodial parents (i.e., do not have legal custody of their children) and nonresidential parents (i.e., do not live in the same household with their children) than might otherwise be expected.

9.16.2.3 Legal and Public Policy Issues

The legal system in the USA has long been hostile to lesbians and gay men, and especially to those who are or who wish to become parents (Editors of the Harvard Law Review, 1990; Falk, 1989; Patterson & Redding, 1996; Polikoff, 1990; Rivera, 1991; Rubenstein, 1991, 1996). Same-sex marriages are currently not legal in any jurisdiction in the USA. Lesbian mothers and gay fathers have often been denied custody and/or visitation with their children following divorce from heterosexual marriages. Although some states now have laws stipulating that parental sexual orientation as such cannot be a factor in determining child custody following heterosexual divorce, in other states gay or lesbian parents may be regarded as unfit. Regulations governing foster care and adoption in many states have also made it difficult for lesbians or gay men to adopt or to serve as foster parents (Ricketts, 1991; Ricketts & Achtenberg, 1990).

One of the central issues underlying judicial decision-making in custody litigation and in public policies governing foster care and adoption has been questions about the fitness of lesbians and gay men to be parents (Patterson, 1995c; Patterson & Redding, 1996). Specifically, policies have sometimes been constructed and judicial decisions have often been made on the assumptions that gay men and lesbians are mentally ill and hence not fit to be parents, that lesbians are less maternal than heterosexual women and hence do not make good mothers, and that lesbians' and gay men's relationships with sexual partners leave little time for ongoing parent–child interactions (Editors of the Harvard Law Review, 1990; Falk, 1989). Because these assumptions have been important ones in limiting gay and lesbian parental rights, and because they are open to empirical evaluation, they have guided much of the research on lesbian and gay parents that is discussed below.

In addition to judicial concerns about parents themselves, three principal kinds of fears regarding the effects of gay and lesbian parents on children have also been reflected in judicial decision-making about child custody and in public policies such as regulations governing foster care and adoption (Patterson, 1992, 1995c; Patterson & Redding, 1996). One of these concerns is that development of sexual identity among children of lesbian and gay parents will be impaired. For example, judges may fear that children will themselves grow up to be gay or lesbian, an outcome which they generally view as negative. Another is that gay and lesbian parents will have adverse effects on other aspects of their children's personal development. For example, judges may fear that children in the custody of gay or lesbian parents will be more vulnerable to behavior problems or to mental breakdown. A third concern is that these children will have difficulties in social relationships. One version of such concern is the fear that children will be teased or stigmatized by peers because of the sexual orientation of their parents. Because such concerns have often been explicit in judicial determinations when lesbian or gay parents" custody or visitation rights have been denied or curtailed, and because these assumptions are open to empirical test, they have provided an important impetus to research.

9.16.3 RESEARCH ON LESBIAN AND GAY COUPLE RELATIONSHIPS

Research on lesbian and gay couples is a phenomenon of the last 25 years, and has addressed a number of interrelated issues. In this section, descriptive information from a large-scale survey of lesbian and gay couples is first summarized, then research findings on love and commitment, power and the division of labor, sexual behavior, problems and conflict in relationships, and the ending of couple relationships are described.

9.16.3.1 Descriptions of Lesbian and Gay Couples

The earliest large-scale study of lesbian and gay couples was conducted by Blumstein and Schwartz (1983). In their study, 1576 lesbian, 1938 gay, 3656 heterosexual married, and 653 heterosexual cohabiting (but unmarried) couples completed extensive questionnaires about topics ranging from money and jobs to love and sexual behavior. Despite the large sample, carefully drawn from diverse geographic areas in the USA, the lesbian and gay respondents were almost all white, between 21 and 40 years of age, mostly well educated, and mostly employed in professional or managerial occupations. Similar demographics have characterized other large-scale studies of lesbian and gay couples as well (Bryant & Demian, 1994; McWhirter & Mattison, 1984).

Bryant and Demian's (1994) study of lesbian and gay couples involved questionnaires completed by 709 lesbian and 557 gay couples who were drawn from 48 of the 50 states, Puerto Rico, and the District of Columbia. This group of respondents averaged about 35 years of age, had on average more than 16 years of education, and was heavily Caucasian. The average couple had been together more than five years, and lived together full-time. Most respondents reported having had one or two major lesbian/gay relationships before the one in which they were currently involved.

Bryant and Demian (1994) reported a number of interesting differences between lesbian and gay couples. For example, lesbian couples were more likely than gay couples to have incorporated ritual and symbolism into their relationships by having a commitment ceremony or by wearing rings to symbolize their commitment to one another. Gay couples reported having sexual relations more frequently than did lesbian couples. Most couples of both genders reported that their relationships were sexually exclusive or monogamous, and very few couples reported sex with partners outside their primary relationships. However, men were over-represented among those who were not monogamous or who reported sex outside the relationship. Further, men were more likely than women to call the person with whom they were involved their "lover," and women were more likely than men to call this person their "partner"; not one person of either gender mentioned using the term "longtime companion" to describe their lover or partner.

When asked to whom they looked for support, couples reported that friends and co-workers were more supportive than relatives (Bryant & Demian, 1994; Kurdek & Schmidt, 1987). Among relatives, siblings were considered most supportive, followed by mothers, and other relatives. Fathers were considered the least supportive among the couples" relatives. Overall, relatives were rated as the third most important challenge to couple relationships, after "communication" and "career."

A minority of couples reported acting in parental or other caregiving roles for children. Altogether, 21% of lesbian and 9% of gay couples reported caring for children. Most (75% of lesbians and 79% of gay men) had children from a previous marriage, and only 13% of lesbian mothers had children who had been conceived via DI. Interestingly, however, children were planned or under consideration by fully a third of all respondents under 35 years of age.

In summary, couples who participated in large-scale surveys have been largely well educated, relatively young, mainly Caucasian, and usually living together full-time. They were likely to regard friends and co-workers (rather than relatives) as their principal sources of support, and unlikely to have children living with them in the home. In addition, some differences between lesbian and gay couples have been described, notably in the couples" use of ritual and symbolism, sexual behavior, and in their names for one another. The extent to which these findings would be true of a more general cross-section of the population of lesbian and gay couples is unknown.

9.16.3.2 Love and Commitment

Many, if not most, lesbians and gay men express the desire for an enduring love relationship with a partner of the same gender. Indeed, research findings suggest that many are successful in creating such relationships. Survey data suggest that 40–60% of gay men and 45–80% of lesbians are currently involved in steady romantic relationships (see Peplau & Cochran, 1990; Peplau, Veniegas, & Campbell, 1996). Because most surveys involve many young adults, who may not have yet committed to stable romantic relationships, these numbers may underestimate the true figures.

When asked about their current relationship, lesbians and gay men report as much satisfaction with their relationships as do heterosexual couples; the great majority describe themselves as happy (Cardell, Finn, & Maracek, 1981; Kurdek & Schmidt, 1986a, 1986b; Peplau, Padesky, & Hamilton, 1982). For example, Peplau and Cochran (1990) surveyed 50 lesbians, 50 gay men, 50 heterosexual women, and 50 heterosexual men who were currently involved in romantic relationships. Both lesbians and gay men reported very positive feelings about their partners and rated their relationships as very satisfying. There were no differences as a function of sexual orientation on any of the measures of relationship quality (Peplau & Cochran, 1990).

Of course, not every lesbian or gay couple is happy together, and research has also focused on factors related to differences in relationship satisfaction. The correlates of relationship quality for lesbian and gay couples include feelings of having equal power, perceiving many attractions and few alternatives to the relationship, endorsing few dysfunctional beliefs about the relationship, placing a high value on the relationship, and engaging in shared decision-making (Blumstein & Schwartz, 1983; Kurdek, 1994, 1995).

9.16.3.3 Power and Division of Labor

How should power be allocated in a couple? The great majority of lesbian and gay couples feel that an equal balance of power would be desirable (Peplau & Cochran, 1990), but not all report that they achieve this ideal state. In Peplau and Cochran's study, only 59% of lesbians, 38% of gay men, 48% of heterosexual women, and 40% of heterosexual men reported that the balance of power in their current relationship was exactly equal. Others have found that majorities of gay as well as lesbian couples report equal power (see Peplau et al., 1996).

When power is unequal in a relationship, which partner has more power in an intimate relationship, and why? Social exchange theory predicts that the partner with greater personal resources (e.g., income, education) should have greater power (Peplau, 1991), and results of a number of studies have supported this view. For example, Harry (1984) found that older, wealthier men tended to have more power in their intimate relationships, and Caldwell and Peplau (1984), in a study of young lesbians, reported that wealthier and better educated women tended to have more power than their partners. Blumstein and Schwartz (1983) found that the partner with greater financial resources had more power in money management issues in gay, married heterosexual, and unmarried (but cohabiting) heterosexual couples, but not in lesbian couples. Whether or not relative financial resources affect the balance of power in lesbian couples remains an open question (see Peplau et al., 1996).

Other predictions from exchange theory have also received support from empirical research (Kurdek, 1995; Peplau, 1991; Peplau et al., 1996). In social exchange theory, the principle of least interest states that when one person is more dependent or involved than the other, the more dependent partner is expected to have less power (Peplau, 1991). In other words, the person who is less interested in continuing the relationship has more power. Consistent with this view, Caldwell and Peplau (1984) found correlations between unequal involvement and unequal power among lesbian couples. Overall, as predicted by social exchange theory, the woman who was less involved in the relationship had more power.

Although many people who are unfamiliar with lesbian and gay couples assume that, in same-sex couples, one partner plays a male and one a female role, research has consistently found that this is only rarely the case (Kurdek, 1995; Peplau et al., 1996). For example, Bell and Weinberg (1978) reported that the majority of lesbians and gay men they studied reported that they shared domestic tasks equally. When they were asked whether one partner does the feminine tasks while the other does the masculine tasks, about 90% of lesbians and gay men said that this was not the case in their households. Kurdek (1993) reported egalitarian divisions of labor among lesbian and gay couples, and Patterson (1995d) reported that in a sample of lesbian couples with children, most family and household tasks were shared equally.

9.16.3.4 Sexual Behavior

Sexual behavior among lesbian and gay couples has been found to be strongly shaped by gender. First, the frequency of genital sexual behavior has been reported to decline with the duration of a relationship, and this is true of reports given by lesbian, gay, heterosexual married, and unmarried (but cohabiting) heterosexual couples (Kurdek, 1995). These declines are less pronounced among gay and more pronounced among lesbian couples than among heterosexual couples, whether married or not. The frequency of genital sexual relations, as reported in surveys, thus appears to increase with the number of men in a couple. Critics of these findings have suggested that lesbian couples may have different definitions of sexual behavior than do gay men or heterosexual couples. As Blumstein and Schwartz (1983) reported, lesbians may place greater value on non-genital expressions of love, such as hugging and other affectionate behaviors.

Another area in which strong gender differences emerge is in the degree of a couple's desire for and accomplishment of sexual exclusivity. Lesbians and heterosexual couples have generally been found to be more supportive than gay men of monogamy in their relationships, and their reported behavior corresponds to these views. For example, Blumstein and Schwartz (1983) reported that, among couples who had been together between 2 and 10 years, most lesbian and heterosexual couples preferred and experienced monogamous sexual relationships, whereas most gay couples did not. These data were collected before the HIV/AIDS epidemic had attracted attention. However, data collected during 1988–89, after HIV infection and AIDS had become widespread in the USA, revealed the same pattern of results (Bryant & Demian, 1994).

Despite differences in preferences and in actual behavior, however, lesbian, gay, heterosexual married, and unmarried (but cohabiting) heterosexual couples all reported similar satisfaction with their sexual relationships (see also Bryant & Demian, 1994). Blumstein and Schwartz (1983) reported that, in each of the four couple types they studied, satisfaction with sex was associated with a perception of equality in initiating and refusing sexual overtures. Thus, although gender differences in sexual attitudes and sexual behavior would appear to be substantial, reported sexual satisfaction within couple relationships has not varied as a function of sexual orientation or gender of partners (Kurdek, 1995; Peplau, 1991).

9.16.3.5 Problems and Conflict in Couples

When lesbian and gay couples experience problems in their relationships, some of these stem from the same roots as those from which difficulties in heterosexual relationships also arise. As in heterosexual relationships, problems can arise due to different religious, racial, ethnic, or socioeconomic backgrounds of partners, and due to the different values that these backgrounds may have inculcated. Relationship difficulties can also arise as a result of problems at either partner's job, financial pressures on the couple, friction with members of extended family networks, and so forth, just as they do in heterosexual relationships. Kurdek (1992) found that the top five areas of conflict for lesbian and gay couples were finances, driving style, affection/sex, being overly critical, and division of household tasks.

There are some conflicts that are probably unique to lesbian and gay couples, and prominent among these are problems generated by negative social attitudes toward homosexuality. Because of prejudice and discrimination directed toward lesbians and gay men, many are unwilling to disclose their sexual identities to family members, neighbors, co-workers, and even to friends. When a couple disagrees about the extent to which they should disclose the lesbian or gay nature of their relationship, real difficulties in the couple relationship can ensue. Resolution of such conflicts may be central to the success of the couple relationship over time (Peplau et al., 1996).

The longevity of lesbian and gay relationships has also been a topic of some research. Blumstein and Schwartz (1983) found that for couples who had already been together 10 years, break-up rates over the 18 months of their study were low; only 6% of lesbian couples, 4% of gay couples, and 4% of married couples separated during this period. For couples who had been together less than 2 years, 22% of lesbian couples, 16% of gay couples, 17% of cohabiting (but unmarried) heterosexual couples, and only 4% of heterosexual married couples had separated. Thus, being married was associated with low break-up rates, but otherwise there were no significant differences. A later study (Kurdek, 1992, cited in Kurdek, 1995) also found no differences in break-up rates between lesbian and gay couples. In this study, break-up rates over a four-year period were 22% for lesbians and 12% for gay men, and this difference was not significant.

Kurdek and Schmidt (1986a) compared the attractions that a relationship held for the partners and also the barriers to exiting a relationship for lesbian, gay, unmarried (but cohabiting) heterosexual, and married heterosexual couples. They found no differences among these four types of couples in the strength of attractions toward their relationships, but did find significant differences in barriers to leaving the relationships. Specifically, married heterosexual spouses reported more obstacles to exiting the relationship than did members of the other three types of couples.

In addition to all the usual ways in which relationships can end, gay relationships in particular have, since the early 1980s, been subject to unusual stresses due to the HIV/AIDS epidemic (Mattison & McWhirter, 1994; Paul, Hays, & Coates, 1995). In particular, the severe illnesses and premature deaths suffered by many gay men have precipitated relationship losses for surviving partners, who may themselves be relatively young. A study of a large group of gay men in New York City found that nearly one-third had suffered the loss of a close friend or lover to AIDS (Martin & Dean, 1993). Such men may not only have cared for a dying partner through a long illness, but also suffer bereavement without the institutionalized supports that would be available to surviving members of heterosexual couples. Despite such stresses, gay communities have responded with tremendous energy by creating volunteer service agencies and memorials such as the AIDS Quilt, which have supported those suffering with illness and helped survivors to keep memories of loved ones alive (D'Augelli & Garnets, 1995; Paul et al., 1995).

In summary, research has begun to describe various characteristics of lesbian and gay relationships. Most lesbians and gay men want to have intimate romantic relationships, and most do in fact create such relationships. The majority of couples surveyed have reported a desire for egalitarian relationships, although many acknowledge that their current relationship may not attain that ideal. What inequalities in power do exist in lesbian and gay couple relationships can often be traced to differences in the personal resources and/or commitment of each partner to maintaining the relationship. Rarely do lesbian or gay couples rely on gendered divisions of labor in their relationships; they are more likely to divide family and household labor in more flexible and egalitarian ways. Although sexual practices diverge to some degree as a function of gender, both lesbian and gay couples are likely to express satisfaction with their sexual relationships. Many lesbian and gay relationships endure for a period of years, and there are no significant differences in break-up rates between lesbian, gay, and unmarried (but cohabiting) heterosexual couples. When problems arise, these are likely to be in areas common to heterosexual couples as well, although a few areas of special concern to lesbian and gay couples also exist due to the effects of societal prejudice and discrimination. Many gay relationships have been terminated prematurely because of the illness and death of a partner from HIV/AIDS, but other gay and lesbian relationships endure. Despite notable advances, research on lesbian and gay couples is still quite new, and much remains to be learned.

9.16.4 RESEARCH ON LESBIAN MOTHERS AND GAY FATHERS

Systematic research on lesbian mothers, gay fathers, and their children is a phenomenon of the last 25 years. Despite the diversity of lesbian and gay parenting communities, however, research has with few exceptions been conducted with relatively homogeneous groups of participants. Like research on couples, research on parents has generally employed samples composed of white, middle or upper middle class, well-educated individuals, generally living in the USA. In this chapter, any studies that provide exceptions to this rule are specifically noted. In this section, research on those who became parents in the context of heterosexual relationships, before coming out as lesbian or gay, is presented first, followed by a description of studies of lesbians who became parents after coming out.

9.16.4.1 Lesbians and Gay Men Who Became Parents in the Context of Heterosexual Relationships

One important impetus for research in this area has come from extrinsic sources, such as judicial concerns about the psychological health and well-being of lesbians compared with heterosexual mothers. Other work has arisen from concerns that are more intrinsic to the families themselves, such as what and when children should be told about their parents" sexual orientation.

9.16.4.1.1 Divorced lesbian mothers

Because it has often been raised as an issue by judges presiding over custody disputes (Falk, 1989), a number of studies have assessed the overall mental health of lesbians compared with heterosexual mothers. Consistent with data on the mental health of lesbians in general (Gonsiorek, 1991), research has revealed that divorced lesbian mothers score at least as high as divorced heterosexual mothers on assessments of psychological health. For instance, studies have found no differences between lesbian and heterosexual mothers on self-concept, happiness, overall adjustment, or psychiatric status (Falk, 1989).

Another area of judicial concern has focused on maternal sex-role behavior and its potential impact on children (Patterson, 1995a, 1995b). Stereotypes cited by the courts suggest that lesbians might be overly masculine and/or that they might interact inappropriately with their children. However, neither lesbian mothers" reports about their sex-role behavior nor their self-described interest in child-rearing have been found to differ from those of heterosexual mothers. Reports about responses to child behavior and ratings of warmth toward children have been found not to differ significantly between lesbian and heterosexual mothers.

Differences between lesbian and heterosexual mothers have also been reported. Among the most straightforward of these are the reports by Lyons (1983) and Pagelow (1980) that divorced lesbian mothers in their samples had more fears about loss of child custody than did divorced heterosexual mothers. Similarly, Green, Mandel, Hotvedt, Gray, and Smith (1986) reported that lesbian mothers were more likely than heterosexual mothers to be active in feminist organizations. Given the environments in which these lesbian mothers were living, findings like these are not surprising.

A few other scattered differences seem more difficult to interpret. For example, Miller, Jacobsen, and Bigner (1981) found lesbian mothers to be more child-centered than heterosexual mothers in their discipline techniques. In a sample of African-American lesbian and heterosexual mothers, Hill (1987) found that lesbian mothers reported being more flexible about rules, more relaxed about sex play and modesty, and more likely to have nontraditional expectations for their daughters. Pending confirmation and replication with different samples, these findings are best viewed as suggestive.

Several studies have also examined the social circumstances and relationships of lesbian mothers. Divorced lesbian mothers have consistently been reported to be more likely than divorced heterosexual mothers to be living with a romantic partner (Harris & Turner, 1985/86; Kirkpatrick, Smith, & Roy, 1981; Pagelow, 1980). Whether this represents a difference between lesbian and heterosexual mother-headed families, or reflects nothing more than sampling biases of the research, cannot be determined from current research studies. Information is sparse about the impact of such relationships in lesbian mother families, but that which has been published suggests that, like heterosexual step-parents, coresident lesbian partners of divorced lesbian mothers can be important sources of conflict as well as support in the family.

Relationships with the fathers of children in lesbian mother homes have also been a topic of study. Few differences in the likelihood of paternal financial support have been reported for lesbian and heterosexual families with children. Kirkpatrick et al. (1981) reported, for example, that only about one-half of heterosexual and about one-half of lesbian mothers in their sample received financial support from the fathers of their children. Findings about frequency of contact with the fathers are mixed, with some (e.g., Kirkpatrick et al., 1981) reporting no differences as a function of maternal sexual orientation and others (e.g., Golombok, Spencer, & Rutter, 1983) reporting more contact among lesbian than among heterosexual mothers.

Although most research has involved assessment of possible differences between lesbian and heterosexual mothers, a few studies have reported other types of comparisons. For instance, in a study of divorced lesbian mothers and divorced gay fathers, Harris and Turner (1985/86) found that gay fathers were likely to report higher incomes and that they encouraged more sex-typed toy play among their children, whereas lesbian mothers were more likely to see benefits for their children (e.g., increased empathy and tolerance for differences) as a result of having lesbian or gay parents. In comparisons of relationship satisfaction among lesbian couples who did or did not have children, Koepke, Hare, and Moran (1992) reported that couples with children scored higher on overall measures of relationship satisfaction and of the quality of their sexual relationship. These findings are intriguing, but much more research will be needed before their interpretation will be clear.

Another important set of questions, as yet little studied, concerns the conditions under which lesbian mothers experience enhanced feelings of well-being, support, and ability to care for their children. Rand, Graham, and Rawlings (1982) reported that psychological health of lesbian mothers was associated with the mothers" openness about her sexual orientation with her employer, ex-husband, children, and friends, and with her degree of feminist activism. Kirkpatrick (1987) found that lesbian mothers living with partners and children had greater economic and emotional resources than those living alone with their children. Much remains to be learned about determinants of individual differences in psychological well-being among lesbian mothers.

Many other issues that have arisen in the context of divorced lesbian mother families are also in need of study. For instance, when a mother is in the process of coming out as a lesbian to herself and to others, at what point in that process should she address the topic with her child, and in what ways should she do so—if at all? And what influence ought the child's age and circumstances to have in such a decision? Reports from research and clinical practice suggest that early adolescence may be a particularly difficult time for parents to initiate such conversations, and that disclosure may be less stressful at earlier or later points in a child's development (Patterson, 1992, 1995a), but systematic research on these issues is just beginning. Similarly, many issues remain to be addressed regarding step-family and blended family relationships that may emerge as a lesbian mother's household seeks new equilibrium following her separation or divorce from the child's father.

9.16.4.1.2 Divorced gay fathers

Although considerable research has focused on the overall psychological adjustment of lesbian mothers as compared with that of heterosexual mothers, no published studies of gay fathers make such comparisons with heterosexual fathers. This may be attributable to the greater role of judicial decision-making as an impetus for research on lesbian mothers. In jurisdictions where the law provides for biases in custody proceedings, these are likely to favor female and heterosexual parents. Perhaps because, other things being equal, gay fathers are unlikely to win custody battles over their children after divorce, fewer such cases have reached the courts. Consistent with this view, only a minority of gay fathers have been reported to live in the same households with their children (Bigner & Bozett, 1990; Bozett, 1980, 1989).

Research on the parenting attitudes of gay versus heterosexual divorced fathers has, however, been reported. Bigner and Jacobsen (1989a, 1989b) compared gay and heterosexual fathers, each of whom had at least two children. Their results revealed that, with one exception, there were no significant differences between gay and heterosexual fathers in their motives for parenthood. The single exception concerned the greater likelihood of gay than heterosexual fathers to cite the higher status accorded to parents as compared with nonparents in the dominant culture as a motivation for parenthood (Bigner & Jacobsen, 1989a).

Bigner and Jacobsen (1989b) also asked gay and heterosexual fathers in their sample to report on their behavior when interacting with their children. Although no differences emerged in the fathers" reports of involvement or intimacy, gay fathers reported that their behavior was characterized by greater responsiveness, more reasoning, and more limit-setting than heterosexual fathers. These reports by gay fathers of greater warmth and responsiveness, as well as greater control and limit-setting, are strongly reminiscent of findings from research with heterosexual families, and would seem to raise the possibility that gay fathers are more likely than their heterosexual counterparts to exhibit authoritative patterns of parenting behavior (Baumrind, 1967; Baumrind & Black, 1967). Caution must be exercised, however, in the interpretation of results which stem entirely from paternal reports about their own behavior.

In addition to research comparing gay and heterosexual fathers, a handful of studies have made other comparisons. For instance, Robinson and Skeen (1982) compared sex-role orientations of gay fathers with those of gay men who were not fathers, and found no differences. Similarly, Skeen and Robinson (1985) found no evidence to suggest that gay men's retrospective reports about relationships with their own parents varied as a function of whether or not they were parents themselves. As noted above, Harris and Turner (1985/86) compared gay fathers and lesbian mothers, reporting that while gay fathers had higher incomes and were more likely to report encouraging their children to play with sex-typed toys, lesbian mothers were more likely to believe that their children received positive benefits such as increased tolerance for diversity from having lesbian or gay parents.

Much research in this area has also arisen from concerns about the gay father identity and its transformations over time. For example, Miller (1978, 1979) and Bozett (1980, 1981a, 1981b, 1987) sought to provide a conceptualization of the processes through which a man who considers himself to be a heterosexual father may come to identify himself, both in public and in private, as a gay father. Based on extensive interviews with gay fathers in the USA and Canada, these authors emphasized the pivotal nature of identity disclosure itself and of the reactions to disclosure by significant people in a man's life. Miller (1978) suggested that while a number of factors, such as the extent of occupational autonomy and amount of access to gay communities, may affect how rapidly a gay man discloses his identity to others, the most important of these is likely to be the experience of falling in love with another man. It is this experience, more than any other, Miller argued, that leads a man to integrate the otherwise compartmentalized parts of his identity as a gay father.

9.16.4.2 Lesbians and Gay Men Choosing to Become Parents

Although for many years lesbian mothers and gay fathers were generally assumed to have become parents in the context of previous heterosexual relationships, both men and women are believed increasingly to be undertaking parenthood in the context of pre-existing lesbian and gay identities (Crawford, 1987; Patterson, 1994a, 1994b). Although a substantial body of research addresses the transition to parenthood among heterosexuals, little research has explored the transition to parenthood for gay men or lesbians. While many issues that arise for heterosexuals also face lesbians and gay men (e.g., concerns about how children will affect couple relationships, economic concerns about supporting children), lesbians and gay men must also cope with many additional issues because of their situation as members of stigmatized minorities. These issues are best understood by viewing them against the backdrop of pervasive heterosexism and antigay prejudice.

Antigay prejudice is evident in institutions involved with health care, education, and employment that often fail to support and, in many cases, are openly hostile to lesbian and gay families (Casper, Schultz, & Wickens, 1992; Pollack & Vaughn, 1987). Lesbian and gay parents may encounter antigay prejudice and bigotry even from their families of origin. Many, if not most of, the special concerns of lesbian and gay parents and prospective parents stem from problems created by such hostility.

A number of inter-related issues are often faced in particular by lesbians and gay men who wish to become parents (Crawford, 1987; Patterson, 1994b). One of the first needs among this group is for accurate, up-to-date information on how lesbians and gay men can become parents, how their children are likely to develop, and what supports are available to assist them. In addition to these educational needs, lesbians and gay men who are seeking biological parenthood are also likely to encounter various health concerns, ranging from medical screening of prospective birthparents to assistance with DI techniques, prenatal care, and preparation for birth. As matters progress, a number of legal concerns about the rights and responsibilities of all parties are likely to emerge. Associated with all of these will generally be financial issues; in addition to the support of a child, auxiliary costs of medical and legal assistance may be considerable. Finally, social and emotional concerns of many different kinds are also likely to emerge (Patterson, 1994b; Pies, 1985, 1990; Pollack & Vaughn, 1987; Rohrbaugh, 1988).

The earliest studies of childbearing among lesbian couples were reported by McCandlish (1987) and Steckel (1987). Both investigators reported research based on small samples of lesbian couples who had given birth to children by means of DI. Their focus was primarily on the children in such families, and neither investigator provided much in the way of systematic assessment of mothers. McCandlish did, however, highlight events and issues that were significant among families in her sample. For instance, she noted that, regardless of their interest in parenting prior to birth of the first child, nonbiological mothers in each couple unanimously reported an "unexpected and immediate attachment" to the child (McCandlish, 1987, p. 28). Although both mothers took part in parenting, they reported shifting patterns of caretaking responsibilities over time, with the biological mother taking primary responsibility during the earliest months, and the nonbiological mother's role increasing in importance after the first year. Couples also reported changes in their relationships following the birth of the child, notably a reduction or cessation in sexual intimacy. Though the best interpretation of results from these pioneering studies is by no means clear, the work raises important issues and questions.

A study by Hand (1991) examined the ways in which 17 lesbian and 17 heterosexual couples with children under two years of age shared childcare, household duties, and occupational roles. Her principal finding was that lesbian couples shared parenting more equally than heterosexual couples. Lesbian nonbiological mothers were significantly more involved in childcare and regarded their parental role as significantly more salient than heterosexual fathers. Lesbian biological mothers viewed their maternal role as more salient than did any of the other mothers, whether lesbian or heterosexual. Fathers viewed their occupational roles as more salient than any of the mothers, whether lesbian or heterosexual. Hand's major result was, however, that lesbian couples were more likely than heterosexual couples to share childcare relatively evenly.

Another study (Osterweil, 1991) involved 30 lesbian couples with at least one child between 18 and 36 months of age. Consistent with Hand's (1991) results for parents of younger children, Osterweil reported that biological mothers viewed their maternal role as more salient than nonbiological mothers. In addition, although household maintenance activities were shared about equally, biological mothers reported somewhat more influence in family decisions and somewhat more involvement in childcare. Osterweil also reported that the couples in her study scored at about the mean for normative samples of heterosexual couples in overall relationship satisfaction. These studies suggest that lesbian couples who have chosen to bear children are likely to share household and childcare duties to a somewhat greater degree than heterosexual couples, and that lesbians are relatively satisfied with their couple relationships. Similar findings have also been reported by Patterson (1995d).

Two studies of men who have become fathers after identifying themselves as gay have been reported. Sbordone (1993) studied 78 gay men who had become parents through adoption or through surrogacy, and compared them with 83 gay men who were not fathers. Consistent with Skeen and Robinson's (1985) findings for divorced gay fathers, Sbordone found no differences between fathers and nonfathers on reports about relationships with the men's own parents. Gay fathers did, however, report higher self-esteem and fewer negative attitudes about homosexuality than gay men who were not fathers.

An interesting observation of Sbordone's (1993) study was that most gay men who were not fathers indicated that they would like to rear a child. The men who said that they wanted children were younger than those who said they did not, but the two groups did not differ on income, education, race, self-esteem, or attitudes about homosexuality. Given that fathers had higher self-esteem than either group of nonfathers, Sbordone suggested the possibility that the higher self-esteem of gay fathers might have been a result, not a cause or simple correlate, of parenthood.

McPherson (1993) reported a study of gay couples who had chosen to become parents. He studied division of labor, satisfaction with division of labor, and satisfaction with couple relationships among 28 gay and 27 heterosexual parenting couples. Consistent with evidence from lesbian couples with children (Hand 1991; Osterweil, 1991; Patterson, 1995d), McPherson found that gay couples reported a more even division of responsibilities for household maintenance and childcare than did heterosexual couples. Gay couples also reported greater satisfaction with their division of childcare tasks than did heterosexual couples. Finally, gay couples also reported greater satisfaction with their couple relationships, especially in the areas of cohesion and expression of affection.

In summary, studies of lesbians and gay men who have chosen to become parents are still sparse. Most research has been conducted on a relatively small scale and many important issues have yet to be addressed. Much remains to be learned about the determinants of lesbian and gay parenting and about its impact on lesbian and gay parents themselves.

9.16.5 RESEARCH ON CHILDREN OF LESBIAN AND GAY PARENTS

Research on children in lesbian and gay families has, with few exceptions, been conducted with relatively homogeneous groups of white, well-educated, middle class, largely professional families living in or around urban centers in the USA or in other Western countries. Unless otherwise specifically noted, these characteristics apply to the research described in this section. Research on children born in the context of heterosexual relationships is presented first, followed by a description of work with children born to or adopted by lesbian and gay parents.

9.16.5.1 Research on Children Born in the Context of Heterosexual Relationships

As with research on lesbian mothers, much of the impetus for research in this area has come from judicial concerns about the welfare of children residing with gay or lesbian parents. Research in each of three main areas of judicial concern—namely, children's sexual identity, other aspects of children's personal development, and children's social relationships—is summarized here.

9.16.5.1.1 Sexual identity

Following Money and Ehrhardt (1972), research on three aspects of sexual identity is considered. Gender identity concerns a person's self-identification as male or female. Gender-role behavior concerns the extent to which a person's activities and occupations are regarded by the culture as masculine, feminine, or both. Sexual orientation refers to a person's choice of sexual partners (e.g., heterosexual, homosexual, or bisexual).

Research on gender identity has failed to reveal any differences in the development of children as a function of their parents" sexual orientation. For example, Kirkpatrick et al. (1981) compared development among 20 children of lesbian mothers with that among 20 same-aged children of heterosexual mothers. In projective testing, most children in both groups drew a same-sex figure first, a finding that fell within expected norms. Of those who drew an opposite-sex figure first, only three (one with a lesbian mother, and two with heterosexual mothers) showed concern about gender issues in clinical interviews. Similar findings have been reported in projective testing by other investigators (e.g., Green et al., 1986) and studies using more direct methods of assessment have yielded similar results (e.g., Golombok et al., 1983).

Research on gender-role behavior has also failed to reveal difficulties in the development of children with lesbian mothers. For instance, Green (1978) reported that 20 of 21 children of lesbian mothers in his sample named a favorite toy consistent with conventional sex-typed toy preferences, and that all reported vocational choices that fell within typical limits for conventional sex roles. In interviews with 56 children of lesbians and 48 children of heterosexual mothers, Green et al. (1986) found no differences with respect to favorite television programs, television characters, games, or toys. They reported that daughters of lesbian mothers were more likely to be described as taking part in rough and tumble play or as playing with masculine toys such as trucks or guns, but found no comparable differences for sons. In all studies, the behavior of lesbian mothers" children was seen as falling within normal limits.

Rees (1979) administered the Bem Sex Role Inventory to 12 young adolescent offspring of lesbian mothers and 12 same-aged youngsters of heterosexual mothers. Although children of lesbian and heterosexual mothers did not differ on masculinity or androgyny, adolescent offspring of lesbian mothers reported greater psychological femininity than their same-aged peers with heterosexual mothers. This result runs counter to expectations based on stereotypes of lesbians as lacking in femininity and, although provocative, it should be interpreted cautiously pending replication. Overall, research has failed to reveal any notable difficulties in the development of sex-role behavior among children of lesbian mothers.

A number of investigators have also studied sexual orientation, the third component of sexual identity. For example, Huggins (1989) interviewed 36 teenagers, half of whom were the offspring of lesbian mothers and half of heterosexual mothers. No child of a lesbian mother identified as lesbian or gay, but one child of a heterosexual mother did. Generally similar results have been reported by other investigators (e.g., Golombok & Tasker, 1996; Gottman, 1990). Studies of the offspring of gay fathers have yielded similar results (Bozett, 1987). Based on the results of these studies, there is no reason to believe that the offspring of lesbian or gay parents are any more likely than those of heterosexual parents to become lesbian or gay themselves.

As clear as these results are, it should be recognized that research on the development of sexual identity among the offspring of lesbian and gay parents has been criticized from a number of perspectives. For instance, many lesbian women do not self-identify as lesbians until adulthood (see Brown, 1995); for this reason, studies of sexual orientation among adolescents may count as heterosexual some individuals who will identify themselves as lesbian later in life. Concern has also been voiced that in many studies that compare children of divorced heterosexual mothers with children of divorced lesbian mothers, the lesbian mothers were more likely to be living with a romantic partner; in these cases, maternal sexual orientation and relationship status have been confounded. While these and other methodological issues await resolution, it remains true that no significant problems in the development of sexual identity among children of lesbian mothers have yet been identified.

9.16.5.1.2 Other aspects of personal development

Studies of other aspects of personal development among children of gay and lesbian parents have assessed a broad array of characteristics (Patterson, 1995a; Tasker & Golombok, 1995). Among these have been psychiatric evaluations and assessments of behavior problems, personality, self-concept, locus of control, moral judgment, and intelligence. Concerns about possible difficulties in personal development among children of lesbian and gay parents have not been sustained by the results of research (Patterson, 1992, 1995a, 1997). As was true for sexual identity, studies of other aspects of personal development have revealed no significant differences between children of lesbian or gay parents and children of heterosexual parents. On the basis of existing evidence, fears that children of gay and lesbian parents suffer deficits in personal development appear to be without empirical foundation.

9.16.5.1.3 Social relationships

Studies assessing potential differences between children of gay and lesbian versus heterosexual parents have sometimes included assessments of children's social relationships. Because of concerns voiced by the courts that children of lesbian and gay parents might encounter difficulties among their peers, the most common focus of attention has been on peer relations. Studies have consistently found that children of lesbian mothers report normal peer relations, and that adult observers agree with this judgment (Patterson, 1992). Anecdotal and first-person accounts describe children's worries about being stigmatized as a result of their parents" sexual orientation (e.g., Pollack & Vaughn, 1987), but research findings provide no evidence for the proposition that children of lesbian mothers have difficulties in peer relations (Tasker & Golombok, 1995).

Research has also described children's relationships with adults, especially fathers. For instance, Golombok et al. (1983) found that children of lesbian mothers were more likely than children of heterosexual mothers to have contact with their fathers. Most children of lesbian mothers had some contact with their fathers during the year preceding the study, but most children of heterosexual mothers had not; indeed, almost a third of the children of lesbian mothers reported at least weekly contact with their fathers, whereas only 1 in 20 of the children of heterosexual mothers reported this. Kirkpatrick et al. (1981) also reported that lesbian mothers were more concerned than heterosexual mothers that their children have opportunities for good relationships with adult men, including fathers. Lesbian mothers" social networks have been found to include both men and women and, as a result, their children have contact with adults of both genders. Overall, research suggests that children of lesbian parents have satisfactory relationships with adults of both sexes.

Concerns that children of lesbian or gay parents are more likely than children of heterosexual parents to be sexually abused have also been voiced by judges in the context of child custody disputes (Patterson, 1992). Research in this area shows that the great majority of adults who perpetrate sexual abuse are males and sexual abuse of children by adult women is extremely rare. Lesbian mothers are thus extremely unlikely to abuse their children. Existing research findings suggest that gay men are no more likely than heterosexual men to perpetrate child sexual abuse. Fears that children in custody of gay or lesbian parents might be at heightened risk for sexual abuse are thus without empirical foundation (Patterson, 1992, 1995a, 1995b).

9.16.5.2 Diversity Among Children With Divorced Lesbian or Gay Parents

Despite the great diversity evident with gay and lesbian communities (Blumenfeld & Raymond, 1988), research on differences among children of lesbian and gay parents is, as yet, very limited. One important dimension of difference among gay and lesbian families concerns whether the custodial parent is involved in a romantic relationship, and if so what implications this may have for children. Pagelow (1980), Kirkpatrick et al. (1981), and Golombok et al. (1983) reported that divorced lesbian mothers were more likely than divorced heterosexual mothers to be living with a romantic partner. Huggins (1989) reported that self-esteem among daughters of lesbian mothers whose lesbian partners lived with them was higher than that among daughters of lesbian mothers who did not live with a partner. This finding might be interpreted to mean that mothers who are high in self-esteem are more likely to be involved in romantic relationships and to have daughters who are also high in self-esteem, but many other interpretations are also possible. In view of the small sample size and absence of conventional statistical tests, Huggins" finding should be interpreted with great caution.

Rand et al. (1982) found that lesbian mothers" sense of psychological well-being was related to the extent to which they were open about their lesbian identity with employers, ex-husbands, and children. In their sample, a mother who felt more able to disclose her lesbian identity was also more likely to express a positive sense of well-being. In light of the consistent finding that, in heterosexual families, children's adjustment is often related to indexes of maternal mental health (Sameroff & Chandler, 1975), one might expect factors which enhance mental health among lesbian mothers also to benefit the children of these women, but this possibility has not yet been studied.

Another area of great diversity among families with a gay or lesbian parent concerns the degree to which a parent's sexual identity is accepted by other significant people in children's lives. Huggins (1989) found a tendency for children whose fathers were rejecting of maternal lesbianism to report lower self-esteem than those whose fathers were neutral or positive. Due to small sample size and absence of conventional statistical tests, this finding should be seen as suggestive rather than definitive. Huggins" results raise questions, however, about the extent to which reactions of important adults in a child's environment influence responses to discovery of a parent's gay or lesbian identity.

Effects of the age at which children learn of parents" gay or lesbian identities have also been investigated. Paul (1986) reported that those who were told either in childhood or in late adolescence found the news easier to cope with than did those who first learned of it during adolescence. Huggins (1989) reported that those who learned of maternal lesbianism in childhood had higher self-esteem than those who were not informed until they were adolescents. From a clinical standpoint, early adolescence is a particularly difficult time for children to learn of their parents" lesbian or gay identities (Patterson, 1992).

In summary, existing data favor early disclosure of identity to children, good maternal mental health, and a supportive milieu, but the available data are very limited. No information is yet available on differences attributable to race or ethnicity, family economic circumstances, cultural environments, or related variables. Because none of the published work has employed observational measures or longitudinal designs, little is known about the details of actual behavior in these families or about any changes over time. It is clear that much remains to be learned about differences among gay and lesbian families and about the impact of such differences on children growing up in these homes.

9.16.5.3 Research on Children Born to or Adopted by Lesbian Mothers

In one of the first systematic studies of children born to lesbians, Steckel (1987) compared the progress of separation-individuation among preschool children born via DI to lesbian couples with that among same-aged children of heterosexual couples. Using parent interviews, parent and teacher Q sorts, and structured doll play techniques, Steckel compared independence, ego functions, and object relations among children in the two types of families. Her main results documented impressive similarity in development among children in the two groups. Similar findings, based on extensive interviews with five lesbian mother families were also reported by McCandlish (1987).

Steckel (1987) did, however, report suggestive differences between the two groups. Children of heterosexual parents saw themselves as somewhat more aggressive than did children of lesbians, and they were seen by both parents and teachers as more bossy, domineering, and negativistic. Children of lesbian parents saw themselves as more lovable and were seen by parents and teachers as more affectionate, responsive, and protective toward younger children. In view of the small sample size, and the large number of statistical tests performed, these results must be interpreted with caution. Steckel's work was, however, the first to make systematic comparisons of development among children born to lesbian and to heterosexual couples.

The first study to examine psychosocial development among preschool and school-aged children born to or adopted by lesbian mothers was conducted by Patterson (1994a). Thirty-seven four- to nine-year-old children were studied, using a variety of standardized measures. The Child Behavior Checklist was used to assess children's social competence and behavior problems; five scales from the Eder Children's Self-View Questionnaire were used to assess children's self-concepts, and open-ended interview techniques were used to assess preferences associated with sex-role behavior.

Results showed that children scored in the normal range for all measures. On the Child Behavior Checklist (Achenbach and Edelbrock, 1983), for example, children of lesbian mothers" scores for social competence, internalizing behavior problems, and externalizing behavior problems differed significantly from the scores for a clinical sample, but did not differ from the scores for a large normative sample of American children. Likewise, children of lesbian mothers reported sex-role preferences within the expected normal range for children of this age. On most subscales of the self-concept measure, answers given by children of lesbian mothers did not differ from those given by same-aged children of heterosexual mothers studied in a standardization sample.

On two subscales of the self-concept measure, however, Patterson (1994a) found that children of lesbian mothers reported feeling more reactions to stress (e.g., feeling angry, scared, or upset), but a greater sense of well-being (e.g., feeling joyful, content, and comfortable with themselves) than did the same-aged children of heterosexual mothers in the standardization sample. One possible interpretation of this result is that children of lesbian mothers report greater reactivity to stress because, in fact, they experienced greater stress in their daily lives than did other children. Another possibility is that, regardless of actual stress levels, children of lesbian mothers were better able to acknowledge both positive and negative aspects of their emotional experience.

Some additional findings from this study can be mentioned briefly. Contrary to stereotypes of these families as isolated from families of origin, most reported that children had regular contact with one or more grandparents, as well as with other adult friends and relatives of both sexes (Patterson, Hurt, & Mason, in press). In families headed by lesbian couples, the parents were likely to maintain egalitarian divisions of labor, but when differences occurred, biological lesbian mothers were likely to do somewhat more childcare and nonbiological lesbian mothers were likely to spend somewhat more time engaged in paid employment (Patterson, 1995d). Even within the relatively small range represented in this sample, families in which childcare was divided more evenly were also those in which children exhibited the most favorable adjustment (Patterson, 1995d).

9.16.6 RESEARCH ON OTHER FAMILY RELATIONSHIPS

In addition to parent–child and couple relationships, lesbians and gay men are likely to maintain contacts with parents, siblings, and other members of their families of origin (Cohen & Savin-Williams, 1996; Laird, 1996; Patterson et al., in press). Although, as Laird (1996) has emphasized, many other issues are undoubtedly significant, the largest amount of research has focused on the concerns of young lesbians and gay men about disclosing their sexual identities to members of their families of origin, especially to parents.

After acknowledging lesbian or gay identities, many people begin to wonder whether and how to share such identities with (i.e., "come out to") those who are important to them. Most lesbians and gay men appear to come out first to close friends, and only later—if at all—to family members (Savin-Williams, 1990). Young people are more likely to come out first to mothers rather than to fathers, perhaps because they expect more positive responses from mothers (Bryant & Demian, 1994; Cohen & Savin-Williams, 1996). For example, Remafedi (1987) studied a sample of gay and bisexual teenagers, and found that most were out to their mothers but not their fathers, and almost all were out to at least one friend.

Although it is difficult to predict parental reactions to disclosure of a nonheterosexual orientation by their offspring, the most common initial reactions are usually negative (Cohen & Savin-Williams, 1996). Negative reactions are likely to be more pronounced among older parents, those with less education, and those whose parent–child relationships were troubled before the disclosure. Although interactions between lesbian and gay young people and their parents often suffer difficulties immediately after disclosure, they most often improve again over time, as families assimilate this new information into existing images of the lesbian or gay child. The best predictor of postdisclosure relationships between lesbian and gay young adults and their parents is the quality of their relationships before the disclosure (Cohen & Savin-Williams, 1996; Savin-Williams, 1990).

What are the associations between disclosure of lesbian or gay identity to parents and young adults" self-esteem? Because of the significance of parent–adolescent relationships, one might expect parental acceptance to be associated with favorable self-images among lesbian and gay youth. Consistent with this view, Savin-Williams (1990) found that teenaged and young adult lesbians who said that their parents were accepting of their sexual identities (or would be accepting if they knew) also reported feeling more comfortable with their sexual orientation. This was true for young men, however, only if they also described their parents as important to their self-image (Savin-Williams, 1990). Because the research has been correlational in nature, it cannot be determined whether parental acceptance makes lesbian and gay children feel better about themselves, whether youth who already have high self-esteem are more likely to disclose to parents, or whether a cyclical process may be involved. Identification of causal pathways in this area thus represents an important challenge for future research.

While some research has focused on young adults" disclosure of lesbian and gay identities to parents, other studies have shown that, among samples of older lesbian and gay adults, sizeable proportions have not come out to parents or other family members. When a lesbian or gay identity has not been disclosed, any one of several coping strategies may be employed by the individual and the family (Brown, 1989). A common one is distancing, whether emotionally or geographically (or both) from the family of origin. Another is the unspoken agreement that nobody in the family will discuss the lesbian or gay individual's personal life; this has been termed the "I know you know" strategy. A third approach is to disclose to one family member, who is thought to be supportive, on the condition that no others be told. This approach appears to rely on and perhaps strengthen coalitions among subgroups within a family. Although these strategies may or may not be viewed as problematic by those who employ them, they all block the achievement of true intimacy and add to the stress experienced by lesbian and gay adults (Brown, 1989).

When a family member's nonheterosexual orientation becomes known, Strommen (1989a, 1989b) has described the family's reaction as involving a two-stage process. First, the family members struggle to understand and assimilate this new information about one of its members. The family may then simply reject the lesbian or gay person, or it may reorganize itself over time to accommodate this shift in identity while still including the lesbian or gay person in family activities. Parents in particular often find that the process of reorganization can be difficult, often extending over substantial periods of time. In the end, many discover that the process has brought them unexpected gifts (Bernstein, 1995).

9.16.7 IMPLICATIONS OF RESEARCH FINDINGS

Research on lesbian and gay couples and families with children, though relatively new, has nevertheless yielded results that are worthy of attention. Without denying the consistency of major research findings, it is important also to acknowledge that the research is subject to various criticisms. For instance, much of the research has involved small samples that are predominantly white, well educated, relatively affluent, and American; the degree to which results would hold with other populations is thus difficult or impossible to evaluate at this time. It would also be desirable to have data based on observational methods, collected within longitudinal designs, but studies of this kind have not yet been reported.

Despite shortcomings, however, central results of existing research on lesbian and gay couples and families with children are exceptionally clear. First, research has succeeded in bringing to light the fact that lesbian and gay parents as well as couples exist in large numbers. Given the relative invisibility of lesbian and gay couples, as well as that of parents and their children, this achievement is significant in itself, and should not be overlooked.

Beyond their witness to the sheer existence of lesbian and gay family lives, the results of existing studies, taken together, also yield a picture of families thriving, even in the midst of discrimination and oppression. Certainly, they provide no evidence that psychological adjustment among lesbians, gay men, their children or other family members is impaired in any significant respect. Indeed, the available evidence suggests that relationships of lesbian and gay couples are just as supportive, and that home environments provided by lesbian and gay parents are just as likely as those provided by heterosexual parents to support and enable psychosocial growth among family members.

Much research on lesbian and gay parenting has focused primarily on comparisons between lesbian and gay families, on the one hand, and heterosexual families, on the other. This approach presumably reflects the concern of researchers to address prejudices and negative stereotypes that have been influential in judicial decision-making and in public policies relevant to lesbian and gay couples, parents and their children in the USA. Now that results of research have begun to converge so clearly on answers to questions posed in this way, it would appear that the time has come also to address a broader range of issues in this area.

Many important research questions arise from a focus on the interests of lesbian and gay families themselves. For example, many lesbian and gay couples with children are interested in distinctions between the experiences of biological and nonbiological parents. How important, they ask, is the biological link in influencing experiences of parenthood? Similarly, both lesbian and gay families can benefit from more information about diversity among lesbian and gay couples and families with children. It would seem likely that future scholarship will increasingly concern itself with the study of sources of strength and resilience in lesbian and gay couples as well as among parents and their children.

In the meantime, however, the central results of research have important implications. If psychosocial development among children born to lesbian mothers and gay fathers is essentially normal, then traditional theoretical emphases on the importance of parental heterosexuality need to be reconsidered. Although many possible approaches to such a task are possible (Patterson, 1992), one promising approach is to focus on the significance of family process rather than structure. Thus, structural variables such as parental sexual orientation may ultimately be seen as less important in mediating children's developmental outcomes than the quality of family interactions, relationships, and processes.

Results of research with lesbian and gay parents and their children also have implications for the politics of family life. If, as would appear to be the case, neither parents nor children in lesbian and gay families run any special risk of maladjustment or other psychosocial problems, then a good rationale for prejudice and discrimination is more and more difficult to provide. Without such a rationale, many legal precedents and public policies relevant to lesbian and gay families would require reconsideration. Ultimately, lesbian and gay couples and parents might come to be viewed as normal, and policies might be designed to protect their legitimate interests, as well as those of their family members.

9.16.8 DIRECTIONS FOR RESEARCH

Because research on the family lives of lesbians and gay men is relatively new, there are many promising avenues for further research (Allen & Demo, 1995). From a substantive point of view, a number of issues have gone all but unexplored in the research literature on lesbian and gay family lives. For example, little attention has been devoted to assessment of sexual orientation over time, and the phenomena associated with bisexuality (Paul, 1996) have received little study. Ethnic, racial, and socioeconomic diversity of lesbian and gay family lives have yet to be systematically explored. Little research has been conducted outside of the United States. These gaps all provide important opportunities for future research.

From a methodological perspective, it would be valuable to have more studies that follow couples or parents and their children over time. Longitudinal studies of the relationships between lesbians, gay men, and members of their families of origin over relatively long periods of time could also be helpful in describing predictable sequences of reactions to significant life events (e.g., coming out, having a child) among family members. To avoid the pitfalls associated with retrospective reporting, these studies should involve prospective designs that follow participants over time.

Another methodological issue in the literature is the almost complete absence of observational data. Observational studies of couples, parents, and children, as well as of lesbian and gay adults with members of their families of origin, could provide valuable evidence about similarities and differences between family processes in the family lives of lesbian, gay, and heterosexual adults. Such observational data could be collected from dyads or triads or larger family groups, at home or in the laboratory, in a single visit or in repeated sessions over time; and it could add tremendously to knowledge in this area.

Overall, the study of lesbian and gay family lives provides a context in which to explore the limits of existing theoretical perspectives, and an opportunity to develop new ones. Future research that addresses these challenges has the potential to improve understanding of lesbian and gay family life, increase inclusiveness of theoretical notions about family structure and process, and inform public policies and judicial rulings relevant to lesbian and gay family life.

When lesbian and gay family lives are viewed in these terms, it serves to underline the tremendous signficance of historical factors in shaping both individual and familial experiences. Rapid change in attitudes, social climates, and even legal rulings relevant to lesbian and gay family lives in the USA has, in many ways, transformed the daily lives of lesbians and gay men, and those of their family members as well. Future events, such as decisions in cases like *Baehr v. Miike*, hold the potential to further transform the experiences associated with lesbian and gay family lives.

9.16.9 REFERENCES

Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Allen, K. R., & Demo, D. H. (1995). The families of lesbians and gay men: A new frontier in family research. *Journal of Marriage and the Family, 57*, 1–17.

Badgett, M. V. L. (in press). The economic well-being of lesbian, gay and bisexual adults" families. In C. J. Patterson & A. R. D'Augelli (Eds.), *Lesbian, gay and bisexual identities in families: Psychological perspectives*. New York: Oxford University Press.

Baumrind, D. (1967). Childcare practices anteceding three patterns of preschool behavior. *Genetic Psychology Monographs, 75*, 43–88.

Baumrind, D., & Black, A. E. (1967). Socialization practices associated with dimensions of competence in preschool boys and girls. *Child Development, 38*, 291–327.

Bell, A. P., & Weinberg, M. S. (1978). *Homosexualities: A study of diversity among men and women*. New York: Simon & Schuster.

Bernstein, R. A. (1995). *Straight parents, gay children: Keeping families together*. New York: Thunder's Mouth Press.

Bigner, J. J., & Bozett, F. W. (1990). Parenting by gay fathers. In F. W. Bozett & M. B. Sussman (Eds.), *Homosexuality and family relations* (pp. 155–176). New York: Harrington Park.

Bigner, J. J., & Jacobsen, R. B. (1989a). The value of children to gay and heterosexual fathers. In F. W. Bozett (Ed.), *Homosexuality and the family* (pp. 163–172). New York: Harrington Park.

Bigner, J. J., & Jacobsen, R. B. (1989b). Parenting behaviors of homosexual and heterosexual fathers. In F. W. Bozett (Ed.), *Homosexuality and the family* (pp. 173–186). New York: Harrington Park Press.

Blumenfeld, W. J., & Raymond, D. (1988). *Looking at gay and lesbian life*. Boston: Beacon.

Blumstein, P., & Schwartz, P. (1983). *American couples: Money, work, sex*. New York: Morrow.

Bozett, F. W. (1980). Gay fathers: How and why they disclose their homosexuality to their children. *Family Relations, 29*, 173–179.

Bozett, F. W. (1981a). Gay fathers: Evolution of the gay-father identity. *American Journal of Orthopsychiatry, 51*, 552–559.

Bozett, F. W. (1981b). Gay fathers: Identity conflict resolution through integrative sanctioning. *Alternative Lifestyles, 4*, 90–107.

Bozett, F. W. (1987). Children of gay fathers. In F. W. Bozett (Ed.), *Gay and lesbian parents* (pp. 39–57). New York: Praeger.

Bozett, F. W. (1989). Gay fathers: A review of the literature. In F. W. Bozett (Ed.), *Homosexuality and the family* (pp. 137–162). New York: Harrington Park.

Brown, L. (1989). Lesbians, gay men, and their families: Common clinical issues. *Journal of Gay and Lesbian Psychotherapy, 1*, 65–77.

Brown, L. (1995). Lesbian identities: Concepts and issues. In A. R. D'Augelli & C. J. Patterson (Eds.), *Lesbian, gay and bisexual identities over the lifespan: Psychological perspectives* (pp. 3–23). New York: Oxford University Press.

Bryant, A. S., & Demian (1994). Relationship characteristics of American gay and lesbian couples: Findings from a national survey. *Journal of Gay and Lesbian Social Services, 1*, 101–117.

Bull, C., & Gallagher, J. (1996). *Perfect enemies: The religious right, the gay movement, and the politics of the 1990s*. New York: Crown.

Caldwell, M. A., & Peplau, L. A. (1984). The balance of power in lesbian relationships. *Sex Roles, 10*, 587–599.

Cardell, M., Finn, S., & Maracek, J. (1981). Sex-role identity, sex-role behavior, and satisfaction in heterosexual, lesbian, and gay male couples. *Psychology of Women Quarterly, 5*, 488–494.

Casper, V., Schultz, S., & Wickens, E. (1992). Breaking the silences: Lesbian and gay parents and the schools. *Teachers College Record, 94*, 109–137.

Cohen, K. M., & Savin-Williams, R. C. (1996). Developmental perspectives on coming out to self and others. In R. C. Savin-Williams & K. M. Cohen (Eds.), *The lives of lesbians, gays, and bisexuals: Children to adults* (pp. 113–151). New York: Harcourt Brace.

Crawford, S. (1987). Lesbian families: Psychosocial stress and the family-building process. In Boston Lesbian Psychologies Collective (Eds.), *Lesbian psychologies: Explorations and challenges* (pp. 195–214). Urbana, IL: University of Illinois Press.

D'Augelli, A. R., & Garnets, L. (1995). Lesbian, gay and bisexual communities. In A. R. D'Augelli & C. J. Patterson (Eds.), *Lesbian, gay and bisexual identities over the lifespan: Psychological perspectives* (pp. 293–320). New York: Oxford University Press.

Dunlap, D. W. (1996). Reform rabbis vote to back gay marriage. *New York Times*, March 29, 1996, p. B-9.

Editors of the Harvard Law Review (1990). *Sexual orientation and the law*. Cambridge, MA: Harvard University Press.

Falk, P. J. (1989). Lesbian mothers: Psychosocial assumptions in family law. *American Psychologist, 44*, 941–947.

Gibbs, E. D. (1988). Psychosocial development of children raised by lesbian mothers: A review of research. *Women and Therapy, 8*, 55–75.

Golombok, S., Spencer, A., & Rutter, M. (1983). Children in lesbian and single-parent households: Psychosexual and psychiatric appraisal. *Journal of Child Psychology and Psychiatry, 24*, 551–572.

Golombok, S., & Tasker, F. (1996). Do parents influence the sexual orientation of their children? Findings from a longitudinal study of lesbian families. *Developmental Psychology, 32*, 3–11.

Gonsiorek, J. C. (1991). The empirical basis for the demise of the illness model of homosexuality. In J. C. Gonsiorek & J. D. Weinrich (Eds.), *Homosexuality: Research implications for public policy* (pp. 115–136). Beverly Hills, CA: Sage.

Gottman, J. S. (1990). Children of gay and lesbian parents. In F. W. Bozett & M. B. Sussman (Eds.), *Homosexuality and family relations* (pp. 177–196). New York: Harrington Park.

Green, R. (1978). Sexual identity of 37 children raised by homosexual or transsexual parents. *American Journal of Psychiatry, 135*, 692–697.

Green, R., Mandel, J. B., Hotvedt, M. E., Gray, J., & Smith, L. (1986). Lesbian mothers and their children: A comparison with solo parent heterosexual mothers and their children. *Archives of Sexual Behavior, 7*, 175–181.

Hand, S. I. (1991). *The lesbian parenting couple*. Unpublished doctoral dissertation, Professional School of Psychology, San Francisco.

Harris, M. B., & Turner, P. H. (1985/86). Gay and lesbian parents. *Journal of Homosexuality, 12*, 101–113.

Harry, J. (1984). *Gay couples*. New York: Praeger.

Harry, J., & DeVall, W. B. (1978). *The social organization of gay males*. New York: Praeger.

Henry, T. (1993). Making gays and lesbians part of the rainbow curriculum. *USA Today*, January 12, 1993, D, 4:1.

Hill, M. (1987). Child-rearing attitudes of black lesbian mothers. In Boston Lesbian Psychologies Collective (Ed.), *Lesbian psychologies: Explorations and challenges* (pp. 215–226). Urban, IL: University of Illinois Press.

Huggins, S. L. (1989). A comparative study of self-esteem of adolescent children of divorced lesbian mothers and divorced heterosexual mothers. In F. W. Bozett (Ed.), *Homosexuality and the family* (pp. 123–135). New York: Harrington Park.

James, S. E., & Murphy, B. C. (in press). Gay and lesbian relationships in a changing social context. In C. J. Patterson & A. R. D'Augelli (Eds.), *Lesbian, gay and bisexual identities in families: Psychological perspectives*. New York: Oxford University Press.

Kirkpatrick, M. (1987). Clinical implications of lesbian mother studies. *Journal of Homosexuality, 13*, 201–211.

Kirkpatrick, M., Smith, C., & Roy, R. (1981). Lesbian mothers and their children: A comparative survey. *American Journal of Orthopsychiatry, 51*, 545–551.

Koepke, L., Hare, J., & Moran, P. B. (1992). Relationship quality in a sample of lesbian couples with children and child-free lesbian couples. *Family Relations, 41*, 224–229.

Kurdek, L. A. (1992). *Conflict in gay and lesbian cohabiting couples*. Paper presented at the meeting of the American Psychological Association, Washington, DC.

Kurdek, L. A. (1993) The allocation of household labor in homosexual and heterosexual cohabiting couples. *Journal of Social Issues, 49*, 127–139.

Kurdek, L. A. (1994). The nature and correlates of relationship quality in gay, lesbian, and heterosexual cohabiting couples: A test of the contextual, investment, and discrepancy models. In B. Greene & G. M. Herek (Eds.), *Lesbian and gay psychology: Theory, research, and clinical applications* (pp. 133–155). Thousand Oaks, CA: Sage.

Kurdek, L. A. (1995). Lesbian and gay couples. In A. R. D'Augelli & C. J. Patterson (Eds.), *Lesbian gay and bisexual identities over the lifespan: Psychological perspectives* (pp. 243–261). New York: Oxford University Press.

Kurdek, L. A., & Schmidt, J. P. (1986a). Early development of relationship quality in heterosexual married, heterosexual cohabiting, gay and lesbian couples. *Developmental Psychology, 22*, 305–309.

Kurdek, L. A., & Schmidt, J. P. (1986b). Relationship quality of partners in heterosexual married, heterosexual cohabiting, and gay and lesbian relationships. *Journal of Personality and Social Psychology, 51*, 711–720.

Kurdek, L. A., & Schmidt, J. P. (1987). Perceived support from family and friends in members of homosexual, married, and heterosexual cohabiting couples. *Journal of Homosexuality, 14*, 57–68.

Laird, J. (1996). Invisible ties: Lesbians and their families of origin. In J. Laird & R. J. Green (Eds.), *Lesbians and gays in couples and families* (pp. 89–122). New York: Jossey-Bass.

Laumann, E. O., Gagnon, J. H., Michael, R. T., & Michaels, S. (1994). *The social organization of sexuality: Sexual practices in the United States*. Chicago: University of Chicago Press.

Lyons, T. A. (1983). Lesbian mothers" custody fears. *Women and Therapy, 2*, 231–240.

Martin, A. (1993). *The lesbian and gay parenting handbook*. New York: HarperCollins.

Martin, J., & Dean, L. (1993). Effects of AIDS-related bereavement and HIV-related illness on psychological distress among gay men: A 7-year longitudinal study, 1985–1991. *Journal of Consulting and Clinical Psychology, 61*, 94–103.

Mattison, A. M., & McWhirter, D. P. (1994). Serodiscordant male couples. *Journal of Gay and Lesbian Social Services, 1*, 83–99.

McCandlish, B. (1987). Against all odds: Lesbian mother family dynamics. In F. Bozett (Ed.), *Gay and lesbian parents* (pp. 23–38). New York: Praeger.

McPherson, D. (1993). *Gay parenting couples: Parenting arrangements, arrangement satisfaction, and relationship satisfaction*. Unpublished doctoral dissertation, Pacific Graduate School of Psychology, Palo Alto, CA.

McWhirter, D. P., & Mattison, A. M. (1984). *The male couple: How relationships develop*. Englewood Cliffs, NJ: Prentice-Hall.

Miller, B. (1978). Adult sexual resocialization: Adjustments toward a stigmatized identity. *Alternative Lifestyles, 1*, 207–234.

Miller, B. (1979). Gay fathers and their children. *Family Coordinator, 28*, 544–552.

Miller, J. A., Jacobsen, R. B., & Bigner, J. J. (1981). The child's home environment for lesbian versus heterosexual mothers: A neglected area of research. *Journal of Homosexuality, 7*, 49–56.

Money, J., & Ehrhardt, A. A. (1972). *Man and woman, boy and girl: The differentiation and dimorphism of gender identity from conception to maturity*. Baltimore: Johns Hopkins University Press.

Osterweil, D. A. (1991). *Correlates of relationship satisfac-*

tion in lesbian couples who are parenting their first child together. Unpublished doctoral dissertation, California School of Professional Psychology, Berkeley, CA.

Pagelow, M. D. (1980). Heterosexual and lesbian single mothers: A comparison of problems, coping and solutions. *Journal of Homosexuality*, *5*, 198–204.

Patterson, C. J. (1992). Children of lesbian and gay parents. *Child Development*, *63*, 1025–1042.

Patterson, C. J. (1994a). Children of the lesbian baby boom: Behavioral adjustment, self-concepts, and sex-role identity. In B. Greene & G. M. Herek (Eds.), *Contemporary perspectives on lesbian and gay psychology: Theory, research, and applications* (pp. 156–175). Beverly Hills: Sage.

Patterson, C. J. (1994b). Lesbian and gay couples considering parenthood: An agenda for research, service, and advocacy. *Journal of Gay and Lesbian Social Services*, *1*, 33–55.

Patterson, C. J. (1995a). Lesbian mothers, gay fathers, and their children. In A. R. D'Augelli & C. J. Patterson (Eds.), *Lesbian, gay and bisexual identities over the lifespan* (pp. 262–290). New York: Oxford University Press.

Patterson, C. J. (1995b). Gay and lesbian parenthood. In M. H. Bornstein (Ed.), *Handbook of parenting* (pp. 255–274). Hillsdale, NJ: Erlbaum.

Patterson, C. J. (1995c). Adoption of minor children by lesbian and gay adults: A social science perspective. *Duke Journal of Gender Law and Policy*, *2*, 191–205.

Patterson, C. J. (1995d). Families of the lesbian baby boom: Parents" division of labor and children's adjustment. *Developmental Psychology*, *31*, 115–123.

Patterson, C. J. (1997). Children of lesbian and gay parents. In T. Ollendick & R. Prinz (Eds.), *Advances in clinical child psychology* (Vol. 19, pp 235–282). New York: Plenum.

Patterson, C. J., & Chan, R. W. (1997). Gay fathers. In M. E. Lamb (Ed.), *The role of the father in child development* (3rd ed., pp. 245–260). New York: Wiley.

Patterson, C. J., Hurt, S., & Mason, C. (in press). Families of the lesbian baby boom: Children's contacts with grandparents and other adults. *American Journal of Orthopsychiatry*.

Patterson, C. J., & Redding, R. (1996). Lesbian and gay families with children: Public policy implications of social science research. *Journal of Social Issues, 52,* 29–50.

Paul, J. P. (1986). *Growing up with a gay, lesbian, or bisexual parent: An exploratory study of experiences and perceptions*. Unpublished doctoral dissertation, University of California at Berkeley, Berkeley, CA.

Paul, J. P. (1996). Bisexuality: Exploring/exploding the boundaries. In R. C. Savin-Williams & K. M. Cohen (Eds.), *The lives of lesbians, gays and bisexuals: Children to adults* (pp. 436–461). New York: Harcourt Brace.

Paul, J. P., Hays, R. B., & Coates, T. J. (1995). The impact of the HIV epidemic on US gay male communities. In A. R. D'Augelli & C. J. Patterson (Eds.), *Lesbian, gay and bisexual identities over the lifespan: Psychological perspectives* (pp. 347–397). New York: Oxford University Press.

Peplau, L. A. (1991). Lesbian and gay relationships. In J. C. Gonsiorek & J. D. Weinrich (Eds.), *Homosexuality: Research implications for public policy* (pp. 177–196). Newbury Park, CA: Sage.

Peplau, L. A., & Cochran, S. D. (1990). A relationship perspective on homosexuality. In D. P. McWhirter, S. A. Sanders, & J. M. Reinisch (Eds.), *Homosexuality/heterosexuality: Concepts of sexual orientation* (pp. 321–349). New York: Oxford University Press.

Peplau, L. A., Padesky, C., & Hamilton, M. (1982). Satisfaction in lesbian relationships. *Journal of Homosexuality*, *8*, 23–35.

Peplau, L. A., Veniegas, R. C., & Campbell, S. M. (1996). Gay and lesbian relationships. In R. C. Savin-Williams & K. M. Cohen (Eds.), *The lives of lesbians, gays, and bisexuals: Children to adults* (pp. 250–273). New York: Harcourt Brace.

Pies, C. (1985). *Considering parenthood.* San Francisco: Spinsters/Aunt Lute.

Pies, C. (1990). Lesbians and the choice to parent. In F. W. Bozett & M. B. Sussman (Eds.), *Homosexuality and family relations* (pp. 137–154). New York: Harrington Park.

Polikoff, N. (1990). This child does have two mothers: Redefining parenthood to meet the needs of children in lesbian mother and other nontraditional families. *The Georgetown Law Review*, *78*, 459–575.

Pollack, S., & Vaughn, J. (1987). *Politics of the heart: A lesbian parenting anthology*. Ithaca, NY: Firebrand.

Rand, C., Graham, D. L. R., & Rawlings, E. I. (1982). Psychological health and factors the court seeks to control in lesbian mother custody trials. *Journal of Homosexuality*, *8*, 27–39.

Rees, R. L. (1979). *A comparison of children of lesbian and single heterosexual mothers on three measures of socialization.* Berkeley, CA: California School of Professional Psychology.

Remafedi, G. (1987). Male homosexuality: The adolescent's perspective. *Pediatrics*, *79*, 326–330.

Ricketts, W. (1991). *Lesbians and gay men as foster parents.* Portland, ME: National Child Welfare Resource Center, University of Southern Maine.

Ricketts, W., & Achtenberg, R. (1990). Adoption and foster parenting for lesbians and gay men: Creating new traditions in family. In F. W. Bozett & M. B. Sussman (Eds.), *Homosexuality and family relations* (pp. 83–118). New York: Harrington Park.

Riley, C. (1988). American kinship: A lesbian account. *Feminist Issues*, *8*, 75–94.

Rivera, R. (1991). Sexual orientation and the law. In J. C. Gonsiorek & J. D. Weinrich (Eds.), *Homosexuality: Research implications for public policy* (pp. 81–100). Beverly Hills, CA: Sage.

Robinson, B.E., & Skeen, P. (1982). Sex-role orientation of gay fathers versus gay nonfathers. *Perceptual and Motor Skills*, *55*, 1055–1059.

Rohrbaugh, J. B. (1988). Choosing children: Psychological issues in lesbian parenting. *Women and Therapy*, *8*, 51–63.

Rubenstein, W. B. (1991). We are family: A reflection on the search for legal recognition of lesbian and gay relationships. *The Journal of Law and Politics*, *8*, 89–105.

Rubenstein, W. B. (1996). Lesbians, gay men, and the law. In R. C. Savin-Williams & K. M. Cohen (Eds.), *The lives of lesbians, gays, and bisexuals: Children to adults* (pp. 331–344). New York: Harcourt Brace.

Sameroff, A. J., & Chandler, M. (1975). Reproductive risk and the continuum of caretaking casualty. In F. D. Horowitz (Ed.), *Review of child development research* (Vol. 4, pp. 187–244). Chicago: University of Chicago Press.

Savin-Williams, R. C. (1990). *Gay and lesbian youth: Expressions of identity*. New York: Hemisphere.

Sbordone, A. J. (1993). *Gay men choosing fatherhood.* Unpublished doctoral dissertation, Department of Psychology, City University of New York.

Skeen, P., & Robinson, B. (1985). Gay fathers" and gay nonfathers" relationships with their parents. *Journal of Sex Research*, *21*, 86–91.

Steckel, A. (1987). Psychosocial development of children of lesbian mothers. In F. W. Bozett (Ed.), *Gay and lesbian parents* (pp. 75–85). New York: Praeger.

Strommen, E. F. (1989a). "You're a what?" Family members" reactions to the disclosure of homosexuality. *Journal of Homosexuality*, *18*, 37–58.

Strommen, E. F. (1989b). Hidden branches and growing pains: Homosexuality and the family tree. *Marriage and Family Review*, *14*, 9–34.

Tasker, F., & Golombok, S. (1991). Children raised by lesbian mothers: The empirical evidence. *Family Law*, *21*, 184–187.

Tasker, F., & Golombok, S. (1995). Adults raised as children in lesbian families. *American Journal of Orthopsychiatry*, *65*, 203–215.

US Bureau of the Census (1995). *Statistical abstract of the United States*, *1995*. Washington, DC: US Department of Commerce.

Weston, K. (1991). *Families we choose: Lesbians, gays, kinship*. New York: Columbia University Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.17
Families of Divorce

ARNOLD L. STOLBERG, ELIZABETH MULLETT, and
EUGENE V. GOURLEY
Virginia Commonwealth University, Richmond, VA, USA

9.17.1 INTRODUCTION 275
9.17.2 DIVORCE, HOSTILITY, AND CHILD ADJUSTMENT 275
  9.17.2.1 Gender Differences in Response to Marital Conflict 278
9.17.3 IMPLICATIONS FOR INTERVENTION 278
9.17.4 MEDIATION 279
9.17.5 COMMUNICATION 279
9.17.6 DEFINING THE INTERVENTION 280
  9.17.6.1 The Conceptual and Empirical Underpinnings of the Therapeutic Process 280
  9.17.6.2 Understanding the Contributions of Each of the Parents 284
  9.17.6.3 Selecting the Intervention Targets 285
  9.17.6.4 Evaluating Client Motivation 285
  9.17.6.5 Increasing Client Motivation 285
  9.17.6.6 Resolving Disputes and Teaching Coparenting Skills 286
  9.17.6.7 Evaluating and Monitoring Therapeutic Gains 287
9.17.7 CONCLUSIONS 287
9.17.8 REFERENCES 288

9.17.1 INTRODUCTION

Approximately one and a half million US children experience their parents' divorce annually. Of this population, 250 000 children will experience serious adjustment problems attributable to this family event (Ellwood & Stolberg, 1993). Most of these families will take their conflict to the courtroom for resolution. The significance of the problem to mental health delivery systems is further dramatized by the fact that almost half of all children seeking public and private psychological counseling are from families of divorce.

The size of the problem is apparent. The cause of the problem, interparent hostility, is clear. The need for an effective strategy to limit divorce-related hostility and to promote effective coparenting is obvious. The foci of this chapter are to clarify the processes underlying interparent hostility which lead to psychological problems in children, to suggest essential elements for a parent-focused intervention to reduce hostility and to promote coparenting, to review existing treatment strategies, and to suggest a clinical approach which will build on the strengths of existing programs and which will attend to psychopathogenic processes.

9.17.2 DIVORCE, HOSTILITY, AND CHILD ADJUSTMENT

Marital conflict has been associated with a wide array of child behavior problems (Emery, 1982; Grych & Fincham, 1990). These problems include externalizing problems such as conduct disorder (Dadds, Sanders, Morrison, & Rebgetz, 1992; Emery & O'Leary, 1984; Jouriles, Barling, & O'Leary, 1989; Wierson, Forehand, & McCombs, 1988), delinquency/antisocial behavior (Emery, & O'Leary, 1984; Mann, Borduin, Henggeler, & Blaske, 1990; Vuchinich, Bank, & Patterson, 1992), and personality disorders (Porter & O'Leary, 1980) as well as internalizing problems such as anxiety, withdrawal (Porter & O'Leary, 1980; Wierson et al., 1988), and depression (Dadds et al., 1992). Other studies have found correlations between marital distress and cognitive functioning (Wierson et al, 1988), social functioning (Emery & O'Leary 1984; Jouriles et al., 1989; Wierson et al., 1988) and grade point average in the children (Wierson et al., 1988).

Past research has focused primarily on correlations between marital conflict and behavior problems. The direct mechanisms through which conflict becomes detrimental to children have yet to be clearly identified (Katz & Gottman, 1993). However, three primary theories have each proposed different ways in which marital conflict affects children. These are social learning, socialization, and family systems theories.

Social learning theory suggests that children may be learning ineffective problem solving and coping skills by modeling parents in conflicted marriages (Bandura, 1973). Hetherington, Cox, and Cox (1982) found that the degree to which children are exposed to marital conflict is directly related to levels of child behavior problems. Other studies have found that children who hear adults fighting were more likely to engage in aggressive play (Cummings, Iannotti, & Zahn-Waxler, 1985). Johnson and O'Leary (1987) examined girls with conduct disorder and the families of these girls were marked by higher levels of parental aggression and hostility. Furthermore, the girls' responses to conflict were closer to those of their mothers than their fathers, thus supporting social learning theory. It has also been proposed that the behavior of the parent of the same sex may be a more influential model for a child in dealing with conflict and aggression. There is some modest, but not always conclusive, support for this position (Turner, Gervai, & Hinde, 1993).

Socialization theory suggests that marital conflict is detrimental to children because of its effects on parenting practices (Belsky, 1984). Marital conflict may have a "spill over effect," whereby the negative effect in marital disharmony carries over into parent–child relations. Christensen and Margolin (1988) found that in distressed families marital conflict was not contained between spouses, but spread to parent–child relationships. Furthermore, parent–child aggression, which may be a consequence of marital conflict, has been related to higher rates of antisocial behavior (Patterson, 1982; Vuchinich et al., 1992) and behavior problems (Friedrich & Einbender, 1983). Jouriles et al. (1989) found that witnessing interspousal aggression was correlated with parent–child aggression, and that parent–child aggression was related to attention problems, anxiety/withdrawal, motor excess, and conduct problems in children.

Much of the influence of marital conflict on children's adjustment has been found to be mediated by disrupted parenting practices in conflicted families (Belsky, 1984). Marital conflict may weaken unified parental alliances (Christensen & Margolin, 1988; Gilbert & Christensen, 1988) and may prevent parents from presenting a unified front to children (Stoneman, Brody, & Burke, 1989). Several investigators found weak unified parental alliances in maritally distressed families, resulting in inconsistency between parents' child rearing practices (Christensen & Margolin, 1988; Gilbert & Christensen, 1988). In turn, inconsistent discipline is related to conduct problems, aggression, and juvenile delinquency (Patterson, 1977, 1986; Patterson & Stouthamer-Loeber, 1984). Weak parental alliances have been repeatedly found in maritally distressed families (Christensen & Margolin, 1988; Gilbert & Christensen 1988) and are closely related to lower levels of effective family problem solving (Vuchinich, Vuchinich, & Wood, 1993).

Family systems theory suggests that conflict is detrimental to children because of its effects on the functioning of the family system (Kerr & Bowen, 1988). Kerr and Bowen (1988) theorize that each person in the family is emotionally invested in keeping the family together and each person has a role or function in its maintenance. Consequently, all members are likely to be influenced by marital conflict because of the tremendous threat marital conflict poses to the integrity of the family.

Individuals' responses to high levels of conflict may involve coping mechanisms aimed at reducing anxiety and tension among family members, even if it comes at a price for the individual. Grych and Fincham (1993) found that factors such as the intensity, content, and cause of the fight increased children's distress and their desire to intervene in the fight. Children proposed indirect intervention (e.g., distracting techniques) if they thought the marital fight would escalate. If the fight was child related, children reported increased shame and fear of being drawn into the conflict. These children were also more likely to suggest that they would directly intervene in the fight. Children exposed to greater frequencies of marital conflict were also more likely to comfort their mother following an angry confrontation in an experimental session (Cummings, Vogel, Cummings, & El-Sheikh, 1989). Again, children were also more likely to propose that they would directly intervene in the fight. These studies suggest that children may intervene in marital conflict in attempt to decrease tension and anxiety.

Hinde and Stevenson-Hinde (1988) suggested that a child may become involved in marital conflict in several ways. A child may enter the conflict and either mediate or side with one person and against another. A child may also reduce tension by distracting the other two from fighting. In addition, a child may also be drawn into the conflict when parents ally themselves and blame the child for the problem, in a process known as scapegoating (Hinde & Stevenson-Hinde, 1988).

A child's involvement in marital conflict may not only decrease a child's anxiety, but also benefit the parents. By entering a conflict, and siding with one parent, a child may provide support for a parent. Kerr and Bowen (1988) theorize that triads are inherently more stable relationship units than dyads. In the case of marital conflict, the same amount of anxiety or hostility is spread among three people instead of two. Scapegoating and third party manipulations are relationship patterns involving more than two people which clinicians often cite as families' attempt to divert attention away from the primary source of conflict, thereby reducing tension and stabilizing the family.

A child's involvement in marital conflict may temporarily help decrease tension in the family, but, it may have detrimental long-term effects on the child. Minuchin (1974) suggested that a child's involvement in cross-generational alliance may have negative psychological effects on the child. A child who sides against a parent may experience anxiety because he/she also betrays the other parent. Furthermore, the parent's alliance with a child breaks the traditional family hierarchy, creating conflict and confusion for the child. Problems could also develop because parents may become over-involved in the child's life, undermining the development of autonomy and peer relationships.

Katz and Gottman (1993) found that parents' coping styles with marital distress predicted types of behavior problems experienced by children three years later. A father's withdrawal from the family was associated with internalizing problems three years later, while mutual conflict and hostility between the parents predicted externalizing problems in children.

Previous research also has found evidence for the existence of several coping mechanisms proposed by family systems theorists. Vuchinich, Emery and Cassidy (1988) examined third party interventions and patterns of interactions in families with adolescents. They found that in triadic interactions, family members intervened in conflict one-third of the time and alliances were formed approximately half of the time. These alliances were evenly spread between all members, however, parents rarely intervened in conflicts between their spouse and one child. Even more salient was the finding that parents rarely sided against each other. However, this study did not assess differences between conflicted and non-conflicted families.

Unlike Vuchinich et al. (1988), Christensen and Margolin (1988) and Gilbert and Christensen (1988) examined both maritally conflicted and nonconflicted families for the presence of cross-generational alliances, and weak marital alliances. Using sequential analytic techniques, Gilbert and Christensen (1988) found significant levels of cross-generational alliances in maritally distressed families. Christensen and Margolin (1988) assessed cross-generational alliances and weak marital alliances using slightly different methods. Weak marital alliances were also found in maritally conflicted families.

A study by Mann et al. (1990) found mixed support for the existence of cross-generational alliances. Delinquency in adolescents was moderately correlated with the presence of cross-generational alliances. In observations of family dyads, delinquent adolescents were more likely than nondelinquent adolescents to be disengaged from their fathers and aligned with their mothers. The parental dyads in families with delinquent adolescents were also more conflicted than those with well adjusted adolescents. However, within individual families, delinquent adolescents were not more aligned with their mothers and conflicted with their fathers. Also, within delinquent families, parental dyads were not more conflicted than the mother and child dyads. Therefore, Mann et al. suggested that cross-generational alliances may be more illusory than real. However, their results also revealed that in families which received multisystemic therapy, improvements in marital relations were linked to the amelioration of symptoms in both adolescents and their parents. These improvements were not found if adolescents received individual therapy only. Mann et al. stressed the need for researchers to go beyond observing family dyadic relations in order to understand family systems.

9.17.2.1 Gender Differences in Response to Marital Conflict

Gender differences have been found in children's psychological adjustment to marital conflict; however, the results have been inconsistent (Grych & Fincham, 1990). Conduct disorder has been found to be related to marital discord for boys but not for girls (Emery & O'Leary, 1984; Porter & O'Leary, 1980). However, other studies involving nonclinical families have found similar levels of pathology in boys and girls exposed to marital conflict (Emery & O'Leary, 1984; Jouriles et al., 1989). Although the response to conflict in boys and girls may differ, Emery and O'Leary (1984) found that both boys and girls accurately assessed the severity of marital conflict. Therefore, Grych and Fincham (1990) suggested that gender differences in behavior problems may be due to factors associated with marital distress other than marital conflict.

A study by Kerig, Cowan, and Cowan (1993) suggested that gender differences in children's psychological adjustment may be related to the interaction styles of their parents. Overall, parents tended to negate daughters more than sons. However, when marital satisfaction was low, parent's interactions with their children were increasingly gender differentiated. In distressed marriages, fathers were more negative towards their daughters. Mothers were more likely to negate daughters in response to nontraditional behaviors (e.g., daughters' self assertion). The overall level of negative interactions between a mother and her son did not change according to levels of marital quality. However, mothers in distressed marriages were more likely to reciprocate their son's negative affect. These results suggest that, although overall levels of negativity may not differ in maritally distressed and nondistressed families, different patterns of interactions and behavior contingencies may emerge between parents and their sons and daughters under higher levels of marital conflict.

Differences in boys' and girls' adjustment to marital conflict may also be association with gender differences in emotional responses and intervention in marital conflict. Cummings et al. (1989) found that boys and girls responded differently to marital conflict. After the age of 10, boys reported greater anger and girls reported greater fear in response to marital conflict (Cummings et al., 1985, 1989). Girls were also more likely to propose solutions to conflicts and were more likely to suggest direct intervention in the fight (Cummings et al., 1989; Vuchinich et al., 1988). Vuchinich et al. found that girls got more involved in all forms of conflict, except marital conflict. Miller, Danaher, and Forbes (1986) found that girls were more likely to mitigate conflict than boys, although this was not supported in another study (Cummings et al., 1989).

9.17.3 IMPLICATIONS FOR INTERVENTION

Several processes account for much of the deleterious impact of hostility on children's adjustment to divorce. Children may imitate their parents' poor problem solving and coping strategies and may display these aggressive behaviors with their peers. Hostility may also interfere with normal, unified parental alliances and may deteriorate the quality of the parenting which the child receives. Children may enter the conflict between parents. They may attempt to mediate the conflict between their parents and/or may side with one against the other. They may also attempt to reduce the conflict by distracting their parents onto other activities and issues. In all cases, the unusual level of responsibility for marital harmony which children accept leads to significant emotional distress and may subsequently interfere with social, academic, and emotional adjustment.

Interparent hostility associated with divorce appears to negatively impact children through five processes. First, children may model the inappropriate conflict resolution skills they see displayed by their parents and may use these strategies in the playground and in the classroom. Much of the peer and social problems common to children of divorce may result from these displays. Second, the hostility expressed between parents may deteriorate their parenting by redirecting their energies from the needs of the children onto the most pressing interspousal issue. Third, scapegoating, or blaming the children for marital problems, may lead the children to inappropriately accept blame for problems within the family. Poor self esteem common to some children of divorce may result from this process. Fourth, the emotional spillover from the anger conveyed by their parents may lead children to experience significant emotional distress. Finally, children often become engaged in the conflict, either by trying to distract parents from their discord or by joining in alliances with one parent against the other.

Preventing and ameliorating children's adjustment problems associated with divorce and interparent hostility may be accomplished by efforts directly aimed at the parents and by attending to the pathogenic processes in the family. Such strategies include teaching parents effective conflict resolution skills. Mediation and alternative dispute resolution may provide a model for improving these skills. By providing a forum to keep the focus on resolving the primary divorce issues, residual deterioration of necessary parenting may not occur. Further, by providing strategies to develop effective and appropriate conflict resolution skills, children may not experience emotional distress resulting from the spillover from their parents' conflict. Helping parents to maintain their focus on the children's divorce-specific and ongoing developmental needs, with a particular emphasis on directing parents to maintain healthy family alliances, may help to preserve effective parenting roles of the conflicted adults.

No programs exist which are aimed at reducing interparenting following divorce and concurrently promoting parenting effectiveness. Two or three existing strategies may be applied to promote healthy and essential aspects of a post-divorce, developmental environment. Mediation may provide strategies for enhancing conflict resolution skills, for resolving conflicts, and for keeping the attention on problem resolution and away from tangential and undeserved blame for parenting inadequacies. The most effective communication skill training package was probably that developed by Hendrix (1989) as a part of his marriage enhancement program.

9.17.4 MEDIATION

Mediation is an increasingly popular alternative to traditional methods of resolving child-related divorce and separation issues (Emery & Wyer, 1987). Proponents of mediation argue that cooperative, self-determined processes for resolving disputes minimize the reduction in family functioning that often follows divorce (Robinson & Parkerson, 1985). Mediation gives parents the rights and responsibilities to determine the nature of the post-divorce family (Trombatta, 1982). Despite support for mediation and the growth of court-sponsored programs, the disputes best suited for intervention and the impact of the process on the family are not well described in the literature (Emery & Wyer, 1987).

Mediation is a task-oriented, problem solving activity that seeks to resolve clearly defined issues related to the on going relationship between divorcing couples (Coogler, 1978; Kressel & Deutsch, 1977). This short term, prevention oriented process seeks to minimize maladaptive outcomes typically associated with marital discord and divorce by directing parties in negotiations towards a fair agreement between themselves. Mediation differs from traditional methods of resolving disputes (either in court or through attorneys out of court) in that it is a cooperative and self-determined process. Parental status and parenting responsibilities are considered to be unalterable. Self-determined resolutions are considered better than those imposed by external forces (Neville, 1985).

Mediation differs from traditional methods of resolving disputes in the manner in which decision-making power is distributed (Girdner, 1985). In traditional methods, couples relinquish their power to the judge or attorney when disputes are settled in the courtroom or in an attorney's office (Pearson & Thoennes, 1984). On the other hand, mediating couples maintain control over the process because they are required to present mutually determined and beneficial solutions to each other, and not to a judge (Neville, 1985).

Research assessing the effectiveness of mediation in keeping couples out of court and in promoting family adjustment to divorce is encouraging but limited. Subjective reports following mediation indicate that couples are more likely to be satisfied with the outcome and to indicate that their lives have improved than those involved in the adversarial process (Emery & Wyer, 1987). A significant gender bias in outcomes has been found, with men reporting greater satisfaction with the process and a positive impact on all family members. Women reported less satisfaction and said that only the children benefited (Emery & Wyer, 1987). Court-based programs, which are the most often studied, have reported rates of agreement in men's and women's satisfaction with outcomes between 50% and 75% (Emery & Wyer, 1987; McIsaac, 1982; Pearson & Thoennes, 1984). These studies have failed, however, to examine the full range of couples who use mediation and have excluded noncustody disputes.

9.17.5 COMMUNICATION

Communication is a bidirectional activity which requires both active listening and clear and intentional expression of ideas. Hendrix (1989) described a multistep procedure in which each of the participants takes turns articulating his or her concerns and, subsequently, listening to the other's ideas. In the first step, the conveyer states his or her ideas and articulates the content of and then the feelings about the message. In an effort to ensure that the receiver focuses on what is being said and is not preoccupied by formulating his or her response, the receiver is then asked to report the message they heard, the emotional import to the conveyer, and the extent to which the receiver understood and accepted the message. In the third step of the process, the conveyer evaluates the accuracy of the receiver's response and the extent to which he or she feels adequately heard. If an accurate exchange of the conveyer's ideas has not occurred, he or she clarifies what was heard and the receiver conveys his/her understanding of the clarified message. If a clear message has been conveyed and received, the participants change roles and activities. This process continues until both participants have communicated all of their ideas and concerns on the topic.

This structured communication process ensures that communication is a bidirectional process; messages are stated clearly and are heard exactly as they were intended. Procedures are included which keep both participants' attention on the immediate activity and not on formulating their own ideas at the expense of hearing the concerns of the other. Because procedures are included which allow the conveyer to correct the misperceptions of the receiver and because each has a turn at explaining their thoughts, yelling and anger do not result. However, this approach has not been evaluated for its effectiveness.

9.17.6 DEFINING THE INTERVENTION

Clinical experience dictates that any intervention to promote coparenting and to reduce hostility between divorcing parents must be built on a conceptual underpinning which puts the developmental needs of the children above most of the needs of their parents. Cooperation, compromise, objectivity, and confrontation of the parents with their contributions to their children's adjustment problems are essential for success. This is a significant departure from traditional approaches to working with divorcing parents. The therapist is required to be strong and directive. The therapist needs to be knowledgeable about the divorce literature, equally supportive of both parents, and, above all, to be able to gain an objective understanding of the dynamics of the family and the ways in which each of its members contributes to the children's adjustment. It is almost always the case that each parent contributes to the deterioration of the children and that, in most cases, each parent can substantially influence the effectiveness of the other parent and the welfare of the children.

The role of the therapist may be construed as the superego for the divorcing family—as a parent to the conflicting adults. Few outside relationships exist for the couple which require cooperative coparenting behaviors and which encourage for the psychological adjustment of the children. Attorneys, family members, and close friends are the strongest influences on parenting behaviors (Camplair & Stolberg, 1987). They often advocate for the needs of their clients, family members, or friends and promote adversarial stances. In fact, five years after the separation most divorcing parents view the influences of these professionals and friends as destructive of the parenting process. The clinician is required to convey an expectation which considers the long-term needs of each of the family members. When the therapist working with high hostility, divorcing couples must define the parameters of the parents' behaviors toward each other and with their children. The therapist must help the parents develop a perspective which sets as the priority the future adjustment of both of the parents and of the children, and not the immediate emotional needs of any individual.

The manner in which the therapist responds to the clients as individuals, and the therapeutic relationship which is established, is essential to the success of the therapy. Support, validation, love, and respect are essential elements of this relationship. Divorce is one of the most difficult periods in an individual's life. Most divorcing adults feel worthless and unloved by their spouses, the most important people in their adult lives. The support for and validation of the client as a valued person, who can respond to demanding circumstances in competent and decent ways, and who is cared for by another person with societal esteem, facilitates the client to behave in a desirable and health promoting manner. Many therapists feel anger and frustration when spouses are engaged in meaningless conflict. This distress by the therapist must be conveyed in a way that promotes the best in the individuals and not with statements that make them feel foolish or dysfunctional. The therapist must act as the authoritative parent who can lovingly listen to the concerns of the children and can incorporate their concerns in the final decisions that the parent makes. The steps of this intervention process are discussed below.

9.17.6.1 The Conceptual and Empirical Underpinnings of the Therapeutic Process

Working with children of divorce and their families requires a keen understanding of the clinical and empirical research literature describing this family event. A review of the literature on divorce and child adjustment identifies factors within the child, parents, and the broader environment in which the child and parents live. The factors which appear to have the most significant influence on the development of the child include the child's age, the adults' abilities to parent the child and to work together without hostility, the life changes which the divorce brings about, the continued active involvement of both parents with the child, and the suggestions, directions, and demands of family members, close friends, and professionals working with the divorcing family.

A review of these influences may be more clearly accomplished by categorizing them according to Bronfenbrenner's ecological model (Bronfenbrenner, 1977, 1979) and the elements of the child's environment to which they belong. The first level of analysis is the individual or ontogenic level (Belsky, 1980). This viewpoint looks at how divorce affects the individual, and how one's genetic predisposition combines with developmental events in personal and family history to affect available coping mechanisms. In this discussion, the children will always be the center of the analysis, thus the ontogenic system will always refer to the children. Divorce from this level focuses on how the children respond to the intense stressor of family separation. The children must deal with the changing demands this process will make on them over time. They will need to draw on available psychological coping resources, defense mechanisms, and behaviors in order to effectively deal with the demands of family transition.

The effects of divorce on children's adjustment covary with their age, and age related developmental capacities and challenges make this an important variable to consider. Infants and toddlers require high levels of involvement in interaction because they are essentially helpless to care for themselves. Disruption in regular routines caused by instability in the household or parental conflict can often be stressful to many infants and toddlers, resulting in more distress and irritability even at an early age (Kalter, 1987, 1989). Initially, in instances where the father leaves the family home, his absence appears to have a surprisingly small effect on the children, primarily due to their inability to conceptualize relationship roles, and the fact that mothers are usually more involved at this early age. Children may benefit greatly from the father–child relationship in later development (Kalter, 1987, 1989). The research suggests that children require a stable, conflict-free environment where their demands can be met by parent figures. While father absence may not be evident in a child's behavior at this point, it is important that both parents remain involved if they are to form lasting attachments with their child.

During the preschool years, children begin to develop more independent abilities including basic communication skills, increased locomotion, and some self care skills such as dressing, washing, and toileting. However, preschool children still have very limited individual resources for coping with divorce (Kalter, 1989). For example, their ability to understand cause–effect relationships is generally limited by a lack of perspective taking ability. This results in one of the most common problems associated with divorce: self blame (Hetherington, 1991; Kalter, 1989; Stolberg, Camplair, Currier, & Wells, 1987). The child perceives the parental separation as resulting from their lack of love for the child. During parental discussion of divorce or custody issues where the child hears his or her name mentioned contributes to this belief. Children may also harbor reunion fantasies for long periods of time due to their limited understanding and reality testing (Hetherington, 1979, 1991; Kalter, 1987, 1989; Wallerstein & Blakasee, 1989). Children at this developmental stage are more aware of their emotional states being able to label emotions such as mad, sad, or happy, but have limited ability to cope with these emotions, understand their causes, or communicate them to others (Kalter, 1989).

Preschool children of divorce show significant differences in behavior from their peers from intact homes (Biller, 1981; Hetherington, 1979, 1991; Kalter, 1987). These children display more negative behavior including being more irritable, aggressive, complaining, and showing less social maturity. In addition, more anxiety, fear, and less imaginative play is observed in these children than those who do not experience parental separation (Hetherington, 1979; Kurdek, 1981; Wallerstein & Kelly, 1980). Consideration of developmental abilities, and the observed effects of divorce, have led researchers to conclude that preschoolers are a very vulnerable group because of their limited cognitive ability and the developmental importance of emerging social interactions (Kalter, 1989; Kurdek, 1981).

Elementary school age children (6–12 years) make significant gains in abstract reasoning, communication, social maturity and their ability to fantasize about possible futures (Bjorklund & Green, 1992; Flavell, 1985). Children at this age will be better able to tolerate separation from a parent and have increased sources of support from teachers and peers outside the home. In the early elementary years (6–8 years), children may experience an intense longing for the noncustodial parent (most often the father), and this longing is based on fantasy and not related to the type of relationship that previously existed between the parent and child (Fry & Addington, 1985; Hetherington & Clingempeel, 1992).

As children progress through the elementary years, they are vulnerable to conflictual feelings of loyalty to each parent. As they understand more about the divorce and their parent's lives after the divorce, they often begin to look for reasons to blame for the situation (Kalter, 1989; Wallerstein & Blakasee, 1989). In addition parents may exacerbate this by talking about the former spouse negatively, blaming the former spouse for current unhappiness or family difficulties, or by using the child as a source of information about the former spouse's behavior and life. Children at this age are likely to understand that the divorce is permanent, are able to tolerate separation, and have better coping skills. They will, however, have strong conflictual feelings for both parents, particularly when parental conflict is ongoing. Children have limited capacity to understand that mother and father are mutually unhappy, and this leads them to alternate blame from one parent to the other (Fry & Addington, 1985; Hetherington, 1991; Kalter, 1989; Wallerstein & Kelly, 1980).

Divorce during adolescence is a large disruption occurring at a time of tremendous developmental change. Children in adolescence undergo the physical changes to adulthood, becoming acutely aware of their bodies and sexuality. Cognitive changes occur as adolescents develop formal operational thought and begin the development of long-term planning skills (Flavell, 1985). Emotionally, adolescents often experience anxiety concerning their future and in developing an adult identity. Social changes occur as adolescents become more independent from parents and develop adultlike relationships with peers of both genders, making them more vulnerable to peer pressure.

The emotional experience of divorce for adolescents is often very intense as a result of their impulsivity and egocentric concern for their own feelings and development. Adolescents alternate between seeing parents in an idealized and deidealized light as they begin to integrate their positive and negative feelings towards them into a more mature and realistic view of the parent–child relationship (Forehand, Long, & Brody, 1988; Stinson, 1991). Divorce can disrupt this process as the child may experience many strong negative emotions toward both parents and create potentially long-lasting internal conflicts for the adolescent. Increased autonomy and an adult appearance may cause the parents to involve their adolescent children in the divorce process, complain about the divorce, seek social support, and require them to be responsible for their siblings (Forehand et al., 1988). Adolescents of divorce frequently experience one or both parents being highly critical, vindictive, or vengeful when talking about the other parent to them. These events can result in feelings of responsibility and guilt in the adolescent.

The Developmental Vulnerability Model, as described by Kalter (1987), speaks most directly to the child's evolving competencies which develop with age. In this model, developmental dimensions are viewed as continuously evolving sets of skills or capacities whose effectiveness is externally defined and at the time at which these skills are required. The progress of growth along these dimensions can be optimally or deleteriously affected by the five ecological system influences described in Bronfenbrenner's model and by the changing life circumstances brought about by acute transitional events. These developmental competencies are open to significant alterations if the ecological influences undergo substantial changes.

Another major ontogenic factor contributing to children's divorce adjustment is the temperament/personality of the child. Thomas and Chess (1977) found that temperament is an important variable in the child–environment interaction. When a child's temperament is inconsistent with the demands of the environment, the child's psychological health is often threatened (Ainsworth, 1979, 1989). Children with a difficult temperament often have more difficult mother–child interactions and more psychological disorders later in life (Thomas & Chess, 1977). However, parents are often able to provide the high level of stability and consistency these children need. Divorce greatly decreases the parent's ability to provide such a stable environment. Children with easy temperaments may adjust the best to environmental inconsistency and less involvement from one or both attachment figures. Difficult children may require more care, and generate a negative interaction with the parent as their demands for care increase the stress on the primary caregiver (Thomas & Chess, 1977).

As the child matures, more sophisticated behavioral and emotional coping strategies and defenses evolve as their personality develops. Because personality develops as an interaction between the person and the environment, parental separation will have a great effect on the child as they learn to cope with environmental and emotional demands. To the extent that divorce denies the child access to caring parental figures, the child may develop more problematic coping strategies which may result in ineffective or self defeating methods for coping with stress and getting emotional needs met.

Many children have special needs which require special attention or increased vigilance by the parents. Children with limited intellectual abilities, learning disabilities, attention-deficit hyperactivity disorder, or other developmental or psychological difficulties may require a consistent, safe, warm environment which can be threatened by divorce. In many cases, the parents working together may be able to relieve each other's stress and combine to provide a good child rearing environment, but are too overwhelmed or lack support to do so following divorce (Belsky, 1980).

The microsystem is the next level of analysis in the ecological model. This system includes the immediate settings that the individual comes into contact with, such as family, friends, lovers, and coworkers. In the divorce process, this system is very much in flux. Obviously, the available coping resources among family members will change for children, with the loss of one parent and children seeing less of each parent. Children's relationships with peers may also change if they have to move, or visit the other parent away from their original neighborhood and peer network. Parents, friends, and in-laws may no longer be available while new dating partners may be introduced as sources of social support.

The most significant microsystem factor in children's divorce adjustment is the amount of conflict and hostility between parents. Several studies have shown that children who become the focus of parental conflict fare the worst emotionally and behaviorally (Hetherington, 1991; Wallerstein & Kelly, 1980). Interparent conflict is a significant source of problems for children in both intact and divorced homes (Emery, 1982). Clearly, one major area where parents can help to buffer their children from negative effects of divorce are in limiting the amount of interparent conflict they are exposed to. This is most difficult in cases of adjudicated custody when children are involved in the custody dispute through legal and mental health professionals. Limiting the interparent conflict is of prime importance for buffering children from the deleterious effects of parental separation.

Life change events in the child's life are an important microsystem influence for children's adjustment to divorce. Several clinical and research studies indicate that the number of life changes that occur during a parental divorce are associated with symptoms of psychological distress (Kalter, Pickar, & Lesowitz, 1984; Stolberg & Bush, 1985; Stolberg et al., 1987). In general, the more life changes that occur, the more vulnerable the child will be to negative psychological and health consequences. For some children divorce may constitute multiple life changes. When children remain in the same neighborhood and are able to maintain the same friends, school, and teachers, they also maintain a high level of stability. In contrast, the child who must move will face a series of stressful life events, such as changes in school, friends, neighborhood, or even a new city and state. These children are at higher risk for psychological disorders, as stress is increased and coping resources must be developed anew.

One of the keys to the whole divorce process is the extent to which the parents are able to cope with the change and distress they are experiencing and are able to maintain their ability to be available to their children. The stress of divorce diminishes the emotional resources a parent has available for the child, often making a good parent marginal, and a marginal parent bad, at least on a temporary basis (Kalter, 1987). Even the most exceptional parent will have a difficult time maintaining their previous level of parenting competence. Divorced parents are less consistent in their discipline, make fewer maturity demands on their children, and lack control over their children's behavior when compared to intact parents (Bloom, Asher, & White, 1978; Emery, 1982). Fortunately, research points to a readjustment period following one year after divorce when most parents begin to regain their original equilibrium and achieve stability in their lives.

The exosystem is the system into which the individual and microsystems are imbedded. This includes settings that often do not directly involve the individual but affect the person due to their influence on what resources are made available and how the community responds to divorce. These will often involve community resources available for divorcing families and may include school systems, local governments, and the availability of mental health services and other community programs (Fry & Addington, 1985). For example, a parent may not be immediately involved in allocating tax dollars for school-based interventions for children of divorce, but both the parent and children will be indirectly affected by the availability of these resources.

As with children, adults experience life changes as stressful and will be more likely to report psychological symptoms, such as anxiety and depression, as the number of life changes increases. Children may not directly experience many life changes of the adults, such as getting a new job, returning to school, changing friends, or beginning to date again, but they are indirectly influenced by their parents' ability to cope with these changes. Thus, the number of life changes an adult is experiencing will be related to their psychological health and ability to provide stability, nurturance, and support to their children.

A major exosystem factor influencing child adjustment is a drop in family income. Even with a financially supportive noncustodial parent, two incomes now go to support two households instead of one. This temporary drop in income often becomes permanent for the custodial parent, who must cope with increased child care demands while maintaining job performance (Weitzman, 1985). When the drop in income coincides with a change in residence, often the move is to a more impoverished neighborhood, and away from familiar friends, neighbors, and schools (Fry & Addington, 1985; Weitzman, 1985).

Divorced parents have lost their spouse, often the main source of companionship and social support. In order to have social support needs met, kin networks, neighbors, churches, or friends must help to take up the slack and allow the parent to develop social support outside the home. Social support has been linked to increased ability to cope with stress, increased self esteem, and less physical and psychological illness (Sarason, Sarason, & Pierce, 1990; Vaux; 1988). This support is often essential for the parent to effectively adjust to the divorce and be an effective parent figure. Availability of babysitters, daycare, and single parent support groups are all important factors which can aid the divorced parent to recover from the stress of a major life change.

Primary decision influencers, such as attorneys, pediatricians, mental health provides, clergy, and friends and family members, are exosystem influences on adjustment (Camplair & Stolberg, 1987, 1990; Davis & Stolberg, 1988). These individuals all indirectly affect the child's welfare by providing support, encouragement and advice to the divorcing parents. As previously mentioned, much of the direction given by these sources does not promote cooperative coparenting.

The macrosystem which encompasses the ontogenic, microsystem, and exosystem includes the institutions and values of the greater society. These may include policies implemented by the state and federal governments as well as the values and attitudes of American society as a whole. Clearly, state and federal law can affect individual family members due to their impact on divorce laws, child custody awards and enforcement, and economic aid to families in transition (Deed, 1991). Society's cultural values also impact family members by helping to determine how divorce is viewed by the community, society and the government. For example, when the social climate views divorce as a negative event and creates negative attitudes and stereotypes of divorcing families, reaction at each community level may be more negative.

Macrosystem influences fluctuate with social, moral, and political changes in the culture. Divorced families are often pathologized as broken, abnormal, or disintegrated. Sex role attitudes often play a part in the ability of women to have access to the workplace and for men to be a primary caregiver (Hartup, 1989). Society's view of the status of children also plays an important role in divorce. As children's rights have become more accepted, their welfare has become more primary to the divorce and custody process.

The final level of analysis, the mesosystem, concerns the interaction and connections among systems, settings, and people. Interaction between systems can help promote more flexibility and better solutions to problems. Perhaps most important is opening communication between the individuals and others in their microsystem. Children and parents both need to be able to communicate needs and feelings with each other and to friends, teachers, and sources of aid and support. Even though the parents may separate, they must be able to effectively communicate to provide for their children. Communication between individuals, families, teachers, school boards, and state governments can help inform policy makers of effective methods for helping children adjust to divorce. Policy makers and community leaders can make use of education and outreach programs to bridge gaps between the macrosystem and exosystem to help make families more aware of what resources are available and how to access them.

9.17.6.2 Understanding the Contributions of Each of the Parents

Once therapists understand and accept their role in this constructive, therapeutic process and are thoroughly acquainted with the literature describing the influences on children's adjustment to divorce, they may begin the intervention process. The first intervention goal is to gain an understanding of each parent's view of the contributors to the hostility and poor coparenting, and each parent's emotional and parenting competencies. The therapist must gain each parent's trust and perception that the therapist is skilled and objective. This may be most effectively accomplished by meeting with each parent individually to gain the perspective they hold and by conducting a thorough personality and parenting assessment of each of the adults. Parents are generally more aware of the weaknesses of the estranged spouse than they are of their own. Taking the time to gain the perspectives of both parents and to do so with a cautious acceptance of each of the parent's "stories" will yield the breadth and objectivity that is required to proceed.

In addition to understanding the parents' contributions to the children's adjustment problems, the therapist must also understand the extent to which the child's normal and extraordinary emotional, physical, social, and academic demands contribute to the observed difficulties and the extent to which they place demands on the parents which may be beyond their ability to respond. Traditional child and parenting assessment, as well as confirmation from outside sources, is necessary.

9.17.6.3 Selecting the Intervention Targets

The third step of the process, or identifying the targets for the intervention naturally extend from the previous efforts. Before meeting with both parents together to convey the therapeutic goals, it is necessary for the therapist to develop a comprehensive conceptualization of the family dynamics which account for the child's present level of adjustment and which will help to foster his or her future development.

The therapist must clearly, objectively, and supportively articulate the dynamic formulation and convey the essential steps that are necessary to promote the child's adjustment and the effective coparenting of the child. It is important that the therapist does not convey good guy and bad guy roles to the parents. Rather, it is necessary to communicate how each parent helps and hurts the child, and how each parent can help the other to be more effective. Honesty and support are essential.

9.17.6.4 Evaluating Client Motivation

Determining each of the parent's willingness to participate in evolving cooperative problem solving is the fourth step in this therapeutic process. The most difficult aspect of working with high conflict divorces is getting the adults to acknowledge their own contributions to the hostility and to put aside their own deserved or unreasonable anger at their estranged spouse. Increasing each of the parent's motivation to put the developmental needs of the children above his or her own hurts and to actively and constructively work with the other parent is the most essential, and difficult, goal of the therapy.

Many parents will express some skepticism about their ability to work cooperatively, although most will state that they are willing to try. The real test of the parents' willingness to compromise and to set aside differences comes when they are confronted with the changes they must make. The first opportunity for the therapist to observe this behavior comes in this stage of the process.

The problem which has the greatest potential for resolution is the first to be brought up for discussion and resolution. It is generally most prudent to state the problem and ask each of the participants what steps they might individually take to resolve the problem. This gives the parents the opportunity to accept responsibility for the problem and to gain some experience in problem solving and in compromise. It also gives the therapist an opportunity to observe the commitment of each of the participants and to identify resistance.

9.17.6.5 Increasing Client Motivation

Few parents involved in high hostility divorces reflect a willingness to actively cooperate with their former spouse and to view their parenting role in a truly generative light. The therapist will almost always be required to search for vehicles to bring parents to the table and to increase their motivation to behave in responsible ways.

The therapist does, however, have considerable authority to influence client motivation beyond trying to enhance client insight or simple pleading. Interestingly, the professionals who most often exacerbate interparent tensions, the attorneys, can be called upon to promote cooperation and to exert some control over their client. Therapists can share their diagnostic formulations about the nature of the conflict and the observations of client commitment with the adult's attorney in an effort to impose some judicial reality. Most attorneys want their clients to win. The professional requirement of the attorney is to advocate for their client. Thus, conveying information to an attorney about the client's constructive or destructive behaviors in the therapy process and in coparenting efforts provides the attorney with information that can be used to help the client. If the client is actively working to promote cooperation, then that parent has a better chance of being viewed positively by the court if litigation occurs. If the parent being represented is poorly motivated to cooperate and emits destructive coparenting behaviors, then the attorney will be forced to exert some control over the client so they can avoid any negative, judicial evaluations.

It is important to note that when working with high conflict, divorcing parents, the therapist is required to obtain signed release of information forms from each parent to discuss therapeutic progress with that parent's attorney. APA ethical guidelines prohibit sharing information gained in mediation with attorneys for the opposing party. Gaining permission in advance to talk with a parent's counsel should avoid any ethical restrictions.

The potential that the therapists have to demand cooperative coparenting comes from the information that has been gained about parental behavior and the observations the therapist may be ordered to convey when called to testify. Attorneys can anticipate judges' reactions to professional testimony, particularly when the psychologist has had the opportunity to observe the relevant behaviors of both of the litigants. Attorneys can often bring to bear pressure on their clients to behave in constructive ways when they are able to help them anticipate the legal consequences of such actions. While a strong coercive aspect exists to this approach to clinical intervention, it is being done to achieve some healthy functioning in a family that is going through an acute and stressful transition. Many of their parenting behaviors have been impaired by outside influences and by their own distress. The therapist's efforts are made in an effort to help parents to grow into the new demands and circumstances which they face and to help them to develop into effective coparents who have healthy children and who enjoy their parenting roles.

9.17.6.6 Resolving Disputes and Teaching Coparenting Skills

The sixth and seventh steps of the intervention are implemented simultaneously. They include:

(i) resolving specific disputes between parents,

(ii) teaching new skills (communication, compromise and conflict resolution, and parenting skills) and enhancing capacities in existing skills which will help parents work together after divorce, and

(iii) teaching parents about the divorce adjustment process, helping them to understand their family dynamics, and helping them to understand the influence of these two processes on the development of their children.

From a technical perspective, the therapist chooses one or two important, but solvable, problems for resolution. When choosing the target problem, the therapist will have three goals. Resolving the dispute, identifying and enhancing capacities, and shaping the couple's views about parenting after divorce. The therapist first defines the problem and then asks each of the participants to convey what they might do to achieve resolution. The exact task is for each parent to identify what they do to contribute to the problem and to define what they may do to make the situation better. It is imperative that each of the participants respond with what they can do. The therapist must limit all discussion to what each parent can do and to accept personal responsibility for the current issue.

The therapist's responsibility is to diagnose each of the parent's capacities to communicate clearly, to compromise, to set aside their own needs for those of their children, and to respect the importance of the other parent to their children. Once skill deficiencies are identified, the therapist adds the teaching of the absent skills and perspectives to dispute resolution. Directiveness is, again, required of the therapist. The therapist must set the standards for behavior, including both target skill levels as well as attitudes about mutual respect and the importance of both parents to the children.

It is very helpful to resolve several specific disputes as quickly as possible during this stage as it will give parents some hope that improvement can occur. More important, however, is the fact that by helping parents to jointly resolve one issue, you are helping them to understand the processes involved in resolving future conflicts.

Once the therapist has observed the parents' attempt to resolve a specific dispute issue and has drawn some important diagnostic conclusions about their strengths and weaknesses, he or she conveys observations about these capacities to the parents and targets skill areas in need of enhancement. Observations should include clear statements about what each of the parents is doing right and what is being done that either harms the children or interferes with resolution of the problems. In some cases, the therapist is required to comment on sources of resistance that may impair their ability to resolve disputes in mutually beneficial ways. Respect, objectivity, and honesty must characterize such statements. In the absence of directness one or both of the parents can not be expected to know how they are impairing the dispute resolution process.

The therapeutic task is also to evaluate each of the parent's understanding of the impact of the divorce on their children. When the impediment to problem resolution is a lack of understanding of the divorce process and its impact on child adjustment, the therapist must be a conveyor of scientific data and developmental theory. It is also useful to relate the data to the specific and current needs to each of the couple's children. It is not the purpose of this paper to review all that has been written on training packages in these areas. Instead, refer to the writing of Hendrix (1989) or the vast literature on mediation and conflict resolution strategies.

Teaching the requisite competencies as they are needed, rather than at the beginning of the therapy, may be a more efficient and effective process. In some cases, parents possess some of the required skills. Teaching what they already know adds substantial time to what is a very time consuming therapy. Ameliorating deficiencies as they appear increases parents' motivation to acquire the skills. They realize that without the specific capacities, they will be unable to reach the goals that either or both hold. Once requisite proficiencies have been taught and several problems have been resolved, the parents should be encouraged to initiate the resolution of some new problem with limited direction from the therapist.

Three procedural points must be considered. First, many parents engaged in high conflict divorces do not acquire the level of skills, respect and motivation necessary to accomplish independent problem solving. Continued assistance from the therapist over time is generally required. It is important to note, however, that substantial reductions in hostility and increases in cooperative coparenting are achieved by helping parents to resolve their most significant disputes. Second, not all disputes are either resolvable or have to be resolved. Some issues are too difficult or too subtle to resolve, others simply do not matter to effective coparenting. It is necessary for the therapist to categorize issues along the dimensions of solvable and relevant so that sufficient attention can be given to reconcilable and relevant differences. Third, strict limits must be set by the therapist on any excessive expressions of anger or of efforts by one parent to control the other. Both parents must feel safe to participate in this process. If either feels threatened or controlled by the other, progress cannot be made. Strategies to insure a sense of safety include defining rules and consequences at the beginning of the therapeutic process. Short breaks during a session or, in some cases, early termination of a session may be required.

Two disputes common to most divorces involving children are child support and flexibility in predetermined schedules. In many cases, parent and professionals involved with the family suggest unnecessarily complicated strategies to resolve these disputes. Child support is clearly defined by most state legislatures. One simple and fair strategy includes reviewing the most recent federal tax returns and using income statements on the returns for support calculations. Support can be recalculated annually, after federal taxes have been filed, using figures listed on the tax returns. This routine review of support insures that relevant figures are being used in the determination of support and that any inequity resulting from a change in the financial status of either parent lasts no more than one year.

Too much flexibility in a previously defined schedule frequently results in unnecessary anger in the parents and confusion in the children. Parents frequently feel used and like "babysitters" to the other parent. Children cannot predict where they will reside. The most straightforward solution to such problems is one in which requests for changes are strictly limited to one, 24-hour alteration every three months. The priority of parenting commitments over work, social, and professional schedules is essential to the children's feeling of self-worth and to the other parent's feelings of respect.

9.17.6.7 Evaluating and Monitoring Therapeutic Gains

As is the case with all structured interventions, evaluation of therapeutic effectiveness is necessary. Such evaluations should include the acquisition of specific and required skills, the immediate resolution of problematic and pathogenic processes, the current functioning of the children, and the long-term maintenance of skill and adjustive gains. Implementing a practical evaluation of therapeutic effectiveness should include some appraisal of parents' knowledge of the divorce adjustment process, particularly as it relates to their children. It should also include an assessment of the parents' skills in communicating, compromising, and parenting. Changes in observed interparent hostility and the ability to resolve specific disputes follow. Measures of child adjustment, such as the Child Behavior Checklist, must also be collected. In an effort to demonstrate the durability of the intervention, an evaluation of hostility, dispute resolution, and child adjustment should be conducted six months after the end of the therapy.

9.17.7 CONCLUSIONS

The consequences of high hostility divorce on child adjustment is severe and long-lasting and impacts hundreds of thousands of children annually. Developing and evaluating a strategy to promote coparenting and to reduce hostility can substantially reduce the severe consequences of such dysfunctional parenting. A preliminary strategy to accomplish such goals has been suggested in the manuscript. Substantial work is still required in the development of such an intervention. First, more information is needed in understanding which specific coparenting behaviors exacerbate interparent hostility and deteriorate children's adjustment and which behaviors promote cooperation and children's mental health. Second, a formal evaluation of the existing strategy is required. Third, the experience of therapists experienced in working with hostile, divorcing parents must be accumulated to identify problems with the existing strategy and to suggest new procedures to ameliorate them. Finally, training at the graduate and professional level will be required so that a sufficient pool of competent therapists exists to meet the substantial demands of this emotionally taxing population.

9.17.8 REFERENCES

Ainsworth, M. D. S. (1979) Infant–mother attachment. *American Psychologist, 34,* 932–937.

Ainsworth, M. D. S. (1989). Attachments beyond infancy. *American Psychologist, 44,* 709–716.

Bandura, A. (1973). *Aggression: A social learning approach.* Englewood Cliffs, NJ: Prentice-Hall.

Belsky, J. (1980). Child maltreatment: An ecological perspective. *American Psychologist, 35,* 320–335.

Belsky, J. (1984). The determinants of parenting: A process model. *Child Development, 55,* 83–96.

Biller, H. B. (1981). The father and sex role development. In M. E. Lamb (Ed.). *The role of the father in child development* (2nd ed. pp. 319–358). New York: Wiley.

Bjorklund, D. F., & Green, B. L. (1992). The adaptive nature of cognitive immaturity. *American Psychologist, 47,* 46–54.

Bloom, B. L., Asher, S. J., & White, S. W. (1978). Marital disruption as a stressor: A review and analysis. *Psychological Bulletin, 85,* 867–894.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32,* 513–531.

Bronfenbrenner, U. (1979). *The ecology of human development.* Cambridge, MA: Harvard University Press.

Camplair, C. W., & Stolberg, A. L. (1987). Professionals' influences on child custody decisions. *Journal of Divorce, 10,* 43–56.

Camplair, C. W., & Stolberg, A. L. (1990). Court sponsored divorce mediation: A description and evaluation of process and outcome. *Mediation Quarterly, 7,* 199–213.

Christenson, A., & Margolin, G. (1988). Conflict and alliance in distressed and non-distressed families. In R. A. Hinde, & J. Stevenson-Hinde (Eds.), *Relationships within families: Mutual influence* (pp. 333–360). Oxford, UK: Oxford University Press.

Coogler, O. J. (1978) *Structural mediation in divorce settlement: A handbook for marital mediations.* Lexington, MA: Lexington Books.

Cummings, E. M., Iannotti, J. R., & Zahn-Waxler, C. (1985). The influence of conflict between adults on the emotions and aggressions of young children. *Developmental Psychology, 21,* 495–507.

Cummings, E. M., Vogel, D., Cummings, J. S., & El-Sheikh, M. (1989). Children's responses to different forms of expression of anger between adults. *Child Development, 60,* 1392–1404.

Dadds, M. R., Sanders, M. R., Morrison, M., & Rebgetz, M. (1992). Childhood depression and conduct disorder: II An analysis of family interaction patterns in the home. *Journal of Abnormal Psychology, 101,* 505–513.

Davis, K., & Stolberg, A. L. (1988). Influences on professionals' child custody recommendations. *Journal of Divorce, 11,* 207–217.

Deed, M. L. (1991). Court-ordered child custody evaluations: Helping or victimizing vulnerable families? *Psychotherapy, 28,* 76–84.

Ellwood, M. S., & Stolberg, A. L. (1993). The effects of family composition, family health, parenting behavior and environmental stress on children's divorce adjustment. *Journal of Child and Family Studies, 2,* 23–36.

Emery, R. (1982). Interparental conflict and the children of discord and divorce. *Psychological Bulletin, 92,* 310–330.

Emery, R. E., & O'Leary, K. D. (1984). Marital discord and child behavior problems in a nonclinic sample. *Journal of Abnormal Child Psychology, 12,* 411–420.

Emery, R. E., & Wyer, M. M. (1987). Divorce mediation. *American Psychologist, 42,* 472–480.

Flavell, J. (1985). *Cognitive development.* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Forehand, R., Long, N., & Brody, G. (1988). Divorce and marital conflict: Relationship to early adolescent competence and adjustment in early adolescence. In E. M. Hetherington & J. D. Arasteh (Eds.), *Impact of divorce, single parenting and step-parenting on children* (pp. 155–196). Hillsdale, NJ: Erlbaum.

Friedrich, W. N., & Einbender, A. J. (1983). The abused child: A psychological review. *Journal of Clinical Child Psychology, 12,* 244–256.

Fry, P. S., & Addington, J. (1985). Perceptions of parent and child adjustment in divorced families. *Clinical Psychology Review, 5,* 141–157.

Gilbert, R. K., & Christensen, A. (1988). The assessment of family alliances. Advances in *Behavioral Assessment of Children and Families, 4,* 219–252.

Girdner, L. K. (1985). Adjudication and mediation: A comparison of custody decision making processes. *Journal of Divorce, 8,* 33–48.

Grych, J. H., & Fincham, F. D. (1990). Marital conflict and children's adjustment: A cognitive-contextual framework. *Psychology Bulletin, 102,* 267–290.

Grych, J. H., & Fincham, F. D. (1993). Children's appraisals of marital conflict: Initial investigation of the cognitive-contextual framework. *Child Development, 64,* 215–230.

Hartup, W. W. (1989). Social relationships and their developmental significance. *American Psychologist, 44,* 120–126.

Hendrix, H. (1989). *Getting the love you want.* New York: HarperCollins.

Hetherington, E. M. (1979). Divorce: A child's perspective. *American Psychologist, 34,* 851–858.

Hetherington, E. M. (1991). The role if individual differences and family relationships in children's coping with divorce and remarriage. In P. A. Cowan & E. M. Hetherington (Eds.), *Family transitions* (pp. 165–194). Hillsdale, NJ: Erlbaum.

Hetherington, E. M., & Clingempeel, W. G. (1992). *Coping with marital transitions: A systems perspective.* Chicago: University of Chicago Press for the Society for Research in Adolescence.

Hetherington, E. M., Cox, M., & Cox, R. (1982). Effect of divorce on parents and children. In M. E. Lamb (Ed.), *Non-traditional families: Parenting and child development* (pp. 233–288). Hillsdale, NJ: Erlbaum.

Hinde, R. A., & Stevenson-Hinde, J. (1988). *Relationships* *within families: Mutual influences*. New York: Oxford University Press.

Johnson, P. L., & O'Leary, K. D. (1987). Parental behavior patterns and conduct disorders in girls. *Journal of Abnormal Child Psychology, 15,* 573–581.

Jouriles, E. N., Barling, J., & O'Leary, K. D. (1989). Interspousal aggression, marital discord, and child problems. *Journal of Consulting and Clinical Psychology, 57,* 453–455.

Kalter, N. (1987). Long-term effects of divorce on children: A developmental vulnerability model. *American Journal of Orthopsychiatry, 55,* 587–600.

Kalter, N. (1989). *Growing up with divorce: Helping your child avoid immediate and later emotional problems*. New York: Free Press.

Kalter, N., Pickar, J., & Lesowitz, M. (1984) School-based developmental facilitation groups for children of divorce: A preventive intervention. *American Journal of Orthopsychiatry, 54,* 613–623.

Katz, L. F., & Gottman, J. M. (1993). Patterns of marital conflict predict children's internalizing and externalizing behaviors. *Developmental Psychology, 29,* 940–950.

Kerig, P. A., Cowan, P. A., & Cowan, C. P. (1993). Marital quality and gender differences in parent–child interactions. *Developmental Psychology, 29,* 931–939.

Koestner, R., Franz, C., & Weinberger, J. (1990). The family origins of empathic concern: A 26-year longitudinal study. *Journal of Personality and Social Psychology, 58,* 709–717.

Kerr R., & Bowen, R. A. (1988). *Family evaluation*. New York: W. W. Norton.

Kressel, K., & Pruitt, D. G. (1985). Themes in the mediation of social conflict. *Journal of Social Issues, 41,* 179–198.

Kressel, K., & Deutsch, M. (1977). Divorce therapy: An in-depth survey of therapists' views. *Family Processes, 16,* 413–443.

Kurdek, L. A. (1981) An integrative perspective on children's divorce adjustment. *American Psychologist, 36,* 856–866.

Mann, B. J., Borduin, C. M., Henggeler, S. W., & Blaske, D. M. (1990). An investigation of systemic conceptualizations of parent–child coalitions and symptom change. *Journal of Consulting and Clinical Psychology, 58,* 336–344.

McIsaac, H. (1982). Court connected mediation. *Conciliation Courts Review, 21,* 49–56.

Miller, P. M., Danaher, D. L., & Forbes, D. (1986). Sex-related strategies for coping with interpersonal conflict in children ages five to seven. *Developmental Psychology, 22,* 543–548.

Minuchin, S. (1974). *Families and family therapy*. Cambridge, MA: Harvard University Press.

Neville, W. G. (1985). Reflections on the growth and significance of divorce mediation. *Journal of Divorce, 8,* 3–7.

Patterson, G. R. (1977). Accelerating stimuli for two classes of coercive behaviors. *Journal of Abnormal Child Psychology, 5,* 335–350.

Patterson, G. R. (1982). *Coercive family processed: A social learning approach* (Vol. 3). Eugene, OR: Castelia Publishing Company.

Patterson, G. R. (1986). Performance models for antisocial boys. *American Psychologist, 41,* 432–444.

Patterson, G. R., & Stouthamer-Loeber, M. (1984). The correlation of family management practices and deliquency. *Child Development, 55,* 1299–1307.

Pearson, J., & Thoennes, N. (1984). Mitigating and litigating custody disputes: A longitudinal evaluation. *Family Law Quarterly, 12,* 497–454.

Porter, B., & O'Leary, K. D. (1980). Marital discord and child behavior problems. *Journal of Abnormal Child Psychology, 8,* 287–295.

Robinson, M., & Parkerson, L. (1985). A family systems approach to conciliation in separation and divorce. *Journal of Family Therapy, 7,* 357–377.

Sarason, I. G., Sarason, B. R., & Pierce, G. R. (1990). *Social support: An interactional view*. New York: Wiley.

Stinson, K. M. (1991). *Adolescents, family and friends: Social support after parents' divorce or remarriage*. New York: Wiley.

Stolberg, A. L., & Bush, J. (1985). A path analysis of factors predicting children's divorce adjustment. *Journal of Clinical Child Psychology, 14,* 49–54.

Stolberg, A. L., Camplair, C., Currier, K., & Wells, M. J. (1987). Individual, familial and environmental determinants of children's post-divorce adjustment and maladjustment. *Journal of Divorce, 11,* 51–70.

Stoneman, Z., Brody, G. H., & Burke, M. (1989). Marital quality, depression, and inconsistent parenting: Relationship with observed mother–child conflict. *American Journal of Orthopsychiatry, 59,* 105–117.

Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Brunner/Mazel.

Trombatta, D. (1982). Custody evaluation and custody mediation: A comparison of two dispute interventions. *Journal of Divorce, 6,* 65–76.

Turner, P. J., Gervai, J., & Hinde, R. A. (1993). Gender typing in young children: Preferences, behavioral and cultural differences. *British Journal of Developmental Psychology, 11,* 323–342.

Vaux, A. (1988). *Social support: Theory, research and intervention*. New York: Prager.

Vuchinich, S., Bank, L., Patterson, G. R. (1992). Parenting, peers, and stability of antisocial behavior in preadolescent boys. *Developmental Psychology, 28,* 510–554.

Vuchinich, S., Emery, R. E., & Cassidy, J. (1988). Family members as third parties in dyadic family conflict: Strategies, alliances, and outcomes. *Child Development, 59,* 1293–1302.

Vuchinich, S., Vuchinich, R., & Wood, B. (1993). The interparental relationship and family problem solving with preadolescent males. *Child Development, 64,* 389–400.

Wallerstein, J. S., & Blakasee, S. (1989). *Second chances: Men, women and children a decade after divorce*. New York: Tichnor and Fields.

Wallerstein, J. S., & Kelly, J. B. (1980). *Surviving the breakup: How children and parents cope with divorce*. New York: Basic Books.

Wierson, M., Forehand, R., & McCombs, A. (1988). The relationship of early adolescent functioning to parent-reported and adolescent-perceived interparental conflict. *Journal of Abnormal Child Psychology, 16,* 707–718.

Weitzman, L. J. (1985). *The divorce revolution: The unexpected consequences for women and children in America*. New York: Free Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.18
Adult Victims of Child Sexual Abuse

ROGER C. KATZ
University of the Pacific, Stockton, CA, USA
and
PATTI LOU WATKINS
Oregon State University, Corvallis, OR, USA

9.18.1 INTRODUCTION 291

9.18.2 DEFINING CHILD SEXUAL ABUSE 292

9.18.3 EXTENT OF THE PROBLEM 293
  9.18.3.1 Incidence and Prevalence 293
  9.18.3.2 Victim and Perpetrator Characteristics 293
  9.18.3.3 Risk Factors for Child Sexual Abuse 294

9.18.4 SEQUELAE OF CHILD SEXUAL ABUSE 294
  9.18.4.1 Post-traumatic Stress Disorder 295
  9.18.4.2 Problems of Male Survivors 296
  9.18.4.3 Medical Concomitants 296
  9.18.4.4 Positive Outcomes 297

9.18.5 TREATING ADULT VICTIMS 297
  9.18.5.1 Relationship Issues 298
    9.18.5.1.1 Barriers to treatment 298
    9.18.5.1.2 Iatrogenic complications 298
  9.18.5.2 Treatment Models 299
    9.18.5.2.1 Post-traumatic stress disorder model 299
    9.18.5.2.2 Four traumagenic dynamics model 299
    9.18.5.2.3 Eclectic models 299
    9.18.5.2.4 Cognitive-behavioral treatments 300
  9.18.5.3 Methodological Considerations 300

9.18.6 PREVENTION 301

9.18.7 FUTURE DIRECTIONS 303

9.18.8 REFERENCES 304

9.18.1 INTRODUCTION

Child sexual abuse (CSA) is one of the most common forms of child victimization in the USA. Until the 1980s, the sexual abuse of children was regarded as a relatively rare event that was often neglected or trivialized by mental health professionals as a construction of the victim's Oedipal fantasies. It is now recognized as a real and all-too-frequent occurrence—one that researchers believe affects about one out of every four girls in the USA and one out of every ten boys (Walker, Bonner, & Kaufman, 1988). Although the exact number of victims will never be known, the consequences of CSA are far reaching, harmful, and long lasting.

The focus of this chapter is on adult victims of child sexual abuse. By adult victims, we mean individuals over the age of 18 who, at a younger age, were forced to engage in sexual behavior by another person or engaged in sexual acts with someone much older than them even if force was not used (Browne & Finkelhor, 1986). The person or perpetrator could be a family member, relative, neighbor, or stranger. The act could be any type of sexual behavior ranging from relatively unobtrusive sexual touching and seductiveness to violent acts such as rape and ritualized sexual abuse (Lloyd, 1992).

It is difficult to talk about the experience of adult victims of child sexual abuse without addressing the issue of sexual abuse at a more general level. Consequently, this chapter begins with a discussion of definitions of CSA because this has important implications for other aspects of the problem, including prevalence rates and the type of psychological sequelae that are observed in survivors. Following this, the issues of incidence and prevalence, victim and perpetrator characteristics are addressed, as well as contextual factors that researchers believe are associated with an increased risk of CSA. Next, the psychological sequelae of CSA are described. This includes a description of methods used to assess them and factors that seem to be associated with increased harm and resilience. The chapter concludes with a description of psychological interventions that have been developed to treat adult victims of CSA, their theoretical underpinnings, and accompanying empirical support. In this final section, attention is also directed to the need for increased methodological rigor in assessing therapy outcome, and to primary and secondary prevention.

9.18.2 DEFINING CHILD SEXUAL ABUSE

One of the most basic yet controversial issues in the CSA literature is deciding what child sexual abuse is. There are many ways of defining the construct but none of them is universally accepted. This has profound implications for understanding the entire field of CSA because epidemiological and clinical research findings are only generalizable to the extent that researchers agree on the nature of their subject matter. Definitions of CSA are not the same as sexual abuse in adulthood (Wyatt, Guthrie, & Notgrass, 1992) and legal definitions of sexual abuse are not the same as those used by researchers (Haugaard & Reppucci, 1988). Disagreement about what CSA is exists for many reasons. To begin with, sexual abuse can be subtle, sadistic, or overt and involve a myriad of behaviors (Lawson, 1993). It occurs between people of different ages (children and teenagers, teenagers and adults, adults and children) and in different relationships (family member, acquaintance, stranger). It can be a single incident or a repeated occurrence, and it may occur in conjunction with physical and emotional abuse. Moreover, its effects are varied and difficult to disentangle from the sequelae of other types of trauma or family dysfunction (Green, 1993; Polusny & Follette, 1995).

Definitions of sexual abuse can be broad, such as forced sexual contact during the person's lifetime, or narrow, such as a particular type of molestation (e.g., father–daughter incest) (Burnam et al., 1988). Some definitions emphasize the exploitative intention of the perpetrator compared with other definitions which focus on the effects of the perpetrator's actions on the survivor. In their seminal review, Browne and Finkelhor (1986) point out that CSA involves two overlapping but distinct types of interaction: (i) coerced sexual activity that is forced on a child, and (ii) sexual activity between a child and a much older person, whether or not coercion is involved. While many researchers have adopted these guidelines in operationalizing CSA, differences of opinion exist about what constitutes a child (e.g., age <14, 16, or 18) and the age criterion for a person much older (Knutson, 1995). For many investigators, the criterion is at least five years older (Browne & Finkelhor, 1986).

How CSA is defined has important consequences at different levels of analysis. To begin with, criteria used to define CSA affects how people perceive it, how they react to the individuals involved, and the type of sexual behaviors that are reported to the police and child welfare workers (Finkelhor, 1984). The latter bears directly on incidence and prevalence rates which provide an estimate of the severity of the problem at a societal level. Because these rates can vary tremendously depending on how CSA is defined, this impacts the decisions of policy makers who allocate resources for treatment, prevention, and research. At a less appreciated level, definitions of CSA have important implications for the identification and severity of abuse-related psychological problems (Briere, 1992). For example, researchers who restrict themselves to earlier and more intrusive forms of abuse might report more severe sequelae in survivors compared with researchers who rely on a broader definition.

One reason is that the effects of CSA may be mediated by developmental factors (Cole & Putnam, 1992; Downs, 1993). Haugaard and Reppucci (1988) recommended that terms like child sexual abuse are too broad to convey useful information and should be avoided in favor of more specific descriptions of the type of abuse being studied (e.g., children molested by their older siblings) or the type of victims being treated (e.g., adolescent boys molested by their mothers). Their point is that any effects of a specific type of abuse or of abuse on a specific type of individual might go undetected in studies using broad definitions of CSA that result in heterogeneous samples.

It is helpful to distinguish between child molester, child rapist, and pedophile because these terms are sometimes used interchangeably when they have different meanings. A child molester is a person, usually an adult, who entices, entraps, or pressures a child into sexual activity (Walker et al., 1988). A child rapist, on the other hand, is a particular type of molester who uses force or assault in perpetrating their crime which, by definition, involves sexual intercourse. The term pedophile refers to an individual who suffers from pedophilia, a specific mental disorder defined in the *DSM-IV* (American Psychiatric Association, 1994).

9.18.3 EXTENT OF THE PROBLEM

9.18.3.1 Incidence and Prevalence

The 1980s and 1990s have ushered in a number of studies on the incidence and prevalence of CSA in the USA and abroad. As a rule, these studies have yielded discrepant findings (Wynkoop, Capps, & Priest, 1995). Some of the discrepancy is due to sampling error but differences in methodology are equally important. Relevant variables include (i) the group studied, (ii) the type of sampling procedures used, (iii) the rate of participation among possible subjects, (iv) the criteria used to define a sexual experience as abusive, (v) the upper age limit of the victim, and (vi) the methods used to collect the data (Anderson, Martin, Mullen, Romans, & Herbison, 1993; Cappelleri, Eckenrode, & Powers, 1993). Discrepancies in the data notwithstanding, statistics on CSA are widely assumed to underestimate the true extent of the problem for the simple reason that a significant percentage of cases go unreported (Urquiza & Keating, 1990).

Prevalence rates determined by random samples of populations in the USA and abroad shed further light on the problem. For example, Russell (1983) found that 38% of 930 women, age 18 and older and living in the San Francisco area, had an unwanted sexual experience before they were 18. In 28% of the cases, the abuse occurred before the victim was 14. Wyatt (1985) estimated that 59% of her sample of African-American and Caucasian women, age 18 to 36, reported some form of sexual abuse before they were 18. Rates were higher for whites than for African-Americans. Moreover, the prevalence increased to 62% when the definition of sexual abuse was broadened to include unwanted sexual experiences with peers. Somewhat lower rates of abuse (19–22%) were reported by Finkelhor (1979) and Fromouth (1986) in their work with female college students. On the other hand, in a study of women from New Zealand, Anderson et al. (1993) reported that 32% of their sample ($N=3000$) had one or more unwanted sexual experiences before they were 16, and 70% of the abuse involved genital contact or severe abuse, although only 7% of the cases were ever officially reported.

9.18.3.2 Victim and Perpetrator Characteristics

The overwhelming majority of CSA victims are girls who are molested between the ages of 8 and 12 (Anderson et al., 1993; Finkelhor, 1984). The age of maximum risk is between 10 and 11; however, this does not mean that younger children and adolescents are immune to sexual mistreatment. While sexual abuse in boys is more likely to be under-reported (Urquiza & Keating, 1990), for every boy who is molested there are about 2.5 female victims (Finkelhor, 1986). Similar findings were reported by Cappelleri et al. (1993) from the Second National Incidence and Prevalence Study of Child Abuse and Neglect. They found a sexual abuse incidence rate (per 1000) of 3.28 for girls compared with 1.0 for boys. Epidemiological data from other countries show that women are abused 1.5 to 3 times more often than men and that prevalence rates are comparable to those in the USA and Canada (Finkelhor, 1994).

The experiences of victims of CSA vary depending on the group being studied (Green, 1993; Knutson, 1995). For this reason, the origin of the sample needs to be considered when addressing issues like victim characteristics and short- and long-term effects of the abuse. While large-scale surveys have called attention to the prevalence of CSA, information derived from these surveys is not the same as results from clinical samples. Compared to large-scale surveys of college students, for example, groups of previously identified victims who are either in therapy or requesting it report more severe and protracted forms of sexual abuse and a higher incidence of actual or attempted intercourse (Haugaard & Reppucci, 1988). For both male and female victims, there is also a higher percentage of fathers as perpetrators. This has important implications because abuse by father figures or abuse involving force and penetration may be associated with increased psychological harm (Kendall-Tackett, Williams, & Finkelhor, 1993; Mennen, 1993).

Perpetrators of sexual abuse against girls are almost always male. Most of them are in their mid-30s and acquainted with their victim (Anderson et al., 1993; Russell & Trainor, 1984). They include family members, neighbors, and people in positions of trust or authority, such as teachers and baby-sitters. With female perpetrators, the abuse tends to occur within a caretaker role; moreover, the victim is more likely to be younger and, often, a male (Rudin, Zalewski, & Boomer-Turner, 1995). Between 15% and 24% of female victims are molested by strangers (Anderson et al., 1993). Among males, the figure may be a little higher (Finkelhor, Hotaling, Lewis, & Smith, 1990) except when the abuser is a female (Rudin et al., 1995). The most common situation is one in which the child is molested only once, usually fondling by a family member. Sexual intercourse is more common in extrafamilial abuse (Russell, 1983). Finklehor (1979) found that having a stepfather more than doubled a young girl's vulnerability to molestation, and Anderson et al. (1993) reported that stepfathers were 10 times more likely to be abusers than biological fathers.

Like victims of CSA, perpetrators are a heterogeneous group. They include men and women in different role relationships with the victim, young and old, and sometimes with a propensity for many types of deviant sexual behavior (Abel & Rouleau, 1990; Lawson, 1993). Margolin (1992) has noted that a vast majority of sexually abusive grandparents are grandfathers who were sexually abusive as fathers and chose their granddaughters as a victim. Stepgrandchildren tended to be at greater risk. In their study of adult male child molesters and other sex offenders, Abel and Rouleau (1990) found that most of the men had many different paraphilias. This was especially true of incest offenders who molested male children.

Reviewing the literature on child and adolescent perpetrators, Vizard, Monck, and Misch (1995) pointed out that at least 20% of the sexual offenses in the USA are committed by young men under the age of 18, many of whom grow up to become adult offenders. A large number of these youthful perpetrators were physically or sexually abused themselves, and their victims are often younger siblings. Although their offending behavior is sexual, the motivation behind it may have more in common with dominance or bullying behaviors. For example, Katz (1990) reported that social skills deficits and peer isolation predispose some adolescents to commit sexual crimes against younger children.

9.18.3.3 Risk Factors for Child Sexual Abuse

Risk factors for CSA often say more about the victim's environment than the victim *per se* (Knutson, 1995). The most common victim characteristics are gender (female), age (older), and social isolation (Cappelleri et al., 1993). However, seven of the eight risk factors described by Finkelhor (1984) were familial (e.g., had a stepfather, separated from mother, sexually punitive mother, no physical affection from father), as were the major risk factors identified by Benedict and Zautra (1993) —parental absence and conflict. Low family income (Bergner, Delgado, & Graybill, 1994; Cappelleri et al., 1993), family history of alcoholism (Windle, Windle, Scheidt, & Miller, 1995), and having a disability that requires residential placement are also associated with an increased risk of abuse (Sullivan & Knutson, 1993).

9.18.4 SEQUELAE OF CHILD SEXUAL ABUSE

A decade after Browne and Finkelhor (1986) first reviewed the long-term effects of CSA, research continues to show that its sequelae are highly diverse. While much of the research in this area has been atheoretical, several reviewers have attempted to place their findings within a theoretical framework. For instance, Cole and Putnam (1992) discussed the sequelae of father–daughter incest within a developmental framework. They contended that adults victimized as preschoolers may experience psychopathology involving denial and dissociation which are common coping strategies during this developmental phase. Indeed, dissociation in clinical or subclinical forms is among the established concomitants of CSA (Polusny & Follette, 1995). Focusing on abusive parent–child relationships, Alexander (1992) discussed CSA sequelae in the context of Bowlby's (1973) attachment theory. While these attempts are somewhat limited in that they focus primarily on incestuous father–daughter relationships, they are to be lauded for their efforts to explain CSA sequelae in theoretical terms. Perhaps because of such efforts, researchers have begun to test theoretical models linking CSA to particular symptom presentations. For example, Mallinckrodt, McCreary, and Robertson (1995) provided empirical support for an association between CSA and eating disorders that may be mediated by impaired maternal attachment and lack of social competencies, such as interoceptive self-awareness and the ability to solicit social support.

Polusny and Follette (1995) have suggested that many of the sequelae of CSA transcend traditional diagnostic categories. They conceptualized CSA-related problems as manifestations of emotional avoidance, a theory that is not incompatible with the developmental and attachment-based perspectives previously described. They suggested that CSA survivors use coping strategies, such as substance abuse and self-mutilation, that are negatively reinforced by the reduction of painful affect associated with the abuse experience and related stimuli. Although such strategies may enhance a child's ability to cope, excessive reliance on them in adulthood can compromise adjustment.

Polusny and Follette's (1995) review also revealed that adult survivors exhibit higher levels of general psychological distress, as well as more specific problems including major depression, anxiety disorders, substance abuse, eating disorders, dissociative disorders, somatization, and personality disorders. A history of CSA was associated with increased dissatisfaction with partner and parenting roles, less sexual satisfaction, high-risk sexual practices, more sexual dysfunction, and an increased risk of sexual revictimization in adulthood.

The remainder of this section discusses the current status of post-traumatic stress disorder (PTSD) (American Psychiatric Association, 1994) as a diagnosis for the problems experienced by CSA survivors, then focuses on the relationship of CSA to mental health problems experienced by male survivors because, compared to female survivors, this is a relatively neglected topic. Next, some of the medical concomitants of CSA are examined, and from there a brief discussion of resiliency factors and potential adaptive or beneficial effects of being a CSA survivor.

9.18.4.1 Post-traumatic Stress Disorder

Rowan and Foy (1993) addressed the applicability of PTSD as a diagnosis for the constellation of symptoms displayed by CSA survivors. Historically, PTSD was first used to describe the psychological problems of Vietnam combat veterans. Later it was applied to victims of rape and domestic violence, as well as adult survivors of CSA who were experiencing psychological difficulties. Indeed, as Rowan and Foy contend, other DSM disorders "fail to encompass the full range of primary symptoms frequently experienced by CSA survivors" (p. 4). A PTSD diagnosis includes criteria which characterize impairments in both self (e.g., psychological distress) and social functioning (e.g., feelings of detachment or estrangement from others) that are commonly observed among CSA survivors. Rowan and Foy reviewed research studies which examined the prevalence of PTSD among CSA survivors as well as those which examined the relationship of assorted exposure variables to PTSD. While they concluded "that the PTSD model appears to describe the core features of the psychological difficulties of many survivors" (p. 17), they tempered this statement with cautions concerning the methodologies employed. In particular, they noted a lack of clarity and/or variability in definitions of CSA, along with an absence of standardized measures of PTSD.

Rowan, Foy, Rodriguez, and Ryan (1994) conducted a study in which they operationalized CSA exposure as well as employed standardized measures, both interview and questionnaire, of PTSD. Results revealed that between 64% and 69% of their clinical sample of CSA survivors met full criteria for PTSD, with another 19% meeting partial criteria. Furthermore, duration of abuse was significantly related to PTSD, as was frequency according to interview-based diagnoses.

McNew and Abell (1995) compared CSA survivors and Vietnam veterans on self-report measures of PTSD symptomatology, anger, social intimacy, and anxiety. Both groups experienced moderately severe levels of the PTSD symptoms, intrusion and avoidance, along with high levels of anger and anxiety. Furthermore, both groups evidenced patterns of anger expression likely to interfere with optimal functioning and an impaired sense of self in relation to others. The veterans obtained significantly higher intrusion scores. Otherwise, the two groups did not differ in symptom severity which led the researchers to conclude that relational disturbances and anxiety may manifest themselves regardless of the type of trauma experienced. McNew and Abell concluded that CSA survivors are similar to Vietnam veterans in their experience of PTSD symptoms, emphasizing that anger might best be considered as a central, rather than an associated, feature of PTSD.

Newman and Peterson (1996) investigated the relatively neglected experience of anger in CSA survivors. They found that female incest survivors were angrier (in general, and toward both parents in particular) than a control group of women without a history of CSA. Rather than view these findings in pathological terms, Newman and Peterson suggested that it may be "legitimate and perhaps beneficial for incest survivors to be angry about specific individuals and occurrences" (p. 471). They cautioned, however, that anger experienced by CSA survivors might be counterproductive if it is expressed in ways that undermine current relationships.

9.18.4.2 Problems of Male Survivors

Compared with female survivors of CSA, male survivors have been neglected in the research literature. This neglect has extended to opportunities for support and treatment (Lisak, 1994). Jumper (1995) conducted a meta-analysis examining the relationship of CSA to adult psychological adjustment based on research published between 1978 and 1991. While he concluded that male and female CSA survivors are similar in terms of psychological symptomatology, only four of the 23 studies analyzed included male samples. Further work has seen a much-needed increase in studies aimed at understanding the psychological sequelae of CSA for men.

In Lisak's (1994) study, 26 adult male survivors completed a psychological symptom inventory which revealed elevated global severity scores and elevated PTSD-subscale scores relative to normative samples. Providing even more insight into the plight of men who were molested as children were the participants' responses to audio-taped autobiographical interviews. These were transcribed verbatim, then subjected to a content analysis which identified 15 different themes. Fear, commonly involving intrusions, was the one theme identified by all members of the sample. Lisak noted that "often the words and expressions used by these men to describe their feelings and experiences are identical to those used by women survivors" (p. 546). Indeed, six of the themes (i.e., helplessness, isolation and alienation, legitimacy, loss, self-blame/guilt, and shame/humiliation) were originally derived from an interview-based study of women rape survivors. While men were similar to women in terms of quantitative and qualitative measures of anxiety, their distress was manifested in some gender-specific ways. Among them were struggles with masculinity and fear of homosexuality. Another way in which sexually abused men and women often differ concerns the expression of anger or efforts to regain a sense of control over their lives. We previously noted that female CSA survivors experience anger, as did 90% of Lisak's male sample. While men often attempted to suppress angry feelings, many ultimately snapped which culminated in violent behavior. In fact, nearly one-third of Lisak's sample admitted victimizing others through physical assault of adult men as well as female partners, rape of adult women, and sexual abuse of children.

This last form of victimization has been the primary focus of other studies investigating CSA sequelae in males. Freund and Kuban (1994) tested the abused-to-abuser theory that CSA predisposes men toward pedophilia. Their sample was somewhat restricted in that it included only heterosexual pedophiles and excluded cases of incest. Furthermore, the authors questioned the accuracy of participants' self-reports of CSA, given the criminal nature of their sample. Nevertheless, pedophilia was significantly associated with CSA, while sex-offender status, in general, was not. Bagley, Wood, and Young (1994) took a different approach to examining the abused-to-abuser theory. They conducted a community survey of male CSA survivors which paralleled an earlier study investigating the psychological sequelae of CSA among women (Bagley, 1991). Specifically, this study included 750 men between the ages of 18 and 27 years, 65 of whom reported a single instance of CSA and 52 who reported multiple or ongoing episodes. As in previous studies of women, male survivors of long-term CSA experienced more severe symptomatology than survivors of short-term CSA or nonabused controls. Moreover, men with long CSA histories, accompanied by histories of emotional abuse, were the most likely candidates to engage in pedophilic behavior. As in the case of female CSA survivors, Bagley et al. speculated that attachment problems may underlie the psychological sequelae of CSA.

9.18.4.3 Medical Concomitants

Although researchers have focused on the psychological and interpersonal sequelae of CSA, they have also turned their attention to physical consequences of this experience. For example, Lechner, Vogel, Garcia-Shelton, Leichter, and Steibel (1993) assessed the self-perceived medical problems of over 500 women visiting a family practice clinic, 26% of whom reported a history of CSA. Apart from indicating they had experienced a significantly greater number of sexually-related medical problems (e.g., surgical evaluation of pelvic pain), the CSA survivors reported more medical complaints in general, including respiratory and gastrointestinal distress. Based on interviews of female college students with CSA histories, Hulme and Grove (1994) found that six physical symptoms, including insomnia, sexual dysfunction, overeating, drug abuse, severe headache, and two or more major surgeries, occurred in over half of their sample. In addition, over one-third of their sample reported undergoing a single hospitalization, and a similar proportion reported experiencing ulcers. Obviously, both of these studies can be criticized on the basis of failing to provide objective verification of medical complaints. Nevertheless, these data suggest that CSA survivors who experience physical symptoms, whether they represent a manifestation of a somataform disorder or an organically-based malady, may well present to medical, rather than mental health, facilities.

Hulme and Grove (1994) contend that health professionals in all clinical areas must be alert to symptom presentations suggestive of CSA. Lechner et al. (1993) placed the burden for detecting such a history on the physician rather than the patient because only 5.1% of their sample of CSA survivors disclosed the experience to their doctors. They believe that physician identification of CSA sequelae might reduce unnecessary repeat visits to primary care settings and facilitate entry into appropriate mental health treatment. In their opinion, the high incidence of sexual abuse warrants this approach.

9.18.4.4 Positive Outcomes

While most researchers have focused on the adverse effects of CSA, some have turned their attention to the positive attributes of survivors. These include resiliency factors—coping styles that seem to contribute to healthy adult functioning—as well as positive consequences of the abuse experience itself. For example, Valentine and Feinauer (1993) recruited 22 adult women who defined themselves as well-functioning, despite their CSA histories, to participate in an interview aimed at eliciting factors important to their survival. The authors substantiated participants' self-perceptions of adjustment, indicating that all were currently employed and/or living without benefit of government assistance and that none had been institutionalized in hospitals, prisons, or shelters for homeless or battered women. From these interviews, six resiliency themes emerged, including the ability to build supportive relationships outside of the abusive environment in settings such as school and church. Furthermore, these contacts helped to develop participants' self-regard, allowing them to embark on careers and other positive life paths. In addition to providing sources of social support, religious settings appeared to foster a sense of purpose in the participants' lives, while alleviating misplaced guilt surrounding their abuse.

Valentine and Feinauer (1993) observed several cognitive coping strategies in their participants, including an internal locus of control and a philosophy of life in which the abuse was reframed as a learning experience. The participants were also able to externalize blame for the abuse, recognize that "it won't always be this way," and understand that the abuse "did not have to affect everything they did" (p. 221). In essence, these well-functioning survivors employed an optimistic explanatory style characterized by external, unstable, and specific causal beliefs about negative events. Peterson and Bossio (1991) have ascertained that such thinking is related to adaptive functioning among various populations other than CSA survivors. Similarly, Feinauer and Stuart (1996) found that blaming oneself for CSA was an impediment to recovery and associated with increased symptomatology.

Morrow and Smith (1995) used a similar approach to gather qualitative information from a clinic sample of 11 women in order to learn how they coped with their abuse. Some common themes emerged. For instance, many participants reframed their abuse to minimize its emotional impact and create an illusion of control or power. They also developed other areas of their life and sought out the consolation of others.

McMillen, Zuravin, and Rideout (1995) interviewed 222 low-income mothers who were CSA survivors and related their responses to self-report measures of adjustment. Nearly one-half of the sample reported some benefits from their abuse which tended to fall into one of four categories: (i) increased ability to protect one's children, (ii) better protection of self, (iii) greater strength and self-sufficiency, and (iv) improved understanding of CSA, along with enhanced empathy for other survivors. Perceived benefit was positively related to self-esteem and comfort depending on others, while negatively related to interpersonal anxiety. The findings of McMillen and his coworkers have implications for treating adult survivors, especially in terms of how the abuse experience is processed intellectually and emotionally.

9.18.5 TREATING ADULT VICTIMS

The long-term effects of CSA vary from person to person and are moderated by different factors. Some of these factors are abuse related, such as its frequency and duration (Kendall-Tackett et al., 1993), and others involve the developmental stage when the abuse occurred (Cole & Putnam, 1992; Downs, 1993). Still others comprise a dynamic interaction between the person, the event, and the surrounding milieu—the "ecology of the individual's experience" (Lebowitz, Harvey, & Herman, 1993, p.379). What adult victims of CSA share is a common event, not a specific syndrome. Many victims grow up relatively unscathed (Finkelhor, 1990), while some reported they even benefited from the experience (McMillen et al., 1995). Differential effects of CSA depend on many factors. Because the long-term effects of CSA are nonspecific and varied, it is not surprising that many different techniques have evolved to treat them. What remains to be determined is how well these treatments work because outcome studies with appropriate control groups and long-term follow-ups are nonexistent in this area (Beutler & Hill, 1992; Briere, 1992).

Treatment is designed to help victims cope with the long-term effects of their abuse. Improved coping means lessening the destructive impact of the experience and any subsequent trauma that may have resulted from sexual revictimization (Wyatt et al., 1992). Beyond this, the specific goals of treatment depend on the theoretical orientation of the clinician and the identified needs of the patient. In planning treatment, one of the first tasks is to evaluate the severity of the individual's symptoms. Contextual variables and host characteristics (e.g., personality attributes, coping skills, co-morbid psychiatric conditions) often interact with the abuse experience to buffer or magnify its impact (Meiselman, 1994). From an ecological perspective, many variables affect the course of treatment which underscores the importance of a careful assessment and the realization that multiple treatment modalities may be needed. An important but neglected topic in this regard is treatment matching—that is whether better outcomes are achieved when survivors are matched with an appropriate intervention based on their unique features and identified needs (Lebowitz et al., 1993). Logic says this would be true but the requisite studies have not been done.

9.18.5.1 Relationship Issues

An important prelude to treatment is rapport building and developing a therapeutic relationship. This is especially important in treating survivors because, during the course of treatment, they can expect to partially reexperience memories and affects that were too painful to be processed normally when their abuse occurred (Meiselman, 1994). In order for patients to confront their trauma, they must first develop trust in their therapist.

9.18.5.1.1 Barriers to treatment

Before a therapeutic relationship is established, barriers may need to be overcome that prevent victims from seeking help. Walker (1994) has described four of these barriers: (i) denial, minimization or complete repression of the abuse; (ii) self-blame, where victims hold themselves responsible for their ordeal; (iii) shame or guilt which can be so powerful that victims are unable to discuss the event with anyone else, and (iv) fear of punishment or retaliation from the perpetrator. Walker also described some obstacles to effective treatment, including the patient's fear of the therapist's reaction and the possibility that his or her story will not be believed. Another barrier is the fear of opening up which results when victims try to cope with their abuse by avoiding the issue or keeping their feelings walled off and under tight control. A third barrier is excessive efforts to please the therapist—a coping strategy some victims learned to reduce the amount of abuse they received. When this generalizes to the therapist's office, it generates the facade of being helped but in reality it only interferes with the real business of treatment. According to Walker, it is important for the therapist not to pressure the person seeking treatment or hold them responsible for their abuse. Either of these can discourage the victim from seeking further help.

9.18.5.1.2 Iatrogenic complications

Adult survivors seeking treatment face the possibility of iatrogenic complications and unwanted effects. One possibility is that disclosing sexual abuse may trigger an outpouring of psychological distress brought on by painful memories of the ordeal. Directly or indirectly, this can disrupt relationships with significant others and increase vulnerability to psychiatric disorder (McNulty & Wardle, 1994).

The controversy surrounding repressed memories of sexual abuse, recovered memory therapy, and the so-called false memory syndrome is another concern (Loftus, 1993). In a recent California court case, the plaintiff was the first accused parent to sue therapists for implanting false abuse memories in his teenage daughter—a malpractice suit he won, along with half a million dollars (*Ramona vs. Isabel*, 1994). Without denying the reality of CSA, Loftus (1993) has questioned the reliability of a particular class of recovered memories that emerge in adulthood after extensive memory work with a therapist. In some cases, these memories may be subtly and unwittingly implanted by overzealous counselors who are convinced that if their patients cannot remember an abuse experience, they probably repressed it. This line of reasoning is not only destructive to the accused but potentially damaging to the patient. McCarthy (1992) raises an additional concern about treating abuse victims, especially in 12-step recovery groups. He believes some survivors get stuck in the role of the angry victim which then becomes the focal point of their existence.

9.18.5.2 Treatment Models

9.18.5.2.1 Post-traumatic stress disorder model

According to this model, abuse-related symptoms of adult survivors result from an interplay of cognitive and learning variables—faulty information processing and distorted assumptions about the abuse on the one hand, conditioned emotional reactions and learned avoidance behavior on the other (Resnick & Newton, 1992; Smucker, Dancu, Foa, & Niederee, 1995). While this model has led to important therapeutic advances, it does not apply to all victims nor does it explain the symptoms that result from many types of sexual abuse (Finkelhor, 1988). For Polusny and Follette (1995), different symptoms of CSA, like binge eating, hypersexuality, drug abuse, and dissociation, are functionally equivalent ways of avoiding emotional pain caused by the person's memories of their victimization. According to this view, gradual exposure to anxiety-eliciting stimuli is necessary for therapy to be effective.

9.18.5.2.2 Four traumagenic dynamics model

Finkelhor (Finkelhor, 1988; Finkelhor & Browne, 1985) has described an alternative to the PTSD model which seems to explain a wider range of symptoms while also providing a useful guide for developing comprehensive intervention strategies. Called the four traumagenic (or trauma-causing) dynamics model, the basic idea is that CSA has different effects depending on its impact on four important areas of children's development: sexuality, the ability to trust, self-esteem, and beliefs about self-efficacy or personal mastery. The four traumagenic dynamics corresponding to these areas are traumatic sexualization, betrayal, stigmatization, and powerlessness.

Traumatic sexualization results from sexual stimulation and reinforcement of the child's sexual responses. One of the consequences of this is that the child learns to use sexual behavior to satisfy nonsexual needs. This can lead to deviant arousal patterns, confused sexual identity, and hypersexual behavior. Betrayal occurs when the child is sexually exploited by adults who are supposed to love and protect them. This can produce generalized feelings of distrust, anger, and hostility. Stigmatization refers to the child's sense of being damaged, defective, or blamed for the molestation. This often generates guilt, shame, and low self-esteem. Powerlessness stems from feelings of helplessness during the assault and leads to problems involving fear, anxiety, and depression. Difficulties in each of these areas may require different treatments modalities.

9.18.5.2.3 Eclectic models

Walker (1994) described an eclectic approach called survivor therapy which has its origins in feminist therapy theory and trauma theory. Her emphasis is on treating female victims. The primary goal of the approach is meeting the individual's safety needs. Once this is accomplished, the therapist moves on to other objectives, including reinstilling a sense of control, validating the person's feelings by supportive and empathic means, exploring options and their associated consequences, helping the person make decisions and exercise good judgment and, finally, healing the acute symptoms of abuse. The last often involves the use of cognitive-behavioral interventions such as cognitive restructuring, assertiveness training, and relaxation training.

Meiselman (1994) described another eclectic approach for treating adult survivors called reintegration therapy. Her approach is similar to the stage-by-dimension model of Lebowitz et al. (1993). Both of these models put a premium on insuring the client's safety which is broadly defined as physical well-being, establishing a safe living environment, and the reduction of self-destructive behaviors and PTSD-like symptoms. Following this, the focus of therapy is on achieving a personal understanding of the effects of the abuse and working on a healthy realignment of important social relationships. Lebowitz et al. advocate an ecological orientation attuned to the totality and variability of the survivor's experience. In their view, a beneficial therapy outcome requires not only effective treatment methods but also a willingness on the part of the therapist to remain flexible so there can be a good fit between the type of treatment offered and the particular needs of the patient.

9.18.5.2.4 Cognitive-behavioral treatments

Several researchers have written about the use of cognitive-behavioral approaches for treating adult survivors. One such approach is called imagery rescripting (Smucker et al., 1995). Designed for patients with PTSD and utilizing imaginal exposure and cognitive restructuring, this approach replaces victimization imagery with mastery imagery. The goal is to help the survivor feel strong and empowered instead of trapped in a childlike state of helplessness. Treatment involves identifying and challenging maladaptive self-schemas and replacing them with more adaptive schemas as viewed through the eyes of an adult. This is a relatively brief structured therapy that can be carried out in as few as nine sessions.

McCarthy (1986, 1990) has described a cognitive-behavioral intervention for treating victims of sexual trauma including CSA and incest. To his credit, he pays particular attention to the problems of male victims. The focus of his work is on the treatment of sexual dysfunctions that result from abuse as well as the maladaptive consequences of the victimization syndrome. The latter occurs when the person adopts a victim's role and acts it out sexually or otherwise until it becomes the defining event in their life. In the process, personal and sexual self-esteem plummet. For McCarthy, the goal of treatment is to teach the patient to be a survivor, not a victim, and to feel relaxed and in control in sexual situations. This often entails cognitive restructuring techniques. For example, instead of dwelling on their suffering and victim status, patients are encouraged to abide by the affirmation that "Living well is the best revenge." Relaxation training, guided imagery, and systematic desensitization are also used, especially if sexual fears or desire problems are present. If the person has a sexual problem and is in a relationship, the problem is dealt with as a couple issue. In many cases, the person is encouraged to confront the perpetrator with the idea of improving self-esteem by not reverting to the role of a helpless victim.

Resnick and Newton (1992) described a multifaceted treatment for survivors of sexual assault with PTSD. Their approach draws heavily on exposure techniques combined with stress-inoculation training. Patients are provided with a learning-based explanation of their fears and the treatment they receive. This includes a discussion of stimulus generalization, higher-order conditioning, avoidance and its role in maintaining anxiety, plus the different channels through which anxiety is expressed (i.e., physiological, cognitive, and behavioral). Among the skills taught are controlled breathing, deep muscle relaxation, thought stopping, and guided self-dialogue or self-instructions. Role playing and covert modeling are used to address inappropriate avoidance behavior.

Evidence for the usefulness of cognitive therapy is provided by Jehu (1989). Participants in his study were 34 women who were molested as children and suffered from mood-related problems. Jehu hypothesized that the mood problems resulted from self-blame and other distorted beliefs associated with their abuse. A treatment package consisting mainly of cognitive restructuring techniques was developed to modify the distortions in hopes of alleviating the associated mood disturbance. Results showed reductions in distorted beliefs that were accompanied by improvements in the patient's self-esteem and mood. This is one of the few treatment studies with a relatively large sample and objective data. While the results are encouraging, their importance is weakened by the absence of an appropriate control group.

Gazan (1986) used relaxation training combined with a cognitive restructuring approach similar to Jehu's (1989) to improve the sexual satisfaction of five women who were sexually abused as children. Results showed that treatment improved the sexual functioning of the women and helped them modify erroneous beliefs about their victimization.

Pearson (1994) conducted a thorough review of treatment techniques for adult female survivors of CSA. Focusing on techniques used in individual therapy, he divided them into 11 categories including cognitive and behavioral techniques. Some of the other categories were life skills training, role playing, gestalt work, and psychodrama; relationship building techniques, transactional analysis and inner child work, family of origin techniques, and hypnotherapy and guided imagery. Summing up his findings, Pearson points out that while most of these techniques seem to produce some success, there is virtually no data to suggest which techniques work best for treating which symptoms or underlying issues common to survivors.

9.18.5.3 Methodological Considerations

The studies we reviewed were weakened by the absence of comparison groups. None of them were controlled studies that allowed for unambiguous interpretations about the effectiveness of treatment. In many cases, they were anecdotal accounts of global therapies that were used with vaguely described and heterogeneous groups of persons. At best, the outcome research in this area has been characterized by the use of single group designs with some objective data in the form of scores on fairly generic psychological tests. Under these conditions, it is impossible to determine if the results attributed to treatment were any better than changes that could have resulted from the passage of time, therapist attention, or placebo effects.

There is a clear need for controlled outcome studies in this area. In an era of managed care, with its emphasis on empirically validated therapies, this need will only increase. While ethical and practical considerations preclude the use of randomized group designs, this should not discourage individual practitioners from using systematic assessment and evaluation procedures in their clinical practice. For therapists who regularly treat CSA survivors, causal inferences about treatment effects can be made using case study methods as long as treatment procedures are clearly specified and objective measures of the problem are taken continuously, before and during the course of therapy. With systematic replications across patients who have somewhat different characteristics, a compelling case can be made that the patients' improvement was due to the treatment they received and not some extraneous variable. The work of Maletzky (1980) with exhibitionists, Blanchard et al. (1985) with headache sufferers, and Morin, Stone, McDonald, and Jones (1994) with insomniacs are examples of how clinical replication studies can be used to gain valuable knowledge about the effectiveness of particular treatments for particular problems and patients. Similar methodologies can be used with CSA survivors.

In the absence of specific information about the type of subjects being treated, how they were recruited, attrition rates, and the way child sexual abuse was defined, it is impossible to know if, and to what extent, the results are generalizable to other populations, or if persons with certain characteristics benefit more than others.

Briere (1992) raises an interesting question: do adult survivors require abuse-specific treatments that are based on precipitating events and a presumed common etiology (i.e., their abuse), or would they fare just as well with treatments that focus on presenting symptoms and cut across etiologic boundaries? Most of the symptoms seen in CSA survivors are common to other disorders. Consequently, they might be amenable to problem-specific therapies that are widely available and whose efficacy has been established (Task Force on Promotion and Dissemination of Psychological Procedures, 1995).

The treatment studies we reviewed were also limited by the absence of follow-up data. Even if a treatment works in the short run, its long-term effects still need to be evaluated. In a similar vein, the specific methods of treatment in the studies discussed were often vaguely described which makes replication difficult, if not impossible. The availability of treatment manuals would be an important step for improving future studies and a benefit to clinicians as well (Beutler, 1993).

9.18.6 PREVENTION

Efforts to prevent CSA go back to the women's movement of the 1960s and 1970s. Shortly thereafter, concerns about protecting adult women from sexual abuse were extended to child populations. Thus, the 1980s saw a proliferation of programs geared toward protecting children from molestation.

Research suggests that primary prevention programs based on the acquisition of relevant safety skills hold the most promise (Wurtele & Miller-Perrin, 1992). One such program, *Safety with strangers* (Gipson et al., 1984), consists of a series of slides, with accompanying audio tape, that depicts 20 situations in which a stranger approaches a child. Scenes and verbal prompts were derived from interviews with incarcerated child molesters in an effort to present realistic stimuli to the target audience, primarily grade school children in a classroom setting. The perpetrators depicted reflect the demographic characteristics of convicted child molesters. As such, they feature primarily Caucasian men (although some women and minorities are included), ranging in age and style of dress, thus addressing Doyle's (1995) concern that "Children often think that a 'stranger' is an ugly man" and are not taught to be "on guard against woman or good looking, charming men" (p. 166).

Scenes include playing alone in the yard, becoming separated from parents in a store, and visiting the local video parlor. Verbal inducements range from bribes (e.g., "Come over to my house and I'll let you play with my kittens."), to trickery (e.g., "Your mother is in the hospital—hurry, come with me!") and physical threats if the child does not comply. Children then actively role play their response to these situations, with feedback and prompting from the program administrator. Correct responses include ignoring or saying no to the stranger, exiting the situation, and reporting the incident to a trusted adult.

Recognizing that CSA perpetrated by strangers constitutes the minority of cases relative to those perpetrated by family members, acquaintances, and caretakers, other programs address this possibility by teaching children to identify inappropriate advances regardless of their relationship to the perpetrator (e.g. Wurtele, Saslawsky, Miller, Marrs, & Brichter, 1986). In addition, these programs explicitly acknowledge the sexual nature of the situation. Rather than requiring children to discriminate good and bad touching based on emotional or physical pleasure, Wurtele (1987) suggested that children be taught not to let "any one bigger or older" (p. 12) than themselves look at or touch their private parts and vice versa. Wurtele's recommendation seems well grounded in that a study found that CSA survivors described some pleasurable aspects of the contact, at least in the early stages (Briggs & Hawkins, 1995). Once sexually-threatening situations have been identified, similar role-play methodology is used to teach children to resist these attempted violations.

Consumers must be aware that, unlike the CSA prevention programs reviewed above, many are available that have not been empirically tested. Furthermore, several criticisms have been made of those subjected to empirical scrutiny. For example, dependent measures have largely consisted of knowledge tests, rather than direct observation of behavioral skills, particularly in real-life situations. In addition, rates of CSA among treated versus nontreated children have yet to be ascertained. A second criticism is that programs are often not developmentally sensitive which may render them ineffective. In fact, some researchers contend that expecting preschool children to protect themselves from potential abuse is unreasonable (Peterson, 1993). However, Wurtele (1990) demonstrated improvements in knowledge and skills which lasted up to a month when she adapted her program to four-year-old preschoolers. Wurtele also addressed another criticism that participation in such programs will produce deleterious side effects. On the contrary, she encountered no program-related behavior problems according to either teachers or parents, and parents did not report increased fear in their children subsequent to participation in the program.

While skills-based primary prevention programs hold promise, they are limited in their impact considering the broader context in which abuse occurs. Researchers warn about imparting a false sense of security to child participants and their parents, recognizing that successful prevention efforts must reach beyond intervention with the children themselves (Wurtele & Miller-Perrin, 1992). Winett, King, and Altman (1989) have suggested that problems in prevention be approached simultaneously on multiple levels of analysis. While individual interventions may have merit, their effects might be enhanced by interventions targeting small groups such as the family, community systems, and sociocultural institutions. Wurtele and Miller-Perrin (1992) invoked a similar ecological model in their consideration of CSA, calling for a comprehensive prevention response which emphasizes person–environment interactions. Furthermore, they stressed the importance of secondary prevention which entails recognition of prodromal signs and symptoms that abuse has or is occurring, along with early intervention to terminate the trauma. Secondary prevention efforts seem crucial, considering that survivors who endured prolonged abuse suffer more serious consequences than those subjected to isolated instances of abuse or abuse of brief duration.

Focusing on the family system, Wurtele and Miller-Perrin (1992) suggested that primary prevention programs be extended from the classroom to the home, with parents implementing strategies such as those presented in the school. Parents are also advised to discuss sexuality and the potential for abuse in a direct and open manner with their children, disarming the secrecy which many perpetrators use to their advantage. These authors stressed the importance of CSA prevention programs occurring in the context of sound and accurate sexual education. Taking a different tack to primary prevention, Peterson (1993) reviewed studies in which regular home visits by various health care personnel resulted in fewer instances of parental physical abuse among at-risk families. Perhaps a similar strategy, designed to increase parental accountability, might reduce the chances of sexual abuse occurring as well.

With regard to secondary prevention, parents might become more vigilant for warning signs of sexual abuse. Wurtele and Miller-Perrin (1992) list behavioral, emotional, and physical indicators of sexual abuse by developmental level. Hernandez (1995) provided evidence that both male and female adolescents experiencing CSA manifest eating-disordered behavior. Wurtele and Miller-Perrin provide information concerning age-appropriate sexual knowledge as another means of detecting potential abuse. For example, preschool children who use correct terminology for genitalia or girls who are knowledgeable about penile erections prior to age 12 may have acquired this information through abuse experiences. Another way in which parents can minimize the impact of abuse, once it is detected, involves their subsequent reaction to the child. Wurtele and Miller-Perrin provide guidelines regarding appropriate and inappropriate verbal, emotional, and behavioral responses. Indeed, many children who bring their abuse to the attention of an adult are met with negative reactions (e.g., Briggs & Hawkins, 1995). Doyle (1995) warns that "The phrase 'Children should be seen and not heard' is manna to perpetrators" (p. 167), and stresses the importance of responding to children's reports of abuse in a nonpunitive manner.

On a community level, primary prevention programs might continue to be sponsored through school settings as well as other community bodies such as churches and day care settings. In addition, well-publicized, organized community involvement might deter would-be perpetrators from acting on their intentions. Of course, community members who have contact with and authority over children must be carefully screened regarding their potential to perpetrate abuse. Wurtele and Miller-Perrin (1992) strongly suggest background checks for staff and volunteers in youth organizations as well as foster care and juvenile detention settings.

Secondary prevention efforts in the community might parallel those described for parents. That is, various professionals throughout the community might be taught to better recognize the early signs of abuse and to communicate their concerns in ways that facilitate entry into appropriate treatment. Wurtele and Miller-Perrin (1992) contend that such resources are underutilized, citing several barriers such as fears associated with reporting CSA and lack of support from professional supervisors and agencies.

9.18.7 FUTURE DIRECTIONS

The past decade has seen a phenomenal increase in attention paid to CSA among health care professionals and the general public alike. While this focus has spurred the development of prevention and treatment programs, much work remains to be done. For example, many interventions which have not been subjected to empirical scrutiny are readily available for public consumption. These include preventive programs disseminated through schools as well as self-help groups and literature for CSA survivors. Given the highly sensitive nature of CSA, concerns abound that consumers may experience emotional and relationship distress without adequate support to cope with these potentially adverse reactions. Enns, McNeilly, Corkery, and Gilbert (1995) have encouraged researchers to work in conjunction with clinicians who treat CSA survivors, as well as self-help organizations comprised of survivors, in an effort to enhance the quality of existing interventions. With the continuing growth of managed care, self-help modalities may become important supplements to formal therapy because they have the potential to reinforce coping skills as well as a sense of empowerment.

Qualitative research in which CSA survivors act as participants may also serve to empower survivors while generating hypotheses to be tested by quantitative methods. Certainly, a need exists for more quantitative research, including controlled treatment outcome studies that use objective measures of specific CSA-related problems.

Assessment begins with ascertaining an individual's abuse status. No longer are gross determinations of abuse history accepted in light of research implicating variables such as duration, type of contact, and relationship to the perpetrator as important factors in symptom presentation. Continuing efforts to specify, if not to quantify, the abuse experience are essential. While researchers endeavor to operationally define CSA, many health care practitioners are reluctant to ask clients about their abuse histories. Enns et al. (1995) implicate the false-memory debate as a contributor to this reluctance. That is, clinicians may be wary of accusations that they have implanted inaccurate recollections of abuse in their clients. To counter this fear, they suggest that questions about abuse be integrated within a comprehensive assessment, perhaps using open-ended questions about negative childhood experiences and family member's roles. In this way, clinicians may be less likely to be perceived as leading their clients to describe nonexistent events.

In this chapter, overview of the current literature on adult victims of CSA is provided. In doing so, problems defining CSA were discussed, and the impact they have on epidemiological and treatment-oriented research. It was noted that the sexual mistreatment of children is all too common, and that its long-term effects are varied and moderated by ecological variables such as the duration and type of abuse, the relationship between victim and perpetrator, and the way in which the abuse is processed by the victim, both intellectually and emotionally. It was pointed out that different theoretical models and techniques are available for treating adult victims of CSA, but all of them are based more on conjecture and anecdotal support than on empirical evidence. Controlled outcome studies in this area, with appropriate control groups and long-term follow-up, are nonexistent. Finally, ways to prevent CSA and minimize its deleterious effects by early detection and treatment were described. The study of adult victims of CSA is relatively new. While much has been learned since interest in this area developed in the 1980s, researchers are still in the early stages of understanding what can be done to prevent the problem and mitigate its damaging effects.

ACKNOWLEDGMENTS

Appreciation is extended to Mr. James Wilson for his generous support, and to Ms. Shannon Witcher for her clerical assistance.

9.18.8 REFERENCES

Abel, G., & Rouleau, J. (1990). The nature and extent of sexual assault. In W. Marshall, D. Laws, & H. Barbaree (Eds), *Handbook of sexual assault: Issues, theories, and treatment of the offender*. New York: Plenum.

Alexander, P. (1992). Application of attachment theory to the study of sexual abuse. *Journal of Consulting and Clinical Psychology, 60,* 185–195.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders*. Washington, DC: Author.

Anderson, J., Martin, J., Mullen, P. Romans, S., & Herbison, P. (1993). Prevalence of childhood sexual abuse experiences in a community sample of women. *Journal of American Academy of Child and Adolescent Psychiatry, 32,* 911–919.

Bagley, C. (1991). The prevalence and mental health sequelae of child sexual abuse in a community sample of women aged 18 to 27. *Canadian Journal of Community Mental Health, 10,* 103–116.

Bagley, C., Wood, M., & Young, L. (1994). Victim to abuser: Mental health and behavioral sequelae of child sexual abuse in a community survey of young adult males. *Child Abuse and Neglect, 18,* 683–697.

Benedict, L., & Zautra, A. (1993). Family environment characteristics as risk factors for child sexual abuse. *Journal of Consulting and Clinical Psychology, 22,* 365–374.

Bergner, R., Delgado, L., & Graybill, D. (1994). Finkelhor's risk factor checklist: A cross validation study. *Child Abuse and Neglect, 18,* 331–340.

Beutler, L. (1993). Designing outcome studies: Treatment of adult victims of sexual abuse. *Journal of Interpersonal Violence, 8,* 402–414.

Beutler, L., & Hill, C. (1992). Process and outcome research in the treatment of adult victims of childhood sexual abuse: Methodological issues. *Journal of Consulting and Clinical Psychology, 60,* 204–212.

Blanchard, E., Andrasik, F., Evans, D., Neff, D., Appelbsum, K., & Rodichok, L. (1985). Behavioral treatment of 250 chronic headache patients: A clinical replication series. *Behavior Therapy, 16,* 308–327.

Bowlby, J. (1973). *Attachment and loss: Vol 1. Attachment.* New York: Basic Books.

Briere, J. (1992). Methodological issues in the study of sexual abuse effects. *Journal of Consulting and Clinical Psychology, 60,* 196–203.

Briggs, F., & Hawkins, R. (1995). Protecting boys from the risk of sexual abuse. *Early Child Development and Care, 110,* 19–32.

Browne, A., & Finkelhor, D. (1986). The impact of child sexual abuse: A review of the research. *Psychological Bulletin, 99,* 66–77.

Burnam, M., Stein, J., Golding, J., Seigel, J., Sorenson, S., Forsythe, A., & Telles, C. (1988). Sexual assault and mental disorders in a community population. *Journal of Consulting and Clinical Psychology, 56,* 843–850.

Cappelleri, J., Eckenrode, J., & Powers, J. (1993). The epidemiology of child abuse: Findings from the Second National Incidence and Prevalence Study of Child Abuse and Neglect. *American Journal of Public Health, 83,* 1622–1624.

Cole, P., & Putnam, F. (1992). Effect of incest on self and social functioning: A developmental psychopathology perspective. *Journal of Consulting and Clinical Psychology, 60,* 174–184.

Downs, W. (1993). Developmental considerations for the effects of childhood sexual abuse. *Journal of Interpersonal Violence, 8,* 331–345.

Doyle, C. (1995). *Helping strategies for child sexual abuse.* London: Whiting & Birch.

Enns, C., McNeilly, C., Corkery, J., & Gilbert, M. (1995). The debate about delayed memories of child sexual abuse: A feminist perspective. *Counseling Psychologist, 23,* 181–279.

Feinauer, L., & Stuart, D. (1996). Blame and resilience in women sexually abused as children. *American Journal of Family Therapy, 24,* 31–40.

Finkelhor, D. (1979). *Sexually victimized children.* New York: Free Press.

Finkelhor, D. (1984). *Child sexual abuse: New theory and research.* New York: Free Press.

Finkelhor, D. (1986). *A source book on child sexual abuse: New theory and research.* Beverly Hills, CA: Sage.

Finkelhor, D. (1988). The trauma of child sexual abuse: Two models. In G. Wyatt & G. Powell (Eds.), *The lasting effects of child sexual abuse.* Newbury Park, CA: Sage.

Finkelhor, D. (1990). Early and long-term effects of child sexual abuse: An update. *Professional Psychology: Research and Practice, 21,* 325–330.

Finkelhor, D. (1994). The international epidemiology of child sexual abuse. *Child Abuse and Neglect, 18,* 409–417.

Finkelhor, D., & Browne, A. (1985). The traumatic impact of child sexual abuse: A conceptualization. *American Journal of Orthopsychiatry, 55,* 530–541.

Finkelhor, D., Hotaling, G., Lewis, A., & Smith, C. (1990). Sexual abuse in a national survey of adult men and women: Prevalence, characteristics, and risk factors. *Child Abuse and Neglect, 14,* 19–28.

Freund, K., & Kuban, M. (1994). The basis of the abused abuser theory of pedophilia: A further elaboration on an earlier study. *Archives of Sexual Behavior, 23,* 553–563.

Fromouth, M. (1986). The relationship of childhood sexual abuse with later psychological and sexual adjustment in a sample of college women. *Child Abuse and Neglect, 10,* 5–15.

Gazan, M. (1986). An evaluation of a treatment package designed for women with a history of sexual victimization in childhood and sexual dysfunctions in adulthood. *Canadian Journal of Community Mental Health, 5,* 85–102.

Gipson, M., Green, J., Telch, M., Watkins, P., Griffith, L., Agras, A., & Ballard, B. (1984). *Safety with strangers: A manual for teachers.* Fort Lauderdale, FL: Adam Walsh Center.

Green, A. (1993). Child sexual abuse: Immediate and long-term effects and intervention. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 890–902.

Haugaard, J., & Reppucci, N. (1988). *The sexual abuse of children.* San Francisco: Jossey-Bass.

Hernandez, J. (1995). The concurrence of eating disorders with histories of child abuse among adolescents. *Journal of Child Sexual Abuse, 4,* 73–85.

Hulme, P., & Grove, S. (1994). Symptoms of female survivors of child sexual abuse. *Issues in Mental Health Nursing, 15,* 519–532.

Jehu, D. (1989). Mood disturbances among women clients sexually abused in childhood. *Journal of Interpersonal Violence, 4,* 164–184.

Jumper, S. A. (1995). A meta-analysis of the relationship of child sexual abuse to adult psychological adjustment. *Child Abuse and Neglect, 19,* 715–728.

Katz, R. (1990). Psychological adjustment in adolescent child molesters. *Child Abuse and Neglect, 14,* 567–575.

Kendall-Tackett, K., Williams, L., & Finkelhor, D. (1993). Impact of sexual abuse on children: A review and synthesis of recent empirical findings. *Psychological Bulletin, 113,* 164–180.

Knutson, J. (1995). Psychological characteristics of maltreated children: Putative risk factors and consequences. *Annual Review of Psychology, 46,* 401–431.

Lawson, C. (1993). Mother–son sexual abuse: Rare or underreported? A critique of the research. *Child Abuse and Neglect, 17,* 261–269.

Lebowitz, L., Harvey, M., & Herman, J. (1993). A stage by dimension model of recovery from sexual trauma. *Journal of Interpersonal Violence, 8,* 378–391.

Lechner, M., Vogel, M., Garcia-Shelton, L., Leichter, J., & Steibel, K. (1993). *Journal of Family Practice, 36,* 633–638.

Lisak, D. (1994). The psychological impact of sexual abuse: Content analysis of interviews with male survivors. *Journal of Traumatic Stress, 7,* 525–547.

Lloyd, D. (1992). Ritual child abuse: Definitions and assumptions. *Journal of Child Sexual Abuse, 1,* 1–14.

Loftus, E. (1993). The reality of repressed memories. *American Psychologist, 48,* 518–537.

Maletzky, B. (1980). Self-referred versus court-referred sexually deviant patients: Success with assisted covert sensitization. *Behavior Therapy, 11,* 306–314.

Mallinckrodt, B., McCreary, B., & Robertson, A. (1995). Co-occurrence of eating disorders and incest: The role of attachment, family environment, and social competencies. *Journal of Counseling Psychology, 42,* 178–186.

Margolin, L. (1992). Sexual abuse by grandparents. *Child Abuse and Neglect, 16,* 735–741.

McCarthy, B. (1986). A cognitive-behavioral approach to understanding and treating sexual trauma. *Journal of Sex & Marital Therapy, 12,* 322–329.

McCarthy, B. (1990). Treating sexual dysfunction associated with prior sexual trauma. *Journal of Sex and Marital Therapy, 16,* 143–146.

McCarthy, B. (1992). Sexual trauma: the pendulum has swung too far. *Journal of Sex Education and Therapy, 18,* 1–10.

McMillen, C., Zuravin, S., & Rideout, G. (1995). Perceived benefit from child sexual abuse. *Journal of Consulting and Clinical Psychology, 63,* 1037–1043.

McNew, J., & Abell, N. (1995). Post-traumatic stress symptomatology: Similarities and differences between Vietnam veterans and adult survivors of childhood sexual abuse. *Social Work, 40,* 115–126.

McNulty, C., & Wardle, J. (1994). Adult disclosure of sexual abuse: A primary cause of psychological distress? *Child Abuse & Neglect, 18,* 549–555.

Meiselman, K. (1994). Treating survivors of child sexual abuse: A strategy for reintegration. In J. Briere (Ed.), *New directions in mental health services: Assessing and treating victims of violence* (pp. 91–100). San Francisco: Jossey-Bass.

Mennen, F. (1993). Evaluation of risk factors in childhood sexual abuse. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 934–939.

Morin, C., Stone, J., McDonald, K., & Jones, S. (1994). Psychological management of insomnia: A clinical replication series with 100 patients. *Behavior Therapy, 25,* 291–310.

Morrow, S., & Smith, M. (1995). Constructions of survival and coping by women who have survived childhood sexual abuse. *Journal of Counseling Psychology, 42,* 24–33.

Newman, A., & Peterson, C. (1996). Anger of women incest survivors. *Sex Roles, 34,* 463–473.

Pearson, Q. (1994). Treatment techniques for adult female survivors of childhood sexual abuse. *Journal of Counseling & Development, 73,* 32–37.

Peterson, C., & Bossio, L. (1991). *Health and optimism.* New York: Free Press.

Peterson, M. P. (1993). Physical and sexual abuse among school children: Prevalence and prevention. *Educational Psychology Review, 5,* 63–86.

Polusny, M., & Follette, V. (1995). Long-term correlates of child sexual abuse: Theory and review of the empirical literature. *Applied and Preventive Psychology, 4,* 143–166.

Ramona vs. Isabel, No. C61898 (Superior Court, Napa County, CA, 1994).

Resnick, H., & Newton, T. (1992). Assessment and treatment of post-traumatic stress disorder in adult survivors of sexual assault. In D. Fox (Ed.), *Treating PTSD: Cognitive behavioral strategies*. New York: Guilford.

Rowan, A., & Foy, D. (1993). Post-traumatic stress disorder in child sexual abuse survivors: A literature review. *Journal of Traumatic Stress, 6,* 3–20.

Rowan, A., Foy, D., Rodriguez, N., & Ryan, S. (1994). Post-traumatic stress disorder in a clinical sample of adults sexually abused as children. *Child Abuse and Neglect, 18,* 51–61.

Rudin, M., Zalewski, C., & Boomer-Turner, J. (1995). Characteristics of child sexual abuse victims according to perpetrator gender. *Child Abuse and Neglect, 19,* 963–973.

Russell, A., & Trainor, C. (1984). *Trends in child abuse and neglect: A national perspective*. Denver, CO: The American Association.

Russell, D. (1983). The incidence and prevalence of intrafamilial and extrafamilial sexual abuse of female children. *Child Abuse and Neglect, 7,* 133–146.

Smucker, M., Dancu, C., Foa, E., & Niederee, J. (1995). Imagery rescripting: A new treatment for survivors of childhood sexual abuse from post traumatic stress. *Journal of Cognitive Psychotherapy 9,* 3–17.

Sullivan, P., & Knutson, J. (1993). *The relationship between child abuse and neglect and disabilities: Implications for research and practice*. Presented at the National Conference on Child Abuse and Neglect, Pittsburgh, PA.

Task Force on Promotion and Dissemination of Psychological Procedures (1995). Training and dissemination of empirically-validated treatments. *The Clinical Psychologist,* 48, 3–23.

Urquiza, A., & Keating, L. (1990). The prevalence of sexual victimization in males. In M. Hunter (Ed.), *The sexually abused male: Prevalence, impact, and treatment* (Vol. 1, pp. 89–103). New York: Lexington.

Valentine, L., & Feinauer, L. (1993). Resilience factors associated with female survivors of childhood sexual abuse. *American Journal of Family Therapy, 21,* 216–224.

Vizard, E., Monck, E., & Misch, P. (1995). Child and adolescent sex abuse perpetrators: A review of the research literature. *Journal of Child Psychology and Psychiatry, 36,* 731–756.

Walker, C., Bonner, B., & Kaufman, K. (1988). *The physically and sexually abused child: Evaluation and treatment*. New York: Pergamon.

Walker, L. (1994). *Abused women and survivor therapy: A practical guide for the psychotherapist*. Washington, DC: American Psychological Association Press.

Windle, M., Windle, R., Scheidt, D., & Miller, G. (1995). Physical and sexual abuse and associated mental disorders among alcoholic inpatients. *American Journal of Psychiatry, 152,* 1322–1328.

Winett, R., King, A., & Altman, D. (1989). *Health psychology and public health: An integrative approach.* Elmsford, NY: Pergamon.

Wurtele, S. (1987). School-based sexual abuse prevention programs: A review. *Child Abuse and Neglect, 10,* 85–92.

Wurtele, S. (1990). Teaching personal safety skills to four-year old children: A behavioral approach. *Behavior Therapy, 21,* 25–32.

Wurtele, S., & Miller-Perrin, C. (1992). *Preventing child sexual abuse: Sharing the responsibility*. Lincoln, NE: University of Nebraska Press.

Wurtele, S., Saslawsky, D., Miller, C., Marrs, S., & Britcher, J. (1986). Teaching personal safety skills for potential prevention of sexual abuse: A comparison of treatments. *Journal of Consulting and Clinical Psychology, 54,* 688–692.

Wyatt, G. (1985). The sexual abuse of Afro-American and White-American women in childhood. *Child Abuse and Neglect, 9,* 507–519.

Wyatt, G., Guthrie, D., & Notgrass, C. (1992). Differential effects of women's child sexual abuse and subsequent sexual revictimization. *Journal of Consulting and Clinical Psychology, 60,* 167–173.

Wynkoop, T., Capps, S., & Priest, B. (1995). Incidence and prevalence of child sexual abuse: A critical review of data collection procedures. *Journal of Child Sexual Abuse, 4,* 49–66.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.19
Adult Victims of Physical Violence

RON ACIERNO, CHRISTINA BYRNE, HEIDI S. RESNICK,
and DEAN G. KILPATRICK
Medical University of South Carolina, Charleston, SC, USA

9.19.1 INTRODUCTION 307

9.19.2 PREVALENCE 308
  9.19.2.1 Measurement Problems 308
  9.19.2.2 Prevalence Rates 308

9.19.3 NATURE 309
  9.19.3.1 Domestic Violence 309
  9.19.3.2 Etiological Theory 310
  9.19.3.3 Trauma Sequelae: PTSD 310
  9.19.3.4 Trauma Sequelae: Depression 311
  9.19.3.5 Trauma Sequelae: Substance Abuse 312
  9.19.3.6 Trauma Sequelae: Panic 312

9.19.4 ASSESSMENT 313
  9.19.4.1 Trauma Assessment 313
    9.19.4.1.1 Trauma assessment: interviews 314
  9.19.4.2 Post-trauma Assessment of PTSD: Interviews 315
  9.19.4.3 Post-trauma Assessment of PTSD: Questionnaires 316
  9.19.4.4 Post-trauma Assessment of Related Symptomatology: Interviews 317
  9.19.4.5 Post-trauma Assessment of Related Symptomatology: Questionnaires 318

9.19.5 TREATMENT 318
  9.19.5.1 PTSD 318

9.19.6 FUTURE DIRECTIONS 321

9.19.7 SUMMARY 321

9.19.8 REFERENCES 322

9.19.1 INTRODUCTION

Violence is an unavoidable facet of human experience. Virtually no one escapes at least vicarious exposure to severe interpersonal conflict, and over 10% of the population endures direct aggression in the form of simple, aggravated, or domestic assault at some point in their lives (Resnick, Kilpatrick, Dansky, Saunders, & Best, 1993). Tangible effects of physical assault are far-reaching and devastating. The US Department of Justice reported that annual violent crime results in at least 2.2 million known injuries, requiring over 700 000 hospitalization days, with medical and lost productivity expenses in excess of $6 000 000 000 (Harlow, 1989). It is, therefore, not surprising that violence is considered both a social and public health emergency in the USA and other countries (Koop, 1992; Novello, 1992). Importantly, the overt economic and medical impact of violence is complemented by less visible, albeit equally significant psychological and emotional difficulties endured by victims, including extreme levels of anxiety, depression, substance abuse, and impaired interpersonal and vocational functioning (Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995; Kilpatrick, Saunders, Veronen, Best, & Von, 1987; Resnick et al., 1993).

9.19.2 PREVALENCE

9.19.2.1 Measurement Problems

Initial estimates of physical assault prevalence were based on indirect indicators of violence, such as police reports or hospital records. However, because most crime is not reported, both the relevance and validity of these data were severely limited (Kilpatrick et al., 1987; Kilpatrick, Resnick, Saunders, & Best, in press). The ongoing National Crime Victimization Survey (NCVS) (Bachman, 1994), a primary source of assault and victimization data for the US Bureau of Justice Statistics, and the Epidemiologic Catchment Area Survey (Helzer, Robins, & McEvoy, 1987), were two initial attempts to address these early assessment problems through use of interview, rather than archival data. Weaknesses in NCVS assessment methodology continue to exist, however, and NCVS assault prevalence estimates are low compared to those reported by epidemiological researchers in the social sciences.

Disparity between past survey findings and those of contemporary social scientists results from differences in assessment techniques. Specifically, initial surveys employed legalistic terminology to define assault variables to the exclusion of behavioral descriptions of assault events. Consequently, reporting rates were greatly affected by both idiotypic interpretation of that which defined assault and psychological factors that impinged upon reporting. In response to these short-comings, criminal justice researchers directing the NCVS altered their survey techniques in 1992 and incorporated the relatively more sensitive assessment methodology utilized by social scientists (e.g., Kilpatrick, Resnick, Saunders, & Best, 1989). Specifically, the redesigned survey included behavioral descriptions of assault variables in the form of closed-ended questions. Expectedly, estimates of assault prevalence increased dramatically.

9.19.2.2 Prevalence Rates

Prevalence rates of physical assault appear to vary by gender, race, and age. Across studies, physical assault is either the most common, or among the most common, forms of experienced trauma in the general population (Breslau, Davis, Andreski, & Petersen, 1991; Kessler et al., 1995; Resnick et al., 1993). Investigating a nationally representative sample of 4009 women, Resnick et al. estimated that 9.9 million female Americans (10.28%) had experienced severe physical assault at some point in their lives. Consistent with these findings, Kessler et al. (1995) reported that 6.9% of the women and 11.9% of the men in their stratified random sample of 5877 participants had been physically assaulted. These are rates directionally similar to those reported by Norris (1992) for women (11.7%) and men (18.7%). Comparable rates of assault were found by Kilpatrick et al. (1987), who studied a random sample of 391 women and by Breslau et al. (1991), who assessed 1007 randomly sampled members of a large Health Maintenance Organization (HMO), with overall prevalence ranging from 9.7% to 15% of participants. Not surprisingly, rates of assault were lower in the NCVS survey, in which Bachman (1994) noted that approximately 3.8 million American women and 4.7 million American men reported being physically attacked.

The 1992–1993 NCVS annual data revealed that men reported suffering violent crime at the hands of strangers at a rate 11 times that of known perpetrators, whereas women experienced assault primarily by known assailants. In terms of lifetime prevalences, men were equally likely to be attacked by known or unknown perpetrators, while women were far more frequently attacked by known individuals such as husbands or ex-husbands (31%), boyfriends (16%), other relatives (27%), and acquaintances (9%) (Kilpatrick et al., in press). Direct sex-based comparisons revealed that men were about twice as likely to be assaulted by strangers as women, and were more likely to be assaulted by multiple assailants (Bachman, 1994). Further, Kessler et al. (1995) noted that while men were more likely to experience trauma in general, women were more likely to experience severe trauma (e.g., rape).

In addition to gender differences, Norris (1992) found that a significantly greater proportion of Caucasians experienced physical assault (18.4%) than non-Caucasians (11.6%), although the latter group appeared relatively more susceptible to post-traumatic psychopathology. Kessler et al. (1995) also noted comparable racial differences, irrespective of income level. Although no race-based differences were discovered by Breslau et al. (1991), these investigators reported that assault was associated with being male, having less than a college education, or having a history of early misconduct.

Age and income differences have also been identified, with youth and poverty associated with increased risk of being assaulted. Norris

(1992) reported assault prevalence rates of 16.8% in younger adults vs. 8.6% in older adults. Similarly, Kilpatrick et al. (in press) noted that 23.7% of their sample of women reported being assaulted before age 17, 46.8% experienced assault between ages 17 and 29, and only 26.8% reported that their first attack occurred after age 30. These investigators also found that individuals with annual incomes below $10 000 were at increased risk of victimization. In addition to age and income, experience of past assault was associated with future assault. Of those with a positive assault history, 30–54% had been victimized at least twice (Kilpatrick et al., 1987; Resnick et al., 1993; Steketee & Foa, 1987).

9.19.3 NATURE

Psychological sequelae of trauma are often severe and debilitating. "Importantly, several relevant aspects of victimization are strongly related to subsequent development and/or exacerbation of psychopathology. In determining factors that increase risk of negative outcome, it is useful to define potentially traumatic events along objective and subjective lines (Weaver & Clum, 1995). Objective parameters of assault include whether a weapon was used, duration and location of the attack, level of force imposed, resultant physical injury, and relationship of perpetrator to the victim. By contrast, subjective aspects of physical assault include perceptions of life threat or potential for injury, appraisal of responsibility and self-blame or guilt, and perceptions of controllability and attacker intentionality (Kilpatrick, Resnick, et al., 1989).

Subjective factors appear to be relatively more important in determining post-trauma psychological distress than objective factors (Weaver & Clum, 1995). For example, Kilpatrick, Resnick et al. (1989) noted that perceived threat accounted for 6.4% of the unique variance in prediction of post-traumatic stress disorder (PTSD), whereas actual injury accounted for 4.9%. Indeed, risk of post-traumatic psychopathology is greatest in individuals who report that during the trauma they feared they would be seriously injured or die, or actually were injured (Green, 1990; Kilpatrick, Resnick et al., 1989; Resnick et al., 1993; Wirtz & Harrell, 1987a). Specifically, 30.8% of those assault victims who reported both perceived life threat and actual physical injury developed PTSD, compared to 20.6% with only life threat, and 25% with only injury (Kilpatrick, Resnick et al., 1989). Because perceived threat of serious injury and actual injury are both risk factors for post-traumatic psychopathology, it is important to identify which forms of trauma are associated with these events. Along these lines, Resnick et al. (1993) demonstrated that threat of injury or actual injury were most often observed in physical assault victims (90.8%), followed by rape victims (67.9%). Further, both perceived threat of injury and actual injury were reported by 61.1% of physical assault victims, compared to 36.7% of rape victims.

These findings are hardly surprising when considering that the intention of most assailants during physical assault is to inflict pain and injury. In addition, these results are all the more impressive when trauma characteristics and outcomes of physical assault are compared to those of severe sexual assault. Resnick et al. (1993) found that women with a history of physical trauma were more likely to meet criteria for both life (38.5%) and current (17.8%) PTSD, than victims of all other trauma, including rape (for which lifetime and current rates of PTSD were 32% and 12.4%, respectively). Similarly, in cases of domestic violence, relatively larger proportions of victims of physical assault feared that they would be killed (100%) than victims of spousal rape (82%) (Riggs, Kilpatrick, & Resnick, 1992). Finally, in their meta-analysis of studies measuring psychological distress associated with interpersonal violence, Weaver and Clum (1995) found no difference between physical and sexual assault.

Other factors that may be positively related to development of post-trauma psychopathology include: intensity of anxiety and distress during an attack, past history of psychopathology, past history of trauma, familial psychopathology, presence and intensity of dissociation during trauma, and low age, education, and socioeconomic status (Breslau et al. 1991; Freedy, Resnick, Kilpatrick, Dansky, & Tidwell, 1994; Norris, 1992). By contrast, stable sources of social support appear to buffer against development of post-traumatic problems in psychological functioning (Burgess & Holstrom, 1974; Wilson & Kraus, 1985).

9.19.3.1 Domestic Violence

Since the late 1970s, the problem of relationship violence has been formally acknowledged as a major social problem. Survey data indicate that violence occurs in 20–30% of marriages in the USA (Straus & Gelles, 1990; Straus, Gelles, & Steinmetz, 1980) and that approximately 1.8 million women are physically abused by their partners each year (Straus & Gelles, 1986). Furthermore, nearly 1700 women die each year as a direct result of their partner's abuse

(Steinmetz, 1978), and women are more likely to be victimized by current or former male partners than by other perpetrators (Koss et al., 1994).

9.19.3.2 Etiological Theory

Psychological models proposed to explain the process by which assault and other forms of traumatic stress potentiate psychopathology include conditioning, cognitive, and social learning theories. An integration of these models appears best to account for post-traumatic symptom outcome. In such a comprehensive etiological model, post-traumatic fear, anxiety, and avoidance behaviors are presumed to be conditioned responses that may be modulated by cognitive styles and social context, and further influenced by individual characteristics such as age or intelligence. Kilpatrick, Veronen, and Resick (1979), and Keane, Zimering, and Caddell (1985) adapted Mowrer's (1960) two-factor model of phobic avoidance to post-traumatic psychological reactions, particularly PTSD. In this conceptualization, assault is the unconditioned stimulus that leads to the unconditioned responses of avoidance, hyperarousal, and intrusive ideation (i.e., the symptom constellation of PTSD). Initial responses might also include symptoms of depression and panic. These responses become associated with salient stimuli present in the environment during the attack (e.g., perpetrator race and build, location, time of day) that then serve to elicit conditioned fear responses in the future. Thus, when a victim is exposed to an individual of the same sex or race as the perpetrator, or, for example, finds herself walking to her car in the darkness, a conditioned response identical or very similar to the unconditioned response will occur. In order to reduce or eliminate extreme aversiveness of this learned anxiety response, individuals will seek to escape from conditioned stimuli. Because escape results in diminution of the fear response, avoidance behavior is negatively reinforced and the probability of its future occurrence is increased. Moreover, because repeated, extended exposure to conditioned stimuli is avoided, extinction of fear responses does not occur. Thus, according to two-factor theory, conditioned post-traumatic anxiety responses are initially acquired through classical conditioning and maintained (in the form of avoidance) through operant conditioning.

Some indirect support for conditioning theory, particularly as it relates to avoidance behavior, has been provided by Wirtz and Harrell (1987a, 1987b). They demonstrated that victims who avoided fear stimuli (i.e., engaged in behavior that would, predictably, maintain conditioned fear responses) experienced more psychological distress than those who exposed themselves to relevant fear-producing stimuli, irrespective of initial distress level. Moreover, success of exposure-based treatments for PTSD (e.g., Fairbank & Keane, 1982; Foa, Hearst-Ikeda, & Perry, 1995; Foa, Rothbaum, Riggs, & Murdock, 1991) provides additional indirect support for the model. Refinement of the conditioning theory of post-traumatic psychopathology has included relatively greater consideration of cognitive factors and individual characteristics. Foa, Steketee, and Rothbaum (1989) provided a thorough discussion of this point and illustrated the major role cognitive components, such as perceptions of controllability over past and future events, and estimates of the risk of danger associated with specific stimuli, appear to play in determining intensity and duration of anxiety responses. Moreover, Keane (1985) highlighted the need to consider individual characteristics, such as family history of psychopathology, level of social support, and age, when assessing and treating post-traumatic symptomatology.

9.19.3.3 Trauma Sequelae: PTSD

According to the *Diagnostic and statistical manual of mental disorders* (4th ed.; *DSM-IV*) (American Psychiatric Association, 1994), a diagnosis of PTSD is assigned only when an individual is exposed to a traumatic event that both presents actual or threatened death or serious injury to oneself or others, and elicits intense fear, helplessness, or horror. Obviously, physical assault sufficiently meets this criterion. Additional parameters of the PTSD diagnosis include symptoms of re-experiencing, avoidance, and hyperarousal that persist for at least one month and cause functional impairment in interpersonal or vocational spheres. Re-experiencing may take the form of recurrent recollections of the event, nightmares, flashbacks, or reactivity upon exposure to traumatic cues. Avoidance is typically manifest as behavioral or cognitive escape from thoughts, feelings, individuals, or places associated with the trauma, as well as the experience of feelings of detachment, foreshortened future, and restricted affect. Finally, hyperarousal is indicated by elevated startle response, sleep disturbances, hypervigilance, and concentration difficulties.

Overall, population prevalence estimates of lifetime PTSD range from 7.8% to 12.3%, and prevalence estimates of past-year PTSD range from 2.3% to 4.2% (Breslau et al., 1991; Kessler et al., 1995; Resnick et al., 1993). Groups at increased risk for developing PTSD resemble those at greater risk of assault, including individuals low in social and economic resources (Freedy, Shaw, Jarrell, & Masters, 1992; Hobfoll, 1989; Kilpatrick et al., in press). Also at increased risk are non-Caucasians (Kulka et al., 1990), and those with familial or personal histories of psychopathology (Davidson, Smith, & Kudler, 1989). Moreover, PTSD is significantly more prevalent in women, who are 2–15 times more likely to develop the disorder in response to physical attack than men (21.3% of women vs. 1.8% of men develop PTSD after physical assault) (Kessler et al., 1995; Norris, 1992). PTSD prevalence rates increase with age for men, but not for women, indicating that women generally experience trauma and suffer its psychological after-effects earlier.

Kessler et al. (1995) reported that the most common precipitants of PTSD in men are combat and witnessed violence, whereas for women sexual and physical assault/abuse are primary index events. Rates of PTSD in women following physical assault and abuse are comparable to those observed with rape. Indeed, Kilpatrick et al. (1987) noted that lifetime rates of postrape and postassault PTSD in women were 57.1% and 36.8%, respectively, while Resnick et al. (1993) found rates of postrape and postassault PTSD of 32% and 38.5%, respectively. Finally, Kessler et al. observed rates of postsexual and physical assault PTSD of 48.0% and 39.1%, respectively.

Data regarding the course of PTSD have been presented for rape victims but not for victims of physical assault. However, given the comparability of outcomes related to physical and sexual assault, it is reasonable to assume that similar proportions of seriously physically assaulted victims will suffer from the disorder over time. Rothbaum, Foa, Riggs, Murdock, and Walsh (1992) found that 90% of rape victims met symptom criteria for PTSD within two weeks of rape, and about 50% continued to meet criteria three months later. Moreover, Kilpatrick et al. (1987) showed that the disorder was present 17 years after assault in 16.5% of cases. Overall, about 50% of those who do meet criteria for PTSD recover within two years, while almost 33% do not experience full remission of symptoms, even after several years (Kessler et al., 1995).

9.19.3.4 Trauma Sequelae: Depression

Although often considered secondary to PTSD, depression is frequently observed in crime victims (Breslau et al., 1991; Kilpatrick et al., 1985). The co-occurrence of depression and anxiety following trauma is well documented, and significant overlap exists between PTSD and the diagnostic criteria for depression (Breslau et al., 1991; Davidson, Hughs, Blazer, & George, 1991; Helzer et al., 1987; Uddo, Allain, & Sutker, 1996). Breslau et al. found that 36.6% of those diagnosed with PTSD concurrently met criteria for major depression. Similarly, Kessler et al. (1995) reported that 47.9% of men and 48.5% of women with PTSD in their sample also suffered from major depression. These rates reflect estimates of comordity following any type of trauma, including, but not limited to physical assault.

Relatively few studies have investigated the relationship between depression and physical assault in isolation. However, data indicate that severity of depression produced by physical assault approximates that produced by rape. Riggs et al. (1992) compared SCL-90-R depression subscale scores of victims of domestic violence and nondomestic violence with nonvictimized women. Women raped or assaulted by their husbands had SCL-90-R depression scores of 0.66 and 0.61, respectively, while women raped or assaulted by strangers had scores of 0.72 and 0.87, respectively. By contrast, nonvictimized women scored 0.47 on the depression subscale.

The National Womens Study (NWS; Kilpatrick, Edmunds, & Seymoor, 1992) also examined effects of physical assault on affect. The NWS was a three-wave longitudinal research project in which a national household probability sample of adult women ($N = 4009$) was interviewed using random digit dialing methodology. Of these women, 3006 also completed Wave III interviews two years later. Interview content included queries about lifetime and recent victimization, post-traumatic stress disorder, depression, and substance use. Variables relevant to physical violence included lifetime and recent (past two years) physical assault, lifetime and recent (past month) depression, and lifetime and recent (past year) substance use/abuse. The following analyses involving depression and assault were conducted: lifetime depression predicted by lifetime physical assault status, current depression predicted by lifetime physical assault status, and current depression predicted by recent assault status. Women with life histories of physical assault were 3.31 times as likely to demonstrate life histories of depression. Similarly, past occurrence of assault increased odds of current depression by a factor of 2.9. Most importantly, recently assaulted women were 5.5 times as likely to present with depression as women who were not recently assaulted, with 12.6% of assaulted vs. 2.6% of nonassaulted women meeting criteria for active depression.

9.19.3.5 Trauma Sequelae: Substance Abuse

Recent investigations also reveal high comorbidity between assault-related PTSD and substance use disorders. Breslau et al. (1991) demonstrated that approximately 45% of participants with PTSD also met criteria for substance use or abuse, with 31% meeting criteria for alcohol abuse or dependence. Kessler et al. (1995) found that 35% of men and 27% of women in their sample diagnosed with PTSD experienced drug abuse or dependence, and over 50% of men and 28% of women with PTSD reported concurrent alcohol abuse or dependence. Following a violent assault, victims may increase substance use to cope with resultant symptoms of PTSD and depression. According to this conceptualization, extremely high levels of negative affect produced by assault cause individuals to engage in behaviors that rapidly reduce negative emotions (Beckham et al., 1995), such as situational escape and ingestion of alcohol or drugs. Objective support for this model of postassault drug and alcohol use has been provided by Levenson, Oyama, and Meek (1987), who demonstrated that alcohol ingestion successfully reduced both psychophysiological and subjective indices of negative affect during stress tasks in individuals at particular dispositional and familial risk of alcoholism.

Despite theory-based evidence that victimization facilitates substance abuse, directionality of the assault/substance abuse relationship remains unclear. Several investigators (Breslau et al., 1991; Cottler, Compton, Mager, Spitznagel, & Janca, 1992; Scribner, Mackinnon, & Dwyer, 1995) have reported an increased risk of trauma given a proximate association with substance abuse. Indeed, victims use substances at significantly elevated rates relative to nonvictims and this substance use appears to facilitate additional violence (Cottler et al., 1992; Kulka et al., 1990). For example, Buss et al. (1995) determined, on the basis of self-report and toxicology screens, that 69.7% of seriously assaulted individuals seeking hospital treatment were under the influence of drugs or alcohol at the time of their attack, and 60% reported that their attackers were drinking or were on drugs.

Data illustrating origins of the substance abuse/assault cycle remain sparse and are derived primarily from long-term retrospective reports. However, prospective data from the NWS demonstrated that rape and physical assault may lead to substance use and abuse in previously nonusing women (Kilpatrick, Acierno, Resnick, Saunders, & Best, 1997). That is, likelihood of progression to substance use/abuse following assault in previously nonusing women was double that of nonassaulted women. Thus, the position that substance use follows physical and sexual assault (in addition to preceding victimization) is strengthened.

As is evident from the above reviewed studies, relatively more specific information is available on risk of substance use following sexual assault than physical assault. The NWS (Kilpatrick et al., 1992) data indicated that, compared to noncrime victims, victims of rape were 13.4 times as likely to have two or more major alcohol-related problems and 26 times more likely to have two or more serious problems related to drug abuse. In addition, rape victims were 3.4 times as likely to use marijuana, 6 times more likely to have used cocaine, and 10.1 times as likely to use hard drugs other than cocaine. Given similarities in outcome between sexual and physical victimization outlined above, analogous findings would be expected in victims of physical assault. Data from the NWS demonstrate that this is precisely the case. Specifically, logistic regression was employed to estimate associated risk of use and abuse of substances following lifetime and recent physical assault. Odds ratios and prevalence rates are given in Table 1. As was the case with rape, physical assault greatly increased the risk of using all substances. Specifically, likelihood of active cigarette use was doubled and risks of active alcohol abuse and drug use were more than tripled in women who had been assaulted in the previous two years. Increased risk was also observed when considering lifetime history of physical assault.

9.19.3.6 Trauma Sequelae: Panic

In addition to PTSD, depression, and substance use, assault appears to facilitate panic in some women. This is not surprising when one considers that the symptom overlap between PTSD and panic is great, and includes palpitations, chest pain, dizziness, shortness of breath, nausea, derealization, sweating, trembling, and numbness (Falsetti, Resnick, Dansky, Lydiard, & Kilpatrick, 1995). The etiological role of stressful events (e.g., trauma) in the development of panic is consistent with the two-factor theory of PTSD and depression outlined above. As applied to panic, initial anxiety attacks are conceptualized as appropriate and logical responses to a life-threatening situation. These responses then become associated with salient stimuli that serve to trigger additional panic in the future (Barlow, 1988). Some empirical evidence for this relationship has been provided by Falsetti et al. who reported that 94.4% of 391 women meeting criteria for panic disorder had

Table 1 Odds ratios and prevalence rates of lifetime and current substance abuse, predicted by lifetime and current assault status.[a]

Predicted variable	*Predictor variable*	*Rate if positive for predictor variable* (%)	*Rate if negative for predictor variable* (%)	*Odds ratios and CIs*
Lifetime cigarette use	Lifetime assault	51.8	41.7	1.5 (1.2–1.8)
Lifetime alcohol abuse	Lifetime assault	28.3	13.9	2.4 (1.9–3.0)
Lifetime marijuana use	Lifetime assault	44.3	20.0	3.2 (2.6–3.9)
Lifetime hard drug use	Lifetime assault	15.1	5.1	3.3 (2.5–4.5)
Current cigarette use	Lifetime assault	30.2	18.9	1.9 (1.5–2.3)
Current alcohol abuse	Lifetime assault	12.7	5.8	2.4 (1.7–3.2)
Current marijuana use	Lifetime assault	14.3	4.9	3.3 (2.4–4.4)
Current hard drug use	Lifetime assault	4.3	0.9	4.8 (2.7–8.4)
Current cigarette use	Recent assault	37.0	19.0	2.4 (1.6–3.5)
Current alcohol abuse	Recent assault	15.1	2.8	6.2 (3.7–10.7)
Current marijuana use	Recent assault	17.6	4.0	5.1 (3.1–8.4)
Current hard drug use	Recent assault	2.5	0.8	3.1 (0.9–10.3)

[a] Lower-limit confidence intervals greater than 1 indicate significance at $p < 0.05$.

positive victimization histories. In addition, Kessler et al. (1995) reported that odds of lifetime comorbid panic disorder in female respondents with PTSD were increased by a factor of three, with 12.6% of PTSD women also meeting criteria for panic, compared to 4.3% of non-PTSD women. Relatedly, Falsetti and Resnick (1997) noted that 68.9% of a clinical sample of 62 participants seeking treatment for trauma-related problems experienced four or more symptoms of panic attacks, and 67.6% of those diagnosed with PTSD reported suffering from panic. Breslau and Davis (1987) have also demonstrated associations between stressor events and panic. Stress-related panic is particularly prevalent in instances where trauma is life-threatening, extremely uncontrollable, and results in significant reductions in self-esteem (Roy-Byrne, Geraci, & Uhde, 1986). Although physical assault meets these criteria, no data are currently available describing the specific relationship between physical violence and panic.

9.19.4 ASSESSMENT

Two distinct areas of measurement are relevant to the study of interpersonal violence: first, assessment of the trauma itself, and second, assessment of post-traumatic psychological functioning. Of course, accurate assessment of the latter presupposes adequate assessment of the former. Examination of studies of interpersonal violence reveals, however, that such methodological adequacy is not always achieved. Early investigations employed flawed strategies to measure rates of physical and sexual assault, such as the use of legal terminology rather than behavioral description of events to determine assault status. In addition, these studies imposed unidimensional structure on inherently multidimensional constructs, with relatively little consideration given to interaction effects of multiple vs. isolated traumatic events, occurring across the lifespan vs. only in adulthood, and committed by known vs. unknown (or multiple vs. single, or peer vs. nonpeer) perpetrators (Kilpatrick et al., in press). Fortunately, advances in measurement of both trauma and post-traumatic psychopathology have been made. As Resnick, Falsetti, Kilpatrick, and Freedy (1996) noted, commonalities of sensitive, reliable, and valid trauma event assessment include: (i) appropriate prefacing statements, (ii) use of behaviorally specific queries, (iii) questions and analyses that consider multiple trauma parameters and relate multiple traumatic events to symptoms, and (iv) queries about qualitative/subjective aspects of trauma.

9.19.4.1 Trauma Assessment

Traumatic events are, by definition, associated with extremely aversive cognitions and affective states. Therefore it is important, both experimentally and ethically, to create an environment in which self-disclosure is met with acceptance, empathy, knowledge, and encouragement. Clearly, victims must feel that their participation in assessment interviews for clinical or research purposes will be worth their effort and discomfort if they are to engage in frank and open discussion of exceedingly personal and often humiliating material. Simply asking an individual if he or she has ever been raped or physically assaulted in the absence of appropriate preface statements does little to convey concern for respondents' well-being, or remove psychological defenses employed to protectively minimize personal relevance or recall of the assault. Rather, a preface statement should provide a simple context within which interviewees can comfortably respond. Resnick et al. (1996) suggested brief prefacing statements that demonstrate concern on the part of interviewers and understanding of the nonrare nature of *DSM-IV* PTSD Criterion A traumatic events.

The second aspect of sensitive and accurate trauma assessment is perhaps the most important. Use of behaviorally specific closed-ended questions increases likelihood that traumatic events will be both reported and operationally definable. This is because behavior-based queries are relatively less affected by cultural and individual differences of respondents. In addition, it may also be less aversive for some victims to answer yes or no to questions about particular behavioral events than to generate descriptive information about their assault. As such, sources of error produced by personal definitions of assault or rape or abuse, or respondent inability or hesitancy to disclose information relating to victimization, are minimized.

The third factor in comprehensive trauma assessment involves measurement of multiple events in context, rather than single events in isolation. Psychological responses and treatment needs of individuals who have been previously assaulted may differ significantly from those who have not experienced such victimization. This point is particularly relevant when considering that approximately 50% of physically and sexually assaulted individuals have prior victimization histories (e.g., Kilpatrick, 1990). Therefore, behaviorally specific close ended questions and postassessment case conceptualization should address childhood physical and sexual assault status; adult physical and sexual assault status; relationship of victim to perpetrator during all assaults; and frequency, intensity, and type of assaults across the lifespan. In addition to assault parameters, it is also essential to assess other contextual factors that influence post-traumatic outcome, including familial and personal history of psychopathology, social and vocational adjustment, and level of social support.

The powerful effects exerted by several qualitative aspects of trauma on post-assault psychological functioning also argue for their inclusion in any comprehensive trauma assessment. Subjective aspects of assault that have proven particularly predictive of post-traumatic outcome include degree of perceived life threat, level of incurred injury, intensity of distress during attack, perceived level of force employed by perpetrator, and level of dissociation experienced. Inclusion of these factors in assessment reduces heterogeneity within broad assault typologies and increases the precision with which intervention techniques can be employed. For example, repeated assault with a handgun by a spouse (objective trauma characteristics) involving high perceived force, high perceived life threat, and intense distress (qualitative or subjective trauma characteristics) expectedly produces different results and requires different treatment components than (objectively) identical assault in which perceived force, life threat, and distress are relatively low.

A final, controversial area of trauma assessment that may prove extremely informative to researchers, clinicians, and diagnosticians alike is victim behavior related to risk of assault. While no individual is responsible for being attacked by another person, several specific behaviors, such as substance use and abuse, have been shown to increase likelihood of victimization (e.g., Kessler et al., 1995; Kilpatrick et al, 1997). It is not clear whether victims' perceptions of responsibility are associated with relatively improved or impaired functioning; however, arguments can be made for both outcomes. That is, individuals who think that their behavior contributed to their victimization may feel empowered to change, or to exert relatively greater control over their environment. By contrast, victims might instead experience intense guilt and self-flagellation for allowing themselves to enter into risky situations.

Assessment of assault-related traumatic events is currently achieved, with varying degrees of success, through standardized questionnaires and structured clinical interviews. As Resnick et al. (1996) noted, inquiries that produce dichotomous data regarding event occurrence/nonoccurrence should be complemented by questions that generate qualitative data regarding event characteristics. In this manner, multivariate relationships between trauma factors can be assessed and targeted by interventions. The next subsection of the chapter is devoted to brief description of some well-known trauma assessment measures that yield both dichotomous and qualitative information, and have at least limited empirical evaluation; later subsections are devoted to assessment of post-traumatic psychopathology.

9.19.4.1.1 Trauma assessment: interviews

The PTSD module of the Structured Clinical Interview for *DSM-III-R* (SCID-R) and the newer version for *DSM-IV* (SCID-IV) contains a very brief preface describing possible PTSD Criterion A events, followed by a short trauma priming list. No behavioral definitions of assault are provided, and respondents are required to generate descriptions of any experienced trauma that are not on the list. Although the interviewer documents all traumatic events, only the subjectively defined "worst" trauma is considered in the diagnostic section. No attempt is made formally to identify or evaluate subjective aspects of trauma that relate to outcome, such as perceived risk of severe injury or death. Similarly, no specific prompts are provided to elicit relevant objective event characteristics, such as number of assailants, duration of assault, or presence of a weapon. Overall, lack of behaviorally defined close-ended assault questions; weaknesses of the preface statement, event documentation, and prompting; and the extent to which objective and subjective parameters of assault are overlooked lessen the likelihood that SCID-based assessments of assault trauma will be of sufficient sensitivity to detect all instances of interpersonal aggression.

The Diagnostic Interview Schedule-III-Revised (DIS-III-R, Robins, Helzer, Cottler, & Goldring, 1988) is a structured clinical interview similar in emphasis and format to the SCID. Specifically, a brief preface statement is followed by a listing of traumatic events intended to prime respondents to describe their own trauma history. As with the SCID, no behavioral descriptions of assault events are provided. In addition to the index question, respondents are prompted to report up to two additional "terrible" or "shocking" experiences they have had. In contrast to the SCID, the DIS-R qualitatively assesses whether injury resulted in response to the traumatic event, but does not specifically measure other objective or subjective aspects of assault.

Both the SCID and the DIS-R emphasize assessment of PTSD symptomatology over that of trauma *per se*. While achieving acceptable diagnostic reliability, both interviews fail to employ sufficiently sensitive or valid techniques to measure trauma. This is somewhat ironic and distressing in that no diagnosis of PTSD is possible without preliminary positive identification of a traumatic event. Hence, in a diagnostic sense, these interviews put the cart before the horse. This is particularly problematic when considering that the overwhelming majority of treatment outcome and epidemiological studies, in addition to hundreds of clinic practices, employ these structured interviews to assess psychopathology. An effective solution to this sensitivity problem involves pairing these symptom-focused interviews with a trauma-focused assessment instrument. Such a hybrid has been developed by Kilpatrick et al. (1989) for the NWS.

The NWS Event History-PTSD Module represents the synthesis of several trauma assessment devices refined through years of epidemiological research on physical and sexual assault by the National Crime Victims Research and Treatment Center at the Medical University of South Carolina. This interview comprehensively assesses lifetime occurrence of civilian crime, and includes a preface statement that provides contextual orientation to the trauma victim, along with accurate information regarding criterion A event prevalence. Moreover, the preface statement details the interviewer's interest in any assault event, not just those perpetrated by strangers or reported to police or individuals in a position of authority.

In addition to physical assault, the NWS Event History PTSD Module assesses other typically stressful life events (e.g., automobile accidents, natural disasters). Past-year occurrence of common life stressors (e.g., marital discord, financial difficulty) are addressed in another section of the interview. Sensitivity is further enhanced and popular stereotypical definitions of trauma terminology avoided through use of behaviorally specific closed-ended questions about physical and sexual assault. Both objective and subjective trauma-related data are obtained, including number of lifetime interpersonal violence episodes, age at first incident of each trauma type, relationship to perpetrator, personal and perpetrator use of substances preceding attack, most serious incident of each trauma type, and perceptions of imminent injury or death from the attack. Although all victimization events are addressed, qualitative information about physical assault history is gathered for only one event (as opposed to rape, for which information is gathered in response to the first, most recent, and "worst" victimization). Further, the measure is geared toward severe assault and may produce overly conservative estimates of some forms of domestic violence. Convergent validity of the instrument has been established, with estimated rates of sexual and physical assault approximating those of other studies (e.g., Norris, 1992).

9.19.4.2 Post-trauma Assessment of PTSD: Interviews

The intended focus of the SCID and the DIS-R is on symptom rather than trauma assessment. Not surprisingly, therefore, stability of diagnostic conclusions based on these instruments is quite high. The SCID-IV PTSD module provides specific prompts and follow-up inquiries, intended to be read verbatim to respondents, to assess presence or absence of each *DSM-IV* PTSD symptom. Symptom presence is rated on a three-point confidence scale based on the interviewer's interpretation of the victim's responses to prompts. Respondents are asked to frame symptoms in terms of their "worst" trauma experience. As with other modules of the SCID, diagnostic cutoff criteria are readily available on scoring forms, along with decision trees and skip-out instructions. Although specific instructions for probe questions and guidelines for response interpretation are provided, the instrument is intended for use by clinicians and highly trained interviewers. While psychometric qualities of the SCID PTSD diagnosis following physical assault have not been investigated, reliability and validity of the PTSD diagnosis in general has been established. Kappas for SCID PTSD range from 0.68 to 0.93 (Kulka et al., 1990; Skre, Onstad, Torgersen, & Kringlen, 1991). In addition, Kulka et al. demonstrated convergent validity with the IES (Horowitz, Wilner, & Alvarez, 1979) and the Mississippi Scale (Keane, Caddell, & Taylor, 1988).

In contrast to the SCID-IV, the DIS-R was specifically designed for use by nonprofessional interviewers. The DIS-R also differs from the SCID in that it allows the PTSD diagnosis to be made with reference to three specific Criterion A events, whereas the SCID considers only the "worst" event. Diagnosis is assigned on the basis of responses to specific questions relating to each *DSM-III-R* defined PTSD symptom. No behavioral anchors are provided to establish interrespondent consistency of symptom severity. Psychometric properties of the DIS PTSD diagnosis are mixed. Initial sensitivity and specificity ratings were quite good (0.87 and 0.73, respectively), as was overall kappa agreement with criterion diagnosis (0.64). However, with a large-scale community sample, sensitivity of the instrument was unacceptably low (0.22) as was diagnostic agreement (kappa = 0.26) (Kulka et al., 1990). Low DIS sensitivity may be a function of its weak trauma screen. For this reason, Resnick et al. (1996) recommended that both the SCID and the DIS PTSD modules be preceded by a sensitive, behaviorally specific trauma screen.

In contrast to the broad-band focuses of the SCID and the DIS, the Clinician Administered PTSD Scale-1 (CAPS-1, Blake et al., 1990) was specifically designed to yield both continuous (i.e., severity) and dichotomous (i.e., diagnostic) information exclusively for PTSD. The CAPS-1 is intended for use by professionals and trained paraprofessionals, and addresses the 17 symptoms of PTSD that comprise the *DSM-IV* diagnosis. Thirteen additional questions target associated symptoms not part of the formal PTSD diagnosis, as well as functional impact of symptoms on social and vocational spheres. Separate ratings of frequency and intensity of each symptom are made on five-point (0 to 4) scales that employ specific behavioral reference anchors to maximize standardization and consistency of interviews. Both past-month and lifetime presence of the disorder are measured. The CAPS-1 has excellent psychometric properties. Blake et al. (1990) reported that test–retest diagnostic reliability ranged from 0.90 to 0.98, and internal consistency was 0.94. Using the SCID as criterion reference, the CAPS-1 achieved acceptable sensitivity (0.84), high specificity (0.95), and good agreement (kappa = 0.78) (Blake et al., 1990). Convergent validity was also demonstrated by a strong correlation between the CAPS-1 total severity score and scores on the Mississippi Scale for combat related PTSD ($r = 0.91$), and the MMPI PK scale ($r = 0.77$). Importantly, the CAPS-1 does not address Criterion A trauma events, and must be accompanied by a trauma screen.

9.19.4.3 Post-trauma Assessment of PTSD: Questionnaires

In addition to structured clinical interviews, several PTSD rating scales enjoy widespread use. Among the most frequently employed are the PTSD Symptom Scale Self-report version (PSS-SR, Foa, Riggs, Dancu, & Rothbaum, 1993), the Impact of Event Scale (IES, Horowitz et al., 1979), and the Trauma Symptom Inventory (TSI, Briere, 1995). The PSS-SR is the self-report version of Foa's structured interview for PTSD by the same name. The 17 items on this scale are identical in content to the interview, but contain simplified wording, and directly correspond to *DSM-IV* symptoms. As such, this self-report scale permits diagnosis of the disorder. Diagnoses based on the PSS-SR are slightly more conservative than those based on its interview counterpart, the PSS-I.

In contrast to the SCID and the DIS, the PSS-SR is intended for use with individuals who have a known assault history, and should thus be accompanied by a trauma screen when assessing individuals for whom basic background information is lacking. For all items, symptom frequency over the preceding two weeks is reported on a four-point scale. A total score is obtained by summing each symptom rating. Subscale scores are calculated by summing symptoms in the re-experiencing (four items), avoidance (seven items), and arousal (six items) clusters. Foa et al. (1993) evaluated the psychometric properties of the PSS-SR with subsets of 46 female recent rape victims and 72 female recent nonsexual assault victims. They obtained a total score alpha of 0.91 and subscale alphas (for re-experiencing, avoidance, and arousal) ranging from 0.78 to 0.82. In addition, one-month test–retest reliability for the total score was 0.74, while test–retest reliability for subscales ranged from 0.56 to 0.71. Convergent validity of the PSS-SR, with the IES and State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lusbene, 1970) was also demonstrated, with correlations ranging from 0.52 to 0.81. Moreover, using the SCID as the criterion reference, PSS-SR diagnostic sensitivity was 62%, specificity was 100%, positive predictive power equaled 100%, negative predictive power was 82%, and the total hit rate was 86%. Overall, the PSS-SR is an easily administered, psychometrically sound measure of PTSD symptom severity and diagnostic status, and is thus an appropriate tool for both clinical and empirical endeavors.

The IES (Horowitz et al., 1979) has a longer history of use than the PSS-SR, but provides relatively less diagnostic specification. The IES is a 15 item self-report measure constructed to assesses intrusion and avoidance symptoms of PTSD experienced over the past week. Symptoms are phrased in the first person and respondents rate their frequency on a four-point scale. The scale focuses on cognitive PTSD symptomatology (as opposed to somatic arousal), and is an excellent indicator of subjective distress. Moreover, the IES is sensitive to effects of treatment (Kilpatrick & Amick, 1985). Although the IES does not permit diagnosis of PTSD along *DMS-IV*-defined criteria, severity cutoff scores have been used to predict diagnostic status. Psychometric evaluation of the scale's reliability by Horowitz et al. produced alphas of 0.78 for the intrusion subscale and 0.80 for the avoidance subscale. Split-half reliability of the total scale was 0.86, and one-week test–retest reliability equaled 0.87.

A somewhat more comprehensive scale than the IES, the TSI (Briere, 1995) is an expanded version of the Trauma Symptom Checklist (Briere & Runtz, 1989). In addition to core symptoms of PTSD, this 100-item questionnaire taps areas of impaired functioning that are not subsumed under the *DSM-IV* definition of PTSD, but are nonetheless related to victimization. Specifically, the TSI is divided into 10 subscales assessing anxious arousal, depression, anger/irritability, intrusive experience, defensive avoidance, dissociation, sexual concerns, dysfunctional sexual behavior, impaired self-reference, tension-reduction behavior. Symptoms are listed in brief descriptive phrases. Respondents rate the level at which they experienced each symptom over the previous six months on a four-point scale. The TSI total score can be used as an index of overall PTSD severity, and total scores in each subscale can be employed to illustrate focal severity of symptoms. The scale also contains 12 "critical items" that serve as flags for particularly severe psychopathology.

Psychometric properties of the TSI have been determined with both clinical and general community samples. In nonclinical samples, reliability for the 10 clinical scales was quite high (average alpha = 0.85) and the hit rate for PTSD diagnostic status was 91% (Briere, 1995). Similar levels of reliability among the 10 scales were observed in clinical samples (alpha range across scales: 0.74 to 0.90) (Briere, Elliott, Harris, & Cotman, 1995). Construct validity of the TSI was evident from higher scores for individuals with positive victimization histories, even when controlling for demographic factors. The instrument also includes three validity scales (atypical response, response level, and inconsistent response), for which clinical psychometric data are not currently available. In light of its ease of administration and completion, and its excellent psychometric properties, the TSI is an appropriate tool for clinical and empirical realms.

9.19.4.4 Post-trauma Assessment of Related Symptomatology: Interviews

Depression, substance abuse, and panic should also be assessed in crime victims. However, these diagnoses are typically considered secondary to PTSD in traumatized individuals, and relatively less coverage of measurement indices is offered here. It is useful, in both an assessment and treatment sense, to frame additional psychopathology in recipients of assault in the context of the traumatic event. For example, if assessment reveals that a victim abuses alcohol in order to reduce intrusive ideation, then standard substance abuse treatment must be complemented by strategies to develop alternative coping techniques. In this example, a diagnosis of alcohol abuse provides only limited information for treatment planning. By contrast, placing substance use in context greatly increases the treatment utility (McKnight, Nelson, Hayes, & Jarrett, 1984) of assessment. Similarly, if a woman's depression is related to cessation of all interpersonal activity after dusk because darkness has been associated with strong negative affect, strategies to increase evening activity should be accompanied by techniques to extinguish the anxiety response.

In addition to PTSD, the SCID-IV contains modules to assess major depressive disorder, dysthymia, substance use disorders, and panic disorder. The structure of SCID modules for these diagnoses is identical to that for PTSD. As with PTSD, diagnostic questions parallel *DSM-IV* classification criteria, and each criterion set of symptoms for each disorder is rated on a three-point confidence scale from "absent" to "present." Both recent and lifetime occurrence of disorders are assessed; however, data are exclusively dichotomous. Although psychometric properties of the SCID-IV will be forthcoming, reliability and validity ratings of the previous version (SCID-III-R) have been reported by several investigators. For major depression, reliability kappas range from 0.69 to 0.93 (Riskind, Beck, Berchick, Brown, & Steer, 1987; Skre et al, 1991; Williams et al., 1992). For alcohol abuse or dependence, reliability ratings ranged from 0.75 to 0.96, and for substances other than alcohol, reliability ratings ranged from 0.73 to 0.89 (Malow, West, Williams, & Sutker, 1989; Skre et al., 1991; Williams et al., 1992). Finally, for panic disorder, reliability ratings ranged from 0.58 to 0.88 (Skre et al., 1991; Williams et al., 1992). Validity of SCID-III-R diagnoses is evident from their frequent role as criterion reference in demonstrations of convergent validity with newly developed instruments.

The Addiction Severity Index (ASI, McLellan, Parikh, & Bragg, 1990) is a standardized clinical interview with an embedded rating scale that provides a continuous measure of the functional impact of substance use. In addition to providing data regarding substance use in the previous 30 days, the ASI illustrates functional impact and use severity in seven specific areas, comprising drug use, alcohol use, family-social functioning, medical functioning, psychological functioning, employment, and legal involvement. Concurrent and interrater reliabilities are impressive (0.74–0.93) and validity is acceptable (Kosten, Rounsaville, & Kleber, 1983; McLellan, Luborsky, & Woody, 1983).

9.19.4.5 Post-trauma Assessment of Related Symptomatology: Questionnaires

Dichotomous measurements of post-traumatic psychological sequelae should be complemented by continuous data to determine symptom severity and describe diagnostic subtypes. The most widely used measure of depressive severity is the Beck Depression Inventory (BDI). (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). Each of the 21 items on this scale contains four statements reflecting current manifestations of depression in increasing intensity, from neutral to severe. Scoring of items is on a 0–3 scale and total scores range from 0 to 63, with higher scores indicating greater depressive severity. Of the 21 items on the scale, 13 assess depressive symptoms that are primarily psychological in nature, while eight measure symptoms that are somatically oriented. Beck and Steer (1984) and Gallagher, Nies, and Thompson (1982) demonstrated that the BDI has high internal consistency ($\alpha = 0.86$ and $\alpha = 0.91$, respectively). Moreover, Olin, Schneider, Eaton, Zemansky, and Pollack (1992) found excellent concurrent validity between the BDI and the GDS ($r = 91$).

There are no standardized measures to assess panic. However, panic diaries, in which ratings of symptom type, frequency, and intensity are recorded daily and summed weekly or biweekly, are standard dependent measures in panic disorder research. These may be effectively complemented by the Beck Anxiety Inventory (BAI; Beck & Steer, 1991) to provide both standardized and descriptive information about somatic responses in crime victims. The BAI is a 21-item self-rating scale of anxiety symptomatology. Patients are asked to rate on a four-point scale the degree to which specific symptoms. Specific symptom clusters have been identified by Beck and Steer (1991) reflecting neurophysiological, subjective (cognitive), panic, and autonomic dimensions. Beck and Steer (1991), Fydrich, Dowdall, and Chambless (1992) and Steer, Ranieri, Beck, and Clark (1993) demonstrated the instrument's concurrent validity with the Hamilton Anxiety Rating Scale. Internal consistency of the BAI has been illustrated by Steer et al. (1993) and Fydrich et al. (1992).

9.19.5 TREATMENT

Comprehensive description of treatments available for all postassault psychopathology is beyond the scope of this chapter. However, existing empirically validated treatments for PTSD will be reviewed below.

9.19.5.1 PTSD

Exposure to fear-producing cues in a relatively safe environment is a central component of all effective treatments for PTSD (Keane, Gerardi, Quinn, & Litz, 1992). Techniques used in addition to exposure, such as muscle relaxation, serve a primarily palliative function in that they prevent anxiety from exceeding tolerable levels during exposure exercises.

Veronen and Kilpatrick (1983) and Foa et al. (1991) evaluated one of the earliest successful psychological treatments for PTSD in assault victims. This treatment is well tolerated and has received strong empirical support. The overall technique is in large part an adaption of Meichenbaum's Stress Inoculation Training (Meichenbaum & Jaremko, 1983) in which several behavioral and cognitive-behavioral procedures are combined to reduce or eliminate PTSD symptomatology. Primary components include: applied relaxation training, breathing retraining, thought stopping, stress inoculation training through imaginal exposure and guided self-dialogue, *in vivo* exposure assignments, and cognitive restructuring. These strategies are preceded by an education module, in which patients are familiarized with the conditioning model of fear acquisition and its appropriate treatment.

Progressive Muscle Relaxation (PMR, Jacobson, 1938), is most frequently taught to patients according to the standard procedure outlined by Bernstein and Borkovec (1973) and Wolpe (1990). In this procedure, patients are trained to detect and replace muscle tension with relaxation. This is accomplished through systematic constriction and release of several specific sets of muscles throughout the body. Typically, therapists will first model the procedure, tensing an isolated group of muscles for about seven seconds, then relaxing that group of muscles for 15 to 20 seconds. Patients are then instructed to tense their muscles in a manner similar to that demonstrated by the therapist, followed by complete relaxation of those muscles. During tension and release cycles, the patient's attention is repeatedly directed toward the qualitatively distinct muscular sensation associated with each phase of the procedure. Once patients achieve proficiency in relaxing specific sets of muscles, more diffuse groupings are used, until finally patients are able to relax their entire body at will. In this manner, patients learn to recognize and eliminate somatic correlates of anxiety.

Diaphragmatic breathing and guided positive imagery are often employed to augment relaxation. In younger adults, relaxation training routinely involves four or five weekly therapist-directed sessions complemented by 30 minutes of daily self-directed practice. Importantly, lower "doses" of training appear to be ineffective (e.g., Borkovec & Sides, 1979). In the treatment of PTSD and other disorders characterized by anxiety upon exposure to specific fear-stimuli, simple relaxation exercises should be augmented by Applied Relaxation Training (Öst, 1988; Öst, Sternes, & Lindahl 1984), in which patients are taught not only to relax their muscles, but actively to apply this state of relaxation to those circumscribed situations and stimuli known to produce anxiety and fear. In order to accomplish this, patients first practice inducing a state of relaxation under a variety of conditions, such as waiting in a line, walking, sitting in an uncomfortable chair, talking to others, or any other situation in which they might experience the fear response. This active form of relaxation is then gradually applied, imaginally or in role-play format, to specific fear-producing situations associated with the traumatic event. Finally, applied relaxation during exposure to actual anxiety-producing stimuli is encouraged and practiced.

As mentioned, diaphragmatic breathing (Clark, Salkovskis, & Chalkley, 1985) frequently complements relaxation during exposure-based treatments and is employed to reduce respiration rate. It is particularly useful when treating victims of interpersonal violence, who are at greater risk of experiencing panic and its associated hyperventilatory symptoms. The procedure is relatively simple and involves replacement of shallow or "upper chest" breathing with deep diaphragmatic inhalation and exhalation. In addition to altering "location" of their breathing, patients are taught to monitor and pace the rate at which they inhale and exhale, so that exhalations last approximately five seconds and inhalations last about four seconds. This has the effect of minimizing hyperventilation, and, if the procedure has been extensively paired with progressive muscle relaxation, inducing a relaxed state. In SIT treatment of PTSD, the procedure is employed during exposure trials to limit exacerbation of somatic arousal.

Thought stopping "is a method of changing cognitive behavior that ... does not depend on changing the meaning of words or interpretation of situations" (Wolpe, 1990, p. 129). This procedure may aid in efforts to reduce intrusive ideation characteristic of PTSD by breaking cognitive chains of aversive thought through abrupt, patient-initiated incompatible behaviors. Interestingly, the technique is somewhat analogous to the behavioral procedure of response prevention/interruption used to eliminate compulsive behavior. In thought stopping, patients are trained to interrupt incipient thoughts as soon as they occur by exclaiming aloud "Stop!" or some other similar vocal command, followed immediately by the substitution of a competing benign thought for the intrusive cognition. As the occurrence of intrusive thoughts is diminished, self-commands are faded to the subvocal level. Thought stopping can be effective in reducing distress following exposure trials, but should not be used as a cognitive avoidance strategy to reduce stimulus salience during exposure. The relatively nonaversive nature of this technique, and the ease with which it is applied to repetitive anxiety-producing thoughts, justifies its inclusion in treatments for PTSD.

The primary component of stress inoculation treatment for PTSD is exposure. Exposure treatments have as their theoretical basis the process of extinction, or the diminution of the probability of occurrence of a conditioned response following repeated nonreinforced exposure to a conditioned stimulus. Exposure typically varies in terms of the following parameters: (i) medium of stimulus presentation (i.e., imaginal vs. *in vivo*), (ii) intensity of stimulus presentation (graded vs. full), (iii) purity of stimulus presentation (pure extinction vs. addition of coping strategies). In stress inoculation treatment for PTSD, steps are taken to maximize exposure while minimizing aversive arousal. Thus, exposure is imaginal, graded, and accompanied by the coping techniques outlined above.

The initial step in a graduated exposure treatment is construction of anxiety hierarchies. Anxiety hierarchies (Wolpe, 1990, p. 160) may involve internal (e.g., feelings of panic) or external (e.g., men who resemble the perpetrator) stimulus situations, or combinations of the two. Detailed and specific hierarchy items are preferable to general items (e.g., "being with unknown men") and facilitate relatively more powerful extinction effects by increasing the salience of anxiety-producing stimuli. Since this specificity may restrict therapeutic generalization, multiple hierarchies are regularly required to overcome all conditioned fear-producing stimuli in a particular patient. Spacing between successive hierarchy items should be even, and progression through the hierarchy tolerable. This spacing is accomplished by assigning a relative value of anxiety to each item via the Subjective Units of Discomfort Scale (SUDS; Wolpe, 1990), or some similar "anxiety thermometer." During hierarchy construction, patients are asked to generate the most anxiety-producing stimulus possible and assign it a SUDS value of 100. Complete relaxation is given a SUDS value of 0. As patient and therapist generate new hierarchy items, a relative SUDS value is assigned to each.

Following construction and rank ordering of hierarchies, patients engage in imaginal exposure to the least aversive item. Exposure trials may include audiotaped descriptions of the trauma event. In order to maximize the opportunity for extinction of conditioned fear (i.e., maximize duration of exposure trials), exposure is accompanied by cognitive and muscular (relaxation) coping techniques. While these coping responses may "dilute" extinction during the session by lessening stimulus salience, they increase the likelihood that patients will continue to engage in exposure within and across sessions by increasing perceptions of self-efficacy and control.

Once imaginal exposure to anxiety hierarchy items can be endured with only minimal distress, *in vivo* exposure should be initiated when possible and safe. *In vivo* exposure is used in addition to imaginal exposure because the level of stimulus salience achieved by the former typically exceeds that of the latter. Hence *in vivo* exposure results in more "efficient" extinction of conditioned anxiety responses. As with imaginal exposure, *in vivo* exposure is performed in a graded format, with introduction to mildly anxiety-producing stimuli on the hierarchy preceding exposure to stimuli that elicit great fear. The degree to which stimulus intensity is varied (i.e., highly graded to full exposure) is determined, in large part, by individual characteristics of each patient. Although high "doses" of stimulus exposure (and hence anxiety) typically produce relatively more rapid treatment response than lower doses, high-intensity exposure is also associated with greater attrition (Barlow, 1988). Thus, exposure should be graded so as to produce significant, but not overwhelming levels of anxiety. As with imaginal exposure, applied relaxation, and diaphragmatic breathing are employed during *in vivo* trials to reduce levels of experienced anxiety during progression through the hierarchy. Anxiety during exposure may also be graded by the proximity of a significant other or therapist. This form of graduation routinely incorporates positive reinforcement (provided by significant other or therapist) into the extinction process and increases the likelihood that exposure exercises will be practiced between sessions. Additional modulation of exposure is achieved simply by altering the duration and frequency of extinction trials.

Complementing both imaginal and *in vivo* exposure is guided self-dialogue (Meichenbaum & Jaremko, 1983) in which each step in the behavioral chain of exposure is broken down and verbally rehearsed to successful completion immediately prior to the actual act. This cognitive rehearsal "sets the stage" for successful completion of the exposure trial. Self-dialogue is first modeled aloud by the therapist during simulated exposure to anxiety hierarchy items. In addition to analysis of required behaviors and rehearsal of successful completion of these acts, positive self-statements are made. Overtly verbal self-dialogue is practiced prior to exposure to hierarchy items. Once proficiency is achieved, patients, aided by therapist prompts, engage in exposure exercises, utilizing similar overtly verbal coping self-statements. After patients begin to demonstrate competency using self-statements, clinician prompts are faded. Finally, patients fade their verbal "self-prompts" as well, first to whispers, and then to completely covert self-talk. In this manner, patients adopt a coping style that enables them to endure relatively greater lengths of exposure with gradually increasing independence. Finally, cognitive interventions are employed to eliminate or alter cognitions that impede treatment progress.

Alternative treatments for PTSD exist, and have been subject, in varying degrees, to empirical evaluation. For example, Keane, Fairbank, Caddell, and Zimering (1989) developed and assessed a flooding intervention that employs highly intense imaginal exposure to reduce symptoms of PTSD in combat veterans. In contrast to the technique outlined above, little or no effort is made to reduce anxiety during extinction trials, and the overall intervention is quite aversive, albeit effective. In flooding, patients are instructed to imagine, in extremely vivid detail, the most aversive aspects of their anxiety-producing stimuli. Often, patient-narrated audiotapes depicting traumatic stimuli are employed to enhance visual imagery. Exposure trials typically last from 30 minutes to several hours. Although exceptionally high levels of anxiety are produced during these trials, repeated exposure results in a diminution or extinction of fear responses. This treatment may not be appropriate for all assault victims (e.g., children, older adults) and therapists and their patients should be cognizant of the fact that participants typically experience initial exacerbations, rather than reductions in anxiety. Obviously, high potential for attrition exists with this procedure, and treatment retention may be lower than that of the multicomponent treatment described above.

9.19.6 FUTURE DIRECTIONS

Progress has been made to develop and refine procedures used to identify and delineate trauma and its sequelae. Similarly, effective treatments have been generated to reduce the wide variety of symptomatology produced by victimization. However, relatively little is known about the effects of physical assault *per se*, particularly when considered in terms of possible permutations of trauma context. That is, we are unaware of the relative effects of assault on adults who were victims of repeated childhood physical assault, compared to adults with no childhood trauma history. Likewise, we have no data regarding differential response of assault victims who have or have not previously experienced nonassault-related trauma, or responses of physical assault victims who have or have not previously been sexually assaulted. Each of these combinations of variables also must be considered in terms of demographic factors (e.g., age, education, ethnicity) and perpetrator status (e.g., known or unknown). Although we have identified risk factors for psychopathology and variables that appear to buffer against it, we cannot predict, with any significant degree of accuracy, who will or will not experience significant post-traumatic emotional difficulties.

Therefore, future research should address the complexity of trauma history across the lifespan and in a variety of interpersonal contexts. Effects of physical assault must be examined in terms of childhood trauma history, adult trauma history, sexual trauma history, relation of perpetrator to victim, and victim age, education, gender, and ethnicity. Next, risk factors for development of postviolence psychopathology must be delineated for each of these possible assault permutations. Finally, research to facilitate effective prescriptive matching of intervention or interventions to trauma subtype must be conducted and disseminated.

9.19.7 SUMMARY

Interpersonal violence is considered a major public health problem in most countries of the world. Assessment of postassault psychopathology is clearly a complex, but not necessarily overwhelming task for the clinician or empirical investigator. Importantly, several core components should be part of any postassault assessment. These include a behaviorally based, sensitive trauma screen with appropriate and accurate preface statements, a psychometrically validated structured clinical interview, and standardized self-report measures. The preface statement should provide contextual reference to the respondent and maximize the likelihood that traumatic events will be reported; the trauma screen should produce consistent and operationally defined information regarding event type and its subjective and objective characteristics, the structured interview should produce reliable and valid diagnoses across a variety of diagnostic classes associated with victimization; finally, standardized questionnaires should yield descriptive information about the intensity, frequency, and quality of postassault psychopathology.

Adequate treatment of postassault psychopathology is contingent upon accurate, multidimensional assessment of symptomatology. Such assessment must address the broad scope of victims' emotional responses to trauma, including PTSD, depression, substance abuse, and panic. Likewise, individual interventions should be selected and implemented within the context of the traumatic event (e.g., if substances are used to cope with anxiety, removal of substance abuse should be accompanied by alternative, more adaptive stress-reduction strategies) in order to maximize treatment relevance.

9.19.8 REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington DC: Author.

Bachman, R. (1994). *Violence against women. A national crime victimization survey report*. Washington, DC: US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.

Barlow, D. H. (1988). *Anxiety and its disorders: The nature and treatment of anxiety and panic*. New York: Guilford.

Beck, A. T., & Steer, R. A. (1984). Internal consistencies of the original and revised Beck Depression Inventories. *Journal of Clinical Psychology, 40*, 1365–1367.

Beck, A. T., & Steer, R. A. (1991). Relationship between the Beck Anxiety Inventory and the Hamilton Anxiety Rating Scale with anxious outpatients. *Journal of Anxiety Disorders, 5*, 213–223.

Beck, A. T., Ward, C. H., Mendelsohn, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561–571.

Beckham, J. C., Rodman, A. A., Shipley, R. H., Hertzberg, M. A., Cunha, G. H., Kudler, H. S., Levin, E. D., Rose, J. E., & Fairbank, J. A. (1995). Smoking in Vietnam combat veterans with post-traumatic stress disorder. *Journal of Traumatic Stress, 8*, 461–472.

Bernstein, D. A., & Borkovec, T. D. (1973). *Progressive relaxation training*. Champaign, IL: Research Press.

Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Klauminzer, G., Charney, D. S., & Keane, T. M. (1990). A clinician rating scale for assessing current and lifetime PTSD: The CAPS-1. *The Behavior Therapist, 13*, 187–188.

Borkovec, T. D., & Sides, J. K. (1979). Critical procedural variables related to the physiological effects of progressive relaxation: A review. *Behaviour Research and Therapy, 17*, 119–125.

Breslau, N., & Davis, G. C. (1987). Posttraumatic stress disorder: The etiological specificity of wartime stressors. *American Journal of Psychiatry, 144*, 578–583.

Breslau, N., Davis, G. C., Andreski, P., & Petersen, E. (1991). Traumatic events and posttraumatic stress disorder in an urban population of young adults. *Archives of General Psychiatry, 48*, 216–222.

Briere, J. (1995). *Trauma Symptom Inventory professional manual*. Odessa, FL: Psychological Assessment Resources.

Briere, J., Elliott, D. M., Harris, K., & Cotman, A. (1995). Trauma symptom inventory: Psychometrics and association with childhood and adult victimization in clinical samples. *Journal of Interpersonal Violence, 10*, 387–401.

Briere, J., & Runtz, M. (1989). The Trauma Symptom Checklist (TSC-33): Early data on a new scale. *Journal of Interpersonal Violence, 4*, 151–163.

Burgess, A. W., & Holstrum, L. (1974). The rape trauma syndrome. *American Journal of Psychiatry, 131*, 981–986

Buss, T. F., Abdu, R., & Walker, J. R. (1995). Alcohol, drugs, and urban violence in a small city trauma center. *Journal of Substance Abuse Treatment, 12*, 75–83.

Clark, D., Salkovskis, P., & Chalkley, A. (1985). Respiratory control as a treatment for panic attacks. *Journal of Behavior Therapy and Experimental Psychiatry, 16*, 23–30.

Cottler, L. B., Compton, W. M., Mager, D., Spitznagel, E. L., & Janca, A. (1992). Posttraumatic stress disorder among substance users from the general population. *American Journal of Psychiatry, 149*, 664–670.

Davidson, L. M., Hughes, D., Blazer, D., & George, L. K. (1991). Post-traumatic stress disorder in the community: An epidemiological study. *Psychological Medicine, 21*, 713–721.

Davidson, L. M., Smith, R., & Kudler, H. (1989). Familial psychiatric illness in posttraumatic stress disorder. *Comprehensive Psychiatry, 30*, 338–345.

Fairbank, J. A., & Keane, T. M. (1982). Flooding for combat-related stress disorders: Assessment of anxiety reduction across traumatic memories. *Behavior Therapy, 13*, 499–510.

Falsetti, S.A., & Resnick, H. S. (1997). Frequency and severity of panic attack symptoms in a treatment seeking sample of trauma victims. *Journal of Traumatic Stress, 10*, 683–689.

Falsetti, S. A., Resnick, H. S., Dansky, B. S., Lydiard, R. B., & Kilpatrick, D. G. (1995). The relationship of stress to panic disorder: Cause or effect. In C. M. Mazure (Ed.), *Does stress cause psychiatric illness?* (pp. 111–147). Washington, DC: American Psychiatric Press.

Foa, E. B., Hearst-Ikeda, D., & Perry, K. J. (1995). Evaluation of a brief cognitive-behavioral program for the prevention of chronic PTSD in recent assault victims. *Journal of Consulting and Clinical Psychology, 63*, 948–955.

Foa, E. B., Riggs, D. S., Dancu, C. V., & Rothbaum, B. O. (1993). Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *Journal of Traumatic Stress, 6*, 459–473.

Foa, E. B., Rothbaum, R., Riggs, D., & Murdock, T. (1991). Treatment of PTSD in rape victims: A comparison between cognitive behavioral procedures and counseling. *Journal of Consulting and Clinical Psychology, 59*, 715–723.

Foa, E. B., Steketee, G., & Rothbaum, B. O. (1989). Behavioral/cognitive conceptualizations of post-traumatic stress disorder. *Behavior Therapy, 20*, 155–176.

Freedy, J. R., Resnick, H. S., Kilpatrick, D. G., Dansky, B. S., & Tidwell, R. P. (1994). The psychological adjustment of recent crime victims in the criminal justice system. *Journal of Interpersonal Violence, 9*, 450–468.

Freedy, J. R., Shaw, D. L., Jarrell, M. P., & Masters, C. R. (1992). Towards an understanding of the psychological impact of natural disasters: An application of the conservation resources stress model. *Journal of Traumatic Stress, 5*, 441–454.

Fydrich, T., Dowdall, D., & Chambless, D. L. (1992). Reliability and validity of the Beck Anxiety Inventory. *Journal of Anxiety Disorders, 6*, 55–61.

Gallagher, D., Nies, G., & Thompson, L. (1982). Reliability of the Beck Depression Inventory with older adults. *Journal of Consulting and Clinical Psychology, 50*, 152–153.

Green, B. L. (1990). Defining trauma: Terminology and generic stressor dimensions. *Journal of Applied and Social Psychology, 20*, 1632–1642.

Harlow, C. W. (1989). *Injuries from crime*. NCJ-116811. Washington, DC: US Department of Justice.

Helzer, J. E., Robins, L. N., & McEvoy, L. (1987). Posttraumatic stress disorder in the general population. *New*

England Journal of Medicine, *317,* 1630–1634.

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, *44,* 513–524.

Horowitz, M., Wilner, N., & Alvarez, W. (1979). Impact of Events Scale: A measure of subjective stress. *Psychosomatic Medicine, 41,* 209–218.

Jacobson, E. (1938). *Progressive relaxation.* Chicago: University of Chicago Press.

Keane, T. M. (1985). Defining traumatic stress: Some comments on the current terminological confusion. *Behavior Therapy, 16,* 419–423.

Keane, T. M., Caddell, J. M., & Taylor, K. L. (1988). Mississippi Scale for combat-related posttraumatic stress disorder. Three studies in reliability and validity. *Journal of Consulting and Clinical Psychology*, *56,* 85–90.

Keane, T. M., Fairbank, J. A., Caddell, J. M., & Zimering, R. T. (1989). Implosive (flooding) therapy reduces symptoms of PTSD in Vietnam combat veterans. *Behavior Therapy*, *20,* 245–260.

Keane, T. M., Gerardi, R. J., Quinn, S. J., & Litz, B. T. (1992). Behavioral treatment of post-traumatic stress disorder. In S. M. Turner, K. S. Calhoun, & H. E. Adams (Eds.), *Handbook of clinical behavior therapy* (pp. 87–97). New York: Wiley.

Keane, T. M., Zimering, R. T., & Caddell, J. M. (1985). A behavioral formulation of posttraumatic stress disorder in Vietnam veterans. *The Behavior Therapist*, *8,* 9–12.

Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C. B. (1995). Posttraumatic stress disorder in the National Comorbidity Survey. *Archives of General Psychiatry*, *52,* 1048–1060.

Kilpatrick, D. G. (1990, August). *Violence as a precursor of women's substance abuse: The rest of the drug-violence story.* Presented at Topical Mini-convention on Substance Abuse and Violence at the 98th Annual Convention of the American Psychological Association, Boston.

Kilpatrick, D. G., Acierno, R., Resnick, H., Saunders, B., & Best, C. (1997). A two-year longitudinal analysis of the relationships among assault and alcohol and drug abuse in women. *Journal of Consulting and Clinical Psychology*, *65,* 834–847.

Kilpatrick, D. G., & Amick, A. E. (1985). Rape trauma. In M. Hersen & C. G. Last (Eds.), *Behavior therapy casebook* (pp 86–103). New York: Springer.

Kilpatrick, D. G., Best, C. L., Veronen, L. J., Amick, A. E., Villeponteaux, L. A., & Ruff, G. A. (1985). Mental health correlates of criminal victimization: A random community survey. *Journal of Consulting and Clinical Psychology*, *53,* 866–873.

Kilpatrick, D. G., Edmunds, C. N., & Seymour, A. K. (1992). *Rape in America: A report to the nation.* Arlington, VA: National Victim Center & Medical University of South Carolina.

Kilpatrick, D. G., Resnick, H. S., Saunders, B. E., & Best, C. L. (1989). *The national women study PTSD module.* Charleston, SC: Crime Victims Research and Treatment Center, Department of Psychiatry, Medical University of South Carolina.

Kilpatrick, D. G., Resnick, H. S., Saunders, B. E., & Best, C. L. (in press). Rape, other violence against women, and posttraumatic stress disorder: Critical issues in assessing the adversity–stress–psychopathology relationship. In B. P. Dohrenwend (Ed.), *Adversity, stress and psychopathology*. Washington, DC: American Psychiatric Press.

Kilpatrick, D. G., Saunders, B. E., Amick-McMullan, A., Best, C. L., & Veronen, L. J., & Resnick, H. S. (1989). Victim and crime factors associated with the development of crime-related post-traumatic stress disorder. *Behavior Therapy, 20,* 199–214.

Kilpatrick, D. G., Saunders, B. E., Veronen, L. J., Best, C. L., & Von, J. M. (1987). Criminal victimization: Lifetime prevalence, reporting to police, and psychological impact. *Crime and Delinquency, 33*(4), 479–489.

Kilpatrick, D. G., Veronen, L. J., & Resick, P. A. (1979). The aftermath of rape: Recent empirical findings. *American Journal of Orthopsychiatry, 49*, 658–669.

Koop, C. E. (1992). Violence in America: A public health emergency. *Journal of the American Medical Association, 267*, 3075–3076.

Koss, M. P., Goodman, L. A., Browne, A., Fitzgerald, L. F., Keita, G. P., & Russo, N. F. (1994). *No safe haven: Male violence against women at home, at work, and in the community*. Washington, DC: American Psychological Association.

Kosten, T. A., Rounsaville, B. J., & Kleber, H. D. (1983). Concurrent validity of the Addiction Severity Index. *Journal of Nervous and Mental Disease, 171*, 606–610.

Kulka, R. A., Schlenger, W. E., Fairbank, J. A., Hough, R. L., Jordan, B. K., Marmar, C. R., & Weiss, D. S. (1990). *Trauma and the Vietnam War generation: Report of findings from the National Vietnam Veterans Readjustment Study*. New York: Brunner/Mazel.

Levenson, R. W., Oyama, O. N., & Meek, P. S. (1987). Greater reinforcement from alcohol for those at risk: Parental risk, personality risk, and sex. *Journal of Abnormal Psychology, 96*, 242–253.

Malow, R. M., West, J. A., Williams, J. G., & Sutker, P. B. (1989). Personality disorder classification and symptoms in cocaine and opioid addicts. *Journal of Consulting and Clinical Psychology, 57*, 765–767.

McKnight, D., Nelson, R. O., Hayes, S. C., & Jarrett, R. (1984). Importance of treating individually assessed response classes in the amelioration of depression. *Behavior Therapy, 15,* 315–335.

McLellan, A. T., Luborsky, L., Woody, G. E., (1983). Predicting response to alcohol and drug abuse treatments. *Archives of General Psychiatry, 40*, 620–625.

McLellan, A. T., Parikh, G., & Bragg, A. (1990). *Addiction Severity Index manual*, (5th ed). Philadelphia: University of Pennsylvania Center for the Study of Addiction.

Meichenbaum, D., & Jaremko, M. E. (1983). *Stress reduction and prevention*. New York: Plenum.

Mowrer, O. H. (1960). *Learning theory and the symbolic processes.* New York: Wiley.

Norris, F. H. (1992). Epidemiology of trauma: Frequency and impact of different potentially traumatic events on different demographic events. *Journal of Consulting and Clinical Psychology, 60*, 409–418.

Novello, A. C. (1992). A medical response to violence. *Journal of the American Medical Association, 267*, 3007.

Olin, J. T., Schneider, L. S., Eaton, E. M., Zemansky, M. F., & Pollack, (1992). The Geriatric Depression Scale and the Beck Depression Inventory as screening instruments in an older adult outpatient population. *Psychological Assessment, 4,* 190–192.

Öst, L. G. (1988). Applied relaxation vs. progressive relaxation in the treatment of panic disorder. *Behaviour Research and Therapy, 26,* 13–22.

Öst, L. G., Sterner, V., & Lindahl, J. L. (1984). Physiological responses in blood phobics. *Behaviour Research and Therapy, 22,* 109–117.

Resnick, H. S., Falsetti, S. A., Kilpatrick, D. G., & Freedy, J. R. (1996). Assessment of rape and other civilian trauma-related PTSD: Emphasis on assessment of potentially traumatic events. In T. W. Miller (Ed.), *Theory and assessment of stressful life events* (pp. 235–271). Madison, CT: International Universities Press.

Resnick, H. S., Kilpatrick, D. G., Dansky, B. S., Saunders, B. E., & Best, C. L. (1993). Prevalence of civilian trauma and PTSD in a representative national sample of women. *Journal of Consulting and Clinical Psychology, 61,* 984–991.

Riggs, D. S., Kilpatrick, D. G., & Resnick, H. S. (1992).

Long-term psychological distress associated with marital rape and aggravated assault: A comparison to other crime victims. *Journal of Family Violence, 7,* 283–296.

Riskind, J. H., Beck, A. T., Berchick, R. J., Brown, G., & Steer, R. A. (1987). Reliability of DSM-III-R diagnoses for major depression and generalized anxiety disorder using the Structured Clinical Interview for DSM-III-R. *Archives of General Psychology, 44,* 817–820.

Robins, L., Helzer, J., Cottler, L., & Goldring, E. (1988). *NIMH diagnostic interview schedule, version III, revised (DIS-III-R).* St Louis, MO. Washington University Press.

Rothbaum, B. O., Foa, E. B., Riggs, D. S., Murdock, T., & Walsh, W. (1992). A prospective examination of post-traumatic stress disorder in rape victims. *Journal of Traumatic Stress 5,* 455–475.

Roy-Byrne, P. P., Geraci, M., & Uhde, T. W. (1986). Life events and the onset of panic disorder. *American Journal of Psychiatry, 143,* 1424–1427.

Scribner, R. A., MacKinnon, D. P., & Dwyer, J. H. (1995). The risk of assaultive violence and alcohol availability in Los Angeles County. *American Journal of Public Health, 85,* 335–340.

Skre, S., Onstad, S., Torgersen, E., & Kringlen, E. (1991). High interrater reliability for the Structured Clinical Interview for DSM-III-R axis I (SCID-I). *Acta Psychiatrica Scandinavica, 84,* 167–173.

Spielberger, C. D., Gorsuch, R. L., & Lusbene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory (Self-Evaluation Questionnaire).* Palo Alto, CA: Consultant Psychologists Press.

Steer, R. A., Ranieri, W. F., Beck, A. T., & Clark, D. A. (1993). Further evidence for the validity of the Beck Anxiety Inventory with psychiatric outpatients. *Journal of Anxiety Disorders, 7,* 195–205.

Steinmetz, S. K. (1978). Violence between family members. *Marriage and Family Review, 1,* 1–16.

Steketee, G., & Foa, E. B. (1987). Rape victims. Post-traumatic stress responses and their treatment: A review of the literature. *Journal of Anxiety Disorders, 1,* 69–86.

Straus, M. A., & Gelles, R. J. (1986). Societal change and change in family violence from 1975 to 1985 as revealed by two national surveys. *Journal of Marriage and the Family, 48,* 465–479.

Straus, M. A., & Gelles, R. J. (1990). *Physical violence in American families: Risk factors and adaptation to violence in 8145 families.* New Brunswick, NJ: Transaction.

Straus, M. A., Gelles, R. J., & Steinmetz, S. K. (1980). *Behind closed doors: Physical violence in the American family.* New York: Doubleday/Anchor.

Uddo, M., Allain, A. N., & Sutker, P. B. (1996). Assessment of posttraumatic stress disorder: A conceptual overview. In T. W. Miller (Ed.), *Theory and assessment of stressful life events* (pp. 235–271). Madison, CT: International Universities Press.

Veronen, L. J., & Kilpatrick, D. G. (1983). Stress management for rape victims. In D. Meichenbaum & M. E. Jaremko (Eds.), *Stress reduction and prevention* (pp. 341–374). New York: Plenum.

Weaver, T. L., & Clum, G. A. (1995). Psychological distress associated with interpersonal violence: A meta-analysis. *Clinical Psychology Review, 15,* 115–140.

Williams, J. B., Gibbon, M., First, M. B., Spitzer, R. L., Davies, M., Borus, J., Howes, M. J., Kane, J., Pope, H. G., Rounsaville, B., & Wittchen, H. (1992). The structured clinical interview for DSM-III-R (SCID): II Multisite test–retest reliability. *Archives of General Psychiatry, 49,* 630–636.

Wilson, J. P., & Kraus, G. E. (1985). Predicting post-traumatic stress disorders among Vietnam veterans. In W. E. Kelly (Ed.), *Post-traumatic stress disorder and the war veteran patient* (pp. 102–147). New York: International Universities Press.

Wirtz, P. W., & Harrell, A. V. (1987a). Victim and crime characteristics, coping responses, and short- and long-term recovery from victimization. *Journal of Consulting and Clinical Psychology, 55,* 866–871.

Wirtz, P. W., & Harrell, A. V. (1987b). Effects of post-assault exposure to attack-similar stimuli on long-term recovery of victims. *Journal of Consulting and Clinical Psychology, 55,* 10–16.

Wolpe, J. (1990). *The practice of behavior therapy* (4th ed.). Elmsford, NY: Pergamon.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.20
Victims of Domestic Violence

AMY HOLTZWORTH-MUNROE and NATALIE SMUTZLER
Indiana University, Bloomington, IN, USA
and
ERNEST N. JOURILES and WILLIAM D. NORWOOD
University of Houston, TX, USA

9.20.1 INTRODUCTION 326
9.20.2 IMPACT OF MALE VIOLENCE ON INTIMATE FEMALE PARTNERS 326
- *9.20.2.1 Psychological Problems Among Battered Women* 326
  - *9.20.2.1.1 Post-traumatic stress disorder* 326
- *9.20.2.2 Depression* 327
- *9.20.2.3 Self-esteem* 327
- *9.20.2.4 Factors that Mediate a Woman's Reactions to Husband Violence* 328
  - *9.20.2.4.1 Coping skills* 328
  - *9.20.2.4.2 Cognitions* 328
  - *9.20.2.4.3 Dependency* 328
  - *9.20.2.4.4 Family of origin* 328

9.20.3 WHY BATTERED WOMEN STAY 328
- *9.20.3.1 Do Battered Women Stay in Abusive Relationships?* 328
- *9.20.3.2 Is Leaving an Abusive Relationship Safe?* 329
- *9.20.3.3 Factors Contributing to a Woman's Decision to Leave or Stay* 329

9.20.4 PSYCHOLOGICAL ABUSE 330
9.20.5 PSYCHOLOGICAL INTERVENTIONS WITH BATTERED WOMEN 330
9.20.6 FUTURE DIRECTIONS AND METHODOLOGICAL CONSIDERATIONS 331
9.20.7 MARITAL VIOLENCE AND CHILDREN 332
9.20.8 PREVALENCE 332
9.20.9 CHILD ADJUSTMENT PROBLEMS 332
- *9.20.9.1 Concurrent Problems* 332
- *9.20.9.2 Long-term Problems* 333

9.20.10 FACTORS ASSOCIATED WITH DETRIMENTAL IMPACT 334
- *9.20.10.1 Parameters of Violence* 334
- *9.20.10.2 Family Variables* 334
- *9.20.10.3 Child Variables* 335

9.20.11 INTERVENTIONS 336
9.20.12 CONCLUSIONS 336
9.20.13 REFERENCES 336

9.20.1 INTRODUCTION

Male violence against intimate female partners, whether it occurs in dating, cohabitating, or marital relationships (and which we will refer to as "husband violence"), is a serious problem. Data from nationally representative surveys in the USA suggest that each year one out of every eight women will experience physical aggression and nearly two million women will be severely assaulted by their male partners (Straus & Gelles, 1986). But women are not the only victims; each year, over three million children are exposed to interparental violence (Carlson, 1984). This chapter reviews the negative consequences of husband violence on women and children.

The first half of the chapter addresses the impact of male violence on intimate female partners. Although both men and women engage in physical aggression against partners, an accumulating body of evidence demonstrates that male violence has more detrimental effects than female violence. Male violence is more likely to result in physical injury and psychological problems (Holtzworth-Munroe, Smutzler, & Bates, 1997). Thus, the focus is on the problem of husband violence. The prevalence and correlates of the most common psychological problems among battered women and factors that may mediate a woman's responses to husband violence are reviewed. Also addressed is the issue of whether battered women stay in abusive relationships, and the problem of psychological abuse is introduced. Interventions with battered women are briefly discussed and some methodological considerations in this research area are mentioned.

The second half of the chapter addresses the impact of male violence against female partners on children living in such homes. First, the prevalence of the problem is discussed. Then, the problems experienced by children in these homes are reviewed, including externalizing problems, internalizing problems, and social competence/skills. Also reviewed are the factors associated with the detrimental impact of marital violence on children, including parameters of the violence, family variables, and child variables. Finally, the clinical implications of the available data are briefly discussed.

9.20.2 IMPACT OF MALE VIOLENCE ON INTIMATE FEMALE PARTNERS

9.20.2.1 Psychological Problems Among Battered Women

The study of battered women's psychological symptoms and diagnoses is controversial. Historically, the application of diagnostic labels to battered women often focused on the "shortcomings" of these women, implying that they were responsible for their own victimization. For example, early accounts described the masochistic tendencies of abused women (e.g., Reynolds & Siegle, 1959) and, more recently, Walker (1991) noted that the *DSM-III-R* diagnosis of self-defeating personality disorder reifies the notion that battered women are somehow responsible for their partners' aggression.

In addition, feminists "argue that diagnosing victims deflects energy from alleviating societal deficits" (Sato & Heiby, 1991, p. 151). According to feminist theory, husband violence is the result of a patriarchal society in which violence is simply one of many means used by men to ensure that women remain subordinate to them. Thus, to stop wife abuse, emphasis must be placed on building a more egalitarian society and holding men accountable for male violence. Given this view, focusing on the psychological characteristics of the victim is not only irrelevant but also counterproductive, as it may allow men to continue to ignore their role in the problem.

In contrast, clinicians often value diagnostic labeling as it can facilitate treatment planning. For example, therapists who are aware of the symptoms suffered by battered women will consider the possibility of marital violence in their female clients' lives, rather than assuming that the psychological symptoms alone should be the focus of treatment (Browne, 1993). In addition, understanding the negative impact of physical aggression on women can help demonstrate that marital violence is a serious problem, deserving of attention and scarce resources.

Despite disagreements about the potential usefulness of studying the psychological symptoms of battered women, researchers from divergent backgrounds have begun to agree that the symptoms of battered women should be viewed as the result or consequence of abuse rather than as stable characteristics predating abuse or as causes of abuse. Indeed, Browne (1993, p. 1077) warns that "psychologists continue to risk pathologizing women victims and [fail] to identify potential danger by not integrating what is known about assaults by relational partners with the rich literature on human responses to trauma." Thus, a focus on the traumatic impact of husband violence may help bridge the gap between feminists and clinicians because it allows for psychological diagnoses that do not blame the victim.

9.20.2.1.1 Post-traumatic stress disorder

For battered women, post-traumatic stress disorder (PTSD) involves symptom development following exposure to an extreme traumatic stressor (i.e., battering) that includes actual or threatened death or serious injury. PTSD symptoms can include increased arousal, avoidance responses, numbing of general responsiveness, re-experiencing of the trauma as well as "perception of life threat . . . extreme fear or terror, and a sense of helplessness at the time of the incident" (Browne, 1993, p. 1081). Researchers have examined the prevalence rates of PTSD symptoms among samples of battered women. Across different samples and using different measures of PTSD, the reported rates of PTSD among battered women range from 30% to 80%, with most researchers reporting a prevalence rate of 40–60% (e.g., Astin, Lawrence, & Foy; 1993; Gleason, 1993; Houskamp & Foy, 1991; Saunders, 1994). As would be expected, rates of PTSD are higher among battered than among nonbattered women (e.g., Kemp, Green, Hovanitz, & Rawlings, 1995). However, Riggs, Kilpatrick, and Resnick (1992) found no differences in PTSD across four groups of women who were victims of violent crime. Thus, their results demonstrate that violence towards a woman that is perpetrated by her husband is no less traumatizing than violence from others.

Consistent with the notion that PTSD is a consequence of the abuse experienced by women, recent studies have demonstrated a relationship between the severity and recency of the abuse experienced and risk for PTSD. When Houskamp and Foy (1991) categorized battered women as having been exposed to either high or low degrees of perceived life threat, 60% of the women in the high exposure group met criteria for PTSD compared with 14% in the low exposure group; in addition, the extent and severity of the abuse and length of relationship were positively related to the severity of PTSD symptoms. Similarly, Astin et al. (1993) found that severity of exposure to violence, as well as the recency of the abuse, were positively related to PTSD symptomatology. Kemp et al. (1995) compared battered women who met PTSD criteria to those who did not, finding that the PTSD group scored higher on measures of extent of physical and verbal abuse, number of injuries, forced sex, and greater sense of threat; they had also been out of the abusive relationship a shorter time.

These researchers also examined the relationship between PTSD symptoms and other correlates. Astin et al. (1993) found that social support and positive life events occurring since the battering were negatively correlated with PTSD symptoms, whereas negative life events and developmental family stressors were positively related to PTSD severity. Kemp et al. (1995) found that negative life events and a history of child abuse predicted PTSD; perceived social support was related to lower levels of PTSD, while disengaged coping was one of the strongest predictors of PTSD symptoms.

9.20.2.2 Depression

Researchers have also reported that a high percentage of battered women, usually about 50%, experience high levels of depression (e.g., Cascardi & O'Leary, 1992; Christian, O'Leary, & Vivian, 1994; Sato & Heiby, 1991). Researchers generally find a higher prevalence of depression among battered women than among nonbattered women (Andrews, 1995; Stets & Straus, 1990), including women experiencing nonviolent serious relationship problems (Campbell, 1989; Cascardi & O'Leary, 1992).

Examining the correlates of depression among battered women, Sato and Heiby (1991) found that battered women who realistically assessed their battering relationship and who reported poor self-reinforcement skills, personal or economic losses, and past histories of depression appeared to be at particular risk for depression. Mitchell and Hodson (1983) found that the severity and frequency of the aggression was related to increases in depression. In addition, less active cognitive coping and more avoidance coping were related to more severe depression, minimal personal resources, less help from institutional sources, and higher levels of avoidance responses from friends were also related to depression. Campbell (1989) was able to account for about 57% of the variance in depression scores among battered women, using a combination of low self-esteem, low perceived ability to care for oneself, and feelings of less control in the relationship predicting depressive symptoms.

9.20.2.3 Self-esteem

The majority of studies conducted suggest that battered women have lower self-esteem than nonbattered women (e.g., Aguilar & Nightingale, 1994; Cascardi & O'Leary, 1992; Perilla, Bakeman, & Norris, 1994). Mitchell and Hodson (1983) found that severity and frequency of husband violence was strongly related to lowered self-esteem. Frisch and MacKenzie (1991) found that chronically abused women scored significantly lower on measures of self-esteem than formally abused women. Similarly, Aguilar and Nightingale (1994) found that battered women who were currently involved with an abusive partner had lower self-esteem than women who had not been involved with their abusive partner for at least one year.

9.20.2.4 Factors that Mediate a Woman's Reactions to Husband Violence

9.20.2.4.1 Coping skills

How a woman copes with the abuse she is suffering may be related to her reactions to the abuse. For example, Mitchell and Hodson (1983) found that less active cognitive coping and more avoidance coping were related to more severe depression, as were minimal personal resources, less help from institutional sources, and higher levels of avoidance from friends.

9.20.2.4.2 Cognitions

Cognitions of a battered woman may influence the steps she takes to deal with the violence in her life. Although a series of studies describe the attributions battered women offer for husband violence (Holtzworth-Munroe, 1988), only a few researchers have related these attributions to womens' responses to husband violence. Andrews and Brewin (1990) found that women who made internal and stable attributions of blame for abuse were more likely, than women who did not, to have suffered persistent depression after the violent relationship ended. In contrast, Campbell (1989) found that womens' attributions of blame for relationship problems did not add significantly to the prediction of depression among the women, once self-esteem, self-care agency, and feelings of control in the relationship were used as predictors. Thus, it is unclear how a woman's attributions for husband violence relate to her reactions to this violence.

Researchers have also investigated the relationship between traditional sex role beliefs of women and husband violence, because a battered woman who holds traditional sex role beliefs may be less able to protect and care for herself (Warren & Lanning, 1992). Hotaling and Sugarman (1986) found that only 50% of studies reported a positive relationship between traditional sex role beliefs of women and marital violence. Subsequently, researchers have generally found that battered women are less masculine or androgynous and more feminine than nonbattered women (Apt & Hurlbert, 1993; Schwartz & Mattley, 1993; Warren & Lanning, 1992). However, whether such beliefs are related to women's reactions to the violence is still a matter of speculation.

9.20.2.4.3 Dependency

Some theorists have hypothesized that wives with high marital dependency may be more tolerant of marital abuse. For example, Kalmuss and Straus (1990) found that women experiencing objective (i.e., economic) or subjective (i.e., fear of negative effects resulting from loss of the marriage) dependency experienced more abuse than nondependent women.

9.20.2.4.4 Family of origin

It has been proposed that women who grew up in violent homes are at risk of becoming adult victims of spouse abuse. For example, girls growing up in violent homes may come to believe that marital violence is normal and acceptable, thus tolerating such abuse when it occurs in their own marriages. Thus, a large number of researchers have asked battered and nonbattered women to retrospectively recall whether or not they experienced violence in their childhood family of origin (e.g., Doumas, Margolin, & John, 1994; Langhinrichsen-Rohling, Neidig, & Thorn, 1995; Simons, Johnson, Beaman, & Conger, 1993). However, these studies have produced mixed findings, so that, at this time, it cannot be definitively concluded that having grown up in a violent home is a risk factor for becoming a victim of husband violence.

9.20.3 WHY BATTERED WOMEN STAY

9.20.3.1 Do Battered Women Stay in Abusive Relationships?

One of the most commonly asked questions in this area is why battered women stay with abusive partners. Schwartz (1988, p. 241) suggests that a more appropriate question may be "why do we assume that abused women stay in their marriages?" Based on interviews with women in the National Crime Survey (NCS), Schwartz (1988) found that 32% of victims of spousal assault were divorced and 47% were separated at the time of the interview. These rates were significantly higher than those found among women who had been victims of a stranger assault (19.1% divorced, 11.2% separated) and women who were not victimized (8.5% divorced, 3.6% separated). These data suggest that the majority of physically abused women do not stay in an abusive relationship, and that our research focus might productively switch to an examination of the ways in which abused women take steps to improve their lives (Schwartz, 1988).

A number of researchers have reported similar data, finding that 50% or more of battered women leave their partner within two years (Okun, 1986; Strube & Barbour, 1984). However, Okun (1986) also found that women may leave many times before finally ending the relationship, suggesting that the decision to leave an abusive relationship is an ongoing process requiring multiple attempts. Okun (1986, p. 118) summarizes this idea: "the process of separating and recohabitating by battered women is not an endless back-and-forth shuttle in static equilibrium, but a progressive process toward change in the violent relationship, including the termination of that relationship."

9.20.3.2 Is Leaving an Abusive Relationship Safe?

Asking why women stay implies that leaving is a safe option for battered women. However, this may not be the case. Comparing rates of spousal homicide among cohabitating and separated couples, Wilson and Daly (1993) found that wives were at higher (i.e., 2–4 times) risk of being killed by their husbands following separation than when living with their partner. Saltzman and Mercy (1993) found that over one-fifth of the fatal incidents occurred between people whose relationships had terminated. These data raise the possibility that an individual cannot always stop the violence by terminating an abusive relationship.

9.20.3.3 Factors Contributing to a Woman's Decision to Leave or Stay

Given the comprehensive Strube (1988) review addressing the question of why women stay in abusive relationships, here we only briefly note a few of the major factors believed to influence a woman's decision to stay in or leave an abusive relationship. Strube and Barbour (1983, 1984) found that economic dependence was related to a woman's decision to stay in an abusive relationship. Women who were employed were more likely to leave, and women who remained attributed economic hardship as at least one factor for staying. Similarly, Aguirre (1985) found that if the husband was the sole source of income, a woman was more likely to return to him after seeking help from a shelter. Okun (1986) also found that 55% of women who had a higher income than their husbands immediately terminated the relationship, compared to 28% of the women whose husbands earned equal or greater pay. In addition, if the husband was not gainfully employed, 48% of the wives immediately terminated the relationship, whereas 27% of women whose husbands were full-time or self-employed terminated cohabitation. Thus, it appears that financial status influences a woman's decision to remain in an abusive relationship. However, these results must be interpreted with caution because the majority of the women in these studies were in battered women shelters which often provide services to lower income women. There is little information as to how women from higher income brackets cope with wife abuse.

Studies have provided mixed and contradictory evidence for the notion that the more severe the marital violence, the more likely the woman is to leave. Ferraro and Johnson(1983, pp. 331–332) speculated that it may not be the frequency or severity of violence that serves as an impetus for a woman's decision to leave a relationship but rather "a sudden change in the relative level of violence. Women who suddenly realize that battering may be fatal may reject rationalizations in order to save their lives." In addition, "as battering progresses, periods of remorse may shorten, or disappear, eliminating the basis for maintaining a positive outlook on marriage." Thus, focusing exclusively on the severity and frequency of violent episodes may be too simplistic. Rather, it may be a significant variation in the pattern of abuse that instigates a separation.

Loseke and Cahill (1984, p. 304) challenged the notion that staying in an abusive relationship is a deviant response. They argued that terminating any type of intimate relationship is difficult, noting that "even a cursory review of the sociological literature on marital stability and instability suggests that, at least in regard to their reluctance to leave their mates, battered women are quite similar to both other women and to men."

Interestingly, researchers often approach the question of why battered women stay as if the woman was the only person participating in the process. Only recently have investigators begun to consider how the help obtained by women may impact the choice of remaining in or leaving a relationship. Similarly, researchers rarely consider the batterer's influence over his partner's decision to return to the relationship. However, this may be an important factor to examine. For example, Schutte, Malouff, and Doyle (1988) found that all the women in their shelter sample reported that their partner had attempted to persuade them to return to the relationship. Similarly, Gondolf (1988) found that the best predictor of returning to an abusive husband was the husband's having sought help or entering counseling, suggesting that battered women place their hope in the notion that their husbands will change, although not all batterers are helped by therapy (Rosenfeld, 1992).

In summary, battered women appear to have many reasons for remaining with an abusive partner. However, such findings must be placed in the context that many battered women do eventually leave their abusers. Rather than continuing to narrowly examine why battered women stay, future research should consider broader issues, including what factors help women to end the abuse in their lives.

9.20.4 PSYCHOLOGICAL ABUSE

Although the primary focus of this chapter is husband physical aggression, researchers have recently begun to recognize the importance of psychological abuse, which co-occurs with physical abuse. Researchers have delineated different types of psychological abuse. For example, Rodenburg and Fantuzzo (1993) assessed sexual, psychological, and verbal abuse, and Shepard and Campbell (1992) assessed social isolation, threats, use of male privilege, and economic control. Tolman (1989) found that the items on his measure of psychological abuse loaded on to two factors—emotional–verbal abuse and dominance–isolation. A consistent theme emerges across these studies, suggesting that psychological abuse involves an attempt to control or dominate one's partner (Murphy & Cascardi, 1993).

Psychological abuse is prevalent among couples experiencing physical aggression. For example, Follingstad, Rutledge, Berg, Hause, and Polek (1990) found that 99% of the battered women they studied had experienced some type of emotional abuse and 72% experienced four or more types. Hudson and McIntosh (1981), examining a sample from a battered women's shelter, found a correlation of 0.86 between the physical abuse and psychological abuse. Using national survey data, Stets (1990) found that less than 1% of respondents indicated that physical aggression had occurred without verbal aggression.

Psychological aggression is damaging to battered women. In their study of women who had been victims of physical abuse, Follingstad et al. (1990) found that 72% of the women believed that emotional abuse had a more severe negative impact on them than physical aggression. This is not surprising given the frequent occurrence of various forms of emotional abuse—over 60% of the women reported that multiple forms of psychological abuse occurred weekly. It is thus possible that many of the effects of marital violence observed in battered women are actually due to psychological abuse or to the combination of psychological and physical abuse (Murphy & Cascardi, 1993).

Recent research suggests that psychological aggression precedes, and is a risk factor for, physical aggression. For example, using national survey data, Stets (1990) found that in over 50% of the cases, verbal aggression occurred without physical aggression. However, less than 1% of respondents indicated that physical aggression occurred without verbal aggression. Prospective studies have led to similar conclusions. For example, Malamuth, Linz, Heavey, Barnes, and Acker (1995) demonstrated that men's use of verbal aggression was the only direct predictor of men's use of physical nonsexual aggression toward women, with verbal aggression accounting for over 30% of the variance in physical aggression. In a study of newly married couples, Murphy and O'Leary (1989) found that husbands' use of psychological aggression at 18 months after marriage significantly predicted husband physical aggression at 30 months after marriage.

In summary, only recently have researchers begun to study husbands' psychological abuse of women. These initial studies indicate that psychological abuse is a major problem. It is highly prevalent among violent couples and has a negative impact on battered women. In addition, psychological abuse may precede or predict husband physical violence.

9.20.5 PSYCHOLOGICAL INTERVENTIONS WITH BATTERED WOMEN

There is insufficient empirical data to make recommendations about what type of psychotherapy is beneficial to battered women; indeed, little research has addressed this issue. This lack of research may reflect the fact that most marital violence experts do not believe that traditional psychotherapy, at least alone, is the most appropriate intervention for battered women (Dutton, 1996). Due to a lack of awareness of the problem of husband violence, psychologists have often mistreated battered women, focusing on their symptoms without fully understanding how the abuse in their lives causes such symptoms. However, focusing solely on psychological symptoms is unlikely to benefit a woman who leaves each therapy session to go home to an abusive partner. In addition, recent research on the correlates of psychological symptoms among battered women suggests that psychotherapy may not be necessary to help many battered women overcome their psychological symptoms; more simply, length of time out of an abusive relationship can lead to a decrease in symptoms. Furthermore, battered women often require help that is not part of a traditional psychotherapy approach and which most psychologists have not been trained to provide. For example, careful safety planning, to protect the woman from further abuse, is often the most important first intervention. Battered women often also require economic and legal assistance to deal with questions of restraining or protective orders, divorce and custody, prosecution of the batterer, how to obtain housing, and financial support. Often, these other needs must be addressed before a battered woman is able to deal with more traditional, psychological issues.

Given these concerns, most experts on battered women recommend therapeutic interventions that address the multiple and individual needs of women who are living in, or have left, an abusive relationship. For example, Dutton (1996, p. 2) proposed that "the task for the therapist is to address the configuration of the client's needs derived from the battering experience, in the context of other life issues or problems that she may also face, and for which she may or may not be seeking help." In addition, psychologists should coordinate their efforts with local womens' shelters and other community agencies. Indeed, Pence (1989) has argued that comprehensive community involvement is required; therapists should collaborate with law enforcement officials, the judicial system, and women advocacy services. Only through such coordinated efforts will therapists be able to adequately address the issues facing a battered woman. Given such emphases, recent developments in intervention efforts reflect the complexity of problems faced by battered women. For example, Dutton (1996) has presented an integrated intervention model based on social learning, cognitive-behavioral, and feminist theories. The model includes crisis intervention (e.g., safety planning), choicemaking (e.g., increasing awareness of options and skill building), and healing (e.g., assessing how a particular woman has reacted to the violence and implementing an appropriate treatment plan).

In summary, traditional psychotherapies may be inadequate to address the immediate needs of a woman living with her batterer or having recently left him. Safety, economic, and legal needs are likely to take precedence. However, as a battered woman obtains an abuse-free life, she may wish to focus on the psychological sequelae of the trauma she has endured. Unfortunately, little is known about the best way to help such women, although supportive therapies and therapy focusing on PTSD are promising places to start.

9.20.6 FUTURE DIRECTIONS AND METHODOLOGICAL CONSIDERATIONS

In early studies in this area, many researchers did not include comparison groups; instead, they simply tallied the percentage of battered women in a sample who showed psychological symptoms. Although more recent studies have included nonabused comparison samples, they rarely include women in nonviolent but maritally distressed relationships. Thus, it is difficult to establish which of the demonstrated effects are unique to violence vs. which are more generally related to living in an unhappy marriage. Another sampling problem involves the large number of studies examining only battered women in shelters or therapy. Such women are seeking help and/or have left their partner. Whether such findings will generalize to battered women in the community has not been extensively examined.

In addition, most researchers assign all battered women to one subject group, comparing them with nonbattered women; however, two typology studies make it clear that battered women are a heterogeneous group. Snyder and Fruchtman (1981) and Follingstad, Laughlin, Polek, Rutledge, and Hause (1991) identified five clusters of battered women. In both studies, women varied along such dimensions as the severity, frequency, and chronicity of the violence they experienced, how stable and satisfactory their relationships were, whether or not they returned to the batterer, to whom or what they attributed the violence, whether they accepted violence in their relationship, and whether their children were also the victims of violence. These typologies demonstrate that it is important for researchers to begin considering such differences in the experiences of battered women.

Studies examining how battered and nonbattered women differ have been controversial because the findings can easily be used to point to the deficits of battered women, implying that they are to blame for their abuse. However, researchers in this area are increasingly sensitive to such issues. For example, studies of PTSD clearly link the psychological distress found among battered women to theories of trauma, and findings relating degree of symptomatology to the amount of abuse experienced support the notion that the problems of battered women are the consequence of violence. These data support Hotaling and Sugarman's (1986, p. 118) conclusion: "The review of victim characteristics makes it clear that the most influential victim precipitant is being female ... These findings do not augur well for theoretical models of victimization that focus upon characteristics of victims." However, in reality, the correlational nature of the current research leaves questions about temporal sequence and causality unanswered and will continue to lead to theoretical debates about these issues. For example, it can be asked why virtually all of the correlates of violence found among husbands are assumed to be causes of violence, while many of the same correlates of violence found among battered women are assumed to be the consequences of violence. Future researchers need to address such questions.

Recent work has begun to examine the correlates of these psychological problems. Unfortunately, however, there are few, if any, standardized measures of these correlates and mediators of the relationship between husband violence and its psychological impact on women; this makes it difficult to draw conclusions across studies. In addition, many potentially important correlates have received little attention. For example, there is a need to understand resiliency among battered women and to study the strengths of battered women who do *not* experience problems. Also, researchers have only begun to examine how leaving an abusive relationship affects a woman's psychological reactions to the abuse. Future longitudinal research should examine the longer term consequences for husband violence.

9.20.7 MARITAL VIOLENCE AND CHILDREN

Women are not the only victims of marital violence; many children of battered women suffer devastating consequences as well. Although the empirical investigation of the effects of spousal violence on children is still in its infancy, one conclusion has received almost universal support: compared to children of nonviolent community families, the children of battered women are at substantially increased risk for behavior problems.

9.20.8 PREVALENCE

Prevalence estimates of children's exposure to marital violence are frequently based on estimates of the presence or absence of such violence in homes which include children. Exposure may thus refer to direct exposure (i.e., hearing or seeing the violence first hand), or to indirect exposure (i.e., living in a household characterized by violence but not directly witnessing it). This distinction is important for a number of reasons but, perhaps most importantly, because direct exposure is likely to have different effects on children than indirect exposure. Unfortunately, distinguishing empirically between direct and indirect exposure has been very difficult. There is little consistency between what parents believe their children have witnessed and what their children report having observed (Jaffe, Wolfe, & Wilson, 1990) and, indeed, only fair agreement between parents about whether their children have seen or heard interparent aggression (O'Brien, John, Margolin, & Erel, 1994). Clinical investigators typically report that the children in these families witness more marital violence than their parents realize (Hughes, 1988; Rosenberg, 1987).

The most widely cited prevalence estimates of children living in households in which marital violence has occurred are based on US demographic data and the Straus, Gelles, and Steinmetz (1980) national survey of community households. Carlson (1984) estimated that at least 3.3 million US children (aged 3–17 years) are exposed directly or indirectly to interparent violence each year. Pagelow (1982) estimated that more than 25 million children in the US are exposed to interparent violence at some point during their childhood. It is also important to note that both Carlson and Pagelow stated that their estimates are most likely lower boundaries, with the true number of children exposed to marital violence being much higher.

Data on children's exposure to marital violence within clinical samples are limited, but the few existing studies suggest substantial rates of exposure. For example, over 40% of boys admitted to a child psychiatric clinic lived in families characterized by wife abuse (Stewart & deBlois, 1981). In a study of couples seeking marital therapy with children between 5 and 12 years, over 50% of the families reported recent spouse abuse (Jouriles, Murphy, & O'Leary, 1989). In sum, it appears that a large percentage of the children who eventually come into contact with mental health professionals live in maritally violent homes.

9.20.9 CHILD ADJUSTMENT PROBLEMS

9.20.9.1 Concurrent Problems

Most controlled investigations of the adjustment problems suffered by children of battered women have assessed the level of child problems that are evident after a recent exposure to marital violence, often during a child's stay at a battered women's shelter. In addition, most of these controlled investigations have assessed two broad-band dimensions of child problems: (i) externalizing problems, such as aggressive and oppositional behavior; and (ii) internalizing problems, such as anxiety and depression. Although there have been some inconsistencies in the specific results across studies, children of battered women have repeatedly been found to exhibit higher levels of problems when compared to children in community comparison families (Christopoulos et al., 1987; Holden & Ritchie, 1991; Wolfe, Jaffe, Wilson, & Zak, 1985). Moreover, the available evidence generally indicates that marital violence is a more potent risk factor for both externalizing and internalizing problems than are verbal conflict or general marital discord (Fantuzzo et al., 1991; Jouriles et al., 1989; Rossman & Rosenberg, 1992), although there are at least two exceptions to this finding (Hershorn & Rosenbaum, 1985; Rosenbaum & O'Leary, 1981). Certainly an important and consistent finding is that a substantial proportion of the children of battered women score in the "clinical" range on standard measures of child problems, that is, at a level considered to warrant clinical intervention. Estimates of the percentage of children of battered women at such levels range from 25% to 70%, compared to 10–20% for children in community comparison groups (McDonald & Jouriles, 1991).

Differences on externalizing and internalizing problems between children exposed to marital violence and community children have been documented with a variety of samples of children. Most commonly, differences have been documented in samples of wide age ranges with boys and girls grouped together (Hughes, 1988; Hughes, Parkinson, & Vargo, 1989), but they have also been found in more homogeneous samples of preschoolers (Fantuzzo et al., 1991), young school-aged children (Rossman & Rosenberg, 1992; Sternberg et al., 1993), and adolescents (Carlson, 1990). There is also some evidence that differences may be larger for younger as opposed to older children (Hughes, 1988), but this is not yet a firmly established finding. When results are reported separately by child gender, differences have emerged for both boys and girls (Christopoulos et al., 1987; Jaffe, Wolfe, Wilson, & Zak, 1986), but the types of problems experienced by boys and girls has not been consistent across studies. For example, some data indicate that boys exposed to marital violence exhibit more externalizing problems than community boys (Jaffe et al., 1986), whereas other data do not (Christopoulos et al., 1987; Rosenbaum & O'Leary, 1981). Similarly, some studies indicate that boys exposed to marital violence exhibit greater internalizing problems than community boys (Jaffe et al., 1986), whereas other studies indicate no such differences (Christopoulos et al., 1987; Rosenbaum & O'Leary, 1981). Similar inconsistent findings have emerged for girls on internalizing and externalizing problems. Nevertheless, at a more general level, the available data suggest that children exposed to marital violence are at increased risk for behavior problems, and the levels of these problems for many of these children are such that they warrant immediate attention.

Other areas of functioning that have received some attention in the empirical literature include social competence, cognitive functioning, and physical health, but existing data on these problems remains preliminary. Several controlled studies indicate that children of battered women exhibit lower levels of social competence than children in community comparison families (Rossman & Rosenberg, 1992; Wolfe et al., 1985; Jaffe et al., 1986). Controlled research also suggests that the children of battered women do not differ on measured intelligence when compared to community children (Christopoulos et al., 1987; Rossman & Rosenberg, 1992). However, most studies of children's social competence, cognitive functioning, and physical health have not included a comparison group, and many have collected these data unsystematically and with instruments of uncertain reliability and validity. Although the findings from such studies often suggest that many children of battered women display problems in the domains of social competence, cognitive functioning, and physical health (Davis & Carlson, 1987; Stagg, Wills, & Howell, 1989). The methodological problems with these studies and conflicting results with some of the better controlled studies suggest that these be interpreted very tentatively.

Clinical reports also suggest a number of additional problems suffered by children of battered women (Jaffe et al., 1990). These include symptoms associated with the diagnosis PTSD, the development of attitudes which support violence as an appropriate and effective means of obtaining power and control, and the development of beliefs that children are themselves responsible for their parents' marital violence. Also, clinical reports suggest that male children of battered women tend to be aggressive toward female targets in particular (Jaffe et al., 1990). Obviously, this is not an exhaustive list of problems derived from clinical data on the children of battered women, but it does help illustrate further the breadth of problems experienced by these children and the subtlety of some of the effects of violence exposure.

9.20.9.2 Long-term Problems

Most investigations of child behavior problems among children of battered women have focused on the level of problems evident during an assessment at one point in time. At present little is known about the duration or developmental course of these problems. However, what little evidence there is indicates that the problems exhibited by children of battered women remain stable over time without specific intervention. In a cross-sectional study comparing a group of children who had recently been exposed to marital violence to a group of children who had not been exposed to violence for at least six months, the mean levels of externalizing and internalizing problems were very similar across the two groups (Wolfe, Zak, Wilson, & Jaffe, 1986). One long-term follow-up study indicates that boys' exposure to wife abuse is one of the few childhood factors which predict adult criminal activity (McCord, 1979), but this has yet to be tested by subsequent research. Retrospective data consistently indicate that boys' exposure to fathers' marital violence relates to spouse abuse in adult intimate relationships (Hotaling & Sugarman, 1986), and adolescents who report a history of parents' marital violence have been found to have higher "child abuse potential" scores than those who do not report a background of domestic violence (Miller, Handal, Gilner, & Cross, 1991). Retrospective data also suggest that childhood exposure to marital violence relates to adult levels of internalizing problems, such as depression, low self-esteem, and trauma (Forsstrom-Cohen & Rosenbaum, 1985; Henning, Leitenberg, Coffey, Turner, & Bennett, 1996; Silvern et al., 1995). Thus, evidence from several sources converges to suggest that childrens' exposure to violence has potential long-term negative consequences.

9.20.10 FACTORS ASSOCIATED WITH DETRIMENTAL IMPACT

Despite convincing evidence of increased risk, there is also evidence of substantial variability in the level of behavioral and emotional adjustment displayed by the children of battered women. Although a large percentage of these children are reported to be exhibiting problems at clinical levels, a subgroup of these children, even in families characterized by extreme marital violence, have been described as well-adjusted (Hughes & Luke, 1998). The empirical literature on children's response to stress suggests that a number of factors may alter or deflect the impact of marital violence on children. These include parameters of the violence (e.g., frequency, severity), family variables (e.g., presence of child abuse, family socioeconomic status, quality of the parent–child relationship, family social support, psychological health of parents), and child variables (e.g., intelligence, problem-solving skills), among others. Although few of these variables have been investigated systematically among the children of battered women, several studies in the area provide important leads.

9.20.10.1 Parameters of Violence

The empirical literature on marital conflict (as distinguished from marital violence *per se*) and child adjustment has uncovered several dimensions of conflict which appear useful in the prediction of child behavior problems. These include the frequency, intensity, resolution, and content of the conflict. In general, the more frequent and intense the marital conflict, the greater the likelihood of child difficulties; witnessing poorly resolved or unresolved marital conflicts are more detrimental to children than witnessing resolved conflicts; and overt marital conflicts about children or childrearing issues appear more detrimental than overt marital conflicts about other issues (Davies & Cummings, 1994; Grych & Fincham, 1990).

The empirical literature on parameters of marital violence and child adjustment has focused primarily upon the frequency and intensity of violence. In most of this research, the frequency and intensity of the marital violence is confounded, perhaps inextricably (e.g., children exposed to frequent violence are also exposed to more severe violence). Similar to the results of research on marital conflict and child adjustment, the frequency and intensity of marital violence relates positively to both externalizing and internalizing child behavior problems (Fantuzzo et al., 1991; Jouriles & Norwood, 1995; Jouriles, Norwood, McDonald, Vincent, & Mahoney, 1996; O'Keefe, 1994; Spaccarelli, Coatsworth, & Bowden, 1995).

Although the impact of content of conflicts which lead to marital violence on child adjustment has not yet been investigated, it has been found that arguments about children and childrearing are more likely to precede marital violence than conflicts about other topics (Straus et al., 1980). Such findings are noteworthy given that marital conflicts over childrearing are more often observed by children than other conflicts, and frequency measures of discord over childrearing relate more consistently to child problems than frequency measures of global marital discord or marital conflicts about other topics (Jouriles et al., 1991).

9.20.10.2 Family Variables

Marital violence typically occurs in the context of other family risk factors for child problems. Marital discord, child abuse, poor parent–child relations, family economic disadvantage, and parent distress/psychopathology are examples. Some of these risk factors are probably a result of marital violence, whereas others are hypothesized to play key roles in the development of violence. It is also likely that some risk factors are neither a direct cause nor a consequence of violence, but simply coexist with it. Researchers are just beginning to investigate how marital violence interacts with other family variables in the development of child problems among these families.

As indicated above, children are at greater risk for problematic adjustment when living in a family marked by recent marital violence than when living in a maritally discordant, nonviolent family (Fantuzzo et al., 1991; Jouriles et al., 1989; Rossman & Rosenberg, 1992). However, the frequency of some behaviors, such as insults or the kicking of objects, which often occur in the context of marital conflict but that do not involve direct physical assaults upon a spouse, still appear to be important in the prediction of concurrent child adjustment problems within maritally violent families. For example, in two separate studies, the frequency of spousal aggression other than physical marital violence has been found to contribute uniquely to the prediction of child problems after controlling statistically for physical marital violence (Jouriles et al., 1996). Consistent with this finding, laboratory data indicate that children with a history of exposure to physical marital violence respond more deviantly to simulated, nonviolent conflicts than children without such a history (Cummings, Pellegrini, Notarius, & Cummings, 1989; O'Brien, Margolin, John, & Krueger, 1991). Thus, children's exposure to physical marital violence does not appear to vitiate the impact of exposure to nonviolent marital conflict; in fact, exposure to physical violence may actually exacerbate the negative effects of exposure to nonviolent conflict.

Physical marital violence is a very strong correlate of parental physical aggression towards children (Straus et al., 1980; Straus & Gelles, 1990), especially aggression towards male children (Jouriles & LeCompte, 1991; Jouriles & Norwood, 1995). There is also some indication that rates of child sexual abuse are extremely high in families characterized by physical marital violence (McCloskey, Figueredo, & Koss, 1995). As one might expect, children who are both victims of parental physical abuse and witnesses of interparent violence are at greater risk for problems than those who only witness interparent violence (Davis & Carlson, 1987; Hughes et al., 1989; O'Keefe, 1995; Silvern et al., 1995). Some research suggests that the effects of these two forms of family violence are additive (i.e., both the witnessing of interparent violence and the experiencing of parental aggression contribute unique variance in the prediction of child adjustment problems (O'Keefe, 1995). Other studies, however, emphasize the greater risk associated with the experience of physical parent–child aggression (Jouriles, Barling, & O'Leary, 1987; McCloskey et al., 1995). It is also possible that the experience of child abuse potentiates the effects of exposure to marital violence.

Marital violence appears to be associated with poor mother–child relationships (McCloskey et al., 1995; O'Keefe, 1995). In fact, O'Keefe reported that within maritally violent homes, the mother–child relationship is more problematic than the father–child relationship. This finding is not only important in its own right, but it is also potentially important to our understanding of how violence affects children. Reviews of childhood stressors and protective factors (e.g., Garmezy, 1983; Luthar & Zigler, 1991; Rutter, 1987) suggest that a good parent–child relationship can buffer children from the negative consequences of stressful events. Thus, a child who has a poor relationship with his or her mother is perhaps more vulnerable to the negative effects of marital violence. It might be hypothesized that one of the pathways linking marital violence to child problems is its alteration of the quality of the mother–child relationship.

Many other family variables have also been linked to marital violence, with varying support for marital violence's unique contribution to child behavior in the context of these other difficulties. Spaccarelli, Sandler, and Roosa (1994) found in a sample of inner city families that children's witnessing of marital violence correlated with parental divorce and single-parent status, lower family income, lower maternal education, frequent family moves, parental drinking, and father's incarceration. After statistically controlling for these variables, the witnessing of parental violence still contributed unique variance to the prediction of girls' conduct problems, and both boys' and girls' depression. In contrast, Wolfe et al. (1985) found marital violence to be correlated with maternal distress and the frequency of family moves, but exposure to marital violence did not contribute uniquely to the prediction of child problems after controlling for these variables.

9.20.10.3 Child Variables

In the above review of the relationship between marital violence and child adjustment, research was discussed that has considered certain child variables such as age and gender. Very few studies have investigated other child variables. There are some data on child ethnicity which suggest that Caucasian boys may exhibit higher levels of externalizing problems in response to interparent violence than African-American boys (O'Keefe, 1994; Stagg et al., 1989). There is also some evidence that children's cognitions about their parents' violence influences the effects of that violence. Specifically, children's ability to distract themselves and their ability to comfort themselves both appear to mitigate to some extent the negative impact of exposure to marital violence (Rossman & Rosenberg, 1992). However, methodological limitations make the reported data difficult to interpret precisely.

9.20.11 INTERVENTIONS

The professional literature on children of battered women offers a number of suggestions for intervention and includes descriptions of several specific intervention programs. It is important to note, however, that many of these suggestions and programs are directed toward children of battered women as a group. That is, a child is targeted for intervention on the basis of the behavior of his or her mother's batterer rather than on the specific needs of the child. Moreover, with a few notable exceptions, systematic evaluation of the efficacy of these suggestions and programs has not been undertaken.

With these problems in mind, we recommend the careful and thorough assessment of child and family functioning and then the customization of services to meet the specific needs of each family. Children respond in a variety of ways to their parents' marital violence, and this variability must be considered during assessment and intervention. For many of these children, the stress caused by factors associated with the sequelae of marital violence (e.g., changes in residence and schools) may be as or more important to the child's current level of functioning than their having witnessed martial violence (Jaffe, Wilson, & Wolfe, 1989).

One area in need of assessment and potential intervention particularly worthy of note is children's safety. Children in maritally violent homes are at risk for child abuse, and its presence or potential must be considered. Children's knowledge of relevant safety steps has been identified as a potentially important area to address among children of battered women in that these children often require a greater repertoire of skills (e.g., finding a safe place for themselves during a violent incident, the ability to summon help or police without endangering themselves) than that required of most other children their age (Jaffe, Sudermann, & Reitzel, 1992). Other potentially important areas for assessment or intervention include identifying and using social supports, coping with feelings of responsibility for family violence, coping with uncertainty about their own and their family's future, childrens' management of their own anger and other emotions, and the potential development of gender stereotypes based on their experience of violently coercive family interactions. These and other potential targets of intervention clearly warrant further research attention.

9.20.12 CONCLUSIONS

Millions of women and children are exposed to husband violence each year in the USA, and the negative consequences of husband violence are many and varied. Not all women and children exposed to male violence suffer negative consequences or psychological problems; rather, the severity of consequences experienced is affected by a variety of factors, most importantly the level of violence experienced.

Unfortunately, social scientists have only recently begun to study the problem of domestic violence. Thus, many therapists have received little formal training on this topic, and they may not be adequately prepared to deal with the problem of domestic violence as it affects the lives of their adult and child clients. For example, Hansen, Harway, and Cervantes (1991) mailed several hundred marital and family therapists case descriptions of couples seeking therapy and asked therapists to write a case conceptualization of the couple. The cases mentioned incidents of husband violence. However, very few of the therapists explicitly identified violence or battering as a problem, and therapists did not indicate that lethality was a possible concern, despite the fact that one of the cases actually ended in the husband's murder of the wife.

These findings suggest that many therapists could benefit from learning more about the problem of domestic violence, the impact that aggression has on women and children, and the relationship between psychological symptoms and abuse experienced by women and children. Therapists owe it to the women and children who are the survivors of male violence to learn about this problem, respect their experiences, and do their best to help them end the violence in their lives.

9.20.13 REFERENCES

Aguilar, R. J., & Nightingale, N. N. (1994). The impact of specific battering experiences on the self esteem of abused women. *Journal of Family Violence, 9,* 81–95.

Aguirre, B. E. (1985). Why do they return? Abused women in shelters. *Social Work, 30,* 350–354.

Andrews, B. (1995). Bodily shame as a mediator between abusive experiences and depression. *Journal of Abnormal Psychology, 104,* 277–285.

Andrews, B., & Brewin, C. R. (1990). Attributions of blame for marital violence: A study of antecedents and consequences. *Journal of Marriage and the Family, 52,* 757–767.

Apt, C., & Hurlbert, D. F. (1993). The sexuality of women in physically abusive marriages: A comparative study. *Journal of Family Violence, 8,* 57–69.

Astin, M. C., Lawrence, K. J., & Foy, D. W. (1993). Posttraumatic stress disorder among battered women: Risk and resiliency factors. *Violence and Victims, 8,* 17–27.

Browne, A. (1993). Violence against women by male partners prevalence, outcomes, and policy implications. *American Psychologist, 48,* 1077–1087.

Campbell, J. C. (1989). A test of two explanatory models of women's responses to battering. *Nursing Research, 38,* 18–24.

Carlson, B. E. (1984). Children's observations of interpersonal violence. In A. Roberts (Ed.), *Battered women and their families* (pp. 147–167). New York: Springer.

Carlson, B. E. (1990). Adolescent observers of marital violence. *Journal of Family Violence, 5,* 285–299.

Cascardi, M., & O'Leary, K. D. (1992). Depressive symptomatology, self-esteem, and self-blame in battered women. *Journal of Family Violence, 7,* 249–25.

Christian, J. L., O'Leary, K. D., & Vivian, D. (1994). Depressive symptomatology in maritally discordant women and men: The role of individual and relationship variables. *Journal of Family Psychology, 8,* 32–42.

Christopoulos, C., Cohn, D. A., Shaw, D. S., Joyce, S., Sullivan-Hanson, J., Kraft, S. P., & Emery, R. E. (1987). Children of abused women: Adjustment at time of shelter residence. *Journal of Marriage and the Family, 49,* 611–619.

Cummings, J. S., Pellegrini, D. S., Notarius, C. I., & Cummings, E. M. (1989). Children's responses to angry adult behavior as a function of marital distress and history of interparent hostility. *Child Development, 60,* 1035–1043.

Davies, P. T., & Cummings, E. M. (1994). Marital conflict and child adjustment: An emotional security hypothesis. *Psychological Bulletin, 116,* 387–411.

Davis, L. V., & Carlson, B. E. (1987). Observation of spouse abuse: What happens to children? *Journal of Interpersonal Violence, 2,* 278–291.

Doumas, D., Margolin, G., & John, R. (1994). The intergenerational transmission of aggression across three generations. *Journal of Family Violence, 9,* 157–175.

Dutton, M. S. (1996). Working with battered women. *In Session: Psychotherapy in Practice, 2,* 63–80.

Fantuzzo, J. W., DePaola, L. M., Lambert, L., Martino, T., Anderson, G., & Sutton, S. (1991). Effects of interparental violence on the psychological adjustment and competencies of young children. *Journal of Consulting and Clinical Psychology, 59,* 258–265.

Ferraro, K. J., & Johnson, J. M. (1983). How women experience battering: The process of victimization. *Social Problems, 30,* 325–339.

Follingstad, D. R., Laughlin, J. E., Polek, D. S., Rutledge, L. L., & Hause, E. S. (1991). Identification of patterns of wife abuse. *Journal of Interpersonal Violence, 6,* 187–204.

Follingstad, D. R., Rutledge, L. L., Berg, B. J., Hause, E. S., & Polek, D. S. (1990). The role of emotional abuse in physically abusive relationships. *Journal of Family Violence, 5,* 107–120.

Forsstrom-Cohen, B., & Rosenbaum, A. (1985). The effects of parental marital violence on young adults: An exploratory investigation. *Journal of Marriage and the Family, 47,* 467–472.

Frisch, M. B., & MacKenzie, C. J. (1991). A comparison of formally and chronically battered women on cognitive and situational dimensions. *Psychotherapy, 28,* 339–344.

Garmezy, N. (1983). Stressors of childhood. In N. Garmezy & M. Rutter (Eds.), *Stress, coping, and development in children* (pp. 43–84). New York: McGraw-Hill.

Gleason, W. J. (1993). Mental disorders in battered women: An empirical study. *Violence and Victims, 8,* 53–66.

Gondolf, E. W. (1988). The effect of batterer counseling on shelter outcome. *Journal of Interpersonal Violence, 3,* 275–289.

Grych, J. H., & Fincham, F. D. (1990). Marital conflict and children's adjustment: A cognitive-contextual framework. *Psychological Bulletin, 108,* 297–290.

Hansen, M., Harway, M., & Cervantes, N. (1991). Therapist's perceptions of severity in cases of family violence. *Violence and Victims, 6,* 225–234.

Henning, K., Leitenberg, H., Coffey, P., Turner, T., & Bennett, R. T. (1996). Long-term psychological and social impact of witnessing physical conflict between parents. *Journal of Interpersonal Violence, 11,* 35–51.

Hershorn, M., & Rosenbaum, A. (1985). Children of marital violence: A closer look at the unintended victims. *American Journal of Orthopsychiatry, 55,* 260–266.

Holden, G. W., & Ritchie, K. L. (1991). Linking extreme marital discord, child rearing, and child behavior problems: Evidence from battered women. *Child Development, 62,* 311–327.

Holtzworth-Munroe, A. (1988). Causal attributions in marital violence: Theoretical and methodological issues. *Clinical Psychology Review, 8,* 331–344.

Holtzworth-Munroe, A., Smutzler, N., & Bates, L. (1997). A brief review of the research on husband violence. Part III: Sociodemographic factors, relationship factors, and differing consequences of husband and wife violence. *Aggression and Violent Behavior, 2,* 285–307.

Hotaling, G., & Sugarman, D. (1986). An analysis of risk markers in husband to wife violence: The current state of knowledge. *Violence and Victims, 1,* 101–124.

Houskamp, B. M., & Foy, D. W. (1991). The assessment of posttraumatic stress disorder in battered women. *Journal of Interpersonal Violence, 6,* 367–375.

Hudson, W. W., & Mcintosh, S. (1981). The assessment of spouse abuse: Two quantifiable dimensions. *Journal of Marriage and the Family, 43,* 873–885.

Hughes, H. M. (1988). Psychological and behavioral correlates of family violence in child witnesses and victims. *American Journal of Orthopsychiatry, 58,* 77–90.

Hughes, H. M., & Luke, D. A. (1998). Heterogeneity in adjustment among children of battered women. In G. W. Holden, R. Geffner, & E. N. Jouriles (Eds.), *Children exposed to marital violence: Theory, research, and applied issues* (pp. 185–221). Washington, DC: American Psychological Association.

Hughes, H. M., Parkinson, D., & Vargo, M. (1989). Witnessing spouse abuse and experiencing physical abuse: A "double whammy"? *Journal of Family Violence, 4,* 197–209.

Jaffe, P. G., Sudermann, M., & Reitzel, D. (1992). Child witnesses of marital violence. In R. T. Amemrman & M. Hersen (Eds.), *Assessment of family violence: A clinical and legal sourcebook* (pp. 313–331). New York: Wiley.

Jaffe, P. G., Wilson, S. K., & Wolfe, D. A. (1989). Specific assessment and intervention strategies for children exposed to wife battering: Preliminary empirical investigation. *Canadian Journal of Community Mental Health, 7,* 157–163.

Jaffe, P. G., Wolfe, D. A., & Wilson, S. K. (1990). *Children of battered women.* Newbury Park, CA: Sage.

Jaffe, P. G., Wolfe, D. A., Wilson, S. K., & Zak, L. (1986). Family violence and child adjustment: A comparative analysis of girls' and boys' behavioral symptoms.

American Journal of Psychiatry, 143, 74–77.

Jouriles, E. N., Barling, J., & O'Leary, K. D. (1987). Predicting child behavior problems in maritally violent families. *Journal of Abnormal Child Psychology, 15,* 165–173.

Jouriles, E. N. & LeCompte, S. H. (1991). Husbands' aggression toward wives and mothers' and fathers' aggression toward children: Moderating effects of child gender. *Journal of Consulting and Clinical Psychology, 59,* 190–192.

Jouriles, E. N., Murphy, C. M., Farris, A. M., Smith, D. A., Richters, J. E., & Waters, E. (1991). Marital adjustment, parental disagreements about child rearing, and behavior problems in boys: Increasing the specificity of the marital assessment. *Child Development, 62,* 1424–1433.

Jouriles, E. N., Murphy, C. M., & O'Leary, K. D. (1989). Interspousal aggression, marital discord, and child problems. *Journal of Consulting and Clinical Psychology, 57,* 453–455.

Jouriles, E. N., & Norwood, W. D. (1995). Physical aggression toward boys and girls in families characterized by the battering of women. *Journal of Family Psychology, 9,* 69–78.

Jouriles, E. N., Norwood, W. D., McDonald, R., Vincent, J. P., & Mahoney, A. (1996). Physical violence and other forms of marital aggression: Links with children's behavior problems. *Journal of Family Psychology, 10,* 223–234.

Kalmuss, D., & Straus, M. A. (1990). Wife's marital dependency and wife abuse. In M. A. Straus & R. J. Gelles (Eds.), *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families* (pp. 369–382). New Brunswick, NJ: Transaction Publishers.

Kemp, A., Green, B. L., Hovanitz, C., & Rawlings, E. I. (1995). Incidence and correlates of posttraumatic stress disorder in battered women: Shelter and community samples. *Journal of Interpersonal Violence, 10,* 43–55.

Langhinrichsen-Rohling, J., Neidig, P., & Thorn, G. (1995). Violent marriages: Gender differences in levels of current violence and past abuse. *Journal of Family Violence, 10,* 159–175.

Loseke, D. R., & Cahill, S. E. (1984). The social construction of deviance: Experts on battered women. *Social Problems, 31,* 296–310.

Luther, S. S., & Zigler, E. (1991). Vulnerability and competence: A review of research on resilience in childhood. *American Journal of Orthopsychiatry, 61,* 6–22.

Malamuth, N. M., Linz, D., Heavey, C. L., Barnes, G., & Acker, M. (1995). Using the confluence model of sexual aggression to predict men's conflict with women: A ten year follow-up study. *Journal of Personality and Social Psychology, 69,* 353–369.

McCloskey, L. A., Figueredo, A. J., & Koss, M. P. (1995). The effects of systemic family violence on children's mental health. *Child Development, 66,* 1239–1261.

McCord, J. (1979). Some child rearing antecedents to criminal behavior in adult men. *Journal of Personality and Social Psychology, 37,* 1477–1486.

McDonald, R., & Jouriles, E. N. (1991). Marital aggression and child behavior problems: Research findings, mechanisms, and intervention strategies. *The Behavior Therapist, 14,* 189–192.

Miller, T. R., Handal, P. J., Gilner, F. H., & Cross, J. F. (1991). The relationship of abuse and witnessing violence on the Child Abuse Potential Inventory with Black adolescents. *Journal of Family Violence, 6,* 351–363.

Mitchell, R. E., & Hodson, C. A. (1983). Coping with domestic violence: Social support and psychological health among battered women. *American Journal of Community Psychology, 11,* 629–654.

Murphy, C., & Cascardi, M. (1993). Psychological aggression and abuse in marriage. In R. L. Hampton, T. P. Gollotta, G. R. Adams, E. H. Potter, & R. P. Weissberg (Eds.), *Family violence: Prevention and treatment* (pp. 86–112). Newbury Park, CA: Sage.

Murphy, C. M., & O'Leary, K. D. (1989). Psychological aggression predicts physical aggression in early marriage. *Journal of Consulting and Clinical Psychology, 57,* 579–582.

O'Brien, M., John, R. S., Margolin, G., & Erel, O. (1994). Reliability and diagnostic efficacy of parents' reports regarding children's exposure to marital aggression. *Violence and Victims, 9,* 45–62.

O'Brien, M., Margolin, G., John, R. S., & Krueger, L. (1991). Mothers' and sons' cognitive and emotional reactions to simulated marital and family conflict. *Journal of Consulting and Clinical Psychology, 59,* 692–703.

O'Keefe, M. (1994). Linking marital violence, mother-child/father-child aggression, and child behavior problems. *Journal of Family Violence, 9,* 63–78.

O'Keefe, M. (1995). Predictors of child abuse in maritally violent families. *Journal of Interpersonal Violence, 10,* 3–25.

Okun, L. E. (1986). *Woman abuse: Facts replacing myths.* Albany, NY: State University of New York Press.

Pagelow, M. D. (1982). Children in violent families: Direct and indirect victims. In S. Hill & B. J. Barnes (Eds.), *Young children and their families* (pp. 47–72). Lexington, MA: D. C. Health.

Pence, E. (1989). Batterer programs: Shifting from community collusion to community confrontation. In P. L. Caesar & L. K. Hamberger (Eds.), *Treating men who batter* (pp. 24–50). New York: Springer.

Perilla, J. L., Bakeman, R., & Norris, F. H. (1994). Culture and domestic violence: The ecology of abused Latinas. *Violence and Victims, 9,* 325–339.

Reynolds, R., & Siegle, E. (1959). A study of casework with sado-masochistic marriage partners. *Social Casework, 40,* 545–551.

Riggs, D. S., Kilpatrick, D. G., & Resnick, H. S. (1992). Long-term psychological distress associated with marital rape and aggravated assault: A comparison to other crime victims. *Journal of Family Violence, 7,* 283–296.

Rodenburg, F., & Fantuzzo, J. (1993). The measure of wife abuse: Steps toward the development of a comprehensive assessment technique. *Journal of Family Violence, 8,* 203–228.

Rosenfeld, B. D. (1992). Court-ordered treatment of spouse abuse. *Clinical Psychology Review, 12,* 205–226.

Rosenbaum, A., & O'Leary, K. D. (1981). Children: The unintended victims of marital violence. *American Journal of Orthopsychiatry, 51,* 692–699.

Rosenberg, M. (1987). Children of battered women: The effects of witnessing violence on their social problem-solving abilities. *The Behavior Therapist, 4,* 85–89.

Rossman, B. B. R., & Rosenberg, M. (1992). Family stress and functioning in children: The moderating effects of children's beliefs about their control over parental conflict. *Journal of Child Psychology and Psychiatry, 33,* 699–715.

Rutter, M. (1987). Psychosocial resilience and protective mechanisms. *American Journal of Orthopsychiatry, 57,* 316–331.

Saltzman, L. E., & Mercy, J. A. (1993). Assaults between intimates: The range of relationships involved. In A. V. Wilson (Ed.), *Homicide: The victim/offender connection* (pp. 65–74). Eastern Kentucky U, OH: Anderson Publishing.

Sato, R. A. & Heiby, E. M. (1991). Correlates of depressive symptoms among battered women. *Journal of Family Violence, 7,* 229–245.

Saunders, D. G. (1994). Posttraumatic stress symptom profiles of battered women: A comparison of survivors in two settings. *Violence and Victims, 9,* 31–43.

Schutte, N. S., Malouff, J. M., & Doyle, J. S. (1988). The relationship between characteristics of the victim, persuasive techniques of the batterer, and returning to a battered relationship. *The Journal of Social Psychology, 128,* 605–610.

Schwartz, M. D. (1988). Marital status and women abuse theory. *Journal of Family Violence, 3,,* 239–248.

Schwartz, M. D., & Mattley, C. L. (1993). The battered women scale and gender identities. *Journal of Family Violence, 8,* 277–287.

Shepard, M. F., & Campbell, J. A. (1992). The abusive behavior inventory A measure of psychological and physical abuse. *Journal of Interpersonal Violence, 7,* 291–305.

Silvern, L., Karyl, J., Waelde, L., Hodges, W. F., Starek, J., Heidt, E., & Min, K. (1995). Retrospective reports of parental partner abuse: Relationships to depression, trauma symptoms and self-esteem among college students. *Journal of Family Violence, 10,* 177–201.

Simons, R. L., Johnson, C., Beaman, J., & Conger, R. (1993). Explaining Women's Double Jeopardy: Factors that mediate the association between harsh treatment as a child and violence by a husband. *Journal of Marriage and the Family, 55,* 713–723.

Snyder, D. K., & Fruchtman, L. A. (1981). Differential patterns of wife abuse: A data-based typology. *Journal of Consulting and Clinical Psychology, 49,* 878–885.

Spaccarelli, S., Coatsworth, J. D., & Bowden, B.S. (1995). Exposure to serious family violence among incarcerated boys: Its association with violent offending and potential mediating variables. *Violence and Victims, 10,* 163–182.

Spaccarelli, S., Sandler, I. N., & Roosa, M. (1994). History of spouse violence against mother: Correlated risks and unique effects in child mental health. *Journal of Family Violence, 9,* 79–98.

Stagg, V., Wills, G. D., & Howell, M. (1989). Psychopathology in early childhood witnesses of family violence. *Topics in Early Childhood Special Education, 9,* 73–87.

Sternberg, K. J., Lamb, M. E., Greenbaum, C., Cicchetti, D., Dawud, S., Cortes, R. M., Krispin, O., & Lorey, F. (1993). Effects of domestic violence on children's problems and depression. *Developmental Psychology, 29,* 44–52.

Stets, J. E. (1990). Verbal and physical aggression in marriage. *Journal of Marriage and the Family, 52,* 501–514.

Stets, J.E., & Straus, M. A. (1990). Gender differences in reporting marital violence and its medical and psychological consequences. In M. A. Straus & R. J. Gelles (Eds.), *Physical violence in American families: Risk factors and adaptation to violence in 8,145 families* (pp. 151–166). New Brunswick, NJ: Transaction Publishers.

Stewart, M. A., & deBlois, C. S. (1981). Wife abuse among families attending a child psychiatry clinic. *Journal of the American Academy of Child Psychiatry, 20,* 845–862.

Straus, M. A., & Gelles, R. J. (1986). Societal change and change in family violence from 1975 to 1985 as revealed by two national surveys. *Journal of Marriage and the Family, 48,* 465–479.

Straus, M. A., & Gelles, R. J. (1990). *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families.* New Brunswick, NJ: Transaction Publishers.

Straus, M. A., Gelles, R. J., & Steinmetz, S. K. (1980). *Behind closed doors: Violence in the American family.* Garden City, NY: Doubleday Press.

Strube, M. J. (1988). The decision to leave an abusive relationship: Empirical evidence and theoretical issues. *Psychological Bulletin, 104,* 236–250.

Strube, M. J., & Barbour, L. S. (1983). The decision to leave an abusive relationship: Economic dependence and psychological commitment. *Journal of Marriage and the Family, 45,* 785–793.

Strube, M. J., & Barbour, L. S. (1984). Factors related to the decision to leave an abusive relationship. *Journal of Marriage and the Family, 46,* 837–844.

Tolman, R. M. (1989). The development of a measure of psychological maltreatment of women by their male partners. *Violence and Victims, 4,* 159–177.

Walker, L. E. (1991). Post-traumatic stress disorder in women: Diagnosis and treatment of battered women syndrome. *Psychotherapy, 28,* 21–29.

Warren, J., & Lanning, W. (1993). Sex role beliefs, control, and social isolation of battered women. *Journal of Family Violence, 7,* 1–8.

Wilson, M., & Daly, M. (1993). Spousal homicide risk and estrangement. *Violence and Victims, 8,* 3–16.

Wolfe, D. A., Jaffe, P. G., Wilson, S. K., & Zak, L. (1985). Children of battered women: The relation of child behavior to family violence and maternal stress. *Journal of Consulting and Clinical Psychology, 53,* 657–665.

Wolfe, D. A., Zak, L., Wilson, S. K., & Jaffe, P. G. (1986). Child witnesses to violence between parents: Critical issues in behavioural and social adjustment. *Journal of Abnormal Child Psychology, 14,* 95–104.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.21
Victims of Child Maltreatment

SANDY K. WURTELE
University of Colorado, Colorado Springs, CO, USA

9.21.1 INTRODUCTION 341
9.21.1.1 Defining Child Maltreatment 341
9.21.1.2 Nature of the Problem 342
9.21.1.3 Incidence and Consequences 342
9.21.2 ASSESSMENT 342
9.21.2.1 Purpose and Process 342
9.21.2.2 Assessing Developmental Domains 343
9.21.2.2.1 Physical domain 343
9.21.2.2.2 Cognitive domain 345
9.21.2.2.3 Socioemotional domain 346
9.21.2.2.4 Behavioral domain 348
9.21.2.2.5 Psychosexual domain 348
9.21.3 TREATMENT 348
9.21.3.1 Purpose and Process 348
9.21.3.1.1 Physical domain 349
9.21.3.1.2 Cognitive domain 350
9.21.3.1.3 Socioemotional domain 350
9.21.3.1.4 Behavioral domain 351
9.21.3.1.5 Psychosexual domain 352
9.21.4 PREVENTION 352
9.21.5 FUTURE DIRECTIONS 353
9.21.6 SUMMARY 354
9.21.7 REFERENCES 354

9.21.1 INTRODUCTION

Child maltreatment is a complex, tragic, and common social problem. Shamefully, ours is a society where children are hit, strangled, shaken, abandoned, starved, and raped. Every day, at least three children die from abuse and neglect in the USA. Numerous studies have implicated child maltreatment in the etiology of a broad range of physical, psychological, and social problems for children that can follow them into adulthood. Clearly, child maltreatment represents a significant physical and mental health concern for researchers and practitioners.

The purpose of this chapter is to provide an overview of the nature and prevalence of the three major types of child maltreatment: physical abuse, sexual abuse, and neglect. The chapter summarizes child-focused assessment strategies and psychosocial interventions with victims of child maltreatment. Finally, the chapter concludes by offering suggestions for moving research and practice forward.

9.21.1.1 Defining Child Maltreatment

It is important to distinguish between types of child maltreatment because the definition of child maltreatment will affect how cases are classified, how eligibility for social services is determined, and how treatment decisions are made. Definitions will also impact research findings. There is great variability among definitions of physical abuse, sexual abuse, and neglect, and the task of determining when adult actions are ultimately abusive remains complex.

In this chapter, physical abuse involves acts of commission, which usually involve inflicting bodily injury on the child. Physical abuse can include scalding, shaking, beating the child with an object, severe physical punishment, and a rare form called Munchausen by proxy syndrome, where an adult feigns or induces illness in a child. Sexual abuse has been defined as contacts or interactions between a child and an adult when the child is used for the sexual stimulation of the perpetrator or another person (National Center on Child Abuse and Neglect, 1978). Offenders may include older children (usually five years age difference) and adults other than parents (e.g., teachers, coaches, strangers). Acts defining sexual abuse include penile penetration (oral, anal, or genital), fondling of genital areas, exposure to indecent acts, or involvement in child pornography. Neglect relates to acts of omission by the caregiver, and is evidenced by failing to meet the child's needs for adequate food, clothing, shelter, hygiene, supervision, education, and/or medical care. The most extreme form of neglect is failure to thrive or growth deficiency, where the infant fails to gain weight due to malnutrition. It has become clear that maltreatment subtypes overlap and that pure subtypes are rare. In this chapter, the term child maltreatment is used to reflect the variability of this phenomenon.

9.21.1.2 Nature of the Problem

Who are the victims of these forms of maltreatment? Victims of child maltreatment can be of any age from zero to 18 years. According to the US Department of Health and Human Services (1996), in 1994 about 27% of victims were three years old or younger. Children aged four to six years represented about 20% of the victims, children aged seven to nine years represented 17%, children aged 10 to 12 years represented 15%, and teenagers (13–18 years old) made up about 21% of the victims. The highest risk for death from physical abuse and neglect is among children less than two years old; 41% of child abuse fatality victims are under the age of one (US Advisory Board on Child Abuse and Neglect, 1995). The number of victims of neglect rapidly decreases with increasing age for both male and female cases, whereas sexual abuse cases increase with age.

Victims of maltreatment include both boys and girls of different races. According to the US Department of Health and Human Services (1996), in 1994, 52% of the victims of maltreatment were female, and 47% were male (with 1% unknown). Victims of sexual abuse are predominantly female. Maltreated children are primarily white (56%), with African-American children representing 26%, Hispanic 9%, and native American victims making up about 2%. African-American children are relatively more likely to be fatally injured or neglected, and less likely to be sexually abused. Caucasian families are more likely to engage in physical abuse, sexual abuse, or a combination of physical abuse and neglect.

9.21.1.3 Incidence and Consequences

In 1994, an estimated 2.9 million children were reported to public child protective service agencies as alleged victims of child maltreatment (US Department of Health and Human Services, 1996). Neglect was the most common form of reported child maltreatment, accounting for 53% of cases, followed by physical abuse (26%), and sexual abuse (14%). The incidence and tragic consequences of child maltreatment prompted the US Advisory Board on Child Abuse and Neglect (1990) to claim that the problem of child abuse and neglect has reached epidemic proportions in the USA and that it constitutes a national emergency.

Child maltreatment is all too common, and much too tragic. Each day in the USA, three to five children die as a result of either physical abuse, neglect, or both (Daro; 1995; US Advisory Board on Child Abuse and Neglect, 1995). Child maltreatment is now the leading cause of death for children under the age of four (US Advisory Board on Child Abuse and Neglect, 1995). Each year, 150 000 children are seriously injured and 18 000 are permanently disabled by abuse or neglect. According to estimates provided by the General Accounting Office, the services required for children who have been abused or neglected cost more than $500 million annually. In addition to the economic costs, the human costs are also distressing.

9.21.2 ASSESSMENT

9.21.2.1 Purpose and Process

Conducting a thorough assessment of victims of child maltreatment is important for both researchers and practitioners. Researchers need to obtain a reliable and valid assessment of children's experiences and their current level of functioning to better understand the psychological sequelae of child maltreatment. Practitioners likewise need to understand children's experiences and reactions, but additional objectives for assessment are to devise appropriate treatment interventions and monitor treatment progress. Assessment is also important for forensic purposes, but this type of evaluation is complex and beyond the scope of this chapter.

There are several measurement issues that need to be considered throughout the assessment process. The assessment process should be grounded in the field of developmental psychopathology, which is a perspective that examines the evolution of psychological disturbances in the context of development. This approach studies the developmental tasks that are compromised by stressful events, and utilizes measures that are linked to developmental tasks. Given that child maltreatment is associated with a variety of psychological sequelae, assessment in each developmental domain (i.e., physical, cognitive, socioemotional, behavioral) is recommended.

What one chooses to measure is also determined by one's theoretical framework. For example, an ecological/transactional model of child maltreatment (Cicchetti & Lynch, 1993) suggests that assessment needs to occur at all levels of children's ecologies (i.e., the macrosystem, exosystem, microsystem, and ontogenic level). In addition to assessing areas of deficits, the assessment should also elucidate areas of strength. Understanding the child's and family's strengths should inform one's choice of effective treatment interventions. Another general measurement issue is from whom information about the case is obtained. Although many researchers and clinicians rely exclusively on reports provided by child protective service workers, there is much to be gained by using multiple sources of data, including parents, teachers, and children.

The assessment process begins with the clinical interview, during which the clinician makes informal observations of the child. The purpose of assessment should be explained to the family, and the clinician must discuss the limits of confidentiality. Both parents and children should be asked questions about the circumstances surrounding the maltreatment. Behaviorally oriented clinicians often use a functional analysis model; asking about the antecedents to the abuse, the abusive behavior itself, and the consequences for the abusive behavior. Gathering information about the nature and extent of the abuse or neglect is important in order to gain insight into the clinical issues most salient to the case and to prescribe the most appropriate therapeutic intervention.

Risk assessment should occur throughout this initial assessment process and on an ongoing basis. Children cannot benefit from therapy when they continue to be victimized outside the clinic setting. The family's willingness to discuss the maltreatment is important to monitor. With defensive parents, more time needs to be spent developing rapport before trying to conduct assessments. The clinician should relate to the parents and their children in a warm and supportive manner, but not condone the abusive or neglectful behavior. In addition to the identified abuse, it is critically important to determine whether other types of maltreatment or family violence are occurring.

9.21.2.2 Assessing Developmental Domains

A number of assessment devices have been utilized with victims of child maltreatment. A list of measures used to assess the effects of child maltreatment for each developmental domain is contained in Table 1. Professionals will not need to use all of these measures. Hopefully, this list will serve to supplement one's current assessment battery, and to encourage assessment of each developmental domain. Readers are urged to select those measures that best fit the characteristics of their clients, time constraints, and their own expertise.

9.21.2.2.1 Physical domain

Children's medical records should be reviewed to obtain health history data and results of medical or forensic evaluations. Physical or neurological damage may relate to academic and social functioning and may affect treatment outcome. Given the frequent finding of developmental delays, a developmental assessment may be necessary, and several instruments are available to measure motor, communication, self-care, and physical development (see Table 1). Somatic complaints have been mentioned frequently among sexually abused children, which may be related to a heightened focus on the physical self (Friedrich & Schafer, 1995). Somatic complaints have usually been assessed by using a checklist of symptoms or by employing the Child Behavior Checklist (CBCL; Achenbach & Edelbrock, 1983) which has a somatic complaints subscale.

Very little is known about the psychophysiological consequences of child maltreatment. Researchers in the physical abuse area have sampled maltreated childrens' saliva to measure cortisol secretion (e.g., Hart, Gunnar, & Cicchetti, 1995), and have monitored childrens' heart rates in response to affect-ladened stimuli

Table 1 Measures used to assess effects of child maltreatment.

Developmental domain/tests
Physical
Bayley Scales of Infant Development
Denver Developmental Screening Test
Learning Accomplishment Profile
McCarthy Scales of Children's Abilities
Minnesota Child Development Inventory
Vineland Adaptive Behavior Scales-Revised
Child Behavior Checklist
Cognitive
Intelligence
WPSSI-R
WISC-III
Stanford–Binet
K-ABC
Language
Peabody Picture Vocabulary Test-Revised (Dunn & Dunn; 1981)
Test of Auditory Comprehension of Language-Revised (Carrow-Woolfolk, 1985)
Expressive One-word Picture Vocabulary (Gardner, 1979)
Academic Achievement
WRAT-R
Woodcock Mastery Tests
Iowa Test of Basic Skills
Child Behavior Checklist (Achenbach & Edelbrock, 1983)
Self-esteem
Piers–Harris Children's Self-concept Scale (Piers & Harris, 1984)
Pictorial Scale of Perceived Competence and Social Acceptance for Young Children (Harter & Pike, 1984)
Self-perception Profile for Children (Harter, 1985)
Coopersmith Self-esteem Inventory (Coopersmith, 1986)
Tennessee Self-concept Scale (Fitts, 1965)
Rosenberg Self-esteem Scale (Rosenberg, 1979)
Self-perception Profile for Adolescents (Harter, 1988)
Attributions
Children's Attributional Style Questionnaire (Kaslow, Tannenbaum, & Seligman, 1978)
Hopelessness Scale for Children (Kazdin et al., 1983)
Locus of Control Scale for Children (Nowicki & Strickland, 1973)
Stanford Internal–External Locus of Control Scale (Mischel, Zeiss, & Zeiss, 1974)
Children's Impact of Traumatic Events Scales (Wolfe et al., 1991)
Children's Attributions and Perceptions Scale (Mannarino et al., 1994)
Coping Styles
Ways of Coping Checklist-Revised (Folkman & Lazarus, 1980)
Child Dissociative Checklist (Putnam et al., 1993)
Socioemotional
Social Behavior/Skills
Child Behavior Checklist
Preschool Behavior Questionnaire (Behar & Stringfield, 1974)
Teacher–Child Rating Scale (Hightower, Spinell, & Lotyczewski, 1985)
Behavior Problem Checklist (Quay & Peterson, 1987)
Matson Evaluation of Social Skills with Youngsters (Matson, Esveldt-Dawson, & Kazdin, 1983)
Self-perception Profile for Children (Harter, 1985)
Friendship Questionnaire (Bierman & McCauley, 1987)
Social Problem-solving Measure (Pettit et al., 1988)
Depression
Children's Assessment Schedule (Hodges, Kline, McKnew, & Cytryn, 1992)
Diagnostic Interview for Children and Adolescents (Welner, Reich, Herjanic, Jung, & Amado, 1987)
Diagnostic Interview Schedule for Children-Revised (Shaffer et al., 1993)
Interview Schedule for Children (Kovacs, 1983)
Schedule of Affective Disorders and Schizophrenia-Child Version (Chambers et al., 1985)
Children's Depression Inventory (Kovacs, 1981, 1992)
Beck Depression Inventory (Beck, 1976)
Reynolds' Adolescent Depression Scale (Reynolds, 1987)

Table 1 (continued)

Developmental domain/tests
Anxiety
Fear Survey Schedule for Children-Revised (Ollendick, 1983)
Revised Children's Manifest Anxiety Scale (Reynolds & Richmond, 1985)
State–Trait Anxiety Inventory for Children (Spielberger, 1973)
Children's Impact of Traumatic Events Scale-Revised (Wolfe et al., 1991)
Children's Post-traumatic Stress Disorder Inventory (Saigh, 1989)
Behavioral
Child Behavior Checklist (Achenbach & Edelbrock, 1983)
Louisville Behavior Checklist (Miller, 1979)
Eyberg Child Behavior Inventory (Eyberg & Ross, 1978)
Revised Behavior Problem Checklist (Quay & Peterson, 1987)
Preschool Behavior Questionnaire (Behar & Stringfield, 1974)
Psychosexual
Child Sexual Behavior Inventory (Friedrich et al., 1992)
Children's Impact of Traumatic Events Scale-Revised (Wolfe et al., 1991)
Trauma Symptom Checklist for Children (Briere, 1996)

(Hill, Bleichfield, Brunstetter, Hebert, & Steckler, 1989). There is some evidence that sexual abuse may elevate cortisol levels and advance the onset of puberty among girls (Putnam & Trickett, 1993), and disrupt the hypothalamic–pituitary–adrenal axis (De Bellis et al., 1994). The lack of sensory experience (as during neglect), or excess exposure to traumatic experiences (as during physical or sexual abuse), can alter neuronal neurochemistry, cytoarchitecture, and, ultimately, brain structure and function in severely traumatized children (Perry, Pollard, Blakley, Baker, & Vigilante, 1995). Although continued research in this area will undoubtedly shed light on abuse-related symptomatology and help us to understand how children adapt to various traumatic experiences, the methodological limitations associated with biological measures will most likely limit their use among practicing clinicians.

9.21.2.2.2 Cognitive domain

Physically abused and neglected children often demonstrate deficits on measures of intelligence (Crouch & Milner, 1993). Intellectual impairments may be due to direct insult, poor nutrition, or insufficient stimulation. General intellectual functioning can be assessed by an individually administered psychometric test of intelligence (see Table 1). A language assessment is often important to conduct, given that maltreated children have been found to exhibit delays in both expressive and receptive language (Vondra, Barnett, & Cicchetti, 1990), as well as deficits in their abilities to maintain dialogues and discuss or identify their own feelings (Coster, Gersten, Beeghly, & Cicchetti, 1989).

Maltreated children perform poorly on academic achievement tests, receive low grades, and exhibit problematic classroom conduct (Eckenrode, Laird, & Doris, 1993; Kendall-Tackett & Eckenrode, 1996; Leiter & Johnsen, 1994). A standard achievement test, along with teacher reports and school records, can provide information on academic functioning of school-aged children. Teachers themselves can provide a wealth of information about the child's academic functioning in the classroom and his/her social competence with peers. Children's perceptions of their academic competence can be assessed using the cognitive competence subscale of the Pictorial Scale of Perceived Competence and Social Acceptance for Young Children (Harter & Pike, 1984).

Research has shown that maltreated children evidence deficits in self-concept and self-esteem (Kaufman & Cicchetti, 1989; Toth, Manly, & Cicchetti, 1992), although the findings with sexually abused children have been mixed, with some researchers finding lower self-esteem among incest victims (Hotte & Rafman, 1992) and others finding no evidence for lower self-esteem (Mannarino, Cohen, Smith, & Moore-Motily, 1991). These conflicting results may be due to differences in the ages of the victims and/or the types of abuse experienced.

Another concern is whether victims overestimate the amount of danger in the world and underestimate their self-efficacy or self-worth. Some researchers have assessed the ways in which children attribute causality for good and bad events using the Children's Attributional Style Questionnaire (CASQ; Kaslow, Tannenbaum, & Seligman, 1978). Negative expectations about the future can be assessed by using the Hopelessness Scale for Children (HPLS; Kazdin, French, Unis, Esveldt-Dawson, & Sherrick, 1983). Table 1 also contains measures of generalized locus of control used with maltreated children.

Attributions that victims make about themselves and the abuse are also important for understanding children's reactions and for guiding treatment interventions. Sexual abuse victims frequently have distorted cognitions about themselves and the abuse. The Children's Impact of Traumatic Events Scale (CITES; Wolfe, Gentile, Michienzi, Sas, & Wolfe, 1991) contains subscales designed to measure children's perceptions and attributions concerning their sexual abuse. The Children's Attributions and Perceptions Scale (CAPS; Mannarino, Cohen, & Berman, 1994) was also designed to measure abuse-related perceptions and attributions among sexually victimized children, although it has also been used with physically abused children (Kolko, 1996). Other ways to explore a child's attributions of responsibility include direct questioning, asking children what they would do differently if they were molested again, and using projective techniques to help children talk about their attributions indirectly (Celano, 1992).

Several researchers have suggested that the coping styles victims use play important roles as mediators of sexual abuse sequelae (e.g., Spaccarelli, 1994). For example, Johnson and Kenkel (1991) assessed coping strategies of sexually abused adolescent females by using a modified version of the Ways of Coping Checklist-Revised (WCCL-R; Folkman & Lazarus, 1980). Teenagers were asked to rate their use of such strategies as problem-focused coping, wishful thinking, distancing/detachment, emphasizing the positive, tension-reduction, self-isolation, and support seeking. Dissociation as a coping strategy has also been assessed. Putnam, Helmers, and Trickett (1993) developed the Child Dissociative Checklist (CDC) which measures several domains of dissociative behaviors (e.g., amnesias, trance states, hallucinations). Research into the coping processes of children exposed to traumatic experiences has only begun (e.g., Compas, 1987; Wertlieb, Weigel, & Feldstein, 1987) and will hopefully inform the child maltreatment field as to the impact coping behaviors might have on the development and course of abuse-related symptomatology.

9.21.2.2.3 Socioemotional domain

According to Cicchetti (1987), childhood maltreatment should be conceived of as relational psychopathology, arising from a dysfunction in the parent–child–environment transactional system. This type of relational disorder can be seen during the first year of life, when the development of a secure attachment relationship with the primary caregiver occurs. Several investigators have found that maltreated infants and toddlers are significantly more likely to form insecure attachment relationships (e.g., Carlson, Cicchetti, Barnett, & Braunwald, 1989). There have been no studies to date looking at attachment in a sexually abused population, although Alexander (1992) has argued that sexual abuse often develops out of the context of an insecure parent–child attachment.

In addition to negative relationships, it is also important to determine if the child has any positive relationships (e.g., with a grandparent, teacher, or sibling). If positive relationships exist, this suggests that the child may be more open to forming a therapeutic alliance, a positive outcome may be optimized, and it also suggests the possibility that this significant other might provide assistance during the therapeutic process. Kaufman (1991) used a modified form of the Arizona Social Support Interview Schedule (ASSIS; Barrera, 1980) to assess the extent and quality of the child's social supports.

These early attachment difficulties are likely to be carried forward through development, affecting the subsequent developmental task of forming effective peer relations. Maltreated children, especially physically abused children, are at risk for poor peer relations (Conaway & Hansen, 1989). Compared with nonmaltreated children, physically abused or neglected preschool-aged children tend to be more aggressive, less prosocial, have fewer positive interactions, and are more likely to respond to distressed peers with aggressive behaviors instead of concern, empathy, or sadness. Maltreated preschoolers also show developmental delays in their play behavior (Alessandri, 1991). School-aged maltreated children have been shown to have lower peer status, are viewed by their peers as aggressive and uncooperative, and evidence less reciprocity in friendship relations (Dodge, Pettit, & Bates, 1994; Rogosch & Cicchetti, 1994; Salzinger, Feldman, Hammer, & Rosario, 1993). With peers, physically abused children also show a negative attributional bias toward inferring negative and hostile intent, which increases the likelihood of aggressive responding (Weiss, Dodge, Bates, & Pettit, 1992), but evidence is mixed regarding whether they lack effective social problem-solving skills (Levendosky, Okun, & Parker, 1995). Neglected children tend to withdraw from interaction with peers (Crittenden, 1992), making them at risk for increasing isolation and rejection.

Problems relating to peers have been assessed using a variety of measures. For example, children have been asked to nominate peers who are cooperative, disruptive, and aggressive, nominate classmates to portray specified roles in a class play, such as "picks on other kids," or "makes new friends easily" (Rogosch & Cicchetti, 1994), or to rate on five-point scale how much they like and dislike their classmates (Dodge et al., 1994). Play and social behavior in maltreated children have also been observed and then coded into types of play (Alessandri, 1991) or scored according to how much aggressive or prosocial behavior is exhibited during structured or unstructured play (Howes & Eldredge, 1985). Teachers can also provide valuable information about children's behaviors in the classroom and when interacting with peers. Teacher-completed and self-report measures of social behaviors and social skills are included in Table 1. In addition, maltreated children's social problem-solving skills have been assessed using the Social Problem-Solving measure (SPS; Pettit, Dodge, & Brown, 1988) which presents hypothetical peer scenarios and asks children to generate solutions to each scenario.

A related impairment within this domain is that maltreated children often have affect-regulation deficits. They have difficulty understanding or accepting emotions in themselves and others. Physically abused (and neglected) children perform more poorly than controls on measures of affective and cognitive role-taking, social sensitivity, the ability to discriminate emotions in others, and in their ability to respond empathically to others in distress (Howes & Eldredge, 1985; Main & George, 1985). The problem with assessing this developmental impairment is the absence of standardized measures. Instead, researchers have used audiotaped skits, puppet play procedures, and direct observation of children; procedures which most clinicians will find too time intensive for routine use. The development of a comprehensive assessment battery of emotional regulation problems of maltreated children is most needed.

Research evidence also supports a relationship between maltreatment in childhood and such internalizing psychological problems as depression and anxiety. Physically abused children evidence higher levels of depressive symptomatology compared with nonmaltreated children (Toth et al., 1992), although the findings with sexually abused children are less consistent, with some researchers finding sexually abused children to be more depressed (e.g., Wozencraft, Wagner, & Pellegrin, 1991), and others finding no support for elevated levels of depression (Mannarino, Cohen, & Gregor, 1989). To elicit information about the child's depressive symptoms, along with other aspects of the child's behavior, thinking, and mood, various structured interviews have been developed. The most common structured interviews used with maltreated children are contained in Table 1, which also includes instruments for measuring self-reported depression. Depressive symptomatology can also be revealed in children's Rorschach responses (Shapiro, Leifer, Martone, & Kassem, 1990). Maltreated (especially sexually abused) children also evidence heightened levels of anxiety. General anxiety problems can be assessed using the instruments listed in Table 1.

Given that sexually abused children often exhibit symptoms of re-experiencing the trauma, avoidance of abuse-related stimuli, diminished responsiveness to the environment, and increased arousal/anxiety, many researchers have conceptualized sexual abuse as possibly leading to post-traumatic stress disorder (PTSD; Miller-Perrin & Wurtele, 1990). Indeed, significant percentages of sexually abused children have met the diagnostic criteria for PTSD. For example, in their extensive literature review of the impact of sexual abuse on children, Kendall-Tackett, Williams, and Finkelhor (1993) found that 53% of sexually abused children exhibited PTSD symptomatology.

To assess PTSD symptoms via parent report, Wolfe, Gentile, and Wolfe (1989) selected 20 PTSD-like behaviors from the CBCL to comprise a CBCL-PTSD scale. These items measure such problems as difficulty concentrating, difficulty sleeping, somatic complaints, and obsessive thoughts. The Children's Impact of Traumatic Events Scale-Revised (CITES-R; Wolfe et al., 1991) assesses children's perceptions and attributions concerning their own sexual abuse. It includes a subscale which measures such PTSD symptoms as intrusive thoughts, avoidance, hyperarousal, and sexual anxiety. A checklist of PTSD symptoms was used by McLeer, Deblinger, Atkins, Foa, and Ralphe (1988). Raters indicate whether the child exhibits any behaviors listed in the three subcategories: (i) re-experiencing behaviors (e.g., flashbacks, nightmares); (ii) avoidant behaviors (e.g., avoidance of people or places associated with the abuse); and (iii) autonomic hyperarousal (e.g., irritability, difficulty falling asleep). Finally, the Children's PTSD Inventory has been developed by Saigh (1989) to be used in a structured interview format. Existing measures rely heavily upon verbal and reading skills. With younger children, evidence of PTSD symptoms can be gleaned from children's drawings or observation of the child's play (Pynoos & Eth, 1986). Given that assessing stress-related reactions in victimized children is an important step in the assessment process, it is hoped that researchers will continue their efforts to validate these measures and observational approaches.

9.21.2.2.4 Behavioral domain

Behavior problems appear to be very common among maltreated (especially physically abused) children, including heightened aggressive behavior, noncompliance, impulsivity, and conduct problems. The aggressiveness of physically abused children often extends into delinquency during adolescence and also violent criminal acts during adulthood (Widom, 1989a, 1989b). Physical aggression and antisocial behavior have been assessed using observational methods, or more frequently by having parents or staff complete instruments measuring behavior problems. Several well-validated instruments are available to identify problem areas within the behavioral domain; five of them are listed in Table 1.

9.21.2.2.5 Psychosexual domain

Effects of maltreatment on this domain have only been examined with respect to sexually victimized children. Over the course of abuse, a child's sexuality becomes distorted; a process Finkelhor and Browne (1985) refer to as traumatic sexualization. Sexually abused children are much more likely to display behaviors indicative of abnormal or precocious sexual development. There is also a strong relationship between child sexual abuse and adolescent pregnancy (Boyer & Fine, 1992; Rainey, Stevens-Simon, & Kaplan, 1995). In their review of the literature, Kendall-Tackett et al. (1993) noted that sexuality problems, in addition to PTSD symptomatology, were the only sets of symptoms that consistently discriminated sexually abused children from nonabused, clinic-referred children. These researchers estimated that 38% of sexually abused children exhibited sexuality problems.

In order to assess sexuality problems, researchers have most often used either the sexual behavior subscales of the CBCL, or the Child Sexual Behavior Inventory (CSBI), a parent-report measure for children aged two to 12 (Friedrich et al., 1992). Other researchers have created lists of sexual symptoms, and then asked parents to indicate if the behaviors have occurred (e.g., White, Halpin, Strom, & Santilli, 1988).

In addition to questionnaires, sexuality problems have also been assessed using drawings, doll play, and self-report. Human figure drawings of sexually abused children have been characterized by either exaggerated or minimized genitalia (e.g., Burgess, 1988; Hibbard, Roghmann, & Hoekelman, 1987). Regarding play behaviors, sexually abused children are more likely to demonstrate sexual behaviors with anatomically detailed dolls (Jampole & Weber, 1987), although nonabused children have also been observed to show some sexual play with these dolls (Everson & Boat, 1990). Finally, two self-report instruments tap sexual issues: CITES-R (Wolfe et al., 1991) and the Trauma Symptom Checklist for Children (TSCC; Briere, 1996).

9.21.3 TREATMENT

9.21.3.1 Purpose and Process

Treating victimized and neglected children presents several unique challenges. Given the often abhorrent acts done to children, the first challenge is being able to form a therapeutic alliance with the parent who has abused or neglected the child. Then the characteristics of many abusive and neglecting families can make this work difficult, as they may be poorly educated, deny that they need help or be mandated to participate in treatment, or have chaotic lifestyles that can disrupt the treatment process. Traditional models of therapy often fail to meet their needs.

Maltreated children present unique challenges, as they may be anxious, withdrawn, aggressive, inattentive, overly dependent, or sexually provocative. Given the likelihood of attachment problems and their previous experiences of being abused or manipulated, it often takes several sessions for maltreated children to trust or engage with a clinician. Managed care's limitations on the number of sessions can impede this therapeutic process. Once they do engage, the relationship often becomes very significant for the child, which can make termination difficult.

This work is also challenging because it often requires collaborating with professionals from very different systems (i.e., legal, social services, and mental health); each with different methods, goals, and perspectives. Some of these professionals may not value mental health services in general or may not believe that the psychological treatment of children is necessary. Many clinicians find the oftentimes necessary involvement in the legal system to be quite stressful.

Finally, treatment decisions are difficult to make given the paucity of literature evaluating treatment interventions with maltreated children. For physical abuse and neglect, most clinical work and empirical research focus largely on the maltreating parents, with relatively little attention paid to the child victims.

Children's involvement in treatment has usually occurred in the context of family-based services, making it difficult to determine the specific contribution of child treatment to outcomes. When victims are targeted, most studies include both physically abused and neglected children; few specify the numbers in each category (Oates & Bross, 1995) and even fewer differentiate between abused and neglected in terms of outcome.

Within the sexual abuse area, there is a great deal of clinical literature describing the treatment of child victims of father daughter incest, but less has been written about children who have been abused by nonfamily members or about male victims. Even in the sexual abuse field, the effectiveness of sexual abuse treatment has not yet been proven (Becker et al., 1995; Finkelhor & Berliner, 1995). Even though there are over 2000 programs in the USA for the treatment of child sexual abuse, few programs use objective measures for assessment and even fewer conduct follow-up evaluations (Keller, Cicchinelli, & Gardner, 1989). Of those studies that used quantitative measures, most were pre–post designs, which cannot tell us if the improvement was due to the treatment or simply the passage of time. Thus, there are few, if any, guidelines for matching victims to types of treatments. Given the limited empirical knowledge about the differential effectiveness of various forms of treatment, along with the heterogeneity of child maltreatment victims, it may be most parsimonious to focus on specific symptoms or problems exhibited by victims of child maltreatment. Heterogeneity of child maltreatment requires heterogeneity of treatment. The following section presents treatment suggestions drawn from research with abused and nonabused children which addresses many of their typical presenting problems.

In terms of process, clinicians should begin with a comprehensive assessment of the child and then, with input from the client(s), devise a therapeutic approach targeted at the identified problem areas. Treatment should be aimed at facilitating growth and competence in all developmental domains. Child-focused treatment modalities have included individual therapy, play therapy, drama therapy, art therapy, group therapy, cognitive-behavioral therapy, and behavior modification. While the child is in individual therapy, it is important to remember that she/he is being affected by home, school/daycare, and the broader community. As noted by Toth and Cicchetti (1993), "therapy with a maltreated child cannot be limited to the confines of the playroom, but must transcend the safety of the office to confront and impact upon the factors that may continue to assault the child in his or her daily life" (p. 415). Child maltreatment does not occur in a vacuum, nor should treatment. The effectiveness of treatment is bound to be strongly influenced by family and school contexts, and therapeutic interventions need to be directed at the system(s) in which the child is living.

9.21.3.1.1 Physical domain

Within the physical domain, specific target behaviors might include developmental delays, motor delays, self-help skills, or toileting problems. For example, Project 12-Ways, an ecobehavioral approach to the treatment and prevention of child abuse and neglect, has evaluated treatments for child neglect dealing with hygiene, nutrition, home safety, cleanliness, infant stimulation, and teaching health-related skills (Lutzker, 1990). Some of the training has been directed at the children, for example, by teaching them proper hygiene and toilet training.

The problem of developmental delays often seen in physically abused and neglected children has also been addressed by enrolling the children in a therapeutic preschool. Often, these programs focus on developing normal gross and fine motor, language, and social skills. For example, Parish, Myers, Brandner, and Templin (1985) demonstrated that a family development center, which combined services to parents (e.g., parenting skills classes) and preschool classes for children, produced substantial developmental skills gains, especially in fine motor and language skills.

Developmental gains in abused preschoolers were also found by Oates, Gray, Schweitzer, Kempe, and Harmon (1995). Further, developmental gains relative to a control group have been reported in other studies (Culp, Heide, & Richardson, 1987; Culp, Little, Letts, & Lawrence, 1991). Unfortunately, the multifaceted nature of treatment and absence of follow-up assessments limit the conclusions that can be made about therapeutic day-care programs.

Clinicians can also teach parents how to stimulate their infant's development. Walker, Bonner, and Kaufman (1988) suggested that clinicians assign parents stimulation tasks to carry out at home. These tasks should be assigned to remedy the specific deficits identified during the assessment process. Research is needed to determine the impact of such infant stimulation programs on children victimized by neglect.

Cases of extreme neglect, or failure to thrive (FTT), require medical diagnosis and management, nutrition intervention, and psychosocial intervention, often in a hospital setting. Drotar

(1995) also emphasized that clinical care of FTT should involve not only the management of current physical, psychological, and behavioral disturbances, but also the prevention of future problems.

9.21.3.1.2 Cognitive Domain

Given the findings documenting academic difficulties among maltreated children (Eckenrode et al., 1993), addressing this problem would appear warranted and should occur early (e.g., in elementary school; Kendall-Tackett & Eckenrode, 1996). Therapists can attend to academic problems in various ways. Clinicians can: (i) consult with teachers to help them understand the relationship between the child's behaviors and abuse-related experiences; (ii) conduct a thorough assessment of the child in order to formulate an intervention plan; (iii) design an educationally beneficial environment for the child; (iv) provide individual and family therapy; and (v) ensure that the child receives special programs or services to promote successful school adaptation and prevent school failure.

Depending on the extent of child maladaptation, educational intervention may best take place in a specialized school or preschool setting. Suggestions for making preschool classrooms function as therapeutic environments can be found elsewhere (Ayoub, 1991; Crittenden, 1989; Koplow, 1996). In this setting, mental health professionals can provide therapeutic services to victims, maltreating parents, and serve as educational consultants. They can also assist the child in making the transition to more conventional learning environments. Clearly, the role of the educational system in addressing the special needs of maltreated children must be stressed.

Findings that abuse-related attributions and perceptions are predictive of psychological symptoms in sexually abused girls (e.g., Mannarino & Cohen, 1996) suggest that these factors need to be addressed in therapy. It is believed that group therapy is particularly effective in helping victims deal with issues of blame, responsibility, and feeling different from peers. One advantage of the group format over individual therapy or family therapy is that victimized children have the opportunity to interact with peers who have experienced similar abuse. They can form positive images of other abused children, receive feedback and support from their peers, and have their distorted beliefs about themselves and the abuse challenged by group members. Several descriptions of group treatment models for sexually abused children and adolescents are available, but the reader should note that the relative efficacy of group therapy compared with other modalities has not been determined (e.g., Hack, Osachuk, & De Luca, 1994; McGain & McKinzey, 1995; Sinclair et al., 1995; Zaidi, & Gutierrez-Kovner, 1995).

During individual sessions, cognitive therapy can be employed to identify and correct any distortions or irrational beliefs a child has related to the abuse. Berliner and Wheeler (1987) emphasized the importance of identifying why children believe they were responsible for the abuse and then helping them understand the distinction between characterological self-blame, which can be maladaptive, and behavioral self-blame, which may be an adaptive coping response. They also encourage therapists to have the children recall the events or behaviors that preceded the actual abuse, with the goal being to help the child client recognize indications of abuse early to avert another incident. This strategy would appear to help reduce the child's feelings of powerlessness, while simultaneously empowering child victims to act responsibly and assertively in the future.

9.21.3.1.3 Socioemotional Domain

Although many theorists agree that attachment problems are central to all forms of child maltreatment, effective interventions to ameliorate attachment disorders are few in number and are mostly directed toward the periods of infancy and toddlerhood. For example, dysfunctional relationships during toddlerhood have been treated using infant-parent psychotherapy (Lieberman, 1992). Sessions are conducted with the parents and child in the office and at home. The therapist provides developmental guidance, helps the mother explore her perceptions, feelings, and reactions to the child, assists her in understanding how her own attachment and developmental experiences affect her relationship with her child, and stimulates changes in her internal representational models. Significant improvements in child, mother, and dyadic functioning have been found using this approach with high-risk families (Lieberman, Weston, & Pawl, 1991).

To treat peer relationship difficulties, behavioral approaches have been utilized to decrease aggressiveness and social withdrawal. For example, Fantuzzo et al. (1988) used peer-initiated social interaction to increase the positive social behavior of withdrawn, maltreated preschool children. More recently, Kolko (1996) noted the limited effectiveness of teaching interpersonal skills with physically abused children involved in individual therapy, and suggested the use of group or peer-based interventions to enhance social competence.

There is also an extensive literature on social skills interventions with nonabused children and adolescents that could be applied to abused children with social interaction deficits (Beelmann, Pfingsten, & Losel, 1994; La Greca, 1993). Aggressive and impulsive behavior has also been treated successfully using cognitive behavioral strategies (Lochman, Dunn, & Klimes-Dougan, 1993; Pepler & Rubin, 1991). Inclusion of teachers and parents in intervention planning is important because of their role in arranging opportunities for peer interactions, for monitoring peer behaviors, and providing appropriate consequences.

To treat anger and aggression, many clinicians encourage sexual and physical abuse victims to talk about their maltreatment and to express their angry feelings during play or verbal therapy. This strategy is believed to serve as a helpful form of catharsis and as an opportunity for gaining insight and to learn alternative social behaviors (e.g., Walker & Bolkovatz, 1988). Unfortunately, there is no empirical support for this practice, and there is the concern that making child maltreatment victims discuss their victimization may be potentially harmful (Weinbach & Curtiss, 1986). The impact of this practice must be empirically determined.

Given the frequent finding of lack of empathy among physically abused and neglected children, it would appear very important to provide empathy training in therapy. Unfortunately, problems with empathy, attachment, trust, and affect regulation are difficult to measure and persistently resistant to therapeutic intervention experiences (Carlson et al., 1989). Friedrich (1996) suggested that group treatment should be an effective therapeutic modality to learn empathy, but this assumption needs to be tested empirically.

Therapeutic interventions for anxiety will depend on how the child expresses the anxiety. For children who experience intrusive thoughts, anxiety, or fears/phobias, interventions such as systematic desensitization, guided imagery, relaxation training, controlled breathing, self-hypnosis, and thought-stopping have been recommended (Lipovsky, 1991; Walker et al., 1988). Children who exhibit PTSD symptoms have been found to benefit from participating in cognitive-behavioral treatment programs that include modeling, cognitive restructuring, coping skills training, gradual exposure to feared stimuli, and sexual abuse education and body safety skills training (Berliner & Saunders, 1996; Deblinger, Lippmann, & Steer, 1996). Others have found anxiety levels among sexually abused girls to decrease significantly as a function of group therapy (McGain & McKinzey, 1995) and abuse-focused individual therapy (Lanktree & Briere, 1995). Severely traumatized children may require adjunctive pharmacotherapy to reduce their levels of anxiety (Bernstein, 1994). Finally, an additional therapeutic focus should be to encourage support networks for the child, and to teach nonoffending parents/caretakers appropriate responses to the child's expressions of distress (Deblinger et al., 1996; Lipovsky, 1991).

For children suffering from depressive symptomatology, individual or group therapy can be conducted (e.g., using cognitive-behavioral therapy, interpersonal therapy, self-control therapy, or increasing response-contingent positive reinforcement via pleasant events). Antidepressant medication may be considered if depressed youth exhibit very significant symptomatology or have not responded to nonpharmacologic interventions.

9.21.3.1.4 Behavioral Domain

As discussed earlier, maltreated children as a group have more behavioral problems than nonmaltreated children. The types of problems range from excessive dependency, withdrawal, and clinginess, to aggression and delinquency. In the general child clinical area, behavioral approaches have been shown to be effective for symptoms such as aggressiveness, social withdrawal, and conduct problems. Within the sexual abuse area specifically, anger and aggressiveness have been difficult to reduce, suggesting that some therapeutic approaches (e.g., group therapy) may not be particularly effective for treating this type of problem (Finkelhor & Berliner, 1995).

Given that aversive behaviors displayed toward parents (e.g., aggression, impulsivity) may serve as "triggers" for abusive behavior, Walker et al. (1988) suggested that abused children be taught anger management skills to reduce their angry outbursts and increase their self-control. They also suggested that abused children should learn to obey commands promptly and earn rewards for adaptive behaviors. Also recommended is making sure that the child is given opportunities to actually practice these techniques in anger-arousing situations. Altering behaviors that contribute to the abusive cycle may lower the likelihood for subsequent abuse.

Aggressive and impulsive children may also benefit from training in problem-solving strategies. Kendall and Braswell (1985) discussed commonly used approaches with children and provide a useful intervention manual. There is also a large body of literature on cognitive and behavioral self-control training paradigms for teaching impulse control (e.g., Schweitzer & Sulzer-Azaroff, 1988) and anger management (e.g., Dangel, Deschner, & Rasp, 1989). For children who exhibit anger, impulsivity, or aggression in school, Cohen and Fish (1993) described interventions with demonstrated or potential value for school settings.

9.21.3.1.5 Psychosexual Domain

Children who exhibit sexually inappropriate behavior, especially if there is a compulsive, aggressive, or coercive quality to the behavior, are at risk for victimizing others, being revictimized themselves, and if they exhibit these behaviors at school or day care setting, then they may be in jeopardy of being suspended or expelled. Treatment of sexual behaviors in sexually abused children is thus highly recommended, but extremely challenging (Friedrich, 1990). For example, Finkelhor and Berliner (1995) found sexualized behavior problems, along with aggressiveness, to be particularly resistant to change. Given the serious consequences of this symptom, research clearly is needed to determine what works to eliminate inappropriate sexual behavior and simultaneously enhance normal sexual development.

Clinicians have suggested several therapeutic goals when working with children who have been traumatized sexually (Cunningham & MacFarlane, 1991; Gil & Johnson, 1993). One goal is to correct the child's thinking about sexual behavior, and this can be achieved by teaching age-appropriate sexual knowledge. Clinicians can help children understand the concepts of consent and mutuality, know the reasons why people are sexual, and define readiness to be sexually active. A second goal is to help increase children's internal inhibitions and external controls against unacceptable sexual behavior. Children can be helped to identify the behaviors, emotions, and thoughts that precede acting out sexually, and can be taught strategies to challenge cognitions and generate alternative behaviors to avoid repeating the behavior. A third goal is to attend to, and if necessary, modify the parents' attitudes and behaviors related to their children's sexual behavior. A fourth goal, if appropriate, would be to address the youth's own history of victimization. Given the relationship between child sexual abuse and adolescent pregnancy, an additional goal would be to discuss pregnancy prevention issues.

Clinicians treating abused children who engage in sexual activities with other children may benefit from the work being done at treatment programs for child perpetrators (see review by Pithers et al., 1995). For example, in the Support Program for Abuse Reactive Kids (SPARK; Johnson & Berry, 1989), group therapy focuses on helping children learn about their behavior patterns as well as to be aware of their feelings and actions in relation to others. They are also taught problem-solving and social skills, they learn impulse control and perspective taking, develop victim empathy, and increase frustration tolerance. Unfortunately, a paucity of empirical data exists to support the efficacy of intervening with young sexual offenders (see Chapter 28, this volume).

9.21.4 PREVENTION

In addition to providing direct intervention with child maltreatment victims, prevention efforts aimed at reducing the incidence of child maltreatment are critical to ultimately eliminating this problem. Along with targeting communities and high-risk families, prevention programs should also target children. For example, in the sexual abuse field, several reviews have concluded that behaviorally oriented educational programs can help children gain in knowledge about sexual abuse and in personal safety skills thought to be effective in preventing abuse (Tutty, 1993; Wurtele & Miller-Perrin, 1992).

Child-focused educational programs intended to prevent physical abuse and neglect are fewer in number and even fewer have been evaluated. One reason for the dearth of physical abuse prevention programs is the concern that telling children that corporal punishment is wrong may conflict with parents' views on this issue. There have, however, been efforts directed at teaching children home safety skills, which could be relevant to the abused and neglected population (Mori & Peterson, 1986; Peterson, 1984a, 1984b). Clinicians can play a role in promoting and recommending educational programs. Continued efforts to prevent abuse and neglect are clearly justified, as the key to the solution of the child maltreatment problem is prevention.

In addition to targeting presenting symptoms, clinicians should also focus their efforts on the prevention of reabuse among their child maltreatment victims. Abused children, perhaps even more than nonabused children, need to be educated about prevention concepts and skills. There are numerous books, games, videos, and pamphlets designed to educate children and their parents about personal safety. Such education can take place either during individual or group work with the children. Topics to be covered might include a discussion of the "triggers" for abuse or neglect, safe strategies to resist abuse, and helpful people/agencies to contact in a crisis.

9.21.5 FUTURE DIRECTIONS

Despite numerous methodological advances related to measurement, sample selection, and research design, a major methodological issue to be addressed in future child maltreatment research is the operationalization of the independent variable; the maltreatment itself. Although there is consensus that maltreatment is a heterogeneous phenomenon, much of the research continues to view maltreatment as a dichotomous variable. A categorization methodology obscures differences in the severity of abuse and ignores the fact that different types of maltreatment often co-occur.

There have been limited attempts at measuring the diverse range of maltreatment experiences present in the histories of victimized children. For example, Barnett, Manly, & Cicchetti (1993) developed a maltreatment classification system to quantify such dimensions as subtype, severity, frequency, chronicity, perpetrator characteristics, the developmental period(s) during which the maltreatment occurred, and out-of-home placement history. In addition to more completely assessing the child's experiences, assessment strategies need to be broadened to include the whole child. Across studies, researchers have considered the impact of maltreatment on several developmental domains, including physical, socioemotional, cognitive, behavioral, and psychosexual. Yet within studies, few researchers have explored the effects on multiple domains; most limit their exploration to the behavioral domain.

The effects of maltreatment on children would be better understood through examining additional domains of functioning. Within the cognitive domain, more research on the coping and appraisal processes of maltreated children is needed. Further examination of maltreated children's general attributional styles, along with their abuse-specific attributions and coping strategies, should have significant research and practice implications. Future studies could focus on whether attributional factors affect symptomatology, and on whether cognitive therapy approaches can be utilized to alter victims' attributions and enhance their recovery. Within the socioemotional domain, research efforts must continue to develop methods to assess the quality of children's attachments to both parents and to a variety of relationship figures. Knowing the extent and quality of support in the child's environment will help elucidate the impact of abuse and also suggest intervention approaches. We also need to learn a great deal more about how maltreatment affects the immunological, neuroendocrinological, and psychophysiological systems of maltreated children. This research will undoubtedly demonstrate the profound and pervasive impact of child maltreatment.

The choice of measures needs to be determined by hypotheses about maltreatment effects and with consideration of the expenses involved (including burden on families). The recommendation for an expanded assessment strategy is contingent upon the development of reliable, valid, and clinically useful instruments, especially instruments that are sensitive enough to detect abuse-specific effects, along with assessing changes over time. Better measurement will allow researchers to unravel the processes contributing to maladaptive outcomes, and to explicate the processes that serve to moderate the adverse consequences of maltreatment. Thus, future directions in clinical practice will depend on the advances made in the measurement area.

Although research directed toward assessing the harmful consequences of child maltreatment has burgeoned, treatment outcome studies for all three types of maltreatment are much fewer in number. In order to convince others of the importance of providing intervention for maltreated children, data on the effectiveness of such services must be collected. Granted, there are a host of difficulties confronting clinicians and researchers in this area. Despite the difficulties inherent in implementing and evaluating therapy for maltreatment victims, it is crucial that children be provided with effective services. And it is important that interventions be methodologically rigorous, with all variables operationalized, including the type(s) of maltreatment experienced and the exact nature of treatment. Basic recommendations for future research include identifying the type(s) of maltreatment experienced, ensuring adequate sample sizes, performing routine assessments both before and after treatment, randomly assigning victims to treatment (when possible), specifying the type of treatment employed and consider using manuals to ensure treatment integrity, and conducting follow-up assessments to determine the stability of therapeutic change.

Clinicians and researchers need to collaborate in identifying effective treatment modalities for victims of child neglect, given its widespread occurrence and demonstrated link to serious negative consequences. With school-aged children, effective interventions to increase their chances of academic success are sorely needed, especially for physically abused and neglected children. In general, research is desperately needed to determine the applicability of intervention strategies with older children and adolescent victims of child maltreatment. Another treatment question that needs to be addressed is whether a symptom-focused treatment approach is more effective than more generic therapies. Clearly, much empirical work needs to be conducted to be able to answer the important question, "What type of therapy works best with which types of children who exhibit what types of problems and who have experienced which types of maltreatment?" Focusing on this complicated question holds exciting possibilities for collaboration between researchers and clinicians.

9.21.6 SUMMARY

It is clear that being neglected or abused negatively affects children's development. But maltreated children are a heterogeneous group; they exhibit a wide variety of physical, cognitive, behavioral, and emotional responses. There is no "syndrome" or specific set of symptoms that characterize all victims of maltreatment. Nor is there a standardized instrument available that adequately evaluates the diverse symptomatology associated with child maltreatment. This characterization highlights the need for clinicians to: (i) use multiple instruments; (ii) obtain data from multiple informants in multiple settings; (iii) use multiple times of assessment (initial assessment, pretreatment, posttreatment, follow-up); (iv) assess across the principle domains of development; (v) measure both global adjustment and problems specific to the abuse; and (vi) consider the multiple influences on the child. Although a number of instruments and approaches are currently available to evaluate the diverse symptomatology associated with child maltreatment, there is clearly a need to develop additional instruments to measure abuse-specific effects.

The documentation of a diverse set of psychological problems in maltreated children, which can persist into adulthood, dramatizes the need to develop effective therapeutic interventions with child victims. Although many clinicians provide therapeutic services to maltreated children and their families, systematic evaluation of these efforts has been limited. Thus, our knowledge about treatment has lagged behind that of our knowledge of the effects of abuse and neglect on children. Further development of treatment outcome research will require empirically based definitions of child maltreatment, better descriptions of the treatment, more abuse-specific assessment methods, and methodologically sound studies. To this aim, collaboration between psychotherapy researchers, practitioners, and researchers in developmental psychopathology will be needed to realize the true potential of psychosocial treatment for child maltreatment victims.

Though fraught with difficulties, the potential value of therapy to child maltreatment victims warrants its conduct. Services for victims of child maltreatment have untested potential for treating the immediate problems, preparing the child for successful resolution of future developmental challenges, and for interrupting the intergenerational transmission of child maltreatment. Child maltreatment impacts the victims, and ultimately us all. Providing effective therapeutic services to maltreatment victims will benefit the children, as well as society.

9.21.7 REFERENCES

Achenbach, T. M., & Edelbrock, C. S. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, VT: Queen City Printers.

Alessandri, S. M. (1991). Play and social behavior in maltreated preschoolers. *Development and Psychopathology, 3,* 191–205.

Alexander, P. C. (1992). Application of attachment theory to the study of sexual abuse. *Journal of Consulting and Clinical Psychology, 60,* 185–195.

Ayoub, C. (1991). Physical violence and preschoolers: The use of therapeutic day care in the treatment of physically abused children and children from violent families. *The Advisor, 4,* 1, 15–16, 18.

Barnett, D., Manly, J. T., & Cicchetti, D. (1993). Defining child maltreatment: The interface between policy and research. In D. Cicchetti & S. L. Toth (Eds.), *Child abuse, child development, and social policy* (pp. 7–73). Norwood, NJ: Ablex.

Barrera, M. (1980). A method for the assessment of social support networks in community survey research. *Connections, 3,* 8–13.

Beck, A. (1976). *Cognitive therapy and the emotional disorders*. New York: New American Library.

Becker, J. V., Alpert, J. L., BigFoot, D. S., Bonner, B. L., Geddie, L. F., Henggeler, S. W., Kaufman, K. L., & Walker, C. E. (1995). Empirical research on child abuse treatment: Report by the Child Abuse and Neglect Treatment Working Group, American Psychological Association. *Journal of Clinical Child Psychology, 24* (Suppl.), 23–46.

Beelmann, A., Pfingsten, U., & Losel, F. (1994). Effects of training social competence in children: A meta-analysis of recent evaluation studies. *Journal of Clinical Child Psychology, 23,* 260–271.

Behar, L., & Stringfield, S. (1974). A behavior rating scale for the preschool child. *Developmental Psychology, 10,* 60–10.

Berliner, L., & Saunders, B. E. (1996). Treating fear and anxiety in sexually abused children: Results of a controlled 2-year follow-up study. *Child Maltreatment, 1,* 294–309.

Berliner, L., & Wheeler, J. R. (1987). Treating the effects of sexual abuse on children. *Journal of Interpersonal Violence, 2,* 415–434.

Bernstein, G. A. (1994). Psychopharmacological interventions. In T. H. Ollendick, N. J. King, & W. Yule (Eds.), *International handbook of phobic and anxiety disorders in* *children and adolescents* (pp. 439–452). New York: Plenum.

Bierman, K., & McCauley, E. (1987). Children's descriptions of their peers" interactions: Useful information for clinical child assessment. *Journal of Clinical Psychology, 16*, 9–18.

Boyer, D., & Fine, D. (1992). Sexual abuse as a factor in adolescent pregnancy and child maltreatment. *Family Planning Perspectives, 24*, 4–11.

Briere, J. (1996). *The Trauma Symptom Checklist for Children*. Odessa, FL: Psychological Assessment Resources.

Burgess, E. J. (1988). Sexually abused children and their drawings. *Archives of Psychiatric Nursing, 2*, 65–73.

Carlson, V., Cicchetti, D., Barnett, D., & Braunwald, K. (1989). Disorganized/disoriented attachment relationships in maltreated infants. *Developmental Psychology, 25*, 525–531.

Carrow-Woolfolk, E. (1985). *Test for Auditory Comprehension of Language* (Rev. ed.). Allen, TX: DLM Teaching Resources.

Celano, M. P. (1992). A developmental model of victims' internal attributions of responsibility for sexual abuse. *Journal of Interpersonal Violence, 7*, 57–69.

Chambers, W. J., Puig-Antich, J., Hirsch, M., Paez, P., Abrosini, P. J., Tabrizi, M. A., & Davies, M. (1985). The assessment of affective disorders in children and adolescents by semistructured interview. *Archives of General Psychiatry, 42*, 696–702.

Cicchetti, D. (1987). Developmental psychopathology in infancy: Illustration from the study of maltreated youngsters. *Journal of Consulting and Clinical Psychology, 55*, 837–845.

Cicchetti, D., & Lynch, M. (1993). Toward an ecological/transactional model of community violence and child maltreatment: Consequences for children's development. *Psychiatry, 56*, 996–1118.

Cohen, J. J., & Fish, M. C. (1993). *Handbook of school-based interventions*. San Francisco: Jossey-Bass.

Compas, B. E. (1987). Coping with stress during childhood and adolescence. *Psychological Bulletin, 101*, 393–403.

Conaway, L. P., & Hansen, D. J. (1989). Social behavior of physically abused and neglected children: A critical review. *Clinical Psychology Review, 9*, 627–652.

Coopersmith, S. (1986). *Manual for the Self-Esteem Inventories*, Palo Alto, CA: Consulting Psychologists Press.

Coster, W. J., Gersten, M. S., Beeghly, M., & Cicchetti, D. (1989). Communicative functioning in maltreated toddlers. *Developmental Psychology, 25*, 1020–1029.

Crittenden, P. M. (1989). Teaching maltreated children in the preschool. *Topics in Early Childhood Special Education, 9*, 16–32.

Crittenden, P. M. (1992). Children's strategies for coping with adverse home environments: An interpretation using attachment theory. *Child Abuse and Neglect, 16*, 329–343.

Crouch, J. L., & Milner, J. S. (1993). Effects of child neglect on children. *Criminal Justice and Behavior, 20*, 49–65.

Culp, R. E., Heide, J., & Richardson, M. T. (1987). Maltreated children's developmental scores: Treatment versus nontreatment. *Child Abuse and Neglect, 11*, 29–34.

Culp, R. E., Little, V., Letts, D., & Lawrence, H. (1991). Maltreated children's self-concept: Effects of a comprehensive treatment program. *American Journal of Orthopsychiatry, 61*, 114–121.

Cunningham, C., & MacFarlane, K. (1991). *When children molest children*. Orwell, VT: Safer Society Press.

Dangel, R. F., Deschner, J. P., & Rasp, R. R. (1989). Anger control training for adolescents in residential treatment. *Behavior Modification, 13*, 447–458.

Daro, D. (1995). Current trends in child abuse reporting and fatalities. NCPCA's 1994 Annual Fifty State Survey. *The APSAC Advisor, 8*, 5–6.

De Bellis, M. D., Chrousos, G. P., Dorn, L. D., Burke, L., Helmers, K., Kling, M. A., Trickett, P. K., & Putnam, F. W. (1994). Hypothalamic–pituitary–adrenal axis dysregulation in sexually abused girls. *Journal of Clinical Endocrinology and Metabolism, 78*, 249–255.

Deblinger, E., Lippmann, J., & Steer, R. (1996). Sexually abused children suffering posttraumatic stress symptoms: Initial treatment outcome findings. *Child Maltreatment, 1*, 310–321.

Dodge, K. A., Pettit, G. S., & Bates, J. E. (1994). Effects of physical maltreatment on the development of peer relations. *Development and Psychopathology, 6*, 43–55.

Drotar, D. (1995). Failure to thrive (growth deficiency). In M. C. Roberts (Ed.), *Handbook of pediatric psychology* (2nd ed., pp. 516–536). New York: Guilford.

Dunn, L. M., & Dunn, L. (1981). *Manual for the Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN: American Guidance Service.

Eckenrode, J., Laird, M., & Doris, J. (1993). School performance and disciplinary problems among abused and neglected children. *Developmental Psychology, 29*, 53–62.

Everson, M. D., & Boat, B. W. (1990). Sexualized doll play among young children: Implications for the use of anatomical dolls in sexual abuse evaluations. *Journal of the American Academy of Child and Adolescent Psychiatry, 29*, 736–742.

Eyberg, S., & Ross, A. (1978). Assessment of child behavior problems: The validation of a new inventory. *Journal of Clinical Child Psychology, 7*, 113–116.

Fantuzzo, J. W., Jurecic, L., Stovall, A., Hightower, A. D., Goins, C., & Schachtel, D. (1988). Effects of adult and peer social initiations on the social behavior of withdrawn, maltreated preschool children. *Journal of Consulting and Clinical Psychology, 56*, 34–39.

Finkelhor, D., & Berliner, L. (1995). Research on the treatment of sexually abused children: A review and recommendations. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*, 1408–1423.

Finkelhor, D., & Browne, A. (1985). The traumatic impact of child sexual abuse: A conceptualization. *American Journal of Orthopsychiatry, 55*, 530–541.

Fitts, W. R. (1965). *Manual for the Tennessee Self-Concept Scale*. Nashville, TN: Counselor Records and Tests.

Folkman, S., & Lazarus, R. S. (1980). An analysis of coping in a middle-aged community sample. *Journal of Health and Social Behavior, 21*, 219–239.

Friedrich, W. N. (1990). *Psychotherapy of sexually abused children and their families*. New York: Norton.

Friedrich, W. N. (1996). An integrated model of psychotherapy for abused children. In J. Briere, L. Berliner, J. A. Bulkley, C. Jenny, & T. Reid (Eds.), *The APSAC handbook on child maltreatment* (pp. 104–118). Thousand Oaks, CA: Sage.

Friedrich, W. N., Grambsch, P., Damon, L., Hewitt, S. K., Koverola, C., Lang, R. A., Wolfe, V., & Broughton, D. (1992). Child Sexual Behavior Inventory: Normative and clinical comparisons. *Psychological Asssessment, 4*, 303–311.

Friedrich, W. N., & Schafer, L. C. (1995). Somatic symptoms in sexually abused children. *Journal of Pediatric Psychology, 20*, 661–670.

Gardner, M. (1979). *Expressive One-Word Picture Vocabulary Test*. Novato, CA: Academic Therapy Publications.

Gil, E., & Johnson, T. C. (1993). *Sexualized children: Assessment and treatment of sexualized children and children who molest*. Rockville, MD: Launch.

Hack, T. F., Osachuk, T. A. G., & De Luca, R. V. (1994). Group treatment for sexually abused preadolescent boys. *Families in Society, 75*, 217–228.

Hart, J., Gunnar, M., & Cicchetti, D. (1995). Salivary cortisol in maltreated children: Evidence of relations between neuroendocrine activity and social competence. *Development and Psychopathology, 7,* 11–26.

Harter, S. (1985). *Manual for the Self-Perception Profile for Children*. Denver, CO: University of Denver.

Harter, S. (1988). *Manual for the Self-Perception Profile for Adolescents*, Denver, CO: University of Denver.

Harter, S., & Pike, R. (1984). The pictorial scale of perceived competence and social acceptance for young children. *Child Development, 55,* 1969–1982.

Hibbard, R. A., Roghmann, K., & Hoekelman, R. A. (1987). Genitalia in children's drawings: An association with sexual abuse. *Pediatrics, 79,* 129–137.

Hightower, A. D., Spinell, A., & Lotyczewski, B. S. (1985). *Primary mental health project teacher-child rating scale (T-CRS) guidelines*. Rochester, NY: Unpublished manuscript.

Hill, S. D., Bleichfield, B., Brunstetter, R. D., Hebert, J. E., & Steckler, S. (1989). Cognitive and physiological responsiveness of abused children. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 219–224.

Hodges, K., Kline, J., McKnew, D. H., & Cytryn, L. (1992). The development of a child assessment interview for research and clinical use. *Journal of Abnormal Child Psychology, 10,* 173–180.

Hotte, J. P., & Rafman, S. (1992). The specific effects of incest on prepubertal girls from dysfunctional families. *Child Abuse and Neglect, 16,* 273–283.

Howes, C., & Eldredge, R. (1985). Responses of abused, neglected and non-maltreated children to the behaviors of their peers. *Journal of Applied Developmental Psychology, 6,* 261–270.

Jampole, L., & Weber, M. K. (1987). An assessment of the behavior of sexually abused and nonsexually abused children with anatomically correct dolls. *Child Abuse and Neglect, 11,* 187–192.

Johnson, B. K., & Kenkel, M. B. (1991). Stress, coping, and adjustment in female adolescent incest victims. *Child Abuse and Neglect, 15,* 293–305.

Johnson, T. C., & Berry, C. (1989). Children who molest: A treatment program. *Journal of Interpersonal Violence, 4,* 185–203.

Kaslow, N. J., Tannenbaum, R. L., & Seligman, M. E. P. (1978). *The KASTAN-R: A Children's Attributional Style Questionnaire (KASRAN-R-CASO)*. Unpublished manuscript. University of Pennsylvania.

Kaufman, J. (1991). Depressive disorders in maltreated children. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 257–265.

Kaufman, J., & Cicchetti, D. (1989). The effects of maltreatment on school-aged children's socioemotional development: Assessments in a day camp setting. *Develomental Psychology, 25,* 516–524.

Kazdin, A. E., French, M. H., Unis, A. S., Esveldt-Dawson, K., & Sherrick, R. B. (1983). Hopelessness, depression, and suicidal intent among psychiatrically disturbed inpatient children. *Journal of Consulting and Clinical Psychology, 51,* 504–510.

Keller, R. A., Cicchinelli, L. F., & Gardner, D. M. (1989). Characteristics of child sexual abuse treatment programs. *Child Abuse and Neglect, 13,* 361–368.

Kendall, P. C., & Braswell, L. (1985). *Cognitive behavioral therapy for impulsive children*. New York: Guilford.

Kendall-Tackett, K. A., & Eckenrode, J. (1996). The effects of neglect on academic achievement and disciplinary problems: A developmental perspective. *Child Abuse and Neglect, 20,* 161–169.

Kendall-Tackett, K. A., Williams, L. M., & Finkelhor, D. (1993). Impact of sexual abuse on children: A review and synthesis of recent empirical studies. *Psychological Bulletin, 113,* 164–180.

Kolko, D. J. (1996). Individual cognitive behavioral treatment and family therapy for physically abused children and their offending parents: A comparison of clinical outcomes. *Maltreatment, 1,* 322–342.

Koplow, L. (1996). *Unsmiling faces: How preschools can heal*. New York: Teachers College Press.

Kovacs, M. (1981). Rating scales to assess depression in school-aged children. *Acta Paedopsychiatrica, 46,* 305–315.

Kovacs, M. (1983). *Interview Schedule for Children (ISC): Form C*. Pittsburgh, PA: Western Psychiatric Institute and Clinic.

Kovacs, M. (1992). *Children's Depression Inventory manual.* North Tonawanda, NY: Multi-Health Systems.

LaGreca, A. M. (1993). Social skills training with children: Where do we go from here. *Journal of Clinical Child Psychology, 22,* 288–298.

Lanktree, C. B., & Briere, J. (1995). Outcome of therapy for sexually abused children: A repeated measures study. *Child Abuse and Neglect, 19,* 1145–1155.

Leiter, J., & Johnsen, M. C. (1994). Child maltreatment and school performance. *American Journal of Education, 102,* 154–189.

Levendosky, A. A., Okun, A., & Parker, J. G. (1995). Depression and maltreatment as predictors of social competence and social problem-solving skills in school-age children. *Abuse and Neglect, 19,* 1183–1195.

Lieberman, A. F. (1992). Infant–parent psychotherapy with toddlers. *Development and Psychopathology, 4,* 559–574.

Lieberman, A. F., Weston, D., & Pawl, J. H. (1991). Preventive intervention and outcome with anxiously attached dyads. *Child Development, 62,* 199–209.

Lipovsky, J. A. (1991). Posttraumatic stress disorder in children. *Community Health, 14,* 42–51.

Lochman, J. E., Dunn, S. E., & Klimes-Dougan, B. (1993). An intervention and consultation model from a social cognitive perspective: A description of the anger coping program. *School Psychology Review, 22,* 458–471.

Lutzker, J. R. (1990). Behavioral treatment of child neglect. *Behavior Modification, 14,* 301–315.

Main, M., & George, C. (1985). Responses of abused and disadvantaged toddlers to distress in age-mates: A study in the day-care setting. *Developmental Psychology, 21,* 407–412.

Mannarino, A. P., & Cohen, J. A. (1996). Abuse-related attributions and perceptions, general attributions, and locus of control in sexually abused girls. *Journal of Interpersonal Violence, 11,* 162–180.

Mannarino, A. P., Cohen, J. A., & Berman, S. R. (1994). The Children's Attributions and Perceptions Scale: A new measure of sexual abuse-related factors. *Journal of Clinical Child Psychology, 23,* 204–211.

Mannarino, A. P., Cohen, J. A., & Gregor, M. (1989). Emotional and behavioral difficulties in sexually abused girls. *Journal of Interpersonal Violence, 4,* 437–451.

Mannarino, A. P., Cohen, J. A., Smith, J. A., & Moore-Motily, S. (1991). Six- and twelve-month follow-up of sexually abused girls. *Journal of Interpersonal Violence, 6,* 494–511.

Matson, J. L., Esveldt-Dawson, K., & Kazdin, A. (1983). Evaluating social skills with youngsters. *Journal of Clinical Child Psychology, 12,* 174–180.

McGain, B., & McKinzey, R. K. (1995). The efficacy of group treatment in sexually abused girls. *Child Abuse & Neglect, 19,* 1157–1169.

McLeer, S. V., Deblinger, E., Atkins, M. S., Foa, E. B., & Ralphe, D. L. (1988). Post-traumatic stress disorder in sexually abused children. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 650–654.

Miller, L. C. (1979). *Louisville Behavior Checklist*. Los Angeles: Western Psychological Services.

Miller-Perrin, C. L., & Wurtele, S. K. (1990). Reactions to childhood sexual abuse: Implications for post-traumatic stress disorder. In C. Meek (Ed.), *Post-traumatic stress disorder: Assessment, differential diagnosis, and forensic evaluation* (pp. 91–135). Sarasota, FL: Professional Resource Exchange.

Mischel, W., Zeiss, R., & Zeiss, A. (1974). Internal–external locus of control and persistence: Validation and implications of the Stanford Preschool Internal–External Scale. *Journal of Personality and Social Psychology, 29*, 265–278.

Mori, L., & Peterson, L. (1986). Training preschoolers in home safety skills to prevent inadvertent injury. *Journal of Clinical Child Psychology, 15*, 106–114.

National Center on Child Abuse and Neglect (1978). *Child sexual abuse: Incest, assault, and sexual exploitation, a special report*. Washington, DC: Author.

Nowicki, S., & Strickland, B. (1973), A locus of control scale for children. *Journal of Consulting and Clinical Psychology, 40*, 148–154.

Oates, R. K., & Bross, D. C. (1995). What have we learned about treating child physical abuse? A literature review of the last decade. *Child Abuse and Neglect, 19*, 463–473.

Oates, R. K., Gray, J., Schweitzer, L., Kempe, R. S., & Harmon, R. J. (1995). A therapeutic preschool for abused children: The KEEPSAFE project. *Child Abuse and Neglect, 19*, 1379–1386.

Ollendick, T. H. (1983). Reliability and validity of the revised Fear Survey Schedule for Children (FSSC-R). *Behaviour Research and Therapy, 21*, 685–692.

Parish, R. A., Myers, P. A., Brandner, A., & Templin, K. H. (1985). Developmental milestones in abused children, and their improvement with a family-oriented approach to the treatment of child abuse. *Child Abuse and Neglect, 9*, 245–250.

Pepler, D. J., & Rubin, K. H. (1991). *The development and treatment of childhood aggression*, Hillsdale, NJ: Erlbaum.

Putman, F. W., & Trickett, P. K. (1993). Child sexual abuse: A model of chronic trauma. *Psychiatry, 56*, 82–95.

Perry, B. D., Pollard, R. A., Blakley, T. L., Baker, W. L., & Vigilante, D. (1995). Childhood trauma, the neurobiology of adaptation, and "use-dependent" development of the brain: How "states" become "traits." *Infant Mental Health Journal, 16*, 271–291.

Putnam, F. W., Helmers, K., & Trickett, P. K. (1993). Development, reliability, and validity of a child dissociation scale. *Child Abuse and Neglect, 17*, 731–741.

Pettit, G. S., Dodge, K. A., & Brown, M. M. (1988). Early family experience, social problem solving patterns and children's social competence. *Child Development, 59*, 107–120.

Pynoos, R. S., & Eth, S. (1986). Witness to violence: The child interview. *Journal of the American Academy of Child and Adolescent Psychiatry, 25*, 306–319.

Rainey, D. Y., Stevens-Simon, C., & Kaplan, D. W. (1995). Are adolescents who report prior sexual abuse at higher risk for pregnancy? *Child Abuse and Neglect, 19*, 1283–1288.

Pithers, W. D., Becker, J. V., Kafka, M., Morenz, B., Schlank, A. & Leombruno, T. (1995). Children with sexual behavior problems, adolescent sexual abusers, and adult sex offenders: Assessment and treatment. In J. M. Oldham & M. B. Riba (Eds.), *Review of Psychiatry* (Vol. 14, pp. 779–818). Washington, DC: American Psychiatric Press.

Peterson, L. (1984a). The "Safe at Home" game: Training comprehensive prevention skills in latchkey children. *Behavior Modification, 8*, 474–494.

Peterson, L. (1984b). Teaching home safety and survival skills to latch-key children: A comparison of two manuals and methods. *Journal of Applied Behavior Analysis, 17*, 279–294.

Piers, E. V., & Harris, D. B. (1984). *Piers–Harris Children's Self-Concept Scale (Revised manual)*. Los Angeles: Western Psychological Services.

Quay, H., & Peterson, D. (1987). *Manual for the Revised Behavior Checklist*. Coral Gables, FL: University of Miami, Department of Psychology.

Reynolds, W. M. (1987). *Reynolds Adolescent Depression Scale: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Reynolds, C. R., & Richmond, B. O. (1985). *Revised Children's Manifest Anxiety Scale* (RCMAS) Manual. Los Angeles: Western Psychological Services.

Rogosch, F. A., & Cicchetti, D. (1994). Illustrating the interface of family and peer relations through the study of child maltreatment. *Social Development, 3*, 291–308.

Rosenberg, M. (1979). *Conceiving the self*. New York: Basic Books.

Saigh, P. A. (1989). The development and validation of the children's posttraumtic stress disorder inventory. *International Journal of Special Education, 4*, 75–84.

Salzinger, S., Feldman, R. S., Hammer, M., & Rosario, M. (1993). The effects of physical abuse on children's social relationships. *Child Development, 64*, 169–187.

Schweitzer, J. B., & Sulzer-Azaroff, B. (1988). Self-control in teaching tolerance for delay in impulsive children. *Journal of the Experimental Analysis of Behavior, 50*, 173–186.

Shaffer, D., Schwab-Stone, M., Fisher, P., Cohen, P., Piacentini, J., Davies, M., Conners, C. K., & Regier, D. (1993). The Diagnostic Interview for Children-Revised version (DISC-R). I. Preparation, field testing, interrater reliability, and acceptability. *Journal of the American Academy of Child and Adolescent Psychiatry, 32*, 643–650.

Shapiro, J. P., Leifer, M., Martone, M. W., & Kassem, L. (1990). Multimethod assessment of depression in sexually abused girls. *Journal of Personality Assessment, 55*, 234–248.

Sinclair, J. J., Larzelere, R. E., Paine, M., Jones, P., Graham, K., & Jones, M. (1995). Outcome of group treatment for sexually abused adolescent females living in a group home setting: Preliminary findings. *Journal of Interpersonal Violence, 10*, 533–542.

Spaccarelli, S. (1994). Stress, appraisal, and coping in child sexual abuse: A theoretical and empirical review. *Psychological Bulletin, 116*, 340–362.

Spielberger, C. D. (1973). *Manual for the State–Trait Anxiety Inventory for Children*. Palo Alto, CA: Consulting Psychologist Press.

Toth, S. L., & Cicchetti, D. (1993). Child maltreatment: Where do we go from here in our treatment of victims? In D. Cicchetti & S. L. Toth (Eds.), *Child abuse, child development, and social policy* (pp. 399–437). Norwood, NJ: Ablex Publishing.

Toth, S. L., Manly, J. T., & Cicchetti, D. (1992). Child maltreatment and vulnerability to depression. *Development and Psychopathology, 4*, 97–112.

Tutty, L. M. (1993). Are child sexual abuse prevention programs effective? A review of the research. *Revue Sexologique, 1*, 93–114.

US Advisory Board on Child Abuse and Neglect (1990). *Child abuse and neglect: Critical first steps in response to a national emergency*. Washington, DC: US Government Printing Office.

US Advisory Board on Child Abuse and Neglect (1995). *A nation's shame: Fatal child abuse and neglect in the United States*. Washington, DC: US Government Printing Office.

US Department of Health and Human Services (1996). *Child maltreatment 1994: Reports from the states to the National Center on Child Abuse and Neglect*. Washington, DC: US Government Printing Office.

Vondra, J. A., Barnett, D., & Cicchetti, D. (1990). Self-concept, motivation, and competence among preschoolers from maltreating and comparison families. *Child Abuse and Neglect, 14*, 525–540.

Walker, C. E., Bonner, B. L., & Kaufman, K. L. (1988). *The physically and sexually abused child: Evaluation and treatment*. New York: Pergamon.

Walker, L. E., & Bolkovatz, M. A. (1988). Play therapy with children who have experienced sexual assault. In L. E. Walker (Ed.), *Handbook on sexual abuse of children* (pp. 249–269). New York: Springer.

Weinbach, R. W., & Curtiss, C. R. (1986). Making child abuse victims aware of their victimization: A treatment issue. *Child Welfare, 65,* 337–346.

Weiss, B., Dodge, K. A., Bates, J. E., & Pettit, G. S. (1992). Some consequences of early harsh discipline: Child aggression and a maladaptive social information processing style. *Child Development, 63,* 1321–1335.

Welner, Z., Reich, W., Herjanic, B., Jung, K. G., & Amado, H. (1987). Reliability, validity, and parent–child agreement studies of the Diagnostic Interview for Children and Adolescents (DICA). *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 649–653.

Wertlieb, D., Weigel, C., & Feldstein, M. (1987). Measuring children's coping. *American Journal of Orthopsychiatry, 57,* 548–560.

White, S., Halpin, B. M., Strom, G. A., & Santilli, G. (1988). Behavioral comparisons of young sexually abused, neglected and nonreferred children. *Journal of Clinical Child Psychology*, *17,* 53–61.

Widom, C. S. (1989a). The cycle of violence. *Science, 244,* 160–165.

Widom, C. S. (1989b). Does violence beget violence? A critical examination of the literature. *Psychological Bulletin, 106,* 3–28.

Wolfe, V., Gentile, C., Michienzi, T., Sas, L., & Wolfe, D. (1991). The Children's Impact of Traumatic Events Scale-Revised: A measure of post sexual abuse PTSD symptoms. *Behavioral Assessment, 14,* 359–383.

Wolfe, V. V., Gentile, C., & Wolfe, D. A. (1989). The impact of sexual abuse on children: A PTSD formulation. *Behavior Therapy, 20,* 215–228.

Wozencraft, T., Wagner, W., & Pellegrin, A. (1991). Depression and suicidal ideation in sexually abused children. *Child Abuse and Neglect, 15,* 505–511.

Wurtele, S. K., & Miller-Perrin, C. L. (1992). *Preventing child sexual abuse: Sharing the responsibility*. Lincoln, NE: University of Nebraska Press.

Zaidi, L. Y., & Gutierrez-Kovner, V. M. (1995). Group treatment of sexually abused latency-age girls. *Journal of Interpersonal Violence, 10,* 215–227.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.22
Victims of Hate Crimes

CHERYL S. AL-MATEEN, DAWN K. LEWIS, and NIRBHAY N. SINGH
Virginia Commonwealth University, Richmond, VA, USA

9.22.1 INTRODUCTION 359
9.22.2 HATE CRIMES 360
  9.22.2.1 Definition 360
  9.22.2.2 Categorization and Typology 361
  9.22.2.3 Reporting of Hate Crimes 361
9.22.3 PREVALENCE 362
9.22.4 REACTIONS OF VICTIMS 364
9.22.5 CULTURAL FACTORS AND HATE CRIMES 365
  9.22.5.1 Gays, Lesbians, and Hate Crimes 366
  9.22.5.2 College Students and Hate Crimes 367
9.22.6 INTERVENTIONS 369
  9.22.6.1 Mental Health Treatment 369
  9.22.6.2 Advocacy 369
9.22.7 PREVENTION 371
9.22.8 FUTURE DIRECTIONS 371
9.22.9 SUMMARY 372
9.22.10 REFERENCES 372

9.22.1 INTRODUCTION

Hate crimes are an international problem. During the twentieth century, mass hatred has resulted in genocide in several countries, including Turkey, Germany, Nigeria, Bangladesh, Bosnia, and Cambodia, among others (Kressel, 1996). It also occurs on a daily basis and takes various forms. In several countries, violent offenses against foreigners have increased related to issues of marginalism, ethnocentrism, and economics (Schneider, 1994). For example, skinhead youths have assaulted foreigners in Germany and urged the eviction of immigrants from France. West European countries have had their highest unemployment rates since World War II, and many people in these countries have blamed immigrants for job shortages. In France, the strong right wing has enormous hostility towards immigrants, particularly dark-skinned Muslims, Arabs, and Africans. Similar hostility has been evident in other countries, such as in Italy and Austria.

Violence due to hatred develops from prejudice and bigotry. In terms of a developmental pathway, it is thought that violence due to hate begins with avoidance, leads to negative speech, and is followed by discrimination and physical attack, ultimately ending in ethnic cleansing and extermination (McLaughlin, Brilliant, & Lang, 1995). Hate crimes are not restricted to people of other racial or ethnic groups and may involve people with various cultural differences. For example, in some segments of the United States, hate crimes are perpetrated on people who are gay and lesbian because they are seen as being sexually deviant. In this chapter, we focus on the victims of violence due to hate crimes.

9.22.2 HATE CRIMES

Hate crimes are characterized by the symbolic status of the victim (Schneider, 1994). The victim of a hate crime belongs to an "out-group," the existence of which symbolizes what the "in-group" does not want to be. The in-group can be defined as any group of people who share a sense of belonging and common identity. The out-group is perceived as being distinctly different in some way from the in-group (McLaughlin et al., 1995) and often is subjected to negative stereotypes (Craig & Waldo, 1996). Weiss (1990) has used the term *ethnoviolence* to describe acts directed against an individual, group, or institution because of race, religion, ethnicity, or sexual orientation. In this context, the in-group does not have to be in the majority and the out-group does not necessarily have to be in the minority (Hidden Hate Crime, 1996).

9.22.2.1 Definition

The US Federal Government has defined hate crimes as, "Crimes motivated by hatred against a victim based on his or her actual or perceived race, color, religion, national origin, ethnicity, gender, or sexual orientation" (United States Congress, 1992). The US Attorney General is required to collect data and publish an annual summary on crimes that manifest prejudice based on race, religion, sexual orientation, and ethnicity (McLaughlin et al., 1995). In 1994, physical and mental disabilities were also recognized as a basis for hate crimes in the United States (United States Department of Justice, Federal Bureau of Investigation, 1997b). Indeed, people with disabilities are over-represented as victims of violence and hate crimes when compared to their numbers in the general population. For example, the incidence of child abuse, domestic violence, rape, and other violent crimes are far higher in people with disabilities than in people without physical and mental disabilities (Waxman, 1991).

Although crimes committed because of gender are not included in the US Federal definition, it must be noted that many crimes against women (e.g., stranger rape) share some of these characteristics (Levin & McDevitt, 1993). For example, a few years ago, a Caucasian female jogger was attacked in Central Park in New York City by a group of young men. It was initially thought that the attack was related to racial hatred because the attackers were African-American and Hispanic-American teenagers. However, it was later discovered that these teenagers had thrown rocks and bottles at cars, cyclists, joggers, and homeless people. They had yelled racial slurs at a couple on bicycles, attacked several Latinos, and then they had decided to "get a woman jogger" (Levin & McDevitt, 1993), suggesting that it was a deliberate act of violence perpetrated on a person because of her gender.

A review of the phases of ethnopolitical activity shows that in the advanced stages of development, groups which intend to overthrow an existing regime become increasingly violent. In the first and second phases of development, leadership develops and begins to attract a following, isolating these supporters from competing organizations. In the third phase, ethnically- and politically-based marches and demonstrations take place. In the fourth phase, sporadic, spontaneous ethnically-based riots, murders, or bombings occur and, in the final phase, premeditated ethnically-based riots, assassinations, terrorism, and coups occur (Dutter, 1987). Ethnic violence often occurs when material interests are a significant concern, economic problems are externally imposed, and interethnic disputes are not successfully resolved (Denemark, 1994).

A hate incident is an event that may involve the use of "harmful words or actions motivated by prejudice against a person or property which do not fall into any criminal category" (McLaughlin & Brilliant, 1997). Even though the physical sequelae of the hate incident may not be as great and the perpetrator is not punishable by law, hate incidents may have the same impact as hate crimes on the individual and the community. However, there may be an emotionally intensifying factor because legal recourse is not available for the victims. Although this chapter is primarily about victims of hate crimes, it should not be overlooked that victims of hate incidents have a similar sequelae.

Verbal harassment is a symbolic form of violence which reminds the victim of the potential of physical assault. It reinforces the sense of being a member of an out-group, and emphasizes the fact that the individual is a socially sanctioned target for violence and hate. Despite the fact that there is no physical injury, the victim's sense of security is violated and, although some victims of verbal assault may minimize its effects, the resultant emotional scarring may be more difficult to heal than physical wounds. This minimization may make it more difficult for the victim to understand any ensuing feelings of fear or self-hatred that he or she may experience. In instances of physical assault, the fact that there is some objective measure of how severely the victim has been injured can serve as a concrete springboard for understanding its emotional sequelae.

The factors involved in the victim's response to hate speech are complex and variable (Cowan & Hodge, 1996). Further, frequent experience with verbal assault can lead the victim to restrict behaviors and become isolated, just as an actual incident of physical assault might, because of the resultant fear that physical violence will soon follow (Garnets, Herek, & Levy, 1990). Further, actual or perceived tolerance of verbal assaults by law enforcement officials "can easily lead to more serious forms of assault. Youths who get away with screaming racial epithets and urinating on the lawn of a family that has recently moved into a neighborhood may well move on to engaging in physical violence and arson" (Wexler & Marx, 1986, p. 220).

9.22.2.2 Categorization and Typology

There are several characteristics of hate crimes which distinguish them from other types of crimes. Hate crimes most often are excessively brutal. There is a greater likelihood of personal violence in a hate crime as opposed to other types of crimes. Further, force is used in a hate crime beyond that which is needed to subdue, disarm, rob, or ensure compliance of the victim. For example, while approximately 75% of victims of hate-related assaults sustain some physical injury and 30% require hospital treatment, only 29% of victims of assault not related to hate sustain injury and only 7% require hospital treatment (Levin & McDevitt, 1993).

In general, victims of hate crimes are total strangers, picked at random, and are seen by the perpetrators of violence as interchangeable objects rather than as people like themselves. The global characteristic of the person is the focus of the attack, whether it be racial (i.e., African-American, Chinese, Asian-Indian), religious (e.g., Jewish, Amish), or based on gender or sexual orientation. Hate crimes are usually perpetrated by multiple offenders who have come together for the purpose of attacking the members of an out-group. For example, reviews show that about two-thirds of all reported hate crimes involve two or more perpetrators, and many with four or more perpetrators (Levin & McDevitt, 1993).

There are several forms of hate crimes, including verbal harassment, posting symbols or slogans of hate, the destruction of religious property, cross burnings, arson, personal assault, and homicide. For example, in a study of victims of hate violence, Barnes and Ephross (1994) reported that 44% of the victims had experienced only one attack, another 10% had experienced multiple unrelated attacks, and the remaining 46% had experienced multiple attacks that were connected in some way. The most frequently reported crimes were physical assault, verbal harassment, and mail or telephone threats, followed by the posting of symbols of hate on or near the victim's property, and attacks on victim's homes, including firing shots and throwing things on the property. In another study, Weiss, Erlich, and Larcom (1991) found that 27% of all "prejudiced-based" incidents occurred in the workplace. These incidents included property damage, robbery, harassing language, physical assault, sexual harassment, and rapes. Further, Weiss et al. (1991) reported that 34% of those who experienced such incidents at work did not know the perpetrator, and an additional 10% had never met the perpetrator before the incidents occurred.

Hate crimes can be categorized according to the perpetrators' motivation for committing the crime. Approximately two-thirds of hate crimes are motivated by thrill-seeking, in which the perpetrators are seeking excitement through terrorizing and bashing. The majority of these crimes are perpetrated by teenagers and young adult males who do not have a criminal record and who live at a distance from the scene of the hate crime. The other third of the hate crimes, categorized as defensive crimes, are perpetrated by people who are reacting to a particular set of outsiders that has infiltrated their turf (e.g., neighborhood or university). These perpetrators are likely to be adults who live or work near the scene of the hate crime. Neighbors or coworkers may share the sympathies of these perpetrators and hinder police investigations. Fewer than 1% of hate crimes, categorized as true believer crimes, are perpetrated by members of an organized group who are carrying out a mission to rid the world of a target group of people (Levin & McDevitt, 1995).

9.22.2.3 Reporting of Hate Crimes

The reporting of hate crimes is related to several factors, including the local definition of hate crime. Different countries classify a hate crime as one which is based on one or more of these characteristics: race, religion, ethnicity, sexual orientation, gender, or disability, or that the motivation for the crime is based on hate. For example, the Ottawa Police Service in Canada uses the criterion that the crime is motivated in whole or in part by hate. In Great Britain, crimes are classified on the basis of racial hatred and do not include other factors (Roberts, 1995).

It is estimated that only 10–30% of hate crimes are reported to law enforcement (Roberts, 1995). It is possible that victims or witnesses fear retaliation by the perpetrators. In the case of property crimes, only increased police surveillance can prevent further episodes. However, data from several European countries suggest that reporting property crimes only serves to provoke further episodes of vandalism because police cannot provide continual surveillance (Roberts, 1995). In some cases, members of groups being victimized may feel that their concerns regarding the hate crimes committed against them will not be taken seriously. In other cases, immigrants may have only negative experiences with law enforcement in their home countries and be afraid to approach the police about hate crimes committed against them. In addition, many new immigrants may experience communication barriers with the police and be unable to report hate crimes. Finally, members of some groups being victimized, such as closeted members of the gay community, may not wish to report hate crimes against them for fear that their secret may be revealed (Weiss, 1990).

In terms of prejudice-based incidents which occur in the workplace, Weiss et al. (1991) found that only 38% of their respondents reported the incident to an authority. Reporting of such incidents was even less likely if the perpetrator was known to the victim or if the perpetrator was the victim's supervisor. Many respondents who did not report these incidents noted that reporting "couldn't help the situation" (Weiss et al., 1991, p. 26), although of those who did report these incidents, about three-quarters of them indicated satisfaction with how the situation was handled. This study indicated that, like hate crimes, prejudice incidents are also under-reported.

It is sometimes difficult to determine if a crime is truly motivated by hatred. In some instances, racial or other epithets are used, or symbols of hatred are posted, enabling law enforcement authorities to assume hatred as the motivation for the crimes committed. In instances where this assumption cannot be made, and assault without specific denigrating phrases occurs, the motivation is not often readily apparent. For example, in the case of the Central Park jogger, there is ambiguity in whether she was attacked because of her race, her gender, or both, or whether she was simply in the wrong place at the wrong time (Levin & McDevitt, 1993). Further, in some instances, an individual may not be aware that he or she has been the victim of hate because other events, such as robbery or burglary, have occurred simultaneously (Kelly, Maghan, & Tennant, 1993).

Certain characteristics of the crime are used as indicators to determine that a hate crime has been committed. For instance, if the perpetrator and the victim belong to different racial, ethnic, or religious groups, or have a different sexual orientation, the likelihood of a hate crime is very high. If there is historical animosity between the two groups represented by the perpetrator and victim, this may indicate that hate was involved. Similar conclusions can be drawn if hate-related comments are uttered or written, gestures are made, or drawings or markings are left at the scene of the crime. Hate is strongly considered if the victim is in a neighborhood which is overwhelmingly populated by people from a different racial, ethnic, religious, or sexual orientation. If the victim or members of the community in which the crime was committed perceive the crime to be related to hate, this is considered a likely motivation for the crime. Another indicator is that the incident or crime occurred on a date of particular significance to the out-group, such as a religious holiday. If the victim was engaged in promotional activities for his or her group, or was a member of an advocacy group that supports the victim's group, this could also indicate hate as the motivation. Finally, if the perpetrator has a history of involvement in hate crimes or is a member of an organized hate group, then the motivation for the crime or incident can be considered to be based on hate (Levin & McDevitt, 1993; McLaughlin et al., 1995). In the USA, it is only necessary for hate to be a contributing factor to the incident or crime for it to be defined as a hate crime (McLaughlin et al., 1995).

9.22.3 PREVALENCE

The US Federal Bureau of Investigation compiles annual statistics on hate crimes (Bishop & Slowikowski, 1995; United States Department of Justice, Federal Bureau of Investigation, 1997a, 1997b). The number of states and law enforcement agencies reporting data on hate crimes continues to grow. For example, in 1995, 45 states were included in the FBI database and, by 1996, this had increased to 49 states and the District of Columbia. These statistics indicate that there is clearly an increase in the reporting of these crimes.

In 1993, there were 7587 hate incidents, 8987 hate offenses, 9372 victims, and 8610 known offenders. Of the hate incidents, at least 16 were murders and 15 were rapes, and the majority of offenses were categorized as intimidation, crimes against property, vandalism, and assaults. The numbers in 1995 were 7947 hate incidents, 9895 hate offenses, 10 469 victims, and 8433 known offenders, indicating a slight decrease in the number of known offenders but an increase in the number of incidents, offenses, and victims. There were a total of 10 702 hate crimes reported in 1996. In all years, over 60% of the offenses were motivated by racial hatred and about 15% by religious hatred. Sexual orientation was the motivation in about 13%, and ethnicity or national origin accounted for about 12% of the hate crimes (Bishop & Slowikowski, 1995; United States Department of Justice, Federal Bureau of Investigation, 1997a, 1997b). A similar pattern of hate crimes was reported in Canada (Roberts, 1995).

In 1993, 85% of victims of hate crimes were individuals, with the remainder being mostly businesses and organizations (Bishop & Slowikowski, 1995). These figures were very similar in 1995, with 83% of these crimes being against individuals. In 1995, 61% of the victims of race-motivated hate crimes were African-American and 24% were Caucasian. About 67% of the victims of crimes based on ethnicity were Hispanic. Of crimes related to religion, about 76% of the victims were Jewish. Of crimes related to sexual orientation, about 72% were directed specifically towards gay males (United States Department of Justice, Federal Bureau of Investigation, 1997a, 1997b). These numbers were consistent with those reported for 1993 (Bishop & Slowikowski, 1995). In 1995, the race of offenders was reported for 62% of the registered hate crimes; 59% of these perpetrators were Caucasian and 27% were African-American. The remaining perpetrators were from other racial groups or had a multiracial heritage. Unfortunately, data on crimes related to disability have only been collected since January 1997, and summary statistics are unavailable at this time (United States Department of Justice, Federal Bureau of Investigation, 1997b).

About 17–26% of all hate crimes can be attributed to juveniles (i.e., those under the age of 18 years), and approximately half of the individuals arrested for hate crimes are between 16 and 25 years of age (Bishop & Slowikowski, 1995). Further, there may be a correlation between involvement in hate crimes and other gang-related crime; this relationship appears most clearly with skinhead gangs and other hate groups (Bishop & Slowikowski, 1995).

Since the late 1980s there has been a significant increase in hate crimes in high schools and on college campuses (Chun & Zalokar, 1992). For example, a Harris poll in 1990 found that over half of high school students witnessed racial confrontations either "very often" or "once in a while," with 25% of the students having been a target of racial confrontation (Center for the Study of Sport in Society, 1990). In another study, it was found that violent hate-related incidents had occurred in 34% of elementary schools, 47% of middle schools, and 42% of high schools in Los Angeles county (Merl, 1989). These incidents had been reported against students of all races and ethnic groups, as well as against Jewish, gay, lesbian, and immigrant students. Only 5% of reported incidents involved members of formalized hate groups, such as white supremacist organizations, and 47% of all incidents had involved "racial slurs and name calling." Physical violence occurred in 25% of the incidents. Some hate-based incidents in schools and colleges are retaliatory, perpetrated by victim groups. For example, some Asian-American and Hispanic-American youth have either joined or formed gangs in order to defend themselves from hate crimes (Chun & Zalokar, 1992). While it is understandable that victims of hate crimes need to feel safe in school and colleges, retaliation and gang membership often serve only to reinforce racial and ethnic stereotypes and to escalate the tension between opposing groups.

The prevalence of violence in the United States is related to its apparent glorification in television, movies, and contact sports, as well as in other aspects of mainstream American culture (Chen, 1994). This cultural preoccupation with violence often contributes to anti-ethnic and other forms of hate crimes. Other societal factors contributing to violent hate crimes include the presence of overt and institutional racism, poverty, and inequality which may contribute to internal and interpersonal conflict (Chen, 1994).

Although it is difficult to determine the exact prevalence of hate crimes in the USA, it is likely that this type of crime has increased in the past several years. In her review, Weiss (1990) noted six factors which may have contributed to the increased rate of hate incidents and crimes in the United States. These include: (i) increased migration of foreign-born people to the United States since the late 1980s; (ii) increased pace of social changes, such as those due to the second wave of the women's movement, the gay–lesbian rights movement, and the implementation of affirmative action programs; (iii) elimination of many federally funded social programs that benefitted the poor, a disproportionate number of whom are from culturally diverse populations; (iv) housing segregation has continued, and violence has been directed against people of diverse cultures new to a neighborhood; (v) increasing gap between the rich and the poor, leading to intense competition for scarce resources and violence among minority groups in inner cities; and (vi) increased reporting in the media of ethnoviolence as dramatic and spectacular events rather than everyday phenomena which occur when large numbers of people live together. Individually and interactively, these factors may contribute to increased violence due to hate crimes.

9.22.4 REACTIONS OF VICTIMS

There are several known reactions that a person who is a victim of personal crimes might experience. These include intense rage or anger, fear of injury, death, and future victimization, sadness, depression, and lowered self-esteem. These reactions have been found in victims of personal crimes, such as assault and rape (Burgess & Holmstrom, 1974). Although victims of different types of crimes use different coping methods, many are also used by victims of hate crimes (Wirtz & Harrell, 1987). However, one major difference is the absence of lowered self-esteem in victims of hate crimes as a result of the attack. Victims of hate crimes often associate the crime with the perpetrator's prejudice and racism, and not with their own personal characteristics.

In general, anyone who is the victim of violence questions a certain basic life assumption—the everyday denial of vulnerability which allows the person to feel safe and secure. The survivor of violence asks: "why me?" and may internalize the stigma attached to violence and question his or her own self-worth (Garnets et al., 1990). There are three stages of reaction to victimization, which may take years to complete (Bard & Sangrey, 1979; Sales, Baum, & Shore, 1984). The first stage is impact, in which the victim feels vulnerable, helpless, and dependent on others. The second stage is recoil, in which the victim may experience mood swings, the waxing and waning of feelings of fear, rage, and revenge fantasies, as well as displacement of anger. In the third stage, reorganization, the experience of victimization is assimilated and put into perspective; the person is able to move on from the experience.

In studies of victims of sexual assault, a victim's personal characteristics, including age, socioeconomic status, race, psychological problems, life stressors, and relationships are all factors which may contribute to the individual's reaction to victimization. A higher degree of violence in the assault is positively related to increased emotional recovery time after the assault. Similarly, the use of weapons increases the strength of the reaction. The role of the victim's relationship to the perpetrator is unclear. The postassault factor of criminal justice involvement and reactions of the victim's support system are also contributing factors (Sales et al., 1984).

The process of pressing charges against the perpetrator can result in a decrease in post-assault symptoms, although there may be a recrudescence as the legal process proceeds (Sales et al., 1984). Schultz (1987) noted that the legal system is a means of channeling "natural vengeance and retaliation" (p. 819). The role of family support may differ with the progression of time after an assault and with the severity of the assault. Further, as Sales et al. (1984) noted the "... impact of violent assault must be viewed within the context of the victim's total life experience. If other conditions are benign, the victim can be expected to weather the crisis more easily. If they have never stockpiled resources, or if their psychic capital is temporarily depleted, they may be more vulnerable" (p. 132).

Victims of various types of crimes have similarities in symptomatology, but may also experience diverse stressful feelings. For example, Garnets et al. (1990) noted that "common reactions to victimization include sleep disturbances and nightmares, headaches, diarrhea, uncontrollable crying, agitation, and restlessness, increased use of drugs, and deterioration in personal relationships" (p. 365). In addition to these reactions, victims of hate crimes also experience a significant amount of psychological distress. For most victims, there is an experience of severe psychopathological aftermath. For people of color, gays and lesbians, and members of some religious groups, there is often the problem of society's continued hostility toward them. First, the victimization of the person involves interference with daily activities. The individual believes that the world is unpredictable and unsafe. Bard and Sangrey (1979) noted that because their victimization did not occur as a result of accidental or natural forces but with malevolent intentions perpetrated against them, victims developed a significant reduction in basic trust.

Second, the "experience of victimization interferes with perceptions of the world as an orderly and meaningful place" (Garnets et al., 1990, p. 368). Often, victims blame themselves for the act. Third, the victim's view of him or herself is important in order to deal with the victimization. The individuals may view themselves as weak, which may reinforce a negative self-image. For all victims of hate crimes, the difficulties in dealing with vulnerabilities, meaningfulness, and positive or negative self-perception are important factors in how he or she will deal with inevitable psychological stressors.

In terms of the perpetrator's motivation for hate crimes, the victim of a "thrill seeking" crime feels more vulnerable to random attack. The typical victim of the "defensive" crime is a person of color. This victim is aware that his or her particular situation may have precipitated the attack and that he or she could be attacked again. This victim feels frustrated because he or she knows exactly why the incident occurred (she "took a stand" when she moved into a new home, accepted a promotion, transferred to a new school, or dated someone of a different race, ethnicity, or religion). For these individuals, "taking a stand" may result in a cross burning on their lawn, a student getting beaten up at school, or an employee getting harassed at the workplace (Levin & McDevitt, 1993). There is no specific literature on the victim of the "true believer" crime. It is reasonable to assume that because these can be the most radical of hate crimes, the surviving victims will endure sequelae that are of the same nature, but more extreme than other victims of hate crime.

Barnes and Ephross (1994) studied victims of hate violence from several cities in the USA. They found that the most frequent emotional response was anger at the perpetrator, followed by fear of injury, sadness about the incident, powerlessness, suspicion, and fear of being killed. About one-third of the victims reported that they changed their behavior in some way to cope with, as well as to avoid, future victimization. These behaviors included changing jobs, moving out of the neighborhood, decreasing social participation, purchasing a gun or feeling more ready to use a gun, getting more home security, and increasing their children's safety precautions.

The initial, or primary, victimization suffered by the victim of the hate crime is multidimensional. These victims recognize that because of the nature of the motivation of a hate crime, there is almost nothing that can be done to alter their vulnerability to the situation as the crime is a result of a core aspect of their identity (McLaughlin et al., 1995; Roberts, 1995). As the attack is based on an immutable aspect of their identity, this can lead to a sense of increased vulnerability and resultant hopelessness. Further, there may be a temporary rejection of that aspect of their personality. As such, there can be a more devastating psychological impact when compared to those who are victims of other crimes. These crimes are meant to send the message that fear and terror can follow at random. The message is received by the individual victim as well as members of the victim's cultural group. The victim may perceive the perpetrator as a representative of the opinion of society at large and feel like an outcast. Members of other groups that are often targeted for hate crimes are reminded of their own vulnerability and may feel victimized as well (Craig & Waldo, 1996).

The message can be compounded if the hate crime is inadequately handled by law enforcement officials, or by the general community, who can unwittingly exacerbate the trauma (McLaughlin et al., 1995). The victim may encounter institutional prejudice or individuals who share similar biases or stereotypes as the perpetrator. Unfortunately, society has responded more slowly to hate crimes than to other forms of crime. This can result in the victim feeling a loss of a sense of belonging to the community and a feeling of betrayal by "the system." The victim may not receive the kind of support he or she might have expected had the crime not been related to hate. This secondary victimization results from the victim's perceived rejection by the community as well as by the lack of expected support. Members of victimized groups are generally under-represented among criminal justice professionals (Roberts, 1995). If the criminal justice system does not recognize that a crime is related to hate, then the true extent of harm to the victim and the community cannot be taken into account, thus failing the community as well as the victim.

9.22.5 CULTURAL FACTORS AND HATE CRIMES

Cultural factors determine whether a crime victim chooses to report a hate crime to the police or other law enforcement officials and whether the victim connects the crime to any psychological sequelae from the event. Culture also affects how the victim views pain that is experienced, including a sense of hopefulness or hopelessness. Further, it affects the victim's decision as to whether, and from whom, to seek assistance. Certain cultural groups have a sense that life is predetermined and unalterable, with an external locus of control of incidents that happen in their life. Members of these groups react differently from those who have an internal locus of control and a sense of being masters of their own destiny.

Victims may have different reactions to hate crimes related to their own demographic characteristics (Craig & Waldo, 1996). Members of certain racial groups have a strong fear of being victims of crimes. For example, Parker, McMorris, Smith, and Murty (1993) found that both Latinos and African-Americans reported significantly high levels of feeling victimized. They also found that Latinos and women tend to have the highest levels of fear of crime. Explanations for these findings may include being previous victims, location of neighborhoods with high crime rates, and racial discrimination.

Following a racially divisive court verdict, the Korean-American victims of rioting in South Central Los Angeles in 1992 experienced financial loss and sustained severe emotional trauma (Kim-Goh, Suh, Blake, & Hiley-Young, 1995). Four months after the riot, the Korean-Americans suffered symptoms similar to post-traumatic stress disorder (PTSD). In addition, there was a gender difference in their reaction to the riots; the women reported high rates of somatization and the men reported a significant increase in alcohol intake after the riots.

Janson and Ryder (1983) found that, although the elderly are less victimized than younger people, they too are highly fearful. Miethe and Lee (1984) noted that the elderly consider fear of crime their most serious personal problem before issues of income, health, and housing. Another common research finding is that women are more afraid than men of crime, although men are more likely than women to be victims of all types of crime, except for sexual assault (Clemente & Kleimen, 1976; Ferraro, 1996). Despite gender differences in the prevalence of various crime victimizations, women are more afraid of all types of crime. For example, once a woman is raped, she fears future face-to-face types of crimes (e.g., burglary, robbery), including the fear of being raped again. Ferraro (1996) also found that fear of rape is very high among younger women, reflecting national victimization data. Ferraro noted that it did not matter whether females had been personally victimized or not; the imagined horror of sexual assault is sufficient to increase such fear.

Snell and Godwin (1993) reported that men are quick to blame women for being raped, particularly if it is an acquaintance rape. Victims of rape are often reluctant to report to law enforcement officials or to disclose their victimization to family members, due in part, to their "concern that their allegations will be discredited or that they will be blamed for the assault" (Foley, Evancic, Karnik, King, & Parks, 1995, p. 7). Acquaintance and date rapes are even less likely to be reported and the assailants less likely to be sentenced by the criminal justice system than cases of stranger rape. As in classic forms of hate crime, it appears that women in our society are less likely to report rape crimes due to their fear of being "persecuted" again in the criminal justice system

9.22.5.1 Gays, Lesbians, and Hate Crimes

Gays are the most frequent victims of thrill-motivated attacks and hate crimes (Levin & McDevitt, 1993). For thrill seekers, gays are easy to find as gay and lesbian establishments are often concentrated in certain districts of major cities. Gay males, who are more frequently victimized than lesbians, are often seen as a sort of psychosexual affront or threat to the sexual identity development of young adult men, who are frequently the perpetrators of thrill-motivated crimes.

Gays may be less likely to report an attack because of the fear that their sexual orientation may be revealed as a result of publicity of the crime. Berrill and Herek (1992) used the term secondary victimization to "highlight negative treatment of gay people that is legal and widely condoned" (p. 290). Reports of the event may "out" the person with their family, friends and employer, resulting in rejection because of their sexual orientation. This is a form of secondary victimization specific to the gay community (Levin & McDevitt, 1993). A victim who is "outed" is vulnerable to secondary victimization in housing, the military, employment, foster care, and child custody, among others (Berrill, 1991). Often there is no legal recourse if an employer learns of the victim's sexual orientation and fires him. Further, members of the criminal justice system may actively victimize individuals in the gay community due to their own biases against gays. Gays may feel vulnerable to a criminal justice system that is often against them. For example, Berrill and Herek (1992) cited an example in which an Asian-American gay man was beaten to death and the Judge jokingly asked the prosecuting attorney, "that's a crime now, to beat up a homosexual?" (p. 294).

Similar to rape victims, gay victims report feeling blamed by the judicial system for the hate crime. "Homosexual panic defense" is an example of blaming the gay victim (Berrill & Herek, 1992). Homosexual panic defense alleges that the defendant's violent actions were committed in self-defense against the victim's unwelcome and aggressive sexual overtures or were part of an acute psychological panic resulting from those overtures (Chuang & Addington, 1988). The blame is then shifted from the defendant to the victim for causing the hate crime. The homosexual panic defense has resulted in lenient sentences for those charged with assaulting or murdering gays.

The judicial system's biased stance against gay people also affects whether gays will report crimes committed against them. Gays may feel that it is better to remain quiet than cause any further problems for themselves because the system appears to be against them. Because many gay and lesbian victims do not report hate crimes related to their stigmatized status, they are viewed as "easy targets" by perpetrators of hate crimes (Berrill, 1991). Clearly, many gays who are victims of hate crimes feel that they are unsupported by the judicial system, and a judicial system that does not protect all individuals regardless of their race, religion, or sexual orientation is in effect contributing to their victimization.

Although we might expect that universities and colleges would value diversity on their campuses, antigay harassment flourishes in these environments. Verbal harassment, intimidation, and vandalism are the primary forms of hate, although physical assault is not uncommon, occurring in 5% of cases (Berrill, 1991). Up to 25% of gay college students have been threatened with violence and up to 75% have experienced verbal harassment (Berrill, 1991). A survey of gays and lesbians found that 20% had experienced an antigay assault and another 20% had been victims of a property crime (Herek, Gillis, Cogan, & Glunt, 1997). Half of the respondents had suffered verbal harassment or threats in the year before the study, with the majority having experienced at least one episode of harassment since age 16. Those who survived an assault had high levels of anxiety, anger, and symptoms of depression and PTSD.

In terms of perpetrators of hate crimes against gays and lesbians, "activists" are the smallest group of individuals who will seek gays out, even in out of the way places, in order to assault them. "Opportunists" will assault gays when the opportunity presents itself. The victims in these cases are most likely people who are overtly homosexual and conform to the perpetrator's image of how a gay person "should" appear (i.e., effeminate or hypermasculine appearance or behavior).

Unlike other target groups, gays and lesbians experience hate and violence in the home as well as in society generally (Hunter, 1992). Surveys have shown that up to 41% of gays have been the victim of insults or intimidation by family members and about 8% have experienced physical violence (Berrill, 1991). In a review of the victimization literature, Herek et al. (1997) reported that victimization has deleterious effects on the mental health of gay and lesbian youth, both directly and through effects on family support and self-acceptance.

Antigay violence is highly concentrated in prisons and jails. This form of hate crime is difficult to avoid in prisons because many law enforcement officials see it as an inevitable consequence of incarceration (Berrill, 1991).

The risk of violence to gays and lesbians of color is higher because of what seems to be a higher rate of homophobia in minority populations when compared to people in society generally (Mays, Cochran, & Rhue, 1993). As most of the gays and lesbians of color who have been studied are African-American or Latino, it is also possible that these higher rates of victimization are related to the higher risk of violent crime experienced by these groups in the United States (Berrill, 1992a). For example, Comstock (1989) found that gays and lesbians of color were more likely to be chased, followed, have objects thrown at, and physically assaulted than the majority population. They were also more likely to be victimized in gay-identified areas.

Lesbians of color are more likely to experience physical violence, rape, threats, and vandalism, while White lesbians are more likely to experience verbal assault (von Schulthess, 1992). As adolescents are more frequently victims of violent crime than adults, adolescent gays and lesbians have a higher risk for being victimized because of their sexual orientation. Gay and lesbian adolescents are also at a higher risk of violence because of the endemic homophobia that is seen in the teen population. Some of this homophobia is related to the epidemic of human immunodeficiency virus (HIV) infection and the association of contagion with homosexual behaviors. This population of gays and lesbians is also at greater risk for victimization because of their increased visibility in society.

Gays and lesbians face an additional challenge, often referred to as "heterosexism" (Garnets et al., 1990). This has been defined as an "ideological system that denies, denigrates, and stigmatizes any nonheterosexual form of behavior, identity, relationship, and community" (Herek, 1990, p. 316). The gay individual who is attacked must deal with our society's heterosexism in organized religion, law, media, and the military.

9.22.5.2 College Students and Hate Crimes

Racism, sexism, heterosexism, ethnocentrism, and other similar biases are prevalent on college campuses (Berrill, 1992b; Herek, 1993). As in society at large, majority vs. minority and minority vs. majority conflicts are present on college campuses. Incidents may include racial jokes, as well as "racist slurs and posters, racial harassment, intimidation; anti-Semitic remarks, graffiti, posters; and threatening statements toward lesbians and gays" (US Commission on Civil Rights, 1990a, p. 4). The US Commission on Civil Rights (1990a) has noted four categories of contributing factors to hate incidents on campus: (a) deficiencies on college campuses that exacerbate existing tensions; (b) society's failure to keep up with the change; (c) competition for limited resources; and (d) extremist speakers" (p. 7). This commission noted that isolation, perceived ethnically- and racially-based hostility on campuses, and defunding of federal programs directed towards minorities are among the factors considered deficiencies on campus. Deficiencies also include the perceived inadequacies of the institution in recruitment and retention of minority students and faculty.

As college campuses become more multicultural, ignorance of others' cultures becomes more unacceptable in our society. Insensitive comments and behaviors are a result of society's failure to keep up with this change. As the size of any group of people grows, demands for recognition of the group's needs by the institution are more likely to occur, with ensuing conflict (US Commission on Civil Rights, 1990a).

In the United States, there is often an oversimplification of racial differences into a "black vs. white" dichotomy, neglecting other ethnic groups which are also minorities on campus (Nebraska Advisory Committee to the United States Commission on Civil Rights, 1990). When financial resources on campuses decrease and the number of groups who qualify for these funds increase, competition for these funds develops, adding fuel to the fire of racial bigotry. In addition, high school students often enter college without a familiarity with their country's civil rights history, which resulted in the appropriation of funds to correct for institutional racism. Based on this limited knowledge, they develop a perception of being discriminated against in favor of a minority student population, thus contributing to campus tension between majority and minority students. Many campus hate incidents occur not only after the presentation of a socially extremist speaker, but also in the months preceding the presentation. During this time, the polarized controversy that is generated may increase both negative and positive publicity for the speaker and the campus, heightening the tension between opposing cultural groups (US Commission on Civil Rights, 1990).

There are at least three levels of racism on college campuses: "(1) students who, out of ignorance or deliberate insensitivity use different comments, actions, and hold a particular attitude towards [a minority group]; (2) an administration which is sometimes not sympathetic to [minority] groups and rights; and (3) law enforcement actions" (California Advisory Committee to the United States Commission on Civil Rights, 1991, p. 9). Each of these factors, individually and collectively, contribute to hate incidents and crimes.

Farrell and Jones (1988) described three types of racial incidents on campus: White insensitivity, environmental racism, and minority and majority student characteristics. They noted that a "sizable proportion of white students, white staff, and white faculty at predominantly white institutions of higher education have experienced minimal contact with and exposure to minorities, particularly blacks, Hispanics, Asians, and American-Indians. In addition, the general white population has not been encouraged or required to expand its social and cultural vistas in order to be sensitive to the social, economic, and cultural backgrounds of minority students" (p. 215).

In addition, they noted that business establishments in the university's immediate surroundings may be racially intolerant, thereby contributing to increased discrimination within the university. Examples include campus bars deterring minority students by not playing "minority-oriented music" or not hiring minority students as workers, restaurants being slow to serve or outwardly hostile towards minority students, local police "closely monitoring" minority male students when they are in the surrounding community's predominantly White establishments or neighborhoods, and local public officials rarely speaking out against racism in the community.

Chih (1994) studied undergraduate Asian-American students to assess their coping and symptomatology following an episode of racially motivated violence on campus. He interviewed the 10 victims of a combined physical and verbal assault and 11 other students who were not present but were very familiar with the incident. About 80% of those surveyed had experienced some sort of hate incident because of their race. Half of the students had been stared at because of their race or had been mistaken for another person of their race.

In the incident, an Asian-American student organization sponsored a dance on campus. As 10 Asian students were cleaning up after the dance, four young White men approached them, yelling ethnic slurs. As the incident escalated, one Asian-American student was assaulted and required hospital treatment. Chih (1994) reported that those who were present at the time of the incident experienced more somatic symptoms than those who were not. They felt helpless and sad, experienced headaches and nightmares, and had difficulty concentrating on their studies. They were more concerned with loss of face than the nonvictims. This group used more internal coping methods to overcome the effects of racial victimization. Over half of the victims did not use their family members as a support, with some indicating that their parents would not understand because of intergenerational differences. Some students felt partially responsible for the incident because it had occurred at their dance on campus, and they expected that their parents would feel that this social event should not have interfered with their education. These students also indicated that they did not want to worry their parents. Most students used their friends for emotional support.

Some students indicated that they felt partially responsible for the incident because they yelled back or did not simply walk away from the assailants. They noted that "Asians are targets of this sort of violence because it's sort of expected that we will not fight back" (Chih, 1994, p. 35). Some students had never been exposed to such blatant racism and were psychologically unprepared for it. Those subjects who were not present at the incident displayed more passion, indignation, and anger about the incident, and wrote letters to the campus paper deploring the racially motivated incident.

Farrell and Jones (1988, p. 220) listed the following several risk factors for hate incidents on predominantly White college and university campuses:

(i) Large enrollments in general terms or in residential enrollment as a percentage of the student population.

(ii) Campuses located in isolated college towns.

(iii) Unclear rules for student conduct and weak or nonexistent sanctions for racist or sexist behavior.

(iv) Substantial numbers of low-income, working-class majority students whose previous residence has been in majority or nearly majority cities and towns.

(v) Weak leadership at the central administrative level of the university with respect to a consistent and clear commitment to racial and gender equity.

(vi) Failure by the university administration to act swiftly and decisively once a racial incident has occurred.

(vii) Ineffective minority student support programs.

(viii) Small numbers of minority faculty and staff at the institution.

(ix) Lack of university programming or initiatives to bridge the cultural and social gap between majority and minority students.

Clearly, one or more of these risk factors are present in most, if not all, university campuses.

9.22.6 INTERVENTIONS

9.22.6.1 Mental Health Treatment

Specific interventions can be targeted for specific phases of a victim's reaction to hate crimes (Garnets et al., 1990). During the impact phase, the mental health professional engages in crisis intervention and assessment of the victim's coping resources. Crisis intervention is focused on ensuring safety, including referral for any required acute medical treatment. This intervention also involves assessing the emotional meaning the victim gives to the event, and the victim's self-assessment in general. The coping resources which should be assessed include learned methods of coping and support networks (i.e., family, friends, significant other). The mental health professional may also connect the individual with community networks, such as victim assistance programs.

During the recoil phase, the mental health therapist's role is to allow ventilation and to provide support. The victim needs to ventilate the horror and terror he or she has experienced, express anger towards his or her assailants in a constructive manner, and progress towards understanding the incident or crime in terms of global hatred. Encouraging involvement in self-defense classes or activist groups formed to prevent violence or support members of a particular outgroup is often a constructive means of directing anger. For certain individuals, the issue of whether or not to report the hate crime is discussed at this stage in therapy, and the issue of potential secondary victimization is addressed. Emotional support is needed as the victim regains his or her self-confidence and processes feelings of guilt, shame, and helplessness. Building additional support networks is beneficial at this stage. Initial treatment of specific psychiatric syndromes, such as PTSD, also occurs during this phase. At times, work with the victim's significant other, as well as family members, may be necessary in order to help support the victim's psychological and social well-being. This may take the form of education about the causes of violence and hate crimes, and how victims cope with the violence and hate crime they have been subjected to.

In the final phase, the reorganization period, the mental health therapist's goal is to help the victims integrate the traumatic experience into his or her world view, and to move on with his or her life.

9.22.6.2 Advocacy

The US Department of Justice and other such agencies compile statistics which show that crime is a major worldwide problem. Greenberg and Ruback (1984) noted that our society has become increasingly aware of and concerned by the problems encountered by victims of crimes. As a response, victim service programs have been developed (Schultz, 1987). These are general victim service programs, with no particular attention paid to victims of hate crimes (Barnes & Ephross, 1994). We review some of the services provided by these programs, with the understanding that additional services may be required by those who are victims of hate crimes.

Victim assistance programs have several responsibilities in working with victims of hate crimes. In her review, Young (1988) noted eight stages in victim support services: emergency response, victim stabilization, resource mobilization, after arrest, precourt appearance, court appearance, presentence, and postsentence. Emergency response includes ensuring the provision of first aid and medical services, as well as the initial availability of food and shelter. Acute emotional support includes comforting statements, counseling, and home visits immediately after the person has been victimized (McLaughlin et al., 1995). Young noted three important messages that must be conveyed to the victim in the first stage: "I'm sorry it happened. I'm glad you weren't killed. It wasn't your fault" (p. 334) and suggested that these statements can be used as a "psychological first aid."

Victim stabilization generally occurs at the crime scene or within the first few days. During this time, the victim is struggling with the initial feelings of being out of control while police are gathering statements. Thus, enabling control by the victim, by offering choices whenever available, is helpful: "what would you like me to call you—Mr. Mertz or Fred?" "Is there anyone I can reach to help you now?" During resource mobilization, practical needs are met, such as providing new locks on doors, replacing windows, or assisting in the retrieval of stolen property from police custody. Short-term counseling services may be offered through the advocacy program, and referral may ultimately be made for longer term treatment. Similarly, referrals may be made, at any phase, for legal and social service assistance on an as-needed basis.

The final five stages are only experienced by those victims whose assailants are arrested. Increasingly, victim services programs are available to provide services through these last painful and often confusing stages. In the postarrest stage, victims want to know as much as possible about bail, pretrial hearings, and plea bargaining, with the hope that their experience at the hands of the defendant will not be discounted for the defendant's rights. In the precourt appearance stage, the victim still has the "need to know." Frustration occurs when court dates are changed without notice when the victim is to testify, when the victim does not clearly understand what is expected of him or her in the courtroom, and with the difficulty of cross-examination.

At the time of court appearance, the victim must "relive the facts of the criminal incident itself. The very recounting of the trauma, especially in that setting, often triggers a re-experiencing of the crisis and all its manifestations" (Young, 1988, p. 342). The confrontation of the assailant is also an emotional hurdle, particularly if the victim is fearful of retaliation should the defendant be released or receive a brief sentence. The arguments of the defense attorney presenting the assailant as deserving of release also may be difficult for the victim to deal with. Advocacy may include ensuring that there are few, if any, repercussions from the employer as the victim spends time in court. The victim assistance program may have the victim observe other trials in order to become familiar with the system or provide an escort on the day he or she is to testify to help him or her through the process. In the presentence phase, the programs may assist the victim in preparing impact statements for the court. There may be compensation and restitution forms with which the victim may need assistance.

In the postsentence phase, there may again be an intense need for mental health services (McLaughlin et al., 1995). Young (1998) has aptly noted that, "Many victims essentially put themselves and their needs aside while they pursue justice through prosecution and conviction. Their sense of mission, of seeking the restoration of fairness in their lives, overwhelms other aspects of their emotional lives. Hence, it is common for victims to return from a sentencing hearing, even when the sentence is exactly what the victims wanted, feeling depressed and isolated from the world. They suddenly feel alone and detached from other things. The flurry and frustration of the trial are over. There is nothing more to be done. Their purpose either has been accomplished or has been flouted. And the activity that kept their grief from surfacing has ceased" (pp. 344–345).

The advocate for the victim is aware that the victim will want information over time about the parole and probation status of the assailant. Some victims may be allowed to give input at parole hearings, and there may be civil suits after criminal proceedings are concluded. Victims of heinous crimes and the survivors of murder victims must continue the process as appeals are filed, with repeated calls to testify, and a prolonged inability to find closure to the legal sequelae of their experience.

In terms of campus hate crimes, the welfare of the victims of the incident should be the foremost focus of the advocacy group, and this should include the decision whether or not to publicize the incident. The advocacy group should support the victim regardless of the decision made, particularly in the case of antigay and lesbian incidents. The use of a facilitated discussion group may be helpful for members of the victim's group on campus. Emotional support, including reminding the student of the availability of counseling, is also important because of the sense of isolation felt after a hate incident.

Scrupulous documentation of discrimination serves to prove to campus administration that the event is not an isolated one. This helps to build a case that official policies, if they do not exist or are not effective, must be set in place and fully implemented. Such documentation includes careful recording of each incident, including names or descriptions of perpetrators, keeping photocopies of hate mail, and taping hate messages on answering machines. The victim's written permission should be obtained if there is any identifying information.

The advocacy group must clearly set goals and establish strategic planning in order that group members agree and are able to work together effectively. Developing alliances with other minority or women's groups are important, as all these groups have similar concerns. It is recommended that the advocacy group have a clear understanding of the institution's procedures for setting policy on hate concerns. The comprehensive documentation gathered must be presented at any meeting with the administration, including presentations by any victims who are willing to describe their experiences. If the administration is not responsive, it may be necessary to seek legal action, involve the media, or have a demonstration (Berrill, 1992b). In general, college and university administrations must prohibit any form of discrimination, encourage the reporting of hate incidents, have clear policies and procedures for addressing these incidents, developing a taskforce if necessary, collect data about incidents, assist victims, discipline perpetrators, and educate the campus and the community at large (Berrill, 1992b). There would be far fewer hate crimes if all colleges and universities were to follow these guidelines.

9.22.7 PREVENTION

An excellent means of helping victims of hate incidents and crime is through prevention. School prevention programs are an important component in teaching children about prejudice and its alternatives early in life. The US Federal Government has sponsored research towards the development of a curriculum, "Healing the Hate," designed specifically for this purpose (McLaughlin & Brilliant, 1997). Pilot-tested in several settings, this middle school curriculum is directed towards young adolescents and focuses on the following general principles:

(i) *Violence and prejudice are preventable.* In these lessons, the teacher fosters the knowledge that prejudice and violence are learned and, as such, can be prevented.

(ii) *Early intervention.* This principle addresses the knowledge that early education and intervention can prevent prejudice.

(iii) *Empathy building and perspective taking.* For these principles, through the promotion of respect for diversity, the curriculum helps students develop empathy skills. Empathy skills have been shown to be effective in reducing prejudice.

(iv) *Awareness and appreciation of differences.* In the process of helping youth understand and appreciate their own identity, the ability to also respect and appreciate others is inculcated.

(v) *Cooperative learning, critical thinking, and interactivity.* With these principles, several students work together in small groups using critical thinking principles to understand the other concepts.

(vi) *Media literacy.* Students learn critical viewing skills while exploring how relevant issues are presented in the media.

(vii) *Inequality and institutional violence and prejudice.* Recognizing that injustice and inequality are often at the root of hate crime, the curriculum explores institutional racism and other such factors.

(viii) *Social responsibility.* The final principle involves encouraging students to make a difference in attempting to reduce hate incidents and crimes.

These principles provide holistic guidelines for teaching our children to nonjudgementally accept others who may be different from them in some way. Diversity and acceptance training are imperative in the prevention of hate crimes. Education about bigotry and resultant hate crimes should occur at all levels: elementary, middle, and high schools, as well as at colleges, universities, and continuing education sites.

9.22.8 FUTURE DIRECTIONS

Much progress has been made over the last few decades in our knowledge of hate victims and their treatment. However, the sum total of our knowledge about every facet of hate victims is still meager. Some areas of future research include the following:

(i) *Issues of definition*. There is little agreement across legal and service agencies, as well as advocacy groups, as to who is a victim of hate crimes. Federal, state, and advocacy data-gathering agencies need to develop an operationally defined set of criteria that can be used to identify victims of hate crimes. As with the definition of a hate victim, all agencies involved in the collection of data on hate victims need to agree on the criteria to be used for classifying a crime as an instance of hate crime.

(ii) *Minimum data set*. All agencies involved with hate crimes should agree on the minimum data set that should be collected on the characteristics of perpetrator(s) and the victim(s) because such an agreement would allow for aggregation and comparison of data across agencies and geographical boundaries.

(iii) *Interventions*. There is a paucity of literature on interventions with victims of hate crimes. First, there are few well-tested treatment methods that have been empirically validated as effective in reducing the psychological trauma associated with being a victim of a hate crime. Second, few interventions have been described fully enough for independent replication of the procedures. Third, few current interventions sufficiently address the cultural issues that the majority of the victims bring with them. There is an urgent need to understand cultural norms and differences in the victims' methods of coping with the hate crime. Unless psychologists and other mental health professionals adequately deal with the cultural nuances of the victims, no intervention will fully address the psychological and other mental health needs of the victims. Further, there is a need to obtain long-term follow-up data on the effects of hate crimes on the victims, regardless of whether the victims have been in treatment or not.

(iv) *Prevention*. Broad-based, culturally appropriate, prevention efforts with young children are urgently needed. Prevention efforts should begin with children as early as possible so that they do not develop the precursors to violence generally and hate crimes in particular.

Although this is not an exhaustive list, these four general areas cover the core issues that psychologists need to deal with. Fortunately, psychologists have the training and expertise to provide leadership in researching these issues in a systematic and scientific manner.

9.22.9 SUMMARY

Victims of hate crimes are found in every nation and virtually anyone can be the victim of a hate crime. This area of research and intervention has been fraught with problems. The literature is unclear as to the criteria that can be used to classify a crime as an instance of hate crime, thereby making it very difficult to aggregate data on the prevalence of hate crimes. The literature is also unclear as to who is a victim of a hate crime, and what constitutes a "victim." These are critical issues in determining the extent of the problem, and unless we have a fairly good understanding of the extent of the problem, policymakers generally will not take the initiative to formulate policies that will protect and fund treatments for victims of hate crimes. What we do know is that regardless of how a hate crime is defined, it is increasing in our society and better methods for assisting the victims of hate crimes are needed.

9.22.10 REFERENCES

Bard, M., & Sangrey, D. (1979). *The crime victim's book*. New York: Basic Books.

Barnes, A., & Ephross, P. H. (1994). The impact of hate violence on victims: Emotional and behavioral responses to attacks. *Social Work, 39,* 247–251.

Berrill, K. T. (1991). Anti-gay violence: Causes, consequences, and responses. In R. J. Kelly (Ed.), *Bias crime: American law enforcement and legal responses* (pp. 151–164). Chicago: University of Illinois at Chicago.

Berrill, K. T. (1992a). Anti-gay violence and victimization in the United States: An overview. In G. M. Herek & K. T. Berrill (Eds.), *Hate crimes: Confronting violence against lesbians and gay men* (pp. 19–45). Newbury Park, CA: Sage.

Berrill, K. T. (1992b). Organizing against hate on campus: Strategies for activitsts. In G. M. Herek & K. T. Berrill (Eds.), *Hate crimes: Confronting violence against lesbians and gay men* (pp. 259–269). Newbury Park, CA: Sage.

Berrill, K. T., & Herek, G. M. (1992). Primary and secondary victimization in anti-gay hate crimes: Official response and public policy. In G. M. Herek & K. T. Berrill (Eds.), *Hate crimes: Confronting violence against lesbians and gay men* (pp. 289–305). Newbury Park, CA: Sage.

Bishop, E., & Slowikowski, J. (1995). *Hate crime*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.

Burgess, A. W., & Holmstrom, L.L. (1974). Rape trauma syndrome. *American Journal of Psychiatry, 131,* 981–986.

California Advisory Committee to the United States Commission on Civil Rights (1991). *Fair and open environment? Bigotry and violence on college campuses in California*. Washington, DC: Author.

Center for the Study of Sport in Society (1990). *Youth attitudes on racism*. Boston: Northeastern University.

Chen, S. A. (1994). Sociocultural factors in anti-ethnic violence. In J. Y. Fong (Ed.), *Proceedings of the Asian American Psychological Association 1994 Convention* (pp. 22–31). Atascadero, CA: Gold Shadow Press.

Chih, D. W. (1994). College student's coping with anti-Asian violence. In J. Y. Fong (Ed.), *Proceedings of the Asian American Psychological Association 1994 Convention* (pp. 32–37). Atascadero, CA: Gold Shadow Press.

Chuang, H. T., & Addington, D. (1988). Homosexual panic: A review of its concept. *Canadian Journal of Psychiatry, 33,* 613–617.

Chun, K., & Zalokar, N. (1992). *Civil rights issues facing Asian-Americans in the 1990s*. Washington, DC: US Commission on Civil Rights.

Clemente, F., & Kleiman, M. B. (1976). Fear of crime among the aged. *The Gerontologist, 16,* 207–210.

Comstock, G. D. (1989). Victims of anti-gay/lesbian violence. *Journal of Interpersonal Violence, 4,* 101–110.

Cowan, G., & Hodge, C. (1996). Judgments of hate speech: The effects of target group, publicness, and behavioral responses of the target. *Journal of Applied Social Psychology, 26,* 355–374.

Craig, K. M., & Waldo, C. R. (1996). "So, what's a hate crime anyway?" Young adults' perceptions of hate crimes, victims, and perpetrators. *Law and Human Behavior, 20,* 113–129.

Denemark, R. A. (1994). Toward a theory of ethnic violence. *Humboldt Journal of Social Relations, 20,* 95–120.

Dutter, L. E. (1987). Ethno-political activity and the psychology of terrorism. *Terrorism 10,* 145–163.

Farrell, W. C., & Jones, C. K. (1988). Recent racial incidents in higher education: A preliminary perspective. *The Urban Review, 20,* 211–226.

Ferraro, K. F. (1996). Women's fear of victimization: Shadow of sexual assault? *Social Forces, 75,* 667–690

Foley, L. A., Evancic, C., Karnik, K., King, J., & Parks, A. (1995). Date rape: Effects of race of assailant and victim and gender of subjects on perceptions. *Journal of Black Psychology, 21,* 6–18.

Garnets, L., Herek, G. M., & Levy, B. (1990). Violence and victimization of lesbians and gay men. *Journal of Interpersonal Violence, 5,* 366–383.

Greenberg, M. S., & Ruback, R. B. (1984). Criminal victimization: Introduction and overview. *Journal of Social Issues, 40,* 1–7.

Herek, G. M. (1990). The context of anti-gay violence. *Journal of Interpersonal Violence, 5,* 316– 333.

Herek, G. M. (1993). Documenting prejudice against lesbians and gay men on campus: The Yale sexual orientation survey. *Journal of Homosexuality, 24*(4), 15–30.

Herek, G. M., Gillis, J. R., Cogan, J. C., & Glunt, E. K. (1997). Hate crime victimization among lesbian, gay, and bisexual adults. *Journal of Interpersonal Violence, 12,* 195–215.

Hidden Hate Crime (1996, April 8). *National Review, 48,* 18.

Hunter, J. (1992). Violence against lesbian and gay male youths. In G. M. Herek & K. T. Berrill (Eds.), *Hate crimes: Confronting violence against lesbians and gay men* (pp. 76–82). Newbury Park, CA: Sage.

Janson, P., & Ryder, L. K. (1983). Crime and the elderly: The relationship between risk and fear. *The Gerontologist, 23,* 207–212.

Kelly, R. J., Maghan, J. & Tennant, W. (1993). Hate crimes: Victimizing the stigmatized. In R. J. Kelly (Ed.), *Bias crime: American law enforcement and legal responses* (pp. 23–47). Chicago: The University of Illinois at Chicago.

Kim-Goh, M., Suh, C., Blake, D. D., & Hiley-Young, B. (1995). Psychological Impact of the Los Angeles riots on Korean-American Victims. *American Journal of Orthopsychiatry, 65,* 138–146.

Kressel, N. J. (1996). *Mass hate: The global rise of genocide and terror*. New York: Plenum.

Levin J., & McDevitt J. (1993). *Hate crimes: The rising tide of bigotry and bloodshed.* New York: Plenum.

Levin, J., & McDevitt, J. (1995, August 4). The research needed to understand hate crime. *The Chronicle of Higher Education, 41,* B1–B2.

McLaughlin, K. A., & Brilliant, K. J. (1997). *A national bias crime prevention curriculum for middle schools.* Newton, MA: Education Development Center, Inc.

McLaughlin, K. A., Brilliant, K., & Lang, C. (1995). *National bias crimes training for law enforcement and victim assistance professionals*. Newton, MA: Education Development Center, Inc.

Mays, V. M., Cochran, S. D., & Rhue, S. (1993). The impact of perceived discrimination on the intimate relationships of black lesbians. *Journal of Homosexuality, 24*(4), 1–13.

Merl, J. (1989, Oct. 26). Survey of Schools Finds Hate Crime Widespread. *Los Angeles Times,* A1, S43.

Miethe, T. D., & Lee, G. R. (1984). Fear of crime among older people: a reassessment of the predictive power of crime-related factors. *The Sociological Quarterly, 25,* 397–415.

Nebraska Advisory Committee to the United States Commission on Civil Rights (1990). *Bigotry and violence on Nebraska's college campuses.* Washington, DC: US Government Printing Office.

Parker, K. D., McMorris, B. J., Smith E., & Murty, K. S. (1993). Fear of crime and the likelihood of victimization: A bi-ethnic comparison. *Journal of Social Psychology, 133,* 723–732.

Roberts, J. (1995). *Disporportionate harm: Hate crime in Canada.* Http//www1.denizkor.org/nizkor/hweb/orgs/canadian/canada/justice/disproportionate-harm/.

Sales, E., Baum, M., & Shore, B. (1984). Victim readjustment following assault. *Journal of Social Issues, 40,* 117–136.

Schneider, H. J. (1994). Violence against foreigners in the Federal Republic of Germany. *International Journal of Offender Therapy and Comparative Criminology, 38,* 179–182.

Schultz, L. G. (1987). Victimization programs and victims of crime. In A. Minahan (Ed.), *Encyclopedia of social work* (18th ed., Vol. 2, pp. 817–822). Silver Spring, MD: National Association of Social Workers.

Snell, W. E., & Godwin, G. (1993). Social reactions to depictions of casual and steady acquaintance rape: the impact of AIDS exposure and stereotypic beliefs about women. *Sex Roles, 29,* 599–616.

United States Commission on Civil Rights (1987). *Recent activities against citizens and residents of Asian descent.* Washington, DC: US Government Printing Office.

United States Commission on Civil Rights (1990a). *Bigotry and violence on American college campuses*. Washington, DC: US Goverment Printing Office.

United States Commission on Civil Rights (1990b). *Intimidation and violence: Racial and religious bigotry in America*. Washington, DC: US Government Printing Office.

United States Congress (1992). H.R. 4797. 102d Congress, 2d Session.

United States Department of Justice, Federal Bureau of Investigation (1997a). *Uniform crime reports: Hate crime, 1995.* Washington, DC: United States Government Printing Office.

United States Department of Justice, Federal Bureau of Investigation (1997b). *Crime in the United States, 1996: Uniform crime reports for the United States*. Washington, DC: United States Government Printing Office.

von Schulthess, B. (1992). Anti-lesbian assault and harassment in San Francisco. In G. M. Herek & K. T. Berrill (Eds.), *Hate crimes: Confronting violence against lesbians and gay men* (pp. 65–75). Newbury Park, CA: Sage.

Waxman, B. F. (1991). Hatred: The unacknowledged dimension in violence against disabled people. *Sexuality and Disability, 9,* 185–199.

Weine, S., Becker, D. F., McGlashan, T. H., Vojvoda, D., Hartman, S., & Robbins, J. P. (1995). Adolescent survivors of ethnic cleansing: Observations on the first year in America. *The Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 1153–1159.

Weiss, J. C. (1990). Ethnoviolence: Violence motivated by bigotry. In L. Ginsberg (Ed.), *Encyclopedia of social work* (18th ed., 1990 Suppl., pp. 307–319). Silver Spring, MD: National Association of Social Workers.

Weiss, J. C., Ehrlich, H. J. & Larcom, B. E. K. (1991). Ethnoviolence at work. *The Journal of Intergroup Relations, 18,* 21–33

Wexler, C., & Marx, G. T. (1986). When law and order works: Boston's innovative approach to the problem of racial violence. *Crime and Delinquency, 32,* 205–223.

Wirtz, P. W., & Harrell, A. V. (1987). Victim and crime characteristics, coping responses, and short- and long-term recovery from victimization. *Journal of Consulting and Clinical Psychology, 55,* 866–871.

Young, M. A. (1988). Support services for victims. In F. M. Ochberg (Ed.), *Post-traumatic therapy and victims of violence* (pp. 330–351). New York: Brunner/Mazel.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.23
Victims of Disasters

NIGEL LONG, KEVIN RONAN, and JOYCE PEREIRA-LAIRD
Massey University, Palmerston North, New Zealand

9.23.1 INTRODUCTION 375
9.23.2 NATURE OF DISASTER 376
9.23.2.1 Definition of Disasters 376
9.23.2.2 Classification of Disasters 377
9.23.2.3 Definition of Victims 378
9.23.3 METHODOLOGICAL AND ASSESSMENT ISSUES 378
9.23.3.1 Sampling Issues and Design 379
9.23.3.2 Instruments of Assessment 379
9.23.3.3 Timing of Assessments 380
9.23.3.4 Assessment Variables 381
9.23.4 PREVALENCE AND MENTAL HEALTH CONSEQUENCES OF DISASTERS 381
9.23.4.1 Prevalence of Disasters and Other Traumatic Events 381
9.23.4.2 Prevalence of Psychological Symptoms and Disorders 382
9.23.5 TREATMENT APPROACHES 383
9.23.5.1 Multiple Gating and Methodology 384
9.23.5.2 Supportive Therapeutic Environment and Normative Information 384
9.23.5.3 Critical Elements Across Therapy Approaches 385
9.23.5.3.1 Exposure and related techniques 385
9.23.5.3.2 Cognitive intervention components 386
9.23.5.4 Research on Efficacy and Integrated Treatment Description 386
9.23.6 CONCLUSIONS 387
9.23.7 REFERENCES 388

9.23.1 INTRODUCTION

Humans have been victims of disasters throughout recorded history. Tornadoes, fires, floods, earthquakes, airplane crashes (and other transportation accidents), hazardous material spills, building collapses, nuclear reactor leaks, terrorist bombings, hostage taking, ice storms, blizzards, and many other accidental, natural, or technological disasters occur throughout the world each year. Disasters, therefore, present one of the most significant concerns among people today and they are a predominant contributing factor to the development of health and mental health problems for millions of people.

A number of studies have investigated factors that contribute to the long-term mental heath outcomes following disasters. This chapter reviews the findings from a diverse group of studies and suggests that the absence of unifying themes in disaster research is associated with difficulties in delineating the salient dimensions of the event and the victims. The chapter then discusses the various definitions and classifications of disasters and victims and shows how the wide variation in methodology and measurement and assessment instruments can impact research findings. These variations also make estimations of the prevalence levels of mental health very difficult to determine and some of these issues are also reflected in the consequent treatment approaches. Many of these are not comprehensive or well controlled and the effectiveness of any treatment approach is also difficult to evaluate. Finally, it is suggested that a multiple gate methodology coupled with a cognitive behavioral approach has the best potential for dealing with the wide range of contingencies which disasters present.

9.23.2 NATURE OF DISASTER

Research into the effects of disaster on humans has advanced since the mid-1980s, however, there are inconsistencies in the research findings reporting the psychological consequences of disaster on victims (Paton & Long, 1996). This is because most individuals when exposed to traumatic events react with a wide range of physiological and psychological responses which are often transient and dissipate over time. While the number and intensity of symptoms may vary over time, some disaster victims remain more vulnerable to future stressors and may be prone to reactivation of psychological disorders (Long, Chamberlain, & Vincent, 1994).

Other inconsistencies in the research may be due to (i) the plethora of definitions used to operationalize the key concepts of disaster, victims, and psychological consequences (Baum, 1991; Warheit, 1988); (ii) diverse research orientations (e.g., psychologists and psychiatrists focus on the psychic components of trauma and mental health of the victims, while sociologists tend to focus on macro-level variables such as general symptoms of distress and "problems of living") (Bolin, 1988; Green, 1991; Warheit, 1988); (iii) confusion between agent-generated (disaster) and response-generated (organizational response efforts) consequences (Warheit, 1988); and (iv) the absence of an integrated theoretical framework among different research groups (Hartsough, 1985; Warheit, 1988). In addition, some of the diversity in disaster research can be attributed to researchers failing to systematically investigate fundamental issues over an extended period of time.

9.23.2.1 Definition of Disasters

Although many people share a common understanding of what is meant by the term disaster, operationally defining disaster for research purposes has proved difficult (Quarantelli, 1985). It is reasonably clear that the term disaster implies a diverse set of circumstances involving physical destruction and psychological distress, property loss, serious injury, and death. Other more specific definitions have been advanced but many of these are imprecise as they involve qualititative statements. Problems of definition are further exacerbated by the inconsistent use of terminology (e.g., catastrophe, calamity, crisis, trauma, cataclysm, stress, and emergency). Clearly, the field lacks consensus on the use of terminology, a definition of disaster, as well as a definite typology of disaster characteristics and consequences.

In earlier definitions, the focus was either on the physical agents of the disaster or on its impact in terms of the physical or social damage it caused. Many of these had a singular focus and did not consider interaction effects (Harshbarger, 1974). For example, by considering the effect of impact alone, there is the danger of encompassing so many situations that the concept of disaster is rendered meaningless.

Although the nature of the event must be considered in any formulation of a definition, it should not be exclusively focused on as there are a set of unique characteristics which are common to all disasters. Damage caused by any sudden, powerful event that is beyond the realm of everyday experience could be considered a disaster, but this narrow definition presents a problem as it could encompass individual stressors which are not what is generally considered as a disaster (e.g., assault). A number of researchers have suggested that the concept of magnitude or scope must be added to the definition so that in order for an event to be a disaster, a substantial number of people must be involved (Baum, 1991; Quarantelli, 1985).

An alternative approach in formulating a definition is to view a disaster as a specific cataclysmic event, that is, a stressor depicted by immense power, large scope, suddenness, and placing excessive demands on individual coping (Lazarus & Cohen, 1978). Such a characterization of disaster has emerged as a unifying theme in discussions of disaster (Baum & Davidson, 1985; Bolin, 1988). In conceptualizing disasters as stressful events, one can use theory and research on stress to understand and predict the effects of disasters.

A review of definitions of disaster has revealed that researchers have tended to focus on one or more of the following factors: (i) the specific events or physical agents of disasters themselves; (ii) the peculiar nature of the event, especially its suddenness, power, and enormity; (iii) the potential of the event to disrupt many factors in the everyday life of an individual; (iv) the way in which victims respond; (v) the loss of daily life structure; (vi) the large degree of destruction and human suffering; (vii) the collective nature of the stress situation; and (viii) the degree of social disruption on group and community functioning.

Baum's (1991) conceptualization of disaster appears to encompass many of these factors: "events or impacts that cause substantial disruption of psychological or physical functioning. They may be thought of as stressors, tending to be more intense, more powerful, and universal than most threatening or demanding events. They affect relatively large numbers of people and pose heavy demands on coping. In addition, they threaten property and life, whether directly or by virtue of long-term threat" (p. 44). This definition is fairly comprehensive and has an underlying theoretical framework which is useful in conceptualizing disaster research.

9.23.2.2 Classification of Disasters

Accurately classifying disasters has several advantages as it allows comparison between different types of events, enables one to assess the differential psychological effects, and helps with treatment planning (Smith, North, & Price, 1988). In an effort to identify the critical elements of disaster, a variety of taxonomies have been proposed (Barton, 1969; Berren, Santiago, Beigel, & Timmons, 1989). Examination of these classification systems suggests that there are a variety of possible dimensions that might assist in predicting which disasters are likely to have the most impact on mental health. However, these approaches are not consistently used and researchers still focus mainly on the differences between human-made and natural disasters. Although this distinction may represent a parsimonious classification scheme, it may be limited to answering the question of who or what is responsible for the disaster. On the other hand, the natural vs. human-made distinction may be a potentially important factor in determining mental health effects (Baum, Fleming, & Davidson, 1983; Green, Lindy, & Grace, 1994). Natural disasters, often referred to as acts of God, can eliminate human agency and culpability from consideration as they may be assumed to be uncontrollable. Conversely, attributions of responsibility following human-made disasters (acts of man) often implicate human negligence, fallibility, or deliberation, and offer victims a target for blame.

Baum (Baum 1991; Baum et al., 1983) has made attempts to delineate specific characteristics which distinguish natural from human-made disasters. These characteristics include suddenness of onset, predictability of the event, extent of damage, persistence of effects, perceptions of control, and absence of an identifiable low point. The notion of a low point denotes that from the victim's perspective, the "worst is over" and this assumption now allows attention to be turned to the recovery process. Researchers have claimed that natural disasters have an identifiable low point (Baum, 1991; Baum et al., 1983; Bolin, 1988). In contrast, some human-made disasters, especially those which occur over a protracted period (nuclear accidents, toxic and chemical spills) have an indistinguishable low point with hardly any noticeable or short-term negative physical effects on victims.

Some researchers have also attempted to provide specific classification schemes within the natural or human-made category. A number of differences between types of natural disasters have been identified using a variety of factors (Alexander, 1993; Belter & Shannon, 1993). It is important to note that although most researchers use the terms human-made and technological disasters synonymously, others have made attempts to delineate the difference between technological related accidents involving chemical and toxic spills and nuclear accidents from other human-made disasters such as those involving structural failures (bridge or building collapses) or transport accidents (Hodgkinson & Stewart, 1991; Taylor, 1996). Beigel and Berren (1985) classified human-made disaster acts as acts of omission (negligence) such as building collapses due to poor infrastructure, and intentional acts of commission, such as terrorism or hostage taking.

It is clear that although there may be important differences between natural and human-made disasters, there are also significant differences between disasters within each of these classes. Although disasters within the same class may have many elements in common, no two disasters present the same logistical and psychological challenges to their victims. For example, although hurricane Hugo in 1989 and hurricane Andrew in 1992 were comparable in terms of the damage and loss produced, they were distinct in that they varied on cultural, socioeconomic, geographical, political, and even meteorological factors. As a result, some researchers emphasize the importance to approach each disaster environment so as to acknowledge its uniqueness (Saylor, 1993).

One shortcoming in the various definitions of disasters is that there are no clear boundaries between natural and human-made disasters (Beigel & Berren, 1985; Green & Solomon, 1995). In many disasters, technology and/or human failure and natural forces all combine to create the disaster. For example, while most people would tend to classify the 1985 Mexico City earthquake as a natural disaster, aspects of it might be regarded as having human-made characteristics, such as the ineffective anti-earthquake structural architecture which led to the collapse of a major hospital where so many died (Smith et al., 1988). This highlights the difficulty in examining the differential effects of disasters because disasters frequently contain mixed typological elements.

An important aspect of disaster research is delineating the salient dimensions of the event which the subjects are responding to. A number of researchers have emphasized the importance of understanding and identifying these dimensions in order to predict the scope and magnitude of psychological and psychosocial dysfunction resulting from the experience of a disaster (Bromet & Dew, 1995). Earlier discussion suggests that the distinction between natural and human-made disasters is often artificial and may not be adequate in predicting disaster consequences. In addition, it is difficult to draw valid conclusions about disasters because they vary on a large number of other crucial dimensions (e.g., scope, severity) that could affect outcome. Without a clear understanding of the salient dimensions that differentiate disasters and how these variables can act individually or in collaboration, it is difficult to predict the psychological and psychosocial consequences of disasters. Although some attempts have been made to develop disaster classification typologies consisting of salient dimensions to provide a conceptual framework for better understanding the concept of disaster (Berren et al., 1989), these classification schemes are rarely used in discussing the psychological effects of disasters.

9.23.2.3 Definition of Victims

At the centre of the notion of disaster is the concept of victims, the individuals who are affected by the event and its consequences. It is not surprising to find a diversity of opinions on how best to conceptualize a disaster victim given the wide disagreement as to what constitutes a disaster. Early efforts at classification categorized victims according to the extent to which they suffered bereavement, injury or sickness, and loss of property; or to the methods used for coping; or to their experiences of social disruption (e.g., Barton, 1969). Researchers have extended their definition of victims to include emergency personnel and also to those victims who were not directly involved but who may have experienced adverse reactions from the disaster (Hodgkinson & Stewart, 1991).

A number of models have been developed to define the different types and degrees of victimization following disasters. For example, Taylor (1996) proposed a six-fold classification of victims which defined victims as any people who suffered an adverse and severe reaction to the disaster, irrespective of the level of exposure. However, the simplest definition differentiates primary victims (directly exposed) from secondary victims (indirectly affected) (Bolin, 1985).

Raphael (1986) suggested that the specific dynamics and outcomes of a disaster event depend on a number of factors. These include the disaster itself, the individual victims, the role expectations for that specific victim situation, and the circumstances of the victimization. Although there is literature suggesting that disaster helpers also experience psychological trauma (Taylor, 1996), due to space limitations, this chapter is limited to primary and secondary victims.

9.23.3 METHODOLOGICAL AND ASSESSMENT ISSUES

Accurate assessment of disasters is important given the range of possible psychological, physiological, and psychosocial outcomes which occur as a result of the disaster. The emphasis and nature of assessment depends on the phase of the disaster (e.g., impact vs. recovery) and on its objective regardless of whether it is prevention, treatment, or a compensation claim. McFarlane (1995) suggested that assessment of the disaster impact is crucial given the variability in response and trauma of individuals and communities involved.

Three broad approaches have been identified in the field of disaster research (Baum, 1991; Solomon, 1989). These include: (i) clinical-descriptive studies which document the symptomatology of disaster victims; (ii) epidemiological studies which measure levels of impairment in population groups and communities following a disaster; and (iii) quasi-experimental studies which focus on psychological variables as mediators of response and subsequent effects of disasters. For an excellent review of the goals and the advantages and disadvantages of these three theoretical approaches, the reader is referred to Solomon (1989).

These competing theoretical approaches are accompanied by divergences in the methodologies used. The differences include variation in sample size and sampling procedures, research designs, timing of assessments, and choice of measurement instruments. It is difficult to draw valid conclusions regarding the extent of impairment resulting from different kinds of disaster experience as the methodological diversity across studies hinders cross-study comparisons, replication, and generalization

(Baum et al., 1993; Green, 1991; Smith et al., 1988).

The study of disasters is further confounded by the nature of the phenomenon under investigation. For instance, predisaster data are rarely available because of the unpredictable and infrequent nature of many types of disasters and this hinders accurate or systematic assessment of preimpact behavior and health (Baum, 1991). Research has rarely obtained pre- and postdisaster data prospectively (Baum, 1991; Baum et al., 1993; Bromet & Dew, 1995; Smith et al., 1988). Most studies that have examined predisaster information have relied upon retrospective reports, and these may present problems regarding reliability. Other useful predisaster information such as prior psychiatric records may also be unreliable due to inexact diagnostic criteria and incompleteness, or may have been destroyed in the disaster (Baum, Schaeffer, Lake, Fleming, & Collins, 1986).

In addition, "the population under study as well as the phenomena under study are widely dispersed and must be sought out rather than screened at a treatment facility" (Baum et al., 1993, p. 126), and this makes access to disaster victims one of the greatest practical problems (Baum et al., 1993; Solomon, 1989). For instance, the disruptive events of a disaster might cause substantial dislocation of the victims. Also, potential subjects may be reluctant to recall their painful experiences or they may be involved in pending legal litigation. Conducting blind studies with regard to traumatized and untraumatized subjects is often impossible, given that the comparison group of untraumatized individuals are unlikely to live in the vicinity of the disaster area (Baum et al., 1993). For a more complete examination of methodological and conceptual issues concerning disaster assessment, the reader is referred to several reviews (Baum et al., 1993; Bromet & Dew, 1995; Green, 1991).

In the next section, a brief examination of methodological issues concerning sampling, design, timing of assessments, instruments of assessment, and what to assess follows. The major purpose in delineating these conceptual and methodological problems is not to denigrate existing research efforts, but rather to raise these issues in the hope that the reader will consider the adequacy of research studies when considering the mental health impact of disasters.

9.23.3.1 Sampling Issues and Design

Many disaster studies in the last decade have utilized a modified cohort design and most have been cross-sectional retrospective studies (Bromet & Dew, 1995). Although longitudinal studies have also been conducted, they varied in their timing of initial and follow-up assessments and this makes any cross-study comparisons difficult.

One important methodological aspect of disaster research, which affects psychological impairment estimates, concerns the group of victims who are investigated, as it is often difficult to differentiate between a disaster-affected population and an unaffected one. This difficulty arises because a variety of victims may result from the same disaster and the main differentiating feature is their exposure levels to the disaster experience (e.g., injured victims and their close relatives, rescue workers, and body handlers). Most disaster research selects individuals that fall into the primary and secondary victim classes reported earlier (Baum et al., 1993).

Although traditional descriptive studies of disaster generally have not employed a control group (e.g., Lifton & Olson, 1976), others have attempted this (e.g., Dew & Bromet, 1993) so that credible conclusions can be reached concerning disaster-associated differences in individual psychological and psychosocial outcomes. Due to the difficulty of obtaining an appropriate control group, researchers have often included less affected individuals rather than unaffected individuals in their comparison group (Baum et al., 1993). This presents a problem as research has indicated that individuals indirectly exposed to disasters also experience adverse psychological experiences (Melick, 1985).

Another important aspect of disaster research concerns sample size. Several reviews (e.g., Bromet & Dew, 1995; Hartsough, 1985; Smith & North, 1993; Smith et al., 1988; Wilson & Raphael, 1993) suggest that sample size varies markedly across disaster studies ($N = 17\,160$, Ahearn, 1981; $N = 7$, Henderson & Bostock, 1977). In a review, Bromet and Dew (1995) suggested that the majority of sample sizes in studies in the 1990s continue to be small and of questionable representativeness.

9.23.3.2 Instruments of Assessment

A review of the existing literature on disasters reveals the use of a wide range of instruments to measure the effects of disasters. Several studies have utilized home-made scales, open and nonstructured interviews, structured diagnostic interviews, retrospective medical record reviews, telephone surveys, standardized diagnostic instruments, and behavioral, self-report, and biologically based measures (Baum et al., 1993; Hartsough, 1985; Smith et al., 1988). The failure to use comparable instruments across studies has hindered cross-study comparisons and has made it difficult to draw valid conclusions from the available literature.

Fortunately, the field of disaster is converging with regard to instrumentation for assessing psychological symptomatology. Since the mid-1980s, the use of home-made instruments designed as one-off assessment tools for a particular study has been on the decline. A number of researchers have stressed the importance of using standardized instruments (Baum et al., 1993; Green, 1991; Hartsough, 1985) to facilitate generalization and to improve accuracy of findings. For instance, some researchers have suggested that higher rates of psychological impairment are found in studies employing self-report or in-depth, open format interviews than those using standardized instruments (Melick, 1985; Smith et al., 1988).

Several disaster investigators have listed and/or described a number of measures that have been used in disaster studies with the aim of recommending ppropriate instrumentation and in the hope of encouraging consistency of measures across studies to facilitate replication of results (e.g., Baum et al., 1993; Bolin, 1988; Hartsough, 1985; Smith et al., 1988). The instruments recommended include diagnostic instruments such as the Diagnostic Interview Schedule/Disaster Supplement (DIS/DS; Robins & Smith, 1983) and the Structured Clinical Interview for *DSM-IV* (SCID; First, Spitzer, Gibbons, & Williams, 1994), standardized symptom checklists such as the Symptom Checklist-90-R (SCL-90-R; Derogatis, 1983), or standardized self-report scales such as the Beck Depression Inventory (BDI; Beck & Steer, 1984) and the Impact of Event Scale (IES; Horowitz, Wilner, & Alvarez, 1979).

However, the application of standardized instruments is not without problems. For instance, it has been shown that the correlation between standardized instruments measuring related constructs within a single study is low, suggesting that the same instrument might have to be used on both populations for cross-study comparisons. In addition, the specificity and sensitivity of standardized instruments may be unacceptably low even within the same study population. These problems can be addressed by using multiple instruments and multimethod approaches for obtaining data.

In their review, Baum and co-workers (1993) stated that although most researchers used measures of psychological functioning that were psychiatric symptom-oriented, some researchers have attempted to measure family, social, and vocational functioning as well as biological correlates of exposure. Since the mid-1980s, there has been much focus on post-traumatic stress disorder (PTSD) measurement in disaster studies. However, one problem with the use of diagnostic PTSD instruments is that the diagnostic criteria for PTSD are frequently changing (Green, 1991), and research findings of cross-study comparisons may be unreliable or invalid.

Although disaster research and assessment of functioning has often incorporated either self-report or clinical assessment of psychological symptoms (e.g., symptom experience, ratings of distress, depression), there has been a trend toward complementing these subjective indices with more objective physiological (e.g., sympathetic arousal), biochemical (e.g., pituitary–adrenal axis arousal—cortisol), and behavioral measures (coping, motivation). Such multimethod assessment approaches have been advocated by a number of researchers (e.g., Baum et al., 1993; Green, 1991; Smith et al., 1988). It has been argued that the concurrent assessment of self-report, performance, and biologically based responses constitutes an optimal measurement strategy that would enhance our knowledge and understanding about the sequelae of disasters (Baum, Grunberg, & Singer, 1982). The utility of such an approach has been demonstrated by a number of researchers (Baum, 1990; Davidson & Baum, 1986).

9.23.3.3 Timing of Assessments

No agreement on the most appropriate time for assessment has been established, and studies vary widely on the timing of data collection (Bromet & Dew, 1995). Reviews (Bromet & Dew, 1995; Green, 1982; Smith & North, 1993; Smith et al., 1988) show that timing of assessments have varied from immediately after the disaster (Friedman & Linn, 1957) to up to five years later (Gleser, Green, & Winget, 1981), with the majority of assessments occurring within one year.

Early, long-term, and multiple measurement is essential for the understanding of the course of responses to disaster stress (Baum et al., 1993; Smith et al., 1988; Solomon, 1989). However, the timing of assessment is often dependent on practical (e.g., funding) rather than scientific considerations (Green, 1982). In addition, measuring victim reactions in the immediate aftermath of a disaster is also ethically questionable and the accuracy of victim's answers may be in question since individuals in acute distress are unlikely to provide complete and accurate responses (Solomon, 1989).Research has indicated that both the nature and intensity of psychological sequelae appear to vary as a function of the time at which they are measured. For instance, the prevalence of anxiety and depression varied fourfold from month to month following the Three Mile Island accident, with the largest peak occurring immediately after the event (Bromet, Parkinson, Schulberg, Dunn, & Gondek, 1980). Overall, there have been few long-term studies and even fewer systematic studies providing assessments of psychological sequelae at multiple points in time.

Multiple-method assessments are important as the different methods tap various aspects of disaster effects. In addition, we suggest these different methods should be used in conjunction with multiple assessment periods which employ a multiple gate strategy, particularly with treatment-related assessment because assessment is part of early intervention efforts. The multiple gate approach involves the use of a less accurate and less extensive screening test (e.g., a self-report inventory) to identify people at risk who could be followed up later for more accurate diagnostic testing. The initial use of self-report questionnaires is helpful in screening populations and identifying those who are at risk for mental health problems and those who may benefit from intervention. Such an approach may be necessary in situations where intervention cannot be provided to all who are exposed due to lack of resources or practical problems, and where it may be premature to use time-consuming diagnostic tests. More comprehensive assessments can be used at the later gates if increasing psychopathology is demonstrated.

9.23.3.4 Assessment Variables

Research attempts involved in identifying mechanisms, predictors, and mediators of response to trauma in disasters are likely to yield more informative and useful findings than research which is purely descriptive. Researchers suggest that to achieve a full understanding of the impact of disaster on subsequent adjustment it is important to measure both the generic and situationally specific features of disaster events (Baum et al., 1993; Bolin, 1988), as well as the anticipatory, acute, and long-term responses to the disaster (Ursano, Fullerton, & McCaughey, 1994).

Ursano et al. (1994), who used a stress model for studying trauma and disasters, suggested that evaluation of psychological, physiological, and sociocultural responses to a disaster is important for the accurate assessment of mental health outcomes. A detailed assessment of the nature and severity of the stressors to which an individual was exposed is necessary, in addition to personal factors (e.g., demographic factors, predisaster psychological functioning, coping style, and appraisal) that might affect post-trauma processing of the event. It is also important to determine whether the victim has suffered prior trauma, because multiple traumatic events may put them at higher risk for developing PTSD and other related symptomatology (Kramer & Green, 1991). Assessment of environment-recovery factors (e.g., social support, extent of network disruption caused by the disaster, societal attitudes about the disaster event) that may inhibit or enhance recovery of the victims is also essential (Ursano et al., 1994).

Assessment represents the first step in any intervention effort. In addition, the assertion made in this chapter and supported by data is that continuing use of assessment-related techniques throughout a treatment application will often translate into more beneficial outcomes. Of course, treatment-related assessment should include multiple methods and multiple gating (see Section 9.23.5).

9.23.4 PREVALENCE AND MENTAL HEALTH CONSEQUENCES OF DISASTERS

The psychological response following exposure to trauma or disaster is a complex interaction of the event and the individual (Paton & Long, 1996). The interaction may be modified by an individual's previous history, his or her physiological response, and the social context of the event (Ursano et al., 1994), and these can alter significantly over life span (Alexander, 1993) and time of assessment (Davidson & Foa, 1991). Many of the difficulties concerning problems with definitions of terms, methodology, and assessment are ultimately reflected in wide variations in the prevalence levels reported. The next section reviews some of the rates reported and suggests that caution should be exercised when extrapolating prevalence estimates from affected populations.

9.23.4.1 Prevalence of Disasters and Other Traumatic Events

The occurrence of disasters and traumatic events may vary across geographic location, in terms of severity and risk, and in their physical attributes (Seaman, 1984). Internationally, there is growing evidence which suggests that the prevalence of death and damage following natural disasters is increasing and that 95% of all deaths occur in countries of the third world (Alexander, 1993). Major floods and hurricanes are the most frequently reported disasters, followed by earthquakes, tornados, and severe storms and the mortality rate of these is also rising (Ursano et al., 1994). Other low-frequency natural catastrophes, such as famines, have been shown to cause millions of deaths over a short period of time (Hodgkinson & Stewart, 1991; Seaman, 1984). Similar to many of the third world disasters, catastrophe can also have severe impact in developed countries, (e.g., the Kobe earthquake), and even though the total number of people affected has not been relatively large, the psychological disorders have been extensively researched (Breslau, Davis, Andreski, & Peterson, 1991; Bromet & Dew, 1995; Norris, 1992). In all situations, the prevalence of psychological disorders has varied across the variety of natural diasters regardless of whether life-time or current rates were reported (Breslau et al., 1991; Norris, 1992).

Community surveys reporting life-time prevalence levels of exposure to at least one major natural disaster have ranged between 1% and 28% (Breslau et al., 1991; Norris, 1992; Vrana & Lauterbach, 1994) and many people had been exposed to a wide range and number of other disasters or traumatic events (Breslau et al.). Within these groups the prevalence levels for exposure to *any* traumatic event was between 39% and 82% (Breslau et al.; Norris, 1992; Vrana & Lauterbach, 1994). Higher prevalence levels (e.g., 92%) have also been reported for clinical samples (Barnard, Hankins, & Robbins, 1992).

Prevalence levels of technological disorders have not been widely reported. However, it has been suggested that chemical pollution, transportation accidents, explosions, structural failures, terrorist acts, and other major threats to the environment do represent increasing concern to physical and mental health (Green & Solomon, 1995).

9.23.4.2 Prevalence of Psychological Symptoms and Disorders

Prior to *DSM-III* (American Psychiatric Association, 1980), research into prevalence levels centered mainly on evaluating the nature and extent of the psychological response following exposure to a wide range of disasters. The wide variability in prevalence levels reported across studies results from a variety of factors, including severity and nature of the disaster, and differences in sampling and methodology (Smith & North, 1993; Ursano et al., 1994).

Earlier studies focused on evaluating the prevalence of a wide range of symptoms and disorders following exposure to one trauma within psychiatric populations or community samples. Although initially many of these studies concentrated on PTSD (Green, 1993) and PTSD symptoms (Norris, 1992; Smith & North, 1993), studies in the 1990s have evaluated other psychological disorders including acute stress disorder (Koopman, Classen, Cardena, & Speigel, 1995). Other studies have evaluated many disorders over varying periods of time (Bland, O'Leary, Farinardo, Jossa, & Trevisan, 1996) and some have studied the impact on children and families (Belter & Shannon, 1993). Rubonis and Bickman's (1991) review of disasters which had a sudden onset, showed that disasters, on average, increased the prevalence rate of consequent psychopathology by 17%. Higher levels of disaster psychopathology was also predicted by clinical interviews, being female, and immediacy of data collection. In many of the studies reviewed, the magnitude of the disaster was correlated with symptom severity and some of these can be considered generic while others are disaster specific.

Results from clinical and community studies have shown that many victims of disasters meet the criteria for PTSD as defined in *DSM-IV* (American Psychiatric Association, 1994) and other victims may have many of the symptoms of PTSD but fail to meet the full diagnostic criteria (Breslau et al., 1991; Bromet & Dew, 1995; Green, 1993; Norris, 1992). Some studies have shown that in addition to PTSD, victims may also meet the full criteria for other co-occurring *DSM-IV* disorders such as generalized anxiety disorder, obsessive-compulsive disorder, dysthymia, depression, and phobic disorders (Breslau et al.). Other outcomes have also been investigated and these included other anxiety or somatic disorders, substance abuse, poor health, treatment seeking, and other psychosocial variables (Bland et al., 1996).

Studies of disasters have reported prevalence rates for both physiological and psychological symptoms, clusters of symptoms, precisely defined psychological disorders, and anecdotal syndromes (e.g., disaster syndrome, and survivor guilt) and many have reported combinations of these (Seaman, 1984). Accurate estimations of prevalence levels are therefore often confounded by confusion between comorbidity (separate disorders) and consanguinity (both disorders one and the same) (Tyrer, 1996).

In addition, estimating psychiatric morbidity following disasters is not the same as estimating the number of cases of particular psychiatric disorders. This is highlighted when the number of symptoms which only partially meet the full set of diagnostic criteria for a specific disorder also contribute to the level of another disorder and this is extremely common especially when generalized anxiety and major depression are present (Long, Chamberlain, & MacDonald, 1996).

In evaluating the prevalence levels of psychiatric disorders, researchers have used a wide range of standard health assessment tools to assess a variety of clinical and subclinical mental health effects of disasters. Few studies were designed with representative samples, comparison groups, systematic and reliable outcome measures, and/or multivariate analysis of exposure effects that took into account other known risk factors, moderators, or mediators of the disorders under study (Bromet & Dew, 1995). These difficulties make comparisons of psychological outcomes among and between disasters particularly problematic and many of these issues have been reviewed (Bromet & Dew, 1995; Green, 1994). In an attempt to overcome some of these difficulties, one meta-analytic study has summarized the impact of different stressors across many studies (Rubonis & Bickman, 1991).

Disaster victims seeking treatment in clinical settings also inflate prevalence levels because they often present a major diagnostic problem because chronic PTSD nearly always occurs in conjunction with a range of psychiatric disorders and especially with depression and anxiety (Davidson & Foa, 1991). In addition, other research in clinical settings has shown that comorbidity increases the likelihood of treatment utilization and reported symptoms (Galbaud du Fort, Newman, & Bland, 1992).

It is therefore important to evaluate prevalence levels in populations not seeking treatment because levels are less likely to be inflated by other comorbid disorders. Comparisons of the psychological impact resulting from natural and human-made disasters have been made (Green & Solomon, 1995). However, these need to be interpreted cautiously since they do not involve equivalent comparisons. It has been argued that human-made disasters (e.g., chronic technological-related events involving toxic and radiation contamination) are inherently more stressful than natural disasters (Baum et al., 1983; Solomon & Smith, 1994; Weisaeth, 1994), because of their unpredictability, uncontrollability, culpability, and their lack of a low point. In their comparison of disaster studies, Green and Solomon (1995) found that victims of acts of commission disasters experienced higher disorder rates than victims of high-death-rate natural disasters.

The implication of these comparisons is that natural disasters may be more readily accepted than technological failures. On the other hand, human-made disasters especially those deliberately perpetuated may have an even more significant impact on ensuing psychopathology and would require long-term medical and psychotherapeutic intervention.

9.23.5 TREATMENT APPROACHES

Treatment approaches assume a familiarity with the factors concerning definitions, assessment, and prevalence levels. Knowledge of data-based literature addressing normative effects of disasters on individuals and communities is an important step before any intervention process. Familiarity with the nature of the traumatic experience, developmental course of the disorder, personal background factors, cognitive processing, coping style and ability, and the relationship among these interpersonal factors, and the social milieu are important variables when the treatment of any individual victim is considered (Wilson & Raphael, 1993). This knowledge will allow the therapist to apply treatments in flexible ways so as to meet the individual needs of the client.

While a considerable data base is available concerning individual treatment factors, it is important to note that methodologically treatment outcome-based research is the exception rather than the rule. Treatments for victims of disasters range from those focused on predisaster readiness and primary prevention to a myriad of treatment approaches focused on postdisaster adaptation and coping. In this latter cluster, approaches range from community-based prevention in the immediate aftermath to longer-term group and individual approaches for those who are not adapting effectively. In addition, individual and group treatment approaches vary tremendously. For example, one comprehensive text included treatment approaches with an emphasis on a wide range of factors: psychobiological, neurocognitive, psychoanalytic, behavioral, holistic, group, family, peer-mediational, and other approaches (Wilson & Raphael, 1993). For the majority of these treatments, the results are rarely reported from controlled clinical trials. Discussion of treatment first focuses on issues important to the development of both practice and research. This includes multiple gating and methodology followed by a focus on elements common to a variety of approaches. Following this discussion, the focus moves to a review of controlled evaluation research assessing the effectiveness of cognitive-behavioral approaches in helping those affected by traumatic events.

9.23.5.1 Multiple Gating and Methodology

The issue of multiple gating raised earlier is an important consideration within treatment applications. Treatment provided to victims of disasters should also follow a gating methodology. The first step is primary prevention. Predisaster readiness, when possible, has been shown to be effective in helping people cope more effectively when a disaster does strike (Wilson & Raphael, 1993). While some disasters can be anticipated (e.g., high-risk areas), others happen without warning. In these latter cases early intervention is vital (Vernberg & Vogel, 1993).

Early intervention can help victims reestablish a sense of control which is instrumental in preventing future problems (e.g., PTSD, major depression). In addition, with some planning these interventions can be implemented to maximize coverage of victims. For example, school-based and community-based interventions (e.g., postincident debriefing) can provide rapid dissemination of information, coping skills education, a chance for the normalizing of reactions to the disaster, and a forum for the activation of social support networks (Hodgkinson & Stewart, 1991). For example, a brief school-based cognitive-behavioral large group intervention was shown to reduce symptoms of distress and increase perceived coping ability in 7- to 14-year-old children following a series of volcanic eruptions (Ronan & Johnston, 1996). These children continued to show reduced distress at follow-up seven months after the initial disaster. For those who do not benefit from initial efforts, however, the next intervention gate would focus on secondary prevention (i.e., intervention at early signs of disorder). At this stage, group or individual interventions can provide additional help that focuses on premorbid and other factors impinging on postdisaster coping and resolution.

Finally, if required, additional intervention would focus on helping reduce symptoms of disorder at the same time as attending to increasing adaptive functioning. Further research that utilizes such a gating methodology is needed to assess effectiveness of a continuum of services following a disaster. Of course, appropriate focus on other aspects of treatment outcome methodology is also warranted in this area (Rothbaum & Foa, 1993).

9.23.5.2 Supportive Therapeutic Environment and Normative Information

While support is a given in psychotherapeutic approaches, some attention to this nonspecific factor is warranted as victims of disasters need support. One element of support following disasters includes helping clients become aware of important information as soon as possible after the disaster. The nature of support given depends on the nature of the particular intervention. However imparted, information provision needs to be a part of predisaster education campaigns, crisis-related postincident debriefing (Busuttil et al., 1995), peer-mediated groups (Keane, Fairbank, Caddell, Zimering, & Bender, 1985), psychoanalytic (Brom, Kleber, & Defares, 1989), and cognitive-behavioral approaches (Foa, Hearst-Ikeda, & Perry, 1995). With the information provided, clients can begin to regain control as they begin to understand the normal effects of trauma on their psychological well-being and coping skills, and how to access assistance and multiple levels of support.

Within supportive environments, the client can learn about, and begin to cope with personal and material loss in a way that is individually appropriate. The information provided should contain up-to-date research material that includes myths associated with coping with loss. For example, despite some reports to the contrary, some beliefs continue to be held by mental health professionals concerning one normal process of coping that includes the necessity of working through the tasks of grieving. These entail accepting the reality of the loss, experiencing the pain of the loss, adjusting after the loss, withdrawing from that which is lost (e.g., loved one) and beginning to reinvest in new relationships and activities. However, at least three other normative patterns of coping have been identified (Wortman & Silver, 1989). Misinformation presented to clients mourning in ways other than that prescribed by the professional has the potential to exacerbate, rather than relieve, feelings of distress. Regarding loss, "there is no prescription for how to grieve properly ... , and no research-validated guideposts for what is normal versus deviant mourning" (Zisook & Schuchter, 1986, p. 288). As Wortman and Silver (1989, p. 355) point out, "recognition of this variability (in coping with loss) is crucial" in order that people can feel supported as they mourn in individual ways without the added concern that their grieving is in some way deviant. However, some do develop pathological grief and related conditions (major depression), and additional intervention in these cases may be warranted (e.g., cognitive or interpersonal therapy for depression).

9.23.5.3 Critical Elements Across Therapy Approaches

The variety of approaches that have been advocated for use with victims of disaster tend to use terminology common to the particular theoretical paradigm. However, common elements tend to find their way across the wide range of approaches. Within a supportive environment, treatment is aimed at helping the client cope with trauma by looking back at the traumatic event itself (Kendall & Hammen, 1995). For example, a psychodynamic approach (Horowitz, 1986) tends to begin with a focus on establishing a relationship between therapist and client while emphasizing an increasing self-awareness. This awareness specifically includes recognizing the impact of the traumatic event, increasing acceptance of its effects, and a consequent reduction in symptoms of distress. Psychodynamic techniques tend to explore the reasons underlying avoidance with the aim of helping the client release the need to avoid. These approaches have been used with some success following traumatic events (Brom et al., 1989). A cognitive-behavioral approach (Foa et al., 1995) specifically emphasizes directly breaking through avoidance by more direct confrontations with trauma-related images, feelings, and stimuli. Despite differences, both of these approaches, together with many others (Hodgkinson & Stewart, 1991; Wilson & Raphael, 1993), emphasize an increase in well-being by looking back at the traumatic event and dealing with the distressing thoughts and feelings associated with it.

9.23.5.3.1 Exposure and related techniques

An essential issue underlying treatment is the length of time one is exposed to the traumatic stimulus and to treatment as well as the interval between treatments. Rather, efficacious treatments tend to involve exposure intervals that are sufficiently long enough to reduce arousal and distress. Briefer exposure intervals, particularly when they elicit high levels of arousal, have the potential to exacerbate problematic functioning. Apart from this issue, variation exists along the dimensions of level of arousal (high vs. low) and type of exposure (imaginal vs. *in vivo*) (Rothbaum & Foa, 1993). For example, systematic desensitization has the client arrange a number of fears along a hierarchy and this procedure is usually carried out in an office setting. The essential features of this approach include imaginal exposure, low levels of arousal, and relatively brief exposure intervals (though long enough to reduce anxiety). By direct contrast, *in vivo* flooding has the client exposed to the feared stimulus in the real environment. The aim is to provoke maximal (flooded) anxiety. With continued and prolonged exposure intervals, the anxiety then begins to diminish.

Of course, any approach using exposure can vary along these dimensions, depending on the nature of the problem (e.g., severity of trauma), availability of real-world stimuli, safety, and pragmatic concerns. For example, intervention following a disaster cannot utilize *in vivo* exposure in the classic sense. Despite this, exposure to stimuli related to the disaster can take place in a variety of ways (e.g., Rothbaum & Foa, 1993). For example, video excerpts have potential in exposure approaches following particular disasters. Ronan and Johnston (1996) used a videotape of a series of volcanic eruptions in an exposure- and school-based cognitive-behavioral prevention program. Finally, with respect to variations in approach, an examination of the available literature indicates that desensitization appears to be used more than flooding in treatments following traumatic events. This could be associated with a sensitivity to the potential of "too much too soon" in the wake of a disaster (Wilson & Raphael, 1993).

Interventions that utilize exposure often include other components. These components are generally designed to provide the individual with coping skills which adds to effects of the exposure or to counteract excess arousal. For example, systematic desensitization employs relaxation (e.g., breathing relaxation, deep muscle relaxation) as an alternative, and competing, response to excess arousal. Imaginal and *in vivo* exposure approaches often use modeling (overt and covert) and role play to help the client practice coping more effectively. Self-reinforcement is used to buttress success and encourage continued attempts at coping. Homework exercises are aimed at helping the client focus on practicing skills learned, continuing to manage activities of daily living, and engaging in pleasurable activities. These prescribed exercises can be designed to help the client reduce avoidance and withdrawal and, importantly, generalize the benefits of treatment to everyday living. Finally, as discussed below, cognitive-behavioral approaches include cognitive coping techniques. In summary, it is essential to emphasize that all variations of exposure treatments, whether graded or flooded, imaginal or *in vivo,* require the client to remain in the presence of the feared, or traumatic, stimulus until arousal and distress reduce to acceptable levels. While prolonged exposure appears necessary, it is still unclear what actual psychological mechanisms (e.g., counterconditionig, habituation) account for the beneficial effects of such exposure. For example, systematic desensitization, which is based on the concept of gradual counterconditioning, has been shown to work without the relaxation component and without a hierarchy (Kazdin & Wilcoxon, 1976). These techniques have shown some promise with traumatized victims (Brom et al., 1989).

9.23.5.3.2 Cognitive intervention components

Suggestions have been made that one process underlying exposure-based treatments includes a cognitive component which involves a growing sense of mastery, or self-efficacy, in fearful situations. In similar manner, cognitive techniques equip the clients with a means by which they can gain an increased sense of control when faced with distressful thoughts and feelings and these include techniques which are aimed at both cognitive distortions and cognitive deficiencies.

In treating distorted processing, cognitive restructuring approaches help the client examine unrealistic beliefs and negative automatic thoughts in a collaborative therapeutic relationship. The essential concept is to help the client anticipate, identify, and modify arousal and thinking that is largely colored by the expectation that bad things will happen (Kendall et al., 1992). For example, following a disaster, a person's cognitions can be dominated by themes of ongoing threat, loss and grief, guilt, and anger. These reactions can keep the disaster alive as well as negatively influencing future events. In most circumstances, distress and coping following disasters unfold in normal and healthy ways and this information is important for clients to understand. It is stressed again that knowledge of the developmental course of postdisaster reactions and normal variation is a primary issue in any pre- or postdisaster intervention. For those who develop conditions such as PTSD, pathological grief, and depression, related problems may be linked with distorted thinking. Combined with normative information, the therapist and client(s) focus on this thinking through the process of collaborative empiricism. That is, they work together to examine the realistic nature of beliefs and automatic thoughts associated with the traumatic event. Thus, explanations based on realities is the continuing theme. Cognitive restructuring can be particularly valuable for those whose trauma includes maladaptive anxiety and depression.

Deficiencies in thinking are addressed through problem-solving approaches. The concept advanced is that cognitive skill-building which helps the client learn systematic strategies can be used to deal with a wide variety of conditions (D'Zurilla, 1986) including trauma. Problem-solving is generally divided into logically-related steps that involve the following components: identifying the nature of the problem, generating alternative solutions and evaluating consequences, deciding on a response, and evaluating the merits of the response. Following traumatic events, this problem-solving sequence will involve identifying realistic and irrational stimuli that are related to the traumatic event and resultant distress. Problem-solving then can be combined with other techniques including exposure to help the client reduce distress by directly facing and systematically coping with problematic situations. For example, in working with anxiety-disordered children, some of whom have experienced traumatic events, Kendall et al. (1992) use a FEAR acronym to help children remember to address the following issues: identifying arousal as the first step to coping and relaxation (*F*eeling frightened?), engaging in cognitive restructuring (*E*xpecting bad things to happen?), implementing cognitive and behavioral coping plans (*A*ttitudes and actions that can help), and practicing self-reinforcement (*R*esults and rewards). This problem-solving approach, once learned, is then applied in a variety of hierarchically organized imaginal and *in vivo* situations. A variety of other techniques including, coping modeling, reinforcement, homework, and parent education are used to help the child learn and generalize these skills. Such an approach may have some promise for children traumatized by disasters (Ronan & Johnston, 1996).

Other techniques are designed to help the client cope with intrusive symptoms. Thought stopping is a technique designed to help the client cope with intrusive thoughts more effectively. Other distracting techniques have been utilized (Hodgkinson & Stewart, 1991). For example, the use of personal stereos helped children get to sleep at night following the Zeebrugge disaster by distracting them from intrusive and distressing reminders of the trauma (Yule & Williams, 1990). While distraction and temporary avoidance can be effective, the client may benefit most by learning additional skills which are designed to assist in a sense of longer-term control.

9.23.5.4 Research on Efficacy and Integrated Treatment Description

The treatment literature in the area of disasters is sparse. In addition, much of the research on the effectiveness of treatment following disasters is limited to case reports or, if using defined groups, does not utilize random assignment or adequate controls. As a result, generalizability of results is often limited. However, some controlled research has begun to examine the effectiveness of treatments for PTSD following a range of trauma (Rothbaum & Foa, 1993; Wilson & Raphael, 1993). For example, Brom et al. (1989) compared trauma desensitization with behaviorally-oriented hypnotherapy, brief psychodynamic therapy (Horowitz, 1986), and a wait-list control condition for clients diagnosed with PTSD. Randomly assigned clients were first taught relaxation followed by imaginal exposure to a hierarchy of situations related to the traumatic event (e.g., accidents). During exposure, clients were encouraged to confront directly (imaginally) the fear-arousing stimulus. The three active treatments were found superior to the wait-list control condition. However, contrary to predictions, the intervention was found to have most beneficial impact on intrusive symptoms and lesser impact on avoidance symptoms. While speculative, the suggestion is that imaginal, rather than some form of *in vivo,* exposure was utilized may have contributed to this pattern of results.

Another study examined the effectiveness of what has been considered by some to be a controversial treatment (Wilson, Becker, & Tinker, 1995). Eye-movement desensitization and reprocessing (EMDR; Shapiro, 1995) works as follows:

> . . . the client is asked to hold in mind an image of the trauma, a negative self-cognition, negative emotions, and related physical sensations about the trauma. While doing so, the client is instructed to move her or his eyes quickly and laterally back and forth for about 15 to 20 seconds, following the therapist's fingers. Other forms of left–right alternating stimulation (auditory, tactile) are sometimes used. The client then reports the images, cognitions, emotions, and physical sensations that emerged. This recursive procedure continues until desensitization of troubling material is complete and positive self-cognitions have replaced the previous negative self-cognition. (Wilson et al., 1995, p. 928)

This therapy has reported positive results but these findings have been criticized on methodological grounds (e.g., Shapiro, 1989). However, a study using a randomized, controlled methodology found three-session EMDR to be more effective than a wait-list condition in a number of areas which included presenting complaints, anxiety, and positive cognition (Wilson et al., 1995). Wait-list participants evidenced no improvement during the control condition on any outcome measure. However, once treated, these same participants evidenced improvement similar to those initially treated on all measures. While the underlying mechanisms of EMDR are unknown, the procedure does incorporate sound behavioral and cognitive techniques and appears to show some promise in treating traumatic memories (Wilson et al.).

Foa, Rothbaum, Riggs, and Murdock (1991) have looked at brief cognitive-behavioral programs for treating established PTSD following a traumatic event (e.g., assaults, rapes). They found a cognitive-behavioral treatment and a prolonged exposure condition both to be more effective than a wait-list control and an alternative treatment for persons who had recently experienced a traumatic event. A brief cognitive-behavioral program was found to be superior to an assessment control condition on a variety of indicators (Foa et al., 1995). This included treated (vs. control) clients hastening recovery as indicated by a significantly reduced frequency of diagnosed PTSD at two-month follow-up (i.e., 70% vs. 10% no longer diagnosed with PTSD). This approach, with modification appropriate to the event, appears to hold promise for use in early intervention with disaster victims.

With the advent of controlled studies, evidence has appeared that supports the use and continued examination of cognitive-behavioral approaches. It must be noted, however, that following disasters as defined earlier in this chapter, there have not been any comprehensive, controlled studies examining the effectiveness of any treatment approach. Thus, while the cognitive-behavioral approach discussed in this section clearly has promise, controlled clinical trials will prove to be the final arbiter.

9.23.6 CONCLUSIONS

Some important methodological and conceptual considerations in the design of disaster studies have been highlighted. For example, problems with sampling, procedure, consistency of measures, and inclusion of appropriate control groups are common. Although knowledge about disasters has greatly increased in the last decade because of methodological improvement in research studies, integration of findings from this line of research has been hampered by its largely atheoretical nature. Many disaster studies focus on the disaster event itself rather than on the theoretical underpinnings of the research. The field needs to base itself on a conceptual framework and utilize a multimethod approach to obtaining data. In addition, there is a need for greater emphasis on the underlying generic and unique dimensions that differentiate responses to different disasters being studied.

We live in increasingly complex societies in which large-scale technological and human-induced disasters are more likely to occur. Research has shown a growing respect for the physical dimensions and effects of disasters, but this knowledge often fails to give the insights into their psychosocial costs. A comprehensive typology of victims and disasters is a prerequisite for the development of this field of study, as it provides the framework within which any systematic evaluation may proceed. Considerable attention has been given to the development and delivery of support and treatment interventions but this has not been matched by a corresponding effort on evaluation, and this needs to be addressed. Professionals working in this area will find increasing external pressures to justify the services they provide. The evaluation of these interventions will require a greater understanding of the relationships between environmental factors, impact, and recovery. Intervention in the future will need to move away from the clinical approach that treats all survivors in relative isolation to one in which treatment is managed by accommodating the social, organizational, and community influences prevailing in the recovery environment. Finally, many disasters cannot be prevented but the magnitude of their physical and psychosocial impact can be reduced by better community preparedness and response.

9.23.7 REFERENCES

Ahearn, F. L., Jr. (1981). Disaster mental health: A pre- and post-earthquake comparison of psychiatric admission rates. *Urban and Social Change Review, 14,* 22–28.

Alexander, D. (1993). *Natural disasters.* London: UCL Press.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Barnard, G. W., Hankins, G. C., & Robbins, L. (1992). Prior life trauma, post-traumatic stress symptoms, sexual disorders, and character traits in sex offenders: An exploratory study. *Journal of Traumatic Stress, 5,* 393–420.

Barton, A. H. (1969). *Communities in disaster.* Garden City, NY: Doubleday.

Baum, A. (1990). Stress, intrusive imagery, and chronic distress. *Health Psychology, 9,* 653–675.

Baum, A. (1991). Toxins, technology, and natural disasters. In G. R. VandenBos & B. K. Bryant (Eds.), *Cataclysms, crises, and catastrophies: Psychology in action* (pp. 9–53). Washington, DC: American Psychological Association.

Baum, A., & Davidson, L. M. (1985). A suggested framework for studying factors that contribute to trauma in disaster. In B. J. Sowder (Ed.), *Disasters and mental health: Selected contemporary perspectives* (pp. 29–40). Rockville, MD: US Department of Health and Human Services, National Institute of Mental Health.

Baum, A., Fleming, R., & Davidson, L. M. (1983). Natural disaster and technological catastrophe. *Environment and Behavior, 15,* 333–354.

Baum, A., Grunberg, N. E., & Singer, J. E. (1982). The use of psychological and neuroendocrinological measurements in the study of stress. *Health Psychology, 1,* 217–236.

Baum, A., Schaeffer, M. A., Lake, C. R., Fleming, R., & Collins, D. L. (1986). Psychological and endocrinological correlates of chronic stress at Three Mile Island. In R. William (Ed.), *Perspectives on behavioral medicine* (pp. 201–217). New York: Academic Press.

Baum A., Solomon, S. D., Ursano, R. J., Bickman, L., Blanchard, E., Green, B. L., Keane, T., Laufer, R., Norris, F., Reid, J., Smith, E., & Steinglass, P. (1993). Emergency/disaster studies: Practical, conceptual, and methodological issues. In J. P. Wilson & B. Raphael (Eds.), *International handbook of traumatic stress syndromes* (pp. 125–133). New York: Plenum.

Beck, A. T., & Steer, R. A. (1984). Internal consistencies of the original and revised Beck Depression Inventory. *Journal of Clinical Psychology, 40,* 1365–1367.

Beigel, A., & Berren, M. R. (1985). Human induced disasters. *Psychiatric Annals, 15,* 143–150.

Belter, R. W., & Shannon, M. P. (1993). Impact of natural disasters on children and families. In C. F. Saylor (Ed.), *Children and disasters* (pp. 85–103). New York: Plenum.

Berren, M. R., Santiago, J. M., Beigel, A., & Timmons, S. A. (1989). A classification scheme for disasters. In R. Gist & B. Lubin (Eds.), *Psychosocial aspects of disaster* (pp. 40–58). New York: Wiley.

Bland, S. H., O'Leary, E. S., Farinardo, E., Jossa, F., & Trevisan, M. (1996). Long-term psychological effects of natural diasters. *Psychosomatic Medicine, 58,* 18–24.

Bolin, R. (1985). Disaster characteristics and psychosocial impacts. In B. J. Sowder (Ed.), *Disasters and mental health: Selected contemporary perspectives* (DHHS Publication No. ADM 85–1421, pp. 3–28). Rockville, MD: US. Department of Health and Human Services, National Institute of Mental Health.

Bolin, R. (1988). Response to natural disasters. In M. Lystad (Ed.), *Mental health response to mass emergencies: Theory and practice* (pp. 22–51). New York: Brunner/Mazel.

Breslau, N., Davis, G. C., Andreski, P., & Peterson, E. (1991). Traumatic events and posttraumatic stress disorder in an urban population of young adults. *Archives of General Psychiatry, 48,* 216–222.

Brom, D., Kleber, R. J., & Defares, P. B. (1989). Brief psychotherapy for posttraumatic stress disorders. *Journal of Consulting and Clinical Psychology, 57,* 607–612.

Bromet, E. J., & Dew, M. A. (1995). Review of psychiatric epidemiological research on disasters. *Epidemiologic Reviews, 17,* 113–119.

Bromet, E. J., Parkinson, D. K., Schulberg, H. C., Dunn, L., & Gondek, P. C. (1980). *Three Mile Island: Mental health findings.* Pittsburgh, PA: Western Psychiatric Institute.

Busuttil, W., Turnbull, G. J., Neal, L. A., Rollins, J., West, A. G., Blanch, N., & Herepath, R. (1995). Incorporating psychological debriefing techniques with a brief group psychotherapy programme for the treatment of post-traumatic stress disorder. *British Journal of Psychiatry, 167,* 495–502.

Davidson, L. M., & Baum, A (1986). Chronic stress and posttraumatic stress disorders. *Journal of Consulting and Clinical Psychology, 54,* 303–308.

Davidson, J. R., & Foa, E. B. (1991). *Posttraumatic stress disorder: DSM-IV and beyond*. Washington, DC: American Psychiatric Press.

Derogatis, L. R. (1983). *SCL-90-R version: Manual I*. Baltimore: John Hopkins University.

Dew, M. A., & Bromet, E. J. (1993). Predictors of temporal patterns of psychiatric distress during 10 years following the nuclear accident at Three Mile Island. *Social Psychiatry and Psychiatric Epidemiology, 28,* 49–55.

D'Zurilla, T. (1986). *Problem-solving therapy. A social competence approach to clinical intervention*. New York: Springer.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1994). *Structured clinical interview for axis I DSM-IV disorders—patient edition* (SCID-I/P, version 2), New York: Biometrics Research Department.

Foa, E. B., Hearst-Ikeda, D., & Perry, K. J. (1995). Evaluation of a brief cognitive-behavioral program for the prevention of chronic PTSD in recent assault victims. *Journal of Consulting and Clinical Psychology, 63,* 948–955.

Foa, E. B., Rothbaum, E. O., Riggs, D., & Murdock, T. (1991). Treatment of PTSD in rape victims: A comparison between cognitive-behavioral procedures and counselling. *Journal of Consulting and Clinical Psychology, 59,* 715–723.

Friedman, P., & Linn, L. (1957). Some psychiatric notes on the Andrea Doria disaster. *American Journal of Psychiatry, 114,* 426–432.

Galbaud du Fort, G., Newman, S. C., & Bland, R. C. (1992). Psychiatric comorbidity and treatment seeking: Sources of selection bias in the study of clinical populations. *Journal of Nervous and Mental Diseases, 181,* 467–474.

Gleser, G., Green, B., & Winget, C. (1981). *Prolonged psychosocial effects of disaster: A study of Buffalo Creek*. New York: Academic Press.

Green, B. L. (1982). Assessing levels of psychosocial impairment following disaster: Consideration of actual and methodological dimensions. *Journal of Nervous and Mental Disease, 17,* 544–552.

Green, B. L. (1991). Evaluating the effects of disasters. *Psychological Assessment, 3,* 538–546.

Green, B. L. (1993). Disasters and posttraumatic stress disorder. In J. R. Davidson & E. B. Foa (Eds.), *Posttraumatic stress disorder: DSM-IV and beyond* (pp. 75–97). Washington, DC: American Psychiatric Press.

Green, B. L. (1994). Psychosocial research in traumatic stress. An update. *Journal of Traumatic Stress, 7,* 341–362.

Green, B. L., Lindy, J. D., & Grace, M. C. (1994). Psychological effects of toxin contamination. In R. J. Ursano, C. S. Fullerton, & B. G. McCaughey (Eds.), *Individual and community responses to trauma and disaster: The structure of human chaos* (pp. 154–176). New York: Cambridge University Press.

Green, B. L., & Solomon, S. D. (1995). The mental health impact of natural and technological disasters. In J. R. Freedy & S. E. Hobfoll (Eds.), *Traumatic stress: From theory to practice* (pp. 163–180). New York: Plenum.

Harshbarger, D. (1974). Picking up the pieces: Disaster intervention and human ecology. *Omega, 5,* 55–59.

Hartsough, D. M. (1985). Measurement of the psychological effects of disaster. In J. Laube & S. A. Murphy (Eds.), *Perspectives on disaster recovery* (pp. 22–60). Northwalk, CT: Appleton-Century-Crofts.

Henderson, S., & Bostock, T. (1977). Coping behavior after shipwreck. *British Journal of Psychiatry, 131,* 15–20.

Hodgkinson, P. E., & Stewart, M. (1991). *Coping with catastrophe: A handbook of disaster management*. London: Routledge.

Horowitz, M. J. (1986). *Stress-response syndromes* (2nd ed.). Northvale, NJ: Jason Aronson.

Horowitz, M. J., Wilner, N., & Alvarez, W. (1979). Impact of events scale. A measure of psychosomatic stress disorder. *Psychosomatic Medicine, 41,* 209–218.

Kazdin, A. E., & Wilcoxon, L. (1976). Systematic desensitization and nonspecific treatment effects: A methodological evaluation. *Psychological Bulletin, 83,* 729–758.

Keane, T. M., Fairbank, J. A., Caddell, J. M., Zimering, R. T., & Bender, M. (1985). A behavioral approach to assessing and treating post-traumatic stress disorder in Vietnam veterans. In C. R. Figley (Ed.), *Trauma and its wake: The study and treatment of post-traumatic stress disorder* (pp. 257–294). New York: Brunner/Mazel.

Kendall, P. C., Chansky, T. E., Kane, M. T., Kim, R., Kortlander, E., Ronan, K. R., Sessa, F. M., & Siqueland, L. (1992). *Anxiety disorders in youth: Cognitive-behavioral interventions*. Needham, MA: Allyn & Bacon.

Kendall, P. C. & Hammen, C. (1995). *Abnormal psychology*. Boston: Houghton-Mifflin.

Koopman, C., Classen, C., Cardena, E., & Speigel, D. (1995). When disaster strikes, acute stress disorder may follow. *Journal of Traumatic Stress, 8,* 29–46.

Kramer, T. L., & Green, B. L. (1991). Posttraumatic stress disorder as an early response to sexual assault. *Journal of Interpersonal Violence, 6,* 160–173.

Lazarus, R. S., & Cohen, J. B. (1978). Environmental stress. In I. Altman & J. F. Wohlwill (Eds.), *Human behavior and the environment: Current theory research* (Vol. 2, pp. 89–127). New York: Plenum.

Lifton, R. J., & Olson, E. (1976). The human meaning of total disaster: The Buffalo Creek experience. *Psychiatry, 39,* 1–18.

Long, N., Chamberlain, K., & MacDonald, C. (1996). Prevalence of posttraumatic stress disorder, depression and anxiety in a community sample of New Zealand Vietnam war veterans. *Australian and New Zealand Journal of Psychiatry, 30,* 253–256.

Long, N., Chamberlain, K., & Vincent, C. (1994). Effects of the Gulf war on reactivation of adverse combat related memories in Vietnam veterans. *Journal of Clinical Psychology, 50,* 138–144.

McFarlane, A. C. (1995). Helping the victims of disasters. In J. R. Freedy & S. E. Hobfoll (Eds.), *Traumatic stress: From theory to practice* (pp. 287–314). New York: Plenum.

Melick, M. E. (1985). The health of postdisaster populations: A review of literature and case study. In J. Laube & S. A. Murphy (Eds.), *Perspectives on disaster recovery* (pp. 179–209). Northwalk, CT: Appleton-Century-Crofts.

Norris, F. H. (1992). Epidemiology of trauma: Frequency and impact of different potentially traumatic events on different demographic groups. *Journal of Consulting and Clinical Psychology, 60,* 409–418.

Paton, D., & Long, N. (1996). *Psychological aspects of disasters: Impact, coping and intervention*. Palmerston North, New Zealand: Dunmore.

Quarantelli, E. L. (1985). What is disaster? The need for clarification in definition and conceptualization in research. In B. J. Sowder (Ed.), *Disasters and mental health: Selected contemporary perspectives* (pp. 41–73). Rockville, MD: Department of Health and Human Services, National Institute of Mental Health.

Raphael, B. (1986). *When disaster strikes: A handbook for the caring professions*. London: Hutcheson.

Robins, L. N., & Smith, E. M. (1983). *Diagnostic interview schedule/Disaster supplement*. St. Louis, MO: Washington University School of Medicine, Department of Psychiatry.

Ronan, K. R., & Johnston, D. M. (1996). *The impact of volcanic eruptions on childhood emotional functioning.*

Vancouver, BC: Proceedings of the Pan Pacific Hazards '96 Conference [CD-Rom].

Rothbaum, B. O., & Foa, E. B. (1993). Cognitive behavioral treatment of post-traumatic stress disorder. In P. A. Saigh (Ed.), *Posttraumatic stress disorder: A behavioral approach to assessment and treatment* (pp. 85–110). New York: Macmillan.

Rubonis, A. V., & Bickman, L. (1991). Psychological impairment in the wake of disaster: The disaster–psychopathology relationship. *Psychological Bulletin, 109,* 384–399.

Saylor, C. F. (1993). *Children and disasters.* New York: Plenum.

Seaman, J. (1984). *Epidemiology of natural disasters.* Basel, Switzerland: Karger.

Shapiro, F. (1989). Efficacy of the eye movement desensitization procedure in the treatment of traumatic memories. *Journal of Traumatic Stress Studies, 2,* 199–223.

Shapiro, F. (1995). *Eye movement desensitization and reprocessing: Basic principles, protocols and procedures.* New York: Guilford.

Smith, E. M., & North, C. S. (1993). Posttraumatic stress disorder in natural disasters and technological accidents. In J. P. Wilson & B. Raphael (Eds.), *International handbook of traumatic stress syndromes* (pp. 405–419). New York: Plenum.

Smith, E. M., North, C. S., & Price, P. C. (1988). Response to technological accidents. In M. Lystad (Ed.), *Mental health response to mass emergencies: Theory and practice* (pp. 52–95). New York: Brunner/Mazel.

Solomon S. D. (1989). Research issues in assessing disaster's effects. In R. Gist & B. Lubin (Eds.), *Psychosocial aspects of disaster* (pp. 308–340). New York: Wiley.

Solomon, S. D., & Smith, E. M. (1994). Social support and perceived control as moderators of response to dioxin and flood exposure. In R. J. Ursano, C. S. Fullerton, & B. G. McCaughey (Eds.), *Individual and community responses to trauma and disaster: The structure of human chaos* (pp. 179–200). New York: Cambridge University Press.

Taylor, T. (1996). Disaster and victim classification. In D. Paton & N. Long (Eds.), *Psychological aspects of disasters: Impact, coping, and intervention* (pp. 26–39). Palmerston North, New Zealand: Dunmore.

Tyrer, P. (1996). Cormorbity or consanguinity. *British Journal of Psychiatry, 168,* 669–671.

Ursano, R. J., Fullerton, C. S., & McCaughey, B. G. (1994). Trauma and disaster. In R. J. Ursano, C. S. Fullerton, & B. G. McCaughey (Eds.), *Individual and community responses to trauma and disaster: The structure of human chaos* (pp. 3–27). New York: Cambridge University Press.

Vernberg, E. M., & Vogel, J. (1993). Interventions with children after disasters. *Journal of Clinical Child Psychology, 22,* 484–498.

Vrana, S., & Lauterbach, D. (1994). Prevalence of traumatic events and post-traumatic psychological symptoms in a nonclinical sample of college students. *Journal of Traumatic Stress, 7,* 289–302.

Warheit, G. H. (1988). Disasters and their mental health consequences: Issues, findings, and future trends. In M. Lystad (Ed.), *Mental health response to mass emergencies: Theory and practice* (pp. 3–21). New York: Brunner/Mazel.

Weisaeth, L. (1994). Psychological and psychiatric aspects of technological disasters. In R. J. Ursano, C. S. Fullerton, & B. G. McCaughey (Eds.), *Individual and community responses to trauma and disaster: The structure of human chaos* (pp. 72–102). New York: Cambridge University Press.

Wilson, S. A., Becker, L. A., & Tinker, R. H. (1995). Eye movement desensitization and reprocessing (EMDR): Treatment for psychologically traumatized individuals. *Journal of Consulting and Clinical Psychology, 63,* 928–937.

Wilson, J.P., & Raphael, B. (Eds.) (1993). *International handbook of traumatic stress syndromes* (pp. 395–404). New York: Plenum.

Wortman, C. B., & Silver, R. C. (1989). The myths of coping with loss. *Journal of Consulting and Clinical Psychology, 57,* 349–357.

Yule, W., & Williams, R. (1990). Posttraumatic stress reactions in children. *Journal of Traumatic Stress, 3,* 279–295.

Zisook, S., & Schuchter, S. R. (1986). The first four years of widowhood. *Psychiatric Annals, 15,* 288–294.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.24
Victims of War

NORMAN A. MILGRAM
Tel Aviv University, Israel

9.24.1 INTRODUCTION 392

9.24.2 STRESS AND COPING PARADIGMS 392

- *9.24.2.1 Person–Environment Encounters* 393
- *9.24.2.2 Cognitive Appraisals* 393
- *9.24.2.3 Normal Reactions to Extraordinary Stressors* 393
- *9.24.2.4 Stress-related Reactions and Disorders* 394

9.24.3 WAR-RELATED STRESSORS 394

- *9.24.3.1 Qualitative and Quantitative Features* 394
- *9.24.3.2 Levels of Professional Intervention* 395

9.24.4 ASSESSMENT AND TREATMENT IN DEVASTATED SOCIETIES 396

- *9.24.4.1 Assessment* 396
- *9.24.4.2 Cross-cultural Considerations in Treatment* 396
- *9.24.4.3 Mental Health of Child Refugees* 397
- *9.24.4.4 Victims of Torture* 398
- *9.24.4.5 Prevalence* 398
- *9.24.4.6 Factors Affecting Adjustment* 398
- *9.24.4.7 Sequelae of Torture* 399
- *9.24.4.8 Treatment Guidelines* 399
- *9.24.4.9 Treatment Centers* 399

9.24.5 THE EFFECT OF WAR ON CIVILIAN POPULATIONS IN INTACT SOCIETIES 400

- *9.24.5.1 The USA and the Gulf War* 400
- *9.24.5.2 Israel and the Gulf War* 401

9.24.6 LONG-TERM ADJUSTMENT OF TRAUMATIZED HOSTAGES 401

9.24.7 PROFESSIONAL RESPONSIBILITIES OF CLINICAL PSYCHOLOGISTS 402

- *9.24.7.1 Ten Precautions* 402
  - *9.24.7.1.1 Problem of empathy with victims of human malevolence* 402
  - *9.24.7.1.2 Lack of concern for the client's cultural background* 403
  - *9.24.7.1.3 Misapplication of the psychodynamic orientation to traumatized clients* 403
  - *9.24.7.1.4 Nonmatching and inappropriate attributions* 403
  - *9.24.7.1.5 Professional bias against war victims* 403
  - *9.24.7.1.6 Discomfort with expressions of revenge* 403
  - *9.24.7.1.7 Reluctance to use lay support groups in treatment* 404
  - *9.24.7.1.8 Reluctance to accept the role of the outsider* 404
  - *9.24.7.1.9 Caution against professional burnout* 404
  - *9.24.7.1.10 Danger of being overly cautious* 404
- *9.24.7.2 Responsibilities of Clinical Psychologists* 404

9.24.8 REFERENCES 405

9.24.1 INTRODUCTION

War adversely affects the lives of millions of men, women, and children. There have been 471 wars since 1700 resulting in the death of over 100 million people, and more wars and casualties in the twentieth century than in any previous one. Since the end of World War II, there have been 150 wars, with 22 million dead and millions of survivors whose lives were disrupted by armed conflict (Goldson, 1993). Wars or armed conflicts take many forms. Armed conflicts have taken place between sovereign states (e.g., USA–Vietnam, Iran–Iraq), between warring ethnic and religious factions within a state (e.g., Hutu–Tutsi tribes in Rwanda, Serbs–Croats–Moslems in the former Yugoslavia), and between governments and their opposition (e.g., Greece and Argentina during the rule of the military juntas). There are 20 major ongoing conflicts (defined as over 1000 victims per year) in the world today, the majority in Africa and Asia. There are 25 intermediate and minor conflicts in these continents, Europe, and the Americas (International Federation of Red Cross and Red Crescent Societies, 1995).

Many people suffer the consequences of an armed conflict, starting with the soldiers who fight in it and ending with children born long after the armed conflict has ceased. The first immediate consequence of war is that soldiers die or are maimed. The second is that unarmed civilians also die or are maimed, and increasingly so in the twentieth century. Cities are bombed as legitimate military targets because their inhabitants produce weapons for their army or because the destruction of cities will destroy the morale of the home front and eventually the morale of the armed forces of the country. The third is the displacement and flight of people from their homes to other areas in their country or to other countries. Their number is estimated at this time to be in excess of 18 million, with the largest number being African refugees, followed by Asian, Near and Middle Eastern, and European. The fourth is the predicament of the survivors of war who remain in their native countries and labor to repair the damage caused by war and to rebuild their homes and lives. The fifth is the severe psychological and psychiatric damage suffered by those exposed to war.

The fallout of war is extensive. Innocent civilians are suddenly caught up in war when they are seized as hostages by one group and used as pawns in a deadly game of chess played before the eyes of the world. War reaches secondary victims, the families and friends of the primary victims who were directly affected by hostilities. War also affects the millions who read, hear, and view ongoing conflicts and their bitter fruits via newspapers, radio networks, TV channels, historical and fictional novels, and films.

The legacy of war extends well into the future. It affects children born after the war to primary and secondary victims. The material costs incurred by a country and its citizens for having engaged in war are paid in installments for 50 years or more. The books of World War I have not yet been closed, as veterans of that war continue to draw compensation and benefits of one kind or another. The books of World War II will not close for another 50 years. The psychological costs of wars are enshrined in the historical consciousness of a people and are manifest in its history books, cemeteries, war memorials, statues, emblems, and commemorative celebrations.

The literature on victims of war is voluminous. The present chapter restricts itself to civilian populations who personally suffered traumatic experiences during armed conflict and/or subsequent to it: (i) civilians who remain in their war-devastated country; (ii) civilians forced to leave their country who seek refuge elsewhere; (iii) civilians in countries not devastated in the war in which their country was engaged; and (iv) civilian hostages released after harrowing experiences.

9.24.2 STRESS AND COPING PARADIGMS

In order to understand the effects of war, it is necessary to become acquainted with the stress and coping paradigms used in the investigation and treatment of stress-related reactions and disorders in general. Mental health professionals who work with victims of natural or manmade disasters are familiar with the theories and concepts used in two complementary stress and coping paradigms (Hobfoll, 1989; Lazarus & Folkman, 1984). The Lazarus and Folkman (1984) paradigm deals with objective parameters of external stressors (e.g., intensity, duration, and recurrence of stressful life events), cognitive appraisals of these stressors, cognitive appraisals of the intrapersonal and interpersonal resources available to cope with these stressors, the consequent stress reactions, emotion-focused efforts to cope with these reactions, and problem-focused efforts to deal with the environmental stressors. The Hobfoll (1989) paradigm defines stress as a threat of loss of resources, loss of resources, or an insufficient return from the investment of resources. Resources refer to material possessions, personal characteristics, values, status, time, and money. The paradigm posits that resource loss has more adverse effect than the positive effect of resource gain, and that a rapid depletion of resources leaves individuals, communities, and societies unable to protect their remaining resources.

The first paradigm focuses on the demand character of a stressful situation (what kinds of adaptations and adjustments people are called upon to make). The second focuses on what is threatened or lost in a stressful situation. These paradigms have many points of similarity with psychodynamic and interpersonal theories of psychopathology, but their differences are of concern here.

9.24.2.1 Person–Environment Encounters

The concepts of these paradigms are applied to three different kinds of person–environment encounters. The first is characterized by demands on people that are easily handled (e.g., a garage mechanic has to fix a flat tire). People have developed competencies that operate as automatic or habitual patterns in dealing with these kinds of situations and are not threatened when they have to deal with them again. If stress refers to the discrepancy between environmental demands and personal supplies or competencies (Lazarus & Folkman, 1984), these encounters are minimally stressful. Alternatively, if stress is defined as the threat of loss or the actual loss of valued resources (Hobfoll, 1989), these recurring encounters contain no threat of loss or actual loss, and are resolved with dispatch.

The second kind of encounter is characterized by heavy demands on people's problem-solving competencies and the threat of loss of important resources (e.g., danger of bankruptcy). The demands of the encounter and the potential or actual losses evoke emotional reactions that depend on people's cognitive assessment of the situation. For example, there may be anxiety at the prospect of failing to handle the situation properly, depression when convinced one will fail and suffer dire consequences, anger at oneself for personal inadequacies or at others for being placed in this untenable situation, or any combination of these emotions. Conversely, there may be keen anticipation at the prospects of meeting the challenge successfully. Two kinds of perceived discrepancy affect one's subjective sense of well being or of distress, and thereby facilitate or interfere with one's efforts to deal effectively with the stressful situation at hand. A favorable discrepancy between situational demands and available resources is reassuring and facilitates optimal performance; an unfavorable discrepancy has the opposite effect. Favorable and unfavorable discrepancies between one's perceived and one's expected coping behavior have corresponding effects. Given favorable appraisal outcomes, people are likely to cope well with the demands of the situation and to remain free of intense, aversive emotions that are known to disrupt functioning in difficult circumstances.

The third kind of encounter refers to extraordinary, highly aversive, or undesirable events (e.g., a drunken driver causes irreversible brain damage to his loved one). Such events are beyond the range of personal experience of most people and call upon cognitive and behavioral competencies that were hitherto not developed. Furthermore, these unfortunate events constitute grave threats to valued resources or cause the outright loss of valued resources (e.g., loss of financial security). Most people lack the competencies required for dealing immediately and effectively with these threats and losses. They are unable to assimilate these shattering experiences to preexisting assumptions and experiences (Janoff-Bulman, 1988), experience a sense of helplessness (Seligman, 1975), become overwhelmed, and suffer acute and possibly chronic stress reactions.

9.24.2.2 Cognitive Appraisals

Emotional and behavioral responses to highly stressful situations are mediated by three interrelated kinds of cognitive appraisal: primary, secondary, and tertiary. Primary cognitive appraisals ascertain the parameters of the threatening situation, its imminence, anticipated impact, and consequences. Secondary cognitive appraisals ascertain what is required to avert a disaster altogether or to minimize its effects if it cannot be averted. Secondary appraisals also ascertain whether the coping person has the wherewithal to cope with the disaster or whether one can recruit others to work in one's behalf. Tertiary cognitive appraisal takes place in the course of a prolonged stressful situation (Milgram, 1998; Taylor & Lobel, 1989) and answers the question as to how well the coping person is handling the situation. People engage in interrelated primary, secondary, and tertiary appraisals many times over the course of an evolving crisis and its changing circumstances.

9.24.2.3 Normal Reactions to Extraordinary Stressors

An implicit assumption of the stress formulations is that many well functioning people become distressed when exposed to stressful situations or become even more distressed and behave in a dysfunctional manner during and after exposure to extraordinary situations. Typically, stress reactions are more severe and long-lasting when the extraordinary situations are of human design.

War is the epitome of a malevolent human encounter because it mobilizes all members of society in a united effort to kill other human beings, the enemies of the society (Weisath & Eitinger, 1993). War shatters cherished assumptions about people and human existence—that people are good, that events in the world are meaningful and not random, and that the person making these assumptions is a person of worth (Janoff-Bulman, 1988). This unique character of war places both victimizers and victims of war at high risk for severe, chronic stress disorders and other comorbid disorders. This is not surprising because war imposes demands and requires adjustments for which people lack prior experience and competence and because war compels them to contend with actual or potential losses of the most important things in their lives. Given the enormity of war-related stressors, severe stress reactions to these consensually extraordinary stressors may be regarded as commensurate and, in a sense, normal (Milgram & Hobfoll, 1986).

9.24.2.4 Stress-related Reactions and Disorders

Clinical psychologists who work with war victims must become as familiar with the various stress-related disorders as they are with the kinds of family-related disorders they encounter in their conventional clinical practice. The *Diagnostic and statistical manual of mental disorders* (4th ed., *DSM-IV*) (American Psychiatric Association, 1994) lists five major stress-related disorders: adjustment disorder, acute stress disorder, post-traumatic stress disorder (PTSD), complex PTSD, and disorders of extreme stress. The first three have been formally incorporated into the *DSM-IV* diagnostic system. The fourth was proposed by Herman (1992) and the fifth by van der Kolk, Roth, Pelcovitz, and Mandel (1993). These five disorders range from a relatively mild disturbance in response to a common, mildly upsetting stressor to a profoundly severe disorder in response to an extraordinary stressor.

Adjustment disorders occur in response to developmental crises or other predictable difficulties that emerge in the course of life. Symptoms and impairment arising from exposure to war-related stressors do not receive this diagnosis. Acute stress disorder and PTSD may appear following exposure to war-related stressors. They have several criteria in common: (i) exposure to a traumatic event; (ii) experience of intense fear, helplessness, or horror in response to the above exposure; (iii) dissociative experiences; (iv) reexperience of the original traumatic event in recurrent images, thoughts, dreams, illusions, flashbacks, *déjà vu* phenomena; (v) marked avoidance of stimuli that arouse recollection of the traumatic event and distress when exposed to reminders; (vi) marked anxiety and arousal patterns; and (vii) event-related distress or impairment in social, occupational, or other important areas of functioning. The two disorders differ in time of onset following the traumatic event, intensity and duration of symptoms, and prognosis. The former disorder is associated with less traumatic experiences, less severe symptoms, and greater probability of spontaneous and more rapid recovery. Many war victims, however, are anxious, or depressed, or easily startled without qualifying for either of these formal diagnoses.

Complex PTSD and disorders of extreme stress are found in individuals who have experienced prolonged and repetitive exposure to traumatic stressors. Examples of war-related situations are periods of time characterized by starvation, deprivation, torture, and constant threat to life and limb. The *DSM-IV* clinical field trial indicated that interpersonal traumatization in the first decade of life gave rise to more complex PTSD symptomatology than later traumatization and that Herman's complex PTSD appears to be a combination of disorders of extreme stress and PTSD. Evidence of widespread somatization was also found in the field trial of disorders of extreme stress (van der Kolk et al., 1993). It took the form of manifestations of and preoccupation with a wide variety of somatic complaints including cardiopulmonary symptoms, digestive difficulties, chronic pain, sexual dysfunction, and conversion symptoms. Many psychiatric disorders co-occur with disorders of extreme stress as well as with the other stress-related disorders. These include substance abuse, recurrent depressive episodes, generalized anxiety, panic attacks, and explosive outbursts.

9.24.3 WAR-RELATED STRESSORS

9.24.3.1 Qualitative and Quantitative Features

Some armed conflicts are restricted in scope and intensity leaving entire groups and regions relatively unaffected. The USA has fought several wars in the twentieth century with no harm to its civilian population or to the infrastructure of the country itself. Other countries (e.g., Cambodia, Vietnam, Afghanistan) have been devastated by war, with huge loss of life of civilians and heavy damage to the essential services of the society.

Milgram (1986) proposed a brief checklist for assessing the cumulative stressful features of armed conflicts for individuals, specific groups within the society, and for the society as a whole. This included: (i) personal injury; (ii) loss of life and limb of loved ones; (iii) destruction of one's home, property, possessions; (iv) prolonged lack of food, water, shelter, communication, transportation, medical care, and other amenities; (v) incarceration, torture, threat to one's life; (vi) participation in atrocities or observation of atrocities inflicted by others; (vii) shelling, bombing, destruction of private dwellings as well as public buildings; and (viii) injury and death of strangers.

Refugees may also be exposed to the following stressors: (i) expulsion from one's country or flight from it; (ii) conflict over leaving one's homeland and immigrating to an alien land and culture; (iii) the peril and the pain in the passage to the new country; (iv) the refugees' encounter with the language, people, culture, and economic system of the country of asylum; (v) relocation in the new country may entail new interpersonal loses as family members and friends become physically and psychologically separated; (vi) the loss of status in the home; (vii) loss of status in the work place. Many refugees may be forced to work at occupations far less attractive and prestigious than in their country of origin.

To offer professional consultation or treatment to war victims remaining in their country of origin, it is necessary to assess the presence and intensity of these stressors in their war experiences, and to become familiar with these parameters in the group, community, and society to which these victims belong. Prognosis for one's rehabilitation, recovery, and return to prewar or optimal functioning is more favorable if one has been exposed to fewer stressors, if the society is largely intact and functioning effectively during and after the armed conflict, and if it is able and willing to identify its war victims and provide them needed support (Jerusalem, Kaniasty, Lehman, Ritter, & Turnbull, 1995). The reverse—the physical and psychological devastation of a society—constitutes a grievous loss, personal and collective, by placing in jeopardy the society's way of life, its values, social and political system, and its national and religious observances.

9.24.3.2 Levels of Professional Intervention

Kaplan and Sadock (1985) distinguish three levels of mental health intervention. Primary intervention aims to eliminate a disease or disorder before it occurs. Secondary intervention aims to shorten the course of an acute disorder by early identification and rapid, crisis intervention. Tertiary intervention aims to reduce chronicity through prevention of complications and active rehabilitation programs. Mental health treatment has typically focused on the tertiary level in the mental health intervention spectrum, that is treating the symptomatic patient or group of symptomatic clients (US Committee on Prevention of Mental Disorders, 1994). It is tempting for the clinical psychologist accustomed to individual psychotherapy to use the same approach with war victims, and it may even be appropriate in some settings with some clients. An example would be mental health centers in Western countries dealing with resettled refugees who continue to be distressed and/or unable to function productively in their new homes.

Multimethod hierarchical intervention models have been developed by groups working with people in Third World countries devastated by war. The model developed by the Cross-cultural Division of the Institute for Psychosocial and Socioecological Research in The Netherlands (de Jong, 1995) describes primary and secondary intervention for refugees awaiting repatriation at the national, community, family, and individual levels.

Intervention at the primary level takes the form of advocating and assisting in the implementation of political, social, psychological, and economic policies designed to reduce the probability of armed conflict, the adverse consequences of armed conflict during the conflict and after it has ended, and the numbers of displaced people seeking asylum elsewhere as a consequence of the conflict. Well-conceived and well-executed policies at the national and community levels may prepare groups at risk for displacement to withstand pressure to flee, to adapt better if and when they do flee or are forcibly displaced, and encourage repatriation after hostile actions have ceased.

Secondary intervention consists of early identification of groups and individuals at high risk for traumatization, and rapid intervention. Early identification requires the establishment of a network of community resources that aid in early identification, prompt referral, and treatment. To expedite early identification, indigenous relief workers must be taught to recognize people in need of their services. In addition, all must be given information in advance about the availability of help if and when the need arises (Milgram, Sarason, Schonplug, Jackson, & Schwarzer, 1995). Secondary intervention provides self-help groups for war victims with common problems, fosters self-reliance, and ensures more effective utilization of the resources allocated by government or nongovernment agencies.

The form that tertiary intervention takes with war victims depends on where they are at the time, and the ratio of potential clients to therapists. One may offer an intensive, highly individualized kind of treatment to a Southeast Asian refugee living in Boston (Mollica, 1988) and a far different kind of treatment to people in a crowded refugee camp in a hostile or indifferent Third World country (McIntyre & Ventura, 1996; Somasundaram, 1993). In the former instance, one attempts to restore the individual to optimal functioning and, in the latter, one attempts to strengthen survival skills and minimize interfering symptoms so that the afflicted people can move on to the next way station in their journey to a permanent residence.

9.24.4 ASSESSMENT AND TREATMENT IN DEVASTATED SOCIETIES

9.24.4.1 Assessment

Historically, the investigation and treatment of war victims in general and refugees in particular began with Jews who survived the Nazi Holocaust in Europe. A vast literature has accumulated on the postwar adjustment of Jewish survivors from vastly different wartime settings: working with resistant partisan groups; physically hiding for years; living with concealed identities; working in labor camps; and surviving concentration camps and death camps. These settings undoubtedly differed in the kind, frequency, and intensity of stressful and traumatic events encountered, but there were also wide individual differences within each kind of setting. Moreover, many survivors had been in more than one setting. There were also major differences in the postwar settings in which these people found themselves. It is not surprising, therefore, with the best of methodological rigor available since the 1940s, that conclusions from one set of studies on Holocaust survivors are not necessarily consistent with those of another.

Lomranz (1995) has provided a systematic analysis of this research and has drawn some cautious conclusions:

(i) A salutogenic assessment (Antonovsky, 1979) that examines superficial indices of healthy functioning shows that the vast majority of these people lived reasonably productive lives.

(ii) A pathogenic assessment that surveys the prevalence of subtle as well as overt psychiatric symptoms and disorders shows that a significant number were anxious, angry, depressed, prone to somatic complaints, confined their social contacts to other survivors, and were otherwise isolated and alienated from society for most of their lives, notwithstanding their overall satisfactory adjustment (Harel, Kahana, & Kahana, 1993).

(iii) Each person's experience during the Holocaust probably affected their family life and parenting behavior. These behaviors produced the complex phenomenon known as the transgenerational effect (Danieli, 1981, 1993).

(iv) The postwar adjustment of Holocaust survivors was affected not only by the events that preceded their liberation but also by the events that followed it: where they settled after the war; the reception they received from Jews and Christians; the cultural and religious milieu in which they lived; and their own impact over time on the society in which they lived. All of these factors affected their life adjustment.

(v) With advancing age there were more frequent intrusions into conscious awareness of frightening thoughts and images from the Holocaust. These intrusions were elicited by stimuli reminiscent of the Holocaust (e.g., the Demaniuk trials, the wearing of gas masks by all Israelis during the Gulf war) as well as by age-related events that evoked feelings of helplessness.

Many of these conclusions also apply to refugee groups who immigrated to Western countries following devastating wars in Vietnam and Cambodia (Kinzie, 1993). A recent three-year follow-up of Vietnamese refugees in Norway showed that very severe traumatic experiences were associated with mental health problems (Hauff & Vaglum, 1995). It was also found that stressful events preceding or following forced migration had an independent impact on subsequent mental health. Similar findings were noted in a recent study of Ethiopian Jews who migrated to Israel (Arieli & Aycheh, 1992). They showed a high frequency of stress-related symptoms associated both with the physical and psychological hardships endured in reaching Israel as well as with difficulties in adjusting to the secular life-style prevailing in Israel.

9.24.4.2 Cross-cultural Considerations in Treatment

Cultural differences must be taken into account by mental health workers planning and implementing treatment policies for war victims belonging to cultural groups other than those of the planners and implementers (de Jong, 1995; Westermeyer, 1989). Western mental health workers who become familiar with the cultural practices of the groups with whom they are working are able to introduce traditional healing and rehabilitating practices into their program. For example, a group of adolescent boys from the Dingka of South Sudan wrote poems and composed songs about their traumatic experiences, and analyzed one another's dreams—all activities honored in that culture (de Jong, 1995). Unaccompanied Cambodian children brought to the USA were placed with families of the same ethnic origin and were encouraged to use Buddhist rituals to honor the dead, and were brought in contact with specially recruited Buddhist religious leaders (Duncan & Kang, 1985).

Culturally inappropriate interventions are common. For example, adolescent Somali males from a persecuted ethnic group were encouraged by Western mental health workers in refugee camps to "bring their experiences out in the open," and talk about the atrocities they had witnessed. This activity neither achieved the expected emotional catharsis nor did it promote a peaceful resolution of the ethnic conflict that led to the atrocities. On the contrary, because Somali society emphasizes collective responsibility over individual, the use of ventilation techniques increases the desire for revenge and gives it further legitimacy (Zarowsky, 1995). When adolescents in El Salvador were encouraged to talk about the death of their father (killed by the government) and to come to understand the circumstances of his death—a reasonable therapeutic strategy in many Western countries—local psychologists labeled this approach foolhardy. They pointed out that the government was likely to interpret this approach as a threat to its authority and to react aggressively against the mental health workers as well as their patients (Marsella, Bornemann, Ekblad, & Orley, 1994).

Western mental health workers may be more sensitive to cultural problems when working in the countries in which the armed conflict occurred. They may be more tempted to use "orthodox" Western approaches to trauma when they treat refugees who have come from these war-torn countries to Western countries of asylum. Urging their clients to talk at length about their traumatic experiences, for example, is ill-advised for refugees for whom denial and silence were effective survival strategies in their own country and are not easily put aside in their new country.

The use of a native interpreter is recommended not only in the country of origin whose language Western mental health workers have not mastered, but also in the country of asylum whose language the refugees have not mastered. The use of an interpreter is even recommended for initial interviews with refugees who have already mastered the language of the asylum country because certain terms possess different connotations in different languages and cultures. An interpreter is necessary, but not sufficient. The interpreter's presence may delude the psychologist into thinking that cultural barriers to therapeutic communication have been removed. In order to work around these barriers and take into consideration culturally bound attitudes toward pain, suffering, loss, and communication—with strangers in general and practitioners of a strange healing art in particular—it is necessary for mental health workers themselves to become knowledgeable about the cultural concepts, rituals, and amenities of their clients (Richman, 1993).

The diagnostic categories used in the stress field have been challenged as culture-laden. Eisenbruch (1992) has argued that the PTSD criteria are based on a Western ethnocentric view of health that prescribes how non-Western refugees should express stress and how their distress should be ameliorated. He suggested that the post-traumatic reaction of many cultural groups should be regarded as a normal and constructive existential response to their predicament rather than as a psychiatric disorder. Ong (1995) has argued that conventional mental health treatment invalidates refugee patients' cultural understanding of their own lives and thereby causes them harm. Most workers in the field do not take this extreme position and cautiously use pharmacological and psychotherapeutic treatments, and at the same time encourage the use of the treatments and rituals indigenous to the clients' cultural background.

9.24.4.3 Mental Health of Child Refugees

Children make up half of the world's refugee population. Some researchers consider them at greater risk for physical and psychiatric disability than adults for several reasons:

(i) They are developmentally more vulnerable to the adverse effects of malnutrition, on the one hand, and traumatic experiences, on the other.

(ii) They may experience cultural uprooting before the process of acculturation in their native culture is completed.

(iii) They often experience traumatic separations, temporary or permanent, from their parents, extended families, and familiar surroundings.

(iv) They are less able to fend for themselves and are more easily exploited by friend and foe alike.

(v) As a consequence, they may be implicated in the commission of violent acts, including murder, torture, and other atrocities.

(vi) They are correctly regarded as the future of their respective societies, as a consequence of which organizations with different ideological agendas may compete for their education and indoctrination (Goldson, 1993; Rousseau, 1995).

Respect for children's cultural background is not always consistent with Western democratic traditions. Some interventions may place Western mental health workers in a quandary: how to react, if at all, to an ideologically undemocratic tradition promoting mental health. For example, Centlivres (summarized in Rousseau, 1995, pp. 320–321) compared two kinds of school programs established in Afghan refugee camps. The United Nations schools provided the children with materials taken from a world far removed from their own. The traditional Moslem schools immersed the children in the military and political atmosphere of Afghan society. Not surprisingly, the students from the latter schools appeared more confident of themselves and their role in a future Afghanistan. The greater success of the latter schools must be seen in the context of the calamity that befell their country—one and a half million dead, hundreds of thousands injured, thousands of towns and villages destroyed, five million living as refugees in neighboring countries. The Moslem schools held out the promise of reconstruction of all aspects of Afghan life—physical, social, economic, and religious—and of the lives of these students in the process (Wardak, 1993). Western mental health workers may not be ideologically comfortable with this particular alternative. It may, however, be more conducive to the restored mental health of these children than what was offered by the United Nations program.

9.24.4.4 Victims of Torture

The complex character of torture is better appreciated when the act is defined comprehensively. The United Nations Convention Against Torture and Other Cruel, Inhuman, and Degrading Treatment or Punishment provided the following definition:

> Any act by which severe pain or suffering, whether physical or mental, is intentionally inflicted on a person for such purposes as obtaining from him or a third person information or confession, punishing him for an act he has committed or is suspected of having committed, or intimidating him or a third person, or for any reason based on discrimination of any kind, when such pain or suffering is inflicted by or at the instigation of or with the consent or acquiescence of a public official or other person acting in an official capacity. Torture does not include pain or suffering arising only from—inherent in or incidental to—lawful sanctions. (Vesti & Kastrup, 1995a, p. 214)

There are some key differences between torture and other traumatic events of human design: (i) the torturer has complete control over the victim; (ii) the victim is alone and has no recourse, legal or otherwise, to prevent the torturer from doing what he wishes; (iii) from the victim's perspective, what transpires in torture is wholly unpredictable; and (iv) torture is a more deliberate malevolent act with clearly defined goals than other acts of human malevolence. Its purpose may range from making the victim conform to the will of the torturer to destroying the humanity and personality of the victim. Torture may be used not only to obtain information but also to punish, guarantee the victim's future cooperation, forcibly indoctrinate the victim, and control an entire society by letting the public become aware of the use of torture.

9.24.4.5 Prevalence

It is estimated that at least 80 nations systematically employ torture and that between 5% to 15% of the world's refugees have suffered torture. Consequently, mental health workers are likely to encounter post-traumatic phenomena associated with torture in people who continue to live in their own country, notwithstanding their torture experience, as well as in refugees who migrated to Western countries of asylum (Vesti & Kastrup, 1995a).

9.24.4.6 Factors Affecting Adjustment

In order to prepare treatment programs for those seeking help, it is important to consider the major factors that determine the long-term effects of torture.

(i) *Parameters of the torture experience.* These include the nature and duration of the torture, intensity of the pain, the extent of permanent physical damage and dysfunction (sensory, motor, sexual), the inclusion of sexual torture and rape of male and female victims, and the psychological manipulation of the victim during torture and incarceration.

(ii) *Availability of intrapersonal resources during and after torture and incarceration.* These include a commitment to a political ideology, religious faith, the placing of the experience in some rational context, and the search for meaning in this painful experience.

(iii) *Immediate aftermath of torture*. Depending on the political climate of their country, some victims will experience social isolation and be stigmatized as subversives with whom it is dangerous to associate. Others may enjoy the support of family and friends, extended social networks, and access to rehabilitative agencies.

(iv) *Subsequent course of the torture victim's life*. Some achieve a degree of rehabilitation in their own country and others must seek asylum elsewhere with varying degrees of success.

(v) *Pretorture factors*. Prior positive adjustment and low levels of prior trauma exposure facilitate eventual recovery.

9.24.4.7 Sequelae of Torture

Many torture victims present avoidance and intrusion patterns consistent with torture-based PTSD. These symptoms lend themselves to the various learning therapies and to cognitive behavior modification therapies used to treat PTSD symptoms acquired in other kinds of traumatic situations (Dattilio & Freeman, 1994; Kanfer & Goldstein, 1991; Meichenbaum, 1994). Other torture victims may suddenly come to professional attention when they behave in a bizarre manner in reaction to innocuous situations. For example, some people have frightening flashbacks in reaction to standard medical examinations or encounters with people that are in some way similar to the original traumatic experiences. Still others experience brief psychotic episodes after taking anesthesia in preparation for surgery. Given knowledge of patients' backgrounds, these inexplicable reactions are seen as conditioned reactivation of feelings of terror associated with these earlier experiences. Reorienting the distraught person to people, time, and place is often sufficient to alleviate these acute symptoms. At the very least, knowledge of the history of torture prevents misdiagnosis and inappropriate treatment.

9.24.4.8 Treatment Guidelines

The adverse effects of torture do not disappear with time. The prevalence of PTSD-related symptoms as well as generalized anxiety and depression remain well above the control baseline. The willingness of torture victims to divulge their torture experiences to professional mental health workers varies widely. Many do not seek professional help because they do not believe they can be helped. Others are too ashamed of what was done to them or what they did under duress (e.g., implicated others, signed false confessions) to confide in others. Still others adjust by shutting down emotional expression altogether (psychic numbing), putting memories of these experiences out of mind, and assiduously avoiding the many situational cues that might activate the primal emotions associated with the torture trauma.

Some refugees may freely acknowledge torture in their initial contact with a mental health worker because evidence of torture endows people with certain rights by international law, not the least of which is the right not to be sent back to the country in which they were tortured. Many government officials are unfamiliar with the widespread character of torture and are inclined to reject these accounts out of hand. Mental health workers with a background in this field are able to provide expert opinion in support of the victims' claims, to prepare them to give credible testimony on these painful matters, and to refer them to legal representatives. They are also able to help these people to come to terms with their torture experiences and to deal more effectively with some of the intrapersonal and interpersonal fallout of torture.

Some refugees may ask for symptomatic relief of their distress or for help in becoming better adjusted, without mentioning their torture experiences. Interviewers should be aware of this possibility and gently follow up leads, but should not press the issue. When clients have developed trust in the therapist and comfort in the relationship, they will divulge this experience on their own.

Sexual torture and/or rape is more traumatizing than other forms of torture because of the special nature of sexual violation as perceived by torturer and victim alike, and because of the consequences of divulging this information to family and friends. Raped women are treated as damaged goods in some societies, divorced by their husbands, ostracized by their families, and in some cases even killed or forced into suicide. Some raped women commit suicide on their own. Consequently, these revelations are subject to stringent confidentiality; otherwise, the structure of an entire family could be destroyed.

9.24.4.9 Treatment Centers

Vesti and Kastrup (1995b) describe a multimodal treatment program for torture victims that involves concurrent physical and psychological therapy, involvement of the victim's family, and establishing a therapeutic milieu that contains few of the physical cues and situations that might remind victims of their torture experiences (e.g., lengthy waiting time, abrupt noise, sharp light shining directly in one's face). Physiotherapy and orthopedic rehabilitation are mandatory for injured limbs, medication is prescribed for reduction of anxiety or depression, and psychotherapy for psychic injury and behavioral dysfunction. Vesti and Kastrup cited numerous contraindications to insight-oriented psychotherapy (the presence of a paranoid or antisocial personality disorder, weak ego boundaries and poor reality testing, severe substance abuse).

The treatment of torture victims is so new and the range of treatment modalities required by these people is so broad that professional treatment should only be done in or in conjunction with a recognized center. Clinical psychologists who work with refugees, but have little or no clinical experience in treating torture victims, should be prepared to raise the question of referral tactfully with refugees whose symptoms are not consistent with the account they gave of their preimmigration experiences. Clients who acknowledge having been tortured should be encouraged, but not pressured, to seek help in a treatment center.

Inexperienced therapists may harm their clients by applying inappropriate treatment principles (e.g., pressuring the torture victim to tell all). Inexperienced therapists are also likely to develop strong countertransference reactions to torture victims. Therapists who react with horror on the one hand and pity on the other can offer little to the rehabilitation of torture victims. Therapists who are uncomfortable with the topic and discourage clients from talking about it are no more helpful than therapists who display a morbid interest in the details *per se*, hence the importance of referring these clients to people trained in this complex problem.

9.24.5 THE EFFECT OF WAR ON CIVILIAN POPULATIONS IN INTACT SOCIETIES

9.24.5.1 The USA and the Gulf War

How American society was affected by the mobilization and movement of half a million military service men and women in Operation Desert Storm (1990–1991) is a question that is being evaluated and will continue to be evaluated for many years. The USA was not bombed or invaded, but the lives of the families and friends of the men and women who liberated Kuwait and forced the withdrawal of the Iraqi army were undoubtedly affected by the absence of their loved ones, concern for their safety, and the changed behavior of some military service personnel on their return. Tens of thousands complained in recent years about frequent headaches, stiffness of joints, nausea, and mood changes. Evidence is now emerging that these soldiers may have been exposed to injurious chemicals due to the allied bombing of an Iraqi ammunition storage depot (NIH Technology Assessment Workshop Panel, 1994; Persian Gulf Veterans Coordinating Board, 1995).

A task force on war-related stress was convened immediately after the cessation of hostilities, to develop strategies for the prevention and treatment of psychological, psychosocial, and psychosomatic disorders in the military and in the civilian population 6000 miles away from the hostilities (Hobfoll et al., 1991). Their report highlighted the major assumptions held by researchers and interventionists on war-related stress in well-functioning societies that were not devastated by war. Their list of groups at risk for war-related distress and behavioral impairment contained some surprising choices: adults with prior histories of traumatic experiences, adults with prior histories of prewar psychopathology, and children in general.

Several implications follow from these inclusions.

(i) Media coverage of the Gulf war and contact with people directly affected by the Gulf war may have been upsetting to vulnerable groups.

(ii) Vulnerable groups can be inoculated against stressful encounters by means of primary and secondary intervention directed toward the general public and toward professional workers (e.g., primary care physicians, psychiatrists, clergy, police) who come in contact with vulnerable people.

(iii) Children may be especially vulnerable because, like adults, they are exposed to frightening images and accounts in the media but, unlike adults, they lack the cognitive ability and background information to place what they see and hear in a nonthreatening context.

For these reasons, the task force made a number of recommendations to meet the special needs of children. Children should be given accurate information about the war that is appropriate to their developmental level. They should not be protected from potentially upsetting information, nor should they be asked to shoulder their parents' burden during family crises pertaining to the war. Children merit an attentive, nonjudgmental adult ear if they elect to talk about their war-related concerns. Children's reactions in war as in any other stressful situation tend to mirror the reactions of their parents and other significant adults. The best thing adults can do to help children cope with a stressful situation is to behave themselves in a responsible, reassuring manner, despite their own apprehensions. Children, like adults, should be provided with positive expectancies, that they will be able to cope effectively with whatever is bothering them. Children in general, but especially older children, should be recruited to perform useful tasks that are relevant to the war.

9.24.5.2 Israel and the Gulf War

Israel was far more strongly affected by the Gulf war than the USA. If the casualties and damage incurred by 39 missiles launched from Iraq against Israel are translated into American terms (the ratio of population in the two countries is 50:1), the comparable losses in American terms would be 450 dead from direct causes and 3000 dead from indirect causes, over 11 000 injured, and 16 000 with severe stress reactions, and half a million dwellings damaged or destroyed (Bleich, Dycian, Koslowsky, Solomon, & Wiener, 1992). Over 14 000 people left the country just before the war broke out and over 100 000 left their homes in the targeted metropolitan areas for safer havens in the countryside.

Studies of public morale and anxiety were consistent. There was high anxiety and even panic after the first few missiles, a leveling off thereafter, and a gradual decline as the war drew to a close. Citizens living in the vicinity of direct hits were far more apprehensive than those living elsewhere. There was a maximal use of hot-lines for many purposes, including how to don the gas mask, what to tell the children, how to handle one's own anxiety. The media recruited large numbers of psychologists to provide answers to practical questions, and subsequent polls of public satisfaction with these talk shows showed that it was high (Milgram, 1994). Many mental health workers spoke to teachers and members of other professional groups that came in contact with the general public. There was professional criticism about the exaggerated emphasis of many psychologists on ventilation and on maintaining emotional control and an insufficient emphasis on problem-solving behaviors during stressful periods (Milgram, 1993).

Some unique features of this war induced a collective sense of learned helpless that affected individuals during the war and the national outlook after the war. Death, injury, and damage to property were all confined to the civilian sector. The Israel defense forces were not mobilized and the protection of the citizens was placed in the hands of American crews unsuccessfully trying to shoot Scud missiles out the sky with Patriot missiles. Adults in general and men in particular were unable to play their customary protective wartime roles and were correspondingly upset. All family members sat in sealed rooms during air raids wearing their gas masks, but were unable to do anything about the air raids themselves. The end of the war did not dispel uncertainty about future missile attacks and their potential payload (chemical or biological agents) and uncertainty about the ability of the Israeli defense force to predict and avert such attacks in the future. This uncertainty affected people's sense of well being, their attitudes toward their collective and personal security, and their willingness to make far-reaching concessions to their Arab neighbors in the hope of achieving peace (Milgram, 1994).

9.24.6 LONG-TERM ADJUSTMENT OF TRAUMATIZED HOSTAGES

The taking of hostages by ideologically driven terrorists has escalated in the latter half of the twentieth century. Ordinary citizens are seized without warning by terrorists who agree to release them unharmed if their government meets the demands of the terrorists (e.g., release their fellow terrorists in prison and grant them safe passage out of the country). Hostages are typically taken in Western countries that place a high value on human life and may be responsive to terrorists' demands. When hostages are released, they return home after a perfunctory debriefing to resume their normal undertakings. Whether they are able to resume their normal routines without suffering the chronic effects of PTSD symptoms and related interpersonal and intrapersonal difficulties is questionable (Ochberg & Soskis, 1982). The extent of their subsequent psychopathology depends on the stressful parameters of the hostage experience. When the experience involves long periods of deprivation, threats to and the injury and death of hostages, it may change a person's perceptions, feelings, cognitions, motivations, and corresponding behaviors permanently. Consider the consequences of the comparatively brief hostage situation described below.

The account of the most famous hostage-taking episode in Israel provides a perspective on the long-term effects of a brief, terrifying hostage episode and the consequences of not providing comprehensive, long-term, multimodal treatment to liberated hostages. Over 100 Israeli high school students on a school outing away from home were seized and held hostage by three Palestinian terrorists in western Galilee in May, 1974. The terrorists had already killed a man, his pregnant wife, and their four-year-old child before coming upon the students. When the terrorists first entered the school building where the students were staying, nearly 20 escaped, including all of the responsible supervising adults. The remaining students found themselves abandoned by adult authority figures and threatened repeatedly by the terrorists for the next 16 hours. The terrorists mined the building with explosives and fired live ammunition between the legs of the hostages. During the rescue operation the terrorists killed 22 and injured more than 50 hostages. Seventeen years later, 59 survivors participated in a comprehensive follow-up study, along with a carefully selected control group (Gal, Syna-Desivilya, & Ayalon, 1994).

The vast majority had married, were raising families, and were gainfully employed. Their parenting skills and family adjustment were comparable to the levels of the control group. All of the physically capable men served in the Israel defense forces. Many survivors stated that the experience made them appreciate life more, enhanced their ability to cope with subsequent adversity, and their desire to help other people.

However, there was evidence of nonspecific maladjustment as well as situation-specific chronic symptoms. Nearly half said that the tragedy had disrupted their lives, caused them great mental distress, and impeded their academic and career development. They reported more psychiatric symptoms than controls. They also reported a high frequency of PTSD symptoms reflecting features of the original traumatic experience. Those who suffered the greatest degree of victimization (more severe physical injury and greater exposure to traumatic events) were the most disturbed. As might be expected, the survivors were more apprehensive about being abandoned and had less trust in people than controls. After all, they had been abandoned by others during the crisis and many felt abandoned during their long recovery when many well-meaning people came briefly and then disappeared. This lack of trust in people may also account for their lower level of intimacy in marriage. That most of the men avoided doing their military service in combat units may be attributed to their heightened sensitivity to weapons and the sounds of weapons, a sensitivity acquired during the hostage episode. Most reported anxiety and fear reactions during the Gulf war because of the uncomfortable associations that the war elicited.

Analysis of the kind and number of psychotherapy sessions given to the survivors indicated that there was no sustained effort by the mental health profession to provide systematic, long-term therapeutic intervention for them. Less than half received any psychotherapy immediately after the episode and, in most cases, it amounted to one or two sessions. Those who did receive psychotherapy indicated that it was ineffective.

Looking back today at their needs then, it is certain that had these people been given a broad range of learning and behavior modification treatments soon after their harrowing experiences—treatments not available in 1974—they would be suffering fewer and less intense PTSD symptoms and related adjustment problems today. People who have been exposed to life-threatening situations, such as being taken hostage, and were fortunate enough to survive have suffered psychic injuries that justify their being offered a wide range of treatments—pharmacological, educational, behavioral, and psychotherapeutic. A group debriefing experience shortly after being freed from the hostage episode is a quick, but insufficient, fix.

9.24.7 PROFESSIONAL RESPONSIBILITIES OF CLINICAL PSYCHOLOGISTS

9.24.7.1 Ten Precautions

Working with war victims, at any intervention level, is a highly complex and demanding professional endeavor. In undertaking to work with war victims, one should be mindful of Hippocrates's admonition: First, do no harm. There are many reasons why psychologists with the best of intentions may not help and even cause harm to these kinds of clients.

9.24.7.1.1 Problem of empathy with victims of human malevolence

Most psychologists living in Western countries have encountered the kinds of war-time suffering experienced by Asians and Africans only in books and in the cinema. They are no more equipped to believe or to empathize with the litany of malevolence and degradation told by a solitary refugee than they are to deal with the sight of 30 000 injured, traumatized, and terrified survivors of the Rwanda massacres, mostly women and children, living in a crowded refugee camp on the border of an indifferent or hostile country.

The mental health professional's initial reaction to the tales of the Holocaust survivors was denial of the message and rejection of the messenger. To overcome these understandable but antitherapeutic reactions it was necessary not only to acquire knowledge about the Holocaust but also to seek supervision from and collaboration with more experienced colleagues, if one hoped to conduct effective psychotherapy with Holocaust survivors (Danieli, 1981). The same cautions apply to war victims today.

9.24.7.1.2 Lack of concern for the client's cultural background

Many psychologists are ignorant of or indifferent to the cultural background of many clients. Some psychologists may simply disregard or reject it out of hand. Cultural imperialism has its adherents and champions in many Western countries and in many walks of life. Clinical psychologists who hope to do effective work with war victims from other cultures must learn and listen first.

9.24.7.1.3 Misapplication of the psychodynamic orientation to traumatized clients

Many clinical psychologists are wedded to a psychodynamic orientation to psychopathology and regard all current symptomatology as stemming from unresolved intrapsychic conflicts associated with the circumstances of their home and upbringing. War victims also carry with them the unresolved conflicts and vulnerabilities of childhood. They have been exposed, however, to extraordinary stressors in war that adversely affect all prior acquisitions, however successful.

9.24.7.1.4 Nonmatching and inappropriate attributions

When clients enter therapy they may attribute their personal difficulties and distress to their own failings and mistakes (an internal orientation) or attribute them to the failings, mistakes, and malevolence of others or to fate (an external orientation). They may also attribute responsibility for resolving their difficulties and achieving mental health to their own efforts (internal) or to the efforts of the therapist or changes in the behavior of significant others (external). The combinations of internal vs. external attributional proclivity for problems and their resolution introduce complex considerations in psychotherapy (Brickman et al., 1982). That therapists also bring their own assumptions about proper attributions to their therapeutic endeavors adds further complications. Many therapists foster an internal orientation in their clients for the circumstances that brought them into treatment and most foster an internal orientation for what must be done to bring about desired change. Milgram (1986) has shown that if the client possesses one attributional style and the therapist another, the lack of correspondence can be a serious obstacle to therapeutic progress. This already complex problem becomes more complex when Western therapists deal with refugees from non-Western countries.

It is inappropriate in most instances for therapists to encourage their clients' adopting an internal orientation for symptomatology and personal problems that stem from war and emigration. It may even be inappropriate for therapists to encourage their refugee clients' adopting an internal orientation for solving their problems. Clients who have been reared in the Western tradition of individualism and personal responsibility for one's successes or failures are more willing to adopt an internal orientation for resolving their difficulties. People reared in a more fatalistic tradition and who may attribute what happens to them as a reward or punishment for their actions in earlier reincarnations (e.g., Buddhist tradition) require a different attributional approach. There is no easy answer here, but it is imperative that therapists ascertain what are the attributional orientations of their clients toward their personal difficulties and the resolutions of these difficulties as they plan their therapeutic approach.

9.24.7.1.5 Professional bias against war victims

Clinical psychologists, like other citizens, have distinct ideological preferences. It follows that clinical psychologists may find themselves treating refugees with whom they do not necessarily empathize. The professional behavior required of the clinical psychologist is to put aside ideological issues and allegiances and treat the client to the best of one's ability. If this is not possible, the psychologist must make an appropriate referral (Milgram & Toubiana, 1988).

9.24.7.1.6 Discomfort with expressions of revenge

Many therapists are uncomfortable about expressions of retribution, retaliation, or revenge voiced by clients who have suffered physical and psychological injury at the hands of others. One can easily understand that many victims of the human malevolence that characterizes war may talk about these desires. However, many therapists regard talk about revenge as unworthy of their clients and forcefully discourage such talk. They may even suggest that their clients forgive and forget. This discomfort with the desire for restitution, retaliation, or revenge runs counter to the prevailing view that any human impulse or thought can be bared in therapy without fear of censure. It also runs counter to the moral outrage experienced by people when guilty parties go unpunished and the moral satisfaction when they are punished. Therapists must carefully think through their attitude toward the revenge theme when it arises in war victims and use it as a channel for enhanced communication and personal growth.

9.24.7.1.7 Reluctance to use lay support groups in treatment

Many clinical psychologists prefer to see their clients in individual or in group therapy, but discourage their involvement in self-help groups run by lay leaders or in other lay support groups. This preference has its roots in the desire to keep the client centered on the therapeutic relationship and to discourage competing encounters that might dilute the relationship or reduce symptomatic distress at the expense of acquiring insight into one's difficulties. Community interventionists, however, encourage clients to receive support from significant lay people (Nietzel, Guthrie, & Susman, 1991). They justify the use of family, friends, and self-help groups to enhance empowerment, to facilitate change in clients who live and work in an open system that must change to accommodate itself to changes in their clients. They assert that it is easier to instigate and maintain change in a client when there are environmental supports. Moreover, if significant others are not involved, they may derail treatment out of resentment, ideological conflict, or any other reason. A comprehensive presentation of the many issues inherent in primary and secondary intervention in disasters may be found in Hobfoll and deVries (1995).

9.24.7.1.8 Reluctance to accept the role of the outsider

Clinical psychologists who elect to work with people living in war-devastated societies (e.g., in refugee camps, or in the countries of origin) are likely to feel that they are outsiders looking in. This is understandable and essentially correct. They are outsiders, but do have professional skills and knowledge to impart to the insider leadership. They serve as consultants with the local leadership, offer to train them in skills relevant to the needs of the local populations, and maintain a watchful brief (Caplan, 1974). Clinical psychologists do not become insiders when they work in their own country and treat refugees coming in for treatment. They are still outsiders to the culture, language system, and the experiences of these people. They should acknowledge and accept their outsider role in good grace on their own home turf as in overseas assignments.

9.24.7.1.9 Caution against professional burnout

The potential for burnout or what Figley (1994) has termed compassion fatigue, is very great in psychologists who immerse themselves for long periods of time, without respite, in their work with victims of war. This is especially true when psychologists work in refugee camps in Third World countries, but it may also occur when they work in their own countries. It is imperative to adopt the various work and life strategies that mitigate against burnout.

9.24.7.1.10 Danger of being overly cautious

Some therapists may become overly cautious in their work with war victims. They may take to heart the caveats summarized above and adopt a passive, waiting orientation towards their clients. Western clients often welcome this orientation as a clear indication that therapy is their responsibility to make of it what they may. This orientation would be unfortunate in working with many war victims from non-Western cultures because such people expect their healers to take positive action. Therapists who have a caring and respectful attitude toward their clients can take small risks in therapy if they have made their clients their partners in the enterprise. It is their partner's responsibility to provide feedback about the therapists' overtures.

9.24.7.2 Responsibilities of Clinical Psychologists

Western-trained clinical psychologists have important responsibilities to perform with war victims. The psychologist is the major source of knowledge about assessment. Clinical psychologists are expert in the construction, validation, administration, and interpretation of assessment scales and interviews. Space does not permit citing the many scales that assess affect, adjustment, general symptoms of maladjustment, and specific scales of symptoms associated with exposure to traumatic episodes. Clinical psychologists are responsible for selecting the appropriate assessment techniques to use with a given population in a given situation, for adapting existing scales to the requirements of special groups, for constructing new scales for new groups, new situations, or new features of known situations, and for validating these scales.

Clinical psychology shares with other professions the responsibility to intervene therapeutically on behalf of clients traumatized or otherwise impaired by their war experiences. Clinical psychologists with a background in community psychology will find the transition to primary and secondary intervention easier than those trained exclusively in one-to-one or group therapy. If clinical psychologists have cultivated a strong critical faculty about their work, they may be able to assess more quickly and correctly what is right and what is not in their work with these clients.

Psychologists are involved in the training of professional and lay groups. These groups need to deepen their understanding of the many experiences, problems, and needs of war victims. Psychologists can make a major contribution by writing, conducting workshops, and consulting with these interventionist groups.

The training of clinical psychologists in empirical research makes them critical consumers of research findings and of authoritative conclusions promulgated by experts in the field. It also qualifies them to engage in correlational clinical research as well as in conducting clinical trials on the efficacy of interventions.

Clinical psychology accepts an advocacy function. War victims are at high risk for chronic psychopathology if needed interventions are not provided. Many of the necessary interventions are dependent on political decisions and support. Politicians with the best of intentions and with generous sums of money may accomplish little good and do much harm, if not given objective information about given problems and feedback about the efficacy of treatments to solve these problems.

The training and talents of many clinical psychologists qualify them to analyze organizational endeavors in an effort to diagnose those variables that contribute to achieving the professed goals of the organizations and the variables that detract from achieving these goals, and to convey their findings and conclusions to others. The numbers and needs of war victims are staggering. To meet big needs, it is necessary to launch big programs. By their nature big programs can make big mistakes. Trouble shooting and whistle blowing are heavy responsibilities and fraught with risks for shooters and blowers. No one person can get the job done right, but no one should be exempt from trying.

9.24.8 REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Antonovsky, A. (1979). *Health, stress, and coping.* San Francisco: Jossey-Bass.

Arieli, A., & Aycheh, S. (1992). Psychopathology among Jewish Ethiopian immigrants to Israel. *Journal of Nervous and Mental Disease, 180,* 465–466.

Bleich, A., Dycian, A., Koslowsky, M., Solomon, Z., & Wiener, M. (1992). Psychiatric implications of missile attacks on a civilian population. *Journal of the American Medical Association, 268,* 613–615.

Brickman, P., Rabinowitz, V. C., Karuza, J., Coates, D., Cohn, E., & Kidder, L. (1982). Models of helping and coping. *American Psychologist, 37,* 368–384.

Caplan, G. (1974). *Support systems and community mental health.* New York: Human Sciences.

Danieli, Y. (1981). Families of survivors of the Nazi Holocaust: Short and long-term effects. In C. S. Spielberger, I. W. Sarason, & N. A. Milgram (Eds.), *Stress and anxiety* (Vol. 8, pp. 405–421). Washington, DC: Hemisphere.

Danieli, Y. (1993). Diagnostic and therapeutic use of the multigenerational family tree in working with survivors and children of survivors of the Nazi Holocaust. In J. P. Wilson & B. Raphael (Eds.), *The international handbook of traumatic stress syndromes* (pp. 889–898). New York: Plenum.

Dattilio, F. M., & Freeman, A. (1994). *Cognitive-behavioral strategies in crisis intervention.* New York: Guilford.

de Jong, J. T. V. M. (1995). Prevention of the consequences of manmade or natural disaster at the inter(national), the community, the family, and the individual level. In S. E. Hobfoll & M. W. deVries (Eds.), *Extreme stress and communities: Impact and intervention* (pp. 207–228). Dordrecht, The Netherlands: Kluwer.

Duncan, J., & Kang, S. (1985). *Using Buddhist ritual activities as foundation for a mental health program for Cambodian children in foster care.* Unpublished manuscript. Mountainlake Terrace, WA: Lutheran Social Services.

Eisenbruch, M. (1992). Toward a culturally sensitive *DSM*: Cultural bereavement in Cambodian refugees and the traditional healer as taxonomist. *Journal of Nervous and Mental Disease, 180,* 8–10.

Figley, C. R. (Ed.) (1994). *Compassion fatigue: Secondary traumatic stress from treating the traumatized.* New York: Brunner/Mazel.

Gal, R., Syna-Desivilya, H., & Ayalon, O. (1994). *Coping and long-term adjustment: The case of the Ma'alot Survivors.* Zichron Ya'akov, Israel: The Carmel Institute for Social Studies,

Goldson, E. (1993). War is not good for children. In L. A. Leavitt & N. A. Fox (Eds.), *The psychological effects of war and violence on children* (pp. 3–22). Hillsdale, NJ: Erlbaum.

Harel, Z., Kahana, B., & Kahana, E. (1993). Social resources and the mental health of aging Nazi Holocaust survivors and immigrants. In J. P. Wilson & B. Raphael (Eds.), *The international handbook of traumatic stress syndromes* (pp. 241–252). New York: Plenum.

Hauff, E., & Vaglum, P. (1995). Organized violence and the stress of exile: Predictors of mental health in a community cohort of Vietnamese refugees three years after resettlement. *British Journal of Psychiatry, 166,* 360–367.

Herman, J. L. (1992). *Trauma and recovery.* New York: Basic Books.

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist, 44,* 513–524.

Hobfoll, S. E., & de Vries, M. W. (Eds.) (1995). *Extreme stress and communities: Impact and intervention.* Dordrecht, The Netherlands: Kluwer.

Hobfoll, S. S., Spielberger, C. D., Breznitz, B., Figley, C., Folkman, S., Lepper-Green, B., Meichenbaum, D., Milgram, N. A., Sandler, I., & van der Kolk, B. (1991). War-related stress: Addressing the stress of war and other traumatic events. *American Psychologist, 46,* 848–855.

International Federation of Red Cross and Red Crescent

Societies (1995). *World Disaster Report 1995*. Geneva, Switzerland: Author.

Janoff-Bulman, R. (1988). Victims of violence. In S. Fisher & J. Reason (Eds.), *Handbook of life stress, cognition and health* (pp. 101–113). New York: Wiley.

Jerusalem, M., Kaniasty, K., Lehman, D. R., Ritter, C. & Turnbull, G. (1995). Individual and community stress: Integration of approaches at different levels. In S. E. Hobfoll & M. W. deVries (Eds.), *Extreme stress and communities: Impact and intervention* (pp. 105–130). Dordrecht, The Netherlands: Kluwer.

Kaplan, H. I., & Sadock, B. J. (1985). *Comprehensive textbook of psychiatry* (Vol. IV). Baltimore: Williams and Wilkins.

Kanfer, F. H., & Goldstein, A. P. (1991). *Helping people change: A textbook of methods* (2nd ed.). Boston: Allyn and Bacon.

Kinzie, J. D. (1993). Posttraumatic effects and their treatment among Southeast Asian refugees. In J. P. Wilson & B. Raphael (Eds.), *The international handbook of traumatic stress syndromes* (pp. 311–319). New York: Plenum.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.

Lomranz, J. (1995). Endurance and living: Long-terms effects of the Halocaust. In S. E. Hobfoll & M. W. deVries (Eds.), *Extreme stress and communities: Impact and intervention* (pp. 325–352). Dordrecht, The Netherlands: Kluwer.

Marsella, A. J., Bornemann, T., Ekblad, S., & Orley, J. (1994). *Amidst peril and pain: The mental health and well being of the world's refugees*. Washington, DC: American Psychological Association.

McIntyre, T. M., & Ventura, M. (1996). Civil war haunts Angola children. *Psychology International, 7,* 1–4.

Meichenbaum, D. (1994). *A clinical handbook of PTSD: Practical therapist manual for assessing and treating adults with PTSD*. Watertown, Canada: Institute Press.

Milgram, N. A. (1986). Attributional analysis of war-related stress: Modes of coping and helping. In N. A. Milgram (Ed.), *Stress and coping in time of war: Generalizations from the Israeli experience* (pp. 9–25). New York: Brunner/Mazel.

Milgram, N. (1993). Stress and coping in Israel during the Gulf War. *Journal of Social Issues, 49,* 103–123.

Milgram, N. (1994). Israel and the Gulf War: The major events and selected studies. *Anxiety, Stress, and Coping, 7,* 205–215.

Milgram, N. A. (1998). Children under stress. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of child psychopathology* (pp. 505–536). New York: Plenum.

Milgram, N. A., & Hobfoll, S. E. (1986). Generalizations from theory and practice to war-related stress. In N. A. Milgram (Ed.), *Stress and coping in time of war: Generalizations from the Israeli experience* (pp. 316–352). New York: Brunner/Mazel.

Milgram, N. A., Sarason, B., Schonplug, U., Jackson, A., & Schwarzer, C. (1995). Catalyzing social support. In S. E. Hobfoll & M. W. deVries (Eds.), *Extreme stress and communities: Impact and intervention* (pp. 473–488). Dordrecht, The Netherlands: Kluwer.

Milgram, N. A., & Toubiana, Y. H. (1988). Bias in identifying and treating high-risk groups. *Professional Psychology: Research and Practice, 19,* 21–25.

Mollica, R. F. (1988). The trauma story: The psychiatric care of refugee survivors of violence and torture. In F. M. Ochberg (Ed.), *Post-traumatic therapy and victims of violence* (pp. 295–314). New York: Brunner/Mazel.

Nietzel, M. T., Guthrie, P. R., & Susman, D. T. (1991). Utilization of Community and Social Support Resources. In F. H. Kanfer & A. P. Goldstein (Eds.), *Helping people change: A textbook of methods*. (2nd ed., pp. 396–421). Boston: Allyn & Bacon.

NIH Technology Assessment Workshop Panel (1994). The Persian Gulf experience and health. *Journal of the American Medical Association, 272,* 391–396.

Ochberg, F. M., & Soskis, D. A. (1982). *Victims of terrorism*. Boulder, CO: Westview.

Ong, A. (1995). Making the biopolitical subject: Cambodian immigrants, refugee medicine, and cultural leadership in California. *Social Science and Medicine, 40,* 1243–1257.

Persian Gulf Veterans Coordinating Board (1995). Unexplained illnesses among Desert Storm veterans: A search for causes, treatment, and cooperation. *Archives of Internal Medicine, 155,* 262–268.

Richman, N. (1993). Children in situations of political violence. *Journal of Child Psychology and Psychiatry, 34,* 1286–1302.

Rousseau, C. (1995). The mental health of refugee children. *Transcultural Psychiatric Research Review, 32,* 299–332.

Seligman, M. E. P. (1975). *Helplessness: On depression, development and death*. San Francisco: Freeman.

Somasundaram, F. J. (1993). Psychiatric morbidity due to war in Northern Sri Lanka. In J. P. Wilson & B. Raphael (Eds.), *The international handbook of traumatic stress syndromes* (pp. 333–348). New York: Plenum.

Taylor, S. E., & Lobel, H. (1989). Social comparison activity under threat: Downward evaluations and upward contacts. *Psychological Review, 96,* 569–575.

US Committee on Prevention of Mental Disorders (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington, DC: National Academy Press.

van der Kolk, B., Roth, S., Pelcovitz, D., & Mandel, F. (1993). *Complex PTSD: Results of the DSM-IV Field Trial for PTSD*. Unpublished paper. Cambridge, MA: Harvard Medical School.

Vesti, P., & Kastrup, M. (1995a). Refugee status, torture, and adjustment. In J. R. Freedy & S. E. Hobfoll (Eds.), *Traumatic stress: From theory to practice* (pp. 213–235). New York: Plenum.

Vesti, P., & Kastrup, M. (1995b). Treatment of torture survivors: Psychosocial and somatic aspects. In J. R. Freedy & S. E. Hobfoll (Eds.), *Traumatic stress: From theory to practice* (pp. 339–363). New York: Plenum.

Wardak, A. W. H. (1993). The psychiatric effects of war stress on Afghanistan society. In J. P. Wilson & B. Raphael (Eds.), *The international handbook of traumatic stress syndromes* (pp. 349–364). New York: Plenum.

Weisath, L., & Eitinger, L. (1993). Posttraumatic stress phenomena: Common themes across wars, disasters, and traumatic events. In J. P. Wilson & B. Raphael (Eds.), *The international handbook of traumatic stress syndromes* (pp. 69–77). New York: Plenum.

Westermeyer, J. (1989). Cross-cultural care for PTSD: Research, training and service needs for the future. *Journal of Traumatic Stress, 2,* 515–536.

Zarowsky, C. (1995). Resources for a critical synthetic approach to refugee mental health. *Transcultural Psychiatric Research Review, 32,* 333–338.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.25
Adult Sexual Offenders

WILLIAM L. MARSHALL
Queen's University, Kingston, ON, Canada

9.25.1 INTRODUCTION 407
9.25.2 NATURE AND PREVALENCE OF THE OFFENSES 408
9.25.2.1 Exhibitionism 408
9.25.2.2 Voyeurism 408
9.25.2.3 Frotteurism 408
9.25.2.4 Child Molestation 408
9.25.2.5 Sexual Assaults 409
9.25.2.6 Summary of Prevalence 409
9.25.3 DIAGNOSIS AND ASSESSMENT 409
9.25.3.1 Diagnosis 409
9.25.3.2 Assessment 410
9.25.3.2.1 Sexual interests/history 410
9.25.3.2.2 Social functioning 411
9.25.3.2.3 Cognitive processes 412
9.25.3.2.4 Personality 412
9.25.3.2.5 Substance abuse 412
9.25.4 TREATMENT 413
9.25.4.1 Cognitive-behavioral Programs 413
9.25.4.1.1 Acceptance of responsibility 413
9.25.4.1.2 Cognitive distortions 413
9.25.4.1.3 Empathy 413
9.25.4.1.4 Social functioning 414
9.25.4.1.5 Deviant fantasies 414
9.25.4.1.6 Relapse prevention 415
9.25.4.2 Overall Value of Treatment 415
9.25.5 FUTURE DIRECTIONS 415
9.25.5.1 Research 416
9.25.5.2 Practice 417
9.25.6 SUMMARY 417
9.25.7 REFERENCES 418

9.25.1 INTRODUCTION

The legal and diagnostic systems differ in their identification of sexual offenders, with the legal system being far more inclusive. The American Psychiatric Association's (1994) *Diagnostic and statistical manual of mental disorders* (*DSM-IV*) lists under the category of "Paraphilias" a limited set of behaviors that are legally proscribed (i.e., exhibitionism, voyeurism, frotteurism, and pedophilia). *DSM-IV* also lists under "Sexual disorders not otherwise specified" additional behaviors that are illegal: necrophilia (sex with a corpse), bestiality (sex with animals), and scatologia (obscene phone calls). In addition, the category of sadism may involve behaviors that break the law (where the putative masochistic partner is an unwilling participant). Unfortunately, many sexual offenders do not meet the diagnostic criteria for paraphilia and this presents a dilemma for clinicians who typically see a need to treat all sexual offenders whether or not they can be given a diagnosis. Worse than that, this conflict between legal proscriptions and diagnostic criteria presents serious problems for researchers, and particularly for reviewers attempting to assimilate the literature, in identifying the population samples. In this chapter the behaviors considered are described in generic terms as follows: exhibitionism, voyeurism, frotteurism, child molestation, and sexual assault (i.e., sexual attacks on adults). In this respect, sexual offenses are construed as sexual (or partly sexually motivated) acts perpetrated on an unwilling victim or on a person who is considered unable to give consent (e.g., a child or a severely disabled person).

9.25.2 NATURE AND PREVALENCE OF THE OFFENSES

There is so little known about some sexual offenses (e.g., obscene telephone calls, bestiality, and necrophilia) that it makes little sense to discuss them. Consequently, the focus is on only those for which there is sufficient available evidence.

9.25.2.1 Exhibitionism

Exhibitionism involves the deliberate and unsolicited exposure of the genitals to an unwilling audience. It is the most common of all sexual offenses, accounting for over one-third of all convictions for sexual crimes in the USA, Canada, England, Germany, and Hong Kong (Rooth, 1973). Abel et al. (1987) reported that exhibitionists attending their clinic admitted to an average of 514 acts per offender, and Marshall, Eccles, and Barbaree (1991) found that 57% of untreated exhibitionists reoffended within four years of being convicted. Up to 50% of adult women report having been the victim of an exhibitionist at least once in their life (Di Vasto et al., 1984). Exposing is obviously a very persistent and frequent behavior among these offenders.

While there are some reports of female exhibitionists, almost all of those who are identified by the legal system, and who receive clinical attention, are males (Blair & Lanyon, 1981). Although some men may expose themselves to children, those who do so on a consistent basis are generally construed by clinicians as child molesters rather than exhibitionists. It appears that the majority of exhibitionists remain just that, although the boldness of their acts typically increases over time as a result of desensitization to the risk of being caught (victims rarely report them). Some exhibitionists, however, do engage in other offensive behaviors. Freund (1990), for example, found that exhibitionists are also quite commonly voyeurs and that 12% had also raped a woman.

9.25.2.2 Voyeurism

Voyeurism is the act of peeping in windows for the purpose of watching unsuspecting persons (usually women) who are undressing, already naked, or who are engaging in sexual acts. Again, it is mostly males who are identified as voyeurs. Abel, Becker, Cunningham-Rathner, Mittelman, and Rouleau (1988) found that voyeurs often display other sexually offensive behaviors, and Langevin, Paitich, and Russon (1985) found that every one of their sample of over 600 voyeurs had engaged in other sexual offenses, including rape. The actual frequency of voyeurism, however, is not well documented.

9.25.2.3 Frotteurism

Frotteurism refers to touching or rubbing against an unwilling person in order to generate sexual excitement. The frotteur typically frequents crowded places so that he can engage in his offensive behavior. Some attempt to make their frotteurism appear accidental, while others are so emboldened that they do not appear to care what an observer might think. These latter frotteurs engage in such aggressive and intrusive behaviors that Langevin (1983) has considered them to belong to the same category as rapists. Again, the frequency of these behaviors is not well documented as frotteurs have rarely been identified in sufficient numbers to provide a basis for research.

9.25.2.4 Child Molestation

Child molestation is a generic term used to describe all types of sexual acts perpetrated by an adult on a child regardless of the gender of the child or the relationship between victim and offender. There are, not surprisingly, important differences between those men who molest their own children (incest offenders) and those who molest unrelated children (nonfamilial offenders), and between those who abuse boys and those who assault girls, but these issues are left until Section 9.25.3. The majority of child molesters are males, but in the past decade there have been increasing reports of female child molesters (Allen, 1991; Elliot, 1993). Nevertheless, male offenders very considerably outnumber females.

The sexual molestation of children involves a range of behaviors from fondling to vaginal or anal intercourse. In our research, we found that 36% of child molesters had intercourse with their child victims and, although this was more likely to occur when the victims were aged 10 years or older, many younger victims (as young as nine months in one case) were forced to submit to vaginal or anal intercourse (Christie, Marshall, & Lanthier, 1979; Marshall & Christie, 1981). Further, over 50% of incarcerated child molesters and 23% of child molesters attending an outpatient clinic had used unnecessary force when sexually assaulting a child. Obviously, child molestation is a far more dangerous and damaging behavior than was once thought to be the case.

Sexual assaults on children also occur with far greater frequency than most people wish to admit. Finkelhor and Lewis (1988), for example, asked a national sample of US adult males whether they had ever molested a child: 17% said yes. If we take this to provide an estimate of how many adult males have molested at least one child, then there are several million children in the USA who have been sexually assaulted. The Committee on Sexual Offenses against Children and Youths (1984) conducted national surveys in Canada to determine the frequency of child molestation. That committee concluded that one-half of Canadian females and one-third of Canadian males had been sexually victimized, and that over four-fifths of these assaults happened before the victim was aged 18 years. Seventy percent of the males and 62% of the females were prepubescent when the assaults occurred. Similar surveys in other countries have revealed similar figures and the effects on the victims are extensive and devastating (see Chapter 18, this volume).

9.25.2.5 Sexual Assaults

In this chapter, the term sexual assaults is used interchangeably with rapes and the offenders are referred to as rapists. Sexual assaults on adults are almost exclusively enacted by males, and by far the majority of the victims are female, although there have been reports of male victims (Stermac, Sheridan, Davidson, & Dunn, 1996). Those assaults that lead to a conviction are typically quite aggressive and involve humiliating and degrading features. Christie et al. (1979) found that over 70% of rapists engaged in gratuitous violence against their victims, and Darke (1990) reported that 60% of rapists engage in behaviors that are intended to degrade their victims. Not surprisingly, the effects on the victims of sexual assault are profoundly damaging (Koss & Harvey, 1991).

The Federal Bureau of Investigation's Uniform Crime Reports revealed dramatic increases in rape up to 1980, at which time 71 out of every 100 000 females at risk were assaulted. From 1980 on, there was a consistent slight decrease until 1989 when the rate returned to 1980 levels. Estimating rates of sexual assault from large-scale surveys reveals far higher frequencies, with the rates typically being four or more times greater than those derived from official figures (Ennis, (1967). Russell (1984), for example, found that 44% of the 930 women she surveyed reported having been sexually assaulted, but only 8% had reported the incident. Other surveys have produced similar findings (e.g., Koss, Gidcyz, & Wisniewski, 1987).

9.25.2.6 Summary of Prevalence

From a consideration of all these data on the various types of sexual offenses, it is clear that these behaviors occur at an alarmingly high rate and cause considerable damage, not only directly to the victims, but also to their families and to the innocent members of the offenders' families.

9.25.3 DIAGNOSIS AND ASSESSMENT

9.25.3.1 Diagnosis

In *DSM-IV*, the paraphilias are said to be characterized by "recurrent, intense, sexually arousing fantasies, sexual urges, or behaviors" (American Psychiatric Association, 1994, p. 522), having to do with the specific interests of the paraphiliac. These fantasies, urges, or behaviors must also "cause clinically significant distress, or impairment in social, occupational, or other important areas of functioning" (p. 523). Thus, a person may be diagnosed as having a paraphilia even if they have never acted on it. On the other hand, if they have enacted the behavior but it has not caused distress or impairment, then they should not be diagnosed as paraphilic.

For pedophilia the diagnostic criteria include "sexual activity with a prepubescent child or children" (American Psychiatric Association, 1994, p. 528), where the child is aged 13 years or younger. Just why this age limit was chosen is not clear. This criterion implies that a person who repeatedly molests postpubescent children should not be diagnosed as a pedophile. There is, unfortunately, no diagnostic category for what some clinicians have called ephebophilia (i.e., the molestation of young teenagers). Accordingly, those who sexually abuse children over age 14 years, even if they do so in a persistent and apparently preferential way, would not meet the diagnostic criteria for any paraphilia and yet they just as certainly would be considered sufficiently disordered by most clinicians to require treatment. As a result, clinicians and researchers do not always use the term "pedophilia" in strict accordance with the diagnostic guidelines.

Very few rapists meet diagnostic criteria for a paraphilia, primarily because there is no specific category in *DSM-IV* for this behavior. Those few rapists who are said to be sadistic would qualify for a *DSM-IV* diagnosis, but there are problems in arriving at such a diagnosis for sexual offenders. According to the diagnostic criteria, a sadist is a person who is sexually excited by the psychological or physical suffering of another person, where the psychological suffering includes humiliation. In our examination of the motives of rapists (Darke, 1990; Marshall & Darke, 1982), we selected only those behaviors displayed by rapists during their assaults that were rated by independent judges to be unequivocal signs of intended humiliation. We found that in 60% of rapes there were clear indications of attempts to humiliate the victim and that 60% of rapists declared that it was their intention to humiliate their victims. If sexual excitement was produced by humiliation in these men, then as many as 60% of rapists would meet the diagnostic criteria for sadism. This seems to contradict current diagnostic practices where far fewer rapists are identified as sadistic.

There are, then, problems in relating diagnostic criteria as defined by *DSM-IV*, to both clinical and research practice. Clinicians typically take the position that anyone who has committed a sexual offense is in need of assessment and treatment. Thus, clinicians tend to be guided by legal rather than by diagnostic criteria. Similarly, researchers generally have access only to adjudicated sexual offenders. Unfortunately, however, researchers have tended to use the terms child molester and pedophile rather loosely, so it is at times difficult to tell whether a study is examining all available child molesters or only those who meet the *DSM* criteria. Future revisions to the diagnostic manual should consider expanding the categories of paraphilia to approximate a closer match to clinical practice. Child molesters, with and without deviant preferences, and rapists with and without sadism, might be more useful categories.

9.25.3.2 Assessment

The assessment of sexual offenders has typically been aimed not at arriving at a diagnosis but rather at determining the man's risk to reoffend, his treatment needs, and whether or not treatment produced the hoped-for result. In pursuit of these aims, an almost unlimited array of evaluation procedures have been adopted. While the assessment measures may differ, the content areas evaluated reveal some consistencies, with the targets identified in Table 1 being the most common. Table 1 also indicates some of the measures that have been used. Projective tests are not listed in this table, although a number of clinics routinely employ them, because the limited evidence available does not encourage confidence in their value with sexual offenders (Levin & Stava, 1987) and because there are more general questions concerning the psychometric properties of projective tests (Kline, 1993).

Other alternatives that have been used to assess the accuracy of the self-reports of sexual offenders include conducting interviews assisted by polygraphy (Abrams, 1991; Emrick & Dutton, 1993), using tests that have built-in indicators of faking (e.g., the Minnesota Multiphasic Personality Inventory [MMPI], Graham, 1990; and the Multiphasic Sex Inventory, Nichols & Molinder, 1984), or adding a measure of the propensity to present in a socially desirable manner (e.g., the Marlowe–Crowne Social Desirability Scale, Crowne & Marlowe, 1980). Unfortunately, there are limits to the value of each of these assessment measures. Although the use of polygraphy in the evaluation of sexual offenders appears to be widespread, its value has been challenged (Lalumière & Quinsey, 1991) and there is evidence that sexual offenders can dissimulate on the MMPI without the faking measures detecting this (Langevin, Wright, & Handy, 1990). Also, there are some serious doubts about whether measures of social desirability do in fact evaluate this proclivity (Holden & Fekken, 1989).

There is not room here to evaluate all the approaches to assessment, but comment is made on some because the areas they examine represent topical issues with sexual offenders.

9.25.3.2.1 Sexual interests/history

It is clear that most offenders can fake their phallometric responses so as to appear to have normal preferences and while some of the more obvious attempts at faking can be detected, cognitive strategies cannot (Murphy & Barbaree, 1994). While some assessors have developed ways to circumvent faking (Malcolm, Davidson,

Table 1 Assessment targets and some typical measures.

Target	*Typical measures*
Sexual interests/history	Phallometric testing (Murphy & Barbaree, 1994)
	Sexual Interest Card Sort (Abel & Becker, 985)
	Clarke Sexual History Questionnaire (Langevin, 1983)
	Multiphasic Sex Inventory (Nichols & Molinder, 1984)
Social functioning	Anxiety (Watson & Friend, 1969)
	Assertiveness (Rathus, 1973)
	State-Trait Anger Expression Inventory (Speilberger, 1988)
	Child Molester Empathy Measure (Fernandez, Marshall, Lightbody, & O'Sullivan, 1997)
	Social self-esteem (Lawson, Marshall, & McGrath, 1979)
	Social intimacy (Miller & Lefcourt, 1982)
Cognitive processes and attitudes and beliefs	Abel's Cognitive Distortions Measure (Abel et al., 1989)
	Rape Myth Acceptance Scale (Burt, 1980)
	Attitudes toward women (Spence, Helmreich, & Stapp, 1973)
	Hostility toward women (Check, 1984)
Personality	Hare's Psychopathy Checklist—Revised (Hare, 1991)
	Minnesota Multiphasic Personality Inventory (Graham, 1990)
Substance abuse	Michigan Alcoholism Screening Test (Selzer, 1971)
	Drug Abuse Screening Test (Skinner, 1982)

& Marshall, 1985; Quinsey & Chaplin, 1988), others have not consistently found these approaches to be effective (Proulx, Coté, & Achille, 1993). Evidence on the reliability of phallometric assessments is quite limited and at least one study reported very poor test–retest reliability (Barbaree, Baxter, & Marshall, 1989).

There are also other problems with the use of phallometry. For example, while the evidence suggests that phallometric measures can distinguish nonfamilial child molesters from other men, these evaluations have not proved so successful in identifying incest offenders, rapists, or exhibitionists (Murphy & Barbaree, 1994). Nevertheless, alternative assessment procedures (e.g., self-reported preferences or card-sort tests) seem even more vulnerable to dissimulation and even less reliably distinguish sexual offenders from other males (Marshall, 1996a). For the foreseeable future, then, phallometric measures will continue to represent the best approach to identifying deviant sexual interests.

It is also essential to determine the sexual history of these men, partly to identify the range and persistence of their deviant practices, but also to determine their experience with appropriate sexual behaviors. Both the Clarke Sexual History Questionnaire (Langevin, 1983) and the Multiphasic Sex Inventory (Nichols & Molinder, 1984) are useful measures in this regard but so also is the clinical interview.

9.25.3.2.2 Social functioning

The primary areas of interest here appear to be the offender's self-confidence, his capacity for empathy, his ability to form intimate relationships, and his more general social skills. There is clear evidence that sexual offenders lack self-esteem (Marshall, Anderson, & Champagne, 1996), have a limited capacity for empathy (Marshall, Hudson, Jones, & Fernandez, 1995), are deficient in intimacy (Marshall, 1993), and are socially deficient (McFall, 1990). Clearly, these aspects of their functioning need to be assessed in order to make clinical decisions about sexual offenders.

Most measures of self-esteem treat the concept as a trait manifest across situations and time, and yet there is clear evidence that self-confidence is significantly affected by general circumstantial features (Marshall, Anderson et al., 1996). For example, we have found that the self-esteem of individual sexual offenders varies according to whether we are asking them about confidence in social or sexual interactions, or about their appearance or their occupational functioning (Anderson & Marshall, 1996). Measures of self-esteem, then, need to be specific to the domain of functioning in which we are particularly interested.

Hanson (1997) has reviewed empathy measures that have been used with sexual offenders. He notes that empathy is, in fact, a multicomponent process and as such we need to employ more than a simple general measure. Davis (1980, 1983) has developed such a measure and it has been usefully employed with sexual offenders. However, more specific measures have also been developed for child molesters (Fernandez, Marshall, Lightbody, & O'Sullivan, 1997) and rapists (Fernandez, 1996), and for the critical perspective-taking aspect of empathy (Hanson & Scott, 1995).

Relevant measures of the capacity of sexual offenders to form adult romantic/sexual relationships have been adopted from other fields of research. For example, various measures of intimacy have been applied to the assessment of sexual offenders and they have revealed significant deficits in these men compared with other males (Bumby & Marshall, 1994; Seidman, Marshall, Hudson, & Robertson, 1994). An alternative and perhaps more productive approach to relationship issues has been developed by Ward, Hudson, and McCormack (1997). They assess attachment styles rather than the attainment of intimacy, since it is the style that maximizes or minimizes the possibility of developing effective relationships, whereas intimacy can be the chance outcome of various other influences as well. Ward et al.'s measure of attachment style relies on a categorization developed by Bartholomew and Horowitz (1991) which unfortunately is rather crude and seems not to fully capture the essence of the issue. However, this appears to be a very promising approach and Bartholomew's measure may be useful in combination with a thorough interview.

Attempts to evaluate other aspects of social functioning include questionnaire assessments of assertiveness, social anxiety, and anger, as well as role-play measures of social skills (Bellack, 1983).

9.25.3.2.3 Cognitive processes

Ward, Hudson, Johnston, and Marshall (1997) have shown that sexual offenders characteristically have inappropriate views of the sexuality of women and children, and distorted perceptions of both their own behavior during their offenses and the reactions of their victims. It has also been shown that child molesters construe children in more appealing ways than they do adults (Horley, Quinsey, & Jones, 1997; Howells, 1978).

Abel et al.'s (1989) measure of cognitive distortions in child molesters evaluates their tendency to distort the world in a self-serving way, but is restricted to their view of children and sexuality. Until the early 1990s, the majority of assessments of rapists' perceptions concerned the evaluation of their attitudes toward women. However, Bumby (1996) developed two excellent measures of cognitive distortions: one for child molesters and one for rapists. Unfortunately, all of these measures rely on self-report, and many offenders who score in the appropriate range later in therapy, or at interview, reveal quite inappropriate or distorted perceptions.

9.25.3.2.4 Personality

Much has been made in the literature of the possibility that sexual offenders may have distinctive personalities. However, despite persistent attempts to discern a personality profile that characterizes these men, little in the way of consistent findings has been revealed.

Among the measures that have been used, the two most popular have been the MMPI and Hare's (1991) Psychopathy Checklist-Revised (PCL-R). While there have been inconsistent results using the MMPI (Marshall & Hall, 1995), it may, nevertheless, be useful to determine whether or not a particular offender displays any other sign of psychopathology. Hare's measure has consistently revealed psychopathic tendencies in a reasonable number of sexual offenders, although the percentage of psychopaths tends to be higher among those held in psychiatric facilities (Quinsey, Lalumière, Rice, & Harris, 1995) than among imprisoned sexual offenders (Serin, Malcolm, Khanna, & Barbaree, 1994). There is no doubt that Hare's measure is a useful assessment tool, particularly for determining a sexual offender's risk to reoffend.

9.25.3.2.5 Substance abuse

A substantial number of sexual offenders are chronic abusers of either alcohol or some other drug, or use one or another drug prior to or during their offensive behavior. Using the Michigan Alcoholism Screening Test (Selzer, 1971), Langevin and Lang (1990) found that 55.8% of sexual offenders scored within the alcoholic range. On the Drug Abuse Screening Test (Skinner, (1982), Langevin and Lang (1990) found that 17.8% had a serious problem in the abuse of other drugs. These two measures adequately evaluate problems with alcohol or drug use, but it is also important for clinicians to attempt to discern the relationship between nonproblematic substance use and offending behavior. Many sexual offenders are more likely to abuse when they are somewhat intoxicated even if they do not characteristically have a substance-use disorder.

9.25.4 TREATMENT

Prior to the 1970s, treatment for sexual offenders, when it was provided, was based on a variety of theoretical views and each program tended to be distinctly unique. Very few of these programs provided sufficient details to allow replication and even less offered estimates of the effectiveness of their program. In the late 1960s, reports began to appear of treatments that were guided by a behavioral view of sexual offending. These programs took the position that sexual offending was driven by acquired deviant sexual preferences and that all that was necessary in treatment was to change these preferences and the problem behaviors would disappear. It did not take long, however, for behavior therapists to recognize that this was a naive conceptualization and elements having to do with social skills were quickly added. These behavioral programs evolved into progressively more comprehensive cognitive behavioral programs targeting an expanded range of social skills and including cognitive distortions and relapse prevention procedures.

Along with the expansion of the content of cognitive behavioral programs, the use of pharmacologic interventions increased. Antiandrogen and hormonal agents were shown to reduce overall drive levels or, if administered in appropriate dosages, they reduced only deviant responding. Since the mid-1980s, the serotonin reuptake inhibitors have been found to reduce the strength of urges sufficiently for the offender to be able to exercise control. The use of these pharmacologic agents, however, has typically been as complementary to psychological interventions used as adjuncts to comprehensive cognitive-behavioral programs (Greenberg & Bradford, 1997).

9.25.4.1 Cognitive-behavioral Programs

Cognitive-behavioral programs (or relapse prevention programs) have become the most popular approaches in North America, as well as in other countries, for the treatment of sexual offenders (Marshall, Fernandez, Hudson, & Ward, in press). The focus of 98% of these programs in North America is group therapy, although individual therapy is often used as an adjunct (Cellini, 1995). Although the content of these programs varies somewhat across the different settings, the following targets appear to be addressed in all cognitive-behavioral approaches: acceptance of responsibility, cognitive distortions and pro-offending attitudes; empathy for victims; social functioning; deviant fantasies; and the development of relapse prevention plans. In addition, most programs address a variety of other features, such as substance abuse, sex education, the childhood abuse of the offender, and whatever other idiosyncratic problems seem to be related to the client's offense propensity.

9.25.4.1.1 Acceptance of responsibility

Many treatment programs appear to expect offenders to accept some degree of responsibility for having offended, prior to entry into treatment. Accordingly, these programs refuse to accept those clients who deny having committed an offense. Other programs consider extensive minimization (e.g., claiming that other people and external circumstances were responsible for the offending, or minimizing the nature of the offense) as evidence of poor motivation and, therefore, exclude them from treatment. However, refusing deniers and minimizers means excluding not only a substantial number of clients who have been convicted of sexual offenses but also many of the most dangerous offenders. All clients should be accepted into treatment because it is the job of therapists to motivate clients and get them to accept responsibility. By challenging them in a firm but supportive manner, and by making clear the costs to them for maintaining denial or minimization and the benefits of changing their position, therapists should be able to shift the sexual offenders' views to an acceptance of responsibility (Barbaree, 1991; Marshall, 1994).

9.25.4.1.2 Cognitive distortions

This term describes both the distorted way in which sexual offenders perceive the responses of their victims and their attitudes that are offense supportive. Many sexual offenders perceive their victims as cooperative and as not distressed or otherwise harmed by the abuse. Rapists and exhibitionists frequently hold negative views of women and as a consequence feel anger toward them. Child molesters typically see children as either more friendly and less threatening than adults, or as simply available for their use. Using variations on what has been called cognitive restructuring procedures (Murphy, 1990), these views are challenged and alternative views are encouraged. However, no evaluations of the specific effects of these interventions are available.

9.25.4.1.3 Empathy

The aims are to sensitize the offender to the harm their abuse has caused and to enhance his feelings of concern for those he has harmed. While the focus in this component of treatment is on the harm to the victim and on the offender's remorse for having caused this harm, the distress experienced by the victim's family and by the offender's own family is also emphasized (Pithers, 1994, in press).

Since empathy has been conceptualized as a four-stage, unfolding process (Marshall et al., 1995), treatment efforts focus on these stages in order. First attempts are made to train offenders in recognizing emotions in others. Next, assistance is given to the offender in identifying harm to the victim by having him adopt the victim's perspective. As part of this process, offenders repeatedly report how they are feeling when they think about the victim and, finally, the offender indicates what he should do to avoid causing harm or to terminate distress when it occurs. Throughout these procedures, all group members are required to evaluate the target offender's responses and indicate their own feelings while attending to the target offender's responses. Evaluation of this treatment component has shown substantial improvement in sexual offenders (Marshall, O'Sullivan, & Fernandez, 1996).

9.25.4.1.4 Social functioning

A broad range of issues are typically addressed in this component. Conversational skills, anxiety, anger management, assertiveness, and problem solving are among the many social problems targeted here. Treatment programs have also included procedures for enhancing self-esteem and improving relationship skills.

Research across a variety of human problems has demonstrated that low self-esteem impedes progress in treatment and increases the risk of post-treatment relapse (Marshall, Anderson et al., 1996). Accordingly, therapists attempt to establish a working environment and a therapeutic process that will facilitate the enhancement of self-esteem, as well as employing specific procedures to increase self-confidence. An evaluation of this approach has revealed marked improvements in self-esteem (Marshall, Champagne, Sturgeon, & Bryce, 1997) and strong relationships between these improvements and positive changes in other treatment targets, including reductions in deviant preferences (Marshall, 1997) and increases in empathy and intimacy skills (Marshall, Champagne et al., 1997).

Attachment style refers to the characteristic way in which people engage in, or avoid, intimate relationships. Treatment programs that address this issue train clients to develop more effective attachment styles and related skills by addressing issues such as the benefits of good attachments, developing an understanding of effective sexual functioning, improving reciprocity and communication in relationships, increasing conflict-resolution skills, learning how to deal with being alone, and, finally, focusing on the nature of jealousy and the problems it can cause. Again, it has been demonstrated that these procedures are effective in reducing feelings of loneliness and in enhancing intimacy in the lives of sexual offenders (Marshall, Bryce, Hudson, Ward, & Moth, 1996).

9.25.4.1.5 Deviant fantasies

Procedures have long been available for reducing an attraction to deviant sexual acts and for increasing the attractiveness for appropriate sexual behaviors. Early treatment programs employed some variation on aversion-conditioning procedures using either mild electric shocks, nausea-inducing drugs, or foul odors as the aversive event. These procedures have, for the most part, been abandoned and replaced by one of several alternatives.

Covert sensitization is a technique that associates imagined unpleasant consequences with imagining a presently preferred, but maladaptive, behavior. Abel, Blanchard, and Becker (1978) provided a detailed account of the application of this procedure to the treatment of a rapist, and Maletzky (1980) described its use with exhibitionists. Covert sensitization, in various forms, has enjoyed considerable popularity ever since. However, the research supporting its use is not substantial, and it has not been unequivocally shown to alter sexual preferences (Quinsey & Earls, 1990; Quinsey & Marshall, 1983).

Masturbatory reconditioning combines two procedures: (i) masturbation to appropriate fantasies up to the point of orgasm (Maletzky, 1985) and (ii) satiation procedures which involve repeated verbalizations (aloud or subvocal) of deviant fantasies during the refractory period immediately after orgasm (Marshall, 1979). While there is reasonable empirical support for the effectiveness of satiation, the evidence bearing on the first part of the procedure (i.e., masturbating to appropriate fantasies) is quite limited (Laws & Marshall, 1991).

Despite these limitations in the evidence supporting procedures that attempt to modify disordered sexual interests, they continue to be employed (Dougher, 1995), apparently for two primary reasons. In the first place, most practitioners appear to believe that deviant sexual fantasies play a significant role in motivating deviant sexual behavior and secondly, it appears to be assumed that most sexual offenders have deviant sexual fantasies even if these are not evident at assessment. Acceptance of these views, of course, would direct treatment attention to changing sexual preferences.

However, as noted earlier, Marshall (1997) has found that modifying other aspects of the offenders' functioning produces indirect and beneficial changes in sexual preferences, and Pithers (in press) has reported similar findings. It seems reasonable to assume then, that, except for those limited number of cases where deviant arousal is very strong, it may not be necessary to directly target sexual preferences in treatment. In those cases where deviant arousal does seem to overwhelm the offender's sense of control, it may be wiser to employ either antiandrogen medication (Bradford, 1990) or serotonin reuptake inhibitors (Fedoroff, 1993) as an interim step until the offender begins to develop a personal sense of control.

In fact, a focus on strictly sexual fantasies may miss the fact that sexual offenders have numerous other offense-related fantasies that need to be addressed in treatment. For example, many rapists are preoccupied with fantasies of control and with fantasies involving the humiliation and degradation of women. These fantasies frequently occur in the absence of sexual arousal and yet they appear to play an important role in motivating and maintaining a persistent disposition to attack women. Similar nonsexual fantasies seem to characterize both child molesters and exhibitionists.

9.25.4.1.6 Relapse prevention

This approach to treatment was adapted from Marlatt and Gordon's 1985) description of their way of enhancing post-treatment maintenance of behavior in addicts and alcoholics. Pithers (1990) has detailed the structure and content of this approach, and Laws (1989) has provided clinical descriptions of each of the elements in a relapse-prevention approach with sexual offenders. Essentially, clients are trained to identify the factors that put them at risk, including precipitating factors (e.g., low self-esteem, deviant fantasizing, intoxication, anger, etc.), circumstantial risks (e.g., being alone with a child, cruising aimlessly in a car), and grooming techniques or strategies for gaining access to a victim. These risk factors are then addressed by generating plans to either avoid such risks or to escape from them should they inadvertently arise. Pithers (1990) adds to these so-called "internal management" components of treatment, an "external supervisory dimension" that requires extensive post-treatment involvement in a maintenance program. This maintenance program includes careful supervision by a person trained in relapse-prevention techniques, the provision of an informed support group of friends or family, and regular participation in a treatment maintenance group.

Attempts were made to evaluate, on the basis of the very limited evidence available, whether or not either the internal management or the external supervision components of relapse prevention added any therapeutic efficacy to the already comprehensive cognitive/behavioral treatment programs (Marshall & Anderson, 1996). It appears that the internal management component does increase efficacy, but the evidence was not sufficient to determine the value of the extensive external management procedures.

9.25.4.2 Overall Value of Treatment

Although there is a lack of satisfactorily controlled treatment outcome studies, some authors have been satisfied that there are grounds for optimism based on the available, albeit limited, evidence (e.g., Marshall, Jones, Ward, Johnston, & Barbaree, 1991). On the other hand, Quinsey and his colleagues (e.g., Quinsey, Harris, Rice, & Lalumière, 1993) have consistently expressed pessimism regarding the value of treatment. Two meta-analyses of the available evidence have taken a more conservative view, suggesting that there does appear to be an effect for treatment but that the effect size is rather small (Hall, 1995; Hanson & Bussière, 1996). In part, this difficulty in discerning clear effects for treatment is the result of the characteristically low base-rate of recidivism among untreated sexual offenders. However, the fact that programs clearly vary, in the content of treatment, the process of treatment delivery, and the length of treatment provided, will continue to make it problematic to compare different programs and, thereby, preclude any general statement about treatment efficacy. On a more positive note, it has been shown that even minimally effective treatment of sexual offenders can reduce the future abuse of innocent women and children, and can save taxpayers a substantial amount of money (Marshall, 1992; Prentky & Burgess, 1991).

9.25.5 FUTURE DIRECTIONS

Despite the difficulties or impracticalities of distinguishing research from practice, this chapter tries to identify those future directions considered to be most important within each of these domains.

9.25.5.1 Research

A primary failing that has beset this area throughout its history remains a stumbling block to attempts at integrating findings from various studies. This failing concerns the lack of sufficient information provided in reported studies on the subjects, procedures, and settings in which the research was conducted. Presumably, no one would expect offenders with a chronic history of sexually abusing numerous victims in a vicious manner to display the same characteristics, and be as responsive to the same treatment, as would single-victim offenders who have engaged in only mildly intrusive, nonforceful acts. Freund and Watson (1991), for example, found that the phallometric test sensitively identified child molesters with multiple victims (80% accurate with nonfamilial offenders who molested female children, and 90% for those with male children victims) but did not do at all well with single-victim offenders (only 45% accuracy). Very few other reports of phallometric evaluations have made this distinction among their subjects, and yet it is obviously very important.

On this issue of phallometric testing, far more research is needed on its psychometric properties, as the present data do not provide a satisfactory basis to defend the use of these procedures. For example, the reliability of phallometry is questionable with very few studies being available and the findings across studies are inconsistent (Marshall, 1996a). In various other respects, although there is an extensive body of literature on phallometric test results, the available data can only be integrated into a supportive argument if a whole range of important inconsistencies between studies are ignored.

Related to this, there is very little support for the value of therapeutic procedures aimed at modifying deviant sexual preferences (Laws & Marshall, 1991; Quinsey & Earls, 1990; Quinsey & Marshall, 1983). In fact, some data suggest that these procedures may be unnecessary (Marshall, in press). Further research is clearly needed to either substantiate the value of these procedures (i.e., do they, in fact, modify deviant arousal) or, more importantly, to determine the necessity of targeting deviant preferences in treatment.

The prediction of recidivism is a critical area that only began to receive the necessary empirical attention in the mid-1990s. While Quinsey et al. (1995) made an excellent start to this process, the research needs to be extended to broader populations of sexual offenders and, most importantly, needs to include more dynamic factors that are open to therapeutic modification. One of the problems that obstructs such research concerns the identification of reoffending. In Canada, satisfactory systems are accessible to researchers that identify all charges, reconvictions, and parole lapses across the country. In the USA and other countries, researchers do not have access to such a complete set of data, making recidivism studies something of a hit-or-miss gamble. In order for effective research to proceed, a sound basis on which to estimate recidivism must be developed. If governments wish to develop a comprehensive understanding of sexual offenders, they must be willing to establish a national pool of information on reoffending that is available to researchers.

Treatment evaluations have been limited in number and have rarely been able to provide an estimate of recidivism in the absence of treatment. This presents a very difficult problem because the opportunity to conduct an outcome study, where subjects are randomly allocated to treatment or no treatment, is rare. In most cases, researchers cannot get the cooperation of sexual offenders because their possible release on parole would be blocked if they remained untreated. Also, very few organizations would be willing to endorse such an endeavor because of the potential for extremely negative publicity that would result if a deliberately untreated sexual offender were to attack an innocent person after discharge. Unfortunately, in those rare cases where rigorous evaluations have been conducted, the funding agencies seem more concerned with imposing experimental rigor than they are with evaluating appropriate clinical applications of treatment. This, of course, represents an ever-present clash of ideologies about how to effectively evaluate clinical processes. While the absence of a concern for controlled evaluations has been most unfortunate, the overconcern for experimental rigor is not necessarily a step forward.

On the issue of treatment evaluation, one of the features that needs examination is the value of the extensive, and expensive, aftercare procedures that are advocated by relapse-prevention enthusiasts. Some degree of post-release supervision and maintenance seems necessary, if only to convince the public that every precaution has been exercised. However, it has been simply assumed that the need for extensive aftercare is obvious. Sometimes what seems obvious is not supported by careful examination, and that needs to be done in this case, particularly considering the very expensive nature of maintenance programs being advocated.

Finally, more work needs to be done on identifying features of sexual offenders that indicate the extent of their treatment needs. This ties in with the need to have available clear indices of reoffending. Limited treatment resources must be put where they are most needed. It is pointless offering extensive treatment to all sexual offenders when some of them are at quite low risk to reoffend, and have limited problems that need treatment attention. However, allocating offenders to one or another level of treatment involvement relies on guesswork based on limited evidence. This basis of evidence needs to be extended. However, generating typologies, as they have been created in the past, has not been particularly useful in solving the problem of the appropriate allocation of treatment resources. While Knight and Prentky's (1990, 1993) typology is a marked improvement over previous classificatory systems, it is, nevertheless, not designed to identify likely treatment responsiveness or needs, and has not, in the late 1990s, yet been evaluated for its utility in this regard.

9.25.5.2 Practice

Parallel with attempts to discern those sexual offenders in need of extensive or less extensive treatment, intervention programs need to develop strategies for allocating resources to clients with different needs so that resources can be maximally utilized. We have outlined an approach for dealing with this problem within both our prison-based and community-based programs (Marshall, Eccles, & Barbaree, 1993). Essentially, we allocate offenders on the basis of assessment and risk determined needs, to one of three levels of extensive/intensive treatment. High risk/high needs offenders receive the maximum allocation of our resources, while the low risk/low needs clients get far less extensive treatment and a lower frequency of treatment sessions. So far, the results from this allocation of resources encourage confidence that our decisions are correct, although more sophisticated procedures for identifying needs based on long-term outcome studies must be developed. Other programs all too frequently do not make such practical use of their limited resources.

In addition to this allocation of resources, it seems sensible to develop specific programs to deal with particular types of offenders, such as minority or culturally distinct groups, disabled offenders, psychiatrically disturbed offenders, professional or clergy offenders, etc. Work on these types of programs has been developed, but most modifications meant to fit these special populations have been based on the practitioners' assumptions of the unique problems facing these offender types rather than on an empirical evaluation of the need for specific approaches.

There are, of course, many features of assessment and treatment that need further development. We are only at the beginning of determining the best way to address the issues of empathy, denial and minimization, attachment styles, and attitudes and belief systems. Conceptualizations of these problems have been inadequate and the consequent lack of clarity has meant that various practitioners use different assessment and treatment approaches to these problems. Not surprisingly, this has meant that, in the late 1990s, assessment procedures for each of these topics are rather inadequate and need improving, and the treatment of these problems has, to date, been either based on the good sense of the practitioner or simply adopted from other areas of treatment without modifications to meet the peculiar needs of sexual offenders. Clarification of, and some degree of agreement about, what constitutes empathy, intimacy, denial, minimization, and dysfunctional attitudes and beliefs, is essential, and practitioners cannot afford to be complacent about this problem.

The final features suggested in this chapter which need a good deal more clinical attention and research are the processes of treatment and assessment. The advent of behavioral therapy, and its offspring cognitive/behavioral therapy, brought about a quite proper concern for specifying procedures, but unfortunately, it also brought with it the belief that procedures alone could produce satisfactory assessment and treatment. There is now a growing sentiment within cognitive/behavioral therapy that process variables, such as the assessor's or therapist's style and the nature of the relationship established between assessors/therapists and clients, are critical to effective evaluation and treatment (Miller & Rollnick, 1991; Schaap, Bennun, Schindler, & Hoogduin, 1993). An examination of these features in the context of work with sexual offenders is long overdue, but some possible avenues for developing clinical practice have been suggested (Marshall, 1996b). Far more work on these presumably critical variables is needed both in clinical practice and in research.

9.25.6 SUMMARY

This chapter has described what appear to be the main issues concerning adult sexual offenders. The nature and extent of sexual abuse is of such significant magnitude, and damages so many innocent lives, that those who work in this field must apply themselves diligently to the task of addressing this serious social problem in the most effective and cost-beneficial way possible. Significant progress has already been made but there is a long way to go and researchers cannot afford to be complacent. Our research and clinical practice needs to be expanded so that the capacity to identify the highest-risk offenders can be increased and the likelihood that they will not reoffend maximized. Evidence available encourages optimism that we are on the right track, but we need to examine procedures more thoroughly and, most importantly, develop new ones as concepts are refined and the areas of evaluation and treatment are expanded.

9.25.7 REFERENCES

Abel, G. G., & Becker, J. V. (1985). *Sexual interest cardsort*. Atlanta, GA: Behavioral Medicine Laboratory, Emory University.

Abel, G. G., Becker, J. V., Cunningham-Rathner, J., Mittelman, M. S., & Rouleau, J. L. (1988). Multiple paraphilic diagnoses among sex offenders. *Bulletin of the American Academy of Psychiatry and the Law, 16,* 153–168.

Abel, G. G., Becker, J. V., Mittelman, M. S., Cunningham-Rathner, J., Rouleau, J. L., & Murphy, W. D. (1987). Self-reported sex crimes of nonincarcerated paraphiliacs. *Journal of Interpersonal Violence, 2,* 3–25.

Abel, G. G., Blanchard, E. B., & Becker, J. V. (1978). An integrated treatment program from rapists. In R. T. Rada (Ed.), *Clinical aspects of the rapists* (pp. 161–214). New York: Grune & Stratton.

Abel, G. G., Gore, D. K., Holland, C. L., Camp, N., Becker, J. V., & Rathner, J. (1989). The measurement of the cognitive distortions of child molesters. *Annals of Sex Research, 2,* 135–152.

Abrams, S. (1991). The use of polygraphy with sex offenders. *Annals of Sex Research, 4,* 239–263.

Allen, C. (1991). *Women and men who sexually abuse children: A comparative analysis*. Brandon, VT: Safer Society Press.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anderson, D., & Marshall, W. L. (1996). *A multidimensional self-esteem inventory related to sexual offending*. Paper presented at the 15th Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, Chicago.

Barbaree, H. E. (1991). Denial and minimization among sex offenders: Assessment and treatment outcome. *Forum on Corrections Research, 3,* 300–333.

Barbaree, H. E., Baxter, D. J., & Marshall, W. L. (1989). The reliability of the rape index in a sample of rapists and nonrapists. *Violence and Victims, 4,* 299–306.

Bartholomew, K., & Horowitz, L. M. (1991). Attachment style among adults: A test of a four category model. *Journal of Personality and Social Psychology, 61,* 226–244.

Bellack, A. S. (1983). Recurrent problems in the assessment of social skills. *Behaviour Research and Therapy, 27,* 29–42.

Blair, D. C., & Lanyon, R. I. (1981). Exhibitionism: Etiology and treatment. *Psychological Bulletin, 89,* 439–463.

Bradford, J. M. W. (1990). The antiandrogen and hormonal treatment of sex offenders. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 297–310). New York: Plenum.

Bumby, K. M. (1996). Assessing the cognitive distortions of child molesters and rapists: Development and validation of the MOLEST and RAPE scales. *Sexual Abuse: A Journal of Research and Treatment, 8,* 37–54.

Bumby, K., & Marshall, W. L. (1994). *Loneliness and intimacy dysfunction among incarcerated rapists and child molesters*. Paper presented at the 13th Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, San Francisco.

Burt, M. (1980). Cultural myths and supports for rape. *Journal of Personality and Social Psychology, 38,* 217–230.

Cellini, H. R. (1995). Assessment and treatment of the adolescent sexual offender. In B. K. Schwartz & H. R. Cellini (Eds.), *The sexual offender: Corrections, treatment and legal practice* (pp. 6.1–6.12). Kingston, NJ: Civic Research Institute.

Check, J. V. (1984). *The hostility towards women scale*. Unpublished doctoral dissertation, University of Manitoba, Winnipeg, MB.

Christie, M. M., Marshall, W. L., & Lanthier, R. D. (1979). *A descriptive study of incarcerated rapists and pedophiles*. Report to the Solicitor General of Canada, Ottawa, ON.

Committee on Sexual Offenses against Children and Youths (1984). *Report on the Committee on Sexual Offenses against Children and Youths, Vols. 1–11, and Summary* (Badgley Report), Ottawa, ON: Department of Supply and Services.

Crowne, D. P., & Marlow, D. (1980). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24,* 349–354.

Darke, J. L. (1990). Sexual aggression: Achieving power through humiliation. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 55–72). New York: Plenum.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10,* 85.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44,* 113–126.

Di Vasto, P. V., Kaufman, L. R., Jackson, R., Christy, J., Pearson, S., & Burgett, T. (1984). The prevalence of sexually stressful events among females in the general population. *Archives of Sexual Behavior, 13,* 59–67.

Dougher, M. J. (1995). Behavioral techniques to alter sexual arousal. In B. K. Schwartz & H. R. Cellini (Eds), *The sex offender: Corrections, treatment and legal practice* (pp. 15.1–15.8). Kingston, NJ: Civic Research Institute.

Elliot, M. (1993). *Female sexual abuse of children*. New York: Guilford Press.

Emrick, R. L., & Dutton, W. A. (1993). The effect of polygraphy on the self report of adolescent sex offenders: Implications for risk assessment. *Annals of Sex Research, 6,* 83–103.

Ennis, P. H. (1967). *Criminal victimization in the United States: A report of a national survey*. National Opinion Research Center (NORC), University of Chicago. Washington, DC: US Government Printing Office.

Fedoroff, J. P. (1993). Serotonic drug treatment of deviant sexual interests. *Annals of Sex Research, 6,* 105–121.

Fernandez, Y. M. (1996). *The rapist empathy measure*. Paper presented at the 15th Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, Chicago.

Fernandez, Y. M., Marshall, W. L., Lightbody, S., & O'Sullivan, C. (1997). *The child molester empathy measure*. Manuscript submitted for publication.

Finkelhor, D., & Lewis, I. (1988). An epidemiologic approach to the study of child molestation. *The Annals of the New York Academy of Science, 528,* 64–67.

Freund, K. (1990). Courtship disorder. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 195–207). New York: Plenum.

Freund, K., & Watson, R. J. (1991). Assessment of the sensitivity and specificity of a phallometric test: An update of phallometric diagnosis of pedophilia. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 254–260.

Graham, J. R. (1990). *MMPI-2: Assessing personality and psychopathology.* New York: Oxford University Press.

Greenberg, D. M. & Bradford, J. M. W. (1997). Treatment of the paraphilic disorders: A review of the role of the selective serotonin reuptake inhibitors. *Sexual Abuse: A Journal of Research and Treatment, 9,* 349–360

Hall, G. C. N. (1995). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology, 63,* 802–809.

Hanson, R. K. (1997). Invoking sympathy: Assessment and treatment of empathy deficits among sexual offenders. In B. K. Schwartz & H. R. Cellini (Eds.), *The sex offender: New insights, treatment innovations and legal developments* (Vol. 11, pp. 1.1–1.12). Kingston, NJ: Civic Research Institute.

Hanson, R. K., & Bussière, M. (1996). *Predictors of sex offender recidivism: A meta-analysis.* Report to the Solicitor General of Canada, Ottawa, ON.

Hanson, R. K., & Scott, H. (1995). Assessing perspective taking among sexual offenders, nonsexual criminals and nonoffenders. *Sexual Abuse: A Journal of Research and Treatment, 7,* 259–277.

Hare, R. D. (1991). *Manual for the revised psychopathy checklist.* Toronto, ON: Multi-Health Systems.

Holden, R. R., & Fekken, G. C. (1989). Three common social desirability scales: Friends, acquaintances, or strangers? *Journal of Research in Personality, 23,* 180–191.

Horley, J., Quinsey, V. L., & Jones, S. (1997). Incarcerated child molesters' perceptions of themselves and others. *Sexual Abuse: A Journal of Research and Treatment, 9,* 43–55.

Howells, K. (1978). Some meanings of children for pedophiles. In M. Cook & G. Wilson (Eds.), *Love and attraction* (pp. 57–82). London: Pergamon.

Kline, P. (1993). *The handbook of psychological testing.* New York: Routledge.

Knight, R. A., & Prentky, R. A. (1990). Classifying sexual offenders: The development and corroboration of taxonomic models. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories and treatment of the offender* (pp. 23–52). New York: Plenum.

Knight, R. A., & Prentky, R. A. (1993). Exploring characteristics for classifying juvenile sexual offenders. In H. E. Barbaree, W. L. Marshall, & S. M. Hudson (Eds.), *The juvenile sex offender* (pp. 45–83). New York: Guilford.

Koss, M. P., Gidycz, C. A., & Wisniewski, N. (1987). The scope of rape: Incidence and prevalence of sexual aggression and victimization in a national sample of higher education students. *Journal of Consulting and Clinical Psychology, 55,* 162–170.

Koss, M. P., & Harvey, M. R. (1991). *The rape victim: Clinical and community interventions* (2nd ed.). Newbury Park, CA: Sage.

Lalumière, M. L., & Quinsey, V. L. (1991). Polygraph testing of child molesters: Are we ready? *Violence Update, 1,* 3–11.

Langevin, R. (1983). *Sexual stands: Understanding and treating sexual anomalies in men.* Hillsdale, NJ: Erlbaum.

Langevin, R., & Lang, R. A. (1990). Substance abuse among sex offenders. *Annals of Sex Research, 3,* 397–424.

Langevin, R., Paitich, D., & Russon, A. E. (1985). Are rapists sexually anomalous, aggressive, or both? In R. Langevin (Ed.), *Erotic preference, gender identity, and aggression in men: New research studies* (pp. 130–138). Hillsdale, NJ: Erlbaum.

Langevin, R., Wright, P., & Handy, L. (1990). Use of the MMPI and its derived scales with sex offenders: 1. Reliability and validity studies. *Annals of Sex Research, 3,* 245–291.

Laws, D. R. (Ed.) (1989). *Relapse prevention with sex offenders.* New York: Guilford.

Laws, D. R., & Marshall, W. L. (1991). Masturbatory reconditioning: An evaluative review. *Advances in Behaviour Research and Therapy, 13,* 13–25.

Lawson, J. S., Marshall, W. L., & McGrath, P. (1979). The social self-esteem inventory. *Educational and Psychological Measurement, 39,* 803–811.

Levin, S. M., & Stava, L. (1987). Personality characteristics of sex offenders: A review. *Archives of Sexual Behavior, 16,* 57–79.

Malcolm, P. B., Davidson, P. R., & Marshall, W. L. (1985). Control of penile tumescence: The effects of arousal level and stimulus content. *Behaviour Research and Therapy, 23,* 273–280.

Maletzky, B. M. (1980). Assisted covert sensitization. In D. J. Cox & R. J. Daitzman (Eds.), *Exhibitionism: Description, assessment, and treatment* (pp. 187–285). New York: Garland STPM Press.

Maletzky, B. M. (1985). Orgasmic reconditioning. In A. S. Bellack & M. Hersen (Eds.), *Dictionary of behavior therapy techniques* (pp. 157–158). New York: Pergamon.

Marlatt, G. A., & Gordon, J. R. (Eds.) (1985). *Relapse prevention.* New York: Guilford.

Marshall, W. L. (1979). Satiation therapy: A procedure for reducing deviant sexual arousal. *Journal of Applied Behavior Analysis, 12,* 10–22.

Marshall, W. L. (1992). The social value of treatment and sexual offenders. *Canadian Journal of Human Sexuality, 1,* 109–114.

Marshall, W. L. (1993). The role of attachment, intimacy, and loneliness in the etiology and maintenance of sexual offending. *Sexual and Marital Therapy, 8,* 109–121.

Marshall, W. L. (1994). Treatment effects on denial and minimization in incarcerated sex offenders. *Behaviour Research and Therapy, 32,* 559–564.

Marshall, W. L. (1996a). *Phallometric testing with sexual offenders: Limits to its value.* Manuscript submitted for publication.

Marshall, W. L. (1996b). The sexual offender: Monster, victim, or everyman? *Sexual Abuse: A Journal of Research and Treatment, 8,* 317–335.

Marshall, W. L. (1997). The relationship between self-esteem and deviant sexual arousal in nonfamilial child molesters. *Behavior Modification, 21,* 86–96.

Marshall, W. L. (in press). Sexual preferences: Are they useful in the assessment and treatment of sexual offenders? In D. Fisher, M. Cardgo, & B. Print (Eds.), *Sex offenders: Toward improved practice.* London: Whiting & Birch.

Marshall, W. L., & Anderson, D. (1996). An evaluation of the benefits of relapse prevention programs with sexual offenders. *Sexual abuse: A Journal of Research and Treatment, 8,* 209–221.

Marshall, W. L., Anderson, D., & Champagne, F. (1996). The importance of self-esteem in sexual offenders. *Psychology, Crime and Law, 3,* 81–106.

Marshall, W. L., Bryce, P., Hudson, S. M., Ward, T., & Moth, B. (1996). The enhancement of intimacy and the reduction of loneliness among child molesters. *Journal of Family Violence, 11,* 219–235.

Marshall, W. L., Champagne, F., Sturgeon, C., & Bryce, P. (1997). Increasing the self-esteem of child molesters. *Sexual Abuse: A Journal of Research and Treatment, 9,* 321–333.

Marshall, W. L., & Christie, M. M. (1981). Pedophilia and aggression. *Criminal Justice and Behavior, 8,* 145–158.

Marshall, W. L., & Darke, J. (1982). Inferring humiliation as motivation in sexual offenses. *Treatment for Sexual Aggressives, 5,* 1–3.

Marshall, W. L., Eccles, A., & Barbaree, H. E. (1991). Treatment of exhibitionists: A focus on sexual deviance versus cognitive and relationship features. *Behaviour Research and Therapy, 29,* 129–135.

Marshall, W. L., Eccles, A., & Barbaree, H. E. (1993). A three-tiered approach to the rehabilitation of incarcerated sex offenders. *Behavioral Sciences and Law, 11,* 441–445.

Marshall, W. L., Fernandez, Y. M., Hudson, S. M., & Ward, T. (Eds.) (in press). *Sourcebook of treatment programs for sexual offenders.* New York: Plenum.

Marshall, W. L., & Hall, G. C. N. (1995). The value of the MMPI in deciding forensic issues in accused sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 7,* 203–217.

Marshall, W. L., Hudson, S. M., Jones, R. L., & Fernandez, Y. M. (1995). Empathy in sex offenders. *Clinical Psychology Review, 15,* 99–113.

Marshall, W. L., Jones, R. L., Ward, T., Johnston, P., & Barbaree, H. E. (1991). Treatment outcome with sex offenders. *Clinical Psychology Review, 11,* 465–485.

Marshall, W. L., O'Sullivan, C., & Fernandez, Y. M. (1996). The enhancement of victim empathy among incarcerated child molesters. *Legal and Criminological Psychology, 1,* 95–102.

McFall, R. M. (1990). The enhancement of social skills: An information processing analysis. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 311–330). New York: Plenum.

Miller, R. S., & Lefcourt, H. M. (1982). The assessment of social intimacy. *Journal of Personality Assessment, 46,* 514–518.

Miller, W. R., & Rollnick, S. (1991). *Motivational interviewing: Preparing people to change addictive behavior.* New York: Guilford.

Murphy, W. D. (1990). Assessment and modification of cognitive distortions in sex offenders. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 331–342). New York: Plenum.

Murphy, W. D., & Barbaree, H. E. (1994). *Assessments of sex offenders by measures of erectile response: Psychometric properties and decision making.* Brandon, VT: Safer Society Press.

Nichols, H. R., & Molinder, I. (1984). *Multiphasic sex inventory manual.* Tacoma, WA: Author.

Pithers, W. D. (1990). Relapse prevention with sexual aggressors: A method for maintaining therapeutic gain and enhancing external supervision. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 343–361). New York: Plenum.

Pithers, W. D. (1994). Process evaluation of a group therapy component designed to enhance sex offenders' empathy for sexual abuse survivors. *Behaviour Research and Therapy, 32,* 565–570.

Pithers, W. D. (in press). Empathy: Definition, enhancement, and relevance to the treatment of sexual abusers. In W. L. Marshall & J.Frenken (Eds.), *North American and European approaches to research and treatment of sexual offenders.* Newbury Park, CA: Sage Publications.

Prentky, R. A., & Burgess, A. W. (1991). Rehabilitation of child molesters: A cost-benefit analysis. *American Journal of Orthopsychiatry, 60,* 108–117.

Proulx, J., Coté, G., & Achille, P. A. (1993). Prevention of voluntary control of penile response in a homosexual pedophile during phallometric testing. *Journal of Sex Research, 30,* 140–147.

Quinsey, V. L., & Chaplin, T. C. (1988). Penile responses of child molesters and normals to descriptions of encounters with children involving sex and violence. *Journal of Interpersonal Violence, 3,* 259–274.

Quinsey, V. L., & Earls, C. M. (1990). The modification of sexual preferences. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 279–295). New York: Plenum.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Lalumière, M. L. (1993). Assessing treatment efficacy in outcome studies of sex offenders. *Journal of Interpersonal Violence, 8,* 512–523.

Quinsey, V. L., Lalumière, M. L., Rice, M. E., & Harris, G. T. (1995). Predicting sexual offenses. In J. C. Campbell (Ed.), *Assessing dangerousness: Violence by sexual offenders, batterers, and child abusers* (pp. 116–137). Thousand Oaks, CA: Sage.

Quinsey, V. L., & Marshall, W. L. (1983). Procedures for reducing inappropriate sexual arousal: An evaluation review. In J. G. Greer & I. R. Stuart (Eds.), *The sexual aggressor: Current perspectives on treatment* (pp. 267–289). New York: Van Nostrand Reinhold.

Rathus, S. A. (1973). A thirty-item schedule for assessing assertive behavior. *Behavior Therapy, 4,* 398–406.

Rooth, F. G. (1973). Exhibitionism outside Europe and America. *Archives of Sexual Behavior, 2,* 351–363.

Russell, D. E. H. (1984). *Sexual exploitation: Rape, child sexual abuse, and workplace harassment.* Newbury Park, CA: Sage.

Schaap, C., Bennun, I., Schindler, L., & Hoogduin, K. (1993). *The therapeutic relationship in behavioral psychotherapy.* New York: Wiley.

Seidman, B. T., Marshall, W. L., Hudson, S. M., & Robertson, P. J. (1994). An examination of intimacy and loneliness in sex offenders. *Journal of Interpersonal Violence, 9,* 518–534.

Selzer, M. L. (1971). The Michigan Alcoholism Screening Test (MAST): The quest for a new diagnostic instrument. *American Journal of Psychiatry, 127,* 1653–1658.

Serin, R. C., Malcolm, P. B., Khanna, A., & Barbaree, H. E. (1994). Psychopathy and deviant sexual arousal in incarcerated sexual offenders. *Journal of Interpersonal Violence, 9,* 3–11.

Skinner, H. A. (1982). The drug abuse screening test. *Addictive Behaviors, 7,* 363–371.

Speilberger, C. D. (1988). *State-trait anger expression inventory: Professional manual, research edition.* Odessa, FL: Psychological Assessment Resources.

Spence, J. T., Helmreich, R., & Stapp, J. (1973). A short version of the Attitudes Towards Women scale (ATW). *Bulletin of the Psychonomic Society, 2,* 219–220.

Stermac, L. E., Sheridan, P. M., Davidson, A., & Dunn, S. (1996). Sexual assault of adult males. *Journal of Interpersonal Violence, 11,* 52–64.

Ward, T., Hudson, S. M., & McCormack, J. (1997). Attachment style, intimacy deficits, and sexual offending. In B. K. Schwartz & H. R. Cellini (Eds.), *The sex offender: New insights, treatment innovations and legal developments* (Vol. 11, pp. 2.1–2.14). Kingston, NJ: Civic Research Institute.

Ward, T., Hudson, S. M., Johnston, L., & Marshall, W. L. (1997). Cognitive distortions in sex offenders: An integrative review. *Clinical Psychology Review, 17,* 479–507.

Watson, D., & Friend, R. (1969). Measurement of social evaluation anxiety. *Journal of Consulting and Clinical Psychology, 33,* 448–457.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.26

Perpetrators of Physical Violence and Abuse

ABRAM ROSENBLATT and JENNIFER A. ROSENBLATT
University of California, San Francisco, CA, USA

9.26.1 INTRODUCTION 422
9.26.2 NATURE 422
  9.26.2.1 Types of Violence 423
  9.26.2.2 Predisposing Factors 423
    9.26.2.2.1 Biological factors 423
    9.26.2.2.2 Developmental factors 423
    9.26.2.2.3 Environmental factors 424
9.26.3 PREVALENCE 424
  9.26.3.1 Violence and Mental Disorder 425
  9.26.3.2 Mental Disorder Among Violent Offenders 425
  9.26.3.3 Violence and the Mentally Ill 425
  9.26.3.4 Summary 426
9.26.4 ASSESSMENT 427
  9.26.4.1 Offender Risk Assessments 427
    9.26.4.1.1 Violence Risk Appraisal Guide 427
    9.26.4.1.2 Hare Psychopathy Checklist-Revised 427
    9.26.4.1.3 Statistical Information on Recidivism Scale 427
    9.26.4.1.4 HCR-20 427
  9.26.4.2 Child Abuse Assessment Instruments 428
    9.26.4.2.1 Child Abuse Potential Inventory 428
    9.26.4.2.2 Michigan Screening Profile of Parenting 428
    9.26.4.2.3 Parenting Stress Index 428
  9.26.4.3 Summary 428
9.26.5 TREATMENT 428
  9.26.5.1 Characteristics of Effective Programs 429
  9.26.5.2 Characteristics of Ineffective Programs 429
  9.26.5.3 Programs for Treating Violence 429
    9.26.5.3.1 Therapeutic community program 429
  9.26.5.4 Medications 430
  9.26.5.5 Psychotherapy for Violent Outpatients 430
  9.26.5.6 Behavioral Techniques 430
  9.26.5.7 Treatment of Child Abusers 430
  9.26.5.8 Summary 431
9.26.6 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE 431
  9.26.6.1 Research on the Nature of Violence 431
  9.26.6.2 Research on the Prevalence of Violence 431
  9.26.6.3 Research on the Assessment of Violence 432
  9.26.6.4 Treatment of Violent Offenders 432
9.26.7 SUMMARY 432
9.26.8 REFERENCES 433

9.26.1 INTRODUCTION

Violence is a vexing problem that conjures a host of public perceptions, both real and imagined. A remarkable amount of the public debate concerning the delivery of human welfare services is driven by widespread fear of physical violence against both adults and children. The need to protect children from abuse clearly dominates the agenda of social welfare services. Parents, concerned about the safety of their children, may move from one community to the next in the hopes of finding a place where children can play without fear. For better or worse, it is the fear of potential violent behavior on the part of persons with mental disorders that mobilizes much of the public debate on mental health services. Current social policies regarding the incarceration of criminal offenders are largely driven by a desire to get violent criminals off the streets.

The prevalence rates of different types of physical violence vary over time. In the USA, reported murder rates increased from 5.1% in 1965 to 9.8% in 1991 (Federal Bureau of Investigation, 1965–1991). This particular crime statistic is alarming to both policy makers and the general public (Lowenstein, 1991). More recently, however, data from the Federal Bureau of Investigation (1996) revealed a 9% decrease in murder rates and an overall decrease in violent crime of 6% in 1996. In spite of such potentially encouraging figures, the rate of violent crime in the USA remains high, relative to other industrialized countries.

It remains to be seen whether decreasing rates of physical violence are actually decreasing as part of a longer-term trend. The perception on the part of the public of increasing violence is pushing lawmakers to enact new legislation designed to tackle the problem of violent crime. The most salient example of such legislation, the Violent Crime Control and Law Enforcement Act, was passed in 1994. This legislation allocated $22 billion to increase the number of police officers, augment prison budgets, increase the duration of prison sentences, and fund prevention programs. The prevention program dollars (US$5 billion), however, were shifted to law enforcement and prison expansion in a revision of the original act by the House of Representatives (Petersilia, 1995).

Although these measures are certainly popular, and rates of violent crime appear to be dropping slightly, most criminal justice experts agree that these tough on crime efforts will do little to decrease crime in the long term. Even some professionals who might be expected to support this legislation are skeptical. A national survey of prison wardens, for example, revealed that 85% of participants agreed that the new crime policy would not be effective (Petersilia, 1995). One reason for the lack of optimism regarding such federal efforts to curb crime is the predominant focus on prisons as the primary mode of intervention. A majority of offenders are not imprisoned; they remain in the community under probation or parole supervision. Less than 20% of all offenders are imprisoned and the majority of violent offenders are not in jail (US Bureau of Justice Statistics, 1995). Approximately 372 000 individuals convicted of a violent offense in 1991 were incarcerated, whereas approximately 590 000 persons convicted of a violent crime resided in the community.

Consequently, it is important to understand the problem of physical violence both because it is desirable to reduce the occurrence of such acts and because the fear of physical violence currently receives considerable attention in the creation of social policy. In the absence of empirical information, social policy is determined largely by public perceptions and the existing political climate. This chapter summarizes the empirical literature, the nature, prevalence, assessment, and treatment of persons who commit violent acts. The chapter focuses both on adults who commit violence toward other adults and on adults who commit violence towards children (commonly referred to as abuse). The problems of domestic violence and juvenile crime are addressed elsewhere in this volume. Finally, attention is focused on the special concerns posed by persons with mental disorders who commit violent acts because this is the group most likely to be seen by mental health professionals.

9.26.2 NATURE

There are potentially many types of violence and factors that influence whether a person may become violent. Understanding the nature of violence requires an understanding of the types of violence and the factors that might predispose an individual to commit violence. Although the end result of violence is the harming of another person, the legal system certainly recognizes that violence committed under different circumstances may be justifiable. Premeditated murder, manslaughter, and self-defense may all result in the death of a person, but punishments for these different types of violence range from the death penalty to being absolved of any blame.

9.26.2.1 Types of Violence

Much of the reason why an act of violence is considered justifiable can be traced to whether the violent act is reactive or instrumental in nature. Reactive aggression and violence are committed in response to a perceived threat or provocation, whereas instrumental acts of violence and aggression are goal directed (Dodge, 1991; Parke & Slaby, 1983). Studies suggest that individuals engaging in instrumental violence can be distinguished from those committing reactive violent crimes. Two studies examining violent offenders suggested that reactive offenders were more likely to harm familiar victims, be angry with their victims, and perceive their victims as provoking the violent act. Instrumental offenders tend to seek a clearly discernible goal (Cornell et al., 1996). Although offenders often engage in reactive and instrumental acts of violence, reactive incidents appear to be more pervasive. Cornell et al. (1996) suggested that individuals who engage in instrumental crime are more pathological than those who engage in reactive violence. These researchers contend that goal-directed acts of violence stem from a lack of concern and respect for others, and an absence of internalized social principles. Instrumental violence appears to be more associated with psychopathy than reactive violence.

9.26.2.2 Predisposing Factors

Although different types of people may commit different types of violence, a considerable body of research examines the range of factors that influence the likelihood that an individual may engage in any type of violent act. Much of this research can be grouped into three classes of factors that may increase violence potential: (i) biological factors; (ii) developmental factors; and (iii) environmental factors. Although examined separately here, violence is generally thought to derive from some combination of all three of these factors.

9.26.2.2.1 Biological factors

Biological factors, such as neurotransmitter levels, brain structure, and neurophysiology, may contribute to the potential for aggression and violence. Research on neurotransmitter levels indicates that increased levels of norepinephrine and dopamine (Eichelman, Elliot, & Barchas, 1981) and decreased levels of serotonin (Brown et al., 1982; Lidberg, Tuck, Asberg, Scalia-Tomba, & Bertilsson, 1985) are associated with aggression and impulsive violence. Abnormalities in brain structure, specifically in the temporal-frontal lobe, have been discovered in violent patients through the use of computer tomography and magnetic resonance imaging (Tonkonogy, 1991). Patients with extremely violent behavior also show metabolic abnormalities in the temporal lobes (Seidenwurm, Pounds, Globus, & Valk, 1997). Similarly, a study using brain single photon emission computerized tomography (SPECT) imaging found differences in brain structures in aggressive individuals as compared to individuals without reported violent behavior (Amen, Stubblefield, Carmicheal, & Thisted, 1996).

Historically, neurophysiological factors, particularly the limbic system, were thought to be associated with violent behavior (Mark & Ervin, 1970). Some more recent controlled studies, however, failed to corroborate these findings by discovering no differences in the level of violence for persons with and without epilepsy (Hermann & Whitman, 1984). This conclusion is supported in a collaborative large-scale study demonstrating that violent episodes were rare for epileptics during seizures (Delgado-Escueta et al., 1981). Although violence potential may be associated with biological factors, recent reviews of literature on genetic abnormalities suggested that violent behavior cannot be accounted for by chromosomal anomalies (Tardiff, 1996).

9.26.2.2.2 Developmental factors

In addition to the biological antecedents of violence, developmental factors also appear to be associated with increased violence. A large body of research suggests a positive relationship between childhood abuse and becoming a physically abusive adult (Kempe & Helfer, 1980; Rivera & Widom, 1990; Widom, 1989b). For example, a study comparing individuals with an official record of neglect or abuse to a matched control group of individuals with no record of abuse found that individuals victimized as children were at increased risk for committing a violent offense (Widom, 1989a). The relationship, however, between abuse and violence may be mediated by other factors. Weiler and Widom (1996), for instance, examined the relationship between childhood victimization, psychopathy, and violence and found that although childhood victimization predicted violence, psychopathy appeared to mediate this relationship. Children witnessing domestic violence may also have an increased chance for becoming violent adults (Jaffe, Wolfe, Wilson, & Zak, 1986; Pelcovitz & Kaplan, 1994). Although many studies suggest the existence of an intergenerational transmission of violence, Widom (1989b) has stressed that understanding of long-term consequences of abusive home environments is circumscribed by a paucity of acceptable empirical evidence.

9.26.2.2.3 Environmental factors

A range of environmental factors are associated with increased violence, including: low socioeconomic status and marital disruption (Messner & Tardiff, 1986); physical crowding and heat (Bell & Baron, 1981); and racial and economic inequality (Blau & Blau, 1982). Longitudinal research suggests that a range of individual, family, school, and peer factors may constitute antecedents of violent behavior. These factors include: (i) history of violence, threatening behavior, and the use of weapons; (ii) exposure to violence as a child; (iii) previous history of alcohol or drug abuse; (iv) history of psychological disorders (e.g., anxiety and depression); (v) antisocial behavior problems; (vi) low self-esteem and feelings of helplessness; (vii) social withdrawal, isolation and lack of social support; (viii) jealousy; (ix) recent aversive life event (e.g., death of a spouse, child, or parent); (x) financial difficulties (e.g., unemployment); and (xi) job dissatisfaction and stress (Browne & Herbert, 1996).

More temporally remote antecedents of violent behavior may exist in the link between childhood aggression and adult violence. The literature suggests that antisocial and aggressive behaviors in childhood are strong predictors of aggression in adulthood (Caspi & Moffit, 1995; Sampson & Laub, 1993). Aggressive children tend to engage in more violent acts as adults than do their nonaggressive counterparts (Jessor, Donovan, & Costa, 1991; Laub & Lauritsen, 1993; Reiss & Roth, 1994). Several factors associated with childhood aggression may, consequently, be associated with adult violence including: child temperament (White, Moffitt, Earls, Robins, & Silva, 1990); poor family functioning (Loeber & Stouthamer-Loeber, 1986); weak social bonds (Laub & Sampson, 1993); academic failure and truancy (Farrington, 1991); and peer rejection (Cairns & Cairns, 1992).

Protective factors may mediate the impact of risk factors. For example, most antisocial youths do not become violent adults. It may be the case that a subgroup of boys with the most extreme behavior problems are those most likely to remain aggressive and antisocial in adulthood (Caspi & Moffitt, 1995). This general hypothesis was explored by Sampson and Laub (1993) who examined the incidence of violent crimes in adulthood for a group of delinquent youths and a group of nondelinquent youths. They found that 33% of a sample of delinquent youths were arrested for a violent crime at ages 17–25, and 18% were arrested at ages 25–32. None of the nondelinquent controls were arrested for a violent crime as adults. Interestingly, these researchers also found that strong work and family connections in adulthood led to decreased crime in both the delinquent and control groups. Thus, social ties, such as marital attachment and job stability, may be mitigating factors in determining whether aggression and antisocial behavior in childhood leads to violent crime in adulthood (Sampson & Laub, 1993). These types of studies have led some researchers to suggest that prevention of antisocial behavior in childhood can reduce violent behavior in adulthood (Farrington, 1991; Laub & Lauritsen, 1993).

9.26.3 PREVALENCE

An examination of the literature regarding the prevalence of violence reveals that although violent crime committed against adults receives considerable attention because of its frightening implications for personal safety, a relatively small proportion of adults engage in violent crimes against adult family members, strangers, and acquaintances. The homicide rate in the USA is 10 per 100 000 individuals. Homicide rates are substantially lower in other countries: 1.1 per 100 000 in Japan; 1.3 per 100 000 in the UK; and 2.3 per 100 000 in Australia (Browne & Howells, 1996). However, homicide is the least frequent type of violent offense. Arrest data from the US Department of Justice reveals that less than 3% of the arrests for violent crime in 1994 were for murder and manslaughter. The most common violent crime was burglary, followed by robbery and aggravated assault. Arrests for violent crimes as a whole represented only 5.4% of the offenses charged for that year (US Department of Justice, 1994).

Sadly, violence against children in the form of abuse is relatively more common. According to a Gallup survey, three million of the approximately 77 million children in the USA (4%) were victims of physical abuse in 1995 (Moore, 1995). Reported rates of abuse do, however, vary considerably. A recent survey of 9953 individuals in Ontario, Canada, revealed that 31% of men and 21% of women self-reported physical abuse and 11% of those men and 9% of those women reported severe physical abuse by an adult during their childhood (MacMillan et al., 1997).

Rates of violent offenses vary by demographic characteristics. Males have much higher rates of violent offending than females (Federal

Bureau of Investigation, 1993). Of those arrested for violent crime in 1994, 84% were male (US Department of Justice, 1995). Over half of those arrested for violent crimes were Caucasian (54%) and virtually all of the rest were African-American (43%). Very few American-Indians or Asian/Pacific Islanders were arrested for violent crimes (less than 1% for both groups). Fewer Caucasians were arrested for robbery (38%) and murder (42%) than African-Americans (US Department of Justice, 1995). These figures must be considered in the context of the ethnic distribution in the USA as a whole, where 83% of the residents were Caucasian in 1994 and 12% were African-American (US Bureau of the Census, 1994).

9.26.3.1 Violence and Mental Disorder

The relationship between psychiatric disorders and violence is especially important to mental health professionals. There are considerable misperceptions about the violent nature of persons with psychiatric disorder. The psychopathic killer freshly escaped from the mental institution is a popular character in movies and television. Contrary to some popular portrayals, all people with mental disorders are not violent and all people who commit violent acts do not have a mental disorder. Fortunately, there is a growing knowledge base regarding the prevalence of mental disorders among violent offenders and the prevalence of violence among individuals with mental disorders.

9.26.3.2 Mental Disorder Among Violent Offenders

The estimated prevalence rates of psychiatric disorders among criminals vary. Brodsky (1972), in a review of studies of prison and court samples from 1918 to 1970, found the rate of psychiatric disorder in criminals to range from 16% to 95%. Hodgins and Cote (1990) used the *Diagnostic and statistical manual of mental disorders,* third edition (*DSM-III*; American Psychiatric Association, 1980) and found the prevalence of schizophrenia in penitentiary inmates to be seven times higher than individuals in the general population. However, in a review of 18 studies of mental disorders among jail inmates, Teplin (1990) found rates of psychotic mental disorders to range from 5% to 16%. Steadman, Fabisiak, Dvoskin, and Holohean (1987) determined that 8% of 3000 prisoners had severe mental illness and another 16% of inmates had significant mental disabilities.

Variation in rates of mental disorders in these prison samples are due, in part, to unstandardized methods of diagnosis (Monahan, 1992). More recent studies utilize structured diagnostic interviews such as the Diagnostic Interview Schedule (DIS). Teplin (1990) compared results from the DIS of 728 men from a county jail and 362 male prison inmates (Hodgins & Cote, 1990) to data from the National Institute of Mental Health's epidemiological catchment area (ECA) study (Robins & Regier, 1991). Results indicated that the prevalence of a severe disorder was 3–4 times greater in the prison and jail samples than in the general population. Specifically, the prevalence of major depression was 3–4 times higher, mania or bipolar disorder was 7–14 times higher, and schizophrenia was three times higher in inmates than in individuals from the general population.

Although these findings indicate a high prevalence of mental disorder among males in jails and prisons, sampling biases exist. Violent offenders with mental illness are more likely to be arrested and imprisoned than offenders without mental illness. Robertson (1988), for example, found that perpetrators with schizophrenia had increased likelihood for arrest at a crime scene than individuals without disorders. Inmates may be a subset of perpetrators and, thus, are not necessarily representative of the population of violent offenders.

9.26.3.3 Violence and the Mentally Ill

Findings from 11 studies of violence among individuals with mental disorders suggest that 10–40% of individuals sampled committed a violent act against another individual two weeks prior to hospitalization (Otto, 1992; Wessely & Taylor, 1991). In addition, Klassen and O'Connor (1990) found that 25–30% of male psychiatric subjects with at least one past incident of violence committed a violent act within a year of discharge from a hospital. The MacArthur Risk Assessment Study, an ongoing longitudinal project analyzing 1000 individuals admitted to psychiatric hospitals, reported that 33% of patients reported one or more violent incidents within a period of six months after hospital discharge (Steadman et al., 1994). Lidz, Mulvey, and Gardner (1993) found 45% of 712 patients engaged in violence within six months of a psychiatric emergency room visit.

Although these studies suggest a positive relationship between violence and mental illness, fundamental methodological problems exist. For instance, very few studies assess rates of violence in comparative samples of nonhospitalized individuals. Selection bias also exists in studies assessing violence in hospitalized samples in that these patients are only a subset of the population of mentally ill. Patients with violent behavior are likely to be hospitalized due to the danger to others standard for civil commitment (Monahan, 1992). Violence may well be the primary reason why people with mental illness are hospitalized.

There is some evidence that rates of violence among psychiatric patients may be rising, especially for females. Disturbingly, Tardiff, Marzuk, Leon, Portera, & Weiner (1997) found that, compared to a methodologically similar study conducted a decade earlier at the same clinic (Tardiff, 1984), the rate of violence among psychiatric patients increased 40% for males and 150% for females. The increase in the violence rates of psychiatric patients may be related to increased substance abuse (Tardiff et al., 1997).

A study by Swanson, Holzer, Ganju, and Jono (1990) overcame several of the weaknesses found in prior studies by investigating the relationship between violence and mental illness in a community-based sample. Swanson et al. (1990) assessed psychiatric disorder and incidents of violence in 10 000 residents in three American cities using data from the National Institute of Mental Health's ECA study. *DSM-III* diagnoses were derived from a standardized interview and participants self-reported violent acts committed in the previous year. Higher rates of violence were positively associated with all psychiatric disorders. Specifically, 2.1% of individuals without a psychiatric diagnosis reported at least one violent incident. The percentage of individuals with a psychiatric diagnosis who reported at least one violent episode in the prior year was much higher, ranging from 11% to 35% depending on the type of diagnosis. Persons receiving a diagnosis of drug or alcohol abuse or dependence had the highest incidence of self-reported violence (35% for drug and 24.6% for alcohol). Rates of self-reported violence in other diagnostic groups were much lower and fairly consistent, including the diagnoses of schizophrenia (12.7%), panic disorder (11.6%), major depression (11.7%), and mania or bipolar disorder (11.0%).

A study by Link, Andrews, and Cullen (1992) corroborated the findings of Swanson et al. (1990). Link et al. (1992) compared arrest rates and self-reported violence in individuals formerly hospitalized and treated, with arrest and violence data from a sample of 400 adults who had never been treated for a mental illness. After controlling for age, gender, ethnicity, socioeconomic status, educational level, homicide rate where the participant lived, and social desirability, researchers found that the treated sample were 2–3 times more violent than the never-treated group. Link et al. (1992) also controlled for psychotic symptoms using the psychotic symptomatology scale. They found that almost all of the differences in rates of violence were accounted for by active psychotic symptoms: the rates of violence in treated patients were similar to untreated individuals when they were not experiencing any psychotic symptoms. In addition, treated and untreated individuals who were actively experiencing psychotic symptoms had a higher rate of violence than those with no psychotic symptoms. These data suggest that individuals who are actively experiencing psychotic symptoms have a greater likelihood for engaging in violent behavior than those nondisordered individuals in the community.

Other research findings also point to a link between mental illness and dangerousness, especially with reference to psychopaths (Kosson, Smith, & Newman, 1990; Serin, 1991, 1996). Psychopathy may be a marker for violence and violent reoffending. Although not every psychopath is a violent offender and, conversely, not every violent offender is a psychopath, research suggests that psychopathic violent offenders have higher rates of violent recidivism (Quinsey, 1995).

The preponderance of violence among persons who abuse alcohol and other drugs has also been found for violence against children (Chaffin, Kelleher, & Hollenberg, 1996; Dinwiddie & Bucholz, 1993; Kelleher, Chaffin, Hollenberg, & Fischer, 1994; Murphy et al., 1991). Kelleher et al. (1994) using ECA data found that 43% of parents who physically abused their children met the criteria for a diagnosis of substance abuse disorder.

9.26.3.4 Summary

Overall, rates of violent offenses are low relative to other crimes, although abuse appears more common. Violence does vary by gender and age, with men and African-Americans having higher rates of arrests for violent offenses. Research does indicate that mental disorder may be a significant risk factor for violent offenses (Mullen, 1997; Wessely & Taylor, 1991). It is important to note that a psychiatric diagnosis alone may not necessitate risk; existence of active psychotic symptoms may be the link to violence (Link et al., 1992). Further, the majority of people with psychiatric disorders are not violent: only 10% of individuals with psychiatric diagnoses self-reported a violent incident in the previous year in the ECA study (Robins & Regier, 1991). Monahan

(1992) provided an eloquent summary of the relationship between violence and mental disorder:

> Evidence now indicates that mental disorder may be a consistent, albeit modest, risk factor for the occurrence of violence. Dire implications for mental patient advocacy, for mental health law, and for the provision of mental health treatment need not follow from candidly acknowledging the possibility of a limited connection between disorder and violence. (p. 511)

9.26.4 ASSESSMENT

Data regarding the prevalence of mental disorder does not, alone, predict the violence potential of any given individual, just as no single characteristic studied at the epidemiologic level can predict future violence. Nonetheless, mental health professionals may be asked by correctional and parole boards to assess the risk of violent reoffending for known violent perpetrators (Serin, 1996), or to assess the violence potential of persons receiving mental health care. There has been some hope that reliable and valid psychiatric and psychological assessments could accurately predict violent behavior. Historically, methods for predicting violent behavior have been poor (Lidz et al., 1993; Menzies, Webster, McMain, Staley, & Scaglione, 1994; Quinsey, 1995; Quinsey & Maguire, 1986). Recent research, however, shows more promise for accurate prediction of violent reoffending (Bonta, Harman, Hann, & Cormier, 1996; Harris, Rice & Quinsey, 1993). The following instruments aid in assessing an offender's propensity to commit a violent offense, including child physical abuse (Blackburn, 1993).

9.26.4.1 Offender Risk Assessments

9.26.4.1.1 Violence Risk Appraisal Guide

The Violence Risk Appraisal Guide (VRAG; Harris et al., 1993) is an empirically derived instrument used for assessing risk of future violence. The 12 items include: psychopathy; separation from parents at age 16 or younger; victim injury in index offense; schizophrenia; marital status; elementary school maladjustment; female victim in index offense; failure on prior conditional release; property offense history; age at index offense; alcohol abuse history; and a diagnosis of personality disorder. The VRAG was developed on a sample of individuals with at least one documented serious violent offense (Borum, 1996). The VRAG classification accuracy rate is approximately 75% and Rice (1997) has argued that this actuarial instrument can fairly accurately predict long-term criminal violence.

9.26.4.1.2 Hare Psychopathy Checklist-Revised

The Psychopathy Checklist-Revised (PCL-R) is frequently used to predict recidivism in criminals (Serin, Peterson, & Barbaree, 1990) and psychiatric patients (Harris et al., 1993; Quinsey, Rice, & Harris, 1995). Empirical analyses have yielded two correlated factors. The first factor encompasses personality characteristics including lack of remorse, pathological lying, conning, and callousness. The second factor includes lifestyle indicators such as criminal versatility, impulsiveness, and boredom (Hare et al., 1990). Serin (1996) found that the PCL-R predicted violent recidivism more accurately than prior history of violence and actuarial risk scales. Findings from several other studies attest to this instrument's predictive validity (e.g., Harris, Rice, & Cormier, 1991; Serin & Amos, 1995). The PCL-R does require more time, training in scoring, and clinical skills than many other prediction scales.

9.26.4.1.3 Statistical Information on Recidivism Scale

The Statistical Information on Recidivism (SIR) Scale was developed to inform parole boards of the future violence potential of federally sentenced inmates in Canada (Nuffield, 1989). Fifteen variables were empirically derived and related to general recidivism in a sample of over 2500 inmates with a follow-up period of three years. The SIR scale yields five categories of risk, ranging from "very good risk" to "poor risk." The SIR scale was recently revalidated using 3267 male offenders from federal penitentiaries (Bonta et al., 1996). The SIR scale was used to assess violent risk as well as risk of violence in general. Results indicated that 18.6% of the offenders committed a violent crime during the follow-up period. A statistically significant relationship between the risk categories and violent recidivism was detected. A univariate correlation of 0.20 was found for SIR scores and future violent behavior. Bonta et al. (1996) suggested that the SIR predicts violent behavior beyond levels of chance.

9.26.4.1.4 HCR-20

The HCR-20 is a three-level checklist designed to assess risk of future violent behavior among criminal and psychiatric individuals (Webster, Eaves, Douglas, & Wintrup, 1995). The first 10 items pertain to historical variables (H scale) including previous violence, age at first violent offense, relationship stability, employment stability, alcohol or drug abuse, mental disorder, psychopathy, early home and school maladjustment, personality disorder, and prior release or detention failure. The second five items are clinical in nature (C scale), including insight, attitude, symptomatology, stability, and treatability. The final five items pertain to risk variables (R scale) in the community or situations in which the individual may be involved: plan feasibility, access, support and supervision, compliance, and stress (Webster et al., 1995). Preliminary data investigating the psychometric properties of the HCR-20 appear promising (Harris et al., 1993).

9.26.4.2 Child Abuse Assessment Instruments

9.26.4.2.1 Child Abuse Potential Inventory

The Child Abuse Potential Inventory (CAP) is a self-report questionnaire widely used for screening suspected physical child abusers (Milner, 1986, 1994). The 160 agree–disagree items comprise a physical abuse scale and six descriptive factor scales. The factor scales include the following: distress, rigidity, unhappiness, problems with child and self, problems with family, and problems from others. The CAP Inventory also contains three validity indexes: faking-good index, faking-bad index, and random response index. A significant relationship between elevated abuse scores on the CAP Inventory and later physical child abuse has been reported (Milner, Gold, Ayoub, & Jacewitz, 1984). In addition, overall abuse scale classification rates are reported to range from 80 to 90% (Caliso & Milner, 1994; Milner, 1989, 1991; Milner & Robertson, 1989).

9.26.4.2.2 Michigan Screening Profile of Parenting

The Michigan Screening Profile of Parenting (MSPP; Helfer, Hoffmeister, & Schneider, 1978) was designed as a screening tool for child abuse and/or neglect. The MSPP consists of 77 items and four factors: emotional needs met, relationship with parents, expectations of children, and coping. The "emotional needs met" factor is used to detect physically abusive parents. Convergence analysis classifies parents as abusive through measurement of inconsistent responding (Hefler et al., 1978). Although sensitivity rates for child physical abusers are adequate, Schneider (1982) reported high rates (20%) of false positive classifications. Hefler et al. (1978) suggested that the MSPP is more suited for identification of problems in parenting than a screen for child abuse.

9.26.4.2.3 Parenting Stress Index

The Parenting Stress Index (PSI; Abidin, 1990) is used to measure child- and parent-related stress in physically abusive parents and those parents at risk for physical abuse, although the measure was not specifically designed for this purpose (Milner & Murphy, 1995). The PSI consists of 120 items that assess parent-related stress, child-related stress, and general life stress. PSI scores, specifically the "parent domain stress" scores, have been correlated with child abuse potential (Milner, 1986). Data suggest that parents at risk for physical abuse as well as parents who are physically abusive score higher on the PSI than comparison parents (Abidin, 1990; Mash, Johnston, & Kovitz, 1983).

9.26.4.3 Summary

A wide range of instruments are available for assessing violence potential, some with impressive psychometric properties. Scales with classification rates above 75% exist for predicting both violence and abuse. One of the difficulties in utilizing such scales is that consequences for false negatives can be extremely high if a person is deemed to be at little risk for future violence and still commits a violent offense. Nonetheless, reasonable tools for the assessment of violence and violence potential do exist.

9.26.5 TREATMENT

The literature on the treatment of violent offenders is not as encouraging as the literature on assessment. The existing literature on the treatment of violent offenders is relatively limited. More extensive literature exists regarding the treatment of criminal offenders in general. Consequently, many of the treatment principles and methodological problems that are discussed in this section derive at least in part from research stemming from a wide range of criminals.

There is a prevailing belief that imprisonment is the only viable solution for criminals involved in certain types of crime. Yet, jail time is known to have deleterious effects, frequently increasing the likelihood that the individual will reoffend (Lipsey, 1992). In fact, even the penalty of death does not appear to dramatically reduce serious crime. A study involving states with and without the death penalty found no differences in terms of serious crime rates (Hollin, 1996).

During the 1970s, researchers suggested that existing treatments were not effective on the recidivism of offenders (Brody, 1976; Greenberg, 1977; Martinson, 1974). Indeed, Lipton, Martinson, and Wilks (1975) reviewed 231 studies of the treatments for offenders and concluded that interventions studied had no differential effects. This conclusion was substantiated by a rigorous article review commissioned by a National Academy of Sciences panel (Sechrest, White, & Brown, 1979).

Recent research, however, points to more positive findings regarding the effectiveness of treatment for criminal offenders (Andrews et al., 1990; Gendreau, 1996; Gendreau & Ross, 1987). Treatments do exist that appear to reduce the risk of recidivism at higher rates than punishment alone (Bonta, 1996). These treatments are enhanced through the use of empirically based risk/need assessment instruments designed to aid in planning specific treatment programs that match the level of offender risk (Bonta, 1996).

9.26.5.1 Characteristics of Effective Programs

Several recent reviews identify the characteristics of successful treatment programs. Lipsey (1992) reviewed 443 programs and found that 64% of the studies reported higher reductions in recidivism rates for the treatment group as compared to the control group. The average recidivism rate was 10%. A review of approximately 500 control group studies on treatment programs for offenders yielded several common characteristics for interventions that have reduced recidivism by 25–80% (Andrews et al., 1990). Gendreau and Paparozzi (1995) provided a summary of the principles of effective interventions for criminal offenders:

(i) 3–9 months of intensive services based on cognitive and social learning theories;

(ii) services target antisocial attitudes and values;

(iii) treatment style and mode is matched to offender's learning style;

(iv) positive reinforcement is used more frequently than punishment;

(v) clinicians are appropriately trained and supervised to relate to offenders in a sensitive and constructive manner;

(vi) offenders are exposed to prosocial activities that interrupt the criminal network.

Gendreau (1996) has contended that effective interventions should include intensive behavioral services that account for 40–70% of an offender's time. Suggested behavioral programs include token economies, modeling, and cognitive behavioral therapy. Criminogenic needs of high-risk offenders should be targeted. Criminogenic needs are dynamic risk factors that are amenable to change and predict criminal behavior (Andrews & Bonta, 1994). If criminogenic factors can be changed, the probability of future criminal behavior can be reduced (Bonta, 1996). These needs include antisocial attitudes, peer associations, issues of self-control, chemical dependencies, and behavioral styles (Andrews & Bonta, 1994).

9.26.5.2 Characteristics of Ineffective Programs

Gendreau and Paparozzi (1995) and Gendreau (1996) have summarized the characteristics of programs that failed to reduce recidivism:

(i) treatment interventions based on Freudian psychodynamic and Rogerian nondirective theories;

(ii) medical model approaches such as dietary alterations, plastic surgery, and pharmacological treatments;

(iii) boot camps;

(iv) services targeting noncriminological issues such as anxiety, self-esteem, and depression;

(v) "punishing smarter" strategies, including intensive supervision programs (ISPs). ISPs typically include home confinement, drug testing, restitution, shock incarceration, boot camps, and electronic monitoring.

9.26.5.3 Programs for Treating Violence

Although the characteristics of effective programs for criminal offenders can be applicable to violent offenders, there are several programs designed specifically for the subgroup of criminal offenders who are violent. Several of the most visible and well-researched programs for violent offenders are reviewed below. It is evident that the data on the effectiveness of these programs are mixed.

9.26.5.3.1 Therapeutic community program

The Oak Ridge Division of the Mental Health Centre in Ontario, Canada began a treatment program in the 1960s and 1970s that attempted to reduce the incidence of violence of participants after release. This program treated psychotic and nonpsychotic patients with up to 80 hours of weekly structured therapeutic activity. The program was designed to increase cooperation, caring, empathy, responsibility, and insight (Rice, 1997). The evaluation study included 146 men treated in the therapeutic program for at least two years and 146 men matched on age, offense type, and criminal history who were in prison. Violent recidivism was examined after 10 years and, unfortunately, was 40% for both untreated and treated participants.

A significant finding was discovered, however, when data from psychotic and nonpsychotic participants were analyzed. A large interaction effect indicated that the treated psychopaths had significantly higher recidivism rates than the untreated psychopaths, and the treated nonpsychopaths had significantly lower recidivism rates than the untreated nonpsychopaths (Harris, Rice, & Cormier, 1994). The program evaluators speculated that the treatment program taught psychopaths how to better manipulate and deceive others while increasing their self-esteem, leading to increased levels of aggression (Baumeister, Smart, & Boden, 1996). These findings suggest that therapeutic programs are less effective in decreasing violent behavior in psychopathic offenders than other criminals.

9.26.5.4 Medications

Prescribed medication is frequently used for violent psychiatric patients. The type of medication prescribed is dependent on the etiology of the violent behavior (Tardiff, 1996). For example, antipsychotic medications (haloperidol, clozapine, and fluphenazine) are used to treat delusional thinking and aid in the management of violence stemming from organic disorders. Clozapine (Buchanan, 1995) and carbamazepine (an anticonvulsant; Neppe, 1983) have both been shown to reduce symptoms of violence and aggression in schizophrenics. Beta blockers have been effective in managing violent behavior in individuals with traumatic brain injury and psychosis (Barratt, 1993). Lithium has been found to be effective in reducing violence in patients with bipolar disorder or mental retardation (Tardiff, 1996). Data suggest that serotonergic medications (e.g., tryptophan and fluoxetine) may reduce violent episodes of patients with schizophrenia and personality disorders (Goldman & Janecek, 1990; Wilcock, 1987).

9.26.5.5 Psychotherapy for Violent Outpatients

Family and group therapy are the most utilized forms of psychotherapy for violent patients. Family therapy is imperative because most violence occurs at home. Group therapy allows patients to discuss issues and interact with individuals experiencing similar violent and aggressive behaviors. Tardiff (1996) has delineated eight objectives of psychotherapy for violent patients:

(i) evaluate the patients' motivations and reasons for participating in psychotherapy;

(ii) facilitate expression of patients' problems and issues in a nonjudgmental and empathetic manner;

(iii) encourage patients' self-control;

(iv) continual awareness of feelings of transference and countertransference;

(v) develop patients' ability to recognize their own anger;

(vi) increase awareness of dynamics of violence;

(vii) increase awareness of consequences of behavior;

(viii) aid in restructuring patients' environments to prevent violence.

9.26.5.6 Behavioral Techniques

Behavior modification is known to be an effective tool in the management of mentally impaired violent patients. Effective behavioral programs include adequately trained staff, a specific treatment program, clear definitions of target behaviors and responses, positive reinforcement, improvement of social skills, withdrawal of positive reinforcement, sensory extinction, time-outs, and contingent restraint (Liberman & Wong, 1984; Tardiff, 1996). Social skills training programs, specifically, have been found to increase interpersonal and problem solving skills and social functioning (Liberman & Corrigan, 1993).

Programs involving behavioral techniques coupled with prescribed medication appear promising for individuals with mental illness. An example is biobehavioral therapy (BBT) which combines behavioral assessment, social learning techniques, family participation, and social skills training with pharmacotherapy (Liberman, 1994; Liberman, Kopelowicz, & Young, 1994).

9.26.5.7 Treatment of Child Abusers

There is a fairly extensive literature on the treatment of child abuse. A range of family therapies and other interventions are designed, at least in part, to prevent or remediate child abuse. In general, positive results have been achieved through parent-focused interventions (including behavioral and cognitive-behavioral programs) that focus on improving childrearing practices and management of stress (reviewed by Wolfe & Wekerle, 1993). Positive outcomes have also been found for physically abusive parents involved in individual behavior therapy and parenting classes, coupled with in-home behavioral coaching (Howing, Wodarski, Gaudin, & Kurtz, 1989). Family preservation services are frequently used as an intervention for families where child abuse exists. Although early research in this area supported the efficacy of these services, more recent research questions the effectiveness, especially in the case of parents with psychopathology and substance use issues (Dore & Alexander, 1996).

9.26.5.8 Summary

The literature on the effectiveness of treatment programs for criminal offenders, in general, and violent offenders, in particular, is plagued by a number of methodological and other problems. Research in the field of criminology does not always build from a strong empirical foundation (Andrews & Bonta, 1994). Andrews and Wormith (1989) stated that research findings not supportive of preferred theories of criminology have been frequently ignored, whereas weak findings supporting traditional theories have been well received. There is also a dearth of evaluation studies using robust experimental designs (Petersilia & Turner, 1993). Random assignment of offenders into treatment groups is difficult in that criminal justice professionals lose control of their decision-making regarding placement. This is especially problematic when the experimental condition is a less severe form of punishment (Petersilia & Turner, 1993).

Another methodological issue that has plagued criminology research is the evaluation of program success. Outcomes are usually measured through perpetrator recidivism, typically assessed by arrest statistics and parole violations. These outcome variables are reliant on the capability and efficiency of law officials in apprehending offenders. Evaluation of community programs is especially problematic in that offenders are subjected to increased surveillance. Indeed, offenders may be arrested more frequently simply due to the increased chances of detection of transgressions (Cullen, Wright, & Applegate, 1996).

Unfortunately, few studies have evaluated community-based programs for violent offenders due, in part, to the more serious nature of the crimes committed. Fortunately, some programs, such as home confinement and electronic monitoring, are beginning to admit more high-risk offenders, including individuals involved in crimes against persons (Cullen et al., 1996). Evaluation of these programs for violent offenders is still in its infancy, but preliminary data suggest that high-risk individuals may be adequately maintained in the community (Austin & Hardyman, 1991).

It is generally agreed that community-based programs rather than incarceration hold the answer to reducing crime in the future (Reiss & Roth, 1994). However, considerable additional research is required before communities will feel confident of the potential of these programs to curb violent behavior.

9.26.6 FUTURE DIRECTIONS FOR RESEARCH AND PRACTICE

Over the last decade, a range of research effort has shed light on many aspects of violence and abuse. Models regarding the nature and etiology of violence are increasingly demonstrating the interplay between environmental and biological factors to violent behavior. Epidemiological research suggests that there is a relationship between violence and mental disorder. Assessment tools exist that predict violence and abuse with relatively high levels of accuracy. Treatment programs for criminal offenders are appearing more promising than they were a decade ago. Nonetheless, much remains to be learned about violent behavior in each of the central areas elucidated in this chapter.

9.26.6.1 Research on the Nature of Violence

Research on the nature of violence is characterized by advances in the understanding of biological and environmental risk factors for violent behavior. However, relatively limited research has examined the interactions between biological and environmental risk factors. Specifically, little is known about the relative contributions of biology and environment to violent behavior. Much more work is needed on the long-term consequences of abusive home environments for the development of violent behavior later in life (Widom, 1989b). Although risk factors for violent behavior are increasingly researched, protective factors are receiving relatively less attention. Understanding these protective factors is especially important for developing successful interventions for preventing violent behavior. Monahan and Steadman (1994) have contended that future violence prediction models should include the disaggregation of the violence predictor variables into (i) the risk factors, (ii) the amount and type of violence predicted, and (iii) the likelihood that a harmful event will occur.

9.26.6.2 Research on the Prevalence of Violence

The epidemiological studies on violence that address the problem of violence and mental illness derive largely from one major epidemiological study conducted a decade ago (the National Institute of Mental Health's ECA study; Robins & Regier, 1991). That study did not extensively sample individuals in prisons or institutional care. Consequently, assessments of past and risk of future violence in large representative samples of individuals in the community, psychiatric hospitals, and prisons are needed to augment extant epidemiological studies. Although such studies can be costly, epidemiological research can be conducted across a range of special populations (e.g., the homeless mentally ill, prison inmates, psychiatric inpatients, incarcerated mentally disordered offenders). Findings can help to more fully develop the relationships between mental disorder, violence, and a range of risk and protective factors. With regard to child abuse, understanding of the actual prevalence rates of physical abuse remains relatively unknown. Rates currently vary depending on the sample and the manner in which questions regarding abuse are asked.

9.26.6.3 Research on the Assessment of Violence

Adequate instruments exist for assessing violence potential. However, the utility of these tools largely rests on their sensitivity rather than their specificity. In general, it is better to intervene with an individual who is assessed as having a high potential for violence but does not actually commit a violent act, than to fail to intervene with an offender who commits an act of violence. Of course, false positives also have negative consequences with regard to social stigma for the person assessed as a high risk for violence, especially in the case of abuse. Nonetheless, more research on developing tests sensitive enough to merit confidence in their use remains a viable avenue for future investigation.

9.26.6.4 Treatment of Violent Offenders

More research is needed on how to treat violent offenders. Existing literature focuses on criminal offenders, in general, with relatively less attention on violent offenders. Clearly, this lack of research is due in part to the desire to punish rather than to treat violent offenders. Furthermore, the consequences for ineffective treatment of violence are dire in that failure results in physical assault or abuse. Nonetheless, the large numbers of violent offenders who remain in the community, paired with high rates of recidivism among those who have been imprisoned, begs for more effective community-based interventions. The only other alternative is longer-term incarceration, which is costly and appears ineffective. Finally, very little is known about the process of breaking the cycle of abuse and violence found in many families (Widom, 1989b). Rigorous, creative, controlled studies need to be conducted to examine how to effectively treat the problem of violence for those living in the community as well as those who are incarcerated. Of all the areas discussed in this chapter, this topic requires the most thorough attention.

9.26.7 SUMMARY

The facts regarding violent behavior appear to be clouded by misperceptions. Accurate data on violent criminals are particularly necessary in light of frequent public misconceptions. Partially due to the media's sensationalism of violence, the public tends to associate violence with the majority of offenders (Bonta & Hanson, 1994) and individuals with mental disorders. In reality, the base rate for violence is less than 10% in criminals (Bonta et al., 1996) and individuals with psychiatric disorders (Monahan, 1992).

Nonetheless, the negative consequences of violent behavior exceed those of virtually any other type of behavior and extend beyond those who are actually harmed. Fear of violence and perceptions of the loss of safety intrude on the lives of those who have never been victimized by a violent act. In this type of environment, there is a strong desire to treat violence through punishment. However, there is little evidence that punishment will actually solve the problem of violence. Consequently, punishment should be combined with other interventions to reduce recidivism.

Perhaps more important are the implications that derive from the assertion of Farrington (1990, p. 110) that "problem children tend to grow up into problem adults, and that problem adults tend to produce problem children." The longer-term solutions to the problem of violence may constitute preventing violent behavior by identifying and intervening with those most at risk. Perpetrators of violence have heterogeneous characteristics. Interventions tailored to the characteristics of individual offenders could lead to more successful outcomes. As practitioners and researchers, understanding of the differences inherent in offenders, knowledge of risk assessment instruments, and awareness of the efficacy of interventions will aid in improving treatment and decreasing recidivism and incidence of violence.

ACKNOWLEDGMENTS

Preparation of this chapter was supported in part by Training and Center Grants from the National Institute of Mental Health (MH43694 and MH18261). The authors wish to acknowledge the assistance of Rachel DuPre and Karla Kruse.

9.26.8 REFERENCES

Abidin, R. R. (1990). *Parenting Stress Index-Manual* (3rd ed.). Charlottesville, VA: Pediatric Psychology.

Amen, D. G., Stubblefield, M., Carmicheal, B., & Thisted, R. (1996). Brain SPECT findings and aggressiveness. *Annals of Clinical Psychiatry, 8,* 129–137.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.) Washington, DC: Author.

Andrews, D., & Bonta, J. (1994). *The psychology of criminal conduct*. Cincinnati, OH: Anderson.

Andrews, D., & Wormith, J. (1989). Personality and crime: Knowledge destruction and construction in criminology. *Justice Quarterly, 6,* 289–309.

Andrews, D., Zinger, I., Hoge, R., Bonta, J., Gendreau, P., & Cullen, F. (1990). Does correctional treatment work? A clinically-relevant and psychologically-informed meta-analysis. *Criminology, 28,* 369–404.

Austin, J., & Hardyman, P. (1991). *The use of early parole with electronic monitoring to control prison crowding: Evaluation of the Oklahoma Department of Corrections pre-parole supervised release with electronic monitoring*. San Francisco: National Council on Crime and Delinquency.

Barratt, E. S. (1993). The use of anticonvulsants in aggression and violence. *Psychopharmacological Bulletin, 29,* 75–81.

Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review, 103,* 5–33.

Bell, P. A., & Baron, R. A. (1981). Ambient temperature and human violence. In P. F. Brain & D. Benton (Eds.), *Multidisciplinary approaches to aggression research* (pp. 263–280). Amsterdam: Elsevier.

Blackburn, R. (1993). *The psychology of criminal conduct: Theory, research and practice*. Toronto, Canada: Wiley.

Blau, J. R., & Blau, P. M. (1982). The cost of inequality: Metropolitan structure and violent crime. *American Sociological Review, 47,* 114–129.

Bonta, J. (1996). Risk-needs assessment and treatment. In A. T. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 18–32). Thousand Oaks, CA: Sage.

Bonta, J., & Hanson, R. K. (1994). Gauging the risk for violence: Measurement, impact and strategies for change. *User report 1994–09*. Ottawa, Canada: Ministry of the Solicitor General.

Bonta, J., Harman, W. G., Hann, R. G., & Cormier, R. B. (1996). The prediction of recidivism among federally sentenced offenders: A re-validation of the SIR scale. *Canadian Journal of Criminology, 38,* 61–79.

Borum, R. (1996). Improving the clinical practice of violence risk assessment: Technology, guidelines, and training. *American Psychologist, 51,* 945–956.

Brodsky, S. L. (1972). *Psychologists in the Criminal Justice System*. Carbondale, IL: American Association of Correctional Psychologists.

Brody, S. R. (1976). *The effectiveness of sentencing*. Home Office Research Study No. 35. London: HMSO.

Brown, G. L., Ebert, M. H., Goyer, P. F., Jimerson, D. C., Klein, W. J., Bunney, W. E., & Goodwin, F. K. (1982). Aggression, suicide, and serotonin: Relationships to CSF amine metabolites. *American Journal of Psychiatry, 139,* 741–746.

Browne, K. D., & Herbert, M. (1996). *Preventing family violence*. Chichester, UK: Wiley.

Browne, K., & Howells, K. (1996). Violent offenders. In C. R. Hollin (Ed.), *Working with offenders: Psychological practice in offender rehabilitation* (pp. 188–210). Chichester, UK: Wiley.

Buchanan, R. W. (1995). Clozapine: Efficacy and safety. *Schizophrenia Bulletin 21,* 579–591.

Cairns, R. B., & Cairns, B. D. (1992). The sociogenesis of aggressive and antisocial behavior. In J. McCord (Ed.), *Facts, frameworks, and forecasts* (pp. 157–191). New Brunswick, NJ: Transaction.

Caliso, J. A., & Milner, J. S. (1994). Childhood physical abuse, childhood social support, and adult child abuse potential. *Journal of Interpersonal Violence, 9,* 27–44.

Caspi, A., & Moffitt, T. E. (1995). The continuity of maladaptive behavior: From description to understanding in the study of antisocial behavior. In D. Cicchetti & D. Cohen (Eds.), *Developmental psychopathology, Vol. 2: Risk, disorder, and adaptation*. (pp. 472–511). New York: Wiley.

Chaffin, M., Kelleher, K., & Hollenberg, J. (1996). Onset of physical abuse and neglect: Psychiatric, substance abuse, and social risk factors from prospective community data. *Child Abuse & Neglect, 20,* 191–203.

Cornell, D. G., Warren, J., Hawk, G., Stafford, E., Oram, G., & Pine, D. (1996). Psychopathy in instrumental and reactive violent offenders. *Journal of Consulting and Clinical Psychology*, *64*, 783–790.

Cullen, F. T., Wright, J. P., & Applegate, B. K. (1996). Control in the community: The limits of reform? In A. T. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 69–116). Thousand Oaks, CA. Sage.

Delgado-Escueta, A. V., Mattson, R. H., King, L., Goldensohn, E. S., Spiegel, H, Madsen, J., Crandall, P., Dreifuss, F., & Porter, R. J. (1981). The nature of aggression during epileptic seizures. *New England Journal of Medicine*, *305*, 711–716.

Dinwiddie, S. H., & Bucholz, K. K. (1993). Psychiatric diagnoses and self-reported child abusers. *Child Abuse & Neglect*, *17*, 465–476.

Dodge, K. A. (1991). The structure and function of reactive and proactive aggression. In D. J. Pepler & K. H. Rubin (Eds.), *The development and treatment of childhood aggression* (pp. 201–218). Hillsdale, NJ: Erlbaum.

Dore, M. M., & Alexander, L. B. (1996). Preserving families at risk of child abuse and neglect: The role of the helping alliance. *Child Abuse & Neglect*, *20*, 349–361.

Farrington, D. P. (1990). Implications of criminal career research for the prevention of offending. *Journal of Adolescence*, *13*, 93–130.

Farrington, D. P. (1991). Childhood aggression and adult violence: Early precursors and later-life outcomes. In D. J. Pepler & K. H. Rubin (Eds.), *The development and treatment of childhood aggression* (pp. 5–29). Hillsdale, NJ: Erlbaum.

Federal Bureau of Investigation. (1965–1991). *Crime in the United States: The uniform crime reports*. Washington, DC: Department of Justice.

Federal Bureau of Investigation. (1993). *Crime in the United States: The uniform crime reports*. Washington, DC: Department of Justice.

Federal Bureau of Investigation. (1996). *Crime in the United States: The uniform crime reports*. Washington, DC: Department of Justice.

Gendreau, P. (1996). The principles of effective intervention with offenders. In A. T. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 117–130). Thousand Oaks, CA: Sage.

Gendreau, P., & Paparozzi, M. A. (1995). Examining what works in community corrections. *Corrections Today*, *57*, 28–31.

Gendreau, P., & Ross, R. (1987). Revivification of rehabilitation: Evidence from the 1980s. *Justice Quarterly*, *4*, 349–407.

Goldman, M. B., & Janecek, H. M. (1990). Adjunctive fluoxetine improves global function in chronic schizophrenia. *Journal of Clinical Neuroscience*, *2*, 429–431.

Greenberg, D. (1977). The correctional effects of corrections: A survey of evaluation. In D. Greenberg (Ed.), *Corrections and punishment* (pp. 111–148). Beverly Hills, CA: Sage.

Hare, R. D., Harpur, T. J., Hakstian, A. R., Forth, A. E., Hart, S. D., & Newman, J. P. (1990). The revised Psychopathy Checklist: Reliability and factor structure. *Psychological Assessment*, *2*, 338–341.

Harris, G. T., Rice, M. E., & Cormier, C. A. (1991). Psychopathy and violent recidivism. *Law and Human Behavior*, *15*, 625–637.

Harris, G. T., Rice, M. E., & Cormier, C. A. (1994). Psychopaths: Is a therapeutic community therapeutic? *Therapeutic Communities*, *15*, 283–299.

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior*, *20*, 315–335.

Helfer, R. E., Hoffmeister, J. K., & Schneider, C. J. (1978). *MSPP: A manual for the use of the Michigan Screening Profile of Parenting*. Boulder, CO: Express.

Hermann, B. P., & Whitman, S. (1984). Behavioral and personality correlates of epilepsy: A review, methodological critique and conceptual model. *Psychological Bulletin*, *95*, 451–497.

Hodgins, S., & Cote, G. (1990). Prevalence of mental disorders among penitentiary inmates in Quebec. *Canada's Mental Health, 8*(1), 1–4.

Hollin, C. R. (1996). *Working with offenders: Psychological practice in offender rehabilitation*. Chichester, UK: Wiley.

Howing, P. T., Wodarski, J. S., Gaudin, J. M., & Kurtz, P. D. (1989). Effective interventions to ameliorate the incidence of child maltreatment: The empirical base. *Social Work*, *34*, 330–338.

Jaffe P., Wolfe, D., Wilson, S. K., & Zak, L. (1986). Family violence and child adjustment: A comparative analysis of girls' and boys' behavioral symptoms. *American Journal of Psychiatry*, *143*, 74–77.

Jessor, R., Donovan, J. E., & Costa, F. M. (1991). *Beyond adolescence: Problem behavior and young adult development*. Cambridge, UK: Cambridge University Press.

Kelleher, K., Chaffin, M., Hollenberg, J., & Fischer, E. (1994). Alcohol and drug disorders among physically abusive and neglectful parents in a community-based sample. *American Journal of Public Health, 84,* 1586–1590.

Kempe, C. H., & Helfer, R. (1980). *The battered child syndrome* (3rd ed.). Chicago, IL: University of Chicago Press.

Klassen, D., & O'Connor, W. (1990). Assessing the risk of violence in released mental patients: A cross-validation study. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1,* 75–81.

Kosson, D. S., Smith, S. S., & Newman, J. P. (1990). Evaluation of the construct validity of psychopathy in Black and White male inmates: Three preliminary studies. *Journal of Abnormal Psychology, 99,* 250–259.

Laub, J. H., & Lauritsen, J. L. (1993). Violent criminal behavior over the life course: A review of the longitudinal and comparative research. *Violence and Victims, 8,* 235–252.

Laub, J. H., & Sampson, R. J. (1993). Turning points in the life course: Why change matters to the study of crime. *Criminology, 31,* 301–325.

Liberman, R. P. (1994). Psychosocial treatments for schizophrenia. Special Issue: Psychosocial treatments: Psychotherapy and experience: Science, subjectivity, and society. *Psychiatry: Interpersonal & Biological Processes, 57,* 104–114.

Liberman, R. P., & Corrigan, P. W. (1993). Designing new psychosocial treatments for schizophrenia. *Psychiatry: Interpersonal & Biological Processes, 56,* 238–249.

Liberman, R. P., Kopelowicz, A., Young, A. S. (1994). Biobehavioral treatment and rehabilitation of schizophrenia. *Behavior Therapy, 25,* 89–107.

Liberman, R. P., & Wong, S. E. (1984). Behavior analysis and therapy procedures related to seclusion and restraint. In K Tardiff (Ed.), *The psychiatric uses of seclusion and restraint* (pp. 36–67). Washington DC: American Psychiatric Press.

Lidberg, L., Tuck, J. R., Asberg, M., Scalia-Tomba, G. P., & Bertilsson, L. (1985). Homicide, suicide and CSF 5-HIAA. *Acta Psychiatrica Scandinavica, 71,* 230–236.

Lidz, C. W., Mulvey, E. P., & Gardner, W. (1993). The accuracy of predictions of violence to others. *Journal of the American Medical Association, 269,* 1007–1011.

Link, B. G., Andrews, H., & Cullen, F. T. (1992). The violent and illegal behavior of mental patients reconsidered. *American Sociological Review, 57,* 275–293.

Lipsey, M. W. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In T. D. Cook, H. Cooper, D. S. Cordray, H. Hartmann, L. V. Hedges, R. J. Light, T. A. Louis, & F. Mosteller (Eds.), *Meta-analysis for explanation* (pp. 83–127). New York: Russell Sage Foundation.

Lipton, D., Martinson, R., & Wilks, J. (1975). *The effectiveness of correctional Treatment*. New York: Praeger.

Loeber, R., & Stouthamer-Loeber, M. (1986). Family factors as correlates and predictors of juvenile conduct problems and delinquency. In N. Morris & M. Tonry (Eds.), *Crime and justice* (pp. 29–149). Chicago: University of Chicago Press.

Lowenstein, L. F. (1991). The psychological understanding of crime. *Contemporary Review, 259,* 27–33.

MacMillan, H. L., Fleming, J. E., Trocme, N., Boyle, M. H., Wong, M., Racine, Y. A., Beardslee, W. R., & Offord, D. R. (1997). Prevalence of child physical and sexual abuse in the community: Results from the Ontario Health Supplement. *Journal of the American Medical Association, 278,* 131–136.

Mark, V. H., & Ervin, F. R. (1970). *Violence and the brain*. New York: Harper & Row.

Martinson, R. (1974). What works? Questions and answers about prison reform. *The Public Interest, 35,* 22–54.

Mash, E. J., Johnston, C., & Kovitz, K. (1983). A comparison of the mother–child interactions of physically abused and non-abused children during play and task situations. *Journal of Clinical Child Psychology, 12,* 337–346.

Menzies, R., Webster, C. D., McMain, S., Staley, S., & Scaglione, R. (1994). The dimensions of dangerousness revisited. *Law and Human Behavior, 18,* 1–28.

Messner, S., & Tardiff, K. (1986). Economic inequality and levels of homicide: An analysis of urban neighborhoods. *Criminology, 24,* 297–317.

Milner, J. S. (1986). *The Child Abuse Potential Inventory: Manual* (2nd ed.). Webster, NC: Psytec Corporation.

Milner, J. S. (1989). Additional cross-validation of the Child Abuse Potential Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1,* 219–223.

Milner, J. S. (1991). Medical conditions and the Child Abuse Potential Inventory specificity. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 208–212.

Milner, J. S. (1994). Assessing physical child abuse risk: The Child Abuse Potential Inventory. *Clinical Psychology Review, 44,* 547–583.

Milner, J. S., Gold, R. G., Ayoub, C., & Jacewitz, M. M. (1984). Predictive validity of the Child Abuse Potential Inventory. *Journal of Consulting and Clinical Psychology, 52,* 879–884.

Milner, J. S., & Murphy, W. D. (1995). Assessment of child physical and sexual abuse offenders. *Family Relations, 44,* 478–488.

Milner, J. S., & Robertson, K. R. (1989). Inconsistent response patterns and the prediction of child maltreatment. *Child Abuse & Neglect, 13,* 59–64.

Monahan, J. (1992). Mental disorder and violent behavior: Perceptions and evidence. *American Psychologist, 47,* 511–521.

Monahan, J., & Steadman, H. (1994). Toward the rejuvenation of risk research. In J. Monahan & H. Steadman (Eds.), *Violence and mental disorder: Developments in risk assessment* (pp. 1–17). Chicago: University of Chicago Press.

Moore, D. W. (1995). Three million children victims of physical abuse last year. *Gallup Poll Monthly, 363,* 4–16.

Mullen, P. E. (1997). A reassessment of the link between mental disorder and violent behaviour, and its implications for clinical practice. *Australian & New Zealand Journal of Psychiatry, 31,* 3–11.

Murphy, J. M., Jellinek, M., Quinn, D., Smith, G., Poitrast, F. G., & Goshko, M. (1991). Substance abuse and serious child maltreatment: Prevalence, risk, and outcome in a court sample. *Child Abuse & Neglect, 15,* 197–211.

Neppe, V. M. (1983). Carbamazepine as adjunctive treatment in nonepileptic chronic inpatients with EEG temporal lobe abnormalities. *Journal of Clinical Psychiatry, 44,* 326–331.

Nuffield, J. (1989). The SIR scale. Some reflections on its application. *Forum on Corrections Research, 1,* 19–22.

Otto, R. (1992). The prediction of dangerous behavior: A review and analysis of "second generation" research. *Forensic Reports, 5,* 103–133.

Parke, R. D., & Slaby, R. G. (1983). The development of aggression. In E. M. Hetherington (Ed.), *Handbook of child psychology: Socialization, personality, and social development* (4th ed., Vol. 4, pp. 547–642). New York: Wiley.

Pelcovitz, D., & Kaplan, S. J. (1994). Child witnesses of violence between parents: Psychosocial correlates and implications for treatment. *Child & Adolescent Psychiatric Clinics of North America, 3,* 745–758.

Petersilia, J. (1995). A crime control rationale for reinvesting in community corrections. *Spectrum: the Journal of State Government, 68*(3), 16–28.

Petersilia, J., & Turner, S. (1993). Intensive probation and parole. In M. Tonry (Ed.), *Crime and justice: An annual review of research.* Chicago: University of Chicago Press.

Quinsey, V. L. (1995). The prediction and explanation of criminal violence. *International Journal of Law and Psychiatry, 18,* 117–127.

Quinsey, V. L., & Maguire, A. (1986). Actuarial and clinical prediction of dangerousness concerning maximum security psychiatric patients. *Journal of Interpersonal Violence, 1,* 143–171.

Quinsey, V. L., Rice, M. E., & Harris, G. T. (1995). Actuarial prediction of sexual recidivism. *Journal of Interpersonal Violence, 10,* 85–105.

Reiss, A. J., & Roth, J. A. (1994). Understanding and preventing violence. Washington, DC: National Academy Press.

Rice, M. E. (1997). Violent offender research and implications for the criminal justice system. *American Psychologist, 52,* 414–423.

Rivera, B., & Widom, C. S. (1990). Childhood victimization and violent offending. *Violence & Victims, 5,* 19–35.

Robertson, G. (1988). Arrest patterns among mentally disordered offenders. *British Journal of Psychiatry, 153,* 313–316.

Robins, L., & Regier, D. (1991). *Psychiatric disorders in America: The epidemiological catchment area study.* New York: Free Press.

Sampson, R. J., & Laub, J. H. (1993). *Crime in the making: Pathways and turning points through life.* Cambridge, MA: Harvard University Press.

Schneider, C. J. (1982). The Michigan Screening Profile of Parenting. In R. H. Starr, Jr. (Ed.), *Child abuse prevention: Policy implications* (pp. 157–174). Cambridge, MA: Ballinger.

Sechrest., L., White, S. O., & Brown, E. D. (1979). *The rehabilitation of criminal offenders: Problems and prospects.* Washington, DC: National Academy of Sciences.

Seidenwurm, D., Pounds, T. R., Globus, A., & Valk, P. E. (1997). Abnormal temporal lobe metabolism in violent subjects: correlation of imaging and neuropsychiatric findings. *American Journal of Neuroradiology, 18,* 625–631.

Serin, R. C. (1991). Psychopathy and violence in criminals. *Journal of Interpersonal Violence, 6,* 423–431.

Serin, R. C. (1996). Violent recidivism in criminal psychopaths. *Law and Human Behavior, 20,* 207–217.

Serin, R. C., & Amos, N. L. (1995). The role of psychopathy in the assessment of dangerousness. *International Journal of Law and Psychiatry, 18,* 231–238.

Serin, R. C., Peterson, R. D., & Barbaree, H. E. (1990). Predictors of psychopathy and release outcome in a criminal population. *Psychological Assessment, 2,* 419–422.

Steadman, H. J., Fabisiak, S., Dvoskin, J., & Holohean, E. (1987). A survey of mental disability among state prison inmates. *Hospital and Community Psychiatry, 38,* 1086–1090.

Steadman, H. J., Monahan, J., Appelbaum, P. S., Grisso, T., Mulvey, E. P., Roth, L. H., Robbins, P. C., & Klassen, D. (1994). Designing a new generation of risk assessment research. In J. Monahan & H. Steadman (Eds.), *Violence and mental disorder: Developments in risk assessment* (pp. 297–318). Chicago: University of Chicago Press.

Swanson, J. W., Holzer, C. E., Ganju, V. K., & Jono, R. (1990). Violence and psychiatric disorder in the community: Evidence from the epidemiologic catchment area surveys. *Hospital and Community Psychiatry, 41,* 761–770.

Tardiff, K. (1984). Characteristics of assaultive patients in private hospitals. *American Journal of Psychiatry, 141,* 1232–1235.

Tardiff, K. (1996). *Concise guide to assessment and management of violent patients* (2nd ed). Washington, DC: American Psychiatric Press.

Tardiff, K., Marzuk, P. M., Leon, A. C., Portera, L., & Weiner, C. (1997). Violence by patients admitted to a private psychiatric hospital. *American Journal of Psychiatry, 154,* 88–93.

Teplin, L. (1990). The prevalence of severe mental disorder among male urban jail detainees: Comparison with the epidemiologic catchment area program. *American Journal of Public Health, 80,* 663–669.

Tonkonogy, J. M. (1991). Violence and temporal lobe lesion: Heat CT and MRI data. *Journal of Neuropsychiatry and Clinical Neuroscience, 3,* 189–196.

US Bureau of the Census (1994). *United States population estimates.* Washington, DC: Author.

US Bureau of Justice Statistics (1995). *Correctional populations in the United States 1992.* BJS Bulletin NCJ-146413. Washington, DC: Author.

US Department of Justice (1994). *Sourcebook of criminal justice statistics.* Washington, DC: Author.

US Department of Justice (1995). *Crime in the US.* Washington, DC: Author.

Webster, C. D., Eaves, D., Douglas, K., & Wintrup, A. (1995). *The HCR-20 Scheme: The assessment of dangerousness and risk.* Burnaby, Canada: Simon Fraser University and Forensic Psychiatric Services Commission of British Columbia.

Weiler, B. L., & Widom, C. S. (1996). Psychopathy and violent behavior in abused and neglected young adults.

Criminal Behaviour & Mental Health, 6, 253–271.

Wessely, S., & Taylor, P. (1991). Madness and crime: Criminology versus psychiatry. *Criminal Behaviour and Mental Health, 1,* 193–228.

White, J. L., Moffitt, T. E., Earls, F., Robins, L., & Silva, P. (1990). How early can we tell? Predictors of childhood conduct disorder and adolescent delinquency. *Criminology, 28,* 507–534.

Widom, C. S. (1989a). Child abuse, neglect, and adult behavior: Research design and findings on criminality, violence, and child abuse. *American Journal of Orthopsychiatry, 59,* 355–367.

Widom, C. S. (1989b) Does violence beget violence? A critical examination of the literature. *Psychological Bulletin, 106,* 3–28.

Wilcock, G. (1987). Trazodone/tryptophan for aggressive behavior. *Lancet, 1,* 929–930.

Wolfe, D. A., & Wekerle, C. (1993). Treatment strategies for child physical abuse and neglect: A critical progress report. *Clinical Psychology Review, 13,* 473–500.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.27
Perpetrators of Domestic Violence

ETIONY ALDARONDO
Boston College, Chestnut Hill, MA, USA

9.27.1 INTRODUCTION 437
9.27.1.1 Definitions 437
9.27.2 RESEARCH ON PERPETRATORS OF DOMESTIC VIOLENCE 438
9.27.2.1 Incidence and Prevalence of Intimate Violence 438
9.27.2.1.1 Heterosexual relationships 438
9.27.2.1.2 Homosexual relationships 439
9.27.2.1.3 Continuity and cessation of wife assault 439
9.27.2.1.4 Intimate violence in clinical populations 439
9.27.3 CHARACTERISTICS OF PERPETRATORS OF DOMESTIC VIOLENCE 440
9.27.3.1 Risk Markers for Wife Assault 440
9.27.3.2 Types of Men Who Batter 441
9.27.4 THE ASSESSMENT OF MALE BATTERERS 441
9.27.4.1 Interview Assessment 442
9.27.4.2 Self-report Measures of Intimate Violence 443
9.27.4.3 Assessment of Dangerousness 445
9.27.5 CLINICAL INTERVENTIONS WITH PERPETRATORS OF DOMESTIC VIOLENCE 445
9.27.5.1 Treatment Programs for Male Batterers 445
9.27.5.1.1 Group therapy approaches 446
9.27.5.1.2 Couple therapy 447
9.27.6 SUMMARY AND FUTURE DIRECTIONS 448
9.27.7 REFERENCES 449

9.27.1 INTRODUCTION

Perpetrators of domestic violence are a heterogeneous group. Violent partners do not fit conventional molds of gender, deviancy, and psychopathology. An understanding of the variability within this group is essential to the development of viable theories, sound assessment instruments, assessment procedures, and effective treatments. Further, various forms of intimate violence, patterns of assault, and typologies of men who batter their female partners will require examination. Hence, the chapter has five sections. The first section examines the nature of research on perpetrators of violence. The second section examines the issue of diversity in national, community, and clinical samples of violent couples. The next two sections focus on assessment and treatment issues, and the final section discusses some implications of the heterogeneity perspective for research and practice.

9.27.1.1 Definitions

Domestic violence may include intimidation, harassment and persecution, verbal aggression, denial of access to resources, sexual coercion and assault, physical assault, and torture (e.g., Walker, 1984). Although each form of violence can have devastating effects on the victim, the chapter will focus on physical aggression in adult intimate relationships. Terms such as abuse, violence, family violence, and intimate violence are used to connote acts of aggression with the potential to cause physical harm or death to a partner. When the issue being addressed is violence by men against female partners, it will be named *wife assault*. Violence in gay relationships will also be regarded as *man-to-man* violence. Violence perpetrated by women against men will be called *woman-to-man* violence. Similarly, violence perpetrated by women in lesbian relationships will be called *woman-to-woman* violence.

9.27.2 RESEARCH ON PERPETRATORS OF DOMESTIC VIOLENCE

Given the explosive growth of family violence research, one might assume that much is known about perpetrators of violence. However, research on domestic violence reveals a tenuous understanding of violent partners and many inconsistent findings on issues of theoretical and practical importance, such as sex role attitudes, alcohol use, recidivism following treatment, and psychopathology. While some of these inconsistencies are the result of methodological deficiencies, there is a growing consensus among family violence researchers that many discontinuities in the literature are also a consequence of the failure to recognize the heterogeneity among perpetrators of violence and to modify our understanding of intimate violence (Aldarondo & Sugarman, 1996; Holtzworth-Munroe & Stuart, 1994; Letellier, 1994; O'Leary, 1993).

9.27.2.1 Incidence and Prevalence of Intimate Violence

9.27.2.1.1 Heterosexual relationships

Typically, intimate violence occurs under a veil of secrecy that makes it difficult to estimate its incidence and prevalence. National surveys, surveys of community samples, and homicide records suggest that violence against women by men is without doubt the most pervasive and consequential form of intimate violence. Male violence against women is a major social problem, a serious public health problem affecting the physical and mental health of women, and a serious obstacle to development and peace in the world (Goodman, Browne, Fitzgerald, Keita, & Russo, 1994; Russo, Koss, & Goodman, 1995). The available evidence for the USA indicates that: (i) an estimated two to three million women are severely assaulted by their husbands or boyfriends each year (Straus & Gelles, 1990); (ii) the rates of depression, suicide attempts, psychosomatic symptoms, and stress are four times greater among female victims of severe violence than among women who are not victims of violence (Stets & Straus, 1990); (iii) most sexual assault of women is perpetrated by male intimates (Finkelhor & Yllö, 1985; Kilpatrick, Edmunds, & Seymour, 1992); and (iv) close to one third of all female homicides are perpetrated by male partners (Browne & Williams, 1993).

The main sources of data on the incidence of intimate violence in the USA are three national surveys conducted in 1975 (Straus, Gelles, & Steinmetz, 1980), in 1985 (Straus & Gelles, 1990), and in 1992 (Kaufman, Jasinski & Aldarondo, 1994). Concerning severe violence only, wife assault rates on the three surveys were 3.8%, 3.0%, and 2.0%, respectively. Straus and Kaufman-Kantor (1994) reported continued significant decrease of severe violence from 1975 to 1992 regardless of the gender of the respondent. Moreover, these authors argued that the decrease in severe wife assaults "is parallel to the 18% decrease in the rate of homicides of wives by husbands during this period" (p. 8). In contrast, in 1992, the rate of minor violence against women as reported by women increased. Whether the decrease in severe violence by men reflects an actual reduction in wife assault or unwillingness to report is unknown, although reductions in the reports given by women suggest the former to be the case.

Women in heterosexual relationships frequently report physical aggression against their male partners. Evaluations of national data have found comparable rates of physical violence by men and women (Brush, 1990; Morse, 1995; Straus, 1993; Straus & Gelles, 1990). Moreover, Browne and Williams' (1993) analysis of homicide records in the USA from 1976 to 1987 showed that over a third of all cases in which one partner killed another were men killed by female partners. The 1985 survey showed little change from the rates of woman-to-man violence reported in 1975—severe violence rates were 4.6% and 4.4%, respectively. The 1992 survey, however, found a significant increase in severe woman-to-man violence (5.9%), as reported by the wives themselves (Straus & Kaufman-Kantor, 1994).

The national rates are much lower than those found in selected community samples (e.g., McLaughlin, Leonard, & Senchak, 1992; O'Leary et al., 1989). There is, however, substantial variability in the prevalence rates from community surveys (Fagan & Browne, 1994; Straus & Gelles, 1990). A review of family violence research concluded that "the larger samples—over 1000 women or couples—have rates around 10 in 100 reporting some form of family violence as measured by the Straus and

Gelles scale" (Reiss & Roth, 1993, p. 227). Community samples have also found that rates of woman-to-man assaults are about the same as the rates of physical violence against women by men (Straus, 1993; Straus & Gelles, 1990).

9.27.2.1.2 Homosexual relationships

In the 1990s, a growing number of studies of physical violence in lesbian and gay relationships (e.g., Kahuna, 1990; Margolies & Leeder, 1995) have challenged traditional feminist and sociocultural perspectives on intimate violence (Coleman, 1994; Letellier, 1994). These studies suggest that the prevalence of physical violence in homosexual relationships may be as high as or higher than in heterosexual couples.

Lie and Gentlewarrior's (1991) survey of 1099 self-identified lesbian women showed that 52% had been victims of violence and 57% had been perpetrators of violence. In another study, Lie, Schilit, Bush, Montague, and Reyes (1991) evaluated the history of violence in the lives of 174 self-identified lesbian women. Forty-seven percent of the women had been victims of physical aggression and 38% had been perpetrators of violence in previous same-sex relationships. By comparison, relationships with current partners were considerably less aggressive; 10% of the women reported being victims and 9% reported being perpetrators of violence in their current relationships.

Lockhart, White, Causby, and Isaac (1994) surveyed a self-selected sample of 284 women attending a women's music festival. The Conflict Tactics Scale (CTS) was used to measure acts of intimate violence directed towards the respondents during the past year. Almost a third of the women (31%) in same-sex relationships reported experiencing one or more incidents of physical abuse in the past year. Among the victimized women, 38% reported that they had experienced severe forms of violence.

Concerning male homosexual relationships, Kelly and Washafsky (1987) found that 62% of the men reported some form of physical aggression in their relationships. Bologna, Waterman, and Dawson (1987) reported that 18% and 14% of gay men reported being victims or perpetrators of violence in their current or most recent relationship, respectively.

9.27.2.1.3 Continuity and cessation of wife assault

Variability in patterns of violence over time is another important aspect of the incidence of violence in intimate relationships. Although there is some information about female perpetrators of violence and violent partners in homosexual relationships, most information about perpetrators of violence deals with men who abused their female partners. Therefore, unless otherwise specified, the discussions that follow will focus on male batterers. Many men assault their partner once in a period of several years, some men use physical aggression occasionally over time, and a smaller group of men persist in their use of violence over time (Aldarondo, 1996).

Aldarondo (1996) evaluated patterns of cessation and persistence of wife assault over a three-year period, using data from the 1985 National Family Violence Survey as baseline. This analysis showed that over half of the men who assaulted their partners ceased or interrupted the use of physical violence for at least a year. After a year of cessation, these men were more likely to remain nonviolent than to resume the use of physical violence against their female partners. More than a third of the men recorded as violent in the first year of the study ceased the violence for the next two years.

Data show that more than half of the violent men in the general population are involved in either persistent or intermittent patterns of wife assault (Aldarondo, 1996). This supports the common belief that once violence becomes scripted in many relationships, it is difficult to change (Bowker, 1983). Women who have ended abuse by a spouse and remained in the relationship have been found to have endured abuse for an average of 10 years (Horton & Johnson, 1993). Thus, although the relatively high rate of wife assault cessation brings hope for change in the lives of couples in which men are violent, the available evidence suggests that a considerable number of men who use physical violence against their female partners are likely to continue using violence in the future (Follingstad, Laughlin, Polek, Rutledge, & Hause, 1991; Woffordt, Mihalic & Menard, 1994). In regard to this issue, then, violent men in the general population resemble the clinical population of men who batter their female partners.

9.27.2.1.4 Intimate violence in clinical populations

Although clinicians have been concerned about reports of physical aggression in mental health clinics for some time, the prevalence of intimate violence in the clinics and mental health centers has received scant attention. Gondolf, Mulvey, and Lidz (1989) assessed the prevalence of family violence in a psychiatric emergency room. Among 389 patients interviewed over a six-month period, 19% reported having perpetrated some act of physical aggression against their intimate partners. No information was offered about the gender of the perpetrators and the victims of violence.

Tham, Ford and Wilkinson (1995) reported that 24% of people presenting acutely for psychiatric services at a clinic had been the victims of physical violence. Thirty-five percent of the women and 13% of the men reported suffering domestic violence at some point in their life. Among those with a history of domestic violence, 24% of the women and 16% of the men had suffered the violence within the last six months. In another study, Cascardi, Langhinrichsen, and Vivian (1992) found that 71% of couples seeking treatment at a marital therapy clinic reported acts of intimate violence in the year prior to coming to therapy. When reports by both men and woman were combined, 27% and 34% of the men were found to engage in minor and severe forms of violence, respectively. According to reports given by women, 54% of the men used physical aggression against them in the past year, and 62% of these violent men used severe forms of violence. Men's reports yielded comparable rates of overall wife assault (55%) but considerably less severe violence (40%) than their female partners reported. Considering woman-to-man violence, combined couple data indicated that 12% used minor violence, while 50% used severe forms of violence. Forty-five percent of the women described themselves as physically aggressive. Of these women, 79% reported having perpetrated severe acts of violence against their male partners in the past year. According to the men, 57% of the women had been physically aggressive in past year and, in 80% of these cases, women relied on severe violence.

O'Leary, Vivian, and Malone (1992) evaluated the prevalence rates of wife assault in a sample of couples in treatment at the same marital therapy clinic mentioned above. Only 1.5% of the men and 6.0% of the women mentioned physical assaults in their written descriptions. Acts of physical aggression were mentioned by 46% of the men and 44% of the women during the interview. In contrast, 53% of the men, 53% of the women, and 67% of the couples reported wife assaults when assessed with the CTS. Severe forms of assault were reported by 17% of the men, 21% of the women, and 25% of the couples.

9.27.3 CHARACTERISTICS OF PERPETRATORS OF DOMESTIC VIOLENCE

9.27.3.1 Risk Markers for Wife Assault

One way in which family violence researchers approach the study of perpetrators of domestic violence is through the identification of risk markers. Risk markers are not conceived as causes of violence but their presence makes violence more likely under certain conditions. "Risk factors are individual or group characteristics as well as social circumstances that are associated with the occurrence of family violence. Rarely does one factor alone determine overall risk; rather a few factors typically appear in combination" (American Psychological Association, 1996, p. 18).

Research on men who batter often relies on reports by female victims recruited from shelters for battered women, community samples, and mental health clinics. Typically, the women's input is obtained through self-report questionnaires and interviews designed to describe the men's use of violence, articulate the victims' experience, and generate a profile of the perpetrators. Research on men who batter is also done with men in court-mandated treatment programs for men who batter, couples in treatment, and with volunteer couples from the community. With such diversity in samples, it should not come as a surprise that an array of individual, social, and demographic variables have been submitted as potential risk factors for wife assault.

Hotaling and Sugarman (1986) evaluated the consistency of 97 risk markers for wife assault across 52 studies of intimate violence. They found that witnessing parental violence during childhood or adolescence, sexual aggression towards the wife, use of violence towards children, high alcohol consumption, low income level, occupational status, low educational level, and lack of assertiveness were all consistent risk markers for men's use of physical aggression against their female partners. On the basis of this evaluation, Hotaling and Sugarman stated that, "Batterers are exposed early in life to family violence, are less assertive, and possess fewer educational and occupational resources than nonviolent men. This review also finds that batterers are much more likely to engage in other forms of antisocial behavior than men who are not violent towards their wives" (p. 114).

Subsequently, Sugarman and Hotaling (1989) examined which risk markers best differentiated among nonviolent men, verbally aggressive men, men who engaged in minor forms of violence, and severely physically violent men in a national probability sample of 2143 families. They found that the higher the level of conflict in the relationship, the greater the probability that men would assault their female partners. Severely violent men were found to have lower socioeconomic status (SES) and greater frequency of witnessing violence in their family of origin than all other groups of men. Aldarondo and Sugarman (1996) found that these conditions also increase the risk that men would continue to act violently against their female partners over time.

9.27.3.2 Types of Men Who Batter

In addition to identifying risk markers for wife assault, family violence researchers are very interested in developing reliable and valid classification systems that would help differentiate between the various subtypes of men who batter. Such classifications are of potential theoretical and practical importance. Typologies of men who batter could help identify different processes leading to violence—what is important for the development of one type of batterer may be irrelevant for the development of other types. Such typologies could also support the development of specific treatment practices for different male batterers. As illustrated by research on the effects of arrest in cases of wife assault, interventions that are effective with some men have no effect on others, and, unfortunately, in some cases are followed by increased levels of violence against women (Buzawa & Buzawa, 1996; Sherman, 1992).

In a review of the literature on typologies of male batterers, Holtzworth-Munroe and Stuart (1994) identified three dimensions that have been found consistently to differentiate among subtypes of batterers: severity of violence, generality of violence, and psychological functioning. On the basis of these dimensions they suggested that there are three main types of male batterers: *family-only, dysphoric/borderline*, and *generally violent/antisocial.* According to Holtzworth-Munroe and Stuart (1994), as many as 50% of men who batter their female partners may be classified as family-only batterers. These men engage in the least severe violence and are the least likely to engage in psychological and sexual abuse. Men in this group generally direct the violence to family members and are the least likely to perpetrate acts of violence outside the home and to have legal problems for this reason. The family-only batterer shows little evidence of psychopathology or personality disorder.

Dysphoric/borderline batterers are estimated to constitute approximately 25% of batterer samples. These men engage in moderate to severe wife abuse, including psychological and sexual abuse. Like family-only batterers, dysphoric/borderline batterers direct their violence primarily to family members. However, unlike the family-only batterers, they may engage in violence outside the family and may engage in other criminal behavior. In terms of psychological functioning, these men are considered to be the most dysphoric, distressed, and emotionally volatile among male batterers. Their personality organization may include borderline and schizoid features. Alcohol and drug abuse may be common in this group of men. The dysphoric/borderline batterer is similar to what Dutton (1994) calls the abusive personality characterized by a cycle of fear and rage directed towards the women to whom they are emotionally connected.

The remaining 25% of batterers are classified as generally violent/antisocial although the number of generally violent/antisocial batterers may be considerably higher in treatment programs with court-mandated batterers (Hart, Dutton, & Newlove, 1993). These men engage in moderate to severe violence, including psychological and sexual abuse. Frequently they engage in violence outside their intimate relationship and have extensive criminal records. In addition, they are likely to have substance abuse problems, antisocial personality disorder, or psychopathology.

Research by Gottman et al. (1995) suggest that there may be a physiologically-based typology of male batterers that is consistent with the general distinction made above between male batterers who are abusive primarily at home and batterers who are also violent outside the marriage. They found that some male batterers (Type 1) lower their heart rate during arguments with their female partners, whereas other men (Type 2) increase their heart rate under similar conditions of conflict with their intimate partners. Although Type 1 men are not more violent against their female partners than Type 2 men, they are generally violent outside the family. They are also more likely to be assessed as antisocial, drug-dependent, and aggressive–sadistic.

9.27.4 THE ASSESSMENT OF MALE BATTERERS

Although there are no common guidelines for an integrated assessment of male batterers, family violence researchers and clinicians agree that women are more reliable sources of information about the violence they have experienced than the batterers. However, one of the major difficulties in the assessment of wife assault in clinical settings is that both perpetrators and victims of violence often fail to report the occurrence of violence (O'Leary et al., 1992).

In order for clinicians to adequately evaluate perpetrators of wife assault, they need to be aware of several "roadblocks" with the potential to divert and change the direction of the interview away from the issue of violence. A partial list of such roadblocks includes the following.

(i) Some people believe that violence is tolerable provided it does not "go too far." They may not be aware of the negative impact violence has in the quality of their life and in their personal health and tend to regard occasional instances of slapping or shoving as only a minor problem that is not important enough to bring up for discussion (O'Leary & Murphy, 1992).

(ii) Millions of people grow up witnessing violence between their parents (Jaffe & Wolfe, 1990; Straus, 1992). This experience validates the use of physical violence as a means of resolving conflicts in the family when other alternatives do not achieve the desired end.

(iii) Some people consider violent acts as merely symptomatic of more fundamental problems and therefore do not mention the violence.

(iv) Often victims and perpetrators of intimate violence do not have the ability to report accurately about highly emotional and traumatic events (Walker, 1995).

(v) Many people find it hard to disclose private aspects of their lives. Often, batterers are ashamed, embarrassed, and guilt-ridden regarding their violent behavior. They may choose to conceal or minimize the physical assaults to protect themselves or their loved ones from public condemnation and humiliation.

(vi) Batterers are known to deny and minimize the violence to escape criminal responsibility (Sonkin, 1987).

(vii) Denial and minimization of the violence by women are normal responses to abuse (Ferraro & Johnson, 1983). Female victims of violence may not trust therapists with knowledge of their victimization for fear their partner may learn about the disclosure and attack them, hurt the children, or take the children away. Women may also be reluctant to disclose acts of violence to avoid both potential retribution against their partners and the threat of separation.

9.27.4.1 Interview Assessment

It is very important for clinicians to facilitate the development of a safe, structured, and supportive relationship that would reduce the need for self-protection and dissimulation on the part of perpetrators and victims of violence. There is no substitute for direct training and supervision to develop these skills. Experts in the field suggest that interview assessment of male batterers should be informed by the following considerations.

Many men who enter treatment are not familiar with the therapy process. They tend to enter treatment to appease their spouses or because they are being told to do so by the courts. Prior to the first meeting it is important to contact the referring person to learn about why the man is being referred for treatment and if there are any strings attached to his participation in therapy (e.g., separation from partner, incarceration). At the start of the assessment interview, it is important to explain the assessment process to the batterer and to discuss the limits of confidentiality. The latter issue is particularly important because, depending on the circumstances of the case and the philosophy of the treatment program, the need may arise for counselors to exchange information with probation departments, other therapists involved with the family, the female partners, shelters for battered women, and child protection agencies.

It is also important for therapists to be mindful of the language they use to describe violent behavior. For many individuals there is a great difference between having pushed their intimate partner and having been violent. The latter is potentially an affront to their self-concept and sense of self-worth and, for that reason, may not be perceived as applying to someone who "only shoved her." Thus, although some therapists may regard using the term violence as more professional than pushing or shoving, only those clients who recognize their behavior as violent and who can tolerate defining themselves as such will respond, at least initially.

Violence may not be revealed unless both partners are asked specifically (Szinovacz, 1983). Men in treatment for wife assault report less violence and less severe acts of wife assault than their partners report for them (Browning & Dutton, 1986). Moreover, questioning both partners reveals incidents that do not come to light when only one partner is questioned (Bohannon & Dosser, 1995). However, it is important to note that interviewing the partners together generates fewer reports of wife assault than separately obtained reports, especially from women (Cantos, Neidig, & O'Leary, 1994).

Finally, many counselors working with men who batter agree that structured instruments should be used in conjunction with interview assessments (Saunders, 1995). The combined use of structured instruments and interview assessment yields additional acts of aggression and intimidation and a better understanding of the context for the physical aggression and its consequences than either technique alone (O'Leary et al., 1992).

9.27.4.2 Self-report Measures of Intimate Violence

Even therapists trained to do interview assessment of intimate violence obtain considerably less information about wife assault than can be obtained with the use of brief self-report measures (O'Leary & Murphy, 1992; Rhodes, 1992). That is why most treatment programs for men who batter their female partners use some self-report measure of violence to assess the history and severity of the violence (Saunders, 1995). Several instruments have been developed to measure physical aggression and psychological abuse in intimate relationships. Most measures are designed to be used with female victims but some are used with both partners and still others with male batterers only. The following list is organized in alphabetical order.

(i) *The Abusive Behavior Inventory (ABI).* Developed by Sheppard and Campbell (1992), this is a 30-item self-report measure of the frequency of abusive behaviors during a six-month period. The ABI includes 20 psychological abuse items concerning emotional abuse, intimidation, threats, use of male privilege, and economic abuse. It also includes 10 psychological abuse items, including forcing the woman to have sex. Initial evaluations of this measure show adequate reliability and validity (Sheppard & Campbell, 1992).

(ii) *The Aggression Scale (AGG).* This is a self-report measure of physical and psychological aggression in intimate relationships (Snyder & Snow, 1995). It has 19 true–false items, including nine physical violence items and 10 psychological aggression items. In an initial evaluation, the AGG had a test–retest reliability of 0.81 and an alpha coefficient of internal consistency of 0.84 for the full scale. The physical violence and the psychological aggression subscales had reliability and internal consistency coefficients of 0.69 and 0.88 and of 0.82 and 0.94, respectively. Moreover, the scale was found to correlate moderately with the CTS verbal aggression and physical violence subscales and to discriminate between partners involved in intimate violence and violence-free couples (Snyder & Snow, 1995).

(iii) *The Conflict Tactics Scales (CTS).* This is the most widely used measure of physical violence in intimate relationships. It is composed of 19 items that assess tactics used in interpersonal conflicts. The items are presented on a continuum from nonviolent to severely violent tactics. The scale has two stable factors, physical aggression and verbal aggression (Barling, O'Leary, Jouriles, Vivian, & MacEwen, 1987), and high internal reliability (Straus, 1979; 1990; 1996).

In addition to the availability of normative tables, the CTS has several features that make it suitable for the practical needs of clinical psychologists and other mental health professionals. First, the CTS takes only a few minutes to administer, either within an interview or as a self-administered questionnaire. Thus, it can be added to standard intake procedures with relative ease. Second, the CTS targets specific behavior and thus minimizes the demand for respondents to recognize their behavior as violent in order to respond. Third, the CTS is a brief means of identifying the occurrence of violence as well as evaluating the severity and chronicity of that violence. Thus, the therapist is able to know whether acts of physical violence occurred prior to treatment and how many times those violent behaviors occurred over the course of the specified target date. The CTS also has some important limitations. The physical violence scale measures only the frequency and severity of physical attacks and does not indicate the extent of injury, if any, resulting from those attacks and it is unable to accurately measure sexual abuse within the relationship.

(iv) *The Danger Assessment Instrument (DA).* This is a statistical risk factor assessment of homicide in couple relationships (Campbell, 1995). The scale has 20 items. First, women are asked to indicate on a calendar the approximate dates and the duration of five specific violent acts experienced by them. Then, they have to say (*yes*) or (*no*) to 15 questions covering issues such as increases in frequency of physical violence during the past year, availability of guns in the house, and threats of suicide. Initial evaluations of the DA using women victims of violence show an alpha reliability of 0.71 and test–retest reliability ranging from 0.89 to 0.94. Moreover, the scale reportedly has moderately strong correlations with the CTS and the Index of Spouse Abuse. Campbell (1995) has suggested that until appropriate tests of the DA's predictive validity support its utility, the measure could be used for informal prediction discussions with battered women and concerned professionals but should not be used for formal prediction of homicide in counseling and court sentencing situations.

(v) *The Index of Spouse Abuse (ISA).* This was developed by Hudson and McIntosh (1981) to measure both physical and nonphysical abuse. The ISA consists of 30 items, three of which ask women about acts of physical violence, and one refers to sexual coercion that is not necessarily physical. The remaining items in the physical abuse subscale appear to measure physical intimidation and domination. Although the ISA provides useful information about the psychological maltreatment of women, the scale offers limited information about specific acts of physical violence. It also mixes physical violence with other forms of maltreatment. Thus, a subject could have a high score on the violence subscale without ever having been physically assaulted. In addition, the ISA contains no items on minor forms of physical violence and therefore is not sensitive to the most frequently occurring type of physical violence against women.

(vi) *The Measure of Wife Abuse (MWA)*. This is a 60-item questionnaire designed to measure different dimensions of wife abuse (Rodenburg & Fantuzzo, 1993). The MWA asks women to declare if during the past six months of their intimate relationship their male partners engaged in specific acts of violence and then asks them to indicate how much this experience hurt or upset them. Initial analysis of the MWA yields factors for physical abuse, sexual abuse, psychological abuse, and verbal abuse. Evaluations of the subscales" internal consistency indicates reasonable reliability, with an average subscale alpha of 0.83 and a total alpha of 0.93. Moderate correlations with the CTS verbal and physical violence scales support the concurrent validity of the measure (Rodenburg & Fantuzzo, 1993).

(vii) *The Propensity for Abuse Scale (PAS)*. This is a self-report scale created to assess the propensity for men to abuse their female partners (Dutton, 1995a). It was constructed to circumvent the tendency by abusive or potentially abusive men to respond in socially desirable ways. The scale does not contain explicit reference to intimate violence. Instead, it includes 29 items on personality traits, anger, trauma symptoms, and recollections of parental treatment. Factor analysis of the PAS yield three factors: negative parental treatment, affective liability, and trauma symptoms. The PAS has good reliability (alpha = 0.88). Moreover, the scale correlates significantly (0.30) with women's reports of physical assaults using the CTS and with women's reports of domination/isolation tactics (0.51) and emotional abuse (0.47) using the Psychological Maltreatment of Women Inventory. Dutton (1995a) cautions that "prior to clinical use of the scale as a potential predictive instrument, more cross-validation work is required, with clinical groups who do not present with an abuse problem" (p. 215).

(viii) *The Psychological Maltreatment of Women Inventory (PMWI)*. This is a 58-item scale including both physical and psychological aggression against women (Tolman, 1989). The scale, which can be completed by both male batterers and female victims, asks respondents to report on the relative frequency of each behavior. It has two stable factors, dominance–isolation and verbal–emotional, with internal consistency reliability ranging from 0.91 to 0.95. Although the measure successfully discriminates between battered and nonbattered women, it remains to be determined if it discriminates between abusive and nonabusive men. Accordingly, Tolman (1989) has suggested that the PMWI "can be reliably used as a measure of psychological maltreatment women are experiencing in their relationships" (p. 175) but should be considered a reliable tool to use in the evaluation of programs for men who batter.

(ix) *The Relationship Conflict Inventory (RCI)*. This was developed by Bodin (1992) and consists of 114 items measuring levels of verbal and physical conflict among couples in treatment. Four items are used to measure physical violence. The psychometric properties of the RCI remain unknown.

(x) *The Revised Conflict Tactic Scales (CTS2)*. This is an enlarged version of the CTS which addresses some of the major criticisms of the original scale (Straus, Hamby, Boney-McCoy, & Sugarman, 1996). The CTS2 has 39 items. It includes a negotiation scale with six items, a psychological aggression scale including four items from the old verbal aggression scale and four new items, a 12-item physical assault scale with three new items, a sexual coercion scale with seven items, and a six-item injury scale. An initial evaluation of the psychometric properties of the CTS2 indicates that the scales have good internal consistency reliability (range 0.79–0.95). There is also preliminary evidence of construct and discriminant validity. However, since these psychometric findings are based on administration to a sample of college students, additional analyses based on samples of the general population and relevant clinical populations are needed to determine the utility of the CTS2 for studying these populations (Straus et al., 1996).

(xi) *The Severity of Violence Against Women Scales (SVAWS)*. This is a 49-item instrument created to assess the severity of violence perpetrated by men on women (Marshall, 1992). Factor analysis of ratings made by college students and women in the community yielded nine factors (symbolic violence, mild threats, moderate threats, serious threats, minor violence, mild violence, moderate violence, sexual violence, and serious violence). There is no information on the validity and reliability of the SVAWS.

(xii) *The Spouse Specific Aggression Scale (SSA)*. This is a 29-item scale designed to measure psychological aspects of abuse in intimate relationships (O'Leary & Curley, 1986). The SSA has subscales for aggression and assertiveness. Respondents are asked to rate themselves on 12 items concerning passive aggressive and verbally aggressive behaviors. The SSA has been found to be associated with spouse abuse (Rosenbaum & O'Leary, 1981). In addition, the scale correlates highly with the verbal aggression subscale of the CTS. In terms of psychometric properties, the SSA has reliability coefficients for clinical samples of men and women of 0.74 and 0.85, respectively.

(xiii) *The Wife Abuse Inventory (WAI)*. A scale developed by Lewis (1985), it was designed to identify abused women or women at risk of being abused by their spouses. The scale asks women to rate their partners and themselves on a number of family management matters. As Saunders (1992b) pointed out, only two of these items measure physical violence and the rest focus not on abusive behaviors but on the causes and consequences of abuse. Thus, the WAI appears to measure the risk of abuse rather than detecting whether violence has occurred or is occurring (Sedlak, 1988).

9.27.4.3 Assessment of Dangerousness

Data on the partners of battered women in shelters (Giles-Sims, 1983), on abusers who have been murdered by their victim (Browne, 1987), on court-mandated abusers (Saunders, 1992a), and on men who killed their spouses (Goetting, 1989, 1991) suggest that self-report measures of violence are not sufficient to identify the life-threatening and often lethal levels of violence that characterize these men. Previous evaluations of dangerousness in intimate relationships (Aldarondo & Straus; 1994; Campbell, 1995; Saunders, 1995; Sonkin, 1987) suggest paying attention to the following risk markers for life-threatening violence by men against women: (i) high frequency of violence; (ii) alcohol or drug abuse; (iii) history of dependency and/or jealousy; (iv) history of causing physical injuries on partner; (v) arrest history or violent behavior outside as well as inside the home; (vi) history of rape or of forcing partner to have sexual intercourse against her will; (vii) history of marital violence in the family of origin; (viii) possession or use of weapons; (ix) threats to hit and/or kill partner; (x) history of psychological maltreatment and coercive behaviors; and (xi) sociopathic and narcissistic features.

It is important for therapists working with intimate violence to be familiar with the factors that increase the likelihood of such acts. However, it is also important to know that the prediction of homicides and potentially lethal violence always results in a high number of false-positive cases. Therapists should be mindful that "the prediction of dangerousness is not without its own dangers" (Saunders, 1995, p. 87). The high false-positive rate is inevitable because these are relatively rare events (Sedlak, 1988; Straus et al., 1980). Therefore, the presence of one or more of these risk markers does not necessarily mean impending lethal violence. It means that the situation must be monitored and that the possibility of lethal violence should be addressed specifically with the batterer, his female victim, and other concerned parties. Moreover, therapists should be prepared to deal with that possibility, including facilitating the connection with family and community resources, such as shelter services, she could use if necessary. In fact, clinical wisdom suggests that it is never too early to help a woman take responsibility to secure safety for herself and her children (Hamberger & Holtzworth-Munroe, 1994). As Rosenbaum and Maiuro (1990) state clearly. "There is, as yet, no scientific formula for determining when a client has crossed the threshold into the realm of dangerousness. It remains a matter of clinical judgment guided by experience and knowledge of the research literature on interpersonal violence" (p. 292).

9.27.5 CLINICAL INTERVENTIONS WITH PERPETRATORS OF DOMESTIC VIOLENCE

9.27.5.1 Treatment Programs for Male Batterers

Since the 1970s there has been a proliferation of treatment programs for men who abuse their female partners (Eddy & Meyers, 1984; Jennings, 1987). These programs vary in the duration of treatment, the number of sessions required of participants, the specialty and level of training of the counselors, the agency or agencies sponsoring the treatment program, the referral source, and the source of funding (Edleson & Tolman, 1992; Pirog-Good & Stets-Kealy, 1985). Many programs intervene only with perpetrators of violence, and some programs work with both men and their female victims (Jennings & Jennings, 1991). Accordingly, the two dominant treatment modalities are group therapy and couple therapy (Feldman & Ridley, 1995; Margolin & Burman, 1993).

9.27.5.1.1 Group therapy approaches

Most treatment programs for men who batter their female partners use some version of group therapy. The group format offers several practical and therapeutic advantages over other treatment formats. It is a cost-effective operation that allows a relatively small number of specialized professionals to serve a large number of clients needing intervention (Carden, 1994; Ganley, 1987). It is an affordable form of treatment, especially for poor or financially strapped clients (Rosenbaum & Maiuro, 1990). Moreover, group therapy is a viable sentencing option for judges, especially for first-offense cases of wife assault.

In terms of therapeutic advantages, groups are presumed to: (i) reduce batterers' social isolation; (ii) provide a safe environment where they can deal with potentially volatile issues and challenge fellow group members who directly or indirectly engage in abusive behavior; and (iii) expose men to alternative coping models. Furthermore, group members offer the social support necessary for the batterer to stop minimizing his problems, address the shame he may feel about the abusive behavior, become accountable for his actions, and consider nonabusive ways of acting (Edleson & Tolman, 1992; Rosenbaum & Maiuro, 1990).

In general, group treatment programs follow a cognitive-behavioral model informed by social learning theory. Many also have an underlying profeminist orientation to the issue. From a social learning perspective, wife assault is viewed as a pattern of behavior familiar to men in part because they have witnessed such behaviors in their families of origin (Hamberger & Lohr, 1989; O'Leary, 1988). Moreover, the men's use of violence against their female partners is presumed to have a functional significance for the men and becomes a regular part of their interpersonal relationships. In accordance with feminists' ideas about the social construction of gender and power in intimate relationships, many treatment programs for men who batter also understand violence against women as "one of several forms of controlling behaviors men use against women. Such control is socialized early in a young boy's childhood and is maintained into adulthood through a continued societal message that legitimizes male violence against women" (Eisikovits & Edleson, 1989, pp. 392–393).

This conceptualization of men's violence against women as learned, functional, and maintained by interpersonal and social dynamics has favored the development of highly structured treatments focused on the cessation of violent behaviors over unstructured growth-oriented approaches (see, e.g., Jennings, 1987; Wallace & Nosko, 1993). Most group treatment programs for wife assault try to: (i) educate men about the causes, dynamics, and consequences of violence; (ii) teach them how to recognize and deal with anger without resorting to violence; (iii) reduce their general level of arousal through relaxation and self-control techniques; (iv) teach them to use time-out techniques to avoid emotionally arousing situations and promote rational problem solving; (v) correct the distorted views and attitudes about intimate relationships that the men have used to justify the abusive behaviors; and (vi) teach them appropriate conflict resolution skills (Edleson & Tolman, 1992; Geffner & Rosenbaum, 1990; Rosenbaum & Maiuro, 1990; Saunders, 1989). Many treatment programs also focus on consciousness-raising and changing sexist attitudes (e.g., Gondolf & Russell, 1986; Pence & Paymar, 1993). These programs put a special premium on the integration of community resources to prevent and eradicate the maltreatment of women by men. They emphasize the men's accountability for the violent behaviors, explore sex roles, and advocate for an egalitarian ethic of intimate relationships (Pence & Paymar, 1993).

Despite these multiple goals, the success of treatment programs for men who batter is judged typically by the reduction in the number of men who engage in violence against their female partners following treatment. However, accurate estimates of wife assault recidivism and cessation following treatment are difficult to obtain. Moral reproach for violent behavior and possible negative social and legal consequences of treatment failure can heighten people's tendency to underreport violence. Rosenfeld (1992) lists the following additional methodological difficulties in obtaining unbiased estimates of wife assault cessation: (i) most outcome studies report data only for men who completed treatment and were available at follow-up time; (ii) there is a low response rate to all follow-up data-gathering techniques; (iii) people who complete post-treatment assessments compare favorably with those who do not in important risk markers for wife assault such as economic resources, alcohol problems, violence in their family of origin, and severity of violence against women; and (iv) many treatment programs include different types of male batterers in the same treatment group while excluding difficult or inappropriate men from treatment. Considering these issues, it seems reasonable to assume that the available data may overestimate the rate of wife assault cessation and underestimate the rate of violence recidivism following treatment.

Feazell, Mayers, and Deschner's (1984) survey of 90 treatment programs for male batterers revealed that between 66% and 75% of batterers ceased their violence for at least a year following treatment. Another survey of 16 group programs for men who batter (Pirog-Good & Stets-Kealy, 1985) found a four-month average cessation rate of 84%. However, as Edleson (1990) pointed out, neither of these studies indicates how the follow-up data were collected, who provided the information, and how the programs computed the cessation rates.

Edleson and Tolman (1992) reviewed the results of 19 outcome studies on group treatment for male batterers, including 16 cognitive-behavioral programs. Most follow-up assessments relied on the CTS to measure violence. A small number of programs used police records and interviews. The follow-up time ranged from three months to three years, with two-thirds of programs completing the assessment at least six months following treatment. Wife assault cessation rates for men who completed treatment ranged from 59% to 87%. Lower violence cessation rates "tended to occur in programs with lengthier follow-up and when success was based on women's reports rather than arrest or self-reports" (Tolman & Bennett, 1990, p. 104). Moreover, approximately two-thirds of men continued their use of threats and emotionally abusive behavior following treatment.

Other evaluations of treatment effectiveness have compared follow-up reports of violence among treatment dropouts and untreated men. As the name suggests, dropouts are men who abandoned treatment. Untreated men were referred for counseling but never started treatment because of practical issues or lack of motivation. At least two studies have found that men who complete treatment and men who do not have comparable rates of wife assault cessation at post-treatment (Edleson & Grunszki, 1988; Hamberger & Hastings, 1988). Rosenfeld's (1992) review of this literature reported overall wife assault cessation rates (based either on police or victim's reports) of 61% and 86% among treatment dropouts and untreated men, respectively. According to Rosenfeld (1992), the comparable rates of wife assault cessation among dropouts, untreated men, and treated men "cast doubt on the specific effect of any spouse abuse treatment program" (p. 218). However, other researchers maintain that expectations for radical change in the behaviors of violent men may be unrealistic (Aldarondo, 1996; Jennings, 1990). As Dutton (1995b) has stated, "It seems overly optimistic to expect a 16-week treatment group to stop all abuse immediately after treatment" (p. 220).

9.27.5.1.2 Couple therapy

There is a growing literature in couple therapy on wife assault (e.g., Balcom, 1991; Gutsche & Murray, 1991; Hansen & Harway, 1993; Jenkins, 1990). Although many couple therapists differ in their theoretical and political orientation towards wife assault, most of them view violent behaviors not so much in terms of social and psychological determinants, as in terms of the interpersonal dynamics that promote and sustain the abuse of women by men. Mantooth, Geffner, Franks, and Patrick (1987) have argued that couple therapy offers a safe and structured environment in which partners can express their feelings, discuss emotionally charged issues, and learn about violence and how to deal with it together. In addition, Margolin and Burman (1993) have suggested that couple therapy: (i) gives partners the opportunity to identify and attempt to alter interactional patterns that promote and sustain the violence; (ii) helps men learn to monitor their emotions and to communicate to their female partners their desire for physical and emotional distance; and (iii) empowers women by teaching them to identify the cues that signal the man's escalation of anger and aggression.

Unlike group therapy, which has been embraced uncritically by those committed to working with batterers, couple therapy for wife assault has been the focus of debate. Critics argue that interactional theories and interventions characteristic of couple therapy are ill-equipped to deal with issues of coercion and violence (Bograd, 1992), that abusive behaviors disappear in systemic formulations (Avis, 1992), and that couple therapy diffuses men's responsibility for their actions by implying that the women should be working on how to stop their victimization (Adams, 1988). Above all, critics argue that couple therapy can cause emotional and physical harm for women who are called to bring their partners' violence into the open (Bograd, 1984; Kaufman, 1992). Unfortunately, as Bograd (1994) has pointed out, although this is possibly one of the most controversial topics in the field, no study has explored the safety of the women when couple therapy is recommended or sought by the batterer.

The apparent incongruity between couple therapy and group treatment for male batterers masks the fact that in practice couple therapy for wife assault relies heavily on short-term, cognitive-behavioral strategies akin to those used in group treatment. Some of these strategies include: (i) teaching the partners to identify cues of violence early on; (ii) teaching them when and how to use time-outs; (iii) using communication training and problem-solving techniques to help partners deal with anger and conflict; (iv) role playing nonviolent interactions, new communication skills, or particular disagreements; (v) teaching men to identify and change negative self-talk and other irrational beliefs that may be related to their use of violence; and (vi) training the couple to use relaxation techniques to reduce stress and promote problem-solving (Geffner, Mantooth, Franks, & Rao, 1989; Neidig & Friedman, 1984; Rosenbaum & O'Leary, 1986).

In addition to traditional couple therapy, some programs use a couples' group format (Deschner, McNeil, & Moore, 1986; Mantooth et al., 1987; Neidig, Friedman, & Collins, 1985). Geffner et al. (1989) maintain that this format has an advantage over working with individual couples because it provides the opportunity for the partners to interact with others who have similar problems and to share information about what has helped them deal with similar issues. In addition, partners in couples' group counseling often establish friendships during the counseling process that may help reduce the partner's isolation and estrangement from others.

In terms of outcome studies, there is a paucity of research on the effectiveness of couple therapy for wife assault. Lindquist, Telch, and Taylor (1985) found that 50% of the couples reported at least one incident of violence within six weeks following treatment. Moreover, all couples reported acts of violence after six months. Harris (1986) reported an undefined success rate of 73% among 30 couples assessed somewhere between two months and three years after treatment. Taylor (1984) reported a six-month wife assault cessation rate of 65% among 50 couples. However, Taylor failed to report both the source of data and the techniques used to obtain this information.

Concerning outcome data on couples' group counseling, Neidig et al. (1985) reported that of 100 couples, 87% of all participants (men and women) were nonviolent four months following treatment. They also found that everyone evaluated reported positive changes in measures of locus of control and marital interaction. Unfortunately, the authors omitted the source of follow-up data, and the report is unclear about the criteria for inclusion in the follow-up. Similarly, Deschner et al. (1986) reported that couples who completed treatment reported significant increases in marital satisfaction and significant reductions in the frequency of argument episodes of anger. Eight months after treatment, 8 out of 15 couples (53%) reported no violence.

9.27.6 SUMMARY AND FUTURE DIRECTIONS

Intimate violence is an extensive and serious problem in both heterosexual and homosexual relationships. While wife assault often increases in frequency and severity over time, it is also characterized by the cessation or interruption of violence for prolonged periods of time. Moreover, different types of men who batter can be identified on the basis of both violent behavior and psychological characteristics.

The heterogeneity among perpetrators of violence has important theoretical and practical implications. In terms of theory, the above considerations point to the limitations of what some researchers have called single-factor explanations of intimate violence (Dutton, 1995b, 1995c; Gelles, 1993). As O'Leary (1993, p. 13) has stated bluntly, "attempts to explain and change various acts of physical aggression among family members through singular accounts, be they feminist, constructionist, sociological, biological, or psychological are doomed." Instead, there is a need to continue to develop multifactor theories of intimate violence such as Dutton's (1995c) nested ecological theory of wife assault, Edleson and Tolman's (1992) ecological framework, Holtzworth-Munroe and Stuart's (1994) developmental model of marital violence, Letellier's (1994) multidimensional theory of battering among gay and bisexual men, and O'Leary's (1988, 1993) social learning theory of physical aggression between spouses. These theories could help to understand how different forms of violence arise and evolve and what specific risk markers may be associated with each.

On the other hand, there is often the assumption that multifactor or multidimensional theories are going to reconcile many of the inconsistencies in the literature and will eventually improve the efficacy of treatment approaches (see, e.g., Carden, 1994; Feldman & Ridley, 1995). Such a premise ignores the methodological and clinical limitations in the research and treatment of physical violence and other forms of abuse in intimate relationships. Appropriate control groups are difficult to find and other experimental manipulations raise serious ethical dilemmas. Furthermore, broad theories often do not have clear clinical applications. Those committed to working with intimate violence need not only comprehensive maps but also specific theories of practice to deal with individual clients under particular conditions.

9.27.7 REFERENCES

Adams, D. (1988). Treatment models for men who batter: A profeminist analysis. In K. Yllö & M. Bograd (Eds.), *Feminist perspectives on wife abuse* (pp. 176–199). Newbury Park, CA: Sage.

Aldarondo, E. (1996). Cessation and persistence of wife assault: A longitudinal analysis. *American Journal of Orthopsychiatry, 66,* 141–151.

Aldarondo, E., & Straus, M. A. (1994). Screening for physical violence in couple therapy: Methodological, practical, and ethical considerations. *Family Process, 33,* 425–439.

Aldarondo, E., & Sugarman, D. (1996). Risk makers analysis of cessation and persistence of wife assault. *Journal of Consulting and Clinical Psychology, 64,* 1010–1019.

American Psychological Association (1996). *Violence and the family: Report of the American Psychological Association presidential task force on violence and the family*. Washington, DC: Author.

Avis, J. M. (1992). Where are all the family therapists? Abuse and violence within families and family therapy's response. *Journal of Marital and Family Therapy, 18.* 225–232.

Balcom, D. (1991). Shame and violence: Considerations in couples' treatment. *Journal of Independent Social Work, 5,* 165–181.

Barling, J. N., O'Leary, K. D., Jouriles, E. V., Vivian, D., & MacEwen, C. (1987). Factor similarity of the Conflict Tactics Scales across samples, spouses, and sites: Issues and implications. *Journal of Family Violence, 2,* 37–53.

Bodin, A. M. (1992). *Relationship Conflict Inventory*. Palo Alto, CA: Mental Research Institute.

Bograd, M. (1984). Family systems approaches to wife battering: A feminist critique. *American Journal of Orthopsychiatry, 54,* 558–568.

Bograd, M. (1992). Values in conflict: Challenges to family therapists' thinking. *Journal of Marital and Family Therapy, 18,* 245–256.

Bograd, M. (1994). Battering, competing clinical models, and paucity of research: Notes to those in the trenches. *The Counseling Psychologist, 22,* 593–597.

Bohannon, J. R., & Dosser, D. A. (1995). Using couple data to determine domestic violence rates: An attempt to replicate previous work. *Violence and Victims, 10,* 133–141.

Bologna, M. J., Waterman, C. K., & Dawson, L. J. (1987, July). *Violence in gay and lesbian relationships: Implications for practitioners and policy makers*. Paper presented at the 3rd National Family Violence Research Conference, Durham, NH.

Bowker, L. H. (1983). *Beating wife-beating*. Lexington, MA: Lexington Books.

Browne, A. (1987). *When battered women kill*. New York: The Free Press.

Browne, A., & Williams, K. R. (1993). Gender, intimacy, and lethal violence: Trends from 1976 through 1987. *Gender and Society, 7,* 78–98.

Browning, J., & Dutton, D. (1986). Assessment of wife assault with the Conflict Tactics Scale: Using couple data to quantify the differential reporting effect. *Journal of Marriage and the Family, 48,* 375–379.

Brush, L. D. (1990). Violent acts and injurious outcomes in married couples: Methodological issues in the National Survey of Families and Households Gender and Society, 4, 56–67.

Buzawa, E. S., & Buzawa, C. G. (Eds.) (1996). *Do arrests and restraining orders work?* Newbury Park, CA: Sage.

Campbell, J. C. (1995). Prediction of homicide of and by women. In J. C. Campbell (Ed.), *Assessing dangerousness: Violence by sexual offenders, batterers, and child abusers* (pp. 96–113). Newbury Park, CA: Sage.

Cantos, M., Neidig, P. H., & O'Leary K. D. (1994). Injuries of women and men in a treatment program for domestic violence. *Journal of Family Violence, 9,* 113–124.

Carden, A. D. (1994). Wife abuse and the wife abuser: Review and recommendations. *The Counseling Psychologist, 22,* 539–582.

Cascardi, M., Langhinrichsen, J., & Vivian, D. (1992). Marital aggression: Impact, injury, and health correlates for husbands and wives. *Archives of Internal Medicine, 152,* 1178–1184.

Coleman, V. E. (1994). Lesbian battering: The relationship between personality theory and the perpetuation of violence. *Violence and Victims, 9,* 139–152.

Deschner, J. P., McNeil, J. S., & Moore, M. G. (1986). A treatment model for batterers. *Social Casework, 67,* 55–60.

Dutton, D. G. (1994). The origin and structure of the abusive personality. *Journal of Personality Disorders, 8,* 181–191.

Dutton, D. G. (1995a). A scale for measuring propensity for abusiveness. *Journal of Family Violence, 10,* 203–221.

Dutton, D. G. (1995b). Intimate abusiveness. *Clinical Psychology: Science and Practice, 2,* 207–224.

Dutton, D. G. (1995c). *The domestic assault of women: Psychological and criminal justice perspectives*. Boston: Allyn & Bacon.

Eddy, M. J., & Myers, T. (1984). *Helping men who batter: A profile of programs in the US*. Austin, TX: Texas Department of Human Resources.

Edleson, J. L. (1990). Judging the success of interventions with men who batter. In D. J. Besharov (Ed.), *Family violence: Research and public policy issues* (pp. 130–145). Washington, DC: AEI Press.

Edleson, J. L., & Grusznski, R. J. (1989). Treating men who batter: Four years of outcome data from the Domestic Abuse Project. *Journal of Social Service Research, 12,* 3–22.

Edleson, J. L., & Tolman, R. M. (1992). *Intervention for men who batter: An ecological approach*. Newbury Park, CA: Sage.

Eisikovits, Z. C., & Edleson, J. L. (1989). Intervening with men who batter: A critical review of the literature. *Social Service Review, 63,* 384–414.

Fagan, J. L., & Browne, A. (1994). Violence between spouses and intimates: Physical aggression between women and men in intimate relationships. In A. J. Reiss & J. A. Roth (Eds.), *Understanding and preventing violence* (Vol. 3, pp. 115–292). Washington, DC: National Academy Press.

Feazell, C. S., Mayers, R. S., & Deschner, J. (1984). Services for men who batter: Implications for programs and policies. *Family Relations, 33,* 217–223.

Feldman, C. M., & Ridley, C. (1995). The etiology and treatment of domestic violence between adult partners. *Clinical Psychology: Science and Practice, 2,* 317–348.

Ferraro, K. J., & Johnson, J. M. (1983). How women experience battering: The process of victimization. *Social Problems, 30,* 325–339.

Finkelhor, D., & Yllö, K. (1985). *License to rape: Sexual abuse of wives*. New York: Brunner/Mazel.

Follingstad D. R., Laughlin, J. E., Polek, D. S., Rutledge, L. L., & Hause, E. S. (1991). Identification of patterns of wife assault. *Journal of Interpersonal Violence, 6,* 187–204.

Ganley, A. L. (1987). Perpetrators of domestic violence: An overview of counseling the court-mandated client. In D.

J. Sonkin (Ed.), *Domestic violence on trial: Psychological and legal dimensions of family violence* (pp. 155–173). New York: Springer.

Geffner, R., Mantooth, C, Franks, D., & Rao, L. (1989). A psychoeducational, conjoint therapy approach to reducing family violence. In P. L. Caesar & L. K. Hamberger (Eds.), *Treating men who batter* (pp. 103–133). New York: Springer.

Geffner, R., & Rosenbaum, A. (1990). Characteristics and treatment of batterers. *Behavioral Sciences and the Law, 8,* 131–140.

Gelles, R. J. (1993). Through a sociological lens: Social structure and family violence. In R. Gelles & D. Loseke (Eds.), *Current controversies on family violence* (pp. 31–46). Newbury Park, CA: Sage.

Giles-Sims, J. (1983). *Wife battering: A systems theory approach*. New York: Guilford.

Goetting, A. (1989). Men who kill their mates: A profile. *Journal of Family Violence, 4,* 285–296.

Goetting, A. (1991). Female victims of homicide: A portrait of their killers and the circumstances of their death. *Violence and Victims, 6,* 159–168.

Gondolf, E. W., Mulvey, E. P., & Lidz, C. W. (1989). Family violence reported in a psychiatric emergency room. *Journal of Family Violence, 4,* 249–258.

Gondolf, E. W., & Russell, D. (1986). The case against anger control treatment programs for batterers. *Response to the Victimization of Women and Children, 9,* 2–5.

Goodman, L. A., Browne, A., Fitzgerald, L., Keita, G. P., & Russo, N. F. (1994). *No safe haven: Violence against women at home, at work, and in the community*. Washington, DC: APA.

Gottman, J. M., Jacobson, N. S., Rushe, R. H., Shortt, J. W., Babcock, J, La Taillade, J. J., & Waltz, J. (1995). The relationship between heart rate reactivity, emotionally aggressive behavior, and general violence in batterers. *Journal of Family Psychology, 9,* 227–248.

Gutsche, S., & Murray, M. (1991). The feminist meets the cybernetician: An integrated approach to spousal violence. *Journal of Strategic and Systemic Therapies, 10,* 76–91.

Hamberger, L. K., & Hastings, J. E. (1988). Skill training for treatment of spouse abusers: An outcome study. *Journal of Family Violence, 3,* 121–130.

Hamberger, L. K., & Holtzworth-Munroe, A. (1994). Partner violence. In F. M. Dattilio & A. Freedman (Eds.), *Cognitive-behavioral strategies in crisis intervention* (pp. 302–324). New York: Guilford.

Hamberger, L. K., & Lohr, J. M. (1989). Proximal causes of spouse abuse: A theoretical analysis of cognitive-behavioral interventions. In P. L. Caesar & L. K. Hamberger (Eds.) *Treating men who batter* (pp. 53–76). New York: Springer.

Hansen, M., & Harway, M. (Eds.) (1993). *Battering and family therapy: A feminist perspective*. Newbury Park, CA: Sage.

Harris, J. (1986). Counseling violent couples using Walker's model. *Psychotherapy, 23,* 613–621.

Hart, S. D., Dutton, D. G., & Newlove, T. (1993). The prevalence of personality disorder amongst wife assaulters. *Journal of Personality Disorders, 7,* 329–341.

Holtzworth-Munroe, A., & Stuart, G. L. (1994). Typologies of male batterers: Three subtypes and the differences among them. *Psychological Bulletin, 116,* 476–497.

Horton, A. L., & Johnson, B. (1993). Profile and strategies of women who have ended abuse. *Families in Society: The Journal of Contemporary Human Services, 74,* 481–492.

Hotaling, G. T., & Sugarman, D. B. (1986). An analysis of risk markers in husband to wife violence: The current state of the knowledge. *Violence and Victims, 1,* 101–124.

Hudson W. W., & McIntosh, S. R. (1981). The assessment of spouse abuse: Two quantifiable dimensions. *Journal of Marriage and the Family, 43,* 873–888.

Jaffe, P. G., & Wolfe, D. A. (1990). *Children of battered women*. Newbury Park, CA: Sage.

Jenkins, A. (1990). *Invitations to responsibility: The therapeutic engagement of men who are violent and abusive*. Adelaide, Australia: Dulwich Centre.

Jennings, J. P. (1987). History and issues in the treatment of battering men: A case for unstructured group therapy. *Journal of Family Violence, 2,* 193–213.

Jennings, J. L. (1990). Preventing relapse versus stopping domestic violence: Do we expect too much too soon from battering men? *Journal of Family Violence, 5,* 43–60.

Jennings, J. P., & Jennings, J. L. (1991). Multiple approaches to the treatment of violent couples. *The American Journal of Family Therapy, 19,* 351–362.

Kahuna, V. (1990). Compounding the triple jeopardy: Battering in lesbians of color relationships. *Women and Therapy, 9,* 169–184.

Kaufman, G., Jr. (1992). The mysterious disappearance of battered women in family therapists' office: Male privilege colluding with male violence. *Journal of Marital and Family Therapy, 18,* 233–243.

Kaufman, G. K., Jasinski, J. L., & Aldarondo, E. (1994). Sociocultural status and incidence of marital violence in Hispanic families. *Violence and Victims, 9,* 207–222.

Kelly, C. E., & Washafsky, L. (1987, July). *Partner abuse in gay and lesbian couples*. Paper presented at the 3rd National Family Violence Research Conference, Durham, NH.

Kilpatrick, D. G., Edmunds, C. S., & Seymour, A. K. (1992). *Rape in America: A report to the nation*. Arlington, VA: National Victims Center and Medical Center of South Carolina.

Letellier, P. (1994). Gay and bisexual domestic violence: Challenges to feminist theory and responses to violence. *Violence and Victims, 9,* 95–106.

Lewis, B. Y. (1985). The Wife Abuse Inventory: A screening device for the identification of abused women. *Social Work, 30,* 32–35.

Lie, G., & Gentlewarrior, S. (1991). Intimate violence in lesbian relationships: Discussion of survey findings and practice implications. *Journal of Social Service Research, 15,* 41–59.

Lie, G., Schilit, R., Bush, J., Montague, M., & Reyes, L. (1991). Lesbians in currently aggressive relationships: How frequently do they report aggressive past relationships? *Violence and Victims, 6,* 121–135.

Lindquist, C. U., Telch, C. F., & Taylor, J. (1985). Evaluation of a conjugal violence treatment program: A pilot study. *Behavioral Counseling and Community Interventions, 3,* 76–90.

Lockhart L. L., White, B. W., Causby, V., & Isaac, A. (1994). Letting out the secret: Violence in lesbian relationships. *Journal of Interpersonal Violence, 9,* 469–492.

Mantooth, C. M., Geffner, R., Franks, D., & Patrick, J. (1987). *Family preservation: A treatment manual for reducing couple violence*. Tyler, TX: University of Texas at Tyler Press.

Margolies, L., & Leeder, E. (1995). Violence at the door: Treatment of lesbian batterers. *Violence against Women, 1,* 139–157.

Margolin, G., & Burman, B. (1993). Wife abuse versus marital violence: Different terminologies, explanations, and solutions. *Clinical Psychology Review, 13,* 59–73.

Marshall, L. L. (1992). Development of the Severity of Violence Against Women Scales. *Journal of Family Violence, 7,* 103–121.

McLaughlin, I. G., Leonard, K. E., & Senchak, M. (1992). Prevalence and distribution of premarital aggression among couples applying for a marriage license. *Journal of Family Violence, 7,* 309–319.

Morse, B. J. (1995). Beyond the conflict tactics scale: Assessing gender differences in partner violence. *Violence and Victims, 10,* 251–272.

Neidig, P. H., & Friedman, D. H. (1984). *Spouse abuse: A treatment program for couples*. Champaign, IL: Research Press.

Neidig, P. H., Friedman, D. H., & Collins, B. S. (1985). Domestic conflict containment: A spouse abuse treatment program. *Social Casework, 66,* 195–204.

O'Leary, K. D. (1988). Physical aggression between spouses: A social learning theory perspective. In V. Van Hasselt, R. Morrison, A. Bellack, & M. Hersen (Eds.), *The handbook of family violence* (pp. 31–55) New York: Plenum.

O'Leary, K. D. (1993). Controversies regarding psychological explanations of family violence. In R. J. Gelles & D. Loseke (Eds.), *Current controversies on family violence* (pp. 7–30). Newbury Park, CA: Sage.

O'Leary, K. D., Barling, J., Arias, I., Rosenbaum, A., Malone, J., & Tyree, A. (1989). Prevalence and stability of physical aggression between spouses: A longitudinal analysis. *Journal of Consulting and Clinical Psychology, 57,* 263–268.

O'Leary, K. D., & Curley, A. D. (1986). Assertion and family violence: Correlates of abuse. *Journal of Marital and Family Therapy, 12,* 281–289.

O'Leary, K. D., & Murphy, C. (1992). Clinical issues in the assessment of spouse abuse. In R. T. Ammerman & M. Hersen (Eds.), *Assessment of family violence: A clinical and legal source book* (pp. 26–46). New York: Wiley.

O'Leary, K. D., Vivian, D, & Malone, J. (1992). Assessment of physical aggression against women in marriage: The need for multimodal assessment. *Behavior Assessment, 14,* 5–14.

Pence, E., & Paymar, M. (1993). *Education groups for men who batter: The Duluth model.* New York: Springer.

Pirog-Good, M., & Stets-Kealy, J. (1985). Male batterers and battering prevention programs: A national survey. *Response to the Victimization of Women and Children, 8,* 8–12.

Reiss, A. J., & Roth, J. A. (Eds.) (1993). *Understanding and preventing violence Vol. 2,* Washington, DC: National Academy Press.

Rhodes, N. R. (1992). The assessment of spousal abuse: An alternative to the Conflict Tactics Scale. In E. C. Viano (Ed.), *Intimate violence: Interdisciplinary perspectives* (pp. 27–36). Philadelphia: Hemisphere.

Rodenburg, F. A., & Fantuzzo, J. W. (1993). The Measure of Wife Abuse: Steps toward the development of a comprehensive assessment technique. *Journal of Family Violence, 8,* 203–228.

Rosenbaum, A., & Maiuro, R. (1990). Perpetrators of spouse abuse. In R. T. Ammerman & M. Hersen (Eds.), *Treatment of family violence: A source book* (pp. 280–309). New York: Plenum.

Rosenbaum, A., & O'Leary, K. D. (1981). Marital violence: Characteristics of abusive couples. *Journal of Consulting and Clinical Psychology, 49,* 63–71.

Rosenbaum, A., & O'Leary, K. D. (1986). The treatment of marital violence. In N. S. Jacobson & A. S. Gurman (Eds.), *Clinical handbook of marital therapy* (pp. 385–405). New York: Guilford.

Rosenfeld, B. D. (1992). Court-ordered treatment of spouse abuse. *Clinical Psychology Review, 12,* 205–226.

Russo, N. F., Koss, M. P., & Goodman, L. (1995). Male violence against women: A global health and development issue. In L. L. Adler & F. L. Denmark (Eds.), *Violence and the prevention of violence* (pp. 121–128). Westport, CT: Praeger.

Saunders, D. G. (1989). Cognitive and behavioral interventions with men who batter: Application and outcome. In P. L. Caesar & L. K. Hamberger (Eds.), *Treating men who batter* (pp. 77–99). New York: Springer.

Saunders, D. G. (1992a). A typology of men who batter: Three types derived from cluster analysis. *American Journal of Orthopsychiatry, 62,* 264–275.

Saunders, D. G. (1992b). Woman battering. In R. T. Ammerman & M. Hersen (Eds.), *Assessment of family violence* (pp. 208–235). New York: Wiley.

Saunders, D. G. (1995). Prediction of wife assault. In J. C. Campbell (Ed.), *Assessing dangerousness: Violence by sexual offenders, batterers, and child abusers* (pp. 68–95). Newbury Park, CA: Sage.

Sedlak, A. J. (1988). Prevention of wife abuse. In V. Van Hasselt, R. Morrison, A. Bellack, & M. Hersen (Eds.), *The Handbook of family violence* (pp. 319–358). New York: Plenum.

Sheppard, M. E., & Campbell J. A. (1992). The Abusive Behavior Inventory: A measure of psychological and physical abuse. *Journal of Interpersonal Violence, 7,* 291–305.

Sherman, L. W. (1992). *Policing domestic violence: Experiments and dilemmas.* New York: Free Press.

Snyder, D. K., & Snow, A. (1995). *Evaluating couples' aggression in marital therapy.* Paper presented at the annual meeting of the American Psychological Association, New York.

Sonkin, D. J. (1987). The assessment of court-mandated male batterers. In D. J. Sonkin (Ed.), *Domestic violence on trial: Psychological and legal dimensions of family violence* (pp. 174–196). New York: Springer.

Stets, J. E., & Straus, M. A. (1990). Gender differences in reporting marital violence and its medical and psychological consequences. In M. A. Straus & R. J. Gelles (Eds.). *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families* (pp. 151–165). New Brunswick, NJ: Transaction Publishers.

Straus, M. A. (1979). Measuring intrafamily conflict and violence: The Conflict Tactics Scales. *Journal of Marriage and the Family, 41,* 75–88.

Straus, M. A. (1990). The Conflict Tactics Scale and its critics: An evaluation and new data on validity and reliability. In M. A. Straus & R. J. Gelles (Eds.), *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families* (pp. 49–73). New Brunswick, NJ: Transaction Publishers.

Straus, M. A. (1992). Children as witness to marital violence: A risk factor for life long problems among a nationally representative sample of American men and women. In D. F. Schwarz (Ed.), *Children and violence: Report of the 23rd Ross roundtable on critical approaches to common pediatric problems* (pp. 98–104). Columbus, OH: Ross Laboratories.

Straus, M. A. (1993). Physical assaults by wives: A major social problem. In R. J. Gelles & D. Loseke (Eds.), *Current controversies on family violence* (pp. 67–87). Newbury Park, CA: Sage.

Straus, M. A. (1996). *Manual for the Conflict Tactics Scales.* Durham, NH: University of New Hampshire, Family Research Laboratory.

Straus, M. A., & Gelles, R. J. (1990). *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families.* New Brunswick, NJ: Transaction Publishers.

Straus, M. A., & Gelles, R. J., & Steinmetz, S. K. (1980). *Behind closed doors: Violence in the American family.* Garden City, NY: Anchor Press.

Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The revised Conflict Tactics Scales (CTS2). *Journal of Family Issues, 17,* 283–316.

Straus, M. A., & Kaufman-Kantor, G. (1994, June). *Change in spouse assault rates from 1975 to 1992: A comparison of three national surveys in the United States.* Paper presented at the 13th World Congress of Sociology, Bielefeld, Germany.

Sugarman, D. B., & Hotaling, G. T. (1989). Violent men in intimate relationships: An analysis of risk markers. *Journal of Applied Social Psychology, 19,* 1034–1048.

Szinovacz, M. E. (1983). Using couple data as a methodological tool: The case of marital violence. *Journal of Marriage and the Family, 45,* 663–644.

Taylor, J. W. (1984). Structured conjoint therapy for spouse abuse cases. *Social Casework, 65,* 11–18.

Tham, S. W., Ford, T. J., & Wilkinson, D. G. (1995). A survey of domestic violence and other forms of abuse. *Journal of Mental Health, 4,* 317–321.

Tolman, R. M. (1989). The initial development of a measure of psychological maltreatment of women by their male partners. *Violence and Victims, 4,* 159–178.

Tolman, R. M., & Bennett L. W. (1990). A review of quantitative research on men who batter. *Journal of Interpersonal Violence, 5,* 87–118.

Walker, L. E. (1984). *The battered woman syndrome*. New York: Springer.

Walker, L. E. (1995). Current perspectives on men who batter women. *Journal of Family Psychology, 9,* 264–271.

Wallace, B., & Nosko, A. (1993). Working with shame in the group treatment of male batterers. *International Journal of Group Psychotherapy, 43,* 45–61.

Woffordt, S., Mihalic, D. E., & Menard, S. (1994). Continuities in marital violence. *Journal of Family Violence, 9,* 195–225.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.28
Children with Aggressive Sexual Behavior

BARBARA L. BONNER
University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

and

WILLIAM E. FAHEY
Virginia Commonwealth University, Richmond, VA, USA

9.28.1 INTRODUCTION 454
9.28.2 NATURE 454
9.28.3 PREVALENCE 456
9.28.4 ASSESSMENT 457
  9.28.4.1 Child Evaluation 458
    9.28.4.1.1 Diagnosis and overall psychopathology 458
    9.28.4.1.2 Behavior and affective problems 458
    9.28.4.1.3 Self-esteem 458
    9.28.4.1.4 Sexual behavior problems 458
  9.28.4.2 Parent/Caretaker Evaluation 459
    9.28.4.2.1 Psychological status 459
    9.28.4.2.2 Level of stress 459
  9.28.4.3 Family Evaluation 459
9.28.5 TREATMENT 459
  9.28.5.1 Children with Sexual Behavior Problems Program 461
    9.28.5.1.1 The cognitive-behavioral group therapy approach 461
    9.28.5.1.2 Case study: cognitive-behavioral treatment approach 462
    9.28.5.1.3 Dynamic play group therapy approach 462
    9.28.5.1.4 Case study: dynamic play therapy treatment approach 464
    9.28.5.1.5 Treatment 464
  9.28.5.2 STEP Program 464
9.28.6 FUTURE DIRECTIONS 465
9.28.7 SUMMARY 465
9.28.8 REFERENCES 465

9.28.1 INTRODUCTION

Sexual abuse and aggressive sexual behavior continue to be serious and widespread problems. Prevalence rates for being sexually abused as a child in the USA have been reported at 27% for females and 16% for males (Finkelhor, Hotaling, & Lewis, 1990). The short- and long-term psychological effects of sexual abuse have been well documented in the literature.

Historically, it has been believed that adult males were primarily responsible for sexual aggression and assault. However, more recent studies indicate that not only adult males, but also adult females, commit sexual crimes against children. Of even more concern are crime statistics showing that these crimes are committed also by juveniles. The 1990 uniform crime reports indicated that 15% of all arrests for forcible rape were of youth under 18 years of age (Federal Bureau of Investigation, 1990).

In the 1990s, professional attention has begun to focus on sexual behavior problems in young children. Given that many people have difficulty accepting the premise that children are sexual beings, there is an even more extreme reaction to the idea that children can have problematic or aggressive sexual behavior. In the mental health field, professionals have devoted many years of research and clinical treatment to aggressive sexual behaviors in adolescents and adults. As a field, we have now reached the point of needing to better understand children who engage in aggressive sexual behavior.

Until recently, sexual interactions involving children have been typically ignored with the explanation that the children were "playing doctor" or just being curious. However, we have begun to realize that while the sexual behavior of some children may be "normal" or developmentally expected, the sexual behavior of other children is of great concern and requires professional assessment and mental health intervention. Only since the late 1980s has the mental health field begun to provide clinical interventions and conduct research for children with aggressive sexual behavior.

This chapter will review what has been reported in the literature on children with aggressive sexual behavior including definitions, typologies, characteristics of the children and their families, assessment of children with aggressive sexual behavior, and current treatment approaches.

9.28.2 NATURE

There is no definition of children with aggressive sexual behavior that is widely accepted and utilized by the various professions that are involved with these children, that is, mental health, medicine, law enforcement, and child protective services. The children have been referred to as mini-perps, children who molest (Gil & Johnson, 1993), abuse-reactive children (Cunningham & MacFarlane, 1991), sexually reactive children (Friedrich, 1990), pre-ads (preadolescent sex offenders), child perpetrators (Johnson, 1989; Okami, 1992), and very young perpetrators (Cantwell, 1988).

The proliferation in terminology is most likely a reflection of the lack of resolution by the field as to whether the children should be defined by the supposed etiology of their behavior, that is, reacting to their own abuse or by more legal terms such as perpetrator. It should be noted that only in rare cases does the juvenile justice system adjudicate a child under 12 as delinquent for illegal sexual behavior.

Friedrich (1990) defined sexually reactive children as those who display sexual behaviors seemingly in direct response to their own abuse experiences. Resultant behaviors may "include masturbation, increased sexual exploration, exhibitionism, and a temporary breakdown in the children's interpersonal boundaries" (p. 254). By sexually reactive, Okami (1992) spoke of "child perpetrators of sexual abuse," or the very young juvenile sexual perpetrator.

A more comprehensive definition was suggested by Berliner and Rawlings (1991), who defined sexual behavior problems in children as those "sexual behaviors which are disturbing to others or have negative consequences for the child" (p. 1). Given the young age of these children and the lack of clarity regarding the degree to which sexual abuse is related to later aggressive sexual behavior in children, clinicians and researchers tend to use a more behavioral and less pejorative definition, that is, children with sexual behavior problems or children with aggressive sexual behavior, to describe this population.

Sexual behaviors in children range from those that are developmentally expected and considered normal, to those that are inappropriate and cause concern, to those that are aggressive and highly problematic. For example, masturbation is a common sexual behavior in children. If and when it constitutes a problem is largely determined by family, societal, and cultural norms. At the other extreme, aggressive sexual behavior in children is (i) not responsive to parental limit-setting, (ii) interferes with developmentally appropriate activities, and (iii) may be indicative of serious mental health difficulties (Gordon & Schroeder, 1995). A thorough assessment and understanding of the behavior, its antecedents, and its consequences are necessary to better discriminate healthy, developmentally expected sexual behavior from abnormal sexual behavior that demands therapeutic intervention.

A major issue for the field is how to discriminate between developmentally expected or "normal" sexual play and aggressive sexual behavior. Lamb and Coakley (1993) asked where should the line be drawn between normal childhood sexual play and abuse. Historically, society has viewed sexual contact involving children as "playing doctor" and other such seemingly innocent interactions. As with many relatively new phenomena in the mental health field, there has been both an underresponse and an overresponse to the evaluation of children's sexual behavior, and it has been suggested that professionals need to avoid overreacting to childhood sex play (Okami, 1992).

It is not clear to what extent sexual play among children is normal (Cantwell, 1988). However, two studies with similar findings provide an empirical basis for comparison. From their study of sexual behaviors in 800 children ages two through 12, Friedrich, Grambsch, Broughton, Kuiper, and Beilke (1991) found that a variety of sexual behaviors in children appear to be normal and that their frequency decreases with age. Overt sexual behaviors by children were found to decline with age after peaking at ages three to five. Lindblad, Gustafsson, Larsson, and Lundin (1995) studied 251 preschoolers' sexual behavior at daycare centers in Sweden. As with the Friedrich et al. (1991) study, more serious sexual behaviors were uncommon. Behaviors such as touching an adult's genitals, attempting to make the adult touch the child's genitals/anus, and masturbating compulsively without pleasure or in a way that caused pain were unusual behaviors in very young children.

Several theories have been suggested to explain the etiology of aggressive sexual behavior in children (Ryan, 1997b). Learning theory is the basis of several explanatory models of sexual offending in adolescents and adults (e.g., Freeman-Longo, 1982; Groth, 1979) and has been applied to children. This theory hypothesizes that sexually aggressive behavior develops as a result of learned behavior and interpersonal reactions. This theory would support the idea that aggressive sexual behavior can come about as a direct result of being sexually abused, that is, abuse-reactive children (Cunningham & MacFarlane, 1991). Support for this theory is found in the high rates of sexual abuse in the histories of children with sexual behavior problems (Adams, McClellan, Douglass, McCurry, & Storck, 1995; Bonner, Walker, & Berliner, 1997; Gray, Busconi, Houchens, & Pithers, 1997). However, children with no known history of sexual abuse display aggressive sexual behavior and numerous sexually abused children have no sexual behavior problems. Garland and Dougher (1990) have cautioned against overreliance on learning theory alone: "The belief that sexual abuse causes sexual abuse, the so-called 'abused–abuser hypothesis' is simplistic and misleading" (p. 505).

The trauma model has also been posited to explain sexual behavior problems in children (van der Kolk, 1989). Following a traumatic event (e.g., sexual abuse), a child may compulsively repeat the trauma as it may provide a temporary sense of mastery. The traumatized child may reenact the event, going from a helpless victim to a powerful perpetrator, in an attempt to integrate the experience or discharge affect associated with the trauma (Gil & Johnson, 1993). It is generally accepted that in cases of sexually aggressive children, mastery should be accomplished through other means and the repetition compulsion should be addressed therapeutically.

Aggressive sexual behavior may be the result of other psychopathology in the child and/or the family. Children with serious mental health difficulties, such as disruptive behavior disorders, thought disorders, or affective disorders, can demonstrate problems with sexual aggression toward other children. In addition, factors in the family dynamics, such as inappropriate boundaries or a lack of nurturance, could contribute to the child's sexual behavior problems. At this point, it appears that there are several pathways to inappropriate or aggressive sexual behavior in children and treatment programs now subscribe to an integrated theory with attention to child, family, and contextual factors (Ryan, 1997b).

Several typologies or classification systems for children with sexual behavior problems have been developed to assist clinicians in differentiating levels of problematic behavior and in planning appropriate treatment interventions. A system developed by Berliner and Rawlings (1991) defined three levels of sexual behavior:

(i) Inappropriate sexual behavior, which includes a number of less serious although highly atypical sexual behaviors, such as persistent and/or public masturbation, highly sexualized behavior or play, and preoccupation with sexual matters;

(ii) Precocious sexual behavior, which includes oral–genital contact or intercourse between pre-adolescents with no evidence of coercion or force involved; and

(iii) Coercive sexual behavior, which refers to sexual acts in which coercion or force is threatened or used, or a significant disparity in development or size exists.

A research study on boys and male adolescents aged 6.5–14 in Canada, the USA, and Bermuda found four categories of sexual behavior in these boys (Hall, Mathews, Pearce, Sarlo-McGarvey, & Gavin, 1996):

(i) Developmentally "expected" sexual behavior, in which children engage in sexual behavior which is developmentally appropriate;

(ii) "Sexualized" behavior, referring to children who are "self-focused/sexualized" in their interests or behaviors but who do not engage in problematic sexual behavior with other children;

(iii) Sexually "intrusive" behavior, refers to children who engage in problematic sexual behavior, including touching or contact, but without planning or the use of force; and

(iv) Sexually "offending" behavior in which children are involved in problematic sexual behavior that includes planning and does/does not include force.

Two national research studies funded by the National Center on Child Abuse and Neglect in the Department of Health and Human Services (NCCAN/DHHS) are nearing completion in the USA (Bonner et al., 1997; Gray et al., 1997). These studies were designed to establish empirically-based typologies of children ages six to 12 with sexual behavior problems and to test the effectiveness of two approaches to treatment. One of these studies generated five subtypes of children with sexual behavior problems: abuse reactive, rule breakers, highly traumatized, nondisordered, and sexually aggressive (Pithers, Gray, Busconi, & Houchens, in press). The preliminary findings from the Bonner et al. (1997) study delineated three groups of children with sexual behavior problems:

(i) Inappropriate sexual behavior, such as public masturbation or highly sexualized talking and gestures, but no direct contact with other children;

(ii) Inappropriate sexual behavior, including minimal contact with other children, such as touching other children's genitals; and

(iii) Aggressive sexual behavior, including full contact with other children, such as oral–genital contact and intercourse.

Although the nature of sexual behavior problems in children is not yet clearly defined and understood, consensus on some aspects is being reached. It is now accepted that children do engage in sexual behavior and for some children, the behavior crosses the line of what is normal or developmentally expected for this age. For those children whose behavior is aggressive and coercive, steps need to be taken immediately to provide appropriate intervention for the aggressive child and protection for other children.

9.28.3 PREVALENCE

The incidence and prevalence of sexual behavior problems in children is unknown and will be difficult to establish in the future. This is due to various reasons. First, sexual behavior problems or when the behavior becomes aggressive is not well-defined and professionals do not agree on what constitutes sexually aggressive behavior in children. Ryan (1997a) pointed out that the evaluation of inappropriate sexual behavior between children becomes more difficult as the ages of the participants become younger and closer. Ryan indicated that in evaluating a sexual interaction to determine whether that contact is abusive, one must assess three factors: equality, consent, and coercion. The following should be evaluated for equality: Differences in physical, cognitive, and emotional development; passivity and assertiveness; power and control; and authority. The ability to give consent implies full knowledge, understanding, and choice. Coercion describes the overt or covert pressures that prevent a child from making a choice about participation. It is obvious that these factors are difficult to determine in cases involving young children. Second, the child's abusive behavior, like other forms of sexual offenses, is often unacknowledged and not reported to authorities even when it is clearly aggressive. It has been noted that the identification and reporting of child offending was almost nonexistent before 1985 (National Adolescent Perpetrator Network, 1993). Throughout the USA different criteria are used to determine when children are sexually aggressive and there is no systematic approach to collecting this type of data across states by child protective services or by law enforcement authorities.

There are, however, indicators that the number of children with aggressive sexual behavior is not insignificant. For example, children under age 12 were found to be responsible for about 13% of all child sexual abuse cases substantiated in the state of Vermont in 1991 (Gray & Pithers, 1993). In addition, surveys conducted by the National Adolescent Perpetrator Network and the Safer Society Foundation in the early and mid 1990s showed a substantial increase in the number of professionals and programs providing services to children with sexual behavior problems, from 25 in 1991 to almost 400 in 1995 (Freeman-Longo, Bird, Stevenson, & Fiske, 1995; Interchange, 1991). While it is clear that aggressive sexual behavior by children is increasingly recognized and treatment provided, it is also clear that uniform definitions and data collection on the national level are necessary to establish reliable estimates of incidence and prevalence.

9.28.4 ASSESSMENT

In the field of child sexual abuse, two distinct areas of assessment or evaluation are emerging, forensic and clinical. The goals, purposes, and uses of these two forms of evaluation are distinctly different and it has been recommended that they be conducted separately by different professionals. In cases where inappropriate or aggressive sexual behavior has occurred, a forensic evaluation is conducted to determine whether or not a child committed the behavior or whether or not a particular child has been sexually abused. These interviews are typically conducted by Child Protective Services (CPS) personnel, law enforcement officers, or a member of a multidisciplinary team at a children's advocacy center. In some cases, mental health professionals are involved in conducting forensic evaluations. It is imperative that the clinician is clear as to which type of assessment is being conducted and that this is conveyed to the parent, child, and other professionals involved with the child prior to beginning the evaluation. A clinical evaluation is designed to obtain information for treatment planning and is not used to determine whether a child committed a particular act.

A comprehensive clinical assessment of the child and parents or caretakers is necessary in order to develop an effective treatment plan. The assessment should be conducted by a professional who is knowledgeable about sexual abuse and children with sexual behavior problems and who is trained in the administration and interpretation of psychological tests. Prior to the assessment, the following information should be obtained and reviewed: reports of previous testing or treatment, school records, and any information from child protective services, law enforcement, or juvenile court regarding the child or the reported current incidents of sexual behavior. A comprehensive assessment includes clinical interviews and the use of psychological instruments to assess both the child's and parent/caretaker's current affective and behavioral functioning.

Before beginning the interviews with the parent and child, the limits of confidentiality should be explained clearly. For some clients, it may be helpful to have a written copy of the limits of confidentiality for later reference. If child and parent completely deny that the child has had a sexual behavior problem, the clinician would need to decide whether to terminate the session and discuss the issue with the referral agency (school, law enforcement, CPS), complete the evaluation and discuss the findings and recommendations with the family, or some other course of action.

The clinician can then proceed to gather information through separate clinical interviews with the parent and child. This would include standard information obtained from clinical interviews for other problems in childhood, including a developmental and medical history, school performance, history of psychological or psychiatric problems, alcohol or drug use by family members, the child's social relationships with peers and siblings, and family functioning and previous changes or moves by the child or family. This information is relatively standard in mental health settings and should not provoke resistance by the child or parent. If the child is in foster care, this information may need to be obtained from the child's caseworker.

In addition, the assessment should include an in-depth history of the child's sexual behavior, keeping in mind that (i) a great deal of sexual behavior between children is seen as normal and not highly problematic, that is, playing doctor with or showing their genitals to same-age peers or siblings; (ii) children and their parents may be quite reluctant to discuss the child's sexual behavior, and (iii) it is unlikely that all of the child's sexual behavior will be disclosed during the initial interview. If a court has ordered that reports be sent to the CPS or the school, for example, this must be clearly explained to the parents and child. If the parent agrees that the reports can be sent, we recommend that the parent sign a release of information to document their consent. In many cases with children under 12, the court is not involved and the typical limits of confidentiality for children can be observed.

To begin the initial interview, it is suggested that the therapist say: "Hello Jason, Ms. Jackson. I understand that you are here today because you (your son) has had a problem with his sexual behavior." This takes a more direct approach, labels the problem behavior, and appears to take some of the pressure off the clients. In our experience, children are quite reluctant to discuss their inappropriate or aggressive behavior and this can be interpreted as denial. This can also be reflected in their parents' responses to inquiries about their children's behavior. We view it more as being reluctant, embarrassed, afraid, or ashamed of the behavior and the current consequences. The therapist could then explain that the session will entail talking to the child and the parent and the completion of some paper and pencil tests to help plan the treatment program.

The assessment should include questions on the child's normal or expected sexual behavior and problematic behavior, including: (i) masturbation; (ii) playing doctor; (iii) looking at others' private parts; (iv) the types, duration, and frequency of the child's inappropriate or aggressive sexual behavior; (v) any use of threats, coercion, tricks, or bribes to involve other children sexually; and (vi) the child's response to discipline or correction regarding their sexual behavior, for example, did he/she stop or become more secretive; and (vii) the overall attitude toward sexuality in the family.

After the clinical interviews are completed, it is recommended that a thorough psychological assessment be conducted. This will provide important information on the child's diagnosis and general intelligence, overall psychopathology, behavior and affective problems, and sexual behavior. The following instruments can be used to assess the child's and parents' psychological state, level of stress, and family functioning.

9.28.4.1 Child Evaluation

In this section we provide a brief overview of the instruments that can be used in evaluating children who engage in aggressive sexual behavior.

9.28.4.1.1 Diagnosis and overall psychopathology

The Child Assessment Schedule (CAS; Hodges, 1993) is a 226-item structured diagnostic interview developed for the assessment of school-age children in clinical or research settings. The responses are summed across the 226 items for a general pathology score and subscales generate assessments of depression, anxiety and fear, self-image, conduct disorders, and somatic complaints. The CAS should be conducted with a child by a trained interviewer.

9.28.4.1.2 Behavior and affective problems

The Child Behavior Checklist (CBCL; Achenbach, 1991) is a 134-item standardized checklist of children's behavior problems and social competence in children ages four to 18. It has forms that can be completed by the child, parent, and/or teacher about the child. The CBCL measures factors such as depression, somatic complaints, hyperactivity, aggressiveness, sexual behavior, anxiety, and delinquent behavior. In addition, the CBCL yields overall scores on a child's externalizing (aggressive and antisocial behavior) and internalizing (withdrawal, fearfulness, inhibition, and depression) behaviors. The test reflects whether or not a child's level of behavior falls into the clinical range for children of the same age and gender.

The Revised Children's Manifest Anxiety Scale (RCMAS; Reynolds & Richmond, 1985) is a 37-item inventory that assesses a variety of anxiety symptoms. The instrument is completed as a self-report by the child. Reliability and validity studies indicate that the instrument acts as a satisfactory measure of chronic anxiety.

The State–Trait Anxiety Inventory for Children (STAIC; Speilberger, 1973) is a 40-item instrument developed as a research tool to evaluate anxiety in elementary school children. The items are divided into two, 20-item, self-report scales. The Anxiety–State subscale reflects transitory anxiety that tends to fluctuate over time. The Anxiety–Trait subscale reflects relatively stable differences in children's levels of anxiety.

9.28.4.1.3 Self-esteem

The Harter Perceived Self-Competence Scale (Harter, 1982) is a 28-item questionnaire in which a child self-reports his/her perceived competence in social, cognitive, physical, and global esteem. The child is asked to select which of two described children he/she is most like and the degree to which he/she is similar.

9.28.4.1.4 Sexual behavior problems

The Child Sexual Behavior Inventory (CSBI; Friedrich et al., 1992) is a 36-item instrument completed by a parent or caretaker to rate sexual behaviors shown by children ages two to 12 during the prior six months. The behaviors are scored 0 to 4 to indicate the frequency of the problematic behaviors. Clinicians should note that this version of the CSBI does not contain items that assess highly aggressive sexual behavior in children ages two to 12. These should be covered in the clinical interview or through other paper and pencil items or questions designed by the clinician.

There is no standardized measure to assess the levels of appropriateness and/or aggressiveness in children's sexual behavior. Two Likert-type scales have been designed and utilized to measure these aspects of children's sexual behavior (Bonner et al., 1997). The 7-point scales are from 1 (very appropriate) to 7 (very inappropriate) and 1 (not at all aggressive) to 7 (very aggressive). The raters, which can be the parents, the referral source, and/or the clinician, take into account the child's age, gender, and the total referral sexual behaviors in making the ratings.

9.28.4.2 Parent/Caretaker Evaluation

9.28.4.2.1 Psychological status

The Brief Symptom Inventory (BSI; Derogatis, 1991) is a 53-item self-report version of the Symptom Checklist-90 (Derogatis, 1977). The symptom items are rated on a five-point scale to indicate the degree of distress within the last week. Scores are obtained for nine primary symptom dimensions and three global indices of distress.

9.28.4.2.2 Level of stress

The Parenting Stress Index (PSI; Abidin, 1990) is a 120-item parent self-report instrument designed to measure the degree of stress in the parent-child system and to identify the sources of distress. Three major sources of stress, child characteristics, mother characteristics, and situational-demographic life stress, are assessed by the instrument.

9.28.4.3 Family Evaluation

The Family Environment Scale (FES; Moos & Moos, 1981) is a 90-item, true/false instrument that measures the social-environmental attributes of various kinds of families. The test assesses three dimensions of family functioning: relationship, with subscales for cohesion, expressiveness, and conflict; personal growth, with subscales for independence, achievement orientation, intellectual–cultural orientation, and moral–religious emphasis; and system maintenance, with subscales for organization and control.

The Child Version of the Family Environment Scale (CVFES; Pino, Simmons, & Slowski, 1984) is a 30-item instrument presented in pictorial form to the child. Children ages five through 12 are shown cartoon images depicting family scenes, with only one dimension changing across multiple choice response options. The child selects the cartoons that best represent how he/she feels about his/her family. The CVFES provides information similar to the adult version described above.

Information from the child's and parents/caretakers' evaluation can be used to develop a comprehensive treatment plan which may include individual, group, and family therapy. The child's sexual behavior problems will be addressed in individual and/or group therapy, along with parent/caretaker support and education. If the child has other significant affective or behavioral problems, these will need to be addressed in individual therapy. The parent/caretaker may also need to be seen or referred for additional treatment for other psychological issues. Family sessions may need to be scheduled to address boundary issues, safety within the home, and other problems that emerge during the evaluation.

9.28.5 TREATMENT

The literature contains descriptions of several treatment approaches for children with inappropriate or aggressive sexual behavior (e.g., Ballester & Pierce, 1995; Becker & Bonner, 1998; Cunningham & MacFarlane, 1991, 1996; Friedrich, 1990; Gil & Johnson, 1993; Gray & Pithers, 1995; Johnson, 1991). In general, the treatment approaches have several common objectives for treatment, such as reducing the child's inappropriate sexual behavior, improving the child's self-esteem, and addressing family problems. Depending on the length of treatment, that is, time-limited or open-ended, there are other issues that may be addressed, such as the child's history of abuse and other problematic behaviors.

In a book of therapeutic exercises for child abuse victims and children with sexual behavior problems, Johnson (1995) listed 10 goals for treatment. These include:

(i) decrease children's problematic sexual behavior;

(ii) understand and integrate issues related to prior victimization, including maltreatment, family separations, and deaths;

(iii) increase children's understanding of their unhealthy beliefs about sex and sexuality;

(iv) increase children's understanding of healthy sexuality;

(v) increase children's understanding of personal and family patterns related to sexual abuse and other problematic behavior;

(vi) assist children to understand their needs and develop their goals and internal resources to reach the goals;

(vii) help children assess and understand the precipitants to their own behavior and to think of the consequences before they act;

(viii) increase children's ability to meet their social needs in appropriate ways;

(ix) increase children's empathy for others; and

(x) increase children's positive relationships and attachment to others and build internal objects that enhance future development.

These goals are quite comprehensive and treatment typically will utilize individual, group, and family therapy as modalities.

While treatment programs may differ in length and degree of comprehensiveness, typically they cover a set of general topics related to the child's sexual behavior. In general, these programs have incorporated, but significantly modified, aspects of treatment techniques or issues addressed in the treatment of adolescent or adult sex offenders. These include sex education and sexuality, social skills, impulse control, and problem-solving skills. Specific focus is placed on topics related to the inappropriate or aggressive sexual behavior, such as acknowledging and discussing their sexual behavior, taking responsibility for their sexual behavior, identifying risk factors, and developing a plan to prevent the abuse from happening again. The literature supports the necessity of focusing specifically on the inappropriate or aggressive sexual behavior as it indicates that children's sexual behavior problems may be more resistant to change and respond better to a targeted treatment intervention (Finkelhor & Berliner, 1995).

Several treatment programs were established in the mid-to-late 1980s to provide treatment for abuse reactive children, for example, children under age 13, most of whom had been sexually abused and who were referred for inappropriate sexual behavior with other children. One of these early treatment programs has been described in the literature. The Support Program for Abuse Reactive Kids (SPARK) was established at Children's Institute International in Los Angeles in 1985 (Ballester & Pierce, 1995). It utilizes the concept of monster therapy to treat boys ages four to 13 who have been sexually aggressive toward their siblings, peers, and other younger children. Group therapy for the children and their parents or caretakers is the primary treatment modality, although other treatment interventions are also utilized.

Monster therapy is described as a systematic, play therapy approach that is used to increase self-esteem, impulse control, anger management, and a sense of mastery in the children. The children's impulsive sexually inappropriate behavior is personified as monster behavior. The therapists report that with this approach the children see their problem, or their monster, as something outside of themselves that can be brought under control and conquered. The therapy sessions and group activities are built on the monster theme and therapy is depicted as a battle to conquer the monster. The therapists note that it is the children's impulsive behavior that is labeled, not the children themselves.

SPARK has six stages of therapy although the stages do not necessarily proceed in a linear fashion.

(i) *Creating a Therapeutic Environment*. Addresses shame, guilt, assertiveness, secrecy, control, safety, and introduces the monster metaphor.

(ii) *Enhancing Self-Awareness*. Addresses identification of feelings, disclosure, self-blame, homophobia, power and control, denial, cognitive distortions, responsibility, projection, and risk factors (time, circumstances, behavior, and feelings).

(iii) *Learning Alternative Behaviors and Skills*. Addresses impulse control through cognitive restructuring, self-talk techniques, and high-risk strategies; social skills such as appropriate affection and socialization, managing anger and anxiety, and communication skills; empathy; and sex education.

(iv) *Identifying and Using Resources*. Addresses identifying resources and support systems, assertiveness skills, and problem solving.

(v) *Putting Skills into Practice*. Addresses empowerment, locus of control, and internalizing skills.

(vi) *Skill Mastery and Relapse Prevention*. Addresses application of skills and concepts to other situations, foreseeing potential problems, and techniques to prevent relapse.

Various group activities are used to enhance the children's learning of the concepts and skills. The children earn small plastic monsters as tokens each week that they put into a cup with their name on it. Other groups make a monster cage from small wood sticks to hold their monster tokens. The monsters are earned each week for learning new skills or rules and are exchanged every three to four weeks for treats from a goody bag. This technique has been found useful in increasing the boys' self-esteem, providing a visual means of self-assessment of progress in therapy, and giving the parents a way to gauge their child's progress.

The group treatment program runs for 14 weeks followed by a two- to four-week break. Children can graduate at this time or continue to participate in the next cycle with new participants and their parents. Families are typically in treatment from 12 to 18 months. Termination is planned when the child has successfully completed the stages of treatment and met all of the treatment goals.

A parallel group treatment program is conducted for the parents or caretakers of the children attending SPARK (Griggs & Boldi, 1995). Parental participation is mandatory for a child to be accepted into SPARK as the home environment is seen as a crucial factor in reducing the child's sexual behavior problems. The group is directed by a male–female cotherapy team that provides education, support, and therapeutic intervention. The group focuses on assisting the parents to resolve their own victimization issues, strengthening the relationship between the parent and child, and providing education on child sexual abuse and human sexuality. The two main goals of parental treatment are to (i) strengthen the family through increasing healthy communication and interactions, and (ii) decrease the pathology by encouraging healthy family dynamics.

The following issues are addressed in the parents' program:

(i) assisting parents to negotiate the various systems involved in these cases, that is, legal, law enforcement, child protective services, school, and child care systems;

(ii) providing information on general child development, child sexual abuse issues, children's sexual development, and distinguishing between normal and sexually reactive behavior;

(iii) assisting parents to provide supervision of the reactive child and the victim;

(iv) teaching skills in parenting and behavior management;

(v) helping parents understand how family dynamics contribute to reactive behaviors;

(vi) assisting parents to normalize their responses and to strengthen the parents' support and resources; and

(vii) helping parents to develop appropriate expectations and treatment goals.

At present, two national treatment outcome studies to verify the effectiveness of different treatment approaches with this population of children are nearing completion (Bonner et al., 1997; Gray et al., 1997). These two studies are particularly important in that they were implemented early in the development of treatment interventions for these children. The results from these studies can provide direction for future clinical programs providing services to this population. Each of the research studies utilizes standardized assessment instruments and is gathering two-year follow-up data on the children.

9.28.5.1 Children with Sexual Behavior Problems Program

The Oklahoma Program, conducted through the Center on Child Abuse and Neglect in the Department of Pediatrics at the University of Oklahoma Health Sciences Center, compared two approaches to time-limited group therapy (Bonner et al., 1997). The groups are for children ages six to 12 and separate educational groups are conducted for the parents or caretakers. The children and the caretakers are randomly assigned to one of two treatment approaches, cognitive-behavioral group therapy or dynamic play therapy. The treatment program is time-limited and consists of 12 one-hour group sessions for both the children and the parents. The primary goal of both approaches is to reduce the children's inappropriate or aggressive sexual behavior. However, the techniques used to reduce the behavior are markedly different in the two treatment approaches.

9.28.5.1.1 The cognitive-behavioral group therapy approach

The cognitive-behavioral therapy (CBT) approach presumes that the inappropriate or aggressive sexual behavior has been learned and can, therefore, be unlearned with new information, practice, and rewards for more appropriate behavior. It is a highly-structured approach in which the two therapists, one male and one female, are active and directive. Girls and boys from age six to 12 participate together in the group. In the CBT approach, the children's inappropriate behavior is addressed directly in the first session. After the children and therapists introduce themselves to each other and some initial rapport is established, the children are told that they are in the group because they broke a sexual behavior rule. They are told that the group will help them learn the sexual behavior rules and how not to break the rules in the future.

As the treatment is time-limited, the program focuses on the most relevant topics related to the children's inappropriate sexual behavior, that is, learning and implementing the sexual behavior rules (see Table 1), learning basic sex education, improving impulse control, and preventing future abuse. These topics are introduced, repeated frequently over the sessions, and positively reinforced. In order to assess the cognitive or learning aspects of the program, a five-item assessment is administered at the end of each session to determine if the children are learning the didactic content. The child's behavior is assessed weekly by child and parent report.

The group sessions for parents also utilize a cognitive-behavioral approach in which the therapists take an active, directive role. The

Table 1 Sexual behavior rules.

1. It's OK to touch your private parts when you are alone
2. It's not OK to touch other people's private parts
3. It's not OK for other people to touch your private parts
4. It's not OK to show your private parts to other people

topics from the children's group are explained and discussed in the parents' group the week before they are introduced to the children. The parents' group covers the sexual behavior rules, the turtle technique (used to decrease impulsive behavior), sex education materials, safety rules in the home, and prevention of reabuse. Other topics are covered in the parents' group, such as developmentally expected sexual behavior in children, supervision of children with sexual behavior problems, family rules about sexual behavior, talking about sexuality with children, and general behavior management (see Table 2 for an outline of sessions).

9.28.5.1.2 Case study: cognitive-behavioral treatment approach

VS was a 10-year-old boy referred by a community psychologist for aggressive sexual behavior. The referral behaviors included attempted penetration of his four-year-old sister, mutual masturbation with his sister, performing oral sex on his 12-year-old male cousin, and exposing himself and grabbing the genitals of two nine-year-old friends who were spending the night.

VS had recently reported that his father had anal and oral intercourse with him when he was three and again when he was six. During a previous evaluation in which his father, mother, he, and his sister participated, VS recanted the accusations. His parents had been separated for over a year and a divorce was pending. The CPS later conducted an investigation and confirmed VS's sexual abuse by his father as valid.

VS was cooperative during the assessment. He reported that his grades had gone from "straight As" to "straight Fs" because of his parents' divorce. He stated that he was afraid of being away from his mother. VS said that he wanted to commit suicide last year because of all the fighting due to the divorce but that he did not have thoughts of suicide at the current time.

VS's test scores indicated that he was functioning in the normal range of intelligence (IQ = 91). His scores were in the clinical range for sexual behavior problems on both the CBCL and the CSBI. On a scale of 1 (very appropriate) to 7 (very inappropriate), his sexual behavior had an average rating of a 6.4. The level of aggressiveness was 6.0 on a scale of 1 (not at all aggressive) to 7 (very aggressive). On the CBCL, VS was seen to have significant difficulties in internalizing problems, including being withdrawn, having somatic complaints, showing thought problems, and having difficulty attending. His mother rated VS in the clinical range on internalizing and total scores on the CBCL. Ratings on the PSI indicated that his mother had difficulty accepting VS and that she was experiencing significant life stress.

VS and his mother attended all 12 treatment sessions. VS was compliant, agreeable, and participated well throughout the program. He learned the sexual behavior rules and the self-control procedures (turtle technique). Over the course of treatment, VS acknowledged his sexual behaviors that had been reported. He was seen as having successfully completed the treatment program. The mother was advised to monitor his behavior closely and contact the therapists if further sexual behavior problems arose.

At the end of the 12 treatment sessions, and at a one-year follow-up, VS's mother reported that he had adjusted to the divorce, his visitation with his father was supervised, and that no further sexual behavior problems had been reported to her.

9.28.5.1.3 Dynamic play group therapy approach

The second treatment approach utilized by the Oklahoma Program was chosen to contrast with the cognitive-behavioral approach. The dynamic play therapy group was less structured and utilized group play activities as the intervention. The children's inappropriate or aggressive behavior was not directly addressed by the therapists unless it emerged during the group sessions. The group was designed to provide a safe, positive atmosphere in which the children could express their thoughts and feelings, develop a better understanding of themselves and others, and interact appropriately with other children. The therapists were nondirective and provided reflection and interpretation to the children. It was a more permissive, child-centered intervention than the cognitive-behavioral approach.

Various play and expressive activities were utilized by the therapists to encourage the expression of feelings and group interaction. These included the use of clay, puppets, group story telling, costumes, group board games, and other art materials and activities.

The parents' group was conducted from a dynamic approach, which utilized the members of the group and the group process as the modality of change. The therapists were again nondirective and used reflection and interpretation as primary interventions. The group processed problems and issues that the parents or caretakers raised each week. The therapists did not set the agenda but followed the direction set by the parents.

Table 2 Cognitive-behavioral therapy session outline.

Children's group		*Parents' group*	
Session 1			
Introduce:	Group purpose	Introduce:	Group purpose
	Group rules		Group rules
	Describe activities		Group format
Assessment:	(5 items)		Sexual behavior rules
Session 2			
Review:	Purpose of group, rules	Review:	Sexual behavior rules
Introduce:	Sexual behavior rules	Introduce:	Parental expectations
	Private parts		Discussing feelings
	Applying sexual behavior rules		
Assessment:	(5 items)		
Session 3			
Review:	Group purpose, rules	Introduce:	Rules in your home
	Private parts		Supervision guidelines
	Sexual behavior rules		
Introduce:	Discussing feelings		
Assessment:	(5 items)		
Session 4			
Review:	Group purpose	Introduce:	Turtle technique
	Sexual behavior rules		Expected sexual behavior in children
Introduce:	Disclosure		
Assessment:	(5 items)		
Session 5			
Review:	Sexual behavior rules	Review:	Turtle technique
	Disclosure		Sexual behavior in children
Introduce:	Turtle technique		
Assessment:	(5 items)		
Session 6			
Review:	Sexual behavior rules	Introduce:	Behavior management techniques
	Turtle technique		
	Disclosure		
Assessment:	(5 items)		
Session 7			
Review:	Turtle technique	Introduce:	Managing sexual misbehavior
Assessment:	(5 items)		Sex education
Session 8			
Review:	Private parts	Review:	Managing sexual misbehavior
Introduce:	Sex education	Introduce:	Abuse prevention
	Talking about sex		
Assessment:	(5 items)		
Session 9			
Review:	Turtle technique	Introduce:	Talking with your child about sex
	Sex education		
Assessment:	(5 items)		
Session 10			
Review:	Turtle technique	Review:	Materials from previous sessions
	Sex education	Introduce:	Open discussion
Introduce:	Abuse prevention		
Assessment:	(5 items)		
Session 11			
Review:	Sexual behavior rules Turtle technique Abuse prevention	Individual meetings between therapists and parents to discuss information on child's progress and recommendations	
Assessment:	(5 items)		
Session 12			
Saying goodbye		Postassessment session for parents to complete instruments	

9.28.5.1.4 Case study: dynamic play therapy treatment approach

NH was an eight-year-old girl who participated in the dynamic play therapy group. She was referred by a psychologist who was seeing NH for a brain injury she had sustained during infancy. NH was referred after she was observed licking the genitals of a six-year-old girl in a bathroom stall at school; it was also reported that another adult had seen NH with her mouth on her four-year-old cousin's penis.

NH was the only survivor of a murder/suicide in which her father had shot her mother and an older brother, and then killed himself. NH had been diagnosed with learning disabilities and was on medication for seizures. She had been living with a maternal aunt and uncle since infancy and had no reported history of abuse. Mrs. H. reported that she thought an adult had possibly allowed NH to see a pornographic video but she did not know who it could be, and that NH would not discuss the issue. Mrs. H's view was that NH was acting out of curiosity but she acknowledged the need for treatment regarding the sexual behavior. She reported that NH had not asked many questions about sex. Mr. and Mrs. H. had often told NH never to let anyone touch her private parts but they were unaware of the extent of NH's sexual knowledge.

NH was attentive and cooperative during the testing session. When the items became more difficult, she readily said, "That's too hard for me," or "I can't do that one." She reported that she sometimes had difficulty with schoolwork but that she made good grades in the learning lab. Her intelligence scores indicated that she was functioning in the Below Average range (IQ = 80). On the CBCL (completed by her aunt), NH was found to be in the clinical range on social problems, delinquent behavior, sex problems, aggressive behavior, and attention problems. Her Externalizing and Total scores were in the clinical range. The CSBI score was not in the clinical range. The ratings of her sexual behaviors by four national experts were 6.0 (7 = very inappropriate) on level of appropriateness and 4.0 (7 = very aggressive) on level of aggressiveness.

The aunt's responses on the PSI indicated significant stress in her relationship with NH, her parenting, and her relationship with her spouse. All six categories in the parent scale and the total stress score were in the high range. On the BSI, all of her scores were within normal limits.

In the dynamic play group, NH was initially quiet but pleasant and friendly to the other group members. She participated in the group activities and appeared to prefer the company of the other children to that of the therapists. When the therapists paid attention to her, she tended to giggle nervously and move away. Over time, NH sought attention from other children by encouraging disruptive behavior and became more oppositional with the therapists. At times, she flatly refused to comply until pressured by the group members.

NH and her aunt and uncle attended 11 of the 12 scheduled sessions. The adults participated actively in the parents' group and openly discussed their concerns about NH. They understood the necessity of supervising NH closely to prevent any further inappropriate sexual behavior and the need for her to interact with other children as much as possible in a supervised setting. Over the course of the group, they both discussed how to communicate with NH about appropriate and inappropriate sexual behavior. No additional treatment was recommended.

At the end of the treatment program, no additional inappropriate sexual behavior was reported and no further treatment was recommended. Similarly, the one- and two-year follow-up reports indicated that NH's overall behavior had improved and that no inappropriate or aggressive sexual behavior had occurred.

9.28.5.1.5 Treatment

The research design for this study included random assignment to either cognitive-behavioral or dynamic play therapy, the use of standardized instruments for assessment, and follow-up at the end of treatment and at one and two years post-treatment. The initial results show that both treatment approaches were effective in reducing the children's inappropriate sexual behavior. The incidence of repeated inappropriate or aggressive behavior was approximately 15% for children who completed treatment in both treatment approaches.

9.28.5.2 STEP Program

The second research project funded by the National Center on Child Abuse and Neglect was conducted in Vermont by Pithers et al. (in press). This study examined the effectiveness of two treatment models, both of which utilized 32 weeks of 90 minute parallel treatment groups for children and their caregivers. One approach was Sex Abuse Specific Treatment (SAST), a modified relapse prevention approach that emphasized the involvement of caregivers as external monitors of their child's behaviors. The second approach, Abuse Prevention Treatment (APT), was an expressive approach that focused on improving self esteem, reducing anxiety, and increasing emotional regulation.

The effects of each treatment were examined across five types of children with sexual behavior problems that had been identified through cluster analytic procedures. The data indicated that SAST provided more immediate benefits for highly traumatized children, suggesting that the modified relapse treatment approach helped children and caregivers manage sequelae associated with the child's history of abuse. However, over time, the differences between treatment approaches eroded, a finding thought to be attributable to the later onset of benefits from the less structured expressive treatment approach.

In summary, treatment programs for children with sexual behavior problems have increased dramatically since the early 1990s. In general, the programs have similar approaches and utilize adapted concepts from treatment programs for adolescents and adults. The treatment outcome research currently being conducted will provide useful information on effective approaches to treating this population of children.

9.28.6 FUTURE DIRECTIONS

The results of initial treatment outcome research with children with sexual behavior problems is promising. It appears that several treatment interventions are useful and that the rates of repeated aggressive sexual behavior are between 10–15%. While we lack an empirically derived theory to explain the etiology of this behavior in children, it appears, both from existing literature and clinical practice, that multiple factors rather than a single cause are involved.

As this area of clinical practice is in its infancy, there are numerous factors that need to be addressed. There is a clear need for the development of a set of agreed upon definitions or criteria for what constitutes aggressive sexual behavior in children and when a system outside the family, that is, CPS, law enforcement, or the courts, needs to become involved. Future research could be instrumental in clarifying a set of definitions and establishing a data base on incidence and prevalence.

There is also a need for the development of standardized instruments to measure the level and severity of aggressive sexual behavior in children. The well-normed CSBI developed by Friedrich et al. (1991) has added a set of questions assessing the presence of more aggressive sexual acts by children. This type of measure could be utilized as a pre-post measure for future treatment outcome studies.

It is also recommended that communities provide a range of services for children with inappropriate or aggressive sexual behavior, including outpatient treatment for children and their families, therapeutic foster care, and when necessary, specialized group homes and residential treatment centers.

The present studies funded by the National Center on Child Abuse and Neglect (Bonner et al., 1997; Gray et al., 1997) will add valuable information to the knowledge base about children with aggressive sexual behavior and the effectiveness of the treatment approaches utilized in the studies, that is, cognitive-behavioral, dynamic play therapy, modified relapse prevention, and expressive therapy, in group settings for children and parents. Additional research is needed on other approaches and modalities of treatment, such as individual and family therapy. It is imperative that future research use sufficient numbers of subjects, standardized instruments, and rigorous methodology in order to reliably establish the findings. The importance of well-designed longitudinal studies of children with sexual behavior problems cannot be understated.

9.28.7 SUMMARY

The extant literature and clinical reports indicate that problematic sexual behavior, including aggressive sexual assaults by children under age 12, is increasing. As this behavior has been reported to occur in schools and other public facilities, numerous professionals are concerned for the safety of children under their supervision. To date, several group approaches have been found effective in reducing aggressive sexual behavior in children, including cognitive-behavioral, dynamic play therapy, modified relapse prevention, and expressive therapy. Future research should focus on the development of treatment models that are both time- and cost-efficient and can be implemented easily in community mental health centers.

9.28.8 REFERENCES

Abidin, R. R. (1990). *Parenting Stress Index.* Charlottesville, VA: Pediatric Psychology Press.

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist 4–18 and 1991 Profile.* Burlington, VT: University of Vermont, Department of Psychiatry.

Adams, J., McClellan, J., Douglass, D., McCurry, C., & Storck, M. (1995). Sexually inappropriate behaviors in seriously mentally ill children and adolescents. *Child Abuse and Neglect, 19,* 555–568.

Ballester, S., & Pierce, F. P. (1995). Monster therapy: The use of a metaphor in psychotherapy with abuse reactive children. In M. Hunter (Ed.), *Child survivors and perpetrators of sexual abuse* (pp. 125–146). Thousand Oaks, CA: Sage.

Becker, J. V., & Bonner, B. L. (1998). Sexual and other abuse of children. In R. J. Morris & T. R. Kratochwill (Eds.), *The practice of child therapy* (3rd ed., pp. 367–389). Boston: Allyn & Bacon.

Berliner, L., & Rawlings, L. (1991). *A treatment manual: Children with sexual behavior problems*. Seattle, WA: Harborview Sexual Assault Center.

Bonner, B. L., Walker, C. E., & Berliner, L. (1997, March). *Group treatment approaches for children with sexual behavior problems*. Presented at the Spurwink Institute, Newport, RI.

Cantwell, H. B. (1988). Child sexual abuse: Very young perpetrators. *Child Abuse and Neglect, 12,* 579–582.

Cunningham, C., & MacFarlane, K. (1991). *When children molest children: Group treatment strategies for young sexual abusers*. Orwell, VT: Safer Society Press.

Cunningham, C., & MacFarlane, K. (1996). *When children abuse: Group treatment strategies for children with impulse control problems*. Brandon, VT: Safer Society Press.

Derogatis, L. R. (1991). *Brief Symptom Inventory*. Baltimore: Clinical Psychometric Research.

Derogatis, L. R. (1977). *SCL-90: administration, scoring, and procedures manual for the revised version*. Baltimore: Clinical Psychometric Research.

Federal Bureau of Investigation (1990). *Uniform crime reports of the United States*. Washington, DC: US Government Printing Office.

Finkelhor, D., & Berliner, L. (1995). Research on the treatment of sexually abused children: A review and recommendations. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 1408–1423.

Finkelhor, D., Hotaling, G. T., & Lewis, F. T. (1990). Sexual abuse in a national survey of adult men and women: Prevalence, characteristics, and risk factors. *Child Abuse and Neglect, 14,* 19–28.

Freeman-Longo, R. E. (1982). Sexual learning and experience among adolescent sexual offenders. *International Journal of Offender Therapy and Comparative Criminology, 26,* 235–241.

Freeman-Longo, R. E., Bird, S., Stevenson, W. F., & Fiske, J. (1995). *1994 nationwide survey of treatment programs and models: Serving abuse reactive children, adolescent, and adult sex offenders*. Brandon, VT: Safer Society Press.

Friedrich, W. N. (1990). *Psychotherapy of sexually abused children and their families*. New York: Norton.

Friedrich, W. N., Grambsch, P., Broughton, D., Kuiper, J., & Beilke, R. L. (1991). Normative sexual behavior in children. *Pediatrics, 88,* 456–464.

Friedrich, W. N., Grambsch, P., Damon, L., Koverola, C., Hewitt, S. K., Lang, R. A., & Broughton, D. (1992). Child Sexual Behavior Inventory: Normative and clinical comparisons. *Psychological Assessment, 4,* 303–311.

Garland, R. J., & Dougher, M. J. (1990). The abused/abuser hypothesis of child sexual abuse: A critical review of theory and research. In J. Feierman (Ed.), *Pedophilia: Biosocial dimensions* (pp. 488–509). New York: Springer.

Gil, E., & Johnson, T. C. (1993). *Sexualized children: Assessment and treatment of sexualized children and children who molest*. Rockville, MD: Launch Press.

Gordon, B. N., & Schroeder, C. S. (1995). *Sexuality: A developmental approach to problems*. New York: Plenum.

Gray, A., Busconi, A., Houchens, P., & Pithers, W. (1997). Children with sexual behavior problems and their caregivers: Demographics, functioning, and clinical patterns. *Sexual Abuse: A Journal of Research and Treatment, 9,* 267–290.

Gray, A., & Pithers, W. (1993). Relapse prevention with sexually aggressive adolescents and children: Expanding treatment and supervision. In H. E. Barbaree, S. Hudson, & W. Marshall (Eds.), *The juvenile sexual offender* (pp. 289–319). New York: Guilford Press.

Griggs, D. R., & Bolbi, A. (1995). Parallel treatment of parents of abuse reactive children. In M. Hunter (Ed.), *Child survivors and perpetrators of sexual abuse* (pp. 147–165). Thousand Oaks, CA: Sage.

Groth, A. N. (1979). Sexual trauma in the life histories of rapists and child molesters. *Victimology, 4,* 10–16.

Hall, D., Mathews, F., Pearce, J., Sarlo-McGarvey, N., & Gavin, D. (1996). *The development of sexual behavior problems in children and youth*. Toronto, Canada: Central Toronto Youth Services.

Harter, S. (1982). The perceived competence scale for children. *Child Development, 53,* 87–97.

Hodges, K. (1993). Structured interviews for assessing children. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 34,* 49–68.

Interchange (1992, Fall). *Results of a survey regarding the treatment of preadolescent sexually abusive behavior*. Denver, CO: Kempe National Center.

Johnson, T. C. (1989). Female child perpetrators: Children who molest other children. *Child Abuse & Neglect, 13,* 571–585.

Johnson, T. C. (1991). Children who molest children: Identification and treatment approaches for children who molest other children. *APSAC Advisor, 4*(4), 9–11, 23.

Johnson, T. C. (1995). *Treatment exercises for child abuse victims and children with sexual behavior problems*. South Pasadena, CA: Author.

Lamb, S., & Coakley, M. (1993). "Normal" childhood sexual play and games: Differentiating play from abuse. *Child Abuse and Neglect, 17,* 515–526.

Lindblad, F., Gustafsson, P. A., Larsson, I., & Lundin, B. (1995). Preschoolers' sexual behavior at daycare centers: An epidemiological study. *Child Abuse & Neglect, 19,* 569–577.

Moos, R. H., & Moos, B. S. (1981). *Manual for the Family Environment Scale*. Palo Alto, CA: Consulting Psychologists Press.

National Adolescent Perpetrator Network (1993). The revised report for the National Task Force on Juvenile Sexual Offending. *Juvenile and Family Court Journal, 44,* 1–120.

Okami, P. (1992). "Child perpetrators of sexual abuse": The emergence of a problematic deviant category. *Journal of Sex Research, 29,* 109–130.

Pino, C. J., Simmons, N., & Slowski, M. J. (1984). *Children's version of the Family Environment Scale*. East Aurora, NY: Slosson.

Pithers, W., Gray, A., Busconi, A., & Houchers, P. (in press). Five empirically-derived subtypes of children with sexual behavior problems: Characteristics potentially related to juvenile delinquency and adult criminality. *Irish Journal of Psychology*.

Reynolds, C. R., & Richmond, B. O. (1985). *Revised Children's Manifest Anxiety Scale (RCMAS)*. Los Angeles: Western Psychological Services.

Ryan, G. (1997a). Sexually abusive youth: Defining the population. In G. D. Ryan & S. L. Lane (Eds.), *Juvenile sexual offending* (pp. 3–9). Lexington, MA: Lexington Books.

Ryan, G. (1997b). Theories of etiology. In G. D. Ryan & S. L. Lane (Eds.), *Juvenile sexual offending* (pp. 19–35). Lexington, MA: Lexington Books.

Speilberger, C. D. (1973). *Manual for the State–Trait Anxiety Inventory for Children*. Palo Alto, CA: Consulting Psychologists.

van der Kolk, B. (1989). The compulsion to repeat the trauma: Re-enactment, revictimization, and masochism. *Psychiatric Clinics of North America, 12,* 389–411.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

9.29
Perpetrators of Racial Violence and Hatred

ALAN S. W. WINTON
Massey University, Palmerston North, New Zealand
and
NIRBHAY N. SINGH and KRISTEN McALEAVY
Virginia Commonwealth University, Richmond, VA, USA

9.29.1 INTRODUCTION 467

9.29.2 NATURE 468
  9.29.2.1 Race, Ethnicity, and Culture 468
  9.29.2.2 Prejudice and Hatred 469
  9.29.2.3 Stereotypes 470

9.29.3 VIOLENCE 470
  9.29.3.1 Theories of Violence 471
  9.29.3.2 Factors in Violence 472
  9.29.3.3 Violence and Hatred 473
  9.29.3.4 Collective Violence 473

9.29.4 PERPETRATORS 474
  9.29.4.1 Influences on Violence 475
  9.29.4.2 Influences on Racial Violence 475
  9.29.4.3 Problems of Measurement of Racial Violence and Hatred 476
  9.29.4.4 Extent of Racial Violence and Hatred 477

9.29.5 INTERVENTIONS 478
  9.29.5.1 Treatment 478
  9.29.5.2 Prevention 478
  9.29.5.3 Tolerance of Violence 478

9.29.6 FUTURE DIRECTIONS 479

9.29.7 SUMMARY 479

9.29.8 REFERENCES 480

9.29.1 INTRODUCTION

Nearly every day there are reports in the news media of acts of harassment and hatred perpetrated on people because of their race or ethnicity. These include not only overtly violent attacks, but also threats of violence, intimidation, and damage to property. Further, many people from racial and ethnic minorities report multiple instances of "lower-level" harassment, such as verbal abuse, where their race or ethnicity is seen as a crucial factor. These are not simply instances of bigotry, prejudice, and discrimination, but constitute real threats to their safety and well-being.

Problems of racial violence and hatred are worldwide; while the form and extent of the problems vary widely, no country is free of them. In some parts of the world, racial violence and hatred are part of conflicts between countries or regions within a country, threatening the lives of all people within these countries and often of those from other countries as well. The conflicts may involve other nations directly or indirectly due to the displacement of large numbers of people from neighboring countries, creating racial and ethnic tensions. In most nations, even without such conflicts, tensions and conflicts between different racial and ethnic groups exist which have the potential to develop into violence and other overt expressions of hatred.

Racial violence and hatred are increasing in many parts of the world because of increasing population and rapid political and social changes. Since the United Nations began in 1945, there have been over 150 small-scale wars with deaths estimated at over 20 million, and disablement and displacement of over 80 million people (Hamburg, 1993). Even without such conflicts, widespread movements of large numbers of people necessitate the sharing of finite resources, not only with more people, but also with people of diverse races and cultures. Successful ways of dealing with the inevitable racial tensions that arise are needed. The destructive power of modern warfare has made the necessity to find solutions to control violence and hatred between people most urgent.

Given the diverse causes of violence, it is necessary to balance theory and practice to find effective ways to lessen violence (National Research Council, 1993). The important focus for clinical psychologists wishing to reduce racial violence and hatred should be on the individual perpetrators, and factors in their lives that can change their attitudes and behavior. However, psychologists can also play an important role by increasing our understanding of the nature of the factors that produce racial violence and hatred, and by researching methods which may prevent their development.

This chapter focuses on the nature of racial violence and hatred. It begins with a brief discussion on some key concepts related to race and hatred. This is followed by a brief overview of violence, including factors that may lead to racial violence, and issues related to collective violence, such as riots and mob violence. Then there is a discussion of the characteristics of perpetrators of violence in general and racial violence in particular. Questions about our knowledge of the number of perpetrators of racial violence and hatred are raised because long-term policy and intervention decisions are dependent on accurate information about the extent of the problem. Finally, issues of treatment and the need for further research are discussed.

9.29.2 NATURE

A variety of analyses have been offered to explain the nature of racial violence and hatred. One major distinction has been made between essentialist explanations and historical-deterministic explanations (Baird & Rosenbaum, 1992a). Essentialist explanations point to causes that are intrinsic to our human nature, involve deep-seated tendencies, and are difficult, if not impossible, to overcome. Historical-deterministic explanations emphasize factors in our environment that change over time and situation, and often are reasonably amenable to manipulation, such as inadequate housing or poverty. To understand the nature and development of racial violence, we must first define several key concepts.

9.29.2.1 Race, Ethnicity, and Culture

Use of the term race, and its derivatives racial, racism, racist, raises a number of potential problems. The term is based on distinguishing amongst peoples on the basis of physical characteristics and traits that are biologically inherited, although the construct of race no longer has any scientific status. It has been defined in a variety of ways that have included biological, national, ethnic, and cultural variables (Solomos, 1993). Many people have argued that it is, at best, a term of limited value which is better avoided altogether because any racial categorizations invariably involve surplus meaning (Gadlin, 1994). However, notions of race continue to be used widely across societies and discourses and can be salient even in situations where race is said to be irrelevant.

Racism has been broadly defined to cover ideologies and social processes which discriminate against others on the basis of racial categories. This has been extended to cover discrimination based on assumed superiority, which is itself based on assumed biological differences (Solomos, 1993). In many societies, such racism is mediated in a variety of ways, including political discourse and the media, and through educational and other institutions. Racism usually has unspoken implications of domination and oppression or subordination of minority cultures (Gadlin, 1994).

Researchers have classified racism in a number of ways. For example, Dovidio and Gaertner (1986a) distinguished three types of racism: individual racism, institutional racism, and cultural racism. Individual racism is similar to racial prejudice but emphasizes biological considerations and encompasses discriminatory acts. Institutional racism refers to institutional policies, whether intentional or not, that unfairly restrict opportunities for certain groups of people. Cultural racism involves elements of the other two types that argue for the superiority of one race's cultural heritage over another's. A related term, ethnocentrism, can be distinguished from racism. Racism involves discriminating against groups of people on the basis of assumed racial characteristics, including behavior interpreted as being due to their race. Ethnocentrism involves discriminating against groups on the basis of their ethnicity and emphasizes a cultural group's unique characteristics.

Ethnicity has had a number of definitions, ranging from those where it is largely a euphemism for race through to broader definitions that include language and customs. In broader definitions of ethnicity, physical characteristics are usually part of the definition, but are not the preeminent defining aspect. Regardless of how it is defined, ethnicity is an important concept because a substantial amount of violence around the world involves groups who define themselves on the basis of ethnicity (Nagel, 1995).

Culture has been given a very wide range of meanings extending from very broad universalistic ones to quite narrow ones (Carter & Qureshi, 1995). Culture can be thought of as " an ordered system of meaning and symbols in terms of which individuals define their world, express their feelings, and make their judgements" (Geertz, 1962, p. 724). Broadly defined, culture includes "the shared values, traditions, arts, history, folklore, and institutions of a group of people that are unified by race, ethnicity, nationality, language, religious beliefs, socioeconomic status, social class, sexual orientation, politics, gender, age, disability, or any other cohesive variable" (Singh, 1996, p. 123).

Most definitions of prejudice emphasize negativity towards a group or an individual based on their membership in that group. Dovidio and Gaertner (1986a) defined prejudice as an *attitude* with cognitive (e.g., irrationally based beliefs), affective (e.g., dislike), and conative (e.g., avoiding) components. Prejudice is seen as highly embedded in psychological and social processes, developing from fundamental biases in the ways in which humans process information. Further, Dovidio and Gaertner (1986a) distinguished motivational from cognitive perspectives. Motivational perspectives assume that prejudice is caused and maintained to achieve goals and needs, or because of negative feelings and beliefs acquired through socialization. Cognitive perspectives emphasize information processing biases in the form of stereotyping and the consequences on intergroup interactions, with more favorable attributions and greater differentiation towards members of in-groups when compared to those from out-groups. Discrimination is regarded as selective, unjustified *behavior* towards members of a group that denies them equality.

9.29.2.2 Prejudice and Hatred

The literature on the nature and causes of racial prejudice is large and varied. For example, psychological theories about the nature of prejudice stress psychosocial explanations that emphasize the role of in-groups and out-groups, and the development of stereotyping.

There are a number of different accounts that emphasize different causes for the development of prejudice. Some involve religious notions, such as"original sin," as being causative; others stress biologically determined factors; while many provide psychosocial explanations. Even with psychosocial explanations, there are differences both in the actual factors stressed and whether they have an inherent basis or are largely ontogenic. For instance, a lot of the literature stresses the role of stereotyping in the development of prejudice, but the tendency to stereotype seems to be considered as being inherent in human nature (Albó, 1994; Dovidio & Gaertner, 1986b).

Not only do the explanations differ in terms of the nature of prejudice, they also differ in terms of the remedy. For instance, if the only cure for the effects of original sin is redemption through a religious conversion, those not subscribing to the need for such a change will remain prejudiced. The outlook for overcoming prejudice is even less promising when biologically determined factors are considered to be of prime importance because there is little that can be done about these factors. Psychosocial viewpoints, although varying in the nature and importance of various factors, stress learning in a social context. Psychosocial viewpoints provide a much more positive prognosis, because if something can be learned it can also be unlearned, or alternative learning can overcome the negative effects of prejudice.

Fundamentalist religious, biological, and social psychological views are examples of essentialist thinking that have been used to explain prejudice and hatred (Baird & Rosenbaum, 1992b). The fundamentalist religious view holds that prejudice and hatred are part of a person's inherently sinful nature, and these can be overcome only by a radical transformation of the self through God's mercy. Another type of essentialist thinking is found in biological views which suggest that evolution has selected the tendency to maintain the genetic integrity of our own group, and it is this tendency that produces prejudice and hatred. This will occur inevitably under certain conditions, even without awareness, and little can be done to avoid its occurrence if these eliciting conditions are present. While most psychosocial explanations point to factors in the environment of the individual as the cause of prejudice and hatred, many researchers regard stereotyping as critical. Further, other researchers have argued that we have a deep-seated need to simplify our interactions with others and this produces a tendency to stereotype people of other cultures. Prejudice and hatred are considered a side effect of this intrinsic tendency.

9.29.2.3 Stereotypes

Various theories point to the tendency of humans to stereotype other humans who are different from them. Stereotypes involve the process of generalization, an important and unavoidable part of learning, where potentially discriminable differences are treated as the same. But stereotypes are generalizations with two important features: they totally obscure individual differences among members of a group, and they tend to resist any contradictory evidence, no matter how compelling it may be. Individuals who do not fit the stereotype will be seen as the exception to the rule and may even be used to justify the group stereotype by arguing how few in that group possess those individuals' characteristics. If stereotypes are repeated, even in jokes, they become accepted and, once accepted, are used as the basis for dealing with individuals from the group being stereotyped.

Stereotypes usually involve negative evaluations of perceived attributes, such as aggression rather than motivation, or positive ones that are used for the purpose of being critical, such as being good at music (meaning not much else) or being powerful (meaning too powerful). Stereotypes often categorize the group as inherently inferior. For instance, if members of a group are seen as an economic threat, they may be categorized as greedy, having obtained their wealth in suspect ways, or not in need of any assistance.

A stereotype tends to be held as true over long periods because it is neither tested nor even thought necessary to be tested. Hence, distortions of social realities can be demonstrated by the considerable misinformation that many people have about members of other groups. This distortion has considerable ramifications, because decisions affecting the lives of members of other groups are often made on the basis of the stereotypes.

9.29.3 VIOLENCE

There is no universally accepted definition of violence (Bradby, 1996). The concept of violence covers a diverse range of behaviors and situations and can involve various levels of explanation (National Research Council, 1993). Aggression is an important concept in many analyses of violence and a lot of research, both basic and applied, has investigated causes and control of aggressive behavior. Studies on aggression are frequently cited in discussions of the development, maintenance, and control of violence, especially those emphasizing psychosocial perspectives (e.g., Pepler & Slaby, 1994). Some caution is warranted, however, because aggression is defined in some studies differently from the way violence is usually defined, and the relationship may not always be clear. It is also true that some definitions of violence include types of violence where overt aggression may be absent, such as in structural violence.

Violence, like culture, has no single agreed definition (Rupesinghe, 1994). Some definitions are very broad, such as that of Galtung (1990), who defined violence as anything avoidable that impedes human self-realization. Galtung identified three types of violence: direct violence, structural violence, and cultural violence (see MacGregor & Rubio, 1994). Direct violence refers to the use of physical force that is directed at other humans. It involves face-to-face aggression where the perpetrator can or may be identified. Structural violence refers to the indirect violence caused by social structures, not persons, although the structures are created and changed by persons and affect all persons in the society, especially the relationships between them. Structural violence includes broad differences in wealth distribution; inequitable administration of justice; authoritarian family socialization; discrimination based on such things as age, gender, race, and religion; and the repression of speech. Institutional violence is a form of structural violence embedded in institutions and accepted, or at least tolerated, by complicity. It can either be formal, such as an army explicitly trained to repress subversion, or true, where it exists despite being not formally acknowledged or even prohibited. Cultural violence which, according to Galtung (1990), can be defined as the denial of identity, security, and symbolic meaning, are those aspects of a culture that legitimize direct or structural violence. This includes societal rules that are not strictly cultural but which have a cultural layer, such as economic policies that are theoretically fair but in reality disadvantage certain groups.

While the three types of violence are distinguishable, they are closely related and not mutually exclusive. For instance, factors resulting in structural violence may not of themselves produce direct violence but, in the presence of other factors, may contribute to direct violence. One distinction between the three types of violence is in their relation to time. Direct violence is a present fact, structural violence is a process which may change over time, while cultural violence is nonvariant over very long periods of time. In addition, it is also important to distinguish violence from its effects. For instance, an injury suffered in a brawl is the effect of violence inflicted by a person but is not direct violence *per se*. Similarly, widespread poverty is not structural violence *per se* but the rules organizing unequal distribution of wealth are.

Direct violence is of major concern, because perpetrators of this type of violence are likely to be the ones with whom clinical psychologists are directly involved. Structural and cultural violence, however, are also important because of their long-term and widespread effects on culturally diverse groups. In trying to deal with the causes of violence, psychologists should be concerned with dealing with short- and long-term factors involved in violence. Further, there is strong evidence that structural factors set the conditions that increase the likelihood of outbreaks of direct, and possibly widespread, violence.

Toplin (1994) has provided another typology of violence. The first type is collective and political violence, which includes group activities, such as protests, demonstrations, strikes, and riots. The second type is criminal violence, which includes acts of violence by individuals or small groups. The third type is psychopathic violence, which includes assassinations, mass gratuitous violence, and serial killing. The final type is foreign policy violence. These distinctions may be very useful from a sociological viewpoint, but they do not define violence in terms of the perpetrator.

9.29.3.1 Theories of Violence

It is clear that violence is multidetermined and can be best understood by considering not only a variety of psychological studies that have investigated the nature and causes of human aggression, but also studies from other disciplines, in particular criminology and sociology (Berkowitz, 1993; Eron & Slaby, 1994). As with prejudice, there are a variety of theories as to how violence develops. Some theories stress inherent factors which are largely genetically determined or, as in psychodynamic theories, inherent structures of the mind. Many psychological explanations stress psychosocial factors. In an overview of current research on violence, Eron and Slaby (1994) concluded that there is overwhelming evidence that human aggression is largely learned and not mainly (and hence, inevitably) determined by biology. Biological factors may interact with learning processes and predispose a person to violence, but they do not make violence inevitable. In particular, violence is not the direct and inevitable result of extreme anger or inadequate impulse control, although inadequate control may put some individuals at greater risk for violence if they have previously learned to be violent. Because violence is largely learned behavior, it should be possible to reduce or prevent violence through learning.

Berkowitz (1993), among others, distinguished two forms of aggression, instrumental aggression and hostile aggression. Instrumental aggression refers to violence perpetrated on the victim largely as a by-product of achieving some goal, such as attaining money or social status. In hostile aggression, however, inflicting violence on the victim is a primary purpose. Berkowitz (1993) argued that hostile aggression is often accompanied by strong emotion (called emotional aggression) but may be inflicted without emotion.

Behavior can be modified by two fundamental conditioning procedures: respondent (or classical) and operant (or instrumental). In respondent conditioning, events in the environment (including the internal and social environments) come to elicit behaviors by association with reflexes or other genetically determined environment–behavior relationships. In operant conditioning, behaviors are modified by their consequences. Thus, both instrumental and hostile aggression are produced by operant conditioning; what distinguishes them is the reinforcing consequences that maintain them. This is a difference in motivation, an important consideration in explaining violent behavior (Tsytsarev & Callahan, 1995).

In hostile aggression, the infliction of injury or pain on the victim is an important consequence. Emotional aggression involves, at least in part, respondent conditioning, but this would be an accompaniment of the hostile aggression rather than a cause. Berkowitz (1993) argued that anger is an emotion that may or may not accompany an act of aggression and, even when it does accompany it, is not the cause of the aggression. It is likely that some forms of hostile aggression would involve respondent conditioning, with it being elicited by someone or something in the environment rather than it being maintained by its effects on the victim.

9.29.3.2 Factors in Violence

While most reviews of the causes of violence have not focused specifically on racially motivated violence, it seems clear that many of the factors identified as important for understanding violence are also important for understanding racially motivated violence. Studies looking at hate crimes in general and racially motivated hatred in particular will be considered in the next section.

A number of factors have been implicated for understanding the psychology of violence and its development (Berkowitz, 1993; Hill, Soriano, Chen, & LaFromboise, 1994; National Research Council, 1993; Pepler & Slaby, 1994). These can be grouped in various ways but include social factors (both macro- and microsocietal), individual characteristics (both psychosocial and biological), and factors labeled as predisposing (including processes), situational, and activating (triggers).

Social and cultural factors increase the probability that members of certain groups will be perpetrators. Major structural factors that have been implicated include low socioeconomic status, ethnic heterogeneity, and residential mobility. Research has further identified a wide range of specific factors, including concentrations of poor families in particular geographical areas, differential income equality, differential social organization affecting supervision of youth, high population turnover, community transitions, family disruption, housing and population density, lack of legitimate business opportunities, lack of legitimate social activities and contacts to gain social status, easy availability of drugs, presence of opportunities associated with violence (e.g., illegal markets in drugs and firearms), and inadequate child care and schooling.

The effects of poverty on patterns of violent behavior have not been well researched and are interwoven with racial differences, making it difficult to untangle their differential effects. Research evidence suggests, however, that group violence is predicted by economic inequality rather than absolute level of material deprivation and also by the absence of available opportunities (Eron & Slaby, 1994).

Other major societal influences on violence that are not formal ecological and sociological variables are the media, guns, mobs, and gangs. The influence of the media has been greatly underestimated despite clear evidence that violence can be acquired through social learning and not just through personal experiences (Berkowitz, 1994; Eron & Slaby, 1994). Repeated viewing of violence can have long-term cumulative effects that include a propensity to engage in violence, an indifference to violence perpetrated by others, and an exaggerated view of the prevalence and inevitability of violence in the real world. Despite some conflicting findings, there is overwhelming evidence that TV violence leads to aggressive behavior and that TV and films can spread the potential for violence rapidly through the community (National Research Council, 1993). However, TV and other media can also play a strong positive role in reducing the potential for violence.

The call for greater gun control to reduce violence has given rise to strong debate, especially in countries like the United States, where the right to carry a gun has long been in existence and where the ownership of guns is widespread. Berkowitz (1993, 1994) has argued that, while the number of violent incidents may not be greater due to the availability of guns, the severity of injuries and the number of deaths are. He contended that any protection guns may afford is strongly offset by the greater probability of death due to guns, especially for innocent people, and any reduction in gun numbers will lower this probability. He further argued that even the sight or presence of a gun can increase the probability of aggression, especially during situations when great emotion is coupled with violent action.

Other factors vary in terms of their temporal proximity to a violent event. Some individual characteristics occur before birth and in early childhood (e.g., temperamental disposition), while others occur later in life (e.g., experiences in adulthood increasing frustration and anger). Some societal factors occur after centuries (e.g., cultural acceptance of violence), while others occur over comparatively shorter periods (e.g., poverty).

While various factors have been implicated in violent behavior, it is important to understand that none by itself will inevitably lead to violence. Violence is multidetermined, with each risk factor contributing to the likelihood of its occurrence (Eron & Slaby, 1994). When several factors converge, violence is likely but still not inevitable. This means that the more factors are present, the greater is the probability of violence occurring. (Berkowitz, 1993). In each violent event, there are variations in the role of each of these factors and in the potential for violence of the individuals involved. Hence, there are many different opportunities to intervene to reduce or even prevent violence from occurring.

9.29.3.3 Violence and Hatred

While our focus is on racial and ethnic hate violence, some useful insights can be derived from considering how hate violence may be distinguished from violence in general. Some studies have looked at hate crimes perpetrated on individuals characterized not by race or ethnicity (e.g., Herek & Berrill, 1992a) or not solely by race or ethnicity (e.g., Levin & McDevitt, 1993). These studies show that many perpetrators of hate violence show a number of common characteristics, irrespective of what groups their victims belong to, and many of them exhibit violence and hatred towards people from more than one cultural group (Levin & McDevitt, 1993).

In analyzing hate crimes in the United States, Levin and McDevitt (1993) noted that the distinguishing feature of hate violence is that it is motivated by hate rather than greed or opportunity. It may be executed on people directly or on their property, but it is usually excessively brutal. Hate violence towards people typically involves using force beyond that necessary to simply subdue the victim. It tends to be executed on strangers, emphasizing that anyone from the group the victim belongs to could be a target and should be fearful because victims are interchangeable. This is tantamount to domestic terrorism, because it sends out signals of hatred and threats of violence and harm to all people with the characteristics by which the victim is being identified and not merely to the actual victims. This type of violence tends to be perpetrated by multiple offenders rather than individuals acting alone.

Levin and McDevitt (1993) distinguished reactive hate violence from mission hate violence. Reactive hate violence is based on the perceived need to eliminate the threat of encroaching outsiders. This violence usually involves narrowly focused attacks aimed at particular persons; it ceases once the threat is perceived as gone. An example might be the destruction of a shop owned by a person of a minority race who has recently come into a community and set up a business that threatens the livelihood of one or more members in the community. Mission hate violence is violence directed towards groups perceived to be a threat to some general concept such as a country's way of life, or to basic moral standards. Perpetrators of this kind of violence often perceive these groups to be part of a conspiracy, but it is not always clear why they perceive a particular victim as being a representative of the hated group.

9.29.3.4 Collective Violence

Forms of collective violence that have been distinguished include group violence, gang violence, mob violence, riots, terrorism, and various forms of state-sanctioned violence, such as that by police and armies. Staub and Rosenthal (1994) used the term group violence to describe organized violence by the state or a dominant group against a minority race or culture. Mob violence, on the other hand, can be enacted by both majority and minority groups and may occur with either organized or spontaneous groups. A riot is an instance of mob violence, and a mob is a crowd that is likely to engage in illegal acts and violence. A crowd is a group of people who share a common interest and whose emotions may be easily aroused; whether they will become a mob depends on the presence of factors that can increase their susceptibility to violence.

Riots clearly illustrate the danger of not dealing with structural problems associated with violence (Staub & Rosenthal, 1994). The 1968 President's Commission on Civil Disorders listed five powerful ingredients that are catalysts for riots (Toch, 1992). These included: (i) frustrated hope generated in the civil rights struggle; (ii) a climate that approved and encouraged violence; (iii) frustration due to feeling powerless to change the system; (iv) a new mood, especially in the young, of self-esteem and racial pride; and (v) the police being viewed as a symbol of white power, racism, and oppression. In concert, these factors lead to the view that riots occur at the end of a chain of events involving unrealized attempts to redress a number of grievances by more acceptable means; riots occur as a last resort after an accumulation of despair, anger, and frustration, although they can be triggered by a relatively minor event that is unrelated to the major grievances of the group(s) involved.

Major themes that can be identified from interviews with riot participants include: (i) police brutality, which is seen as largely racist; (ii) a retaliation against White exploitation; (iii) a chance to get back at Whites, which is sometimes seen as worthwhile at any cost; (iv) an expression of hopelessness; (v) anonymity, which allows members of an oppressed minority group to be collectively noticed and taken seriously; and (vi) a finding of identity and even pride in their own cultural group.

Research suggests that the following are possible precipitating factors for riots: (i) persons being predisposed to action by many unredressed grievances; (ii) a situation that allows the event to be interpreted as a further act of insult or humiliation, even if legitimate and innocent; (iii) ambiguity or a development within the incident; (iv) temperature, especially a heat wave; (v) a change in background factors (e.g., hopes recently raised by changes elsewhere); and (vi) the recent occurrence of riots elsewhere. All of these factors have the potential for inciting racial and ethnic riots, as well as violence.

One important factor in many riots is that it involves perpetrators who would not normally engage in violent behavior. In analyzing race riots in the United States and England, Grimshaw (1969a) has argued that racial violence is due mainly not to conscious policy decisions but to reactions of the dominant group to real or perceived threats to its accommodative structure. Throughout the history of the United States, when African-Americans accepted a submissive role, Black–White violence was relatively rare; when the status quo was challenged, race conflict often ensued. Grimshaw has suggested that Blacks in England have largely accepted or tolerated their situation because of their shorter history and because there have been fewer of them. He argued that even when prejudice and discrimination have given rise to strong social tensions, the occurrence of social violence has depended on the extent of constraint by external forces.

Social violence is not always preceded by a gradual build up of social tension. Grimshaw (1969b) has identified a number of macro- and microsocial factors that contribute to the build up of the social tensions underlying social conflicts. Macrosocial factors include changes in the intellectual milieux, in the economic organization, in international relationships, and urbanization and industrialization. Microsocial factors include specific incidents that provide occasions for social disturbances but which depend on a sufficient level of social tension and a sufficiently low level of external control. He argued that deteriorating relations between conflicting groups leads to violence. The usual sequence is the identification of an opposing group, relationships with that group become strained, social tensions gradually increase, which eventually leads to social and physical violence.

9.29.4 PERPETRATORS

In general, much less is known about perpetrators than victims of violence (see Chapter 23 this volume). The National Research Council (1993) has summarized the characteristics typical of perpetrators of violence. The perpetrators usually: (i) are from mainstream society rather than deviant types in organized hate groups; (ii) share similar demographic profiles to victims; (iii) are males (89% of all arrested) below 30 years of age; (iv) are disproportionately drawn from racial and ethnic minorities; (v) are somewhat older than their victims; (vi) are multiple offenders in 25% of nonfatal offenses; and (vii) commit violent crimes infrequently.

In summarizing the features of many young violent perpetrators in the United States, Eron and Slaby (1994) noted that the form and absolute level of individual aggression may vary over a lifetime but, for an individual, their level of aggression relative to peers remains remarkably constant and predictable. For most, the pattern of violence at an early age is a good predictor of their violence decades later. Many young perpetrators tend to share common social experiences, especially strong rejection by peers, but they differ in a variety of ways in personality characteristics and how they deal with social problems. Their cognitions seem to play an important role. Violent individuals tend to perceive environmental events and responses by others differently from their nonviolent peers. They frequently show a history of ineffective parenting, drug use, school failure, and poor employment history. They report thoughts that demonstrate poor problem-solving skills, greater hostility, and a greater acceptance of aggression and violence than their nonviolent peers. These thoughts typically develop early, become habitual, and are associated with violent behavior. Therefore, if their thoughts can be modified, it may be possible to modify their overt violent behavior (Berkowitz, 1993).

Toch (1992) has also analyzed the characteristics of perpetrators of violence. He suggested that the two types of orientations that are most especially likely to produce violence are when a person sees others as tools to serve his needs, and when a person is feeling vulnerable to diminishment. Both of these are based on the premise that human relationships are power-centered and that efforts at self-assertion have a desperate, feverish quality suggestive of self-doubt. Toch argued that all violent people may not initially fit this categorization, but when they engage in violence they learn to respond in this way. The propensity to violence appears to be due to a failure of socialization and positive self-perception; this probably starts with experiences early in life, when it was largely elicited behavior that was subsequently reinforced and later maintained by others.

9.29.4.1 Influences on Violence

It is clear that there is no one set of factors, whether individual characteristics, family and other interpersonal interactions, or social situations that can account for the development of the potential for violence. Also, some factors may only put individuals at risk if certain conditions are present. For instance, some studies have shown that prenatal and perinatal complications are related to later aggression but only for those raised in unstable homes (National Research Council, 1993). However, the mechanisms by which these factors work to encourage the development of aggression is not clear.

There also appear to be protective factors which may actually reduce the effect of other risk factors. The National Research Council (1993) reported on several studies that had focused on why many children at risk for violence do not offend. These studies found evidence that shyness, nervousness, or social isolation, being firstborn, having a high IQ, having parents who regularly attended religious services, and coming from smaller families with low discord are factors associated with low or reduced risk. These factors, however, may not work in a simple additive way; for example, in one study, shyness predicted later offending only if coupled with early aggressiveness.

Individual personality characteristics found to be present in violent offenders during childhood include hyperactivity-impulsivity-attention deficit, restlessness and lack of concentration, high risk-taking, poor at delaying gratification, low empathy, and (more often than other offenders) low IQ (although it may not be IQ *per se* that is predictive of aggression, but rather the aggression that can develop at a young age with low IQ being a factor).

There is some evidence that early temperament correlates highly with later temperament, especially for those who are very inhibited or very uninhibited (National Research Council, 1993). A very uninhibited temperament is a risk factor for later aggressiveness, while a very inhibited temperament may be a protective factor against later aggressiveness, and these effects may be especially true in children from homes which have a high risk for violence. Because aggression usually appears earlier, the effect of puberty is probably not a major cause of aggression. It may, however, affect the frequency or intensity of aggression.

Family factors that have been implicated are poor parental child rearing methods, including physical abuse, lack of warmth, rejection and hostility, poor supervision, and no caring adult or friend for support. Many violent people have alcoholic or criminal parents and many have been separated from parents as children. They often come from low-income, large-sized families, with poor housing in deprived inner city locations.

Many violent offenders have been victims of physical abuse and this has been implicated by many as a causative factor. A causal connection, however, has not been established. The methodology of many studies in this area has been weak, and the causal link did not hold strongly for females and Whites. Many violent adults were neglected in childhood but were not subjected to undue violence. Further, parental abuse may have been a reaction to the child's aggressiveness. The best predictor of adult aggressiveness is child aggressiveness, irrespective of parental behavior. What may be passed on is a parental genetic predisposition to aggression rather than any modeling or reinforcement of aggression.

There is clear evidence that people from ethnic minorities in the United States are at much greater risk of being both perpetrators and victims of violence (Eron, Gentry, & Schlegel, 1994). This may not be due to any inherent propensity or a result of any direct encouragement of violence by the culture, but rather due to societal factors interacting with these cultures to put them at much greater risk (National Research Council, 1993). By the time children are five to six years old, they have learned to identify or share psychological characteristics with their parents, and ethnic and class groups. If these are defined as undesirable and evoke hostility by the dominant culture, and are seen as barriers to achieving desired goals, these create strong psychological conditions for making these children disproportionately at risk for aggressive and violent behavior.

9.29.4.2 Influences on Racial Violence

Racial violence appears to have its roots in an individual's learned prejudices against particular groups of people. While much is known about the development of prejudice, little is known about the relationship between prejudice and racial or hate violence. Some generalizations can be made about perpetrators of racial violence and hatred, although these do not apply to all individuals. The perpetrators of different types of hate violence show somewhat different characteristics, and even perpetrators of similar types of violence are not all identical.

Levin and McDevitt (1993) noted that many perpetrators of racial violence live on the margins of their community. They have been designated as cowards because they often offend in groups, where they may find safety in numbers and achieve some anonymity. Being in a group often encourages mutual incitement to increase the intensity of hatred, with some individuals seemingly out to prove themselves by surpassing previous atrocities. Many who perpetrate hate violence against one group also seem to perpetrate it against other groups (Berrill, 1992; Herek & Berrill, 1992b).

Perpetrators usually see their actions as justified and even necessary; resentment against the group the victim belongs to often forms a significant part of this justification. Many perpetrators appear to have lost self-esteem or status, often with economic loss, and they blame the presence or action of out-groups for this. These factors are often present in both reactive and mission hate crimes. Reactive hate crimes often involve adult perpetrators who see the need to defend themselves against the encroachment by outsiders. However, they tend to stop committing violent acts when the perceived threat is no longer present and may never be violent again. Perpetrators of mission hate crimes tend to be extremists and incredibly bigoted. Although they are often seriously paranoid and delusional, not all are considered mentally ill. Other perpetrators, usually young males, appear to commit hate crimes largely for the fun of it. For them, acceptance by a group is often very important and provides a major motivation for committing the crime.

9.29.4.3 Problems of Measurement of Racial Violence and Hatred

It is difficult to assess prevalence rates of acts of racial violence and hatred. For example, exactly how do we define racially motivated acts? How can racially motivated violence be distinguished from other forms of violence? Clearly, all acts of violence by one person towards a person of a different race or ethnicity, such as in a robbery, should not be automatically considered to be racially motivated. Even where there is clear evidence that a violent act is motivated by hatred, how can this be distinguished from hate violence because of other factors, such as gender or sexual orientation?

Another set of problems concerns whom the statistics are collected on. In many countries, people from diverse cultural backgrounds are not distinguished on the basis of race and ethnicity, or are lumped into single large groupings. In addition, while figures are being kept in some countries on the prevalence of violence among different racial or ethnic groups, these are not usually broken down further on the basis of the race or ethnicity of the victim, so it is difficult to distinguish intergroup from intragroup acts of violence.

Further, there are serious concerns regarding how violent events are defined and the data on them collected. There is no single way to define, classify, and measure violent events, and most national systems vary in what they count and how they count it. Many measures are derived from police records, which tend to deal only with the most serious acts that are likely to lead to criminal prosecution. Police officers often have discretion in classifying and counting violent crimes and the definitions of violent behavior as crimes often change over time. Some classifications include social constructions which provide more complete information than those that do not, but these can cause distortions as they involve various persons' perceptions and classifications which can vary in such things as to what secondary information is included. Finally, a society's interpretation of violent events as crimes, changes over time.

These issues are particularly true of racial violence, which is distinguished from other violence by the presumed role of social relationships in the motivation of the perpetrator (National Research Council, 1993). Determining unequivocally that a motive is racist can be difficult, especially in instances where there is an absence of clear signals, such as graffiti or organizational identity. Classifying a violent act as a racially motivated crime can depend very much on the view of the law enforcement officers dealing with the complaint. This makes accurate documenting and analyzing of patterns of racial violence very difficult. This is particularly true in instances where some crime, like robbery, has been committed, but where a hate component may be present as well.

Also important is that racially motivated events may not be recognized as crimes by the police and other law enforcement officials and therefore not reported. Hence, there is value in examining other statistics, such as those that have been kept by advocacy organizations. However, these data may also provide distorted prevalence figures because these organizations often lack the resources for classifying and counting according to rigorous criteria, may interpret ambiguous events as evidence of racial hatred and violence, and may include crimes other than violent ones (National Research Council, 1993). Figures based on interviews among samples of particular populations do tend to show much higher rates of hate-motivated violence than those reported in official figures, probably because they may include a much wider range of incidents and not just criminal offenses (Eron & Slaby, 1994).

These may include acts of racial violence and hatred labeled as "low-level harassment." Related to this is the increasing use of broader definitions of racial violence which include incidents caused by structural violence, but which may not result in personal violence.

With increasing acceptance of the need to collect data on hate crimes, there may be a greater tendency for people to report these than previously. For instance, in the United Kingdom the early police figures showed a much lower prevalence of hate crimes than were indicated by more informal data collection (Virdee, 1995). Following social pressure for the police to keep hate crime statistics, reported hate crimes showed a considerable increase for some time, while informal methods of collecting data showed them to be fairly static. This indicates a greater willingness to officially report hate offenses and, therefore, the increase in official figures was more of an increase in reporting than an actual increase in these offenses.

A related problem is the evidence that many racially motivated acts, even violent acts, are not reported. When members of minority groups are interviewed by a researcher, many will report being victims of acts of violence which they have never told anyone about, often out of fear. This fear may be due to a belief that authorities would not act on the complaint, fear of losing a job or perhaps being deported, and a fear of retaliatory violence for making a complaint. Other reasons include shame and guilt felt by victims, especially in relation to their families.

Finally, it must be remembered that most of the available figures report on the prevalence of the violence in general and do not provide systematic data on the perpetrators. Again, this is particularly true of perpetrators of racial violence and hatred. Even when some characteristics of the perpetrators are given, the data are not detailed enough to address a number of important issues. These include whether the perpetrators are repeat offenders, vary for different forms of racial violence, vary from country to country, or can be distinguished from those who engage in other kinds of violence.

9.29.4.4 Extent of Racial Violence and Hatred

In the United States, there is evidence that much of the overt racial violence of the past has been declining although new forms of racism are already developing (Baird & Rosenbaum, 1992a; Dovidio & Gaertner, 1986a). This has resulted in new inter-racial tensions, because events regarded as racially discriminatory by one group are not so regarded by another. The former group may propose solutions to deal with this that the latter group regard as being racially discriminatory. There are also increasing reports of new levels of racial conflict occurring in neighborhoods, workplaces, and university campuses, these are predicted to worsen (Levin & McDevitt, 1993).

A sharp increase has been reported in hate crimes which, with increasing diversity and complexity, is likely to get worse (Levin & McDevitt, 1993). Other forms of racism are also on the increase which, while seemingly a small problem in comparison to problems such as increasing illicit drug use, have the potential to disrupt communities and spark violence. For example, Levin and McDevitt (1993) have documented evidence of increasing racism in humor aimed at groups that are vulnerable and sometimes defenseless, humor which, by being used repeatedly, helps reinforce stereotypes about these groups. Hate rhetoric aimed at different groups and peoples, particularly minorities, is increasing in pop music (sometimes intertwined with overt themes of violence), and also in public statements by religious leaders, politicians, and media commentators. On college campuses, racial hatred is occurring increasingly, usually involving harassment but sometimes also including physical violence. This upsurge of racist expression is seen as reflecting a deeper strain of prejudice throughout American society. There is a new rising tide of conservatism that is becoming more vociferous in its attacks on minorities, multiculturalism, and diversity (Singh, 1996).

Levin and McDevitt (1993) have argued that not all promoters of hate intend to harm others, but that hatred is becoming more accepted; one indication of this acceptance is the increasingly wide variety of audiences who enjoy media depicting hate messages. One of the dangers in this is that it allows racist people to justify venting hatred on those being stereotyped. This is a particular problem for the younger generation of a society who have not yet developed their own positions and who have the highest representation among the perpetrators of hate violence.

In the United States, the most important official data on hate violence are provided by the Hate Crime Statistics Act enacted in 1990. This act requires annual national reporting on crimes perpetrated on persons or property that manifest evidence of prejudice based on several characteristics, including race and ethnicity. It is designed to remedy inconsistent reporting by individual states by having all information collected in a single format by a single agency (i.e., the FBI). It has resulted in some positive effects in improving data collection and sensitivity of police forces to hate crimes, although there is still some way to go to achieving overall consistency and eliminating problems of reliability and validity due to variations in systems across states (Levin & McDevitt, 1993).

Other national sources of data on violent behaviors considered crimes are the National Crime Statistics (NCS), which are derived from national household samples, and Uniform Crime Reporting (UCR) based on crimes detected by or reported to police. However, it is difficult to discover the exact figures on the involvement of interethnic violence from these data. Vontress (1995) quoted figures showing that homicide is the primary cause of death for members of African-American communities aged between 15 and 34 years—121 times that of Whites in the same age group. While they do show that there is a disproportionate perpetration of violence by people from ethnic minorities, the greatest proportion of this is intra-ethnic (Eron et al., 1994).

9.29.5 INTERVENTIONS

9.29.5.1 Treatment

Few studies have dealt specifically with perpetrators of racial violence and hatred, but an increasing number have looked at treating violent people; these provide some relevant data on how one might deal with perpetrators of racial violence and hatred. Measures for dealing with violence have varied, again, in part, based on the discipline of the investigators. Although there may be considerable overlap, criminologists have looked at the effects of legislation and legal sanctions, public health professionals and educators have investigated the effects of institutional changes and education programs, sociologists at the effects of broad changes in social structures and conditions, and psychologists at the effects of various therapies on individual perpetrators. Each has provided important insights into dealing with the problem holistically, supporting the need for a multidisciplinary approach (Berkowitz, 1993; Hammond & Yung, 1994).

Risk assessment studies suggest that the strongest indicator of the risk of an individual perpetrating violence is aggressive behavior as a child; hence intervention should begin as early as possible. While intervening with violent teenagers is difficult, it is even harder with adults, as learned violent behavior is resistant to extinction. However, because of the seriousness of the behavior it is probably never too late to intervene. Whether or when to intervene requires a good knowledge of the role of all the risk factors and characteristics of the perpetrators; however, this knowledge is still lacking. Further, a lot of predictive accuracy concerns risk factors associated with aggressive behavior rather than violence in general or racial violence in particular. Consequently, there are few good outcome studies on the treatment of perpetrators of racial violence.

Treatment strategies for both perpetrators and victims should be directed at community, family, and individual levels. These strategies would include targeting those groups of individuals who are at highest risk for engaging in racial violence.

9.29.5.2 Prevention

The prognosis for successfully treating violent offenders with psychological procedures ranges from hopeful (Eron et al., 1994) to pessimistic (Berkowitz, 1993). However, there is agreement that much can be done to lessen the probability of violent offending. For example, educational interventions can be aimed at reducing prejudice, such as encouraging positive intergroup interactions. Aronson (1992) described an interesting set of studies in classrooms where the usual independent and competitive behaviors gave way to interdependent and cooperative behaviors which had the effect of increasing academic performance and self-esteem in many students, increasing empathy with other students, and in reducing prejudice between students. The interventions had the greatest effect on young children.

Working with children at the earliest stages has been recommended by others. Racism is learned at a very early age and therefore it is important to deal with instances of it among the very young, even when the racism may be being maintained at home (Siraj-Blatchford, 1994). It is harder to change racism and bigotry at an older age, and prevention is much easier than treatment.

There is no question that changes to legislation with accompanying enforcement methods are important approaches for reducing racial violence and hatred. Most of these methods, however, do little to reduce the underlying hatred and prejudice.

9.29.5.3 Tolerance of Violence

In most societies, violence *per se* is not condemned, only certain expressions of violence. People are not being taught to be nonviolent, but rather when to be violent. In some cultures the range and extent of violence encouraged, or at least tolerated, is considerable. In many societies, there are times when hatred and violence are positively encouraged or even glorified, such as during wars and intercountry conflicts. However, even in peacetime, many societies tolerate high levels of violence by taking action only against some forms of violence and under some circumstances. Many allow sports, movies, and television programs which present both explicit and implicit violence in ways which, if not directly encouraging, are accepting and made entertaining. Although many viewers will not engage in the violence witnessed or portrayed, the widespread acceptance of it is an encouragement to those with a tendency to violence. Often there is an implicit acceptance of the inevitability and, indeed, the justification of violence as a viable and even singular solution. Even acts of terrorism are not universally condemned but are accepted or deplored depending on the particular commentators and the particular group committing the terrorist act (Takooshian & Verdi, 1995).

Perhaps even more widespread is the tolerance of racial harassment and hatred where there may be no overt acts of physical violence. As has been pointed out, words may not simply constitute racial harassment, but ultimately constitute an act of racial violence (MacKinnon, 1993). In a number of countries (e.g., the United States, United Kingdom, Canada, Australia and New Zealand), specific laws are being enacted to deal with this type of racism. A lot more needs to be done to create a climate where such expressions of racism will be actively discouraged.

9.29.6 FUTURE DIRECTIONS

The causes of racial violence and hatred are both widespread and complex. A great deal of work is required both to gain a full understanding of these phenomena, and to find real and long-lasting methods for preventing and reducing them. While there are no easy solutions, there are indications that real progress is possible. What is clear, however, is that further research into finding solutions is urgent because of the threat to the very fabric of many societies. Racial conflicts show no sign of abating and will likely increase with increasing populations and greater mobility. Effective solutions will require the input of a wide range of disciplines. The recommendations from a number of studies on violence are clearly relevant (e.g., Eron et al., 1994; National Research Council, 1993). Of equal importance for long-term prevention are recommendations emerging from studies on the causes of prejudice, bigotry, and discrimination. If these are not reduced, hatred and harassment will be inevitable and violence will remain a high probability.

We do not have a clear indication of the extent of racial violence and hatred even though these are common events in most of our lives. There is a need for longitudinal studies that provide prevalence figures within and between culturally diverse groups (National Research Council, 1993). Current reporting of racial violence is fraught with definition problems, mechanisms for data gathering, and the specificity of the data gathered. We need to know not only that an incident of racial violence occurred, but also details of the incident, including sociodemographic characteristics of the perpetrators and the victims. At present, we have more information on the victims of racial violence than on its perpetrators.

There is a paucity of good studies on the various risk factors for racial violence. There is little information on how developmental processes in ethnically and socially diverse communities affect the emergence, maintenance, and prevention of violence. Public policy response to the rise in racial violence and hatred has largely been driven by fear and anxiety rather than a genuine concern to do the right thing. Central government funding of research on risk factors for racial violence in both the dominant and minority cultures is long overdue.

With immigration increasing in many countries, there will be increasing tensions as more and more of the traditional power will be challenged, and more and more minority groups will demand a share in the nation's wealth and opportunities. Resentment and hostility will increase for a variety of reasons based on perceptions of the migrants' financial success or sharing of the community's limited resources (Peach, 1994; Wickberg, 1994), willingness to work for lower wages (Daniels, 1994), lack of commitment to the host country (Wickberg, 1994), or abuse of welfare provisions (Spoonley, 1993; Werbner, 1994). Finding ways of welcoming new immigrants without heightening social and economic tensions in the community is a priority.

9.29.7 SUMMARY

What are the solutions to preventing racial violence and hatred? While it is clear that there are no simple answers, the only real solutions lie in dealing with structural problems in society. There needs to be stronger legal sanctions against not only racial violence but also against all forms of racism and racial discrimination in all areas of our society. Comparisons with societies that have less violence and hatred may give some clues as to what needs to be done. In looking at traditional Buddhist societies with very little violence and hatred, Gielen (1995) has suggested that the emphasis on more individual rather than collective freedoms may be part of the problem rather than the solution of the social disorganization seen in Western nations, such as the United States.

Some researchers have argued that psychologists, teachers, social workers, and others operating on the individual level can play an important part in finding solutions. Learning more about the nature and causes of racial violence and hatred identifies possibilities for reducing them and preventing their development. Clinical psychologists need to be strong advocates for reducing all forms of violence and should be able to help provide alternatives to social violence. Others have less faith in the ability of psychologists to deal effectively with violence and have argued for solutions at the personal level. Kappeler (1995), for example, stated that there is an urgent need to deal with the numerous individual examples of everyday violence that are condoned by our failure to treat them as a problem. She argued further that using violence to combat violence ultimately is not a solution. Certainly, one or two people inspired to work for nonviolent solutions (e.g., Ghandi, Martin Luther King) have had profound influences on the occurrence of racial violence and hatred.

We need to find ways to celebrate our differences rather than use them as a basis for prejudice, discrimination, and conflict. But we must go further than resolving our interpersonal conflicts or even celebrating our differences. We need to understand that we are interdependent (Aronson, 1992; Singh, 1995) and that true harmony will come only when we realize our need for each other, and that in hating or harming another we are hating and harming ourselves. With regard to those who engage in violent behavior we need to "link them to the rest of humanity, while separating them from their violent acts" (Toch, 1992, p. xiv). The ultimate solution, however, is to go beyond even interdependence and find oneness. Only at the deepest level of ourselves can we discover true oneness; a level where we unite with others at their deepest level. The ways to this are prescribed in the wisdom traditions of all cultures.

9.29.8 REFERENCES

Albó, X. (1994). Ethnic violence: The case of Bolivia. In K. Rupesinghe & M. Rubio (Eds.), *The culture of violence* (pp. 119–143). Tokyo: United Nations University Press.

Aronson, E. (1992). Stateways can change folkways. In R. M. Baird & S. E Rosenbaum (Eds.), *Bigotry, prejudice and hatred: Definitions, causes & solutions* (pp. 185–201). Buffalo, NY: Prometheus Books.

Baird, R. M., & Rosenbaum, S. E. (1992a). *Bigotry, prejudice and hatred: Definitions, causes & solutions.* Buffalo, NY: Prometheus Books.

Baird, R. M., & Rosenbaum, S. E. (1992b). Introduction. In R. M. Baird & S. E Rosenbaum (Eds.), *Bigotry, prejudice and hatred: Definitions, causes & solutions* (pp. 9–17). Buffalo, NY: Prometheus Books.

Berkowitz, L. (1993). *Aggression: Its causes, consequences and control.* New York: McGraw-Hill.

Berkowitz, L. (1994). Guns and youth. In L. D. Eron, J. H. Gentry, & P. Schlegel (Eds.), *Reason to hope: A psychosocial perspective on violence and youth* (pp. 251–279). Washington, DC: American Psychological Association.

Berrill, K. (1992). Anti-gay violence and victimization in the United States: An overview. In G. M. Herek & K. T. Berrill (Eds.), *Hate crimes: Confronting violence against lesbians and gay men* (pp. 19–45). Newbury Park, CA: Sage.

Bradby, H. (1996). Introduction. In H. Bradby (Ed.), *Defining violence* (pp. 1–8). Aldershot, UK: Avebury.

Carter, R. T., & Qureshi, A. (1995). A typology of philosophical assumptions in multicultural counseling and training. In J. G. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander (Eds.), *Handbook of multicultural counseling* (pp. 239–262). Thousand Oaks, CA: Sage.

Daniels, R. (1994). The Indian diaspora in the United States. In J. M. Brown & R. Foot. (Eds.), *Migration: The Asian experience* (pp. 83–103). New York: St. Martin's Press.

Dovidio, J. F., & Gaertner, S. L. (1986a). Prejudice, discrimination, and racism: Historical trends and contemporary approaches. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 1–34). Orlando, FL: Academic Press.

Dovidio, J. F., & Gaertner, S. L. (1986b). *Prejudice, discrimination, and racism.* Orlando, FL: Academic Press.

Eron, L. D., & Slaby, R. G. (1994). Introduction. In L. D. Eron, J. H. Gentry, & P. Schlegel (Eds.), *Reason to hope: A psychosocial perspective on violence and youth* (pp. 1–22). Washington, DC: American Psychological Association.

Eron, L. D., Gentry, J. H., & Schlegel, P. (Eds.) (1994). *Reason to hope: A psychosocial perspective on violence and youth.* Washington, DC: American Psychological Association.

Gadlin, H. (1994). Conflict resolution, cultural differences, and the culture of racism. *Negotiation Journal, 10,* 33–47.

Galtung, J. (1990). Cultural violence. *Journal of Peace Research, 27,* 291–305.

Geertz, C. (1962). The growth of culture and the evolution of mind. In J. M. Schon (Ed.), *Theories of the mind.* Glencoe, IL: The Free Press.

Gielen, U. P. (1995). Traditional Buddhist Ladakh: A society at peace. In L. L. Adler & F. L. Denmark (Eds.), *Violence and the prevention of violence* (pp. 191–203). Westport, CT: Praeger.

Grimshaw, A. D. (1969a). Factors contributing to color violence in the United States and Britain. In A. D. Grimshaw (Ed.), *Racial violence in the United States* (pp. 254–269). Chicago: Aldine.

Grimshaw, A. D. (1969b). *Racial violence in the United States.* Chicago: Aldine.

Hamburg, D. A. (1993). Preventing contemporary intergroup violence. *Report of the President/1993.* NY: Carnegie Corporation.

Hammond, W. R., & Yung, B. R. (1994). African

Americans. In L. D. Eron, J. H. Gentry, & P. Schlegel (Eds.), *Reason to hope: A psychosocial perspective on violence and youth* (pp. 105–118). Washington, DC: American Psychological Association.

Herek, G. M., & Berrill, K. T. (Eds.) (1992a). *Hate crimes: Confronting violence against lesbians and gay men*. Newbury Park, CA: Sage.

Herek, G. M., & Berrill, K. T. (1992b). Introduction. In G. M. Herek & K. T. Berrill (Eds.), *Hate crimes: Confronting violence against lesbians and gay men* (pp. 1–10). Newbury Park, CA: Sage.

Hill, H. M., Soriano, F. I., Chen, S. A., & LaFromboise, T. D. (1994). Sociocultural factors in the etiology and prevention of violence among ethnic minority youth. In L. D. Eron, J. H. Gentry, & P. Schlegel (Eds.), *Reason to hope: A psychosocial perspective on violence and youth* (pp. 59–97). Washington, DC: American Psychological Association.

Kappeler, S. (1995). *The will to violence: The politics of personal behaviour*. Cambridge, UK: Polity Press.

Levin, J., & McDevitt, J. (1993). *Hate crimes: The rising tide of bigotry and bloodshed.* New York: Plenum.

MacGregor, F. E., & Rubio, M. (1994). Rejoinder to the theory of structural violence. In K. Rupesinghe & M. Rubio (Eds.), *The culture of violence* (pp. 42–58). Tokyo: United Nations University Press.

MacKinnon, C. A. (1993). *Only words.* Cambridge, MA: Havard University Press.

Nagel, J. (1995). Resource competition theories. *American Behavioral Scientist, 38,* 442–458.

National Research Council (1993). *Understanding and preventing violence*. Washington, DC: National Academy Press.

Peach, C. (1994). Three phases of South Asian emigration. In J. M. Brown & R. Foot (Eds.), *Migration: The Asian experience* (pp. 38–55). New York: St. Martin's Press.

Pepler, D. J., & Slaby, R. G. (1994). Theoretical and developmental perspectives on youth and violence. In L. D. Eron, J. H. Gentry, & P. Schlegel (Eds.), *Reason to hope: A psychosocial perspective on violence and youth* (pp. 27–58). Washington, DC: American Psychological Association.

Rupesinghe, K. (1994). Forms of violence and its transformation. In K. Rupesinghe & M. Rubio (Eds.), *The culture of violence* (pp. 14–41). Tokyo: United Nations University Press.

Singh, N. N. (1995). In search of unity: Some thoughts on family-professional relationships in service delivery systems. *Journal of Child and Family Studies. 4,* 3–18.

Singh, N. N. (1996). Cultural diversity in the 21st century: Beyond *e pluribus unum*. *Journal of Child and Family Studies, 5,* 121–136.

Siraj-Blatchford, I. (1994). *The early years: Laying the foundations for racial equality*. Stoke-on-Trent, UK: Trentham Books.

Solomos, J. (1993). *Race and racism in Britain* (2nd ed.). London: Macmillan.

Spoonley, P. (1993). *Racism and ethnicity* (2nd ed.). Auckland, New Zealand: Oxford University Press.

Staub, E., & Rosenthal, L. H. (1994). Mob violence: Cultural-societal sources, instigators, group processes and participants. In L. D. Eron, J. H. Gentry, & P. Schlegel (Eds.), *Reason to hope: A psychosocial perspective on violence and youth* (pp. 281–313). Washington, DC: American Psychological Association.

Takooshian, H., & Verdi, W.M. (1995). Assessment of attitudes toward terrorism. In L. L. Adler & F. L. Denmark (Eds.), *Violence and the prevention of violence* (pp. 33–41). Westport, CT: Praeger.

Toch, H. (1992). *Violent men: An inquiry into the psychology of violence*. Washington, DC: American Psychological Association.

Toplin, R. B. (1994). Violence and culture in the United States. In K. Rupesinghe & M. Rubio (Eds.), *The culture of violence* (pp. 237–256). Tokyo: United Nations University Press.

Tsytsarev, S. V., & Callahan, C. V. (1995). Motivational approach to violent behavior: A cross-cultural perspective. In L. L. Adler & F. L. Denmark (Eds.), *Violence and the prevention of violence* (pp. 3–10). Westport, CT: Praeger.

Virdee, S. (1995). *Racial violence and harassment*. London: Policy Studies Institute.

Vontress, C.E. (1995). The breakdown of authority: Implications for counseling young African American males. In J. G. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander (Eds.), *Handbook of multicultural counseling* (pp. 457–473). Thousand Oaks, CA: Sage.

Werbner, P. (1994). Renewing an industrial past: British Pakistani entrepreneurship in Manchester. In J. M. Brown & R. Foot (Eds.), *Migration: The Asian experience* (pp. 104–130). New York: St. Martin's Press.

Wickberg, E. (1994). The Chinese as overseas migrants. In J. M. Brown & R. Foot (Eds.), *Migration: The Asian experience* (pp. 12–37). New York: St. Martin's Press.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.01

Introduction to Diversity in Clinical Psychology

HARRY C. TRIANDIS
University of Illinois, Urbana, IL, USA

10.01.1 INTRODUCTION 2

10.01.2 DEFINITIONS 2

- *10.01.2.1 Culture* 2
- *10.01.2.2 Cultural Syndromes* 3
- *10.01.2.3 Identity* 3
- *10.01.2.4 Acculturation* 3
- *10.01.2.5 Ecology* 4
- *10.01.2.6 Psychological Processes* 4
- *10.01.2.7 Cultural Distance* 4
- *10.01.2.8 Culture Shock* 4
- *10.01.2.9 Ethnocentrism* 4
- *10.01.2.10 Racism* 4
- *10.01.2.11 Prejudice* 5
- *10.01.2.12 Emic and Etic Aspects of Culture* 5
- *10.01.2.13 An Integrative Example of These Definitions* 5

10.01.3 WORLD VIEWS 6

- *10.01.3.1 Cultural Complexity* 6
- *10.01.3.2 Tightness–Looseness* 6
- *10.01.3.3 Collectivism* 7
- *10.01.3.4 Individualism* 7
- *10.01.3.5 Horizontality* 7
- *10.01.3.6 Verticality* 7
- *10.01.3.7 Active–Passive* 7
- *10.01.3.8 Honor* 8
- *10.01.3.9 Universalism–Particularism* 8
- *10.01.3.10 Diffuseness–Specificity* 8
- *10.01.3.11 Ascription–Achievement* 8
- *10.01.3.12 Instrumental–Expressive* 8
- *10.01.3.13 Value Orientations* 8
- *10.01.3.14 Dionysian vs. Apollonian Cultures* 9

10.01.4 METHODOLOGIES FOR THE STUDY OF WORLD VIEWS 9

- *10.01.4.1 Some General Strategies* 11
- *10.01.4.2 Good Theory Can Eliminate Some Rival Hypotheses* 12
- *10.01.4.3 Methods that Can Establish Measurement Equivalence Across Cultures* 12
- *10.01.4.4 Emics and Etics in Cross-cultural Research* 12
- *10.01.4.5 How to Use Emic Measurements of Etic Constructs* 14
  - *10.01.4.5.1 General recommendations* 15
- *10.01.4.6 Examples of Multimethod Measurements* 15
  - *10.01.4.6.1 Ethnographic work* 15
  - *10.01.4.6.2 Establishing shared cognitions* 16
  - *10.01.4.6.3 Conclusions* 17

10.01.4.7 Translations 18
10.01.4.8 Ethics of Cross-cultural Studies 19
10.01.4.9 Ethnocentric and Androcentric Bias of Researchers 19
10.01.4.10 Summary of Criteria of Good Cross-cultural Research 20
10.01.5 ACCULTURATION 20
10.01.6 IDENTITIES OF AFRICAN- AND EUROPEAN-AMERICANS AS AN EXAMPLE OF THE IDENTITIES OF CLINICIANS AND CLIENTS 21
10.01.7 METHODS FOR THE DEVELOPMENT OF AN UNDERSTANDING OF THE CULTURE OF THE CLIENT 23
10.01.7.1 Methods of Culture Learning 24
10.01.7.1.1 Culture-general vs. culture-specific training 25
10.01.7.1.2 Self-insight 26
10.01.7.1.3 Experiential training 26
10.01.7.1.4 Exposure to many local cultures 26
10.01.7.1.5 Field trips 26
10.01.7.1.6 Culture assimilators or intercultural sensitizers 26
10.01.7.1.7 Behavior modification cross-cultural training 29
10.01.7.1.8 Some summary points about training 29
10.01.8 FUTURE DIRECTIONS 30
10.01.9 SUMMARY AND CONCLUSIONS 30
10.01.10 REFERENCES 30

10.01.1 INTRODUCTION

Clinical psychologists deal with many kinds of diversity. While diversity due to culture may be the most imporant kind that affects their work, it is useful to remember that diversity in social class, sexual orientation, gender definitions, disabilities, political perspectives, esthetic preferences, religion, and the definition of what is "true" may also affect their practice.

Diversity may be defined as the condition of interacting with another person who is perceived to be "different." It is a condition which is increasingly more frequent, since most societies are multicultural and a number of factors favor interactions among the cultures. In the US the clients of clinical psychologists are becoming more different from the white middle-class which was the usual clientele of these professionals a few years ago.

It is impossible to deal with all the aspects of diversity in one chapter, so the focus will be on culture. However, some of the other kinds of diversity (e.g., in religion or social class) do interact with cultural differences, and some of the experience which clinicians have acquired in dealing with these other kinds of diversity may prove useful in dealing with cultural diversity.

This chapter will begin with definitions of concepts such as culture, and the relevance of ethnocentrism, prejudice, and racism in making judgments about the behavior of other humans. It will also examine the terminology for dealing with what is culture-common (etic) and culture-specific (emic). Section 10.01.3 will describe some of the diversity of world views and values, and will examine methodologies which can provide reliable information concerning such world views. The focus will be on helping clinicians discriminate reliable from unreliable cross-cultural studies, so that they can begin developing a body of knowledge about cultural differences. Then, the chapter will examine the relationship of acculturation to varieties of cultural identity (in both the majority and minority culture). Finally, it will describe methods and procedures that clinicians can use to develop more detailed understandings of other cultures, and to improve the effectiveness of their behavior when interacting with diverse clients.

10.01.2 DEFINITIONS

10.01.2.1 Culture

Culture is to society what memory is to individuals (Kluckhohn, 1954, p. 967). It includes all the things which have "worked" in the history of an interacting group of humans, and have become unstated assumptions, standard operating procedures, standards of perceiving, judging, deciding, and acting. Cultural anthropologists (e.g., Kroeber & Kluckhohn, 1952) have developed more than 100 definitions of "culture," such as that culture is the human-made part of the environment (Herskovits, 1955). Such a definition is too broad, but it does have the advantage that we can ask: Is this particular element human-made? For instance, is this belief human-made? Then it is an aspect of culture. The Herskovits' definition can become more useful if we distinguish material culture (tools, roads, houses) from subjective culture (categorizations, beliefs, attitudes, self-definitions, role definitions, norms, values (Triandis, 1972).

Most contemporary anthropologists (e.g., Shweder & Levine, 1984) agree that culture is a system of *shared* meanings. Hence my favorite definition:

> Culture is a set of human-made objective and subjective elements that in the past have increased the probability of survival and resulted in satisfactions for the participants in an ecological niche, and thus became *shared* among those who could communicate with each other because they had a common language and they lived in the same time period and place. (Triandis, 1994, p. 22)

Note that this definition, in addition to focusing on what is shared, specifies that a common language is important, and communication can take place because people live in the same time period and geographic region. Thus, in principle, people who live next door to each other but speak different languages (say, in Los Angeles, Hispanics and Thais), or during different time periods (say, in Los Angeles in 1950 and in 1996), or in different geographic regions (say, in Texas and New York) may have different cultures. Physicians have their own language in the hospital, and thus physicians-on-the-job may constitute a culture, by this definition. According to this definition the number of cultures is enormous.

Humans regardless of culture have many common attributes. All humans have language, food habits, art, myths, religious practices, family structures, economic systems, "truth," government, war, kinship, shelter, training systems, hygiene, and incest taboos (Brown, 1991). However, note that these categories are extremely broad. Take an easy one to think about: food habits. What people eat, when, where, with whom, and to what effect are clearly diverse. If we are to pay attention to that level of detail, it is obvious that the diversity of cultures is enormous.

10.01.2.2 Cultural Syndromes

Shared patterns of beliefs, attitudes, self-definitions, norms, roles, and values are sometimes organized around a common theme. When they are and found among members of a culture, they constitute a cultural syndrome. The logic of shared elements of subjective culture allows us to identify such syndromes (Triandis, 1996). For example, we can examine the response distributions obtained from samples in a particular culture, and if an arbitrary 90% of the participants give a response on the same side of the neutral point, on a nine-point scale that measures an element of subjective culture, then that element is an aspect of the cultural syndrome. Another method (see Triandis, Bontempo, Leung, & Hui, 1990) is to form (say 50) triads of representatives of each culture, and ask them to reach a group agreement on an element of subjective culture, for example, "Do you agree that widows should not eat chicken?" Agreement in less than 60 seconds, among say 90% of the triads, would indicate that the belief is widely shared, and thus it is an element of the culture syndrome. In sum, both the time to reach agreement and the percentage of triads which reach agreement can be used as criteria for determining the presence of an element of a cultural syndrome. When the cultural elements, identified by such procedures, which represent attitudes, beliefs, norms, role-definitions, self-definitions, and values are interrelated, we have defined empirically the cultural syndrome.

10.01.2.3 Identity

People make statements that include the words "me," "mine," "myself," "I," and the like. All these statements taken together constitute the *self*. An important aspect of the self concerns statements about belonging to some group. For example "I am a Korean" specifies a Korean identity. "I am a bridge-player" is also an identity. Brewer (1991) has argued that each individual experiences forces toward differentiation from groups as well as assimilation to groups. These opposing forces are in equilibrium at the point of "optimal distinctiveness" which is characteristic of each person. This point is determined by cultural norms, individual socialization, and recent experiences with the group. For instance, in traditional cultures, which tend to be collectivist, the forces of assimilation are stronger than the forces of differentiation, and thus the optimal distinctiveness point is closer to the group than in modern cultures.

10.01.2.4 Acculturation

Some individuals, especially in multicultural cities, are exposed to numerous cultures simultaneously. They may use more than one identity, for example, "I am Mexican" and "I am American." The more elements of one culture are attached to an identity (e.g., Since I am Mexican I speak Spanish most of the time), the more influential is that identity in determining behavior. An individual may be more or less bicultural, include elements from both cultures in an identity, assimilated, include mostly elements of another culture in an identity, separated, that is, even though the individual lives in an environment which includes many cultures the identity is separate and distinct and monocultural, or alienated, may reject all cultures.

10.01.2.5 Ecology

Ecology refers to where people live. It consists of objects, resources, the geography of the environment, and the ways of making a living and surviving in that environment. Ecology provides schedules of reinforcement (Skinner, 1981), which shape both the elements of subjective culture and the behavior of those who live in these ecologies. Thus, culture emerges from the ecology.

10.01.2.6 Psychological Processes

Clinicians are, of course, focusing on psychological processes, such as personality, behavior patterns, and the beliefs, attitudes, and values of their clients. It is obvious from the above definitions that culture influences many of these psychological processes. Poor clinical analyses occur when the clinician and the client come from different cultures, because the very same behavior may have different meanings in the two cultural systems and very frequently these meanings influence important judgments, such as diagnoses.

Take one simple example: in collectivist cultures, such as those found in the Far East, communications that are unclear and indirect which save someone's face are expected and are very common. In individualistic cultures, such as in the West, people are expected to provide "straight talk" and not "beat around the bush." Similarly, a lie is morally reprehensible in the West much more than in the East. Authenticity (Trilling, 1972) is valued in the West, and that includes "telling it as it is" and telling the truth. Thus, the same behavior can be interpreted differently in the two kinds of cultures. Indirect communications may be seen as "tactful and polite" and a lie can be seen as saving face in the East, while the same behaviors may be seen as "confused and confusing" and reprehensible in the West. The interpretation is an aspect of an integrated conceptual system, and is often so compelling that it is difficult for either party to easily use the other's way of thinking.

10.01.2.7 Cultural Distance

Cultural distance is an important concept in understanding diversity, and the way it affects human relationships. Cultural distance can reflect differences in (i) language (e.g., language families, such as the client speaks a tonal language and the clinician an Indo-European language), (ii) family structure (say, monogamous vs. polygamous marriages can create distance), (iii) religions (e.g., animist and Christian religions), (iv) wealth and life style (e.g., the difference between a wealthy jetsetter and a member of a culture of hunters and gatherers), (v) values (e.g., the difference between conservative, very traditional values and self-actualizing, hedonistic values).

10.01.2.8 Culture Shock

Triandis and Gelfand (1994) presented a model that links cultural distance and culture shock (Oberg, 1960). The larger the cultural distance the larger the culture shock. The model posits that cultural distance leads to low perceived similarity. When people perceive each other as dissimilar and are also in contact, the relationship is tense, and punishing, resulting in distortions such as extreme stereotypes, inaccurate attributions (e.g., clinician and client attribute different causes to the same behavior), which lower the sense of control (ability to act so as to obtain positive outcomes from relationships). A low sense of control results in culture shock. The model also includes other variables, which need not concern us here.

10.01.2.9 Ethnocentrism

Most humans are ethnocentric (Brewer & Campbell, 1976; Campbell & Levine, 1968; Triandis, 1994). Ethnocentrism follows from the fact that most humans are only exposed to their own culture. Only humans who live in multicultural environments, or have traveled extensively, are moderately ethnocentric. The extent our culture influences our perceptions of the behavior of others is large and subtle. We are usually not aware of our ethnocentrism. Those who have been exposed to other cultures learn to appreciate aspects of these cultures which are not included in their own cultures, and thus become less ethnocentric. Clinicians must control their ethnocentrism by learning to appreciate other cultures. Yet that is especially difficult when the cultural distance between their culture and the client's culture is large. In that case it may be best to admit that they are not able to help the client, and it is desirable to find colleagues whose cultural distance is smaller and who might be able to help the client.

10.01.2.10 Racism

Racism is a special case of ethnocentrism, where the standards of one's own racial group are used to judge other racial groups.

10.01.2.11 Prejudice

Prejudice is judgment of the attributes and behavior of members of another culture on the basis of negative preconceived ideas and inaccurate stereotypes. Such judgments are usually associated with negative emotional reactions toward members of the other culture, as well as with self-instructions to keep some social distance from the members of the other culture (rejection of positive and endorsement of negative behavioral intentions toward the members of the other culture). People who are very ethnocentric tend to be prejudiced. Even people who score as nonprejudiced on attitude scales, when responding to ethnic stimuli tachistoscopically, so that they do not have the opportunity to "control" their prejudice, show themselves to use unfavorable stereotypes. In short, the unprejudiced have learned to control their prejudice and not show it to researchers, but the negative stereotypes are nevertheless present in their cognitive system (Devine, 1989).

10.01.2.12 Emic and Etic Aspects of Culture

Phon*etics* deals with sounds which occur in all languages. Phon*emics* deals with sounds which occur in only one language. This led the linguist Pike (1967) to coin the words *etics* and *emics* to refer to the culture-general and culture-specific elements of culture. *Etics* reflect constructs which apply to phenomena that occur in all cultures. *Emics* are constructs which occur in only one culture.

For example, in all cultures ingroup members (family, tribe, co-workers, co-religionists) are treated better than outgroup members (enemies, strangers, outsiders). That is an etic. However, there are also emic patterns. For example, when Hispanics and non-Hispanics were presented with several hundred situations and asked if they expected people in such a situation to use positive (e.g., admire, respect, support) or negative (e.g., criticize, dominate, fight with) behaviors, there was a response pattern which was called the "*simpatia* cultural script" (Triandis, Marin, Lisansky, & Betancourt, 1984) that emerged when the Hispanic responses were compared to the non-Hispanic responses to the same situations. Specifically, Hispanics compared to non-Hispanics expected more positive behaviors and fewer negative behaviors in a wide range of situations. They expected people to try to be *simpatico*, that is, to be nice and pleasant, even in situations where non-Hispanics expected neutral or negative behaviors. This is similar, but not identical, to the greater importance of face in collectivist than in individualistic cultures that has been discussed by East Asian social scientists (e.g., Ting-Toomey, 1988). The simpatia cultural script is an emic.

In other words, when dealing with ingroup members or even acquaintances, Hispanics tend to be more positive in their behaviors than non-Hispanics. The practical significance of this pattern is that, relative to non-Hispanics, Hispanics will expect a clinician to be more positive in situations that ordinarily will be ambiguous or call for somewhat negative behaviors. For instance, criticism or other negative behaviors may be more upsetting to Hispanics than to non-Hispanics.

For another example of an emic motif, consider the *amae* pattern found in Japan (Doi, 1986). Japanese often presume that they may depend on others, the way a child depends on a parent. Mutual dependence is an emic aspect of that culture.

The significance of the emic and etic distinction is that it alerts us to which specific cultures may have idiosyncratic ways of cutting the pie of experience. If we use only etics, we miss some of that information and do not identify many cultural differences (Marin & Marin, 1991).

There are ways to combine emics and etics (see Triandis, 1992), making our measurements culture-sensitive but also equivalent across cultures. These techniques will be discussed in Section 10.01.4.5. Many techniques specific to cultural studies have been identified (see Triandis & Berry, 1980). Some of these use methods (e.g., Szalay, 1970; Szalay & Deese, 1978) such as word association which are also used by some clinical psychologists.

10.01.2.13 An Integrative Example of These Definitions

The ethnocentrism mentioned above is an etic. It is accentuated by two other phenomena, which are also probable etics: (i) we have a tendency toward *naive realism* which limits our capacity to appreciate the extent to which our construals are subjective (Robinson, Keltner, Ward, & Ross, 1995); and (ii) we have a tendency toward *false consensus*, which is to think that other humans agree with our positions more than is true, and disagree with our position less than is true (Krueger & Clement, 1994).

In short, we tend to think that the way we see the world is both valid and universal. Our culture provides the "lenses" for seeing the world in a particular way and that way of seeing is so obvious that it is not questioned.

Consider this example. In Orissa, India, most of the population believes that widows must not eat chicken (Shweder, Mahapatra, & Miller, 1990). When asked if this behavior should be universal they say: "Of course. It is a great sin for widows to eat chicken." When told that Americans do not believe this, they look down upon Americans, and explain this "moral deficiency" by noting that America is a young country which has not yet reached the level of moral maturity found in India.

Now, consider what happens when Illinois participants are asked whether widows must not eat chicken. They say that this belief is silly. When asked if this rule should be universal, they object vehemently. When told that people in Orissa, India, strongly believe that widows must not eat chicken, they look down upon these Indians, and point out that they are not sufficiently developed to have "correct views."

Do you see the ethnocentrism in *both* perspectives? The Indian view stems from the basic assumption that people are interdependent. Married individuals are supposed to be linked to each other for ever. For a widow to eat chicken is a sin because they believe that eating chicken makes one sexually aroused, and such arousal will result in the widow having sexual relations with someone, and thus breaking the eternal bond with her husband. Note that cultures often have sets of beliefs which are supportive of each other.

Now consider the American view. The basic assumption is consistent with American individualism (Triandis, 1995) which assumes that people are autonomous entities. Widows can do their own thing. If Indians have an idea that is different from the American idea, it is because they are not sufficiently developed!

Thus, both cultures conclude that their views are superior and the views of the other culture are inferior. Ethnocentrism leads to prejudice, and attempts to impose the subjective culture of one's own culture on other cultural groups.

This is not the place to debate the merits of individualism and collectivism. Hofstede (1980) linked these concepts with many ecological variables. Triandis (1995) has suggested that individualism is associated with high levels of achievement, creativity, self-actualization, and democracy, but also with high levels of crime, divorce, and child abuse. Collectivism is associated with high levels of social support, cooperation, interpersonal sensitivity, and pleasantness in social relationships, but also with extreme conformity, low creativity, and ethnic cleansing. In short, both cultural patterns have both positive and negative elements, and it is natural for most people from all types of cultures to prefer their cultures. As scientists we can examine the links of ecology, culture, and social pathology, but preferences for social pathology are matters of taste, not scientific judgment!

10.01.3 WORLD VIEWS

Many cultural syndromes have been identified thus far. The exact number of syndromes is not known, but a guess is that a score may be sufficient to describe the most important cultural differences. Here we present some of the most important syndromes.

10.01.3.1 Cultural Complexity

Simple cultures such as hunters and gatherers have few roles, simple social structures, few levels of social and political stratification, simple religious beliefs, simple understandings of the nature of the world, few domains of esthetic activity, and simple methods of economic exchange (e.g., barter). Specific simple cultures may be very complex in one domain, such as people may have an extensive knowledge of the history of their ancestors, as is found among Australian aborigines, but in most domains they are simple. Many attributes of culture are associated with cultural simplicity. For example, simple cultures are characterized by small communities (often bands of 50 individuals), the use of barter, walking as the only means of transportation, and so on.

Complex cultures, such as information societies, have many roles, for example, a quarter of a million occupations in the *Dictionary of Occupational Titles*, complex social structures, for example, General Motors, complex social and political stratification (e.g., federal, state, and local governments), a multiplicity of religious beliefs, a complex understanding of the nature of the world, for example, modern science, multiple types of esthetic activities, and numerous methods of economic exchange, such as, money, bonds, stock, certificates of deposit, etc. They include mostly urban social patterns. Again, in one domain there might be considerable simplicity, such as a family structure that consists of only a mother and her children, but on the whole there is complexity in most domains.

In complex cultures subjective cultures also tend to be more complex, for example, one finds large vocabularies (100 000+ terms in English) and complex ways of classifying experience, as well as complex beliefs.

10.01.3.2 Tightness–Looseness

Some cultures have (i) many rules, norms, and detailed specifications about how one is to act, and (ii) members who react quite strongly (e.g., criticize or even kill those who deviate from the norms) to minor deviations from the rules, norms, and "proper behavior" (Triandis, 1989, 1994). Such cultures are called "tight." Other cultures are loose, that is, have few rules, norms, and do not specify how one is to behave in very many situations, and when a person deviates from these few rules, norms, or conceptions about proper behavior, people tolerate such deviations.

Tight cultures provide criticism and severe sanctions (e.g., 50 lashes) for deviation; loose cultures criticize only major deviations and provide mild sanctions (e.g., a fine) for most deviations.

Tightness and looseness are domain specific. For example, a culture may be very tight about the use of time, but be rather loose about the way one chooses friends. For example, the US, in 1996, is very tight about writing bad checks and rather loose about who one lives with. However, across domains, tight cultures tend to be tight, and loose cultures tend to be loose.

Japan is an example of a tight culture, and Thailand an example of a loose culture. The US is fairly loose. For example, one extreme incident that occurred in Japan in 1990 was that a teacher killed a student who was two minutes late for class. Japanese opinion was divided about this teacher, most arguing that he was "deranged" even though they understood his efforts to impose discipline. However, some psychopathology is an extreme expression of a cultural norm (Draguns, 1990)!

Rural Thailand is loose. An employee may come to work, or not come to work, and no explanations are offered (Phillips, 1965). After living in Thailand for several months, Phillips (1965) had difficulty understanding "the mystifying American capacity to conform to the expectations of others" (p. 204).

10.01.3.3 Collectivism

The collectivism syndrome has been studied more extensively than most of the other syndromes (Kim, Triandis, Choi, Kagitcibasi, & Yoon, 1994; Markus & Kitayama, 1991; Triandis, 1990, 1995). It has received a central position in most reviews of cultural differences in social behavior (e.g., M. Bond & Smith, 1996). Triandis (1994) has speculated that maximum collectivism occurs in simple tight cultures, such as theocracies. He also argued (Triandis, 1995) that in collectivist cultures individuals define themselves as aspects of groups (i.e., I am a member of ...), place the goals of their collectives ahead of their personal goals, behave under the influence of norms and roles imposed by their groups, and relate to ingroup members by paying attention to the needs of others rather than to their own needs. Many East Asian societies are collectivist, as are many traditional societies in Africa and Latin America. In the US, Hispanics and Asians tend to be collectivists.

10.01.3.4 Individualism

Triandis (1994) speculated that this pattern is maximal in complex, loose cultures, and is characterized (Triandis, 1995) by individualists perceiving themselves as autonomous from their groups, giving priority to their personal rather than to their ingroup's goals, behaving according to their attitudes and feelings of enjoyment, and dealing with members of their ingroups by paying attention to the "profit" and "loss" incurred from the relationship, as specified by exchange theory (Thibaut & Kelley, 1959). Many Western societies are individualistic, especially the US, Australia, and Britain.

10.01.3.5 Horizontality

Cultures that are egalitarian, such as the Eskimos, and tend to emphasize tit-for-tat, one person one vote and the like are horizontal. In such cultures, people do not share or distribute resources according to need, contributions (equity), or status, but equally. There is little social stratification; most people carry out most of the tasks of the society.

10.01.3.6 Verticality

The cultural pattern of verticality emphasizes hierarchy. It is similar to Hofstede's (1980) "power distance," a cultural pattern in which people at the top of the power hierarchy seem to be very distant from people at the bottom of the hierarchy. Resources are divided according to status. There is much deference for those who have status. High levels of stratification are present.

10.01.3.7 Active–Passive

This syndrome was described by Diaz-Guerrero (1979) when he compared US and Mexican data. Those who are active, such as Americans, thrive on competition, action, and emphasize self-fulfillment. They try to change the world, rather than change themselves to fit into the world. Those who are passive, are cooperative, emphasize the experience of being, and tend to change themselves to fit into situations.

10.01.3.8 Honor

Honor is a narrow syndrome, found in environments in which property is mobile, and thus individuals develop attitudes, values, and behaviors that are designed to scare others, so as to ensure that their property is not taken away from them. For example, cattle are more easily stolen than agricultural property, so that people in cattle cultures are fierce, aggressive, hypersensitive to affronts, stress defense of one's honor, and socialize their children to react to challenges more than occurs in agricultural societies (Cohen & Nisbett, 1994; Nisbett & Cohen, 1996). Several kinds of studies support the view that the American South is high on this cultural syndrome. Cohen, Nisbett, Bowdle, and Schwartz (1996) reported three experiments where University of Michigan students from the North or South of the US were insulted by a confederate. They reported that Southerners were more likely to think that their masculine reputation was threatened, became more upset (as shown by the rise of cortisol levels), became more physiologically primed for aggression (as shown by testosterone levels), were more cognitively primed for aggression, and more likely to engage in aggressive and dominant behavior than students from the North. Cohen (1996) showed that this cultural syndrome was reflected in social policies. An analysis of laws relating to guns, defense of self and home, corporal punishment, capital punishment, and attitudes toward foreign policy issues showed greater approval of violence in the South than in the North.

While the Nisbett and Cohen research focuses on the higher levels of honor found in the American South than the North, extreme emphasis on honor can be found also in most of the Mediterranean societies, for example, Corsica, the Serbs. Campbell's (1964) ethnography of a culture of Northern Greece (the Karakatzanoi) provides details of this syndrome.

10.01.3.9 Universalism–Particularism

Parsons and Shils (1951) contrasted cultures where people categorize other people on the basis of universal criteria, and cultures in which people categorize other people by taking into account their particular characteristics and the unique situation within which they are interacting with them.

10.01.3.10 Diffuseness–Specificity

Parsons and Shils (1951) contrasted cultures in which people respond in a diffuse, holistic manner (e.g., I do not like your report implies I do not like you) with cultures where people respond with specificity.

10.01.3.11 Ascription–Achievement

Parsons and Shils (1951) discussed cultures in which people judge others primarily on the basis of ascribed attributes (e.g., sex, race, family membership) with those which use achieved attributes (e.g., has received this award).

10.01.3.12 Instrumental–Expressive

Parsons and Shils (1951) contrasted cultures in which most interaction has instrumental purposes (get task done) with cultures in which most interactions are expressive (friendship, enjoyment).

10.01.3.13 Value Orientations

Kluckhohn and Strodtbeck (1961) identified several value orientations which may well be the bases of additional cultural syndromes. Specifically, they identified societies where innate human nature is assumed to be (i) evil, neutral, or good; (ii) mutable or immutable. When human nature is assumed to be evil, many rules are introduced to control people and make them behave appropriately; when it is assumed to be good there is a tendency towards minimal government; when it is assumed to be mutable institutions are created (e.g. educational systems) which can change them in the appropriate direction, but when they are assumed to be immutable there is little support for such institutions.

(iii) Humans are assumed to be subjugated to nature, in harmony with nature, or allowed to master nature. In cultures which emphasize mastery over nature, beliefs, attitudes, values, and behaviors favor changes of the environment. Where harmony is emphasized, there is more concern about not changing the environment (e.g., the "greens" of the world). When subjugation is emphasized there are many beliefs, attitudes, and behavior showing reverence for nature (e.g., in many Native American cultures).

(iv) Varieties of time orientation, such as emphases on the past, present, or future. Time orientation also differs because some societies use linear time (e.g., the West) and others use circular time (e.g. the ancient Mayans). Monochronic use of time (i.e., people can carry out only one conversation at a time) or polychronic use of time (i.e., people can carry out several conversations at the same time) are also important cultural differences in the use of time.

(v) The emphasis is on doing, being (the experience), or becoming (growing, changing, improving). For example, in doing cultures, such as the US, a person's worth is measured by a record of accomplishments, income, previous actions. In cultures which emphasize experiences, a person's worth depends on the experiences that the person has had. In cultures which emphasize becoming, the focus is in how much and how fast the person is changing.

10.01.3.14 Dionysian vs. Apollonian Cultures

Again this is a narrow syndrome focused on the way people are expected to deal with emotions. In Dionysian cultures they are taught to express emotions. Loud speech is expected. Beliefs that one should state a view by using extreme language are common (e.g., as found among some Arabs), and attitudes such as the positive evaluation of touching, kissing, close positioning of the bodies, looking into the eye, and other behaviors are often associated with this pattern.

The Apollonian pattern requires people to control their emotions (e.g., East Asians), to stand far from others, not touch, not look into the eye. Corresponding beliefs, attitudes, and values support these behaviors. For example, it is impolite, and an indication of low social class, to speak loudly. Related to the above pattern is: high contact vs. low contact cultures, which reflects high or low level of voice, eye contact, frequency of gestures, touching, distance between the bodies. Interrelationships of the syndromes. Since the concept of cultural syndromes is relatively new, there has not been much research on their number or a clear indication of how many syndromes there are or how they are interrelated.

In addition, there will be emic patterns, which will characterize each individual culture. To identify cultural syndromes it is necessary to do research, based on the fact that syndromes reflect shared psychological processes (see Triandis, 1996) to see exactly what cultural elements are to be included in each cultural syndrome. We need to develop reliable and valid methods for the measurement of cultural syndromes, and then correlate these measurements to discover how the syndromes are related to each other.

10.01.4 METHODOLOGIES FOR THE STUDY OF WORLD VIEWS

The general perspective of this section is not to teach the reader how to do cross-cultural research, but how to tell the difference between a "good" and a "bad" study. I assume that clinicians read the literature, but the literature includes many "bad" studies. The most common error is to take an instrument that has been developed with middle-class mainstream undergraduates and use it with some ethnic group that is quite different from the standardization sample, without any attempt to check the reliability and construct validity of the instrument in this new sample.

Those wishing to do cross-cultural studies can consult a more advanced text. The most complete one is the volume by Triandis and Berry (1980). A shorter volume is by Lonner and Berry (1986). Special issues concerning the equivalence of cross-cultural measurement, and the unstated assumptions that are made by people who collect cross-cultural data are discussed in Hui and Triandis (1985).

Good cross-cultural research is like ordinary social research, only more difficult because of the complexity of the issues of translation, equivalence of measurement, and the like. Those who want to do such research should become familiar with the writings of Campbell (1988), as well as the publications mentioned above.

A central issue in cross-cultural research is whether we have a "real" cultural difference or an "apparent" difference due to an artifact or a "rival hypothesis." For example, it may be that people in one culture respond to a particular method of data collection differently from the way people in another culture respond to it. In such a case we do not have a substantive finding but rather a finding that is best explained by the way people react to the method of data collection. Or, it may be that we have an apparent "race" effect, but careful examination would reveal that it is a "social class" effect.

To establish a "real" cultural difference we need to eliminate all plausible "rival hypotheses" (Malpass & Poortinga, 1986) that may account for the observed difference.

Consider the following specific case. We tested two samples and found a difference in their level of anxiety. What are some of the plausible rival hypotheses?

(i) Anxiety may mean different things in the two cultures. The perceived antecedents of anxiety may be different. For example, in one culture it may be seen as caused by the prospect of lack of food and in the other by the possibility of downsizing the corporations most people work for.

To eliminate this rival hypothesis one needs to study the meaning of constructs, in the relevant cultures, independently of their measurement. For example, one can ask samples of participants from the various cultures: What makes you anxious? When you are anxious, what happens after that? Content analyses of the responses to such questions obtained in the relevant cultures can reveal both similarities and differences in the meaning of the construct. Associations of events with anxiety, semantic differential judgments involving the words "Fear," "Anxiety," and their synonyms, can result in an understanding of the these similarities and differences in the meaning of the concept.

More generally, when we compare two cultures we are comparing entities that differ in a myriad of ways. The conclusion that the difference reflects a particular variable (e.g., anxiety) requires that we "control" all the interactions of anxiety with other plausible variables. Furthermore, we need to sort out which aspect of culture is the relevant one. For example, in a comparison of African-Americans and White-Americans, we need to distinguish a difference due to race/color, *per se*, from differences due to nutrition, social class, neighborhood, historical influences, and so on. If all the observed differences are due to social class, it is scientifically irresponsible to report them as due to race. If both social class and race are relevant, one must report that fact also.

Since any two cultures differ in many ways, it becomes apparent that a two-culture comparison will have almost no scientific value, since in science we want to find relationships among well-defined, clear, reliably measured variables. Suppose we want to check the relationship of nutrition to anxiety. Ideally (assuming we have the funds), we should try to include a large sample of cultures in our study that differ in measurable ways on the nutrition variable. Ideally, the cultures that we select should differ in many ways, except that they should be ranked on one clear variable, nutrition. Then the countless other variables will not be systematically associated with nutrition, and we will be able to say something reliable about the relationship of nutrition to anxiety (Leung, 1989).

Campbell (1988) makes a valid point: the comparison of any two cultures is essentially useless (except for preliminary, hypotheses-generating work). This point requires us to study many cultures (a minimum of two "high" and two "low" on the variable of interest). The more cultures we can include in our analyses, the better. Furthermore, we must select cultures which are similar on the variable of interest but very different on all other variables, because of the phenomenon of cultural diffusion. It is well known that cultural elements are copied in adjacent cultures. If there is a lot of diffusion we do not have two independent cases, but actually only one case.

In addition, we must replicate the study with many different methods, since each method has its own meaning in each culture, but it is unlikely that an unsound hypothesis will be supported with very different methods. It is especially important to replicate the study with both nonreactive (e.g., participant observations) and more reliable (but usually reactive) methods, such as tests or questionnaires.

In sum, one must be skeptical about the equivalence of the meaning of concepts. For example, Diener (personal communication) asked samples in different parts of the world whether they were "happy." One respondent in India said: "I do not know; ask my husband."

In what follows I will identify some of the more obvious additional rival hypotheses which may produce an "apparent" difference when a "real" difference is not present.

(ii) The instructions may not be understood the same way. This is especially likely if the members of one culture have much less experience with a method or task than the members of the other. Related to this variable is the extent the two populations are familiar with the measurement method. For example, multiple choice tests are widely used in the US and a measure which uses such a format can give US subjects an advantage over a population from a country where such tests are not used.

A special test which measures the "understanding of the instructions" should be used routinely to eliminate this rival hypothesis. Familiarity with the measurement format and content has to be equated. This can be done by using formats which are equally familiar (and contents which have been pretested to be equally familiar) in the two cultures.

(iii) The level of motivation to respond to the particular method of the two samples may be different. Americans are likely to be motivated if an experimenter tells them that their "anxiety" will be measured to give an accurate response. They may want to know "their score" and how it compares to the scores of their peers. But in other cultures, where "anxiety" is perceived as a private matter not to be revealed to outsiders, or where it is mandatory to lie to outsiders, or where anxiety is conceived as something shameful, there may be strong motivation to give an inaccurate answer.

Independent measurements of the levels of motivation should be used to eliminate this rival hypothesis.

(iv) The reactions to the experimenter may be different. For example, in some cultures it is against the norms to cooperate with "outsiders." In some cultures it is mandatory to "lie" to outsiders or trick them! In such cultures villagers compete among themselves in telling lies about their culture, and later discuss among themselves what lies they told and laugh at the researcher for having believed their lies. To overcome such difficulties it is necessary to use several researchers, some "insiders," and some "outsiders" to obtain some estimate of the importance of this factor.

(v) The meaning of the test situation is not always the same. For example, in some studies (Bond & Yang, 1982; Bond & Cheung, 1984; Yang & Bond, 1980) the language of the instructions influenced the results. When instructions were given to Hong Kong subjects in Mandarin, Cantonese, or English, the results differed. Apparently, the language of the instructions suggested to the subjects who was interested in the results of the study—the Beijing authorities, the Hong Kong authorities, or the British colonial authorities.

(vi) Response sets differ across cultures (Hui & Triandis, 1989; Triandis, 1972). For example, in some cultures if one is "sincere" one gives a strong answer (e.g., "I strongly agree."), but if one is not sincere, one might use the middle of a Likert scale. In Japan it is arrogant to use an extreme response, and it is appropriate (modest, polite) to use the middle of the scale ("perhaps I agree with it"). In some cultures people only answer the questions which they are absolutely sure of; in others, they answer all the questions.

Marin, Gamba, and Marin (1992) found that Hispanics in the US, relative to non-Hispanics, are more likely to use extreme responses (e.g., I strongly agree) and to show acquiescence (agree with every question). The more acculturated (spent a lot of time in the US, are comfortable when using English) the Hispanics are to the US, the less they do that. In some cultures people are expected to agree when asked questions; in others they are expected to disagree when asked questions by "outsiders."

(vii) The samples, in the relevant cultures, may not have been strictly equivalent. For example, there might have been differences in social class, age, sex, religion, or some other demographic attributes which are the "real" cause of the difference, yet the difference is attributed to language/culture. One should analyze the data separately for each demographic category. Furthermore, the two populations under study may not be stable over time (e.g., sampling in 1960 and assuming that the results apply in 1999 may well be an error). When a population is heterogeneous, sampling in one area and hoping that the results are valid in another area (e.g., sampling in New York and assuming that the results apply to the whole of the US) may be misleading. For that reason, ideally, one should obtain samples in different parts of a country, and from different occupational, sex, and age groups, and check the consistency of the findings across these samples.

(viii) The ethical acceptability of the method may not have been the same in the various cultures. For example, in some cultures some samples (e.g., women) are not supposed to have an opinion, let alone an opinion which is different from their husband's. If one asks about that opinion, people can become embarrassed or angry because the experimenter "dared" to ask such a question. Pretests are needed in which samples are asked after the pretest "what did you think about this method?" It is useful to use a set of scales (such as good vs. bad, active vs. passive, strong vs. weak) for the subjects to rate the test situation itself (was the test situation pleasant or unpleasant?) One of these scales could be moral vs. immoral. If the study appears moral in one culture and immoral in the other, it would be necessary to investigate the matter further, before collecting much data.

It is clear from the above that there are many ways in which a difference may appear. A difference is easy to obtain across cultures, but the question is: Is the difference a substantive finding, or is the apparent difference due to something which is indirectly associated with the measurement? We can never be sure that we "controlled" all rival hypotheses, so any conclusions about cultural differences must remain tentative.

The discussion presented above does not mean that we must reject all forms of measurement. One can control statistically variables which are correlated with the dependent variable of interest. For example, if there is a response set, one way to control it is to "standardize" the data statistically. One method is to standardize the data obtained from a single participant (within participant standardization), thus eliminating individual differences, but focusing of cultural differences. If one is interested in individual differences, the best strategy is to measure them within culture, not across cultures.

10.01.4.1 Some General Strategies

When studying a phenomenon across cultures it is highly desirable to work with local social scientists who understand their culture. Such scientists are likely to insist that culture-specific aspects of the phenomenon be included in the study. Measurement can be made sensitive to the cultural context. Interpretation of the results can be much more sound when the perspectives of many cultures are used during data interpretations.

In addition, it is necessary to use more than one method because each method has a different meaning in each culture. If the results of several different methods converge, it is unlikely that the same bias accounts for all the results.

One must also do separate "construct validations" within each culture. In a construct validation, one has a theory and tests the theory. If the data hang together the way the theory predicts, then both the theory and the measurements of the constructs must be valid.

Suppose you have a theory which predicts that there are certain "antecedents" of anxiety in every environment and certain "consequences." If you measure anxiety with different methods in each culture and you find that the correlations between your measures of anxiety and the several antecedents and consequences specified by your theory are approximately the same in all cultures, then you have validated your measures in those cultures. If the patterns of correlations are similar, and if certain statistical tests (discussed by Irvine & Carroll, 1980) indicate that the measures are equivalent, then you can compare the results across the cultures.

10.01.4.2 Good Theory Can Eliminate Some Rival Hypotheses

If one has a well-developed theory that makes several predictions, and the theory is supported by *the* data, some of the rival hypotheses become less plausible (Malpass, 1977). Suppose my theory predicts that in a collectivist culture people will be low in anxiety when they are with members of their ingroups and higher in anxiety when they are far away from their ingroups. In an individualistic culture my theory predicts no difference when the individual is tested when with or away from members of the ingroup.

Suppose I collected data in rural China and rural America and I found what I predicted. In that case many of the rival hypotheses—different definitions of the constructs, different levels of motivation, response sets—become less plausible.

This does not mean we should not worry about rival hypotheses when we have good theory. Clearly, if the theory and a rival hypothesis make the same predictions, we cannot use the theory to eliminate that rival hypothesis. But in many cases the theory and the rival hypothesis will make different predictions and then if the data support the theory that is reassuring.

The main point to remember from this discussion is that cultural comparisons which indicate a cultural difference require a lot of work, checking, elimination of rival hypotheses, multimethod measurements, and the like, before the presence of cultural difference becomes plausible. As a sophisticated consumer of cross-cultural research, the reader should be looking for evidence that this extra work, checking, and elimination of rival hypotheses was actually undertaken by the researcher.

10.01.4.3 Methods that Can Establish Measurement Equivalence Across Cultures

Several methods may be used to establish cross-cultural equivalence.

(i) Suppose one has measured several variables in each culture and submits the correlations among these variables to a factor analysis. If the same factor structure appears in each culture (e.g., Tucker coefficients of 0.90 + for each factor) that is one indication that the variables have the same meaning in the relevant cultures.

(ii) Item response theory (Hulin, Drasgow, & Parsons, 1983) provides a statistical procedure which indicates if a particular item has the same meaning in two or more cultures. The procedure is based on the similarity in the way people respond to that item and to all the other items.

(iii) Psychophysical methods allow some physical variable (e.g., distance between physical objects) to be linked to a psychological variable (how do you feel about social entities) (for an example, see Triandis, McCusker, & Hui, 1990).

10.01.4.4 Emics and Etics in Cross-cultural Research

I introduced above the idea of etic (universal) and emic (culture specific) cultural elements. Here I will elaborate on the meaning of these constructs and explain how they might be used in the measurement of constructs.

Emics, roughly speaking, are ideas, behaviors, items, and concepts, which are culture-specific. Etics, roughly speaking, are ideas, behaviors, items, and concepts which are culture-general—that is, universal.

Emic concepts are especially useful in communicating within a culture, where one word can sometimes be used to convey a very complex idea. For example, a geographer who has studied the inhabitants of the Tierra del Fuego in the Southern tip of South America, told me that they have a word, *mamihlapinatapei*, which means "looking at each other hoping that either one will offer to do something that both desire but are unwilling to be the first to do."

One can almost see the scenario when boy meets girl and they mamihlapinatapei! One learns quite a bit about that culture by knowing the definition of this word. Similarly, by learning particular words, we get to know more about a culture.

Emic concepts are essential for understanding a culture. However, since they are unique to the particular culture, they are not useful for cross-cultural comparisons.

Let us consider an analogy, the comparison of apples and oranges. Apples and oranges have some common attributes, such as size, weight, price, and availability. They also have some unique attributes, such as unique flavor and aroma. Clearly, if we are going to understand what an orange is, we need to know about orange flavor. But we can not compare apples and oranges on orange flavor, except to say that apples do not have it. On the other hand, if we want to talk about price, we can compare apples and oranges. If we want to compare price to size or weight we can certainly do that, too. So, we can have a "theory" of size–weight–price relationships and see if it holds as well for apples as it does for oranges, and we can even extend it to other fruits. Emics are like apple or orangeflavor; etics are like size, weight, and price. So, for certain purposes, such as comparisons, we must use etic concepts; for other purposes, such as getting a real "taste" of the culture, we must use emic concepts.

More formally, emics are studied within the system in one culture, and their structure is discovered within the system. Etics are studied outside the system, in more than one culture, and their structure is theoretical. To develop "scientific" generalizations about relationships among variables, we must use etics. However, if we are going to "understand" a culture we must use emics.

Many anthropologists work with emics and think that etics are silly. They would say: "you do not know about apples by just knowing about their price, weight, and size." Psychologists want to make generalizations about people, so they do not want to get into the details of a single culture. Cross-cultural psychologists try to both understand and compare cultures. They work with both emics and etics.

The important point is to find convergence between different methods of understanding reality. I do not agree with the view expounded by some "deconstruction" humanists, or even some who pretend to be scientists, who argue that nature is constructed, not discovered—that truth is made, not found (e.g., Haraway, 1991). I think that this is an extreme position which is not science, since it cannot be discomfirmed. On the other hand, I do grant that our subjective responses to reality are often constructed. I am a realist, so I do not believe that women are less good as Chief Executive Officers (CEOs), just because that is the male-dominated "construction" and conventional wisdom of current CEOs. But I would agree with this view if someone showed me a convincing study, with hard criteria, which indicated that this view is correct.

"Constructionism" can result too easily in an image of reality which is found in only one person's mind. It might be good literature, but it is terrible science. The essence of science is a conversation between scientists and nature. It requires probing with multiple methods, to replicate important findings, and establish convergence between observations and measurements. If such convergence is broad, that is, includes humanistic "findings" as well as findings obtained through scientific methods, we can be much more certain that we have identified an important phenomenon than if we have only one person's argument or a single set of findings obtained in one place, with only one method. An important distinction between the humanistic and the scientific method is the contrast between the subjective and the objective. If the humanistic insight is to be taken seriously it has to converge at some point with other evidence; it cannot remain entirely subjective.

If we take a construct generated in our culture and use it in another culture, we may have a pseudoetic (false etic) construct. We must get empirical evidence that the construct operates the same way in the other culture, that is, it is a true etic, before we use it to compare the cultures. Remember the discussion about construct validity above? That is what needs to be done to establish a true etic. Then we have to make sure that this etic construct is measured in ways that are culturally sensitive. That often requires the use of local terms and ideas. In short, we must use both etic constructs and emic ways to measure them. That is why cross-cultural psychologists advocate the use of *both* emics and etics (Berry, Poortinga, Segall, & Dasen, 1992; Triandis, 1972, 1992).

One advantage of such a strategy is that one can obtain more sensitive information about the relevant cultures, which implies that one can obtain more cultural differences if one uses both emics and etics in the design of cross-cultural studies. In one study (Triandis & Marin, 1983), samples of Hispanics and non-Hispanics responded to either a questionnaire based on ideas generated in previous studies with American samples (pseudoetic) or a questionnaire which used ideas obtained from focus groups (informal discussion groups which generated items to be used in the questionnaire) of Hispanics and non-Hispanics. The latter questionnaire included both emic items which were spontaneously generated only by the Hispanic focus groups and etic items which were generated by both the Hispanic and non-Hispanic focus groups. Since the questionnaires in that study had 600 questions, one could get six cultural differences by chance (at the 0.01 level of significance). The number obtained by the pseudoetic questionnaire was 14 and the number obtained by the emic plus etic questionnaire was 50. In other words, one can miss a lot of subtle information about the way cultures differ if one does not use emics.

The importance of emics cannot be overestimated. There are emic concepts that are extremely difficult to understand by people who use an etic framework.

Consider this example: What is a geisha? Most Westerners are likely to freely associate "prostitute" with that concept, and as a result they will be quite wrong. A 500+ page ethnography about geishas (Dalby, 1983) shows that the closest association to "geisha" should be "jester." Just like the jester in a king's court had the function of diverting the king and his guests, so a geisha has as her chief function entertaining the clients of an establishment. Most geishas spend their time reciting poetry, singing, dancing, and serving their clients while they eat and drink.

Furthermore, there are many kinds of geishas, and only a minority are actual prostitutes (yujo). Most do other things, including greeting people at the door or playing musical instruments. The most highly sought geishas are older, more experienced, and thus able to act more appropriately. President Ford, the first US President to visit post-war Japan, had dinner with several geishas as part of his touring experience. The most beautiful were placed next to him; but the ones who did most of the entertaining were the older, more experienced ones. Of course, without knowledge of Japanese, President Ford may not have been able to appreciate the entertainment.

The contrast between "geisha" and "wife," in terms of the free associations of these words among Japanese males, is instructive. The former is seen as sexy, artistic, witty, and economically self-sufficient; the latter as sober, humdrum, serious, and economically dependent. The former is likely to be well read, know the latest poems and songs; the latter is likely to talk about the problems of the children.

Now that you have a glimpse of the meaning of geisha, do you see that it is a culturally specific role, that corresponds to an emic Japanese term, and the Western view that she is a "prostitute" is a pseudoetic, and becomes a true etic only when it uses the idea of "jester?" The general point is that using our own emic terms to understand other cultures will generally result in misunderstandings. We need to obtain etic terms which correspond to the emics.

10.01.4.5 How to Use Emic Measurements of Etic Constructs

Any theoretical construct of some generality in psychology is likely to have both etic and emic aspects. For example, the concept of "social distance" makes sense in every culture, and even among animals where territoriality is a well-established phenomenon. You allow some people to come close while you keep others away from you. The original concept, developed by Bogardus (1925), was operationalized by asking people if they would like to "marry," "live in the same neighborhood," "exclude from the country," etc. certain groups of people. It was found that Americans, in the 1920s, showed little social distance toward Western Europeans, especially those who came from the UK, and much social distance toward "Negroes," "Turks," and "Jews."

While the social distance concept is etic, the way it is operationalized can vary with culture. In India, the idea of "ritual pollution" results in social distance, and while a person may not mind living in the same neighborhood with a member of the lower castes, especially if that person is a servant, one can mind very much if that person "touches my earthenware." Thus, in India, social distance is indexed differently. Touching earthenware is an Indian emic; social distance is a theoretical concept, an etic.

Let me give you an example of how to use emic items to measure etic constructs. In a study I did a long time ago (Triandis & Triandis, 1962) I asked focus groups to generate items which measure social distance in Athens, Greece, and in Champaign, Illinois. I obtained at least 100 ideas from each sample. I then subjected the items to Thurstone scaling, which eliminates items which are ambiguous in each culture, and provides scale values for each item on an equal interval scale. I found that some items were etic, for example, "I would accept this person as an intimate friend" had a scale value of 11.1 in the US, and "I would accept this person as a best friend" had a scale value of 13.5 in Greece. But "I would accept this person as a roommate" was an American emic, with a scale value of 29.5. "I would accept this person as a member of my *parea*" was a Greek emic, with a scale value of 31.1. *Parea* means, roughly, regularly-meeting peer group. Thus, these two items mean more or less the same thing, but the American item does not make sense in Greece, and the Greek item does not make sense in the US. The method allows for culture-sensitive measurement, since Greeks rarely live in residential quarters in college, and thus rarely have roommates, while Americans are used to that concept. On the other hand, Greeks do have regular-meeting peers and a term for that. “I would accept this person as my family’s friend” had a scale value of 40.9 in the US and 24.0 in Greece, indicating that while that item makes sense in both cultures it implies much less social distance in Greece than in the US. It apparently was more significant to be a friend of the family in Greece (a collectivist culture at that time, where the family was the most important social group) than in the US. Similarly, “I would rent a room from this person” implied less social distance in Greece (42.8) than in the US (57.5). On the other hand, “I would exclude this person from my country” had about the same meaning (95 in US, 82.6 in Greece).

Thurstone scaling uses judges to obtain the scale values of the items. The judges examine each item and place it on an 11-point scale, according to how much social distance is implied when a person agrees with that item. About 30 judges seem to be sufficient for each culture. The distributions of the 30 judgments concerning each item are examined. When these distributions are broad or bimodal, the item is rejected because it is ambiguous, or does not measure the quality of interest. These judgments are then treated statistically to obtain the scale values for each item. In this case the top and bottom of the scale were called 100 and zero for easy comprehension.

When the actual people are studied, the items of the social distance scale are presented to them in a scrambled order. They are asked to pick three statements which best represent how they would act toward a particular stimulus person, for example, a “50-year-old unskilled Chinese laborer.” The respondent’s social distance is the average social distance value of the three items which were endorsed. One can generate stimuli systematically, by varying, for instance, the age, nationality, and occupation (or whatever one is interested in studying) of the stimulus person and then obtain the social distance each respondent feels toward that stimulus combination. This allows the researcher to do analyses of variance that show the relative importance of the attributes, such as age, nationality, and occupation, as determinants of the social distance judgments. Note that the data come from the scale values of the statements that have been generated and standardized in each culture separately. So, the fact that the statements have different scale values in different cultures does not produce a problem. They are emic values, but the construct is etic, allowing cross-cultural comparison.

The same strategy can be used with most concepts and any attitude. Concepts may have a universal meaning that most researchers understand the same way, yet they must be operationalized and measured differently in each culture, because behavior has different meanings in each culture (Triandis, Vassiliou, & Nassiakou, 1968). Those interested in a more detailed discussion of this approach may find Triandis (1992) useful.

10.01.4.5.1 General recommendations

In summary, for good cross-cultural work we must start with the theoretical construct and discuss it with local informants to see how it is best operationalized. It is then useful to use “focus groups,” that is, small groups that do some brainstorming, and come up with around 100 ideas about how to measure the construct. While these ideas will probably be different in each culture, the chances are that some of them will be generated in both cultures.

Next, we should take each of these 100 ideas and “scale them” separately in each culture. This means local standardization of the items. The technique developed by Thurstone (1931) (see, Edwards, 1957, Chapter 5) seems most helpful, but other scaling strategies may also be useful. With this method a small set of 15 or so scaled items can be identified in each culture that have been locally standardized but link to the etic continuum under study.

This method provides excellent locally standardized and equivalent equal-interval scales in the two cultures. Then the 15 items are scrambled and presented to the subjects under study in each culture with the instructions: “Pick three of the following 15 items which describe your most likely behaviors toward (the social category of interest).”

10.01.4.6 Examples of Multimethod Measurements

We have already emphasized the desirability of using multimethod measurements in cross-cultural studies. The following list of methods suggests the range of possibilities.

10.01.4.6.1 Ethnographic work

This is based on observations, with some questioning of informants and occasionally an experiment or survey (see Goodenough, 1980, for details). In such work an anthropologist spends one or two years among a group of people as a participant observer. After learning the local language, the scientist becomes a member of the culture, and often assumes one of the existing roles within the culture (e.g., the chief’s son). After that, observations can be done informally or formally (see Longabaugh, 1980, for details) using video tapes or films which are later coded by several coders to establish inter-rater reliability.

10.01.4.6.2 Establishing shared cognitions

This approach includes the measurement of a psychological construct with several methods in each culture, and checking if the measurements converge. However, it is important to be able to distinguish a psychological from a demographic or cultural construct. Culture consists of shared elements. When we measure a psychological construct we do not know if it is shared so it may or may not be part of culture.

Triandis, Bontempo, Leung, and Hui (1990) have developed a method which allows a researcher to sort the personal, demographic, and cultural constructs. The basic idea is to have triads (three respondents) hear a question and answer it while the time it takes for them to agree on the answer is measured. Cultural elements do not require debate. For example, if you ask three Americans whether "fairness" is important or unimportant in everyday life, they are likely to supply a "Yes" in less than two seconds, and over 90% of the triads which are tested will agree. But if you ask: Is "fame" important or unimportant, they engage in a debate. Being famous is not a widely shared element of American culture; some want to be famous and others do not. For those who want to be famous this is a personal construct.

Using these two criteria (length of time to respond, percent of the triads who agree), the researchers can pick those elements of the culture which are widely shared.

This study was done in Hong Kong and Illinois as a demonstration of what can be done to measure cultural elements. The same could be done with demographic elements (men vs. women, old vs. young, rich vs. poor) by assembling appropriate triads and measuring how much time it takes for them to agree on how to answer a question, and what percentages of the triads do agree. Also, this study was limited to "values," but it could be done with any element of subjective culture.

Unfortunately, this method is participant-intensive. To have 50 triads, which is not an especially large number for a somewhat heterogeneous culture, 150 people are needed in each culture.

It is possible to compare the data of this method with the results obtained by questioning one person at a time with other methods. If the results are similar, one has evidence about the convergent validity of both methods. The Hong Kong/Illinois study did that and convergent validity was obtained.

Another strategy which takes seriously the fact that culture consists of shared elements is to examine the distributions of the responses of people from each culture to any psychological measure. If an arbitrary 90% of the respondents give the same answer, that is a cultural element. Clearly, the identification of cultural elements is important in studying cultures, but is of little interest when the focus is on individual differences. The two strategies are quite different. When studying individual differences we should do it within culture.

Interviews and surveys are often used (Pareek & Rao, 1980). Some of these are informal and allow for questions to be asked as the interview proceeds, while others follow a rigid schedule. These interviews can be based on limited or representative samples of the culture.

One might also use tests (Irvine & Carroll, 1980), attitude scales, personality scales, projective tests (Holtzman, 1980), and psychophysical tests.

Experiments can be done in more than one culture, but they present special difficulties (see Brown & Secrest, 1980; Ciborowski, 1980). For example, it is extremely difficult to ensure that the same degree of manipulation of the independent variables has been used in each culture.

However, some interesting results can be obtained when the same experimental procedure is used in each culture. Strodtbeck (1951) provides an interesting example. He examined whether culture is related to the probability that the wife or the husband will "win" an argument. He tested 10 husband-wife pairs in each of three cultures: Navajos (where the custom was for the husband to live with his wife's relatives, and hence one might expect husbands to have less power), Texans, and Mormons (where traditional male supremacy was used). The study began by asking each spouse a number of questions separately. The researcher noted on which questions the spouses disagreed. He then put them together and tape-recorded their interactions while they discussed their answers to the questions and attempted to reconcile their differences. If the couple finally agreed with the husband's original position, that indicated husband dominance; if the couple finally agreed with the wife's original position, it indicated wife dominance. The results were as follows:

Number of decisions "won" by		
	Husband	Wife
Navajos	34	46
Texans	39	33
Mormons	42	29

The hypothesis was supported. Culture predicts who wins, with the Navajo more likely to agree with the wife and the Mormons with the husband.

Finally, I will mention two methods which can obtain data from texts or other kinds of published materials:

(i) Content analyses (see Brislin, 1980) of children's stories, newspaper stories, myths, formal and informal communications, speeches, or movies produced in several cultures have been used to measure attitudes, motives, opinions, values, and other attributes. For example, one rates the materials on two variables and sees if the ratings are correlated.

(ii) The human relations area files (Barry, 1980) is a very useful data set. It consists of photocopies of much of the world's ethnographic record classified alphabetically according to the content of each paragraph. Thus, for instance, a researcher who is interested in checking if the "age of weaning" of a child is related to "level of anxiety" in the culture's adult population would look in the files under "weaning" and would find the ages of weaning which were reported by ethnographers who had studied several cultures. Similarly, this researcher would find some categories which might index the extent adults in the culture are high or low in anxiety. Then the researcher would either plot the data, or use correlations or chi-squares to check the relationship of age of weaning to anxiety.

Of course, ethnographies do not include every type of information, so it may be that there is no information relevant to anxiety in a particular culture. The result is "missing data."

In these analyses the number of cultures is the number of observations on the basis of which the correlation is computed, or *N* cultures are classified as "high" or "low" on each of the variables to compute the chi-square test. In other words, one can establish an association between two variables, based on holocultural data (all cultures on which there is information; see, Naroll, Michik, & Naroll, 1980 for details).

10.01.4.6.3 Conclusions

It is ideal to test a hypothesis with as many of these methods as feasible. Some of these methods are "operant" in the sense that the researcher provides a minimal stimulus and the subjects provide many responses (operate on the stimulus as they see fit when responding). For example, sentence completion or projective techniques, ethnographies and the human relations area files, observations, unobtrusive measures (see Bochner, 1980, for details), and content analyses are operant techniques. In most cases the researcher does nothing to stimulate the production of the data. In some cases a minimal stimulus, such as "Please write 20 statements which begin with the words *'I am,'*" is provided.

By contrast, experiments, surveys, and interviews are "respondent" methods; the subject is responding to stimuli presented by the researcher.

Respondent methods are more obtrusive, and they are more likely to be distorted by reactivity. The respondents are more likely to distort their answers, so they will appear to be socially desirable people to the researcher, their peers, the authorities in their culture, or from the point of view of their culture's ideal. In short, respondent methods are more likely to result in cultural differences due to the method. However, it is easier to "control" artifacts, when using respondent methods because the study can be replicated under a variety of conditions.

Operant methods are generally less reliable. Many observations or responses are obtained and it is not clear which ones are the most important. Operant methods also have a high "dross rate" (irrelevant information is a high proportion of the total information obtained). Furthermore, the observer may be biased and only see or hear what is consistent with a hypothesis.

Both kinds of methods should be used at different points in the research sequence. At the initial stages of the research, when one knows little about the culture, lacks good hypotheses, and is dealing with respondents who are not familiar with social science methodologies, it is best to use operant methods. Such methods are especially good when investigating complex relationships. These relationships can be kept in mind while making additional observations. However, since these methods are not sufficiently reliable, and are difficult to check for reliability and validity, one should not reach definitive conclusions by using only these methods. Rather, it is best to refine existing hypotheses, develop new ones, and keep an open mind about the culture while using them.

At a later stage in the research process when more is known about the culture, the hypotheses are more likely to be supported and then it might be possible to design experiments, questionnaires, and interview schedules which are appropriate for the problem. It also helps if the participants are familiar with the researcher's methods, since they may then be more likely to give the same meaning as the researcher to the testing situation.

In other words, no method is perfect; each method has both advantages and disadvantages. The sophisticated consumer of cross-cultural research will give more weight to findings which have been supported by more than one method, particularly if the methods were very different.

Separate tests of hypotheses within cultures and between cultures, and with more than one method, increase the confidence in any finding.

However, we must not assume that a test at the between cultures level will necessarily give the same results as a test at the within cultures level. For example, consider the variable "degree of industrialization" and correlate it with the variable "probability that a worker will vote for the communist party." In India and the US, there is a greater probability of a communist vote in a highly industrialized voting district than in a less industrial one. In other words, within countries there is a positive relationship between the variables of interest. But when we compare the two countries we see a negative relationship between these two variables. In the US there is a higher level of industrialization and also a low probability that anyone will vote for the communist party; in India, there is less industrialization, but a relatively high probability that a worker will vote for the communists. In short, there is a reversal of the sign of the relationship between the two variables when we study the phenomenon across vs. within cultures.

This means that while we should do our analyses at both the between and within cultures levels, we need not become discouraged when these results are inconsistent. Of course, consistency boosts our confidence that the relationship is robust. There are examples (see Rohner, 1986; Schwartz, 1994) in which a hypothesis has been tested at both levels and was supported consistently.

10.01.4.7 Translations

It is ideal to gather the data in each culture independently by using the same procedures but without translation of specific items. This view is supported by the realization that translation is at best approximate.

There are some hilarious examples of poor translation that suggest how difficult it is to translate. For example, outside a Hong Kong tailor shop: "Ladies may have a fit upstairs." In a Greek tailor shop: "Order your suits here; because of a big rush we will execute customers in strict rotation." An Italian laundry: "Ladies, leave your clothes here and spend the afternoon having a good time." In a Japanese hotel: "You are invited to take advantage of the chambermaid." A Norwegian cocktail lounge: "Ladies are requested not to have children in the bar." At the office of an Italian physician: "Specialist in women and other diseases." A Japanese hotel air conditioner: "Cools and Heats: If you want just condition of warm in your room, please control yourself."

In short, one can make the case that translation should be avoided. But that is not always practical, so a word about translation is necessary.

Translation of single words is unwise. It is extremely difficult to establish exact equivalence in meaning for single words. Only if the study focuses on the meaning of single words and uses several methods to tap that meaning is translation of single words acceptable.

The greater the context of a text, idea, or concept, the more likely it is that it can be translated properly. If one expresses an idea in more than one way, translation has a chance of reaching linguistic equivalence. Thus, it is helpful to introduce redundancy, synonyms, and context.

Brislin (1980) provides many useful suggestions on how to maximize linguistic equivalence. A good approximation to the ideal translation uses the Werner and Campbell (1970) method of double translation with decentering. This method is based on the realization that there are many ways to say the same thing, and careful adjustments to the original language of the research project may not produce any difficulties for the research but may facilitate the translation.

For example, the researcher starts with an English text (E), and asks a bilingual to translate it into Japanese (J). Then, he/she asks another bilingual to translate it back into English to obtain E′. A comparison of E and E′ indicates where problems may exist. If the discrepancies of E and E′ are significant, the researcher can "decenter" the text by producing a new text (E″ which is satisfactory for the purposes of the research, but is closer to E′ than to E). Such a text is likely to translate into Japanese more easily, since E′ was an English version of a Japanese text. The process of translation, back-translation, and decentering continues until two English texts emerge that are more or less identical.

Back translation seems like a good way to obtain linguistic equivalence but it has problems. First, many words in languages of the same language family (e.g., Indo-European languages cover a wide area from India to Europe to the Americas) have the same roots but different meanings. For example, in the Latin-based languages, the term *sympatique* or *simpatico* means "pleasant or agreeable person." In English *sympathetic* means "someone who feels like the other person does." Clearly, the meanings are not the same. Yet, bilinguals are apt to translate it and back-translate it into the original language when in fact the translation is not correct.

Second, skilled bilinguals are good at imagining what the original text might have looked like. So, suppose a Japanese text is being translated, the interpreter may be able to guess what the original E or E′ was and produce an E″ which is just like it. The researcher would be reassured, but the translation is imperfect. In short, while back-translation is desirable, it does not guarantee linguistic equivalence.

One technique which is useful, continuing with our example of English and Japanese, is administration of both the English and Japanese version of a questionnaire to a sample of Japanese bilinguals. In such a case one can compare the responses of the same person when answering in the two languages. This can provide a check on the translation, since the same answer would indicate good translation. However, even this technique is problematic because bilinguals have a tendency to present themselves in a more socially desirable way to "outsiders" than to "insiders" (Marin, Triandis, Betancourt, & Kashima, 1983). Thus, it is likely that a Japanese bilingual will present a more favorable set of answers when answering in English than when answering in Japanese.

In sum, to be safe do not translate. Basically, you can carry out the same data gathering operations in each culture, and the only text that needs to be translated is the instructions, which usually have a good deal of context. Then, standardization is done separately in each culture. Comparison of distributions, computation of Tucker coefficients of the factor structures that emerge from the correlations of the variables of the study, item response theory checks, etc. can be done to establish equivalence before any comparisons are made.

Finally, studies which report a large number of cultural similarities, and show the cultural differences as being embedded in these similarities, are more dependable than studies which do not report similarities. If all the data from one culture is different from all the data from another, it is very likely that the tasks were not communicated adequately from one culture to another, and the cultural differences are due to some artifact.

10.01.4.8 Ethics of Cross-cultural Studies

The sophisticated consumer of cross-cultural research will want to see that the research has been carried out ethically. Good research is ethical. Unfortunately, not all researchers carry out their work ethically. Some study what is easy to study, and as a result we know a lot more about the disadvantaged than the advantaged segments of most cultures. When two groups are in conflict and a researcher offers to study the conflict, the weak are more likely than the strong to accept the study. The powerful do not submit to research very easily; that means our picture of the powerful members of a culture may be distorted.

Some researchers do not provide their subjects "informed consent." Ethical research requires that the participants are able to avoid participation in studies they consider unethical or disadvantageous to them. But in many nonliterate societies the concept of "research" is nonexistent, and it is impossible for the prospective participants to provide meaningful informed consent.

With most operant methods the chances are that people will not be exposed to stress or risks. But with respondent methods some risks can be significant. If risks beyond those of ordinary life are involved in a research method it is essential that the risks be explained to the individuals, and they should have the opportunity to decline participation. This, of course, produces other problems, such as distorted samples consisting mostly of volunteers. Nevertheless, if risk is involved the subjects must be informed.

Since risk is perceived differently in each culture, it is important to bring research collaborators into the decision-making process at the earliest points of a project. If a method which appears risk free in one culture is not risk free in another, the method may have to be changed. There are many ways to collect data, so it does not follow that because we changed the method we cannot test an important theory.

There are special problems of research collaboration across cultures that are discussed in detail by Tapp, Kelman, Triandis, Wrightsman, and Coelho (1974). For example, research collaborators may be harmed by collaborating with someone from a culture that is politically taboo in their country. It is generally believed that researchers must leave something of value in the culture they studied to avoid "intellectual colonialism." What is of value will, of course, vary with culture but it can be information, procedures, material goods, or payments.

10.01.4.9 Ethnocentric and Androcentric Bias of Researchers

Most cross-cultural researchers are Western men; we all have difficulties in escaping our ethnocentric (my culture is the standard of comparison) and androcentric (my gender offers the only valid perspective on an issue) biases. We can try to control such biases, but the choices of problems, theories, and methods are likely to reflect such biases. For example, we cannot be sure that we have controlled such biases when we evaluate whether or not gender inequalities are similar or different across cultures. Thus, when evaluating cross-cultural research it is wise to ask whether such biases may have colored the reported findings, interpretations, and conclusions. A useful way to overcome such biases is to collaborate with researchers from other cultures and the other sex.

10.01.4.10 Summary of Criteria of Good Cross-cultural Research

(i) Does the study include many similarities across the cultures and are cultural differences embedded in the similarities?

(ii) Did the study use multimethod procedures that converged (Fiske, 1986; Fiske & Shweder, 1986)?

(iii) Did the study use tests both within (psychological level) and between (cultural level) cultures? Were the results consistent, or when inconsistent theoretically meaningful?

(iv) Were rival hypotheses checked and eliminated?

(v) Did the study use etic constructs measured emically, that is, with emic items generated in each culture to tap the etic constructs. Ideally, researchers should generate items from both men and women (Harding, 1987; Nielson, 1990; Reinharz, 1992), as well as from all important groups that make the culture heterogeneous (e.g., social class, religion, language, age, family structure). Were measurements standardized independently in each culture?

(vi) Was the study conducted ethically?

(vii) Did the researchers make an effort to control ethnocentric and androcentric biases?

10.01.5 ACCULTURATION

The discussion above assumed that cultures are isolated entities. However, in fact, they are in constant interaction. Individuals are often influenced by many cultures.

When two cultures come in contact with each other, individuals have the tendency to adopt some of the attributes of members of the other culture (accommodation), and sometimes they are so enthusiastic in adopting elements of the other culture that they overshoot. For example, individuals who come from a culture which uses time flexibly, for example, Mexico, may become superpunctual in a culture where "time is money."

However, migrants who are rejected sometimes react to the culture contact by emphasizing their own culture (ethnic affirmation). For example, they may become even more flexible about their use of time, and advocate that one should not own a watch, or bother to be on time, because that reduces one's freedom and makes a person like an automaton.

A complication is that in some domains people show accommodation or overshooting, and in other domains they show ethnic affirmation. Triandis, Kashima, Shimada, and Villareal (1986) hypothesized that if newcomers in the US are accepted they will show accommodation or overshooting, but if they are rejected they will show ethnic affirmation. In addition, accommodation and overshooting are more likely on visible traits, such as clothing, and ethnic affirmation is more likely on subjective traits, such as beliefs, stereotypes, or values.

A socio-political issue of great importance, when two cultures meet, is whether individuals should adopt or reject the other culture, and also try to maintain or reject their own culture. Berry (1990) identified four outcomes of the acculturation process: he used the terminology integration for the situation when the individual adopts the other culture while maintaining his/her own culture; assimilation for the situation when the individual adopts the other culture and rejects his/her own culture, separation for the situation where the individuals maintain their own culture and reject the other culture, and marginalization for the situation where both cultures are rejected.

Given that in the US, integration, as used in the press, means something different from Berry's integration, it may be best to call that condition biculturalism. In any case, Berry, Kim, Power, Young, and Bujaki (1989) found that people who adopt the integration or biculturalism solution to the acculturation situation, and secondarily the assimilation pattern, are better adjusted, both physically and in terms of their mental health, than those who use the other two acculturation patterns.

If we note that learning about another culture is similar to learning another language, we can see the advantages of knowing about more than one culture. Segalowitz (1980) identified many cognitive advantages associated with bilingualism, for example, cognitive flexibility, creativity. Similarly, a multicultural person has more than one way of interpreting reality. However, as Lambert (1992) has noted, bilingualism can be both negative and positive. It is negative when in the process of learning another language one forgets one's own language; it is positive when one is actually comfortable in both languages. Similarly, Triandis (1976a) has argued that multiculturalism can be both negative (assimilation) and positive (biculturalism). Adjustment is likely to be superior when positive multiculturalism is present.

Social identity (Tajfel, 1982) is strongly linked with the processes reflected in acculturation. An interesting example is provided by Rhee, Uleman, Lee, and Roman (1995). They asked Korean and American participants to complete 20 statements that started with "I am . . . " They coded the responses into nine categories, one of which was "pure traits." Previous work (e.g., Triandis et al., 1990) had found that in collectivist cultures the content of the responses to the 20 statements is more social than in individualistic cultures. Clearly, pure traits is an individualistic category. European-Americans gave pure trait responses 29% of the time, Koreans did so only 12% of the time.

Asian-Americans were classified according to whether when completing "I am . . ." sentences they had identified their ethnicity "never," "once," or "twice." Those who never mentioned their ethnicity gave 39% pure trait responses—a case of overshooting the European-American norm. They must have been the assimilated Asian-Americans. Those who identified their ethnicity once (e.g., I am Asian or I am Korean-American) averaged 25%, almost the same as the European-Americans. Those who identified their ethnicity twice averaged 17%, which is very similar to the 12% of the Koreans in Korea. In short, those with a strong Asian-American identity responded like the Koreans.

The implications for clinical work of the observation that biculturalism is best is that far from pressuring clients to discard their culture, clinicians should urge them to maintain parts of their culture as well as to acquire those aspects of the mainstream's culture which will help their economic advancement. That may at times seem difficult to achieve. However, there are opportunities for "cultural synergy" (Adler, 1991). Adler argues that when one culture solves a problem one way and the other culture solves it in a different way, rather than become frozen in the opposing ways to solve the problem, one should seek creative solutions which are acceptable to both cultures. In short, rather than the parochial "our way is the only way," or the ethnocentric "our way is the best way," there should be the synergistic, that is, the "creative combinations of our way and their way may be the best way" (p. 107).

To achieve synergy one needs to understand the point of view of each culture first. Second, one needs to explore and understand the cultural assumptions which underpin the particular beliefs. Then one can create new alternatives which are based on, but are not limited to, the cultures involved. For example, in the clash of collectivist and individualist perspectives, one might find a perspective which is compatible with both sets of assumptions. In our previous example about widows not eating chicken, one could serve a food which is not perceived as having the attributes of increasing sexual desire (as seen in Orissa) and is interesting and desirable (as seen in the US). The next step is to check if the solution fits the assumptions of both cultures. One must discuss the solution with members of each culture to make sure that the solution is acceptable.

10.01.6 IDENTITIES OF AFRICAN- AND EUROPEAN-AMERICANS AS AN EXAMPLE OF THE IDENTITIES OF CLINICIANS AND CLIENTS

The US includes 125 ethnic groups with nontrivial numbers (Thernstrom, 1980). Thus, a clinical psychologist is likely to encounter clients from many ethnic groups. Of course, a lot will depend on location. There are more Albanians or Zairians in one city than another. Also, fortunately, there are similarities among cultures. Anthropologists identify "cultural regions" within which there is considerable cultural similarity. Generally speaking, Africa South of the Sahara is one region; the cultures around the Mediterranean, including most of Europe, have much in common; the cultures of East Asia, South Asia, and the Pacific Islands have similarities. Native Americans (of course, they include tribes in South, Central, as well as North America) are among the most diverse cultures. For example, according to the criterion of social structure, they belong to five cultural regions (Burton, Moore, Whiting, & Romney, 1996). Clearly, clinicians cannot be expected to know much about all 125 cultures, or even about most of the cultural regions. However, there are some ethnic groups which are sufficiently common to require such knowledge. African-Americans are by far the most important, both because of their numbers and the long history of cross-cultural contact between them and the European cultures.

Each pair of cultures has a unique history of contact. For example, African- and European-Americans have had a history of contact which is not the same as the history of Asian-Americans with African-Americans. It is important to realize that the history of contact is an important factor in the kinds of identity which each group develops, and the kinds of "contact models" which each group generates.

Carter (1995) argues that people in the US develop a racial identity which affects their thoughts, feelings, and behaviors. Any valid psychotherapeutic process in this country must include race. He believes that this identity is important even in White–White interactions because so much of the content of social behavior in the US reflects inter-racial experiences in school, work, dating, neighborhood, or public policy.

Carter (1995) also argues that each ethnic group (Native Americans, Asian-Americans, Hispanics, African-Americans) goes through a series of stages in developing an ethnic identity which is bound to enter into the clinician–client relationship. Ethnic identity reflects how the individuals think about their own group as well as how they view the relationship between their group and the other ethnic groups.

Ethnic identity, of course, also reflects the kinds of acculturation patterns which are operating, as discussed above.

Carter's (1995) discussion of Black racial identity specifies four stages, each of which includes two levels. I will only describe the four stages.

Stage 1 is the pre-encounter stage during which there is dependency on White society for self-definition and approval. Racial identity attitudes towards Blackness are negative and White culture and society are seen as the ideal. There is an idealization of whites: much anxiety, low self-esteem, and a weak ego (in Loevinger's, 1976, 1979, sense). Most Blacks in this stage have higher than average incomes and do not report that they experience discrimination in the workplace. The client in this stage prefers a white counselor.

Stage 2, the encounter stage, begins when the individual has a personal and challenging racial experience. It includes confusion, emotions which are bitter, hurt, angry; there is high anxiety, low psychological well-being, high depression, and preference for a Black counselor.

Stage 3, the immersion stage, is characterized by positive feelings toward Blacks and idealization of Black culture, while having intense negative feelings toward Whites and White culture. The emotions include rage and depression. Individuals have low self-esteem and low self-actualization tendencies, are characterized by emotional dependence, but people in this stage have a strong ego in Loevinger's sense. This stage is most common among those of low education, who are separatists with Afro-centric values and anti-White/system attitudes.

Stage 4 is internalization. Here individuals realize that both Blacks and Whites have both strengths and weaknesses. Black identity is experienced as positive, important, and a valued aspect of the self. Individuals in this stage respect Whites and tolerate differences between Whites and Blacks. The emotions are self-confidence and security. At the most developed level of this stage there is individual activity to promote the welfare of Black people.

Carter (1995) describes five stages of White racial identity: stage 1, contact, includes persons who are unaware of others in racial terms, and deny the importance of race. However, they experience discomfort with unfamiliar people and situations. They are characterized by low anxiety, poor self-image, dependence, and are immature interpersonally. Such people have Euro-American values, but support organizations that promote equity.

In Stage 2, disintegration, people become aware of social norms and pressures associated with cross-racial interactions, see negative reactions by Whites to inter-racial associations, and attempt to find some balance between human decency and external pressures to be negative toward Blacks. Typical emotions are anger and guilt. They are still interpersonally immature, prefer a White counselor, are confused by Blacks, have Euro-American values, and may have negative views of racial issues in the workplace.

Stage 3 is reintegration. People at this stage are fearful, angry and hostile toward Blacks. They are anti-Black and pro-White, and see Blacks stereotypically. Individuals idealize Whiteness, devalue other races, and are anti-Black. Their emotions include fear and anger. They feel high anxiety and hold racist views; they have Euro-American values and negative views of the racial issues in the workplace.

Stage 4 is called pseudo-independence. At this stage people understand that there are similarities and differences between the races and accept the differences. However, these people understand the differences only intellectually, and remain emotionally distant from Blacks. They may be interpersonally mature, may accept inter-racial dating, may be comfortable in cross-racial settings, but nevertheless prefer White company.

Stage 5, autonomy, is the most accepting and flexible of the White racial identities. People at this stage accept race as a positive part of the self, are calm, secure, self-confident, mature interpersonally, self-actualized, inner-directed, support racial integration, have balanced views of inter-racial relations, and positive racial views of the workplace.

Carter's theory is that if the therapist (Black or White) and client (Black or White) are at equivalent levels of racial identity development, there will be little improvement in the client's condition. Even worse, if the therapist is at a lower stage than the client, there can be regression. Only when the therapist is at a higher stage than the client will there be progress in psychotherapy. Thus, he argues that racial identity is crucial in understanding the therapeutic relationship. In fact, it is not race as such but racial identity that is most important in understanding the course of psychotherapy.

Therapists who are not aware of the proper way of understanding Black behavior may misinterpret as pathological behaviors which are normal in the other culture, or clients might feel anxious for not being what the dominant culture expects them to be. Thus, Carter (1995) advocates that racial identity must be incorporated in clinical training and practice and he criticizes models of intercultural training which do not explicitly deal with racial identity. The second part of his book reports empirical studies which support some of these arguments.

This theory may have merit, but perhaps Carter (1995) overemphasizes the importance of race, and also does not seem to take seriously the importance of the situation. There is at least one study which suggests that identity may switch rather readily. Sussman and Rosenfeld (1982) asked Venezuelan and Japanese bilingual students to place their chairs anywhere in an empty room and chat for 10 minutes in either their own language or in English. The researchers measured the distance between the chairs. They found that when these students spoke in their own languages the Venezuelans placed their chairs close to each other and the Japanese placed them far from each other. But when these participants spoke in English there were no cultural differences. This study would suggest that the situation (speaking in own language or in English) determined the participants' identity and behavior. It would seem that any theory about the way racial identity influences behavior should take the situation into account.

10.01.7 METHODS FOR THE DEVELOPMENT OF AN UNDERSTANDING OF THE CULTURE OF THE CLIENT

Analysis of the subjective culture of the clients (Triandis, 1972) requires examination of the way clients categorize experience, the meaning of key terms, associations, beliefs, evaluations, expectations, norms, role definitions, self-definitions, ideals, and values (see Triandis, 1994, Chapter 4). This is often done most easily by reading ethnographies about the client's culture, or empirical studies which have measured these cultural elements.

All 125 or so of the American ethnic groups (Thernstrom, 1980) are quite heterogeneous, so that any statement made about them must be used very carefully. For example, Hispanics are quite different depending on whether they have come from Cuba, Mexico, South America, or Puerto Rico. Social class differences, the relative influence of African or Native American cultures, the genetic mixture of European- and Native-American populations, and factors reflecting the history of the relationship between the ethnic group and the mainstream populations of the US, have influenced the kind of culture that has emerged.

A major difference between Asian- and Hispanic-Americans, on the one hand, and European-Americans, on the other, is that the former are more collectivist than the latter. Many of the attributes of these two kinds of cultural patterns (see Triandis, 1995, for details) are reflected in the greater emphasis on politeness, being a nice person, saving face (Ting-Toomey, 1988), and the like which are associated with collectivism and increasingly neglected among extreme individualists. Asian-Americans are also said to be less assertive than European-Americans (e.g., Zane, Sue, Hu, & Kwon, 1991). An excellent comparison of Chinese and Americans can be found in Hsu (1981).

Hispanics have been found to be rather collectivist (Marin & Triandis, 1985), especially familist (see Triandis, 1994, p. 247 for references), concerned with smooth interpersonal relationships (Triandis et al., 1984), interpersonal connections, and personalistic relationships (dealing with each person differently). They are especially concerned with "respect," "dignity," "loyalty," and "cooperation" (Albert, 1996).

African-Americans value expression in movement, sound, and the visual modalities more than European-Americans. Expression as personal style and movement, and the use of a "cool pose" (Majors & Mancini-Billson, 1992), as well as different patterns of communication in both language (Landis, McGrew, Day, Savage, & Saral, 1976) and paralinguistically, require that for effective interaction with African-Americans, European-Americans must learn a good deal about the African-American subcultures. Some analysts have argued that African-Americans when compared to European-Americans are more spiritual, higher in harmony with nature, more present oriented, use time more flexibly, emphasize oral expressions, movement, surprise, improvisation (see jazz), rejection of routine, have an identity that is defined by expression, style, and spontaneous activity, are more gregarious, flexible, easygoing, affectively driven, and high in affect (Boykin, 1983; Jones,1986, 1988).

The largest differences between European- and African-Americans have been found among those African-Americans who have never had a job (Triandis, 1976b). While those with jobs often have a very positive self-concept and a subjective culture that is not very different from the subjective culture of mainstream Americans, those who have never had a job often have neutral self-concepts; they accept the conditions of the ghettos (e.g., high crime rates) as natural, have strong antiestablishment attitudes, and see little connection between what one does and what one gets. They can be characterized as having "ecosystem distrust," that is, they do not trust people in their environment, and see events in their environment as unpredictable.

Some of these perceptions may reflect the realities of life in segregated ghettos. For example, if discrimination prevents people from getting a job, not seeing a link between finishing school and a job may actually be a veridical judgment rather than an aspect of ecosystem distrust. It may be the case that people in those environments are untrustworthy. What is striking, however, is that the great distrust is uncovered across the board, in the way they rate stimuli like mother, father, teachers, Black women, Black professionals, Black men, and so forth (Triandis, 1976b, p. 123).

When clinicians interact with ethnics, such as those described above, they may well experience "culture shock" (Oberg, 1960). Culture shock is a consequence of not having as much control (Langer, 1983) over the social environment as one is used to having. If the way clinicians usually approach their clients does not result in anticipated changes, the clinicians may feel that they have lost control of their clients. To regain control clinicians need to put themselves into the shoes of their clients (Brislin, 1993). That means learning to make "isomorphic attributions" (Triandis, 1975). That is, when they explain the behavior of their clients, they must use approximately the same causal attributions as their clients use to explain their own behavior.

10.01.7.1 Methods of Culture Learning

There are numerous ways clinicians can learn about another culture: they can spend time with members of the other culture; they can read about them; they can ask people who have worked with members of the other culture to describe that culture. Each of these methods can be effective, but each has limitations.

One of the major problems in culture learning is convincing the clinician that culture training is worthwhile. After all the clinician is busy, and culture learning is like learning another language. Can the time it takes to learn about the other culture be justified?

Also, most people are ethnocentric and feel that others must learn about their culture rather than they should learn about other cultures. After all, if clients are going to be successful in American culture they must learn how to negotiate situations within that culture, that is, learn the culture of the clinician. So why should the clinician bother to learn their culture?

Many clinicians also feel that they already know enough. Thus, very little cross-cultural learning is likely to take place.

However, if one is to do a good job with a client from another culture it is important to learn as much as possible about that culture. Changing a behavior which is central to the other culture can do more harm than good. For example, clients from collectivist cultures often appear to be too dependent on their families. A clinician may be tempted to make them more autonomous. Yet that is likely to upset many beliefs, feelings, and behavioral patterns if it is attempted too quickly and while the client is still young. Changing a client without understanding that behaviors which appear pathological are normal in the client's culture is extremely dangerous.

Qureshi (1989) has written a whole book about misdiagnoses which are likely when the health provider and the client come from different cultures. For example, if the clinician does not know that in Pakistan it is inappropriate to say "I do not know," or in India it is rare that one says "thank you," major misunderstandings might occur. If the clinician does not know that in the Eastern Mediterranean there is often a culture of honor (Campbell, 1964), and many people show an "Othelo syndrome" which consists of excessive concern about the chastity of spouse, jealousy, and delusions that the wife is unfaithful, he might interpret "normal behavior" as abnormal. Differences in gestures, such as using the gesture that in the US says "no" to say "yes," as in Bulgaria and Southern India, can cause serious problems for interaction. Culture-linked diets can result in misinterpretations. Differences in what is normal social structure (e.g., equality or inequality of roles), conceptions of beauty (e.g., fat is beautiful in some cultures), taboos (what parts of the body can a clinician not touch), conceptions about what it means to be "sick" and what one is supposed to do about it (e.g., go to a shaman or take a drink which contains mercury that is supposed to be a therapy for impotence), beliefs about the proper behavior after a birth, marriage, death, and bereavement can also result in misunderstandings.

Psychopathology often is an exaggeration or caricature of normal culture (Draguns, 1988). It may reflect culture-linked myths, religious beliefs, political views, esthetic standards, economic conditions which are unknown to the clinician, and result in misunderstandings. It often reflects the basic personality which is most frequent in the culture, social representations found in that culture, the inability to deal with specific environments, the values of the culture, specific forms of the self, methods of self-control, and culture-linked lifestyles. Even highly educated samples may use indigenous methods of dealing with illness. For example, Cook (1994) asked university level samples of Chinese, Indian, and Anglo-Canadians about their beliefs. Three kinds of methods of dealing with 16 illnesses were provided: (i) biomedical (e.g., consult a physician), (ii) psychosocial (e.g., use a remedy recommended by the family), and (iii) indigenous (consult a religious leader). There were highly significant differences among the three cultural groups in the extent of endorsement of the indigenous and psychosocial beliefs, but their endorsement of the biomedical methods was equivalent. For example, the Chinese indicated that in many cases they would not go beyond the family for help. Issues of "face" are linked with going for help outside the family in Chinese culture. There is also much more somatization of psychological problems in that culture.

There are many forms of cross-cultural training. Here I will describe them briefly. Landis and Bhagat (1996) have provided a Handbook which discusses them in detail.

10.01.7.1.1 Culture-general vs. culture-specific training

There are many topics relevant to understanding other cultures that are not specific to any culture. These include:

(i) Know that we are all ethnocentric. Try to catch yourself when you are ethnocentric.

(ii) Know that the attributions that clients will make about their own behavior are likely to be different from the attributions which you will make about their behavior. Look for discrepancies in these attributions.

(iii) Learn to sort what is personal from what is ethnic. When clients react to you they are often seeing you as a "representative" of your own ethnic group. They may be hostile, not because they do not like you, but because they do not like your ethnic group, which may have dominated their ethnic group. Also, people may be reacting to a cultural trait (e.g., they may not like professionals who wear white coats) rather than to a personal trait.

(iv) Learn that it is most difficult to act correctly in an environment where the situation looks just as it does in your own culture, but where the required behavior is very different. Most cross-cultural errors occur in such situations. You are in a situation called psychotherapy and have tendencies to behave as that situation specifies, not taking into account the special attributes of the client that may require some different behaviors.

(v) Learn how to initiate conversations which will teach you about the client's culture. For instance, ask questions like these: What do you think has caused your problem? When did it start? Who is likely to be involved in causing your problem? Why do you think it started when it did? What do you think your problem is doing to you? How does it do that? How severe is your difficulty? Will it have a short or a long course? What kinds of treatment do you think you should receive? What are the most important results which you hope to receive from this treatment? What are the chief side effects which your difficulty has caused you? What do you fear most about your condition?

From such questions the clinician can construct the client's "cultural model of the problem." Especially important is to obtain clear definitions of the meaning of terms, associations between terms and other terms, and the perceived antecedents and perceived consequences of the clients "problem." For instance, one can ask "In your culture, most people experience what conditions which result in the kind of problem you are having?" "In your culture, people who have the condition you are having experience what kinds of consequences?"

The cultural model that is identified may be quite different from any model encountered in the clinician's culture. To place it in context, it may be useful to inquire about religious beliefs, myths, norms, values, and the like which may have influenced the definition of the problem and the expected outcomes of the treatment.

(vi) Learn to see a positive aspect in every one of the client's cultural traits.

(vii) Learn to suspend judgment, to live with ambiguity, to categorize broadly (e.g., while the client's response is different from the typical response of most clients, it is not that different).

(viii) Learn how people react when they are confronted by differences in attitudes and values: they are likely to (a) ignore the difference; (b) bolster their position (I am obviously correct and you are stupid; they are likely to find additional arguments to strengthen their own position, or get social support from others who agree with them); (c) differentiate—it is okay for people in the other culture to do it that way, and for us to do it our way; and (d) transcend—both are correct under some conditions; our way is best under conditions X and their way is best under conditions Y. Once you know these four ways of reacting to what is different, try to suppress the first two and use the last one more frequently.

10.01.7.1.2 Self-insight

The emphasis in this kind of training is the development of an understanding of how culture affects one's behavior. This technique uses an actor to behave in the opposite way from the way which is prescribed by the culture of the trainee. For example, in training Americans, the actor behaves like a "contrast-American" (Stewart, 1966).

Topics include the American emphasis on material goals vs. stress on spiritual goals (e.g., as in India), the emphasis on achievement vs. ascribed status based on family prestige; Americans see competition as desirable, whereas in many cultures it is undesirable; Westerners in general emphasize planning, while traditional cultures emphasize fate; Westerners also place reliance on self, while in high power cultures reliance on superiors is more likely. Similarly, in low power distance cultures, the emphasis is on equality, while status differences are most important in high power distance cultures. Other contrasts: the belief that knowledge through observation is superior to knowledge received from authorities; the belief that thoughts cannot influence events vs. the belief that they can do so.

The actor interacts with the trainee and the session is videotaped. The trainer then spends time with the trainee going over the tape and explaining how the trainee's behavior is determined by culture.

This method is good because the trainee learns about his/her own culture. However, the trainee does not learn anything specific about another particular culture.

10.01.7.1.3 Experiential training

Experiential training involves bringing the trainees into contact with members of the host culture in situations where they can make mistakes which will not hurt long-term relationships.

For example, in the early 1960s, the Peace Corps used an exact replica of a South Pacific village (no electricity, no movies, no running water, etc.) in a valley on the Big Island of Hawaii (Brislin, 1993). Trainees spent several weeks learning the languages and customs of such villages while interacting with trainers from South Pacific cultures. Many trainees exposed to the rigors of that environment decided to drop out, saving the Peace Corps the expense of early repatriation and the embarrassment of failure on the job.

This method, however, is expensive, and depends on trial and error learning, which is slow. Nevertheless, it produces results if the training organization can spend the money and the trainees have the time.

10.01.7.1.4 Exposure to many local cultures

It is assumed that the more experience trainees have had in entering, learning about, and leaving cultures, the better equipped they will be to deal with new cultures. This type of training was developed by the late psychiatrist Wedge. It encourages the trainees to join different urban subcultures, such as the police, the fire department, unions, top management of corporations, school boards, religious groups, political groups, pickpockets, prostitutes, and so on. Each of these groups has a unique culture, and the skills required to join it should be valuable. One has to learn how to scout, enter, explore, terminate, evaluate the new skills and transfer them to the next culture. The trainer helps the trainee get in and out of each subculture, and discusses the experiences of the trainees once a week in review sessions where each trainee's problems in getting in and out of each culture are examined.

This method has a number of practical difficulties. It is difficult for a trainee to get into some of these groups (e.g., in one case, when a trainee wanted to join some stevedores they attacked him with a knife), and while there is much to be said for the process skills which one acquires, trainees do not learn much content about the host culture.

10.01.7.1.5 Field trips

The clinician may visit the culture of the client for some weeks to become familiar with the environment. This method teaches very little in depth about the local culture.

10.01.7.1.6 Culture assimilators or intercultural sensitizers

The introduction to this section mentioned the importance of learning to make isomorphic attributions (Triandis, 1975). In order to train people to make isomorphic attributions this technique was developed and validated (e.g., Fiedler, Mitchell, & Triandis, 1971) and found to work. It is the only cross-cultural training method that has been evaluated so far, with random assignment of trainees to experimental and control groups and has been shown to be effective (Albert, 1983). This does not mean that the other methods do not work; it only means that rigorous studies have not yet been reported.

The culture assimilator is a programmed learning approach to cultural training. It consists of a set of 100–200 critical incidents, that is, scenarios where people from two cultures interact. Each episode is followed by four or five explanations of why the member of the other culture has acted in a specific way. The trainee selects one of the explanations and is asked to turn to another page (or computer screen) where feedback is provided concerning the chosen explanation.

Let us look at a simple example. Suppose you are teaching a clinical psychologist to understand the behavior of a lower class Hispanic. If the psychologist understands the Hispanic's viewpoint, this is considered successful training.

Episode: A Hispanic lower class client looks down when spoken to.

Question: Why did the client look down?

Attributions:

(i) He was distracted. Turn to page 50.
(ii) He was fearful. Turn to page 51.
(iii) He was respectful. Turn to page 52.
(iv) He was hostile. Turn to page 53.

When trainees turn to pages 50, 51, or 53, they find negative feedback, along lines such as these: "No, this is incorrect; try another explanation." When the trainees turn to page 52, they get feedback such as: "Excellent! That is correct. When we presented this question to a sample of 80 Hispanics and 95 non-Hispanics, 85% of the Hispanics considered this answer to be the correct one, and only 36% of the non-Hispanic thought that this was correct."

Note that the construction of assimilator training is culture-specific. It requires the use of samples of people from the two cultures who study the episode and the attributions and select the attribution they consider to be correct. Thus, the training is validated as it is being constructed.

In addition, the feedback gives the percentages of samples who agree or disagree with each answer. This avoids teaching stereotypes. Instead, one learns that judgments about why people acted in a given way are probabilistic, and the probabilities differ across cultures.

Of course, in constructing assimilators it is necessary to have many more episodes than will be used in the final training, because some will not discriminate the two samples of subjects, and thus must be discarded.

Construction of the episodes is based on interviews with people who have wide experience in the two relevant cultures, or on research which has identified differences in the subjective culture of the two relevant samples. Any item of cultural difference can be included. For example, if the trainer wants to teach what it means to be *simpatico* she would have one correct option and several incorrect ones and let the trainee pick options and get feedback.

In training persons from an individualistic culture such as the US to go to a collectivist culture such as the Far East, one might include a variety of items such as:

(i) Differences in norms, such as do not bring a present of a certain color.

(ii) Differences in roles, such as the almost sacred parent–child role in some collectivist cultures (so that if one's spouse disagrees with one's parents, it is obligatory to take the side of one's parents in such a dispute).

(iii) Differences in the way behaviors can express intentions, so that "no" must be expressed most indirectly, and in a very subtle way, for instance during a visit by serving two substances (e.g., tea and bananas) which are usually not served together.

(iv) Differences in self-concepts, for example, collectivist selves are more likely to be appendages of groups than to be autonomous.

(v) Differences in what behaviors are valued, for example, that in collectivist cultures people use a modest introduction of their lecture, such as "what I have to tell you is not very important;" or they may say "I thank you for spending your valuable time with me." In individualist cultures they are more likely to present a self-assured front.

(vi) Differences in the kinds of associations which people make, for instance, in collectivist cultures people are more likely to associate the word "progress" with national rather than with personal progress.

(vii) Differences in the kinds of differentiations (sensitivity to hierarchies based on minor cues) which people typically make, for example, in collectivist cultures there will be more differentiations on the vertical axes than on the horizontal axes of social relationships.

(viii) Differences in the important determinants of behavior—in collectivist cultures behavior is more likely to reflect norms than attitudes.

(ix) Differences in the kinds of reinforcements which people expect for particular behaviors in particular situations—in collectivist cultures one is supposed to give a gift in many situations in which in individualist cultures one would pay.

Triandis, Brislin, and Hui (1988) advised individualists to pay attention to group attributes more than they do in their own culture, to learn more about the ingroups and outgroups of collectivists and to expect sharp differences in behavior when the collectivist interacts with members of such groups, to expect more harmony within the ingroup, for example, no criticism of high status people, than is typical of their own culture. They also suggested that trainees should cultivate long-term relationships, be modest when presenting a lecture, stress equality and need when distributing resources, and give more gifts than is customary in their own culture.

Construction of a typical assimilator begins with 200 or so episodes extracted from discussions with people who know both cultures, or by analyzing relevant ethnographies to identify a set of four or five attributions for each episode. Then the episodes and the attributions are presented to samples from the trainee's and the host's culture, and differences in the response patterns are examined. Those episodes which do not produce significant chi-squares are discarded. Then the assimilator is printed in a book or placed in a computer.

Administration can be done with each trainee working through the book or computer program alone, or the trainer may select some of the more interesting episodes and present them for discussion. Trainees may also role play the content of some of the episodes.

In principle, this method can be expanded for use with an interactive disk technology, so that trainees see videotaped actors performing the episode and participate by pushing buttons that indicate what responses they think are appropriate. Feedback from the computer in the form of praise or criticism can make this task most interesting.

In evaluations of this kind of training, trainees were assigned randomly to two groups: they received or did not receive the training (e.g., Weldon, Carlston, Rissman, Slobobin, and Triandis, 1975). Such studies have shown that trained people learn to make isomorphic attributions, they expand the range of explanations they give for specific behaviors, they become less ethnocentric, and develop more accurate expectations concerning appropriate behavior.

Assimilator training increases cognitive complexity. Cognitive complexity makes it possible to consider the subjective culture of the other cultural group as "valid" and thus lessens prejudice (Gardiner, 1972).

However, culture assimilator training does not strongly increase liking for the other group or reduce social distance toward it. Liking depends on the number of pleasant experiences one has shared with the other group. Just knowing how the other group thinks does not change emotions.

Stephan and Stephan (1984) used a culture assimilator to increase "knowledge of Chicano culture" among Anglo students in New Mexico. This study indicated that Anglo students' attitudes toward Chicanos in the Southwest of the US are related to three factors: (i) how much they know about Chicano culture, (ii) how much contact they have with Chicanos, and (iii) their parents' attitudes toward Chicanos. Contact with Chicanos was determined by the attitudes of their friends toward Chicanos. Finally, contact increased knowledge of Chicano culture, and the assimilator training boosted that knowledge a bit more, so that those who were trained had more positive attitudes than those who were not. The Stephans concluded their study as follows:

> Simple intergroup contact, such as the contact that typically occurs in desegregated schools, is not likely to improve intergroup relations. However, specially designed educational programs, designed to reduce ignorance of the outgroup, do appear to improve intergroup relations. (p. 249)

Social distance often depends on the norms of our own group. If our own group urges us to get along with the other group, we are more likely to do so than if it opposes such friendliness. As in the Stephans' study, when the friends favor contact, there is likely to be an improvement in relationships.

For these reasons culture assimilator training is only one component of cross-cultural training. It is not sufficient by itself. It needs to be supplemented with other methods, such as the self-insight method which is likely to increase the motivation of the trainees to learn about the other culture, and experiential training which can change emotions. Also, the norms of social interaction between the two cultures need to be changed to modify social distance.

Another limitation of assimilator training is that it does not change behavior. It is one thing to know how one is supposed to behave, and quite another to behave correctly. To achieve the latter goal one needs to have behavior modification training.

The development of assimilators is time-consuming. It is obvious that any one clinician cannot afford to develop an assimilator for particular clients. However, a group of clinicians, such as the APA's Division 12, could undertake the development of relevant assimilators.

In addition to assimilators which focus on a specific culture, there is a general assimilator (Brislin, Cushner, Cherrie, & Yong, 1986). It deals with the fact that it is natural and to be expected to: (i) feel anxious abroad, (ii) experience disconfirmed expectancies, (iii) be unable to feel that one belongs to the local culture, (iv) experience ambiguity about what one should do, and (v) be exposed to local prejudices. It teaches that one must learn to control one's own prejudices.

The general assimilator also teaches trainees to expect differences in the way people view: (i) work, (ii) the relationship of work and social interaction, (iii) time, (iv) space, (v) language, (vi) roles, (vii) groups, (viii) rituals, (ix) hierarchies, and (x) values.

It also helps the trainees to understand differences in: (i) categorization, (ii) differentiation (e.g., some people know more about some subjects than you do), (iii) the importance of the ingroup–outgroup distinction (e.g., between collectivists and individualists), (iv) learning styles (not everyone learns best the same way), and (v) attributions (how to make isomorphic attributions).

The general assimilator has been successfully evaluated (Cushner, 1989). Twenty-eight trained adolescents from the Pacific Rim visiting New Zealand were better adjusted than 22 control adolescents hosted by New Zealanders. The trained adolescents completed a number of tasks measuring cross-cultural sensitivity better than the control subjects.

A final point about assimilators: clinicians who have been warned that there will be problems when interacting with members of a specific culture are more likely to be able to deal with these problems than naive clinicians. This phenomenon is parallel to findings with hospital preoperative patients: those who were told that they will feel quite uncomfortable after an operation dealt better with the postoperative pain than those who had been told nothing. In general, the more the training creates realistic expectations about events, the better.

10.01.7.1.7 Behavior modification cross-cultural training

Behavior modification techniques require that people be rewarded for desirable behaviors and be made to do an incompatible behavior whenever they have the urge or tendency to do an undesirable behavior. For example, in Latin America it is customary among friends to give an *abrazo*, a kind of embrace with the arms. Also, holding hands, touching, and other behaviors between good friends of the same sex are common. Yet North Americans have to be trained to carry out such behaviors. If they behave as is expected by the Latinos they will be liked better when interacting with Latinos than if they behave as they do at home. If they do not behave as expected the locals will think of them as cold and undemonstrative "Gringos." But North Americans who behave this way without proper training will feel that they are behaving inappropriately.

Conversely, some behaviors have to be eliminated. For example, in some cultures (e.g., Greece), to show the palm of your open hand is an insult. It is called a "moutza" and reflects utmost contempt. However, many people are in the habit of greeting others by waving an open hand. For these people to stop doing this requires some training. Most people wave without thinking. A trainer must catch them in the act, and tell them not to open their hand. Self-correction after such acts can change the habit, so that the person waves with the palm closed and turned toward themself.

Especially difficult is to change habits such as nodding to say "no" and shaking the head to say "yes" (Bulgaria, Southern India). Here one needs a lot of practice to get over well-established habits.

10.01.7.1.8 Some summary points about training

Cross-cultural training can be effective. Black and Mendelhall (1990) reviewed 29 studies that measured effectiveness. They found that all the studies that measured how people felt about the training had positive results, all the studies which measured effective interpersonal relationships found improvements, all the studies which measured changes in perception found them, and all research which tried to reduce culture shock succeeded in doing so. Two-thirds of the studies which measured performance obtained improved performance.

In general these effects were found in field studies but not in laboratory studies. This may be due to the limited time one can use in the laboratory. A few hours of training are not enough to produce major changes. Furthermore, field studies involve volunteers who are motivated to succeed. If we tried to train the general public perhaps we would not have positive results. No matter how good the training or the trainer, if the trainees do not wish to change, they will not. Nevertheless, the studies with control groups and longitudinal designs reviewed by Black and Mendelhall (1990) show some impressive results which should encourage more cross-cultural training. A meta-analysis which statistically combines the results from many studies has also found that cross-cultural training is effective (Deshpande & Viswesvaran, 1992).

The scientific understanding of how culture affects social behavior and the use of this understanding to improve relationships across cultures has just begun, but it is a promising area of activity.

10.01.8 FUTURE DIRECTIONS

Future research should determine how many cultural syndromes are needed for an adequate understanding of cultural differences. Learning about these cultural syndromes should allow clinicians to improve their understanding of the cultures of their clients.

10.01.9 SUMMARY AND CONCLUSIONS

Culture penetrates our judgments and behaviors in subtle often uncontrollable ways. Becoming familiar with the extent culture shapes our perceptions, standards, and standard operating procedures can be useful. The aim of this chapter was to increase the sensitivity of the readers to the concepts which may help clinicians control the cultural dimensions of their interaction with their clients. It examined the major ways in which cultures differ and explored the attributes of methodologically sound studies which can determine precisely how cultures differ. It then reviewed ways in which clinicians can learn about the culture of their clients.

10.01.10 REFERENCES

Adler, N. (1991). *International dimensions of organizational behavior*. Boston: PWS-Kent Publishing.

Albert, R. D. (1983) The intercultural sensitizer or culture assimilator: A cognitive approach. In D. Landis & R. W. Brislin (Eds.), *Handbook of intercultural training* (Vol. 2, pp. 186–217). New York: Pergamon.

Albert, R. D. (1996). A framework and model for understanding Latin American and Latino/Hispanic cultural patterns. In D. Landis & R. S. Bhagat (Eds.), *Handbook of intercultural training* (2nd ed., pp. 327–348). Thousand Oaks, CA: Sage.

Barry, H. (1980). Description and uses of the Human Relations Area Files. In H. C. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 445–478). Boston: Allyn and Bacon.

Berry, J. W. (1990). The psychology of acculturation. In *Nebraska Symposium on Motivation, 1989* (pp. 201–234). Lincoln, NE: University of Nebraska Press.

Berry, J. W., Kim, U., Power, S., Young, M., & Bujaki, M. (1989). Acculturation attitudes in plural societies. *Applied Psychology, 38,* 185–206.

Berry, J. W., Poortinga, Y., Segall, M., & Dasen, P. (1992). *Cross-cultural psychology*. New York: Cambridge University Press.

Black, J. S., & Mendenhall, M. (1990). Cross-cultural training effectiveness: A review and theoretical framework for future research. *Academy of Management Review, 15,* 113–136.

Bochner, S. (1980). Unobtrusive methods in cross-cultural experimentation. In H. C. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 319–388). Boston: Allyn and Bacon..

Bogardus, E. S. (1925). Measuring social distance. *Journal of Applied Sociology, 9,* 299–308.

Bond, M. H., & Cheung, M. (1984). Experimenter language choice and ethnic affirmation by Chinese trilinguals in Hong Kong. *International Journal of Intercultural Relations, 8,* 347–356.

Bond, M., & Smith, P. B. (1996). Cross-cultural social and organizational psychology. *Annual Review of Psychology, 47,* 205–235.

Bond, M., & Young, K. S (1982). Ethnic affirmation vs cross-cultural accommodation: The variable impact of questionnaire language on Chinese bilinguals in Hong Kong. *Journal of Cross-Cultural Psychology, 13,* 169–185.

Bond, R., & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin, 119,* 111–137.

Boykin, A. W. (1983). The academic performance of Afro-American children. In J. Spence (Ed.), *Achievement and achievement motives* (pp. 328–371) New York: Freeman.

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin, 17,* 475–485.

Brewer, M., & Campbell, D. T. (1976). *Ethnocentrism and intergroup attitudes: East African evidence*. New York: Halsted/Wiley.

Brislin, R. W. (1980). Translation and content analysis of oral and written materials. In H. C. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 389–444). Boston: Allyn and Bacon.

Brislin, R. (1993). *Understanding cultures influence on behavior*. Fort Worth, TX: Harcourt Brace Jovanovich.

Brislin, R., Cushner, K., Cherrie, C., & Yong, M. (1986). *Intercultural interactions: A practical guide*. Beverly Hills, CA: Sage.

Brown, D. E. (1991). *Human universals*. Philadelphia: Temple University Press.

Brown, E. D., & Secrest, L. (1980). Experiments in cross-cultural research. In H. C. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 297–318). Boston: Allyn & Bacon.

Burton, M. L., Moore, C. C., Whiting, J. W. M., & Romney, A. K. (1996). Regions based on social structure. *Current Anthroplogy, 37,* 87–123.

Campbell, D. T. (1988). *Methodology and epistemology for social science*. Chicago: University of Chicago Press.

Campbell, D. T., & Levine, R. (1968). Ethnocentrism and intergroup relations. In R. P. Abelson et al. (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 551–564). Chicago: Rand McNally.

Campbell, J. K. (1964). *Honor, family and patronage*. Oxford: Clarendon Press.

Carter, R. T. (1995). *The influence of race and racial identity in psychotherapy: Toward a racially inclusive model*. New York: Wiley.

Ciborowski, T. (1980). The role of context, skill and transfer in cross-cultural experimentation. In H. C. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 279–296). Boston: Allyn & Bacon.

Cohen, D. (1996). Law, social policy, and violence: The impact of regional cultures. *Journal of Personality and Social Psychology, 70,* 961–978.

Cohen, D., & Nisbett, R. E. (1994). Self-protection and the culture of honor: Explaining Southern violence. *Personality and Social Psychology Bulletin, 20,* 551–567.

Cohen, D., Nisbett, R. E., Bowdle, B. F., & Schwartz, N. (1996). *Journal of Personality and Social Psychology, 70,* 945–960.

Cook, P. (1994). Chronic illness beliefs and the role of social networks among Chinese, Indian, and Angloceltic Canadians. *Journal of Cross-Cultural Psychology, 25,* 452–465.

Cushner, K. (1989) Assessing the impact of a culture-general assimilator. *International Journal of Intercultural Relations, 13,* 125–146.

Dalby, L. C. (1983). *Geisha.* Berkeley, CA: University of California Press.

Deshpande, S. P., & Viswesvaran, C. (1992). Is cross-cultural training of expatriate managers effective: A meta analysis. *International Journal of Intercultural Relations, 16,* 295–310.

Devine, P. G. (1989). Stereotypes and prejudice: The automatic and control components. *Journal of Personality and Social Psychology, 56,* 5–18.

Diaz-Guerrero, R. (1979). The development of coping style. *Human Development, 22,* 320–331.

Doi, T. (1986). *The anatomy of conformity: The individual versus society.* Tokyo: Kadansha.

Draguns, J. (1988). Personality and culture: Are they relevant for the enhancement of quality of mental life? In P. R. Dasen, J. W. Berry, & N. Sartorius (Eds.), *Health and cross-cultural psychology* (pp. 141–161). Newbury Park, CA: Sage.

Draguns, J. (1990). Normal and abnormal behavior in cross-cultural perspective: Specifying the nature of their relationship. In J. Berman (Ed.), *Nebraska symposium on motivation, 1989* (pp. 235–278). Lincoln, NE: University of Nebraska Press.

Edwards, A. L. (1957). *Techniques of attitude scale construction.* New York: Appleton-Century-Crofts.

Fiedler, F. E., Mitchell, T., & Triandis, H. C. (1971). The culture assimilator: An approach to cross-cultural training. *Journal of Applied Psychology, 55,* 95–102.

Fiske, D. W. (1986). Specificity of method and knowledge in social science. In D. W. Fiske & R. A. Shweder (Eds.), *Metatheory in social science* (pp. 61–82). Chicago: University of Chicago Press.

Fiske, D. W., & Shweder, R. A. (1986). *Metatheory in social science.* Chicago: University of Chicago press.

Gardiner, G. S. (1972). Complexity training and prejudice reduction. *Journal of Applied Social Psychology, 2,* 326–342.

Goodenough, W. H. (1980). Ethnographic field techniques. In H. C. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 29–56). Boston: Allyn & Bacon.

Haraway, D. J. (1991). *Simians, cyborgs, and women.* New York: Routledge.

Harding, S. (1987) (Ed.). *Feminism and methodology.* Bloomington, IN: Indian University Press.

Herskovits, M. J. (1955). *Cultural anthropology.* New York: Knopf.

Hofstede, G. (1980). *Culture's consequences.* Beverly Hills, CA: Sage.

Holtzman, W. H. (1980). Projective techniques. In H. C. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 245–278). Boston: Allyn & Bacon.

Hsu, F. L. K. (1981). *Americans and Chinese: Passage to differences.* Honolulu, HW: University of Hawaii Press.

Hui, C. H., & Triandis, H. C. (1985). Measurement in cross-cultural psychology: A review and comparison of strategies. *Journal of Cross-Cultural Psychology, 16,* 131–152.

Hui, C. H., & Triandis, H. C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural Psychology, 20,* 296–309.

Hulin, C. L., Drasgow, F., & Parsons, C. K. (1983). *Item response theory: Applications to psychological measurement.* Homewood, IL: Irwin.

Irvine, S., & Carroll, W. K. (1980). Testing and assessment across cultures: Issues in methodology and theory. In H. C. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 181–244). Boston: Allyn & Bacon.

Jones, J. M. (1986). Racism: A cultural analysis. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination and racism.* Orlando, FL: Academic Press.

Jones, J. M. (1988). Racism in black and white. In P. A. Katz & D. A. Taylor (Eds.), *Eliminating racism* (pp. 117–135). New York: Plenum.

Kim, U., Triandis, H. C., Choi, S. Kagitcibasi, C., & Yoon, G. (1994). *Individualism and collectivism: Theory, method and applications.* Thousand Oaks, CA: Sage.

Kluckhohn, C. (1954). Culture and behavior. In G. Lindzey (Ed.), *Handbook of social psychology* (Vol. 2, pp. 921–976). Cambridge, MA: Addison-Wesley.

Kluckhohn, F., & Strodtbeck, F. (1961). *Variations in value orientations.* Evanston, IL: Row Peterson.

Kroeber, A. L., & Kluckhohn, C. (1952). *Culture: A critical review of concepts and definitions* (Vol. 147, No. 1). Cambridge, MA: Peabody Museum.

Kruger, J., & Clement, R. W. (1994). The truly false consensus effect: An ineradicable egocentric bias in social perception. *Journal of Personality and Social Psychology, 67,* 596–610.

Lambert, W. E. (1992). Challenging established views on social issues: The power and limitations of research. *American Psychologist, 47,* 533–542.

Landis, D., & Brislin, R. (1983). *Handbook of intercultural training* (in 3 vols). New York: Pergamon.

Landis, D., & Bhagat, R. S. (1996). *Handbook of intercultural training* (2nd ed.). Thousand Oaks, CA: Sage.

Landis, D., McGrew, P., Day, H., Savage, J., & Saral, T. (1976). Word meaning in black and white. In H. C. Triandis (Ed.), *Variations in black and white perceptions of the social environment* (pp. 45–80). Urbana, IL: University of Illinois Press.

Langer, E. J. (1983). *The psychology of control.* Beverly Hills, CA: Sage.

Leung, K. (1989). Cross-cultural differences: Individual-level vs. cultural level analysis. *International Journal of Psychology, 24,* 703–719.

Loevinger, J. (1976). *Ego development.* San Francisco: Jossey-Bass.

Loevinger, J. (1979). The idea of the ego. *Counseling Psychologist, 8,* 3–5.

Longabaugh, R. (1980). The systematic observation of behavior in naturalistic settings. In H. C. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 57–126). Boston: Allyn & Bacon.

Lonher, W., & Berry, J. (1986). *Field methods in cross-cultural research.* Beverly Hills, CA: Sage.

Malpass, R. S. (1977). Theory and method in cross-cultural psychology. *American Psychologist, 32,* 1069–1079.

Malpass, R. S., & Poortinga, Y. H. (1986). Strategies for design and analysis. In W. J. Lonner & J. W. Berry (Eds.), *Field methods in cross-cultural research* (pp. 47–84). Beverly Hills, CA: Sage.

Majors, R., & Mancini-Billson, J. (1992). *Cool pose.* New York: Lexington Books.

Marin, G., & Marin, B. V. (1991). *Research with Hispanic populations.* Newbury Park, CA: Sage.

Marin, G., Gamba, R. J., & Marin, B. V. (1992). Extreme response style and acquiescence among Hispanics. *Journal of Cross-Cultural Psychology, 23,* 498–509.

Marin, G., & Triandis, H. C. (1985). Allocentrism as an important characteristic of the behavior of Latin Americans and Hispanics. In R. Diaz-Guerrero (Ed.), *Cross-cultural and national studies in social psychology* (pp. 85–104). Amsterdam: North-Holland.

Marin, G., Triandis, H. C., Betancourt, H., & Kashima, Y. (1983). Ethnic affirmation versus social desirability: Explaining discrepancies in bilinguals' responses to a questionnaire. *Journal of Cross-Cultural Psychology, 14,* 173–186.

Markus, H., & Kitayama, S. (1991). Culture and self:

Implications for cognition, emotion and motivation. *Psychological Review, 98,* 224–253.

McClelland, D. C. (1980). Motive dispositions: The merits of operant and respondent measures. In L. Wheeler (Ed.), *Review of personality and social psychology* (pp. 10–41). Beverly Hills, CA: Sage.

Naroll, R., Michik, G. L., & Naroll, F. (1980). Holocultural research methods. In H. C. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 479–522). Boston: Allyn & Bacon.

Nielsen, J. (1990). *Feminist research methods.* Boulder, CO: Westview Press.

Nisbett, R. E., & Cohen, D. (1996). *The culture of honor.* Boulder, CO: Westview Press.

Oberg, K. (1960). Culture shock: Adjustment to new cultural environments. *Practical Anthropology, 7,* 177–182.

Pareek, U., & Rao, T. V. (1980). Cross-cultural surveys and interviewing. In H. C. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 127–180). Boston: Allyn & Bacon.

Parsons, T. & Shils, E. A. (1951). *Toward a general theory of action.* Cambridge, MA: Harvard University Press.

Pike, K. L. (1967). *Language in relation to a unified theory of the structure of human behavior.* The Hague: Mouton.

Phillips, H. P. (1965). *Thai peasant personality: The patterning of interpersonal behavior in the village of Bang Chan.* Berkeley, CA: University of California Press.

Qureshi, B. (1989). *Transcultural medicine: Dealing with patients from different cultures* (2nd ed.). Dordrecht: Kluwer.

Reinharz, S. (1992). *Feminist methods in social research.* New York: Oxford University Press

Rhee, E., Uleman, J. S., Lee, H. K., & Roman, R. J. (1995). Spontaneous self-descriptions and ethnic identities in individualistic and collectivist cultures. *Journal of Personality and Social Psychology, 69,* 142–152.

Robinson, R. J., Keltner, D., Ward, A., & Ross, L. (1995). Actual versus assumed differences in construal: "Naive realism" in intergroup perception and conflict. *Journal of Personality and Social Psychology, 68,* 404–417.

Rohner, R. P. (1986). *The warmth dimension: Foundations of parental acceptance-rejection theory.* Newbury Park, CA: Sage.

Schwartz, S. H. (1994). Beyond individualism/collectivism: New cultural dimensions of values. In U. Kim, H. C. Triandis, C. Kagitcibasi, S.-C. Choi, & G. Yoon (Eds.), *Individualism and collectivism: Theory, method and applications* (pp. 85–122). Thousand Oaks, CA: Sage.

Segalowitz, N. S. (1980). Issues in the cross-cultural study of bilingual development. In H. C. Triandis & A. Heron (Eds.), *Handbook of cross-cultural psychology: developmental* (Vol. 4, pp. 55–92). Boston: Allyn & Bacon.

Shweder, R. A., Mahapatra, M. & Miller, J. G. (1990). Culture and moral development. In J. W. Stigler, R. A. Shweder, & G. Herdt (Eds.), *Cultural psychology* (pp. 130–204). New York: Cambridge University Press.

Shweder, R. A., & LeVine, R. (1984). *Culture theory.* Chicago: University of Chicago Press.

Skinner, B. F. (1981). Selection by consequences. *Science, 213,* 501–504.

Sussman, N. M., & Rosenfeld, H. M. (1982). Influence of culture, language, and sex on conversational distance. *Journal of Personality and Social Psychology, 42,* 66–74

Stephan, W. G., & Stephan, C. W. (1984). The role of ignorance in intergroup relations. In N. Miller & M. B. Brewer (Eds.), *Desegregation: Groups in contact* (pp. 229–256). New York: Academic Press.

Strodtbeck, F. (1951). Husband–wife interaction over revealed differences. *American Sociological Review, 16,* 468–473.

Stewart, E. (1966). The simulation of cultural differences. *Journal of Communication, 16,* 291–304.

Szalay, L. B. (1970). *A communication dictionary of cultural meanings.* Washington, DC: American Institutes for Research.

Szalay, L. B. (1985). Psychocultural findings. In H. W. Sinaiko, P. M. Curran, B. T. King, & J. M. Schneider (Eds.), *Hispanic subpopulations and Naval service.* Technical Report SI/MRAS/TR-11. Washington, DC: Office of Naval Research.

Szalay, L. B., & Deese, J. (1978). *Subjective meaning and culture: An assessment through word associations.* New York: Wiley.

Tajfel, H. (1982). *Social identity and intergroup relations.* New York: Cambridge University Press.

Tapp, J. L., Kelman, H. C., Triandis, H. C., Wrightsman, L., & Coelho, G. (1974). Continuing concerns in cross-cultural ethics: A report. *International Journal of Psychology, 9,* 231–249.

Thernstrom. S. (1980). *Harvard encyclopedia of American ethnic groups.* Cambridge, MA: Harvard University Press.

Thurstone, L. L. (1931). The measurement of social attitudes. *Journal of Abnormal and Social Psychology, 26,* 249–269.

Ting-Toomey, S. (1988). A face-negotiation theory. In Y. Kim & W. Gudykunst (Eds.), *Theories of intercultural communication.* Newbury Park, CA: Sage.

Thibaut, J., & Kelley, H. (1959). *The social psychology of groups.* New York: Wiley

Triandis, H. C. (1972). *The analysis of subjective culture.* New York: Wiley.

Triandis, H. C. (1975). Cultural training, cognitive complexity, and interpersonal attitudes. In R. W. Brislin, S. Bochner, & W. J. Lonner (Eds.), *Cross-cultural perspectives on learning* (pp. 39–77). Beverly Hills, CA: Sage.

Triandis, H. C. (1976a). The future of pluralism. *Journal of Social Issues, 32,* 179–208.

Triandis, H. C. (1976b). *Variations in black and white perceptions of the social environment.* Urbana, IL: University of Illinois Press.

Triandis, H. C. (1989). Self and social behavior in differing cultural contexts. *Psychological Review, 96,* 269–289.

Triandis, H. C. (1990). Cross-cultural studies of individualism and collectivism. In J. Berman (Ed.), *Nebraska symposium on motivation, 1989* (pp. 41–133). Lincoln, NE: University of Nebraska Press.

Triandis, H. C. (1992). Cross-cultural research in social psychology. In D. Granberg & G. Sarup (Eds.), *Social judgment and intergroup relations: Essays in honor of Muzafer Sherif* (pp. 229–244). New York: Springer.

Triandis, H. C. (1994). *Culture and social behavior.* New York: McGraw-Hill.

Triandis, H. C. (1995). *Individualism and collectivism.* Boulder, CO: Westview Press.

Triandis, H. C. (1996). The psychological measurement of cultural syndromes. *American Psychologist, 51,* 407–415.

Triandis, H. C., & Berry J. W. (1980). *Handbook of cross-cultural psychology* (Vol. 2). Boston: Allyn and Bacon.

Triandis, H. C., Bontempo, R., Leung, K., & Hui, C. H. (1990). A method for determining cultural, demographic, and personal constructs. *Journal of Cross-Cultural Psychology, 21,* 302–318.

Triandis, H. C., Brislin, R. & Hui, C. H. (1988). Cross-cultural training across the individualism–collectivism divide. *International Journal of Intercultural Relations, 12, 269–289.*

Triandis, H. C., & Gelfand, M. (in press). Converging measurement of horizontal and vertical individualism and collectivism. *Journal of Personality and Social Psychology.*

Triandis, H. C., Kashima, Y., Shimada, E., & Villareal, M.

(1986). Acculturation indices as a means of confirming cultural differences. *International Journal of Psychology, 21,* 43–70.

Triandis & Marin (1983).

Triandis, H. C., Marin, G. Lisansky, J., & Betancourt, H. (1984). *Simpatia* as a cultural script of Hispanics. *Journal of Personality and Social Psychology, 47,* 1363–1375.

Triandis, H. C., McCusker, C., & Hui, C. H. (1990). Multimethod probes of individualism and collectivism. *Journal of Personality and Social Psychology, 59,* 1006–1020.

Triandis, H. C., & Triandis, L. M. (1962). A cross-cultural study of social distance. *Psychological Monographs, 76*(21) (Whole issue of No. 540).

Triandis, H. C., Vassiliou, V., & Nassiakou, M. (1968). Three cross-cultural studies of subjective culture. *Journal of Personality and Social Psychology, Monograph Suppl. 8,* No, 4, 1–42.

Trilling, L. (1972). *Sincerity and authenticity*. London: Oxford University Press.

Weldon, D. E., Carlston, D. E., Rissman, A. K., Slobodin, L., & Triandis, H. C. (1975). A laboratory test of effects of culture assimilator training. *Journal of Personality and Social Psychology, 32,* 300–310.

Werner, O., & Campbell, D. T. (1970). Translation, working through interpreters and the problem of decentering. In R. Naroll & R. Cohen (Eds.), *A handbook of method in cultural anthropology* (pp. 398–422). New York: American Museum of Natural History.

Yang, K-S., & Bond, M. H. (1980). Ethnic affirmation by Chinese bilinguals. *Journal of Cross-Cultural Psychology, 11,* 411–425.

Zane, N. W., Sue, S., Hu, L., & Kwon, J.-H. (1991). Asian-American assertion: A social learning analysis of cultural differences. *Journal of Counseling Psychology, 38,* 63–70.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.02
Cross-cultural Psychopathology

FANNY M. CHEUNG
The Chinese University of Hong Kong, Hong Kong

10.02.1 INTRODUCTION 35
10.02.2 ETHNOCENTRIC APPROACH OF CLINCIAL PSYCHOLOGY 36
10.02.3 CULTURE-BOUND SYNDROMES 37
10.02.4 KORO 38
10.02.5 ALTERNATIVE VIEWS ON CBS 38
10.02.6 TRANSFORMATION OF NEURASTHENIA 39
10.02.7 ILLNESS EXPERIENCE 41
10.02.8 SOMATIZATION AS A CATEGORY FALLACY 41
10.02.9 SOMATIZATION AS ILLNESS EXPERIENCE 42
10.02.10 CULTURE AND PSYCHOPATHOLOGY 44
  10.02.10.1 Role of Culture in Psychopathology 44
    10.02.10.1.1 Culture-produced stress 44
    10.02.10.1.2 Culture-related problems 44
    10.02.10.1.3 Culture-inherited vulnerability 45
    10.02.10.1.4 Cultural contribution to choice of psychopathology 45
  10.02.10.2 Cultural Explanations 45
10.02.11 CROSS-CULTURAL PSYCHOPATHOLOGY 45
  10.02.11.1 Universalist vs. Relativist Approaches 45
  10.02.11.2 Recent Research on Cross-cultural Psychopathology 47
  10.02.11.3 Future Directions 48
10.02.12 SUMMARY 48
10.02.13 REFERENCES 48

10.02.1 INTRODUCTION

While psychopathology is embedded in the wider cultural context, theories of clinical psychology have generally been ethnocentric in nature. Cross-cultural studies in psychopathology often started with the assumption of the universality of these theories and attempted to compare the similarities and differences between cultural groups on aspects of these theories. Earlier interests in culture-specific aspects of psychopathology were related to peripheral curiosities about bizarre phenomena in exotic cultures. Because of their peripheral status, culture-bound syndromes were identified without much systematic research into the cultural dynamics of these syndromes. While they may have clinical meaning to the specific cultures, there is little relevance to the mainline theories. In this chapter, some of the culture-bound syndromes are demystified. The roles played by cultural factors in the presentation and interpretation of psychopathology are examined. Recent research is incorporating the universalist and relativist approaches of cross-cultural psychology and bringing cultural dimensions of psychopathology into the mainstream.

10.02.2 ETHNOCENTRIC APPROACH OF CLINCIAL PSYCHOLOGY

Clinical psychology has always considered itself to be a scientific field and has placed a strong emphasis on research (Routh, 1994). The major content areas of research include psychopathology, assessment, and intervention. Despite the empirical orientation of psychology, a major deficiency in the prominent models of psychopathology is the lack of cross-cultural awareness in both theory and research. The ethnocentric approach in Anglo-American clinical psychology may be reflected in the emphasis on the individual or the intrapsychic self, be it the biological paradigm, the psychoanalytic paradigm, the learning paradigm, the cognitive paradigm, or the phenomenological paradigm. In none of these models is the wider cultural context considered to be an important factor in understanding the cause, course, and consequences of psychopathology.

Lewis-Fernandez and Kleinman (1994) pointed out three culture-bound assumptions which biased the professional concepts of mental health and illness in North America. The first is the "egocentricity of the self" which is seen as "a self-contained, autonomous entity, characterized by a unique configuration of internal attributes that determine behavior" (p. 67). The focus on the self, individual experience, and internal attributes of personality in understanding psychopathology misses the sociocentric ideology and interpersonal contexts which most of the rest of the world espouses.

The second culture-bound assumption is the philosophical roots of mind–body dualism, which leads to the division of psychopathology into the organic disorders and psychological problems. This dualistic professional model would misinterpret the cultural expressions of patients in most parts of the world who experience human suffering as simultaneous mind and body distress. The integrated somatopsychological idioms of expressions are often viewed as a reflection of a lack of introspection or psychological-mindedness.

The third culture-bound assumption that biases professional concepts is the view that culture is "an arbitrary superimposition on a knowable biological reality" (p. 67). This assumption views culture as "epiphenomenal." Cultural beliefs about disease categories, illness experiences, and healing practices of people in other cultures are reduced to the status of misinformed or superstitious obstacles in the diagnosis, treatment, and outcome of real diseases which are based on an invariant reality of biology.

These culture-bound assumptions result in an ethnocentric psychology based on "culturally specific ways of viewing individuals and their personality development," "diagnostic categories of personality disorder that ignore the fundamental influence of social context and cultural norms on human behavior," and "a psychiatric nosology that claims to be universal but does not take seriously the great cross-cultural diversity of somatic and psychological symptoms" (Lewis-Fernandez & Kleinman, 1994, p. 67).

The lack of a cross-cultural perspective in clinical psychology cannot be explained purely on the grounds of the limited diversity of the ethnic backgrounds of trained clinical psychologists. The ethnic backgrounds of clinical psychologists in the USA have diversified in the past two decades and the population of clinical psychologists in different parts of the world has vastly expanded. The little research found in cross-cultural clinical psychology has focused mainly on applicability of Western or Caucasian middle-class models of treatment or assessment procedures.

In contrast, since the 1960s the field of psychiatry has developed a substantial, albeit peripheral, interest in cultural psychiatry, raising sensitivity to cultural aspects of psychopathology and treatment. It has been pointed out that modern psychiatry was "moulded by specifically Western philosophical and scientific traditions" and should overcome the "inertia of ethnocentricism" (Lin, 1982, p. 235). Large-scale international collaborative research in psychopathology and diagnosis is evident in the activities of international psychiatric organizations and publications of cultural psychiatry. For example, a sociocultural lexicon for psychiatry (Mental Health Division, WHO, in press) will be published to accompany the tenth revision of the *International Classification of Diseases* (*ICD-10*; World Health Organization, 1992). Global collaborative research projects have been set up to compare the etiology, symptomatology and treatment of specific psychiatric disorders such as schizophrenia (World Health Organization, 1973, 1975, 1979, 1991).

Within American psychiatry, there are attempts to make the role of cultural analysis more of a mainstream than a secondary activity. Where the *Diagnostic and statistical manual of mental disorder* (*3rd ed.*) (*DSM-III* American

Psychiatric Association, 1980) has been criticized for ignoring cultural diversity, efforts have been made to represent cultural concerns in *DSM-IV* (American Psychiatric Association, 1994; Good, 1996). A Group on Culture and Diagnosis consisting of anthropologists and cross-cultural psychiatrists was formed to advise the *DSM-IV* Task Force on how to make culture more central to *DSM-IV*. One of the specific aims of the Group was "to devise a mechanism that would facilitate the application of a cultural perspective to the process of clinical interviewing and diagnostic formulation in psychiatry" (Lewis-Fernandez, 1996, p. 133). The Group on Culture and Diagnosis proposed an outline for cultural formulation to supplement multiaxial diagnostic assessment in a multicultural environment. As summarized by Lewis-Fernandez:

> the cultural formulation provides a systematic review of the individual's cultural background, the role of the cultural context in the expression and evaluation of symptoms and dysfunction, and the effect that cultural differences may have on the relationship between the individual and the clinician. (p. 137)

Eventually, however, only an edited and shortened version of this outline was included as an appendix instead of the central text of *DSM-IV*. Nevertheless, compared with previous editions, cultural concerns are represented in a much more significant manner in the text of *DSM-IV* and are included in the introduction, in the context of particular categories, as well as in a glossary of cultural terms (Good, 1996).

No such systematic approach to mainstream cultural analysis is found in the field of clinical psychology. Therefore, reference to the knowledge base developed in cultural psychiatry is needed to review research in cross-cultural psychopathology and culture-bound syndromes.

10.02.3 CULTURE-BOUND SYNDROMES

The interest in culture-bound syndromes (CBS) in cultural psychiatry in itself illustrates a process of transformation in the approach to cultural dimensions of psychopathology. Fascination with CBS has its sources in the "bizarre characteristics of the behavior displayed" (Hughes, 1985). Originating from the voyeuristic interest on folk illnesses in "exotic" cultures, CBS are now included in the glossary of cultural terms in *DSM-IV*. The study of culture-bound syndromes has gained respectability from the scholarly works of pioneers such as Yap who alleged that:

> It has long been known that there are, in certain cultural groups, peculiar aberrations of behavior which are regarded by themselves as abnormal. Over the years a number of terms taken from indigenous languages have crept into the psychiatric literature to denote these conditions, but many of them do not point to novel or distinct forms of disorder unknown elsewhere. Some are simply generic terms for "mental disorders" without definite meaning, others refer only to healing rituals, and still others to supernatural notions of disease causation. ... To avoid stagnation in this field, it is essential to apply the concepts of clinical psychopathology to the analysis of these disorders, to integrate them into recognized classifications of disease if possible, or to broaden the classification if necessary. (1974, p. 86)

Initial interests in CBS have given rise to a host of "folk" taxonomies from different cultures (Simons & Hughes, 1985). More common CBS include the following:

(i) *Latah*. Originating from Malaysia and Indonesia, *Latah* refers to the exaggerated startle response followed by odd behaviors. The afflicted person typically responds to a frightening stimulus with an exaggerated startle or jump, sometimes throwing or dropping a held object, uttering some improper word, or matching the words or movements of people nearby (Simons, 1985, p. 43). The pattern is a highly stereotypic, culturally labeled state which, though contravening the social norm, is differentiated from insanity.

(ii) *Amok*. *Amok* refers to the sudden mass assault taxon indigenous to Malayo-Indonesians, but may find parallel instances of indiscriminate homicide in other parts of the world. It is defined as "an acute outburst of unrestrained violence associated with homicidal attack, preceded by a period of brooding, and ending with exhaustion and amnesia" (Carr, 1985, p. 199). It is believed to have originated from the cultural training for warfare of the early Javanese and Malays which was intended to terrify the enemy into believing that they could expect no mercy and could save themselves only by flight.

(iii) *Pibloktoq*. The term is also labeled as arctic hysteria, polar hysteria, or transitional madness, referring to a group of hysterical symptoms among the Polar Eskimo of Greenland brought on by the depressing or monotonous effects of the Arctic's winter climate. The syndrome is composed of a series of reactive patterns with different combinations of the features in different cases. Disturbance of consciousness during the seizure and amnesia for the attack are the central clinical features. Some frequent behavioral symptoms are tearing off of clothing resulting in nudity, glossolalia, fleeing, and running across or rolling in snow. Gussow (1985) suggested that *Pibloktoq* is a reaction of "the basic Eskimo personality" to "situations of unusually intense, but culturally typical stress" (p. 282).

(iv) *Susto*. Illness arising from spirit separating from the body due to fright is widespread in one form or another in Latin America. The folk interpretation of *Susto* is based on the belief that a person comprises spiritual and organic elements. The spiritual element may be detachable from the host organism under unsettling experiences that disturb the equilibrium or when the unsettled spiritual element is seized and controlled by the spirits of the natural environment. The situation results in loss of appetite, weight, and strength; restlessness in sleep; listlessness when awake; and depression and introversion. "For recovery, the captured spirit must be retrieved from its captor, ransomed in the folk treatment process, and 'led back' to be reincorporated into the victim's body" (Rubel, O'Nell, & Collado, 1985, p. 334).

Most of the initial interest in CBS has surrounded the description of these phenomena. According to Prince and Tcheng-Laroche's (1987) definition of CBS, the culturally distinct meaning of the illness for individuals and cultures affects the occurrence of signs and symptoms of diseases in some cultures but not in others. The difference in occurrence depends on the psychosocial features of those cultures. CBSs have also been called culture-reactive syndromes, which suggests that the psychopathology is deemed to be rooted in social environmental forces and/or situations. The signs and symptoms reflect an attempt by the patients to adapt to major problems in their lives.

In the following section, one of the more popular syndromes in the literature, *Koro*, will be discussed in greater detail to illustrate how the cultural context of CBS affects their manifestation.

10.02.4 KORO

Koro arises from the belief about the retraction of the sexual organs (including penis, breast, nipples) into the body which leads to eventual death. Koro was believed to originate from the Malay word *koro* which means "shrink." An alternative source is from the word *kura* which means a tortoise. The "head of the tortoise" is often used as an expression for the penis by the Malays as well as the Chinese (Edwards, 1985). A similar condition is also found among Chinese under the name of Suoyang, another native terminology for genital retraction complaints. The syndrome was considered an emic or locally defined construct unique to these cultures with culture-specific etiologies.

Cultural meanings have been offered to explain individual and mass episodes of Koro. For example, Bartholomew (1994) argued that Koro is "a rational attempt at problem-solving that involves conformity dynamics, perceptual fallibility, and the local acceptance of koro-associated folk realities" (p. 46). Gwee (1985) offers the interpretation that Koro is an "acute hysterical panic reaction, brought on by auto- or hetero-suggestion and conditioning by the cultural background" (p. 159). The indigenous ethnographic literature describes Koro attacks as "unpredictable, but [they] usually appear after a shock which made the patient anxious or frightened, after performing strenuous manual labor or no labor at all, or as a result of immoderate nocturnal partying" (Edwards, 1985, p. 171). The etiology is attributed to nonconformity with community norms in the cultural context of competitiveness. In Chinese societies, these fears are associated with beliefs about sexual activities. Suoyang appears as "a product of intra- and inter-personal violations of medico-sexual regulations" (Edwards, p. 183).

Originally believed to be indigenous to Southeast Asia, Koro has also been detected in Africa, South Asia, and North America since the late 1970s (Chowdhury, 1996; Fishbain, Barsky, & Goldberg, 1989; Holden, 1987). Attempts have been made to fit Koro into the *etic* or externally defined psychiatric nosology, including acute hysterical panic reaction, anxiety states, depersonalization, conversion disorder, or atypical somatoform disorder (Simons & Hughes, 1985). Hughes (1985) justified the application of conventional diagnostic concepts to culture-bound phenomena in that they provided "the basis for more informed judgment regarding the ontological status and (eventually) etiology of the culture-bound syndromes" (p. 21). He went on to point out that a culturally informed use of the conventional diagnostic categories required knowing the culture of the person being evaluated to make a valid assessment of the behaviors, as well as being skeptical of the implicit values and assumptions underlying these categories. He warned that these categories should be tested against the observed primary behavior patterns and not be taken for granted.

10.02.5 ALTERNATIVE VIEWS ON CBS

In a reversed direction in the search for culture-bound syndromes, some Western-based syndromes such as anorexia nervosa, obesity, or even adolescence have been postulated as Western cases of CBS (Hill & Fortenberry, 1992; Prince, 1985; Ritenbaugh, 1982; Swartz, 1985). It was argued that anorexia nervosa fits the criteria of CBS, namely, that "the syndrome cannot be understood apart from its specific cultural or subcultural context," that "the etiology summarizes and symbolizes core meanings and behavioral norms of that culture," that "diagnosis relies on culture-specific technology as well as ideology," and that "successful treatment is accomplished only by participants in that culture" (Swartz, p. 725). Notwithstanding the need for empirical evidence to demonstrate that anorexia nervosa is indeed restricted to Western cultures, such an encapsulated direction to search for CBS misses the conceptual advances provided by CBS research in understanding the contextual nature of all forms of psychopathology.

Prince and Tcheng-Laroche 1987) justifiably pointed out that Western diagnostic classification systems, including both *ICD* and *DSM*, are based on psychiatric conceptualization of Western medical science. Western medical models are based on universal theories of disease. CBS, on the other hand, do not fit well into these universal models. According to Prince and Tcheng-Laroche, the inherent emic–etic diversity and semantic specificities of the CBS hinder their incorporation into prototype approaches. Instead, they suggested that the descriptive classification consisting of inventories of signs and symptoms based on the meaning of illness both for individuals and for cultures rather than the etiological classification would be more practical for clinical use.

"Culture-bound" syndromes are now generally viewed as "culture-related" syndromes (Mental Health Division, WHO, in press). While these syndromes have been found with differing incidence or prevalence rates across societies, they are considered to be more "overdetermined" or "related" to certain cultures than they are "bound" to those cultures. These culture-related syndromes serve to illustrate both the "core elements" of a psychiatric disorder, as well as "those elements that can be modified by individual psychology, social context, and culture." By examining how cultural contexts affect the conceptualization and manifestation of these syndromes, we not only expand our knowledge of aberrant phenomena unfamiliar to Western models of psychopathology, but also gain a better understanding of the cultural contexts of the conventional models of psychopathology themselves. For example, neurasthenia, once classified as a form of neurosis in Western psychiatric nosology, has assumed the identity of a culture-related syndrome in Chinese societies. The transplantation and transformation of neurasthenia as a disease entity outside Western societies provides an apt illustration of the pathoplasticity of culture-related syndromes.

10.02.6 TRANSFORMATION OF NEURASTHENIA

Neurasthenia as a diagnostic category has a vicissitudinary career in Western psychiatry. Coined by the American neurologist George Beard (1869) to describe the exhaustion of the nervous system, it was found to be common among patients attending general practitioners in the early 1990s. It was, however, ignored as a diagnostic category until being classified among the neuroses in *DSM-II* (American Psychiatric Association, 1968) as a condition characterized by weakness, fatigue, lack of stamina, and exhaustion, as distinguished from hysterical neurosis, anxiety neurosis, and depressive neurosis. Literally, it connotes a condition of "weak nerves" in which the sufferer is "thought to be easily overwhelmed by the ordinary stresses of life, with resultant symptoms of somatization, anxiety, and depression" (Mental Health Division, WHO, in press). It was subsequently dropped from the later revisions of the *DSM* but is still retained in *ICD-10* (World Health Organization, 1992). Because of the somatic focus of the symptoms, it is common for patients complaining of neurasthenia to be diagnosed by physicians for some physical dysfunctions. The vagueness of the nosology, however, does not fit in with the increasing emphasis on objectivity and functionality in the American classification system. Despite its American origin, neurasthenia has become a nondisease in America (Kleinman, 1982).

On the other hand, neurasthenia is now the most widely utilized psychiatric diagnosis in China (Mental Health Division, WHO, in press), and is a popular folk nomenclature in Chinese societies and Japan (Lin, 1989). Literally translated as shenjing shuairuo (weakness of the nervous system), this Western term was believed to be "imported" into China via Japan in the early 1990s (Lin). Popular books on Chinese folk medicine published in the past few decades have included neurasthenia as a condition that could benefit from traditional medical treatment, as if it were an indigenous form of disorder. The term is also commonly used among traditional doctors, physicians, psychiatrists, and lay persons to refer to a range of neurotic, psychosomatic, and psychotic disorders (Cheung, 1989; Lee & Wong, 1995). Etiology and treatment of neurasthenia are also elaborated using lay concepts of mental health and Chinese folk medicine. Cheung (1989) suggested that the term's

> ambiguity as well as the overlap with other psychoneurotic disorders may be reasons leading to the rare use of this diagnosis in Western psychiatry ... [but] due to its popularity among the lay public that new life has been injected into this imported medical term. Neurasthenia has been incorporated into Chinese folk medicine, and a set of beliefs has grown up around the original definitions of the disease outlined by George Beard in 1880. (pp. 235–236)

Cheung proposed two possible reasons to explain the popularity of neurasthenia as a diagnostic term among the Chinese. First, neurasthenia provides an acceptable paradigm of health and medicine. From its original Western definitions, neurasthenia has been "indigenized" and incorporated into the paradigm of traditional Chinese folk medicine which is based on a holistic concept of health and illness. Neurasthenia as a multidimensional illness provides a conceptualization in terms of both psychological and somatic symptoms and contexts. Among Chinese patients, somatic symptoms form an important part of the phenomenology of psychiatric illness. The illness experience of neurasthenia in the form of what Kleinman (1982) described as a "socially and culturally shaped type of somatization" may be affected by the patients" course of help-seeking. The somatic symptoms legitimize the sick role whereby the patient can get sympathy and attention from medical professionals and yet not be rejected or stigmatized.

The second reason is the use of the term neurasthenia as a euphemism by psychiatrists and the lay public alike to destigmatize psychiatric disorders. As many as 30% of Chinese psychiatric patients describe themselves as suffering from neurasthenia when presenting themselves to psychiatric clinics (Kleinman, 1982; Wong & Chan, 1984). The distinction between neurasthenia and other forms of psychiatric disorders may stem from its detachment from the notion of insanity and the stigma of madness. Neurasthenia is attributed to somatic depletion, overwork, irregular lifestyle, and extended intellectual activities. It is an acceptable pretext for seeking help for a variety of conditions including psychological symptoms such as overworry, irritability, and memory difficulties, which when presented on their own accord, may be considered inappropriate or insufficient cause for medical consultation.

Even during the early days of the Great Leap Forward and the Cultural Revolution in the People's Republic of China between the 1950s and 1960s, the problem of neurasthenia as a form of mental illness was readily accepted, whereas other Western theories of psychopathology were deemed mentalistic and therefore inconsistent with the communist ideology. The concept of neurasthenia in the People's Republic of China was originally based on the Pavlovian theory of neurophysiology (Chin & Chin, 1969) during the era of Sino-Soviet cooperation when Soviet neuropsychology had a strong influence on Chinese psychology. Neurasthenia is attributed to "tension in the higher nervous system in excess of its capacity, thus causing a weakening in the functioning capacity of the brain tissues and a lack of balance or confusion in nervous activity" (p. 70). Serious disturbances occur in the activity of the nervous system relating to work, study, society, and family. Its "politically correct" etiology and manifestation allowed the disease entity to survive the Cultural Revolution during which other forms of Western diagnoses tended to be denied (Cheung, 1989). Cadres and professionals were spontaneous in their admission of the high prevalence of neurasthenia especially among their peers since they regarded neurasthenia as the outcome of overwork and not as a form of mental illness.

The manifestation of neurasthenia as an illness has also undergone changes within Chinese societies. Earlier studies in China and Hong Kong (Kleinman, 1982; Wong & Chan, 1984) showed that the presenting complaints of neurasthenic patients were predominantly somatic, especially weakness, fatigue, insufficiency of *qi* (vital energy), poor appetite, backache, and weak limbs. A more recent study with young university students (Lee & Wong, 1995) found a more psychological construal of neurasthenia. Moreover, there is a disparity between lay and professional definitions of neurasthenia. According to Beard's (1869) or the *ICD-10*'s (World Health Organization, 1992) definition, fatigue is considered a core symptom of neurasthenia but is infrequently reported in Taiwan or Hong Kong (Lee & Wong, 1995; Rin & Huang, 1989). The de-emphasis on fatigue distinguishes the Chinese experience from the Western condition of chronic fatigue syndrome which may be considered as a contemporary form of "revived" neurasthenia in the West (Abbey & Garfinkel, 1991). Such a disparity points to the importance of examining cross-cultural differences not only in disease entities, but also in the conceptualization of the illness experience.

10.02.7 ILLNESS EXPERIENCE

Medical sociologists and cultural psychiatrists have distinguished between illness and disease (Mechanic, 1982). Kleinman (1988a) makes the distinction succinctly by referring to illness as "the patient's perception, experience, expression, and pattern of coping with symptoms" and to disease as "the way practitioners recast illness in terms of their theoretical models of pathology" (p. 7). The illness experience is a culturally shaped phenomenon that is interpreted by the patients and their families resulting in a particular course of action. These interpretations are mediated by language, illness beliefs, personal significance of suffering, and learned illness behavior.

The experiences of distress are multifaceted. The distressed persons choose their form of expression according to their own sensitivity to those facets that make sense to them or that conform to their understanding. Folk understanding of psychological distress seldom conforms to the professional classification of disorders, especially in cultures where formalized mental health services are not popular or are inaccessible. Thus, these expressions are necessarily selective, being bound by the personal and cultural repertoire of distress idioms. The repertoire of coping and illness behavior is similarly acquired in the sociocultural context. The illness experience encompasses not only the phenomenology of the individual which focuses on the intrapsychic dimension, but also on the interaction of the person with significant members and aspects of the environment. These social and cultural dimensions form an integral part of the dynamics involved in the process of illness.

A psychiatric diagnosis, on the other hand, is a professional's interpretation of the person's experience which becomes formalized as signs of particular disease states in a static medical model. These diagnostic categories are often reified as real entities by the professionals who try to fit the reported experiences into one of the categories. Through exposure, patients learn to report what the professionals expect of them. However, when the diagnostic categories developed in one culture are projected onto patients in another culture, their validity needs to be established. Mental health professionals who are mostly trained in Western medical models have often imposed these categories in non-Western cultures even when they lack folk coherence. Some of these categories may be more valid or relevant than others, whereas there is strong evidence that psychiatric diagnoses such as organic brain disorders and schizophrenia are valid worldwide (Kleinman, 1988a; World Health Organization, 1973, 1979), other diagnoses such as dysthymic disorder or anorexia nervosa are questioned as examples of "category fallacy" which lack cross-cultural validity. When the culturally imposed categories are found to differ from the illness experience found within the culture, ethnocentric professionals tend to interpret the discrepancy as a problem of the individual or the culture, instead of questioning the relevance of the category itself.

10.02.8 SOMATIZATION AS A CATEGORY FALLACY

The problem of category fallacy can be illustrated by the way in which mental health professionals from Western cultures initially interpreted the tendency among Chinese psychiatric patients to somatize their psychological distress. In the wake of early interest in cultural psychiatry, a number of studies on the patterns of symptom presentation among Chinese psychiatric patients noted that these patients tend to somatize their problems and delay psychiatric treatment (Kleinman, 1977; Lin, 1982; Lin, Tardiff, Donetz, & Goresky, 1978; Marsella, Kinzie, & Gorson, 1973; Tseng, 1975). Somatization has been emphasized as the cultural feature of neurotic disorders, especially depression, among the Chinese. Various aspects of the Chinese culture were then put forth to explain the somatization tendency (Cheung, 1982, 1985, 1995; Cheung & Lau, 1982).

One group of cultural explanations attributes the somatization tendency to denial, suppression, or repression of emotions. The typical Chinese is described as being reserved in expressing feelings and prone to avoiding open display of emotions, especially negative ones. Another group of explanations argues that the use of concrete physical terms in expressing one's distress demonstrates that the Chinese language lacks an adequate vocabulary to express emotions explicitly. So physical metaphors or terms related to body organs are used to represent affective states. The third group of explanations is based on the assumption of mind–body dualism as the standard of epistemology, so that the lack of distinction between psychological and somatic systems among the Chinese was likened to the simple, undifferentiated, and concrete level of thinking among primitive cultures.

These cultural assumptions have been challenged both on the grounds of cultural specificity as well as interpretation. It has been questioned whether somatization is indeed specific to the Chinese as originally claimed. Even so, the explanations for this culturally related observation need further scrutiny. Cheung (1982, 1985) criticized these cultural attributions as *post hoc* explanations based on cultural generalizations without empirical verification. In an extensive review of the literature, Singer (1975) argued that in the first place, the phenomenon of somatization is not exclusive to Chinese patients, but is also commonly found among less educated and lower social-class patients in Western countries. Kleinman (1986) later reckoned that somatization is also very common in the West although the fact may be ignored or de-emphasized there. Cheng (1989) compared survey results on the symptomatology of minor psychiatric morbidity in Taiwan and in Britain and found the rates of somatic symptoms for both countries to be comparable.

There is also evidence to show that although somatic complaints are the dominant features in Chinese patients' symptom presentation, psychological symptoms are by no means suppressed or repressed. Cheung, Lau, and Waldmann (1980–81) found that while depressed patients reported mostly somatic complaints to the medical doctor in general practice, they were very ready to admit to having a range of psychological and emotional symptoms when directly asked by the practitioner. Subsequent studies with psychiatric patients attending psychiatric services for the first time also revealed that psychological symptoms were frequently reported along with somatic complaints (Cheung, Lau, & Wong, 1984). Studies with the general population and with university students in Hong Kong also confirm that psychological problems were commonly reported (Cheung, 1982; Cheung, Lee, & Chan, 1983). However, psychological problems may not be readily reported to the medical practitioner. Instead, the Chinese would turn to their friends and family members for help if they perceive their problems to be psychological in nature.

One of the problems of these early discussions on the somatization tendency suffered from variations in the definitions of the term used. Somatization may be referred to as a psychiatric disorder, various patterns of illness behavior, forms of help-seeking related to bodily symptoms, or a combination of these concepts. Based on the formistic and mechanistic model of biomedicine (Schwartz, 1982), somatic presentation of distress is assumed to be due to an underlying psychological problem masked by the somatic presentation of symptoms often equated with somatization. Somatization, thus, is presumed to be pathological with maladaptive consequences. In the formistic world view, a person is categorized as either sick or well. In the mechanistic world view, a specific cause is linked to a single effect. These approaches ignore the contextual nature of illness involving both the organism and the environment, in which there may be interaction among multiple causes in the organism and the environment.

The focus on somatization as a pathological disorder misses the opportunity to understand in a systemic way the various patterns of illness behavior among Chinese patients, including their subjective experience, problem presentation, communication, and form of help-seeking.

10.02.9 SOMATIZATION AS ILLNESS EXPERIENCE

While studies of Chinese patients in Chinese societies have consistently found that somatic complaints are the dominant features in their descriptions of discomfort, it is also noted that psychological symptoms are readily reported by both clinical and nonclinical samples, especially if directly asked (Cheng, 1988, 1995; Cheung, 1982, 1987; Cheung et al., 1980–81, 1984). An attempt to understand the phenomenon of somatization among the Chinese should take into account the phenomenology of discomfort and suffering, the process of communication, ways of coping and help-seeking, and the patient–doctor relationship.

In his recent works, Kleinman (1986, 1988b) has provided an insightful perspective to understand somatization as human suffering contextualized in the personal, interpersonal, and cultural meanings of illness. He proposed that a "dialectical relationship exists between symptoms and society" (p. 2). The examination of how social factors produce the connection between neurasthenia-depression and pain which may be perceived as the cause of *disease* provides the illustration as to the way "culture shapes the strategic interpretation and negotiated experience of *illness*" (p. 2). As an idiom of distress, somatization is an illustration of the ethnography of the suffering experience. While the narrative of human suffering is not confined to bodily sensations, this mode of expression is highlighted in the medical context especially in non-Western cultures.

The meaning of bodily symptoms for the Chinese has been examined in terms of the Chinese language and traditional Chinese medical practice (Cheng, 1989; Ots, 1990; Tung, 1994). In traditional Chinese medicine, emotional problems are associated with different body organs. Dysharmonic emotions are pathogenic etiological factors which produce somatic dysfunctions, which, in turn, should be treated by harmonizing bodily functions (Ots, 1990). Tung also illustrated that the Chinese use body-related verbal expressions in an all-embracing fashion to express personal and social aspects of human concerns. Body-related words form expressions that are idioms of human conditions. These forms of expression are indigenous and often expressed in local dialects that escape the attention of Western-trained psychiatrists (Cheng, 1989). On the other hand, psychiatrists who try to fit the patients" expressions into their professional taxonomy would miss the rich information that accompanies the patients' expressions.

As illustrated earlier, the use of the term neurasthenia provides culturally legitimized somatic idioms to sidestep the Western framework of mental diseases which are highly stigmatized. However, somatic idioms are more than just euphemistic substitutes for Western diagnostic labels. They encompass schema to make sense of life problems and to direct courses of action. Illness meanings are physiological as well as social. According to Kleinman (1986), "the lived experience of social reality mediated by the body . . . is the symbolic bridge, the socio-somatic reticulum that ties failure to headaches, anger to dizziness, loss and demoralization to fatigue" (p. 146).

Zheng, Xu, and Shen 1986) analyzed the styles of verbal expression of emotional and physical experiences of depressed patients and normal controls in China. Both emotional and physical expressions are used as narratives of human suffering, depending on the nature of the experience. For example, a purely psychological verbal style with words such as "unhappy" and "worried" was used to describe abstract emotional states such as indecisiveness, self-pity, guilt, or helplessness. Purely somatic expressions such as "heart beating" would be used for the concrete emotional state of fear. The neutral mode including both psychological and somatic features (e.g., "uncomfortable inside the heart" or "cannot use my brain") was used to express anxiety. For depression, both the somatic and the neutral modes were frequently used. In the expression of suicidal ideation, however, the person was often unable or unwilling to choose an expression, which the authors labeled as the deficient expressive style. The somatic style in expressing emotions among the Chinese patients in this study, however, was not related to symptom manifestation as reported on a symptom checklist.

Cheung (1995) pointed out that "somatization as an idiom of distress is contextualized in a process of communication" (p. 163). This process includes both the discourse between Chinese patients and doctors as well as other interpersonal interactions. Cheung and Lau (1982) showed that situational variations elicited different expectations and behaviors from patients who behave according to what they believe is expected of them in the situation. When Chinese psychiatric patients perceived their consultation to be medical in nature, such as attending a psychiatric clinic, they reported their physical as well as psychological symptoms. On the other hand, when they perceived their consultation to be related to a judicial nature, such as being referred to the psychiatrist at the correctional services by the judge, prisoners would recount the events leading to their arrest instead of presenting their symptoms.

Studies based on experiences of psychiatric patients may be biased by the self-selection of people who have chosen to consult psychiatric services, leading to the conclusion that Chinese patients tend to somatize their psychological problems. To understand the illness experience of the Chinese people, their coping approaches and the course of help-seeking prior to the psychological consultation need to be included. Studies of the patterns of help-seeking among different ethnic groups have shown that non-Caucasian patients generally delay seeking psychiatric treatment and approach other indigenous resources instead, at least initially (Cheung, 1987; Cheung et al., 1984; Lin et al., 1978). In Hong Kong, psychiatric consultation may be delayed by as long as 30 years after the initial onset of symptoms if they were perceived to be purely psychological in nature (Cheung, 1987). The Chinese prefer to endure their suffering, rely on themselves, or seek help from their friends or family for psychological problems. They would be more likely to seek medical attention for physical symptoms which would be perceived as illness needing medical attention (Cheung et al., 1983, 1984). The first medical consultation is predominantly with the general practitioner. Owing to the unfamiliarity with mental health services, the patients generally do not make a distinction among medical and mental health specialties which are more demarcations defined by the professionals.

The dominant role played by the general practitioner in directing psychiatric patients to mental health services has been highlighted by Goldberg and Huxley (1980). They showed that even in developed countries, only a small minority of psychologically disturbed persons reached psychiatric services. Many psychiatric illnesses passed unrecognized by general practitioners who prolonged medical attention on the physical complaints. Cheung (1991) likened the function of the general practitioner to that of "a roundabout where routes may converge and be redirected. When most travelers were unfamiliar with the direct route to recovery, the roundabout was the hub they would return to before embarking on alternative treatment paths" (p. 67). The patients' initial conceptualization of the causes and severity of their illness would direct their coping and help-seeking behavior. They would only seek medical attention if the problem is perceived to be somatic and serious. Once they embark on the medical course, the experience of medical consultation, in turn, would have shaped the patients' interpretation of their suffering. Patients who finally arrived at the psychiatric clinic are likely to have traveled through detours of medical consultation which focus their attention on their somatic complaints. They would be likely to present to the mental health professional, to them another medical doctor, that which they deemed as expected and appropriate (Cheung, 1987).

The reconceptualization of somatization in terms of the individual's culturally embedded illness experience and narratives of suffering advances the theoretical understanding from that of a category fallacy to a useful framework to interpret the phenomena. Psychopathology is after all a scientific attempt to organize and explain the observed phenomena systematically. Recognition of cross-cultural diversities poses new challenges to the traditional ethnocentric ideologies that have encapsulated Western theories of psychopathology and approaches to clinical psychology.

10.02.10 CULTURE AND PSYCHOPATHOLOGY

The role of culture in counseling and psychotherapy has received more attention from psychologists than the role of culture in the area of psychopathology (Leong, 1986; Pedersen, 1983, 1987, 1990, 1991; Pederson & Marsella, 1982; Pedersen, Draguns, Lonner, & Trimble, 1981; Sue, D. W., 1977, 1978, 1989, 1990, 1991; Sue, Arredondo, & McDavis, 1992; Sue & Kirk, 1975; Sue & Sue, 1977, 1990; Sue & Zane, 1987; Vontress, 1974). Culture has gained the status of a generic theory to explain counseling relationships in general, and not just for understanding exotic people. Pedersen (1991) proposed *multiculturalism* as a "fourth force" in counseling, complementary to psychodynamic, behavioral, and humanistic explanations of human behavior. The multicultural perspective "combines the extremes of universalism and relativism by explaining behavior both in terms of those culturally learned perspectives that are unique to a particular culture and in the search for common-ground universals that are shared across cultures" (p. 6). While acknowledging the pervasive nature of multiculturalism in complex modern society, Pedersen's theory is focused only on counseling approaches.

A few attempts have been made to apply specific psychological theories to mental health issues among ethnic subgroups in the USA. For example, S. Sue (1977) and Zane, Sue, Hu, and Kwon (1991) examined cultural differences in learned helplessness and assertion among Asian-Americans using a social learning analysis. Specific interest has been directed to the identity development of ethnic minority groups and how minority identity development may affect mental health problems (Sue, 1989). These efforts result in better understanding of the cultural diversities and cultural differences among specific ethnic groups. There is much less written on the general role of culture in psychopathology or abnormal psychology.

10.02.10.1 Role of Culture in Psychopathology

Tseng and McDermott (1981) gave a more comprehensive introduction to the role of culture in psychiatry. Culture affects psychopathology in many ways, including producing stress, creating specific problems, predisposing vulnerability, and contributing to the choice of psychopathology (pp. 14–24).

10.02.10.1.1 Culture-produced stress

Stress may be created by culturally formed anxiety or culturally demanding performance. Cultural attitudes and beliefs prescribe expectations and standards which if not met would produce mental and emotional stresses. Many societies have established taboos, rules, or rites which, if contravened, would create anxieties. Individuals living in cultures that believe spells can cause death would become ill or even die if they think that they are under a spell. Similarly, cultural demands of individuals cause stress to those who are unable to meet those demands. For example, many cultures demand that a woman produce a male child. The woman would be blamed and would feel guilty if she continued to produce only girls. She would face a loss of status or the threat of replacement by another woman. These demands produce fear, anxiety, and shame.

10.02.10.1.2 Culture-related problems

Specific problems may be created by a culturally determined limitation in behavior range, cultural changes, changing roles, and loss of source of support, as well as sociocultural discrimination. Within a cultural setting where rules are excessive, but the behavior range allowed is very limited, individuals who are not prepared for these restrictions would encounter problems. Similarly, rapid cultural changes may result in a confused value system to individuals who are used to a stable and traditional culture. Changing roles for women, for example, while improving women's social status and sense of independence, may also render them more vulnerable to loneliness and depression when the protection and support women previously obtained from their families have diminished. Sociocultural groups who face discrimination within a society suffer scorn, isolation, and disadvantage, resulting in stress and deficits.

10.02.10.1.3 Culture-inherited vulnerability

Vulnerability to stress may be predisposed by culturally prescribed child-rearing practices which influence the child's personality. In some cultures, parents emphasize independence and early decision making. This push toward independence may put undue pressure on the young person who needs a longer period of dependency. In other cultures, intimate dependency on mothers is extended for a much longer time. This dependency is extended into adult relationships as well which may lead to later problems in socialization with others.

10.02.10.1.4 Cultural contribution to choice of psychopathology

Culture plays a part in how its vulnerabilities find expression by influencing model solutions to problems that may themselves be pathological. The form of psychopathology found in a culture tends to be linked with the stresses within that culture. For example, excessive self-depreciation, a result of compliance to strong external control, may become a cause for depression. There is also a suggestion that low social cohesion in a community is related to the frequency of depression.

10.02.10.2 Cultural Explanations

Once psychological disorders are detected, culture also plays a role in labeling these disorders based on its concepts of mental illness. These labels describe the observed behavior, attribute causes, provide explanations, and convey social acceptance or rejection of the afflicted. Explanations of the nature and cause of mentally disturbed behaviors are grouped into four categories by Tseng and McDermott (1981, pp. 29–33):

(i) *Supernatural explanations*. The disturbance is explained as resulting from spirit possession, soul loss, divine wrath, sorcery, or violation of taboo.

(ii) *Natural explanations*. Illness, misfortune, and unhappiness are related to the underlying principles of the universe which govern all nature. Disharmony of the natural elements within the human body, incompatibility with natural principles in the environment, and noxious factors in the environment upset the natural principles and cause physical and mental disturbance.

(iii) *Physical-medical explanations*. The reasons for illness are viewed as physical or physiological originating from the individual. The causes may be physical-physiognomy problems (such as facial features, body build), physiological imbalance or insufficiency (such as badly balanced diet, excessive sexual activity), or disease.

(iv) *Sociopsychological explanations*. Mental disorder is seen as a psychological reaction to the stress of internal or external maladjustment.

Tseng and McDermott, however, did not make any distinction between the professional or "authoritative" explanations of abnormal behaviors and the folk conceptualizations and attributions of the experience of distress in these categories. In non-Western cultures that have not adopted Western models of medicine and psychology, explanatory models are often based on folk concepts. In cross-cultural comparisons of explanatory models, there may be biases when Western professional models of psychiatric classification are compared with local folk concepts of distress.

10.02.11 CROSS-CULTURAL PSYCHOPATHOLOGY

10.02.11.1 Universalist vs. Relativist Approaches

Cross-cultural psychopathology is premised on two competing orientations. The biological approach assumes that the cultural invariance of mental disorders, that is, the same disorders are prevalent in all cultures (Fernando, 1988, p. 60). In this approach, the "universalist" would use the psychiatric classification system developed in the West as a basis of identifying disorders in other parts of the world. The social-anthropological approach, on the other hand, emphasizes the cultural relativity of psychopathology, with each culture having its own disorders. Culture plays an important role in determining the behavior, thinking, and emotions of the individuals resulting in different forms of mental disorders. The "relativist" looks at illness as a part of the total cultural context deriving its meaning from the specific culture.

The two orientations have led to different research directions in cross-cultural psychopathology. The universalist or etic orientation has focused on cross-cultural comparisons in the rates of major psychiatric disorders, based on classification systems developed in the West. These studies have made use of population surveys with results showing a range of rates of major mental disorders in different countries (cf. Kleinman, 1988a, pp. 34–41, or Draguns, 1985, 1986, 1989 for a review of the findings). However, it has been noted that a narrower range of rates is found when diagnostic criteria are standardized and more homogenous samples of cases are used. International studies have been carried out to compare the prevalence and symptomatology of common psychiatric diseases such as schizophrenia (Sartorius et al., 1986; World Health Organization, 1973, 1979) or depression (Jablensky, Sartorius, Gulbinat, & Ernberg, 1981; World Health Organization, 1983). These studies identified core syndromes of the disorders. For example, across the nine research sites which represented both developed and developing countries, the core syndrome of schizophrenia included "restricted affect, poor insight, thinking aloud, poor rapport, incoherent speech, unrealistic information, and bizarre and/or nihilistic delusions" (Draguns, 1989, p. 242). The more modest project on depression involving four countries (World Health Organization, 1983) also identified a small number of symptoms such as vegetative symptoms which were present in most of the countries. On the other hand, the experience of guilt is much less universal in the symptomatology of depression.

Similar to psychiatric epidemiological studies comparing specific disorders, cross-cultural assessment studies have applied the major assessment instruments developed in the West to other cultures and identified cross-cultural similarities and differences. For example, one personality test that has been translated and tried out in the most number of countries is the Minnesota Multiphasic Personality Inventory (MMPI; Butcher & Clark, 1979; Butcher & Pancheri, 1976; Butcher & Spielberger; 1985, Hathaway & McKinley, 1967). These studies confirmed the clinical utility of the MMPI in other countries although cultural differences are identified and cultural adaptations may have to be made (e.g., Cheung & Song, 1989; Cheung, Song, & Butcher, 1991; Cheung, Song, & Zhang, 1996; Cheung, Zhao, & Wu, 1992).

The relativist or emic orientation has focused on culture-specific phenomena, such as culture-bound syndromes discussed in an earlier section. Research based on in-depth studies within a single culture provides a rich source of information about the phenomena. However, as pointed out earlier, attributions to cultural phenomena were often based on *post hoc* and generalized assumptions about the culture involved without elaboration on the functional relationship between cultural processes and psychopathology.

The shortcoming of the universalist approach lies in its ethnocentricism which assumes that the Western culture is more developed and superior. Cross-cultural differences are interpreted as a reflection of cultural underdevelopment, as in the early days of psychiatry when non-Western cultures were often referred to as "primitive" and "savages" (Fernando, 1988, p. 60). It is assumed that diagnostic categories developed in a Western cultural setting are applicable universally. Standardized research methods are developed to establish reliable diagnoses across cultures in Western terms. However, the cross-cultural validity of the disease as an illness or a phenomenon in those other cultures may not be established (Fernando).

Similarly, the universalist approach to assessment may lead to misleading clinical interpretations. For example, in the application of the MMPI to the Chinese people, several clinical scales, including the Depression scale and the Schizophrenia scale, were found to be consistently elevated even among normal samples when the American norms were used (Cheung, 1995; Cheung, Song, & Zhang, 1996). Both in Hong Kong and the People's Republic of China, the mean T-scores for normal adults on these two scales usually approach the clinical cut-off point used in the USA, although the mean scores for clinical samples are even higher. If cultural universality is assumed, there would be a danger of overinterpreting clinical pathology in the normal Chinese population.

The relativist approach, on the other hand, may not get beyond describing culture-bound syndromes and constructing intracultural explanations, thereby not contributing much to cross-cultural psychopathology. By assuming cultural specificity, no comparison can be made between cultural units. Without a cross-cultural framework, the notion of culture, however, becomes irrelevant. The determination of the cultural unit itself is also a subject of controversy. Cultural groups are usually operationally defined in terms of race, ethnicity, geographic location, or religion, assuming cultural similarity among members of the same groups. Subcultural groups further complicate the assumption of cultural similarity.

An alternative approach to get out of the impasse created by these two extreme positions is to reconstruct cross-cultural psychopathology from the basic units of clinical psychology: observation of the phenomenological experience of individuals from their cultural contexts, description of these expressions and manifestations on the basis of the individuals' perspectives, prescription of labels which are culturally relevant or meaningful based on indigenous and cross-cultural studies, selection of appropriate treatment alternatives, and prediction of outcome. While the world is becoming more multicultural, this contextualized analysis may become a "universal" approach.

10.02.11.2 Recent Research on Cross-cultural Psychopathology

Recent research on cross-cultural differences in psychopathology has attempted to explain the similarities and differences in terms of the cultural contexts. For example, while a pancultural nucleus of symptoms was found for schizophrenia and depression, a variety of cultural transformations was also identified. Draguns (1989) proposed a range of possible relations between culture and the manifestations of the disorders. These include:

(i) magnification or exaggeration—a culturally characteristic behavior being caricatured and reduced to absurdity;

(ii) violation of cultural norms—doing what is not culturally permissible;

(iii) cultural differences in values affecting the characteristic modes of expressing psychopathology—symptom choice and expression being conceptualized as social transactions.

However, there are few studies that examine the functional relationship between cultural characteristics and cross-cultural differences in psychopathology. Draguns (1989) attempted to apply Hofstede's (1980) worldwide study of work-related values to manifestations of abnormal behavior across cultures. In Hofstede's original study, which involved employees from over 40 nationalities, four cultural dimensions were found to be capable of accounting for most of the cultural differences. These four dimensions were individualism/collectivism, uncertainty avoidance, power distance, and masculinity/femininity. Hofstede (1983) had linked the dimension of individualism/collectivism to the distinction made earlier by Benedict (1946) between "guilt" (individualist) and "shame" (collectivist) cultures. However, the relationship between this value dimension and psychopathology involves a more subtle process. Rather than assuming that patients from some Asian cultures do not experience guilt, cultural differences may be found more in the way that guilt is expressed. In individualistic cultures, guilt may be expressed in absolute and abstract terms involving self-accusation and condemnation. In collectivistic cultures, guilt may be expressed in more concrete and interpersonal modes.

To date, there are few studies on cross-cultural psychopathology using the other three dimensions from Hofstede's study. Draguns (1989) further noted that psychopathology varied across cultures even when there was no substantial difference on Hofstede's four dimensions. He pointed to the need for more multicultural studies involving both normal and abnormal samples in order to formulate the functional relationship between cultural characteristics and psychopathology.

The series of studies on somatization among the Chinese described in earlier sections also illustrates how cross-cultural differences in the expression of symptomatology could be understood in terms of the culturally relevant cognitive schema and interpersonal contexts. Through these cultural studies, important dimensions that have not been examined in North American and Western European theories could be incorporated to increase the cultural sensitivity of these theories. This cultural sensitivity is becoming more important not only because ethnic and cultural pluralism is becoming more prevalent in many societies. What are thought to be indigenous categories can in fact be extended to explain the complexity of human personality and psychopathology within and across cultures.

One example of this combined emic-etic approach is the development of the Chinese Personality Assessment Inventory (CPAI; Cheung, Leung, Fan, et al., 1996) which integrated Western methods of personality assessment with folk concepts of personality in the Chinese culture. In addition to etic personality constructs, emic constructs which were not covered in translated personality inventories were included in the CPAI. Scales were developed for personality characteristics such as harmony, relationship orientation, face, modernization, and thrift, which are salient constructs for person descriptions among the Chinese. These scales were loaded on a factor labeled Chinese Tradition. The Chinese Tradition factor was found to be a relevant predictor of mental health measures especially for Chinese males. It enhanced life satisfaction and prevented the acting out of antisocial behavior particularly when the stress level was high (Cheung & Gan, 1996). The Chinese Tradition factor was also found to be culturally distinct from the five-factor structure of personality which is currently accepted as a universal model of personality (Cheung, Leung, Law, & Zhang, 1996). These emic constructs may introduce new dimensions of the personality structure to Western models of personality and psychopathology.

10.02.11.3 Future Directions

As Lewis-Fernandez and Kleinman (1994) concluded, "we might very well come across local indigenous categories, such as face and favor, that can be used to reformulate our leading models of personality formation and their relationship to psychopathology" (p. 70). The study of cross-cultural psychopathology is expanding beyond the exotic frontier into the mainstream of psychology.

Research in cross-cultural psychopathology, as in other fields of cross-cultural psychology, is incorporating the combined emic–etic approach. There is an increasing recognition that "culture-bound" syndromes may not be unique to a specific culture. Instead, culture-related phenomena such as neurasthenia and somatization help to illustrate the dynamics of cultural forces affecting illness behavior in different societies.

Psychopathology is manifested in a cultural context, and should be interpreted via a contextualized analysis. Cross-cultural psychopathology research can help to identify convergent and divergent dimensions in explaining normal and abnormal behaviors across cultures, enhancing our awareness as researchers and practitioners of the cultural context of our analyses. As our global community becomes more multicultural, the dialectic process of synthesizing new observations and constructs from other cultural contexts poses a challenge to clinical psychologists to recognize culture as an integral part of their theoretical framework.

10.02.12 SUMMARY

Clinical psychology has suffered from an ethnocentric approach blinded by culture-bound assumptions. Early research in cross-cultural psychopathology was borrowed from cultural psychiatry which has focused its interest on the identification of culture-bound syndromes. In this chapter, an alternative approach to study the interplay of cultural dynamics in the manifestation of these "culture-bound syndromes" is illustrated through the discussion of Koro, neurasthenia, and somatization. The transplantation and transformation of neurasthenia from a Western diagnostic category to a culture-related syndrome in Chinese societies demonstrate the fluidity of diagnostic labels in incorporating cultural meanings. By reexamining somatization as an illness experience and narratives of suffering, a more useful framework for studying the relationship between culture and psychopathology is also identified.

Culture affects psychopathology by producing stress, creating specific problems, predisposing vulnerability, and selecting the form of psychopathology. It also provides folk explanations of the nature and cause of aberrant behavior. Earlier research methods have adopted either the emic or the etic approach. The emic approach focused on the relativistic aspects of culture such as culture-bound syndromes. The etic approach, on the other hand, assumes that Western-based constructs are applicable in other cultures and attempts to replicate these constructs in cross-cultural studies. Recent studies attempt to incorporate the emic–etic approaches. For example, the development of an emic–etic instrument such as the CPAI has introduced new dimensions of the personality structure to Western models of psychopathology. Other work such as that by Draguns (1989) are uses multicultural studies to formulate functional relationships between cultural characteristics and psychopathology. These new developments have pointed to the importance of the interpersonal dimension in psychopathology which has often been neglected in Western models of psychopathology. Future directions for study in cross-cultural psychopathology need to bring culture from its marginal status into the mainstream of clinical psychology.

10.02.13 REFERENCES

Abbey, S. E., & Garfinkel, P. E. (1991). Neurasthenia and chronic fatigue syndrome: The role of culture in the making of a diagnosis. *American Journal of Psychiatry, 148,* 1638–1646.

American Psychiatric Association (1968). *Diagnostic and statistical manual of disorders* (2nd ed.). Washington, DC: Author.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association (1994) *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Bartholomew, R. E. (1994). The social psychology of "epidemic" koro. *International Journal of Social Psychiatry, 40,* 46–60.

Beard, G. M. (1869). Neurasthenia, or nervous exhaustion. *Boston Medical and Surgical Journal, 3,* 217–221.

Benedict, R. (1946). *The chrysanthemum and the sword.* Boston: Houghton Mifflin.

Butcher, J. N., & Clark, L. A. (1979). Recent trends in cross-cultural MMPI research and application. In J. N. Butcher (Ed.), *New developments in the use of the MMPI* (pp. 69–111). Minneapolis, MN: University of Minnesota Press.

Butcher, J. N., & Pancheri, P. (1976). *A handbook of cross-national MMPI research.* Minneapolis, MN: University of Minnesota Press.

Butcher, J. N., & Spielberger, C. D. (Eds.) (1985). *Advances in personality assessment* (Vol. 4). Hillsdale, NJ: Erlbaum.

Carr, J. E. (1985). Ethno-behaviorism and the culture-bound syndromes: The case of *Amok*. In R. C.. Simons & C. C.. Hughes (Eds.), *The culture-bound syndromes: Folk illnesses of psychiatric and anthropological interest* (pp. 199–223). Dordrecht, The Netherlands: Reidel.

Cheng, T. A. (1988). A community study of minor psychiatric morbidity. *Psychological Medicine, 18,* 697–708.

Cheng, T. A. (1989). Symptomatology of minor psychiatric morbidity: A cross-cultural comparison. *Psychological Medicine, 19,* 697–708.

Cheng, T. A. (1995). Neuroses in Taiwan: Findings from a community survey. In T. Y. Lin, W. S. Tseng, & E. K. Yeh (Eds.), *Chinese societies and mental health* (pp. 167–175). Hong Kong: Oxford University Press.

Cheung, F. M. (1982). Psychological symptoms among Chinese in urban Hong Kong. *Social Science and Medicine, 16,* 1339–1344.

Cheung, F. M. (1985). An overview of psychopathology in Hong Kong with special reference to somatic presentation. In W. S. Tseng & D. Y. H. Wu (Eds.), *Chinese culture and mental health* (pp. 287–304). Orlando, FL: Academic Press.

Cheung, F. M. (1987). Conceptualization of psychiatric illness and help-seeking behavior among Chinese. *Culture, Medicine, and Psychiatry,* 11, 97–106.

Cheung, F. M. (1989). The indigenization of neurasthenia in Hong Kong. *Culture, Medicine, and Psychiatry, 13,* 227–241.

Cheung, F. M. (1991). Health psychology in Chinese societies in Asia. In M. A. Jensen & J. Weinman (Eds.), *The International development of health psychology* (pp. 63–74). Chur, Switzerland: Harwood.

Cheung, F. M. (1995). Facts and myths about somatization among the Chinese. In T. Y. Lin, W. S. Tseng, & E. K. Yeh (Eds.), *Chinese societies and mental health* (pp. 156–166). Hong Kong: Oxford University Press.

Cheung, F. M., & Gan, Z. Q. (1996). Personality traits as predictors of mental health in Chinese: Cultural and gender issues. Unpublished manuscript.

Cheung, F. M., & Lau, B. W. K. (1982). Situational variations in help-seeking behavior among Chinese patients. *Comprehensive Psychiatry, 23,* 252–262.

Cheung, F. M., Lau, B. W. K., & Waldmann, E. (1980–81). Somatization among Chinese depressives in general practice. *International Journal of Psychiatry in Medicine, 10,* 361–374.

Cheung, F. M., Lau, B. W. K., & Wong, S. W. (1984). Paths to psychiatric care in Hong Kong. *Culture, Medicine, and Psychiatry, 8,* 207–228.

Cheung, F. M., Lee, S. Y., & Chan, Y. Y. (1983). Variations in problem conceptualizations and intended solutions among Hong Kong students. *Culture, Medicine, and Psychiatry, 7,* 263–278.

Cheung, F. M., Leung, K., Fan, R. M., Song, W. Z., Zhang, J. X., & Zhang, J. P. (1996). Development of the Chinese Personality Assessment Inventory. *Journal of Cross-Cultural Psychology, 27,* 181–199.

Cheung, F. M., Leung, K., Law, J. S., & Zhang, J. X. (1996, August). Indigenous Chinese personality constructs. Paper presented at the 26th International Congress of Psychology, Montreal, Canada.

Cheung, F. M., & Song, W. Z. (1989). A review on the clinical applications of the Chinese MMPI. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1,* 230–237.

Cheung, F. M., Song, W. Z., & Butcher, J. N. (1991). An infrequency scale for the Chinese MMPI. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 648–653.

Cheung, F. M., Song, W. Z., & Zhang, J. X. (1996). The Chinese MMPI-2: Research and applications in Hong Kong and the People's Republic of China. In J. N. Butcher (Ed.), *International adaptations of the MMPI-2: A handbook of research and applications* (pp. 137–161). Minneapolis MN: University of Minnesota Press.

Cheung, F. M., Zhao, J. C., & Wu, C. Y. (1992). Chinese MMPI profiles among neurotic patients. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 4,* 214–218.

Chin, R., & Chin, A. S. (1969). *Psychological research in communist China: 1949–66*. Cambridge, MA: MIT Press.

Chowdhury, A. N. (1996). The definition and classification of Koro. *Culture, Medicine, and Psychiatry, 20,* 41–65.

Draguns, J. G. (1985). Psychological disorders across cultures. In P. Pedersen (Ed.), *Handbook of cross-cultural counseling and therapy* (pp. 55–62). Westport, CT: Greenwood.

Draguns, J. G. (1986). Culture and psychopathology: What is known about their relationship? *Australian Journal of Psychology, 38,* 329–338.

Draguns, J. G. (1989). Normal and abnormal behavior in cross-cultural perspective: Specifying the nature of their relationship. In J. J. Berman (Ed.), *Nebraska Symposium on Motivation 1989,* (Vol. 37, pp. 235–277). Lincoln, NE: University of Nebraska Press.

Edwards, J. G. (1985). Indigenous Koro, a genital retraction syndrome of insular Southeast Asia: A critical review. In T. C. Simons & C. C. Hughes (Eds.), *The culture-bound syndromes: Folk illnesses of psychiatric and anthropological interest* (pp. 169–192). Dordrecht, The Netherlands: Reidel.

Fernando, S. (1988). *Race and culture in psychiatry*. London: Croom Helm.

Fishbain, D. A., Barsky, S., & Goldberg, M. (1989). "Koro" (genital retraction syndrome): Psychotherapeutic interventions. *American Journal of Psychotherapy, 43,* 87–91.

Goldberg, D., & Huxley, P. (1980). *Mental illness in the community: The pathway to psychiatric care*. London: Tavistock.

Good, B. J. (1996). Culture and DSM-IV: Diagnosis, knowledge and power. *Culture, Medicine, and Psychiatry, 20,* 127–132.

Gussow, Z. (1985). *Pibloktoq* (hysteria) among the Polar Eskimo: An ethnopsychiatrc study. In R. C. Simons & C. C. Hughes (Eds.), *The culture-bound syndromes: Folk illnesses of psychiatric and anthropological interest* (pp. 271–287). Dordrecht, The Netherlands: Reidel.

Gwee, A. L. (1985). Koro—A cultural disease. In T. C. Simons & C. C.. Hughes (Eds.) *The culture-bound syndromes: Folk illnesses of psychiatric and anthropological interest* (pp. 155–160). Dordrecht, The Netherlands: Reidel.

Hathaway, S. R., & McKinley, J. C. (1967). *The Minnesota Multiphasic Personality Inventory Manual*. New York: Psychological Corporation.

Hill, R. F., & Fortenberry, J. D. (1992). Adolescence as a culture-bound syndrome. *Social Science and Medicine, 35,* 73–80.

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: Sage.

Hofstede, G. (1983). The cultural relativity of organizational practices and theories. *Journal of International Business Studies, 14,* 75–89.

Holden, T. J. (1987). Koro syndrome associated with alcohol-induced systemic disease in a Zulu. *British Journal of Psychiatry, 151,* 695–697.

Hughes, C. C. (1985). Culture-bound or construct-bound: The syndromes and DSM-III. In T. C. Simons & C. C. Hughes (Eds.), *The culture-bound syndromes: Folk illnesses of psychiatric and anthropological interest* (pp. 3–24). Dordrecht, The Netherlands: Reidel.

Jablensky, A., Sartorius, N., Gulbinat, W., & Ernberg, G. (1981). Characteristics of depressive patients contacting psychiatric services in four cultures. *Acta Psychiatrica Scandinavia, 68,* 367–383.

Kleinman, A. (1977). Depression, somatization and the "new cross-cultural psychiatry." *Social Science and Medicine, 11,* 3–10.

Kleinman, A. (1982). Neurasthenia and depression: A study of somatization and culture in China. *Culture, Medicine and Psychiatry, 6,* 117–190.

Kleinman, A. (1986). *Social origins of distress and disease: Neurasthenia, depression and pain in modern China.* New Haven, CT: Yale University Press.

Kleinman, A. (1988a). *Rethinking psychiatry: From cultural category to personal experience.* New York: The Free Press.

Kleinman, A. (1988b). *The illness narratives: Suffering, healing, and the human condition.* New York: Basic Books.

Lee, S., & Wong, K. C. (1995). Rethinking neurasthenia: The illness concepts of Shejing Shuairuo among Chinese undergraduates in Hong Kong. *Culture, Medicine, and Psychiatry, 19,* 91–111.

Leong, F. T. L. (1986). Counseling and psychotherapy with Asian-Americans: Review of the literature. *Journal of Counseling Psychology, 33,* 196–206.

Lewis-Fernandez, R. (1996). Cultural formulation of psychiatric diagnosis. *Culture, Medicine, and Psychiatry, 20,* 133–144.

Lewis-Fernandez, R., & Kleinman, A. (1994). Culture, personality, and psychopathology. *Journal of Abnormal Psychology, 103,* 67–71.

Lin, T. Y. (1982). Culture and psychiatry: A Chinese perspective. *Australian and New Zealand Journal of Psychiatry, 16,* 235–245.

Lin, T. Y. (1989). Neurasthenia revisited: Its place in modern psychiatry. *Culture, Medicine, and Psychiatry, 13,* 105–129.

Lin, T. Y., Tardiff, K., Donetz, G., & Goresky, W. (1978). Ethnicity and patterns of help-seeking. *Culture, Medicine, and Psychiatry, 2,* 3–13.

Marsella, A. J., Kinzie, D., & Gorson, P. (1973). Ethnic variation in the expression of depression. *Journal of Cross-cultural Psychology, 4,* 435–358.

Mechanic, D. (ed.) (1982). *Symptoms, illness behavior and help seeking.* New York: Prodist.

Mental Health Division, WHO (in press). *A sociocultural lexicon for psychiatry.* Geneva, Switzerland: World Health Organization.

Ots, T. (1990). The angry liver, the anxious heart and the melancholy spleen: The phenomenology of perceptions in Chinese culture. *Culture, Medicine, and Psychiatry, 14,* 21–58.

Pederson, P. P. (1983). The cultural complexity of counseling. *International Journal for the Advancement of Counselling, 6,* 177–192.

Pedersen, P. P. (1987). Ten frequent assumptions of cultural bias in counselling. *Journal of Multicultural Counseling and Development, 15,* 16–24.

Pedersen, P. P. (1990). The multicultural perspective as a fourth force in counseling. *Journal of Mental Health Counseling, 12,* 93–95.

Pedersen, P. P. (1991). Multiculturalism as a generic approach to counseling. *Journal of Counseling and Development, 70,* 6–12.

Pedersen, P. P., Draguns, J. G., Lonner, W. J., & Trimble, J. E. (Eds.) (1981). *Counseling across cultures* (Rev. and expanded ed.). Hawaii: East-West Center, University of Hawaii.

Pedersen, P. P. & Marsella, A. J. (1982). The ethical crisis for cross-cultural counseling and therapy. *Professional Psychology: Research and Practice, 13,* 492–500.

Prince, R. (1985). The concept of culture-bound syndromes: Anorexia nervosa and brain-fag. *Social Science and Medicine, 21,* 197–203.

Prince, R., & Tcheng-Laroche, F. (1987). Culture-bound syndromes and international disease classification. *Culture, Medicine, and Psychiatry, 11,* 3–20.

Rin, H., & Huang, M. G. (1989). Neurasthenia as nosological dilemma. *Culture, Medicine, and Psychiatry, 13,* 215–226.

Ritenbaugh, C. (1982). Obesity as a culture-bound syndrome. *Culture, Medicine, and Psychiatry, 6,* 347–361.

Routh, D. K. (1994). *Clinical psychology since 1917: Science, practice, and organization.* New York: Plenum.

Rubel, A. J., O'Nell, C. W., & Collado, R. (1985). The folk illness called *Susto.* In R. C. Simons & C. C. Hughes (Eds.), *The culture-bound syndromes: Folk illnesses of psychiatric and anthropological interest* (pp. 33–350). Dordrecht, The Netherlands: Reidel.

Sartorius, N., Jablensky, A., Korten, A., Ernberg, G., Anker, M., Cooper, J. E., & Day, R. (1986). Early manifestation and first contact incidence of schizophrenia: A preliminary report on the initial evaluation phase of the WHO Collaborative Study on Determinants of Outcome of Severe Mental Disorders. *Psychological Medicine, 16,* 909–928.

Schwartz, G. E. (1982). Testing the biopsychosocial model: The ultimate challenge facing behavioral medicine? *Journal of Consulting and Clinical Psychology, 50,* 1040–1053.

Simons, R. C. (1985). The resolution of the *Latah* paradox. In R. C. Simons & C. C. Hughes (Eds.), *The culture-bound syndromes: Folk illnesses of psychiatric and anthropological interest* (pp. 43–62). Dordrecht, The Netherlands: Reidel.

Simons, R. C. & Hughes, C. C. (Eds.) (1985). *The culture-bound syndromes: Folk illnesses of psychiatric and anthropological interest.* Dordrecht, The Netherlands: Reidel.

Singer, K. (1975). Depressive disorders from a transcultural perspective. *Social Science and Medicine, 9,* 289–301.

Sue, D. W. (1977). Counseling the culturally different: A conceptual analysis. *Personnel and Guidance Journal, 55,* 422–425.

Sue, D. W. (1978). Eliminating cultural oppression in counseling: Toward a general theory. *Journal of Counseling Psychology, 25,* 419–428.

Sue, D. W. (1989). Racial/cultural identity development among Asian-Americans: Counseling theory implications. *Asian-American Psychological Association Journal, 13,* 80–86.

Sue, D. W. (1990). Culture-specific strategies in counseling: A conceptual framework. *Professional Psychology: Research and Practice, 21,* 424–433.

Sue, D. W. (1991). A model for cultural diversity training. *Journal of Counseling and Development, 70,* 99–105.

Sue, D. W., Arredondo, P., & McDavis, R. J. (1992). Multicultural counseling competencies and standards: A call to the profession. *Journal of Multicultural Counseling and Development, 20,* 64–88.

Sue, D. W., & Kirk, B. A. (1975). Asian-Americans: Use of counseling and psychiatric services on a college campus. *Journal of Counseling Psychology, 22,* 84–86.

Sue, D. W., & Sue, D. (1977). Barriers to effective cross-cultural counseling. *Journal of Counseling Psychology, 24,* 420–429.

Sue, D. W., & Sue, D. (1990). *Counseling the culturally different: Theory and practice* (2nd ed.). New York: Wiley.

Sue, S. (1977). Psychological theory and implications for Asian Americans. *Personnel and Guidance Journal, 55,* 381–389.

Sue, S., & Zane, N. (1987). The role of culture and cultural techniques in psychotherapy: A critique and reformulation. *American Psychologist, 42,* 37–45.

Swartz, L. (1985). Anorexia nervosa as a culture-bound syndrome. *Social Science and Medicine, 20,* 725–730.

Tseng, W. S. (1975). The nature of somatic complaints among psychiatric patients: The Chinese case. *Comprehensive Psychiatry, 16,* 237–245.

Tseng, W. S., & McDermott, J. F., Jr. (1981). *Culture, mind and therapy: An introduction to cultural psychiatry.* New York: Brunner/Mazel.

Tung, M. P. M. (1994). Symbolic meanings of the body in Chinese culture and "somatization." *Culture, Medicine, and Psychiatry, 18,* 483–492.

Vontress, C. E. (1974). Barriers in cross-cultural counseling. *Counseling and Values, 18,* 160–165.

Wong, C. K., & Chan, T. S. C. (1984). Somatic symptoms among Chinese psychiatric patients. *Hong Kong Journal of Mental Health, 13,* 5–21.

World Health Organization (1973). *Report of the international pilot study of schizophrenia.* Geneva, Switzerland: Author.

World Health Organization (1975). *Schizophrenia, a multinational study: A summary of the initial evaluation phase of the international pilot study of schizophrenia.* Geneva, Switzerland: Author.

World Health Organization (1979). *Schizophrenia: An international follow-up study.* Chichester, UK: Wiley.

World Health Organization (1983). *Depressive disorders in different cultures: Report of the WHO collaborative study of standardized assessment of depressive disorders.* Geneva, Switzerland: Author.

World Health Organization (1991). *Evaluation on methods for treatment of mental disorders: Report of a WHO scientific group on the treatment of psychiatric disorders.* Geneva, Switzerland: Author.

World Health Organization (1992). *The ICD-10 classification of mental and behavioral disorders: Clinical descriptions and diagnostic guidelines.* Geneva, Switzerland: Author.

Yap, P. M. (1974). *Comparative psychiatry: A theoretical framework.* Toronto, Canada: University of Toronto Press.

Zane, N. W., Sue, S., Hu, L. T., & Kwon, J. H. (1991). Asian-American assertion: A social learning analysis of cultural differences. *Journal of Counseling Psychology, 38,* 63–70.

Zheng, Y. P., Xu, L. Y., & Shen, Q. J. (1986). Styles of verbal expression of emotional and physical experiences: A study of depressed patients and normal controls in China. *Culture, Medicine, and Psychiatry, 10,* 231–243.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.03

Cultural Bias in Testing of Intelligence and Personality

CECIL R. REYNOLDS
Texas A&M University, College Station, TX, USA

10.03.1 INTRODUCTION 53

10.03.2 THE CONTROVERSY OVER BIAS IN PSYCHOLOGICAL TESTING: WHAT IT IS AND WHAT IT IS NOT 54

10.03.3 THE NATURE OF PSYCHOLOGICAL TESTING ADDS TO THE CONTROVERSY 57

10.03.4 WHAT ARE POSSIBLE SOURCES OF BIAS? 59

10.03.5 MEAN SCORE DIFFERENCES AS TEST BIAS 60
- *10.03.5.1 Culture-free Tests, Culture Loading, and Culture Bias* 61
- *10.03.5.2 The Question of Labeling Effects* 62

10.03.6 THE PROBLEM OF DEFINITION 64

10.03.7 RESEARCH STRATEGIES AND RESULTS 65
- *10.03.7.1 Bias in Content Validity* 65
- *10.03.7.2 Bias in Construct Validity* 72
- *10.03.7.3 Bias in Predictive or Criterion-related Validity* 77

10.03.8 CROSS-CULTURAL TESTING WHEN TRANSLATION IS REQUIRED 85

10.03.9 SUMMARY AND FUTURE DIRECTIONS 85

10.03.10 REFERENCES 86

10.03.1 INTRODUCTION

The issues of bias in psychological testing have been a source of intense and recurring social controversy throughout the history of mental measurement. In England, the issue was raised by Burt (1921) early in the twentieth century. The first investigations into cultural bias, however, can be traced to Binet, originating around 1910 in France (Binet & Simon, 1916) and to Stern (1914) shortly therafter. In the USA, discussions pertaining to test bias are frequently accompanied by emotionally laden polemics decrying the use of mental tests with any minority group member, since ethnic minorities have not been exposed to the cultural and environmental circumstances and values of the so-called White middle class. Intertwined within the general issue of bias in tests has been the more specific question of whether intelligence tests should be used for educational purposes. Although scientific and societal discussion pertaining to differences among groups on measures of cognitive or intellectual functioning in no way fully encompasses the broader topic of bias in mental measurement, there is little doubt that the so-called "IQ controversy" has received the lion's share of public scrutiny over the years. It has been the subject of numerous publications in the more popular press (see Gould, 1981;

Herrnstein & Murray, 1994; or Jensen, 1980, chap. 1), and court actions and legislation have addressed the use of IQ tests within schools and industry.

From Binet to Jensen, many professionals have addressed the problem, with varying and inconsistent outcomes. Unlike the pervasive and polemical nature–nuture argument, the bias issue was until the 1970s largely restricted to the professional literature, except for a few early discussions in the popular press (e.g., Freeman, 1923; Lippmann, 1923a, 1923b). Of some interest is the fact that one of the psychologists who initially raised the question was the then-young Cyril Burt (1921), who even in the 1920s was concerned about the extent to which environmental and motivational factors affected performance on intelligence tests. Within the last 30 years, however, the questions of cultural test bias have burst forth as major contemporary problems far beyond the bounds of scholarly academic debate in psychology. The debate over bias has raged in both the professional and the popular press for multiple decades (e.g., Fine, 1975). Entangled in the larger issues of individual liberties, civil rights, and social justice, the bias issue has become a focal point for psychologists, sociologists, politicians, and the public. Increasingly, the issues have become political and legal ones, as reflected in numerous court cases and passage in the state of New York and consideration elsewhere of what is popularly known as "truth-in-testing" legislation. The magnitude—and the uncertainty—of the controversy and its outcome is shown in two highly publicized US Federal district court cases. The judiciary's answer to the question "Are the tests used for pupil assignment to classes for the educably mentally retarded biased against cultural and ethnic minorities?" initially was "Yes" in California (Larry P. et al. v. Wilson Riles et al., 1979) and "No" in Illinois (PASE v. Hannon et al., 1980), although the Larry P. finding has now been overturned, giving a consistent nature to court findings in the USA.

Unfortunately, we are all prisoners of our language. The word bias has several meanings, not all of which are kept distinct. In relation to the present issue, bias as "partiality towards a point of view or prejudice" and bias as "a statistical term referring to a constant error of a measure in one specific direction as opposed to random error" frequently become coalesced. If the latter meaning did not drag along the excess baggage of the former, the issue of bias in mental testing would be far less controversial and emotional than it is. However, as indicated in the *Oxford English Dictionary*, bias as partiality or prejudice can be traced at least to the sixteenth century and clearly antedates the statistical meaning. Nevertheless, the discussion of bias in psychological testing as a scientific issue should concern only the statistical meaning, whether or not there is systematic error in the measurement of a psychological attribute as a function of membership in one or another cultural or racial subgroup (Reynolds, 1982b). This definition, defined more technically as required later, will be followed throughout this chapter.

10.03.2 THE CONTROVERSY OVER BIAS IN PSYCHOLOGICAL TESTING: WHAT IT IS AND WHAT IT IS NOT

Systematic group differences on standardized intelligence and aptitude tests occur as a function of socioeconomic level, race or ethnic background, and other demographic variables throughout the various countries of the world. Black–White differences on IQ measures in the USA have received extensive investigation since the 1930s. The preponderance of these studies have been reviewed by Shuey (1966), Tyler (1965), Jensen (1980), and Willerman (1979). Results have not changed fundamentally in the last century. Although the results occasionally differ slightly, depending on the age groups under consideration, random samples of Blacks and Whites in the USA show a mean score difference of about one standard deviation, with the mean score of the Whites consistently exceeding that of the Black groups. The differences have persisted at relatively constant levels for quite some time and under a variety of methods of investigation. The exception to this is the reduction of the Black–White IQ difference on the Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983) to about 0.5 standard deviations on the intelligence portion of the scale, a controversial and poorly understood finding (see Kamphaus & Reynolds, 1987, for a discussion). These findings are consistent only for the US Black population, however, and other, quite diverse findings appear for African and other Black populations (e.g., see Jensen, 1980).

When a number of demographic variables are taken into account (most notably socioeconomic status), the size of the mean Black–White difference in the USA reduces to 0.5–0.7 standard deviations (e.g. Jensen, 1980; Kaufman, 1973; Kaufman & Kaufman, 1973; Reynolds & Gutkin, 1981), but remains robust in its appearance. All studies of racial and ethnic group differences on ability tests do not show higher levels of performance by Whites. Although not nearly as thoroughly researched as Black–White groups, Oriental groups have been shown to perform consistently as well as, or better than, White groups (Pintner, 1931; Tyler, 1965; Willerman, 1979). Depending on the specific aspect of intelligence under investigation, other racial and ethnic groups show performance at or above the performance level of White groups. There has been argument over whether any racial differences in intelligence are real or even researchable (e.g., Schoenfeld, 1974), but the reliability across studies is very high, and the existence of the differences is now generally accepted. It should always be kept in mind, however, that the overlap among the distributions of intelligence test scores for the different races is much greater than the degree of differences among the various groups. There is always more within-group variability than between-group variability in performance on psychological tests. The differences are nevertheless real ones and are unquestionably complex (e.g., Reynolds & Jensen, 1983).

The issue at hand is the explanation of those group differences. It should be emphasized that both the lower scores of some groups and the higher scores of others need to be explained, although not necessarily, of course, in the same way. The problem was clearly stated by Eells in his classic study of cultural differences:

> Do the higher test scores of the children from high socioeconomic backgrounds reflect genuine superiority in inherited, or genetic, equipment? Or do the high scores result from a superior environment which has brought about real superiority of the child's "intelligence"? Or do they reflect a bias in the test materials and not any important differences in the children at all? (Eells, Davis, Havighurst, Herrick, & Tyler, 1951, p. 4).

Eells et al. also concisely summarized the cultural-test-bias hypothesis as it applied to differences in socioeconomic status (SES):

> If (a) the children from different social-status levels have different kinds of experiences and have experiences with different types of material, and if (b) the intelligence tests contain a disproportionate amount of material drawn from the cultural experiences with which pupils from the higher social-status levels are more familiar, one would expect (c) that children from the higher social-status levels are more familiar, one would expect (d) that children from the higher social-status levels would show higher IQs than those from the lower levels. This argument tends to conclude that the observed differences in pupil IQs are artifacts dependent upon the specific content of the test items and do not reflect accurately any important underlying ability in the pupils. (Eells et al., 1951 p. 4)

Eells was aware that his descriptions were oversimplifications and that it was unlikely that all of the observed group differences could be explained by any one of the three factors alone. Loehlin, Lindzey, and Spuhler (1975) concluded that all three factors were probably involved in racial differences in intelligence as have a myriad of other researchers (e.g., Bouchard & Segal, 1985; Flynn, 1991). In its present, more complex form, the cultural test bias hypothesis considers other factors than culture-loaded items, as will be seen below. But the basics of Eell's summary of the cultural-test-bias hypothesis still hold: group differences stem from characteristics of the test or from aspects of test administration. Because mental tests are based largely on middle-class White values and knowledge, they are more valid for those groups and are biased against other groups to the extent that these groups deviate from those values and knowledge bases.

This position has been reframed slightly over the years, principally by Mercer (e.g., 1979) and reinforced by Helms (1994), who argue the lower scores of ethnic minorities on aptitude measures can be traced to the Anglocentrism (degree of adherence to White, middle-class value systems of USA families) of aptitude measures. Mercer (1979) developed an entire system of assessments designed to provide complex demographic corrections to IQs obtained by ethnic minorities that had the effect of equating these groups on mean scores. (This system, known as the SOMPA, had quite a bit of popularity for several years but is rarely used today because of its conceptual and psychometric inadequacies.) Lonner (1985) discusses similar issues under the rubric of cultural isomorphism in testing and assessment. Helms (1994) makes similar criticisms of ability tests, rejects most psychometric research on these issues, and posits (quite similar to Mercer's position) it is the Eurocentric nature of aptitude tests that produces artifactual differences in mean levels of performance across ethnic lines, focusing especially on the performance of Black Americans. In all of those conceptual models, which are essentially contemporaneous even with arguments of Burt as early as 1921, ethnic and other group differences in mean levels of performance on aptitude measures are seen to result from flawed psychometric methodology and not from actual differences in aptitude (see also Harrington, 1975, 1976). As described below, this hypothesis reduces to one of differential validity, the hypothesis of differential validity for mental tests being that tests measure intelligence more accurately, and make valid predictions about the level of intellectual functioning for individuals from the groups on which the tests are mainly based than for those from other groups. Artifactually low scores on an aptitude test could lead to denial of employment or in schools to pupil misassignment to educational programs and unfair denial of admission to college, graduate school, or other programs or occupations in which such test scores are an important decision-making component. This is the issue over which most legal cases have been fought in the USA.

Further, there would be dramatic implications for whole areas of psychological research and practice if the cultural-test-bias hypothesis is correct. The principal research of the twentieth century in the psychology of human differences would have to be dismissed as confounded and largely artifactual because much of the work is based on standard psychometric theory and testing technology. The result would be major upheavals in the practice of applied psychology, as the foundations of clinical, school, counseling, and industrial psychology are strongly tied to the basic academic field of individual differences. The issue, then, is crucial not only to the science of psychology but to practice (Lonner, 1985; Reynolds, 1980c).

On the other hand, if the cultural-test-bias hypothesis is incorrect, then group differences are not attributable to the tests and must be due to one of the other factors mentioned by Eells et al. (1951) or to some combination of them. That group differences in test scores reflect real group differences in ability should be admitted as a possibility, and one that calls for scientific study.

The controversy over test bias should not be confused with that over the etiology of any obtained group differences in test scores (see Reynolds & Kaiser, 1990, for a review) other than to rule bias in or out as an explanation. Unfortunately, it has often been inferred that measured differences themselves indicate genetic differences, and therefore the genetically based intellectual inferiority of some groups. Jensen has himself consistently argued since 1969 that mental tests measure, to a greater or lesser extent, the intellectual factor *g* which has a large genetic component, and that group differences in mental test scores may then reflect group differences in *g*. Unless one reads Jensen's statements carefully, it is easy to overlook the many qualifications that he makes regarding these differences and conclusions.

But, in fact, Jensen's or anyone else's position on the genetic basis of actual group differences should be seen as irrelevant to the issue of test bias. However controversial, etiology is a separate issue. It would be tragic to accept the cultual test bias hypothesis as true if it is, in fact, false. In that case, measured differences would be seen as not real, and children, for example, might be denied access to the educational environment best suited to them, and job applicants may end up in occupations for which they are ill-suited or for which others are more qualified. Further, research on the basis of any group differences would be stifled, as would the implementation of programs designed to remediate any deficiencies. The most advantageous position for the true White racist and bigot would be to favour the test bias hypothesis! Acceptance of that hypothesis inappropriately would eventually result in inappropriate pupil assignment, less adaptive education for some groups, and less implementation of long-range programs to raise intellectual performance. Inappropriate confirmation of the test bias hypothesis would appear to maintain, not break down, the poverty cycle (Birch & Gussow, 1970).

The controversy is also not over the blatantly inappropriate administration and use of mental tests. The administration of a test developed and published in English to an individual for whom English is a second language and whose English language skills are poor is inexcusable, regardless of any bias in the tests themselves. It is just as inexcusable to translate the test into the examinee's first or dominant language and assume the validity of the results. Translating a test from its original language is a wholly inadequate process and creates many problems of reliability and validity. Instead, a test must be redeveloped, preferably following the steps outlined and discussed by Hambleton and Kanjee (1995). It is of obvious importance that tests be administered by skilled and sensitive professionals who are aware of the factors that may artifactually lower an individual's test scores. Considering the use of tests to assign pupils to special education classes or other programs, a question needs to be asked: What would one use instead? Teacher recommendations are notoriously less reliable and less valid than standardized test scores, which increases the probability of bias occurring. For college admissions, this question was asked long ago and the use of grades and subjective letters of recommendation and interviews found to be unreliable and biased. This research in the early 1900s led to the formation of the College Entrance Examination Board and the Scholastic Achievement Test (SAT).

The controversy over the use of mental tests is further complicated by the fact that resolution of the cultural test bias question in either direction will not resolve the problem of the role of nonintellective factors that may influence the test scores of individuals from any ethnic group. Regardless of any group differences, it is individuals who are tested and whose scores may or may not be accurate. Similarly, it is individuals who are assigned to classes, chosen for universities, placed in jobs or vocations, and who are accepted or rejected. As indicated by Wechsler (1975) and others, nonintellective factors, informational content, and emotional–motivational conditions may be reflected in performance on mental tests. The extent to which these factors influence individual as opposed to group performance is difficult to determine. Perhaps with more sophisticated multivariate designs, we will be better able to identify individuals with characteristics that are likely to have an adverse effect on their performance on mental tests. Basically, outside the major thrust of the issue of bias against groups, potential bias against individuals is a serious problem itself, and merits research and analysis. Sternberg (1980), also concerned about individual performance, observed a still true matter: research on bias has concentrated on status variables such as ethnicity rather than on functional variables such as cognitive styles and motivations.

10.03.3 THE NATURE OF PSYCHOLOGICAL TESTING ADDS TO THE CONTROVERSY

The question of bias in mental testing arises largely because of the nature of psychological processes and the measurement of those processes (Reynolds & Brown, 1984). Psychological processes, by definition internal and not directly subject to observation or measurement, must be inferred from behavior. Theoretically, in the classic discussion by MacCorquodale and Meehl (1948), a psychological process has the status of an "intervening variable" if it is used only as a component of a system that has no properties beyond those that operationally define it, but it has the status of a "hypothetical construct" if it is thought actually to exist with properties beyond the defining ones. A historical example of a hypothetical construct is gene, which had meaning beyond its use to describe the cross-generational transmission of characteristics. Intelligence, from its treatment in the professional literature, has the status of a hypothetical construct as does personality.

As even beginning psychology students know, it is difficult to determine one-to-one relationships between observable events in the environment, the behavior of an organism, and hypothesized underlying mediational processes. Many classic controversies over theories of learning have revolved around constructs such as expectancy, habit, and inhibition (e.g., Goldstein, Krantz, & Rains, 1965; Hilgard & Bower, 1975; Kimble, 1961). Disputes among different camps in learning have been polemical and of long duration. Indeed, there are still disputes about the nature and the number of processes such as emotion and motivation (e.g., Bolles, 1975; Mandler, 1975). One of the major areas of disagreement has been over the measurement of psychological processes. It should be expected that intelligence, as one of the most complex of psychological processes, involves definitional and measurement disputes that prove difficult to resolve.

Assessment of intelligence and personality, like that of many other psychological processes in humans, is accomplished by standard psychometric procedures that are the focus of the bias issue. These procedures, described in detail in general assessment texts (e.g., Anastasi, 1982; Linn, 1989), are only briefly summarized here in relation to the issue of bias. Problems specific to validity are discussed in Section 10.03.7.

Similar procedures are used in the development of any standardized psychological test. First, a large number of items are developed that for theoretical or practical reasons are thought to measure the construct of interest. Through a series of statistical steps, those items that best measure the construct in a unitary manner are selected for inclusion in the final test battery. The test is then administered to a sample, which should be chosen to represent all aspects of the population on whom the test will be used. Normative scales based on the scores of the standardization sample then serve as the reference for the interpretation of scores of individuals tested thereafter. Thus, as has been pointed out numerous times, an individual's score is meaningful only relative to this normative base and is a relative, not an absolute, measure. Charges of bias frequently arise from the position that the test is more appropriate for the groups heavily represented in the standardization sample. Whether bias does, in fact, result from this procedure is a specific question to be addressed empirically.

As has also been frequently pointed out, there are few charges of bias of any kind relating to physical measures that are on absolute scales, whether interval or ratio. Group differences in height, as an extreme example, are not attributed by anyone we know of to any kind of cultural test bias or specific methods of measuring height. There is no question concerning the validity of measures of the height or weight of anyone in any culture. Nor is there any question about one's ability to make cross-cultural comparisons of these absolute measures (which variables, such as height, are clearly subject to genetic by environmental interactions).

The whole issue of cultural bias arises because of the procedures involved in psychological testing. Psychological tests measure traits that are not directly observable, that are subject to differences in definition, and that are measurable only on a relative scale. From this perspective, the question of cultural bias in mental testing is a subset—obviously of major importance—of the problems of uncertainty and of possible bias in psychological testing generally. Bias may exist not only in mental tests but in other types of psychological test, including personality, vocational, and psychopathological tests. Making the problem of bias in mental testing even more complex is the fact that not all mental tests are of the same quality. Like Orwell's pigs, some may be more equal than others. There is a tendency for critics and defenders alike to overgeneralize across tests, lumping virtually all tests together under the heading "mental tests" or "intelligence tests." As reflected in the *Mental measurements yearbook* (e.g., Buros, 1978), professional opinions of mental tests vary considerably, and some of the most used tests are not well respected by psychometricians. Thus, unfortunately, the question of bias must eventually be answered on a virtually test-by-test basis.

In 1969, the Association of Black Psychologists (ABP), a USA-based professional association, adopted the following official policy on educational and psychological testing:

> The Association of Black Psychologists fully supports those parents who have chosen to defend their rights by refusing to allow their children and themselves to be subjected to achievement, intelligence, aptitude and performance tests which have been and are being used to A. Label Black people as uneducable. B. Place Black children in "special" classes and schools. C. Perpetuate inferior education in Blacks. D. Assign Black children to educational tracts. E. Deny Black students higher educational opportunities. F. Destroy positive growth and development of Black people. (quoted in Reynolds, 1982a, p. 179)

Since 1968, the ABP has sought continuously a moratorium on the use of all psychological and educational tests with the culturally different (Samuda, 1975, and Williams, Dotson, Dow, & Williams, 1980, have provided a more detailed history of these efforts). The ABP, an organizational leader in protesting test bias for all ethnic groups, carried its call for a moratorium to other professional organizations in psychology and education. In direct response to the ABP call, the Board of Directors of the American Psychological Association (APA) requested its Board of Scientific Affairs to appoint a group to study the use of psychological and educational tests with disadvantaged students. The committee report (Cleary, Humphreys, Kendrick, & Wesman, 1975) was subsequently published in the official journal of the APA, *American Psychologist*.

Subsequent to the ABP's policy statement, other groups adopted policy statements on testing. These groups included the National Association for the Advancement of Colored People (NAACP), the National Education Association (NEA), the National Association of Elementary School Principals (NAESP), the American Personnel and Guidance Association (APGA), and others (Committee on Testing, 1974; Williams et al., 1980). The APGA called for the Association for Measurement and Evaluation in Guidance (AMEG), a sister organization, to develop and disseminate a position paper stating the limitations of group intelligence tests particularly and generally of standardized psychological, educational, and employment testing for low socioeconomic and underprivileged and non-White individuals in educational, business, and industrial environments.

The APGA also resolved that, if no progress was made in clarifying and correcting the current testing of minorities, it would also call for a moratorium, but only on the use of group intelligence tests with these groups.

The NAACP adopted a more detailed resolution and joined in the call for a moratorium on standardized testing of minority groups at its annual meeting in 1974. The text of the NAACP resolution was:

> Whereas a disproportionately large number of Black students are being misplaced in special education classes and denied admissions to higher educational opportunities,
>
> Whereas students who fail to show a high verbal or numerical ability, score low on the Scholastic Achievement Test (SAT), the Law School Admissions Test (LSAT), the Graduate Record Examination (GRE), etc., and are routinely excluded from college and graduate or professional education,
>
> Be it resolved that the NAACP demands a moratorium on standardized testing whenever such tests have not been corrected for cultural bias and directs its units to use all administrative and legal remedies to prevent the violation of students" constitutional rights through the misuse of tests, and
>
> Be it further resolved that the NAACP calls upon the Association of Black Psychologists to assert leadership in aiding the College Entrance Examination Board to develop standardized tests which have been corrected for cultural bias and which fairly measure the amount of knowledge retained by students regardless of his or her individual background.
>
> Be it finally resolved that the NAACP directs its units to use all administrative remedies in the event of violation of students' constitutional rights

> through the misuse of tests and directs National Office staff to use its influence to bring the CEEB and ABP together to revise such tests.

Also in 1974, the Committee on Testing of the ABP issued a position paper on the testing of Blacks that described their intent as well as their position:

> (1) To encourage, support and to bring action against *all* institutions, organizations and agencies who continue to use present psychometric instruments in the psychological assessment of Black people;
>
> (2) To continue efforts to bring about a cessation of the use of standard psychometric instruments on Black people until culturally specific tests are made available;
>
> (3) To establish a national policy that in effect gives Black folk and other minorities the right to demand that psychological assessment be administered, interpreted, and supervised by competent psychological assessors of their own ethnic background;
>
> (4) To work toward and encourage efforts to remove from the records of all Black students and Black employees that data obtained from performance on past and currently used standard psychometric, achievement, employment, general aptitude and mental ability tests;
>
> (5) To establish a national policy that demands the appropriate proportional representation of competent Black psychologists on all committees and agencies responsible for the evaluation and selection of tests used in the assessment of Black folk;
>
> (6) To establish a national policy that demands that all persons engaged in the evaluation, selection and placement of Black folks undergo extensive training so they may better relate to the Black experience;
>
> (7) To demand that all Black students improperly diagnosed and placed into special education classes be returned to regular class programs;
>
> (8) To encourage and support all suits against any public or private agency for the exclusion, improper classification and the denial of advancement opportunities to Black people based on performance tests.

It should be noted that the statements by these organizations have assumed that present tests are biased, and that what is needed is the removal of that assumed bias. These assumptions continue in the work of Helms (e.g., 1992), Mercer (1979), Padilla and Medina (1996), and others (e.g., Guilford Press, 1997).

10.03.4 WHAT ARE POSSIBLE SOURCES OF BIAS?

Many potentially legitimate objections to the use of educational and psychological tests with minorities have been raised by Black and other minority psychologists. Unfortunately, these objections are frequently stated as facts on rational rather than empirical grounds (e.g., Council for Exceptional Children, 1978; Chambers, Barron, & Sprecher, 1980; Dana, 1996; Helms, 1992; Hilliard, 1979). The most frequently stated problems fall into one of the following categories:

(i) *Inappropriate content.* Black and other minorities have not been exposed to the material involved in the test questions or other stimulus materials. The tests are geared primarily toward the majority class homes, vocabulary, and values. Different value systems among cultures may lead to cognitively equivalent answers scored as incorrect based on prejudicial value judgments, not on differences in ability.

(ii) *Inappropriate standardization samples.* Ethnic minorities are underrepresented in standardization samples used in the collection of normative reference data. Population proportionate sampling with stratification by ethnicity is the herald for standardization samples for tests and is done to enhance the accuracy of parameter estimation for scaling purposes. As such, although presented proportionately, ethnic minorities may appear in test standardization samples in small absolute numbers. This may bias item selection (e.g., Harrington, 1975, 1976) and fails to have any impact of significance on the tests themselves from these ethnic groups. In earlier years, it was not unusual for standardization samples to be all White (e.g., the 1937 Binet and 1949 WISC).

(iii) *Examiner and language bias.* Since most psychologists in the USA are White and speak only standard English, they may intimidate Black and other ethnic minorities. They are also unable accurately to communicate with minority children—to the point of being insensitive to ethnic pronunciation of words on the test. Lower test scores for minorities, then, may reflect only this intimidation and difficulty in the communication process or motivational differences (Zigler & Butterfield, 1967), not lower ability. In other countries, professionals such as psychologists tend to be from the dominant culture and similar issues appear.

(iv) *Inequitable social consequences.* As a result of bias in educational and psychological tests, minority group members, already at a disadvantage in the educational and vocational markets because of past discrimination and thought unable to learn, are disproportionately relegated to dead-end educational tracks. Labeling effects also fall under this category.

(v) *Measurement of different constructs.* Related to (i), this position asserts that the tests measure different attributes when used with children from other than the majority culture of a country, the culture on which the tests are largely based, and thus do not measure minority intelligence or personality validly.

(vi) *Differential predictive validity*. Although tests may accurately predict a variety of outcomes for members of a majority culture within a population, they do not predict successfully any relevant behavior for cultural minority group members Further, there are objections to use of the standard criteria against which tests are validated with minority cultural groups. For example, in the USA, scholastic or academic attainment levels in White middle-class schools are themselves considered by a variety of Black psychologists to be biased as criteria (e.g., see discussion in Reynolds, 1982a, pp. 179–180).

(vii) *Minority and majority aptitude and personality are qualitatively distinct*. Championed by Helms (e.g., 1992), this position would lead to the conclusion that ethnic minorities and the majority culture are so different as to require different conceptualizations of ability and personality. Helms (1982), for example, argues the potential existence of a "White *g*" factor that is separate from an "African *g*" factor (p. 1090) that would necessitate separate tests for these groups.

Contrary to the situation of the late 1960s and 1970s, when the current controversies resurfaced after some decades of simmering, research now exists that examines the above areas of potential bias in assessment. Except for the still unresolved issue of labeling effects, the least amount of research is available on the long-term social consequences of testing, although some limited data are available (e.g., Lambert, 1979). Both of these problems are aspects of testing in general and are not limited to minorities. The problem of the social consequences of educational tracking is frequently lumped with the issue of test bias. Those issues, however, are separate. Educational tracking and special education should be treated as problems of education, not assessment. These are going to become more heavily contested areas for psychologists in the future, however, as the revision of the *Joint Technical Standards for Educational and Psychological Tests and Manuals* now underway is likely to add the issue of consequential validity to considerations in test use, reflecting concerns by minority groups within the USA over the consequences of test use with ethnic minorities. Educational tracking as a result of test performance is a prominent problem for ethnic minorities in many countries applying such competitive schooling practices.

The call by Helms (1992) and earlier by Williams (1974) for separate tests for minority groups within the USA deserves comment and raises an interesting set of legal and conceptual issues. The concept of separate but equal programs, facilities, and the like was protested vehemently by minorities for many years and ultimately rejected by the US Supreme Court. It is of interest that some minority scholars would now seek "separate but equal" tests. This could lead to arguments that if the cognitive structures of various ethnic groups are so different as to require distinct tests of intelligence and personality for accurate assessment, then schooling, job training, mental health interventions, and the like might also need to be different—an argument long attacked and rejected by many minority scholars and spokespersons.

10.03.5 MEAN SCORE DIFFERENCES AS TEST BIAS

A popular lay view, and one promoted in the popular media, has been that differences in mean levels of scoring on cognitive, achievement, or personality tests among groups constitute bias in tests; however, such differences alone clearly are not evidence of test bias. A number of writers in the professional literature have taken this position as well (e.g., Adebimpe, Gigandet, & Harris, 1979; Alley & Foster, 1978; Chinn, 1979; Guilford Press, 1997; Hillard, 1979; Jackson, 1975; Mercer, 1976; Padilla, 1988; Williams, 1974; Wright & Isenstein, 1977). Those who support this definition of test bias correctly state there is no valid *a priori* scientific reason to believe that intellectual or other cognitive performance levels should differ across race. It is the inference that tests demonstrating such differences are inherently biased because there can in reality be no differences that is fallacious. Just as there is no *a priori* basis for deciding that differences exist, there is no *a priori* basis for deciding that differences do not exist. From the standpoint of the objective methods of science, *a priori* or premature acceptance of either hypothesis ('differences exist" vs. "differences do not exist") is untenable. As stated by Thorndike (1971), "The presence (or absence) of differences in mean score between groups, or of differences in variability, tells us nothing directly about fairness" (p. 64). Some adherents to the "mean score differences as bias" viewpoint also require that the distribution of test scores in each population or subgroup be identical before one can assume that the test is fair: "Regardless of the purpose of a test or its validity for that purpose, a test should result in distributions that are statistically equivalent across the groups tested in order for it to be considered nondiscriminatory for those groups" (Alley & Foster, 1978,

p. 2). Portraying a test as biased regardless of its purpose or validy is psychometrically naive. Mean score differences and unequivalent distributions have been the most uniformly rejected of all criteria examined by sophisticated psychometricians involved in investigating the problems of bias in assessment. Ethnic group differences in mental test scores are among the best-documented phenomena in psychology, and they have persisted over time at relatively constant levels (Reynolds & Gutkin, 1980a, 1981).

Jensen (1980) has discussed the "mean score differences as bias" position in terms of the egalitarian fallacy. The egalitarian fallacy contends that all human populations are in fact identical on all mental traits or abilities. Any differences with regard to any aspect of the distribution of mental test scores indicates that something is wrong with the test itself. Such an assumption is totally scientifically unwarranted. There are simply too many examples of specific abilities and even sensory capacities that have been shown to differ unmistakably across human populations. The result of the egalitarian assumption, then, is to remove the investigation of population differences in ability from the realm of scientific inquiry. Logically followed, this fallacy leads to other untenable conclusions as well. Torrance (1980), an adherent of the cultural bias hypothesis, pointed out that disadvantaged Black children in the USA occasionally earn higher scores on creativity tests—and therefore, have more creative ability—than many White children because their environment has forced them to learn to "make do" with less and with simpler objects. The egalitarian assumption would hold that this is not true, but rather that the content of creativity tests is biased against White or high SES children.

The attachment of minorities to the "mean score differences as bias" definition is probably related to the nature–nurture controversy at some level, and has been difficult for even databased, minority research scientists to abandon (e.g., Dana, 1993). Certainly data reflecting racial differences on various aptitude measures have been interpreted to indicate support for a hypothesis of genetic differences in intelligence and to imply that one race is superior to another. However, as discussed previously, the so-called nature–nurture issue is not an inextricable component of bias investigation. Assertions as to the relative impact of genetic factors on group ability levels step into a new arena of scientific inquiry, with differing bodies of knowledge and methods of research. Suffice it to say that in the arena of bias investigation, mean differences on aptitude, achievement, or personality measures among selected groups are not evidence *per se* that the measures are biased.

10.03.5.1 Culture-free Tests, Culture Loading, and Culture Bias

A third area of bias investigation that has been confusing in both the professional (e.g., Alley & Foster, 1978; Chinn, 1979) and the lay literature has been the interpretation of culture loading and culture bias. A test can be culture loaded without being culturally biased. Culture loading refers to the degree of cultural specificity present in the test or individual items of the test. Certainly, the greater the cultural specificity of a test item, the greater the likelihood of the item being biased when it is used with individuals from other cultures. The test item "Who was the first president of the United States?" is a culture-loaded item. However, the item is general enough to be considered useful with school-aged children attending school since first grade in the USA. The cultural specificity of the item is too great, however, to allow the item to be used on an aptitude measure of 10-year-old children from other countries. Virtually all tests in current use are bound in some way by their cultural specificity. Culture loading must be viewed on a continuum from general (defining a culture in a broad, liberal sense) to specific (defining a culture in narrow, highly distinctive terms).

A variety of attempts have been made to develop a "culture-free" (sometimes referred to as "culture-fair") intelligence test (e.g., Cattell, 1979). However, the reliability and validity of these tests are uniformly inadequate from a psychometric perspective (Anastasi, 1982; Ebel, 1979). The difficulty in developing a culture-free measure of intelligence lies in the test being irrelevant to intellectual behavior within the culture under study. Intelligent behavior is defined within human society in large part on the basis of behavior judged to be of value to the survival and improvement of the culture and the individuals within the culture. A test that is "culture-blind," then, cannot be expected to predict intelligent behavior within a variety of cultural settings. Once a test has been developed within a culture (a culture-loaded test), its generalizability to other cultures or subcultures within the dominant societal framework becomes a matter for empirical investigation, and tests should not be used cross-culturally without demonstrative evidence for the validity of inferences to be drawn from such tests. The same holds true for personality tests and other techniques designed to assess affective disorders.

Although there exists a universality to certain psychopathological disorders (e.g., schizophrenia, bipolar disorder) in both taxonomy and manifestation, the existence of certain disorders independent of a cultural context can be questioned readily (e.g., dependent personality disorder) and the symptoms or manifestations of other disorders may well vary across cultures as they do across age (e.g., depression, anxiety disorders).

Jensen (1980) admonishes that when one is investigating the psychometric properties of culture-loaded tests across differing societies or cultures, one cannot assume that simple inspection of the content will determine which tests or items are biased against those cultures or societies not represented in the tests or item content. Tests or items that exhibit characteristics of being culturally loaded cannot be determined to be biased with any degree of certainty unless objective statistical inspection is completed. Jensen refers to the mistaken notion that anyone can judge tests and/or items as being "culturally unfair" on superficial inspection as the "culture-bound fallacy." The issue of item bias is revisited in some detail in Section 10.03.7.1.

10.03.5.2 The Question of Labeling Effects

The relative impact of placing a label on a person's behavior or developmental status has also been a hotly discussed issue within the field of psychometrics in general and bias investigation in particular. The issue undoubtedly has been a by-product of the practice of using intellectual measures for the determination of mental retardation. Although the question of labeling effects is a viable and important one, it requires consideration in bias research only in much the same way as does the ongoing debate surrounding the nature–nurture question. As the concept of consequential validity grows, this issue will likewise grow in importance. However, there are some important considerations regarding bias in referral for services, diagnosis, and labeling, which no interested student of the diagnostic process in psychology can afford to ignore.

Rosenthal is the researcher most closely associated with the influence of labeling upon teachers' and parents' perceptions of a child's ability and potential. Even though his early studies had many methodological and statistical difficulties, labeling effects have been shown in some subsequent experimental studies (e.g., Critchley, 1979; Foster & Ysseldyke, 1976; Jacobs, 1978), but not in others (e.g., MacMillan, Jones, & Aloia, 1974; McCoy, 1976). However, these studies have generally been of a short-term nature, and have usually been conducted under quite artificial circumstances. Typically, participants are asked to rate the behavior or degree of pathology of a child seen on videotape. Categorical labels for the child are systematically varied while the observed behaviors remain constant. The demand characteristics of such a design are substantial. Long-term effects of labeling and diagnoses resulting in special education placement or mental health treatment in real-life situations have been examined less vigorously. For example, comparisons of the effects of formal diagnostic labels with the informal, often cursory, personal labeling process that occurs between teachers and children over the course of a school year, and that is subsequently passed on to the next grade via the teachers' lounge (Dworkin & Dworkin, 1979), need to be made. Although Reynolds (1982b) called for this research in past decades, this important question has not been addressed. The strict behaviorist position (Ross, 1974, 1976) also contends that formal diagnostic procedures are unnecessary and potentially harmful because of labeling effects. However, whether or not the application of formal labels has detrimental effects remains an open question now, much as it did at the conclusion of a monumental effort to address these important questions throughout the USA in the mid-1970s (Hobbs, 1975).

Even without the application of formal, codified labels by psychologists or psychiatrists, the mental labeling, classification, and appraisal of individuals by people with whom they come into contact are common, constant occurrences (Reynolds, 1979a). Auerbach (1971) found that adults often interpret early learning difficulties of children as primarily emotional disturbances, unrelated to learning problems. According to Bower (1974), children who start the first grade below the mean age of their classmates and are below average in the development of school readiness skills or have behaviour problems are more likely to be regarded as emotionally disturbed by school staff and are more likely to be referred to residential placement than their peers. The American Psychological Association (1970) acknowledges that such constant appraisal of individuals occurs at the informal level, and in an official position statement takes the stance that specialized, standardized psychological techniques have been developed to supersede our informal, often casual approach to the appraisal of others. The specialized psychological techniques available to the trained examiner add validity and utility to the results of such appraisals. The quantification of behaviour permits systematic comparisons of individuals' characteristics with those of a selected reference or norm group. It is not unreasonable to anticipate that the informal labeling of children so often indulged in by teachers and parents is substantially more harmful than accurate psychoeducational diagnostics intended to accrue beneficial activity toward the child. Should noncategorical funding for services to exceptional children become a reality (Gutkin & Tieger, 1979), or should the use of normative assessment ultimately be banned, the informal labeling process will continue and in all likelihood will exacerbate children's problems.

From the standpoint of cultural test bias issues, the question of labeling persons or not labeling persons is moot. Cultural test bias is concerned with the accuracy of such labels across some nominal grouping system (typically, race, sex, and SES have been the variables of interest). It is a question of whether race, sex, or any other demographic variable of interest influences the diagnostic process or the placement of an individual in special treatment programs, independent of the individual's cognitive, emotional, and behavioral status. Several well-designed studies have investigated the influences of race and SES on diagnosis and the placement recommendations of school psychologists (i.e., bias in test interpretation). One of the studies investigated teacher bias as well.

Frame (1979) investigated the accuracy of school psychologists' diagnoses and consistency of treatment plans within the USA, with regard to bias effects associated specifically with race and SES. In Frame's study, 24 school psychologists from a number of school districts diagnostically rated and provided treatment plans for hypothetical cases in which all information except race, SES, and the achievement level of the child's school was held constant. No differences in the accuracy of diagnosis (as defined by interrater reliability) occurred as a function of race or SES. Differences did occur with regard to treatment recommendations, however. With all other data held constant, lower SES Black children were less likely to be recommended for special education placement than their White counterparts or higher SES Black children. A more general trend was for higher SES children to be recommended for special class placement more often than children of lower SES.

In a similar vein, Matuszek and Oakland (1979) asked whether SES and race influenced teacher or psychologist placement recommendations, independent of other characteristics such as adaptive behavior, IQ, and classroom achievement levels. Matuszek and Oakland concluded that "The data from this study clearly indicate that they [psychologists] did not make different recomendations on the basis of race." Consistent with the results of Frame (1979), psychologists were more likely to recommend special class placement for high SES status children than for low SES children when other varibles were held constant. Teachers showed no bias in regard to special education placement recommendations on the basis of race or SES. Upon investigating special education placement recommendations as a function of minority group status (Black, native American, or Oriental), Tomlinson, Acker, Canter, and Lindborg (1977) reported that psychologists recommended special education resource services more frequently for minority than for White children. Placement in a special education class, however, was recommended more frequently for White than minority children. A rather extensive study of placement in classes for the educable mentally retarded (EMR) in California also failed to find any racist intent in the placement of minority children in special classes (Meyers, MacMillan, & Yoshida, 1978). In fact, the tendency was not to place Black children in special education classes, even though they might be failing in the regular classroom. An even earlier study by Mercer (1971), one of the major critics of IQ testing with minorities, reached the same conclusion.

The general tendency not to label Black children also extends to community mental health settings. Lewis, Balla, and Shanok (1979) reported that when Black adolescents were seen in mental health settings, behaviors symptomatic of schizophrenia, paranoia, and a variety of psychoneurotic disorders were frequently dismissed as only "cultural aberrations" appropriate to coping with the frustrations created by the antagonistic White culture. Lewis et al. (1979) further noted that White adolescents exhibiting similar behaviours were given psychiatric diagnoses and referred for therapy and/or residential placement. Lewis et al. contended that this failure to diagnose mental illness in the Black population acts as bias in the denial of appropriate services. A tendency for psychologists to reward depressed performance on cognitive tasks by Blacks and low SES groups as a "cultural aberration" has also been shown. An early empirical study by Nalven, Hofmann, and Bierbryer (1969) demonstrated that psychologists generally rated the "true intelligence" of Black and lower SES children higher than that of White and middle-class children with the same Wechsler Intelligence Scale for Children (WISC) IQ. This tendency to "overrate" the intellectual potential of Black and low SES children probably accounts, at least in part, for psychologists' reluctance to recommend special education placement for these children; it could also be viewed as a discriminatory denial of services, depending on whether the provision of services is considered beneficial or harmful to the individual. Despite the outcome of these and other studies, it is common to read that psychologists overdiagnose psychopathology in cultural groups outside the mainstream culture (e.g. Dana, 1993).

Bias by ethnicity in diagnosis of specific forms of affective disturbance has been the subject of fewer examinations in the literature. However, there is little evidence to suggest that ethnicity or SES independently influence clinical diagnoses such as autistic disorders or attention-deficit hyperactivity disorder (e.g., Cuccaro, Wright, Rhownd, & Abramson, 1996).

These studies clearly indicate that the demographic variables of race and SES do not, independent of other pupil characteristics, influence or bias psychologists' diagnostic or placement behavior in a manner that would cause Blacks or lower SES children to be labeled inaccurately or placed inappropriately or in disproportionate numbers in special education programs. The empirical evidence, rather, argues in the opposite direction: in the USA, Black and low SES children are less likely to be diagnosed as having a form of psychopathology than their White or higher SES peers with similar cognitive, behavioral, and emotional characteristics. The data simply do not support William's (1970) and others (e.g., Dana, 1993; Guilford Press, 1997; Padilla, 1988) charges that ethnic minorities are more likely to be overly diagnosed with a variety of psychopathological disorders.

10.03.6 THE PROBLEM OF DEFINITION

The definition of test bias has produced considerable continuing debate among measurement and assessment experts (Angoff, 1976; Bass, 1976; Bernal, 1975; Bond, 1981; Cleary et al., 1975; Cole & Moss, 1989; Cronbach, 1976; Dana, 1993; Darlington, 1978; Einhorn & Bass, 1970; Flaugher, 1978; Gordon, 1984; Gross & Su, 1975; Helms, 1992; Humphreys, 1973; Hunter & Schmidt, 1976, 1978; Jackson, 1980; Linn, 1976; McNemar, 1975; Moreland, 1996; Novick & Petersen, 1976; Padilla, 1988; Petersen & Novick, 1976; Reschly, 1980; Reynolds, 1978; 1982b, 1995; Reynolds & Brown, 1984; Sawyer, Cole, & Cole, 1976; Schmidt & Hunter, 1974; Thorndike, 1971). Although the resulting debate has generated a number of selection models with which to examine bias, selection models focus on the decision-making system and not on the test itself. The various selection models are discussed at some length in Hunter and Schmidt (1974), Hunter, Schmidt, and Rauschenberger (1984), Jensen (1980), Peterson and Novick (1976), and Ramsay (1979). The choice of a decision-making system (especially a system for educational decision-making) must ultimately be a societal one; as such, it will depend to a large extent on the value system and goals of the society. Thus, before a model for test use in selection (whether ultimately selection is for a treatment program, a job, a college, etc.) can be chosen, it must be decided whether the ultimate goal of selection is equality of opportunity, equality of outcome, or representative equality (these concepts are discussed in more detail in Nichols, 1978).

"Equality of opportunity" is a competitive model wherein selection is based on ability. As more eloquently stated by Lewontin (1970), under equality of opportunity, "true merit ... will be the criterion of men's earthly reward" (p. 92). "Equality of outcome" is a selection model based on ability deficits. Schooling provides a good model to illustrate these concepts that are also applicable to mental health. Compensatory and remedial programs are typically constructed on the basis of the equality-of-outcome model. Children of low ability or children believed to be a high risk for academic failure are selected for remedial, compensatory, or other special educational programs. Adults vying for jobs may be placed in specialized job training programs. In a strictly predictive sense, tests are used in a similar manner under both of these models. However, under equality of opportunity, selection is based on the prediction of a high level of criterion performance; under equality of outcome, selection is determined by the prediction of "failure" or a preselected low level of criterion performance. Interestingly, it is the failure of compensatory and remedial education programs to bring the disadvantaged learner to "average" levels of performance that resulted in the charges of test bias now in vogue.

The model of "representative equality" also relies on selection, but selection that is proportionate to numerical representation of subgroups in the population under consideration. Representative equality is typically thought to be independent of the level of ability within each group; however, models can be constructed that select from each subgroup the desired proportion of individuals (i) according to relative ability level of the group, (ii) independent of group ability, or (iii) according to some decision rule between these two positions. Even under the conditions of representative equality, it is imperative to employ a selection device (test) that will rank order individuals within groups in a reliable and valid manner. The best way to ensure fair selection under any of these models is to employ tests whose scores are equally reliable and equally valid for all groups concerned. The tests employed should also yield the most reliable and most valid scores for all groups under consideration. The question of test bias *per se* then becomes a question of test validity. Test use (i.e., fairness) may be defined as biased or nonbiased only by the societal value system; at present, this value system within the USA is leaning strongly toward some variant of the representative-equality selection model. In other sociopolitical structures, other models may be more appropriate. As noted above, all models are facilitated by the use of a nonbiased test. That is, the use of a test with equivalent cross-group validities makes for the most parsimonious selection model, greatly simplifying the creation and application of the selection model that has been chosen.

This leads to the essential definitional component of test bias. “Test bias” refers in a global sense to systematic error in the estimation of some “true” value for a group of individuals. The key word here is “systematic”; all measures contain error and in all cultural settings, but this error is assumed to be random unless shown to be otherwise. Bias investigation is a statistical inquiry that does not concern itself with culture loading, labeling effects, or test use/test fairness. Concerning the last of these, Jensen (1980) comments,

> [U]nbiased tests can be used unfairly and biased tests can be used fairly. Therefore, the concepts of bias and unfairness should be kept distinct ... [A] number of different, and often mutually contradictory, criteria for fairness have been proposed, and no amount of statistical or psychometric reasoning per se can possible settle any arguments as to which is best. (pp. 375–376)

There are three types of validity as traditionally conceived: content, construct, and predictive (or criterion-related). Test bias may exist under any or all of these categories of validity. Though no category of validity is completely independent of any other category, each is discussed separately here for the purposes of clarity and convenience. (All true evidence of validity is as likely as not to be construct validity, and other, more detailed divisions including this one are for convenience of discussion.) Frequently encountered in bias research are the terms “single-group validity” and “differential validity.” Single-group validity refers to the phenomenon of a score interpretation being valid for one group but not another. Differential validity refers to a condition where an interpretation is valid for all groups concerned, but the degree of validity varies as a function of group membership. Although these terms have been most often applied to predictive or criterion-related validity (validity coefficients are then examined for significance and compared across groups), the concepts of single-group and differential validity are equally applicable to content and construct validity.

10.03.7 RESEARCH STRATEGIES AND RESULTS

The methodologies available for research into bias in mental tests have grown rapidly in number and sophistication since the 1970s. Extensive reviews of the questions to be addressed in such research and their corresponding methodologies are available in Berk (1982), Camilli and Shepard (1994), Jensen (1980), Reynolds (1982b, 1995), and Reynolds and Brown (1984). The most popular methods are reviewed below, along with a summary of findings from each area of inquiry (these summaries are based almost entirely on research with ethnic subcultures within the USA). The sections are organized primarily by methodology within each content area of research (i.e., research into content, construct, and predictive validity).

10.03.7.1 Bias in Content Validity

Bias in the item content of intelligence tests is one of the favorite topics of those who decry the use of standardized tests with minorities (e.g., Hilliard, 1979; Jackson, 1975; Williams, 1972; Wright & Isenstein, 1977). The earliest work in cultural test bias centered around content. Typically, critics review the items of a test and select specific items as being biased because: (i) the items ask for information that ethnic minority or disadvantaged persons have not had equal opportunity to learn; and/or (ii) the scoring of the items is improper, since the test author has arbitrarily decided on the only correct answer and ethnic minorities are inappropriately penalized for giving answers that would be correct in their own culture but not that of the test maker; and/or (iii) the wording of the questions is unfamiliar, and an ethnic minority person who may “know” the correct answer may not be able to respond because he or she does not understand the question. Each of these and related criticisms, when accurate, has the same basic empirical result: the item becomes relatively more difficult for ethnic minority group members than for the majority population, for example, an ethnic minority and a member of the majority culture with the same standing on the construct in question will respond differently to such biased items. This leads directly to a definition of content bias for aptitude tests that allows empirical assessment of the phenomenon. An item or subscale of a test is considered to be biased in content when it is demonstrated to be relatively more difficult for members of one group than for members of another when the general ability level of the groups being compared is held constant and no reasonable theoretical rationale exists to explain group differences on the item (or subscale) in question.

With regard to achievement tests, the issue of content bias is considerably more complex. Exposure to instruction, general ability level of the group, and the accuracy and specificity of the sampling of the domain of items are all important variables in determining whether the content of an achievement test is biased (see Schmidt, 1983). Research into item (or content) bias with achievement tests has typically, and perhaps mistakenly, relied on methodology appropriate for determining item bias in aptitude tests. Nevertheless, research examining both types of instruments for content bias has yielded quite comparable results. Items on personality tests may be perceived differently across cultures as well or appropriate responses may vary dramatically and quite properly deserve different interpretations cross-culturally. If so, the items will behave differently across groups for individuals with the same relative standing. This too is detectable through analyses of item response data across groups.

One method of locating "suspicious" test items requires that item difficulties be determined separately for each group under consideration. If any individual item or series of items appears to be exceptionally difficult for the members of any group, relative to other items in the test, the item is considered potentially biased and removed from the test. An early, widespread approach to identifying biased items involved analysis of variance (ANOVA) and several closely related procedures wherein the group X item interaction term is of interest (e.g., Angoff & Ford, 1973; Cardall & Coffman, 1964; Cleary & Hilton, 1968; Plake & Hoover, 1979; Potthoff, 1966; Stanley, 1969).

The definition of content bias set forth above actually requires that the differences between groups be the same for every item on the test. Thus, in the ANOVA procedure, the group X item interaction should not yield a significant result. Whenever the differences in items are not uniform (a significant group X item interaction does exist), one may contend that biased items exist. Earlier in this area of research, it was hoped that the empirical analysis of tests at the item level would result in the identification of a category of items having similar biased, and that such items could then be avoided in future test development (Flaugher, 1978). Very little similarity among items determined to be biased has been found. No one has been able to identify those characteristics of an item that cause the item to be biased. It does seem that poorly written, sloppy, and ambiguous items tend to be identified as biased with greater frequency than those items typically encountered in a well-constructed standardized instrument. The variable at issue then may be the item reliability. Item reliabilities are typically not large, and poorly written or ambiguous test items can easily have reliabilities approaching zero. Decreases in reliability are known to increase the probability of the occurrence of bias (Linn & Werts, 1971). Informal inventories and locally derived tests are much more likely to be biased than professionally written standardized tests that have been scrutinized for bias in the items and whose item characteristics are known.

Once items have been identified as biased under the procedures described above, attempts have been made to eliminate "test bias" by eliminating the offending items and rescoring the tests. As pointed out by Flaugher (1978) and Flaugher and Schrader (1978), however, little is gained by this tactic. Mean differences in performance between groups are affected only slightly, and aptitude and achievement tests become more difficult for everyone involved, since the eliminated items typically have moderate to low difficulty. When race X item interactions have been found, the interaction typically accounts for a very small proportion of variance. For example, in analyzing items on the WISC-R, Jensen (1976), Sandoval (1979), and Mille (1979) found the group X item interaction to account for only 2–5% of the variance in performance. Using a similar technique with the Wonderlic Personnel Test, Jensen (1977) found the race X item interaction to account for only about 5% of the test score variance. Thus, elimination of the offending items can be expected to have little, if any, significant effect. These analyses have been of a *post hoc* nature (i.e., after the tests have been standardized), however, and use of empirical methods for determining item bias during the test development phase (as with tests like the K-ABC and the TOMAL) is to be encouraged.

The ANOVA methodology is appealing conceptually but has some significant problems, even though it was the dominant methodological approach to the issue of item bias through the

1980s. Camilli and Shepard (1987) have provided convincing examples, albeit using contrived data, that ANOVA methods often miss biased items, in both directions, and identify some items as biased which are not. An algebraic demonstration of the reasons for this is provided in Camilli and Shepard (1994) who conclude that ANOVA should no longer be used.

Based upon their thorough and compelling analysis of methods for detecting biased items, Camilli and Shepard (1994) recommend methods derived from item response theory (IRT) to detect what has come to be known as differential item functioning (DIF). IRT models are conceptually similar to what other models such as ANOVA attempt. IRT is concerned principally with the probability of a particular response to a test item as a function of the examinee's relative position on the latent trait assessed by the scale to which the item belongs. IRT models to detect DIF are primarily superior to prior methods because they are less sample dependent and they allow one to estimate multiple item statistics more precisely than a technique such as ANOVA. Using item characteristic curves, DIF is more accurately and readily detected when the probability of a particular response changes as a function of some nominal variable (e.g., ethnicity or gender) for individuals with the same relative standing on the latent trait being assessed.

Early studies using other approaches for the detection of DIF using a partial correlation procedure developed independently by Stricker (1982) and Reynolds, Willson, and Chatman (1984) have found no systematic bias against American Blacks or against women on measures of English vocabulary. Willson, Nolan, Reynolds, and Kamphaus (1989), using the same partialling methodology, examined DIF of the mental processing scales of the Kaufman Assessment Battery for Children, concluding " ... there appears to be little evidence of systematic race or gender bias ... " (p. 289).

With multiple-choice tests, another level of complexity is added to the examination of content bias. With a multiple-choice question, three or four distractors are typically given in addition to the correct response. Distractors may be examined for their attractiveness (the relative frequency with which they are chosen) across groups. When distractors are found to be disproportionately attractive for members of any particular group, the item may be defined as biased. When items are constructed to have an equal distribution of responses to each distractor for the total test population, then chi-square can be used to examine the distribution of choices for each distractor for each group (e.g., Burrill, 1975).

Jensen (1976) investigated the distribution of wrong responses for two multiple-choice intelligence tests, the Peabody Picture Vocabulary Test (PPVT) and Raven's Progressive Matrices (the Raven). Each of these tests was individually administered to 600 White and 400 Black children between the ages of six and 12. The analysis of incorrect responses for the PPVT indicated that the errors were distributed in a nonrandom fashion over the distractors for a large number of items. However, no racial bias in response patterns occurred, since the disproportionate choice of distractors followed the same pattern for Blacks and Whites. On the Raven, Blacks make different types of errors than Whites, but only on a small number of items. Jensen followed up these items and compared the Black response pattern to the response pattern of White children at a variety of age levels. For every item showing differences in Black–White response patterns, the Black response could be duplicated by the response patterns of Whites approximately two years younger than Blacks.

Veale and Foreman (1983) have advocated inspecting multiple-choice tests for bias in distractor or "foil" response distributions as a means of refining tests *before* they are finalized for the marketplace. They note that there are many instances whereby unbiased external criteria (such as achievement or ability) or culturally valid tests are not readily accessible for detecting bias in the measure under study. Veale and Foreman add that inspection of incorrect responses to distractor items can often lead to greater insight concerning cultural bias in any given question than would inspection of percentage of correct responses across groups. Veale and Foreman (1983) provide the statistical analyses for their "overpull probability model" along with the procedures for measuring cultural variation and diagraming the source of bias within any given item.

Investigation of item potential sources of bias during test development is certainly not restricted to multiple-choice items and methods such as those outlined by Veale and Foreman (1983). The possibilities are numerous (see Camilli & Shepard, 1994; Jensen, 1980, chap. 9). For example, Scheuneman (1987) has used the results of linear methods on Graduate Record Examination (GRE) item data to show interesting influences on Black–White performance when specific item characteristics (e.g., vocabulary content, one true or one false answer to be selected, diagrams to be used or not used, use of antonym items, etc.) are uniformly investigated. Although Scheuneman indicates that future research of this type should reduce the number of variables to address (there are 16 hypotheses therein), the results nonetheless suggest that bias methodology is a viable way in which to determine whether differential effects can "be demonstrated through the manipulation of relatively stable characteristics of test items" (p. 116). Scheuneman presented pairs of items, with the designated characteristic of a question format under study present in one item and absent or modified in the other. Paired experimental items were administered in the experimental section of the GRE General Test, given in December 1982. Results indicated that certain "item elements'—common in general form to a variety of questions—appeared to have a differential impact on Black and White performance. For example, significant group version interactions were seen for one correct true vs. one correct false response and for adding/modifying prefixes/suffixes to the stimulus word in antonym items. The question is thus raised as to whether the items showing differential impact are measuring the content domain (e.g., veral, quantitative, or analytial thinking) as opposed to an aspect of "element" within the presentation to some degree.

Another approach to the identification of biased items has been pursued by Jensen (1976). According to Jensen, if a test contains items that are disproportionately difficult for one group of examinees as compared to another, the correlation of *P* decrements between adjacent items will be low for the two groups. ("*P* decrement" refers to the difference in the difficulty index, *P*, from one item of a test to the next least or most item. Typically, ability test items are arranged in ascending order of difficulty.) Jensen (1974, 1976) also contends that if a test contains biased items, the correlation between the rank order of item difficulties for one race with another will also be low. Jensen (1974, 1976, 1977) calculated cross-racial correlations of item difficulties for large samples of Black and White children on five major intelligence tests: the PPVT, the Raven, the Revised Stanford–Binet Intelligence Scale Form L-M, the WISC-R, and the Wonderlic Personnel Test. Cross-racial correlations of *P* decrements were reported for several of the scales. Jensen's results are summarized in Table 1, along with the results of several other investigators also employing Jensen's methodology.

As is readily apparent in Table 1, little evidence to support any consistent content bias within any of the scales investigated was found. The consistently large magnitude of the cross-racial correlations of *P* decrements is impressive and indicates a general lack of content bias in the instuments as a whole. As previously noted, however, some individual items were identified as biased, yet they collectively accounted for only 2–5% of the variance in performance differences and showed no detectable pattern in item content.

This method has proved popular with some test publishers who desire to look at the items on a test as a group, despite the fact this approach may be overly sensitive due to the instability of estimates of *P* values (which would cause the correlation to be spuriously low). Using the most recent version of the Detroit Tests of Learning Aptitude (DTLA-3; Hammill, 1991), Hammill reported correlations of *P* decrements exceeding 0.90 for all subtests with most exceeding 0.95. Similar results have been reported for other aptitude measures. On the 14 subtests of the Test of Memory and Learning (TOMAL; Reynolds & Bigler, 1994), Reynolds and Bigler report correlations across *P* decrements by gender and ethnicity that all exceed 0.90 with most again above 0.95.

Another approach to this question is to use the partial correlation between a demographic or other nominal variable and item score, where the correlation between total test score and the variable of interest has been removed from the relationship. If a significant partial correlation exists, say, between race and an item score after the race–total test score relationship has been partialed, then the item is performing differentially across race within ability level. Bias has been demonstrated at this point under the definition offered above. Use of the partial correlation (typically a partial point-biserial *r*) is a simple yet powerful item bias detection approach, but its development is relatively recent and its use not yet common. The partial correlation approach also does not have the problems of ANOVA methods. An example of its application may be found in Reynolds, Willson, and Chatman (1984).

A common practice has been a return to including expert judgment by professionals and members of minority groups in the item selection for new psychological and educational tests. This approach was used in development of the K-ABC, the revision of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-R), the PPVT-R, PPVT-III, and a number of other contemporary tests. The practice typically asks for an "armchair" inspection of individual items as a means of locating and purging biased items in the measure under development. Since, as previously noted, no detectable pattern or common characteristic of content of individual items statistically shown to be biased has been observed (given reasonable care in the item-writing stage), it seems reasonable to question the "armchair" approach to determining biased

Table 1 Cross-racial analysis of content bias for five major intelligence scales.

Scale	*Cross-racial correlation of rank order of item difficulties*[a] *Black–White correlations*[b]	 *White–Mexican–American correlations*[b]
Peabody Picture Vocabulary Test (Jensen, 1974)	0.99 (0.79), 0.98 (0.65)	0.98 (0.78), 0.98 (0.66)
Raven's Progressive Matrices (Jensen, 1974)	0.99 (0.98), 0.99 (0.96)	0.99 (0.99), 0.99 (0.97)
Stanford-Binet Intelligence Scale (Jensen, 1976)	0.96	
Wechsler Intelligence Scale for Children-Revised		
(Jensen, 1976)	0.95	
(Sandoval, 1979)[c]	0.98 (0.87)	0.99 (0.91)
(Mille, 1979) (1949 WISC)	0.96, 0.95	
Wonderlic Personnel Test (Jensen, 1977)	0.94 (0.81)	

[a] Correlation of P decrements across race is included in parentheses if reported. [b] Where two sets of correlations are presented, data were reported separately for males and females and are listed males first. The presence of a single correlation indicates that data were pooled across gender. [c] Median values for the 10 WISC-R subtests excluding Digit Span and Coding.

items. The bulk of scientific data since the pioneering work of McGurk (1951) has not supported the position that anyone can—upon surface inspection—detect the degree to which any given item will function differentially across groups (Shepard, 1982). Several researchers since McGurk's time have identified items as being disproportionately more difficult for minority group members than for members of the majority culture and have subsequently compared their results with a panel of expert judges. The data have provided some interesting results.

Although examples of the failure of judges to identify biased items now abound (Camilli & Shepard, 1994) and show that judges are right about an item about as often as they are wrong, two studies demonstrate this failure most clearly. After identifying the eight least racially discriminating items on the Wonderlic Personnel Test, Jensen (1976) asked panels of five Black psychologists and five White psychologists to sort out the eight most and eight least discriminating items when only these 16 items were presented to them. The judges sorted the items at a level no better than chance. Sandoval and Mille (1980) conducted a somewhat more extensive analysis, using items from the WISC-R. These two researchers had 38 Black, 22 Mexican-American, and 40 White university students from Spanish, history, and education classes identify items from the WISC-R that would be more difficult for a minority child than a White child and items that would be equally difficult for each group. A total of 45 WISC-R items were presented to each judge; these items included the 15 most difficult items for Blacks as compared to Whites, the 15 most difficult items for Mexican-Americans as compared to Whites, and the 15 items showing the most nearly identical difficulty indices for minority and White children. The judges were asked to read each question and determine whether they thought the item was: (i) easier for minority than for White children, (ii) easier for White than for minority children, or (iii) of equal difficulty for White and minority children. Sandoval and Mille's (1980) results indicated that the judges were not able to differentiate accurately between items that were more difficult for minorities and items that were of equal difficulty across groups. The effects of the judges' ethnic background on the accuracy of item bias judgments were also considered. Minority and nonminority judges did not differ in their ability to identify accurately biased items, nor did they differ with regard to the type of incorrect identification they tended to make. Sandoval and Mille's (1980) two major conclusions were that "(1) judges are not able to detect items which are more difficult for a minority child than an Anglo child, and (2) the ethnic background of the judge makes no difference in accuracy of item selection for minority children" (p. 6). In each of these studies, the most extreme items were used, which should have given the judges an advantage.

Anecdotal evidence is also available to refute the assumption that armchair analyses of test bias in item content are accurate. The most widely cited example of a biased intelligence test item is item 6 of the WISC-R Comprehension subtest: "What is the thing to do if a boy (girl) much smaller than yourself starts to fight with you?" This item is generally considered to be biased against US Black children in particular, because of the scoring criteria. According to the item's critics, the most logical response for a Black child is to "fight back," yet this is a 0-point response. The correct (2-point) response is to walk away and avoid fighting with the child—a response that critics claim invites disaster in the Black culture, where children are taught to fight back and would not "know" the "correct White response." Black responses to this item have been investigated empirically in several studies, with the same basic results: the item is relatively easier for Black children than for White children. When all items on the WISC-R are ranked separately according to difficulty level for Blacks and Whites, this item is the 42nd least difficult item (where 1 represents the easiest item) for Black children and the 47th least difficult for White children (Jensen, 1976). Mille (1979), in a large *N* study of bias, reached a similar conclusion, stating that this item "is relatively easier for Blacks than it is for Whites" (p. 163). The results of these empirical studies with large samples of Black and White children in the USA are unequivocal: when matched for overall general intellectual skill, more Black than White children will get this item correct—the very item most often singled out as a blatant example of the inherent bias of intelligence test against Blacks, but so selected using anecdotal impressions and sterotypical views of Black cultural even by Black judges (see also Reynolds & Brown, 1984).

Even without empirical support for its accuracy, a number of prestigious writers support the continued use of the "face validity" approach of using a panel of minority judges to identify "biased" test items (Anastasi, 1986; Kaufman, 1979; Sandoval & Mille, 1979). Those who support the continued use of this technique see it as a method of gaining greater rapport with the public. As pointed out by Sandoval and Mille (1979), "Public opinion, whether it is supported by empirical findings, or based on emotion, can serve as an obstacle to the use of a measurement instrument" (p. 7). The elimination of items that are offensive or otherwise objectionable to any substantive segment of the population for whom the test is intended seems an appropriate action that may aid in the public's acceptance of new and better psychological asessment tools. However, the subjective-judgment approach should not be allowed to supplant the use of more sophisticated analyses in the determination of biased items. The subjective approach should serve as a supplemental procedure, and items identified through this method (provided that some inter-rater agreement can be obtained—an aspect of the subjective method yet to be demonstrated) as objectionable can be eliminated when a psychometrically equivalent (or better) item can be obtained as a replacement and the intent of the item is kept intact (e.g., with a criterion-referenced measure, the new item must be designed to measure the same objective).

Researchers such as Tittle (1982) have stressed that the possibility of and need for cooperation between those advocating statistical validity and those advocating face validity in nonbiased test construction is greater than one might think, given the above-cited research. Judgmental analysis allows for the perception of fairness in items, tests, and evaluations, and this perception should not be taken lightly. Tittle (1982) argues that "judgmental methods arise from a different, nonstatistical ground. In examining fairness or bias primarily on statistical grounds, we may again be witnessing a technical solution to a problem that is broader than the technical issues" (p. 34). Cronbach (1980) does not find the issue of fairness as determined by subjective judgment to be outside the realm of test validation. Cronbach states, "The politicization of testing ought not [to] be surprising. Test data influence the fortunes of individuals and the support given to human service programs" (p. 100). Tittle (1975, 1982) argues that the general field of test development requires greater consensus regarding specific, multidimensional steps taken in formulating "fair" measures, because "fairness" in testing will never be realistically viewed by the public from a unidimensional statistical standpoint.

Considerably less work has been conducted in all areas of bias relative to personality testing, where there would appear to be greater opportunity for cultural, social, and ethnic factors to act to produce bias. Research on item bias of personality measures, though less extensive than with aptitude measures, has produced results similar to those with aptitude measures (see especially Moran, 1990; Reynolds, in press; Reynolds & Harding, 1983).

Research evaluating behavior rating scales is also meager but at present supports the use of parent ratings in the diagnosis of childhood psychopathology independent of the child's ethnic background (e.g., see Mayfield & Reynolds, in press). A common set of items seems to measure consistently a variety of personality and behavioral traits for Whites, Blacks, and various Hispanic and Latin population residing in the USA (James, 1995; Mayfield & Reynolds, in press; Reynolds & Kamphaus, 1992).

Thus far, this section has focused on the identification of biased items. Several studies evaluating other hypotheses have provided data that are relevant to the issue of content bias of psychological tests, specifically the WISC-R (although now largely superseded in practice by WISC-III, little data regarding bias are available specifically on this new scale).

Jensen and Figueroa (1975) investigated Black–White differences in mental test scores as a function of differences in Level I (rote learning and memory) and Level II (complex cognitive processing) abilities. These researchers tested a large number of Blacks and Whites on the WISC-R Digit Span subtest and then analyzed the data separately for digits forward and digits backward. The content of the digits forward and digits backward procedures is the same. Thus, if score differences are due only to bias in content of the item validity, score differences across race should remain constant for the two tasks. On the other hand, since the information-processing demands of the two tasks are quite different (Reynolds, 1997b), the relative level of performance on the two tasks should not be the same, as Blacks and Whites differ in their ability to process information according to the demands of the two tasks. Jensen and Figueroa (1975) found the latter to be the case. The Black–White score difference on digits backward was more than twice the magnitude of the difference for digits forward. Granted, this methodology can provide only indirect evidence regarding the content validity of an instrument; however, its importance is in providing a different view of the issues and an alternative research strategy. Since the Jensen and Figueroa results do not indicate any content bias in the Digit Span subtest, they add to a growing body of literature that strongly suggests the lack of cultural bias in well-constructed, standardized tests, and their generalization within a common language and with some cultural diversity, at least for Black and for White US cultures.

Another study (Reynolds & Jensen, 1983) examined each of the 12 WISC-R subtests for cultural bias against Blacks using a variation of the group X item ANOVA methodology discussed earlier. Reynolds and Jensen matched 270 Black children with 270 White children from the WISC-R standardization sample on the basis of gender and WISC-R Full Scale IQ. Matching the two groups of children on the bases of the Full Scale IQ essentially equated the two groups for *g*. Therefore, examining Black–White differences in performance on each subtest of the WISC-R made it possible to determine which, if any, of the subtests were disproportionately difficult for Blacks or Whites. A summary of the Reynolds and Jensen (1983) results is presented in Table 2. Blacks exceeded Whites in performance on two subtests: Digit Span and Coding. Whites exceeded Blacks in performance on three subtests: Comprehension, Object Assembly, and Mazes. A trend was apparent for Blacks to perform at a higher level on the Arithmetic subtest, while Whites tended to exceed Blacks on the Picture Arrangement subtest. Although these results can be interpreted to indicate bias in several of the WISC-R subtests, the actual differences were very small (typically of the order of 0.10–0.15 standard deviation), and the amount of variance in performance associated with ethnic group membership was less than 5% in each case. The results are also reasonably consistent with Jensen's theory on mental test score differences and their relationship to Level I and Level II abilities. The Digit Span and Coding subtests are clearly the best measures of Level I abilities on the WISC-R, while Comprehension, Object Assembly, and Mazes are more closely associated with Level II abilities. Digit Span and Coding are also tasks adversely affected by increases in an examinee's anxiety level. The relatively higher level of performance by the Black children on these tasks is also inconsistent with arguments that the testing circumstance is more unfamiliar and more anxiety provoking to minority children relative to their White counterparts.

From a large number of studies employing a wide range of methodologies, a relatively clear picture emerges: content bias in well-prepared standardized tests is irregular in its occurrence, and no common characteristics of items that are found to be biased can be ascertained by expert judges (minority or nonminority). The variance in group score differences on mental tests associated with ethnic group membership when content bias has been found is relatively small (typically ranging from 2% to 5%). Even this small amount of bias has been seriously questioned, as Hunter (1975) describes such findings basically as methodological artifacts.

Harrington (1975, 1976) argued that traditional statistical methods would not detect DIF and that tests are designed in a psychometrically flawed manner so as to always favor the numerically superior group in a population. Although Harrington's research and logic are impressive, his results have not held and the Harrington Effect is not seen in a variety of works (e.g., Bayley, 1969; Beauchamp, Sammuels, & Griffore, 1974; Hickman & Reynolds, 1986; James, 1995).

Although the search for common "biased" item characteristics will continue, and psychologists must pursue the public relations issues of face validity, "armchair" claims of cultural bias in aptitude tests have found no empirical support in a large number of actuarial studies contrasting the performance of a variety of racial groups on items and subscales of the most

Table 2 Means, standard deviations, and univariate *F*s for comparison of performance on specific WISC-R subtests by groups of Blacks and Whites matched for WISC-R full scale IQ.

	Blacks		*Whites*				
WISC-R variable	$\bar{X}$	*SD*	$\bar{X}$	*SD*	D^a	F^b	*p*
Information	8.40	2.53	8.24	2.62	−0.16	0.54	NS
Similarities	8.24	2.78	8.13	2.78	−0.11	0.22	NS
Arithmetic	8.98	2.62	8.62	2.58	−0.36	2.52	0.10
Vocabulary	8.21	2.61	8.27	2.58	+0.06	0.06	NS
Comprehension	8.14	2.40	8.58	2.47	+0.44	4.27	0.05
Digit Span	9.51	3.09	8.89	2.83	+0.62	6.03	0.01
Picture Completion	8.49	2.88	8.60	2.58	+0.11	0.18	NS
Picture Arrangement	8.45	2.92	8.79	2.89	+0.34	1.78	0.01
Block Design	8.06	2.54	8.33	2.76	+0.27	1.36	NS
Object Assembly	8.17	2.90	8.68	2.70	+0.51	4.41	0.05
Coding	9.14	2.81	8.65	2.80	−0.49	4.30	0.05
Mazes	8.69	3.14	9.19	2.98	+0.50	3.60	0.05
Verbal IQ	89.63	12.13	89.61	12.07	−0.02	0.04	NS
Performance IQ	89.29	12.22	90.16	11.67	+0.87	0.72	NS
Full Scale IQ	88.61	11.48	88.96	11.35	+0.35	0.13	NS

NS, not significant.
[a] White $\bar{X}$–Black $\bar{X}$ difference. [b] Degrees of freedom = 1538.

widely employed intelligence and achievement scales in the USA; neither differential for single-group validity has been demonstrated. However, these results apply only to groups where a common language exists among examinees. Personality and interest measures fare likewise but there is less good research available for such tests. Test translations are a different issue and required a different line of study altogether.

10.03.7.2 Bias in Construct Validity

There is no single method for the accurate determination of the construct validity of educational and psychological tests. Defining bias in construct validity thus requires a general statement that can be researched from a variety of viewpoints with a broad range of methodologies. The following rather parsimonious definition is proffered: bias exists in regard to construct validity when a test is shown to measure different hypothetical traits (psychological constructs) for one group than for another; that is differing interpretations of a common performance are shown to be appropriate as a function of ethnicity, gender, or another variable of interest, one typically but not necessarily nominal.

As befits the concept of construct validity, many different methods have been employed to examine existing tests for potential bias in construct validity. One of the most popular and necessary empirical approaches to investigating construct validity is factor analysis (Anastasi, 1982; Cronbach, 1970). Factor analysis as a procedure identifies clusters of test items or clusters of subtests of psychological or educational tests that correlate highly with one another, and less so or not at all with other subtests or items. It thus allows one to determine patterns of interrelationships of performance among groups of individuals. For example, if several subtests of an intelligence scale load highly on (are members of) the same factor, then if a group of individuals score high on one of these subtests, they would be expected to score at a high level on other subtests that load highly on that factor. Psychologists attempt to determine, through a review of the test content and correlates of performance on the factor in question, what psychological trait underlies performance; or, in a more hypothesis-testing approach, they will make predictions concerning the pattern of factor loadings. Dana (1993) notes that factor analysis across various cultural groups can be a useful means of examining the cross-cultural validity of tests of personality, behavior, or cognitive skill. Dana argues that when "... the factor dimensions resulting from different factor analytic methods are stable, and therefore present an invariant structure across cultures, then cross-cultural validity may be inferred" (1993, p. 101). Reynolds (1982a, 1982b) has made similar albeit more detailed arguments as have others in specific domains such as intelligence. Hilliard (1979), one of the more vocal critics of IQ tests on the basis of cultural bias, has pointed out one of the potential areas of bias in comparisons of the factor-analytic results of tests across races:

> If the IQ test is a valid and reliable test of "innate" ability or abilities, then the factors which emerge on a given test should be the same from one population to another, since "intelligence" is asserted to be a set of mental processes. Therefore, while the configuration of scores of a particular group on the factor profile would be expected to differ, logic would dictate that the factors themselves would remain the same. (p. 53)

Although researchers do not necessarily agree that identical factor analytic results across groups speak to the innateness of the traits being measured, consistent factor-analytic results across populations do provide strong evidence that whatever is being measured by the instrument is being measured in the same manner and is, in fact, the same latent construct within each group. The information derived from comparative factor analysis across populations is directly relevant to the use of educational and psychological tests in diagnosis and other decision-making functions. Psychologists, in order to make consistent interpretations of test score data, must be certain that a test measures the same variable across populations.

Two basic approaches, each with a number of variations have been employed to compare factor-analytic results across populations. The first and more popular approach asks how similar the results are for each group; the second and less popular approach asks whether the results show a statistically significant difference between groups. The most sophisticated approach to the latter question has been the foundational work of Jöreskog (1969, 1971) in simultaneous factor analysis in several populations and now basically represented in the LISREL series of computer programs. However, little has been done with the latter approach within the context of test bias research. Mille (1979) has demonstrated the use of a simpler method (actually developed by Jensen and presented in detail in Jensen, 1980) for testing the significance of the difference between factors for two populations.

A number of techniques have been developed to measure the similarity of factors across groups. The two most common methods of determining factorial similarity or factorial invariance involve the direct comparison of factor loadings across groups. The two primary techniques for this comparison are (i) the calculation of a coefficient of congruence (Harman, 1976) between the loadings of corresponding factors for two groups, and (ii) the simple calculation of a Pearson product-moment coefficient of correlation between the factor loadings of the corresponding factors. The latter technique, though used with some frequency, is less satisfactory than the use of the coefficient of congruence, since in the comparison of factor loadings certain of the assumptions underlying the Pearson r may be violated. When one is determining the degree of similarity of factors, a value of 0.90 or greater is typically, though arbitrarily, taken to indicate equivalent factors (factorial invariance). However, the most popular methods of calculating factorial similarity produce quite similar results (Reynolds & Harding, 1983), at least in large n studies.

In contrast to Hilliard's (1979) strong statement that studies of factorial similarity across race have not been reported in the technical literature on IQ, a number of such studies have appeared dealing with a number of different tests. The focus here is primarily on studies comparing factor-analytic results across ethnic groups in the USA for aptitude tests, since these are the most controversial of tests.

Because the WISC (Wechsler, 1949) and its successor, the WISC-R (Wechsler, 1974; now superseded by the WISC-III, but as yet no such bias research is available on the latter scale), have been the most widely employed individual intelligence tests with school-age children in much of the world, it is appropriate that the cross-race structure of these two instruments has received extensive investigation for both normal and referral populations of children. Using a large, random sample, Reschly (1978) compared the factor structure of the WISC-R across four racially identifiable groups: Whites, Blacks, Mexican-Americans, and native American Papagos, all from the southwestern USA. Consistent with the findings of previous researchers with the 1949 WISC (Lindsey, 1967; Silverstein, 1973), Reschly (1978) reported substantial congruency of factors across races when the two-factor solutions were compared (the two-factor solution typically reiterated Wechsler's *a priori* grouping of the subtests on to a Verbal and a Performance, or nonverbal, scale). The 12 coefficients of congruence for comparisons of the two-factor solution across all combinations of racial groupings ranged only from 0.97 to 0.99, denoting factorial equivalence of this solution across groups. Reschly also compared three-factor solutions (three-factor solutions typically include Verbal Comprehension, Perceptual Organization, and Freedom from Distractibility factors), finding congruence only between Whites and Mexican-Americans. These findings are also consistent with previous research with the WISC (Mille, 1979; Semler & Iscoe, 1966), an intelligence scale originally developed and normed on an all-White sample.

The *g* factor present in the WISC-R was shown to be congruent across race, as was also demonstrated by Mille (1979) for the WISC. Reschly (1978) concluded that the usual interpretation of the WISC-R Full Scale IQ as a measure of overall intellectual ability appears to be equally appropriate for Whites, Blacks, Mexican-Americans, and Native American Papagos. Jensen (1985) has presented compelling data indicating that the Black-White discrepancy seen in major tests of aptitude reflects primarily the g factor. Reschly also concluded that the Verbal-Performance scale distinction on the WISC-R is equally appropriate across race and that there is strong evidence for the integrity of the WISC-R's construct validity as a measure of intelligence for a variety of populations.

Support for Reschly's (1978) conclusions is available from a variety of other studies of the WISC and WISC-R. Applying a hierarchical factor-analytic method developed by Wherry and Wherry (1969), Vance and Wallbrown (1978) factor analyzed the intercorrelation matrix of the WISC-R subtests for 150 referred Blacks from the Appalachian region of the USA. The two-factor hierarchical solution determined for Vance and Wallbrown's (1978) Blacks was highly similar to hierarchical factor solutions determined for the standardization samples of the Wechsler scales generally (Blaha, Wallbrown, & Wherry, 1975; Wallbrown, Blaha, & Wherry, 1973). Vance and Wallbrown's (1978) results with the WISC-R are also consistent with a previous hierarchical factor analysis with the 1949 WISC for a group of disadvantaged Blacks and Whites (Vance, Huelsman, & Wherry, 1976).

Several studies comparing the WISC-R factor structure across races for normal and referral populations of children have also provided increased support for the generality of Reschly's (1978) conclusions and the results of the other investigators cited above. Oakland and Feigenbaum (1979) factor analyzed the 12 WISC-R subtests' intercorrelations separately for stratified (race, age, sex, SES) random samples of normal White, Black, and Mexican-American children from an urban school district of the northwestern USA. Pearson *r*'s were calculated between corresponding factors for each group. For the *g* factor, the Black–White correlation between factor loadings was 0.95, the Mexican-American White correlation was 0.97, and the Black Mexican-American correlation was 0.96. Similar comparisons across all WISC-R variables produced correlations ranging only from 0.94 to 0.99. Oakland and Feigenbaum concluded that the results of their factor analyses "do not reflect bias with respect to construct validity for these three racial-ethnic ... groups" (1979, p.973).

Gutkin and Reynolds (1981) determined the factorial similarity of the WISC-R for groups of Black and White children from the WISC-R standardization sample. This study is particularly important to examine in determining the construct validity of the WISC-R across races, because of the sample employed in the investigation. The sample included 1868 White and 305 Black children obtained in a stratified random sampling procedure designed to mimic the 1970 US census data on the basis of age, sex, race, SES, geographic region of residence, and community size. Similarity of the WISC-R factor structure across race was investigated by comparing the Black and White groups for the two- and three-factor solutions on (i) the magnitude of unique variances, (ii) the pattern of subtest loadings on each factor, (iii) the portion of total variance accounted for by common factor variance, and (iv) the percentage of common factor variance accounted for by each factor. Coefficients of congruence comparing the unique variances, the *g* factor, the two-factor solutions, and the three-factor solutions across races all achieved a value of 0.99. The portion of total variance accounted for by each factor was the same in both the two- and three-factor racial groups. Gutkin and Reynolds (1981) concluded that for White and Black children the WISC-R factor structure was essentially invariant, and that no evidence of single-group or differential construct validity could be found. Subsequent studies comparing the WISC-R factor structure for referral populations of White and Mexican-American children have also strongly supported the construct validity of the WISC-R across races (e.g., Dean, 1979b; Gutkin & Reynolds, 1980).

DeFries et al. (1974) administered 15 mental tests to large samples of Americans of Japanese and Chinese ancestry. After examining the pattern of intercorrelations among the 15 tests for each of these two ethnic groups, DeFries et al. concluded that the cognitive organization of the two groups was virtually identical. In reviewing this study, Willerman (1979) concluded that "The similarity in factorial structure [between the two groups] suggests that the manner in which the tests are constructed by the subjects is similar regardless of ethnicity and that the tests are measuring the same mental abilities in the two groups" (p. 468). At the adult level, Kaiser (1986) and Scholwinski (1985) have analyzed the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) and reported substantial similarity between factor structures for Black and White samples obtained from the WAIS-R standardization data.

At the preschool level, factor-analytic results also tend to show consistency of construct validity across races, though the results are less clear-cut. In a comparison of separate factor analyses of the McCarthy Scales of Children's Abilities (McCarthy, 1972) for groups of Black and White children, Kaufman and DiCuio (1975) concluded that the McCarthy Scales showed a high degree of factorial similarity between the two races. The conclusion was not straightforward, however. Four factors were found for the Blacks and three for the Whites. Kaufman and DiCuio based their conclusion on factorial similarity on the finding that each "White" factor had a coefficient of congruence of 0.85–0.93 with one "Black" factor. One Black factor on the McCarthy Scales had no White counterpart with a coefficient of congruence beyond 0.74 (the Memory factor), and the Black and White Motor factors showed a coefficient of congruence of only 0.85.

When investigating the factor structure of the WPPSI across race, Kaufman and Hollenbeck (1974) found much "cleaner" factors for Blacks and Whites than with the McCarthy Scales. The two factors, essentially mirroring Wechsler's Verbal and Performance scales, were virtually identical between the races. Both factors also appear closely related to the hierarchical factor solution presented by Wallbrown at al. (1973) for Blacks and Whites on the WPPSI. When comparing factor analyses of the Goodenough–Harris Human Figure Drawing Test scoring item, Merz (1970) found highly similar factor structures for Blacks, Whites, Mexican-Americans, and native Americans.

Other investigators have found differences across races in the factor structures of several tests designed for preschool and primary-grade children. Goolsby and Frary (1970) factor-analyzed the Metropolitan Readiness Test (MRT) together for separate groups of Blacks and Whites, finding differences in the factor structure of this grouping of tests across races. When evaluating the experimental edition of the Illinois Test of Psycholinguistic Abilities, Leventhal and Stedman (1970) noted differences in the factor structure of this battery for Blacks and Whites. Two more studies have clarified somewhat the issue of differential construct validity of preschool tests across race.

The MRT (Hildreth, Griffith, & McGauvran, 1969) is one of the most widely employed of all preschool screening measures in the USA, and its 1969 version is composed of six subtests: Word Meaning, Listening, Matching, Letter Naming, Numbers, and Copying. Reynolds (1979b) had previously shown this to be essentially a one-factor (general Readiness) instrument. In a subsequent study, Reynolds (1979c) compared the general factor making up the MRT across races (Blacks and Whites) and genders. Substantial congruence was noted: coefficients of congruence across each pair of race–sex groupings ranged only from 0.92 to 0.99, with the lowest coefficient derived from the intraracial comparison for White females and White males. Eigenvalues, and subsequently the proportion of variance accounted for by the factor, were also highly similar for the race–sex groupings. The lack of differential or single-group construct validity across sex has also been demonstrated with aptitude tests for school-age children (Reynolds & Gutkin, 1980c).

In a more comprehensive study employing seven major preschool tests (the McCarthy Draw-a-Design and Draw-a-Child subtests, the Lee-Clark Reading Readiness tests, The Tests of Basic Experiences Language and Mathematics subtests, the Preschool Inventory-Revised Edition, and the MRT), Reynolds (1980a) reached a similar conclusion. A two-factor solution was determined with this battery for each of the four race–sex groups as above. Coefficients of congruence ranged only from 0.95 to 0.99 for the two factors, and the average degree of intercorrelation was essentially the same for all groups, as were eigenvalues and the percentage of variance accounted for by the factors. Reynolds (1980a) again concluded that the abilities being measured were invariant across race and that there was no evidence of differential or single-group construct validity of preschool tests across races or genders. The clear trend in studies of preschool tests' construct validity across race (and sex) is to uphold validity across groups. Such findings add support to the use of existing preschool screening measures with Black and White children of both sexes in the very necessary process of early identification (Reynolds, 1979a) of potential learning and behavior problems.

Taken individually but especially as a whole, these various cross-group factor analytic studies contradict sharply Helms' (1992) assertion that Black and White groups within the USA have significantly different latent cognitive structures. These results argue in favor of a common human organization of abilities and neuropsychological processes among the many ethnic groups studied thus far. Although majority–minority culturally related differences in level of performance differ in various ways across many ethnic groups, a common organization of latent ability structures seems clearly evident at this time.

As is appropriate for studies of construct validity, comparative factor analysis has not been the only method of determining whether single-group or differential validity exists. Another method of investigation involves comparing internal-consistency reliability estimates across groups. Internal-consistency reliability is determined by the degree to which the items are all measuring a similar construct. To be unbiased with regard to construct validity, internal-consistency estimates should be approximately equal across races. This characteristic of tests has been investigated with Blacks, Whites, and Mexican-Americans for a number of popular aptitude tests within the US population.

With groups of Black and White adults, Jensen (1977) calculated internal-consistency estimates (using the Kuder–Richardson 21 formula) for the Wonderlic Personnel Test (a frequently used employment/aptitude test). Kuder–Richardson 21 values of 0.86 and 0.88 were found, respectively, for Blacks and Whites. Using Hoyt's formula, Jensen (1974) determined internal-consistency estimates of 0.96 on the PPVT for each of three groups of children: Blacks, Whites, and Mexican-Americans. When children were categorized by gender within each racial grouping, the values ranged only from 0.95 to 0.97. On Raven's Progressive Matrices (colored), internal-consistency estimates were also quite similar across race and sex, ranging only from 0.86 to 0.91 for the six race–sex groupings. Thus, Jensen's (1974, 1977) research with three popular aptitude tests shows no signs of differential or single-group validity with regard to homogeneity of test content or consistency of measurement across groups.

Sandoval (1979) and Oakland and Feigenbaum (1979) have extensively investigated internal consistency of the various WISC-R subtests (excluding Digit Span and Coding, for which internal-consistency analysis is inappropriate) for Whites, Blacks, and Mexican-Americans. Both of these studies included large samples of children, with Sandoval's (1979) including over 1000. Sandoval found internal-consistency estimates to be within 0.04 of one another for all subtests except Object Assembly. This subtest was most reliable for Blacks (0.95), while being about equally reliable for Whites (0.79) and Mexican-Americans (0.75). Oakland and Feigenbaum (1979) reported internal-consistency estimates that never differed by more than 0.06 among the three groups, again with the exception of Object Assembly. In this instance, Object Assembly was most reliable for Whites (0.76), with about equal reliabilities for Blacks (0.64) and Mexican-Americans (0.67). Oakland and Feigenbaum also compared reliabilities across sex, finding highly similar values for males and females. Dean (1977) examined the internal consistency of the WISC-R for Mexican-American children tested by White examiners. He reported internal-consistency reliability estimates consistent with, although slightly exceeding, values reported by Wechsler (1974) for the predominantly White standardization sample. The Bender–Gestalt Test has also been reported to have similar internal-consistency estimates for Whites (0.84), Blacks (0.81), and Mexican-Americans (0.72), and for males (0.81) and females (0.80) (Oakland & Feigenbaum, 1979).

Several other methods have also been used to determine the construct validity of popular psychometric instruments across races. Since intelligence is considered a developmental phenomenon, the correlation of raw scores with age has been viewed as one measure of constrict validity for intelligence tests. Jensen (1976) reported that the correlations between raw scores on the PPVT and age were 0.79 for Whites, 0.73 for Blacks, and 0.67 for Mexican-Americans. For Raven's Progressive Matrices (colored), correlations for raw scores with age were 0.72 for Whites, 0.66 for Blacks, and 0.70 for Mexican-Americans. Similar results are apparent for the K-ABC (Kamphaus & Reynolds, 1987) and the Test of Memory and Learning (Reynolds & Bigler, 1994). Thus, in regard to increase in scores with age, the tests behave in a highly similar manner for Whites, Blacks, and Mexican-Americans. Similar findings occur in male–female comparisons.

In the review work of Moran (1990) and in a search for more recent work, it is apparent that only a few studies of the differential construct validity of personality tests have been undertaken, despite large mean differences across ethnicity and gender on such popular measures as the stalwart MMPI. Dana (1993) sharply criticizes even the meager work to date on the MMPI, stating that "[C]omparative MMPI research studies of Anglo-American and other cultural groups have typically not only used inappropriate statistics but also failed to equate groups adequately on socioeconomic criteria or even to define ethnicity" (p. 98). The lack of extensive research using method designed to detect test bias for such widely used scales as the MMPI is nothing short of appalling. A look at the newer MMPI-2 Manual suggests gender differs in construct validity but provides no real evidence either way. A few studies of factorial similarity of instruments such as the Revised Children's Manifest Anxiety Scale show little bias and high degrees of similarity by ethnicity and gender (Moran, 1990; Reynolds & Paget, 1981).

Constuct validity of a large number of popular psychometric assessment instruments has been investigated across races and genders with a variety of populations of minority and White children and with a divergent set of methodologies (see Reynolds, 1982b, for a review of methodologies). All roads have led to Rome: no consistent evidence of bias in construct validity has been found with any of the many tests investigated. This leads to the conclusion that psychological tests (especially aptitude tests) function in essentially the same manner, that test materials are perceived and reacted to in a similar manner, and that tests measure the same construct with equivalent accuracy for Blacks, Whites, Mexican-Americans, and other American minorities of both sexes and at all levels of SES. Single-group validity and differential validity have not been found and probably do not exist with regard to well-constructed and well-standardized psychological and educational tests of intellect, personality, or behavior although data are not as thorough for the latter two categories and much work is needed here.

10.03.7.3 Bias in Predictive or Criterion-related Validity

Evaluating bias in predictive validity of educational and psychological tests is less closely related to the evaluation of group mental test score differences than to the evaluation of individual test scores in a more absolute sense. This is especially true for aptitude (as opposed to personality or diagnostic) tests, where the primary purpose of administration is the prediction of some specific future outcome or behavior. Internal analyses of bias (such as in content and construct validity) are less confounded than analyses of bias in predictive validity, however, because of the potential problems of bias in the criterion measure. Predictive validity is strongly influenced by the reliability of criterion measures, which frequently is poor. The degree of relationship between a predictor and a criterion is restricted as a function of the square root of the product of the reliabilities of the two variables.

Arriving at a consensual definition of bias in predictive validity is also a difficult task, as previously discussed. Yet, from the standpoint of the practical applications of aptitude and intelligence tests, predictive validity is the most crucial form of validity in relation to test bias. Much of the discussion in professional journals concerning bias in predictive validity has centered around models of selection. These issues have been discussed in Section 10.03.6. Since this section is concerned with bias in respect to the test itself and not the social or political justification of any one particular selection model, the Cleary at al. (1975) definition, slightly rephrased here, provides a clear and direct statement of test bias with regard to predictive validity: a test is considered biased with respect to predictive validity if the inference drawn from the test score is not made with the smallest feasible random error or if there is constant error in an inference or prediction as a function of membership in a particular group. This definition is a restatement of previous definitions by Cardall and Coffman (1964), Cleary (1968), and Potthoff (1966), and has been widely accepted (though certainly not without criticism; e.g., Bernal, 1975; Linn & Werts, 1971; Schmidt & Hunter, 1974; Thorndike, 1971).

Oakland and Matuszek (1977) examined procedures for placement in special education classes under a variety of models of bias in prediction, and demonstrated that the smallest number of children are misplaced when the Cleary et al (1975) conditions of fairness are met. (However, under "quota" system requirements, Oakland and Matuszek favor the Thorndike, 1971 conditions of selection.) The Cleary et al. definition is also apparently the definition espoused in US government guidelines on testing and has been held in at least one court decision (Cortez v. Rosen, 1975) to be the only historically, legally, and logically required condition of test fairness (Ramsay, 1979), although apparently the judge in the Larry P. v. Riles (1979) decision (but overturned on appeal) in the US Federal court system adopted the "mean score differences as bias" approach. A variety of educational and psychological personnel long have adopted the Cleary et al. regression approach to bias, including: (i) noted psychological authorities on testing (Anastasi, 1986; Cronbach, 1970; Humphreys, 1973); (ii) educational and psychological researchers (Brossard, Reynolds, & Gutkin, 1980; Kallingal, 1971; Pfeifer & Sedlacek, 1971; Reynolds & Hartlage, 1978, 1979; Stanley & Porter, 1967; Wilson, 1969); (iii) industrial/organizational psychologists (Barlett & O'Leary, 1969; Einhorn & Bass, 1970; Gael & Grant, 1972; Grant & Bray, 1970; Ramsay, 1979; Tenopyr, 1967); and (iv) even critics of educational and psychological testing (Goldman & Hartig, 1976; Kirkpatrick, 1970; Kirkpatrick, Ewen, Barrett, & Katzell, 1968).

The evaluation of bias in prediction under the Cleary et al. (1975) definition (the regression definition) is quite straightforward. With simple regression, predictions take the form of $\hat{Y}_{\mathrm{i}} = aX_{\mathrm{i}} + b$, where a is the regression coefficient and b is a constant. When this equation is graphed (forming a regression line), a represents the slope of the regression line and b the Y intercept. Since our definition of fairness in predictive validity requires errors in prediction to be independent of group membership, the regression line formed for any pair of variables must be the same for each group for whom predictions are made. Whenever the slope or the intercept differs significantly across groups, there is bias in prediction if one attempts to use a common equation for all groups. However, if the regression equations for two (or more) groups are equivalent, prediction is the same for all groups. This condition is referred to variously as "homogeneity of regression across groups," "simultaneous regression," or "fairness in prediction." Homogeneity of regression across groups is illustrated in Figure 1. In this case, the single regression equation is appropriate with all groups, any errors in prediction being random with respect to group membership (i.e., residuals uncorrelated with group membership). When homogeneity of regression does not occur, for "fairness in prediction" to occur, separate regression equations must be used for each group.

In actual clinical practice, regression equations are seldom generated for the prediction of future performance. Instead, some arbitrary or perhaps statistically derived cutoff score is determined, below which "failure" is predicted. For school performance, IQs two or more standard deviations below the test mean are used to infer a high probability of failure in the regular classroom if special assistance is not provided for the student in question. Essentially, then, clinicians are establishing mental prediction equations that are assumed to be equivalent across ethnicity, gender, and so on. Although these mental equations cannot be tested readily across groups, the actual form of criterion prediction can be compared across groups in several ways. Errors in prediction must be independent of group membership. If regression equations are equal, this condition is met. To test the hypothesis of simultaneous regression, slopes and intercepts must both be compared. An alternative method is the direct examination of residuals through ANOVA or a similar design (Reynolds, 1980b).

In the evaluation of slope and intercept values, two basic techniques most often have been employed in the research literature. Gulliksen and Wilks (1965) and Kerlinger (1973) describe methods for separately testing regression coefficients and intercepts for significant differences across groups. Using separate, independent tests for these two values considerably increases the probability of a decision error and unnecessarily complicates the decision-making process. Potthoff (1966) has described a useful technique that allows one to test simultaneously the equivalence of regression coefficients and intercepts across K independent groups with a single F ratio (the Potthoff equations may be found also in Reynolds, 1982b). If a significant F results, the researcher may then test the slopes and intercepts separately if information concerning which value differs is desired. When homogeneity of regression does not occur, there are three basic conditions that can result: (i) intercept constants differ, (ii) regression coefficients (slopes) differ, or (iii) slopes and intercepts differ. These conditions are depicted pictorially in Figures 2, 3, and 4, respectively.

The regression coefficient is related to the correlation coefficient between the two variables and is one measure of the strength of the relationship between two variables. When intercepts differ and regression coefficients do not, a situation such as that shown in Figure 2 results. Relative accuracy of prediction is the same for the two groups (*a* and *b*), the use of a regression equation derived by combining the two groups results in bias that works against the group with the higher mean criterion score. Since the slope of the regression line is the same for all groups, the degree of error in prediction remains constant and does not fluctuate as a function of an individual's score on the independent variable. That is, regardless of group member *b*s score on the predictor, the degree of underprediction in performance on the criterion is the same. As illustrated in Figure 2, the use of the common score of Y_c for a score of X overestimates how well members of group *a* will perform and underestimates the criterion performance of members of group *b*.

In Figure 3, nonparallel regression lines illustrate the case where intercepts are constant across groups but the slope of the line is different for each group. Here, too, the performance of the group with the higher mean criterion score is typically underpredicted when a common regression equation is applied. The amount of bias in prediction that results from using the common regression line is the distance of the score from the mean. The most difficult, complex case of bias is represented in Figure 4 where the result of significant differences in slopes and intercepts is shown. Not only does the amount of bias vary but even the direction can reverse, depending on the location of the individual's score in the distribution of the independent variable. Only in the case of Figure 4 do members of the group with the lower mean criterion score run the risk of having their performance on the criterion variable underpredicted by the application of a common regression equation.

A considerable body of literature has developed regarding the differential predictive

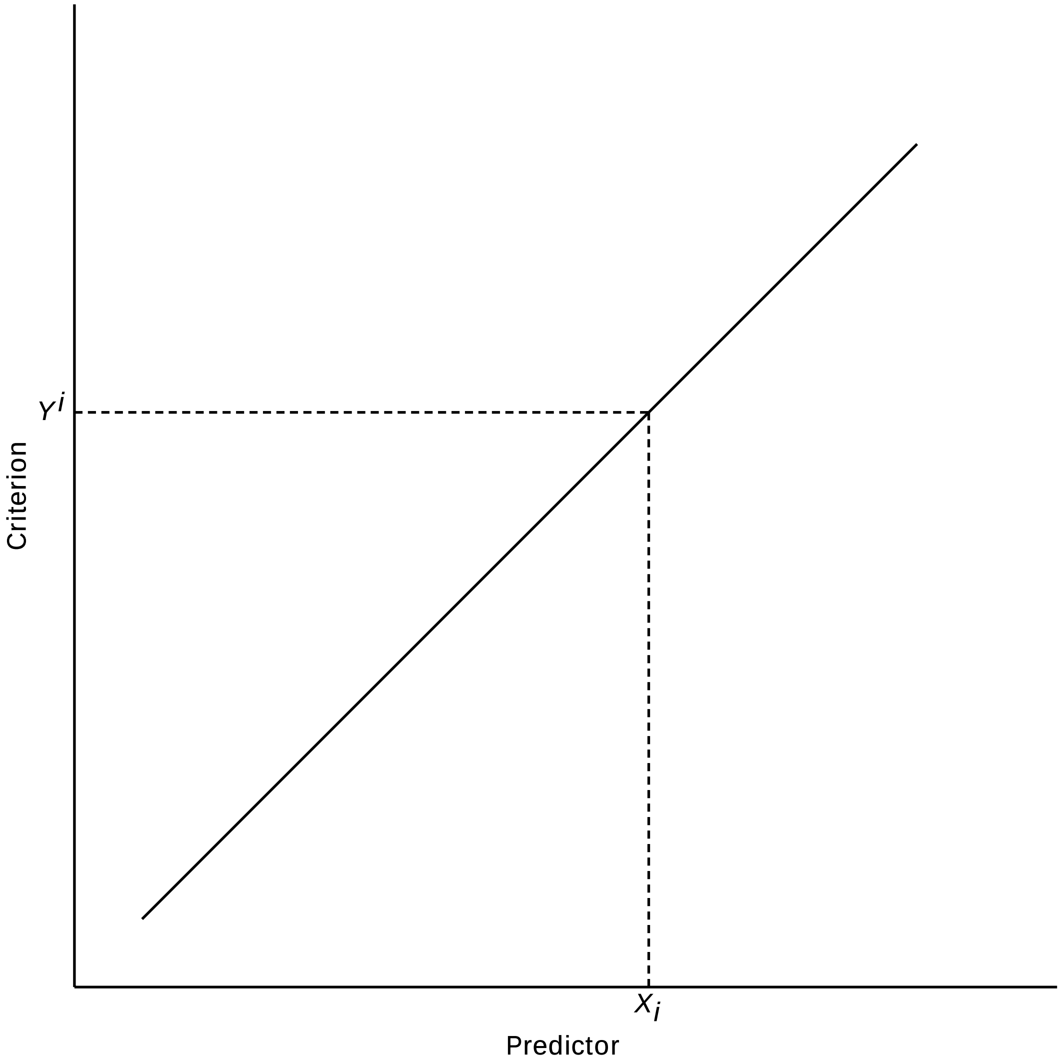

Figure 1 Equal slopes and intercepts result in homogeneity of regression that causes the regression lines for group *a*, group *b*, and the combined group *c* to be identical.

validity of IQ tests and for tests for employment selection and college admissions. However, virtually nothing appears in this regard with reference to personality tests (Dana, 1993; Moran, 1990) and this is a major weakness in the literature.

In a review of 866 Black–White test validity comparisons from 39 studies of test bias in personnel selection, Hunter, Schmidt, and Hunter (1979) concluded that there was no evidence to substantiate hypotheses of differential or single-group validity with regard to the prediction of job performance across races for Blacks and Whites. A similar conclusion was reached by Jensen (1980), O'Conner, Wexley, and Alexander (1975), and Reynolds (1982a, 1995) among others. A number of studies have also focused on differential validity of the Scholastic Aptitude Test (SAT) in the prediction of college performance (typically measured by grade point average, GPA). In general, these studies have found either no differences in the prediction of criterion performance for Blacks and Whites or a bias (underprediction of the criterion) against Whites (Cleary, 1968; Cleary et al., 1975; Goldman & Hewitt, 1976; Kallingal, 1971; Pfeifer & Sedlacek, 1971; Stanley, 1971; Stanley & Porter, 1967; Temp, 1971). When bias against Whites has been found, the differences between actual and predicted criterion scores, although statistically significant, have been quite small.

Reschly and Sabers (1979) evaluated the validity of WISC-R IQs in the prediction of Metropolitan Achievement Tests (MAT) performance (Reading and Math subtests) for Whites, Blacks, Mexican-Americans, and native American Papagos. The choice of the MAT as a criterion measure in studies of predictive bias is particularly appropriate, since item analysis

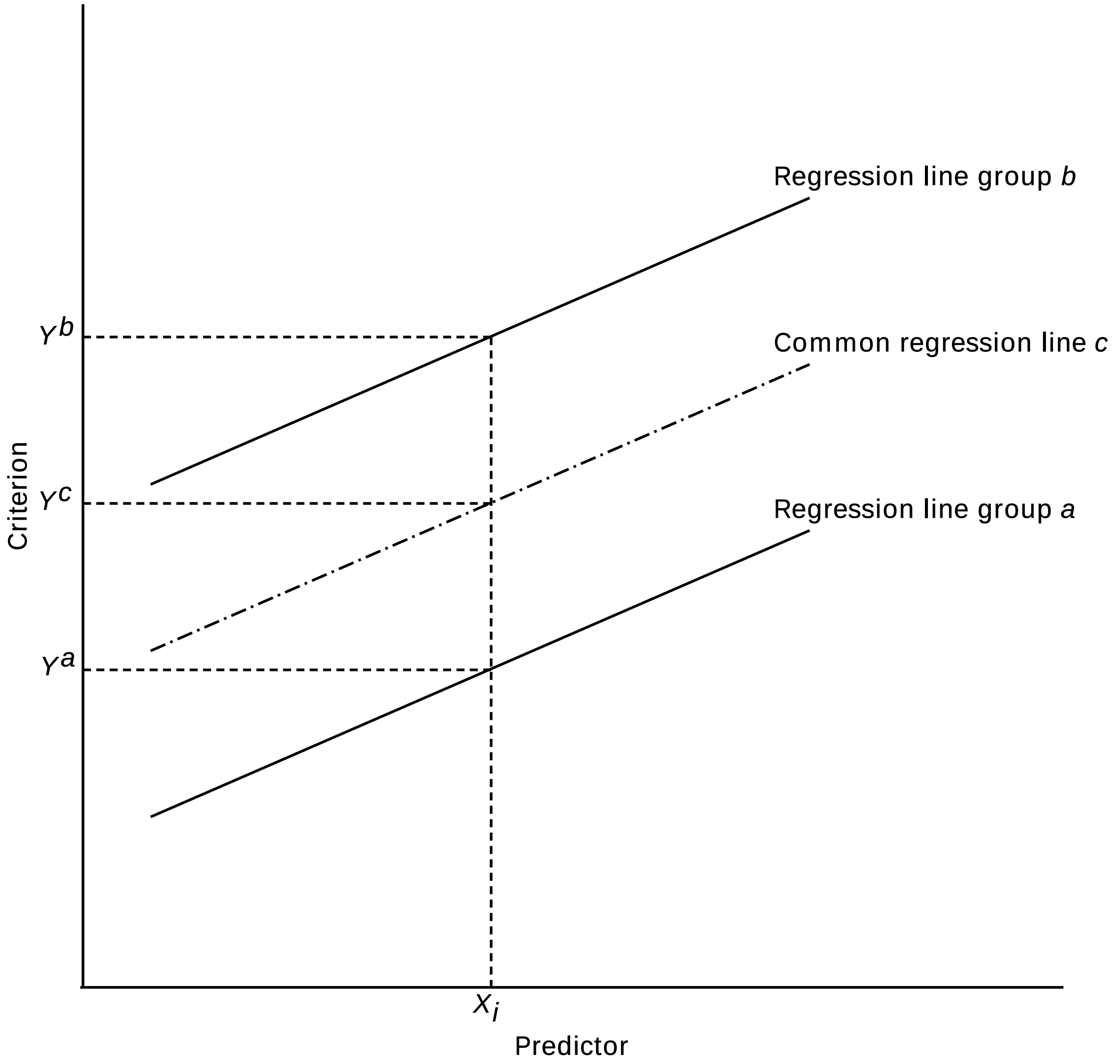

Figure 2 Equal slopes with differing intercepts result in parallel regression lines and a constant bias in prediction.

procedures were employed (as described earlier) to eliminate racial bias in item content during the test construction phase. Anastasi (1986) has described the MAT as an excellent model of an achievement test designed to reduce or eliminate cultural bias. Reschly and Saber's (1979) comparison of regression systems indicated bias in the prediction of the various achievement scores. Again, however, the bias produced generally significant underprediction of White performance when a common regression equation was applied. Achievement test performance of the native American Papago group showed the greatest amount of overprediction of all non-White groups. Using similar techniques, but including teacher ratings, Reschly and Reschly (1979) also investigated the predictive validity of WISC-R factor scores with the samples of White, Black, Mexican-American, and native American Papago children. A significant relationship occurred between the three WISC-R factors first delineated by Kaufman (1975) and measures of achievement for the White and non-White groups, with the exception of the Papagos. Significant correlations occurred between the WISC-R Freedom from Distractibility factor (Kaufman, 1975) and teacher ratings of attention for all four groups. Reschly and Reschly concluded that "These data also again confirm the relatively strong relationship of WISC-R scores to achievement for most non-Anglo as well as Anglo groups" (1979, p. 239). Reynolds and Hartlage (1979) investigated the differential validity of Full Scale IQs from the WISC-R and its 1949 predecessor, the WISC, in predicting reading and arithmetic achievement for Black and White children who had been referred by their

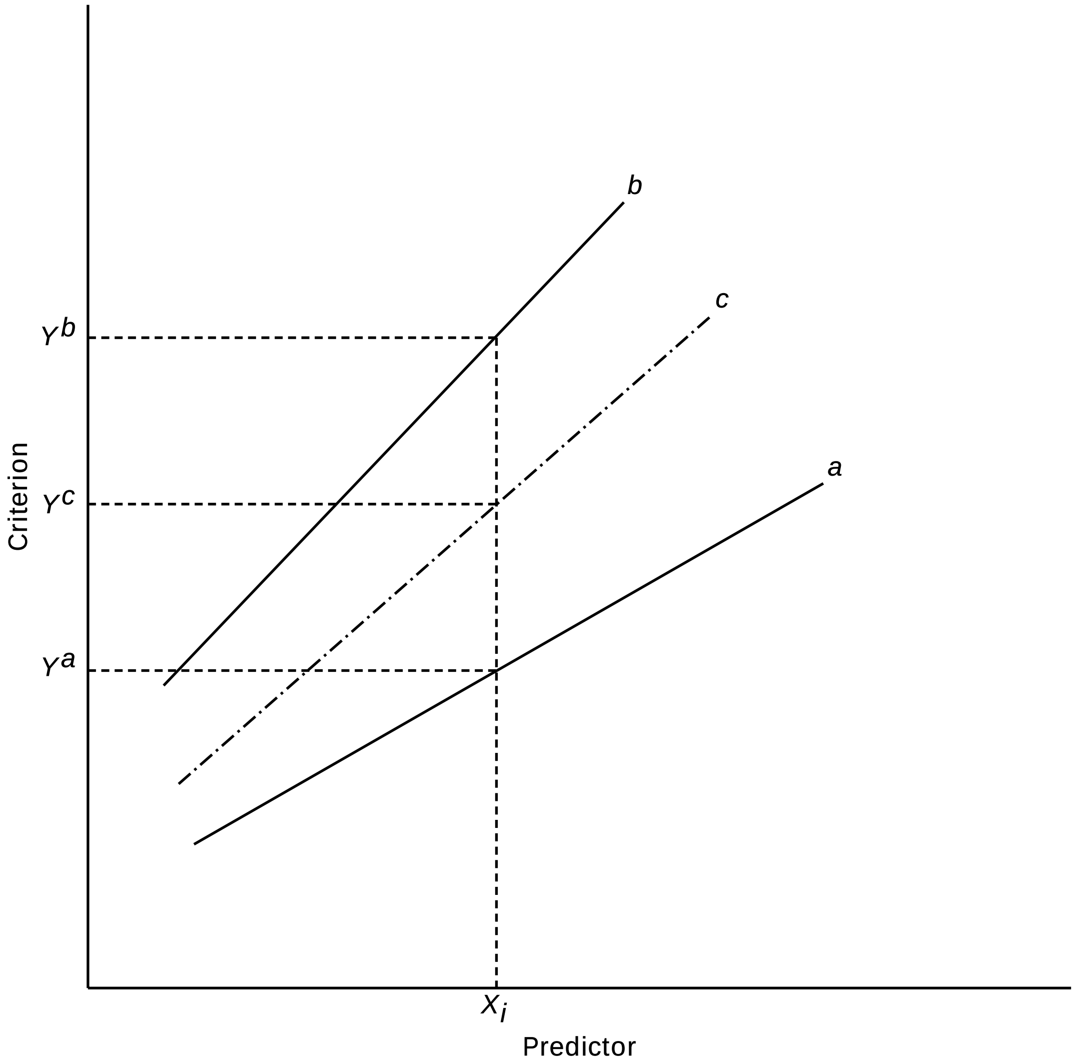

Figure 3 Equal intercepts and differing slopes result in nonparallel regression lines with the degree of bias dependent on the distance of the individual's score (x_i) from the origin.

teachers for psychological services in a rural Southern school district. Comparisons of correlations and a Potthoff (1966) analysis to test for identity of regression lines revealed no significant differences in the ability or function of the WISC and WISC-R to predict achievement for these two groups. Reynolds and Gutkin (1980b) replicated this study for the WISC-R with large groups of White and Mexican-American children from the Southwest. Reynolds and Gutkin contrasted regression systems between WISC-R Verbal, Performance, and Full Scale IQs and the "academic basics" of reading, spelling, and arithmetic. Only the regression equation between the WISC-R Performance IQ and arithmetic achievement differed for the two groups. The difference in the two equations was due to an intercept bias that resulted in the overprediction of achievement for the Mexican-American children. Reynolds, Gutkin, Dappen, and Wright (1979) also failed to find differential validity in the prediction of achievement for males and females with the WISC-R.

In a related study, Hartlage, Lucas, and Godwin (1976) compared the predictive validity of what they considered to be a relatively culture-free test (Raven's Progressive Matrices) with a more culture-loaded test (the 1949 WISC, a test developed and standardized on all White sample) for a group of low-SES, disadvantaged rural children. Hartlage et al. (1976) found that the WISC had consistently larger correlations with measures of reading, spelling, and arithmetic than Raven's Matrices. Although it did not make the comparison with other groups that is necessary for the drawing of firm conclusions, the study does support the validity of the WISC,

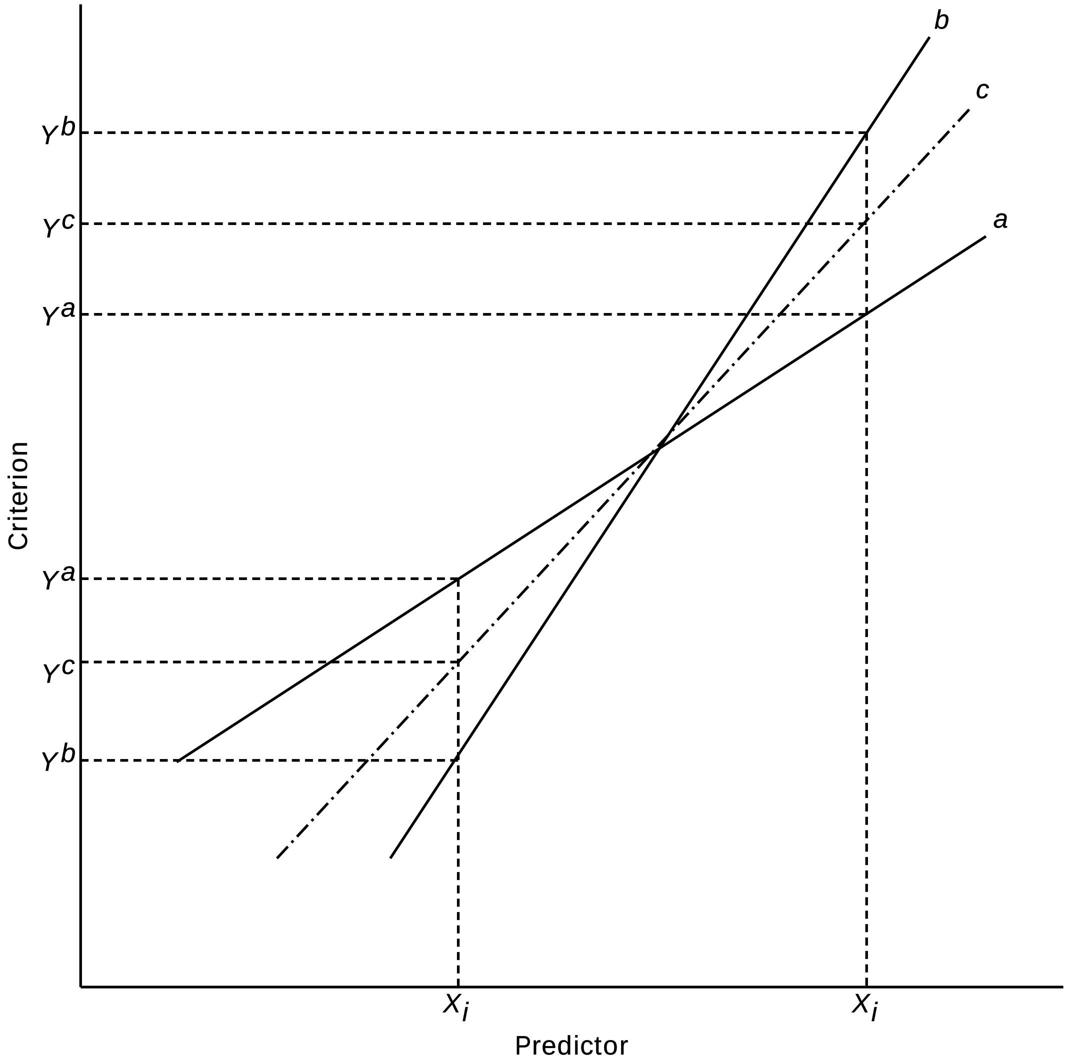

Figure 4 Differing slopes and intercepts result in the complex condition where the amount and the direction of the bias are a function of the distance of an individual's score from the origin.

which has been the target of many of the claims of bias in the prediction of achievement for low-SES, disadvantaged rural children. Henderson, Butler, and Goffeney (1969) also reported that the WISC and the Bender–Gestalt Test were equally effective in the prediction of reading and arithmetic achievement for White and non-White groups, though their study had a number of methodological difficulties, including heterogeneity of the non-White comparison group. Reynolds, Willson, and Chatman (1985) evaluated the predictive validity of the K-ABC for Blacks and Whites. Occasional evidence of bias was found in each direction, but mostly in the direction of overprediction of the academic attainment levels of Blacks.

Bossard et al. (1980) published a regression analysis of test bias on the 1972 Stanford–Binet Intelligence Scale for separate groups of Black and White children. Neither regression systems nor correlations differed at $p < 0.05$ for the prediction of the basic academic skills of reading, spelling, and arithmetic achievement for these two groups of referred children. An earlier study by Sewell (1979), a Black opponent of testing, did not compare regression systems, but also found no significant differences in validity coefficients for Stanford–Binet IQs predicting California Achievement Test (CAT) scores for Black and White first-grade children.

A series of studies comparing the predictive validity of group IQ measures across races has been reviewed by Jensen (1980) and Sattler (1974). Typically, regression systems have not been compared in these studies: instead, researchers have compared only the validity coefficients across races—a practice that tells only whether the test is potentially nonbiased. The comparison of validity coefficients is nevertheless relevant, since equivalence in predictive validities is a first step in evaluating differential validity. That is, if predictive validities differ, then regression systems must differ; the reverse is not necessarily true, however, since the correlation between two variables is a measure of the strength or magnitude of a relationship and does not dictate the form of a relationship. Although the number of studies evaluating group IQ tests across ethnicity is small, they have typically employed extremely large samples. The Lorge–Thorndike verbal and nonverbal IQs have been most often investigated. Jensen (1980) and Sattler (1974) concluded that the few available studies suggest that standard IQ tests in current use have comparable validities for Black and White children at the elementary school level.

Guterman (1979) reported on an extensive analysis of the predictive validity of the Ammons and Ammons Quick Test (QT; a measure of verbal IQ) for adolescents of different social classes. Social class was determined by a weighted combination of Duncan's SES index and the number of years of education of each parent. Three basic measures: (i) the Vocabulary subtest of the General Aptitude Test Battery (GATB); (ii) the test of Reading Comprehension from the Gates Reading Survey; and (iii) the Arithmetic subtest of the GATB were used. School grades in academic subjects for 9th, 10th, and 12th grades were also used to examine for bias in prediction. Guterman reached similar conclusions with regard to all criterion measures across all social classes: slopes and intercepts of regression lines did not differ across social class for the prediction of any of the criterion measures by the IQ derived from the QT. Several other social knowledge criterion measures were also examined. Again, slopes were constant across social class, and, with the exception of sexual knowledge, intercepts were also constant. Guterman concluded that his data provided strong support for equivalent validity of IQ measures across social class. In reanalyzing the Guterman (1979) study, Gordon and Rudert (1979) reached even stronger conclusions. Certainly with school-age children and adults, there is compelling evidence that differential and single-group predictive validity hypotheses must be rejected.

As with constrict validity, at the preschool level the evidence is less clear and convincing but points toward a lack of bias against minorities. Because of doubts expressed about the usefulness of customary readiness tests with students of certain racial and ethnic backgrounds and with low-SES children, Mitchell (1967) investigated the predictive validity of two preschool readiness tests used in the US Office of Education (Cooperative First-Grade Reading Study, 1964–1965 revision) and the Murphy–Durell Reading Readiness Analysis (1964 revision). Mitchell concluded that the two readiness tests performed their functions as well with Black as with White children and that the general level of predictive validity was similar. This overstates the case somewhat, since only validity coefficients and not regression systems were compared, but Mitchell's (1967) study does support the predictive validity of these readiness tests across race.

Oakland (1978) assessed the differential predictive validity of four readiness tests (the MRT, the Tests of Basic Experiences battery, the Slosson Intelligence Test, and the Slosson Oral Reading Test) across races (Black, White, and Mexican-American) for middle- and lower-SES children. The MAT, the CAT, and the California Test of Mental Maturity (CTMM) served as criterion variables. Since the CTMM is an IQ test, prediction of CTMM scores by the various readiness tests is excluded from the following discussion. Although Oakland (1978) did not use any test of statistical significance to compare the correlations between the independent and dependent variable pairs across ethnicity and SES, a clear pattern was found, showing higher levels of prediction for White as opposed to non-White groups. Oakland also did not compare regression systems, limiting his study to the report of the various validity coefficients for each race–SES grouping. Oakland's (1978) results clearly indicate potential bias in the prediction of early school achievement by individual readiness or screening tests. The lower correlations for non-White groups, however, given their lower mean criterion scores, led to anticipation of bias favoring non-Whites in the prediction of early school achievement.

To investigate this possibility, Reynolds (1978) conducted an extensive analysis of predictive bias for seven major preschool tests (the Draw-a-Design and Draw-a-Child subtests of the McCarthy Scales; the Mathematics and Language subtests of the Tests of Basic Experiences; the Preschool Inventory-Revised Edition; and the Lee–Clark Reading Readiness Test) across race and gender for large groups of Blacks and Whites. For each preschool test, validity coefficients, slopes, and intercepts were compared, with prediction of performance on four subtests of the MAT (Word Knowledge, Word Discrimination, Reading, and Arithmetic) as the criterion measures. The general advantage of the MAT as a criterion in external studies of bias has previously been pointed out.

In the Reynolds (1978) study, the MAT had the added advantage of being chosen by the teachers in the district: data were gathered on a large number of early achievement tests, and the teachers selected the MAT as the battery most closely measuring what was taught in their classrooms. Regression systems and validity coefficients were compared for each independent–dependent variable pair for White females (WF) vs. White males (WM), Black females (BF) vs. Black males (BM), WF vs. BF, and WM vs. BM, resulting in 112 comparisons of validity coefficients and 112 comparisons of regression systems. Although the mean correlations were slightly lower for Blacks, the 112 comparisons of pairs of correlations revealed only three significant differences, a less-than-chance occurrence with this number of comparisons. Using the Potthoff (1966) technique for comparing regression lines produced quite different results. Of the 112 comparisons of regression lines, 43 (38.4%) showed differences. For comparisons with race as the major variable (and gender controlled), 31 (55.2%) of the 56 comparisons showed significantly different regression lives. Clearly, racial bias was significantly more prevalent than gender bias ($p < 0.01$) in prediction. In comparing the various pretests, bias occurred most often with the Preschool Inventory and the Lee–Clark, whereas none of the comparisons involving the MRT showed bias. Though race clearly influenced homogeneity of regression across groups, the bias in each case acted to overpredict performance of lower-scoring groups; thus the bias acted against Whites and females and in favor of Blacks and males. A follow-up study (Reynolds, 1980b) has indicated one potential method for avoiding bias in the prediction of early school achievement with readiness or screening measures.

Brief screening measures, especially at the preschool level, typically do not have the high level of reliability obtained by such instruments as the WISC-R or the Stanford–Binet. Linn and Werts (1971) have demonstrated convincingly that poor reliability can lead to bias in prediction. Early screening measures, as a rule, also assess a very limited area of functioning, rather than allowing the child to demonstrate its skills in a variety of areas of cognitive functioning. The one well-researched, reliable, broad-based readiness test, the MRT, has failed to show bias with regard to internal or external criteria. Comprehensive and reliable individual preschool instruments such as the WPPSI and the McCarthy Scales, while showing no internal evidence of test bias, have not been researched with regard to predictive bias across race. Reynolds (1980b) examined the predictive validity of the seven preschool measures described previously when these were combined into a larger battery, thus increasing the scope and reliability of the assessment.

Since the definition of predictive bias noted earlier requires that errors in prediction be independent of group membership, Reynolds (1980b) directly examined residuals (a "residual term" is the remainder when the predicted score for an individual is subtracted from the individual's obtained score) across race and gender when the seven-test battery was used to predict MAT scores in a multiple-regression formula. Subtests of the seven-test battery were also examined. Results of a race X sex ANOVA of residuals for each of the MAT subtests when the seven-test battery was employed revealed no significant differences in residuals across races and genders, and no significant interactions occurred. When a subset of the larger battery was submitted to the same analysis, racial bias in prediction did not occur; however, a significant F resulted for gender effects in the prediction of two of the four MAT subscores (Word Discrimination and Word Knowledge). Examination of the residuals for each group showed that the bias in prediction was again against the group with the higher mean criterion scores: there was a consistent underprediction of performance for females. The magnitude of the effect was small, however, being of the order of 0.13–0.16 standard deviations. Thus, at the preschool level, the only convincing evidence of bias in predictive validity is a gender effect, not a race effect.

Kamphaus and Reynolds (1987) reviewed the available literature on predictive bias with the K-ABC and concluded that overprediction of Black children's performance in school is more common with the K-ABC, particularly the K-ABC Sequential Processing scale, than with other tests. The effects are small, however, and are mitigated in large part by using the K-ABC Mental Processing Composite. Some bias also occurs against Blacks, but when the extensive nature of the bias research with the K-ABC is considered, results with the K-ABC are not substantially different from the results with the WISC-R (with the exception of overprediction of Black academic performance by the K-ABC Sequential Processing scale).

Keith and Reynolds (1990) have suggested the use of path analysis as an alternative model for assessing bias is predictive validity. In such a path model, ability would be proposed to predict achievement, and group membership would be assessed as a moderator variable. A diagrammatic representation of a biased and an unbiased model is shown in Keith and Reynolds (1990). Bias in prediction would exist in such a model when group membership affects measured ability independent of true ability, that is, errors of measurement in testing of ability would be correlated with group membership. With regard to bias in predictive validity, the empirical evidence suggests conclusions similar to those regarding bias in content and construct validity. There is no strong evidence to support contentions of differential single-group validity. Bias occurs infrequently and with no apparently observable pattern, except when instruments of poor reliability and high specificity of test content are examined. When bias occurs, it is most often in the direction of favoring low-SES, disadvantaged ethnic minority children, or other low-scoring groups.

10.03.8 CROSS-CULTURAL TESTING WHEN TRANSLATION IS REQUIRED

When a test is translated from one language to another, the research findings discussed thus far do not hold. It is inappropriate simply to translate a test and apply it in a different linguistic culture. A test must be redeveloped from scratch (although constructs may be retained) basically before any such application would be appropriate. New items, new normative data, and new scaling would all be required. This has been known since the early days of psychological assessment and testing. In the early 1900s, when the Binet–Simon tests were brought to the USA from France, approximately 30 different versions of the test were developed in the USA by various researchers. However, most of these were mere translations or contained minor modifications to adapt to American culture. The Stanford–Binet Intelligence Scale, in its various incarnations, however, became the standard bearer for measurement of intelligence for nearly 60 years and was even more popular in France at one time than the original French Binet–Simon scales. The reason for the domination of the Stanford–Binet series was Lewis Terman's insight and tenacity in redeveloping the test in the USA. After determining that Binet's theory of intelligence applied, new items were written, tried out, and a new scale devised for norming that was conceptually consistent with the Binet–Simon scales but in its practical application was a new and different test.

The problems in translating verbal and nonverbal concepts across linguistic cultures are difficult but in any event the redevelopment of tests in such circumstances seems required. Cronbach and Drenth (1972) provide a book length treatment of these problems and various experiences with proposed solutions to cross-cultural adaptation of psychological tests stemming from some 30 nations from throughout the world. The various contributors describe both the strengths and limitations of adapting tests cross-culturally from one country to another, providing perspectives from such diverse discipline psychometrics, cognitive development, psychology, and anthropology. More recent guidelines and reviews of the issues involved in cross-cultural adaptation of psychological and educational tests can be found in Hambleton (1994), Hambleton and Kanjee (1995), and Van de Vijver and Hambleton (1996).

10.03.9 SUMMARY AND FUTURE DIRECTIONS

There is little question that the issue of bias in mental testing is an important one with strong historical precedence in the social sciences and, ultimately, formidable social consequences. Because the history of mental measurement has been closely wed from the outset to societal needs and expectations, testing in all forms has remained in the limelight, subjected to the crucible of social inspection, review, and (at times) condemnation in various cultures throughout the world. However, the fact that tests and measures of human aptitude and achievement continue to be employed in most modern cultures indicates strongly that the practice has value, despite the recurring storms of criticism over the years. The ongoing controversy related to test bias and the "fair" use of measures will undoubtedly remain with the social sciences for at least as long as we intertwine the nature–nuture question with these issues and affirm differences between–among groups in mean performance on standardized tests. Numerous scholars in the field of psychometrics have been attempting to separate the nature–nurture issue and data on mean score differences from the more orderly, empirically-driven specialty of bias investigation, but the separation will undoubtedly not be a clean one. A sharp distinction has developed between the popular press and scientific literature with regard to the interpretation of mental measurement research. The former all too often engenders beliefs that biased measures are put into use for socially pernicious purposes and psychology and education are often accused of courting political, social, and professional ideologies. The former appears to have created confusion in public opinion concerning the possibility of "fair" testing, to say the least. The latter—reported in this chapter—has been demonstrating through a rather sizable body of data that the hypothesis of cultural bias in tests is not a particularly strong one at present, at least in cultures with a common language and some degree (the extent or qualitative features of which are as yet indeterminant) of common experience. In any event, societal scrutiny and ongoing sentiment about testing have without question served to force the psychometric community to refine its definitions of bias further, to inspect practices in the construction of nonbiased measures, and to develop statistical procedures to detect bias when it is occurring. We can argue whether the social sciences have from the outset overstepped their bounds in implementing testing for social purposes before adequate data and methods were developed, but the resulting advancements made in bias technology in response to ongoing public inspection are undeniable.

Data from the empirical end of bias investigation do suggest several guidelines to follow in order to ensure equitable assessment. Points to consider include: (i) investigation of possible referral source bias, as there is evidence that persons are not always referred for mental health or special eductional services on the basis of impartial, objective rationales; (ii) inspection of test developers' data for evidence that sound statistical analyses for bias across groups to be evaluated with the measure have been completed; (iii) assessment with the most reliable measures available; and (iv) assessment of multiple abilities with multiple methods. In other words, psychologists need to view multiple sources of accurately derived data prior to making decisions concerning individuals. We may hope that this is not too far afield from what has actually been occurring in the practice on psychological assessment, though one continues to hear isolated stories of grossly incompetent diagnostic decisions being made (e.g., Mason, 1979). This does not mean that psychologists should be blind to a person's environmental background. Information concerning the home, community, and other environmental circumstances must all be evaluated in the individualized decision-making process. Exactly how this may be done is addressed in all other chapters and volumes of this work. Neither, however, can the psychologist ignore the fact that ethnic minority group members who score at levels indicative of psychopathology are just as likely to have problems and need intervention as are majority-class individuals. Indeed, it is the purpose of the assessment process to beat the prediction—to provide insight into hypotheses for environmental and biological interventions that will prevent the identified pathology from continuity to exert a negative influence on the person's life.

A philosophical perspective is emerging in the bias literature that is requiring test developers not only to demonstrate whether their measures demonstrate differential content, construct, and predictive validity across groups *prior* to publication, but also to incorporate in some form content analyses by interested groups to ensure that offensive materials are omitted. Although there are no sound empirical data to suggest that persons can determine bias upon surface inspection, the synergistic relationship between test use and pure psychometrics must be acknowledged and accommodated in an orderly fashion before tests gain greater acceptance within society. Ideally, a clear consensus on "fairness" (and steps taken to reach this end) is needed between those persons with more subjective concerns and those interested in gathering objective bias data during and after test construction. Accommodation along this line will ultimately ensure that all parties interested in any given test believe that the measure in question is nonbiased and that the steps taken to achieve "fairness" can be held up to public scrutiny without reservation. Given the significant and reliable methods developed over the last several decades in bias research, it is untenable at this point to abandon statistical analyses in favor of "armchair" determinations of bias. Test authors and publishers need to demonstrate factorial invariance across all groups for whom the test is designed in order to make the instrument more readily interpretable. Comparisons of predictive validity across races and genders during the test development phase are also needed. With the exception of some recent achievement tests, this has not been common practice, yet it is at this stage that tests can be altered through a variety of item analysis procedures to eliminate any apparent racial and sexual bias. As scientists, we must also inform the media of our results, whatever they may be.

Bias research in the area of personality testing must be expanded. Little has been done and this represents a major weakness in the literature. Only recently (e.g., Reynolds & Kamphaus, 1992) have publishers begun to give appropriate attention to this problem. Researchers in personality and psychodiagnostics must move ahead in this area of concern. Similar problems exist in the growing area of neuropsychological testing (e.g., Reynolds, 1997a).

10.03.10 REFERENCES

Abebimpe, V. R., Gigandet, J., & Harris, E. (1979). MMPI diagnosis of Black psychiatric patients. *American Journal of Psychiatry, 136,* 85–87.

Alley, G., & Foster, C. (1978). Nondiscriminatory testing of minority and exceptional children. *Focus on Exceptional Children, 9,* 1–14.

American Psychological Association (1970). Psychological assessment and public policy. *American Psychologist, 25,* 264–266.

Anastasi, A. (1982). *Psychological testing* (5th ed.). New York: Macmillan.

Anastasi, A. (1986). *Psychological testing* (4th ed.). New York: Macmillan.

Angoff, W. H. (1976). Group membership as a predictor variable: A comment on McNemar. *American Psychologist, 31,* 612.

Angoff, W. H., & Ford, S. R. (1973). Item-race interaction on a test of scholastic aptitude. *Journal of Educational Measurement, 10,* 95–106.

Auerbach, A. G. (1971). The social control of learning disabilities. *Journal of Learning Disabilities, 4,* 369–378.

Bartlett, C. J., & O'Leary, B. S. (1969). A differential prediction model to moderate the effect of heterogeneous groups on personnel selection and classification. *Personnel Psychology, 22,* 1–19.

Bass, A. R. (1976). The "equal risk" model: A comment on McNemar. *American Psychologist, 31,* 611–612.

Bayley, N. (1969). *Bayley Scales of Infant Development*. New York: Psychological Corporation.

Beauchamp, D. P., Samuels, D. D., & Griffore, R. J. (1979). WISC-R Information and Digit Span scores of American and Canadian children. *Applied Psychological Measurement, 3,* 231–236.

Berk, R. A. (Ed.) (1982). *Handbook of methods for detecting test bias*. Baltimore: Johns Hopkins University Press.

Bernal, E. M. (1975). A response to "Educational uses of tests with disadvantaged students." *American Psychologist, 30,* 93–95.

Binet, A., & Simon, T. (1905). Methods nouvelles pour le diagnositic du nivea intellectual des anormaux. *L'AnnCe Psychologique, 11,* 191–244.

Binet, A., & Simon, T. (1916). *The development of intelligence in children*. New York: Arno (English translation reprinted, 1973).

Birch, H. G., & Gussow, J. D. (1970). *Disadvantaged children: Health, nutrition, and school failure*. New York: Saunders.

Blaha, J., Wallbrown, F., & Wherry, R. J. (1975). The hierarchical factor structure of the Wechsler Intelligence Scale for Children. *Psychological Reports, 35,* 771–778.

Bogen, J. E., DeZure, R., Tenhouten, N., & March, J. (1972). The other side of the brain: IV. The A/P ratio. *Bulletin of the Los Angeles Neurological Society, 37,* 49–61.

Bolles, R. C. (1975). *Theory of motivation* (2nd ed.). New York: Harper Collins.

Bond, L. (1981). Bias in mental tests. In B. F. Green (Ed.), *Issues in testing: Coaching, disclosure, and ethnic bias*. San Francisco: Jossey-Bass.

Bossard, M., Reynolds, C. R., & Gutkin, T. B. (1980). A regression analysis of test bias on the Stanford-Binet Intelligence Scale. *Journal of Clinical Child Psychology, 9,* 52–54.

Bouchard, T. J., & Segal, N. L. (1985). Environment and IQ. In B. Wolman (Ed.), *Handbook of intelligence: Theories, measurement, and applications* (pp. 391–464). New York: Wiley.

Bower, E. M. (1974). The three-pipe problem: Promotion of competent human beings through a preschool, kindergarten program and other sundry elementary matters. In G. J. Williams & S. Gordon (Eds.), *Clinical child psychology: Current practices and future perspectives*. New York: Behavioral Sciences Press.

Buros, O. K. (Ed.) (1978). *Eighth mental measurements yearbook*. Highland Park, NJ: Gryphon Press.

Burrill, L. (1975). *Statistical evidence of potential bias in items and tests assessing current educational status*. Paper presented at the annual meeting of the Southeastern Conference on Measurement in Education, New Orleans.

Burt, C. (1921). *Mental and scholastic tests*. London: King.

Camilli, G., & Shepard, L. A. (1987). The inadequacy of ANOVA for detecting test bias. *Journal of Educational Statistics, 12,* 87–99.

Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: Sage.

Cardall, C., & Coffman, W. E. (1964). *A method of comparing the performance of different groups on the items in a test* (RB-64–61). Princeton, NJ: Educational Testing Service.

Cattell, R. B. (1979). Are culture fair intelligence tests possible and necessary? *Journal of Research and Development in Education, 12,* 3–13.

Chambers, J. S., Barron, F., & Sprecher, J. W. (1980). Identifying gifted Mexican-American students. *Gifted Child Quarterly, 24,* 123–128.

Chinn, P. C. (1979). The exceptional minority child: Issues and some answers. *Exceptional Children. 46,* 532–536.

Cleary, T. A., (1968). Test bias: Prediction of grades of negro and White students in integrated universities. *Journal of Educational Measurement, 5,* 118–124.

Cleary, T. A., & Hilton, T. L. (1968). An investigation of item bias. *Educational and Psychological Measurement, 28,* 6–75.

Cleary, T. A., Humphreys, L. G., Kendrick, S. A., & Wesman, A. (1975). Educational uses of tests with disadvantaged students. *American Psychologist, 30,* 15–41.

Cole, N. S., & Moss, P. (1989). Bias in test use. In R. Linn (Ed.), *Educational measurement* (3rd ed., pp. 201–219). New York: Macmillan.

Council for Exceptional Children (1978). Minorities position policy statements. *Exceptional Children, 45,* 57–64.

Critchley, D. L. (1979). The adverse influence of psychiatric diagnostic labels on the observation of child behavior. *American Journal of Orthopsychiatry, 49,* 157–160.

Cronbach, L. J. (1970). *Essentials of psychological testing*. New York: Harper & Row.

Cronbach, L. J. (1976). Equity in selection—where psychometrics and political philosophy meet. *Journal of Educational Measurement, 13,* 31–41.

Cronbach, L. J. (1980). Validity on parole: How can we go straight? In W. B. Schraeder (Ed.), *New directions for testing and measurement: Vol 5. Measuring achievement: Progress over a decade*. San Francisco: Jossey-Bass.

Cronbach, L. J., & Drenth, P. J. D. (Eds.) (1972). *Mental tests and cultural adaption*. The Hague, Netherlands: Mouton.

Cuccaro, M. L., Wright, H. H., Rownd, C. V., & Abramson, R. (1996). Professional perceptions of children with developmental difficulties: The influence of race and socioeconomic status. *Journal of Autism and Developmental Disorders, 26,* 461–469.

Dana, R. H. (1993). *Multicultural assessment perspectives for professional psychology*. Needham Heights, MA: Allyn & Bacon.

Dana, R. H. (1996). Culturally competent assessment practices in the United States. *Journal of Personality Assessment, 66,* 472–487.

Darlington, R. B. (1978). Cultural test bias: Comments on Hunter and Schmidt. *Psychological Bulletin, 85,* 673–674.

Dean, R. S. (1977). Reliability of the WISC-R with Mexican-American children. *Journal of School Psychology, 15,* 267–268.

Dean, R. S. (1979). *WISC-R factor structure for Anglo and Hispanic children*. Paper presented at the Annual Meeting of the American Psychological Association, New York.

DeFries, J. C., Vandenberg, S. G., McClearn, G. E., Kuse,

A. R., Wilson, J. R., Ashton, G. C., & Johnson, R. C. (1974). Near identity of cognitive structure in two ethnic groups. *Science, 183,* 338–339.

Dworkin, N., & Dworkin, Y. (1979). The legacy of Pygmalion in the classroom. *Phi Delta Kappan, 61,* 712–715.

Ebel, R. L. (1979). Intelligence: A skeptical view. *Journal of Research and Development in Education, 12,* 14–21.

Eells, K., Davis, A., Havighurst, R. J., Herrick, V. E., & Tyler, R. W. (1951). *Intelligence and cultural differences: A study of cultural learning and problem-solving*. Chicago: University of Chicago Press.

Einhorn, H. J., & Bass, A. R. (1970). Methodological considerations relevant to discrimination in employment testing. *Psychological Bulletin, 75,* 261–269.

Fine, B. (1975). *The stranglehold of the IQ*. Garden City, NY: Doubleday.

Flaugher, R. L. (1978). The many definitions of test bias. *American Psychologist, 33,* 671–679.

Flaugher, R. L., & Schrader, W. B. (1978). *Eliminating differentially difficult items as an approach to test bias* (RB-78-4). Princeton, NJ: Educational Testing Service.

Flynn, J. R. (1991). *Asian-Americans: Achievement Beyond IQ*. Hillsdale, NJ: Erlbaum.

Foster, G., & Ysseldyke, J. (1976). Expectancy and halo effects as a result of artificially induced teacher bias. *Contemporary Educational Psychology, 1,* 37–45.

Frame, R. (1979, September). *Diagnoses related to school achievement, chlient's race, and socioeconomic status*. Paper presented at the Annual Meeting of the American Psychological Association, New York.

Freeman, F. N. (1923). A referendum of psychologists. *Century Illustrated Magazine, 107,* 237–245.

Gael S., & Grant, D. L. (1972). Employment test validation for minority and non-minority telephone company service representatives. *Journal of Applied Psychology, 56,* 135–139.

Goldman, R. D., & Hartig, (1976). The WISC may not be a valid predictor of school performance for primary-grade minority children. *American Journal of Mental Deficiency, 80,* 583–587.

Goldman, R. D., & Hewitt, B. N. (1976). Predicting the success of black, Chicano, Oriental, and white college students. *Journal of Educational Measurement, 13,* 107–117.

Goldstein, H., Krantz, D. L., & Rains, J. D. (Eds.) (1965). *Controversial issues in learning*. New York: Appleton-Century-Crofts.

Goolsby, T. M., & Frary, R. B. (1970). Validity of the Metropolitan Readiness Test for white and Negro students in a Southern city. *Educational and Psychological Measurement, 30,* 443–450.

Gordon, R. A. (1984). Digits backward and the Mercer–Kamin law: An empirical response to Mercer's treatment of internal validy of IQ tests. In C. R. Reynolds & R. T. Brown (Eds.), *Perspectives on bias in mental testing* (pp. 357–506). New York: Plenum.

Gordon, R. A. & Rudert, E. E. (1979). Bad news concerning IQ tests. *Sociology of Education, 52,* 174–190.

Gould, S. J. (1981). *The mismeasure of man*. New York: Norton.

Grant, D. L., & Bray, D. W. (1970). Validation of employment tests for telephone company installation and repair occupations. *Journal of Applied Psychology, 54,* 7–14.

Gross, A. L., & Su, W. (1975). Defining a "fair" or "unbiased" selection model: A question of utilities *Journal of Applied Psychology, 60,* 345–351.

Guilford Press (1997). Culturally sensitive assessment: Paying attention to cultural orientation. *Child Assessment News, 6,* 8–12.

Gulliksen, H., & Wilks, S. S. (1965). Regression tests for several samples. *Psychometrika, 15,* 91–114.

Guterman, S. S. (1979). IQ tests in research on social stratification: The cross-class validity of the test as measures on scolastic aptitude. *Sociology of Education, 52,* 163–173.

Gutkin, T. B., & Reynolds, C. R. (1980). Factorial similarity of the WISC-R for Anglos and Chicanos referred for psychological services. *Journal of School Psychology, 18,* 34–39.

Gutkin, T. B., & Reynolds, C. R. (1981). Factorial similarity of the WISC-R for White and Black children from the standardized sample. *Journal of Educational Psychology, 73,* 227–231.

Gutkin, T. B., & Tieger, A. G. (1979). Funding patterns for exceptional children: Current approaches and suggested alternatives. *Professional Psychology, 10,* 670–680.

Hambleton, R. K. (1994). Guidelines for adapting educational and psychological test: A progress report. *European Journal of Psychological Assessment, 10*(3), 229–244.

Hambleton, R. K., & Kanjee, A. (1995). Increasing the validity of cross-cultural assessments: Use of improved methods for test adaptations. *European Journal of Psychological Assessment, 11,* 147–157.

Hammill, D. (1991). *Detroit tests of learning aptitude* (3rd ed.). Austin, TX: PRO-ED.

Harman, H. (1976). *Modern factor analysis* (2nd ed.). Chicago: University of Chicago Press.

Harrington, G. M. (1975). Intelligence tests may favour the majority groups in a population. *Nature, 258,* 708–709.

Harrington, G. M. (1976, September). *Minority test bias as a psychometric artifact: The experimental evidence*. Paper presented at the Annual Meeting of the American Psychological Association, Washington, DC.

Hartlage, L. C., Lucas, T., & Godwin, A. (1976). Culturally biased and culturally—fair tests correlated with school performance in culturally disadvantaged children. *Journal of Consulting and Clinical Psychology, 32*, 325–327.

Helms, J. E. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing? *American Psychologist, 47*, 1083–1101.

Helms, J. E. (1994). The conceptualization of racial identity and other "racial" constructs. In E. Trickett, R. Watts, & D. Birman (Eds.), *Human diversity* (pp. 285–311). San Francisco: Jossey-Bass.

Henderson, N. B., Butler, B. B., & Goffeney, B. (1969). Effectiveness of the WISC and Bender–Gestalt test in predicting arithmetic and reading achievement for white and non-white children. *Journal of Clinical Psychology, 25,* 268–271.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve*. New York: Free Press.

Hickman, J. A., & Reynolds, C. R. (1986). Are race differences on mental tests an artifact of psychometric methods: A test of Harrington's experimental model. *Journal of Special Education, 20,* 409–430.

Hildreth, G. H., Griffiths, N. L., & McGauvran, M. E. (1969). *Metropolitan Readiness Tests*. New York: Harcourt-Brace Jovanovich.

Hilgard, E. R., & Bower, G. H. (1975). *Theories of learning* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Hilliard, A. G. (1979). Standardization and cultural bias as impediments to the scientific study and validation of "intelligence." *Journal of Research and Development in Education, 12*, 47–58.

Hobbs, N. R. (1975). *The futures of children*. San Francisco: Jossey-Bass.

Humphreys, L. G. (1973). Statistical definitions of test validity for minority groups. *Journal of Applied Psychology, 58*, 1–4.

Hunter, J. E. (1975, December). *A critical analysis of the use of item means and item-test correlations to determine the presence or absence of content bias in achievement test items*. Paper presented at the National Institute of

Education Conference on Test Bias, Annapolis, MD.

Hunter, J. E., & Schmidt, F. L. (1976). Critical analysis of the statistical and ethnical implications of various definitions of test bias. *Psychological Bulletin, 83*, 1053–1071.

Hunter, J. E., & Schmidt, F. L. (1978). Bias in defining test bias: Reply to Darlington. *Psychological Bulletin, 85*, 675–676.

Hunter, J. E., Schmidt, F. L., & Hunter, R. (1979). Differential validity of employment tests by race: A comprehensive review and analysis. *Psychological Bulletin, 86*, 721–735.

Hunter, J. E., Schmidt, F. L., & Rauschenberger, J. (1984). Methodological, statistical, and ethical issues in the study of bias in psychological tests. In C. R. Reynolds & R. T. Brown (Eds.), *Perspectives on bias in mental testing* (pp. 41–100). New York: Plenum.

Jackson, G. D. (1975). Another psychological view from the Association of Black Psychologists. *American Psychologist, 30*, 88–93.

Jackson, R. (1980). The scholastic aptitude test: A response to Slack and Porter's "Critical appraisal." *Harvard Educational Review, 50*, 382–391.

Jacobs, W. R. (1978). The effect of the learning disability label on classroom teachers' ability objectively to observe and interpret child behaviors. *Learning Disability Quarterly, 1,* 50–55.

James, B. J. (1995). *A test of Harrington's experimental model of ethnic bias in testing applied to a measure of emotional functioning in adolescents*. Unpublished doctoral dissertation. Texas A&M Univeristy, College Station, TX.

Jensen, A. R. (1974). How biased are culture loaded test? *Genetic Psychology Monographs, 90,* 185–224.

Jensen, A. R. (1976). Test bias and construct validity. *Phi Delta Kappan, 58,* 340–346.

Jensen, A. R. (1977). An examination of culture bias in the Wonderlic Personnel test. *Intelligence, 1,* 51–64.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Jensen, A. R. (1985). The nature of the Black–White differences on various tests: Spearman's hypothesis. *Behavioral and Brain Sciences, 8,* 193–263.

Jensen, A. R., & Figueroa, R. (1975). Forward and backward digit span interaction with race and IQ: Predictions from Jensen's theory *Journal of Education Psychology, 67,* 882–893.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika, 34,* 183.

Jöreskog, K. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 30,* 409–426.

Kaiser, S. (1986). *Ability patterns of Black and White adults on the WAIS-R independent of general intelligence and as a function of socioeconomic status*. Unpublished doctoral dissertation, Texas A&M University, College Station, TX.

Kallingal, A. (1971). The prediction of grades for black and white students at Michigan State University. *Journal of Educational Measurement, 8,* 263–265.

Kamphaus, R. W., & Reynolds, C. R. (1987). *Clinical and research applications of the K-ABC*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S. (1973). Comparison of the performance of matched groups of Black children and White children on the Wechsler Preschool and Primary Scale of Intelligence. *Journal of Consulting and Clinical Psychology, 4,* 186–191.

Kaufman, A. S. (1975). Factor analysis of the WISC-R at 11 age levels between 6½ and 16½ years. *Journal of Consulting and Clinical Psychology, 43,* 135–147.

Kaufman, A. S. (1979). *Intelligent testing with the WISC-R*. New York: Wiley-Interscience.

Kaufman, A. S., & DiCiuo, R. (1975). Separate factor analyses of the McCarthy Scales for groups of Black and White children. *Journal of School Psychology, 13,* 10–18.

Kaufman, A. S., & Hollenbeck, G. P. (1974). Comparative structure of the WPPSI for Black and Whites. *Journal of Clinical Psychology, 30,* 316–319.

Kaufman, A. S., & Kaufman, N. L. (1973). Black–White differences at age 21⁄2–81⁄2 on the McCarthy Scales of Children's Abilities. *Journal of School Psychology, 11,* 196–206.

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman assessment battery for children*. Circle Pines, MN: American Guidance Service.

Keith, T. Z., & Reynolds, C. R. (1990). Measurement and design issues in child assessment research,. In C R. Reynolds & R. W. Kamphaus (Eds.). *Handbook of psychological and educational assessment of children* (Vol. 1, pp. 29–61). New York: Guilford Press.

Kerlinger, F. N. (1973). *Foundations of behavioral research*. New York: Holt, Rinehart & Winston.

Kimble, G. A. (1961). *Hilgard and Marquis' conditioning and learning* (2nd ed.). New York: Appleton-Century-Crofts.

Kirkpatrick, J. J. (1970, September). *The psychological testing establishment: Vested interest versus responsibility*. Paper presented at the annual meeting of the American Psychological Association, Miami Beach, FL.

Kirkpatrick, J. J., Ewen, R. G., Barrett, R. S., & Katzell, R. A. (1968). *Testing and fair employment*. New York: New York University Press.

Lambert, N. M. (1979). *Adaptive behavior assessment and its implications for educational programming*. Paper presented to the Fourth Annual Midwestern Conference on Psychology in the Schools, Boys Town, NE, October 1979.

Larry, P. et al. v. Wilson Riles et al., C 71 2270 (United States District Court for the Northern District of California, October, 1979, slip opinion).

Leventhal, D. S., & Stedman, D. J. (1970). A factor analytic study of the Illinois Test of Psycholinguistic Abilities. *Journal of Clinical Psychology, 26,* 473–477.

Lewis, D. O., Balla, D. A., & Shanok, S. S. (1979). Some evidence of race bias in the diagnosis and treatment of the juvenile offender. *American Journal of Orthopsychiatry, 49,* 53–61.

Lewontin, R. C. (1970). Race and intelligence. *Bulletin of the Atomic Scientists, 26,* 2–8.

Lindsey, J. (1967). *The factorial organization of intelligence in children as related to the variables of age, sex, and subculture*. Unpublished doctoral dissertation, University of Georgia.

Linn, R. L. (1976). In search of fair selection precedures. *Journal of Educational Measurement, 13,* 53–58.

Linn, R. L. (Ed.) (1989). *Educational Measurement* (3rd ed.). New York: Macmillan.

Linn, R. L., & Werts, C. E. (1971). Considerations for studies of test bias. *Journal of Educational Measurement, 8,* 1–4.

Lippmann, W. (1923a). A judgment of the tests. *New Public, 34,* 322–323.

Lippmann, W. (1923b). Mr. Burt and the intelligence tests. *New Republic, 34,* 263–264.

Loehlin, J. C., Lindzey, G., & Spuhler, J. N. (1975). *Race differences in intelligence*. San Francisco: W. H. Freeman.

Lonner, W. J. (1985). Issues in testing and assessment in cross-cultural counseling. *The Counseling Psychologist, 13,* 599–614.

MacCorquodale, K., & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review, 55,* 95–107.

MacMillan, D. L., Jones, R. L., & Aloia, G. F. (1974). The mentally retarded label: A theoretical analysis and review of research. *American Journal of Mental Deficiency, 79,* 241–261.

Mandler, G. (1975). *Mind and emotion.* New York: Wiley.

Mason, E. J. (1979). A blessing dressed up like the plague? *The School Psychologist, 35,* 6.

Matuszek, P., & Oakland, T. (1979). Factors influencing teachers' and psychologists' recommendations regarding special class placement. *Journal of School Psychology, 17,* 116–125.

Mayfield, J. W., & Reynolds, C. R. (in press). Are ethnic differences in diagnosis of childhood psychopathology an artifact of psychometric methods? An experimental evaluation of Harrington's hypothesis using parent reported symptomotology. *Journal of School Psychology.*

McCarthy, D. (1972). *McCarthy Scales of Children's Abilities.* San Antonio, TX: The Psychological Corporation.

McCoy, S. A. (1976). Clinical judgments of normal childhood behaviors. *Journal of Consulting and Clinical Psychology, 44,* 710–714.

McGurk, F. C. (1951). *Comparison of the performance of negro and White high school seniors on cultural and noncultural psychological test questions.* Washington, DC: Catholic University of America Press.

McNemar, Q. (1975). On so-called test bias. *American Psychologist, 30,* 848–851.

Mercer, J. R. (1971). The meaning of mental retardation. In R. Koch & J. Dobson (Eds.), *The mentally retarded child and his family: A multidisciplinary handbook.* New York: Brunner/Mazel.

Mercer, J. R. (1976). *Cultural diversity, mental retardation, and assessment: The case for nonlabeling.* Paper presented to the Fourth International Congress of the International Association for the Scientific Study of Mental Retardation, Washington, DC, August, 1976.

Mercer, J. R. (1979). *System of multicultural pluralistic assessment: Conceptual and technical manual.* San Antonio, TX: The Psychologcial Corporation.

Merz, W. R. (1970). *A factor analysis of the Goodenough–Harris drawing test across four ethnic groups.* Unpublished doctoral dissertation, University of New Mexico.

Meyers, C. E., MacMillan, D. L., & Yoshida, R. K. (1978). Validity of psychologists' identification of EMR students in the perspective of the California decertification experience. *Journal of School Psychology, 16,* 3–15.

Mille, F. (1979). Cultural bias in the WISC. *Intelligence, 3,* 149–164.

Mitchell, B. C. (1967). Predictive validity of the Metropolitan Readiness Tests and the Murphy-Durrell Reading Analysis for white and for Negro pupils. *Educational and Psychological Measurement, 27,* 1047–1054.

Moran, M. P. (1990). The problem of cultural bias in personality assessment. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children. Vol. 2. Personality, behavior and context* (pp. 524–545). New York: Guilford Press.

Moreland, K. L. (1996). Persistent issues in multicultural assessment of social and emotional functioning. In L. A. Suzuki, P. J. Meller, & J. G. Ponterrotto (Eds.), *Handbook of multicultural assessement: Clinical, psychological, and educational applications.* San Francisco: Jossey-Bass.

Nalven, F. B., Hoffman, L. J., & Bierbryer, B. (1969). The effects of subject's age, sex, race, and socioeconomic status on psychologists' estimates of "True IQ" from WISC scores. *Journal of Clinical Psychology, 25,* 271–274.

Nichols, R. C. (1978). Policy implications of the IQ controversy. In L. S. Schulman (Ed.), *Review of research in education* (Vol. 6). Itasca, IL: F. E. Peacock.

Novick, M. R., & Petersen, N. S. (1976). Towards equalizing educational and employment opportunity. *Journal of Educational Measurement, 13,* 77–88.

Oakland, T. (1978). Predictive validity of readiness tests for middle and lower socioeconomic status Anglo, black and Mexican American children. *Journal of Educational Psychology, 70, 574–582.*

Oakland, T., & Feigenbaum, D. (1979). Multiple sources of test bias on the WISC-R and Bender–Gestalt Test. *Journal of Consulting and Clinical Psychology, 47,* 968–974.

Oakland, T., & Matuszek, P. (1977). Using tests of nondiscriminatory assessment. In T. Oakland (Ed.), *Psychological and educational assessment of minority group children.* New York: Brunner/Mazel.

O'Conner, E. J., Wexley, K. N., & Alexander, R. A. (1975). Single group validity: Fact or fallacy. *Journal of Applied Psychology, 60,* 352–355.

Padilla, A. M. (1988). Early psychological assessment of Mexican-American children. *Journal of the History of the Behavioral Sciences, 24,* 113–115.

Padilla, A. M., & Medina, A. (1996). Cross-cultural sensitivity in assessment: Using tests in culturally appropriate ways. In L. A. Suzuki, P. J. Miller, & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment: Clinical psychological, and educational applications* (pp. 3–28). San Francisco: Jossey-Bass.

PASE: Parents in action on special education et al. v. Hannon et al. No. 74 C 3586 (United States District Court for the Northern District of Illinois, Eastern Division, July, 1980, slip opinion).

Petersen, N. S., & Novick, M. R. (1976). An evaluation of some models for culture—fair selection. *Journal of Educational Measurement, 13,* 3–29.

Pfeifer, C. M., & Sedlacek, W. E. (1971). The validity of academic predictors for black and white students at a predominately white university. *Journal of Educational Measurement, 8,* 253–261.

Pintner, R. (1931). *Intelligence testing.* New York: Holt, Rinehart & Winston.

Plake, B., & Hoover, H. (1979, September). *A methodology for identifying biased achievement test items that removes the confounding in a items by groups interaction due to possible group differences in instructional level.* Paper presented at the annual meeting of the American Educational Research Association, Toronto.

Potthoff, R. F. (1966). *Statistical aspects of the problem of biases in psychological tests.* (Institute of Statistics Mimeo Series No. 479). Chapel Hill: University of North Carolina, Department of Statistics.

Ramsey, R. T. (1979). *The testing manual: A guide to test administration and use.* Pittsburgh: Author.

Reschly, D. J. (1978). WISC-R factor stuctures among Anglos, blacks, Chicanos, and Native American Papagos. *Journal of Consulting and Clinical Psychology, 46,* 417–422.

Reschly, D. J., & Reschly, J. E. (1979). Validity of the WISC-R factor scores in predicting achievement and attention for four sociocultural groups. *Journal of School Psychology, 17,* 355–361.

Reschly, D. J., & Sabers, D. (1979). Analysis of test bias in four groups with the regression definition. *Journal of Educational Measurement, 16,* 1–9.

Reynolds, C. R. (1978). *Differential validity of several preschool assessment instruments for Blacks, Whites, males, and females.* Unpublished doctoral dissertation, University of Georgia.

Reynolds, C. R. (1979a). Should we screen preschoolers? *Contemporary Educational Psychology, 4,* 175–181.

Reynolds, C. R. (1979b). A factor analytic study of the Metropolitan Readiness Test. *Contemporary Educational Psychology, 4,* 315–317.

Reynolds, C. R. (1979c). The invariance of the factorial validity of the Metropolitan Readiness Tests for blacks, whites, males and females. *Educational and Psychological Measurement, 39,* 1047–1052.

Reynolds, C. R. (1980a). Differential construct validity of intelligence as popularly measured: Correlations of age with raw scores on the WISC-R for Blacks, White, males and females. *Intelligence, 4,* 371–379.

Reynolds, C. R. (1980b). An examination for test bias in a preschool battery across race and sex. *Journal of Educational Measurement, 17,* 137–146.

Reynolds, C. R. (1980c). In support of "Bias in Mental Testing" and scientific inquiry. *The Behavioral and Brain Sciences, 3,* 352.

Reynolds, C. R. (1982a). The problem of bias in psychological assessment. In C. R. Reynolds & T. B. Gutkin (Eds.), *The handbook of school psychology*. New York: Wiley.

Reynolds, C. R. (1982b). Construct and predictive bias. In R. A. Berk (Ed.), *Handbook of methods for detecting test bias* (pp. 199–227). Baltimore: Johns Hopkins University Press.

Reynolds, C. R. (1990). Conceptual and technical problems in learning disability diagnosis. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children* (Vol. 1, pp. 571–592). New York: Guilford Press.

Reynolds, C. R. (1995). Test bias and the assessment of intelligence and personality. In D. Saklofske & M. Zeidner (Eds.), *International handbook of personality and intelligence* (pp. 545–573). New York: Plenum.

Reynolds, C. R. (1997a). Measurement and statistical problems in neuropsychological assessment of children. In C. R. Reynolds & E. Fletcher-Janzen (Eds.), *Handbook of clinical child neuropsychology* (2nd ed., pp. 180–203). New York: Plenum.

Reynolds, C. R. (1997b). Forward and backward memory span should not be combined for clinical analysis. *Archives of Clinical Neuropsychology, 12,* 29–40.

Reynolds, C. R. (in press). Need we measure anxiety differently for males and females. *Journal of Personality Assessment*.

Reynolds, C. R., & Bigler, E. D. (1994). *Test of memory and learning*. Austin, TX: PRO-ED.

Reynolds, C. R., & Brown, R. T. (1984). Bias in mental testing: An introduction to the issues. In C. R. Reynolds & R. T. Brown (Eds.), *Perspectives on bias in mental testing* (pp. 1–39). New York: Plenum.

Reynolds, C. R., & Gutkin, T. B. (1980a). A regression analysis of test bias on the WISC-R for Anglos and Chicanos referred for psychological services. *Journal of Abnormal Child Psychology, 8,* 237–243.

Reynolds, C. R., & Gutkin, T. B. (1980b, September). *WISC-R performance of Blacks and Whites matched on four demographic variables*. Paper presented at the Annual Meeting of the American Psychological Association, Montreal.

Reynolds, C. R., & Gutkin, T. B. (1980c). Stabillity of the WISC-R factor structure across sex at two age levels. *Journal of Clinical Psychology, 36,* 775–777.

Reynolds, C. R., & Gutkin, T. B. (1981). A multivariate comparison of the intellectual performance of Black and White children matched on four demographic variables. *Personality and Individual Differences, 2,* 175–180.

Reynolds, C. R., Gutkin, T. B., Dappen, L., & Wright, D. (1979). Differential validity of the WISC-R for boys and girls referred for psychological services. *Perceptual and Motor Skills*, *48,* 868–870.

Reynolds, C. R., & Harding, R. E. (1983). Outcome in two large sample studies of factorial similarity under six methods of comparison. *Educational and Psychological Measurement*, *43,* 723–728.

Reynolds, C. R., & Hartlage, L. C. (1978, March). *Comparison of WISC and WISC-R racial regression lines*. Paper presented at the annual meeting of the Southeastern Psychological Association, Atlanta.

Reynolds, C. R., & Hartlage, L. C. (1979). Comparison of WISC and WISC-R regression lines for academic prediction with Black and with White referred children. *Journal of Consulting and Clinical Psychology, 47,* 589–591.

Reynolds, C. R., & Jensen, A. R. (1983). *Patterns of intellectual performance among Blacks an Whites matched on "g."* Paper presented to the Annual Meeting of the American Psychological Association, Montreal, September, 1983.

Reynolds, C. R., & Kaiser, S. (1990). Test bias in psychological assessment. In T. B. Gutkin & C. R. Reynolds (Eds.), *The handbook of school psychology* (2nd ed., pp. 487–525) New York: Wiley.

Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior assessment system for children: Manual*. Circle Pines, MN: American Guidance Service.

Reynolds, C. R., & Paget, K. (1981). Factor analysis of the Revised Children's Manifest Anxiety Scale for Blacks, Whites, males, and females with a national normative sample. *Journal of Consulting and Clinical Psychology, 49,* 352–359.

Reynolds, C. R., Willson, V. L., & Chatman, S. P. (1984). Item bias on the 1981 revisions of the Peabody Picture Vocabulary Test using a new method of detecting bias. *Journal of Psychoeducational Assessment, 2,* 219–224.

Reynolds, C. R., Willson, V. L., & Chatman, S. P. (1985). Regression analyses of bias on the K-ABC. *Journal of School Psychology, 23,* 195–204.

Ross, A. O. (1974). A clinical child psychologist "examines" retarded children. In G. J. Williams & S. Gordon (Eds.), *Clinical child psychology: Current trends and future perspectives*. New York: Behavioral Publications.

Ross, A. O. (1976). *Psychological aspects of learning disorders of children*. New York: McGraw-Hill.

Samuda, A. J. (1975). *Psychological testing of American minorities: Issues and consequences*. New York: Dodd, Mead.

Sandoval, J. (1979). The WISC-R and internal evidence of test bias with minority groups. *Journal of Consulting and Clinical Psychology, 47,* 919–927.

Sandoval, J., & Mille, M. (1980). Accuracy judgments of WISC-R item difficulty for minority groups. *Journal of Consulting and Clinical Psychology, 48,* 249–253.

Sawyer, R. L., Cole, N. S., & Cole, J. W. (1976). Utilities and the issue of fairness in a decision theoretic model for selection. *Journal of Educational Measurement, 13,* 59–76.

Scheuneman, J. D. (1987). An experimental, exploratory study of causes of bias in test items. *Journal of Educational Measurement, 24,* 97–118.

Schmidt, W. H. (1983). Content biases in achievement tests. *Journal of Educational Measurement, 20, 165–178.*

Schmidt, F. L., & Hunter, J. E. (1974). Racial and ethnic bias in psychological tests: Divergent implications of two definitions of test bias. *American Psychologist, 29,* 1–8.

Schoenfeld, W. N. (1974). Notes on a bit of psychological nonsense: "Race differences in intelligence." *Psychological Record, 24,* 17–32.

Scholwinski, E. (1985). *Ability patterns of White and Black adults as determined by the subscales of the WAIS-R.* Unpublished doctoral dissertation, Texas A&M University, College Station, TX.

Semler, I., & Iscoe, I. (1966). Structure of intelligence in Negro and White children. *Journal of Educational Psychology, 57,* 326–336.

Sewell, T. E. (1979). Intelligence and learning tasks as predictors of Scholastic achievement in Black and White first–grade children. *Journal of School Psychology, 17,* 325–332.

Shepard, L. S. (1982). Definitions of bias. In R. A. Berk (Ed.), *Handbook of methods for detecting test bias.*

Baltimore: Johns Hopkins University Press.

Shuey, A. M. (1966). *The Testing of Negro intelligence* (2nd ed.). New York: Social Science Press.

Silverstein, A. R. (1973). Factor structure of the Wechsler Intelligence Scale for Children for three ethnic groups. *Journal of Educational Psychology, 65,* 408–410.

Stanley, J. C. (1971). Predicting college success of the educationally disadvantaged. *Science, 171,* 640–647.

Stanley, J. C., & Porter, A. C. (1967). Correlation of scholastic aptitude test scores with college grades for Negroes vs. whites. *Journal of Educational Measurement, 4,* 199–218.

Stern, W. (1914). *The psychological methods of testing intelligence.* Baltimore: Warwick & York.

Sternberg, R. J. (1980). Intelligence and test bias: Art and science. *Behavioral and Brain Sciences, 3,* 353–354.

Stricker, L. J. (1982). Identifying test items that perform differentially in population subgroups: A partial correlation index. *Applied Psychological Measurement, 6,* 261–273.

Temp, G. (1971). Validity of the SAT for blacks and whites in thirteen integrated institutions. *Journal of Educational Measurements, 8,* 245–251.

Tenopyr, M. L. (1967, September). *Race and socioeconomic status as moderators in predicting machine-shop training success.* Paper presented at the annual meeting of the American Psychological Association, Washington, DC.

Thorndike, R. L. (1971). Concepts of culture-fairness. *Journal of Educational Measurement, 8,* 63–70.

Tittle, C. K. (1975). Fairness in educational achievement testing. *Education and Urban Society, 8,* 86–103.

Tittle, C. K. (1982). Use of judgemental methods in item bias studies. In R. A. Berk (Ed.), *Handbook of methods for detecting test bias.* Baltimore: Johns Hopkins University Press.

Tomlinson, J. R., Acker, N., Canter, A., & Lindborg, S. (1977). Minority status, sex, and school psychological services. *Psychology in the Schools, 14,* 456–460.

Torrance, E. P. (1980). Psychology of gifted children and youth. In W. M. Cruickshank (Ed.), *Psychology of exceptional children and youth.* Englewood Cliffs, NJ: Prentice-Hall.

Tyler, L. E. (1965). *Psychology of human differences.* New York: Appleton-Century-Crofts.

Van de Vijver, F., & Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist, 1,* 89–99.

Vance, H. B., & Wallbrown, F. H. (1978). The structure of intelligence for Black children: A hierarchical approach. *Psychological Record, 28,* 31–39.

Vance, H. B., Huelsman, C. B., & Wherry, R. J. (1976). The hierarchical factor structure of the Wechsler Intelligence Scale for Children as it relates to disadvantaged White and Black children. *Journal of General Psychology, 95,* 287–293.

Veale, J. R., & Foreman, D. I. (1983). Assessing cultural bias using foil response data: Cultural variation. *Journal of Educational Measurement, 20,* 249–258.

Wallbrown, F. H., Blaha, J., & Wherry, R. J. (1973). The hierarchical factor structure of the Wechsler Adult Intelligence Scale. *British Journal of Educational Psychology, 44,* 47–56.

Wechsler, D. (1949). *Wechsler Intelligence Scale for Children.* New York: Psychological Corporation.

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children-Revised.* New York: The Psychological Corporation.

Wechsler, D. (1975). Intellegence defined and undefined: A relativistic appraisal. *American Psychologist, 30,* 135–139.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised.* San Antonio, TX: The Psychological Corporation.

Wherry, R. J., & Wherry, R. J., Jr. (1969). WNEWH program. In R. J. Wherry (Ed.), *Psychology department computer program.* Columbus, OH: Ohio State University.

Willerman, L. (1979). *The psychology of individual and group differences.* San Francisco: W. H. Freeman.

Williams, R. L. (1970). Danger: Testing and dehumanizing Black children. *Clinical Child Psychology Newsletter, 9,* 5–6.

Williams, R. L. (1972, September). *The BITCH-100: A culture specific test.* Paper presented at the annual meeting of the American Psychological Association, Honolulu.

Williams, R. L. (1974). From dehumanization to black intellectual genocide: A rejoinder. In G. J. Williams & S. Gordon (Eds.), *Clinical child psychology: Current practices and future perspectives.* New York: Behavioral Publications.

Williams, R. L., Dotson, W., Dow, P., & Williams, W. S. (1980). The war against testing: A current status report. *Journal of Negro Education, 49,* 263–273.

Willson, V. L., Nolan, R. F., Reynolds, C. R., & Kamphaus, R. W. (1989). Race and gender effects on item functioning on the Kaufman Assessment Battery for Children. *Journal of School Psychology, 27,* 289–296.

Wilson, K. M. (1969). *Black students entering CRC colleges: Their characteristics and their first year academic performance* (Research memo No. 69-1). Poughkeepsie, NY: College Research Center.

Wright, B. J., & Isenstein, V. R. (1977). Psychological tests and minorities. Rockville, MD: NIMH, DHEW Publication No. (ADM) 78–482 (reprinted 1978).

Zigler, E., & Butterfield, E. C. (1967). Motivational aspects of changes in IQ test performance of culturally deprived nursery school children. *Child Development, 39,* 1–14.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.04
Cross-cultural Clinical Interventions

JOHN E. LEWIS
Nova Southeastern University, Fort Lauderdale, FL, USA

10.04.1 INTRODUCTION 94

10.04.2 PROXIMAL VS. DISTAL VARIABLES 95

- *10.04.2.1 Proximal Variables* 95
- *10.04.2.2 Distal Variables* 96

10.04.3 CLINICAL INTERVENTION PROCESSES 97

- *10.04.3.1 Cross-cultural Competence and Sensitivity* 97
- *10.04.3.2 Intercultural Communication Style* 98
- *10.04.3.3 Cross-cultural Verbal Communication* 100
- *10.04.3.4 Nonverbal Communication* 100
  - *10.04.3.4.1 Proxemics* 100
  - *10.04.3.4.2 Kinesics* 101
  - *10.04.3.4.3 Paralinguistics* 103
- *10.04.3.5 Use of Translators* 103

10.04.4 THERAPIST–CLIENT MATCH 104

10.04.5 TREATMENT OUTCOME 107

- *10.04.5.1 Definition* 107
- *10.04.5.2 Cross-cultural Outcome Research* 107

10.04.6 METHODOLOGICAL CONSIDERATIONS 109

- *10.04.6.1 Reliability and Validity in Cross-cultural Research* 109
- *10.04.6.2 Recommendations* 110

10.04.7 FUTURE DIRECTIONS 111

- *10.04.7.1 Terminology* 111
- *10.04.7.2 Ethical and Professional Issues* 112
- *10.04.7.3 Future Research* 112
  - *10.04.7.3.1 Research agenda* 113
  - *10.04.7.3.2 Assessment techniques* 114
  - *10.04.7.3.3 Theory of multicultural counseling and therapy* 114
  - *10.04.7.3.4 Specific research topics* 114
- *10.04.7.4 Global Psychology* 115

10.04.8 SUMMARY 116

10.04.9 REFERENCES 116

10.04.1 INTRODUCTION

While attempts have been made to trace the evolution of multicultural counseling from ancient civilizations to the present day (Thompson, 1989), current thinking supports the notion that multicultural counseling is only several decades old. This is evidenced by the increased intercultural focus in publications from 1960 to 1995 (Jackson, 1995).

Early in the 1990s multicultural counseling and therapy (MCT) was hailed as the fourth force in psychology (Pedersen, 1990, 1991b, 1994; Pedersen & Ivey, 1993). As a result, the movement has been to formulate a framework for MCT in an attempt to describe and explain the impact of cultural diversity on psychotherapeutic interventions. This movement spawned a definitive explication of MCT theory in a conceptual treatise (Sue, Ivey, & Pedersen, 1996). The work provides a metatheory for cross-cultural interventions that espouses six basic propositions, encompassing 47 corollary suppositions.

The propositions and corollaries were originally proposed by Sue (1995), but were refined by Sue et al. (1996) to help psychologists move from a culture-bound theory of intervention to a theory that could encompass a variety of existing world views. The propositions and corollaries acknowledge that (i) both therapist and client identities are formed and embedded within a complex set of experiences that have a fluid dimension and a contextual salience, (ii) the development of one's cultural identity encompasses attitudes to one's own culture and other cultures; (iii) psychotherapeutic outcomes are more successful when goals and therapy modalities are consistent with a client's values, goals, and cultural experiences; and (iv) individual counseling may need to be balanced with models that incorporate family and societal units.

This new framework spawned critical debate (Casas & Mann, 1996; Corey, 1996) centered on the lack of adequate definitions within the field of MCT. The terms "race," "ethnicity," "minority," and "culture" have not been clarified and, thus, are constantly confused by professionals (Atkinson, Morten, & Sue, 1993). In 1994, Pedersen defined these terms. He stated that race is a classification based on genetic or biological differences and includes a shared genetic history and/or physical characteristics such as skin color; ethnicity refers to a common sociocultural history that can include political history, similar genealogy, and religion; minority is used to identify a group that has been treated unfairly or has suffered from collective discriminating practices due to the minorities' lack of power in the dominant culture; finally, culture has taken on broader implications in the literature of the 1990s to include demographic variables (e.g., age, gender, and region), status variables, affiliation variables, and the traditional variables of race and ethnicity. This view of culture focuses on its nonstatic nature, as culture also denotes patterns of behavior, family life, and beliefs (Fernando, 1995). Poortinga (1992) has argued that culture constitutes a set of shared constraints that limit individual members of groups in terms of their behavioral responses.

Goodchilds (1991) described race and ethnicity as dimensions not categories. Phinney (1996) argued that one needs to explore three dimensions that differ between and within ethnic groups: "First, cultural norms and attitudes that may be influential in psychological processes. ... Second, the strength, salience, and meaning of individuals' ethnic identities. ... Third, individuals' experiences as a member of a minority group with lower status and power" (p. 925). Readers are diverted to more complete explanations of the terminology concerning culture, race, and ethnicity (Gaw, 1993; Jackson, 1995; Marsella & Pedersen, 1981; Marsella & White, 1982; Pedersen, 1991b; Phinney, 1996).

This new emphasis has spawned concern about the terminology and definitions used by researchers and practitioners alike. In particular, current terms such as "multicultural counseling," "culture-centered counseling," cross-cultural intervention," "cross-cultural communication," "minority counseling," and "pluralism" have tended to obfuscate an already complex area of interest.

Pedersen (1991b) has suggested that psychologists use a broad definition of culture that includes demographic variables such as economics and education as well as informal and formal affiliation variables to supplement the traditional cultural variables of ethnicity, race, nationality, religion, and linguistic group. This wide view of culture seeks to balance culture-specific with universal perspectives that underlie a client's behavior, cognition, and affect. Pedersen (1991a) has also called for the term "multicultural counseling and therapy" to be used since it "implies a wide range of multiple groups without grading, comparing, or ranking them as better or worse than one another and without denying the very distinct and complementary perspective that each group brings with it" (p. 4).

Keeping in the forefront the concepts and definitions integral to MCT, this chapter will focus on issues pertaining to cross-cultural interventions. First, proximal and distal factors will be described. Second, interviewing and clinical interventions will be examined, by summarizing the research on effective intercultural communication, and verbal/nonverbal language factors in counseling and psychotherapy. Third, the issue of client and therapist match will be explored. This discussion will evaluate the notion that therapists and clients from similar backgrounds have more accurate communication and greater potential for successful therapeutic outcomes. Fourth, the treatment outcome literature will be examined with a view to evaluating the efficacy of particular intervention strategies with culturally different clients. Fifth, present and future methodological issues that relate to the clinical interventions with a culturally diverse clientele will be examined. Finally, future issues in this area will be discussed.

10.04.2 PROXIMAL VS. DISTAL VARIABLES

Two prerequisites of culturally responsive behavior are attitudes and beliefs (Ridley, Mendoza, Kanitz, Angermeier, & Zenk, 1994). These two constructs can be discussed by using Poortinga's (1992) notion that behavioral responses are dependent upon internal and external constraints which are governed by a person's sociocultural history. These constraints have been delineated into two variables: internal–proximal and external–distal. Proximal variables, also called proximal levels, include aspects of the immediate setting and behavior of the client, and the "related flux of affect and cognition during a particular interaction" (Fletcher & Fitness, 1990, p. 464). Distal variables, or levels, are generally "stable dispositional variables that predate the immediate proximal context" (Fletcher & Fitness, 1990, p. 464). Distal variables include personality characteristics, levels of effect, and client expectations. They also include ethnicity and race (Sue, Zane, & Young, 1994).

10.04.2.1 Proximal Variables

The effect of proximal variables on relationships and communication processes has emerged in studies originating in the field of social cognition (Bradbury & Fincham, 1988; Fletcher & Fitness, 1990; Kelly et al., 1983). The social cognition notion of proximal level can be expanded into cross-cultural interactions. The therapist must identify the thoughts and feelings experienced by the client immediately preceding a therapy session, especially values and beliefs (Bond & Smith, 1996).

The study of values within a cross-cultural framework is best exemplified in the work of Kluckhohn and Strodtbeck (1961). This model recognized that cultures differ in their perception of time, attitudes towards activity, their lineal versus collateral relationships, and their notion of the nature of people. These value orientations have been modified and expanded by Ibrahim, (1985) and Kohls (1979) to provide social/cultural interpretations and examples. Carter (1991) reviewed the research on the Kluckhohn–Strodtbeck model and concluded that a counselor needed to understand a culture's value structure if he or she hoped to understand both the unique and universal attributes of that culture.

Schwartz (1992, 1994) has examined the universal content and structure of values. He identified, through the results of a self-report values instrument, 10 universal value domains at the individual level and seven at the global cultural level. Scores were obtained for 38 different cultural groups (Schwartz, 1994). These culture-level scores have been compared to the results obtained in Hofstede's studies. Hofstede (1980, 1991) described a pattern of universal and culture-specific variables in 55 countries. He identified measures that related to thinking, feeling, and behaving and claimed that these patterns, or mental programs, were constructed differently by different cultures at different times. In addition, he described four dimensions that varied among the countries: small/large power distance, collectivism vs. individualism, femininity vs. masculinity, and weak/strong uncertainty avoidances. More recently, researchers (Kim, Triandis, Kagitcibasi, Choi, & Yoon, 1994; Smith, Peterson, Akande, Callan, & Cho, 1995) have examined specific value orientations in different countries and have illustrated clear patterns of cultural difference with respect to external–internal control and collectivism, respectively.

Whereas values exemplify what a culture or individual holds as important, beliefs encompass an individual's or culture's sense of how the world is construed. The most often studied variable in this area is locus of control. This concept was formulated by Rotter (1966) and incorporated the notion that people have differential beliefs about how outcomes are obtained in relation to personal power. Rotter described two dimensions, internal control and external control. Internal control referred to the idea that people can control their own outcomes. External control referred to the notion that outcomes are more determined by chance environmental circumstances rather than personal power. Through the socialization process, people learn one of these two views. These locus-of-control beliefs of clients are related to how a client experiences "control harmony, and submission with respect to the environment" (Bond & Smith, 1996, p. 211).

Locus of responsibility is closely related to locus of control. This belief, which has emerged largely from attribution theories in social cognition (Fiske & Taylor, 1991), is concerned with how much blame is placed on individuals or systems. Sue and Sue (1990) have described these emphases as important for a client's life orientation development. They believe that a person-oriented, person-blame individual tends to focus on individual responsibility attributes. He or she tends to focus on a person's motivations, values, feelings, and goals and tends to attribute successful or unsuccessful outcomes to a person's strengths and weaknesses. These individuals see a strong relationship between individual effort and success. This stance allows some members of the dominant culture to view minorities as totally responsible for their behavior and ignore contextually relevant information that may contribute to or cause a person's behavior. Some clients, on the other hand, tend to be situation-blaming. They tend to attribute successful and unsuccessful outcomes to variables outside of themselves, in the sociocultural environment. Sue and Sue believe that therapists need to view the individual–system blame continuum in an effort to understand their clients and to see the culturally relevant factors that may impact on a person's situation.

Other beliefs that are not universal, but culturally specific, have been suggested. They include the notion of justice (Furnham, 1993), global cooperation (Der-Kerabetian, 1992), and beliefs and attitudes about the work ethic (Furnham, Bond, Heaven, Hilton, & Lobel, 1993). Further research is needed to clarify and verify these variables.

10.04.2.2 Distal Variables

There has been a notable lack of empirical studies that have focused on external–distal factors. Bond and Smith (1996) have concluded that "almost all current models of cultural difference are thus proximal rather than distal" (p. 211). Sue and Zane (1987) discussed the proximal–distal issue by evaluating whether cultural knowledge or culture-specific techniques were linked to psychotherapeutic success. They formulated a proximal–distal model that suggested cultural knowledge and technique-oriented methods in therapy are of a distal nature and, therefore, may not contribute significantly to effective counseling and psychotherapy. Rather, they believe that "it may be wiser to focus on the proximal process of therapist credibility" (p. 39). This credibility involves a therapist's ability to conceptualize the client's problem, sensitivity to culturally appropriate means of problem resolution, and client–therapist goal compatibility. In other words, high achieved and ascribed credibility is likely to produce an effective therapeutic process, whereas low ascribed and achieved credibility will result in early termination of therapy.

Akutsu, Lin, and Zane (1990) stated that the proximal–distal model predicts that counselor credibility is related to continued involvement in treatment, whereas the distal variables, cultural knowledge, and techniques are mediated through the credibility that exists in the therapeutic relationship. They tested the proximal–distal model with Chinese students and concluded that counselor credibility was related to the client's intention to utilize therapy and that relationship variables and therapist style were mediated through the subject's evaluation of the counselor's credibility. Proximal factors, therefore, were seen as essential for continued intentions for therapy by culturally diverse clients.

Berry (1990) has studied distal factors with respect to ethnic group relationships. He identified a map of relationships that included three sociocultural factors: collective stereotypes, history–economics–politics, and collective behaviors. These categories influenced various psychological variables, including perception, attribution, attitudes, and behavior. Fiske (1992) has described sociocultural factors by proposing a four-component model, stating that social relationships in cultures include communal sharing, authority, ranking, equality matching, and market pricing. Berry (1990) and Triandis (1994) have called for more studies that are sensitive to external, distal ethnic variables.

The preceding discussion has encompassed the notion that universal and culture-specific thoughts, feelings, and behaviors affect client–therapist interaction. It was Frank (1961) that stated a "psychotherapist in any culture works to assist in the definition of subtle gradations of implicit rules governing behavior which derive from a vast array of norms, value systems, and ideologies present in the shared assumptive worlds of the patient and therapist" (p. 28). Therapists with culturally diverse clients are required to decode implicit rules that are culturally determined and to make them more situationally appropriate in therapy. It necessitates a therapist's close attention to what individual and sociocultural variables are salient at any given moment during the cross-cultural interaction. Failure to do so will result in cultural encapsulation, ineffective client–therapist relationships, and possible ineffective treatment outcomes.

The issues involved with distal and proximal variables are at the core of interviewing skills and clinical interventions. The act of interviewing and therapy involves the formation of a close relationship between client and therapist. This relationship, inevitably, exposes the proximal and distal factors present in the lives of both participants. Sensitivity to proximal and distal factors, as well as culture-specific and universal expressions of behavior from a client, are fundamental concerns in clinical interventions.

10.04.3 CLINICAL INTERVENTION PROCESSES

The area of clinical intervention process encompasses topics as diverse as cross-cultural sensitivity and competence, intercultural communication style, cross-cultural verbal communication, nonverbal communication, and the role of translators during therapy.

10.04.3.1 Cross-cultural Competence and Sensitivity

Cross-cultural counseling has been defined as "any counseling relationship in which two or more participants differ in cultural background, values, and lifestyle" (Sue et al., 1982, p. 47). This definition has led many counselors and therapists to the conclusion that *all* counseling could be considered to have a multicultural dimension (Pedersen, 1994; Speight, Myers, Cox, & Highlen, 1991; Sue, Ivey, & Pedersen, 1996). In fact, it is argued that within-group differences have at least equal impact as between-group differences on the clarification of cultural issues (Speight et al., 1991).

These definitions and statements have demanded the attention of organizations and policy makers concerned with the provision of psychological services. In 1992, the Association for Multicultural Counseling and Development (AMCD), a branch of the American Counseling Association (ACA), published a document that outlined multicultural counseling competencies and standards for inclusion in psychological service provider programs. This document espoused a three-by-three matrix for organizing and developing cross-cultural competencies. This matrix identified three areas of concern: (i) beliefs and attitudes, (ii) knowledge, and (iii) skills in relation to counselor self-awareness, counselor understanding of client's world view, and the development of culturally appropriate intervention strategies and techniques. Of particular relevance were the recommendations for interventions, which included the need for more accurate verbal communication, increased understanding of nonverbal communication, greater proficiency in dealing with clients' diverse behavioral and affective response styles, and increased understanding of the role of societal and familial issues in therapy.

In 1993, the American Psychological Association (APA) established guidelines for the provision of services to ethnic, linguistic, and culturally diverse populations. These guidelines stated: "Psychologists need a sociocultural framework to consider diversity of values, interactional styles, and cultural expectations in a systematic fashion" (p. 45). In response to this need for frameworks for effective cross-cultural interactions, psychologists have examined cultural differences and universal expressions of client behavior (Sue & Sue, 1990).

Culture tends to have a limiting or restructuring effect on universal expression. These cultural differences create stress on the therapeutic relationship. Problems can occur when the therapist becomes too enamored of his or her own cultural values. Wrenn (1962, 1985) created the term "cultural encapsulation" to describe the notion that some therapists: (i) substitute stereotypes for the real world, (ii) ignore cultural variations across cultures, and (iii) adhere to a technique-oriented, static notion of the therapeutic process. The result of cultural encapsulation is that the therapist's role becomes rigidly defined with an ensuing universal notion of health, normality, and pathology.

This notion of cultural encapsulation led Sue and Sue (1990) to outline the characteristics of the culturally skilled counselor and therapist. These characteristics involve the development of attitudes and beliefs that have understanding of the self and the client's world view at their core. The culturally skilled counselor (pp. 167–169):

(i) has moved from being culturally unaware to being aware and sensitive to his or her own cultural heritage and to valuing and respecting differences;

(ii) is aware of his or her own values and biases and how they might affect minority clients;

(iii) is comfortable with differences that exist between themselves and their clients in terms of race and beliefs;

(iv) is sensitive to circumstances (personal biases, stage of ethnic identity, sociopolitical influences, etc.) that may dictate referral of a minority client to a member of his or her own race or culture or to another counselor in general;

(v) is aware of his or her own racial attitudes, beliefs, and feelings;

(vi) must possess specific knowledge and information about the particular group he or she is working with;

(vii) must have a good understanding of the sociopolitical systems' operation with respect to its treatment of minorities;

(viii) has clear and explicit knowledge of the generic characteristics of counseling and therapy;

(ix) is aware of the institutional barriers that prevent minorities from using mental health services.

These attributes of the skilled multicultural therapist have been discussed more elaborately in the cross-cultural competencies and standards outlined by Sue, Arredondo, and McDavis (1992). Attempts to measure these competencies have been made with two rating scales: the Cross-cultural Competence Inventory (Hernandez & LaFromboise, 1985) and the Cross-cultural Competence Inventory–Revised (LaFromboise, Coleman, & Hernandez, 1991) and three self-report measures: the Multicultural Awareness–Knowledge–Skills Survey (D'Andrea, Daniels, & Heck, 1991), the Multicultural Counseling Awareness Scale (Ponterotto, Sanchez, & Magids, 1991), and the Multicultural Counseling Inventory (Sodowski, Taffe, Gutkin, & Wise, 1994).

Cultural sensitivity has been defined in various ways: cross-cultural competence, cross-cultural expertise, cross-cultural effectiveness, culture responsiveness, cultural awareness, and cultural skill. This has led to confusion in the vast literature surrounding cultural sensitivity (Ridley, 1984; Ridley et al., 1994) as authors have tended to use the terms interchangeably. Speigal and Papajohn (1986) have accordingly questioned how one can operationalize the terminology in order to construct treatment models.

Ridley et al. (1994) have called for cultural sensitivity that is linked to culturally responsive interventions. They formulated a model that proposed that cultural sensitivity can be viewed as being measured on a continuum which has underlying prerequisite behaviors, culturally responsive behavior, and the effects of this culturally responsive behavior. The assumptions that underlie this model are fivefold. First, the authors believe that cultural sensitivity depends on the personal meaning that a client has about events that depend on the broad spectrum of cultural attitudes available to the client (Pedersen, 1994). Second, therapists need to adopt a "naive" posture in processing information (p. 129). Third, cultural sensitivity is a prerequisite for interventions that are culturally responsive.

Fourth, cultural differences that exist between client and therapist may interfere with a therapist's information processing. Fifth, cultural sensitivity is grounded in the information processing theory of perceptual schemata; mental schemata develop likely as a natural process during client and therapist interactions. The authors have called for empirical studies that can verify or expand this model.

Culturally sensitive counselors, therefore, are seen as implicitly facilitating therapy that has more effective outcomes for diverse clients, while culturally insensitive counselors are seen as ineffective change agents. Culturally insensitive counselors have been viewed as a main contributor to problems of ethnocentrism, lack of empathy, misdiagnosis, and treatment selection errors (Gim, Atkinson, & Kim, 1991). In addition, the underutilization of mental health services and early termination by minorities have been attributed to the lack of sensitive therapy services (Romeo, 1985; Sue & Sue, 1990).

Pedersen and Lefley (1986), Sue and Sue (1990), and Sue et al. (1992) state that, for effective counseling to occur, the counselor and the client must also be able to appropriately and accurately send and receive both verbal and nonverbal messages. The next section of this chapter looks at the various factors that lead to effective interviewing and counseling. In order to understand effective interviewing and counseling strategies, however, one must first examine the elements of effective intercultural communication and the barriers present in noneffective communication. This necessitates a discussion of how communication styles differ between client and therapist.

10.04.3.2 Intercultural Communication Style

Communication style refers to those aspects that go beyond the literal content presented in verbal interactions. It includes the cadence of speech, inflection, tone, syntactical structure, fluidity, and depth. Communication styles are strongly correlated with race, ethnicity, and culture (Sue & Sue, 1990). For example, Yum (1991) has suggested that Confucianism has strongly impacted on communication styles with Asians; the main function of communication is to initiate, foster, and maintain social relationships. Communication, therefore, encourages these relationships and sets the framework for an infinite interpretation (Cheng, 1987).

In a multicultural setting the therapist usually conducts therapy without the familiar cultural guidelines that are normally present in counseling sessions with mainstream clients (Pedersen, 1994). Garza, (1981) focused on issues that emerge during initial diagnostic interviews with minorities. He discussed the notion that therapists, when confronted with a culturally different client, depart from their usual ways of interviewing. These departures include the use of cultural stereotypes as a distancing device. The therapist also tends to use generalizations which produce dehumanizing consequences. Furthermore, therapists are guilty of simplifying the task of diagnosis by overinterpreting, Garza stated:

> Overdiagnosing issues around cultural difference, precise and accurate as it may be, can at times serve the purpose not so much of enhancing the diagnostic process through the lens of a bicultural perspective, but of avoiding the task of really getting into the complexities of another's individual life. (p. 15)

Sue and Sue (1990) have delineated generic characteristics in Western counseling and psychotherapy. These characteristics incorporate culture, class, and language issues. Specifically, these generic characteristic include standard English, verbal communication, individual-centered therapy, verbal/behavioral/emotional expressiveness, openness, cause and effect attributions, and clear mental–physical dimensions. In addition, counseling models and techniques have traditional middle-class use of language, verbal communication, rigid time frames, long-range goals, and ambiguity. Sue and Sue (1990) describe these variables as being largely used in Western models of counseling and therapy. Non-Western cultures, on the other hand, have different modes of expressing affect, using silence, different language usage, differing time perspectives, nonstandard English, and a focus on individualism versus collectivism.

Sue (1977, 1981) has outlined in detail the barriers to effective cross-cultural counseling that have evolved as a natural consequence of utilizing generic modes of counseling with culturally diverse clients. Culture-bound values encompass several areas of concern. First, most forms of therapy are individually focused, emphasizing the importance of the client–counselor relationship. Some Western cultures, such as are found in the USA, have been identified as cultures that are based on individualism and competition between members. Individualism and competitiveness are seen as virtues within the culture. Not all cultures are based on this notion of individualism. Many cultures are socialized in a system of communality and codependence. The values in non-Western cultures differ markedly in that some might view individualism as an impediment to true mental health. Thus, counselors who do not deal with individualism versus group orientation will potentially create problems in the interview. An extensive review of cooperation and individualism in various cultures has provided support for this viewpoint (McGoldrick, Pearce, & Giordano, 1982).

Second, many therapists create an atmosphere within an interview that promotes verbal, emotional, and behavioral expressiveness. It is felt that such openness will allow a more accurate exploration of the problem and concomitant issues. Also, that it will provide a more focused emphasis on intrapsychic issues. This belief is especially important since it can lead to misperceptions and labeling of the client's reactions during the interview. Clients are in danger of being labeled as apathetic, repressed, or resistant if they do not exhibit expressiveness and openness. An open, expressive communication style is not predominant in Hispanic, Asian, or Aboriginal cultures (Sue & Sue, 1990). Studies have shown, for instance, that Chinese and Japanese cultures value restraint of feeling and circumspectual approaches when communicating (Samovar & Porter, 1991). Additionally, openness and expressiveness are antithetical to Native American communication styles (Dignes, Trimble, Manson, & Pasquale, 1981; Thomason, 1991).

Many interviews and psychotherapeutic interventions encourage and foster self-disclosure from the client. The problem with this approach is that self-disclosure has been determined to be a negative value trait among many cultures (Sue & Sue, 1990). Self-disclosure is a Western psychotherapeutic value. It has been called the *sine qua non* of most approaches to therapy (Poston, Craine, & Atkinson, 1991). Therapists may misinterpret a client's nondisclosure, as different cultures teach and validate different levels of self-disclosure (Pedersen, 1994).

Another culture-bound barrier involves the notion of "insight" in psychotherapy. Client insight is generally assumed to have positive or beneficial consequences, but clients from different cultures may place lower value on insight (Sue & Sue, 1990). This dichotomy can impede accurate communication. In addition, Sue and Sue described counseling and therapy within a Western context as linear and analytic with an emphasis on cause-and-effect explanations of healthy psychological functioning and psychological maladjustment. This reliance on scientific method and rationalism has contributed to client–counselor barriers in therapy, since different attributions about the etiology of psychological problems may exist between client and therapist.

Finally, the culture-bound barrier of ambiguity is concerned with the level of structure that is present in an interview or counseling session. Sue and Sue (1990) state that lack of definitive structure may be confusing to clients from Hispanic backgrounds and may impede the communication process. They go on to state that the patterns of communication among culturally diverse clients are dependent upon one's cultural upbringing.

The focus on differing communication styles has led to an ever-increasing list of models, prescriptives, and resources for effective interviewing and counseling with specific populations (ACA, 1990; APA, 1994; Eisenbruch, 1990; Fiksdal, 1990; Gonzalez-Lee & Simon, 1990; Hollingsworth, 1987; Pedersen & Ivey, 1993; Song & Parker, 1995; University of Wisconsin, 1991). In particular, these differences in communication styles have led to a detailed examination of language issues that encompass verbal and nonverbal aspects of communication.

10.04.3.3 Cross-cultural Verbal Communication

Language issues constitute one of the most important differences and one of the greatest barriers to effective communication between cultures (Argyle, 1991). Labov and Fanshel (1977) state that the most important data during the counseling session are the words spoken by the client and counselor and, thus, whatever is construed from these words. Further, Ivey (1981) has stated that any person-to-person encounter must consider the language of helping. This is especially crucial in multicultural counseling since different cultures use different words, syntax, and constructs to convey meaning.

Ivey called for an examination of language patterns as central to effective cross-cultural interviewing and therapy. Language in therapy can be idiosyncratic and misunderstood by therapists who are not familiar with a particular culture (Draguns, 1981). Meaning is lost in therapy when the counselor is not attuned to the subtleties of specific cultural uses of affect, diagnosis, and treatment (Vontress, 1981; Yamamoto, Silva, Justice, Chang, & Leong, 1993).

Yum (1991) has indicated that east Asian communication patterns differ from those in North America. East Asian languages are very complex and there are honorific linguistic systems present in Asia which differentiate the use of words based on formality or informality of context (Ogino, Misono, & Fukushima, 1985). Dillard (1983) and McWhirter and Ryan (1991) have described the enormous number of dialects and language systems in North American native clients. Cross (1995), Dillard (1983), Pennington (1979), and Sue and Sue (1990) have described the language characteristics of African-Americans and their impact on therapy.

Language plays a critical role in psychotherapy, especially when the therapist and client do not share a common language. Westermeyer (1993) has stated that most cultures have a *lingua franca* or single major language. Comas-Diaz and Griffith (1988) have described the differential impact of a client using nonstandard English. They claimed that clients are diagnosed and assessed primarily on the basis of their ability to use standard English. They stressed the importance of accurate and meaningful client–therapist communication.

10.04.3.4 Nonverbal Communication

Effective intercultural counselors need to be aware of the total process of communication. They must be sensitive to the fact that communication is both verbal and nonverbal (Wolfgang, 1985). The study of nonverbal behavior has been divided into three areas: proxemics, kinesics, and paralinguistics. Overviews of the early work done in nonverbal communication can be found in Harper, Wiens, & Matarazzo (1978), Henley (1977), and LaFrance and Mayo (1978), and a comprehensive bibliography of the nonverbal literature from 1859 to 1983 has been compiled by Wolfgang and Bhardwaj (1984).

10.04.3.4.1 Proxemics

Hall (1959) initiated the study of proxemics. He demonstrated the impact of ethnicity and culture on various nonverbal behaviors such as eye contact, gestures, and personal space. The term "personal space" was introduced by Sommer (1959) to denote the area that separates ourself from others. The size of this area varies with situations and contexts. Hall (1969) delineated four zones of interpersonal distance that characterize Western culture: intimate (up to 18 inches), personal (18–48 inches), social (48 inches to 12 feet), and public (greater than 12 feet). Extensive research into proxemics has revealed that conversational distances vary cross-culturally and that these differences have an impact on formal and informal counseling interactions (Wolfgang, 1985).

Ramsay (1979) examined the studies that evolved from the proxemic inventory described by Hall (1959). These studies examined Arabs, Latin Americans, southern Europeans, northern Europeans, Indians, Pakistanis, east Asians, and Puerto Ricans. This body of literature served to illustrate that different cultures have different culturally determined ideas of appropriate social distance. Hanna (1984) studied the interactional distances between blacks and whites in the USA and discovered substantial variation in conversational distance. Jensen (1985) discovered smaller acceptable conversational distances for Latin-Americans, Africans, African-Americans, Indonesians, Arabs, and South Americans than Anglos.

Some studies have extended the investigation to include cultural attitudes toward crowding and its resultant effect on intimacy (Altman, 1975; Altman & Chemers, 1980; Pandey, 1978, 1990, 1996; Worchel & Teddlie, 1976). These studies illustrate that attitudes towards crowding differ around the world with respect to individual comfort level, as every culture has different tolerance levels for personal space. Further research is needed to clarify the relationship of these findings to the dynamics of interviewing.

10.04.3.4.2 Kinesics

The study of kinesics was originally described by Birdwhistell (1970) and focused on body movements. Kinesics include facial expressions, gestures, eye behavior, and posture. Kinesics are influenced strongly by the norms of one's culture (Smith, 1984). Rosenthal, Hall, Archer, Dimatteo, and Rogers (1979) discovered that subjects from outside of the USA, in countries that were culturally and/or linguistically similar (e.g., Canada, New Zealand, and the UK), scored higher than other nations on tests measuring the nonverbal accuracy of meaning. In other words, the subjects were able to interpret the body movements and gestures accurately. American subjects received the highest accuracy scores, indicating that culturally specific cues are present in a person's body language. Key (1975) described nonverbal acts within an overall structure that related kinesic qualities to communication channel, time, distance relationship, cultural affiliation, environmental considerations, and individual differences. She believed the preceding determinants of nonverbal language should be considered as interactional in nature.

Sue and Sue (1990) state that eye contact is "perhaps the non-verbal behavior most likely to be addressed by mental health providers ... counselors attribute negative traits to the avoidance of eye contact: shy, unassertive, sneaky, or depressed" (p. 55). Studies have illustrated that higher levels of gaze and eye contact invite more positive reciprocal nonverbal behavior (Smith, 1984). Studies on gaze behavior in Europe have been analyzed and summarized by Ellgring (1984); these studies showed cross-national differences among subjects.

Ramsey (1984) discovered differences between Eastern and Western cultures in eye gazing behavior largely through an analysis of studies conducted in Japan. Graham (1985) discovered differences in gazing behavior among Brazilian, Japanese, and American negotiators. Researchers have investigated the role of gaze-directed behavior and reciprocal interactions (Argyle & Cook, 1976; Argyle & Dean, 1965; Exline, Gray, & Schuette, 1985). Johnson (1972) discovered differences between black and white clients with respect to eye movements and discovered eye-rolling behavior in black American speakers. These eye movements were used to convey disapproval, impudence, and hostility.

Ivey (1994) identified the dimensions of attending. The first dimension concerned the use of the eyes. He stated that eye contact is desirable in some cultures, while in others it has negative connotations. He claimed that eye behavior had the universal characteristic of being used to convey meaning. Ivey concluded that eye contact should be modeled by the counselor as early in the relationship as possible. Pedersen and Ivey (1993) offer practical suggestions on improving one's eye contact and other attending skills for use in cross-cultural counseling and therapy.

Gestures refer to the types of hand movements that people make. Although there has been a great deal of disagreement as to what constitutes a gesture (Krauss, Chen, & Chawla, 1996), a typology of gestures has been proposed by Kendon (1983). This classification system describes symbolic, conversational, and adaptational gestures. The differentiation of gestures grew largely out of early anthropological studies with Darwin and included the work of Kleinberg and LaBarre (Efron, 1972; Ramsey, 1979). Morris, Collette, Marsh, and O'Shaughnessy (1979) studied the meaning given to various hand movements in British, Greek, Spanish, Turkish, and Italian subjects. They studied 20 forms of gestures and discovered that only two had universal use, and five had widespread use. The occurrences of the remaining gestures were found to be culturally specific. They concluded that people from various regions in Europe use substantively different gestures when communicating.

The most often studied area in cross-cultural interaction is conversational gestures, which are hand movements that accompany speech and seem to be related to the meaning of that speech. LaBarre (1985) documented the different gestures of politeness in south Asian, African, European, Pacific islander, east Asian, and American cultures. He also identified cross-cultural differences in gestures of contempt, beckoning behavior, kissing, and sticking out of one's tongue. These culturally determined nonverbals have been analyzed extensively by Krauss, Chen, and Chawla (1996). They have indicated that gestures contribute to comprehension and communicative intent in conversations and interviews. Touching another person has been studied from a cross-cultural perspective. It has been claimed that touching holds special significance in Anglo cultures because of its infrequency of occurrence (Henley & LaFrance, 1984). Touching behavior studies, even though generally focused on gender issues, have been reviewed by Knapp (1984).

Eakins and Eakins (1985) have gone beyond looking at hand movements to an examination of head movements. Jensen (1985) has illustrated that cultural differences exist in the meaning attached to head movements. A most detailed focus on head movements can be found in Erickson and Schultz (1982). The authors used a microethnographic approach to study social interactions. Head-nodding behavior was analyzed within a cultural context. It was discovered that head nodding was part of a conversational rhythm that contained culturally specific patterns. They described the impact of cultural organization or shared standards and communication traditions on nonverbal social interactions. They stated that "both the ends and means of communication are culturally defined as appropriate, effective, and intelligible. When persons meet who have learned different communicative traditions regarding intelligibility, effectiveness, and appropriateness, troubles can result in the social organization of their interaction" (p. 100).

Pearson (1985) believed counseling interpretations can be based on the facial expressions of a client. Early studies of facial expression focused on the universality of expressions and the culturally specific interpretations of facial movement (Ekman & Friesen, 1975; Izard, 1971). Ekman (1993) claimed that such psychological luminaries as Allport, Brunswik, Hull, Maslow, and Titchner all studied facial expressions early in their careers, each focusing on the dilemma of whether or not facial expressions are universal or culture-specific. Ekman concluded after considering the history of kinesics, "I found more than one answer. Different aspects of expression are both universal and culture-specific" (p. 391). In particular he studied facially active vs. facially inactive subjects.

The first issue of concern was specificity. This refers to any differences that exist in the manifestation of emotions. Are clients facially active with some emotions, yet facially inactive with others? Ekman (1993) concluded that individual differences are present with respect to specificity and that future research on these differences may shed light on the issue of universality. Yamamoto and Kubota (1983) described the differences in facial expressions between Asian and European clients. In particular, restraint of facial expressions was valued in Asian culture.

The second issue is threshold. Do individuals have a lower threshold for emotional experience than facial expression? Are clients able to experience psychological distress without showing physical signs? Izard and Izard (1980) theorized there was a cultural universality of emotion that was strongest in the area of primary emotions such as fear, sadness, and happiness. He argued the further one was removed from these basic primary emotions, the more cultural specificity occurred in the expression of emotion. Ekman (1993) has suggested further research is needed in this area.

Choosing one postural stance from the many options that are available appears to be culturally determined (Hewes, 1955; Ramsey, 1979). Mehrabian (1972) described posture as having two dimensions. These dimensions, immediacy and relaxation, were found to be related to how one communicates one's attitudes. Leaning forward increases immediacy of contact and communicates acceptance and liking. Leaning backwards tends to indicate relaxation and sometimes disinterest. These dimensions have been studied by Bond and Shirashi (1974) with Japanese subjects. They discovered preferences for counselors to lean forward, since flexibility and politeness were attributed to this postural movement. Scheflen (1964) linked the notion of posture congruence to successful psychotherapy. He claimed that if a therapist wished to communicate congruent feelings, he or she needed to have a relaxed posture that was reflective of his or her state of mind.

From the field of neurolinguistic programming there have begun to emerge studies on posture that are cross-cultural in scope. Sandhu (1984) found that therapists' mirroring of Choctaw adolescents in the USA increased adolescents' perceptions of empathy present in the therapists. Sandhu, Reeves, and Portes (1993) further studied the effects of mirroring on empathy with Native Americans. They concluded that postural mirroring may significantly aid in the building of rapport with culturally different clients.

10.04.3.4.3 Paralinguistics

Paralinguistics is the study of vocal cues that clients use to communicate meaning. Poyatos (1984) defined paralanguage as "the ever present co-occurrent voice modifications or alternating independent sounds of perfectly lexical value" (p. 433). These aspects of language include the higher volume used in conversation with Latins and Arabs, high-pitched or breathy intonations of some clients, tongue clicking as seen in South Africa, and sighing. It includes the use of silence, hesitations or pauses, rate of speech, inflections, and expressiveness (Sue & Sue, 1990).

Lass, Mertz, and Kimmel (1978), Mehrabian and Ferris (1976), and Poyatos (1976, 1984) concluded that paralanguage is culture specific. Wolfgang (1985) has described the use of extraverbal elements associated with speech, from a cross-culture perspective. He has described these elements as operating at the unconscious level in clients. Weitz (1972) outlined the effects of paralanguage in black American and white American communications. Crystal (1975) has looked at the various cultural uses of paralanguage. Apple, Streeter, and Krauss (1979) indicated that pitch and speech rate can significantly affect a person's attributions. Erickson and Shultz (1982) provided a detailed exploration of interactional rhythmic styles in social interactions. They explored cultural differences with regard to pauses, hesitation, and silences. The authors concluded "behavioral regularity, especially rhythmic regularity, may be *prima facie* evidence of shared interpretative frameworks among those engaged in interaction" (p. 143).

Nonmembers of the culture may display interactional arrhythmia and fail to engage a client in a smooth conversation or verbal interaction. Erickson and Shultz (1982) have claimed that when counseling occurs with ethnically and racially diverse clients, interactional arrhythmia occurs which causes discomfort in the one-on-one interview. These interactional differences may cause counselors who are unaccustomed to the interactional rhythm to be "out of sync in communicating" (Kim, 1988, p. 91) with different language proficient clients. In addition, Kim (1988) indicated that stuttering, accents, and other paralinguistic patterns which differ from the language used in therapy could also stifle the communication process.

Thomas (1995) has stated that, in Great Britain, English is the predominant language used in intercultural counseling. He believes that therapists experience difficulty in attending to the message delivered by the client because of differential usages of the English language. These paralinguistic patterns also include the notion of monochronic and polychronic time systems (Hall, 1976). Hall stated that monchronic time patterns are prevalent in the USA, where punctuality is valued, and the pace of the conversation differs from the polychronic time orientation found in Mexico and other Latin American countries. Kim (1988) stated that these vocal patterns help to establish the emotional and attitudinal foundations of social interaction and clarify the relationship between client and therapist.

10.04.3.5 Use of Translators

Translation has been described as "the exchange of the demonstrative meaning of a word, phrase, or sentence in one language for the same meaning in another language" (Westermeyer, 1993, p. 129). Some therapists call for literal translations when working with linguistically different clients; however, many therapists move beyond literalism to the interpretation of a client's words through the discovery of connotative meanings. The latter is attempted in order to achieve accuracy of meaning. De Figueiredo (1976) and de Figueiredo and Lemkau (1980) stressed the need for accurate translation when interviewing in Portuguese and Knokani in Goa, India. Kline, Acosta, Austin, and Johnson (1980) interviewed Spanish-speaking patients with the aid of an interpreter. They discovered increased patient satisfaction and greater client understanding.

Therapy with the use of translators has been shown to be time-consuming, inaccurate, and difficult (Carillo, 1982). Marcos (1976) has studied bilingual therapists and patients and discovered that language switching can be used by the patient to distance themselves from emotional issues in therapy. This process, labeled "code switching" (Pitta, Marcos, & Alpert, 1978), has been shown to occur in Hispanic clients (Price & Cuellar, 1981; Rivas-Vazquez, 1989). Russell (1988) determined that code switching during an interview or therapy session is a clear signal from the client that clinicians should ignore; it will have impact on the credibility of the therapist. In cultures where bilingualism is present, code switching can occur. The result is that communication accuracy is minimized (Berkanovic, 1980; Kline et al., 1980). Moreover, when bilingual clients switch to their mother tongue, an increase in emotional response has been observed (Marcos, Urcuyo, Kesselman, & Alpert, 1973). Santiago-Rivera (1995) has concluded that therapists need to assess a client's language dominance and preference prior to assessing their psychological and physical health.

Difficulties in translating mental health concepts into different languages occur when attempting to apply Western methods of therapy in other parts of the world. Westermeyer (1993) illustrated these issues with Thai-Lao language translators. Nichter (1981) discovered similar issues when working in south India where undesirable connotations occurred with translated terminology. Edgerton and Karno (1971) discovered differences between Mexican-American English speakers and Mexican-American Spanish speakers with respect to describing symptomatology. Jackson and Moto (1996) translated various forms of psychopathology and therapy terms from English into Chichewa, one of the main languages in Malawi, Africa. They discovered that some words lacked accurate English equivalents.

Acosta, Yamamoto, and Evans (1982) have called for non-Spanish-speaking therapists to use trained interpreters when conducting interviews with Spanish-speaking clients to ensure accuracy of communication. Trained bilingual interpreters have been used with success in community clinics in the USA (Acosta & Cristo, 1981). Baker (1981) advocated the use of interpreters with culturally diverse clients. He worked with south Asian refugees and recommended the creation of a group of interpreters who understood psychological and social processes. Additionally, Owan (1985) extensively discussed the successful use of interpreters when interviewing south east Asian clients.

Freed (1988) found that interpreters who had backgrounds in social work and psychology were effective largely due to the fact they could facilitate a therapeutic alliance with the therapist. She listed a series of translator considerations deemed important for the establishment of accurate and meaningful communication. These factors included role definition, assurance of confidentiality, cultural differences, and the individual characteristics of the translator.

While skillful translators appear to be desirable, evidence suggests that low socioeconomic clients use nonstandardized English which can create difficulties for middle-class formal language users. Forrest, Ryan, and Lazar (1978) showed that foreign medical graduates were more deficient in slang than in standard English. This factor caused them difficulty when communicating with clients. This prompted the creation of a course "American for Foreign Medical Graduate Psychiatric Residents" designed to equip therapists with slang, colloquial language, and subtleties of humor. Follow-up indicated that greater communication skills were present in the trained experimental group than in those in a nontrained control group.

The use of bilingual or bicultural counseling has received attention in the literature. Ruiz and Casas (1981) indicated that bicultural counseling can be considered as a blending of majority and minority approaches to communication. There is evidence that bicultural/bilingual therapists have increased flexibility, greater tolerance for ambiguity, and serve well as cultural mediators (Szapocznik, Rio, Perez-Vidal, Kurtines, & Sanisteban, 1986).

Song and Parker (1995) focused on bicultural Chinese-English and Korean-American young adults. They discovered their interviews were filled with shifting multiple positions as well as identifications rather than a single unitary focus on identification. This introduction of complexity into the interview has been deemed to have a positive impact on meaningful communication (Pedersen, 1991c, 1994). Rogler (1993) has examined the impact of bilingualism on the development of diagnostic errors with culturally diverse clients. He proposed a research framework designed to facilitate a culturally sensitive diagnostic interview that encompasses the assessment of symptoms from a culturally appropriate viewpoint.

Salzma (1995) discovered that cross-cultural interactions were replete with attributional discrepancies, partially based on linguistic factors. He studied Navajo and Athabascan Native Americans and discovered many assessment errors occurred during diagnostic interviews. These errors were due to linguistically based attribution biases which occurred because of differences between client and therapist. These differences included social distance, taciturnity, length of pause, and initiation of speech.

Ho (1987, 1992) calls for care in the assessment of a client's fluency in English and their dominant language when choosing a treatment modality. He delineated guidelines in this area for working with Asian-American families. Similarly, Nguyen (1992) examined Asian-American communication patterns and called for therapists to consider the metaphoric nature of Japanese, Chinese, and Vietnamese languages.

10.04.4 THERAPIST–CLIENT MATCH

There has been a long-standing debate over the issue of whether a client should be given therapy by a culturally similar therapist or one who is not of the same cultural background. Atkinson 1983, 1985) provided comprehensive reviews of the studies that occurred in the 1970s and early 1980s. Atkinson and Lowe (1995) updated the earlier reviews with a more contemporary discussion of preference for counselor ethnicity. They stated that early research on client–therapist match focused on the preferences of African-American clients. Most studies (e.g., Bernstein, Wade, & Hofmann, 1987; Greene, 1982; Harrison, 1975; Sattler, 1977) revealed that African-American clients preferred African-American counselors and that Asian-Americans preferred Asian-American therapists over white therapists. These studies were predicated on the work of Carkhuff and Pierce (1967), LeVine and Campbell (1972), Mitchell (1970), and Stanges and Riccio (1970) which indicated that clients prefer counselors who are culturally similar to them.

Higginbotham (1977) and Higginbotham, West, and Forsyth (1988) provided an alternative point of view by concluding that race or ethnicity may not be the most important variable in cross-cultural counseling. They believed that situational, rather than cultural, variables could evoke client outcome expectancies. In addition, Atkinson (1983) concluded that, even though African-American clients showed a preference for racially similar counselors, other ethnic groups showed no such preference. Instead, these studies have stressed that other variables of counselor effectiveness have a greater impact with ethnically diverse clients. Parloff, Waskow, and Wolfe (1978) have pointed out the need to include the sex, socioeconomic status, and attitude of the counselor in client preference studies. Some studies have concluded that clients actually prefer ethnically dissimilar counselors (Gamboa, Tosi, & Riccio, 1976). In addition, Flaskerud (1990) reviewed the literature on client match from 1970 to 1990 and concluded that there was no strong support for the notion that client–therapist match in ethnicity, language, and gender had positive effects on the process or outcome of therapy.

Many studies in the 1980s focused on the impact of racial identification on preference for counselors (Parham & Helms, 1981, 1985; Sanchez & Atkinson, 1983). These studies concluded that the more one identified with one's own cultural group, the greater likelihood of preference for an ethnically/racially similar therapist. In addition, some researchers have focused on the issue of clinical judgment as it is related to the ethnicity of the therapist (Lopez & Hernandez, 1987; Malgady, Rogler & Constantino, 1987; Pakov, Lewis & Lyons, 1989). They discovered that clients perceive ethnically/racially similar counselors as possessing greater clinical judgment skills than nonsimilar counselors.

In the 1990s, the majority of studies have tended to support the notion that ethnically/racially diverse clients prefer ethnically/racially similar therapists (Coleman, Wampold, & Casali, 1995). The following three studies, however, have failed to support this notion. Steward, Gimenez, and Jackson (1995) studied the preferences of academically successful, culturally diverse students and concluded that ethnic match was not a well-defined preference among this group. Sue, Fujino, Hu, Takeuchi, and Zane (1991) failed to show any treatment effects when counselors were matched for ethnicity. Gottheil, Sterling, Weinstein, and Kurtz (1994) matched therapists and patients with respect to gender and race. The matching did not increase the proportion of cocaine-dependent patients that returned for a second session.

The following studies, however, have examined client preferences, and results support the notion that clients prefer same-race/ethnic counselors. These studies focus on Asian-American, Hispanic-American, African-American, and Native American subjects.

Yeh, Eastman, and Cheung (1994) investigated the effects of language and ethnic match with Asian-American clients and therapists. Ethnic match was a significant predictor of the dropout rate after the first session, while language match failed to predict early termination of counseling. Takeuchi, Mokuau, and Chun (1992) investigated the counselor matching among Asian-American and Pacific islander clients. They discovered that ethnic matching significantly reduced premature termination of therapy. In fact, when clients were matched with ethnically similar therapists they were five times less likely to prematurely leave therapy than clients matched with an ethnically similar therapist. They found, however, no subject differences in global functioning, an outcome measure of psychological, social, and occupational functioning as determined by the Global Assessment Scale.

Yeh, Takeuchi, and Sue (1994) discovered that Asian-American children and adolescents who received treatment at facilities with ethnically similar therapists were more likely to remain in therapy and achieve higher functioning upon termination than those at facilities with mainstream professionals. Flaskerud and Liu (1991) discovered that ethnicity match with Asian clients and therapists significantly reduced the early termination rate but did not affect global functioning scores. The researchers also examined the effect of counselor language and ethnicity on the outcome of therapy with southeast Asian clients (Flaskerud & Liu, 1990). Ethnicity match and language match significantly increased the number of client sessions but not the overall global functioning upon discharge. Flaskerud and Hu (1994) further studied the process of therapy for Asian-Americans suffering from major depression. They concluded that treatment by an ethnically similar therapist within an ethnically matched facility was related to the duration of treatment, not to the clients's global functioning upon termination of therapy. Fujino, Okazaki, and Young (1994) examined therapist–client match with Asian-American female clients. They concluded that ethnic match was significantly related to lowered premature termination, increased treatment duration, and higher functioning upon termination.

Tedeschi and Willis (1993) examined the attitudes of Asian international students toward counseling. Asian subjects were found to prefer an older counselor of similar ethnicity. Gim et al. (1991) examined the effects of counselor ethnicity on clients' perceptions of counselors credibility and efficacy. Asian-Americans were shown to attribute greater effectiveness and credibility to Asian counselors.

Atkinson and Matsushita (1991) discovered that Japanese-Americans were most apt to seek counseling from a directive Japanese-American counselor. Ngugen (1992) discovered that Asian-American patients extended therapy with Asian-American therapists. Several disadvantages were noted in the study with therapist matching. Some Vietnamese-American and Chinese-American clients preferred Western therapists, because of a fear of encountering breaches of confidentiality with Asian-American therapists, a concept apparently not totally adhered to in Asian-to-Asian encounters.

Lopez, Lopez, and Fong (1991) reviewed client preference among Mexican-Americans. They discovered that all studies supported the notion that Mexican-American clients prefer ethnically similar counselors. The authors stated they had conducted two studies of their own confirming this preference. A third study showed that medium and highly acculturated Mexican-Americans indicated greater preference for same-ethnicity counselors than Mexican-Americans with low acculturation.

Atkinson and Wampold (1993) critiqued the methodology used to obtain these conclusions. They called for the avoidance of simple choice methods to assess preference for counselors with Mexican-Americans. Further, Atkinson, Casas, and Abreu (1992) examined low-, medium-, and high-acculturated Mexican-Americans in relation to perceived counselor competence. They concluded that cultural competence was related to perceived cultural responsiveness, regardless of counselor ethnicity or client acculturation. Hess and Street (1991) studied Mexican-American high school students with respect to the effects of counselor ethnicity. No significant differences were found in the ratings of effectiveness in relation to ethnicity. Yeh et al. (1994) reported a lower therapy dropout rate when Mexican-American clients were matched with ethnically similar therapists. Sue et al. (1991) showed that ethnic match between Mexican clients and therapists increased the length of treatment and improved the effectiveness of therapy.

While most research indicates problems with matching of black clients with white therapists (Krause & Miller, 1995), problems associated with black therapists working with black clients have been observed (Montalvo & Gutierrez, 1988; Thomas, 1995). These matching studies have focused on the attractiveness and credibility of the therapist.

Tomlinson-Clarke and Cheatham (1993) matched ethnically/racially similar counselors with black and white clients. They discovered no significant differences in counselor judgments (i.e., clinical disposition, severity of diagnostic rating, need for psychological treatment, and predicted number of sessions) upon intake when clients were matched with similar counselors. Goldberg and Tidwell (1990) investigated perceived counselor–client similarity and its relationship to the clients' perceptions of attractiveness and the satisfaction derived from therapy. They investigated both gender and race with black and white counselors and students. They concluded that racial and gender differences appeared not to function as barriers to effective communication. They suggested instead that other counselor variables such as attitudes and personality were more important characteristics for perceived competence.

Redfern, Dancey, and Dryde (1993) examined the role of counselor ethnicity with English black and white undergraduate students. Black counselors were rated higher than white counselors on scales of attractiveness, competency, and trustworthiness. Tomlinson and Cheatham (1989) examined the effects of ethnicity match on intake interviews with black clients. They concluded that the race of the therapist affected the clinical judgments at the time of client intake.

Race and ethnicity have been studied within the psychoanalytic relationship (Leary, 1995). Three patients were paired with an African-American therapist. Qualitative descriptions showed the importance of client–therapist negotiation when racial characteristics differ. In addition, White (1994) investigated the impact of race and ethnicity on transference and countertransference in interracial group psychotherapy and discovered differential transference issues related to the race or ethnicity of the therapist.

Bennett and Bigfoot-Sipes (1991) examined Native American and white college student preferences for counselor ethnicity. They concluded that both American Indian and white students preferred counselors to have similar values, but Native Americans showed greater preferences for ethnically similar counselors, while white students showed no such preference. In addition, Johnson and Lashley (1989) examined the counselor preferences of Native American undergraduate students. They discovered that clients with a strong cultural commitment to Native American values placed great importance on the ethnicity of the counselor by preferring counselors with similar ethnic backgrounds.

Comas-Diaz and Jacobsen (1987) proposed that a client's ethnocultural identity be assessed as an auxiliary therapeutic tool. One aspect of their identity assessment focuses on the background of the therapist relative to that of the client. This assessment then aids in matching client and counselor. Hall and Maloney (1983) suggested that highly ethnoculturally dominant therapists should be matched with same-ethnicity clients while less ethnoculturally dominant therapists be matched with culturally different clients. Coleman et al. (1995) conducted a meta-analysis of studies that investigated perceptions and preferences of ethnically diverse clients. They concluded that ethnic minorities tended to prefer ethnically similar therapists and rated these therapists as more competent than European-American therapists.

The client–therapist match literature in the 1990s, therefore, has focused on the preferences for counselors of Asian-Americans, Hispanic-Americans, African-Americans, and Native Americans. Various topics in relation to client preference have been investigated, most notably: racial/ethnic identification, acculturation, counselor credibility, early termination from therapy, and outcome measures of global functioning. Investigation of these and other variables will continue into the next decade.

10.04.5 TREATMENT OUTCOME

10.04.5.1 Definition

Treatment outcome research was defined by Mowrer (1953) as a situation whereby the "emphasis is upon measuring significant aspects of personality before and after treatment and noting the nature and extent of the resulting changes" (p. 4). Sue and Zane (1987) have stated that therapeutic outcome is "the cumulative product of many discrete dynamics between client and therapist" (p. 44). Orlinski, Grawe, and Parks (1994) have asserted that the term "treatment outcome" has had a history of divergent meanings that have ranged from observational perspective to analysis levels. The authors pointed out that the definition and criteria for outcomes differ depending on who is completing the assessments before and after treatment is received. That is, differences are present if the assessment is completed by therapists, clients, nonparticipant observers, or nonprofessionals. They indicated that level of analysis also had critical issues in need of clarification. These issues included the outcomes occurring within, and external to, the therapy sessions and the use of evaluative or descriptive assessment instruments. The authors concluded that "treatment outcome should refer to changes in condition (psychological, somatic, physical, social, and cultural) reflecting favorable or adverse effects on the patients well being" (p. 284).

10.04.5.2 Cross-cultural Outcome Research

VandenBos (1996) presented a succinct history of outcome research beginning in the 1950s. He pointed out that most contemporary treatment outcome research has focused on the effectiveness of comparative treatment interventions and/or has focused on specific treatment methods targeting specific clinical problems. Goldfried and Wolfe (1996) examined the difficult relationship between psychotherapy practice and research. They called for a new outcome research paradigm that seeks to increase clinician and researcher collaboration. Other recent studies have examined many issues surrounding the outcome assessment of psychotherapy, pointing to methodological and conceptual problems (Barlow 1996; Hollon, 1996; Howard, Moras, Brill, Martinovich, & Lutz 1996; Jacobson & Christensen, 1996; Newman & Tejeda, 1996; Sechrest, McKnight, & McKnight 1996; Seligman 1996; Strupp, 1996).

While these studies examined many of the contemporary issues surrounding treatment outcome research, they conspicuously omitted any mention of ethnic or racial factors that contribute to the efficacy of treatment. Sue et al. (1994) have comprehensively analyzed the research on American minority clients. Specifically, they reviewed the outcome research using African-American, Asian-American, Native American, and Latino clients. The authors addressed the question: Is there any research evidence that reveals the efficacy of treatment with minority clients? Specifically, their paper addressed whether minority clients demonstrated differences in pre- and post-therapy assessment. It also investigated whether ethnic minority treatment outcomes differed from outcomes seen in majority clients, and whether various ethnic groups had differential treatment outcomes. They pointed out that research in these areas has historically been lacking and has not specifically addressed the outcome of psychotherapy. Sue et al. asserted that the relationship between ethnic/cultural/racial variables and psychotherapy outcome has not been determined with empirical evidence.

Zane and Sue (1991) stated that ethnic differences involve group membership, while cultural differences "constitute a host of cognitive variables which are limited to different cultural lifestyles and perspectives" (p. 32). They stated that these cognitive variables, rather than ethnic membership, have been responsible for the lack of conclusive psychotherapy outcome studies.

Sue et al. (1994) identified the reasons why outcome studies in cross-cultural psychology have been problematic. They pointed to four areas of concern: (i) the use of analogue studies, (ii) the heterogeneity of the sample, (iii) the insensitive or inappropriate use of dependent measures, and (iv) the lack of within-group design strategies. These methodological deficiencies have caused both researchers and clinicians to dismiss outcome research as being of limited practical value. With these limitations in mind, researchers have most often focused on treatment outcomes of African-Americans, Asian-Americans, Latino-Americans, and Native Americans. Research studies with these four groups will now be examined.

With regard to studies conducted with African-American subjects, Sue et al. (1994) state that, generally, "in no studies have African-Americans been found to exceed white Americans in terms of favorable treatment outcomes. Some investigators have revealed no ethnic differences and some studies have supported the notion that outcomes are less beneficial for African Americans" (p. 788). These conclusions were based on studies conducted in the 1970s and 1980s (Brown, Joe, & Thompson, 1985; Griffith & Jones, 1978; Jones, 1978, 1982; Lerner, 1972; Parloff et al., 1978; Sattler; 1977).

Latino/Hispanic-American studies have often been evaluated in terms of the utilization and retention rate of patients in therapy. Studies using more direct measures have been scarce in the literature. However, discussion on culturally appropriate modes of therapy and the effectiveness of therapies tailored for Latino/Hispanic clients have been studied (Sue et al., 1994; Szapocznik et al., 1989). Cuento therapy, utilizing Puerto Rican folktales, was examined by Constantino, Malgady, and Rogler (1986), and results from this study confirmed that culturally sensitive forms of therapy can reduce trait anxiety scores. Malgady, Rogler, & Constantino (1990) discovered that using culturally appropriate modeling therapy in conjunction with Cuento (folktales) therapy reduced anxiety symptoms and aggression in Puerto Rican children. Rey-Perez (1996), using a qualitative approach with five Hispanic families, constructed a comprehensive, culturally sensitive treatment program for Hispanic children and their families. Therapeutic gains were noted in each child; the subjects were Cuban, Puerto Rican, Nicaraguan, and Peruvian.

There have been a limited number of outcome studies using Asian-American subjects. Many studies have not studied specific groups of Asian-Americans (e.g., Vietnamese, Chinese, Japanese); rather they have viewed Asians as monolithic. Kinzie and Leung (1989), Mollica, Wyshak, Lavelle, Truong, Tor, and Yang (1990), and Snodgrass et al. (1993) found outcome gains in Vietnamese, Laotian, and Cambodian patients. Comparative studies with Asian and non-Asian clients have revealed few differences (Sue et al. 1994).

The authors stated that:

> any conclusions about the effectiveness of treatment for Asians would be premature given the limited data (four outcome studies), but that several empirical trends should be noted. First, some evidence suggests that certain Asian groups improve with psychotherapy and/or adjunct treatments. Second, with respect to differential outcome, divergent trends are found and these are associated with the type of outcome measure used. . . .In summary, Asian clients appear to be deriving less positive experiences from therapy than whites but it is unclear if this difference in client satisfaction actually reflects ethnic differences in actual treatment outcomes (i.e., symptom reduction). (p. 798)

Native American populations, while culturally heterogenous, have been generally studied as an aggregate population. The number of outcome studies is limited in this area, and Manson, Shore, and Bloom (1985) and Neligh (1988) have reported that no research has been conducted to compare the efficacy of various treatment modalities with this population. Most outcome studies have focused on alcohol and drug use reduction and prevention (Bobo, Gilchrist, Cvetkovich, Trimble, & Schinke, 1988; Manson, 1982; Query, 1985; Schinke et al., 1988). Sue et al. (1994) have concluded that "it is apparent that research on intervention (i.e., treatment and prevention) has proceded very slowly, and it would be premature to try to address the question of the efficacy of mental health interventions with American Indians at this time" (p. 795).

The paucity of outcome studies in the literature suggest that more ethnically/racially specific studies on the efficacy of treatment are needed. In addition, methodological weaknesses need to be examined and corrected by future researchers. These methodological issues are explored in the next section.

10.04.6 METHODOLOGICAL CONSIDERATIONS

10.04.6.1 Reliability and Validity in Cross-cultural Research

The last decade has seen a focus on racial/ethnic research issues as they relate to shedding light on effective clinical interventions. Much of the literature has centered on the validity and usefulness of research studies that investigate racial and ethnic variables. The APA (1993) has outlined six areas that have been the focus of mental health research:

> (a) the impact of ethnic/racial similarity in the counseling process, (b) minority underutilization of mental health services, (c) relative effectiveness of directed versus nondirected styles of therapy, (d) the role of cultural values in therapy, (e) appropriate counseling and therapy models, and (e) competency in skills for working with specific ethnic populations. (p. 45).

Research designs using cross-cultural variables, as with all research designs, are expected to have adequate reliability and validity.

Ponterotto and Casas (1991) have carefully delineated the major criticisms leveled at racial/ethnic minority research. They have listed (pp. 78–79) the top 10 criticisms as:

(i) lack of conceptual/theoretical frameworks to guide research;

(ii) overemphasis on simplistic counselor/client process variables and a disregard for important psychosocial variables, within and outside the culture, that might impact counseling;

(iii) overreliance on experimental analog research;

(iv) disregard for within-group or intracultural differences;

(v) the use of easily accessible college student populations;

(vi) reliance on culturally encapsulated psychometric instrumentation;

(vii) failure adequately to describe one's sample in terms of socioeconomic status;

(viii) failure to delineate the study's limitations;

(ix) lack of adequate sample sizes; and

(x) overreliance on paper and pencil outcome measures.

In addition to the identification of the methodological weaknesses in ethnic/social research, the authors reviewed articles from the five journals that most often publish research in this area. They discovered that only about one-third of the studies had an adequate conceptual foundation. With regard to the second criticism, the authors discovered that, although many variables were investigated, key areas lacked systematic attention. Those areas included communication styles, socializing practices, learning styles, effects of poverty and discrimination, and acculturation. With respect to the third criticism, it was discovered that survey research was used in 72.5% of the studies, while analog research accounted for 12.5%, true experimental designs 8.8%, and archival designs 6.3% (p. 89). The authors concluded that researchers, rather than relying on analog studies, have overutilized a survey methodology.

The authors found overwhelming evidence to justify the veracity of their claim in the fourth criticism. They discovered that two-thirds of the literature did not include within-group designs or intracultural comparisons. The fifth criticism was justified, as it was discovered that over one-half of the studies used college and high-school student samples. The sixth criticism was also supported, as the authors discovered that only one-third of the studies used instruments that were designed for minority populations. The authors concluded this to be one of the most justified criticisms in ethnic/racial research studies.

Criticism (vii) was justified, as two-thirds of the studies failed to report social economic variables. The eighth criticism was not wholly justified, as 60% of the studies listed their limitations. Criticism (ix) was not justified, as sample size depends on many factors, including type of design and sampling characteristics. Finally, the tenth criticism was fully justified, as 90% of the studies utilized outcome measures that were paper-and-pencil instruments.

The authors concluded, therefore, that criticisms (i), (iv), (v), (vi), (vii) and (x) are justified; criticisms (ii) and (ix) partially justified; and criticisms (iii) and (viii) unjustified.

Stanfield (1993) has criticized cross-cultural research methods from an epistomological perspective. He examined the fallacies of homogeneity and monolithic identity, and the presumptions that underlie research ethics. He believed that the failure of researchers adequately to clarify these issues has led to exploitative research with ethnically and racially diverse subjects.

10.04.6.2 Recommendations

The comprehensiveness of Ponterotto and Casas's (1991) work can be seen in the methodological recommendations that they offer. These suggestions cover five areas of concern. First, because culture is complex (Pedersen 1991c), a variety of research methods need to be used. In particular, the authors called for researchers to "embrace qualitative methods that have been effectively employed in the related disciplines of ethnology, cultural anthropology, and sociology" (p. 96). Second, the authors suggest that studies focus on within-group differences rather than between-group differences. In so doing, research could avoid comparing racially and ethnically different clients to white or mainstream clients, a situation that usually implies majority behavior to be normal and minority behavior to be abnormal. Third, the authors suggest research that focuses on variables that "transcend culture" (p. 99). Those variables could include some of the universals of human experience such as economic deprivation, family loss, serious medical and psychological illness, and change in environmental conditions. Fourth, culture-specific instrumentation is recommended, especially for such psychological constructs as depression, self-concept, and assertiveness. Finally, the authors believe that more studies need to focus on the strengths of ethnically/racially different clients, rather than on their weaknesses. This suggestion also includes a focus on the positive aspects of biraciality and biculturality.

Mio and Iwamasa (1993) vociferously attacked the methodologies used in conducting ethnic minority research. They posed the question of whether or not researchers from a majority culture could and should conduct research with minority clients. Reactions to this position can be found in Atkinson (1993), Ponterotto (1993), Casas and San Miguel (1993), Helms (1993), Ivey (1993), Parham (1993), Pedersen (1993), and Sue (1993). These reactions suggested ways to improve the quality of ethnic/racial minority research.

Berry, Poortinga, Segall, and Dasen (1992) also explored methodological issues in cross-cultural psychology. They indicated that the two major methodological concerns in any empirical research study are the design and the analysis of results. The authors indicated that results from studies which focus on racially diverse subjects are more difficult to interpret validly than studies in which race and ethnicity are not considered as variables. They delineated four measures for accepting alternative hypotheses indicated by the data.

One method for increasing confidence is by selecting subjects on the basis of the variables race, ethnicity, and culture. Groups need to be included in a study based on their ethnic group inclusion. The authors point out that many studies involving ethnically/culturally diverse subjects have problems during the selection phase of the study. Differences between the groups can be due to factors that are present in the majority culture, the minority culture, or an interaction of factors between the two cultures.

Second, Berry et al. (1992) recommend the dependent variable be "expressed as a function of two or more separate scores. An example is the score that can be obtained by taking the difference betwween scores on two measurements" (p. 223); this strategy can help to eliminate rival hypotheses.

Third, they advocate the minimization of the effect of extraneous variables through statistical methods such as analyses of covariance and multiple regression analysis. Finally, they advocate that more than one dependent measure be used, through methods that include self-report measures, interviews, and archival information. This method increases the amount of data one has to use in the formulation of conclusions and interpretations.

Like Ponterotto and Casas (1991), Berry et al. (1992) show concern about the instrumentation utilized in cross-cultural studies. "Emic" approaches study behaviors of one culture from within the subjects' cultural system and use criteria that are related to the internal characteristics of that culture. "Etic" approaches study behaviors from outside of the culture, by comparing one or more cultures, using dependent measures that are considered universally applicable. The authors indicate that one must consider the emic–etic distinction when designing and interpreting research studies. Berry (1969), based on the work of Pike (1967), argued that researchers need to distinguish between culturally specific variables (emic variables) and culture-general, universal variables (etic variables). The authors advocate emic approaches to research. These recommendations reflect Ponterotto and Casas's suggestions that culturally relevant instruments be used for measuring variables in a study.

Berry et al. (1992), Hui and Triandis (1985), and Triandis (1978) have advocated a combined emic and etic approach to research. This approach identifies an etic construct to study, then operationalizes that construct through the development of emic instrumentation which evolves out of the culture studied. Bond and Smith (1996) have concluded that if future research is to proceed in a useful manner the "methodological problems stemming from the etic–emic dilemma must be more clearly addressed" (p. 226). Berry (1989) has suggested that research should proceed by investigating indigenously two or more cultures in order to arrive at a universally derived etic. This approach is in stark contrast to the more typical approach in cross-cultural psychology that uses etic measures from Western countries.

Mrinal, Mrinal, and Takooshian (1994) have outlined the various research methods that have historically been utilized by cross-cultural researchers. They discuss the use of experimentation, observation, sampling, and assessment. In particular they analyze the use of dependent measures by comparing various assessment methods and instruments. They call for flexibility and adaptation to be used when conducting research cross-culturally.

Stanfield (1993) has questioned the epistomological and methodological perspectives in contemporary cross-cultural studies. He believes that researchers have been pursuing the wrong question, or have answered incomplete questions concerning ethnically and racially diverse topics. He advocates a reorientation among researchers to re-evaluate and improve qualitative, quantitative, and comparative/historical research methods.

Qualitative researchers have advocated the improvement of cross-cultural research by focusing on class, race, or gender relations from within the culture (Andersen, 1993), participant observation (Dennis, 1993), ethnographic methods (Facio, 1993; Spradley, 1979; Williams, 1993), discourse analysis (Van Dijk, 1993), and personal narrative approaches (Jordan, 1995; Madrid, 1995; Moraga 1995). Quantitative suggestions range from re-evaluating demographic statistics (Marks 1993), to new ways of measuring and detecting discrimination (Myers 1993), to the assessment of outcomes (Patton 1993; Sue et al., 1994), and to survey research (Smith 1993). Comparative/historical methods have been suggested and outlined by Champagne (1993), Ragin and Hein (1993), and Stanfield (1993).

10.04.7 FUTURE DIRECTIONS

In 1989 three trends were predicted by Kagitcibasi and Berry. They predicted there would be an increasing focus on individualism and collectivism, on indigenous psychotherapies, and on the search for behaviors and values common to every culture (cultural universals). In 1996, Bond and Smith concluded that three trends have emerged during the 1990s in cross-cultural psychology. They suggested individualism and collectivism have been studied extensively but the other two areas have been plagued by methodological problems. They state that the emic–etic (i.e., culture-specific–universal) dilemma must be more clearly delineated. This clarification could provide new frameworks for research with culturally diverse subjects.

In addition, psychologists will have to focus on resolving issues of terminology and definition, clarifying professional and ethical concerns, addressing methodological weaknesses in research, and generating new topics for study. These areas of concern will be examined in this section of the chapter.

10.04.7.1 Terminology

The terms "ethnicity," "race," and "culture" have had a long history of clarification (Jackson, 1995; Jones, 1991). This history has been fraught with confusion and misunderstanding (Phinney, 1996). Ridley et al. (1994) and Westbrook and Sedlacek (1991) have criticized the inconsistent usage of terminology and pointed to the negative effects of labeling and stereotyping. LaFromboise, Foster, and James (1996) have suggested that the term "race" may be eliminated from future research studies, based on the conclusions of Yee, Fairchild, Weizmann, and Wyatt (1993), who indicated the lack of uniform use of the terms "race" and "ethnicity."

Alvidrez, Azocar, and Miranda (1996) have suggested methods to conceptualize, measure, and interpret ethnic or racial variables in studies of clinical interventions. The recommendations were that researchers should identify their definitions of race and ethnicity, provide more comprehensive information about subjects other than ethnicity or race (e.g., demographic information and language), identify social class variables, and indicate occupation level. The recognition that ethnicity is multifaceted has led to a re-examination of the terminology. Ragin and Hein (1993) have described the term "ethnicity" as interactive and contextual. Ethnicity and race have recently been viewed as dimensions rather than categories (Goodchilds, 1991). Phinney (1996) has suggested these dimensions "clearly cluster together in ways that make ethnicity a highly salient and meaningful construct" (p. 925).

Additionally, Phinney (1996) has also suggested that if psychologists wish to generate adequate explanations about treatment outcomes when ethnicity is a factor, they need to investigate three dimensions within and across ethnic groups. These are: significant cultural norms and attitudes, factors involved with a client's ethnic identification, and individual minority client experiences of decreased status and power.

Pedersen (1995) advocates the future use of the term "culture" as formulated by Betancourt and Lopez (1993). These authors have described culture as a learned, socially shared set of variables, incorporating individual, family, and societal variables. Fernando (1995) has examined the issue from the perspective of ethnic groups in Great Britain and has argued for a more dynamic conception of culture to be formulated in cross-cultural psychology. Pedersen and Ivey (1993) have outlined the complexity and dynamic nature of culture and have focused on the need to identify salient ethnic, racial, and cultural variables. The importance of these variables varies from moment to moment, and they include "ethnographic, demographic, status, and affiliation characteristics . . . particularly gender, ethnicity, life-style or other affiliations" (p. 188).

The consideration of these variables will allow researchers and practitioners to dissolve stereotypes of client behaviors based only on ethnic, cultural, and racial group membership. Pedersen and Ivey (1993) concluded that one needs to address the dynamics of culture by focusing on the salience and complexity of culture. This new focus will have greater relevance in future individual counseling discussions and is a foundation for an examination of ethical considerations in cross-cultural counseling (LaFromboise et al., 1996; Pedersen, 1995).

10.04.7.2 Ethical and Professional Issues

A 20 year historical analysis of cross-cultural counseling ethics, beginning with the 1973 Vail Conference, has been completed by Pedersen (1994). He traced the formulation of ethical principles in the professional associations of psychology. Ethical principles for conducting cross-cultural research were adopted in the 1970s by the International Association of Cross-Cultural Psychology (Tapp, Kelman, Triandis, Wrightsman, & Coelho, 1974). Ethical principles have also been formulated by the APA (1992) and the ACA (1988). The APA principles address the issues of race, ethnicity, and culture in an effort to minimize and dispel bias in therapeutic interventions, arising from stereotyping or discrimination. The ACA ethical principles indicate the need to respect the client's individuality and dignity with regard to cultural difference. In addition, guidelines have been provided for clinical interventions with ethnically, culturally, and linguistically different clients (APA, 1993).

Several future trends seem imminent in the area of ethics. Ethical principles are being revised and more specifically delineated. The ethics of caring and responsibility are beginning to receive more attention in the cross-cultural literature (Pedersen, 1995). Ethics in the area of multiculturalism will be expanded to delineate more fully the impact of psychotherapeutic training models on a counselor's multicultural competence (Sue et al., 1996). In addition, ethical issues concerning the types of, and approaches to, research with diverse populations will be elaborated in greater detail (Pedersen, 1995).

The issue of white American researchers conducting the majority of the research with minority clients is an area that will be examined with greater focus (Mio & Iwamasa, 1993). This issue will be expanded to incorporate the recruitment and retention of minorities in psychotherapy research, as there has been an under-representation of minorities in clinical research (Miranda, 1996). This under-representation has led to the formation of guidelines from the National Institutes of Health (NIH) in the USA concerning the inclusion of minorities and women in psychology research (NIH, 1994). These new NIH guidelines for the inclusion of minorities in clinical research will be evaluated and studied in greater depth (Hohmann, & Parron, 1996). The focus on recruitment and retention has centered on Native Americans (Norton & Manson, 1996), African-Americans (Thompson, Neighbors, Munday, & Jackson, 1996), Latinos (Miranda, Azocar, Organista, Munoz, & Lieberman, 1996), and elderly ethnic clients (Arean & Gallagher-Thompson, 1996). It is believed that the NIH guidelines will help encourage collaborative efforts between practitioners and researchers (Miranda, 1996). Specific recommendations will need to be formulated for recruiting and retaining other specific, racially, ethnically, and culturally diverse research subjects.

10.04.7.3 Future Research

Future cross-cultural research will address methodological, theoretical, and content issues. In particular, an examination of research agendas, assessment methods, multicultural counseling and therapy theory, and specific research topics, appear to be areas of future concern.

10.04.7.3.1 Research agenda

Sue and Sundberg (1996) have proposed 15 specific hypotheses to be tested in future research on individual counseling. These hypotheses are based on questions that range from client–counselor similarity, to worldview differences, to emic–etic issues. They are as follows (pp. 329–343).

(i) Entry into the counseling system will be affected by cultural conceptualization of mental disorders and socialization towards seeking help.

(ii) The more similar the expectations of the intercultural client to the goals and process of counseling the more effective the counseling will be.

(iii) Of special importance in intercultural effectiveness is the degree of congruence between the counselor and client in their orientations in philosophical values and views towards dependency, authority, power, openness of communication, and other special relationships inherent in counseling.

(iv) The more the aims and desires of the client can be appropriately simplified and formulated as objective behavior or information (such as university course requirements or specific tasks), the more effective the intercultural counseling will be.

(v) Cross-cultural empathy and rapport are important in establishing a working alliance between the counselor and the culturally different client.

(vi) Effectiveness will be enhanced by the counselor's general sensitivity to communications, both verbal and nonverbal. The more personal and emotionally laden the counseling becomes, the more the client will rely on words and concepts learned early in life, and the more helpful it will be for the counselor to be knowledgeable about socialization and communication styles in the client's culture.

(vii) The less familiar the client is with the counseling process, the more the counselor or the counseling program will need to instruct the client in what counseling is and in the role of the client.

(viii) Culture-specific modes of counseling will be found that will work more effectively with certain cultural ethnic groups than others.

(ix) Ethnic similarity between counselor and client increases the probability of a positive outcome.

(x) Within-group differences on variables such as acculturation and stage of racial identity may influence receptivity to counseling.

(xi) Credibility can be enhanced through acknowledgment of cross-cultural factors in cross-cultural encounters.

(xii) In general, women respond more positively than men to Western-style counseling.

(xiii) "The person who acts with intentionality has a sense of capability. She or he can generate alternative behaviors in a given situation and approach a problem from different vantage points. The intentional, functioning individual is not bound to one course of action but can respond in the moment to changing life situations" (cited from Ivey, Ivey, & Simek-Morgan, 1993, p. 8).

(xiv) Identity-related characteristics of white counselors can influence their reaction to ethnic minority clients.

(xv) Despite great differences in cultural contexts, in language, and in the implicit theory of the counseling process, a majority of the important elements of intercultural counseling are common across cultures and clients.

The authors have labeled their 15 hypotheses as protohypotheses because they acknowledge that the constructs and propositions need to be more fully sharpened, focused, and developed before effective research can be started. Their hope is to stimulate further thinking in this area. The testing of their hypotheses is proposed by the authors as an agenda for future research in the area of cross-cultural counseling and psychotherapy. Sue and Sundberg (1996) also have suggested that future research should broaden the criteria for inclusion into a study by using marginal and at-risk clients. They also advocate a detailed examination of the emic–etic controversy.

Ponterotto and Casas (1991) have identified areas of research that should constitute the agenda for the 1990s and beyond. The areas include: accurate epidemiological studies of psychological problems with various ethnic and racial identities, impact of Eurocentric political and social systems on clients, systemic racism, a focus on the strengths present in low socioeconomic groups, biracial identity development, ethnic minority youth research, primary prevention, assessment, and combination emic–etic studies.

Berry et al. (1992) predict a future focus on examining emic–etic issues. They envisage cross-cultural research that examines psychopathology from various perspectives. These perspectives are: invariate or present in all cultures (absolute); present in all cultures, but culturally determined (universal); and unique to specific cultures and emically definable (culturally relative). Helms (1994) has argued that the new focus on multiculturalism and pluralism has resulted in a neglect of racial factors that contribute to psychotherapy interventions. She has called for a greater focus on racial factors for the future.

10.04.7.3.2 Assessment techniques

Future research will inevitably include the development and validation of assessment techniques that are culturally appropriate. The use of traditional assessment techniques has been criticized extensively in the literature (APA, 1993; Dana, 1993; Lonner, 1990; Paniagua, 1994; Ponterotto & Casas, 1990; Samuda, 1975; Samuda, Kong, Cummins, Pascual-Leone, & Lewis, 1991; Sodowski & Impara, 1996). The issues of cross-cultural assessment and appraisal have most recently been examined by Lonner and Ibrahim (1996). The authors present a general overview of the issues involved when assessing culturally diverse clients by focusing on the impact of response sets and emic–etic issues. They have examined the use of cognitive ability tests and personality measures. In addition, they have provided a historical analysis of the assessment of the therapeutic process and the assessment of refugees. Subsequent to their extensive review they concluded that these areas need to have a greater focus in the future.

In addition to the investigation of traditional standardized cognitive and personality measures, researchers will need to develop future instruments, in various areas of concern, from an emic perspective (e.g., biracial identity and worldview). These instruments will need to be validated and thoroughly researched. Process-oriented techniques of assessment, such as those in the dynamic assessment movement (Feuerstein, 1979), will need to be considered for cross-cultural utility and feasibility (Samuda & Lewis, 1992). The area of assessment constitutes one of the most significant methodologic weaknesses (Ponterotto & Casas, 1991) and will, therefore, continue to occupy psychologists in the next decade.

10.04.7.3.3 Theory of multicultural counseling and therapy

Future research will focus on the theory of multicultural counseling and therapy (MCT). This new framework for delivery of psychotherapeutic services to culturally diverse clients, proposed by Sue et al. (1996), has six major propositions and 47 corollary suppositions. Research will undoubtedly focus on the specifics proposed by these authors.

The authors have advocated a future agenda that has researchers: (i) conduct a metareview of historical and current studies from a MCT perspective, (ii) examine the shift from a focus on the individual to a study of the individual within a context or in relation to others, and (iii) examine the positive attributes of individuals and cultures, rather than focusing on stereotyped behaviors and pathologies. Additionally, they have advocated the investigation of traditional research methods that generate epistemological assumptions that may not be universal. This future focus could incorporate a detailed examination of differing worldviews and how these differences impact on cross-cultural interventions. Finally, the authors recommend the use of qualitative research methodologies when working with a culturally diverse population. This approach, as well as quantitative methods, will need to be used effectively to “explore new methods that are culture centered rather than consider culture as peripheral” (p. 36).

The reactions to MCT theory (Arredondo, 1996; Ballou, 1996; Casas & Mann, 1996; Corey, 1996; Daniels, & D’Andrea, 1996; Highlen, 1996; LaFromboise & Jackson, 1996; Lee, 1996; Leong, 1996; Parham, 1996; Pope-Davis & Constantine, 1996; Vasquez-Nuttal, Webber, & Sanchez, 1996) and the counter-reactions (Sue et al., 1996) indicate that the dialogue will continue into the twenty-first century, with a refinement of the theory and an empirical validation of the underlying constructs.

10.04.7.3.4 Specific research topics

The methodological issues previously outlined and the research suggestions offered (APA, 1992; Berry et al., 1992; Ponterotto & Casas, 1990; Sue et al., 1996; Sue & Sundberg, 1996) provide a foundation for specific topic areas that appear to be receiving attention in the literature. Topics such as underserved populations, biraciality, the vocational behavior of diverse clients, and indigenous therapy, need more examination in the future. These four potential topics are now examined in greater detail.

In the 1990s, there has been a move to focus on specific ethnic/cultural/racial groups, rather than the more commonly studied four groups, in the USA: African-American, Asian-Americans, Latino-Americans, and Native Americans (Sue et al., 1994). This new focus on specific underserved populations and within-group populations could be a major focus in the next decade. In particular, the treatment of specific populations has begun to appear in the literature: traditional Arabic clients (Dwairy & Van Sickle, 1996), Puerto Ricans (Malgady, Rogler, & Rogler, & Constantino, 1990), Haitians (Gies, 1990), Guatemalan Mayans (Alger, 1996), Vietnamese (Snodgrass et al., 1993), Cambodians (Bemak & Greenberg, 1994), and Venezuelans (Villegas-Reimers, 1996). These authors have concluded that researchers need to go beyond the traditional ethnic/racial group categories to focus on the heterogeneity that exists within each ethnic/racial group. Future studies that examine these and other minority groups need to focus on both between- and within-group comparisons. Such a focus will provide an evaluation of the impact of the culture-specific versus universal variables on counseling and psychotherapy.

The focus on heterogeneity has revealed the presence of biracial/biethnic and multiracial/multiethnic variables. As the demographics change in countries, so do the number of intercultural and interracial marriages. Interracial, interethnic, and biracial families have become a reality in the 1990s and have received increasing attention in the psychological literature (Root, 1992). The term "biraciality" has been used to describe the children of racially different parents and is akin to biculturality and bilingualism (Kerwin & Ponterotto, 1995).

This area has evolved from early research on the biracial identity models as outlined by Kerwin and Ponterotto. The authors have suggested five research areas for future study: (i) effects of peer pressure on adolescent identity development, (ii) impact of cultural focus on identity development, (iii) dichotomous choices of identifying as multiracial/multicultural rather than choosing one racial/ethnic categorization, (iv) development of bicultural competence for biracial adolescents and adults, and (v) role of parents and educators in developing an adequate sense of identity in biracial children. This complex area of concern will receive greater study in the future.

During the 1980s and 1990s there has been an increasing focus on the vocational concerns of racial and ethnic minorities (Leong, 1995). This focus on career development of the culturally diverse has led to an exploration of the issues with African-Americans (Bowman, 1995; Brown, 1995), Hispanic-Americans (Arbona, 1995; Fouad, 1995), Asian-Americans (Leong & Gim-Chung, 1995; Leong & Serafica, 1995), and Native Americans (Johnson, Swartz, & Martin, 1995). This examination of career behavior has promulgated a multicultural theory of career development (Osipow & Littlejohn, 1995) and recommendations for vocational assessment and career counseling (Betz & Fitzgerald, 1995). Future literature on cross-cultural vocational counseling interventions will elucidate and attempt to validate this theory.

Another area of concern is the role of indigenous therapies. In many countries there is an acceptance of both standard, formal psychotherapeutic intervention methods and traditional healing methods. The role of traditional, or indigenous, healing practices has been examined by Lefley (1994) and Hiegel (1994). A comprehensive treatment of the subject is available in Kim & Berry (1993). This area of psychotherapy will inevitably continue to grow as emic approaches become more acceptable in both Western and non-Western cultures. Indigenous methods for investigation could include Morita, Naikan, Voodoo, Santoria, and Espiritismo therapy. Bemak, Chung, and Bornemann (1996) have proposed a multilevel model for counseling refugees that integrates Western therapies and indigenous methods of intervention. This integration of approaches has resulted in more effective psychotherapy outcomes (World Health Organization, 1992) and is a harbinger of future work in cross-cultural psychology.

10.04.7.4 Global Psychology

It has been suggested by Moghaddam (1987), that three worlds of psychology exist: the body of research derived from American research studies in the USA, research generated in other industrialized countries, and a third body of knowledge that has emerged from developing countries. Mays, Rubin, Sabourin, and Walker (1996) have stated: "Within this triumvirate, US psychology has been imported and serves as an important source of influence for a number of developed nations—the European communities, in particular—as well as the developing nations" (p. 485). Pawlik and d'Ydewalle (1996) have predicted the rise of international cooperation and exchange in psychology, occurring through international organizations, conferences, and publications. Gergen, Gulerce, Lock, and Misra (1996) have examined the global context of psychological practice. They delineated various indigenous issues present in India, New Zealand, and Turkey. They argue for a multicultural psychology that down plays the dominant influence of Western psychological models of theory and practice. Lunt and Poortinga (1996) have advocated the increase in European psychology studies that are based on the multicultural diversity present in European countries. The trend towards internationalizing psychology, therefore, appears to be one which will receive greater attention in the future and may have an impact on the area of cross-cultural intervention research.

10.04.8 SUMMARY

Multicultural counseling has been hailed as the fourth force in psychology (Pedersen, 1990). The area of cross-cultural interventions has evolved into a multidimensional field that encompasses several specific interest areas. First, the impact of proximal and distal variables upon behavior has been examined. This area has provided evidence that a therapist needs to understand both universal and culture-specific aspects of a client's functioning. Second, research on cross-cultural intervention processes has been reviewed. These processes include topics as diverse as counselor competence and sensitivity, communication style, and verbal and nonverbal communication. The research suggests that cross-cultural interventions can be facilitated and improved by a therapist having an understanding of the racial/ethnic/cultural differences that exist between the client and therapist. Third, the research on therapist–client match was examined. The majority of studies suggest that clients prefer ethnically/racially/culturally similar therapists and judge these therapists as more competent than dissimilar therapists. Other research, however, repudiates these notions, pointing to other therapist variables (e.g., sensitivity) as important. Fourth, the treatment outcome research has been examined. It was discovered that few methodologically sound studies exist in the literature and, also, that most studies have focused on minority groups in the USA. Fifth, methodological research concerns were explored. The weaknesses of cross-cultural studies were examined, and recommendations for future research were reviewed. Finally, this chapter has looked to future trends in cross-cultural intervention research. It predicted that terminology and ethical considerations will be more clearly delineated. In addition, future trends include the elucidation and validation of the theory of multicultural counseling and therapy, refinement and construction of culturally appropriate assessment and outcome measures, and a focus on specific within-group research topics. It is suggested that greater emphasis on global issues in psychology, and an increase in international research, will further the study of cross-cultural clinical interventions.

10.04.9 REFERENCES

Acosta, K., & Cristo, M. (1981). Development of a bilingual interpreter: An alternative model for Spanish-speaking services. *Professional Psychology, 12,* 474–482.

Acosta, K., Yamamoto, J., & Evans, L. (1982). *Effective psychotherapy for low-income and minority patients.* New York: Plenum.

Akutsu, P., Lin, C., & Zane, N. (1990). Predictors of utilization intent of counseling among Chinese and White students: A test of the proximal–distal model. *Journal of Counseling Psychology, 37*(4), 445–452.

Alger, M. (1996). *Counseling Guatemalan Mayans in South Florida.* Unpublished manuscript, Boca Raton, FL: Florida Atlantic University.

Altman, I. (1975). *The environment and social behavior.* Monterey, CA: Brooks/Cole.

Altman, A., & Chemers, M. M. (1980). *Culture and environment.* Monterey, CA: Brooks/Cole.

Alvidrez, J., Azocar, F., & Miranda, J. (1996). Demystifying the concept of ethnicity for psychotherapy researchers. *Journal of Consulting and Clinical Psychology, 64*(5), 903–908.

American Counseling Association (1988). *Ethical standards.* Alexandria, VA: Author.

American Counseling Association (1990). *Counseling the Black/African American client; Counseling the Mexican client; Counseling the Native American client; Counseling the Vietnamese client* [videocassettes]. (Available from the American Counseling Association, Alexandria, VA).

American Psychological Association (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47*(12), 1597–1611.

American Psychological Association (1993). Guidelines for providers of psychological services to ethnic, and culturally diverse populations. *American Psychologist, 48,* 45–48.

American Psychological Association (1994). *Ethnocultural psychotherapy: Comas-Diaz* [videocassette]. (Available from the American Psychological Association, Washington, DC)

Andersen, M. (1993). Studying across difference: Race, class, and gender in qualitative research. In J. Stanfield II & R. Dennis (Eds.), *Race and ethnicity in research methods* (pp. 39–52). Newbury Park, CA: Sage.

Apple, W., Streeter, L., & Krauss, R. (1979). Effects of pitch and speech rate on personal attribution. *Journal of Personality and Social Psychology, 37,* 715–727.

Arbona, C. (1995). Theory and research on racial and ethnic minorities: Hispanic Americans. In F. Leong (Ed.), *Career development and vocational behavior of racial and ethnic minorities* (pp. 37–66). Mahwah, NJ: Erlbaum.

Arean, P., & Gallagher-Thompson, D. (1996). Issues and recommendations for the recruitment and retention of older ethnic minority adults into clinical research. *Journal of Consulting and Clinical Psychology, 64*(5), 875–880.

Argyle, M. (1991). Understanding intercultural communication. In L. Samovar & R. Porter, *Intercultural communication: A reader* (6th ed., pp. 33–45). Belmont, CA: Wadsworth.

Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze.* Cambridge, UK: Cambridge University Press.

Argyle, M., & Dean, J. (1965). Eye-contact, distance, and affiliation. *Sociometry, 28,* 289–304.

Arredondo, P. (1996). MCT theory and Latina(o)-American populations. In D. W. Sue, A. Ivey, P. Pedersen (Eds.), *A theory of multicultural counseling and therapy* (pp. 217–235). Pacific Grove, CA: Brooks/Cole.

Atkinson, D. R. (1983). Ethnic similarity in counseling psychology: A review of research. *The Counseling Psychologist, 11*(3), 79–92.

Atkinson, D. R. (1985). A meta-review of research on cross-cultural counseling and psychotherapy. *Journal of Multicultural Counseling and Development, 13,* 138–153.

Atkinson, D. R. (1993). Who speaks for cross-cultural counseling research? *The Counseling Psychologist, 21*(2), 218–224.

Atkinson, D., Casas, A., & Abreu, J. (1992). Mexican-American acculturation, counselor ethnicity, and cultural sensitivity and perceived counselor competence.

Journal of Counseling Psychology, 39(4), 515–520.

Atkinson, D., & Lowe, S. (1995). The role of ethnicity, cultural knowledge, and conventional techniques in counseling and psychotherapy. In J. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander (Eds.), *Handbook of multicultural counseling* (pp. 387–414).

Atkinson, D., & Matsushita, Y. (1991). Japanese-American acculturation, counseling style, counselor ethnicity, and perceived counselor credibility. *Journal of Counseling Psychology, 38*(4), 473–478.

Atkinson, D., Morten, G., & Sue, D. (1993). *Counseling American minorities: A cross-cultural perspective* (4th ed.). Dubuque, IA: Brown and Benchmark.

Atkinson, D., & Wampold, B. (1993). Mexican Americans' initial preferences for counselors: Simple choice can be misleading: Comment on Lopez, Lopez, and Fong (1991). *Journal of Counseling Psychology, 40*(2), 245–248.

Baker, N. (1981). Social work through an interpreter. *Social Work, 26,* 391–397.

Barlow, D. (1996). The health care policy, psychotherapy research and the future of psychotherapy. *American Psychologist, 51* (10), 1050–1058.

Ballou, M. (1996). MCT theory in women. In D. W. Sue, A. Ivey, & P. Pedersen (Eds.), *A theory of multicultural counseling and therapy* (pp. 236–246). Pacific Grove, CA: Brooks/Cole.

Bemak, F., Chung, R. C., & Bornemann, T. (1996). Counseling and psychotherapy with refugees. In P. Pedersen, J. Draguns, W. Lonner, & J. Trimble (Eds.), *Counseling across cultures* (4th ed., pp. 243–265). Thousand Oaks, CA: Sage.

Bemak, F., & Greenberg, B. (1994). Southeast Asian refugee adolescents: Implications for counseling. *Journal of Multicultural Counseling and Development, 22*(4), 115–124.

Bennett, S., & Bigfoot-Sipes, D. (1991). American Indian and White college student preferences for counselor characteristics. *Journal of Counseling Psychology, 38*(4), 440–445.

Berkanovic, E. (1980). The effect of inadequate language translation on Hispanics' responses to health surveys. *American Journal of Public Health, 70*, 1273–1276.

Bernstein, B. L., Wade, P., & Hofmann, B. (1987). Students' race and preferences for counselor's race, sex, age, and experience. *Journal of Multicultural Counseling and Development, 15*, 60–70.

Berry, J. (1969). On cross-cultural comparability. *International Journal of Psychology, 4*, 119–128.

Berry, J. (1989). Imposed etics–emics–derived etics: The operationalization of a compelling idea. *International Journal of Psychology, 24*, 721–735.

Berry, J. (1990). The role of psychology in ethnic studies. *Canadian Ethnic Studies, XXII*(1), 8–21.

Berry, J., Poortinga, Y., Segall, M., Dasen, P. (1992). *Cross-cultural psychology: Research applications*. New York: Cambridge University Press.

Betancourt, H., & Lopez, S. (1993). The study of culture, ethnicity, and race in American psychology. *American Psychologist, 48*, 629–637.

Betz, N., & Fitzgerald, L. (1995). Career assessment and intervention with racial and ethnic minorities. In F. Leong (Ed.), *Career development and vocational behavior of racial and ethnic minorities* (pp. 263–279). Mahwah, NJ: Erlbaum.

Birdwhistell, R. (1970). *Kinesics and contexts*. Philadelphia, PA: University of Pennsylvania Press.

Bobo, J. K., Gilchrist, L. D., Cvetkovich, G. T., Trimble, J. E., & Schinke, S. P. (1988). Cross-cultural service delivery to minority communities. *Journal of Community Psychology, 16*, 263–272.

Bond, M. H., & Shirashi, D. (1974). The effect of body lean and status of an interviewer on the non-verbal behavior of Japanese interviewees. *International Journal of Psychology, 9*, 117–128.

Bond, M. H., & Smith, P. (1996). Cross-cultural social and organizational psychology. In J. Spence, J. Darley, & G. Foss (Eds.), *Annual review of psychology* (pp. 205–235). Palo Alto, CA: Annual Reviews.

Bowman, S. (1995). Career intervention strategies and assessment issues for African Americans. In F. Leong (Ed.), *Career development and vocational behavior of racial and ethnic minorities* (pp. 137–164). Mahwah, NJ: Erlbaum.

Bradbury, T., & Fincham, F. (1988). Individual difference variables in close relationships: A contextual model of marriage as an integrative framework. *Journal of Personality and Social Psychology, 54*(4), 713–721.

Brown, B. S., Joe, G. W., & Thompson, P. (1985). Minority group status and treatment retention. *International Journal of the Addictions, 20*, 319–335.

Brown, M. T. (1995). The career development of African Americans: Theoretical and empirical issues. In F. Leong (Ed.), *Career development and vocational behavior of racial and ethnic minorities* (pp. 7–36). Mahwah, NJ: Erlbaum.

Carillo, C. (1982). Changing norms of Hispanic families. In E. E. Jones & S. J. Korchin (Eds.), *Minority mental health* (pp. 250–256). New York: Praeger.

Carkhuff, R., & Pierce, R. (1967). Differential effects of therapist race and social class upon patient depth of self-exploration in the initial clinical interview. *Journal of Consulting Psychology, 31*, 632–634.

Carter, R. (1991). Cultural values: A review of empirical research and implications for counseling. *Journal of Counseling and Development, 70*(1), 164–173.

Casas, J., & Mann, D. (1996). MCT theory and implications for research. In D. Sue, A. Ivey, & P. Pedersen (Eds.), *A theory of multicultural counseling and therapy* (pp. 99–111). Pacific Grove, CA: Brooks/Cole.

Casas, J., & SanMiguel, S. (1993). Beyond questions and discussions, there is a need for action: A response to Mio and Iwamasa. *The Counseling Psychologist, 21*(2), 233–239.

Champagne, D. (1993). Toward a multidimensional historical comparative methodology: Context, process, and causality. In J. Stanfield II & R. Dennis (Eds.), *Race and ethnicity in research methods* (pp. 233–253). Newbury Park, CA: Sage.

Cheng, C. Y. (1987). Chinese philosophy and contemporary communication theory. In D. L. Kincaid (Ed.), *Communication theory: Eastern and western perspectives.* (pp. 39–58). New York: Academic Press.

Coleman, H. L., Wampold, B. E., & Casali, S. L. (1995). Ethnic minorities' ratings of ethnically similar and European American counselors: A meta-analysis. *Journal of Counseling Psychology, 42*(1), 55–64.

Comas-Diaz, L., & Griffith, E. (Eds.) (1988). *Clinical guidelines in cross-cultural mental health*. New York: Wiley.

Comas-Diaz, L., & Jacobsen, F. (1987). Ethnocultural identification in psychotherapy. *Psychiatry, 50*, 232–241.

Constantino, G., Malgady, R., & Rogler, L. (1986). Cuento therapy: A culturally sensitive modality for Puerto Rican children. *Journal of Counseling and Clinical Psychology, 54*, 639–645.

Corey, G. (1996). Theoretical implications of MCT theory. In D. W. Sue, A. Ivey, & P. Pedersen (Eds.), *A theory of multicultural counseling and therapy* (pp. 99–111). Pacific Grove, CA: Brooks/Cole.

Cross, W. Jr. (1995). The psychology of nigrescence: Revising the Cross model. In J. Ponterotto, J. Casas, L. Suzuki, & C. Alexander (Eds.), *Handbook of multicultural counseling* (pp. 93–122). Thousand Oaks, CA: Sage.

Crystal, D. (1975). Paralinguistics. In J. Benthall & T.

Polhemus (Eds.), *The body as a medium of expression* (pp. 49–74). New York: Dutton.

Dana, R. (1993). *Multicultural assessment perspectives for professional psychology*. Boston: Allyn & Bacon.

D'Andrea, M., Daniels, J., & Heck, R. (1991). Evaluating the impact of multicultural training. *Journal of Counseling and Development, 70,* 143–150.

Daniels, J. & D'Andrea, M. (1996). MCT theory and ethnocentrism in counseling. In D. W. Sue, A. Ivey, & P. Pedersen (Eds.), *A theory of multicultural counseling and therapy* (pp. 155–173). Pacific Grove, CA: Brooks/Cole.

de Figueiredo, J. (1976). Interviewing in Goa: Methodological issues in the study of a bilingual culture. *Journal of Social Science and Medicine, 10,* 503–508.

de Figueiredo, J., & Lemkau, P. (1980). Psychiatric interviewing across cultures: Some problems and prospects. *Social Psychiatry, 15,* 117–121.

Dennis, R. (1993). Participant observations. In J. Stanfield II & R. Dennis (Eds.), *Race and ethnicity in research methods* (pp. 53–74). Newbury Park, CA: Sage.

Der-Kerabetian, A. (1992). World-mindedness and the nuclear threat: Multinational study. *Journal of Social and Behavioral Personality, 7,* 293–303.

Dignes, N., Trimble, J., Manson, S., & Pasquale, F. (1981). Counseling and psychotherapy with American Indians and Alaska Natives. In A. Marsella & P. Pedersen (Eds.), *Cross-cultural counseling and psychotherapy: Foundations, evaluation and cultural considerations* (pp. 243–276). Elmsford, NY: Pergamon.

Dillard, J. M. (1983). *Multicultural counseling*. Chicago: Nelson.

Draguns, J. G. (1981). Cross-cultural counseling and psychotherapy: History, issues, current status. In A. J. Marsella & P. B. Pedersen (Eds.), *Cross-cultural counseling and psychotherapy* (pp. 3–27). New York: Pergamon.

Dwairy, M., & Van Sickle, T. (1996). Western psychotherapy in traditional Arabic societies. *Clinical Psychology Review, 16*(3), 231–249.

Eakins, B., & Eakins, R. (1985). Sex differences in nonverbal communication. In L. Samovar & R. Porter (Eds.), *Intercultural communication: A reader* (pp. 192–217). Belmont, CA: Wadsworth.

Edgerton, R. B., & Karno, M. (1971). Mexican-American biligualism and the perception of mental illness. *Archives of General Psychiatry, 24,* 286–290.

Efron, D. (1972). Gesture, race and culture. In T. A. Sebeok (Ed.), *Approaches to semibiotics* (pp. 44–59). The Hague, The Netherlands: Mouton.

Eisenbruch, M. (1990). The Cultural Bereavement Interview: A new clinical research approach for refugees. *Psychiatric Clinics of North America, 13*(4), 715–735.

Ekman, P. (1993). Facial expression and emotion. *American Psychologist, 48*(4), 384–392.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial expressions*. Englewood Cliffs, NJ: Prentice-Hall.

Ellgring, H. (1984). The study of nonverbal behavior and its applications: State of the art in Europe. In A. Wolfgang (Ed.), *Nonverbal behavior* (pp. 115–138). Lewiston, NY: Hogrefe.

Erickson, F., & Shultz, J. (1982). *The counselor gatekeeper: Social interaction in interviews*. New York: Academic Press.

Exline, R. A., Gray, D., & Schuette, D. (1985). Visual behavior in a dyad as affected by interview content and sex of the respondent. *Journal of Personality and Social Psychology, 1,* 201–209.

Facio, E. (1993). Ethnography as personal experience. In J. Stanfield II & R. Dennis (Eds.), *Race and ethnicity in research methods* (pp. 75–93). Newbury Park, CA: Sage.

Fernando, S. (Ed.) (1995). *Mental health in a multi-ethnic society*. London: Routledge.

Feuerstein, R. (1979). *The dynamic assessment of retarded performers*. Baltimore, MD: University Park Press.

Fiksdal, S. (1990). *The right time and pace: A microanalysis of cross-cultural gatekeeping interviews*. Norwood, NJ: Ablex.

Fiske, A. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review, 99,* 689–723.

Fiske, S., & Taylor, S. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.

Flaskerud, J. (1990). Matching client and therapist ethnicity, language, and gender: A review of research. *Issues in Mental Health Nursing, 11*(4), 321–336.

Flaskerud, J., & Hu, L. (1994). Participation in and outcome of treatment for major depression among low income Asian-Americans. *Psychiatry Research, 53*(3), 289–300.

Flaskerud, J., & Liu, P. (1990). Influence of therapist ethnicity and language on therapy outcomes of Southeast Asian clients. *International Journal of Social Psychiatry, 36*(1), 18–29.

Flaskerud, J., & Liu, P. (1991). Effects of an Asian client–therapist language, ethnicity, and gender match on utilization and outcome of therapy. *Community Mental Health Journal, 27*(1), 31–42.

Fletcher, G., & Fitness, J. (1990). Occurrent social cognition in close relationship interaction: The role of proximal and distal variables. *Journal of Personality and Social Psychology, 59*(3), 464–474.

Forrest, D., Ryan, J., & Lazar, V. (1978). American familiar language and the FMG psychiatric resident. *The Journal of Psychiatric Education, 2*(1), 68–81.

Fouad, N. A. (1995). Career behavior of Hispanics: Assessment and career intervention. In F. Leong (Ed.), *Career development and vocational behavior of racial and ethnic minorities* (pp. 165–191). Mahwah, NJ: Erlbaum.

Frank, J. D. (1961). *Persuasion and healing: A comparative study of psychotherapy*. Baltimore, MD: Johns Hopkins University Press.

Freed, A. O. (1988). Interviewing through an interpreter. *Social Work, 46,* 315–319.

Fujino, D., Okazaki, S., & Young, K. (1994). Asian-American women in the mental health system: An examination of ethnic and gender match between therapist and client. *Journal of Community Psychology, 22*(2), 164–176.

Furnham, A. (1993). Just world beliefs in twelve societies. *Journal of Social Psychology, 133,* 317–329.

Furnham, A., Bond, M., Heaven, P., Hilton, D., & Lobel, T. (1993). A comparison of Protestant work ethic beliefs in thirteen nations. *Journal of Social Psychology, 133,* 185–197.

Gamboa, A., Tosi, D., & Riccio, A. (1976). Race and counselor climate in the counselor preference of delinquent girls. *Journal of Counseling Psychology, 23,* 160–162.

Garza, A. C. (1981). Potential pitfalls in the diagnosis and treatment of minority groups. *The Journal of Social Psychology, 114,* 9–22.

Gaw, A. C. (Ed.) (1993). *Culture, ethnicity, and mental illness*. Washington, DC: American Psychiatric Press.

Gergen, K. J., Gulerce, A., Lock, A., & Misra, G. (1996). Psychological science in cultural context. *American Psychologist, 51*(5), 496–503.

Giles, H. C. (1990). Counseling Haitian students and their families: Issues and interventions. *Journal of Counseling and Development, 68,* 317–320.

Gim, R. H., Atkinson, D., & Kim, S. (1991). Asian American acculturation, counselor ethnicity, and cultural sensitivity and ratings of counselors. *Journal of Counseling Psychology, 38,* 57–62.

Goldberg, B., & Tidwell, R. (1990). Ethnicity and gender similarity: The effectiveness of counseling for adolescents. *Journal of Youth and Adolescents, 19*(6), 598–603.

Goldfried, M., & Wolfe, B. (1996). Psychotherapy practice and research: Repairing a strained alliance. *American Psychologist, 51*(10), 1007–1016.

Gonzalez-Lee, T., & Simon, H. J. (1990). *Medical Spanish: Interviewing the Latino patient: A cross-cultural perspective*. Englewood Cliffs, NJ: Prentice-Hall.

Goodchilds, J. (Ed.) (1991). *Psychological perspectives on human diversity in America*. Washington, DC: American Psychological Association.

Gottheil, E., Sterling, R., & Weinstein, S. (1995). Generalizing from controlled treatment outcome studies: Sample data from a cocaine treatment program. *American Journal on Addictions, 4*(4), 331–338.

Gottheil, E., Sterling, R., Weinstein, S., & Kurtz, J. (1994). Patient/therapist matching and early treatment dropout. *Journal of Addictive Diseases, 13*(4), 169–176.

Graham, J. L. (1985). The influence of culture on business negotiations. *Journal of International Business Studies, 16,* 81–96.

Greene, C. J. (1982). *A study of community-college students' initial counselor preferences based on the nature of the personal problem and on the age, ethnicity, and sex of the counselor and student*. Unpublished doctoral dissertation, University of San Francisco.

Griffith. M. S., & Jones, E. E. (1978). Race and psychotherapy: Changing perspectives. In J. H. Masserman (Ed.), *Current psychiatric therapies* (Vol. 4, pp. 225–235). New York: Grune & Stratton.

Hall, E. T. (1959). *The silent language*. Garden City, NJ: Doubleday.

Hall, E. T. (1969). *The hidden dimension*. Garden City, NJ: Doubleday.

Hall, E. T. (1976). *Beyond culture*. Garden City, NJ: Anchor Press.

Hall, G., & Maloney, H. (1983). Cultural control in psychotherapy with minority clients. *Psychotherapy: Theory, Research, & Practice, 20*(2), 131–142.

Hanna, J. (1984). Black/white nonverbal differences, dance, and dissonance: Implications for desegregation. In A. Wolfgang (Ed.), *Nonverbal behavior: Perspectives, applications, intercultural insights* (pp. 373–410). Lewiston, NY: Hogrefe.

Harper, R. G., Wiens, A. N., & Matarazzo, J. D. (1978). *Nonverbal communication: The state of the art*. New York: Wiley.

Harrison, D. K. (1975). Race as a counselor–client variable in counseling and psychotherapy: A review of the research. *The Counseling Psychologist, 5,* 124–133.

Helms, J. E. (1993). I also said, "White racial identity influences White researchers". *The Counseling Psychologist, 21*(2), 240–243.

Helms, J. E. (1994). How multiculturalism obscures racial factors in the therapy process: Comment on Ridley et al. (1994), Sodowsky et al. (1994), Ottavi et al. (1994), and Thompson et al. (1994). *Journal of Counseling Psychology, 41*(2), 162–165.

Henley, N. (1977). *Body politics: Power, sex, and nonverbal communication*. Englewood Cliffs, NJ: Prentice-Hall.

Henley, N., & LaFrance, M. (1984). Gender as culture: Difference and dominance in nonverbal behavior. In A. Wolfgang (Ed.), *Nonverbal behavior*: (pp. 351–372). Lewiston, NY: Hogrefe.

Hernandez, A., & LaFromboise, T. (1985, August). *The development of the Cross-Cultural Counseling Inventory*. Paper presented at the 93rd Annual Convention of the American Psychological Association, Los Angeles, CA.

Hess, R., & Street, E. (1991). The effect of acculturation on the relationship of counselor ethnicity and client ratings. *Journal of Counseling Psychology, 38*(1), 71–75.

Hewes, G. (1955). World distribution of certain postural habits. *American Anthropologist, 52*(2) 231–244.

Hiegel, J. P. (1994). Use of indigenous concepts and healers in the care of refugees: Some experiences from the Thai border camps. In A. J. Marsella, T. Bornemann, S. Ekblad, & J. Orley (Eds.), *Amidst peril and pain: The mental health and well-being of the world's refugees* (pp. 293–310). Washington, DC: American Psychological Association.

Higginbotham, H. N. (1977). Culture and the role of client expectancy in psychotherapy. *Topics in Culture Learning, 5,* 107–124.

Higginbotham, H. N., West, S., & Forsyth, D. (1988). *Psychotherapy and behavior change: Social, cultural and methodological perspectives*. New York: Pergamon.

Highlen, P. (1996). MCT theory and implications for organizations/systems. In D. W. Sue, A. Ivey, & P. Pedersen (Eds.), *A theory of multicultural counseling and therapy* (pp. 65–85). Pacific Grove, CA: Brooks/Cole.

Ho, M. (1987). *Family therapy with ethnic minorities*. Newbury Park, CA: Sage.

Ho, M. (1992). *Minority children and adolescents in therapy*. Newbury Park, CA: Sage

Hofstede, G. (1980). *Culture's consequences: International differences in work related values*. Beverly Hills, CA: Sage.

Hofstede, G. (1991). *Cultures and organizations: Software of the mind*. London: McGraw-Hill.

Hohmann, A., & Parron, D. (1996). How the new NIH guidelines on inclusion of women and minorities apply: Efficacy trials, effectiveness trials and validity. *Journal of Consulting and Clinical Psychology, 64*(5), 851–855.

Hollingsworth, D. M. (1987). *The interview: A cross-cultural model, strategies and evaluative measures*. Unpublished doctoral dissertation, University of Arizona.

Hollon, S. (1996). The efficacy and effectiveness of psychotherapy relative to medications. *American Psychologist, 51*(10), A1020–1025.

Howard, K., Moras, K., Brill, P., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist, 51*(10), 1059–1064.

Hui, C., & Triandis, H. (1985). Measurement in cross-cultural psychology: A review and comparison of strategies. *Journal of Cross-Cultural Psychology, 16(2),* 131–152.

Ibrahim, F. A. (1985). Effective cross-cultural counseling and psychotherapy. *The Counseling Psychologist, 13,* 625–638.

Ivey, A. E. (1981). Counseling and psychotherapy: Toward a new perspective. In A. Marsella & P. Pedersen (Eds.), *Cross-cultural counseling and psychotherapy* (pp. 279–311). New York: Pergamon.

Ivey, A. E. (1993). On the need for reconstruction of our present practice of counseling and psychotherapy. *The Counseling psychologist, 21*(2), 225–228.

Ivey, A. E. (1994). *Intentional interviewing in counseling: Facilitating client development*. Pacific Grove, CA: Brooks/Cole.

Ivey, A. E., Ivey, N. B., & Simek-Morgan, L. (1993). *Counseling and psychotherapy* (3rd ed.). Boston: Allyn & Bacon.

Izard, C. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.

Izard, C., & Izard, B. (1980). Expressions of emotions as transcultural language in social interaction and theatrical performances. In W. von Raffer-Engel (Ed.), *Aspects of nonverbal communication* (pp. 78–102). Lisse, the Netherlands: Swets & Zeitlinger.

Jackson, K., & Moto, F. (1996, August). *Mental health concepts in Chichewa and English*. Poster session presented at the XIIIth Congress of the International Association of Cross-Cultural Psychology, Montreal, PQ, Canada.

Jackson, M. (1995). Multicultural counseling: Historical perspectives. In J. Ponterotto, J. Casas, L. Suzuki, & C. Alexander (Eds.), *Handbook of multicultural counsel-*

ing (pp. 3–16). Thousand Oaks, CA: Sage.

Jacobson, N., & Christensen, A. (1996). Studying the effectiveness of psychotherapy: How well can clinical trials do the job? *American Psychologist, 51*(10), 1031–1039.

Jensen, J. V. (1985). Perspective on non-verbal communication. In L. A. Samovar & R. E. Porter (Eds.), Intercultural communication: A reader (pp. 165–189). Belmont, CA: Wadsworth.

Johnson, D. R. (1972). Black kinesics: Some nonverbal communication patterns in the black culture. In L. A. Samovar & R. E. Porter (Eds.), *Intercultural communication: A reader* (pp. 231–254). Belmont, CA: Wadsworth.

Johnson, M., & Lashley, K. (1989). Influence of Native-Americans' cultural commitment on preferences for counselor ethnicity and expectations about counseling. *Journal of Multicultural Counseling and Development, 17*(3), 115–122.

Johnson, M. J., Swartz, J., & Martin, W. Jr. (1995). Applications of psychological theories for career development with Native Americans. In F. Leong (Ed.), *Career development and vocational behavior of racial and ethnic minorities* (pp. 103–133). Hillsdale, NJ: Erlbaum.

Jones, E. E. (1978). Effects of race on psychotherapy process and outcome: An exploratory investigation. *Psychotherapy: Theory, Research, and Practice, 15*(3), 226–236.

Jones, E. E. (1982). Psychotherapists' impressions of treatment outcome as a function of race. *Journal of Clinical Psychology, 38*(4), 722–731.

Jones, J. (1991). Psychological models of race: What have they been and what should they be? In J. Goodchilds (Ed.), *Psychological perspectives on human diversity in America* (pp. 7–46). Washington, DC: American Psychological Association.

Jordan, J. (1995). Report from the Bahamas. In N. R. Goldberger & J. B. Veroff (Eds.), *The culture and psychology reader* (pp. 605–617). New York: New York University Press.

Kagitcibasi, C., & Berry, J. W. (1989). Cross-cultural psychology: Current research and trends. *Annual Review of Psychology, 40,* 493–451.

Kelly, H. H., Berscheid, E., Christensen, A., Harvey, J. H., Huston, T. L., Levinger, G., McClintock, E., Peplau, L. A., & Peterson, D. (1983). *Close relationships*. San Francisco: Freeman.

Kendon, A. (1983). Gesture and speech: How they interact. In J. M. Weimann, & R. P. Harrison (Eds.), *Nonverbal interaction*. Beverly Hills, CA: Sage.

Kerwin, C., & Ponterotto, J. (1995). Biracial identity development: Theory and research. In J. Ponterotto, J. Casas, L. Suzuki, & C. Alexander (Eds.), *Handbook of multicultural counseling* (pp. 199–217). Thousand Oaks, CA: Sage.

Key, M. L. (1975). *Paralanguage and kinesics*. Metuchen, NJ: Scarecrow Press.

Kim, U., & Berry, J. (Eds.) (1993). *Indigenous psychologies: Research and experience in context*. Newbury Park, CA: Sage.

Kim, U., Triandis, H., Kagitcibasi, C., Choi, S., & Yoon, G. (Eds.) (1994). *Individualism and collectivism: Theory, method and applications*. Newbury Park, CA: Sage.

Kim, Y. Y. (1988). *Communication and cross-cultural adaptation*. Clevedon, UK: Multilingual Matters.

Kinzie, J. D, & Leung, P. (1989). Clonidine in Cambodian patients with post-traumatic stress disorder. *Journal of Nervous and Mental Disease, 177*, 546–550.

Kline, F., Acosta, F., & Austin, V., & Johnson, R. (1980). The misunderstood Spanish-speaking patient. *American Journal of Psychiatry, 137*(12), 1530–1533.

Kluckhohn, F. R., & Strodtbeck, F. L. (1961). *Variations in value orientation*. New York: Harper & Row.

Knapp, M. (1984). The study of nonverbal behavior vis-à-vis human communication theory. In A. Wolfgang (Ed.), *Nonverbal behavior* (pp. 15–40). Lewiston, NY: Hogrefe.

Kohls, L. R. (1979). *Survival kit for overseas living*. Chicago: Intercultural Press.

Krause, I., & Miller, M. C. (1995). Culture and family therapy. In S. Fernando (Ed.), *Mental health in a multi-ethnic society* (pp. 148–171). New York: Routledge.

Krauss, R. M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 389–450). San Diego, CA: Academic Press.

LaBarre, W. (1985). Paralinguistics, kinesics and cultural anthropology. In L. Samovar & R. Porter (Eds.), *Intercultural communication: A reader* (pp. 172–189). Belmont, CA: Wadsworth.

Labov W., & Fanshel, D. (1977). *Therapeutic discourse*. New York: Academic Press.

LaFrance, M., & Mayo, C. (1978). Cultural aspects on nonverbal communication: A review essay. *International Journal of Intercultural Relations, 2*, 71–89.

LaFromboise, T., Coleman, H., & Hernandez, A. (1991). Development and factor structure of the Cross-Cultural Counseling Inventory–Revised. *Professional Research and Practice, 22*, 380–388.

LaFromboise, T., Foster, S., & James, A. (1996). Ethics in multicultural counseling. In P. B. Pedersen, J. G. Draguns, W. J. Lonner, & J. E. Trimble (Eds.), *Counseling across cultures* (4th ed., pp. 47–72). Thousand Oaks, CA: Sage.

LaFromboise, T., & Jackson, M. (1996). MCT theory and Native-American populations. In D. W. Sue, A. Ivey, & P. Pedersen (Eds.), *A theory of multicultural counseling and therapy* (pp. 192–203). Pacific Grove, CA: Brooks/Cole.

Lass, N. J., Mertz P. J., & Kimmel, K. (1978). The effect of temporal speech alterations on speaker race and sex identification. *Language and Speech, 21*(3), 279–290.

Leary, K. (1995). "Interpreting in the dark": Race and ethnicity in psychoanalytic psychotherapy. *Journal of Psychoanalytic Psychology, 12*(1), 127–140.

Lee, C. C. (1996). MCT theory and implications for indigenous healing. In D. W. Sue, A. Ivey, & P. Pedersen (Eds.), *A theory of multicultural counseling and therapy* (pp. 86–98). Pacific Grove, CA: Brooks/Cole.

Lefley, H. (1994). Mental health treatment and service delivery in cross-cultural perspective. In L. Adler & U. Gielen (Eds.), *Cross-cultural topics in psychology* (pp. 179–199). Westport, CT: Praeger.

Leong, F. T. (Ed.) (1995). *Career development and vocational behavior of racial and ethnic minorities*. Mahwah, NJ: Erlbaum.

Leong, F. T. (1996). MCT theory and Asian-American populations. In D. W. Sue, A. Ivey, & P. Pedersen (Eds.), *A theory of multicultural counseling and therapy* (pp. 204–216). Pacific Grove, CA: Brooks/Cole.

Leong, F. T., & Gim-Chung, R. (1995). Career assessment and intervention with Asian Americans. In F. Leong (Ed.), *Career development and vocational behavior of racial and ethnic minorities* (pp. 193–226). Mahwah, NJ: Erlbaum.

Leong, F. T., & Serafica, F. (1995). Career development of Asian Americans: A research area in need of a good theory. In F. Leong (Ed.), *Career development and vocational behavior of racial and ethnic minorities* (pp. 67–102). Mahwah, NJ: Erlbaum.

Lerner, B. (1972). *Therapy in the ghetto: Political impotence and personal disintegration*. Baltimore, MD: Johns Hopkins University Press.

LeVine, R., & Campbell, D. (1972). *Ethnocentrism: Theories of conflict, ethnic attitudes and group behavior*. New York: Wiley.

Lonner, W. (1990). An overview of cross-cultural testing and assessment. In R. W. Brislin (Ed.), *Applied cross-cultural psychology* (pp. 56–76). Newbury Park, CA: Sage.

Lonner, W., & Ibrahim, F. (1996). Appraisal and assessment in cross-cultural counseling. In P. Pedersen, J. Draguns, W. Lonner, & J. Trimble (Eds.), *Counseling across cultures* (4th ed., pp. 223–267). Thousand Oaks, CA: Sage.

Lopez, S., & Hernandez, P. (1987). When culture is considered in the evaluation and treatment of Hispanic patients. *Psychotherapy, 24*, 120–126.

Lopez, S., Lopez, A., & Fong, K. (1991). Mexican Americans' initial preferences for counselors: The role of ethnic factors. *Journal of Counseling Psychology, 38*(4), 487–496.

Lunt, I., & Poortinga, Y. H. (1996). Internationalizing psychology: The case of Europe. *American Psychologist, 51*(5), 504–508.

Madrid, A. (1995). Diversity and its discontents. In N. R. Goldberger & J. B. Veroff (Eds.), *The culture and psychology reader* (pp. 617–626). New York: New York University Press.

Malgady, R., Rogler, L., & Constantino, G. (1987). Ethnocultural and linguistic bias in mental health evaluation of Hispanics. *American Psychologist, 42*, 228–234.

Malgady, R., Rogler, L., & Constantino, G. (1990). Culturally sensitive psychotherapy for Puerto Rican children and adolescents: A program of treatment outcome research. *Journal of Counseling and Clinical Psychology, 58*, 704–712.

Manson, S. M. (1982). *New directions in prevention among American Indian and Alaskan Native communities*. Portland, OR: Oregon Health Sciences University.

Manson, S. M., Shore, J. H., & Bloom, J. D. (1985). The depressive experience in American Indian communities: A challenge for psychiatric theory and diagnoses. In A. Kleinman & B. Good (Eds.), *Culture and depression* (pp. 331–368). Berkeley, CA: University of California Press.

Marcos, L. (1976). Bilinguals in psychotherapy: Language as an emotional barrier. *American Journal of Psychotherapy, 30*, 552–560.

Marcos, L., Urcuyo, L., Kesselman, M., & Alpert, M. (1973). The language barrier in evaluating Spanish-American patients. *Archives of General Psychiatry, 29*(5) 655–659.

Marks, C. (1993). Demography and race. In J. Stanfield II & R. Dennis (Eds.), *Race and ethnicity in research methods* (pp. 159–171). Newbury Park, CA: Sage.

Marsella, A. J., & Pedersen, P. B. (Eds.) (1981). *Cross-cultural counseling and psychotherapy*. New York: Pergamon.

Marsella, A. J., & White, G. M. (Eds.) (1982). *Cultural conceptions of mental health and therapy*. Dordrecht, The Netherlands: Reidel.

Mays, V. M., Rubin, J., Sabourin, M., & Walker, L. (1996). Moving toward a global psychology: Changing theories and practice to meet the needs of a changing world. *American Psychologist, 51*(5), 485–487.

McGoldrick, M., Pearce J., & Giordano, J. (Eds) (1982). *Ethnicity and family therapy*. New York: Guilford.

McWhirter, J. J., & Ryan, C. (1991). Counseling the Navajo: Cultural understanding. *Journal of Multicultural Counseling and Development, 19*, 74–82.

Mehrabian, A. (1972). *Nonverbal communication*. Chicago, IL: Aldine-Atherton.

Mehrabian, A., & Ferris, S. R. (1976). Influence of attitudes from nonverbal communication in two channels. *Journal of Consulting Psychology, 31*, 248–252.

Mio, J. S., & Iwamasa, G. (1993). To do, or not to do: That is the question for White cross-cultural researchers. *The Counseling Psychologist, 21*(2), 197–212.

Miranda, J. (1996). Introduction to the special section on recruiting and retaining minorities in psychotherapy research. *Journal of Consulting and Clinical Psychology, 64*(5), 848–850.

Miranda, J., Azocar, F., Organista, K., Munoz, R., & Lieberman, A. (1996). Recruiting and retaining low-income Latinos in psychotherapy research. *Journal of Consulting and Clinical Psychology, 64*(5), 868–874.

Mitchell, H. (1970). The Black experience in higher education. *The Counseling Psychologist, 2*, 30–36.

Moghaddam, F. (1987). Psychology in the three worlds: as reflected by the crisis in social psychology and the move towards indigenous Third World psychology. *American Psychologist, 42*(10), 912–920.

Mollica, R. F., Wyshak, G., Lavelle, J., Truong, T., Tor, S., & Yang, T. (1990). Assessment symptom change in Southeast Asian refugee survivors of mass violence and torture. *American Journal of Psychiatry, 147*(1), 83–88.

Montalvo, G., & Gutierrez, M. (1989). Nine assumptions for work with ethnic minority families. *Journal of Psychotherapy and the Family, 6*(1–2), 35–52.

Moraga, C. (1995). La guera. In N. R. Goldberger & J. B. Veroff (Eds.), *The culture and psychology reader* (pp. 596–604). New York: New York University Press.

Morris, D., Collette, P., Marsh, P., & O'Shaughnessy, M. (1979). *Gestures: Their origins and distributions*. London: Cape.

Mowrer, O. (Ed.) (1953). *Psychotherapy: Research and practice*. New York: Ronald Press.

Mrinal, N., Mrinal, U., & Takooshian, H. (1994). Research methods for studies in the field. In L. Adler & U. Gielen (Eds.), *Cross-cultural topics in psychology* (pp. 25–40). Westport, CT: Praeger/Greenwood.

Myers, S. (1993). Measuring and detecting discrimination in the post-civil rights era. In J. Stanfield II & R. Dennis (Eds.), *Race and ethnicity in research methods* (pp. 172–197). Newbury Park, CA: Sage.

Neligh, G. (1988). Major mental disorders and behavior among American Indians and Alaska Natives [Monograph]. In *Behavioral health issues among American Indians and Alaska Natives: Explorations on the frontiers of the biobehavioral sciences. American Indian and Alaska Mental Health Research, 1*, 116–159.

Newman, F., & Tejeda, M. (1996). The need for research that is designed to support decisions in the delivery of mental health services, *American Psychologist, 51*(10), 1040–1049.

Nguyen, N. (1992). Living between two cultures: Treating first-generation Asian Americans. In L. Vargas & J. Koss-Chioino (Eds.), *Working with culture* (pp. 204–224). San Francisco: Jossey-Bass.

Nichter, M. (1981). Idioms of distress, alternatives in the expression of psychosocial distress: A case study from South India. *Culture and Medical Psychiatry, 5*, 379–408.

National Institutes of Health (1994). NIH guidelines on the inclusion of women and minorities as subjects in clinical research. 59, Fed. Reg. 14, 508 (Document no. 94-5435).

Norton, I., & Manson, S. (1996). Research in American Indian and Alaska Native communities: Navigating the cultural universe of values and process. *Journal of Consulting and Clinical Psychology, 64*(5), 856–860.

Ogino, T., Misono, Y., & Fukushima, C. (1985). Diversity of honorific usage in Tokyo: a sociolinguistic approach based on a field survey. *International Journal of Sociology of Language, 55*, 23–39.

Orlinski, D., Grawe, K., & Parks, P. (1994). Process and outcome in psychotherapy–noch einmal. In A. Bergin & S. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 270–376). New York: Wiley.

Osipow, S., & Littlejohn, E. (1995). Toward a multicultural theory of career development: Prospects and dilemmas.

In F. Leong (Ed.), *Career development and vocational behavior of racial and ethnic minorities* (pp. 251–261). Mahwah, NJ: Erlbaum.

Owan, T. C. (Ed.) (1985). *Southeast Asian mental health: Treatment, prevention, services, training, and research*. Rockville, MD: National Institute for Mental Health.

Pakov, T., Lewis, D., & Lyons, J. (1989). Psychiatric diagnosis and racial bias: An empirical investigation. *Professional Psychology: Research and Practice, 20,* 364–368.

Pandey, J. (1978). Effects of crowding on human social behavior. *Journal of Social and Economic Studies, 6,* 85–95.

Pandey, J. (1990). The environment, culture, and behavior. In R. Brislin (Ed.), *Applied cross-cultural psychology* (Vol. 14, pp. 254–277). Newbury Park, CA: Sage.

Pandey, J. (1996, August). *Perceptions of crowding in India*. Paper presented at the XIII Congress of the International Association for Cross-Cultural Psychology, Montreal, PQ.

Paniagua, F. A. (1994). *Assessing and treating culturally diverse groups: A practical guide*. Thousand Oaks, CA: Sage.

Parham, T. (1993). White researchers conducting multicultural counseling research: Can their efforts be "mo betta"? *The Counseling Psychologist, 21*(2), 250–256.

Parham, T. (1996). MCT theory and African-American populations. In D. W. Sue, A. Ivey, P. Pedersen (Eds.), *A theory of multicultural counseling and therapy* (pp. 177–191). Pacific Grove, CA: Brooks/Cole.

Parham, T., & Helms, J. (1981). The influence of Black student's racial identity attitudes on preference for counselor's race. *Journal of Counseling Psychology, 28*(3) 250–257.

Parham, T., & Helms, J. (1985). Relation of racial identity attitudes to self-actualization and affective states of Black students. *Journal of Counseling Psychology, 32*(3), 431–440.

Parloff, M., Waskow, E., & Wolfe, B. (1978). Research on therapist variables in relation to process and outcome. In S. Garfield & A. Bergin (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis* (pp. 233–282). New York: Wiley.

Patton, J. (1993). Psychoeducational assessment of gifted and talented African Americans. In J. Stanfield II & R. Dennis (Eds.), *Race and ethnicity in research methods* (pp. 198–216). Newbury Park, CA: Sage.

Pawlik, K., & d'Ydewalle, G. (1996). Psychology and the global commons: Perspectives of international psychology. *American Psychologist, 51*(5), 488–495.

Pearson, J. C. (1985). *Gender and communication*. Dubuque, IA: Brown.

Pedersen, P. B. (1990). The multicultural perspective as a fourth force in counseling. *Journal of Mental Health Counseling: 12*(1), 93–95.

Pedersen, P. B. (1991a). Multiculturalism as a generic approach to counseling. *Journal of Counseling and Development: 70*(1), 6–12.

Pedersen, P. B. (Ed.) (1991b). Multiculturalism as a fourth force in counseling (Special issue). *Journal of Counseling and Development 70*(1).

Pedersen, P. B. (1991c). Complexity and balance as criteria of effective multicultural counseling. *Journal of Counseling and Development 68*, 550–554.

Pedersen, P. B. (1993). The multicultural dilemma of White cross-cultural researchers. *The Counseling Psychologist, 21*(2), 229–232.

Pedersen, P. B. (1994). *A handbook for developing multicultural awareness* (2nd ed.). Alexandria, VA: American Psychological Association.

Pedersen, P. B. (1995). Culture-centered ethical guidelines for counselors. In J. Ponterotto, J. Casas, L. Suzuki, & C. Alexander (Eds.), *Handbook of multicultural counseling* (pp. 34–49). Thousand Oaks, CA: Sage.

Pedersen, P. B., & Ivey, A. (1993). *Culture-centered counseling and interviewing skills*. Westport, CT: Greenwood/Praeger.

Pedersen, P. B., & Lefley, H. P. (Eds.) (1986). *Cross-cultural training for mental health professionals*. Springfield, IL: Thomas.

Pennington, D. L. (1979). Black–White communication: An assessment of research. In M. Asante, E. Newmark, & C. Balke (Eds.), *The handbook of intercultural communication* (pp. 383–401). Beverly Hills, CA: Sage.

Phinney, J. (1996). When we talk about American ethnic groups, what do we mean? *American Psychologist, 51*(9), 918–927.

Pike, K. (1967). *Language in relation to a unified theory of the structure of human behavior*. The Hague, The Netherlands: Mouton.

Pitta, P., Marcos, L., & Alpert, M. (1978). Language switching as a treatment strategy with bilingual patients. *American Journal of Psychoanalysis, 38*(3), 255–258.

Ponterotto, J. (1993). White racial identity and the counseling professional. *The Counseling Psychologist, 21*(2), 213–217.

Ponterotto, J., & Casas, M. (1991). *Handbook of racial/ethnic minority counseling research*. Springfield, IL: Thomas.

Ponterotto, J., Sanchez, C., & Magids, D. (1991, August). *Initial development and validation of the Multicultural Awareness Scale*. Paper presented at the 99th Annual Convention of the American Psychological Association, San Francisco, CA.

Poortinga, Y. (1992). Towards a conceptualization of culture for psychology. In S. Iawaki, Y. Kashima, & K. Leung (Eds.), *Innovations in cross cultural psychology* (pp. 3–17). Amsterdam: Swets and Zeitlinger.

Pope-Davis, D., & Constantine, M. (1996). MCT theory and implications for practice. In D. W. Sue, A. Ivey, & P. Pedersen (Eds.), *A theory of multicultural counseling and therapy* (pp. 112–122). Pacific Grove, CA: Brooks/Cole.

Poston, W. S., Craine, M., & Atkinson, D. L. (1991). Counselor dismilarity confrontation, client cultural mistrust, and willingness to self-disclose. *Journal of Multicultural Counseling and Development, 19*(2), 65–73.

Poyatos, F. (1976). *Man beyond words: Theory and methodology of nonverbal communication*. Oswego, NY: New York State English Council.

Poyatos, F. (1984). Linguistic fluency and verbal–nonverbal cultural fluency. In A. Wolfgang (Ed.), *Nonverbal behavior: Perspectives, applications, and cultural insights* (pp. 431–460). Lewiston, NY: Hogrefe.

Price, C., & Cuellar, I. (1981). Effects of language and related variables on the expression of psychopathology in Mexican psychiatric patients. *Hispanic Journal of Behavioral Science, 32*(2), 145–160.

Query, J. M. (1985). Comparative admission and follow-up study of American Indians and Whites in a youth chemical dependency unit on the North Central Plains. *The International Journal of the Addictions, 20,* 489–502.

Ragin, C., & Hein, J. (1993). The comparative study of ethnicity: Methodological and conceptual issues. In J. Stanfield II & R. Dennis (Eds.), *Race and ethnicity in research methods* (pp. 254–272). Newbury Park, CA: Sage.

Ramsey. S. (1979). Nonverbal behavior: An intercultural perspective. In M. Asante, E. Newmark, & C. Balke (Eds.), *The handbook of intercultural communication* (pp. 105–143). Beverly Hills, CA: Sage.

Ramsey, S. (1984). Double vision: Nonverbal behavior: Perspectives, applications, cultural insights. In A. Wolfgang (Ed.), *Nonverbal behavior* (pp. 139–167). Lewiston, NY: Hogrefe.

Redfern, S., Dancey, C., & Dryden, W. (1993). Empathy: Its effects on how counsellors are perceived. *British*

Journal of Guidance and Counselling, 21(3), 300–309.

Rey-Perez, A. (1996). *A culturally sensitive psychotherapeutic program for Hispanic children and families*. Unpublished doctoral directed study project, Nova Southeastern University, Fort Lauderdale, FL.

Ridley, C. (1984). Clinical treatment of the nondisclosing Black client: A therapeutic paradox. *American Psychologist, 39,* 1234–1244.

Ridley, C., Mendoza, D., Kanitz, B., Angermeier, L., & Zenk, R. (1994). Cultural sensitivity in multicultural counseling: A perceptual schema model. *Journal of Counseling Psychology, 41*(2), 125–136.

Rivas-Vazquez, R. (1989). *Assessing differential levels of affecting expression in bilingual subjects using the Thematic Apperception Test*. Unpublished doctoral dissertation, Nova Southeastern University, Fort Lauderdale, FL.

Rogler, L. (1993). Culturally sensitizing psychiatric diagnosis: A framework for research. *The Journal of Nervous and Mental Disease, 181*(7), 401–408.

Romeo, D. (1985). Cross-cultural counseling: Brief reactions for the practitioner. *The Counseling Psychologist, 13*(4), 665–671.

Root, M. (1992). Within, between, and beyond race. In M. Root (Ed.), *Racially mixed people in America* (pp. 3–11). Newbury Park, CA: Sage.

Rosenthal, R., Hall, J., Archer, D., Dimatteo, R., & Rogers, P. (1979). Measuring sensitivity to nonverbal communication: The PONS Test. In A. Wolfgang (Ed.), *Nonverbal behavior: Applications and cultural implications* (pp. 222–241). New York: Academic Press.

Rotter, J. (1966). Generalised expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80,* 1–28.

Ruiz, R., & Casas, M. (1981). Culturally relevant and behavioristic counseling for Chicano college students. In P. Pedersen, J. Draguns, W. Lonner, & J. Trimble (Eds.), *Counseling across cultures: Revised and expanded edition* (pp. 203–226). Honolulu, HI: University of Hawaii Press.

Russell, D. (1988). Language in psychotherapy: The influence of nonstandard English in clinical practice. In L. Comas-Diaz & E. Griffith (Eds.), *Clinical guidelines in cross-cultural mental health* (pp. 33–68). New York: Wiley.

Salzma, M. (1995). Attributional discrepancies and bias in cross-cultural interactions. *Journal of Multicultural Counseling and Development, 23*(3), 181–193.

Samovar, L., & Porter, R. (Eds.) (1991). *Intercultural communication: A reader* (6th ed.). Belmont, CA: Wadsworth.

Samuda, R. J. (1975). *Psychological testing of American minorities: Issues and consequences*. New York: Harper & Row.

Samuda, R. J., Kong, S., Cummins, J., Pascual-Leone, J., & Lewis, J. E. (1991). Assessment and placement of minority students. Kingston, ON: Hogrefe/ISSP.

Samuda, R. J., & Lewis, J. E. (1992). Evaluation practices for the multicultural classroom. In C. Diaz (Ed.), *Multicultural education in the 21st century* (pp. 97–111). Washington, DC: National Education Association.

Sanchez, A., & Atkinson, D. (1983). Mexican American cultural commitment preference for counselor ethnicity and willingness to use counseling. *Journal of Counseling Psychology, 30*(2), 215–220.

Sandhu, D. (1984). *The effects of mirroring versus non-mirroring of clients' nonverbal behaviors on empathy, trustworthiness, and positive interaction in cross-cultural counseling dyads (NLP, neurolinguistic programming)*. Unpublished doctoral dissertation, Mississippi State University, MS.

Sandhu, D., Reeves, T., & Portes, P. (1993). Cross-cultural counseling and neurolinguistic mirroring with Native American adolescents. *Journal of Multicultural Counseling and Development, 21*(2), 106–118.

Santiago-Rivera, A. L. (1995). Developing a culturally sensitive treatment modality for bilingual Spanish-speaking clients: Incorporating language and culture in counseling. *Journal of counseling and development*, 74, 12–17.

Sattler, J. (1977). The effects of therapist–client racial similarity. In A. Gurman & A. Razin (Eds.), *Effective psychotherapy: A handbook of research* (pp. 252–290). New York: Pergamon.

Scheflen, A. (1964). The significance of posture in communication systems. *Psychiatry, 27,* 316–331.

Schinke, S. P., Orlandi, M. A., Botvin, G. J., Gilchrist, L. D., Trimble, J. E., & Locklear, V. B. (1988). Preventing substance abuse among American-Indian adolescents: A bicultural competence skills approach. *Journal of Counseling Psychology, 35,* 87–90.

Schwartz, S. (1992). The universal content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna (Ed.), *Advanced experimental social psychology* (pp. 1–65). New York: Academic Press.

Schwartz, S. (1994). Beyond individualism-collectivism: New cultural dimensions of values. In U. Kim, H. Triandis, C. Kagitcibasi, S. Choi, & G. Yoon (Eds.), *Individualism and collectivism: Theory, method, and applications* (pp. 85–119). Newbury Park, CA: Sage.

Sechrest, L., McKnight, P., & McKnight, K. (1996). Calibration of measures for psychotherapy outcome studies. *American Psychologist, 51*(10), 1065–1071.

Seligman, M. (1996). Science as an ally of practice. *American Psychologist, 51*(10), 1072–1079.

Smith, A. W. (1993). Survey research on African Americans: Methodological innovations. In J. Stanfield II & R. Dennis (Eds.), *Race and ethnicity in research methods* (pp. 217–232). Newbury Park, CA: Sage.

Smith, H. (1984). State of the art of nonverbal behavior in teaching. In A. Wolfgang (Ed.), *Nonverbal behavior: Perspectives, applications, and cultural insights* (pp. 171–202). Lewiston, NY: Hogrefe.

Smith, P., Peterson, M., Akande, D., Callan, V., & Cho, N. (1994). Organizational event management in 14 countries: A comparison with Hofstede's dimensions. In A. Bouvy, F. van de Vijver, P. Boski, & P. Schmitz (Eds.), *Journeys into cross-cultural psychology* (pp. 212–229). Amsterdam: Swets & Zeitlinger.

Snodgrass, L. L., Yamamoto, J., Frederick, C., Ton-That, N., Foy, D., Chan, L., Wu, J., Hahn, P. H., Shinh, D., Nguyen, N. H., de Jonge, J., & Fairbanks, L. (1993). Vietnamese refugees with PTSD symptomatology: Intervention via a coping skill model. *Journal of Traumatic Stress, 6*(4), 569–575.

Sodowski, G., & Impara, J. (Eds.) (1996). *Multicultural assessment in counseling and clinical psychology*. Lincoln, NE: Buros Institute of Mental Measurements.

Sodowski, G., Taffe, R., Gutkin, T., & Wise, S. (1994). Development of Multicultural Counseling Inventory: Self-report measure of multicultural competencies. *Journal of Counseling Psychology, 41,* 137–148.

Sommer, R. (1959). Studies in personal space. *Sociometry, 22,* 247–260.

Song, M., & Parker, D. (1995). Commonality, difference and the dynamics of disclosure in in-depth interviewing. *Journal of Sociology, 29*(2), 241–256.

Speigal, J., & Papajohn, J. (1986). Training program in ethnicity and mental health. In H. Lefley & P. Pedersen (Eds.), *Cross-cultural counseling for mental health professionals* (pp. 49–71). Springfield, IL: Thomas.

Speight, S. L., Myers, L. J., Cox, C., & Highlen, P. S. (1991). A redefinition of multicultural counseling. *Journal of Counseling and Development, 70*(1), 29–36.

Spradley, J. P. (1979). *The ethnographic interview*. New York: Holt, Rinehart, & Winston.

Stanfield, J. II (1993). Methodological reflections: An introduction. In J. Stanfield II & R. Dennis (Eds.), *Race and ethnicity in research methods* (pp. 3–15). Newbury Park, CA: Sage.

Stanges, B., & Riccio, A. (1970). A counselee preference for counselors: Some implications for counselor education. *Counselor Education and Supervision, 10,* 39–46.

Steward, R., Gimenez, M., & Jackson, C. (1995). A study of personal preferences of successful university students as related to race/ethnicity and sex: Implications and recommendations for training, practice, and future research. *Journal of College Student Development, 36*(2), 123–131.

Strupp, H. (1996). The tripartite model and the Consumer Reports study. *American Psychologist, 51*(10), 1017–1024.

Sue, D., & Sundberg, N. D. (1996). Research and research hypotheses about effectiveness in intercultural counseling. In P. B. Pedersen, J. G. Draguns, W. J. Lonner, & J. E. Trimble (Eds.), *Counseling across cultures* (4th ed., pp. 323–352). Thousand Oaks, CA: Sage.

Sue, D. W. (1977). Barriers to effective cross-cultural counseling. *Journal of Counseling Psychology, 24,* 420–429.

Sue, D. W. (Ed.) (1981). *Counseling the culturally different.* New York: Wiley.

Sue, D. W. (1990). Culture specific techniques in counseling: A conceptual framework. *Professional Psychology, 21*(6), 424–433.

Sue, D. W. (1991). A model for cultural diversity training. *Journal of Counseling and Development, 70,* 99–105.

Sue, D. W. (1993). Confronting ourselves: The White and racial/ethnic-minority researcher. *The Counseling Psychologist, 21*(2), 244–249.

Sue, D. W. (1995). Toward a theory of multicultural counseling and therapy. In J. A. Banks & C. A. McGee Banks (Eds.), *Handbook of research on multicultural education* (pp. 647–659). New York: Macmillan.

Sue, D. W., Arredondo, P., & McDavis, R. J. (1992). Multicultural competencies and standards: A call to the profession. *Journal of Counseling and Development, 70,* 477–486.

Sue, D. W., Bernier, J. E., Durran, A., Feinberg, L., Pedersen, P., Smith, E. J., & Vasquez-Nuttal, E. (1982). Position paper: Cross-cultural competencies. *The Counseling Psychologist, 1,* 45–52.

Sue, D. W., Ivey, A., & Pedersen, P. (Eds.) (1996). *A theory of multicultural counseling and therapy.* Pacific Grove, CA: Brooks/Cole.

Sue, D. W., & Sue, D. (1990). *Counseling the culturally different: Theory and practice* (2nd ed.). New York: Wiley.

Sue, S., Fujino, D. C., Hu, L. T., Takeuchi, D. T., & Zane, N. W. (1991). Community mental health services for ethnic minority groups: A test of the cultural responsiveness hypothesis. *Journal of Counseling Psychology, 59,* 533–540.

Sue, S., & Zane, N., (1987). The role of culture and cultural techniques in psychotherapy: A critique and reformulation. *American Psychologist, 42*(1), 37–45.

Sue, S., Zane, N., & Young, K. (1994). Research on psychotherapy with culturally diverse populations. In A. Bergin & S. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 783–820). New York: Wiley.

Szapocznik, J., Rio, A., Perez-Vidal, A., Kurtines, W., & Sanisteban, D. (1986). Family effectiveness training for Hispanic families. In H. Lefley & P. Pedersen (Eds.), *Cross-cultural training for mental health professionals* (pp. 245–261). Springfield, IL: Thomas.

Takeuchi, D., Mokuau, N., & Chun, C. (1992). Mental health services for Asian Americans and Pacific Islanders. *The Journal of Mental Health Administration, 19*(3), 237–244.

Tapp, J. L., Kelman, H., Triandis, H., Wrightsman, L., & Coelho, J. (1974). Advisory principles for ethical considerations in the conduct of cross-cultural research: Fall 1973 revision. *International Journal of Psychology, 9*(3), 240–249.

Tedeschi, G., & Willis, F. (1993). Attitudes toward counseling among Asian international and native caucasian students. *Journal of College Student Psychotherapy, 7*(4), 43–54.

Thomas, L. (1995). Psychotherapy in the context of race and culture: An inter-cultural therapeutic approach. In S. Fernando (Ed.), *Mental health in a multi-ethnic society* (pp. 172–190). London: Routledge.

Thomason, T. (1991). Counseling Native Americans: An introduction for non-Native American counselors. *Journal of Counseling and Development, 69*(4), 321–327.

Thompson, E., Neighbors, H., Munday, C., & Jackson, J. (1996). Recruitment and retention of African American patients for clinical research: An exploration of response rates in an urban psychiatric hospital. *Journal of Consulting Psychology, 64*(5), 861–867.

Thompson, L. (1989). *Romans and Blacks.* Norman, OK: University of Oklahoma Press.

Tomlinson, S., & Cheatham, H. (1989). Effects of counselor intake judgements on service to Black students using a university counseling center. *Counselling Psychology Quarterly, 2*(2), 105–111.

Tomlinson-Clarke, S., & Cheatham, H. (1993). Counselor and client ethnicity and counselor intake judgements. *Journal of Counseling Psychology, 40*(3), 267–270.

Triandis, H. (1978). Some universals of social behavior. *Personality in Social Psychology Bulletin, 4,* 1–16.

Triandis, H. (1994). *Culture and social behavior.* New York: McGraw-Hill.

University of Wisconsin (1991). *Cross-cultural interviewing in social work practice: The case of the Hmong* [videocassette]. (Available from the Social Work Department of the University of Wisconsin, Eau-Claire, WI)

VandenBos, G. (1996). Outcome assessment of psychotherapy. *American Psychologist, 51*(10), 1005–1006.

Van Dijk, T. (1993). Analyzing racism through discourse analysis: Some methodological reflections. In J. Stanfield II & R. Dennis (Eds.), *Race and ethnicity in research methods* (pp. 92–134). Newbury Park, CA: Sage.

Vazquez-Nuttall, E., Webber, J., & Sanchez, W. (1996). MCT theory and implications for training. In D. W. Sue, A. Ivey, & P. Pedersen (Eds.), *A theory of multicultural counseling and therapy* (pp. 123–138). Pacific Grove, CA: Brooks/Cole.

Villegas-Reimers, E. (1996). Self-development of Venezuelan adolescents. *Journal of Cross-Cultural Psychology, 27*(1), 25–36.

Vontress, C. (1981). Racial and ethnic barriers in counseling. In P. Pedersen, J. Draguns, W. Lonner, & J. Trimble (Eds.), *Counseling across cultures: Revised and expanded edition* (pp. 87–107). Honolulu, HI: University of Hawaii Press.

Weitz, S. (1972). Attitude, voice, and behavior: A repressed affect model of interracial interaction. *Journal of Personality and Social Psychology, 24,* 14–21.

Westbrook, F. D., & Sedlacek, W. E. (1991). Forty years of using labels to communicate about nontraditional students: Does it help or hurt? *Journal of Counseling and Development, 70*(1), 20–28.

Westermeyer, J. J. (1993). Cross-cultural psychiatric assessment. In A. C. Gaw (Ed.), *Culture, ethnicity and mental illness* (pp. 125–144). Washington, DC: American Psychiatric Press.

White, J. (1994). The impact of race and ethnicity on transference and countertransference in combined individual–group therapy. *Group, 18*(2), 89–99.

Williams, M. (1993). Urban ethnography: Anther look. In

J. Stanfield II & R. Dennis (Eds.), *Race and ethnicity in research methods* (pp. 135–158). Newbury Park, CA: Sage.

Wolfgang, A. (1985). Intercultural counselling and nonverbal behavior: An overview. In R. Samuda & A. Wolfgang (Eds.), *Intercultural counselling and assessment* (pp. 33–48). Lewiston, NY: Hogrefe.

Wolfgang, A., & Bhardwaj, A. (1984). Bibliography of 100 years of nonverbal study. In A. Wolfgang (Ed.), *Nonverbal behavior* (pp. 461–469). Lewiston, NY: Hogrefe.

Worchel, S., & Teddlie, C. (1976). The experience of crowding: A two factor theory. *Journal of Personality and Social Psychology, 34*, 30–40.

World Health Organization (1983). *Depressive disorders in different cultures*. Geneva, Switzerland: Author.

World Health Organization (1992). *Refugee mental health: Draft manual for field testing*. Geneva: Author.

Wrenn, C. (1962). The culturally encapsulated counselor. *Harvard Educational Review, 32*, 444–449.

Wrenn, C. (1985). Afterward: The culturally encapsulated counselor revisited. In P. Pedersen (Ed.), *Handbook of cross-cultural counseling and therapy* (pp. 323–329). Westport, CT: Greenwood.

Yamamoto, J., & Kubota, M. (1983). The Japanese American family. In J. Yamamoto, A. Romero, & A. Morales (Eds.), *The psychosocial development of minority group children* (pp. 65–84). New York: Brunner/Mazel.

Yamamoto, J., Silva, J., Justice, L., Chang, C., & Leong, G. (1993). Cross-cultural psychotherapy. In A. Gaw (Ed.), *Culture, ethnicity, and mental illness* (pp. 101–124). Washington, DC: American Psychiatric Press.

Yee, A. H., Fairchild, H. H., Weizmann, F., & Wyatt, G. (1993). Addressing psychology's problems with race. *American Psychologist, 48*, 1132–1140.

Yeh, M., Eastman, K., & Cheung, M. (1994). Children and adolescents in community mental health centers: Does the ethnicity or the language of the therapist matter? *Journal of Community Psychology, 22*, 153–163.

Yeh, M., Takeuchi, D., & Sue, S. (1994). Asian-American children treated in the mental health system: A comparison of parallel and mainstream outpatient service centers. *Journal of Clinical Child Psychology, 23*(1), 5–12.

Yum, O. (1991). The impact of Confucianism on interpersonal relationships and communication patterns in East Asia. In L. A. Samovar & R. E. Porter (Eds.), *Intercultural communication: A reader* (6th ed., pp. 71–78). Belmont, CA: Wadsworth.

Zane N., & Sue, S. (1991). Culturally-respective mental health services for Asian Americans: Treatment and training issues. In H. Myers, P. Wohlford, P. Guzman., & R. Echemendia (Eds.), *Ethnic minority perspectives on clinical training and services in psychology* (pp. 49–58). Washington, DC: American Psychological Association.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.05

Cultural Competence Training in Clinical Psychology: Assessment, Clinical Intervention, and Research

FELIPE G. CASTRO

Arizona State University, Tempe, AZ, USA

10.05.1 INTRODUCTION 127
10.05.2 DEFINITIONS AND KEY CONCEPTS 128
  10.05.2.1 Why the Need for Cultural Competence Training? 128
  10.05.2.2 Cultural Competence as Good Science and Practice 128
  10.05.2.3 Key Aspects of Cultural Competence 129
  10.05.2.4 The Concept of Culture 130
  10.05.2.5 Levels of Cultural Capacity 131
10.05.3 EVALUATING LEVELS OF CULTURAL CAPACITY 133
  10.05.3.1 Evaluation Relevant to a Reference Group 133
  10.05.3.2 Cultural Capacity in Assessment 135
  10.05.3.3 Cultural Capacity in Clinical Intervention 136
  10.05.3.4 Cultural Capacity in Research 137
  10.05.3.5 A Life-long Professional Odyssey 138
10.05.4 FUTURE DIRECTIONS 139
10.05.5 SUMMARY 139
10.05.6 REFERENCES 139

10.05.1 INTRODUCTION

This chapter presents a three-factor model for describing and rating the capacity of a clinical psychologist or other health professional to conduct culturally effective assessments, clinical interventions, and research with members of ethnic minority populations. The overall aim of this model is to provide a specific and clinically useful system that guides skills training towards the development of cultural competence. Developing cultural competence in a new cohort of health professionals aims to enhance quality in the health services and research that is delivered to members of various ethnic minority groups, including African-Americans, Latinos/Hispanics, Asian-Americans and Pacific Islanders, and Native Americans.

This training model introduces a distinct minority perspective based on at least five assumptions. First, the model assumes that "culture" cannot be ignored, and that it must be well understood in order for a clinical psychologist or for a health program to work effectively with members of ethnic/racial populations. Second, this model assumes that the clinical psychologist must examine sociocultural factors in order to conduct effective assessments, treatments, and research with ethnic minority clients and communities. Third, it assumes that attaining the highest level of cultural capacity, which is cultural proficiency, requires the clinical psychologist to decenter himself or herself culturally, meaning that he or she must learn empathically to adopt the client's *emic* (cultural insider's) world view. A fourth assumption is that cultural proficiency in research and in community-based health service programs requires a social action research approach, an approach which invites the participation of ethnic minority community folks in the development and implementation of any intervention that is conducted within that ethnic minority community (Flores, Castro, & Fernandez-Esquer, 1995). Finally, within this perspective, a fifth assumption is that the culturally proficient clinical psychologist should exercise leadership by way of mentorship and an active involvement in advancing minority issues.

10.05.2 DEFINITIONS AND KEY CONCEPTS

10.05.2.1 Why the Need for Cultural Competence Training?

Today, the concept of cultural competence has emerged as the result of several converging social trends. First, the demographic face of the USA has changed since the early 1970s as several ethnic/racial populations have grown in size and significance. Several of the major ethnic/racial populations in the USA have shown growth rates that surpass the growth rate of the mainstream USA population. As these trends continue, for the year 2000, the projected sizes of the major ethnic racial populations of the USA in terms of size and percentage are: African-Americans, about 33.7 million (about 12.2%); Latinos/Hispanics, about 31.1 million (about 11.3%); Asian-Americans and Pacific Islanders, about 11.4 million (about 4.1%); and Native Americans/American-Indians, about 2.2 million (about 0.7%) (Department of Commerce, 1994).

As a second factor, the 1985 *Secretary's Task Force Report on Black and Minority Health* highlighted in greater detail what was known previously, which is that within the USA there exists a major gap in health status between the white mainstream (primarily middle class) population and the ethnic/racial minority and mostly lower class populations (Department of Health and Human Services, 1985). Moreover, a social process involving several interrelated behavioral, sociocultural, environmental, and economic factors has contributed to this disparity (Anderson, 1995). Thus, clinical psychologists and other providers of health services must comprehend the manner in which these various sociocultural factors have operated to influence the health status of these minority populations. Clearly then, clinical psychologists must acknowledge and understand the detrimental effects of certain sociocultural factors when conducting assessments, when developing clinical interventions, and when designing and conducting research that involves ethnic minority individuals, families, or communities. Within this context and at the level of the individual client, it is clear that ethnic minority clients have complex, and at times unique, needs that must be well understood within a cultural context in order to engage them successfully in treatment and/or research. This emerging recognition of the effects of sociocultural factors on physical and mental health underscores the need for cultural competence training, in order to improve the clinical psychologist's capacity to work effectively with members of these minority populations.

Third, the current need to develop cultural competence in clinical psychologists has also emerged as a result of the "pipeline issue." Trends regarding the educational progress of students of ethnic/racial backgrounds reveal low numbers of ethnic minority graduates who will be able to provide culturally relevant health services to members of their ethnic/racial group. Only a limited number of these minority graduates will enter the professional arena in order to address the large and growing need for services offered by minority professionals who understand the needs of ethnic/racial minority clients. Given this trend, and the growing need for culturally relevant health services, a more viable solution is to train and educate mainstream service providers, administrators, and researchers on essential aspects of the culture and health needs of clients from various ethnic/racial populations.

10.05.2.2 Cultural Competence as Good Science and Practice

Developing cultural competence in clinical psychologists of all backgrounds should not be seen as a "politically correct" gesture to demonstrate token responsiveness to a contemporary social concern. On the contrary, developing the cultural competence of clinical psychologists should be seen as an important training goal that offers them a broader range of capabilities for working more effectively clinically and in research with a broader and more diverse group of clients and research participants. Moreover, the American Psychological Association's Ethical Principles of Psychologists and Code of Conduct indicates that psychologists should "provide services, teach, and conduct research only within the boundaries of their competence, based on their education, training, supervised experience, or appropriate professional experience" (American Psychological Association, 1992, p. 1600; Standard 1.04—Boundaries of Competence).

In addition, these guidelines indicate that,

> in those emerging areas in which generally recognized standards for preparatory training do not yet exist, psychologists [should] nevertheless take reasonable steps to ensure the competence of their work and to protect patients/clients, students, research participants, and others from harm.

Clearly, in the late 1990s and in the early years of the twenty-first century, the guidelines that define cultural competence in clinical interventions with ethnic/racial populations are still evolving. Thus, in response to the American Psychological Association's Ethical Principles, all psychologists who work with clients from ethnic/racial backgrounds or who anticipate working with these clients should continually seek ways to enhance their cultural competence in order to conduct effective assessments, interventions, and/or research with these minority clients.

Moreover, from an organizational perspective, it can be noted that practically all organizations have a "culture," which consists of organizational operating characteristics. Based on these operating characteristics, some organizations can be described as having a healthy, accepting, collegial, supportive, comfortable, consumer-friendly, and culturally competent culture. By contrast, other organizations can be described as having a paranoid, judgmental, patriarchal, controlling, stressful, consumer-unfriendly, and culturally incompetent culture. The operating characteristics that govern a particular organization's culture affect the way that consumers will react to the services that the organization delivers, and this is especially true for the ethnic minority consumers who seek services within that organization. Human service organizations that have institutional policies, procedures, and structures that discount or devalue the importance of cultural competence are not likely to deliver culturally competent services to their minority clients, and they are likely to induce tension and demoralization among members of their staff who see the need for cultural competence in services that would be delivered to minority clients.

Finally, cultural competence training is also relevant to the conduct of good research and good science. In March 1994, the National Institutes of Health (NIH) published new guidelines for the inclusion of women and minorities in clinical research, guidelines which reflected the requirements of the NIH Revitalization Act 1993 (Public Law 103-43) (National Institutes of Health, 1994). The release of these guidelines gave formal recognition to the importance of gender and ethnic cultural factors in scientific clinical research (Hohmann & Parron, 1996). These guidelines indicate that studies of outstanding scientific merit cannot be conducted if such studies systematically or arbitrarily exclude a large and diverse sector of the population from their operating plans. Such studies are flawed if, without a sound rationale, they exclude women and/or minorities from their theoretical framework, their sampling design, their methodology, and thus from their findings and their conclusions. Here, cultural competence in research is demonstrated by an investigator who understands and can conceptualize the hypothesized effects of specific sociocultural factors on health outcomes, and who can design and implement studies that effectively recruit and retain women and minority participants (Areán & Gallagher-Thompson, 1996; Miranda, Azocar, Organista, Muñoz, & Lieberman, 1996; Thompson, Neighbors, Munday, & Jackson, 1996).

10.05.2.3 Key Aspects of Cultural Competence

Generally, cultural competence may be conceptualized as a capacity to work effectively with individuals from special populations, including clients from ethnic/racial communities. Cultural competence is expressed by communicating acceptance, deep understanding, and responsiveness to the needs and concerns of members of special populations. Cultural competence may also be conceptualized as a capacity level that supersedes cultural sensitivity, given that cultural competence involves a level of understanding and skills that moves beyond a basic awareness of the needs and concerns of members from a special population.

Cultural competence has been defined in various ways, although several definitions reveal certain common themes. For example, the Arizona Department of Health Services has defined cultural competence at the institutional level as

> knowledge, attitudes, and policies within an agency which allows individuals to work effectively in cross cultural situations. This requires the willingness and ability to utilize community-based values, traditions, and practices in developing and evaluating interventions, communication, and other activities. (Arizona Department of Health Services, 1995)

A definition that applies at the individual and institutional levels was developed by a behavioral health workgroup. This definition indicates that cultural competence

> includes the attainment of knowledge, skills, and attitudes to enable practitioners and systems of care to provide care for diverse populations, i.e., to work within the person's reality conditions, in particular minority populations, acknowledges and incorporates variance in normative acceptable behaviors, beliefs, and values in determining an individual's mental wellness/illness and incorporates those variables into assessment and treatment. (The National Latino Behavioral Health Workshop, 1996)

A few common themes have been stated or implied in these and other definitions of cultural competence as defined at both the macro institutional level, and the micro individual levels. These themes include the notions that cultural competence:

(i) consists of a complex set of knowledge, attitudes, and skills;

(ii) is characterized by an empathic and culturally relevant understanding of the client, using an insider's view of the client's presenting problem;

(iii) requires that the clinical psychologist or program exhibit cultural responsiveness, that is, a commitment towards understanding the client within the context of the client's own culture;

(iv) requires the use of culture-based insights in planning an intervention (individual/family therapy, or a program) that is culturally relevant and effective in alleviating the client's problem; and

(v) requires careful attention to the process of intervention, which is as important as the outcome, because in minority cultures expressions of respect and acceptance are indispensable aspects of the client–therapist relationship and will greatly influence client participation and treatment outcomes.

In addition, there are other aspects of cultural competence, primarily at the micro clinical level, that are less salient in the literature, but that are important in cultural competence training for clinical psychologists. These are:

(i) that cultural competence is a value-added capability where, under the current managed care environment, the culturally competent clinical psychologist is capable of delivering culturally responsive interventions that surpass the effectiveness of conventional interventions (The National Latino Behavioral Health Workshop, 1996);

(ii) that level of cultural competence is specific to a given cultural reference group, where, based on training and experience, a given clinical psychologist may exhibit cultural competence in working with one cultural group, for example with inner city African-American adolescents, but may only exhibit cultural sensitivity in addressing the needs of a different cultural group, such as elderly rural Mexican Americans;

(iii) that the concepts of culture and subculture are not synonymous with demographic ethnic labels (e.g., Hispanic), where cultural competence in program planning requires a segmentation of the targeted ethnic population into subcultural groups (Balcazar, Castro, & Krull, 1995), the members of which have distinct lifestyles, needs, and problems (e.g., Chicano gang adolescents from East Los Angeles vs. Navajo elderly who live on the reservation);

(iv) that culturally responsive interventions should be guided by culturally relevant intervention models and defining criteria, and by a management information system that gathers local data and that uses culturally relevant variables; and

(v) that truly effective training to develop cultural competence requires continuing education that systematically upgrades skills across levels of the cultural capacity continuum towards the ultimate goal of cultural proficiency (Bernal & Castro, 1994).

10.05.2.4 The Concept of Culture

The concept of "culture" is often difficult to grasp operationally, because, as a concept, culture is a rich, complex, and multidimensional concept. Indeed, one metaphor of culture as it relates to the heart and soul of ethnic minority peoples is the concept of "humidity" as it relates to the weather. You cannot see humidity, but you can definitely feel it when it is there. The same is true for culture. Given its rich complexity, well over 100 definitions of "culture" have been proposed (Murphy, 1986, cited in Baldwin & Lindsley, 1994). Indeed, culture has been defined broadly as

> the total body of tradition borne by a society and transmitted from generation to generation. It thus refers to norms, values, and standards by which people act, and it includes the ways distinctive in

> each society of ordering the world and rendering it intelligible. (Baldwin & Lindsley, 1994; Murphy, 1986).

Across various definitions of culture, certain common themes have emerged. These are:

(i) that culture consists of the totality of learned behaviors of a people;

(ii) that it is transmitted from generation to generation, that is, from elders to children;

(iii) that people construct their own culture in order to give meaning to life by creating a world view that helps to explain their own reality;

(iv) that it consists of a people's shared beliefs, values, ways of making things, customs, behaviors, traditions, and lifestyles;

(v) that it offers a community of people with a set of social norms and moral values on how life should be lived;

(vi) that, via a culture's art, music, folklore, and other forms of creative expression, culture captures the soul, character, and essence of a people; and

(vii) that culture is important to a group of people because it promotes a sense of kinship, belonging, and group identity.

Given the breadth and richness of the concept of culture, for cultural competence training a more narrow concept and definition is needed as a reference point for the operationalization of levels of cultural capacity (cultural sensitivity, cultural competence, and cultural proficiency). The approach that we will use is to begin by identifying a specific targeted reference group, which is the group of clients that is targeted for a given intervention or program. This targeted group may be defined either broadly (e.g., African-Americans, Asian-Americans) or more narrowly (e.g., African-American elderly aged 65 years and older, low-acculturated Chinese adolescents). The more broadly defined groups are typically populations that share a common culture. By contrast, the more narrowly defined groups are typically subpopulations and subcultures that share a culture with the parent group, but that are also governed by additional and perhaps unique norms, values, and other local sociocultural conditions. Thus, a clinical psychologist's cultural capacity level can be defined more accurately, by describing specific capacity levels in the areas of assessment, clinical intervention, or research that are defined in relation to the culture of a specific targeted reference group.

10.05.2.5 Levels of Cultural Capacity

Figure 1 shows that the capacity for cultural competence varies along a graded continuum. This concept of a cultural capacity continuum has been proposed previously by various scholars (Cross, Bazron, Dennis, & Isaacs, 1989; Kim, McLeod, & Shantis, 1992; Orlandi, 1992). In our current model, we are taking this concept of a cultural capacity continuum and are modifying and expanding it.

On this cultural capacity continuum, the lowest level is "cultural destructiveness" (− 3), which consists of an ethnocentric orientation that harbors openly negative destructive attitudes that emphasize the superiority of one's own culture and the inferiority of other cultures. Attitude has been previously defined as a "readiness to respond," a learned "predisposition to respond in a particular way to a specific attitude object" (Oskamp, 1991). Within a health service organization, a given staff member's condescending attitude can serve as an expression of cultural destructiveness, and this attitude will destroy the client–provider relationship. Similarly, at the institutional level, an agency's policies or practices that promote stereotypes and a discrimination of lower class clients is an expression of institutional cultural destructiveness. Here a stereotype refers to a mental image or a set of beliefs about a group of people, beliefs that are highly simplistic (and often inaccurate), highly evaluative (often derogatory), and rigidly resistant to change (and thus not modified by corrective information) (Oskamp, 1991). Ironically, in contemporary American society, culturally destructive attitudes in individuals and culturally destructive policies within institutions still exist.

"Cultural incapacity"(−2) refers to a professional's or an organization's orientation that emphasizes separate but equal treatment. While this may seem to be a viable and fair approach, in reality it often results in an institutional incapacity to provide equal and effective treatment of those clients who are culturally different. Similarly, at the next level, an orientation describable as "cultural blindness" (− 1) emphasizes that all cultures and people are alike and equal. While this orientation appears unbiased and accepting, its subtle yet injurious effect on members of special populations is that this approach discounts the importance of culture and the need to consider multicultural perspectives. Indeed, in a well-meaning but culturally insensitive effort to be fair and to treat everyone the same, an administrator may use this orientation to discount or discredit programmatic efforts that aim to respond affirmatively and in a culturally sensitive manner to the unique needs of clients from special populations.

The first level of a positive cultural capacity is "cultural sensitivity" (+ 1). Cultural sensitivity is characterized by the presence of a basic

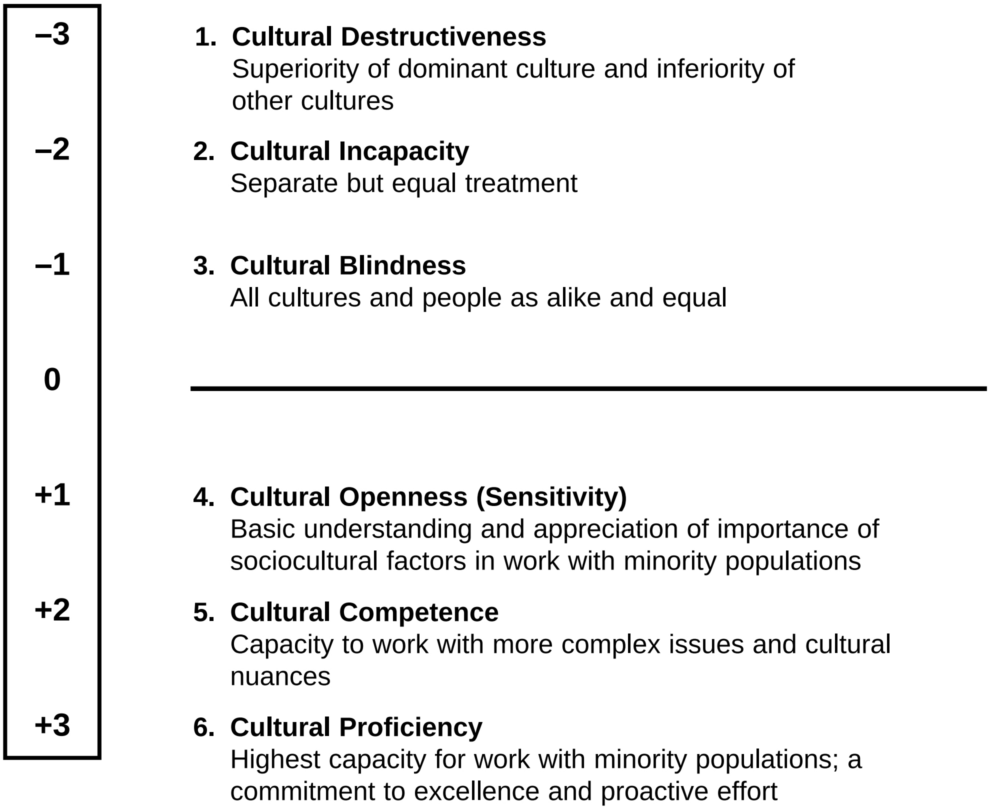
–3
1. Cultural Destructiveness
Superiority of dominant culture and inferiority of other cultures
–2
2. Cultural Incapacity
Separate but equal treatment
–1
3. Cultural Blindness
All cultures and people as alike and equal
0
+1
4. Cultural Openness (Sensitivity)
Basic understanding and appreciation of importance of sociocultural factors in work with minority populations
+2
5. Cultural Competence
Capacity to work with more complex issues and cultural nuances
+3
6. Cultural Proficiency
Highest capacity for work with minority populations; a commitment to excellence and proactive effort

Figure 1 Levels of cultural capacity. Adapted from Kim et al. (1992).

understanding and appreciation of the importance of sociocultural factors as these affect the client and the most viable treatment that may be offered. For example, in a health promotion program for Latinos, cultural sensitivity is reflected by the program's attention to the role of linguistic factors, interpersonal factors (*personalismo*, *respeto*, *confianza*), and of familal factors (traditionalism, familism), as these affect the health-related behaviors of various Latinos.

Cultural sensitivity is also characterized by the presence of an elementary, but nonetheless critical, appreciation of the within-group variability, that is, of the diversity that exists within a given ethnic minority population. For example, among Latinos, scholars have regularly asserted that Hispanics/Latinos are a heterogeneous population characterized by large variability among its members in terms of racial background, educational attainment, household income, family size and composition, and even in the level of ethnic pride evident among its members (Montgomery, 1994).

"Cultural competence" (+2) represents a higher level of capacity to work with members of a special population. Relative to cultural sensitivity, cultural competence is characterized by a greater depth of skills and experience that allows the clinical psychologist to understand and work effectively with cultural nuances. By understanding cultural nuances, the clinical psychologist is able to detect the ethnic client's subtle communications, to interpret their meanings, and to do so within the client's specific cultural context. Based on this deeper level of understanding, the clinical psychologist is thus more capable of planning culturally potent interventions. Such interventions would appeal to the ethnic client, elicit the client's participation, and motivate sustained participation.

Similarly, in a community-based health promotion program, the cultural competence of the program's staff is characterized by their capacity to work actively with members of a community. This capacity includes an understanding of the cultural and political nuances that helps these staff members to develop trust and credibility with members of the local community, and to elicit their support and participation in the design and implementation of a health promotion program. Here also, cultural competence is characterized by the clinical psychologist's knowledge of the motivational dynamics that prevail within subgroups of a given population, as these motivational dynamics may be mediated by age, gender, cultural orientation, and other aspects of the group's subculture.

"Cultural proficiency" (+3) is the highest expression of cultural capacity. It serves as an ideal, rather than necessarily as a state that is attainable by all clinical psychologists. Cultural proficiency consists of a state of high mastery, a commitment to excellence, and a proactive attitude that facilitates the design and delivery of therapeutic services for members of a specific population. Here it is recognized that a given clinical psychologist or health program may exhibit cultural proficiency in service delivery to members of one targeted population (e.g., Latinos), but may only exhibit cultural sensitivity in service delivery to members of a different population. Thus, complete cultural proficiency with all special populations would be a rare accomplishment for many clinical psychologists, health programs, or human service agencies. Nonetheless, constantly striving to develop cultural proficiency in work with one or more special population is a laudable goal for any clinical psychologist, health program, or agency, although this calls for a commitment to life-long learning with the ultimate aim of developing depth and breadth in skills for working with members of one or more special population.

10.05.3 EVALUATING LEVELS OF CULTURAL CAPACITY

This section presents a more detailed discussion of key aspects of cultural capacity relevant to the training of clinical psychologists. Table 1 lists the specific knowledge, attitudes, and skills that serve as indicators of varying levels of cultural capacity in the areas of assessment, clinical intervention, and research. The table presents a progressive increase in the knowledge, attitude, and skills that serve as specific operational criteria to help gauge a clinical psychologist's progression across levels of cultural capacity from cultural sensitivity, to cultural competence, to cultural proficiency.

The criteria in Table 1 that describe levels of cultural capacity are presented as illustrative descriptors. They have not been validated empirically, although they are derived from the contemporary literature on cultural competence and from clinical and research experience (Allison, Crawford, Echemendia, Robinson, & Knepp, 1994; American Psychiatric Association, 1994; Dillard et al., 1992; La Fromboise & Foster, 1992; Szapocznik & Kurtines, 1993). Accordingly, these statements offer a viable system to evaluate levels of cultural capacity where these descriptive statements serve as guideposts for conceptualization and evaluation when conducting cultural capacity training in the areas of assessment, clinical intervention, and research. When using Table 1, cultural capacity should be evaluated in relation to a specific reference group, that is a specific and targeted cultural group or subgroup.

10.05.3.1 Evaluation Relevant to a Reference Group

One major problem in prior conceptualizations of cultural competence is that previous approaches have defined it in global and nonspecific terms. As a result, for the purpose of clinical training, cultural competence has remained a vague, unintelligible concept that is disassociated from critical clinical and organizational referents. Thus, as used previously, the concept of cultural competence has had limited practical usefulness to clinicians and to program planners. Missing is a conceptualization that defines cultural competence more accurately, as it is linked to a specific reference group. The present model aims to add greater precision and specificity to the conceptualization and measurement of cultural competence.

To operate as a construct that has practical applicability in the clinical, organizational, and community settings, cultural competence must therefore be defined in terms of skills for service delivery to members of a specific reference group. For example, through training and experience, a clinical psychologist may acquire the knowledge, attitudes, and skills that afford him or her the capacity to conduct effective therapeutic work with various African-American clients. Thus, this clinical psychologist could be rated as being culturally competent in the area of clinical interventions with African-American clients. By contrast, this same clinical psychologist may have never worked with a Asian-American client, and thus would not be culturally competent in service delivery with Asian-American clients. However, that clinical psychologist's knowledge about aspects of Asian-American cultural norms and values, his or her accepting attitude towards Asian-Americans and their culture, and his or her assessment skills for differential interpretation of tests conducted with various Asian

American subgroups, including a familiarity with scales of acculturation, could garner for this clinical psychologist a rating of cultural sensitivity for work with Asian-American clients in the area of assessment.

Based on these considerations, cultural capacity can be defined more accurately by

Table 1 General guidelines for evaluating levels of cultural capacity in assessment, clinical intervention, and research.

Assessment

Sensitivity (+1)

- Is familiar with major cultural characteristics found among members of one or more cultural group, and understands basic aspects of the within-group variability that exists for a given cultural group
- Is accepting of the concept of diversity, and is aware of the need to decenter oneself culturally in order to understand the world view of people from other cultures
- Is capable of administering and interpreting tests and assessment instruments with a consideration of sources of bias or adjustment needed in order to interpret test results in a fair manner. Can formulate culturally relevant treatment recommendations

Competence (+2)

- Is knowledgeable of and can understand apparent paradoxes and cultural nuances in the thoughts and behaviors of members of the cultural group
- Has an affective attachment to the culture and a deeper understanding of issues within that culture. Has the capacity for empathic decentering
- Is capable of understanding and working with nuances and apparent paradoxes in assessing, diagnosing, and interpreting more complex cultural aspects of thought and behavior. Can design a comprehensive treatment plan

Proficiency (+3)

- Has developed a deep, integrated knowledge of a culture based on years of experience
- Feels a deep attachment and appreciation for that culture
- Is capable of understanding nuances and apparent paradoxes. Can distil deeper cultural meaning to obtain a deep understanding of cultural thoughts and behaviors in assessing or diagnosing a problem. Can plan an optimally effective treatment plan

Clinical intervention

Sensitivity (+1)

- Is familiar with basic aspects of a culture and with the within-group variability in social norms that govern interpersonal relationships and the accepted manner in which various people communicate with one another
- Is accepting of these social norms
- Has developed ease and comfort in engaging persons from that culture in a culturally appropriate manner. Is able to decenter empathically and thus to gain an awareness of the world view of persons from the cultural reference group

Competence (+2)

- Has greater knowledge of complex aspects of social norms and patterns of communication within the culture, and is able to understand nuances present in that communication
- Has an affective bond with the culture and has a finer appreciation and acceptance of nuances and of apparent paradoxes. Is able to engage in advanced empathic decentering by understanding his or her own prejudices, as well as by being able to understand the reference culture from the “world view” perspective of members of that culture
- Is able to conceptualize the client’s problem using the most relevant therapeutic mode

Proficiency (+3)

- Has deep, rich, culturally relevant knowledge of the culture and about the rich complexity that governs that culture
- Feels a strong affective bond with the culture and expresses this in appreciation and respect for that culture and its people
- Has the capacity to detect and to work with clinical nuances and to implement one of various therapeutic approaches as is most relevant to a client’s central problem. In doing so, knows clearly where the case is going and can intervene efficiently and effectively. Is able to conduct potent culturally relevant interventions and/or design potent culturally relevant programs

Research

Sensitivity (+1)

- Is familiar with variables such as race and ethnicity as important sociocultural factors in research design and methodology. Attends to issues involved in approaching and informing individuals and communities in a manner that protects their rights and fosters trust, respect, and mutual collaboration
- Has an attitude of respect for the rights and sovereignty of individuals or communities from the culture who may participate in a proposed study
- Is able to engage individuals or communities in an ongoing dialogue that promotes a partnership in research

Table 1 (continued).

Competence (+2)

Has a knowledge of the complexities of race, ethnicity, acculturation, and other sociocultural factors as these must be conceptualized and at times reconceptualized, in order to understand the more complex aspects of a culture. Can engage in theoretical decentering by evaluating the fit or lack of fit of Eurocentric models as applied to a cultural reference group

Has a commitment to establishing a partnership with individuals or communities from a culture in a manner that offers them a voice and active ownership of the research that is to be conducted

Is capable of working actively with individuals or communities from the culture in order to ensure their participation and active ownership of the study. Develops social action research

Proficiency (+3)

Has a deep and rich knowledge of the complexities and apparent paradoxes that exist and may emerge within a culture

Has a strong affection for the core values and traditions of a culture and has a strong commitment to ensuring that these values and traditions are acknowledged and respected in the design of proposed research. This respect includes ensuring that members of the culture have a voice and active ownership of the study

Is capable of establishing a strong and enduring collaboration with individuals or communities from the culture based on earned respect and trust. This strong research bond is also used to design research that has social action value, social action research, such that its findings or discoveries can be used to benefit the individuals or communities that have participated in that research. Also, takes a leadership role in offering scholarly critique on various minority issues and in serving as a role model and mentor for junior faculty and graduate students

conceptualizing it according to three dimensions: (i) capacity level (none, sensitivity, competence, proficiency; (ii) in a specialty area (assessment, clinical intervention, research); and (iii) in relation to a specific cultural reference group (African-Americans, Latinos, Asian-Americans, Native Americans). Thus, in assessing training-related cultural capacity, a clinical psychologist may be described as having *attained* cultural competence (capacity level) *in the area of* clinical intervention (specialty area) *for work with* African-Americans (reference group). This three-factor descriptive model provides a specific and meaningful approach to guide clinical training and professional communications about the relevant cultural capacity of a given clinical psychologist as he or she proceeds through a training program.

Regarding the dimension of reference group, a targeted reference group can be defined more generally, as a primary reference group, such as for example an entire ethnic/racial population that has a definable culture such as African-Americans or Latinos. Alternatively, this reference group may be defined more specifically as a secondary reference group, that is a smaller more specific subgroup—a subpopulation, such as African-American adolescents or Latino elderly. Referring to a more specific reference group may influence the level of cultural capacity that can be rated. As indicated earlier, a clinical psychologist may have attained cultural proficiency in conducting clinical interventions with African-American adolescents, but may only be rated as having attained cultural sensitivity in conducting clinical interventions with African-American elderly.

Similarly, this three-factor model can be applied at the organizational level, where an intervention program can be rated for its capacity to deliver culturally relevant services. Based on the service components of an intervention program, the skills of its providers, its institutional policies, and its clinical environment, a program may be rated in terms of cultural capacity (none, sensitivity, competence, proficiency), in a specialty area (assessment, clinical intervention, research, or other), with a given reference group (e.g., Native American elderly).

10.05.3.2 Cultural Capacity in Assessment

Assessment refers to the use of standardized tests, interview approaches, and other diagnostic methods for systematically learning about the client (Dana, 1993). A clinical psychologist's progressive increase in knowledge, attitudes, and skills characterizes his or her increased levels of cultural capacity in the area of assessment.

At the lowest capacity level, cultural sensitivity in assessment, the clinical psychologist should demonstrate an elementary knowledge of major cultural issues that affect the life experiences of members of a special population, such as members of one of the major ethnic/racial groups (African-Americans, Latinos, Asian-Americans, Native Americans). The clinical psychologist should also be familiar with the need to decenter himself or herself from a solely Eurocentric world view in order to adopt an insider's *emic* view (i.e., one that takes the perspective of the indigenous client). This decentering allows the exploration of culturally specific concepts; these are concepts that may be unique to the indigenous culture, or concepts that are defined differently in the indigenous culture relative to how they are defined within the mainstream culture.

At the level of cultural sensitivity in assessment, the clinical psychologist should have an elementary knowledge of the within-group variability that exists within a given reference group, and should know about the moderator variables, such as the level of acculturation, that describe this within-group variability (Dana, 1993). This basic knowledge ideally facilitates an accepting attitude towards variability in culture, where variability is conceptualized as consisting of differences and not as deficits. Furthermore, this orientation should aid in reducing biased and stereotypical thinking (Oskamp, 1991), thus facilitating a more accurate assessment of clinical issues which are examined within the context of the client's cultural background.

In addition, this sensitivity in orientation should contribute to overall skills for conducting a basic assessment of the client's needs and clinical problems. This assessment should take into account the influence of cultural and social factors in drawing valid conclusions about the client's strengths and weaknesses, while also observing conflicts, relative skills deficits, and psychopathology if these are present. Attaining cultural sensitivity in assessment should enhance the clinical psychologist's capabilities for developing preliminary treatment recommendations that take into account the client's social, familial, and personal resources and barriers, as these may facilitate or impede participation in a treatment or intervention program.

As compared with cultural sensitivity, cultural competence in assessment constitutes a higher level of skills development. As contrasted with cultural sensitivity, attaining cultural competence in assessment requires a higher knowledge level, wherein the clinical psychologist understands more complex interactions and nuances as observed within the context of the client's culture. Apparent paradoxes when observed from a Eurocentric perspective may not necessarily appear as paradoxical or as psychopathology when observed from an indigenous cultural perspective. This capacity for empathic decentering when interpreting complex cultural meanings characterizes cultural competence in the area of assessment. The clinical psychologist's affective attachment to the reference group's culture would enhance his or her skills for working with cultural conflicts and with more subtle issues. Having advanced skills for interpreting manifest and more latent content, for organizing content, and for processing information would help the clinical psychologist to design an insightful and comprehensive treatment plan or intervention program.

As compared with cultural competence, at the level of cultural proficiency in assessment, the clinical psychologist would exhibit a deep and integrative knowledge of the client's culture. This knowledge would be coupled by a deep appreciation and bonding with that culture, based on having a deep personal involvement and years of experience. The culturally proficient clinical psychologist would have masterful skills for detecting and understanding subtle and complex cultural communications and for interpreting these accurately and in culturally appropriate ways. Such deep insight would allow the clinical psychologist to plan an optimally effective treatment plan or intervention program.

10.05.3.3 Cultural Capacity in Clinical Intervention

In a manner parallel with assessment, a level of cultural sensitivity in clinical intervention is characterized by a clinical psychologist's demonstrated knowledge about the prevailing social norms and patterns of interpersonal communication for a given cultural reference group. This includes having a basic understanding of the within-group variability that exists in that group. This knowledge would be supplemented by the clinical psychologist's accepting attitude towards the group's culture and values. This positive orientation towards that cultural group would also facilitate the clinical psychologist's ease and comfort in conducting therapeutic work with persons from that cultural reference group. The clinical psychologist's basic capacity for empathic decentering would also help in understanding the ethnic experience of members of that cultural reference group.

Cultural competence in clinical intervention would be characterized by the clinical psychologist's greater depth of knowledge about complex aspects of a culture's social norms, forms of communication, and types of interpersonal behavior, thus fostering the capacity to understand nuances and apparent cultural contradictions. Moreover, developing an affective bond and an appreciation for the reference culture and its people would foster the caring attitude that characterizes cultural competence in the clinical setting. This advanced level of knowledge and a caring attitude would enhance the clinical psychologist's understanding of the life conditions and the major ways of coping found among members of the reference group, and would aid in relating effectively to the client. Moreover, the clinician's ability to understand his or her own prejudices as one aspect of advanced empathic decentering would further enhance his or her cultural competence for conducting clinical interventions. This enhanced capacity to understand self and client as a dyad would also facilitate the clinical psychologist's capacity to conceptualize the client's problem using the most relevant therapeutic model, as it may apply within the client's life context. In addition, the culturally competent clinical psychologist would be able to plan and implement insightful and effective interventions or programs that address the ethnic client's most pressing needs.

As an even more advanced stage of clinical capability, cultural proficiency in clinical intervention would be characterized by a clinical psychologist's deep and rich knowledge of the reference group's culture and of the complex norms, values, and other factors that govern behavior within that culture. At this stage, the clinical psychologist would develop a strong affective bond with the reference group and its culture and would have a deep respect for its people. This combination of rich knowledge and a strong positive attitude would facilitate the clinical psychologist's capacity to comprehend the life situation of various clients from that reference group, to conceptualize a given client's problem from various theoretical perspectives, and to visualize clearly where the case is going. Thus, the culturally proficient clinician would also have the capacity to plan and implement potent clinical interventions and/or to design and implement culturally effective treatment programs.

10.05.3.4 Cultural Capacity in Research

In parallel with the knowledge, attitudes, and skills that serve as indicators of cultural capacity in a given specialty area, a clinical psychologist's cultural sensitivity in research would be characterized by his or her basic familiarity with sociocultural variables, including moderator variables, that must be incorporated into the research design and methodology of studies conducted with participants from a given cultural group. This knowledge-based sensitivity would include an awareness of procedures for the protection of potential research participants that would be recruited from that reference group. An attitude of respect for members of that culture and an awareness of their special linguistic and other needs would enhance the clinical psychologist's ability to safeguard the rights of these potential participants. In addition, cultural sensitivity would be reflected in the clinical psychologist's basic capacity to engage individuals or communities from the reference group with respect, and in his or her ability to develop a research partnership with members of that reference group.

As a more advanced stage of research capacity, cultural competence in research would be characterized by the clinical psychologist's greater knowledge of cultural interactions and of complex sociocultural processes, such as the process of acculturation, as these affect various members of the cultural reference group. This includes the clinical psychologist's capacity for theoretical decentering, which involves taking a critical examination of current Eurocentric models and their fit, or lack of fit, when applied with members of the cultural reference group. In cases in which model fit is poor, the culturally competent clinical psychologist is able to consider alternate constructivist models that feature an emic, insider's view of the ethnic person's life experiences and of the experience of being a minority person.

Moreover, the clinical psychologist's commitment to establishing a research partnership with individuals and communities in a manner that offers them a voice in the design and conduct of the proposed research would also characterize cultural competence in research (Flores et al., 1995; Norton & Manson, 1996). This in-depth knowledge and a proactive attitude that favors research partnerships and social action research would facilitate the clinical psychologist's capacity to work actively with individuals and communities from the cultural reference group in the design and implementation of social action research.

Beyond cultural competence, the more advanced level defined as cultural proficiency in research would involve advanced knowledge about the cultural reference group. This proficiency would be demonstrated by the clinical psychologist's deep and rich knowledge about the complexities and apparent paradoxes observed within the cultural reference group. This depth of knowledge would yield insights that facilitate the clinical psychologist's capacity to design significant new research. In principle, such research would be insightful and integrative, and would propose new culturally relevant models that could be adequately defined and tested. These models would include sociocultural variables, including moderator and culturally specific variables that capture, describe, and explain various sociocultural processes that affect the health of clients from the given cultural reference group.

In addition, cultural proficiency in research would be characterized by the clinical psychologist's strong appreciation for the core values and traditions of the cultural reference group and by a deep and abiding respect for these values and traditions. These proactive attitudes would be reflected in the clinical psychologist's strong commitment to safeguard the rights of prospective participants that are recruited from the cultural reference group.

The culturally proficient researcher would also exhibit research leadership in conceptualizing, designing, and implementing innovative research that examines critical issues of concern to members of the reference community. In conducting social action research, the culturally proficient researcher would pursue active solutions to contemporary social problems that affect the reference community. In such research, the investigator would exert leadership in building community partnerships and in building a research team that includes members of the targeted community in positions that contribute to the planning, design, and implementation of that social action research project.

Finally, the culturally proficient researcher would serve as a role model and as a mentor to junior faculty and graduate students, thus contributing to the training of future culturally competent investigators and teachers. The culturally proficient researcher would also adopt a leadership role in conducting critical yet helpful reviews of publication manuscripts and research proposals. This criticism would not only address issues of scientific merit, but would also address ethnic/racial issues, as well as issues of social inequality and scientific inaccuracy in manuscripts that examine ethnic minority communities and members of other special populations. In short, the culturally proficient researcher would take a proactive and leadership role in advancing a variety of minority research and training issues in the fields of clinical psychology, minority mental health, minority health, health promotion, and other related fields.

10.05.3.5 A Life-long Professional Odyssey

The road to cultural proficiency begins not with knowledge or skills, but with attitude. Given that "attitude" is defined as a predisposition to respond in a particular manner to a particular social situation (Oskamp, 1991), it is the clinical psychologist's positive attitude, an open-mindedness and willingness to learn about new cultures and people which characterizes that clinical psychologist's preparedness to embark on the journey towards cultural proficiency. Thus, instilling an attitude of openness in those who do not already have it is the first task in cultural competence training.

In addition, ethical practice in work with diverse populations requires a commitment on the part of the clinical psychologist continually to upgrade his or her cultural capabilities. To meet this need, clinical training programs should be proactive in developing curricula and resources that help clinical psychology trainees to develop their cultural competence. Here, Standard 6.01 of the Ethical Principles of Psychologists and Code of Conduct asserts that

> psychologists who are responsible for education and training [should] seek to ensure that the programs are competently designed to provide the proper experiences, and meet the requirements for licensure, certification, or other goals for which claims are made by the program. (American Psychological Association, 1992, Standard 6.01, Design of Education and Training Programs)

Relevant to this point, a study of 104 clinical training programs that examined changes in service and research with ethnic/racial populations across a 10-year period found some improvements among many of the 104 clinical training programs surveyed (Bernal & Castro, 1994). However, these improvements consisted primarily of introductory-level capabilities training, and did not offer the depth in training needed to develop cultural competence among program trainees. Bernal and Castro noted that, despite the tangible improvements in minority training in clinical psychology that have occurred since the mid-1980s relevant structural shifts have not taken place. These structural shifts refer to institutionalized programmatic changes that involve concerted plans to integrate key aspects of minority training across the clinical training curriculum, coupled with a commitment of resources to implement these plans. Indeed, only four of the 104 programs surveyed, all at professional schools, exhibited training program profiles for an integrated curriculum that would foster cultural proficiency among its trainees. Whereas many programs may express an in-principle endorsement of the importance of multicultural training, few have developed the infrastructure (coursework, minority faculty recruitment and retention, community links) for promoting cultural competence in clinical training. Given that departments and clinical training programs can also be rated on levels of cultural capacity, most programs nationally would fare no higher than being culturally sensitive. Clearly, based on the findings of the Bernal and Castro study, it can be concluded that most clinical psychology training programs in the USA can and should do more in order to develop their future capacity to promote cultural competence within their own clinical training program.

Within the context of our previous statements, we also recognize that not all clinical psychologists or students are necessarily interested in or prepared to participate in cultural competence training. Similarly, not all clinical training programs give sufficient value to cultural competence training, or do not seek to offer their students training in cultural competence as compared with other important clinical training needs. Unfortunately, those who could benefit most from cultural competence training are often those who are least interested in it. However, for those who do value learning about other cultures, and for those who can be convinced to try it, developing the cultural capacity to work with persons from various cultures is an odyssey of a professional lifetime that offers intellectual, affective, social, and spiritual rewards. The invitation to undertake this journey is there for those who are willing to take it.

10.05.4 FUTURE DIRECTIONS

The overview of the levels of cultural capacity that has been presented offers specific guidelines to help define cultural sensitivity, culture competence, and culture proficiency in the areas of assessment, clinical intervention, and research. However, in many ways, the present model only constitutes an early step towards a more detailed articulation of levels of cultural capacity. Additional work can and should be conducted to relate the guidelines presented here to a variety of specific cases with the aim of enhancing quality in assessment, treatment, and research with a diverse group of ethnic/racial clients.

Moreover, clinical training programs can and must do more to add depth of content and diversity in their clinical training activities, as these additions facilitate greater competence in assessment, clinical intervention, and research with members of ethnic/racial populations. One potential outcome of presenting the current three-factor model and its guidelines is that it may serve as a template from which to rate the level of cultural capacity of given clinical training program; that is, to rate the current capacity of the program to train culturally proficient clinical psychologists.

10.05.5 SUMMARY

In this chapter a three-factor model has been presented that aims to define operationally the various levels of cultural capacity (sensitivity, competence, proficiency) as related to three specialty areas (assessment, clinical intervention, research), and as related to a specific cultural reference group. It is recognized that attaining true cultural competence in the areas of assessment, intervention, and research for each of several ethnic/racial populations is a high ideal that will be achieved by only a few clinical psychologists in a lifetime of professional practice. However, this chapter has also identified and defined graded steps towards this ideal, and encourages the clinical psychologist to commit a sustained effort towards enhancing his or her level of cultural capacity. Here also, clinical training programs have a professional obligation to upgrade their training program by adding depth of content and activities in order to enhance their programmatic capacity to offer true depth in cultural competence training. It is recognized that the present three-factor model constitutes only an early step in the task of further articulating the characteristics of various levels of cultural capacity, as this articulation aids in clarifying further what it means to become culturally competent.

10.05.6 REFERENCES

Allison, K. W., Crawford, I., Echemendia, R., Robinson, L. V., & Knepp, D. (1994). Human diversity and professional competence: Training in clinical and counseling psychology revisited. *American Psychologist*, *49*, 792–796.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist*, *47*, 1597–1611.

Anderson, N. B. (1995). Behavioral and sociocultural perspectives on ethnicity and health: Introduction to the special issue. *Health Psychology*, *14*, 589–591.

Areán, P. A. & Gallagher-Thompson, D. (1996). Issues and recommendations for the recruitment and retention of older ethnic minority adults into clinical research. *Journal of Consulting and Clinical Psychology*, *64*, 875–880.

Arizona Department of Health Services (1995). *Cultural competency in the administration and delivery of behavioral health services*. Phoenix, AZ: Author.

Balcazar, H., Castro, F. G., & Krull, J. L. (1995). Cancer risk reduction in Mexican American women: The role of acculturation, education, and health risk factors. *Health Education Quarterly*, *22*, 61–84.

Baldwin, J. R., & Lindsley, S. L. (1994). *Conceptualizations* *of culture*. Tempe, AZ: Urban Studies Center.

Bernal, M., & Castro, F. G. (1994). Are health professionals prepared for service and research with ethnic minorities? Report of a decade of progress. *American Psychologist*, *49*, 797–805.

Cross, T. L., Bazron, B. J., Dennis, K. W., & Isaacs, M. R. (1989). *Toward a culturally competent system of care*. Washington, DC: Georgetown University Child Development Center.

Dana, R. H. (1993). *Multicultural assessment perspectives for professional psychology*. Boston: Allyn and Bacon.

Department of Commerce (1994). *Statistical abstract of the United States: 1994* (114th ed.). Washington, DC: US Government Printing Office.

Department of Health and Human Services (1985). *Report of the Secretary's Task Force on Black and Minority Health*. Washington, DC: US Government Printing Office.

Dillard, M., Andonian, L., Flores, O., Lai, L., MacRae, A., & Shakir, M. (1992). Culturally competent occupational therapy in a diversely populated mental health setting. *The American Journal of Occupational Therapy*, *46*, 721–726.

Flores, E. T., Castro, F. G., & Fernandez-Esquer, M. (1995). Social theory, social action, and intervention research: Implications for cancer prevention among Latinos. *Journal of the National Cancer Institute Monographs*, *18*, 101–108.

Hohmann, A. A., & Parron, D. L. (1996). How the new NIH guidelines on inclusion of women and minorities apply: Efficacy trials, effectiveness trials, and validity. *Journal of Consulting and Clinical Psychology*, *64*, 851–855.

Kim, S., McLeod, J. H., & Shantis, C. (1992). Cultural competence for evaluators working with Asian American communities: Some practical considerations. In M. A. Orlandi, R. Weston, & L. G. Epstein (Eds.), *Cultural competence for evaluators* (pp. 203–260). Rockville, MD: Office of Substance Abuse Prevention.

La Fromboise, T. D. & Foster, S. L. (1992). Cross-cultural training: scientist–practitioner model and methods. *The Counseling Psychologist*, *20*, 472–489.

Miranda, J., Azocar, F., Organista, K. C., Muñoz, R. F., & Lieberman, A. (1996). Recruiting and retaining low-income Latinos in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *64*, 868–874.

Montgomery, P. A. (1994). The Hispanic population in the United States: March 1993. *Current Population Reports—Population Characteristics* (Series P20-475). Washington, DC: US Government Printing Office.

Murphy, R. F. (1986). *Cultural and social anthropology: An overture*. Englewood Cliffs, NJ: Prentice-Hall.

National Institutes of Health. (1994). *Outreach notebook for the NIH guidelines on inclusion of women and minorities as subjects in clinical research*. Bethesda, MD: Author.

Norton, I. M., & Manson, S. M. (1996). Research in American Indian and Alaska Native communities: Navigating the cultural universe of value process. *Journal of Consulting and Clinical Psychology*, *64*, 856–860.

Orlandi, M. A. (1992). Defining cultural competence: an organizing framework. In M. A. Orlandi, R. Weston, & L. G. Epstein (Eds.), *Cultural competence for evaluators* (pp. 293–299). Rockville, MD: Center for Substance Abuse Prevention.

Oskamp, S. (1991). *Attitudes and opinions* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Szapocznik, J., & Kurtines, M. (1993). Family psychology and cultural diversity: opportunities for theory, research and application. *American Psychologist*, *48*, 400–407.

The National Latino Behavioral Health Workshop (1996). *Cultural competence standards in managed care mental health services for Latino populations*. Phoenix, AZ: Author.

Thompson, E. E., Neighbors, H. W., Munday, C., & Jackson, J. S. (1996). Recruitment and retention of African American patients for clinical research: An exploration of response rates in the urban psychiatric hospital. *Journal of Consulting and Clinical Psychology*, *64*, 861–867.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.06

Reconstructing Race, Rethinking Ethnicity

MARIA P. P. ROOT
University of Washington, Seattle, WA, USA

10.06.1 INTRODUCTION 141
10.06.1.1 Race and Ethnicity 143
10.06.1.2 Racial Identity Development as Racial Healing 144
10.06.2 PROCESS OF IDENTITY CONSTRUCTION 145
10.06.2.1 Process of Differentiation 145
10.06.3 RACIAL IDENTITY THEORIES 147
10.06.3.1 Nigrescence and Minority Racial Identity Stage Model Theories 147
10.06.3.2 White Racial Identity 148
10.06.4 ETHNIC IDENTITY MODELS 149
10.06.4.1 Comparison and Critique of Racial and Ethnic Identity Models 150
10.06.5 RETHINKING RACE AND ETHNICITY 151
10.06.6 CONTEMPORARY MODELS FOR RETHINKING THE RELATIONSHIP BETWEEN RACE AND ETHNICITY 152
10.06.6.1 Ecological Model of Race and Ethnicity 152
10.06.7 CLINICAL IMPLICATIONS 155
10.06.8 FUTURE DIRECTIONS: RETHINKING RACE AND ETHNICITY 157
10.06.9 SUMMARY 157
10.06.10 REFERENCES 158

10.06.1 INTRODUCTION

Ethnicity and race in the USA have become synonymous concepts in everyday language although they signify different experience. This chapter, which considers identity development in persons of mixed racial heritage, necessitates we differentiate these concepts. Increasingly, ambiguous phenotypes, proliferating since the biracial baby boom post-1967, defy instantaneous recognition or accurate assignment of ethnic group belonging (1967 marked the repeal of the last antimiscegenation laws in the USA forbidding and penalizing interracial marriage. Since this period of time the percentage of interracial marriages has approximately doubled with each subsequent decade. Correlated with this increase in legalized interracial unions has been a significant increase in the number of biracial and multiracial children and young adults, particularly in large cities of the USA (Root, 1992, 1996b)). These constructs seemingly collide when we consider the identities of people who claim multiple group memberships. This chapter uses the presence and experience of a growing presence of racially mixed people in the USA whose simultaneous, dynamic, and multiple identies require reexamination of assumptions about race and ethnicity.

How are racial and ethnic identity different? This has become an increasingly difficult differentiation as racial logic is somewhat circular and the distinctions between their historical constructions are suppressed (Gordon, 1995). Both exist because diversity in race and ethnicity have a social significance used to stratify society (Phinney, 1990). Much of the work on ethnic and racial identity stems from notions of the effects of marginality (Oetting & Beauvais, 1990–1991) and the observations and hypothesizing about the hybrids of race and culture (Park, 1931; Stonequist, 1964; Young, 1995).

Race and ethnicity become significant statuses relevant to experience of self and others when they are differentiated within a society. In his classic work on the roots of discrimination, Allport (1994, p. 221) lists 10 conditions which when present are accompanied by an increase in prejudiced personalities:

> Where the social structure is marked by heterogeneity
> Where vertical mobility is permitted
> Where rapid social change is in progress
> Where there are ignorance and barriers to communication
> Where the size of a minority group is large or increasing
> Where direct competition and realistic threats exist
> Where exploitation sustains important interests in the community
> Where customs regulating aggression are favorable to bigotry
> Where traditional justifications for ethnocentrism are available
> Where neither assimilation nor cultural pluralism is favored.

If one carries a visible status such as race, its socially constructed meaning results in a labeling process. Visible identities, whether they connote positive meaning or stigmatized status, effect the opportunities we encounter (Goffman, 1963). The process of incorporating a stigmatic status requires a personal understanding of human competition, oppression, and resilience. When these identities carry stigmatic statuses, the process by which they are incorporated into a positive aspect of self requires ego strength, resilience, and a repertoire of coping. Thus, identity is a very central concept for clinical, developmental, and social psychology.

The construction of race in the USA is not unique. Several scholars have elucidated the construction of racial and ethnic others as a colonial strategy (whether it be the Spanish, Dutch, Japanese, British, American, etc.) for justifying maltreatment of people (Foucault, 1978; Young, 1995). The construction of the racial or ethnic other sets the stage for justifying genocide whether it be in Bosnia, World War II Germany, the USA, or Africa at various points in history.

Because the mechanics of oppression have universal methods, the example of US history of race relations to examine current theories of racial identity development is used. In doing so, it becomes clear that our psychology exists within a political tangle of propaganda and oppression—past and present. The effects of our history of racial and ethnic oppression becomes evident in the critical examination of current racial identity theories from the experience of a growing minority in this country, persons of multiracial origin, who declare a biracial or multiracial heritage.

DuBois (1903) noted that the affirmation of a devalued self can create a depth of conscious awareness that possibly expands one's humanity. Almost three-quarters of a century later, the development of ethnic and racial identity theories elucidate the phenomenological process largely affected by a political history (Ramirez, 1983, 1997; Root, in press). Allport's (1954) conditions of almost a half century ago work well in understanding how sociology, politics, and psyche are interdependent. The 10 conditions he outlines shape the social intercourse of identity. Identity development will be necessarily a conscious and more complex status for racial and ethnic others. In this chapter, the biracial or multiracial person who chooses not to adhere to centuries-old rules of racial classification and insists upon simultaneous or blended identities is chosen as the example with which to explore race and ethnicity.

Differentiating between race and ethnicity becomes more important at this moment in history because of the trend in increasing miscegenation and the subsequent, significant biracial baby boom. This young, visible cohort, contests the virulent antimiscegenistic foundation upon which racial rules for belonging and identification derived and even Stonequist's hypothesized marginality (1964). Furthermore, recent research using multiracially identified individuals suggests that race is neither a necessary nor sufficient condition for establishing ethnic identity (Hall 1980, 1992; Stephan, 1992). The psychological models guiding and capturing racial identity are now dated and limited. Current and future directions must be plotted against the backdrop of a post-civil rights era that has spawned a biracial baby boom.

10.06.1.1 Race and Ethnicity

Some scholars suggest that the racially mixed person embodies one of the most tangible threats to the current racial order. In order to understand this threat, an historical understanding of the derivation of race is necessary. Without this examination, culture and ethnicity are misleading euphemisms for a construct laden with centuries of harm.

Colonists striving for an economic and political stronghold in new lands needed to create a system with which to distinguish themselves from the indigenous people. The construction of race was intrically aligned with economic systems of expansion. This system racialized gender and engendered race (Foucault, 1978; Stoler, 1995; Zack, 1997b). Whether this be through Dutch, Spanish, British, Japanese, or US colonization, the mechanisms have manufactured similar oppressions. Ethnocentric interpretations of differences in cultural practices are coded through the colonized group's bodies (Fanon, 1967; Foucault, 1978; Stoler, 1995; Young, 1995; Zack, 1997a). Subsequently, notions of group distinctions on moral and behavioral bases are attributed to lineage which eventually becomes accepted scientific folklore as biological heritability. A mythology and rationale for oppressive treatment of the colonized completes an irrational circle of logic to ensure that the physical demarcations provide clear markers of insider and outsider (Spickard, 1992). Utilizing paternalistic attitudes coupled with the belief in biological heritability of morality, intelligence, and civility (Young, 1995), the colonized group is considered helpless to change its outlook or status, which justifies paternalistic and missionary interventions introduced with imperial expansion.

Whereas racial classification systems are determined by those in power, those subjects disenfranchised by this system often unwittingly internalize it; subsequently, groups agree upon assignments between the colonizer and the colonized. For example, in the USA, hypodescent, a system by which a mixed race person is assigned to the group with the lowest social value, is now an agreed upon system. Using an extreme example, many black people will insist a person is black even if they grow up with a single white parent in a predominantly white neighborhood with little African-American ethnic group knowledge or shared customs, behaviors, or attitudes.

Weinreich (1986) suggests that ethnicity requires internal recognition and affiliation with the ethnic reference group. Mutual agreement between participant and group regarding group membership is based upon behavioral practices, rituals, customs, and values. However, because we are such an ethnically diverse society, and ethnicity is centrally defined by the previous variables and not defined phenotypically, ethnic ambiguity is reduced through using racialized bodies as signifiers of ethnic group membership. Whereas the enactment of ethnicity is dynamic over time and generation, race is repeatedly enacted within the constricture of laws of hypodescent and, until recently, laws against miscegenation. Among other purposes, these latter laws kept phenotypical signifiers of group belonging more visible. Given our country's history of racial tragedy, dialogue about race is avoided through the use of terms as ethnicity and culture that have come to be used interchangeably for race despite their essential difference.

Both ethnicity (Tajfel, 1981) and race (Cross, 1971; Helms, 1990) are salient identities of the self that provide a basis for interpreting the actions of others and one's internal experiences. Their similarities are perhaps bound both by the sense of reference group or belonging that one psychologically internalizes. One might construct their difference by viewing race as constructed by negative forces which subsequently require connectivity between members of a disenfranchised group for survival. In contrast, the connectivity between people derived from ethnicity includes positive experiences of sharing rituals and customs and language on daily basis. When negative forces drive definitions of ethnicity from outside the group, such as in anti-semitism, Jewishness is also constructed in racial rather than only ethnic terms, and the negative connectivity resulting from race also binds this ethnic group for survival.

Ironically, colonization and physical proximity between different populations inevitably leads to miscegenation of populations (Spickard, 1989; Stoler, 1995; Young, 1995). Initially the offspring may have the privileges, but when a critical mass starts to develop, these bodies may become viewed as the enemy within (Stoler, 1995), and this new group is relegated to another position. This system is challenged and breaks down when a critical mass of people refuse to collude with the system. This is the case with a portion of the contemporary generation of mixed racial heritage. The numbers of people born out of a biracial baby boom who do not wish to participate in conventional societal racial antagonisms has reached a critical mass.

In summary, race is based upon a delusional system of hierarchical difference marked by phenotypical differences. Systemically, race is not initially an agreed upon system. Assigned by those in power, group acceptance and belonging is based upon recognition and acceptance by the disenfranchised group. Whereas white society and groups of color are congruent in their assignments and acceptance of racial group membership, this system breaks down contemporarily with the refusal of many people of mixed heritage to agree to participate according to the rules of the system.

We must reexamine the assumptions upon which racial identity has been based as we reach the centennial anniversary of DuBois' observation that the color line would be one of the most significant issues of the twentieth century. We must shift our concern at the end of this millennium to the microcosm of this American dilemma—the individual who embodies multiple heritages and affirms multiple allegiances. What are the psychological implications of this for us as a nation and for individual identity development (Root, 1996b)? Current data suggest that the biracial baby boom is accelerating an unraveling of race as it has been practiced in the USA for centuries. These changes have implications for psychological constructions of race and racial identity through the dilemas posed by demographic changes.

10.06.1.2 Racial Identity Development as Racial Healing

Racial identity is necessarily defined in terms of a psychological process of healing from the insidious wounds of racism. In the broadest sense, the damage caused by justifying racial inferiority of nonwhite people has similar dynamics to hostage-taking during war. Interestingly, Banks' (1988) ethnic identity development theory posits the first stage as "ethnic psychological captivity." The tactics of isolation of the individual or group, deprivation, monopolizing of perceptions, and brainwashing result in a dependence on the captor, a belief in one's inferiority, and an aspiration to be more like the captor in hopes of survival, or even better treatment.

Furthermore, contemporary literature makes the case that race and gender are co-constructed (Zack, 1997a). Lerner (1986) suggest that men of color are constructed in the female gender. Historically, she makes the case for observing that desirable women have been constructed in the image of children. Paternalistic models of racism also construct people of color and indigenous people in images and caricatures of children.

Similarly, the dynamics of race relations have shared similar aspects to the chronic abuse of children or women in domestic households. In these positions women and children are often held hostage economically, materially, and psychologically. Their survival and welfare depends on developing a keen sensitivity to the whims and wishes of the captor or head of household. Feelings of inferiority, wishes to be other than female, and devaluing other women may follow (Griffin, 1992).

With the metaphors of war and the insidious psychological violence wreaked by the abuse of dominance, a healing process is required to establish a constructive sense of self which affirms the possibility of a positive racial and ethnic identity. In the meantime, however, what is the cost to individuals?

Rates of depression, schizophrenia, and stress reactions have been associated with race and social class which are confounded in this society (Dowrenhend & Dowrenhend, 1974). Although depression and schizophrenia exist across cultures, their rates vary across countries suggesting mediation by environmental influences. The US literature points to the stress caused by self-devaluation by racial system in terms of obstacles to economic advancement through glass ceilings and limited employment opportunities, and stress caused by living in unsafe neighborhoods that are also associated with impoverished levels of income, joblessness, and disillusionment by young people (Wilson, 1987). Recent literature investigates the relationship of stress caused by racial position and impact on mental health such that seeming personality disorders must be viewed in cultural contexts (Alarcon & Foulks, 1995). Unfortunately, the standard reference group unchallenged until more recent times has been similar to the researchers: white, male, and middle to upper middle class. Clients or students whose behaviors or psychological profile differed significantly tended to be pathologized or problematized. For example, a significant body of literature exists on the personality problems of the person of African-American heritage, particularly in the use of the Minnesota Multiphasic Personality Inventory. A debate over the intelligence of minority group children, particularly those of African heritage or of bilingual backgrounds, resurfaces every 15–20 years (Hilliard, 1996).

The civil rights movement, and legislation that accompanied it to extend equal rights guaranteed by the US constitution to all people, beginning in the 1950s and extending two more decades, might best mark the beginning of public awareness of the healing oppressed persons continued despite prevailing conditions in which prejudices thrived. Legislation attempted to correct previous misapplications of the US constitution and the Bill of Rights through formal desegregation of schools, workplaces, and neighborhoods. In time, a new body of literature emerged on ethnic and racial identity, produced primarily by people of color who were now part of a small and growing cohort of scholars. It was no coincidence that the primary originators of racial identity theories were people who were deeply influenced by the racial pride movements of the 1960s and 1970s, which occurred during their young adult years.

During this period in history, ethnicity and race essentially became the same; ethnic solidarity meant racial solidarity (Omi & Winant, 1986). Ironically, notions of ethnic/racial solidarity invoked an inside–outsider politic akin to the process outlined for individuals in racial identity theories; the oppressed person, in an attempt to selvage their racial self-esteem, separates themselves from the majority to seek refuge and absorb the positive aspects of identifying as a person of color. This period of refuge may also reflect a hatred of what is white and dominant, accompanied by an uncritical acceptance of racial group signifiers that one idealistically reveres. Phenotypic identification by styles of dress and physical appearance take on significant meaning as ethnic markers and symbols of racial pride and solidarity. Thus, a literature emerged on racial identity that was virtually synonymous with African heritage and ethnicity (Cross, 1971). Simultaneously, within the ethnic pride movements of the American Indian Movement (AIM), the Chicano Movement (La Raza), and the Asian American Movement, other theorists emerged who focused on the meaning and importance of biculturality (Ramirez, 1983).

10.06.2 PROCESS OF IDENTITY CONSTRUCTION

Human nature observes difference and leads us to differentiate ourselves at a basic level from "the other," someone different from us on a dimension that has meaning. Thus, part of building self-definition is through contrasting ourselves with the other (Sartre, 1976). This process of differentiation is initially primitive rather than complex, casting difference in binary, oppositional terms establishing poles of privilege and disenfranchisement (Zack, 1993): male, female; white, not-white; able bodied, disabled; same sex orientation, different sex orientation. Gordon (1995) outlines the logic of this transformation as it has been applied to race. A superior group is boundless from constraints and derives its justification from the proof of the inferior. In turn, the inferior is treated and constructed in less than human form. This oppressive process is the process of colonization of a people and colonization of the mind.

The body of the individual, through assigned social meaning, is stigmatized (Goffman, 1963; Young, 1995). In more recent work, Gordon (1997) uses a philosophical perspective to suggest that to have a race, in an antiblack world, is to be assigned the pole of raced, that is, black. To be white is raceless. A similar physical marker is assigned to gender. To have a gender is to be female; male is genderless. In contrast, Zack (1993) suggests that the polar opposites on race are to have a race (black, white, etc.) and its opposite is hybridity. Thus, racial mixing derived from either system of logic has the negative valence and has the most significant meaning attached to its embodiment. Racial mixing also brings in the topic of control of sexuality through privileged positions of bodies, both raced, and gendered.

In patriarchal societies the ideal self is referenced as male, white, able-bodied, and heterosexual. Sociological theories of anomy (Durkheim, 1966) and marginality (Stonequist, 1964) conclude that to possess other than the ideal and privileged position is to be inherently less powerful and more likely to be marginal. Activists and scholars concerned with the impact of colonialism make similar conclusions (Fanon, 1967; Friere, 1970). To function in the other position is to have more obstacles to overcome and to have to fight harder to derive a positive self-identity. Erikson's (1968) theory on the importance of identity and conscious identity process suggests that to be defined as other will predict at the least a more arduous task of positive self-identification. (Whereas he tackles the topic of race he does it also as a representative of the times in which he was produced; gender identity development receives relatively little attention (Burman, 1994).) The ethnic and racial identity theories that developed in the 1970s and 1980s are very much a consequence of this arduous task.

10.06.2.1 Process of Differentiation

Deriving a sense of self from contrast with the other begins a process of differentiation (Allport, 1954). Much contemporary literature has focused on the social construction of race and virtually little on the process of differentiation that gives rise to meanings assigned to differences. This process has different paths it can take. Root (in press) notes that the process of differentiation can be constructive or destructive, which promotes very different cognitive schemas around differences.

Tolerance for ambiguity of meaning and ability to operate fearlessly in the face of difference allows one to engage in the process of constructive differentiation. Tolerance for ambiguity allows the individual to suspend stereotyping and to refrain from applying conceptual frameworks that do not fit the situation. Differences may be observed, but neither a binary schema nor a valence is assigned. Thus, the other is observed side by side with the self rather than in competition or opposition. The other does not have to be inferior. Constructive differentiation observes difference outside of a hierarchical schema and allows one to expand one's world view, and even possibly find oneself in this difference. The ability to do this, particularly in light of observable physical differences, such as by color, by gender, or by height, is stifled early on by external inputs that impart hierarchical, competitive schemas and schemas of opposition. Categorization seems to be an important part of the cognitive process that allows the world to be made simpler. However, categorization is not necessarily synonymous with hierarchical stratification. It is the step from categorization to hierarchicalization which may shift the process of understanding difference into destructive differentiation.

Unfortunately, by the age of seven or eight, the process of destructive differentiation around race is well assimilated. Remember the song in the 1950s musical, *South Pacific* set during World War II: "you've got to be taught/before it's too late/before you are six, or seven, or eight/ to hate all the people your relatives hate." Underlying such a process is a tendency to interpret the other as threatening and to subsequently counter anxiety or threat with negative evaluation. Through a process of negative stereotyping, threatening differences are stratified (Berger, Cohen, & Zelditch, 1966) and stigmatized (Goffman, 1963).

Negative differentiation employs parsimonious, reductionistic strategies too early. Subsequently, this process yields primitive rather than complex understanding, though, complex and convoluted rationalization may develop to sustain these reductionistic schemas, as has been the case in using race as a reductionistic scheme. Racial epithets are an extreme example of destructive differentiation. With a few words, an overtly hostile act of differentiation relegates the target person to a less desirable position. In this act, the instigator defines themselves as superior in contrast to their target.

Ultimately, negative differentiation results in self-definition by what one is not—a distancing from feared association with the other. For example, a person may have ideas that they are not like Filipinos, not like Africans, but still not have a constructively defined notion of a racial self as a white person, that is, whiteness as the privileged position is inherently defined as raceless or without stigma. (In the USA, many white persons no longer have an ethnic affiliation, so that the primary identity is through white racial status rather than through an ethnic identity). Many researchers note that, unlike the identity process that persons who have been defined as other experience in this country, white persons do not have to think about themselves on a daily basis in terms of their race or ethnic belonging (Frankenburg, 1993; Helms & Carter, 1990). Destructive differentiation leaves little room for exploration on the very dimension one might use to expand one's world view. The mixed heritage person, socially located as other, and invisible, but feared and mistrusted, must struggle to create positive self-definition in the face of few positively valued reference groups or role models.

Destructive differentiation produces an identity that is inherently fragile, dependent upon maintaining distance from the other, and believing in the inferiority of the other in order to maintain a sense of self. Its unresolved conflicts regenerate itself; the oppressed may become the oppressor (identifying with the captor) as a survival strategy, yielding again, a fragile sense of identity. For example, given that this country's racial system is predicated on pure race, mixed-race people experience gatekeeping around racial belonging. Acceptance requires a singular belonging. Thus, persons who declare a biracial identity are often put to racial authenticity tests by other people of color. If they are also of white European origin, centuries-old rules of hypodescent automatically fail them in white authenticity. Multiple allegiances and affiliations of multiracial or multiethnic people may be vilified through the process of negative differentiation. On another level, women may be more oppressed than dictated by cultural rules as a means for men who feel oppressed to achieve a position of superiority. Unfortunately, this yields a fragile self construction. This juxtaposition of the inferiorizing of racial and gendered identity can provide the basis for some racial antagonisms between parents and mixed heritage children in some families (Kich, 1992).

Ironically, political consciousness pertaining to racial and ethnic identity is a response to negative differentiation in which being the other has simultaneous negative meaning and invisibility (Freire, 1970). Hurtado and Gurin (1995) discuss three interrelated aspects of racial and ethnic identity that originate out of a process of negative differentiation for individuals of

Mexican heritage. These issues are also relevant to other groups of people who may locate themselves in a multiracial and/or multiethnic context. First, the group is indignant over the lack of power; for example they might protest over lower wages than other ethnic groups in the same work environment performing the same work. Second, the group demonstrates that this social location and lack of power results in suffering healthwise, economically, and/or educationally. Third, collective action is necessary to improve the social standing of the group, such as strike action by groups of persons who occupy niches of wage work that capitalize on their lower status, but are necessary for the profit of a capitalist society. All three aspects of this increased political consciousness are directly implicated in the process of racial awareness and the personal meaning of racial identification.

Diminishing denigrating defensive strategies is difficult in a society that has colluded and condoned negative stereotyping such as that based on colorism, race, body size, gender, and so on. Bennett (1995, p. 36) suggests that confrontation of negative stereotyping or denigrating tactics may result in more overt protests that are denial defenses. In contrast to Bennett's suggestion that superiority or grandiosity is a defense that does not require overt denigration of the group constructed as other, it is a stratification defense against fragility. This defense, too, is a form of destructive differentiation through stratification driven by anxiety and primitive threat.

10.06.3 RACIAL IDENTITY THEORIES

10.06.3.1 Nigrescence and Minority Racial Identity Stage Model Theories

Numerous theories blend racial and ethnic identity. The most numerous models are found in the nigrescence models of racial identity offered by African-American researchers (Banks, 1988; Cross, 1971; Gay, 1984; Jackson, 1975; Milliones, 1980; Parham, 1989; Thomas, 1971; Vontress, 1971). Most of these models emerged in the context of working with clients of African-American heritage in counseling or educational settings. In both settings, the standard for behavior and mental health used a middle-class, white, male reference group similar to the majority of researchers. African-American researchers started depathologizing and renorming behavior in the social and historical context of healing from the wounds of racism. Significant reviews of this literature have been provided over the course of its development (Carter, 1995; Cross, 1978; Helms, 1990).

Although all these models imply some process that responds to racial traumatization, some might best be described as typology models and others as stage models. Helms (1990) summarizes the stages or categories proposed by each model in the early body of identity models. The models that have had most impact on contemporary thinking about racial identity are the stage models.

These models, particularly the early ones, were developed with little influence from each other. Thus, it is all the more remarkable how similar they are in process, varying in the number of stages from four to six. The assumption is that an individual proceeds linearly through the stages with the final stage synonymous with optimum mental health. Crises, which may be personally defined, are the catalysts to rethink one's experience and assumptions about race and how they fit into that schema.

First stages of these models typically suggest that the individual is hostage to the prevailing racial system. Both a racial and ethnic identity model, Banks' (1988) first stage is labeled "ethnic psychological captivity," connotating the culmination of destructive differentiation. In the first stage of these models, the individual holds the prevailing stereotypes which denigrate the group to which one is assigned. As a result, one may blame the self for plights and difficulties, ignoring systemic obstacles and barriers to success. They may insist that history has changed so that opportunities for equitable treatment exist. Thus, they are able to hold a just world hypotheses view of the world regarding race: People basically get what they deserve and that right behavior is rewarded and the individual is valued (Janoff-Bulman, 1992). When one encounters a significant challenge to this perception of the world, their assumptions are shattered and they must recreate a new understanding of how the world works.

A typical second stage in these models suggests that the individual withdraws into the black culture to reconstruct their world; simultaneously, they may employ negative differentiation by relabeling the binary poles of opposition. That which is associated with the dominant culture is worthless and reprehensible. The default racial label for dominance is white and engendered as male. Labels such as "encounter," "immersion," "separation," and "confrontation" characterize the process at work in this stage. Typical of the process of quick change which swings to an opposite pole, this stage might be characterized by over-idealization of the new world view, at times exaggeration of the evils of the dominant world view, and exaggeration of the flawlessness of one's racial/ethnic group. The racial pride movements of the 1960s and 1970s typified this stage of ethnic identity. Everything black, Chicano, Indian, or Asian was idealized by the respective groups. Those aspects of the culture that may not be positive may be ignored. Theorists reflected what they observed in the process of clients and peers. In order to counter the abundant negative messages one must denigrate the practices and identities associated with dominance. Out of this stage emerges a new basis for an identity that rebuffs the negative images of self. What has received little attention is that the default gender is black male, or male of color. Female gender and the co-constructed process of identity is not addressed.

With labels such as "internalization," "post-encounter," and "affirmation mode," the next stage suggests that one works to be able to internalize positive images of blackness that are not necessarily at the expense of denigrating whiteness or any other group. Thus, negative differentiation is less prevalent as an organizing strategy. One continues to build pride in one's self associated with ethnic and racial heritage.

Some models originally offered a stage that suggested it was possible to move towards a multicultural position and appreciation of diversity (Banks, 1988; Cross, 1971, 1991; Milliones, 1980; Thomas, 1971). Labeled as "integration," "internalization–commitment," or "global competency," individuals in this stage attempt to use experience and knowledge to fight oppression around them. Solidarity is defined less stereotypically as behaviors; symbols are not used for quick judgments or categorizations of other people. Constructive differentiation is an organizing schema. This stage implies a level of healing of rage and anger that requires life experience and some luck of positive encounters.

In more recent work and reworking of these models, a last stage seems to be informed by the life stage of the researchers. This stage is one in which there seems to be a transcendence of the constrictions of race, while maintaining an awareness that race does make a difference (Atkinson, Morten, & Sue, 1979; Carter, 1995; Cross, 1985, 1991).

Most of these models exist within a dichotomous racial framework of white and not white, thus, capturing the prevalent racial paradigm and the prevailing racial politics of the time in which they emerged. They assume that the two major operating rules of race (pure race and hypodescent) abound so that there is a racial hierarchy and the races are separate. Atkinson et al. (1979) developed a minority identity model to reflect similar process in ethnic identity and racial identity development across groups. However, this model as the other racial identity models, does not explain the process of identity development when a multiracial paradigm exists and other aspects of identity coconstruct identity.

10.06.3.2 White Racial Identity

Starting in the late 1970s through the 1980s researchers conceptualized white identity stage models (Carney & Kahn, 1984; Ganter, 1977; Helms, 1984). Again, Helms (1990) provides a summary of these models. An identifiable process of white racial identity formation has been offered, that is, moving from assuming a raceless position to acknowledging that whiteness comes with inherent privileges and power.

These models suggests that a white person's linear movement through the stages progresses towards establishing egalitarian relationships with persons different from oneself. Such an accomplishment would again require an active use of constructive differentiation to interpret differences. These models vary between offering three to six stages. These descriptions of the process through which an individual moves are not as uniform as the black identity models. The labels are less descriptive in these models, some preferring to refer to stages as phase or stage 1, 2, 3, and so on.

Although it is hard to typify the stages across models, the first and last stages bear similarities across models. The first stage typically captures the acceptance of the white dominant position in society with an obliviousness to this being a racial identity. Without acknowledgment of white racial identity at work, an individual can protest the notion that the status quo associated with white domination economically and politically promotes racism—or that these white individuals serve as pawns in the racial machinery governing the social structure of this country.

In the last stage of the models, the individual acknowledges the significance of white racial identity and accepts responsibility for the losses associated with white superiority and domination. The individual moves to broaden their standards and values of reference beyond their own. Unlike the nigrescence or minority models, none of the models reviewed by Helms (1990) suggest that fighting oppression is inherent to transcending the constrictures of race as constructed in this country.

The racial identity models, that developed out of the counseling and educational contexts, have been conceptually linked to mental health: well-being, beliefs, and behaviors. However, the empirical testing is just beginning. No conclusive data is available.

10.06.4 ETHNIC IDENTITY MODELS

Outside the bipolar black–white racial framework, multiracial populations exist which are primarily defined by ethnicity (e.g., Mexican, Latino, Puerto Rican, Native Hawaiian, Native American, and Filipino) rather than race (e.g., black, white). Although the embodiment of this ethnicity may be physically blended with non-European features, these groups fall outside of the most contrasted borders of black–white. Examination of the process of identity that occurs for these groups of people is important when considering the question of how race and ethnicity might be different despite phenotypic variation within a population (Ramirez, 1983, 1997). The ethnic identity models do not contend with race directly as much as they do with cultural difference and shared cultural elements that define an ethnic community.

Sociological theories suggest that several interactional factors affect ethnic identity: the individual, the society, and the group to which an immigrant is assigned reactions, positive or negative, and the interaction of these three social locations (Mittelberg & Waters, 1992). The symbolic interactionist theories highlight the interaction between the individual or collective and the environment from the actor's point of view (Lal, 1995). These models become precursors for multidimensional psychological models which emphasize the ecology of the environment within which the individual acts. Furthermore, many sociological concepts of ethnicity explore which factors influence the boundaries of ethnic identity. Within this framework, certain visible characteristics of an individual restrict options for ethnic identification, such as race and language facility (Mittelberg & Waters, 1992; Waters, 1990).

Psychological models of ethnic identification introduce concepts of acculturation and assimilation as source of stress, directing the outcome of attempts to integrate a bicultural world view (Olmedo, 1979; Padilla, 1980; Ramirez, 1983, 1997). Much of this work emerges from research with youth (Oetting & Beauvais, 1990–1991; Padilla, 1980; Phinney, 1989; Ramirez, 1983; Trimble, 1996, in press). Gurin, Hurtado, and Peng (1994) concluded the macrosocial conditions, such as language dominance, length of residence, geographic dispersal, and the diversity of work settings, that influenced the opportunities for group contact which in turn influence the formation of social ethnic identities in Chicanos and Mexicanos in the Southwest and Chicago. Ramirez (1983, 1997) uses the term mestizo(a) as a world view perspective rather than a racial term for bicultural individuals who are capable of blending distinct cultural influences. More recently, Huang (1994) provides a model for identity formation for Asian Americans. Acknowledging the interplay between internal and external identities, she notes that an important aspect of ethnic identity is the degree to which the individual perceives it to have salience in life and the degree to which the person's external identity is mediated by the congruence or conflict between their acceptance and belonging to an in group and out group. The centrality of a salience hierarchy, and the flexibility posed by separating internal from external identities, provides for a host of profiles of identity which can accommodate the range of ethnic identity outcomes discussed in the literature. Thus, her model moves towards a newer wave of multidimensional models.

Oetting and Beauvais (1990–1991) suggest that the precursors to ethnic or cultural identification models stem from Park's (1931) and his student's, Stonequist's (1964), seminal work on hybridity and marginality, respectively. Similar to the racial identity models, these ethnic or cultural identification models emerge out of attempts to explain problematic behavior in persons who are culturally different to the dominant reference group, deemed American. They provide a summary of models that have prevailed from the linear continuum models to more complex multidimensional models. Finally, they propose the orthogonal cultural identification theory. The dominant majority models suggest movement away from the culture of origin to the new. Furthermore, Anglo culture is good and to be emulated. Simple transitional models suggest that, in the process of movement there is inherent stress that is always difficult and may result in problematic behaviors such as drinking. Another form of transitional model, the alienation model, suggests that in the transition there may be a good or a bad outcome, depending on coping. Good coping results in movement towards the adoption of Anglo ways, whereas failed coping results in alienation. To some degree these models are invoked at various times in the research literature or in everyday life explanations for the success and failure of people who are originally culturally different from the dominant, European-derived culture.

Three multidimensional models are also summarized, including their own. Multidimensional models allow multiple values to co-exist at different points along the continuum from culture of origin to European derived culture. Bicultural models are the first models to suggest that an individual can be simultaneously identified with two cultures. Although they explain and predict that people who are adept and involved in both cultures become more flexible, these models do not generally have explanatory power for the individual who is marginal to both cultures, who demonstrates low involvement or identification with both cultures. Lastly, the orthogonal identification model suggests that identification with one culture is independent of another. Thus, unlike other models, this one allows for any and all combinations of identifications.

It is important with the multidimensional models not to believe that the bicultural models are inferior to the orthogonal models, but to examine which model fits which contexts best. Ramirez (1983) offers a bicultural model born out by research which strongly suggests that biculturality does seem to increase some cognitive flexibility. In educational settings, this model may be advantageous (Banks, 1988). With increasing trends for some populations to live in two countries, the notions of biculturality will continue to be reinforced and be a reality for many populations of Latino and Asian origin.

On a process level, Bennett (1995) suggests that the defenses of denial, minimization, acceptance, adaptation, and integration are employed as people move through a process of interethnic and intercultural valuation related to identity. This model is unique in that process is discussed in terms of developmental strategies with an emphasis on the role of intrapsychic defenses against inferior status.

10.06.4.1 Comparison and Critique of Racial and Ethnic Identity Models

The ethnic identity models have grappled with reference group orientation multiple alliances, and more recently with the transnational identities. Often referred to as a process of ethnogenesis, immigrants reformulate their identity after entering the USA. This reformulation often requires accomodating a new definition of race and being relegated to a minority status. This process is impacted by the way in which bodies are racially assigned meaning. The destructive differentiation based on race in the process of ethnogenesis is illustrated succinctly in a recent study by Mittelberg and Waters (1992) of identity formation in two immigrant groups: middle-class Haitians and secular kibbutz-born Israelis. The authors observe, unsurprisingly, that the Israelis have fewer constraints in reformulating their identities than do Haitians, because of the primacy of the meaning of blackness as a race in this country.

The racial identity models seem the more limited in contemporary context because they are still embedded in an undimensional model of racial belonging, as centuries-old rules remain unquestioned. The linear nature of the models of nigrescence and minority racial identity development assume that an individual will retreat from white society's denigration and refuge in the community of color. However, biracial people, particularly of European heritage, do not have a guaranteed refuge because of their marginalized status by multiple groups. With other people in this stage of development, such as high school and college age peers, they may not be viewed as authentic members due to their biracial status, particularly if they are not willing to denigrate and denounce in an exaggerated way that which is white. To do this is to potentially denounce part of one's heritage and people one loves (Root, 1990). Thus, the biracial person may be subjected to racial or ethnic authenticity tests which are charicatures of socially constructed race and ethnicity. Many biracial persons find this neither a natural nor only process for achieving a positive racial and ethnic identity (Root, 1992a, 1996b). Many young adults enact multidimensional models of race which are more similar to models of ethnicity. Identities can be dynamic fluctuating in salience contextually and be orthogonal (Duffy, 1978; Hall, 1980, 1992; Stephan, 1992; Williams, 1992). Contextually, the salient identity of someone of black and Asian heritage may sometimes be black, at other times Asian. An orthogonal experience of identity would be manifested in the declaration of being both black and Asian.

However, despite some of the strengths of applicability of multidimensional models of ethnicity to mixed heritage individuals, neither the racial nor ethnic identity development models consider other simultaneous influences of other salient statuses integral to identity, such as gender or sexual orientation. Huang's (1994) model can acccomodate this through the identity salience hierarchy and the consideration that other identities interact and inform personal internal identity. Even the diversity of social economic class orientation, often erroneously assumed homogeneous within ethnic and racial groups, may be critical to the construction of identity. For example, Singelis, Triandis, Bhawuk, and Gelfand (1995) observe that individualism and collectivism vary cross-culturally within a culture according to class: wealthy and impoverished classes are much more individualistic than middle classes. Thus, class location may affect value systems as much or more man race; class and race are often confounded.

Stage models of racial identity do not explain the fluidity of exchange of different aspects of identity between background and foreground, such as gender and race, race and class and gender, sexual orientation and race. These models assume to some degree that either gender has negligible effect or—consistent with much research of the time—did not critically assess how gender informs and shapes life experience and worldview. Notably, most of these researchers being male did not have access to the gendered experience of femaleness and its secondary status within their ethnic and racial groups. Also with a more privileged gendered status, they were not as compelled to consider simultaneously the role of female experience as a female researcher might be.

It is also very likely that the models which prevail have been implicity generation specific; it was no coincidence that early racial and ethnic identity models emerged out of the movements of the 1960s and 1970s. However, the contemporary generation of teenagers and adults have inherited many of the benefits of the civil rights movement of the third quarter of this century. Subsequently, they have not experienced the extremes of racial injustice experienced by persons a generation or more before them. There is also a large segment of the Asian American community who is adjusting to the social construction of race in the USA, having moved from racially homogenous cultures to a racially diverse society that stratifies according to race. Many of these immigrants are learning to identify the racism that blocks their opportunities or explains the differential treatment they receive.

When primary identities are without significant ties and markers specific to ethnicity, more fragile identities are constructed through race For example, consider the genre of US movies and trade books that describe someone who is white discovering they have a black ancestor and consider the deep emotional reaction someone has to this revelation. Why should this be such an unravelling experience if one already knows who one is? Such reactions suggest that the primary identity occurs through race and whiteness rather than through ethnicity.

10.06.5 RETHINKING RACE AND ETHNICITY

Both race and ethnicity are socially constructed. Thus, they are dynamic and influenced by many factors including history, gender, and other aspects of identity that are socially salient and specific to the macro- and microsocial compositions of regions or neighborhoods and ethnic group histories. With the legacy of civil rights legislation in the third quarter of this century, many young people grow up with more latitude to cross racial boundaries in friendship and intimacy. Technological advances in transportation and communication have increased the numbers of persons involved in international and cross-cultural marriages. Again, we can expect that the products of these unions will have the options to be multiethnically, multiracially, and even multinationally identified. This forecast necessitates that we rethink the meaning of race and ethnicity and their relationship.

Consider these demographic trends (Root, 1996b). Since the repeal of antimiscegenation laws in 1967, the rates of racial intermarriage have almost doubled with each subsequent decade. The number of babies of mixed racial heritage has boomed since the late 1980s. Also, the number of persons writing in racial identifiers to elaborate on the other choice has caused the US Bureau of the Census enough concern to entertain a multiracial way of identifying for the year 2000 census. In July 1997, the Office of Management and Budget which oversees the Bureau of the Census decided after much research that the year 2000 decennial census will allow people to mark more than one box in response to the race question. This is an historical change in the accounting of population in the USA.

Whereas the folklore sustaining racial boundaries has purported that racially mixed people will be marginally located and accepted in society, this marginality is being explored and challenged in the academic literature (Anzaldua, 1987; Root, 1992, 1996a; Zack, 1993; 1995), literary essays (Funderburg, 1994; Jones, 1994), novels, and mainstream cinema. The social location of an increasing number of multiracially and multiethnically identified people is often positively constructed, in contemporary time and particular on the west coast of the USA.

More than ever, current research on ethnicity and race suggests that we must attempt to disengage these constructs so that, although at times they admittedly remain confounded, they are not synonymous. For example, Stephan's (1992) research in Hawaii and New Mexico points out geographic differences in how race and parents' ethnicity effects the identification of multiracial Asians and multiracial Mexican American young adults. Her conclusion is that race is neither a necessary nor sufficient condition for making an assumption about ethnic identification. She found that when offered an open ended response option to ethnicity in different situations, few respondents identified themselves consistently across all situations. The conclusion was that identity has contextually constructed Hall (1980) found that neither self-assessed phenotypical resemblance to either black or Japanese reference groups or language facility predicted ethnic identity in her black Japanese sample.

The person of mixed heritage experiences many of the psychological assaults and injuries that their generation cohort of people of color experience. Furthermore, many assaults and insults stem from and are specifically directed to their racial ambiguity. These assaults come from groups of people with whom the biracial person shares heritage and identifies. Table 1 provides a summary of 40 racial experiences commonly experienced by many persons whose racial classification is ambiguous or interpreted by others as mixed. Experienced repeatedly throughout a lifetime, these items potentially influence one's understanding of race relations, race rules, and racial identity. Inherent to this inventory is the negative stereotyping of mixed race, boundary policing of authentic racial behavior, physical objectification of phenotype, and denigration of whiteness. Therapists, teachers, and school counselors might use such a list to provoke discussion. Therapists might further use such an inventory to generate discussion about coping, identity formation, and psychological defenses against exclusion, authenticity testing, denigration, and idealizing that the person of mixed heritage may experience.

10.06.6 CONTEMPORARY MODELS FOR RETHINKING THE RELATIONSHIP BETWEEN RACE AND ETHNICITY

Current trends to reformulate racial formation or ethnogenesis should not ignore the conceptual and empirical information gleaned from earlier research. However, current trends necessitate that we move beyond bipolar conceptualizations of race, reexamine assumptions about racial identity formation, and differentiate race and ethnicity. Contextual spheres of influence from large macrodimensions of geographic region to microdimensions of family interaction and personality must be considered in identity formation.

10.06.6.1 Ecological Model of Race and Ethnicity

Figure 1 represents an ecological identity model of influences on identity development. It evolved out of observing the identity process and differential outcomes for persons of multiracial ancestry, often of multiethnic background, in both clinical and nonclinical settings. The model can accommodate several different outcomes of identity for both monoracially and multiracially identified people from very fragile identities to well-grounded and stable identities. It also accomodates some previous work which states that identity is indeed dynamic and can change over a lifetime due to changing contexts and developmental issues over the lifespan (Root, 1990, in press). However, unlike the stage models which may reflect an underlying linear process, this change of identity does not necessarily invoke a linear movement. Social identities are informed by the interaction of micro and macro influences on identity. This model breaks away from the longstanding constraints of bipolar racial frameworks and assumed rules of hypodescent. In line with the symbolic interactionist theories and models, this one proposes reflexive interaction between variables.

In the proposed model, all boxes are lenses and are present to some degree. The drawing serves as a conceptual organizing tool with which to consider influences in identity. The lenses recognize individual factors, familial factors, community influences—both imagined and factual such as historical influence and experience that transcends generations. Through these lenses, the inherited influences, traits, and social interactions with community determine identity development. The inherited influences include biological and environmental inheritance. Biological inheritance, such as sexual orientation and phenotype, significantly influence life. The other inherited influential lenses are environmental: languages spoken at home; parent's identification ethnically, racially, and nationally; nativity; presence and socialization influence of extended family; given names and nicknames; and home values. All of these variables have been documented in different literatures as having influence on identity. All the environmental influences provide cultural markers to a lesser or greater degree of distinctiveness. The lenses with inherited influences interact with one another.

These inherited influences interact with traits: temperament, social skills, talents, and coping skills. These influences may have both elements of nature and nurture. The traits are some of those aspects of the individual that together are often summarized as personality. The traits are differentiated from inherited influences through what the individual has control over. This model assumes the individual has little control over their inherited influences, whether they be biological or environmental. In contrast, the traits, are deemed to be majorly subject to social influence and learning, even when there may be different degrees of natural talents, coping, sociability and its skills, and even temperament.

Table 1 Items from the racial experiences questionnaire.

1. Told, "You have to choose; you can't be both"
2. Your ethnicity misidentified
3. People assume your race to be different by phone and in person
4. Accused of not acting or wanting to be ... (Latino, black, Asian, etc.)
5. Told, "Mixed race people are so beautiful/handsome"
6. Strangers look between you and your parent(s) to figure out if you are related
7. Told, "You don't look ... (Asian, black, native, etc.)"
8. Asked, "What are you?"
9. People might not say certain things in front of you if they knew how you think about race
10. Asked, "Where are you from?"
11. Stared or looked at by strangers
12. Told, "You look exotic"
13. Your choice of friends is interpreted as your selling out or not being authentic
14. Accused of acting or wanting to be white
15. Judgments of your racial authenticity are based upon your partner's race
16. Comments are made about your physical attributes (hair/hairstyle, skin color, eyes, shape, etc.)
17. You are subjected to jokes about people of mixed heritage
18. Told, "You think you're too good for your own kind"
19. Grandparent(s) or relatives don't accept you because of your parents' interracial relationship
20. Your parents/relatives compete to claim you for their own racial or ethnic group
21. Told, "You have the best of both worlds"
22. Asked about your racial heritage
23. Upon meeting you, people seem confused by your name(s)
24. People assume you are confused about your racial identity
25. People speak to you in foreign languages because of how they intepret your physical appearance
26. Told, "Society doesn't recognize mixed race"
27. Told, "you aren't really ... (Asian, black, Hispanic/Latino, native, etc.)"
28. Mistaken for another person of mixed heritage who does not resemble you
29. Told you must be full of self-loathing or hatred because of how you racially identify yourself
30. Told, "You are a mistake"
31. People's racial identification of you varies and is colored by the race of the people you are with
32. The race people assume you are varies in different parts of the USA
33. You have difficulty filling out forms requiring you to check one race only
34. You identify your race differently to how others identify your race
35. Told, "You aren't like other ... (blacks, Latinos, native Americans, Asians, etc.)"
36. Your siblings identify their race differently than you identify yours
37. Called racial slurs of groups who aren't part of your heritage
38. When friends suggest dating partners for you, they stick within a single racial or ethnic group
39. Your parents identify your race differently than you identify yourself
40. Told, "You aren't ... (native, Asian, black, Latino, etc.) enough"

These interactions takes place in contexts. Such interpersonal contexts allow for the communications about rules of belonging whether it be to the outgroup or ingroup. Reflections of who we are outside of our families serves as frame for further reconciling the private and public experience of identity. For most of us, five social contexts encapsulate most of our social interactions: home; school and/or work; community groups; friendships, and communities in which we are strangers. This last social context is important because it may challenge an individual's construction of self when community environments differ from one another or a person's identity is in some way ambiguous. In the case of phenotypically ambiguous people, whether or not they are multiracially or multiethnically identified, a community to which they are a stranger may provide them with feedback that affects how they experience themselves in relationship to others—and it may be different to what they are used to. The most dramatic stories emerge from recent biographical stories of persons raised white who change living situation or state and then, though phenotypically white, must assume the life of a black person (Williams, 1995).

Lastly, the summation of this interactive life experience filtered through the interaction of the different lenses posed by inherited influences, traits, and social contexts helps us understand identity—both ethnic and racial identity in the context of history and gender. These identities, in turn, provide a lens through

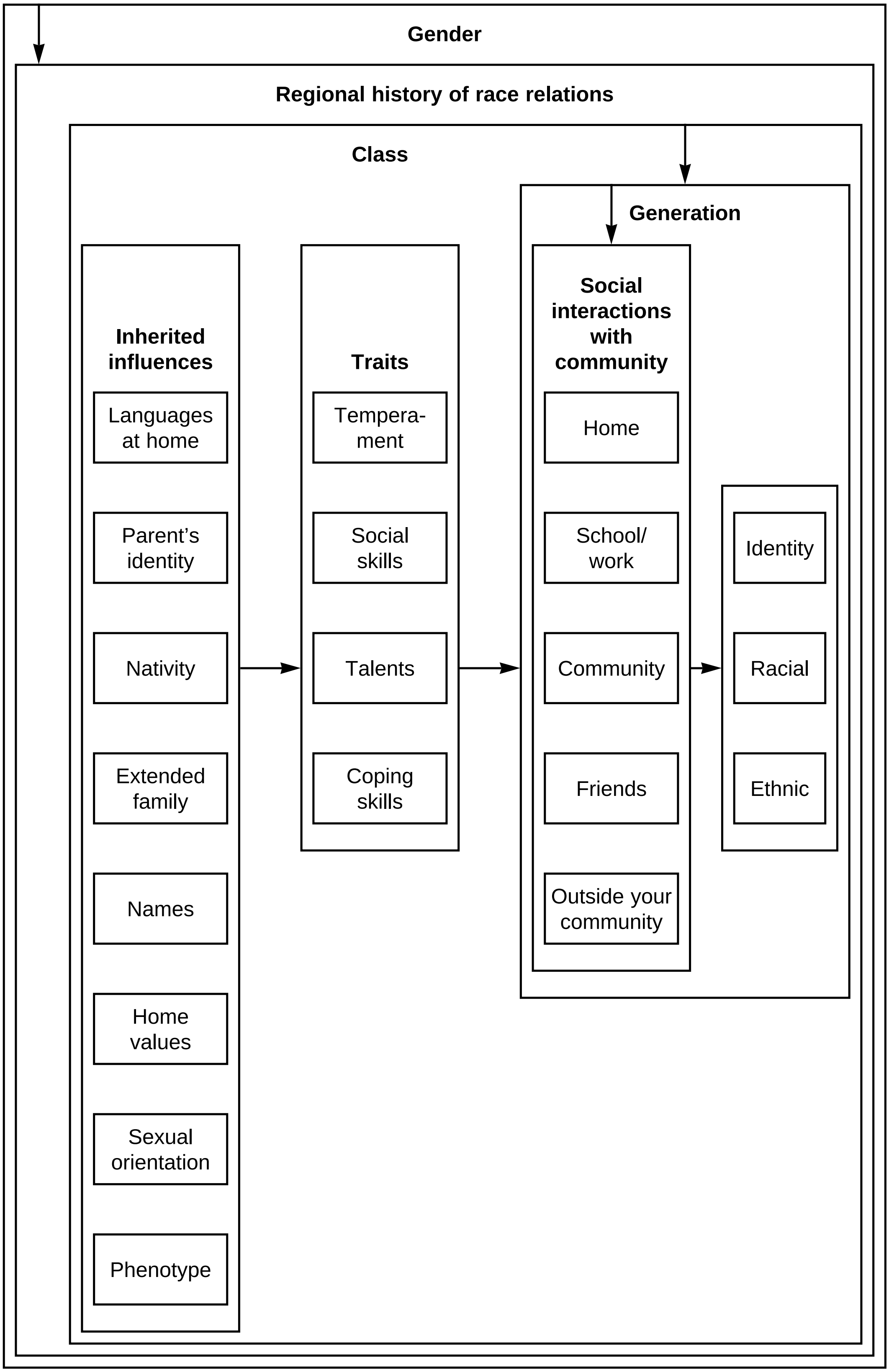
Gender
Regional history of race relations
Class
Generation
Inherited influences
Languages at home
Parent's identity
Nativity
Extended family
Names
Home values
Sexual orientation
Phenotype
Traits
Tempera-ment
Social skills
Talents
Coping skills
Social interactions with community
Home
School/work
Community
Friends
Outside your community
Identity
Racial
Ethnic

Figure 1 Ecological identity model; all boxes are interactive.

which life is experienced. As one moves through milestones in one's life course, racial and ethnic identity can influence what is passed on as inherited influences to the next generation.

This model suggests that the sociohistorical construction of gender and race are lenses through which most of our life experiences are filtered. However, the construction and connotations of race and gender are also informed by geographical regional context and the historical generational cohort. These additional filters allow us to understand how race and gender have been constructed through paradigms of domination and submission (Lerner, 1986; Stoler, 1995; Zack, 1997b). Patriarchal and imperialistically driven cultures construct racial and ethnic hierarchies with imperialist or the colonial schemes (Freire, 1970; Stoler, 1995; Young, 1995). Therefore, in this model gender and ultimately racial identity are viewed as lenses that are layered upon one another. These two lenses interact dynamically, exchanging salience in relationship to one another as foreground or background filters. This conceptual relationship also explains why some persons assert that gender is the primary organizing factor in life experience, whereas others assert that race is the primary organizing factor in life. However, at all times there is an interaction as described above that makes it certain that the social address of men and women differ when we filter their location through the lenses of gender and race.

Other secondary statuses are significant dimensions or lenses through which life is experienced, for example, class and sexual orientation. Both of these locations may be accompanied by denigrating and oppressive or affirming reflections of self. Again, psychological defenses and coping skills are necessary to be able to incorporate these aspects of experience, and ultimately identity into the larger concept of self.

Likewise, racial and ethnic experience in this country, by region, yield different experiences between two individuals despite the similarity of a host of other life variables. For example, the rates of intermarriage between whites and blacks is very different in the south compared to the west. The histories of these regions are also very different. The prominence of lynchings of black men who dared to interact or even look at white women is well documented in the south, particularly prior to this century. Through the lens of race, the socially constructed groups in the USA are Asian and Pacific Islander American, American Indian, African-American, and white. Whereas persons of Hispanic origin may be any race, this country's history of race relations has constituted this category as a race based on attitudes (Root, 1996b). There will be points at which the experiences of each of these groups will overlap. However, if the gender lens is more prominent, an African-American woman may share more in common in a given context with a Chicana than with an African-American man. For example, the threat to one's body through sexual violence is an experience shared by women and much less so by men. The lense of historical race relations suggest that women of color have been less protected by the legal system in matters of violations of their body, whether it be child abuse, domestic violence, or rape.

This model can accommodate the process described in racial identity development models described earlier in this chapter. However, it does not predict or require that an individual will go through the first stage of negative feelings about their minority racial or ethnic group; it can accomodate such a process, though. This is important because in contemporary times, perhaps due to some of the benefits of the civil right movement, many young people do not feel badly about their ethnic or racial reference group which society ascribes to them.

The ecological identity model also accommodates typology models. It evolved from an earlier work (Root, 1990) in which types of identity were proposed for multiracial people. The process by which people might get to those identities is subsumed in this model. Thus, it is possible to change identity over time, and sometimes this move is simply precipitated by a geographic move, such as from mainland USA to Hawaii where a multiracial and multiethnic identity is the norm. It also provides an increased number of variables with explanatory power for differences in identity absent in other racial identity models.

10.06.7 CLINICAL IMPLICATIONS

The ecological identity model may be useful in clinical settings. It acknowledges trauma which in fact determines some of the salient lenses through which one experiences life (Root, 1992). If a patient has been sexually violated because she is female, and links her suffering to this status, the lens of gender will be prominent. If life chances have been compromised by skin color, and this link has been made, the lens of race will operate prominently. But both lenses will interact. Therapists and counselors can understand the possibility and impact of transgenerational trauma as part of inherited influence: stories of genocide, slavery, incarceration, colonization, and domination imparted from an early age are part of a parent's or relative's identity. The stories may be passed on in the inherited influences of names and family interactions within the community. Parenting is extremely important in the passing on of both coping with a secondary status and the formation and interpretation of both racial and ethnic identity.

Kich (1992) provides one of the best clinical explanations of the processes and outcomes that the biracial person may experience through his three stages. One of his most significant contributions is that his model provides for an understanding of how internal family dynamics may interact with lessons around race and ethnicity. Given that society stratifies race, this hierarchy may be replicated in family dynamics. A child may develop a distance and ambivalent relationship with the parent of color if emotional denigration or the process of destructive differentiation occurs around cultural differences.

The model has implications for different developmental periods that can be further elucidated by the ecological framework suggested in this chapter. For example, if the parent with higher racial or ethnic status denigrates the other parent, a child may feel simultaneously victimized or denigrated. Simultaneously, if this child feels helpless to help, she or he may develop an acute awareness of self which is akin to the vigilance typical of post-traumatic responding. Several aspects of his or her existence may serve as subsequent environmental cues which trigger vigilance in certain settings. For those children who have been raised in psychologically or emotionally denigrating households, the statements, questions, and experiences in Table 1 may all serve as cues leading to experiences of marginalization and rejection.

Therapists and clinicians may want to explore how these experiences have had impact on a person at different stages in their life, and how their coping deflects or defends against society's unresolved anger stemming from the consequences of the racial system. In this system anger and anxiety is projected onto the body of the mixed-race person and subsequently creates a gestalt of experience unique to the mixed-race individual's experience. Major themes are rejection, marginalization, societal anxiety, and objectification as represented by the items of Table 1. These experiences may be specific to certain social environments which feed into definition of self through acceptance and rejection experiences which also seem critical to ethnic identity of racially mixed people (Hall, 1992; Root, 1990; Stephan, 1992). Each of these themes is explained below.

Rejection is captured by items such as, "You are a mistake," "Grandparents or relatives don't accept you because of your parents' interracial relationship," and "Told society doesn't recognize mixed race." It also encompasses the accusations of passing into a dominant desired group with statements such as "You think you're to good for your own kind," "Ethnicity misidentified," or "Accused of acting white." Ironically, because of the past history of racially mixed people passing into a desired group group because their phenotype allowed this, many racially mixed persons are struggling for acceptance from an ethnic group or racial group of their heritage because their phenotype creates suspicion about their intent to belong and their trustworthiness.

Marginalization relates to rejection through authenticity tests and applications of rigid boundaries. Cues for this experience may occur with statements such as, "You have to choose, you can't be both" or "Told you must be full of self-loathing or hatred because of how you racially identify" or "Your racial authenticity is based upon your partner's race."

Much societal anxiety about race is projected onto the mixed race person who refuses to comply with the implicit rules of race which guide daily interactions. This anxiety manifests in questions and statements such as, "Asked about your racial heritage," "Asked what are you?" and "Accused of not wanting to be Latino/black/Asian, etc."

Objectification encompasses the deindividualization and sexualization projected onto race and gender and their interaction. Such a process sets the stage for cruel treatment and relegating persons of mixed heritage to otherwise invisibility. Experiences such as being "told you look exotic" and "stared or looked at by strangers" signal this process.

In increasing attempts at individuation, particularly during adolescence in western cultures, anger at the parent with perceived lower status may be a dysfunctional attempt to attempt to avoid taking on the denigrated status. Also of this racial group, the adolescent is caught in a cycle of anger or even hate that precludes them from realistic affirmation of their whole self. Whereas adolescence in the USA is replete with anger, racial dynamics and gender dynamics may become entwined in a way that makes it very difficult for a child to cope and determine psychologically adaptive ways of responding that may serve them into adulthood. For example, if a white father repeatedly denigrates an Asian mother, not only does this set up the possible dynamics above, it may perpetuate misogyny and denigration of female status, particularly women of color, and set up an idealization of white women.

The status of parents' relationship to one another and society may also be formative in identity. Stephan (1992) discusses the possible influence of the gender of parents in patriarchally organized societies as it may differentially influence identity formation in Hawaii versus New Mexico. Results from her study suggest that for persons of mixed heritage, having an Hispanic father in New Mexico may be more of a determining factor in New Mexico, where such status is lower than having a Japanese American father in Hawaii, where Japanese Americans have relatively high social status. Kich (1982, 1992) further suggests that the absence of a parent and how that absence occurs and is explained may be very critical to how race is interpreted in the family. More recently, with high rates of divorce, increasing attention is paid to white single parenting of mixed heritage children so that they can defend themselves in a society that has been antimiscegenistic and anxious about their existence. Similarly, children of single parents of color must also be aware of the intense discrimination towards racially mixed people that may occur in some closed communities of color. McRoy and Hall's (1996) clinical work on transracial adoption is applicable here. They essentially point out that basic needs of a child must be met even before the tasks of racial defense are relevant, that is, a child must feel loved and secure.

Root (1994) identifies the issues in therapy for mixed race women, although most of the issues are applicable to men. Whereas few people enter therapy to work on their mixed racial heritage, the experiences society has dealt around this status can manifest in issues around self-definition, acceptance and belonging, interpretations of power, sources of self-esteem, and independence.

Gibbs and Hines (1992) suggest that adolescent biracial identity resolution revolves around several developmental questions: Who am I? Where do I fit? What is my social role? Who is in charge of my life? Where am I going? These questions are windows into several areas of developmental conflicts including, racial/ethnic identity, social marginality, sexuality, autonomy and independence, and education and career aspirations. Bradshaw (1992) cautions therapists to take the time to understand the contemporary multiracial experience to contextualize the acute and at times everpresent self-consciousness present in people's lives. Lastly, Root (1994) cautions therapists not to assume that being of mixed heritage is an unresolved issue for persons coming into therapy.

10.06.8 FUTURE DIRECTIONS: RETHINKING RACE AND ETHNICITY

There are several factors driving the need for a reexamination of how we think about race and ethnicity. Unless we rethink race, the process of differentiation between ourselves and those who look different from us will often times be a destructive process of differentiation driven by unresolved conflicts and fraught with stigmatizing of bodies we have inherited from this country's history of race relations. Unless we reopen the dialogue—and a new one that considers a contemporary meaning for mixed race—we may not be able to understand contemporary ethnicity in America. Without separating the concepts of ethnicity and race, dialogues will remain confused as discomfort with the construct of race, can be evaded by changing the dialogue to one that shifts emphasis to the other construct of ethnicity, despite the negative differentiating driving the construct.

Likely the racial identity models of the 1970s and 1980s were necessary to understand and be able to educate people oppressed by the racial system in a way that promoted healing. We are now challenged to see if we can undertake a next step in healing. Can we develop models that allows us to explore the fluidity of identity in the social contexts of current and past historical moments?

10.06.9 SUMMARY

A biracial baby boom became possible with the repeal of the last antimiscegenation laws in the USA in 1967, changes in immigration laws, and global mobility of people and populations. Changes in the psychology of race and ethnicity in this last part of the century are integrally linked to these demographic changes. Nevertheless, the 10 conditions Allport (1954) hypothesized correlate with prejudicial attitudes still exist. A more critical understanding of how racial politics derive from a process of negative rather than constructive differentiation is highlighted by examining the current experience, identity process, and social location of the person of mixed racial heritage. An ecological framework for thinking about the different outcomes of racial identity is presented. Gender, geographical history, class, and generation are significant, dynamic lenses which filter many life experiences which have not been put forth in previous racial identity models. The ecological identity model can accomodate the processes proposed by past identity models. However, it accomodates increasingly more common identity development processes that do not involve traumagenic experiences despite prevailing conditions towards prejudicial attitudes. Psychological processes in resolution and integration of a multiracial location and identity are discussed in this chapter with cautions that therapists should not assume that a multiracial identity is inherently problematic. For many young people, the work of the generation before has allowed them to open a different dialogue on race. This is a challenge to society and to therapists trying to understand the process by which these people negotiate their daily lives.

10.06.10 REFERENCES

Alarcon, R. D., & Foulks, E. (1995). Personality disorders and culture: Contemporary clinical views. *Cultural Diversity and Mental Health, 1,* 3–17.

Allport, G. W. (1954). *The nature of prejudice.* Reading, MA: Addison-Wesley.

Anzaldua, G. (1987). *Borderlands/La Frontera: The New Mestiza.* San Francisco: Spinsters/Aunt Lute.

Atkinson, D. R., Morten, G., & Sue, D. W. (Eds.) (1979). *Counseling American minorities: A cross-cultural perspective.* Dubuque, IA: William C. Brown.

Banks, J. A. (1988). The stages of ethnicity: Implications for curriculum reform: In J. A. Banks (Ed.), *Multi-ethnic education: Theory and practice* (pp. 129–139). Boston: Allyn & Bacon.

Bennett, M. J. (1995). A developmental model of intercultural sensitivity. In R. M. Paige (Ed.), *Education for the intercultural experience* (pp. 29–71). Washington, DC: Intercultural Press.

Berger, J., Cohen, B. P., & Zelditch, M. J. (1966). Status characteristics and expectation states. In J. Berger, M. Zelditch, & B. Anderson (Eds.), *Sociological theories in progress 1.* Boston: Houghton Mifflin.

Bradshaw, C. K. (1992). Beauty and the beast: On racial ambiguity. In. M. P. P. Root (Ed.), *Racially mixed people in America* (pp. 77–90). Thousand Oaks, CA: Sage.

Burman, E. (1994). *Deconstructing developmental psychology.* New York: Routledge.

Carney, C. G., & Kahn, K. B. (1984). Building competencies for effective cross-cultural counseling: A developmental view. *The Counseling Psychologist, 12*(1), 111–119.

Carter, R. T. (1995). *The influence of race and racial identity in psychotherapy: Toward a racially inclusive model.* New York: Wiley.

Cross, W. E., Jr. (1971). The Negro-to-Black conversion experience: Toward a psychology of Black liberation. *Black World, 20* (9), 13–27.

Cross, W. E., Jr. (1978). Models of psychological nigrescence: A literature review. *Journal of Black Psychology, 5*(1), 13–31.

Cross, W. E., Jr. (1985). Black identity: Rediscovering the distinction between personal identity and reference group orientation. In M. Spencer, G. Brookins, & W. Allen (Eds.), *Beginnings: The social and affective development of Black children* (pp. 155–171). Hillsdale, NJ: Erlbaum.

Cross, W. E., Jr. (1991). *Shades of black: Diversity in African–American identity.* Philadelphia: Temple University Press.

Dohrenwend, B. P., & Dohrenwend, B. S. (1974). Social and cultural influences on psychopathology. *Annual Review of Psychology, 25,* 417–452.

DuBois, W. E. B. (1903, 1989). *The souls of Black folks.* New York: Bantam.

Duffy, L. K. (1978). *The interracial individuals: Self-concept, parental interaction, and ethnic identity.* Unpublished master's thesis, University of Hawaii, Honolulu.

Durkheim, E. (1966, c1951). *Suicide: A study in sociology.* (Translated by John A. Spaulding and George Simpson.) New York: Free Press.

Erikson, E. (1968). Race and the wider identity. In E. H. Erikson, *Identity, youth and crisis.* New York: Norton.

Fanon, F. (1967). *Black skin: White masks.* New York: Grove Press.

Foucault, M. (1978). *The history of sexuality: An introduction.* New York: Random House.

Frankenberg, R. (1993). *The social construction of whiteness: White women, race matters.* Minneapolis, MN: University of Minnesota Press.

Friere, P. (1970). *Pedagogy of the oppressed.* New York: Seabury.

Funderburg, L. (1994). *Black, white, and other: Biracial Americans talk about race and identity.* New York: William Morrow.

Ganter, G. (1977). The socio-conditions of the White practitioner: New perspectives. *Journal of Contemporary Psychotherapy, 9*(1), 26–32.

Gay, G. (1984). Implications of selected models of ethnic identity development for educators. *The Journal of Negro Education, 54*(1), 43–52.

Gibbs, J. T., & Hines, A. M. (1992). Negotiating ethnic identity: Issues for black–white biracial adolescents. In M. P. P. Root (Ed.), *Racially mixed people in America* (pp. 223–238). Thousand Oaks, CA: Sage.

Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity.* Englewood Cliffs, NJ: Prentice-Hall.

Gordon, L. R. (1995). Critical "Mixed Race"? *Social Identities, 1*(2), 381–395.

Gordon. L. R. (1997). Race, sex, and matrices of desire in an antiblack world: An essay in phenomenology and social role. In N. Zack (Ed.), *Race/sex: Their sameness, difference and interplay* (pp. 117–132). New York: Routledge.

Griffin, S. (1992). *A chorus of stones: The private life of war.* New York: Doubleday.

Gurin, P., Hurtado, A., & Peng, T. (1994). Group contacts and ethnicity in the social identities of Mexicanos and Chicanos. *Personality and Social Psychology Bulletin, 20*(5), 521–532.

Hall, C. C. I. (1980). *The ethnic of racially mixed people: A study of Black-Japanese.* Unpublished doctoral dissertation, University of California, Los Angeles.

Hall, C. C. I. (1992). Please choose one: Ethnic identity choices for biracial individuals. In M. P. P. Root (Ed.), *Racially mixed people in America* (pp. 250–264). Thousand Oaks, CA: Sage.

Helms, J. E. (1984). Toward a theoretical explanation of the effects of race on counseling: A black and white model. *The Counseling Psychologist, 12*(4), 153–165.

Helms, J. E. (1990). An overview of Black racial identity theory. In J. E. Helms (Ed.), *Black and White racial identity: Theory, research, and practice* (pp. 9–32). New York: Greenwood Press.

Helms, J. E., & Carter, R. T. (1990). Development of the White Racial Identity Inventory. In J. E. Helms (Ed.), *Black and White racial identity: Theory, research, and practice* (pp. 67–80). New York: Greenwood Press.

Hilliard, A. G., III (1996). Either a paradigm shift or no mental measurement: The nonscience and the nonsense of the Bell curve. *Cultural Diversity and Mental Health, 2,* 1–20.

Huang, L. N. (1994). An integrative view of identity formation: A model for Asian Americans. In E. P. Salett & D. R. Koslow (Eds.), *Race, ethnicity and self* (pp. 42–59). Washington, DC: National Multicultural Institute.

Hurtado, A., & Gurin, P. (1995). Ethnic identity and bilingualism attitudes. In A. M. Padilla (Ed.), *Hispanic psychology: Critical issues in theory and research* (pp. 89–103). Thousand Oaks, CA: Sage.

Jackson, B. (1975) Black identity development. In L. Golubchick & B. Persky (Eds.), *Urban social and educational issues* (pp. 158–164). Dubuque, IA: Kendall-Hall.

Janoff-Bulman, R. (1992). *Shattered assumptions: towards a new psychology of trauma*. New York: The Free Press.

Jones L. (1994). *Bulletproof diva: Tales of race, sex and hair*. New York: Praeger.

Kich, G. K. (1982). *Eurasians: Ethnic/racial identity development of biracial Japanese/white adults*. Unpublished doctoral dissertation, Wright Institute Graduate School of Psychology, Berkeley, CA.

Kich, G. K. (1992). The developmental process of asserting a biracial, bicultural identity. In M. P. P. Root (Ed.), *Racially mixed people in America* (pp. 304–320). Thousand Oaks, CA: Sage.

Lal, B. B. (1995). Symbolic Interaction Theories. *American Behavioral Scientist, 38*(3), 421–441.

Lerner, G. (1986). *The creation of patriarchy*. New York: Oxford University Press.

McRoy, R. G., & Hall, C. C. I. (1996). Transracial adoptions: In whose best interest? In M. P. P. Root (Ed.), *The multiracial experience: Racial borders as the new frontier* (pp. 63–78). Thousand Oaks, CA: Sage.

Milliones, J. (1980). Construction of a Black consciousness measure: Psychotherapeutic implications. *Psychotherapy: Theory, Research, and Practice, 17*(2), 175–182.

Mittelberg, D., & Waters, M. C. (1992). The process of ethnogenesis among Haitian and Israeli immigrants in the United States. *Ethnic and Racial Studies, 15*(3), 412–435.

Oetting, E. R., & Beauvais, F. (1990–91). Orthogonal cultural identification theory: The cultural identification of minority adolescents. *The International Journal of Addictions, 25* (5A & 6A), 655–685.

Olmedo, E. L. (1979). Acculturation: A psychometric perspective. *American Psychologist, 34,* 1061–1070.

Omi, M., & Winant, H. (1986). *Racial formation in the United States from the 1960s to the 1980s*. New York: Routledge & Kegan Paul.

Padilla, A. M. (Ed.) (1980). *Acculturation: Theory, models and some new findings*. Boulder, CO: Westview.

Parham, T. (1989). Cycles of nigrescence. *The Counseling Psychologist, 17,* 187–226.

Park, R. E. (1931). Mentality of racial hybrids. *American Journal of Sociology, 36,* 534–551.

Phinney, J. (1989). Stages of ethnic identity development in minority group adolescents. *Journal of Early Adolescence, 9,* 34–39.

Phinney, J. (1990). Ethnic identity in adolescents and adults: A review of research. *Psychological Bulletin, 108,* 499–514.

Ramirez, M., III (1983). *Psychology of the Americas: Mestizo perspectives on personality and mental health*. New York: Pergamon.

Ramirez, M., III (1997). *Multicultural/multiracial psychology: Mestizo perspectives in personality and mental health*. Northvale, NJ: Jason Aronson.

Root, M. P. P. (1990). Resolving "Other" status: Identity development of biracial individuals. In L. S. Brown & M. P. P. Root (Eds.), *Diversity and complexity in feminist therapy*. New York: Harrington Park Press.

Root, M. P. P. (Ed.) (1992). *Racially mixed people in America*. Thousand Oaks, CA: Sage.

Root, M. P. P. (1994). Mixed race women. In L. Comas Diaz & B. Greene (Eds.), *Women of color and mental health: The healing tapestry*. New York: Guilford Press.

Root, M. P. P. (Ed.) (1996a). *The multiracial experience: Racial borders as the new frontier*. Thousand Oaks, CA: Sage.

Root, M. P. P. (1996b). The multiracial experience: Racial borders as a significant frontier in race relations. In M. P. P. Root (Ed.), *The multiracial experience: Racial borders as the new frontier*. Thousand Oaks, CA: Sage.

Root, M. P. P. (in press). The biracial baby boom: Understanding ecological constructions of racial identity in the twenty-first century. In E. Hollins & R. H. Sheets (Eds.), *Racial, ethnic, cultural human development: Implications for education*. New York: Erlbaum.

Sartre, J.-P. (1976) translation by A. Sheridan-Smith). *Critique of dialectical reasoning: Theory of practical ensembles*. London: New Left Books.

Singelis, T. M., Triandis, H. C., Bhawuk, D. P. S., & Gelfand, M. J. (1995). Horizontal and vertical dimensions of individualism and collectivism: A theoretical and measurement refinement. *Cross-cultural Research, 29*(3), 240–275.

Spickard, P. R. (1989) *Mixed blood: Intermarriage and ethnic identity in twentieth-century America*. Madison, WI: University of Wisconsin Press.

Spickard, P. R. (1992). The illogic of American racial categories. In M. P. P. Root (Ed.), *Ratially mixed people in America* (pp. 12–23). Thousand Oaks, CA: Sage.

Stephan, C. W. (1992) Mixed-heritage individuals: Ethnic identity and trait characteristics. In M. P. P. Root (Ed.), *Racially mixed people in America* (pp. 50–63). Thousand Oaks, CA: Sage.

Stoler, A. (1995). *Race and the education of desire: Foucault's history of sexuality and the colonial order of things*. Durham, NC: Duke University Press.

Stonequist, E. V. (1964). The marginal man: A study in personality and culture conflict. In E. Burgess & D. J. Bogue (Eds.), *Contributions to urban sociology*. Chicago: University of Chicago Press.

Tajfel, H. (1981). *Human groups and social categories*. Cambridge, UK: Cambridge University Press.

Thomas, C. (1971). *Boys no more*. Beverly Hills, CA: Glencoe Press.

Trimble, J. E. (1996). Accumulturation, ethnic identification, and the evaluation proces. In A. H. Buyer, F. L. Brisbane, & A. Romirez (Eds.), *Advanced methodological issues in culturally competent evaluation for substance abuse prevention*. Center for Substance Abuse Prevention, Cultural Competence Series 6, Department of Human and Health Services, Publication (SMA) 96-3110

Trimble, J. E. (in press). Social psychological perspectives on changing self-identification among American Indian and Alaska natives. In R. H. Dana (Ed.), *Handbook of cross-cultural/multi-cultural personality assessment*. Mahweh, NJ: Erlbaum.

Vontress, C. E. (1970). Counseling Blacks. *Personnel and Guidance Journal, 48,* 713–719.

Waters, M. S. (1990). *Ethnic options: Choosing identities in America*. Berkeley, CA: University of California Press.

Weinreich, P. (1986). The operationalisation of identity theory in racial and ethnic relations. In J. Rex & D. Mason (Eds.), *Theories of race and ethnic relations* (pp. 299–320). Cambridge, UK: Cambridge University Press.

Williams, G. H. (1995). *Life on the color line: The true story of white boy who discovered he was black*. New York: Dutton.

Williams, T. K. (1992). Prism Lives: Identity of Binational Amerasians. In M. P. P. Root (Ed.), *Racially mixed people in America* (pp. 280–303). Thousand Oaks, CA: Sage.

Wilson, W. J. (1987). *The truly disadvantaged: The inner city, the underclass, and public policy*. Chicago: University of Chicago Press.

Young, R. (1995). *Colonial desire: Hybridity in theory, culture and race*. New York: Routledge.

Zack, N. (1993). *Race and mixed race*. Philadelphia: Temple University Press.

Zack, N. (Ed.) (1995). *American mixed race: The culture of microdiversity*. Lanham, MD: Rowman & Littlefield.

Zack, N. (Ed.) (1997a). *Race/Sex: Their sameness, their difference*. New York: Routledge.

Zack, N. (1997b). The American sexualization of race. In N. Zack (Ed.), *Race/Sex: Their sameness, difference, and interplay* (pp. 145–156). New York: Routledge.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.07
The Psychology of Gender and Health

RICHARD M. EISLER
Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

10.07.1 INTRODUCTION 161
10.07.2 GENDER AND EPIDEMIOLOGY OF HEALTH PROBLEMS 162
10.07.3 HISTORICAL PERSPECTIVE ON GENDER DIFFERENCES 163
10.07.4 THEORETICAL OVERVIEW OF GENDER DIFFERENCES 164
  10.07.4.1 Biological Perspective 164
  10.07.4.2 Social Psychological Perspective 165
  10.07.4.3 Cross-cultural Perspective 165
  10.07.4.4 Summary of Gender and Health Perspectives 166
10.07.5 THE ROLE OF PSYCHOLOGICAL STRESS IN WOMENS' AND MENS' HEALTH 166
  10.07.5.1 Gender and Biological Vulnerability to Stress 166
  10.07.5.2 Gender and Cognitive Vulnerability to Stress 167
  10.07.5.3 Gender Role Stress 167
  10.07.5.4 Gender Roles and Coping with Stress 168
10.07.6 GENDER DIFFERENCES IN SELF-DISCLOSURE AND HELP SEEKING 168
  10.07.6.1 Gender and Emotional Expression 169
  10.07.6.2 Gender and Social Support 169
  10.07.6.3 Gender and Help Seeking 170
10.07.7 FUTURE DIRECTIONS 170
10.07.8 SUMMARY 171
10.07.9 REFERENCES 171

10.07.1 INTRODUCTION

Why is it important for a discourse on the psychology of health to focus on gender issues? Most importantly, there has been a shift in the conception of healthcare practices over the past few decades from the biomedical model to the biopsychosocial model. The biomedical model had limited the assessment and treatment of disease to the biological components of organs, biochemicals, and infectious agents. In this model biochemical problems caused by faulty organs or infectious agents resulted in illness. The biomedical model served us well 100 years ago, when the leading causes of death were tuberculosis, pneumonia, and influenza. The leading causes of death today, however, such as cardiovascular disease and lung cancer, have psychosocial etiologies that may be more comprehensively evaluated and treated by understanding cultural risk factors, including gender determinants.

In the current biopsychosocial model of health, biological disease processes are acknowledged, but psychosocial factors that predispose one to illness are strongly emphasized. The psychosocial components encompass engaging in high-risk behaviors, appraising and coping with stress, utilizing social support systems, and the seeking of appropriate help from the healthcare system. In all the aforementioned psychosocial areas, recent research has shown that there are important sex or gender distinctions, the evaluation of which is crucial for the prevention of illness and the maintenance of health. These psychosocial issues are the focus of this chapter.

It has been evident that many health problems occur with different frequencies among men and women. Paradoxically, while women are more likely than men to have most of the nonfatal chronic illnesses and acute medical conditions, more men than women die of each of the 12 leading causes of death (Wingard, 1984). There are also important differences in the types of stress to which each gender is subjected. For example, research has shown that working women have greater gender role stresses than working men, in that they have to manage both family and work roles (Frankenhaeuser, Lundberg, & Chesney, 1991). Further, it appears that marriage itself contributes differentially to the health of women as contrasted to men. Married men report lower rates of mental and physical health problems than single men. Married women, on the other hand, report higher rates of mental health problems than single women. Finally, women are far more vulnerable to sexual and physical abuse than men in their intimate relationships (Koss et al., 1994).

Another gender-related health issue is the frequency and quality of each sex's interaction with the professional healthcare system. Across most age categories men have a lower frequency of physician office visits than women, and less than a third as many visits to a mental health professional. Pennebaker (1982) theorized that underutilization of the healthcare system by men compared to women may be the result of men's tendency to disregard the importance of their symptoms. Another possibility could be that men are reluctant to visit caregivers because it conflicts with masculine gender role imperatives to appear strong and tolerate pain (Eisler & Blalock, 1991). Not only do women and men utilize healthcare systems differently, but healthcare systems may also respond differently to each gender. If similar kinds of symptoms are reported by each sex, women are more likely than men to be diagnosed with a mental or nervous condition (Travis, 1988). In addition, physicians may be less likely to diagnose women with diseases that they consider to be most prevalent in men, such as heart disease.

Thus, we may conclude that there are very significant reasons for including gender in studying the psychology of health. More than half of the mortality from the 10 leading causes of death may be traced to lifestyle risk factors on which women and men differ significantly (Verbrugge, 1989). Divergent gender roles produce large variations in exposure to and coping with stress, which result in different vulnerabilities to various disorders. Also, men and women have different values regarding the importance of social support and preventive healthcare practices. Finally, existing research protocols have primarily been designed to assess the causes of health problems in white males. It is clear that this is inadequate for understanding treatment issues for women and other cultural groups.

It is instructive to note here that the current convention is to use the term "sex" when referring to reliable biological distinctions between males and females. Sex refers to biological characteristics including chromosome pattern, and genital structure. However, "gender" is the term preferred by psychologists to denote the idea that men and women are socialized to display "masculine" and "feminine" characteristics and behavior patterns. The term sex will be used henceforth to denote the biological aspects of individuals, and gender will be used to refer to the socially constructed roles deemed socially appropriate for either men or women.

10.07.2 GENDER AND EPIDEMIOLOGY OF HEALTH PROBLEMS

Recent epidemiological data on the association between gender and health have documented gender differences in "premature mortality" based on more hazardous lifestyles and behavior patterns engaged in by men compared to women (Harrison, Chin, & Ficarrotto, 1989). In the USA, women live, on average, about seven years longer than men. The death rate for men is higher than for women at all ages and for all leading causes of death (Verbrugge, 1985). Men have twice the rate of premature death from coronary artery disease, and they are three times more likely to die in motor vehicle accidents than are women. In addition, Waldron and Johnson (1976) have noted that mens' death rates from lung cancer have been nearly six times that of women, and twice as high from cirrhosis of the liver.

In terms of psychological dysfunction, men are much more prone to a higher incidence of personality and behavior disorders. In a review of epidemiological studies, Cleary (1987) noted that men have been three times more likely to be problem drinkers and four times as likely to be diagnosed with alcoholism than are women. It was also evident in this review that men were more likely than women to abuse drugs and to be diagnosed with antisocial personality disorder. In addition, a review by Widom (1984) cited evidence that men, far more than women, are involved in violent crime, including armed robbery and homicide. Widom also noted that men are much more likely than women to be involved in sexual deviance and more frequently are perpetrators of child and spouse abuse. Taken together, these statistics suggest that men more than women tend to employ antisocial and aggressive means in dealing with conflicts and personal problems. This is consistent with the evidence reported by Huselid and Cooper (1994) that men are more inclined than women to "externalize" their distress through aggressive and hazardous coping strategies.

On the other hand, there is convincing evidence that women may suffer from more "internalized" psychological distress than men (Huselid & Cooper, 1994). Women's stereotypic gender roles associated with passivity and self-blame may be responsible for their greater vulnerability to internalizing disorders such as anxiety and depression. In an early study, Weissman and Klerman (1977) reviewed 40 studies in over 30 countries before concluding that women were two to three times more likely than men to suffer from depression. The greater incidence of depression in women has been confirmed in a large-scale community study of the mental health of over 20 000 residents in different US cities (Robins et al., 1984). With respect to the anxiety disorders, Robins et al. showed that women had higher prevalence rates of agoraphobia, panic disorder, and obsessive-compulsive disorders than men. In an extensive review, Attie and Brooks-Gunn (1987) described weight concerns and chronic dieting as the focus of constant stress in women as being directly responsible for the extraordinarily high ratio of eating disorders (anorexia and bulimia) in women compared to men.

From the above discussion, we may conclude that there have been significant gender differences in incidence of different disorders, including psychological dysfunction. It is also apparent that gender roles mediate male vs. female differences in expressions of psychopathology in response to stress. In the remaining sections of this chapter, we will be concerned with the historical and theoretical studies of gender differences, gender roles and one's vulnerability to stress, and how each gender handles and expresses emotional distress. In addition, we discuss how each gender seeks help for their psychological problems through informal support systems and professional counseling. Finally, we explore how these perspectives may influence the future directions of the investigations concerned with gender and health.

10.07.3 HISTORICAL PERSPECTIVE ON GENDER DIFFERENCES

While philosophers and social scientists have deliberated for ages about the essential qualities of women compared to men, the first systematic attempt by psychologists to synthesize a large body of research data about the psychological differences between the sexes may be traced to Maccoby and Jacklin's classic work *The Psychology of sex differences* published in 1974. Maccoby and Jacklin's volume reviewed over 1400 of the available psychological studies which conducted statistical comparisons of the sexes on their intellectual abilities and social behavior. Their purpose, as stated in the introduction, was to "sift the evidence" to determine which of the many (stereotypic) beliefs about presumed differences between the sexes had a solid basis in fact and which did not. Overall, they concluded from the data that the sexes differed much less than commonly held cultural stereotypes had indicated. Based on the evidence, the authors found sex differences in a few intellectual abilities: verbal, quantitative, and spatial skills. However, Maccoby and Jacklin argued that the evidence was negative or inconclusive with regard to differences in social behavior, with the notable exception that boys tended to display more aggression than girls. Finally, they drew attention to the fact that differences among individuals within each gender were often as great as the average differences between genders. Once again, the point was made that the ability of the individual, not the individual's gender, was most important.

With respect to the potential sources of verified sex differences, such as aggression, Maccoby and Jacklin acknowledged that biological predispositions based on sex could interact with psychosocial experiences to determine the psychological and social characteristics of the person. However, the authors indicated that the biological bases for these psychological characteristics of women and men could not be specified with any degree of certainty.

Interestingly, and probably not coincidently, the publication of Maccoby and Jacklin's (1974) volume coincided with the early phases of the feminist movement in the USA. At that time feminist theory had a political as well as a scientific stake in gender research (Eagly, 1995). Two of the major feminist political agendas were to discover scientific data to show that: (i) women were intellectually just as capable as men, and (ii) gender stereotypes that depicted women socially or emotionally at a disadvantage compared to men were false. This agenda required an interpretation of the data that showed there were no meaningful differences between the basic abilities of men and women, which would then give women equal opportunity with men to succeed in society. While feminist ideology may have biased gender research of the 1970s toward the conclusion that there was little evidence of sex differences, research in the mid-1980s and 1990s returned to more refined theoretical questions about the origin and health implications of gender similarities and differences.

A thorough critique of Maccoby and Jacklin's (1974) synthesis regarding the evidence of significant gender differences in intellectual and social abilities is beyond the scope of this chapter. Suffice it to say that subsequent to their review, more sophisticated quantitative analyses comparing gender differences in a variety of skills and social behavior have followed. These have included meta-analytic studies of gender differences in decoding nonverbal cues (Hall, 1978), conformity and influenceability (Becker, 1986), and helping and aggressive behaviors (Eagly & Crowley, 1986), to name just a few. While the significance of these newer findings is hotly debated, the conclusions reached by more recent work have been that there are in fact functional gender distinctions in such areas as empathy (Eisenberg & Lennon, 1983), leadership (Eagly & Karau, 1992), and the incidence of psychological disorders such as depression (Nolen-Hocksema, 1987). In addition, nontrivial gender differences in important social behaviors have been reviewed and discussed by gender researchers (Daly & Wilson, 1988; Deaux & Major, 1987; Eagly & Wood, 1991).

10.07.4 THEORETICAL OVERVIEW OF GENDER DIFFERENCES

Having established that gender differences exist in certain psychological and behavioral domains, we can turn to questions more pertinent to the present chapter on the relationship between gender and health psychology. For instance, why do males and females sometimes differ considerably, moderately, or not at all in their vulnerability to physical and psychological disorders? To what extent are the sources of these distinctions biological, environmental, or an interaction between both?,Furthermore, how does the health of the individual depend on her or his social roles and abilities? Shifting from the question of whether gender differences exist to explaining why males and females are predisposed to different health problems requires testable theories about how gender differences come about and what biopsychosocial forces maintain them.

Most modern theories about the origin of gender difference adopt an "interactional" perspective, whereby the biological features of each sex interact with the social environment to produce the observed gendered behavior of men and women. On the one hand, however, it is fair to say that theories that give prominence to "biological" features treat gendered behavior as arising primarily from characteristics that are built into the person prior to social experience. On the other hand, social psychological theories regard gendered behavior as primarily arising out of the interaction between the person and his or her social environment.

10.07.4.1 Biological Perspective

Few modern theories attempt to explain the differences in the behavior of men and women solely in terms of biological forces. Recently, Buss and Schmitt (1993) and Kenrick (1994) have advocated what has been termed "evolutionary psychological theory" to account for some different features in the behavior of the sexes. Evolutionary psychology, which was first applied to nonhuman species, is based on the premise that biologically based sex differences have evolved over time, employing the Darwinian notion of sexual selection (Buss, 1995). These biologically oriented psychologists have postulated that there are sex differences in genetic make-up or differences in the structural features of females' and males' brains. Presumably, these biological distinctions have evolved over time to produce functionally useful aspects of masculine behavior for men, and feminine behavior for women. Evolutionary psychology predicts that males and females will be similar in those domains in which the sexes have faced similar adaptive problems. For example, both sexes have similar preferences for fats, sugars, and salt because both sexes have faced similar food consumption problems. However, in some domains the sexes will have faced different problems. For instance, men have evolved to have greater skills in spatial visualization than women, because men have had a greater need to develop these skills for hunting. Women, on the other hand, have faced the adaptive challenge of securing sufficient food to carry them through pregnancy and lactation; therefore, evolution might have favored selection of women who developed the ability and features to attract mates to help them through this period. Thus, evolutionary psychology predicts distinctive male and female behavior in areas where sexual selection has favored the development of the physically and psychologically different adaptive characteristics for men and women. Presently, the evolutionary account of sex differences is regarded by many gender researchers as highly speculative. However, this perspective has presumably generated testable hypotheses (Eagly, 1995).

10.07.4.2 Social Psychological Perspective

Over the past several decades, a diverse group of scientists has contributed to the social psychology of gender, which in its simplest description investigates how gendered behavior is defined, created, and maintained for each sex through social norms and social interaction. Social psychological theories that feature discussions about the acquisition of disparate gender roles and how "sex-typed" gender roles create stress-related problems for men compared to women have generated much of the recent data on gender and behavior for clinical health psychologists. For example, the masculine gender imperative for men always to appear invulnerable may be related to their reluctance to seek medical or psychological care. Similarly, the feminine gender imperative for women to be ever youthful and attractive may contribute to excessive dieting among women, which may in turn render women especially vulnerable to eating disorders.

The following are a few examples of the contributions of social psychologists that view gender distinctions more as a social construction than a biological reality. Bem (1981, 1982) contributed to gender role theory by supposing that children develop cognitive schema to organize information about the world and themselves in terms of socially desirable masculine or feminine traits. That is, at an early age, boys and girls learn to judge their adequacy as people based on whether their thoughts and behavior are in conformity with social norms for their gender. For example, women may be reluctant to be assertive about their needs because it conflicts with their gender schema about appearing appropriately conciliatory. A man's gender schema might direct him not to express his true worries and concerns because it is inconsistent with masculine values to show one's emotional vulnerability. "Never let them see you sweat" was an often quoted line from a men's deodorant commercial. Indeed, an individual's own behavior is constantly being scrutinized by his or her own cognitive commitments to being in conformity with masculine or feminine ideology.

In addition to what we as women or men consider is acceptable behavior for our gender, social psychological research has shown that a great deal of our behavior is based on what others expect from us solely on the basis of our being male or female. Geiss (1993) and Maccoby (1990) have reviewed research which showed that our sex-typed expectations of each sex tends to generate "self-fulfilling prophecies" in producing the behavior expected of our gender. For example, if we expect a woman to be submissive in the presence of a dominant male colleague, she may well conform to this "feminine" expectation in his presence. However, she might not act submissively with her husband who does not expect her to act in a stereotypically feminine way. Because we expect males not to show their pain, we might see a little leaguer get up smiling defiantly after being knocked down by a fast ball at the plate. Moreover, it is evident that our own beliefs about how each gender should think and act, as well as others' expectations about how we as women or men should think and act are major determinants of our behavior. The notion that gendered behavior is socially constructed and enforced by society to provide distinctive roles and behaviors for men and women has important implications for people who live in diverse cultures. If the behavior of women is largely the result of cultural imperatives, we would expect the behavior of women to differ depending on the culture in which they live.

10.07.4.3 Cross-cultural Perspective

Cross-cultural research on gender has been concerned with the degree to which each gender's attitudes and behaviors are either invariant across cultures or may be different depending on the distinctive features of one's cultural environment. Let us frame the question in another way. Are men less expressive of tender emotions than women in all cultures? If the answer to this question is yes, it would mean that inexpressiveness is intrinsic to all individuals who are biologically male, no matter what their social environment. On the other hand, if a man's emotional expressiveness depends on the particular culture in which he is found, then we would have evidence that expressiveness in men is largely determined by cultural forces.

While the study of cultural differences in gendered behavior is still in its infancy, Williams and Best (1990) have collected some interesting data on sex role ideology, sexual stereotypes, and work values in 14 different countries. Their results showed a high degree of agreement between men and women living within a given country about expected gender roles. However, there were important differences in social expectations about distinctive gender roles across different cultures. For example, within developed Protestant countries, both genders emphasized the importance of men behaving assertively and having power and possessions to a greater extent than women and men in the less economically developed countries. Also, men and women were seen as necessarily more distinctive in their gender roles by people in developed countries, while people from less developed countries tended to see the gender roles of males and females as more similar. Williams and Best (1990) concluded that the effect of culture was greater than the effect of biological sex in determining the expected behavior of women compared with men.

10.07.4.4 Summary of Gender and Health Perspectives

Taken as a whole, research examining the basis for gendered behavior seems to demonstrate that biological differences between the sexes may establish predispositions for certain kinds of health problems. Female hormones have been found to offer protection from heart disease prior to women's menopause. Greater cardiovascular reactivity in men may predispose males to coronary artery disease. However, most researchers would agree that sociocultural forces appear to have the most profound effect on predisposing each gender to different health risks. In the following sections we hope to show that the mechanisms linking gender to health involves, to a great extent, gender differences in beliefs about health, exposure to certain health risks, gender determined roles, and gender relevant stresses and coping styles.

10.07.5 THE ROLE OF PSYCHOLOGICAL STRESS IN WOMENS' AND MENS' HEALTH

At this point we have some theoretical perspectives about the acquisition and maintenance of socially constructed gender roles related to socially accepted behavior for women as distinct from acceptable behavior for men. We now move to a discussion about how each gender evaluates and copes with the stressors to which each may be exposed. This is crucial in our understanding of the nature of the relationship between gender and vulnerability to specific health problems. We begin with an orientation to the biological and psychological features of stress and coping as they relate to health.

10.07.5.1 Gender and Biological Vulnerability to Stress

Modern theories of how stressful situations produce health problems for each sex are based on the notion that psychological responses to stress produce aberrant physiological or cognitive responses that may be different for women and men. For example, it has been demonstrated that men produce larger increases in blood pressure in response to some forms of stress than women (Stoney, Davis, & Matthews, 1987). The question of how differences in the sex hormones estrogen, progesterone, and testosterone affect gender differences in health has also been a novel area of study. There has been some evidence that the female hormone estrogen is protective against atherosclerosis for women, whereas testosterone may enhance the risk of this disease for men (Polefrone & Manuck, 1987). In addition, there appear to be consistent differences between women and men in healthy immune function, which is suppressed by gender related stresses (Kiecolt-Glaser & Glaser, 1987). Women, for example, exhibit higher immunoglobulin levels than men, and produce larger antibody response to infection than do men (Michaels & Rogers, 1971).

Psychologists and other scientists who have accumulated large quantities of data over the years about stress have concluded that stress is prominently involved in the production of a myriad of physical and psychological disorders (see Goldberger & Breznitz, 1982; Neufield, 1989). During the latter part of the twentieth century, conceptions of stress have evolved and have repeatedly shown that biological responses and the psychological make-up of a person interact with environmental pressures to produce deleterious somatic and mental health problems. Psychological stress has been acknowledged to pose health risks for both medical and psychological types of disorders (Goldberger & Breznitz, 1982; Neufield, 1989). The fact that psychological stressors may be the origin of either mental or somatic health problems, and the fact that biological and psychological systems have been shown to interact have tended to blur earlier distinctions between the two categories of health problems.

Currently, sophisticated views about stress describe an interaction between biological and psychosocial factors. First, there are gender differences in the probability that each gender will be exposed to a particular stressful situation. For example, women are more likely to be exposed to the stress of sexual harassment and rape than men. Divorced women compared to divorced men are more likely to be exposed to the dual stresses of having to earn a living and manage a family. However, men may be more stressed than women at some work sites because men have been expected to deal with physically hazardous situations (e.g., working in a coal mine) to a greater extent than women. In addition to the nature of the environmental events, modern views about stress focus on the individual's cognitive appraisal of the threat and her or his psychological coping strategies.

10.07.5.2 Gender and Cognitive Vulnerability to Stress

Mediating the relationship between traumatic events and physiological stress arousal are cognitive processes that assess the nature of the threat (Lazarus & Folkman, 1984; Neufield, 1989). According to Lazarus and Folkman (1984), psychological stress involves the relationship between the person and the environment that is appraised by the person as taxing or exceeding his or her coping resources and endangering his or her well-being. Thus, the interaction between our cognitive evaluations of particular situations and our assessment of our abilities to manage those problems physically and psychologically defines the stress process for us as individuals.

Furthermore, according to Lazarus and Folkman, the strength of one's commitment to a particular outcome increases vulnerability to stress. Therefore, should a woman become extremely committed to losing weight, her vulnerability to stress will be proportional to the strength of her commitment to weight loss. These concepts about stress have been adopted by most clinical health psychologists, whether they work with psychological disorders such as anxiety and depression, or with the effects of stress on biological systems, such as in the cases of gastric ulcers and coronary heart disease.

10.07.5.3 Gender Role Stress

A fair amount of evidence has linked sex differences in the incidence of various health problems to gender differences in appraisal of stress and utilization of coping responses (see Barnett, Biener, & Baruch, 1987; Baum & Grunberg, 1991; Collins & Frankenhaeuser, 1978). Building on the previous cognitive conceptualizations of stress, the present author and his colleagues have developed a paradigm describing gender role stress to explain how one's commitment to gender role ideology produces stress that may lead to gender-related health problems.

Incorporating Pleck's (1981) initial notions about the stress imposed by gender roles, we have developed a model that independently evaluates masculine gender role stress (MGRS) for men and feminine gender role stress (FGRS) for women. This gender role stress model reflects our attempt to account for some of the vulnerability of women and men to contrasting health problems (Eisler, 1995; Eisler & Blalock, 1991; Eisler & Skidmore, 1987; Eisler, Skidmore, & Ward, 1988; Gillespie & Eisler, 1992; Lash, Eisler, & Schulman, 1990; Lash, Gillespie, Eisler, & Southard, 1991; Martz, Handley, & Eisler, 1995; Watkins, Eisler, Carpenter, Schectman, & Fisher, 1991).

The gender role stress paradigm includes several major components. The first of these is the development of gender schema through social learning environments, which tend to reward the development of masculine conceptions of self for boys and feminine conceptions of self for girls. Opposite gender conceptions of self tend to be strongly discouraged. For example, boys are rewarded for acting tough and girls for being nice. Boys are punished for acting feminine (sissy) and girls for acting aggressively (bitchy). These gender schema are then utilized to evaluate whether or not a particular situation may be threatening. Expressing feelings of vulnerability tends to be far more stressful for men than women. On the other hand, many women are much more concerned about their looks than men so that a substantial weight gain is more stressful to women. Finally, one's gender schema guide each person's choice of coping responses. A man whose masculine gender schema promotes highly vigorous means of coping with stress may rely extensively on aggression to cope with stress, whereas a woman might feel that aggressive displays were inconsistent with her feminine gender schema of how women should act.

According to this paradigm, gender role stress may result from faulty gender determined appraisals of the situation, in that a man's expression of fear (feminine) need not lead to environmental stress for him, whereas acting in a forceful (masculine) manner might reduce the environmental pressures felt by a woman. Alternatively, gender role stress may result from excessive reliance on socially approved masculine or feminine coping styles. For example, a man's reliance on a hostile or combative style of coping with stress may be dysfunctional in that it produces chronic physiological arousal leading to a heart attack. A woman who copes with stresses by the continual expression of dysphoric emotions and rumination about her difficulties could end up feeling helpless and depressed. Consequently, gender role norms have helped create gender role coping strategies for an individual based purely on his or her gender. Living up to these culturally defined masculine or feminine expectations may, in itself, be difficult for each gender, and produce stress or dysfunctional appraisal and coping behavior.

10.07.5.4 Gender Roles and Coping with Stress

A major issue linking gender role stress and health has involved gender differences in coping with stress. As we alluded to earlier, consistent with their gender roles, men are more likely to externalize and women internalize their stress (Huselid & Cooper, 1994). Therefore, masculine ideology specifies different kinds of coping behavior for men than feminine ideology does for women. Because of social norms, men are more likely to utilize instrumental aggressive behaviors and displays of dominance to cope with their problems than women. However, women may be more likely to ruminate and engage in self-blame in response to crises in their lives because this behavior is more consistent with feminine gender role expectations.

When we look at gender-related coping styles, we can see that some health problems may result from dysfunctional coping styles associated with one's gender role. For example, most psychological disorders involving antisocial and abusive behavior reflect exaggerated aggressive coping styles typically used by men. This might explain why there is a strikingly high incidence of men compared to women who commit crimes of violence (Widom, 1984). Also, women make more suicide attempts than men, yet men are more likely to succeed in killing themselves. Psychological disorders utilizing passive, stereotypical feminine coping styles are disorders that show higher prevalence rates for women, such as anxiety disorders and obsessive-compulsive disorders. Learned helplessness (Seligman, 1975), a major cognitive theory of depression, is more consistent with enacting the feminine rather than the masculine role. Admittedly, much research remains to be done to demonstrate that there are direct relationships between masculine or feminine gender roles and gender role coping strategies responsible for particular disorders. However, the possibilities are intriguing, and we have conducted a number of preliminary studies to evaluate the possible association between gender role stress and coping and potentially adverse health effects.

Utilizing the MGRS scale (Eisler and Skidmore, 1987), we found that there were significant associations between the MGRS scale measure and self-reported anger, anxiety, anger, and poor health habits in men (Eisler & Skidmore, 1987; Eisler, Skidmore, & Ward, 1988). In addition, Lash, Eisler, and Schulman (1990) showed that men with a high score on the MGRS scale, compared to men with a low score, were subject to greater increases in blood pressure when the stressor dealt with masculine relevant tasks in which men were expected to perform well. No differences between men with high and low MGRS scale scores were noted when the task was concerned with feminine behaviors or tasks not particularly relevant to expected male gender roles. Additional studies by Lash et al. (1991) showed that women had greater increases in their blood pressure responses than men if the stress threatened areas that were inherent in the female gender role, including nurturance and child rearing. Watkins et al. (1991) showed that high gender role stress was associated with coronary-prone type A behavior and elevated blood pressure in both men and women. Martz et al. (1995) found that FGRS scores could distinguish women who had eating disorders from women who had other psychiatric conditions. Gillespie and Eisler (1992) showed that women who had high scores on the FGRS scale reported more depression than did women low in feminine role stress.

As a whole, these studies have demonstrated that responses to stress have gender components in that women are more likely than men to be stressed by events that the culture has specified as especially important to women's gender roles. Similarly, men are more subject to stress when the situation requires expected masculine behaviors. These studies have also tended to indicate that gender differences in favored coping styles may be predictive of the types of disorder that men compared with women are prone to develop.

10.07.6 GENDER DIFFERENCES IN SELF-DISCLOSURE AND HELP SEEKING

In this section we consider gender issues in the expression of emotion, the ability to elicit or utilize social support, and the inclination to seek help for personal problems. These are gender differences that may have profound implications for the health of women compared to men. The majority of verbal psychological therapies used in Western countries rely on the assumption that expressing one's feelings about past traumatic or painful events has a healing effect. Pennebaker (1995) has been foremost among researchers attempting to understand the beneficial effects of emotional disclosure on psychological and physical health. In the introductory chapter of his book *Emotion, disclosure, and health*, Pennebaker (1995) has reviewed the evidence that full expression of emotion reduces rumination and worry and promotes cognitive resolution of disturbing events. He also cites laboratory studies which show that disclosure of upsetting events has produced reductions in sympathetic nervous system arousal, including blood pressure and muscle tension. Other studies have shown that improvement in immune function is another benefit of self-disclosure (Petrie, Booth, & Davison, 1995). In reviewing the research detailed in chapters in Pennebaker's book by experts on psychopathology, psychophysiology, and counseling, one is struck by the evidence that the tendency to share emotional experiences has great power in increasing interpersonal intimacy, which in turn tends to buffer stress for the individual.

How do the genders compare on the tendency to express emotion and create social support networks to buffer stress? Is the ability to express innermost feelings related to an individual's tendency to seek help for emotional turmoil? These questions about emotional expression, social support, and the willingness to seek help for psychological and other health problems are very much interrelated issues. Hence we now turn to a discussion of gender differences in emotional expression, gender distinctions in the ability to develop social support networks, and the tendency for women as contrasted to men to seek professional help for health problems.

10.07.6.1 Gender and Emotional Expression

In a review of literature on gender differences in emotional expression Sauer and Eisler (1990) found that, almost without exception, women were more self-disclosing than men. However, it did appear that men self-disclosed more to women than to other men. Eagly and Wood (1991) explained these gender differences in terms of the widespread belief in our culture that women are more emotionally sensitive, more socially skilled, and more concerned with maintaining personal relationships than are men. Eisler and Blalock (1991) proposed an alternative possibility: that expressing feelings other than anger is perceived by men as "feminine" behavior, which is contrary to their masculine role. Finally, while developing a measure of MGRS, Eisler and Skidmore (1987) found that men found it significantly more stressful than women to express feelings of affection or fear. However, this might not be true for men in other cultures.

10.07.6.2 Gender and Social Support

The evaluation of social support in psychology has been hampered by a lack of agreement on a single definition of support and a broad range of measures. However, most researchers agree that the functions of social support include providing assistance in obtaining resources, problem solving, and bolstering self-worth (Belle, 1987). Research has shown that other mechanisms by which social support is associated with health are: (i) decreasing the supported persons's feelings of isolation and loss of self-esteem; (ii) attenuating the perception of certain events as highly threatening; and (iii) reducing physiological reactivity to stress (Shumaker & Hill, 1991).

Throughout the lifespan, evidence indicates that females are more likely to have supportive confidants than men. While the relative size of womens' and mens' support networks varies, womens' investment in supportive disclosing relationships consistently tends to exceed those of men (Belle, 1987). While mens' supportive relationships center around shared activities and experiences, womens' relationships emphasize shared emotional intimacy. Females are more likely both to seek and to receive social support in times of stress than men (Wolchik, Sandler, & Braver, 1984). Women are also more likely than men to be supportive as friends. Finally, the evidence indicates that men have much more restricted social networks than women. In fact, married men tend to rely almost exclusively on their spouses for support, whereas married women utilize other relatives and friends for support (Belle, 1987). The potential downside of this for women is that large social networks may be emotionally stressful as well as being supportive.

The impact of social support on the overall health of women compared to men is far from clear. We do know from the above data that women are more likely than men to participate in stress-buffering social networks and are more able to mobilize social support in times of stress. Women also provide more frequent and more effective support to others of both sexes.

Schumaker and Hill (1991) have concluded that the role of social support in womens' and mens' health awaits greater numbers of women to add power to the comparative health studies and more sophisticated measures of social support. We now turn to some data from the counseling literature on the relationship between gender and the tendency to seek help for health problems.

10.07.6.3 Gender and Help Seeking

A crucial determinant of preventing disorder and maintaining health involves the desire and ability to seek help for problems at an early stage of difficulty. Nearly 70% of all clients seeking psychological help are female. Furthermore, one in three women, as compared to one in seven men, has sought help from a mental health professional at some point during her life (Collier, 1982). Psychologists concerned with gender issues have written extensively about a strong association between the masculine gender role and the comparative reluctance of men to seek help for psychological problems (Eisler, 1990; Good, Dell, & Mintz, 1990; Scher, 1990). Until recently, most forms of psychotherapy were developed by men to be practiced on women by male psychotherapists. This unfortunate evolution of the provider–client relationship has created problems in the delivery of health services to both women and men.

In order to understand why men seek psychological and other health services as a last resort, we need to appreciate the demands of the masculine gender role and the fact that most forms of counseling have required behavior that may be regarded as feminine. Writers on mens' issues have theorized that client characteristics required by most forms of counseling and psychotherapy are completely antithetical to the masculine gender role (Eisler, 1990; Eisler & Blalock, 1991; Scher, 1990). Conventional wisdom is that, for counseling to be successful, the client must acknowledge personal problems, be willing to self-disclose, tolerate feelings of interpersonal vulnerability, and be willing to submit control of the central issues in one's life to an outsider. All these therapeutic expectations are completely contrary to what males have been taught is essential to perform adequately as a man.

Men are raised to see their world as an adversarial place where a man must attain mastery of the world by being independent and cope with stress by wielding power and not by revealing his vulnerabilities. Men perceive a basic incongruity between what is expected from them as men in society and what is expected of them as consumers of health services. The result of this is that men typically perceive going for help with personal problems as an intolerable loss of personal control and power. Psychotherapists perceive the behavior of most of their male clients as defensive, emotionally constricted, and insensitive because the male role of "invincibility" appears inconsistent with seeking help. We must conclude that, for many men, entrenched masculine values continue to be an important obstacle to seeking and receiving help for health problems.

10.07.7 FUTURE DIRECTIONS

Expanding our understanding of health from a biomedical enterprise to a biopsychosocial paradigm requires a greater understanding of health beyond assessment and treatment of disease processes and infectious agents. More comprehensive views of health require an understanding of psychological and sociocultural factors. An individual's beliefs and attitudes, and his or her perception of and exposure to stress that differentially affect the risk of illness have been shown to have important gender and cultural components. How these psychosocial components interact with biological systems continues to be a fascinating area for further theory and research. Research on health should endeavor to employ a more equal distribution of women and men, as well as to sample more individuals from different ethnic groups.

Many of the health problems of the twenty-first century that are reflected in premature mortality and morbidity will be preventable by attention to changes in lifestyle risk factors and creation of lower stress environments for people of different gender and cultural backgrounds. For example, preliminary research has shown that cultural and gender differences in dietary patterns, smoking, alcohol consumption, and high-risk sexual and aggressive behaviors are significantly related to differential risks of mortality and morbidity in different cultural groups (Anderson, 1995).

More research is needed to understand the different kinds of stress factors to which women and men and people from different cultural groups are exposed. Clearly, racism, sexism, and economic hardship variously constitute major stress and health risks for different groups. Presently, health research has focused on specific health problems such as heart disease, the effectiveness of smoking cessation programs, cancer screening tests, and blood pressure monitoring programs. Perhaps this should be supplemented by comprehensive studies examining the health risks to which diverse populations are exposed. For example, how are the dietary patterns and smoking habits of diverse populations including Asian, Black, White, and Latino women and men related to different rates of acute and chronic illness?

Based on these analyses, it appears that preventive health interventions are inadequately guided if they do not consider individuals as inseparable from their gender and cultural beliefs and roles. Men experience, appraise, and cope with stress differently than women, and these cognitions and behaviors undoubtedly vary for women and men belonging to different cultural groups as well. Therefore, to be successful in providing health interventions to diverse populations, we will need to address the social and cultural differences among these diverse groups of women and men.

10.07.8 SUMMARY

As stated in the introduction, research and practice in healthcare delivery has recently focused our attention on psychosocial risk factors in addition to the traditional emphasis on biology. A major aspect of this shift has been to study gender and cultural variables that affect group differences in prevention practices, health maintenance, activities, and differences in outcome. Many of our conclusions have been based on differences in cultural responses to stress and studying lifestyle risk factors for different groups.

It was apparent from our review that women and men experience different rates of both somatic and mental health problems, including anxiety disorders, excessive anger problems, certain types of cancer, and premature heart disease to cite just a few examples. Taken as a whole, the bases for differences in womens' and mens' health were found to be related to an interaction of biological differences with psychosocial risk factors or lifestyles. More specifically, gender differences in appraisal and coping with stress was discussed as a major factor accounting for a significant amount of variance in mens' and womens' health problems.

For example, gender differences in expressing emotion and tendency to seek help from the healthcare system were shown to have major implications for gender variations in health outcome. Gender differences in social support were difficult to evaluate because of problems of definition for empirical research. We concluded that the improved health care of the twenty-first century will be more attuned to research on the psychosocial characteristics of individuals who have different risk profiles for different medical and psychological disorders throughout their lifespan.

10.07.9 REFERENCES

Anderson, N. B. (1995). Behavioral sociocultural perspectives on ethnicity and health: Introduction to the special issue. *Health Psychology, 14,* 589–591.

Attie, I., & Brooks-Gunn, J. (1987). Weight concerns as chronic stressors in women. In R. C. Barnett, L. Biener, & G. K. Baruch (Eds.), *Gender and stress* (pp. 219–254). New York: Free Press.

Barnett, R. C., Biener, L., & Baruch, G. K. (Eds.) (1991). *Gender and stress.* New York: Free Press.

Baum, A., & Grunberg, N. E. (1991). Gender, stress, and health. *Health Psychology, 10*(2), 80–85.

Becker, B. J. (1986). Influence again: An examination of reviews and studies of gender differences in social influence. In J. S. Hyde & M. C. Linn (Eds.), *The psychology of gender: Advances through meta-analysis* (pp. 178–209). Baltimore: Johns Hopkins University Press.

Belle, D. (1987). Gender differences in the social moderators of stress. In R. C. Barnett, L. Biener, & G. K. Baruch (Eds.), *Gender and stress* (pp. 257–277). New York: Free Press.

Bem, S. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review, 88,* 354–364.

Bem, S. (1982). Gender schema theory and self-schema theory compared: A comment on Markus, Crane, Bernstein, and Siladi's "Self-schemas and gender." *Journal of Personality and Social Psychology, 43,* 1192–1194.

Buss, D. M. (1995). Psychological sex differences: Origins through sexual selection. *American Psychologist, 50,* 164–168.

Buss, D. M., & Schmitt (1993). Sexual strategies theory: An evolutionary theory on human mating. *Psychological Review, 100,* 204–232.

Cleary, P. D. (1987). Gender differences in stress related disorders. In R. C. Barnett, L. Biener, & G. K. Baruch (Eds.), *Gender and stress* (pp. 39–72). New York: Free Press.

Collier, H. V. (1982). *Counseling women: A guide for therapists.* New York: Macmillan.

Collins, A., & Frankenhaeuser, M. (1978). Stress responses in male and female engineering students. *Journal of Human Stress, 4,* 43–48.

Daly, M., & Wilson, M. (1988). *Homicide.* New York: Aldine de Gruyter.

Deaux, K., & Major, B. (1987). Putting gender into context: An interactive model of gender-related behavior. *Psychological Review, 94,* 369–389.

Eagly, A. H. (1995). The science and politics of comparing women and men. *American Psychologist, 50,* 145–158.

Eagly, A. H., & Crowley, M. (1986). Gender and helping behavior: A meta-analytic view of the social psychological literature. *Psychological Bulletin, 100,* 283–308.

Eagly, A. H., & Karau, S. S. (1991). Gender and the emergence of leaders: A meta-analysis. *Journal of Personality and Social Psychology, 60,* 685–710.

Eagly, A. H., & Wood, W. (1991). Explaining sex differences in social behavior: A meta-analytic perspective. *Personality and Social Psychology Bulletin, 17,* 306–315.

Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin, 94,* 100–131.

Eisler, R. M. (1990). Gender role issues in the treatment of men. *Behavior Therapist, 13,* 57–60.

Eisler, R. M. (1995). The relationship between masculine gender role stress and men's health risk: The validation of a construct. In R. F. Levant & W. S. Pollack (Eds.), *A new psychology of men.* New York: Basic Books.

Eisler, R. M., & Blalock, J. (1991). Masculine gender role stress: Implications for the assessment of men. *Clinical*

Psychology Review, 11, 45–60.

Eisler, R. M., & Skidmore, J. R. (1987). Masculine gender role stress: Scale development and component factors in the appraisal of stressful situations. *Behavior Modification, 11,* 123–136.

Eisler, R. M., Skidmore, J. R., & Ward, C. H. (1988). Masculine gender role stress: Predictor of anger, anxiety, and health risk behaviors. *Journal of Personality Assessment, 52,* 133–141.

Frankenhaeuser, M., Lundberg, U., & Chesney, M. (1991). *Women, work, and health: Stress and opportunities.* New York: Plenum.

Geiss, F. L. (1993). Self-fulfilling prophecies: A social psychological view of gender. In A. E. Beall & R. S. Sternberg (Eds.), *The psychology of gender* (pp. 9–55). New York: Guilford Press.

Gillespie, B. L., & Eisler, R. M. (1992). Development of the feminine gender role stress scale. *Behavior Modification, 16,* 426–438.

Goldberger, L., & Breznitz, S. (Eds.) (1982). *Handbook of stress: Theoretical and clinical aspects.* New York: Free Press.

Good, G. E., Dell, D. M., & Mintz, L. B. (1989). Male role and gender role conflict: Relations to help seeking in men. *Journal of Counseling Psychology, 36,* 295–300.

Hall, J. A. (1978). Gender effects in decoding non-verbal cues. *Psychological Bulletin, 85,* 845–857.

Harrison, J., Chin, J., & Ficarrotto, T. (1989). Warning: Masculinity may be dangerous to your health. In M. S. Kimmel & M. A. Messner (Eds.), *Men's lives* (pp. 296–309). New York: Macmillan.

Huselid, R. F., & Cooper, M. L. (1994). Gender roles as mediators of sex differences in expressions of pathology. *Journal of Abnormal Psychology, 103,* 595–603.

Kenrick, D. T. (1994). Evolutionary social psychology: From sexual selection to social cognition. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 26, pp. 75–121). San Diego, CA: Academic Press.

Kiecolt-Glaser, J., & Glaser, R. (1987). Psychosocial moderators of immune function. *Annals of Behavioral Medicine, 9*(2), 16–40.

Koss, M. P., Goodman, L. A., Browne, A., Fitzgerald, L. F., Keita, G. P., & Russo, N. F. (1994). *No safe haven: Male violence against women at home, at work, and in the community.* Washington, DC: American Psychological Association.

Lash, S. J., Eisler, R. M., & Schulman, R. S. (1990). Cardiovascular reactivity to stress in men. *Behavior Modification, 14,* 3–20.

Lash, S. J., Gillespie, B. L., Eisler, R. M., & Southard, D. R. (1991). Sex differences in cardiovascular reactivity: Effects of the gender relevance of the stressor. *Health Psychology, 6,* 392–398.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping.* New York: Springer.

Maccoby, E. E. (1990). Gender and relationships: A developmental account. *American Psychologist, 45,* 513–520.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.

Martz, D. M., Handley, K. B., & Eisler, R. M. (1995). The relationship between feminine gender role stress, body image, and eating disorders. *Psychology of Women Quarterly, 19,* 493–508.

Michaels, R. H., & Rogers, K. D. (1971). A sex difference in immunologic responsiveness. *Pediatrics, 47,* 120–123.

Neufield, R. W. (Ed.) (1989). *Advances in the investigation of psychological stress.* New York: Wiley.

Nolen-Hocksema, S. (1987). Sex differences in unipolar depression: Evidence and theory. *Psychological Bulletin, 101,* 259–282.

Pennebaker, J. W. (1982). *The psychology of physical symptoms.* New York: Springer.

Pennebaker, J. W. (Ed.) (1995). *Emotion, disclosure, & health.* Washington, DC: American Psychological Association.

Petrie, K. J., Booth, R. J., & Davison, K. P. (1995). Repression, disclosure, and immune functioning: Recent findings and methodological issues. In J. W. Pennebaker (Ed.), *Emotion, disclosure, and health.* New York: Springer.

Pleck, J. (1981). *The myth of masculinity.* Cambridge, MA: MIT Press.

Polefrone, J. M., & Manuck, S. B. (1987). Gender differences in cardiovascular and neuroendocrine response to stressors. In R. C. Barnett, L. Biener, & G. K. Baruch (Eds.), *Gender and stress* (pp. 13–38). New York: Free Press.

Robins, L. N., Helzer, J. E., Weissman, M. M., Orvaschel, H., Gruenberg, E., Burke, J. D., & Regier, D. A. (1984). Lifetime prevalence of specific psychiatric disorders in three sites. *Archives of General Psychiatry, 41,* 949–958.

Sauer, M. H., & Eisler, R. M. (1990). The role of masculine gender role stress in expressivity and social support network factors. *Sex Roles, 23,* 261–271.

Seligman, M. E. (1975). *Helplessness: On depression, development, and death.* San Francisco: Freeman.

Scher, M. (1990). Effect of gender role incongruities on men's experience as clients in psychotherapy. *Psychotherapy, 27,* 322–326.

Shumaker, S. A., & Hill, D. R. (1991) Gender differences in social support and physical health. *Health Psychology, 10(2),* 102–111.

Stoney, C. M., Davis, M. C., & Matthews, K. A. (1987). Sex differences in physiological responses to stress and in coronary heart disease: A causal link? *Psychophysiology, 24,* 127–131.

Travis. C. (1988). *Women and health psychology.* Hillsdale, NJ: Erlbaum.

Verbrugge, L. M. (1985). Gender and health: An update on hypothesis and evidence. *Journal of Health and Social Behavior, 26,* 156–182.

Verbrugge, L. M. (1989). The twain meet: Empirical explanations of sex differences in health and mortality. *Journal of Health and Social Behavior, 30,* 282–304.

Waldron, I., & Johnson, S. (1976). Why do women live longer than men? *Journal of Human Stress, 2,* 19–29.

Watkins, P. L., Eisler, R. M., Carpenter, L., Schectman, K. B., & Fisher, E. B. (1991). Psychosocial and physiological correlates of male gender role stress among employed adults. *Behavioral Medicine, 17,* 86–90.

Weissman, M., & Klerman, G. L. (1977). Sex differences in the epidemiology of depression. *Archives of General Psychiatry, 34,* 98–111.

Widom, C. S. (Ed.) (1984). *Sex roles and psychopathology.* New York: Plenum.

Williams, J. E., & Best, D. L. (1990). *Measuring sex stereotypes: A thirty-nation study* (Rev. ed.). Beverley Hills, CA: Sage.

Wingard, D. (1984). The sex differential in morbidity, mortality, and lifestyle. *Annual Review of Public Health, 5,* 433–458.

Wolchik, S. A., Sandler, I. N., & Braver, S. L. (1984). *The social support networks of children of divorce.* Paper presented at the American Psychological Association meeting, Toronto.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.08

A Cultural Perspective on Families Across the Life Cycle: Patterns, Assessment, and Intervention

Outline

NADINE J. KASLOW and KEITH A. WOOD
Emory University School of Medicine, Atlanta, GA, USA
and
MONICA R. LOUNDY
Georgia State University, Atlanta, GA, USA

10.08.1 INTRODUCTION 174

10.08.2 DEFINITION OF TERMS AND CONCEPTS 175
- *10.08.2.1 Culture* 175
- *10.08.2.2 Family* 175
  - *10.08.2.2.1 Normal and dysfunctional family processes* 176

10.08.3 CULTURE AND THE FAMILY LIFE CYCLE 177
- *10.08.3.1 Family Life Cycle* 177
- *10.08.3.2 Significance of Family Life Transitions Across Cultures* 177
- *10.08.3.3 Rituals Denoting Family Transitions Across Cultures* 177
- *10.08.3.4 Migration: A Life Cycle Transition Unique to Immigrant Families* 178

10.08.4 FAMILY LIFE CYCLE STAGES: A CULTURAL PERSPECTIVE 179
- *10.08.4.1 The Joining of Families Through Marriage: The New Couple* 179
  - *10.08.4.1.1 Marriage: a life cycle transition* 179
  - *10.08.4.1.2 Intermarriage* 181
- *10.08.4.2 Becoming Parents: Families with Young Children* 181
  - *10.08.4.2.1 Childbirth* 181
  - *10.08.4.2.2 Childrearing* 181
  - *10.08.4.2.3 Attachment* 182
  - *10.08.4.2.4 Developing a personal identity* 182
  - *10.08.4.2.5 Developing an ethnic identity* 183
  - *10.08.4.2.6 Developing effective coping strategies* 184
- *10.08.4.3 Families of Adolescents* 184
  - *10.08.4.3.1 Separation-individuation* 185
  - *10.08.4.3.2 Ethnic identity development* 185
  - *10.08.4.3.3 Achievement* 186
  - *10.08.4.3.4 Dating and sexuality* 186
- *10.08.4.4 Launching Children, Single Young Adults Leaving Home, and the Middle Generation Moving On* 187
  - *10.08.4.4.1 Cultural values and norms on launching* 188

10.08.4.4.2 Leaving home: single young adults 188
10.08.4.4.3 Conflicts experienced by the middle generation 189
10.08.4.4.4 Dealing with parental illness and death 189
10.08.4.5 Families in Later Life 189
10.08.4.5.1 Definition of the elderly 190
10.08.4.5.2 Issues confronting the elderly 190
10.08.4.5.3 Immigrant elderly 190
10.08.4.5.4 Ethnic identity 191
10.08.5 CULTURALLY SENSITIVE FAMILY ASSESSMENT AND INTERVENTION 191
10.08.5.1 Values 191
10.08.5.2 Help-seeking Patterns 192
10.08.5.3 Adherence 193
10.08.5.4 Culturally Sensitive Techniques 193
10.08.5.4.1 Stance of the therapist 193
10.08.5.4.2 Development of the working alliance 194
10.08.5.4.3 Assessment process 194
10.08.5.4.4 Intervention process 195
10.08.5.5 Culturally Sensitive Techniques Across the Life Cycle 197
10.08.5.5.1 The joining of families: culturally sensitive assessment and intervention 197
10.08.5.5.2 Families with young children: culturally sensitive assessment and intervention 198
10.08.5.5.3 Families with adolescents: culturally sensitive assessment and intervention 199
10.08.5.5.4 Launching children, single young adults leaving home, and the middle generation moving on: culturally sensitive assessment and intervention 200
10.08.5.5.5 Families in later life: culturally sensitive assessment and intervention 200
10.08.6 FUTURE DIRECTIONS 201
10.08.7 SUMMARY 202
10.08.8 REFERENCES 202

10.08.1 INTRODUCTION

Family theory and therapy and the multicultural perspective offer complementary views on assessment and intervention (Gushue & Sciarra, 1995). Both approaches share the assumption that the individuals must be understood within their larger familial and sociocultural context (Gushue & Sciarra, 1995). Thus, a comprehensive assessment and intervention should attend to the complex interactions among the individual, the family, and the culture across the family life cycle (Carter & McGoldrick, 1989; Szapocznik & Kurtines, 1993).

Family theorists and therapists increasingly have underscored the importance of the cultural context in which family therapy is conducted (e.g., Falicov, 1983; Ho, 1987; McGoldrick, Giordano, & Pearce, 1996; McGoldrick, Pearce, & Giordano, 1982). They have come to appreciate therapeutic models as reflecting the culture within which they are developed and practiced. Family theorists have become attuned to the role of culture and values in the conceptualization of interactive processes. Consistent with this, family therapists have become more cognizant of the impact of culture, values, and theoretical stance on the assessment and intervention process and outcome. These changes in perspective have led more family theorists and therapists to view each individual as a member of a family system that exists within a context of cultural diversity (Szapocznik & Kurtines, 1993).

After defining terms and concepts (e.g., culture, family) central to a discussion of a cultural perspective on families across the life cycle, a detailed presentation of cultural characteristics of families across the family life cycle is offered. Adapting the work of Carter and McGoldrick (1989), information is provided on cultural group differences across the major family life cycle stages. The aforementioned sections lay the groundwork for a discussion of culturally sensitive family assessment and intervention across the family life cycle. The theoretical underpinnings of family therapy from a culturally informed perspective are reviewed and culturally sensitive family assessment and intervention approaches in the United States are delineated. Hopefully, the information provided will inform psychologists about ways to adapt and modify current practices to enhance the cultural fit between the assessment and intervention approach and the family. Additionally, given the considerable intra- and intercultural variation in family patterns, it is our goal to present information that will help psychologists better treat families with whom they share a common cultural heritage and families from a different cultural background than their own.

We hope the reader will bear in mind the following caveats regarding our presentation. First, there is a lack of clarity in the literature regarding the distinction between culture and ethnicity. We use the term culture to refer broadly to values, attitudes, and behaviors that characterize groupings of people who are influenced by their culture of origin, religion, race, and socioeconomic status (SES). Second, the cultural characterizations presented are not meant as stereotypes of specific groups. Instead, the depictions reflect generalizations regarding common family patterns associated with a given cultural group (Falicov, 1983). These generalizations require validation at the level of the individual family. Third, and in a related vein, this chapter incorporates an intercultural perspective in which the definitions of what is appropriate or normal family structure, values, communication patterns, and interactional processes are seen as varying across different cultures. Although this chapter does not attend to individual differences between families within a given culture (intracultural perspective), we want to underscore the importance of attending to the unique characteristics of each family, which are often affected by such variables as the phase of acculturation, the stage of ethnic identity, and each members' choice of a primary language(s) both within the family system and in the larger community (Gushue & Sciarra, 1995). Fourth, due to space considerations, only a limited number of examples of some family cultures are provided. For in-depth reviews of families from specific cultures, and for coverage of a greater breadth of families, the reader is referred to McGoldrick, Pearce, and Giordano (1982) and McGoldrick, Giordano, and Pearce (1996). Fifth, it is not our intention to suggest that culture is the only or primary factor that impacts the family life cycle. Culture is one variable that interacts with multiple factors in influencing a family's presentation and effective assessment and intervention approaches. The salience of cultural issues differs across families, as does the significance of cultural factors in assessment and intervention. Thus, it behoves clinicians and researchers to incorporate culture as a key, albeit not sole, element in their work with families.

10.08.2 DEFINITION OF TERMS AND CONCEPTS

To set the stage for a discussion of culture and the family life cycle and culturally sensitive assessment and intervention, a number of terms warrant definition.

10.08.2.1 Culture

For this chapter, culture is defined as the totality of tacit behaviors, ideas and knowledge, attitudes, values, beliefs, customs, language, and technology shared by a group of people and transmitted by families from one generation to the next (Bullrich, 1989). Cultural patterns are influenced by multiple factors, including ethnicity, race, SES, education, gender roles, sexual orientation, country of origin, current residence, migration, religious and political affiliations, and family life cycle stage. Cultural norms and values inform family members' roles and rules regarding family membership and structure, rituals, help seeking behavior, problem definition and resolution, communication styles, expression of affect, and feelings about life, death, and illness (Carter & McGoldrick, 1996). In sum, culture is a dynamic and evolving process reflecting the interaction between a family's and a group's collective history and ongoing experiences in a particular context.

10.08.2.2 Family

Historically, the dominant American definition of family has focused on the intact, nuclear, heterosexual family unit, connected by blood or legal bonds, the members of which typically reside in one domicile. During the past few decades, the definition of the American family has expanded to include myriad constellations of emotional subsystems, such as step and remarried families, foster and adopted families, and cohabitating heterosexual, lesbian, gay, and bisexual couples.

With the multifaceted complexion of America, the definition of "family" varies among cultural groups (McGoldrick et al., 1996). For example, African-Americans' definition of family often refers to a wide network of kin and community and extended family members (e.g., grandparents, aunts, uncles, cousins) and nonfamily members (e.g., very close neighbors and friends, church leaders, boyfriends) who function as part of one's psychological family. It is understood that all members of this family will provide assistance and guidance in various aspects of family life including child rearing, caring for the elderly, providing material resources, and offering support during crises and transitions (Billingsley, 1992; Boyd-Franklin, 1989).

In the Latino community, the term family frequently connotes a close network of blood relatives that spans three or four generations. This network often includes nonblood members (e.g., godparents, legally adopted children whose adoptions have not been formalized). The traditional concept of the nuclear family has little meaning or applicability in the Latino subculture (Garcia-Preto, 1996).

Historically, the definition of family in certain Asian cultures (e.g., Chinese) included ancestors and descendants as well as living nuclear and extended family members. These families were patriarchal in structure. However, the Chinese-American family has evolved such that the contemporary family consists primarily of members of the nuclear family in which there is a more egalitarian sharing of roles across gender lines. This shift in family structure may reflect an accommodation to the dominant culture, the reality that only certain family subgroups immigrated to the US, and an increased emphasis on functional relationships (Lee, 1996).

In some Native American cultures, the primary parental functions are played by the grandparents and thus the key bond is between the children and the grandparents. The term grandparent in these families may refer to grandparents by blood, godparents, maternal or paternal great aunts, and so on. In many Native American families, no distinctions are made between one's family of origin and in-law family; these families are blended, rather than joined by marriage (Sutton & Nose, 1996).

10.08.2.2.1 Normal and dysfunctional family processes

Historically, family theories (for review, see N. J. Kaslow, Celano, & Dreelin, 1995) have portrayed the family as structured according to clearly defined, static, gender-based roles and family composition and have viewed the stability of the family household as the major family unit (Wamboldt & Reiss, 1991). These models also delineated the primary tasks to be accomplished by the typical family (Wamboldt & Reiss, 1991). For the most part, however, until recently, when detailing normal family structure and processes, these models failed to consider the effects of gender, social class, ethnicity, race, and sexual orientation (N. J. Kaslow et al.). This is unfortunate given that a Eurocentric model of normalcy is not applicable across cultures (e.g., Jamaican-Americans, Japanese-Americans, African-Americans) (e.g., Boyd-Franklin, 1989; Dechesnay, 1986; Hsu, Tseng, Ashton, McDermott, & Char, 1985).

For instance, while the role of the parental child is often viewed as indicative of dysfunction in the family structure, in many low-income African-American families, parentification of children may reflect a family's adaptive response to their economic conditions (Boyd-Franklin, 1989). In these families, the presence of a parentified child is dysfunctional only when these young people are inundated with abdicated, rather than delegated, responsibilities at the expense of their development (Boyd-Franklin, 1989). As another example, the extent to which triangulation (two opposing family members demand that a third person support him or her against the opposing party) within the family should be interpreted as reflecting maladaptive family patterns depends on the cultural context. In nuclear family structures common to EuroAmericans, triangulation reflects family pathology and thus may destabilize the family unit. In extended family structures common to many African-American, Asian, and Latino families, in which cross-generational executive subsystems (e.g., mother and grandmother) are normative, triangulating a child is less likely to affect the marital dyad negatively because authority lines are clear and the exclusion of one parent is not perceived as rejection (Falicov, 1983).

These examples highlight the fact that the cultural validity of various family theories and theoretical constructs (e.g., kinship, division of labor, proximity and hierarchy, triangulation, parentification) require empirical examination prior to making claims about universal v. culturally specific models of family normality and dysfunction (Boyd-Franklin, 1989; Fisek, 1991; Tseng, 1985). In addition, differences in family interactions between cultural groups (e.g., Japanese-American vs. Caucasian families in Hawaii) may be attributed to cultural factors, thus, it is essential that culture-specific profiles of healthy family functioning be devised (Hsu et al., 1985). If distinct, culture-specific portrayals of normal family functioning are not developed and validated empirically, normal interactions among family members that are disparate from those observed in the majority culture may be misconstrued as pathological (Hsu et al.).

Theories on normal and abnormal family functioning are socially constructed and shaped by the evolving cultural context (Walsh, 1993). Even the value placed on distinguishing between functional (i.e., normal) and dysfunctional family processes varies across cultures. While such a distinction has been of paramount importance in mainstream American culture, the conceptual dichotomy associated with the distinct categories of normal and abnormal is antithetical to the underlying assumptions of human behavior characteristic of many non-western societies (Foulks, 1991). Given that definitions of normal behavior vary across cultural groups, and the utility of the construct of normality differs between Western and non-Western societies, it behoves therapists to refrain from classifying certain family patterns as dysfunctional solely because they deviate from the norms of the dominant culture (Walsh, 1993). Rather, only families who manifest interactional processes discordant with the cultural context in which they are embedded and/or are associated with elevated levels of distress or problems establishing relationships with other cultural group members, require intervention (Foulks, 1991; Wamboldt & Reiss, 1991). For example, a person raised in a Southern African-American family affiliated with a Pentecostal Church who believes that their symptoms can be attributed to rootwork (a malevolent form of magic) may not be psychotic, particularly if these beliefs are shared by the family (Hillard, 1982). Although mental health services may be required, the patient and family may benefit from the assistance of a rootdoctor, who can remove a hex from the patient (Hillard, 1982). This example highlights that it is a challenge for therapists both to refrain from falsely ascribing dysfunction to a culturally normative pattern or exaggerating the significance of culture at the cost of failing to identify dysfunctional patterns (Falicov, 1983).

10.08.3 CULTURE AND THE FAMILY LIFE CYCLE

To present coherently the wealth of material on cultural characteristics of families across the life span, we organize our discussion using a family life cycle framework (Carter & McGoldrick, 1989), which provides a breadth and depth of insight into family functioning in all stages of development. This rubric is applicable to families from myriad cultures.

10.08.3.1 Family Life Cycle

The family life cycle, a circular model of normal family development, identifies key transitional stages of the family. First applied to intact, middle class, mainstream Euro-American families, this model divides the life span into six stages reflecting significant events in the lives of family members: unattached young adulthood, young couples, transition to parenthood, families with adolescents, launching, and families in later life (Carter & McGoldrick, 1989). As the composition of family units has evolved, the model has been revised to incorporate additional life cycle stages observed in families who experience divorce, single parenthood, and remarriage (Carter & McGoldrick, 1989). An additional life cycle transition pertinent to immigrant families, and thus relevant to this chapter, is that of migration.

Significant intra- and extrafamilial stress are evidenced by many families during family life cycle transitions. Elevated stress levels are associated with disruption in the emotional processes identified with each stage, as well as with the requisite second-order changes in family status that accompany each family life cycle stage. Thus, families often seek or require mental health services during key transition points.

10.08.3.2 Significance of Family Life Transitions Across Cultures

Cultural factors impact families' definitions of the nature, timing, tasks, and rituals of life cycle transitions, as well as the significance the family attaches to each transition (Carter & McGoldrick, 1989). For example, in Amish-American families, births are understood as gifts from God and associated with increased standing in the community; thus, births represent a transition of particular importance in the Amish community (Emery, 1996). Births are major events in many other groups, including most Latino cultures.

The shift from adolescence to adulthood, marked by the ritual of the Bar or Bar Mitzvah, is the most salient transition in Jewish-American tradition (Davis, 1988). The transition to adulthood is also one of importance in the Amish culture, denoted by the individual's decision to join the church and marriage (Emery, 1996). In many Latino cultures, *quinceaneras* (at 15, girls are given a party in honor of their coming of age) mark the adolescence to adulthood transition. Depending on the country of origin and the social class, the elaborateness of these festivities and the associated religious ceremonies vary.

Marriage is a highly valued family life cycle stage that carries considerable status in many cultural groups (e.g., Italian-American, Puerto Rican, West Indian). Finally, in some cultures, death is the most significant transition. This is the case for Irish-American families who emphasize rituals associated with mourning (e.g., wake).

10.08.3.3 Rituals Denoting Family Transitions Across Cultures

Cultures also differ in the rituals used to connote major life events (Carter & McGoldrick, 1989). For example, there are notable differences in the rituals associated with a child's birth. In the Jewish-American tradition, the birth of a son includes a celebration (*bris*) when the circumcision takes place. In response to the increased value placed upon females in the Jewish culture, baby naming ceremonies and parties have become more common for girls. In the Roman Catholic community, birth is marked by the baptism. In some Latino Catholic subcultures, baptism is accompanied by parties marking the beginning of a life long relationship among the child, parents, and godparents.

Marriage and wedding rituals are strikingly different across ethnic groups. An African tradition, adopted by many African-Americans, is jumping the broom, in which the couple jump together over a broom to mark their union. The broom signifies the beginning of homemaking for the couple. Consistent with ancient Chinese tradition, Chinese-American women may wear red wedding gowns. Red symbolizes prosperity, fertility, and good luck. In addition, before the wedding, the groom and bride give tea to elder family members as a token of appreciation for having been nurtured. Family members bless them with red packets of money that represent support for having a good life. In the Iranian community, the wedding is a spontaneous occasion, held at the bride's family's home in the presence of close relatives. Family members stand and clap during the ceremony, and the bride and groom are seated as they take their marriage vows.

Family and community behaviors that mark death differ significantly (McGoldrick, Hines, Lee, & Preto, 1986). For example, for Irish-American families, funerals are times of drinking and story and joke telling. In the African-American community, funerals are marked by a public expression of grief and moving spirituals. The funerals of Puerto Rican families are characterized by crying, screaming, and *ataques* (i.e., hysterical, convulsive reactions). In Chinese-American families, survivors are expected to cry without inhibition at the wake and funeral. To facilitate the deceased person's transition to the afterlife, and to insure that the next life is a good one, loved ones burn paper money and clothing and present elaborate spreads of food. Other dealth-related rituals that differ across cultural groups relate to the timing of the funeral (e.g., African-American Muslims and Jews bury the dead person quickly, whereas Hindus in America allow many days to elapse between one's death and burial) (McGoldrick et al., 1996) and the preparation of the deceased for the burial (e.g., in the African-American Muslim and Orthodox Jewish communities, often the deceased person's corpse is cleansed by individuals of the same sex as the deceased and given a ritual ablution) (McGoldrick et al., 1996).

Consistent with the evolving nature of the American family, rituals have been designed to mark transitions in a broader array of family constellations, such as welcoming adopted members, addressing remarriage, supporting lesbian and gay couples in making public their commitment to one another, and acknowledging divorce (Imber-Black, Roberts, & Whiting, 1988; F.W. Kaslow, 1993; Scrivner & Eldridge, 1995). Although historically, certain cultures (e.g., Orthodox Jews) have had divorce ceremonies for the purpose of freeing each partner to remarry and for acknowledging the legitimate birth of subsequent children, only recently has such a ritual been recommended for individuals across cultures. F. W. Kaslow (1993) presents a framework for conducting a divorce ceremony to help each partner heal from the pain of the termination of a failed union. In this ceremony, each partner affirms the positive aspects of the relationship and acknowledges that the children represent an important product of their union. The children are encouraged to process the impact of the divorce and to articulate expectations regarding future involvements with each parent.

10.08.3.4 Migration: A Life Cycle Transition Unique to Immigrant Families

Migration, a unique life cycle transition of immigrant families, reflects a developmental process of family adjustment with impact across multiple generations. The process of the immigrant family's transition to a new culture occurs on a continuum, and includes a number of specific stages: preparatory, active migration, overcompensation, decompensation, and transgenerational phenomena (Sluzki, 1979). Successful immigration is associated with both maintenance of the family's cultural identity and integration into the new culture; both tasks require reorganization of the family unit and the development of complex elaborations of family structure and processes (Bullrich, 1989). Although the migration process varies across cultures and is affected by the family's stage of development, cultural patterns, and the characteristics of the new culture, there are similarities in family patterns associated with the migratory process across cultures (Carter & McGoldrick, 1989; Sluzki, 1979).

Many refugee families seek mental health services in response to the stress associated with immigration and acculturation, in addition to the traumas experienced in their country of origin. Typical presenting complaints relate to intergenerational conflicts, marital discord, guilt and anxiety mixed with anger secondary to the tensions experienced between family members and relatives left behind, and conflicts with the sponsor/sponsor organization (McGoldrick et al., 1996). Often, a family assessment reveals that intergenerational difficulties are a screen for a more fundamental problem, namely a clash between the family's culture of origin and the host culture. In such cases, it is recommended that the therapist reframe the subsystem conflicts as culture conflicts, which ameliorates the fixed subsystem–culture alliances (Szapocznik, Santisteban, Kurtines, Perez-Vidal, & Hervi, 1984).

10.08.4 FAMILY LIFE CYCLE STAGES: A CULTURAL PERSPECTIVE

The following sections address cultural influences on each major family life cycle stage. Due to space considerations, only key stages are detailed. (i) The joining of families through marriage: The new couple. (ii) Families with young children. (iii) Families with adolescents. (iv) Launching children, single young adults leaving home and the middle generation moving on. (v) Families in later life. We deviate some from the family life cycle sequence provided by Carter and McGoldrick (1989), since the family life cycle is a circular process that can be conceptualized as begining at any stage. One of our modifications reflects our decision to initiate the family life cycle with the marriage of the young couple, which is the begining of a new family unit. Our second modification is our decision to combine two stages, launching children and moving on (Stage 5 in the Carter and McGoldrick model) with the leaving home: single young adult phase (Stage 1 in the Carter and McGoldrick model) because these processes often occur simultaneously in the younger and middle generations. Although cultural variations in patterns and belief systems across the family life cycle have been delineated, there is a dearth of empirical data on this topic (N. J. Kaslow et al., 1995). Relevant empirical data will be cited when available.

10.08.4.1 The Joining of Families Through Marriage: The New Couple

The decision to marry represents a commitment to a new family unit and a marital subsystem. This entails maintaining an intense attachment to someone outside of one's family of origin and renegotiating one's family of origin and peer group relationships. Such renegotiation includes redefining the structure and process of the relationships in order to incorporate a new member, and a change in the family member's identity. Forming a new family (family of creation, family of choice) also requires that the couple make decisions regarding routines and rituals, and customs and traditions.

Historically, in the majority of cultures, the ritual of marriage was the primary marker of the transition to adulthood. The significance and universality of this life cycle stage have shifted with changes in the onset of sexual involvements (earlier), increased incidence of cohabitation among nonmarried partners, increased reliance on nontraditional marital arrangements (e.g., same-sex couples, couples who cohabitate for long periods without having a legal ceremony marking their commitment), greater acceptance of people's choices to remain single or divorce or remarry, and later onset of marriages due to changes in educational and career paths. Also, although many groups prohibit intermarriages, there has been a significant rise in intermarriage rates since the mid-1970s.

10.08.4.1.1 Marriage: a life cycle transition

Virtually every cultural group has traditions and rituals associated with one's choice of a partner and the marital process. Between-group differences are noted on such factors as: process of mate selection, age at time of marriage, presence of a formal marital contract, meaning of marriage, marital roles and responsibilities, and expectations regarding monogamy and marital fidelity.

In the past, many cultures advocated arranged marriages. In some groups, arranged marriages are still common (Iranians, Greeks). In many Islamic Arab-American families, marriage is a family matter, in which partners are chosen by one's family and not on the basis of romantic love, as is the norm in most Western cultures (McGoldrick et al., 1996). In other communities (Chinese-American), arranged marriages of two young adults residing in their parents' homes used to be the norm. However, cultural shifts have resulted in many more marriages occurring after the adult children leave the family home and form a romantic attachment with a partner (McGoldrick et al., 1996).

In terms of age at time of marriage, in some cultures (e.g., Amish) it is expected that individuals marry young (late adolescence or young adulthood), whereas in other cultures (e.g., Jewish), people tend to marry at an older age. Further, sex differences exist within cultures regarding the expected age of marriage. For instance, among Iranian-Americans, females typically marry during adolescence or young adulthood, whereas males tend to be 10–15 years older than their partners (McGoldrick et al., 1982). In addition, the expectations regarding one's financial status often influence the typical age at the time of marriage. As an example, Jamaican males often marry 10 years after their Euro-American male counterparts, due to the expectation that they be financially secure prior to marriage (Comas-Diaz & Greene, 1994).

In the majority of cultural groups, heterosexual couples have a formal legal marital agreement. However, this is not true for couples of all socioeconomic levels in a given cultural group. As an example, formalized marriages are the norm among middle and upper class Haitians, but the equivalent of common-law marriages (i.e., *placage*) are common among low-income Haitian couples (McGoldrick et al., 1996). In the West Indian culture, although legal marriages are valued, many couples cohabit for lengthy periods of time to test their compatibility and they often procreate during this period (McGoldrick et al., 1982). Among some cultural groups, two types of formal marital contracts are obtained. Many Jews have a Jewish marriage certificate (*ketubah*) and a state-approved marriage certificate, with separate rituals associated with the receipt of each document.

There are myriad ways in which the meaning of the marital bond differs across cultural groups. For some groups, the formation of a marital unit indicates the creation of a new family designed to enable individuals to meet individual needs (e.g., Anglo-American); for others, it marks a continuation of the man's family line (e.g., Asian-American), and for still others it represents an alliance between two extended families (e.g., Nigerian, Puerto Rican) (McGoldrick et al., 1982, 1996). In some communities, getting married is a sign of status. Within the West Indian culture, a man only marries once he is capable of providing his spouse with economic security and material goods (e.g., house). In other cultures, getting married is an expectation and is not associated with status (Anglo-Americans). Finally, in some communities, marriage is conceived of as a symbolic rite of passage into adulthood (e.g., Amish), whereas other nodal events mark the transition from adolescence to adulthood in other communities.

With regard to marital roles and responsibilities, significant intercultural variations, often reflecting differences in gender roles, have been noted. In many cultural groups, the traditional marriage is one in which the male is the dominant member who has key roles outside the family system, whereas the women's major role is at home and with the family. Within many Mexican-American families, the ideal of machismo is expected from the husband; that is, the male member of the marital dyad is expected to have considerable sexual experience, and to be brave, pugnacious, and sheltering of his partner. Normative behavior of married women within Mexican-American homes is consistent with submission and humility. However, despite these gender role based social presentations, within the home, the domineering and patriarchal male member of the marital dyad often depends on his female partner for major family-related decisions (Falicov, 1996). In the Puerto Rican community and in many Dutch families, women have the primary responsibility for home and family, whereas men are responsible for protecting and providing for the family (Garcia-Preto, 1996). However, among Dutch partners, the husband often solicits his wife's input on business, career, and family financial matters. A reverse pattern of gender roles in marriage is observed among many Hungarian families (Smith, 1996). Hungarian women often are powerful in politics, business, and community affairs, yet have a clearly delineated role in the family typically confined to childrearing decisions and domestic tasks. In these families, the husband has ultimate power regarding domestic decisions. Among Amish-Americans, marriage is viewed as a practical matter and thus partners are expected to cooperate and be respectful of one another, but there is no expectation of romantic love or physical affection (Emery, 1996).

There are also within-culture variations regarding marital roles that often reflect the era of one's birth. As an example, Japanese-American couples in which both partners were born prior to 1935 tend to exhibit traditional, patriarchal, and hierarchical family structures in which the overt expression of affection and love is minimal, whereas couples with members born subsequent to 1935 more often assume an egalitarian relationship in which physical displays of affection are normative. Although these younger couples appear to have marriages more similar to that observed in many Euro-American families, they typically retain their cultural heritage by continuing to hold traditional Japanese attitudes, values, and beliefs (Matsui, 1996). Similar generational differences are exhibited among other cultural groups (e.g., Korean-Americans).

Fidelity is a value shared by members of virtually all cultural groups. However, the consequences of infidelity differ significantly. Among West Indians, adultery is an acceptable reason to divorce. In Vietnamese families, divorce is only an option if adultery is committed by the women. Although polygamy is illegal in America and thus is illegal for Vietnamese individuals residing in the US, the practice of men having female sexual partners in addition to their wives is somewhat acceptable and overlooked by the wife if her position in her extended family is secure and her children receive adequate provisions (McGoldrick et al., 1996). In the Jamaican community, although monogamy is supported, many Jamaican men maintain a mistress and their mutual children in a separate household (practice of "twin households") (Comas-Diaz & Greene, 1994).

10.08.4.1.2 Intermarriage

More than 50% of people in the US marry someone who is not a member of the same cultural group and 33 million adults reside in homes in which at least one other adult has a different religious background (Mayer & Kosmin, 1994). In 1990, 1.5 million children lived in families in which one caregiver was Caucasian American and another caregiver was of African, Asian, or Native American descent. The rates of African-American–Caucasian marriages have tripled since the mid-1960s, even though the actual numbers of such marriages remain relatively low (Wright, 1995). Within the Asian-American community, inter-racial marriages constitute 10–15% of marriages, with such marital unions being most common among Japanese-Americans and least evidenced by Korean-Americans (Karnow & Yoshihara, 1992). The increasing rates of intimate partnerships between individuals of different ethnic, religious, and racial backgrounds suggest that American society is becoming more tolerant of differences. However, couples from intercultural marriages often need therapy to address their cultural differences.

Inter-racial marriage complicates the successful negotiation of this family life cycle phase. Conflicts regarding values, religious beliefs and practices, communication styles, childrearing, and relations with in-laws are common (McGoldrick et al., 1996). The more salient the cultural differences between the partners, the more difficulty they are likely to experience in accommodating to married life (McGoldrick et al., 1996). It has been hypothesized that ethnic intermarriages reflect an effort on the part of each partner to redefine the most important ethnic characteristics and to abandon those ethnic traits that are valued least (McGoldrick et al., 1996).

10.08.4.2 Becoming Parents: Families with Young Children

Within the Euroamerican culture, the transition to this life cycle stage requires that the new parents develop a primary identity as caretakers for the next generation. Typical issues addressed are associated with childbearing, childrearing, and parenting (Carter & McGoldrick, 1989). Young children in families form attachments, develop increased competencies, and begin to develop a sense of self and an ethnic identity. This section reviews the major tasks for the adults associated with this family life cycle phase, including childbirth and childrearing. Then, attention is paid to issues focal in the child's development: attachment, developing a personal identity, developing an ethnic identity, and learning effective coping strategies.

10.08.4.2.1 Childbirth

Variations are noted in birth practices across cultures. In the US, the birth process has been considered a medical event, one that usually occurs in the hospital. With the influence of other cultures on American customs, more American families choose naturalistic approaches (e.g., home births, midwives, involvement of multiple family members) (McGoldrick et al., 1996).

10.08.4.2.2 Childrearing

Cultural differences in childrearing, parenting styles, and socialization practices are significant and often reflect the conditions (e.g., economics, work demands) and values (e.g., independence vs. dependence) of the culture in which they are formed (Matsumoto, 1994). For example, Mexican-American and Puerto Rican parents value family closeness and thus often socialize their children to be obedient and respectful and to prefer familial support over self-reliance (e.g., Zayas & Solari, 1994). As another example, Caucasian American mothers are more likely than African-American or Haitian mothers to use modeling and reassurance to help their children cope with fearful situations; Haitian mothers are more likely than either Caucasian American or African-American mothers to use force when their children confront anxiety-producing stimuli (Reyes, Routh, Jean-Gilles, Sanfilippo, & Fawcett, 1991). These differential parenting patterns could be construed as mothers using parenting styles that encourage adaptive survival in their home country.

The structure of the family affects childrearing and caregiving practices. In Euroamerican culture, the primary responsibility for childrearing has been the purview of the mother, and more recently the parental dyad, while in many non-Anglo-American cultures, children have been raised by extended family members (e.g., grandparents, aunts, and uncles) older siblings, nannies, wet nurses, or in group settings (e.g., kibbutz) (Matsumoto, 1994; McGoldrick et al., 1996). For example, African-

American children living in poverty are often raised in multigenerational families in which maternal grandmothers are the primary caregivers. Frequently, these grandmothers coreside with their daughter and her children, often with no adult male in the home (Harrison, Wilson, Pine, Chan, & Buriel, 1990). Although this family constellation differs from what has been considered the typical American family (i.e., heterosexual, married parents, who coreside with their biological children), it is often associated with healthy African-American child and family development (Canino & Spurlock, 1994; Wilson, 1989). In addition, a significant percentage of minority single-parent families are successful (Lindblad-Goldberg, 1989). As another example, in a number of cultures (e.g., Latino, Filipino), godparents serve as important models for young children and as sources of support for the young child's parents (Matsumoto, 1994). Although the composition of the extended family varies across groups, extended family networks are all characterized by a sharing of resources, emotional support, and caregiving responsibilities (Matsumoto, 1994).

In addition to cross-cultural differences in the identity of the primary caretaker(s), there are variations between cultures in terms of how caretaking occurs and the values communicated to young children within the family. These variations reflect differences in both beliefs regarding childrearing and in standards of living (e.g., family economic status, country of origin). In many Western European cultures, teaching some independence early is emphasized. Toddlers are raised in their own rooms, some crying is tolerated so the child learns not to cry manipulatively, strong verbal feedback during childhood is accepted, and the reward–punishment system emphasizes the gain or loss of possessions or privileges (Matsumoto, 1994). In many other cultures (e.g., Asian-American, African-American, Native-American) communicating a value on dependence and interdependence is emphasized (Matsumoto, 1994). For some, the young child may sleep in the parents' bed for several years and may be breast fed until age three or four. Often in cultures in which dependence and interdependence are emphasized, crying is not tolerated and thus may result in positive (e.g., picking up, cuddling) or negative (e.g., spanking) attention, disrespect in the form of talking back to elders (parental figures) is unacceptable, and the reward–punishment system frequently means the presence or absence of corporal consequences (e.g., Greek-American, Puerto Rican). These differential rearing and parenting practices prepare youth to live in the cultural context in which they are embedded.

10.08.4.2.3 Attachment

Cultural values in childrearing practices impact upon the nature and perception of the parent–child attachment bond. While many attachment behaviors are found across cultures, the selection, shaping, and interpretation of these attachment behaviors are culturally determined (IJzendoorn, 1990). Empirical investigations reveal frequency differences of the major forms of attachment (anxious avoidant, anxious resistant, securely attached, disorganized/disoriented) between and within cultures (across regional and socioeconomic groups) (Bretherton & Waters, 1985; IJzendoorn, 1990). These data underscore the importance of ascertaining the contextual meanings of attachment behavior, as well as whether or not attachment patterns are differentially adaptive across cultures (Harwood, Miller, & Irizarry, 1995).

Studies examining attachment behavior in Anglo vs. Puerto Rican mother–infant dyads indicate that Anglo-American mothers value child qualities associated with individualism (e.g., autonomy, self-control, activity), whereas Puerto Rican mothers prefer child factors consistent with interpersonal relatedness (e.g., affection, dignity, respectfulness, responsiveness to others, proximity seeking) (Harwood, 1992; Harwood & Miller, 1991). These preferences mirror the differences in values within the two cultures as individuation is promoted in Anglo-American culture, whereas relatedness is considered optimal in Puerto Rican culture. Despite these differences in valued child characteristics, both Anglo-American and Puerto Rican mothers rate the securely attached child as more positive than the insecurely attached child. However, the basis for this preference differs between the two groups of women. Anglo mothers note the confidence and independence of the securely attached child; Puerto Rican mothers find the child's demeanor, obedience, and relatedness most appealing (Harwood & Miller, 1991).

10.08.4.2.4 Developing a personal identity

The challenge for young children is learning, through interactions with key attachment figures, the adaptive patterns of their social group (Canino & Spurlock, 1994). Early in life, a child begins negotiating different affiliation styles, communication strategies, coping patterns, and family hierarchy and approaches to discipline. As the young children struggle with these issues and conflicts, they begin to form a personal identity.

The development of a personal identity is complicated when the child has to adapt to various, changing, and/or conflicting social rules and expectations. This is particularly true when the child must meet the demands of different cultures or adapt to a culture in rapid transition. Understandably, the child may have difficulty finding that adaptive skills in one culture may be a liability in another. For example, the openness valued in the Euro-American culture may be shamed in the Asian-American family context, and the gender-associated aggresiveness praised in the Latino culture may be punished in the Euroamerican culture. Many first-generation immigrant children become confused and overwhelmed by the need for their family to adapt to the dominant cultural demands. They may feel rejected by their peer group when their family engages in behaviors consistent with their culture of origin (e.g., afternoon siesta) and rejected by their family if they fail to conform with culturally normative rituals. In addition, these children often receive contradictory messages from their family as the family tries to meet the demands of the American culture (e.g., forego the afternoon siesta).

10.08.4.2.5 Developing an ethnic identity

As part of a personal identity, each child develops an ethnic identity, that is, a sense of belonging to an ethnic group with shared values, attitudes, language, behavior, perceptions, and social interaction patterns (Phinney, 1990a). The relevance of ethnic identity and awareness of this aspect of a person's identity are particularly important when children realize that their personal ethnic or cultural identity differs from the culture in power (Erikson, 1958). Therefore, being a "minority" or a member of a nondominant culture automatically predisposes a child to different identity development contingencies. For example, the dark, slanted eyes common in Asian-American children are not dark or slanted (they are normal) when everyone else's eyes are that way; but when a child realizes that beautiful and normal equals bluer and rounded eyes, that child has to deal with different identity-related issues than blue, round-eyed children. How this difference impacts identity is significant, and the family's role in helping the child develop a positive identity is crucial.

Accepting differences in personal characteristics is a developmental process. Children proceed from wanting to be something other than who they are (pre-encounter phase) to only wanting to be with and trusting others who are like themself (encounter and immersion–emersion stages), to accepting and appreciating variety (internalization stage) (Parham & Helms, 1981). This is a challenging task, especially when the socialization process (adult guided) is not clear on such issues. Older siblings and caretakers must model for younger children the internalization phase and help them negotiate earlier phases to facilitate their attainment of a more mature ethnic identity.

The best known empirical studies on ethnic identity in youth focus on minority children's preference for stimuli associated with Caucasian culture (e.g., white dolls) vs. their own culture (e.g., Banks, 1976; Clark & Clark, 1940). These findings initially were interpreted as indicating low self-esteem and a preference for the dominant culture vs. one's own culture. Many researchers criticized these early studies on methodological grounds (Banks, 1976). Also, the findings are dated and with the increased Afrocentric pride evidenced in most African-American families (Mahon, 1976) and the greater availability of black dolls in stores, African-American children are more likely to prefer black dolls. The other well-researched topic relates to children's self-identification, or self-definition, and capacity to correctly label their own ethnic identity (Aboud, 1987). Some suggest that incorrect self-definition vis-à-vis one's ethnicity indicates a negative self-concept, and thus may merit attention.

Much of ethnic identity is developed in the communication style in which the child is surrounded. For example, in many African-American social groups, the meaning of a message may be more in how (rhythm, tonal range) something is said than what is said. Part of the identity of the Italian language is the use of hand motions which are absent in Asian cultures. For certain African groups, much of the meaning of a message is in the nonword sounds that accompany the language, whereas for some Native Americans certain sounds are their words.

Cultures take possession of their unique ways of communication and generally are intolerant of deviations. In each culture, there are idiosyncratic words, phrases, grammar patterns, accents, and nonverbal behaviors that are understood by group members. For example, Jewish phrases help identify one's Jewishness. Children learn these language patterns and are reinforced for using these words in their family and with their Jewish peers, but may be teased if they incorporate these phrases into their communication with their non-Jewish peers. Black children sounding White to sound intellectual in the school environment may be perceived as rejecting their cultural communication identity, which could result in them having to deal with derogatory perceptions and comments from peers and family members. Cultural communication is also developed in musical types and unique sounds are associated with each culture. Learning these patterns of communication is a crucial aspect of children developing their ethnic identity.

Communication conflicts occur when one group defines the "right" way to speak and views certain cultural subtypes as "wrong" or inferior. Children struggle with their language identity when they learn different rules at school and at home. For some immigrant families, attempting to fit in with the major culture's language style has been critical and, as a result, their native language or dialects are not taught to their children. As an example, many Japanese families who were in internment camps during World War II discourage the use of Japanese by their children and grandchildren. This is in an effort to reduce their own pain and to be more "American," but may also function to hinder the development of a positive ethnic identity as a Japanese-American in their children or grandchildren (Vargas & Koss-Chiono, 1992).

Issues around appearance are another domain of ethnic socialization often associated with parent–child tensions. The young female whose family is from East India who wants to cut her hair in a stylish fashion (rather than never to cut it as is the cultural custom) like her friends poses issues for herself (identity formation), and is likely to meet with resistance from her parents who may feel distressed that their ability to control her ethnic identity is threatened. The Native American boy who only wants to perform the "in" (i.e., Macarena) is engaging in behavior that is likely to conflict with the traditional cultural patterns being enforced at home. What could be interpreted as "rebellious" may more appropriately be framed as the child's efforts to fit in.

Children must learn the subtleties of how to act appropriately within each unique cultural context. Much of this learning is acquired through parental modeling. Mothers and fathers struggling to negotiate their cultural practices in an environment intolerant of certain customs (corporal punishment is encouraged in some cultures, but is illegal in America) often send confusing messages to their children. This learning process is complicated further by possible cultural clashes at school, where appropriate behavior, modeled and reinforced by teachers and students, can be drastically different from the home environment (e.g., asking questions and challenging authority may be the norm in school, whereas deference to authority is expected in Asian-American families).

10.08.4.2.6 Developing effective coping strategies

A major challenge of childhood is developing a coping style for handling the obstacles further developmental stages present. If the child looks or sounds different from the majority of Americans, strategies must be formed for coping with being a minority and facing discrimination. Effective use of such tactics enables the child to establish a healthy self-view and a respect for others. While myriad coping strategies are potentially efficacious, it is the child's capacity to use a range of tactics flexibly that determines the overall coping ability. It is helpful to the child if these strategies are taught by elders. For example, African-American children can learn to cope with racism if their caretakers share African cultural values and practices and instil in them a sense of pride in their racial background (Vargas & Koss-Chioino, 1992).

When developing strategies for handling discrimination, it is often a challenge for children to find an approach that is respected in their own culture and the dominant culture. Withdrawing from or attacking the majority culture may be praised by one's own culture, but may result in rejection from the majority culture. Thus, this approach is inconsistent with developing bicultural competence.

The child must also form strategies for negotiating differences between the family's culture and the majority culture. Applying traditions associated with the environmental, social, or economic demands of a non-American country may have limited applicability to successful matriculation the US. One common coping strategy for maintaining important elements from one's culture of origin, while accommodating to one's culture of choice, is culture blending. This involves the Americanization of traditional practices, thoughts, and perceptions. Culture blending is noted throughout the US, including among the Creole population in New Orleans (French and American) and the Gullah population on the coast of South Carolina and Georgia (African and American). Culture blending is often manifested in the use of amalgamated language forms, such as Patwah (Caribbean dialect and English), Gullah (African dialect and English), and Spanglish (Spanish and English).

10.08.4.3 Families of Adolescents

The transition to this stage in the family life cycle often brings marked changes in the roles that adults and children take in the family. In Euroamerican culture, this stage is marked by the establishment of boundaries between adolescents and younger children, as well as parents relinquishing complete authority over the adolescent (Carter & McGoldrick, 1989). In addition, the married couple begin to address midlife marital and career issues and start to care for the older generation. This section focuses on cultural differences in negotiating the major issues of adolescence, separation-individuation, ethnic identity development, achievement, and dating and sexuality.

10.08.4.3.1 Separation-individuation

Families who experience problems in this stage often have difficulty negotiating flexible boundaries in which the adolescent is encouraged to be both independent and dependent. The shift in parenting patterns from childhood to adolescence is often problematic. As an example, some Irish-American families move from a structured, controlling parenting style in childhood to a permissive, laissez-faire style in adolescence, a shift often associated with increased behavior problems in the teenager (Carter & McGoldrick, 1989). Parents may be confused about how much autonomy is appropriate and norms on the balance between dependence and independence vary across cultural groups. In Arab-American families, adolescents are not expected to challenge parental authority, to engage in behaviors that conflict with parental expectations, or to place individual concerns over family interests (Timimi, 1995). Those Arab-American adolescents who rebel to attain increased autonomy are engaging in culturally unacceptable behavior and are likely to meet with intense parental resistance (Timimi, 1995). In a related vein, whereas Euroamerican families believe that healthy development in adolescence is associated with the young people's capacity to both love and leave their parents, in Japanese American families, the value is on family cohesion and support; leaving home is not viewed as an important aspect of adolescent development (McGoldrick et al., 1996).

In the most cultures, adolescent males are afforded more latitude to exert independence than their female counterparts, who often experience rigid and restricted rules. For example, an Italian-American family presents for help because their 16-year old daughter is "disrespectful, rejecting, and unappreciative." They report that she asserts her plan to stay out after midnight following her prom, intends to apply to and attend an out-of-state college, and prefers to buy her own clothes rather than wear clothes made by her mother. The 16-year-old claims her parents are old-fashioned and controlling. If the therapist holds the dominant value orientation, valuing autonomy more than responsibility to a collectivity (Papajohn & Spiegel, 1975), the cultural basis of the conflict, may not be acknowledged. As a result, the therapist will have problems supporting the family's negotiation of a mutually acceptable solution.

During this family life cycle stage, adolescents often fight to spend more time away from the family as peer relations become paramount. Similarly, as adolescents seek to fine tune their personal identity, they will often bring friends home who may expose the family to new values and ideas. While including nonblood relatives is welcomed in some cultures (e.g., Irish-Americans) it may be frowned upon as an intrusion in other cultures (e.g., Italian-Americans).

10.08.4.3.2 Ethnic identity development

Given the centrality of identity development concerns during adolescence, ethnic identity is of great importance during this developmental stage (Phinney, 1990a). Whereas the primary question for young children regarding ethnic identity relates to the accuracy of self-definition, during adolescence the major issues relate to the labels one chooses to use to define oneself, the degree to which one examines one's ethnic identity, and the degree and nature of group identification (Phinney, 1990a). In terms of labels, do adolescents whose parents migrated from Cuba refer to themselves as Cuban, Cuban-American, Hispanic, or Latino? Do adolescents whose parents grew up in Poland and moved to the US prior to their birth refer to themselves as Polish-American or simply American? Such decisions are complex for an adolescent, particularly if the parents are of different heritages, as parental heritage may not be consistent with the adolescent's ethnic self-identification (Phinney, 1990a). Further, each ethnic label has a different connotation regarding the adolescent's ethnic identity and group identification.

Research reveals that adolescents with a well-defined ethnic identity manifest more positive psychological adjustment than their counterparts with limited ethnic self-identification. In a group of African-American, Asian-American, and Mexican-American high school students, those who revealed more mature stages of ethnic identity during an interview had higher self-esteem, a greater sense of mastery, and more positive relationships than their peers who endorsed lower levels of ethnic identity (Phinney, 1989).

The degree and nature of group identification can often be inferred from the degree to which individual adolescents manifest positive vs. negative attitudes toward their ethnic group, and their involvement in the social activities and cultural practices of their ethnic group. The adolescent's involvement in such activities is evident via language choice, friends' ethnic background, religious practices, and participation in structured ethnic social groups (Phinney, 1990a). It is interesting to note that there are sex differences in adolescents' ethnic group identification; research suggests that females are more involved with their ethnic heritage and manifest more mature levels of ethnic identity than males (Phinney, 1990a; Plummer, 1995).

One important aspect of personal and ethnic identity is the expression of gender role, which is influenced by cultural factors. For Latino males, becoming a man often involves learning to physically fight, and taking a dominant role with females. These behaviors could be considered chauvinistic and threatening to those from a Euro-American culture that takes a more egalitarian view regarding gender differences. Traditional *Chicano* ways by a Mexican-American adolescent male toward a White girl could be interpreted as rude or discriminatory and result in a very negative social interaction, despite the fact that the behaviors reflect cultural differences in communication between the sexes.

10.08.4.3.3 Achievement

The arenas in which achievement is valued differ across cultural groups (Gibbs & Huang, 1989). Native American youth may become estranged from their community if they receive recognition for their academic successes, particularly if they choose education over tribal activities. For many Asian-American families, strong academic success is the focus of the family and the source of praise and attention. Excellence in sports, for example, is less valued among many Asian-American families compared to African-American families.

One aspect of achievement that has received considerable attention is academic achievement. A burgeoning body of research underscores ethnic differences in school performance among children and adolescents (Slaughter-Defoe, Nakagawa, Takanishi, & Johnson, 1990; Steinberg, Dornbusch, & Brown, 1992). Data indicate that African-American and Latino youth perform more poorly in academic settings than their Caucasian counterparts in terms of grades obtained and amount of education completed (e.g., Mickelson, 1990). Conversely, there is evidence that the academic performance of Asian-American students surpasses that of their Caucasian, African-American, and Latino peers (Sue & Okazaki, 1990). Environmental factors proposed to explain the impact of culture on adolescents' school performance focus on parenting practices, familial values regarding education, and beliefs communicated within the family regarding the benefits of education. The data on each of these variables are complex and do not provide a clear picture (Steinberg et al., 1992). Evidence suggests, however, that although the authoritarian parenting patterns often found in Asian-American families may not be optimal for superior academic performance, these parenting styles are typically counterbalanced by high levels of peer support for superior academic performance. Caucasian adolescents tend to have optimal academic performance when their parents use an authoritative parenting style and they receive support from their peer group regarding academic achievement. In addition, difficulties in academic success among both African-American and Latino youth may be attributed to low levels of peer support for academic excellence. This low level of peer support is often accompanied by authoritarian parenting practices in the homes of Latino youth, a combination of factors that often interferes with academic success. These data underscore the importance of focusing on cultural values and expectations, as well as family and peer relations, in assessing academic difficulties and devising intervention programs. The development and implementation of such preventive intervention programs must be sensitive to the child's economic conditions.

10.08.4.3.4 Dating and sexuality

Another important task of adolescence is managing dating and sexuality. The major issues that differ across cultures regarding dating relate to the presence of a chaperon and who chooses who the adolescent will date (and ultimately marry). The tradition in many Cuban-American families continues in which chaperons accompany the dating couple (McGoldrick et al., 1982). The chaperon's presence may be a major source of parent–adolescent conflict for adolescents who are more acculturated than their parents. The custom of parental mate selection persists in a number of cultures outside the US. Families from some of these cultures continue this practice upon immigrating to America (e.g., Arab-Americans, East Indians). Again, this practice may cause disagreement between acculturated adolescents and their families. A major issue among many cultural groups relates to parental and extended family concerns regarding the adolescent's decision to date someone outside the family's culture, race, or religion. Interethnic, interracial, and interfaith dating on the part of the adolescent is often a controversial family matter.

Many immigrant populations (e.g., Portuguese, Amish) have strong beliefs that sex is reserved for marriage. However, other cultural groups are permissive of sexual experiences prior to marriage. Italian-American families often encourage and expect their adolescent sons (not daughters) to become sexually proficient as a sign of masculinity and mastery of interpersonal relations. Given that premarital sex is the norm in present day American culture, it is common for a family whose heritage views premarital sex as unacceptable and deserving of punishment to seek mental health services for the adolescent engaged in sexual relations. With this presenting problem, a family intervention rather than individual therapy for the adolescent is warranted. The goal of the family intervention is to help all members address the conflicting norms and understand one another's behavior and attitudes within the larger sociocultural context. In families such as Chinese-American families in which sexuality is a taboo subject, the therapist must form a strong working alliance with the family prior to addressing issues of sexuality (Gibbs & Huang, 1989).

One possible consequence of sexual behavior is pregnancy. The rates of pregnancy and teenage parenthood differ significantly across cultural groups. For example, adolescent childbearing and parenthood are more likely to occur in low-income African-American samples than in demographically matched Caucasian samples or middle-income African-American groups (Franklin, 1987). Also, the rates of teenage pregnancy in the US are increasing, most notably among African-Americans and Latinas (Jacobs, 1994). Multiple psychosocial factors have been proposed to explain differential pregnancy rates among adolescents from different ethnic groups. It has been suggested that African-American teenagers from low-income or single-parent homes become more active sexually than their Caucasian peers and at a younger age. Their increased exposure to sexual behavior in the overcrowded environments in which they reside, the greater level of peer socialization associated with an enhanced awareness of sexuality at an early age, and the higher levels of involvement in adult responsibilities including sexual roles are hypothesized as factors to explain this discrepancy (Staples & Johnson, 1993). In addition to higher rates of sexual activity, low-income African-American adolescents are less likely than their demographically matched Caucasian peers to use contraception or to get abortions (Franklin, 1987). Further, there is increased tolerance of teenage pregnancy in many low-income African-American families and communities and such a pattern has a long history (Boyd-Franklin, 1989). In addition, extended family members are more likely to assist in raising children of teen mothers in the African-American community than in the Caucasian community (Boyd-Franklin, 1989).

10.08.4.4 Launching Children, Single Young Adults Leaving Home, and the Middle Generation Moving On

We have modified Carter and McGoldrick's (1989) family life cycle stages by combining the "Launching Children and Moving On" and the "Leaving Home: Single Young Adults" phases; these typically occur simultaneously in the family's life. This family life cycle stage for EuroAmericans refers to family functioning once children become of age and begin to leave the home.

For the middle generation, the major tasks of this transition are the negotiation and acceptance of adult children leaving the family unit and welcoming their partners and children into the extended family system (Carter & McGoldrick, 1989). The launching of adult children facilitates restructuring of the parent subsystem back to the marital dyad. In addition, it is likely that the couple will have to endure the disability and death of their parents. Families at this stage often present for therapy when parents, particularly mothers who worked as homemakers, experience depression and the "empty nest syndrome" at the loss of the family unit as defined since the mid-1970s. The couple's role as parents is no longer paramount and, as such, they may experience difficulties finding meaningful new life activities. This period may be the first time the couple has lived alone since their early marriage, and thus, they may need to become reacquainted as husband and wife. This may be a period of reinvestment in the relationship or a time when the couple move toward dissolving the marriage. For those who renew their marital commitment, this phase of the family life cycle often brings greater financial freedom, allowing the couple to engage in activities that may have been postponed earlier in the relationship (e.g., vacations, new careers, relocation).

For the younger generation of young adults from Euroamerican backgrounds, during the leaving home phase, the major emotional tasks are that of accepting responsibility (emotional, financial) for themselves, differentiating as an adult from their family of origin, developing intimate relationships, and gaining increased academic and/or occupational self-definition (Carter & McGoldrick, 1989). The individuation characteristic of this phase of development entails clear articulation of personal and ethnic identities (Phinney, 1990a).

10.08.4.4.1 Cultural values and norms on launching

There is a continuum of normative launching behavior across cultural groups, ranging from no expectation of launching to a high degree of expectation regarding launching at an early age and to a significant extent (Carter & McGoldrick, 1989). For example, in many Italian-American families, older adolescent and young adult children are not launched but are expected to remain in the family home or neighborhood, and the family integrates their partners and children into the family. In these families, interdependence is valued and autonomy strivings are experienced as disloyalty to the family unit. In some cultural groups (Iranian-American), the middle generation consider their adult children, even those who are married with offspring, to be children (McGoldrick et al., 1996). Further, in many refugee families, especially those who were traumatized in their homeland (e.g., Cambodian), the launching phase of the family cycle is threatening. The middle generation feels isolated and abandoned when the younger generation attempts to separate, particularly if their extended kin and social support network have remained in their homeland (McGoldrick et al., 1996).

Conversely, in many White Anglo Saxon Protestant families, value is placed on launching older adolescents and young adults and in their assuming significant independence. In such families, an 18-year-olds' inability or unwillingness to leave the family home and function with limited family support is perceived as pathological. Many cultural groups fall in the middle of this continuum, although the specific characteristics of this middle group vary. As an example, many Jewish-American families give mixed messages about separation during this family life cycle phase. Increased autonomy and success is valued highly on the one hand, but high levels of contact and intimate sharing continue to be expected on the other. These conflictual messages are often associated with guilty distancing or ambivalent closeness (Carter & McGoldrick, 1989).

10.08.4.4.2 Leaving home: single young adults

For single young adults, this stage of development includes continued separation–individuation, self-definition, and definition of ethnic identity, processes begun earlier in life. The recently "launched" young adults must grapple with the degree and nature of contact with their family of origin. Young adults from certain ethnic backgrounds (e.g., Jewish, Greek) maintain frequent contact (telephone, letters, email, visits) with their family of origin, yet attain a significant degree of autonomy to pursue success in their chosen endeavors. When they fail to attain such success, increased family involvement occurs (Carter & McGoldrick, 1989). Conversely, young adults from ethnic backgrounds that value independence beginning in older adolescence or young adulthood (White Anglo Saxon Protestants, Germans, Scandinavians) may be rejected by their family of origin if they cannot function autonomously due to personal limitations (e.g., disabilities) or need for continued financial support to pursue life goals (e.g., advanced training and education) (Carter & McGoldrick, 1989). Young adults often need to find a way to negotiate their personal desires for autonomy, their family's and culture's expectations about independent functioning, and the majority population's values on age-appropriate adult behavior. This negotiation is quite complex for young adults who leave their ethnic community and engage actively in mainstream culture.

Young adults' ethnic involvement during this phase is manifested in their choice of friends and partners, religious practices, structured ethnic social groups, political ideology and activity, area of residence, and career path (Phinney, 1990a). Studies examining ethnic identity among college students fail to yield consistent results regarding the association between positive or negative feelings about individual's ethnic identity, self-esteem, and psychological adjustment (Phinney, 1990a). However, research on the stage model of ethnic identity reveals an association between college students' achieved level of ethnic identity and self-esteem (Parham & Helms, 1985a, 1985b; Phinney & Alipuria, 1990). This relation is particularly evident in African-American, Asian-American, and Mexican-American young adults, and less obvious in Caucasian American college students (Phinney & Alipuria, 1990). Data also indicate that college students' stage of ethnic or racial identity is associated with their preferences regarding the race or ethnicity of their therapist (Parham & Helms, 1981). Specifically, older African-American adolescents and young adults who endorse early stages of racial identity (i.e., pre-encounter) often prefer Caucasian therapists and are nonaccepting of African-American counselors, whereas African-American adolescents and young adults who exhibit higher levels of racial identity (i.e., encounter, immersion–emersion, internalization) manifest varying degrees of both preference for same-race therapists and lack of interest in working with Caucasian counselors (Parham & Helms, 1981). These data suggest an association between identity formation and self-esteem; the more young adults accept their ethnicity, the more positive they feel about themselves. These data also underscore the need to consider the young adults' level of ethnic identity in making decisions regarding therapist assignments.

A major foci of the young adults' development relates to their choice of friends and intimate partners, and decisions about whether to be involved in a relationship or remain single. Once older adolescents and young adults leave the family home, they often feel an increased sense of personal choice regarding the ethnicity of their friends and partners and their sexual orientation. While this may be a relief for young adults, it often forces them to become more clear about their preferences and values regarding interpersonal relationships as distinct from those held by the family. Should young adults choose to engage in relationships not acceptable to their family's cultural norms and expectations (e.g., interracial, interfaith, or interethnic dating; same-sex relationships), the risk of being ostracized or criticized by the family of origin may be considerable and certainly depends to some extent on the family's ethnic heritage. For instance, since the African-American family often buffers its members from racism, these families are unlikely to reject a lesbian young adult family member, despite the fact that a lesbian life style may be considered unacceptable (Comas-Diaz & Greene, 1994). This stance does not imply acceptance of the young adult's sexual orientation but rather reflects tolerance and a desire to protect the young person from additional discrimination. This stance of tolerance, however, is often relinquished should the lesbian young adult openly define herself as a lesbian (Comas-Diaz & Greene, 1994). In the Native American community, young adult women who choose a lesbian life style are most often accepted in the lesbian community within the mainstream culture and rejected by their family and community on the reservation (Comas-Diaz & Greene, 1994). As a result, these young women often experience a loss of support if they explore a nontraditional sexual orientation.

10.08.4.4.3 Conflicts experienced by the middle generation

Realignment of the marital subsystem in response to the launching of young adult children and the manner in which marital dyads deal with the illness and death of their parents are influenced by the middle generation's culture of origin. Sex roles, an aspect of one's culture of origin, impact the negotiation of these family life cycle tasks (Carter & McGoldrick, 1989). For families whose cultural traditions emphasize the primacy of the woman's role as mother and wife (e.g., Italian-American, Puerto Rican), the launching phase may be associated with a crisis in the marriage if the woman becomes involved in activities outside the home. Conversely, in communities in which the woman's role outside the home is valued or accepted (e.g., Irish-American, African-American, White Anglo Saxon Protestant), the likelihood of this type of marital crisis at this stage is minimal. This is particularly true for those cultural groups in which limited emphasis is placed on intimacy within the marital dyad (e.g., Irish-American) or in which parent–child connectedness is valued more highly than the husband–wife bond (e.g., Mexican-Americans) (McGoldrick et al., 1996).

10.08.4.4.4 Dealing with parental illness and death

There are differences among ethnic groups about the responsibility the middle generation experiences regarding caring for their aging and dying parents (Carter & McGoldrick, 1989). In some ethnic groups (e.g., Greeks, Italians, Chinese, African-Americans), the middle generation expects to be the primary caretakers for their elderly parents and placing a loved one in a nursing home is rarely considered an option. Among certain groups (e.g., Germans, Scandinavians, Jews), the decision to place an elderly parent in a nursing home is commonly accepted. For some of these families, it is considered essential to provide aging parents with a familiar cultural environment (e.g., Jewish nursing home) in which familiar ethnic rituals and customs are practiced.

10.08.4.5 Families in Later Life

The elderly face many pressing challenges in the US, and as such the role of culture has often been overlooked by clinicians, researchers, and, at times, the elderly. As individuals get older, there is a greater emphasis on changing roles in the intergenerational family, the "culture of poverty," age discrimination, victimization, retirement, meaningfulness of life, health and vitality, loss of significant relationships and death of loved ones, and preparation for their own death. Such issues entail life review and integration. Families with elderly loved ones generally function best when they acknowledge and experience the wisdom of the elderly and support the older generation without overfunctioning for them (Carter & McGoldrick, 1989). The multitude of issues experienced by the elderly and their family often obscure the importance of culture in this period. This section begins by defining the term elderly. This is followed by a review of issues confronting the elderly, with specific attention paid to the immigrant elderly. Then, ethnic identity in the elderly is examined. It should be noted that cultural information regarding the elderly is minimal and the data are limited to a few subcultures (American Psychiatric Association, 1994a).

10.08.4.5.1 Definition of the elderly

In general, old is defined as 65–84, whereas old-old refers to age 85 and above. However, in some groups, being elderly starts at 55. In the US, some benefits associated with age begin at 55, others begin at 65 or 70. These numbers have changed over the years, partially due to the fact that the average life span has increased.

The onset of old age varies among cultures, in part because individuals of various cultures die at different ages (American Psychiatric Association, 1994a). The definition of old age is complicated further by the within-group variability in typical age of death. For instance, African-Americans tend to die younger than Euro-Americans. However, African-American individuals who survive past 75 have a longer life expectancy than Caucasian persons of the same age (Baker, 1994). Differences in life expectancies and health status raise questions about when a family and cultural group views an older loved one as elderly.

10.08.4.5.2 Issues confronting the elderly

The elderly confront myriad issues during the stage referred to as "Families in Later Life." First, non-Euro-American elderly individuals often experience discrimination, based on their age and cultural background. Latino elders often are the object of prejudice because of their Spanish surnames and inability to speak English (Jimenez & de Figueiredo, 1994). Due to their limited command of English, combined with the fact that interpreters often fail to capture accurately the perspective of the elderly, Latino elderly are often unable to defend themselves against discrimination and their behaviors and attitudes are often misinterpreted.

Appearance (e.g., clothing, accessories) and diet are variables associated with discrimination against minority elderly. For instance, because it is most common for the elderly to wear traditional dress, many elderly Asian-Indian women are ridiculed if they don a sari. Elderly people who retain their culture's traditions may feel particularly ostracized if their younger family members criticize their choices to retain their language and/or dress of origin.

The maintenance of beliefs from the homeland may also increase the elderly's vulnerability to oppression. The Chinese-American value placed on stoicism often conflicts with the American medical system, as well as the family system (Sakauye & Chacko, 1994). Medical personnel and younger family members often express frustration and confusion regarding the severity of an elderly Chinese-American's pain given their tendency to refrain from articulating the seriousness of their problems.

Many minority elderly face a sense of isolation and alienation due to their memories of traumatic experiences in the US and/or their country of origin. Elderly African-Americans who talk about slavery, elderly European Jews who reflect upon the Holocaust, and elderly Japanese-Americans who discuss the internment camp experiences of World War II may feel that their pain is discounted by younger family members who fail to appreciate the ongoing significance of such traumatic times. This sense of rejection by family members, who are unlikely to ever encounter such tragedies, further compounds the suspiciousness often associated with postslavery, the Holocaust, or internment camp experiences.

Some elderly people struggle to maintain a sense of autonomy and independence, despite increased family dependence due to economic limitations or physical or mental deterioration. This conflict is most pronounced in cultural groups in which caretaking of parents is not the norm (e.g., Germans, Scandinavians, Jews). Conversely, for those groups (e.g., African-Americans, Latinos, Native Americans, Asian-Americans) in which it is common for older family members to reside with, and be cared for by, younger members, independence–dependence struggles may be less central.

10.08.4.5.3 Immigrant elderly

From a cultural perspective, the immigrant elderly deserve specific attention. With the exception of African-Americans, Native Americans, Alaskan, and Hawaiian Natives, and those from cultures whose families immigrated to the US hundreds of years ago (e.g., White Anglo Saxon Protestants), many minority elders are relatively new immigrants to the country (e.g., Soviet Jews) (American Psychiatric

Association, 1994a). These elderly immigrants must adapt to a novel environment with new expectation while losing much of their accumulated experience from their culture of origin. As a consequence, they may lose much of their value within the family. Superimposed on this are the other issues surrounding getting older, resulting in the immigrant elderly feeling extremely challenged and overwhelmed. The intensity of these struggles appears correlated with the recency of immigration, with less recent immigrants manifesting fewer problems (American Psychiatric Association, 1994a). In spite of the aforementioned challenges, the elderly, particularly those who are immigrants, tend to be survivors who are resilient and exhibit a pleasure in life that is not quelled by age, suffering, or loss (Myerhoff, 1979).

10.08.4.5.4 Ethnic identity

Elders often assume the role of maintaining their own and their family's ethnic identity despite the family's assimilation into mainstream culture and associated loss of cultural knowledge, language, and traditional values and beliefs (American Psychiatric Association, 1994a). In the Native American community, this may be particularly problematic given that much of the tradition is maintained through oral history. A second aspect of ethnic identity of concern to many minority elderly is that they are no longer being as honored because of their strong ethnic identification (Thompson, 1994). Again, in some Native American families, as the younger generations become more Americanized, the elderly are less valued despite their wealth of experience. A similar process is evident in other groups (e.g., Asian-Americans).

10.08.5 CULTURALLY SENSITIVE FAMILY ASSESSMENT AND INTERVENTION

The conduct of culturally sensitive assessments and interventions is predicated on the incorporation of a conceptual model that emphasizes the behaviors, perceptions, beliefs, and values of different groupings of individuals. Clinicians and service delivery systems must be knowledgeable about, and attentive to, their own cultural backgrounds and those of the client and their family. Awareness of these cultural factors must inform the assessment and intervention process with families from all cultural groups (i.e., cultural competence) (Dana, 1993; Zayas, Torres, Malcolm, & DesRosiers, 1996).

According to Zayas et al. (1996) and Dana (1993), therapists transverse a number of stages in becoming culturally sensitive practitioners:

(i) unawareness and/or denial, defensiveness and minimization of the importance of cultural issues;

(ii) heightened awareness and acceptance of culture;

(iii) burden of considering culture and doing therapy and adapting behavior and thinking accordingly; and

(iv) integration and synthesis of culture into the assessment and intervention process.

Psychologists who engage in culturally informed assessments and interventions consider myriad variables, including the family's group identity, the family's identity as a subsystem, and the identity of each individual member. Also, the clinician focuses on the beliefs, values, and language of each family member and the family unit as a whole (Dana, 1993).

The following section reviews the conceptual issues and assumptions that guide family assessment and intervention in a culturally informed manner (Kaslow et al., 1995; Odell, Shelling, Young, Hewitt, & L'Abate, 1994). The topics covered include values, help-seeking behavior, and adherence issues. These issues are of paramount importance given the considerable variation in the presenting problem, nature and rate of help-seeking, views of mental health services, and the nature and rate of premature termination among diverse cultural groups (Kazdin, Stolar, & Marciano, 1995). Next, techniques for assessment and intervention appropriate across cultural groups and family life cycle stages are described. This information provides the background for the later section on culturally sensitive assessment and intervention across each specific phase of the family life cycle.

10.08.5.1 Values

To conduct culturally sensitive assessments and interventions, psychologists must be cognizant of their own biases and prejudicial reactions regarding various cultural groups, their personal and culturally based values, and the cultural values of the family with whom they are working. Since families often seek help when their norms are disparate from the cultural values represented by other family members and/or the community, values play a central role in family work.

If the clinician and family differ in their values about the presenting problem, but these differences do not reflect significant conflict or indicate pathological functioning in the family's cultural context, it is recommended that the therapist incorporates the family's cultural framework throughout the assessment and intervention. This recommendation assumes that the therapist's and family's values are equally valid. Another possible stance is for the therapist and family to acknowledge their differences of opinion and the acceptability of such differences.

If during evaluation or treatment the practitioner discovers that the family's value system leads them to engage in behaviors considered illegal or unethical in the relevant jurisdiction, the clinician must inform the family and take appropriate steps. Harsh physical punishment of children may be normative in certain cultural groups (e.g., Jamaican families), but such behavior is illegal in the US and thus requires that the therapist report the family to the appropriate authorities (e.g., Gibbs & Huang, 1989). In such instances, the therapist needs to work with the family to help them modify their disciplinary practices in accordance with the law. This needs to be done sensitively, without the therapist invalidating the family's cultural practices.

If the behaviors associated with the family's values are contrary to normative practices for the relevant cultural groups in the current ecological context, but the behaviors are not illegal or unethical, the therapy helps the family modify these values and associated behaviors in an effort to foster healthy development and enhance family functioning. To accomplish this goal, the psychologist must help the family examine the cultural origins of these values, ascertain the positive and negative consequences of engaging in behaviors associated with the values, and learn culturally sanctioned alternative practices. It is helpful if the therapist underscores that the family's cultural values may lead to problems in some contexts and be beneficial in others, and that values developed in one historical context may be maladaptive in a new sociocultural context.

When there are conflicts among family members regarding key values, and the therapist shares the values of only some family members, the therapeutic process becomes complicated. Often, the therapist will be most aligned with those family members who have the most similar acculturation status and/or cultural identity to the therapist. It is common for the therapist working with immigrant or refugee families to conflict with the grandparent generation regarding gender roles in the family system and to support the gender roles desired by the adult children who are likely to have acculturated more rapidly.

Classes among family members and/or generations regarding values are a necessary consideration in the assessment and treatment of minority families (Gibbs & Huang, 1989). Failure to attend to the processes of value orientation and value clashes may result in an overemphasis on intrapsychic or family dynamics, and may limit the treatment efficacy. Thus, an ecologically oriented approach, that attends to contextual issues as these influence the values of each family member, must be incorporated in all assessment and intervention endeavors with families from nondominant cultural groups.

10.08.5.2 Help-seeking Patterns

The family's cultural background often influences the type of help the family seeks for a loved one's problems. Many families from minority cultures prefer to receive help from both trained and not formally trained members of their own community. This preference often results in their being less likely to seek medical or mental health services from qualified professionals, given the relative dearth of well-trained minority providers and culturally sensitive service programs. For example, African-American and Latino families are less likely than Euroamerican families to seek support from agencies and professionals in the initial stages of responding to the a family members' difficulties. (McMiller & Weisz, 1996). Different help-seeking patterns appear to reflect more negative views toward professionals by majority families and the increased availability of alternative social networks (e.g., extended family, clergy) in many minority communities as compared to Euroamerican communities. In a related vein, in many minority communities there is considerable stigma associated with revealing personal and/or cultural values and beliefs to someone outside the family unit, particularly to someone outside the culture.

These differential help-seeking patterns must be considered when forming an alliance with minority group families, assessing the index person's and family's problems, and designing, implementing, and evaluating culturally based community outreach programs (McMiller & Weisz, 1996). Specifically, during the initial stages of the evaluation process with African-American and Latino families, the psychologist must address the family's concerns and reluctance about mental health services and educate the family about mental health services. Only by doing so will the clinician earn the family's trust sufficiently to conduct a thorough, accurate assessment (McMiller & Weisz, 1996).

10.08.5.3 Adherence

Given the data on minority families' reluctance to use professional mental health services and the high rates of attrition among culturally diverse groups, it behoves the therapist to become familiar with risk factors associated with dropping out of treatment. While several factors of premature termination (e.g., parental stress, antisocial behavior of the parent and/or child, adverse parenting styles, low SES) are common among most families, specific differences among ethnic groups in the rate and nature of intervention drop out have been found (Kazdin et al., 1995). As an example, African-American families leave treatment more frequently and earlier than their Caucasian counterparts (Kazdin et al.). Predictors of African-American drop out also differed, with compliance often predicated on child academic functioning.

There are a number of techniques that may reduce attrition rates, however, many of these strategies are ineffective with families from culturally diverse groups (Paniagua, 1994). To effectively prevent premature termination of family therapy, the psychologist's interventions must attend to the family's culture and the impact of the family's culture on attrition. Separate guidelines have been proposed to prevent attrition in African-American, Latino, Asian-American, and Native American families. For example, when working with African-American families, it is suggested that racial differences be addressed, mental health problems not be linked to parental behaviors, medication not be recommended as the initial treatment of choice, and the psychologist "not give the impression that he or she is the protector of the race when discussing racial issues" (Paniagua, 1994, p. 93). To minimize the likelihood of attrition when working with Native American families, the psychologist should emphasize listening and collaboration rather than talking and authority, de-emphasize the issue of time, discuss the administration rather than the control of the problem, and avoid personalism.

10.08.5.4 Culturally Sensitive Techniques

Cultural issues influence the therapeutic stance, the development of the working alliance, and the assessment and intervention process (Sue & Zane, 1987). In addition to attending to cultural factors as these influence family dynamics, the psychologist conducting the family assessment and intervention must consider the unique aspects of each family to distinguish culturally based and idiosyncratic family relational patterns (Goldenberg & Goldenberg, 1994).

10.08.5.4.1 Stance of the therapist

Nonspecific therapist characteristics associated with the conduct of competent assessments and interventions (e.g., empathy, warmth, genuineness, acceptance, respect) are necessary but not sufficient to assess adequately cultural issues (Sue & Sue, 1990). Culturally informed clinicians must also: (i) be knowledgeable about their own cultural background and its influence on clinical practice; (ii) comprehend the impact of culture on the lives of the families with whom they work; (iii) be cognizant of the variability of normative family structures and functioning across and within cultural groups to avoid stereotyping; (iv) be attuned to subtle cultural variations in language, nonverbal behavior, and expressions of distress; (v) adjust their style to be compatible with the family's culturally influenced patterns of relating to one another and to outsiders in authority; (vi) modify the therapeutic strategies used to facilitate behavior change in a manner consistent with the family's cultural values; (vii) recognize and use the family's natural support systems in the community; and (viii) be willing to incorporate other culturally acceptable clinicians from multiple disciplines into the assessment and intervention process (London & Devore, 1988; Odell et al., 1994; Sue & Sue, 1990).

The American Psychological Association's guidelines for ethical practice for the provision of psychological services to ethnic, linguistic, and culturally diverse populations (American Psychological Association, 1993) maintain that it is only ethical to assess and intervene with individuals and families from the nondominant culture if the professional is trained in multicultural assessment and intervention. According to these guidelines, the therapist's stance must reflect their recognition of cultural diversity, understanding of the role that culture, ethnicity, and race play in psychosocial development, awareness that economic and political factors influence psychosocial development, and knowledge of the impact of culture, gender, and sexual orientation on behavior. Psychologists working with culturally diverse families must communicate a value on helping families understand, maintain, and resolve their own sociocultural identification (American Psychological Association, 1993).

10.08.5.4.2 Development of the working alliance

Joining, a key component of forming a working alliance between the therapist, each family member, and the family system as a whole, is crucial to the conduct of effective assessment and intervention. This process is facilitated when the therapist's space (office, agency) includes culturally familiar objects and symbols to the family (Vasquez-Nuttall, Avila-Vivas, & Morales-Barreto, 1984). Through the joining process, a trusting relationship is forged between the therapist and the family.

Through the process of joining, the clinician becomes cognizant of the culturally influenced rules, roles, structure, communication and problem-solving patterns, and traditional sources of help that impact on family interactional processes. To enhance the working alliance with a family from a cultural group unfamiliar to the psychologist, or when the family engages in culturally driven family patterns to which the therapist is unaccustomed, it is suggested that the therapist assume a position of humility coupled with benign curiosity (Odell et al., 1994). Such a stance can be assumed most readily if the clinician requests the family's aid in learning relevant cultural norms and patterns. If the family is unable to provide adequate information to insure an empathic awareness of the family's interactions within their cultural context, the therapist must glean such data from others knowledgeable about the particular family's culture.

Credibility is associated with effective joining (Sue & Zane, 1987). Professionals' credibility, a reflection of their status, may be determined by the position assigned by the family (ascribed credibility) or by their competencies (achieved credibility). When therapists exhibit culturally consistent assessment and intervention strategies, they may be considered credible by a given family system. The lack of ascribed credibility of formalized mental health services may account for low utilization rates among certain cultural groups, and the lack of achieved credibility may contribute to explanations of nonadherence.

10.08.5.4.3 Assessment process

Once psychologists have begun to form a working alliance with the family, the assessment may begin. The psychologist who gathers family assessment data in a culturally sensitive manner communicates a respect of the cultural diversity of families (Hanson, Lynch, & Wayman, 1990). In conducting a family evaluation, the ways in which the family's culture impacts on their symptom presentation and interactional patterns must be considered (Zayas et al., 1996). Evaluators must be mindful of the ways in which their own culture may affect the assessment process. In addition to cultural biases influencing the assessment of intellectual, psychiatric, and behavioral functioning, these biases may impact upon the evaluator's perception and evaluation of family functioning (Canino & Spurlock, 1994; Dana, 1993). Given the growing literature on cultural considerations in assessing psychological symptoms and psychiatric conditions in people across the life span (e.g., Alarcon, 1995; American Psychiatric Association, 1994a; Canino & Spurlock, 1994; Gaw, 1993), and in diagnosing psychiatric disorders in the recent *Diagnostic and statistical manual of mental disorders* (*4th ed.; DSM-IV*) (American Psychiatric Association, 1994b), our comments on assessment only address considerations pertinent to families.

During the evaluation, the clinician must gather information about the ethnic, racial, and religious background of the family and data about the historical and current political, social, and economic conditions of the family's cultural group. The construct of a genogram may be useful (McGoldrick & Gerson, 1985). Genograms, which graphically provide personal data regarding individual members (e.g., ethnicity, religion, current residence), and information about family structure, relationships, and patterns, enable the psychologist and family to develop systemic hypotheses about family functioning and its connection to the larger family and sociocultural context.

Some therapists question the appropriate timing and nature of genogram construction with families from certain cultural groups. Boyd-Franklin (1989) asserted that with African-American families, genogram work may be most meaningful later in the treatment after trust has been established, rather than during the evaluation. Similarly, Odell et al. (1994) have noted that since constructing a genogram may elicit painful memories for many immigrant and refugee families, this technique may sabotage the development of the initial working alliance and thus may best be introduced during the intervention. For many families who have lost a loved one due to sociopolitical conditions (e.g., Holocaust survivors, Bosnian refugees), projective genogramming elicits effectively distressing) yet important information and feelings (F. W. Kaslow, 1995). Thus, while genograms may be an important assessment tool with families from most cultural groups, the family's cultural background should inform the timing of this task.

An assessment strategy related to the genogram, but devised for work with diverse family groups, is the culturagram (Congress, 1994). With the culturagram, psychologists glean information from the family regarding: reasons for immigration, length of time in the community, citizenship status, language spoken at home and in the community, health beliefs, major holidays, and values on family, education, work, gender roles, religion, and money. Constructing a culturagram enables clinicians to ascertain the effects of culture on the family system and to individualize ethnically similar families. As a result, clinicians are more culturally empathic and more able to empower the families with whom they work.

Enumerable tools and methodologies have been devised to assess multiple aspects of family functioning. These include self-report scales, micro- and macroanalytic coding schemas to code interactional patterns, and projective techniques (for review, see Fredman & Sherman, 1987; Jacob, 1987; L'Abate & Bagarozzi, 1993). Unfortunately, few of these measures or coding schemas have been developed, normed, or empirically tested for specific cultural groups, and thus their utility across groups remains questionable (Dana, 1993). In addition, generalizations about findings gleaned from using standard assessment protocols must be limited and made with caution. Given the scarcity of culture-free assessment strategies, the assessor must take the family's cultural context into account when interpreting findings from assessment protocols not standardized with the family's cultural group. This requires including an evaluation of culturally valid constructs (e.g., racial or ethnic identity, level of acculturation, belief systems, culture-specific syndromes) (Lasry & Sayegh, 1992; Phinney, 1990b). For example, the psychologist may use Gushue's (1993) recent adaptation of Parham and Helms (1985b) Black and White interaction model for assessing and working with families. This assessment strategy enables the clinician to incorporate cultural identity data in making an initial family assessment.

The final phase of the assessment process is problem definition and goal-setting. Cultural variables that may impact this phase include the role of authority, preferred decision-making strategies, the view of psychological problems and potential solutions, and culturally based values. Thus, with Latino families, the goals should address immediate and concrete concerns (Ho, 1987). These families prefer to focus on goals that affect family subsystems, particularly the parent–child subsystem, rather than individual family members or the marital dyad. When working with many Asian-American families, it is recommended that the goals be well-defined, objective, and address practical matters. Asian-American families prefer the therapist to be confident and active in the goal-setting process, while simultaneously communicating respect for the family (Ho, 1987). With Native American families, goal-setting should be collaborative. Given the value on interdependence in the Native American community, all relevant nuclear and extended family members should be included in setting intervention goals (Ho, 1987). When setting goals with low-income, African-American families, a mutual process should be assumed, with a focus on survival needs and the incorporation of an ecostructural framework (Ho, 1987).

In addition to setting goals specific to the given family, many culturally competent practitioners recommend that cultural intentionality and bicultural competence be defined as therapeutic goals for families across cultural groups who seek mental health services (Boyd-Franklin, 1989; Ivey, Ivey, & Simek-Morgah, 1993; Szapocznik et al., 1984). Cultural intentionality refers to the ability to communicate competently with others within the cultural group and with individuals from multiple cultural backgrounds. Bicultural competence connotes the simultaneous processes of accommodating to the host culture and retaining aspects of the culture of origin.

10.08.5.4.4 Intervention process

Throughout the intervention phase, the psychologist is a cultural interpreter or culture broker, helping the family recognize and resolve conflicts between the demands and values of the minority and majority cultures (Canino & Spurlock, 1994; Glordano & Giordano, 1995). Issues that may require interpretation include: kinship relations, respect for roles and family hierarchy, personalization in the therapeutic relationship, and the need to involve additional people from the family's community (e.g., clergy). To assume the role of cultural interpreter, the therapist may incorporate Boyntons (1987) ESCAPE model, according to which the therapist, as cultural interpreter, must: Engage the family within its context; be sensitive to the family's culture; communicate an awareness of the culture's potential and positives; and know the environment in which culture clashes are likely to occur.

When intervening with families from either the same or a different cultural background as themselves, therapists must be mindful that interpersonal behaviors may be interpreted differently. For example, whereas listening may be experienced by some as reflecting empathy and concern, families from certain cultural backgrounds are likely to perceive listening as not caring. Thus, it is crucial that the family and the therapist discuss their differing perceptions and experiences throughout the course of the family work. Awareness of these differences should inform subsequent interventions.

A major focus of family interventions is the introduction and utilization of adaptive problem-resolution tactics. Problem-resolution strategies are most effective when system and cultural variables are taken into account. Since culturally specific intervention approaches for problem resolution have been offered to families from multiple cultural groups (for detailed discussions, see Boyd-Franklin, 1989; Falicov, 1983; Ho, 1987; McGoldrick et al., 1982, 1996; Paniagua, 1994), this section reviews, briefly, techniques for problem resolution with families from the most often encountered cultural groups in the US.

The optimal family intervention approach for African-American families, regardless of SES, is time-limited, problem-focused, and present-oriented with all significant family members and members of the kinship and caregiving networks included (Boyd-Franklin, 1989; Ho, 1987; McGoldrick et al., 1996). The psychologist, who needs to assume an active stance, should inform the family about the goals, process, structure, and limits of family therapy, and serve as a role model, educator, and advocate. Effective family therapy underscores and utilizes family strengths and resources to empower the family to solve its own problems, exert more control over the environment, cope constructively with racism and oppression, and extricate itself from the victim system (Boyd-Franklin, 1989; Pinderhughes, 1989). Psychoeducation and skills training, role clarification, and boundary setting are useful strategies. The unique power and value conflicts of middle-class African-American families may require additional treatment modifications, such as the therapist relabeling behaviors that appear maladaptive as manifestations of the family's determination to negotiate effectively between cultures with competing demands (Corner-Edwards & Spurlock, 1988).

When intervening with Asian-American families, the psychologist must respect key values in the community (e.g., the central role of family solidarity, the emphasis on connectedness and integration rather than separation and individuation) and must be sensitive to the families' culturally sanctioned reluctance to expose conflict to outsiders because it reflects disloyalty (Gaw, 1993; Tamura & Lau, 1992). Although these family values and patterns of relating characterize Asian-American families from multiple countries of origin, effective interventions must take into account the family's specific ethnic background and level of acculturation (Gaw, 1993). When working with Asian-American families, it is suggested that the intervention approach be relatively formal and structured, and focus on pragmatic concerns rather than underlying affects. The therapist must balance showing respect for parental authority and the family's culture of origin with the assumption of an authoritative stance. Family members should be encouraged to communicate verbally and the value of democratic communication should be highlighted (Gaw, 1993). Since many Asian-American families prefer negotiation of differences rather than direct conflict, it may help to hold separate sessions with individual members prior to addressing family conflict directly (Berg & Jaya, 1993; Ho, 1987).

The mental health services literature on work with Latino populations has focused on improving the accessibility of mental health services, choosing interventions according to the cultural characteristics of Latinos, and modifying traditional intervention approaches based on a familiarity and evaluation of ethnic characteristics or creatively deriving the intervention program in response to the cultural milieu (Rogler, Malgady, Constantino, & Blumenthal, 1987). For example, home visits may be an essential component of family interventions with Latino families (Ho, 1987). In addition, or when home visits are not possible, the psychologist should become acquainted with extended family members and assimilated into the family's power hierarchy in order to exert the influence necessary to help the family change. The therapist must appreciate the positive aspects of roles within Latino families that may have negative connotations within middle class, Anglo American culture (e.g., machismo in the father figure). It is essential that the therapist function as cultural translator, mediator, and role model to facilitate the family's active engagement in the community.

It is often the case that Native American families' explanation of a family problem (e.g., marital discord, parent–child conflict) is that the problem is a response to the family's inability to provide for essential needs (e.g., food, shelter) (Ho, 1987). As such, the early stages of a family intervention must address these basic needs and the psychologist should provide advice and information to enhance the family's capacity to fulfill these basic needs. Such a stance communicates the therapist's sensitivity to the family's basic needs and willingness to provide immediate services (Panigua, 1994). Such a stance is likely to foster the development of a working alliance with the family. When undertaking the problem-resolution phase of family interventions with Native American families, the therapist should emphasize group decision-making by including nuclear and extended family members and key community members (medicine man or woman, tribal leaders) (Ho, 1987), and must be mindful of this culture's holistic, interdependent world view and the fact that relationships with kin are of prime significance (Ho, 1987). Suggestions must be delivered to the family in a thoughtful, pragmatic, and calm manner (Ho, 1987; Panigua, 1994). One of the major goals of family interventions with Native American families is the promotion of interdependence. This can be achieved most readily when the therapist functions as a role model, committed to engaging all relevant parties in problem assessment and resolution. Involving all family members in the intervention process typically entails allowing the family to determine the issues to be addressed (Ho, 1987; Panigua, 1994). The use of structural and strategic interventions such as reframing and relabeling are likely to be particularly effective. Strategies experienced by the family as indicating the family's causal role in the identified patient's symptoms are likely to meet with considerable opposition.

Given the consistency between Jewish family values (family membership and loyalty, egalitarianism with regard to roles, intellectual achievement, verbal expression of feelings) and those of psychotherapy, Jewish families often appear to engage readily in family therapy (Rosen & Weltman, 1996). However, the culture's pride in suffering and emphasis on insight over behavior change may complicate the family therapy process. Jewish families often seek services when there is a conflict within the system regarding the degree of closeness vs. separateness. When working with these families, the psychologist must communicate to the family a respect for their value on family cohesion, while simultaneously addressing the ways in which closeness in the family system may interfere with the development and maintenance of family members' personal boundaries and age-appropriate separation–individuation.

Irish-Americans tend to minimize problems, have difficulty articulating their feelings, and use humor and alcohol to manage stress (McGoldrick, 1996). These coping strategies are often associated with a reluctance to seek psychological services. When Irish-American families do enter therapy, the psychologist is likely to encounter guilt, resentment, and self-reproach. To intervene effectively with Irish-American families, family clinicians should ask specific questions to elicit all relevant information, maintain a friendly distance, and avoid exploration of painful effects and embarrassing topics (Jalali, 1988; McGoldrick, 1996). Thus, a structured approach, focusing on the presenting problem, is suggested (McGoldrick, 1996).

The family is of prime importance in Italian-American communities. Families in these communities are often reluctant to support members' differentiation. In addition, many Italian-American families attempt to resolve conflicts within the nuclear and extended family. The therapist, therefore, must respect the family as a cohesive unit. Attempts to decrease family closeness by strengthening internal and weakening external boundaries are likely to be resisted and met with the family's refusal to let the therapist enter the family system. This impedes the family work, which can only be effective when the therapist is treated as a member, albeit distant, of the family. Because these families often seek therapy during life cycle transitions associated with separation, it may be difficult for the therapist to refrain from challenging the family's strong emotional ties. Thus, family therapists can help most by coaching individuals to remain close to the family, without becoming engulfed (Giordano & McGoldrick, 1996). Given the Italians' penchant for "histrionics," the therapist must ascertain the actual severity of the problem (Jalali, 1988). Despite their emotionally expressive presentation, they frequently deny family conflicts and secrets until they trust the therapist. It is important, therefore, to provide timely advice regarding the chief complaint, while patiently waiting for the family to examine underlying patterns (Giordano & McGoldrick, 1996).

10.08.5.5 Culturally Sensitive Techniques Across the Life Cycle

10.08.5.5.1 The joining of families: culturally sensitive assessment and intervention

Given the significant cultural variations in the nature of marital relationships among diverse families, an assessment of a couple must include attention to those cultural variables most likely to impact upon the functioning of the dyad. In addition, when conducting an assessment, the psychologist must be mindful of accepted cultural practices, many of which may be unfamiliar to the clinician (e.g., arranged marriages, lack of emphasis placed on love). When assessing newly formed couples, whether or not they are married, the psychologist must examine each person's perspective on the meaning of an intimate partnership (values on intimacy, fidelity, divorce), appropriate gender roles within a couple, and the degree of autonomy vs. interdependence expected by each partner both within the dyad and between the dyad and the extended family. Attention must be paid to each person's notions regarding the degree to which each party has an equal role and value in the partnership and the responsibilities to be assumed by each person. It is also crucial that the psychologist investigates physical abuse in all couples, regardless of their ethnic heritage, as domestic violence is prevalent in all cultures (Levinson, 1989). Finally, given the increasing rates of intermarriage and the fact that intimate relationships between people of different ethnic, religious, and racial backgrounds compound the issues faced by couples from a single ethnic group, similarities and differences between partners based upon their cultural heritage must be examined.

The joining of two families through marriage requires that the two individuals choosing to marry renegotiate with one another about areas of concern. Difficulties in negotiating and compromising on such issues as core values, the role of the family, communication styles, roles and responsibilities, level of intimacy, religion, and politics often lead couples to seek couples' therapy (Carter & McGoldrick, 1989). When intervening with newly formed couples, the psychologist must first help each partner to feel comfortable in sharing their perceptions, opinions, and wishes regarding various areas of disagreement (e.g., Falicov, 1983). One way in which a nonjudgemental context can be created to facilitate this sharing is by the therapist encouraging both members of the dyad to examine how their views are influenced by their cultural background. During this process of disclosure, it behoves the therapist to validate for the couple the complexities of negotiating different life styles and different cultural stereotypes. The couple also needs help in realizing the importance of integrating their differences, rather than choosing between polar opposites. The process of working through similarities and differences enables the couple to form their own identification as a new family unit, both related to and separate from their respective families of origin.

10.08.5.5.2 Families with young children: culturally sensitive assessment and intervention

In working with culturally diverse families with young children, psychologists need to choose and administer assessment devices and interpret the results from a culturally informed perspective (Canino & Spurlock, 1994). This involves using and producing culturally and developmentally sensitive instruments, techniques, and explanations. Direct observation of children playing, drawing, and interacting with their family can provide metaphors for the child's thoughts and feelings (Grizenko, Sayeghm, & Migneault, 1991). Such metaphors are particularly helpful when language differences complicate the assessment process. The psychologist must guard against any personal biases and shed cultural stereotypes when conducting and interpreting a family assessment of a young child (Grizenko et al., 1991). When evaluating a child's presenting problem, the psychologist must ascertain the meaning of the symptoms in the child's culture and the degree to which the symptoms are viewed as pathological in the child's culture (Grizenko et al.). An assessment of a young child, individually as well as in the context of the family, must consider the cultural context of the child's socialization. For example, deferential behavior may raise questions about the Caucasian child's self-confidence and or a rigid family structure; such behavior is expected in many Latino cultures and thus is most appropriately understood as reflecting the child's efforts to conform to cultural expectations.

When intervening with families with young children, the therapist must consider the family's cultural values on various aspects of childrearing and socialization. For example, when working with many traditional Latino families, it may be inappropriate for the therapist to emphasize the need for the child to be behaviorally and verbally assertive, as such a recommendation may be perceived by the family as undermining of the parents' role (Zayas & Solari, 1994). The therapist must join such family systems by acknowledging, not challenging, the parents' authority. Gradually, this will enable the family to trust the therapist, which in turn will facilitate the family's increased comfort with the expression of a full range of emotions from the children (Zayas & Solari, 1994).

One complicated dilemma that may be encountered when working with families with young children occurs when the child is the only family member fluent in English (Vasquez-Nuttall et al., 1984). For example, for a Cuban-American family in which Spanish is the sole language of the adults, a Spanish speaking therapist may be optimal. If such a therapist is unavailable or if the family prefers an English speaking therapist who can introduce them to the majority culture, a translator may be an option. Unfortunately, however, translations often misrepresent the actual meaning and cultural significance of family expressions and fail to consider implications of nonverbal behaviors. While some families prefer using the young child as an interpreter, this is problematic in families from cultural groups in which respect for parental authority is of utmost importance. In such instances, the child as interpreter may be perceived as being insubordinate, disrespectful, and intruding on adult concerns.

One particular form of intervention that impacts significantly on the family of a young child is that of early interventions (Hanson et al., 1990). Because early interventions focus on the young child with a disability, these programs interact with cultural views and values. Early interventionists often work with the young child with a disability in the family's home. This experience offers an inside view of family rituals, communication, and caretaking patterns, and a perspective on children and childrearing, and family and family roles. In addition, it typically is sanctioned for early interventionists to consult with the family regarding health and mental health care. Such involvements with the family heighten the early interventionist's awareness of different cultural perspectives and the need to respect the family's boundaries while simultaneously imparting the views of the larger culture so that they can receive optimal care for their young child (Hanson et al., 1990).

10.08.5.5.3 Families with adolescents: culturally sensitive assessment and intervention

Defining deviant adolescent behavior in the context of culturally diverse families is complicated. The degree to which rebellion, deviance, and separation–individuation strivings are tolerated varies with different ethnic and cultural groups. For example, the Latino family's awareness that their adolescent daughter is sexually active with her boyfriend may lead to their seeking services for their child. Once the psychologist is familiar with the family's cultural norm that premarital sex among adolescent females is unacceptable, the therapist must assess the adolescent's sexual activity (solely with significant other vs. with multiple partners, safe sex practices) and other potentially maladaptive or self-destructive behaviors. Based on these data, the psychologist can ascertain if the adolescent's behavior is inconsistent with her family's cultural norms but in keeping with the practices of the majority culture, or deviant from the expectations of both cultures. This information should be shared with the family during the assessment.

Among immigrant families from such diverse cultures as Mexico, Nigeria, and Kuwait, adolescents are expected to seek, respect, and obey their parents' guidance (Baptiste, 1993). After living in the US, adolescents from these families often begin to behave in a more individualistic fashion, place a greater value on autonomy than is accepted in their family and culture of origin, and seek advice from nonfamily members. Such actions are perceived by older family members as a reflection of the family's inability to cope effectively with immigration and are experienced as threatening and disrespectful. As a consequence, a rigidification of family rules is common. If such a response fails to reduce the adolescent's autonomy strivings, the family may seek a mental health evaluation. It is incumbent upon the psychologist evaluating such families to bear in mind the conflictual values regarding family hierarchy and dependence vs. independence manifested by members of different generations. Feedback from such an assessment must incorporate the therapist's evaluation of the extent to which the adolescent's behaviors and attitudes are consistent with, or discrepant from, the different cultures in which the adolescent embedded. If based on this feedback, the therapist recommends treatment and the family concurs, the therapist should serve as a family intermediary, whose primary function is to translate cultural behavior in a developmental perspective. In working with some families (e.g., Mexican-Americans), this may entail conducting separate sibling and parent sessions to discuss sexual matters, prior to holding sessions with all family members present (Falicov, 1996). Such an approach demonstrates the therapist's respect for the parents' view that discussion of intimate matters with their children is inappropriate and undermining of their authority (Garcia-Preto, 1996).

There is a burgeoning body of research on racial and ethnic similarities and differences in adolescent substance abuse patterns and risk factors (e.g., Gottfredson & Koper, 1996). Data from these studies have been used to guide the development and implementation of culturally informed preventive intervention programs, such as the State-wide Indian Drug Prevention Program (IDPP) (Bobo, Gilchrist, Cvetkovich, Trimble, & Schinke, 1988). In addition, these data have laid the foundation for treatment outcome studies. For example, an excellent example of culturally sensitive treatment outcome research has been conducted with the families of Cuban-American adolescents with substance abuse problems (e.g., Kurtines & Szapocznik, 1996; Szapocznik et al., 1984; 1986). Bicultural effectiveness training (BET), a brief family intervention model that incorporates culture as a key content area upon which to base family interventions, has been found to be equally effective in improving family functioning as structural family therapy, a standard family approach that is not necessarily culturally based. These findings support the use of a culturally informed family intervention for enhancing the functioning of adolescents with substance abuse problems from families that experience conflicts associated with cultural and generational differences (Szapocznik et al., 1986).

10.08.5.5.4 Launching children, single young adults leaving home, and the middle generation moving on: culturally sensitive assessment and intervention

Given cultural variations in what is considered normative behavior during the launching stage of the family life cycle, it is crucial that the psychologist takes cultural norms into account when ascertaining whether or not family conflicts during this time reflect significant individual or family dysfunction, differences in values across generations about independence v. interdependence, differences in perspectives on healthy behavior during this phase between the family's cultural group and the majority culture, or simply normal family processes. It is useful to learn if previous generations of the family migrated during this phase, as these family of origin experiences would intensify the significance of this period (Carter & McGoldrick, 1989). When assessing the single young adult leaving home, it is recommended that information be gleaned regarding how the person's parents managed this phase of their own development. Questions should be asked about the young adult's internal struggles and conflicts with cultural norms regarding separation and individuation. Also, data should be gathered regarding the individual's view of acceptable gender roles vs. those held by the family and cultural group (e.g., views on remaining single as a woman). Finally, given that this process requires a restructuring of the marital relationship, the parent's marital history and current marital status deserve attention.

It is often recommended that family work with single young adults or with young adults in the launching phase be conducted primarily with the young adult, using a family of origin approach (Bowen, 1978). It is useful for the psychologist to help young adults examine the relation between multigenerational patterns of relating and their own personal development. To facilitate this awareness, the psychologist can "coach" or guide young adults in both the emancipation process and in renegotiating family relationships in order to develop a healthier balance between autonomy and attachment that meets the needs of both the individual and the family, and that takes cultural considerations into account. The coaching process must address cultural norms when focusing on the reduction of extreme enmeshment, radical disengagement, and the active use of triangulation. Specific strategies that may facilitate this process include: detriangling, person-to-person contact, reversals, and reconnecting (Carter & McGoldrick, 1989). These strategies enable the family system to alter patterns of relating that are dysfunctional given the family's current cultural context and foster mature communication among family members.

When couples seek mental health services during this family life cycle stage, culturally sensitive interventions must take at least a three-generation perspective. A multigenerational view allows the psychologist to help the couple deal with their independence from their offspring and their increased commitment to the older generation. In addition, they may need help in dealing with the separation and loss associated with the death of older family members who often serve as culture bearers. The negotiation of new relationships with children and parents is often complicated by the need to renegotiate the marital dyad. Thus, family interventions must address all relevant relationships, with particular attention to the ways in which the cultural values of individuals across the generations impact on the interpersonal difficulties that emerge. Further, for individuals from many cultural groups, this may be the first family life cycle stage in which their own individual development is paramount. Thus, personal difficulties camouflaged by the demands of family life at other stages in the family life cycle may come to the fore, and may require individual attention.

10.08.5.5.5 Families in later life: culturally sensitive assessment and intervention

For psychologists to assess accurately the functioning among culturally diverse families with elderly members, they must first build a therapeutic relationship with all family members. This may be complicated when working with elderly family members, as these individuals from different cultural groups are often suspicious of professionals who they perceive as having little in common with them. They suspect that mental health and health care professionals may have biases against their age, race, sex, and culture.

Hays' (1996) framework for the conduct of culturally responsive assessments underscores the importance of attending to a number of influences that impact on the functioning of diverse older adults. She uses the acronym ADRESSING (albeit incorrectly spelled) to organize the key influences to be taken into account when assessing elderly family members: age, disability, religion, ethnicity, social status, sexual orientation, indigenous heritage, national origin, and gender. Attention to these variables facilitates the building of rapport, enhances the psychologist's understanding of the elderly family member's cultural identity and heritage, and improves the psychologist's awareness of elderly family members' cultural environment, needs, and strengths (Hays, 1996).

When assessing elderly clients in a family context, it behoves the psychologist to use strategies that enhance rapport and communication. Possible techniques that may facilitate positive interactions include sitting closer and speaking clearly, talking to "the good ear," using titles (Mr, Mrs, Dr, etc.) before names, using respectful and culturally sensitive interpreters when language differences exist, being equipped with conveniences for physical disabilities, and having culturally diverse artifacts in the office (Hays, 1996). When gathering a history from elderly family members, the psychologist should be cognizant of the importance of gathering extensive background information and encouraging the older family members to tell their life stories. During the assessment process, it is recommended that some time be spent with all family members present and some time should be devoted to separate meetings with the older people.

During the multigenerational meetings, the psychologist can communicate respect for the elderly by encouraging the eldest family member to speak first and by valuing the input offered by elderly family members (Kim, 1985). When conducting diagnostic interviews with the elderly or administering and interpreting test results, the psychologist must bear in mind the pernicious problem of misdiagnosis of psychiatric disorders in minority older adults (American Psychiatric Association, 1994a). Elderly clients are likely to view testing as humiliating, confusing, too time-consuming; and not relevant to them, and thus may respond with excessive anxiety and a lack of focus and effort to the process. Frequently, the instruments used and the process of assessing intelligence or personality have been created and indexed around younger, Euroamerican cultural values and constructs (Berry, Poortinga, Segall, & Dasen, 1992), and thus may not be valid or reliable for use with diverse older adults. When such instruments or processes are used with elderly ethnic minority individuals, misdiagnoses, confusion, and inappropriate interventions can occur (Hilliard, 1992). The likelihood of misdiagnosis can be minimized if the psychologist takes the time to use observations and interviews, and to look for "intrapersonal strengths related to specific cultural identities as well as interpersonal or social supports related to minority group membership" (Hays, 1996, p. 192).

There is a dearth of information on clinical interventions with the ethnic minority elderly and their families (American Psychiatric Association, 1994a). Little data exist regarding whether or not efficacious interventions for nonminority elderly will be effective for minority elderly. The extant literature underscores the importance of engaging as many extended family members as possible in the care of older family members. Unfortunately, however, for many immigrants, family support may be inadequate or absent and family relations are often marked by significant intergenerational conflicts (Goldstein, 1989). Thus, adequate mental health services must include outreach efforts and the availability of mental health resources in conjunction with medical care (Sakauye, 1989).

Family interventions with diverse older adults should be primarily directive (Paniagua, 1994) and include the provision of education to the patient and family. The psychologist should assume an active stance and incorporate a problem focus. All interventions should communicate respect for the elderly and an appreciation of the elderly's perspective on the role of the therapist. For example, Native American elderly typically view the therapist as akin to the medicine man, whereas Latino elderly often perceive the therapist as a folk healer (Paniagua, 1994). Further, family interventions should be geared toward the acceptance of intergenerational differences and the resolution of associated conflicts.

10.08.6 FUTURE DIRECTIONS

Given the relative dearth of empirical literature from a culturally sensitive perspective regarding family patterns, family-oriented assessment, and family interventions, it is incumbent upon psychologists to conduct research on all of these topics. One specific area that deserves immediate attention is the delineation, based on research findings, of family patterns among different ethnic groups across the various family life cycle stages. A second research topic that merits further investigation is the development and collection of normative data on culturally sensitive assessment tools (e.g., questionnaires, observational coding schemas, semistructured interviews) for various aspects of family functioning across a multitude of cultural settings. The need for assessment methodologies specific to each given culture requires special attention. A third avenue for research exploration relates to the development, implementation, and evaluation of family interventions designed to be culturally and developmentally sensitive. The conduct of assessment and intervention research must include a focus on the validity, reliability, and efficacy of various techniques based upon the culture of the therapist and the family.

Due to the rapidly changing nature of the family in America and throughout the world, as well as increased levels of multiculturalism, clinicians and researchers must devote efforts to understanding more about families embedded in cultures in transition. This will require more attention to such issues as cultures within cultures (e.g., a lesbian couple in which one member is Latino and her partner is Asian-American), and intergenerational conflicts within families associated with levels of acculturation and/or sense of connection to the culture of origin.

Fortunately, diversity training for clinicians and researchers has advanced significantly. However, it is incumbent upon us as psychologists to insure that cultural considerations be given paramount importance in coursework, supervision, and continuing education programs related to family assessment and intervention. This entails the utilization of a cultural formulation for family case conceptualization, increased and ongoing discussion of the family's cultural context as it impacts on family relationships, and enhanced sensitivity to cultural issues in all aspects of the family assessment and intervention process. Such a shift in our orientation as psychologists and family therapists will result in the delivery of more ethical and effective interventions.

10.08.7 SUMMARY

Since the 1970s, family theory and therapy has come to recognize the significant impact that a family's cultural context has on assessment and intervention across the family life cycle. Increasingly, family theorists and therapists view each individual as apart of a family system embedded in a larger culture, the structure and process of which depends on the family's stage of development. The current chapter focused on a cultural perspective on families across the life cycle and implications for the conduct of culturally sensitive assessment and intervention. While conceptualizations of culture often include ethnicity, race, SES, gender, sexual orientation, country of origin, and religious and political affiliations, we have chosen to concentrate on cultural issues relevant to ethnic groups living in the US. A thorough review of the literature suggests that effective family assessment and intervention is characterized by:

(i) the therapists' acknowledgement of their own cultural origins and the impact of this background on the assessment and intervention process;

(ii) a sensitivity to the family's cultural background;

(iii) an appreciation of the unique interactional patterns, attitudes, feelings, and behaviors of each family;

(iv) a recognition of the family's location in the family life cycle;

(v) validation and strengthening of the ethnic identity of each family member and the family unit as a whole; and

(vi) an incorporation of relevant support systems (Giordano & Giordano, 1995; N.J. Kaslow et al., 1995).

To provide culturally sensitive family assessments and interventions, the clinician must strive to overcome cultural barriers while demonstrating respect for the cultural identity and integrity of each family member, the family system, and the therapist (Tyler, Brome, & Williams, 1991). As the dramatic increase in cultural diversity in the US is expected to continue into the twenty-first century, there is a strong need for empirically validated multicultural and culture specific family theories and therapies. However, until such theories and therapies are available, psychologists are charged with helping families negotiate the reconnection with their cultural heritage in order to establish or preserve a sense of belonging and cultural identity, while concurrently adapting to an evolving cultural context (N.J. Kaslow et al., 1995).

10.08.8 REFERENCES

Aboud, F. (1987). The development of ethnic self-identification and attitudes. In J. Phinney & M. Rotheram (Eds.), *Children's ethnic socialization: Pluralism and development* (pp. 32–55). Newbury Park, CA: Sage.

Alarcon, R. D. (Ed.) (1995). Cultural psychiatry (special issue). *The Psychiatric Clinics of North America, 18*(3).

American Psychiatric Association (1994a). *Ethnic minority elderly: A task force report of the American Psychiatric Association.* Washington, DC: Author.

American Psychiatric Association (1994b). *Diagnostic and* *statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychological Association (1993). Guidelines for providers of psychological services to ethnic, linguistic, and culturally diverse populations. *American Psychologist, 48,* 45–48.

Baker, F. M. (1994). Issues in the psychiatric care of African American elders. In American Psychiatric Association (Ed.), *Ethnic minority elderly: A task force report of the American Psychiatric Association* (pp. 21–62). Washington, DC: American Psychiatric Association.

Banks, W. (1976). White preference in Blacks: A paradigm in search of a phenomenon. *Psychological Bulletin, 83,* 1179–1186.

Baptiste, D. A., Jr. (1993). Immigrant families, adolescents and acculturation: Insight for therapists. *Marriage and Family Review, 19,* 341–363.

Berg, I. K., & Jaya, A. (1993). Different and same: Family therapy with Asian American families. *Journal of Marital and Family Therapy, 19,* 31–38.

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (1992). *Cross-cultural psychology: Research and applications.* New York: Cambridge University Press.

Billingsley, A. (1992). *Climbing Jacob's ladder: The enduring legacy of African American families.* New York: Simon and Schuster.

Babo, J. K., Gilchrist, L. D., Cvetkovich, G. T., Trimble, J. E., & Schinke, S. P. (1988). Cross-cultural service delivery to minority communities. *Journal of Community Psychology, 16,* 263–272.

Bowen, M. (1978). *Family therapy in clinical practice.* New York: Jason Aronson.

Boyd-Franklin, N. (1989). *Black families in therapy: A multi-systems approach.* New York: Guilford Press.

Boynton, G. (1987). Cross-cultural family therapy: The ESCAPE model. *American Journal of Family Therapy, 15,* 123–130.

Bretherton, I., & Waters, E. (Eds.) (1985). Growing points of attachment theory and research. *Monographs of the Society for Research in Child Development, 50 (1–2, Serial No. 209).*

Bullrich, S. (1989). The process of immigration. In L. Combrinck-Graham (Ed.), *Children in family contexts: Perspectives on treatment* (pp. 482–501). New York: Guilford Press.

Canino, I., & Spurlock, J. (1994). *Culturally diverse children and adolescents: Assessment, diagnosis and treatment.* New York: Guilford Press.

Carter, E., & McGoldrick, M. (1989). *The changing family life cycle: A framework for family therapy* (2nd ed.). Boston: Allyn and Bacon.

Clark, K. B., & Clark, M. P. (1940). Skin color as a factor in racial identification of Negro preschool children. *Journal of Social Psychology, 11,* 156–169.

Comas-Diaz, L., & Greene, B. (Eds.) (1994). *Women of color: integrating ethnic and gender identities in psychotherapy.* New York: Guilford Press.

Coner-Edwards, A. F., & Spurlock, J. (Eds.) (1988). *Black families in crisis: The middle class.* New York: Brunner Mazel.

Congress, E. P. (1994). The use of culturagrams to assess and empower culturally diverse families. *Families in Society, 75,* 531–540.

Dana, R. H. (1993). *Multicultural assessment perspectives for professional psychology.* Boston: Allyn and Bacon.

Lavis, J. (1988). Mazel tov: The Bar Mitzvah as a multigenerational ritual of change and continuity. In E. Imber-Black, J. Roberts, & R. Whiting (Eds.), *Rituals in families and family therapy* (pp. 177–208). New York: Norton.

Dechesnay, M. (1986). Jamaican family structure: The paradox of normalcy. *Family Process, 25,* 293–300.

Emery, E. (1996). Amish families. In M. McGoldrick, J. Giordano, & J. Pearce (Eds.), *Ethnicity and family therapy* (2nd ed., pp. 442–450). New York: Guilford Press.

Erikson, E. H. (1958). *Identity: Youth and crisis.* New York: Norton.

Falicov, C. J. (Ed.) (1983). *Cultural perspectives in family therapy.* Rockville, MD: Aspen.

Falicov, C. J. (1996). Mexican families. In M. McGoldrick, J. Giordano, & J. Pearce (Eds.), *Ethnicity and family therapy* (2nd ed., pp. 169–182). New York: Guilford Press.

Fisek, G. O. (1991). A cross-cultural examination of proximity and hierarchy as dimensions of family structure. *Family Process, 30,* 121–133.

Foulks, E. F. (1991). Transcultural psychiatry and normal behavior. In D. Offer & M. Sabshin (Eds.), *The diversity of normal behavior* (pp. 207–238). New York: Basic Books.

Franklin, D. L. (1987). Black adolescent pregnancy: A literature review. *Child and Youth Services, 9,* 15–39.

Fredman, N., & Sherman, R. (1987). *Handbook of measurements for marriage and family therapy.* New York: Brunner Mazel.

Garcia-Preto, N. (1996). Latino families: An overview. In M. McGoldrick, J. Giordano, & J. Pearce (Eds.), *Ethnicity and family therapy* (2nd ed., pp. 141–154). New York: Guilford Press.

Gaw, A. C. (Ed.) (1993). *Culture, ethnicity, and mental illness.* Washington, DC: American Psychiatric Press.

Gibbs, J. T., & Huang, L. N. (1989). *Children of color: Psychological interventions with minority youth.* San Francisco: Jossey Bass.

Giordano, J., & Giordano, M. A. (1995). Ethnic dimensions in family therapy. In R. Mikesell, D. Lusterman, & S. McDaniel (Eds.), *Integrating family therapy: Handbook of family psychology and systems theory* (pp. 347–356). Washington, DC: American Psychological Association.

Giordano, J., & McGoldrick, M. (1996). Italian families. In M. McGoldrick, J. Giordani, & J. Pearce (Eds.), *Ethnicity and family therapy* (2nd ed., pp. 567–582). New York: Guilford Press.

Goldenberg, H., & Goldenberg, I. (1994). *Counseling today's families* (2nd ed.). Pacific Grove, CA: Brooks/Cole.

Goldstein, M. Z. (1989). *Family involvement in treatment of the frail elderly.* Washington, DC: American Psychiatric Press.

Gottfredson, D. C. & Koper, C. S. (1996). Race and sex differences in the prediction of drug use. *Journal of Consulting and Clinical Psychology, 64,* 305–313.

Grizenko, N., Sayegh, L., & Migneault, P. (Eds.) (1991). *Transcultural issues in child psychiatry.* Verdun, PQ: Editions Douglas.

Gushue, G. (1993). Cultural-identity development and family assessment: An interaction model. *Counseling Psychologist, 21,* 487–513.

Gushue, G. V., & Sciarra, D. T. (1995). Culture and families: A multidimensional approach. In J. G. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander (Eds.), *Handbook of multicultural counseling.* Thousand Oaks, CA: Sage.

Hanson, M. J., Lynch, E. W., & Wayman, K. I. (1990). Honoring the cultural diversity of families when gathering data. *Topics in Early Childhood Special Education, 10,* 112–131.

Harrison, A. O., Wilson, M. N., Pine, C. J., Chan, S. Q., & Buriel, R. (1990). Family ecologies of ethnic minority children. *Child Development, 61,* 347–362.

Harwood, R. L. (1992). The influence of culturally derived values on Anglo and Puerto Rican mothers' perceptions of attachment behavior. *Child Development, 63,* 822–839.

Harwood, R. L., & Miller, J. G. (1991). Perceptions of attachment behavior: A comparison of Anglo and Puerto Rican mothers. *Merrill-Palmer Quarterly, 37,* 583–599.

Harwood, R. L., Miller, J. G., & Irizarry, N. L. (1995). *Culture and attachment: Perceptions of the child in context.* New York: Guilford Press.

Hays, P. A. (1996). Culturally responsive assessment with diverse older clients. *Professional Psychology: Research and Practice, 27,* 188–193.

Hillard, J. R. (1982). Diagnosis and treatment of the rootwork victim. *Psychiatric Annals, 12,* 705–714.

Hilliard, A. S. (1992). IQ and the courts: Larry P. v. Wilson Riles and PASE v. Hannon. In A. K. H. Burlew, W. C. Banks, H. P. McAdoo, & D. A. Azibo (Eds.), *African American psychology* (pp. 199–217). Newbury Park, CA: Sage.

Ho, M. K. (1987). *Family therapy with ethnic minorities.* Newbury Park, CA: Sage.

Hsu, J., Tseng, W. S., Ashton, G., McDermott, J. F., & Char, W. (1985). Family interaction patterns among Japanese-American and Caucasian families in Hawaii. *American Journal of Psychiatry, 142,* 577–581.

IJzendoorn, M. H. Van (Ed.) (1990). Cross-cultural validity of attachment theory (special issue). *Human Development, 33.*

Imber-Black, E., Roberts, J., & Whiting, R. (Eds) (1988). *Rituals in families and family therapy.* New York: Norton.

Ivey, A. E., Ivey, M. B., & Simek-Morgan, L. (1993). *Counseling and psychotherapy: A multicultural perspective* (3rd ed.). Boston: Allyn and Bacon.

Jacob, T. (Ed.) (1987). *Family interaction and psychopathology: Theories, methods, and findings.* New York: Plenum Press.

Jacobs, J. L. (1994). Gender, race, class, and the trend toward early motherhood: A feminist analysis of teen mothers in contemporary society. *Journal of Contemporary Ethnography, 22,* 442–462.

Jalali, B. (1988). Ethnicity, cultural adjustment and behavior: Implications for family therapy. In L. Comas-Diaz & E. Griffith (Eds.), *Clinical guidelines in cross-cultural mental health* (pp. 9–32). New York: Wiley.

Jimenez, R. G., & Figueiredo, J. M. de (1994). Issues in the psychiatric care of Hispanic American elders. In American Psychiatric Association (Ed.), *Ethnic minority elderly: A task force report of the American Psychiatric Association* (pp. 63–90). Washington, DC: American Psychiatric Association.

Karnow, S., & Yoshihara, N. (1992). *Asian Americans in transition.* New York: The Asia Society.

Kaslow, F. W. (1993). The divorce ceremony: A healing strategy. In T. Nelson & T. Trepper (Eds.), *101 favorite family therapy interventions* (pp. 341–345). New York: Haworth Press.

Kaslow, F. W. (1995). *Projective genogramming.* Sarasota, FL: Professional Resource Press.

Kaslow, N. J., Celano, M., & Dreelin, E. D. (1995). A cultural perspective on family theory and therapy. *Psychiatric Clinics of North America, 18,* 621–633.

Kazdin, A. E., Stolar, M. J., & Marciano, P. L. (1995). Risk factors for dropping out of treatment among White and Black families. *Journal of Family Psychology, 9,* 402–417.

Kim, S. C. (1985). Family therapy for Asian Americans: A strategic structural framework. *Psychotherapy, 22,* 342–348.

Kurtines, W. M., & Szapocznik, J. (1996). Family interaction patterns: Structural family therapy in contexts of cultural diversity. In E. D. Hibbs & P. S. Jensen (Eds.), *Psychosocial treatments for child and adolescent disorders: Empirically based strategies for clinical practice* (pp. 671–697). Washington, DC: American Psychological Association.

L'Abate, L., & Bagarozzi, D. A. (1993). *Sourcebook of marriage and family evaluation.* New York: Brunner Mazel.

Lasry, J. C., & Sayegh, L. (1992). Developing an acculturation scale: A bidimensional model. In N. Grizenko, L. Sayegh, & P. Migneault (Eds.), *Transcultural issues in child psychiatry* (pp. 67–86). Verdun, PQ: Editions Douglas.

Lee, E. (1996). Chinese families. In M. McGoldrick, J. Giordano, & J. Pearce (Eds.), *Ethnicity and family therapy* (2nd ed., pp. 249–267). New York: Guilford Press.

Levinson, D. (1989). *Family violence in cross cultural perspective.* Newbury Park, CA: Sage.

Lindblad-Goldberg, M. (1989). Successful minority single-parent families. In L. Combrinck-Graham (Ed.), *Children in family contexts: Perspectives in treatment* (pp. 116–134). New York: Guilford Press.

London, H., & Devore, W. (1988). Layers of understanding: Counseling ethnic minority families. *Family Relations, 37,* 310–314.

Mahon, J. (1976). Black and White children's racial identification and preference. *Journal of Black Psychology, 3,* 47–58.

Matsui, W. T. (1996). Japanese families. In M. McGoldrick, J. Giordano, & J. Pearce (Eds.), *Ethnicity and family therapy* (2nd ed., pp. 268–280). New York: Guilford Press.

Matsumoto, D. (1994). *People: Psychology from a cultural perspective.* Pacific Grove, CA.: Brooks/Cole.

Mayer, E., & Kosmin, B. (1994). *National survey of religious identification.* New York: University of New York.

McGoldrick, M. (1996). Irish families. In M. McGoldrick, J. Giordano, & J. Pearce (Eds.), *Ethnicity and family therapy* (2nd ed., pp. 544–566). New York: Guilford Press.

McGoldrick, M., & Gerson, R. (1985). *Genograms in family assessment.* New York: Norton.

McGoldrick, M., Hines, P., Lee, E., & Preto, N. G. (1986). Mourning rituals: How culture shapes the experience of loss. *Family Therapy Networker, November/December,* 28–36.

McGoldrick, M., Pearce, J. K., & Giordano, J. (Eds.) (1982). *Ethnicity and family therapy.* New York: Guilford Press.

McGoldrick, M., Pearce, J. K., & Giordano, J (Eds.) (1996). *Ethnicity and family therapy* (2nd ed.). New York: Guilford Press.

McMiller, W. P., & Weisz, J. R. (1996). Help-seeking preceding mental health clinic intake among African-American, Latino, and Caucasian youths. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 1086–1094.

Mickelson, R. (1990). The attitude-achievement paradox among Black adolescents. *Sociology of Education, 63,* 44–61.

Myerhoff, B. (1979). *Number our days: Culture and community among elderly Jews in an American ghetto.* New York: Meridian.

Odell, M., Shelling, G., Young, K. S., Hewitt, D. H., & L'Abate, L. (1994). The skills of the marriage and family therapist in straddling multicultural issues. *American Journal of Family Therapy, 22,* 145–155.

Panigua, F. A. (1994). *Assessing and treating culturally diverse clients.* Thousand Oaks, CA: Sage.

Papajohn, J., & Spiegel, J. P. (1975). *Transactions in families: A modern approach for resolving cultural and generational conflict.* San Francisco: Jossey-Bass.

Parham, T. A., & Helms, J. E. (1981). The influence of Black students' racial identity attitudes on preferences for counselor's race. *Journal of Counseling Psychology,* *28*, 250–257.

Parham, T. A., & Helms, J. E. (1985a). Attitudes of racial identity and self-esteem of Black students: An exploratory investigation. *Journal of College Student Personnel, 26*, 143–147.

Parham, T. A., & Helms, J. E. (1985b). Relation of racial identity attitudes to self-actualization and affective states of Black students. *Journal of Counseling Psychology, 32*, 431–440.

Phinney, J. S. (1989). Stages of ethnic identity in minority group adolescents. *Journal of Early Adolescence, 9*, 34–49.

Phinney, J. S. (1990a). Ethnic identity in adolescents and adults: Review of research. *Psychological Bulletin, 108*, 499–514.

Phinney, J. S. (1990b). The multi-group ethnic identity measure. A new scale for use with diverse groups. *Journal of Adolescent Research, 7,* 156–176.

Phinney, J. S., & Alipuria, L. (1990). Ethnic identity in older adolescents from four ethnic groups. *Journal of Adolescence, 13,* 171–183.

Pinderhughes, E. (1989). *Understanding race, ethnicity, and power*. New York: Guilford Press.

Plummer, D. L. (1995). Patterns of racial identity development of African American adolescent males and females. *Journal of Black Psychology, 21,* 168–180.

Reyes, M. B., Routh, D. K., Jean-Gilles, M. M., Sanfilippo, M. D., & Fawcett, N. (1991). Ethnic differences in parenting children in fearful situations. *Journal of Pediatric Psychology, 16,* 717–726.

Rogler, L. H., Malgady, R. G., Constantino, G., & Blumenthal, R. (1987). What do culturally sensitive mental health services mean? The case of Hispanics. *American Psychologist, 42,* 565–570.

Rosen, E. J., & Weltman, S. F. (1996). Jewish families: An overview. In M. McGoldrick, J. Giordano, & J. Pearce (Eds.), *Ethnicity and family therapy* (2nd ed., pp. 611–645). New York: Guilford Press.

Sakauye, K. M. (1989). Ethnic variations in family support of the frail elderly. In M. Zucker-Goldstein (Ed.), *Family involvement in treatment of the frail elderly* (pp. 63–106). Washington, DC: American Psychiatric Press.

Sakauye, K. M., & Chacko, R. C. (1994). Issues in the psychiatric care of Asian/Pacific American elders. In American Psychiatric Association (Ed.), *Ethnic minority elderly: A task force report of the American Psychiatric Association* (pp. 115–148). Washington, DC: American Psychiatric Association.

Scrivner, R., & Eldridge, N. S. (1995). Lesbian and gay family psychology. In R. Mikesell, D. D. Lusterman, & S. H. McDaniel (Eds.), *Integrating family therapy: Handbook of family psychology and systems theory* (pp. 327–345). Washington, DC: American Psychological Association Press.

Slaughter-Defoe, D. T., Nakagawa, K., Takanishi, R., & Johnson, D. J. (1990). Toward cultural/ecological perspectives on schooling and achievement in African and Asian American children. *Child Development, 61,* 363–383.

Sluzki, C. E. (1979). Migration and family conflict. *Family Process, 18,* 379–390.

Smith, D. (1996). Hungarian families. In M. McGoldrick, J. Giordano, & J. Pearce (Eds.), *Ethnicity and family therapy* (2nd ed., pp. 530–543). New York: Guildford Press.

Staples, R., & Johnson, L. B. (1993). *Black families at the crossroads: Challenges and prospects.* San Francisco: Jossey-Bass.

Steinberg, L., Dornbusch, S. M., & Brown, B. B. (1992). Ethnic differences in adolescent achievement: An ecological perspective. *American Psychology, 45,* 913–920.

Sue, D. W., & Sue, D. (1990). *Counseling the culturally different: Theory and practice* (2nd ed.). New York: Wiley.

Sue, S., & Okazaki, S. (1990). Asian American educational achievements: A phenomenon in search of an explanation. *American Psychology, 45,* 913–920.

Sue, S., & Zane, N. (1987). The role of culture and cultural techniques in psychotherapy: A critique and reformulation. *American Psychologist, 42,* 37–45.

Sutton, C. T., & Nose, M. A. B. (1996). American Indian families: An overview. In M. McGoldrick, J. Giordano, & J. Pearce (Eds.), *Ethnicity and family therapy* (2nd ed., pp. 31–44). New York: Guilford Press.

Szapocznik, J., & Kurtines, W. M. (1993). Family psychology and cultural diversity. *American Psychologist, 48,* 400–407.

Szapocznik, J., Rio, A., Perez-Vidal, A., Kurtines, W., Hervis, O., & Santisteban, D. (1986). Bicultural effectiveness training (BET): An experimental test of an intervention modality for families experiencing intergenerational/intercultural conflict. *Hispanic Journal of Behavioral Sciences, 8,* 303–330.

Szapocznik, J., Santisteban, D., Kurtines, W., Perez-Vidal, A., & Hervis, O. (1984). Bicultural effectiveness training: A treatment intervention for enhancing intercultural adjustment in Cuban American families. *Hispanic Journal of Behavioral Sciences, 6,* 317–344.

Tamura, T., & Lau, A. (1992). Connectedness versus separateness: Applicability of family therapy to Japanese families. *Family Process, 31,* 319–340.

Thompson, J. W. (1994). Issues in the psychiatric care of American Indian and Alaska Native elders. In American Psychiatric Association (Ed.), *Ethnic minority elderly: A task force report of the American Psychiatric Association* (pp. 91–114). Washington, DC: American Psychiatric Association.

Timimi, S. B. (1995). Adolescence in immigrant Arab families. *Psychotherapy: Theory, Research, Practice, and Training, 32,* 141–149.

Tseng, W. (1985). Cultural aspects of family therapy. *International Journal of Family Psychiatry, 6,* 19–31.

Tyler, F. B., Brome, D. R., & Williams, J. E. (1991). *Ethnic validity, ecology, and psychotherapy: A psychological competence model.* New York: Plenum Press.

Vargas, L. A., & Koss-Chiono, J. D. (1992). *Working with culture: Psychotherapeutic interventions with ethnic minority children and adolescents.* San Francisco: Jossey-Bass.

Vazquez-Nuttall, E., Avila-Vivas, A., & Morales-Barreto, G. (1984). Working with Latin American families. *Family Therapy Collections, 9,* 74–90.

Walsh, F. (Ed.) (1993). *Normal family processes* (2nd ed.). New York: Guildford Press.

Wamboldt, F. S., & Reiss, D. (1991). Task performance and the social construction of meaning: Juxtaposing normality with contemporary family research. In D. Offer & M. Sabshin (Eds.), *The diversity of normal behavior* (pp. 164–206). New York: Basic Books.

Wilson, M. N. (1989). Child development in the context of the Black extended family. *American Psychologist, 44,* 380–385.

Wright, L. (1995, November 6). One drop of blood. *The New Yorker.*

Zayas, L. H., & Solari, F. (1994). Early childhood socialization in Hispanic families: Context, culture, and practice implications. *Professional Psychology: Research and Practice, 25,* 200–206.

Zayas, L. H., Torres, L. R., Malcolm, J., & DesRosiers, F. S. (1996). Clinicians' definitions of ethnically sensitive therapy. *Professional Psychology: Research and Practice, 27,* 78–82.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.09
Sexual Orientation

BEVERLY GREENE
St. John's University, Jamaica, NY, USA

10.09.1 INTRODUCTION 207
10.09.2 DEFINING AND CONCEPTUALIZING SEXUAL ORIENTATION: CHALLENGES 209
10.09.3 DEMOGRAPHICS 210
  10.09.3.1 Etiology Theories 213
10.09.4 UNIQUE PSYCHOLOGICAL AND DEVELOPMENTAL TASKS CHALLENGING LGB PERSONS 215
  10.09.4.1 Identity Development 216
  10.09.4.2 Resilience in LGB People 217
10.09.5 UTILIZATION OF PSYCHOLOGICAL SERVICES 218
  10.09.5.1 Psychotherapy with LGB Persons 219
10.09.6 LGB PERSONS AND FAMILY ISSUES 222
  10.09.6.1 Gay and Lesbian Couples 223
  10.09.6.2 Gay and Lesbian Parents 224
  10.09.6.3 Gay and Lesbian Youth 225
10.09.7 TRAINING 226
  10.09.7.1 Hetrosexist Bias in the Delivery of Psychological Services 226
10.09.8 FUTURE DIRECTIONS 228
10.09.9 SUMMARY 229
10.09.10 REFERENCES 229

10.09.1 INTRODUCTION

Sexual orientation is generally defined by the sex of the person that an individual is sexually and emotionally attracted to (Bohan, 1996) and includes heterosexual, lesbian, gay, and bisexual orientations. The meaning of sexual orientation and what it represents, like other aspects of sexuality, is contextual. The context in which sexual orientation is defined and a brief history of its conceptualization is required to assess its meaning appropriately both in our society and as an aspect of human identity.

Despite the fact that sexual orientation conceptually encompasses all orientations, the study of sexual orientation in American psychology has focused generally on individuals with lesbian, gay, and bisexual (LGB) sexual orientations. Heterosexuality traditionally has been considered normative and because of this it has not been the object of study or the focus of the kinds of questions that LGB sexual orientations are subject to. The presumption of heterosexuality as the only form of sexual orientation that is psychologically normal and legitimate, and the subsequent depiction of other sexual orientations as deviant or pathological, is referred to as heterosexism (Bohan, 1996; Gonsiorek & Weinrich, 1991; McWhirter, 1990).

The failure of psychological research to ask relevant questions about the origins and determinants of all sexual orientations, rather than selective, nontraditional orientations, obscures information about all of them and leads to a range of erroneous assumptions about sexual orientation in general. Traditionally, nontraditional sexual orientations have been assumed to represent the presence or manifestation of a psychological disturbance or pathological developmental arrest.

Hooker (1957) was among the first empirical studies in psychology to broadly contradict the pathology model of homosexuality in mental health (McWhirter, 1990). Hooker administered batteries of projective tests to 30 homosexual male and 30 heterosexual male subjects. Results were reviewed by experts who could not distinguish homosexual and heterosexual subjects on the basis of their test results. Gonsiorek (1977, 1991), Meredith and Reister (1980), and Reiss (1980) provide detailed reviews of psychometric studies on homosexuality. These studies generally conclude that homosexuality *per se* is not pathological, nor is it related to psychological adjustment. The presence of such evidence, however, has not led to a complete abandonment of the pathology model in mental health practice.

An intense period of political activism and advocacy for social justice for lesbians and gay men, and the efforts of lesbian and gay mental health professionals, led to formal challenges to pathological views and subsequent changes in the diagnostic nomenclature. The details of that struggle are beyond the scope of this chapter, however, the reader is referred to Greene (1994b) and Giusti and Katz (1992) for a more detailed description of them.

In 1973, the American Psychiatric Association (APA) removed the diagnosis homosexuality from the *Diagnostic and statistical manual of mental disorders* (*DSM*), replacing it with the diagnosis of "ego dystonic homosexuality" (Haldeman, 1991, 1999). This diagnosis was used to describe individuals who were unhappy with their lesbian or gay sexual orientation and who wanted therapy to make them heterosexual. The absence of a diagnostic category for people who were dissatisfied with their heterosexuality was another manifestation of the presumed normative nature of heterosexuality. Opponents to the use of ego dystonic homosexuality as a diagnostic entity argued that it continued to pathologize homosexuality, albeit indirectly (Bohan, 1996; Gonsiorek, 1991; Haldeman, 1991, 1994, 1999). Diagnosing this predictable response to society's intolerance as if it were a condition that required repair fails to address the pathology of homophobia (Weinberg, 1972), heterosexism, and the unfair treatment of lesbians and gay men. Rather, it unfairly stigmatized the victims of unfair treatment. Furthermore, it is unlikely that individuals who face hostility and stigma as a result of an LGB sexual orientation would not at different stages of their sexual identity development manifest levels of ego dystonia or distress. In a climate that is hostile to LGB, persons, it would be likely that some would wish to be heterosexual or express distress about their sexual orientation. Bohan (1996) and Haldeman (1994, 1999) argue that a period of ego dystonia does not represent the individual's distress with their sexual orientation *per se*, but with the attitudes of a society that stigmatize that orientation and make its adoption painful. Acknowledging an LGB sexual orientation means not only the loss of heterosexual privilege but also carries the realistic danger of becoming the target of hostility, derision, potential rejection from loved ones, loss of employment, loss of child custody, and even violence (Herek, 1989). The ego dystonia observed should be regarded as a predictable and understandable part of the evolution of an LGB identity and a "normal aspect of the process of managing denigrating attitudes" associated with those identities rather than as a condition that warrants attempts at conversion to heterosexual sexual orientation (Bohan, 1996, p. 19). The diagnostic entity, ego dystonic homosexuality, was removed from the *DSM* in 1988.

In 1975, the American Psychological Association (APA) adopted an official policy statement that: homosexuality *per se* implies no impairment in judgment, stability, reliability, or general social and vocational capabilities (Committee on Lesbian and Gay Concerns [CLGC], 1986; Morin & Rothblum, 1991). At that time, APA urged all mental health professionals to take the lead in removing the stigma of mental illness that was associated historically with lesbian and gay sexual orientations in Western psychology. After years of lobbying efforts, in 1980 the CLGC was approved as a standing committee and an official part of APA's governance. The Committee, whose name has been recently changed to include Bisexual concerns, (CLGBC), was charged with establishing standards for eliminating heterosexist bias in psychological research and practice with LGB persons, as well as developing policy statements on lesbian and gay issues for the Association. Since its adoption, the committee has developed a range of guidelines for affirmative practice and research with LGB persons. Prior to the adoption of this policy, however, reproductive sexuality was legitimized as the only form of sexual orientation that was psychologically normal and morally correct. This assumption widely pervaded research, training, and practice in psychology. Some would argue that despite many changes, the association between pathology and LGB sexual orientations still lingers broadly in various corners of this discipline.

In the 23 years since the introduction of APA policy changes regarding LGB sexual orientations, LGB persons assume a higher level of visibility among those who seek professional psychological services (Liddle, 1997). However, despite major changes in the diagnostic nomenclature aimed at removing the stigma of psychopathology from gay and lesbian sexual orientations, negative bias and misinformation about lesbians and gay men continues and is still observed in clinical practice, research, and training efforts (Brown, 1996; CLGC, 1990; Garnets, Hancock, Cochran, Goodchilds, & Peplau, 1991; Garnets & Kimmel, 1991; Markowitz, 1991; Rothblum, 1993).

In 1986, the CLGC conducted a survey investigating bias in psychotherapy with lesbians and gay men. Two thousand five hundred and forty-four psychologists were surveyed to elicit themes suggestive of biased and sensitive practice. The results suggested the presence of a wide range of variance in the degree to which psychologists conform to unbiased practice standards with lesbian and gay clients (CLGC, 1990; Garnets et al., 1991). Data from this study and other research (DeCrescenzo, 1985; Rudolph, 1988) suggests that practice does not conform to APA policy standards. This suggests that despite major strides in depathologizing lesbian and gay sexual orientations *per se*, there is still much to be done with respect to educating psychologists about sexual orientation and that attitudes within organized psychology toward sexual orientation remain ambivalent (CLGC, 1990; Garnet et al., 1991; Rothblum, 1993).

10.09.2 DEFINING AND CONCEPTUALIZING SEXUAL ORIENTATION: CHALLENGES

"Homosexual" was the traditional, clinical term used to refer to people whose erotic, sexual, and affectional attractions were to individuals of the same sex. Many objections to the use of this term came from lesbians and gay men themselves as it was originally used to describe a form of psychiatric disorder or psychopathology. Other objections focused on the term's perceived emphasis on the sexual component of lesbian and gay men's experiences in isolation from other complex and integral aspects of their identities. Still other objections focused on the gender neutrality of the term and its masking of the differences between lesbians' and gay men's experiences and issues based on gender (Bohan, 1996; Gonsiorek, 1991). Since most early psychological and medical studies on sexual orientation focused on males, the continued use of the term homosexual was deemed methodologically imprecise in its application to both men and women. In the 1990s, LGB sexual orientations or lesbian, gay man, and bisexual man and woman are the terms preferred by APA reflected in their 1994 publication standards (APA, 1994).

Lesbian and gay sexual orientation is defined as an erotic and or affectional disposition to the same sex (Bohan, 1996; Gonsiorek & Weinrich, 1991). Bisexuality is defined as an affectional and/or erotic attraction to members of both sexes, serially or simultaneously (Bohan, 1996). While distinctions are made between members of these groups for the sake of descriptive clarity, human sexuality exists along a continuum in most persons, rather than as dichotomous and discrete categories. Any form of sexual orientation represents an interaction of biological, cultural, historical, and psychosocial influences (Bohan, 1996; Garnets & Kimmel, 1991; McWhirter, 1990).

The definition of sexual orientation in Western cultures is based explicitly on the biological sex of the person an individual is sexually and emotionally attracted to (Ames, 1996; Bohan, 1996). In this context, there is an inextricable link between the sociopolitical meanings of gender and sexual orientation in Western culture (Ames, 1996; Bohan, 1996; Greene, 1994a, 1996a, 1999; Kaschak, 1992; Kitzinger, 1987). Sexual attraction to members of the other gender is a central part of the way that being a normal man or woman has always been defined in American society (Ames, 1996; Bem, 1993; Bohan, 1996; Greene, 1994a, 1994b, 1994c, 1999). It is not surprising that in this context, lesbians and gay men are presumed to want to be members of the other sex or are viewed as defective examples of their own sex.

Bohan (1996) discusses the extent to which certain questionable assumptions about sexual orientation are embedded in psychological theories and paradigms that are also a function of societal gender and sex roles. Lesbian or gay sexual orientation is assumed to entail cross-gender behavior, with the assumption that gender roles are and should be inextricably linked to and defined by a person's biological sex. Bohan (1996) reviews a range of studies and scales in the psychological literature that serve as illustrations of these assumptions. The first psychological scale designed to measure masculinity and femininity assumed that lesbians and gay men would have M–F scores that differed from their biological sex. M–F scores measure the degree to which a person's behavior is consistent with that of male vs. female gender roles. The assumption is that a person's behavior and therefore their score should be consistent with their biological sex. Therefore, a basic assumption of the scale was that adherence to sex role stereotypes defined heterosexual sexual orientation. Departures from those stereotypes marked a person lesbian or gay. These kinds of assumptions are prevalent among lay persons as well as mental-health professionals. They are more of a reflection of what society values and wants people to be rather than an accurate reflection or measure of who they are. In other studies, when animal or human behavior was not consistent with traditional gender role stereotyped behavior, the presence of homosexuality or the potential for its development was presumed (Bohan, 1996; Haumann, 1995; Parker & DeCecco, 1995). The latter is reflected in the assumption that children who behave in gender atypical ways will become lesbian or gay. There is some evidence to suggest a link between extreme gender atypical behavior and later gay sexual orientation in boys. It does not, however, explain the development of lesbian sexual orientation in women, nor does it explain the presence of heterosexual sexual orientations in adults who were gender atypical children (Bohan, 1996). Another assumption related to the latter is expressed in the belief that if you are able to inhibit gender atypical behavior in children you will prevent them from becoming lesbian or gay. Of course there is no evidence to support this belief. All of these assumptions highlight the contextual nature of sexual orientation as a concept. Gender and sex role behaviors and expectations differ across cultures and differ over time within the same culture. Because of these variations, the concept of sexual orientation would vary as well. However, the ethnocentric nature of American psychological research has obscured important differences in gender and sex role expectations across cultures and in doing this has also obscured the effect of those differences on the psychological conceptualization of human sexual orientation.

Gonsiorek (1991) goes on to discuss the problems defining lesbian or gay sexual orientations that contribute to methodological challenges and flaws in empirical research. Problems establishing precise definitions of sexual orientation also affect the degree to which even our estimates of the number of LGB persons and heterosexual persons in the general population can be considered accurate. The concept of sexual orientation may be viewed from essentialist or social constructionist perspectives. Essentialist perspectives view sexual orientation as an intrinsic characteristic of a person, that endures over time, whether it can be observed by the individual possessing it, by others, or not. From this perspective, sexual orientation is an element of identity that has always existed in every person, in every culture, and in every point in time. For the most part, psychology has studied LGB sexual orientations as if they were enduring traits of people whose determinants could be discovered, quantified, and measured objectively and understood. The social constructionist perspective views sexual orientation as a construct that varies over time and place and has meaning only in the context of a particular culture, in a specific point in time. Sexual orientation from this perspective is viewed as contextual. It is a category that has meaning only because in Western culture we choose to imbue it with specific meaning. This meaning of sexual orientation is created out of the importance we give to the sex of a person that an individual is romantically attracted to. As previously discussed, that meaning is also a function of the meaning we give to gender and sex roles. In the absence of such "constructs," sexual orientation *per se* has no special meaning. In cultures where gender and sexuality have different meanings, sexual orientation may not even exist as an entity to be studied or deemed important enough to label (Tafoya, 1997).

10.09.3 DEMOGRAPHICS

Gonsiorek and Weinrich (1991) estimate that the number of lesbians and gay men in the USA ranges from 4 to 17% of the general population depending on the sampling methods and sources used. These estimates may be deceptive, however, given the imprecision of the definitions used and the hidden nature of this population. Both of these problems make obtaining representative samples a major challenge in conducting empirical research.

Kirk and Madsen (1996) observe that one of the seven myths that straight Americans believe is that all LGB persons can be easily identified by their outward appearance. This is based in part on gendered definitions of heterosexual and LGB orientations. Despite these beliefs, for the most part, LGB persons are invisible. They may not be identified solely on the basis of observation as they are a part of every social group in society. Unlike people with physical distinctions or challenges that readily identify them, LGB persons who are subjects in research are deemed to be LGB because they self-identify or say that they are. Just as we regard individuals in research studies to be LGB because they say they are, we do not assume that someone is LGB unless they actually tell us so. Therefore, we cannot be certain about the numbers of heterosexual or LGB participants in such studies. Unfortunately, for a variety of reasons this problem is not completely resolved by simply asking subjects to identify their sexual orientation.

Gonsiorek (1991), Bohan (1996), and McWhirter (1990) assert that problems defining LGB orientations make it difficult to derive accurate population estimates and to obtain appropriate samples. Attempts to define sexual orientation based on sexual behavior are less than reliable. Bohan (1996) observes that a person may experience themselves as or may be appropriately labeled lesbian or gay before they have ever had sexual relationships. Other individuals may have both male and female sexual partners and/or relationships. Some of these individuals would consider themselves LGB while others would not self-label in this way. Still other people may have sexual partners that are not consistent with the content of their fantasies. As observed in the previous example, some would self-label as LGB while others would not.

Gonsiorek (1991) observes that defining and obtaining a representative sample of LGB persons is the largest methodological problem in the scientific study of LGB sexual orientation. If subjects are considered LGB because they say that they are, it is important that the individual subjectively identifies themselves as having an LGB orientation if they are LGB. It is equally important that they do not identify themselves as such if they are not. This approach cannot account, however, for persons who do not yet acknowledge an LGB orientation, or those who are in an ego dystonic period. It also fails to account for individuals who maintain intimate relationships with members of the same sex but do not consider themselves or conceptualize their relationships (for cultural and other reasons) as lesbian or gay (Tafoya, 1997). There are other problems as well. Acknowledging an LGB sexual orientation is often associated with negative consequences. As such, many individuals who may be LGB may not openly acknowledge that they are. Class, ethnic, and other distinctions may also compound the consequences of disclosure and affect the degree to which many people will or will not take those risks. On the one hand, many people who are successful professionally but who are not out will not risk the embarrassment or loss of stature that may be associated with disclosure (Gonsiorek, 1991). On the other hand, many LGB persons who are economically marginal may be unable, for example, to risk the loss of their jobs, child custody, or visitation rights if their sexual orientation is discovered (Bradford, Ryan, & Rothblum, 1994; Fassinger, 1991; Greene 1994a, 1999). Many members of ethnic minority groups may conceal their LGB orientation from their families and ethnic communities to a greater degree than their white counterparts (Chan, 1992; Greene, 1990, 1994b, 1996a, 1997; Morales, 1992). In fact, in many families of people of color, a lover or life partner may be accepted as long as they are not labeled, as long as the family member does not label themselves lesbian or gay, and as long as they do not attempt to discuss their sexual orientation openly. The failure to make such statements about LGB sexual orientation to family members must be understood in the context of the ethnic group the person belongs to, the level of tolerance within that group, and traditional ways of managing members who deviate from family norms without causing family strife (Chan, 1990; Croom, in press; Fygetakis, 1997; Greene, 1994a, 1994b, 1999; Liu & Chan, 1996; Morales, 1992). Fygetakis (1997) observes that many white lesbians and gay men who belong to ethnic or religious groups that are intolerant of LGB sexual orientations may not be “out” to their families or in their ethnic/religious communities for reasons that are similar to those expressed by lesbians and gay men of color. She warns that it is inappropriate to use Western models of being out (particularly to family) as a measure of LGB self-acceptance or well-being for all LGB persons (Fygetakis, 1997). Accordingly, such individuals may not openly acknowledge their LGB orientation as research subjects, despite assurances of confidentiality.

Studies of LGB sexual orientations based on samples of psychiatric patients, prison inmates, psychoanalytic patients, and patients seeking aversion therapy, the primary subject pool for early studies, are not considered representative of LGB persons and therefore are not generalizable to them (Gonsiorek, 1991). Their continued use in making such generalizations is considered methodologically inappropriate. In early studies of sexual orientation, the use of psychiatric patients, patients in psychotherapy, and psychoanalysis was widespread. Many of these studies compared nonpatient heterosexual controls to homosexual patients. Individuals drawn from clinical populations by virtue of their status as patients are not representative of nonpatient LGB samples (Bohan, 1996; Gonsiorek, 1991). Even if this population were representative, comparisons to nonpatient heterosexual controls lack validity. Gonsiorek (1991) observes that samples drawn from individuals in psychoanalysis are atypical of the general population of LGB persons and may be atypical of psychotherapy patients as well. On average they are better educated and must be highly motivated to engage in a form of therapy that requires frequent visits over a protracted period of time and as such is expensive to maintain. Samples drawn from the armed forces, prison, or other law enforcement sources are also dubious. The presence of the same sex sexual behavior in prison or in places where heterosexual behavior is not available does not mean that its participants are LGB persons.

Gonsiorek (1991) also warns that individuals whose sexual behavior has come to the attention of law enforcement officials may include a higher proportion of individuals whose reality testing, impulse control, and capacity for good judgement are impaired. Often, such individuals may have been involved in sexual activity that was not conducted in private, was not consenting, was not with another adult, or was the specific result of entrapment (Gonsiorek, 1991).

Still, there are problems with other samples as well. Gonsiorek (1991) observes that studies whose samples are drawn from LGB bars and clubs are skewed toward people who drink alcohol, are young, able-bodied, extroverted, urban, lack a consistent sexual partner, and who can economically afford to socialize in those venues (Bradford et al., 1994; Croom, in press; Greene, 1999). Other samples drawn from LGB organizations will over-represent people who are out and who do not fear disclosure of their LGB orientation. Those samples obtained via social, friendship networks, experimenter contacts, and the like tend to be biased in their homogeneity (Gonsiorek, 1991).

Generally, samples of LGB persons are sought in venues that tend to be labeled as such, an approach that would seem to make sense. However, the diversity of LGB persons warrants understanding that many of them will not be found in such venues for a variety of reasons, some of which have been addressed here. Whether or not they would respond to requests for participation in such research is questionable. We know that there are many persons who would never be asked to participate by virtue of their absence in the kinds of venues targeted for obtaining samples. Hence, they are invisible to us. Largely, research is based on samples that are predominantly white, middle-class, able-bodied, North American, English speaking, and well educated (Bradford et al., 1994; Croom, in press; Gonsiorek, 1991; Greene, 1999). Because of these limitations, results of LGB research should be approached with caution as its generalizability is often limited to very specific subgroups. Similarly, suspicion is warranted when evaluating population estimates of LGB persons. It is likely that those estimates are an under-representation of the total population of LGB persons in the USA (Bohan, 1996). It is wise to ask how we know what we propose to know about the group in question, how that information was acquired, and who would have been omitted by virtue of the sampling methods chosen.

LGB persons are often presumed to be a part of a monolithic community. This obscures the wide range of diversity within them as a group that crosses all cultural, ethnic, racial, economic, social, age, geographic locale, and other lines. There is little research on or about LGB realities outside of the USA. Brown (1989) suggests that there is no unitary lesbian or gay reality; rather, there are multiple realities as diverse as lesbians and gay men themselves (Bell & Weinberg, 1978). While generalizations are made about a range of psychological realities for LGB persons, they are always embedded in other aspects of human experience and cannot be understood realistically when examined in isolation. LGB persons with multiple identities may experience their sexual orientation and heterosexism very differently than those who do not have multiple identities (Croom, in press; Greene, 1994b; 1999; Potgieter, 1997).

For the most part, American psychology as well as LGB psychology has understood human behavior from decidedly Western cultural perspectives. It is important to view LGB psychology from diverse perspectives within and outside of the USA. There is often a glaring absence of international perspectives on LGB persons' lives. Assumptions about what is helpful or healthy in one context may not have the same meaning or implications in another. For example, in American LGB psychology, the demise of pathology models of LGB sexual orientation is viewed as a major step toward affirming LGB sexual orientations as equal in status to heterosexual sexual orientations. It is difficult to consider that this would not be the case for LGB men and women around the world, however, this may not be as simple as it appears. Potgieter (1997) interviewed a sample of 30 Black South African lesbians from rural and urban settings and from a wide range of class and educational backgrounds in South Africa. Many of the women she interviewed had never even spoken to another lesbian. Potgieter's (1997) study documents, with the exception of dissertations, the absence of a single published article on homosexuality by any South African psychology journal, as well as the prominent role of South African psychologists in supporting the entrenchment of apartheid, and the view that homosexuality is both a sickness (that can be cured) and a sin (that the individual is responsible for) that warrants criminalization.

It would seem on the face of things that the eradication of the pathology model would benefit South African lesbians and gay men of all ethnicities. Potgieter (1997), however, points out that among Black South Africans, families are required to maintain contact with and provide family support for any member who is deemed ill. Despite the fact that lesbian or gay sexual orientation is viewed as both pathological and sinful and is disapproved of, the LGB family member is explicitly entitled to support and is not evicted to the degree that LGB persons in Western culture might be treated by disapproving family members. Potgieter (1997) asserts that we must consider what would happen, in Black South African culture, if the lesbian or gay family member were not viewed as sick, but simply sinful, particularly in a country where the gay and lesbian rights movement *per se* is fledgling at best. The entire basis of an entitlement to family support could be questioned. Hence, the impact of rejecting the pathology model poses different consequences for LGB persons in different cultures. This highlights the importance of viewing LGB psychologies and imperatives from a wide range of perspectives.

10.09.3.1 Etiology Theories

The "cause" of LGB sexual orientations has long been a focus of interest in psychological and medical research in the USA. However, the very question suggests that we know the origins of heterosexuality. The reality is that we presume to know because it is a preferred orientation and whatever is believed to be characteristic of it is deemed normative. When the question of causation of LGB sexual orientations arises, it prompts us to ask why its origins are important, particularly since similar questions are not raised about heterosexual sexual orientation. The latter is presumed to be an evolutionary and adaptive means of continuing the species. This assumption prevails despite what we know about the complexity of human sexuality and relationships that goes far beyond propagation.

The importance of determining a cause for LGB orientations assumes that if a cause is identified, interventions may be designed to prevent the development of LGB orientations. This is consistent with the prevalence of the disease and pathology models of homosexuality, although some proponents of conversion would say that their position does not rest on the presumption of pathology. This group of advocates of conversion would suggest that LGB sexual orientation is problematic for individuals because it is stigmatized and because it results in the LGB person being disadvantaged. The logic is that it is kinder to convert the individual to something that society will more readily accept. Despite the formal changes in the diagnostic nomenclature, inquiries about causation are embedded in the need to justify conversion. Altering LGB sexual orientations, in this context, is couched in what appears to be a benign assumption. If being a member of a stigmatized group is difficult, removing the stigma will make a child or adult's life easier. LGB sexual orientations, in this context, are still deemed the cause of the problem. Alter it and the problem goes away.

Bohan (1996) points out another way that the pathology model continues to inform the way that LGB persons are understood. She writes that when explanations for LGB orientations are sought, it is with the assumption that whatever markers are used or traits that are found will be, by definition, abnormal in LGB persons. Conversely, whatever is found in heterosexual persons will be considered normal. For example, she observes that the research sought to find hormonal imbalances in LGB samples to explain the presence of their orientation. This approach, however, presumed that whatever levels were measured in heterosexuals were normal. Bohan (1996) suggests that the presence of bisexuality in different societies in different historical times gives credence to the bisexual potential in all persons and suggests that legitimizing only one sexual orientation is arbitrary.

Explanations for LGB sexual orientations range from one extreme in which biological underpinnings are presumed to totally account for them, and at the other, that environmental determinants are deemed causative. There is also a range of approaches that postulate the presence of combinations of determinants that fall between the polarities of these extremes and contain elements from both of them. These theories fall roughly into three categories.

Explanations based on psychodynamic theory are well known and longstanding. Overall traditional psychoanalytic explanations of LGB sexual orientation view them as a symptom of a developmental conflict and early fixation (Isay, 1990). Isay (1990) observes that traditional analysts believe that homosexuality represents an unfavorable unconscious solution to a developmental conflict and that the entire personal life of the individual reflects the effects of the early fixation. While this view is based on Freud's (1922/1959) early models, Isay (1990) notes that post-Freudian analysts ignore the ambiguities in his theory of object choice. Freud suggested that all human beings were innately bisexual, and supported decriminalizing and depathologizing homosexuality (Bohan, 1996; Freud, 1922/1959; Isay, 1990; Jones, 1965). Despite this, American psychoanalysts viewed LGB sexual orientations as illnesses in need of a cure. A range of explanations developed that posit the presence of dysfunctional relationships between parents and children accounting for the development of LGB orientations. Their common themes are based on the assumption that a key task in psychosexual development is to identify with a parent of the same sex and transform oneself psychologically into that parent. Based primarily on studies with psychoanalytic patients as subjects, psychoanalytic research alleged that the fathers of gay men were aloof, distant, and uninvolved with them, while their mothers were smothering and overprotective (Isay, 1990). Lesbianism was deemed a manifestation of incest fears, an expression of a wish to adopt a masculine role, or "a girl's anger at her father for rejecting her causing her to hate men" (Bohan, 1996, p. 76). Critiques of these explanations cite their reliance on patient samples and the presence of other methodological problems. There is also too much variation in the families and family dynamics of lesbians and gay men to make such generalizations (Bohan, 1996; Isay, 1990). Furthermore, the same kinds of dynamics ascribed to the families of lesbians and gay men may be found among heterosexual men and women. There are no adequate explanations in these paradigms for why heterosexual men and women with like family dynamics do not become lesbian or gay.

Biological explanations enjoy great popularity in the 1990s and tend to look for genetic and/or hormonal markers of lesbian or gay sexual orientation. There is some evidence that there may be a genetic element in the development of LGB orientations for some individuals, but it is far from conclusive (Bohan, 1996; Diamond, 1993; McGuire, 1995). Bohan (1996) provides a detailed review of concordance studies that are summarized briefly. Concordance studies seek to determine the degree to which there is agreement between two individuals on a particular trait. In this case, the trait would be LGB sexual orientations. We would expect, if sexual orientation were 100% genetically determined, to find a near 100% concordance rate between monozygotic twins, and a 50% rate on measures between fraternal siblings. Studies in this area (Bailey & Pillard, 1991) produce results that show higher concordance rates for identical twin brothers than for fraternal siblings, however, not at the levels we would expect to find to demonstrate 100% genetic determination. Similar findings for women are not reported. Other biological theories analyze the role of hormones in prenatal development (Bailey, 1995; Hamer, Hu, Magnusson, Hu, & Pattatucci, 1993 Meyer-Bahlberg, 1995) and sex differences in brain organization. In the former theories, sex differences in behavior are attributed to hormonally related brain organization. Despite contradictory findings in studies attempting to replicate this work (Byne, 1995), these models seem to have some measure of validity (when sexual behavior is the variable measured) in research with lower animals. However, making the leap to encompass the complexity of an LGB orientation is questionable. Gendered assumptions about LGB orientations are evident in these approaches. It is presumed that brain development is influenced by the presence of prenatal hormones and that sex-typical behaviors are a reflection of brain organization rather than cultural influences. Lesbians are presumed to have brain organizations similar to those of heterosexual men, and gay men are presumed to have brain organizations similar to those of heterosexual women. A familiar problem in this schema is that gendered behaviors may have specific meaning in Western culture that they do not have in other cultures. The most recent study conducted by LeVay (1991) is based on autopsied brains of heterosexual and gay men. LeVay (1991) hypothesized that a forward region of the hypothalamus would be smaller in the brains of gay men (as it would in heterosexual women) than in heterosexual men. These findings were confirmed. Despite this, the meaning of these findings is unclear. LeVay's (1991) sample was small and involved the use of subjects who died as a result of complications from AIDS. Furthermore, there was no evidence that the differences in size of these brain structures were prenatal. Overall, the results of these studies are intriguing but far from definitive. Bohan (1996) and McWhirter (1990) observe that research in the biology of sexual orientation has failed to produce any conclusive findings.

Another group of explanations are based on social learning or learning theories. These theories suggest that sexual orientation is a constellation of complex, learned behavior based on prior experiences that have been reinforced. One example posits the existence of previous heterosexual encounters that are unpleasant or traumatic, or same-sex sexual encounters early in life that are seductive in nature. Still others suggest that "gender deviance," particularly in boys, goes unpunished and evolves into an LGB orientation. Bohan (1996) explains that it is possible that some children may cope with their gender atypical identity by internalizing the identity that the culture provides them with to explain their difference, therefore defining themselves as lesbian, gay, or bisexual. These are the labels they are given to choose from. There is little to support perspectives based solely on social learning models. This is particularly so given the degree to which LGB persons often identify themselves as such long before they have any same-sex or heterosexual contact. There is also evidence that individuals who engage in same-sex sexual behavior do not necessarily adopt an LGB orientation. Furthermore, same-sex sexual contact is more likely to be delayed because of the stigma associated with LGB orientations. The notion that there is reinforcement for them seems questionable. Clinicians observe clients who want nothing more than to avoid the societal rejection that accompanies the adoption of an LGB identity.

In a shift toward a more interactionist approach, Bem (1996) proposes that biological determinants predispose temperaments and that it is temperaments that predispose behavior which is sex-typical or atypical. Bem's theory asserts that a child who is sex-atypical will feel different from his or her same-sex peers and may be rejected by them. Children whose temperament predisposes them to be sex-typical experience themselves as different from peers of the other sex. The rejection from peers and a sense of difference from them is presumed to create anxiety and discomfort which then triggers physiological arousal. He proposes that this is the crucible in which an association between same-sex peers and physiological arousal leads to erotocizing those peers. Bohan (1996) suggests that this theory fails to explain why in this situation anxiety becomes eroticized rather than leading to avoidance of same-sex peers.

The question of causality in LGB sexual orientation research is intriguing but leaves us with no definitive conclusions. McWhirter (1990) observes that there does not seem to be a single cause or a simple developmental path that determines sexual orientation of any type. Bell and Weinberg (1978) suggest that our understanding of sexual orientation would be greatly enhanced by recognizing and considering that there is great diversity among LGB persons, just as there is among heterosexual men and women; that there may be many homosexualities as well as many multiple paths to any single form of sexual orientation (Garnets & Kimmel, 1991).

Sexual orientation is likely established by adolescence, usually before sexual activity begins, often preceded by a subjective awareness of same gender attraction (Bell, Weinberg, & Hammersmith, 1981; Garnets & Kimmel, 1991; Gonsiorek & Weinrich, 1991). Still other LGB persons may not be aware of same-sex attractions until later in life. It is difficult to determine how, at what stage of development, or chronological age LGB sexual orientation would emerge were it not for the pervasive societal messages condemning it. Many individuals who have same-sex attractions may suppress or repress an awareness of them because of the ways that LGB identity is stigmatizing. This is reflected in the tendency among many LGB persons to reject or experience confusion about their orientation initially. Sexual orientation appears to be a stable characteristic over the life span for some individuals; for others, one orientation may be adopted after lengthy experience with the other as an adult. Money (1988) refers to the latter as sequential bisexuality (Bohan, 1996; Garnets & Kimmel, 1991; Golden, 1996).

Because LGB individuals are not readily identifiable on the basis of physical characteristics, they are often presumed to be and treated as if they were heterosexual. Such treatment forces LGB persons to make conscious decisions about whether or not to reveal their sexual orientation routinely. This raises one of the core psychological tasks confronting LGB persons, "coming out."

10.09.4 UNIQUE PSYCHOLOGICAL AND DEVELOPMENTAL TASKS CHALLENGING LGB PERSONS

Coming out has been defined at its simplest as the realization or conscious acknowledgment of one's LGB sexual orientation and the subsequent disclosure of that orientation to others (Bohan, 1996; Dworkin & Gutierrez, 1992; Garnets & Kimmel, 1991; Isay, 1990). While this definition describes the salient features of coming out as a developmental event, it is a deceptively simplistic depiction of an extremely complex lifelong task. Because of the presumption of heterosexuality in our society, most LGB persons will be presumed to be heterosexual unless they say otherwise. For members of many ethnic minority groups, skin color and other features make their group membership visible and a given that is beyond their control. Unlike visible minority group members, LGB persons must consciously decide whether or not to disclose their sexual orientation to others. This also means that to some extent, they have some degree of control over who has information about their orientation and who does not, affording a measure of safety under certain circumstances. This control, however, has benefits as well as disadvantages. The need to make decisions about coming out surfaces whenever the individual is in a new situation such as a new job with new colleagues, and is usually stressful. Whether or not to disclose, to whom, at what point in time, and weighing the risks vs. benefits, is a lifelong task for most LGB persons.

Perhaps the most anxiety provoking stage of coming out is the initial stage of self-awareness. Bell, Weinberg, and Hammersmith (1981) report that 70% of lesbians and gay men in their sample reported having a vague sense of feeling different from same-sex peers as young as age four or five and that the vague feeling of difference had or developed a sexual component. Savin-Williams (1996) reports that LGB youth move from the stage of awareness to that of attempting to name or define the sense of difference that they experience. This period of self-definition, that for some but not all LGB people begins in early childhood, is usually resolved by early adolescence or young adulthood. The resolution of self-definition, however, may be postponed from several years to decades depending on the degree to which negative attitudes toward LGB people are internalized (Savin-Williams, 1996).

Following the period of self-awareness, another anxiety-provoking period may be marked by the desire to come out to family members, about not wanting to do so, or the uncertainty of which course to take. Many individuals, however, endure constant concern about being "outed" before they have chosen to personally disclose their sexual orientation to someone. Individuals who do not disclose their sexual orientation are considered closeted (Bohan, 1996). However, being closeted, like being out, is not a dichotomous, either/or state. Most LGB persons are out to some people in their lives and closeted to others.

Coming out is a major source of stress for LGB persons and is frequently cited as a major source of anxiety among LGB clients entering psychotherapy (Bell & Weinberg, 1978; Bradford & Ryan, 1988). Coming out is, in reality, a process that incorporates an important element of a person's identity, an element that the individual may not have always been aware of. The LGB individual is required to incorporate this perhaps newly discovered aspect of themselves, in affirmative ways, in the midst of a culture which is anything but affirming. This is reflected in its legitimized intense, negative reactions to LGB persons (Bohan, 1996; Garnets & Kimmel, 1991; Gonsiorek, 1991; Isay, 1990).

There is research that documents the benefits of coming out. Decreased feelings of loneliness and guilt, identity synthesis, integration, and commitment (Cass, 1979; Coleman, 1981/1982), healthy psychological adjustment and positive self-esteem (Savin-Williams, 1990), positive gay identity, a greater sense of freedom to be oneself, of not living a lie, and experiencing genuine acceptance were reported as benefits of disclosure (Savin-Williams, 1996). Despite the benefits of disclosure, for some LGB persons negative reprisals may outweigh benefits depending on the unique aspects of an individual's life, that may change over time and in different situations. For some LGB persons, reluctance to come out may be attributed to a reluctance to confront fears of the unknown, an avoidance of rejection, or harassment or abuse by society and/or loved and trusted figures in the individual's life (Savin-Williams, 1996). Some individuals structure their lives around two completely separate worlds, their LGB friends and acquaintances with whom they can be out, and their co-workers, families, friends, or other persons with whom they are closeted.

It is important to remember that coming out is a contextual process, and that it will be experienced, timed, and understood by the individual in concert with that person's gender, ethnicity, socioeconomic class, age, and other distinctions. For example, there may be differences in the process that are a function of gender that make it different for gay men and lesbians. Gonsiorek (1988) suggests that coming out may be more abrupt for males who come out during the adolescent period than for females, and that men may tend to act on their sexual feelings at an earlier stage than women. Women in Western cultures are permitted to display a broader range of emotions and behaviors with other women that do not violate their traditional gender role expectations. Physical contact and physical displays of affection between women are not necessarily presumed to mean that such contact is sexual or that they are lesbians. Their male counterparts do not enjoy the same latitude of behaviors and in the expression of emotions. Men engaging in physical displays of affection with other men would be deemed suspicious as such conduct is not in keeping with the Western male stereotype of masculinity. However, there are subcultural differences in the latitude men are allowed for such conduct. For example, African-American males who have close emotional or familial ties are very likely to hug one another or exchange unique handshakes or other displays on greeting one another. This can be understood as a means of displaying solidarity as well as affection (Boyd-Franklin, 1990). In the context of African-American and Latino communities, such behavior is not presumed to represent suspicious conduct, but rather may represent a more culturally syntonic value of touch. This underscores the importance of understanding the cultural reference point of the person as it will play a significant role in that person's response to their own as well as other's behaviors.

10.09.4.1 Identity Development

Theories of sequential stages in identity development in LGB persons are similar in structure to those for ethnic group members (Bohan, 1996; Cass, 1984; Coleman, 1981/1982; Fassinger, 1991; Weinberg, 1983). However, smooth transitions through the stages of any of these models will be complicated by the pervasive presence of negative attitudes toward and discrimination against group members. Members of ethnic minority groups, who come from healthy families and affirming communities, usually receive positive cultural mirroring during the course of their development that LGB people do not receive. That is, ethnic minority individuals are often socialized to actively challenge rather than accept the dominant culture's assertions about them from their families and community. LGB persons, however, learn a range of negative stereotypes and attitudes about LGB persons, not simply from the dominant culture but from loved and trusted figures (Greene, 1994a, 1994b; Greene & Boyd-Franklin, 1996). To complicate matters, they are likely exposed to such attitudes long before they are aware of their own sexual orientation. When LGB persons have uncritically accepted such values and attitudes, moreover, when they have internalized them, the process of self-acceptance becomes complicated and often fraught with ambivalence. It is not uncommon during this period to observe what has been previously described as ego dystonic feelings about one's LGB orientation. The internalization of hateful and denigrating attitudes toward LGB persons has been defined as internalized homophobia (Bohan, 1996; Gonsiorek, 1982; Shidlo, 1994; Weinberg, 1972). Internalized homophobia may be manifested in a hatred or contempt for any quality or characteristic attributed to LGB persons. It may also be manifested in depression, suicidal ideation or attempts, verbal or physical aggression toward other LGB persons, substance abuse, and other self-destructive behaviors. Addressing internalized homophobia and eliminating it is an important aspect of psychotherapy with LGB persons whenever it is present.

Successfully moving through the different stages of any model may take years or, for some individuals, a lifetime to accomplish. The precise length of time required, however, will also vary from person to person depending on other circumstances and other psychological tasks they must perform. There is frequently a delay in time between the discovery of an LGB sexual orientation and its acceptance as a healthy and acceptable part of an individual's identity (Bohan, 1996; Garnets & Kimmel, 1991; Gonsiorek, 1991).

10.09.4.2 Resilience in LGB People

D'Augelli (1994) suggests that development in LGB persons should be approached from the perspective of exceptionality given the superior coping strategies required for them to negotiate developmental tasks successfully in an antagonistic environment. Such an approach, he suggests, will be more likely to yield important information about adaptive and perhaps even exceptional strategies that might otherwise be overlooked. The concept of resilience in LGB persons is rarely explored. Jones (1997) attempts to analyze the components of resilience by examining the paradoxical life of jazz composer Billy Strayhorn. Strayhorn, an African-American gay man in a viciously racist and homophobic climate, was also the victim of childhood physical abuse and other troubled family dynamics. Despite those obstacles, he came to be known as a major figure in American jazz music idioms, and developed what appeared to be satisfying adult relationships of all types. His adult life was not entirely free of the residual psychological effects of his childhood struggles with adversity. In many ways the effects of those struggles were compounded by the double layers of societal discrimination that were routine features of American society at that time. Jones (1997) identifies a range of key ingredients in the psychological resilience found in members of many groups, including sexual minorities, who have been forced to endure "patterned injustice" (p. 13). Among those ingredients, Jones (1997) observes that carefully constructed, segregated communities can facilitate the development of self-constructed psychological realities and a form of psychological independence observable in members of disadvantaged groups. People who move within those communities can develop a high level of resilience in part because these communities provide them with the opportunity to engage in a process of independent self-construction. LGB subcultures have always existed. They have offered LGB persons the opportunity to have small corners of the world in which they may safely be themselves and experience other people who are like them directly. Despite their often secretive nature, these subcultures nevertheless provide members with the important opportunity to develop the alternative self-images needed to challenge the negative labels applied to them from members of the dominant group (Greene, 1999; Greene & Boyd-Franklin, 1996; Jones, 1997). The opportunity to engage in a process that allows group members to define themselves rather than accept negative definitions of them is a salient ingredient in their mental health.

The process of coming out is further complicated by the frequent absence of routinely available and explicit role models that socially affirm LGB identities. Media images of LGB persons, with rare exceptions, routinely depict them as seriously flawed, tragic, or comedic characters, when they are visible at all. Such depictions leave many persons struggling with their new found identity, reluctant to identify with a group which is perceived to have many negative features that are deemed to characterize them. As a result, many individuals must develop their own personal framework for identity and for maintaining self-esteem (Garnets & Kimmel, 1991; Peplau, 1991).

10.09.5 UTILIZATION OF PSYCHOLOGICAL SERVICES

Lesbians and gay men are more likely to see a therapist or seek psychological services than their heterosexual counterparts (Bell & Weinberg, 1978; Bradford et al., 1994; Morgan, 1992; Morgan & Eliason, 1992). This is not surprising as group members are vulnerable to oppression and discrimination, the denial of many basic legal rights, as well as exposure to the same life stressors that their counterparts must negotiate (Fassinger, 1991; Greene, 1999). Some seek therapy to have a place where they can safely discuss their feelings about their sexual orientation, where their "secret" cannot be divulged to anyone else without their permission. Over 50% of lesbians in a national sample reported being the target of verbal attacks, 13% reported losing their jobs, and others reported discrimination against or being stereotyped by mental health professionals because of their sexual orientation (Bradford et al., 1994). This kind of treatment often results in fears of disclosure even to mental health professionals. Such fears of disclosure may represent another factor that indirectly affects our estimates of the numbers of LGB persons receiving mental health services, as some remain invisible to us, even in the therapy hour. Bradford et al. report that three-quarters of their sample reported seeking mental health services at some time, the majority reported finding it helpful.

Most LGB persons seek psychological services for many of the same reasons that other people do, normal life transitions as well as unexpected trauma. However, they must manage those stressors in the context of an environment that is hostile to them. They also tend to see therapists for longer periods than their heterosexual counterparts (Liddle, 1997; Morgan, 1992). Many of their presenting problems, however, focus on or are related to the presence of societal and internalized homophobia, issues that heterosexual persons do not have to manage, often exposing them to chronic, high levels of stress (Morgan, 1992). These issues are compounded for LGB persons who are members of ethnic minority or other societally disadvantaged groups as they must manage multiple levels of discrimination (Chan, 1989; Greene, 1999; Greene & Boyd-Franklin, 1996). Liddle (1997) observes that LGB clients tend to choose LGB therapists 41% of the time; gay men tend to choose gay or bisexual male therapists; lesbians tend to choose lesbian or bisexual female therapists (Modrcin & Wyers, 1990). When the sexual orientation of the therapist is unknown, lesbians are more likely to see female therapists, with 89% of Liddle's (1997) lesbian sample choosing female therapists (Modrcin & Wyers, 1990). Morgan (1992) finds that lesbians hold unusually positive attitudes about seeking professional help. However, this is consistent with the demographics of lesbians that tend to be sampled. Members of groups sampled tended to be young, White, and better educated than their heterosexual counterparts (Morgan, 1992; Trippet, 1994). Higher levels of education are associated with positive attitudes toward seeking professional help.

Presenting mental health concerns in therapy include homophobia and stress (Niesen, 1990), depression (Rothblum, 1990), relationship issues (not being out or coming out to family members, limited support for or disparaging attitudes toward relationships), and conflicts about being closeted or out (e.g., being closeted at work and out socially; whether or not to come out and to whom). Bradford et al. (1994) found depression present in half of their sample, in fact depression was often the factor precipitating a referral. They also determined that the degree to which a subject was out affected their mental health. In a national survey of lesbian health care and mental health concerns, Bradford et al. found that relationships were a major focus of attention in therapy. Problems with lovers were reported as a major focus by 44% of participants, with family (34%), and with friends (10%), respectively. Loneliness was also a significant focus of attention (21%). In the Bradford et al. sample, 68% of respondents reported experiencing a range of mental health problems in the past. Those problems included long-term depression and sadness, constant anxiety and fear, as well as other mental health concerns. However, only 23% of that group reported being in therapy prior to the present time. 49% of that group, however, reported having been in therapy for one year or less. These patterns suggest the presence of coping and survival skills, reliance on friends, and the presence of social supports as alternatives to therapy (Bradford et al., 1994; Fassinger, 1991).

Concerns about money and financial well-being were reported by 57% of the Bradford et al. (1994) sample. The percentages of African-American lesbians and lesbians aged 54 and younger were somewhat higher in the group reporting financial concerns. While most respondents reported a preference for private therapy, the lack of third-party payments and insufficient personal resources impinged on those preferences. One-third of the sample reported seeking professional services for "personal growth," however, this was less frequently cited as a reason for referral among African-American and Latin, lesbians. Rates of eating disorders and sexual abuse among the respondents in this group were similar to those of their heterosexual counterparts. Higher percentages of suicidal ideation and attempts were reported, however, among the younger aged groups, suggesting that the potential for suicide among LGB adolescents warrants serious attention (Bradford & Ryan, 1989; Bradford et al., 1994). Researchers caution that the samples in the research studies cited tended to be young, White, better educated, and accessible to major cities. Older, non-White, financially less well off LGB persons in rural or isolated communities are not well represented. In isolated settings, people may be at even higher risk for distress and the development of mental disorders, however, they may be more likely to be cut off from a supportive and sympathetic community when they are in greatest need of one.

10.09.5.1 Psychotherapy with LGB Persons

Prior to changes in the diagnostic nomenclature, most psychotherapy with lesbian and gay clients focused on changing their sexual orientation. Despite changes in the nomenclature and organized mental health's rejection of pathology theories, efforts to change the sexual orientation of gay and lesbian clients have persisted and are referred to as conversion therapies (Haldeman, 1991, 1994; 1999; McWhirter, 1990). There is no credible empirical evidence to warrant the assumption that therapies aimed at altering sexual orientation are successful. Haldeman (1999) documents psychology's evolution from this position to its current policy and offers guidance to clinicians who deliver services to clients who may be distressed by their LGB sexual orientation.

Groups under the auspices of fundamentalist Christian denominations touting the sanctity of literal biblical interpretations as superior to scientific studies on LGB sexual orientations are the "main purveyors" of conversion therapies. In addition to these groups, the National Association for Research and Therapy of Homosexuality (NARTH), founded by Charles Socarides and Joseph Nicolosi, is committed to the pathology model of LGB sexual orientation. Its members believe that the removal of homosexuality as a diagnostic entity from the *DSM* was the result of the power of gay rights lobbyists and reject any scientific studies challenging the pathology model. Group members are committed to the pathology model and use of reparative therapies to change LGB sexual orientations to heterosexual orientations. Generally, the use of conversion therapies rests on the assumption that LGB sexual orientations are expressions of psychopathology. Gonsiorek (1991) and Haldeman (1994, 1999) provide comprehensive discussions of the weaknesses of the pathology models. Briefly summarized, these studies suffer from a range of methodological problems. Among them, it is not clear if the subjects who are labeled homosexual in fact are homosexual, as opposed to bisexual or even heterosexual. As previously discussed, they often compare lesbian or gay subjects drawn from patient populations to nonpatient heterosexual controls, and frequently use outcome measures of questionable validity. Furthermore, there is data to support the contention that there is no difference between LGB persons and their heterosexual counterparts on instruments measuring psychological adjustment. Haldeman (1999) further documents the methodological weaknesses in studies that purport to demonstrate that sexual orientation can be altered. Selection criteria, subject classification, and outcome measures are cited as major flaws (Haldeman, 1991, 1994, 1999).

As previously examined in the discussion of ego dystonic homosexuality, there are serious questions about the origins of requests from LGB clients to alter their sexual orientation. In this context, Haldeman (1999) raises an even more important question. Whether or not LGB sexual orientations can be changed, should they be? In a climate of hostility and derision, where acknowledging an LGB sexual orientation results in a loss of societal privilege, at the very least, the offer of "treatment" or "conversion" may serve to reinforce negative societal attitudes. It would be difficult to conduct such treatment without supporting the idea that something is wrong with the LGB person that warrants correction, and in doing so, reinforcing heterosexist bias. The LGB person becomes the focus of change, rather than the pathology of heterosexism and homophobia and its advocates. Furthermore, Haldeman (1994, 1999) argues that such treatments are not benign. In some studies, patients complain of significant discomfort following attempts to convert their sexual orientation. Such reactions include chronic depression, anxiety, intimacy avoidance, and sexual dysfunction.

APA, in an attempt to address ethical questions raised by the use of conversion therapies, approved (in a near unanimous vote) a resolution, "Appropriate Therapeutic Responses to Sexual Orientation" at its 105th Annual Convention in August 1997. The resolution holds clinicians responsible for rejecting depictions of LGB persons as mentally ill and acknowledges the role of societal prejudice in prompting many LGB persons to seek conversion therapies (Haldeman, 1999). Another aspect of the resolution requires clinicians to disclose the theoretical underpinnings and scientific basis for their interventions. As conversion therapies do not enjoy the support of empirical evidence, it would make it more difficult to ethically offer them as suitable responses to an LGB client's distress about their sexual orientation. APA commentary on conversion therapies states, "these findings suggest that efforts to repair homosexuals are nothing more than social prejudice garbed in psychological accoutrements" (Welch, 1990, in Haldeman, 1991, p. 160).

Different stages of the coming out process may be extremely anxiety provoking. The discovery or confirmation of a lesbian or gay sexual orientation may be experienced as frightening, ego alien, and/or a source of great subjective distress (Gonsiorek, 1982, 1988). It is also a frequently cited reason LGB persons give for seeking psychological services. As such, it may lead to the expression of behaviors or feelings which resemble symptoms of severe psychopathology. However, the presence of this behavior, particularly under these circumstances, does not warrant the immediate assumption that an underlying psychiatric disorder is present (Gonsiorek, 1982). For many clients, the clinician is observing the individual's difficult struggle confronting the real nature of their sexual orientation, perhaps for the first time. For other clients, however, the intense stress inherent in this process may, like any other stressor, precipitate the expression of serious underlying psychiatric disorders. In either case, a person is required to manage this realistically stressful life event. Clients who have underlying psychopathology must do this with fewer emotional resources. Another diagnostic variation is observed in the client who is consumed with great anxiety or fears of delusional proportions that they may be gay or lesbian despite the absence of any rational foundation for such beliefs. It is important to attend to diagnostic distinctions in these situations. A comprehensive discussion of differential diagnostic issues can be found in Gonsiorek (1982).

The relative invisibility of gay men and lesbians allows them to "pass" as heterosexual. Passing in the LGB community is also referred to as being closeted. Both "passing" and being "out" have their own distinct negative and positive consequences. Passing may conceal group members from other LGB persons. Both passing and being out are accompanied by varying types and degrees of psychological demands and the stress that is a result of those demands. Gay men and lesbians pass or are closeted when they do not challenge the assumption that they are heterosexual or when they actively conceal their sexual orientation.

Passing can be an adaptive coping strategy when used strategically. It has been used historically by ethnic minorities in threatening situations and can be an adaptive survival tool. It was often effective in helping its users avoid or escape imminent harm or to obtain goods, services, or jobs which would be otherwise inaccessible to them because of discrimination (Greene, 1994a, 1994b, 1994c). When used as a long-term survival tool, however, it deprives its user of the spontaneity required for authenticity in interpersonal relationships. Individuals live with a consistent pressure to conceal major aspects of their lives and may often live with the constant dread of being discovered (Bradford et al., 1994; Greene, 1992). When passing is accompanied by the belief that being gay or lesbian is a sign of inferiority or pathology, it represents an expression of internalized homophobia. Lesbians and gay men who pass or are forced to remain closeted, particularly when it is dangerous not to do so, are confronted with a chronic stressor which can leave them at risk for negative psychological outcomes (Bradford et al., 1994; Fassinger, 1991; Trippet, 1994).

Not only are lesbians and gay men unprotected by legislation which prohibits discrimination based on group membership, some legislation exists which actually requires discrimination against them, for example, military regulations.

Bisexual persons are frequently the objects of hostility from gay men, lesbians, and heterosexuals as well. Bohan (1996) uses the term biphobia to describe the discomfort people experience about bisexuality. The negative treatment bisexual men and women receive from the lesbian and gay community can include their exclusion from LGB social events, as well as a denial of the legitimacy of bisexual identity. They may be perceived by members of the gay and lesbian community as persons who are really gay or lesbian, but who conceal or deny their true identity to avoid the stigma of being gay, and to allow them to benefit from heterosexual privilege (Bohan, 1996; Garnets & Kimmel, 1991).

Their bisexuality may also be perceived as an expression of their internalized homophobia. Bohan (1996) argues that many lesbians and gay men consider the presence of a singular commitment to one's own sex as a statement of loyalty to lesbians and gay men, the absence of such a commitment as disloyal, and perhaps the fear that such a person will ultimately reject their lesbian or gay sexual orientation if it is opportune to do so. Conversely, they may be viewed by heterosexuals as less normal or inferior to persons who are exclusively heterosexual (Garnets & Kimmel, 1991).

Heterosexism and its concomitant negative stereotypes of gay men and lesbians has been so much a part of the definition of psychological normalcy that it predisposes practitioners to make a range of erroneous assumptions about LGB clients. Many of these assumptions represent ideas of questionable validity, which they have not been sensitized to recognize during their training (Brown, 1996; CLGC, 1990, 1991; Garnets et al., 1991; Glassgold, 1992; Gonsiorek, 1991; Liddle, 1997; Markowitz, 1991; Morin, 1977).

One example of such errors is the presumption that clients are heterosexual until proven otherwise. This can be particularly problematic for individuals who are in the early stages of coming out, who lack a sense of clarity about their sexual orientation, who are distressed by the possibility that they may have an LGB sexual orientation, or who have internalized the dominant culture's heterosexist bias. The client may not necessarily disclose their LGB orientation to the therapist at all or they may do so at any point in treatment (Bradford et al., 1994; Fassinger, 1991; Garnets & Kimmel, 1991; Greene, 1994b; Markowitz, 1991; Liddle, 1997; Youngstrom, 1991). Like other factors which vary from person to person, sexual orientation may be centrally or distantly related to the presenting problem. Many individuals, however, may depend on the therapy process and the therapist's acceptance to help them navigate this uncharted and frequently anxiety-ridden course to self-acceptance (Greene, 1994b).

A former therapy client recounted: "not one of them (therapists) was prepared or brave enough to ask me the one question that might have saved me and two ex-wives a lot of pain ... Do you think you might be gay?" (Woolley, 1991, p. 30).

Errors in treatment can occur at either of two extremes and, of course, at any point along the continuum between those extremes. On the one hand, the therapist may minimize the importance of the client's sexual orientation and the negative impact of heterosexism on the client's life (Garnets et al., 1991; Greene, 1994b; Youngstrom, 1991). On the other hand, the therapist may tend to focus on the client's lesbian or gay sexual orientation as pathological and the source of all of the client's problems, ignoring matters that are more troubling to the client (Garnets et al., Markowitz, 1991). Similarly, therapists may focus on the client's sexual orientation in a voyeuristic manner, wishing to hear unnecessary details about the client's sexual practices for the therapist's own titillation, neglecting other areas of inquiry important to the client's treatment (Dworkin & Gutierrez, 1992; Garnets & Kimmel, 1991; Gonsiorek & Weinrich, 1991; Markowitz, 1991).

Many clients report finding that when they raise questions or attempt to explore their confusion about their sexual orientation, they find themselves confronted with the therapist's anxiety about the topic. This anxiety may take the form of the therapist's failure to ask questions or assist the client in explorations of these issues. Other clients report the therapist's outright denial that the client could be gay. This may be accompanied by the therapist actively discouraging the client from having or adopting a gay or lesbian sexual orientation (Brown, 1996; Garnets et al., 1991; Youngstrom, 1991) or selectively encouraging heterosexual relationships. On these occasions, many clients simply drop the issue. Others move on to another therapist. Still others leave therapy, sometimes the worse for the experience.

Generally, therapists who work with LGB persons must appreciate the individual client's dilemma in the context of real and not fantasised prejudice and discrimination. Heterosexism often assumes the form of physical violence directed at LGB persons (Fassinger, 1991; Herek & Berrill, 1992). Data on bias crimes suggest that 92% of lesbians and gay men report being targets of antigay verbal abuse or threats, and that 24% report physical attacks, of which some result in death (Herek, 1989). Gay men and lesbians face the routine tasks of assessing realistic dangers associated with divulging the nature of their sexual orientation. Many endure painful isolation from families of origin who do not accept their identity, as well as the painful process of rejection which often precedes it. They are challenged by their own internalized homophobia, which can negatively affect their psychological adjustment (Fassinger, 1991; Trippet, 1994; Weinberg, 1972).

Those who are out face the formidable task of negotiating nontraditional relationships and family structures with few models and little support for doing so in a ubiquitously hostile environment. In such a climate, many people who would consider expressions of racism or ethnic bigotry wholly inappropriate are quite comfortable expressing heterosexist bigotry and find support for doing so. These issues are further complicated for gay men and lesbians who are members of ethnic minority or other marginalized groups (Chan, 1992; Greene, 1994a, 1994b, 1994c, 1996a, 1996b, 1999; Morales, 1992).

Therapists must begin to assess the impact of a legacy of negative stereotypes about LGB persons on their own thinking, and they must do so before a client ever appears before them. Markowitz (1991) warns that therapists must acknowledge and understand the extent to which they have internalized society's negative depictions of gay men and lesbians. It is not sufficient to simply not believe the stereotypes and to be unbiased without an inquiry into their personal effects on deeper levels (Brown, 1996; Liddle, 1997; Markowitz, 1991).

Unexamined fears in the therapist may be triggered by a range of interactions with LGB clients. The client may, if angry, accuse the therapist of being homophobic. Conversely, if the therapist is supportive and understanding, some LGB clients may assume that they share sexual orientations. The therapist may assume a defensive posture, rather than exploring the client's feelings in an uncritical manner. The therapist may respond to the client's accusation that the therapist is homophobic by failing to set or maintain appropriate limits, as if this disproves the client's assumptions. Such behavior on the therapist's part is not done with the client's interests in mind, rather it is often the therapist's response to their own feelings of guilt or discomfort (Greene, 1994a, 1994b, 1994c).

This does not mean that whenever a client perceives homophobia in the therapist, it must be a distortion on the client's part. It means that it is always the clinician's responsibility to be familiar with their own personal feelings and attitudes toward LGB persons and the ways they may be manifested in the therapy process. The therapist must then assess the client's complaint in this context, while simultaneously exploring the range of conscious and unconscious purposes which may be served by the client's beliefs about the therapist (Brown, 1996; Greene, 1994b, 1999; Liddle, 1997).

The issue of therapist neutrality, common in psychodynamic therapies, must be managed delicately. Because LGB persons are frequently the objects of intense disapproval from both loved ones and society, a client may read a therapist's neutrality as disapproval. The therapist in these situations must also be careful not to advance their own agenda for the client, no matter how well intentioned. It is inappropriate to push the client toward acknowledging a gay or lesbian sexual orientation before the client is psychologically prepared to manage it. Rather, the therapist must be affirming and supportive of the client, regardless of the client's choice about their sexual orientation or the management of challenges associated with adopting LGB orientations. Furthermore, therapy with members of oppressed groups must address the pathology of the oppressive behavior rather than exclusively focusing on victims' responses to oppressive behavior. In this context, helping the client understand negative responses to LGB persons (heterosexist bias) as an example of societal pathology, and not the client's flaw or defect, must be an explicit part of the therapeutic work (Brown, 1996; Dahlheimer & Feigal, 1991; Greene, 1994b).

Just as the authentic inclusion of any previously excluded minority group forces the larger group to transform itself into a realistic reflection of the diversity of its members, the visible presence of LGB persons who seek psychological services forces us to rethink our traditional notions about personality development, developmental tasks and stressors, and about what kinds of constellations of persons constitute a couple, a marriage, or a family (Greene, 1994b).

10.09.6 LGB PERSONS AND FAMILY ISSUES

LGB persons come from diverse family constellations as do their heterosexual counterparts. They also require the same tangible and emotional support that all people need, usually provided for by family members. It may be said that strong family ties are even more crucial to LGB persons given the hostility and rejection they routinely confront in the outside world. They may not, however, presume the support that heterosexual family members facing any challenge might take for granted. Therapists should be attuned to the possibility that the absence of such support has more negative consequences for LGB persons than for their heterosexual counterparts. Some research suggests that White lesbians report that they are as much as three times more likely to depend on friends for support over family members (Bradford et al., 1994; Kurdek & Schmidt, 1987; Morgan, 1992). Other research suggests that ethnicity may play a role in the extent to which family or friends are more salient sources of support. Mays and Cochran (1986; 1988) report that African-American lesbians tend to depend on family members for tangible and emotional support to a greater degree than their White counterparts. This may contribute to making African-American lesbians more reluctant to come out to family members than their White counterparts.

Many lesbian and gay clients struggle at some point in their lives with crises related to maintaining secrecy about their sexual orientation, the true nature of their relationships, and the ongoing problems that are a consequence of that decision. An expected developmental issue involves the decision to come out to family members, and the consequences of that decision. Many clients may feel very conflicted about what they want to do and will express great confusion in making this decision. Some may even feel guilty about their disclosure if it upsets family members. The therapist in these situations must be careful not to reinforce the client's guilt by discouraging disclosure when it is appropriate or by suggesting that the family's rejection is justified.

While it may be expected that a family member's disclosure may upset some family members or disrupt family functioning, there is no uniform way in which families respond. Each client's family is different and the range and intensity of their responses will vary from acceptance, or perhaps disappointment, to outrage and outright rejection. Ethnicity, socioeconomic class, religious belief, age at the time of coming out, and other variables will affect the responses of family members as well as the consequences of those responses to the LGB person. Previous family dynamics, rivalries, close ties, and strengths will also affect the process differentially. Over time, some families will come to accept the lesbian or gay member's sexual orientation to varying degrees and others will not. Responses of families will be as diverse as the racial, ethnic, and class groups they represent. Clients may need to be reminded of the often lengthy and difficult process of coming out that they have previously negotiated. This process required that they reevaluate their previously accepted attitudes about LGB persons and come to accept and embrace their identity over time. Just as they required time and understanding to do so, as painful or infuriating as the process may be for them personally, their family members will require time as well. Strommen (1989), Liu and Chan (1996), Greene and Boyd-Franklin (1996), and Brown (1988) provide a detailed discussion of the complexities inherent in the coming out processes for diverse LGB persons and their families of origin

10.09.6.1 Gay and Lesbian Couples

LGB persons seek, form, and sustain relationships in a context where there is little support for them. Most lesbians and gay men grew up in families where their parents were heterosexual and where they had few useful role models for understanding what normal transitions and developmental periods in gay and lesbian relationships would be like (Garnets & Kimmel, 1991). For some couples, there may be a tendency to idealize relationships with members of the same gender. This belief may have some of its origins in the assumption that their gender similarity makes them the same, and that sameness presumes an ease in relating. While this may be true for some couples, it is certainly not generalizable to all of them. Gender similarity does not preclude conflicts, disagreements, or dysfunction in relationships and in some cases may even intensify it. Hence, many gay and lesbian couples are surprised and disappointed when they encounter problems or conflicts within their relationships, even when they are representative of the normal range of challenges for all couples. They may harbor the notion that they are the exception to their idealized image of lesbian or gay relationships, or conversely that all lesbian or gay relationships are doomed to fail. For example, some women believe that battering does not occur in lesbian relationships. They believe that women, unlike men, are not physically abusive. Hence, lesbians in battering relationships may be confused and unable to appropriately label such behavior out of the belief that it does not occur between women, despite the fact that it is happening to them. Renzetti (1992) documents the problem of partner abuse in lesbian relationships and offers a helpful analysis. Kanuha (1990) offers special insights into this problem as it occurs in lesbian women of color.

If lesbian and gay couples are out, their relationship may be constantly minimized and challenged. If they are not out the invisibility of their relationship can be just as problematic. The invisible relationship goes unnoticed. While this is in part the idea, it means that the legitimate status accorded relationships and the social and economic privileges that accompany them are lost. The special pressures and stressors that accompany them remain unacknowledged as well. Family members who perceive a relative to be single, or without the responsibility of a family of their own, may be more demanding of a relative's time or resources. They may fail to understand the responsibility to a partner or relationship which they do not see or one which they choose not to recognize. On this informal level, when the couple has a relationship problem, loss or termination of the relationship, responses to such problems from heterosexual peers or family members may be inconsistent with the realistic magnitude and significance of the loss, or with the attentiveness that would be given to a marital partner in a heterosexual relationship. Thus, members of couples may not derive appropriate support during critical times from resources commonly available to heterosexual couples. Lack of legal status for gay and lesbian relationships in most cities in the USA has direct economic consequences reflected in part in the loss of employee benefits that would normally be extended to a marital partner.

Therapists who are unfamiliar with gay and lesbian couples, and some couples themselves, may attempt to impose male–female models on them, as there are no similar models for same-sex couples (Markowitz, 1991). The myth that in lesbian and gay couples gender roles are reversed is not supported, rather it is contradicted by empirical research on couples. Peplau (1991) and Peplau and Cochran (1990) find that in most lesbian and gay relationships surveyed participants actively reject traditional gender roles as models. Appropriate models would need to take into account differences between male and female socialization and its effects on relationships in which both persons received the same gender socialization (Markowitz, 1991; Schreurs, 1993).

Peplau (1991) discusses a range of commonly believed stereotypes about gay and lesbian couples. Among them, gay men and lesbians neither wish to have, nor are they capable of forming, enduring relationships, most lesbians and gay men grow old unhappy and alone, their relationships are inferior imitations of heterosexual relationships, or that traditional husband and wife and other traditional gender roles are reversed. Such stereotypes are still believed by many therapists despite the absence of any credible evidence to support them and in the presence of a growing body of evidence which disputes such contentions. Bell and Weinberg (1978) report that 40–60% of the participants in their research were in steady relationships. Peplau and Cochran (1990) suggest that the invisibility of long-term couples in research can be attributed in part to sampling bias. LGB research subjects are often recruited in venues that over-represent young, White, urban participants, often in bars. They suggest that couples, particularly long-term partners, do not frequent bars as often as their younger, single counterparts. If older LGB people are not sufficiently sampled, we would not expect to see longer relationships among young subjects. The fact that LGB persons cannot legally marry also means that there are no marriage records that would document the existence of long-term partnerships rendering them less visible (Peplau & Cochran, 1990). Clunis and Green (1988) provide a detailed review of issues facing lesbian couples which may be particularly helpful to the therapist with no experience or training in this area. Kurdek (1994), Kurdek and Schmidt (1987), Peplau (1991), Peplau and Cochran (1990), and Modrcin and Wyers (1990) report on extensive studies of cohabiting gay, lesbian, and heterosexual couples. Findings suggest that lesbian and gay couples are no more prone to relationship dissatisfaction or problems than their heterosexual counterparts. While they may have special problems and challenges, they are no more likely to have marital problems or impoverished social support networks than their heterosexual counterparts and generally report satisfaction in their relationships.

10.09.6.2 Gay and Lesbian Parents

It has been traditionally, albeit incorrectly, presumed that gay men and lesbians do not wish to have children and that they do not make appropriate parents (Greene, 1990, 1994b). Nonetheless, gay men and lesbians become parents in many different ways, as do heterosexual parents. However, they do so amidst the pervasive assumption that their sexual orientation makes them inappropriate parenting figures. This assumption is often based on the false belief that they will increase the likelihood that their child will be gay, lesbian, or psychologically defective. Research on children of LGB parents does not support these contentions (Bigner & Bozett, 1989; Patterson, 1994, 1996). Patterson (1994) provides a systematic study of the behavioral adjustment, self concepts, and sex role identity of children of the lesbian baby boom.

Many LGB persons choose to limit their social world to other LGB people whenever possible. Those who have children, however, are forced to both interact with and negotiate systems, for example, schools, that they might otherwise choose to avoid. They find themselves presented with the dilemma of how to present their unique family constellation in ways that maintain its integrity, protect its privacy, and are sensitive to the child's needs as well. Those who are involved in shared custody arrangements with a former heterosexual spouse often face the realistic danger of having the courts remove their children from their custody if the true nature of their relationship is exposed. The problem of maintaining secrecy in families where a gay or lesbian parent is not out is complex. Care must be taken to determine how to create the privacy or secrecy the family needs without leaving children feeling so fearful or burdened that it interferes with their own emotional needs and development (Greene, 1994b).

Disclosure of lesbian or gay sexual orientation to children is an anxiety-laden process for the parent and child. Many of the dynamics discussed in Strommen (1989) and Brown (1988) can be applied to an understanding of problems which may arise in such disclosures to children. Bigner and Bozett (1987), Crawford (1987), Falk (1989), and Green and Bozett (1991) provide detailed reviews of the issues faced by gay and lesbian parents relevant to disclosure, family dynamics, clinical strategies, and pertinent research.

10.09.6.3 Gay and Lesbian Youth

Gay and lesbian youth force us to take a second look at traditional assumptions about developmental tasks and tensions during the adolescent period. Despite the highly sexualized nature of popular culture in the USA, most adolescents have little accurate information about sexuality at a time when they may be in greatest need of it. They have even less knowledge about gay and lesbian sexual orientation which is not pejorative.

Gonsiorek (1988), Herdt (1989), Savin-Williams (1989, 1990, 1996) describe a range of intense social and personal pressures facing gay and lesbian adolescents. All are complicated by the normative rigidity and intolerance of differences adolescents maintain toward themselves and others. Adolescents who disclose their concerns about their sexual orientation are almost sure to find themselves in the midst of conflict with other family members and peers as well. The negative stereotypes associated with gay and lesbian sexual orientations and the dominant culture's denials of sexuality during this developmental period intensify this conflict. This denial is manifested in a conspicuous absence of institutional support for these youngsters. Coping strategies may also include withdrawal from social activities, denial, overcompensation, emotional constriction, and self-destructive behavior (Gonsiorek, 1988; Savin-Williams, 1989, 1990, 1996).

Savin-Williams (1990, 1996) and Golden (1996) caution that there are important gender differences in the process of self-awareness and disclosure of LGB orientation. Golden (1996) tells us that self-labeling is more fluid in women over the lifespan than men and that the concept of sexual orientation may not be as stable for women as it seems to be for their male counterparts. Women are more likely to initially disclose their awareness of a lesbian sexual orientation in the context of an emotional and affectionate relationship with another woman, and are less likely to disclose prior to coming to a more personal, internal resolution for themselves (Savin-Williams, 1996). Men, on the other hand, are more likely to disclose earlier and to use external social and sexual activity to develop a sense of self-acceptance, a more external process (Savin-Williams, 1990, 1996).

Many adolescents do not reveal their concerns about their sexual orientation to their families or other people who are close to them out of realistic fears of rejection as well as punishment (Hersch, 1991). When they do, Savin-Williams (1996) observes that they are likely to tell a close friend first rather than a family member. They may tell family members as much as decades later or for some LGB persons, never. Some may often withdraw at a time when they are in critical need of support. The high rate of suicide among adolescents believed to be gay and lesbian underscores the dire implications of this problem and warrants the serious attention of mental health practitioners. Support groups, accessible health and social services, the availability of healthy role models, advocacy and education, and education with respect to sexually transmitted diseases, particularly AIDS, are suggested as needed services for gay and lesbian adolescents, a vulnerable population (Gonsiorek, 1988; Hersch, 1991).

The visibility of LGB persons as consumers of psychological services requires practitioners to rethink traditional definitions of family constellations. Practitioners in the 1990s may find themselves confronted by the family in crisis after an adolescent member has disclosed that they are lesbian or gay, or by the adolescent who is in the throes of deciding whether or not, or who, to tell. Similarly, they may be consulted by the wife or husband who has chosen or struggles with the decision about whether or not to leave a heterosexual marriage to come out. They may also be consulted by the confused and angry heterosexual spouse.

Practitioners may also be confronted with family and couple constellations with which they are totally unfamiliar. Aside from issues of custody and marital conflict in parents, it is not uncommon to be asked to treat a child or children, in both the public and private sectors, who must in these arrangements acknowledge this previously feared or unknown aspect of their parent's life. This may occur in conjunction with the sudden upheaval and breakup of their family unit. As gay men and lesbians actively embrace their lifestyles and choose to form their own families, therapists are frequently consulted to assist them in addressing difficulties in their own relationships, as well as making decisions to have and raise children in nontraditional situations. It is incumbent on therapists in these scenarios to be aware of the unwarranted but realistic level of prejudice and discrimination, particularly in custody hearings and court proceedings, against gay and lesbian parents that persists.

10.09.7 TRAINING

10.09.7.1 Hetrosexist Bias in the Delivery of Psychological Services

Educational institutions are not exempt from pervasive heterosexist bias in Western culture. In fact, they have often communicated and legitimized that bias. Institutions responsible for the training of clinical psychologists are no exception. Practicing psychologists, therefore, are not immune to the effects of the pervasive themes of heterosexist bias in the dominant culture. Such bias has often served as an integral part of the underpinnings of theoretical and research paradigms. Hence, this bias insidiously pervades research and practice despite the intentions of many well-meaning research scientists and practitioners. For example, psychoanalytic theories evolved out of traditional culturally-bound views of gender roles. Its theory of etiology of sexual orientation has its origins in views which privilege reproductive sexuality as the only healthy outcome of psychosexual development. Simoni (1999) reports on the review of 24 college-level textbooks in social, abnormal, and developmental psychology. Overall her study reveals inadequate coverage and segregated treatment of the topic of lesbian and gay sexual orientation. Notably, there was little attention given to the systematic incorporation of lesbians and gay men in contexts other than sexual orientation or sexual behavior, such as commitment, parenting, and other routine aspects of human behavior and relationships.

The work of Hooker (1957), Kinsey, Pomeroy and Martin (1948), Kinsey, Pomeroy, Martin, and Gebhard (1953) and others established that lesbian and gay sexual orientations were not synonymous with poor psychological adjustment. Despite this, ignorance and heterosexist bias, which formal training rarely addresses directly, continues to significantly influence the delivery of psychological services to gay and lesbian clients, as well as research in areas relevant to developing a better understanding of this group (Greene, 1994b).

Buhrke's (1989a) findings suggest that students receive little exposure to information relevant to the delivery of services to gay and lesbian clients during the course of their formal training. Garnets et al. (1991), Graham, Rawlings, Halpern, and Hermes (1984), Liddle (1997), and Youngstrom (1991) recommend that current research and findings on lesbians and gay men should be discussed in graduate programs, continuing education and inservice training, as well as in undergraduate courses. Liddle (1997) asserts that being unbiased is not sufficient to demonstrate competence or expertise in clinical work with LGB persons.

Since the late 1980s we have witnessed the growth of a significant body of psychological literature which appropriately addresses the aforementioned concerns. It has not, however, found its way into the mainstream of clinical psychology training programs and curriculums. Browning and Kain (1999), Buhrke (1989b), Chan (1996), Greene and Croom (1999), King (1988), and Simoni (1999) review a range of resources and strategies (see also Dworkin & Gutierrez, 1989; Stone, 1991) which may be useful in the design, implementation, and supplementation of traditional graduate and undergraduate psychology courses, rendering them more sensitive and relevant to LGB persons.

LGB sexual orientations are topics which most persons socialized in the USA have intense feelings about. Hence, it is important that psychologists learn to explore and understand these issues appropriately in both their clients and themselves. Few topics, however, are more scrupulously avoided in the formal training of clinical psychologists. Just as members of ethnic minority groups have been harmed by a legacy of racially stigmatizing psychological folklore, gay and lesbian clients are also harmed by negative heterosexist bias and the misinformation which pervades psychotherapy practice and has similarly pervaded psychological research. Just as ethnic minority clients can be harmed by the unexamined racism in the therapist or inherent in a research design, gay and lesbian clients can be similarly harmed by unexamined and untransformed heterosexist bias in those areas.

As greater numbers of gay and lesbian individuals and families seek counseling and psychotherapy, many therapists find themselves ill-equipped to provide services sensitive to what may be seen as a distinct cultural group. It is important that therapists be aware of the ways in which routine life stressors may be intensified for individuals who are actively discriminated against by the dominant culture and who, unlike members of oppressed ethnic groups, may not find mentoring, support from, or identification with family members. Such identification and sense of shared struggle with family members has been an important coping mechanism in the adaptive "banding together against the outside oppressor" for many members of oppressed ethnic groups. While family members of ethnic minority group members often teach their children to actively challenge the dominant culture's views of them, they are more likely to embrace heterosexist attitudes and join in the dominant culture's rejection of LGB persons.

The ethnic minority LGB person depends on the protective armoring of their family and ethnic community more than their White counterparts but may find themselves on the outskirts of that protective buffer against racism. In addition to the dominant culture's racism, LGB people of color are challenged by the additional stress of coping with the LGB community's racism, the dominant culture's heterosexism, as well as the heterosexist bias of their own ethnic group. This results in a complex inter-relationship of loyalties and estrangements (Chan, 1992; Greene, 1994a, 1994b, 1994c, 1996; Greene & Boyd-Franklin, 1996; Morales, 1992).

Clinicians in these encounters are confronted with the difficult and challenging task of disentangling characterological issues from the personal distress that results from the pressures of societal inequities and their effects on mental health. This cannot be accomplished without a realistic portrayal of the institutional barriers which regularly confront LGB people. This includes an understanding of the role of heterosexist bias in many psychological treatment and research paradigms. Despite the extreme hostility and discrimination that LGB people routinely encounter, they are not inevitable psychological cripples. The resilience and unique skills which many out gay men and lesbians develop, despite the hostility and barriers they face, must be more fully explored as well as the negative outcomes (Greene 1994b, 1999). While the realistic need for secrecy, and often justifiable suspicion LGB persons hold for mental health practitioners and research scientists often complicates research endeavors, they are not impossible to negotiate.

Garnets et al. (1991) suggest the continued use of survey research as a means of reducing sampling bias found in many studies on lesbians and gay men. Caution must exercised when making generalizations from samples which are nonrepresentative and replication of results is important (Gonsiorek, 1991).

While APA has charged psychologists with the task of challenging heterosexist bias in both research and practice, it has provided little in their formal training to assist them in making the necessary shifts in practice as well as attitude. Just as clients are presumed to be heterosexual, faculty members and psychologists in training are similarly regarded. Heterosexist bias in training programs continue to make it unsafe for some faculty members and lesbian and gay students to divulge their sexual orientation. The locus and levels of danger assume different forms.

D'Augelli (1989) observes that LGB students and faculty members continue to suffer from discrimination in the workplace. Three-quarters of the lesbians and gay men in his sample in university environments reported experiences of verbal abuse, 26% reported being threatened with violence, and 17% reported having their personal property damaged. Kitzinger (1996) reports that British studies have similar findings. Tierney (1992), in a report from the University of Oregon, finds the university environment neither consistently safe, tolerant, nor academically inclusive of LGB persons or of research relevant to them. The study concurred with D'Augelli (1989) and Kitzinger (1996) who report higher incidences of verbal and physical attacks on LGB students than on their heterosexual counterparts. Kitzinger (1996) reports that the British Psychological Society has refused on three occasions to consider the establishment of a specific forum or unit devoted to the psychological study of LGB issues.

For many LGB students in graduate programs, disclosing their sexual orientation before their formal training is completed puts them at risk for discrimination that may be subtle but nonetheless can adversely affect their training status within their programs. The requirement to manage the heterosexism of faculty members, supervisors, and fellow students can leave LGB students with additional challenges that are neither appropriate nor shared by their heterosexual counterparts. It also deprives them of the opportunity to authentically explore their own feelings, ideas, and appropriate use of self as a clinical instrument in their work with clients. Hence, there is a failure to appropriately explore countertransference problems as well as issues which may arise when the therapist and client are both LGB persons (Greene, 1994b).

Dahlheimer and Feigal (1991) suggest that therapists may use a variety of techniques in conjunction with didactic training to develop greater sensitivities toward gay and lesbian clients. They suggest the use of role-playing and the review of case material for insensitive approaches. Case material may be used that requires students or supervisees to review data or client histories and determine how they might view the case if the heterosexual client were an LGB person (Greene, 1994b).

Another suggestion directs therapists as a part of their training to purchase a gay or lesbian magazine (which is clearly labeled as such), to carry it visibly throughout the course of their day, to monitor their feelings about doing so as well as their internal and overt responses to the inquiries and responses of others, and to discuss these experiences in supervision. It is suggested that these exercises be used to develop a subjective sense of the level of self-consciousness, shame, fear, vulnerability, or anger that many of their clients must manage routinely (Dahlheimer & Feigal, 1991). Participation in such exercises does not mean that doing so will give a heterosexual therapist a complete understanding of what life is like for their LGB client. Rather it is intended to provide the therapist with some realistic, albeit minimal, understanding of what their client must manage on an ongoing basis without the luxury of terminating the "exercise" when they are uncomfortable or in danger (Greene, 1994b). Browning and Kain (1999) present a range of exercises and other strategies that may be useful in training.

Practitioners who lack formal training in or experience working with LGB clients and their families are advised to use cogroup and cofamily therapist arrangements; others may use peer and/or other forms of supervision with practitioners who have this professional experience and training as an additional means of augmenting their skills (Liddle, 1997).

David Scasta, editor of the *Journal of Gay and Lesbian Psychotherapy* (Markowitz, 1991), observes that the treatment of gay men and lesbians is a specialized field that requires a heightened level of self-awareness in the clinician and a commitment to being educated about gay and lesbian issues. It is important that all current and prospective therapists examine their conscious and unconscious levels of heterosexist bias to assure that they do not intrude into their client's therapy (Brown, 1996; Liddle, 1997).

10.09.8 FUTURE DIRECTIONS

As we approach the twenty-first century, clinical psychology is confronted with the need to be more inclusive if it is to serve the needs of consumers of psychological services appropriately. It is also challenged with the need to create theoretical paradigms and develop research that is a reflection of a wider range of human behavior and realities.

Chan (1996) asserts that academic institutions play a major role in perpetuating heterosexist bias and can play a major role in eliminating it. Listing sexual orientation in the nondiscrimination policies of institutions and seriously implementing that policy is one step. The assumption of a proactive, rather than reactive, affirmative stance among the leadership of academic institutions is an important means of establishing and implementing such policies.

A new line of inquiry must include acknowledgment, exploration, and understanding of the role of heterosexist bias in psychological practice, research, and the development of theoretical paradigms. It must also include international perspectives on assessments of the role of negative attitudes toward LGB sexual orientations on the development of identities in lesbians, gay men, bisexual, and heterosexual persons. LGB research must develop the means of sampling a wider range of LGB persons if studies are to be more representative of the population of LGB persons. This information is critical to the establishment of a reliable database as well as to the development of culturally and contextually sensitive methods of understanding the complex issues that arise when treating a diverse range of LGB clients (Croom, in press).

It is essential that the treatment of LGB clients and research with this population be given formal and explicit attention in the training of clinical psychologists. Initially, the invisibility of LGB issues that should be integrated in and across disciplines must be addressed. This can assume many different forms including the incorporation of required course work and clinical supervision (Buhrke, 1989a; Greene, 1994b; Liddle, 1997). The active recruitment of psychologists with research, supervision, teaching, and other forms of expertise and interest in this area and the inclusion of openly gay or lesbian faculty and staff members in academic and training institutions must also be accomplished in order to meet the goals of competent practice.

This may also be accomplished by supporting the creation of LGB scholarship by rewarding (rather than punishing or failing to prioritize) course development, providing tangible resources, supporting the acquisition of state-of-the-art publications and collections in university libraries, and providing support for interdisciplinary LGB research and teaching that develops scholarship from diverse perspectives (Chan, 1996; Simoni, 1999). Simoni (1999) advocates the identification, definition, and confrontation of heterosexism in psychology as the explicit responsibility of an effective instructor in the academy.

Accreditation teams must be serious in their efforts to hold training institutions, practicum, externship, and internship sites accountable for upholding meaningful standards of competence. Establishing requirements for demonstrating proficiency as an integral requirement for graduate comprehensive examinations, licensing, and certification should be undertaken as well. This can be accomplished by including relevant items on LGB issues on licensing, certification, and advanced practice specialty examinations (i.e., American Board of Professional Psychology).

10.09.9 SUMMARY

Sexual orientation is an important and complex psychological variable that has long been an object of scientific scrutiny. Its study has undergone critical change since the late 1970s. However, there is still much to be done. Potgieter (1997) documents the long history of struggle against racism and heterosexism in South Africa. In the fall of 1995, the first gay and lesbian colloquium took place in South Africa. In its final constitution, South Africa became the first country to include sexual orientation in its nondiscrimination clause. Since the late 1980s, the South African National Gay and Lesbian Coalition has become a prominent and active lobby for lesbian and gay rights. Even in quarters where change is slow, change takes place nonetheless. To collectively promote more accurate understandings of LGB sexual orientations, dichotomous and categorical attempts to classify it must be abandoned. Our task is the development of broader, more inclusive questions that explore the complex origins and determinants of all affectional and erotic attractions as they interact with other salient aspects of human beings as collective reflections of a healthy range of human diversity.

10.09.10 REFERENCES

American Psychological Association (1994). *Publication Manual of the American Psychological Association* (4th ed). Washington, DC: Author.

Ames, L. J. (1996). Homo-phobia, homo-ignorance, homo-hate: Heterosexism and aids. In E. Rothblum & L. Bond (Eds.), *Preventing heterosexism and homophobia* (pp. 239–265). Thousand Oaks, CA: Sage.

Bailey, J. M., & Pillard, R. C. (1991). A genetic study of male sexual orientation. *Archives of General Psychiatry, 48,* 1089–1096.

Bell, A. P., & Weinberg, M. S. (1978). *Homosexuality: A study of diversity among men and women.* New York: Simon & Schuster.

Bell, A. P., Weinberg, M. S., & Hammersmith, S. K. (1981). *Sexual preference: Its development in men and women.* Bloomington, IN: Indiana University Press.

Bem, D. J. (1996). Exotic becomes erotic: A developmental theory of sexual orientation. *Psychological Review, 103,* 320–335.

Bem, S. J. (1993). *The lenses of gender: Transforming the debate on sexual inequality.* New Haven, CT: Yale.

Bigner, J. J., & Bozett, F. W. (1989). Parenting by gay fathers. *Marriage and Family Review, 14,* 155–175.

Bohan, J. (1996). *Psychology and sexual orientation: Coming to terms.* New York: Routledge.

Boyd-Franklin, N. (1989). *Black families in therapy: A multisystems approach.* New York: Guilford Press.

Bradford, J., & Ryan, C. (1988). *The National Lesbian Health Care Survey: Final Report.* National Lesbian and Gay Health Foundation.

Bradford, J., Ryan, C., & Rothblum, E. (1994). National lesbian health care survey: Implications for mental health care. *Journal of Consulting and Clinical Psychology, 62,* 228–242.

Brown, L. S. (1988). Lesbians, gay men and their families: Common clinical issues. *Journal of Gay and Lesbian Psychotherapy, 1,* 65–77.

Brown, L. (1989). New voices, new visions: Toward a lesbian and gay paradigm for psychology. *Psychology of Women, 13,* 445–458.

Brown, L. (1996). Preventing heterosexism and bias in psychotherapy and counseling. In E. Rothblum & L. Bond (Eds.), *Preventing heterosexism and homophobia* (pp. 36–58). Thousand Oaks, CA: Sage.

Browning, C., & Kain, C. (1999). Teaching lesbian, gay and bisexual psychology: Issues and resources. In B. Greene & G. L. Croom (Eds.), *Psychological perspectives on lesbian, gay and bisexual issues Vol 5—Education, research and practice in lesbian, gay, bisexual and transgendered psychology: A resource manual.* Thousand Oaks, CA: Sage.

Buhrke, R. A. (1989a). Female student perspectives on training in lesbian and gay issues. *The Counseling Psychologist, 17,* 629–636.

Buhrke, R. A. (1989b). Incorporating lesbian and gay issues into counselor training: A resource guide. *Journal of Counseling and Development, 68,* 77–80.

Byne, W. (1995). Science and belief: Psychobiological research and sexual orientation. *Journal of Homosexuality, 28,* 303–344.

Cass, V. C. (1984). Homosexual identity formation: A theoretical model. *Journal of Sex Research, 20,* 143–167.

Chan, C. S. (1989). Issues of identity development among Asian-American lesbians and gay men. *Journal of Counseling and Development, 68,* 16–20.

Chan, C. S. (1992). Cultural considerations in counseling Asian American lesbians and gay men. In S. Dworkin & F. Gutierrez (Eds.), *Counseling gay men and lesbians: Journey to the end of the rainbow* (pp. 115–124). Alexandra, VA: American Association for Counseling and Development.

Chan, C. S. (1996). Combating heterosexism in educational institutions: Structural changes and strategies. In E. Rothblum & L. Bond (Eds.), *Preventing heterosexism and homophobia* (pp. 20–35). Thousand Oaks, CA: Sage.

Clunis, D. M. & Green, G. D. (1988). *Lesbian couples.* Seattle, WA: Seal Press.

Coleman, E. (1981/1982). Developmental stages of the coming out process. *Journal of Homosexuality, 7,* 31–43.

Committee on Lesbian and Gay Concerns (1986). APA policy statement on lesbian and gay issues. Washington, DC: American Psychological Association.

Committee on Lesbian and Gay Concerns (1990). Final report of the Task Force on Bias in Psychotherapy with Lesbians and Gay Men. Washington, DC: American Psychological Association.

Crawford, S. (1987). Lesbian families: Psychosocial stress and the family building process. In Boston Lesbian Psychologies Collective (Ed.), *Lesbian psychologies* (pp. 195–214). Chicago: University of Illinois Press.

Croom, G. L. (in press). Lesbian, gay and bisexuals of color: A challenge to representative sampling in empirical research. In B. Greene & G. L. Croom (Eds.), *Psychological perspectives on lesbian, gay and bisexual issues, Vol 5—Education, research and practice in lesbian, gay, bisexual and transgendered psychology: A resource manual.* Thousand Oaks, CA: Sage.

Dahlheimer, D., & Feigal, J. (1991, Jan.–Feb.). Bridging the gap. *The Family Therapy Networker, 15,* 26–29; 31–35.

D'Augelli, A. (1989). Lesbian and gay men's experiences of discrimination and harassment in a university community. *American Journal of Community Psychology, 17,* 317–321.

D'Augelli, A. (1994). Lesbian and gay male development: Steps toward an analysis of lesbians' and gay men's lives. In B. Greene & G. Herek (Eds.), *Lesbian and gay psychology: Theory, research and clinical applications* (pp. 118–132). Thousand Oaks, CA: Sage.

DeCrescenzo, T. (1985). Homophobia: A study of attitudes of mental health professionals toward homosexuality. In R. Schoenberg, R. Goldberg, & D. Shore (Eds.), *With compassion toward some: Homosexuality and social work in America* (pp. 115–136). New York: Harrington Park Press.

Diamond, M. (1993). Some genetic considerations in the development of sexual orientation. In M. Huang, R. E. Whalen, C. Aron, & K. Olsen (Eds.), *The development of sex differences and similarities in behavior* (pp. 291–309). Dordrecht, The Netherlands: Kluwer.

Dworkin, S., & Gutierrez, F. (Eds.) (1989, Sept./Oct.). Special Issue: Gay, lesbian and bisexual issues in counseling. *Journal of Counseling and Development, 68,* 6–96.

Dworkin, S., & Gutierrez, F. (1992). Opening the closet door. In S. Dworkin & F. Gutierrez (Eds.), *Counseling gay men and lesbians: Journey to the end of the rainbow* (pp xvii–xxvii). Alexandria, VA: American Association of Counseling and Development.

Falk, P. J. (1989). Lesbian mothers: Psychosocial assumptions in family law. *American Psychologist, 44,* 941–947.

Fassinger, R. E. (1991, April). The hidden minority: Issues and challenges in working with lesbian women and gay men. *The Counseling Psychologist, 19,* 157–176.

Freud, S. (1922/1959). *Group psychology and the analysis of the ego* (James Strachey, Ed. and trans.). London: Hogarth Press. (Original work published in 1922).

Fygetakis, L. (1997). Greek American Lesbians: Identity odysseys of honorable good girls. In B. Greene (Ed.), *Ethnic and cultural diversity among lesbians and gay men* (pp. 152–190). Thousand Oaks, CA: Sage.

Garnets, L., Hancock, K. A., Cochran, S. D., Goodchilds, J., & Peplau, L. A. (1991). Issues in psychotherapy with lesbians and gay men: A survey of psychologists. *American Psychologist, 46,* 964–972.

Garnets, L., & Kimmel, D. (1991). Lesbian and gay male dimensions in the psychological study of human diversity. In J. Goodchilds (Ed.), *Psychological perspectives on human diversity in America: Master lectures* (pp. 143–192). Washington, DC: American Psychological Association.

Giusti, I., & Katz, R. (1992). *The history of the organization of gay and lesbian psychology: How it all began.* New York: Unpublished manuscript.

Glassgold, J. (1992). New directions in dynamic theories of lesbianism: From psychoanalysis to social constructionism. In J. Chrisler & D. Howard (Eds.), *New directions in feminist psychology: Practice, theory, and research* (pp. 154–164). New York: Springer.

Golden, C. (1996). Whats in a name? Sexual self-identification among women. In R. C. Williams & K. M. Cohen (Eds.), *The lives of lesbians, gays and bisexuals: Children to adults* (pp. 229–249). Fort Worth, TX: Harcourt Brace.

Gonsiorek, J. (1977). Psychological adjustment and homosexuality. *Social and Behavioral Sciences Documents, MS 1478.* San Raphael, CA: Select Press.

Gonsiorek, J. (1982). The use of diagnostic concepts in working with gay and lesbian populations. In J. Gonsiorek (Ed.), *Homosexuality and psychotherapy: A practitioners handbook of affirmative models* (pp. 9–20). Thousand Oaks, CA: Sage.

Gonsiorek, J. (1988). Mental health issues of gay and lesbian adolescents. *Journal of Adolescent Health Care, 9,* 114–122.

Gonsiorek, J. (1991). The empirical basis for the demise of the illness model of homosexuality. In J. Gonsiorek & J. Weinrich (Eds.), *Homosexuality: Research implications for public policy* (pp. 115–136). Thousand Oaks, CA: Sage.

Gonsiorek, J., & Weinrich, J. (1991). The definition and scope of sexual orientation. In J. Gonsiorek & J. Weinrich (Eds.), *Homosexuality: Research implications for public policy* (pp. 1–12). Thousand Oaks, CA: Sage.

Graham, D., Rawlings, E. I., Halpern, H. S., & Hermes, J. (1984). Therapists' needs for training in counseling lesbians and gay men. *Professional Psychology, 15,* 482–496.

Green, G. D., & Bozett, F. (1991). Lesbian mothers and gay fathers. In J. Gonsiorek & J. Weinrich (Eds.), *Homosexuality: Research implications for public policy* (pp. 197–214). Thousand Oaks, CA: Sage.

Greene, B. (1990, December). African American lesbians. *BG Magazine, 26,* 6.

Greene, B. (1992). Black feminist psychotherapy. In E. Wright (Ed.), *Feminism and psychoanalysis: A critical dictionary* (pp. 34–35). Oxford, UK: Basil Blackwell.

Greene, B. (1994a). Lesbian women of color. In L. Comas-Diaz & B. Greene (Eds.), *Women of color: Integrating ethnic and gender identities in psychotherapy* (pp. 389–427). New York: Guilford Press.

Greene, B. (1994b). Lesbian and gay sexual orientations: Implications for clinical training, practice and research. In B. Greene & G. Herek (Eds.), *Psychological Perspectives on lesbian and gay issues. Vol. 1—Lesbian and gay psychology: Theory, research & clinical applications* (pp. 1–24). Thousand Oaks, CA: Sage.

Greene, B. (1994c). Ethnic minority lesbians and gay men: Mental health and treatment issues. *Journal of Consulting and Clinical Psychology, 62*(2), 243–251.

Greene, B. (1996a). Lesbians and gay men of color: The legacy of ethnosexual mythologies in heterosexism. In E. Rothblum & L. Bond (Eds.), *Preventing heterosexism and homophobia* (pp. 59–70). Thousand Oaks, CA: Sage.

Greene, B. (1996b). Psychotherapy with African American women: Considering diverse identities and societal barriers. *Annals of the New York Academy of Sciences-Women & Mental Health, 789,* 191–210.

Greene, B. (1997, June). Psychotherapy with African-American women: Integrating feminist and psychodynamic models. *Journal of Smith College Studies in Social Work—Theoretical, Research, Practice and Educational Perspectives for Understanding and Working with African-American Clients, 67*(3), 299–322.

Greene, B. (1999). Beyond heterosexism and across the cultural divide: Privilege, invisibility and future lesbian, gay and bisexual psychology. In B. Greene & G. L. Croom (Eds.), *Psychological perspectives on lesbian, gay and bisexual issues. Vol. 5—Education, research & practice in lesbian, gay, bisexual and transgendered psychology: A resource manual.* Thousand Oaks, CA: Sage.

Greene, B., & Boyd-Franklin, N. (1996). African American lesbians: Issues in couples therapy. In J. Laird & R. J. Green (Eds.), *Lesbians and gay men in couples and families: A handbook for therapists* (pp. 251–271). San Francisco: Jossey-Bass.

Greene, B., & Croom, G. L. (Eds.) (1999). *Psychological Perspectives on Lesbian, Gay, and Bisexual and Issues, Vol. 5—Education, Research and Practice in Lesbian, Gay, Bisexual and Transgendered Psychology: A Resource Manual.* Thousand Oaks, CA: Sage.

Haldeman, D. (1991). Sexual orientation conversion therapy for gay men and lesbians: A scientific examination. In J. Gonsiorek & J. Weinrich (Eds.), *Homosexuality: Research implications for public policy* (pp. 149–160). Newbury Park, CA: Sage.

Haldeman, D. (1994). The practice and ethics of sexual orientation conversion therapy. *Journal of Consulting and Clinical Psychology, 62,* 221–227.

Haldeman, D. (1999). Therapeutic responses to sexual orientation: Psychology's evolution. In B. Greene & G. L. Croom (Eds.), *Psychological perspectives on lesbian, gay and bisexual issues, Vol. 5—Education, research and practice in lesbian, gay, bisexual and transgendered psychology: A resource manual*. Thousand Oaks, CA: Sage.

Hamer, D. H., Hu, S., Magnuson, V. L., Hu, N., & Pattatucci, A. (1993). A linkage between DNA markers on the X chromosome and male sexual orientation. *Science, 261,* 321–327.

Haumann, G. (1995). Homosexuality, biology and ideology. *Journal of Homosexuality, 28,* 57–77.

Herdt, G. (1989). Gay and lesbian youth. *Journal of Homosexuality, 17,* 1–4.

Herek, G. M. (1989). Hate crimes against lesbians and gay men. *American Psychologist, 44,* 948–955.

Herek, G. M., & Berrill, K. (Eds.) (1992). *Hate crimes: Confronting violence against lesbians and gay men*. Newbury Park, CA: Sage.

Hersch, P. (1991, Jan.–Feb.). Secret lives: Lesbians and gay teens in fear of discovery. *The Family Therapy Networker, 15,* 36–39; 41–43.

Hooker, E. (1957). The adjustment of the male overt homosexual. *Journal of Projective Techniques, 21,* 18–31.

Isay, R. A. (1990). Psychoanalytic theory and psychotherapy of gay men. In D. P. McWhirter, S. A. Sanders, & J. M. Reinisch (Eds.), *Homosexuality/heterosexuality* (pp. 283–301). New York: Oxford University Press.

Jones, E. (1965). *The life and work of Sigmund Freud*. New York: Basic Books.

Jones, F. (1997, March). *Eloquent anonymity*. [Review of the book *Lush Life: A Biography of Billy Strayhorn*]. *Readings: A Journal of Reviews and Commentary in Mental Health, 12*(1), 10–14.

Kanuha, V. (1990) Compounding the triple jeopardy: Battering in lesbian of color relationships. In L. S. Brown & M. P. P. Root (Eds.), *Diversity and complexity in feminist therapy* (pp. 169–184). New York: Haworth Press.

Kashak, E. (1992). *Engendered lives: A new psychology of women's experience*. New York: Basic Books.

King, N. (1988). Teaching about lesbians and gays in the psychology curriculum. In P. A. Bronstein & K. Quina (Eds.), *Teaching a psychology of people: Resources for gender and sociocultural awareness* (pp. 168–174). Washington, DC: American Psychological Association.

Kinsey, A., Pomeroy, W. B., & Martin, C. E. (1948). *Sexual behavior in the human male*. Philadelphia: Saunders.

Kinsey, A., Pomeroy, W. B., Martin, C. E., & Gebhard, P. H. (1953). *Sexual behavior in the human female*. Philadelphia: Saunders.

Kirk, M., & Madsen, H. (1996). A field trip to straight America. In K. Rosenblum & T.-M. C. Travis (Eds.), *The meaning of difference: American constructions of race, sex and gender, social class, and sexual orientation* (pp. 400–412). New York: McGraw-Hill.

Kitzinger, C. (1987). *The social construction of lesbianism*. London: Sage.

Kitzinger, C. (1996). Speaking of oppression: Psychology, politics and the language of power. In E. Rothblum & L. Bond (Eds.), *Preventing heterosexism and homophobia* (pp. 3–19). Thousand Oaks, CA: Sage.

Kurdek, L. A. (1994). The nature and correlates of relationship quality in gay, lesbian and heterosexual cohabiting couples. In B. Greene & G. Herek (Eds.), *Lesbian and gay psychology: Theory, research and clinical applications* (pp. 133–155). Thousand Oaks, CA: Sage.

Kurdek, L. A., & Schmidt, J. P. (1987). Perceived emotional support from family and friends in members of homosexual, married, and heterosexual cohabiting couples. *Journal of Homosexuality, 14,* 57–68.

LeVay, S. (1991). A difference in hypothalamic structure between heterosexual and homosexual men. *Science, 253,* 1034–1037.

Liddle, B. J. (1997). Gay and lesbian clients' selection of therapists and utilization of therapy. *Psychotherapy, 34,* 11–18.

Liu, P., & Chan, C. S. (1996). Lesbian, gay and bisexual Asian Americans and their families. In J. Laird & R. J. Green (Eds.), *Lesbians and gays in couples and families* (pp. 137–152). San Francisco: Jossey-Bass.

Markowitz, L. M. (1991, Jan.–Feb.). Homosexuality: Are we still in the dark? *The Family Therapy Networker, 15,* 26–29; 31–35.

Mays, V. M., & Cochran, S. D. (1986, August). Relationship experiences and the perception of discrimination by Black lesbians. Paper presented at the 94th Annual Convention of the American Psychological Association, Washington, DC.

Mays, V.M., & Cochran, S. D. (1988). The Black women's relationship project: A national survey of Black lesbians. In M. Shernoff & W. Scott (Eds.), *The sourcebook on lesbian/gay healthcare* (2nd ed., pp. 54–62). Washington, DC: National Lesbian and Gay Health Foundation, Inc.

McGuire, T. R. (1995). Is homosexuality genetic? A critical review. *Journal of Homosexuality, 28,* 115–145.

McWhirter, D. P. (1990). Homosexuality/Heterosexuality: An overview. In D. P. McWhirter, S. A. Sanders, & J. M. Reinisch (Eds.), *Homosexuality/heterosexuality* (pp. xiv–xxvii). New York: Oxford University Press.

Meredith, R. L., & Reister, R. W. (1980). Psychotherapy responsibility and homosexuality: Clinical examination of socially deviant behavior. *Professional Psychology, 11,* 174–193.

Meyer-Bahlberg, H. (1995). Psychoneuroendocrinology and sexual pleasure: The aspect of sexual orientation. In P. R. Abramson & S. D. Pinkerton (Eds.), *Sexual nature/sexual culture* (pp. 135–153). Chicago: University of Chicago Press.

Modrcin, M. J., & Wyers, N. L. (1990). Lesbian and gay couples: Where they turn when help is needed. *Journal of Gay and Lesbian Psychotherapy, 1,* 89–104.

Money, J. (1988). *Gay, straight and in between: The sexology of erotic orientation*. New York: Oxford University Press.

Morales, E. (1992). Counseling Latino gays and Latina lesbians. In S. Dworkin & F. Gutierriez (Eds.), *Counseling gay men and lesbians: Journey to the end of the rainbow* (pp. 125–139). Alexandria, VA: American Association for Counseling and Development.

Morgan, K. S. (1992). Caucasian lesbians use of psychotherapy: A matter of attitude? *Psychology of Women Quarterly, 16,* 127–130.

Morgan, K. S., & Eliason, M. J. (1992). The role of psychotherapy in Caucasian lesbians' lives. *Women & Therapy, 13,* 27–52.

Morin, S. (1977). Heterosexual bias in psychological research on lesbianism and male homosexuality. *American Psychologist, 32,* 629–637.

Morin, S., & Rothblum, E. (1991). Removing the stigma: Fifteen years of progress. *American Psychologist, 46,* 947–949.

Niesen, J. P. (1990). Heterosexism: Redefining homophobia for the 1990s. *Journal of Gay and Lesbian Psychotherapy, 1,* 21–23.

Parker, D. A., & DeCecco, J. P. (1995). Sexual expression: a global perspective. *Journal of Homosexuality, 28,* 427–430.

Patterson, C. (1994). Children of the lesbian baby boom. In B. Greene & G. Herek (Eds.), *Lesbian and gay psychology: Theory, research and clinical Applications* (pp. 156–175). Thousand Oaks, CA: Sage.

Patterson, C. (1996). Lesbian mothers and their children:

Findings from the Bay Area Study. In J. Laird & R. J. Green (Eds.), *Lesbians and gays in couples and families* (pp. 420–437). San Francisco: Jossey-Bass.

Peplau, L. A. (1991). Lesbian and gay relationships. In J. Gonsiorek & J. Weinrich (Eds.), *Homosexuality: Research implications for public policy* (pp. 177–196). Newbury Park, CA: Sage.

Peplau, L. A., & Cochran, S. (1990). A relationship perspective on homosexuality. In D. P. McWhirter, S. A. Sanders, & J. M. Reinisch (Eds.), *Homosexuality/heterosexuality* (pp. 321–349). New York: Oxford University Press.

Potgieter, C. (1997). From apartheid to Mandela's constitution: Black South African Lesbians in the Nineties. In B. Greene (Ed.), *Ethnic and cultural diversity among lesbians and gay men* (pp. 88–116). Thousand Oaks, CA: Sage.

Reiss, B. F. (1980). Psychological tests in homosexuality. In J. Marmor (Ed.), *Homosexual behavior: A modern reappraisal* (pp. 293–311). New York: Basic Books.

Renzetti, C. (1992). *Violent betrayal: Partner abuse in lesbian relationships*. Newbury Park, CA: Sage.

Rothblum, E. (1990). Depression among lesbians: An invisible and unresearched phenomenon. *Journal of Gay and Lesbian Psychotherapy, 1,* 67–87.

Rothblum, E. (1993). Gay and lesbian faculty face issues in academia. *APA Monitor, September*, 4.

Rudolph, J. (1988). Counselors attitudes toward homosexuality: A selective review of the literature. *Journal of Counseling and Development, 67,* 165–168.

Savin-Williams, R. (1989). Gay and lesbian adolescents. *Marriage and Family Review, 14,* 197–216.

Savin-Williams, R. (1990). *Gay and lesbian youth: Expressions of identity*. Washington, DC: Hemisphere.

Savin-Williams, R. C. (1994). Verbal and physical abuse as stressors in the lives of lesbian, gay male and bisexual youths: Associations with school problems, running away, substance abuse, prostitution and suicide. *Journal of Consulting and Clinical Psychology, 62,* 261–269.

Savin-Williams, R. (1996). Self labeling and disclosure among gay, lesbian and bisexual youths. In J. Laird & R. J. Green (Eds.), *Lesbians and gays in couples and families* (pp. 153–182). San Francisco: Jossey-Bass.

Schreurs, K. M. G. (1993). Sexuality in lesbian couples. *Annual Review of Sex Research, 4,* 49–66.

Shidlo, A. (1994). Internalized homophobia: Conceptual and empirical issues in measurement. In B. Greene & G. Herek (Eds.), *Lesbian and gay psychology: Theory, research and clinical applications* (pp. 176–205). Thousand Oaks, CA: Sage.

Simoni, J. (1999). Confronting heterosexist bias in the teaching of psychology. In B. Greene & G. L. Croom (Eds.), *Education, research and practice in lesbian, gay, bisexual and transgendered psychology: A resource manual.* Thousand Oaks, CA: Sage.

Stone, G. L. (Ed.) (1991, April). Counseling lesbian women and gay men—Special issue. *The Counseling Psychologist, 19,* 155–248.

Strommen, E. (1989). Hidden branches and growing pains: Stressful aspects of negotiating their lives. Homosexuality and the family tree. *Marriage and Family Review, 14,* 9–34.

Tafoya, T. (1997). Native gay and lesbian issues: The two spirited. In B. Greene (Ed.), *Ethnic and cultural diversity among lesbians and gay men* (pp. 1–10). Thousand Oaks, CA: Sage.

Tierney, W. G. (1992). Enhancing diversity: Toward a better campus climate. Report of the Committee on Lesbian and Gay Concerns. University Park: Pennsylvania State University.

Trippet, S. E. (1994). Lesbians' mental health concerns. *Health Care for Women International, 15,* 317–323.

Weinberg, G. (1972). *Society and the healthy homosexual.* New York: St. Martins.

Weinberg, T. S. (1983). *Gay men, gay selves: The social construction of homosexual identities.* New York: Irvington.

Welch, B. (1990, January 26). Statement on homosexuality [Available from American Psychological Association, 750 First St., NE, Washington, DC. 20002–4242].

Woolley, G. (1991, Jan.–Feb.). Beware the well intentioned therapist. *The Family Therapy Networker, 15,* 30.

Youngstrom, N. (1991, July). Lesbians and gay men still find bias in therapy. *APA Monitor, 22,* 24–25.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.10

Diversity Matters: Religion and the Practice of Clinical Psychology

DONALD W. PREUSSLER, RICHARD E. BUTMAN,
and STANTON L. JONES
Wheaton College, IL, USA

10.10.1 INTRODUCTION 233

10.10.2 THE DIVERSITY OF THE WORLD'S RELIGIONS 236
- *10.10.2.1 Cognitive Dimension* 237
- *10.10.2.2 Ritualistic and Symbolic Dimension* 237
- *10.10.2.3 Moral Dimension* 237
- *10.10.2.4 Institutional Dimension* 238
- *10.10.2.5 Community and Lifestyle Dimension* 238
- *10.10.2.6 Experiential Dimension* 238

10.10.3 RELIGION AND/OR VS. SPIRITUALITY 239

10.10.4 PHILOSOPHICAL BASIS FOR APPRECIATING RELIGIOUS DIVERSITY IN CLINICAL CARE 239
- *10.10.4.1 The Philosophical Conflict Between Clinical Care and Religious Belief* 240
- *10.10.4.2 A Philosophy of Clinical Practice that Appreciates Religious Belief* 240
- *10.10.4.3 A Paradigm for Dialog on Religious Beliefs in Clinical Care* 241

10.10.5 CLINICAL ASSESSMENT AND DIAGNOSIS: APPRECIATING RELIGIOUS DIVERSITY IN THE PROCESS OF MEASUREMENT 243

10.10.6 THE EMPIRICAL BASIS FOR APPRECIATING RELIGIOUS DIVERSITY IN CLINICAL CARE 245
- *10.10.6.1 Religious Persons and Mental Health* 245
- *10.10.6.2 Psychotherapy and Values* 246
- *10.10.6.3 Religion in Psychotherapy* 246
- *10.10.6.4 Religion and Clinical Judgment, Technique and Behavior* 247
- *10.10.6.5 Religion and Clinical Research* 248

10.10.7 THE RELIGIOUS COMMUNITY AS A RESOURCE FOR SUPPORT 248

10.10.8 TRAINING CLINICAL PSYCHOLOGISTS IN RELIGIOUS DIVERSITY 249

10.10.9 FUTURE DIRECTIONS 250

10.10.10 SUMMARY 252

10.10.11 REFERENCES 252

10.10.1 INTRODUCTION

In this chapter, the aim is to promote the inclusion of spirituality and religion as potentially important and salient factors in the assessment, diagnosis, and treatment of clients by mental health professionals, particularly clinical psychologists. To do this, an attempt will be made to lay the philosophical, empirical, and clinical bases for such a treatise. The chapter will conclude with thoughts about future directions for the inclusion of spirituality and religion in clinical care.

"Probably nothing in human history has sparked more controversy and debate than religion" (Paloutzian, 1996, p. 2). The etymology of the word religion is interesting. It comes from the Latin word *legare,* which means that which binds or connects. As Paloutzian (1996) suggests, religion is a process of rebinding or reconnection. In the psychology of religion, it is less clear whether that binding or connecting is to "God, Nature, a state of mind, a cosmic force, each other as individuals or their communities" (Paloutzian, 1996, p. 7).

Historically, religion has not been a prominent diversity issue for the field of clinical psychology. Most often the field of psychology has embraced tolerance of diversity in the world of ideas and experience, with the notable exception of religious ideas and spirituality (Kauffmann, 1991). It is remarkable that an issue of *The Clinical Psychologist* devoted to diversity (Comas-Díaz & Striker, 1993) did not focus on, and barely mentioned, religion as a diversity variable. Bergin (1991) has asserted that "Psychologists' understanding and support of cultural diversity has been exemplary with respect to race, gender, and ethnicity but the profession's tolerance and empathy has not adequately reached the religious client."

There is historical precedent for the inclusion of religion in the practice of clinical psychology. The pioneering psychologist William James wrote in 1910 on the pragmatics of religious diversity in his essays on the philosophical construct of pluralism (James, 1910/1963). In his classic text on the psychology of religion, *The varieties of religious experience*, James wrote that "Religion is an essential organ of our life, performing a function which no other portion of our nature can so successfully fulfill." Beyond James, there is a long and established tradition of studies in the psychology of religion that lends support to the notion of "interest" in the psychological study of religion. Unfortunately, the psychology of religion literature conceptualizes religion primarily as a psychosocial variable or examines the functional utility of religion and rarely speaks to the issue of religion as a clinically relevant diversity variable that may transcend a unidimensional psychological or functional analysis.

Clearly, religion and spirituality are important to most people served by health care providers. Gallup polls in 1990 indicated that 95% of Americans believe in God and 85% believe the Bible to be the word of God. Polls also suggest that approximately 75% of Americans view their religious faith to be the most important influence in their life and large numbers (approximately 42%) would embrace the label of being "born again" (as cited in Larson, Lu, & Swyers, 1996, p. 1). There is also evidence that there is no difference between the religious beliefs and practices of mental health patients and the general population (Larson et al., 1996).

These religious believers often see their faith as relevant to their health and lifestyle concerns. Studies (e.g., Matthews, 1997) have found that 77% of patients in hospitals wanted their physician to consider their spiritual needs and 64% wanted their physician to pray for them if they requested it. Larson et al. (1996) review studies suggesting that most patients view spiritual health as relevant and important to their general health. Further, studies cited by Larson et al. suggest that religious coping is a major asset to elderly patients with psychiatric and mental disorders, and that there is a significant relationship between church attendance and lower rates of mental illness, as well as a beneficial relationship between religious commitment and lower suicide rates.

Research documents the positive contribution of religion to mental health and quality of life. A study of over 1000 university students found that religious students had better overall health and fewer injuries as well as less frequent use of tobacco, drugs, and alcohol, leading the researchers to conclude that religion had a positive effect on healthy lifestyle and behavioral choices (Matthews, 1997, p. 13). Worthington, Kurusu, McCullough, and Sandage (1996) concluded, after an exhaustive review of the empirical literature, that religious clients cannot reliably be labeled as having poor mental health. Many draw on their religion to cope with stress and the challenges of everyday living, especially when crisis strikes and options are limited (Hood, Spilka, Hunsberger, & Gorsuch, 1996). Hood et al. (1996, p. 378) summarize the work of Pargament and others saying, "People do not face stressful situations without resources. They rely on a system of beliefs, practices, and relationships which affects how they deal with difficult situations. In the coping process, this orienting system is translated into concrete situation-specific appraisals, activities, and goals. Religion is part of this general orienting system. A person with a strong religious faith who suffers a disabling injury must find a way to move from the generalities of belief to the specifics of dealing with that injury." In particular, prayer and social support may be viewed as positive coping strategies (Kauffmann, 1991).

Worthington et al.'s (1996, p. 457) analysis of empirical findings in the field from the last 10 years suggests that there are seven positive effects religion may have on mental health: (i) religion may produce a sense of meaning; (ii) religion may stimulate hope; (iii) religion may give religious people a sense of control by a beneficent God, which compensates for reduced personal control; (iv) religion prescribes a healthier lifestyle that yields positive health and mental health outcomes; (v) religion may set positive social norms that elicit approval, nurturance, and acceptance from others; (vi) religion may provide a social support network; (vii) religion may give a person a sense of the supernatural that is certainly a psychological boost but may also be a spiritual boost that cannot be measured phenomenologically.

Further, Hood et al. (1996) suggest that religion has the potential to meet the needs for meaning, control, and self-esteem. No doubt there are some positive and proactive ways this can take place, while other efforts might best be understood as avoidant or self-defeating (Malony, 1991). But to view prayer and social support as only means to control emotions seems potentially reductionistic and even patronizing. Perhaps naturalistic assumptions limit full appreciation of more adaptive "problem-focused" or "emotion-focused" possibilities? Even speculations of sociobiologists would suggest that religion has the potential to be a species-wide coping mechanism that has aided humans to cope successfully with life, and has enhanced their chances for physical survival (Wright, 1994).

Hood et al. (1996) have also concluded that the research evidence would "appear to be quite strong that religion, through offering a sense of meaning, control and self-esteem, does support an optimistic outlook. This in turn helps people deal constructively with life, and seems to have long-range beneficial effects." Especially pronounced are the potential for coping with death and the stresses often associated with the natural aging process. This is not to say that religion, in and of itself, "guarantees" stress inoculation; however, religiously based coping strategies appear to be quite effective for large numbers of persons in our society. Beliefs, rituals, prayer, and social support are important resources in the coping and adjustment of many people; religion may actually change a person since new interpretations are offered that might make problems less distressing and threatening.

As Hood et al. (1996) suggest, for those individuals whose religious orientation is intrinsic-committed rather than extrinsic-consensual:

> Religion probably helps because it provides individuals with personally useful meanings for upsetting circumstances. Concurrently, it offers, through a variety of avenues, opportunities for an enhanced sense of power and control over what is taking place. The result of both these tendencies and of faith itself is buttressing of self-esteem. Things no longer seem as bad as they once were, and since the individuals now believe they are doing the best that is possible, they can feel good about themselves. For the overwhelming majority of North Americans, the message of the 46th Psalm thus holds: "God is our refuge and strength, a very present help in time of trouble. (p. 401)

Still, there can be no doubt that in certain expressions of psychopathology, there is overt religious content and symbolism. Whether the connection is causal or consequential is far less clear (Paloutzian, 1996). Considering the fact that the concept of mental health itself is hard to pin down precisely, it is exceedingly difficult to draw sweeping conclusions in this vast and rapidly growing literature. Hood et al. (1996) note that the research has not been organized along productive theoretical lines. Further, they assert that antireligious biases were especially evident in earlier studies, and that serious methodological flaws limit the generalizability of those tentative conclusions. They remark:

> Personal religious expression may still reflect underlying mental disturbance, and for some, institutional faith remains a danger to their mental health. In most instances, however, faith buttresses people's sense of control and self-esteem, offers meanings that oppose anxiety, provides hope, sanctions social facilitating behavior, enhances personal well-being, and promotes social integration. All of these possibilities work to the benefit of distressed persons; ideally, they will be increasingly employed by mental health professions, to the advantage of those who seek their help. (p. 436)

Friedman and Benson (1997) suggest, "Many spiritual and religious individuals believe that the positive relationship that may exist between spirituality, religion, and health involves more than psychology and behavior" (pp. 1–12). Unfortunatley, as a study by Bergin and Jensen (1990) points out, of all the providers of psychotherapy, clinical psychologists are the least religious of the group (as measured by formal institutional involvement and investment) and they may also therefore be the least informed about religious behavior, particularly as it is represented in its institutional forms. They also appear to be significantly less likely than the general population to view religion as having an important role in their lives. It could therefore be argued that there is not only a lack of religious diversity in the profession but there may also be a lack of personal understanding and investment in the important religious concerns of our clients.

The American Psychiatric Association (cited in Larson et al., 1996, pp. 5–6) issued "Guidelines Regarding Possible Conflict Between Psychiatrists' Religious Commitments and Psychiatric Practice" in which they stated explicitly that psychiatrists should not only "maintain respect for their patient's beliefs" but also "obtain information on the religious or ideological orientation and beliefs of their patients so that they may properly attend to them in the course of treatment." However, as Larson et al. point out in their psychiatric residency curriculum, "While well intended and vital as an important first step, these guidelines may prove difficult for mental health professionals to follow due to their lack of familiarity with religious issues in the clinical environment."

These important guidelines are germane to the practice of clinical psychology as well. The American Psychological Association's revision (1992) of its "Ethical Principles of Psychologists and Code of Conduct" includes religion in the list of other matters of diversity as an area that may require special competence (Standard 1.08). It is suggested that clinicians need to not only respect but also increase competency by being sensitive to this multicultural diversity issue through awareness, knowledge, and skills related to addressing religion as a diversity variable in treatment (cf. Brems, 1993, p. 74). A lack of understanding and appreciation for the role that religion plays in the lives of clients may reduce the clinician's effectiveness in assessing and aiding clients to change, and may significantly influence the ability to build rapport and trust. At a minimum, clinicians need to avail themselves of opportunities to learn more about religious traditions different from their own.

It is also important to acknowledge that religion has provided some unique challenges as clinicians have embraced diversity in all its important expressions. A commitment to diversity is associated in the minds of most with the presumed notion that diversity should be celebrated in a pluralistic spirit. By implication, if I recognize and respect your religion as on par with mine, you should reciprocate, a notion not necessarily endorsed by all religions (Kung, 1985). Some of the world's religions do embrace diversity which in turn may even lead to syncretism or universalism (respectively, the combination or reconciliation of two belief systems). However, some of the world's major religious systems are "intolerant" and "authoritarian," and are compelled to embrace their own notions *a priori* and reject, or at least devalue, those of other religions (exclusivism). Indeed, the study of world religions would suggest that it is easier to internalize exclusive claims to truth, a reality which can complicate any search for common ground. Therefore, religion, by its nature, at times precipitates a philosophical conflict between those who embrace exclusive religious systems and those who anticipate reciprocity in the celebration of diversity.

10.10.2 THE DIVERSITY OF THE WORLD'S RELIGIONS

A chapter on the proper appreciation of religion as a powerful diversity variable ought to contain a terse summary of the major religions and their major distinctives, but such a summary would fill an entire chapter itself and do an injustice to the exquisitely complex realities of the religious faiths. Instead an attempt will be made to mention some of the most important dimensions on which religions vary.

Several caveats are in order. First, it should be noted that there is the need for humility in attempting to understand different religions; few people are experts in the world religions and all their variants (for orienting surveys, see Nielsen; 1993; Noss & Noss, 1993; Smart, 1989), and even less can people truly be appreciative supporters of all religions equally. There should be a readiness to acknowledge the limits of knowledge, and of the limited attitudinal flexibility which can be mustered in confronting beliefs that are different. Of particular danger to psychologists is the temptation to confuse their personal synthesis of religions, often via some sort of psychological functional analysis, with a genuine appreciation of all religions. Such a synthesis is necessarily a variant on religious belief itself and hence in tension with other religious beliefs; for example, an analysis of all religions as "particularistic cognitive renderings of the universal human pursuit of transcendent purpose, the ethical good, and of community" is a competing definition of, rather than an apt summary of, any particular religion.

Second, often there is greater diversity within religious categorizations than across them. For example, conservative Catholics and conservative Protestants have, on many dimensions, more in common than liberal and conservative Protestants. Hence, some very diverse religious groups are able to build remarkable consensus on certain foundational issues.

There have been many attempts to do what amounts to a conceptual factor analysis of religion, with varying outcomes (there have been empirical attempts as well; see Gorsuch, 1984). This chapter draws on the work of Glock (1962), Smart (1989), and others and discusses the multidimensionality of the religions in terms of their cognitive dimension (religious beliefs), ritualistic and symbolic dimension (religious practice), moral dimension (religious action), institutional dimension (religious organization), community and lifestyle dimension (religious community), and experiential dimension (religious feelings).

10.10.2.1 Cognitive Dimension

Religions vary cognitively in a number of ways. Myths play a central role in most religions, where myth is understood not in the general use of the term as a fantastic fictional tale, but rather as a set of religious stories that "quiver with special or sacred meaning" (Smart, 1983, p. 7). The importance of these are made clear in the care and honor given to the sacred texts that record them: the Christian Bible, the Hebrew Torah, the Hindu Bhagavad-Gita, Islam's Koran. The abiding power (i.e., the value of the narrative whether oral or written) of such sacred stories vibrates in the communities which have been transformed and sustained by such stories as the Jewish Exodus from Egypt or the visions of Lao-tzu of the early Taoist movement. For some believers, the historicity of the founding myths is vital, while others regard them as symbols pointing to meanings not tied to specific events in history. Christians, for example, have traditionally insisted on the historical reality of the death and resurrection of Jesus Christ (literalism); some continue that tradition while others regard that story as a nonhistorical emblem of the ability to overcome evil and adversity through the transforming power of God's love (symbolism). Recognition of the literalism vs. symbolism hermeneutic typically reaches beyond just one dogma or belief of the group but often pervades into other areas of interpretation within the group's religious belief system.

Religions vary according to the content of their founding myths, and also according to the place of doctrine in the religious community, its sophistication, and degree of elaboration. Smart (1983, p. 8) defines religious doctrine as "an attempt to give system, clarity and intellectual power to what is revealed through the mythological and symbolic language of religious faith and ritual." In essence, doctrine is an attempt to systematize divine revelation and render it applicable to everyday life. Some religious traditions have given rise to extensive systematized literatures (e.g., Christianity, Judaism, Islam, and Hinduism), while others have not (e.g., Animism and Shintoism).

Myth and doctrine together can contribute to the formation of the world views of religious adherents. Some religions have limited general application; some animistic faiths prescribe rituals by which to appease or petition their gods, but have few broader implications. But a religion can give broad definition to the world which its faithful inhabit. In essence, a religious faith can constitute the lenses (i.e., cognitive-perceptual "style," world view, or control beliefs) through which believers see the world, and those lenses are clearly different from those through which the unfaithful peer. The central focus around which a religious world organizes is the sacred or divine. This understanding of the divine is so striking that it literally shapes how an individual, indeed, how a community, understands the greater order of existence. For a Christian this picture orients around a creative, redemptive, and present God. To Shinto believers, the kami, a spirit or divinity, completes their understanding of holiness in this world. The Muslim believer finds certainty in existence in Allah's Five Pillars of Faith.

10.10.2.2 Ritualistic and Symbolic Dimension

Most obvious to an outside observer are the differing roles played by ritual in the world religions. Common forms of religious ritual are worship, singing, fasting, and prayer. In general these are "some form of outer behavior coordinated to an inner intention to make contact with, or take part in, the invisible world" (Smart, 1968, p. 6). Rituals can be daily practices such as the yoga of Hindus, the prayers of the Shinto or the purity rituals of Orthodox Judaism, weekly participation in services such as Catholic mass or Jewish temple, or annual celebrations such as Islam's Ramadan or the Hindu Divali. Each is a unique attempt to connect with, through discipline and remembrance, the divine which provides an orientation to self, others, and moral good.

10.10.2.3 Moral Dimension

Religions vary in the degree of elaboration of their accompanying ethical systems, but such systems are connected vitally to religious faith and the consequences of belief. At its most basic level, ethics is the way in which religious systems answer the challenge of evil in the world and deal with the profane (Paden, 1988, p. 144). Inherent in religious ethics is a call to live in a manner which reflects one faith in an unbelieving world. Thus for a Muslim, love for Allah will be reflected in distributing wealth among those in need, while for a Sikh it involves, but is not exhausted by, wearing uncut hair, a dagger, breeches, a comb, and iron bangle (Smart, 1989, p. 98). Religious ethics systems vary according to their overt applicability to society; Islam and ancient Judaism have ethical systems seemingly designed for implementation on a societal basis, while New Testament Christianity articulated an ethical code for members of a disenfranchised and powerless subculture. Religious ethical systems typically have individual, interpersonal, and communal implications.

10.10.2.4 Institutional Dimension

The religions differ in terms of their formalization as enduring human institutions. At one end of the continuum, imagine an autonomous American who distills a set of idiosyncratic religious beliefs which can be embraced, and who then quietly, privately, and with dignity lives consistently with those beliefs without the formation of an organization at all. At the other extreme contemplate the Roman Catholic Church with its high degree of institutionalization. While some degree of institutionalization is probably inevitable with growing size of the adherent body, religions differ in terms of how readily they engender institutionalization.

10.10.2.5 Community and Lifestyle Dimension

World understandings formed and framed by religious belief serve social functions. They draw boundaries which allow believers to understand themselves (the insider) in contrast to others (the outsiders); thus, a Jew is able to clearly define themselves as distinct from a Muslim or a Buddhist, both in terms of differences and commonalties. A communal consensus on "who we are," common understandings of proper and improper behavior and values, the power of shared rituals, language and symbols, common engagement with religious institutions, and even an emphasis on the importance of community itself all contribute to a sense of belonging to a religious community and a cohesive sense of lifestyle. Some religious believers have a diffuse sense of engagement with their religious community, while others are deeply engaged with a highly visible and formalized community. Williams notes that it has been common in the sociology of religion to distinguish between church and sect according to the degree of rejection of the dominant social environment, with members of sects disengaged from majority culture. "Compared with members of churches, members of sects are poorer, less educated, contribute more money to their religious organizations, attend more services, hold stronger and more distinctive religious beliefs, belong to smaller congregations, and have more of their friends as members of their denomination" (Williams, 1993, p. 127). This is a helpful distinction as long as the term sect is understood in a nonpejorative sense. Such a distinction may also be an important indicator of potential individual differences in ego-strength or assertiveness between groups as well as relevant to understanding counterculture tendencies and conformity pressures for individuals within the sect.

10.10.2.6 Experiential Dimension

Religious experience is often regarded as the *sine qua non* of religious life and its goal. Historically, such experience has often occurred at the founding moments of a religious tradition: the Koran tells of Mohammed's overwhelming, painful experience of receiving revelation from Allah; Buddhists" honor the light that filled Buddha's mind under the Bodhi tree, allowing him to see the antidote for the suffering of this world; and Christians recall the blinding, life altering vision the apostle Paul received on the road to Damascus. Believers across the spectrum regularly celebrate and search after the same. Thus, many Christians refer to their entrance in the faith as being "born again," Taoists search for inner illumination that will lead to a "mystic union," and Hindu's practice yoga to catch a glimpse of Nirvana. Such desires recognize the unique and holy place of the divine in the world of a believer and may in fact be one dimension that drives clients in psychotherapy.

Psychology as a discipline has often attempted to understand religious experience, but in doing so has often imperialistically presumed that only certain types of experiences are properly religious. Lash (1988), for example, argued against the account of religious experience of William James (articulated in his classic *The varieties of religious experience*) on the basis that it is an exclusivist account of religious experience, one which looked for a particular and peculiar type of experience as qualifying as religious experience and which ruled out as "true" religious experience on *a priori* grounds any experience which was tightly connected to a particular religious tradition. Such an *a priori* definition misses the reality that there is no such thing as generic and pure religious experience, and that the forms of religious experience are intimately connected with and vary according to the faith systems in which they occur (Lash, 1988). A variety of religious experiences within Christianity—of shame and guilt for sin, of repentance, of gratitude for God's mercy—have no direct parallel in other faiths such as Buddhism, a reality which may produce some unique challenges for engagement of the two religious systems.

In attempting to understand the religion of a client and its impact upon their presenting concerns, understanding these dimensions of religion—cognitive, ritual and symbol, moral, institutional, community and lifestyle, and experiential—can serve as a guide for exploration of a client's particular religious faith. Awareness of these factors can assist psychologists to catch a general glimpse of how faith affects and is interwoven into the lives of clients, and the unique differences between the many faiths which will be encountered.

10.10.3 RELIGION AND/OR VS. SPIRITUALITY

Even a cursory reading of standard texts in the psychology of religion (Hood et al., 1996; Kauffmann, 1991; Malony, 1991; Meadow & Kahoe, 1984; Paloutzian, 1996; Wulff, 1991) would suggest that there is no consensus in the field about the definition of either religion or "spirituality." The use of the term "spirituality" is almost a "hot" topic in clinical psychology although the term is being historically redefined in the psychology of religion literature. As Worthington et al. (1996) have noted, the fascination has more to do with an interest in mysticism and very private spiritualities rather than with the more institutionalized, corporate, and communal expressions of institutional worship, fellowship, and service. This focus on spirituality, in deference to the more classical understandings of the psychology of religion, trivializes more well-defined constructs like intrinsic-committed or extrinsic-consensual religiosity or the quest orientation as well as the major religious orientations derived from decades of research. Many clinicians who focus strictly on the notion of spirituality lack significant familiarity with institutionalized religion. This has a tendency to make it difficult for the clinicians to fully appreciate the complex and subtle ways that persons of corporate and institutionalized faith interact within their communities.

Zinnbauer, Pargament, Cowell, and Scott (1996), in a paper entitled "Religion and Spirituality: Unfuzzying the Fuzzy," recognized the inherent definitional difficulty in attempting to study "spirituality" in contrast to "religion." They conducted a study to measure how individuals define religiousness and spirituality. They were interested in how the individual's definition might be associated with different demographic, religio/spiritual, and psychosocial variables. They concluded, first, that religiousness and spirituality are probably different concepts. They found that religiousness was associated with higher levels of such things as "authoritarianism, religious orthodoxy, intrinsic religiousness, parental religious attendance, self-righteousness, and frequency of prayer," while spirituality was more closely associated with "mystical experiences, new age beliefs and practices, higher SES, and frequency of prayer" (p. 9). Whereas spirituality seemed to be more closely aligned with both personal and experiential dimensions, religion seemed to connect specifically with organizational or institutional beliefs as well.

On the other hand, they also concluded that individuals tend to integrate both spirituality and religion into their lives. In particular, they found that for the majority of individuals, spirituality and religion were irrevocably tied together. Their third conclusion was that "most believers approach the sacred through the personal, subjective, and experiential path of spirituality, but they differ in whether they also include organizational or institutional beliefs and practices" (p. 10). The researchers described their study as having particular importance for mental health workers, who as a group tend not to integrate their spirituality into religiousness. Thus, mental health workers may have a potential bias toward spirituality but against religiousness when for the general population they are typically linked. In light of the fact that their study found most people to identify themselves as both spiritual and religious, mental health workers prone to defining themselves as spiritual but not religious may fail to be sensitive to this integration. They may fail to appreciate the religious side of the majority of their clients.

Finally, it should be noted that formal and informal attempts have been made throughout history to engage in acts of kindness and healing under the auspices of spiritual direction. In contrast, disciplines like clinical psychology have a relatively recent history of such work. Unfortunately, only a minority of the mental health providers are even remotely aware of the rich and centuries-old traditions of pastoral care and spiritual formation (Benner, 1988; Browning, 1987; Coles, 1990; Groeschel, 1992; Jones & Butman, 1991; Malony, 1991; Miller & Jackson, 1995; Shafranske, 1996; Worthington et al., 1996). In many ways, psychotherapy is an heir to a rich tradition of altruistic service and compassion performed by spiritual directors.

10.10.4 PHILOSOPHICAL BASIS FOR APPRECIATING RELIGIOUS DIVERSITY IN CLINICAL CARE

It could be argued that religion and mental health care have much in common metaphysically. This is particularly true in the broad philosophical areas of ultimate meaning and morality. In fact, there are those in the field of clinical psychology and mental health care who would argue that the constructs and application of psychological theory are "religious" in nature and form a pragmatic basis for the replacement of institutional and/or personal religion in the lives of "believers" (Gross, 1978; Szasz, 1977; Zilbergeld, 1983). Perhaps no one has made the case more clearly than London (1986) in regard to the moral and religious nature of psychology, particularly applied psychology, and the enterprise of psychotherapy. In many ways it would appear that religion and psychotherapy are both activities that involve the search for a coherent world view and attempt to link beliefs with behaviors.

10.10.4.1 The Philosophical Conflict Between Clinical Care and Religious Belief

There is a potential for philosophical conflict between clinical psychology, as a presuppositionally religious enterprise attempting to answer broad questions of ultimate reality and morality, and the control beliefs of other religious systems. Thinking about clinical work as controlled by a belief system that influences how humans view themselves and their world makes it particularly important to consider how the clinicians" view of religion impacts the practice of clinical psychology. The clinician's control belief system clearly impacts the lives of religious clients.

For example, for the client who presents with marital difficulties and comes with a religious background that contains strict teachings about gender roles within the family or community, the control beliefs of the client based on the teachings of their religion may come into conflict with the control beliefs of the clinician regarding gender roles. For the clinician, this difference may be based on the theoretical orientation which forms their frame of reference or world view. Hypothetically, the hierarchical view of gender roles embraced by the clinician may come into direct conflict with the client's own chosen response to the egalitarian control beliefs of their religion. In fact, for the clinician, the client's religious beliefs about gender roles may in fact suggest presence of psychopathology or at least unhealthy thinking. The client is left with the dilemma of either embracing the control beliefs of the clinician (and getting "well" or becoming "healthy" in the process) or maintaining their own beliefs and being "sick" or "unhealthy." Further, it is important to recognize the religious context from which the client comes and the role of their control beliefs in maintaining important psychosocial resources in their life. These resources are often a direct result of the corporate and/or communal dimensions of their religion. To fail to do so may force the client into a painful choice between accepting the control beliefs of the clinician and losing important psychosocial resources or rejecting the control beliefs of the clinician and being abandoned.

Another example might involve questions of ultimate reality. The religious client who comes to the clinician family a number of stresses as well as feelings of depression in the aftermath of the death of a spouse may find religious beliefs about an afterlife minimized or even pathologized. Confrontation might be faced over being in "denial," or a challenge made to express suppressed anger toward religious deity, or to abandon a fantasy about reincarnation or reunion with the spouse. Such interventions by the clinician fail to appreciate the positive role that religion may be playing in the life of the client. In the worst cases, the clinician may also assume that in some way the religious control beliefs of the client are inferior to the psychologized control beliefs of the clinician. In these cases, the clinician presupposes that the client's religious beliefs are totally ineffective in helping the client manage his life.

10.10.4.2 A Philosophy of Clinical Practice that Appreciates Religious Belief

A starting point for the dialog about clinical practice and religious belief may be to think about clinical psychology as a postpositivistic science that is not value free. Further, in the practice of clinical care, it should also be noted that the scientific enterprise involves metaphysical presuppositions that reflect the world view and belief system of the scientist. Obviously, the potential exists for the clinician's control beliefs about ultimate reality and morality to be different from the presuppositional control beliefs of a client's religious system. It is further suggested (cf. Jones, 1994) that the relationship between religion and clinical work is in part an issue of addressing the religious diversity that may exist between the clinician and the client. An explicit acknowledgment of control beliefs, is recommended, on the part of both the client and clinician, as an important point of dialogical contact.

Following the "Ethical Principles of Psychologists and Code of Conduct" (American Psychological Association, 1992) with respect to "Principle A: Competence" and "Principle D: Respect for People's Rights and Dignity" would, on a very practical level, seem to involve the clinician acknowledging awareness (or lack thereof) of the person's particular religion, articulating a level of knowledge in regard to the control beliefs of the religion, and finally describing any skills or previous experience in working with individuals with similar religious beliefs. It would also seem logical that such a discussion would involve an articulation of the contrasting or similar control beliefs of the clinician as well as the clinician's theoretical orientation. Obviously, the role of religion in clinical work may be very different depending on the clinician's theoretical orientation. Further, the extent of the discussion regarding religious diversity may ebb and flow with the importance of the issue in the presenting problem of the client as well.

10.10.4.3 A Paradigm for Dialog on Religious Beliefs in Clinical Care

Dialogical contact between clinical psychology and religion may be enhanced by an articulation of the nature of the dialog, particularly if it is understood and applied in terms of multicultural notions of religious diversity. In an article on religion and multiculturalism, Dell'Olio (1996) attempted to address three philosophical constructs related to religious diversity: exclusivism, pluralism, and inclusivism. If the control beliefs of the clinician, whether personal or based on her theoretical orientation, are considered as overlapping metaphysically (e.g., issues of ultimate meaning, morality, etc.) with the religious control beliefs of the client, Dell'Olio's (1996) proposal for the management of religious diversity may contain merit for understanding the relationship between the clinician's perspective on these overlapping metaphysical issues and the client's religious control beliefs.

Dell-Olio (1996) rejects the construct of religious exclusivism on moral grounds. He believes it is immoral to reject totally the beliefs of another simply because they are different or ancillary to one's own. In addition, he rejects the religious pluralism of John Hick, arguing that in his underlying assumptions Hick's "perspective presumes to know more about what the religions themselves know about what they know, and thus refuses to recognize the legitimacy of the other's perspective regarding religious truth." Dell'Olio (1996) argues for "inclusivism" which allows for limited reciprocity on the basis of what he refers to as the multicultural mandate of the "morality of recognition." Limited reciprocity refers to a process whereby individuals seek common ground in regard to their individual control beliefs and gain an understanding of the distinctives of each others control beliefs and, even when their control beliefs are irreconcilable respecting the control beliefs of the other without abandoning one's own control beliefs. To Dell'Olio, (1996), limited reciprocity or "inclusivism" is the moral "high ground" or what he refers to as the "morality of recognition." In other words, the individual is free to acknowledge the sincerity of another's perspective even when they would sincerely disagree with some, much, or all of it.

Dell'Olio's (1996) conceptualization (morality of recognition) applied to the relationship between religion and clinical psychology would suggest that clinicians need to recognize with sincerity the client's perspective regarding religious truth even if the clinician disagrees with some, much, or all of it. This is particularly important when clinical work addresses the broader questions of ultimate reality and morality. However, the religious perspective of the client may not only be very important to the client's understanding of ultimate reality and morality but also the client's sense of self.

Dell'Olio's (1996) notion of inclusivism holds promise as a paradigm for the dialogical relationship between religious beliefs and the practice of clinical psychology. Inclusivism allows the person to remain committed to their religious beliefs while recognizing that religion does not have exclusive claim to all "truth." In this way, the religious client can remain open to the truth-claims of the clinician. On the other hand, the clinician can remain committed to the religious dimensions of their truth-claims while recognizing that they do not have exclusive claim to all "truth." Inclusivism allows the religious person to appreciate the truth-claims of the clinician and also be free to note where and when the truth-claims of the clinician are inadequate in describing or appreciating their religious experience. Inclusivism allows the clinician to appreciate the truth-claims of the client's religion as well as to note where and when the client's religion is inadequate in describing or appreciating the nature of the client's problems or their clinical work.

The morality of recognition principle in the care of clients means that the clinician is not obligated to embrace all of the religious control beliefs of the client in order to work with the client, nor is the client obligated to embrace all of the control beliefs of the clinician in order to benefit from therapy. Rather, it is the open acknowledgment on the part of both the client and clinician of control belief similarities and differences that may impact their work together, particularly as it relates to the presenting problem and the experience of the clinician.

For example, a religiously liberal Christian client being treated with cognitive therapy for depression may have little difficulty with the morality of recognition principle when being cared for by a religiously liberal Jewish or Muslim clinician or visa versa. In contrast, a conservative Christian being treated for a generalized anxiety disorder who is prescribed an Eastern meditation technique will likely have difficulty embracing the acceptableness of that approach in light of their religious teachings, regardless of the clinician's beliefs or intent. The clinician's awareness of such a conflict based on knowledge of the religious beliefs of the client may cause the clinician to consider alternative treatments in light of the morality of recognition principle. However, in some cases, the client or clinician may feel that the clinician's awareness, knowledge, and skill related to the client's religious diversity issues are inadequate for their work to be as productive as it might be if the client was referred to another clinician more aware, knowledgeable, or skillful in the religious diversity issues of the client. In all three situations, there is an expression of the morality of recognition principle.

Paradigmatically, inclusivism avoids the imperialism of exclusivism but in contrast to the pluralism, inclusivism "realizes that judgments must be made from a particular perspective" and that neither the clinician nor the client needs to "give up their particular (religious) perspective" when engaged in the process of clinical care. Clinicians are not required to give up their control beliefs in appreciating their own and the client's experience of religion and spirituality.

It is necessary for those who practice clinical psychology to seek an understanding of religion and spirituality in the context of an inclusive diversity while appreciating the contribution of both psychology and religion to the welfare of the human family. While applied postpositivistic psychology (as relativistic "science") and religion (as absolutist "divine or natural" law) may share very similar subject matter, they come to the human family with very different epistemological and methodological notions. The morality of recognition principle goes beyond just the awareness and knowledge of different control beliefs and the skills to address those differences. The morality of recognition is also an awareness and knowledge of different epistemologies and methods for arriving at those control beliefs. For the clinician this involves the development of skills to navigate these different epistemological and methodological approaches.

For example, for the client who endorses the "absolute authority" of a sacred text, anything the "science" of psychology may conclude from empirical investigation or theoretical development that would oppose or contradict the sacred text would be discounted because of the primary authority given to the text. Conversely, teachings of a sacred text are often discounted by "scientific" findings. The lack of awareness and knowledge of both epistemologies will be a deficit to the clinician who requires skill to navigate these issues in therapy. The clinician needs to recognize how the client understands the authority of religious truth and the role the client ascribes to "scientific" truth and work with the tensions that may arise in both the client and clinician as a result.

Religion and clinical psychology share similarities in subject matter and are human enterprises (Jones, 1994), and in their pragmatic forms both attempt to understand and interpret the behavior and experience of human beings. They also share constructs such as cognition, consciousness, emotion, motivation, and relation to name just a few (Tisdale, 1980). As clinical psychology and religion attempt to share the same constructs, their contrasting differences are pronounced by the fact that they may come from different epistemological and methodological bases. The result may often lead to different conclusions about the same matters. This affects the practice of clinical psychology because it affects not only how the clinician and client view the nature of the client's problem but also the resources that should be accessed in treatment of the problem and the validity and reliability of the treatment outcome.

For example, many clinicians have wondered about the immediate positive effects of antidepressant medicine in clients who have been suffering from depression. This is particularly true when psychiatric colleagues explain the pharmacology of the drug. Many clinicians may have also wondered about the client who informed them that their depressed mood had lifted as a result of prayer or some other prescribed spiritual discipline. In the first case, there is no pharmacological basis to adequately explain the effect. However, in the second there is often a clearly defined religious control belief or teaching for the religious intervention effect. Both situations could be potentially explained psychologically as the classic "placebo" effect, but it may be a lack of epistemic humility that makes the latter is so hard to accept. In fact, even though there may be no pharmacological explanation for the immediate effect of the antidepressant in some clients, most clinicians still refer severely depressed clients for medical treatment. Is it any less appropriate to refer a religious client for a religious intervention that is consistent with the client's religious belief system?

Clinical psychology and religion share similar concerns about outcomes. It could be argued that both disciplines are attempting to interpret and manipulate human experience in meaningful ways that will produce "ultimate good." Again, the disagreement often comes in the epistemology and methodology used (e.g., divine revelation vs. materialistic determinism; Stevenson, 1987). In fact, what ultimately may matter to both is what or who gets the credit for any "good" outcome.

Clinicians have also witnessed the lack of efficacy of medical treatments for a number of mental disorders. Should the practice of referring clients for medical treatment for that reason be abandoned totally? Why should it be any different for religious interventions? In fact, many religious systems have developed systems for explaining the limited efficacy, or the conditions of efficacy, for their religious interventions. It is a lack of epistemic humility that keeps clinicians from taking full advantage of the religious interventions found in the religious beliefs of many clients.

Even in light of these commonalties and significant metaphysical differences, we have an obligation to do "good" psychology, while recognizing the important role that religious presuppositions play in all of our lives. The multicultural mandate of the morality of recognition principle is important as is the fact that religion is an important multicultural variable in the lives of both clinicians and their clients.

10.10.5 CLINICAL ASSESSMENT AND DIAGNOSIS: APPRECIATING RELIGIOUS DIVERSITY IN THE PROCESS OF MEASUREMENT

There is a rich tradition in psychology and sociology of measuring and assessing religion and religiosity in its many complexities. For an introduction to this vast literature, the interested reader is urged to turn to Gorsuch (1984) and Williams (1993), or to one of the excellent survey texts in the psychology of religion, such as Hood et al. (1996), Paloutzian (1996), or Wulff (1995). Additionally, MacDonald, LeClair, Holland, Alter, and Friedman (1995) provide an analysis of instruments which attempt to measure spiritual experience and spirituality disconnected from traditional forms of institutionalized Western religion; for example, "spiritual orientation," mysticism and mystical experience, peak experiences, self-transcendence, paranormal beliefs and experiences, altered states of consciousness, holistic living, and so forth.

Religious beliefs and behaviors are complex, multidimensional variables. Attempts to measure religious variables have two principle historical roots: Allport's intrinsic and extrinsic orientations and Batson's quest approach (Batson, Schoenrade, & Ventis, 1993). In the broadest sense, an intrinsic person appears to live the faith whereas an extrinsic person uses the faith (see any of the above noted overviews for more in-depth discussions). The assumption is that an extrinsic orientation is a less mature or developed orientation. The quest orientation is one in which an individual adopts a critical but open-ended approach to existential questions. Other formulations tend to combine elements of the quest and intrinsic orientations. These are hardly "pure" categorizations.

Researchers and theoreticians alike assert that the overwhelming majority of people in the Anglo-American context show elements of all three tendencies in the ebb and flow of everyday life. It has been noted that all three orientations stress "process" more than "content," as do measures of moral, cognitive, and psychosocial development. Perhaps the most helpful formulation is to use an attributional perspective that is respectful of the phenomena studied, appreciates the potential contributions of the intrinsic, extrinsic, and quest orientations, and strives to find connections between motivational and cognitive patterns that are at the heart of religious experience and behavior.

For decades, attempts have been made to measure aspects of the dimensions of religious commitment. Hundreds of decent measures are available. Gorsuch (1984) has argued that there is a need to refine existing measures rather than create new ones. Hill, Butman, and Hood (1997) have published a collection of such measures; this should help address the fact that few clinicians seem to be aware of the existence of such measures or where to find them. Familiarity with a range of assessment instruments can enrich an understanding of the range of religious experience. Further, not all of the assessment effort has been targeted at individual differences; Pargament and his colleagues have done innovative work in assessing the differing organizational climates of religious congregations (Pargament, Silverman, Johnson, Echemendia, & Snyder, 1983; Pargament, Tyler, & Steele, 1979), finding that they differ on such dimensions as order and clarity, sense of community, openness to change, social concern, autonomy, stability, activity, expressiveness, problem solving, and participation.

The utility of existing measures is, however, limited. There is little convincing data on their "relevance" beyond the majority culture in the Anglo-American context. They tend to stress behaviors (e.g., church attendance) more than religious beliefs, or often contain a mixture of questions about belief content, attitudes, values, morality, actions, and experiences which makes interpretation of the resulting scores difficult. Gorsuch (1984) criticizes their heavy reliance on self-report and tendency to self-distorting bias or impression management. They are usually confined to an objective format and, thus, do not allow for a more complete understanding of the motivational and cognitive patterns that undergird so many of our decisions to act, qualities that perhaps can only be fully explored in a more open-ended interview format. Many instruments tend only to measure broad dimensions of religious orientation; it would be helpful if these measures could evaluate the more specific and concrete aspects of religiosity for descriptive/prescriptive purposes. Finally, psychological measurement of religion appears to often be contaminated by the theoretical commitments of the researcher; Gartner (1996, pp. 187–203) has noted that "hard measures" of health and well-being (such as death rate) almost without exception show the positive health benefits of religion, while "soft measures" of personality traits and other theoretically-grounded psychological variables often show negative outcomes to be associated with religion, suggesting that the soft measures are biased or contaminated in some way.

Drawing from this vast assessment literature, it is suggested that intake questionnaires in clinical practice be used to increase the clinician's awareness of religious variables and issues in the life of the client. Profitably they could include a few initial items regarding religion to set the stage for further examination of these issues if merited. Brief items querying how frequently the person attends religious services and functions, and how frequently the person engages in personal religious practices and activities such as prayer, meditation, or bible reading, can both be answered on a five- or seven-point scale from daily to never. The vast assessment literature indicates that the action or behavioral dimension of religion is important to assess, as there is a direct tie between what people do and their degree of commitment to religion; participation in personal or institutionalized religious activities may be regarded as a proxy for religious commitment generally. Clients can be asked to rate how important or significant religious faith is to them, from extremely important and at the center of life to not important at all. Enquiries can be made about their current religious affiliation, and what other important religious affiliations have had a major impact upon them in the past. The degree to which they believe their religious faith is relevant to their presenting concerns can be rated.

As the assessment of religion and religious faith in considered for the sake of better understanding clients and the difficulties they present, however, the limitations of questionnaire assessment methods are confronted. In the area of sexuality, a variety of measures exist which are well suited for research but which have little clinical utility and which can never duplicate the richness of data which can be generated by competent clinical assessment; so also in this area.

As in all areas of clinical interview assessment, the clinician will only get useful information if religion is queried sensitively and with respect for whatever answer may result. Clinicians should be wary of the functional equivalent of the "You aren't still masturbating, are you?" question in the area of religion. Introductory questions can include: "People in our community hold to a wide variety of religious faiths and beliefs, and sometimes a person's religion is quite important to the issues and concerns being addressed in the counseling relationship. Where are you on the matter of religion? Do you see your religious faith as being related to this problem in any way?"

In responding to client output, clinicians must be prepared for religion to serve a complex array of roles, none of which are mutually exclusive: motivation for change ("God wants me to overcome this problem"), cause ("My fundamentalist upbringing is what made me sexually unresponsive"), effect ("This problem has brought my faith to life as I have realized how much I need God" or "I gave up believing as I have suffered through this problem for years and found no relief in my church"), potential obstacle to change ("But the people in my church would ostracize me if I were to be more assertive with them"), coping strategy ("When I get stressed out, it helps to meditate"), potential resource for change ("A number of men in the synagogue are very willing to serve to keep me on target in this change plan"), apparent locus of pathology (religious delusions and religious content to depressogenic cognitions), and others.

It is best to look at the client's religious faith developmentally and as a dynamic reality rather than as a static one. The various religious or faith development theories are not regarded as normative in this area, though they can be helpful. For example, Fowler's (1981) theory suggests that faith development parallels cognitive and moral development as understood from the work of Piaget and Kohlberg as the individual moves from highly concrete and literalistic faith to a religion of universal abstraction. Such an analysis, while engaging, may overestimate the role of rationality in religious growth and may impose an inappropriate normative model that psychologizes faith and obscures differences across religions. A more open-ended querying of religious change over the life span is suggested. If a client indicates or comes to understand that their religious faith is relevant to the clinical concerns, it may be fruitful to explore the major epochs in the religious life, asking for information about those periods when religious faith was most central to the person, the major factors producing change in religious belief and experience, and important formative persons and events in the life of faith. The clinician's ability to show respect and empathy for the client's journey will pay dividends in rich information which can inform intervention. The clinician's ability to empower the client to believe that it is acceptable to talk about religion/religious issues and even to probe the clinician's religious perspectives is a very helpful way to build rapport and facilitate the assessment process.

Religion serves a variety of functions in the lives of participants, and psychological interpretation of those functions can enrich clinical understanding. Interpretations, however, may be distorted by prejudices and theoretical systems, and such a functional analysis can be alien to the fundamental views and instincts of the client. The shape of a functional analysis is driven by the basic theoretical assumptions which the clinician brings to the case, and so functional analysis from a classic psychoanalytic perspective will come up with a fundamentally different portrait of the dynamics of religious faith than that from a behavioral or a cognitive perspective. Psychological analyses also usually presume that there is no supernatural dimension to religious experience, a presumption which goes well beyond what science can demonstrate. Religiously grounded guilt may mediate avoidance of and restraint of unacceptable sexual impulses, or may represent the internalization of interpersonal patterns of approval and disapproval in the social environment, but it may also be the natural outcome of doing something wrong and a sign of a supernatural presence quickening one's need for repentance. Collaborative exploration of the variety of meanings of religious phenomena with a client can facilitate growth in awareness and understanding in the client and the clinician.

It is important to remember that much of the research on the psychology and sociology of religion has been done on white middle-class populations, and may not tap dimensions of relevance to minority populations, whether racial/ethnic (the experience of African-Americans) or religious (the experience of Jehovah's Witnesses) minorities. This should highlight how important it is for the clinician to be sensitive to all minority issues, particularly when the individual's life is diverse in many ways from the life of the clinician. As in expressing empathy with affective responses of clients, so also in the area of religion should clinicians check their understandings of the client's experiences regularly and with a spirit of willingness to be instructed and corrected by our clients.

Diagnostically, the *Diagnostic and statistical manual of mental disorders,* 4th ed. (1994) devotes only a single paragraph to "religious or spiritual problem." This would hardly seem to demand the development or use of diagnostic assessment instruments measuring religious constructs. Even within the multiaxial system, there is little or no place to address religious or spiritual factors in formulation of the client's "clinical picture." Considering the importance of the religious diversity for the majority of persons, this relative neglect is striking.

10.10.6 THE EMPIRICAL BASIS FOR APPRECIATING RELIGIOUS DIVERSITY IN CLINICAL CARE

Our brief and selective overview of empirical research on religion and clinical care will be organized around the findings of the excellent review by Worthington et al. (1996) of empirical research since the late 1980s.

10.10.6.1 Religious Persons and Mental Health

Worthington et al. (1996, p. 451) recall that religion has been viewed by many (such as Albert Ellis) as associated with irrationality and psychopathology. A large number of studies have produced relatively consistent findings that religion *per se* does not negatively impact mental health in general, and that the mental health of religious persons is often positively effected by their religiosity in a variety of ways as discussed earlier.

A number of studies discussed in Worthington et al. (1996, p. 451) have found superior mental health outcomes for intrinsically religious individuals; extrinsically oriented persons may even experience some negative impact from their religiosity. Therefore, assessing the religious orientation (intrinsic vs. extrinsic) of the client may provide some important data for the clinician as to the potential impact of religion on the mental health of the client in the process of clinical care. Bergin (1991, cited in Worthington et al., p. 457) found that intrinsically oriented religious persons, while they may tend to frustrate the clinician with their religious explanations for their behavior and experiences, are more likely to be open to therapeutic change than the extrinsically oriented person. Pargament (1987, cited in Worthington et al.) found that religiously conservative people may be more open to therapeutic change than people who typically associate themselves with more mainline religious groups. As in other matters of diversity, stereotypes (in this case the stereotype that highly religious individuals are defensive and rigid) can be very dangerous.

Intrinsically and extrinsically oriented people differ on other life dimensions as well. Worthington et al. (1996, pp. 451, 457) cited the findings of Hood et al. (1996) that there are significant differences in how intrinsically religious students, extrinsically oriented students, and proreligious students (both intrinsically and extrinsically oriented) describe identical sensory experiences. Intrinsically oriented students give religious descriptions of their experience spontaneously, while extrinsically oriented students do not use religious descriptions of their experience even when prompted to do so, and proreligious students tend to only mention religion in their descriptions if prompted. Clinically understanding the different ways in which the religious orientation of the client may predispose the client's use of religious explanations in therapy could be of significant help in the clinical care of the client.

10.10.6.2 Psychotherapy and Values

It is widely recognized that successful psychotherapy often entails a certain degree of convergence in values, with the values of the client moving to be more like those of the clinician. It appears that psychotherapy does not typically result in movement in basic religious values of clients, but the close articulation of so many values with religious faith makes this an area of concern for many religious persons seeking counseling and psychotherapy (Worthington et al., 1996). Schwartz and Huismans (1995) report in their article on value priorities and religiosity that there is a negative correlation between religiosity and values such as universalism (understanding, appreciation, tolerance, and protection for the welfare of all people and for nature), stimulation (excitement, novelty, and challenge in life), and self-direction (independent thought and action; choosing, creating, exploring). Their research also suggested that "valuing openness to change and free self-expression inclines people to become less religious." Conversely, they found that "valuing certainty, self-restraint and submission to superior external verities inclines people to become more religious in general."

Given that psychotherapists are disproportionately nonreligious compared to their clientele, and that psychotherapy is a value-changing relationship, such findings can be seen as justifying concerns about value influence for religious clients; it becomes obvious that value differences based on religiosity may significantly impact the work of clinicians who are not sensitive to the values reflected in the religious beliefs of their clients. It may be for this very reason that the data suggests that the world views of religious people may lead them to prefer religious counselors (Worthington et al., 1996). Highly religious persons appear to prefer therapists of very similar religious beliefs. They may also have clear expectations of religious counseling. Highly religious individuals may tend to view the world with religious schema and may view psychotherapy differently as a result.

Shafranske and Malony (1990) discovered in their study of the nature of clinical psychologists" religiousness that, while their sample of clinical psychologists appeared to value the role of religion in human experience in general terms, they were also less likely to be involved with religious institutions. "Less than one in five declared organized religion to be their primary source of spirituality." Approximately 25% reported negative feelings regarding past religious experiences. They also found support for the findings of previous studies that personal attitudes appear to play a more important role than clinical training when it comes to therapeutic interventions related to religion. In fact, they suggest that "it may be that religious beliefs function as a meta-theory that significantly influences psychotherapy in both implicit and explicit ways." As Jones (1994) points out, " ... it seems that the concerns presented by clients often push the practitioner beyond the limits of what consensually validated scientific research has established. 'Given that research supplies only a small fraction of the information needed to completely understand the psychotherapeutic process, we are often compelled to rely on our tacit, background metaphysical notions' " (p. 191).

10.10.6.3 Religion in Psychotherapy

Religion can be best viewed as a multidimensional variable that includes facets such as what people believe, feel, do, know, and how they respond to their beliefs. Thinking in terms of the many important dimensions of religious experience helps to avoid the possibility of reductionistic thinking (e.g., their religion is "nothing but . . . ") and encourages more holistic and integrated thinking, that is, the type of thinking that should characterize the psychotherapy process.

The "bias" among clinical psychologists appears to be overwhelmingly functional. In other words, when they speak about the religiosity of a client, they are most often referring to the intersection of psychodynamics and faith, that is, how it operates in an individual's life (Groeschel, 1992). It is suspected that it is exceedingly rare for most clinical psychologists to be able to speak directly to religion as both a process and as a set of particular beliefs (Shafranske, 1996). Clearly, institutional identification with a particular religious tradition is not generally seen as a top priority of the majority of clinical psychologists (Worthington et al., 1996). If the majority of clients are religious and if the majority of providers do not identify with institutional religion or are unable to think about religiosity beyond the functional, there is a clear possibility that clinical psychologists and their clients may find it difficult to connect on issues which give them focus and meaning in their lives (see for further discussion Hood et al., 1996, Kauffmann, 1991; Paloutzian, 1996).

Worthington et al. (1996) summarize their findings by saying, "highly religious people may prefer religious counselors and explicitly religious counseling .... Despite preferring religious counselors, people do not want their counseling to focus mainly on religion" (p. 460). While it may be a truism, it is difficult to appreciate the impact of the experience of another human being upon themselves or upon clinicians if they have no reference point with which to compare or have no particular knowledge with which to interpret that experience. Clearly, if clinical psychologists are going to do therapy with religious persons, it would seem logical that they should have some reference point or knowledge of religious experience. It is contended that the religious knowledge and experience of the clinician plays an important role in clinical work with religious persons even if the client is not in therapy to address religious or spiritual issues.

10.10.6.4 Religion and Clinical Judgment, Technique and Behavior

Worthington et al. (1996, p. 467) refer to previous studies more than 10 years old which found that the religious orientation of the clinician did not seem to influence the diagnosis given to either religious or nonreligious clients. They also found that religious orientation did not appear to influence the ability of clinicians to discriminate between religious experiences that were either real or fabricated as well as the pathological or responsible use of religion. However, more recent research does not appear to be as clear. According to Worthington et al. (1996), studies done since the mid-1980s have had more mixed results. In fact, they cite a particularly methodologically sound study by Gartner, Harmatz, Hohmann, and Larson (1990) which found that clinician ratings were affected by patient ideology. It seems that in light of the "Ethical Principles of Psychologists and Code of Conduct" (APA, 1992), it would be important for clinicians to be aware of this potential bias.

Religious variables can affect the formation of a therapeutic relationship. Worthington et al. (1996) summarize the research by noting that "When counselors disclose their religious beliefs or values, their disclosure will likely affect both the client's behavior and expectations about the counseling process and outcome. Disclosing a counselor's religious beliefs and values can facilitate counseling if the counselor and client are quite similar in beliefs and values and if the counseling does not focus mainly on religion" (p. 460). In the light of managed care and the concerns of the public about accountability in the practice of psychotherapy, it may be well for us to be concerned about the values clients bring to therapy and their dissonance with the clinician's.

Research studies cited by Worthington et al. (1996) have investigated the role religion plays in other clinical behaviors. It appears that there is a positive correlation between clergy exposure to mental health issues and training and the propensity of clergy to do counseling and to refer. There would also appear to be a relationship between the clinician's theoretical orientation and the referral behaviors of clergy. It appears that mainline Protestant clergy are more likely to refer to humanistic or behavioral therapists and that clergy from more conservative fundamentalist or orthodox faiths are more reluctant to refer to therapists of a psychodynamic orientation. Interestingly, they also state that "Secular professionals rarely refer to clergy, even when difficult spiritual issues arise in counseling" (p. 468). It would seem important for clinicians to give clergy the same referral treatment we do other helping professionals. Perhaps it is an imperialistic attitude or a lack of relationships that prevents such from happening.

Holden et al. (cited in Worthington et al., 1996, p. 468) found that both counselors and clergy demonstrated equally developed skills in assessing the interpretation accuracy of Judeo-Christian principles in the religious ideation of a depressed client. Further, counselors were more reluctant than clergy to challenge those religious beliefs. The question might be asked; if clinicians are reluctant to challenge religious beliefs that they accurately perceive as distorted, why don't they refer the client to someone who may help the client by doing just that?

Some explicitly or implicitly religious techniques are used regularly in psychotherapy. Worthington et al. (1996) state, "The use of religious techniques by explicitly religious therapists stands in some contrast to the general field of clinical psychology" (p. 469). They cite an important study by Shafranske and Malony (1990) in which they did a national survey of clinical psychologists. They reported that 59% of the clinicians surveyed supported the use of religious language in psychotherapy but 55% opposed the use of scriptures. Further, only 19% of the clinician found it acceptable to pray with a client while 68% believed it inappropriate. However, as previously noted, the public seems to feel prayer is important to their physical health and the majority would like their physician to pray for them. Clearly, the issue of prayer in mental health care needs to be investigated further. Worthington et al. (1996) state, "Prayer appears to be the most common form of religious coping by most religious people, and even nonreligious people often turn to prayer in the throes of suffering" (p. 474).

Another religious behavior that has found its way frequently into clinical work is forgiveness. Worthington et al. (1996, p. 475) note that forgiveness has been almost as popular in the psychological literature as the religious journals. Unfortunately, they make a similar assessment of research on forgiveness as other religious interventions. While forgiveness is used with self-reported efficacy in case studies, little empirical attention has been given to its potential efficacy in clinical populations. Also, while forgiving may have efficacy as an intervention, few studies have investigated the potential effect of seeking forgiveness in clinical populations. This may have particular interest for those investigating personality disorders such as antisocial personality disorder.

A particularly problematic clinical intervention for clients of the Western religious traditions has been the Eastern Hindu and Buddhist oriented meditation techniques. Because of their close connection with specific practices of a religious system, many highly religious persons have objected to their use or have been surprised, given their religious connections, at their acceptance by the clinical community. A study by Carlson, Bacaseta, and Simanton (1988, cited in Worthington et al., 1996) found progressive relaxation and Christian meditation equal in their efficacy. Worthington et al. (1996) state, "Most of what can be accomplished therapeutically with meditation can be accomplished with relaxation training which is generally easier and avoids religious associations of meditation" (p. 475). This is an example of an exclusive religious practice that has mental health benefits and can be used inclusively when modified to meet the religious diversity concerns of others.

10.10.6.5 Religion and Clinical Research

Worthington et al. (1996) note that religion has become an increasingly acceptable topic for research in counseling and psychotherapy in part due to the "fourth force" in psychology, i.e. multiculturalism. According to the authors, religion and spirituality has moved into the mainstream of clinical care in many ways. This has been demonstrated by the increase in religiously oriented professional organizations in the field, the development of doctoral level training programs that have a religious orientation, the number of conferences and workshops presented in mainstream marketplaces such as the APA preconvention workshops, and the publishing of a "plethora of theoretical, polemic, and conceptual works" (p. 448). They suggest that there has been an improvement in the scientific study of counseling for the religious and of religious counseling since the mid-1980s. They conclude that "religious counseling by religious counselors of religious clients has recently assumed an increased prominence" (p. 449), and note that the changes occurring in the mental health care marketplace will have a continuing effect on how religious clients and religious counselors experience mental health care (p. 480).

10.10.7 THE RELIGIOUS COMMUNITY AS A RESOURCE FOR SUPPORT

It seems clear that there are many signals in contemporary North American society that persons are looking for the kind of social support that can only be found in a community (Kauffmann, 1991). This phenomenon appears to be particularly relevant to the Christian church. When the church is at its best, there is concern for the total welfare of its members and friends: "By prayer and petition for the Spirit's leading, by identifying the resources it possesses, and by framing programs consistent with the best empirical findings, congregations can provide a much needed service to both believers and nonbelievers" (Kauffmann, 1991, pp. 134–135).

Particularly in the era of decreased resources due to managed care, it might be useful to think about a religious group as a potential therapeutic community. Indeed, the data (Miller & Jackson, 1995) would support the observation that religious professionals, rather than mental health or health care providers, are often the first persons contacted and sought out when individuals or families are in a crisis. A failure to recognize this pattern and to consider networking with local religious groups might be one way to guarantee that even the best effort in the consulting room will not generalize or maintain beyond the immediate professional involvement. Religious professionals and members of local religious communities should be viewed as resources and as potential collaborators. Indeed, professional isolation (Guy, 1987) may be hazardous to clinicians" well-being as well as that of clients.

10.10.8 TRAINING CLINICAL PSYCHOLOGISTS IN RELIGIOUS DIVERSITY

It is hard to find truly balanced and informed treatments of creative and even curative possibilities of the resources of faith and religious communities in theoretical and research literature, although there are some encouraging trends in the field (Browning, 1987; Hood et al., 1996; Jones & Butman, 1991; Malony, 1995; McLemore, 1982; Meyer & Deitsch, 1996; Miller & Jackson, 1995; Shafranske, 1996; Worthington et al., 1996; Wulff, 1991). The lack of epistemic humility when studying the traditions of others is a truly disturbing characteristics of training programs and clinical practice.

Shafranske and Malony (1990) concluded in their study of the religiousness of clinical psychologists that only one-third of their subjects "expressed personal competence" to intervene in the religious aspects of their clients" lives. This was contrasted with the finding that the majority felt they had the knowledge and ability to deal with religious issues. Further, 85% reported the frequency of discussion related to religious issues in their training experiences to be rare or never. Along with other past reviewers, they concluded that clinical psychologists should "receive limited training respective of religious and spiritual issues" and that:

> in light of the limited training opportunities, the profession may have failed to heed the admonition of Perry London that while modern psychotherapists have grown up in the tradition of medicine, the nature of the ailments they deal with and the way they treat them, make them function much like clergy.

The profession's concern about diversity in all its form is to be applauded. It can be argued that religious diversity is an important multicultural factor that has not been treated equally with others in addressing the need for training. Petersen (1988, cited in Brems, 1993, pp. 72–86) proposes a threefold approach to increase multicultural sensitivity: awareness, knowledge, and skills. Petersen's model is developmental in that each of the areas builds on the other and thereby assumes a process of growth.

Awareness would be the first step in becoming sensitive to the issue of religious diversity as a multicultural issue. Brems (1993, p. 74) describes seven characteristics of an aware clinician. First, an aware clinician would have an awareness of their own cultural heritage. The clinician could participate in this by self-reflection on their own religious or spiritual journey and seek to share that story with a colleague. Second, Brems proposes that the clinician be conscious and embracing of all the diversity memberships in their life. This may involve the clinician making an intentional connection between ethnic heritage, gender identification, and religious background and how these issues in concert may impact her clinical work with others, particularly in light of the fact that they may not share all of the same diversity issues with the client. Third, the clinician should "value and respect" the diversity of others. This may involve the intentional exposure to other religious groups, orientations, and variety of religious experiences through attending institutional functions or reading popular literature or even dialoging with a colleague or friend who has a religious experience very different from one's own. Fourth, the clinician should become aware of their own values and potential bias and the potential effect they may have on therapy. This could come as a result of dialoging with a colleague or being supervised by a colleague with more knowledge and expertise about the particular religious affiliation of a client. This would hopefully enable the clinician to become more aware of the salient differences and their potential impact on therapy.

Fifth, Brems (1993) would suggest that the aware clinician should be careful not to "overemphasize" or "underemphasize" the differences from the client. This might be observed by the clinician or the client. The clinician might consider seeking common religious ground with the client by comparing and contrasting religious ideas. They might ask about the client's concern with regard to religious symbols in their office, for example, or might seek clarification of the religious language used by the client. Sixth, the clinician should seek to be comfortable with and diversity that exists between themself and their client. This might be accomplished by the clinician asking to be "educated" by the client with regard to the client's religious background while expressing genuine interest in understanding the impact religion has for the client. The clinician might read about the client's religion independently and dialog about that with the client. Finally, the clinician should seek to be sensitive to the need for referral should that be in the best interest of the client and the diversity issues present.

Brems (1993) also describes the further development of knowledge in the process of becoming sensitive to multicultural issues in therapy. The development of knowledge involves gaining accurate information about the client's religion. It also involves being sensitive to the fact that, even within very structured religious groups, individuals within the group may think, feel, and act differently than their cohorts as it relates to their religious expression. Accurate information should not be limited to just the writing of psychologists but should also involve the writings of both supporters and critics of the religion. First-hand experience can also be a powerful source of imagery as it relates to understanding the client's experience. Finally, the clinician should become aware of the past injustices and inadequacies of the mental health system to adequately address the needs of the religious client. The clinician may need to rethink their conclusions about psychopathology when religious variables predominate.

Brems (1993) describes the skills necessary to successfully deal with multicultural issues in therapy. This includes the matching of communication and therapeutic orientation to meet the client's specific diversity needs, being careful not to stereotype or categorize, and remaining flexible in meeting the client's needs. This may involve being an agent for social change by using language that is appropriate and not prejudicial. Most importantly for the religious client, a therapist may need to have the contacts and relationships within the client's religious community to promote or participate in institutional interventions or referrals. This process is no different to that of the establishment of referral relationships with other mental health care providers who intervene in ways unique to their discipline.

10.10.9 FUTURE DIRECTIONS

According to *The encyclopedia of American religions* (1993), there are 1730 "primary religious bodies" in America alone (cited in Paloutzian, 1996, p. 7). Collectively, these groups are described in the psychology of religion as churches, denominations, sects, or cults. As Paloutzian (1996) has noted, if to this the full scope of world religions is added including varieties of Buddhism, Hinduism, New Age religions, Spiritualist, Wisdom religions, and others, the numbers are staggering. Even a full-time student of domestic or world religions would have to devote decades of serious study in order to fully appreciate the complexities and subtleties of the ideological, ritualistic, experiential, intellectual, and consequential dimensions of all these religious commitments. Certainly, this would be an unrealistic task for even the most interested and committed clinician or clinician-in-training.

It is not proposed that clinicians become experts in religion or spirituality any more than it is expected that clinicians be experts in law (although some are experts in forensic psychology) or medicine (although some are very knowledgeable about psychopharmacology). Even if limiting oneself to Anglo-American religions (assuming one is from an Anglo-American religious and cultural heritage), it would be an unnecessary and daunting task. What is being argued in this chapter is a need for sensitivity, awareness, and acknowledgment of the religious diversity (or lack thereof) of clinicians and of the religious diversity of clients. Further, it has been the intention to make the case for understanding the role of religion in the experience of clients not only from a psychosocial perspective but from a transcendent one as well.

It is suggested that religion be treated with the same respect and concern for competency as other areas of diversity and that clinicians recognize their abilities as well as lack of ability to deal with religious issues in the assessment, diagnosis, and treatment of their clients. A model is proposed for such clinical activity that is consistent with other models of competency including forensic psychology, or even more appropriately, psychopharmacology.

What is required is an Assess–Treat–Refer model of clinical care as it relates to clinicians" religious diversity and that of their clients. The "Assess" aspect of this model involves identifying control beliefs, religious beliefs, and commitments, and the impact they may have on the care of clients. The clinician should not keep this totally to themself, but rather in keeping with ethical guidelines be explicit with the client about differences that may impact clinical care.

The clinician might attempt to articulate spirituality/religiosity through a self-assessment by reflecting on the dimensions of religion noted (cognitive, ritualistic and symbolic, moral, institutional, community and lifestyle, and experiential). Further, the clinician may need to increase sensitivity to the limits of knowledge and the limits of attitudinal flexibility. They may also need to reflect on the implicit assumption that nobody can support all religions equally and assess the level of empathy with different traditions. Finally, in the assessment of their own control beliefs about religion and spirituality the clinician should be careful not to make personal synthesis of religion a "reality" but rather recognize it as a variant of religious belief itself.

Additionally, the assessment aspect of this model would involve exploring the client's religious background and control beliefs in an attempt to not only understand the psychosocial aspects of their religious experience but also the potential resource that religion might be in their clinical care. The clinician should be careful not to only make a functional analysis of the client's religious beliefs and behaviors (as helpful as that can be), but rather consider the implications that religion has for the client's interpretation of ultimate reality.

While there are many instruments available for measuring the religious aspects of a client's experience, the assessment of religion/spirituality as part of the regular clinical interview may be the most helpful place to start. As previously noted, the clinical interview should explore the various dimensions of religious experience and behavior as well as serve as a helpful understanding of the role of religion in the life of the client. The assessment process should be sensitive to bias that may be inherent in soliciting such information. Clinicians should pay careful attention to and encourage the continuing explosion of research on clinical processes and outcomes which includes careful study of religion and spirituality and how they interact with clinical variables.

The "treat" aspect of the model recognizes that some clinicians may identify readily with a particular religious tradition in ways that make them uniquely qualified to care for particular clients. These clients may present with religious issues that are very significant, in not only formalizing their case, but also in treating them.

It is important to note that the standard of care has changed and continues to evolve in mental health care. The focus on managed care and utilization review has produced a system of care that is multidisciplinary and interdisciplinary. This model of care is held out to be not only in the best interest of the client, but it also recognizes that the clinician cannot provide all of the beneficial care that a client deserves or that may be available. While some clinicians may be relative experts in a particular religious context or identify with particular religious systems, even then the benefits of accessing clergy or other religious professionals may be more appropriate in meeting the client's religious and/or spiritual needs as part of a multidisciplinary approach to mental health care.

Certainly, the competence to treat involves not just multicultural awareness, knowledge, and skills but will ultimately be decided on by effectiveness in dealing with a particular client with a particular problem and in a particular context. With the increasing call for accountability and responsibility for clinical care, it would appear to be important to focus on outcomes as a particular measure in the care of the religiously diverse client. Outcome measures should certainly have clinical utility (changes in symptoms, behaviors, mood, relationships, and so on), but also need to measure consumer satisfaction. Broadly understood, this could involve the client, significant others in the client's life, and, for the religiously diverse client, the religious community. A potentially important measure of the effectiveness in treating the religiously diverse client may be the evaluation by the religious community of the impact of work on the client in the context of their faith.

Finally, the "Refer" aspect of the model would suggest the obvious. There are times when competencies are clearly limited in cases that involve religious issues. We should seek to network with religious professionals similar to networks we may have with other healers. Clinicians should be open and clear with clients that deference in caring for them is not a function of inability to morally recognize their religious diversity but rather it is a matter of limited competence and the specialized ability of others to whom they can be referred.

A case example may illustrate the model. An 18-year-old male was referred for psychological evaluation by his public school because of "aggressive and violent behavior" towards peers and staff. The young man was a new immigrant from Bosnia. His primary language was Bosnian, and so there was a language barrier that would require the assistance of a translator. An initial hypothesis was that this "language barrier" in other relationships and the resulting isolation and frustration may have played a role in this young man's "aggressive and violent behavior."

While the examiner was aware of a "religious war" in the former Yugoslavia, the "religious" aspect of the war was not well understood by the clinician. In taking the clinical history of this young man it was discovered that he had immigrated to the US through the auspices of a "Christian" organization. Prior to that, he had been held in a refugee camp which he described as very undesirable. While at the camp, he was identified as one who could immigrate because of his "circumstances," those circumstances being the murder of his entire family, in his presence, by "Christians." They were killed, he stated, because they were "Muslims."

He was a Muslim. He stated that he would die for Islam. He also stated that his chief end in life was to revenge the death of his family and that this was dictated by his Muslim religion. Having only a rudimentary understanding of Islam and noting that revenge was not one of the five pillars of faith, the clinician was compelled to seek the assistance of a Muslim cleric who could clarify the issues and counsel this young man about this issue in particular. It seemed better to the client to discuss his religious beliefs with a Muslim cleric than with a Christian clinician.

It became clear in conceptualizing the nature of the "aggressive and violent behavior" that this young man was very angry over the loss of his family, had struggled to understand and be understood in a culture with a language very different from his own, and that the client was conflicted over his hatred for "Christians" and their apparent goodwill toward himself. Obviously, the assessment of this young man's presenting problem hinged upon an understanding of his religious and sociocultural background. The treatment of his problem required the special expertise of a Muslim cleric as well as a recognition of the therapist's identification as a "Christian" Ultimately, his care involved the multidisciplinary involvement of a foreign language teacher, a Muslim cleric, and a clinical psychologist.

10.10.10 SUMMARY

Clinical psychologists are nonrepresentative of the general population in terms of their lack of commitment to traditional religious faiths. Religious faith, like other vital diversity variables such as race, gender, ethnicity, age, sexual orientation, and others, shapes and contextualizes the clinical concerns of persons who present for psychological assessment and treatment. Clinical psychologists have an ethical obligation to become more knowledgeable, aware, and skilled in managing the impact which religious variables, broadly conceived, can have on the process and outcome of psychological service delivery. They must also become more aware of their own biases and limitations in dealing with such material, more aware of the degree to which their own religious assumptions color their theories and practices in clinical psychology, more tolerant of the wide array of religious beliefs and practices which characterize the people they serve, more respectful of the resource which religious beliefs, practices, communities, and institutions are to many clients, and more cognizant of the extensive and growing research base on the role of religion in human life generally and in psychological practice in particular. Contrary to the prevailing mindset in psychological circles, for most of the population, to be spiritual is to be religious, and psychologists must avoid the tendency to "reduce" the religious faith of those they serve to either a bland, generic spirituality or to a set of sociopsychological processes.

10.10.11 REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington: American Psychiatric Association.

American Psychiatric Association Task Force on Religion and Psychiatry (1975). *Psychiatrists viewpoints on religion and their services to religious institutions and the ministry.* Washington, DC: American Psychiatric Association.

American Psychological Association (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47,* 1597–1611.

Batson, C. D., Schoenrade, P., & Ventis, W. L. (1993). *Religion and the individual: A social-psychological perspective,* New York: Oxford University Press.

Benner, D. G. (Ed.) (1988). *Psychology and religion.* Grand Rapids, MI: Zondervan.

Bergin, A. E. (1991). Values and religious issues in psychotherapy and mental health. *American Psychologist, 46,* 394–403.

Bergin, A. E., & Jensen, J. (1990). Religiosity of psychotherapists: A national survey. *Psychotherapy, 27,* 3–7.

Brems, C. (1993). *A comprehensive guide to child psychotherapy.* Boston: Allyn & Bacon.

Browning, D. S. (1987). *Religious thought and the modern psychologies: A critical conversation in the theology of culture.* Philadelphia: Fortress Press.

Coles, R. (1990). *Harvard diary: Reflections on the sacred and the secular.* New York: Crossroad.

Comas-Díaz, L., & Striker, G. (1993). Special issue: Diversity in clinical psychology. *The Clinical Psychologist, 46*(2), 88–89.

Dell'Olio, A. (1996). Multiculturalism and religious diversity: A Christian perspective. *Christian Scholars Review, 25,* 459–477.

Fowler, J. (1981). *Stages of faith.* San Francisco: Harper & Row.

Friedman, R., & Benson, H. (1997). Spirituality and medicine. *Mind/Body Medicine, 2*(1), 1.

Gartner, J. (1996). Religious commitment, mental health, and prosocial behavior: A review of the empirical literature. In E. P. Shafranske (Ed.), *Religion and the clinical practice of psychology* (pp. 187–214). Washington, DC: American Psychological Association.

Gartner, J., Harmatz, M., Hohmann, A., & Larson, D. (1990). The effect of client and counselor values on clinical judgment. *Counseling and Values, 35,* 58–62.

Glock, C. (1962). On the study of religious commitment: Review of recent research bearing on religious and character formation. *Religious Education, 57*(4), S98–S110.

Gorsuch, R. L. (1984). Measurement: The boon and bane of investigating religion. *American Psychologist, 39*(3), 228–236.

Groeschel, B. J. (1992). *Spiritual passages: The psychology of religious development*. New York: Crossroad.

Gross, M. L. (1978). *The psychological society*. New York: Random House.

Guy, J. D. (1987). *The personal life of the psychotherapist*. New York: Wiley.

Hill, P. C., Butman, R. E., & Hood, R. W., Jr. (Eds.) (1997). *Measures of religious behavior*. Birmingham, AL: Religious Education Press..

Hood, R. W., Spilka, B., Hunsberger, B., & Gorsuch, R. (1996). *The psychology of religion* (2nd ed.). New York: Guilford Press.

James, W. (1910/1963). Pragmatism. In P. R. Reynolds (Ed.) *William James: Pragmatism and other essays* (pp. 119–132). New York: Simon & Schuster.

Jones, S. L. (1994). A constructive relationship for religion with the science and profession of psychology: perhaps the boldest model yet. *American Psychologist, 49*(3), 184–199.

Jones, S. L., & Butman, R. E. (1991). *Modern psychotherapies: A comprehensive Christian appraisal.* Downers Grove, IL: InterVarsity Press.

Kauffmann, D. (1991). *My faith's ok—your faith's not: Reflections on the psychology of religion*. Goshen, IN: Goshen College.

Kung, H. (1985). *Christianity and world religions*. Maryknoll, NY: Orbis Books.

Larson, D., Lu, F., & Swyers, J. (1996, May). *Model curriculum for psychiatric residency training programs: religion and spirituality in clinical practice* (available from first author and used by permission. Published by The National Institute for Healthcare Research, Rockville, MD).

Lash, N. (1988). *Easter in ordinary: Reflections on human experience and the knowledge of God.* Notre Dame, IN: University of Notre Dame Press.

London, P. (1986). *The modes and morals of psychotherapy* (2nd ed.). Washington, DC: Hemisphere.

MacDonald, D. A., LeClair, L., Holland, C. J., Alter, A., & Friedman, H. L. (1995). A survey of measures on transpersonal constructs. *Journal of Transpersonal Psychology, 27*(2), 171–235.

Malony, H. N. (1991). *The psychology of religion: Personalities, problems, possibilities.* Grand Rapids, MI: Baker.

Malony, H. N. (1995). *The psychology of religion for ministry*. Pasadena, CA: Integration Press.

Matthews, D. A. (1997). Religion and spirituality in primary care. *Mind/Body Medicine, 2*(1), 9–19.

McLemore, C. W. (1982). *The scandal of psychotherapy*. Wheaton, IL: Tyndale.

Meadow, M. J., & Kahoe, R. D. (1984). *Psychology of religion: Religion in individual lives*. New York: Harper and Row.

Meyer, R. G., & Deitsch, S. E. (1996). *The clinician's handbook: Integrated assessment in adult and adolescent psychopathology* (4th ed.). Boston: Allyn and Bacon.

Miller, W. R., & Jackson, K. A. (1995). *Practical psychology for pastors* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Nielsen, N. C. (1993). *Religions of the world* (3rd ed.). New York: St. Martins Press.

Noss, J. B., & Noss, D. S. (1993). *A history of the world's religions* (9th ed.). New York: Macmillan.

Paden, W. E. (1988). *Religious worlds: The comparative study of religion.* Boston: Beacon Press.

Paloutzian, R. G. (1996). *Invitation to the psychology of religion* (2nd ed.). Boston: Allyn and Bacon.

Pargament, K., Silverman, W., Johnson, S., Echemendia, R., & Snyder, S. (1983). The psychosocial climate of religious congregations. *American Journal of Community Psychology, 11,* 351–381.

Pargament, K., Tyler, F. B., & Steele, R. E. (1979). The church/synagogue and the psychosocial competence of the member: An initial inquiry into a neglected dimension. *American Journal of Community Psychology, 7,* 649–664.

Schwartz, S. H., & Huismans, S. (1995). Value priorities and religiosity in four western religions. *Social Psychology Quarterly, 58*(2), 88–107.

Shafranske, E. P. (Ed.) (1996). *Religion and the clinical practice of psychology*. Washington, DC: American Psychological Association.

Shafranske, E. P., & Malony, H. N. (1990). Clinical psychologists" religious and spiritual orientations and their practice of psychotherapy. *Psychotherapy, 27*(1), 72–78.

Smart, N. (1968). *The religious experience of mankind.* New York: Charles Scribner.

Smart, N. (1983). *Worldviews: Cross-cultural explorations of human beliefs.* New York: Charles Scribner.

Smart, N. (1989). *The world's religions.* Cambridge, UK: Cambridge University Press.

Stevenson, L. (1987). *Seven theories of human nature* (2nd ed.). New York: Oxford University Press.

Szasz, T. (1977). *The theology of medicine*. Syracuse, NY: Syracuse University Press.

Tisdale, J. R. (1980). *Growing edges in the psychology of religion*. Chicago: Nelson-Hall.

Williams, D. R. (1993). The measurement of religion in epidemiologic studies: Problems and prospects. In J. S. Levin (Ed.), *Religion in aging and health: Theoretical formulations and methodological frontiers* (pp. 125–148). Thousand Oaks, CA: Sage.

Worthington, E. L., Kurusu, T. A., McCullough, M. E., & Sandage, S. J. (1996). Empirical research on religion and psychotherapeutic processes and outcomes: A 10-year review and research prospectus. *Psychological Bulletin, 119*(3), 448–487.

Wright, R. (1994). *The moral animal: The new science of evolutionary psychology*. New York: Pantheon.

Wulff, D. M. (1991). *Psychology of religion: Classic and contemporary views*. New York: Wiley.

Zilbergeld, B. (1983). *The shrinking of America: Myths of psychological change.* Boston: Little, Brown, and Company.

Zinnbauer, B. J., Pargament, K. I., Cowell, B., & Scott, A. B. (1996). *Religion and spirituality: unfuzzying the fuzzy*. Paper pressented at the American Psychological Association Annual conference. Toronto, Canada. August, 1996.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.11
Mental Health in Rural Society

MICHAEL MURRAY
Memorial University of Newfoundland, St. John's, NF, Canada
DAVID S. HARGROVE
University of Mississippi, MS, USA
and
MICHAEL BLANK
University of Virginia, Charlottesville, VA, USA

10.11.1 INTRODUCTION 256

10.11.2 STRUCTURE OF RURAL SOCIETY 256
- *10.11.2.1 Changing Demographics* 256
- *10.11.2.2 Definitions of Rural* 256
- *10.11.2.3 Diversity* 257
- *10.11.2.4 Age and Education* 257
- *10.11.2.5 Work* 257
- *10.11.2.6 Income* 258
- *10.11.2.7 Living Conditions* 258

10.11.3 LIFE IN RURAL SOCIETY 258
- *10.11.3.1 Social Life* 258
- *10.11.3.2 Women and Family Life* 259
- *10.11.3.3 Working Life* 259
- *10.11.3.4 Change in Rural Society* 260

10.11.4 HEALTH ISSUES 261
- *10.11.4.1 Stress in Rural Society* 261
- *10.11.4.2 Impact of Crisis* 261
- *10.11.4.3 Social Cohesion* 261
- *10.11.4.4 Mental Illness* 262

10.11.5 MENTAL HEALTH SERVICES 263
- *10.11.5.1 Healthcare Utilization* 263
- *10.11.5.2 Specialty Mental Health Services* 263
- *10.11.5.3 Public Policy* 264
- *10.11.5.4 Empirical and Theoretical Basis for Planning Services* 264
- *10.11.5.5 Healthcare Reform* 265
- *10.11.5.6 Primary Care* 265
- *10.11.5.7 Telemedicine* 266

10.11.6 PSYCHOLOGY AND RURAL MENTAL HEALTH 267
- *10.11.6.1 Historical Background* 267
- *10.11.6.2 Roles of Psychologists in Rural Communities* 267
- *10.11.6.3 Training Psychologists for Rural Practice* 268

10.11.7 TYPES OF INTERVENTION 269
10.11.7.1 Fitting the Intervention to the Context 269
10.11.7.2 The Rural Psychologist as a Generalist 270
10.11.7.3 The Concept of Community and Psychological Intervention 270
10.11.7.4 Ethical Dilemmas in Rural Practice 271
10.11.8 FUTURE DIRECTIONS 272
10.11.9 SUMMARY 272
10.11.10 REFERENCES 273

10.11.1 INTRODUCTION

The twentieth century has seen a dramatic shift in the distribution of the population over most parts of the globe, including North America. Whereas in the early part of this century the majority of the world's population lived in small communities relatively isolated from population centers, the world has now become a predominantly urban society. Large metropolitan centers have not only become the location for an increasing proportion of the world's population but also of service provison and, indeed, of human research interest.

This change has led to an orientation towards understanding and providing for the needs of urban society and a lesser interest in the nature of rural society and of the various problems it faces. The purpose of this chapter is to focus attention specifically on rural society in North America. It begins by outlining the demographic structure of rural society, proceeds to describe some of the problems rural society is currently experiencing, describes some of the main mental health problems, and then considers the character of health service provision, and the role of psychologists in rural mental health care.

10.11.2 STRUCTURE OF RURAL SOCIETY

10.11.2.1 Changing Demographics

According to the most generous estimates, approximately 25% of North Americans currently live in rural areas. This proportion has declined steadily throughout the twentieth century. However, the absolute number of people living in rural settings has actually increased. Between 1930 and 1990 the rural population in the USA increased by 7.9 million (15%) to 53 million (22%), whereas in Canada it increased from 4.4 million (51%) in 1921 to 6.4 million (23%) in 1991 (45% increase) (Statistics Canada, 1993). This increase in population confirms what Cordes (1990) describes as the myth of a shrinking rural America.

Admittedly, these overall figures conceal a decline in the population of certain rural areas. This decline in population is due to a steady increase in outmigration from those rural regions that are experiencing economic difficulties. Part of the reason for the absolute increase in the rural population is the movement of urban residents to rural or semirural areas from which they commute to urban centres for employment. These numbers are also supplemented by those former urban residents who retire to rural areas.

10.11.2.2 Definitions of Rural

In attempting to describe the character of rural society one endemic problem is the actual definition of rural. There has been ongoing debate about the term. The three most common dimensions used in definitions are (see Bealer, Willis, & Kuvlesky, 1965):

(i) *Ecological.* This concerns the spatial distribution of the population. According to this definition a rural area is one where the population is small, widely distributed, and remote from more urban settings.

(ii) *Occupational.* This suggests that rural refers to a limited number of occupations, specifically those that are involved with agriculture and, to a lesser extent, with fishing, forestry, and mining.

(iii) *Sociocultural.* This refers to the complex of values and behaviors that typify a particular group of people. Stereotypically, rural people are considered to be socially conservative, provincial, and resistant to change.

Admittedly, most studies do not make use of all three dimensions (see Bosack & Perlman, 1982). In compiling statistics it is important to have a common definition. In the USA the favored definition is that developed by the Bureau of Census. It defines urban areas as central cities of 50 000 or more and adjacent territory of more than 2500 residents living outside the geopolitical boundary. From this definition, rural is then defined as the residual category.

Statistics Canada (1993) uses a similar definition describing an urban area as having attained a population concentration of at least

1000 and a population density of at least 400 per square kilometer at the previous census. All territory lying outside urban areas is considered rural. Recently a more sophisticated definition has been developed which classifies US Census returns according to both the population of the locality and the proximity to urban centers. This scheme produces a 17 category continuum from core counties with a population greater than one million to nonmetropolitan nonadjacent counties with no community greater than 2500. This classification scheme has recently been used to clarify the demographics of the USA (Miller, Farmer, & Clarke, 1994). Admittedly, even when using a classification as sophisticated as this, there remains the problem of the substantial variations that exist between and within regions.

10.11.2.3 Diversity

An overall characteristic of rural society in North America is the broad sociocultural diversity of its population. From Southern Blacks, Hispanics in the Southwest, through American indians and Appalachian whites, to Canadian natives and Innuit in the Arctic north, the diversity of cultures, languages, and traditions is extensive.

It is estimated that about 8.3% of the rural population in the USA is Black as opposed to 11.8% in urban areas. Similarly, less than 2.5% of the rural population is Hispanic compared with over 7.5% in urban areas (Cordes, 1989). However, these minorities are concentrated in certain regions. Snipp (1996) argues that racial segregation is as much a feature of rural as of urban America. Besides the enduring poverty of rural society, these minorities also experience discrimination and the threat of physical harm.

While the overall level of poverty is higher in rural than in urban areas, it peaks among the ethnic minorities. According to Lichter (1989), rural blacks "remain among the most economically disadvantaged groups in the United States" (p. 444). Although there was an improvement in the economic position of these minorities during the 1960s and 1970s the evidence suggests that there has been a deterioration since the 1980s (Jensen & Tienda, 1989).

10.11.2.4 Age and Education

The ongoing economic difficulties facing rural America have led to an outflow of young adults and a consequent aging of the rural population. According to 1990 US figures, whereas 11.3% of the population of core urban areas was 65 years and older, 16.2% of those resident in the most rural counties were of a similar age (Miller et al., 1994). In certain pockets, the median age can be much higher. The exodus of many young people in search of employment has led in certain areas to a graying of rural America. Johnson and Beale (1992) note that in many rural communities more than one in four of the population is over 65 years.

These elderly people are a mixture of those who have lived all or most of their lives in rural areas and urban seniors who have retired to a rural setting. The latter tend to be wealthier and to have fewer immediate family connections. There is also evidence that elderly rural residents will migrate to small towns to retire (Li & MacLean, 1989).

The general pattern of education is one with rural residents having poorer educational qualifications than urban residents. For example, whereas only 6.2% of the population of the fringe suburban counties in the USA had less than nine years of schooling, the figure for the most rural counties was 17.7% (Miller et al., 1994). In Canada, while 26% of the general working population has a university education, only 13% of farm opeators have that level of education (Statistics Canada, 1996).

10.11.2.5 Work

While the traditional image of rural work is that of farming, in fact farmers represent only a small proportion of workers in rural areas. According to the 1990 US Census only 11% of the workforce in rural areas was involved in farming (Miller et al., 1994). Admittedly, agriculture may form a central component of the economy of certain rural regions as do the fishery, forestry, and mining industries in others. The central position of these industries in local communities can result in substantial dislocation of the whole community if that industry is threatened. This has occurred in many rural areas since the mid-1970s.

With the decline in the relative importance of these natural resource-based industries, increasingly the largest proportion of the rural workforce is now employed as private wage and salary workers in service and manufacturing industries. In addition, farmworkers will seek temporary nonfarm employment. For example, Bollman and Smith (1988) report that between 1951 and 1981 the number of off-farm work days reported annually by farm operators in Canada increased from 75 to 171. Also during the period 1941 to 1981 the proportion of farm operators reporting working full-time off the farm increased from 3% to 14%.

Increasingly, residents of rural areas have become more dependent upon other sources of income. According to Bender et al. (1985) only

29% of US nonmetropolitan counties are farming-dependent, 28% are manufacturing-dependent, 21% are retirement-dependent, and 8% are mining- and energy-dependent. Further, while the farming-dependent counties only account for 13% of the rural population, the manufacturing-dependent counties account for 40%. Indeed, as some large firms relocate to rural areas in search of cheaper rents and labor, the proportion of rural jobs now attributed to manufacture is similar to that in urban centers (Bluestone & Daberkow, 1986).

The job crisis affecting urban America has its counterpart in rural America. It has been estimated that approximately half the workforce in rural America either does not have a job or has an inadequate one (DeLeon, Wakefield, Schultz, Williams, & VandenBos, 1989). Young people without jobs will move to urban centers in search of employment, while older unemployed workers will tend to remain in the small communities.

10.11.2.6 Income

The average income of rural areas is generally lower than that in urban areas. Analysis across the 17 US census categories revealed that while the median income in core urban areas was \$32 000 and in the more suburban fringe counties it reached \$38 000, it was only \$20 000 in the most rural counties (Miller et al., 1994). In comparison with urban centers a greater proportion of rural dwellers are classified as poor—17% in rural areas compared with 12% in urban areas (Cordes, 1989).

Again, these figures conceal substantial variations even within a single area. For example, in Canada it was estimated that 30% of rural households had an income less than \$20 000 but 25% had an income greater than \$45 000 (Bolaria, Dickinson, & Wotherspoon, 1991). Similarly, a study of the Atlantic fishery (Task Force, 1993) found that while two-thirds of fishermen made less than \$20 000 per annum and nearly one-quarter made less than \$10 000, 11% had average incomes of over \$35 000. Thus while rural areas are generally poorer than urban areas, there also exist pockets of relative affluence and of substantial poverty within these areas.

Further, the evidence suggests that along with the rest of America, rural America is experiencing increasing inequalities in the distribution of wealth (Tolbert & Lyson, 1992). This increasing inequality has implications for the health of the community over and above that due to the adverse effects of the absolute level of poverty (Wilkinson, 1996).

10.11.2.7 Living Conditions

In general, housing in rural areas is of a lower standard than in urban areas. According to White House (1979) figures, rural people in the USA are three times more likely to live in substandard housing than those in urban areas. Similarly, the 1981 Canadian census found that while 1 in 15 homes overall required major repairs in order to meet basic accommodation standards, the comparable figure for rural areas was less than 1 in 10. Further, while 90% of urban homes had central heating, the comparable figure for rural homes was 80% (Bolaria et al., 1991).

Again, these figures conceal substantial variation within rural areas. The same 1981 Canadian Census found that 21% of native people in rural areas lived in housing that required major repairs, 46% lived in housing that lacked central heating, and 27% did not have a bathroom.

10.11.3 LIFE IN RURAL SOCIETY

10.11.3.1 Social Life

There has been much discussion in the rural sociology literature about the so-called duality between rural and urban life (see Labao, 1996). The popular image of rural life is that of pastoral tranquillity. Short (1991) summarized the features of this image as follows:

> a less-hurried lifestyle where people follow the seasons rather than the stock market, where they have more time for one another and exist in more organic community where people have a place and an authentic role. The countryside has become the refuge from modernity. (p. 34)

This bucolic image of rural life is often contrasted with the murder and mayhem that is supposed to typify urban America. Srole (1972) traced this distinction back to the Old Testament with the city life of Sodom and Gomorrah being portrayed as an example of evil. Herzlich (1974) in her study of contemporary social representations (popular beliefs) of health and illness found that lay people tend to characterize urban life as the major source of ill health.

To investigate this image further, Melton and Hargrove (1987) asked a sample of American university students to write a short paragraph describing an urban scene, an urban person, a rural scene, and a rural person. Content analysis of their replies revealed that their descriptions of a rural person were generally positive and frequently referred to warmth, friendliness, and simplicity of lifestyle. Less frequently mentioned were slowness of pace, family centeredness, and conformity. These findings led Melton and Hargrove to concur with the following comment of Coward and Jackson (1983):

> there is little empirical support for this folklore [of a family immersed in a strong and pervasive social support network of kin, friends, and church. To the extent that kin, friends, and neighbors can serve to support families during periods of stress, the conclusion to be drawn from the research literature is that rural families are not particularly advantaged. (p. 196)

In North America, the decline of the family farm, the rise of agribusiness, and the integration of rural areas into urban society has contributed to the undermining of the traditional communitarian culture (Cordes, 1990). Murray and Kupinsky (1982) found that now fewer rural Americans participate in community activities, such as granges, church functions, and civic groups, than in the past. They suggest that changes in communication patterns and the geographic dispersal of extended families have led to strains on traditional sources of social support.

One seemingly enduring attribute of rural societies is a high value placed on self-sufficiency and self-reliance (Dengerink & Cross, 1982). A consequence of this can be a reduced demand for health care. Bigbee (1990) found less evidence of health problems in a rural community and suggested that this was due to under-reporting by residents who wished to maintain a sense of self-reliance.

10.11.3.2 Women and Family Life

Women have traditionally played a central role in rural society. Although there are substantial variations, there is evidence that rural women are more likely than urban women to be married, have more children, and live in larger families (Mansfield, Preston, & Crawford, 1988). The greater frequency of larger extended families is due to a variety of social and economic factors. Many young rural couples initially establish their families in their original family homes and some single-parent mothers return to their home of origin. This living arrangement can be a source of social support at times of difficulty, but can also be a source of conflict (Bushy, 1990).

There is evidence that rural women more frequently adopt traditional gender roles in terms of family care and household responsibilities (see Bushy, 1993). In addition, those who are farmers' wives have the dual responsibilities of assisting with a wide range of farming tasks. Despite these increased demands, rural women have fewer sources of support. For example, Bushy (1993), in a study of the health needs of rural women found that they were less likely than urban women to identify a large network of social support that could be of benefit during times of difficulty. The rate of female participation in the labor force also tends to be lower in rural areas (Swanson & Butler, 1987).

10.11.3.3 Working Life

Rural occupations are not confined to farming and related activities. Indeed, increasingly there is a widespread range of occupations in rural areas, although admittedly there are fewer professionals and white-collar workers than in urban centres. Despite this change, an image persists that working life in rural areas is gentle and easygoing. Behind this image there exist quiet desperation and much hardship.

Farming would seem to be an inherently stressful occupation (Keating, 1987). There are a variety of reasons for this. An important factor is the very uncertainty of the job in terms of the market and the weather. This lack of control would be expected to lead to feelings of unease (see Belyea & Labao, 1990). Another factor is the multiple responsibilities a farmer has in terms of work on and off the farm and with respect to the family. Finally, there is the changing rural scene which threatens the whole farming way of life.

Keating (1987) considered the level of stress among farming couples on Canadian grain farms. She found that personal resources, which was a measure of perceived mastery, was the best predictor of stress. In a comparable study of Ohio farmers, Belyea and Labao (1990) found that relatively young farmers with a large number of children and a large proportion of acreage in grain crops were most vulnerable economically as measured by the debt-to-asset ratio and the net family income. Further, those who were most vulnerable reported more feelings of economic hardship and stress. In turn, perceived hardship and stress were related to depression.

One factor that deters many rural residents from leaving despite difficulties is the sense of family and community history. Many of the rural residents will come from families who have lived in a specific community for generations. Their whole lives are defined by the character of their work and their community and it is difficult for them to imagine another way of living. Often, they will have been working on the farm or in the family business before they left school and expect to do so until retirement as their parents did before them. Their work is not just a job but a way of life. It is integrated into their daily lives. It is for this reason that a threat to their jobs becomes a threat to their whole way of life (Schroeder, Fliegel, & VanEs, 1985).

Another feature of rural employment, at least that which is based upon the exploitation of natural resources, is its seasonal nature. This has given rise to large numbers of migrant workers in both the USA and Canada. These workers are largely drawn from ethnic and racial minorities (Burawoy, 1976). For example, most farmworkers in British Columbia are East Indians, Chinese, native Indians, Francophones, and migrant youth (Sharma, 1983). These workers often work in atrocious conditions and are clearly exploited by their employers. A 1973 Canadian Task Force study of migrant farmworkers found evidence of "child labor, sick, pregnant, and otherwise unfit adults working with only the head of the family being paid" (Sanderson, 1974, p. 405). Hopefully, conditions have improved since then.

Finally, a large proportion of rural workers are without regular employment. These individuals would be expected to experience the various deleterious effects of loss of employment (see Jahoda, 1972). Admittedly there is some evidence that this impact may be ameliorated in rural areas owing to the presence of social support (Harding & Sewel, 1992; Murray & Dolomount, 1995). However, the apparent decline of such support leaves these individuals at increased risk of psychological distress.

10.11.3.4 Change in Rural Society

Rural communities have been adversely affected by both economic and natural problems over the past decade or so. One major change has been the rapid decline of the medium-sized family farm. There are various economic and political explanations for this but the result has been that many small farmers have had their farms seized by the banks and large lending institutions. According to the US Department of Agriculture (1985) about one-third of American farmers with sales over $100 000 are at risk of losing their operations. Between 1981 and 1986 650 000 farms in the USA were foreclosed. About half a million rural jobs were lost between 1981 and 1983 (Human & Wasem, 1991). According to the US Office of Technology over a million farmers will leave the land by the year 2000 (Rosemann & Delworth, 1990).

It would seem that the farmers who have been most vulnerable to this economic threat have been those with medium-sized farms and those who are younger and well educated. These are the farmers who took advantage of loans in the 1970s to expand their operations but found it difficult to maintain their payments when the market changed in the 1980s.

A survey of Canadian prairie provinces by the National Farmers' Union (1989) concluded that "over the past ten years ... thousands of farm families have been dispossessed" (p. 4). The report further adds that:

> the *corporate ownership* of land means that an increasing amount of the value of farm production leaves the community in the form of payment to corporate owners. This means less spending power will remain with rural communities which, in turn, will influence their future viability. (p. 8)

In view of these economic difficulties it is perhaps not surprising that farmers generally view financial problems as the principal source of stress (Olson & Schellenberg, 1986).

Throughout the twentieth century there has been a steady decline in the number of farms and what has been described as the rise of a dualistic system of agriculture (Albrecht & Murdock, 1988). On the one hand are the small family-run farms which are increasingly economically nonviable and where the farmer often seeks part-time employment off the farm in order to survive. On the other hand are the large industrial-type farms owned by corporations. These large farms are run as capitalist enterprises with more mechanized methods of cultivation and lesser requirement for human labor. In Canada the total number of farms has dropped from a peak of 732 832 in 1941 to 280 043 in 1990 (Statistics Canada, 1996). Currently 8% of Canadian farms account for 43% of Canadian farmland.

The crisis hitting small farmers and fishermen has an impact beyond the immediate worker and threatens the survival of the worker's family and of the local community. Hoyt, O'Donnell, and Mack (1995) note that the farm crisis had two major impacts on the immediate community. First, there is the direct economic impact which removes money from the local economy, threatens small businesses, reduces employment opportunites, and accelerates the exodus of young people in search of alternative employment. Davidson (1989) compared the effects of this rural upheaval to the more visible degradation of urban ghettos.

A second, more indirect, consequence of the farm crisis is the decline of support resources. The decline in the population base is followed by the loss or consolidation of formal support services such as hospitals, community organizations, and churches. Thus, the remaining elderly population has more difficulty gaining access to these services. Human and Wasem (1991) describe this process as a vicious cycle: "when times are bad, the need for health services—particularly mental health services—is greater, but because times are bad, the ability to purchase services is lower, as is the ability of the communities to provide services" (p. 234).

It is not just the natural resource-based industries that have been restructured but also the small rural-based service and manufacturing industries (Fitchen, 1991). Brooks, Stucker, and Bailey (1986) argue that as the small farms disappear so, too, do the small shops, to be relaced by fewer and larger firms with a total of fewer employees. Like the large farms, these large firms are often owned by nonrural corporations whose primary interest is extracting profit from their investment and not with the economic maintenance of rural society.

These changes have had a widespread impact on the character of traditional rural society. Naples (1994) found that rural residents frequently report increasing dissatisfaction with community life and fewer feelings of community cohesiveness. These feelings would be expected to contribute to feelings of social isolation and hopelessness and subsequent psychological distress.

10.11.4 HEALTH ISSUES

10.11.4.1 Stress in Rural Society

As already mentioned, farm life has a variety of associated stressors. Rosmann and Delworth (1990) distinguished between social stressors and more work-related stressors. The former are the range of social pressures from commercial and government agents that the farmer has to deal with on a regular basis. The work-related pressures derive directly from the uncertainty of farm life. These concern the variability of prices and production costs, the variability of the weather, and the variability of government and other regulations.

The psychological consequences of such economic uncertainty affects the whole farming family and not just the farmer. Rosenblatt and Keller (1983) studied a small sample of Minnesota farming couples. They found that economic vulnerability was related to perceptions of economic distress and also to evidence of interpersonal conflict.

Some research has explored gender differences in the character of psychological distress experienced by farmworkers. The results of this work would suggest that women farmworkers experience higher levels of stress (Walker & Walker, 1987). This may be because they are the ones who shoulder more of the problems that affect rural families, including those experienced by male farmworkers.

10.11.4.2 Impact of Crisis

Several studies have examined the impact of farm crises and foreclosure on the health of farming families. This research has shown that economic hardship is followed by a rise in psychological distress (e.g., Armstrong & Schulman, 1990; Belyea & Labao, 1990).

One important longitudinal study was conducted in Nebraska (see Ortega, Johnson, Beeson, & Craft, 1994). A sample of farm families was followed throughout the 1980s and it was found that the level of self-reported depression significantly increased following the economic downturn but improved when there was a recovery in the economy. This would suggest that the negative mental health effects were short term.

Hoyt et al. (1995) found greater evidence of psychological distress among residents living in small rural communities than among those living on farms or in larger towns. This distress was particularly pronounced among those with low levels of social support. This would suggest that the effect of place on distress may be due to the erosion of communal identity and decreasing sense of collective concern.

10.11.4.3 Social Cohesion

One of the most important factors in understanding mental health, and indeed physical health, in rural society is the degree of social cohesion. Traditionally this has been considered a defining aspect of rural societies. Brody (1973) argues that mutual aid is virtually a defining characteristic of small farming communities. He quotes the Russian anarchist Kropotkin (1939):

> the mutual-aid tendency in man has so remote an origin, and is so deeply interwoven with all the past evolution of the human race, that it has been maintained by mankind up to the present time, notwithstanding all vicissitudes of history. (p. 180)

It is the undermining of this social solidarity that is possibly the greatest threat to the mental health of rural society.

Social interdependence leads to substantial social interaction as a requirement both for work and also for entertainment. This was so in traditional rural communities. It was based upon the large degree of social equality that existed within these communities. This meant that people shared in times of need and celebrated together in times of plenty. The breakdown of the family farm and small rural businesses has meant the rise of a social hierarchy in rural communities that undermines this basis of social solidarity. A classic example of this change is the case of Roseto, the small Italian community in Pennsylvania. The reason for initial research interest in this community was the low death rates despite the presence of the standard behavioral risk factors. It was noticed that Roseto was a very egalitarian and close-knit community. Bruhn and Wolf (1979) summarized the impact as follows:

> The sense of common purpose and the camaraderie precluded ostentation or embarrassment to the less affluent, and the concern for neighbors ensured that no one was ever abandoned. This pattern of remarkable cohesion … provided security and cohesion against any catastrophe [and] was associated with the striking absence of myocardial infarction and sudden death. (p. 136)

Sadly, the death rate in the community rose as the degree of social integration declined (Egolf, Lasker, Wolf, & Potvin, 1992; Wilkinson, 1996).

The decline of social solidarity is, perhaps, one of the most negative consequences of the social changes in rural America. It has also been argued that the change in the social composition of rural America away from an interdependent society towards one that in many ways is more comparable to mainstream America has introduced more individualistic attitudes and values to the detriment of rural mental health (Zahner, Jacobs, Freeman, & Trainor, 1993).

10.11.4.4 Mental Illness

The results of studies comparing mental illness in urban and rural communities are confusing. One problem with early work was that it was limited to rates of admission to mental institutions and did not control for differences in the regional provision of such facilities (see Cochrane, 1983, for commentary). For example, Srole (1972) refers to the early work which suggested that black slaves in the southern states had lower rates of admission to mental hospital than freed slaves in the north. These studies neglected to point out that there were no mental hospital beds for blacks in the south.

Another problem is that early studies neglected to consider sociodemographic differences between urban and rural areas. More recent work that attempted to control for such differences, for example, Neff (1983) and Scheidt (1985), concluded that there were few differences between urban and rural mental health. Zahner et al. (1993) found that urban–rural differences in emotional and behavioral problems among children largely disappeared after statistically adjusting for mobility and cultural and economic differences.

The Epidemiological Catchment Area study (ECA; Robins & Regier, 1991) and National Co-morbidity Study (NCS; Kessler et al., 1994) found only minor differences in prevalence rates among some diagnostic categories of mental illness among individuals living in rural and urban areas. Prevalence data for rural populations in general and particularly for rural African-American and poor populations are limited and have provoked some controversy. The NCS failed to find differences based on race or rurality for the disorders studied. It has been suggested that conclusions about psychiatric disorders in these populations are inaccurate and related to the poor representation of the samples of rural populations included in them and in the ECA study, as well as the insensitivity of measures used.

There is some evidence for differences in the prevalence of certain mental illness among the rural elderly. General estimates for the prevalence of mental illness among the rural elderly range from 23–25% (Scheidt & Windle, 1982; Rosen et al., 1981), compared with 15–25% in the general elderly population (Weber, 1990). The suicide rate is also high among rural elderly; however, the prevalence of depression is the same if not lower than among urban elderly (US Congress, Office of Technology Assessment, 1990). The suicide rate among the general elderly population is 19.8 per 100 000, compared with 12.6 for the general population. Little is known about the dynamics of suicide among rural elderly or the consistent rise in suicide among African-American elderly.

Many nursing and adult homes exist in rural areas. Between 60 and 90% of the residents of these institutions have been diagnosed as having a mental illness. Many of these homes are not licensed, are poorly staffed, maintain inadequate records, and do not provide adequate programs for their residents.

Taken as a whole, these epidemiological studies do not provide compelling evidence for major differences in the incidence and prevalence of mental disorders between urban and rural populations. Admittedly there are specific social and cultural issues peculiar to rural communities that must be taken into consideration in the design of intervention programs. A key question is how rural persons conceptualize their symptoms, and whether or not they access services. How rural persons interpret physical illness symptoms, seek confirmation of the presence and meaning of symptoms, and make decisions about where to go for help is not well understood. Even less well understood is the experience and meaning of symptoms of psychiatric disorders among rural minority and poor persons. To understand help-seeking for mental health problems among these individuals, issues surrounding availability, accessibility, and acceptability of care must be examined (Blank, Fox, Hargrove, & Turner, 1995).

10.11.5 MENTAL HEALTH SERVICES

10.11.5.1 Healthcare Utilization

Despite claims that the disparity between rural and urban health service availability is narrowing (Freeman et al., 1987) and that differences in access to healthcare in rural and urban areas are less, patterns of utilization still reflect tremendous inequity (Gesler & Ricketts, 1992). Although rural residents are more likely to suffer from chronic health problems and limited functional status, they are more likely to be without a regular source of health care (Robert Wood Johnson Foundation, 1987). They are also less likely to see a physician, and those without Medicaid or private insurance are the least likely to access health services.

Inequities in health service utilization are further exacerbated among cultural and ethnic minorities in rural areas. In the USA, rural African-Americans are twice as likely as Whites to be without a regular source of healthcare (Davis et al., 1987). Further, service access and utilization for poor rural African-Americans in the south continue to lag even further behind service access in other geographical regions. Some have suggested that lack of insurance and social isolation among these rural poor minorities contribute substantially to the absence of a regular source of care (Lewin-Epstein, 1991). Given these circumstances, it is even more important to design interventions that take into account the barriers to care, community residents' beliefs about illness, and help-seeking patterns.

10.11.5.2 Specialty Mental Health Services

In the USA most inpatient mental health care (43%) is provided by general hospitals, with only 35% provided by state and county hospitals (Narrow et al., 1993). Compared with several decades ago, the total number of mental health facilities has increased but the number of inpatient beds has decreased. Thus the average facility has downsized while the overall number of facilities has increased.

Over the same period of time, downsizing and closing of state and county hospitals has been a well-documented change in the specialty mental health system. Inpatient, outpatient, and partial care services have declined in the public sector while the availability of private services has increased dramatically. There has also been a parallel increase in the number of private psychiatric hospitals and psychiatric units in general hospitals. The number of staff available to these facilities has also increased dramatically (Redick et al., 1992). Unfortunately, these changes have not substantially affected the mental health care of rural persons, since these facilities are most frequently located in urban areas, or in close proximity to urban hubs.

A major concern for healthcare planners is the distribution of mental health services across different regions of North America and specifically within rural areas. It has been estimated that 50% of rural persons with mental disorders do not seek any help (Lee & Bowles, 1974). No doubt, numerous factors contribute to this reluctance to seek assistance, but the fact is that there are 1682 counties in the USA without any psychiatrists, psychologists, or social workers. All of these counties are rural. Inpatient, outpatient, and partial hospitalization programs are much less likely to be located in rural than urban areas. For example, only 13% of nonmetropolitan counties have psychiatric inpatient units (Wagenfeld et al., 1988). Many poor rural residents do not qualify for Medicaid because they own land (the "landed poor") and they are more likely to be uninsured. It is unlikely that a market for private psychiatric services will develop in rural communities in the near future. The evidence that rural persons are more likely to be admitted to public mental hospitals probably reflects the absence of alternative mental health services in rural areas.

In the USA, specialty mental health service organizations have expanded rapidly since 1980. Documented increases in service delivery would indicate that these agencies are providing important services to individuals with mental disorders. Unfortunately, rural areas most frequently do not have access to adequate mental health services or service providers. Therefore, rural individuals with mental disorders are most likely to go without appropriate care. If they do access care, they do so later in the course of the illness, which results in increased cost and length of treatment. Also, because of the shortage of specialty mental health providers in rural areas, they are more likely to access care provided by poorly trained or more entry-level professionals. Further, rural residents are more likely to access care at a great distance from their home community which can result in loss of work and a severing of community ties (National Association for Rural Mental Health, 1993).

10.11.5.3 Public Policy

Rural mental health is a field that has been historically neglected both by researchers and policy makers, rendering the development of coordinated and empirically based services even more difficult (Fox, Blank, Kans, & Hargrove, 1994). The economic realities of providing public rural mental health services also serves as an obstacle to implementing theory-driven service delivery systems. For example, despite frequent claims of the cost effectiveness of model prevention and treatment programs, there are declining allocations for new research in rural mental health. Further, most of the funding of public services for community-dwelling mentally ill persons currently flows from entitlement programs such as Medicaid which are administered at the state level. This shift in the locus of authority from a federal to state operation of mental health services reduces the opportunities both for national initiatives as well as for more carefully tailored local efforts in rural systems development (Hargrove & Melton, 1987). What is needed is an empirical base and the development of sound theoretical models of health and illness in rural areas, which can then develop into the deliberate development and testing of interventions designed to fit within a rural social ecology.

Continuing healthcare reform efforts have emphasized a lack of health and mental health services for rural poor and minority persons. This lack of services can be conceptualized as including problems of availability, accessibility, acceptability, and accountability (Blank et al., 1995). Service providers and health policy analysts are beginning to address how mental healthcare needs are identified among rural poor and minority persons, and how healthcare providers in existing rural healthcare sites address these needs. The impact of inadequate diagnosis and ineffective treatment of mental disorders on utilization patterns and costs of other types of healthcare for rural minority and poor persons needs further research.

Most rural persons with mental illness receive services in the general medical sector of the healthcare system and through other nonspecialty providers including social support networks rather than through formal mental health specialist services. This commonly existing loose and fragmented collection of mental health services has been described as the "*de facto* mental health service system." The *de facto* system includes specialty mental health services, general medical services including primary care and nursing homes, other human service providers such as ministers and counselors who are not principally mental health specialists, and voluntary services including self-help groups, families, and friends.

Healthcare reform efforts have been directed toward extending access to health services for previously uninsured and undeserved persons while simultaneously controlling overall costs for care. Rowland and Lyons (1989) demonstrated that a large number of rural, minority, and poor persons are included in the underserved and uninsured population. In order for public policy to change to improve efficacy and cost containment a better understanding of mental health needs, as well as current modes of service access and utilization, are imperative.

10.11.5.4 Empirical and Theoretical Basis for Planning Services

Rural mental health service delivery is frequently described as costly, inefficient, and/or ineffective (Aviram, 1990), although there is a paucity of data regarding cost, effectiveness, or outcomes. There is no doubt that the "boundarylessness" (Bachrach, 1983) of mental health service delivery in rural areas care not only requires management of psychiatric symptoms, but also frequently requires attention to physical health and to long-term needs including housing, transportation, and safety (Bigelow, McFarland, & Olson, 1991).

Presently, public care for rural mentally ill consumers is unable to provide consistent and reliable service focused on the complex requirements of effective community care. The lack of data regarding the distribution of providers and the characteristics, needs, and outcomes of rural dwellers with regard to mental health services makes planning for policy dependent on idiosyncratic beliefs of powerful decision makers. In rural areas, the increasing dependency upon local authority makes public mental health policy even more dependent upon a few individuals who may or may not be well informed (Blank et al., 1995). As a result, rural service delivery systems are more vulnerable to changing political forces (Fox et al., 1994).

In order to provide better service delivery systems in rural areas and evaluate them, more consistent theoretical perspectives and more uniform standards of care are needed. Social isolation and greater geographic distances in rural environments contribute to the variability and inconsistency of service systems. Because of the lack of attention to rural areas by the scientific community as well as policy makers, there is relatively little to guide practice in this area. The emerging literature of rural mental health care is influenced by assumptions about the nature of human service delivery in rural contexts which are frequently contradictory (Murray & Keller, 1991).

The importance of understanding the influence of informal sources of care in rural areas cannot be overemphasized. Fox et al. (1994) have proposed a model for linkage of formal and informal care-givers for mental health service provision among seriously mentally ill consumers in rural areas. This model is based on the Balance Theory of Coordination developed by Litwak and Meyer (1966) and is applied to service delivery systems in rural areas. Critical features of this model are the formation of linkages between formal and informal care providers which are characterized by neither too much social distance nor overly enmeshed relationships, and the fit of task to structure. By testing the applicability of such a model to rural case management services, it may be possible to ascertain the factors necessary to provide optimal care for seriously mentally ill persons who live in rural areas.

10.11.5.5 Healthcare Reform

Rural persons are disadvantaged by services that are almost exclusively conceptualized as facility-based. Given the characteristics of rural populations and rural areas, the mental health of rural communities may be better served by outreach treatment modalities and the development of in-home services. That is, services cannot and should not be limited to the confines of a particular place, and need to utilize existing community structures better, and expand upon informal systems of care. Ironically, more community-responsive models of service such as home-based services and smaller population-specific clinics are emerging in urban areas (clinics specifically for women, homosexuals, minorities, etc.). However, perhaps owing to more severe financial constraints in rural areas, parallel specialty services are not being developed in rural areas. In fact, Mermelstein and Sundet (1988) found that rural community mental health centers were less likely to create new and innovative services during the rural crisis in the 1980s.

Another important feature in planning rural mental health service delivery systems is that frequently the entry points into the mental health system are existing rural organizations. An increasingly frequent example of this is the use of churches in collaboration with mental health professionals in providing for their congregations (Pargament et al., 1991). Churches provide naturally occurring, convenient gathering places for the surrounding community, and through true collaboration and mutual help, congregations can perceive ownership of innovative programs which can reduce suspiciousness, fear, and stigma, and increase participation.

Rural mental health services can also learn from work in developing countries. Susser, Schanzer, Varma, and Gittelman (1996) note that in these countries the family often plays a central role in the care of patients with mental illness. It has been known that despite the lack of availability of hospital care, mental patients in developing countries have been found to suffer from less impairment and disability than those in developed countries (Jablensky et al., 1992). Susser et al. (1996) suggest that this apparent advantage is due to the involvement of the family who are encouraged to understand the mental health problem and to develop ways of coping with the patient's needs and demands. With such a shortage of hospital care in rural areas, the potential involvement of the family in patient care offers much promise.

10.11.5.6 Primary Care

In both urban and rural settings most individuals with mental health problems only receive care in the general medical care system. The National Co-morbidity Survey reported that of the 42% of people with psychiatric problems who received professional help, only 26% obtained help from mental health specialists (Rich, 1994). An overwhelming majority of people with psychiatric problems are being managed through the general health sector. Furthermore, persons with mental health problems make about twice as many visits to primary care providers as do primary care patients without mental health problems (Cleary, 1987). Approximately 22% of persons utilizing primary care suffer from a mental health problem (Narrow et al., 1993). Despite considerable concern about the co-occurrence of mental illnesses or emotional dysfunction and physical illnesses in primary care populations, there has been little consideration of the impact of the separate health and mental healthcare services for individuals accessing primary care (Coulehan et al., 1990).

Lack of integration of general medical and mental health services is an often cited barrier to more effective referral from primary care to mental health specialists (Kamerow, Pincus, & MacDonald, 1986). Widespread concern has been expressed about primary care providers' lack of recognition, and treatment or referral of patients with mental disorders (Jones et al., 1987). Approximately one-third of primary care patients with psychiatric disorders remain unidentified and untreated and almost 40% of persons receiving treatment for mental disorders receive all their treatment through general medical services (Jones et al.).

According to Morlock (1989), primary care physicians record a primary or secondary psychiatric diagnosis for only about 4.4% of patient visits despite a significantly higher prevalence of mental disorders in the primary care patient population. In a study of 1000 primary care patients, Spitzer et al. (1994) found that 26% of those primary care patients met full diagnostic criteria for mental disorder with an additional 13% meeting conditions for subthreshold diagnosis. However, approximately half of those patients (48%) had not been recognized by their physicians as having a diagnosis. It is alarming that despite healthcare reform efforts, no reports have been published documenting the capacity of the general medical sector to identify, treat, or refer minority or poor persons for mental disorders.

Limited attention to mental health training of primary care providers, negative provider and patient attitudes about mental illness, and practice constraints have been highlighted as barriers to more effective mental health treatment and referral in primary care practice. Numerous investigators and clinicians have proposed that psychiatric diagnostic instruments and psychiatric epidemiological research instruments (the American Psychiatric Association's *Diagnostic and statistical manual of mental disorders* Diagnostic Interview Schedule) are not appropriate for use in primary care sites and there is a need for development of sensitive and specific mental health screening instruments for general healthcare sites (Spitzer et al., 1993). Other investigators highlight resistance of care providers and patients in primary care sites to refer to or comply with recommended treatment of mental health problems identified through screening. Little is known about which mental health approaches may be more or less culturally acceptable and thus promote compliance in primary care patients needing mental health services.

There are few studies on the management of psychiatric conditions in primary care practices, factors that influence care providers' decisions about mental health interventions, and how and why these decisions differ for patients of various cultural and ethnic backgrounds and geographic regions. Service systems have historically been resistant to providing integrated health and mental health services. The stigma associated with mental illness and treatment of mental disorders has also been blamed for underdevelopment of mental health screening and intervention in primary healthcare settings. Additionally, little is known about the delivery of standard effective mental health treatments in different types of primary care sites or how these treatments are related to health outcomes. The absence of research on outcomes of mental health treatment in primary care contributes to general healthcare providers' reluctance to incorporate mental health screening and treatment more rigorously in primary care encounters (Davenport, Goldberg, & Millar, 1987).

In view of the greater availability of primary care physicians than specialists in rural areas (Dor & Holahan, 1990), they have potentially an even more important role than their urban counterparts to play in the care of patients with mental health problems. Greater cooperation between psychologists and primary care physicians will ensure a more accessible service for people with mental health problems.

10.11.5.7 Telemedicine

Interest in telemedicine and communications technologies to improve health and mental healthcare in rural areas has increased dramatically since 1985. The introduction of telemedicine offers great potential to enable the patient/client to overcome the physical/geographical barriers to accessing specialist healthcare and to enable the health professional to continue his or her education (Preston, Brown, & Hartley, 1992).

Zelman (1995) has described a typology for innovative technologies using POTS (Plain Old Telephone Service) and PANS (Pretty Amazing New Stuff). The PANS consist of two-way interactive televideo systems which typically utilize high-speed dedicated access (such as T1, ISDN, or switched-56 lines), and are dependent upon sophisticated (and expensive) equipment. The US Office of Rural Health Policy has funded over a dozen demonstration projects for telemedicine since 1992, many of which include mental health treatment as part of their goal. To date, there have been no mental health services research studies concerning the cost, effectiveness, or attitudes of consumers and providers toward innovative technologies. One study did examine the reliability of assessments using this technology for persons with obsessive-compulsive disorder (OCD; Baer et al., 1995). These researchers concluded that standardized assessments for OCD could be administered reliably using interactive televideo systems.

Mental health service delivery seems particularly well suited for dissemination using two-way interactive televideo systems. Unlike many of the medical specialties, mental health services need to be delivered in real time. Further, most assessments and treatments rely on the type of face-to-face interaction that is supported by these systems. Given that most of the diagnostic information typically relied upon is visual and verbal, these systems hold great promise for extending our reach into underserved rural areas.

There are several public policy pitfalls awaiting telemedicine. Perhaps most obvious is the question of whether services delivered through these media will be eligible for third party reimbursement. It seems clear that health services research will need to be conducted to determine the quality and effectiveness of specific procedures such as diagnosis, other assessment, and treatments delivered at a distance through new technologies. Another difficult policy issue concerns licensure of health professionals. Typically, licenses to practice are issued by states, yet these televideo transmissions may frequently cross state lines. Recognition of state licensure in other jurisdictions presents a number of legal and administrative problems. There are also concerns about practice liability and confidentiality of video transmissions.

10.11.6 PSYCHOLOGY AND RURAL MENTAL HEALTH

10.11.6.1 Historical Background

The discipline and profession of psychology has played a variable role in the history of mental health services in the rural areas of the North America. As a part of the community mental health centers movement in the USA in the early 1960s, psychology was by mandate integrally involved in the planning, development, and implementation of services in the catchment areas. But a further, more refined focus on rural services within that movement was left to various multidisciplinary groups and psychology as a profession took little interest in rural communities or the rural context of the practice of psychology. As Keller and Prutsman (1982) pointed out:

> Psychology has traditionally been an urban profession. Most psychologists are trained in, and subsequently remain in, large metropolitan areas or atypical university communities. Consequently, psychologists have largely failed to consider the special mental health needs of more than one-quarter of the nation's population. (p. 190)

Professional groups such as the Rural Social Work Caucus and the National Association for Rural Mental Health were active in the 1960s and 1970s, typically as a part of the community mental health center movement. The American Psychological Association (APA) demonstrated interest in the late 1980s with the appointment of a Rural Task Force to address the issues of psychologists practicing in rural communities. While psychologists had been involved in the rural mental health movement for a number of years, this was the first time that organized psychology devoted resources and gave recognition to an important issue for many of its members.

As one of the four core professions designated by the National Institutes of Mental Health (NIMH), psychologists were typically well represented in the early mental health centers and continued until the system experienced major structural changes in most states in the early 1980s. Prior to the development of these centers, mental health services were typically not available to rural residents, who had to rely on whatever types of assistance may have been available in communities from other agencies or professionals or travel to the nearest urban centers.

10.11.6.2 Roles of Psychologists in Rural Communities

The roles of psychologists in rural mental health programs have included provider of clinical services, program and agency administrator, program evaluator, clinical supervisor, and program and service planner. From the number and diversity of roles that psychologists play in mental health agencies in rural settings, it is clear that psychologists must be able to perform a wide range in a number of different aspects of the service delivery system. Berry and Davis (1978) pointed out that "rural mental health workers must be a practitioner–generalist, able to handle a variety of problems because the rural community cannot afford a large selection of specialists or referral resources" (p. 677).

Keller and Prutsman (1982) argue that rural psychologists "must be flexible generalists who comfortably serve a wide range of human needs." They added:

> In a time when much of psychology is becoming highly specialized, the rural psychologists will need to fulfill a broad range of functions partially because there exists a lack of specialists and persons from other disciplines who are available to meet unusual needs. (p. 1992)

Hargrove and Howe (1981) identified the generalist style as an objective of clinical psychology training for rural service delivery. Hargrove (1983) further refined the concept of generalist for psychologists and identified the various components of that style of practice in the light of professional identity.

When the role of generalist is interpreted within the clinical domain of practice, the psychologist must be able to call upon a broad range of assessment and intervention strategies. A doctoral-level psychologist may be required to conduct assessments for school systems, courts, law enforcement agencies, social service agencies, as well as for the mental health system. In addition to these uniquely psychological functions, the psychologist is also likely to be expected to do case management and other programmatic and clinical functions as needed.

It has been noted that doctoral-level psychologists are likely to rise rapidly in the administrative structure of programs and agencies. It was not infrequent in the early days of community mental health centers for doctoral-level psychologists to become directors of centers with responsibilities for recruitment, personnel management, systems development, fiscal operations, management information systems, contract negotiation, and board administration in addition to the clinical and programmatic operations.

The generalist concept, then, reached far beyond serving a broad range of clients with a wide range of problems. It included involvement in the internal administrative operations of the agency, clinical and program management, and community linkages. Few psychologists were trained to assume such a broad range of responsibilities at such significant levels in the organizations.

10.11.6.3 Training Psychologists for Rural Practice

Psychology training has largely taken place in urban settings. The greatest contextual influence on training, then, has been an urban one. This created problems for psychologists who wished to work in rural settings because they did not have appropriate conceptual frameworks or experiences to successfully work there.

As this urban domination of training programs and models became known, the NIMH in the 1980s began its focus on prioritized rural populations, among others, for special attention in the funding of training programs in psychology, psychiatry, social work, and nursing for rural practice. Several university and free-standing programs for psychological training for rural practice developed in conjunction with this initiative.

An important recent initiative has been taken by the APA to develop a common training program for psychologists, social workers, and nurses working in rural mental health. The development of the common curriculum (APA Office of Rural Health, 1994) not only signaled the increasing awareness of the need for training materials for those working in this area but of the necessity for collaboration in training. Unlike urban practitioners, rural health workers in general find that there is less value in maintaining strict lines of professional demarcation and more need for interprofessional collaboration to overcome the inadequacy of resources and the breadth of mental health problems.

This common curriculum emphasized the need for rural psychologists to consider: (i) the social, economic, political, and religious influences affecting rural communities; (ii) the importance of ethnic and cultural influences in rural communities and the importance of oral tradition; and (iii) the uniqueness of each region and community. This report also refers to the reluctance of rural residents to seek professional support, often turning instead to the clergy and family members in times of distress. Further, the report reminds the psychologists to be reflexive in their practice and to recognize the impact of their own culture on the delivery of care and on their sensitivity to the client.

Several other issues are important for professional psychology's rural training agenda. First is training for personal and family adaptation to the rural environment. Frequently it is difficult to assist in the personal preparation of professional persons in academic settings because of the limited scope of traditional activities. Second is whether doctoral-level psychologists are necessary for rural service delivery or whether persons with masters degrees are adequate for this work. Jerrell and Herring (1983) studied psychologists in 20 rural Pennsylvania counties and found that there was little difference in job function between doctoral- and masters-level providers. Doctoral persons were slightly more likely to assume administrative roles. The entry-level credentials have been and continue to be a controversial topic for organized professional psychology.

Training for personal adaptation is a difficult task under any circumstances; much less in an academic context. Personal adaptation to the rural environment involves the individual and intimate choices that are made by persons and families, including choices of living environments, and the associations with friends. They also include the preferences and needs for certain types of resources. The unexpected characteristics and demands of the rural environment, particularly when the new recruits are not accustomed to rural lifestyles, can be devastating to individuals and families who are not able to anticipate them. Even students who have grown up in rural areas had difficulties assuming the role of a professional person while adapting to the lifestyle in the small community. Personal and family adaptation to the rural community is essential if the psychologist is to be successful. Training for this adaptation is typically lacking in psychology training programs. Further, Murray (1990) recommends that "rural recruiters [should] seek people who really understand rural life and who value its virtues and are not just intent on escaping urban irritants" (p. 18).

The second important training issue for psychologists is the level of training. Typically, in North America, the doctoral degree is the entry credential. However, many states and several Canadian provinces license persons with masters degrees. Indeed, persons with masters degrees in psychology have begun to demand parity with doctoral providers both in public agency service as well as in independent practice. Strong positions have been taken on both sides and the issue is far from settled.

The importance for the rural area is that most human resource distribution studies show that doctoral providers are not practicing in rural communities. Several investigators (e.g., Sladen & Mozdzierz, 1989) have studied various aspects of the distribution of psychologists in different parts of the country and typically found a lack of doctoral providers.

10.11.7 TYPES OF INTERVENTION

10.11.7.1 Fitting the Intervention to the Context

The focus of the practice of psychology and other mental health professions in rural communities has been on the alleviation of the suffering of individuals and families who live in relatively remote areas, removed from the traditional resources associated with mental health services. Since no consistent types of psychopathology have been reported to characterize rural people or rural environments, there are also no consistently recommended types of interventions that are especially designed for people in rural settings. While there have been attempts to identify "ruralness" as a trait of individuals who live in rural environments (Flax, Wagenfeld, Ivens, & Weiss, 1979; Melton, 1983), and consistent perceptions of the environment among rural and urban people (Melton & Hargrove, 1987), no reliable support for such constructs has been found.

It has been argued, however, that since peoples' values determine how they interpret their experience, values also influence how they define problems in mental health terms. This definition has an impact on their choice of whether and how individuals choose to seek help for problems (Wagenfeld & Wagenfeld, 1981). Mazer (1976) adds that the values and definitions influence the expectations that people have when they seek assistance for problems. The rural context is characterized by certain attitudes that mitigate against seeking assistance for psychological problems (Bachrach, 1977; Kenkel, 1986). An awareness of these characteristics and of the broader changing social context within which people live and work is essential for the effective delivery of psychological services.

The manner in which the intervention is implemented also depends on the context in which it occurs. Espousing the behavioral-ecological perspective, Jeger and Slotnick (1982) point out that:

> individual-level community mental health interventions should aim to provide learning opportunities to consumers that will increase their ability to influence their environments (i.e., promote competence). Behavioral training to facilitate coping with stress, developing social skills, and improving general problem-solving capacities are compatible with this value. (p. 11)

They document Rappaport's caution that:

> competence should not become a substitute label for psychotherapy, which maintains a hierarchical doctor–patient relation. Furthermore, experts should not offer canned behavioral packages to teach specific skills under the guise of competence training. Instead, the broader community context within which competence training takes place must be considered, since the community provides the meaning system and values of what constitutes adaptive behavior.

The person suffering from schizophrenia who has supportive family members in the immediate vicinity will be treated quite differently from the transient person who appears in town from unknown places. While the technology of the therapeutic intervention may not differ substantially between rural and urban environments, the choices and application of interventions are heavily influenced by the contexts in which they are applied. The rural psychologist must be aware of the changing family, social, and working life of rural communities and especially attuned to the impact of ongoing social crises. The person who loses his farm that has been in his family for generations and that provides a livelihood for an extended family needs to be approached differently from a person who loses his job in a context of expanding urban opportunity.

10.11.7.2 The Rural Psychologist as a Generalist

Psychologists, like most mental health workers, must function as generalists within the rural context (Hargrove, 1983, Keller & Prutsman, 1982). The demand of the generalist role counters the tendency toward specialization that characterizes the training and development of professional psychologists. Generalist practice requires the psychologist to be skillful in a broad range of psychological functions with an even broader range of potential clients.

The mental health system in the rural context is characterized by few professionally trained practitioners who are separated from one another by either long distances or natural barriers, or both (Hargrove, 1982). There are fewer potential clients in the community and they will probably reside and work at great distances from the providers. In most cases, the public delivery system, whether an outpost of a larger community mental health program or a specialized mental health unit of a health department, is responsible for the delivery of care. There are relatively few doctoral psychologists who practice privately in rural areas. Furthermore, there are relatively few practitioners in other mental health or mental health-related professions in rural areas.

The result of this scarcity is that those providers who are available in the rural community must be able to respond to the mental health needs that present themselves. Since there are no data to indicate that rural environments are exempt from the broad range of mental disorders, it is reasonable to assume that rural mental health providers will probably encounter most, if not all, of the major mental disorders at one time or another. It is not likely, on the other hand, that they will encounter them in great numbers.

This characteristic has several consequences. First, the mental health providers in rural areas must be prepared to respond to the broad range of mental disorders without a significant amount of collateral professional resources and support. Second, these providers must respond to needs without the benefit of being able to aggregate the clientele in either clinically or economically beneficial ways. Finally, providers must be at least familiar with a broad range of mental, emotional, and behavioral disabilities to be responsive to the needs of the population that they serve.

10.11.7.3 The Concept of Community and Psychological Intervention

It has been suggested that the best psychological interventions are not in the clinic, but in the community. The concept of "community" is deeply embedded in the sense of ruralness both in definitional as well as experiential terms. Virtually all definitions of rural contain references to small communities of people that are some distance from other people. The idea of ruralness is not limited to the population base of small communities, but is applicable to the experience of small communities with their limited populations and greater knowledge of each other's affairs (see Section 10.11.4.3).

Bell and Newby's (1972) use of Tonnies' concepts of *Gemeinschaft* and *Gesellschaft* provides a model for a general understanding of the rural community. *Gemeinschaft* is described in terms of the intimacy of relationships, kinship, and one's place in the social and community structure. The terms "blood," "kinship," and "friendship" are used to characterize this concept of community. *Gesellschaft*, on the other hand, represents large, depersonalized communities in which relationships might best be categorized as contractual. *Gemeinschaft* is closer to the traditional rural community in which the rural mental health worker may be located.

Heyman (1982), noting the lack of models for rural service delivery and the dominance of urban models for training and service, suggested that there was some freedom in the lack of constraints of rural practice:

> Every rural worker must deal with the absence of a "general" model by developing situation-specific models. There may be similarities in the implementation of these models, but the dissimilarities, in response to the varying situations, are likely to be critical to their actual functioning. The absence of a model provides a freedom and flexibility that complements the unique nature of each community. (p. 36)

Heyman points out the opportunities for rural service that are grounded in the *Gemeinschaft* orientation to community. They include visibility within the community and access to various dimensions of the community, including those in which power is embedded.

Psychological and social intervention are conducted within these contexts of social intimacy and knowledge. They are also carried out within the context of a high level of visibility and access. Clearly, there are both positive and negative consequences of these features. First, the results of one's work is available and known throughout the community. Clinical failures as well as successes may well be common knowledge and may be topics of discussion in both personal and professional circles. Second, the mental health worker may well have access to people and resources to bring to bear healing forces that otherwise may be unavailable. Third, common knowledge of people, families, and groups within communities may be useful in programmatic and clinical activities.

While Heyman (1982) certainly advocates intense involvement of the rural mental health worker in the community in which he or she lives and/or practices, he acknowledges that this involvement must be a cautious one. He calls upon the participant–observer model as one that would be satisfactory. Precaution must be taken to be involved in a broad, flexible way, active but not over involved and certainly not in a polarizing manner.

Psychologists and other mental health workers who work in rural settings do not have the luxury of limiting their perspectives of intervention to face-to-face contact between clients and clinicians. Some of the characteristics of rural environments preclude reliance on clinical service delivery as the only means of providing mental health care to the population. Distances, for example, prevent people from being able to meet at frequent intervals. The infrequency of the occurrence of similar types of psychopathology, sometimes the criterion for grouping consumers into group-oriented treatment programs, prevents the use of these types of intervention strategies.

As a result of these characteristics, rural mental health service providers, researchers, and administrators have utilized strategies of prevention and collaboration with other agencies and indigenous groups to enhance the quality of care available in the rural environments. D'Augelli and Vallance (1981) reported on the Community Helpers Project in which local residents were trained in basic helping skills who, in turn, trained others to provide informal help to persons in crisis. Libertoff (1980) presented a case study in which the informal system of care of a small New England community was utilized in developing youth and family services. Bergstrom (1982) summarized the research and made recommendations for collaboration with natural helpers for rural mental health service delivery.

A number of programs that utilize paraprofessionals have been described (e.g., Connors & Gabel, 1983; D'Augelli, 1982; D'Augelli & Vallance, 1981; Echterling, 1981; Heyman, 1982; Sundberg, 1986) These programs typically involve mental health professionals training paraprofessionals, collateral workers, and indigenous persons to be responsible to the mental health needs of people in rural areas and small towns.

An example of an initiative in community mental health service delivery was the Community Support Program (CSP) developed by the NIMH in the USA (Turner & TenHoor, 1978). The CSP, which has been successfully applied in rural settings, has a number of components including building social support for the client. A similar program developed by the New York State Office of Mental Health (Baker & Intagliata, 1984) also placed emphasis on developing opportunities for work and social interaction, competency skills training, social clusters, evening, and outreach programs, and so on.

Comprehensive psychological intervention requires going beyond the traditional clinical treatment approach to consider prevention and rehabilitation. In the rural context this requires that the psychologist not only be prepared for crisis intervention in the face of both personal and social crises but also to intervene to prepare and enhance people's resources to deal with and to combat such assaults on their way of life (see Trainor, Pope, & Pomeroy, 1997).

10.11.7.4 Ethical Dilemmas in Rural Practice

When the expectations of professional helpers encounter the intimacy and complexity of relationships that customarily exist in a rural community, the fabric of professional practice is strained. This is particularly true in the ethical guidelines that govern the conduct of practice of most mental health providers. The emerging literature on ethical complexities in rural mental health practice (Hargrove, 1986) has identified three problematic areas. These are dual relationships, limits of competence, and confidentiality.

Small communities have a finite number of people and potential relationships. Further, people relate to one another on various levels and in various domains of the community. For example, the physician may be an elder in the church, on the school board, and a director on the local bank's board. This places the psychologist in a special relationship to other members of the church, the minister and family, teachers and administrators of the school system, and employees and customers of the bank. These relationships give the character of the therapist–client relationship a special personal quality that can potentially be compromised if considerable caution is not taken.

Many people play many roles in rural communities. Therefore many people are involved with several, different aspects of each others' lives. The prohibition against dual relationships contained in the ethical codes of mental health professionals presents potentially difficult constraints on the professional and personal lives of rural practitioners.

The second problematic area consists of limits of competence. The mental health professional is in a difficult situation when he or she is the only practitioner in a community and does not have requisite skills to handle certain clinical problems. Many problems of course are so far outside the practitioner's competence that there is no question but that a referral must be made. The need for neuropsychological assessment or rehabilitation is clearly outside the competence of the average, clinically trained psychologist. But, as specialties develop and techniques are refined, it is frequently difficult for the general practitioner to know whether the practice is within the limits of competence.

Third, confidentiality may be problematic because of the high level of visibility and intimacy that is characteristic of rural communities. Certain professional relationships may carry the expectation of information exchange regardless of whether releases of information have been executed. The physician, for example, who refers a patient to the mental health center may expect a report from the center on the progress of the patient. The release of information, a necessary formality, may not be important to this physician because of the history of practice and involvement with people in the community. If the release is not signed and the mental health professional refuses to report to the physician, the political and community consequences could be quite problematic.

The high levels of visibility and access of mental health professionals in the rural community require certain precautions and preventive measures to avoid unnecessary ethical conflicts. Typically, frank and open discussion with clients in the clinical setting about boundaries and suggested courses of action if and when awkward situations develop is sufficient to avoid difficult situations.

10.11.8 FUTURE DIRECTIONS

Rural America is changing, and for a large number of its residents the future is not bright. There is an urgent need for the careful development of mental health services to address the many problems faced by rural people. The discipline of psychology is also changing, embracing a broader range of activities but also continuing toward increased specialization.

Three areas appear to characterize the immediate future in the profession and discipline of psychology as it might appear in rural America. First, the tendencies toward behavioral health and, in some venues, becoming a primary care provider have expanded the roles of psychologists beyond that of mental health provider. The development of the family psychology specialty and the harmonious relationships with family physicians have given psychologists new perspectives of research, training, and practice.

Second, the development of collaborative relationships with other professionals and with nonprofessionals and indigenous workers characterizes the work of psychologists in rural settings. Collaboration is a relative newcomer to the psychology lexicon, but the complexity of the contexts in which psychologists practice requires functional reciprocity. Fox et al. (1994) have provided a theoretical framework for coordinating rural mental health services based on Litwak's Balance Theory of Coordination and have provided guidelines for collaboration with at least one population of providers.

Finally, in the USA managed care is oozing into rural areas at a slower pace than it has flowed into urban and suburban. There appear to be several reasons for this. First, there are fewer providers in rural areas. Second, the market is considerably smaller. Third, provision of adequate health and mental health services is considerably more difficult in rural areas. It is not clear how managed care companies will respond to the health and mental health problems of rural environments.

10.11.9 SUMMARY

Rural society represents a large, although neglected, part of North American society. It is extremely diverse in its makeup. However, one consistent pattern is poverty and aging. Admittedly, there are certain pockets of wealth, but overall rural society is deprived with reference to urban society. The character of rural society is changing from one of full employment based on exploitation of natural resources to one of underemployment and greater employment in service and manufacturing industries. As a consequence, social and working life in rural America is undergoing dramatic change. Formerly it was a cohesive and supportive society but now there is less evidence of social cohesion.

In view of the diversity and changing character of rural society it is not surprising that it is difficult to identify some consistent patterns of health. However, it is apparent that there is a considerably lower degree of provision of health, including mental health, services. This lowered degree of availability is coupled with difficulties in accessibility. There have been some limited initiatives in community mental health services but there is a need not only for more resources but also for careful planning and evaluation of service provision.

Psychologists have had a lengthy but checkered involvement in rural mental health work. Even today, few doctoral-level psychologists practice in rural areas. There is greater need for advanced training in rural practice. Several recent initiatives have developed interdisciplinary training programs. In particular, the psychologist needs to be trained to become more involved in primary care and in collaborating with primary care physicians and other health professionals so as to provide integrated mental health services. Further, a psychologist working in a rural setting will confront many personal problems and ethical dilemmas. The rural psychologist should be aware of these issues and how best to deal with them.

Working in a rural setting offers many opportunities for a psychologist to treat a variety of mental health problems. For this reason the prychologist's orientation must be that of a generalist. In addition, rural psychologists need to consider the full spectrum of psychological services from health promotion through crisis intervention to long-term care of the chronically mentally ill. Further, intervention must be matched to the context to be effective. In this respect rural psychologists need to be aware of the sociocultural patterns of rural life and of the many changes rural society is currently experiencing.

10.11.10 REFERENCES

Albrecht, D. E., & Murdock, S. H. (1988). The structural characteristics of US agriculture: Historical patterns and precursors of producers' adaptations to crisis. In S. H. Murdock and F. L. Leistritz (Eds.), *The farm financial crisis: Socioeconomic dimensions and implications for producers and rural areas.* Boulder, CO: Westview.

American Psychological Association Office of Rural Health (1994). *Caring for the rural community: An interdisciplinary curriculum.* Washington, DC: American Psychological Association.

Armstrong, P. S., & Schulman, M. (1990). Financial strain and depression among farm operators: the role of perceived economic hardship and personal control. *Rural Sociology, 55,* 475–493.

Aviram, U. (1990). Community care of the seriously mentally ill: Continuing problems and current issues. *Community Mental Health Journal, 26,* 69–88.

Bachrach, L. L. (1977). Deinstitutionalization of mental health services in rural areas. *Hospital and Community Psychiatry, 28,* 669–672.

Bachrach, L. L. (1983). Psychiatric services in rural areas: a sociological overview. *Hospital and Community Psychiatry, 34,* 215–226.

Baer, L., Cukor, P., Jenike, M. A., Leahy, L., O'Laughlen, J., & Coyle, J. T. (1995). Pilot studies of telemedicine for patients with obsessive-compulsive disorder. *American Journal of Psychiatry, 152,* 1383–1385.

Baker, F., & Intagliata, J. (1984). Rural community support services for the chronically mentally ill. *Journal of Rural Community Psychology, 5,* 3–14.

Bealer, R. C., Willis, F. K., & Kuvlesky, W. (1965). The meaning of rurality in American society: some implications of alternative definitions. *Rural Sociology, 30,* 255–266.

Bell, C., & Newby, H. (1972). *Community studies: An introduction to the sociology of the local community.* New York: Praeger.

Belyea, M. J., & Labao, L. M. (1990). Psychosocial consequences of agricultural transformation: The farm crisis and depression. *Rural Sociology, 55,* 58–75.

Bender, L., Green, B., Hady, T., Kuehn, J., Nelson, M., Perkinson, L., & Ross, P. (1985). *The diverse social and economic structure of nonmetropolitan America* (Rural Development Research Report No. 49). Washington, DC: US Government Printing Office.

Bergstrom, D. A. (1981) Collaborating with natural helpers for delivery of rural mental health services. *Journal of Rural Community Psychology, 3*(2), 26.

Berry, B., & Davis, A. E. (1978). Community mental health ideology: A problematic model for rural areas. *American Journal of Orthopsychiatry, 48,* 673–675.

Bigbee, J. L. (1990). Stressful life events and illness occurrence in rural versus urban women. *Journal of Community Health Nursing, 7,* 105–113.

Bigelow, D. A., McFarland, B. H., & Olson, M. M. (1991). Quality of life of community mental health program clients: Validating a measure. *Community Mental Health Journal, 27*(1), 43–57.

Blank, M. B., Fox, J. C., Hargrove, D. S., & Turner, J. T. (1995). Critical issues in reforming rural mental health service delivery. *Community Mental Health Journal, 31*(6), 511–524.

Bluestone, H., & Daberkow, S. G. (1986). Employment growth in non-metro America: Trends and prospects to 1990. *Rural Development Perspective, 20,* 34–37.

Bolaria, B. S., Dickinson, H. D., & Wotherspoon, T. (1991). Rural issues and problems. In B. S. Bolaria (Ed.), *Social issues and contradictions in Canadian society* (pp. 393–416). Toronto, Canada: Harcourt Brace Jovanovich.

Brody, H. (1973). *Inishkillane: Change and decline in the West of Ireland.* London; Faber and Faber.

Bollman, R. D., & Smith, P. (1988). Integration of Canadian farm and off-farm markets and the off-farm work of farm women, men, and children. In G. S. Basran & D. A. Hay (Eds.), *The political economy of agriculture in Western Canada* (pp. 85–202). Toronto: Garamand Press.

Bosack, J., & Perlman, B. (1982). A review of the definition of rural. *Journal of Rural Community Psychology, 3,* 3–34.

Brooks, N. L., Stucker, T. A., & Bailey, J. A. (1986). Income and well-being of farmers and the farm financial crisis. *Rural Sociology, 51,* 391–405.

Bruhn, J. G., & Wolf, S. (1979). *The Roseto story: An anatomy of health.* Norman, OK: University of Oklahoma Press.

Burawoy, M. (1976). The functions and reproduction of migrant labor: comparative material from Southern Africa and the United States. *American Journal of Sociology, 81,* 1050–1087.

Bushy, A. (1993). Rural women: Lifestyle and health status. *Nursing Clinics of North America, 28,* 187–197.

Cleary P. D. (1987). Gender differences in stress related outcomes. In R. Barnett, L. Biener, & G. Baruch (Eds.), *Women and stress* (pp. 39–74). New York: Free Press.

Cochrane, R. (1983). *The social creation of mental illness.* London: Longman.

Connors, G. J., & Gabel, H. (1983). Effect of paraprofessional contacts in rural parent group recruitment. *Journal of Rural Community Psychology, 4*(1), 35–42.

Cordes, S. M. (1989). The changing rural environment and the relationship between health services and rural development. *Health Services Research, 23,* 757–784.

Cordes, S. M. (1990). Come on in, the water's just fine. *Academic Medicine, 65,* S1–S9.

Coulehan, J.L., Schulberg, H. C., Block, M. R., et al. (1990). Depressive symptomatology and medical comorbidity in a primary care clinic. *International Journal of Psychiatry in Medicine, 20*(4), 335–347.

Coward, R. T., & Jackson, R. W. (1983). Environmental Stress: the rural family. In H. I. McCubbins & C. R. Tigley (Eds.), *Stress and the family. vol. 1: Coping with normative transitions* (pp. 188–200). New York: Brunner/Mazel.

D'Augelli, A. R. (1982). Future directions for paraprofessionals in rural mental health, or how to avoid giving indigenous helpers civil service ratings. In P. A. Keller & J. D. Murray (Eds.), *Handbook of rural community mental health* (pp. 210–222). New York: Human Sciences Press.

D'Augelli, A. R., & Vallance, T. R. (1981). The helping community: Promoting mental health in rural areas through informal helping. *Journal of Rural Community Psychology, 2*(1), 3–16.

Davenport, S., Goldberg, D., & Millar, T. (1987). How psychiatric disorders are missed during medical consultations. *Lancet, 2*(8556), 439–441.

Davidson, O. G. (1989). *Broken heartland: the rise of America's rural ghetto.* New York: Free Press.

Davis, K., Lillie-Blanton, M., Lyons, B., et al. (1987). Health care for black Americans: The public sector role. *The Milbank Quarterly, 65*(1), 213–247.

DeLeon, P. H., Wakefield, M., Schultz, A. J., Williams, J., & VandenBos, G. R. (1989). Rural America: Unique opportunities for health care delivery and health services research. *American Psychologist, 44,* 1298–1306.

Dengerink, H., & Cross, H. (Eds.) (1982). *Training professionals for rural mental health.* Lincoln, NE: University of Nebraska Press.

Dor, A., & Holahan, J. (1990). Urban–rural differences in medicare physician expenditures. *Inquiry, 27,* 301–318.

Echterling, L. G. (1981). A rural program for enhancing prevention and intervention skills. *Journal of Rural Community Psychology, 2*(1), 32–38.

Egolf, B., Lasker, J., Wolf, S., & Potvin, L. (1992). The Roseto effect: A 50-year comparison of mortality rates. *American Journal of Public Health, 82,* 1089–1092.

Fitchen, J. M. (1991). *Endangered spaces, enduring places: Change, identity, and survival in rural America.* Boulder, CO: Westview.

Flax, J. W., Wagenfeld, M. O., Ivens, R. E., & Weis, R. J. (1979). Mental health and rural America: An overview and annotated bibliography (DHEW Publication No. ADM 78-753). Washington, DC: US Government Printing Office.

Fox, J. C., Blank, M. B., Kane, C. F., & Hargrove, D. S. (1994). Balance theory as a model for coordinating delivery of rural mental health services. *Applied and Preventive Psychology, 3*(1), 121–129.

Freeman, H., Blendon, R., Aiken, L., Sudman, S., Mullinix, C. F., & Corey, C. R. (1987). Americans report on their access to health care. *Health Affairs, 6,* 6–17.

Gesler, W. M., & Ricketts T. C. (Eds.) (1992). *Health in rural North America: The geography of health care services and delivery.* New York: Rutgers University Press.

Harding, L., & Sewel, J. (1992). Psychological health and employment status in an island community. *Journal of Occupational and Organizational Psychology, 65,* 269–275.

Hargrove, D. S. (1982). An overview of professional considerations in the rural community. In P. A. Keller & J. D. Murray (Eds.), *Handbook of rural community mental health* (pp. 169–189). New York: Human Sciences Press.

Hargrove, D. S. (1983). The rural psychologist as generalist: A challenge for professional identity. *Professional Psychology, 13*(2), 302–308.

Hargrove, D. S. (1986). Ethical issues in rural mental health practice. *Professional Psychology: Research and Practice, 17,* 20–23.

Hargrove, D. S., & Howe, H. E. (1981). Training in rural mental health delivery: a response to prioritized needs. *Professional Psychology, 12,* 722–731.

Hargrove, D. S., & Melton, G. B. (1987). Block grants and rural mental health services. *Journal of Rural Community Psychology, 8,* 4–11.

Herzlich, C. (1974). *Health and illness: A social psychological analysis.* London: Academic Press.

Heyman, S. R. (1982). Capitalizing on unique assets of rural areas for community interventions. *Journal of Rural Community Psychology, 3*(1), 35–48.

Hoyt, D. R., O'Donnell, D., & Mack, K. Y. (1995). Psychological distress and size of place: The epidemiology of rural economic distress. *Rural Sociology, 60,* 707–720.

Human, J., & Wasem, C. (1991). Rural mental health in America. *American Psychologist, 46,* 232–239.

Jablensky, A., Sartorius, N., Ernberg, G., et al. (1992). Schizophrenia: manifestation, incidence, and course in different cultures: A World Health Organization ten country study. *Psychological Medicine Supplement, 20,* 97.

Jahoda, M. (1972). *Employment and unemployment: A social psychological analysis.* Cambridge, UK: Cambridge University Press.

Jeger, A. M., & Slotnick, R. S. (1982). Community mental health: Toward a behavioral-ecological perspective. In A. M. Jeger and R. S. Slotnick (Eds.), *Community mental health and behavioral-ecology: A handbook of theory, research, and practice* (pp. 7–26). New York: Plenum.

Jensen, L., & Tienda, M. (1989). Nonmetropolitan minority families in the United States: Trends in racial and ethnic economic stratification, 1959–1986. *Rural Sociology, 54,* 509–532.

Jerrell, J., & Herring, J. (1983). The role of psychologists in rural community mental health facilities. *Journal of Rural Community Psychology, 4*(2), 3–17.

Johnson, K. M., & Beale, C. L. (1992). Natural population decrease in the United States. *Rural Development Perspectives, 8,* 20–27.

Jones, L. R., Badger, L. W., Ficken, R. P., et al. (1987). Inside the hidden mental health network examining mental healthcare delivery of primary care physicians. *General Hospital Psychiatry, 9,* 287–293.

Kamerow, D. B., Pincus, H. A., & MacDonald, D. I. (1986). Alcohol abuse, other drug abuse and mental disorders in medical practice: Prevalence, costs, recognition, and treatment. *Journal of the American Medical Association, 255* (15), 2054–2057.

Keating, N. C. (1987). Reducing stress of farm men and women. *Family Relations, 36,* 358–363.

Keller, P. A., & Prutsman, T. D. (1982). Training for professional psychology in the rural community. In P. A. Keller & J. D. Murray (Eds.), *Handbook of rural community mental health.* New York: Human Sciences Press.

Kenkel, M. B. (1986). Stress-coping-support in rural communities: A model for primary prevention. *American Journal of Community Psychology, 14*(5), 457–478.

Kessler, R. C., McGonagle, C. A., Zhao, S. et al. (1994) Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: Results from the National Comorbidity Survey. *Archives of General Psychiatry, 51,* 8–19.

Kropotkin, P. (1939). *Mutual aid.* Harmondsworth, UK: Penguin.

Labao, L. (1996). A sociology of the periphery versus a peripheral sociology: Rural sociology and the dimension of space. *Rural Sociology, 61,* 77–102.

Lee, A. S., & Bowles, G. K. (1974). *Policy implications of the movement of blacks out of the rural south.* Washington, DC: Economic Research Service, National Science Foundation.

Lewin-Epstein, N. (1991). Determinants of regular source of health care in black, Mexican, Puerto Rican, and non-Hispanic white populations. *Medical Care, 29*(6), 543–557.

Li, P. S., & MacLean, B. D. (1989). Changes in the rural elderly population and their effects on the small town economy: The case of Saskatchewan 1971–1986. *Rural Sociology, 54,* 213–226.

Libertoff, K. (1980). Natural helping networks in rural youth and family services. *Journal of Rural Community Psychology, 1*(1), 4–17.

Lichter, D. T. (1989). Race, unemployment hardship, and inequality in the American nonmetropolitan south. *American Sociological Review, 54,* 436–446.

Litwak, E., & Meyer, H. J. (1966). A balance theory of coordination between bureaucratic organizations and community primary groups. *Administrative Science Quarterly, 2*(1), 31–58.

Mansfield, P., Preston, D., & Crawford, C. (1988). Rural–urban differences in women's well-being. *Health Care for Women International, 9,* 289–304.

Mazer, M. (1976). *People and predicaments: Of life and distress on Martha's Vineyard.* Cambridge, MA: Harvard University Press.

Melton, G. B. (1983). Ruralness as a psychological construct. In A. W. Childs & G. B. Melton (Eds.), *Rural psychology*. New York: Plenum.

Melton, G. B., & Hargrove, D. S. (1987). Perceptions of urban and rural communities. *Journal of Rural Community Psychology, 8,* 3–13.

Mermelstein, J., & Sundet, P. (1988). Factors influencing the decision to innovate: The future of community responsive programming. *Journal of Rural Community Psychology, 9,* 61–75.

Miller, M. K., Farmer, F. L., & Clarke, L. L. (1994). Rural populations and their health. In J. E. Beaulieu & D. E. Berry (Eds.), *Rural health services: A management persepective*. Ann Arbor, MN: Alpha Press/Health Administration Press.

Morlock, L. L. (1989). Recognition and treatment of mental health problems in the general health sector. In C. A. Taube, D. Mechanic, & A. A. Hohmann (Eds.), *The future of mental health services research* (pp. 39–61). Rockville, MD: National Institute of Mental Health.

Murray, J. D. (1990). Professional survival and success in the rural community. *The Clinical Psychologist, 10,* 16–21.

Murray, J. D., & Keller, P. A. (1991). Psychology and rural America: Current status and future directions. *American Psychologist, 46,* 220–231.

Murray, J. D., & Kupinsky, S. (1982). In P. A. Keller & J. D. Murray (Eds.), *Handbook of rural community mental health* (pp. 62–73). New York: Human Sciences Press.

Murray, M., & Dolomount, M. (1995). *Accidents in the inshore*. Report submitted to the Occupational Health and Safety Branch, Department of Employment and Labour Relations, Government of Newfoundland and Labrador.

Naples, N. A. (1994). Contradictions in agrarian ideology: Restructuring gender, race-ethnicity, and class. *Rural Sociology, 59,* 150–165.

Narrow, W., Regier, D., Rae, D., et al. (1993). Use of services by persons with mental and addictive disorders. *Archives of General Psychiatry, 50,* 95–107.

National Association for Rural Mental Health (1993). *Where we stand: Health care reform, rural mental health.* Washington, DC: Authors.

National Farmers' Union (1989). *Statement on the subject of the foreclosure of farm land in the prairie region.* Saskatoon, Canada: Authors.

Neff, J. A. (1983). Urbanicity and depression reconsidered. *Journal of Nervous and Mental Disease, 171,* 546–552.

Olson, K. R., & Schellenberg, R. P. (1986). Farm stressors. *American Journal of Community Psychology, 14,* 555–569.

Ortega, S. T., Johnson, D. R., Beeson, P. G., & Craft, B. J. (1994). The farm crisis and mental health: A longitudinal study of the 1980s. *Rural Sociology, 59,* 598–619.

Pargament, K., Falgout, K., Ensing, D. S., Reilly, B., et al. (1991). The Congregation Development Program: Data-based consultation with churches and synagogues. *Professional Psychology Research and Practice, 22*(5), 393–404.

Preston, J., Brown, F. W., & Hartley, B. (1992). Using telemedicine to improve health care in distant areas. *Hospital and Community Psychiatry, 43,* 25–32.

Redick, R. W., Witkin, J. J., Atay, J. E., et al. (1992). Specialty mental health system characteristics. In Center for Mental Health Services and National Institute of Mental Health, R. W. Manderscheid & M. A. Sonnenschein (Eds.), *Mental health, United States.* DHHS Pub. No. (SMA) 92–1942. Washington, DC: Supt. of Docs., US Government Printing Office.

Regier, D. A., Goldberg, I. D., & Taube, C. A. (1978). The *de facto* US mental health services system: A public health perspective. *Archives of General Psychiatry, 35,* 685–693.

Rich, S. (1994). Study: Half of US adults have had mental illness. *Washington Post,* January 14, p. 2.

Robert Wood Johnson Foundation (1987). *Access to health care in the United States: Results of a 1986 Survey,* No. 2. Princeton, NJ: Authors.

Robins, L. N., & Regier, D. A. (1991). *Psychiatric disorders in America: The epidemiologic catchment area study.* New York: Free Press.

Rosemann, M. R., & Delworth, U. (1990). Clinical and community perspectives on the farm crisis. *The Clinical Psychologist, 10,* 10–16.

Rosenblatt, P. C., & Keller, L. O. (1983). Economic vulnerability and economic stress in farm couples. *Family Relations, 32,* 567–573.

Rowland, D., & Lyons, B. (1989). Triple jeopardy: Rural, poor, and uninsured. *Health Services Research, 23*(6).

Sanderson, G. (1974). The sweatshop legacy: still with us in 1974. *The Labour Gazette, 74,* 400–417.

Scheidt, R. J. (1985). The mental health of the aged in rural environments. In R. T. Coward & G. R. Lee (Eds.), *The elderly in rural society.* New York: Springer.

Scheidt, R. J., & Windle, P. G. (1982). Well-being profiles of small town elderly in differing rural contexts. *Community Mental Health Journal, 18,* 257–267.

Short, J. (1991). *Imagined country.* London: Routledge.

Sladen, B. J., & Mozdzierz, G. J. (1989). Distribution of psychologists in underserved areas: Changes over time, 1970–1981. *Professional Psychology, 20,* 244–247.

Snipp, C. M. (1996). Understanding race and ethnicity in rural America. *Rural Sociology, 61,* 125–142.

Schroeder, E. H., Fliegel, F. L., VanEs, J. C. (1985). Measurement of the lifestyle dimensions of farming for small-scale farmers. *Rural Sociology, 50,* 305–322.

Spitzer, R. L., Williams, J. B. W., Kroenke, K., Linzer, M., de Gruy, F. V., Hahn, S. R., Brody, D., Johnson, I. G.

(1994). Utility of a new procedure for diagnosing mental disorders in primary care: The PRIME-MD 100 study. *Journal of American Medical Association, 272,* 1749.

Srole, L. (1972). Urbanization and mental health: A reformulation. *Psychiatric Quarterly, 44,* 449–461.

Statistics Canada (1993). *1991 Census of Canada: profile of urban and rural areas, Part A: Canada, Provinces and Territories.* Ottawa: Industry, Science and Technology Canada.

Statistics Canada (1996). *Trends and highlights of Canadian agriculture and its people (Cat. 96-303E).* Ottawa: Industry, Science and Technology Canada.

Sundberg, N. D. (1986). Prevention and promotion—mental health in rural and small town settings. *Journal of Rural Community Psychology, 7*(2), 3–6.

Susser, E., Schanzer, B., Varma, V. K., & Gittelman, M. (1996). Topics for our times: Can we learn from the acre of persons with mental illness in developing countries? *American Journal of Public Health, 86,* 926–927.

Swanson, L., & Butler, M. (1987). Human resource base of rural economies. In Department of Agriculture, Economic Research service. *Rural economic development in the 1980s: preparing for the future.* Washington, DC: US Government Printing Service.

Task Force on Incomes and Adjustment in the Atlantic Fishery (1993). *Charting a new course: Towards the fishery of the future.* Cat. No. Fs 23–243/1993. Ottawa, Canada: Minister of Supply and Services.

Tolbert, C. M., & Lyson, T. A. (1992). Earnings inequality in the nonmetropolitan United States: 1967–1990. *Rural Sociology, 57,* 494–511.

Trainor, J., Pope, B., & Pomeroy, E. (1997). Critical challenges for Canadian mental health policy. *Canadian Review of Social Policy/Revue Canadienne de Politique Sociale, 39,* 55–64.

Turner, J. C., & TenHoor, W. J. (1978). The NIMH community support program: Pilot approach to a needed social reform. *Schizophrenia Bulletin, 4,* 319–348.

US Congress, Office of Technology Assessment (1990). *Health care in rural America* (OTA-H-434). Washington, DC: Government Printing Office.

US Department of Agriculture (1985). The current financial conditions of farmers and farm lenders. *Economic Research Service, Agriculture Information Bulletin, No. 490.* Washington, DC: Author.

Wagenfeld, M. O., & Wagenfeld, J. K. (1981). Values, culture, and the delivery of mental health services in rural areas. In M. O. Wagenfeld (Ed.), *Perspectives on rural mental health* (New Directions for Mental Health Services Series, No. 9). San Francisco: Jossey-Bass.

Wagenfeld, M. O., Goldsmith, H. F., Stiles D., et al. (1988). Inpatient mental health services in metropolitan and non-metropolitan counties. *Journal of Rural Community Psychology, 9*(2), 13–28.

Walker, L. S., & Walker, J. L. (1987). Stressors and symptoms predictive of distress in farmers. *Family Relations, 36,* 374–378.

Weber, S. (1990). Mental health and rural aging. *The Rural Elderly Networker, 2*(2), 1–5.

Wilkinson, R. (1996). *Unhealthy societies: The Afflictions of inequality.* London: Routledge.

Zahner, G. E. P., Jacobs, J. H., Freeman, D. H., & Trainor, K. F. (1993). Rural–urban child psychopathology in a northeastern US state: 1986–1989. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 378–387.

Zelman, M. R. (1995). Telemedicine, POTS and PANS technology and rural health care in Texas. *Texas Journal of Rural Health, 14*(2), 1–4.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.12

Objective Personality Assessment: Computer-based Minnesota Multiphasic Personality Inventory-2 Interpretation in International Clinical Settings

JAMES N. BUTCHER
University of Minnesota, Minneapolis, MN, USA
ELLEN BERAH
Monash Medical Centre, Clayton, Vic, Australia
BJORN ELLERTSEN
University of Bergen, Norway
PATRICIA MIACH
Monash Medical Centre, Clayton, Vic, Australia
JEEYOUNG LIM
Samsung Group, Jeongdun-Maul, Sungnam-City, South Korea
ELAHE NEZAMI
USC/IPR, Los Angeles, CA, USA
PAOLO PANCHERI
5a Cattedra di Clinica Psichiatrica, Rome, Italy
JAN DERKSEN
University of Nijmegen, The Netherlands
and

MOSHE ALMAGOR
University of Haifa, Israel

10.12.1 INTRODUCTION 278
10.12.1.1 Cultural Factors in Mental Health Assessment 279
10.12.1.2 Application of Psychological Tests Across Cultural Boundaries 279
10.12.1.3 Development of Automated Interpretive Strategies for Objective Personality Tests 280
10.12.2 EVALUATION OF COMPUTER-BASED MMPI-2 REPORTS IN INTERNATIONAL CONTEXTS: RESULTS FROM AUSTRALIA, FRANCE, NORWAY, AND THE USA 280
10.12.2.1 The Research Protocol 280
10.12.2.2 General Testing Procedures 280
10.12.2.3 Clinical Rating Procedures 281
10.12.2.4 The Test Sites 281
10.12.2.4.1 Australian clinical sample 281
10.12.2.4.2 French clinical sample 281
10.12.2.4.3 Norwegian clinical sample 282
10.12.2.4.4 US clinical sample 282
10.12.2.5 Results of the Studies 282
10.12.2.5.1 Validity considerations 282
10.12.2.5.2 Symptomatic pattern 283
10.12.2.5.3 Interpersonal relations 283
10.12.2.5.4 Diagnostic considerations 288
10.12.2.5.5 Treatment considerations 288
10.12.2.5.6 Overall accuracy 289
10.12.2.6 Discussion of Overall Accuracy 289
10.12.2.7 Limitations of this Research 290
10.12.3 CASES 290
10.12.3.1 Nijmegen, The Netherlands: The Case of Susan 290
10.12.3.1.1 The Minnesota Report on Susan 292
10.12.3.1.2 Comment 294
10.12.3.1.3 DSM-IV classification 294
10.12.3.1.4 Treatment 294
10.12.3.2 Rome, Italy: The Case of Mario P. 294
10.12.3.2.1 The Minnesota Report on Mario P. 295
10.12.3.2.2 Comment 298
10.12.3.2.3 Treatment 298
10.12.3.3 Iran, United States: The Case of Ms. B. 298
10.12.3.3.1 The Minnesota Report on Ms. B. 300
10.12.3.3.2 Diagnostic impression 302
10.12.3.3.3 Treatment 302
10.12.3.4 Seoul, Korea: The Case of Kim 302
10.12.3.4.1 Family history 303
10.12.3.4.2 Behavioral observations 303
10.12.3.4.3 Symptomatic behavior 303
10.12.3.4.4 The Minnesota Report on Kim 303
10.12.3.4.5 Diagnosis 306
10.12.3.5 Haifa, Israel: The Case of Yuri Z. 306
10.12.3.5.1 The Minnesota Report on Yuri Z. 307
10.12.3.5.2 Treatment 310
10.12.3.6 Cross-national Generality of Psychopathology: What We Have Learned from the Cases 310
10.12.4 FUTURE DIRECTIONS 310
10.12.5 SUMMARY 311
10.12.6 REFERENCES 311

10.12.1 INTRODUCTION

The use of psychological testing in clinical assessment situations has been expanding from its origins within Europe and the USA to numerous other countries around the world. The rapidly increasing use of Western-derived clinical tests results from several factors including: the expansion of available mental health services in many countries since the mid-1980s, the growth in the number of assessment-trained psychologists in other countries, an increased recognition that psychological assessment can provide valuable information in mental health contacts, and as a result of increased professional communication through international congresses. Many psychologists have become aware of the effectiveness of psychological tests in countries like the USA and have adapted these procedures for use in their own countries.

Can psychological assessment and clinical diagnostic procedures developed in one language and culture be readily adapted and applied in another? Do stimuli employed in objective personality tests, such as the items in the most widely used clinical measure, the Minnesota Multiphasic Personality Inventory (MMPI-2), transcend the gulf of language and culture and apply effectively in different cultural contexts? Do psychological tests that are translated into other languages require special test development methods? Can norms from one culture be "transplanted" to another? These questions have been addressed extensively in the cross-cultural testing literature and are briefly summarized below (Butcher, 1996; Kazarian & Evans, in press; Lonner & Berry, 1986).

10.12.1.1 Cultural Factors in Mental Health Assessment

Although language and customs, as reflected in familial and social practices, vary between countries, there are also many common features. Psychological disorders appear to be generally comparable across different cultures and, although there have reportedly been some "exotic" or culture-specific conditions (Yap, 1951), many commonalities exist with similar symptoms and common manifestations across diverse groups. For example, schizophrenia appears to occur in all known cultures and societies through history (where records permit comparison) although the symptoms may vary somewhat and rates may differ (Butcher, Narikiyo, & Bemis-Vitousek, 1992).

Commonalities across cultures in mental disorder have allowed for the development of the international classification of mental disorders (World Health Organization, 1992) referred to as ICD-10 (the most up-to-date version). The ICD-10 diagnostic system parallels the diagnostic system developed in the USA, the *Diagnostic and statistical manual of mental disorders* (4th ed.); *DSM-IV*), which is also in wide use in several other countries. Many psychiatrists employ *DSM-IV* instead of ICD-10. The very existence of a common language for describing these psychological disorders and their apparent relevance around the world suggests that the elements of mental disorders, that is, the symptoms, features, established course, and so forth, are relatively stable if not universal across cultures.

10.12.1.2 Application of Psychological Tests Across Cultural Boundaries

The use of clinical personality tests across cultures has a history dating back to the 1920s. There were several early studies devoted to exploring the generalizability of psychoanalytic test constructs across cultures by administering the Rorschach test to individuals from different cultural backgrounds and comparing their responses to those of Westerners (see Adcock & Richie, 1958). The use of the Rorschach in anthropological research was, however, replete with problems and this approach had pretty well run its course by the late 1950s. Lindzey (1958) discussed the problems of the projective method in cross-cultural research and summarized these difficulties.

The overall use of psychological tests across cultural boundaries was, however, not decreased by any means. In fact, the practical use of translated and adapted personality tests was only beginning. The 1950s and 1960s witnessed a number of projects to translate and adapt objective personality instruments across cultures, not for the purposes of anthropological study, but for the emphatically practical aim of making clinical decisions in a more objective manner than was currently available.

One of the most widely employed and internationally adapted personality inventories used in clinical assessment has been the MMPI, which was subjected to a number of early translations, including Italian (Reda, (1948), Japanese (Abe, (1955), and German (Sundberg, 1956). In the years that followed, over 150 translations of the MMPI were developed and the inventory came to be widely employed in over 46 countries (see Cheung & Song, 1989; Cheung, Song, & Butcher, 1991; Cheung, Zhao, & Wu, 1992; Kim, 1988; Risetti, Himmel, Maltes, & Gonzalez, 1989; Strassberg, Clutton, & Korboot, 1991; Strassberg, Tilley, Bristone, & Tian, 1992; Zou & Zhao, 1992). In 1976, Butcher and Pancheri published an international handbook on using the MMPI across cultures that described model translation projects, provided substantial psychometric equivalence data, and illustrated clinical validation efforts in international contexts.

With the redevelopment of the MMPI and the publication of MMPI-2 in 1989, a new wave of test translation projects was initiated. A number of recent studies have explored the use of the MMPI-2 in other cultures and a broad collection of clinical researchers, from several different countries, have described their translation and adaptation research efforts and highlighted clinical usage in a compendium of papers (Butcher, 1996). This international handbook includes three chapters on test translation methods and recommended procedures for evaluating translation equivalence. Central to using the MMPI-2 in cross-national settings is the establishment of sound translation of the items into the target language and culture and demonstration of test equivalence, but the core of the volume is 56 contributions from psychologists and psychiatrists from around the world.

10.12.1.3 Development of Automated Interpretive Strategies for Objective Personality Tests

Progress in civilization has often involved the evolution of labor-saving techniques or inventions that free human beings from routine or laborious tasks. Assessment psychologists have long been interested in the use of automated or mechanical methods for assisting in the routine and often onerous task of scoring and processing responses to psychological tests. Mechanical test-scoring devices have been employed in psychology since the 1940s, and in the 1950s, mainframe computers were first used to process large batches of test protocols quickly and accurately.

Paralleling this technical effort to develop more efficient data-processing techniques came conceptual progress into methods of making clinical decisions more objective. With several publications by Meehl (1954, 1956) and his followers (Gilberstadt & Duker, 1965; Halbower, 1955; Marks, Seeman, & Haller, 1974) the automatic interpretation of one test, the MMPI by mechanically combining established test correlates, received broad attention. The MMPI, being an instrument that was developed according to an empirical scale-validation approach, had acquired by 1960 a very substantial research base supporting the interpretation of scales and combinations of scales (profile codetypes). In the 1960s psychologists began to experiment with the actual interpretation of psychological test scores using computers. The first of these approaches, an MMPI computer interpretation program developed at the Mayo Clinic in Rochester, Minnesota, demonstrated the effectiveness and accuracy of objective interpretation with the computer (Rome et al., 1962). Computer test interpretation programs have evolved substantially over the succeeding decades and have become an accepted strategy for interpreting psychological tests in the USA (American Psychological Association [APA] 1986; Butcher, 1987; Eyde, Kowal, & Fishburne, 1991; Fowler, 1969; Ziskin, 1981).

With computerized test interpretation programs for the MMPI/MMPI-2 becoming widely available for clinical use in the USA (Butcher, 1995), psychologists in other countries also began to explore computer-derived personality assessments (Fowler & Butcher, 1987; Gillet et al., 1996; Pancheri & Biondi, 1987; Pancheri, Sirigatti, & Biondi, 1996). An early study of the utility of a computer-based report (Roche Psychiatric Service Institute) on the original version of the MMPI in Australia produced promising results (Fowler, 1978). More recently, research has confirmed that computer-based MMPI-2 reports provided accurate and useful information when applied with Australian psychiatric patients (Berah et al., 1993; Berah, Miach, & Butcher, 1995).

In this chapter we explore the objective use of psychological tests in cross-cultural settings in two ways. We describe an empirical evaluation of the generalizability and accuracy of computer-derived MMPI-2 reports describing the symptoms and behavior of patients in four countries—Australia, France, Norway, and the USA. This discussion will be followed by several case examples from other countries.

10.12.2 EVALUATION OF COMPUTER-BASED MMPI-2 REPORTS IN INTERNATIONAL CONTEXTS: RESULTS FROM AUSTRALIA, FRANCE, NORWAY, AND THE USA

10.12.2.1 The Research Protocol

The research design for this project was similar in all four countries. Clinicians who were evaluating their patients administered the MMPI-2 to each client using a booklet format. The appropriate language version was used in each country. A broad range of clients were tested in a variety of settings: inpatient, outpatient, court-ordered, neuropsychological, and gender reassignment assessment. No effort was made to obtain homogeneous research groups but to simply evaluate any available patients being assessed in an ongoing process, so that the psychologist would have sufficient information about the patient to be able to rate the adequacy of the computer-based report.

10.12.2.2 General Testing Procedures

The MMPI-2 testing was conducted by psychologists who were seeing the patients for clinical evaluations or psychological treatment.

Following an explanation to the patient of the reasons for the assessment, testing took place in individual sessions using standardized test instructions. Softcover booklet format was used, with patients completing the test at a private table under supervision in the psychologist's office. The MMPI-2 was usually administered in one long session but, depending on the needs of the particular patients, it was occasionally administered in two or three shorter sessions. When there was any doubt about a patient's reading ability, a test to help determine reading comprehension was usually first administered. The completed MMPI-2 answer sheets and relevant demographic information were sent to James Butcher in Minneapolis, Minnesota for scoring and data processing. The item responses of each patient were key-entered or optically scanned to a computer disk and scored by personal computer. The Minnesota Report, a computer-based narrative interpretation program (Butcher, 1993), was generated by microcomputer. (It is important to note that the American norms were used in all of the countries to compare the cases.) The computer report printout was sent back to the test location for the next phase of the project: the ratings of report adequacy.

10.12.2.3 Clinical Rating Procedures

Ratings were completed by the same psychologists who were seeing the patients in the context of psychological evaluation or therapy at the time of the study. The Rating Form required the following information on each report: Patient Information (gender, age, education, marital status, clinical setting, and clinician's hours spent with the patient); Rater Information (degree, profession, years of experience); Report Rating (the validity considerations, symptomatic pattern, interpersonal relations, diagnostic considerations, and treatment considerations sections of each report were each rated as providing "insufficient," "some," "adequate," "more than adequate" or "extensive" information. Raters also indicated the percentage of statements in the report they considered to be accurate descriptions of the patient, using the categories of less than 20, 20–39, 40–59, 60–79, or 80–100%). Finally, open-ended questions were asked for symptomatic and diagnostic information, and for ways in which the report could have been improved.

The raters participating in the evaluation program were clinicians who were seeing the patient in the professional diagnostic context or for therapy at the time of the study. Each rater was given a copy of the computer report output and asked to evaluate the adequacy and accuracy of the narrative report in describing the behavior and symptoms of the client and to complete the Rating Form.

10.12.2.4 The Test Sites

10.12.2.4.1 Australian clinical sample

The Australian patients were clients in the Adult Psychiatry Service of Monash Medical Centre, Melbourne, Australia, a 700-bed general hospital affiliated to Monash University. The Adult Psychiatry Service provides treatment for patients from its local and regional communities, from other services of the hospital and, for its specialist programs, from the wider community. It has a 36-bed inpatient unit which admits informal and involuntary patients, several community clinics, a mobile community treatment team, and extensive consultation–liaison services to the divisions of obstetrics and gynecology, medicine and surgery. Specialist inpatient and outpatient assessment and treatment programs are provided through the eating disorders and the mothers and babies clinics. A specialist gender dysphoria clinic, the only such clinic in Australia, accepts referrals from the whole country for assessment of gender dysphoria and admission to the program for sexual reassignment surgery. The Adult Psychiatry Service works on a multidisciplinary basis with a staff of 11 clinical psychologists.

Patients were referred to the Adult Psychology Section by their consultant psychiatrist for psychological assessment to help in differential diagnosis and treatment planning in the early stage of admission or outpatient presentation. The two exceptions to this were patients from the eating disorders and the gender dysphoria clinics who are routinely administered the MMPI-2. The patients (N = 167) included 67 men and 100 women with an age range of 16 to 75; 94 were being evaluated as outpatients, 69 were inpatients, and four were from other settings. The clinical diagnoses of the patients covered a broad range of problems. The amount of time the clinician-raters spent with patients is an important consideration in the study. Nearly all patients (97%) had been seen by the psychologist for three or more hours and half had been seen for five or more hours. The amount of contact between raters and patients, in all likelihood, provided raters with sufficient knowledge of the patients' problems and symptoms to be an appropriate source of external information about the patients.

10.12.2.4.2 French clinical sample

The French clinical study involved testing 100 patients (54 inpatients and 46 outpatients) in two specialized psychiatric departments in Paris: the Clinique des Maladies Mentales et de l'Encéphale (CMMD), which is part of the Hôpital Sainte Anne, and the Hôpital International de l'Université de Paris (see Gillet et al., 1996). The test was administered using the French language MMPI-2 booklet. A total of 22 mental health specialists, all of them psychiatrists, evaluated the 100 computerized reports. Their clinical experience varied as follows: 33 reports were analyzed by five practitioners with 4–6 years of experience in psychiatry, 33 reports were analyzed by 10 practitioners with 7–10 years' experience, and 33 reports were analyzed by 10 practitioners with more than 10 years' experience. The time spent with each patient varied from two to six hours (or more). Demographic data for the population studied include the following: four patients were between the ages of 18 and 19 years; 71 patients were between the ages of 20 and 40 years; 22 patients were between the age of 41 and 60 years; and three patients were older than 60 years. The number of years of schooling for the sample was as follows: 21 patients had less than 10 years; 26 patients had between 10 and 12 years; 23 patients had between 13 and 14 years; and 30 patients had 15 years or more.

10.12.2.4.3 Norwegian clinical sample

The Norwegian study involved 99 valid patient protocols. Among these, 60 patients were in psychotherapy (outpatients), eight were psychiatric inpatients, 10 were medical (neurology) inpatients, nine were substance abuse inpatients and 10 were in counseling settings. A total of 10 clinicians with PhDs, situated in different counties covering all parts of Norway, evaluated the protocols. Nine of these clinicians evaluated 10 protocols whereas one evaluated nine protocols. Among the clinicians, nine had more than four years of clinical experience whereas one had three years. The time spent with the patients being evaluated was more than six hours in 75% of the cases, five or six hours in 10%, three or four hours in 8% and two hours in 7%. The clinical sample consisted of 56 males (mean age 35, range 10–62) and 43 females (mean age 35, range 18–58).

10.12.2.4.4 US clinical sample

The patients included in the US study were 263 people who were being evaluated using the Minnesota Report in the Minnesota Psychotherapy Project (Butcher, 1996). A total of 44 psychotherapists from several states and from diverse clinical settings (including independent practice, mental health clinics, hospital practice, court setting, and so forth) were asked to rate their patients on a number of personality and symptom variables. They were also provided with a copy of the Minnesota Report derived from the patient's responses to the MMPI-2. The therapists had a broad range of backgrounds, although the majority had PhDs (46%) or PsyDs (20.5%) or MAs (23%) in psychology; 50% of these were trained in clinical and 23% in counseling psychology, and 2.3% were trained in social work. Finally, 63.6% considered themselves cognitive-behavioral; 11.4% were behavioral; 31.8% were psychodynamic; and 6.5% were client-centered. The mean number of years the therapists had been conducting a practice in psychotherapy was 11.5 years.

10.12.2.5 Results of the Studies

The results of the four projects will be reported for each section of the Minnesota Report. For example, under the heading "Validity considerations" the findings from each country will be presented. In addition, the results from the four countries will be shown in graphic form to give the reader a visual comparison of the summaries of the clinicians' ratings in each country.

10.12.2.5.1 Validity considerations

(i) Australia

Valid, interpretable profiles were produced by 95% of the sample ($N = 59$). While the validity considerations sections of the computerized reports were provided and rated for all patients, the remaining sections of the reports were unavailable for patients with invalid profiles. Over 90% of the computer-based reports were rated as providing adequate-to-extensive information in their appraisal of patients' approach to the testing (Figure 1); fewer than 5% of the reports were considered to have insufficient information about protocol validity.

(ii) France

Of the 100 MMPI-2s completed, only 17 profiles were found to be invalid and were not included in the study. Of these 17 invalid MMPI-2 inventories, 11 of the patients had been diagnosed as psychotic, three had personality disorders, one was suffering from an eating disorder, another from an affective disorder, and the last from a mixed disorder (Axis I and Axis II on the *DSM-III-R*). The raters evaluated 83 profiles, from 25 males and 58 females. Only 13.3% of the cases were thought to provide insufficient information; 61.3% were viewed as adequate, more than adequate, or as providing extensive information (Gillet et al., 1996; Figure 2).

(iii) Norway

All of the Norwegian sample produced valid, interpretable profiles. Clinicians evaluating the assessment adequacy of the Minnesota Report validity in appraising their patient's approach to the testing rated the reports as providing adequate-to-extensive information (82.8%). Only 4% of the reports rated were considered to have insufficient information about the patient (Figure 3).

(iv) USA

Fewer than 1% of the reports were considered to provide insufficient information, and 92% of the reports were considered to provide adequate (50.2%), more than adequate (27.2%), or extensive information (14.8%) (Figure 4).

10.12.2.5.2 Symptomatic pattern

(i) Australia

The description of patients' symptom patterns provided by the computerized reports was considered by the clinician-raters to provide adequate, more than adequate, or extensive information in 77% of cases (Figure 1). Information was rated as insufficient in 8% of reports.

(ii) France

According to Gillet et al. (1996), the clinicians' opinions of the quality of descriptions in the symptomatic pattern section were favorable in 59% of the cases, if the categories "extensive," "adequate," and "more than adequate" and more than sufficient information were pooled (Figure 2).

(iii) Norway

In terms of symptom pattern exhibited by the patients in the Norwegian study, 78.9% of the Minnesota Reports were rated as providing adequate, more than adequate, or extensive information, and only 2.0% of the narrative reports were considered insufficient by the judges (Figure 3).

(iv) USA

Only 1.5% of the Minnesota Reports were considered to provide insufficient information on the symptomatic pattern section of the report. On the other hand, 89.5% of the Minnesota Reports were rated as having adequate (39.7%), more than adequate (31%) or extensive information (17.6%) (Figure 4).

These findings are consistent with the results of other studies of the generalizability of MMPI-2 descriptors to patients in other countries (Butcher & Pancheri, 1976; Manos, 1984; Savasir & Erol, 1990) and provide support for the general clinical impression reported by the raters that the MMPI-2 correlates can be confidently applied to patients in other countries.

10.12.2.5.3 Interpersonal relations

(i) Australia

Of the reports, 78% were considered to provide adequate, more than adequate, or extensive information on patients' interpersonal relationships, while 6% were rated as providing insufficient information (Figure 1).

(ii) France

The clinicians' ratings of the interpersonal relations section of the 83 patient reports were favorable in 70% of the cases when the categories "extensive," "adequate," and "more than adequate" information were pooled (Gillet et al. 1996) (Figure 2).

(iii) Norway

The MMPI-2 report also provided substantial information about the interpersonal behavior of clients, according to practitioner ratings. Only 4% of the reports were judged to provide insufficient information on this variable; however, 76% of the narrative reports were considered adequate, more than adequate, or extensive in information provided (Figure 3) and 21% of the raters stated that the reports provided some information.

(iv) USA

In the US sample, only 2.7% of the patient reports were considered to provide insufficient information (Figure 4) while 85.4% were considered to be adequate (39.2%), more than adequate (29.2%), or to provide extensive information (16.9%).

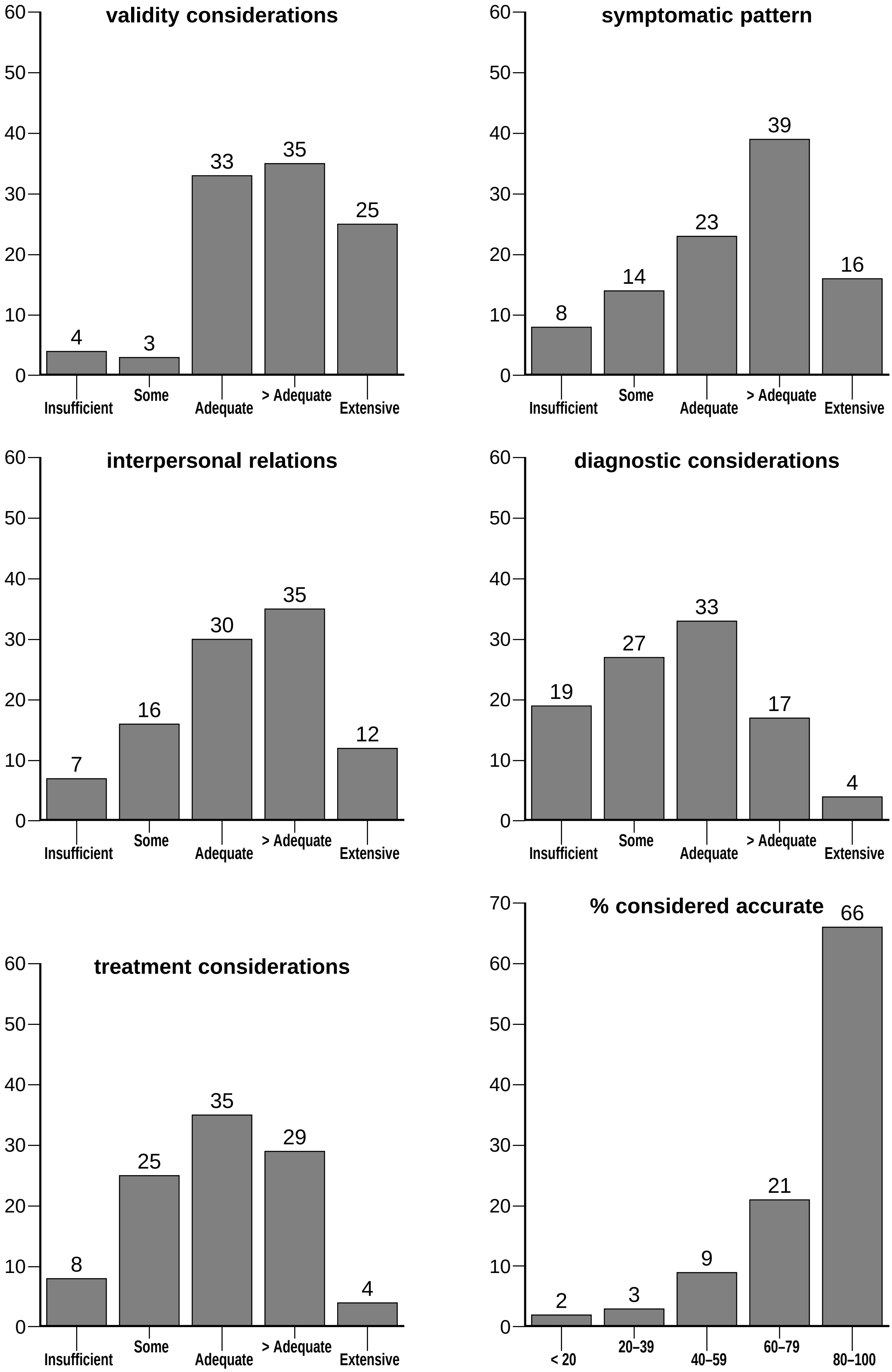

Figure 1 Evaluations of computer-based MMPI-2 reports: Australian sample.

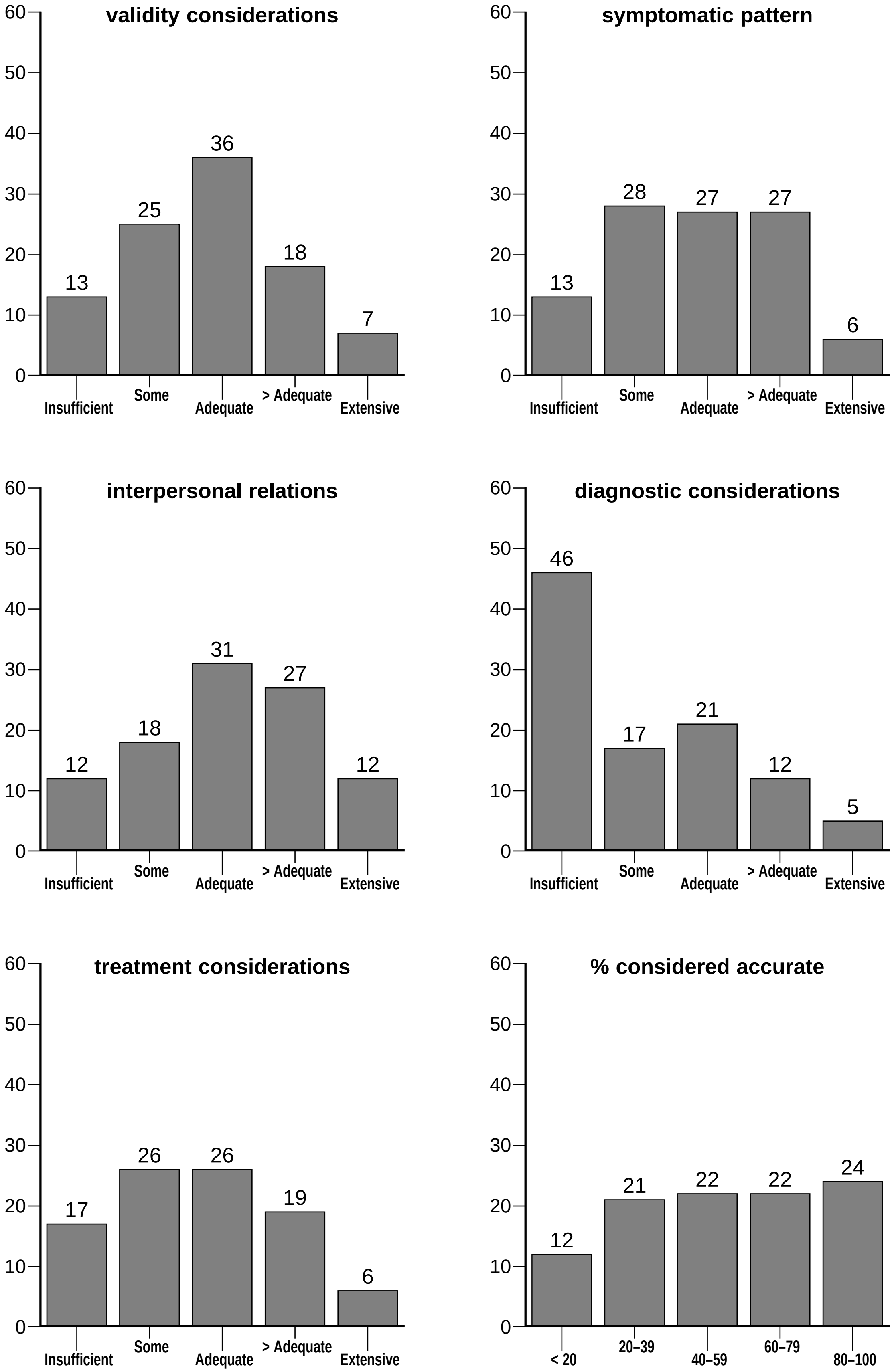

Figure 2 Evaluations of computer-based MMPI-2 reports: French sample.

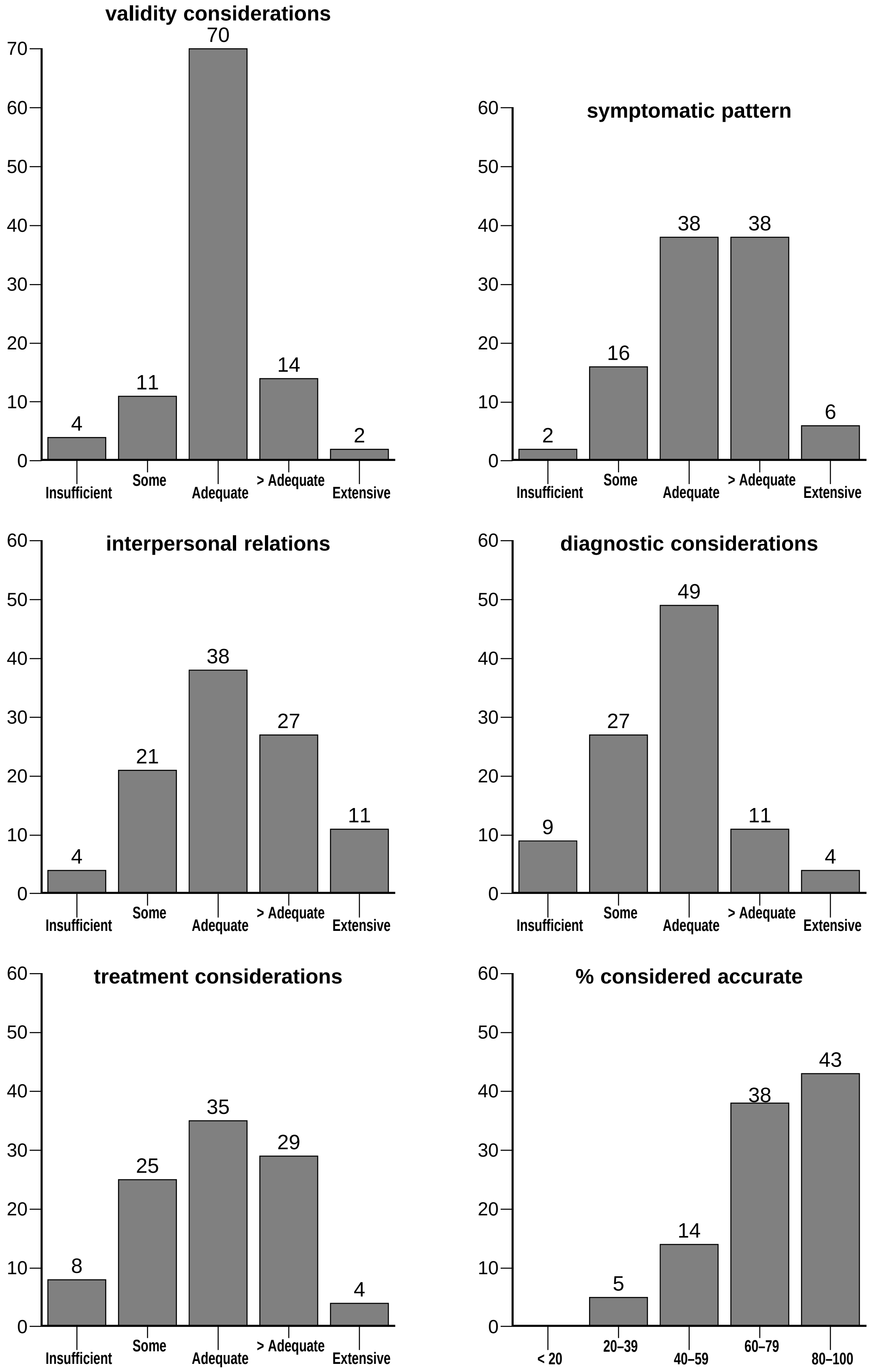

Figure 3 Evaluations of computer-based MMPI-2 reports: Norwegian sample.

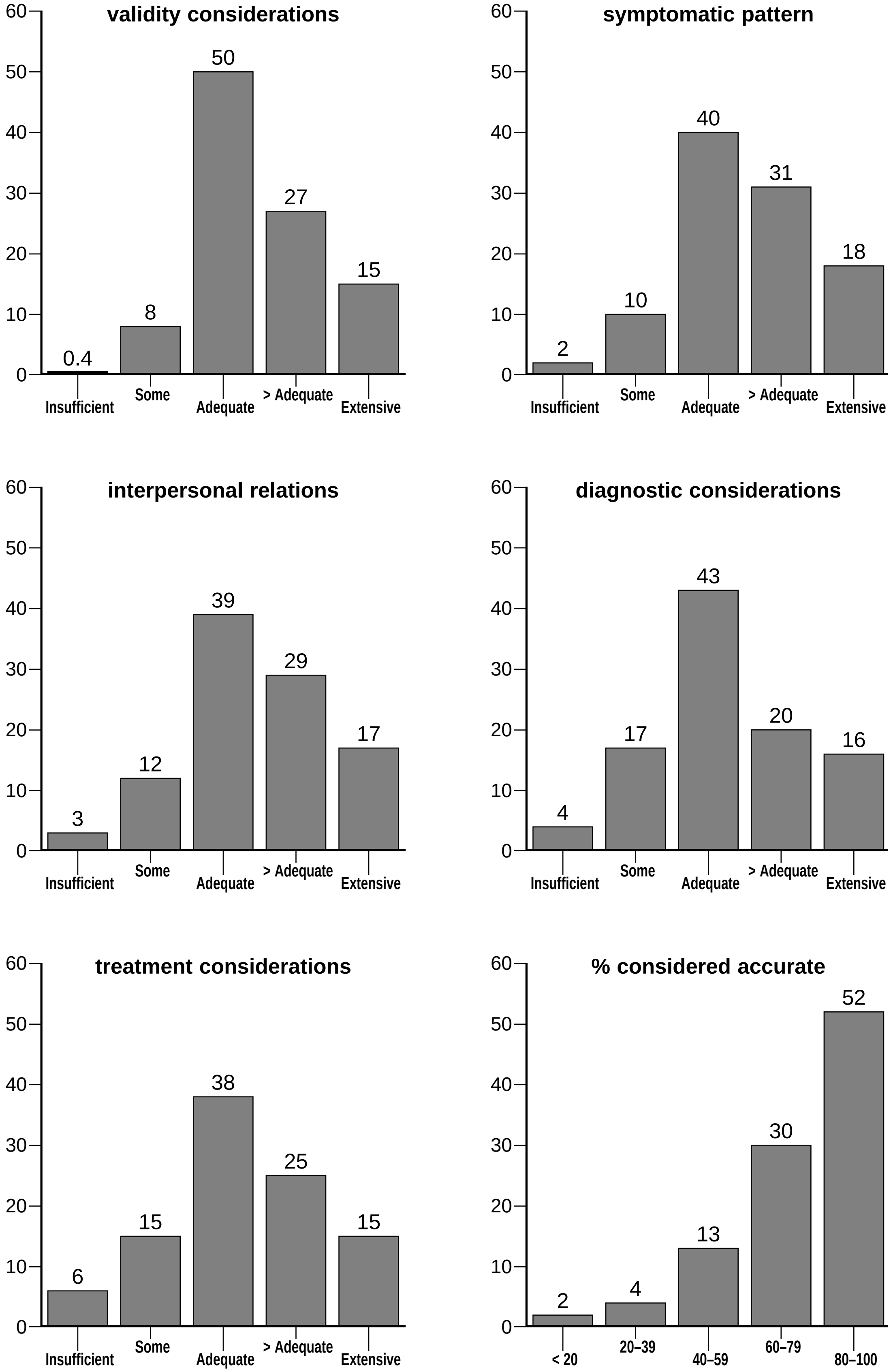

Figure 4 Evaluations of computer-based MMPI-2 reports: US sample.

Thus reports from all four countries were thought to provide substantial information about, the interpersonal behavior, of patients.

10.12.2.5.4 Diagnostic considerations

(i) Australia

This section of the computerized reports was considered by raters to be less adequate than the above sections. Even here, however, over half (55%) of the narrative reports were rated as containing at least adequate information on diagnosis, while 18% were thought to provide insufficient information (Figure 1).

(ii) France

According to Gillet et al. (1996), the clinicians' views of the diagnostic considerations section of the 83 reports were favorable in 38% of the cases when the categories "extensive," "adequate," and "more than adequate" and more than sufficient information were pooled (Figure 2). This finding of less relevant diagnostic information in the ratings probably results from the reliance upon a different diagnostic system in France than in the USA.

(iii) Norway

The diagnostic sections of the computer report were considered by raters to be slightly less comprehensive than other sections of the report (Figure 3). Nevertheless, over 60% of the narrative reports were judged to contain adequate, more than adequate, or extensive diagnostic information about their clients. Only 8.1% of the protocols were considered to provide insufficient information in the area of clinical diagnosis.

(iv) USA

The clinicians from the USA considered 4.2% of the reports as providing insufficient diagnostic information. However, 78.9% of the reports were thought to contain adequate (42.5%), more than adequate (20.1%) or extensive information (16.2%) (Figure 4).

Several factors may account for the variation in results in the rating of clinical diagnosis:

(i) Clinical diagnosis is a task that usually requires more information than is available through an objective self-report questionnaire. Practitioners usually need to obtain a detailed history and behavioral descriptions in order to obtain a clinical diagnosis.

(ii) The diagnostic procedures differed somewhat between countries with *DSM-IV* being used in the USA, Australia, and Norway and the ICD-10 in France.

(iii) Another difference to consider in the cross-cultural comparisons is that in France the ratings were completed by psychiatrists and in the other countries psychologists were the predominant professional performing the ratings.

The above differences aside, the majority of raters in all four countries considered the reports to have adequate, more than adequate, or extensive information available for the task of clinical diagnosis from the MMPI-2.

10.12.2.5.5 Treatment considerations

(i) Australia

As with diagnosis, 8% of reports were deemed to provide insufficient information regarding treatment planning. Two-thirds (68%) were considered to provide adequate, more than adequate or extensive information (Figure 1).

(ii) France

The clinicians' ratings of the treatment considerations section of the 83 reports were favorable in 51% of the cases when the categories "extensive," "adequate," and "more than adequate" and more than sufficient information were pooled (Figure 2). The lower agreement found between the clinicians and the computer-based reports, compared with other countries, probably resulted from different treatment orientations and practices in France.

(iii) Norway

Similarly to diagnosis, the treatment considerations section of the Minnesota Report was rated as providing valuable information on clients. Over 64% of the patient MMPI-2 reports were considered to provide adequate-to-extensive information for treatment planning (Figure 3). Only 8% of the patient reports were thought to provide insufficient information for treatment planning. The relatively high rate of positive evaluations of the treatment considerations section of the report suggests that treatment planning in Norway and the USA may rely upon similar types of clinical information, much of which is available through the MMPI-2.

(iv) USA

The therapists rating the American patients considered the reports to be helpful in treatment planning. Only 5.8% considered the reports to provide insufficient information while 78.9%

thought the reports provided adequate (38.1%), more than adequate (25.4%), or extensive information (15.4%).

In general, as with clinical diagnosis, therapists like to have extensive information for treatment planning, and test-based information is not the sole source of data required. However, in all four countries the majority of the clinicians rated the reports as adequate or better in their evaluations.

10.12.2.5.6 Overall accuracy

(i) Australia

In addition to rating the adequacy of the information provided in the narrative reports, raters estimated the percentage of statements they considered to be accurate descriptions of their patients (Figure 1). Two-thirds (66%) of reports on Australian patients were rated as having 80–100% accuracy, while 87% were rated as having more than 60% accuracy. Only 2% were rated as having less than 20% accuracy. As noted above, a large number of patients in the study (46) were gender-dysphoric clients undergoing evaluation for possible gender reassignment surgery. It might be hypothesized that their inclusion may have lowered the overall adequacy and accuracy ratings because of possible attempts by these patients to present themselves in an overly positive manner. This was not the case. There was no significant difference between the ratings of gender-dysphoric and other patients on the symptomatic pattern ($t(157) = -0.53$), the interpersonal relations ($t(157) = 1.14$), the diagnostic considerations ($t(156) = 1.79$) and the treatment considerations ($t(154) = 0.46$) sections of the reports. With respect to the two other ratings, the reports on the gender-dysphoric patients were actually rated more positively than those of the other patients (validity considerations, $t(165) = -2.84$, $p < 0.01$; overall accuracy, t for unequal variances $(110) = -2.15$, $p < 0.05$).

(ii) France

The overall relevance of the report to French clinicians was considered high when categories D and E were met and 67.5% if we consider categories C, D, and E.

(iii) Norway

In addition to the question of the thoroughness of the information provided by the narratives, the study addressed the accuracy of computer-generated personality descriptions at describing patients. As shown in Figure 3, the clinician ratings provided strong support for using the MMPI-2 in clinical assessment in Norway. Overall, 92.9% of the reports were rated in the adequate, more than adequate, or extensive information categories . Over 42.4% of the patient reports were classified as being in the 80–100% range of accuracy and 38.4% were judged to possess between 60 and 79% accuracy. Interestingly, only 5% of the narrative reports were considered low in accuracy. Overall, the accuracy at describing clinical behavior and symptoms in a diverse clinical sample was considered to be high.

(iv) USA

In terms of overall accuracy, the American clinicians rated the reports as being quite accurate. Only 1.9% were considered to be inaccurate while 81.6% were considered to be over 60% accurate with over half (52.1%) being rated in the highest category (Figure 4).

10.12.2.6 Discussion of Overall Accuracy

The results of these studies are generally consistent with the impressions of practitioners using computerized reports in other countries, namely that the automated MMPI-2 interpretations generalize well to other locations and that the descriptors about patients apply well in countries other than the USA. These results support the conclusion that computer-based reports of the MMPI-2 can be a valuable aid in clinical assessment in other countries.

Although the overall acceptability of the Minnesota Report in France was good, the ratings provided by French mental health professionals on the utility of the reports were somewhat lower than those obtained by their counterparts in Australia (Berah et al., 1993), Norway, and the USA. These findings might reflect differences in clinical practice in different countries, different attitudes toward computers, or genuine personality differences between the French and patients from other countries. Further research will be needed to clarify this finding.

Relatively few of the practitioners in any of the four countries studied found the reports inappropriate or inaccurate. The sections for which most of the reports were considered as valuable and to provide in-depth information were the validity considerations, the symptomatic patterns, and the interpersonal relations sections and, to a lesser degree, the diagnostic and the treatment considerations sections.

10.12.2.7 Limitations of this Research

The limited information that was available on each patient from the clinicians prevented a fuller exploration of the descriptive power of the MMPI-2. As with any field study conducted in service-oriented clinics or hospitals, the amount of extra work that can be asked of clinicians is limited. While more extensive ratings would have been desirable, the cooperation of clinicians may well have been sacrificed if the task was seen as overly burdensome.

Second, it should be kept in mind that the study employed US norms for the computer-derived interpretations. It is possible that the use of country-specific norms, particularly in France, might have added a higher degree of accuracy to the reports and greater congruence in the ratings. French norms are now available for the MMPI-2. Future research might be developed with country-specific norms, which gives different results.

Finally, we were not able, in this study, to provide an estimate of inter-rater reliability. It would have been desirable to have had a number of cases rated by at least two clinicians, so that interrater reliability could have been assessed. It is not known whether use of a different group of clinicians in completing the ratings would have produced similar results leading to the present conclusions. Future research might be directed toward obtaining a broader base of information.

Because of these limitations the results of this study can be taken as only suggestive; it is felt, nonetheless, that the data indicate that the computer-based Minnesota Report can be a useful aid to assessment in other countries to aid clinicians in interpreting MMPI-2 results. Even with these limitations, the results of the present study were very informative and encouraging with respect to the use of computer-based MMPI-2 reports in other countries. Overall, the reports were considered to provide a considerable amount of valuable information.

We turn next to illustrations of the use of the computer-based report in clinical settings from several other countries. We provide clinical case material from several other countries along with the MMPI-2 computerized report to provide the reader with an example of the kinds of information available in the report and its relevance for other countries.

10.12.3 CASES

10.12.3.1 Nijmegen, The Netherlands: The Case of Susan

In our outpatient psychological practice, patients with many types of complaints are treated, usually on a short-term basis. The mental health facility is integrated into the primary health care system and wide-ranging referrals are accepted. Patients usually come in for an assessment initially and treatment starts immediately. The mean number of sessions (including assessment) is 16 and the complaints include: anxiety disorders, mood disorders, somatoform disorders, adjustment disorders, family problems, and personality disorders. A group of eight clinical psychologists and two psychological assistants staff the facility. The practice is in a rural area near two cities, Nijmegen and Arnhem, and we work closely with 15 family practice physicians for referrals. Every week about six new patients are referred to the clinic. The treatment orientation is eclectic: we do behavioral therapy, rational emotive therapy, family therapy, short-term psychodynamic therapy, group therapy, and assertiveness training.

The case described here is a 23-year-old woman named Susan who came to our practice on referral from her physician. Initial impressions suggested an attractive young woman, a bit shy, but pleasant in the interpersonal sphere. Her chief complaints were obsessive ruminations and a feeling of uncertainty about what she wants to do and can do with her life. She expressed feelings of depression and feelings of inferiority which she connects with her negative self-image. A direct reason for asking for help could be found in her suicidal thoughts that made her panic and feel ashamed. She sleeps well, has no weight loss, and does not complain about fatigue. Besides the symptoms mentioned earlier, she has had an eating disorder and problems with the acceptance of her body which she considers too fat. Although she is slightly overweight at present, this is only to a minor degree.

The initial assessment did not result in a diagnosis of major depressive disorder and psychotic symptoms were lacking. She does show elements of dysthymia. Her personality organization, in terms of Kernberg criteria, is neurotic.

Susan was born in the eastern part of The Netherlands and about six months before she came for treatment she had moved to a rural area near the practice. This happened at the time that her boyfriend took over his father's business, a nursery for growing roses. Susan and her boyfriend started living together at that time and she began working in this family enterprise. Her chief complaints started when, in summer, they had a lot of work to do and lots of responsibilities; amongst other things she had to manage the staff. Especially when alone, she began to feel unhappy. She responded by trying to avoid these feelings as long as possible and then by binge eating. In addition she often had crying spells. She connected the binge eating with a need to pamper herself.

During the assessment phase her boyfriend was invited for a session and he explained that Susan had chosen to live with him but was not eager to work in the rose nursery. He himself has worked in this business since he was 16 years old and is aware that the work is demanding and stressful at times. He would like her to enjoy this work because he wants to share everything with her. The possibility of Susan choosing a different career path is felt by him as a threat.

Susan is the youngest in a family with two daughters. She reports some positive early childhood memories. She experienced her family as very well attached; many family activities were shared together. Her mother was, like Susan, rather introverted and also reports problems with accepting her body and difficulties with bodily contact. Susan never felt secure with her mother. She described her father as a warm and sensitive but somewhat prudish man. He works for an insurance company and was always busy and stayed away from home a lot. Her father used to call Susan the "sunshine in the house." Her mother was responsible for the children's upbringing and stopped working as a typist when she gave birth to the oldest. It is striking that, in the interview, when Susan made a critical comment about her parents she immediately covered it up with a positive one.

During puberty she had difficulties with the development of female secondary sex characteristics because she really wanted to be a boy. There were some precursors for this in the family because her father often treated her as though she was a boy. She was always quiet, introverted, well-adapted, nice, and sensitive but she never knew what to do with her life. Her sister, who is two years older, however, always manifested herself in a rebellious manner, refused to accept societal and family norms and values, and was extreme in her dress, for example she often dressed as a "punk." Her parents devoted a lot of attention to her older sister. Susan solved her problems herself and asked little from her parents.

During her adolescence, her mother criticized her quite often because of her weight and her tendency to be introverted. This criticism resulted in her developing some irritation and resentment. Her mother thought she was stubborn if Susan did not do what she asked her to do. During adolescence Susan always felt inferior to her sister and thought that her mother preferred her sister. When she was 16 she had a brief relationship with her first boyfriend; however, she refused sexual intercourse with him. Prior to this relationship her interest in boys was purely nonsexual and she was afraid that boys might take advantage of her. The relationship with her first boyfriend ended because he felt that she could not express any feelings toward him. She always wanted to be strong and tough but the end of the relationship was difficult for her and she never really worked through her feelings about it.

After school Susan attended a lower professional school (training as a hospital orderly) but stopped in the second year of her education because she was unable to maintain good interpersonal contact with co-workers and the patients. During this period, when she was 19 years old, she met her present boyfriend who is also her employer. She characterizes him as also being introverted. After he had known her for one-and-a-half months he asked her whether she would like to work in the rose nursery and she accepted his offer. She believed that their relationship would have ended if she had rejected his offer. Their relationship started very cautiously and the sexual part of it has been problematic from the beginning.

From the earliest days of their relationship she did not stand up for her needs and wants. In addition, because she was afraid of losing him, she demanded perfection from herself. She feels that she must be steady and very adaptable to the demands of her boyfriend's family. She also feels quite stressed over the great demands placed on her by the work in the nursery. Part of her difficulty comes from the dual relationship she has with her employer. Her boyfriend is also her boss and he often makes administrative decisions about, for instance, when she can have a day off. She often becomes emotional when this complex relationship is discussed and she tends to protect him immediately indicating that "he is so patient with her."

Her binge eating usually becomes manifest when she is alone. She often stores candies and then eats them all at once. She feels ashamed over her behavior and was especially guilty when her boyfriend discovered her problem. One major cause of hurt for her was when her boyfriend once told her that he was not primarily attracted to her body but rather her character. She told her mother, her best friend and her sister about her problems and was surprised that they were so supportive of her.

During the evaluation and also later in therapy she impressed the psychologist as playing the role of a victim in the mixed relationships in which she was involved with her boyfriend and her own family. Everyone seems to know what was best for her and she had borne these burdens but was feeling unhappy now. She seemed unable to assert herself with others and to express her own anger in a productive way.

The following psychological tests were administered as well as the MMPI-2. On the Wechsler Adult Intelligence Scale she had a verbal IQ of 103 and a performance IQ of 104 which is in agreement with her school education. On the "house–tree–person" test she drew a person with only a head. It seems as if her thinking is most important and she likes to control things rationally. Several other projective techniques were administered, such as the Thematic Apperception Test and the Sentence Completion Test. Themes that appeared prominent were her sexual problems, hidden aggression in relation to her parents, idealized pictures of other people, not knowing where to go in life, and a lack of perseverance.

10.12.3.1.1 The Minnesota Report on Susan

On the MMPI-2 she produced the profiles shown in Figures 5 and 6 and the following interpretation.

Profile validity

This is a valid MMPI profile. The client's attitude toward the testing was appropriate. She responded to the items in a frank and open manner, freely admitting to some psychological problems, which are described below.

Symptomatic patterns

The behavioral correlates included in the narrative report are likely to provide a good description of the client's current personality functioning. The clinical scale prototype used in the report, which incorporates correlates of D and Pt, is based on scores with high profile definition. The client's MMPI-2 clinical profile reflects much psychological distress at this time. She has major problems with anxiety and depression. She tends to be high-strung and insecure, and she may also be having somatic problems. She is probably experiencing loss of sleep and appetite and a slowness in personal tempo.

Individuals with this profile often have high standards and a strong need to achieve, but they feel that they fall short of their expectations and then blame themselves harshly. This client feels quite insecure and pessimistic about the future.

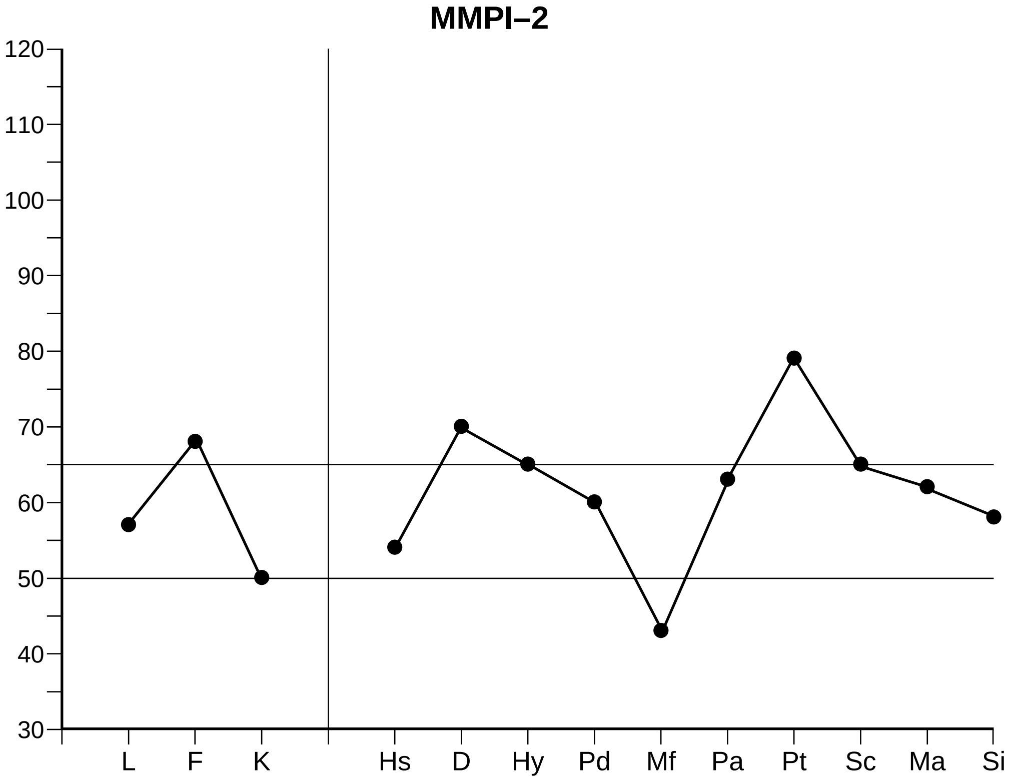

	L	F	K	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma	Si
Raw score	5	9	15	7	30	29	21	39	14	28	20	21	34
K correction				8			8			15	15	3	
T score	57	68	50	54	70	65	60	43	63	79	65	62	58

Figure 5 MMPI-2 profile of Susan (1).

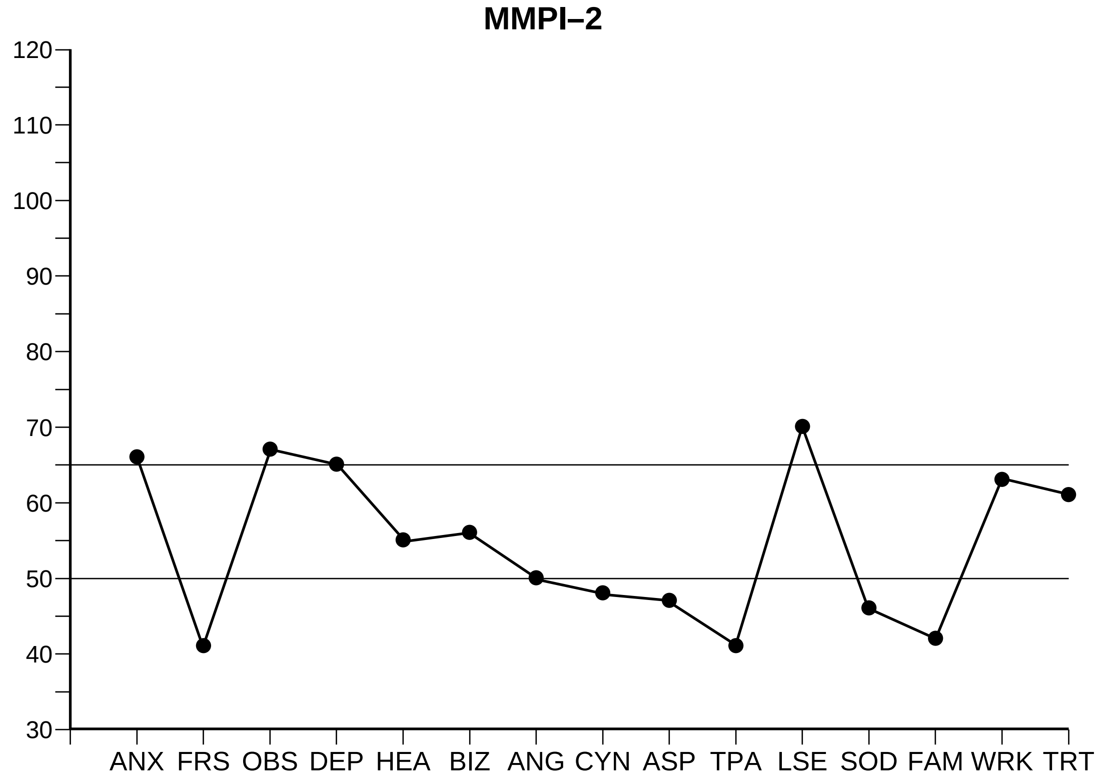

	ANX	FRS	OBS	DEP	HEA	BIZ	ANG	CYN	ASP	TPA	LSE	SOD	FAM	WRK	TRT
Raw score	14	3	11	14	8	3	6	8	5	4	14	5	3	16	10
T score	66	41	67	65	55	56	50	48	47	41	70	46	42	63	61

Figure 6 MMPI-2 profile of Susan (2).

She also feels quite inferior, has little self-confidence, and does not feel capable of solving her problems.

In addition, the following description is suggested by the content of the client's item responses. She has endorsed a number of items suggesting that she is experiencing low morale and a depressed mood. She reports a preoccupation with feeling guilty and unworthy. She feels that she deserves to be punished for wrongs she has committed. She feels regretful and unhappy about life, and she seems plagued by anxiety and worry about the future. She feels hopelessness at times and feels that she is a condemned person. She views her physical health as failing and reports numerous somatic concerns. She feels that life is no longer worthwhile and that she is losing control of her thought processes.

According to her response content, there is a strong possibility that she has seriously contemplated suicide. The client's recent thinking is likely to be characterized by obsessiveness and indecision.

Profile frequency

It is usually valuable in MMPI-2 clinical profile interpretation to consider the relative frequency of a given profile pattern in various settings. The client's MMPI-2 high-point clinical scale score (Pt) is found in only 5.1% of the MMPI-2 normative sample of women. Only 2.1% of the women have Pt as the peak score at or above a T score of 65, and less than 1% have well-defined Pt spikes. This high MMPI-2 profile configuration (2–7/7–2) is very rare in samples of normals, occurring in 1.2% of the MMPI-2 normative sample of women.

Her MMPI-2 high-point score is relatively uncommon in various outpatient settings. In the National Computer Systems (NCS) outpatient sample, only 6.4% of the females have this MMPI-2 high-point clinical scale score (Pt). Moreover, only 5.3% of the female outpatients have the Pt scale spike at or above a T score of 65, and only 2.1% have well-defined Pt peaks. Her elevated MMPI-2 profile configuration (2–7/7–2) is found in 7.4% of the women in the NCS outpatient sample. This is the second most frequent elevated two-point code in the sample.

Profile stability

The relative scale elevation of the highest scales in her clinical profile reflects high profile definition. If she is retested at a later date, the peak scores on this test are likely to retain their relative salience in her retest profile pattern. Her high-point score on Pt is likely to remain stable over time. Short-term test–retest studies have shown a correlation of 0.88 for this high-point score.

Interpersonal relations

She appears to be quite passive and dependent in interpersonal relationships and does not speak up for herself even when others take advantage of her. She avoids confrontation and seeks nurturance from others, often at the price of her own independence. She forms deep emotional attachments and tends to be quite vulnerable to being hurt. She also tends to blame herself for interpersonal problems. Individuals with this profile are often experiencing psychological distress in response to stressful events. The intense feelings may diminish over time or with treatment.

Diagnostic considerations

Individuals with this profile tend to be considered neurotic and receive diagnoses such as Dysthymic Disorder or Anxiety Disorder. They may also receive Axis II diagnosis of Dependent or Compulsive Personality Disorder

Treatment considerations

Individuals with this MMPI-2 pattern are usually feeling a great deal of discomfort and tend to want help for their psychological problems. The client's self-esteem is low and she tends to blame herself too much for her difficulties. Although she worries a great deal about her problems, she seems to have little energy left over for action to resolve them.

Symptomatic relief for her depression may be provided by antidepressant medication. Psychotherapy, particularly cognitive behavioral treatment, may also be beneficial.

The passive, unassertive personality style that seems to underlie this disorder might be a focus of behavior change. Individuals with these problems may learn to deal with others more effectively through assertiveness training.

If psychological treatment is being considered, it may be profitable for the therapist to explore the client's treatment motivation early in therapy. The item content she endorsed includes some feelings and attitudes that could be unproductive in psychological treatment and in implementing change.

Note. This MMPI-2 interpretation can serve as a useful source of hypotheses about clients. This report is based on objectively derived scale indices and scale interpretations that have been developed in diverse groups of patients. The personality descriptions, inferences, recommendations contained herein need to be verified by other sources of clinical information because individual clients may not fully match the prototype. The information in this report should most appropriately be used by a trained, qualified test interpreter. The information contained in this report should be considered confidential.

10.12.3.1.2 Comment

Susan produced a valid profile. The higest scale is scale 7 (79), followed by scale 2 (70), and 3 (65). The profile is essentially a neurotic one; she seems to be controlling feelings by way of rationalizing and ruminating. Her subjective suffering is clearly present (scale 2). Scale 7 assesses abnormal fears, self-criticism, difficulties in concentration and guilt feelings. Further, the item content reflects a characterologic basis for a wide variety of psychasthenic symptoms. Anxiety is a persistent trait. This seems in agreement with her complaints and attitudes. On the content scales her highest scale score was on Low Self Esteem (70) and her second highest was Obsessions (67) and then Anxiety (66). Depression is 65 and Problems in Relation to Work 63.

10.12.3.1.3 DSM-IV classification

This was as follows:
Axis I: 313.82 Identity Problem
307.50 Eating Disorder Not Otherwise Specified
Axis II: 301.60 Traits of the Dependent and Obsessive-Compulsive Personality Disorder.

10.12.3.1.4 Treatment

Individual insight-oriented sessions were initiated with Susan to give her the opportunity to explore her needs and wishes in relation to her work, study, sexual behavior, and partner choice. Although her intelligence is only average, she is well motivated to work things out for herself. An attempt has been made to focus on her conflicts between dependency and autonomy. The aim is to facilitate the process of individualization and to provide her with the opportunity to discover her feelings of anger that she holds toward others close to her.

Treatment progress has thus far been effective. During the course of her treatment she has gradually discovered better her own needs and preferences. She has started to understand her ambivalent feelings towards her parents and also towards her boyfriend. At the present time in therapy she is exploring her fantasies of leaving him and starting a different life. Her characterologic basis for psychasthenic traits, indicated by the MMPI-2, has appeared very strong and difficult to influence so far.

10.12.3.2 Rome, Italy: The Case of Mario P.

Mr. P. is a 36-year-old male from a low socioeconomic background and educational level. At the age of 21 he married an 18-year-old woman and they now have two children aged 14 and six. He was referred to our outpatient psychiatry service by an otolaryngologist whom he consulted for a "lump in his throat" when his physician could find no organic basis to his symptoms. He has a family history of depression: his mother suffered from depression in the past and one of his three sisters is presently depressed. His wife, who recently left her job to dedicate herself to housekeeping, had suffered in the past from what seems to be somatic complaints of depression.

Until recently the patient was working as a manual laborer until 2.00 a.m. every night because he held two jobs. However, he had to give up one of his jobs three months ago, just prior his first visit to the clinic, because he developed nervousness, apathy, tension, and irritability. At this visit he reported poor sleep quality, reduced appetite, and excessive increase in cigarette smoking. He lacked a drive for living; he could not face life either at home or at work and saw no "way-out." He felt guilty about not being able to care for his children, he seemed discouraged and "switched-off." His impulse control is poor; he reported several acts of violence (having beaten his children). Fearing the consequences of his impulsive behavior and being afraid of his lack of control over his acts, he often leaves his home for hours, walking with no precise destination, seeking isolation to escape from difficulties. He was prescribed 1.5 mg/die oral desmethyldiazepam and took this medication for some time but had to reduce it to 0.5 mg/die due to the development of marked sedation that impaired his daytime performance.

During the interview, the patient spoke spontaneously, with a monotonous, low voice. Speech content and form were normal, however, and he expressed guilt, pessimism, and inadequacy. He showed no evidence of free-floating anxiety. His mood was markedly depressed and he frequently sighed aloud. He appeared to have good insight with respect to his psychopathological problems and was willing to accept treatment.

10.12.3.2.1 The Minnesota Report on Mario P.

The patient completed the MMPI-2 at the end of his first visit to the clinic producing the profiles shown in Figures 7 and 8 and the following narrative report.

Profile validity

The client responded to the MMPI-2 items in an unusual manner. He claimed an unrealistic amount of virtue while also endorsing a great number of psychological difficulties. This infrequent response pattern reflects some unconventional and possibly bizarre beliefs.

Careful evaluation of the individual's response attitudes should be undertaken to explain this unusual validity scale pattern. The following hypotheses might be explored. He may have consciously distorted the test responses to create a particular impression, or he may be generally unsophisticated. The resulting clinical pattern should be interpreted with caution.

The client's responses to items in the latter portion of the MMPI-2 were somewhat exaggerated in comparison to his responses to earlier items. There is some possibility that he became more careless in responding to these later items, thereby raising questions about that portion of the test. Although the standard validity and clinical scales are scored from items in the first two-thirds of the test, caution should be taken in interpreting the MMPI-2 Content Scales and supplementary scales, which include items found throughout the entire pool.

Symptomatic patterns

This report was developed using the D, Pt, and Sc scales as the prototype. A pattern of chronic psychological maladjustment characterizes individuals with this MMPI-2 clinical profile. The client is overwhelmed by anxiety, tension, and depression. He feels helpless and alone, inadequate and insecure, and he believes that life is hopeless and that nothing is working out right. He attempts to control his worries through intellectualization and unproductive self-analysis, but he has difficulty concentrating and making decisions. This is a rather chronic behavior pattern. Individuals with this profile typically live a disorganized and pervasively unhappy existence. They may have episodes of more intense and disturbed behavior resulting from an elevated stress level.

He is functioning at a lower level of efficiency. He tends to overreact to even minor stress, and he may show rapid behavioral deterioration. He also tends to blame himself for his problems. His lifestyle is chaotic and disorganized, and he has a history of poor work and achievement. He may be preoccupied with obscure religious ideas.

The client seems to have a rather limited range of cultural interests and tends to prefer stereotyped masculine activities to literary and artistic pursuits or introspective experiences. Interpersonally, he may be somewhat intolerant and insensitive.

Profile frequency

It is usually valuable in MMPI-2 clinical profile interpretation to consider the relative frequency of a given profile pattern in various settings. The client's MMPI-2 high-point clinical scale score (Pt) is found in only 4.9% of the MMPI-2 normative sample for men. Only 3.1% of the sample have Pt as the peak score at or above a T score of 65, and only 1.6% have well-defined Pt spikes. This

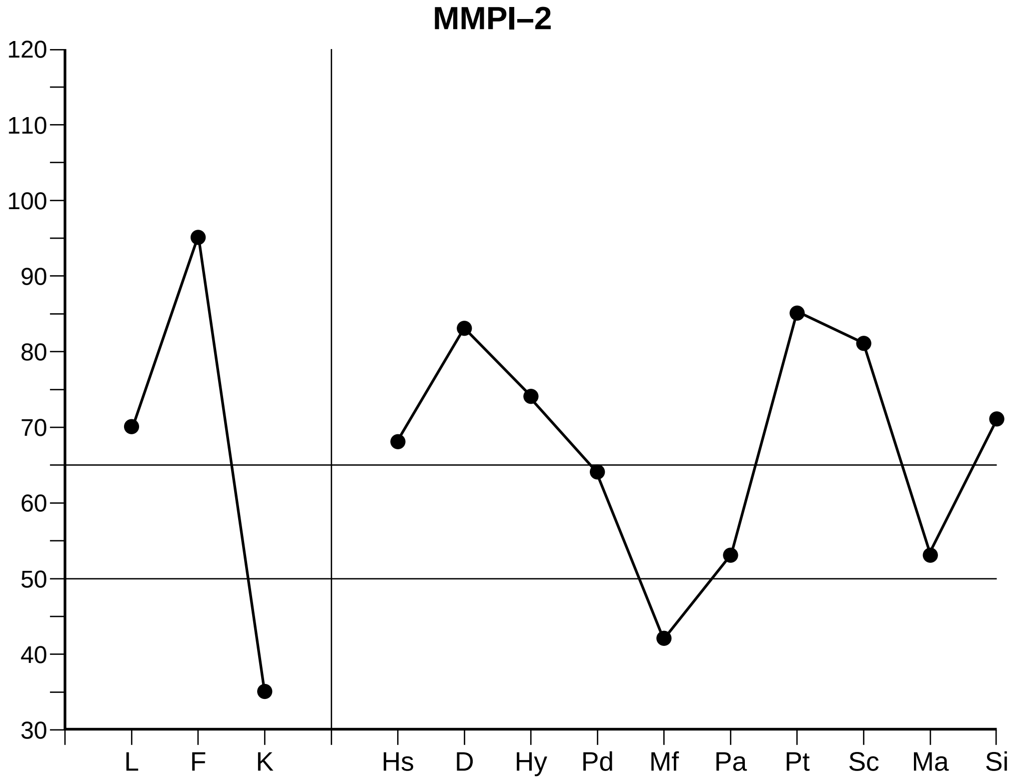

	L	F	K	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma	Si
Raw score	8	19	8	16	35	31	26	22	11	35	36	20	43
K correction				4			3			8	8	2	
T score	70	95	35	68	83	74	64	42	53	85	81	53	71

Figure 7 MMPI-2 profile of Mario P. (1).

elevated MMPI-2 profile configuration (2–7/7–2) is very rare in samples of normals, occurring in less than 1% of the MMPI-2 normative sample of men.

The relative frequency of this MMPI-2 high-point score is informative. In the NCS outpatient sample, 7.7% of the males have this MMPI-2 high-point clinical scale score (Pt). Moreover, 6.5% of the male outpatients have the Pt scale spike at or above a T score of 65, and 3.3% have well-defined Pt spike scores in that range. His elevated MMPI-2 profile configuration (2–7/7–2) is relatively common in outpatient men. It occurs in 4.8% of the men in the NCS outpatient sample. The 2–7 profile code is the second most frequent two-point code in outpatient men when both scales are at or above a T score of 65.

He scored relatively high on MAC-R, suggesting the possibility of a drug or alcohol abuse problem. The base rate data on his profile type among residents in alcohol and drug programs should also be evaluated. His MMPI-2 profile code, including D and Pt, is the second most frequent two-point code among men in alcohol- and drug-abusing populations. Over 13% of the men in substance-abuse programs have this pattern (McKenna & Butcher, 1987).

Profile stability

The relative scale elevation of the highest scales in his clinical profile reflects high profile definition. If he is retested at a later date, the peak scores on this test are likely to retain their relative salience in his retest profile pattern. His high-point score on Pt is likely to show high stability over time. Short-term test–retest studies have shown a correlation of 0.89 for this high-point score. Spiro, Butcher, Levenson, Aldwin, and Bosse (1993) reported a test–retest stability index of 0.65 in a large study of normals over a five-year test–retest period.

Interpersonal relations

Problematic personal relationships are also characteristic of such clients. He seems to lack basic social skills and is behaviorally withdrawn. He may relate to others ambivalently, never fully trusting or loving anyone. Many individuals with this profile never establish lasting, intimate relationships. His marital situation is likely to be unrewarding and impoverished. He seems to feel inadequate and insecure in his marriage.

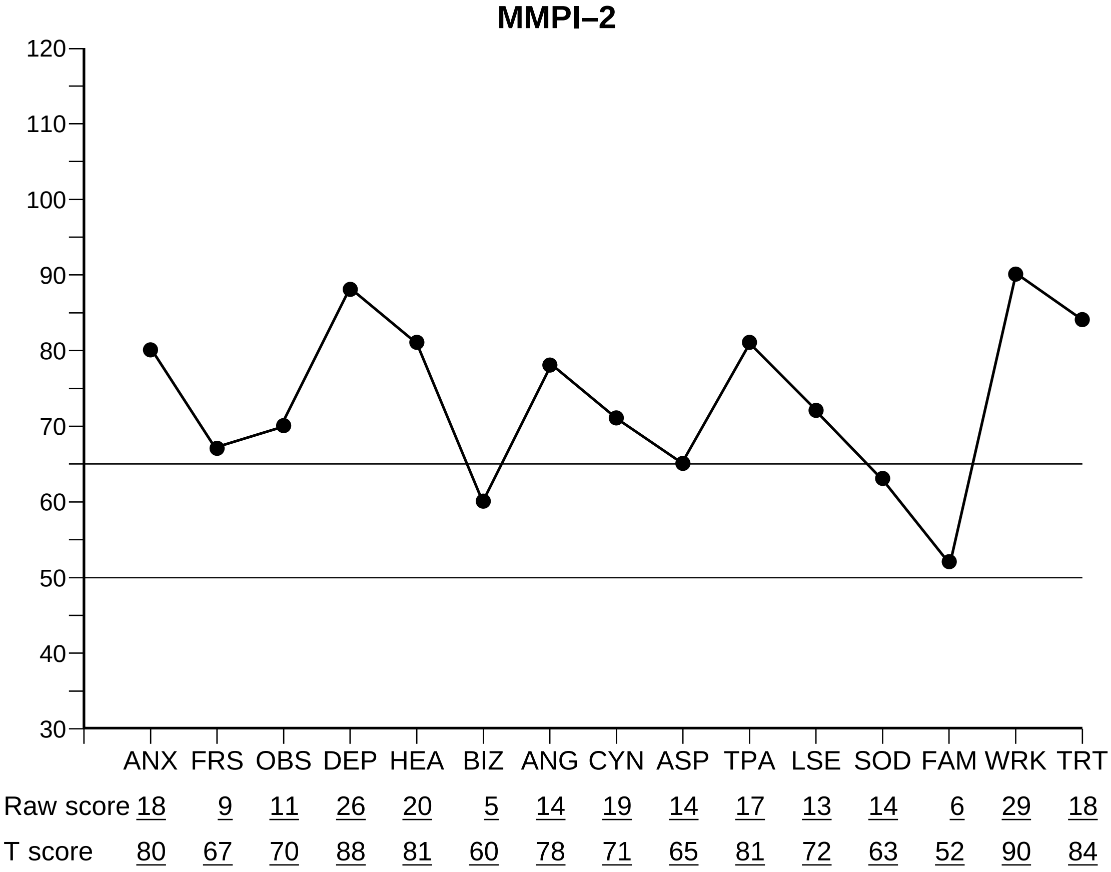

Figure 8 MMPI-2 profile of Mario P. (2).

He is a very introverted person who has difficulty meeting and interacting with other people. He is shy and emotionally distant. He tends to be very uneasy, rigid, and overcontrolled in social situations. His shyness is probably symptomatic of a broader pattern of social withdrawal. Personality characteristics related to social introversion tend to be stable over time. His generally reclusive behavior, introverted lifestyle, and tendency toward interpersonal avoidance may be prominent in any future test results.

His very high score on the Marital Distress Scale suggests that his marital situation is quite problematic.

Diagnostic considerations

Individuals with this profile have a severe psychological disorder and would probably be diagnosed as severely neurotic with an Anxiety Disorder or Dysthymic Disorder in a Schizoid Personality. The possibility of a more severe psychotic disorder, such as Schizophrenic Disorder, should also be considered, however.

His extremely high scores on the addiction proneness indicators suggest the possible development of an addictive disorder. Further evaluation of substance use or abuse problems is strongly recommended.

Treatment considerations

Individuals with this MMPI-2 clinical profile often receive psychotropic medications for their depressed mood or intense anxiety. Many individuals with this profile seek and require psychological treatment for their problems along with any medication that is given. Because many of their problems tend to be chronic, an intensive therapeutic effort might be required in order to bring about any significant change. Patients with this profile typically have many psychological and situational concerns; consequently, it is often difficult to maintain a focus in treatment.

He probably needs a great deal of emotional support at this time. His low self-esteem and feelings of inadequacy make it difficult for him to get energized toward therapeutic action. His expectation for positive change in therapy may be low. Instilling a positive, treatment-expectant attitude is important for him if treatment is to be successful.

Individuals with this profile tend to be overideational and given to unproductive rumination.

They tend not to do well in unstructured, insight-oriented therapy and may actually deteriorate in functioning if they are asked to be introspective. He might respond more to supportive treatment of a directive, goal-oriented type. Individuals with this profile present a clear suicide risk; precautions should be taken.

Note. This MMPI-2 interpretation can serve as a useful source of hypotheses about clients. This report is based on objectively derived scale indices and scale interpretations that have been developed in diverse groups of patients. The personality descriptions, inferences, recommendations contained herein need to be verified by other sources of clinical information because individual clients may not fully match the prototype. The information in this report should most appropriately be used by a trained, qualified test interpreter. The information contained in this report should be considered confidential.

10.12.3.2.2 Comment

Mario P. obtained high scores on the D scale, compatible with a depressive state, high scores on the Pt and Sc scales, compatible with the symptoms of irritability, anxiety, avoidance, and tension, and marginally high social introversion that can explain his pursuit of isolation. His scores on the Pd scale reflect his violent acts and impulsiveness. Higher than average scores on the Hs and Hy scales are consistent with the presence of the lump in the patient's throat (globus hystericus) and with the tendency to use bodily symptoms to resolve his conflicts. The mood axis is oriented towards depression. The presence of an 8–7–2 codetype reflects his pessimism, weakness and fatigue, lack of initiative, perception of loneliness, poor description of his own emotional state, self-blame, worthlessness, and impulsiveness. The personality profile, with the T scores of all neurotic scales being over 65, was compatible with the presence of a neurotic disorder, and the diagnosis was made of severe, single episode major depression.

10.12.3.2.3 Treatment

The patient was started on oral chlomipramine, 100 mg/die, titrated upwards. This drug is marketed in Europe as an antidepressant; it is also used for obsessive-compulsive disorder and other disorders of the obsessive-compulsive spectrum, including those where impulsive behavior predominates. The patient returned regularly every 2–4 weeks for his follow-up visits. At the second visit, the patient was switched to slow-release chlomipramine, 150 mg/die, titrated upwards to 187.5 mg/die. Sodium valproate, 1 mg/die, was added to control impulsiveness. He reported outbursts of aggression and impulsiveness, with mood swings during the day, but overall, he felt slightly better. By the third visit, one month after the beginning of drug treatment, the patient started to report better control of his impulses and satisfactory mood improvement. He further improved with time, returning to work after three months of therapy. By this time, the quality of sleep had improved and his appetite had increased, especially for carbohydrates. Impulsiveness and dysthymic mood were no longer reported.

He developed a fine tremor that subsided the next month. After four months of treatment, his mood was stabilized and the patient was more active. He returned to work but still felt the presence of a lump in his throat, although only occasionally. About this time he developed hypersomnia. He tried to suspend treatment for a week or so, but his symptoms worsened. He had his chlomipramine reduced to 112.5 mg/die and valproate to 500 mg. Improvement continued and after seven months the patient seems to have completely recovered and is maintained on the above drug combination.

10.12.3.3 Iran, United States: The Case of Ms. B.

The patient is a 30-year-old female Iranian, who is currently separated. She has an AA degree in fine arts and is currently unemployed. She was self-referred for therapy. Ms. B.'s initial complaints included a wide range of psychosomatic problems including headaches, dizziness, weakness, sore throat, pressure in the chest cavity, and upset stomach. She added that she felt sad, anxious, and angry and worried that she might lose her mind. Ms. B. indicated that no one understands or accepts her. This feeling generates a tremendously high level of stress, sadness, and anger. As a result of the lack of understanding, she feels lonely and tries to justify herself and goes through lengthy explanations to prove that she really is a good woman and worthy of love. This urge to prove her virtues is dated back to the time that her family found out about her sexual intimacy with an older man when she was 16 years old. The patient got herpes from this two-year relationship. Ms. B. considers herself a victim of rape in this relationship and believes that the trauma and the consequent problems continue to have an impact on her present life.

Ms. B. stated that her parents had a turbulent marriage. According to the patient her mother married her father because of coercion, not love.

She added that her mother was depressed when she was pregnant with the patient. During her childhood she witnessed family discord and conflicts with repeated discussions of divorce. According to the patient her father used to beat her mother. Consequently, the patient would stay home to provide some sort of protection for her mother. Her parents were separated for a couple of years but never got divorced. Her mother repeatedly told her that she did not get a divorce because of her and blamed the patient for her unhappiness.

The patient's emotional status has deteriorated following early separation after a short-lived marriage a couple of months before the present testing session. Ms. B. has been experiencing more problems, including depression and frustration, following her separation.

The patient denied a history of previous psychiatric problems, but later mentioned that she had received short-term counseling in the past. The patient has been diagnosed with herpes since the age of 16. No other medical problems are indicated. Ms. B. denies a history of substance abuse.

The patient is an Iranian and the oldest of three children. One of her younger brothers has been diagnosed with schizophrenia. The patient denied any history of other psychiatric problems in her family. She also denied any history of alcohol or substance abuse in her family. However, during therapy the presence of personality problems and antisocial characteristics among the family members was detected. As a child, she reported being emotionally abused. Her parents, especially her mother, blamed her for their unhappy marriage and she witnessed repeated family fights as she was growing up. Her mother repeatedly told her that she was an unwanted child.

The patient reported having sex with an older man at age 16. She remained intimate with this man for a couple of years. When her family found out about her relationship they severely chastised her and told her that she was not suitable for marriage any longer. She has always had difficulty in maintaining relationships with men, particularly when there was a hint that it might become serious. At age 29, the patient was married for a very short period of time. This

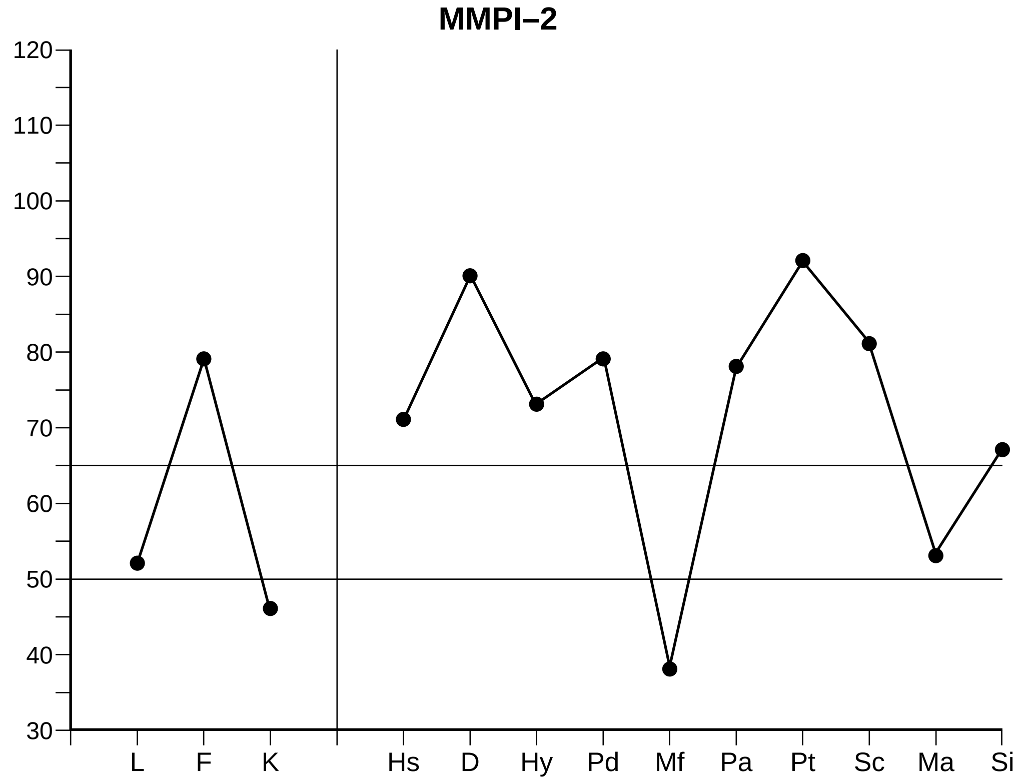

	L	F	K	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma	Si
Raw score	4	12	13	16	39	32	29	41	18	37	33	18	43
K correction				7			5			13	13	3	
T score	52	79	46	71	90	73	79	38	78	92	81	53	67

Figure 9 MMPI-2 profile of Ms. B. (1).

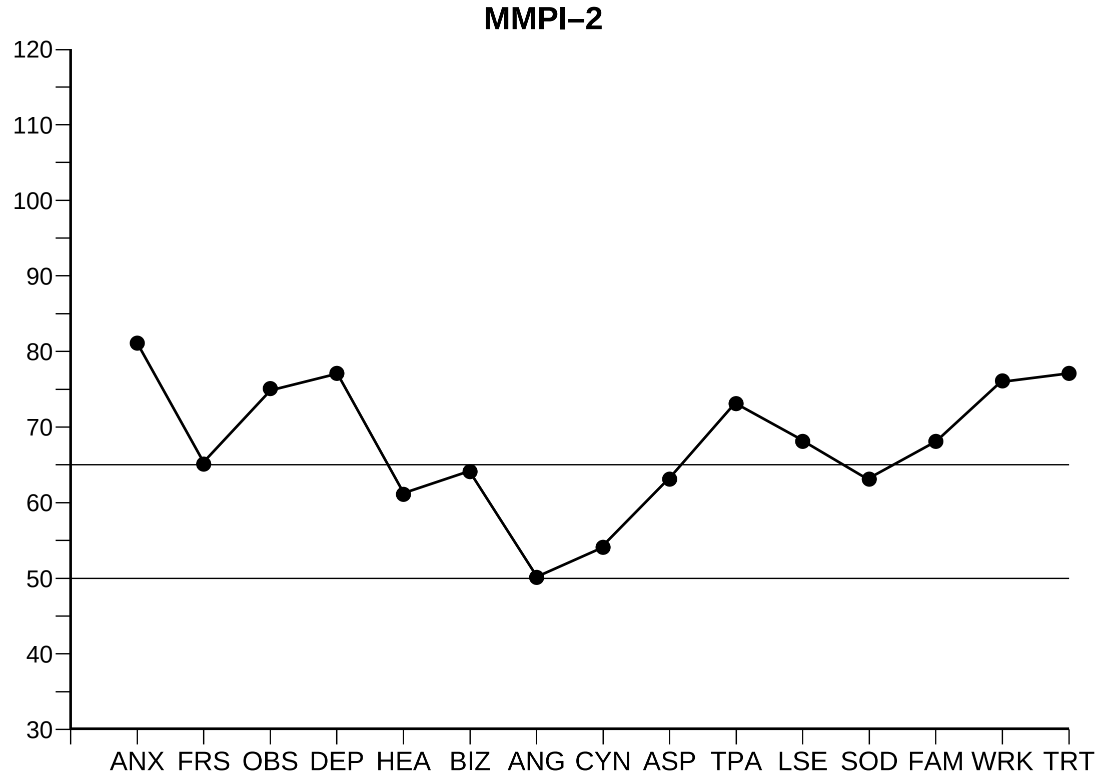

	ANX	FRS	OBS	DEP	HEA	BIZ	ANG	CYN	ASP	TPA	LSE	SOD	FAM	WRK	TRT
Raw score	20	12	13	21	11	6	6	12	11	14	13	14	13	22	16
T score	81	65	75	77	61	64	50	54	63	73	68	63	68	76	77

Figure 10 MMPI-2 profile of Ms. B. (2).

short marriage was marked by mistrust, ambivalence, verbal abuse, and recurrent arguments, resulting in a separation.

The assessment was conducted in an outpatient setting. At the time of testing, Ms. B. presented as well-groomed and fashionably dressed. She was oriented and her speech was fluent and intelligible. No suicidal or homicidal thoughts were indicated. She denied having any hallucinations and no delusions were noted. Remote memory and immediate memory for recall were intact.

10.12.3.3.1 The Minnesota Report on Ms. B.

When the MMPI-2 was administered, Ms. B. produced the profiles shown in Figures 9 and 10 and the following narrative report.

Profile validity

This client has endorsed a number of psychological problems, suggesting that she is experiencing a high degree of stress. Although the MMPI-2 clinical scale profile is probably valid, it may show some exaggeration of symptoms.

She endorsed the items at the end of the booklet in an extreme or exaggerated manner, producing a high score on F(B). This elevated score could result from a number of conditions such as confusion, exaggerated symptom checking, or consistently misrecording her responses on the answer sheet. The scores on the MMPI-2 Content Scales, supplementary scales, and content component scales could be influenced by this tendency.

Symptomatic patterns

This report was developed using the D, Pt, and Sc scales as the prototype. A pattern of chronic psychological maladjustment characterizes individuals with this MMPI-2 clinical profile. The client is overwhelmed by anxiety, tension, and depression. She feels helpless and alone, inadequate and insecure, and she believes that life is hopeless and that nothing is working out right. She attempts to control her worries through intellectualization and unproductive self-analysis, but she has difficulty concentrating and making decisions. This is a rather chronic behavioral pattern. Individuals with this profile typically live a disorganized and pervasively unhappy existence. They may have episodes of more intense and disturbed behavior resulting from an elevated stress level.

She is functioning at a very low level of efficiency. She tends to overreact to even minor stress, and she may show rapid behavioral deterioration. She also tends to blame herself for her problems. Her lifestyle is chaotic and disorganized, and she has a history of poor work and achievement. She may be preoccupied with obscure religious ideas.

The client does not report great concerns about her sex-role identity. However, the extent to which her rather "traditionally feminine" feelings and attitudes are integrated into her life at this point may need to be evaluated further.

In addition, the following description is suggested by the content of the client's item responses. She has endorsed a number of items suggesting that she is experiencing low morale and a depressed mood. She reports a preoccupation with feeling guilty and unworthy. She feels that she deserves to be punished for wrongs she has committed. She feels regretful and unhappy about life, and she seems plagued by anxiety and worry about the future. She feels hopeless at times and feels that she is a condemned person. She endorsed response content that reflects low self-esteem and long-standing beliefs about her inadequacy. She views her physical health as failing and reports numerous somatic concerns. She feels that life is no longer worthwhile and that she is losing control of her thought processes.

According to her response content, there is a strong possibility that she has seriously contemplated suicide. The client's recent thinking is likely to be characterized by obsessiveness and indecision. She feels somewhat self-alienated and expresses some personal misgivings or a vague sense of remorse about past acts. She feels that life is unrewarding and dull, and she finds it hard to settle down. She views the world as a threatening place, sees herself as having been unjustly blamed for others' problems, and feels that she is getting a raw deal out of life. She is rather high-strung and believes that she feels things more, or more intensely, than others do. She feels quite lonely and misunderstood at times. The client attests to having more fears than most people do. Her high endorsement of general anxiety content is likely to be important to understanding her clinical picture.

Profile frequency

It is usually valuable in MMPI-2 clinical profile interpretation to consider the relative frequency of a given profile pattern in various settings. The client's MMPI-2 high-point clinical scale score (Pt) is found in only 5.1% of the MMPI-2 normative sample of women. Only 2.1% of the women have Pt as the peak score at or above a T score of 65, and less than 1% have well-defined Pt spikes. This high profile configuration (2–7/7–2) is very rare in samples of normals, occurring in 1.2% of the MMPI-2 normative sample of women.

Her MMPI-2 high-point score is relatively uncommon in various outpatient settings. In the NCS outpatient sample, only 6.4% of the females have this MMPI-2 high-point clinical scale score (Pt). Moreover, only 5.3% of the female outpatients have the Pt scale spike at or above a T score of 65, and only 2.1% have well-defined Pt peaks. Her elevated MMPI-2 profile configuration (2–7/7–2) is found in 7.4% of the women in the NCS outpatient sample. This is the second most frequent elevated two-point code in the sample.

Profile stability

The relative elevation of her clinical scale scores suggests that her profile is not as well defined as many other profiles. That is, her highest scale or scales are very close to her next scale score elevations. There could be some shifting of the most prominent scale elevations in the profile code if she is retested at a later date. The difference between the profile type used to develop the present report and the next highest scale in the profile code was 2 points. So, for example, if the client is tested at a later date, her profile might involve more behavioral elements related to elevations on Pd. If so, then on retesting, acting-out, aggressive, and irresponsible behavior might become more pre-eminent.

Interpersonal relationships

Problematic personal relationships are also characteristic of such clients. She seems to lack basic social skills and is behaviorally withdrawn. She may relate to others ambivalently, never fully trusting or loving anyone. Many individuals with this profile never establish lasting, intimate relationships.

She is quite shy and inhibited in social situations, and she may avoid others for fear of being hurt. She is emotionally alienated from others. She is likely to have very few friends and to be considered by others as distant and hard to get to know. She is quiet and submissive and lacks self-confidence in dealing with other people. Individuals with this passive and withdrawing lifestyle are unable to assert themselves appropriately and are frequently taken advantage of by others. Personality characteristics related to social introversion tend to be stable over time. Her generally reclusive behavior, introverted lifestyle, and tendency toward interpersonal avoidance may be prominent in any future test results.

Her very high score on the Marital Distress Scale suggests that her marital situation is quite problematic at this time. She has reported a number of problems with her marriage that are possibly important to understanding her current psychological symptoms.

Diagnostic considerations

Individuals with this profile have a severe psychological disorder and would probably be diagnosed as severely neurotic with an Anxiety Disorder or Dysthymic Disorder in a Schizoid Personality. The possibility of a more severe psychotic disorder, such as Schizophrenic Disorder, should also be considered, however. Her self-reported tendency toward experiencing a depressed mood should be taken into consideration in any diagnostic formulation.

Treatment considerations

Individuals with this MMPI-2 clinical profile often receive psychotropic medications for their depressed mood or intense anxiety. Many individuals with this profile seek and require psychological treatment for their problems along with any medication that is given. Because many of their problems tend to be chronic, an intensive therapeutic effort might be required in order to bring about any significant change. Patients with this profile typically have many psychological and situational concerns; consequently, it is often difficult to maintain a focus in treatment.

She probably needs a great deal of emotional support at this time. Her low self-esteem and feelings of inadequacy make it difficult for her to get energized toward therapeutic action. Her expectation for positive change in therapy may be low. Instilling a positive, treatment-expectant attitude is important for her if treatment is to be successful.

Individuals with this profile tend to be overideational and given to unproductive rumination. They tend not to do well in unstructured, insight-oriented therapy and may actually deteriorate in functioning if they are asked to be introspective. She might respond more to supportive treatment of a directive, goal-oriented type. Individuals with this profile present a clear suicide risk; precautions should be taken.

The client endorsed item content that seems to indicate low potential for change. She may feel that her problems are not addressable through therapy and that she is not likely to benefit much from psychological treatment at this time. Her apparently negative treatment attitudes may need to be explored early in therapy if treatment is to be successful.

Her item content suggests some family conflicts that are causing her considerable concern at this time. She feels unhappy about her life and resents having an unpleasant home life. Psychological intervention could profitably focus, in part, on clarifying her feelings about her family.

In any intervention or psychological evaluation program involving occupational adjustment, her negative work attitudes could become an important problem to overcome. She has a number of attitudes and feelings that could interfere with work adjustment.

Note. This MMPI-2 interpretation can serve as a useful source of hypotheses about clients. This report is based on objectively derived scale indices and scale interpretations that have been developed in diverse groups of patients. The personality descriptions, inferences, recommendations contained herein need to be verified by other sources of clinical information because individual clients may not fully match the prototype. The information in this report should most appropriately be used by a trained, qualified test interpreter. The information contained in this report should be considered confidential.

Ms. B. presents with a long-lasting discontent with her life accompanied by mild depression which has worsened following her separation. She reports numerous physical complaints which might indicate somatization in the face of stress or health concerns in accordance with a high level of general anxiety and the belief that nothing is working for her. While depressed the patient has low self-esteem and feels hopeless. According to the patient, she was emotionally abused as a child and sexually abused at the age of 16. She feels tremendous guilt and a need to explain the past. There is a strong need for approval and unconditional love. She blames herself for her parents' unhappiness in addition to bringing shame on her family. While blaming herself for everything that has gone wrong in her parents' life as well as her own, she feels helpless and insecure in the face of any challenges or problems. Following her separation, she reports that life is full of difficulties and devoid of any happiness or reinforcers.

10.12.3.3.2 Diagnostic impression

According to the psychiatrist, Ms. B.'s clinical diagnosis, at least for insurance purposes, was considered to be major depression.
Axis I: Major Depression
Axis II: Mixed Personality (borderline, passive-aggressive, narcissistic, antisocial, histrionic, and dependent)
Axis III: Herpes.

10.12.3.3.3 Treatment

Ms. B. was in therapy for a total of three years. She was prescribed Prozac, 40 mg, which was later reduced to 20 mg. Her treatment plan consisted of weekly individual therapy sessions to address current stress due to her separation and divorce proceedings. Supportive therapy in a safe environment was provided to allow her to discuss issues related to her childhood trauma and acceptance of the past. According to the therapist, Ms. B. was always demanding and never satisfied with anything including therapy.

10.12.3.4 Seoul, Korea: The Case of Kim

Kim is a 36-year-old unemployed, unmarried Korean man with two years of college education. Kim has been obsessed with suicidal ideas (disembowelment) since 1992. He feels extremely guilty and depressed. Kim feels that he has to die, but he doesn't know what to blame for his problems.

In 1987, at the age of 27, Kim went to Australia for his college education. In 1990, he stopped taking regular courses because of language difficulties and came back to Korea. In 1992, he went to Australia again to continue his studies, but had a lot of difficulty in getting along with people, because he felt that people didn't like him and spoke ill of him behind his back. Kim decided to go to a psychiatric clinic but tried to jump off the roof of the hospital because he felt that some racists had followed him to the hospital to kill him. A month later (in September, 1992), Kim came back to Korea and was extremely anxious because he thought the Mafia would kill him and his family. A few weeks later, Kim tried to cut his wrist and was brought to a psychiatric hospital in Korea. He was hospitalized with the diagnosis of schizophrenia, and remained at the hospital until January, 1993.

After being released from the hospital, he stayed home in a relatively isolated state, receiving outpatient treatment for a year. Kim moved from job to job until the summer of 1995. In November, 1995, Kim's brother suggested to him that they start a small construction company together. Kim was supposed to do the jobs after his brother had made the business arrangements. However, three days before starting work, Kim began to feel extremely anxious about the new business and suddenly told his family that he would kill himself. Kim was then brought to the hospital by his family and was hospitalized.

10.12.3.4.1 Family history

Kim's father is 80 years of age and owns a lot of land in the southern part of Korea. However, he is extremely stressed about money and exists on only minimal living expenses. He had considerable trouble with his wife and children because he was very stubborn, mean with money, and did not listen to anybody. He does not get along well with people around him and is very indifferent to his children. Kim's mother, who died in 1994, had received only six years of elementary school education. She was very introverted and passive, and had much trouble with her husband who was verbally and physically abusive toward her. She relied on her children and often told them how her husband abused her. Her children were always on their mother's side. Kim has one sister (age 54) who is a pharmacist who worked at this occupation and helped support her siblings until she got married. Kim has a brother who is a 47-year-old obstetrician. He got married when he was in medical school and has a daughter. He was divorced after a few years, because he got a girlfriend whom he married. He also has troubles with his new wife. Kim's brother also had some psychiatric problems when he was a high school student. At that time, he was worried that he would die of lung cancer even though he had no health problems. He was also very afraid that the roof might fall down. Since then, he has not had any mental problems. Kim also has a second brother, age 33, who runs a construction company and who is very close to the patient.

10.12.3.4.2 Behavioral observations

Kim does not talk easily with people and prefers to keep a distance from others. He always worries about what other people think of him or are saying about him, and is afraid that he might make a mistake in front of people. Kim spends most of the time staring at the wall or lying in his bed. At times, he can be heard talking to himself in a very loud voice.

10.12.3.4.3 Symptomatic behavior

Kim's symptoms and behavior reflect severe psychopathology as is shown below.

(i) *Thought control.* He feels as though his thoughts are being controlled by others, and that somebody is telling him to jump off a mountain or a building. He feels as though other people are controlling his thinking processes through the television or through birds.

(ii) *Thought broadcasting or leaking.* Kim feels that other people know what he thinks even though he says nothing to them. He feels that, in spite of precautions, what he thinks just leaks out and he can not control it.

(iii) *Idea of reference and paranoid ideas.* Kim feels that someone overhears his talking on the phone, and he suspects people around him are trying to do him harm. He thinks that all Koreans know that he has committed grave offenses and are backbiting him.

(iv) *Feels guilty and depressed.* He says that he feels depressed whenever he thinks about all of his faults. He also reported that he feels more depressed since he failed to kill himself. Kim is worried now about how to kill himself.

10.12.3.4.4 The Minnesota Report on Kim

The problems of Kim's psychopathology are clearly described in the MMPI-2 profiles of Figures 11 and 12 and the following narrative report.

> *Profile validity*
>
> This MMPI-2 profile should be interpreted with caution. There is some possibility that the clinical report is an exaggerated picture of the client's

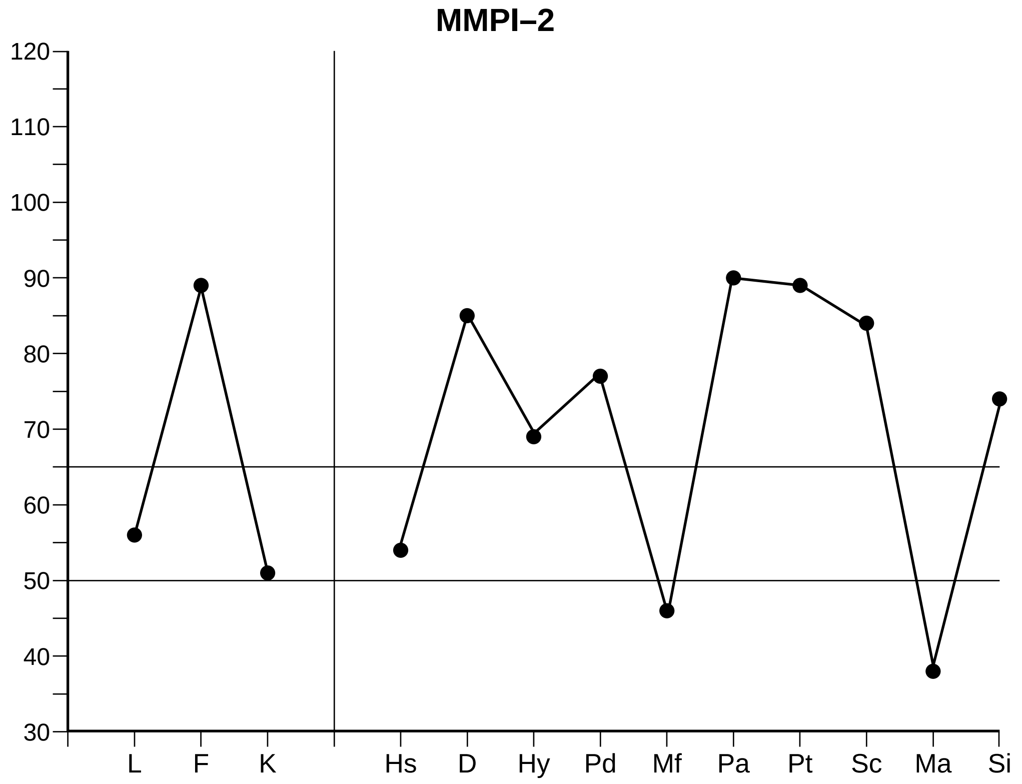

	L	F	K	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma	Si
Raw score	5	17	16	6	36	29	28	24	21	29	30	11	46
K correction				8			6			16	16	3	
T score	56	89	51	54	85	69	77	46	90	89	84	38	74

Figure 11 MMPI-2 profile of Kim (1).

present situation and problems. He is presenting an unusual number of psychological symptoms. This response set could result from poor reading ability, confusion, disorientation, stress, or a need to seek a great deal of attention for his problems.

His test-taking attitudes should be evaluated for the possibility that he has produced an invalid profile. He may be showing a lack of cooperation with the testing or he may be malingering by attempting to present a false claim of mental illness. Determining the sources of his confusion, whether conscious distortion or personality deterioration, is important because immediate attention may be required. Clinical patients with this validity profile are often confused and distractible and have memory problems. Evidence of delusions and thought disorder may be present. He may be exhibiting a high degree of distress and personality deterioration.

The client's responses to items in the latter portion of the MMPI-2 were somewhat exaggerated in comparison to his responses to earlier items. There is some possibility that he became more careless in responding to these later items, thereby raising questions about that portion of the test. Although the standard validity and clinical scales are scored from items in the first two-thirds of the test, caution should be taken in interpreting the MMPI-2 Content Scales and supplementary scales, which include items found throughout the entire item pool

Symptomatic patterns

This report was developed using the Pa and Pt scales as the prototype. The client's MMPI-2 clinical profile suggests that he is experiencing many psychological problems at this time. He appears to ruminate a great deal and may manifest obsessional or compulsive behavior. He holds beliefs that others are not likely to accept and tends to obsess about them to the point of alienating others. He appears to be quite intense, anxious, and distressed. Individuals with this profile may be overreacting to environmental situations with intense anxiety, suspicion, and concern. He feels insecure and inadequate when dealing with his problems. He may feel very angry with himself and others, and he may feel very guilty about his fantasies or beliefs. Often rather rigid, he may have problems controlling and directly expressing his anger.

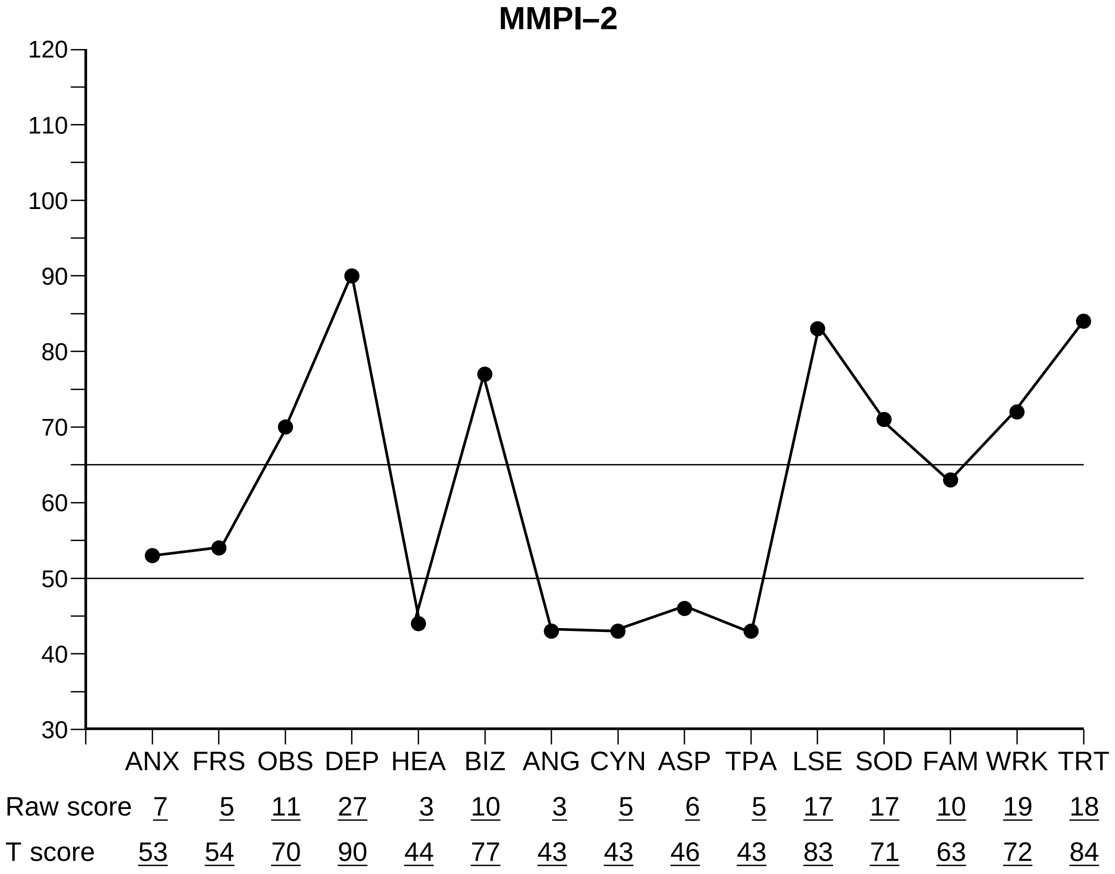

Figure 12 MMPI-2 profile of Kim (2).

Profile frequency

Profile interpretation can be greatly facilitated by examining the relative frequency of clinical scale patterns in various settings. The client's high-point clinical scale score (Pa) occurs in 9.6% of the MMPI-2 normative sample of men. However, only 3% of the sample have Pa as the peak score at or above a T score of 65, and only 2.2% have well-defined Pa spikes. This elevated MMPI-2 pattern (6–7/7–6) is very rare in samples of normals, occurring in less than 1% of the MMPI-2 normative sample of men.

The frequency of this MMPI-2 high-point Pa score is relatively high in various inpatient settings. In the Graham and Butcher (1988) sample of psychiatric inpatients, this profile peak is the second most frequent peak score (15.7%) for males, with 12.6% of the cases scoring in the clinically significant range (8.2% are well defined). In the large NCS inpatient sample, this high-point clinical scale score (Pa) is the third most frequent peak score, occurring in 14.3% of the men. Moreover, 12.1% of the males in the inpatient sample have this high-point scale spike at over a T score of 65, and 7.5% are well defined in that range. Male inpatients in a Veterans Administration setting (Arbisi & Ben-Porath, 1993) produce this high-point peak score with 10.1% frequency; 9.1% of the cases have Pa elevated above a T score of 65, and 6.1% of the cases are well defined.

This elevated MMPI-2 pattern (6–7/7–6) is found in less than 1% of the males in the Graham and Butcher (1988) sample, in 2.6% of the males in the NCS inpatient sample, and in 1.7% of the men in a Veterans Administration inpatient sample (Arbisi & Ben-Porath, 1993).

Profile stability

The relative scale elevation of his highest clinical scale scores suggests some lack of clarity in profile definition. Although his most elevated clinical scales are likely to be present in his profile pattern if he is retested at a later date, there could be some shifting of the most prominent scale elevations in the profile code. The difference between the profile type used to develop the present report and the next highest scale in the profile code was 4 points. So for example, if the client is tested at a later date, his profile might involve more behavioral elements related to elevations on D. If so, then on retesting, pronounced complaints of depressed mood and low morale might become more prominent.

Interpersonal relations

Individuals with this profile tend to be experiencing some interpersonal distress. The client seems somewhat shy and may have excessively high moral standards by which he judges others. His inflexibility in interpersonal situations may place a great strain on close relationships because he seems to test other people to reassure himself. He appears to be rather touchy or hostile interpersonally, and he may brood over what he imagines others have done to him. He tends to feel insecure in personal relationships, is hypersensitive to rejection, and may become jealous at times. He tends to need a great deal of reassurance.

He is a very introverted person who has difficulty meeting and interacting with other people. He is shy and emotionally distant. He tends to be very uneasy, rigid, and overcontrolled in social situations. His shyness is probably symptomatic of a broader pattern of social withdrawal. Personality characteristics related to social introversion tend to be stable over time. His generally reclusive behavior, introverted lifestyle, and tendency toward interpersonal avoidance may be prominent in any future test results.

Diagnostic considerations

His excessive anxiety and obsessive behavior should be taken into consideration in the final diagnosis. In addition, the possibility of a Paranoid Disorder or Paranoid Personality should be evaluated.

Treatment considerations

Inpatients with this MMPI-2 clinical profile usually have severe psychological problems as well as physical health concerns. Psychiatric treatment should focus on the client's anxiety and self-doubts. The use of medication to relieve his intense tension should be considered. The possibility that he has suspicious or paranoid ideas should be kept in mind when considering psychological treatment options. He may have difficulty forming a therapeutic relationship. He is quite rigid and intellectualizes a great deal; therapeutic progress is likely to be slow. Individuals with this MMPI-2 Clinical profile tend to have unrealistic expectations of themselves and perfectionistic ideals that may require some challenging if their personal vulnerability is to be diminished. A therapeutic approach such as rational-emotive therapy might enable him to acquire more self-acceptance.

Note: this MMPI-2 interpretation can serve as a useful source of hypotheses about clients. This report is based on objectively derived scale indices and scale interpretations that have been developed in diverse groups of patients. The personality descriptions, inferences, recommendations contained herein need to be verified by other sources of clinical information because individual clients may not fully match the prototype. The information in this report should most appropriately be used by a trained, qualified test interpreter. The information contained in this report should be considered confidential.

10.12.3.4.5 Diagnosis

Possibly schizophrenia or delusional disorder. Rule out depression with psychotic feature.

10.12.3.5 Haifa, Israel: The Case of Yuri Z.

The patient, Yuri Z., is a 30-year-old male who was born in Israel of Polish immigrant parents. He is married and he and his wife are expecting their second child. He has completed 13 years of education and at present he is a student in a large university and works part-time as a security guard. He was referred by the student counseling center of the university for psychological evaluation and possible treatment.

Yuri came to the clinic following a brief course of couple therapy where he and his wife were treated for severe marital problems. Following this family therapy Yuri decided to seek further individual therapy even though he still sees that separation from his wife is a possibility because their problems, though lessened, are still present.

Yuri is the youngest child in a family of four siblings. His older sister is doing her residency in psychiatry at Harvard. He initially described his family home life as "loving, encouraging, and wonderful' but later in the session changed this interpretation to more rigid, Spartan, encouraging to "hold on," and generally a "struggle" much of the time. Yuri's parents are Holocaust survivors and the father is described as saying that he is not a victim of the Holocaust but a victor of the Holocaust. The father has tried to cut himself and his children off from the terrible family history. The daughter, unlike Yuri, felt that the father did not allow her to grow and develop her own identity and she disconnected from the father a long time ago. Yuri sees his father as emotional, kind, and highly protective. He described his mother as a model mother, a wonderful person. His mother is a nurse who works with babies and who has always devoted herself to helping others. Later in the session he described his mother as distant, cold, and "frozen."

Yuri is a very bright person who had been an excellent student all through his earlier school years. However, he had school behavior problems all the time. Yuri considers himself to have been an undiagnosed and untreated hyperactive child. Due to behavior problems he was suspended from junior high school for a month (this was very unusual in the school he was attending). The reason for the suspension, after several reprimands, was his unexcused absences from school. After a period of time he was allowed to return to school. However, he lost his motivation to attend the regular school system and he transferred to another school. But when he learned about the possibility of going to a new kibbutz, he was excited, and left home despite his parents' objections, especially those of his mother who was afraid he wouldn't complete school. Yuri describes his years at the kibbutz as the best time in his life. He felt as though he had gained freedom, and enjoyed the open land and young peers.

He enlisted in the army, in a unit that consisted mainly of persons who wished to live on a kibbutz. He served in an elite unit in that corps. His mother died the day that he enlisted in the army. He served in the army for five years in what was described as a rigorous and very hard duty, mostly in the field and in the front line, because "I had nothing waiting for me outside the army." He saw many of his friends die or be wounded and he participated in several killings of enemy terrorists. Images of scenes from that period still recur in his dreams. He was a superb soldier in the army, but even there he had disciplinary problems because of his irresponsible behavior at times. His underlying personality problems persisted.

10.12.3.5.1 The Minnesota Report on Yuri Z.

Yuri Z's personality problems are illustrated in the MMPI-2 profiles of Figures 13 and 14 and the following narrative report.

Profile validity

This MMPI-2 profile should be interpreted with caution. There is some possibility that the clinical report is an exaggerated picture of the client's present situation and problems. He is presenting an unusual number of psychological symptoms. This response set could result from poor reading ability, confusion, disorientation, stress, or a need to seek a great deal of attention for his problems.

His test-taking attitudes should be evaluated for the possibility that he has produced an invalid profile. He may be showing a lack of cooperation with the testing or he may be malingering by attempting to present a false claim of mental illness. Determining the sources of his confusion, whether conscious distortion or personality deterioration, is important because immediate attention may be required. Clinical patients with this validity profile are often confused and distractible and have memory problems. Evidence of delusions and thought disorder may be present. He may be exhibiting a high degree of distress and personality deterioration.

Symptomatic patterns

This report was developed using the Pd and Sc scales as the prototype. A somewhat mixed symptom picture is reflected in this profile. Individuals with this MMPI-2 clinical scale configuration tend to show a pattern of chronic psychological maladjustment. The client appears to be unconventional and nonconforming, tending to act impulsively without regard for the consequences. Apparently he is emotionally troubled, and his behavior is unpredictable. He may be aloof and withdrawn. He may also appear hedonistic and may engage in dangerous or deviant behavior for the thrill of it. He may have a history of unreliable and irresponsible actions, for which he often blames others. He tends to have low self-esteem, a history of underachievement, and problems with authority. Excessive alcohol or drug use is often found among individuals with this clinical pattern.

In addition, the following description is suggested by the content of the client's item responses. He had endorsed a number of items suggesting that he is experiencing low morale and a depressed mood. Although he may be socially assertive and may project a positive image to others, his response content indicates a rather negative self-image. He reports some antisocial beliefs and attitudes, admits to rule violations, and acknowledges antisocial behavior in the past.

Profile frequency

Profile interpretation can be greatly facilitated by examining the relative frequency of clinical scale patterns in various settings. The client's high-point clinical scale score (Pd) occurs in 9.1% of the MMPI-2 normative sample of men. However, only 3.3% of the normative men have Pd as the peak score equal to or greater than a T score of 65, and only 1.0% have well-defined Pd spikes. This elevated MMPI-2 profile configuration (4–8/8–4) is very rare in samples of normals, occurring in less than 1% of the MMPI-2 normative samples of men. A high-point Pd scale score occurs in 10.3% of the sample of military men (Butcher, Jeffrey, et al., 1990). However, only 3% of the sample have Pd equal to or greater than a T score of 65. Only 2% of these high-point peaks are well defined at that level of elevation.

He scored relatively high on AAS, suggesting the possibility of a drug- or alcohol-abuse problem. The base rate data on his profile type among residents in alcohol and drug programs should also be evaluated. This MMPI-2 profile configuration contains the most frequent high point, the Pd score, among alcohol- and drug-abusing populations. Over 24% of the men in substance-abuse treatment programs have this pattern (McKenna & Butcher, 1987).

Profile stability

The relative scale elevation of the highest scales in his clinical profile reflects high profile definition. If he is retested at a later date, the peak scores on

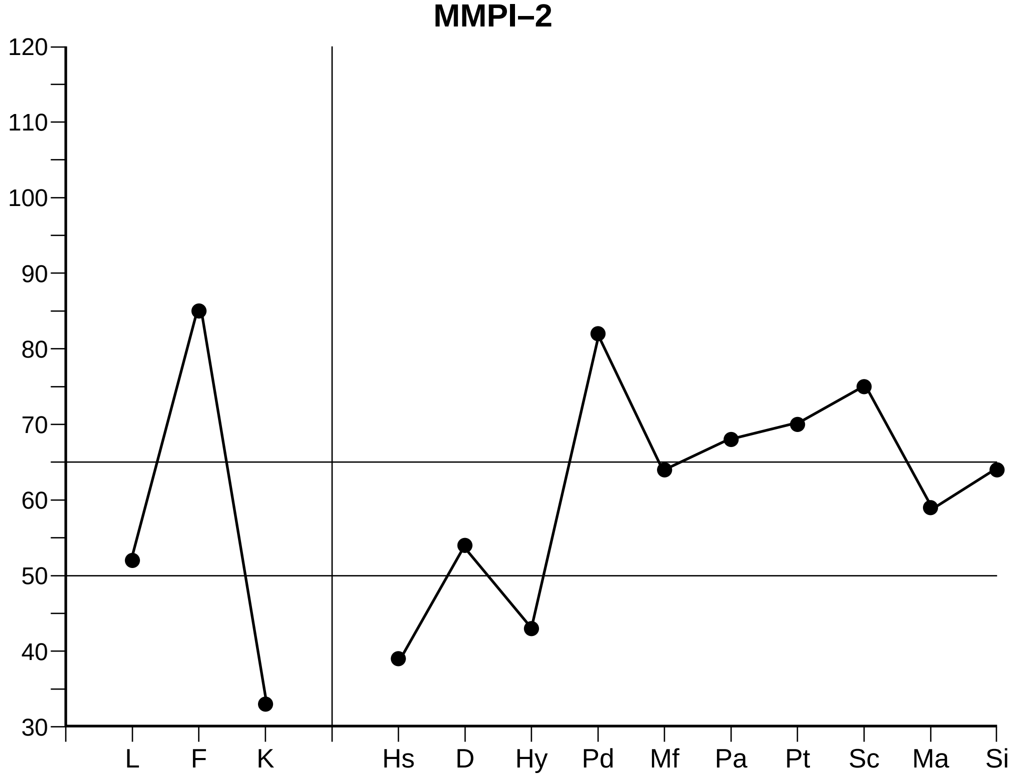

	L	F	K	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma	Si
Raw score	4	16	7	5	20	18	33	33	15	29	34	23	37
K correction				4			3			7	7	1	
T score	52	85	33	39	54	43	82	64	68	70	75	59	64

Figure 13 MMPI-2 profile of Yuri Z. (1).

this test are likely to retain their relative salience in his retest profile pattern. His high-point score on Pd is likely to remain stable over time. Short-term test–retest studies have shown a correlation of 0.81 for this high-point score. Spiro, Butcher, Levenson, Aldwin, and Bosse (1993) reported a moderate test–retest stability index of 0.62 in a large study of normals over a five-year test–retest period.

Interpersonal relations

Although his behavior is frequently seen as manipulative, he apparently lacks social skills. Often misunderstanding the motives of others and overly sensitive to rejection, he may appear aloof and feel insecure in close personal relationships. He may also greatly distrust others. He is probably behaving in unpredictable and erratic ways that may produce a great deal of marital strain. Many individuals with this profile have serious marital problems that require marriage counseling.

He is somewhat shy, with some social concerns and inhibitions. He is a bit hypersensitive about what others think of him and is occasionally concerned about his relationships with others. He appears to be somewhat inhibited in personal relationships and social situations, and he may have some difficulty expressing his feelings towards others.

His very high score on the Marital Distress Scale suggests that his marital situation is quite problematic at this time. He has reported a number of problems with his marriage that are possibly important to understanding his current psychological symptoms.

The content of this client's MMPI-2 responses suggests the following additional information concerning his interpersonal relations. He views his home situation as unpleasant and lacking in love and understanding. He feels like leaving home to escape a quarrelsome, critical situation and to be free of family domination.

Diagnostic considerations

Individuals with this profile usually receive a diagnosis of Personality Disorder. His response content is consistent with the antisocial features in his history. These factors should be taken into consideration in arriving at a clinical diagnosis. His self-reported tendency toward experiencing a depressed mood should be taken into consideration in any diagnostic formulation.

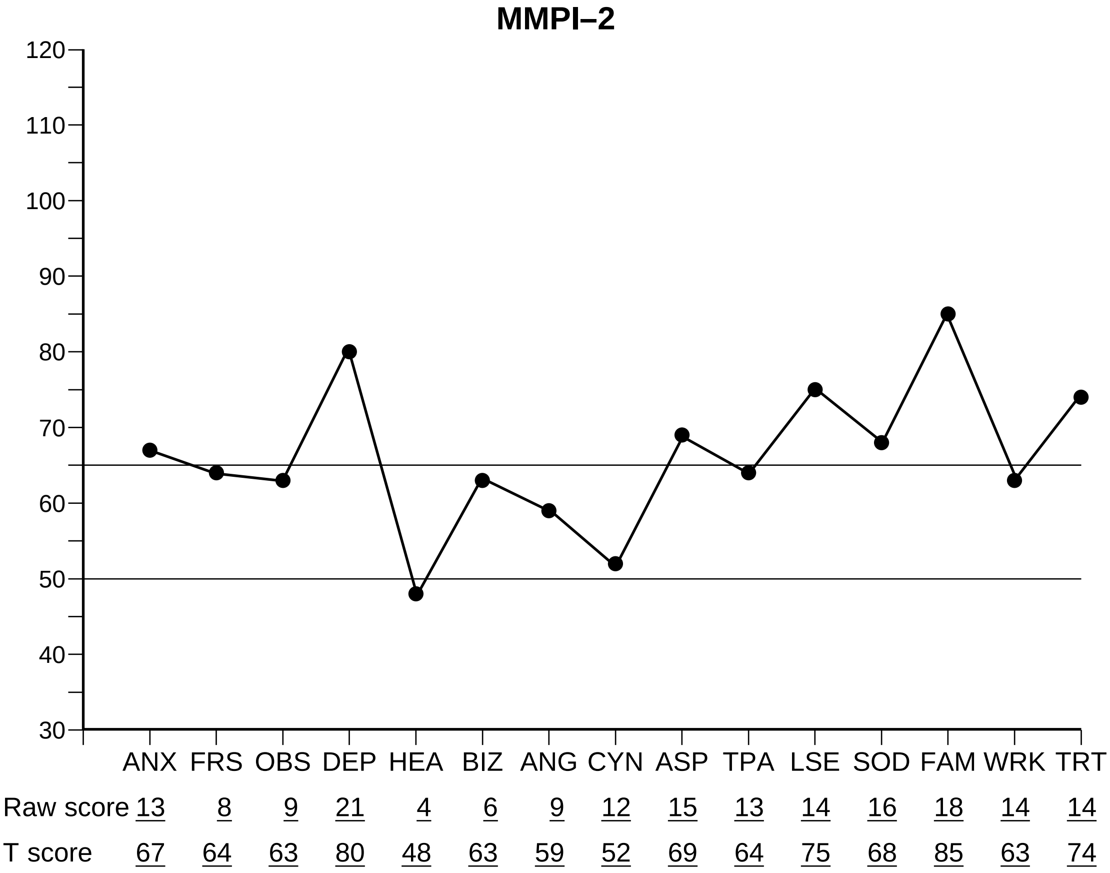

Figure 14 MMPI-2 profile of Yuri Z. (2).

He has a number of personality characteristics that are associated with a substance-use or abuse disorder. The client's scores on the addiction proneness indicators, along with the personality characteristics reflected in the profile, suggests that he resembles some individuals who develop addictive disorders. A substance-abuse evaluation should explore this possibility through a careful review of his personality traits and typical behaviors. In his responses to the MMPI-2, he has acknowledged some problems with excessive use or abuse of addictive substances.

Treatment considerations

Individuals with this profile often live stressful and crisis-filled lives and may need psychological treatment to enable them to structure their lives better. They typically have a number of reality-based problems that complicate treatment on an outpatient basis. Although they may seek temporary symptom relief or may be referred for psychological treatment by another person or agency, their long-standing personality problems make it unlikely that they would profit much from individual, insight-oriented psychotherapy. They are evidently not very reflective or introspective and would also have difficulty establishing a treatment relationship because they mistrust other people. The client is so emotionally and socially alienated that it would be difficult for a therapist to gain his confidence. Such individuals also tend to act out problems rather than attempting to understand and solve them.

Behavioral treatment that focuses on life management may be of value in teaching them more adaptive social functioning. Regardless of the treatment programs offered, their behavior may change quite slowly and there may be many crises during the course of therapy.

Note This MMPI-2 interpretation can serve as a useful source of hypotheses about clients. This report is based on objectively derived scale indices and scale interpretations that have been developed in diverse groups of patients. The personality descriptions, inferences, recommendations contained herein need to be verified by other sources of clinical information because individual clients may not fully match the prototype. The information in this report should most appropriately be used by a trained, qualified test interpreter. The information contained in this report should be considered confidential.

After his discharge from the army he felt very empty and depressed. He locked himself in at home for three months and was using drugs extensively at this time. Afterwards he was admitted to a technical program at the university. Later in that year his father died. A year later he went on a trip, a "root searching' tour to Poland. He described his parents' death as the beginning of his crisis. He feels that he has never really processed or coped with their deaths.

10.12.3.5.2 Treatment

Yuri was admitted to the clinic's outpatient program for individual psychodynamically oriented therapy. He comes quite regularly to sessions and has been in therapy for the last eight months.

10.12.3.6 Cross-national Generality of Psychopathology: What We Have Learned from the Cases

The case studies presented in this chapter illustrate clearly the generalizability of the MMPI-2 across diverse cultures. As one can see by reading the personality descriptions and problem situations described in the clinical cases from different countries, the MMPI-2 variables are what one would expect to find if the patients were American patients assessed in the USA. Moreover, these MMPI-2 profiles appear to be close matches with the cases, in spite of the fact that the profiles were generated on US norms. In addition, the computer-derived narratives, which were developed on research conducted in the USA, show close matches when the reports are applied in other countries.

How can an instrument developed in one language and culture show such robustness when applied to patients in other cultures who responded to translated versions of the items? Several factors account for this correspondence of patient description across cultures. Research in cross-cultural psychopathology has shown that mental disorders and diagnostic systems are comprised of similar symptom patterns in different cultures—a schizophrenic is a schizophrenic regardless of culture. Many similar disorders show common patterns of symptoms across cultural boundaries.

Similarly, the MMPI-2 items sample a broad range of symptoms and problems constituting the abnormal syndromes. Given a well translated item set, resulting in an equivalent form, the symptoms are found to describe well the disorders in each culture. Patients who are depressed, for example, tend to respond to the same types of items in the cultures studied here, producing familiar personality profile patterns. This is not to say that cultural factors are irrelevant to the manifestation of psychopathology. It has been shown that, for example, child-rearing patterns in Italy might result in a "dramatization of symptoms" (Butcher & Pancheri, 1976; Zola, 1966) that is not noted in other cultures. However, such differences do not greatly alter the main symptoms of disorder, thereby resulting in some commonality for disorders across cultures.

These case studies also provide an important example with respect to objective personality assessment. The study on computer-based assessment across Norway, France, Australia, and the USA shows that the empirical description of MMPI-2 patterns has broad generalizability whether interpreted by clinicians or by the computer.

10.12.4 FUTURE DIRECTIONS

Each year more foreign language translations of the MMPI-2 become available, providing clinicians and researchers in other countries the means of assessing clients in ways comparable to those in the USA. The availability of objective parallel methods of assessing psychological problems make it possible for joint international research programs to examine psychopathology cross-culturally. The MMPI-2 has shown considerable adaptability across international boundaries. When careful translation and adaptation procedures are followed to develop a sound translation, the MMPI-2 can be usefully employed in clinical assessment in other languages and cultures. Moreover, the translated version of the MMPI-2 appears to have a high degree of accuracy in characterizing personality and patient problems even with clients being assessed in very diverse settings. What can one conclude from this body of research?

The research reported here supporting the questionnaire approach to psychopathology should assure psychologists from other countries about the transportability of the MMPI-2. The utility of the instrument in making clinical decisions has been amply demonstrated through several international projects, and MMPI-2 profiles can add a salient dimension to clinical decision making in other countries. Psychologists and psychiatrists in other countries who are interested in studying psychopathology across different cultures or national groups might consider employing an objective measuring instrument, like the MMPI-2, to assess symptoms of disorder.

The accuracy of automated test interpretation has been demonstrated in the research reported here, and should provide researchers with an objective tool for cross-cultural research in psychopathology. When a computer-based interpretation, developed on normals and patients in the USA, is applied in international contexts, the resulting reports characterize well the patients in other countries. This finding has great implications for cross-cultural research on psychopathology in that subjectivity in interpretation in cross-cultural psychopathology comparisons can be substantially reduced.

10.12.5 SUMMARY

Psychologists have explored the use of computers to interpret psychological tests for many years. The first psychological test to be fully interpreted by computer was the Minnesota Multiphasic Personality Inventory in 1961 at the Mayo Clinic, Rochester, Minnesota. Since then a number of computer interpretation systems have become available in the USA. Previous studies on computer assessment for the original MMPI and MMPI-2 in Australia have shown it to be a promising interpretive approach. This chapter describes an empirical evaluation of computer-based MMPI-2 reports on Australian, French, Norwegian, and American patients. The patients were administered the MMPI-2 and computerized Minnesota Reports were generated. Clinicians familiar with the patients rated the reports in terms of adequacy of information and accuracy of descriptions and predictions. The computer-based reports were considered to provide valuable information, with two-thirds (66%) judged to be 80–100% accurate and 87% considered to be over 60% accurate in the information provided. The clinical application of computer-based reports was illustrated through case examples from several countries, including The Netherlands, Italy, Korea, Iran/United States, and Israel.

ACKNOWLEDGMENTS

We would like to express our appreciation to Joe Bolza, Shelly Colman, Pam McAsey, and Louise Sellenger for their cooperation in gathering the data for the study. We would especially like to thank Linda Fresquez for her tireless and efficient processing of test information and Steven Rouse and Russ Morfitt for their able assistance in the data processing.

10.12.6 REFERENCES

Abe, M. (1955). On the personality inventory. *Bulletin of Tohoku Institute of Correctional Science, 1,* 161–162.

Adcock, C. J., & Richie, J. E. (1958). Intercultural use of Rorschach. *American Anthropologist, 60,* 881–892.

American Psychological Association (1986). *American Psychological Association guidelines for computer-based tests and intepretations.* Washington, DC: Author.

Arbisi, P., & Ben-Porath, Y. S. (1993, March). *Interpreting the F scales of inpatients: Moving from art to science.* Paper presented at the 28th Annual symposium on Recent Developments in the Use of MMPI/MMPI-2/MMPI-A. St. Petersburg Beach, FI.

Berah, E., Butcher, J., Miach, P., Bolza, J., Colman, S. & McAsey, P. (1993, October). *Computer-based interpretation of the MMPI-2: An Australian evaluation of the Minnesota Report.* Poster presented at the 28th Annual Conference of the Australian Psychological Association, Gold Coast.

Berah, E., Miach, P., & Butcher, J. N. (1995, June). *The MMPI-2 downunder: Validation of computer-based interpretive reports in Australia.* Paper given at the 15th International Conference on Personality Assessment., Haifa, Israel.

Butcher, J. N. (Ed.) (1987). *Computerized psychological assessment.* New York: Basic Books.

Butcher, J. N. (1993). *User's guide for the MMPI-2 Minnesota Report: Adult clinical system.* Minneapolis, MN: National Computer Systems.

Butcher, J. N. (1995). How to use computer-based personality test reports. In J. N. Butcher (Ed.), *Clinical personality assessment: Practical approaches* (pp. 78–94) . New York: Oxford University Press.

Butcher, J. N. (1996). *International adaptations of the MMPI-2: Research and clinical applications.* Minneapolis, MN: University of Minnesota Press.

Butcher, J. N., Jeffrey, T., Cayton, T. G., Colligan, S., DeVore, J., & Minnegawa, R. (1990). A study of active duty military personnel with the MMPI-2. *Military Psychology, 2,* 47–61.

Butcher, J. N., Narikiyo, T., & Bemis-Vitousek, K. (1992). Understanding abnormal behavior in cultural context. In H. Adams & P. Sutker (Eds.), *Comprehensive handbook of psychopathology,* (2nd ed., pp. 83–105). New York: Plenum.

Butcher, J. N., & Pancheri, P. (1976). *A handbook of cross-national MMPI research.* Minneapolis, MN: University of Minnesota Press.

Cheung, F. M., & Song, W. (1989). A review on the clinical applications of the Chinese MMPI. *Psychological Assessment, 1,* 230–237.

Cheung, F. M., Song, W., & Butcher, J. N. (1991). An infrequency scale for the Chinese MMPI. *Psychological Assessment, 3,* 648–653.

Cheung, F. M., Zhao, J., & Wu, C. (1992). Chinese MMPI profiles among neurotic patients. *Psychological Assessment, 4,* 214–218.

Eyde, L., Kowal, D., & Fishburne, F. J. (1991). In T. B. Gutkin & S. L. Wise (Eds.), *The computer and the decision making process* (pp. 75–123). Hillsdale, NJ: Erlbaum.

Fowler, R. D. (1969). Automated interpretation of psychological test data. In J. N. Butcher (Ed.) *MMPI: Research developments and clinical applications* (pp. 105–126). New York: McGraw-Hill.

Fowler, R. D. (1978). *Computer interpretation of the MMPI in Australia.* Unpublished manuscript.

Fowler, R. D., & Butcher, J. N. (1987). International applications of computer based testing and interpretation. *Applied Psychology: An International Review, 36,* 419–429.

Gilberstadt, H., & Duker, J. (1965). *A handbook for clinical and actuarial MMPI interpretation.* Philadelphia: Saunders.

Gillet, I., Simon, M., Guelfi, J., Baun, A., Monier, C. Seunevei, F., & Svarna, L. (1996). The MMPI-2 in France. In J. N. Butcher (Ed.) *International adaptations of the MMPI-2: A handbook of research and clinical* *applications* (pp. 395–415). Minneapolis, MN University of Minnesota Press.

Graham, J. R. & Butcher, J. N. (1988, March) *Differentiating schizophrenic and major affective disorders with the revised form of the MMPI.* Paper presented at the 23rd Annual Meeting of the Symposium on Recent Developments in the Use of the MMPI. St. Petersburg, FI.

Halbower, C. C. (1955). *A comparison of actuarial versus clinical prediction to classes discriminated by MMPI.* Doctoral dissertation, University of Minnesota.

Kazarian, S., & Evans, D. R. (in press). *Cross cultural clinical psychology*. New York: Oxford University Press.

Kim, A. S. (1988). *MMPI—The clinical interpretation.* Seoul, South Korea: National University Press.

Lindzey, G. (1958). *Projective techniques in cross cultural research.* New York: Appleton-Century-Crofts.

Lonner, W., & Berry, J. (1986). *Field methods in cross cultural research.* Beverly Hills, CA: Sage.

Marks, P. A., & Seeman, W. (1963).

Marks, P. A., Seeman, W., & Haller, D. L. (1974). *The actuarial use of the MMPI with adolescents and adults.* Baltimore: William and Wilkins.

Manos, N. (1984). Adaptation of the MMPI in Greece: Translation, standardization, and cross-cultural comparison. In J. N. Butcher & C. D. Spielberger (Eds.), *Advances in personality assessment* (Vol. 4, pp. 159–208). Hillsdale, NJ: Erlbaum.

McKenna, T., & Butcher, J. N. (1987). *Continuity of the MMPI with alcoholics.* Paper presented at the 23rd Annual Symposium on Recent Developments in the Use of the MMPI. Seattle, WA.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence.* Minneapolis, MN: University of Minnesota Press.

Meehl, P. E. (1956). Wanted—a good cookbook. *American Psychologist, 11,* 263–272.

Pancheri, P., & Biondi, M. (1987). Computerized psychological assessment in Italy. In J. N. Butcher (Ed.), *Computerized psychological assessment: A practitioner's guide* (pp. 236–260). New York: Basic Books.

Pancheri, P., Sirigatti, S., & Biondi, M. (1996). The MMPI-2 in Italy. In J. N. Butcher (Ed.), *International adaptations of the MMPI-2: A handbook of research and clinical applications* (pp. 416–441). Minneapolis, MN: University of Minnesota Press.

Reda, G. (1948). Unpublished translation. University of Rome, Italy.

Risetti, F. J., Himmel, E., Maltes, S., & Gonzalez, H. A. (1989). Estandarizacion del Inventario Multifasico de Personalidad de Minnesota (MMPI), en poblacion adulta Chilena [Standardization of the Minnesota Multiphasic Personality Inventory (MMPI) in the adult Chilean population]. *Revista Chilena de Psicologia, 10,* 41–62.

Rome, H. P., Swenson, W. M., Mataya, P., McCarthy, C. E., Pearson, J. S., Keating, F. R., & Hathaway, S. R. (1962). Symposium on automation techniques in personality assessment. *Proceedings of the Staff Meetings of the Mayo Clinic, 37,* 61–82.

Savasir, I., & Erol, N. (1990). The Turkish MMPI: Translation, adaptation, and validation. In J. N. Butcher & C. D. Spielberger (Eds.), *Advances in personality assessment* (Vol. 8, pp. 49–62). Hillsdale, NJ: Erlbaun.

Spiro, R., Butcher, J. N., Levenson, M., Aldwin, C., & Bosse, R. (1993, August). *Personality change over five years: The MMPI-2 in older men.* Paper at the Annual Meeting of the American Psychological Association, Toronto, Canada

Strassberg, D. S., Clutton, S., & Korboot, P. (1991). A descriptive and validity study of the Minnesota Multiphasic Personality Inventory-2 (MMPI-2) in an elderly Australian sample. *Journal of Psychopathology and Behavioral Assessment, 13,* 301–311.

Strassberg, D. S., Tilley, D., Bristone, S., & Tian, P. S. (1992). The MMPI and chronic pain: A cross-cultural view. *Psychological Assessment, 4,* 493–497.

Sundberg, N. (1956). The use of the MMPI for cross-cultural personality study: A preliminary report on the German translation. *Journal of Abnormal and Social Psychology, 58,* 281–283.

World Health Organization (1992). *ICD-10 classification of mental and behavioral disorder: Clinical description and diagnostic procedures.* Geneva, Switzerland: Author.

Yap, P. M. (1951). Mental diseases peculiar to certain cultures: A survey of comparative psychiatry. *Journal of Mental Science, 97,* 313–327.

Ziskin, J. (1981). The use of the MMPI in forensic settings. In J. N. Butcher, W. G. Dahlstrom, M. D. Gynther, & W. Schofield (Eds.), *Clinical notes on the MMPI* (whole #9). Nutley, NJ: Hoffman-LaRoche Laboratories/NCS.

Zola, I. K. (1966). Culture and symptoms—an analysis of patients' presenting complaints. *American Sociological Review, 31,* 615–630.

Zou, Y., & Zhao, C. (1992). The validity of the MMPI in the clinical diagnosis of 1422 Chinese subjects. *Chinese Mental Health Journal, 6,* 211–213.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.13
Mental Health in the Arab World

MARWAN DWAIRY
Nova Southeastern University, Oakland Park, FL, USA

10.13.1 INTRODUCTION 313
10.13.2 ARAB SOCIOCULTURAL HISTORY AND POPULATION 313
10.13.3 PSYCHOCULTURAL FEATURES OF ARABS 314
10.13.4 CLINICAL ASSESSMENT CONSIDERATIONS 315
  10.13.4.1 Manifestations of Psychosocial Disorders 315
  10.13.4.2 Psychometric Characteristics 316
10.13.5 HELP-SEEKING BEHAVIOR 316
  10.13.5.1 Traditional Arab Theory of Psychotherapy 318
10.13.6 EPIDEMIOLOGY OF CLINICAL PROBLEMS 318
  10.13.6.1 Psychiatric Morbidity 318
  10.13.6.2 Distribution of the Disorders 319
10.13.7 HEALTH SERVICE DELIVERY SYSTEM 320
10.13.8 FUTURE DIRECTIONS 321
10.13.9 SUMMARY 322
10.13.10 REFERENCES 323

10.13.1 INTRODUCTION

Compared to most Western societies, Arab societies are more collectivist and authoritarian, and Arab individuals thus adhere more closely to social values and norms. The relationship of the Arab individual to the family is characterized by interdependence, resulting in people who are less individuated than the typical Westerner. How do these sociocultural factors affect the Arabs" personality characteristics? How does the Arabic culture influence the epidemiology and manifestations of psychological disorders? How do Arabs perform on psychological tests? Given the individualistic background of psychotherapy, how do Arabs regard psychotherapy, and how do they deal with their psychological disorders? What kind of mental health system exists in Arabic countries? These points are addressed in this chapter, and suggestions are offered to psychotherapists for tailoring their therapies to the Arab client.

10.13.2 ARAB SOCIOCULTURAL HISTORY AND POPULATION

Arabs are currently living in 21 countries to the east and south of the Mediterranean sea, and in the Arabian peninsula. The vast majority are Muslims; however, a minority of Christians reside in Syria, Lebanon, Palestine, Jordan, and Egypt, and a minority of Druze in Syria, Lebanon, and Palestine. Despite the sociocultural diversity among Arabs, all share a historical experience that has shaped certain sociocultural features of contemporary Arab life.

Before the birth of Islam in 622, Arabs had lived in the Arabian peninsula, Syria, and Iraq, in a tribal social system characterized by collectivism and authoritarianism. From the beginning of Islam until the fifteenth century, Arabs experienced the glory period of the Arab-Islamic empire, in which they were the rulers of a state that expanded from Iran in the east, through North Africa, to Spain in the west. In this period, Arabs combined their tribal cultural heritage with the new values of Islam. Between the sixteenth and nineteenth centuries there was a stagnation period, in which they were governed by other Muslim states such as the Mamlukes and Ottomans (Turks). In the last two centuries, two important processes have taken place in the Arab world: a massive exposure to the democratic, more individualistic values of the West, and the establishment of separate independent states (Fattal & Sokari, 1988; Mansfield, 1991).

Four main historical periods can be identified, each noteworthy for a different collective experience: tribal, rulers, ruled, and westernization, successively. New periods added new layers of values that interacted with the former ones, such that each period is represented in contemporary Arab culture. In this way, the contemporary values of the Arab-Muslim people vary along two main dimensions: (i) collectivism–individualism, the first stemming from the tribal heritage and the second from exposure to the West; and (ii) ruler–ruled, the first with roots in the Arab-Muslim empire and the second from the period of stagnation (Dwairy, 1997b, 1998).

These two-dimensional value structures help to explain the cultural diversity among Arabs today. For example, the first cultural dimension would show that the majority of Arabs have adopted a traditional collectivistic way of life, in which the family is the primary social unit. The identity of the individual is derived from the family (Dwairy & Van Sickle, 1996; Meleis & La Fever, 1984), and family interests take priority over individual ones (Timimi, 1995). Family affairs are kept within the family, away from external intrusion (Meleis & La Fever, 1984). Familial affiliation is central for the majority of Arabs, as it meets the basic needs of the individual in exchange for the individual's adherence to family values and norms (Dwairy, 1997b, 1998). The development of an independent identity is not encouraged (Gorkin, Masalha, & Yatziv, 1985; Timimi, 1995). In addition to the collectivistic majority, there is a portion of Arabs that has to some degree assimilated into the Western individualistic culture, or has adopted a bicultural way of life (Dwairy, 1997b, 1998).

The second cultural dimension, ruler–ruled, helps to explain the hierarchy of authority in most Arabic societies, in which men and elders have exerted power over women and children (Budman, Lipson, & Meleis, 1992). Children are typically expected to obey the rules of the family and to modify their own desires for education, a job, and marriage, to meet family expectations. Females are expected to maintain the family's honor by chaste behavior, marrying through the family's arrangement or approval, and producing male children (Budman et al., 1992). In many Arab families, boys are regarded as an asset, whereas girls are a liability (Sharabi, 1977).

Socioeconomic styles of life explain much of the variance within the two cultural dimensions. Three major socioeconomic groups are found among Arabs: Bedouin (nomadic), rural, and urban. There has been a tremendous urbanization process since the late 1980s. Although the majority of Arabs have lived rural or Bedouin lives for centuries, endorsing a collective lifestyle, by the end of the twentieth century, 70% of Arabs will live in cities. Although living in a city is generally associated with less dependence on the family and less social control, collectivistic values are still esteemed. In fact, most Arabic cities do not meet all of the basic human needs; many urban citizens continue to work in agriculture and to maintain cohesive family relationships (Barakat, 1993; Fattal & Sokari, 1988). Thus, most urban, as well as rural and Bedouin families continue to endorse a traditional collectivistic way of living (Dwairy, 1997b, 1998).

10.13.3 PSYCHOCULTURAL FEATURES OF ARABS

Certain psychological features of Arabic culture are derived from the collectivistic and authoritarian structures in Arabic societies and are typified as follows:

(i) *External locus of control*. As a result of the actual helplessness of the individual who lives in a collective authoritarian society, he or she learns to believe that life events are determined by external powers—family, social leaders, or God. Religious people believe that their destiny is *maktoob* (written). Therefore, they do not feel responsible for their behavior or the events in which they have been involved (Al Khani, Bebbington, Watson, & House, 1986; Bazzoui, 1970).

(ii) *Interpersonal rather than intrapsychic sources of distress*. Arab individuals are conditioned through external threats of sanctions, making external controls more significant than internal (superego) ones. Therefore, shame rather than guilt dominates (Gorkin et al., 1985; Sharabi, 1975, 1977), and interpersonal coping strategies become more effective than intrapsychic defense mechanisms (Dwairy & Van Sickle, 1996).

(iii) *Indirect expression of emotions*. In Arabic socities, people tend to avoid expressing negative emotions such as anger and jealousy towards family members (West, 1987). Individuals are expected to exhibit emotions congruent to societal norms (*mosaiara*), and to hide authentic expressions. Other emotions are expressed through acting out behaviors away from the attention of others (*isligaba*) or through body language (Dwairy, 1997b, 1998; Sharabi, 1975).

(iv) *Unindividuated self*. The psychological autonomy and individuation that many Western psychosocial theories describe bear only limited relevance to the common pattern of psychosocial development in Arabic cultures (Dwairy & Van Sickle, 1996; Gorkin et al., 1985; Timimi, 1995). An Arab's identity is derived from the family, self-concept is enmeshed in the family concept, and an individual's needs, attitudes, and values stem from those of the family.

10.13.4 CLINICAL ASSESSMENT CONSIDERATIONS

Unfamiliarity with the Arabic culture may cause the clinician to misinterpret or pathologize certain cultural behaviors. Overdependence on the family, overprotective parents, avoidance of touching the opposite sex, passivity, ambivalence, reluctance to reveal family matters, hesitation in making personal decisions, and time-related problems are culturally normal and should not be considered as signs of immaturity or resistance (Budman et al., 1992). An Arab male inpatient's reluctance to make his bed or to clean his room is probably not resistance or regression; rather, it may well be his normal way of living (Meleis & La Fever, 1984).

10.13.4.1 Manifestations of Psychosocial Disorders

Because of the sociocultural background of Arabic patients, the manifestations of some psychological disorders may vary from those described in the *Diagnostic and statistical manual of mental disorders* (4th ed.) (*DSM-IV*, American Psychiatric Association, 1994). Due to the holistic cultural attitude toward mind and body adopted by Arabs, conflicts and distress, in general, are expressed in somatic terms (Dwairy, 1997c). This is the legitimate way of expressing personal problems (Budman et al., 1992; Racy, 1980). Sometimes somatic complaints are, indeed, the only expression of psychological distress.

Bazzoui (1970), for example, has identified certain characteristics of depression among patients in Iraq that are different from what is common in the West. Iraqi depressed patients show milder (and sometimes an absence of) mood changes. Mood changes were found in only 34% of depressed cases, and only 37% of manic cases displayed an elated mood. Feelings of guilt and worthlessness were almost absent among the depressed (only 13.8%). Suicidal thoughts and attempts were rare. Physical symptoms and hysterical behavior were the most outstanding features of depression, found in 65.5% of the cases. Fear of breaking the rules or shaming themselves and their families was much more dominant than feelings of individual responsibility and guilt (Bazzoui, 1970). Similar findings were reported for Sudanese depressed patients (Baasher, 1962).

Paranoid ideation, projection of responsibility, aggression, and "noisiness" were prevalent (47.5%) among hypomanic cases, and less common (27.6%) among depressed cases in Iraq. Interestingly, about one-quarter of the depressed population in the Iraqi study expressed a desire to run out of their homes into the wilderness, but this impulse was acted upon in only a few cases (Bazzoui, 1970). Bazzoui attributes these differences to the lack of individuated ego that is necessary to experience worthlessness. She suggests, moreover, that clinicians can diagnose Arab patients with depression even without the presence of miserable feelings (a hallmark feature for depression in the West), and with mania even when elated mood is absent.

According to Okasha and his colleagues, the obsessive-compulsive disorder (OCD) manifestations in Egypt are colored by Islamic religious rituals and beliefs that make the differential diagnosis difficult. Interestingly and in contrast to one of the *DSM-IV* criteria for OCD, most OCD subjects in the Egyptian study did not recognize the absurdity of their symptoms (Okasha, Saad, Khalil, Seif El Dawla, & Yehia, 1994).

Okasha and his colleagues also studied the manifestations of psychosis in Egypt. Delusions, depression, excitement, hallucinations, worrying, delayed sleep, and irritability were the most common presenting symptoms of Egyptian psychotic patients. The onset of acute psychosis occurred within five days in 68% of the cases. A stressor preceded the onset of psychotic symptoms in 74% of the Egyptian cases, and correlated positively with a favorable prognosis at the one year follow-up. During this same follow-up period, 64% of the cases had fully remitted. The social and clinical outcomes of the cases were not correlated with family psychiatric history (Okasha, Seif El Dawla, Khalil, & Saad, 1993).

The results of Okasha and colleagues agree with other studies that show similar psychotic syndromes among other non-Western cultures (Cooper, Jablensky, & Sartorius, 1990), and validate the presence of a special psychotic diagnostic category. This kind of psychosis is normally precipitated by stress, has an acute onset, and most patients recover rapidly. Associated symptoms are rather polymorphic, and include schizophrenic, affective, and neurotic symptoms.

El-Islam (1979) reported that Arab schizophrenics belonging to extended families had less malevolent manifestations, and were found to be less prone to deterioration into affectively blunted and withdrawn states than nuclear families.

Some delusional cultural beliefs, such as the devil, *Jinn*, sorcery, and the evil eye, may mislead the therapist who is not familiar with these beliefs (El-Islam, 1982). Much bizarre behavior which seems to be psychotic may occur: unacceptable wishes, feelings, and acts are liable to be projected onto the devil and appear as *wiswas*, or *jinn* possession. *Wiswas* is a rumination involving aggressive or sexual impulses that are attributed to the devil, enabling people to avoid responsibility and guilty feelings. *Jinn* possession permits acting out of forbidden drives, while claiming to be unaware of and not responsible for these acts (El-Islam, 1982).

Thin lines differentiate pathological delusions from cultural ones. Sometimes hallucinations may lose their perceptual quality and their contents are regarded as devil-induced thoughts. El-Islam (1982) noted that paranoid delusions may become normalized into ideas which are in keeping with beliefs about the evil eye or sorcery.

10.13.4.2 Psychometric Characteristics

Few studies have addressed psychometric aspects among Arabs. Ibrahim (1977, 1979, 1982) found some psychometric features among Egyptians. He found Egyptian students to be more dogmatic than Western students (Ibrahim, 1977), and less extraverted than Americans and British (Ibrahim, 1979). He found that cultural features such as oversensitiveness, superstition, and concern for socially sanctioned behavior and social desirability were more important personality factors among Egyptians than the neuroticism, psychoticism, and extraversion which are more common cultural features in the West (Ibrahim, 1982). The latter three factors do also apply in Arab societies, however (Farrag, 1987).

Based on clinical experience and field research, the present author has found the following psychometric characteristics of Palestinians" performance on some well-known (Western-developed) psychological tests:

(i) *Wechsler Intelligence Scale for Children-Revised* (*WISC-R*). Compared to Jewish students, Arab elementary school students in Israel had significantly lower IQ scores. The most common difficulties were in the similarities and comprehension subtests. The only subtest on which Arab students outperformed Jewish students was information. These differences may be at least partially explained by the culturally alien test items given to Arab students. In addition, these scores reflect the Arabic educational climate that encourages rote learning and discourages analytical thinking and the decision-making process.

(ii) *Bender-Gestalt and draw-a-person*. In both tests, performance of Arab students is similar to that of Americans until age seven, then the Arab scores show a comparative decline. At age 10–11, the performance of Arabs is equivalent to eight-year-old American children. This lag may be explained by the lack of encouragement for graphic activity by Arab parents and educators.

(iii) *Rorschach*. The performance of 50 Arab students, age 16, was compared to a similar age group of Westerners (reported by Levitt & Truumaa, 1972). Only slight differences were found. Arab students showed lower R (17.2 vs. 20.8), higher W (35% vs. 21.9%), lower P (4.5 vs. 5.3), and higher shading scores. All other scores were equivalent.

10.13.5 HELP-SEEKING BEHAVIOR

Generally speaking, Arabs tolerate their psychological distresses for a long time before seeking help. Among the Palestinian clients who sought help in the author's clinic in Nazareth, Israel, 44.2% came to therapy after years of hesitating, and another 44.2% came after months. Only 11% of the clients came after a couple of weeks, and 0.6% came without any delay (Dwairy, 1997b, 1998).

Arabs consider engaging in psychotherapy to be an admission of being crazy. Therefore, distressed people avoid this stigma by avoiding therapy. Furthermore, Arabs believe that physicians, native healers, religious figures, or family members are the proper alleviators of illnesses and distress. Seeing a psychiatrist or psychologist is a last resort (Okasha et al., 1994; Timimi, 1995).

Traditional and religious healers play a major role in primary psychiatric care in Arab countries. They use religious and group psychotherapies, and devices such as amulets and incantations. An estimated 60% of outpatients at the university clinic of Cairo, which serves patients of low socioeconomic levels, have sought treatment from traditional healers before coming to a psychiatrist (Okasha, 1993; Okasha, Kamel, & Hassan, 1968).

Arab families discourage long-term hospitalization or institutionalization for their patients. They feel they have the right and responsibility to be the caretakers of the unfortunate members of their family. Okasha (1993) reported that caring for an elderly demented person outside of the family is considered shameful; Egyptians, therefore, prefer to assimilate chronic mental patients. The parents of retarded or hyperactive children also feel a primary responsibility towards them rather than relinquishing them to institutionalized care (Okasha, 1993). Similar attitudes seem to exist in Qatar (El-Islam, 1982).

During hospitalization, Arab families do not discontinue their supportive role; on the contrary, family members and friends volunteer to attend to the patient. Groups of them stay with the patient for hours in the hospital, comforting him or her, bringing food from home, and offering their help. This cultural system of support seems unusual for the Western practitioner and may be considered intrusive. Meleis and La Fever (1984) reported that nurses and doctors in the USA often label Arab family members "anxious" and "intrusive," consider Arab patients "unpopular" or "chronic," and give them less time and personalized attention.

Males are currently approaching psychotherapy more than females. Among the clients of the author's clinic, 65.3% were males. Similarly, more males than females were represented in outpatient clinics in Cairo (Okasha et al., 1994). It seems that women may be more vulnerable to the stigma of psychotherapy. For example, the family does its best to hide, and manage at home, psychological disorders of females. They may even consider some disorders normal—such as anxiety and phobias. Males, on the other hand, are less vulnerable and could not be hidden at home. They are therefore referred more easily (El-Islam, 1969, 1982). It seems that psychiatric assistance is expedited for women only in severe cases, when the decision is finally made by the family. Therefore, women seem to be referred to hospitals more than men. The ratio between women and men in Saudi psychiatric hospitals was 1.6:1 (Rufaie & Mediny, 1991).

Arab clients typically are unaware of their mood and emotions. Initially, they present very supportive family relationships, and present their complaints in somatic terms, using metaphoric styles, such as "my heart feels like a dark room," or "my heart is dead" (Bazzoui, 1970; Timimi, 1995). Based on the belief that life, as well as the future, is "in the hands of Allah (God)," clients do not assume responsibility for their pathological actions (West, 1987). They place the responsibility for change on the therapist. They often appear silent, expecting the therapist to do all of the work. Because they regard comprehensive assessment as intrusive, they become uncommunicative when asked probing questions, and endeavor to absolve the family from responsibility for the illness. Their answers are often vague and unspecific because they question the relevance of this personal information to their health (Meleis, 1981; Meleis & La Fever, 1984). The concept of working to improve the self is generally nonexistent in Arab cultures (West, 1987). All these culture-bound behaviors are frequently misinterpreted by Western clinicians as resistance or lack of maturity (Dwairy, 1997a, 1998).

Typically, Arab patients expect an instant cure and assume that medication by intrusive methods (injections rather than pills) is most effective (Meleis & La Fever, 1984). They do not believe that talk therapy is worth paying for, unless it is direct advice. Long-term and nondirective therapy may confuse and frustrate the patient (Dwairy & Van Sickle, 1996; Gorkin et al., 1985; West, 1987).

Despite the above, psychotherapists are considered an authority in Arab societies. Influenced by sexist attitudes, however, Arab clients prefer and respect male rather than female therapists. Female therapists experience greater difficulties. Racy (1980) reported that even Saudi Arabian women preferred a male therapist.

He also described in detail the behavior of Saudi female patients. He reported that typically they would come to the clinic veiled and accompanied by a male member or several members of the family. They would limp in, leaning on the arms of a male relative. Dropping heavily into a chair, they may remain quiet and veiled until spoken to. Their demeanor reflects passivity to an extreme. Initially, they may say nothing and refer the question to the accompanying relative. Shoulders are frequently shrugged and "I don't know" is a common answer. They will answer only when pressed. Interpersonal and psychological difficulties are usually denied, and the therapist is assured that all is well and everyone is happy. Anger is transformed into sadness, disappointment, and/or self-blame, particularly when the object of anger is present. The patient may express her problem in somatic terms that prevent her from attending to her household duties and force her to spend much time resting. She does not assume responsibility for her illness nor does she accuse anyone else. She expects to be examined and prescribed a medication. Injections and large, colored pills are preferred (Racy, 1980).

In other Arab countries, especially in urban populations, the behavior of Arab clients may be very different from that of the Saudi example. Racy (1977) reported that among urbanized Arabs, the behavior of depressive patients is approaching that in the West. Among the Palestinian clients of the author's clinic, 29.6% came without informing their families, and 12.9% came in spite of family opposition (Dwairy, 1997b, 1998).

It is difficult to maintain professional relationships and boundaries with Arab clients. They often prefer to discuss problems in a social visit to their homes rather than in a formal session at the clinic. They may come late to the session or appear on the wrong date, or they may invite the therapist to the family home, or to a meal with the family. The therapist's negative response to such behavior is likely to be seen as insulting.

Although many clinicians have reported the difficulties of Arab clients in long-term therapy, no information has been reported about the attrition rate. In the author's clinic, only 15% remained in therapy for more than 10 sessions, and 49% dropped out after they achieved initial relief during 5–10 sessions. The rest were referred to other kinds of treatment (hospitals or psychiatrists) after the intake session or after reaching an impasse (Dwairy, 1997b, 1998).

10.13.5.1 Traditional Arab Theory of Psychotherapy

Germ theory and the evil eye have long provided explanations for illness (Meleis, 1981). Traditional Arabic beliefs about the origin of psychological problems frequently have linked them to supernatural causes: God, sorcery, or the evil eye. Many Arabs believe that evil thoughts could be transferred to another person through witchcraft or directly through the eye. The Arabic word for madness is *Junoon*, derived from the word *Jinn*, meaning evil spirit (Timimi, 1995). "Bad nerves" or "failure to eat" are other common explanations (Budman et al., 1992).

To undo the effects of *Jinn*, sorcery, or the evil eye, some rituals may be performed. Examples are antisorcery and antienvy rituals, including the use of amulets containing verses of the *Qura'an* (Muslim holy book), fumigation with incense, visits to the tombs of religious sheikhs, and purification rituals that involve drinking or washing in water that has been washed off *Qura'anic* verses written on a plate (El-Islam, 1982).

Another common ritual is the *Zar* ceremony. In this ceremony, participants experience states of trance under the influence of drumming, chanting, and dancing, during which they express worries and wishes that are otherwise socially prohibited. Each song is directed by the *Zar* healer to a particular spirit which, if present in the body of a member of the audience, will speak through him. It is to the spirit that strange wishes are attributed. Such wishes must be answered, for example, by sacrifices to placate the spirit and to persuade it to leave the possessed body. Birds, rams, and lambs are often slaughtered as sacrifices, and parts of them are then eaten by the person to be exorcized. Blue beads or figures involving the number five, such as hand symbols, are frequently used to protect people from the evil eye, sorcery, and *Jinn* (El-Islam, 1982).

In adddition to the aforementioned treatment types, marriage is often regarded as "the cure" for Arab patients. Schizoid, retarded, and depressed individuals may be forced into marriage by family elders on the assumption that marriage will take them out of their seclusion and improve their health, an assumption which is often disproved when they experience psychotic episodes soon after marriage (El-Islam & El-Deeb, 1968).

10.13.6 EPIDEMIOLOGY OF CLINICAL PROBLEMS

Epidemiological psychiatric studies are limited or localized in Arab countries, which makes it difficult to reach comprehensive conclusions without the risk of inappropriate generalizations. The main findings currently available are outlined below.

10.13.6.1 Psychiatric Morbidity

Based on Western measures, many studies have reported higher psychiatric morbidity among Arabs than Westerners. Ibrahim and Ibrahim (1993) reported higher neuroticism and lower extraversion scores among Egyptians compared to Americans and British. They also reported higher social anxiety and shyness among Libyan students than a comparable Western cultural group. The high level of anxiety and the tendency to develop neurotic disorders under stress were attributed to the social constraints imposed on individuals in Arabic societies (Ibrahim, 1979; Ibrahim & Ibrahim, 1993).

Based on the clinical interview schedule, the point prevalence rate of psychiatric morbidity among patients of a primary health care center in the United Arab Emirates was 27.6% (El-Rufaie & Absood, 1993). A high prevalence (26%) was also found in Saudi Arabia (El-Rufaie, Albar, & A-Dabar, 1988) and Dubai (22.7%) (Ghobash & Bibbington, 1994; Ghobash, Hamdi, & Bibbington, 1992). These rates were higher than those found in four other developing countries (Colombia, India, Sudan [an Arabic country], and the Philippines), in which the prevalence was between 10.6 and 17.7% (average 13.9%) (Harding et al., 1980). Some studies have indicated that the high morbidity in Saudi Arabia and the Gulf states is associated with sociocultural changes (Ibrahim & Al-Nafie, 1990).

In Dubai, psychiatric morbidity was found to be increased among women when their attitude and behavior departed considerably from each other. For example, psychiatric morbidity was particularly high for women harboring traditional beliefs but behaving less conventionally (Ghobash & Bibbington, 1994; Ghobash et al., 1992).

Conflicts between the traditional collectivistic authoritarian values and the individualistic liberal ones contribute heavily to the distress of Arabs. According to El-Islam (1979, 1982), intergenerational conflicts precipitated 50% of suicide attempts, 20% of neuroses, and 17% of schizophrenic illness among young Arab patients. Among the clients of the author's clinic, 18.1% of them exhibited client–family conflict by bringing it into therapy at the beginning, and 47.6% brought it during the course of therapy. Only in 34.3% of the cases had this familial conflict not emerged at all in therapy (Dwairy, 1997b, 1998).

Women and children were found to be more vulnerable to psychiatric illness than other social groups. Although men approached clinical help easier than women, studies indicate higher morbidity among women in the United Arab Emirates (31.9% female vs. 20.3% male) (El-Rufaie & Absood, 1993). Women reported more symptoms of anxiety and depression (Ibrahim, 1991), and neuroticism (Khalik & Eysenck, 1983). Although the general morbidity in Saudi Arabia was higher among men than women (El-Rufaie et al., 1988), more women were represented among hospitalized patients in Saudi Arabia (Rufaie & Mediny, 1991), suggesting the existence of covert morbidity among Saudi women. The unmarried, widowed, separated, polygamously married, and childless women were at high risk for psychiatric morbidity in Dubai (Ghobash et al., 1992), and for somatization disorder in Saudi Arabia (El-Islam, 1982).

It seems that in less traditional Arabic societies than Saudi Arabia, men rather than women tend to express their complaints in somatic terms. Among the Palestinian clients of the author's clinic, for example, men were more likely to express their distress by physical symptoms, whereas women tended to express their feelings of anxiety, fear, and depression directly. Expressions of anxiety, fear, and depression are normally considered signs of weakness in Arab societies. Therefore, it is more accepted among females, but detracts from the manhood of the Arab male, so somatic complaints may be more legitimate and less "damaging" for him (Dwairy, 1997b, 1998).

Children are also a higher risk group. Children represent 65% of outpatient visits in Saudi Arabia; almost 35% of them showed psychiatric disturbances that warranted intensive care (Ibrahim & Ibrahim, 1993; Tuma, 1989). Children in Sudanese nuclear families had more conduct, emotional, and sleep problems, and were more likely to be overdependent than those living in extended families (Al-Awad & Sonuga-Barke, 1992).

Many areas in the Middle East such as Lebanon, Palestine, Iraq, Kuwait, and Sudan, have been fraught with extreme and continuous stressors in the last decade, because of war and occupation. Many studies have indicated a high prevalence of psychological disorders in those area—especially among children. Of Lebanese children, 20–30% have exhibited at least one of the following types of behavioral problem: hyperactivity, overdependence, aggression, depression, or social misbehavior, and 58% have exhibited psychosomatic problems (Chimienti & Abu-Nasr, 1992–93; Nassar, 1991). A high prevalence of psychological disorders has been reported in Palestine (Baker, 1990), Sudan, and Iraq (Raundalen & Melton, 1994).

10.13.6.2 Distribution of the Disorders

A significant number of studies have shown that somatization is the main form of psychological complaint, and therefore the most common diagnosis, in Iraq (Al-Issa & Al-Issa, 1969), Egypt (Isaui, 1994), Kuwait (Parhad, 1965), Palestine (Dwairy, 1997b, 1998; Gorkin et al., 1985), and Saudi Arabia (El-Islam, 1982; Racy, 1980; West, 1987).

The most common diagnoses encountered in the United Arab Emirates have been neurotic depression (55% of the cases), anxiety-depressive disorders (13.3%), and anxiety disorders (11.7%). Anxiety-depressive disorders have been more common in females. Neurotic depression and anxiety-depression disorders were most prevalent in the middle-aged group (35–54 years), while anxiety disorders were most common in the young age group (15–34 years) (El-Rufaie & Absood, 1993).

Depressive and related symptoms were common among Saudi university students (Ibrahim & Al-Nafie, 1991). Among these symptoms were self-blame (64% of the students), inability to concentrate well (54%), shyness (53%), shivering and shaking in response to interpersonal distress (42%), and academic problems (33%). A similar epidemiological study in Dubai showed that the prevalence of depression (13.7%) was higher than anxiety (7%) or mania and psychotic disorders (1.9%). Depressed mood was present in two-thirds of the anxiety cases, and autonomic anxiety was present in over half the cases of depression (Ghobash & Bibbington, 1994; Ghobash et al., 1992).

In Egypt, Okasha (1993) reviewed the diagnoses given to 800 patients admitted to the psychiatric center in Ain Shams in 1991 and revealed the following rates of disorders: mood disorders (18.3%); schizophrenia (16.1%); somatoform disorders (9.5%); substance-use disorders (8.5%); disorders of infancy, childhood, and adolescence (8.2%); and anxiety disorders (7.9%) (Okasha, 1993). The most frequent diagnosis among the psychotic Egyptian patients was brief reactive psychosis, with polymorphic acute symptoms precipitated by stress, and with rapid recovery (Okasha et al., 1993).

In contrast to the above, Bazzoui (1970) reported low rates of depression (1.3% depressive and 2.9% manic patients) among 1120 patients admitted to a psychiatric hospital in Iraq. These contradictory results may be attributed to demographic differences in the population, to the methodology, or to changes that have occured since Bazzoui's 1970 report, when only severe cases were sent to psychiatric hospitals. Among the Palestinian clients of the author's clinic, 60% complained of anxiety, phobias, somatoform disorders, and obsessive-compulsive disorders, compared to 32.6% complaining of depression (Dwairy, 1997b, 1998).

Eating disorders (anorexia and bulimia) in Arabic societies have been rare or absent in the past (Al-Issa, 1966; El-Sarrag, 1968); however, since the late 1980s these disorders have seemed to be on the increase—especially among Arabs in social transition. Nasser (1986) found that 12% of the female Arab undergraduate students attending London universities were bulimic, but found no anorexia or bulimia cases in a similar female Arab student group in Cairo. The absence of anorexia may be related to the Arabic cultural attitude towards food. Arabs also associate plumpness with attractiveness and health (Timimi, 1995).

Sexual dysfunction has been rarely studied in Arabic societies. Okasha and Demerdash (1975) reported on 68 male Arabs (Kuwaitis, Palestinians, and Egyptians) who had suffered erectile disorders and/or premature ejaculation. Of these, 45% reported a past history of homosexuality. When asked about their own evaluation of their disorder, most of these patients attributed it to undersized or diseased genitalia. Few of them blamed supernatural causes or masturbation. Many of the patients had accompanying somatic complaints, mostly backache.

Among the clients of the author's clinic, 18% of the women, but only 7% of the men, complained of sex-related anxiety. This difference may be attributed to the sexual oppression of Arab women or to the difficulties of the Arab man in admitting his sexual problems (Dwairy, 1997b, 1998).

10.13.7 HEALTH SERVICE DELIVERY SYSTEM

The first mental hospitals in the world were built in Arab countries: first in Baghdad, Iraq, in 705 AD, followed by hospitals in Cairo (800 AD) and Damascus (1270 AD) (Okasha, 1993). Although Arabs were the first to develop psychiatric institutions, their psychiatric problems today are often referred either to native (usually religious) healers (Timimi, 1995), or to medical doctors, due to the somatic manifestations of their distress. When they do reach psychiatrists, they are provided with medications or other medical treatments rather than psychotherapy (Gorkin et al., 1985; Ibrahim & Ibrahim, 1993; West, 1987). Psychiatric treatment then typically occurs in one of two settings: the traditional mental hospital or an outpatient setting. Private psychiatric clinics are also an option for some. A few private corporations, universities, and religious organizations own and operate some health services that also provide psychiatric services (Ibrahim & Ibrahim, 1993); however, in most of these institutions psychotherapy *per se* is not represented. Variations of counseling, behavior therapy, and family therapy are sometimes employed.

Obviously, the mental health system in the Arab world barely reaches a small portion of the population that needs psychological attention. Marital counseling, parent counseling, vocational guidance, crisis intervention, crisis support groups, school psychology, and other mental health services are rare or absent in most Arab countries. Psychologists who are trained in any form of psychotherapy are rare, making the need for research on mental health in general, and training programs in particular, imperative (Ibrahim & Ibrahim, 1993).

In general, there has been a paucity of information about mental health services in Arab countries. An exception is Okasha's (1993) detailed report about the mental health system in Egypt, considered one of the most developed Arab countries. According to this report, only 450 psychiatrists serve 57 000 000 citizens of Egypt (1 per 127 000 citizens). There are about 8000 psychiatric beds (1 bed per 7000 people), 5750 of them in Cairo. Only 250 clinical psychologists were reported to be practicing in Egypt, with hundreds of general psychologists working in fields unrelated to mental health services. There are many social workers practicing in psychiatric facilities, but they are not trained in psychiatric social work (Okasha, 1993).

Out of 24 governorates of Egypt, 19 have some kind of psychiatric clinic and outpatient unit, while five have no psychiatric services. Community care in the form of hostels, day centers, and rehabilitation centers is available only in Egypt's big cities. In the rural areas, community care is implemented without health care workers. Chronic and mild psychiatric patients are "rehabilitated" through cultivating and planting the countryside, under the supervision of family members. The programs for community care in big cities take the form of outpatient clinics, hostels for the elderly, institutions for the mentally retarded, centers for drug abuse, and school and university mental health. In Egypt, there are 13 medical schools, that for 35 years have offered school master's and doctoral programs in psychiatry (Okasha, 1993).

The denial of the role of emotions in physical health is common not only among patients, but among physicians, as well. That somatic symptoms could be of psychological origin is beyond the imagination of the average Arab patient and beyond the knowledge of many medical practitioners. Both physicians and patients are somatically oriented. When doctors fail to locate disease in the body, they typically make the diagnosis of "general physical weakness" and prescribe vitamins and rest (El-Islam, 1982).

Fortunately, some countries are witnessing the nascent growth and dissemination of information bringing to the awareness of practitioners the psychological and social contributors to illness. The psychiatric center in Ain Shams, for example, emphasizes a biopsychosocial approach to outpatient care. In addition, they are encouraging intensive psychiatric outpatient programs in all general hospitals (Okasha, 1993).

The Palestinians in the Gaza strip and the West Bank, as well as inside Israel, have made significant progress in developing mental health services during the last few years. In the Gaza strip (an area of 800 000 citizens), one psychiatric hospital and one community mental health program direct three local clinics. Programs utilize a multidisciplinary approach that employs psychiatrists, psychologists, social workers, nurses, and other specialists. In the West Bank (1 500 000 citizens) there is one large psychiatric hospital (320 beds) and three related local clinics. In 1995, the Palestinian authority initiated professional development courses for physicians to increase their awareness of mental health issues. In addition, a mental health plan was initiated to meet the needs of schools and families adversely affected by the 29 years of Israeli occupation (Palestinian Authority Report, 1994–1995). Among the Palestinians who live in Israel (800 000 citizens), there has been significantly more progress in the psychological services provided to schools. At least 24 centers for psychological services are now providing mental health services to the Palestinian Arab schools in Israel (Israeli Ministry of Education Report, 1997).

Despite limited mental health services and the lack of information about psychotherapy among Arab populations, the number of Arabs seeking psychotherapy has been steadily increasing (Dwairy, 1997b, 1998; Gorkin et al., 1985; Okasha, 1993).

10.13.8 FUTURE DIRECTIONS

Compared to most people living in more individualistic, liberal, Western societies, Arabs are more dependent on their families, their personalities are less individuated, and they do not welcome autonomy or self-actualization. Therefore, an individual's repression is maintained by external factors rather than intrapsychic constructs such as an ego or superego. Implementing psychodynamic or other nondirective therapies, which are designed to deal with intrapsychic processes and to motivate the patient towards self-actualization, may result in serious problems. First, they miss the main source of repression, which is external rather than internal; and second, they activate repressed emotions that may not be allowed to be expressed in the Arabic society, such as sexual drives or aggression towards family members. Therefore, "successful" treatment that makes the unconscious conscious converts the intrapsychic conflict into a conflict between the client and his or her family, and may leave the client with an untreated wound (Dwairy, 1997a; Dwairy & Van Sickle, 1996).

To avoid these problems, therapists are encouraged to be flexible, tailoring the therapy to the client, rather than fitting the client into their favorite treatment approach. Therapists need to be able to alter their personal therapeutic style, and be open and willing to learn new approaches to treatment (Budman et al., 1992). An understanding of the client's cultural identity and family relationships is essential evaluative information in directing the therapist's choice and trajectory of treatment. For Westernized Arabs, such as some of the more educated peoples of major urban areas, a therapist may consider insight-oriented therapies, whereas for traditional Arabs, behavioral, directive, goal-oriented, and short-term therapies are recommended. With traditional clients, psychotherapy may include counseling and academic and career assistance (West, 1987). The client's ego strength is another important factor to evaluate. Before implementing any type of insight therapy that may evoke culturally forbidden needs and emotions, the therapist should be reasonably sure that the client is strong enough to participate meaningfully in the conflicts that will emerge between the client and family.

Because Arab patients have many misunderstandings about psychotherapy, West (1987) suggested that issues of trust, the therapist–patient relationship, and status differentials need to be discussed openly in the early stages of therapy. Orientation to psychotherapeutic processes prior to commencing treatment may facilitate the effectiveness of the treatment, as well (Dwairy & Van Sickle, 1996). After careful consideration of the above factors, a therapist may begin treatment in a directive mode and later shift gradually to a nondirective and insight-oriented made (Gorkin et al., 1985).

When a traditional Arab client is not individuated from his/her family, parental authority is never to be challenged or criticized (West, 1987). Joining the authority of the family and gaining its trust is crucial at the beginning, especially when the client is a woman (Dwairy & Van Sickle, 1996; Racy, 1980). Without the collaboration of the family, therapy may well be sabotaged. A supportive personal approach that acknowledges cultural norms may enhance the trust.

Egalitarian family therapy that focuses on encouraging the direct expression of feeling is frustrating, if not impossible, with traditional families. Such therapy may threaten the authoritatian structure of the Arab family, and terminate treatment. Promotion of separation from the family is considered disrespectful (Budman et al., 1992).

Therapists should learn about the culture (values, norms, prohibitions, taboos, etc.) from the client, without imposing his or her own values on the client. Therapists may enlist parts of the Arabic culture to accomplish therapeutic changes (Dwairy, 1998; Dwairy & Van Sickle, 1996). West (1987) suggested cooperating with folk healers to broaden therapists" repertoire of treatment modalities that have been proven effective in Arab societies.

The therapist should find a way to help the client within the limitations and specifications of the client's culture. Their culture should be considered an immutable, but important, piece of the therapeutic process. Racy (1980) suggested that the therapist avoid both of two extremes: the temptation to be so respectful of tradition that one becomes paralyzed, and the tendency to become a social reformer. A combination of warmth, authority, diplomacy, and verbal facility can go far toward establishing an effective alliance.

Based on the sociocultural background of Arabs, characterized by the somatization of psychosocial distresses, and person–family interdependence, the flexible implementation of a biopsychosocial approach to treatment, involving the client's family, may be promising for the majority of Arab clients (Dwairy, 1997c).

10.13.9 SUMMARY

The Arabic culture is characterized by authoritarian and collective values rather than liberal and individualistic ones. The interpersonal relationship is both supportive and distressing. The locus of control is external, and the expression of emotions is indirect. An individual's self and identity are enmeshed in the collective identity.

Unfamiliarity with Arabic culture may lead to misinterpreting and pathologizing the normal behavior of Arabs. The manifestation of psychological disorders among Arabs is different from that in the West. Depression is characterized by somatic rather than mood changes. Guilt feelings, worthlessness, and suicidal thoughts and attempts among Arab depressive patients are rare or absent.

Obsessive-compulsive symptoms are influenced by and integrated with Islamic religious rituals. Delusional cultural beliefs such as the devil, sorcery, and the evil eye may overshadow the pathological delusions. Many psychotic Arab patients have psychogenic disorders that were precipitated by stress, have an acute onset, and, in most cases, show rapid recovery.

A few psychometric characteristics have been reported. Egyptians have been shown to be less extraverted, and more oversensitive, superstitious, and concerned for socially approved behavior. Due to sociocultural factors, Palestinian children in Israel have performed lower on WISC-R IQ tests compared to Jewish children. They also lag behind American children on graphic tests such as the Bender-Gestalt and Draw-a-person tests.

Arabs consider that there is a stigma attached to approaching psychotherapy. They seek psychotherapy after years of delay or after seeking help from traditional healers or physicians. It seems that men approach outpatient clinics more than women, who are often "helped" at home until they have severely deteriorated and are referred to hospitals.

Arab patients are typically unaware of their mood, do not feel responsible for their situation, and are reluctant to reveal family matters. They expect direct advice or medication. They attribute their symptoms to supernatural powers, germs, or the lack of food or rest.

High morbidity has been reported in Arab countries, especially in the areas of dramatic social transition. Women and children are at high risk. The mental health system in Arab countries barely reaches a small portion of the population. In rural areas it is almost absent. Somatic complaints are the most common presentation, and medication is the most common treatment offered. Psychotherapy is offered by only a few mental health facilities.

Insight and nondirective therapies, that emerged in the more individualistic West, encounter cultural barriers. Cultural discouragement of individuation and autonomy and the expression of emotions make psychotherapies that encourage emotional expression and self-actualization difficult. The cultural identity of the client and his or her family, as well as ego strength, are important factors that may help the therapist to tailor the therapy to the case. Collaboration with the family is vital for successful therapy. Learning the cultural background and the client's expectations is important to the therapeutic contract. Orientation to psychotherapy helps to make the client more engaged in the therapeutic process. Flexible, integrative approaches seem to be suited to the majority of Arab clients.

10.13.10 REFERENCES

Al-Awad, A. M. E., & Sonuga-Barke, E. J. S. (1992). Childhood problems in a Sudanese city. A comparison of extended and nuclear families. *Child Development, 63,* 906–914.

Al-Issa, I. (1966). Psychiatry in Iraq. *British Journal of Psychiatry, 112,* 827–832.

Al-Issa, I., & Al-Issa, B. (1969). Psychiatric problems in a developing country: Iraq. *International Journal of Social Psychiatry, 16,* 15–22.

Al-Khani, M. A. F., Bebbington, P. E., Watson, J. P., & House, F. (1986). Life events and schizophrenia: A Saudi Arabian study. *British Journal of Psychiatry, 148,* 12–22.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Baasher, T. (1962). Some aspects of the history of the treatment of mental disorders in the Sudan. *Sudan Medical Journal, 1,* 44.

Baker, A. M. (1990). The psychological impact of the intifada on Palestinian children in the occupied West Bank and Gaza: An exploratory study. *American Journal of Orthopsychiatry, 60,* 496–505.

Barakat, H. (1993). *The Arab world: society, culture, and state.* Berkeley, CA: University of California Press.

Bazzoui, W. (1970). Affective disorders in Iraq. *British Journal of Psychiatry, 117,* 195–203.

Budman, C. L., Lipson, J. G., & Meleis, A. I. (1992). The cultural consultant in mental health care: The case of an Arab adolescent. *American Journal of Orthopsychiatry, 62*(3), 359–370.

Chimienti, G., & Abu-Nasr, J. (1992–93). Children's reactions to war-related stress. II. The influence of gender, age, and the mother's reaction. *International Journal of Mental Health, 21*(4), 72–86.

Cooper, J. E., Jablensky, A., & Sartorius, N. (1990). WHO collaborative studies on acute psychosis using the SCAAPS schedule. In C. N. Stefanis (Ed.), *Psychiatry: A world-wide perspective.* New York: Elsevier.

Dwairy, M. (1997a). Addressing the repressed needs of the Arabic client. *Cultural Diversity and Mental Health, 3*(1), 1–12.

Dwairy, M. (1997b). *Personality, culture, and Arabic society.* Jersusalem: Alnoor [Arabic.]

Dwairy, M. (1997c). A big psychosocial model of metaphor therapy with holistic cultures. *Clinical Psychology Review, 17*(7), 719–732.

Dwairy, M. (1998). *Cross-cultural counseling: The Arab-Palestinian Case.* New York: Haworth Press.

Dwairy, M., & Van Sickle, T. (1996). Western psychotherapy in traditional Arabic society. *Clinical Psychology Review, 16*(3), 231–249.

El-Islam, M. F. (1969). The educational and occupational correlates of psychiatric disorder. *International Journal of Social Psychiatry, 15,* 288.

El-Islam, M. F. (1979). A better outlook for schizophrenics living in extended families. *British Journal of Psychiatry, 135,* 343.

El-Islam, M. F. (1982). Arabic cultural psychiatry. *Transcultural Psychiatric Research Review, 19,* 5–24.

El-Islam, M. F., & El-Deeb, H. A. (1968). Marriage and fertility of psychotics. *Social Psychiatry, 3,* 24.

El-Rufaie, O. E., & Absood, G. H. (1993). Minor psychiatric morbidity in primary health care: Prevalence, nature and severity. *International Journal of Social Psychiatry, 39*(3), 159–166.

El-Rufaie, O. E., Albar, A. A., & A-Dabal, B. K. (1988). Identifying anxiety and depressive disorder among primary care patients: A pilot study. *Acta Psychiatrica Scandinavica, 77,* 280–282.

El-Sarrag, M. E. (1968). Psychiatry in the Northern Sudan:

A study in comparative psychiatry. *British Journal of Psychiatry, 114,* 946–948.

Farrag, M. F. (1987). Dimensions of personality in Saudi Arabia. *Personality and Individual Differences, 8*(6), 951–953.

Fattal, H., & Sokari, R. (1988). *Contemporary Arab history* Beirut, Lebanon: Gros Bryce. [Arabic].

Ghobash, E. H., & Bibbington, P. (1994). The Dubai community psychiatric survey: Acculturation and the prevalence of psychiatric disorder. *Psychological Medicine, 24,* 121–131.

Ghobash, E. H., Hamdi, E., & Bibbington, P. (1992). The Dubai community psychiatric survey: I. Prevalence and sociodemographic correlates. *Social Psychiatry and Psychiatric Epidemiology, 27,* 53–61.

Gorkin, M., Masalha, S., & Yatziv, G. (1985). Psychotherapy of Israeli-Arab patients: Some cultural considerations. *Journal of Psychoanalytic Anthropology, 8*(4), 215–230.

Harding, T. W., De-Arango, M. V., Baltazar, J., Climent, C. E., Ibrahim, H. H. A., Ignacio, L. L., Murthy, R. S., & Wig, N. N. (1980). Mental disorders in primary health care: A study of their frequency and diagnosis in four developing countries. *Psychological Medicine, 10,* 231–241.

Ibrahim, A. S. (1977). Dogmatism and related personality factors among Egyptian university students. *Journal of Psychology, 95,* 213–215.

Ibrahim, A. S. (1979). Extraversion and neuroticism across cultures. *Psychological Reports, 44,* 799–803.

Ibrahim, A. S. (1982). The factorial structure of the Eysenck personality questionnaire among Egyptian students. *Journal of Psychology, 112,* 221–226.

Ibrahim, A. S., & Al-Nafie, A. (1991). Perception and concern about sociocultural change and psychopathology in Saudi Arabia. *Journal of Social Psychology, 13,* 179–186.

Ibrahim, A. S., & Ibrahim, R. M. (1993). Is psychotherapy really needed in nonwestern cultures? The case of Arab countries. *Psychological Reports, 72,* 881–882.

Ibrahim, R. M. (1991). Sociodemographic aspects of depressive symptomatology: Cross-cultural comparisons. *Dissertation Abstracts International, 51*(12).

Isaui, A. (1994). *Psychosomatic disorders*. Beirut, Lebanon: Dar El-Nahda Al-Arabia. [Arabic].

Israeli Ministry of Education Report. Psychological services department (1997).

Khalik, A. M., & Eysenck, S. B. G. (1983). A cross-cultural study of personality: Egypt and England. *Research in Behavior and Personality, 3,* 215–226.

Levitt, E. E., & Truumaa, A. (1972). *The Rorschach technique with children and adolescents: Application and norms*. New York: Grune & Stratton.

Mansfield, P. (1991). *A history of the Middle East*. New York: Penguin.

Meleis, A. I. (1981, June). The Arab American in the health care system. *American Journal of Nursing,* 1180–1183.

Meleis, A. I., & La Fever, C. W. (1984). The Arab American and psychiatric care. *Perspectives in Psychiatric Care, 22*(2), 72–86.

Nassar, C. (1991). *The reality of war and its reflections on the child*. Tarablus, Lebanon: Gros Bryce. [Arabic].

Nasser, M. (1986). Comparative study of prevalence of abnormal eating attitudes among Arab female students of both London and Cairo universities. *Psychological Medicine, 16,* 621–625.

Okasha, A. (1993). Psychiatry in Egypt. *Psychiatric Bulletin, 17,* 548–551.

Okasha, A., & Demerdash, A. (1975). Arabic study of cases of functional sexual inadequacy. *British Journal of Psychiatry, 126,* 446–448.

Okasha, A., Kamel, M., & Hassan, A. (1968). Preliminary psychiatric observations in Egypt. *British Journal of Psychiatry, 114,* 494.

Okasha, A., Saad, A., Khalil, A. H., Seif El Dawla, A., & Yehia, N. (1994). Phenomenology of obsessive-compulsive disorder: A transcultural study. *Comprehensive Psychiatry, 35*(3), 191–197.

Okasha, A., Seif El Dawla, A., Khalil, A. H., & Saad, A. (1993). Presentation of acute psychosis in an Egyptian sample: A transcultural comparison. *Comprehensive Psychiatry, 34*(1), 4–9.

Palestinian Authority Report. Ministry of Health (1994–1995).

Parhad, L. (1965). The cultural–social conditions of treatment in psychiatric out-patient department in Kuwait. *International Journal of Social Psychiatry, 11,* 14–19.

Racy, J. (1977). Psychiatry in Arab East. In C. L. Brown & N. Itzkowitz (Eds.), *Psychological dimensions of Near Eastern studies (pp. 279–329)*. Princeton, NJ: Darwin.

Racy, J. (1980). Somatization in Saudi women: A therapeutic challenge. *British Journal of Psychiatry, 137,* 212–216.

Raundalen, M., & Melton, G. B. (1994). Children in war and its aftermath: Mental health issues in the development of international law. *Behavioral Science and the Law, 12*(1), 21–34.

Rufaie, O. E. F., & Mediny, M. S. (1991). Psychiatric inpatients in a general teaching hospital: An experience from Saudi Arabia. *Arab Journal of Psychiatry, 2,* 138–145.

Sharabi, H. (1975). *Introductions to studying Arabic Society*. Jerusalem: Salah Eldeen. [Arabic.]

Sharabi, H. (1977). Impact of class and culture on social behavior: The feudal bourgeois family in Arab society. In L. C. Brown and N. Itzkowitz (Eds.), *Psychological dimensions of Near Eastern studies*. Princeton, NJ: Darwin.

Timimi, S. B. (1995). Adolescence in immigrant Arab families. *Psychotherapy, 32,* 141–149.

Tuma, J. M. (1989). Mental health services for children: The state of the art. *American Psychologist, 44,* 188–199.

West, J. (1987). Psychotherapy in the eastern province of Saudi Arabia. *Psychotherapy, 24*(1), 8–10.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.14

Perspectives from Lithuania

DANUTĖ GAILIENĖ
University of Vilnius, Lithuania

10.14.1 INTRODUCTION 325
10.14.2 THE BEGINNINGS OF PROFESSIONAL PSYCHOLOGY 326
10.14.3 THE SITUATION OF PSYCHOLOGY UNDER SOVIET OCCUPATION 327
10.14.4 THE BEGINNINGS OF CLINICAL PSYCHOLOGY 328
10.14.5 REGAINING INDEPENDENCE 331
10.14.6 PRESENT SITUATION AND PROBLEMS 332
10.14.7 REFERENCES 334

10.14.1 INTRODUCTION

The situation and development of clinical psychology in Lithuania is inextricably related to the sociopolitical situation of the country and its rather dramatic history. This chapter is an attempt to show that all the historical and political changes have directly influenced clinical psychology in Lithuania.

Lithuania is situated near the Baltic Sea; it presently has a population of approximately 3.7 million. Around 3000 BC the first wave of Indo-Europeans appeared in these places, and the group of Baltic peoples formed. The Balts included Lithuanians, Latvians, and Prussians. The latter were destroyed by the German Order of the Knights of the Cross or assimilated by them a few centuries ago. Lithuanians now live in the neighborhood of Finno-Ugrians (Estonians, Finns) and Slavs (Russians, Poles). Lithuanians were the only Balt people to found a powerful state of their own in the middle ages, the Grand Duchy of Lithuania. This state survived up to the late eighteenth century. As it grew weaker, it was annexed to the Czarist Russian Empire. The following centuries demanded from Lithuanians a determined resistance to different occupations, threatened assimilation, and dissolution.

Let us consider what was happening in Lithuania at the time when clinical psychology appeared in the West, that is when Lightner Witmer founded a "psychological clinic" at Pennsylvania University, Emil Kraepelin began experimental psychological study in a psychiatric clinic, and Sigmund Freud was working out his theory and practice of psychoanalysis. At that time Lithuania constituted part of the Czarist Russian Empire. The only Lithuanian university, which was actually the oldest in Eastern Europe, founded in 1579 (Bumblauskas, 1994), had been closed for almost a century. This had been the reaction of the Czarist government to the uprising of 1831 against the imperial regime. After an even more powerful Lithuanian and Polish uprising in 1863, the reaction was still grimmer: the Lithuanian press was prohibited, Lithuanian schools were closed, even the Latin characters used in Lithuanian writing were prohibited; only the Cyrillic (Russian) script was allowed. Hundreds of people were murdered or deported to long years of exile in Siberia. The others resisted with determination the threatened national annihilation, until the press ban was abolished in 1904. This was the period of the "book-hawkers," the smugglers who secretly carried over the border Lithuanian periodicals and books published in East Prussia or elsewhere in the West. The so-called "hardship schools" appeared at this time, where country women would secretly teach children to read and write in Lithuanian.

10.14.2 THE BEGINNINGS OF PROFESSIONAL PSYCHOLOGY

It was only after the First World War that Lithuania was re-established as an independent state. The country witnessed a vivid development in all spheres of life: economic, cultural, and scientific. Since the capital city, Vilnius, was occupied by Poland in 1919, the University of Lithuania was founded in the temporary capital of Kaunas. There had been no professional psychologists in Lithuania up to that time, only students of psychology or representatives of other professions, such as physicians, biologists, and the clergy, who showed interest in the field (e.g., M. Reinys, J. Steponavičius, J. Vabalas Gudaitis, V. Lazersonas). The reason was that:

> The Lithuanian intelligentsia were not able to get jobs in Lithuania; most often they had to work in Russia. Only priests and representatives of the free professions—physicians and lawyers—could stay in Lithuania. Therefore Lithuanian intelligentsia would usually choose these professions. (Šapoka, 1989, p. 523)

The new university also had a department of pedagogy and psychology. Some specialists, for example, one of the most prominent pioneers of professional psychology in Lithuania, Alfonsas Gučas, having completed their studies here, went abroad to enrich their qualifications with the leading specialists of the time. In this way the first generation of professional psychologists came into existence. They taught psychology and organized practical and research institutions, created the system of teaching psychology, started experimental studies, actively participated in the press, and published and translated books on psychology. There also appeared what could be called the rudiments of clinical psychology.

In 1931 the Lithuanian Society of Psychotechnics and Vocational Guidance was founded, including a consulting department headed by the psychiatrist, Juozas Blažys (Gučas, 1937).

Some psychiatrists were interested in psychology and psychotherapy, and would apply methods of suggestion, hypnosis, and rational psychotherapy in their practice; they also took an interest in the newest trends and theories of psychotherapy. They applied their efforts to spreading and providing a critical evaluation of the new ideas, theories, and methods of clinical psychology, as well as towards educating society with regard to the prophylaxis of mental disease and to nurturing general psychic health.

As early as 1921 Blažys, who was the first professor of psychiatry in Lithuania and the author of the first Lithuanian psychiatry textbook (*Introduction to psychiatry*; Blažys, 1935), published an article entitled "Psychoanalysis and psychotherapy" in the journal *Medicina* (Blažys, 1921). It was a conceptual, exhaustive study, appearing in the early stages of the development of the new Lithuanian state and its medicine. The author stressed the necessity of psychotherapy, presented in sufficient detail the psychoanalytic theory of Freud and his methods, and maintained that "the time has passed for being content with mere bromides," and that psychological theory and practice concerning mental disorders are indispensable. Psychoanalysis occupied the central place in the treatment of neuroses, being a method based upon the psychogenic theory of the origins of neuroses and the principle of causality. Comparing psychoanalysis to other prominent methods of psychotherapy, the author asserted that in psychoanalysis, as in the rational psychotherapy of Dubois, explaining their situation to the patient is related to persuasion, and the therapeutic effect depends on the moral influence the doctor exerts upon the patient. Even though Blažys was a representative of biological psychiatry, in the *Introduction to psychiatry* he presented not only the biological aspects of psychic diseases, but also the latest tendencies of psychotherapy and clinical psychology: he discussed Freud's psychoanalysis and Adler's individual psychology, pointing out that these developments had illuminated the hitherto unpenetrated "secret" depths of the human mind. They had also demonstrated that not only in psychopathological disorders, but also in normal psychic life a very important role was played by the deeper psychic layers, and human feelings and drives. The author also introduced behaviorist views, at the same time pointing out their limitations: the analysis of psychopathology cannot ignore the subjective factors. As already mentioned, Blažys aligned himself with the neurological trend; however, he criticized physicians for their "medical materialism," and invited them to take a new look at man, to be aware of and seek new ideas, and also encouraged them to take up philanthropy. Eventually, the same author published a book entitled *Tolerance as a cultural principle* (Blažys, 1936), which was to be banned in Lithuania throughout the years of Soviet occupation for its strictly negative evaluation of totalitarian regimes, such as fascism and bolshevism.

Other methods of psychotherapy were also discussed, both in the specialist and popular press: autogenic training, hypnosis, and rational psychotherapy, as well as Jung's system of psychotherapy. Most publications could be characterized as adhering to the pragmatic view that no single theory can explain all of psychopathology, while each may be useful in a certain aspect. The questions widely discussed comprised the training of professional psychotherapists, problems of pediatric psychology, and the task of prophylaxis of mental disorder.

During the 20 years of independence the first professional psychologists were trained in Lithuania, and the period saw what could be called the beginnings of clinical psychology. The treatment of mental disorders did not comprise only biological methods but included psychotherapy as well; Lithuanian society was being informed and its psychological background enriched. Psychological knowledge was spread by people of various vocations—psychologists, medical specialists, Catholic priests and bishops.

All these achievements were doubtless very modest. But the span of time was also brief: 20 years is not very long in the life of a state, especially one which has started from nothing. However, the goal of building a state of their own mobilized enormously the activity and initiative of the people. With amazing determination they sought knowledge, high culture, and a more profound education. Even though the economic situation was difficult, young people studied foreign languages and made efforts to go to the best universities abroad, and upon returning they embarked on the task of creating a scholarship and culture of their own. Obstacles and hardships would foster character and a constant readiness to accomplish something which seemed impossible. Even though short, the period of the independent state helped an entire generation to cherish initiative, feelings of their own value, and responsibility for the state. A rather interesting point was made by the historian Danys in 1986. Writing about the emigrants from Eastern Europe escaping from the horrors of the Second World War, nazism, and communism, she notes that Lithuanians and emigrants from other independent states of the period differed greatly from the Soviet people who had lived for the same span under totalitarian regimes. The former generally had more initiative, and were more responsible and braver people since, in the author's view, they already possessed the experience of building their own state.

All these factors were also important in surviving another disaster that befell the country, that is the 50-year-long Soviet occupation. But even under the conditions of occupation, clinical psychology did gradually develop in Lithuania.

10.14.3 THE SITUATION OF PSYCHOLOGY UNDER SOVIET OCCUPATION

The existence of the independent state was again interrupted by subsequent occupations, first by the Soviets in 1940, then by the Nazis, and once again by the Soviets after the Second World War.

This had detrimental effects for the state in every way, and certainly interfered with the development of psychology; the humanities, as compared to the sciences, were especially strictly restrained. At the beginning of the Soviet period, the specialty of psychology still existed at Vilnius University; there was a department of psychology, and later a joint department of pedagogics and psychology, but these were soon closed, and the specialty abolished. A course for teachers of psychology, logics, and the Lithuanian language still existed for a time at the Vilnius Pedagogical Institute, but from 1956 even this training was stopped.

Soviet reality established itself firmly. A kind of Orwellian world crystallized, of which one of the strictest requirements was that of ideological purity and loyalty. All the humanitarian disciplines—philosophy, history, psychology, literary scholarship—had to be based not upon "idealist conceptions," but upon the foundations of dialectic materialism, whereas the "doctrines alien to our view" had to be abandoned. As in Orwell's world, a special language developed that for an outsider would sometime be very difficult to grasp. "Bourgeois ideology and idealist conceptions" were Western, non-Marxist creations; "obsequiousness before the West" and "views alien and unacceptable to us" were crimes that would not only draw criticism but could also possibly lead to imprisonment, Siberian exile, or even death. Such a perspective also confronted psychologists, and other specialists who contributed to the psychological culture of the country. Some perished, others were deported to Siberia, while some had managed to emigrate to the West before the "Iron Curtain" closed. The psychologists remaining in Lithuanian had to be extremely cautious, like all people working in the humanities. Most often psychological work would be restricted to comparatively "innocent" topics, remote from ideology, such as investigations of perception, psychophysiological measurements, and so on. Psychology was impoverished. It was "sovietized," which means that it was affected by the bureaucratization, ideologization and centralization of learning and teaching (Bagdonas, 1991).

From the beginning of the 1950s, when Nikita Khrushchev became leader of the Soviet Union, a period of "warming" began. First of all, this meant a more or less freer cultural climate.

Thanks to the initiative of Professor Alfonsas Gučas, a psychologist of the prewar period, the specialty of psychology was reopened at Vilnius University, and the department of psychology was refounded within the faculty of history. Officially the department was expected to train specialists in "engineering and work psychology." It soon emerged that under the conditions of a planned economy, the work of psychologists in various enterprises was hardly possible and altogether senseless. Where everything is decided "from above"; where "some pretend to be working while others pretend to be paying them"; where there is no place for practically any initiative; where the principle of equalization is applied everywhere; when it is not so much the quantity, quality, and realization of the produced goods that are important, but rather prestated plans "accomplished and overfulfilled" in papers and reports, as well as the "freely adopted" obligations of "socialist competition"; where the universal and most important aim of everyone's efforts is to cheat Moscow and to "draw" from it as much as possible—then simple, sound common sense and cynical socialist resourcefulness is quite enough. Professional psychologists have nothing to do in these areas.

However, these matters were merely the usual obstacles and absurdities that one would constantly face. Undoubtedly, the re-establishment of the psychology department at Vilnius University was of immense importance to Lithuanian psychology and clinical psychology. The head of the department, Professor Gučas, and its teachers, his former students, were anything but adepts of Marxist psychology and ideology. A positive element was also the fact that Lithuania had been incorporated into the Soviet Union considerably later than other countries. Being closer to the West and of a more European orientation, the country resisted Soviet indoctrination more persistently. The department existed in a rather liberal atmosphere. The teachers sought to present modern psychology and its history objectively, to foster a critical view of all the trends and fashions in psychology. Among the first graduates of the department, a few began to take an interest in clinical psychology (including the present author). The department's attitude was positive toward this, even though, as already mentioned, clinical psychology was not officially included among the spheres of study. Since higher education was strictly centralized, that is, the whole curriculum was compiled by the central Ministry of Higher Education in Moscow, to which all Soviet institutions of higher education were subordinated, the department was not allowed to introduce teaching courses or to plan studies on its own. The only liberty afforded the department was the "special courses" that constituted a very small part of the curriculum. The first course of psychotherapy for students of psychology was introduced in the form of such a special course. It was entitled "Methods of psychocorrection," since "psychotherapy," that is, healing, was a prerogative of physicians alone, and the fact that this type of healing meant "treatment by psychological means" did not interest the term-jugglers. Generally, according to the principle of economy, it was often more worthwhile to evade obviously absurd requirements and prohibitions than to try to prove their absurdity.

10.14.4 THE BEGINNINGS OF CLINICAL PSYCHOLOGY

Those starting to work in the field of clinical psychology had to make efforts to attain the appropriate education and at the same time to achieve professional recognition. Clinical psychology was taught and developed at two centers in the USSR, Moscow and Leningrad.

Incidentally, the term "clinical psychology" was not accepted, for ideological reasons. Everything went under the title "medical psychology," while in Moscow the term "pathopsychology" was also used, defined as being the application of psychological skills in a mental clinic.

> The psychologist assists the psychiatrist in the different stages of the investigation, treatment and examination of the patient, as well as participating in the preparation of sociopsychological recommendations for the integration of the patients into society. The main form of the pathopsychologist's work is the experimental psychological examination of psychiatric patients. (p. 5)

In thus such a complicated way was the psychologist's work defined in an instruction published in Moscow in 1975, *On the work of the pathopsychologist in a mental clinic* (Zuhar, Rubinštein, Poperečnaja, & Portnov, 1975), which was approved by the head of the Chief Coordinating Board for Research Institutes and Research, and which was addressed to "the chief specialists of psychiatric institutions, in order to lead and coordinate the pathopsychologists' work, and to psychologists as methodical instructions for practical work." As can be observed from the quoted passage, these instructions did not mention psychotherapy, since the treatment of mental disorders was supposed to be the prerogative of the medical specialists, and the latter in their turn would treat only by medication and other medical procedures. Incidentally, the "experimental psychological examination" mentioned here is a very characteristic example of the imprecise usage of terms. In this situation, empirical data is usually accumulated at the clinic by means of psychodiagnostic techniques, whereas the instruction presents this process, quite inadequately, as "experimental," attaching to it the title of an investigative strategy (Rimkutė, 1986).

In Moscow the dominating approach was Pavlovian, and altogether unfavorable with regard to psychotherapy. At Moscow University, however, research was carried out concerning the cognitive aspects of mental diseases and the psychological aspects of various somatic diseases.

The attitudes at Leningrad University and the Bekhterev Institute of Psychoneurology were rather more liberal; there, both psychodiagnostics and psychotherapy were developed. Psychological tests and techniques of examination worked out in the West were also applied there.

Thus, those Lithuanian psychologists who wished to work in the field of clinical psychology used to go to the universities and institutes of Leningrad or Moscow to continue their studies or to do postgraduate work. On the whole those who started to work as clinical psychologists would at the same time have to work independently and to seek opportunities to develop their skills. On the one hand, one had to strive for professional recognition, to seek good contacts with medical specialists, to prove that clinical psychology and psychotherapy offered much to medical practice. On the other hand, the information and literature available were insufficient, for Soviet literature, usually in Russian, was as a rule one-sided, while new literature from the West was very hard to obtain. Also there was no chance of breaking through the Iron Curtain and going abroad to study. Other ways and means had to be sought.

Comparatively good psychological literature was available from neighboring Poland, where the situation was more favorable than in Lithuania. Poland was not a Soviet republic, but only a satellite to the Soviet Union, and ideological restraints were not so severe there. Numerous valuable books were translated into Polish, and therefore quite easily obtainable; one had only to learn Polish.

Another possible way of developing skills was through contacts with colleagues in East Germany. In this country, otherwise most strictly socialist, the psychodynamic trend in clinical psychology was not banned. At the end of the 1970s Dr. Kurt Höck, a psychotherapist from East Berlin, began to organize yearly seminars in psychodynamic group psychotherapy for East German psychologists and psychotherapists. One small group in this seminar was international and included German-speaking colleagues from Poland, Hungary, Czechoslovakia and other East European countries. The two participants from Lithuania were the present author and the psychotherapist A. Alekseičikas. Of course, these visits were only a private initiative, unsupported by any institutions, at one's own expense and during one's vacation. But once an invitation "to visit acquaintances" was received, it was relatively simple to get leave to attend.

Participation in this seminar offered not merely the possibility of learning group psychotherapy; it was in a sense an aid to Lithuanian clinical psychology at large. The international group comprised active and creative people from the East European countries. Even after the several years' seminar program had been completed, the group continued to meet in different countries: East Germany, Poland, Lithuania, Hungary. All in all, this communication continued for 15 years. It offered opportunities to learn and to exchange experience, and the possibility of creating joint projects. We did our best to communicate the knowledge and new skills acquired during these sessions to our Lithuanian colleagues, as well as to mediate their new contacts with East European specialists.

This was a very special experience which unified the East European psychologists and psychotherapists, all deprived and isolated to a greater or lesser degree, and exerting immense efforts in order to achieve professional skills. The dramatic political events in East Europe were also experienced together—starting with the Solidarity movement in Poland and the first visit by Pope John Paul II there, the birth of the liberation movements in these countries, and finally, the destruction of the Berlin Wall.

This seminar program directly inspired the yearly seminars for psychotherapists, clinical psychologists and psychiatrists which we began to organize in the Vilnius Psychoneurological Hospital. The first seminar took place in spring 1978. Soon the initiative was taken over by the section of Clinical Psychology, the most active one in the Lithuanian branch of the Society of Psychologists of the USSR. (According to the already mentioned principle of centralization, the Soviet Republics had no right to form professional organizations of their own, only to have so-called branches of the central organizations.) The seminars would include reports on urgent subjects, as well as practical sessions. They were positively valued by most colleagues, and soon became well known throughout the whole of the Soviet Union as the "Vilnius spring seminars." They used to attract specialists from Estonia, Latvia, Georgia, the Ukraine, Russia, and even from Siberia and the Far East. The time came when all those wishing to attend could not possibly be received, so the participants had to be selected. On the whole, in the 1980s, the practical clinical psychology and psychotherapy of Lithuania were considered to be the most developed among Soviet republics. Lithuanians were often invited to other republics to conduct various training courses.

These seminars took place yearly for more than a decade and constituted an important contribution to the professional skills and consciousness of clinical psychologists and psychotherapists. Understandably, with time, the huge common undertaking began to fail to answer the needs of evermore differentiated professional interests and growing specializations. It was necessary to look for other forms also and various more specialized professional associations began to appear.

The pioneers of clinical psychology also received considerable assistance from the clinical psychologists of the Lithuanian emigration. In 1976 Vytautas Bieliauskas, Professor of Clinical Psychology at Xavier University (Cincinnati, Ohio) came to Lithuania for the first time. Since then he has been a frequent visitor and offers much help in training specialists in clinical psychology and raising their qualifications. In the 1990s his son Linas Bieliauskas, Professor of Neuropsychology at the University of Michigan, has also established contacts with Lithuanian psychologists and offers real assistance. During his first visits, Vytautas Bieliauskas delivered cycles of lectures on various questions of clinical psychology, to local clinical psychologists, psychotherapists, psychiatrists, and other medical specialists. This was of great importance not only because of the valuable information content, but also because it helped psychologists to spread psychological knowledge in the sphere of medicine, and contributed to the professional establishment of clinical psychologists. During further visits Vytautas Bieliauskas also conducted special training courses for psychologists and psychotherapists in the theory and practice of family therapy, as well as in clinical psychodiagnostics, presenting in detail the projective personality test of house–tree–person drawings.

For his first stay in Lithuania, Vytautas Bieliauskas had received a scientist's exchange grant from the National Institutes of Health, and visited the USSR as a health service representative. Thus he had an opportunity to also visit Leningrad and Moscow. He described his impressions from this trip in the journal *American Psychologist* (1977). Writing of the situation of clinical psychology and psychotherapy in the USSR and Lithuania, he pointed out that these were young professions in those countries. People working in them were young, enthusiastic, and hopeful. However, "everything requires time and a great deal of patience in the USSR" (p. 379). He also noted the dominating centralization and ideologization of science and learning. "My audience," wrote Professor Bieliauskas of his lectures in Vilnius, "showed a great deal of familiarity with our literature and with the literature of the West in general" (p. 378). He also stated that in Lithuania a wide interest in research and new trends of healing could be encountered, and the local psychiatric service was a little more developed than in other Soviet republics.

There were also occasions for learning from other prominent specialists. Carl Rogers visited Moscow and led seminars there, as did the creator of logotherapy, Viktor Frankl. (It was then that we started investigations of the "feeling of purpose in life" in Lithuania, as well as the adaptation of the Crumbaugh and Maholic purpose in Life test.) Virginia Satir also led family therapy seminars in Vilnius.

The first Lithuanian clinical psychologists began to work at state psychiatric hospitals (there was no other type at the time). Throughout the Soviet Union, psychology, and applied psychology in particular, was regarded rather suspiciously and only mental hospitals were allowed to introduce psychologists into their staff, but not other health institutions.

In Lithuania there were already some more progressive psychiatrists who were interested in psychology and psychotherapy, and applied psychotherapeutic techniques in their work. They included N. Indrašius, A. Alekseičikas, A. Dembinskas, L. Radavičius, and A. Vinkšna. Some of them had already acquired degrees in medical psychology in Leningrad. They were both the first teachers and our first colleagues. Alekseičikas was perhaps the only psychiatrist in Lithuania specifically practicing psychotherapy. The doctor of medical psychology Goštautas had already founded, at Kaunas Cardiology Institute, a laboratory of medical psychology which carried out research and the adaptation of psychodiagnostic methods. Dembinskas and Goštautas formed psychology and psychotherapy sections for medical students in the frame of students' research societies, in Vilnius and Kaunas respectively. Students of psychology joined them later. In this way circles of people were formed who sought to extend psychiatric aid in Lithuania beyond the biological approach, and to make psychotherapy available for patients too.

The pioneer clinical psychologists worked in the field of psychodiagnostics and to an even greater extent occupied themselves with psychotherapeutic work.

One must say that clinical psychology developed rather successfully in Lithuania, insofar as it was possible under the prevailing political conditions. Professional recognition was achieved in psychiatry. The mere fact that quite soon psychologists were working in all Lithuanian mental hospitals was proof of this (given that it was up to the hospitals to decide whether they needed such specialists), and some hospitals even developed departments of psychology, the so-called "psychology laboratories." Even some psychiatrists, having taken special courses, started working in clinical psychology (Juozaitytė, Dembinskienė, etc.).

In spite of various restrictions, the psychology graduates of Vilnius University made efforts to progress in diverse spheres of clinical psychology. This always required considerable initiative, commitment, and dedication. Much depended on the first specialist in a particular field. The prestige of psychology would be enhanced by personalities able to demonstrate competence and the value of psychology, as well as to establish good contacts and collaborate with other professionals; they would prepare the way for other colleagues in that field. Some of these pioneers should be mentioned. Rasa Bieliauskaitė started the development of child clinical psychology in Lithuania; Rūta Sargautytė was the first psychologist in a psychosomatic clinic; Gražina Gudaitė, a pioneer in Lithuania of Jung's analytic psychology, headed a department of a large mental clinic; Kristina Ona Polukordienė began work in the psychotherapy of adolescents, youth, and patients with stutters; Rimantas Kočiūnas and Aleksandras Kučinskas pioneered the humanist psychology and psychotherapy which won great popularity in Lithuania; and Laima Bulotaitė was the first to investigate the psychological aspects of alcoholism and drug addiction.

Alongside practical work, most psychologists took up research; they usually prepared and presented their doctoral theses in the universities of Moscow and Leningrad. In their dissertations they analyzed the cognitive peculiarities of schizophrenic patients (Gailienė), the personal peculiarities of children with a long history of illness (Bieliauskaitė), the effectiveness of group psychotherapy with young stuttering patients (Polukordienė), the possibilities of youth group psychotherapy (Želvys), the psychological problems of alcoholic patients (Bulotaitė), and psychological problems in cardiology (Rugevičius, Palujanskienė).

Clinical psychologists and psychotherapists also took up teaching; they gave various courses in clinical and medical psychology, clinical psychodiagnostics, and psychotherapy courses for students of medicine and psychology, as well as for physicians. They also wrote several textbooks (Dembinskas et al., 1981; Gailienė, 1989; Sargautytė & Želvys, 1988).

The work of popularizing psychology and psychotherapy, and informing the public was also very important. Psychologists wrote popular articles, spoke on radio and television and gave public lectures. In this way the popularity and authority of psychology grew, and psychological help become available to ever wider circles of society.

10.14.5 REGAINING INDEPENDENCE

In the late 1980s perestroika emerged in the Soviet Union, which meant the beginning of the crash of the totalitarian system. In Lithuania all the ties of exaggerated dependence were rapidly torn. Professional organizations were the first to free themselves from enforced centralization, and psychologists were among the very first. In November 1988 the Lithuanian branch of the Society of Psychologists of the USSR held a conference which adopted the decision to close the branch. The constituent conference of January 1989 founded an independent organization, the Lithuanian Union of Psychologists, confirmed its statute, and elected its council and president. An address to the Society of Psychologists of the USSR was adopted which was subsequently read at the Eighth Conference of the Society on January 30, 1989, which was "as usual, pervaded by chaos. We were only observers there" (Bagdonas, 1991, p. 62). The newly formed Lithuanian Union of Psychologists included 410 members. Moreover, clinical psychologists and psychotherapists made haste to enjoy the emerging opportunities and freedom, and created specialized associations. Thus the Lithuanian Association for Application of Psychoanalysis was founded, followed by the Lithuanian Association of Group Psychotherapy, and the Lithuanian Association for Humanistic Psychology, also uniting mainly clinical psychologists, and so on. It was then that the first private psychotherapy consultations appeared.

The same year witnessed radical reforms at Vilnius University. It was the first among the Soviet universities to declare the so-called ideological "sciences" to be pseudosciences, and to dissolve the departments of scientific communism, scientific atheism, and history of the Communist Party. Other Soviet universities followed this example. The oldest faculty of Vilnius University, the Faculty of philosophy, was re-established, including studies in philosophy, sociology, and psychology. Instead of one department of psychology, two were created—the department of clinical and social psychology and that of general and pedagogical psychology.

On March 11, Lithuania proclaimed its Act of Independence, which meant that Lithuania was to be freed from the Soviet empire and was to seek, by persistent peaceful means, *de facto* independence. In the summer of 1991, after the abortive putsch in Moscow, the independent state of Lithuania was recognized by the world.

On June 12, 1990 the Supreme Council of Lithuania confirmed the new Statute of Vilnius University, establishing the autonomy of the university. The university re-established the stages of study that had already been adopted in 1579: the bachelorship, mastership and doctorate. This had positive effects upon the training of clinical psychologists. The department of clinical and social psychology prepared master's and doctorate programs in clinical psychology. The first doctorate theses worked on at Vilnius University and written in Lithuanian have already been presented. (During the Soviet period, particularly in the 1970s and 1980s, an anticonstitutional requirement had been passed that dissertations should be presented only in Russian.) The teachers of the department have written textbooks to be used in these studies (Gudaitė, in press; Kočiūnas, 1995).

10.14.6 PRESENT SITUATION AND PROBLEMS

On the one hand the present situation is characterized by numerous positive changes in various ways. The public attitude towards psychology has changed. It is no longer "some kind of strange and funny profession," but an undoubtedly necessary and respected one. This is proved by the number of applicants each year to the specialties of psychology, the popularity of psychological literature, and the willingness of people in other professions to cooperate with psychologists. Of course, these changes greatly facilitate the work of psychologists. On the other hand, clinical psychologists maintain that fear of stigmatization is still too strong; people still hesitate too much before seeking the help of a psychologist.

For clinical psychologists the present situation opens up countless possibilities for learning. Contacts with other countries have become easier; many new professional contacts are being made. Psychologists go to European countries and the USA to study and to gain experience. Specialists in various trends come to Lithuania with lectures and seminars at the invitation of professional organizations and universities.

Vilnius University trains the clinical psychologists of Lithuania. Psychologists who have a master's degree in clinical psychology are allowed to practice and psychologists with a doctorate may teach students.

The postgraduate training of clinical psychologists and psychotherapists is mainly the business of professional associations. In the mid-1990s there are over 10 professional psychologists' organizations in Lithuania, and almost all of them encompass clinical psychologists and psychotherapists. Most of these organizations have already become members of corresponding international organizations.

The Lithuanian Psychotherapeutical Society is a member of the International Federation for Psychotherapy and the European Association of Psychotherapy. Its aim is that our schemes of teaching psychotherapy and our certification procedures should correspond to European requirements.

The Lithuanian Association for Application of Psychoanalysis maintains close contacts with the psychoanalytic associations of other countries. The Finnish Association organized full instruction for six colleagues from Lithuania. At present, in 1997, they have finished the course, received the status of associate members and are beginning to get back to work in Lithuania. Other members of the Association are participating in long-term programs for teaching psychoanalysis which have been worked out together with the Dutch and German psychoanalytic societies.

The field of group analysis, the concern of the Lithuanian Association of Group Psychotherapy also has long-term teaching programs. This association is a member of the International Association of Group Psychotherapy. It not only seeks study opportunities for its members but also organizes training courses in which colleagues both from Lithuania and neighboring countries willingly participate.

The Lithuanian Association for Humanistic Psychology also pays great attention to training, and has gone so far as to found the Institute of Humanistic and Existential Psychology (headed by Dr. R. Kočiūnas). Training courses are also organized by the Lithuanian Association of Hypnosis and the Lithuanian Society of Jungian Analytical Psychology.

Of the professional associations, the Lithuanian Suicidological Association, which is a member of the International Association for Suicide Prevention, is probably the most interdisciplinary, like suicidology itself. The problem of suicides in Lithuania has grown dramatically during the 1990s, and the specialists in this area are the first to be alarmed. They in their turn are trying to draw the attention of both society and of governmental institutions to the problem, to inform the public, conduct investigations, and create suicide prevention programs (Gailienė, 1996).

At present psychologists and psychotherapists work in all Lithuanian mental hospitals and in the psychiatric departments of various clinics and general hospitals. Both professions are practicing psychotherapy. Even though the formal position of the hospitals has not changed—they are all state mental health hospitals—some shifts have taken place in attitudes to patients and their treatment. The inpatient departments of mental hospitals have often introduced the principles of the therapeutic community, and have opened day-patient departments. Treatment is mostly combined medication and psychotherapy. Three mental hospitals already have inpatient psychotherapeutic departments where the main method of treatment is psychotherapy, while medication is used to a minimal extent. Mental hospitals most often apply group psychotherapy, client-centered psychotherapy, psychodynamic therapy, cognitive-behavioral therapy, and Gestalt therapy.

Certainly it would be an exaggeration to speak of an unproblematic integration of psychiatry and clinical psychology. There are problems and tensions enough, but there is also progress.

The Clinical Psychotherapy Center has been opened in Vilnius, offering specifically psychotherapeutic help. The Center includes inpatient departments for adults and children, an outpatient department, a psychotherapy and rehabilitation department for torture victims, and a department for the prevention of mental disorder in adolescents and young people. The Center is subordinated to the health department of Vilnius municipality, but receives patients from the whole republic, for there are practically no such institutions to be found in other cities. The financial situation of the center is rather typical of the present Lithuanian health service which has not yet turned into a health insurance-based system. Psychotherapeutic help here is paid for, but children, adolescents, invalids, and others needing social support receive free help. The prices are established by the Minister of Health, and are low.

Thanks to the work of Dr. K. O. Polukordienė, the Youth Psychological Aid Center (YPAC) has been opened in Vilnius. This is a nonprofit, nongovernmental organization which gives free psychological help to young people. Psychologists and psychotherapists work there, as well as over one hundred specially selected and trained volunteers. The Center provides the "Youth Line," a voluntary anonymous phone help service from other young people; professional psychological and psychotherapeutic aid; various clubs and groups (e.g., the Stuttering Problems Club), a youth discussion club, an art studio, and so on. The YPAC is a member of the International Federation of Telephonic Emergency Services and maintains close cooperation with Befrienders International, the international volunteer organization of the UK. The Center also assists all the new psychological help services being founded in Lithuania and employs both professionals and volunteers.

As a result of the persistent efforts of the child psychiatrist Dr. D. Pūras, the University Center for Children with Developmental Disorders began in Vilnius in 1991. Children with especially complicated psychosocial problems were, for the first time, treated not only biologically and by means of medication, but also by receiving far wider psychosocial help which was offered to them and their families by clinical psychologists, child psychiatrists, and social workers working as a team. Students of medicine, psychology, and social work are also trained at the Center, and, postgraduate courses are offered to specialists in these fields. The Center cooperates intensively with corresponding institutions in various foreign countries. The Center has created two models of service for children with developmental disorders which are being introduced in all the municipalities. The first is early intervention for children up to four years old with risk factors and for their families. It is carried out by physicians, clinical psychologists, social workers, physiotherapists, and speech therapists working as a team. The second model is that of child psychiatric help which also first involves planning various psychosocial interventions.

Help for alcoholic patients and those with addictive diseases has also broken out of the narrow medical framework. Although the number of institutions providing assistance to these patients is far from sufficient, in those that already exist, psychotherapy and psychosocial interventions are widely applied, and specialists of various professions cooperate.

Clinical psychologists also work in the field of health psychology: they investigate the risk factors for different disorders, organize teaching, write books (Gailenė, Bulotaitė, & Sturlienė, 1996; Lepeškienė, 1996) and edit the magazine *Psychology for You* (Dr. G. Chomentauskas).

However, qualified psychological and psychotherapeutic help is only available to the population of the few largest cities so far. In smaller settlements it is still lacking. Psychology has been introduced only to a small extent into other spheres of medicine beside psychiatry. This is conditioned by different factors. The economic situation of Lithuania is still rather unsteady. In the health service system even the most elementary needs of the patients are sometimes scarcely answered, so a more sophisticated kind of help, such as psychological aid, seems to be an unaffordable luxury. Moreover, the reform of the health service is proceeding very slowly. As already mentioned, there is no insurance-based medical system. Therefore as yet there is almost no private psychotherapy practice, since in this situation each payable service means that the client pays a double fee: once with his obligatory state taxes, and also when paying for the particular service.

Other acute problems also arise from the economic situation. For example, on account of the extremely poor salaries paid to specialists doing research or academic work, some capable people, especially the young, have to abandon these areas.

Bureaucratic reforms are also taking a long time. The legal status of the specialty of psychotherapy remains problematic. The certifying commission at the ministry decides upon the qualifications of psychologists and psychotherapists, assigning qualification categories according to which the salaries of these workers at governmental institutions are established. Previously this was a prerogative of the certifying commission for neurologists and psychiatrists which was far less competent, for that matter, than the present one. But probably in the future the task of certification of specialists will be taken over by professional organizations. For example, a licensing commission has already been formed at the Lithuanian Society of Psychologists which is working upon the licensing regulations for practicing psychologists.

Briefly speaking, the present situation of Lithuanian clinical psychology is more or less hopeful. One can expect that as the general situation of the country normalizes, the position of clinical psychology will improve as well, since the need for psychological help in this society is very great indeed.

10.14.7 REFERENCES

Bagdonas, A. (1991). Lietuvos psichologijos raida [The development of Lithuanian psychology]. *Mokslas ir Lietuva* [Science, arts and Lithuania], *2*, 55–63. (In Lithuanian)

Bieliauskas, V. J. (1977). Mental health care in the USSR. *American Psychologist*, *5*, 376–379.

Blažys, J. (1921). Psichoanalizė ir psichoterapija [Psychoanalysis and psychotherapy]. *Medicina*, *2*, 35–42. (In Lithuanian)

Blažys, J. (1935). *Į vadas į psichiatriją* [Introduction to psychiatry]. Kaunas, Lithuania: Vytauto Didžiojo universiteto Medicinos fakultetas. (In Lithuanian)

Blažys, J. (1936). *Tolerancija kaip kultūros principas* [Tolerance as a cultural principle]. Kaunas, Lithuania: Spaudos fondas. (In Lithuanian)

Bumblauskas, A. (1994). Vilniaus universitetas: politinės ir istoriografinės kolizijos [Vilnius University: political and historiographical collisions]. In *Vilniaus universiteto istorija 1579–1994* [History of Vilnius University 1579–1994] (pp. 7–17). Vilnius, Lithuania: Valstybinis leidybos centras. (In Lithuanian)

Danys, M. (1986). *Lithuanian immigration to Canada after the Second World War*. Toronto, ON: Multicultural History Society of Ontario.

Dembinskas, A., Alekseičikas, A., Gailienė, D., Goštautas, A., Grizickas, A., Pūras, A., & Radavičius, A. (1981). *Psichologija medicinoje* [Psychology in medicine]. Vilnius, Lithuania: Mokslas. (In Lithuanian)

Gailienė, D. (1989). *H. Roršacho metodika asmenybei tirti* [The Rorschach method of personality diagnosis]. Vilnius, Lithuania: Vilniaus universitetas. (In Lithuanian)

Gailienė, D. (1996). *Savižudybė? Ne!* [Suicide? No!]. Vilnius, Lithuania: ia Lietuvos suicidologijos asociacija. (In Lithuanian)

Gailienė, D., Bulotaitė, L., & Sturlienė, N. (1996). *Ašmyliu kiekvieną vaiką. Apie vaikų psichologinio atsparumo ugdymą* [In love each child. On training children's psychological resistance]. Vilnius, Luthiania: Valstybinis leidybos centras. (In Lithuanian)

Gučas, A. (1937). *Pašaukimas ir darbas* [Vocation and work]. Kaunas, Lithuania: Lietuvos psichotechnikos ir profesinės orientacijos draugija. (In Lithuanian)

Gudaitė, G. (in press). *Įvadas į analitinę psichologiją* [Introduction to analytical psychology]. (In Lithuanian)

Kočiūnas, R. (1995). *Psichologinis konsultavimas* [Psychological counseling]. Vilnius, Lithuania: Lumen. (In Lithuanian)

Lepeškienė, V. (1996). *Humanistinis ugdymas mokykloje* [Humanistic training at school]. Vilnius, Lithuania: Valstybinis leidybos centras. (In Lithuanian)

Rimkutė, E. (1986). Psichologijos tyrimo metodai [Methods of psychological investigation]. In A. Gučas (Ed.), *Bendroji psichologija* [General psychology] (pp. 24–36). Vilnius, Lithuania: Mokslas. (In Lithuanian)

Šapoka, A. (Ed.) (1989). *Lietuvos istorija* [History of Lithuania]. Vilnius, Lithuania: Mokslas. (In Lithuanian)

Sargautytė, R., & Želvys, R. (1988). *Medicininės psichologijos pagrindai* [Elements of medical psychology]. Vilnius, Lithuania: Lietuvos TSR aukštojo ir specialiojo vidurinio mokslo ministerija. (In Lithuanian)

Zuhar, V. P., Rubinštein, S. J., Poperečnaja, L. N., & Portnov, A. A. (1975). *O rabote patopsichologa v psichiatričeskoj bolnice* [On the work of the pathopsychologist in a mental hospital]. Moscow: Ministerstvo zdravoohranenija RSFSR. (In Russian)

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.15
From a Monocultural Identity to Diversity Identity: A Psychological Model for Diversity Management in South African Organizations

KEDIBONE LETLAKA-RENNERT
Johannesburg, South Africa
and
WOLFGANG P. RENNERT
University of the Witwatersrand, Johannesburg, South Africa

10.15.1 INTRODUCTION 335
10.15.2 PSYCHOLOGY IN SOUTH AFRICA: A HISTORICAL PERSPECTIVE 336
10.15.3 INDIVIDUAL AND CHANGE 336
10.15.4 A SOUTH AFRICAN MODEL FOR ORGANIZATIONAL DIVERSITY 337
10.15.5 MONOCULTURAL IDENTITY 338
  10.15.5.1 Maintaining Monocultural Identity Through Restrictive Affirmative Action 339
  10.15.5.2 Moving Towards an Identity of Awareness Through Restorative Affirmative Action 339
    10.15.5.2.1 Role of the organizational psychologist: constructive affirmative action 340
10.15.6 IDENTITY OF MULTICULTURAL AWARENESS 340
  10.15.6.1 From the Identity of Awareness to Diversity Identity Through Diversity Management Training 340
10.15.7 DIVERSITY IDENTITY: OUTLOOK FOR THE FUTURE 341
10.15.8 SUMMARY AND CONCLUSION 342
10.15.9 REFERENCES 342

10.15.1 INTRODUCTION

The urban traveler in the South Africa of today is frequently confronted with a sight unknown to his counterpart from older days: white homelessness. Street corners are occupied by weary looking men and women, sunburnt from hours of standing outside, holding up signs asking for food or for any job. The emphasis is on "any." The days of job preservation for whites are clearly over, and with a new political era comes the pressure for individuals as well as institutions and organizations to adapt to what bishop Desmond Tutu has named the "rainbow nation."

Identification as a rainbow nation demonstrates South Africa's will to survive change. Adopting the rainbow nation as a construct, although political and economic in nature, is also understandable psychologically. It is an inclusive construct that provides space for everybody to join a multicultural society. It enables the forging of a new identity, as members of a new nation. Simply stated, diversity has four components in South Africa. It is a unifying mechanism, it provides a preventative strategy, it creates a basis for proactive intervention, and lastly it is a compelling vision. If the nation's entrenched divisions were not to be permitted to run rampant and deepen, some sustainable way forward had to be found. Healing had to take place simultaneously with growing as a new nation. To varying degrees both the private and the public sectors have reacted to the implications of the new dispensation.

10.15.2 PSYCHOLOGY IN SOUTH AFRICA: A HISTORICAL PERSPECTIVE

Psychologists have been involved in the creation of apartheid from the outset and are becoming instrumental in overcoming it today. In 1928 a team of psychologists, sociologists, and educators were sent to South Africa by the Carnegie Commission to study the problems of poor whiteism. Their extensive report (Wilcocks, 1932) stated with great concern that the:

> Long continued economic equality of poor whites and the great mass of non-Europeans and propinquity of their dwellings tend to bring them to social equality. This impairs the tradition which counteracts miscegenation, and the social colour divisions are noticeably weakening. (p. 33)

Wilcocks continued to recommend the creation of employment categories reserved for whites and called for "legislation which inflicts severe penalties on sexual intercourse between the races" (p. 33).

Fellow psychologist and later prime minister H.F. Verwoerd used Wilcocks' recommendations and combined them with his own ideas of racial supremacy. Verwoerd became one of the architects of apartheid and focused particularly on "Bantu education" as a means with which to deprive the indigenous African population of any access to knowledge. He believed that:

> The school must equip the Bantu to meet the demands which the economic life of South Africa will impose on him . . . there is no place for him in the European community above the level of certain forms of labour . . . What is the use of teaching a Bantu child mathematics when it cannot use it in practice? . . . That is absurd. (Harrison, 1981, p. 194)

As late as 1986 South African psychologists applauded the Carnegie Investigation for its "usefulness in the solution of social problems" (Louw, 1986). Many professionals stood on the sidelines of the struggle for liberation and claimed that their disciplines were apolitical in nature, forbidding them to examine scientifically the consequences of racial and economic oppression, while benefitting themselves materially and in the progression of their careers from that very oppression (Biesheuvel, 1958). Few psychologists dared to speak out for a politicization of psychology, and went as far as to suggest that apartheid policies and the dehumanization of the vast majority of the population might actually be detrimental to the mental well-being of the community (Dawes, 1985).

For a while, the struggle became the focus of research for mostly white academic psychologists, who saw the writing on the wall and tried to jump ship before they might lose influence and privilege. Studies on torture, violence, and colonialism filled the journals authored by psychologists, who did not dare to venture out into the townships where their subjects lived, and who kept black professionals out of their own academic units (Letlaka-Rennert, 1991).

Psychologists of color, often against the resistance of their white colleagues, struggled to create an academic discourse of diversity, introducing African writers and theorists into the psychological mainstream (Bulhan, 1990; Manganyi & du Toit, 1990; Nicholas, 1993). It became obvious that the real liberation from apartheid required the transformation of the whole society including all institutions, organizations, and structures of political, social, and economic life. Clinical psychologists understand today that the transformation of the individual has to be accompanied by the transformation of institutions and organizations. It is here that psychologists contribute to the development of a new society, a rainbow nation. Therefore, the authors choose to use the field of organizational psychology and the area of organizational transformation in their discussion of diversity development and psychology in South Africa.

10.15.3 INDIVIDUAL AND CHANGE

In South African organizational psychology today, professionals are as much concerned with the psychological well-being of individuals within organizations as with the organizations themselves. It is not surprising that clinical psychologists find themselves working in institutions and organizations and that there is a demand for clinical skills from organizational psychologists.

What is at play when looking at the individual and change, from a clinical perspective? In order to tackle resistance and assist people seeking personal transformation in the context of diversity, one has to focus on the obstacles and problems related to change. It is one thing to consider the theoretical background of societal transformation. It is another to examine the foreground of personal interactions and behaviors as it presents itself in workshops and training sessions with employees and staff members.

The clinical psychologist working with diversity is confronted with reactions of people under threat. As coping with conflict, vulnerability, and exposure are called upon, issues of safety come to the fore. There is a risk of possible repercussions at the workplace connected to making disclosure in diversity management workshops after the encouraging and supporting psychologists have long gone. We also underestimate the resilience and power of stereotypes working against disclosure. Whites might fail to disclose because of their feeling that blacks are not good enough or sufficiently like them to appreciate them, warts and all. Similarly, blacks fail to disclose because of their feeling that whites cannot be trusted to operate on a basis of equality. Thus, when under pressure and scrutiny, the defense mechanisms of one's choice flare up (Duckitt, 1992). The challenge remains to enable individuals to disclose and to promote self-confidence, such that people cannot only take the risks of change in the workshop, but have the courage to try what they have discovered outside in the real world.

10.15.4 A SOUTH AFRICAN MODEL FOR ORGANIZATIONAL DIVERSITY

Companies and institutions looking for help in facilitating a process of adaptation are still hindered by the resistance of established power structures, prejudice, and a culture of exclusion. Psychologists and human resource strategists in South Africa have no indigenous models to help them conceptualize the state of an organization's cultural identity and to meaningfully assist in the implementation of strategies to facilitate collective and organizational identity development (Table 1). Although models have been developed in the western world, none has been derived from and designed for the specific conditions of Africa in general and South Africa in particular. Typically, models and conceptualizations borrowed from the USA are used to help understand what manifests here and unfortunately, occurances are not sufficiently contextualized (Cross, Katz, Miller, & Seashore, 1994; Thomas, 1996).

This chapter explains why it is in South African organizations' self-interest to become involved in diversity work, in its broadest sense, from merely promoting multiculturalism and pluralism to direct behavioral and mental health interventions. Thus, a model is presented for diversity development that is reflective of the specific historic and economic texture of South Africa. It tries to provide a conceptualization of a possible development from the monocultural philosophy of racial exclusion and oppression, as it was prevalent in the apartheid years, to a multicultural philosophy that acknowledges and respects difference and can provide the basis for the progression of the rainbow nation (Steyn & Motshabi, 1996).

The model is psychological in nature and explains how organizations reflect the cultural state of the society in which they operate. Historically, working environments in the South Africa of the apartheid era were structured and organized to care exclusively for the needs of the white minority. The organizational identity was monocultural. In many institutions and organizations this identity still prevails and will have to be transformed and developed. Changes have to be promoted and facilitated on a level of personal identity management for members of the organization, as well as on a collective level for the organization as a whole.

The initial process will be largely one of awareness building. In a society where, for generations, a small minority has succeeded in keeping a large indigenous population under conditions of total deprivation, psychological strategies of denial, rationalization, and displacement are widely prevalent cognitive processes and learned behaviors. Interventions will have to utilize internal measures of awareness building, together with external measures of reinforcement and control in order to overcome prejudice, ambivalence, and resistance. Once an organization and its members have reached a level of awareness that provides the background for active diversity seeking, an identity of multicultural awareness will have been reached.

From here onwards diversity becomes an integral part of the organization's strategies. External control measures will become redundant and diversity will become organizational policy until a state of diversity identity has been reached. Now that diversity is no longer an objective, it has become part of the organization's fiber. The following sections

Table 1 Concepts of South African organizational identity.

Monocultural identity	*Multicultural identity*	*Diversity identity*
Restrictive affirmative action	Restorative affirmative action	Affirmed environment
External control measures	Awakening internal control measures	Internalized control measures
External regulation	Practicing internal regulation	Internalized regulation, both individual and organizational
Racial superiority	Multicultural awareness	Multicultural advancement and promotion
Entitlement	Appreciation of difference	
Power	Sharing of power	Equality
Defensiveness	Curiosity	Deriving benefits for all from diversity management
Fear	Openness	
Sense of doom	Engagement	Confidence
Prejudice	Open-mindedness	Integrationist orientation
Cultural ignorance	Culture of exploration, both internal and external	Culture of inclusion
Employee composition white	Employee composition in transition	Employee composition diverse and integrated
Organization monocultural	Introduction of different cultures, multiculturalism	

explain the stages of the diversity management model in more detail and make reference to the specific South African context, from which it is derived.

10.15.5 MONOCULTURAL IDENTITY

In the South Africa of the 1940s to the 1980s the institutional and industrial culture was monocultural. In a country of 30 million people, only 4 million of European decent mattered. The indigenous population and, to a lesser degree, Asian immigrants and people of mixed racial origin, were used as industrial cannon fodder in a process that was exclusively geared to the promotion and development of the white minority (Davenport, 1987).

This process required a maximum level of external regulation and control, and promoted on the inside a mentality of racial and intellectual superiority, entitlement, and power, but also one of defensiveness, fear, and a sense of doom. The most powerful army on the African continent focused its patrols on the townships rather than the borders. The enemy was within the society, not outside it. White South Africans who had access to foreign passports were in a permanent state of preparedness to leave, while reaping the benefits of their privileged position as long as possible. White gun ownership far exceeded the statistics published in the USA, commonly known as the "armed society," and the illegal transfer of assets and money offshore totalled billions of rands. On the outside, physical terror and abuse were combined with a system of dehumanization, together with the slow implementation of a slave mentality (Fanon, 1967).

Organizations and institutions were a reflection of the minority rather than integrated into the society as a whole. Group thinking dominated industrial and political strategies, and promotional efforts were limited to members of the white group. At one stage, the mining industry management, down to the levels of foremen on the shaft floors, became so aloof from the black labor force that a new language—Funigalo—was invented, in order to maintain a minimum of communication between foremen and laborers.

Today, one still finds organizational thinking stuck in the same paralyzing framework of prejudice and cultural ignorance. A culture of exclusion has been sustained for decades through schools, universities, families, and the information industry. The basic fibre of organizational thinking has had to be loosened and rewoven into a culture of inclusion. It was interesting to note that, after the first democratic elections in 1994, conservative white circles immediately called for a white homeland, a move destined to maintain on a small scale, what had to be given up on the whole, namely apartheid, or the maintenance of a culture of exclusion. Somewhat more liberal elements, around the former president F. W. deKlerk, called for the inclusion of minority rights into the new constitution, and a white popular movement rallied around the few Afrikaans language schools trying to preserve a pocket of white cultural exclusivity. It was the enforced introduction of the same Afrikaans into the syllabus of black schools in 1976 that had led to township riots, scores of lost lives, and a whole generation of youngsters who refused to acquire even a minimal education. In the South Africa of today a culture of inclusivity rather than exclusivity has to be learned from scratch.

10.15.5.1 Maintaining Monocultural Identity Through Restrictive Affirmative Action

The instrument with which to maintain an institutional culture of exclusivity was what we would call "restrictive affirmative action." The development of a minority was furthered by the restriction of the majority. Pass laws were imposed on black laborers, controlling their influx into the industrial centres. Wages for blacks were kept at a level prohibitive to the support of a family, in order to keep women in rural farm labor. Black children were subjected to Bantu education, designed to restrict knowledge to the following of simple orders. Where, despite institutional disadvancement, black labor met with white labor in a potentially competitive situation, labor laws were designed to preserve certain jobs and job categories for whites. Mine workers handling explosives had to be white, not in the least in order to prevent a potentially dangerous skill transfer. The postal services and other government agencies became a haven for whites who could not otherwise secure employment. Apartheid laws enforcing segregated residential areas (Group Areas Act) and the prohibition of interracial relationships (Immorality Act, Mixed Marriages Act) cemented cultural exclusivity (Davenport, 1987; Mandela, 1994).

It is interesting to observe today, that those parts of society who have benefited from affirmative action then, are the most outspoken critics of affirmative action now. They fear more than anything the competition of well-educated blacks who are competing for jobs that whites have been occupying simply because they were white. Inevitably there will be a shift away from employment on racial grounds only, towards an employment strategy that aims at racial equality on the grounds of equal qualification and competency. More whites holding up signs at street corners will bear witness to the fact that being white alone no longer guarantees advancement.

10.15.5.2 Moving Towards an Identity of Awareness Through Restorative Affirmative Action

The predominant sentiment in the established white institutional and business community today is one of fear and insecurity. Although the new national leadership has made it very clear that it intends to include all communities in the process of national unification, there is an atmosphere of mistrust, partially because many whites fear the maintenance of cultural domination—black culture this time—and partially because there may be deeply rooted feelings, that the acquired privileges were achieved through oppression and violence and are therefore undeserved. The predominant discussion in the public domain today is about crime, about the redistribution of white wealth through black criminals. The call for punishment ranges from castration to the reintroduction of the death penalty, scrapped from the legal instrumentarium by the current government of national unity (Duckitt, 1992).

At this stage, the main need is one of awareness building, and the overcoming of resistance. Many institutions and organizations are hesitant, fearful, or simply ignorant about how to honestly and more fully open themselves to the needs and challenges of a multicultural society. Many companies lack diversity in their employee composition and have still to discover large parts of the population as customers and clients, struggling with the need to modify products and designs to accommodate different tastes (Gordon, 1991; Lau & Shani, 1992).

Here, affirmative action of a different kind, namely restorative affirmative action, is required to facilitate the achievement of an identity of awareness. Resistance and ambivalence can only be overcome with the help of external control measures, this time however, not with the intention of enforcing a system of injustice, but rather with the aim of setting in motion a dynamic of learning that will render the tool redundant in the end. Government legislation as well as organizational and institutional policies need to be set in place to provide the basis for the opening of organizational culture. Enhanced education of disadvantaged students through scholarships and improved educational strategies in the townships should be combined with provisions for improved on-the-job training for employees of color. Labor laws are being revised to abolish job preservation for whites and to guarantee equal opportunity employment and remuneration. Government tenders are becoming linked to affirmative action goals within the companies competing for the job (Jones, 1991; Witt, 1990).

The Truth and Reconciliation Commission is another unique South African initiative attempting to combine awareness of past injustices with reconciliation and a positive outlook on future cooperation between diverse racial and political groups. Despite being criticized by whites for being vindictive and by blacks for letting perpetrators of racial oppression off the hook, the commission has already achieved, long before its conclusion, the initiation of dialogue between the races and the creation of an atmosphere of reflection on the limitations of monoculturalism.

10.15.5.2.1 Role of the organizational psychologist: constructive affirmative action

On an intraorganizational level, companies are seeking help from organizational psychologists to further an atmosphere of awareness of difference and its advantages for obtaining business objectives. Companies are slowly understanding that diverse perspectives will provide greater flexibility, increased creativity and improved problem-solving strategies. What they lack is a system to facilitate communication and to understand different perspectives. Here South African human resource specialists such as psychologists can provide indigenous systems that help to bridge the gap between different cultural backgrounds. The aim is to acquire awareness of difference as something nonthreatening and potentially beneficial to the progression of the organization. There is a shift from a more external restorative affirmative action to a more internally motivate constructive affirmative action (Trompenaars, 1993).

A program has been designed for institutions and organizations to facilitate this process. A multitude of intervention programs, workshop designs, and strategies are marketed by various organizations and companies, none of which take into consideration the specific conditions of the South African context. The legacy of the past and the demands of the future pose unique challenges and obstacles that require an approach that is indigenous and sensitive to the multicultural background of the rainbow nation.

The program has two modules, the first of which is designed to facilitate the progression of an organizational identity from one of monoculturalism to one of cultural awareness. Exercises are introduced to value difference, to understand affirmative action, and to develop concepts for the management of diversity. Personal attitudes, assumptions, prejudices, and behaviors have to be analyzed and modified. The predominant organizational culture is investigated and put into a societal context. On a practical basis, intercultural communication skills are provided, dealing with stereotypes and myths, and methods useful in the overcoming of resistance are practiced. On a further level, personal diversity skills are integrated into an organizational diversity management framework. Teamwork, interpersonal understanding, relationship building, empowerment, and organizational commitment are introduced as competencies that comprise the pillars of diversity management.

A member of the South African Zulu people greets another with the word *sawubona*, (I see you). The reply is *sikhona* (I am here). Essentially, this greeting means "I exist because you see me." To see a person as what they are, rather than as what one's preconceptions of them are, is the aim of intervention programs at this stage. In addition, individuals finding out about and treating colleagues in ways that are meaningful to them, and not only in the way prescribed by the old corporate culture, is another fundamental objective of the program (Letlaka-Rennert, 1996).

10.15.6 IDENTITY OF MULTICULTURAL AWARENESS

The intermediate phase of organizational identity finds companies and institutions in a state of raised awareness and residual ambivalence, but with a positive attitude towards change. Staff members that have joined the company through affirmative action programs become appreciated as valuable additions to the team. There is less fear of retribution and reverse racism, but ambivalence and prejudice may still inhibit processes of further diversification. However, at this stage, institutional leaders and industrial managers will be able to set goals and create visions of complete diversity identity. They will have understood that, for an institution to thrive in a society, it has to become a reflection of that society, in its goals as well as in its composition (Trompenaars, 1993).

10.15.6.1 From the Identity of Awareness to Diversity Identity Through Diversity Management Training

Human resource strategies at this level will rely much less on externally controlled affirmative action programs. Now organizational psychologists will provide organizations with the tools to internally develop their identity towards diversity. Diversity management may affect marketing strategies, the expansion or alteration of product portfolios, the composition of management teams, modifications of managerial and production styles, and communication styles within the company and between the company and the society in general.

The second module of the diversity management program is designed to facilitate this transition. Managers learn how to create environments that allow diverse individuals within the organization to reach their full potential in pursuit of corporate objectives. The impact of multiculturalism on market forces, competition, customer composition, recruiting, decision making, competency, performance management, and conflict resolution within the company are discussed. Crisis management in relation to issues around diversity is exercised, self-awareness is practiced and different management styles are introduced and analyzed. The projection of positive self-fulfilling prophecies is demonstrated. Recruiting, hiring, and promoting staff is discussed in the diversity management context. Performance management, compensation and rewards, communication, training, and education are all discussed as part of integrated and successful diversity management. Finally, the process of diversity management itself is analyzed and reviewed from planning to intervention, monitoring, analysis, and evaluation (Boon, 1996).

Management styles may change from a hierarchical form emphasizing individual power within a set structure, to a communal form, emphasizing consent in decision-making processes. African management styles premised on the philosophy of *ubuntu* ("personhood" in the Xhosa language) are examined as an addition to the available repertoire of management styles. The Southern African principle of *ubuntu* sets the individual in their communal context. *Ubuntu ngumuntu ngabantu* means "a person is a person through other human beings." Consensus building, cooperation, and mutual supportiveness form the basis for personal management that brings all members of the organization together for the benefit of the whole organization. Once the individuals that shape the organization have been included into the process of moving the organization forward through mutual trust and cooperation, they will be able to share the vision of the organization and contribute fruitfully (Boon, 1996).

10.15.7 DIVERSITY IDENTITY: OUTLOOK FOR THE FUTURE

At a level of diversity identity the achievement of diversity is no longer a managerial goal. Diversity has become integral to the fiber of the organization. Difference is appreciated as enriching, and as furthering creativity. Valuing difference has to do with flexibility and conflict resolution, and capitalizing on it leads to market responsiveness and general productivity. Employees are appreciated for their job performance and judged in relation to achievement rather than background. The difference now is the openness of corporate culture to different approaches used to achieve the greater company good. An organization with a strong diversity identity does not try to become a melting pot in which all differences are amalgamated into a new singular identity. It becomes, rather, a salad bowl in which differences are maintained and appreciated but freed from discrimination. Interaction is possible with all aspects of society as they are represented within all structures of the organization (Letlaka-Rennert, 1996).

Diversity identity is the end result of a progression from prejudice in monocultural identity, through openmindedness in multicultural identity to an integrationist orientation. The culture has moved from one of exclusion and ignorance to one of exploration and ultimately one of inclusion and integration. Characteristics of the monocultural identity consist of racial superiority, elitism, entitlement, dominance, defensiveness, fear, aggression, and a siege mentality. Characteristics of the multicultural identity are openness, curiosity, discovery, engagement, and risk-taking. Characteristics of the diversity identity are appreciation of others, self-respect, differentiation, innovation, maturity, clarifying interactions, and ongoing dialogue.

Typically, in the South African monocultural context, the corporate or institutional employee composition was predominantly white. In the multicultural context of today, the employee composition is being addressed or redressed. Race and gender are at the forefront of the diversity drive, more so than sexual orientation or age. Skill shortages and the need for training and staff development are hindrances that have to be addressed. Organized labor is demanding to be included in the process. So far, diversity has not been achieved but it is becoming an organizational goal. In the South Africa of tomorrow, organizations will have a multicultural integrated work force representing diversity in race, gender, sexual orientation, disability, socioeconomic status, lifestyle, and family composition.

Both individuals and organizations will have evolved from the awakening internal control measures and their practice to a naturally internalized regulation. Multiracial advancement and promotion will be optimized for gaining a competitive edge. Organizational diversity will translate into organizational growth. President Nelson Mandela stated that

South Africa's greatest resource is its people. By empowering them to know how to work together in a conducive environment with an accommodating and permissive culture is the best reason to give diversity a chance.

10.15.8 SUMMARY AND CONCLUSION

The history of South Africa is one of racism, exploitation, and oppression. On every level of personal, organizational, and societal life individuals and institutions are faced with overcoming psychological and institutional barriers against change. Clinical and organizational psychologists are asked to help in facilitating a process of societal transformation that allows individuals as well as organizations and institutions to change their culture. Coming from a background of cultural exclusion, moving towards an awareness and appreciation of difference requires disclosure, vulnerability, and the tolerance of conflict. Prejudices and stereotypes have to be overcome and ingrained power structures have to be questioned and loosened.

In the South Africa of today we have a long way to go before our institutions—any of our institutions and organizations—can claim to have reached the diversity identity state. We are dealing with setbacks at every step of the way, we are hindered by old prejudices as much as by a sense of entitlement within the new elites, some of whom forget their constituencies once they have reached positions of influence and power. Corruption is a problem, as it is in every society today. But who would have predicted in 1989 or 1992 that a process of societal transformation could be achieved peacefully and with the full commitment of the vast majority of the people of this country. The road is still long, but South Africans are on the way, and we are determined to arrive together—a rainbow nation.

10.15.9 REFERENCES

Biesheuvel, S. (1958), Objectives and methods of South African psychological research. *Journal of Social Psychology, 47,* 161–168.

Boon, M. (1996). *The African way: The power of interactive leadership.* Johannesburg, South Africa: Zebra Press.

Bulhan, H. A. (1990) Afro-centric psychology: Perspective and practice. In L. J. Nicholas & S. Cooper (Eds.), *Psychology and apartheid* (pp. 66–78). Johannesburg, South Africa: Modiba Publications..

Cross, E. Y., Katz, J. H., Miller, F. A., & Seashore. E. W. (1994). *The promise of diversity.* New York: Irvin.

Davenport, T. R. H. (1987). *South Africa: A modern history.* Johannesburg, South Africa: Macmillan.

Dawes, A. (1985). Politics and mental health: The position of clinical psychology in South Africa. *South African Journal of Psychology, 15,* 55–61.

Duckitt, J. (1992). *The social psychology of prejudice.* Westport, CT: Praeger.

Fanon, F. (1967) *The wretched of the earth.* Harmondsworth, UK: Penguin.

Gordon, J. R. (1991). *Organizational behavior.* Boston: Allyn and Bacon.

Harrison, D. (1981). *The white tribe of Africa.* Los Angeles: University of California Press.

Jones, A. J. (1991). *Affirmative talk, affirmative action: A comparative study of the politics of affirmative action.* Westport, CT: Praeger.

Lau, J. B., & Shani, A. B. (1992). *Behavior in organizations: An experiential approach.* Chicago: Irvin.

Letlaka-Rennert, K. T. (1991). Ambivalence. *Agenda: A Journal about Women and Gender, 9,* 9–10.

Letlaka-Rennert, K. T. (1996). *The relationships among racial and gender identity models and locus of control and self-efficacy with black South African university students.* Unpublished doctoral dissertation, George Washington University, Washington, DC.

Louw, J. (1986). White poverty and psychology in South Africa: The poor white investigation of the Carnegie Commission. *Psychology in Society, 6,* 47–62.

Mandela, N. (1994). *Long walk to freedom: The autobiography of Nelson Mandela.* Randburg, South Africa: Macdonald Purnell.

Manganyi, C., & du Toit, A. (1990). *Political violence and the struggle in South Africa.* Halfway House, South Africa: Southern Book Publishers.

Nicholas, L. J. (1993). *Psychology and oppression: Critiques and proposals.* Johannesburg, South Africa: Skotavi.

Steyn, M. E., & Motshabi, K. (1996). *Cultural synergy in South Africa: Weaving strands of Africa and Europe.* Randburg, South Africa: Knowledge Resources.

Thomas, R. R. Jr. (1996). *Redefining diversity.* New York: AMACOM.

Trompenaars, F. (1993). *Riding the waves of culture: Understanding cultural diversity in business.* London: Ther Economist Books.

Wilcocks, R. W. (1932). *The poor white problem in South Africa: The report of the Carnegie Commission.* Stellenbosch, South Africa: Pro Ecclesia Drukkery.

Witt, S. L. (1990). *The pursuit of race and gender equity in American academe.* Westport, CT: Praeger.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.16
Clinical Psychology in Asia: A Taiwanese Perspective

SUE-HUEI CHEN
National Taiwan University, Taipei, Taiwan
and
EUGENE K. EMORY
Emory University, Atlanta, GA, USA

10.16.1 INTRODUCTION 343
10.16.2 MANIFESTATION OF PSYCHOPATHOLOGY 344
10.16.2.1 Social Orientation and Collectivism as a Potential Cause 344
10.16.2.2 Social Orientation and Collectivism as a Buffer 345
10.16.3 DEVELOPMENT OF THE PROFESSION OF CLINICAL PSYCHOLOGY IN TAIWAN 346
10.16.4 CHOICE OF TREATMENT FOR PSYCHOLOGICAL DISORDERS 346
10.16.4.1 Treatment Choices 346
10.16.4.2 Therapy Models 347
10.16.5 CONCLUSIONS 347
10.16.6 REFERENCES 348

10.16.1 INTRODUCTION

Contrary to some Western conceptions, Asian people, such as those of Chinese descent, comprise a very heterogeneous group. The impact of Western culture on Chinese societies, in particular, Taiwanese people, may have positive and negative effects. One important feature of Westernization in Taiwan is the role it plays in the manifestation of psychological disorders. Cultural factors are imminently important in the identification and interpretation of clinical symptoms, and in the efficacy of certain forms of therapy. The emergence of doctoral level training of Taiwanese, from abroad and in the country itself, will have a profound influence on the future of clinical psychology in this Chinese island nation.

Asian people from different parts of the world are considered by most Westerners as a homogenous group. This tendency to homogenize distinctly heterogenous subcultures can lead to flawed concepts about these people. For example, many Westerners assume that people who live on the China mainland, in Hong Kong, and in Taiwan represent one ethnic group. While there are some obvious cultural and anthropologic similarities, including religious practices and physical features, these groups have markedly different social customs. They include dietary, religious, and even language differences. People in Hong Kong, for example, speak Cantonese, whereas Taiwan uses Mandarin Chinese. These are quite different dialects to what is spoken in various parts of mainland China. Moreover, vastly different ideologies and governmental policies exist. Thus, while similarities within Asian populations may appear obvious to outsiders, they are mostly superficial. Historical, geographic, and political segregation have made Chinese in Taiwan different from Chinese in other parts of the world, such as mainland China.

The transformation of the political system in Taiwan and the remarkably rapid economic changes and influences from the Western culture have altered the beliefs, value systems, and lifestyle of Taiwanese people. Therefore, the main sociocultural theme in Taiwan is a mixture of traditional Chinese culture and contemporary Western ideology. To what extent can one apply a traditional "Chinese" or Western psychological perspective to the people of Taiwan?

In a rapidly changing society, people may face heightened social and political stress. Economic competition and the standard of living have increased, but political conflicts between Taiwan and mainland China are persistent and apparent. Taiwanese people are experiencing higher levels of uncertainty about lifestyle and future well-being than they were in the past. In contrast, the traditional emphasis on family structure and collective social orientation is weakening (Yang, 1985, 1993, 1995). In societies where traditional norms and values change rapidly, new and effective coping strategies are needed. These sociopolitical conditions are associated with increasing psychopathology (Wallace, 1956). The lifestyle in Taiwan exhibits pervasive Western influences that are at variance with traditional Chinese values. The social impact appears to be a greater level of emotional distress. As a result, the manifestation of stress-related psychopathology represents itself in a changing pattern unique to this island. Could the pattern be depicted as a "Westernized" psychopathology?

In addressing the research and professional issues related to clinical psychology in Taiwan, it is appropriate to examine similarities and differences in the manifestation of psychopathology, development of clinical psychology, and choices of treatment for psychological disorders, in comparison to those in mainstream European and American society. There is a paucity of empirical and theoretical literature focusing on clinical psychology from a Taiwanese perspective, due in no small part to the limited number of Taiwanese clinical psychologists.

The goal of this chapter is to clarify a Taiwanese perspective on clinical psychology issues. It is hoped that the perspective from a rapidly changing society—Taiwan—may serve to provoke further discussions and deeper understanding of how such changes can alter social structures in an established society where social collectivism and group identity attempts to co-exist with increasing individualistic ideals. The impact of this change can be measured in psychological terms. There is still not enough empirical evidence to clarify these relationships. One outcome unfortunately may be a rise in the diagnosis of major psychological disorders. Our understanding of these outcomes may shed light on the universe of human psychopathology, not just the depiction of psychopathology and the emergence of clinical psychology in Taiwan.

10.16.2 MANIFESTATION OF PSYCHOPATHOLOGY

Traditionally familial rather than individual goals have been the measures of social success in Taiwan. According to K. S. Yang, a leading theorist of indigenous social and personality psychology, people in Taiwan have been influenced greatly by familial collectivism (Yang, 1972, 1995). On that account, social-oriented cultural features in Taiwan may very likely be seen as both a potential cause, as well as a buffer, of psychopathology. In fact, these concepts are not much different from the social dynamics that exist in the West.

10.16.2.1 Social Orientation and Collectivism as a Potential Cause

Traditionally Taiwanese people have a stronger social–collective rather than individual orientation. That is, the individual was inclined to subordinate her or his personal goals and choices for the sake of the family. More specifically, while dealing with family affairs, one tended to be primarily influenced by family's decision, regardless of one's individuality (Yang, 1972, 1995).

Such social oriented or collectivist attributes may be manifested in the symptoms of psychopathology. For example, concerns or problems about family conflicts outnumbered other psychological problems concerning individual issues in a Taiwanese sample who sought local spiritual help or professional counseling (Yee, 1985). Another interesting example is found in the delusional content of schizophrenics. Spirit possession phenomena (SPP) was noted in about 25% of first-onset schizophrenic patients. Of the variety of symptoms, there was a disproportionate number related to being possessed by ancestors or alive family members, rather than spiritual agents or demons (Wen, Lin, Chen, Chou, & Huang, 1992). Thus, as is the case in other societies, the content of delusional thought may be culturally balanced.

10.16.2.2 Social Orientation and Collectivism as a Buffer

Similar to the influence of social orientation on the manifestations of psychopathology, a family orientation may buffer individuals against psychological distress and thus decrease the frequency and type of psychological disorder. The different prevalence rates of major depression in Taiwanese and Americans patients may also reflect different distributions of psychopathology under the influences of collectivism or individualism, respectively. Several studies of Chinese samples carried out in Taiwan (Cheng, 1989; Compton et al., 1991; Lin, 1953; Lin, Rin, Yeh, Hsu, & Chu, 1969) and in the US (Weissman et al., 1992; Yamamoto, Yeh, Loya, Slawson, & Hurwicz, 1985) have found that the prevalence rate of major depression is lower among Chinese patients in Taiwan than American patients, ranging from 0.07 to 1.1% and from 5.2 to 17.1%, respectively. Furthermore, as reported in several studies, Chinese depressives manifest more somatic symptoms (e.g., Lin, 1982; Kleinman, 1982). This somatization tendency may thus contribute to the possibility of under-reporting or under diagnosing clinical depression in Chinese patients (Kleinman, 1982). What tends to be overlooked by these studies, however, is that somatization may play a role in the illness process. Maintaining the mixed anxiety-depressive state may in turn prevent the less depressed patients from becoming the hopeless depressive (Chen, 1995).

In Taiwan, the whole family tends to take responsibility for a psychological problem in order to protect the individual from being singled out. This buffering and shared responsibility may affect the revelation of a psychological disorder in the Western context. Taiwanese tend to express personal or social distress via somatization, which has been defined by Kleinman and Kleinman (1985) as the expression of personal and social distress in an idiom of physical complaints and medical help-seeking behaviors. Somatization may also be considered a reflection of basic cultural traits that discern Chinese society with a predominantly oral-hypochondriacal quality (Lin, 1982). In addition to this psychosomatic conception of "disorder," the social system shaped by collectivism has played a role in the expression of emotional distress.

Note that within the traditionally social-oriented collectivist Chinese society, the individual is not the unit of social structure. The individual's private emotions may be largely considered as a reflection of family matters. Public expression and discussing one's emotions, especially the negative emotions, may cause more social consequences and conflicts in the social milieu, thereby bringing more shame and guilt feelings to oneself and one's family or relatives. In reality these emotions may be more likely to result from an interpersonal conflict, but expression of depression or negative affect is culturally discouraged (Lin, 1982; Tseng & Hsu, 1969). Thus a more stoic demeanor is generally adopted.

Within the social-oriented system, somatization tends to have positive social consequences to the extent that both care-givers and receivers obtain affirmative evaluations in the helping process. Since psychological complaints did not have the same social efficacy that somatization had in generating support and care (Kleinman, 1982), it is thus more likely to find the somatic expression of depression in Chinese society. In the past, somatization had been commonly seen as an expression of psychological distress in Western society until psychological experience and its rationalization into discrete labels were regarded as one index of modernization. Psychologization, rather than somatization, consequently may have become the way of perceiving and expressing one's own distress in modern Western society (Kleinman & Kleinman, 1985). Could the declining trend of social orientation in Westernized human societies account for the fact that major depression rates are increasing both in Taiwan and in several Western countries?

Another factor that affects the manifestation of psychopathology may be the availability and acceptability of emotional expression. Taking depression as an index, Taiwan is in a progressive Westernization mode, and its rate of major depression is increasingly approaching statistics in the West. However, there has been a dearth of adequate "words" to express "depression" in Taiwan. In Chinese culture people were not used to expressing depression literally or verbally in colloquial speech. Instead, it was much easier to somatize (Tseng & Hsu, 1969), or to refer to those specific psycholinguistic expressions which have been shaped by the more somatic-toned culture (Cheng, 1989). Hence, it is not surprising that an enormous number of Chinese cases of depression have sought physical, rather than psychological and psychiatric, treatment for their illness (Sue & Sue, 1990). Younger people, however, seem to be conflicted between traditional values and contemporary Westernized culture. Again along with the declining trend of social orientation and collectivism among younger generations, one might associate this trend with the earlier onset of major depression in younger cohorts, both in Taiwan and in several Western countries.

10.16.3 DEVELOPMENT OF THE PROFESSION OF CLINICAL PSYCHOLOGY IN TAIWAN

As of 1996, the number of clinical psychologists remained very low in Taiwan. According to the author's own informal surveys, the estimated number is about 200. There are about 10 clinical psychologists who have a Ph.D. in clinical psychology. Among those doctorates, there are three US-trained clinical psychologists (including the first author) and one US-trained clinical neuropsychologist. The remainder received their degrees from the same local clinical psychology program, National Taiwan University. All of them currently hold a faculty position in clinical psychology programs at various universities and are carrying out part time clinical services at their affiliated medical centers. In addition, about a quarter have obtained a master's degree. The remainder hold a bachelor degree in psychology. There is no formal licensure for clinical psychologists in Taiwan.

In terms of subspecialties within Taiwanese clinical psychology, there are various programs established over the years, as is the case in many Western universities. At National Taiwan University, for example, there are traditional clinical psychology and child clinical psychology, health psychology, and clinical neuropsychology, which provide training associated with the application in psychiatric settings, family medicine settings, and neurology departments, respectively. Consistent with the Western trend, health psychology and neuropsychology have received more attention in recent decades. We note that there is only one health psychology program to date, but it has quickly developed and encountered less resistance in Taiwan, thanks, we suspect, greatly to the somatic-toned culture.

Although the first clinical psychology program started about 25 years ago, the number of Taiwanese clinical psychologists grew very slowly. This small number is due solely to the fact that clinical psychology is a recently developing profession in this transitional society, where psychological services were discouraged by traditional values. Although subtle resistance for the development of clinical psychology existed, the number of students now being trained is growing. Equally important, the percentage of master or doctorate level recipients has shown a rapid increase in recent years. This increase may mean that more, albeit still less than enough, research in Taiwanese clinical psychology may be generated by local or international psychologists in the foreseeable future, and more qualified and/or experienced clinical psychologists may be available to serve in both academic and clinical settings.

With regard to psychological assessment, there obviously exists a dearth of adequate tests during this beginning stage in the development of clinical psychology in Taiwan. Two attempts can be employed to "soothe the pang": to develop culturally sensitive tests or to conduct cross-cultural translation of those well-established Western tests. It seems more feasible to adopt and translate well-established clinical tests during this burgeoning period. In fact, some research concerning cross-cultural translation of Western tests such as Beck Depression Inventory (BDI) and Children's Depression Inventory (CDI) has been conducted. Several well-validated personality inventories have been used in Taiwan, such as Ko's Mental Health Questionnaire (Ko, 1977, 1995) and the Health–Personality–Habit Scale (Ko, 1996).

Efforts need to proceed with caution when conducting cross-cultural translation of Western tests. It was reported that Asian-Americans at UCLA tended to endorse more items on the MMPI-2 than their American counterparts (Dr. Hsieh, personal communication 1996). The authors shared similar observations while using the English form of the MMPI-2 with Taiwanese students at Emory University. Those students may not have presented a clear picture of distress at the interview, whereas they endorsed more items on some negative distress scales. A question thus arises as to whether the validity of assessment will be biased by test formats, such as self-report scales or interpersonal interview, across various groups from different cultural backgrounds. Still, such questions need to be answered.

10.16.4 CHOICE OF TREATMENT FOR PSYCHOLOGICAL DISORDERS

The behavioral patterns of people with a collectivism or social orientation may vary greatly from those who possess more individual-oriented characteristics. The same holds true in the manifestation of symptoms and their attribution of causes for psychopathology, and in turn, affects their choices related to treatment. Another factor affecting their choice of treatment models may be the limited number of clinical psychologists in the country.

10.16.4.1 Treatment Choices

We hypothesize that the choice of treatment for psychological disorders varies across social and economic status in Taiwan. The severity and types of the disorders may also affect the choice (Tseng, 1972; Tseng & McDermott, 1975; Wen, Lin, & Chen, 1993; Yee, 1985). Specifically, people with higher socioeconomic status (SES) tend to seek professional help from Western models of treatment including psychiatric medicine and psychotherapy for their mental illnesses and counseling for maladjustment problems. People with median SES have pluralistic help-seeking patterns. That is, they tend to utilize Western models for more severe mental disorders but Chinese medicine or folk therapy for psychosomatic-like disorders, as well as problems related to daily life. However, people with lower SES seldom access the services of Western treatment models. Rather, they tend to visit spirit mediums, fortune-tellers, shaman, and so forth (Yee, 1985). To date, this may suggest both a lack of knowledge of possible causes as well as a scarcity of available professional psychology services. One can see parallels in the utilization of psychological services as it relates to SES in both Taiwanese and American culture.

10.16.4.2 Therapy Models

Given that the availability of psychological services is limited in Taiwan, and the number of clinical psychologists is small, it comes as no surprise that theoretical orientation of clinical psychologists in the country appears relatively homogenous. According to the authors' anecdotal surveys, the cognitive-behavioral approach is the most popular model employed by clinical psychologists in Taiwan in both research and clinical settings. The phenomenological-humanistic approaches are second, followed by psychodynamic or psychoanalytic approaches.

Several factors presumably contribute to this pattern of theoretical orientation. The primary ingredient has been the lack of programs, as stated earlier, that recruit and train enough clinical psychologists to provide for varied theoretical orientations. Moreover, most mainstream clinical psychologists in Taiwan rarely stray from their mentors' orientation. Second, the authoritarian orientation of Chinese culture contributes to the preference for a particular orientation. Chinese tend to respect structure and formality in interpersonal relationships. The same holds true in therapeutic relationships. Yang (1995) pointed out, "as long as the Chinese see authorities as trustworthy and 'almighty,' they will be completely dependent upon those authorities." Respect for authority may take the form of agreeing readily to what the therapist does and proposes. The cognitive-behavioral approach offers relatively more formality and authoritarianism. Perhaps the very acceptability of formality and authoritarian aspects of psychotherapy in Taiwanese is likely to be consistent with the cognitive-behavioral approach. Unfortunately, there is a lack of empirical evidence making comparisons among various approaches applied to Taiwanese patients and clients.

To be sure, there are some commonalities among Asian populations with respect to psychological services. Atkinson, Maruyama, and Matsui (1978) revealed that, while treating Asian-Americans, the therapist may have to be more directive and active than they might otherwise be, given the preference of many Asian-Americans for a structured approach over an unstructured one. As proposed earlier, the younger cohorts have been dropping their social orientation and/or authoritarian attitude and becoming Westernized. Accordingly, what may be informative is whether or not the younger cohorts show an increase in their preference for less-structured therapies. Much research remains to be done with respect to how clinical psychology can be most appropriately applied in non-Western, especially Asian, culture.

10.16.5 CONCLUSIONS

As stated previously, the goal of this chapter is to provoke thought and further discussion on the topic of clinical psychology in Taiwan, and more generally, Asian cultures. Will the development of clinical psychology in Taiwan have similar trends to the US or other non-Western nations, or will its development be unique? Clinical psychology in Taiwan has developed from a circumstance in which scientific psychology has already developed itself as a mature science, whereas clinical psychology in Western countries, such as the US, seems to parallel the entire history of psychology. In some ways the development of clinical psychology in Taiwan seems to be a microcosm of the 100 years of Western clinical psychology. However, there are vast differences between the East and the West. It will be interesting and exciting to observe whether the development of clinical psychology in Taiwan, and Asia more generally, will continue the trend towards the West, or branch off into an entirely new direction.

The following factors will affect the future development of clinical psychology in Taiwan:

(i) diversity of training programs throughout the country;

(ii) the number of clinical psychologists trained locally and abroad;

(iii) the amount of empirical research conducted;

(iv) sociopolitical changes.

These factors, and perhaps others not mentioned here, will influence significantly the course of clinical psychology in Taiwan. Productive cross-cultural collaboration as well as indigenous research will undoubtedly contribute to the success of Taiwanese psychology.

10.16.6 REFERENCES

Atkinson, D. R., Maruyama, M., & Matsui, S. (1978). Effects of counselor race and counseling approach on Asian Americans' perceptions of counselor credibility and utility. *Journal of Counseling Psychology, 25*(1), 76–83.

Chen, S. H. (1995). *The role of certainty of helplessness in comorbidity of anxiety and depression: A prospective study of Chinese patients*. Unpublished doctoral dissertation, Emory University, Atlanta, GA.

Cheng, T. A. (1989). Symptomatology of minor psychiatric morbidity: A cross-cultural comparison. *Psychological Medicine, 19,* 697–708.

Compton W. M., III, Helzer, J. E., Hwu, H. G., Yeh, E. K., McEvoy, L., Tipp, J. E., & Spitznagel, E. L. (1991). New methods in cross-cultural psychiatry: Psychiatric illness in Taiwan and the United States. *American Journal of Psychiatry, 148*(12), 1697–1704.

Kleinman, A. (1982). Neurasthenia and depression: A study of somatization and culture in China. *Culture, Medicine, and Psychiatry, 6,* 117–190.

Kleinman, A., & Kleinman, J. (1985). Somatization: The interconnections in Chinese society among culture, depressive experiences, and the meanings of pain. In A. Kleinman & B. Good (Eds.), *Culture and depression: Studies in the anthropology and cross-cultural psychiatry of affect and disease* (pp. 429–490). Los Angeles: University of California Press.

Ko, Y. H. (1977). *Ko-Shi-Shin-Ger-Leong-Beau-Sou-Tzer (Manual of Ko's Mental Health Questionnaire)* [in Chinese]. Taipei, Taiwan: Chinese Behavioral Sciences Press.

Ko, Y. H. (1995). *Ko-Shi-Shin-Ger-Leong-Beau-Shou-Din-Yen-Gio (Validation Study of Ko's Mental Health Questionnaire)* [in Chinese]. Taipei, Taiwan: Ministry of Education.

Ko, Y. H. (1996). *Gen-Kon-Shin-Ger-See-Guong-Leong-Beau-Sou-Tzer (Manual of Health, Personality, Habit Scale)* [in Chinese]. Taipei, Taiwan: Testing Press.

Lin, T. Y. (1953). A study of the incidence of mental disorders in Chinese and other cultures. *Psychiatry, 16,* 313–336.

Lin, T. Y. (1982). Culture and psychiatry: A Chinese perspective. *Australian and New Zealand Journal of Psychiatry, 16,* 235–245.

Lin, T. Y., Rin, H., Yeh, E. K., Hsu, C. C., & Chu, H. M. (1969). Mental disorders in Taiwan fifteen years later. In W. Caudill & T. Y. Lin (Eds.), *Mental health research in Asia and the Pacific*. Honolulu, HI: East-West Center Press.

Shieh, B. L., & Sue, S. (1996). Deciphering psychopathology from cultural response sets: Asians and MMPI-2 performance. Abstract presented at the Annual Meeting of Chinese Psychological Association, Taipei, Taiwan.

Sue, D. W., & Sue, D. (1990). Counseling Asian Americans. In D. W. Sue & D. Sue (Eds.), *Counseling the culturally different: Theory and practices* (pp. 189–208). New York: Wiley.

Tseng, W. S. (1972). Psychiatric study of shamanism in Taiwan. *Archives of General Psychiatry, 26*(6), 561–565.

Tseng, W. S., & Hsu, J. (1969). Chinese culture, personality formation and mental illness. *International Journal of Social Psychology, 16,* 5–14.

Tseng, W. S., & McDermott J. F., Jr. (1975). Psychotherapy: Historical roots, universal elements, and cultural variations. *American Journal of Psychiatry, 132*(4), 378–384.

Wallace, A. F. C. (1956). Revitalization movements. *American Anthropologist, 58,* 264–281.

Weissman, M. M., & Cross-National Collaborative Group (1992). The changing rate of major depression. *Journal of American Medical Association, 268*(21), 3098–3105.

Wen, J. K., Lin, S. L., Chen, C. C., Chou, W. G., & Huang, H. L. (1992). Lin-Huang-Fu-Shen-Sheng-Seong: Tai-Wan-Ben-Tu-Der-Ea-Li-Ing-Eng-Shin-Wei (Spirit possession phenomena: Stress-Coping Behavior in Taiwan) [in Chinese]. In *Proceedings of Institute of Ethnology* (Vol. 73, pp. 1–32). Taipei, Taiwan: Institute of Ethnology, Academia Sinica.

Wen, J. K., Lin, S. L., & Chen, Y. P. (1993). Lin-Hung-Fu-Shen-Sheng-Seong-Gin-Shen-Gee-Bin-Yu-Shin-Li-Sher-Hui-Win-Hwa-Ing-Shu (Spirit possession phenomena, mental illnesses, and Psycho-socio-cultural factors) [in Chinese]. In K. S. Yang (Ed.), *Culture, psychological disorders and treatments*. Taipei, Taiwan: Indigenous Psychological Research Center Press.

Yamamoto, J., Yeh, E. K., Loya, F., Slawson, P., & Hurwicz, M. L. (1985). Are American Psychiatric outpatients more depressed than Chinese outpatients? *American Journal of Psychiatry, 142*(11), 1347–1351.

Yang, K. S. (1972). Expressed values of Chinese college students [in Chinese]. In Y. Y. Li & K. S. Yang (Eds.), *Symposium on the character of the Chinese: An interdisciplinary* (pp. 257–312). Taipei, Taiwan: Institute of Ethnology, Academia Sinica.

Yang, K. S. (1985). "The change of personality and behavior among the Chinese people in Taiwan" [in Chinese]. In *Proceedings of the Conference on Modernization and Its Problems in Taiwan* (pp. 75–100). Taipei, Taiwan: Institute of Three People's Principles, Academia Sinica.

Yang, K. S. (1993). Can traditional and modern values coexist? [in Chinese]. In K. S. Yang (Ed.), *Chinese values: A social science viewpoint* (pp. 65–120). Taipei, Taiwan: Kwai Kuan.

Yang, K. S. (1995). Chinese social orientation: An integrative analysis. In T. Y. Lin, W. S. Tseng, & Y. K. Yeh (Eds.), *Chinese societies & mental health* (pp. 19–39). Hong Kong: Oxford University Press.

Yee, D. H. (1985). *Tai-Wan-Min-Shu-Chin-Li-Fu-Dau (Folk counseling in Taiwan)* [in Chinese]. Taipei, Taiwan: Teacher Chang Publisher.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.17
Clinical Psychology in Aotearoa/ New Zealand: Indigenous Perspectives

TAIMA M. MOEKE-PICKERING, MAHALIA K. PAEWAI, AMELIA TURANGI-JOSEPH, and AVERIL M. L. HERBERT
University of Waikato, Hamilton, New Zealand

10.17.1 INTRODUCTION 349
10.17.2 A BRIEF HISTORY 350
  10.17.2.1 Maori Identity 350
  10.17.2.2 The Treaty of Waitangi 350
10.17.3 PSYCHOLOGY IN AOTEAROA 350
  10.17.3.1 Historical Context 350
  10.17.3.2 The State of Maori Mental Health 351
10.17.4 DEVELOPING A BICULTURAL MODEL IN AOTEAROA 351
  10.17.4.1 The Need for Change 351
  10.17.4.2 A New Approach to Maori Mental Health 351
10.17.5 PSYCHOLOGY IN A BICULTURAL SETTING 351
  10.17.5.1 The Changing Face of Psychology 351
  10.17.5.2 Training in Psychology 352
  10.17.5.3 National Standing Committee on Bicultural Issues 352
10.17.6 MAORI AND PSYCHOLOGY 352
  10.17.6.1 Research, Training, and Teaching 352
  10.17.6.2 Cultural Safety 353
  10.17.6.3 Cultural Safety and Clinical Practice 353
  10.17.6.4 Maori Psychological Initiatives 353
10.17.7 CONCLUSION 354
10.17.8 REFERENCES 354

10.17.1 INTRODUCTION

This chapter provides an indigenous perspective of psychology in Aotearoa (the Maori name for New Zealand) by examining bicultural developments in the fields of teaching, research, theory, and practice. The key issue for Maori (the indigenous people of Aotearoa) as psychology has evolved from early British to later American influences is the relevance of Western based psychological theory and practice. This raised questions about the need to develop indigenous paradigms in Aotearoa. The Treaty of Waitangi provided a unique basis for a bicultural model to be developed. Bicultural and Maori initiatives have played a significant role in psychology training programs. In particular, the development of Maori-initiated psychological programs demonstrate an increased acceptance of cultural dimensions in psychology.

10.17.2 A BRIEF HISTORY

10.17.2.1 Maori Identity

To understand the development of psychology in Aotearoa as well as its responsiveness to the needs of Maori, the historical and cultural context in which the discipline has developed must be considered. The dominant ethnic group in Aotearoa are Pakeha (mainly New Zealanders of predominantly European–British descent). Like the experience of many indigenous peoples from other settler colonies, Maori continue to strive for self-determination in an environment of Pakeha dominance (Hamerton, Nikora, Robertson, & Thomas, 1995; Jackson, 1988; Walker, 1990).

Prior to Pakeha contact, Maori identified themselves primarily from the strata of their tribal structures, these being *whanau* (extended family), *hapu* (subtribe), *iwi* (tribe), and *waka* (ancestral canoe linking *iwi* and the peoples of the Pacific). Tribal structure and organization were based on descent from a common Maori ancestor (Buck, 1949). Identity was derived from kinship and learning within these tribal structures (Moeke-Pickering, 1996).

Cultural practices were based on a shared system of understanding by a tribal group that were deemed to be important and meaningful to them (Moeke-Pickering, 1996). Language, customs, spirituality, respect for the land, kinship obligations, and traditions were fundamental to the socialization of Maori identities. In essence historical references, cultural practices, kinship structures, and the land served as charters for Maori identities (Ritchie, 1992).

10.17.2.2 The Treaty of Waitangi

During the period of early contact with Pakeha, the sovereign status of Maori was internationally recognized in the Declaration of Independence in 1836. Maori were involved in international trade and were both thriving and flourishing economically. The lawlessness of many early Pakeha settlers, however, was causing concern among Maori leaders. They resolved to enter into a constitutional arrangement with the British Crown which would allow the British to control the unruly behavior of its citizens resident in Aotearoa (Orange, 1987). The vehicle for this was to be the Treaty of Waitangi which was signed on the February 6, 1840 and marked the beginning of a new era for Maori. The treaty prescribed a proposed constitutional relationship between the Crown and Maori that outlined the duties and obligations of both parties. In delegating limited authority to the Crown, the Treaty of Waitangi guaranteed Maori sovereignty or chieftainship over their lands, their settlements, and all other property (Jackson, 1988). Written in both the Maori and English language, ambiguities in interpretation between both texts have caused much debate until the late 1990s (Jackson, 1988; Orange, 1987; Renwick, 1990).

In the intervening 156 years, the principles of the Treaty of Waitangi have been breached repeatedly through Pakeha hegemony and the ethnocentric attitudes that underpin it (Ballara, 1986). Maori have faced the trauma of loss of land, language, and dignity, as well as absolute control over their ability to maintain a separate culture and identity (Vasil, 1988). As a result they suffered cultural alienation, poverty, unemployment, and hardship which ultimately led to further entrenchment and subordination to dominant Pakeha cultural ways (Awatere, 1984; Lawson-Te Aho, 1984; Walker, 1990).

Clearly, both Maori and Pakeha ethnic identities have developed, out of a binary opposition between the two people with Pakeha becoming the dominant ethnic group (Walker, 1989). Today Maori people comprise 12% of the population, control less than 5% of the land, and continue to struggle to maintain a positive cultural identity.

10.17.3 PSYCHOLOGY IN AOTEAROA

10.17.3.1 Historical Context

Psychology modeled on the British system developed as a discipline within Aotearoa (Hamerton et al., 1995). The introduction of psychology was part of the imposition of a colonial knowledge system that systematically undermined Maori ways of knowing in favor of a Western world view. Evidence of this was clear in the way that the Western psychological paradigm was hierarchically positioned over and above Maori epistemology. Psychology had a status as a scientific discipline at a time when Maori knowledge was perceived as an inferior theoretical framework (Lawson-Te Aho, 1984; Stewart, 1995).

At that time, the application of a psychological paradigm foreign to Maori experiences became accepted practice by psychologists and was deemed to be appropriate despite obvious ethnocentric biases inherent within this framework. An example was the use of intelligence tests to stream Maori school children used in the

1950s and 1960s. Predictably, these tests were demonstrated to be both culturally biased and an unfair means of assessing the ability of Maori due to the culturally bound nature of the constructs used (Thomas, 1988).

10.17.3.2 The State of Maori Mental Health

During the 1990s it became clear that although Maori were becoming mentally ill at a similar rate to Pakeha, poorer outcomes were being experienced by Maori during the post-admission phases of treatment (Bridgman, 1993). Drug and alcohol related problems as well as other psychoses were the main categories underpinning Maori admissions. Although the actual statistics need careful examination, the overall trend highlighted the fact that the fastest growth area in terms of hospitalization was the readmission of Maori with serious psychotic illness (Awatere, 1984; Bridgman, 1993). Bridgman (1993) also suggested that psychiatric diagnoses for Maori presented a distorted picture with inaccurate assessments and ineffective hospitalization. Psychiatric statistics for Maori were less indicative of the levels of psychopathology in the population and more an indication of the inability of health providers to understand the social, educational, and economic realities faced by many Maori. Furthermore, attitudes in service agencies had failed to recognize or accommodate cultural differences (Spoonley, 1988).

10.17.4 DEVELOPING A BICULTURAL MODEL IN AOTEAROA

10.17.4.1 The Need for Change

The inability of the health sector to respond appropriately to many Maori clients became increasingly obvious to many health professionals and community workers and led to an examination of different approaches to Maori health (Durie, 1994). While Durie (1994) was able to identify earlier Maori health professionals who had established appropriate cultural practices, there was no widespread recognition for such contributions until the mid-1970s.

At this time the government finally addressed over a century of injustices with the passing of the Treaty of Waitangi Act 1975. This provided for the first time a recognized forum for Maori to air grievances and to advocate Treaty obligations in a bicultural milieu (Kawharu, 1989). This led to Maori people being legally recognized as “tangata whenua” or indigenous people of the land, with rights under the Treaty to their own language and cultural self-determination (Abbot & Durie, 1987). The passing of the 1975 Act and the Treaty itself gave impetus to claims by Maori that health policies and initiatives should reflect their rights to development and self-determination.

10.17.4.2 A New Approach to Maori Mental Health

By the 1980s, challenges to the discipline of psychology to recognize the significance of cultural issues further increased. A bicultural approach which served the health needs of Maori needed to be developed based on increased recognition of cultural differences between Maori and Pakeha. In practice, biculturalism as an approach to Maori health initially attempted to include a Maori perspective into existing health systems.

With the increasing awareness of a range of culturally determined behaviors, some professionals suggested that Maori were misdiagnosed as psychotic and inappropriately referred to psychiatric services. Of particular significance within the health field were Maori assertions that a bicultural model should provide equitably for both Treaty partners. Furthermore, provisions already existed in the articles of the Treaty of Waitangi to justify these claims. Durie (1994) noted that by 1985 the Standing Committee on Maori Health had recommended that the Treaty of Waitangi be regarded as the foundation for good health.

In the mental health setting, the application of a bicultural model in the 1980s relied upon good faith and agreed-upon-goals rather than formal conventions (Durie, 1994). As a result, the parameters of the bicultural model were not clearly defined and the rights and responsibilities of Maori and Pakeha were poorly understood. Some programs gave recognition, but little else, to the cultural traditions of Maori (Mulgan, 1989a) while others favored a redistribution of resources to Maori (Jackson, 1988). At one level there was an implied inclusion of Maori values in mainstream institutions, at another the development of specific Maori institutions to provide for Maori needs (Sharp, 1995). Regardless of these difficulties, the bicultural model served as a framework for other institutions, agencies, and social services to extend upon.

10.17.5 PSYCHOLOGY IN A BICULTURAL SETTING

10.17.5.1 The Changing Face of Psychology

Awatere (1981), a noted Maori psychologist, raised the notion that certain psychological techniques and practices used the privilege of power to further oppress Maori cultural processes. She expanded on this by asserting that:

> My job as a psychologist is to facilitate the process of understanding, confronting and changing; to use psychological techniques to help the powerless gain power instead of what psychologists have always done, which is to help maintain their power. (p. 202)

Waldegrave (1985), in developing therapies for Maori and Samoan families, found that the predominantly monocultural and monoclass approach was inadequate due to the existence of broader social and political issues which impacted on these cultural groups. He pointed out that psychologists were not required to demonstrate any knowledge of Maori culture at all, and that psychological knowledge, including family therapy, was considered sufficient to address the problems of both Maori and Pakeha families.

10.17.5.2 Training in Psychology

In the United States, Yutrzenka (1995) noted that until the 1970s there was little inclusion of cross-cultural material, as well as limited diversity in the client populations with whom trainees received their clinical service training. Increasing awareness in the 1980s and 1990s saw a change towards including more cross-cultural theory and practices in psychology training programs. Many professional psychologists were aware of this need to prepare for a multicultural, multiracial, and multiethnic environment, but as Yutrzenka (1995) commented, the commitment and implementation of cross-cultural understanding was highly variable across these programs.

To some extent this overseas trend was also being reflected in Aotearoa. For example, in 1987, a survey of the nine postgraduate training programs in clinical, educational, and community psychology in Aotearoa was completed (Abbot & Durie, 1987). The findings of this survey confirmed that few programs incorporated "taha Maori" (a Maori dimension). The monocultural training of psychologists compared unfavorably with social work and medical training. Both Abbot and Durie and later Brady (1992) questioned the credibility of the discipline and challenged it to: (i) increase the number of trained Maori psychologists; (ii) develop an increased awareness of entrenched attitudes; and (iii) invest more energy into making changes at all levels of training including format, assessment methods, texts, and ethics.

In 1990, a study conducted by Sawrey (1990) further illustrated the need to include "taha Maori" into training programs and to develop a Maori psychology as well as include Maori psychology programs within mainstream education. In this study of 163 clinical psychologists of whom only one was Maori, over 85% of the sample supported the notion that psychologists' knowledge of Maori culture was an important factor in good psychotherapeutic outcomes with Maori clients. Interestingly, 75% of the sample felt they had inadequate Maori knowledge to effectively work with Maori, while 85% believed that their psychological training had not equipped them sufficiently to effectively work with Maori clients.

10.17.5.3 National Standing Committee on Bicultural Issues

The report of Abbot and Durie (1987) presented a damning critique of psychological training. The report served as the catalyst for change in psychological training and two years later the National Standing Committee on Bicultural Issues (NSCBI) was established under the auspices of the New Zealand Psychological Society.

The NSCBI was formed to monitor the development of bicultural activities, organize symposia, disseminate information on issues relevant to indigenous development, and to develop and encourage policies and practices that reflected New Zealand's cultural diversity (Hamerton et al., 1995). The Committee's broad aims were to initiate social changes that supported recognition and development of Maori psychology, to assist the New Zealand Psychological Society to honor its obligations under the Treaty of Waitangi, and to assist psychologists to develop appropriate teaching, research, and practice (NSCBI, 1994a). These activities have been effective in instituting cultural agendas at both the national and regional levels.

10.17.6 MAORI AND PSYCHOLOGY

10.17.6.1 Research, Training, and Teaching

By 1996 most university psychology departments and psychological service providers have either considered, or put in place, initiatives aimed at supporting Maori students and practitioners. While bicultural initiatives in Aotearoa have served to increase acceptance of cultural issues in psychology, it had also become apparent that there was a need to develop a body of research relevant to Maori and psychology. There has been a considerable amount of research and literature relating to Maori and psychology which is now having some impact on both professional training and mental health services (Awatere, 1981; Durie, 1994; Lawson-Te Aho, 1984; McFarlane-Nathan, 1996; NSCBI, 1994b, 1995; Paterson, 1992; Ritchie & Ritchie, 1978; Stewart, 1995; Thomas & Nikora, 1992). In addition, the experiences of other colonized indigenous people overseas have relevance for Maori by serving to direct attention on similar issues within the context of Aotearoa (McFarlane-Nathan, 1996).

There are two universities in Aotearoa that currently offer courses which focus on Maori topics relating to psychology. In addition, two other universities have made an explicit commitment to developing course material and/or structures relevant to Maori in psychology (Stewart, 1995). These courses have had the effect of increasing Maori participation in psychology. An increase in the number of Maori enrolling in psychology training programs coupled with increased retention of Maori both within the discipline and as practitioners have resulted in a larger pool of Maori with relevant research and practical expertise. With increasing numbers both in practice and research, Maori are better placed to develop more appropriate theoretical frameworks.

The frameworks referred to above contribute to a psychology that services the needs of Maori in a way that is relevant to their needs and consistent with their world view. It has been necessary to develop a Maori resource base from which an understanding and awareness of cultural knowledge, skills, and competence can be drawn. Students are encouraged to read literature describing colonization, Maori history, Maori traditions, as well as tribal and cultural knowledge to gain background information pertaining to Maori people. Durie (1994), Ritchie (1992), and Walker (1990) provide comprehensive information on biculturalism, the impact of colonization, the development of bicultural policies, tribal development and identity, and Maori health.

Current research in the areas of cross-cultural psychology (Thomas, Nikora, & Moeke-Pickering, 1996), identity issues (Durie et al., 1995; Heperi, 1996; Hingston, 1993; Moecke-Pickering, 1996; Nikora, 1995b; Thomas & Nikora, 1996), and cultural practice issues (Paewai, 1996) have helped to shape alternative practices and strategies for working with Maori and provide insights into Maori conceptualizations.

10.17.6.2 Cultural Safety

While psychological literature contributes to a more informed trainee and practitioner, cultural approaches within other disciplines have assisted psychology training to provide strategies to improve cultural care, meaning, and sensitivity when working with Maori. One such approach is that of "cultural safety" which is described as

> ... actions which recognise, respect and nurture the unique cultural identity of tangata whenua, and safely meet their needs, expectations and rights. A culturally unsafe practice is (are) any actions which diminish, demean or disempower the cultural identity and well-being of an individual. (Ramsden, 1991, pp. 7–8)

Cultural safety was initiated by Maori nurses and later developed by the New Zealand Council of Nursing (Ramsden, 1991). As an approach, cultural safety has increased understanding of the different ways that helping professions might develop working guidelines. The focus of these guidelines is to enhance good teaching practices and strategies when educating health professionals. Cultural safety is now generally accepted in professional practice and training programs, including psychology.

10.17.6.3 Cultural Safety and Clinical Practice

Cultural safety incorporates two principles inherent in the Treaty of Waitangi. First, active protection of Maori Treaty rights, and second, recognition of all cultures in Aotearoa. A study by Paewai (1996) described various cultural safety strategies being employed by six Maori clinical psychologists working in Aotearoa. A significant finding in Paewai's study was that psychotherapeutic outcomes were improved if the clinician acknowledged the cultural identity of the client in assessment and treatment procedures.

To this end, some training programs are developing a curriculum that monitors, teaches, and evaluates cultural safety and competence. For example, at Waikato University, courses in cross-cultural psychology, community psychology, and Maori development and psychology highlight cultural safety concepts and approaches. These programs have been developed to assist psychology students to increase their cultural analysis by providing informed consideration of issues of culture, with the intent of avoiding ignorant rejection or violation of cultural norms (Nikora, 1995c).

10.17.6.4 Maori Psychological Initiatives

Three agency-based programs (one hospital-based and two in the Justice Department) developed specifically for working with Maori are presented in this section. The support of local tribes, Maori healers, and resource people have been essential (McFarlane-Nathan, 1996) to the success of these programs.

Whaiora is a ward based at a psychiatric hospital at Tokanui in the central North Island. It emerged from an increasing awareness among Maori nursing staff at Tokanui Hospital that the therapy strategies employed were culturally biased. They often alienated Maori clients (Durie, 1994). The successful development of a parallel Maori cultural therapy unit in 1986 was a direct attempt to redress the spiritual and psychological imbalance prevalent among Maori clients by implementing a more appropriate, holistic approach to mental health (Rankin, 1986). This approach encompassed many elements critical to Maori health and necessitated that a balance be maintained between each dimension for adequate well-being. One of these approaches is the Maori health model that incorporates the *tinana* (physical body), *hinengaro* (thoughts and aspects of the mind), *whatumanawa* (feelings and deep emotions), and *whanau* (extended family). The model provides a framework that assists in the understanding of Maori processes and people. In addition, Whaiora seeks to restore and enhance self-esteem, self-worth, and self-confidence by encouraging and promoting a sense of being and pride in the client's own cultural heritage (Rankin, 1986). The unit places strong emphasis on *whanau* involvement and community support networks which are fundamental to a smoother transition for Maori clients returning to the community and decreasing readmission rates.

Te Piriti Special Treatment Unit is a relapse prevention program for sex offenders based in Auckland. The main goal of the program is to reduce sexual reoffending and assist offenders and *whanau*/support people through the initial rehabilitation process. This organization has a Maori "cultural consultant" who is involved at both the assessment and intervention stage of the treatment. The assessment procedure begins with building trust then moving gradually into Maori identity issues such as place of origin, *whanau*, *hapu*, and *iwi*. Whakapapa (genealogy) is discussed, along with *tikanga* (beliefs and traditions) and *te reo* Maori (Maori language). All of these variables are fundamental to Maori identity and hence, play a key role in treatment procedures.

The Bicultural Therapy Model is also being currently piloted through the office of Psychological Services in Auckland. The Bicultural Therapy Model draws many elements from the work of Renfrey (1992) which recognizes the stressors caused by acculturation and/or deculturation. Cultural supervision of the clinician ensures that the necessary knowledge base is being maintained to enable satisfactory assessment of Maori clients. Clinicians require training which focuses on the cultural context within which Maori operate (McFarlane-Nathan, 1994). They must also develop an awareness of their own limitations for dealing with Maori clients. At the initial assessment stage, the trained clinician is encouraged to consider the possibility of deculturation and/or acculturation stressors in the assessment of etiology. The Bicultural Therapy Model recognizes that cognitive behavior therapy as a therapeutic approach provides a framework for the culturally-competent clinician to investigate the social and behavioral context of the client. The client's level of bicultural competence is also assessed to establish the "degree to which clients possess sufficient cultural resources to enable them to develop successfully as indigenous people in Western society" (McFarlane-Nathan, 1996).

10.17.7 CONCLUSION

Bicultural and Maori initiatives in psychology have contributed to an evolving understanding of cultural variables when working with Maori people. Contributions made by Maori psychologists, Maori students, and Maori-initiated programs have helped to refine a resource base where theory, research, and skills relevant to Maori development and psychology can be developed. As Maori are appointed to senior positions in health agencies and universities, there will be increasing opportunities for Maori initiatives in research and the development and application of indigenous psychological paradigms as well as Maori control of services for Maori.

10.17.8 REFERENCES

Abbot, M. W., & Durie, M. H. (1987). A whiter shade of pale: Taha Maori and professional psychology training. *New Zealand Journal of Psychology, 16,* 58–71.

Awatere, D. (1981). Maori counselling. In F. Donnnelly (Ed.), *A time to talk: Counsellor and counselled* (pp. 198–202). Sydney, Australia: Allen & Unwin.

Awatere, D. (1984). *Maori sovereignty*. Auckland, New Zealand: Broadsheet.

Ballara, A. (1986). *Proud to be white: A survey of Pakeha prejudice in New Zealand.* Auckland, New Zealand: Heinemann.

Brady, E. C. (1992). Cross-cultural issues in clinical psychology training programs. *New Zealand Journal of Psychology, 21,* 56–61.

Bridgman, G. (1993). *Nga ia o te oranga hinengaro Maori: Trends in Maori mental health—a discussion document.* Wellingon, New Zealand: Te Puni Kokiri Ministry of Maori Development.

Buck, P. T. (1949). *The coming of the Maori.* Wellington, New Zealand: Maori Purposes Fund Board and Whitcombe & Tombs.

Durie, M. H. (1994). *Whaiora: Maori health development*. Auckland, New Zealand: Oxford University Press.

Durie, M. H., Black, T. E., Christensen, I. S., Durie, A. E., Taiapa, U. K., Potaka, U. K. J., & Fitzgerald, E. (1995, October). Te hoe nuku roa framework—A Maori identity measure. *ASSR NEWS* (Published Monthly by the Association of Social Science Researchers) 1–6.

Hamerton, H., Nikora, L. W., Robertson, N., & Thomas, D. (1995). Community psychology in Aotearoa/New Zealand. *The Community Psychologist* (Published by The Society for Community Research and Action) *28*(3), 21–23.

Heperi, L. (1996). *Awhi atu, awhi mai—Help seeking and Maori*. Unpublished master's thesis. Hamilton, New Zealand: University of Waikato.

Hingston, O. M. (1993). *Ethnicity and oppression: The perceptions of seven Maori individuals*. Unpublished master's thesis. Hamilton, New Zealand: University of Waikato.

Jackson, M. (1988). *The Maori and the criminal justice system: He whaipaanga hou—A new perspective (Part 2)*. Wellington, New Zealand: Department of Justice, Policy and Research Division.

Kawharu, I. H. (1989). Introduction. In I. H. Kawharu (Ed.), *Waitangi: Maori and Pakeha perspectives of the Treaty of Waitangi* (pp. x–xxiv). Auckland, New Zealand: Oxford University Press.

Lawson-Te Aho, K. (1984). *The masters tools: Maori development inside Pakeha psychology*. Unpublished master's thesis. University of Waikato, Hamilton, New Zealand.

McFarlane-Nathan, G. (1994). *Cognitive behaviour therapy and the Maori client*. Unpublished paper. Department of Justice, Psychological Services Division, Auckland, New Zealand.

McFarlane-Nathan, G. (1996). *The bicultural therapy project*. Unpublished paper. Department of Corrections, Psychological Service, Auckland, New Zealand.

Moeke-Pickering, T. M. (1996). *Maori identity within whanau*. Unpublished master's thesis. University of Waikato, Hamilton, New Zealand.

Mulgan, R. (1989a). *Maori, Pakeha and democracy*. Auckland, New Zealand: Oxford University Press.

National Standing Committee on Bicultural Issues (NSCBI) (1994a). The New Zealand Psychological Society and the Treaty of Waitangi: Proposed implementation plan. *New Zealand Psychological Society Bulletin, 82,* 13–17.

National Standing Committee on Bicultural Issues (NSCBI) (1994b). An interview with Linda Waimarie Nikora, out-going convenor of the NSCBI. *New Zealand Psychological Society Bulletin, 83,* 6–8.

National Standing Committee on Bicultural Issues (NSCBI) (March). How does the Treaty of Waitangi (1840) relate to a science that developed after 1890? *New Zealand Psychological Society Bulletin, 84,* 8–11.

Nikora, L. W. (1995b). *The maintenance of Maori tribal identities in Aotearoa: Rationale and overview of working papers*. Unpublished paper. Waikato University, Hamilton, New Zealand.

Nikora, L. W. (1995c). *Cultural safety*. Unpublished paper. University of Waikato, Hamilton, New Zealand.

Orange, C. (1987). *The Treaty of Waitangi*. Wellington, New Zealand: Allen & Unwin.

Paewai, M. K. (1996). *Cultural safety within clinical psychology—A Maori perspective*. Unpublished paper. Waikato University, Hamilton, New Zealand.

Paterson, K. (1992). *Wahine maia, me te ngawari hoki: Maori women and periodic detention in Kirikiriroa: Their life difficulties and the system*. Unpublished master's thesis. University of Waikato, Hamilton, New Zealand.

Ramsden, I. M. (1991). *Kawa whakaruruhau: Cultural safety in nursing education in Aotearoa (New Zealand)*. Wellington, New Zealand: New Zealand Council of Nursing.

Rankin, J. F. A. (1986). Whaiora: A Maori cultural therapy unit. *Community Mental Health in New Zealand, 2,* 38–47.

Renfrey, G. (1992). Cognitive behavior therapy and the Native American client. *Behavior Therapy, 23*(3), 321–340.

Renwick, W. (1990). *The Treaty now*. Wellington, New Zealand: GP Books.

Ritchie, J. (1992). *Becoming bicultural*. Wellington, New Zealand: Huia/Brasell.

Ritchie, J., & Ritchie, J. (1978). *Growing up in New Zealand*. Auckland, New Zealand: Allen & Unwin.

Sawrey, R. (1990). *A survey of psychologists' opinions and behavior on aspects of Maori mental health*. Unpublished master's thesis. Victoria University, Wellington, New Zealand.

Sharp, A. (1995). Why be bicultural? In M. Wilson & A. Yeatman (Eds.), *Justice and identity: Antipodean practices* (pp. 116–133). Wellington, New Zealand: Williams Books.

Spoonley, P. (1988). *Racism and ethnicity*. Auckland, New Zealand: Oxford University Press.

Stewart, T. R. (1995). *Ka pu te ruha, ka hao te rangatahi: Contributions to indigenous psychology in Aotearoa/New Zealand*. Unpublished master's thesis. University of Auckland, Auckland, New Zealand.

Thomas, D. R. (1988). Development of a test of Maori knowledge. *New Zealand Journal of Psychology, 17,* 59–67.

Thomas, D. R., & Nikora, L. W. (1992). From assimilation to biculturalism: Changing patterns in Maori-Pakeha relationship. In D. R. Thomas & A. Veno (Eds.), *Psychology and social change: Australian and New Zealand perspectives* (pp. 231–256). Palmerston North, New Zealand: Dunmore Press.

Thomas, D. R., & Nikora, L. W. (1996). Maori, Pakeha and New Zealander: Ethnic and national identity among New Zealand students. *Journal of Intercultural Studies, 17,* 29–40.

Thomas, D. R., Nikora, L. W. & Moeke-Pickering, T. M. (1996). *Readings in culture and ethnicity*. Department of Psychology, University of Waikato, Hamilton, New Zealand.

Waldegrave, C. (1985). Mono-cultural, mono-class and so called non-political family therapy. *Australia and New Zealand Journal of Family Therapy, 6,* 197–200.

Walker, R. (1989). Maori identity. In Novitz and Willmott (Eds.), *Culture and identity in New Zealand* (pp. 35–52). Wellington, New Zealand: Government Printer.

Walker, R. (1990). *Ka whawhai tonu matou: Struggle without end*. Auckland, New Zealand: Penguin.

Vasil, R. (1988). *Biculturalism: Reconciling Aotearoa with New Zealand*. Wellington, New Zealand: Victoria University Press.

Yutrzenka, B. A. (1995). Making a case for training in ethnic and cultural diversity in increasing treatment efficacy. *Journal of Consulting and Clinical Psychology, 63*(2), 197–206.

Copyright © 1998 Elsevier Science Ltd. All rights reserved.

10.18
Perspectives from Sub-Sahara Africa

PIUS K. ESSANDOH
Jersey City State College, NJ, USA

10.18.1 MISCONCEPTIONS ABOUT AFRICA 357

10.18.2 SOCIOCULTURAL BELIEFS 358

10.18.2.1 Spirituality 358
10.18.2.2 Collective Responsibility and Cooperation 359
10.18.2.3 Kinship 359
10.18.2.4 Gender Roles 360

10.18.3 OTHER THEORETICAL ISSUES 360

10.18.4 TRADITIONAL COMMUNITIES AND PSYCHOPATHOLOGY 360

10.18.5 THE THERAPEUTIC PROCESS 362

10.18.5.1 Assessment/Diagnosis 362
10.18.5.2 Prognosis 362
10.18.5.3 Treatment/Cure 363

10.18.6 TRAINING IMPLICATIONS 363

10.18.7 CONCLUSION 363

10.18.8 REFERENCES 364

10.18.1 MISCONCEPTIONS ABOUT AFRICA

Despite its size and the role it has played in both ancient and modern civilization, Africa continues to be relatively unknown to the Western world. The little that has been written about Africa in general, and African psychology in particular, is full of myths and misconceptions. For example, it was widely believed that depression, suicide, and affective disorders were either nonexistent or less prevalent in African communities and that in general, Africans experienced fewer psychological problems than peoples from other cultures. Although such myopic views have been challenged by several authors, who point to as many psychological disturbances in Africa as in other cultures (Binitie, 1981; German, 1987; Leighton, Lambo, & Hughes, 1963), there has not been a consistent body of research on the theory of psychopathology and treatment issues.

The primary focus of this chapter is to discuss the sociocultural factors in psychological health in sub-Sahara Africa, bearing in mind that any good definition of mental health and the delivery of mental health services in Africa must take these factors into account. The chapter also discusses issues of diversity (e.g., gender) as they affect the delivery of psychological services. It should be pointed out that psychological therapy, as practiced in the West, is almost nonexistent in most parts of Africa and as such, for many people it is not an option for the resolution of interpersonal conflicts and other psychological problems. It is also important to point out that sub-Sahara Africa is so massive in size and rich in cultural diversity that to attempt to write about Africa as if it contained one homogeneous group would be the greatest act of disservice. Most of the generalizations made in this chapter about Africans or Africa are more impressionistic than empirical and are drawn primarily from my own experiences as an African and my work with African college students in the USA, and also from the works of other indigenous African researchers.

A suggestion frequently made is that African traditional life and culture are primarily variants of folk culture that existed throughout the rest of the world in the preindustrial revolution era and are therefore not relevant for the postindustrial revolution era. This is in spite of the fact that most Western nations are desperately seeking some aspects of traditional life and culture as vital to national and individual well-being and as needed additions to industrial and economic development. It is commendable that despite the fact that sub-Sahara Africa is witnessing major political and economic changes typical of postindustrial revolution, most of the urban communities remain traditional and they continue to have a more stable and homogeneous cultural tradition. Writing about traditional healing in Africa and its implication for cross-cultural counseling, Vontress (1991) pointed out that even when they are outside of their countries, most African students in the USA retain their traditional beliefs; this presumably slows down the process of acculturation. It is reasonable to suggest then that when they are in Africa, urbanized Africans, in spite of some changes in their world views, may continue to hold a culturally pluralistic view of mental health and psychological well-being. Inherent in this cultural pluralism is group identity, collective responsibility, cooperation, interdependence, certain values and beliefs about etiological causes of mental illness, attribution of disturbances, and spirituality. These values are for the most part retained or revisited when individuals have to respond to some psychological symptoms of stress and alienation that industrialization and urbanization present. In most instances, urbanized Africans utilize both systems of care—traditional and Western—sequentially, if not simultaneously. They see these two systems not as contradictory but as reinforcing each other (Onyioha, 1977). The implications for psychological development and services have probably only been inferred or assumed at this time but clinical psychologists must understand some aspects of these sociocultural beliefs before they are able to work with clients from this culture.

10.18.2 SOCIOCULTURAL BELIEFS

African etiological theory of mental illness and psychological disorders is still profoundly rooted in sociocultural and spiritual foundations (Danquah, 1982; Essandoh, 1995; Wyllie, 1983). Enemies use witchcraft and other magical means to afflict their victims with mental illness. Binitie (1991) has indicated that it is the belief of some Africans that witches "can do harm from a distance without being in physical contact with their victims ... They terrorize the victims through their capacity to cause illness, misfortune, childlessness, poverty, and death" (p. 5); thus, presenting problems that traditional healers as well as clinical psychologists have to deal with. The gods and ancestors can afflict individuals with mental illness as well as provide cure and grant success and well-being. As agents of healing power and knowledge, they influence to a very large extent the moral development of individuals as well as the social changes that take place in the community. Binitie (1991) has pointed out further that witches are viewed as "judicial agents" and also as "agents of socialization" who help "to equilibrate society by preventing social greed and unscrupulous behavior" (p. 5). Individuals and communities become aware of the possible social consequences of their behavior and thus they show deep respect for customs and traditions and also work on improving interpersonal, family, and social relationships, very critical components of good mental health. When there are any lapses that lead to mental health or physical problems, traditional healers who serve as intermediaries between the spirit world and the physical world intercede with sacrifices and provide treatment. Traditional healers also serve as interpreters and teachers of social values and beliefs.

10.18.2.1 Spirituality

African traditional beliefs emphasize the importance of spirituality and external supernatural sources of help. There is a sense of reliance and dependence on a supreme being and on nature. People must either be subjugated to nature or live in harmony with it and submit to the authority of God or a deity to whom everybody is accountable. Faith in the supernatural influences most, if not all, healing. Thus, traditional healers and religious/spiritual healers become particularly powerful, not only because they share these same beliefs as their clients, but also because only they have the skills to lock up evil spirits as well as invoke the help of good spirits through prayer, sacrifices, and concoctions.

There is also the belief in reincarnation. This is expressed in ceremonies concerning life and death. For example, a new-born child is named after a dead ancestor and is sometimes believed to be this dead ancestor who has reincarnated. The link between the metaphysical and the physical must remain unbroken and the new-born is expected to develop such altruistic attributes of sharing and interdependence that characterized the life of the deceased relative. For living relatives, it is a great honor to have a new-born baby named after them. This is a testimony of their moral character and the respect enjoyed between the family and community. Young ones learn early in life that if they do good, they can reap the benefit in this current life or the life thereafter.

10.18.2.2 Collective Responsibility and Cooperation

Traditional African families derive their strength from cooperation and mutual helpfulness. Individual growth and development must be seen in the context of family growth and development. This collateral relationship which emphasizes self as part of a group, and people as more important than possessions, is different from other world views where individualism and task productivity are more important than relationships. The African is brought up with values that encourage mutual support in times of good fortune as well as in times of adversity, ill-health, and failure. The only contribution in life which significantly influences development is to cooperate with the family and group. Individual happiness means interdependence (not independence); there is no other means of meaningful life. Interdependence is a life-long virtue and a yardstick for measuring the "good life." Too much independence is interpreted as pride and a haughty attitude, and one is often punished by the gods for this kind of attitude. Within this culturally defined belief system that exalts interdependence above independence, the development of mental illness (and to some extent physical illness) depends on the disturbed person's interpersonal relationship; recognizing the importance of interdependence can ensure good psychological health, but not recognizing this important aspect of human development can lead to poor psychological health.

Collective responsibility and mutual cooperation also demand that parents (including the extended family of uncles, aunts, etc.) provide for the welfare and security of children who in turn will ensure the security of parents in their old age. This view is nicely summarized by a Ghanaian proverb (Akan) which, when literally translated into the English language, states that if you depend on people during infancy when you are growing your teeth, you reciprocate this gesture by helping these individuals during old age when they lose their teeth. This kind of reciprocal expectation, when not properly resolved, could lead to anxiety over one's inability to be a good provider for one's children or parents. For the African parent, this relationship could lead to good mental health and respectability in the community, as well as provide social security during old age. For the African child, failure to live up to this filial obligation has very serious social and mental health consequences.

Even on the international scene, African leaders have often recognized the importance of collateral relationships. For example, Nkrumah of Ghana articulated very strongly and clearly the notion that Ghana's independence (and by extension, development) was meaningless unless it was tied to the total emancipation and development of all Africa. Nkrumah's new republic was not ready for competitive nationalism. Instead, he was ready to cooperate in a harmonious relationship with the rest of Africa for its political as well as economic development. Thus, the African value of interdependence and collective responsibility is clearly demonstrated not only at the individual/family level but also at community, national, and continental levels.

10.18.2.3 Kinship

One of the most distinctive features of traditional African communities is the importance of kinship ties and communal style of living. All activities of social life—economic, political, and religious—are based on such kinship relations. Such ties bring people together for very important family activities like farming and harvesting (e.g., the Ndoboa system among the Akans of Ghana where families help each other in cultivating the land and in harvesting the produce), the birth of a child, puberty rites, and marriage and funeral ceremonies. Feeling closely related to kin offers security in all aspects of family life—economic, social, and psychological—and it gives expression to the sense of interdependence and cooperation. A strong social support network, very instrumental in preventing some psychological problems (e.g., stress, depression, uncomplicated bereavement), develops. Practitioners must recognize and utilize this network in the delivery of psychological services.

10.18.2.4 Gender Roles

A number of researchers have estimated that the ratio of African males to females in terms of psychiatric hospitalization is two to one (Dawson, 1964; Lamptey, 1977; Orley, 1972). Yet, despite this difference in the rate of hospitalization, it appears that African women present with higher incidence of depression and other psychosomatic disorders than their male counterparts. Kisekka (1990) has speculated that the low hospitalization rate among women may well be due to the fact that women use alternative forms of treatment (such as spiritual healing and traditional healing); or that because men are over-represented in the formal employment sector, they have access to free medical care; and that men are over-represented in cases of mandated admission from law enforcement agencies, particularly the police. Kisekka has further hypothesized that gender-role stressful situations such as a "barrenness, failure to have the desired number and sex of children, unwanted pregnancies, divorces, matrifocality, and early and forced marriages" (p. 10) may impact women more significantly than they do men. The etiology and nature of these emotional disturbances in Africa seem to point to sociocultural factors rather than intrapsychic factors. If this assertion is accurate, it will be imperative for psychologists and other mental health professionals to consider sociocultural factors in psychopathology and treatment as an integral part of competent professional practice.

10.18.3 OTHER THEORETICAL ISSUES

Although it is clear from the discussion of the sociocultural beliefs that differences exist between sub-Sahara Africa and the Euro-American culture, it is important to point out that differences exist also in terms of the psychological characteristics of individuals from these different cultures. There is evidence of variation in the definition of some psychological constructs. Traditional psychological theories have often defined development from a Eurocentric world view. This world view is middle class, male, white, and with Protestant values in orientation. Development is conceptualized in terms of self-actualization, separation, differentiation, independence, respect for personal boundaries, and self-development. All this takes precedence over group and community interests. But even in the West, feminist critics and multicultural experts have challenged such Eurocentric developmental theories (Asante, 1987; Atkinson, Morten, & Sue, 1993; Cross, 1971; Gilligan, 1982; Jackson, 1975; Miller, 1976) and they have pointed out that women, as well as the culturally different way of knowing and developing, are somehow different from the stages proposed by these Euro-American theories.

Most developmental theories in psychology and education follow a stage-wise, linear, predictable, and hierarchical progression. They measure development in stages across life-span and this life-span is divided into prenatal, infancy, early childhood, late childhood, adolescence, young adulthood, middle age, old age, and death. To most Africans, life-span development has very few stages—prenatal, childhood, adulthood, and old age. The kind of identity confusion that faces Western children as they make the transition from late childhood to adolescence simply may not exist for the African child, who becomes an adult overnight after certain rites of passage ceremonies have been performed. In a similar vein, using Eurocentric theories to measure constructs like self-esteem, self-actualization, autonomy, doubt, initiative, and guilt do not make sense. Development from an African frame of reference, including even life-span, is cyclical, nonlinear, multigenerational, as well as transgenerational and it measures constructs such as collaboration, communal values, and spirituality. In Africa, development in any facet must include character training (similar to Kohlberg's moral development) and only this is seen as beneficial to society. It is important to understand that if these theoretical constructs have different meaning for the African, then a transcultural description of psychopathology may not be possible after all and that the development of a common psychological language for communication between sub-Saharan African mental health professionals and their Euro-American counterparts is a long way from happening. This has important implications for training, research, and practice.

10.18.4 TRADITIONAL COMMUNITIES AND PSYCHOPATHOLOGY

The suggestion that traditional African life and culture are not relevant for the post-industrial revolution era raises a number of important questions. Should modernization—technological, industrial, and economical—always be at the expense of the social, spiritual, and communal values that have sustained the African continent for hundreds of years? Does the African culture help lay a foundation for a psychologically healthy personality strong enough to withstand the pressures of industrialization and urbanization? Answers to these questions may help explain how psychopathology develops in the sub-Sahara African context. If one adopts the unipolar way of thinking that folk culture is not relevant for development in the postindustrial revolution era, one will fail to recognize that the modern world, when alienated from social, spiritual, and communal values, experiences a breakdown in moral values that leads to the disintegration of families in particular and society in general. The argument for the retention of traditional cultural beliefs and practices also presupposes that Africans will enjoy an improved psychological health if they retain these beliefs. The truth, however, is that there are many aspects of traditional life that could either initiate psychological problems or exacerbate them.

The role of sociocultural factors in the development of psychopathology has not received any serious research consideration in sub-Sahara Africa. While cultural factors could be the source of cohesion, identity, and strength, and be instrumental in the development of good psychological health and the prevention of some serious psychological disorders, in some instances these same factors could be the source of psychopathology. For example, it is often believed that the extended family/kinship system provides personal and social security. However, literature is often silent on the fact that such deep commitment to the extended family could lead to stress, anger, intergenerational conflict, feelings of betrayal and shame, marital problems, and a host of other problems. Writing about bereavement and stress in career women in Nigeria. Kalu (1990) has indicated that friends and relatives take over "the care of children, household chores, and errands connected with funeral announcement and arrangement" (p. 79). This arrangement could prevent "depression and low morale as the result of major life events" (p. 75). She also indicates, however, that the contemporary Nigerian widow "is under stress not so much from the loss of the spouse but from the demands of a large and sometimes unwieldy group of significant others who participate in the funeral process" (p. 81). Is the source of this strain, discordance, and strife clinical or is it a conflict over the ways of achieving social values?

Another area where cultural factors become a source of stress, anxiety, and depression revolves around generativity and the expectation to have children to continue the cycle of life. The African society is family focused and it values child bearing and child rearing. In most instances, singleness and childlessness are not options, especially for the African woman. Kisekka (1990) notes the excessive stigma attached to barrenness and goes on to point out that "both traditional and modern folklore, literature and popular music are replete with depictions of barren women as lonely, malicious, [and] cruel to children" (p. 6). It appears that the feelings of isolation, loneliness, and even depression are the expressions of the problems of living in this culture and are culturally induced rather than being intrapsychically induced.

Although cultural factors could exacerbate and even initiate psychopathology, urbanized Africans may experience psychopathology because of the degree of their alienation from these traditional and cultural values. For example, Africans have become enslaved to the corrupting influences of individualism (the "me first" attitude), and they have made social and cultural adaptations from kin-bound societies and kinship ties that provided the much needed solidarity for physical as well as mental growth. The extent to which these individuals want to assimilate the West culture creates an identity confusion. The Western culture is very different from the African culture and as Taft (1977) has suggested, when the size of the gap between a familiar culture and an unfamiliar culture is very big, it is difficult for individuals to adjust and cope with the changes that they need to make. Thus, caught between retaining most of the traditional African values and beliefs that are incompatible with the values of a competitive, impersonal, and technologically oriented value system, and the desire to be accepted by the West; the social, spiritual, and communal values of Africans become very adulterated and compromised. As Some (1993) has observed, the "sweeping industrial imperialism" (p. 36) brought on Africans puts the African in two worlds: the traditional and the corporate. He further argues that:

> the corporate world dims the light of the traditional world by exerting a powerful magnetic shadow-like pull on the psyche of the individual. Thus the individual feels compelled to respond. But as he or she tries to respond, the individual begins to realize that the source of the pull is elusive. (p. 36)

To maintain a sense of self, there is the need to remain somewhat linked to "essential traditionalism" (Some, 1993). If this does not happen, the development of a healthy personality is affected. The African becomes marginalized and this creates some stressful conditions that may very well be misdiagnosed. When misdiagnosed these stressful conditions may not respond to traditional treatment.

10.18.5 THE THERAPEUTIC PROCESS

10.18.5.1 Assessment/Diagnosis

The issue of lack of trained mental health professionals in sub-Sahara Africa has been well documented (Danquah, 1982; Lamb, 1983; Okpaku, 1991). Thus, the variety of assessment methods described here are ones used primarily by traditional healers. However, it is important to recognize parallels between what traditional healers do and what Western-trained mental health professionals do or should do when working with sub-Sahara African clients. This has implication for training since the use of any assessment methods requires cultural knowledge on the part of practitioners. Recognizing parallels will also encourage cooperation between the two systems of helping.

Assessment in traditional healing often begins with the healer attempting to establish credibility and genuineness. Among some Nigerian tribes, for example, traditional healers initiate the session by offering cola nuts. The belief is that "he who brings cola, brings life" and as Onyioha (1977) further explains, cola suggests that "one is bound to be faithful to a man with whom one has eaten cola—you must not lie against him, you must not plan evil against him" (p. 214). Healers pledge to be faithful and truthful and they expect forthrightness from their clients. When this mutual pledge for truth and respect is established, clients are asked in a life-history or autobiographical approach (biographical approach if a client is not competent or capable of doing so) to tell their story to the healer. Although the African culture does not encourage self-disclosure outside of the family, the client now is ready to talk about everything that the healer is ready to listen to and the healer does not impose any constraint on what should be included in the story. The content of the story provides the data for diagnosis as well as for treatment planning. Traditional healers, therefore, listen very attentively asking very few questions. Essandoh (1995) has suggested that this life-history narrative sometimes includes multigenerational historical accounts and he has compared this to the use of genogram by family therapists in the West. It is amazing that with no recording devices, traditional healers are able to remember many of the details of what clients disclose to them; evidence of their effective listening skills.

The use of divination in diagnosis and treatment provides a link between therapeutic methods and the spiritual and ceremonial life of the African culture. If culture is truly a shared world view, the healer's use of divination respects this world view, in which ancestral spirits and supernatural powers play an important role both in the origin and treatment of diseases. Divination, religion, sorcery, and witchcraft all serve very important psychological and metaphysical functions in this context. The divination system includes the use of beads, amulets, rings, and other paraphernalia needed for the ritual and magical treatment as well as to provide protective charm for the client.

Other assessment techniques include behavior observational techniques by the healer, who will either move in and live with the family or ask the client and sometimes family members to move in and live with them (Vontress, 1991). Again, like the life-history approach, healers have no specific behaviors to observe. Instead they allow a more naturalistic observation to proceed but as difficulties in interpersonal relationship or other conflicts unfold, healers use this diagnostic information to their advantage. Yalom (1995) describes most forms of psychopathology as difficulties in interpersonal relationships, and psychotherapy (especially group psychotherapy) as an exercise to improve or repair interpersonal relationship. This is consistent with what most traditional healers do when they use this naturalistic observation technique.

10.18.5.2 Prognosis

Traditional healers may use any of the diagnostic assessment methods described above or several others to determine the prognosis for a client. In some instances they may use herbs or perform some initial rituals. If the disorder does not respond to these treatments, there is a rediagnosis, and other assessment methods are considered. Essandoh (1995) has described this practice as similar to the use of decision trees and differential diagnosis in clinical practice in the West and has suggested that when traditional healers rule in or rule out causes of disorders, there is the implication that something else (often supernatural) is causing the disease. Danquah (1982) has suggested that some disorders do not respond to initial treatment because traditional healers give instructions that are impossible to follow leading to a high "patient relapse rate" (p. 8) which benefits the healer. Rediagnosis, according to Danquah; involves giving the client "a few simple post-treatment instructions to enable him to carry them out daily with success, and less anxiety" (p. 8). In most instances, significant alleviation of symptoms happens, albeit temporarily leading both the client and the healer to believe that the prognosis is good.

10.18.5.3 Treatment/Cure

Traditional healers use a variety of techniques that have been empirically proven effective in Western psychotherapy. One of their many strengths is their ability to engender faith and hope in the therapeutic process (Frank, 1978; Yalom, 1995). Especially for spiritual healing, it is important for the client to have faith and hope in the process. Doubt and the lack of faith do not lead to significant therapeutic gains. In fact, psychotherapy and medical treatment in all societies benefit from faith and hope long before the other therapeutic factors kick in. Thus, the ability of traditional healers to mobilize their clients" hope becomes an important variable in the therapeutic process.

Traditional healers are also holistic in their approach. Treatment usually combines the physical, social, and spiritual in an effort to restore harmony to the client. Herbal treatment is employed where necessary and social and spiritual interventions are made as appropriate. Whenever necessary (although very infrequently) referral is made to another healer. In all instances, family and close neighbors participate in the treatment, making it easy for the healer to employ all available resources for support and effective treatment. What is good for the client is decided by all of them within the context of the plural medical systems; choosing to use the systems concurrently or simultaneously.

Treatment techniques also include traditional music and dancing, therapeutic rituals and sacrifices, suggestions, hypnosis, and other techniques similar to cognitive-behavioral approaches. The important thing throughout all counseling is the attention given to interpersonal relationships, significant others, and contextual issues.

10.18.6 TRAINING IMPLICATIONS

The preceding paragraphs should have placed issues in their proper perspective not only for clinical recognition by psychologists and mental health professionals in sub-Sahara Africa but also for training and empirical research. The lack of mental health resources—human and infrastructural—makes access to Western-type mental health services very limited. In spite of industrialization and urbanization, sub-Sahara Africans seem to be on a journey to reinvent the traditional African culture. Any training of mental health professionals should recognize this journey by not assuming that mental health problems associated with industrialization will be manifested in sub-Sahara Africa the same way as they are in the West.

One of the first implications is that sub-Sahara Africa must develop its own graduate programs in psychology and mental health to prevent its mental health professionals from receiving graduate training almost exclusively in Europe and North America. Training programs should incorporate traditional medicine and healing into their curriculum in order to be responsive to the needs of a large segment of the population that continues to retain the traditional culture. Courses should be organized around themes that emphasize world views and the sociopolitical nature of differences in the region. Students who choose to train in Europe or North America must be encouraged to consider practicum and internship experiences in Africa supervised either by indigenous clinicians in Africa or culturally competent and sensitive non-Africans. This will increase self-knowledge and understanding of other sociocultural factors in mental health. Graduate programs in Europe and North America where most African governments send their citizens to train must develop multicultural competencies that take traditional healing methods into account. More culture-specific strategies must be taught to sub-Sahara African students who will go back to work within this cultural context.

In terms of assessment measures, sociocultural factors should influence their development very significantly. Assessment should also look at the importance of sociodemograhpic variables as they predict psychopathology. This will help in classification systems that will facilitate communication between sub-Sahara African mental health professionals and their counterparts overseas.

Both faculty and students must be encouraged to consider research into traditional healing methods and also sociocultural factors and psychopathology. Such research should utilize both the empirical and experiential approaches in order to capture the totality of the African reality.

10.18.7 CONCLUSION

The influence of the world view on how counseling is delivered and/or accepted in a given society cannot be overemphasized. If such differences in the world view are important considerations in clinical psychology, then more empirical research is needed to define clearly what mental health means in sub-Sahara Africa. A diversified society in which people hold culturally pluralistic views requires a mental health system that is pluralistic, not monocultural. This is a social reality, the recognition of which will facilitate cooperation between

Western trained providers of mental-health services and traditional healers. More than ever, clinicians and researchers must redouble their efforts to reaffirm their belief in the importance of sociocultural factors in psychotherapy as well as the effectiveness of indigenous modes of helping. Cultural competence requires this kind of collaboration between the different forms of traditional psychotherapy and these indigenous modes of treatment.

10.18.8 REFERENCES

Asante, M. K. (1987). *The Afrocentric idea*. Philadelphia: Temple University Press.

Atkinson, D. R., Morten, G., & Sue, D. W. (1993). *Counseling American minorities: A cross-cultural perspective*. Dubuque, IA: Brown.

Binitie, A. (1981). The clinical manifestation of depression in Africans. In T. A. Ban, R. Gonzales, A. S. Jablensky, N. A. Sartorius, & F. E. Vartanian (Eds.), *The prevention and treatment of depression*. Baltimore: Johns Hopkins University Press.

Binitie, A. (1991). The mentally ill in modern and traditional African societies. In S. O. Okpaku (Ed.), *Mental health in Africa and the Americas today: A book of conference proceedings*. Nashville, TN: Chrisolith.

Cross, W. E., Jr. (1971). Negro-to-Black conversion experience: Toward a psychology of black liberation. *Black World, 20*(9), 13–27.

Danquah, S. A. (1982). The practice of behavior therapy in West Africa: The case of Ghana. *Journal of Behavioral Therapy and Experimental Psychiatry, 13*(1), 5–13.

Dawson, J. (1964). Urbanization and mental health in a West African community. In A. Kiev (Ed.), *Magic, Faith and Healing*. New York: Free Press.

Essandoh, P. K. (1995). Counseling issues with African college students in US colleges and universities. *The Counseling Psychologist, 23*(2), 348–360.

Frank, J. D. (1978). *Psychotherapy and the human predicament*. New York: Schocken.

German, G. A. (1987). Mental health in Africa: I. The extent of mental health problems in Africa today: An update of epidemiological knowledge. *British Journal of Psychiatry, 151, 435–439*.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.

Jackson, B. (1975). Black identity development. *Journal of Educational Diversity and Innovation, 2, 19–25*.

Kalu, W. (1990). Bereavement and stress in career women. *Women and Therapy, 10*(3), 75–87.

Kisekka, M. N. (1990). Gender and mental health in Africa. *Women and Therapy, 10*(3), 1–13.

Lamb, D. (1983). *The Africans*. New York: Vintage.

Lamptey, J. J. (1977). Patterns of psychiatric consultation at the Accra Psychiatric Hospital. *African Journal of Psychiatry, 3, 123–127*.

Leighton, A. H., Lambo, T. A., & Hughes, C. C. (1963). *Psychiatric disorders among the Yoruba*. Ithaca, NY: Cornell University Press.

Miller, J. B. (1976). *Toward a new psychology of women*. Boston: Beacon Press.

Okpaku, S. O. (1991). *Mental health in Africa and the Americas today: A book of conference proceedings*. Nashville, TN: Chrisolith.

Onyioha, K. O. K. (1977). The metaphysical background to traditional healing in Nigeria. In P. Singer (Ed.), *Traditional healing: New science or new colonialism*. New York: Conch Magazine.

Orley, J. (1972). A prospective study of 372 consecutive admissions to Butabika Hospital, Kampala. *East African Medical Journal, 49, 16–26*.

Some, M. P. (1993). *Ritual: Power, healing and community*. Portland, OR: Swan/Raven.

Taft, R. (1977). Coping with unfamiliar cultures. In N. Warren (Ed.), *Studies in cross-cultural psychology* (Vol. 1). New York: Academic Press.

Vontress, C. E. (1991). Traditional healing in Africa: Implications for cross-cultural counseling. *Journal of Counseling and Development, 70, 242–249*.

Wyllie, R. W. (1983). Ghanaian spiritual and traditional healers' explanation of illness: A preliminary survey. *Journal of Religion in Africa, 14, 46–57*.

Yalom, I. D. (1995). *Theory and practice of group psychotherapy*. New York: Basic Books.